**Program Use Notes**

**Introduction**

1. The main program “UniversalPrimerMaker” takes in information about gene sequences and primers .It returns a list highlighting what primers are eligible to be used as universal primers and gives their degenerate counterparts. “UniversalPrimerMaker” is designed to speed up the universal primer design process ; once more than 3 primers are being compared it tends to be quicker to use “UniversalPrimerMaker”.(COL1A2 had >350 primer choices)
2. “UniversalPrimerMaker” has in theory no limit to how many species can be compared, however, is currently being validated for 4 species (3 of same family , 1 of same order) so would recommend using it at this number.
3. For “UniversalPrimerMaker” to work 6 files are required , obtaining the files will be explained by the next sections. The required files are as follows.

-A file containing the Alignment of the gene in each species of interest **(alignment.fasta)**

-A file containing designed primers and their related information **(primerinformation.xls)**

-A file containing just the designed primer names and sequences **(rawprimers.txt)**

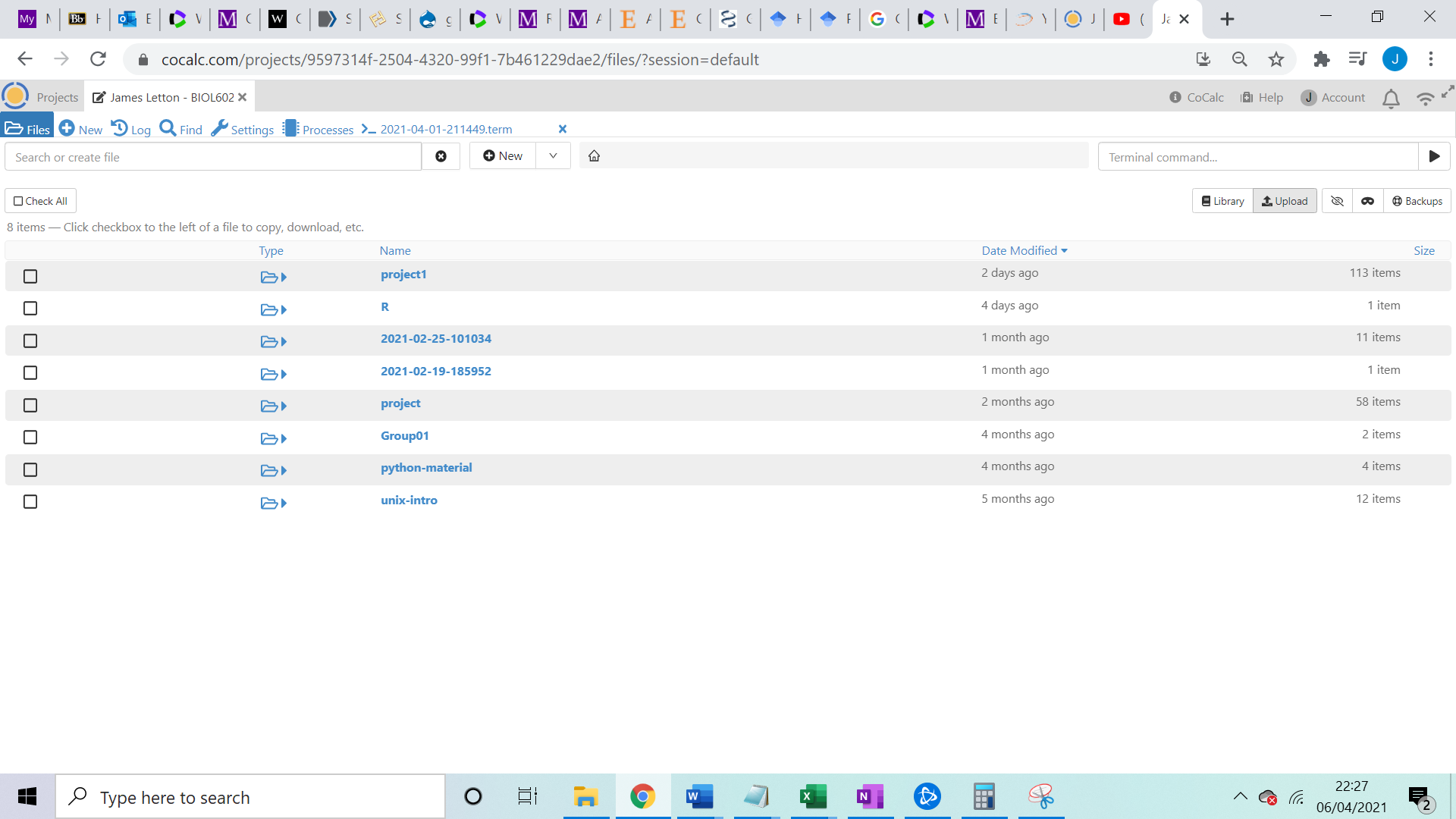
-A file containing all the blast information for all designed primers filter for only the reference species **(blastresults.txt)**

