OXFORD

BIOLOGY Methods & Protocols

Biology Methods and Protocols, 2024, **9(1)**, bpae083

https://doi.org/10.1093/biomethods/bpae083

Advance Access Publication Date: 5 November 2024

Methods Article

Enhancing negative control selection: A comparative analysis of random and targeted sampling techniques for obtaining High-Quality RNA from normal breast tissue

Komal Mehta[1], Archana Sharma[1], Anurag Mehta[2] and Juhi Tayal [1,*]

[1]Biorepository, Department of Research, Rajiv Gandhi Cancer Institute and Research Centre, New Delhi 110085, India
[2]Histopathology, Department of Laboratory Services, Rajiv Gandhi Cancer Institute and Research Centre, New Delhi 110085, India

*Corresponding author. Biorepository, Department of Research, Rajiv Gandhi Cancer Institute and Research Centre, D-18, Sector-V, Rohini, New Delhi, 110085, India. Tel. +91 9871130522; E-mail: juhitayal76@gmail.com

Abstract

Molecular profiling is a crucial aspect of cancer therapy selection, underscoring the necessity for representative sampling of both tumor and normal tissues. While much attention has been given to representative tumor sampling, there has been a notable lack of exploration into the issue of poor RNA quality in normal breast tissue processing. Normal breast tissue from the same patient is often used as a negative control for most "-omics" experiments. RNA extracted from normal breast tissues frequently contains nucleic acids from surrounding adipocytes, endothelial cells, and immune cells, leading to a low representation of ductal elements and skewed results. Therefore, ensuring a complete representation of breast glandular tissue is imperative. The study aimed to investigate the variations in RNA enrichment between a random sampling technique and a targeted sampling approach when visually selecting normal breast tissue sections as negative controls for "-omics" experiments. Fifteen female breast cancer subjects who underwent Modified Radical Mastectomy were selected for the study. Normal Breast tissue was visually examined, and samples were collected from random fat pockets (random sampling) and fibromuscular grey-white streak areas (targeted sampling). RNA was isolated, followed by spectrophotometric analysis, agarose gel electrophoresis and Agilent Tape station analysis. Histopathological assessments and a gene expression study for housekeeping genes were performed on both subsets. Tissues collected through targeted sampling exhibited significantly higher RNA quality than those obtained via random sampling. Histopathological analysis revealed cellular areas abundant in terminal ductular units within the targeted samples, and a final validation qPCR showed that the targeted samples were the most representative of normal breast glandular tissue. The comparative analysis of the two sampling methods clearly indicates that the targeted approach, with its superior accuracy and reliability, is the more practical choice for obtaining representative normal breast glandular tissue for "-omics" experiments.

Keywords: Breast cancer; random and targeted tissue sampling; RNA isolation; formalin-fixed paraffin-embedded (FFPE) samples; triple negative breast cancer (TNBC)

Introduction

Biobanks play a vital role in advancing translational research by managing the acquisition, processing, annotation, storage, and distribution of biological specimens. Regarding cancer tissue biobanking, meticulous handling is crucial to preserve the RNA, DNA, and proteins in surgical specimens. The usability and fit-for-purpose samples are critical for basic and clinical research applications. Our tertiary cancer-focused biorepository, operating within the auspices of the Rajiv Gandhi Cancer Institute and Research Centre, has received numerous breast tissue specimens, comprising both tumor and distant normal tissues, serving as a cornerstone for breast cancer research endeavors. Following the dissemination of these specimens to researchers, feedback has underscored a notable divergence in the quantity and quality of RNA extracted from tumor tissues, compared to normal tissues. This incongruity poses a significant obstacle to downstream transcriptome and real-time PCR analysis.

Female breast has abundant adipose tissue, which plays a crucial role in breast development and changes throughout a woman's life, including puberty, pregnancy, lactation, and involution. Breast adipose tissue interacts with cancer cells, modifying the tumor microenvironment and potentially promoting cancer progression [1, 2]. At a cellular level, nucleic acids extracted from normal breast tissues often contain nucleic acids from surrounding adipocytes, endothelial cells, and immune cells, with a low representation of ductal elements [1]. Pathology best practices must be ensured while selecting negative control sections from the same patient, the rationale for using surrounding breast sections as negative controls is for baseline comparison, Differential Expression Analysis, and validation of findings. Surrounding breast tissue typically contains normal, non-cancerous cells, by comparing the transcriptome of tumor samples with that of adjacent normal tissue, researchers can establish a baseline for gene expression, helping to highlight changes specifically associated

Received: 14 August 2024. **Revised:** 21 October 2024. **Editorial decision:** 22 October 2024. **Accepted:** 1 November 2024

with tumorigenesis. This approach also allows for identifying differentially expressed genes (DEGs), that may contribute to cancer progression. The DEGs are particularly relevant in studies of various breast cancer subtypes, such as triple-negative breast cancer (TNBC), where understanding the differences in gene expression can provide insights into tumor behavior and potential therapeutic targets. Ensuring the quality of RNA is paramount for accurate gene expression and transcriptome analysis, as intact RNA minimizes biases in the measurement of expression levels [3]. Negative controls are essential for validating results. If some genes are found to be upregulated or downregulated in tumor samples compared to adjacent normal tissue, these findings will be more confidently attributed to the pathological state rather than technical artifacts or variations in sample processing [4]. In breast cancer research, it is imperative to confirm the accuracy of the gene expression data collected to reflect the *in vivo* state of the samples. RNA Extraction poses significant challenges being more labile than DNA [5]. The efficacy of RNA-based analysis dramatically depends on the purity and integrity of the RNA, with various pre-analytical factors, such as tissue type, cold ischemia time, sampling sites handling and processing procedures playing a crucial role in maintaining this integrity [6]. The varied composition of adipose, stromal, and epithelial tissue within the breast is integral in defining the tissue's unique characteristics [7]. However, excess adipose tissue in the breast can be an interfering factor, emphasizing the importance of precise sampling methods to ensure accurate and impactful results [8]. Also, the quality of extracted nucleic acids can drastically affect outcomes. Therefore, ensuring the suitability of the normal breast tissues before succumbing it to RNA extraction becomes imperative. One of the methods to confirm the suitability of normal breast tissue sections as negative controls is to perform immunohistochemistry using markers such as cytokeratin's, myoepithelial cell markers (e.g. p63 or calponin), estrogen receptor, progesterone receptor, or HER2 receptor [9]. However, this approach can be costly and labor-intensive, making it impractical to apply in a high-throughput biobank setting before storing all normal breast tissue samples. To address this challenge comprehensively, we have embarked upon an initiative to conduct in-house isolation of RNA, prospectively from distant normal breast tissue specimens from random and targeted region of distant breast to compare both qualitative and quantitative outcomes of extracted RNA and performing histopathological confirmation. Both the samples are procured from the quadrant opposite to the tumor quadrant 5–10 cm away to avoid any discrepancy due to field cancerization effect.

