



## Part II – Circumstantial Evidence

Other possible sources of the infection included the woman's prior sexual contacts and occupational exposure, given that she was a nurse.

All seven of the men that she had been in sexual contact with (including her former boyfriend) were tested and found to be HIV-negative.

Her employment records were examined, and there were no reports of any accidental or occupational exposures other than the saliva that was splashed on her skin sometime in the mid-1980s. Her file did not have any documentation of any needle sticks at work.

As the investigation proceeded, it was found that an HIV-positive patient under the care of the ex-boyfriend/doctor had blood drawn at the physician's offices on August 4, 1994. The paperwork for this procedure was deliberately hidden (the log book was found in a box of "1982 records" in a storage room with other records from the 1980s) and was not filled out in a manner that was consistent with standard office practices.

Based on the circumstantial case against the physician, the reverse transcriptase (RT) sequences from the victim (the nurse/ex-girlfriend) and from the physician's HIV-positive patient (the putative source of the nurse's infection via the injection administered during the ill-fated argument) were analyzed. The reverse transcriptase (pol) gene was used in the actual criminal trial. Reverse transcriptase has a unique set of selective pressures exerted on it and is therefore a useful gene/protein to explore in order to establish relationships between different strains of HIV.

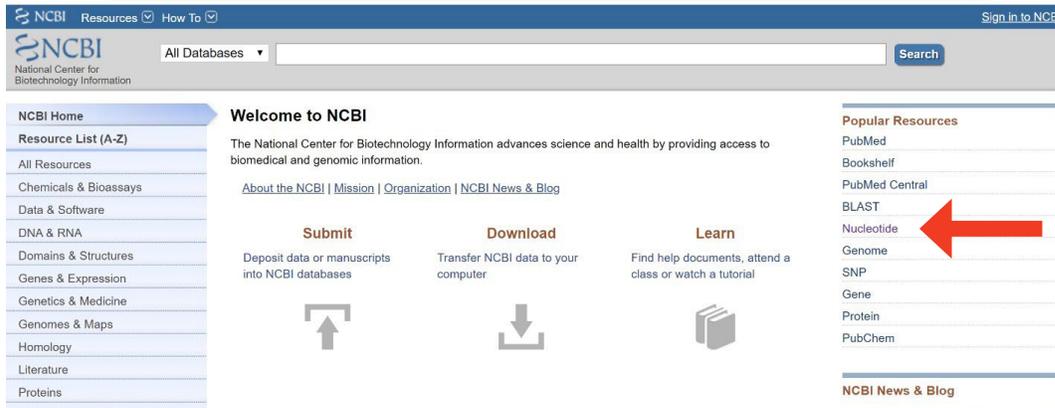
### *Question*

1. HIV-1 mutates rapidly. Based on this, would you expect to find a single, identical sequence in the victim and the patient, or would you expect to find a set of related sequences that share a common ancestor? Explain your answer.

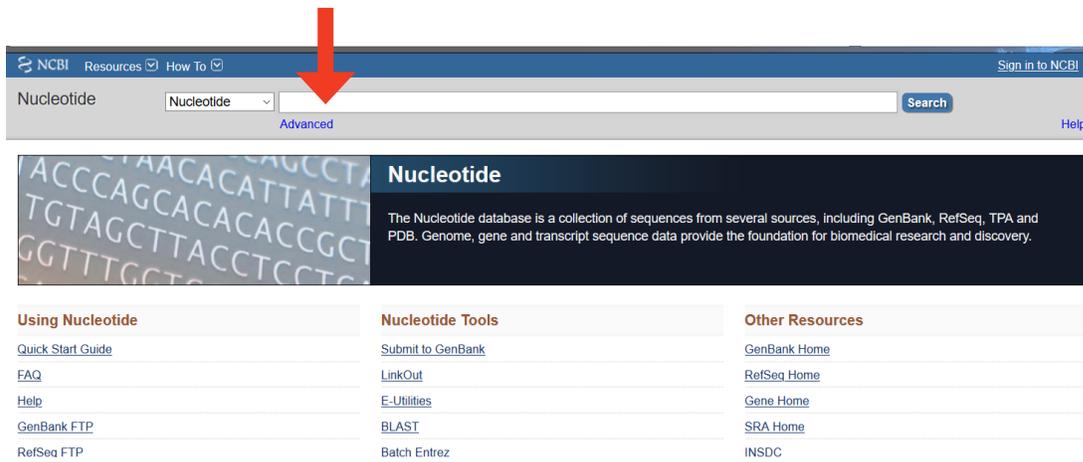
## Part III – Comparing Sequences

You will be examining the relationship between patient and victim sequences from the actual criminal trial. Here's how:

