

Small RNA sequencing analysis identified miR159a as a novel candidate for activity in plant-derived nanovesicles from limon, hassaku, and sudachi

Received: 28 May 2025

Accepted: 2 February 2026

Published online: 09 March 2026

Cite this article as: Takakura H., Miyamoto S., Yamamoto T. *et al.* Small RNA sequencing analysis identified miR159a as a novel candidate for activity in plant-derived nanovesicles from limon, hassaku, and sudachi. *Sci Rep* (2026). <https://doi.org/10.1038/s41598-026-38951-4>

Hideki Takakura, Shingo Miyamoto, Tetsushi Yamamoto, Toshimasa Nakao, Atsushi Taga, Michihiro Mutoh & Keisuke Oda

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

Small RNA sequencing analysis identified miR159a as a novel candidate for activity in plant-derived nanovesicles from limon, hassaku, and sudachi

Hideki Takakura ^{1*}, Shingo Miyamoto ², Tetsushi Yamamoto ³, Toshimasa Nakao ⁴, Atsushi Taga ^{3,5}, Michihiro Mutoh ², and Keisuke Oda ¹

¹ Laboratory of Biopharmaceutics and Pharmacokinetics, Faculty of Pharmaceutical Sciences, Hiroshima International University, Hiroshima 737-0112, Japan.

² Department of Molecular-Targeting Prevention, Kyoto Prefectural University of Medicine, Kyoto 602-8566, Japan.

³ Pathological and Biomolecule Analyses Laboratory, Faculty of Pharmacy, Kindai University, Osaka 577-8502, Japan.

⁴ Division of Renal Surgery and Transplantation, Jichi Medical University, Shimotsuke, Tochigi, Japan.

⁵ Antiaging Center, Kindai University, Osaka, Japan

*Correspondence: Hideki Takakura E-mail: takakura@hiroko-u.ac.jp; Tel.: +81 823-73-8580

Abstract

Owing to their safety, plant-derived nanovesicles (PDNVs), particularly those derived from edible plants, are expected to be natural delivery carriers for microRNAs (miRNAs). Among PDNVs, *Citrus limon*-derived extracellular vesicle-like nanovesicles have been shown to exert antitumor effects. We previously reported that PDNVs derived from *C. limon L.* exert selective inhibitory effects on the proliferation of p53-inactivated colon cancer cells. However, miRNAs in *C. limon L.*-derived nanovesicles have not yet been identified, and it remains unclear which miRNAs are contained in these PDNVs. We herein extracted PDNVs from *C. limon L.*, *C. hassaku*, and *C. sudachi* and identified and compared the miRNAs present within them via next-generation sequencing. A total of 158 miRNAs were identified, of which 109 were known miRNAs and 49 were novel miRNAs. Comparisons of miRNAs expressed by *C. limon L.*, *C. hassaku*, and *C. sudachi* revealed 77 miRNAs that were commonly expressed. Among all the libraries, miR159a had the highest number of reads,

and the same result was obtained when it was converted to TPM. The present results raise the possibility that miR159a could be involved in the biological properties of citrus-derived nanovesicles, given that it was the most abundantly expressed miRNA.

Keywords: miRNA, Plant-derived nanovesicles, miR159a.

Introduction

Extracellular vesicles (EVs) are stable carriers of microRNAs (miRNAs) as bioactive cargo for intercellular communication. EVs are broadly defined as membrane-bound particles released from cells that cannot replicate and are distinguished from free macromolecules; consensus definitions and recommended characterization criteria are provided by the MISEV2023 guidelines¹. In 2007, EVs were shown to encapsulate miRNAs². Furthermore, in 2010, miRNAs were found to move between cells via EVs and have functions for recipient cells³. A recent study reported the secretion of extracellular vesicle-like nanovesicles from plant cells through the cell wall and, thus, they have been attracting increasing attention as carriers of interspecies communication⁴. The interspecies functions of plant miRNAs have been demonstrated, and plant-derived miRNA sequences have been detected in mammalian blood⁵. Plant-derived nanovesicles (PDNVs) were shown to be stable in the intestines, and diet is considered to be a route for the uptake of plant miRNAs⁶. However, limited information is currently available on the miRNAs present in PDNVs, and the effects of plant-derived miRNAs on humans remain largely unknown.

miRNAs are short noncoding RNAs of approximately 21 to 25 bases, and are a representative group of small RNAs that play a role in the endogenous RNA silencing mechanism^{7,8}. miRNAs are involved in the regulation of a number of cellular processes, including proliferation, differentiation, immunity, and metabolism⁷. In recent years, plant miRNAs have been detected in mammalian blood and are suspected to be related to diseases. Western donor serum was found to contain plant miR159a, and its abundance in serum extracellular vesicles inversely correlated with the incidence and progression of breast cancer in patients⁹. Furthermore, miR159a was shown to inhibit the proliferation of

breast and colorectal cancer cells by targeting the TCF7 gene and suppressing the expression of the oncogene MYC, which is downstream of the Wnt signaling pathway^{9,10}. In addition, miR159a was found to exert anti-inflammatory effects by targeting TNF- α receptors¹¹ and functioned as a regulator of cellular cholesterol efflux *in vitro*¹². However, it currently remains unknown whether miR159a is present in PDNVs derived from citrus fruits, including lemons.

PDNVs, particularly those derived from edible plants, are expected to be natural delivery carriers for miRNAs because of their safety. Previous studies have examined the effects of PDNVs on intestinal bacteria¹³, the intestinal epithelium⁶, and immune cells¹⁴. miRNAs in PDNVs have been shown to exert antitumor¹⁵, anti-inflammatory¹¹, and immunostimulatory effects¹⁶ and are being investigated for medical applications and as previously unknown plant bioactive substances. Among PDNVs, *Citrus limon L.*-derived nanovesicles reportedly exert antitumor effects¹⁷, LDL receptor-reducing effects^{18,19}, antioxidant effects^{20,21}, increased tolerance of lactobacilli to bile²², and inhibitory effects on *Clostridioides difficile* infections²³ and have been used to treat Alzheimer's disease²⁴.

We previously demonstrated that PDNVs derived from *C. limon L.* were taken up by p53-inactivated colon cancer cells via the macropinocytosis pathway and selectively inhibited cell proliferation²⁵.

The active component for these effects may be bioactive miRNAs within PDNVs, such as miR159a; however, the miRNAs present in PDNVs derived from citrus fruits, such as *C. limon L.*, have yet to be identified. The selection of *C. limon L.*, *C. hassaku*, and *C. sudachi* was based on their phylogenetic proximity and regional cultivation patterns. This enabled a comparative analysis of PDNVs miRNA expression profiles aimed at distinguishing species-specific miRNA signatures from those conserved across citrus species. Therefore, we herein extracted PDNVs from *C. limon L.*, *C. hassaku*, and *C. sudachi* and identified and compared the miRNAs present within them via next-generation sequencing.

Results

Isolation and identification of nanovesicles from *C. limon L.*, *C. hassaku*, and *C. sudachi*.

