Project title: Implications of Medical Low Dose Radiation Exposure

Grant Agreement Number: 755523

Call identifier: NFRP-2016-2017

Topic: NFRP-9

Deliverable D.5.8

Report on results of the analysis miRNA and lncRNA (SCK CEN)

Lead partner:	SCK CEN
Author(s):	Rafi Benotmane (SCK CEN)
Work Package:	WP5
Estimated delivery:	30 November 2021
Actual delivery:	20 December 2021
Type:	Deliverable
Dissemination level:	Public

This project has received funding from the Euratom research and training programme 2014-2018 under grant agreement No 755523.

Table of contents

Contents

List of figures

List of tables

Abbreviations

miRNA: microRNA

lncRNA: long non-coding RNA

ncRNA: non-coding RNA

EMT: endothelial to mesenchymal transition

UCC: urothelial cell carcinoma

GBM: glioblastoma

rtqPCR: real-time quantitative polymerase chain reaction

PCA: ploly component analysis

1. Introduction

There is a common consensus that new clinical biomarkers are needed to provide powerful tools for monitoring cancer patient health following treatment (radiotherapy) or imaging (CT-scan). These would open new avenues for the early assessment of treatment side effects and for optimal follow up of patients following diagnosis or treatment. It has been previously shown that exposure to radiation induces changes in gene expression in a dose dependent way. In addition, it was observed that genes modulated at high doses are mainly involved in stress response and apoptosis, while low radiation doses induced genes mostly related to inflammatory response and late epigenetic events. Many of the low dose expressed genes correspond to non-coding RNAs (JARED M et al., 2021). Non-coding RNAs "ncRNAs" are short RNAs that have been widely described as being stable in many body fluids (Orozco AF et al., 2010) and are emerging as new regulators of diverse biological functions, including oncogenesis and tumor progression. NcRNAs are classified into two types based on their transcript sizes: small ncRNAs (<200 bp) e.g., miRNAs and SiRNAs, while long ncRNAs (>200 bp) (16). Over the last two decades, ncRNAs such as microRNAs (miRNAs) and long noncoding RNAs (lncRNAs) have been shown to be excellent cancer biomarker candidates and can be detected in surrogate tissues and biofluids (Li Y, Jiang T, et al., 2020; Xiao Y et al., 2020; Ruxandra Volovat S et al., 2020; Baraniskin A et al., 2010; Bhagirath D et al., 2018; Piao X-M et al., 2018). NcRNAs operate as multifunctional epigenetic upstream regulators that can affect many tumor-related signalling pathways at the same time. Many are identified as biomarkers of radiation damage and represents a potential avenue for further exploration. They circulate stably through the body fluids, in addition to their intracellular role as direct/indirect gene expression modulators, preserving their ability to regulate mRNA target(s) (Orozco AF et al., 2010). They are becoming potential predictive indicators in the development and prognosis of cancer. In recent years, ncRNAs, particularly miRNAs, have been shown to be abnormally expressed in human breast cancer and seems to play an important role in the carcinogenesis of breast cancer (Mei J et al., 2020; Crudele F et al., 2020) and cardiovascular diseases. Long non-coding RNAs have been shown in other studies to act as oncogenes or tumor suppressors by targeting genes that promote or inhibit tumor growth and progression. Furthermore, they usually impact cancer cell proliferation, apoptosis, metabolism, metastatic spreading, and treatment resistance, regardless of whether they function as promoters or inhibitors throughout cancer progression (Venkatesh J et al., 2021).

Biomarkers' detection in easily accessible body fluid such as saliva has obvious advantages, namely ease of collection not requiring trained personnel, non-invasiveness reducing the discomfort of sample collection as well as better patient compliance. Saliva-based adjunct tests play a critical role in improving conventional medical assessments of serious systemic diseases over the last decade (Khurshid Z, et al., 2021). As a result, the development of a salivary sensitive assay for the accurate, rapid, and non-invasive identification of biomarkers would be extremely beneficial to the detection and screening of cancer. Several lncRNAs have already been shown to be useful as diagnostic and prognostic cancer biomarkers in clinical practice. They have been shown to enable the detection of different cancer types with high specificity and sensitivity such as bladder (M, Hoefig et al., 2010) and they have been shown able of classifying urothelial cell carcinoma (UCC) patients (Mengual L et al., 2013). In addition, other ncRNAs have been identified in sputum lung squamous cell carcinoma patients, helping thus in the early detection of lung squamous cell carcinomas (Xing L et al., 2010). The rapidly expanding catalogue of salivary ncRNA biomarkers for several cancers, suggests that ncRNAs

will become increasingly important in cancer patient management in the near future (Grillone K et al., 2020).

In the MEDIRAD project within subtask 5.2 we aimed at studying long-term effects of low doses radiation exposure on risk of cancer in children by identifying non-coding RNA biomolecules in saliva of children/young adults previously exposed to low dose radiation (CT-scan) at young age. For more accurate risk estimation, we use a nested case control study in a cohort of CT scan exposed individuals with risk factors for haematological malignancies and/or brain tumours. In addition, we have been conducting biological investigations in this task allowing deciphering genetic and epigenetic factors, which might link low dose radiation with the risk of brain cancer, which would improve medical RP. The main objective is to identify ncRNA biomarkers of low dose radiation sensitivity, which may help for identification of most sensitive individuals at risk of cancer development. These biomarkers might be involved in molecular pathways of early cellular carcinogenesis, which may help in early detection of patient with higher susceptibility to radiation induced cancer risk. Of course, such molecular pathways are modulated differently in the population depending on the exposure dose, age at exposure and the genetic background of the individual, also called individual radiosensitivity.

We started by optimizing the protocol for saliva collection and storage, then different RNA extraction kits have been tested and adapted to be able to extract both long and small RNA species in the same reaction. Finally, good quality RNA with reasonable amount was used for cDNA synthesis and amplification with qPCR of micro RNAs and long non-coding RNAs. The tested miRNA and lncRNA primers, which were previously designed based on a list of biomarkers agreed on with all partners. Unfortunately, the expected saliva samples from the case-control cohort were delayed on several occasions, first due to GDPR issues, then the Covid crisis, and at the end until the fourth year, we did not receive any sample. We have presented a plan B to the project coordination to overcome the delay and to fit the objectives as initially planned. Since an increase in brain cancers was observed in the CT-scan cohort, we decided to conduct identification of non-coding RNA in glioblastoma brain tissues obtained in collaboration with University of Antwerp Biobank (Belgium). MicroRNAs were extracted from the tissue samples and small RNAseq was performed to identify differentially expressed miRNA between glioblastoma and control tissues. On the other hand, meta-analysis using bioinformatics tools was conducted to cluster open-source data of glioblastoma to identify common signatures of long non-coding RNA, out of which some might be secreted in saliva. Finally, validation of the glioblastoma tissue samples data was operated using qPCR on selected miRNAs on a glioblastoma cell line U87 exposed to low and high doses X-rays.

In all, and despite the difficulties encountered to obtain the saliva samples originally scheduled for the optimization of detection of non-coding RNA in saliva to assess cancer progression, we were able to optimize a full protocol for detection of small and long non-coding RNAs in saliva. We identify several miRNAs and lncRNAs specific to brain cancer differentially expressed in glioblastoma tissue and U87 cell lines vs healthy controls. The identified biomarkers require further validations in CT scan patient saliva samples to validate their late expression after exposure at young age to low dose radiation. This deliverable provides an innovative initiative for the detection of noncoding RNAs in saliva as originally stated in the objective of the task, as well as a panel of micro RNAs and long non-coding RNAs modulated in brain cancer tissue and cell lines.