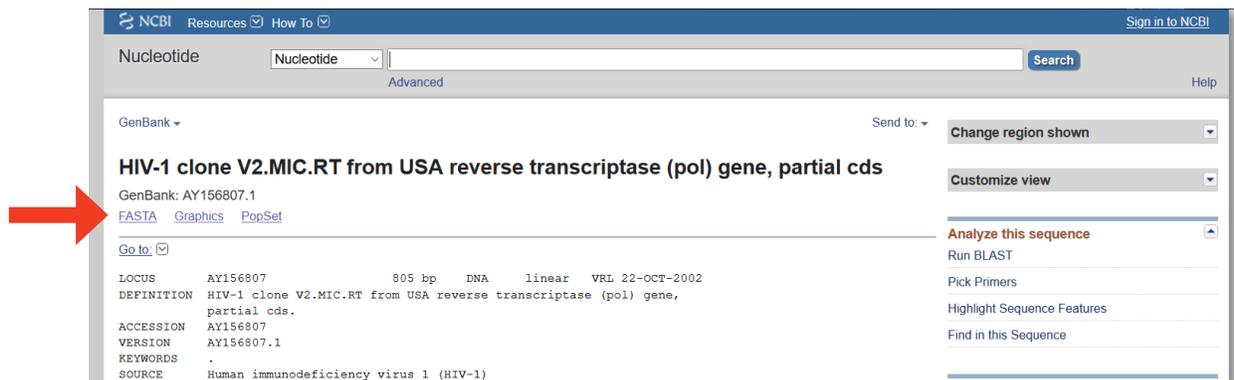
1. Go to the NCBI homepage (<http://www.ncbi.nlm.nih.gov/>). On the right, you will see the link “Nucleotide” for searching for nucleotide sequences. Click on this link (see red arrow in screen capture below).



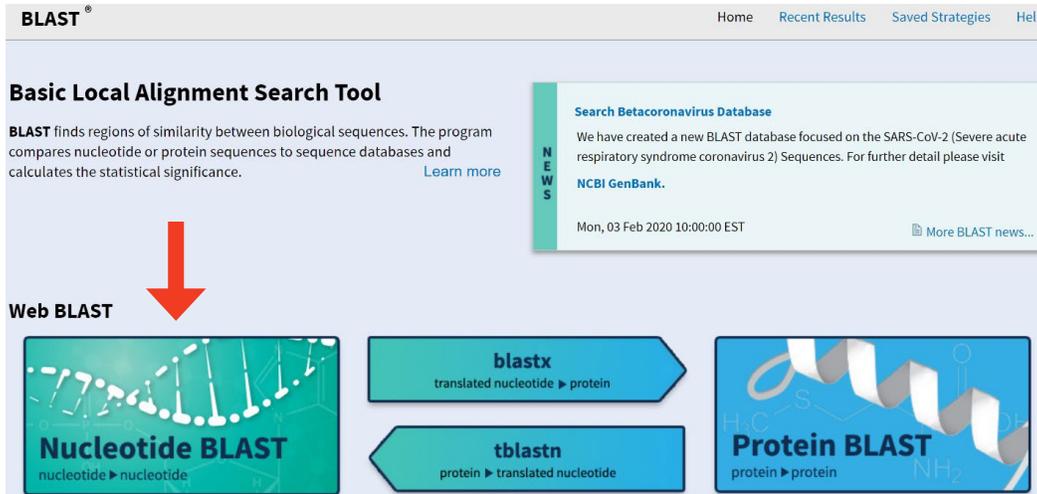
2. You are now prepared to search for a sequence from the actual trial. In order to find sequences from the trial, enter “AY156803” into the search box at the top of the page. This is the ID number for a patient (“P”) sequence from the trial. When you click on the “Search” button, you will find this particular sequence from the trial.



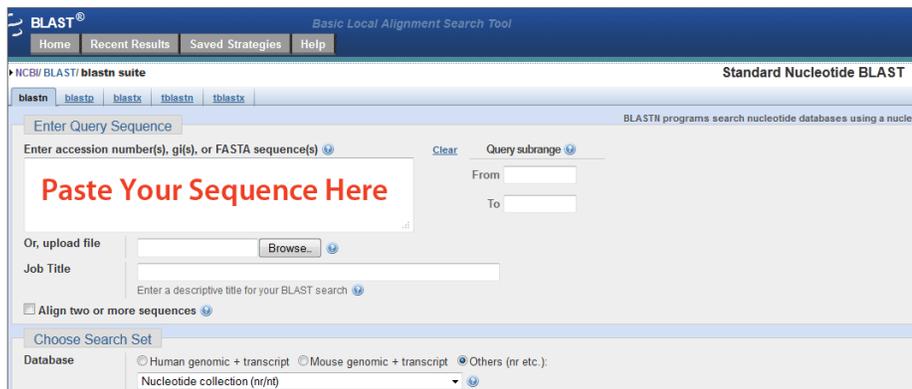
3. Once the search result loads, click on FASTA towards the top (see below). This will open up a new window with that Patient HIV sequence in a format (FASTA) suitable for alignment and tree building in Bioinformatics programs. Copy and paste that sequence (starting with the first nucleotide) into a text file.



