JSR Publications

Available Online

J. Sci. Res. **17** (2), 531-566 (2025)

JOURNAL OF SCIENTIFIC RESEARCH

www.banglajol.info/index.php/JSR

Efficient and Affordable Cancer Screening through Analysis of Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) Gene Expression in Circulating Cell-Free DNA and RNA

H. Ara[1], S. M. S. Al-Din[2], S. A. Tarek[3], M. K. Biswas[4], E. Hoque[1], S. B. Faruque[1], Y. Haque[1], S. M. Sharafuddin[1*]

[1]Department of Physics, Shahjalal University of Science and Technology, Sylhet, Bangladesh

[2]Invent Technologies Ltd, Banani, Dhaka, Bangladesh

[3]American International University-Bangladesh, Dhaka, Bangladesh

[4]Sunamgonj Govt. College, Sunamgonj, Bangladesh

Received 20 November 2024, accepted in final revised form 17 March 2025

Abstract

This research aims to address challenges by proposing a swift and cost-efficient methodology for cancer detection. The study involved 80 (eighty) participants, comprising 40 (forty) diagnosed with cancer and an equivalent number without any known cancer diagnosis, all of whom provided informed consent. The real-time quantitative polymerase chain reaction (RT-qPCR) targeting the human-specific glyceraldehyde 3-phosphate dehydrogenase (GAPDH) gene was employed as a novel method for reliable and precise quantification of circulating cell-free DNA (ccfDNA) in human blood plasma. Additionally, three other housekeeping genes, namely beta-actin (ACTB), ubiquitin C (UBC), and hypoxanthine-guanine phosphoryl transferase (HPRT1), were employed as reference genes for the relative quantification of GAPDH in circulating cell-free mRNA (ccfmRNA). In addition to assessing GAPDH gene expression in ccfDNA, the findings were validated by examining the expression of these three reference genes in circulating cell-free mRNA (ccfmRNA). The results demonstrated a significantly elevated expression ($p < 0.001$) of the GAPDH gene in both ccfDNA and ccfmRNA within plasma samples of cancer patients compared to those individuals without cancer diagnosis. This study underscores the reliability and effectiveness of assessing GAPDH gene expression in ccfDNA using RT-qPCR, establishing it as a promising tool for rapid cancer diagnosis.

Keywords: Cancer screening; RT-qPCR; Circulating cell-free DNA; Circulating cell-free mRNA.

doi: http://dx.doi.org/10.3329/jsr.v17i2.77839 J. Sci. Res. **17** (2), 531-566 (2025)

1. Introduction

Cancer is one of the world's major causes of illness and mortality, offering substantial challenges to public health systems and individual well-being [1–3]. While advances in

* *Corresponding author*: sharif-phy@sust.edu

treatment techniques have increased survival rates for many cancer types, early identification is still critical for attaining favorable results [4–7]. Unfortunately, established cancer screening technologies frequently have drawbacks such as invasiveness, extended timelines, and high financial costs, emphasizing the critical need for novel ways to improve the diagnosis and patient outcomes. Both solid and liquid biopsies serve as essential tools for analyzing cancer biomarkers. Solid biopsies, though effective, are invasive and often painful. In contrast, liquid biopsies from blood offer a non-invasive, painless, and cost-effective alternative, providing valuable insights into disease progression. Various sources are responsible for liquid biopsies, including serum, plasma, swabs, and urine. Several body fluids containing circulating tumor cells (CTCs) and circulating tumor DNA (ctDNA) are examined. Analysis of cell-free DNA derived from blood samples shows the potential to mitigate the lack of cancer screening in marginalized communities. To narrow this disparity, our methodology centers on enhancing the expression of the glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene in cell-free DNA isolated from blood plasma, aiming to lower mortality rates linked to advanced cancer detection.

To evaluate the validity of GAPDH gene expression in circulating cell-free DNA (cfDNA) from blood plasma, it was necessary to perform the relative quantification of the GAPDH gene alongside three other housekeeping genes (ACTB, UBC, and HPRT1). For this, another circulating cell-free nucleic acid, cell-free RNA (cfRNA), in addition to cfDNA quantification analysis, was also considered. Circulating cell-free DNA (ccfDNA) and mRNA (ccfRNA) are nucleic acid molecules that Mandel and Metais first discovered [8]. These molecules in the human body are released from diverse origins [9] and can be found in bodily fluids such as blood plasma [10], urine [11], cerebrospinal fluid [12], and pleural fluid [13]. The predominant source of ccfDNA in the bloodstream is blood stem cells [10], which are found in both blood and bone marrow and are made up of small DNA fragments generated by cells during processes such as apoptosis and necrosis [10,14], Trauma, and surgery [15,16] and potential exogenous source during immune responses or metabolic digestion [17,18].

The importance of this compound in clinical settings has been highlighted as scientists have noticed differences in its levels between healthy individuals and those who are ill [8]. Several scientific studies have identified ccfDNA as a crucial molecular marker for cancer detection [19-21]. Many studies have shown that most ccfDNA molecules are relatively short, around 165 base pairs (bp), although longer fragments exist [22]. Numerous investigations have explored different types of cancer and their levels of circulating cell-free DNA (ccfDNA) [20,21,23-28]. The collective findings of these studies provide strong evidence that individuals with cancer have significantly higher levels of ccfDNA in their blood plasma compared to those without cancer. Our research was comprehensive, covering all types of cancer randomly rather than focusing on one type.

Circulating cell-free nucleic acid can convey genetic information [29] about the cells' physiological or pathological condition, making it important for cancer diagnosis and surveillance [9]. The analysis of circulating cell-free DNA (ccfDNA) often focuses on non-invasive detection of mutations that might lead to chemo-resistance, therapy response, and

disease monitoring in cancer patients [30,31]. The clinical usefulness of ctDNA in metastatic settings includes monitoring tumor evolution, understanding treatment resistance mechanisms, and guiding decisions to switch anticancer therapies [32].

Similarly, the clinical utility of circulating cell-free mRNA (ccfRNA) remains inadequately defined up to this point, numerous studies have shown that circulating cell-free mRNAs (ccfmRNAs) serve as biomarkers for diagnosing various cancers, including those affecting the breast, prostate, pancreas, colon, thyroid, and skin [33-35]. ccfmRNA can be passively released from cells engaged in necrosis and apoptosis, or actively secreted via membrane-bound microvesicles and exosome signaling. These circulating cell-free RNA molecules are not associated with intact cells and originate from various cell types [24,36]. Changes in circulating cell-free messenger RNA (ccfmRNA) profiles, particularly in cancer, reflect changes in gene expression patterns and provide insights into illness onset and progression. A recent review paper demonstrated that circulating cell-free nucleic acids can serve as biomarkers for cancer genotyping and immune phenotyping of the tumor microenvironment [37]. These nucleic acids can convey genetic information about cells' physiological or pathological condition, making them essential for cancer diagnosis and surveillance [38]. Since ccfDNA and ccfRNA are present in the circulation and reflect physiological and pathological processes occurring in the body, including the presence of malignant cells, there is reason to investigate them as potential biomarkers for cancer diagnosis [9].

The choice of the GAPDH gene as the target biomarker for cancer detection is crucial to this study. The housekeeping gene GAPDH has become more and more prominent as a possible cancer biomarker since it is dysregulated in several different forms of cancer [39]. The GAPDH gene produces the GAPDH enzyme, which is essential in glycolysis, the metabolic route that transforms glucose into pyruvate to provide energy in the form of adenosine triphosphate (ATP) [40]. GAPDH catalyzes glycolysis's sixth stage, which involves oxidizing glyceraldehyde-3-phosphate to 1,3-bisphosphoglycerate and reducing NAD^+ (oxidized Nicotinamide Adenine Dinucleotide) to NADH (reduced form of Nicotinamide Adenine Dinucleotide) [41,42]. In addition to glycolysis, GAPDH is involved in DNA repair, RNA transport, and apoptosis [43,44]. It also serves as a housekeeping gene, meaning it is expressed in most cells at generally consistent levels, and is used as a reference gene in gene expression studies to normalize mRNA levels [39,45]. It is becoming more well-acknowledged that the first discovered glycolytic enzyme GAPDH, which was involved in glucose metabolism, is frequently employed as a housekeeping gene for experimental controls in protein, mRNA, and DNA research, has since been shown to carry out several tasks outside energy metabolism [46]. Recent data indicates that some cancer cells have dysregulated GAPDH expression, pointing to wider functions in cellular functions. Elevated GAPDH expression is linked to increased glycolytic activity, which encourages tumor growth. There is growing evidence that the survival strategies of cancer cells in various cancer types, including hepatocellular carcinoma, are linked to the overexpression of GAPDH [41,47]. GAPDH is essential for metabolism and gene transcription, among other cellular processes [48]. Several studies show that GAPDH plays

a major role in cancer metastasis and that malignancies like colon cancer are associated with increased expression of this gene [50]. For example, GAPDH contributes to multiple carcinogenic pathways in hepatocellular carcinoma and is responsive to novel anticancer medications [49]. A recent review emphasized how GAPDH expression is dysregulated in several malignancies, including lung cancer [51], renal [52], breast [53,54] glioma [12], bladder [55], and prostate cancer [56,57]. GAPDH is commonly overexpressed in various cancers, linked to poor prognosis, immune cell infiltration, and immune checkpoint expression, and is regulated by DNA copy number, and methylation [12]. GAPDH mRNA levels correlate with metastatic potential and cell motility in prostate cancer cells by enhancing androgen receptor transcriptional activity [56].

GAPDH is a possible therapeutic target because of its complex role in controlling the fate of cancerous cells, which includes both promoting tumor progression and regulating cell death. GAPDH is closely regulated at both the transcriptional and posttranscriptional levels, while its dysregulated expression and other cellular roles are still poorly understood. Cancer-related variables like insulin, HIF-1, p53, nitric oxide (NO), and acetylated histones affect GAPDH expression and protein activities through common pathways. Furthermore, GAPDH undergoes posttranslational changes (PTMs) in cancer cells, which result in novel functions unrelated to its initial glycolytic function [58].

GAPDH is involved in various cellular activities beyond cancer, notably in diseases like diabetic retinopathy (DR). Shivashankar *et al*. (2021) found that while hyperglycemia or cytokine exposure alone did not affect GAPDH and glutamine synthetase (GS) activity, a decrease was observed in a DR model with combined factors [59]. Further analysis showed GAPDH inhibition increased glucose levels, and cytokines raised lactate and ATP levels while reducing glutamate levels. Additionally, GAPDH can modulate immune responses in allergic conditions, such as in acute lung injury (ALI), by suppressing harmful macrophage activities and altering cytokine production to reduce inflammation [60].

Given the high level of GAPD-related sequences, it is crucial to avoid genomic DNA amplification when using GAPDH as an endogenous control in mRNA quantification. We have developed a pair of GAPDH primers designed to prevent genomic DNA amplification in real-time reverse transcription PCR applications using BRYT-Green Dye. These newly designed primers amplify only a 75 bp DNA sequence. Since the circulating cell-free DNA (ccfDNA) molecules extracted from blood plasma are small fragments, typically around 170 bp, and are obtained using a specialized ccfDNA extraction kit, there was no risk of genomic DNA amplification in our study. Although we acknowledge the potential for a significant portion of GAPDH-related sequences to result from DNA amplification, it is important to note that the GAPDH gene has been successfully utilized in numerous instances for cancer detection [51-57].

In numerous studies, real-time quantitative PCR has been employed to investigate the diagnostic and prognostic implications of various biomarkers [61]. In one study, the real-time quantitative fluorescence polymerase chain reaction (FQ-PCR) method was used to measure human telomerase reverse transcriptase (hTERT) DNA levels in patients with hepatocellular carcinoma [62]. The quantitative real-time polymerase chain reaction

method (qRT-PCR) was used to detect biomarker for gastric cancer [63], ovarian cancer [64], and colorectal cancer [65].

The objective of this work was to measure the levels of GAPDH ccfDNA copy number in plasma samples taken from cancer patients and non-cancer patients by using the RT-qPCR method. A comparative investigation of its expression in different groups has been conducted to confirm GAPDH's value as a biomarker for easy and rapid diagnosis of cancer and evaluate its potential for incorporation into standard screening procedures. By using the capabilities of RT-qPCR analysis of ccfDNA and ccfRNA, this study also aims to address the urgent need for quick, affordable, and non-invasive cancer detection techniques. The outcomes of this examination could have a substantial impact on cancer monitoring and treatment approaches, as well as the overall cancer detection paradigm.

2. Materials and Methods

2.1. *Patients*

This research project enlisted the involvement of 40 individuals diagnosed with cancer from the National Institute of Cancer Research and Hospital (NICRH), Bangladesh. Blood samples were collected randomly from cancer patients, ensuring a diverse representation of various cancer types, ages, and genders. The demography of patients is described in Table 1. A cohort of 40 non-cancerous samples originating from various regions of Bangladesh was included in the study to establish the baseline levels of plasma ccfDNA. Several criteria were considered in the selection of the control group for our investigation, apart from their absence of cancer history: Physiological Condition: Verifying that the controls do not suffer from chronic conditions or diseases that could influence the research outcomes, such as diabetes, cardiovascular disorders, or autoimmune ailments; Behavioral Patterns: Considering variables like alcohol and tobacco consumption, level of physical exercise, and smoking habits to correspond with the characteristics of the patient cohort; Pharmaceutical Intake: Ensuring that none of the controls are consuming medications that could impact the variables under study. The demography of the control group is described in Table 2. All patients and controls (non-cancerous people) provided written informed permission. Before the study's start, institutional ethical committee approval was obtained. Both malignant and non-cancerous samples were collected using the same method. The study gained ethical approval from the National Institute of Cancer Research and Hospital (NICRH), Mohakhali, Dhaka-1212, Bangladesh, under the reference number of NICRH/Ethics/2019/525 and was carried out in compliance with the national laws and regulations of the country as well as the ethical precepts outlined in the WMA Declaration of Helsinki. The Bangladesh Council for Scientific and Industrial Research (BCSIR), in Dhaka, acquired this ethical approval of informed consent (ref: NICRH/Ethics/2019/525) to open a genomic research facility.

