BTE401: Bioinformatics

Summer 2020

Section : 02

QUIZ2

Date of submission: 09-08-20

Submitted to:

Romana Siddique

Senior Lecturer

Biotechnology Program

Department of Mathematics and Natural Sciences

BRAC University.

Submitted by:

Mahfuzul Hasan

ID: 17126024

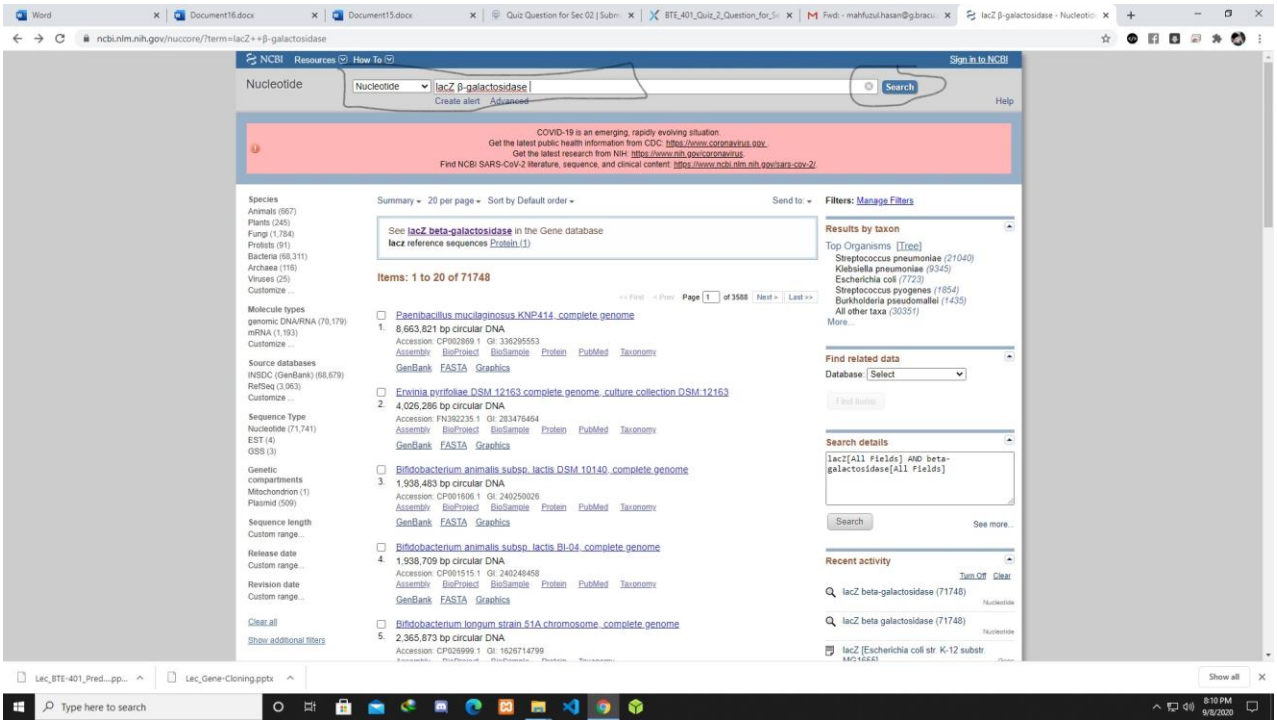
Microbiology program.

Department of Mathematics and Natural Sciences.

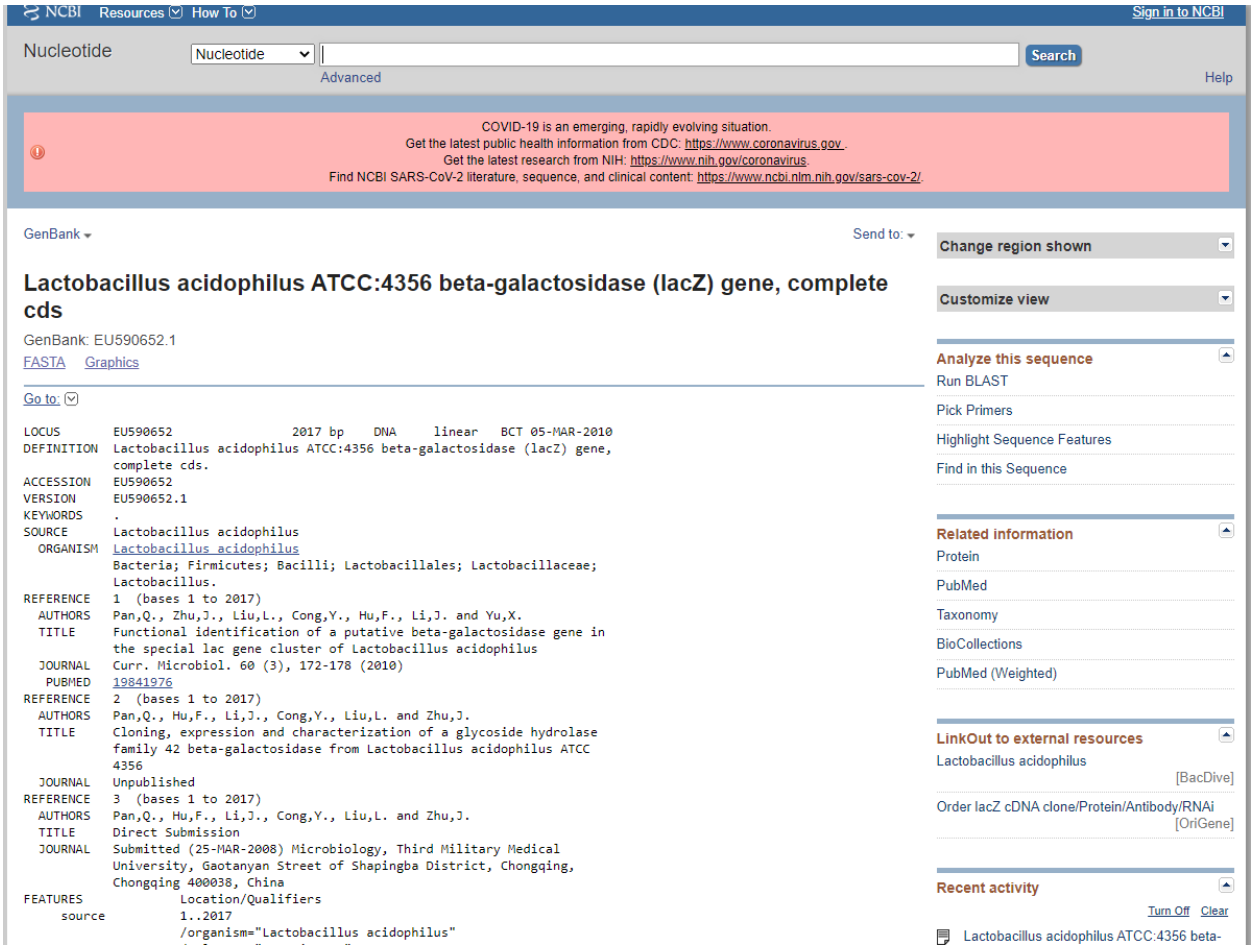
Answer to the question no 1:

Here, we extracted the FASTA sequence of lacZ gene from NCBI GenBank. For this purpose, the following steps were performed:

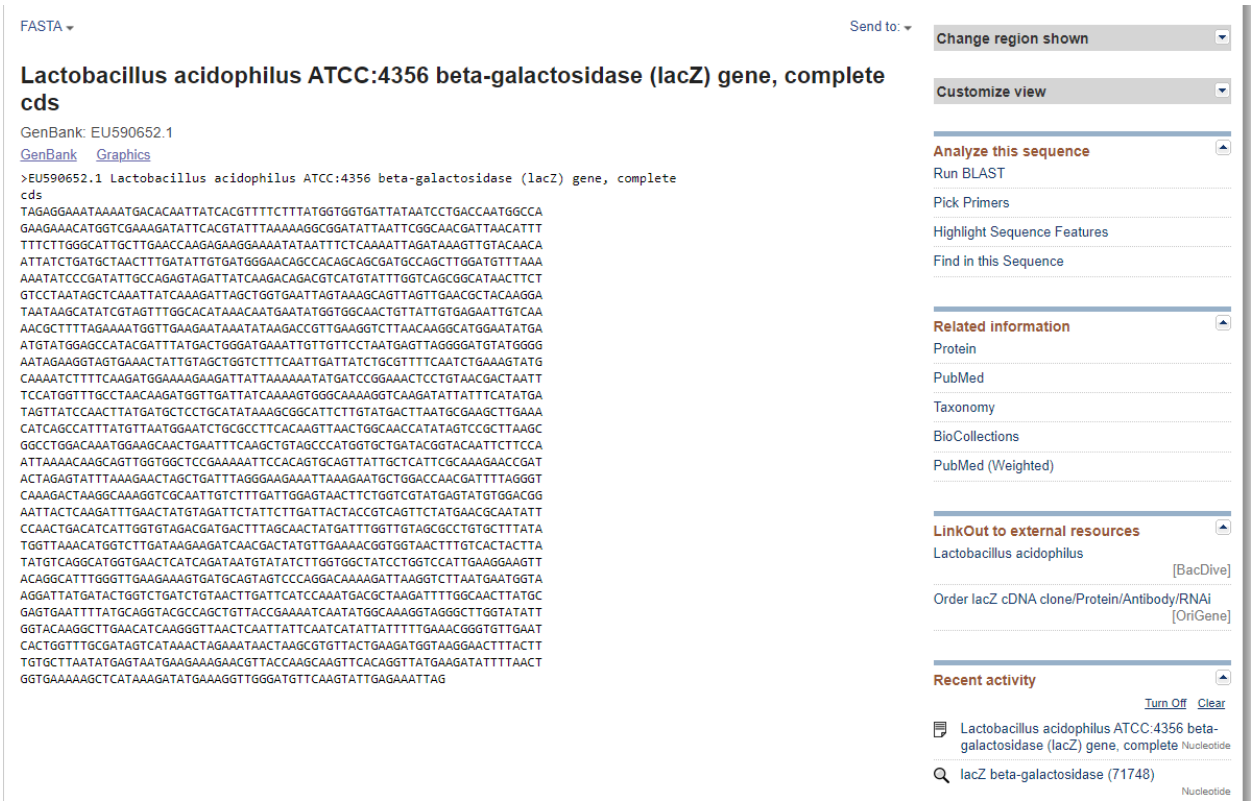
Step 1: searched for lac Z gene in the database:



Step2: after that, selected the desired gene from the database:

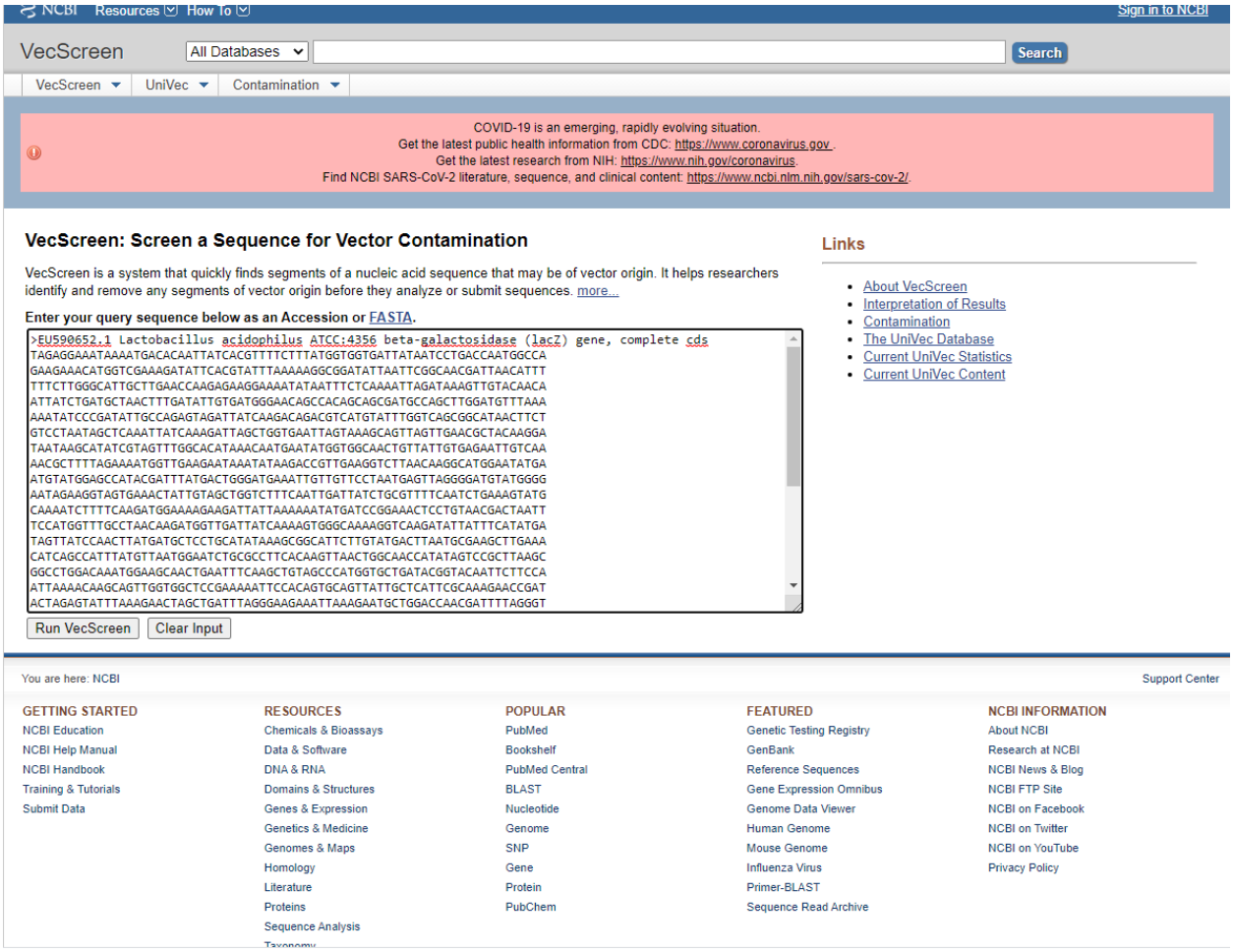


Step 3: finally, took the nucleotide sequence in FASTA format:

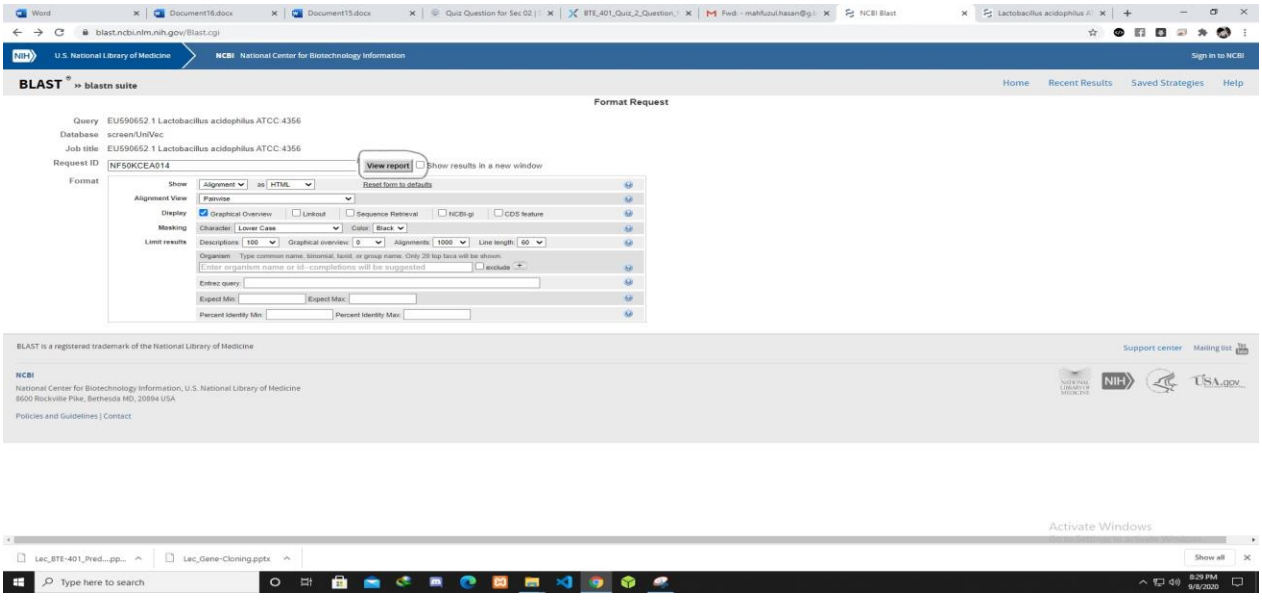


For detection of vector contamination, we used VecScreen (https://www.ncbi.nlm.nih.gov/tools/vecscreen/).