2. Non-coding RNA in saliva

WP5 aims to **improve the direct estimation of cancer risk** following low doses of ionizing radiation (IR) from computed tomography (CT) scanning in childhood and adolescence and to **study the role of factors - including age and genetic and epigenetic variants -, which may modify this risk**. In turn, Task 5.2 has as objective to conduct molecular epidemiological case-control study of haematological malignancies and brain tumours nested within the French and Spanish EPI-CT cohorts in order to **study potential markers of radiation sensitivity**. This task should consist in the conduct of the analysis of 100 cases of leukaemia, lymphoma and brain tumor and 100 counter-matched controls to **identify genetic and epigenetic variants that may increase the risk of cancer** in children and adolescents who have been previously scanned by CT for diagnostic purposes.

2.1. Saliva collection

We use the DNA genotec CP-190 tubes for saliva collection. As described in CP-190 tubes sampling protocol (***Figure 1***), the donor needs to not have had anything to eat, drink and had not chewed gum for 30 minutes before giving the saliva sample. If this is not the case, then the patient must wait for 30 minutes before being allowed to give the saliva sample. After this, the donor needs to carefully spit saliva into the designated tube until it is filled to the line indicating the limit of sample collection (figure 1), which is about 2 ml. After this, the tube cap will be securely closed until hearing a click and be sure that the liquid in the cap is released. The tube will then be left in a vertical position for 5-10 min in a tube holder to ensure that all the liquid is in the bottom part, then should unscrew the upper collection cap and replace it with the regular screw cap provided in the packaging. After this, should shake the tube by inverting it for 5-6 times. Finally, the tube will be disinfected by spraying it with the available lab disinfectant (e.g. Umonium, alcohol) to prevent any possible contamination or infection present in patient saliva (e.g. COVID-19).

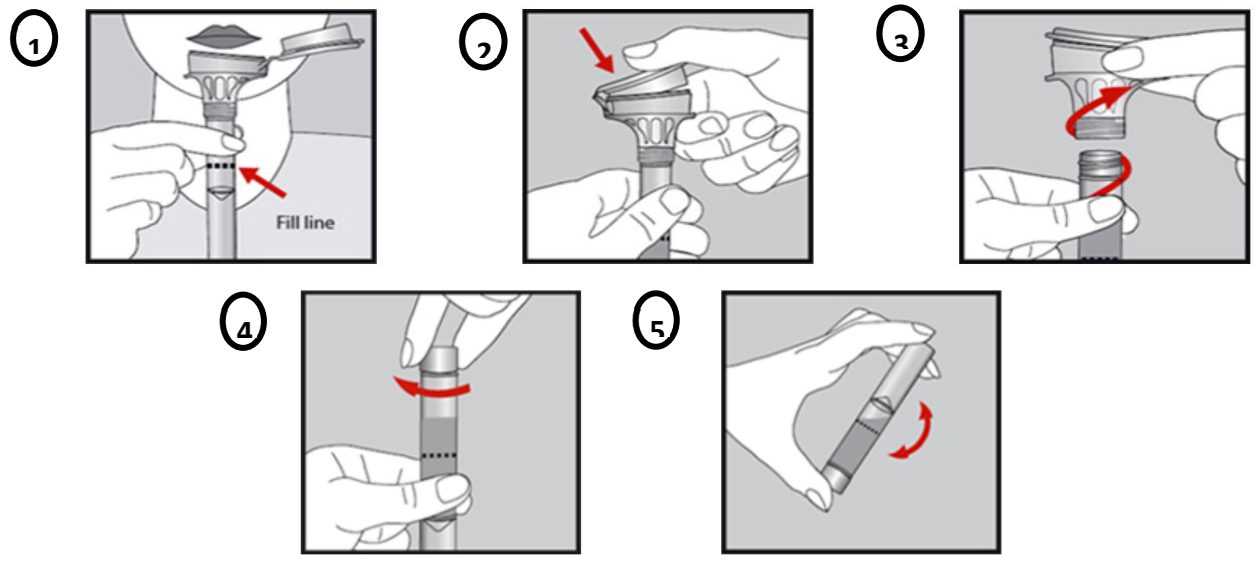


Figure 1: Procedure for saliva collection using DNA genotech CP-190 collection tube for RNA extraction

2.2. Saliva RNA extraction

At SCK CEN, we performed optimization of RNA extraction from saliva to concurrently extract microRNAs (miRNAs) and long noncoding RNAs (lncRNAs). Normally, RNA extraction kits allow to purify RNA molecules > 200 nucleotides. This is not sufficient for our present study as miRNAs have a length of around 22 nucleotides. Therefore, our RNA extraction kit needed to be capable of extracting total RNA including both lncRNAs and miRNAs. We selected Qiagen's miRNeasy Serum/Plasma Kit as it promised miRNA enrichment in addition to total RNA extraction. However, since no kit is optimized for extraction of RNA from saliva. We had to first adapt the protocol to use on saliva samples. Therefore, we applied the protocol on saliva samples using varying starting volumes to select the optimal volume providing the highest RNA concentration. Our results showed that 750 µLs produced the highest amount of extracted RNA and thus this volume was chosen for further experiments (***Figure 2***).

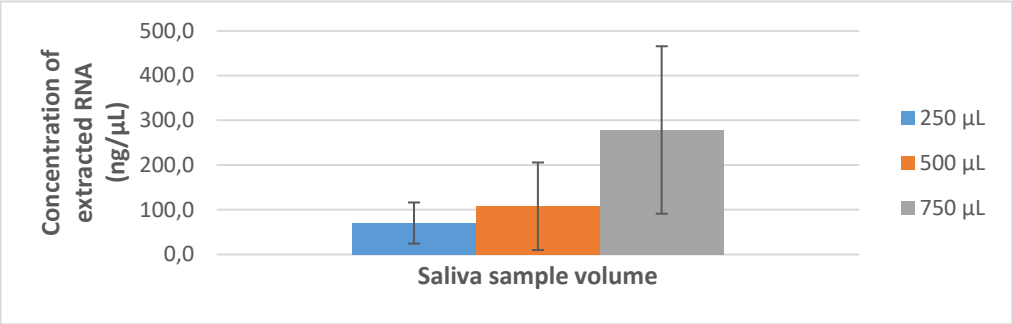
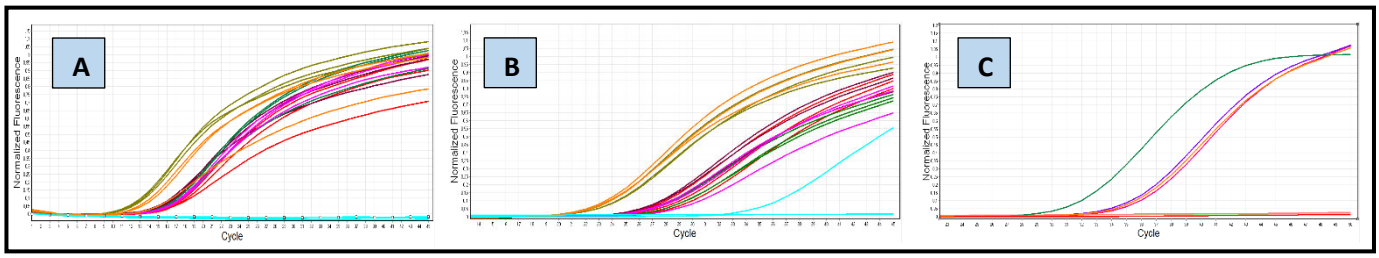


Figure 2: Concentration of RNA extracted from varying volumes of saliva. RNA was extracted using miRNeasy Serum/Plasma Kit using 250, 500 and 750 µLs of saliva respectively according to kit protocol. Concentration of RNA was analyzed using Lunatic by Unchained Labs. Results are from 3 replicates (Mean ± SD)

2.3. Saliva cDNA synthesis and quantitative PCR

Then, we proceeded to optimize the reverse transcription reaction of the extracted RNA. Our aim here was to use a single reverse transcription reaction for both miRNAs and lncRNAs. This is especially difficult because miRNAs are short sequences and thus it is difficult to design appropriate primers. 3 kits were tested for the concurrent reverse transcription of miRNAs and lncRNAs: High-Capacity RNA-to-cDNA, miScript II RT Kit and miRCURY LNA RT Kit. We then proceeded to select miRCURY LNA RT kit as it enabled the concurrent extraction of miRNAs and lncRNAs (c.f. High Capacity RNA-to-cDNA kit) while not necessitating the use of specific downstream kits (c.f. miScript II RT kit).

