

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
22 February 2007 (22.02.2007)

PCT

(10) International Publication Number
WO 2007/020081 A1

(51) International Patent Classification:

C12Q 1/68 (2006.01)

(21) International Application Number:

PCT/EP2006/008097

(22) International Filing Date: 16 August 2006 (16.08.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

GR 20050100430 17 August 2005 (17.08.2005) GR

60/795,149 25 April 2006 (25.04.2006) US

(71) Applicant (for all designated States except US):

MEDEXIS S.A. [GR/GR]; 4, Asklipiou Str., 145 68
Kryoneri, GR-145 68 Athens (GR).

(72) Inventors; and

(75) Inventors/Applicants (for US only): LIANIDOU, Evrik-

lia [GR/GR]; Piperopoulou 19, GR-145 75 Athens (GR).

STATHOPOULOU, Aliko [GR/GR]; Hatziconstanti 31,

GR-115 24 Athens (GR). MAVROUDIS, Dimitrios

[GR/GR]; Dimokratias 27, GR-71 306 Herakleion (GR).

GEORGOULIAS, Vasileios [GR/GR]; Magnisias 12,

GR-71 305 Herakleion (GR).

(74) Agents: HIEBL, Inge et al.; Thomas-Wimmer-Ring 15,

80539 München (DE).

(81) Designated States (unless otherwise indicated, for every

kind of national protection available): AE, AG, AL, AM,
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI,
GB, GD, GE, GH, GM, HN, HR, HU, ID, IL, IN, IS, JP,
KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT,
LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ,
NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU,
SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT,
TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every

kind of regional protection available): ARIPO (BW, GH,
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,
FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT,
RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

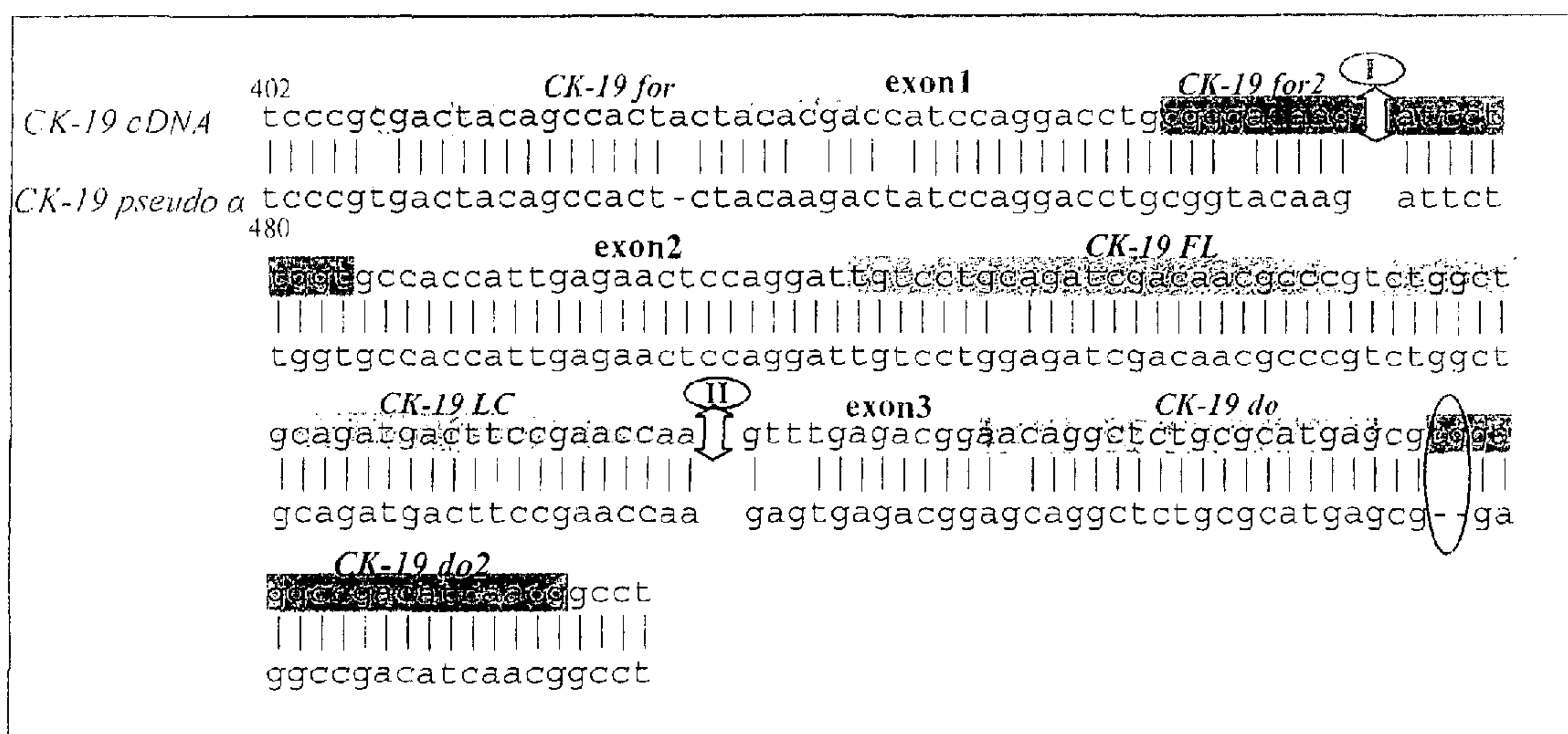
- with international search report
- with amended claims and statement

Date of publication of the amended claims and statement:

5 April 2007

[Continued on next page]

(54) Title: COMPOSITION AND METHOD FOR DETERMINATION OF CK19 EXPRESSION



CK-19 cDNA and CK-19 pseudo a gene 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.

(57) Abstract: Disclosed is a method for quantitative determination of CK-19 mRNA positive cells in a biological sample. The method can be used, for instance, with peripheral blood to detect cancer in a patient. In one embodiment, the method can be used to detect the cancer before initiation of adjuvant treatment, thereby providing information about therapeutic efficacy. Practice of the invention method is sensitive, reliable, and easy to perform.

WO 2007/020081 A1

WO 2007/020081 A1



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 60

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 60

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

COMPOSITION AND METHOD FOR DETERMINATION OF CK19
EXPRESSION

5

The present invention relates to a composition and method for quantitative determination of CK-19 mRNA positive cells in biological samples.

In particular, the invention relates to a
10 method for the detection of circulating tumor cells (CTCs) based on the quantitative determination of the molecular marker CK-19 in biological samples such as those from patients suffering from cancer. Using the method according to the invention detection can take
15 place before the initiation of any adjuvant treatment in order to provide information concerning the effectiveness of the therapy.

BACKGROUND OF THE INVENTION

20 During the last years there is an increasing body of evidence that detection and characterization of tumor cells in bone marrow or peripheral blood of breast cancer patients may be clinically relevant in terms of disease-free interval and overall survival
25 (A. C. Lambrechts *et al*; 1998). Moreover, the prospective evaluation of minimal residual disease (MRD) may give information concerning the effectiveness of adjuvant therapy (K. Pantel *et al*; 2003). Therefore, highly sensitive methods for the
30 early detection of circulating cancer cells are very important for the early diagnosis and more effective

treatment of MRD.

The intermediate filament cytokeratin-19 (CK-19) is stably and abundantly expressed in the majority of epithelial tumor cells and is one of the 5 most frequently used markers for the detection of occult tumor cells in the peripheral blood of patients with breast cancer (S. Braun *et al*; 2000; Y.H.Datta *et al*; 1994; A. Schoenfield *et al*; 1997). The present inventors have recently shown that the 10 detection of CK-19 mRNA positive cells in the peripheral blood represents one of the most powerful determinants of outcome in patients with operable breast cancer before the initiation of any adjuvant treatment, with patients negative for CK-19 mRNA 15 having a better chance of long-term survival and disease free interval (A. Stathopoulou *et al*; 2002).

Furthermore, in a previous study we have developed a quantitative method based on real-time monitoring during PCR of fluorescently-labeled 20 specific hybridization probes for CK-19 mRNA (A. Stathopoulou *et al*; 2003). By applying that method in patients with breast cancer, either stage I/II (operable) or IV (metastatic), as well as, in healthy blood donors we have found positive cells in 70/337 25 (20.77%) and in 2/89 (2.2%) peripheral blood samples, respectively. In this way, we observed a false positive rate (2.2%) for normal blood donors, when a cutoff level of 0.6 MCF-7 cell equivalents/5 μ g RNA (detection limit of the method) was set. By using 30 this statistically calculated cut-off, some peripheral blood samples of patients and healthy

donors were regarded as negative, despite showing an amplification curve for CK-19 at very high crossing points (Cps). These amplification curves were due to amplification of low level illegitimately transcribed CK-19 from hematopoietic cells (J. A. López-Guerrero *et al*; 1997), *CK-19a* and *CK-19b* pseudogenes (P. Ruud *et al*; 1999; E. S. Savtchenko *et al*; 1988) or amplification of contaminating genomic DNA, co extracted with total RNA from our samples. 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 early breast cancer patients (V. Bozionellou *et al*; in press).

Thus there still exists a need for improved primers and methods for quantitative determination of mRNA transcripts in a biological sample. In particular there exists a need for improved primers and methods for determination of CK-19 mRNA positive cells in peripheral blood of operable cancer patients, which methods gives reduced background compared with previously known primers and methods, a high sensibility and a low frequency of false positives.

25 SHORT DESCRIPTION OF THE INVENTION

Thus, in one aspect the invention relates to a primer pair capable of hybridizing to a target sequence of a gene which gene comprises at least one intron, wherein at least one of said primers comprises at least one intron spanning site. Preferably at least one of the primers has a sequence

having a low homology to possible pseudogenes.

In a preferred embodiment the gene is the human CK-19 gene.

In another aspect the present invention
5 provides an improved method for quantitative
determination of mRNA derived from a gene comprising
at least one intron sequence in a sample comprising
the genomic gene, and optionally one or more
pseudogenes, using a pair of primers, wherein at
10 least one of said primers comprises at least one
intron spanning site comprising the steps of

- (i) forming a reaction mixture
comprising nucleic acid
amplification reagents, the primer
15 pair according to the invention and
a test sample;
- (ii) subjecting the mixture to
amplification conditions to generate
at least one copy of a nucleic acid
20 sequence complementary to the target
sequence; and
- (iii) determining the amount of the mRNA
in the sample using real-time
monitoring during PCR.

25

Preferably the gene is the human CK-19 gene and
the sample is a blood sample, a sample from the bone
marrow or a sample derived from the lymph nodes.
Using the CK-19 primers and the method according to
30 the invention the early detection of circulating
tumor cells (CTCs) based on the quantitative

determination of the molecular marker CK-19 in biological samples of patients suffering from cancer may be performed with higher accuracy than by using previously known methods and primers. Surprisingly the observed background and the sensitivity of the method are significantly improved.

In a further aspect the invention relates to a diagnostic method for determining the prospects of adjuvant therapy in a patient suffering from cancer comprising the steps of

- (i) providing a biological sample from the patient;
- (ii) isolating nucleic acids from the biological sample;
- 15 (iii) optionally reverse transcribing the isolated nucleic acids, when the origin of the nucleic acid is RNA;
- (iv) forming a reaction mixture comprising nucleic acid amplification reagents, the primer pair according to the invention and an aliquot of the nucleic acids isolated in step (ii);
- 20 (v) subjecting the mixture to amplification conditions to generate at least one copy of a nucleic acid sequence complementary to the target sequence;
- 25 (vi) quantification of the CK-19 mRNA positive cells in the sample using real-time monitoring during PCR; and
- 30 (vii) based on the amount of CK-19 mRNA positive cells in the sample determining

the prospects of adjuvant therapy.

Using the diagnostic method according to the invention enables the detection of smaller amounts of CK-19 mRNA positive cells in the biological sample of the patients. Furthermore, the frequency of false positive determinations is very low. This means that a reliable determination can be made at an earlier time in the progress of the disease with the consequence that the prospects of the patient is significantly improved.

In another aspect the invention provides a housekeeping primer pair. The housekeeping primer pair hybridizes to a housekeeping gene, which is ubiquitous to a given cell type/organism. The use of housekeeping primer pairs significantly improves the applicability of quantitative real-time PCR, since the incidence of false negatives can be avoided.