We cut *C. limon L.*, *C. hassaku*, and *C. sudachi* into pieces and made juice from the edible parts (Figure 1A). Nanovesicles were isolated from *C. limon L.*, *C.*

hassaku, and *C. sudachi* juices via ultracentrifugation. The size distribution and concentration of nanovesicles determined via nanoparticle tracking analysis (NTA) are shown in Figure 1B. The results revealed that the diameters of citrus fruit-derived nanovesicles (CDNVs) ranged between 100 and 200 nm. CDNVs were extracted from fruit juice at the same concentration ratio (80-fold). We then measured the protein and RNA concentrations in the nanovesicles. The nanoparticle concentration was the highest in *C. limon L.* and the lowest in *C. sudachi* (Table 1). The protein quantification results were not consistent with the nanoparticle concentrations. In contrast, the total RNA concentrations were in accordance with the nanoparticle concentrations (Supplementary Figure 1). We confirmed the presence of detectable RNA from all CDNVs via a bioanalyzer (Figure 1C). On the basis of these results, we confirmed that the three types of CDNVs obtained had similar particle sizes to those of typical extracellular vesicles derived from mammals and retained detectable RNA.

Analysis of sequences from three libraries

We generated a library for each of the three CDNVs to investigate the miRNAs they contained. After sequencing via Illumina NovaSeq 6000 with single-end 50 bp (SE50) reads, we removed low-quality reads and adapter-contaminated sequences. High-throughput sequencing generated 12,348,390 reads (96.09%) for *C. limon L.*, 12,013,778 reads (98.00%) for *C. hassaku*, and 10,665,932 reads (97.80%) for *C. sudachi*. To control the quality of the libraries, we examined the error rate and GC content of the sequencing results, which are shown in Supplementary Tables s1 and 2. The error rate distribution was also established (Supplementary Figure 2).

We also analyzed the number and distribution of small RNA (18~30 nucleotide (nt)) reads. The number of small RNA reads for each sample was 4,926,724 for *C. limon L.*, 7,756,137 for *C. hassaku*, and 4,208,646 for *C. sudachi*. The read length distributions of the three CDNVs revealed that the majority of the reads were 19–24 nt in size (Figure 2), with additional peaks at 21 or 24 nt.

Small RNAs are mapped to a reference genome and analyzed.

Since sequence information for *C. limon L.*, *C. hassaku*, and *C. sudachi* is not publicly available, small RNAs were mapped to the sequence information registered in miRbase22 and the Kyoto Encyclopedia of Genes and Genomes (KEGG) of the closely related species *C. sinensis* (NCBI taxonomy ID: 2711). The

mapping rate of small RNAs in all the libraries exceeded 80%, indicating that the sequencing data were of sufficient quality for downstream analyses. The mapped small RNAs were then unambiguously assigned in the following order of priority: known miRNA > rRNA > tRNA > snRNA > snoRNA > repeats > genes > NAT-siRNA > ta-siRNA > novel miRNA. Less than 60% of the rRNA in the plant samples and less than 40% of the rRNA in the animal samples generally indicated high quality. For all three libraries, the rRNA ratio was <60%, and high-quality RNA was recovered and analyzed. These data are summarized in Table 2.

Known miRNAs

Matched small RNAs were initially checked against known RNAs. A total of 14,376,407 reads were measured for known miRNAs (across all the libraries), with 11,308 reads (0.23%) for *C. limon L.*, 15,490 reads (0.20%) for *C. hassaku*, and 6,668 reads (0.16%) for *C. sudachi*. A total of 109 known miRNA sequences were identified, of which 79 were identified from *C. limon L.*, 91 from *C. hassaku*, and 72 from *C. sudachi*. Among all the libraries, miR159a had the highest number of reads, and the same results were obtained when it was converted to TPM (Figure 3A, Supplementary Figure 3A).

Novel miRNAs

A total of 8,420 reads were measured for novel miRNAs (across all the libraries), with 3,375 reads (0.07%) for *C. limon L.*, 2,905 reads (0.04%) for *C. hassaku*, and 2,140 reads (0.05%) for *C. sudachi*. A total of 49 novel miRNA sequences were identified, of which 39 were identified from *C. limon L.*, 42 from *C. hassaku*, and 25 from *C. sudachi* (Figure 3B, Supplementary Figure 3B).

In this study, we analyzed miRNAs (known and novel miRNAs) and provided detailed information for future studies, including nucleotide sequence information of rRNA, tRNA, snRNA, snoRNA, repeats, genes, NAT-siRNA, and ta-siRNA as well as the structures of precursor miRNAs (Additional Information Files).

Comparative analysis of three types of libraries containing known and novel miRNAs.

The expression of known and novel miRNAs in each library was compared and is shown in Figure 4. We identified 158 known or novel miRNAs, 109 known miRNAs, and 49 novel miRNAs in the three CDNVs. We identified 77 known or novel miRNAs expressed by all three CDNVs, of which 58 were classified as

known miRNAs and 19 as novel miRNAs. Comparisons of the two species revealed distinct miRNA expression patterns (Supplementary Figure 4).

We compared the expression levels of known and novel microRNAs in the three libraries and listed the top 10 (Table 3). Among all the libraries, miR159a-3p was the most highly expressed and was more than twice as highly expressed as the second most highly expressed. The top expressed microRNAs were common among all the libraries. The structural features of the miRNAs were visualized by predicting the secondary structure of MiR159a-3p (Fig. 5).

We created a heatmap of the miRNA sets based on log₂-transformed TPM values, sorting them by the average expression level of all expressed microRNAs and the 77 commonly expressed microRNAs (supplementary figure 5). This heatmap visualization showed that the three species share considerable similarities in their miRNA expression profiles, while also exhibiting distinct expression patterns between species.

Comparisons of the two species were performed via three libraries (Limon vs. Hassaku, Limon vs. Sudachi, and Hassaku vs. Sudachi) containing known and novel miRNA expression levels. Spearman's correlation ρ values between the three libraries were 0.60–0.65 across all samples, with similar correlations and no pairs being homologous (Supplementary Table 3).

We focused on microRNAs showing large interspecies differences in expression and calculated log₂TPM fold changes for each pairwise comparison. To ensure reproducibility, microRNAs with fewer than 100 reads were excluded, and the top 10 microRNAs with the largest log₂TPM fold changes were extracted and are presented in Supplementary Table 4.

The top 10 microRNAs exhibited marked interspecies variation in expression. Notably, novel_63 was highly expressed specifically in Hassaku, whereas members of the csi-miR167 and csi-miR482 families presented preferential expression in Limon. These microRNAs may be involved in species-specific physiological traits or stress responses and represent promising candidates for future functional studies.

As a result of the miRNA family analysis, the same number of miRNA families were identified in *Glycine max* (soybean) and *Solanum tuberosum* (potato) as in the reference genome (Supplementary Figure 6). The results of the miRNA

family analysis have focused mainly on the study of miRNA gene families. miRNAs belonging to the same family generally have similar regulatory functions, and analyzing these families (clusters) will help elucidate the cooperative regulatory mechanisms of multiple miRNAs. Although not all the identified miRNAs are listed in the database yet, the relationships between the above two plants and the three types of citrus fruits may become clear in the future.

Differences in the expression levels of known and novel miRNAs between each library and their target genes were analyzed. These results may be important for clarifying the role of PDEVs in future plant research.