We also assessed the capacity to quantify miRNAs and lncRNAs in volunteer saliva samples whereby the protocol of RNA extraction and reverse transcription was performed and quantification of a number of miRNAs and lncRNAs (e.g. mir-21, miR-100, HOTAIR) was successfully performed (***Figure 3***).



2.4. Target non-coding RNA for the study

We conducted a literature search to select non-coding RNAs (miRNA or lncRNA) involved in the processes of cancer progression (proliferation, invasiveness) and metastasis. The table below summarizes the selected genes with a small description of their functions (table 1). Taqman probes corresponding to the selected biomarkers are available at SCK CEN and will be tested using qPCR to

Figure 3: Raw qPCR expression results of A)miR-21, B)miR-100 and C)HOTAIR

miRNA	lncRNA
miR21 indirectly regulate antiapoptotic Bcl-2 as anti- miR-21 treatment lead to downregulation of Bcl-2 →**increased proliferation**	**HOTAIR** is an lncRNA that plays a role as an oncogenic molecule in different cancer cells, such as breast, gastric, colorectal, and cervical cancer cells. Therefore, *HOTAIR* expression level is a potential biomarker for diagnostic and therapeutic purposes in several cancers
miR128 Upregulation of anti-apoptotic BCL-2 → **increased proliferation**	**RUNXOR (RUNX1 overlapping promoter-derived noncoding RNA) (216 kb)** epigenetically upregulates RUNX1 which is a key regulator of hematopoiesis
miR15b tumor suppressor/oncogene? Tumor suppressor in glioma. Cyclin D1 (also by mir-9) and E1 which leads to interruption of cell cycle→ **decreased proliferation and invasiveness**	**MEG3 Maternally expressed gene 3 (MEG3) (1.6 kb)**: inhibits miR-184 which induces proliferation and invasion, induces and enhances binding of p53, increases RB and leads to decreased E2F and inhibited cell proliferation
miR124 tumor suppressor, inhibit Rac1 from the MAPK pathway→ decreased cell proliferation and metastasis	**CCAT1 Colon cancer-associated transcript-1 (2kb):** inhibits myeloid cell differentiation and promoted cell proliferation by inhibiting miR-155
miR222/221 inhibit CDKN1B and CDKN1C which code for cyclin-dependent-kinase (CDK) inhibitor proteins leading to arrest of cell cycle → **increased proliferation.** Also inhibit PTEN and tissue inhibitor of metalloproteinases (TIMP-3) → **increased proliferation and invasiveness**	**NEAT1 Nuclear paraspeckle assembly transcript 1 inhibits** miR-23a-3p and consequently inhibits cell proliferation
miR9 Overexpressed in only primary and not metastatic brain tumors. Interestingly, overexpression was found only in glioblastoma with oligoneural origins thereby suggesting a possible benefit to miRNA profiling of glioblastoma prior to treatment	**IRAIN (5.4 kb)** orchestrates interaction between IGF1R promoter and enhancer. Maybe used as prognostic marker since the AML low-risk patients had greater *IRAIN* lncRNA expression than did the AML high-risk patients.
miR218, is a tumor supressor, inhibit thehigh mobility group box-1 (HMGB1)→ decreased proliferation and invasiveness	**MALAT1** ***Metastasis-associated lung adenocarcinoma transcript 1 (MALAT1) or also known as nuclear enriched abundant transcript 2*** Assumed to affect activated protein kinase (ERK/MAPK) signaling activity and expression of matrix metalloproteinase 2 (MMP). Sponge inhibition of tumor suppressive miR-101 which has been shown to inhibit proliferation, migration and invasion through inhibition of SOX9, KLF2 AND LMO3

Table 1: Selection of miRNAs and lncRNAs for expression assessment in patients saliva

Given the inability to obtain enough human samples to perform the study on patient saliva (only 3 samples received from the France collaborators), because of the various delays (GDPR, COVID), a new approach was proposed to get as close as we can to the MEDIRAD objectives: **to study potential markers of radiation sensitivity that may increase cancer risk after CT scanning**. Since brain tumors have been observed to increase in the EpiCT cohort, we first assessed the expression of miRNA species in glioblastoma biopsies as well as glioblastoma immortalized cell line (U87), instead of performing an epidemiological case-control study with human samples. Glioblastoma tissue samples as well as control brain tissue samples were received from University of Antwerp biobank after submission and approval of an ethical file describing the purpose of the study.

3. Small RNA sequencing of glioblastoma tissue samples

3.1. RNA extraction

miRNA extraction was performed using miRNeasy Serum/Plasma Kit (Cat. No.: 217184, Qiagen). First, QIAzol lysis reagent was added to 20 ng of tissue and lysed using TissueLyser II (Qiagen) for 2 min at 20 Hz. This lysed sample was then used for the subsequent miRNA extractions following miRNeasy Serum/Plasma kit protocol. In short, 5 volumes QIAzol Lysis Reagent was added to the lysate and incubated at room temperature (15–25°C) for 5 min. Then, an equal volume of chloroform to the starting sample was added, shaken vigorously for 15 seconds, incubated at room temperature for 2–3 min and then centrifuged for 15 min at 12,000g at 4°C. After that, 1.5 volumes of 100% ethanol were added to the upper aqueous phase then this was added in 700 µl aliquots into an RNeasy MiniElute spin column and centrifuged at 8000g for 15 seconds at room temperature. Washing of the column was performed by a series of washing steps involving the addition of different buffers (Buffer RWT, Buffer RPE and 80% ethanol) to the RNeasy MiniElute spin column and centrifugation for 15 seconds, 15 seconds and 2 mins at 8000g, respectively. Extracted RNA was then eluted from the column in 30 µl RNase-free water directly by centrifugation for 1 minute at full speed. Quality control of the extracted RNA was assessed by the Agilent Bioanalyzer 2100. Since the RNA is isolated from biobank frozen tissue for many years, the quality of RNA is quite variable between the different samples.

3.2. Small RNA sequencing

Small RNA sequencing was performed on 20 individual tissue samples (with the highest quality) to identify differentially expressed miRNAs biomarkers in excised glioblastoma tumor tissue. Sequencing libraries were constructed using 2 µg of RNA per sample and Illumina's TruSeq Small RNA Sample Prep Kit with the manufacturer's protocol. Illumina's HiSeq 2000 system was used for 2×150 nucleotides, single-end sequencing, conducted at GENEWIZ (www.genewiz.com). The average number of reads was 16 million reads/sample.

3.3. Small RNA sequencing data analysis

RNAseq reads were quality checked using fastqc (v 0.11.9), while the rest of the analysis was conducted using CLC workbench (v 21.0.3). First, reads were quality trimmed, with 0.05 quality limit, removing reads with any ambiguous nucleotides or less than 15 bp, and length was trimmed to a max of 55 bp (using Trim Reads module). Next, reads were mapped to miRBase (species *Homo Sapiens,* release 22) using the default parameters of Quantify miRNA module. Lastly differential expression analysis was done between the mature miRNA profiles of the control and glioblastoma samples, with TMM as normalization methods (with stable housekeeping genes as reference), using Differential Expression in Two Groups module (***Figure 4***).

