Int. J. Cancer: **119,** 1654–1659 (2006)

A highly specific real-time RT-PCR method for the quantitative determination of CK-19 mRNA positive cells in peripheral blood of patients with operable breast cancer

Aliki Stathopoulou[1,2], Maria Ntoulia[1], Maria Perraki[2], Stella Apostolaki[2], Dimitris Mavroudis[2,3], Nikos Malamos[4], Vassilis Georgoulias[2,3] and Evi S. Lianidou[1]*

[1]*Laboratory of Analytical Chemistry, University of Athens, Athens, Greece*
[2]*Laboratory of Tumor Cell Biology, School of Medicine, University of Crete, Heraklion, Greece*
[3]*Department of Medical Oncology, University General Hospital of Heraklion, Heraklion, Greece*
[4]*Medical Oncology Unit, "Helena Venizelou" Hospital, Athens, Greece*

The aim of the present study was to decrease the incidence of false positives and to better characterize marginally *cytokeratin-19* (*CK-19*) mRNA positive peripheral blood samples from patients with early stage breast cancer. A new set of highly specific primers for *CK-19,* which avoids amplification of contaminating genomic DNA, was designed and evaluated to improve the specificity and sensitivity of the previously described methodology. The primers were specifically designed to avoid amplification of contaminating genomic DNA and *CK-19* pseudogenes. The breast cancer cell line MCF-7 was used as positive control for the development and analytical evaluation of the assay, while peripheral blood samples from 62 healthy female individuals and 160 patients with early breast cancer were used for the evaluation of the sensitivity and specificity of the new primer pair. The novel designed primer pair was highly sensitive, as it detects up to 1 MCF-7 cell, and specific as none of the healthy individuals had detectable *CK-19* mRNA positive cells in their peripheral blood. *CK-19* mRNA positive cells were detected in 33 out of 160 (20.6%) patients with early breast cancer. Results obtained by the proposed optimized real-time RT-PCR protocol correlated well with those obtained in the same samples by our previously reported quantitative real-time RT-PCR [concordance in 198/222 (89.2%), $p = 0.0022$, McNemar test]. The improved method eliminates the incidence of false positives and is highly sensitive and specific. The method could be used in a clinical setting in the near future for continuous monitoring and quantification of circulating epithelial cells in the peripheral blood of patients with operable breast cancer, provided that a quite larger number of clinical samples with a known follow-up will be analyzed.

Key words: breast cancer; CK-19; peripheral blood; real-time RT-PCR; circulating tumor cells (CTCs)

During recent years, there has been an increasing body of evidence that the detection and characterization of tumor cells in bone marrow or peripheral blood of breast cancer patients may be clinically relevant and correlates well with disease-free interval and overall survival.[1] Moreover, the prospective evaluation of minimal residual disease may give information concerning the effectiveness of adjuvant therapy.[2] Therefore, highly sensitive methods for the early detection of circulating tumor cells (CTCs) are very important for the early diagnosis and more effective treatment of minimal residual disease.

The intermediate filament *cytokeratin-19* (*CK-19*) is stably and abundantly expressed in the majority of epithelial tumor cells and is one of the most frequently used markers for the detection of CTCs in the peripheral blood of patients with breast cancer.[3–5] Our group has recently shown that the detection of *CK-19* mRNA positive cells in the peripheral blood before the initiation of any adjuvant treatment represents one of the most powerful determinants of outcome in patients with operable breast cancer, with patients who are negative for *CK-19* mRNA having a better chance of long-term survival and disease free interval.[6]

Furthermore, in a previous study, we have developed a quantitative method based on real-time PCR of fluorescently labeled specific hybridization probes for *CK-19* mRNA.[7] By applying that method in breast cancer patients (stage I/II) and in healthy blood donors, positive cells were detected in 70/337 (20.77%) and in 2/89 (2.2%) peripheral blood samples, respectively. In this way, we observed a false-positive rate of 2.2% for normal blood donors, when a cutoff level of 0.6 MCF-7 cell equivalents/5 μg total RNA (detection limit of the method) was set. By using this statistically calculated cut-off, some peripheral blood samples of patients and healthy donors were regarded as negative, despite showing a late amplification curve for *CK-19* (*i.e.*, at very high crossing points). However, for samples found to be very close to this cut-off, the interpretation of this "grey zone" results was very critical for the treatment of our patients with early breast cancer.[8]

To further optimize the specificity and sensitivity of our assay, we decided to design and evaluate a new set of primers in the same region of the *CK-19 gene*. In designing this new primer pair, an intron spanning site in the forward primer was incorporated, so that nonspecific target transcript detection in any trace of contaminating genomic DNA was completely avoided. Furthermore, the new primers were designed to distinguish the *CK-19* specific sequence from the 2 known *CK-19 pseudogenes*[9,10] by including sequence areas with low homology. The improved primer pair was evaluated and compared with the previous one in a large number of peripheral blood samples obtained from patients with early breast cancer, as well as, healthy blood donors.