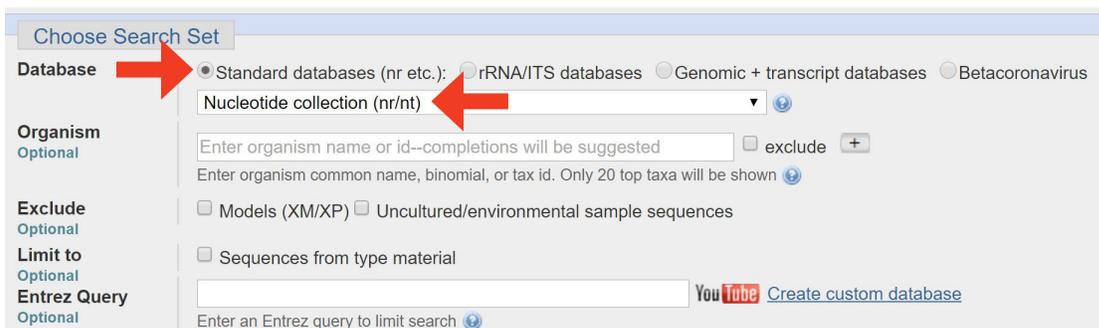
- Now that you have the nucleotide sequence from one HIV strain from this trial, you are going to use that sequence to find related sequences in the database. You will use a program called “BLAST” to do this (BLAST makes sequence comparisons, and finds closely related sequences in the database).
- Navigate to BLAST (back on the NCBI homepage, on the right column, above the first link you clicked—see figure for step 1 above). You will then come to a page that looks like the figure below. Click on Nucleotide BLAST (as you want to compare nucleotide sequences).



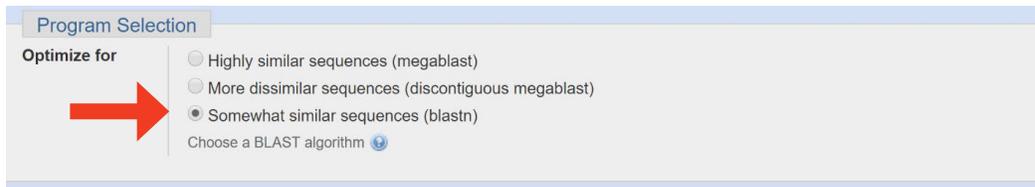
- When you get to the nucleotide blast page, paste your FASTA sequence (from your text file) into the top box (Enter Query Sequence).



- On that same screen, under “Choose Search Set,” select “Standard databases (nr etc.)” for the database and “nucleotide collection (nr/nt)” from the database dropdown menu.



- Under Program Selection (towards the bottom of the same screen), select the button to optimize for somewhat similar sequences (blastn):



- A little further below, click on the BLAST button.



- After a few moments, you will get a list of “hits” that have nucleotide similarities to your Query sequence (from the patient). Not surprisingly, the top hits are patient and victim sequences from this case. You now need to select several sequences to compare and build a phylogenetic tree. Scroll down to the “Alignments” section of this page. This is where individual sequences have been aligned with the patient sequence that you entered. Select several victim and patient sequences by clicking in their check boxes. You should include between five and eight sequences, including both patient and victim sequences. Also select some sequences that are from other sources (e.g., USA sequences or sequences from other countries).
- Once you have checked the sequences you want to compare, click on “Distance tree of results”:

Sequences producing significant alignments:

Select: All None Selected: 10

Alignments Download GenBank Graphics Distance tree of results

	Description	Max score	Total score	Query cover	E value	Max ident	Accession
<input checked="" type="checkbox"/>	HIV-1 clone P6.MIC.RT from USA reverse transcriptase (pol) gene, partial cds	1452	1452	100%	0.0	100%	AY156803.1
<input checked="" type="checkbox"/>	HIV-1 clone V2.MIC.RT from USA reverse transcriptase (pol) gene, partial cds	1429	1429	100%	0.0	99%	AY156807.1
<input checked="" type="checkbox"/>	HIV-1 clone V1.MIC.RT from USA reverse transcriptase (pol) gene, partial cds	1425	1425	100%	0.0	99%	AY156806.1

This will build a phylogenetic tree of the sequences that you have selected.

- The phylogenetic tree will open in a new window. The sequences will have the same labels as they did when you checked them from the Alignments listing in step 10. Remember that sequences from our patient and victim are labeled with a P and V, respectively.

### Questions

- Describe the tree (in general terms). On the next page, draw a quick sketch of the tree.