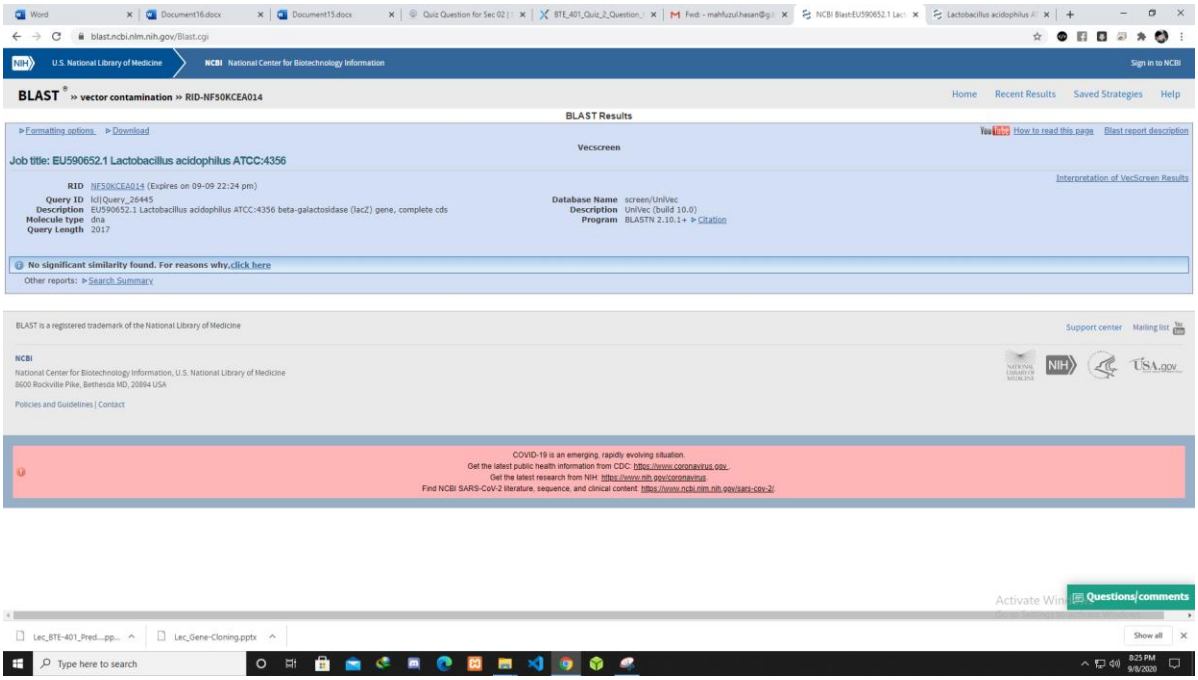
In vecscreen, the FASTA format of the sequence was pasted in query box



Then pressed “Run VecScreen“ and the report was viewed:



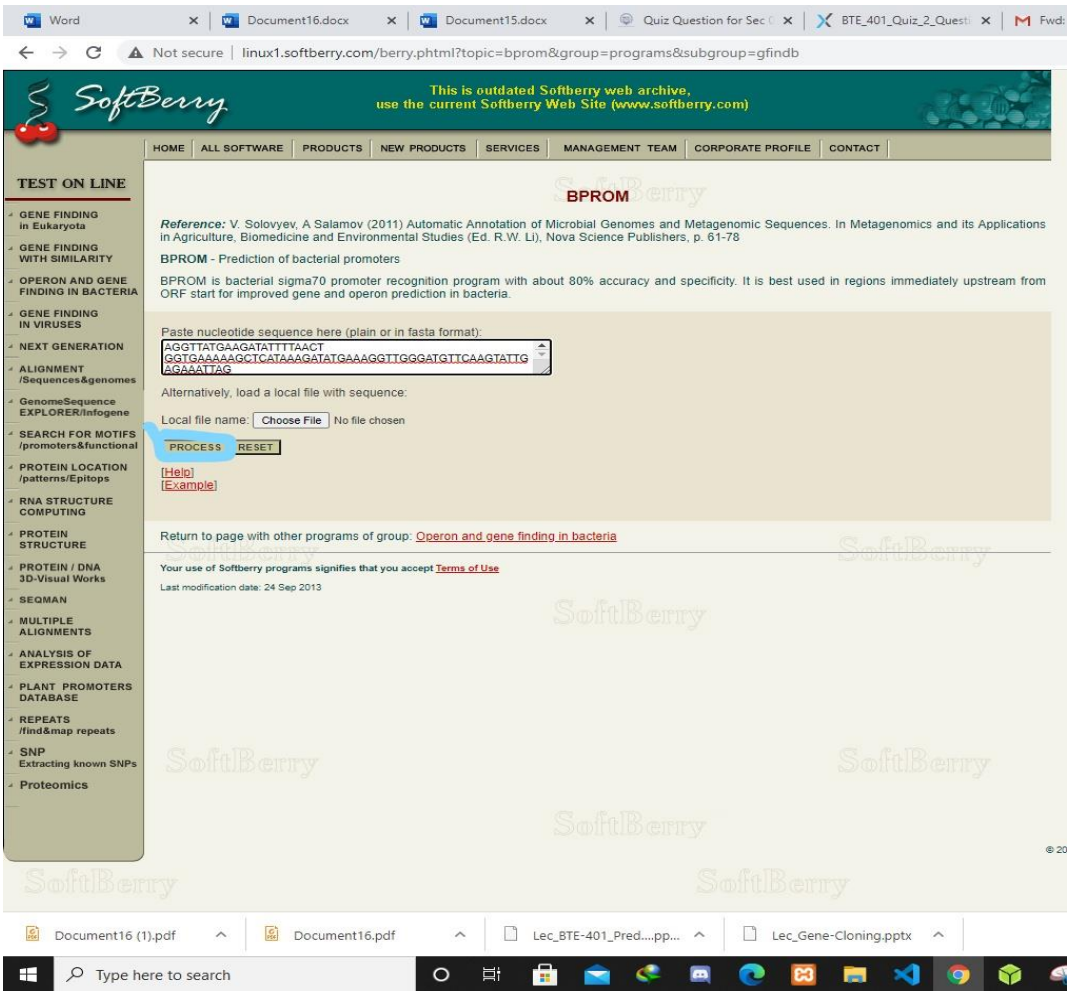
The report showed that the nucleotide sequence did not contain any vector contamination:



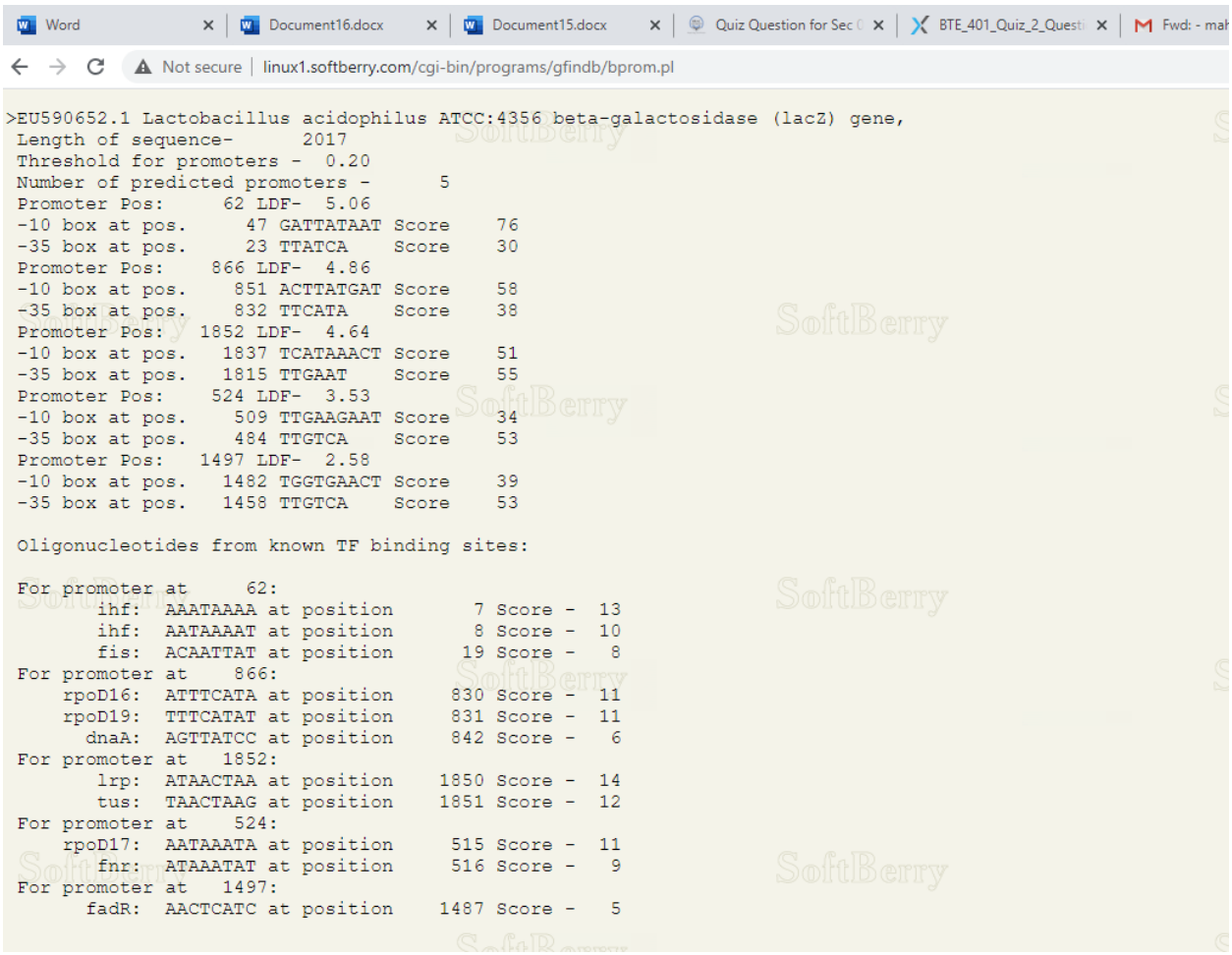
Answer to the question no: 02

For finding promoter of Lactobacillus acidophilus ATCC:4356 beta-galactosidase (lacZ) gene, we used BPROM because the organism is prokaryotic. First, went to the website: http://linux1.softberry.com/berry.phtml
Here, selected BPROM from menu:

Here, we pasted the sequence in FASTA format



Then after pressing ”process” option, we could view the possible promoters for this sequence. Here, the sequence suggested 5 possible promoters. Moreover, the site also shows some TF binding sites, which influence the transcription process.



