Assessment of salivary microRNA by RT-qPCR: Challenges in data interpretation for clinical diagnosis

Marc Van der Hofstadt[1], Anna Cardinal[1], Morgane Lepeltier[1], Jérémy Boulestreau[1], Alimata Ouedraogo[1], Malik Kahli[1], Laurence Molina[1], Franck Molina[1*], Thi Nhu Ngoc Van[1,2*].

[1]Sys2Diag UMR9005 CNRS/ALCEN, Cap Gamma, Parc Euromédecine, 1682 rue de la Valsière, 34184, Montpellier, CEDEX 4, France.
[2]SkillCell, Cap Gamma, Parc Euromédecine, 1682 rue de la Valsière, 34184, Montpellier, CEDEX 4, France.

*Corresponding Authors: franck.molina@sys2diag.cnrs.fr; thi-nhu-ngoc.van@sys2diag.fr

Keywords: Saliva, miRNA, Biomarkers, Diagnostic, RT-qPCR.

Abstract

Human saliva has attracted a great attention to the medical diagnostic field as a potential biofluid that provides access to various types of biomarkers in a non-invasive manner. In particular, the recent focus has been on microRNA (miRNA) due to their extracellular circulation for cell-cell communication and their dysregulation has been associated to diverse diseases. However, their analysis by the golden standard (RT-qPCR) are highly heterogenous, where miRNA signatures differ depending on the study and on the method of normalization. Using 6 miRNA targets and 10 participants through a 3 months study, here we show that a RT-qPCR kit is capable of demonstrating significant differences, yet they have inssuficient comparable rigour since the majority do not fall within the quantification region. In addition, we observed high crosstalk between the miRNA assays, altogether demonstrating that the patterns and differences observed are most likely associated to technical limitations rather than to biological contexts. As a result, we believe that more rigorous data analysis should be done prior extrapolating conclusions, which will help to homogenize results for the relay to molecular diagnostics.

Introduction

In the last few decades, there has been a desire to transition from reactive to preventive medicine. This movement has been largely motivated by the increase in treatment efficiencies, better patient well-being and the reduction in health care cost, as well as its endorsement by the shift towards precision medicine. To achieve such transition, the medical and scientific communities have especially focused on liquid biopsies (*i.e.* blood, urine and saliva) due to the presence of circulating molecules and the reduced invasiveness for sample acquisition. Results have already demonstrated the potential of these biofluids to indirectly access disease stages as well as evaluating patient responses to medical treatments[6]. Saliva has been positioned at the upfront of primary stages due to its easiness in sample collection (allowing for auto-sampling and notably practical in emergency situations such as route accidences or sport related injuries), simplified transportation and stocking conditions[1]. Its high turnover rate opens access to recent biological events, but at the cost of inducing broad heterogeneity of biochemical and physical dynamics among samples[7]. In addition, the major drawback of saliva biofluids is the abundant quantity of buccal mucosa and non-human derived (*i.e.* oral bacteria) biomaterial that masks the circulating biomarkers of interest. This makes the quantification of salivary biomarkers highly challenging, requiring exceedingly robust assays.[8]

Different classes of salivary biomarkers (DNA, RNA, proteins and metabolites) have been identified via numerous clinical studies across a broad range of diseases[9,10,11], demonstrating its potential for early detection of diseases[1]. In particular, human microRNA (miRNA) have emerged as an attractive biomarker due to its involvement in RNA silencing and post-transcriptional regulation of gene expression, hence associating their dysregulation with diverse diseases,[3] and their extracellular circulation for cell-cell communication.[12] For example, signatures of salivary miRNAs have been recently associated to the diagnosis of mid traumatic brain injury (mTBI), [11] cancers, [10] endometriosis, [13] neurodegenerative or systemic diseases.[14,15] However, due to their small size (19-25 nucleotides), their high sequence resemblance and their complex maturation process[16], miRNA characterization still remains non-trivial and of constant debate[3].

To detect such small size and variability, potential miRNA biomarkers are primarily identified by NGS analysis and subsequently validated by RT-qPCR technique. However, upon their identification, miRNA hits are finalized by a validation step using RT-qPCR technique[17,18] prior to their transition to clinics, since it remains the golden standard in clinics due to its practicality, rapidity and cost effectiveness[3]. This has favoured the apparition of a broad range of commercially available adaptations of the RT-qPCR for the detection of miRNAs.[19,20,21,22] In general, these techniques mainly focus on the conjugation of a synthetic oligomer to the miRNA target, which enables the downstream amplification by qPCR at the cost of introducing partial crosstalk due to the shared synthetic oligomer. Moreover, data normalization in miRNA studies remains unclear,[4,5] where miRNA signatures differ depending on the study[23] and on the method of normalization.[24] Altogether, this makes the exploitation of salivary miRNA

biomarkers for clinical diagnostics highly requested, and hence requiring highly robust yet specific quantification assays.[8]

In this study, we aim to decipher the potential use of commercially available RT-qPCR kit relying on poly(A) elongation[25] to detect salivary miRNA for clinical diagnosis. We have chosen a panel of six miRNAs, namely hsa-Let-7a-5p, hsa-Let-7f-5p, has-mir-148a-3p, has-miR-26b-5p, hsa-miR-103a-3p and hsa-miR-107, and analysed their presence in small RNA samples extracted from the saliva of ten healthy volunteers across a three months study. Firstly, we observed chemophysical variations on the saliva samples and notable variations on the concentrations of extracted small RNA, which even varied within the same participant across the study. Secondly, we noticed that although Ct values were within the acceptable detection range considered by the community (ranging from 6 to 35)[18,26], few fell within the quantification region of their respective miRNA assay, reducing their quantitative credibility. In addition, we observed high crosstalk between the six miRNA assays used, convoluting the interpretation of the detected Ct values. Finally, by spiking synthetic hsa-Let-7a-5p on the samples, we could distinguish those endogenous miRNA concentrations that were present at limit of quantification from those that were at significantly lower concentrations, even if values were outside the true quantification region of the miRNA assay.