***C. limon L.*, *C. hassaku*, and *C. sudachi*-derived nanovesicles show inhibitory effect of human colorectal cancer cell proliferation.**

We previously demonstrated that nanovesicles derived from *C. limon L.* suppress the proliferation of human colorectal cancer cells²⁵, suggesting that citrus-derived nanovesicles may contain components with conserved anti-proliferative activity. Notably, several microRNAs were abundantly preserved across all three citrus species, raising the possibility that shared miRNAs may contribute to common biological effects. To further explore this possibility, we extended our analysis to nanovesicles isolated from *C. hassaku* and *C. sudachi*. This approach allowed us to examine whether the inhibitory effect observed in *C. limon L.* is a species-specific property or a broader characteristic shared among CDNVs. Consistent with this hypothesis, nanovesicles from all three species inhibited colorectal cancer cell proliferation in a dose-dependent manner, although the degree of inhibition varied among species (Fig. 6). The observation that all three CDNVs exhibited similar inhibitory activities supports the possibility that commonly preserved miRNAs may underlie this shared effect.

Discussion

The present results revealed the miRNAs present in PDNVs from three types of citrus fruits: *C. limon L.*, *C. hassaku*, and *C. sudachi*. Previous studies reported that the diameter of PDNVs from *C. limon L.* was approximately 100 nm²¹. Furthermore, the majority of studies on PDNVs have shown that the diameter ranges between 100 and 200 nm. Since the number of CDNVs obtained in the present study was proportional to the amount of RNA, we initially speculated

that RNA might represent a major component of these preparations. However, because we did not perform nuclease protection assays or membrane-permeabilization controls, we cannot determine whether the detected RNA is encapsulated within vesicles or is externally associated. Therefore, the conclusion that RNA is a major internal component cannot be rigorously supported. In addition, because RNase $\pm$ detergent assays were not performed, the present study cannot determine whether the detected RNA is intravesicular or externally associated. PDNVs may deliver plant-derived small RNAs to mammalian cells; however, the present study did not determine whether these RNAs are intravesicular or externally bound. Thus, any assumption regarding RNA stability or delivery mechanisms should be interpreted with caution. We previously imaged nanoscale vesicles from citrus-derived fractions via transmission electron microscopy and demonstrated that their diameters are approximately 100–200 nm²⁵. However, definitive characterization of these fractions as extracellular vesicles (EVs) requires additional analyses, such as immunoblotting for established plant EV markers and proteomic profiling²⁶. Because citrus species such as lemon are phylogenetically distant from model plants such as *Arabidopsis*, antibody cross-reactivity and the coverage of sequence databases are currently limited; therefore, the available evidence for EV identity remains constrained. In addition, although our previous work confirmed the EV-like morphology of citrus-derived nanovesicles via transmission electron microscopy, we did not perform TEM analysis of the specific preparations used in the present study. Consequently, we cannot determine whether the nanovesicles analyzed here possess the same structural characteristics. Taken together, these limitations indicate that the vesicles characterized in this study should be interpreted as EV-like nanovesicles rather than fully validated extracellular vesicles, and this point should be considered when interpreting their biological effects.

The length distribution of RNAs differs between plants and animals. Small RNAs often peak at 21 or 24 nt in plants and at 22 nt in animals. Since the distribution of small RNAs in all the libraries peaked at 21 or 24 nt, the plant-derived small RNAs were successfully recovered. The collation of mapped miRNAs revealed that the known miRNA miR159a was the most abundant of the three libraries. However, comparison of the miRNA expression profiles across the three libraries revealed 77 miRNAs that were commonly expressed, while each library

also contained a set of uniquely expressed miRNAs. This result suggests a characteristic expression pattern common to the three types of libraries rather than contamination during CDNVs extraction or next-generation sequencing. Among the identified microRNAs, we focused on miR159a because it presented the highest abundance in all three libraries. Previous studies have shown that miR159 family members regulate transcription factors involved in Wnt signaling, and miR159a in particular has been reported to directly target TCF7 in mammalian cells. Although we did not perform target prediction or functional validation analyses in the present study, these previously reported interactions provide a contextual background for considering how citrus-derived miRNAs, including miR159a, might be linked to biological processes in mammalian cells. Despite these observations, the overall abundance of miRNAs within the vesicle RNA population was extremely low, whereas rRNA fragments dominated the small RNA profiles. This imbalance introduces uncertainty regarding the extent to which miRNAs contribute to the functional properties of citrus-derived nanovesicles. Accordingly, the potential biological roles of these miRNAs should be interpreted with caution, and their low abundance must be considered a key limitation of the present study.

The information obtained on miRNAs in the present study provides promising materials for discussion as an activator of the physiological effects of CDNVs, including *C. limon L.* in mammals. We previously reported the antitumor effects of CDNVs from *C. limon L.* on colon cancer cells in which the Wnt pathway is known to be activated²⁵. miRNAs encapsulated in PDNVs are thought to act on mammalian cells through multiple mechanisms²⁷. Encapsulation in membrane vesicles protects small RNAs from degradation in the gastrointestinal tract and facilitates their uptake into cells by endocytosis or membrane fusion, enabling delivery to distant tissues. Furthermore, endocytic activity has been reported to be affected by food components, which may have a significant effect on the correlation between diet and the incidence of diseases such as colorectal cancer.

Considering that previous studies have reported -inhibitory activity of miR159a for the Wnt pathway, it remains possible, although not demonstrated in the present study, that miR159a could be one of several factors contributing to these effects. However, we did not assess Wnt pathway activation or downstream targets in our experimental system; therefore, the involvement of miR159a-mediated Wnt inhibition remains unclear. Thus, the present study does not

provide evidence of miR159a delivery to SW480 cells or modulation of Wnt pathway components, and no causal relationship can be inferred. We confirmed that CDNVs derived from *C. limon* L., *C. hassaku*, and *C. sudachi* each dose-dependently inhibited human CRC cell proliferation. However, the present study did not include nanovesicles-depleted supernatants or further purified nanovesicles fractions (e.g., density-gradient or size-exclusion chromatography); therefore, we cannot exclude the possibility that copurified soluble metabolites contributed to the observed suppression. Although blank controls consisting of CDNVs plus WST-8 reagent without cells did not show abnormal absorbance, indicating that direct chemical interference with the WST-8 reaction is unlikely, this does not rule out cytotoxic effects of non-vesicular components. Furthermore, our previous work quantified the citric acid concentration in *C. limon* L.-derived nanovesicle fractions and demonstrated that applying the same concentration of citric acid to cells did not affect proliferation, suggesting that acidic components alone are unlikely to account for the observed effects. Nevertheless, without CDNV-depleted and gradient-purified controls, the dose-dependent suppression cannot be attributed exclusively to extracellular vesicles. These findings are consistent with the possibility that miR159a, among other vesicle components, could contribute to the observed effects, although the present data do not establish a direct link to Wnt pathway regulation.

Wang *et al.*²⁸ analyzed PDNVs from grapefruit, ginger, lemon, and grape and reported that only 15 miRNAs were consistently expressed across these species. They also demonstrated that the top 20 miRNAs accounted for the majority of the total miRNA abundance in each nanovesicles sample. In their lemon-derived EVs, miR3952 was identified as the most highly expressed miRNA, followed by miR159a. In contrast, our study, which focused on three citrus species—*C. limon* L., *C. hassaku*, and *C. sudachi*—revealed 77 commonly expressed miRNAs, with miR159a being the most abundant of all three. This discrepancy is likely attributable to differences in the taxonomic background of the plant species analyzed, which may be the primary factor. However, methodological variations, including nanovesicles isolation protocols, small RNA library preparation conditions, and miRNA annotation standards, may also contribute to the observed differences. Because miRNA abundance was derived from single libraries without biological replicates or spike-in normalization, the apparent

predominance of miR159a may reflect technical biases rather than true biological abundance. Thus, miR159a should be interpreted as the most abundant miRNA within the analyzed datasets, not necessarily across different harvests or citrus cultivars.