Material and methods

Cell samples

The human mammary carcinoma cell line MCF-7, which expresses the *CK-19* gene (obtained from the American Type Culture Collection; ATCC), was used as positive control and cultured as previously described.[11]

Clinical samples

Twenty milliliter of peripheral blood in EDTA was obtained from 160 patients with stage I/II (early stage) breast cancer 3–6 weeks after the surgical removal of the primary tumor and 62 female healthy volunteers (aged 18–65 years). To reduce blood contamination by epithelial cells from the skin, the first 5 mL of blood were discarded and the collection tube was disconnected before withdrawing the needle at the end of the blood draw. Peripheral blood samples from healthy donors and patients were collected and processed in the same manner. All patients and blood donors gave their informed consent and the study has been

Grant sponsors: Cretan Association for Biomedical Research (CARB) and Roche Diagnostics (Roche Hellas).
***Correspondence to:** Laboratory of Analytical Chemistry, Department of Chemistry, University of Athens, 15771, Greece. Fax: +30-210-7274750. E-mail: lianidou@chem.uoa.gr
Received 21 July 2005; Accepted after revision 15 March 2006
DOI 10.1002/ijc.22017
Published online 17 May 2006 in Wiley InterScience (www.interscience.wiley.com).

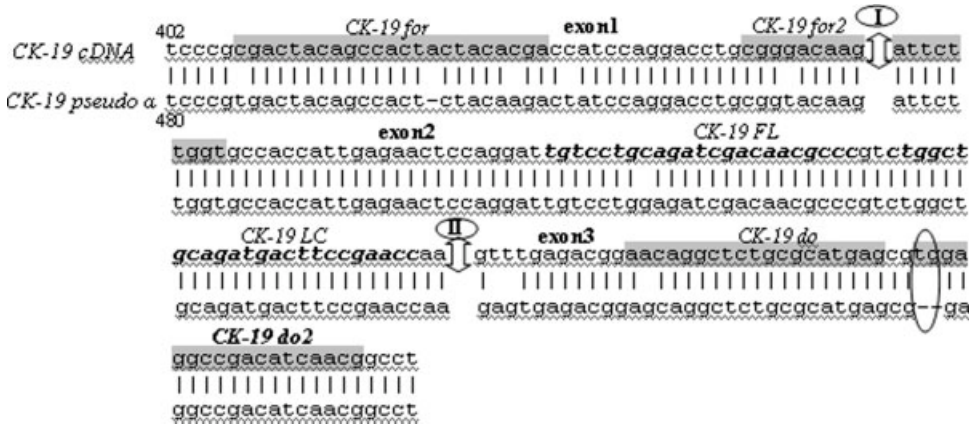

FIGURE 1 – *CK-19* cDNA and *CK-19* α pseudogene sequence alignment and hybridization sites for primers and probes used in protocols A and B. Points I and II represent junctions between exons 1/2 and exons 2/3, respectively.

TABLE I – SEQUENCES OF PRIMERS AND HYBRIDIZATION PROBES USED IN THIS STUDY FOR PROTOCOL B

Gene	Use	Name	Sequence (5′–3′)
CK-19	Forward primer	CK19-for2	CgggACAAgATTCTTggT
	Reverse primer	CK19-do2	CgTTgATGTCggCCTCCA
	Hybridization probe	CK19-FL[1]	TgTCCTgCAgATCgACAACgCCC-FL
	Hybridization probe	CK19-LC[2]	LCRed640-CTggCTgCAgATgACTTCCgAACC

[1]Labeled with fluorescein.–[2]Labeled with LC Red640.

approved by the Ethical and Scientific Committees of the participating Institutions. The peripheral blood mononuclear cells were isolated within 1 hr of venipuncture by gradient centrifugation with Ficoll Hypaque-1077 (Sigma Chemical Company, England), as previously described,[11] and cell pellets were kept at −80°C until total RNA extraction.

Total RNA isolation and cDNA synthesis

Total RNA isolation was performed by using Trizol LS reagent (Invitrogen, USA) according to the manufacturer's instructions. All preparation and handling steps of RNA took place in a laminar flow hood, under RNAse-free conditions. The isolated RNA was dissolved in RNA storage buffer (Ambion, USA) and stored at −70°C until used. RNA concentration was determined using the RiboGreen RNA Quantitation Kit (Molecular Probes, Eugene, OR), with the LightCycler (Roche Diagnostics, Manheim, Germany) serving as a simple fluorimeter. The RNA quantification was performed in the following way: 5 μL of an RNA solution of known concentration supplied with the kit, its dilutions or the unknown sample was added along with 5 μL of the fluorophore RiboGreen in the LightCycler glass capillaries. A standard curve was created by using the fluorescence values of the RNA standard solutions measured using the LightCycler instrument in the Real Fluorimeter Mode (range 5–500 ng/mL). The fluorescence of the samples was measured in triplicate and the RNA concentration was calculated against the standard curve.