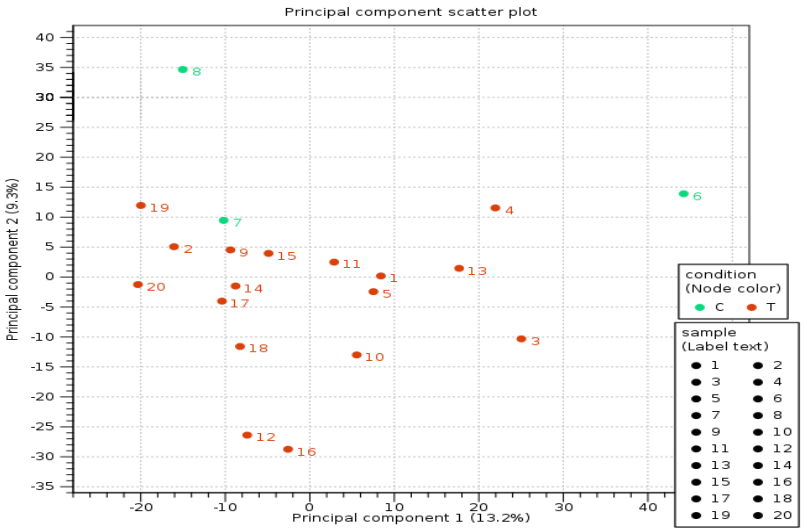


Figure 4: PCA of the 20 assayed samples of 17 glioblastoma tissue and 3 control samples

Six miRNAs exhibited significantly altered expression (FDR adjusted p value<0.05) in glioblastoma tumor tissue relative to control brain tissue: hsa-miR-183-5p, hsa-miR-183-3p, hsa-miR-1246, hsa-miR-182-5p, hsa-miR-549a-3p and hsa-miR-96-5p. ***Figure 5*** shows a heat map of all identified differentially expressed miRNAs in glioblastoma tissue samples regardless of their adjusted p values and a principal component analysis of the samples used, respectively.

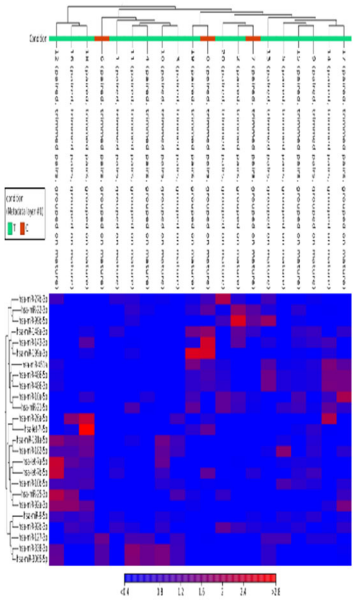


Figure 5: Heat map of differentially expressed miRNAs identified in glioblastoma tissue

3.4. Real time quantitative PCR validation

We then proceeded to validate these identified candidate miRNA biomarkers in the remaining glioblastoma tissue samples using qPCR. miRNA extraction was performed using miRNeasy Serum/Plasma Kit (Cat. No.: 217184, Qiagen). Reverse transcription and cDNA synthesis was performed using miRCURY LNA RT Kit (Cat. No. 339340, Qiagen) and selected miRNAs were quantified using QuantiNova SYBR® Green PCR Kit (Cat. No.: 208056, Qiagen) using Qiagen pre-designed miRNA primers relative to the average expression of 2 reference miRNAs U6 and SNORD91A (miR-183-5p: Cat. No.YP00206030, miR-183-3p: Cat. No.YP00205673, miR-1246: Cat. No.YP00205630, miR-182-5p: Cat.

No.YP00206070, miR-96-5p: Cat. No.YP00204417, miR-549a: Cat. No.YP00205975, U6: Cat. No. YP00203907 and SNORD91A: Cat. No. SBH0265807-200).

3.5. Real time quantitative PCR of glioblastoma tissue

Mir-96-5p was found to have significantly upregulated expression in glioblastoma tumour tissues relative to control brain samples (***Figure 6***). MiR-96-5p is involved in many cancers and was shown to enhance cell proliferation and invasion in gastric cancer, inhibits apoptosis by targeting the caspase-9 gene in in hepatocellular carcinoma (Nato Iwai et al., 2018). In another study, miR-96-5p was shown to promote cell growth and metastasis in glioma (Shoudan Zang., 2019)

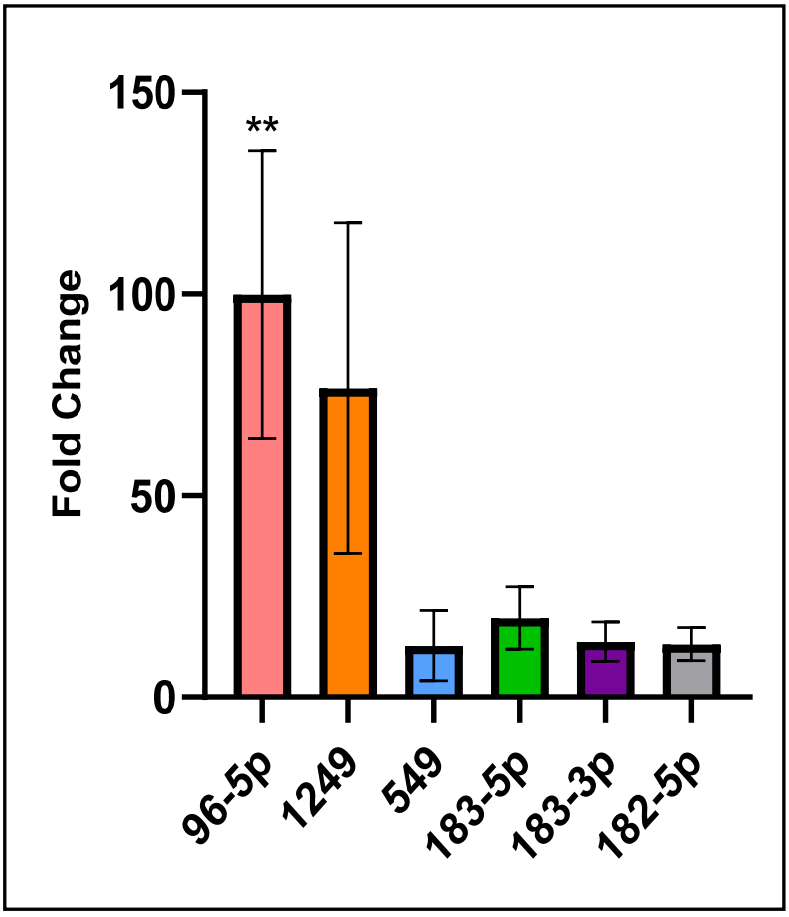


Figure 6: Relative fold change of sequencing-identified candidate glioblastoma miRNA biomarkers in in remaining glioblastoma tissue samples (normal brain n= 3, glioblastoma tissue n=28) (** → p<0.01)

4. Glioblastoma U87 cell line validation

U87-MG cell culture: We then proceeded to validate these glioblastoma tissue-identified miRNA candidate biomarkers in U87 glioblastoma cell line. 5 independent biological replicates were performed on consecutive weeks with 6 technical replicates per condition per week. As can be seen in ***Figure 7***, U87-MG cell lines (passages 11-15) from ATCC were cultured (75,000 cells for 24 hour plates

and 35,000 cells for 72 hour plates) in EMEM medium supplemented with 10% fetal bovine serum and 5% Penicillin Streptomycin solution. Cells were irradiated with two doses (0.1 and 2 Gy at dose rate of 500 mGy/min) along with sham irradiated controls.