In a further aspect the invention relates to a kit comprising the primer pair according to the invention and some or all the reagents necessary for the method according to the invention.

The invention also provides a method of determining the presence of CK-19 mRNA in a biological fluid such as blood. The method includes at least one of and preferably all of the following steps:

- a) separating epithelial mononuclear cells from the biological fluid,
- 30 B) contacting the separated mononuclear cells with an antibody that binds, preferably specifically,

an antigen expressed by the epithelial mononuclear cells. Preferably, the antibody is bound to a solid support and the contacting is sufficient to form a binding complex between the cells, antibody and solid
5 support,

c) separating the binding complex from any unbound material,

d) isolating nucleic acid from endothelial mononuclear cells bound to the complex,

10 e) forming a reaction mixture comprising nucleic acid amplification reagents, a primer pair as disclosed herein and the nucleic acid isolated from the epithelial mononuclear cells,

f) subjecting the mixture to amplification
15 conditions to generate at least one copy of a nucleic acid sequence complementary to the CK-19 target sequence; and

g) determining the amount of the CK-19 mRNA in the biological fluid using PCR, preferably
20 real-time PCR.

SHORT DESCRIPTION OF THE FIGURES

Figure 1 depicts CK-19 cDNA and CK-19 pseudo α gene sequence alignment and hybridization sites for
25 primers and probes used in protocols A and B. Points I and II represent junctions between exons 1/2 and exons 2/3, respectively;

Figure 2 is a real-time PCR for genomic DNA by using four combinations of primers with the same
30 hybridization probes [A) CK19-do2/CK19-for2, B) CK19-do2/CK19-for, C) CK19-do/CK19-for, D) CK19-do/CK19-

for2]; and

Figure 3 is a graph showing CK-19 mRNA positive cell levels expressed as MCF-7 cell equivalents/5 μ g RNA obtained by protocols A and B.

5 Figure 4 is a typical real-time PCR graph for PBGD. The graph shows the real time PCR amplification curves for the housekeeping gene generated using the housekeeping primer pair of the invention detected using a Taqman probe of the invention using
10 biological samples (peripheral blood) from five healthy donors (normal sample 1-5), which efficiently amplify the PBGD gene. In the figure it is observed that no amplification occurs in the two samples containing genomic DNA (DNA isolated from healthy
15 individuals). This is a consequence of the design of the housekeeping primers so that genomic DNA is not amplified. Negative control (NC) corresponds to PCR reaction that does not contain a nucleic acid template.

20 Figure 5 is an agarose gel electrophoresis (2%) for the PBGD PCR products. The actual PCR products were loaded on an agarose gel. 10 μ l of the reactions (half of the total volume) was loaded and detected using standard ethidium bromide staining.
25 Samples 1-5 correspond to normal samples 1-5 in figure 4, whereas negative control corresponds to the negative control in figure 4.

Figure 6 is a schematic drawing showing certain experimental steps for isolating circulating tumor
30 cells (CTCs) from peripheral blood (PB).

Figure 7 is a schematic drawing showing

immunomagnetic enrichment using the monoclonal antibody Ber-EP4 and the magnetic dynabeads epithelial enrich kit.

Figures 8A-C are graphs showing real-time PCR results for three groups of samples referenced below in the Example section. Figure 8A (1st group = PB spiked with known amount of MCF-7 cells, immunomagnetic enrichment), Figure 8B (2nd group= PBS spiked with known amount of MCF-7 cells), Figure 8C (3rd group, same as 1st group except no immunomagnetic enrichment.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides primers and methods for detecting mRNA of genes comprising at least one intron using real-time monitoring during PCR.

The quantitative detection of mRNA is accomplished using any available technique for quantitative determination of PCR products. Preferably this method is real-time PCR, but any other suitable method is within the scope of the invention e.g. competitive PCR. The quantification may be performed by any suitable method and the choice of method is within the skill of the art.

The invention further provides diagnostic methods and kits for detecting the presence of mRNA of a gene comprising at least one intron.

Thus, in a first aspect is provided a primer pair capable of hybridizing to a target sequence of a gene which gene comprises at least one intron,

wherein at least one of said primers comprise at least one intron-spanning site.

The primers described herein may comprise deoxyribonucleic acid (DNA), ribonucleic acid (RNA) 5 or nucleic acid analogs such as uncharged nucleic acid analogs including peptide nucleic acids (PNAs) which are disclosed in WO 92/20702 or morpholino analogs which are described in US 5,185,444; 5,034,506 and 5,142,047. Such sequences can routinely 10 be synthesized using a variety of techniques. In an alternative embodiment the primers comprise labels.

As used herein "target sequence" means a sequence that is detected, amplified, both amplified and detected or is complementary to the sequences 15 provided herein or otherwise has at least one intron in its native state i.e. as genomic DNA or extra chromosomal DNA. While the term target sequence is sometimes referred to as single stranded, those skilled in the art will recognize that the target 20 sequence may be double stranded.

In particular the target sequence is the mRNA-transcript of the CK19 gene. In a preferred embodiment is provided the primer pair having the sequences according SEQ ID NO: 1 25 (5'CGGGACAAGATTCTTGGT-3' FORWARD) and SEQ ID NO: 2 (5'CGTTGATGTCCGCCTCCA-3' REVERSE), respectively, which primers can be employed to amplify the CK19 target sequence.

It should be understood that the sequences of 30 SEQ ID NO: 1 and 2 are preferred embodiments of the primer pair according to the invention. According to

the invention the main feature of the primer pair is that it comprises at least one intron-spanning site. This provides a primer pair that will only bind to a sequence in which the introns have been spliced out, 5 e.g. mRNA, cDNA. It should be understood that said "splicing" may occur naturally i.e. to provide for the detection of mRNA in a biological sample. However, the term also encompasses an engineered sequence having the introns "spliced out" of the 10 sequence, e.g. cDNA. The intron-spanning site may comprise only one base at either site of the intron provided that the primer only binds the sequence without introns under the conditions employed. One or both of the forward and reverse primers may comprise 15 one or more intron-spanning site(s). In a preferred embodiment only one primer comprises an intron-spanning site, and in a particular preferred embodiment the forward primer comprises the intron-spanning site.

20 In the present disclosure the forward primer is the primer that is extended in the same direction as the coding strand of the target nucleic acid. Conversely, the reverse primer is the primer that is extended in the same direction as the non-coding 25 strand of the target nucleic acid. Consequently, the primers align with their 3'-ends facing each other.

Another feature of the primers according to the invention is that the primer sequences do not hybridize to or bind pseudogenes of a particular gene 30 of interest. This second feature is only necessary in cases where pseudogenes exist, and construction of

such sequences requires that the sequence of possible pseudogenes is known or can be found in a sequence database.

In a preferred embodiment of this aspect of the invention the primer pair is designed so that at least one of said primers comprises at least one mismatch at the 3'-end of a possible pseudogene, and preferably 2 or 3 mis-matches. In a particular preferred embodiment of the invention the at least one of the primers comprises at least one mis-match at the 3'-end of the pseudogene of CK-19, CK-19a.

In a second aspect of the invention is provided a method of detecting the presence of a mRNA in a test sample using the primers of the invention comprising the steps of (i) forming a reaction mixture comprising nucleic acid amplification reagents, the primer pair of the invention and a test sample; (ii) subjecting the mixture to amplification conditions to generate at least one copy of a nucleic acid sequence complementary to the target sequence; and (iii) quantification of the mRNA in the sample using real-time PCR monitoring.

In the present description "test sample" and "biological sample" are used interchangeably. In this context "test sample" means anything suspected of containing the target sequence. The test sample can be derived from any biological source, such as for example blood, bone marrow, lymph nodes, bronchial alveolar lavage, saliva, throat swabs, ocular lens fluid, spinal fluid, sweat, sputa, urine, milk, ascites fluid, mucous, synovial fluid, peritoneal

fluid, cerebrospinal fluid, amniotic fluid, tissues such as breast tissues and the like; or fermentation broths, cell cultures, chemical reaction mixtures and the like. Lung cells or tissue may also be used. Most
5 typically the test sample is derived from blood, such as peripheral blood, bone marrow or lymph nodes. The test sample may be used directly as obtained from the source or following a pre-treatment to modify the character of the sample. Thus, the test sample can be
10 pre-treated prior to use by, for example, preparing plasma from blood, disrupting cells, preparing liquids from solid materials, diluting viscous fluids, filtering liquids, distilling liquids, concentrating liquids, inactivating interfering
15 components such as epithelial cells, adding reagents purifying nucleic acids and the like. In a preferred embodiment the pre-treatment is centrifugation.

A "biological fluid" is a biological sample having (or made to have) a liquid form. Examples
20 include peripheral blood, plasma, or an extract obtained from cells or tissue.

The optional reverse transcription step in the method of the invention is included wherever necessary in order to amplify the target sequence,
25 i.e. when the nature of the target sequence is RNA. This process, designated reverse transcription, occurs under the direction of an RNA-dependent DNA polymerase enzyme called a reverse transcriptase. The process furthermore requires buffers and reagents,
30 such as dNTPs, for the reverse transcription. Reverse transcription kits are commercial available and it is

within the skill to perform this process.

The nucleic acid amplification reagents used in the invention includes reagents which are well known and may include, but are not limited to, an enzyme
5 with polymerase activity e.g. heat stable polymerases such as the Taq-polymerase (and, as necessary, reverse transcriptase activity e.g. when monitoring mRNA), enzyme cofactors such as magnesium or manganese; salts and deoxynucleotide triphosphates
10 (dNTPs).

The term "amplification conditions" is generally defined as conditions, which promote hybridizing or annealing of primer sequences to a target sequence and subsequent extension of the
15 primer sequence. It is well known in the art that such annealing is dependant on several parameters, including temperature, ionic strength, sequence length, complementarity and G:C content of the sequences. For example, lowering the temperature in
20 the environment of complementary nucleic acid sequences promotes annealing. For any given set of sequences, melt temperature, or T_m , can be estimated by any of several known methods. Typically, diagnostic applications utilize hybridization
25 temperatures, which are close to (i.e. within 10°C) the melt temperature. Ionic strength or "salt" concentration also impacts the melt temperature, since small cations tend to stabilize the formation of duplexes by negating the negative charge on the
30 phosphodiester backbone. Typical salt concentrations depend on the nature and valency of the cation but

are readily understood by those skilled in the art. Similarly, high G:C content and increased sequence length are also known to stabilize duplex formation because G:C pairings involve 3 hydrogen bonds where 5 A:T pairs have just two, and because longer sequences have more hydrogen bonds holding the sequences together. Thus, a high G:C content and longer sequence lengths impact the hybridization conditions by elevating the melt temperature. Once sequences are 10 selected for a given diagnostic application, the G:C content and length will be known and can be accounted for in determining precisely what hybridization conditions will encompass. Since ionic strength is typically optimized for enzymatic activity the only 15 parameter left to vary is the temperature. Generally, the hybridization temperature is selected close to or at the T_m of the primers or probe. Thus, obtaining suitable hybridization conditions for a particular primer, probe, or primer and probe set is well within 20 ordinary skill of one practicing this art. The amplification product produced as above can be detected during or subsequently to the amplification of the target sequence using any suitable method and a probe disclosed in greater detail below.

25 The invention furthermore discloses any sequence specific probes, such as hybridization probes, Taqman probes or molecular Beacon type probes, for detecting/quantification of the amplification product. Furthermore, the probe may be 30 used to ensure specificity. Said probes may have the sequence according to SEQ ID NO: 3 and 4 for

detecting amplification of the CK19 gene. Construction of probes for detecting amplification of a target sequences is within the skill of the art.

The probes are preferably labelled. The label
5 can be either directly detectable as with for example fluorophores, chemiluminophores, fluorescent particles and the like or indirectly detectable as with specific binding partners and nucleic acids. Preferred labels are directly detectable, and
10 particular preferred labels are fluorescent dyes, such as Sybr Green I, FAM, HEX, VIC, fluorescein LC Red 610, LC Red640 , LC Red670, LC Red 705, and other fluorescent dyes known in the art.