2.2. *Plasma sample preparation*

Whole blood was obtained and placed into commercially accessible K2EDTA tubes. Subsequently, these tubes were either promptly processed or subjected to a temperature range of 2-10 °C for a brief duration to facilitate any required manipulation. The collected whole blood was centrifuged at 4 °C and at 2000 × g for 10 minutes using a Sorvall ST 8R centrifuge (Thermo Scientific) to isolate the plasma fraction. To preserve the buffy coat integrity, the plasma was carefully aspirated using a pipette. To enhance the guarantee of eliminating any probable white blood cells, a subsequent process of centrifugation was conducted on the plasma at (2000 × g) for 10 more min. Following centrifugation, the supernatant was carefully transferred to a fresh tube and maintained at 2-8 °C throughout the procedure. The obtained plasma samples were stored at -20 °C and for long-term storage maintained at -80 °C temperature.

Table 1. Demography of patients.

SL No	Type of Cancer	No of Patients	Sex (%)	Age (on average)
1	Breast Cancer	14	Female 100	39 ± 9 Years
2	Ovary Cancer	3	Female 100	41 ± 10 Years
3	Rectal Cancer	3	Male 100	36 ± 13 Years
4	Leukemia	1	Male 100	28 ± 0 Years
5	Squamous Cell Carcinoma	5	Female 80 Male 20	51 ± 13 Years
6	Lymphoma	2	Female 50 Male 50	33 ± 5 Years
7	Early Stage	2	Female 50 Male 50	33 ± 7 Years
8	Adenocarcinoma	2	Male 100	40 ± 0 Years
9	Pulmonary & Abdominal Mets	1	Male 100	19 ± 0 Years
10	Osteosarcoma	1	Female 100	40 ± 0 Years
11	Esophagus	1	Female 100	40 ± 0 Years
12	Medulloblastoma	1	Female 100	8 ± 0 Years
13	Stomach Cancer	1	Male 100	60 ± 0 Years
14	High-grade Urethral Carcinoma	1	Male 100	52 ± 0 Years
15	Pleural infusion	1	Female 100	50 ± 0 Years

Table 2. Demography of non-cancer control group.

SL No	Category	No of Subjects	Sex (%)	Age (on average)
1	Students	22	Male 95 Female 5	24 ± 1.3 Years 23 ± 3.2 Years
2	Housewife	2	Female 100	39 ± 4 Years
3	Service Holder	15	Male 67 Female 23	42 ± 6 Years 40 ± 8 Years
4	Child	1	Female 100	4 Years

2.3. *Extraction of ccfDNA*

Circulating cell-free DNA (ccfDNA) was extracted using the Maxwell® RSC ccfDNA Plasma Kit (Promega Corporation, Madison, WI, USA; Cat # AS1480) on a Maxwell® 16 Instrument (Promega Corporation, Madison, WI, USA; Cat. # AS1000. This extraction process adhered to the manufacturer's recommendations outlined in the Maxwell® RSC ccfDNA Plasma Kit technical manual (# TM454) [66–68]. The "ccfDNA Plasma" method of the Maxwell® 16 Instrument was designed to extract ccfDNA with this Kit. The ccfDNA was eluted in 60 µL of the supplied solution and stored at -20°C. Using a NanoDrop™ Spectrophotometer (Model: ND2000, Origin: Scientific, USA) [69] the A260/A280 ratios were measured to determine the purity and quantity of the extracted DNA.

2.4. *Extraction of ccfmRNA from cancerous and non-cancerous blood plasma*

In this study, ccfmRNA were extracted from the blood plasma (cancerous and non-cancerous) by using the specific kit named SV Total RNA Isolation System (Promega Corporation, Cat # Z3100, Origin: USA), following the instructions provided in Technical Manual #TM048. For total RNA extraction from human blood plasma, 100 µL of fresh plasma was combined with 175 µL of SV RNA lysis buffer in the initial step. Then, for the remaining steps, the SV Total RNA Isolation System protocol was followed [70,71]. The A260/280 ratio, ensuring RNA purity, was measured using a NanoDrop™ Spectrophotometer and found to be ≥ 1.9. The synthesized cDNA was derived from the extracted RNA using the complete reverse transcriptase kit named GoScript™ Reverse Transcriptase (Promega corporation, Cat#A5003, origin: USA) following the instructions outlined in the manufacturer's manual (GoScript™ Reverse Transcription System TM136) [72,73]. The resulting cDNA exhibited A260/280 ratios ranging from 1.75 to 1.9, confirming the quality of the synthesized complementary DNA.

2.5. *Quantification of extracted ccfDNA by Quantus™ Fluorometer*

The Quantus™ Fluorometer from Promega, USA (Cat. # E6150) [74,75] was used to conduct a fluorescence test to gauge the quantity of ccfDNA present. The analysis followed the manufacturer's recommended protocol. The Quantus™ Fluorometer is specifically designed to function with Promega QuantiFluor® dsDNA Dye Systems [76] and quantifies ccfDNA concentration in ng/µL. For each sample, we conducted three measurements, and the recorded results were averaged to ensure accuracy and reliability. This Fluorometer is a highly sensitive instrument that enhances the reliability and precision of DNA quantification in various research and laboratory applications. The sensitivity of this in measuring Double-stranded DNA (dsDNA) is critical for its performance in accurately quantifying DNA concentrations, particularly when working with limited samples or low DNA concentrations. The Quantus™ Fluorometer utilizes specific fluorescent dyes that bind to dsDNA molecules, allowing for the measurement of fluorescence intensity. The

effectiveness of these dyes in attaching to DNA and the device's capacity to pick up the released fluorescence both affect sensitivity. It has a high signal-to-noise ratio, meaning that the signal from the actual DNA is distinct and easily distinguishable from background noise. This is crucial for accurate and reliable measurements, especially when dealing with minute quantities of DNA.

2.6. *Absolute quantification of ccfDNA by Real Time–Quantitative Polymerase Chain Reaction*

The amount of ccfDNA recovered from the patients and the controls was measured by intercalating dye-based qPCR method using a qTOWER3 thermal cycler (Analytik Jena, Germany) [77,78]. To do this, GAPDH gene amplification was targeted for the absolute quantification of the ccfDNA in the plasma samples, and the results were compared to a standard curve made by taking measurements of DNA controls. To create DNA reference standards, we used the Standard Applied Biosystems™ TaqMan™ Control Human Genomic DNA (Applied Biosystem, catalog number: 4312660; concentration set at 10 ng/µL). The levels of circulating cell-free plasma DNA were assessed by employing meticulously standardized dilution curves. These curves were constructed using a known concentration range of human genomic DNA, spanning from 10^5 to 1 pg/µL, and employing a consistent dilution factor of 6. We employed the GoTaq® qPCR Master Mix as our standard PCR settings to generate consistent curves with an average slope of approximately -3.34 (equating to 99.13 % efficiency). This master mix, designed for RT-qPCR, integrates GoTaq® Hot Start polymerase, a precisely formulated buffer, and the vibrant BRYT Green® fluorescent DNA binding dye. This blend significantly improved the reliability, reproducibility, and sensitivity of our PCR experiments. Tailored specifically for instruments capable of detecting SYBR® Green I or FAM™ Dye, it ensured optimal performance. In our RT-qPCR reaction, all samples and standards were prepared using this master mix. Six dilution factors for the DNA standard (Std1 to Std6) were established using Nuclease-free water: Std1 at 100 ng; Std2, Std3, Std4, and Std5 at 10 ng, 1 ng, 0.1 ng, and 0.01 ng respectively. The 100 ng stock solution was appropriately diluted to achieve these concentrations.

To quantify DNA copy numbers, the human-specific primer is utilized targeting GAPDH with the GenBank Accession No. NC_000012.12:6534517-653837. The forward primer sequence was 5' - TGTAGGAGGGACTTAGAGAAGG -3', and the reverse primer sequence was 5'-ACTCAAAGGGCAGGAGTAAAG-3' [79]. Each 20 µL qPCR reaction mixture contained 10 µL of 2X GoTaq® qPCR Master Mix (Cat. # A6001, Promega, USA), 10 pM each of forward and reverse primers, and either 1 µL of the DNA reference standard or 4 µL of the ccfDNA sample. The volume was adjusted with nuclease-free water. The qPCR was conducted on an Analytik Jena qTOWER3 thermal cycler RT-qPCR detection system (USA) with the following cycling conditions: an initial activation at 95 °C for 2 min, followed by 45 cycles of denaturation at 95 °C for 15 sec, annealing at 59 °C for 15 sec, extension at 60 °C for 20 sec, and a subsequent melt analysis with a ramp from 60 °C to 95

°C (gradient 1 °C) [80,81]. Three replicates of each run were generated for all the samples, and a negative control devoid of DNA was included. The machine's Ct values and amplification plots were then used to build a standard curve, which was then used to calculate the amount of DNA in the sample.

2.7. *Relative quantitation of ccfRNA using human-specific GAPDH gene through RT-qPCR.*

In this study, the human-specific GAPDH gene served as the target for accurate assessment of circulating cell-free DNA (ccfDNA). Three additional housekeeping genes-beta-actin (ACTB), Ubiquitin C (UBC), and hypoxanthine guanine phosphoribosyl transferase 1 (HPRT1)—were employed as reference genes for the relative quantification of GAPDH. The same methodology and thermal cycler were utilized throughout the study. The cDNA sample preparation procedure was conducted following the ccfDNA sample preparation procedure for the qPCR method. The cycling conditions for gene expression were identical to those used for the absolute quantification of GAPDH ccfDNA. The characteristics of these three housekeeping genes (reference gene) are mentioned in Table 3.

Table 3. The RT- concentration levels qPCR primer characteristics with Gene bank accession number.

Name	Type	Accession no	Sequence	Length	Tm	Amplicon
GAPDH (target gene)	Forward	NC_000012.12:6534517-6538371	TGTAGGAGGGACTTAGAGAAGG	22 bp	61.92	75 bp
	Reverse		ACTCAAAGGGCAGGAGTAAAG	21 bp	61.89	
ACTB (reference gene)	Forward	NC_000007	GGATCAGCAAGCAGGAGTATG	21 bp	62.36	96 bp
	Reverse		AGAAAGGGTGTAACGCAACTAA	22 bp	62.24	
HPRT1 (reference gene)	Forward	NC_000023	CAGGAGTATGGAGGTTTGGATG	22 bp	62.11	93 bp
	Reverse		GATCCAAAGGCTACGGTGATAG	22 bp	62.1	
UBC (reference gene)	Forward	NC_000012.12:c124914650-124911646	TGAAGACCCTGACTGGTAAGA	21 bp	62.26	99 bp
	Reverse		GAGGGATGCCTTCCTTATCTTG	22 bp	62.18	

2.8. *Standard curve*

A standard curve was created in this experiment using a series of six-fold serial dilutions of human genomic DNA. These dilutions were prepared to encompass a range of concentrations that were known, ranging from 10^5 to 1 picogram per microliter (pg /µL). The specific concentrations included in the dilution series were 10^5, 10^4, 10^3, 10^2, 10, and 1 pg /µL. The plot of the standard curve was generated using the amplification of the GAPDH gene in the DNA standards. Notably, this plot exhibited a high level of correlation with an R^2 value of 0.996, as demonstrated in Supplement 1, Section 1.

2.9. *Statistical analysis*

Statistical inquiries in this study were conducted using MS Excel software. The mean and standard deviation (SD) were the main statistical metrics used to illustrate the findings of

the quantitative analysis of circulating cell-free DNA (ccfDNA). A single-factor ANOVA test was utilized to compare the cancer and non-cancer groups. Additionally, a two-sample t-test, assuming equal variances, was performed to examine the variation in plasma ccfDNA concentration between the two groups. A p-value of 0.05 or less was considered statistically significant in this study. For Log_{10} DNA copy number calculation the equation $X_0 = E_{AMP}$ *(b-Ct) was employed, as detailed in Supplement 1, Section 2. As stated in Supplement 1, Section 3, probability analysis included the normal distribution and descriptive statistics. The discriminatory ability of ccfDNA between the two patient groups was evaluated through logistic regression and ROC curve analysis. For the study on relative gene expression, the expression level of a target gene was compared to three reference genes. This was accomplished by employing $2^{-\Delta\Delta Ct}$ calculation in Excel for a relative fold change analysis. To normalize the reference genes, Norm Finder software was utilized.

3. Results and Discussion

The study included 80 participants, comprising 40 individuals diagnosed with cancer and an equal number without a cancer diagnosis. Two distinct approaches were employed in our investigation to identify significantly elevated expression of the GAPDH gene in both ccfDNA and ccfRNA within plasma samples from cancer patients compared to those without a diagnosis. The GAPDH gene served as the focal point, while three other housekeeping genes (ACTB, UBC, and HPRT1) were utilized as reference genes for relative quantification. Real-time quantitative PCR (RT-qPCR) was the method of choice for this study. Additionally, for the comparative quantification of ccfDNA, the fluorometric method was also employed.

