Absolute dPCR Quantification of MicroRNAs by Absolute dPCR for the Diagnostic Screening of Colon Cancer

Farid E Ahmed*

Department of Radiation Oncology, Leo W. Jenkins Cancer Center, The Brody School of Medicine at East Carolina University (ECU), USA

Short Mini-Review

Colorectal Cancer (CRC) is the third most common malignancy worldwide, with an estimated one million new cases and half million deaths yearly. Screening for CRC allows early stage diagnosis of malignancy and potentially reduces disease mortality. The convenient and inexpensive Fecal Occult Blood Test (FOBT) screening test has low sensitivity and requires dietary restriction, which impedes compliance. Although colonoscopy. Is the golden screening standard for this cancer, the invasive nature, abdominal pain and high cost have hampered worldwide application of this procedure. A noninvasive sensitive screen for Colon Cancer (CC) requiring no dietary restriction is a more convenient test. CC is more abundant in the USA than Rectal Cancer (RC).

The discovery of small non-coding protein sequences, 17-27 nucleotides long RNAs (such as microRNAs), has opened new opportunities for a non-invasive test for early diagnosis of many cancers. MiRNA functions seem to regulate development and apoptosis, and specific miRNAs are critical in oncogenesis, effective in classifying solid and liquid tumors, and serve as oncogenes or suppressor genes. MiRNA genes are frequently located at fragile sites, as well as minimal regions of loss of heterozygosity, or amplification of common break-point regions, suggesting their involvement in carcinogenesis. Profiles of miRNA expression differ between normal tissues and tumor types, and evidence suggests that miRNA expression profiles can cluster similar tumor types together more accurately than expression profiles of protein-coding mRNA genes. Although exosomal RNA are missed, a parallel carried out on stool miRNAs to compare the extent of loss when colonocytes are only used can be carried out, and an appropriate correction for exosomal loss can be made. To ascertain the validity of a miRNA screening test for CC, it must be validated in a study, using a nested case control epidemiology design and employing a prospective specimen collection, retrospective blind evaluation (PRoBE) of control subjects and test colon cancer patients, as delineated by NCI's Early Detection Research Network (EDRN) http://edrn.nci.nih.gov. Immunoparamagnetic are employed to capture colonocytes from harsh stool environment,

Table 1: Characteristics of fourteen up-/down- regulated microRNAs in human stool.

MiRNA	Up-Regulated	Down-Regulated	Chromosome Location	Known Putative Cancer Target Gene(s)
MiR-19a	Yes	No	13q31.3	Undetermined
MiR-20a	Yes	No	13q31.3	PTEN, TMP1
MiR-21	Yes	No	17q23.1	PTEN, BCL2, PDCD4, TIMP3, SPRY2, REC, T1AM1
MiR-31	Yes	No	9p21.3	T1AM1, AX1N1, FOXC2, FOXP3, H1F1AN
MiR-34a	Yes	No	1p36.22	BCL2, TP53, E2F3, NOTCH1, E2F1, S1RT
MiR-96	Yes	No	17q32.2	KRAS
MiR-106a	Yes	No	Xq26.2	PTEN, E2F1, RB1
MiR-133a	Yes	No	18q11.2/20q13.33	BAX, KRAS
MiR-135b	Yes	No	1q32.1	MSH2
MiR-200c	Yes	No	12p13.31	ZEB1
MiR-224	Yes	No	Xp23	Undetermined
MiR-30a	No	Yes	6q13	RASA1, ERG, SEMA6D, SEMA3A
MiR-143	No	Yes	5q32	KRAS, MAPK7, DNMT3A
MiR-145	No	Yes	5q32	TGFBRE, APC, IRS1, STAT1, YES1, FLI1

Table 2: Absolute quantification of up-/down- regulated miRNAs in stool by QuantStudio™ 3D chip-based digital PCR.

