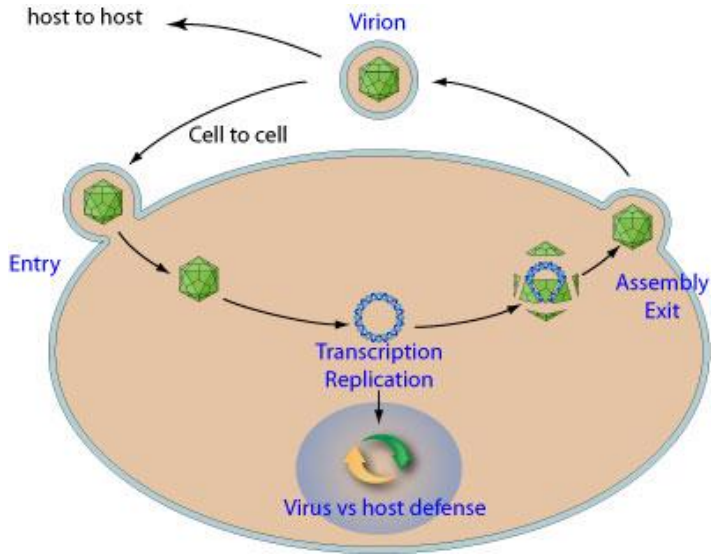
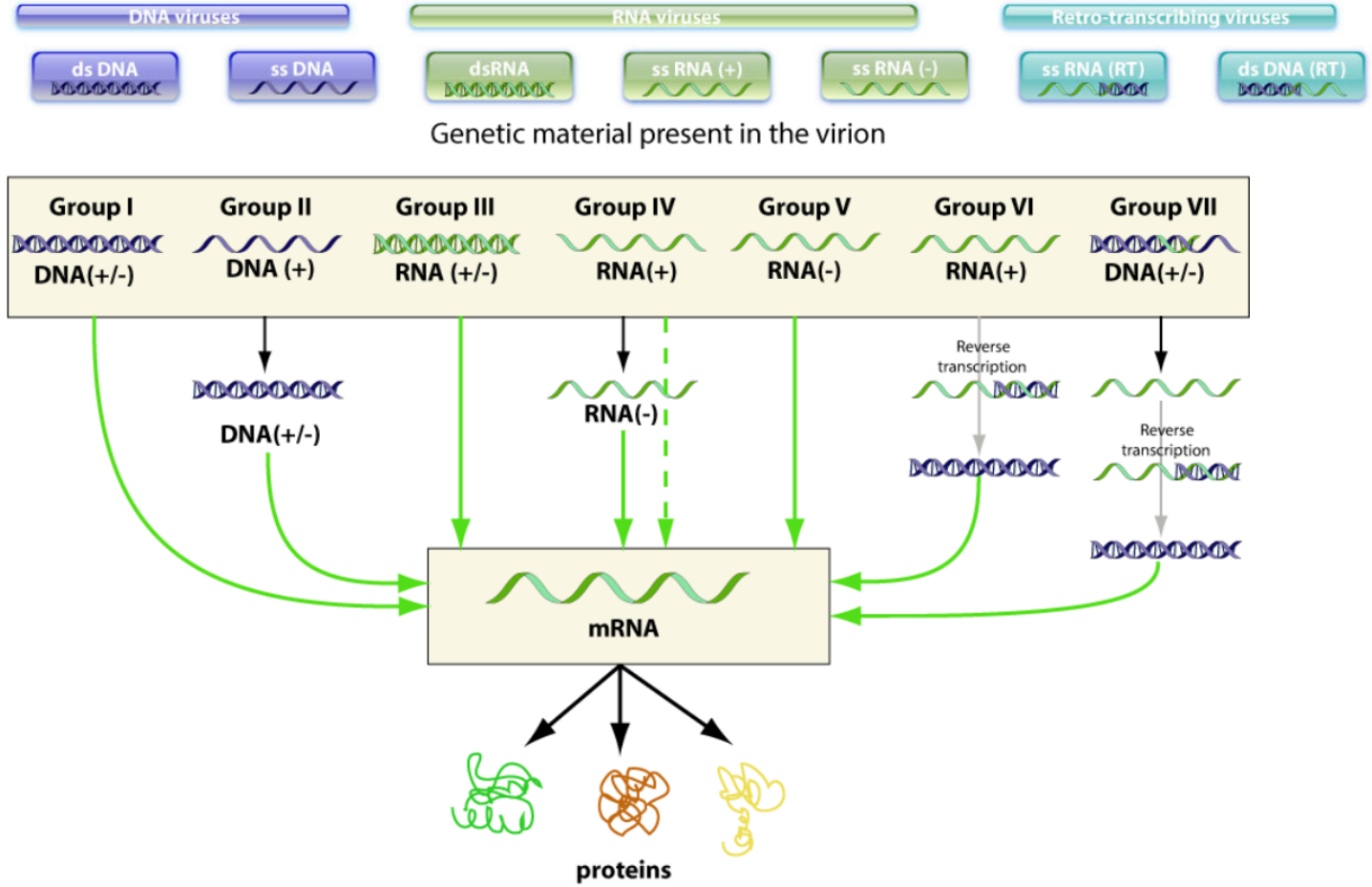
DIVISION OF INFECTION AND IMMUNITY
CENTRE FOR CLINICAL MICROBIOLOGY

Basics concepts of molecular virology

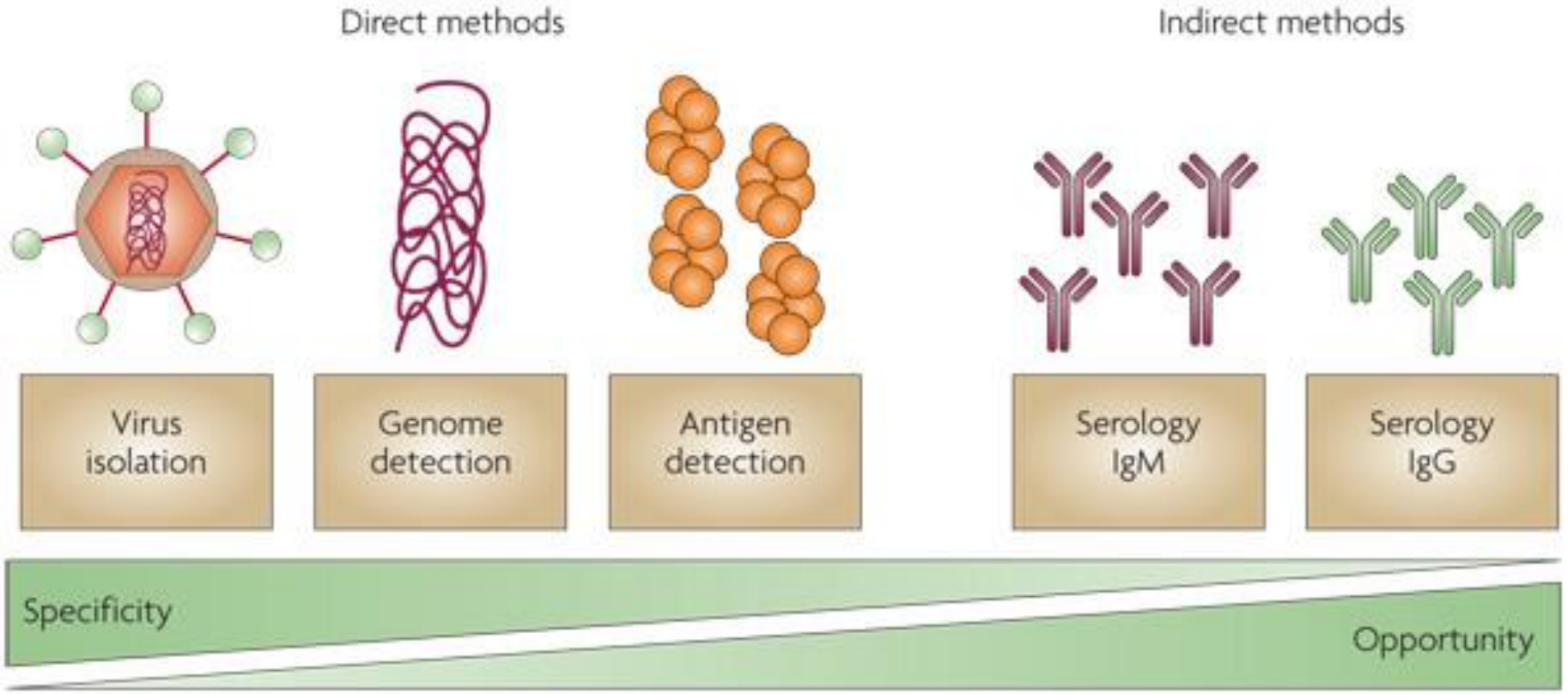
Dr Liã Bárbara Arruda
Postdoctoral Research Associate
Centre for Clinical Microbiology
PANDORA-ID-NET Consortium
14 Feb 2020

Viruses are just the coolests





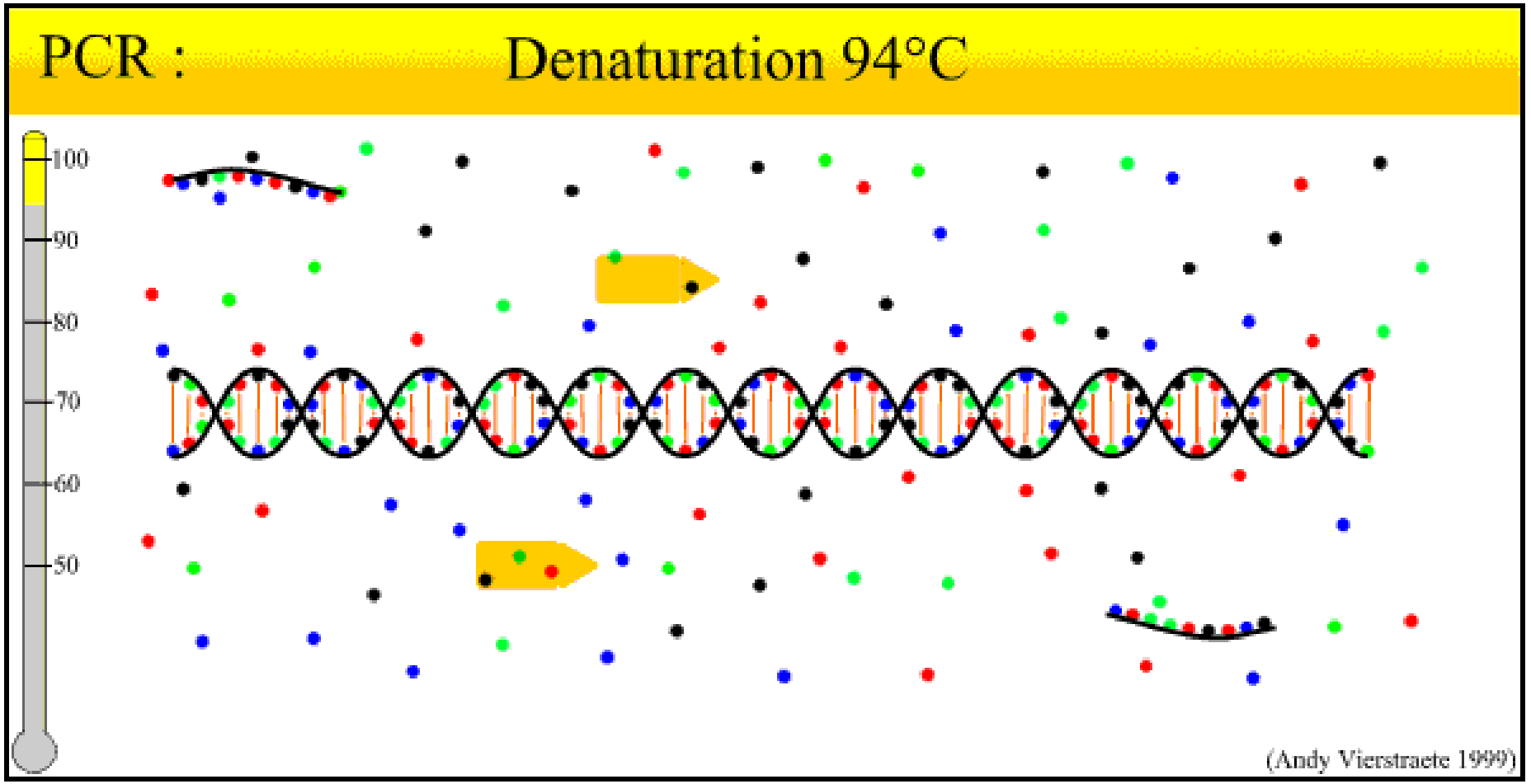
Detection of viruses



https://www.who.int/in-vitro-diagnostic/en/

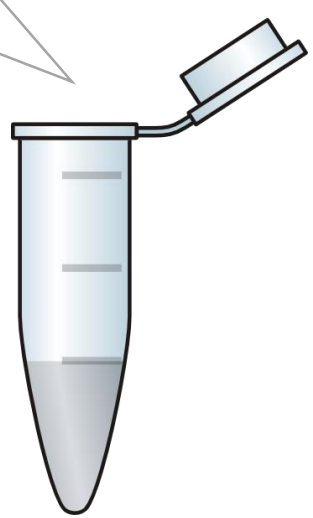
Peeling et al. Nat Rev Microbiol. 2010 Dec;8(12 Suppl):S30-8. doi: 10.1038/nrmicro2459

