DOI: 10.1038/s41598-024-75740-3

scientific reports

Check for updates

OPEN

Data normalization of plasma miRNA profiling from patients with COVID-19

Julia Tiemi Siguemoto[1,5], Carolini Motta Neri[1,5], Nadine de Godoy Torso[2], Aline de Souza Nicoletti[2], Marília Berlofa Visacri[3], Carla Regina da Silva Correa da Ronda[1], Mauricio Wesley Perroud Jr.[2], Leonardo Oliveira Reis[2], Luiz Augusto dos Santos[4], Nelson Durán[2], Wagner José Fávaro[2], Eder de Carvalho Pincinato[2] & Patricia Moriel[1]✉

When using the reverse-transcription quantitative polymerase chain reaction (RT-qPCR) technique for quantitative assessment of microRNA (miRNA) expression, normalizing data using a stable endogenous gene is essential; however, no universally adequate reference gene exists. Therefore, in this study, we aimed to determine, via the RNA-Seq technique, the most adequate endogenous normalizer for the expression assessment of plasma miRNAs in patients with coronavirus disease 2019 (COVID-19). Two massive sequencing procedures were performed (a) to identify differentially expressed miRNAs between patients with COVID-19 and healthy volunteers ($n = 12$), and (b) to identify differentially expressed miRNAs between patients with severe COVID-19 and those with mild COVID-19 ($n = 8$). The endogenous normalizer candidates were selected according to the following criteria: (1) the miRNA must have a fold regulation = 1; (2) the miRNA must have a p-value > 0.990; and (3) the miRNAs that were discovered the longest ago should be selected. Four miRNAs (hsa-miR-34a-3p, hsa-miR-194-3p, hsa-miR-17-3p, and hsa-miR-205-3p) met all criteria and were selected for validation by RT-qPCR in a cohort of 125 patients. Of these, only hsa-miR-205-3p was eligible endogenous normalizers in the context of COVID-19 because their expression was stable between the compared groups.

Keywords MicroRNA, COVID-19, Endogenous normalizer, Reference gene

Coronavirus disease 2019 (COVID-19) is an acute respiratory infection caused by the SARS-CoV-2 coronavirus, a betacoronavirus identified in December 2019 in Wuhan, Hubei Province, China. It has been described as the seventh coronavirus to infect humans[1]. As of March 2024, more than 770 million confirmed COVID-19 cases have been reported worldwide[2], making it one of the most contagious pandemics in human history[3].

Given the global prevalence of this disease, understanding and characterizing the pathways associated with its pathogenesis and severity is crucial. MicroRNAs (miRNAs) are among the molecules considered for this purpose. miRNAs are small, single-stranded RNA molecules consisting of approximately 22 non-coding nucleotides that regulate gene expression post-transcriptionally[4].

Reverse-transcription quantitative polymerase chain reaction (RT-qPCR) is considered as the gold standard for evaluating miRNA expression. For this purpose, data normalization is required; more precisely, the expression of each target miRNA should be normalized to the expression of a stable endogenous miRNA (known as an endogenous normalizer)[5]. This procedure aims to eliminate variations that do not result from the biological conditions[5].

Although data normalization directly influences the accuracy of results, no standardized protocol exists for this procedure. In addition, no specific normalized miRNAs have been identified in patients with COVID-19.

[1]Faculty of Pharmaceutical Sciences, Universidade Estadual de Campinas, Campinas 13083970, Brazil. [2]School of Medical Sciences, Universidade Estadual de Campinas, Campinas 13083894, Brazil. [3]Faculty of Pharmaceutical Sciences, University of São Paulo, São Paulo 05508-000, Brazil. [4]Hospital Municipal de Paulínia, Paulínia, SP, Brazil. [5]Julia Tiemi Siguemoto and Carolini Motta Neri contributed equally to this work. ✉email: patricia.moriel@fcf.unicamp.br

A commonly employed strategy for data normalization involves the use of cel-miR-39, a synthetic miRNA of exogenous origin often added to samples[5]. Therefore, it cannot be considered the best candidate for use as a normalizer; instead, it should be used as a quality control for miRNA extraction, as the efficiency ratio can be determined by comparing the concentration of cel-miR-39 added to the sample to the quantity obtained after extraction[5].

Hence, considering that there are no concrete conclusions regarding which miRNA should be used as a normalizer for miRNA expression assessment in COVID-19 and considering the severity and importance of the disease, further research is required to elucidate this issue. Therefore, in this study, we aimed to determine the most appropriate endogenous normalizer for assessing plasma miRNA expression in patients with COVID-19.

Results

Participants

The criterion adopted for dividing the participants included in the study was the presence or absence of SARS-CoV-2 virus infection, as confirmed by a reference RT-qPCR test conducted using plasma collected on the same day as the test. Two large-scale sequencing procedures were performed.

In the first group, 12 participants were involved and subdivided into two groups: (1) Case group ($n=8$, participants with a positive RT-PCR test for SARS-CoV-2); (2) Control group ($n=4$, with a negative RT-PCR test for SARS-CoV-2).

The second sequencing involved eight participants, all of whom tested positive for SARS-CoV-2 using RT-PCR. The patients were divided into two groups based on COVID-19 severity (classified according to NIH Coronavirus Disease 2019 Treatment Guidelines[6]): (1) Non-severe group ($n=4$ participants with mild COVID-19); (2) Severe/critical group ($n=4$ participants with severe/critical COVID-19) (Fig. 1).

The participants' clinical characteristics are shown in Table 1.

miRNA sequencing

For the sequencing that evaluated miRNA expression between patients with COVID-19 and controls, the GeneGlobe Data Analysis Center (Qiagen) executed the analysis using the following method: fold regulation = miRNA expression in the case group (patients with COVID-19)/miRNA expression in the control group (healthy volunteers) (Table 2). For sequencing that evaluated miRNA expression between patients with severe/critical COVID-19 and those with mild COVID-19, fold regulation = miRNA expression of the severe or critical group/miRNA expression in the mild or moderate group (Table 3). For both tests, p-value were determined using the Wald test. Tables 2 and 3 show the sequencing results for plasma miRNAs with a fold regulation of 1.00.

In accordance with the selection criteria of this study and the sequencing outcomes, four miRNAs were selected as exogenous control candidates. The selected miRNAs, hsa-miR-34a-3p, hsa-miR-194-3p, hsa-miR-17-3p, and hsa-miR-205-3p, are listed in Tables 2 and 3.

Figures 2 and 3 present volcano plots with all plasma miRNAs identified in the sequencing comparing miRNA expression between patients with COVID-19 and controls, and the sequencing comparing miRNA expression between patients with severe/critical COVID-19 and patients with mild COVID-19, respectively.

