The Utility of Pooling Salivary MicroRNA Sample for Oral Disease Related Research: Systematic Review

Sharul Nisha Ali1, Fouad Hussain Al-Bayaty1, Khor Goot Heah2, Mohd Faizal Hafez Hidayat1*, Sharul Nisha Ali1

1. Center for Periodontology Studies, Faculty of Dentistry, Universiti Teknologi MARA, Sungai Buloh Campus, 47000 Sungai Buloh, Selangor, Malaysia.
2. Center of Preclinical Sciences Studies, Faculty of Dentistry, Universiti Teknologi MARA, Sungai Buloh Campus, 47000 Sungai Buloh, Selangor, Malaysia.

Abstract
This review aims to systematically review the utility of salivary microRNA pooling for downstream analysis in oral diseases analysis. A systematic search of the literature was conducted independently from PubMed, SAGE and Science Direct for the full articles published between January 2009 to June 2019. Articles were strictly selected based on the inclusion criteria and eligible studies were assessed for the risk of bias and applicability using the QUADAS-2 tool. PRISMA guideline was followed in conducting this systematic review. A total of 122 titles and abstracts were screened. Of that, 13 articles met the inclusion criteria and were further reviewed. To date, there is still limited study utilizing the usage of salivary microRNA as their study samples. The utility of pooling salivary microRNA before downstream analysis is being studied and in part due to the minute concentration of microRNA in the saliva. The decisions to pool or not pool samples relatively depend upon several factors such as financial constraints, expertise, duration and objectives of the research.

Keywords: Salivary, oral disease, microRNA extraction, human experimentation, pooling.

Received date: 26 August 2020
Accept date: 21 November 2020

Introduction

Oral diseases such as dental caries, oral cancer, and periodontal disease are among the diseases that contribute to the increase in the percentage of oral diseases worldwide and affect general health. Oral health was defined as ‘a state of being free from mouth and facial pain, oral and throat cancer, oral infection and sores, periodontal disease, tooth decay, tooth loss, and other diseases and disorders that limit an individual’s capacity in biting, chewing, smiling, speaking, and psychosocial wellbeing’.

New and novel diagnostic procedures are beginning to appear to detect oral disease. Apart from the conventional full mouth oral examination and biopsy, one of the most advanced technique is by detecting the genetic susceptibility towards the oral disease suffered by the patient. This is because there is an increase of awareness on the relationship between oral health and the genetics components as well as the bidirectional relationship between oral diseases and medical condition has altered our perception of human diseases.

Many studies had utilized the usage of saliva as their sample to detect the patient with susceptibility to oral diseases. Saliva is a good source to identify unique biomarkers that reflect on the association between oral health and systemic health as it is rich with water, mucus, a wide range of enzymes, antibacterial components, electrolytes and mucus. With a minute amount, saliva has sufficient capability to reveal the patient’s health information and the disease status in detail. Saliva collecting procedure is non-invasive, hence it can help in reducing the patient's anxiety, discomfort and simplify the process of repeated samples collection, especially in the longitudinal monitoring study. Prevention of cross infectious diseases such as human immunodeficiency viruses (HIV), is also possible with the use of saliva as a study sample. As compared to blood,
saliva does not clot, cheaper, less hassle on the manipulation process, easily accessible and has a lesser risk to the patient. Body fluid like urine, blood, and saliva is an excellent source for genetic biomarkers like microRNA (miRNA). Description of miRNA and its regulatory function was first conducted by a group of researchers in 1993. Certain disease identification such as oral squamous carcinoma and periodontal disease has been shown to have an expression of specific miRNA biomarkers. Extracellular circulating miRNA is found to be stable and able to survive in an unfavourable physiological condition such as extreme pH, in boiling form, multiple series of freeze-thawing cycles, and extended time of storage. In contrast to miRNA, other RNA species are easily degraded after being placed in a nuclease rich extracellular environment. miRNA was found to be more stable and detectable with a good integrity number even was placed in various types of ribonucleases.

To date, there are still a limited amount of studies describing the pooling of salivary miRNA in studying the relation between miRNA and oral diseases. Thereby, the objective of this systematic review is to systematically review the utility of salivary microRNA pooling for downstream analysis and to discuss the advantages and disadvantages of pooling salivary microRNA.

Materials and methods

Preferred Reporting Items for Systematic Review (PRISMA) checklist, an established guideline in reporting a systematic review and meta-analyses is followed in the process of reporting this systematic review. Study Design

The present systematic review was reviewed and summarized all the published studies on the methods of microRNA extraction from the saliva sample in patients who were diagnosed with oral diseases.

