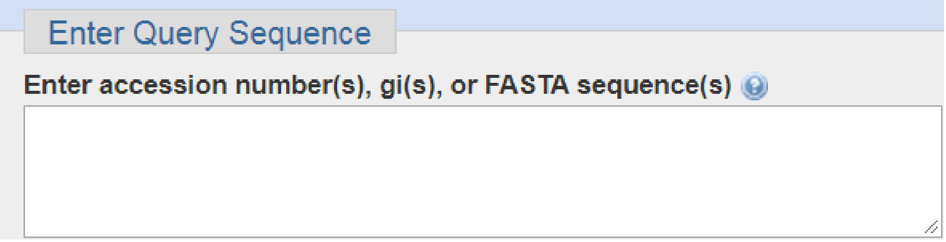
Confirming the presence of HIV-1 envelope gene sequences in the genome of James D. Watson.

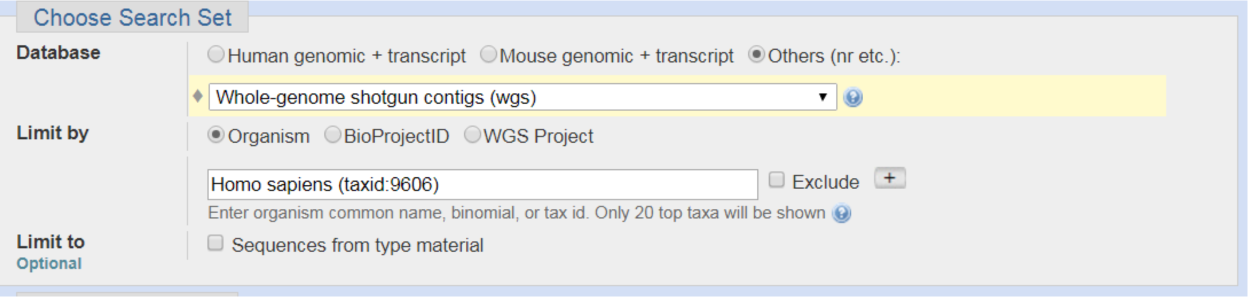
K03455.1 is the accession number of the HIV-1 (HXB2), complete genome; the HIV/HTLV-III/LAV reference genome.

Go to https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch

Enter K03455.1 into this box



In the Choose Search Set box select “Whole-genome shotgun contigs (wgs)” [1] and “Homo sapiens (taxid: 9606) from the dropdown boxes.



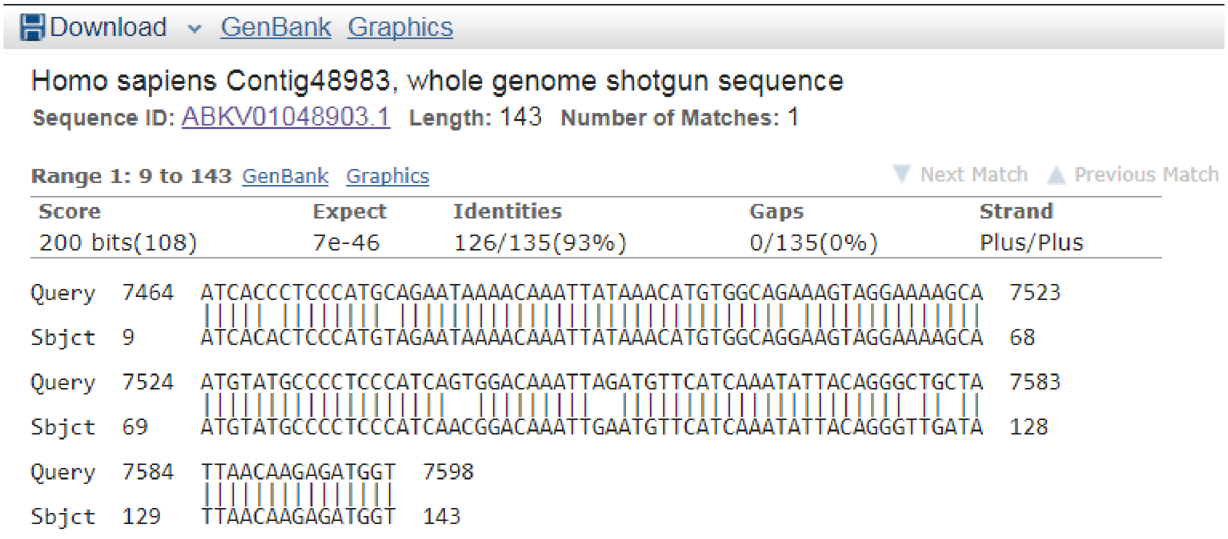
Database ○ Human genomic + transcript ○ Mouse genomic + transcript ● Others (nr etc.):
Whole-genome shotgun contigs (wgs)
Limit by ● Organism ○ BioProjectID ○ WGS Project
Homo sapiens (taxid:9606) ☐ Exclude +
Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown
Limit to Optional ☐ Sequences from type material