Type	miR-19a	miR-20a	miR-21	miR-31	miR-34a	miR-96	miR-106a	miR-133a	miR-135b	miR-200c	miR-224	miR-30a	miR-143	miR-145
Control	9964.23	9724.14	9699.68	9591.16	9580.92	9590.59	9464.64	9574.13	9568.15	9556.85	9631.73	9401.81	9585.54	9683.18
Control	9984.55	9890.38	9795.44	9588.24	9602.9	9587.82	9592.68	9680.24	9515.46	9511.29	9592.62	9580.92	9504.61	9506.12
Control	9950.19	9898.88	9938.74	9791.83	9894.82	9862.24	9875.88	9800.08	9824.18	9843.18	9810.2	9780.74	9699.52	9823.54
Stage 1	7998.16	8011.92	7949.68	7864.18	7880.18	7790.44	7682.74	7687.88	7561.64	7402.8	6994.24	6892.54	1995.92	1884.54
Stage 1	7814.22	7901.24	7890.32	7798.92	7780.28	6849.68	6999.68	6742.6	6640.16	6616.1	6872.54	6640.24	1879.04	1764.92
Stage 1	7764.5	7745.38	7690.32	7549.28	7610.32	6787.62	6870.96	6739.42	6690.82	6584.74	6477.52	6454.44	1799.92	1668.19
Stage 2	7414.42	7569.16	7529.9	7492.68	7384.82	7189.64	6794.88	6690.98	6504.2	5702.16	5464.16	4870.22	1346.48	1040.26
Stage 2	7390.84	7490.96	7501.62	7379.04	7202.28	7102.28	6472.48	6598.24	6242.82	4387.76	5414.08	4189.42	988.14	862.08
Stage 2	7208.16	7378.74	7402.68	7299.76	7124.56	7098.04	6402.18	6401.16	6218.92	4123.18	4098.78	3894.9	872.4	763.14
Stage 3	6850.14	6936.16	6902.04	6890.14	7092.18	6586.18	6319.08	5898.36	5386.66	3821.22	3679.62	3601.4	365.42	256.28
Stage 3	6792.75	6790.29	6776.26	6658.78	6674.54	6560.68	6116.84	5602.16	4999.16	3715.22	3686.92	3570.92	260.14	154.02
Stage 3	6622.84	6662.9	6694.28	6558.84	6554.28	6510.27	6039.84	5404.68	5498.82	3421.22	3614.62	3120.18	194.84	133.37
Stage 4	6506.92	6538.8	6419.02	6227.54	5978.48	5766.32	5686.36	5256.81	4973.28	3327.28	3479.52	2052.38	92.45	88.49
Stage 4	6468.22	6384.12	6397.92	6117.12	5856.66	5681.82	5259.84	4905.76	3840.86	3244.16	3276.42	1096.44	76.88	67.42
Stage 4	6488.38	6434.48	6346.06	5898.78	5466.16	5372.56	4896.36	4812.44	3784.56	3164.8	3186.14	678.56	56.82	49.26

