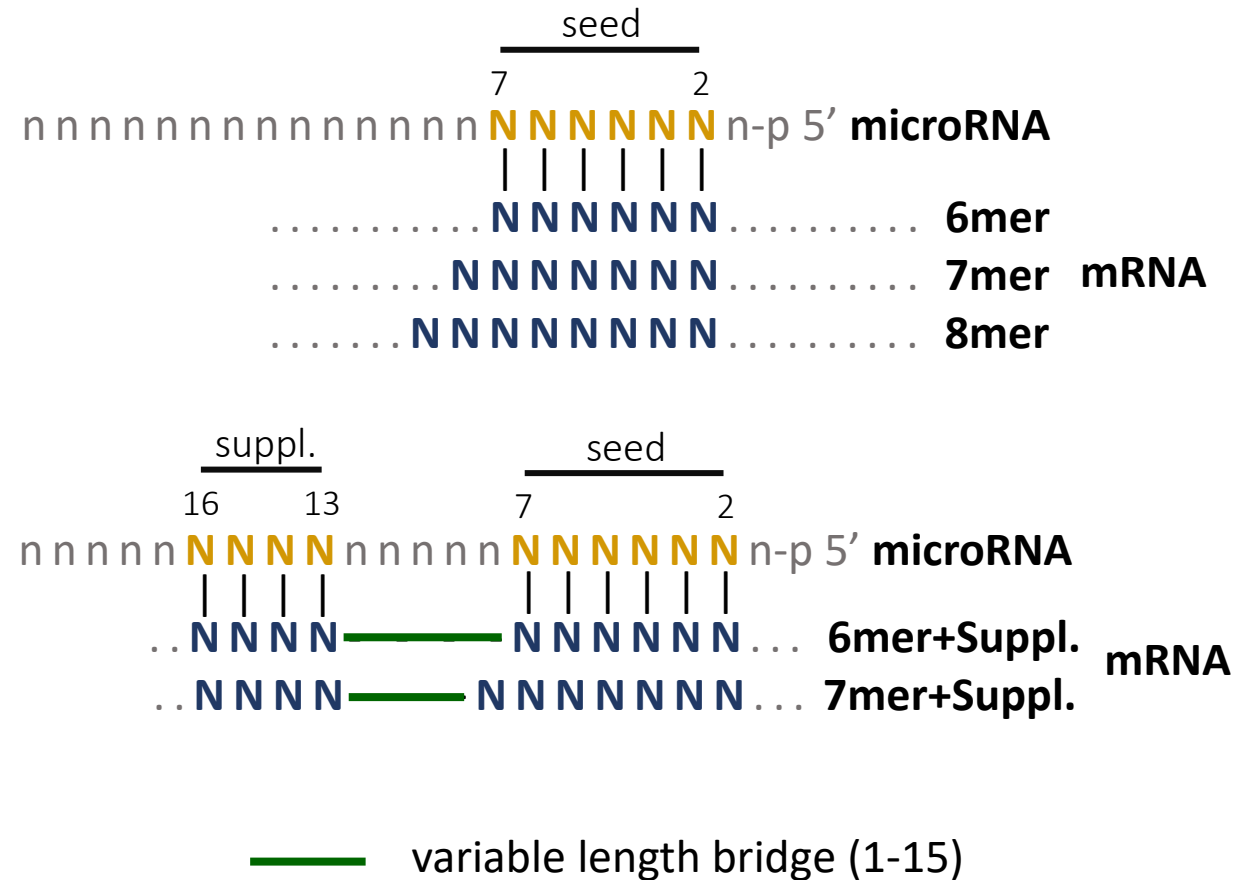


Readme









First download the necessary files for the .exe to run from:

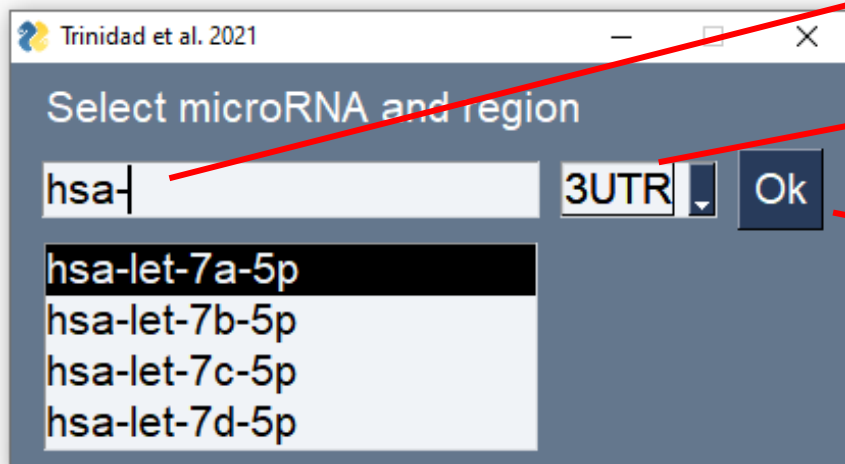
<https://drive.google.com/drive/folders/1aZka4ZUoSU4thmlYFPgqCy8cUrj0WXqD?usp=sharing>

