CANCER GENOMICS & PROTEOMICS *4*: 1-20 (2007)

Transcriptomic Molecular Markers for Screening Human Colon Cancer in Stool and Tissue

FARID E. AHMED[1*], PAUL VOS[2], STEPHANIE IJAMES[3], DONALD T. LYSLE[3], RON R. ALLISON[1], GORDON FLAKE[4], DENNIS R. SINAR[5], WADE NAZIRI[6], STEFAN P. MARCUARD[7] and RODNEY PENNINGTON[8]

Departments of [1]Radiation Oncology, LSB 003, Leo W. Jenkins Cancer Center and [5]Internal Medicine, The Brody School of Medicine (BSOM) at East Carolina University (ECU), Greenville, NC 27858; [2]Department of Biostatistics, School of Allied Health Sciences, ECU, Greenville, NC 27858; [3]Department of Psychology, University of North Carolina, Chapel Hill, NC 27599; [4]Laboratory of Experimental Pathology, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709; [6]Carolina Physicians PA, Greenville, NC 27834; [7]Carolina Digestive Disease, Greenville, NC 27858; [8]Roche Applied Science, Indianapolis, IN 46250, U.S.A.

Abstract. *There is a need for sensitive and specific diagnostic molecular markers that can be used to monitor early patterns of gene expression in non-invasive exfoliated colonocytes shed in the stool, and in situ in adenoma-carcinoma epithelium of the colon. RNA-based detection methods are more comprehensive than either DNA-, protein- or methylation-based screening methods. By routinely and systematically being able to perform quantitative gene expression studies on these samples using less than ten colon cancer genes selected by the enormous resources of the National Cancer Institute's Cancer Genome Anatomy Project, we were able to monitor changes at various stages in the neoplastic process, allowing for reliable diagnostic screening of colon cancer particularly at the early, pre-malignant stages. Although the expression of some of the genes tested in tissue showed less variability in normal or cancerous patients than in stool, the stool by itself is suitable for screening. Thus, a transcriptomic approach using stool or tissue samples promises to offer more sensitivity and specificity than currently used molecular screening methods for colon cancer. A larger prospective clinical study utilizing stool and tissue samples derived from many control and colon cancer patients, to allow for a statistically valid analysis, is now urgently required to determine the true sensitivity and specificity of the transcriptomic screening approach for this preventable cancer.*

Abbreviations: ATCC, American Type Culture Collection; cDNA, copy deoxyribonucleic acid; CD, Crohn's disease; CEA, carcinoembryonic antigen; CGAP, Cancer Genome Anatomy Project; CP, comparative cross point; CRC, colorectal cancer; CT, computed tomography; DEPC, Diethyl pyrocarbonate; DGED, Digital Gene Expression Displayer; E-Method, also referred to as Second Derivative Maximum or CP method; EST, expressed sequence tag; FOBT, fecal occult blood test; GI, gastrointestinal; GLS, Gene Library Summarizer; H&E, Hematoxylin and Eosin staining; IBD, inflammatory bowel disease; LCM, laser capture microdissection; mRNA, messenger ribonucleic acid; NCI, National Cancer Institute; OR, odd ratio; QC, quality control; RT-qPCR, reverse transcriptase quantitative polymerase chain reaction; rRNA, ribosomal ribonucleic acid; SAGE, Serial Analysis of Gene Expression; ss, single stranded; UC, ulcerative colitis; UDG, uracil-DNA glycosylase.

Correspondence to: Dr. Farid E. Ahmed, Clinical Professor of Molecular Oncology, The BSOM at ECU, Greenville, NC 27858, U.S.A. Tel: +252 744 4636, Fax: +252 744 3775, e-mail: ahmedf@ecu.edu

Key Words: Adenocarcinoma, colonocyte, diagnosis, IBD, LCM, genes, QC, RT-qPCR, RNA.