2.1.1 Inclusion criteria

Articles that focused on the salivary microRNA extracted from human saliva were selected. All subjects chosen were diagnosed with various oral diseases such as oral squamous cell carcinoma, oral leukoplakia, and periodontal diseases. Only studies published in English or translated in the English language were reviewed. Novel or primary studies that carried out in the last 10 years to present (2009 to June 2019) with the aim of the discovery of the salivary microRNA related to oral diseases were further reviewed.

2.1.2 Exclusion criteria

The following exclusion criteria were applied. All other than human studies were excluded. Studies that studied in the expression of salivary microRNAs from non-oral diseased patients, as well as microRNA extracted from other sources other than saliva, were also chosen to be reviewed.

2.2 Information sources and search strategy

A detailed individual search of each of the bibliographic online databases was carried out: Sage, MEDLINE/Pubmed, and Science Direct. No partial grey literature search was performed. All published studies from January 2009 to June 2019 across all databases were included.

The search strategy was done using the Boolean operator where terms and free text words were combined: (salivary microRNA) AND ((oral disease OR periodontal disease OR periodontitis OR oral squamous cell carcinoma OR OSCC OR oral leukoplakia OR OLP OR oral cancer)) AND (microRNA extraction) AND (human experimental study) NOT (animal study OR animal subject OR animal subjects)).

All the duplicated search results were removed manually and by the reference manager software (EndNote, Thomson Reuters). Few selected articles from the reference list of the primary search were manually hand-screened for potentially relevant studies that could have been left out during the primary electronic database search.

2.3 Study selection

The search for studies selected was performed in two phases. In phase one, two authors (SNA and MFHH) manually screened and reviewed the titles and abstracts of all the references independently through the online database search. Articles were selected based on the titles and abstracts that met with the inclusion criteria. The disagreement between the two authors was resolved through mutual consensus. In phase two of the study selection, the two authors independently evaluated all the articles that were selected from phase one. The sensitivity and specificity reported from the articles were evaluated. An additional complimentary hand search was also conducted.
on the reference list of all the selected articles. Finally, depending on the full text of the publication, the articles selected were finalized.

2.4 Data collection process

The following information was extracted: characteristic of the study (year of publication, author, country of the study was conducted), population of the study (sample size, types of oral diseases studied), age of the subjects (range and mean), study methodology (types of samples collected, method of microRNA extraction and the downstream analysis), expression of microRNA and the main conclusion. The correspondence authors were contacted following the data extraction to retrieve the missing information.

2.5 Risk of bias in the studies

The risk of bias in the methodology of each of the selected studies was evaluated using the QUADAS-2 tool. QUADAS-2 is a revised tool for use in a systematic review of diagnostic accuracy that is widely used for the quality assessment of a study. Two authors (SNA and MFHH) were responsible for assessing the quality of each selected study and the scoring of each of the four domains in ‘low’, ‘high’ or ‘unclear’. If any disagreement occurs it was resolved by the third and fourth reviewers (KGH and FHA).

2.6 Study summary and planned methods for data analyses

The sensitivity and specificity of the method of salivary microRNA extraction and pooling in identifying the expression of microRNA in various oral diseases were considered as the main outcome. No statistical analyses are carried out.

Results

3.1 Study Selection

The literature search was conducted based on the guideline by PRISMA. In phase one, a total of 1109 papers were identified across three electronic databases (Pubmed, SAGE, and ScienceDirect). 1036 papers were removed due to duplicates, leaving 135 papers to be further screened. Entering the phase two, 122 papers were thoroughly screened through the title and the abstracts. 62 additional studies were selected from the reference list of the papers and only one was included in the next phase.

In the next phase, the full-text review was done on the 122 papers that previously selected from the screening process (phase one). These two phases have led to the final 13 out of 122 papers that were selected to be included in the descriptive analysis. The rejection reasons were due to the studies not meeting the inclusion and exclusion criteria previously explained in the methodology. Two other reasons for the papers to be rejected was due to not being a primary study and if the full paper searched was not available in the database. The retained papers for further review and descriptive analyses are 8,9, 23–25, 15–22. Figure 1 shows the detailed workflow on the process of identification, inclusion, exclusion of the studies selected.

3.2 Study characteristics

All studies selected were based on the inclusion and exclusion criteria as previously mentioned in the methodology. Five out of eleven studies appraised other than saliva to extract the microRNA from the specific oral diseases. They were a biopsy of pathological tissue, serum from the blood and cell line. The microRNA biomarker assessed went through different types of downstream analyses such as the microarray, various types of polymerase chain reaction (PCR), and next generation sequencing (NGS).