Reverse transcription of RNA was carried out with the THERMOSCRIPT RT-PCR System (Invitrogen, USA). Total RNA prepared from the MCF-7 cell line was used as a positive control. cDNA was synthesized from 5 μg of total RNA isolated from peripheral blood mononuclear cells of healthy volunteers and breast cancer patients, according to the manufacturer's instructions. Genomic DNA from peripheral blood of 4 healthy donors was isolated as previously described.[12]

RNA integrity was tested in the cDNA preparations by real-time RT-PCR amplification of the human *hypoxanthine-guanine phosphoribosyl transferase* (*HPRT*) gene using the LightCycler-h-HPRT gene set (Roche Diagnostics), according to the manufacturer's instructions. However, since current scientific data suggest that normalization to single housekeeping gene is inappropriate,[13,14] our results were not normalized to the amount of the *HPRT* gene but rather to the quantity of total RNA that was used for cDNA synthesis, as previously described.[7]

Design of primers for optimized protocol B

The oligonucleotide sequences of the new primer pair CK19-do2 and CK19-for2 used (protocol B) were firstly designed and evaluated in-silico by using the Primer Premier 5 software (Premier Biosoft International, Palo Alto, CA) to avoid primer–dimer formation, false priming sites and formation of hairpin structures. Furthermore, forward primer (CK19-for2) was selected to position on an intron–exon junction, so that hybridization to genomic *CK-19* DNA was completely avoided. Moreover, the primers and probes were designed to differentiate between the homologous *CK-19 a* and *b pseudogenes* according to a search in the BLAST Sequence Similarity Search tool (NCBI, NIH). Especially, the reverse primer (CK19-do2) was designed to a specific location of the *CK-19* mRNA to have 2 mismatches at its 3′-end for *CK-19 a* pseudogene (Fig. 1) and no significant homology with the less homologous *CK-19 b pseudogene,* so that Taq DNA polymerase elongation would not be possible and false-positive results from *CK-19* pseudogenes amplification were avoided. Hybridization probes (TIBmol, Berlin, Germany) were the same as previously described (protocol A).[7] Primers were synthesized at the Lab of Microchemistry (FORTH, Crete, Greece). All primers and hybridization probes sequences are shown in Table I.

In the process of evaluating the specificity of the new primer pair concerning the genomic DNA, we proceeded to the real-time RT-PCR amplification of 4 genomic DNA samples isolated from peripheral blood of healthy donors by using 4 combinations of the previously used (CK19-do and CK19-for) and the newly designed primers (CK19-do2 and CK19-for2). Furthermore, we treated the RNA of 4 samples which were positive with protocol A and negative with protocol B with DNase I RNase-Free enzyme (Ambion, USA), repeated 3 times according to the manufacturer's instructions, prior to cDNA synthesis and retested them with both protocols.

Optimized real-time RT-PCR for CK-19 mRNA (protocol B)

Quantification was based on real-time monitoring during PCR of fluorescently labeled specific hybridization probes for *CK-19*. The point where the fluorescence rises above background noise (crossing point, Cp) is best quantified through the LightCycler

software as the second derivative maximum of the curve. Real-time RT-PCR for *CK-19* mRNA was performed using the LightCycler system (Roche Diagnostics). For protocol A, the primers (CK19-do and CK19-for) and the hybridization probes (CK19-FL and CK19-LC) were used as previously described.[7] For protocol B, our newly designed primers CK19-do2 and CK19-for2, with the same hybridization probes as in protocol A, were used after optimization of their concentration (Table I).

Real-time RT-PCR was performed in a total volume of 20 μL in the LightCycler glass capillaries. For the PCR, 2 μL of cDNA were placed into a 18-μL reaction volume containing 2 μL of the PCR Synthesis Buffer minus Mg^{2+} (10×), 1 μL of $MgCl_2$ (50 mM), 0.4 μL dNTPs (10 mM), 0.3 μL BSA (10 μg/mL), 0.2 μL Taq platinum DNA polymerase (5 U/μL) (Invitrogen, USA), 1 μL of the sense primer CK19-for2 (3 μM), 1 μL of the antisense primer CK19-do2 (3 μM), 1 μL of the hybridization probe CK19-FL (3 μM), 1 μL of the hybridization probe CK19-LC (3 μM) and DEPC-H_2O (added to the final volume). PCR reaction was initiated after a 10 min denaturation at 95°C (hot start PCR) and terminated with a 30 sec cooling step at 40°C. The cycling protocol consisted of denaturation step at 95°C for 10 sec, annealing at 55°C for 20 sec and extension at 72°C for 20 sec and the cycle was repeated for 50 times. Fluorescence detection was performed at the end of each annealing step for 0 sec.

