

IWBBIO 2022

Collecting SARS-CoV-2 encoded miRNAs via text mining

A. Schubö, A. Hadziahmetovic, M. Joppich, R. Zimmer



viral miRNAs involved in disease¹

→ often silence important regulatory factors
and immune response

SARS-CoV-1 shown to produce svRNAs²

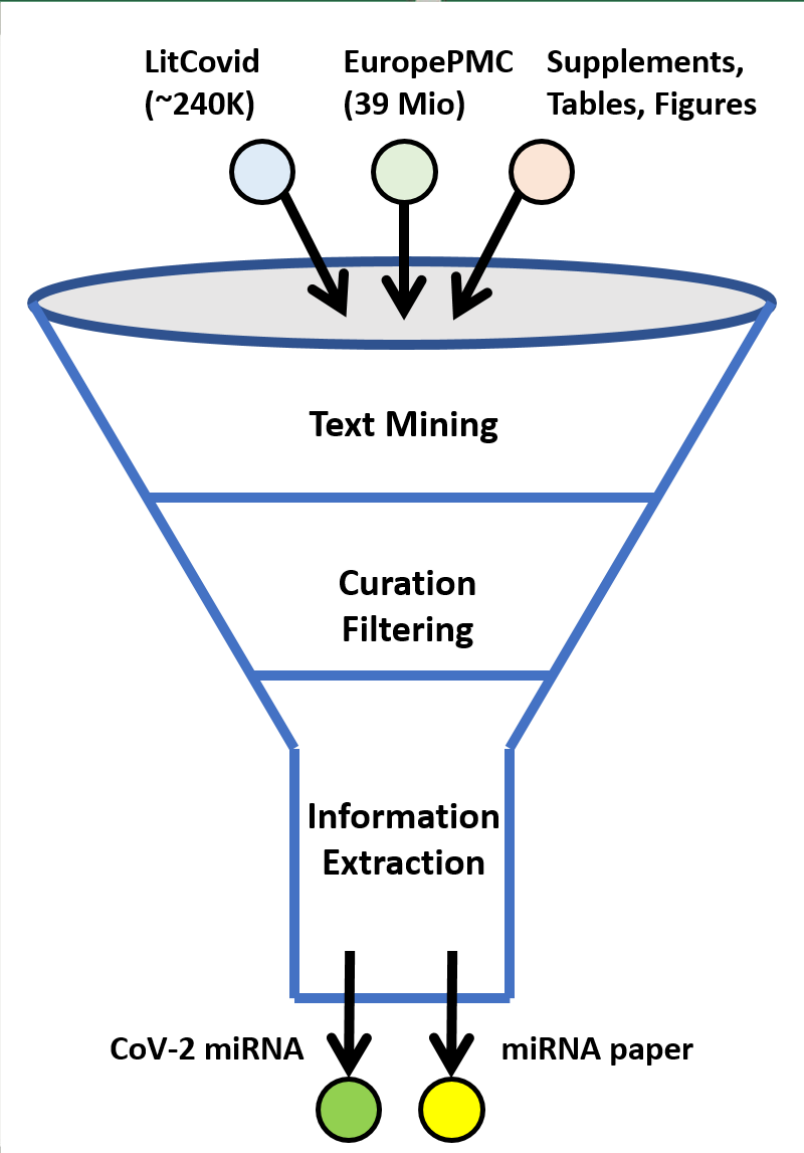
Our task: find all SARS-CoV-2 miRNAs

¹ M. Richa et al., The Interplay Between Viral-Derived miRNAs and Host Immunity During Infection, 2020, *Frontiers in Immunology*, 10.3389/fimmu.2019.03079

² L. Morales et al., SARS-CoV-Encoded Small RNAs Contribute to Infection-Associated Lung Pathology, 2017, *Cell Host & Microbe*, 10.1016/j.chom.2017.01.013