Polymerase chain reaction (PCR)

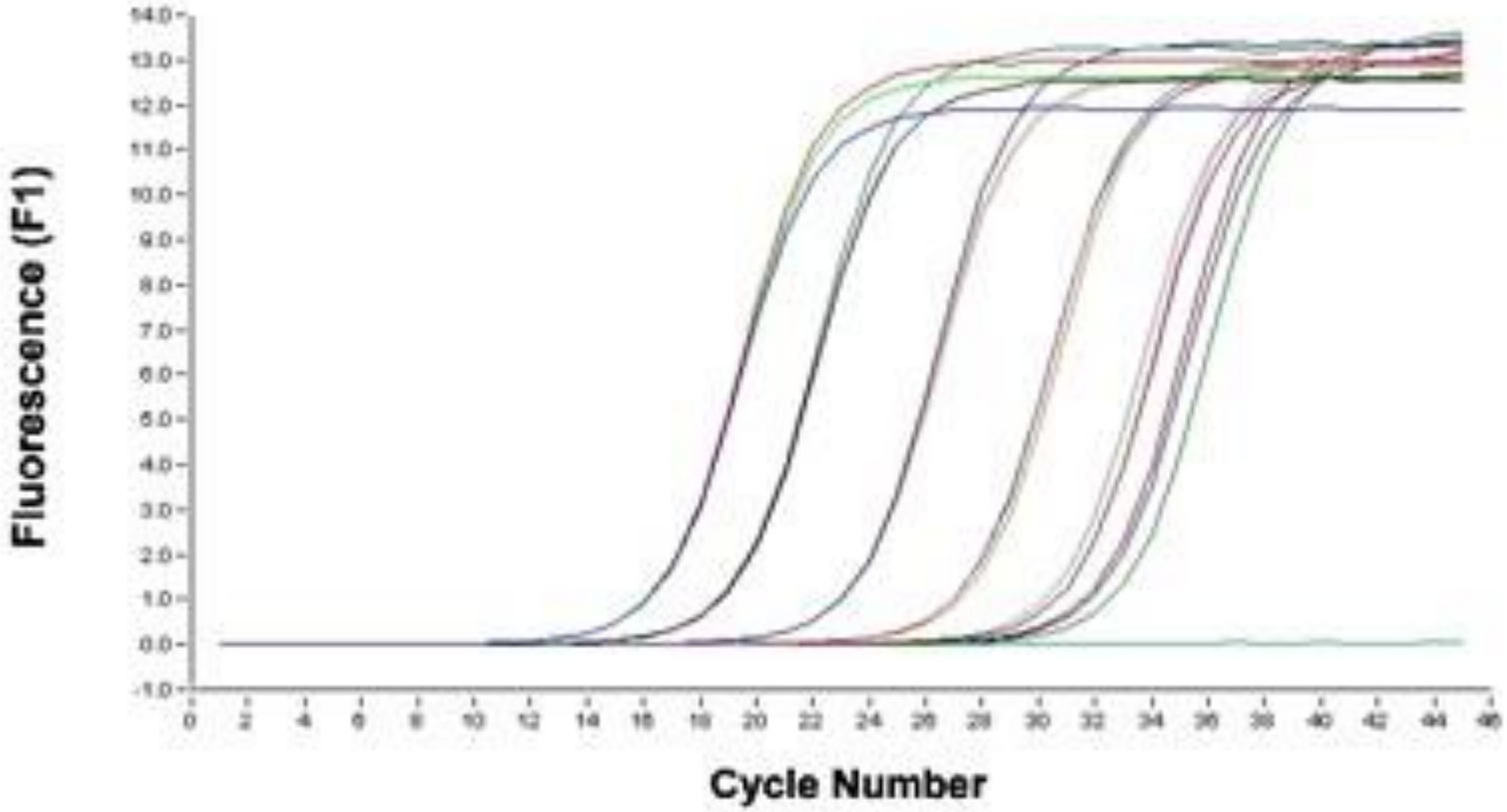


PCR components

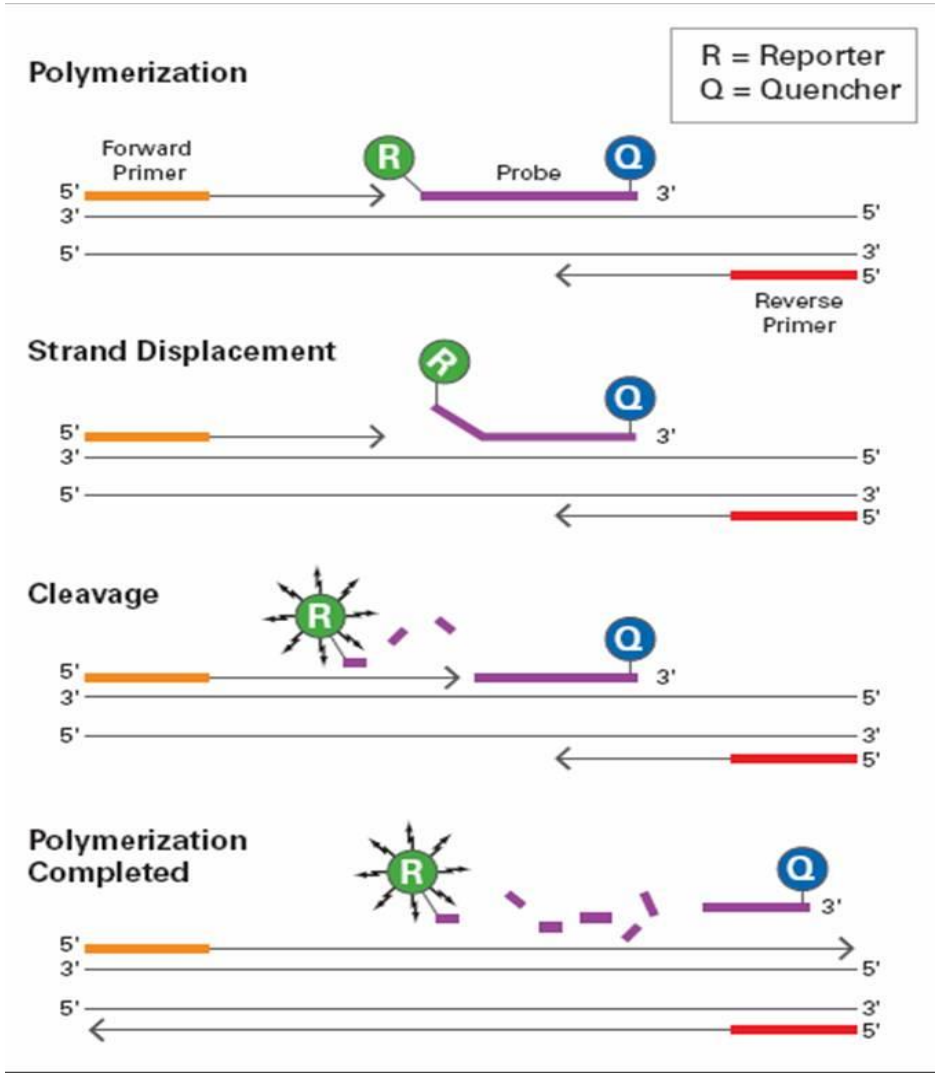
Nuclease-free water
Buffer
Taq DNA Polymerase*
Taq DNA Polymerase co-factor (Mg++)
dNTPs
Oligonucleotides
Template



Real-time PCR (qPCR)



TaqMan system



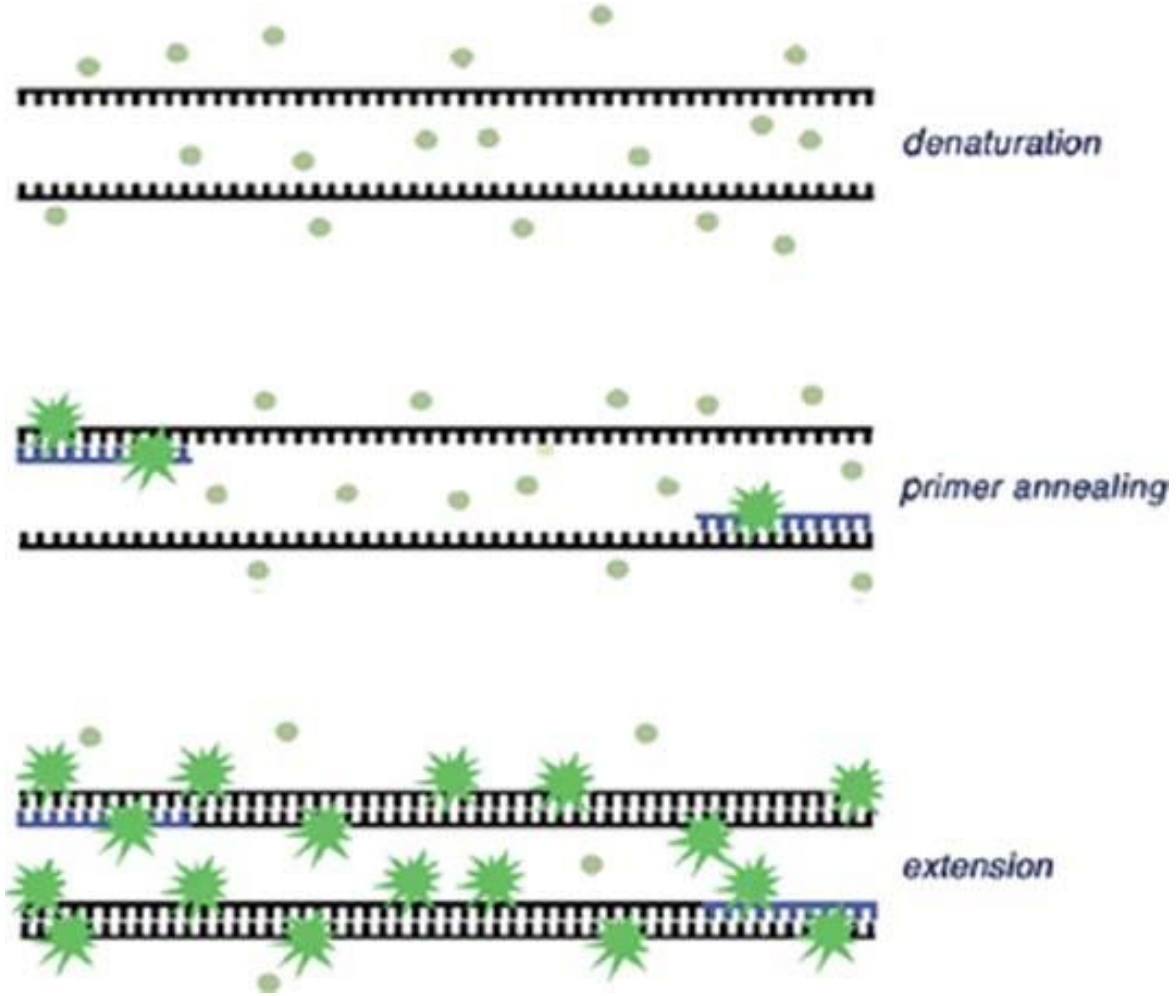
TaqMan labels

5' FLUOROPHORE	Abs [nm]	Em [nm]	3' QUENCHER
FAM [FAM]	495	520	TAM, BHQ1, DAB, Eclip
TET [TET]	521	536	TAM, BHQ1
JOE [JOE]	520	548	TAM, BHQ1, BHQ2
Yakima Yellow [YAKYE]	530	549	BHQ1, Eclip
HEX [HEX]	535	556	TAM, BHQ1, BHQ2, Eclip
Cyanine3 [CY3]	552	570	BHQ1, BHQ2, BBQ650
ATTO 550 [ATTO550]	554	576	TAM, BHQ2
TAMRA [TAM]	544	576	BHQ2
ROX [ROX]	575	602	TAM, BHQ2, BBQ650
Texas Red [TxRed]	583	603	BHQ2, BBQ650
Cyanine3.5 [CY35]	588	604	BHQ2
LightCycler 610 [LC610]	590	610	BHQ2
LightCycler 640 [LC640]	625	640	BHQ2, BBQ650
ATTO 647N [ATTO647N]	644	669	BHQ2, BBQ650
Cyanine5 [CY5]	649	670	BHQ2, BBQ650
Cyanine5.5 [CY55]	675	694	BHQ2, BBQ650
ATTO 680 [ATTO680]	680	700	BBQ650

Table of available dye-quencher combinations.