The sample size was varied, ranging from 24 to 120 patients. These patients were inclusive of oral squamous carcinoma, oral pemphigus vulgaris, oral verrucous leukoplakia, tongue squamous carcinoma, oral potentially malignant disorder with or without the dysplasia and recurrent aphthous stomatitis. MicroRNA is the only biomarker reviewed in this systematic review. A summary of the studies reviewed is depicted in Table 2.

3.3 Risk of Bias within the included studies

QUADAS 2 tool was used to systematically review and assess the risk of bias in each of the studies. The methodology of each of the studies was homogenous and had a high methodological quality. However, none of these included studies fulfilled all of the criteria in QUADAS 2 methodological quality. Domain 1 in section A, one of the criteria (was a case-control design avoided) was scored as “no” due to the design of all the studies where each study recruited patients with oral diseases and healthy control as their subjects. All of the criteria in Domain 3 were scored as “unclear” as there is no description of the reference standard in each of
the study. One study was scored as high risk due to the sample size used is too small (one periodontal disease patient and one healthy individual). Generally, the studies selected were considered as having a low risk of bias and applicability. The result of the risk and bias of each study is depicted in Table 1, while the overall judgment of the included studies is depicted in Figure 2.

3.4 Results of individual studies

There was a heterogeneity of the microRNA evaluated in each of the oral diseases from the studies selected. One type of microRNA (miR-31) is always seen to be expressed in the oral cancer lesion. Various downstream analyses were chosen before analyses of the microRNA. However, to answer the objective of this systematic review, only two studies 15,25 clearly stated the method of pooling the sample before the downstream analyses. Other studies included did not state on the pooling process before analyses but all of the studies mentioned that every step taken was based on the manufacturer’s protocol. Further investigation was done by emailing the corresponding author and from the reply, the researchers clearly stated that no pooling was done before downstream analyses. Overall, each of the studies indeed high quality and valuable for the detection and diagnosis of various oral diseases.

Discussion

This systematic review aimed to evaluate the ideal preparation method of salivary microRNAs before various downstream analysis. Several challenges such as minute amount and short nucleotide length26 were identified in the analysis of circulating salivary microRNA. Quantification of the salivary microRNA in various samples is also influenced by the high degree of similar sequence within the microRNA families and due to the existence of the isomiRs27.

The three most common platforms in profiling salivary microRNAs are the many types of polymerase chain reaction (PCR), microarray and next-generation sequencing (NGS). Each of these platforms bear different advantages and disadvantages mainly on the sensitivity and specificity of biomarker profiling, throughput capability, operating cost, library preparation and the capability in sequencing genome at the population level28-31. The decision in choosing the platform depends upon various factors such as financial, available expertise and outcome measures or objective of the study.

Based on this review, a huge difference of sample size were used in the three included studies 15,24,25. Researchers tend to opt for smaller sample size mainly due to financial constraint but this can lead to a lower power affecting the significance of the study and possibility of false positive and negative error32. The most important criterion is to ensure the proposed study had met the scientific standard33. An appropriate number of samples renders the research to be more reliable, employment of limited resource investment, while able to conform to ethical principles.

Currently, there is limited studies that incorporated the pooling of salivary microRNA. Based on this study, there were only three studies that utilized pooled microRNA sample in clinical diagnostic15,24,25. Three other searched papers claimed that there was no pooling of saliva samples was done without any further explanation 18, 20,23. However, the remaining searched papers did not respond to the email sent nor state in their studies whether pooling or the reasons for not doing so was mentioned8, 9, 16,17, 19,21,22.

One of the major considerations in pooling the samples for downstream analysis is the contamination of subjects with an altered level of expression which were not detected during data pre-processing34. Reduced overall variability due to the poorly designed probe leading to less hybridization or insufficient in cross-hybridization or due to originally small biological variability is also seen in sample pooling35. Therefore, pooling is advantageous in cases with a high level of biological variation but in small biological variability, there is less significant gain34. Another concern is the intraindividual variability that may arise from the environment, diet and other risk factors.36. The effect of sample pooling in oral diseased patients with great intraindividual variability is limited. Intraindividual variability was shown to increase the research bias and underpowered especially if smaller sample size is used37.