- scientific literature

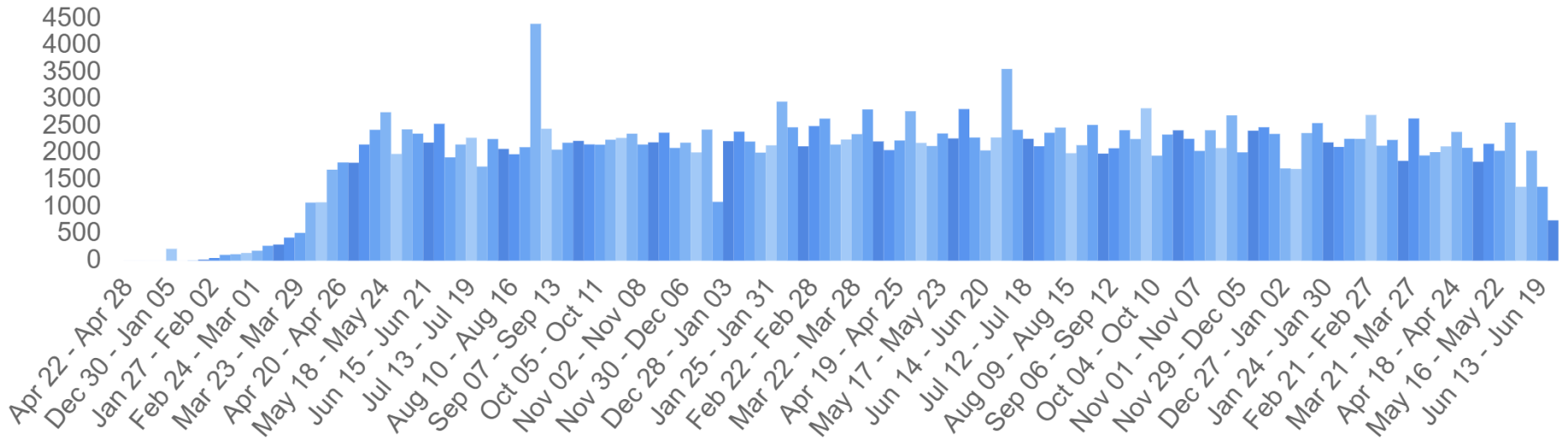
1. miRNA sequences
2. papers of origin



Why?

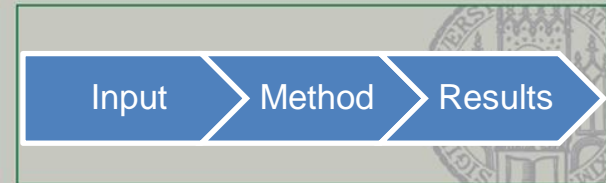
- Literature growing rapidly
- Manual search → huge cost

WEEKLY PUBLICATIONS



LitCovid:

- 240 000 texts*
- XML → sentences
- Named Entity Recognition on synonym list
- attribute of + miRNA



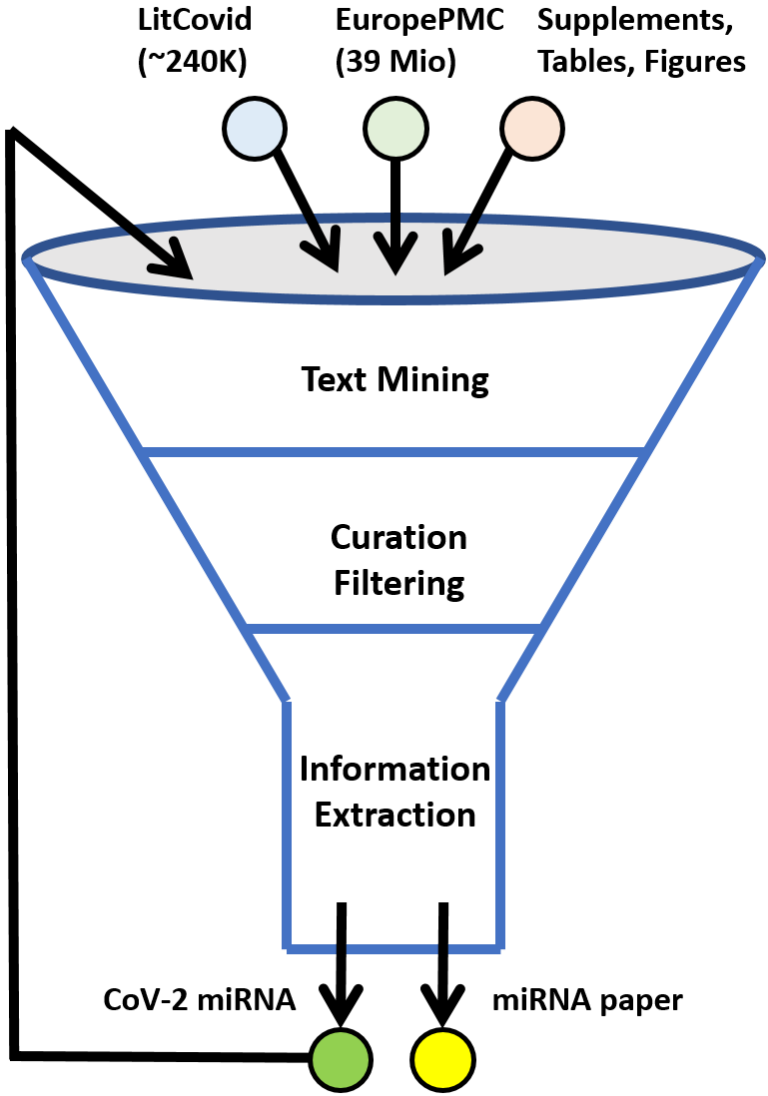
LitCovid:

- 240 000 texts*
 - XML → sentences
 - Named Entity Recognition on synonym list
 - attribute of + miRNA
- cooccurrence search
- Target set: paper of interest
- high sensitivity
 - low specificity

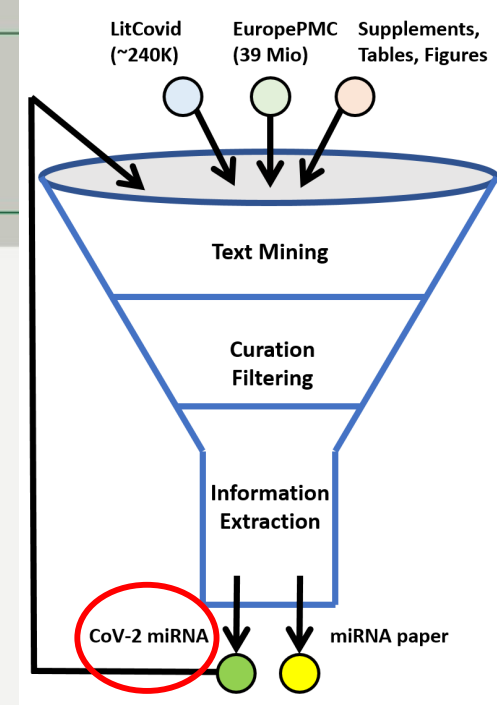
240 000 texts

350 texts of interest

Identified: 19 target publications
→ 560+ unique sequences



CoV-2 miRNA



M.N. Aydemir et al. Gene Reports 22 (2021) 101012

Table 2
MatureBayes analysis of SARS-CoV-2 miRNAs.

MatureBayes results				
Name	5' stem		3' stem	
	Position	Sequence	Position	Sequence
SARS-CoV-2-pre-miR-D7	43-64	AUACUAGUUUGUCUGGUUUUAA	75-96	UGUGUUAUGUAUGCAUCAGCUG

Saini et al., / Mol Biol Res Commun 2020;9(2):83-91 DOI: 10.22099/mbrc.2020.36507.1487 **MBRC**

Table 3: Mature miRNAs length, location and sequence as predicted by MatureBayes

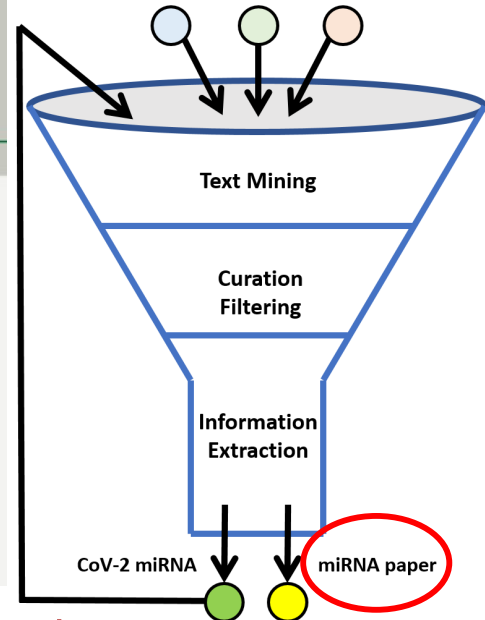
Mature miRNAs	Length (nt)	Location	Mature miRNAs sequence
SARS-CoV-2-MD135 5P	22	5'	AUACUAGUUUGUCUGGUUUUAA
SARS-CoV-2-MD135 3P	22	3'	UGUGUUAUGUAUGCAUCAGCUG