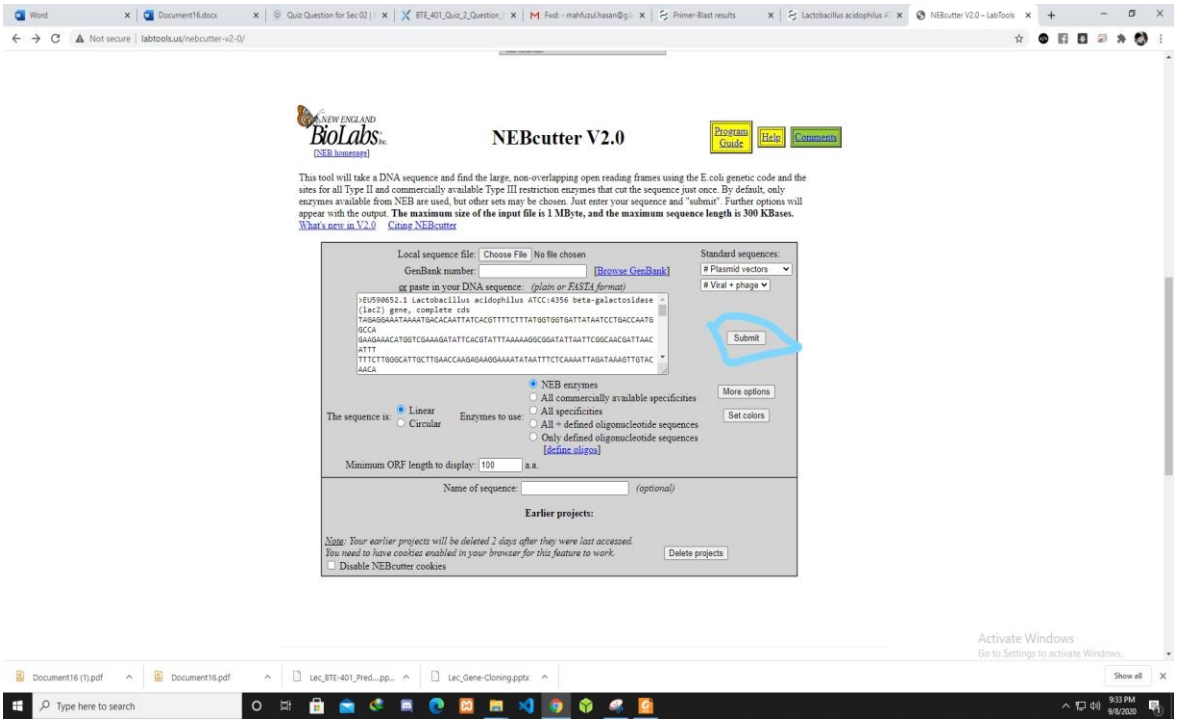
Word | Document16.docx | Document15.docx | Quiz Question for Sec | BTE_401_Quiz_2_Questi | Fwd: - mah

Not secure | linux1.softberry.com/cgi-bin/programs/gfindb/bprom.pl

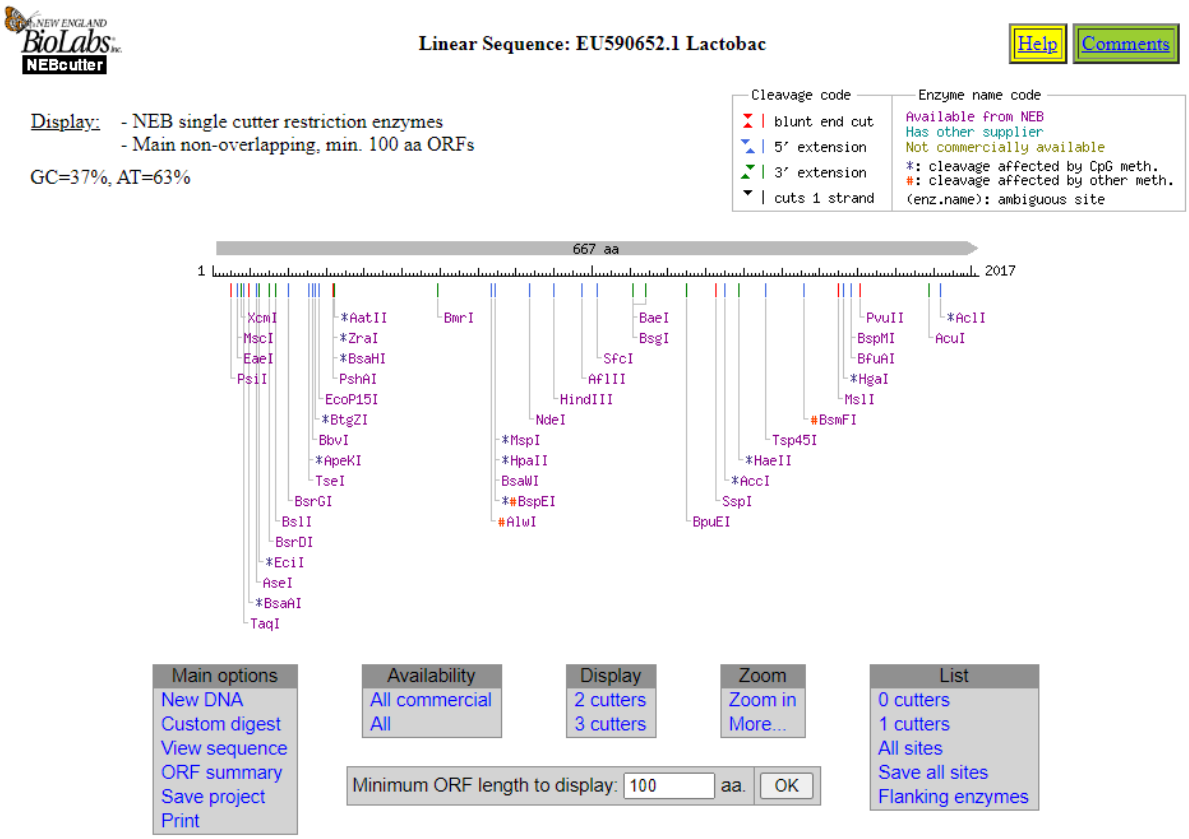
```
>EU590652.1 Lactobacillus acidophilus ATCC:4356 beta-galactosidase (lacZ) gene,
 Length of sequence-      2017
 Threshold for promoters -  0.20
 Number of predicted promoters -      5
 Promoter Pos:     62 LDF-  5.06
 -10 box at pos.     47 GATTATAAT Score    76
 -35 box at pos.     23 TTATCA    Score    30
 Promoter Pos:    866 LDF-  4.86
 -10 box at pos.    851 ACTTATGAT Score    58
 -35 box at pos.    832 TTCATA    Score    38
 Promoter Pos:   1852 LDF-  4.64
 -10 box at pos.   1837 TCATAAACT Score    51
 -35 box at pos.   1815 TTGAAT    Score    55
 Promoter Pos:    524 LDF-  3.53
 -10 box at pos.    509 TTGAAGAAT Score    34
 -35 box at pos.    484 TTGTCA    Score    53
 Promoter Pos:   1497 LDF-  2.58
 -10 box at pos.   1482 TGGTGAACT Score    39
 -35 box at pos.   1458 TTGTCA    Score    53

 Oligonucleotides from known TF binding sites:

 For promoter at     62:
        ihf:  AAATAAAA at position       7 Score -  13
        ihf:  AATAAAAT at position       8 Score -  10
        fis:  ACAATTAT at position      19 Score -   8
 For promoter at    866:
     rpoD16:  ATTTCATA at position     830 Score -  11
     rpoD19:  TTTCATAT at position     831 Score -  11
       dnaA:  AGTTATCC at position     842 Score -   6
 For promoter at   1852:
        lrp:  ATAACTAA at position    1850 Score -  14
        tus:  TAACTAAG at position    1851 Score -  12
 For promoter at    524:
     rpoD17:  AATAAATA at position     515 Score -  11
        fnr:  ATAAATAT at position     516 Score -   9
 For promoter at   1497:
       fadR:  AACTCATC at position    1487 Score -   5
```