Figure 1: Diagram illustrating QuantStudio™ 3D digital PCR system chip; (1) Chip Case Lid; (2) Digital PCR 20K 10 mm² nanofluidic v2 chip, which contains 20,000 reaction wells; (3) QuantStudio™ 3D Digital PCR Chip Case; (4) Chip ID; (5) Fill port; and (6) Reaction wells, the 20,000 physical holes that suspend individual PCR reactions.
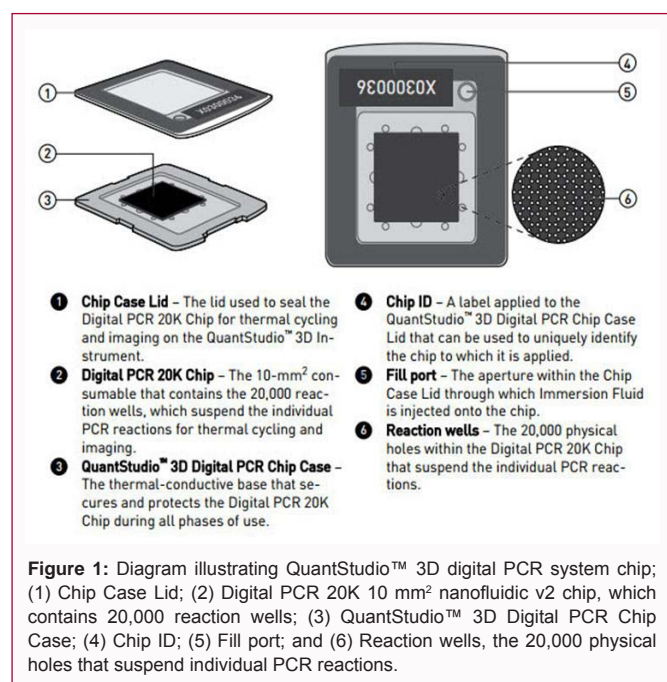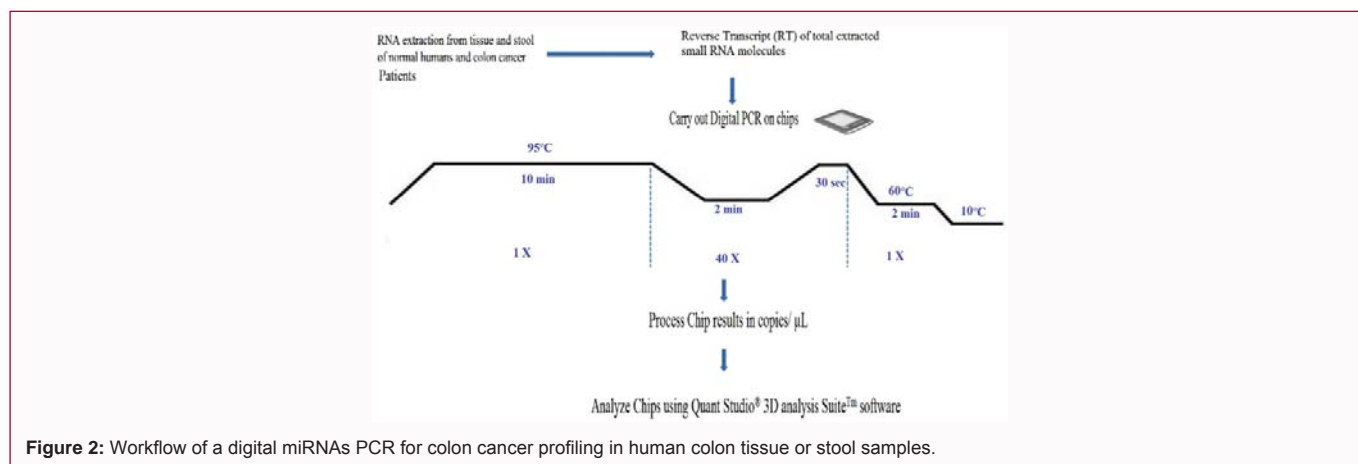
whose extracted fragile total small RNA is stabilized shortly after stool excretion by commercial kits so it does not ever fragment, followed by standardized analytical quantitative miRNA dPCR-chip profiling in non invasive stool samples, to develop a panel of few stable miRNAs for absolute quantitative diagnostic screening of early sporadic colon cancer (stage 0-1), more economically and with higher sensitivity and specificity than other CC screening test on the market today.

A preliminary global microarray expression analysis using an exfoliated colonocytes enrichment strategy, which employed control subjects and various stages (0-4) of CC, using Affymetrix GeneChip miRNA 2.0 Array, showed 180 preferentially expressed miRNA genes that were either increased (124 miRNAs), or reduced (56 miRNAs) in expression in stool samples from CC patients. This allowed careful selection of 14 miRNAs (12 Up-Regulated, miR-19a, miR-20a, miR-21, miR-31, miR-34a, miR-96, miR-106a, miR-133a, miR-135b, miR-206, miR-224 and miR-302; and 2 Down-Regulated, miR-143 and miR-145) for further PCR analysis (Table 1).