Refer to Fig. 1 for a detailed overview of the entire experimental procedure.

Materials and methods

Study design and demographic details

We assessed and compared two sample procurement methods from mastectomy specimens to overcome the common issue of low RNA yield from surrounding breast tissue. The objective was to explore the variations in RNA enrichment between employing a random sampling technique and a targeted sampling approach. The study involved collecting tissue samples from surrounding breast tissue, focusing on a group of 15 female subjects diagnosed with primary breast cancer, including TNBC, who underwent Modified Radical Mastectomy (MRM) (Table 1). Notably, none of the patients had undergone chemotherapy or radiation treatment before the surgery.

Clinicopathological features of breast cancer patients

The study included female patients ($n = 15$) aged between 34 and 78 years, all diagnosed with infiltrating ductal carcinoma (IDC). The disease was unifocal in presentation and affected patients across premenopausal ($n = 6$), perimenopausal ($n = 2$), and postmenopausal ($n = 7$) groups. The histological subtypes were classified as Luminal A (ER+PR+HER2−), triple-negative breast cancer (ER−PR−HER2−), HER2-positive (ER−PR−HER2+), and triple-positive (ER+PR+HER2+). Cancer staging was carried out in accordance with the American Joint Committee on Cancer (AJCC) Cancer Staging Manual, 8th edition. The patient distribution included those with Stage II and Stage III tumors (Fig. 2), with tumor sizes ranging from pT2 to pT3. Clinical staging varied from IIA, IIB, IIIA, to IIIC (Fig. 2). No evidence of distant metastasis (M0) was observed in any of the patients.

Tissue sample collection

A specialized breast oncopathologist from the Rajiv Gandhi Cancer Institute and Research Centre in New Delhi collected research tissue samples. The MRM specimen was first inked on the deep surgical margin. Parallel/Bread loafing of the breast specimen was carried out using a 14'long dissecting knife. (Fig. 3) The tumor was identified, measured, and the quadrant noted. Tumor sections were taken. The opposite quadrant was then located for sampling of normal breast tissue. New scalpel blades were then inserted onto the blade holder and normal tissue was collected. This step ensured that there would be no tumor contamination in the sample collected.

The sampling process involved visually selecting a fat pocket (random sample) (Fig. 3) and selecting tissue from grey-white fibromuscular streak areas (targeted sample) (Fig. 3), all from a quadrant opposite to the tumor-bearing quadrant. The samples were collected using blunt heavy dissecting forceps and a sterile surgical blade size 10. Both random and targeted samples were then sent to the Basic Research Lab for RNA extraction. Additionally, mirror samples were fixed in a 10% neutral buffered formalin solution for 8–24 h and were sent to histopathology lab for further processing. The formalin-processed specimens were processed into FFPE Blocks, and HE stained slides were prepared for examination.

RNA extraction protocol

The samples from random and targeted regions were processed for total RNA extraction, enabling subsequent comparative analysis. Approximately 100–150 mg of tissue was utilized for each extraction. Following the manufacturer's guidelines, the Qiagen All Prep DNA/RNA/miRNA Universal Kit (Cat No. 80224) was employed. This column-based method is specifically designed to extract complex tissues that are difficult to lyse, ensuring high-quality RNA suitable for downstream applications.

Tissue specimens were minced using a scalpel blade, followed by lysis employing the buffer provided in the kit. The lysates were then incubated for 2–2.5 h to facilitate adequate lysis. Subsequently, treatment with DNase I and Proteinase K was administered. RNA concentration (ng/μl) and purity (determined by the 260/280 and 260/230 absorbance ratio) was assessed for each sample using a Nano drop spectrophotometer (Thermo-Scientific, Wilmington, DE). All RNA samples were subsequently stored at −80°C until further processing.