Results

Salivary small RNA variations among samplings and participants

Prior to small RNA extraction, the chemophysical aspect of the saliva samples from the different participants (P) was evaluated (Figure 1A & 1B). Results revealed that chemophysical aspects were constant throughout the study with several of them classifiable into 3 major groups: i) transparent and liquid (P5, P6), ii) translucent and viscous (P1, P4, P7), iii) opaque and viscous (P2, P10). Small RNA extraction from saliva samples revealed variations in RNA concentrations among participants and among different time points for a given participant (Figure 1C). However, no significant difference was observed between time samplings (Figure S1). Compared to the other participants, we noted that participants belonging to group i (P5 and P6) had lower variation across this study and had significantly lower total RNA concentrations (Figure S2). Moreover, to assess if the small RNA extraction efficiency was identical among participants, an equal amount of miRNA UniSP6 was spiked-in during the extraction. Consistent RT-qPCR results (Ct = 21.7 ± 0.07) revealed a comparable extraction efficiency for all participant (Figure 1D), and that it was conserved throughout this study (Figure S3). In particular, upon standardizing total RNA concentration with respect to 1 µL condition, UniSP6 detection behaved linearly (Figure S4).

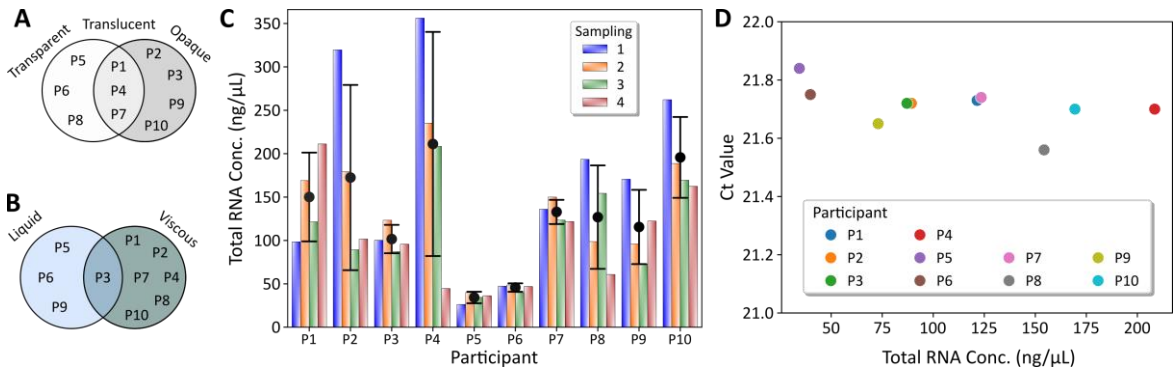

Figure 1: Chemophysical aspect and small RNA extraction of saliva samples from 10 healthy participants. (**A**) Transparency and (**B**) viscosity of the saliva samples prior RNA extraction. (**C**) Total small RNAs extracted from 250 µL of saliva for each participants and for the 4 sampling points. Error bars show one standard deviation. (**D**) Detection of spiked artificial UniSP6 miRNA prior to RNA extraction for sampling three. 1 µL of each extracted RNA sample was used as input. P1-10: Saliva sample from participant 1-10.

Quantification of six miRNAs from saliva samples by RT-qPCR

Six miRNA molecules were chosen as our targeted panel for our study (Table S1), since they have been previously identified in saliva by different authors[27,17]. To assess the suitable working small RNA concentration, we validated the linear response of each chosen miRNA in a ranging concentration from 1 ng up to 300 ng of input salivary small RNA. In particular, we chose a representative participant from each group previously defined, namely P2, P6 and P7. Results demonstrated linear behaviour for all six tested miRNAs within all three participants (Figure S5). However, while participants P2 and P7 showed high correlations coefficients (R^2>0.99, except for hsa-Let-7f-5p of the P7), participant P6 had lower coefficients, ranging from 0.94 to 0.99. In particular, at high concentrations of small RNA (300ng), Ct values are lower than predicted by the linear fit. Moreover, qPCR efficiencies for the six tested miRNAs followed a similar pattern, with efficiencies ranging from 85.6% to 100.7% for participants P2 and P7, while efficiencies for P6 drop down 73.9% to 90.7% (Figure S6A). Although a difference was observed between P6 and P7 (Figures S6B), no significant differences were observed among efficiencies of the six miRNA assays (Figures S6C).

We identified 50 ng of input salivary small RNA as a potential concentration to detect endogenous miRNA. To verify this assessment, we firstly tested that the dilution effect of the sample to achieve the 50 ng concentration had no effect on the linear response, especially on the participants that didn't fall within the three chemophysical groups, namely P3, P8 and P9. To do so, we compared between a fixed input volume (1 µL) and the respective dilution factor to achieve 50 ng. Results revealed that Ct values for the six miRNA and participants tested behaved linearly according to the required dilution factor (Figure S7), indicating that within this working dilution range the efficiency of these assays are not significantly affected. Secondly, we validated that the addition of spiked miRNA during the extraction process had

no influence on the detection of the six target miRNAs. To this aim, we performed the RT-qPCR analysis using 50ng of extracted small RNA from two sets of the three samples (P2, P6 and P7) in which the extraction was carried out with and without the spiking of the artificial miRNA UniSP6 (Figure S8A). We observed that Ct values increased by only 0.2 ± 0.09, which represented an increase of less than 1% in CT value, and spike-in had no differential effect between the different miRNA assessed (Figure S8C).

