miREV: An Online Database and Tool to Uncover Potential Reference RNAs and Biomarkers in Small-RNA Sequencing Data Sets from Extracellular Vesicles Enriched Samples

Alex Hildebrandt [1]*†, Benedikt Kirchner [1]†, Esther N. M. Nolte-'t Hoen [2] and Michael W. Pfaffl [1]

*1 - **Animal Physiology and Immunology,** Technical University of Munich, Freising, Germany*
*2 - **Department of Biochemistry and Cell Biology,** Utrecht University, Utrecht, the Netherlands*

Correspondence to Alex Hildebrandt: Weihenstephaner Berg 3, 85354 Freising, Germany. *alex.hildebrandt@tum.de (A. Hildebrandt)*
https://doi.org/10.1016/j.jmb.2021.167070
Edited by Michael Sternberg

Abstract

Extracellular vesicles (EVs) are nano-sized, membrane-enclosed vesicles released by cells for intercellular communication. EVs are involved in pathological processes and miRNAs in EVs have gained interest as easily accessible biomolecules in liquid biopsies for diagnostic purposes. To validate potential miRNA biomarker, transcriptome analyses must be carried out to detect suitable reference miRNAs. miREV is a database with over 400 miRNA sequencing data sets and helps the researcher to find suitable reference miRNAs for their individual experimental setup. The researcher can put together a specific sample set in miREV, which is similar to his own experimental concept in order to find the most suitable references. This allows to run validation experiments without having to carry out a complex and costly transcriptome analysis priorly. Additional read count tables of each generated sample set are downloadable for further analysis. miREV is freely available at https://www.physio.wzw.tum.de/mirev/.

Introduction

Cells release EVs from endosomal compartments or shed them from the plasma membrane. EVs include exosomes, larger microvesicles, and apoptotic vesicles released upon cell death and even more vesicle subpopulations. They can be found in various types of tissues and body fluids of most eukaryotic species.[1] Different methods are available for EV isolation, with differential ultracentrifugation (dUC) being the most widely used method. Additional isolation techniques include methods based on precipitation, filtration, buoyant density centrifugation or chromatography.[2,3] The purification of EVs from complex biofluids is a challenge. To generate reliable and comparable EV based microRNA expression data, suitable EV isolation and characterization techniques must be applied. The international standard in EV research, the MISEV guidelines, recognize that there is no consensus on an optimal or best isolation method to get pure EVs.[4] The choice of the method depends on the research question at hand and the downstream analyses. A detailed documentation of the applied methods and the characterization of EVs is important for the reproducibility of results. To be able to interpret EV-based meta-analyses correctly, all relevant aspects of a study must be summarized and made available as meta information. By considering this information, studies with different methodologies can be compared. In addition, studies with the same or similar characteristics can be combined so that the statistical power increases. Individual characteristics of EV-studies can also be weighed for different results through exploratory data analysis.

EVs play an important role in intercellular communication through the transfer of cytosolic proteins, lipids and different types of nucleic acids, i.e. miRNA, lncRNA, mRNA, and DNA.[1] In the past, numerous studies have repeatedly pointed out the high potential of EVs as a source of biomarker molecules and how EVs in liquid biopsies could be used to develop minimally invasive diagnostic protocols.[5–11] Transcriptome analysis of EVs has mostly focused on miRNAs.[12] These short non-coding RNAs play an important role in the post-transcriptional regulation of hundreds of biological processes and pathways. An atypical level of specific, single miRNAs or an altered miRNA pattern in body fluids can indicate the presence or progression of pathophysiological processes or diseases.[13–15] In this context, it should be mentioned that extracellular RNA profiles from body fluids are not exclusively due to the RNAs in EVs. Extracellular RNAs (exRNA) can also be present in protein complexes and lipoprotein particles.[16] Most EV isolation methods enrich for EV RNAs but also allow co-isolation of other exRNAs carrier.

Candidate miRNA biomarker signatures discovered in small-RNA sequencing experiments require additional validation by reverse transcription followed by real-time quantitative-PCR (RT-qPCR). The relative expression confirmation by RT-qPCR needs stably expressed miRNA reference transcripts for reliable normalization between patients or treatment groups. These reference miRNAs can be identified either by performing a transcriptome analysis through next generation sequencing (NGS), which is time-consuming and cost-intensive, or by searching for stably expressed transcripts in extensive small-RNA sequencing data sets. When using such published miRNA references, it must be ensured that these are from the same tissue type or disease as used in the new experiment. Additionally, a comparable experimental setup and workflow must be adopted. For EV-RNA analysis, important experimental parameters and confounding methodological factors are the investigated tissue type, sample handling, and the applied EV isolation method, because RNA yield and resulting miRNA profiles strongly depend on these factors.[17,18]

We here present miREV, which can be used as a database and online tool to find stable candidate reference miRNAs within a context of choice. miREV can help scientists to save valuable resources and validate their hypotheses more quickly. We here introduce miREV and show it can be used to search for stable reference miRNA candidates using the 428 small-RNA sequencing data sets of EVs currently present in the database.