Colorectal cancer (CRC) is the second and third most common malignancy in men and women, respectively, in developed and developing countries, including the United States of America (USA) (1, 2). In the USA, an estimated 106,680 cases of colon and 41,930 cases of rectal cancer are expected to occur in 2006, of which 55,170 deaths are estimated to materialize, which account for ~10% of all cancer deaths. Incidence rates decreased by 1.8% per year during 1998-2002 in the USA (Table I), partially reflecting increased screening and polyp removal during colonoscopy screening. CRC incidence rates have been steadily

Table I. *Estimated new and Existing CRC cases, deaths and ncidence rates in the ten States in the USA with the greatest prevalence*[†].

State	New CRC cases in 2006#	Estimated CRC deaths in 2006*	CRC Incidence rates, 1998-2002**		CRC Death rates, 1998-2002***	
			Male	Female	Male	Female
California	14,820	5,500	57.2	42.0	21.0	15.2
Florida	9,970	3,700	64.6	47.9	22.4	15.8
New York	9,540	3,540	71.9	52.8	26.2	18.4
Texas	9,510	3,530	58.5	41.1	23.7	16.2
Pennsylvania	8,000	2,970	74.6	52.3	28.0	19.4
Illinois	6,760	2,510	72.1	51.0	28.0	19.1
Ohio	6,730	2,500	67.1	94.2	27.6	19.2
Michigan	4,930	1,830	64.6	47.9	24.7	17.0
New Jersey	4,850	1,800	75.5	53.7	27.4	19.4
Virginia	3,690	1,370	58.5	41.1	25.0	18.2

[†]Source: Modified from Reference (2). #Estimated new colon cancer cases by sex for all locations in the USA in 2006 are: 106,680 for both sexes (49,220 for males and 57,460 for females). Estimated new colon cancer deaths by sex for all sites in the USA in 2006 are 55,170 for both sexes (27,870 for males and 27,300 for females). All estimates are rounded to the nearest 10. *Rounded to the nearest ten. **Per 100,000, age adjusted to the 2000 USA standard population. ***Per 100,000, age-adjusted to the 2000 US standard population.

decreasing since 1985, from 66 to 52 per 100,000 in 2002 (3). Mortality rates from CRC have also decreased in both genders over the past two decades at an average of 1.8% per year, reflecting declining incidence rates and improved survival. The 1- and 5-year relative survival for CRC for all stages combined is 83% and 64%, respectively. Survival continues to decline beyond 5 years to 5% at 10 years after diagnosis (2); thus, early detection contributes significantly to the prevention of death from this cancer (4-7).

The most commonly used screening tests in the USA for colon adenomas in men and women, aged ≥50 years old, are the fecal occult blood test (FOBT) and colonoscopy. The former, although convenient and relatively inexpensive, suffers from low sensitivity, whereas the latter – considered the gold standard for CRC screening – is expensive and requires cathartic preparation and patient sedation, which has resulted in a low rate of compliance (8). Computed tomographical (CT) colonography (virtual colonoscopy), which does not require bowel preparation may be considered an adequate alternative only for asymptomatic, not at risk individuals, and only if it can effectively improve the detection of small lesions (9, 10).

Cells are continuously shed by colon tumors in the lumen of the gastro intestine (GI) (*i.e.*, approximately 10^{10} normal adult colonic epithelial cells, each having a lifespan of 3-4 days, are shed daily from the lower two thirds of colon crypts) (11, 12), and their detection in the stool has allowed for the employment of mutation or other functional genomic techniques in their study (1, 6, 13-15). Since CRCs exhibit genetic heterogeneity, multitarget approaches employing mutations in *K-ras*, *APC* and *p53*, the microsatellite instability marker Bat-26 and "long" DNA (representing DNA of nonapoptotic colonocytes characteristic of cancer cells exfoliated from neoplasms, but not normal apoptotic colonocytes) have been examined and undergone clinical testing (13, 16). However, DNA alterations were disappointedly detected in only 16 of 31 (51.6%) invasive cancer, 29 of 71 (40.8%) invasive cancer plus adenoma with high-grade dysplasia, and 76 of 418 (18.2%) in patients with advanced neoplasia (tubular adenoma ≥1 cm in diameter, polyps with high grade dysplasia, or cancer) (17). Protein-based methods are not suited for screening and early diagnosis because they are generally not specific, although they may be of more value as prognostic markers (18). More recently, by employing commercial preparations, we have overcome RNA's lability by stabilizing it within a short period of time after samples (*e.g.*, stool, tissue or blood) were removed from the body, resulting in a total RNA that was readily reverse-transcribable by another commercial preparation making a high quality single-stranded (ss) copy (c) DNA suitable for expression profiling (19-21). Therefore, the identification of new transcriptomic molecular markers with high sensitivities and specificities in exfoliated stool samples is now possible.