This study focuses on circulating cell-free DNA extracted from human blood plasma, which is utilized for cancer detection. We quantify the expression of the specific GAPDH gene using real-time quantitative polymerase chain reaction (RT-qPCR). It is well established that DNA amplification techniques, like PCR, can often lead to the generation of a high abundance of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) related sequences. This phenomenon arises from the potential for certain sequences to be preferentially amplified over others, potentially resulting in an overrepresentation of GAPDH-related sequences in the amplified DNA pool. Given the high conservation of the GAPD enzyme across different species, primer designs targeting the GAPDH gene in one species may inadvertently bind to and amplify similar sequences from other species. Therefore, maintaining meticulous care during sample preparation and handling is crucial to minimize the risk of contamination and ensure the accuracy of results.

3.1. *Dynamic range of the Real-Time qPCR*

In supplement-1, Table S1, the GAPDH gene assay in the real-time quantitative PCR (RT-qPCR) experiment demonstrated exceptional sensitivity across a wide range of logarithmic DNA copy numbers. This is further illustrated in Fig. S1 (in Supplement 1). The obtained

results confirmed a significant and reliable linear relationship between the cycle threshold (Ct) readings and the logarithm of the DNA copy number (R^2 = 0.996, intercept = 36.37, slope = -3.34, and E_{amp} = 1.99). Notably, all analyzed plasma samples fell within the range established by the standard curve. Comparison with Previous Studies: Our findings align with previous research by Gal et.al (2004) who observed a strong linear correlation (R^2 > 0.99, slope = -3.39) between Ct values and the logarithm of copy numbers for the beta-globin gene when quantifying circulating cell-free DNA from the blood serum of breast cancer patients [83]. In our study, analysis of variance (ANOVA) revealed a statistically significant difference ($p < 0.001$) in the Ct values obtained by the GAPDH gene amplification on ccfDNA among the study groups, as presented in Table S1 (Supplement 1). This result is similar to the findings reported by Singh *et al.* [83].

3.2. *ccfGAPDH DNA copy number analysis by RT-qPCR*

In our research, we employed Real-Time Quantitative Polymerase Chain Reaction (RT-qPCR) for absolute quantification of GAPDH DNA on circulating cell-free DNA in both cancerous and non-cancerous blood plasma. In the first step, we amplified the GAPDH DNA in ccfDNA, extracted from blood plasma, and then quantified the GAPDH DNA copy number following the equation of DNA copy number calculation (in Supplement 1, section 2). The Ct (cycle of threshold) values obtained from RT- qPCR were converted to Log_{10} DNA copy numbers and the resulting data were documented in Table 5. The statistical analysis comparing results between two groups - 40 patients diagnosed with cancer and 40 healthy individuals without a cancer diagnosis - has been summarized in Table 4. Our study revealed elevated levels of the total GAPDH DNA in circulating cell-free DNA (ccfDNA) solution of cancer blood plasma compared to ccfDNA of non-cancerous blood plasma, as illustrated in Fig. 1A. Utilizing RT-qPCR, we observed a significantly increased GAPDH DNA concentration in ccfDNA of cancer patients, ranging from 340.40 to 1570.68 DNA copy numbers per microliter of eluted ccfDNA solution. Conversely, the healthy control group displayed markedly lower values, with a range of 0.74 to 7.22 DNA copy numbers per microliter of eluted ccfDNA solution (Fig. 1A). A significant difference ($P \leq 0.0$) in Log_{10} DNA copy numbers was observed between the two study groups. A comprehensive summary of the data is provided in Table 4. Similar results were found in a study by Zhong *et al*. (2007), who compared the mean concentrations of cell-free DNA (cfDNA) in plasma samples from three groups: patients with breast cancer, patients with benign breast lesions, and normal controls [54]. The mean concentrations were 2285, 1368, and 1489 genome equivalents (GE) per milliliter, respectively. Another study with breast cancer patients showed similar results in a study focusing on breast cancer and quantifying Rt-qPCR [54]. They found that the median serum DNA concentration was 63 ng/mL (ranging from 5 to 456 ng/mL) for control subjects, whereas for patients, it was 221 ng/mL (ranging from 17 to 3325 ng/mL). In another investigation, it was observed that the average levels of ccfDNA were notably higher in ovarian cancer patients compared to controls [84]. In our study, the levels of GAPDH DNA in plasma ccfDNA were notably higher in the cancer group than

the control group. This result was consistent with numerous studies indicating notably reduced concentrations of circulating cell-free DNA in serum and plasma samples from healthy individuals compared to those with breast cancer or other malignancies, including pancreatic cancer, prostate cancer, and various other cancer types [85] An additional study has demonstrated that the presence of extracellular DNA in the bloodstream presents a substantial opportunity for the identification of novel thyroid-related biomarkers.

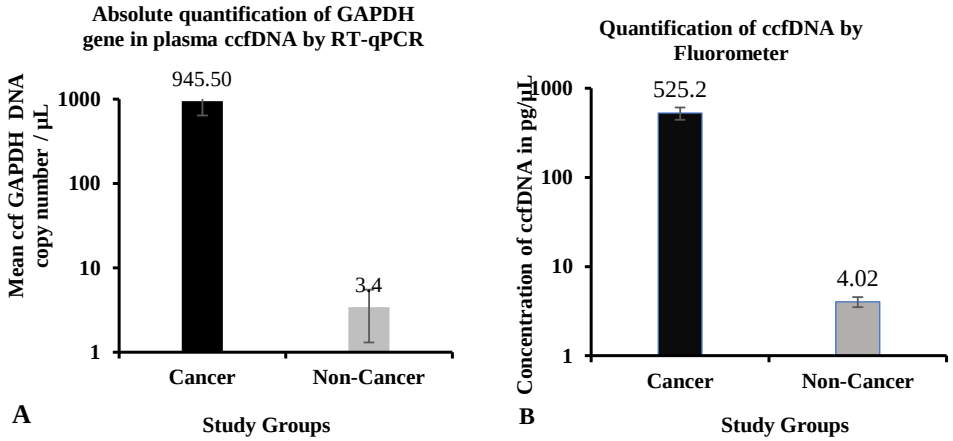

Fig. 1. Results of ccfDNA concentrations in samples extracted from both cancer and non-cancer group; (A) GAPDH DNA copy number in per microliter eluted ccfDNA solution as per RT-qPCR measurements, (B) ccfDNA concentrations as per Quantus™ Fluorometer analysis.

Table 4. Summary of ccfDNA analyzed via two experimental methods.

Experiment Name	Units	Subjects	Cancer	Non-cancer	P value between groups
			40	40	
RT-qPCR Analysis	Log_{10} DNA copy number/µL	Mean	945.5	3.4	$P \leq 0.0$
		Range	340.4–1570.7 DNA copy number/µL	0.74-7.22 DNA copy number/µL	
		Standard deviation	306.60	2.05	
		Skewness	0.15	0.22	
		Kurtosis	-0.41	-1.36	
Fluorometric Analysis	ccfDNA in pg/µL	Mean	525.2	4.02	$P \leq 0.0$
		Standard deviation	82.8	0.51	
		Skewness	0.04	-0.07	
		Kurtosis	-0.65	-0.81	

Table 5. Comparison of DNA copy numbers between two study groups (cancer and non-cancer).

Cancer					Non-Cancer				
Cancer Type	Serial No	Age	Sex	DNA copy number	Cancer Type	Serial No	Age	Sex	DNA copy number
Breast Cancer	1	40	2	1075.37	Non-Cancer	1	40	1	3.83
	2	22	2	1528.00		2	22	1	7.22
	3	50	2	547.51		3	35	2	3.84
	4	38	2	637.10		4	50	2	5.85
	5	39	2	956.54		5	36	2	2.70
	6	50	2	540.02		6	23	1	5.72
	7	45	2	762.05		7	49	1	5.39
	8	32	2	1060.66		8	23	1	4.93
	9	50	2	911.51		9	24	1	5.32
	10	32	2	874.61		10	24	1	3.18
	11	40	2	1268.68		11	25	1	4.79
	12	47	2	1251.33		12	28	1	5.17
	13	38	2	1152.05		13	24	1	5.26
	14	25	2	976.51		14	23	2	3.81
Ovary Cancer	1	35	2	924.16		15	25	1	4.85
	2	50	2	682.53		16	24	1	1.41
	3	30	2	1075.37		17	24	1	1.16
	4	50	2	1090.29		18	23	1	0.98
Rectal Cancer	1	40	1	969.81		19	22	1	1.07
	2	21	1	850.84		20	24	1	1.85
	3	45	1	996.90		21	23	1	0.98
Leukemia	1	28	1	783.34		22	25	1	1.41
Squamous cell Carcinoma	1	71	2	1152.05		23	23	1	2.36
	2	52	2	1359.15		24	24	1	1.18
	3	50	1	799.70		25	40	1	1.18
	4	48	2	637.10		26	33	1	1.18
	5	35	2	1046.15		27	36	2	1.18
Lymphoma	1	28	2	1570.68		28	42	2	3.89
	2	38	1	1003.79		29	46	1	3.89
Early stage	1	28	2	874.61		30	4	2	1.80
	2	45	2	594.69		31	21	1	6.73
Pulmonary	1	19	1	772.62		32	32	1	3.70
Urethral Carcinoma	1	52	1	430.22		33	27	1	6.79
Osteosarcoma	1	40	2	1322.21		34	30	1	5.60
Papillary Cancer	1	40	1	731.20		35	18	2	6.52
Esophagus	1	38	1	454.60		36	65	1	4.45
Pleural effusion	1	50	2	936.98		37	34	2	0.74
Stomach Cancer	1	60	1	1517.51		38	40	2	1.77
Medulloblastoma	1	8	2	340.40		39	36	1	1.10
Adenocarcinoma	1	41	1	1359.15		40	33	1	0.79

The results showed that those with thyroiditis had significantly greater levels of ccfDNA than the control cohorts, with a statistical significance of $p < 0.0001$ [86]. The potential of circulating cell-free DNA (ccfDNA) in the early detection of diseases signifies a revolutionary shift in contemporary diagnostics. The detection of ccfDNA's presence and distinct attributes in the bloodstream can function as premature indicators of various diseases, well in advance of the detection capacity of conventional diagnostic techniques. In the context of cancer diagnostics, for example, deviations in ccfDNA levels or specific genetic alterations can indicate the emergence of malignancies, frequently preceding the

manifestation of any clinical manifestations [87]. This early detection capability of ccfDNA holds significance not only for cancer but also encompasses other conditions like cardiovascular diseases and autoimmune disorders [88,89].

In this study, all cancer types were encompassed within the cancer group, with no consideration given to factors such as sex, age, or specific cancer types. Likewise, the selection process for healthy controls adhered to the same methodology. However, our results remained consistent across different age groups, genders, and cancer types. Our analysis of descriptive statistics revealed differences in the distribution of plasma GAPDH DNA copies per µL of eluted ccfDNA solution between cancer patients and healthy controls. In cancer patients, the skewness value was 0.15, indicating a slightly right-skewed distribution (elongated tail towards higher values). This is further supported by the negative kurtosis value (-0.41), suggesting a flatter distribution compared to a normal curve. In contrast, healthy controls exhibited a higher skewness (0.22) and a more pronounced negative kurtosis (-1.4), indicating a potentially even flatter distribution with a longer tail towards higher values (Tables 4 and Table S2).

To further explore the distribution, a bell curve graph was created for all data assuming a normal distribution (Fig. S2A). This revealed that 96.76 % of the data from cancer patients fell within the expected area. However, the Z-distribution of the non-cancer plasma sample data didn't exhibit a bell pattern (Fig. S2B). Outlier analysis determined the upper and lower limits for all cancer data as 1632.81 and 227.26, respectively, indicating that all values were within the defined limit (Table S3).

To assess the diagnostic utility of GAPDH gene in ccfDNA of blood plasma, we conducted a receiver operating characteristic (ROC) curve analysis, considering the disparity in GAPDH DNA copies per µL of eluted ccfDNA solution between cancer and non-cancerous blood plasma. The graph in Fig. 2 illustrates the ROC curve, with the false positive rate (FPR) displayed on the horizontal x-axis and the true positive rate (TPR) shown on the vertical y-axis. Each point on the curve corresponds to a distinct threshold value. The cut-off point, indicated by the red dot on the graph, is where a false positive rate (FPR) of 0 corresponds to a sensitivity (TPR) of 1. Details are provided in Fig. 2 and Table S4. The ROC curve includes a model point (TPR = 1, FPR = 0) that ascends directly to the upper-left corner, signifying optimal performance. The area under the ROC curve was determined to be 1, indicating perfect discriminatory ability. The model exhibited a sensitivity of 100 % (Fig. 2). Details are provided in Table S4. Another study, employing the SYBR Green I fluorescence dye in qPCR, yielded similar results, achieving an AUC of 0.7679 and a sensitivity of 95 % for qPCR analysis [90]. In this study, this method utilized the BRYT Green® fluorescent DNA binding dye. Comparatively, it achieved superior results.