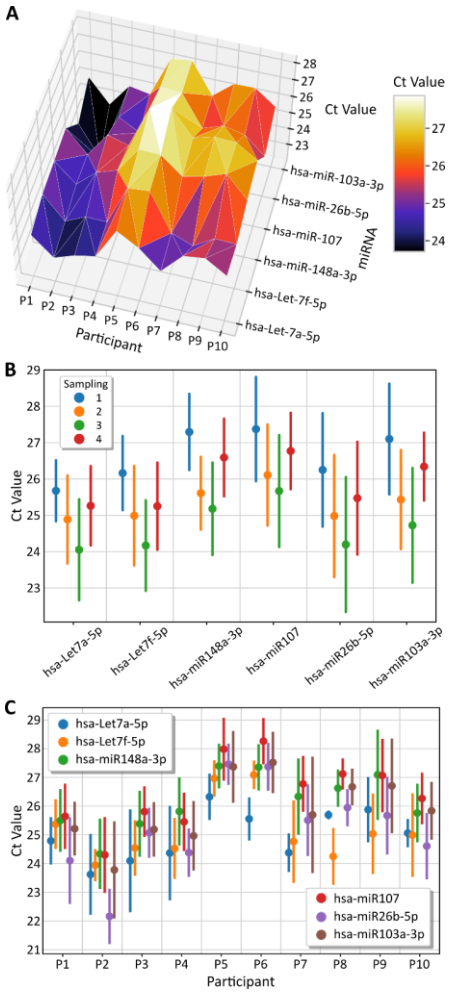

Figure 2: RT-qPCR shows high variability of salivary miRNA signatures within 10 healthy participants. (A) Average Ct values of the four sampling points for 10 participants and for the six assessed miRNA assays. Signatures of the six analyzed miRNAs across (B) time and (C) between different participants. All analysis was perform with 50 ng of total extracted small RNAs. Error bars show one standard deviation.

Figure 2A shows the heterogeneous signatures obtained when assessing the six identified miRNA targets by RT-qPCR on 4 time samplings for 10 different participants when using 50 ng of extracted small RNA as input. Averaged Ct values ranged from 22 to 29 (±1.5), which is within the acceptable detection range considered by the community[18,26] A deeper analysis of the results revealed that the six miRNA targets shared a similar response pattern among the four sampling points (Figure 2B), with a reduction in Ct value from sampling 1 to sampling 3, followed by a restoration of the Ct value for sampling 4 that fell between the first two samplings. Contrary to the total extracted small RNA concentration previously shown, Ct values revealed a significant difference between sampling 1 and 3 for all miRNAs, and between sampling 3 and 4 for half of the miRNA assessed, namely hsa-Let-7a-5p, hsa-miR148a-3p and hsa-miR103a-3p (Figure S9A). Moreover, when evaluating endogenous miRNA presence, statistical analysis revealed the apparition of two major groups, a lower Ct value group containing hsa-Let-7a-5p and hsa-Let-7f-5p, and a greater Ct value containing hsa-miR148a-3p and hsa-miR107 (Figure S9B). In between both groups, we found hsa-miR26b-5p and hsa-miR103-3p, which shared proximity with the lowest and the greatest Ct value, respectively. However, upon evaluation of the Ct value with respect to individual participants (Figure 2C), we observed that the average miRNA profile (Figure 2B) partially masked individual miRNA profiles. Furthermore, statistical analysis revealed significant differences between participants, dividing them into

three major groups: i) high Ct values (P5 and P6), ii) moderate Ct values (P8 and P9) and iii) low Ct values (P1, P2, P3 and P4), with participant P7 and participant P10 between the later groups, being adjacent to group ii and iii respectively (Figure S10). We noted that these groups were in close relation to the dilution factor required to achieve 50 ng of extracted small RNA as input (Table S2). In particular, the first group presented small dilution factors (7x and 9x), the second group intermediate (25x and 23x, followed by P7 with 26.5x) and the third group highest (ranging from 30x to 42x, with the exception of the P3).

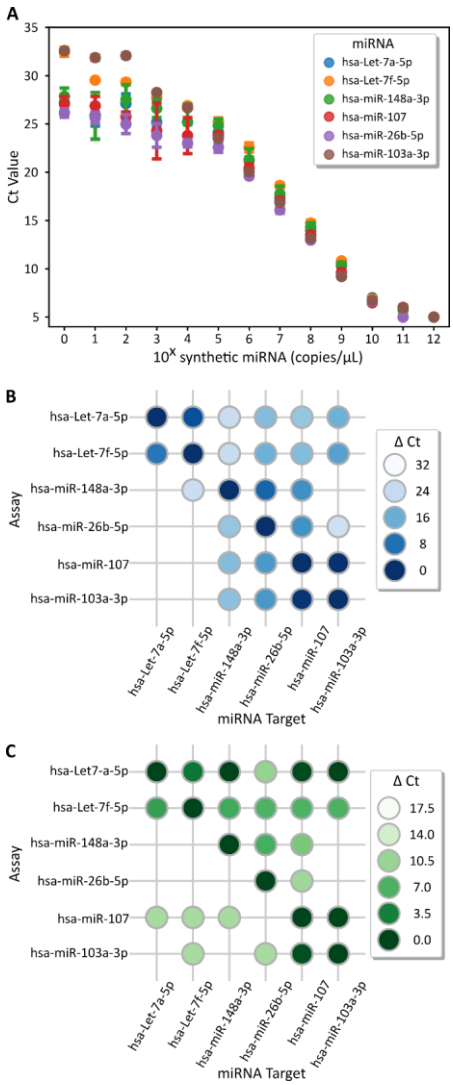

Figure 3: Specificity and sensitivity limitations of the six miRNA assays. (**A**) Serial dilution from 1 to 10^{12} copies/µL with synthetic miRNA target for each miRNA assay to determine their limit of detection (LOD) and their limit of quantification (LOQ). Cross reactions among miRNA assays and synthetic targets at (**B**) 10^9 copies/µL and (**C**) 10^5 copies/µL. ΔCt values are calculate from Figure S13. Empty spaces represent no crosstalk at all (Ct Value >40). Data obtained from a duplicate experiment.