-A file containing all the blast information for all designed primers primers filter for only the human **(blastresults.txt)**

-A file containing the desired settings of the program **(settings.txt)**

**Upload program files to CoCalc (other python environments are valid)**

1. CoCalc is a web application (<https://cocalc.com/app>) that allows code to be run and written. The big advantage to it is that no installation is required.
2. Upload the contents of the program folder to Cocalc .



-follow the same process with the files mentioned below making sure they are all located in the same folder

**Obtain gene sequence (alignment.fasta)**

1. Retrieve gene sequences for each species of interest .Sites such as Ensembl and NCBI are two options for obtaining the sequences (recommend including 3000 bp of untranslated region in each direction) .
2. Align the sequences using Clustal Omega to align the sequences (<https://www.ebi.ac.uk/Tools/msa/clustalo/> ).

-Paste the sequence information into the text box

-When entering the sequences in Clustal make sure the species name is included in the sequence name eg “>cow”

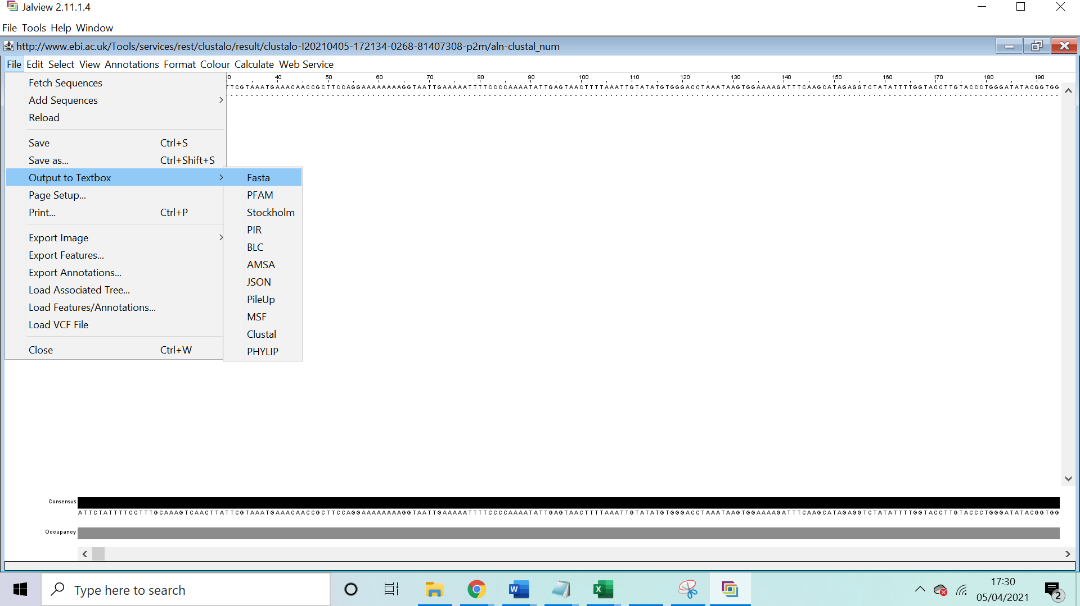
-The output should be ClustalW or ClustalW with characters

1. Export the alignment as a fasta file .The following instructions will show how to do this in Jalview however all alignment programs should have this option .

-if you have not already, download and install jalview.

- In clustal omega open the results viewer tab and then click view results with Jalview (sometimes helps to open it in a separate window and then refresh , OPENING AS TAB DOES NOT WORK)

-The location of the option to download the alignment as a fasta file in Jalview is shown below.



-Once outputted to textbox click file and save.