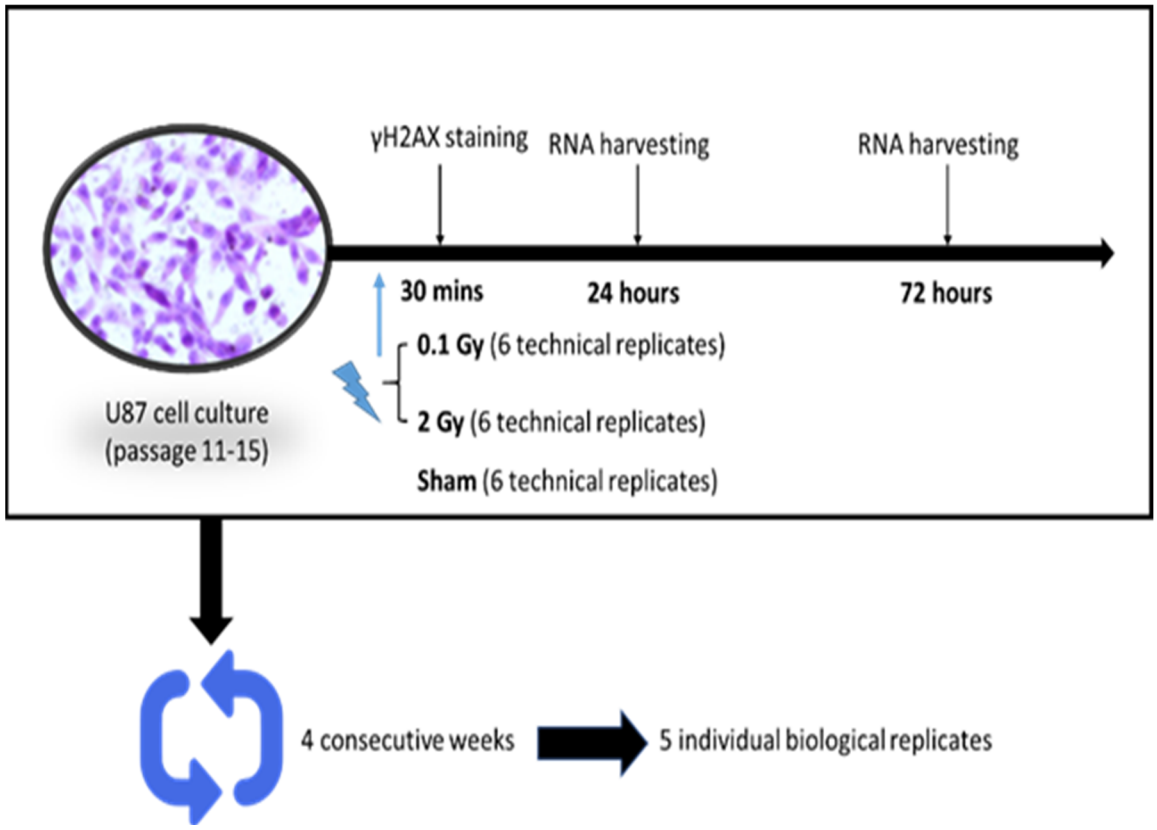


Figure 7: Scheme of U87-MG cell culture experiment. U87-MG cells were either irradiated with 0.1 or 2 Gy. Sham replicates were not exposed to any radiation. Cells were fixed in 8-well Lab-Tek chamber slides for γH2AX staining 30 minutes after irradiation. Then, cells were harvested at 24 and 72 hours after irradiation. This setup included 6 technical replicates per condition and it was repeated to ascertain replicability in the following 4 weeks

4.1. DNA damage analysis:

We performed γH2AX staining of U87 sham, 0.1 Gy and 2 Gy irradiated cultures. γH2AX is a marker of DNA double strand breaks and can therefore be used to monitor DNA repair after irradiation (Nagelkerke & Span, 2016). However, due to problems with cell attachment to the slide chambers, insufficient cells were present for quantification of DNA damage and repair (***Figure 8***).

Cells were fixed 30 minutes for γH2AX staining after irradiation in 8-well Lab-Tek® Chamber slides (Cat. No.: C7182, Sigma Aldrich) with 2% PFA for 30 min at room temperature, and permeabilized with 0.25% Triton X-100. Following blocking with PIG in TNB for 1 hour at room temperature, cells were incubated with 1:300 dilution of mouse monoclonal γ-H2AX antibody (*Millipore 05-636*) and 1:1000 dilution of rabbit polyclonal *53BP1* antibody (Novus Biologicals NB100-304) for 1 hour at 37°C. The cells were then incubated with 1:300 dilution Alexa fluor 488 goat anti-mouse antibody and 1:1000 dilution Alexa fluor 568 goat anti-rabbit antibody for 1 hour in the dark at 37°C. Following staining with DAPI, cells were washed and stored at -20°C for analysis. Cells were analyzed under a fluorescence microscope (×200 magnification; Olympus Corporation, Tokyo, Japan).

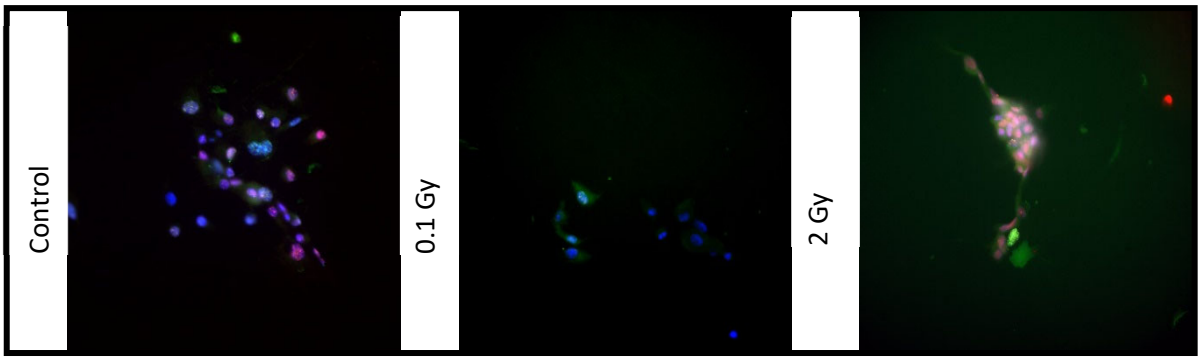


Figure 8: γH2AX and 53BP1 staining of Sham, 0.1 Gy and 2 Gy irradiated U87-MG cell culture

4.2. Real time quantitative PCR

Cells were harvested at 24 and 72 hours after irradiation. RNA was then extracted using miRNeasy Serum/Plasma Kit (Cat. No.: 217184, Qiagen). Reverse transcription and cDNA synthesis was performed using miRCURY LNA RT Kit (Cat. No. 339340, Qiagen) and selected miRNAs were quantified using QuantiNova SYBR® Green PCR Kit (Cat. No.: 208056, Qiagen) using Qiagen pre-designed miRNA primers relative to the average expression of 2 reference miRNAs U6 and SNORD91A.

2 miRNAs (mir-96-5p and mir-182-5p) (***Figure 9A and B***) showed significantly decreased expression 24 hours after irradiation with 0.1 Gy ionizing radiation (IR) relative to sham irradiated controls. However, this decrease seems to be transient with levels returning to normal levels at 72 hours after irradiation. Further investigations are required to understand the whole picture and the role of miR-96-5p in glioblastoma cell line U87.