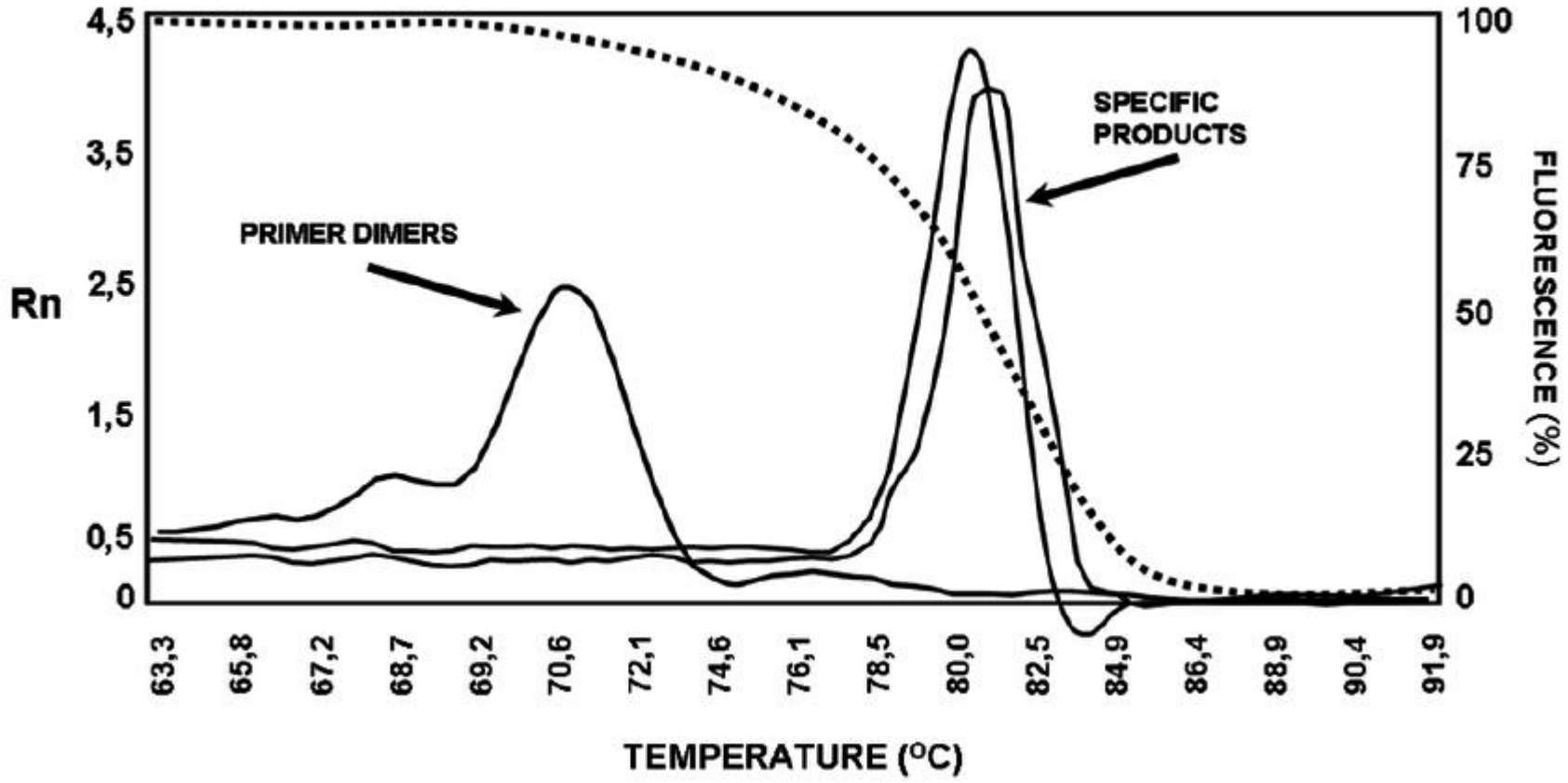
Rotor Gene 5 plex HRM		
Coulour	Source (nm)	Detector (nm)
Green	470	510
Yellow	530	555
Orange	585	610
370/510	530	510
370/555	530	555
Red	625	660
Crimson	680	710
HRM	460	510

SYBR Green system



Melting curve

Temperature at which 50% of the DNA is denaturated



Template

- Ensure high quality of nucleic acid
  - Please prevent inhibitors: excess of phenol, proteinase K, chelanting agents, haemoglobin, SDS, salt
  - Please prevent contamination: amplicons and nucleases
- RNA – cannot be used as template for PCR
  - Synthesis of complementary DNA (cDNA)
  - RT-PCR = reverse transcriptase PCR

Controls for extraction and amplification

- Positive = sample known to have the targeted region
- Negative = sample known to NOT have the target region (Nuclease free water/blank)
- Spike-in = exogenous nucleic acid added in knwon amount into the tested samples and controls before extraction
- Housekeeping = constitutive gene that is always expressed in the tested organism

Controls for amplification

- Standards = serial diluted positive control
- Reverse transcripte negative (RT-) = control of DNA contamination in RNA samples
- No template control (NTC)/blank = water nuclase free instead of the sample. MANDATORY!

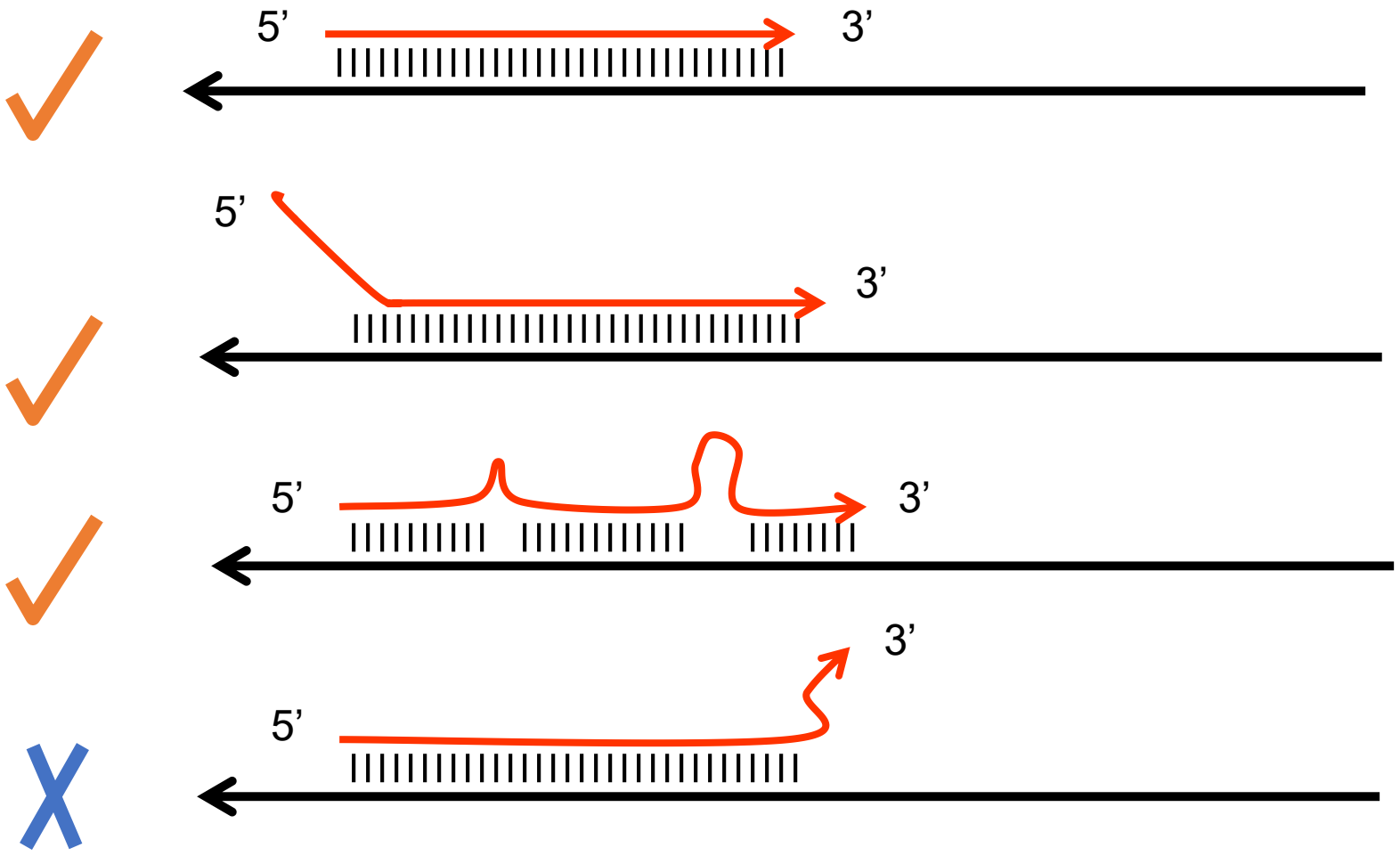
What about the oligos?

- Size ~ 20 base pairs (bp)
- C+G content ~ 50%
- 3’ end should cotain CG
- Similar anneling temperature (Ta)

  Ta = Melting temperature (Tm) – 5 °C

  Tm = [2 x (T+A)] + [4 x (C+G)]

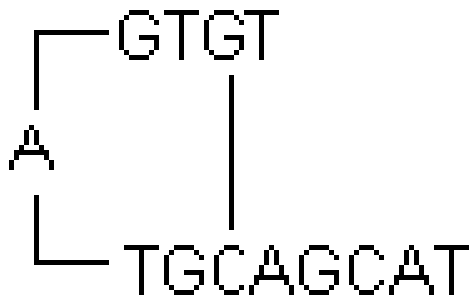
What you want



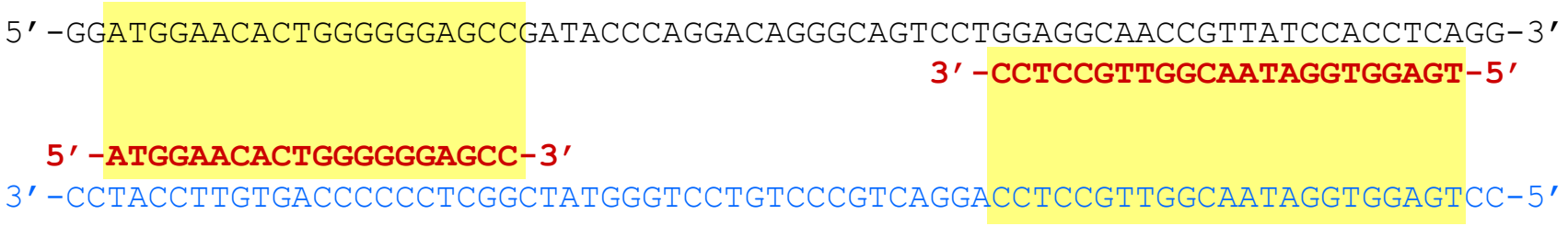
What you want to avoid

- Primer dimer

- Primer hairpin



Reverse complement



Primers (5’ > 3’)

Forward: **ATGGAACACTGGGGGGAGCC**
Reverse: **TGAGGTGGATAACGGTTGCCTCC**

How can I trust in the literature?

Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China

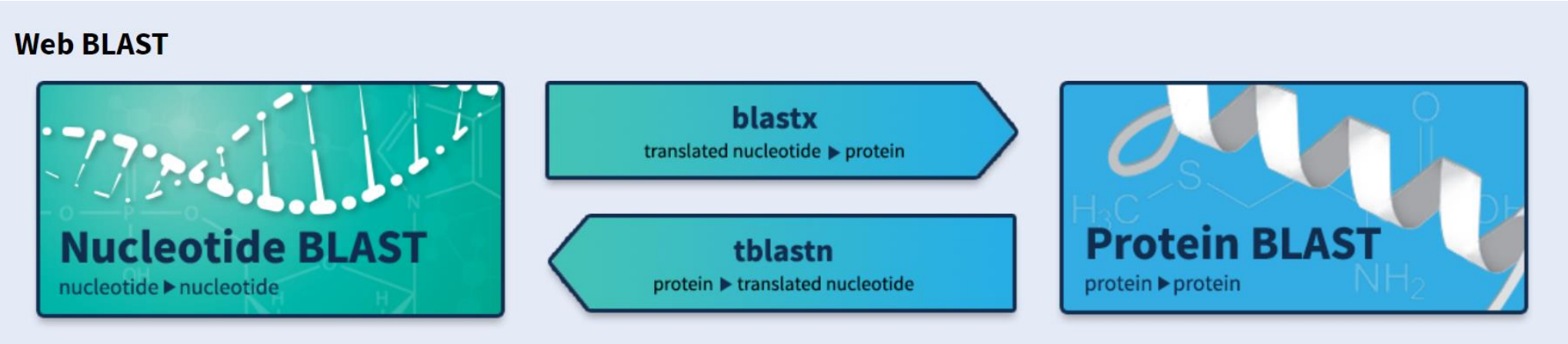
Chaolin Huang, Yeming Wang*, Xingwang Li*, Lili Ren*, Jianping Zhao*, Yi Hu*, Li Zhang, Guohui Fan, Jiuyang Xu, Xiaoying Gu, Zhenshun Cheng, Ting Yu, Jiaan Xia, Yuan Wei, Wenjuan Wu, Xuelei Xie, Wen Yin, Hui Li, Min Liu, Yan Xiao, Hong Gao, Li Guo, Jungang Xie, Guangfa Wang, Rongmeng Jiang, Zhancheng Gao, Qi Jin, Jianwei Wang†, Bin Cao†*

Procedures

Local centres for disease control and prevention collected respiratory, blood, and faeces specimens, then shipped them to designated authoritative laboratories to detect the pathogen (NHC Key Laboratory of Systems Biology of Pathogens and Christophe Mérieux Laboratory, Beijing, China). A novel coronavirus, which was named 2019-nCoV, was isolated then from lower respiratory tract specimen and a diagnostic test for this virus was developed soon after that.[14] Of 59 suspected cases, 41 patients were confirmed to be infected with 2019-nCoV. The presence of 2019-nCoV in respiratory specimens was detected by next-generation sequencing or real-time RT-PCR methods. The primers and probe target to envelope gene of CoV were used and the sequences were as follows: forward primer 5′-TCAGAATGCCAATCTCCCCAAC-3′; reverse primer 5′-AAAGGTCCACCCGATACATTGA-3′; and the probe 5′CY5-CTAGTTACACTAGCCATCCTTACTGC-3′BHQ1. Conditions for the amplifications were 50°C for 15 min, 95°C for 3 min, followed by 45 cycles of 95°C for 15 s and 60°C for 30 s.