1. For the program to recognise the file, save the file in the same folder as “UniversalPrimerMaker” and give the file the exact name “**alignment.fasta”.**

**Primer design (primerinformation.xls)**

1. For this section using Primer3 (not Primer3+ as tends to time out at higher numbers of primers) for design will be described and is recommended. Other methods can be used but will require more formatting which will be explained later.
2. Enter the entire gene sequence of the animal that is being used as a reference into primer3 and change the primer selection found at the top of the page to “pick\_sequencing\_primers”.
3. Depending on the gene ,the settings will have to be altered to maximise the number of eligible primers . The following setting stated below recommended to get a range of primers and are set with the gene repetitive in mind. Rational for some parameters are stated below.

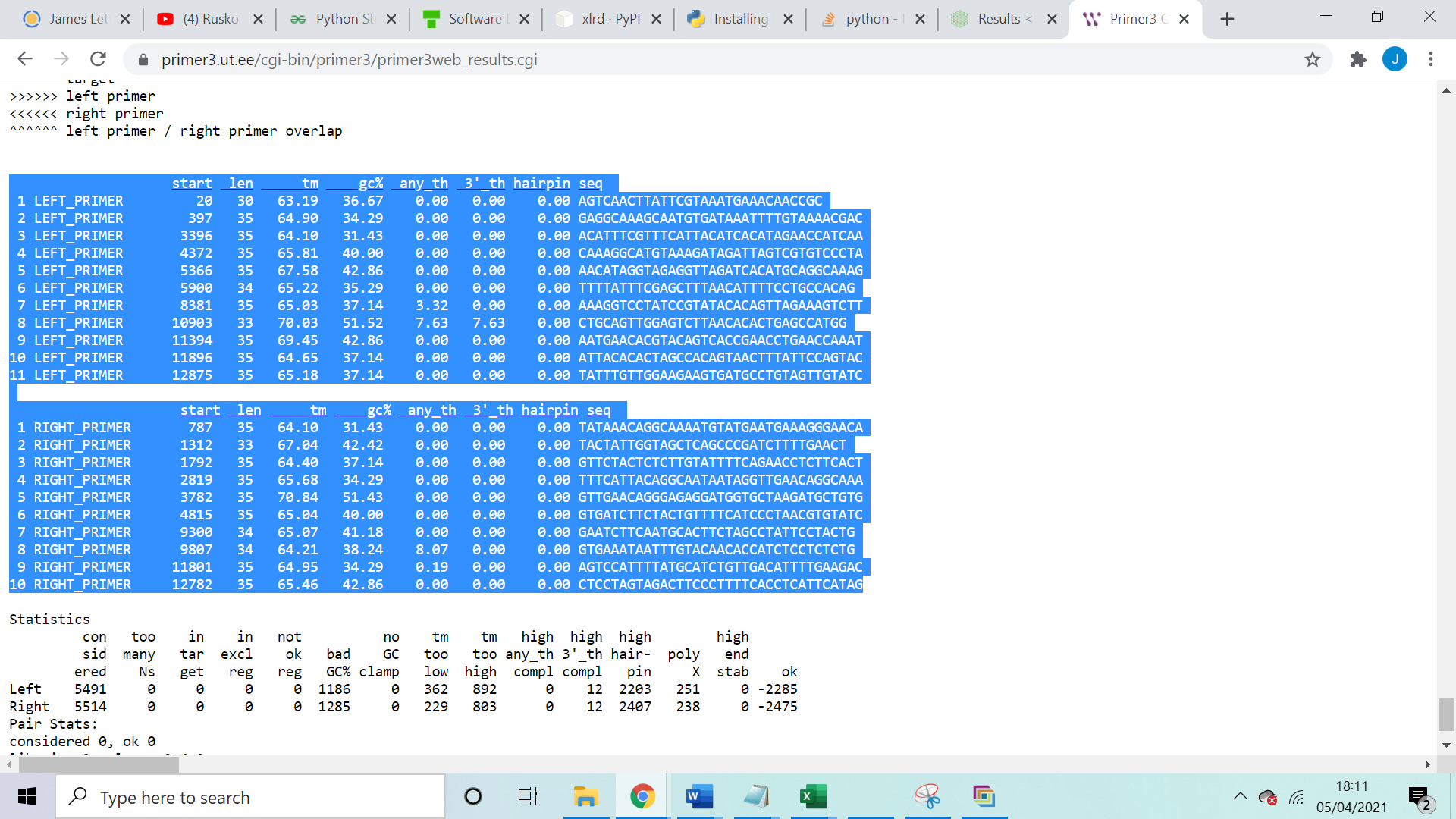
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Parameter | Parameter set 1 value | Parameter set 2 value | Parameter set 3 value | Parameter set 4 value | Parameter set 5 value |
| GC% Min | 40 | 40 | 40 | 40 | 40 |
| GC% Max | 60 | 60 | 60 | 60 | 60 |
| Number to return | 150 | 150 | 150 | 150 | 150 |
| Primer size Min | 30 | 30 | 30 | 30 | 30 |
| Primer size Opt | 35 | 35 | 30 | 35 | 35 |
| Primer size Max | 35 | 35 | 35 | 35 | 35 |
| Primer tm Min | 66 | 66 | 66 | 66 | 66 |
| Primer tm Opt | 66 | 76 | 66 | 66 | 66 |
| Primer tm Max | 76 | 76 | 76 | 76 | 76 |
| Max tm Difference | 7 | 7 | 7 | 7 | 7 |
| Hairpin | Default value | Default value | Default value | 40 | 100 |