On the other hand, in plant production, miR159a has been reported to exert an antidrought effect and be upregulated after drought treatment²⁹. Citrus fruits are plant species that are well suited to drought-prone regions, possibly because miR159a was the most highly expressed miRNA in CDNVs from *C. limon L.*, *C. hassaku*, and *C. sudachi*.

The dynamics of miRNAs in PDNVs have not yet been clarified in detail. The present results revealed that miR159a was present in the highest amounts in CDNVs from *C. limon L.*, *C. hassaku*, and *C. sudachi*; therefore, it may be used as an indicator to clarify the dynamics of miRNAs in PDNVs in the body after their oral ingestion and may also be established as a new nutritional component in the future.

Conclusion

The present results identified miRNAs in PDNVs from three types of citrus fruits. Among the three types of CDNVs, miR159a was the most highly expressed miRNA. This study helps clarify the physiological activities of CDNVs in mammals and identify active substances.

Materials and methods

Preparation of Nanovesicles

C. limon L. was purchased from an organic farmer in Ehime, Japan. *C. hassaku* was a gift from a private farmer in Ehime, Japan. *C. sudachi* was purchased from an organic farmer in Tokushima, Japan. *C. limon L.*, *C. hassaku*, and *C. sudachi* were gently squeezed manually. The obtained juice was centrifuged at $2000\times g$ for 20 min, and the resulting supernatant was centrifuged at $10,000\times g$ for 60 min. The supernatant was filtered through a filter with 0.22- μm pores and centrifuged at $100,000\times g$ for 90 min in a fixed angle rotor P70AT (Eppendorf Himac Technologies Co., Japan). The pellet was suspended in PBS (-) and transferred to a 30% sucrose/D₂O cushion. After centrifugation at $100,000\times g$ for 180 min in a swinging bucket rotor SW 32.1 Ti (Beckman Coulter Inc., Brea,

CA, USA), the nanovesicle-containing fraction was resuspended in PBS (-) and centrifuged at $100,000\times g$ for 90 min two times. The pellet was collected and resuspended in PBS (-) for subsequent experiments.

NTA

NTA was performed as previously described²⁵. The samples were diluted 100-fold with PBS (-) and subjected to NTA. NTA was conducted five times to obtain the average value and standard error (Figure 1B). The measured value was multiplied by the dilution factor to calculate the concentration of nanoparticles in the sample (Table 1).

Protein Quantitation

Nanovesicles were dissolved in PIPA buffer (Nacalai Tesque, Japan), and protein quantitation was performed via the TAKARA BCA Protein Assay Kit (Takara, Japan) according to product information.

RNA Quantification and Qualification.

Filgen (Nagoya, Japan) performed RNA isolation and quantitation. Small RNAs were extracted from nanovesicles using the mirVana™ miRNA Isolation Kit with phenol (Thermo Scientific, Waltham, MA) and Plant RNA Isolation Aid (Thermo Scientific, Waltham, MA). RNA quantitation was performed by NanoDrop™ One (Thermo Fisher Scientific). RNA integrity was assessed using the Agilent RNA 6000 Nano Kit (Agilent Technologies, Inc., cat.# 5067-1151) with the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA).

Library Preparation and Small RNA Sequencing.

Novogene (Beijing, China) performed library preparation and sequencing analyses using 10 ng of RNA per sample. Sequencing libraries were generated using the Small RNA Lib Prep Kit for Illumina (ABclonal, RK20305) following the manufacturer's recommendations, and index codes were added to attribute sequences to each sample. Briefly, the 3' SR adaptor, 5' end adaptor, and M-MuLV Reverse Transcriptase (RNase H-) were used to synthesize first-strand cDNA. PCR amplification was performed using LongAmp Taq 2× Master Mix, the SR Primer for Illumina, and the index (X) primer. The PCR products were purified on an 8% polyacrylamide gel (100 V, 80 min). Recovered DNA fragments were dissolved in 15 µL of elution buffer, and library quality was checked on the

Agilent Bioanalyzer 2100 system via DNA high-sensitivity chips.

Small RNA sequencing was performed via the Illumina NovaSeq 6000 platform with a single-end 50 bp (SE50) read configuration. This platform utilizes sequencing-by-synthesis (SBS) chemistry, which combines bridge amplification and reversible terminator-based fluorescence detection to achieve high-throughput and accurate base calling. The SE50 configuration was selected to match the insert size of the small RNA library (18–45 bp), ensuring optimal coverage and resolution for miRNA profiling.

Data analysis

Raw Data

The original fluorescence image files from the Illumina platform were converted into short reads (raw data) via base calling, recorded in the FASTQ format³⁰, which includes sequence information and corresponding sequence quality information, and processed through fastp software (v0.23.1). Using this software, clean data (clean reads) were obtained by removing reads containing the adaptor, reads containing poly-N, and low-quality reads from the raw data.

Data Quality Control

Fastp³¹ (v0.23.1) was used to perform basic statistics on the quality of the raw reads. We discarded paired reads if either read contained adapter contamination, if more than 10% of the bases were uncertain in either read, and if the percentage of low-quality (Phred quality <5) bases was >50% in either read.

Mapping to the Reference Genome

Small RNA tags were mapped to the reference sequence via Bowtie³² software. In the mapping step of our analysis pipeline, we used Bowtie version 1.0.1 to align small RNA reads to the reference genome. The alignment was performed via the following command-line parameters: bowtie -v 1 -k 1.

Analysis of Known miRNA

Reads that matched reference sequences were compared with sequences in the specified range in miRBase, and small RNA information was obtained from mirdeep2³³, such as the miRNA secondary structure, miRNA sequence, length, and frequency of occurrence. We used miRDeep2 v0.0.5 and performed

alignment to reference miRNA sequences with the mapper.pl script via the parameters -e -h -m -s reads.fa -t reads.arf -p genome.fa.

ncRNA Analysis & Repeat Sequences Alignment

We annotated small RNAs via the ncRNA sequences of the species or selected rRNAs, tRNAs, snRNAs, and snoRNAs from RFAM³⁴ to annotate small RNAs and identify and remove potential rRNAs, tRNAs, snRNAs, and snoRNAs.

Exon and Intron Alignment

To exclude sequences of small RNA sequences that were degraded from mRNA fragments, we used Bowtie³² software, and known exon and intron sequences were compared, after which sequence information from the comparisons was counted.

Plant NAT-siRNA

These species were not present in the database, and we adopted the analysis methods of PlantNATsDB³⁵ for *de novo* predictions and detected NAT-siRNA.

Plant ta-siRNA

TAS gene analysis was performed on known TAS genes from rice and *Arabidopsis* as databases for homology comparisons. Moreover, the TAS gene was predicted via the ta-siRNA identification software UEA sRNA tools³⁶.

Novel miRNA Prediction

To predict new miRNAs, we used miREvo³⁷ and mirdeep2³³ to extract small RNA sequences of a specific length and aligned them with the reference genome. The sequences were analyzed for secondary structures, Dicer cleavage site information, energy characteristics, and other features to identify new miRNAs. We used miREvo version 1.1 with the -i -r -M -m -k -p 10 -g 50000 and performed for integration of miREvo and mirdeep2 for novel miRNA prediction.