Then analysis carried out using absolute miRNAs expression by a chip-based digital PCR by partitioning a sample of DNA or cDNA into many individual, parallel PCR reactions; some of which contain the target molecule (positive), while others do not (negative). A single molecule can be amplified a million-fold or more. During amplification, TaqMan chemistry with dye-labeled probes is used to detect sequence-specific targets. When no target sequence is present, no signal accumulates. Following dPCR analysis, the fraction of negative reactions is used to generate an absolute count of the number of target molecules in the sample, without the need for standards or endogenous controls. In conventional qPCR, the signal from wild-type sequences dominates and obscures the signal from rare sequences. By minimizing the effect of competition between targets, dPCR overcomes the difficulties inherent to amplifying rare sequences and allows for sensitive & precise absolute quantification of the selected miRNAs. Applied Biosystem QuantStudio™ 3D instrument only performs the imaging and primary analysis of the digital chips. The chips themselves must be cycled offline on a Dual Flat Block GeneAmp® 9,700 PCR System or the ProFlex™ 2 × Flat PCR System. The QuantStudio™ 3D Digital PCR System can read the digital chip in less than 1 minute, following thermal cycling (Figure 1). The current QuantStudio™ 3D Digital PCR Chip allows for one sample per chip; although, duplexing allows for analysis of two targets per chip. Sample prep for digital PCR is no different than for real-time PCR, when using the QuantStudio™ 3D Digital PCR System. The concentration of cDNA stock can be estimated by including all of the necessary dilution factors into the AnalysisSuite™ software, which gives the copies/μL in the stock. A critical step in dPCR, is sample partitioning [i.e., division of each sample into thousands of discrete subunits prior to amplification by PCR, each ideally containing either zero or one (or at most, a few) template molecules]. Each partition behaves as an individual PCR reaction-as with real-time PCR—fluorescent FAM probes [or others, as VIC fluorescence. Samples containing amplified products are considered positive (1, fluorescence), and those without product–with little or no fluorescence are negative (0, fluorescence). The ratio of positives to

Figure 2: Workflow of a digital miRNAs PCR for colon cancer profiling in human colon tissue or stool samples.

Table 3: Representation of SDs and r^2 for miRNAs tested by absolute digital PCR.

| Type	miR-19a	miR-20a	miR-21	miR-31	miR-34a	miR-96	miR-106a
sd	92.239	111.10331	99.76355	146.64101	209.04905	278.47558	301.87638
r^2	99.4831	99.18486	99.34603	98.65141	97.63002	96.13899	96.19772
Type	miR-133a	miR-135b	miR-200c	miR-224	miR-30a	miR-143	miR-145
sd	300.06189	409.67168	449.86741	376.84372	424.99723	132.76331	110.89266
r^2	96.85741	95.49454	96.70427	97.61795	97.95389	99.87075	99.91289

negatives in each sample is the basis of amplification. Unlike real-time qPCR, dPCR does not rely on the number of amplification cycles to determine the initial amount of template nucleic acid in each sample, but it relies on Poisson Statistics to determine the absolute template quantity. The unique sample partitioning step of dPCR, coupled with Poisson Statistics, allows for higher precision than both traditional and qPCR methods; permitting for analysis of rare miRNA targets. The use of a nanofluidic chip provides a convenient mechanism to run thousands of PCR reactions in parallel. Each well is loaded with a mixture of sample, master mix, and Applied Biosystems TaqMan Assay reagents are individually analyzed to detect the presence (positive) or absence (negative) of an endpoint signal. To account for wells that may have received more than one molecule of the target sequence, a correction factor is applied using the Poisson model. It features a filter set that is optimized for the FAM™, VIC®, and ROX™ dyes, available from Life Technologies.

Absolute quantification of the 14 miRNAs is shown in Table 2 and 3 is a representation of SDs and and R² for the 14 miRNAs tested by absolute digital PCR. Figure 2 is Workflow of a digital miRNAs PCR for colon cancer profiling in human colon tissue or stool samples. Figure 3 is a graphical representation of the absolute quantification of the 12 up- or 2 down-regulated miRNAs in Human Stool by the QuantStudio™ 3D Digital PCR Chip System. Digital PCR, however, needs a rough estimate of the concentration of miRNAs of interest to carry out first, in order to make appropriate dilutions; Non-template controls and a RT negative control must be set up for each miRNA, when using a "primer pool method" for retro-transcription; a chip-based dPCR method requires less pipetting steps, which reduces potential PCR contamination, and Quant Studio™ 3D chip has 20,000 fixed reaction wells, which reduces variability of dPCR results.