Our investigation is centered on 15 variations of cancer, exhibiting a prevalence of breast cancer patients and a smaller representation of other cancer subtypes. Owing to the restricted number of patients per specific cancer subtype - frequently just one or a few - our dataset is inadequate to confirm that GAPDH serves as a dependable biomarker for the majority of cancers. The comparative analysis of the GAPDH DNA copy number in ccfDNA derived from blood plasma across these 15 cancer types is depicted in Fig. 4 and

Table 8. The findings indicated that the leukemia sample exhibited the highest GAPDH DNA copy number (6282.73). Nonetheless, the adequacy of this data for concluding is limited by the small sample size (only one case). In contrast, early-stage cancer samples displayed the lowest GAPDH DNA copy number (1541.24), yet this information is also insufficient for decision-making due to the very limited sample size (only two cases). As a result, it is imperative to explore the implementation of the suggested methodology in a larger and more diverse cohort for specific cancer types to confirm the efficacy of evaluating GAPDH DNA expression in circulating cell-free DNA. Specifically, our analysis of breast cancer cases compared to healthy female controls revealed a significant difference ($p \leq 0$) in the GAPDH DNA copy number in eluted ccfDNA solution, as shown in Fig. 5.

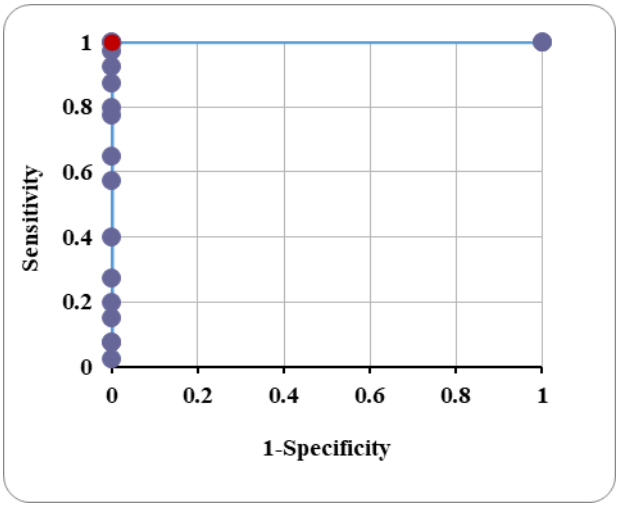

Fig. 2. ROC Curve with sensitivity and specificity.

The assessment of overall levels of cell-free DNA circulating in blood plasma can be enhanced using the RT-qPCR method with the human-specific GAPDH DNA. Although RT-qPCR is not a new technique, it offers a reliable approach to detecting cancer in individuals of any age or gender by quantifying cell-free DNA in blood plasma samples. The study findings demonstrated accurate quantification of total cell-free DNA in plasma samples from cancer patients using RT-qPCR, with the GAPDH DNA showing the highest amplification efficiency as the endogenous control. This methodology provides a non-invasive and cost-effective strategy for cancer screening compared to conventional invasive methods. These results are consistent with those reported by Skrypkina *et al.* [91]. Despite these promising outcomes, further validation with a larger sample size is necessary.

3.3. *Fluorometric results of ccfDNA concentrations in blood plasma*

The study compared fluorometric analysis with RT-qPCR for quantifying circulating cell-free DNA (ccfDNA) in blood plasma samples from both cancer patients and non-cancer

controls. Following the extraction of ccfDNA, its concentration was determined using a Quantus™ Fluorometer along with a double-stranded DNA-specific kit. Cancer patients exhibited a significantly higher of average plasma ccfDNA level (525.2 pg/µL ± 82.8 SD) compared to healthy controls (4.02 pg/µL ± 0.51 SD) (see Table 4, $p<0.001$, two-sample t-test assuming equal variances). A one-way ANOVA confirmed a significant difference in plasma ccfDNA levels between the two groups (F-statistic = 3.96, $p<0.05$). Further details can be found in Supplement 1, Section 5. These findings were consistent with the results obtained through the RT-qPCR method.

3.4. *Relative quantification of GAPDH gene in ccfmRNA of cancerous and non-cancerous blood plasma*

This method allows for simultaneous measurement of gene expression across multiple samples and various genes. Relative quantification assays rely on reference genes to standardize variations. These variations can arise from differing starting material amounts, amplification efficiency, gene expression fluctuations, and overall transcription levels. Accurate gene expression assessment necessitates normalizing qPCR results to a reliable reference gene. Ideally, this reference gene should show minimal change upon experimental treatments and maintain consistent expression across different cancer types. Multiple genes may be used for robust evaluation to identify a single gene with consistent expression across all conditions [92]. The relative quantification of the GAPDH gene against the other three housekeeping genes was analyzed in circulating cell-free RNA (cfRNA) of blood plasma using RT-qPCR have presented in Fig. 3 and Tables 6 & 7. The methodology employed the $2^{-\Delta\Delta Ct}$ method and Relative Fold Change (RFC) analyses to quantify the GAPDH gene. NormFinder, a model-based variation estimation algorithm, was used for normalization. This approach has been previously applied in studies on bladder and colon cancer [93]. Further details regarding the NormFinder algorithm experiment can be found in Supplement 1, Section 6.

3.4.1. *Relative Fold Change (RFC) of GAPDH Gene*

The fold change in gene expression indicated the variance in expression levels of the target gene between two samples or conditions. It was quantified using the $2^{-\Delta\Delta Ct}$ method. A fold change of a target gene greater than one from any reference gene signifies its upregulation, while a value less than one signifies its downregulation. This study investigated the expression of the GAPDH gene in cancerous and non-cancerous blood plasma in terms of three reference genes Beta-Actin, UBC, and HPRT1. The results showed considerable variation in fold changes for GAPDH across all reference genes. For instance, in cancerous blood plasma, the fold changes were (26.06 ± 12.8) for Beta-Actin, (8.89 ± 3.99) for UBC, and (7.34 ± 4.26) for HPRT1 (Table 6). In contrast, non-cancerous blood plasma exhibited distinct fold change values; (4.66 ± 3.3) for Beta-actin, (7.96 ±5.07) for UBC, and (2.91 ±

1.9) for HPRT1 (Table 6). It is noteworthy that, there was significant variation ($p<0.05$) in fold change values even within non-cancerous samples.

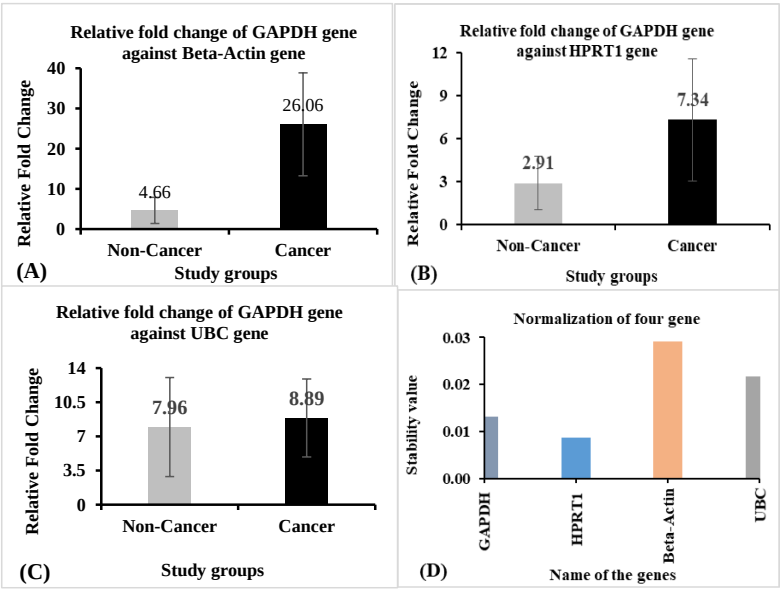

Fig. 3. Relative gene expression analysis of the GAPDH gene in the blood plasma of both cancer patients and non-cancer healthy controls: Relative Fold Change (RFC) of GAPDH gene against (A) Beta-Actin (ACTB) gene, (B) Hypoxanthine guanine phosphoryl transferase (HPRT1) gene, (C) Ubiquitin C (UBC) gene and (D) Analysis of Gene Expression Stability: NormFinder Assessment of Three Reference Genes and Targeted GAPDH Gene in ccfRNA.

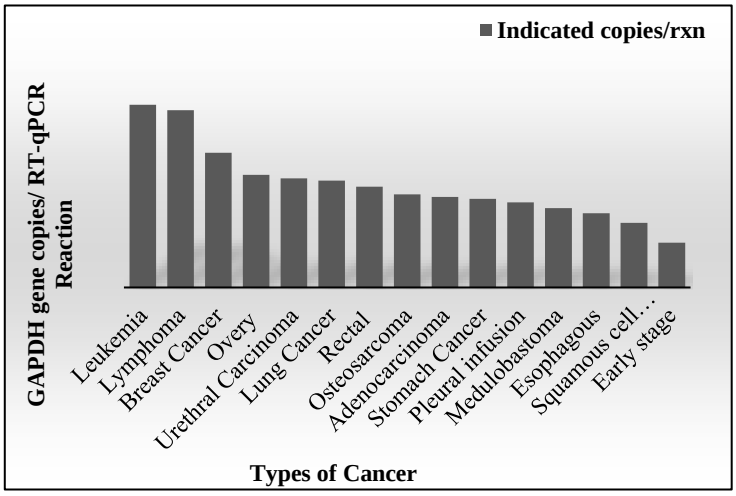

Fig. 4. Comparison of ccfGAPDH DNA copy numbers in blood plasma across fifteen types of cancer, determined by RT-qPCR.

The average $2^{-\Delta\Delta C}t$ values, presented in Table 6, demonstrate variations in the relative expression levels of the GAPDH gene compared to the three reference genes (Beta-Actin, UBC, and HPRT1) in cancer patients. These values (1.16, 1.41, and 3.33) are significantly higher ($p<0.05$) than those observed in non-cancer patients (0.59, 1.26, and 2.92). This indicates that both cancer and non-cancer groups exhibit disparities in GAPDH expression relative to the reference genes.

Fig. 3(A)-(C) further illustrate these findings. The figures depict the average relative fold change of the GAPDH gene compared to each reference gene. Black bars represent the fold change in cancerous blood plasma, while grey bars represent non-cancerous samples. The clear distinction between the bars visually confirms significant variations in GAPDH gene expression levels between cancer and non-cancer patients when compared to the reference genes.

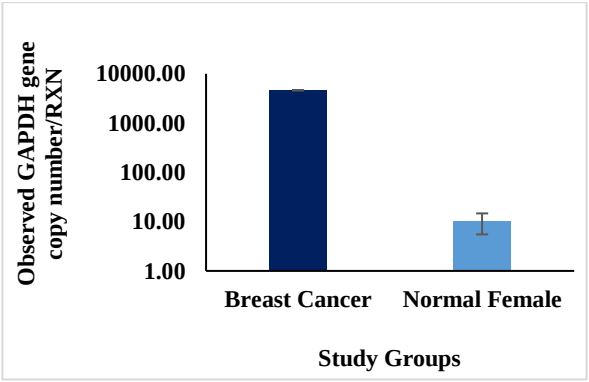

Fig. 5. ccfGAPDH DNA copy numbers variation between breast cancer patients and non-cancerous females.

Table 6. Relative fold change (RFC) and $2^{-\Delta\Delta Ct}$ of GAPDH gene against three reference genes between cancer and non-cancer blood plasma.

Statistical parameter	Gene Name	Cancer			Non-Cancer			P value
		Mean	Stdev	SE	Mean	Stdev	SE	
$2^{-\Delta\Delta Ct}$	ACTB	3.33	1.63	0.30	0.59	0.42	0.08	$P \leq 0.0$
	UBC	1.41	0.64	0.12	1.26	0.82	0.15	$P \leq 0.05$
	HPRT1	2.92	1.72	0.31	1.16	0.76	0.14	$P \leq 0.0$
RFC	ACTB	26.06	12.79	2.34	4.66	3.26	0.60	$P \leq 0.0$
	UBC	8.89	3.99	0.73	7.96	5.07	0.93	$P \leq 0.05$
	HPRT1	7.34	4.26	0.78	2.91	1.89	0.34	$P \leq 0.0$

3.4.2. *Normalization of three reference genes for expression of the GAPDH gene in blood plasma*

RT-qPCR proved to be a dependable, swift, and consistent method for assessing alterations in gene expression. To enhance gene expression investigations and ensure precise RT-qPCR results, normalization against stable reference genes was imperative. Traditionally, housekeeping genes like GAPDH, UBC, β-actin, elongation factor 1, and HPRT1 were used

as reference points for normalization, based on their consistently elevated expression levels [94,95]. In this study, relative gene expression was measured by comparing the expression of a target gene (GAPDH) with three reference genes (HPRT1, beta-actin, and UBC) to select the best reference gene or set of genes for normalization, as shown in Table 7.

Table 7. Normalization of reference gene against targeted GAPDH gene for gene expression.

Gene Type	Target gene	Reference gene		
Gene Name	GAPDH	HPRT1	Beta-Actin	UBC
Stability value	0.013	0.009	0.029	0.022
The best combination of two genes	GAPDH	HPRT1		
Stability value for the best combination of two genes	0.009			
Normalization	HPRT1 is the best reference gene for targeted GAPDH gene			

The NormFinder algorithm was employed to normalize the expression of the target gene (GAPDH) relative to reference genes. In this context, the stability value assigned to each gene by NormFinder reflects how consistently it expresses itself across different experimental conditions. Lower stability values indicate a more uniform expression. HPRT1 emerged as the ideal single reference gene due to its exceptionally low stability value of 0.009 (Table 7). This value signifies a highly consistent expression pattern. The target gene, GAPDH, and the other reference gene candidates (UBC and Beta-Actin) displayed higher stability values of 0.013, 0.029, and 0.022, respectively. While based on Table 7, HPRT1 appears to be the most suitable reference gene for normalizing GAPDH expression, it's important to note that the "best" reference gene can vary depending on the experiment. For instance, a previous study by Weber *et al*. and Santin *et al.* identified ACTB (Beta-Actin) as the most reliable reference gene in their investigation of 14 thyroid specimens and 4 primary culture cells treated with hormones [96,97].