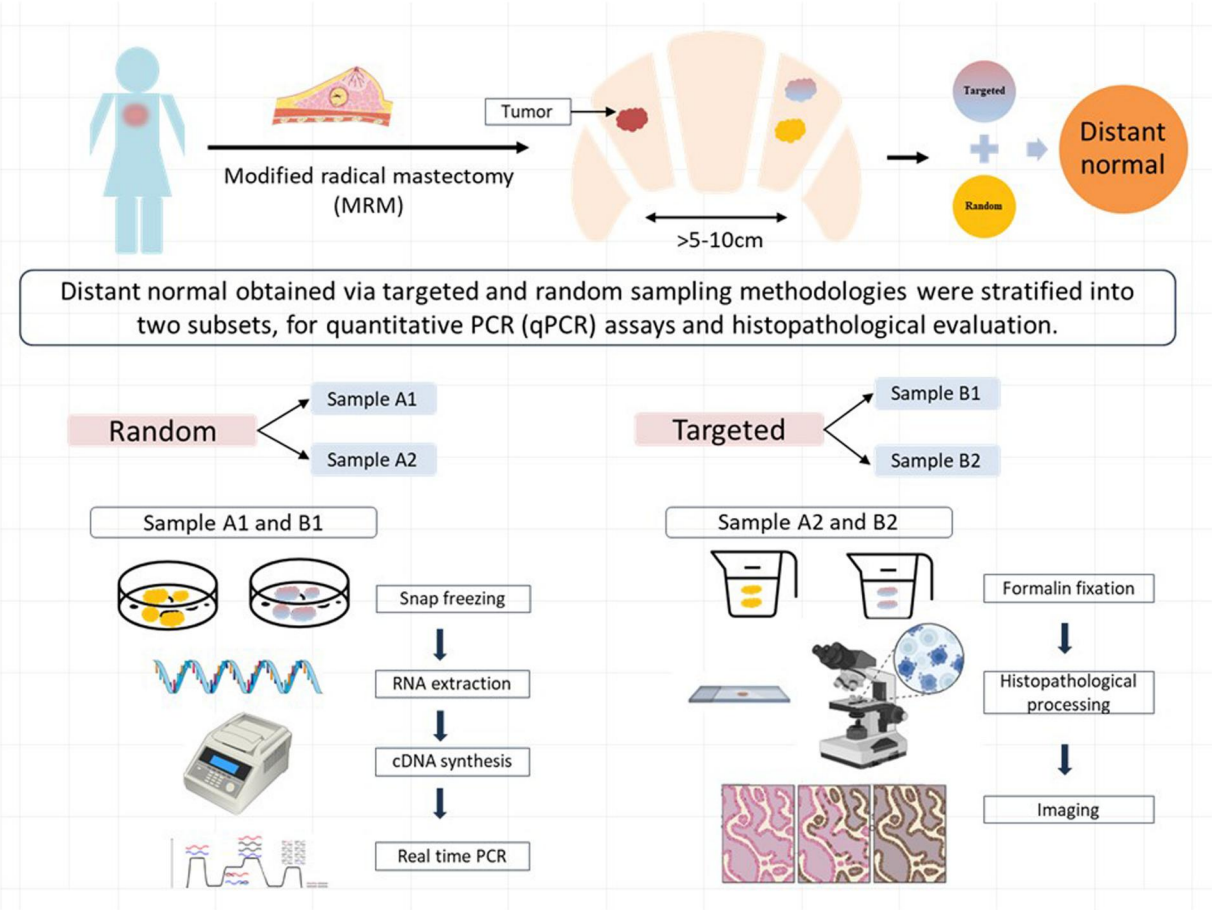

Figure 1. Schematic workflow of sample collection from surrounding breast tissue, followed by stratification into two major groups to compare the RNA integrity obtained from two different sampling sites. Some icons of these illustrations are taken from bio render

Agarose gel electrophoresis

The present study used agarose gel electrophoresis to check isolated RNA's integrity, size, and purity from random and targeted normal breast tissue. A 1% agarose gel was prepared to detect 18S and 28S rRNA.

Agilent tape station

For comparative analysis of RNA integrity by an automated bioanalytical instrument, we utilized the Agilent technologies 4150 Bioanalyzer, along with RNA Screen Tape, following the manufacturer's protocols for total RNA sample preparation. The RNA Integrity Number (RIN) and concentration were recorded automatically by the Bioanalyzer, ensuring accurate and reproducible assessments of RNA quality across samples cDNA and real-time PCR.

To validate the quality of extracted RNA, the samples were converted into cDNA using the High-capacity cDNA Reverse transcription kit (Cat no-4374966 Thermo-fisher scientific).

One microgram of extracted RNA was used for cDNA synthesis. For the real time PCR, 25 ng (A total of 1 µg (1000 ng) of RNA was used in a 20 µl cDNA synthesis reaction. Subsequently, 0.5 µl of the resulting cDNA, equivalent to 25 ng, was utilized for the qPCR reaction) of cDNA was utilized with primer specific for18S rRNA (18sFp-TTCGGAACTGAGGCCATGAT, 18 s Rp-TTTCGCTCTGGTCC GTCTTG), CCSER2 (CCSER2 Fp GACAGGAGCATTACCACCTCAG, CCSER2 Rp- CTTCTGAGCCTGGAAAAAGCGC), RPL13A-(RPL13 Fp-GAGGCCCCTACCACTTCC, RPL13A Rp- AACACCTTGAGACGG TCCAG) and SYBR Green dye was used for fluorescence detection. The real-time PCR for each RNA sample was performed in triplicate. No-template control (NTC) was included in all the RT-PCR reagents except the RNA template. Minus reverse transcriptase (–RT) was used to rule out cross-contamination from genomic DNA as well as from reagents and surfaces.

Histopathology processing and microscopic examination

Formalin-Fixed Paraffin-Embedded (FFPE) blocks were prepared from distant breast samples obtained by random and targeted sampling using the Leica TP1020 Automatic Benchtop tissue processor (Histocore pegasus). The grossed tissue bits were first fixed in 10% Neutral Buffered formalin overnight and dehydrated by immersing the tissue in different ethanol concentrations (50%, 70%, 80%, 90%, 95% (v/v) and absolute ethanol.) for 4.5 h at 35°C. Next, xylene was used as a clearing agent for 2 h at 35°C. Finally, samples were paraffin embedded at 58°C. FFPE blocks were trimmed, and 3 um thick sections were mounted on slides, automatically stained with Hematoxylin and Eosin (H&E) following the H&E staining protocol of Medite Multistainer TST 44C. Histopathology samples were processed and stained at the Department of Histopathology, Rajiv Gandhi Cancer Institute and Research Centre. Reporting of H&E-Stained sections was performed at Biorepository, Rajiv Gandhi cancer Institute and Research Centre.

Table 1. Assessment of random samples RNA yield and the number of terminal duct lobular units on histopathology HE stained slides.

Sample ID	random sample RNA yield (conc.)	260/280	260/230	TDLU's Identified in submitted section on HE
1	59.9	1.80	0.28	No TDLU's identified. Only adipocytes with dilated lymphovascular channels. Occasional lymphocytes.
2	82.8	1.84	0.37	1–2 TDLU's identified. Only adipocytes with dilated lymphovascular channels. Occasional lymphocytes.
3	40.6	1.46	0.91	No TDLU's identified. Only adipocytes with dilated lymphovascular channels. Occasional lymphocytes
4	77.3	1.65	0.91	1–2 TDLU's identified Only adipocytes with dilated lymphovascular channels. Occasional lymphocytes.
5	43.8	1.71	1.04	No TDLU's identified. Only adipocytes with dilated lymphovascular channels. Occasional lymphocytes.
6	73.2	1.80	0.28	No TDLU's identified. Only adipocytes with dilated lymphovascular channels. Occasional lymphocytes
7	70.8	1.63	0.98	1 TDLU identified. Only adipocytes with dilated lymphovascular channels. Occasional lymphocyte seen.
8	40.6	1.95	0.68	No TDLU's identified. Only adipocytes with dilated lymphovascular channels. Occasional lymphocytes.
9	29.4	1.84	0.51	No TDLU's identified. Only adipocytes with dilated lymphovascular channels. Occasional lymphocytes.
10	36.5	1.70	0.43	No TDLU's identified. Only adipocytes with dilated lymphovascular channels. Occasional lymphocytes.
11	85.8	1.75	0.51	No TDLU's identified, Adipocytes with occasional intervening lymphocytes noted. No lymphovascular channels noted.
12	70.9	1.64	0.96	1 TDLU identified. Only adipocytes with dilated lymphovascular channels. Occasional lymphocyte seen.
13	73.6	1.80	1.35	No TDLU's identified. Only adipocytes with dilated lymphovascular channels. Occasional lymphocytes.
14	44	1.72	1.01	1-2 TDLU's identified Only adipocytes with dilated lymphovascular channels. Occasional lymphocytes.
15	45.5	1.48	0.77	No TDLU's seen. Many adipocytes and occasional lymphovascular channel noted. No lymphocyte seen.