- Number to return = 150 . The aim is to return as many primers as possible to increase the chance of there being enough universal primers to over the whole gene .

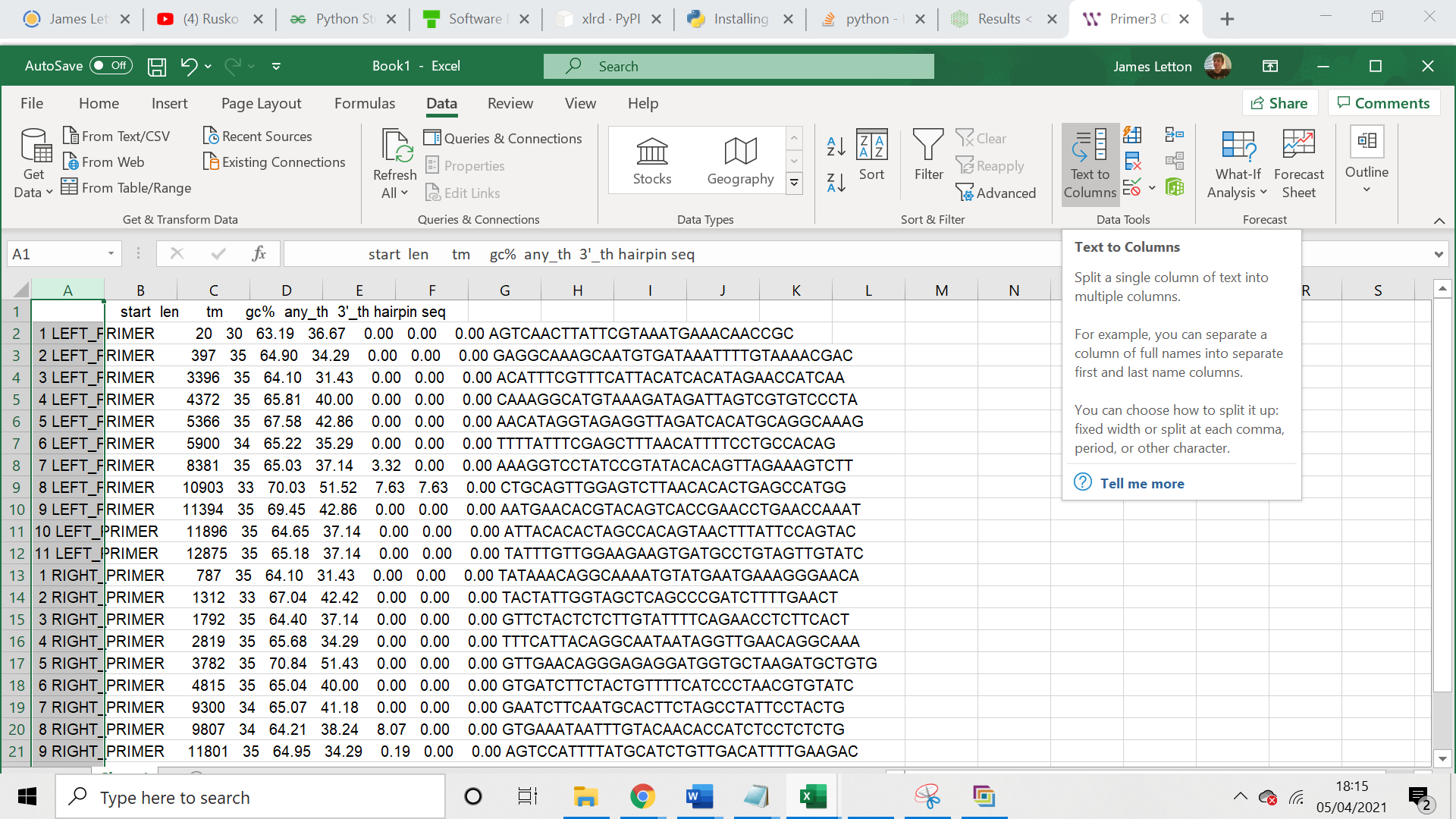
-Primer size set to min= 30 opt=35 max=35 .COL1A1 is highly repetitive and therefore highly likely to match other genes . Lengthing the primer helps overcome this.