Validation of candidates for normalizing miRNAs

Four miRNAs (hsa-miRNA-34a-3p, hsa-miR-194-3p, hsa-miR-17-3p, and hsa-miR-205-3p) selected by next-generation sequencing were confirmed by RT-qPCR in a larger cohort of participants ($n=125$, including all those involved in the sequencing).

Although hsa-miR-194-3p was detected in all samples during sequencing, it was not amplified by RT-qPCR. Different experiments were conducted to determine whether amplification could be confirmed, but all experiments were unsuccessful. Therefore, the use of this miRNA as an endogenous control was ruled out.

For the remaining three miRNAs (hsa-miR-34a-3p, hsa-miR-17-3p, and hsa-miR-105-3p), Tables 4 and 5 show their expression levels according to RT-qPCR. When comparing the expression levels between the case ($n=85$) and control groups ($n=40$), hsa-miR-17-3p was not suitable as an endogenous normalizer, as it exhibited significant differences in expression ($p<0.05$). Similar expression levels of hsa-miR-34a-3p and hsa-miR-205-3p were observed in the two groups ($p=0.3990$ and $p=0.0805$, respectively).

Furthermore, when comparing the expression levels between the group of patients with mild COVID-19 ($n=45$) and critical COVID-19 ($n=39$), a significant difference was observed ($p<0.05$) for hsa-miR-34a-3p, hsa-17-3p, and hsa-205-3p. Therefore, none of the selected miRNAs met the prerequisites for use as endogenous normalizers for different disease severities.

Stability assessment of selected endogenous miRNAs

The endogenous reference gene must be stably expressed in both groups (Case and Control when discussing the first sequencing; Mild COVID-19 and Critical COVID-19 when discussing the second sequencing). Thus, a statistical analysis was performed to evaluate this premise; as all candidate miRNAs showed significant differences when comparing expression between patients with severe/critical COVID-19 and those with mild COVID-19 ($p<0.05$), the analysis was made only for the results comparing miRNA expression between COVID-19 patients and controls.

Figure 4 illustrates the sequencing analysis conducted to compare the differentially expressed miRNAs between patients with COVID-19 and controls (healthy volunteers). RefFinder stability analysis[7,8] indicated that hsa-miR-17-3p was the most stable candidate, followed by hsa-miR-205-3p and hsa-miR-34a-3p. However, as hsa-miR-17-3p exhibited significant differences in expression between the control and case groups, the best endogenous normalizer for miRNA expression assessment was hsa-miR-34a-3p, followed by hsa-miR-205-3p.

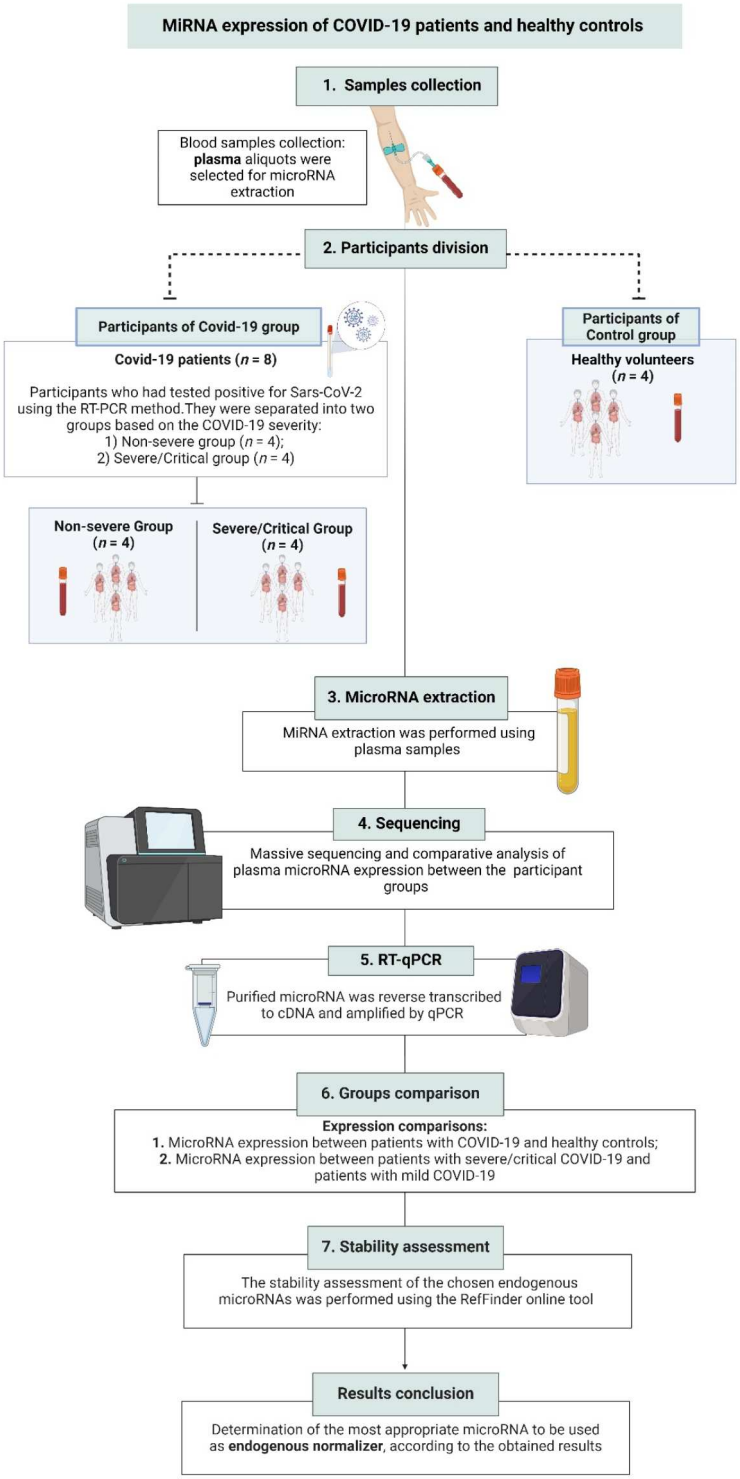

Fig. 1. On the left, the sequencing was performed to identify differentially expressed miRNAs between patients with COVID-19 and controls (healthy volunteers). Participants were divided as follows: (1) Case group ($n=8$, participants who tested positive for SARS-CoV-2 in the reference test RT-PCR); (2) Control group ($n=4$, participants with a negative RT-PCR result for SARS-CoV-2). On the right, the sequencing was performed to identify differentially expressed miRNAs between patients with severe/critical COVID-19 and those with mild COVID-19. Participants were divided as follows: (1) non-severe group ($n=4$, participants with mild/moderate COVID-19); (2) severe/critical group ($n=4$, participants with severe/critical COVID-19). (Created with BioRender.com, accessed on March 17, 2024.)