Table 8. Comparative GAPDH DNA copy number per RT-qPCR reaction of 15 cancer cases samples.

SL no	Cancer Type	Number of patients	Average GAPDH DNA copy number/rxn
1	Leukemia	1	6282.73
2	Lymphoma	2	6091.01
3	Breast Cancer	14	4632.33
4	Ovary Cancer	3	3870.48
5	Urethral Carcinoma	1	3747.91
6	Lung Cancer	2	3671.34
7	Rectal Cancer	3	3466.75
8	Osteosarcoma	1	3198.79
8	Adenocarcinoma	2	3111.93
9	Stomach Cancer	1	3048.21
11	Pleural infusion	1	2924.80
12	Medulloblastoma	1	2730.12
13	Esophagus	1	2548.40
14	Squamous cell carcinoma	5	2219.14
15	Early Stage	2	1541.24

Interestingly, the human-specific GAPDH gene served a dual purpose in this study. While traditionally used as a reference gene for circulating cell-free messenger RNA (ccfmRNA) quantification, this study also explored its application in circulating cell-free DNA (ccfDNA) copy number analysis. This dual biomarker approach allowed for a more comprehensive understanding of both genetic and transcriptional changes.

4. Conclusion

In conclusion, our study presents a promising avenue for rapid and cost-efficient cancer detection through the absolute quantification of circulating cell-free DNA (ccfDNA), with a focus on glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene copy number. For the evaluation of the result of absolute GAPDH gene quantification by quantitative real time polymerase chain reaction, relative quantification was performed using the GAPDH gene in comparison to three additional housekeeping genes. The significantly elevated expression of GAPDH gene in both ccfDNA and ccfRNA from cancer patients, compared to individuals without a cancer diagnosis underscored its potential as a reliable biomarker for rapid cancer detection. By leveraging RT-qPCR to quantify GAPDH gene expression levels, we demonstrated a non-invasive approach with high sensitivity and specificity for cancer screening. This method offers a promising alternative to traditional screening methods, addressing their limitations by providing a faster and more cost-effective option. A recent study highlighted the potential of circulating cell-free DNA (cfDNA) as a biomarker for predicting and diagnosing papillary thyroid cancer (PTC) [97]. The validation of GAPDH gene expression as a biomarker for cancer detection opened new pathways for developing liquid biopsy-based diagnostic approaches. Further research and validation studies are necessary to refine and optimize this methodology for clinical implementation. Additionally, exploring the utility of GAPDH gene expression in combination with other biomarkers may enhance the diagnostic accuracy and reliability of cancer screening protocols. Overall, our findings contribute to advancement in the field of early cancer detection and emphasize the importance of continued research efforts to improve screening methods, ultimately leading to earlier diagnoses, better patient outcomes, and reduced healthcare costs.

Study Limitations

We acknowledge certain limitations in our research. However, we have offered a clear evaluation of the extent and consequences of our study, as well as pinpointing areas for future investigation and enhancement.

Sample Size Limitation: The study's sample size of 80 participants, with 40 diagnosed with cancer and 40 without cancer, may be considered relatively small, potentially limiting the generalizability of the findings. In our study, we worked with 15 types of cancer. Among them, fourteen samples were in the breast cancer group and others consisted of only 1 or a

few patients for most cancers. While our results provide valuable insights, larger sample sizes are warranted for each cancer group to confirm and extend these findings to broader populations.

Validation with Larger Sample Size: Although the research demonstrates promising results regarding the elevated expression of the GAPDH gene in cancer patients, further validation with larger and more diverse sample size is recommended before clinical application. This would enhance the robustness and reliability of our findings.

Focus on GAPDH Gene: The study primarily focuses on the GAPDH gene as a biomarker for cancer detection, potentially overlooking other genetic markers or pathways that could contribute to a more comprehensive understanding of cancer biology and diagnosis. Future research should explore a broader range of biomarkers to capture the complexity of cancer biology.

Focus on Age Group: The variability in age at diagnosis for different types of cancers might affect the observed levels of ccfDNA in the blood plasma. The outcomes of our investigation might lack complete representativeness across all age categories owing to the limited age span of the individuals involved. This constrained variation in age could potentially impact the applicability of the results, given that variations in DNA copy number quantification and gene expression patterns due to age-related biological factors may arise. So, it is necessary to further research with a broader age demographic to validate the findings across different age cohorts.

Lack of Specific Cancer Types or Stages: In a specific group of cancer patients, there is a considerable quantity of heterogeneity, or variation, in the causes (etiology) of their cancers. This variation can arise from various factors such as genetic mutations, environmental influences, lifestyle choices, and more. Due to this heterogeneity, it is important to understand whether there are differences in the amount of circulating cell-free DNA (ccfDNA) present in the blood plasma among different types of cancer. However, due to the limited sample size for individual cancer types, it was impossible to study this in our research. Our investigation does not center on particular cancer categories or phases, factors that could impact the detected Log_{10} DNA copy number examination, and genetic expression trends, along with the applicability of the findings across different cancer subcategories. Subsequent studies ought to contemplate classifying subjects based on the kind and phase of cancer to offer more elaborate perspectives.

Limitations of RT-qPCR: There are limitations regarding scalability and availability when using RT-qPCR to examine gene expression in clinical settings. This is because this procedure requires specialist knowledge and equipment. Examining other methodologies or technological solutions that offer advantages for applicability and practicability could enhance the translational consequences of our findings.

Funding: This research project was financially supported by the Ministry of Education in Bangladesh as part of the HEQEP initiative, specifically under project CP-4044.

Acknowledgments

We extend our gratitude to the National Institute of Cancer Research and Hospital (NICRH) in Bangladesh for their invaluable assistance in collecting all cancer samples. The collection of human blood samples strictly adhered to the guidelines established by the governing body. We would like to express our sincere appreciation to S. Khan from Bangladesh Council of Scientific and Industrial Research (BCSIR) Dhaka, Bangladesh, for facilitating the necessary work permissions under the ethical approval (Ref. NICRH/Ethics/2019/525) granted for the study conducted for the project "Establishment of the Genomic Research Laboratory" at the Bangladesh Council of Scientific and Industrial Research (BCSIR), Dhaka, Bangladesh. Additionally, we are thankful to Invent Technologies Limited for their invaluable assistance and support as it is one of the leading and reputable molecular laboratory service providers in Bangladesh.

References

1. S. Meneguin, I. G. Alves, H. T. F. Camargo, C. F. Pollo, A. V. Z. Segalla, and C. de Oliveira, Eur. J. Investig. Health Psychol. Educ. **14**, 339 (2024). https://doi.org/10.3390/ejihpe14020023
2. S. Craig, Y. Cao, J. McMahon, T. Anderson, P. Stark et al., Healthcare (Basel) **11**, ID 3166 (2023). https://doi.org/10.3390/healthcare11243166
3. J. J. Mao, G. G. Pillai, C. J. Andrade, J. A. Ligibel, P. Basu et al., CA Cancer J. Clin. **72**, 144 (2022). https://doi.org/10.3322/caac.21706
4. R. M. Munshi, PLoS One **19**, ID e0296107 (2024). https://doi.org/10.1371/journal.pone.0296107
5. P. Kumar, S. Gupta, and B. C. Das, Transl. Oncol**. 40**, ID 101827 (2024). https://doi.org/10.1016/j.tranon.2023.101827
6. A. Pulumati, A. Pulumati, B.S. Dwarakanath, A. Verma, and R. V. Papineni, Cancer Rep. **6,** ID e1764 (2023). https://doi.org/10.1002/cnr2.1764
7. E. M. Stoffel, R. E. Brand, and M. Goggins, Gastroenterology **164**, 752 (2023). https://doi.org/10.1053/j.gastro.2023.02.012
8. P. Mandel and P. Metais, CR Seances Soc. Biol. Fil. **142**, 241 (1948).
9. J. Aucamp, A. J. Bronkhorst, C. P. S. Badenhorst, P. J Pretorius, Biol. Rev. Camb. Philos. Soc. **93,** 1649 (2018). https://doi.org/10.1111/brv.12413.
10. E. Heitzer, L. Auinger, and M. R. Speicher, Trends Mol. Med. **26**, 519 (2020). https://doi.org/10.1016/j.molmed.2020.01.012
11. B. Salfer, F. Li, D. T. Wong, and L. Zhang, Clin. Chem. **68**, 1493 (2022). https://doi.org/10.1093/clinchem/hvac122
12. Y. Wang, S. Springer, M. Zhang, K. W McMahon, I. Kinde, L. Dobbyn, and C. Bettegowda, Proc. Natl. Acad. Sci. U. S. A. **112**, 9704 (2015). https://doi.org/10.1073/pnas.1511694112
13. K. B. Sriram, V. Relan, B. E Clarke, E. E. Duhig, M. N. Windsor, K. S. Matar, and K. M Fong, BMC Cancer **12**, 428 (2012). https://doi:10.1186/1471-2407-12-428
14. K. Sun, P. Jiang, K. C. A. Chan, J. Wong, Y. K. Y Cheng, et al., LoProc. Natl. Acad. Sci. U. S. A. **112**, E5503 (2015). https://doi.org/10.1073/pnas.1508736112
15. T. Qi, M. Pan, H. Shi, L. Wang, Y. Bai, and Q. Ge, Int. J. Mol. Sci. **24**, 1503 (2023). https://doi.org/10.3390/ijms24021503

16. M. Gögenur, J. Burcharth, and I. Gögenur, Crit. Care **21**, 14 (2017). https://doi.org/10.1186/s13054-016-1578-9
17. C. M. Lo, H. B. Wang, M. Dembo, and Y. L. Wang, Biophys. J. **79**, 144 (2000). https://doi.org/ 10.1016/S0006-3495(00)76279-5
18. Y. Y. Lui, K. W. Chik, R. W. Chiu, C. Y. Ho, C. W. Lam, and Y. M. Lo, Clin. Chem. **48**, 421 (2002). https://doi.org/10.1093/clinchem/48.3.421
19. O. Caglar, B. Cilgin, M. Eroglu, and A. Cayir, Braz. J. Otorh. **86(3)**, 321 (2020). https://doi.org/10.1016/j.bjorl.2018.12.008
20. A. Szpechcinski, P. Rudzinski, W. Kupis, R. Langfort, T. Orlowski, and J. Chorostowska-Wynimko et al., Int. J. Mol. Sci. **18**, 1350 (2017). https://doi.org/10.1016/j.canlet.2016.02.002
21. R. Klimaite, M. Kazokaite, A. Kondrotiene, D. Dauksiene, R. Verkauskiene, B. Zilaitiene, and A. Dausa, Anti. Res., **42**, 2289 (2022). https://doi.org/10.21873/anticanres.15708
22. M. v. Vaart and P. J. Pretorius, Ann. N. Y. Acad. Sci. **1137**, 18 (2008). https://doi.org/10.1196/annals.1448.022
23. Ó. Rapado-González, L. Muinelo-Romay and M. M. Suarez-Cunqueiro, Oral Onco. **112**, ID 105016 (2021). https://doi.org/ 10.1016/j.oraloncology.2020.105016
24. A. Jamshidi, M. C. Liu, E. A. Klein, O. Venn, E. Hubbell et al., Cancer Cell, **40**, 1537 (2022). https://doi.org/10.1016/j.ccell.2022.10.022
25. A. J. Bronkhorst, V. Ungerer, and S. Holdenrieder, Biomol. Detect. Quantif. **17**, ID 100087 (2019). https://doi.org/10.1016/j.bdq.2019.100087
26. M. Shtumpf, K. V. Piroeva, S. P. Agrawal, D. R. Jacob, and V. B. Teif, Chromosoma **131**, 19 (2022). https://doi.org/10.1007/s00412-021-00766-9
27. K. W. E. Cheung, S. Y. R. Choi, L. T. C. Lee, N. L. E. Lee, H. F Tsang, et al., Expert Rev. Mol. Diagn. **19**, 579 (2019). https://doi.org/10.1080/14737159.2019.1633307
28. P. P. Leong, W. M. Koch, A. Reed, D. Eisele, D. Sidransky, J. Jen, and W. H. Westra, J. Natl. Cancer Inst. **90**, 972 (1998). https://doi.org/ 10.1093/jnci/90.13.972
29. L. Lu, J. Bi, and L. Bao, J. Genet. Genomics **45**, 79 (2018). https://doi.org/10.1016/j.jgg.2017.11.006.
30. S. Mader and K. Pantel, Oncol. Res. Treat**. 40**, 404 (2017). https://doi.org/10.1159/000478018
31. K. Pantel and C. Alix-Panabières, Nat. Rev. Clin. Oncol. **16**, 409 (2019). https://doi.org/10.1038/s41571-019-0187-3
32. P. J. Vellanki, S. Ghosh, A. Pathak, M. J. Fusco, E. W. Bloomquist, et al., J. Immunother. Cancer **11**, ID e005344 (2023). https://doi.org/10.1136/jitc-2022-005344
33. G. F. Barbosa and M. Milas, Expert Rev. Anticancer Ther. **8**, 1415 (2008). https://doi.org10.1586/14737140.8.9.1415
34. C. E. Condrat, D. C. Thompson, M. G. Barbu, O. L. Bugnar, A. Boboc, et al., Cells **9**, 276 (2020). https://doi.org/10.3390/cells9020276
35. J. Darr, A. Tomar, M. Lassi, M. R. Gerlini, L. Berti et al., Cell Rep. **30**, 3183 (2020). https://doi.org/ 10.1016/j.celrep.2020.02.020
36. K. Pantel and C. Alix-Panabières, Nat. Rev. Clin. Oncol. **16**, 409 (2019). https://doi.org/10.1038/s41571-019-0187-3
37. M. Chaddha, H. Rai, R. Gupta, and D. Thakral, Front. Genet. **14**, ID 1138625 (2023). https://doi.org/10.3389/fgene.2023.1138625
38. N. P Tessier, L. M. Hardy, J. F. Deleuze, and A. How-Kit, Front. Genet. **14**, ID 1321280 (2023). https://doi.org/10.3389/fgene.2023.1321280
39. M. Tilli, C. D. S. Castro, J. A. Tuszynski, and N. Carels, BMC genomics **17**, 1 (2016). https://doi.org /10.1186/s12864-016-2946-1
40. I. Naletova, E. Schmalhausen, B. Tomasello, D. Pozdyshev, F. Attanasio, and V. Muronetz, Front. Mol. Biosci. **10**, ID 1256963 (2023). https://doi.org/10.3389/fmolb.2023.125696
41. N.W. Seidler. Adv. in Exp. Med. and Bio. (Springer, Dordrecht, Germany, 2013) **985**, pp. 37. https://doi.org/10.1007/978-94-007-4716-6_2
42. A. Malla, S. Gupta, and R. Sur, Cell Biochem. Biophys. **82**, 351 (2024).