Small RNA Annotation

A single small RNA may match multiple different annotation categories. Therefore, to ensure that each unique small RNA was assigned to only one annotation, we followed the priority order shown below. known miRNA > rRNA > tRNA > snRNA > snoRNA > repeats > genes > NAT-siRNA > ta-siRNA >

novel miRNA.

miRNA Base Edit

To identify potentially mutated miRNAs, small RNA sequences were aligned from each sample with the known and novel mature miRNAs detected as well as their precursors. miRNAs may undergo nucleotide editing at specific positions, leading to changes in the seed sequence with altered target genes³⁸.

miRNA Family Analysis

Known miRNAs were identified by miFam.dat, and their family origin was examined. Novel miRNAs were classified using RFAM to identify their RFAM family.

miRNA Expression and Differential Expression

TPM³⁹ was used to normalize the expression levels of known and novel miRNAs in each sample. $TPM = (\text{read count} \times 1,000,000) / \text{libsize}$ (libsize: total miRNA read count). For samples with biological replicates, we used DESeq2⁴⁰ for differential expression analysis between two comparison groups. DESeq2 detects differential expression in digital gene expression data via a model based on the negative binomial distribution. For samples without biological replicates, we used the edgeR⁴¹ TMM algorithm to normalize the read count data.

Differential expression was assessed via DESeq2 (v1.24.0) for experiments with biological replicates ($\text{padj} \leq 0.05$), while experiments without replicates were analyzed via edgeR (v3.24.3) applying $\text{padj} \leq 0.05$ and $|\log_2\text{foldchange}| \geq 1$.

Target Gene Predictions for Known and Novel miRNA

psRobot⁴² was used to predict known and novel miRNA target genes to obtain the relationships between miRNAs and their target genes. We used psRobot version 1.2 and performed plant target prediction with the -s -t -o -p 1.

Cell culture and cell proliferation assay

The human colorectal cancer cell line SW480 was purchased from the American Type Culture Collection. The cells were cultured in Roswell Park Memorial Institute 1640 medium supplemented with 10% exosome-depleted fetal bovine serum (System Bioscience, Palo Alto, CA, USA). All the cells were cultured at 37°C in an atmosphere containing 5% CO₂. SW480 cells were plated at a density

of 2×10^3 cells per well in 96-well plates. After preincubation for 24 hours, the cells were treated with CDEVs for 96 hours. After treatment, the cells were incubated with WST-8 reagent (Cell Counting Kit-8; Dojindo Laboratories, Kumamoto, Japan) for 4 hours at 37°C, and the optical density of the culture mixture was measured at 450 nm via an ELISA plate reader.

Acknowledgments

This work was supported by the Scholarship Fund for Young/Women Researchers from The Promotion and Mutual Aid Corporation for Private Schools of Japan (for H.T.). We thank M. Eng. Shinji Masuzaki, a gift of organic cultivated Hassaku. Part of the ultracentrifugation method for nanovesicle production was performed at the Hiroshima Prefectural Technology Research Institute Fisheries and Ocean Technologies Center.

Author contributions statement

H.T. conceived the study.

H.T. designed the study.

H.T. performed the experiments and analyzed the data.

S.M. helped with the experiments.

H.T. wrote the original manuscript.

S.M., T.Y., T.N., A.T., M.M., and K.O. supervised.

H.T. was responsible for project administration and funding.

All the authors reviewed and edited the original manuscript.

All the authors read and approved the final manuscript. All authors have read and agreed to the published version of the manuscript.

Additional information

All the authors declare no conflicts of interest.

Data availability

All small RNA expression levels and nucleotide sequence information supporting the findings of this study can be found in the paper and its Supplementary Information. The RNA-seq data generated in this study have been deposited in the DDBJ BioProject database with a link to BioProject accession number PRJDB 35681.

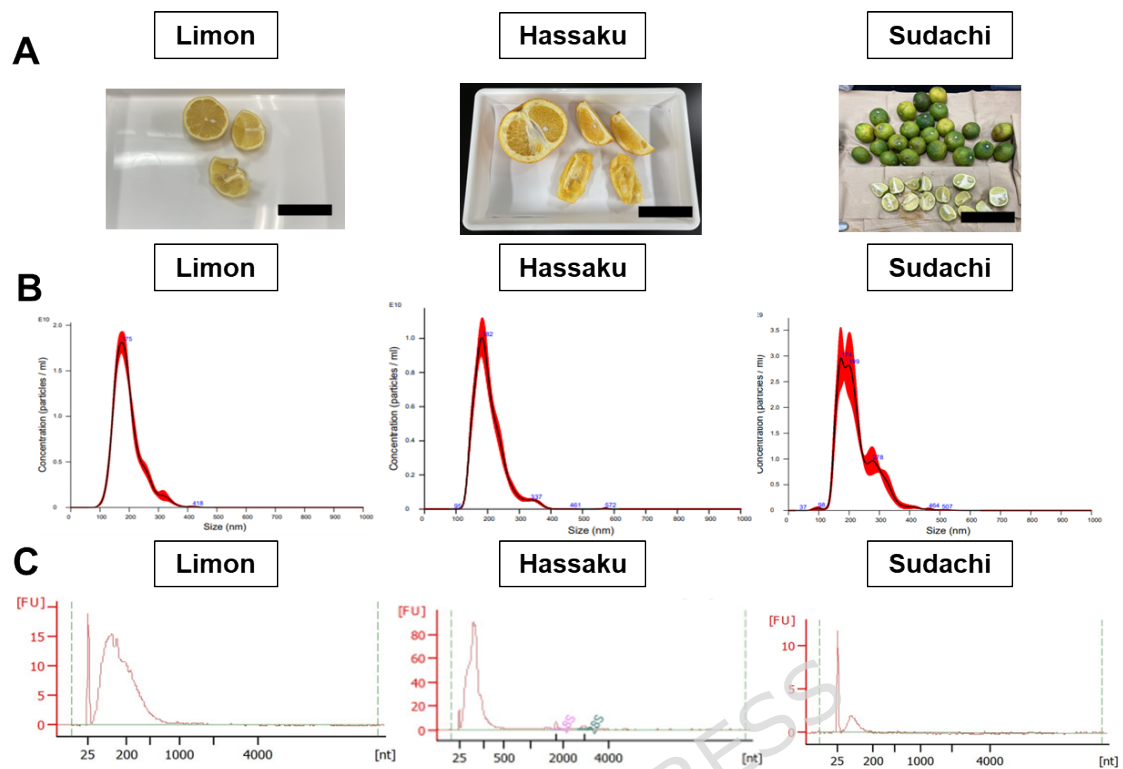


Figure 1. A, Photographs of *C. limon* L., *C. hassaku*, and *C. sudachi* from which juice was extracted. A 1-cm scale bar is shown. B, Size distribution of nanovesicles. Size distributions were analyzed by nanoparticle tracking analysis. C, RNA in the nanoparticles analyzed using the Agilent 2100 Bioanalyzer. Molecular weight markers (in nucleotides (nts)) were indicated at 25 nts. A trace showing the amount of RNA versus size in nts is shown.