Scroll down to the bottom of the page and click the BLAST button on the left hand side. Wait until the program finishes calculating. This may take several minutes.

When the results appear scroll down below the bottom of the page and the “Alignments” heading.

This reveals a 135 bp “Sequence ID: ABKV01048903.1”
Click on the sequence ID link. ABKV01048903.1

This opens a page “Homo sapiens Contig48983, whole genome shotgun sequence” [1]. This is the DNA sequence of a diploid genome of a single individual, James D. Watson [2].



Homo sapiens Contig48983, whole genome shotgun sequence

Sequence ID: ABKV01048903.1 **Length:** 143 **Number of Matches:** 1

Range 1: 9 to 143 GenBank Graphics

Score	Expect	Identities	Gaps	Strand
200 bits(108)	7e-46	126/135(93%)	0/135(0%)	Plus/Plus

```
Query  7464  ATCACCCTCCCATGCAGAATAAAACAAATTATAAACATGTGGCAGAAAGTAGGAAAAGCA  7523
             ||||| |||||||| |||||||||||||||||||||||||||||| ||||||||||||||
Sbjct  9     ATCACACTCCCATGTAGAATAAAACAAATTATAAACATGTGGCAGGAAGTAGGAAAAGCA  68

Query  7524  ATGTATGCCCCTCCCATCAGTGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTA  7583
             |||||||||||||||||||  |||||||||  |||||||||||||||||||||| || ||
Sbjct  69    ATGTATGCCCCTCCCATCAACGGACAAATTGAATGTTCATCAAATATTACAGGGTTGATA  128

Query  7584  TTAACAAGAGATGGT  7598
             |||||||||||||||
Sbjct  129   TTAACAAGAGATGGT  143
```

The 135 bases Query HIV-1 sequence
ATCACCCTCCCATGCAGAATAAAACAAATTATAAACATGTGGCAGAAAGTAGGAAAAGCAATG
TATGCCCCTCCCATCAGTGGACAAATTAGATGTTCATCAAATATTACAGGGCTGCTATTAACA
AGAGATGGT
is nucleotides 7464-7498 in the HIV-1 env gp120/160 protein coding region of HIV-1.

126/135 bases of the Subject James Watson sequence is 93% identical to the HIV-1 sequence. eValue $7X10^{-46}$ with no gaps [3]. eValue is an estimate of the probability of the HIV-1 sequence being a chance finding. eValues $<10^{-6}$ are considered significant. This eValue is approximately zero.

References

1. “Shotgun sequencing: An approach used to decode a genome by shredding ("shotgunning") it into smaller fragments of DNA which can then be individually sequenced. The sequences of these fragments are then ordered, based on overlaps in the genetic code, and finally reassembled into the complete sequence”. http://www.medicinenet.com/script/main/art.asp?articlekey=20667

2. Wheeler DA, Srinivasan M, Egholm M, Shen Y, Chen L, McGuire A, He W, Chen Y-J, Makhijani V, Roth GT. The complete genome of an individual by massively parallel DNA sequencing. *Nature* 2008. 452:872-876. https://www.ncbi.nlm.nih.gov/pubmed/18421352

3. Romero Fernández-Bravo M. Readers comment on genome of Dr. James Watson. *Nature* 2014. 452:872-876.
http://www.nature.com/nature/journal/v452/n7189/full/nature06884.html#comment-64495
http://www.nature.com/nature/journal/v452/n7189/full/nature06884.html#comment-67241