BLAST

- Basic Local Alignment Search Tool
  - Finds regions of similarity between biological samples



Identity

- **The higher the percent identity is, the more significant the match.**

- The percent identity is a number that describes how similar the query sequence is to the target sequence (how many characters in each sequence are identical).

E-value

- **The BLAST E-value is the number of expected hits of similar quality (score) that could be found just by chance.**
- E-value of 10 means that up to 10 hits can be expected to be found just by chance, given the same size of a random database.
- E-value can be used as a first quality filter for the BLAST search result

Score

- **The higher the bit-score, the better the sequence similarity**

- The bit-score is the requires size of a sequence database in which the current match could be found just by chance.

- Bit-score does not depend on database size. The bit-score gives the same value for hits in databases of different sizes and hence can be used for searching in an constantly increasing database.

Testing published primers/probe

Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China

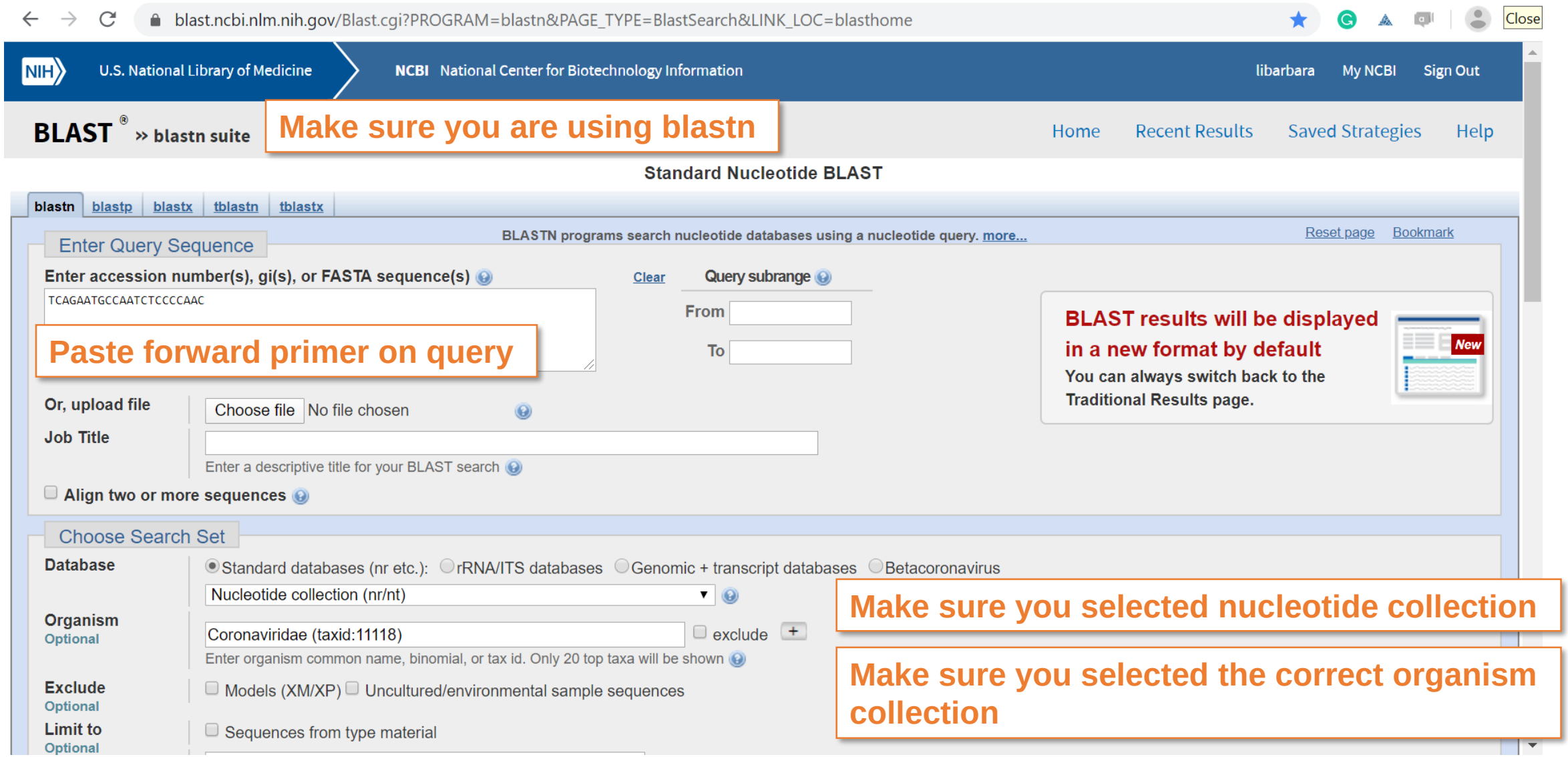
Chaolin Huang, Yeming Wang*, Xingwang Li*, Lili Ren*, Jianping Zhao*, Yi Hu*, Li Zhang, Guohui Fan, Jiuyang Xu, Xiaoying Gu, Zhenshun Cheng, Ting Yu, Jiaan Xia, Yuan Wei, Wenjuan Wu, Xuelei Xie, Wen Yin, Hui Li, Min Liu, Yan Xiao, Hong Gao, Li Guo, Jungang Xie, Guangfa Wang, Rongmeng Jiang, Zhancheng Gao, Qi Jin, Jianwei Wang†, Bin Cao†*

Procedures

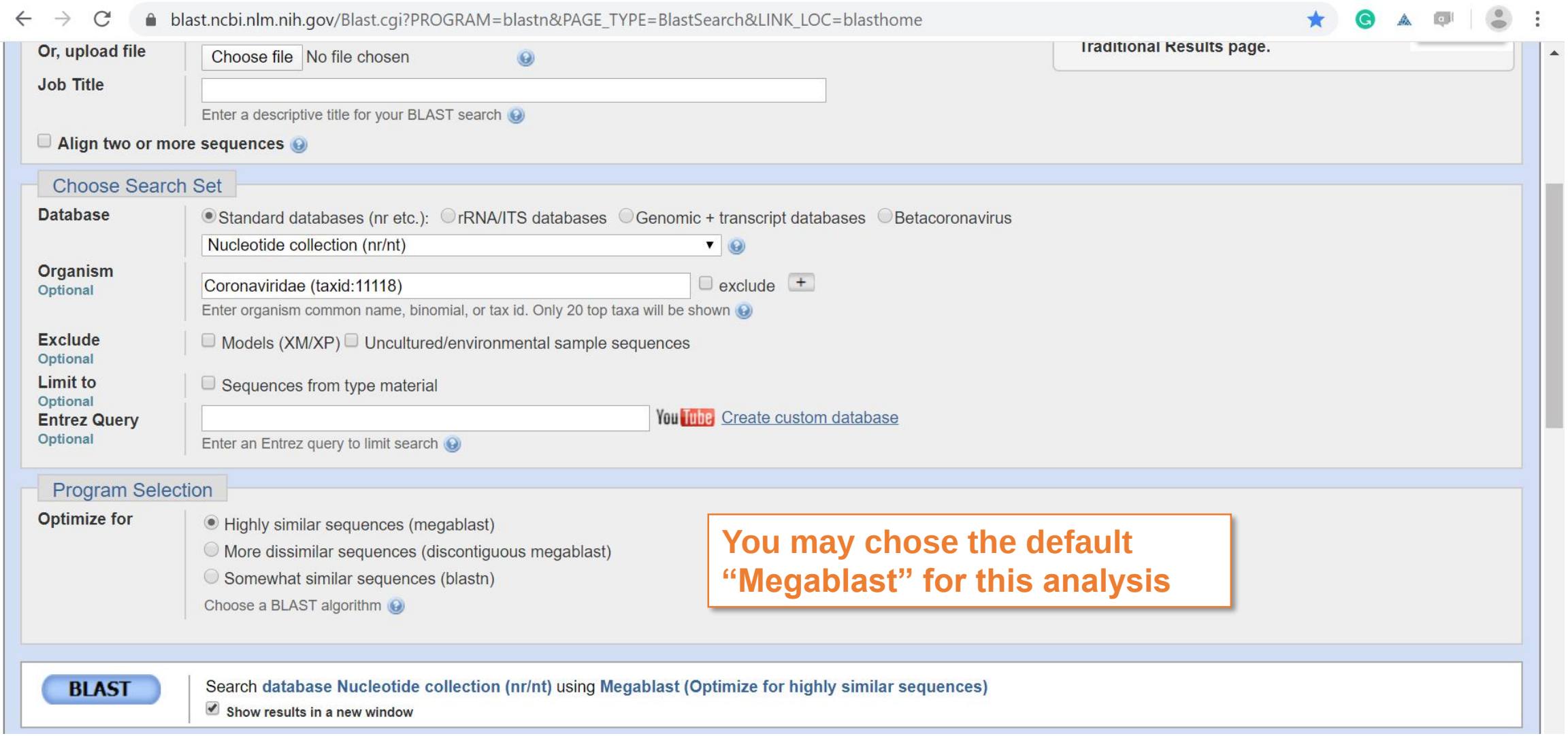
Local centres for disease control and prevention collected respiratory, blood, and faeces specimens, then shipped them to designated authoritative laboratories to detect the pathogen (NHC Key Laboratory of Systems Biology of Pathogens and Christophe Mérieux Laboratory, Beijing, China). A novel coronavirus, which was named 2019-nCoV, was isolated then from lower respiratory tract specimen and a diagnostic test for this virus was developed soon after that.[14] Of 59 suspected cases, 41 patients were confirmed to be infected with 2019-nCoV. The presence of 2019-nCoV in respiratory specimens was detected by next-generation sequencing or real-time RT-PCR methods. The primers and probe target to envelope gene of CoV were used and the sequences were as follows: forward primer 5′-TCAGAATGCCAATCTCCCCAAC-3′; reverse primer 5′-AAAGGTCCACCCGATACATTGA-3′; and the probe 5′CY5-CTAGTTACACTAGCCATCCTTACTGC-3′BHQ1. Conditions for the amplifications were 50°C for 15 min, 95°C for 3 min, followed by 45 cycles of 95°C for 15 s and 60°C for 30 s.

- Copy only the nucleotides
- Forward primer and probe are in sense direction
- Reverse primer is in reverse complementary direction