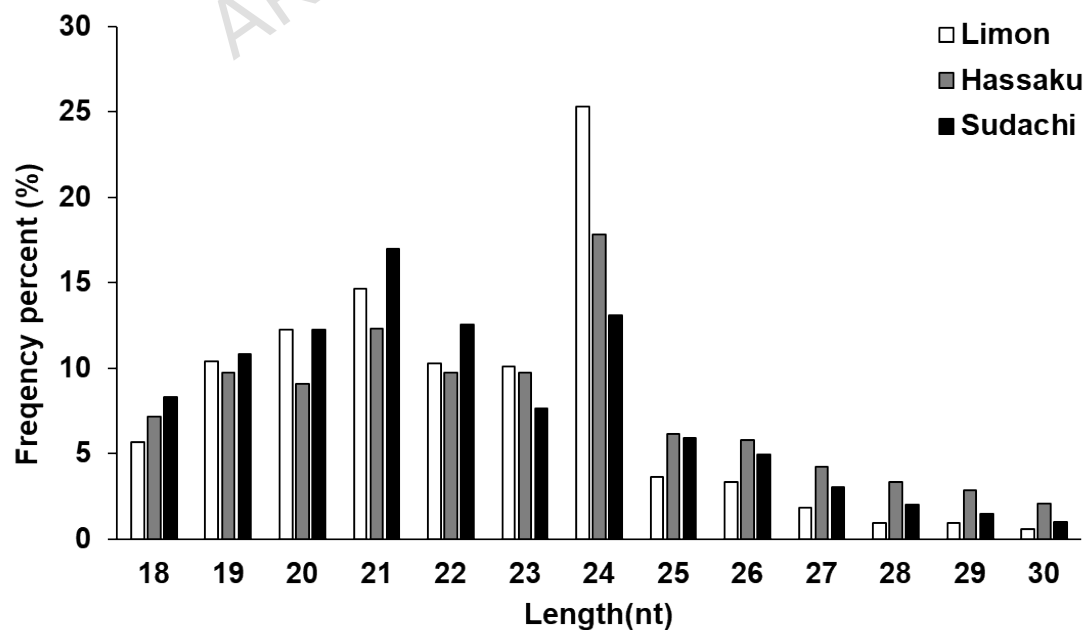


Figure 2. Distribution ratios of small RNA lengths in the three libraries.

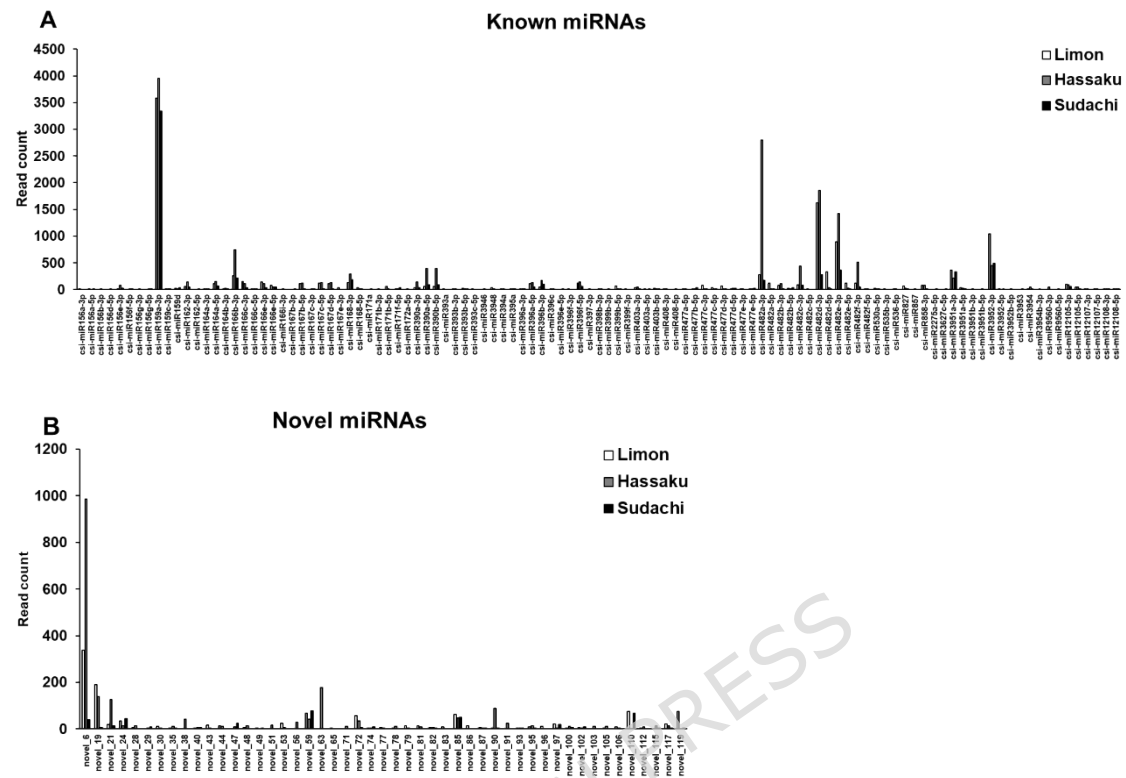


Figure 3. Distribution of member numbers and read count of each known miRNA (A) and novel miRNA (B) in the three libraries.

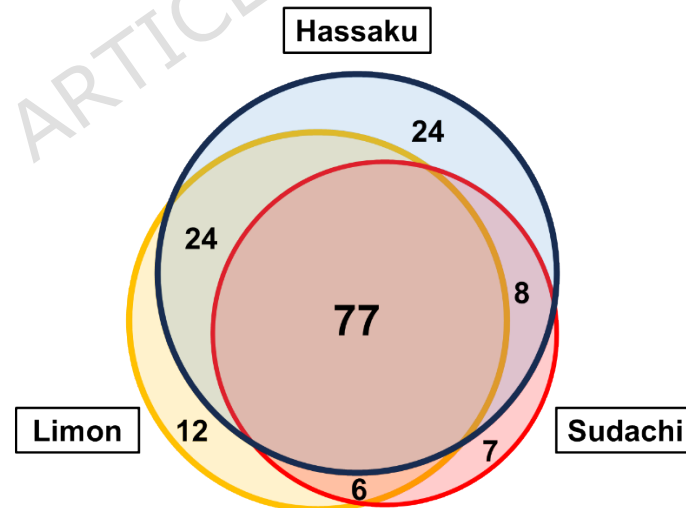


Figure 4. Venn diagram of known and novel miRNAs identified in nanovesicles derived from three types of citrus fruits.

Figure 5. The predicted hairpin structures. The figures illustrate the predicted hairpin structures of csi-MIR159a, and the red parts reflect mature sequences.