Samples pooling also averages out the RNA abundance. In pooled samples, the RNA abundance undergoes a nonlinear transformation and applies to the pool, while for the individual sample the transformation is in linear form and
represents the individual sample itself. Study design with a larger sample size tends to gain accuracy when pooling is done. However, due to limited biological variance, the outlier and the components of the gene-specific variance cannot be identified. Batch and lane effect are the two main sources that contribute to the confounding of effect in RNA-seq leading to variances and biases in the data. To equally distribute the effects throughout the samples, pooling the samples into the same batch and sequence in the same flow cell may be advantageous. By sample pooling, the researchers able to reduce the cost, labour and time in a large scale genotyping study, as well as aid in insufficient or limited starting material up to 10 fold of the original cost. Not only pooled design in analysis of RNA family enable for evaluation of large number of samples using relatively fewer arrays, but it also aid in making a diagnosis or prognosis when the characteristic on the individual is not the primary concern. This is particularly used in identifying certain biological biomarkers or gene expression in a population across the affected or common individuals.

A study comparing standard (without pooling) versus pooling of the RNA-seq samples was conducted recently. The standard design resulted in expression levels with a higher variance and the need for more libraries preparation, hence increasing the cost. Compared to the standard procedure, reduction in the number of replicates in sample pooling leads to expression levels with a lower variance and maximal reduction on experiment cost. The study suggested for researcher to determine the number of pool size, pool groups and the sequencing depth for the pooled design to be as effective as the standard RNA-seq experimental design.

Though presented with disadvantages, few adjustments of sample pooling can be made in the research design. Increasing the number of sample sizes in an array can increase the specificity, improve the power and ability to minimize the bias resulted from sample pooling to detect the differential expression genes. Increasing the number of replicates rather than sequencing depth also aid in reducing the pooling bias and effectively increases the power for differential gene expression. An unnecessary increase in the sequencing depth not only increase cost and limits researcher's ability to sequence more replicates, it also reduces the power in detecting the differential gene expression. For example, in small RNA sequencing, the depth of sequencing is reduced due to it's lack in complexity. Thirdly, validation of the result is essential and the method chosen is very much depends on the research goal and objective. The validation methods are explained elsewhere. Three of the included studies which pooled their sample for downstream analysis was found to validate their genes expression. These studies cost-effectively validate only the selected genes expressed from the result they obtained and reported only the significantly expressed microRNA as suitable biomarker of the oral diseases studied.

A good quality and quantity of miRNA samples yield may affect the final result regardless the samples are pooled or not. Based on the results, only two studies by Fujimori and co-workers in 2019 as well as Byun and co-workers in 2015 had clearly mention on the extraction of miRNA from the salivary exosome where the miRNA yield were expected to be more stable in form and number.

This study is confounded with some limitations. There are few studies utilizing salivary microRNA as biomarker for studying oral diseases. Various extraction kits used in the included studies may affect the quality and quantity of the miRNA extracted. This may affect the researcher's decision of the included studies to either pool or not pool their sample for downstream analysis hence leads to the heterogeneity on the data synthesis. The information on the extraction method could not be retrieved in details as clarification from the corresponding author could not be obtained. The oral diseases included in this review does not present the whole dynamic of oral diseases encountered in daily clinical practise, only certain actively studied oral diseases were included. This may provide search bias as only actively studied oral diseases were searched and included.

Conclusions

Within its limitation, the present review highlights the utility, advantages, and disadvantages of pooling the salivary biomarker samples and possible solutions for proper laboratory work while achieving the most
possible outcome. None of the methods is superior to another provided the study design is conducted based on the protocol and the research design chosen. To date, there are still limited amount of studies utilizing salivary microRNA. Further study is needed to evaluate the effect of pooling the salivary microRNA and comparison of pooling and non-pooling of salivary microRNA samples.

Acknowledgement

The authors would like to thank the Faculty of Dentistry, Universiti Teknologi MARA for the supports while performing the data analysis of this study.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Declaration of Interest

The authors report no conflict of interest.

Figure 1. The flow of literature search and criteria of selection retrieved from the electronic and hand searched.
Figure 2. Risk of bias (a) and applicability (b) concerns graph. Author's judgments concerning each domain presented as a percentage across included studies.
### Table 1.

Risk of bias and concerns of applicability summary: author’s judgment regarding each domain for each included study.