In one embodiment the probe may initially be
15 part of the amplification reaction mixture in which case it is desirable to select conditions such that the probe sequence has a lower melt temperature than the primer sequence. In this way the temperature can initially be chosen so that the probe does not
20 hybridize to the target sequence i.e. over the T_m of the probe. After copies of the target sequence are synthesized the temperature can be lowered in order to let the probe hybridize to the newly synthesized target sequence, provided that this target sequence
25 originally was present in the test sample, and subsequently the possible presence of this target sequence will be detectable. Alternatively the probe is added separately. Preferably the probe does not hybridize to sequences corresponding to the primer
30 sequences.

In another variant of this second aspect of the

invention, step (i) of the method may further comprise a housekeeping primer pair that hybridizes to a housekeeping gene in order to ensure that amplifiable material is present in the test samples 5 and in order to avoid false negative results. Said housekeeping primer pair may be the commercially available housekeeping primer pair for hypoxanthine-guanine phosphoribosyl transferase (HPRT) (purchased from Roche applied Science).

10 Alternatively, a housekeeping primer pair identified by the present inventors may be used in the method of the invention.

Therefore, in a third aspect the present invention provides a housekeeping primer pair having 15 the sequence according to SEQ ID NO 5 and 6 for the forward and reverse primer, respectively.

In the context of the present invention "housekeeping primer pair" and "primer pair" are not the same. In the context of the present invention the 20 term "housekeeping primer pair" is intended to mean a primer pair, which is capable of hybridizing to a target sequence of a gene, which is ubiquitous to a given cell. In other words a "housekeeping primer pair" can be used as an internal control in a method 25 or kit of the invention, i.e. as a negative control.

Said housekeeping primer pair of the invention hybridizes to the housekeeping gene PBGD: Human non-erythropoietic porphobilinogen deaminase (PBGD; hydroxymethylbilane synthase; Accession no: X04808), 30 the third enzyme of the heme biosynthetic pathway, which catalyzes the stepwise condensation of four

porphobilinogen units to yield hydroxymethylbilane, which is in turn converted to uroporphyrinogen III by cosynthetase. A housekeeping gene is a gene that is essential to a cell and thus always present under any
5 conditions. The housekeeping primer pair designed by the present inventors for the PBGD mRNA amplification is: Forward(HGF1) 5'-GGTGGGTGTGCTGCACGAT-3' (SEQ ID NO 5) and Reverse(HGR) 5'-ATCTTCATGCTGGGCAGGGA-3' (SEQ ID NO 6).

10 Said housekeeping primer pair is suitable for the methods and the kit of the present invention. However, the use of the housekeeping primer pair according to the third aspect of the invention is not limited to said methods and the kit, but may be used
15 whenever the samples (cells) to be tested ubiquitously comprise the gene encoding human non-erythropoietic porphobilinogen deaminase.

In Real-time PCR hybridization probes, Taqman probe or a molecular beacon type probe for the
20 visualization of PCR products may be used as described previous in relation to the primer pair of claim 1. One preferred Taqman probe is: 6FAM-ATGAAGGATGGGCAACTGTACCTGACTGG-TMR.

The skilled person will appreciate that instead
25 of the Taqman probe described above, any set of hybridization probes may be used for the detection of the amplified target sequence of the PBGD, or any other suitable housekeeping gene, in the biological sample. Further, taking advantage of the existence of
30 different fluorescent channels available in PCR machines known in the art (machines having 3-6

fluorescent channels are commercially available) the amplification of the housekeeping gene can be done in the same run as the amplification of the CK-19 gene or any other suitable target gene, since a housekeeping primer pair can be used as internal control in many different cases.

The housekeeping primers are preferably designed in a way that avoids amplification of genomic DNA or cDNA in order to avoid non-specific amplification of contaminating genomic DNA in the sample. This may be accomplished using in principle the same criteria for designing the household primers as is used for designing the CK-19 primers according to the invention.

In a fourth aspect of the invention is disclosed a diagnostic method of determining the prospects of adjuvant therapy in a patient suffering from cancer comprising the steps of (i) providing a biological sample from the patient; (ii) isolating nucleic acids from the biological sample; (iii) forming a reaction mixture comprising nucleic acid amplification reagents, the primer pair according to claim 1 and an aliquot of the nucleic acids isolated in step (ii); (iv) subjecting the mixture to amplification conditions to generate at least one copy of a nucleic acid sequence complementary to the target sequence; (v) quantification of the CK-19 mRNA positive cells in the sample using real-time PCR monitoring; and (vi) based on the amount of CK-19 mRNA positive cells in the sample determining the prospects of adjuvant therapy. In a preferred

embodiment of this aspect of the invention the primer pair has the sequence according to SEQ ID NO: 1 and 2.

In another preferred embodiment the biological sample is derived from blood, bone marrow or the lymph nodes, and in a particular preferred embodiment the sample is blood, e.g. a peripheral blood sample.

In yet another preferred embodiment the cancer is breast cancer, preferably operable breast cancer.

10 In general the diagnostic method of the invention may also be used to detect/quantify circulating tumor cells (CTCs) based on the CK-19 marker in cancer types of epithelial origin including but not limited to squamous epithelium, such as
15 squamous cell papilloma and squamous cell carcinoma; transitional epithelium, such as transitional cell papilloma and transitional cell carcinoma; basal cell, such as basal cell carcinoma; glandular epithelium, such as adenoma, cystadenoma and
20 adenocarcinoma; kidney tubules epithelium, such as renal tubular adenoma, renal cell carcinoma and Grawitz tumor; hepatocytes such as hepatocellular adenoma and hepatocellular carcinoma; bile ducts epithelium, such as cholangiocellular adenoma and
25 cholangiocellular carcinoma; and melanocytes, such as melanocytic nevus and malignant melanoma.

In this forth aspect of the invention a sample may be pre-treated similarly to the "test sample" as described earlier. Thus, in a preferred aspect the
30 exemplary blood sample is centrifuged prior to isolation of the nucleic acid in order to isolate the

peripheral mononuclear blood cells (PBMCs). This may be done using any centrifugation technique known in the art, such as Ficoll enrichment, PAX gene blood collection system, immunomagnetic separation and
5 enrichment, and a preferred centrifugation technique is gradient centrifugation. The "nucleic acid amplification reagents" and "amplification conditions" in context of this aspect of the invention are the same as described above.

10 In a fifth aspect the present invention provides a kit for use in the diagnostic method of the fourth aspect of the invention. Said kit comprises the primer pair of the invention, optional further primers that hybridize to other markers on
15 cancer cells and amplification reagents.

Said amplification reagents and the primer pair may either be provided separately or, where appropriate, be mixed.

In a preferred embodiment of this aspect the
20 further primer sequences hybridize to CK19. In another preferred embodiment the further primers hybridize to HER2/neu and cytokeratins, such as CK20, CK8 etc., maspin, GABA A α , B305D-C, PIP, S100A9, S100A14, PSA, mucin, carcinoembryonic antigen, β -
25 subunit of human chorionic gonadotropin, mammaglobin, epidermal growth factor, Ep-CAM and several other mRNA markers known in the art. The choice and combination of additional markers is within the skill of the art. Combination of primers is optional
30 depending on the type of cancer indication. In a particular preferred embodiment the primer pair has

the sequence according to SEQ ID NO: 1 and 2.

Combinations of two or more primers may ensure that the incidence of false negatives is reduced given the fact that more-than-one marker on a cancer
5 cell is detected.

In another preferred embodiment of the invention the kit comprises an internal control in order to avoid false negatives, wherein the internal control preferably is a housekeeping primer pair.
10 Said housekeeping primer pair preferably has the sequences according to SEQ ID NO 5 and 6.

In yet another preferred embodiment of the kit according to the invention all ingredients are lyophilized. In yet another embodiment two or more,
15 e.g. all, lyophilized reagents are mixed. In this case the user, e.g. a clinician, may simply dissolve the mixture in a suitable buffer and add the sample to be tested before the amplification. Besides simplifying the handling procedure, lyophilization
20 makes the reagents more stable for storage.

Additional CK-19 Primers of the Invention

As will be apparent from the foregoing, the
25 present invention encompasses a wide range of suitably "modified" primers (or pair of modified primers) that can detect the human CK-19 gene (including PCR-amplifiable fragments thereof). Preferably, such modified primers or pairs thereof
30 are fully capable of specifically binding a CK-19 target (e.g., the sequence shown in Figure 1) under

one or a combination of the specific primer hybridization conditions disclosed herein. By "specific hybridization" is meant that under a particular hybridization condition, a subject primer
5 or primer pair can form a binding complex with the CK-19 target which complex can be PCR-amplified to produce amplified product. Preferably, the amplified product is in about 90% abundance, preferably about 95% abundance or greater, relative to any other
10 amplified product as determined by standard methods such as quantitative agarose gel electrophoresis. A primer or primer pair is "suitable" if use can achieve one or more objects of the present invention.

15 Before turning to a further discussion about illustrative primer modifications, it is an object of the invention to provide primer(s) having at least about 8 nucleobases(i.e. linked nucleosides) of the sequence shown in SEQ ID NO:1 or SEQ ID NO: 2.
20 Preferably, one or both of the primers include at least about 10 or about 12 of such nucleobases, more preferably at least about 15 up to about 18 of such nucleobases. Primers having the entire sequence of either SEQ ID NO: 1 or SEQ ID NO: 2 will be preferred
25 for many invention applications. Other suitable primers include those having additional sequence up to about 20 to about 30 nucleobases (preferably arranged from the 5' end of the sequences). Still other suitable primers according to the invention
30 include those oligonucleotide sequences spanning from about 8 to about 18 nucleobases in length that

include a stretch of at least eight (8) consecutive nucleobases, preferably at least about 10 to about 15 nucleobases, more preferably about 16, or 17 nucleobases selected from the sequences shown in SEQ ID NO: 1 or SEQ ID NO: 2.

Further illustrative primer pairs according to the invention include one or more suitable primers in which the DNA sequence (or sometimes RNA sequence) has at least four (4) consecutive nucleobases, preferably five (5), six (6), seven (7), eight (8) or nine (9) consecutive nucleobases on either side of the exon 1/2 junction as shown in Figure 1 (about nucleotides 449 to 455 of the CK-19 cDNA). Additionally suitable primers include at least about 1, 2, 3, or 4 nucleotides at their respective 3' ends that do not hybridize (ie., are unable to form hydrogen bonds with) corresponding CK-19 pseudogene alpha sequence as shown in Figure 1. Preferably, such non-hybridizing nucleotides will span about nucleotides 568 to 571 of the CK-19 cDNA. That is, by nucleotide substitution, or in some cases deletion, the 3" end of one or both primers of the primer pair will not fully hybridize to the CK-19 pseudogene under one or combination of hybridization conditions selected.

Additional suitable primers within the scope of the present invention include those having at least about 10 additional nucleotides (e.g, 1, 2, 3, 4, 5, 6, 7, 8 or 9 nucleotides) from the 5'-terminus of one of the sequences represented by SEQ ID NO: 1 or SEQ ID NO: 2. Although less preferred for many uses, the invention also encompasses suitable primers having at

least about 5 (five) additional nucleotides (e.g, 1, 2, 3, 4, or 5 nucleotides) from the 3'-terminus of one of the sequences represented by SEQ ID NO: 1 or SEQ ID NO: 2.

5 Those of skill in the field, having read the instant application, will appreciate that a wide range of other primers and primer pairs are within the scope of the invention. Such embodiments include, without limitation, suitable primers in
10 which one, two or three of the nucleotides shown in SEQ ID NO: 1 or SEQ ID NO: 2 are substituted with A, G, C, T, or U. Also contemplated are suitable deletions of one, two or three sequences (consecutive or non-consecutive) in the sequences represented by
15 SEQ ID NO: 1 or SEQ ID NO: 2.

As will be appreciated, a nucleoside is a base-sugar combination. The base portion of the nucleoside is normally a heterocyclic base. The two most common classes of such heterocyclic bases are the purines
20 and the pyrimidines. Nucleotides are nucleosides that further include a phosphate group covalently linked to the sugar portion of the nucleoside. For those nucleosides that include a pentofuranosyl sugar, the phosphate group can be linked to either the 2', 3' or
25 5' hydroxyl moiety of the sugar. In forming oligonucleotides, the phosphate groups covalently link adjacent nucleosides to one another to form a linear polymeric compound. In turn, the respective ends of this linear polymeric structure can be
30 further joined to form a circular structure, however, open linear structures are generally preferred. In

addition, linear structures may also have internal nucleobase complementarity and may therefore fold in a manner as to produce a double stranded structure. Within the oligonucleotide structure, the phosphate groups are commonly referred to as forming the internucleoside backbone of the oligonucleotide. The normal linkage or * backbone of RNA and DNA is a 3' to 5' phosphodiester linkage.