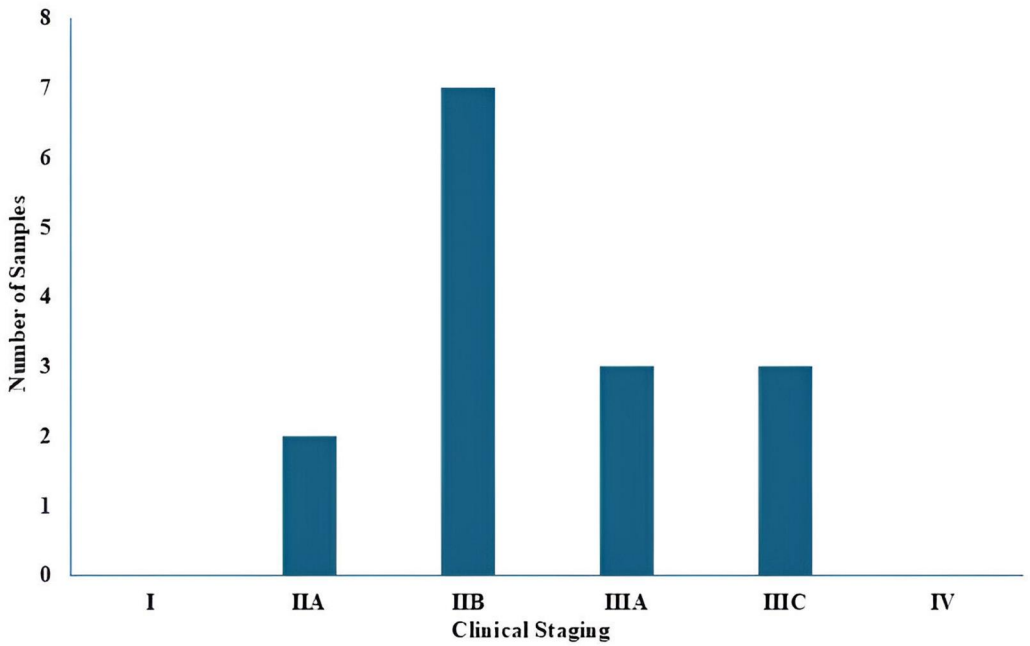

Figure 2. The representative image illustrates the cancer staging of the patients ($n = 15$), with distribution as follows: IIA ($n = 2$), IIB ($n = 7$), IIIA ($n = 3$), and IIIC ($n = 3$)

Results

The study addressed the challenge of low RNA yield from surrounding breast tissue. We strategized sample collection by two different methods and compared the outcomes genomically and histopathologically. A cohort of 15 pilot cases was included into the study for comparison.

Tissue samples were collected from normal breast tissue of female subjects diagnosed with primary breast cancer, who

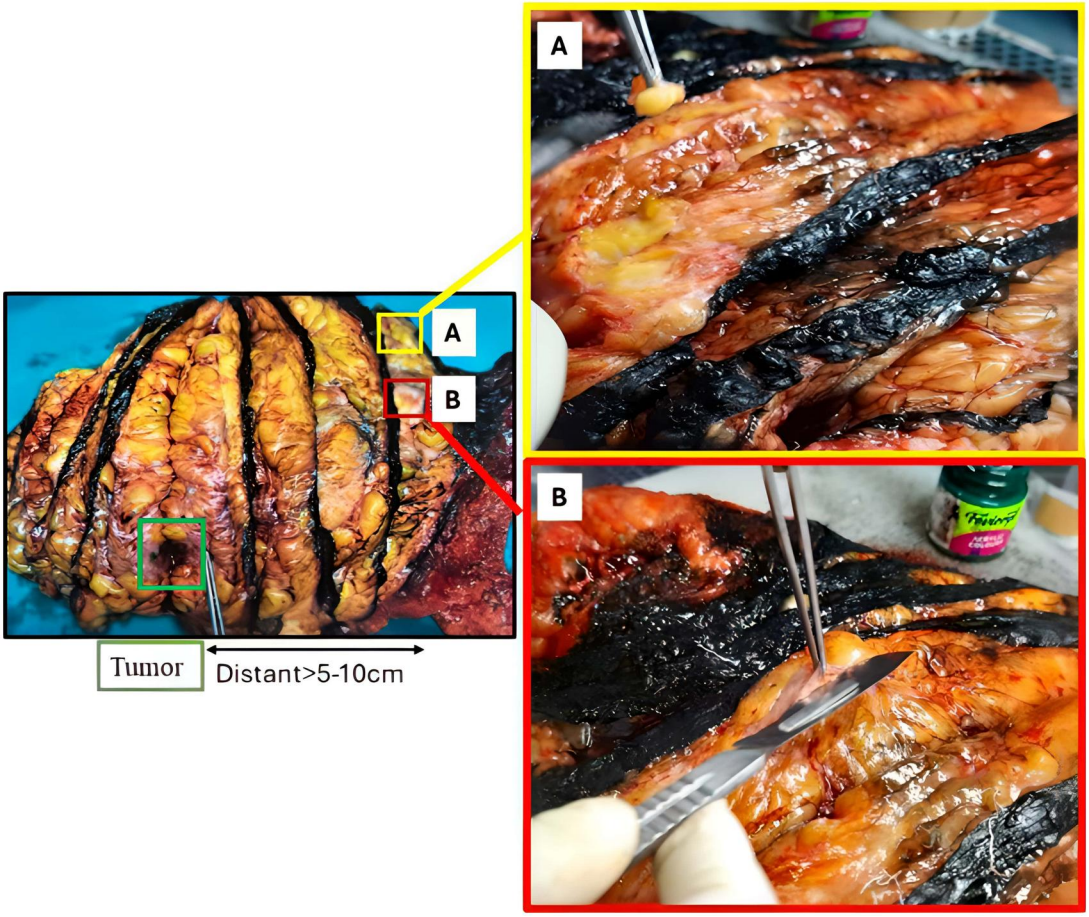

Figure 3. Histopathological sectioning of breast tissue: normal tissue collected from opposite quadrant by random sampling (Inset A) and targeted sampling (Inset B)

underwent MRM. Both random and targeted tissue sampling procedures were employed, with samples acquired from a quadrant opposite to the tumor-bearing quadrant. Matched FFPE blocks were also prepared. RNA was extracted in varying quantities from both random and targeted samples. Subsequently, the epithelial components, represented by the number of terminal duct lobular units (TDLUs) were assessed on histopathology HE stained slides (Tables 1 and 2).

Statistical evaluation

Data normality was evaluated using the Shapiro–Wilk test, while homogeneity of variances was assessed via Levene's test. Subsequent analysis incorporated parametric and non-parametric approaches, including the independent t- test and the Mann–Whitney U test, conducted using SPSS version 29.0.0.0. Statistical significance was established at a *P*-value threshold of less than ($P < .05$). *P*-value outcomes were interpreted to indicate statistically significant difference between sampling sites (Fig. 4A–D).