https://doi.org/10.1007/s12013-023-01213-5
43. S. J. Mallinson, D. Dessaux, S. Barbe, and Y. J. Bomble, ACS Catal. **13**, 11781 (2023). https://doi.org/10.1021/acscatal.3c01452
44. P. Li, R. Zhong, J. Yu, Y. Wang, C. Wang, W. Geng, S. Bao, S. Wang et al., Curr. Eye Res. **48**, 992 (2023). https://doi.org/10.1080/02713683.2023.2241159
45. Z. Chen, J. Song, L. Xie, G. Xu, C. Zheng, et al., Cell Death Differ. **30**, 1957 (2023). https://doi.org/10.1038/s41418-023-01190-5
46. N. Agrawal, M. Khanna, and G. Dhawan, Ind. J. Med. Micro. **42**, 49 (2023). https://doi.org/10.1016/j.ijmmb.2023.01.007
47. G. S. Krasnov, A. A.Dmitriev, A. V. Snezhkina and A. V. Kudryavtseva, Expert Opin. Ther. Targets **17**, 681 (2013). https://doi.org/10.1517/14728222.2013.775253
48. N. Barta, N. Ördög, V. Pantazi, I. Berzsenyi, B. N. Borsos et al., Biomolecule **13**, 1523 (2023). https://doi.org/10.3390/biom1310152
49. S. Ganapathy-Kanniappan, R. Kunjithapatham, and J. F. Geschwind, Oncotarget **3**, 940 (2012). https://doi.org/10.18632/oncotarget.623
50. K. Liu, Z. Tang, A. Huang, P. Chen, P. Liu et al., Int. J. Oncol. **50**, 252 (2017). https://doi.org/10.3892/ijo.2016.3774
51. K. Tokunaga, Y. Nakamura, K. Sakata, K. Fujimori, M. Ohkubo et al., Cancer Res. **47**, 5616 (1987).
52. M. R Vilà, A. Nicolás, J. Morote, I. de Torres, and A. Meseguer, Cancer **89**, 152 (2000). https://doi.org/10.1002/1097-0142(20000701)
53. F. Revillion, V. Pawlowski, L. Hornez, and J. P Peyrat, Eur. J. Cancer **36**, 1038 (2000). https://doi.org/10.1016/S0959-8049(00)00051-4
54. X. Y. Zhong, A. Ladewig, S. Schmid, E. Wight, S. Hahn et al., Arch. Gynecol. Obstet. **276**, 327 (2007). https://doi.org/10.1007/s00404-007-0345-1
55. F. Ohl, M. Jung, A. Radonić, M. Sachs, S. A. Loening, and K. Jung, J. Urol. **175**, 1915 (2006). https://doi.org/10.1016/S0022-5347(05)00919-5
56. D. E. Epner, A.W. Partin, J. A. Schalken, J. T. Isaacs, and D. S. Coffey, Cancer Res. **53**, 1995 (1993).
57. N. Harada, R. Yasunaga, Y. Higashimura, R. Yamaji, K. Fujimoto et al., J. Biol. Chem. **282 31**, 22651(2007). https://doi.org/10.1074/jbc.M610724200
58. J. Y. Zhang, F. Zhang, C. Q. Hong, A. E. Giuliano, X. J. Cui et al., Cancer Biol. Med. **12,** 10 (2015). https://doi.org/10.7497/j.issn.2095-3941.2014.0019
59. G. Shivashankar, J. C. Lim, and M. L. Acosta, Exp. Eye Res. **213**, ID 108845 (2021). https://doi.org/10.1016/j.exer.2021.108845
60. A. Iwamoto, Y. Inoue, H. Tachibana, Cytotechnology **73**, 333 (2021). https://doi.org/10.1007/s10616-020-00438-z
61. G. Carraro, G. Albertin, M. Forneris, and G. G. Nussdorfer, Mol. Cell. Probes **19**, 181 (2005). https://doi.org/10.1016/j.mcp.2004.11.004
62. Y. J. Yang, H. Chen, P. Huang, C. H. Li, Z. H. Dong et al., Clin. Invest. Med. **34**, E238 (2011). https://doi.org/10.25011/cim.v34i4.15366
63. F. Wang, G. P. Sun, Y. F. Zou, J. O. Hao, F. Zhong and W. J. Ren, Cancer Biomark. **11**, 259 (2012). https://doi.org/10.3233/cbm-2012-00284
64. H. Zheng, L. Zhang, Y. Zhao, D. Yang, F. Song et al., PLoS One **8**, ID e77853 (2013). https://doi.org /10.1371/journal.pone.0077853
65. R. Yi, Y. Li, F. L. Wang, G. Miao, R. M. Qi and Y. Y. Zhao, World J. Gastrointest. Oncol. **8**, 330 (2016). https://doi.org/10.4251/wjgo.v8.i4.330
66. L. Z Sorber, K. Deschoolmeester, V. Roeyen, G. Rolfo, C. Pauwels, Belg. J Med. Oncol. **15,** 48 (2021).
67. B. O. V. Emburgh, S. Arena, G. Siravegna, L. Lazzari, G. Crisafulli et al., Nat. Commun. **7**, ID 13665 (2016). https://doi.org/10.1038/ncomms13665
68. E. Y. Lee, E. J. Lee, H. Yoon, D. H. Lee, and K. H. Kim, Diagnostics (Basel) **10**, 234 (2020). https://doi.org/10.3390/diagnostics10040234

69. A. L. Baker, Investigating the Novel Oncogenic Function of SAPCD2 in Pediatric Neuroblastoma (Texas State University, 2021). https://digital.library.txst.edu/bitstreams/d2e7ed7d-f566-4f6b-a827-116405f657a5
70. Ruslan I. Dmitriev, Dmitri B. Papkovsky, FEBS Lett. **589**, 138 (2015). https://doi.org/10.1016/j.febslet.2014.11.038
71. A. Zulueta, A. Caretti, P. Signorelli, F. Dall'Olio, and M. Trinchera, FASEB J. **28**, 946 (2014). https://doi.org/10.1096/fj.13-236273
72. X. Gao, L. Han, N. Ding, Y. Mu, P. Guan et al., J. Antibiot. (Tokyo) **71**, 808 (2018). https://doi.org/10.1038/s41429-018-0066-7
73. D. Posfai, K. Sylvester, A. Reddy, J.G. Ganley, J. Wirth et al., PLoS Pathog. **14**, ID e1007057 (2018). https://doi.org/10.1371/journal.ppat.1007057
74. A. L. Niles, K.R. Kupcho, D. F. Lazar, and J. J. Cali, Cancer Res. **81**, 1933 (2021). https://doi.org /10.1158/1538-7445.AM2021-1933
75. E. Fischer and J. Mook, Promega Corporation, 2800 Woods Hollow Rd, Madison, WI 53711 (2023).
76. L. Roberts, R. Ratnapriya, M. du Plessis, V. Chaitankar, R. S. Ramesar, and A. Swaroop, Invest. Ophthalmol. Vis. Sci. **57**, 6374 (2016). https://doi.org/10.1167/iovs.16-19785
77. K. J. Livak and T. D. Schmittgen, Methods **25**, 402 (2001). https://doi.org/10.1006/meth.2001.1262
78. F. Biermann, J. Mathews, B. Nießing, N. König, and R. H. Schmitt, Processes (Basel) **9**, 966 (2021). https://doi.org/10.3390/pr9060966
79. C. Nusbaum, T. Mikkelsen, M. Zody, S. Asakawa, S. Taudien et al., Nature **439**, 331 (2006). https://doi.org/10.1038/nature04406
80. S. A. Bustin, V. Benes, J. A. Garson, J. Hellemans, J. Huggett et al., Clin. Chem. **55**, 611 (2009). https://doi.org/10.1373/clinchem.2008.112797
81. J. Torra, M. D. Osuna, A. M. Junior, Multiple Herbicide-Resistant Weeds and Non-target Site Resistance Mechanisms: A Global Challenge for Food Production (Frontiers Media SA, 2021).
82. S. Gal, C. Fidler, Y. M. D Lo, M. Taylor, C. Han et al., Br. J. Cancer **90**, 1211 (2004). https://doi.org /10.1038/sj.bjc.6601609
83. S. Singh, R. Goyal, A. Gupta, R. Singh, M. Singh et al., Ind. J. Clin. Biochem. (2024). https://doi.org/10.1007/s12291-024-01181-4
84. A. A. Kamat, A. K. Sood, D. Dang, D. M. Gershenson, J. L. Simpson, and F. Z. Bischoff, Ann. N. Y. Acad. Sci. **1075**, 230 (2006). https://doi.org /10.1196/annals.1368.031
85. D. Allen, A. Butt, D. Cahill, M. Wheeler, R. Popert, and R. Swaminathan, Annals New York Acad. Sci. **1022**, 76 (2004). https://doi.org /10.1196/annals.1318.013
86. O. Caglar, B. Cilgin, M. Eroglu, and A. Cayir, Braz. J. Otorhinolaryngol. **86**, 321 (2020). https://doi.org /10.1016/j.bjorl.2018.12.008
87. B. Altieri, S. Appenzeller, W. Arlt, M. Asia, V. Chortis, Y. S. Elhassan et al., J. Endocr. Soc. **6**, A81 (2022). https://doi.org/10.1210/jendso/bvac150.169
88. M. Odah, Preprints.org (2024). https:// doi. Org/10.20944/preprints202403.0467.v2
89. J. Xie, J. Yang, and P. Hu, The Am J. Med. Sci. **356**, 121 (2018). https://doi.org/10.1016/j.amjms.2018.04.007
90. M. L. Dahn, C. A. Dean, D. B. Jo, K. M. Coyle, and P. Marcato, Mol. Ther. Meth. Clin. Dev. **20**, 398 (2021).
91. I. Skrypkina, L. Tsyba, K. Onyshchenko, D. Morderer, O. Kashparova et al., Dis. Markers 2016, ID 3693096 (2016). https://doi.org/10.1155/2016/3693096
92. C. Buccitelli and M. Selbach, Nat. Rev. Genetics **21**, 630 (2020). https://doi.org/10.1038/s41576-020-0258-4
93. C. L. Andersen, J. L. Jensen, and T. F. Ørntoft, Cancer Res. **64**, 5245 (2004). https://doi.org/10.1158/0008-5472.CAN-04-0496
94. S. Arrowsmith, Mol. Biol. Rep. **48**, 413 (2021). https://doi.org/10.1007/s11033-020-06066-2
95. H. Pan, X. Yang, B. D. Siegfried and X. Zhou, PLoS One **10**, ID e0125868 (2015).

https://doi.org/10.1371/journal.pone.0125868
96. R. Weber, A. P. S. Bertoni, L. W. Bessestil, B. M. D. A. A. Brasil, and T. W. Furlanetto, Biomed Res. Int. 2014, ID 198582 (2014). https://doi.org/10.1155/2014/198582
97. A. P. Santin, A. F. D. Souza, L. S. Brum, and T.W. Furlanetto, Mol. Biotecnol. **54**, 278 (2013). https://doi.org/10.1007/s12033-012-9565-0

Supplementary Materials

1. Standard Curve Generation

The standard curve was created on six-fold serial dilutions of known amounts of human genomic DNA, amplified by the GAPDH gene using Real-Time quantitative PCR. The dilutions ranged from 10^5 to 1 pg/µL, with a dilution factor of 6 (including concentrations of 100,000, 10,000, 1000, 100, 10 and 1 pg/µL). On the graph, the Y-axis shows the threshold cycle (Ct), indicating the cycle number at which the fluorescence signal reaches a predetermined threshold. Meanwhile, the X-axis represents the $\log_{10}$ DNA copies RT-qPCR reaction, presented on a logarithmic scale.