10 Modified CK-19 Primers: Backbone

Additional examples of primers and primer pairs within the scope of the present invention include oligonucleotides with modified backbones or non-natural internucleoside linkages. As defined in this specification, oligonucleotides having modified backbones include those that retain a phosphorus atom in the backbone and those that do not have a phosphorus atom in the backbone. For the purposes of this specification, and as sometimes referenced in the field, modified oligonucleotides that do not have a phosphorus atom in their internucleoside backbone can also be considered to be oligonucleosides.

Accordingly, the invention encompasses primers and primer pairs in which one or both primers include modified oligonucleotide backbones. Such backbones include phosphorothioates, chiralphosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotri-esters, methyl and other alkyl phosphonates including 3'-alkylene phosphonates, 5'-alkylene phosphonates and chiral phosphonates, phosphinates, phosphoramidates including 3'-amino

phosphoramidate and aminoalkylphosphoramidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, selenophosphates and borano-phosphates having normal 3'-5' linkages, 2'-5' 5 linked analogs of these, and those having inverted polarity wherein one or more internucleotide linkages is a 3' to 3', 5' to 5' or 2' to 2' linkage. Additional oligonucleotides having inverted polarity comprise a single 3' to 3' linkage at the 3'-most 10 internucleotide linkage i.e. a single inverted nucleoside residue that may be abasic (the nucleobase is missing or has a hydroxyl group in place thereof). Various salts, mixed salts and free acid forms are also included. See, for example, the following 15 patents 3,687,808; 4,469,863; 4,476,301; 5,023,243; 5,177,196; 5,188,897; 5,264,423; 5,276,019; 5,278,302; 5,286,717; 5,321,131; 5,399,676; 5,405,939; 5,453,496; 5,455,233; 5,466,677; 5,476,925; 5,519,126; 5,536,821; 5,541,306; 5,550,111; 5,563,253; 5,571,799; 5,587,361; 20 5,194,599; 5,565,555; 5,527,899; 5,721,218; 5,672,697 and 5,625,050; as well as references disclosed therein.

It is a further object of the invention to provide suitable primers and primer pairs in which 25 one or both of the primers do not include a phosphorus atom. Such embodiments will have backbones that are formed by short chain alkyl or cycloalkyl internucleoside linkages, mixed heteroatom and alkyl or cycloalkyl internucleoside linkages, or 30 one or more short chain heteroatomic or heterocyclic internucleoside linkages. These include those having

morpholino linkages (formed in part from the sugar portion of a nucleoside as discussed above); siloxane backbones; sulfide, sulfoxide and sulfone backbones; formacetyl and thioformacetyl backbones; methylene
5 formacetyl and thioformacetyl backbones; riboacetyl backbones; alkene containing backbones; sulfamate backbones; methyleneimino and methylenehydrazino backbones; sulfonate and sulfonamide backbones; amide backbones; and others having mixed N, O, S and CH₂
10 component parts. See, for instance, the following patents 5,034,506; 5,166,315; 5,185,444; 5,214,134; 5,216,141; 5,235,033; 5,264,562; 5,264,564; 5,405,938; 5,434,257; 5,466,677; 5,470,967; 5,489,677; 5,541,307; 5,561,225; 5,596,086; 5,602,240; 5,610,289; 5,602,240;
15 5,608,046; 5,610,289; 5,618,704; 5,623,070; 5,663,312; 5,633,360; 5,677,437; 5,792,608; 5,646,269; 5,677,439; and references disclosed therein.

In some invention embodiments, it will be useful to have one or both primers bear novel groups ie.,
20 not associated with naturally-occurring nucleosides. One such oligomeric compound, an oligonucleotide mimetic that has been shown to have excellent hybridization properties, is referred to as a peptide nucleic acid (PNA; see discussion above). In PNA
25 compounds, the sugar-backbone of an oligonucleotide is replaced with an amide containing backbone, in particular an aminoethylglycine backbone. The nucleobases are retained and are bound directly or indirectly to aza nitrogen atoms of the amide portion
30 of the backbone. See, for instance, the following patents: 5,539,082; 5,714,331; and 5,719,262 for

additional information about making and using PNA compounds.

Further suitable primers and primer pairs in accord with the invention include oligonucleotides
 5 with phosphorothioate backbones and oligonucleosides with heteroatom backbones, and in particular --CH₂--NH--O--CH₂--, --CH₂--N(CH₃)--O--CH₂-- (known as a methylene (methylimino) or MMI backbone], --CH₂--O--N(CH₃)--CH₂, --CH₂--N(CH₃)--N(CH₃)--CH₂-- and --O--
 10 N(CH₃)--CH₂--CH₂--, and --O--P--O--CH₂--. Also preferred are oligonucleotides having morpholino backbone structures. See the previous discussion and U.S. Pat. No. 5,034,506. See also U.S. Pat. Nos. 5,489,677, and 5,602,240.

15

Modified CK-19 Primers: Sugar group

In some invention embodiments, it may be useful to have primers and primer pairs in which the oligonucleotides are modified to have one or more
 20 substituted sugar moieties. Preferred oligonucleotides with this modification include one of the following at the 2' position: OH; F; O--, S--, or N-alkyl; O--, S--, or N-alkenyl; O--, S-- or N-alkynyl; or O-alkyl-O-alkyl, wherein the alkyl,
 25 alkenyl and alkynyl may be substituted or unsubstituted C₁ to C₁₀ alkyl or C₂ to C₁₀ alkenyl and alkynyl. Additional modifications include
 O[(CH₂)_nO]_mCH₃, O(CH₂)_nOCH₃, O(CH₂)_nNH₂, O(CH₂)_nCH₃,
 O(CH₂)_nONH₂, and O(CH₂)_nON[(CH₂)_nCH₃]₂, where n and m are
 30 from 1 to about 10. Other exemplary oligonucleotides comprise one of the following at the 2' position: C₁

to C₁₀ lower alkyl, substituted lower alkyl, alkenyl, alkynyl, alkaryl, aralkyl, O-alkaryl or O-aralkyl, SH, SCH₃, OCN, Cl, Br, CN, CF₃, OCF₃, SOCH₃, SO₂CH₃, ONO₂, NO₂, N₃, NH₂, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, poly-alkylamino, substituted silyl, an RNA cleaving group, a reporter group, or a nucleic acid intercalator. Additional modifications include 2'-methoxyethoxy (2'-O--CH₂CH₂OCH₃, also known as 2'-O-(2-methoxyethyl) or 2'-MOE) (Martin et al., *Helv. Chim. Acta*, 1995, 78, 486-504) i.e., an alkoxyalkoxy group. A further illustrative modification preferred includes 2'-dimethylaminoethoxy, i.e., a O(CH₂)₂ON(CH₃)₂ group, also known as 2'-DMAOE, as described in examples hereinbelow, and 2'-dimethylamino-ethoxyethoxy (also known as 2'-O-dimethyl-amino-ethoxy-ethyl or 2'-DMAEOE), i.e., 2'-O--CH₂--O--CH₂--N(CH₃)₂.

Other suitable primers and primer pairs are within the scope of the present invention. These include those primers having modifications that include 2'-methoxy (2'-O--CH₃), 2'-aminopropoxy (2'-OCH₂CH₂CH₂NH₂), 2'-allyl (2'-CH₂-CH=CH₂), 2'-O-allyl (2'-O--CH₂-CH=CH₂) and 2'-fluoro (2'-F). The 2'-modification may be in the arabino (up) position or ribo (down) position. An illustrative 2'-arabino modification is 2'-F. Similar modifications may also be made at other positions on the oligonucleotide, particularly the 3' position of the sugar on the 3' terminal nucleotide or in 2'-5' linked oligonucleotides and the 5' position of 5' terminal nucleotide. Oligonucleotides may also have sugar

mimetics such as cyclobutyl moieties in place of the pentofuranosyl sugar. See, for example, the following patents: 4,981,957; 5,118,800; 5,319,080; 5,359,044; 5,393,878; 5,446,137; 5,466,786; 5,514,785; 5 5;519,134; 5,567,811; 5,576,427; 5,591,722; 5,597,909; 5,610,300; 5,627,053; 5,639,873; 5,646,265; 5,658,873; 5,670,633; 5,792,747; and 5,700,920.

Still further primer pairs according to the invention include one or more primers with a Locked
10 Nucleic Acid (LNA). A preferred LNA features a 2'-hydroxyl group linked to the 3' or 4' carbon atom of the sugar ring thereby forming a bicyclic sugar moiety. The linkage is preferably a methylene (--CH₂--)_n group bridging the 2' oxygen atom and the 4' carbon
15 atom wherein n is 1 or 2. LNAs and preparation thereof are described in International Published Patent Application Nos. WO 98/39352 and WO 99/14226 as well as the following U.S. Patents and patent publications: 6,794,499; 6,670,461; 2003/0082807
20 (Xylo-LNA); 2003/0087230 (L-ribo-LNA); and 2003/0224377.

Modified CK-19 Primers: Nucleobase

As will be appreciated, oligonucleotides may
25 also include nucleobase (often referred to in the art simply as "base") modifications or substitutions. As used herein, "unmodified" or "natural" nucleobases include the purine bases adenine (A) and guanine (G), and the pyrimidine bases thymine (T), cytosine (C)
30 and uracil (U). Modified nucleobases include other synthetic and natural nucleobases such as 5-

methylcytosine (5-me-C), 5-hydroxymethyl cytosine, xanthine, hypoxanthine, 2-aminoadenine, 6-methyl and other alkyl derivatives of adenine and guanine, 2-propyl and other alkyl derivatives of adenine and 5 guanine, 2-thiouracil, 2-thiothymine and 2-thiocytosine, 5-halouracil and cytosine, 5-propynyl ($--C\equiv C--CH_3$) uracil and cytosine and other alkynyl derivatives of pyrimidine bases, 6-azo uracil, cytosine and thymine, 5-uracil (pseudouracil), 4-10 thiouracil, 8-halo, 8-amino, 8-thiol, 8-thioalkyl, 8-hydroxyl and other 8-substituted adenines and guanines, 5-halo particularly 5-bromo, 5-trifluoromethyl and other 5-substituted uracils and cytosines, 7-methylguanine and 7-methyladenine, 2-F-15 adenine, 2-amino-adenine, 8-azaguanine and 8-azaadenine, 7-deazaguanine and 7-deazaadenine and 3-deazaguanine and 3-deazaadenine. Further modified nucleobases include tricyclic pyrimidines such as phenoxazine cytidine (1H-pyrimido[5,4-20 b][1,4]benzoxazin-2(3H)-one), phenothiazine cytidine (1H-pyrimido[5,4-b][1,4]benzothiazin-2(3H)-one), G-clamps such as a substituted phenoxazine cytidine (e.g. 9-(2-aminoethoxy)-H-pyrimido[5,4-b][1,4]benzoxazin-2(3H)-one), carbazole cytidine (2H-25 pyrimido[4,5-b]indol-2-one), pyridoindole cytidine (H-pyrido[3',2':4,5]pyrrolo[2,3-d]pyrimidin-2-one). Modified nucleobases may also include those in which the purine or pyrimidine base is replaced with other heterocycles, for example 7-deaza-adenine, 7-30 deazaguanosine, 2-aminopyridine and 2-pyridone. Further nucleobases include those disclosed in U.S.