RNA Quality assessed using gel electrophoresis showed significant bright bands for 18S rRNA and 28SrRNA in all targeted samples. In contrast, randomly collected samples showed very faint bands for 18SrRNA and almost no bands for 28SrRNA (Fig. 5).

Histopathological examination

H&E-stained slides were reported using Nikon E200 Eclipse microscope. H&E-stained sections from FFPE Blocks at low power (10×) showed numerous adipocytes with intervening blood vessels and none-to-occasional terminal duct lobular units (Fig. 6A) in randomly collected tissues. The H&E section from targeted areas showed 3–5 terminal duct lobular units entrapped in a fibro collagenous stroma with few intervening adipocytes (Fig. 6B).

RNA integrity value

In addition to assessing the RNA integrity by traditional methods, Tape station assessment was performed on five random and five targeted samples. Total RNA was extracted from both the target and random regions of breast tissue, and the difference in RNA integrity between the two sample sites was analyzed using the Agilent 4500 Bioanalyzer and the RNA screen tape. During the analysis one outlier is identified via tape station which was excluded from the final analysis (Fig. 7).

Discussion

RNA is essential for various molecular investigations. Extracting RNA from surrounding breast tissue as a negative control poses challenges due to its lower quality and quantity. RNA transcript degradation within cellular machinery occurs at different rates through various mechanisms, and few studies address the stability of nucleic acids, including DNA and RNA, which varies gradually among surrounding breast [10]. Therefore, it is crucial to optimize the sampling site, handling, and other pre-analytical factors for efficient RNA extraction from breast tissue.

Table 2. Assessment of targeted samples RNA yield and the number of terminal duct lobular units on histopathology HE stained slides.

Sample ID	targeted sample RNA yield	260/280 ratio	260/230	TDLU's identified in the submitted section on HE
1	193.5	2.84	1.97	3–5 TDLU's identified. Few adipocytes and lymphocytes identified.
2	205.1	2.02	1.84	3–4 TDLU's identified. Few adipocytes in fibrocollagenous stroma. No lymphocytes seen.
3	234	1.97	1.86	3–5 TDLU's identified. Few adipocytes In fibrocollagenous stroma. No lymphocytes seen.
4	277.6	1.99	1.76	3–5 TDLU's identified. Few adipocytes in fibrocollagenous stroma. Occasional lymphocyte seen.
5	167.5	1.97	1.97	2–3 TDLU's identified. Few adipocytes in fibrocollagenous stroma. Lymphovascular channels noted.
6	87.8	1.81	1.65	2–3 TDLU's identified. Few adipocytes in fibrocollagenous stroma. Lymphovascular channels noted. Occasional lymphocyte seen.
7	158.4	1.88	1.90	2–3 TDLU's identified. Few adipocytes in fibrocollagenous stroma. Lymphovascular channels noted.
8	143.8	2.00	1.86	3–5 TDLU's identified. Few adipocytes In fibrocollagenous stroma. Few lymphocytes seen.
9	117.6	1.87	1.88	2–3 TDLU's identified. Few adipocytes in fibrocollagenous stroma.
10	119.4	1.97	1.65	3–4 TDLU's identified. Few adipocytes admixed with lymphocytes embedded in fibrocollagenous stroma.
11	145	2.03	1.94	2–3 TDLU's identified. Few adipocytes in fibrocollagenous stroma. Lymphovascular channels noted.
12	258.9	2.02	1.85	2–3 TDLU's identified. Few adipocytes in fibrocollagenous stroma.
13	356.4	2.01	1.81	3–5 TDLU's identified. Few adipocytes in fibrocollagenous stroma. Few lymphocytes seen.
14	308.2	1.79	1.49	3–5 TDLU's identified. Few adipocytes in fibrocollagenous stroma. Occasional lymphocyte seen.
15	124.8	1.98	1.76	2–3 TDLU's identified. Few adipocytes in fibrocollagenous stroma.

The abundance of RNA transcripts expressed in the surrounding breast tissue exhibits variability across different sites of normal breast tissue. By examining the associations between genotypic data and phenotypic pathological images, researchers can gain insights into the distribution and abundance of these transcripts within an individual's breast tissue [11]. However, existing studies predominantly focus on tumors and lesion areas, often neglecting the non-diseased surrounding areas of the breast. These non-diseased areas are crucial as they are negative controls in the breast studies. The RNA expression profiles in these regions demonstrate considerable variation, reflecting the inherent heterogeneity of breast tissue. For transcriptome and microarray analysis, the quality and quantity of RNA should be good enough to yield consistent and accurate output [12]. This variability underscores the need for more comprehensive studies on non-diseased areas to fully understand the complexity of RNA transcript expression in breast tissue. Traditionally, assessments of the breast in various contexts such as surgical interventions, molecular biology, and reconstruction have primarily focused on volume, shape, or weight rather than the detailed tissue composition. Research indicates that the breast typically comprises of approximately 70% fat and 30% glandular tissue [13]. One of the objectives of this study is to elucidate how different sampling sites within the breast (target and random) reveal varying composition of TDLU structures. This variation is linked to the nuclear characteristics of cells and transcriptome-wide RNA gene expression. To confirm these associations, the total RNA integrity from two distinct sampling locations was assessed using agarose gel electrophoresis, Agilent tape station, and real-time quantitative PCR analysis. The analysis focused on consistently expressed housekeeping genes, such as 18S rRNA, CCSER2, and RPL13A, which maintain stable expression levels, provided that an equal quantity of cDNA is used in the PCR reaction. Notably, this expression diminishes significantly—up to 15-fold—only when RNA degradation is extensive [14]. This degradation can be measured through cycle threshold (CT) values derived from intact RNA from targeted regions compared to degraded RNA from randomly selected areas of surrounding breast tissue.

The stability of RNA is the main factor influencing its abundance in nature as, degraded RNA, on the other hand, reduces the number of RNA molecules containing intact amplicons, significantly lowering the expression of DEGs [15]. Using degraded RNA in RNA-Seq may lead to shorter fragmented transcripts and fewer mappable reads. Therefore, maintaining RNA integrity is crucial for accurate measurement and gene expression analysis, as degraded RNA can lead to erroneous interpretations of transcript abundance and gene expression profiles. Quality and quantity assessment of RNA is an additional method to obtain minimal amount to conduct gene expression studies. Nevertheless, it is not always possible to measure 18S and 28S ribosomal RNA bands or perform qPCR so it remains to be determined that the amplified mRNA can reliably be used as a control for future studies. This methodology promotes an unbiased approach for selecting samples from the surrounding breast as controls, where the likelihood of encountering degraded RNA is lower.