Results and discussion

The transcriptional profile of a sample is strongly influenced by the pre-analytical and experimental conditions under which the sample was acquired. Preanalytical variables, EV isolation and RNA extraction methods are as important as tissue of origin and disease state of the host.[17] miREV enables scientists to investigate the extracellular miRNA transcriptome of samples and helps estimating technical as well as biological variances for selected sample sets. Figure 1 shows an example of some result plots that can be created with miREV.

(a) Sample set creation: After user of miREV has created a sample set by the selection of variables the sample composition is displayed through tables as well as through donut plots. **(b) Clustering data according to groups:** PCA plot of a sample set of 112 data sets. 47 datasets were from patients with colon carcinoma, 16 from patients with subclinical depression and 49 from healthy volunteers. Data was normalized by median ratio of expression and transformed with the variance stabilizing transformation method. The 100 microRNAs with the greatest variance were selected for the calculation of the principal components. **(c) Finding suitable reference miRNAs:** A bar plot summarizes the most occurred stably expressed transcripts of the result lists for the sample set chosen. To compare outcome of different stability measure algorithms as well as of different normalization strategies an overlap analysis is available in miREV. **(d) Downstream analysis:** Heatmap of most significant differential expressed microRNAs. This heatmap represents the clustering results based on the significant differential expressed microRNAs found with the DESeq2 package. Results where filtered with these thresholds: adjusted p-value: padj. < 0.1; log 2-fold change: log2FC > |1|. The selected sample set has 71 data sets, including 59 serum samples from healthy patients and 12 serum samples from patients diagnosed with sepsis. None = healthy patients; sepsis = sick patients.

Sample set creation

First, the user of miREV need to create an individual sample set. This can be done by selecting experimental variables. EV-isolation method, biofluid and disease type have several selections available. A sample set is created based on this selection. The composition is shown using four donut plots (Figure 1(a)). The percentage of the sample set related to the three experimental variables and the EV characterization method used for the samples is

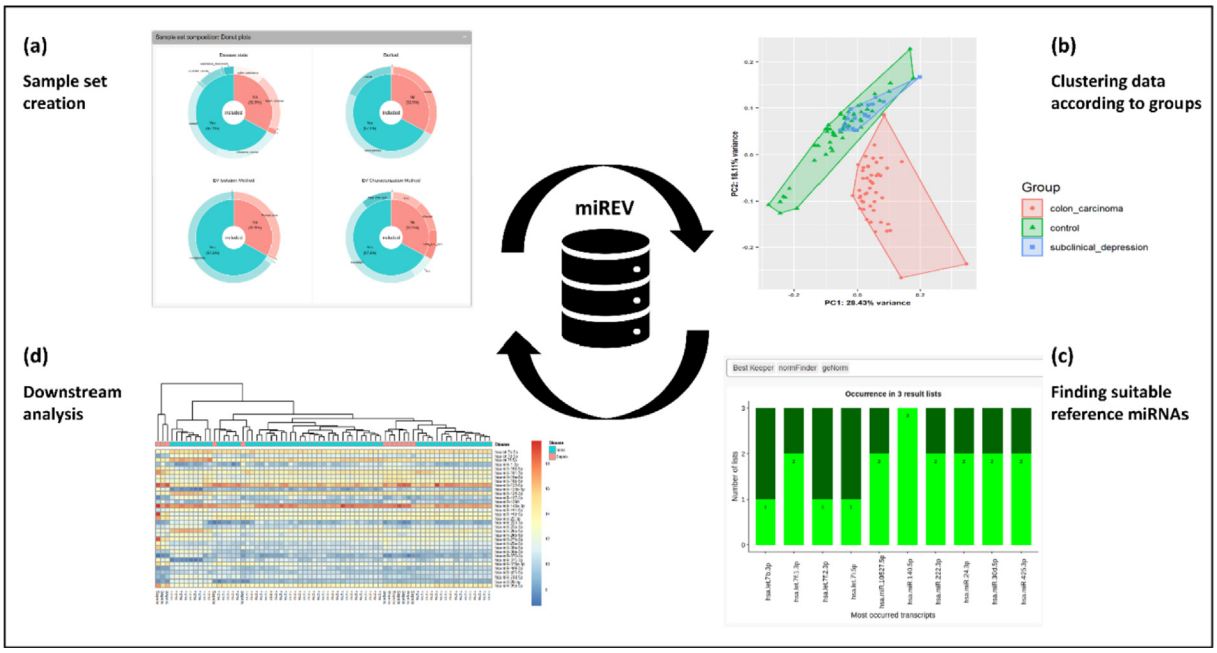

Figure 1. Exemplary result plots of miREV and a downstream analysis.

shown here. The number of samples per selection is also shown as a table in miREV for each experimental variable.