Materials and Methods

Adenocarcinoma cell line and culture conditions. Adenocarcinoma cell line HT-29 is used for validating the range of gene expression measurements in stool spiking studies. Cells were obtained from the American Type Culture Collection (ATCC), Rockville, MD, USA. The cells were propagated in Iscore's Modified Dulbeccos medium (IMDM) (Sigma, St. Louis, MO, USA), supplemented with 100 ml/L fetal calf serum, 105 IU/L penicillin and 0.1 g/L

streptomycin in an atmosphere of 5 % CO_2 in a humidified incubator kept at 37˚C. Cultures were passed twice per week as per ATCC recommendations.

Acquisition of clinical specimens. Stool and tissue samples were obtained from twenty control subjects and thirty patients with various stages of colon adenocarcinoma (Dukes' stages 0 to 3), five patients with ulcerative colitis (UC) and 5 patients with Crohn's disease (CD) according to an approved ECU Medical Center Institutional Review Board (IRB) protocol. All laboratory work was carried out and standardized under blind conditions and in accordance with the guidelines for handling biohazardous material established by ECU's Biological Safety and Hazardous Substance Committee.

Control stool and tissue samples.
(i) Fecal specimens. Control stool samples were collected from consenting individuals visiting our GI Clinic/Endoscopy Lab who did not show any polyps or inflammatory bowel diseases, such as colitis or diverticulitis. Stool samples were either processed immediately to extract RNA, or stored overnight at 4˚C in a bacteriostatic preservative S.T.A.R medium (Roche Diagnostics, Indianapolis, IN, USA), and RNA was extracted within a few days. In situations where longer storage of fecal specimens was desired, the preservative RNA*Later*®-ICE (Ambion, Austin, TX, USA) was added at 2.5 ml per 1 g of stool, followed by freezing of the stool sample at –70˚C.
(ii) Tissue specimens. Normal tissues were usually obtained from a small piece of colon tissue (about 0.5 cm^3) removed >10 cm away from diseased patient tissue at surgery (22), or from biopsies taken during colonscopy from non-diseased areas of consenting individuals. For UC or CD patients, a small piece of tissue taken further away from the inflamed or diseased tissue was considered normal. Tissues were flash frozen in liquid nitrogen and stored at –70˚C for subsequent laser capture microdissection (LCM) work. Longitudinal sectioning of the tissue before LCM use was employed in order to pick up only the epithelial cells that would eventually be shed as colonocytes into the lamina propria from the bottom of epithelial cells among the proliferative enterocyte crypt lineage.

Experimental stool and tissue samples from cancerous or inflamed patients.
(i) Fecal specimens. A 10 g sample of feces (bowel movement) was collected the night before surgery or earlier, before administering any bowel preparation, in a plastic container containing either: a) a bacteriostatic preservative S.T.A.R. medium, which was then covered and either processed immediately to extract total RNA or stored overnight at 4˚C then processed the next day for RNA extraction, or b) RNA*Later*®-ICE to allow longer storage at –70˚C, then the sample followed either RT-PCR processing or storage of the extracted RNA at –70˚C until further manipulation and PCR analysis. Stool processing was standardized for all samples by scraping and employing the surface mucinous layer, which is usually rich in colonocytes (13).
(ii) Tissue specimens. A small piece of tissue sample (about 0.5 cm^3) was obtained after colonoscopy for adenoma, or at surgery for carcinoma. Samples were processed after flash freezing in liquid N_2 and storage at –70˚C for subsequent microdissection. Longitudinal LCM sectioning was performed (Figure 1A, B), and the marked areas of the crypt indicated where the transformed cells (*i.e.*, adenoma, carcinoma) were to be captured by laser microdissection (Figure 1C-E) for subsequent RNA extraction.