Sensibility, specificity and fidelity of the RT-qPCR assay

Given the significant variation in endogenous miRNA quantification, we decided to verify if this variability was associated to biological factors or due to technical artefacts. To this aim, we investigated the sensitivity and specificity of the chosen miRNA assays using synthetic miRNAs. To determine the sensitivity limitation for each miRNAs assay, we performed serial dilutions ranging from 1 to 10^{12} copies/µL (Figure 3A). Results demonstrated high sensitivities, allowing to detect down to 1 copies/µL and with negative controls (Ct > 35). However, we noted that under a given concentration of copies/µL the behaviour was no longer linear, and hence we defined this concentration as the limit of quantification (LOQ). The LOQ concentrations for hsa-Let-7a-5p, hsa-miR-148a-3p, hsa-miR26b-5p and hsa-miR-107 assays where higher (10^5 copies/µL) compared to hsa-Let-7f-5p and hsa-miR-103a-3p assays, which decreased down to 10^4 and 10^2 copies/µL, respectively. Calculation of the RT-qPCR efficiencies between the LOQ and the Ct saturation point (*i.e.* Ct value = 5) revealed high efficiencies for all miRNA assays, ranging from 98.5 and 112.5% (Figure S11 & S12), which fall within the acceptable limits for PCR exponential amplification.

Knowing the quantification regimes for the six miRNA assays, we observed that out of the 240 samples measured a total of 42% were outside the

quantification zone (*i.e.* Ct values higher than at the LOQ), which was the case for the 40 samples for the hsa-miR-107 assay. Subsequently, we also noticed that 29% were obtained at the respective LOQ of each miRNA assay, leaving assays such as hsa-Let-7a-5p and hsa-miR-148a-3p fluctuating either outside or at the LOQ. Yet, 29% of the obtained Ct values were within the linear quantification zone, where the hsa-miR103 assay was the sole to be completely done in the quantification zone (due to the lower LOQ). Results indicated that endogenous hsa-miR103 was present at concentrations between $3.8x10^3$ and $5.4x10^4$ copies/µL, depending on participant.

Due to the short size, and hence high sequence resemblance[2], miRNA detection is prone to crosstalk. To undercover the technical limitations regarding specificity in this study, we compared the detection of each miRNA assay and its corresponding miRNA target (ON target) with the other five miRNAs used in this study (OFF target). Firstly, we assessed the crosstalk contribution at 10^9 copies/µL, a concentration which falls within the middle of the quantification zone (Figure 3B & S13A). Out of the 30 possible crosstalk combinations, only 8 presented no crosstalk at all (Ct Value >40), 4 potential crosstalk (Ct Value = 35) and the rest (60%) presented non-negligible crosstalk. In particular, we noticed that hsa-miR-107 and hsa-miR-103a-3p had a high crosstalk between them, ΔCt of 0.21 and 0.84 respectively, making their discrimination very unlikely. Similarly, hsa-Let7a-5p and hsa-Let7f-5p also presented high crosstalk, ΔCt of 3.9 and 8.7, while the rest of combinations ranged between ΔCt from 6.5 up to 25.4. Since in this study 71% of the Ct values obtained were on the LOQ or lower, we secondly assessed crosstalk contribution at 10^5 copies/µL, which corresponds to the LOQ of 4 of the miRNA assays (Figure 3C & S13B). At this concentration, non-negligible crosstalk reduced down to 40%. However, in the cases where crosstalk had not disappear, decreasing the concentration down to 10^5 copies/µL significantly increased the crosstalk, ranging from ΔΔCt of 7.1 up to 19.0. Interestingly, we observed that crosstalk between hsa-miR-107 and hsa-miR-103a-3p, and hsa-Let7a-5p and hsa-Let7f-5p didn't notably change, ΔΔCt of 0.1 and 1.9, respectively, when decreasing the miRNA concentration.

To verify that for assays hsa-Let7a-5p and hsa-Let7f-5p crosstalk is conserved independently of miRNA concentration, we examined their respective crosstalk throughout the serial dilution of miRNA ranging from 1 to 10^{12} copies/µL (Figure 4). As expected, crosstalk remained generally stable, with an average ΔCt value of 3.78 ±1.02 for hsa-Let7a-5p and 10.29 ± 1.20 for hsa-Let7f-5p. However, we noted two tendencies: firstly, that as the concentration of OFF target decreased the error value increased, and secondly, a transition at the LOQ, where lower concentrations presented erratic responses. Re-calculating average ΔCt within the limit of quantification zone didn't considerably change, being 3.93 ± 1.01 and 10.02 ± 1.1 for hsa-Let7a-5p and for hsa-Let7f-5p, respectively. Knowing the OFF target effect of both assays, we decided to validate the limitation of detection in the presence of OFF target. To do so, we re-assessed the serial dilution of the ON target but in the presence of 10^7 copies/µL of OFF target (Figure 4).