Clustering data based on metadata selection

Initial assessments can be quickly made by visualizing miRNA profiles in a principal component analysis (PCA). Figure 1(b), (Figure S3) is showing an example of the clustering of a sample set in relation to their disease status. 112 data sets are included. 47 were from patients with colon carcinoma, 16 from patients with subclinical depression and 49 from healthy volunteers. For all samples, EVs were isolated by precipitation from serum. The PCA plot shows a separation of the two diseased groups on the second component, responsible for 18.11% of total variance. In comparison, healthy controls and colon carcinoma patients are clearly separated when the first 2 components are considered, highlighting the classification potential of miREV. With a subsequent differential expression analysis (DEA), differentially expressed transcripts can be identified and further investigated. The necessary raw read counts and metadata are available to the user. To avoid misinterpretations of the PCA, one should pay attention to which experimental setup has been chosen. The sample set should differ in only one variable. For example, one cannot say anything about the origin of a variance if one has different isolation methods and different diseases in the one selected sample set. Since the PCA plot only shows one of the variables, no information about the other is visible and the variance is difficult to assign. Another possible application example would be the investigation of batch effects. These can be caused by the use of different isolation methods or biofluids for the samples within the selection. To be able to identify a batch effect with certainty, the largest possible number of data records is required. For now, the database in miREV does not provide enough statistical power to be able to carry out such investigations. However, this will be possible with the addition of new data sets.

Finding stably expressed miRNA as potential endogenous reference candidates

The quantification of miRNA or mRNA gene expression has become a standard procedure to make statements about various cellular and pathological processes. Although NGS has arrived in many laboratories, RT-qPCR is still considered the gold standard for determination of gene expression levels of a sample. The reason for this lies within the sensitivity, accuracy and speed combined with a comparably low price of this technique. Accuracy of RT-qPCR analysis, however, depends on the use of appropriate internal controls. In expression analyses, stably expressed transcripts are utilized as internal controls, according to the MIQE guidelines.[19] The normalization with these endogenous references makes it possible to compare biological samples from different contexts. Different normalization methods can be evaluated for their accuracy, precision and/or over-fitting[20] and the degree to which they influence data analysis depends, amongst others, on the data distribution.[21] miREV offers the opportunity of choosing from 6 common

normalization methods and combining them with 3 stability algorithms. This means that a total of 18 result lists are available for each selected sample set. A result list is a selection of potential reference miRNAs in descending order according to their respective stability values in the sample set. Each result list can be displayed individually or combined to review consensus results of different normalizations or stability measurements. A maximum of 30 miRNAs are listed, displaying only transcripts with the best values for stability. As an example, Figure 1 (c) and Table S4 shows the result of potentially, stably expressed miRNAs for all data sets integrated in miREV so far and lists how often individual miRNAs appear in the 18 result lists. Some of the listed miRNAs have already been described in various human studies as endogenous, stably expressed exosomal or extracellular vesicle derived RNAs. For instance *miR-103*, *miR-221* and *let-7a* have been mentioned in studies working with exosomal RNAs obtained serum with hepatocellular carcinoma, hepatitis B, and healthy patients.[22,23] Members of the *miR-30* family, such as *miR-30a* and *miR-30e*, are mentioned in studies in which the exosomal RNA was obtained from plasma.[24] Also *miR-425* is mentioned as a reference miRNA in EVs from adipose derived mesenchymal stem cells.[25] One could easily assume that the miRNAs with the highest frequency in that table can be used as universal RT-qPCR references. However, it should be noted that the stability algorithms also evaluate how often a miRNA is present within the samples of the sample set. Accordingly, a miRNA that is present in all samples of the selected sample set can be ranked as more stable than a miRNA that is only present in part of the sample set. The expression variance of the poorly ranked miRNA can, however, be smaller than that of the ubiquitous for the part of the sample set in which it is present. For example, some of the most common miRNAs in the result list (Table S4) are mentioned as differentially expressed in a wide variety of contexts. *miR-30d-5p*, for example, is discussed as an EV associated miRNA in connection with rectal cancer[26] and ovarian cancer[27] and as a regulator of conceptus-uterus interactions.[28] These examples should make clear that the result lists in miREV must always be seen in the biological context of the sample set. Top listed, potentially stably expressed miRNAs should be used as a starting point. Validation in each specific experimental setup or study is required. Information on the biological context in which the potentially stably expressed miRNAs are mentioned can give first indications about the usability.