Pat. No. 3,687,808, those disclosed in The Concise Encyclopedia Of Polymer Science And Engineering, pages 858-859, Kroschwitz, J. I., ed. John Wiley & Sons, 1990, those disclosed by Englisch et al.,
5 Angewandte Chemie, International Edition, 1991, 30, 613, and those disclosed by Sanghvi, Y. S., Chapter 15, *Antisense Research and Applications*, pages 289-302, Crooke, S. T. and Lebleu, B. ed., CRC Press, 1993. Additional modifications include 5-substituted
10 pyrimidines, 6-azapyrimidines and N-2, N-6 and O-6 substituted purines, including 2-aminopropyladenine, 5-propynyluracil and 5-propynylcytosine. 5-methylcytosine substitutions have been shown to increase nucleic acid duplex stability by 0.6-
15 1.2.degree. C. (Sanghvi, Y. S., Crooke, S. T. and Lebleu, B., eds., *Antisense Research and Applications*, CRC Press, Boca Raton, 1993, pp. 276-278) and are illustrative base substitutions, even more particularly when combined with 2'-O-
20 methoxyethyl sugar modifications. See, for instance, U.S. Pat. No. 3,687,808, as well as U.S. Pat. Nos.: 4,845,205; 5,130,302; 5,134,066; 5,175,273; 5,367,066; 5,432,272; 5,457,187; 5,459,255; 5,484,908; 5,502,177; 5,525,711; 5,552,540; 5,587,469; 5,594,121,
25 5,596,091; 5,614,617; 5,645,985; 5,830,653; 5,763,588; 6,005,096; and 5,681,941, 5,750,692.

A primer or primer pair in accord with the invention is "modified" if it includes at least one
30 of the foregoing oligonucleotide modifications. As will be readily apparent, certain of the modified

primers and primer pairs will not be optimal for some invention embodiments such as performing real-time PCR. However, the modified primers can be useful as electrophoretic markers, and/or as "antisense" compositions, for instance.

For PCR applications in which increased target affinity and specificity is useful or when enhanced robustness is helpful, one or both of the sequences represented by SEQ ID NO: 1 or SEQ ID NO: 2 can be modified to include at least one LNA, for example, 1 (one), 2 (two), 3 (three), 4 (four) or 5 (five) of such LNAs. See, for example, Vester, B and J. Wengel (2004) *Biochemistry* 43: 13233; and references cited therein, for additional disclosure relating to making and using LNA oligonucleotides.

As discussed above, the invention also provides a method determining the presence of CK-19 mRNA in a biological fluid. In one embodiment, the method includes the following steps (a)-(g):

a) separating any mononuclear cells from the biological fluid. The separation step can be performed using nearly any method capable of separating cells from a biological fluid such as filtration and/or centrifugation. In embodiments in which centrifugation is selected, it will often be preferred to use a Ficoll or other suitable cell separating gradient. Use of the Ficoll Histopaque-1077 system (Sigma Aldrich, St. Louis, MO (USA)) is preferred for many applications such as those in which the biological fluid is peripheral blood.

b) contacting the separated mononuclear cells with a polyclonal or monoclonal antibody (or antigen binding fragment thereof such as Fab, F(ab')₂, single-chain antibodies, and the like) that specifically
5 binds an antigen expressed by the epithelial mononuclear cells. In one embodiment, the antigen is a glycoprotein expressed by cells, for instance on the cell surface or cytoplasm. An illustrative antibody is one that specifically binds the antigen
10 CDC326, for instance, ber-EP4, B302 (323/A3), B29.1 (VU-ID9), VU-1D9, HEA125. These and other suitable antibodies can be obtained from a variety of commercial sources such as Abcam plc (Cambridge, UK);
Dako UK LTD. (Cambridgeshire, UK), and Santa Cruz
15 Biotechnology INC (Santa Cruz, CA (USA)). The antibody (or antigen binding fragment thereof) can be pre-bound to any suitable solid support, for instance, glass fiber filter paper, nitrocellulose,
scintered glass, plastic, synthetic polymer,
20 cellulose, cellulose acetate, polytetrafluoroethylene, polyethylene, polypropylene, or polyvinylidene fluoride. In one embodiment, the solid support is in a bead format, preferably one that includes a magnetic or paramagnetic material. A
25 preferred bead is one manufactured by Dynal. Preferably, the contacting step of the method is sufficient to form a binding complex between the cells, antibody and solid support.

c) separating the binding complex from any
30 unbound material, for instance, by filtration and/or centrifugation,

d) isolating nucleic acid (e.g, RNA such as mRNA) from endothelial mononuclear cells bound to the complex. Typically, and as described above, cDNA will be made from the RNA isolated from the cells,

5 e) forming a reaction mixture comprising nucleic acid amplification reagents, a primer pair as disclosed herein, for instance primers having the sequence represented by SEQ ID Nos. 1 and 2, and the nucleic acid isolated from the mononuclear cells,

10 f) subjecting the mixture to amplification conditions to generate at least one copy of a nucleic acid sequence complementary to the CK-19 target sequence; and

15 g) detecting CK-19 mRNA in the biological sample using PCR, preferably RT-PCR. If desired, the amount of the CK-19 mRNA in the biological fluid can determined.

By the term, "specific binding" or a similar term is meant a molecule disclosed herein which binds
20 another molecule, thereby forming a specific binding pair. However, the molecule does not recognize or bind to other molecules as determined by, e.g., Western blotting ELISA, RIA, mobility shift assay, enzyme-immuno assay, competitive assays, saturation
25 assays or other protein binding assays know in the art. See generally, Harlow and Lane in, *Antibodies: A Laboratory Manual* (1988) and references cited therein for examples of methods for detecting specific binding between molecules.

30 In embodiments of the foregoing method in which the solid support is a magnetic bead, the method will

further include the step of impressing a magnetic field on the binding complex to separate the complex from any unbound material. The separated bead complex can then be manipulated to isolate the cells (and
5 prepare nucleic acid therefrom) using standard procedures. See for instance, information from Dynal (Epithelial Enriched Dynabeads).

The method is flexible and compatible with use of one or a combination of primer pairs as disclosed
10 herein. Use of a particular primer pair will depend on intended use. However for many embodiments, the primers represented by SEQ ID No: 1 and SEQ ID No.2 will be sufficient. A preferred biological fluid is peripheral blood.

15

If desired, the method is readily adapted to include use of one or more suitable control assays such as those mentioned in the Examples. For instance, it will often be useful to prepare a
20 standard curve of CK-19 expressing cells in embodiments in which the user wishes not only to detect but to quantify mononuclear cells in a particular biological sample. The Examples below show how to make an illustrative standard curve in which
25 peripheral blood is spiked with MCF-7 cells. It will be appreciated that other cells can be used to create the standard curve. It will also be appreciated that once the standard curve is prepared, it need not be repeated every time the method is practiced. For
30 instance, in embodiments in which the invention is used in a clinical setting, the standard curve could

be prepared once (or a most a few times) in which one or only a few types of biological samples are assayed such as peripheral blood obtained from patients.

If desired, the foregoing method can also be
5 adapted to include use of one or more of the housekeeping genes disclosed herein. Amplified CK-19 target sequence can be detected and optionally quantified using the probes disclosed herein.

As will be apparent from the foregoing, the
10 present invention is flexible and can be used to detect and optionally quantify CK-19 as expressed in a variety of biological samples including normal and abnormal (e.g., cancerous) tissues. Regarding normal tissues, the following are exemplary: hair follicles,
15 secretory cells of sweat glands, Merckell cells, luminal epithelial cells of breast ducts, surface mucosa and glands of endometrium and endocervix, exocervix, ovary surface mesothelium, Fallopian tube epithelium, cyto- and syncytiotrophoblast cells,
20 amnion, umbilical cord surface epithelium, luminal- and basal cells of prostate, testes rete epithelium, ductuli efferentes, epididymal tubules, Bowman's capsule, proximal-, distal- and collecting tubules of the kidney, urothel, bile duct- and gall bladder
25 epithelium, squamous epithelium-, taste buds-, secretory glandular cells and glandular ducts of tongue, squamous epithelium- and submucosal glands of esophagus, surface mucosa- and glands of stomach, surface mucosa- and crypts of small- and large
30 intestine, pancreas ducts, secretory- and duct cells of salivary glands, thyroid epithelium, surface

mucosa- and glands of trachea, bronchial mucosa and - glands, alveoli, pleura-mesothelium, Hassal's corpuscles and thymus epithelial cells. Regarding abnormal tissues, the following list is illustrative:

5 human breast tumors, fibroadenomas, fibrocystic diseases, cystosarcoma phyllodes, infiltrating ductal carcinomas, infiltrating lobular carcinomas, medullary carcinomas and metastases, invasive carcinoma, intraductal papillomas, pure in situ

10 carcinomas, tissue having Paget's disease, thyroid adenoma, colon-, gastric- and lung adenocarcinomas, ovarian- and urinary bladder carcinomas, teratomas, embryonal carcinomas, testicular cancers, epidermal tumour, squamous- and basal cell carcinomas, and

15 keratocanthomas.

Certain aspects of the forgoing invention have been disclosed in Greek patent application GR 20050100430 as filed on August 17, 2005; and in U.S. Provisional Application No. 60/795,149 as filed on

20 April 4, 2006; each of which is incorporated herein by reference.

The disclosure of all references cited herein are incorporated by reference. The invention has been described in detail with reference to preferred

25 embodiments thereof. However, it will be appreciated that those skilled in the art, upon consideration of this disclosure, may make modifications and improvements within the spirit and scope of the invention.

EXPERIMENTAL

5 Materials and methods

A. 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
10 positive control and cultured as previously described (A. Stathopoulou *et al*; 2001).

B. Clinical samples

Peripheral blood in EDTA was obtained from 160
15 patients with stage I/II (early stage) breast cancer postoperatively and 62 female healthy volunteers (aged 18-65 years). To reduce blood contamination by epithelial cells from the skin, the first 5mL of blood were discarded and the collection tube was at
20 the end disconnected before withdrawing the needle. Peripheral blood samples from healthy donors and patients were collected and processed in the same manner. All patients and donors gave their informed consent and the study has been approved by the
25 Ethical and Scientific Committees of the participating Institutions. The peripheral blood mononuclear cells (PBMC) were isolated within one hour of venipuncture by gradient centrifugation with Ficoll Hypaque-1077 (Sigma Chemical Company, LTD,
30 England), as previously described (A. Stathopoulou *et*

al; 2001), and cell pellets were kept at -80°C until total RNA extraction.

C. Total RNA isolation and cDNA synthesis

5 Total RNA isolation was performed by using Trizol LS reagent (Invitrogen, Corp., Carlsbad, 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
10 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, USA), with the LightCycler (Roche Diagnostics, Mannheim,
15 Germany) serving as a simple fluorimeter. The RNA quantification was performed in the following way: 5 μL of a supplied with the kit RNA solution of known concentration or its dilutions or the unknown sample was added along with 5 μL of the fluorophore
20 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
25 samples was measured in triplicate and the RNA concentration was calculated with the use of the standard curve.

Reverse transcription of RNA was carried out with the THERMOSCRIPT RT-PCR System (Invitrogen,
30 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 PBMC of healthy volunteers and breast cancer patients, according to the manufacturer's instructions.

5 RNA integrity was tested in the cDNA preparations by real-time 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
10 instructions. However, since current scientific data suggest that normalization to single housekeeping genes is inappropriate [C. Tricarico *et al*; 2002 and K. Dheda *et al*; 2004], our results were not
15 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 (A. Stathopoulou *et al*; 2003).

D. Design of primers for optimized protocol B

20 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, USA) in order to avoid
25 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,
30 the primers and probes were designed to differentiate between the highly homologous CK-19a pseudogene

according to a search in the BLAST Sequence Similarity Search tool (NCBI, NIH) (see Figure 1). Especially, the reverse primer (CK19-do2) was designed to a specific location of the CK-19 mRNA in order to have two mismatches at its 3'-end for CK-19a pseudogene (Figure 1) so that Taq DNA polymerase elongation is not possible and false positive results from CK-19a pseudogene amplification are avoided. Hybridization probes (TIBmol, Berlin, Germany) were the same as previously described (protocol A) (A. Stathopoulou et al; 2003). Primers were synthesized at the Lab of Microchemistry (FORTH, Crete, Greece). All primers and hybridization probes sequences are shown in Table 1.

15

Gene	Use	Name	Oligonucleotide sequence (5'-3')
CK-19	Forward primer	CK19-for2	CgggACAAGATTCTTggT
	Reverse primer	CK19-do2	CgTTgATGTCggCCTCCA
	Hybridization probe	CK19-FL ^a	TgTCCTgCAGATCgACAACgCCC-FL
	Hybridization probe	CK19-LC ^b	LCRed640-CTggCTgCAGATgACTTCCgAACC

Table 1. Sequences of primers and hybridization probes used in this study for protocol B. ^a Labeled with fluorescein; ^b Labeled with LC Red640 (TIB MOLBIOL)

20 In the process of evaluating the specificity of the new primer pair concerning the genomic DNA we proceeded to the real-time PCR amplification of a genomic DNA sample isolated from peripheral blood of

a healthy donor by using 4 combinations of the previously used (CK19-do and CK19-for) and the newly designed primers (CK19-do2 and CK19-for2).