For the current study, stool and normal tissue samples were obtained from 20 non-cancerous non IBD control individuals, 20 patients having adenomatous polyps ≥1 cm with high grade dysplasia (stage 0-1), 5 patients with stage 2 carcinoma, five patients with stage 3 carcinoma, five non-cancerous patients with severe UC, and five non-cancerous patients with severe CD. Each subject provided a stool sample for a total of 60 stool samples. Tissue samples were obtained from only one of the UC patients and only one of the CD patients, but were obtained for each of the remaining patients, for a total of 52 tissue samples.

Selection of cancerous or inflamed cells from colon tissue of patients by LCM. LCM was employed as an enrichment technique for tumors isolated from colon adenocarcinoma patients to separate transformed cells from nonneoplastic stromal and inflammatory cells (23). The frozen tissues, embedded in Tissue Tek OCT compound (Sakura, MI, USA) were transported in cold packs to Laboratory of Experimental Pathology, National Institute of Environmental Health Sciences (NIEHS), Research Triangle Park, North Carolina. There, LCM was performed using an Arcturus PixCell II system (Arcturus Engineering, Inc., Mountain View, CA, USA), which employed a 15 μm diameter infrared (IR) laser pulse (220 mV, 49 mW) with a duration of 2.2 ms to microdissect only the tumor cells (24). Approximately 20,000 cells were captured for each preparation. The LCM samples, adhering to the thermoplastic polymer film on the plastic cap (Figure 1C-E), were sectioned at 6-μm in a cryostat and picked up on non-charged microscopic slides (Fisher Scientific, Pittsburgh, PA, USA). The slides were kept in a slide holder on dry ice, fixed for 30 sec in 70% ethanol, dipped in distilled water for 15 sec, stained in Mayer's Hematoxylin for 15 sec, rinsed for 15 sec in 1X automation buffer, pH 7.5 (Biomeda Corp. Foster City, CA, USA), again rinsed in distilled water followed by 70% ethanol for 30 sec each, counterstained in Eosin Y (Cell Point Scientific, Gaithersburg, MD, USA) for 30 sec, followed by dehydration in graded ethanol solutions (95, 100 and 100%), 30 sec each, cleared by two rinses in xylene, 1 min each, air dried for 5 min, and stored in a slide box in a dessiccator for up to 3 h before LCM. Captured cells were fitted to a 0.5 ml sterile microcentrifuge tube, and returned to ECU in cold packs for RNA extraction.

Manual extraction of total RNA from LCM cells and ss-cDNA preparation. This manual procedure was used for extracting RNA from a small number of LCM captured cells was carried out according to manufacture's specifications using the RNeasy isolation Kit® from Qiagen, Valencia, CA, USA, as previously described (19, 21). The quality of RNA was determined on an Agilent 2100 Bioanalyzer (Agilent Technologies, Inc, Palo Alto, CA, USA) utilizing the RNA 6000 Nano LabChip®, or by electrophoresis on Superload (Viagen, Austin, TX, USA) agarose gels (25) and RNA quantitated with RiboGreen quantitation reagent (26) (Molecular Probes, Eugene, OR, USA). The "Sensiscript RT Kit®" from Qiagen was then employed for making a copy of ss-DNA, resulting in 40 μl of ss-cDNA, of which 2-3 μl was subsequently amplified by PCR. One hundred thousand captured cells on 5 plastic LCM caps (each accommodating 20,000 cells) was enough to test all the 11 genes of interest, considering that each cell contains ~20 pg total RNA or 0.4 pg mRNA (equivalent to 0.36 pg ss-cDNA), as only a few picograms of cDNA are needed per PCR reaction (27).