Downstream analysis of raw read counts

miREV can be used for discovery of differentially expressed miRNAs in various diseases. This is facilitated by the fact that all data sets in miREV were processed using a standardized alignment and annotation pipeline.[29] The generated raw read counts can be downloaded for any selected sample set. The additionally available metadata makes it possible to carry out downstream analyses such as DEA or pathway analysis. Due to the uniform processing, a wide variety of sample sets can be compared with one another. There is no need for time-consuming creation of annotated read count tables. Results of an exemplary DEA are shown in Table S5 and Figure 1(d), (Figure S6). The selected sample set has 71 data sets, including 59 serum samples from healthy patients and 12 serum samples from patients diagnosed with sepsis. Since discussing all differences in miRNA abundance in this sample set would go beyond the scope of this paper, only some miRNAs that have already been associated with sepsis are highlighted. For example, *miR-223-3p* was found to be a potential biomarker for sepsis in circulating miRNAs mentioned by Zhang et al.[30] Expression levels of *miR-223* and *miR-122* are described by Wang et al. as significant higher and lower than in controls.[31] miREV does not only offer the possibility to check whether results can be reproduced in a similar context. Additional information relevant to the disease can be found. As a result of the exemplary DEA, further differentiated expressed miRNAs are listed. For example, *miR-26a-5p* is more highly expressed in the selected sample set and *miR-100-5p* is less than in the control group. The results of such a DEA can provide the user with useful additional information for their own studies. In addition, there is the possibility of comparing results and thereby strengthening a hypothesis or re-evaluating it.

Materials & methods

All data sets and calculations available in miREV were processed in advance to ensure quick response times to queries. Individual tasks of the calculation pipeline are described below and shown in Figure 2.

(a) Mapping: The mapping step processed raw NGS data sets to miRNA count tables. Data sets mapped with less than 7% to reference miRNAs were removed. **(b) Sample set creation:** Data sets were split in all possible combinations of the experimental variables. Each combination is one sample set. Resulting sample sets with less than 10 samples were removed. Additionally, miRNAs that did not pop up in at least 95% of each sample set were excluded. **(c) Normalization:** Each sample set was normalized according to six different normalization methods. **(d) Stability Calculation:** Finally, stably expressed miRNAs were determined with three different stability algorithms. For each sample set reference miRNA lists (result lists) were calculated, each representing the top 50% ranked stably expressed miRNAs.

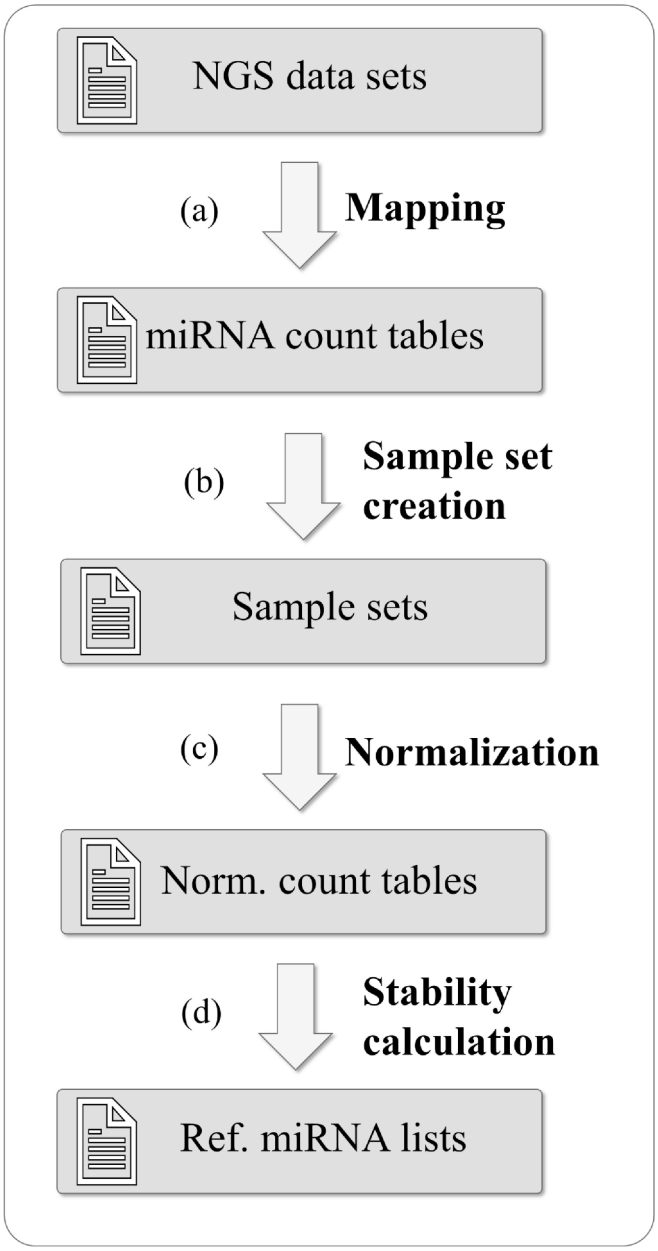

Figure 2. Calculation pipeline workflow, from raw NGS data sets to stably expressed miRNAs.