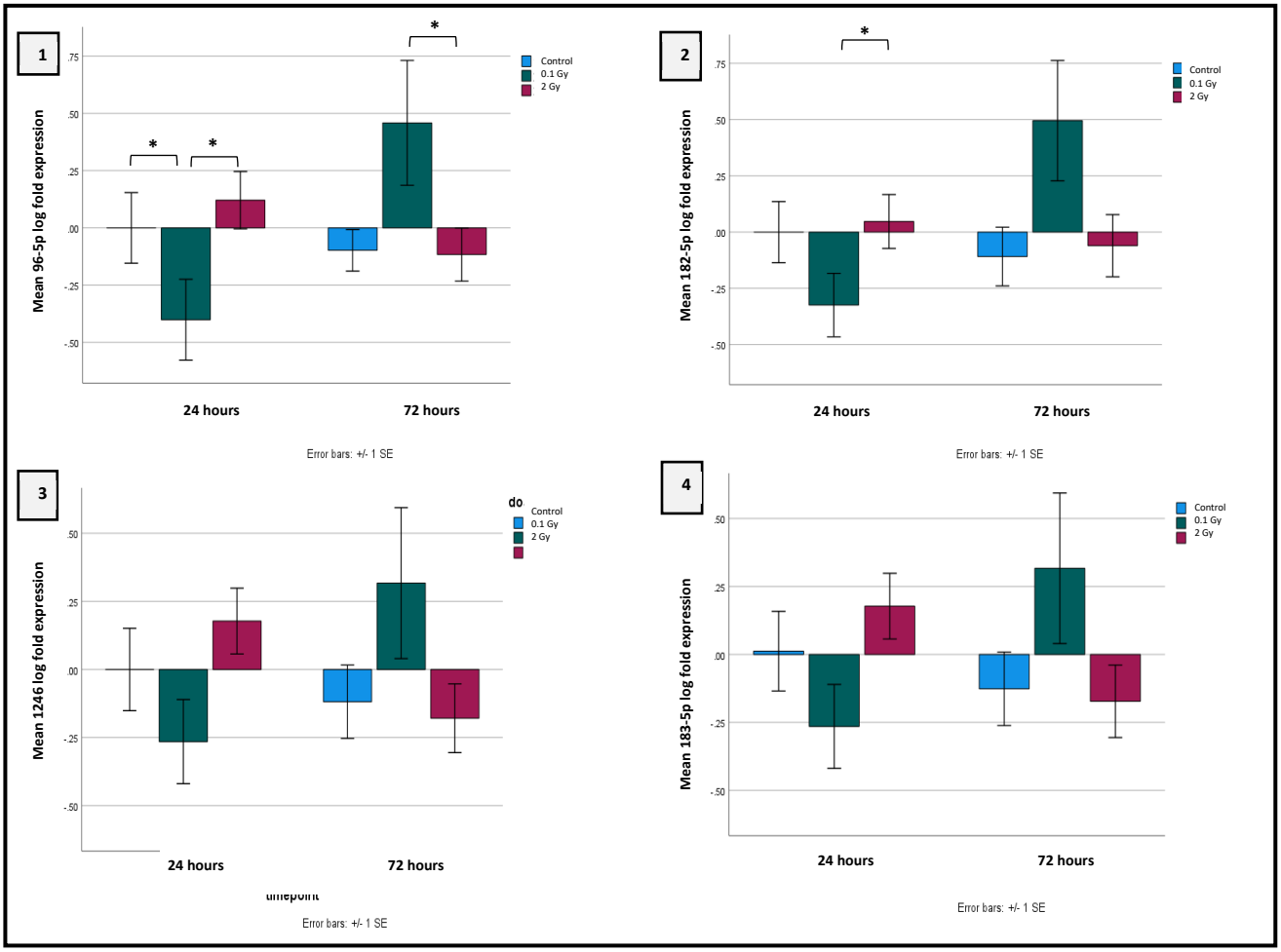


Figure 9: Relative mean log fold change of mir-96-5p (1), mir-182-5p (2), mir-1246 (3) and mir-183-5p (4) in U87-MG cell lines irradiated with 0.1 or 2 Gy and harvested at 24 and 72 hours after irradiation (** → p<0.01)

5. Meta-analysis of lncRNA candidate biomarkers of Glioblastoma:

We also initiated a meta-analysis of freely available glioblastoma tissue RNAseq raw data to identify candidate lncRNA biomarkers for glioblastoma, some of which might be secreted in saliva. Identifying specific glioblastoma non-coding RNA markers in saliva would help in early detection of patient at risk of developing brain cancer. We searched for freely available RNAseq raw data of de novo glioblastoma tissue samples in Gene Expression Omnibus which is a public functional genomics data repository containing both array- and sequence-based data (Barrett et al., 2013). Our search consisted of the following search terms: (((RNA sequencing) OR "Sequence Analysis, RNA"[Mesh])) AND ((((((("Glioblastoma" [Mesh]) OR glioblastoma [Title])) AND expression)). Our search yielded 313 results which underwent further rounds of filtration considering the following criteria: 1) Study analyses glioblastoma patient tissue samples and 2) Studies use case-control. After filtration, four studies were selected for data download and further bioinformatic processing (***Table 2***)

Study	Number of samples
T et al., 2021	15 glioblastoma tissue samples and 15 paired controls
Gill et al., 2014	20 glioblastoma tissue samples and 20 paired controls
Stathias et al., 2014	2 glioblastoma tissue samples and 2 normal brain tissues from epilepsy patients controls
Yuan et al., 2018	3 glioblastoma tissue samples and 3 paired controls

Table 2: Summary of the samples included in the selected studies for meta-analysis

	IQC	AQCg	CQCg
Study1	5.615465	23.07654	240.9925
Study2	3.299765	2.493493	52.77707
Study3	1.30103	2.665496	17.73623
Study4	5.615465	61.24808	145.1773

Table 3: Quality control parameters for meta-analysis

Bioinformatics analysis was performed as described previously (Han, Hua, Xue, & Zhu., 2019). We first downloaded the sequence data of these studies by *Prefetch* and converted them into FASTQ files using *fastq-dump* tool of the SRA Toolkit software. Next, we downloaded the reference sequences of lncRNA and protein coding transcripts in FASTA format from NONCODE and Ensemble respectively and merged the two FASTA format files. Then, we removed the pseudogenes and performed the quantification of the lncRNA and protein-coding transcripts simultaneously by mapping the RNA-seq reads of each study to the merged reference sequence (pseudoalignment) and calculating the count values using *Kallisto* software. For the paired-end sequencing samples, the arguments were set to defaults, i.e., the number of bootstrap samples (-b) equals 0 and the number of threads (-t) equals 1. For the single-end sequencing samples, besides these default parameter settings, we set the estimated average fragment length (-l) and the standard deviation of fragment length (-s) to 200 and 20, respectively, according to Kallisto's recommended parameters. Finally, based on the annotation file "Transcript2Gene," we integrated transcript-level count values of lncRNAs to calculate their corresponding gene-level count values using the R package "tximport". Table 3 shows the quality control parameters of the analysis.