Search by microRNA

If you are interested in looking for all target mRNAs that have sequence complementarity with seeds and seed plus supplementary-region of certain microRNAs, you should use:



 searchByMicroRNA	Aplicación	30.388 KB
 miR_Family_Info_targetScan_2016	Documento de te...	774 KB
 mart_export_CDS_2021_online	Documento de te...	145.433 KB
 mart_export_5UTRs_2021_online	Documento de te...	35.254 KB
 mart_export_3UTRs_2021_online	Documento de te...	116.218 KB



1

Write or select microRNA name of interest

2

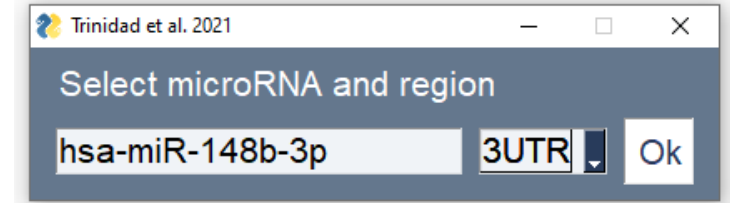
Select mRNA region in to look for

3

Press "OK" bottom to run the process

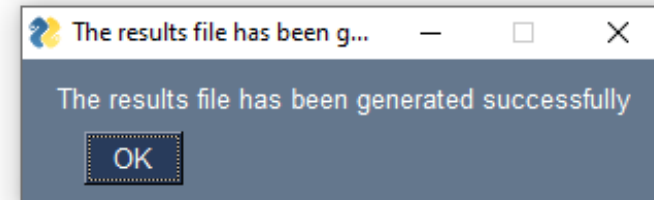
4

While process is running the “OK” bottom will be unavailable (white fill). This will take from few seconds to few minutes depending on the computer power.



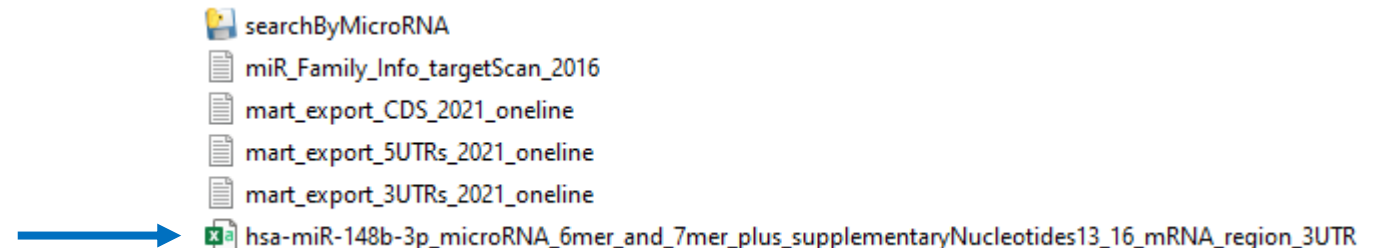
5

When the file is ready this window will pop up.