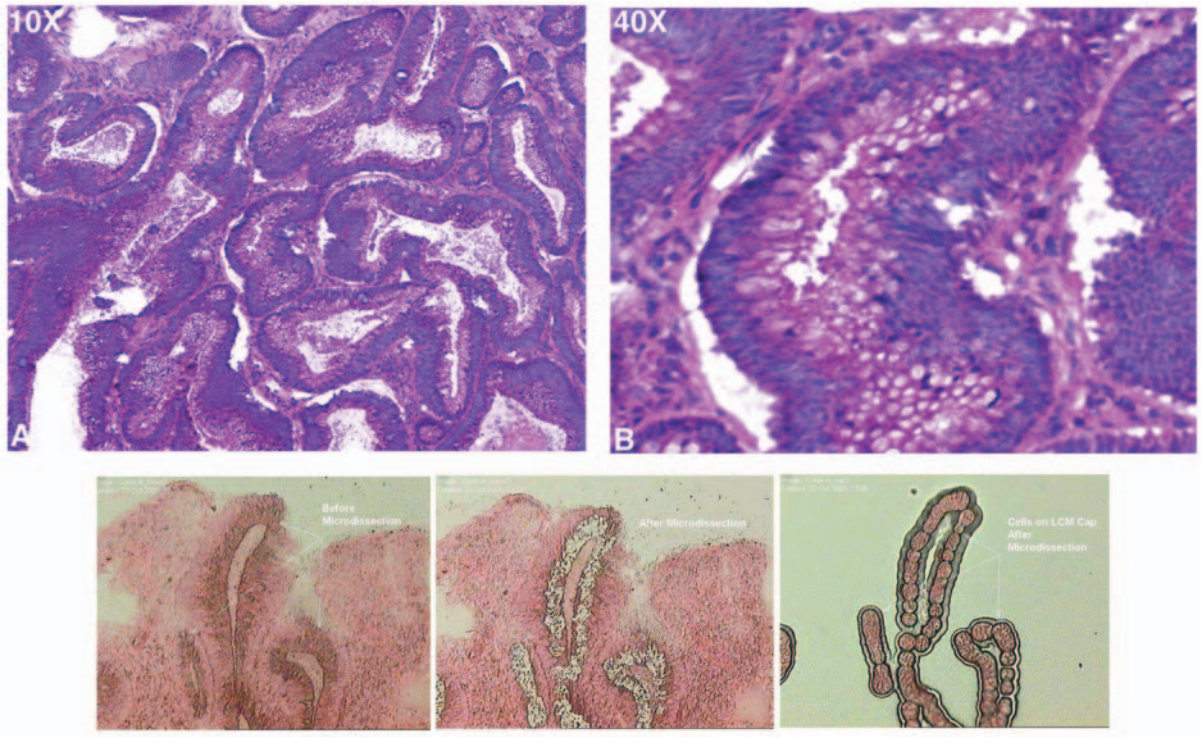

Figure 1. *A) Longitudinal H & E cryostat section of colon adenoma exhibiting high grade dysplasia (i.e., carcinoma in situ, stage 0), 10X. B) As A x 40. C-E: LCM displaying dysplastic cells from the above section before being pulsed by an IR laser (C); the middle panel shows same area with dysplastic cells removed (D); and the right panel shows removed dysplastic regions on a film cap (E).*