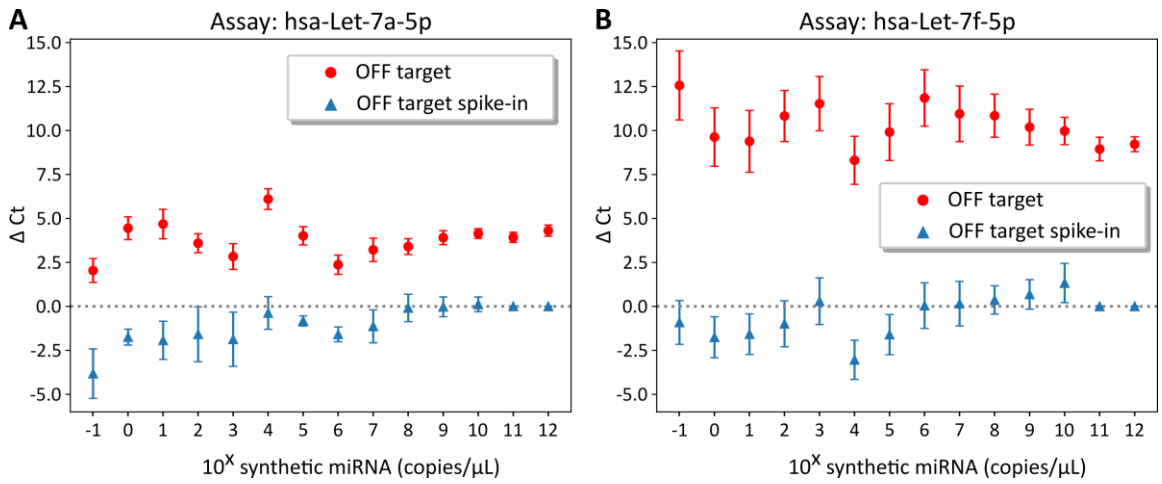

Figure 4: Conserved crosstalk for hsa-Let-7a-5p and hsa-Let-7f-5p in the quantification regime. ΔCt values obtained when comparing in a serial dilution from 1 to 10^{12} copies/µL the corresponding synthetic miRNA (ON target) with its cross reaction (OFF target, red circles) for (**A**) hsa-Let-7a-5p assay and (**B**) hsa-Let-7f-5p assay. The ΔCt by the addition of 10^7 copies/µL of OFF target on the serial dilution of ON target was also calculated (blue triangles). Error bars show one standard deviation from a duplicate experiment. Red circles = $Ct^{OFF\ Target} - Ct^{ON\ Target}$, blue triangles = $Ct^{ON\ Target+\ OFF\ Target\ at\ 10^7\ copies/\mu L} - Ct^{ON\ Target}$.

Since for the hsa-Let7a-5p assay we obtained a ΔCt value of 3.78 (~ 1 log) for the OFF target, we expected greatest crosstalk contribution by the 10^7 copies/µL of the OFF target at 10^6 copies/µL of the ON target (1 log difference), and partial crosstalk at its adjacent concentrations, which was what we observed. Note that due to erratic responses under the LOQ, interpretation at concentrations under LOQ should be handled with care. Similarly, since for hsa-Let7f-5p OFF target we obtained a ΔCt value of 10.3 (~ 3 logs), we observed greatest crosstalk contribution at 10^4 copies/µL, showing a greater resilience of hsa-Let7f-5p to crosstalk compared to hsa-Let7a-5p.

The use of spiked synthetic miRNA for the potential detection of endogenous miRNA

With the hypothetical assumption that the endogenous miRNA present on the extracted salivary samples are slightly under or at the LOQ (given the Ct values obtained for hsa-Let7a-5p analysis), we decided to spike-in synthetic hsa-Let7a-5p to surpass the LOQ value. To this end, we chose 4 spiking concentrations; two which were significantly under the LOQ (1 copy and 10^3 copies), one which was at the LOQ (10^6 copies) and one 10-fold higher (10^7 copies). As expected, the higher the concentration of spiked-in synthetic hsa-Let7a-5p in the endogenous sample, the higher the ΔCt value with respect to the endogenous sample (Figure 5a & S14). However, two distinctive groups appeared; a group that only showed responses upon the addition of 10^6 copies (P1 and P2) and a second group that showed responses directly with the addition of only 1 copy (P4, P5, P6, P8 and P9). To understand these shifts, we

calculated the ΔCt value between the spiked-in salivary samples and the corresponding synthetic serial dilution (Figure 5B). Results showed that only for participants P1 and P2 the addition of 10^7 copies had a negative effect on the detection of hsa-Let7a-5p compared to the synthetic system, while for the remaining participants 10^6 copies was sufficient.

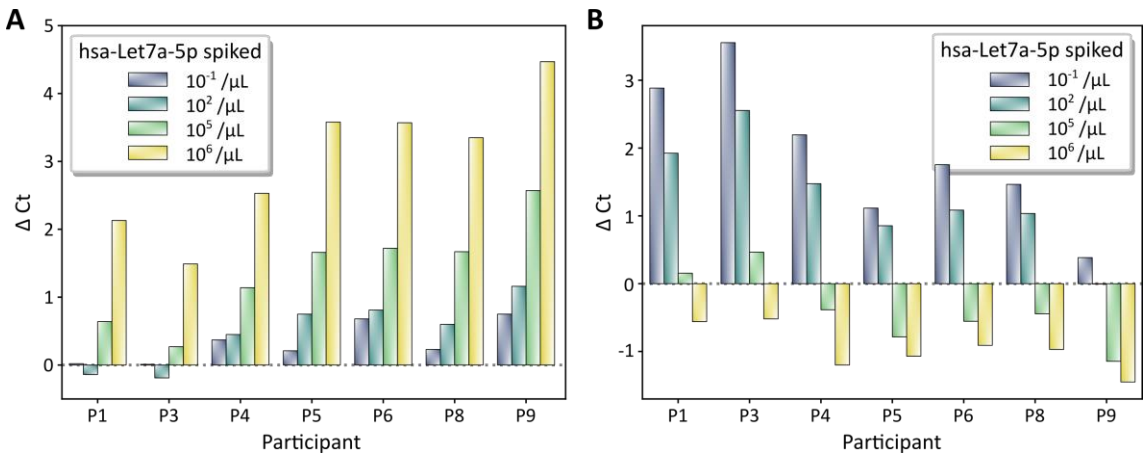

Figure 5: The spiking of synthetic hsa-Let-7a-5p miRNA allows semi-quantification at LOQ concentrations. (**A**) ΔCt values of spike-in synthetic hsa-Let-7a-5p miRNA in 50 ng RNA extract with respect to the non-spike saliva sample in a range of spike-in concentrations. ΔCt were calculated as $Ct^{Endogenous}$ - $Ct^{Endogenous + synthetic}$. (**B**) ΔCt values of spike-in synthetic hsa-Let-7a-5p miRNA in 50 ng RNA extract with respect to same synthetic miRNA concentration in the absence of sample. ΔCt were calculated as $Ct^{synthetic}$ - $Ct^{endogenous + synthetic}$. Data determined from Figure S14.

Discussion

Salivary miRNA biomarkers have become recently relevant to the medical diagnostic community due to the easiness of salivary collection and the deregulation of miRNA in diverse diseases[9,11,15],. We investigated the potential to detect salivary miRNA using a commercially available kit relying in the conventional poly(A) elongation of the miRNA targets prior to RT reaction. We observed significant variability between miRNA targets and participants, albeit these variations cannot be fully attributed to biological fluctuations, reinforcing the need for deeper data judgement prior extrapolating conclusions.