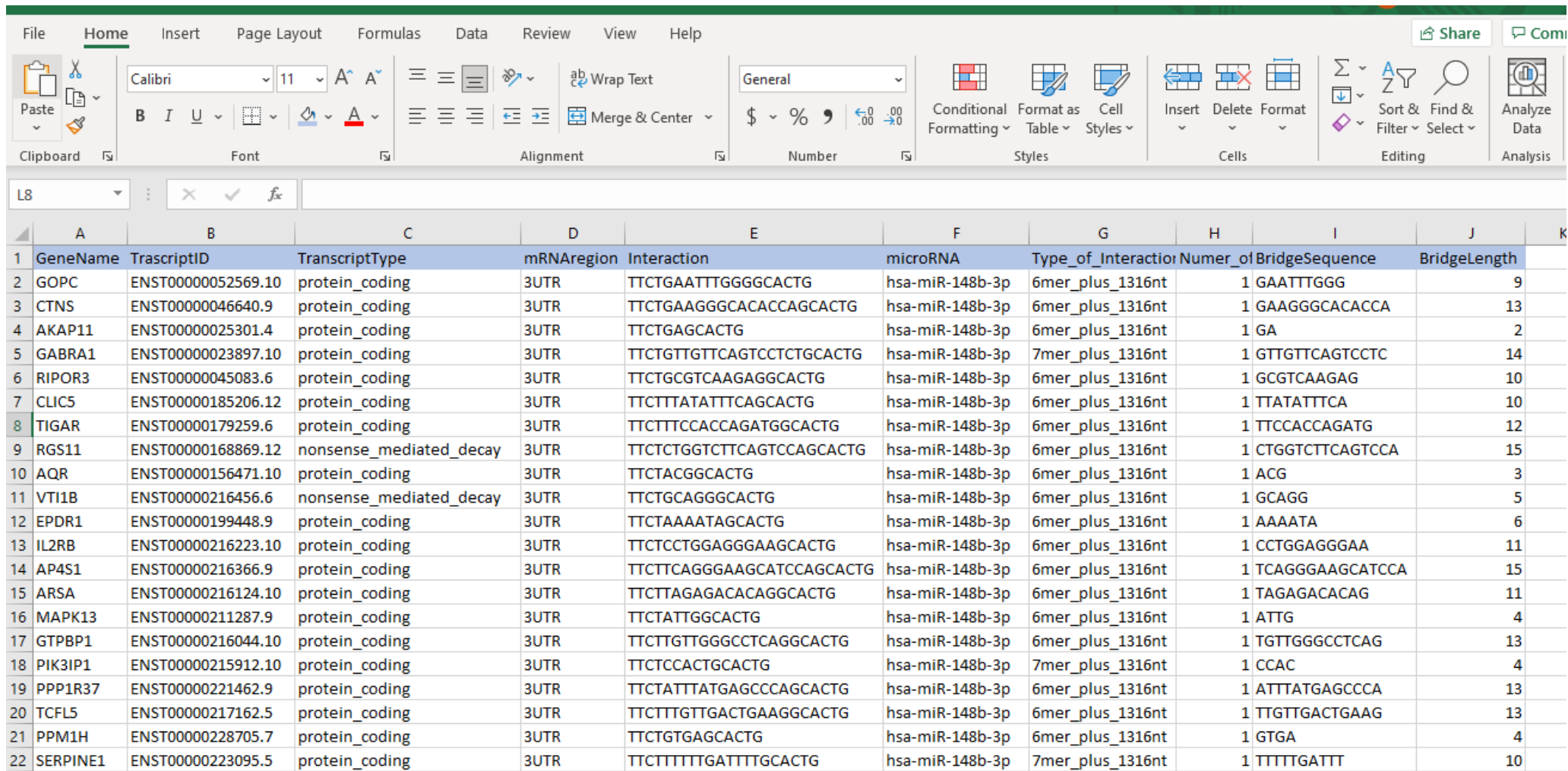


6

Check for the **.txt/.csv** (tab separated) output in the folder where you ran the **.exe**. The name must start with the microRNA name and finish with the region you search for.