For quantification, an external calibration curve was obtained by using external standard cDNAs. Serial dilutions of a known number of MCF-7 cells (as verified by a hemocytometer) corresponding to 1-1000 MCF-7 cells were used for total RNA isolation and cDNA synthesis. These cDNAs were kept in aliquots at −20°C and used throughout the study as external standards. This calibration curve was created by plotting the number of MCF-7 cells corresponding to each external standard cDNA vs the value of its crossing point (Cp). The number of circulating *CK-19* mRNA positive cells for all tested samples was expressed as MCF-7 cell equivalents per 5 μg of total-RNA, as determined by LightCycler software 3.1, according to this external standard calibration curve, so that the quantitative results obtained for our clinical samples by protocol B could be compared with the corresponding data obtained by protocol A, as previously described.[7]

To provide a quantitative molecular handle except for MCF-7 cell equivalents, we prepared a PCR amplicon corresponding to the template amplified by CK19-do2 and CK19-for2 primers, that would serve as external standard, according to a procedure we have recently described.[15] For this purpose, cDNA was synthesized from total RNA extracted from MCF-7 cells and served as a template for the amplification of the target of interest by real-time PCR using the above-described specific primers. PCR product was run on a 2% agarose gel. No other nonspecific products were observed but the band of proper size (142 bp). This specific band was excised and the amplicon was purified by passing through a column of the QIA Quick Gel Extraction Kit (Qiagen, USA). Quantification of the amplicon was done by using the PicoGreen DNA Quantification Kit (Molecular Probes, Eugene, OR), in the following way: 5 μL of a DNA solution of known concentration (ng/mL) provided with the kit (or its dilutions) was added along with 5 μL of the fluorophore PicoGreen in the LightCycler glass capillaries. A standard curve was generated by using the fluorescence values as measured in triplicate with the Real Fluorimeter mode in the LightCycler (range = 5–500 ng/mL). Then the corresponding concentration of the purified amplicon (fluorescence measured in triplicate), expressed as ng/mL, was converted to copies per μL by using the Avogadro constant and their molecular weight [number of bases of the PCR product multiplied by the average molecular weight of a pair of nucleic acids which is 660]. Serial dilutions of that stock amplicon solution ranging from 10^6 to 10 copies/μL were kept in aliquots at −20°C and served as external standards throughout the study. A calibration curve was created by plotting the concentration of each external standard expressed as copy number per μL vs the value of its corresponding crossing point (C_p). The copy numbers for CK-19 in the unknown samples were determined by the LightCycler software 3.5 according to its corresponding calibration curve.

To ensure that amplifiable material was present in all specimens and to avoid false-negative results, real-time amplification of the housekeeping gene *HPRT* (LightCycler-h-HPRT gene set, Roche Applied Science) was performed for all samples.

To reduce the risk of contamination, RNA extraction, cDNA synthesis, preparation of the real-time RT-PCR steps and thermocycling were performed in separate rooms. Preparation of the PCR mixture was set up in a hood (BioTechne Hepa, TECHNE, Cambridge, UK) and for every extraction or synthesis step during the whole procedure, we used filter tips and included a positive and a negative control.

Statistics

The McNemar and Fischer exact test was used to compare real-time RT-PCR results for *CK-19* mRNA detection on the same cDNAs by both sets of primer pairs. The Wilcoxon test for paired nonnormally distributed groups was used to compare the *CK-19* positive cell levels in samples estimated by the 2 protocols ($p < 0.05$ was considered as statistically significant). Data analysis was carried out with the Statmost statistical package (Statmost, DataMost Corp, USA).

Results

Specificity of real-time RT-PCR for CK-19 (protocol B)

The specificity of the optimized protocol B for real-time RT-PCR for *CK-19* was evaluated by applying 4 combinations of the 2 sets of primer pairs [(*i*) CK19-do2/CK19-for2, (*ii*) CK19-do2/CK19-for, (*iii*) CK19-do/CK19-for, (*iv*) CK19-do/CK19-for2] in 4 genomic DNA samples. The primer pair CK19-do2/CK19-for2 showed no amplification of any product, while the other 3 combinations demonstrated amplification with genomic DNA (Fig. 2). This experiment was performed 3 times and showed consistent results. Moreover, the 4 samples that were positive with protocol A and negative with protocol B became negative with protocol A after DNase treatment.