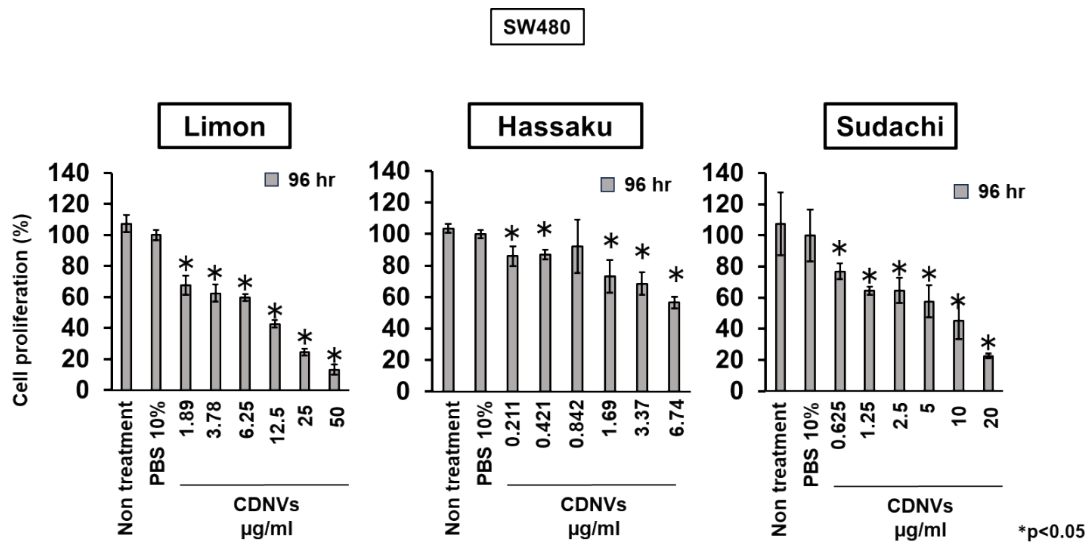


Figure 6. Inhibitory effects of citrus-derived nanovesicles (CDNVs) from limon, hassaku, and sudachi on the proliferation of human colorectal cancer cells. SW480 cells were treated with the indicated concentrations of CDEVs for 96 hours. The data are presented as the means $\pm$ standard errors; $n = 3$; *, $P < 0.05$, relative to the control. P values were calculated via one-way ANOVA followed by Dunnett's post hoc test.

Species	Particle Concentration	Protein Concentration (BCA assay)	Total RNA Concentration	Concentration factor
Limon	1.66e+11 particles/mL	491 µg/ml	14.75 µg/mL	80 times
Hassaku	8.83e+10 particles/mL	67 µg/ml	6.90 µg/mL	80 times
Sudachi	3.23e+10 particles/mL	202 µg/ml	2.49 µg/mL	80 times

Table 1. Citrus-derived nanovesicle concentrations and components.

Read Types	Limon	Hassaku	Sudachi
Total	4926724 (100 %)	7756137 (100 %)	4208646 (100 %)
Sequences mapped to genome	4031254 (81.82 %)	6636892 (85.57 %)	3708261 (88.11 %)
known_miRNA	11308 (0.23 %)	15490 (0.20 %)	6668 (0.16 %)
rRNA	2002213 (40.64 %)	4367588 (56.31 %)	2444951 (58.09 %)
tRNA	93598 (1.90 %)	177321 (2.29 %)	120760 (2.87 %)
snRNA	816 (0.02 %)	1173 (0.02 %)	607 (0.01 %)
snoRNA	15586 (0.32 %)	63850 (0.82 %)	6473 (0.15 %)
repeat	142111 (2.88 %)	114463 (1.48 %)	55836 (1.33 %)
NAT	156168 (3.17 %)	190519 (2.46 %)	68875 (1.64 %)
novel_miRNA	3375 (0.07 %)	2905 (0.04 %)	2140 (0.05 %)
TAS	21 (0.00 %)	9 (0.00 %)	N.D. (0.00 %)
exon	60712 (1.23 %)	106309 (1.37 %)	56305 (1.34 %)
intron	50556 (1.03 %)	55502 (0.72 %)	25324 (0.60 %)
other	1494790 (30.34 %)	1541763 (19.88 %)	920322 (21.87 %)
Un_annotated	895470 (18.18 %)	1119245 (14.43 %)	500385 (11.89 %)

Table 2 Numbers of reads for each small RNA classification identified.

miRNA	Sequence	Limon	Hassaku	Sudachi	Total
csi-miR159a-3p	UUUGGAUUGAAGGGAGCUCUA	3578 (1)	3957 (1)	3334 (1)	10869 (1)
csi-miR482d-3p	UCCCUACUCCACCCAUGCCAUA	1623 (2)	1856 (3)	284 (5)	3763 (2)
csi-miR482a-3p	UCUUCCCUAUGCCUCCCAUUC	277 (8)	2797 (2)	172 (8)	3246 (3)
csi-miR482e-3p	UUGCCAACUCCUCCCAUGCCGA	898 (4)	1425 (4)	368 (3)	2691 (4)
csi-miR3952-3p	UGAAGGGCCUUUCUAGAGCAC	1046 (3)	447 (8)	493 (2)	1986 (5)
novel_6	UCUUACCUAUGCCACCCAUUC	339 (6)	985 (5)	40 (24)	1364 (6)
csi-miR166b-3p	UCUCGGACCAGGCUUCAUUC	263 (9)	742 (6)	220 (6)	1225 (7)
csi-miR3951a-3p	UUUCUCUUAUCGUUAUCUGU	366 (5)	215 (13)	333 (4)	914 (8)
csi-miR482f-3p	UUUCCCAACACCUCCCAUCCC	120 (15)	513 (7)	48 (20)	681 (9)
csi-miR168-3p	CCCGCCUUGCAUCAACUGAAU	132 (13)	292 (12)	188 (7)	612 (10)

Table 3. Number of reads and sequences of the top 10 known or novel microRNAs with the highest read counts identified in the three libraries.

References

- 1 Welsh, J. A. *et al.* Minimal information for studies of extracellular vesicles (MISEV2023): From basic to advanced approaches. *J Extracell Vesicles* **13**, e12404, doi:10.1002/jev2.12404 (2024).
- 2 Valadi, H. *et al.* Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat Cell Biol* **9**, 654-659, doi:10.1038/ncb1596 (2007).
- 3 Pegtel, D. M. *et al.* Functional delivery of viral miRNAs via exosomes. *Proc Natl Acad Sci U S A* **107**, 6328-6333, doi:10.1073/pnas.0914843107 (2010).
- 4 Cai, Q. *et al.* Plants send small RNAs in extracellular vesicles to fungal pathogen to silence virulence genes. *Science* **360**, 1126-1129, doi:10.1126/science.aar4142 (2018).
- 5 Ma, J. *et al.* Exosomal microRNAs in giant panda (*Ailuropoda melanoleuca*) breast milk: potential maternal regulators for the development of newborn cubs. *Sci Rep* **7**, 3507, doi:10.1038/s41598-017-03707-8 (2017).
- 6 Ju, S. *et al.* Grape exosome-like nanoparticles induce intestinal stem cells and protect mice from DSS-induced colitis. *Mol Ther* **21**, 1345-1357, doi:10.1038/mt.2013.64 (2013).
- 7 Bartel, D. P. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell* **116**, 281-297, doi:[https://doi.org/10.1016/S0092-8674\(04\)00045-5](https://doi.org/10.1016/S0092-8674(04)00045-5) (2004).
- 8 Ha, M. & Kim, V. N. Regulation of microRNA biogenesis. *Nature Reviews Molecular Cell Biology* **15**, 509-524, doi:10.1038/nrm3838 (2014).
- 9 Chin, A. R. *et al.* Cross-kingdom inhibition of breast cancer growth by plant miR159. *Cell Res* **26**, 217-228, doi:10.1038/cr.2016.13 (2016).
- 10 Liu, J. *et al.* Soybean-derived miRNAs specifically inhibit proliferation and