Output file




	A	B	C	D	E	F	G	H	I	J	K
	GeneName	TrascriptID	TranscriptType	mRNAregion	Interaction	microRNA	Type_of_Interaction	Numer_of	BridgeSequence	BridgeLength	
1	GOPC	ENST00000052569.10	protein_coding	3UTR	TTCTGAATTTGGGGCACTG	hsa-miR-148b-3p	6mer_plus_1316nt	1	GAATTTGGG	9	
2	CTNS	ENST00000046640.9	protein_coding	3UTR	TTCTGAAGGGCACACCAGCACTG	hsa-miR-148b-3p	6mer_plus_1316nt	1	GAAGGGCACACCA	13	
3	AKAP11	ENST00000025301.4	protein_coding	3UTR	TTCTGAGCACTG	hsa-miR-148b-3p	6mer_plus_1316nt	1	GA	2	
4	GABRA1	ENST00000023897.10	protein_coding	3UTR	TTCTGTTGTTTCAGTCTCTGCACTG	hsa-miR-148b-3p	7mer_plus_1316nt	1	GTTGTTTCAGTCTCTC	14	
5	RIPOR3	ENST00000045083.6	protein_coding	3UTR	TTCTGCGTCAAGAGGCACTG	hsa-miR-148b-3p	6mer_plus_1316nt	1	GCGTCAAGAG	10	
6	CLIC5	ENST000000185206.12	protein_coding	3UTR	TTCTTTATATTTTCAGCACTG	hsa-miR-148b-3p	6mer_plus_1316nt	1	TTATATTTCA	10	
7	TIGAR	ENST000000179259.6	protein_coding	3UTR	TTCTTTCCACCAGATGGCACTG	hsa-miR-148b-3p	6mer_plus_1316nt	1	TTCCACCAGATG	12	
8	RGS11	ENST000000168869.12	nonsense_mediated_decay	3UTR	TTCTCTGGTCTTCAGTCCAGCACTG	hsa-miR-148b-3p	6mer_plus_1316nt	1	CTGGTCTTCAGTCCA	15	
9	AQR	ENST000000156471.10	protein_coding	3UTR	TTCTACGGCACTG	hsa-miR-148b-3p	6mer_plus_1316nt	1	ACG	3	
10	VTI1B	ENST000000216456.6	nonsense_mediated_decay	3UTR	TTCTGCAGGGCACTG	hsa-miR-148b-3p	6mer_plus_1316nt	1	GCAGG	5	
11	EPDR1	ENST000000199448.9	protein_coding	3UTR	TTCTAAAATAGCACTG	hsa-miR-148b-3p	6mer_plus_1316nt	1	AAAATA	6	
12	IL2RB	ENST000000216223.10	protein_coding	3UTR	TTCTCCTGGAGGGAAGCACTG	hsa-miR-148b-3p	6mer_plus_1316nt	1	CCTGGAGGGAA	11	
13	AP4S1	ENST000000216366.9	protein_coding	3UTR	TTCTTCCAGGAAGCATCCAGCACTG	hsa-miR-148b-3p	6mer_plus_1316nt	1	TCAGGGAAGCATCCA	15	
14	ARSA	ENST000000216124.10	protein_coding	3UTR	TTCTTAGAGACACAGGCACTG	hsa-miR-148b-3p	6mer_plus_1316nt	1	TAGAGACACAG	11	
15	MAPK13	ENST000000211287.9	protein_coding	3UTR	TTCTATTGGCACTG	hsa-miR-148b-3p	6mer_plus_1316nt	1	ATTG	4	
16	GTPBP1	ENST000000216044.10	protein_coding	3UTR	TTCTTGTTGGGCTCAGGCACTG	hsa-miR-148b-3p	6mer_plus_1316nt	1	TGTTGGGCTCAG	13	
17	PIK3IP1	ENST000000215912.10	protein_coding	3UTR	TTCTCCACTGCACTG	hsa-miR-148b-3p	7mer_plus_1316nt	1	CCAC	4	
18	PPP1R37	ENST000000221462.9	protein_coding	3UTR	TTCTATTATGAGCCCAGCACTG	hsa-miR-148b-3p	6mer_plus_1316nt	1	ATTATGAGCCCA	13	
19	TCFL5	ENST000000217162.5	protein_coding	3UTR	TTCTTTGTTGACTGAAGGCACTG	hsa-miR-148b-3p	6mer_plus_1316nt	1	TTGTTGACTGAAG	13	
20	PPM1H	ENST000000228705.7	protein_coding	3UTR	TTCTGTGAGCACTG	hsa-miR-148b-3p	6mer_plus_1316nt	1	GTGA	4	
21	SERPINE1	ENST000000223095.5	protein_coding	3UTR	TTCTTTTTGATTTTGCACTG	hsa-miR-148b-3p	7mer_plus_1316nt	1	TTTTTGATT	10	






Remember this is a **tab separated file**, so, you must use Excel Data > Text to Columns. Here is an example:

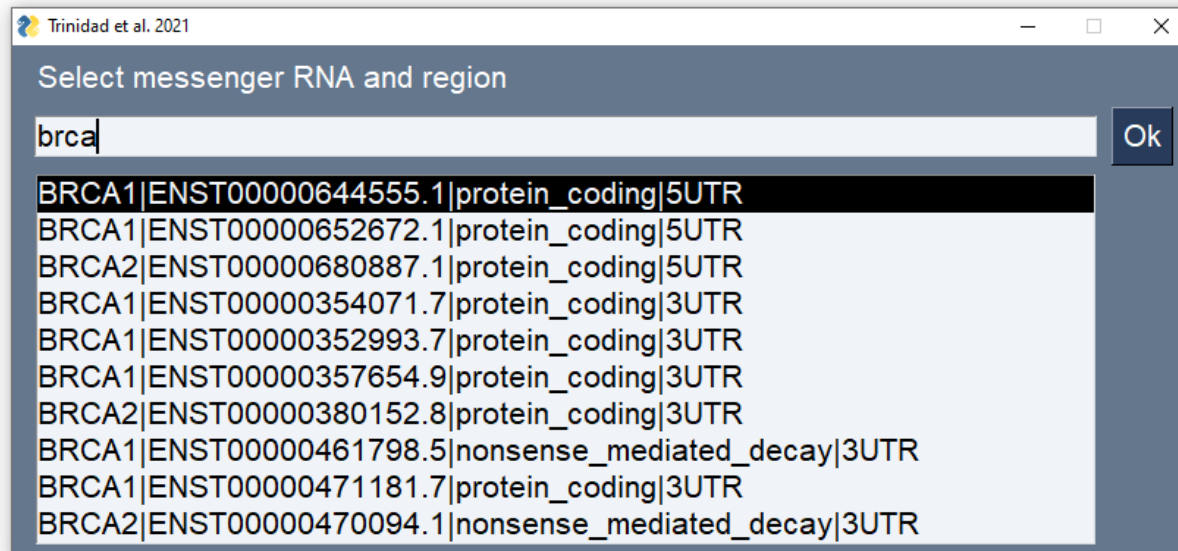

<https://support.microsoft.com/en-us/office/split-text-into-different-columns-with-the-convert-text-to-columns-wizard-30b14928-5550-41f5-97ca-7a3e9c363ed7>

Search by mRNA

If you are interested in looking for with seeds and seed plus supplementary-region interactions of **any microRNA** in **certain mRNAs**, you should use:



 searchByMessengerRNA	Aplicación	30.388 KB
 miR_Family_Info_targetScan_2016	Documento de te...	774 KB
 mart_export_CDS_2021_online	Documento de te...	145.433 KB
 mart_export_5UTRs_2021_online	Documento de te...	35.254 KB
 mart_export_3UTRs_2021_online	Documento de te...	116.218 KB



Select messenger RNA and region

brca

Ok

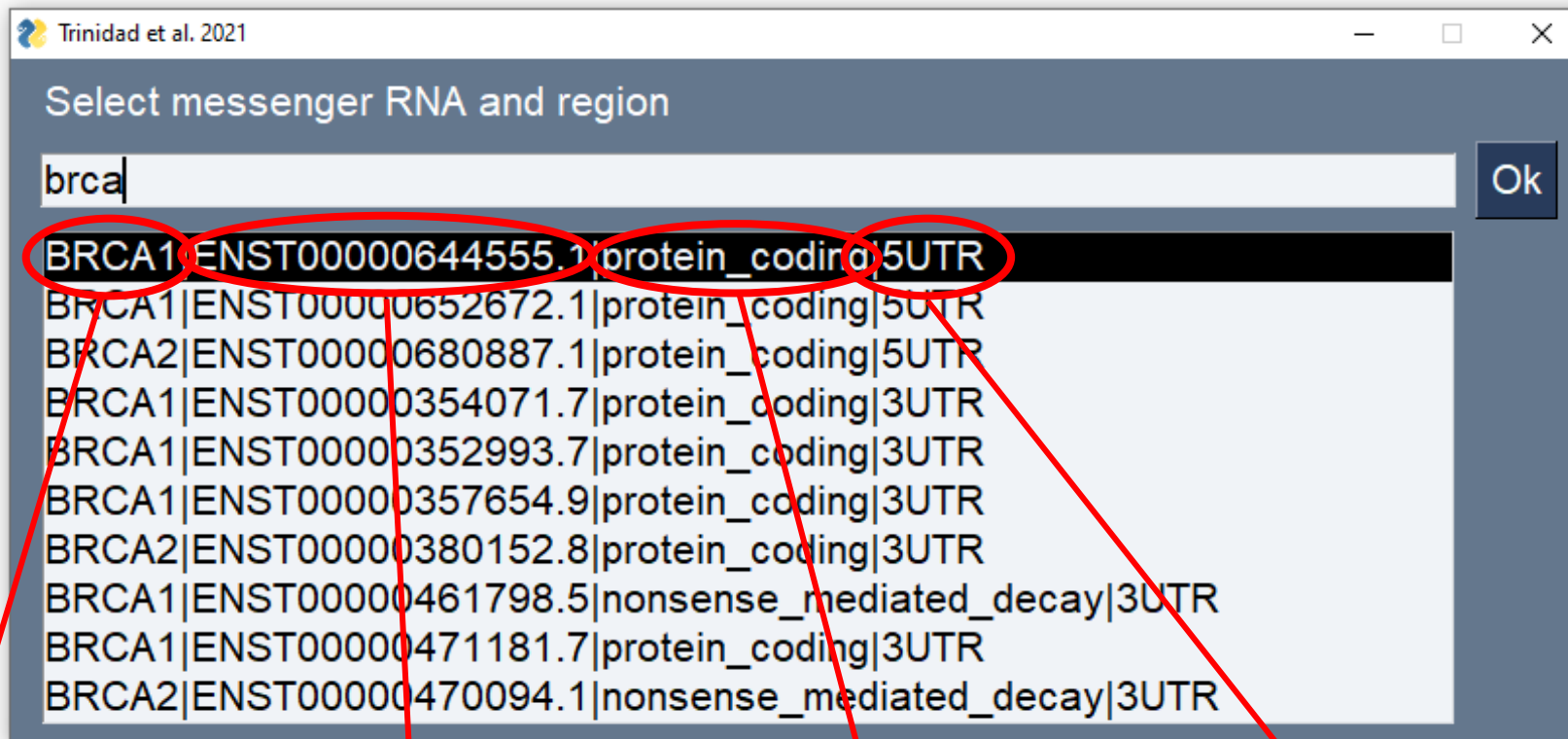
BRCA1|ENST00000644555.1|protein_coding|5UTR
BRCA1|ENST00000652672.1|protein_coding|5UTR
BRCA2|ENST00000680887.1|protein_coding|5UTR
BRCA1|ENST00000354071.7|protein_coding|3UTR
BRCA1|ENST00000352993.7|protein_coding|3UTR
BRCA1|ENST00000357654.9|protein_coding|3UTR
BRCA2|ENST00000380152.8|protein_coding|3UTR
BRCA1|ENST00000461798.5|nonsense_mediated_decay|3UTR
BRCA1|ENST00000471181.7|protein_coding|3UTR
BRCA2|ENST00000470094.1|nonsense_mediated_decay|3UTR