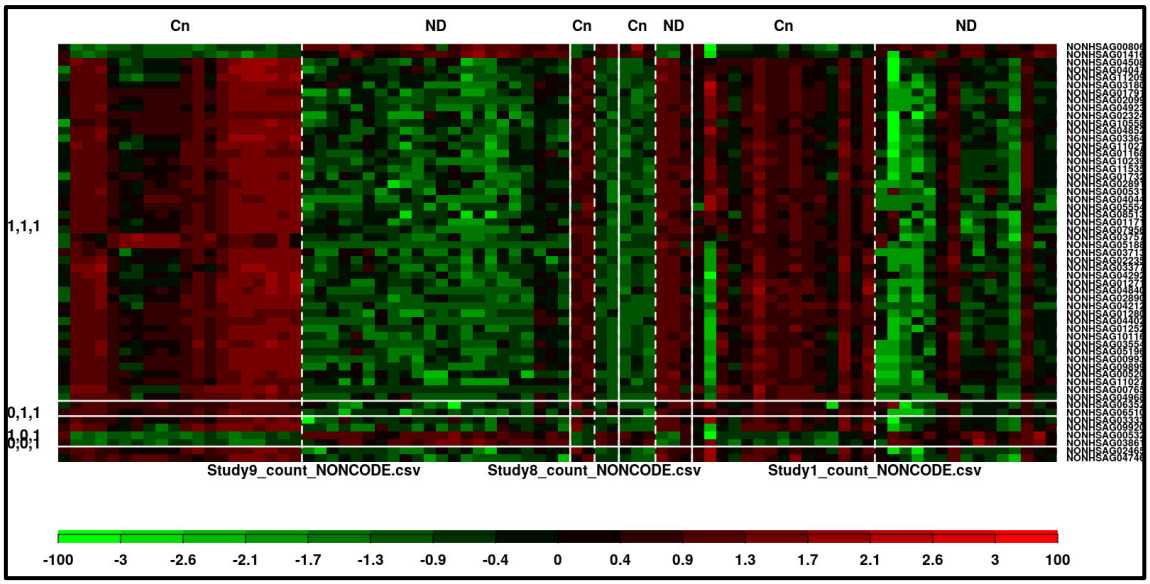


Figure 10: Heat map showing lncRNA expression in glioblastoma tissue samples (ND) and paired control brain tissue (Cn) in the 4 selected studies

Bioinformatics analysis of glioblastoma tissue samples relative to paired controls identified 617 lncRNAs that were significantly differentially expressed in all four selected studies with 176 having Ensemble IDs (listed in ***Supplementary Table 1)***. Further literature review is still needed to identify the common pathways between the identified lncRNA candidate biomarkers and subsequently glioblastoma tumorigeneses.

6. Discussion and Conclusion

Radiation-induced DNA damage is shown to lead to accumulation of genetic abnormalities, which can result in the development of leukaemia or other solid tumours approximately 10 up to 30 years after exposure (Prasad KN et al;, 2004). Identification of specific ncRNAs of low dose radiation induced cancer could potentially allow to develop a ncRNA biomarker signature for early side effect detection and help clinicians to determine preventive measures to treat radiation damage in both cancer patients and civilians exposed to radiation. Several non-coding RNAs have been identified in recent years as key epigenetic players in radiation response (Jared M et al., 201). On the other hand, non-coding RNAs are playing important role in cancers tumorigeneses and metastasis and are also investigated as potential biomarkers and potential therapeutic targets (Banelli et al., 2017). In our experiments, we found miR-183-5p, miR-183-3p, miR-1246, miR-182-5p, miR-549a-3p and miR-96-5p to have significantly altered expression by small RNA sequencing of glioblastoma tissue samples. miR-183-5p has been associated with glioma cell proliferation, migration and invasion as well as being identified as a prognostic biomarker of glioblastoma progression (Pavlakis et al., 2017). In addition, previous research has shown that mir-1246 promotes glioma cell migration and invasion both in cell lines (U87MG, U251, and A172 glioma cells) and in vivo (nude mice bearing U87MG glioma cell xenografts) (Qian et al., 2021). On the other hand, mir-183 is a TGFβ-induced miRNA previously reported to suppress tumour-associated natural killer cells which was found to be upregulated in glioblastoma tumour centres when compared to the peri-tumor areas (Donatelli et al., 2014; Fazi et al., 2015). Previous research also linked mir-182-5p to glioblastoma tumorigenesis, angiogenesis and metastasis (Xue et al., 2016; Zhang et al., 2020). Next to that, mir-549a was found to be upregulated

in cerebrospinal fluid of glioblastoma patients in comparison to controls and was found to be predictive of prognosis in patients with tumours of glial origin (Drusco et al., 2018; Kopkova et al., 2019). Finally, 96-5p was found to be upregulated in glioma with downregulation of miR-96-5p inhibiting growth and metastasis of glioma (Zhang & Guo, 2019). Upregulation of miR-96 was also found to promotes radioresistance in T98G glioblastoma cells (P. Guo et al., 2018). On the other hand, qPCR validation of these miRNAs in remaining glioblastoma tissues samples and 0.1 Gy irradiated U87 cell culture showed significantly altered expression of mir-96-5p in both tested glioblastoma tissues samples and cell culture relative to respective controls. Nevertheless, additional validations are required to understand the contrasting expression patterns between the low and high dose radiation and between 24 h and 72 h in the cell culture experiment. In addition, further analysis of the meta-analysis results is still ongoing to identify specific lncRNA of glioblastoma progression.

Taken together, the initial objectives aimed at studying long-term effects of low doses radiation exposure on risk of cancer in children by identifying non-coding RNA biomolecules in saliva of children/young adults previously exposed to CT-scan radiation at young age. Despite the encountered difficulties to obtain saliva samples from the nested case control cohort of CT scan exposed individuals with risk factors for haematological malignancies and/or brain tumours, we were able to meet the original objectives. First, with optimizing a protocol to detect non-coding RNAs in saliva capable of quantifying both micro RNAs and long non-coding RNAs in the same RNA sample from patients' saliva. Second, we identify a panel of micro RNAs differentially expressed in glioblastoma tissue, as well as long non-coding RNAs from a meta-analysis data, which would provide strong base for screening methods in saliva. The miRNA panel was validated on glioblastoma cell line irradiated with low and high doses radiation. We evaluated the expression of the identified non-coding RNA after 24 h and 72 h in U87 cell line, which does not correspond to long-term effect after exposure. Nevertheless, the tested non-coding RNA were selected from RNAseq data of glioblastoma tissue samples as well as meta-analysis from databases of glioblastoma tissue. The identified non-coding RNAs might not correspond exactly to the long-term radiosensitivity response after low dose exposure, but would perfectly fit our aim to identify radiation susceptibility markers, which may help to identify individuals at higher radiation induced cancer risk. Therefore, further investigations are required to validate the identified non-coding RNAs in a large cohort of CT patients' saliva.