Several studies have demonstrated that gene expression signatures can classify patients into high- risk or low-risk subgroups [16]. Accurate transcriptome profiling, utilizing true negative controls, enhances patient selection and mitigates the risk of both overt and undertreatment, ultimately reducing patient mortality [16–18]. Using surrounding breast tissue as negative controls in transcriptomic studies significantly enhance our understanding of the molecular alterations associated with breast cancer. This methodology not only aids in the identification of potential biomarkers for diagnosis and treatment but also contributes to a more profound comprehension of the tumor microenvironment and its role in cancer progression. By incorporating surrounding breast tissue which yields good quality RNA, as negative controls,

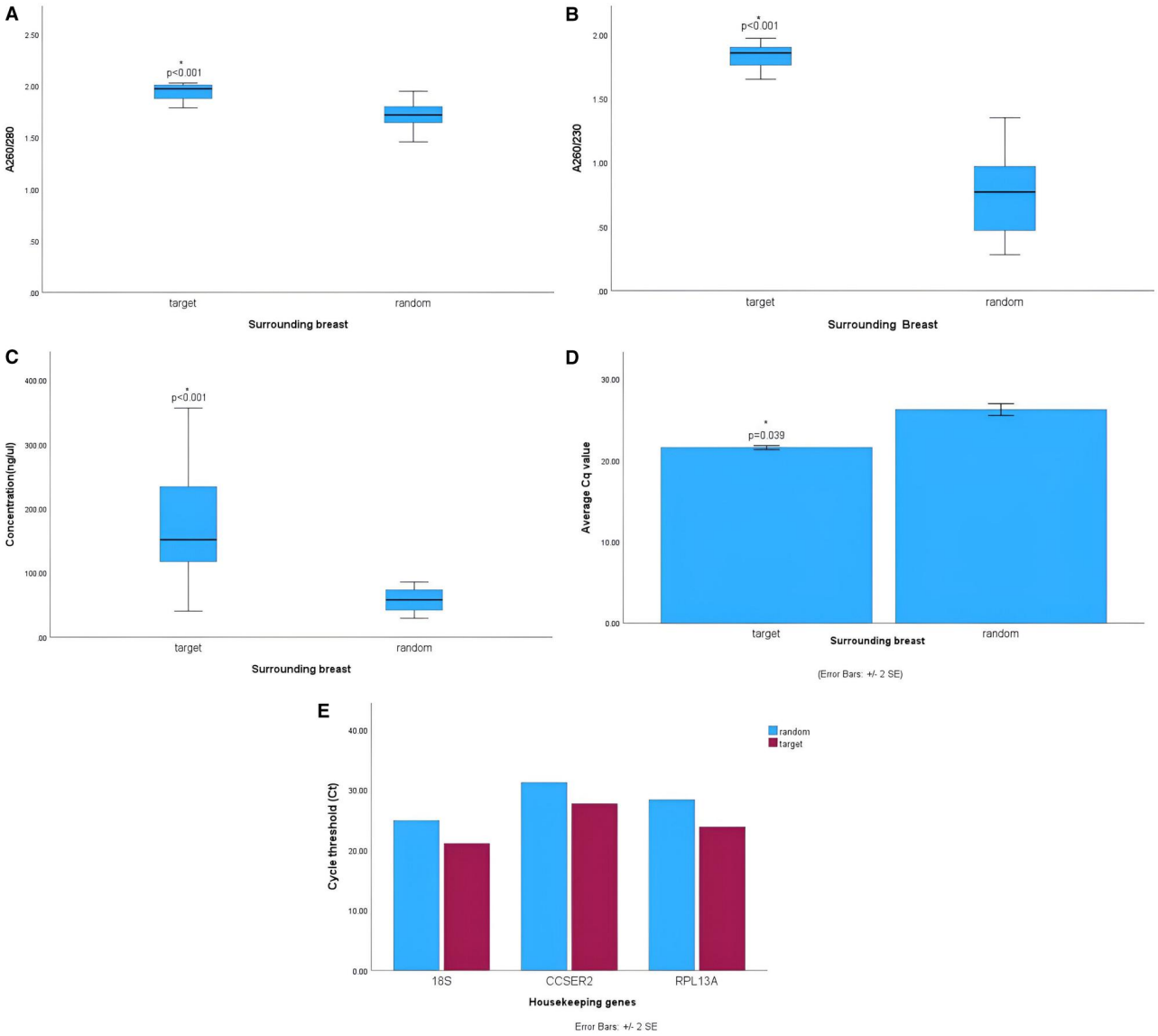

Figure 4. (**A**) The Kruskal–Wallis H test revealed a significant difference in A260/280 values between the target and random groups, H (1) = 17.764, $P < .001$. The target group had a higher mean rank (22.27) compared to the random group (8.73), indicating that the target group had significantly higher A260/280 values than the random group. (**B**) An independent samples t-test revealed a significant difference in A260/230 values between target (M = 1.81, SD = 0.13) and random (M = 0.73, SD = 0.32), t (18.59) =11.96, $P < .001$. The large effect sizes (Cohen's $d = 4.37$, Glass's delta = 3.34) indicate target group A260/230 value was significantly higher than random group. (**C**) Comparison of target and random groups revealed a significant difference in mean values ($P < .001$) with Group 1 (M = 193.20, SD = 78.69) showing higher values than Group 2 (M = 58.31, SD = 18.91). The effect size, calculated using Cohen's d, was found to be 2.36 (95% CI: 1.40–3.29), indicating a large effect. Additionally, Glass's delta was 7.13 (95% CI: 4.42–9.83), further confirming target group has statistically significant concentration values than random group. (**D**) An independent samples t-test is conducted to compare the threshold cycle of 18S mRNA between the target group and the random group. Levene's test for equality of variances is significant, $P = .039$, indicating that the assumption of equal variances is violated. Therefore, the t-test results assuming unequal variances are used. The average Cq value for the target group (M = 21.58, SD = 0.38) is significantly lower than for the random group (M = 26.25, SD = 1.14); t (10.94) = −12.32, $P < .001$, 95% CI [−5.50, −3.83]. The effect size, measured using Cohen's d, is 0.85, indicating a large effect size. This suggests that the target group has a significantly lower cycle threshold compared to the random group. (**E**) This graph illustrates the average threshold cycle (Ct) of 18 s, CCSER2 and RPL13A mRNA isolates from target and random region of breast cancer tissue

researchers can achieve more accurate and reliable gene expression profiles, facilitating the development of more effective therapeutic strategies and improving patient outcomes.

The findings from this study have several implications for future breast cancer research and clinical applications. The study emphasizes the challenges in RNA extraction and the significance of targeted sampling methods in enhancing both the quantity and quality of RNA extracted from distant breast tissue. This advancement is crucial for ensuring accurate gene expression studies and molecular analyses in breast cancer research underscores the critical role of biobanks in managing biological specimens for translational research. Proper handling of tissue samples is essential for preserving RNA, DNA, and proteins, which are vital for advancing cancer research and personalized

therapy. The study also underscores the significance of meticulous sampling techniques to avoid contamination of tumor RNA by distant or adipose tissue RNA, which could impact research and clinical outcomes.