The standard curve showcased a robust and consistent linear relationship between the Ct values and the logarithm of the DNA copy numbers (Fig. S1).

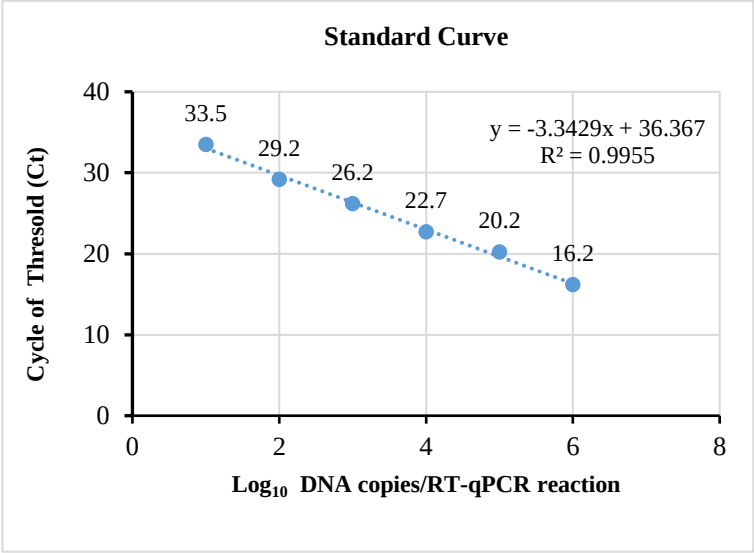

Fig. S1. Standard curve for the GAPDH gene created by the amplification of known concentrated human genomic DNA (10 ng/µL; cat # 4312660, Applied Biosystem) using RT-qPCR technique.

2. Calculation of Log_{10} DNA Copy Number:

In our study, we calculate, Log_{10} DNA copy numbers according to the following equation:

$X_0 = E_{AMP} * (b\text{-}Ct)$ [1]

where $X_0 = Log_{10}$ DNA copy number,

E_{AMP} = efficiency of the standard curve =1.99,

b= intercept of the standard curve = 36.37,

Ct = Cycle of the threshold of each sample.

Table S1. RT-qPCR machine-generated Ct values and their corresponding $\log_{10}$ DNA copies/ RT-qPCR Reaction (rxn).

SL No	Patients diagnosed with Cancer					Healthy Control, not diagnosed with Cancer				
	Ct1	Ct2	Ct3	Avg. Ct	DNA copies/rxn	Ct1	Ct2	Ct3	Avg. Ct values	DNA copies/rxn
1	24.21	24.23	24.22	24.22	4301.48	32.4	32.41	32.4	32.40	15.33
2	23.72	23.72	23.70	23.71	6111.99	31.5	31.5	31.49	31.50	28.60
3	25.20	25.12	25.30	25.20	2190.06	32.42	32.41	32.42	32.42	15.19
4	25.01	24.98	24.89	24.98	2548.40	31.78	31.78	31.78	31.78	23.54
5	24.40	24.38	24.40	24.39	3826.16	32.92	32.92	32.92	32.92	10.74
6	25.23	25.21	25.22	25.22	2160.10	31.82	31.82	31.83	31.82	22.84
7	24.72	24.72	24.72	24.72	3048.21	31.89	31.91	31.9	31.90	21.67
8	24.23	24.25	24.24	24.24	4242.63	32.04	32.04	32.04	32.04	19.68
9	24.50	24.46	24.41	24.46	3646.06	31.91	31.92	31.92	31.92	21.42
10	24.50	24.56	24.50	24.52	3498.44	32.68	32.68	32.68	32.68	12.67
11	23.96	23.98	24.00	23.98	5074.73	32.08	32.08	32.08	32.08	19.15
12	23.98	24.02	24.00	24.00	5005.30	32	32	32	32.00	20.23
13	24.10	24.11	24.12	24.12	4608.22	31.94	31.94	31.94	31.94	21.08
14	24.35	24.34	24.37	24.36	3906.05	32.41	32.41	32.41	32.41	15.26
15	24.43	24.44	24.45	24.44	3696.63	32.06	32.05	32.05	32.05	19.50
16	24.88	24.87	24.89	24.88	2730.12	33.85	33.85	33.85	33.85	5.66
17	24.22	24.22	24.22	24.22	4301.48	34.12	34.11	34.123	34.12	4.71
18	24.18	24.22	24.20	24.20	4361.15	34.38	34.38	34.38	34.38	3.93
19	24.34	24.40	24.37	24.37	3879.24	34.25	34.26	34.25	34.25	4.29
20	24.54	24.56	24.58	24.56	3403.37	33.46	33.46	33.46	33.46	7.41
21	24.32	24.33	24.34	24.33	3987.61	34.38	34.39	34.39	34.39	3.91
22	24.68	24.66	24.70	24.68	3133.37	33.85	33.85	33.85	33.85	5.66
23	24.12	24.11	24.13	24.12	4608.22	33.11	33.12	33.11	33.11	9.40
24	23.88	23.86	23.90	23.88	5436.60	34.11	34.12	34.12	34.12	4.71
25	24.66	24.65	24.66	24.65	3198.79	34.1	34.12	34.11	34.11	4.74
26	24.99	24.97	24.98	24.98	2548.40	34.11	34.11	34.11	34.11	4.74
27	24.26	24.26	24.26	24.26	4184.59	34.11	34.11	34.11	34.11	4.74
28	23.67	23.66	23.68	23.67	6282.73	32.38	32.38	32.38	32.38	15.57
29	24.34	24.32	24.30	24.32	4015.17	32.38	32.39	32.38	32.38	15.54
30	24.50	24.54	24.52	24.52	3498.44	33.5	33.49	33.5	33.50	7.22
31	25.12	25.06	24.98	25.08	2378.77	31.58	31.58	31.58	31.58	27.01
32	24.70	24.70	24.70	24.70	3090.50	32.47	32.47	32.47	32.47	14.64
33	25.54	25.56	25.55	25.55	1720.90	31.57	31.57	31.57	31.57	27.20
34	23.92	23.90	23.94	23.92	5288.86	31.86	31.86	31.86	31.86	22.28
35	24.78	24.77	24.79	24.78	2924.80	31.63	31.64	31.63	31.63	26.04
36	25.47	25.46	25.48	25.47	1818.39	32.19	32.19	32.2	32.19	17.71
37	24.43	24.42	24.41	24.42	3747.91	34.8	34.8	34.8	34.80	2.95
38	23.70	23.73	23.74	23.72	6070.03	33.53	33.53	33.53	33.53	7.06
39	25.90	25.88	25.89	25.89	1361.59	34.21	34.21	34.21	34.21	4.42
40	23.88	23.88	23.88	23.88	5436.60	34.68	34.68	34.68	34.68	3.20

The Ct values obtained from RT-qPCR amplification of the GAPDH gene in the circulating cell-free DNA were measured for both cancerous and non-cancerous blood plasma samples and their corresponding DNA copy numbers were been calculated in Table S1.

3. Probability Analysis of Our Data:

For the probability analysis of our data sets, three descriptive statistics were used in our study.

3.1. Normal distribution:

In this study, we used the normal distribution for the probability analysis of our data sets. Normal distribution, or Gaussian distribution, is a fundamental concept in statistics and probability theory.

In our study three observations are pointed out by the normal distribution:

1. The data clusters around the mean (average) value.
2. The distribution is symmetric, meaning the left and right halves mirror each other.
3. The spread of the data is determined by the standard deviation.

To visually represent the probability distribution as a normal distribution, we generated a bell curve graph. Notably, within this representation, approximately 96.76% of the data points from the cancer patient group were situated within the bounds of the z-distribution. This graphical depiction can be observed in Figure S2 (A). In this figure in panel (A), the graph portrays the typical bell curve characteristic of a normal distribution, often referred to as the Z-distribution. This curve delineated the distribution of GAPDH ccfDNA copy numbers found in the blood plasma of individuals diagnosed with cancer. Conversely, in panel (B), the bell curve pattern was absent when applying the Z-distribution to the dataset comprising blood plasma samples from non-cancer subjects.

These findings underline the distinction in distribution patterns between the data from the blood plasma of cancer patients and the non-cancer peoples, showcasing the potential implications of such disparities in our study's context.

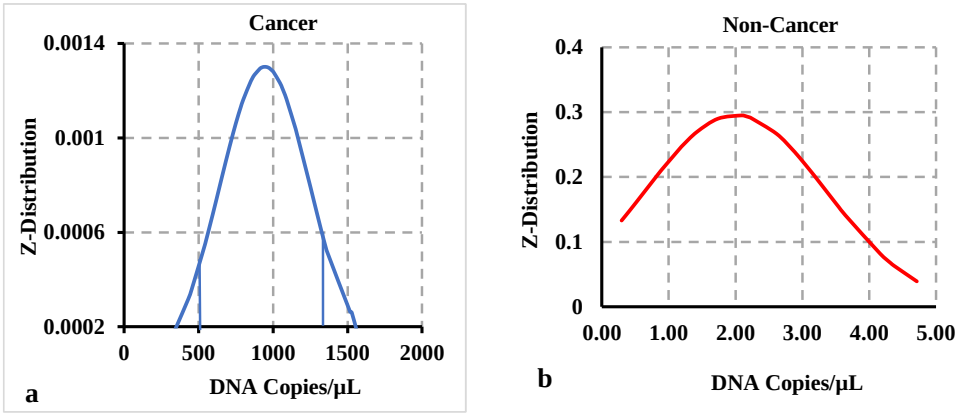

Fig. S2. Normal distribution of DNA copy numbers obtained from blood plasma samples. Fig. S2(a), the graph represents the normal distribution of GAPDH ccfDNA copy numbers found in the blood plasma of cancer-diagnosed patients. Fig. S2 (b), showed the normal distribution of the dataset obtained from non-cancer subjects.

3.2. *Skewness and kurtosis analysis of our data:*

In statistical analysis, examining skewness and kurtosis can provide insights into the shape and characteristics of the data distribution. For instance, they can help in assessing whether the data follow a normal distribution or if there are departures from normality. Additionally,

these measures can guide decisions on the appropriateness of certain statistical methods that assume specific distributional properties.

Skewness quantifies a distribution's asymmetry, where positive skewness suggests a stretched right tail and negative skewness suggests a stretched left tail.

Kurtosis gauges the peakedness or flatness of a distribution, where high kurtosis signals a concentration of data around the mean with heavy tails, while low kurtosis indicates a flatter distribution with fewer outliers (Table S2).

Table S2. Results of descriptive statistics analysis.

Cancer		Healthy Control	
Mean	945.45	Mean	3.39
Standard Error	48.48	Standard Error	0.32
Median	946.76	Median	3.76
Mode	1075.37	Mode	1.18
Standard Deviation	306.60	Standard Deviation	2.05
Sample Variance	94004.56	Sample Variance	4.21
Kurtosis	-0.41	Kurtosis	-1.36
Skewness	0.15	Skewness	0.22
Range	1230.28	Range	6.48
Minimum	340.40	Minimum	0.74
Maximum	1570.68	Maximum	7.22
Sum	37818.02	Sum	135.61
Count	40.00	Count	40.00
Largest(1)	1570.68	Largest(1)	7.22
Smallest(1)	340.40	Smallest(1)	0.74
Confidence Level (95.0%)	98.06	Confidence Level (95.0 %)	0.66

3.3. *Outlier Analysis:*

In the realm of data analysis, outlier analysis holds considerable significance. Outliers refer to data points that exhibit substantial deviations from the general data distribution within a dataset. These outliers wield noteworthy influence over the outcomes of statistical evaluations and can emerge due to diverse factors, including measurement inaccuracies, errors during data input, or instances of genuinely exceptional observations. The process of outlier analysis encompasses the tasks of recognizing, comprehending, and addressing these atypical observations.

In the context of this study, quartile-based statistical analysis was employed to compute and identify outliers within the dataset about cancer patients. Quartiles are values that divide a dataset into four equal parts. The quartiles of a dataset are often denoted as Q1, Q2, and Q3, representing the 25th, 50th, and 75th percentiles, respectively. In our study for outlier analysis, the first quartile (Q1) and third quartile (Q3) were employed, and the interquartile range of the data set was calculated as the difference between the third quartile (Q3) and the first quartile (Q1).

Table S3. Conducting Outlier analysis on data of GAPDH ccfDNA copy number derived from cancer samples via RT-qPCR experimentation.

Data	Outlier	Q1	Q3	Internal QR (IQR)	Upper Limit	Lower limit
340.40	FALSE	754.3406	1105.73	351.389	1632.81	227.26
430.22	FALSE	Results: All the values are within the limit				
454.60	FALSE					
540.02	FALSE					
547.51	FALSE					
594.69	FALSE					
637.10	FALSE					
637.10	FALSE					
682.53	FALSE					
731.20	FALSE					
762.05	FALSE					
772.62	FALSE					
783.34	FALSE					
799.70	FALSE					
850.84	FALSE					
874.61	FALSE					
874.61	FALSE					
911.51	FALSE					
924.16	FALSE					
936.98	FALSE					
956.54	FALSE					
969.81	FALSE					
976.51	FALSE					
996.90	FALSE					
1003.79	FALSE					
1046.15	FALSE					
1060.66	FALSE					
1075.37	FALSE					
1075.37	FALSE					
1090.29	FALSE					
1152.05	FALSE					
1152.05	FALSE					
1251.33	FALSE					
1268.68	FALSE					
1322.21	FALSE					

1359.15	FALSE	
1359.15	FALSE	
1517.51	FALSE	
1528.00	FALSE	
1570.68	FALSE	

4. Generating a Receiver Operating Characteristics (ROC) Curve

To assess the effectiveness of binary classification models, we utilize a visual tool called an ROC curve. This curve provides a graphical representation of the trade-off between two vital metrics: sensitivity (true positive rate) and 1-specificity (false positive rate) across different classification thresholds. In our research experiment, we constructed a ROC curve based on DNA copy numbers derived from the amplification of the GAPDH gene within ccfDNA samples obtained from both cancer and non-cancerous blood plasma, employing RT-qPCR analysis.