Variable	Patients with COVID-19		Patients without COVID-19 ($n=4$)
	Non-severe group ($n=4$)	Severe/critical group ($n=4$)	
Age (mean ± SD, years)	61.8 ± 11.7	61.5 ± 8.5	51.7 ± 16.9
Gender (n, %)			
Male	2 (50.0)	2 (50.0)	2 (50.0)
Female	2 (50.0)	2 (50.0)	2 (50.0)
Ethnicity (n, %)			
Caucasian	0 (0.0)	2 (50.0)	4 (100.0)
Non-Caucasian	0 (0.0)	2 (50.0)	0 (0.0)
Not reported	4 (100.0)	0 (0.0)	0 (0.0)
Comorbidities*			
Yes	2 (50.0)	3 (75.0)	2 (50.0)
No	2 (50.0)	1 (25.0)	2 (50.0)

Table 1. Clinical data of participants whose plasma miRNA expression were evaluated in the sequencing ($n=12$). *n*, absolute number of patients; SD, standard deviation. *Systemic arterial hypertension, diabetes mellitus, congestive heart failure, chronic kidney disease, hypothyroidism, obesity, cardiac arrhythmia, acute myocardial infarction and/or previous stroke.

Discussion

To ensure the success of miRNA expression assessment using the gold standard method RT-qPCR, the most appropriate normalizer is expected to be utilized. For this purpose, a gene must remain stable and exhibit constant expression across different contexts and individuals. However, as there is no ideal miRNA, testing the stability of potential normalizers prior to experimentation is necessary. Thus, we aimed to determine the best plasma miRNA normalizer for patients with COVID-19.

Although miR-16 and miR-191 are some of the most frequently employed miRNAs as normalizers[9], these miRNAs might not consistently exhibit the same pattern of expression in different biological or experimental contexts[10–12]. In a recent study by Belmonte et al.[13], who also investigated endogenous normalizers for COVID-19 severity, miR-16-5p expression showed a significant difference between the COVID and control groups. The authors also mentioned that certain binding sites in the SARS-CoV-2 genome are specific for miR-16-5p, suggesting that miR-16-5p may interfere with viral replication[13]. The high number of these binding sites in the viral genome could potentially reduce the presence of the coronavirus in infected cells, possibly aiding in infection control[13].

Therefore, to ensure the selection of the most appropriate normalizer, and consequently, the quality of the results[14] in the present study, miRNA sequencing was used to identify the best candidates. This high-throughput assay provided a comprehensive view of the presence and expression of miRNAs in a sample[15]. Combining the RNA sequencing results and the criteria established previously in this study (namely, a fold change = 1, a p-value > 0.990 for the miRNA, and the longest previously discovered miRNAs) yielded four candidates for validation as possible endogenous normalizers: hsa-miRNA-34a-3p, hsa-miR-194-3p, hsa-miR-17-3p, and hsa-miR-205-3p. Subsequently, RT-qPCR was conducted with a larger cohort of participants, analyzing the selected miRNAs as potential endogenous normalizers. Furthermore, care was taken to ensure that differences in individual profiles did not interfere with the study of miRNA expression.

During the RT-qPCR step, hsa-miR-194-3p did not show amplification in any sample, despite being detected by RNA sequencing. This could be attributed to various reasons such as inadequate reaction conditions for the specific miRNA or the presence of polymerase inhibitors that affect primer-template annealing[16]. Additionally, this result may be influenced by the fact that sequencing and RT-qPCR techniques are not always directly comparable, along with the potential limitations of RT-qPCR in validating high-throughput sequencing data[5]. Therefore, hsa-miR-194-3p was not an appropriate endogenous normalizer for use in the experiments, as neither amplification nor detection was observed using the gold standard method.

Next, the RT-qPCR results for the remaining three miRNAs were analyzed. To obtain an appropriate normalizer, gene expression would remain the same between the two groups being compared (in this case, case versus control group and non-critical COVID-19 versus critical COVID-19 group). At this stage, hsa-miR-17-3p could be discarded because a significant difference in its expression was observed both in the comparison between the case and control groups and between the mild and severe COVID-19 groups. Furthermore, the miRNAs hsa-miR-205-3p and hsa-miR-34a-3p showed significant differences in expression between the groups with mild and severe COVID-19 and was excluded in the context of disease severity. Thus, the aforementioned miRNAs are not ideal reference normalizers for RT-qPCR, because it requires stable expression across all individuals[17].

RefFinder[7,8] was used to evaluate the stability of hsa-miR-34a-3p, hsa-miR-17-3p, and hsa-miR-205-3p. The results indicate that hsa-miR-17-3p is the most stable miRNA. However, this miRNA exhibited different values when comparing the two groups involved in the analyses and, thus, was disregarded. Among the two remaining candidates, hsa-miR-34a-3p was more stable than hsa-miR-205-3p.

Differential miRNA expression between COVID-19 patients and controls (healthy volunteers)		
miRNA	**Fold regulation**	***p*-Value ***
hsa-miR-34a-3p	**1**	**0.993**
hsa-miR-194-3p	**1**	**0.991**
hsa-miR-374b-3p	1	0.996
hsa-miR-518a-3p	1	0.996
hsa-miR-519b-5p	1	1.000
hsa-miR-520a-5p	1	0.997
hsa-miR-525-5p	1	0.998
hsa-miR-548am-3p	1	1.000
hsa-miR-548ay-5p	1	1.000
hsa-miR-636	1	0.995
hsa-miR-1292-5p	1	0.991
hsa-miR-3650	1	0.996
hsa-miR-3689b-5p	1	1.000
hsa-miR-3936	1	0.996
hsa-miR-3941	1	1.000
hsa-miR-4419a	1	0.992
hsa-miR-4520-5p	1	1.000
hsa-miR-4524a-5p	1	0.995
hsa-miR-4721	1	0.990
hsa-miR-4740-3p	1	0.992
hsa-miR-4766-3p	1	0.998
hsa-miR-5095	1	0.997
hsa-miR-6508-3p	1	0.990
hsa-miR-6516-5p	1	0.993
hsa-miR-6762-3p	1	0.997
hsa-miR-6802-3p	1	0.994
hsa-miR-7160-5p	1	0.994

Table 2. Sequencing results for the plasma miRNAs with fold regulation = 1 and p-value ≥ 0.990, indicating differential miRNA expression between patients with COVID-19 and controls (healthy volunteers). The miRNAs are ordered in the table in ascending numerical order according to their respective nomenclature. The two miRNAs that have been identified as potential normalizers are highlighted in bold. *Wald Test.