1

Write and select mRNA of interest.
Consider that selection include:

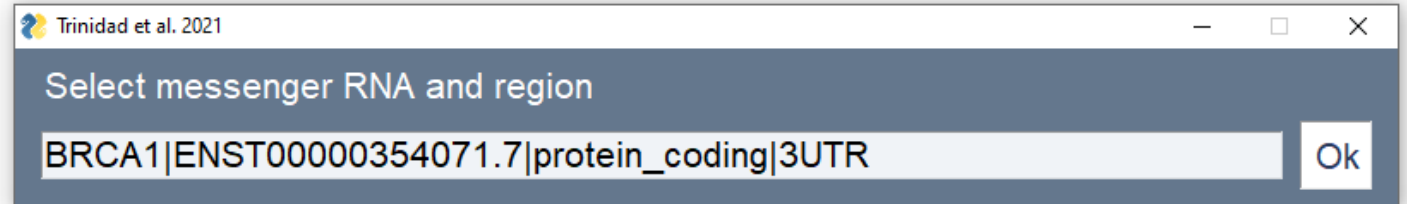
Gene Name | Transcript stable ID version | Gene type |
mRNA region to search for



Gene Name | Transcript stable ID version | Gene type | mRNA region to search for

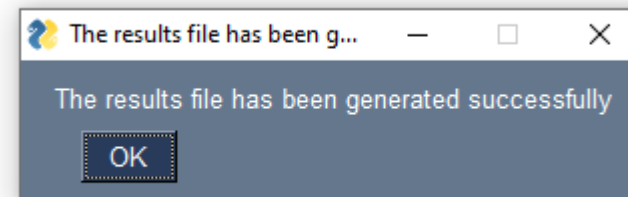
2

While process is running the “OK” bottom will be unavailable (white fill).
This will take from few seconds to few minutes depending on the computer power.