5 *D. Optimized real-time RT-PCR for CK-19 mRNA*
(protocol B)

Quantification is based on real-time monitoring during PCR of fluorescently labeled specific hybridization probes for CK-19. The point where the
10 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
15 protocol A, the primers (CK19-do and CK19-for) and the hybridization probes (CK19-FL and CK19-LC) were used as previously described (A. Stathopoulou *et al*; 2003). For protocol B, our newly designed primers CK19-do2 and CK19-for2 with the same hybridization
20 probes as in protocol A, were used; see table 1.

Real-time PCR was performed in a total volume of 20 μ L in the LightCycler glass capillaries. For the PCR, 2 μ L of cDNA were placed into a 18- μ L reaction volume containing 2 μ L of the PCR Synthesis Buffer
25 minus Mg²⁺ (10x), 1 μ L of MgCl₂ (50 mM), 0.4 μ L dNTPs (10 mM), 0.3 μ L BSA (10 μ g/mL), 0.2 μ L Taq platinum DNA polymerase (5 U/ μ L) (Invitrogen, USA), 1 μ L of the sense primer CK19-for2 (3 μ M), 1 μ L of the antisense primer CK19-do2 (3 μ M), 1 μ L of the
30 hybridization probe CK19-FL (3 μ M), 1 μ L of the hybridization probe CK19-LC (3 μ M) and DEPC-H₂O (added

45

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 5 at 95 °C for 10 sec, annealing at 55 °C for 20 sec and extension at 72 °C for 20 sec and repeated for 50 times. Fluorescence detection was performed at the end of each annealing step for 0 sec.

For quantification, an external calibration 10 curve was obtained by using external standard cDNAs. Total RNA was prepared from 1×10^6 MCF-7 cells (as verified by a hemocytometer). Serial dilutions of this RNA preparation in DEPC-treated water, corresponding to 1-1000 MCF-7 cells, were used for 15 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 20 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 the external standard calibration curve, 25 as previously described (A. Stathopoulou et al; 2003).

To ensure that amplifiable material was present in all specimens and to avoid false negative results, real-time amplification of the housekeeping gene 30 hypoxanthine-guanine phosphoribosyl transferase

(HPRT) (LightCycler-h-HPRT gene set, Roche Applied Science) was performed for all samples.

Following protocol was used for amplification of the housekeeping gene. Real-time PCR was performed in a total volume of 20 μ L in the LightCycler glass capillaries. For the PCR, 2 μ L of cDNA were placed into a 18- μ L reaction volume containing 2 μ L of the PCR Synthesis Buffer minus Mg²⁺ (10x), 1 μ L of MgCl₂ (50 mM), 0.4 μ L dNTPs (10 mM), 0.3 μ L BSA (10 μ g/mL), 0.2 μ L Taq platinum DNA polymerase (5 U/ μ L) (Invitrogen, USA), 1 μ L of each the housekeeping sense and antisense primers (3 μ M), 1 μ L of the hybridization probe CK19-FL (3 μ M), 2 μ L of the Taqman probe (6FAM-ATGAAGGATGGGCAACTGTACCTGACTGG-TMR) (3 μ M) and DEPC-H₂O (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 repeated for 50 times. Fluorescence detection was performed at the end of each extension step for 0 sec.

Precautions

To reduce 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 have used filter tips and included a positive and a negative sample control.

Statistics

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

15 **Results**

A. Protocol B real-time RT-PCR for CK-19 and genomic DNA

The specificity of the optimized protocol B for real-time RT-PCR for CK-19 was evaluated by applying
20 4 combinations of primers [A) CK19-do2/CK19-for2, B) CK19-do2/CK19-for, C) CK19-do/CK19-for, D) CK19-do/CK19-for2] in a genomic DNA sample (see Figure 2). The primer pair CK19-do2/CK19-for2 showed no amplification of any product, while the other three
25 combinations demonstrated amplification.

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

We improved our previously reported real-time
30 assay (A. Stathopoulou et al; 2003) by designing a

new highly specific primer pair for CK-19. Only slight modifications regarding the conditions of the PCR reaction for protocol B were necessary: the amplification temperature was lowered from 60 to 55 °C and the amplification time was increased from 10 to 20 sec.

We evaluated the analytical sensitivity and linearity of the protocol B real-time RT-PCR for CK-19, by analyzing the cDNA external standards (prepared as described above) 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. 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 expressed as $E = [10^{-1/\text{slope}}] - 1$ (I.R. Peters *et al*; 2004) 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 C_p of the first external standard (1 MCF-7 cell equivalent) divided by the mean slope of the calibration curve ($D.L. = 3.3SD/\text{slope}$) was found to correspond to 0.4 MCF-7 cell equivalents.

To determine within-run precision of protocol B, CK-19 mRNA was quantified in four cDNA samples corresponding to 1, 10, 100 and 1000 MCF-7 cells, in the same run, in 6 parallel determinations, in the LightCycler.

Reproducibility of the assay								
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

Table 2. Within-run and between-run precision of the Real-time RT-PCR protocol B for CK-19 mRNA.

5 Table 2 demonstrates within-run CV's 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
10 same cDNA samples were frozen (-20 °C) in aliquots and analyzed over a period of one month on 4 separate assays performed in 4 different days. Table 2 indicates between-run CV's for MCF-7 cells as determined by the calibration curve ranged from 6.7%
15 to 18.9%, while for the corresponding Cp values ranged from 0.76% to 1.12%.

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

The specificity and sensitivity of the 5 optimized protocol B for real-time RT-PCR for CK-19 was evaluated in respect to 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 10 (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 15 cDNA synthesis and for normalization of our quantitative RT-PCR data (A. Stathopoulou *et al*; 2003).

The specificity of the new set of primers was evaluated by re-examining 62 out of 89 peripheral 20 blood samples from the healthy volunteers we had previously analyzed with protocol A (A. Stathopoulou *et al*; 2003). By applying protocol A, 2 out of these 89 samples were considered as positive according to the analytical cut-off of the assay ($C_p = 32.17 \pm$ 25 0.70 , $CV(\%) = 2.2$), while none of the 62 samples (the two positive samples were included) showed any amplification when they were analyzed with protocol B. The sensitivity of the optimized method was evaluated by analyzing 160 peripheral blood samples 30 of operable breast cancer patients with both protocols.

Comparison of Protocol A and Protocol B			
Protocol A	Protocol B		Total
	+	-	
+	29	20	49
-	4	169	173
Total	33	189	222

Table 3. Comparison of protocol A and B for real-time 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)

As can be seen in Table 3, 33 (20.6%) of these 10 samples were found positive. Twenty samples that were in the gray zone and characterized as positive with protocol A, were found negative by protocol B, while 4 samples that were characterized as negative with protocol A since they gave amplification curves with 15 Cps greater than the cutoff, were found positive with protocol B. By including all the peripheral blood samples tested (healthy donors n=69 and breast cancer patients n=160) 29 samples were positive and 169 were negative with both protocols, so there was an 89.2% 20 concordance (198/222) of positivity and negativity between the two protocols (McNemar and Fisher exact test, n=222, P=0.0022) (Table 3). As can be seen in

Figure 3 CK-19 mRNA positive cell levels expressed as MCF-7 cell equivalents/5 µg RNA obtained by these two protocols correlated very well ($r = 0.986$, $n = 29$) as can be seen in Figure 3, and did not differ significantly (Wilcoxon test for paired data, $n=29$, $P=0.164$).

Immunomagnetic enrichment of epithelial cells

10 Three sample groups were tested to evaluate the efficacy of new protocols to isolated circulating tumor cells (CTCs) from peripheral blood. Figure 6 shows each sample group along with subsequent manipulation.

15

1st Group (shown as Group A in Figure 6)

The 1st group consisted of peripheral blood samples, spiked with known amounts of MCF-7 cells. These samples were added to a Ficoll Histopaque-1077 system
20 (Sigma Aldrich, St. Louis, MO) and centrifuged at 1,500 rpm for 30 min. The mononuclear cell layer was removed, washed twice with PBS, diluted to 1 mL with PBS/0.1% bovine serum albumin, and incubated with Epithelial Enriched Dynabeads (1 x 10⁷ beads in a
25 volume of 20 µL) while rocking for 1 hour. The cell suspension was placed on a magnet for at least 6 min and the supernatant was carefully removed. The cells attached to the magnetic beads were washed thrice with 1 mL PBS/0.1% bovine serum albumin and lysed with
30 the lysis binding buffer supplied with the kit. The

lysed cell suspension (with beads attached) was stored at -80°C until processing. The MCF-7 epithelial cells were enriched by immunomagnetic capture using the monoclonal antibody, Ber-EP4, and the magnetic
5 Dynabeads Epithelial Enrich kit according to the manufacturer's instructions (Dynal). The manufacturers showed that up to a 5 log enrichment of epithelial cells and a yield of 70% viable, bead-free tumor cells can be obtained using this kit (Dynal).
10 The Ber-EP4 antibody recognizes two glycoproteins on the surface and in the cytoplasm of epithelial cells except the superficial layers of squamous epithelia, hepatocytes, and parietal cells.

2nd Group (shown as Group B in Figure 6)

15 The 2nd group consisted of samples prepared by spiking known amounts of MCF-7 cells in PBS, and following the same procedure as for the 1st group. This group of samples was used as a reference for the recovery of the MCF-7 cells after Ficoll isolation with (1st
20 group) or without (3rd group) immunomagnetic enrichment.

3rd Group (shown as Group C in Figure 6)

The 3rd group consisted of peripheral blood samples, spiked with known amounts of MCF-7 cells, added to
25 Ficoll Histopaque-1077 (Sigma Aldrich, St. Louis, MO) and centrifuged at 1,500 rpm for 30 minutes. The mononuclear cell layer was removed, washed twice with PBS and PBMCs were stored at -80°C until processing.

mRNA isolation and reverse transcription. Total RNA isolation was performed by using Trizol LS reagent (Invitrogen) 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 -80°C until used. RNA concentration was determined with NanoDrop Spectrophotometer ND-1000 (NanoDrop). Reverse transcription of RNA was carried out with the Superscript III Platinum Two Step qRT-PCR kit (Invitrogen).

Real-time PCR (quantitative PCR). Real-time RT-PCR for CK-19 was performed in a total volume of $20\mu\text{L}$ in the LightCycler glass capillaries. For the PCR, $2\mu\text{L}$ of cDNA were placed into a $18\text{-}\mu\text{L}$ reaction volume containing $2\mu\text{L}$ of the PCR Synthesis Buffer minus Mg^{2+} (10x), $1\mu\text{L}$ of MgCl_2 (50 mM), $0.4\mu\text{L}$ dNTPs (10 mM), $0.3\mu\text{L}$ BSA (10 $\mu\text{g}/\text{mL}$), $0.2\mu\text{L}$ Taq platinum DNA polymerase (5 U/ μL) (Invitrogen, USA), $1\mu\text{L}$ of the sense primer CK19-for2 (3 μM), $1\mu\text{L}$ of the antisense primer CK19-do2 (3 μM), $1\mu\text{L}$ of the hybridization probe CK19-FL (3 μM), $1\mu\text{L}$ 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.

Referring now to Figure 8A-C, it can be seen
5 that high sensitivity was achieved when ficoll
separation of peripheral blood mononuclear cells
(PBMC) spiked with MCF-7 cells was followed by
immunomagnetic enrichment. In these experiments,
detection limits down to 1 MCF-7 cell/ml PB was
10 achievable. See Figure 8A.

Additional Uses of the Invention

The present invention discloses, for instance,
15 methods for the quantitative determination of
circulating tumor cells identified in biological
samples of patients. An example is a patient
suffering from breast cancer. Preferred invention
methods use Real-Time PCR amplification of specific
20 CK-19 mRNA transcripts using a primer pair of the
invention.