To our knowledge, no studies have been published associating these miRNAs with potential endogenous normalizers. However, hsa-miR-34a-3p was previously shown to be associated with cell proliferation and apoptosis in meningioma cells[18] and vascular proliferation in acute pulmonary embolism[19]. Most importantly, it is associated with the pathophysiology of diabetes[20]. As diabetes is a relatively common comorbidity among patients with COVID-19, this miRNA may not be suitable as a normalizer in these patients. By contrast, hsa-miR-205-3p was associated only with bladder[21], lung[22], breast[23], and gastric cancers[24], and none of the patients in our cohort had such conditions.

Therefore, hsa-miR-205-3p is the most suitable normalizer for comparing plasma miRNAs from patients with COVID-19 with those from healthy volunteers. However, when comparing distinct COVID-19 severities, there was no indication that it would be the best endogenous normalizer, thus, leaving it as a subject to be addressed in future research.

Limitations

This study had two significant limitations. First, to avoid potential bias in the interpretation of the results, the control and case groups should differ only with reference to COVID-19 for a better study design. While precautions were taken to match the groups in terms of age and sex, the same was not possible regarding ethnicity and comorbidities, such as diabetes and systemic arterial hypertension. Second, ideally, the number of participants in the validation cohort should have been larger. Therefore, some issues must be addressed to establish hsa-miR-205-3 as an endogenous normalizer.

Conclusions

In the context of COVID-19, there is currently no ideal normalization strategy for applications in studies assessing miRNA expression using the gold standard RT-qPCR method. However, an appropriate endogenous normalizer should possess characteristics such as stability under different conditions and equivalent expression

Differential miRNA expression between severe/critical COVID-19 patients and mild COVID-19 patients		
miRNA	**Fold regulation**	***p*-Value***
hsa-miR-17-3p	**1**	**0.995**
hsa-miR-205-3p	**1**	**0.999**
hsa-miR-513a-5p	1	0.992
hsa-miR-596	1	0.990
hsa-miR-670-3p	1	0.996
hsa-miR-922	1	0.993
hsa-miR-1207-3p	1	0.996
hsa-miR-2682-3p	1	0.997
hsa-miR-3133	1	0.992
hsa-miR-3187-3p	1	0.995
hsa-miR-3605-5p	1	0.999
hsa-miR-3907	1	0.992
hsa-miR-4329	1	0.997
hsa-miR-4443	1	0.999
hsa-miR-4474-5p	1	0.997
hsa-miR-4493	1	0.991
hsa-miR-4650-3p	1	0.990
hsa-miR-4667-3p	1	0.995
hsa-miR-4701-5p	1	0.990
hsa-miR-4799-3p	1	0.998
hsa-miR-5090	1	0.999
hsa-miR-5571-5p	1	0.999
hsa-miR-5693	1	1.000
hsa-miR-6717-5p	1	0.993
hsa-miR-6737-3p	1	0.994
hsa-miR-6746-3p	1	0.999
hsa-miR-6750-3p	1	0.997
hsa-miR-6755-5p	1	0.992
hsa-miR-6847-5p	1	0.994
hsa-miR-7844-5p	1	0.994
hsa-miR-7856-5p	1	0.993

Table 3. Sequencing results for the plasma miRNAs with fold regulation = 1 and p-value ≥ 0.990, indicating differential miRNA expression between patients with severe/critical COVID-19 and those with mild COVID-19. The miRNAs are ordered in ascending numerical order according to their respective nomenclature in the table, with potential normalizers highlighted in bold. *Wald Test.

across all individuals. Thus, this study aimed to determine the best endogenous normalizer for assessing plasma miRNA expression in patients diagnosed with COVID-19.

Based on the experimental results, the miRNA that most closely approximates the ideal conditions for use as an endogenous normalizer when comparing plasma miRNAs from patients with COVID-19 versus healthy volunteers is hsa-miR-205-3p. Future research is expected to define the most suitable normalizer for different COVID-19 severities.

Considering the significance of the topic presented by COVID-19 in recent years and ongoing research, we believe that this work will pave the way for further research and advancements in science and technology within the context of the disease.

Materials and methods

Ethical considerations

This study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the *Universidade Estadual de Campinas* (protocol code: 31049320.7.1001.5404, July 25, 2020, and 36041420.0.000.5404, August 15, 2020). Informed consent was obtained from all participants or their legal guardians.

Participants characterization

This was a case-control study. The participants were subdivided into two groups according to their SARS-CoV-2 results in the reference RT-PCR test. For the first one, the "case group," the inclusion criteria consisted