Data collection and count table generation

Currently the whole database focuses on human blood derived EVs and includes 9 different diseases and 3 different isolation methods from both serum and plasma. Appropriate and well-annotated small-RNA sequencing data sets (one data set represents all small-RNA sequencing reads of one biological sample) of circulating EVs for this meta-analysis were obtained from the gene expression omnibus database GEO.[32] Additionally, various in-house data sets as well as data sets from collaborators were added, resulting in a total of 654, which represent the starting basis of our analysis. S1 summarizes the origin of the data sets including details on EV isolation and characterization. Raw data was processed by an established alignment and annotation pipeline.[29] In brief, raw sequences were trimmed with btrim32[33] (version 0.3.0). Quality of NGS outcome was analysed for each sample with FastQC[34] (v0.11.9). The alignment of reads was performed with bowtie[35] (v.1.2.3) with a cut off for reads set to maximum one mismatch. Furthermore, parameters that limit alignment to the sense strand (--norc) and output to the single best match in terms of mismatch quality (--best) were applied. References of non-coding RNA sequences for rRNA, snRNA, snoRNA and

tRNA were downloaded from RNACentral[36] (release 12). miRNA references were obtained from miRBase[37] (release 22.1).

Sample set creation

The calculation pipeline underlying miREV utilizes read count tables to determine stably expressed transcripts of a user defined sample set. A sample set arises from the selection of experimental variables. Table S2 lists all the different experimental variables and possible selections. In total 21.483 sample sets were created to reflect all possible combinations of experimental setups.

Determination of stably expressed transcripts

Each sample set was processed by 6 popular normalization methods, all well established in next generation sequencing data analysis. Normalizing raw read counts is essential before differential expression analyses as well as for exploratory data analysis. Normalization methods are accounting for different factors such as sequencing depth, gene length and RNA composition. By scaling read counts, the expression level between samples becomes more comparable. The following methods were used: total count normalization (TC), median normalization (Med),[38] full quantile and upper quartile normalization (FQ & UQ),[39] trimmed mean of M-value (TMM)[40] and median of ratios normalization (MoR).[41] Herein applied normalization methods were shown to be the most reliable for comparison of RNA sequencing results.[42–44] All normalized sample sets were subsequently analysed by 3 different stability measure algorithms that each compute a specific stability indicator. In addition to bestKeeper,[45] which calculates the coefficient of variation (CV) across data sets, geNorm[46] and normFinder[47] were applied, which indicate expression stability via the M-value and stability measure rho, respectively. All 3 algorithms were originally designed for detecting stable mRNA or miRNA transcripts in high throughput RT-qPCR experiments, and have been cited in thousands of peer-reviewed publications.[48–50]

Filtering

The data was strictly filtered to guarantee an overall comparable and reliable sequencing data source. In a first quality filtering, data sets with less than 7% of total reads mapping to a miRNA precursor were removed to exclude sequencing results with a low miRNA level. 428 data set out of originally 654 remained. Sample sets with less than 10 samples were removed as well to get a comparable and reliable minimal size for each data set. Accordingly, 12,578 of the original 21,483 sample sets remained after this filtering step. Second, a gene filtering step was included to exclude lowly expressed genes individually for each sample set from further analysis. Genes that were not present in at least 95% of all data sets were removed, resulting in different number of miRNAs in each sample set, ranging from a few dozen to several hundreds of genes. The last filter step was applied to the result lists, which contain the proposed stably expressed miRNA transcripts. Each result list is a selection of potential reference miRNAs in descending order according to their respective stability values. miRNAs that were not amongst the top 50% of ranked transcripts were removed. This filter step ensured that only the most stably expressed transcripts are included in the result lists.

Outlook

In order for miREV to have even more selection criteria available in the experimental setup, a continuous expansion of the database is planned. Existing selection criteria are expanded and new ones, such as species, cell culture or other diseases, are added. We hereby ensure that miREV will remain freely available for at least the next five years.

CRediT authorship contribution statement

Alex Hildebrandt: Investigation, Software, Data curation, Validation, Writing - original draft, Visualization. **Michael W. Pfaffl:** Conceptualization, Writing - review & editing. **Esther N.M. Nolte-'t Hoen:** Conceptualization, Writing - review & editing. **Benedikt Kirchner:** Conceptualization, Supervision, Writing - review & editing.