CK-19, being an epithelial marker abundantly
expressed in tumors, is also a marker (alone or in
25 combination with other markers) for the
identification of circulating tumor cells in
biological samples of patients bearing tumors of
epithelial origin, including endometrial (Ji XQ et
al, Gynecol Oncol. 2006 Feb;100(2):355-60),
30 colorectal (Yeh CS et al, Int J Oncol. 2006
Feb;28(2):411-20; Wang JY et al, World J Surg. 2006

Jun;30(6):1007-13), gastric (Wu CS et al, Int J Cancer. 2006 Jul 15;119(2):373-9), head & neck (Tao L et al, Br J Cancer. 2006 Apr 24;94(8):1164-9), prostate (O'Hara SM et al, Clin Chem. 2004 5 May;50(5):826-35) and malignant pleural effusions caused by various types of tumors (Xe F et al, J Zhejiang Univ Sci. 2004 Oct;5(10):1286-9). Such biological samples may include peripheral blood, bone marrow, lymph nodes, spinal fluid and ocular lens 10 fluid.

To identify CK-19 mRNA positive circulating tumor cells from biological samples derived from patients bearing the aforementioned tumors, one or a combination of the methods disclosed herein can be 15 used. For instance, clinical samples are collected and total RNA prepared using isolated peripheral blood mononuclear cells (PBMCs). The immunomagnetic purification strategy outlined above can be used, for instance. RNA is quantified, if desired, and stored 20 at -70°C for long term storage. Alternatively, the RNA is used (5 µg) to perform a reverse transcription reaction to synthesize cDNA (target sequence). Samples from healthy individuals are used as controls and will be processed in parallel to clinical samples 25 in an identical manner. It will be appreciated that such controls need not be performed if a control sample to be tested has a known CK-19 expression profile. Synthesized cDNAs are used in Real-time PCR reactions using a primer pair and hybridization probe 30 pair as described above to amplify the CK-19 sequence. Use of the primer pair set forth as SEQ ID

Nos. 1 and 2 will be preferred for many applications.

For quantification, an external calibration curve will be prepared by using external standard cDNAs prepared from RNA isolated from 1×10^6 MCF-7 cells as described earlier in the application.

DISCUSSION

The present inventors have developed a specific and sensitive method for quantification of circulating CK-19 mRNA positive cells in peripheral blood samples of breast cancer patients (A. Stathopoulou et al; 2003). Despite the very low false positive rate of this 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 for patient samples showing an amplification curve at a C_p slightly lower than the cut-off has proved to be very difficult and critical. This "gray 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 either genomic DNA contamination or illegitimate expression, 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 pure genomic DNA we have clearly shown that this new primer pair in combination with this

pair of hybridization probes is highly specific and is not affected by the presence of a high concentration of genomic DNA and CK-19 pseudogenes. In retesting the samples from a subgroup (n=62) of
5 the same previously studied healthy volunteers with the new primer pair, we have seen a significant improvement in the specificity of the assay since none of these samples had amplifiable product of CK-19 mRNA.

10 By using this new highly specific pair of primers for the real-time PCR quantification of CK-19 mRNA the present inventors have considerably improved the specificity of this method. In this way, clear
15 distribution between positive and negative samples is achieved and the difficult interpretation of the gray-zone results in the previous assay is completely avoided. For the majority of samples, the two sets of primer pairs give almost the same results. In a total
20 of 222 samples tested, 29 samples were found positive and 169 negative by both primer pairs [concordance of 89.2% (198/222)]. However, for the 10 positive samples whose concentration were very close to the analytical detection limit of the method in protocol
25 A and thus were in the "gray-zone" four were found to be true positives by protocol B (40%), while six were found to be false positives (60%). In the other set of 19 samples that were found negative by protocol A, with concentrations slightly below the analytical detection limit, 2 were found to be false negative by
30 protocol B (10.5%). In this way, a small percentage of patient samples 29/222 (13%) that were in a

"gray-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 co extracted with total RNA.

Review of the following references will enhance appreciation of the present invention.

- 10 A. C. Lambrechts, L. J. Veer, S. Rodenhuis, Ann. Oncol. 9 (1998) 1269-1276.
K. Pantel, V. Muller, M. Auer, N. Nusser, N. Harbeck, S. Braun, Clin. Cancer Res. 9 (2003) 6326-6334.
- 15 S. Braun, K. Pantel, P. Müller, W. Janni, F. Hepp, C. R. Kentenich, S. Gastroph, A. Wischnik, T. Dimpfl, G. Kindermann, G. Riethmuller, G. Schlimok, N. Engl. J. Med. 342 (2000) 525-533.
Y. H. Datta, P. T. Adams, W. R. Drobyski, S. P. Ethier, V. H. Terry, M. S. Roth, J. Clin. Oncol. 12 (1994) 475-482.
- A. Schoenfeld, K. H. Kruger, J. Gomm, H. D. Sinnott, J. C. Gazet, N. Sacks, H. G. Bender, Y. Luqmani, R.C. Coombes, Eur. J. Cancer 33 (1997) 854-25 861.
- A. Stathopoulou, I. Vlachonikolis, D. Mavroudis, M. Perraki, Ch. Kouroussis, S. Apostolaki, N. Malamos, S. Kakolyris, A. Kotsakis, N. Xenidis, D. Reppa, V. Georgoulis, J. Clin. Oncol. 20 (2002) 3404-3412.
- A. Stathopoulou, A. Gizi, M. Perraki, S.

Apostolaki, N. Malamos, D.Mavroudis, V. Georgoulas, E. Lianidou, Clin. Cancer Res. 9 (2003) 5145-5151.

J. A. López-Guerrero, P. Bolufer-Gilabert, M. Sanz-Alonso, E. Barragan-Gonzalez, J. Palau-Perez, J. De la Rubia-Comos, A. Sempere-Talens, S. Bonanad-Boix, Clin. Chim. Acta 263 (1997) 105-116.

P. Ruud, O. Fodstad, E. Hovig, Int. J. Cancer 80 (1999) 119-125.

E. S. Savtchenko, T. A. Schiff, C. K. Jiang, I. M. Freedberg, M. Blumenberg, Am. J. Hum. Genet. 43 (1988) 630-637.

V. Bozionellou, D. Mavroudis, M. Perraki, S. Papadopoulos, S. Apostolaki, E. Stathopoulos, A. Stathopoulou, E. Lianidou, V. Georgoulas, 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. (in press).

A. Stathopoulou, K. Angelopoulou, V. Georgoulas, E.S. Lianidou, Clin Biochem. 34 (2001) 651-659.

C. Tricarico, P. Pinzani, S. Bianchi, M. Paglierani, V. Distante, M. Pazzagli, S. Bustin, C. Orlando, Anal. Biochem. 309 (2002) 293-300.

K. Dheda, J.F. Huggett, S.A. Bustin, M.A. Johnson, G. Rook, A.Zumla, Biotechniques, 37 (2004) 118-119.

I.R. Peters, C.R. Helps, E.J. Hall, M.J. Day, J. Immunol. Methods 286 (2004) 203-217.

DEMANDE OU BREVET VOLUMINEUX

LA PRÉSENTE PARTIE DE CETTE DEMANDE OU CE BREVET COMPREND PLUS D'UN TOME.

CECI EST LE TOME 1 DE 2
CONTENANT LES PAGES 1 À 60

NOTE : Pour les tomes additionels, veuillez contacter le Bureau canadien des brevets

JUMBO APPLICATIONS/PATENTS

THIS SECTION OF THE APPLICATION/PATENT CONTAINS MORE THAN ONE VOLUME

THIS IS VOLUME 1 OF 2
CONTAINING PAGES 1 TO 60

NOTE: For additional volumes, please contact the Canadian Patent Office

NOM DU FICHER / FILE NAME :

NOTE POUR LE TOME / VOLUME NOTE:

AMENDED CLAIMS

received by the International Bureau on 12 February 2007 (12.02.2007)

1. (cancelled)
2. (cancelled)
- 5 3. (cancelled)
4. (cancelled)
5. (cancelled)
6. (amended) Primer pair consisting of one primer having the sequence of SEQ ID NO: 1 and one
10 primer having the sequence of SEQ ID NO: 2.
7. (amended) Primer pair according to claim 6, wherein the primers comprises nucleotides, analogues of nucleotides or labels.
8. (amended) Method of quantitative determining
15 the presence in a test sample of a mRNA derived from a gene comprising at least one intron using a pair of primers, wherein at least one of said primers comprises at least one intron-spanning site comprising the steps of
20 (i) forming a reaction mixture comprising nucleic acid amplification reagents, the primer pair of claim 6 and a test sample;
(ii) . subjecting the mixture to amplification
25 conditions to generate at least one copy of a nucleic acid sequence complementary to the target sequence; and
(iii) determining the amount of the mRNA in the sample using real-time monitoring
30 during PCR.
9. Method according to claim 8, wherein the

AMENDED SHEET (ARTICLE 19)

test sample is subjected to reverse transcription prior to forming the reaction mixture of step (i).

10. Method according to claim 8 or 9, wherein the test sample is selected from a blood sample, a
5 sample from the bone marrow and a sample derived from the lymph nodes.

11. Method according to claim 10, wherein the sample is a blood sample.

12. Method according to claim 11, wherein the
10 blood sample is centrifuged before forming the reaction mixture.

13. Method according to any of the claims 8-12, wherein the reaction mixture of step (i) further comprises a housekeeping primer pair.

14. Method according to claim 13, wherein the
15 housekeeping primer pair has the sequence according to SEQ ID 5 and 6.

15. (amended) Diagnostic method of determining the prospects of adjuvant therapy in a patient
20 suffering from cancer comprising the steps of

- (i) providing a biological sample from the patient;
- (ii) isolating nucleic acids from the biological sample;
- 25 (iii) optionally reverse transcribing the isolated nucleic acids, when the origin of the nucleic acid is RNA;
- (iv) forming a reaction mixture comprising nucleic acid amplification reagents, the
30 primer pair of claim 6 and an aliquot of the nucleic acids isolated in step (ii)

- or the reverse transcribed nucleic acid of step (iii);
- (v) subjecting the mixture to amplification conditions to generate at least one copy of a nucleic acid sequence complementary to the target sequence;
- 5 (vi) quantitating the CK-19 mRNA positive cells in the sample using real-time monitoring during PCR; and
- 10 (vii) based on the amount of CK-19 mRNA positive cells in the sample determining the prospects of adjuvant therapy.

16. Diagnostic method according to claim 15, wherein the cancer is breast cancer, preferably operable breast cancer.

15

17. Diagnostic method according to claim 15 or 16, wherein the biological sample is a blood sample.

18. Diagnostic method according to any of the claims 15-17, wherein the reaction mixture of step (iii) further comprises a housekeeping primer pair.

20

19. Method according to claim 18, wherein the housekeeping primer pair has the sequence according to SEQ ID 5 and 6.

20. (amended) Diagnostic kit for amplifying a CK19 target sequence comprising:

25

- (i) the primer pair of claim 6;
- (ii) optionally sequences hybridizing to additional markers on cancer cells;
- (iii) amplification reagents.

21. Diagnostic kit according to claim 20, wherein the amplification reagents comprise the

30

hybridization probes of SEQ ID NO 3 and 4.

22. Diagnostic kit according to claim 20 or 21, wherein the kit further comprises a housekeeping primer pair, preferably with the sequences according to SEQ ID NO5 and 6.

23. Diagnostic kit according to any of the claims 20 to 22, wherein the amplification reagents and the primer pair(s) are lyophilized.

24. (amended) A primer pair comprising a first primer having at least about 8 nucleobases of the sequence represented by SEQ ID NO: 1 or SEQ ID NO: 2, the first primer being from between about 10 to about 20 nucleobases in length and further being capable of specifically hybridizing with the CK-19 cDNA sequence shown in Figure 1,

and further comprising a second primer having at least about 8 nucleobases of the sequence represented by SEQ ID NO: 1 or SEQ ID NO: 2, the second primer being from between about 10 to about 20 nucleobases in length and further being capable of specifically hybridizing with the CK-19 cDNA sequence shown in Figure 1.

25. (cancelled)

26. The primer pair of claims 24-25, wherein at least one of the first and second primers comprises a modified internucleoside linkage, modified sugar moiety or modified nucleobase.