To avoid bias, and ensure that biomarker selection and outcome assessment will not influence each other, a prospective specimen collection retrospective blinded evaluation (PRoBE) design randomized selection could be employed. An enrichment and exfoliation strategy of colonocytes from stool for miRNA profiling using Dynal superparamagnetic polystyrene beads coated with a mouse IgG1 monoclonal antibody (Ab) Ber-Ep4, specific for an epitope on the protein moiety of the glycopolypeptide membrane antigen Ep-CAM, which is expressed on the surface of colonocytes and colon carcinoma cells, can be used. Comparing the Agilent electrophoretic (18S and 28S) patterns to those obtained from total RNA extracted from stool, and differential lysis of colonocytes by RT lysis buffer (Quagen), could be construed as a validation that the electrophoretic pattern observed in stool (18S and 28S) is truly due to the presence of human colonocytes, and not due to stool contamination with *Escherichia coli* (16S and 23S). While some exosomal RNA can be released from purified colonocytes into stool, correction for that effect can be made. Hence, for CRC screening, miRNA markers are more comprehensive and preferable to a DNA-, epigenetic-, mRNA- or a protein-based marker. An added advantage of the use of the stable, non-degradable miRNAs by PCR expression, or chip-based methods is being automatable, making them more economical and acceptable by laboratory personnel performing these assays.

Stool testing has several advantages over other colon cancer screening methods as it is truly noninvasive and requires no unpleasant cathartic preparation, formal health care visits, or time away from work or routine activities. Unlike sigmoidoscopy, it reflects the full length of the colorectum and samples can be taken in a way that represents the right and left side of the colon. It is believed that colonocytes are released continuously and abundantly into the fecal stream, contrary to blood that is released intermittently as in guaiac FOBT; therefore, this natural enrichment phenomenon partially obviates the need to use a laboratory-enrichment technique to enrich for tumorigenic colonocytes, as for example when blood is used for testing. Furthermore, because testing can be performed on mail-in-specimens, geographic access to stool screening is essentially unimpeded. Results have show that even the presence of bacterial
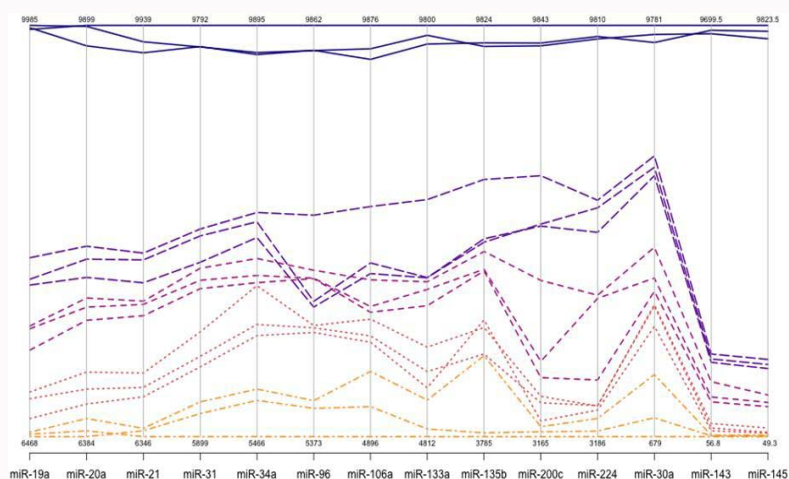
Figure 3: Absolute quantification of Up-/Down- regulated miRNAs in human stool by QuantStudio™ 3D digital PCR chip system.

DNA, non-transformed RNA and other interfering substances in stool does not interfere with measuring miRNA expression, when an enrichment method such as the immunological paramagnetic capture method is used, and when appropriate PCR primers are employed. Besides, stool colonocytes contain much more miRNA than that available in free circulation such as in plasma.

Routine extraction of total small RNA from Stool Samples can be carried out using commercial RNA extraction preparations, which provide the advantage of manufacturer's established validation and quality control standards, increasing the probability of good results to extract high quality total RNA from an environment as hostile as stool; thus, shattering the myth that it is difficult to employ RNA as a screening substrate, as fragmented RNA results in poor cDNA synthesis and ultimately in less than optimal PCR amplification. This step is followed by Reverse Transcription (RT) and preparation of single stranded copy Deoxyribonucleic Acid (ss-cDNA). Compared to qPCR, dPCR clearly offers more sensitive and considerably more reproducible clinical methods that could lend themselves to diagnostic, prognostic, and predictive tests. The technology, however, needs further development to reduce cost and simplify application, and research reported with understanding of associated errors.