Acknowledgments

We thank all further members at Division of Animal Physiology and Immunology, Technical University of Munich for small-RNA Seq. data sets they have provided. We thank Chenna R. Galiveti for an initial assessment of applicable EV studies from various sources and Ming Wu for advice in IT implementation of data pipelines and creation of an online accessible tool.

Disclosure statement

The authors reported no potential conflict of interest.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2021.167070.

Received 11 January 2021;
Accepted 22 May 2021;
Available online 28 May 2021

Keywords:
biofluid;
normalization methods;
stability algorithms;
human blood plasma & serum;
RT-qPCR;
Extracellular vesicles;
Next-generation Sequencing

† These authors contributed equally to this work.

References

1. Raposo, G., Stoorvogel, W., (2013). Extracellular vesicles: Exosomes, microvesicles, and friends. *J. Cell Biol.*, **200**, 373–383.
2. Gardiner, C., Di Vizio, D., Sahoo, S., Théry, C., Witwer, K. W., Wauben, M., et al., (2016). Techniques used for the isolation and characterization of extracellular vesicles: Results of a worldwide survey. *J. Extracell Vesicles*, **5** (1), 32945. https://doi.org/10.3402/jev.v5.32945.
3. Van Deun, J., Mestdagh, P., Sormunen, R., Cocquyt, V., Vermaelen, K., Vandesompele, J., et al., (2014). The impact of disparate isolation methods for extracellular vesicles on downstream RNA profiling. *J. Extracell Vesicles*, **3** (1), 24858. https://doi.org/10.3402/jev.v3.24858.
4. Théry, C., Witwer, K.W., Aikawa, E., Alcaraz, M.J., Anderson, J.D., Andriantsitohaina, R., et al., (2018). Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. *J Extracell Vesicles*, **7** (1)
5. J. Lin, J. Li, B. Huang, J. Liu, X. Chen, XC-T scientific world, et al. Exosomes: novel biomarkers for clinical diagnosis. hindawi.com [Internet]. [cited 2020 May 25]; Available from: https://www.hindawi.com/journals/tswj/2015/657086/abs/.
6. Michael A, Bajracharya S, Yuen P, ... HZ-O, 2010 undefined. Exosomes from human saliva as a source of microRNA biomarkers. Wiley Online Libr [Internet]. [cited 2020 May 25]; Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1601-0825.2009.01604.x
7. F. Properzi, M. Logozzi, Medicine SF-B in, 2013 undefined. Exosomes: the future of biomarkers in medicine. Futur. Med. (2013) [cited 2020 May 25]; Available from: www.futuremedicine.com.
8. H. Ogata-Kawata, M. Izumiya, D. Kurioka, one YH-P, 2014 undefined. Circulating exosomal microRNAs as biomarkers of colon cancer. ncbi.nlm.nih.gov [Internet]. [cited 2020 May 25]; Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3976275/.
9. Michael, A., Bajracharya, S.D., Yuen, P.S.T., Zhou, H., Star, R.A., Illei, G.G., et al., (2010 Jan). Exosomes from human saliva as a source of microRNA biomarkers. *Oral Dis.*, **16** (1), 34–38.
10. D.D. Taylor, C. Gercel-Taylor, MicroRNA signatures of tumor-derived exosomes as diagnostic biomarkers of ovarian cancer. Elsevier [Internet]. 2008 [cited 2020 May 25]; Available from: www.sciencedirect.com.
11. Properzi, F., Logozzi, M., Fais, S., (2013). Exosomes: the future of biomarkers in medicine. *Biomark. Med.*, **7**, 769–778.
12. G. Raposo, WS-J of C Biology, 2013 undefined. Extracellular vesicles: exosomes, microvesicles, and friends. rupress.org [Internet]. [cited 2020 Jun 23]; Available from: https://rupress.org/jcb/article-standard/200/4/373/37234/Extracellular-vesicles-Exosomes-microvesicles-and.
13. A.L.S. Revenfeld, R. Bæk, M.H. Nielsen, A. Stensballe, K. Varming, M. Jørgensen, Diagnostic and prognostic potential of extracellular vesicles in peripheral blood. Vol. 36, Clinical Therapeutics. Excerpta Medica Inc.; 2014. pp. 830–846.
14. Dong, L., Lin, W., Qi, P., Xu, M.D., Wu, X., Ni, S., et al., (2016 Jul 1). Circulating long RNAs in serum extracellular vesicles: Their characterization and potential application as biomarkers for diagnosis of colorectal cancer. *Cancer Epidemiol. Biomarkers Prev.*, **25** (7), 1158–1166.
15. Yuan, T., Huang, X., Woodcock, M., Du, M., Dittmar, R., Wang, Y., et al., (2016). Plasma extracellular RNA profiles in healthy and cancer patients. *Sci. Rep.*,.
16. Murillo, O.D., Thistlethwaite, W., Rozowsky, J., Subramanian, S.L., Lucero, R., Shah, N., et al., (2019). exRNA atlas analysis reveals distinct extracellular RNA cargo types and their carriers present across human biofluids. *Cell*, **177** (2) 463–477.e15.
17. Buschmann, D., Kirchner, B., Hermann, S., Märte, M., Wurmser, C., Brandes, F., et al., (2018). Evaluation of serum extracellular vesicle isolation methods for profiling miRNAs by next-generation sequencing. *J. Extracell Vesicles*, **7** (1), 1481321. https://doi.org/10.1080/20013078.2018.1481321.
18. Alvarez, M.L., Khosroheidari, M., Kanchi Ravi, R., Distefano, J.K., (2012). Comparison of protein, microRNA, and mRNA yields using different methods of urinary exosome isolation for the discovery of kidney disease biomarkers. *Kidney Int.*, **82** (9), 1024–1032.
19. Hellemans, J., Pfaffl, M.W., Wittwer, C.T., Garson, J.A., Huggett, J., Benes, V., et al., (2009). The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.*, **55** (4), 611–622.
20. Argyropoulos, C., Chatziioannou, A.A., Nikiforidis, G., Moustakas, A., Kollias, G., Aidinis, V., (2006). Operational criteria for selecting a cDNA microarray data normalization algorithm. *Oncol. Rep.*, **15** (4), 983–996 http://www.sanger.ac.uk/.
21. Zyprych-Walczak, J., Szabelska, A., Handschuh, L., Górczak, K., Klamecka, K., Figlerowicz, M., et al., (2015). The Impact of Normalization Methods on RNA-Seq Data Analysis. *Biomed. Res. Int.*, **2015**
22. Li, Y., Xiang, G.M., Liu, L.L., Liu, C., Liu, F., Jiang, D.N., et al., (2015). Assessment of endogenous reference gene suitability for serum exosomal microRNA expression analysis in liver carcinoma resection studies. *Mol. Med. Rep.*, **12** (3), 4683–4691.