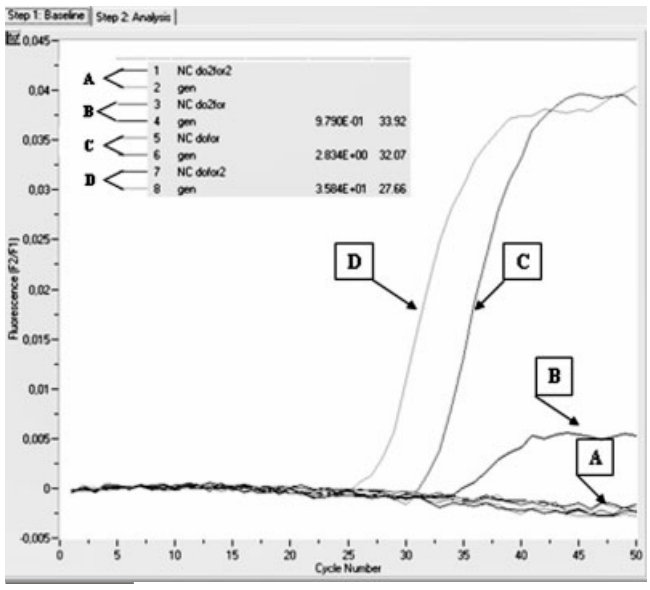

Figure 2 – A characteristic real-time RT-PCR graph for genomic DNA by using 4 combinations of primers with the same hybridization probes [(A) CK19-do2/CK19-for2, (B) CK19-do2/CK19-for, (C) CK19-do/CK19-for, (D) CK19-do/CK19-for2].

TABLE II – WITHIN-RUN AND BETWEEN-RUN PRECISION OF THE REAL-TIME RT-PCR PROTOCOL B FOR *CK-19* mRNA

MCF-7 cell equivalents	Within-run precision ($n = 6$)				Between-run precision ($n = 4$)			
	Crossing point (Cp)		MCF-7 cells		Crossing point (Cp)		MCF-7 cells	
	Mean (SD)	CV%	Mean (SD)	CV%	Mean (SD)	CV%	Mean (SD)	CV%
1	33.6 (0.42)	1.25	1.04 (0.25)	25	32.3 (0.34)	1.05	1.09 (0.15)	13.8
10	29.6 (0.11)	0.37	10.5 (0.7)	6.6	29.1 (0.21)	0.76	9.64 (1.8)	18.9
100	26.0 (0.1)	0.42	86.5 (5.4)	6.3	25.8 (0.24)	0.93	89.5 (6.0)	6.7
1000	21.7 (0.04)	0.21	1084 (31.0)	2.9	22.3 (0.25)	1.12	972 (97.2)	10

Optimization of protocol B real-time RT-PCR for CK-19

The previously reported, real-time assay[7] was improved by using the above novel designed highly specific primer pair for *CK-19* (CK19-do2 and CK19-for2). Since this new set of primers includes the hybridization region of the previously described protocol, it was possible to use the same set of hybridization probes for protocol B as well (CK19-FL and CK19-LC) (Fig. 1). Only slight modifications regarding the conditions of the PCR reaction for protocol B were necessary: the amplification temperature was lowered from 60°C to 55°C and the amplification time was increased from 10 to 20 sec.

The analytical sensitivity and linearity of the protocol B real-time RT-PCR for *CK-19* evaluated by analyzing the cDNA external standards in 4 experiments. Calibration curves from these data showed linearity over the entire quantification range (1-1000 MCF-7 cells) and correlation coefficients greater than 0.99 in all cases, indicating a precise log-linear relationship. To determine within-run precision of protocol B, *CK-19* mRNA was quantified in 4 cDNA samples corresponding to 1, 10, 100 and 1000 MCF-7 cells, in the same run, in 6 parallel determinations, in the LightCycler. Table II demonstrates within-run CVs for MCF-7 cells as determined by the calibration curve ranged from 2.9% to 25%, while for the corresponding Cp values ranged from 0.21% to 1.25%. Furthermore, to determine between-run precision of the assay, the same cDNA samples were frozen (−20°C) in aliquots and analyzed over a period of 1 month on 4 separate assays performed in 4 different days. Table II indicates between-run CVs for MCF-7 cells as determined by the calibration curve ranged from 6.7% to 18.9%, while for the corresponding Cp values ranged from 0.76% to 1.12%.

The mean slope and intercept of the calibration curve was -3.226 ± 0.14 (CV = 4.3%, $n = 4$) and 32.30 ± 0.22 (CV = 0.7%, $n = 4$), respectively, while the PCR efficiency[16] expressed as $E = [10^{-1/\text{slope}}] - 1$ was 1.04 ± 0.06 (CV = 2.9%, $n = 4$). The analytical detection limit of the method defined as 3.3 times the standard deviation of the Cp of the first external standard (1 MCF-7 cell) divided by the mean slope of the calibration curve (D.L. = 3.3SD/slope) was found to correspond to 0.4 MCF-7 cell equivalents.