Limitations of the study

The study is limited by a small sample size, which may limit the generalizability of the findings to a broader population. It also focuses on only one subtype of breast cancer, potentially overlooking variations in RNA yield and quality in different subtypes or stages of breast cancer, thus, lacking diversity. A broader perspective would have been provided if different RNA extraction methods had also been compared. A more extensive study on a more significant number of diverse breast cancer samples can enhance its robustness and applicability in the field of breast cancer research.

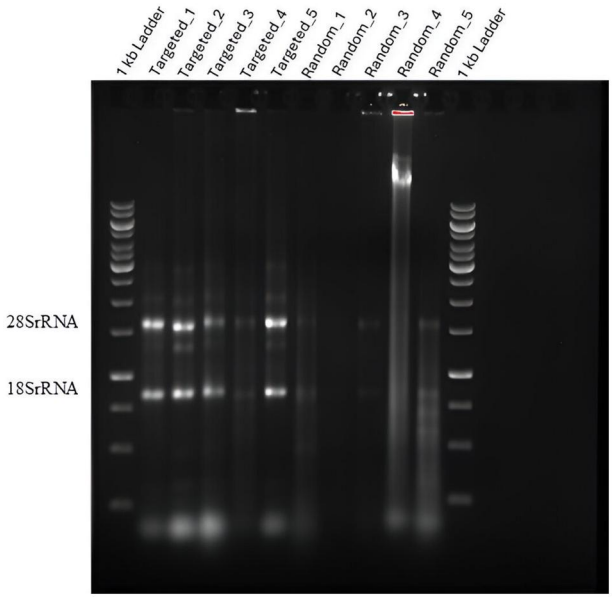

Figure 5. The representative image of Agarose Gel electrophoresis showing targeted and random samples

Conclusion

In conclusion, this study successfully addressed the challenge of low RNA yield from nearby distant breast tissue by employing and comparing two histopathological sample procurement methods: random sampling and targeted sampling. The study involved thorough procedures, such as random and targeted sampling techniques, to ensure diverse representations of surrounding breast tissue. To mitigate the impact of intra-tumoral heterogeneity, we employed a meticulous process of sampling site identification, tissue fragmentation, snap-freezing, and preservation of RNA later before storage at −80°C. Subsequent histopathological assessments of FFPE blocks provided valuable insights into the tissue characteristics. The study's results provide valuable insights for future breast cancer research endeavors, offering guidance on optimizing RNA extraction techniques, employing targeted sampling approaches and enhancing molecular analyses. These insights are essential for advancing our understanding of breast cancer pathogenesis, treatment strategies, and the development of personalized medicine applications.

Acknowledgements

Gratitude is extended to the research scientists at Biorepository for donor recruitment and consent, Ms. Ruchika John, our research Histopathology technician- in charge, for histopathology processing.

Author contributions

komal Mehta (Conceptualization [equal], Data curation [lead], Formal analysis [lead], Investigation [lead], Methodology [lead], Project administration [equal], Resources [equal], Supervision [equal], Validation [equal], Visualization [equal], Writing—original draft [equal], Writing—review & editing [equal]), Archana Sharma (Resources [lead], Visualization [lead], Writing—review & editing [supporting]), Anurag Mehta (Conceptualization [equal], Project administration [equal], Resources [lead], Supervision [supporting], Validation [supporting], Visualization [supporting], Writing—review & editing [equal]), and Juhi Tayal (Conceptualization [lead], Data curation [supporting], Formal analysis [supporting], Funding acquisition [lead], Investigation

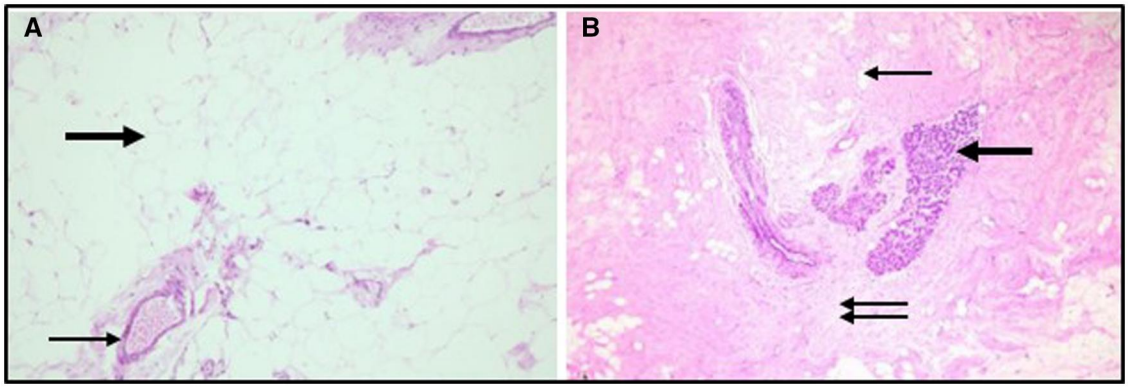

Figure 6. The representative H&E images show the pathological section of breast carcinoma: (**A**) random sampling. H&E-stained section (10× magnification) shows numerous adipocytes (thick arrow) with intervening blood vessels (thin arrow). No terminal duct lobular units identified. (**B**) targeted sampling. H&Estained section (10× magnification) shows a distant terminal duct lobular unit (thick arrow) entrapped in a fibro collagenous tissue (double arrow) with few intervening adipocytes (thin arrow)