23. Y. Li, L. Zhang, F. Liu, G. Xiang, D. Jiang, XP-D markers, et al. Identification of endogenous controls for analyzing serum exosomal miRNA in patients with hepatitis B or hepatocellular carcinoma. hindawi.com [Internet]. [cited 2020 Jul 27]; Available from: https://www.hindawi.com/journals/dm/2015/893594/abs/.
24. X. Huang, T. Yuan, M. Liang, M. Du, S. Xia, RD-E urology, et al. Exosomal miR-1290 and miR-375 as prognostic markers in castration-resistant prostate cancer. Elsevier [Internet]. [cited 2020 Jul 27]; Available from: https://www.sciencedirect.com/science/article/pii/S0302283814006873.
25. Ragni, E., Orfei, C.P., De Luca, P., Colombini, A., Viganò, M., Lugano, G., et al., (2019). Identification of miRNA reference genes in extracellular vesicles from adipose derived mesenchymal stem cells for studying osteoarthritis. *Int. J. Mol. Sci.*, **20** (5), 1108 www.mdpi.com/journal/ijms.
26. Exosomal miR-486-5p, miR-181a-5p and miR-30d-5p from hypoxic tumour cells are candidate circulating markers of high-risk rectal cancer - ProQuest [Internet]. [cited 2020 Aug 12]. Available from: https://search.proquest.com/openview/a70102e6b61c0cc4ad2abae01a5c5672/1?pq-origsite=gscholar&cbl=2030046.
27. Yamamoto, C.M., Oakes, M.L., Murakami, T., Muto, M.G., Berkowitz, R.S., Ng, S.W., (2018). Comparison of benign peritoneal fluid- and ovarian cancer ascites-derived extracellular vesicle RNA biomarkers. *J. Ovarian Res.*, **11** (1), 1–9. https://doi.org/10.1186/s13048-018-0391-2.
28. O'Neil, E.V., Burns, G.W., Spencer, T.E., (2020). Extracellular vesicles: Novel regulators of conceptus-uterine interactions?. *Theriogenology*, **1** (150), 106–112.
29. Buschmann, D., Haberberger, A., Kirchner, B., Spornraft, M., Riedmaier, I., Schelling, G., et al., (2016). Toward reliable biomarker signatures in the age of liquid biopsies - How to standardize the small RNA-Seq workflow. *Nucleic Acids Res.*, **44** (13), 5995–6018.
30. Zhang, W., Jia, J., Liu, Z., Si, D., Ma, L., Zhang, G., (2019). Circulating microRNAs as biomarkers for Sepsis secondary to pneumonia diagnosed via Sepsis 3.0. *BMC Pulm. Med.*, **19** (1)
31. Wang, H.J., Zhang, P.J., Chen, W.J., Feng, D., Jia, Y.H., Xie, L.X., (2012). Four serum microRNAs identified as diagnostic biomarkers of sepsis. *J. Trauma Acute Care Surg.*, **73** (4), 850–854.
32. Edgar, R., (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.*, **30** (1), 207–210.
33. Y.K. Genomics, 2011 undefined. Btrim: a fast, lightweight adapter and quality trimming program for next-generation sequencing technologies. Elsevier [Internet]. [cited 2020 Jun 23]; Available from: https://www.sciencedirect.com/science/article/pii/S0888754311001339.
34. S. Andrews, FastQC: a quality control tool for high throughput sequence data. 2010.
35. B. Langmead, SS-N methods, 2012 undefined. Fast gapped-read alignment with Bowtie 2. nature.com [Internet]. [cited 2020 Jun 23]; Available from: https://www.nature.com/articles/nmeth.1923.pdf?origin=ppub.