Viral miRNAs in SARS-CoV-2 infection 9

Supplementary Table 2 The details of candidate miRNAs

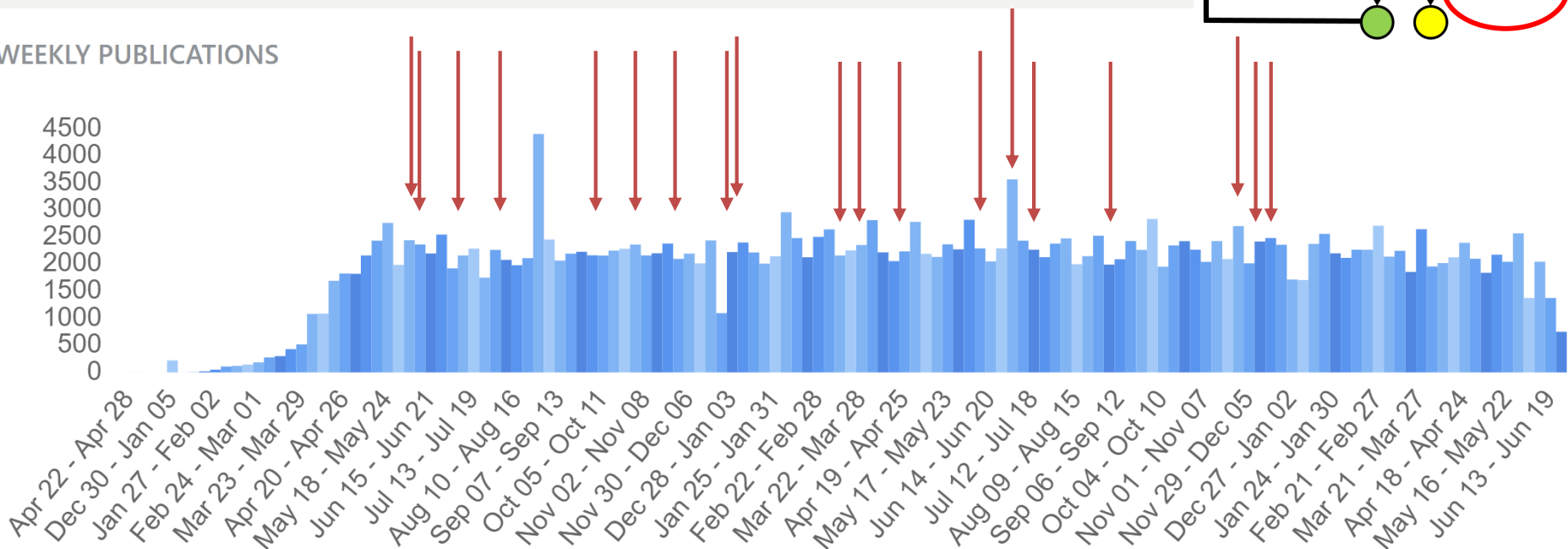
pre-miRNA ID	Start	End	MFE (kcal/mol)	Sequence of pre-miRNA	5p miRNA	3p miRNA
MD172	11234	11334	-36.6	GCUAGUUGGGUGAUGCGUAUUUGACA UGGUUGGAUAUGGUUGAUACUAGUUUG UCUGGUUUUAAGCUAAAAGACUGUGUU AUGUAUGCAUCAGCUGUAGU	AUACUAGU UUGUCUGG UUUUAA	UGUGUUAU GUAUGCAU CAGCUG

Inherent naming ambiguity



miRNA papers

WEEKLY PUBLICATIONS



Comprehensive collection of:

- SARS-CoV-2 encoded miRNA papers
19 papers
- list of all reported viral miRNAs
560+ validated and/or predicted viral encoded candidates

Processing of new literature

- since submission: ~20k new papers in LitCovid

Improve specificity of Text Mining

- add new constrains to reduce false positives

Validate SARS-CoV-2 miRNA candidates



Thank you for your attention