Answer to the question no : 03

For finding out which restriction enzymes can cut this sequence, we used NEBCUTTER version 2 (http://www.labtools.us/nebcutter-v2-0/).



Then after pressing “Submit”, the tool provided a list of restriction enzymes which can be used for cutting this DNA sequence:

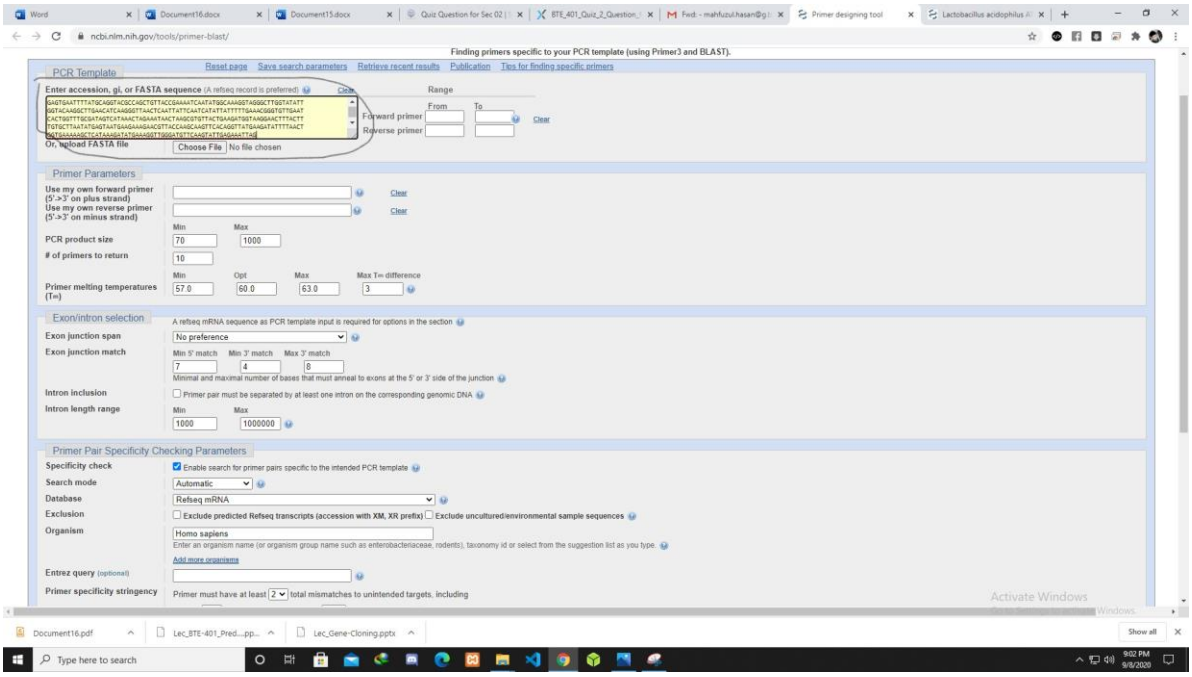


Answer to the question no 04

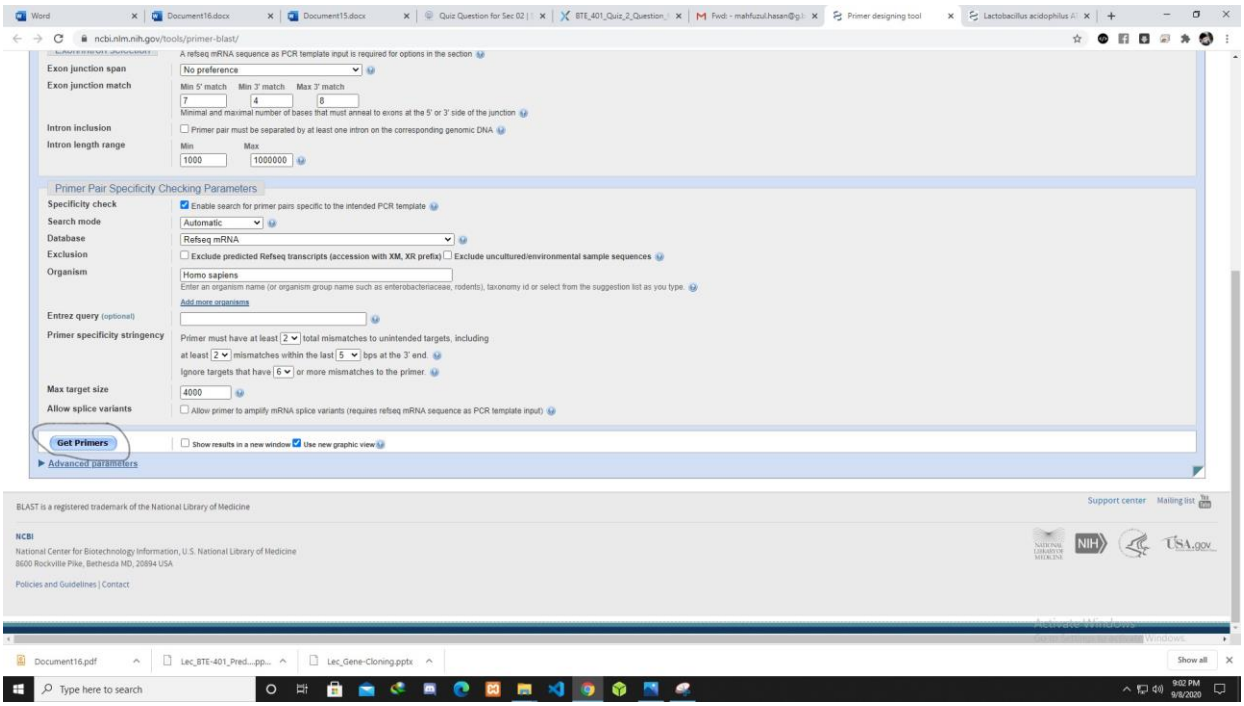
Here, for designing primers for this gene, PRIMER-BLAST was used. First, visited the website:(https://www.ncbi.nlm.nih.gov/tools/primer-blast/)