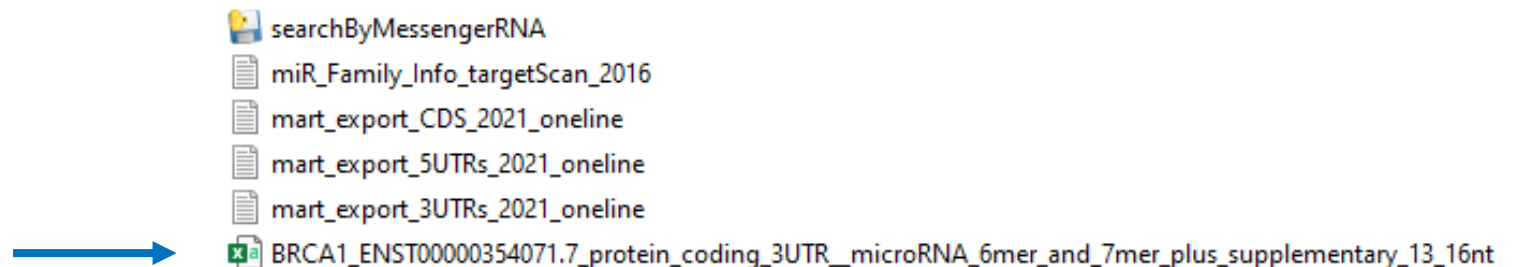
3

When the file is ready this window will pop up.

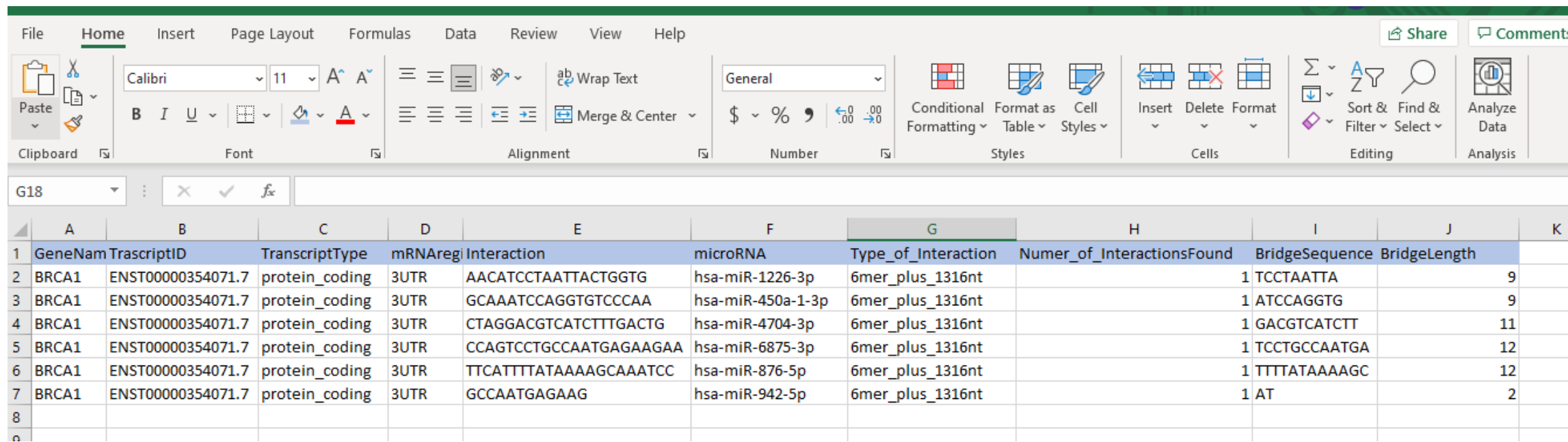


4

Check for the **.txt/.csv** (tab separated) output in the folder where you ran the **.exe**
The name must start with the mRNA name and include all information selected in the first window.



Output file



The screenshot shows the Microsoft Excel interface with the 'Home' tab selected. The ribbon includes options for Clipboard, Font, Alignment, Number, Styles, Cells, Editing, and Analysis. The formula bar shows 'G18'. The data table is as follows:

	A	B	C	D	E	F	G	H	I	J	K
1	GeneNam	TrascriptID	TranscriptType	mRNAregi	Interaction	microRNA	Type_of_Interaction	Numer_of_InteractionsFound	BridgeSequence	BridgeLength	
2	BRCA1	ENST00000354071.7	protein_coding	3UTR	AACATCCTAATTACTGGTG	hsa-miR-1226-3p	6mer_plus_1316nt	1	TCCTAATTA	9	
3	BRCA1	ENST00000354071.7	protein_coding	3UTR	GCAAATCCAGGTGTCCCAA	hsa-miR-450a-1-3p	6mer_plus_1316nt	1	ATCCAGGTG	9	
4	BRCA1	ENST00000354071.7	protein_coding	3UTR	CTAGGACGTCATCTTTGACTG	hsa-miR-4704-3p	6mer_plus_1316nt	1	GACGTCATCTT	11	
5	BRCA1	ENST00000354071.7	protein_coding	3UTR	CCAGTCCTGCCAATGAGAAGAA	hsa-miR-6875-3p	6mer_plus_1316nt	1	TCCTGCCAATGA	12	
6	BRCA1	ENST00000354071.7	protein_coding	3UTR	TTCATTTTATAAAAGCAAATCC	hsa-miR-876-5p	6mer_plus_1316nt	1	TTTATAAAAGC	12	
7	BRCA1	ENST00000354071.7	protein_coding	3UTR	GCCAATGAGAAG	hsa-miR-942-5p	6mer_plus_1316nt	1	AT	2	
8											