<table>
<thead>
<tr>
<th>Study</th>
<th>Flow and Timing</th>
<th>Reference Standard</th>
<th>Index Test</th>
<th>Patient Selection</th>
<th>Patient Selection</th>
<th>Index Test</th>
<th>Reference Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liu et al, 2011</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
</tr>
<tr>
<td>Mehdipour et al, 2018</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
</tr>
<tr>
<td>Hung et al, 2016</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
</tr>
<tr>
<td>Momen-Heravi et al, 2014</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
</tr>
<tr>
<td>Zhang et al, 2017</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
</tr>
<tr>
<td>Aghbari et al, 2018</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Zahran et al, 2015</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Byun et al, 2015</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Liu et al, 2010</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
</tr>
<tr>
<td>Park et al, 2009</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Duz et al, 2015</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Nisha et al, 2019</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Fujimori et al, 2019</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

*+‘ low risk, ‘-’ high risk, ‘?’ unclear risk*
<table>
<thead>
<tr>
<th>Author, Year and country</th>
<th>Sample and oral disease type</th>
<th>Sample (pool/no pool)</th>
<th>Assay</th>
<th>Main conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARK ET AL. 2009</td>
<td>Unstimulated saliva OSCC</td>
<td>Pooled</td>
<td>Real-Time PCR</td>
<td>Lower miR-125a and miR-200a in OSCC</td>
</tr>
<tr>
<td>LIU ET AL. 2010</td>
<td>Plasma and Unstimulated saliva OSCC</td>
<td>Not stated</td>
<td>Real-Time PCR</td>
<td>miR-31 associated with the OSCC tumor</td>
</tr>
<tr>
<td>LIU ET AL. 2011</td>
<td>SAS cell and Unstimulated saliva OSCC, OVL</td>
<td>Not stated</td>
<td>qRT-PCR</td>
<td>Upregulated miR-31 in OSCC. No different between OVL and control</td>
</tr>
<tr>
<td>MOMEN-HERAVI ET AL. 2014</td>
<td>Unstimulated saliva OSCC, OSCC-R, OLP</td>
<td>Not stated</td>
<td>nCounter miRNA expression assay</td>
<td>miRNA-27b as OSCC biomarker. miRNA-191 as control</td>
</tr>
<tr>
<td>ZAHRAEI ET AL. 2015</td>
<td>Unstimulated saliva PMD, OSCC, RAS</td>
<td>No pooling</td>
<td>SYBR green PCR</td>
<td>miR-21, miR-145, miR-184 showed malignant transformation</td>
</tr>
<tr>
<td>BYUN ET AL. 2015</td>
<td>Unstimulated saliva OLP</td>
<td>Pooled</td>
<td>Microarray</td>
<td>Upregulated miR-4484 in OLP patients</td>
</tr>
<tr>
<td>DUZ ET AL. 2015</td>
<td>Unstimulated saliva TSCC</td>
<td>No pooling</td>
<td>Microarray</td>
<td>miR-139-5p and miR-424 in TSCC diagnosis</td>
</tr>
<tr>
<td>HUNG ET AL. 2016</td>
<td>Tissue and Unstimulated saliva OPMD</td>
<td>Not stated</td>
<td>Tissue: miRCURY LNATM Saliva: RT-PCR</td>
<td>Upregulated miR-21 and miR-31 in OPMD. miR-31 indicates progression or recurrence</td>
</tr>
<tr>
<td>WITH ZHANG ET AL. 2017</td>
<td>Stimulated saliva Peripheral blood Oral mucosal tissue OPV</td>
<td>Not stated</td>
<td>qPCR</td>
<td>IL-6 and miR-217 in development of OPV</td>
</tr>
<tr>
<td>AGHBARI ET AL. 2018</td>
<td>Unstimulated saliva and oral mucosa OLP</td>
<td>No pooling</td>
<td>SYBR Green PCR</td>
<td>Downregulated miR-27b and miR-137 in OLP pathogenesis and prognosis</td>
</tr>
<tr>
<td>MEHDIPOUR ET AL. 2018</td>
<td>Unstimulated saliva OLP, OSCC</td>
<td>Not stated</td>
<td>qRT-PCR</td>
<td>miR-200a, miR-31, miR-21 and miR-125a involved in diagnostic and prognosis</td>
</tr>
<tr>
<td>NISHA ET AL. 2019</td>
<td>Unstimulated saliva Periodontal disease</td>
<td>Not stated</td>
<td>HiSeq 2000</td>
<td>miR-143-3p upregulated in periodontitis</td>
</tr>
</tbody>
</table>

Table 2. List of studies included and the extracted data.
References

13. Bitteti MS Di, Ferreras JA. Publish (in English ) or perish: The effect on citation rate of using languages other than English in scientific publications. R Swedish Acad Sci. 2017;46(9):121-127.