Chemophysical aspects of salivary samples, namely transparency and viscosity, revealed 3 major groups. However, we do not disregard the possibility that with sufficient participants all the possible groups may be represented. Nevertheless, we noted that the group which was transparent and liquid had significantly lower variation and total RNA concentration throughout this study (Figure 1C & S2). Since small RNA extraction efficiency was constant (Figure 1D), we discarded the possibility that the chemophysical properties of the sample affected extraction efficiency. Interestingly, serial dilutions of participant P6 also presented slightly lower correlations coefficients (Figure S5) and partially lower efficiencies (Figure S6A) for RT-qPCR. However, we observed that at high concentrations of small RNA (300ng) the Ct values were lower than predicted by the linear fit. We hypothesize that since sample P6

requires higher volumes (3-fold greater compared to participant P7), this may affect the efficiency of the RT-qPCR reaction at high small RNA concentrations (*e.g.* remaining purification reagents). Furthermore, since we observed that correlations coefficients were not related to efficiencies, and only 3 out of the 6 miRNA presented considerable low efficiencies (<75%), we hypothesized that there could be a selective impairment for these 3 miRNAs assays and hence should be taken into consideration when analysed (since the exponential coefficient reduces by ~20%).

Having chosen 50 ng of input salivary small RNA and verified that the dilution factor from the extracted small RNA didn't affect the RT-qPCR outcome (Figure S7), we measured the 6 targeted miRNA across time and for 10 participants. We firstly observed that, while no extracted small RNA concentration pattern was observed across the study (Figure S1), upon normalization of the concentration to 50 ng a significant fluctuation of miRNA detection across time was obtained (Figure 2B). This demonstrates that a direct relation between extracted small RNA concentration of the population across time and miRNA presence cannot be done. Furthermore, since the trend is considerably marked for all the participants, we hypothesize that this observed shift may be attributed to an environmental effect that has homogenously affected all participants throughout the three months of the study, implying the importance of comparing samples collected within the same time period. Secondly, we determined that the six chosen miRNA targets significantly divided into two major groups, with hsa-Let-7a-5p and hsa-Let-7f-5p having lowest Ct values and hsa-miR148a-3p, and hsa-miR107 having the greatest Ct values. However, we noticed that this average miRNA profile partially masked the individual miRNA profile of each participant (Figure 2C). Thirdly, we detected significant differences between participants, subdividing them into 3 groups. We noticed that these three groups were in relation to the dilution factor (Table S2), and hence to the average extracted RNA concentration of the participant. Interestingly, while the participants within two groups had no strong previous relations, we observed again participant P5 and P6 stand out as an individual group (Figure S10). Given that we had previously observed lower efficiencies for hsa-miR-148a-3p, hsa-miR-26b-5p and hsa-miR-103a-3p for participant P6 (Figure S4), we checked for abnormal miRNA patterns. However, the miRNA patterns of P5 and P6 were very similar to the rest of the participants, diminishing the importance of the lower RT-qPCR efficiencies for the mentioned miRNA assays.

The concentration of circulating miRNA is very low, such as for example 10^6 copies/µL in blood plasma. For this reason, we had to make sure that the used RT-qPCR kit was capable of quantitatively measure concentrations even lower. Indeed, results revealed that LOQ for the six miRNA assays were at 10^5 copies/µL or lower (Figure 3A). However, comparison of LOQ values to Ct values obtained from salivary samples revealed that only 29% of the obtained results were within the quantitative region, highlighting the technical limitations of the assays and reducing the quantitative credibility of the Ct values obtained due to insufficient sensitivity. In particular, we could only infer quantitative information from hsa-miR-103a-3p assay, which demonstrated 1.4-fold difference between highest and lowest endogenous

concentration. In addition to sensitivity issues, RT-qPCR also suffer from specificity, especially when amplifying miRNA due to their short size. Our results clearly demonstrated this phenomenon, where the cross-assessment between the six miRNA assays revealed 60% of non-negligible crosstalk at high miRNA concentrations (Figure 3B), which reduced down to 40% but at the cost of significantly increasing crosstalk of the remaining. As expected, miRNA assays that differed by only one nucleotide, namely hsa-Let-7a-5p with hsa-Let-7f-5p and hsa-miR107 with hsa-miR-103a-3p, are almost incapable of discriminating between miRNAs (Figure 3C). Nevertheless, crosstalk can be also advantageously used to increase precision. For example, although the LOQ for hsa-miR107 is 10^5 copies/µL, due to the high crosstalk with hsa-miR-103a-3p, we can discard the hypothesis that the presence of endogenous hsa-miR107 is above the endogenous hsa-miR-103a-3p concentration detected (as low as $3.8x10^3$), decreasing by at least 1 order of magnitude the LOQ of hsa-miR107. Moreover, we observed that for the hsa-Let-7a-5p and hsa-Let-7f-5p pair the crosstalk remained constant thought the serial dilution of the OFF target from 1 to 10^{12} copies/µL (Figure 4). However, as the concentration of OFF target decreased, error increased, until the LOQ where erratic responses were obtained. We attribute this behaviour to the stochastic (*i.e.* non-specific) nature of the enzymes[28], where at miRNA concentrations sufficiently low the stochastic behaviour predominates over the deterministic behaviour, decreasing the robustness and linearity of amplification techniques[29],[30]. This reinforces the importance of assuring obtained Ct values are within the limits of quantification zone.

Since quantification of endogenous miRNA cannot be done with Ct values outside the boundaries of quantification zone (*i.e.* Ct values equal and greater than LOQ), we decided to push the limits to obtain semi-quantitative information by spiking synthetic hsa-Let7a-5p in the endogenous samples. The intention of this methodology is to clearly distinguish samples that are on the LOQ from those that are significantly lower. The rationality behind this statement is that the closer the endogenous concentration is to the added synthetic miRNA, the lower the impact by the addition of synthetic miRNA to the sample (*i.e* lower ΔCt). the lower the impact the addition of synthetic miRNA will have on the Ct value. For example, the presence of 10^4 endogenous copies/µL will not affect the Ct value when adding 10^6 synthetic copies/µL (ratio 1:100), although the presence of 10^5 endogenous copies/µL (ratio 1:10) will have a significant effect, and even greater if there was 10^6 endogenous copies/µL (ratio 1:1). With this rationality, we hypothesize that participant P1 and P3 have a concentration ~10^5 copies/µL (although there is 0.68 ΔCt between them) while the other participants (P4, P5, P6, P8, P9) have concentrations < 10^5 copies/µL, since the addition of 10^6 synthetic copies/µL has low and high effect on the Ct values, respectively (Figure 5B). We do not adventure in predicting the concentration of the latter group (or their order) since the addition of 10^5 synthetic copies/µL is at the LOQ and hence is highly prone to error. Taken together, these results demonstrate the possibility of using commercial RT-qPCR to detect endogenous salivary miRNA but must be supported by technical information (mainly LOQ and crosstalk) to corroborate the robustness of the results and associate variations to biological contexts.