- stimulate apoptosis of human colonic Caco-2 cancer cells but not normal mucosal cells in culture. *Genomics* **112**, 2949-2958, doi:10.1016/j.ygeno.2020.05.011 (2020).
- 11 Aquilano, K. *et al.* Adipocyte metabolism is improved by TNF receptor-targeting small RNAs identified from dried nuts. *Commun Biol* **2**, 317, doi:10.1038/s42003-019-0563-7 (2019).
 - 12 López de Las Hazas, M. D. C. *et al.* Dietary plant microRNAs as potential regulators of cellular cholesterol efflux. *Clin Investig Arterioscler* **36**, 315-324, doi:10.1016/j.arteri.2024.02.004 (2024).
 - 13 Sundaram, K. *et al.* Plant-Derived Exosomal Nanoparticles Inhibit Pathogenicity of *Porphyromonas gingivalis*. *iScience* **21**, 308-327, doi:10.1016/j.isci.2019.10.032 (2019).
 - 14 Wang, B. *et al.* Targeted drug delivery to intestinal macrophages by bioactive nanovesicles released from grapefruit. *Mol Ther* **22**, 522-534, doi:10.1038/mt.2013.190 (2014).
 - 15 Li, Z. *et al.* Arrowtail RNA for Ligand Display on Ginger Exosome-like Nanovesicles to Systemic Deliver siRNA for Cancer Suppression. *Sci Rep* **8**, 14644, doi:10.1038/s41598-018-32953-7 (2018).
 - 16 Teng, Y. *et al.* Grapefruit-derived nanovectors deliver miR-18a for treatment of liver metastasis of colon cancer by induction of M1 macrophages. *Oncotarget* **7**, 25683-25697, doi:10.18632/oncotarget.8361 (2016).
 - 17 Raimondo, S. *et al.* Citrus limon-derived nanovesicles inhibit cancer cell proliferation and suppress CML xenograft growth by inducing TRAIL-mediated cell death. *Oncotarget* **6**, 19514-19527, doi:10.18632/oncotarget.4004 (2015).
 - 18 Raimondo, S. *et al.* The phospholipase DDHD1 as a new target in colorectal cancer therapy. *J Exp Clin Cancer Res* **37**, 82, doi:10.1186/s13046-018-0753-z (2018).
 - 19 Raimondo, S. *et al.* Label-free quantitative proteomic profiling of colon cancer cells identifies acetyl-CoA carboxylase alpha as antitumor target of Citrus limon-derived nanovesicles. *J Proteomics* **173**, 1-11, doi:10.1016/j.jprot.2017.11.017 (2018).
 - 20 Baldini, N. *et al.* Exosome-like Nanovesicles Isolated from Citrus limon L. Exert Antioxidative Effect. *Curr Pharm Biotechnol* **19**, 877-885, doi:10.2174/1389201019666181017115755 (2018).
 - 21 Urzi, O. *et al.* Lemon-derived nanovesicles achieve antioxidant and anti-inflammatory effects activating the AhR/Nrf2 signaling pathway. *iScience* **26**,

- 107041, doi:10.1016/j.isci.2023.107041 (2023).
- 22 Lei, C. *et al.* Lemon exosome-like nanoparticles enhance stress survival of gut bacteria by RNase P-mediated specific tRNA decay. *iScience* **24**, 102511, doi:10.1016/j.isci.2021.102511 (2021).
 - 23 Lei, C. *et al.* Lemon Exosome-like Nanoparticles-Manipulated Probiotics Protect Mice from *C. difficile* Infection. *iScience* **23**, 101571, doi:10.1016/j.isci.2020.101571 (2020).
 - 24 Dolma, L., Damodaran, A., Panonnummal, R. & Nair, S. C. Exosomes isolated from citrus lemon: a promising candidate for the treatment of Alzheimer's disease. *Therapeutic Delivery* **15**, 507-519, doi:10.1080/20415990.2024.2354119 (2024).
 - 25 Takakura, H. *et al.* Citrus limonL-Derived Nanovesicles Show an Inhibitory Effect on Cell Growth in p53-Inactivated Colorectal Cancer Cells via the Macropinocytosis Pathway. *Biomedicines* **10**, doi:10.3390/biomedicines10061352 (2022).
 - 26 Koch, B. L. *et al.* Arabidopsis Produces Distinct Subpopulations of Extracellular Vesicles That Respond Differentially to Biotic Stress, Altering Growth and Infectivity of a Fungal Pathogen. *J Extracell Vesicles* **14**, e70090, doi:10.1002/jev2.70090 (2025).
 - 27 Tembo, K. M. *et al.* Exploring the bioactivity of MicroRNAs Originated from Plant-derived Exosome-like Nanoparticles (PELNs): current perspectives. *J Nanobiotechnology* **23**, 563, doi:10.1186/s12951-025-03602-9 (2025).
 - 28 Wang, F. *et al.* Identification of the Micro-Ribonucleic Acid Profiles of Plant-Derived Extracellular Vesicles and their Potential Crosskingdom Regulation. *Advanced NanoBiomed Research* **5**, 2400193, doi:<https://doi.org/10.1002/anbr.202400193> (2025).
 - 29 Fu, T. *et al.* Function identification of miR159a, a positive regulator during poplar resistance to drought stress. *Hortic Res* **10**, uhad221, doi:10.1093/hr/uhad221 (2023).
 - 30 Cock, P. J. A., Fields, C. J., Goto, N., Heuer, M. L. & Rice, P. M. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Research* **38**, 1767-1771, doi:10.1093/nar/gkp1137 (2010).
 - 31 Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884-i890,

- doi:10.1093/bioinformatics/bty560 (2018).
- 32 Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**, doi:10.1186/gb-2009-10-3-r25 (2009).
 - 33 Friedlander, M. R., Mackowiak, S. D., Li, N., Chen, W. & Rajewsky, N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res* **40**, 37-52, doi:10.1093/nar/gkr688 (2012).
 - 34 Griffiths-Jones, S., Bateman, A., Marshall, M., Khanna, A. & Eddy, S. R. Rfam: an RNA family database. *Nucleic Acids Res* **31**, 439-441, doi:10.1093/nar/gkg006 (2003).
 - 35 Chen, D. *et al.* PlantNATsDB: a comprehensive database of plant natural antisense transcripts. *Nucleic Acids Res* **40**, D1187-1193, doi:10.1093/nar/gkr823 (2012).
 - 36 Stocks, M. B. *et al.* The UEA sRNA workbench: a suite of tools for analysing and visualizing next generation sequencing microRNA and small RNA datasets. *Bioinformatics* **28**, 2059-2061, doi:10.1093/bioinformatics/bts311 (2012).
 - 37 Wen, M., Shen, Y., Shi, S. & Tang, T. miREvo: an integrative microRNA evolutionary analysis platform for next-generation sequencing experiments. *BMC Bioinformatics* **13**, 140, doi:10.1186/1471-2105-13-140 (2012).
 - 38 Wei, Y., Chen, S., Yang, P., Ma, Z. & Kang, L. Characterization and comparative profiling of the small RNA transcriptomes in two phases of locust. *Genome Biol* **10**, R6, doi:10.1186/gb-2009-10-1-r6 (2009).
 - 39 Zhou, L. *et al.* Integrated profiling of microRNAs and mRNAs: microRNAs located on Xq27.3 associate with clear cell renal cell carcinoma. *PLoS One* **5**, e15224, doi:10.1371/journal.pone.0015224 (2010).
 - 40 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
 - 41 Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140, doi:10.1093/bioinformatics/btp616 (2010).
 - 42 Wu, H. J., Ma, Y. K., Chen, T., Wang, M. & Wang, X. J. PsRobot: a web-based plant small RNA meta-analysis toolbox. *Nucleic Acids Res* **40**, W22-28, doi:10.1093/nar/gks554 (2012).