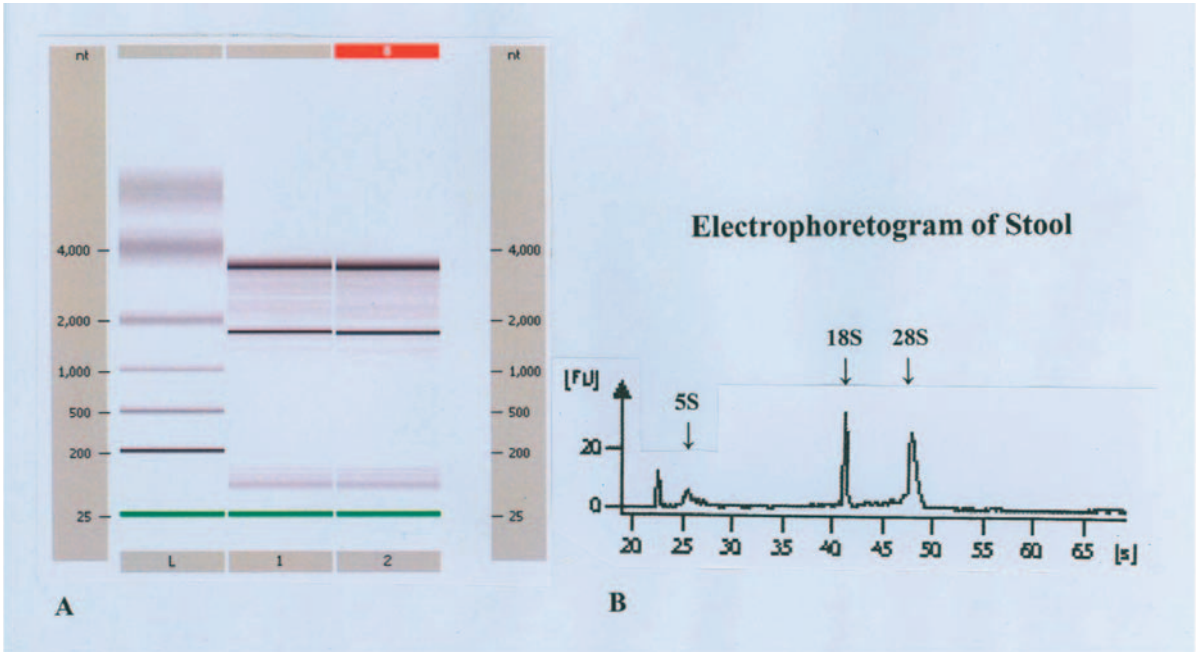

Figure 2. *A) Representative Superload agarose gel protocol for a stool sample showing a nondegraded RNA. Lane L is molecular marker, and lanes 1 and 2 are replicas of a 4 ng total RNA of the same stool sample. B) An Agilent 2100 electrophoretogram showing the 28S, 18S and tRNA, 5.8S and 5S bands for the same stool sample as in panel A.*