Conclusions:

In this work we have demonstrated that a commercially available RT-qPCR kit shows significant differences between six salivary miRNA targets and between healthy participants. However, we demonstrate that the majority of these differences are associated to technical limitations rather than to biological contexts. As a result, we believe that our findings will contribute to more rigorous data analysis before extrapolating conclusions from RT-qPCR results. In particular, since it can be extrapolated to other miRNA biosensors, we believe that it will help to homogenize results for the relay to molecular diagnostics.

Materials and Methods

Study approval and saliva collection

For this study a total of 10 healthy males with ages ranging from 18 to 40 were recruited through the Sys2diag laboratory, UMR 9005, the Centre National pour la Recherche Scientifique (CNRS), based in Montpellier, France according to the CPP with resisted number of 2023-A00188-37. All subjects were informed, signed and consented in accordance with the CPP prior to the recruitment. Saliva collection and analyses were performed with approved protocols. Participants were asked to refrain from eating, drinking or smoking for at least 30 min prior to saliva collection, and performing all at the same time of the day. No drop out associated with sickness or oral infection was observed during the study. Approximately 2 mL of unstimulated saliva was collected from each participant and stored at 4°C for maximum 2H before small RNA extraction.

Salivary miRNA extraction and quantification

Prior small RNA extraction, all salivary samples were homogenized, and when needed, artificial miRNA UniSP6 (5.6 x 10^8 copies) was added as a technical control of the extraction process. All salivary miRNA were extracted using miRNeasy Serum/Plasma Kit (Qiagen) according to the instruction of the manufacturer except for the elution step in which extracted salivary small RNAs were collected in 20 µL nuclease free water. In brief, 250 µL of whole saliva from each participant were homogenized in 1 mL of Qiazol solution followed by chloroform purification, and RNA precipitation by isopropanol. Precipitated RNA was then loaded to miRNeasy column, where only small RNAs will be retained after multiple washes.

Total concentration was quantified by Nanodrop One (Thermo Scientific, Wilmington USA), Qubit™ microRNA Assay Kits (ThermoFisher Scientific) and Labchip small RNA assay (PerkinElmer) according to the manufacturer's instructions.

Synthetic miRNA Targets

Six miRNAs, including hsa-Let7a-5p, hsa-Let7f-5p, hsa-miR-148a-3p, hsa-miR-26b-5p, hsa-miR-107 and hsa-miR-103a-3p, were selected according to their sequence similarity. All the six

synthetic miRNAs were purchased from Integrated DNA technologies (IDT, Europe). All sequences and annotations are available in the supplementary data section.

Reverse transcriptase reaction:

Reverse transcriptase reactions using miRCURY LNA RT Kit (Qiagen) for synthetic and/or purified salivary miRNA were prepared in 96 well plates. Each 10 µL reaction contained 2µL of 2X reaction buffer; 1 µL 10X reverse transcriptase and 1 µL of synthetic or purified salivary miRNA at desired concentrations. The reaction was incubated in a peqSTAR 96X thermocycler (Ozyme, Montigny-le-Bretonneux France) at 42°C for 60 min followed by a denaturation step at 85°C for 5 min and stocked at -20°C until used. Negative controls (no-template and no reverse transcriptase) were included in each run. All samples and control conditions were run in duplicate.

For the analysis using 50ng of extracted small RNAs as input, all extracted samples were normalized to 50 ng/µL and then 1µL was added to the RT-qPCR reaction, except for those whose concentration were inferior to 50 ng/µL (P5 and P6 samples) where higher volumes were required.

Q-PCR quantification and analysis:

Q-PCR reactions using miRCURY LNA SYBR Green PCR Kit and miRCURY LNA miRNA PCR Assay (Qiagen) were prepared in 384 multi-well plates. Each 10 µl reaction volume contained 5µL of qPCR master mix; 1 µL of miRCURY LNA miRNA PCR Assay and 3 µL of 10X diluted RT product. Negative controls (no-template) were included for each assay. All sampled and control conditions were run in duplicate. Real-time Q-PCR thermal cycling reactions were performed by the LightCycler 480 (Roche, Meylan France) directed by the LightCycler 480 Software (version 1.5.1.62). Thermal cycling conditions were: Pre-Incubation for 2 minutes at 95°C, 40 cycles of amplification—(95°C for 10s, 56°C for 60s). For melting curve analysis, we set up 1 cycle at 95°C for 5s, 60°C for 1 minute, 97°C in continuous mode and 1 cooling cycle at 40°C for 10s. Ct values and Melting curves were analysed by Abs Quant/2nd Derivative Max of the same software.

Statistical Analysis

Scipy version 1.11.2 was used on python 3.10.4 to do statistical analysis. All statistical test were done using Mann-Whitney U (scipy.stats.mannwhitneyu) with asymptotic method (*i.e.* p-value calculated by comparing to normal distribution and hence correcting for ties).

Aknowledgements

The authors would like to thank study participants.

Author contributions

Study concept and design: M.VDH., F.M. and TNN.V. Acquisition, analysis, interpretation of data: All authors. Drafting of the manuscript and critical revision of the manuscript for

important intellectual content: M.VDH., F.M. and TNN.V. Statistical analysis: M.VDH. Funding acquisition: F.M. All authors critically reviewed and edited the manuscript.