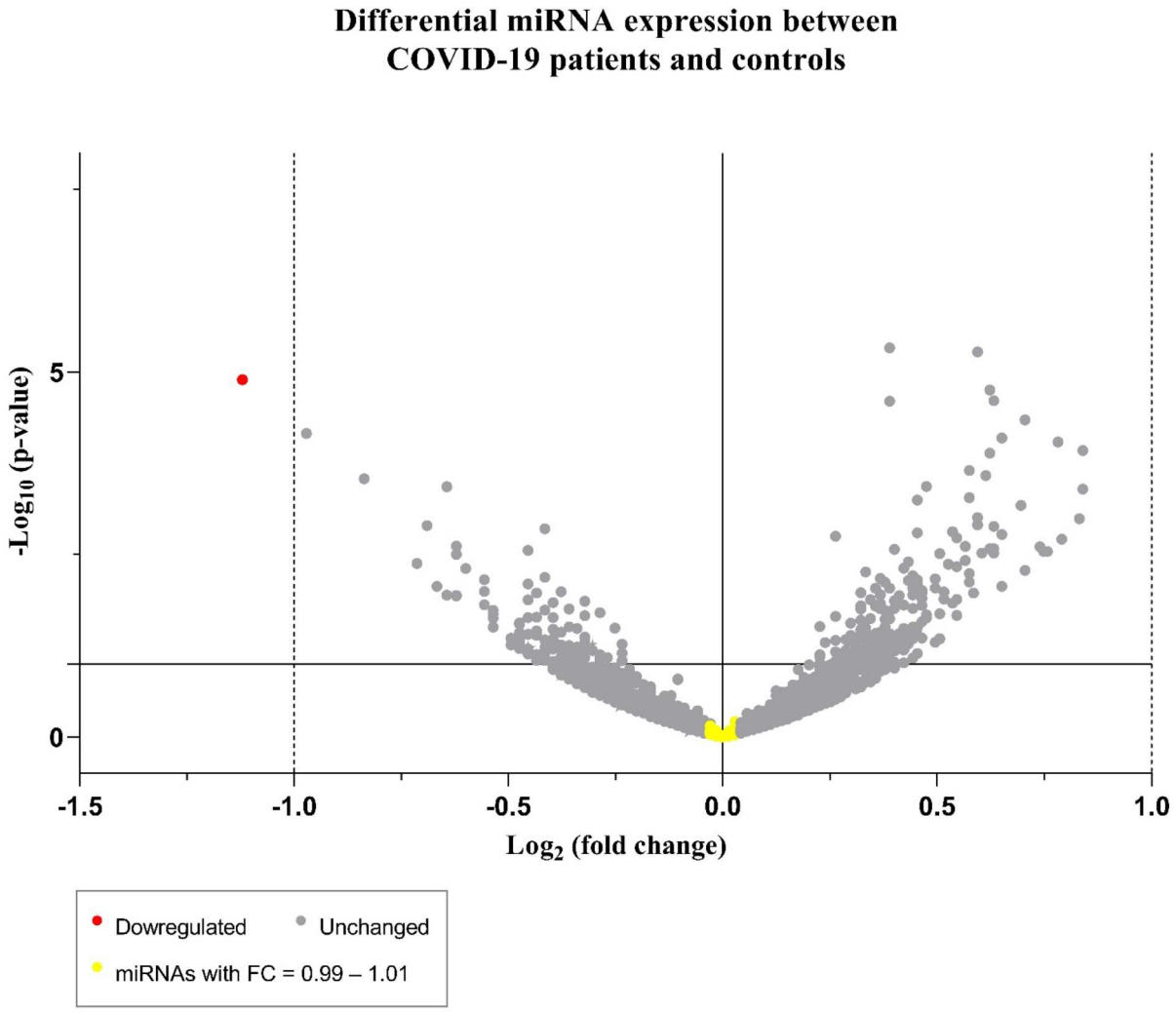

Fig. 2. Volcano plot illustrating plasma miRNA expression according to RNA-sequencing technique. Samples are from case and control groups (respectively, participants with a positive RT-PCR test for SARS-CoV-2 and participants with a negative RT-PCR test for SARS-CoV-2). The miRNAs that had no difference in expression between the two groups are shown as gray dots. miRNAs with a fold change = 0.99–1.01 are shown as yellow dots (the miRNAs mentioned in Table 2). FC, fold change = miRNA expression in the case group (patients with COVID-19)/miRNA expression in the control group (healthy volunteers).

of individuals aged between 18 and 80 years, who tested positive for SARS-CoV-2. These participants were hospitalized at the *Hospital Estadual de Sumaré Dr. Leandro Francheschini* (HES), the *Hospital de Clínicas* of University of Campinas (HC-UNICAMP), the *Centro de Saúde da Comunidade* (CECOM UNICAMP), or the Basic Health Units of Paulínia, all located in the state of São Paulo, Brazil. In the "control group," participants were included according to the following criteria: individuals between 18 and 80 years old and members of the UNICAMP community with a negative RT-PCR result for SARS-CoV-2. These individuals, in addition to having negative test results, had no flu-like symptoms, had no contact with SARS-CoV-2 infected people, did not participate in frontline COVID virus control efforts, and were monitored for 15 days after sample collection to ensure that they did not exhibit COVID-19-related symptoms.

Additionally, a second division was performed within the case group, separating it into "Patients with mild/moderate COVID-19" and "Patients with severe/critical COVID-19." The severity criteria used were based on NIH Coronavirus Disease 2019 Treatment Guidelines[6]: *mild*, individuals with common symptoms of COVID-19 but without shortness of breath, dyspnea, or abnormal chest imaging; *moderate*, patients with evidence of lower respiratory disease or imaging and with oxygen saturation (SpO2) ≥ 94% in room air; *severe*, individuals with SpO2 < 94% in room air, a ratio of arterial partial pressure of oxygen to fraction of inspired oxygen (PaO2/FiO2) < 300 mm Hg, a respiratory rate > 30 breaths/min, or lung infiltrates > 50%; and *critical*, individuals who presented respiratory failure, septic shock, and/or multiple organ dysfunction.