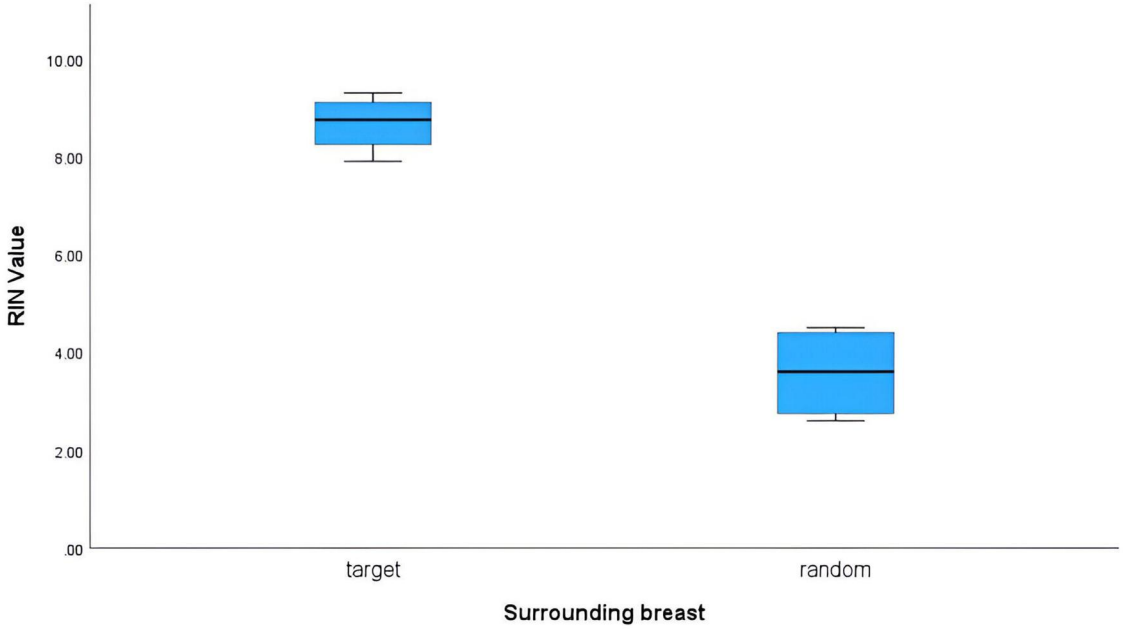

Figure 7. For the same breast tissue samples, the RIN values measured using the Tape Station for the target samples ranged from 8.0 to 9.5, with an average concentration between 160 and 402 ng/μl. In contrast, the RINe values for the random samples ranged from 2.6 to 4.9, with an average concentration between 42 and 150 ng/μl

[lead], Methodology [supporting], Project administration [equal], Resources [supporting], Software [supporting], Supervision [lead], Validation [supporting], Visualization [lead], Writing—original draft [equal], Writing—review & editing [equal])

Conflict of interest statement. None declared.

Funding

None declared.

Data availability

All data are incorporated into the article and its online supplementary material.

Ethics approval and consent to participate

The study was approved by Institutional review board of RGCIRC, Informed consent was obtained from all individual participants included in the study.

Consent for publication

The authors affirm that human research participants provided informed consent for publication of the images in Figures 3 and 6 for utilization of biobank samples.

References

1. Kothari C, Diorio C, Durocher F. The importance of breast adipose tissue in breast cancer. *Int J Mol Sci* 2020; **21**:5760. https://doi.org/10.3390/ijms21165760
2. Somiari SB, Shuss S, Liu J *et al.* Assessing the quality of RNA isolated from human breast tissue after ambient room temperature exposure. *PLoS One* 2022;**17**:e0262654. https://doi.org/10.1371/journal.pone.0262654
3. Samadani AA, Nikbakhsh N, Fattahi S *et al.* RNA extraction from animal and human's cancerous tissues: does tissue matter? *Int J Mol Cell Med* 2015;**4**:54–9.
4. Norollahi SA, Kokhaee P, Pour AR. Comparison of methods of RNA extraction from breast and gastric cancer tissues. *Crescent J Med Biol Sci* 2018; **5**:25–8.
5. Mee BC, Carroll P, Donatello S *et al.* Maintaining breast cancer specimen integrity and individual or simultaneous extraction of quality DNA, RNA, and proteins from Allprotect-stabilized and nonstabilized tissue samples. *Biopreserv Biobank* 2011; **9**:389–98. https://doi.org/10.1089/bio.2011.0034
6. Grizzle WE, Otali D, Sexton KC, Atherton DS. Effects of cold ischemia on gene expression: a review and commentary. *Biopreserv Biobank* 2016; **14**:548–58. https://doi.org/10.1089/bio.2016.0013
7. Nandi Jui B, Sarsenbayeva A, Jernow H *et al.* Evaluation of RNA isolation methods in human adipose tissue. *Lab Med* 2022;**53**:e129–e133. https://doi.org/10.1093/labmed/lmab126
8. Hemmrich K, Denecke B, Paul NE *et al.* RNA isolation from adipose tissue: an optimized procedure for high RNA yield and integrity. *Lab Med* 2010;**41**:104–6. https://doi.org/10.1309/LMFSBPUOA19MH5BV
9. Torlakovic EE, Francis G, Garratt J *et al.* Standardization of negative controls in diagnostic immunohistochemistry: recommendations from the international ad hoc expert panel. *Appl Immunohistochem Mol Morphol.* 2014;**22**:241–52. https://doi.org/10.1097/PAI.0000000000000069
10. Ohashi Y, Creek KE, Pirisi L *et al.* RNA degradation in human breast tissue after surgical removal: a time-course study. *Exp Mol Pathol* 2004;**77**:98–103. https://doi.org/10.1016/j.yexmp.2004.05.005
11. Mou T, Liang J, Vu TN *et al.* A comprehensive landscape of imaging feature-associated RNA expression profiles in human breast tissue. *Sensors (Basel)* 2023;**23**:1432. https://doi.org/10.3390/s23031432
12. Fleige S, Pfaffl MW. RNA integrity and the effect on the real-time qRT-PCR performance. *Mol Aspects Med* 2006;**27**:126–39. https://doi.org/10.1016/j.mam.2005.12.003

13. Vandeweyer E, Hertens D. Quantification of glands and fat in breast tissue: an experimental determination. *Ann Anat* 2002; **184**:181–4. https://doi.org/10.1016/S0940-9602(02)80016-4
14. Ahr A, Karn T, Solbach C *et al.* Identification of high-risk breast-cancer patients by gene expression profiling. *Lancet* 2002;**359**: 131–2. https://doi.org/10.1016/S0140-6736(02)07337-3
15. Antonov J, Goldstein DR, Oberli A *et al.* Reliable gene expression measurements from degraded RNA by quantitative real-time PCR depend on short amplicons and proper normalization. *Lab Invest* 2005; **85**:1040–50. https://doi.org/10.1038/labinvest.3700303
16. van de Vijver MJ, He YD, van't Veer LJ *et al.* A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002;**347**:1999–2009. https://doi.org/10.1056/NEJMoa021967
17. Nelson TR, Cerviño LI, Boone JM, Lindfors KK. Classification of breast computed tomography data. *Med Phys* 2008;**35**:1078–86. https://doi.org/10.1118/1.2839439
18. Strand C, Enell J, Hedenfalk I., Fernö M. RNA quality in frozen breast cancer samples and the influence on gene expression analysis—a comparison of three evaluation methods usingmicrocapillary electrophoresis traces. *BMC Mol Biol* 2007; **8**:38. https://doi.org/10.1186/1471-2199-8-38

Biology Methods and Protocols, 2024, 9, 1–10
https://doi.org/10.1093/biomethods/bpae083
Methods Article