Data availability statement

The authors declare that all relevant data has been provided within the manuscript and its supporting information files.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

1. Lee, Y.-H. & Wong, D. T. Saliva: an emerging biofluid for early detection of diseases. *Am J Dent* **22**, 241–248 (2009).
2. Nemeth, K., Bayraktar, R., Ferracin, M. & Calin, G. A. Non-coding RNAs in disease: from mechanisms to therapeutics. *Nat Rev Genet* (2023) doi:10.1038/s41576-023-00662-1.
3. Saliminejad, K., Khorram Khorshid, H. R., Soleymani Fard, S. & Ghaffari, S. H. An overview of microRNAs: Biology, functions, therapeutics, and analysis methods. *Journal Cellular Physiology* **234**, 5451–5465 (2019).
4. Nuzziello, N. & Liguori, M. Seeking a standardized normalization method for the quantification of microRNA expression. *Muscle and Nerve* **59**, (2019).
5. Donati, S., Ciuffi, S. & Brandi, M. L. Human Circulating miRNAs Real-time qRT-PCR-based Analysis: An Overview of Endogenous Reference Genes Used for Data Normalization. *IJMS* **20**, 4353 (2019).
6. Qiu, J. *et al.* Refining Cancer Management Using Integrated Liquid Biopsy. *Theranostics* **10**, 2374–2384 (2020).
7. Nonaka, T. & Wong, D. T. W. Saliva Diagnostics. *Annual Rev. Anal. Chem.* **15**, 107–121 (2022).
8. Ostheim, P. *et al.* Overcoming challenges in human saliva gene expression measurements. *Sci Rep* **10**, 11147 (2020).

9. Yoshizawa, J. M. *et al.* Salivary Biomarkers: Toward Future Clinical and Diagnostic Utilities. *Clin Microbiol Rev* **26**, 781–791 (2013).

10. Robotti, M., Scebba, F. & Angeloni, D. Circulating Biomarkers for Cancer Detection: Could Salivary microRNAs Be an Opportunity for Ovarian Cancer Diagnostics? *Biomedicines* **11**, 652 (2023).

11. Mavroudis, I. *et al.* Review on the Role of Salivary Biomarkers in the Diagnosis of Mild Traumatic Brain Injury and Post-Concussion Syndrome. *Diagnostics* **13**, 1367 (2023).

12. Bayraktar, R., Van Roosbroeck, K. & Calin, G. A. Cell-to-cell communication: microRNAs as hormones. *Molecular Oncology* **11**, 1673–1686 (2017).

13. Bendifallah, S. *et al.* Salivary MicroRNA Signature for Diagnosis of Endometriosis. *J Clin Med* **11**, 612 (2022).

14. Hindson, J. Salivary miRNA signature for ESCC. *Nat Rev Gastroenterol Hepatol* (2023) doi:10.1038/s41575-023-00837-5.

15. Setti *et al.* Salivary MicroRNA for Diagnosis of Cancer and Systemic Diseases: A Systematic Review. *IJMS* **21**, 907 (2020).

16. Shang, R., Lee, S., Senavirathne, G. & Lai, E. C. microRNAs in action: biogenesis, function and regulation. *Nat Rev Genet* **24**, 816–833 (2023).

17. Di Pietro, V. *et al.* Unique diagnostic signatures of concussion in the saliva of male athletes: the Study of Concussion in Rugby Union through MicroRNAs (SCRUM). *Br J Sports Med* **55**, 1395–1404 (2021).

18. Romani, C. *et al.* Genome-wide study of salivary miRNAs identifies miR-423-5p as promising diagnostic and prognostic biomarker in oral squamous cell carcinoma. *Theranostics* **11**, 2987–2999 (2021).

19. Chen, C. Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Research* **33**, e179–e179 (2005).

20. Androvic, P., Valihrach, L., Elling, J., Sjoback, R. & Kubista, M. Two-tailed RT-qPCR: a novel method for highly accurate miRNA quantification. *Nucleic Acids Research* **45**, e144–e144 (2017).

21. Honda, S. & Kirino, Y. Dumbbell-PCR: a method to quantify specific small RNA variants with a single nucleotide resolution at terminal sequences. *Nucleic Acids Res* **43**, e77–e77 (2015).

22. Gines, G., Menezes, R., Xiao, W., Rondelez, Y. & Taly, V. Emerging isothermal amplification technologies for microRNA biosensing: Applications to liquid biopsies. *Molecular Aspects of Medicine* **72**, 100832 (2020).

23. Ghaith, H. S. *et al.* A Literature Review of Traumatic Brain Injury Biomarkers. *Mol Neurobiol* **59**, 4141–4158 (2022).

24. Faraldi, M. *et al.* Normalization strategies differently affect circulating miRNA profile associated with the training status. *Sci Rep* **9**, 1584 (2019).

25. Shi, R., Sun, Y.-H., Zhang, X.-H. & Chiang, V. L. Poly(T) Adaptor RT-PCR. in *Next-Generation MicroRNA Expression Profiling Technology* (ed. Fan, J.-B.) vol. 822 53–66 (Humana Press, Totowa, NJ, 2012).

26. Bianchi, F. *et al.* A serum circulating miRNA diagnostic test to identify asymptomatic high-risk individuals with early stage lung cancer. *EMBO Mol Med* **3**, 495–503 (2011).

27. Hicks, S. D. *et al.* Diagnosing mild traumatic brain injury using saliva RNA compared to cognitive and balance testing. *Clinical and Translational Medicine* **10**, (2020).

28. Zyrina, N. V., Antipova, V. N. & Zheleznaya, L. A. *Ab initio* synthesis by DNA polymerases. *FEMS Microbiol Lett* **351**, 1–6 (2014).

29. Tan, E. *et al.* Specific versus Nonspecific Isothermal DNA Amplification through Thermophilic Polymerase and Nicking Enzyme Activities. *Biochemistry* **47**, 9987–9999 (2008).

30. Urtel, G., Van Der Hofstadt, M., Galas, J.-C. & Estevez-Torres, A. rEXPAR: An Isothermal Amplification Scheme That Is Robust to Autocatalytic Parasites. *Biochemistry* **58**, 2675–2681 (2019).