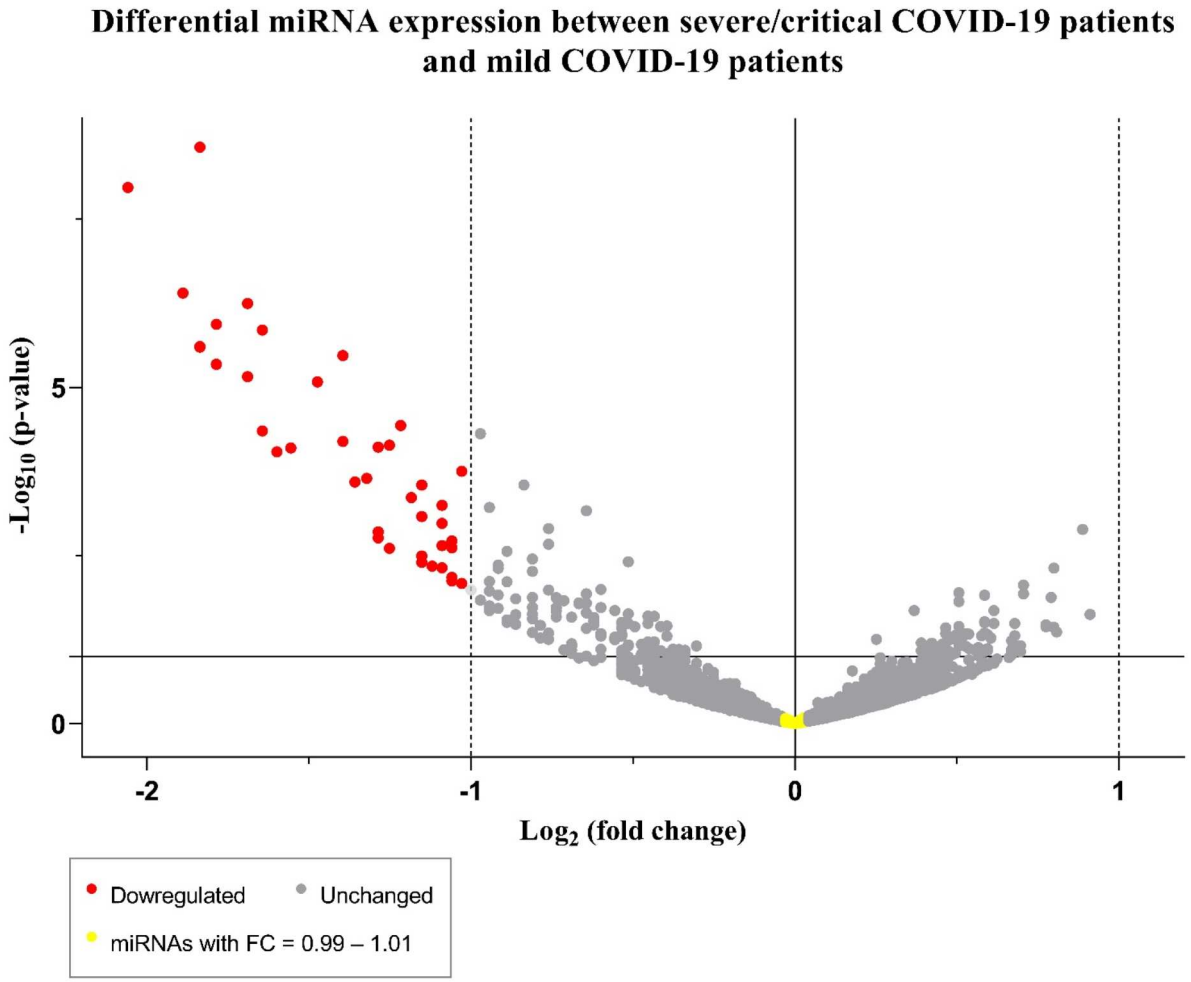

Fig. 3. Volcano plot illustrating plasma miRNA expression according to RNA-sequencing technique. Samples are from severe and mild groups (respectively, participants with severe COVID-19 and mild COVID-19). The least expressed miRNAs in the group of patients with severe COVID-19 are shown as red dots; the miRNAs that had no difference in expression between the two groups are shown as gray dots. miRNAs with a fold change = 0.99–1.01 are shown as yellow dots (the miRNAs mentioned in Table 3). FC, fold change = miRNA expression of severe or critical group/miRNA expression of mild or moderate group.

miRNA	ΔCt (mean ± standard deviation)			Fold-change	*p*-Value*
	Control group (*n* = 40)	Case group (*n* = 85)	*p*-Value*		
hsa-miR-17-3p	15.12 ± 1.79	14.02 ± 1.38	< 0.005	1.06	0.667
hsa-miR-34a-3p	26.95 ± 3.49	26.78 ± 3.50	0.399	1	0.992
hsa-miR-205-3p	25.25 ± 1.89	24.61 ± 3.11	0.080	1.22	0.193

Table 4. Expression of endogenous miRNAs (mean ± standard deviation), comparing the control group versus the case group data. ΔCt (Ct miR of interest - Ct cel-miR-39); Ct, threshold cycle; n, absolute number of participants. *t test for nonparametric data, considering $p < 0.05$ as a significant value.

The participants were characterized by age, sex, and comorbidities. Data were collected from patient records (case group) and questionnaires (control group). Patients with severe/critical COVID-19, mild/moderate COVID-19, and healthy controls were matched for sex and age.

Sample collection

Peripheral venous blood samples were collected from all participants in EDTA tubes. In the case group, all samples were collected during hospitalization within 10 days of symptom onset. Plasma samples were separated

miRNA	ΔCt (mean ± standard deviation)			Fold change	*p*-Value*
	Mild COVID-19 group ($n=45$)	Severe/critical COVID-19 group ($n=39$)**	*p*-Value*		
hsa-miR-17-3p	13.66 ± 1.20	14.45 ± 1.46	0.008	1	0.995
hsa-miR-34a-3p	24.66 ± 2.44	29.22 ± 2.90	< 0.005	0.97	0.953
hsa-miR-205-3p	26.23 ± 2.05	23.07 ± 3.25	< 0.005	1	0.999

Table 5. Expression of endogenous miRNAs (mean ± standard deviation), comparing the patients with mild COVID-19 versus the those with severe/critical COVID-19 data. ΔCt (Ct miRNA of interest - Ct cel-miR-39); Ct, threshold cycle; n, absolute number of participants. *t test for nonparametric data, considering $p < 0.05$ as a significant value. **$n = 39$, as two samples failed on the RT-PCR amplification.

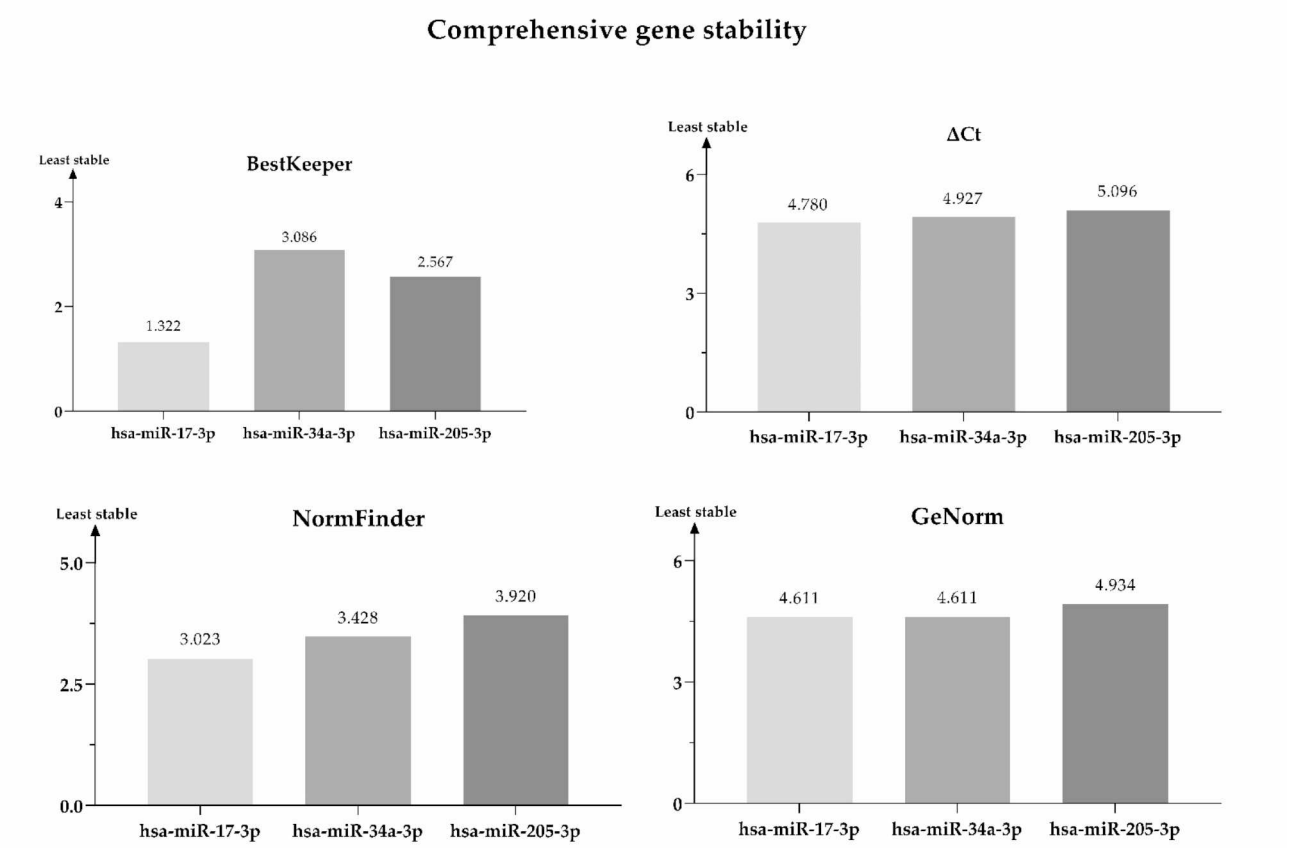

Fig. 4. Stability of candidate miRNAs for normalization chosen by the RT-qPCR results comparing differentially expressed miRNAs between patients with COVID-19 and controls (healthy volunteers), determined by different algorithms (RefFinder, http://blooge.cn/RefFinder/). On the ordinate axis, higher values indicate lower the stability of the miRNA.

from the whole blood by centrifugation at 2500 rpm, at 4 °C for 10 min, and stored in a -80 °C biobank freezer until the experiments were conducted.