Remember this is a **tab separated file**, so, you must use Excel Data > Text to Columns. Here is an example:

<https://support.microsoft.com/en-us/office/split-text-into-different-columns-with-the-convert-text-to-columns-wizard-30b14928-5550-41f5-97ca-7a3e9c363ed7>











Troubleshooting

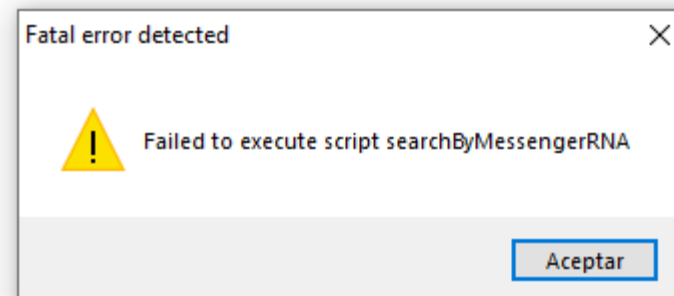
Contact jtrinidad@fcien.edu.uy

Most common problem:

Not running program. Check if **texts files (.txt/.csv)** are in the same folder than **.exe files**

To run the programs, you need these .txt files in the same folder

 searchByMessengerRNA	Aplicación	30.388 KB
 miR_Family_Info_targetScan_2016	Documento de te...	774 KB
 mart_export_CDS_2021_online	Documento de te...	145.433 KB
 mart_export_5UTRs_2021_online	Documento de te...	35.254 KB
 mart_export_3UTRs_2021_online	Documento de te...	116.218 KB
 searchByMicroRNA	Aplicación	30.388 KB
 miR_Family_Info_targetScan_2016	Documento de te...	774 KB
 mart_export_CDS_2021_online	Documento de te...	145.433 KB
 mart_export_5UTRs_2021_online	Documento de te...	35.254 KB
 mart_export_3UTRs_2021_online	Documento de te...	116.218 KB



Interested in running the
program for other species?

1

This algorithm uses the **.txt files** in the folder to create the microRNA interaction and to search in the mRNAs. If you wanted to use this algorithm for other species, you need to change the **.txt files** respecting their **format and names**.



mart_export_5UTRs_2021_online.txt
mart_export_3UTRs_2021_online.txt
mart_export_CDS_2021_online.txt
miR_Family_Info_targetScan_2016.txt

We use two types of files:

- The ones with the **mRNA sequences** which uses a fasta format with the sequence in one-line names mart_export*.txt . If you want to use sequence from other specie you must use in this fasta format and with the same name as the original file.



>GeneInfo
All sequence on one line
>GeneInfo
All sequence on one line

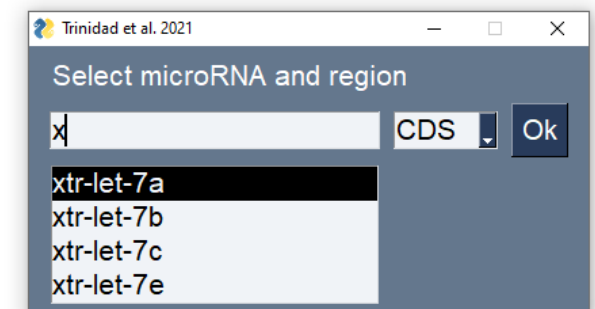
- For microRNA we use TargetScan file which is a table like file, you can check it with excel or text visualizers.

The easy way to add microRNAs of other species if they are in the TargetScan file:

Our script select from this files those microRNAs that have 'Species ID' columns equal to 9606 (Human). So, you can contact me to change this parameter and I will give you back the **.exe** or you can change in the file you 'Species ID' to 9606. The microRNA names will not change so you can't be wrong. Remember this is a tab separated file, if you use excel to change Species ID be aware of saving to this format.

Here is the example of changing Species ID on excel and who microRNA name will be remained.

You must provide the correct mart_export files (one-line fasta files) from the specie of interest otherwise you will be searching in Human mRNA.



Do not hesitate to contact us,
we will be happy to help you!

Contact jtrinidad@fcien.edu.uy