36. A. Bateman, S. Agrawal, E. Birney, E.A. Bruford, J.M. Bujnicki, G. Cochrane, et al., RNAcentral: A vision for an international database of RNA sequences. rnajournal.cshlp.org [Internet]. [cited 2020 Jun 23]; Available from: http://www.rnajournal.org/cgi/doi/10.1261/rna.2750811.
37. A. Kozomara, SG-J-N acids research, 2014 undefined. miRBase: annotating high confidence microRNAs using deep sequencing data. academic.oup.com [Internet]. [cited 2020 Jun 23]; Available from: https://academic.oup.com/nar/article-abstract/42/D1/D68/1057911.
38. Anders, S., Huber, W., (2010). Differential expression analysis for sequence count data. *Nature Preced.*,. 1–1.
39. Hu, J., He, X., (2007). Enhanced quantile normalization of microarray data to reduce loss of information in gene expression profiles. *Biometrics*, **63** (1), 50–59. https://doi.org/10.1111/j.1541-0420.2006.00670.x.
40. M. Robinson, D. McCarthy, G.S. Bioinformatics, 2010 undefined. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. academic.oup.com [Internet]. [cited 2020 May 25]; Available from: https://academic.oup.com/bioinformatics/article-abstract/26/1/139/182458.
41. S. Anders, Analysing RNA-Seq data with the DESeq package [Internet]. genomatix.de. [cited 2020 May 25]. Available from: http://www-huber.embl.de/users/.
42. Garmire, L.X., Subramaniam, S., (2012). Evaluation of normalization methods in mammalian microRNA-Seq data. *RNA*, **18** (6), 1279–1288 http://rnajournal.cshlp.org/content/18/6/1279.full.
43. Meyer, S.U., Pfaffl, M.W., Ulbrich, S.E., (2010). Normalization strategies for microRNA profiling experiments: A "normal" way to a hidden layer of complexity?. *Biotechnol. Lett.*, **32**, 1777–1788. https://doi.org/10.1007/s10529-010-0380-z.
44. Rao, Y., Lee, Y., Jarjoura, D., Ruppert, A.S., Liu, C.G., Hsu, J.C., et al., (2008). A comparison of normalization techniques for microRNA microarray data. *Stat. Appl. Genet. Mol. Biol.*, **7** (1) https://wwwdegruyter.com/view/journals/sagmb/7/1/article-sagmb.2008.7.1.1287.xml.xml.
45. Pfaffl, M.W., Tichopad, A., Prgomet, C., Neuvians, T.P., (2004 Mar). Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper - Excel-based tool using pair-wise correlations. *Biotechnol. Lett.*, **26** (6), 509–515.
46. Etschmann, B., Wilcken, B., Stoevesand, K., Von Der Schulenburg, A., Sterner-Kock, A., (2006). Selection of reference genes for quantitative real-time PCR analysis in canine mammary tumors using the GeNorm algorithm. *Vet. Pathol.*, **43** (6), 934–942.
47. NormFinder software [Internet]. [cited 2020 Apr 20]. Available from: https://moma.dk/normfinder-software.
48. J. Vandesompele, M.K. gene validation software, 2009 undefined. Real-time PCR: current technology and applications. gene-quantification.org [Internet]. [cited 2020 May 25]; Available from: http://gene-quantification.org/hkg.html.
49. M. Pfaffl, J. Vandesompele, M. Kubista, Real-time PCR: current technology and applications. 2009.
50. S.A. Bustin, V. Benes, J.A. Garson, J. Hellemans, J. Huggett, M. Kubista, et al. The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. 2009 [cited 2020 May 25]; Available from: http://www.mibbi.org.