Criteria for defining possible normalizer candidates

Three criteria were employed to select candidate normalizer miRNAs from RNA-Seq data: (1) miRNA must have a fold-change value equal to 1; (2) miRNA must have a *p*-value > 0.990; and (3) considering the first two criteria, the third criterion was to select the miRNAs that were discovered the earliest (i.e., those with the smallest number in their nomenclature). For RT-qPCR validation, the most adequate normalizer miRNA was the one with no difference in expression between the evaluated groups ($p > 0.05$).

miRNA extraction and cDNA synthesis

miRNA extraction was performed using 200 µL of each plasma sample with the miRNeasy Serum/Plasma Kit (Qiagen, Germany), following the manufacturer's instructions. The samples were once again kept in a -80 °C freezer after the experiment was completed.

To ensure an accurate analysis of miRNA expression using the RT-qPCR technique, 15 fmol of the synthetic miRNA cel-miR-39 was added. This exogenous miRNA (spike-in) served as a quality control for the entire technical process. If the sample threshold cycle (Ct) was outside two standard deviations from the mean Ct of the exogenous control, the sample was re-extracted. This was exclusively the case with the miRNA samples used for RT-qPCR validation; the samples extracted for sequencing did not contain cel-miR-39.

cDNA was synthesized using the TaqMan™ Advanced miRNA cDNA Synthesis kit (Applied Biosystems, Waltham, MA, USA) following the manufacturer's instructions.

miRNA sequencing

The sequencing cohort comprised 12 participants ($n=12$). For library construction, 5 µL of each miRNA sample was used, along with the QIAseq® miRNA Library Kit (Qiagen, Germany), following the manufacturer's instructions. Quality control of the samples was also performed by analyzing 1 and 2 µL of each miRNA sequencing library using an Agilent Bioanalyzer and a Qubit fluorometer, according to the manufacturer's specifications. Library preparations were sequenced using an Illumina HiSeq 2500 platform, generating 75 bp single-end reads, which were sent to the Life Sciences Core Facility (LaCTAD) at UNICAMP for sequencing.

RT-q PCR for the complete cohort

RT-qPCR was used to validate the results obtained by sequencing as it is considered the gold standard for assessing miRNA expression. The validation cohort comprised 125 participants ($n=125$), including the 12 participants from the sequencing cohort.

The RT-qPCR was performed at the Clinical Pharmacy Laboratory (CliPhar), Faculty of Pharmaceutical Sciences, UNICAMP. RT-qPCR was performed on a Rotor-Gene Q (Qiagen, Germany) using TaqMan™ Advanced miRNA Assays (Applied Biosystems, Waltham, MA, USA) for the selected miRNAs as possible normalizers, as well as for cel-miR-39 (spike-in). The total reaction volume was reduced to 10 µL, consisting of 5 µL of TaqMan® Fast Advanced Master Mix (2×) (Applied Biosystems, USA), 0.5 µL of TaqMan® Advanced miRNA Assay (20×) (Applied Biosystems, USA), 2 µL of RNase-free water, and 2.5 µL of diluted cDNA (1:10). The reactions were performed in duplicates. Raw data were evaluated using Rotor-Gene Q Series Software 2.1.0.9 (Qiagen, Germany). As part of quality control, samples with cel-miR-39 expression above two standard deviations were excluded from the analysis.

Stability assessment of selected endogenous miRNAs

The stability assessment of the endogenous miRNAs chosen by the RT-qPCR results was performed using the RefFinder online tool[7,8], which encompasses four different normalization tools commonly utilized: BestKeeper, ΔCt, NormFinder, and GeNorm, each employing distinct algorithms to assess gene expression stability, comparing the stability values of other endogenous controls; the lower the stability value, the more appropriate is the endogenous control.

Statistical analysis

For the clinical and demographic data analyses, absolute frequencies/percentages, measures of position (mean), and dispersion (standard deviation) were presented. Data normality was assessed using the Shapiro–Wilk test. During the validation phase, miRNA expression was compared between the COVID-positive and COVID-negative groups using the Mann–Whitney U test. A p-value < 0.05 was considered statistically significant for all analyses. All statistical analyses were performed using the GraphPad Prism v.9.1.0 software for Windows (GraphPad Software, Inc., San Diego, CA, USA).

The stability of endogenous control miRNAs was assessed using the online in silico prediction tool RefFinder (http://blooge.cn/RefFinder/)[7,8], which comprises four normalization methods: BestKeeper, Comparative ΔCt, NormFinder, and GeNorm. Briefly, it employs distinct algorithms concurrently based on different statistical endpoints, and assigns a value to each gene to determine the most consistently expressed gene.

BestKeeper is an Excel-based tool that determines the best internal reference gene using a pairwise correlation analysis between the expression levels of all candidate genes[25]. GeNorm is also an Excel-based approach that determines the most stable candidate from a set based on the average expression stability value, which is the average pairwise variation of a specific gene with all other genes[26]. NormFinder, unlike GeNorm, assesses each candidate's expression stability independently and can also evaluate system-induced errors as it allows a direct measure of the estimated expression variation[27]. Finally, the comparative ΔCt method uses the gene stability indicator, which is the mean of standard deviation values derived from a comparison between a particular reference gene and any other candidate; the lower the arithmetic mean, the more stable is the gene[28].

Data availability

The datasets generated and/or analyzed during the current study are available from the Research Data Repository of the Universidade Estadual de Campinas, https://doi.org/10.25824/redu/WWUNT1. Accessed on 01 June 2023.