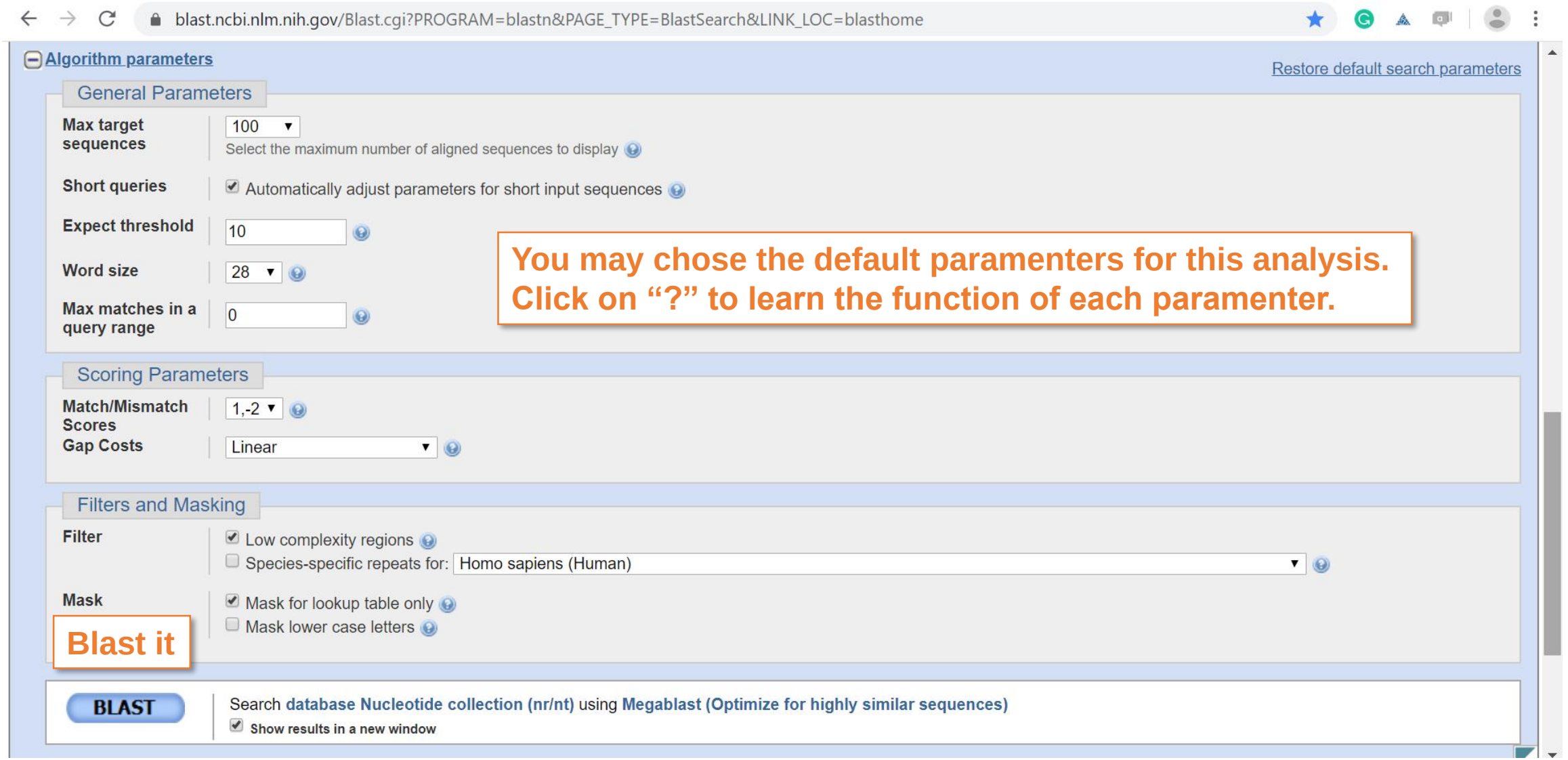
Paste in query



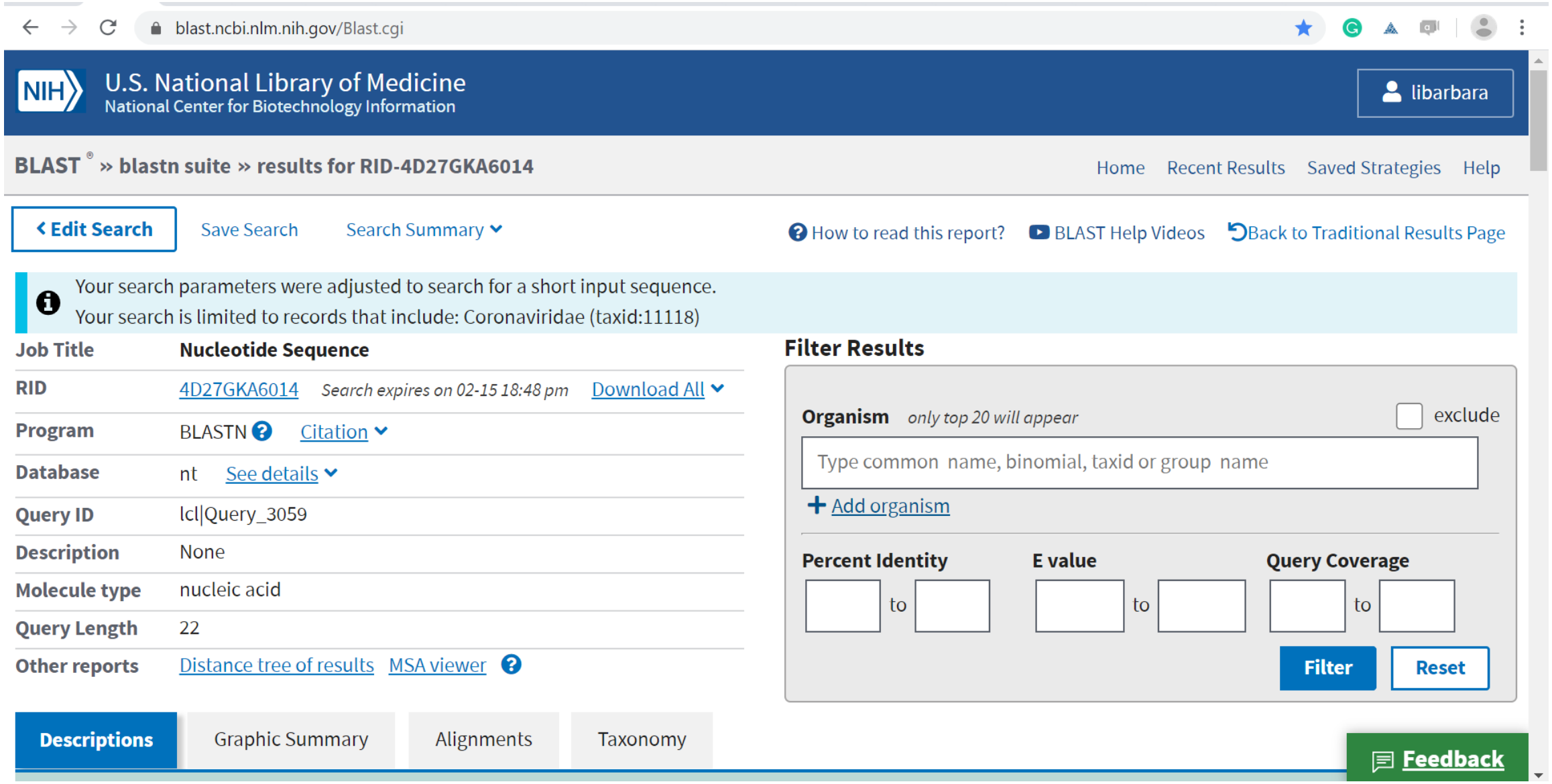
Bottom of the query page



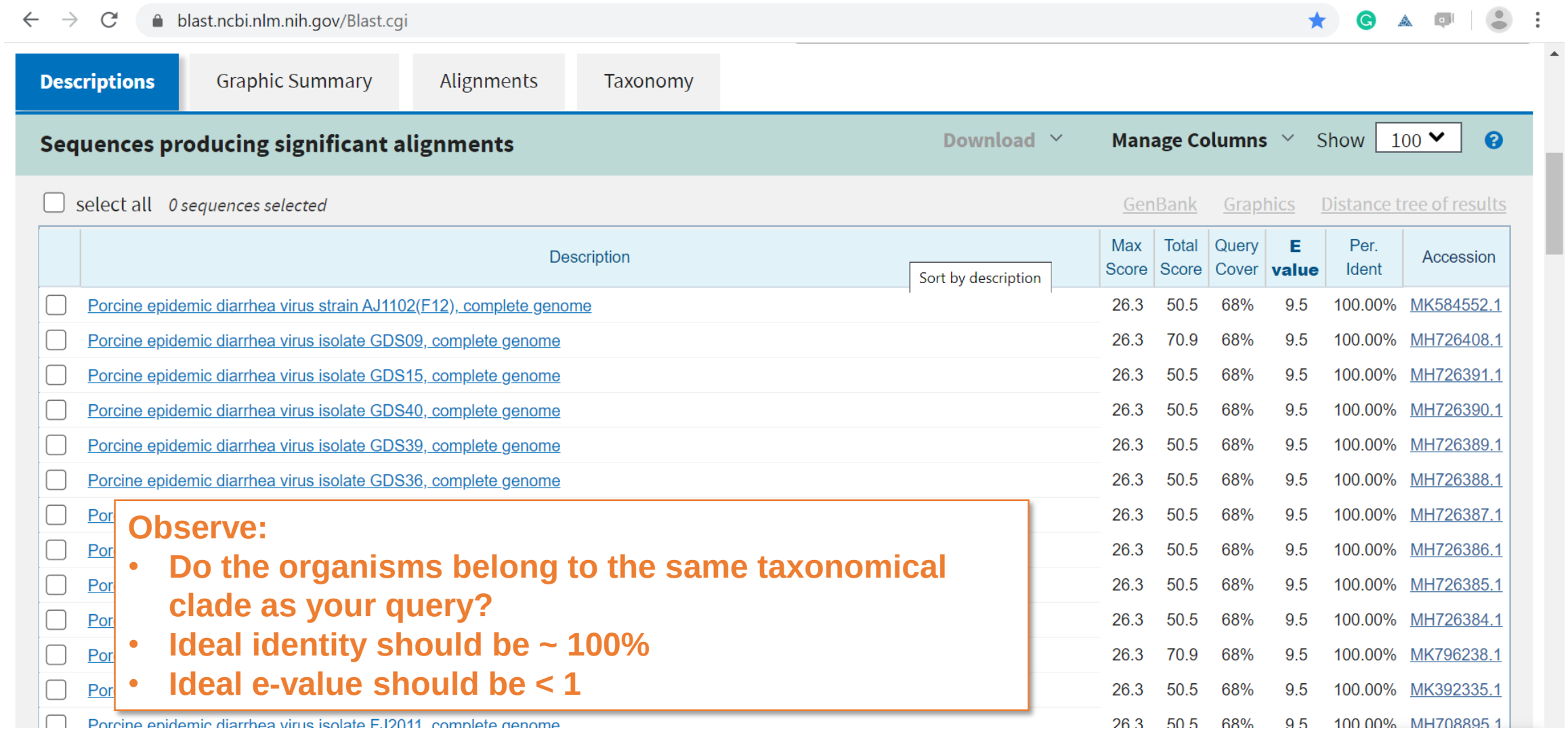
Addional parameters



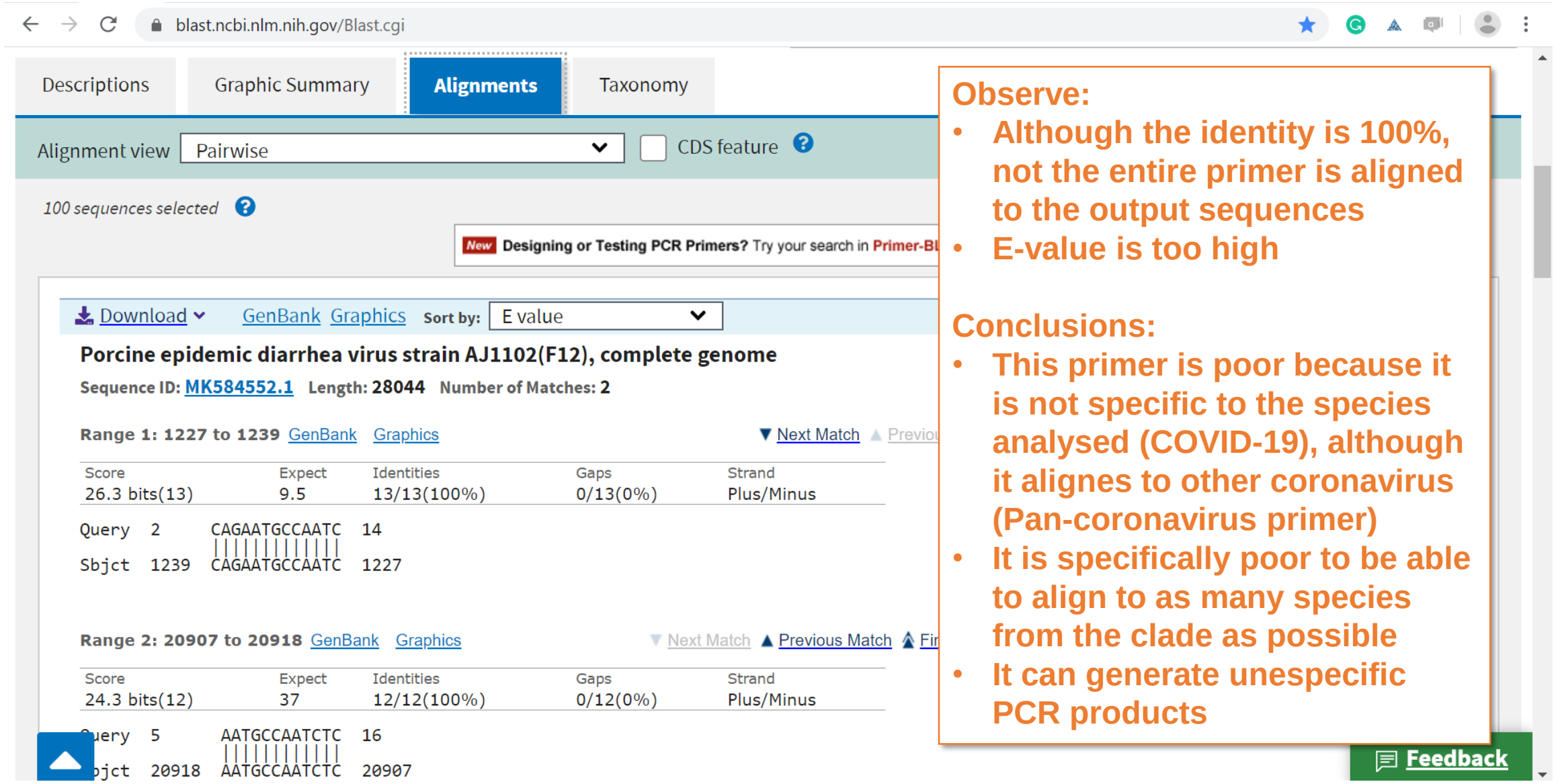
Output - top page (1 of 3)



Descriptions tab (2 of 3)



Alignments tab (3 of 3)



Observe:

- Although the identity is 100%, not the entire primer is aligned to the output sequences
- E-value is too high

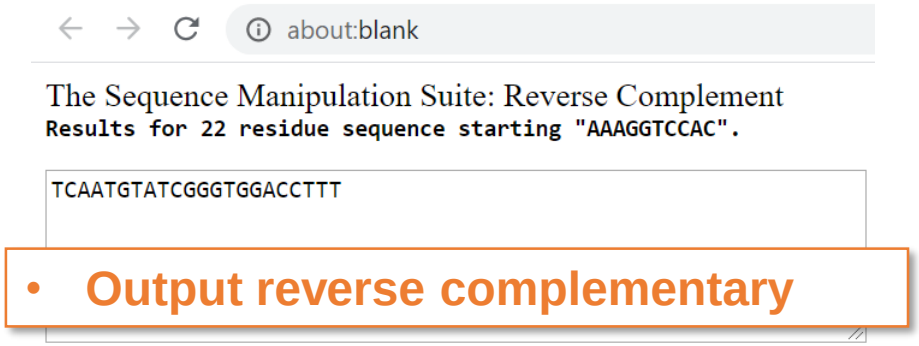
Conclusions:

- This primer is poor because it is not specific to the species analysed (COVID-19), although it alignes to other coronavirus (Pan-coronavirus primer)
- It is specifically poor to be able to align to as many species from the clade as possible
- It can generate unespecific PCR products

Reverse primer

- Must be converted to “reverse complementary” before searching it on the query.