Calculation of True Positive Rate (TPR) and False Positive Rate (FPR):

To generate the ROC curve, we computed TPR and FPR, which can be found in Table S4.

True Positive Rate (Sensitivity) is calculated as the ratio of true positives to the sum of true positives and false negatives.

False Positive Rate (1-Specificity) is determined by the formula False Positives / (False Positives + True Negatives).

Area Under the Curve (AUC) Computation: This metric acts as a quantitative evaluation of the ROC curve's effectiveness. It offers a numerical measure of the likelihood that a randomly chosen positive instance will be ranked higher than a randomly chosen negative instance. In our study, we utilized our dataset to build a classifier with an AUC value of 1, indicating excellent discrimination. Further details are available in Table S3.

Table S4. ROC Table.

Bin of Log10 DNA Copy Number	Cancer Blood Plasma	Non-Cancer	FP	TP	FPR	TPR	AUC
10	0	0	1	1	1	1	0
100	0	40	1	0	1	1	0
190	0	0	1	1	0	1	0
280	0	0	1	1	0	1	0.025
370	1	0	0.975	1	0	1	0.025
460	2	0	0.95	1	0	0.975	0
550	2	0	0.95	1	0	0.925	0.025
640	3	0	0.925	1	0	0.875	-0.05
730	1	0	0.975	1	0	0.8	0.1
820	5	0	0.875	1	0	0.775	-0.05
910	3	0	0.925	1	0	0.65	0.1
1000	7	0	0.825	1	0	0.575	-0.05
1090	5	0	0.875	1	0	0.4	-0.05
1180	3	0	0.925	1	0	0.275	-0.025
1270	2	0	0.95	1	0	0.2	0.025
1360	3	0	0.925	1	0	0.15	-0.075
1450	0	0	1	1	0	0.075	0.05

1540	2	0	0.95	1	0	0.075	-0.025
1630	1	0	0.975	1	0	0.025	0.975
	40	40					1

5. Fluorometric Analysis of ccfDNA Concentration in Blood Plasma of both Study Groups

In this study, two distinct approaches were used to compare cancer patients and non-cancer persons' ccfDNA concentration levels. One of these methods utilized fluorometric analysis, which makes use of a light beam in fluorescence spectroscopy to jolt electrons inside molecules, causing them to emit light. The emitted light is directed through a filter and onto a detector to determine whether these molecules are present and how they have changed. We used the Quantus™ Fluorometer in our study to quantify the ccfDNA. The condensed version of the raw data obtained from the fluorometer is shown in Table S5.

Table S5. Concentration levels of ccfDNA in picogram per microliter obtained by Fluorometric analysis.

	Cancer				Non-Cancer			
Serial	Conc1 (pg./µL)	Conc2 (pg./µL)	Conc3 (pg./µL)	Mean	Conc1 (pg./µL)	Conc2 (pg./µL)	Conc3 (pg./µL)	Mean
1	428	410	407	415.00	4.9	4.5	4.2	4.53
2	627	659	682	656.00	3.2	3.8	3.6	3.53
3	698	688	678	688.00	2.9	3.2	3	3.03
4	488	469	489	482.00	3.9	4.4	4	4.10
5	632	610	621	621.00	4.2	4	4.6	4.27
6	459	472	466	465.67	4.5	4.1	4.2	4.27
7	508	492	499	499.67	4.7	4.4	4.8	4.63
8	624	632	641	632.33	3.8	3.2	3.6	3.53
9	550	537	542	543.00	4	3.9	4	3.97
10	536	542	549	542.33	3.5	3.2	3.3	3.33
11	485	495	478	486.00	4.2	3.9	4.3	4.13
12	469	470	475	471.33	4.6	4.2	4.4	4.40
13	422	433	440	431.67	4.7	4.6	4.3	4.53
14	648	632	630	636.67	3.8	3.6	4	3.80
15	420	428	420	422.67	4.3	4.2	3.9	4.13
16	540	541	535	538.67	4.1	3.8	3.8	3.90
17	410	426	436	424.00	3.6	3.4	3.8	3.60
18	525	536	542	534.33	4.8	4.5	4.4	4.57
19	486	478	500	488.00	4.9	4.5	4.8	4.73
20	566	569	581	572.00	4.5	4	4.6	4.37
21	620	635	647	634.00	3.5	3.6	3.8	3.63
22	631	647	655	644.33	4.2	3.9	4.5	4.20
23	499	502	509	503.33	5.3	5.2	4.9	5.13
24	347	352	338	345.67	3.4	3.2	3.6	3.40
25	362	378	386	375.33	4.6	4.2	4.4	4.40
26	628	635	662	641.67	4.8	4.2	4.4	4.47
27	583	592	569	581.33	4.5	4.1	4.3	4.30
28	462	487	475	474.67	3.2	3	3.6	3.27
29	493	469	486	482.67	3.8	3.6	3.9	3.77
30	591	582	571	581.33	3.4	3.2	2.9	3.17

31	462	442	452	452.00	3.5	3.2	3.8	3.50
32	478	462	459	466.33	3.4	3.3	3.6	3.43
33	487	469	475	477.00	4.7	4.2	4.4	4.43
34	503	532	521	518.67	3.9	3.6	3.9	3.80
35	562	532	545	546.33	4.2	4	4.5	4.23
36	469	448	452	456.33	3.3	3	3.6	3.30
37	592	613	632	612.33	3.8	3.5	3.7	3.67
38	607	588	575	590.00	4.6	4.2	4.3	4.37
39	549	562	532	547.67	4.1	4.2	3.7	4.00
40	514	529	533	525.33	5.2	4.9	4.2	4.77

6. Relative Quantification of ccfRNA

Principle: Relative quantification refers to the comparison of the amount of a specific nucleic acid sequence, typically a gene of interest, in one sample to its amount in another sample. This comparison is typically carried out using methodologies such as quantitative real-time polymerase chain reaction (qPCR) or reverse transcription-qPCR (RT-qPCR) by calculating Ct values for gene expression analysis, where lower Ct values are indicative of higher expression or abundance of the target sequence.

A fundamental part of relative quantification is the incorporation of a reference or control sample as a point of comparison with other samples. The reference sample may be representative of a specific time point, a particular treatment, or a normal condition, depending on the experimental design.

To perform relative quantification of ccfRNA, the expression of the GAPDH gene in samples of cancerous blood plasma was compared to samples of non-cancerous blood plasma. Thirty healthy individuals without cancer provided blood plasma for the control group, which contained ccfRNA, and thirty cancer patients provided ccfRNA for the treatment group. To achieve relative quantification, GAPDH was selected as the target gene, while three housekeeping genes—beta-actin (ACTB), ubiquitin C (UBC), and hypoxanthine-guanine phosphoryl transferase (HPRT1)—served as reference genes.

Rt-qPCR involves the determination of quantities, which frequently rely on cycle threshold (Ct) values. The Ct value represents the number of cycles at which the fluorescence signal reaches a specific threshold. Reduced Ct values are indicative of augmented expression or abundance of the targeted genetic sequence. The machine-generated average Ct values for all samples with specific primers are described in Table S6. Relative fold change and normalization stand as pivotal concepts in molecular biology, specifically within the realm of relative quantification in RT-qPCR.

Relative Fold Change:

Formula: Relative Fold Change=$2^{-\Delta\Delta Ct}$

ΔCt (delta Ct) is the difference in cycle threshold values between the target gene and an internal control or reference gene.

$\Delta\Delta Ct$ (delta-delta Ct) represents the difference in ΔCt values between the experimental condition and a reference or control condition.

Interpretation:

- A fold change of 1 indicates no change in expression.
- Values greater than 1 signify upregulation.
- Values less than 1 indicate downregulation.

Normalization:
In the domain of gene expression analysis, normalization refers to adjusting gene expression data to accommodate any variations that might have occurred during the experimental procedures. This particular stage is of utmost importance as it guarantees the precision and significance of comparisons that are made between various samples. The practice of normalization assists in eliminating predispositions and technical anomalies, thereby enabling researchers to concentrate on authentic biological distinctions.
Aligned with the research objective, diverse normalization methods are employed. This study chooses internal controls, commonly known as housekeeping genes, due to their assumed stable expression across varying conditions. These genes establish a foundation for comparison.

Significance:
Normalization accounts for variations in sample quality, RNA/DNA extraction efficiency, and instrument performance.
It enhances the accuracy and reliability of quantitative measurements.

Table S6. The Ct values acquired through RT-qPCR, utilized for the relative quantification of the GAPDH gene in cancerous blood plasma.

SL no	Sample type	Sample Name	Target gene	Reference gene		
			Ct Mean (GAPDH)	Ct Mean (Beta-Actin)	Ct Mean (UBC)	Ct Mean (HPRT1)
1	Control	N-1	27.25	33.26	32.05	31.38
2	Control	N-2	26.91	32.66	32.07	32.14
3	Control	N-3	27.2	34.23	31.23	31.24
4	Control	N-4	27.24	34.48	34.49	32.41
5	Control	N-5	27.26	33.8	32.3	31.81
6	Control	N-6	27.25	34	31.67	32.5
7	Control	N-7	27.33	34.595	32.22	32.57
8	Control	N-8	27.335	34.755	35.32	33.26
9	Control	N-9	27.14	34.625	33.4	32.64
10	Control	N-10	27.15	33.45	32.38	31.45
11	Control	N-11	26.99	32.259	32.3	32.27
12	Control	N-12	27.02	34.19	31.41	32.2
13	Control	N-13	27.17	32.61	32.23	32.61
14	Control	N-14	26.3	32.28	31.51	32.5
15	Control	N-15	27.16	32.76	31.65	32.76
16	Control	N-16	27.15	31.89	31.6	31.87
17	Control	N-17	27.25	31.39	32.51	31.44
18	Control	N-18	27.12	31.64	32.56	31.64
19	Control	N-19	26.52	31.42	31.28	31.42
20	Control	N-20	26.51	32.01	31.72	32.11
21	Control	N-21	25.97	31.66	32.28	31.6
22	Control	N-22	26.73	32.11	34.54	30.13
23	Control	N-23	27.17	34.15	32.57	32.15
24	Control	N-24	27.27	33.33	31.65	32.33
25	Control	N-25	26.61	31.66	32.78	31.6
26	Control	N-26	27.05	33.19	31.76	32.18
27	Control	N-27	26.30	33.08	32.39	29.13

28	Control	N-28	27.51	34.25	31.4	32.35
29	Control	N-29	27.22	33.15	32.67	32.3
30	Control	N-30	26.48	34.51	33.36	31.53
1	Cancer	C-1	23.03	27.23	28.66	29.25
2	Cancer	C-2	23.55	26.14	28.65	29.6
3	Cancer	C-3	23.48	27.2	28.81	27.27
4	Cancer	C-4	24.10	27.5	28.93	27.54
5	Cancer	C-5	23.50	27.21	29.13	27.35
6	Cancer	C-6	23.90	27.25	28.75	27.2
7	Cancer	C-7	23.02	27.34	28.43	27.44
8	Cancer	C-10	23.44	27.34	29.21	27.54
9	Cancer	C-11	24.34	27.36	28.65	27.48
10	Cancer	C-14	23.91	27.51	28.45	27.53
11	Cancer	C-16	23.62	27.06	30.56	26.99
12	Cancer	C-19	23.53	27.07	28.21	27.13
13	Cancer	C-20	23.10	27.17	28.36	27.34
14	Cancer	C-24	23.02	26.31	28.57	26.45
15	Cancer	C-27	24.4	27.61	28.48	27.6
16	Cancer	C-28	23.03	27.63	28.18	27.81
17	Cancer	C--35	23.95	27.41	28.6	27.25
18	Cancer	C-37	24.43	27.27	28.46	27.43
19	Cancer	C-39	23.79	26.52	28.5	27
20	Cancer	C-9	24.44	26.49	28.43	26.4
21	Cancer	C-8	23.12	25.55	28.6	26.95
22	Cancer	C12	24.19	26.73	28.66	26.73
23	Cancer	C-13	23.8	27.41	28.6	27.17
24	Cancer	C-15	23.73	27.27	28.73	27.17
25	Cancer	C-17	24.05	30.4	28.28	26.61
26	Cancer	C-21	23.5	27.05	28.58	27.05
27	Cancer	C-22	23.09	26.1	29.39	26.3
28	Cancer	C-23	23.5	28.26	29.48	27.81
29	Cancer	C-24	23.58	27.34	28.58	27.22
30	Cancer	C-25	23.62	26.49	28.61	26.5

Reference

1. J. M Gallup, How to determine copy number of a gene via Absolute Real Time PCR? Copy number question RG 2-9-16 J MG.xls (2016). https://www.researchgate.net/post/How_to_determine_copy_number_of_a_gene_via_Absolute_Real_Time_PCR