Received: 13 June 2024; Accepted: 8 October 2024
Published online: 05 November 2024

References

1. Ministério da Saúde do Brasil. Coronavírus, Ministério da Saúde. *Ministério da Saúde* https://www.gov.br/saude/pt-br/assuntos/coronavirus (2022).
2. World Health Organization. Coronavirus (COVID-19) Cases -Dashboard https://data.who.int/dashboards/covid19/cases?n=c (2023).
3. Wayan Suryasa, I., Rodríguez-Gámez, M. & Koldoris, T. The COVID-19 pandemic. *Int. J. Health Sci. (Qassim)*. **5**, VI–IX (2021).
4. Lagos-Quintana, M., Rauhut, R., Lendeckel, W. & Tuschl, T. Identification of novel genes coding for small expressed RNAs. *Science (1979)*. **294**, 853–858 (2001).
5. de Torso, N. G. et al. Data normalization of urine miRNA profiling from Head and Neck Cancer patients treated with cisplatin. *Int. J. Mol. Sci.* **24**, (2023).
6. COVID-19 Treatment Guidelines Panel. Coronavirus Disease 2019 (COVID-19) treatment guidelines. *Natl. Institutes Health* https://www.covid19treatmentguidelines.nih.gov/
7. Xie, F., Wang, J. & Zhang, B. RefFinder: A web-based tool for comprehensively analyzing and identifying reference genes. *Funct. Integr. Genomics*. **23**, 1–5 (2023).
8. Xie, F. et al. A miRNA analysis tool for deep sequencing of plant small RNAs. *Plant. Mol. Biol.* **80**, 75–84 (2012).
9. Peltier, H. J. & Latham, G. J. Normalization of microRNA expression levels in quantitative RT-PCR assays: Identification of suitable reference RNA targets in normal and cancerous human solid tissues. *RNA*. **14**, 844–852 (2008).
10. Costa, M. C., Leitão, A. L. & Enguita, F. J. MicroRNA profiling in plasma or serum using quantitative RT-PCR. In *Methods in Molecular Biology* vol. 1182 121–129 (Humana Press Inc., 2014).
11. Kreth, S. et al. Identification of valid endogenous control genes for determining gene expression in human glioma. *Neuro Oncol.* **12**, 570–579 (2010).
12. Davoren, P. A., McNeill, R. E., Lowery, A. J., Kerin, M. J. & Miller, N. Identification of suitable endogenous control genes for microRNA gene expression analysis in human breast cancer. *BMC Mol. Biol.* **9**, (2008).
13. Belmonte, T. et al. Addressing the unsolved challenges in microRNA-based biomarker development: Suitable endogenous reference microRNAs for SARS-CoV-2 infection severity. *Int. J. Biol. Macromol.* **269**, (2024).
14. Solayman, M. H. M. et al. Identification of suitable endogenous normalizers for qRT- PCR analysis of plasma microRNA expression in essential hypertension. *Mol. Biotechnol.* **58**, 179–187 (2016).
15. Kukurba, K. R. & Montgomery, S. B. RNA sequencing and analysis. *Cold Spring Harb Protoc.* **2015**, 951–969 (2015).
16. Ban, E. & Song, E. J. Considerations and suggestions for the Reliable analysis of miRNA in plasma using qRT-PCR. *Genes (Basel)* **13**, (2022).
17. Kozera, B. & Rapacz, M. Reference genes in real-time PCR. *J. Appl. Genet.***54**, 391–406 (2013).
18. Werner, T. V. et al. MiR-34a-3p alters proliferation and apoptosis of meningioma cells in vitro and is directly targeting SMAD4, FRAT1 and BCL2. *Aging* 9, (2017).
19. Li, Y. et al. MiR-34a-3p suppresses pulmonary vascular proliferation in acute pulmonary embolism rat by targeting DUSP1. *Biosci. Rep.* **42**, (2021).
20. Mone, P. et al. Functional role of miR-34a in diabetes and frailty. *Front. Aging* vol. 3 Preprint at (https://doi.org/10.3389/fragi.2022.949924 2022).
21. Zhenhai, Z. et al. MiR-205-3p suppresses bladder cancer progression via GLO1 mediated P38/ERK activation. *BMC Cancer* **23**, (2023).
22. Xu, L. et al. Mir-205-3p promotes lung cancer progression by targeting APBB2. *Mol. Med. Rep.* **24**, (2021).
23. Qiu, C., Huang, F., Zhang, Q., Chen, W. & Zhang, H. Mir-205-3p promotes proliferation and reduces apoptosis of breast cancer MCF-7 cells and is associated with poor prognosis of breast cancer patients. *J. Clin. Lab. Anal.* **33**, (2019).
24. Zhang, Z. et al. Advantages of restoring mir-205-3p expression for better prognosis of gastric cancer via prevention of epithelial-mesenchymal transition. *J. Gastric Cancer*. **20**, 212–224 (2020).
25. Pfaffl, M. W., Tichopad, A., Prgomet, C. & Neuvians, T. P. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper - Excel-based tool using pair-wise correlations. *Biotechnol. Lett.* **26**, 509–515 (2004).
26. Vandesompele, J. et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, (2002).
27. Andersen, C. L., Jensen, J. L. & Ørntoft, T. F. Normalization of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* **64**, 5245–5250 (2004).
28. Silver, N., Best, S., Jiang, J. & Thein, S. L. Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Mol. Biol.* **7**, (2006).

Author contributions

Research conception and design, C.M.N., J.T.S., N.d.G.T. and P.M.; Data acquisition, C.M.N., J.T.S., A.d.S.N., M.B.V., E.d.C.P. and P.M.; Data analysis and interpretation, C.M.N., J.T.S., N.d.G.T., A.d.S.N., M.B.V. and C.R.d.S.C.d.R.; Statistical analysis, C.M.N., J.T.S., A.d.S.N., and M.B.V; resources, P.M., M.W.P.J., L.O.R., L.A.d.S., N.D., and W.J.F.; Drafing the manuscript, C.M.N., J.T.S., N.d.G.T., P.M. and E.d.C.P.; Critical revision of the manuscript, N.d.G.T. and P.M. All authors have read and agreed to the final version of the manuscript.

Funding

This research was funded by the São Paulo Research Foundation (FAPESP), grant number 2021/04669–9. CRSCR is a recipient of a doctoral scholarship from *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior*—Brazil (CAPES), number 88887.513100/2020–00. JTS is a recipient of a scientific initiation scholarship from FAPESP, number 2021/12359–0. CMN is a recipient of a scientific initiation scholarship from FAPESP, number 2021/12387–3. The content of this article results from the Pharmaceutical Security Nucleus Project, object of Agreement No. 895688/2019, the result of a partnership between the Ministry of Justice and Public Security of Brazil, through the Fund for the Defense of Diffuse Rights, and the *Universidade Estadual de Campinas*. The content does not necessarily reflect the official position of the Ministry or the Fund on the subject under consideration.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the *Universidade Estadual de Campinas* (protocol code: 31049320.7.1001.5404, July 25, 2020, and 36041420.0.000.5404, August 15, 2020).

Declarations

Competing interests
The authors declare no competing interests.

Informed consent
Written informed consent has been obtained from the patient(s) to publish this paper.

Conflict of interest
The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Additional information
Correspondence and requests for materials should be addressed to P.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.