-Primer tm min= 66 opt=65 max =76 . Increasing the primer size means that the temperature needs increasing or very few primers will be possible.

1. Copy the results to excel into the first cell/box .

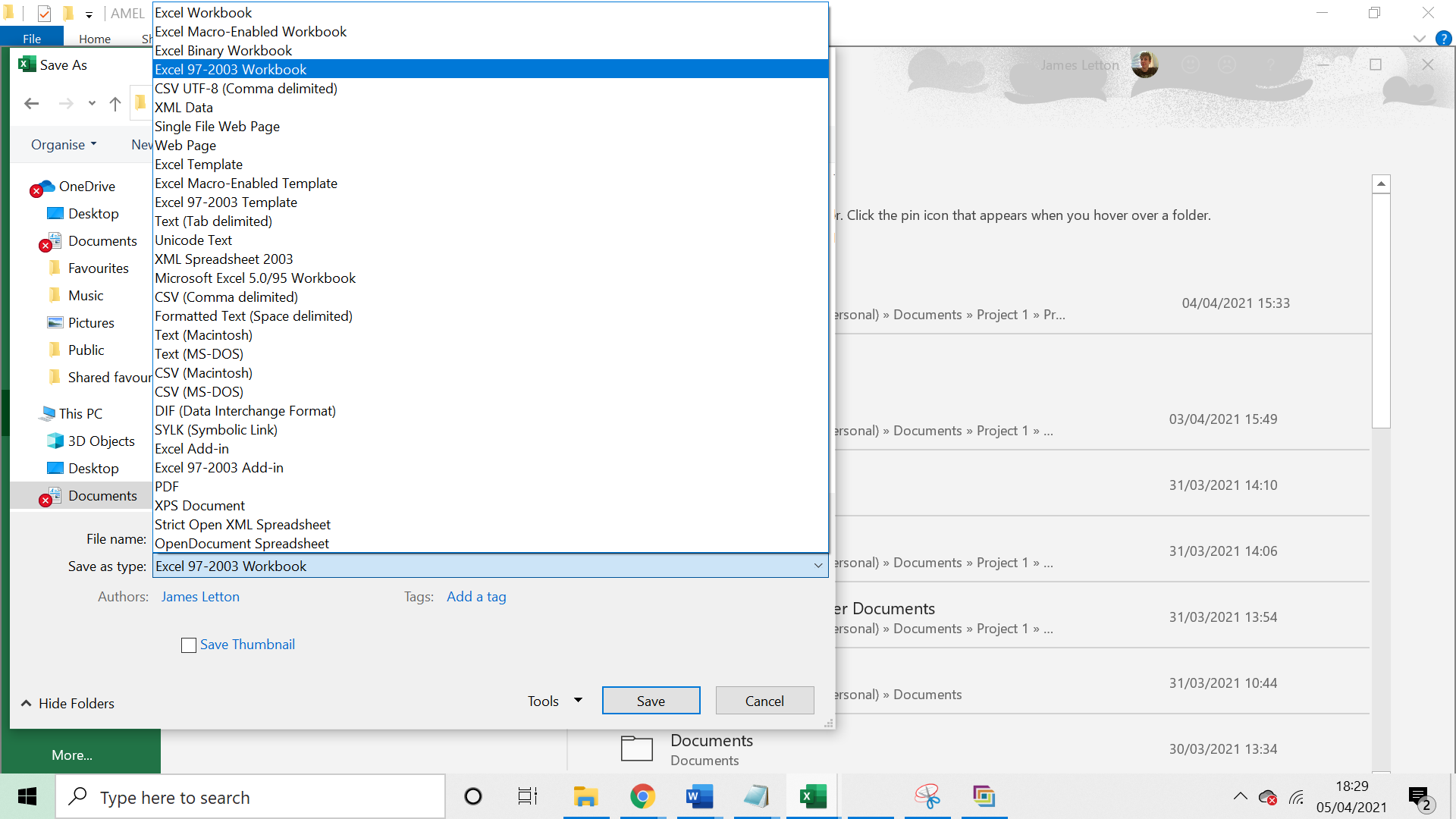


1. Delete the lines separating the two lists and remove the second set of headers.
2. Highlight all of column A and then open the Data tab and click text to column.



-When the options come up click next and then finish.If any headers have merged separate them manually

1. Delete the first column containing the numbers .
2. Name the header column containing “RIGHT\_PRIMER” or “LEFT\_PRIMER” “direction”.
3. Save the file as **primerinformation.xls** and set the save type as excel 97-2003 workbook.