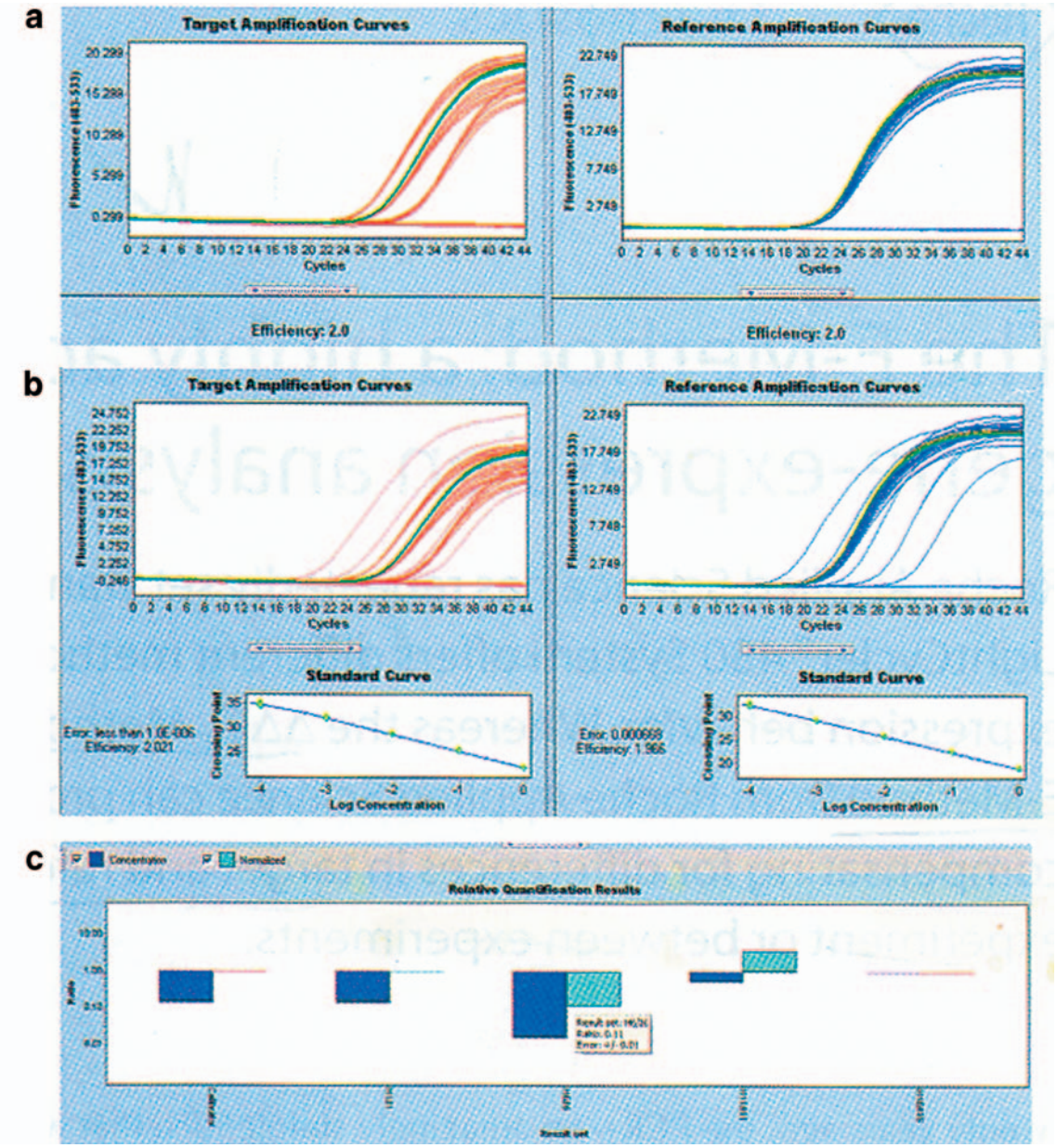

Figure 3. *Two different relative quantification analysis of the same run for the* $2^{-\Delta\Delta C}{}_T$ *method, upper panel (a) or E-method (b). The final quantification results are automatically calculated from the crossing point (CP) values of the target and the reference gene (unknowns and calibrators) as shown in the bottom panel (c). Adapted from reference 29.*

Automatic extraction of total RNA from stool samples. In collaboration with colleagues at Roche Applied Science (Indianapolis, IN, USA), total RNA from stool was automatically extracted by the "MagNA pure LC™" automated system. The MagNA Pure LC™ is a compact benchtop robotic workstation programmed to automatically perform separation of nucleic acids. Propriety magnetic glass particles are used for separation of RNA, followed by transfer of eluted RNA (in a 100 µl volume) into a storage cartridge, which keeps samples cool at 4˚C until removed. The Roche RNA Isolation Kit II™ was used. The machine can automatically pipette purified RNA into borosilicate capillaries of a LightCycler's PCR instrument in 2-3 µl volume (28) for high throughput qPCR analyses.

Stool samples preserved in bacteriostatic S.T.A.R. medium were shipped overnight to Roche in cold packs, and the RNA extracted from the samples (100 µl eluates kept in 0.5 ml sterile Eppendorf tubes) were returned to us overnight in cold packs. The quality and yield of the RNA extraction for each sample was determined in an

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.