The analytical sensitivity of the proposed assay was also determined by using the external CK-19 standard prepared as described above. A calibration curve was generated using serial dilutions ranging from 10^6 to 10 copies of the target per μL and showed linearity over the entire quantification range with correlation coefficients larger than 0.99 indicating a precise log-linear relationship. The analytical detection limit of the method was found to correspond to 1 copy of the target per μL while the quantitation limit was equal to 3 copies per μL. This sensitivity is in accordance with most real-time PCR assays, where even commercially available real-time PCR kits (*e.g.*, Roche) use standards that start from these concentrations.

Quantification of CK-19 transcripts per MCF-7 cell

By using the external standard calibration curve and cDNAs synthesized from RNAs extracted from known number of MCF-7 cells (1–1000 cells), as unknown samples, we quantified the number of CK-19 transcripts per MCF-7 cell. According to our results, each MCF-7 cell contains 2,000 ± 246 ($n = 4$) copies of CK-19 transcript. This was estimated after taking into account all the dilution steps of the initial RNA concentration that each MCF7 cell contains. More precisely, the total RNA from one MCF 7 cell is extracted in 10 μL RNA storage buffer. One microliter of this solution was used to synthesize 20 μL of cDNA, from which only 1 μL was used in the real-time PCR reaction, that makes a 1 to 200 dilution, and for one MCF-7 cell a final target concentration of 10 copies per μL in the real-time reaction.

TABLE III – COMPARISON OF PROTOCOL A AND B FOR REAL-TIME RT-PCR FOR THE DETECTION OF *CK-19* POSITIVE CELLS IN PERIPHERAL BLOOD SAMPLES. CONCORDANCE: 89.2% (198/222), ($p = 0.0022$, MCNEMAR & FISCHER EXACT TEST)

Protocol A	Protocol B		Total
	+	−	
+	29	20	49
−	4	169	173
Total	33	189	222

Comparative quantification of CK-19 mRNA positive cells in peripheral blood samples

The specificity and sensitivity of the optimized protocol B for real-time RT-PCR for *CK-19* was evaluated and compared with protocol A. Both quantitative protocols were applied in a total of 222 peripheral blood samples obtained from 62 healthy female blood donors and 160 patients with operable (stage I/II) breast cancer. All these samples were tested for their RNA quality and cDNA synthesis by the expression of the *HPRT* housekeeping gene. Total RNA in each sample was fluorimetrically quantified by the Ribo Green. The same amount of RNA was used for cDNA synthesis and for normalization of our quantitative RT-PCR data.[7]

The specificity of the new set of primers was evaluated by examining 62 peripheral blood samples from healthy volunteers with both protocols. By applying protocol A, 47 samples gave late amplification curves and 2 out of these 62 samples (3.2%, specificity: 96.8%) were considered as positive according to the analytical cut-off of the previously described assay, while none of them showed any amplification when they were analyzed with protocol B (specificity: 100%). The sensitivity of the optimized method was evaluated by analyzing 160 peripheral blood samples of operable breast cancer patients with both protocols. Table III indicates that 33/160 (20.6%) of these samples were found positive with protocol B. By including all the peripheral blood samples tested (healthy donors $n = 62$ and breast cancer patients $n = 160$), 29 samples were positive and 169 were negative with both protocols, so there was an 89.2% concordance (198/222) of positivity and negativity between the 2 protocols (McNemar and Fisher exact test, $n = 222$, $p = 0.0022$) (Table III). Figure 3 demonstrates that *CK-19* mRNA positive cell levels expressed as MCF-7 cell equivalents/5 μg RNA obtained by these 2 protocols correlated very well ($r = 0.986$, $n = 29$) and did not differ significantly (Wilcoxon test for paired data, $p = 0.164$).

Discussion

During recent years, it has been clearly shown by numerous studies that identification of CTCs in peripheral blood offers a unique opportunity to optimize management of breast cancer patients.[6,8,17] Extensive evaluation of various tumor markers for

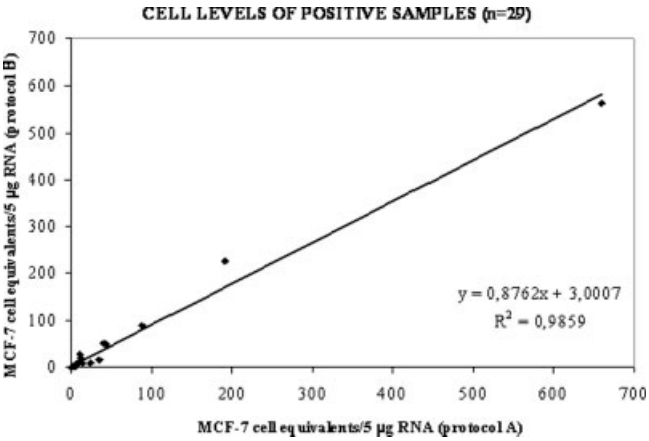

FIGURE 3 – *CK-19* mRNA positive cell levels expressed as MCF-7 cell equivalents/5 μg RNA obtained by protocols A and B.