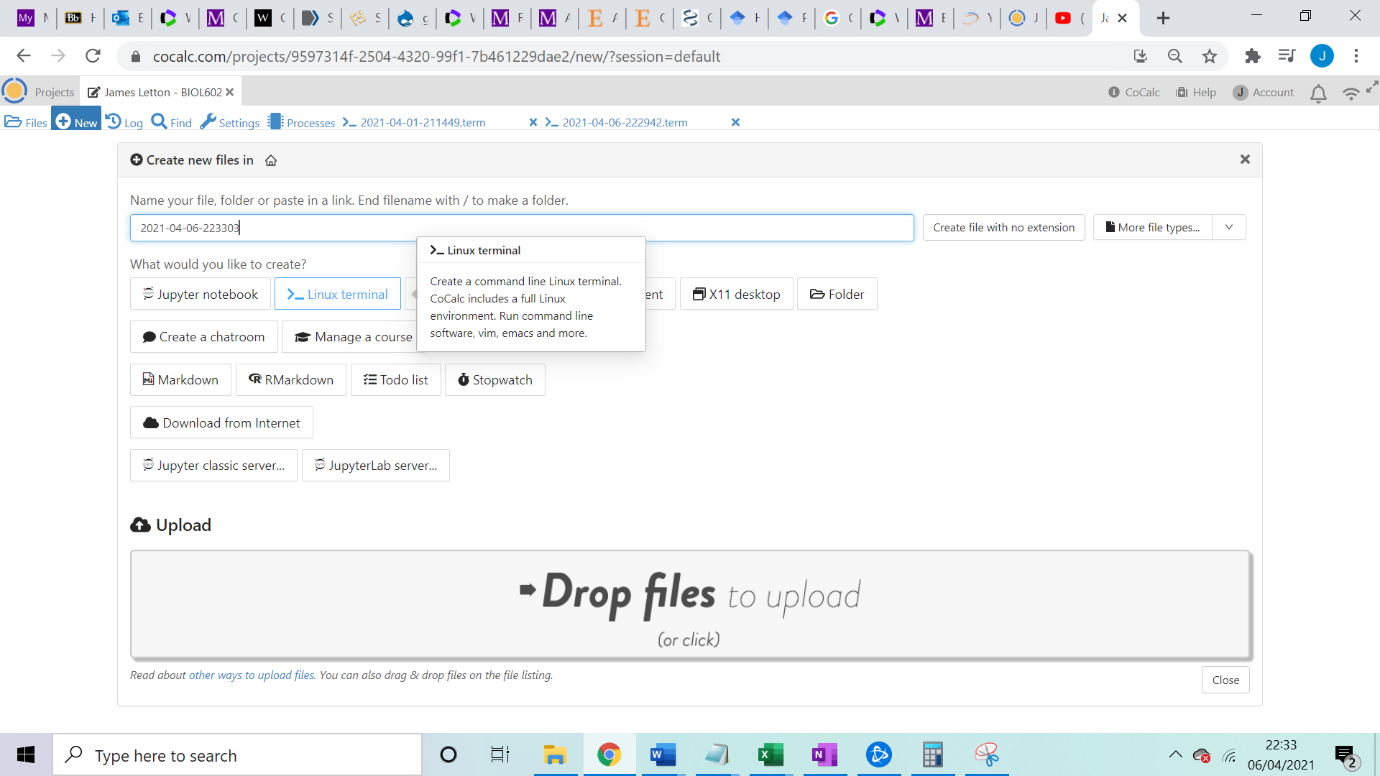


1. You will end up with some repeated primers from using the different parameters. Once the results are all together with just one set of headers you can quickly remove repeats by selecting all and clicking the remove duplicates button , in the data tools section of the data tab.

-If you would prefer to use another design program or add a couple more primers later that is fine if no cells are left blank.

**Primer list (rawprimers.txt)**

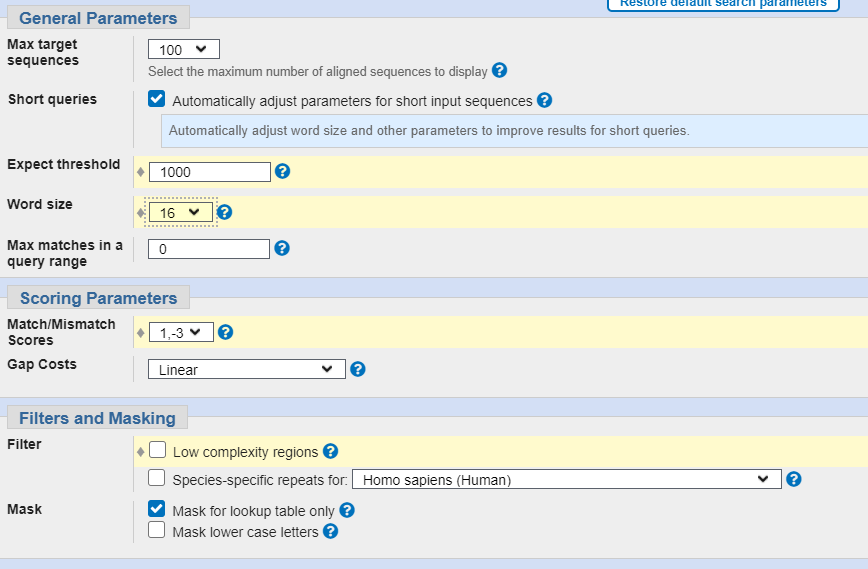
1. Open up a new linex terminal.