The sequence obtained from NCBI-GenBank was placed on the query box.

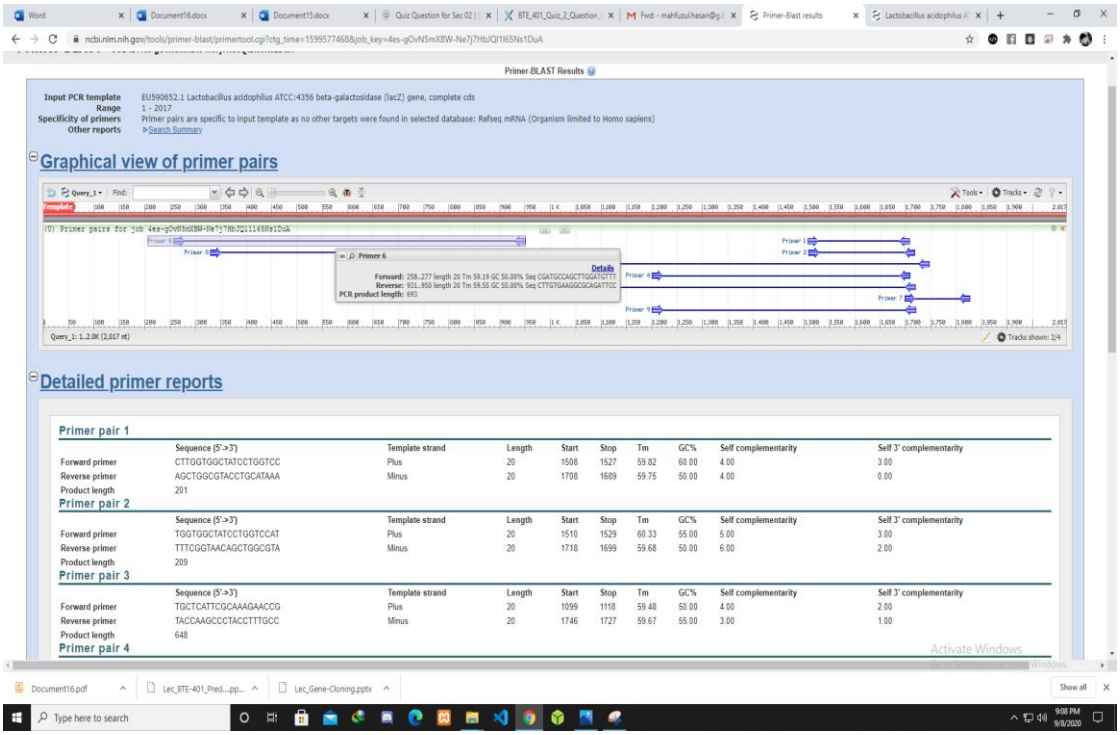
Here, the default conditions provided by the database was used for primer detection,



Then, clicked the option of “get primers”.



The result suggested 10 pairs of primer.



However, among these 10 pairs, it is efficient to choose the one with the highest product length which is 744 base pairs in this case (pair 5). While choosing primer, the product length is an important parameter since it determines the length of the sequence to be amplified.