the molecular detection of CTCs in peripheral blood has shown that *CK-19* is a very sensitive molecular marker for circulating epithelial cells.[3–7] However, the specificity of *CK-19*, as reported so far,[4,5] is not satisfactory to permit its wide spread clinical use. These false positive results could be due to the following reasons: (*i*) contamination with skin epithelial cells during sample collection, (*ii*) presence of genomic DNA contamination in total RNA extracts, (*iii*) amplification of the 2 known *CK-19* pseudogenes or (*iv*) amplification of low level illegitimately transcribed *CK-19* from hematopoietic cells.[18]

Our group recently developed a specific and sensitive method for quantification of circulating *CK-19* mRNA positive cells in peripheral blood samples of breast cancer patients.[7] Despite the very low false-positive rate of this previously published assay, since only 2 in 89 (2.2%) healthy blood donors were found positive for *CK-19* mRNA, there were samples with amplifiable cDNA sequence, considered as negative, since they were detected at very high crossing points below the analytical detection limit of the assay. The evaluation of results from patient samples showing an amplification curve at a Cp slightly lower than the cut-off has proved to be very difficult to interpret and critical in determining the presence of minimal residual disease. This "grey decision zone" had led us to design and evaluate a new set of primers (CK19-do2 and CK19-for2). Our main goal was to avoid false-positive results due to genomic DNA contamination, as well as, false negative, due to a very low initial concentration of *CK-19* mRNA in our samples. By testing the 4 different combinations of the old and the new *CK-19* primer pairs with 4 pure genomic DNA samples isolated from healthy blood donors, we have clearly shown that this new primer pair in combination with our previously reported pair of hybridization probes was highly specific and was not affected by the presence of high concentration of genomic DNA and *CK-19 pseudogenes*. In retesting the 4 samples that were positive with protocol A and negative with protocol B, none of them showed any amplification with both protocols, after DNase treatment. Also, the samples from a subgroup ($n = 62$) of the same previously studied healthy volunteers were retested with both primer pairs. Our results show a significant improvement in the specificity of the assay, since none of these samples had amplifiable product of *CK-19* mRNA by using the new set of primers in contrast to 3.2% (2/62) false positives detected with our previously reported primers.

The new pair of primers evaluated in our study has considerably improved the specificity of *CK-19* mRNA detection by eliminating the false-positive results due to the detection of genomic DNA or *CK-19 pseudogenes*; in this way, clear distribution between positive and negative samples is achieved and the difficult interpretation of the grey-zone results in our previous assay is completely avoided. False-positive results due to the illegitimate transcription of the *CK-19 gene* by nonepithelial cells, *i.e.*, hematopoietic cells, theoretically still remains an unresolved problem; however, we did not detect any CK-19 amplification in the group of 62 healthy volunteers studied.

In contrast with protocol A, any amplification (early or late) with protocol B is considered to be a positive result. For the majority of samples, the 2 sets of primer pairs give almost the same results. In a total of 222 samples tested, 29 samples were found positive and 169 negative by both primer pairs [concordance of 89.2% (198/222)]. However, a small percentage of patient samples 28/222 (13%) that were in a "grey-zone" of *CK-19* detection, as determined by protocol A, could be more definitely characterized as positive or negative by protocol B, since this protocol is not affected by trace amounts of genomic DNA coextracted with total RNA. More specifically, for 10 out of 49 positive samples with protocol A whose concentration was very close to the analytical detection limit of the method, 4 were found to be true positives by protocol B (40%), while 6 were found to be false positives (60%). In the other set of 18 out of 173 samples that were found negative by protocol A, with concentrations slightly below the analytical detection limit ("grey-zone"), 2 were found to be positive by protocol B, thereby false negatives by protocol A (11%).

In conclusion, the proposed new pair of primers is highly specific for *CK-19* and does not hybridize to genomic DNA or *CK-19 pseudogenes*. In this way, all samples that show amplifiable cDNA sequences are considered as true positives. Furthermore, by using this new pair of primers, we avoid the time-consuming DNase treatment step, which could also cause a problem to the RT-PCR translation step of our samples, in case that DNase is not completely inactivated.

The method could be used in a clinical setting in the near future for continuous monitoring and quantification of circulating epithelial cells in the peripheral blood of patients with operable breast cancer, provided that a quite larger number of clinical samples with a known follow-up will be analyzed. Moreover, the method could be also applied for the detection of minimal residual disease in other cancers of epithelial origin such as lung cancer and colorectal cancer.