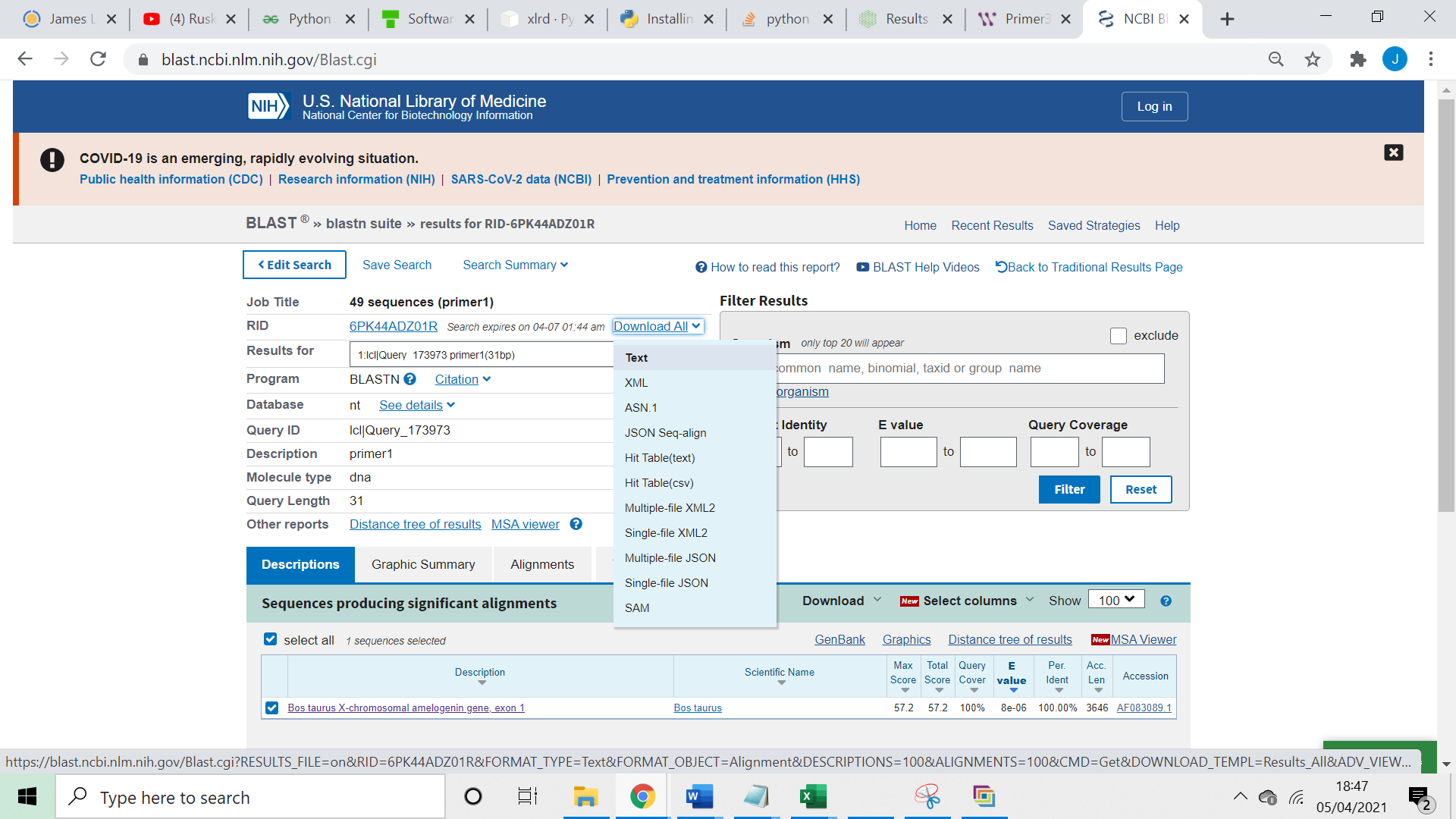
1. Enter “python3 PrimerFormater.py” into the terminal (run PrimerFormater). The file will be created automatically from primerinformation.xls .

**Blast results (blastresults.txt)**

1. Open blast from NCBI (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome>) and select to do a blastn search.
2. Copy the entire rawprimer file into the blast search box and under organism put your reference species .
3. Check at the bottom that blastn is selected in the program selection and then run the blast.Check that the search has been adjusted for short sequences this should occur automatically ,however, if it does not the settings are shown below.



1. Click download all and download as a text. Name the file **blastresults.txt**



1. Repeat but only search for human sequences

**Program settings (settings.txt)**

1. Settings is by default already in the program folder ,however, it will have to be altered depending on what gene and how many species are being looked at .
2. Open settings.txt and input your reference species into SpeciesReferance=????: . Make sure there is no space between the “=” and the species name also makes sure there is no space between the species name and “:”.
3. Input all the species being compared to the reference species . To do this input the names into the lines starting with Species .Enter them the same as in step 28.
4. At the end of the Species reference line it references whether the species should be taken into account for degenerate primer production. Yes means it will be used No means it won’t.

-Note that yes or no does not follow the reference species and should be left blank.

1. When the program is run it ignores the gene of interest in blast and only gives other gene hits . However , to filter against the gene the user has to enter words that would be found in the sequence name .Input the phrases you would like to use after “exclude=”.
2. Settings also has the setting “exemption=”. This is give the option to override the exclusion if there is another word in the hit name . For example, if you used “COL1A1” as an exclusion phrases it would ignore all hit names with COL1A1 . However if there was a reoccurring hit name called COL1A1 like receptor it would be incorrectly ignored. So you could use “receptor” as an exemption phrase which would override COL1A1 so it is not ignored
3. Once settings is filled out save the file (do not change the name).

**Run UniversalPrimerMaker**

1. Open up a new linex terminal(refer to step 20 for reminder).
2. Enter “python3 UniversalPrimerMaker.py”.
3. A file called similarity information should be produced

**Copying txt file to excel**

1. Use control A to copy all of the text file and paste it into cell A1.
2. Highlight the column and go to the data tab and click text to columns (similar to step 14).

-click next then click the deliminter “space” and then click finish