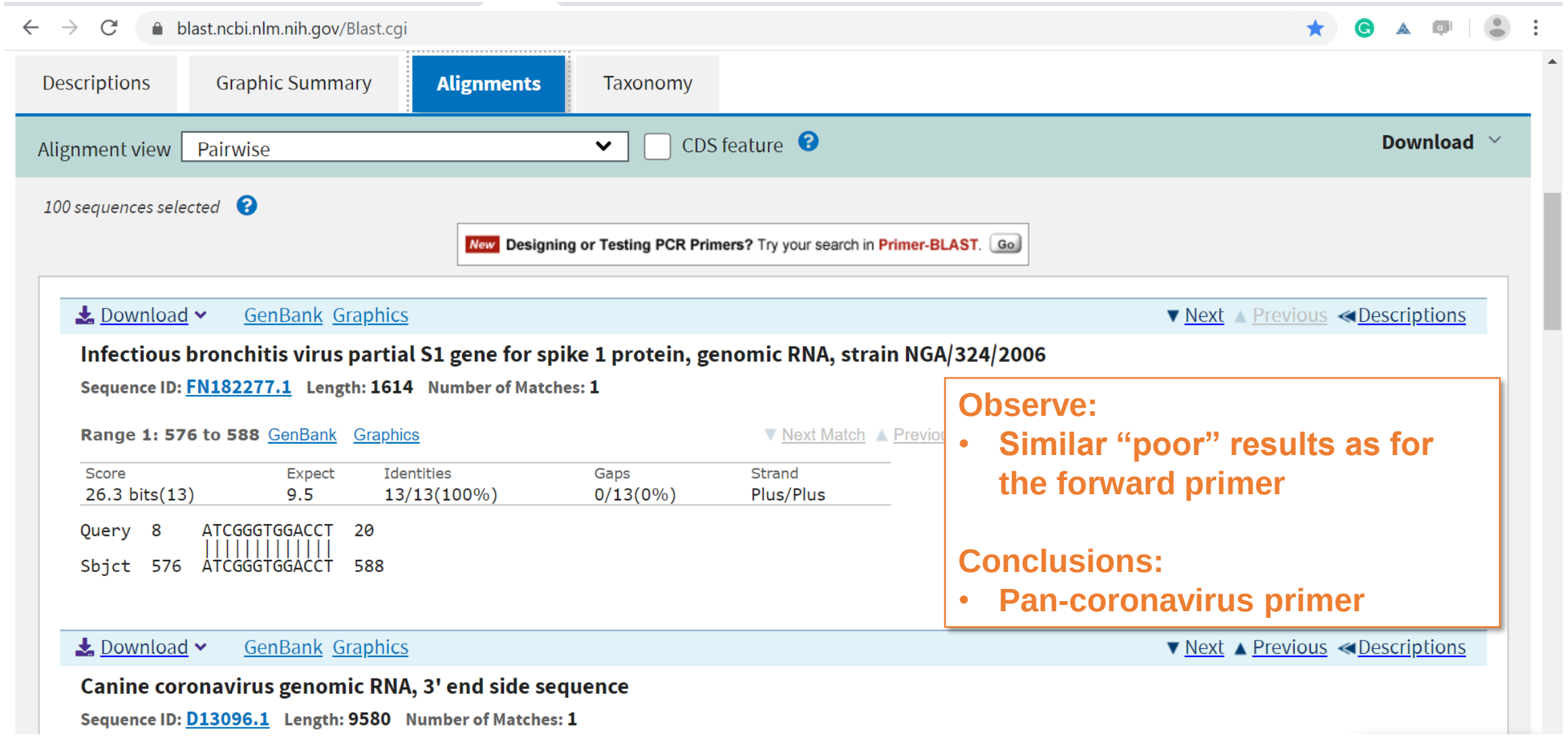
- **Input reverse primer**



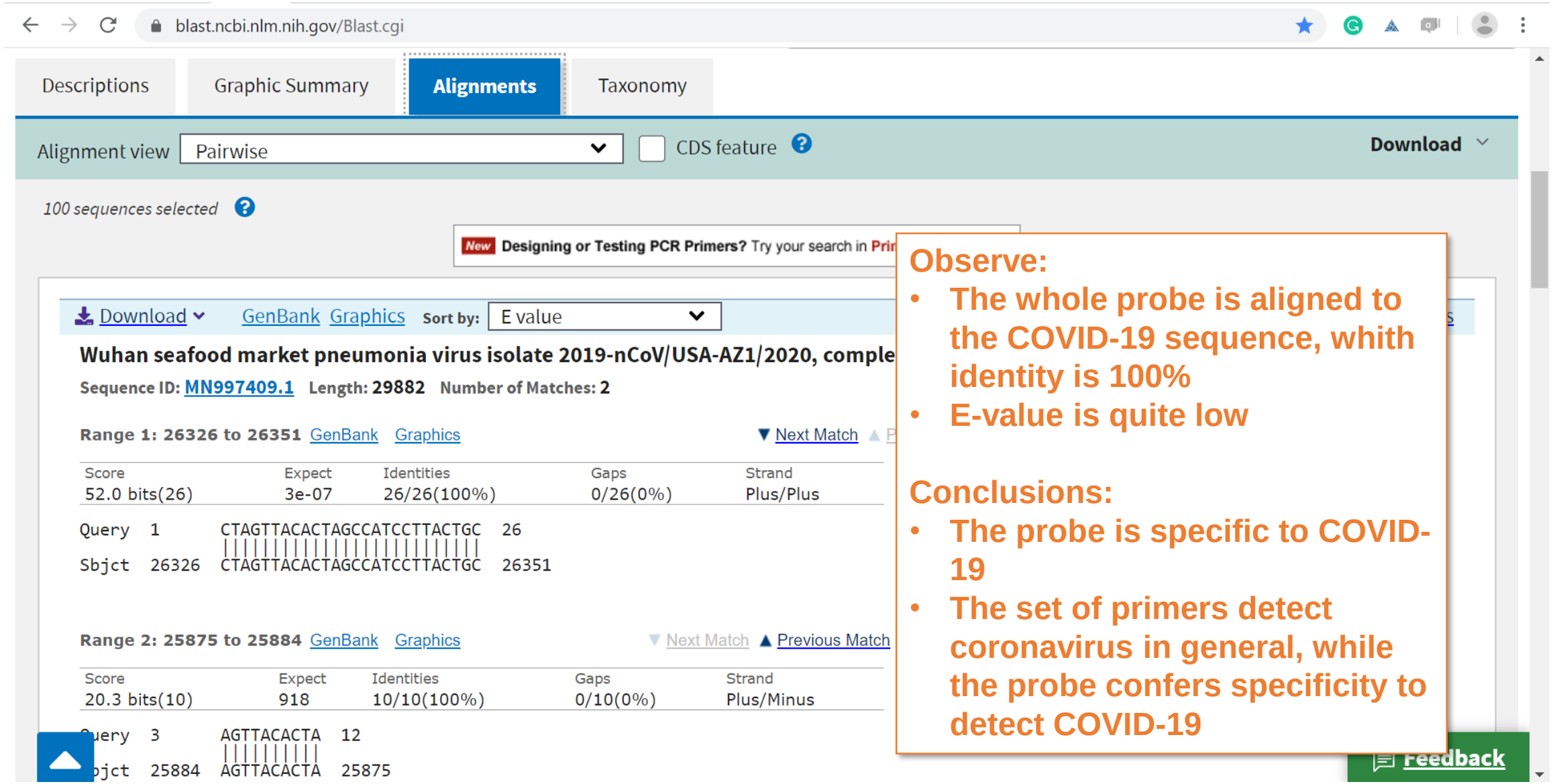
- **Output reverse complementary**

- **Copy/paste the reverse primer and “submit” it**
- **Copy the output and paste on the blastn query like for the forward primer**
- **Repeat the analyses**

Reverse primer output



Repeat the process using the probe



Observe:

- **The whole probe is aligned to the COVID-19 sequence, whith identity is 100%**
- **E-value is quite low**

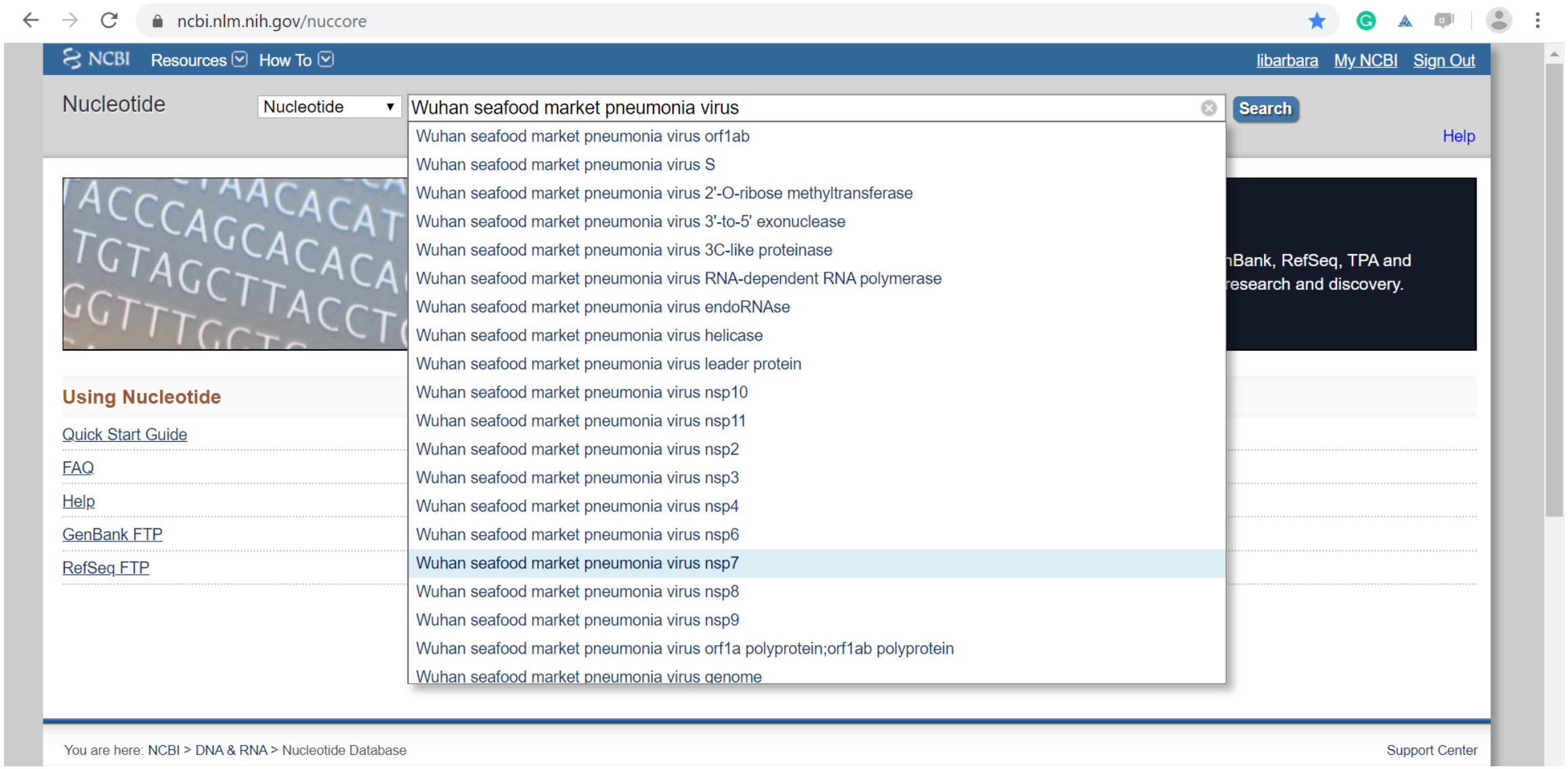
Conclusions:

- **The probe is specific to COVID-19**
- **The set of primers detect coronavirus in general, while the probe confers specificity to detect COVID-19**

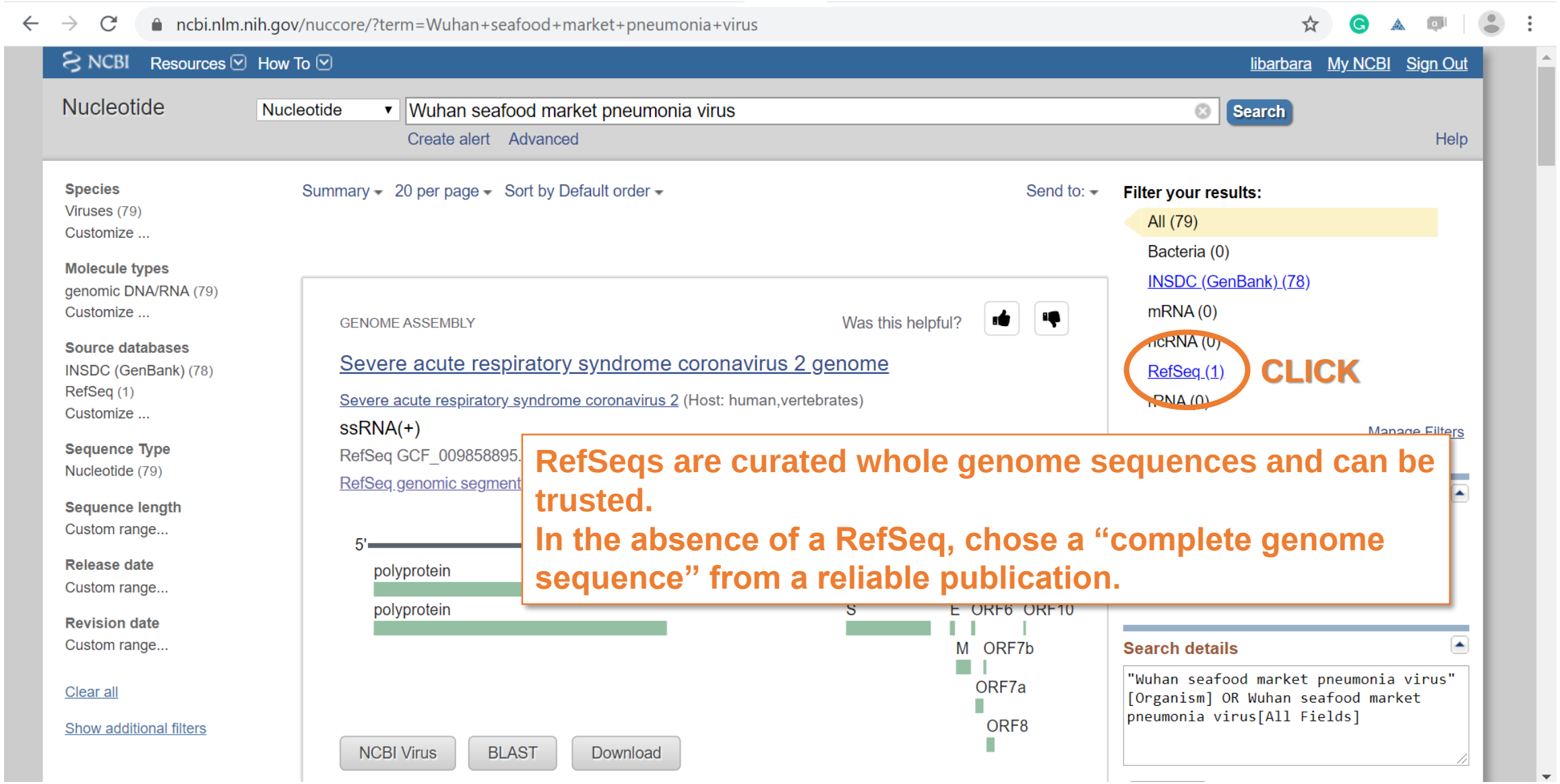
Designing your own primers/probes

- Use primer-BLAST or any other designing tool
- Always use a Reference Sequence (RefSeq) as template to desing your oligos
- Use GenBank tools to get RefSeqs
- Double-check your designed oligos testing them using the same procedures as described previously

Search for sequences on the nucleotide database

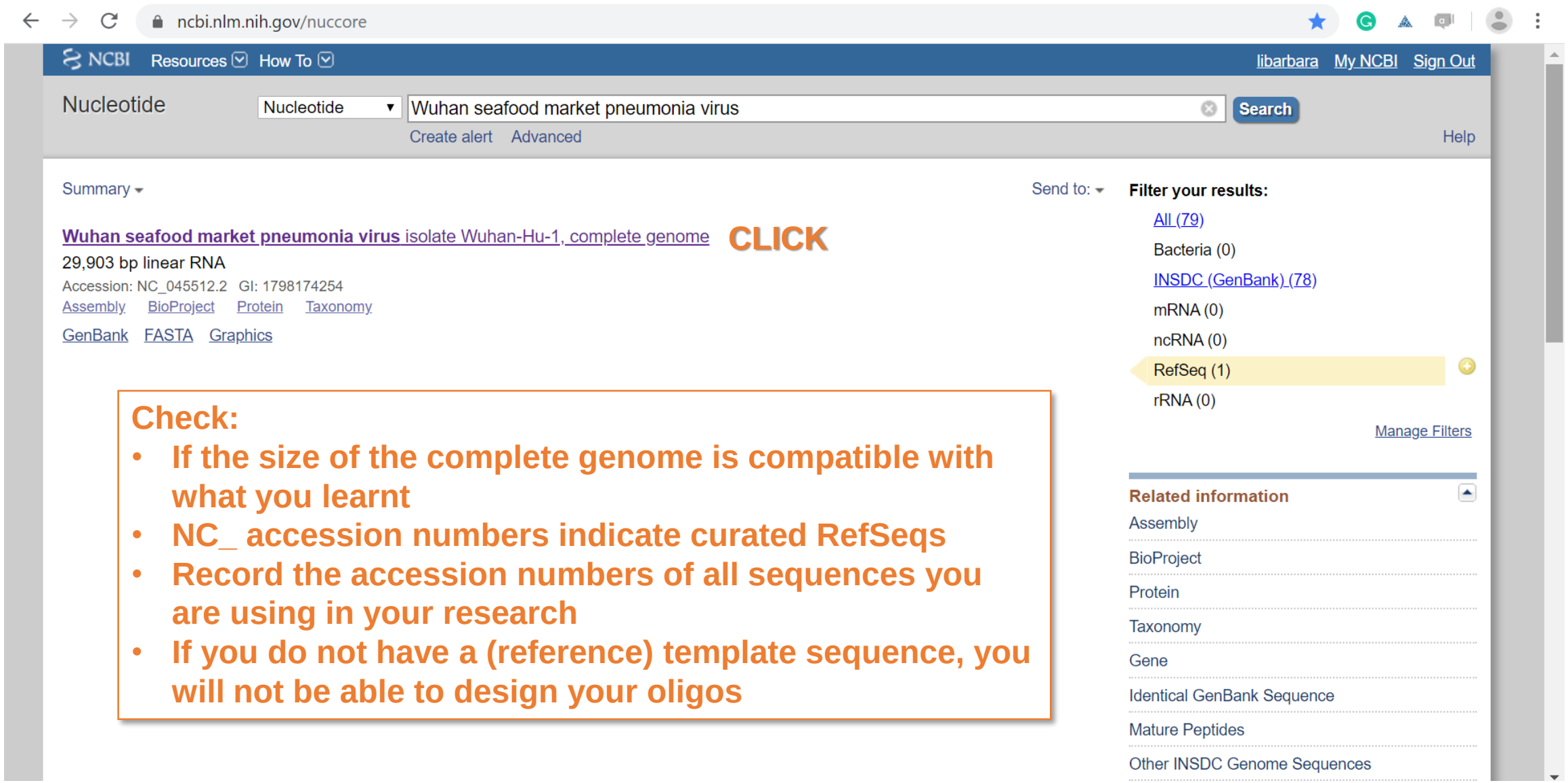


Select preferably RefSeq sequences



RefSeqs are curated whole genome sequences and can be trusted.
In the absence of a RefSeq, chose a "complete genome sequence" from a reliable publication.