References

1. Lambrechts AC, Veer LJ, Rodenhuis S. The detection of minimal numbers of contaminating epithelial tumor cells in blood or bone marrow: use, limitations and future of RNA-based methods. Ann Oncol 1998;9:1269–76.
2. Pantel K, Muller V, Auer M, Nusser N, Harbeck N, Braun S. Detection and clinical implications of early systemic tumor cell dissemination in breast cancer. Clin Cancer Res 2003;9:6326–34.
3. Braun S, Pantel K, Müller P, Janni W, Hepp F, Kentenich CR, Gastroph S, Wischnik A, Dimpfl T, Kindermann G, Riethmuller G, Schlimok G. Cytokeratin-positive cells in the bone marrow and survival of patients with stage I, II, or III breast cancer. N Engl J Med 2000;342:525–33.
4. Ring AE, Zabaglo L, Ormerod MG, Smith IE, Dowsett M. Detection of circulating epithelial cells in the blood of patients with breast cancer: comparison of three techniques. Br J Cancer 2005;92: 906–12.
5. Reinholz MM, Nibbe A, Jonart LM, Kitzmann K, Suman VJ, Ingle JN, Houghton R, Zehentner B, Roche PC, Lingle WL. Evaluation of a panel of tumor markers for molecular detection of circulating cancer cells in women with suspected breast cancer. Clin Cancer Res 2005; 11:3722–32.
6. Stathopoulou A, Vlachonikolis I, Mavroudis D, Perraki M, Kouroussis Ch, Apostolaki S, Malamos N, Kakolyris S, Kotsakis A, Xenidis N, Reppa D, Georgoulias V. Molecular detection of cytokeratin-19 positive cells in the peripheral blood of patients with operable breast cancer: evaluation of their prognostic significance. J Clin Oncol 2002; 20:3404–12.

7. Stathopoulou A, Gizi A, Perraki M, Apostolaki S, Malamos N, Mavroudis D, Georgoulias V, Lianidou E. Real-time quantification of CK-19 mRNA positive cells in peripheral blood of breast cancer patients using the lightcycler system. Clin Cancer Res 2003;9:5145–51.
8. Bozionellou V, Mavroudis D, Perraki M, Papadopoulos S, Apostolaki S, Stathopoulos E, Stathopoulou A, Lianidou E, Georgoulias V. Trastuzumab (herceptin) administration can effectively target chemotherapy-resistant cytokeratin-19 (ck-19) mRNA-positive tumor cells in the peripheral blood and bone marrow of patients with breast cancer. Clin Cancer Res 2004;10:8185–94.
9. Ruud P, Fodstad O, Hovig E. Identification of a novel cytokeratin 19 pseudogene that may interfere with reverse transcriptase-polymerase chain reaction assays used to detect micrometastatic tumor cells. Int J Cancer 1999;80:119–25.
10. Savtchenko ES, Schiff TA, Jiang CK, Freedberg IM, Blumenberg M. Embryonic expression of the human 40-kD keratin: evidence from a processed pseudogene sequence. Am J Hum Genet 1988;43:630–7.
11. Stathopoulou A, Angelopoulou K, Georgoulias V, Lianidou ES. Quantitative RT-PCR luminometric hybridization assay with an RNA-internal standard for cytokeratin-19 mRNA in peripheral blood of patients with breast cancer. Clin Biochem 2001;34:651–9.
12. Miller S, Dykes D, Polesky H. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acid Res 1988; 16:1215.
13. Tricarico C, Pinzani P, Bianchi S, Paglierani M, Distante V, Pazzagli M, Bustin S, Orlando C. Quantitative real-time reverse transcription polymerase chain reaction: normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies. Anal Biochem 2002;309:293–300.
14. Dheda K, Huggett JF, Bustin SA, Johnson MA, Rook G, Zumla A. Validation of housekeeping genes for normalizing RNA expression in real-time. PCR Biotechniques 2004;37:118–9.
15. Zygalaki E, Stathopoulou A, Kroupis C, Kaklamanis L, Kyriakides Z, Hadjisavvas A, Kyriacou K, Kremastinos D, Lianidou ES. Quantitative real time RT-PCR study of the expression of Vascular Endothelial Growth Factor (VEGF) splice variants in breast cancer. Clin Chem 2005;51:1518–20.
16. Peters IR, Helps CR, Hall EJ, Day MJ. Real-time RT-PCR: considerations for efficient and sensitive assay design. J Immunol Methods 2004;286:203–17.
17. Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Matera J, Miller MC, Reuben JM, Doyle GV, Allard WJ, Terstappen LW, Hayes DF. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. N Engl J Med 2004;351:781–91.
18. López-Guerrero JA, Bolufer-Gilabert P, Sanz-Alonso M, Barragan-Gonzalez E, Palau-Perez J, De la Rubia-Comos J, Sempere-Talens A, Bonanad-Boix S. Minimal illegitimate levels of cytokeratin K19 expression in mononucleated blood cells detected by a reverse transcription PCR method (RT-PCR). Clin Chim Acta 1997;263: 105–16.