7. References

1. Orozco AF, Lewis DE. Flow Cytometric Analysis of Circulating Microparticles in Plasma. 2010; Available from: www.interscience.
2. Li Y, Jiang T, Zhou W, Li J, Li X, Wang Q, et al. Pan-cancer characterization of immune-related lncRNAs identifies potential oncogenic biomarkers. Nat Commun [Internet]. 2020;11(1):1–13. Page 5 of 10, http://dx.doi.org/10.1038/s41467-020-14802-2
3. Xiao Y, Xiao T, Ou W, Wu Z, Wu J, Tang J, et al. LncRNA SNHG16 as a potential biomarker and therapeutic target in human cancers. Biomark Res. 2020;8(1):1–11.
4. Ruxandra Volovat S, Volovat C, Hordila I, Hordila D-A, Camil Mirestean C, Miron OT, et al. MiRNA and LncRNA as Potential Biomarkers in Triple-Negative Breast Cancer: A Review. 2020; 10:1. Available from: www.frontiersin.org
5. Baraniskin A, Kuhnhenn J, Schlegel U, Chan A, Deckert M, Gold R, et al. Identification of microRNAs in the cerebrospinal fluid as marker for primary diffuse large B-cell lymphoma of the central nervous system. Blood [Internet]. 2011;117(11):3140–6. Available from: http://dx.doi.org/10.1182/blood-2010-09-308684
6. Bhagirath D, Yang TL, Bucay N, Sekhon K, Majid S, Shahryari V, et al. microRNA-1246 Is an Exosomal Biomarker for Aggressive Prostate Cancer. 2018; Available from: http://cancerres.aacrjournals.org/
7. Piao X-M, Jeong P, Kim Y-H, Joon Byun Y, Xu Y, Won Kang H, et al. Urinary cell-free microRNA biomarker could discriminate bladder cancer from benign hematuria. 2018;
8. Mei J, Hao L, Wang H, Xu R, Liu Y, Zhu Y, et al. Systematic characterization of non-coding RNAs in triple-negative breast cancer. Cell Prolif. 2020;53(5):1–19.
9. Crudele F, Bianchi N, Reali E, Galasso M, Agnoletto C, Volinia S. The network of non-coding RNAs and their molecular targets in breast cancer. Mol Cancer. 2020;19(1):1–18.
10. Venkatesh J, Wasson MCD, Brown JM, Fernando W, Marcato P. LncRNA-miRNA axes in breast cancer: Novel points of interaction for strategic attack. Cancer Lett [Internet]. 2021;509:81–8. Available from: https://doi.org/10.1016/j.canlet.2021.04.002
11. Khurshid Z, Warsi I, Moin SF, Slowey PD, Latif M, Zohaib S, et al. Biochemical analysis of oral fluids for disease detection [Internet]. 1st ed. Vol. 100, Advances in Clinical Chemistry. Elsevier Inc.; 2021. 205–253 p. Available from: http://dx.doi.org/10.1016/bs.acc.2020.04.005
12. Hanke M, Hoefig K, Merz H, Feller AC, Kausch I, Jocham D, et al. A robust methodology to study urine microRNA as tumor marker: MicroRNA-126 and microRNA-182 are related to urinary bladder cancer. Urol Oncol Semin Orig Investig [Internet]. 2010;28(6):655–61. Available from: http://dx.doi.org/10.1016/j.urolonc.2009.01.027
13. Mengual L, Lozano JJ, Ingelmo-Torres M, Gazquez C, Ribal MJ, Alcaraz A. Using microRNA profiling in urine samples to develop a non-invasive test for bladder cancer. Int J Cancer. 2013;133(11):2631–41.
14. Xing L, Todd NW, Yu L, Fang H, Jiang F. Early detection of squamous cell lung cancer in sputum by a panel of microRNA markers. Mod Pathol. 2010;23(8):1157–64.
15. Grillone K, Riillo C, Riillo C, Scionti F, Rocca R, Rocca R, et al. Non-coding RNAs in cancer: Platforms and strategies for investigating the genomic "dark matter." J Exp Clin Cancer Res. 2020;39(1):1–19.
16. Naoto Iwai Kohichiroh Yasui Akira Tomie Yasuyuki Gen Kei Terasaki Tomoko Kitaichi Tomohiro Soda Nobuhisa Yamada Osamu Dohi Yuya Seko Atsushi Umemura Taichiro Nishikawa Kanji Yamaguchi Michihisa Moriguchi Hideyuki Konishi Yuji Naito Yoshito Itoh. Oncogenic miR-96-5p inhibits apoptosis by targeting the caspase-9 gene in hepatocellular carcinoma. Oncology on: April 12, 2018. doi.org/10.3892/ijo.2018.4369
17. Shoudan Zhang Wenshi Guo. Long non-coding RNA MEG3 suppresses the growth of glioma cells by regulating the miR-96-5p/MTSS1 signaling pathway. Mol Med rep, September 10, 2019, p: 4215-4225. doi.org/10.3892/mmr.2019.10659
18. A Nagelkerke, PN Span. Staining against phospho-H2AX (γ-H2AX) as a marker for DNA damage and genomic instability in cancer tissues and cells. Tumor Microenvironment, 2016 – Springer
19. Barbara Banelli, Alessandra Forlani, Giorgio Allemanni, Anna Morabito, Maria Pia Pistillo, Massimo Romani. MicroRNA in Glioblastoma: An Overview. Int J Genomics. 2017;2017:7639084 doi: 10.1155/2017/7639084. Epub 2017 Nov 6.
20. Evangelos Pavlakis Anton B. Tonchev Ara Kaprelyan Yavor Enchev Anastassia Stoykova. Interaction between transcription factors PAX6/PAX6-5a and specific members of miR-183-96-182 cluster, may contribute to glioma progression in glioblastoma cell lines. Oncology reports January 30, 2017, p 1579-1592 https://doi.org/10.3892/or.2017.5411
21. Sarah S. Donatelli, Jun-Min Zhou, Danielle L. Gilvary, Erika A. Eksioglu, Xianghong Chen, W. Douglas Cress, Eric B. Haura, Matthew B. Schabath, Domenico Coppola, Sheng Wei, and Julie Y. Djeu. TGF-β–inducible microRNA-183 silences tumor-associated natural killer cells. PNAS March 18, 2014 111 (11) 4203-4208; https://doi.org/10.1073/pnas.1319269111
22. Jin Qian, Y ingna Xu, Xing Xu, Zhenyu Tao, Yang Luo, Yichang Xu, Yong Zhang, and Chunfa Qian. Hsa_circ_0091581 promotes glioma progression by regulating RMI1 via sponging miR-1243-5p. J Cancer. 2021; 12(11): 3249–3256. doi: 10.7150/jca.55558
23. Jianfei Xue, Aidong Zhou, Yamei Wu, Saint-Aaron Morris, Kangyu Lin, Samirkumar Amin, Roeland Verhaak, Gregory Fuller, Keping Xie, Amy B. Heimberger and Suyun Huang. miR-182-5p Induced by STAT3 Activation Promotes Glioma Tumorigenesis. CancerRes Tumor and Stem Cell Biology, DOI: 10.1158/0008-5472.CAN-15-3073 Published July 2016
24. panelHe Zhang, Liying Yang, Yichao Wang, Wenli Huang, Yang Li Shuchun Chen Guangyao, Song Luping Ren, Oxymatrine alleviated hepatic lipid metabolism via regulating miR-182 in non-alcoholic fatty liver disease. Life Sciences Volume 257, 15 September 2020, 118090
25. Alessandra Drusco, Paolo Faddab, Giovanni Nigitaa, Matteo Fassanc Arianna, Bottonia Marina, Pgardimanc Diana, Sacchic Federica, Calorea Mariantonia, Carosid Anna, Antenuccie Beatrice, Casinid Hesham, Kelania Edoardo, Pescarmonad Gianpiero, Di Levaf Nicola, Zanesia Mitchell, Sbergerg Carlo, M Crocea. Circulating Micrornas Predict Survival of Patients with Tumors of Glial Origin. EbioMedicine. Volume 30, April 2018, Pages 105-112
26. Alena Kopkova, Jiri Sana, Tana Machackova, Marek Vecera, Lenka Radova, Karolina Trachtova, Vaclav Vybihal, Martin Smrcka, Tomas Kazda, Ondrej Slaby,and Pavel Fadrus. Cerebrospinal Fluid MicroRNA Signatures as Diagnostic Biomarkers in Brain Tumors. Cancers 2019, 11(10), 1546; https://doi.org/10.3390/cancers11101546
27. Shoudan Zhang & Wenshi Guo. Long non-coding RNA MEG3 suppresses the growth of glioma cells by regulating the miR-96-5p/MTSS1 signaling pathway. MolMed reports, November-2019, Volume 20 Issue 5 p 4215-4225,
28. Pin Guo, Yanan Yu, Zibin Tian, Yingying Lin, Yongming Qiu, Weicheng Yao, Lijuan Zhang. Upregulation of miR-96 promotes radioresistance in glioblastoma cells via targeting PDCD4. International Journal of Oncology, October-2018, Volume 53 Issue 4 p1591-1600. https://doi.org/10.3892/ijo.2018.4498

29. Prasad KN, Cole WC, Hasse GM. Health risks of low dose ionizing radiation in humans: a review. Exp Biol Med (Maywood) 2004;229:378–82. Long and short non-coding RNA and radiation
30. JARED M. MAY, MICHELLE BYLICKY, SUNITA CHOPRA, C. NORMAN COLEMAN, and MOLYKUTTY J. ARYANKALAYIL. Long and short non-coding RNA and radiation response: a review. Translational Research, Volume 233 p162-179. https://doi.org/10.1016/j.trsl.2021.02.005