The term "absolute quantification" used in dPCR refers to an estimate derived from the count of the proportion of positive partitions relative to the total number of partitions and their known volume. When the sample is sufficiently dilute, most partitions will not contain template and those that do are most likely to contain single molecules. As the sample becomes more concentrated, the chance of more than 1 molecule being present within a positive partition increases. This does not pose too great of a challenge, because the distribution of molecules throughout the partitions approximates a Poisson distribution, and a Poisson correction is applied. The dynamic range of a dPCR assay can extend beyond the number of partitions analyzed but the assay precision deteriorates at each end. In contrast, qPCR precision deteriorates only at low copy numbers.

dPCR benefits from a far more predictable variance than qPCR, but dPCR is susceptible to upstream errors associated with factors like sampling and extraction. dPCR can also suffer systematic bias, particularly leading to underestimation, and internal positive controls are likely to be as important for dPCR as they are for qPCR, especially when reporting the absence of a sequence. Calibration curves are frequently employed to reduce the error associated with qPCR, but they in turn are challenging to select, value assign and apply in a manner that will be reproducible; their application also contains inherent error that is almost never considered. A key problem with applying qPCR to areas such as the discovery of biomarkers that will eventually be translated to clinical care, understanding whether poor reproducibility is biological, or if it is due to issues related the fact the qPCR technique is difficult to perform reproducibly. Implement dMIQUE Guidelines on dPCR data analysis helps to standardize experimental protocol, maximize inefficient utilization of resources, and enhance the impact of this technology. To access test Performance Characteristics (PCT) of the miRNA approach, the copies/μl values of the miRNA gene panel (or a derived miRNA index, IMP) obtained from stool/colonelcy samples of normal subjects and CC patients with high sensitivity and specificity are compared to the commonly used guardian FOBT test and with endoscopy results obtained from patients' medical records to access TPC of the miRNA approach.

False positive discovery rates (expected proportion of incorrect assignment among the accepted assignments) can be assessed by statistical methods, as it could reflect on the cost effectiveness of our test. The number of optimal miRNA genes (whether 14 or less) to achieve an optimum gene panel can be established statistically.

Providing numerical underpinning of dPCR as a function of total small RNA can be ascertained by use of cytological methods on purified colonocytes, employing Papanicolaou and Giemsa staining. A known number of the colonocytes isolated from 1 g stool (from normal and neoplastic preparations), extracting total RNA from them to determine the actual amount of total RNA per stool sample, and determining the average copies/μl value from the panel of selected miRNAs from dPCR using QuantStudio™ 3D Digital PCR Chip instrument, give an average value per a certain amount (pg or ng) of total RNA. Statistical methods, both parametric and non parametric methods can be used for validating the miRNA approach. Cross-validation can be used to protect against over-fitting. Bioinformatics methods can be used to correlate seed miRNA with target mRNA data.

To measure the clinical utility of gene expression testing as a screening test, the diagnostic sensitivity and specificity of the miRNA gene panel selected, or a derived PMI, can be compared to the published sensitivity and specificity of the commonly employed

diagnostic screening test, guaiac FOBT, which for over 3 decades in large adenoma averaged <12%, and in carcinoma averaged ~30%, and the specificity averaged ~95%, and to the gold screening standard "colonoscopy" results obtained from participants' medical records that averaged 87% for sensitivity and 100% for specificity. The limitations of FOBT are biologically inescapable and cannot be reversed by technological advances. On the other hand, we can screen for per-malignant stages of CC, with >90% sensitivity and >95% specificity, employing 14 miRNA genes in a functional assay, which is better than any available noninvasive test. Thus, a large number of patients will be spared the discomfort, risk and expense of screening colonoscopy. Only those patients truly at risk of having CC will need to undergo colonoscopy.

Using the copies/µl results from the panel of genes selected obtained from stool samples of normal, and from stool samples of cancer patients. A 2 × 2 tables can be used to predict a miRNA index to determine the clinical sensitivity and specificity of the miRNA assay from mRNAs stool specimens' results. dPCR is the most practical, least labor-intensive and economical approach to quantify miRNA as a noninvasive diagnostic test to screen for CC, using automatic RNA extraction technology. Although other methods such as qPCR, Next Generation Sequencing (NGS) technologies, plate assay technology to study miRNA expression, and miRNA measurements from exosomes and microvesicles extracted from stool, can provide alternate standardized methods for miRNA quantification, these methods will not be as precise, nor as economical as using dPCR technology.

Acknowledgment