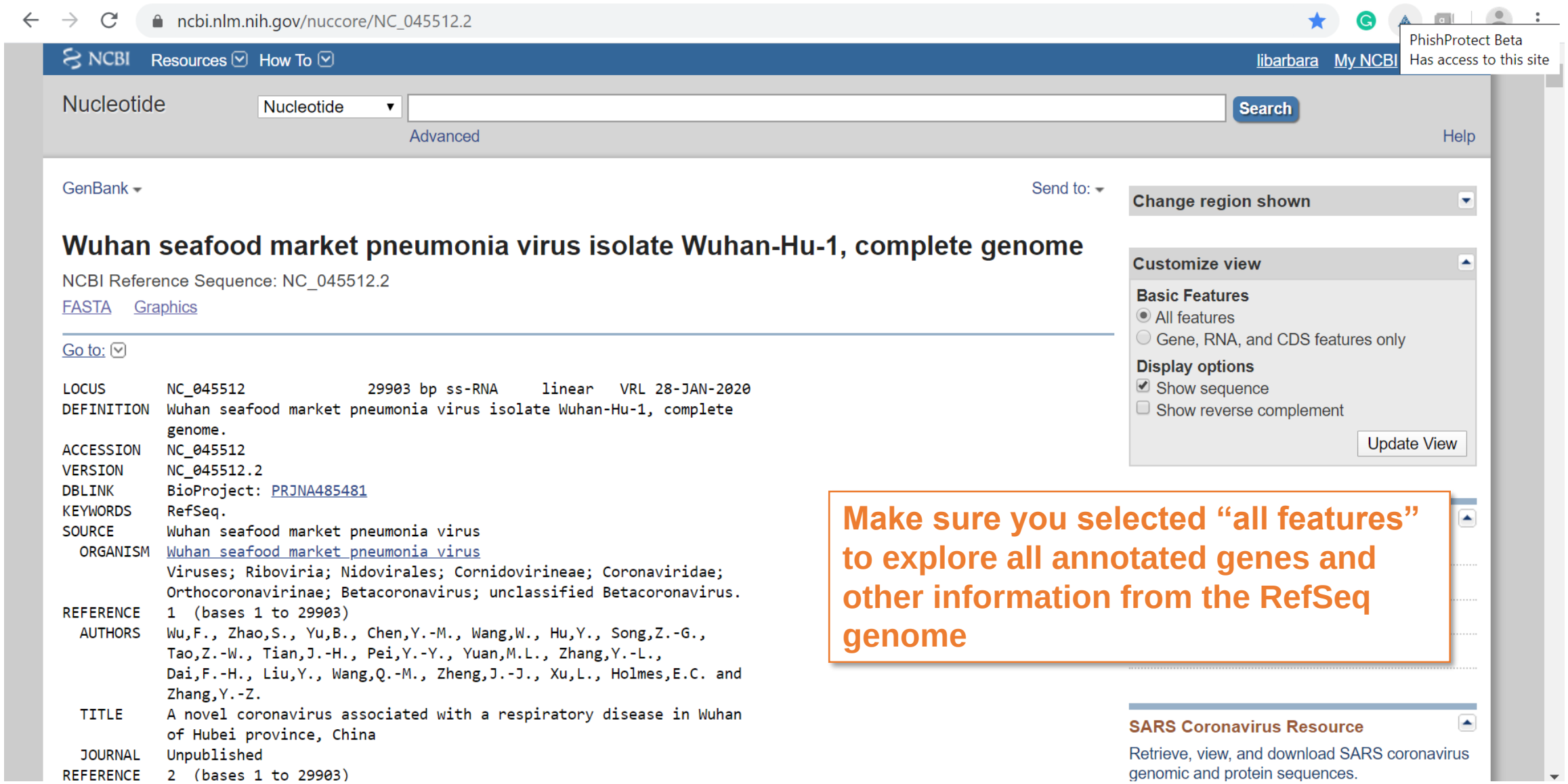
Check the information you know



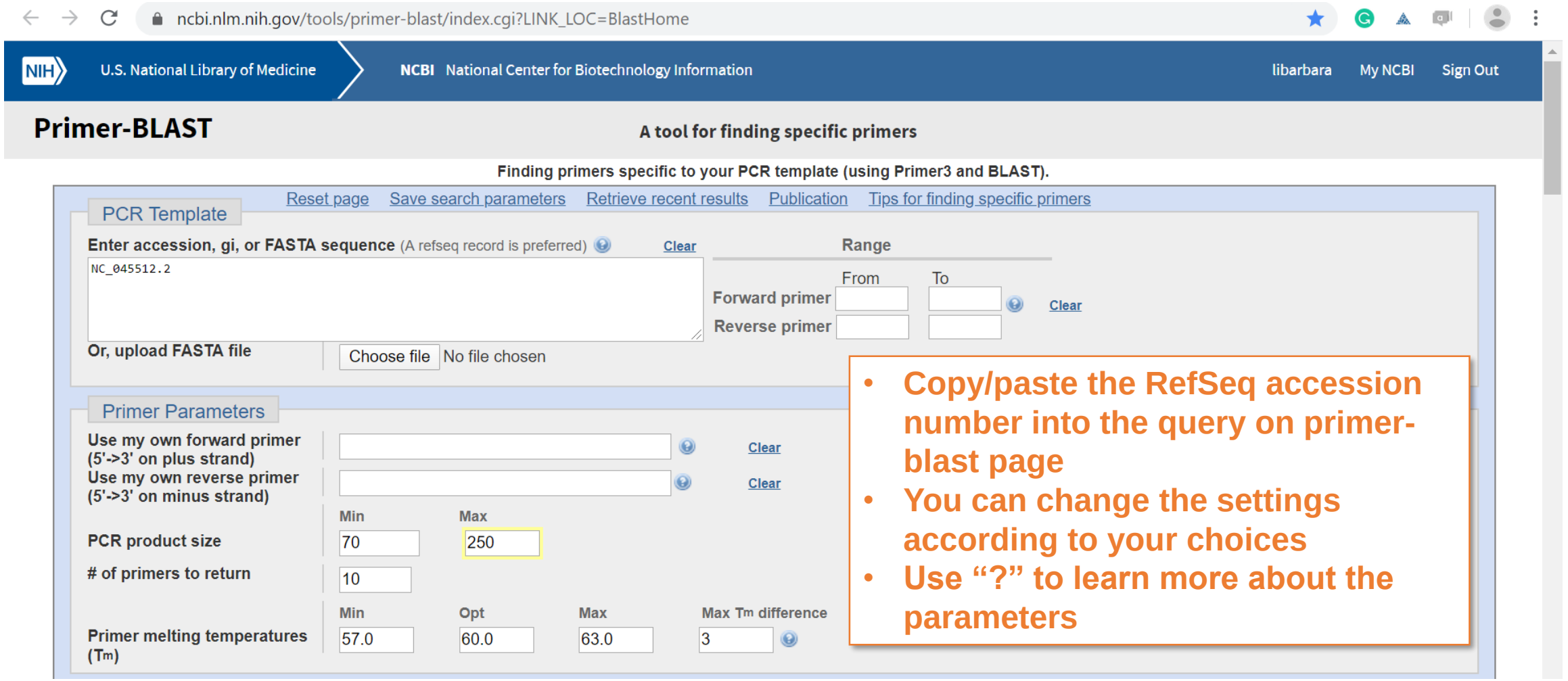
Check:

- **If the size of the complete genome is compatible with what you learnt**
- **NC_ accession numbers indicate curated RefSeqs**
- **Record the accession numbers of all sequences you are using in your research**
- **If you do not have a (reference) template sequence, you will not be able to design your oligos**

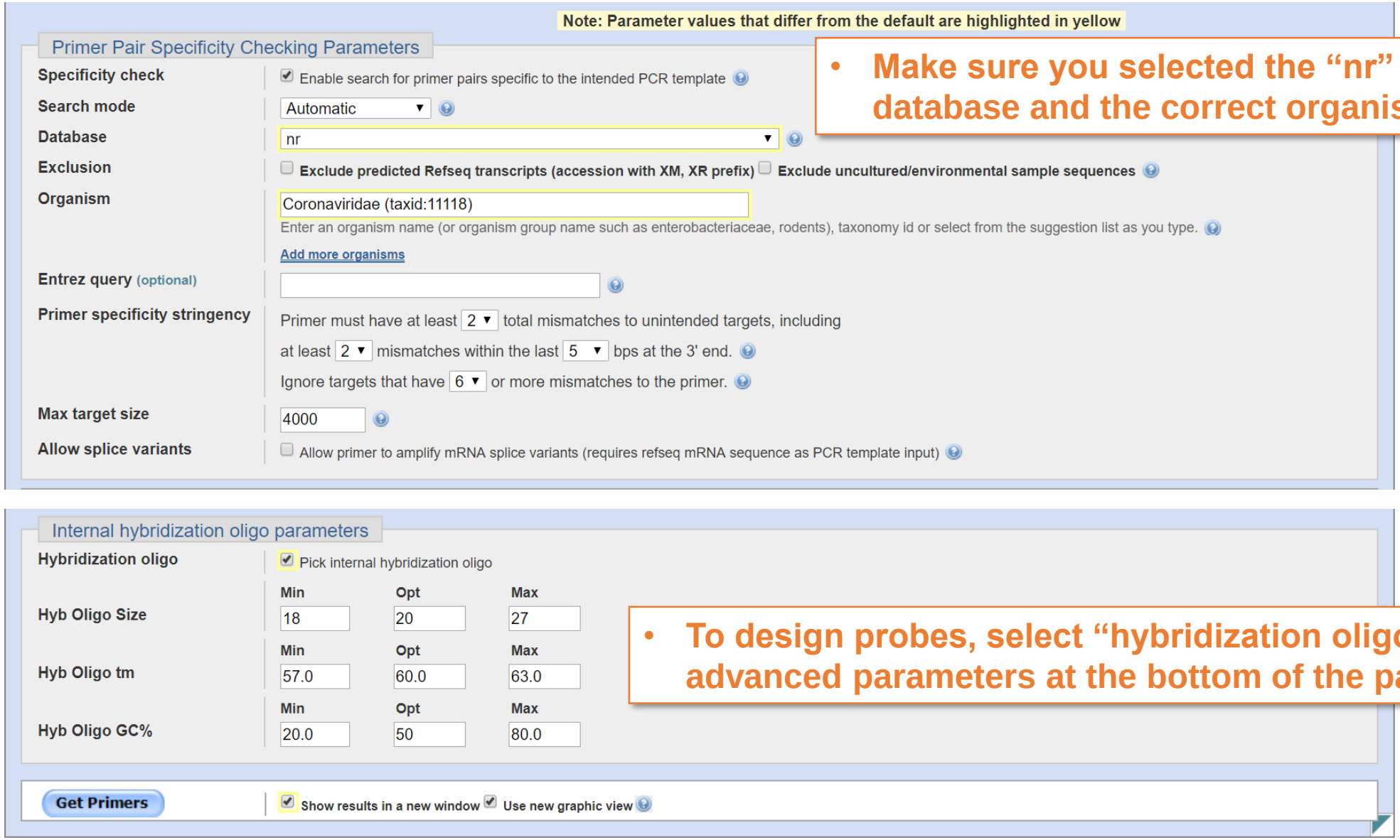
Learn more from the organism genome



Primer-BLAST query (1 of 2)



Primer-BLAST query (2 of 2)



- **Make sure you selected the “nr” database and the correct organism taxid**

- **To design probes, select “hybridization oligo” on advanced parameters at the bottom of the page**

Addional information

ncbi.nlm.nih.gov/tools/primer-blast/primertool.cgi

NIH U.S. National Library of Medicine | NCBI National Center for Biotechnology Information | libarbara | My NCBI | Sign Out

Primer-BLAST

A tool for finding specific primers

Finding primers specific to your PCR template (using Primer3 and BLAST)

Input PCR template NC_045512.2 Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1, complete genome
Range 1 - 29903

Your PCR template is highly similar to the following sequence(s) from the search database. To increase the chance of finding specific primers, please review the list below and select all sequences (within the given sequence ranges) that are intended or allowed targets.

Select: All None Selected:2

Accession	Title	Identity	Alignment length	Seq. start	Seq. stop
MN988669.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV WHU02, complete genome	100%	29881	1	29881
MN988668.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV WHU01, complete genome	100%	29881	1	29881
MN994468.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-CA2/2020, complete genome	99.99%	29883	1	29883
MN975262.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV_HKU-SZ-005b_2020, complete genome	99.98%	29891	1	29891
MN985325.1	Severe acute respiratory syndrome coronavirus 2 isolate 2019-nCoV/USA-WA1/2020, complete genome	99.99%	29882	1	29882
MN997409.1	Severe acute respiratory syndrome corona			1	29882
MN988713.1	Severe acute respiratory syndrome corona			1	29882
MN994467.1	Severe acute respiratory syndrome corona			1	29882
MN938384.1	Severe acute respiratory syndrome corona			1	29838

In some cases primer-BLAST may ask you to confirm addional information to narrow down the analyses

Submit | Show results in a new window

Predicted oligos output (1 of 2)

ncbi.nlm.nih.gov/tools/primer-blast/primertool.cgi?ctg_time=1581693560&job_key=4ug945cjmou9sQq0B9Quhn3PP7RQ3CSpUQ

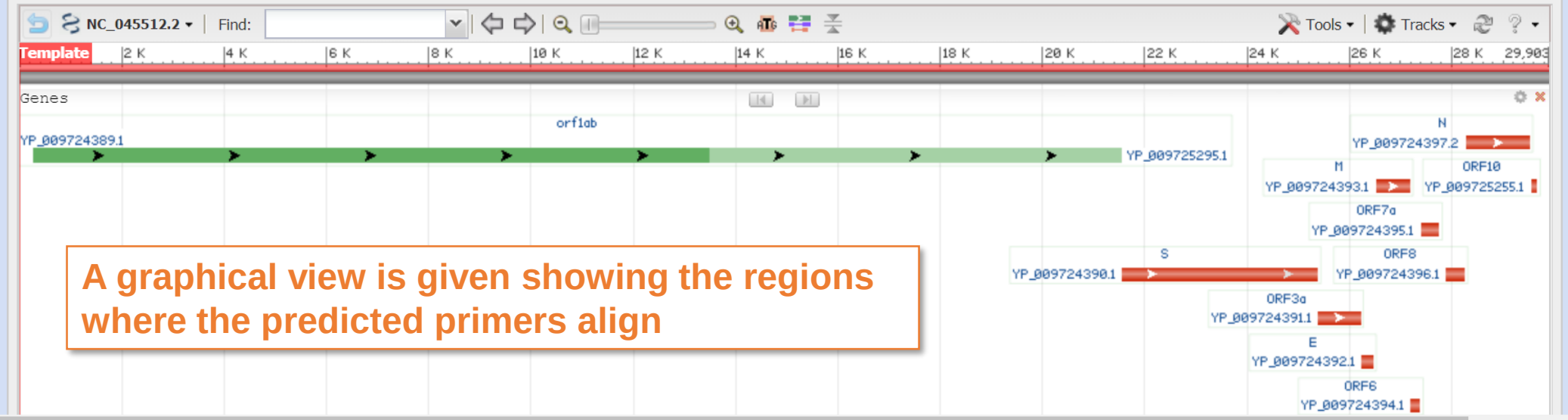
NIH U.S. National Library of Medicine | NCBI National Center for Biotechnology Information | libarbara | My NCBI | Sign Out

Primer-BLAST » JOB ID:4ug945cjmou9sQq0B9Quhn3PP7RQ3CSpUQ

Primer-BLAST Results

Input PCR template	NC_045512.2 Wuhan seafood market pneumonia virus isolate Wuhan-Hu-1, complete genome
Range	1 - 29903
Specificity of primers	Primers may **not** be specific to the input PCR template as targets were found in selected database:Nucleotide collection (nt) (Organism limited to Coronaviridae)...help on specific primers
Other reports	Search Summary

Graphical view of primer pairs



A graphical view is given showing the regions where the predicted primers align

Predicted oligos output (2 of 2)

ncbi.nlm.nih.gov/tools/primer-blast/primertool.cgi?ctg_time=1581693560&job_key=4ug945cjmou9sQq0B9Quhn3PP7RQ3CSpUQ

Detailed primer reports

You can re-search for specific primers by accepting some of the unintended targets, check the box(es) next to the ones you accept and try again to re-search for specific primers Submit

Primer pair 1

	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	ACCTTCCCAGGTAACAAACCA	Plus	21	14	34	59.15	47.62	5.00	0.00
Reverse primer	ACTCGTGTCCTGTCAACGAC	Minus	20	168	149	59.97	55.00	4.00	3.00
Internal oligo	GTGGCTGTCACTCGGCTGCA	Plus	20	88	107	59.63	65.00		
Product length	155								

Products on intended targets

>MN988669.1 Wuhan seafood market pneumonia virus isolate 2019-nCoV WHU0

```
product length = 155
Forward primer  1   ACCTTCCCAGGTAACAAACCA  21
Template        13  ....................  33

Reverse primer  1    ACTCGTGTCCTGTCAACGAC  20
Template        167  ....................  148
```

>MN988668.1 Wuhan seafood market pneumonia virus isolate 2019-nCoV WHU0

```
product length = 155
Forward primer  1   ACCTTCCCAGGTAACAAACCA  21
Template        13  ....................  33

Reverse primer  1    ACTCGTGTCCTGTCAACGAC  20
Template        167  ....................  148
```

A few oligo set options will be given. Chose the best result.

A good oligo set should have:

- **GC% (GC content) > 50%**
- **Similar Tm (melting temperature) for all oligos in the set**
- **Self complementarity score near 0.00**

Your new best friends

- ViralZone https://viralzone.expasy.org/
- International Committee on Taxonomy of Viruses https://talk.ictvonline.org/taxonomy/
- Pubmed https://www.ncbi.nlm.nih.gov/pubmed/
- GenBank https://www.ncbi.nlm.nih.gov/genbank/
- Blast https://blast.ncbi.nlm.nih.gov/Blast.cgi