27. The primer pair of claim 26, wherein at least one of the first and second primers includes at least one locked nucleic acid (LNA) unit.

28. A method of determining the presence in a

test sample of a mRNA derived from a gene comprising at least one intron using a pair of primers, wherein at least one of said primers comprises at least one intron-spanning site comprising the steps of

- 5 (iv) forming a reaction mixture comprising nucleic acid amplification reagents, the primer pair of claims 24-27 and a test sample;
- 10 (v) subjecting the mixture to amplification conditions to generate at least one copy of a nucleic acid sequence complementary to the target sequence; and
- 15 (vi) determining the amount of the mRNA in the sample using real-time monitoring during PCR.

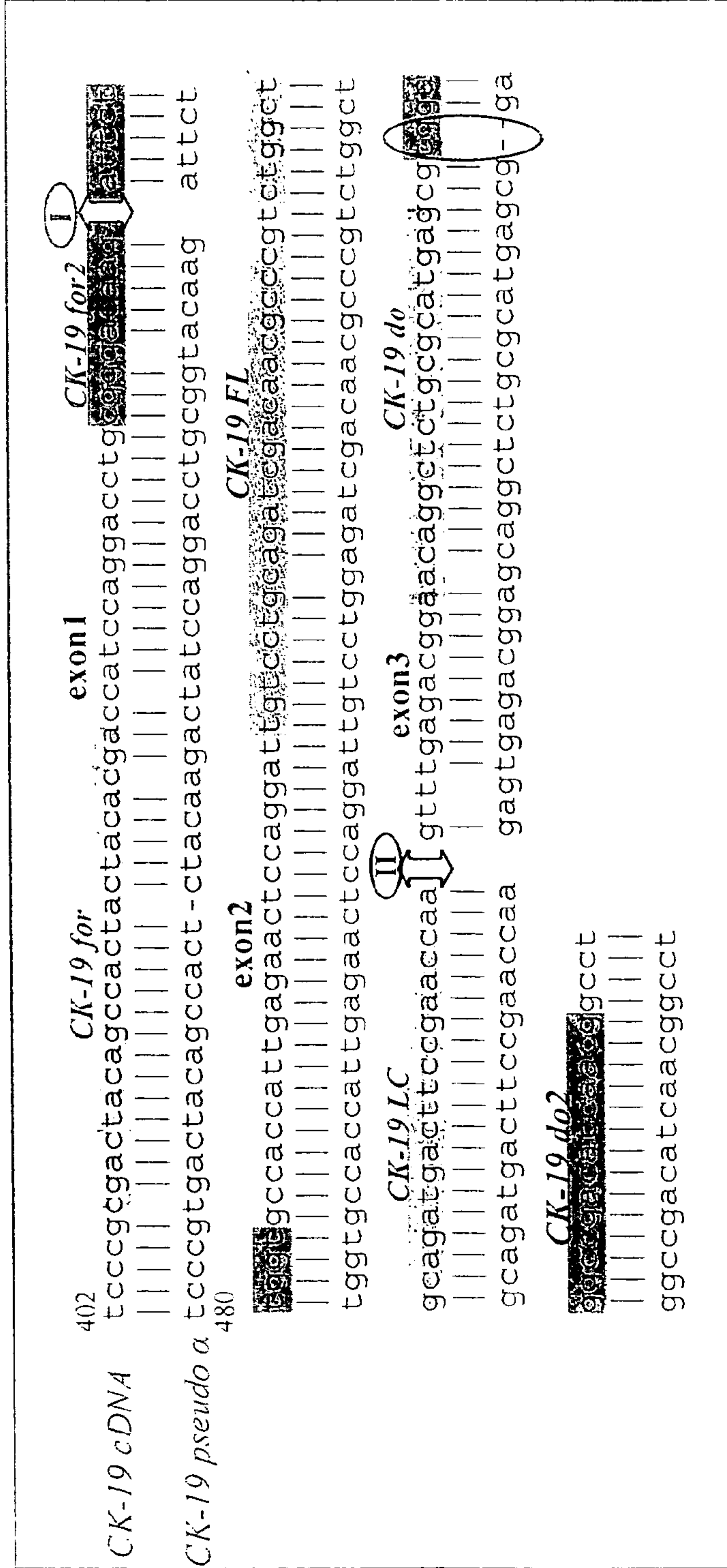
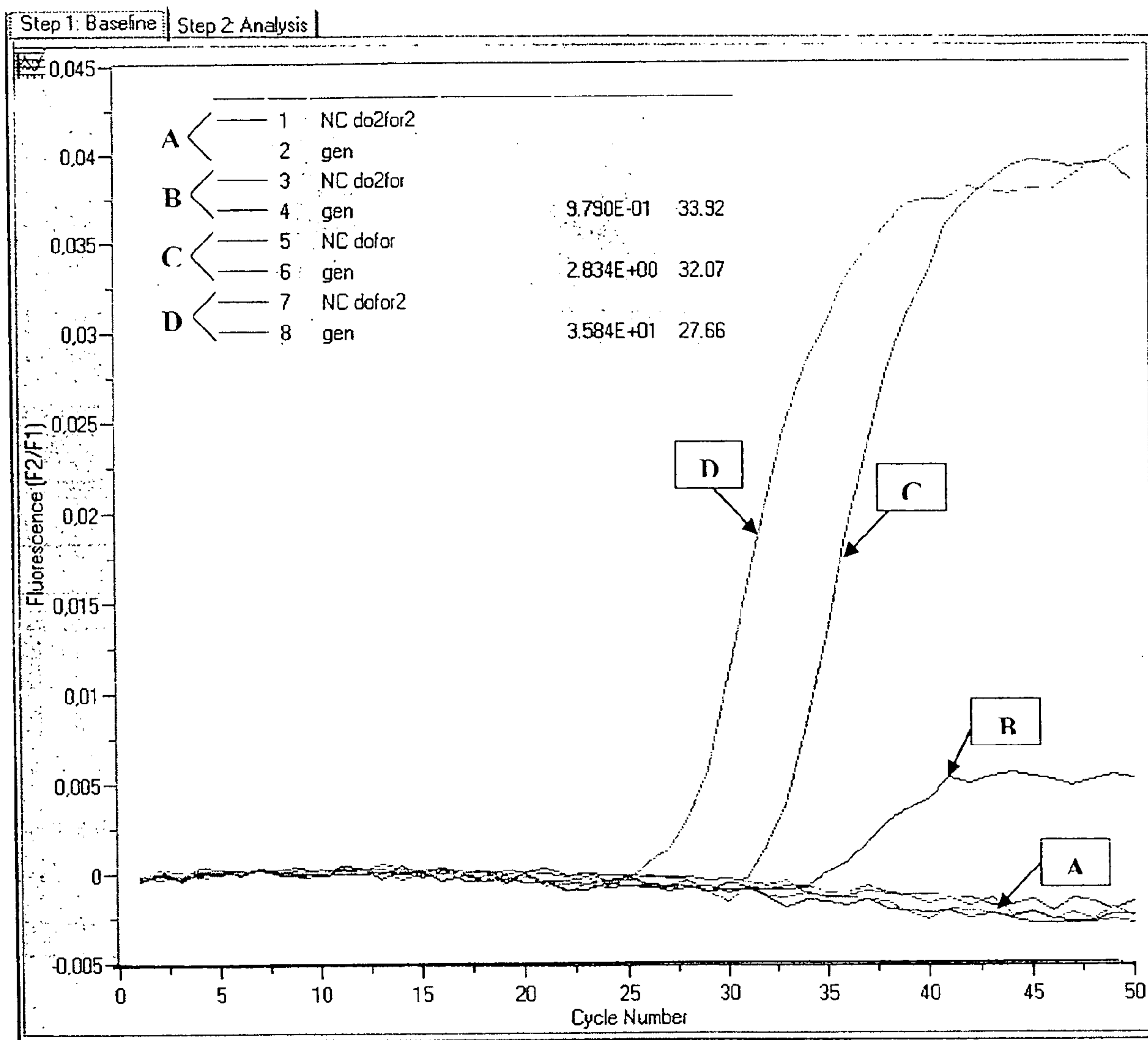


Figure 1 CK-19 cDNA and CK-19 pseudo α gene sequence alignment and hybridization sites for primers and probes used in protocols A and B. Points I and II represent junctions between exons 1/2 and exons 2/3, respectively.

2/10

**Figure 2**

Real time PCR for genomic DNA by using four 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]

3/10

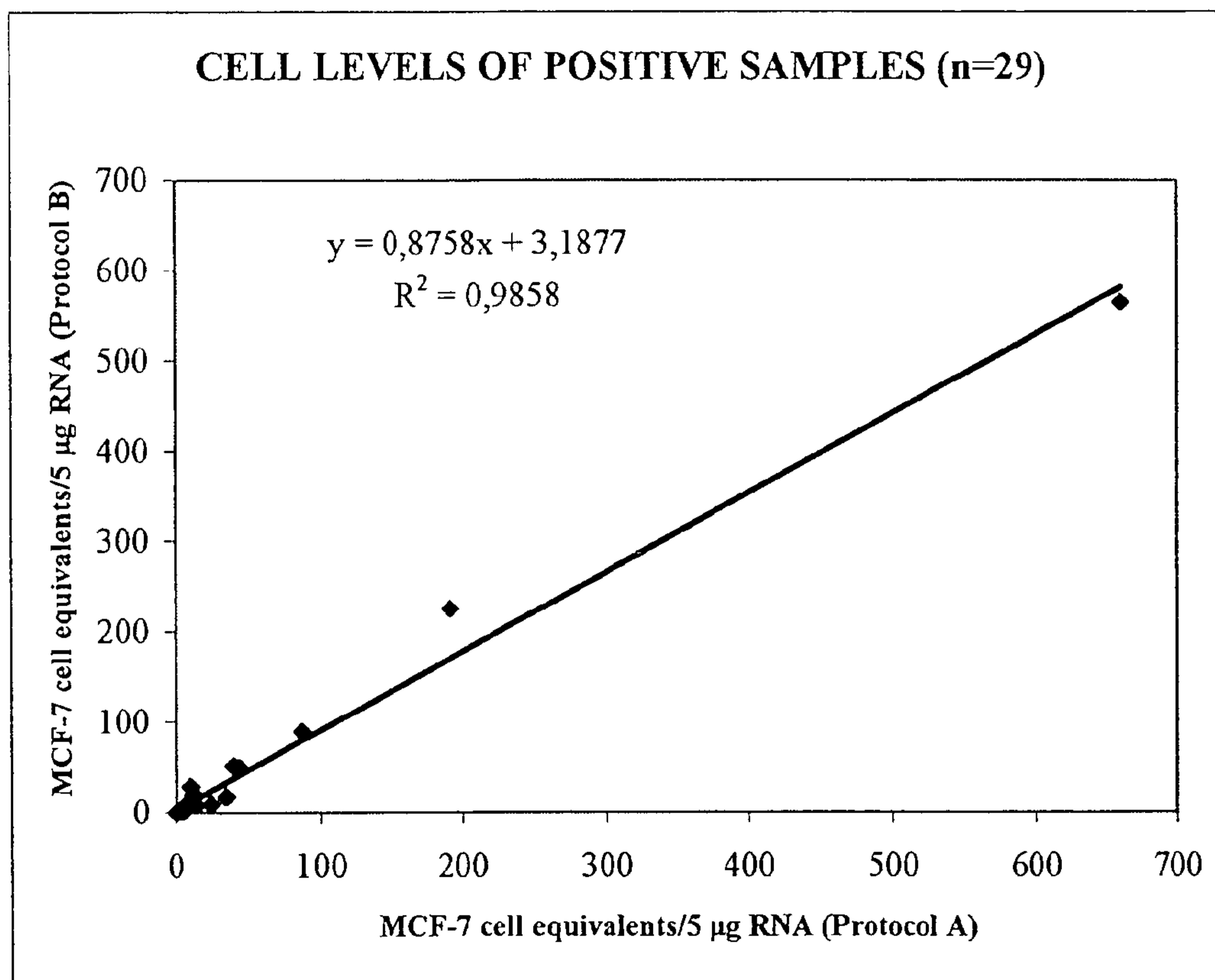


Figure 3
CK-19 mRNA positive cell levels expressed as MCF-7 cell equivalents/5µg RNA obtained by protocols A and B.

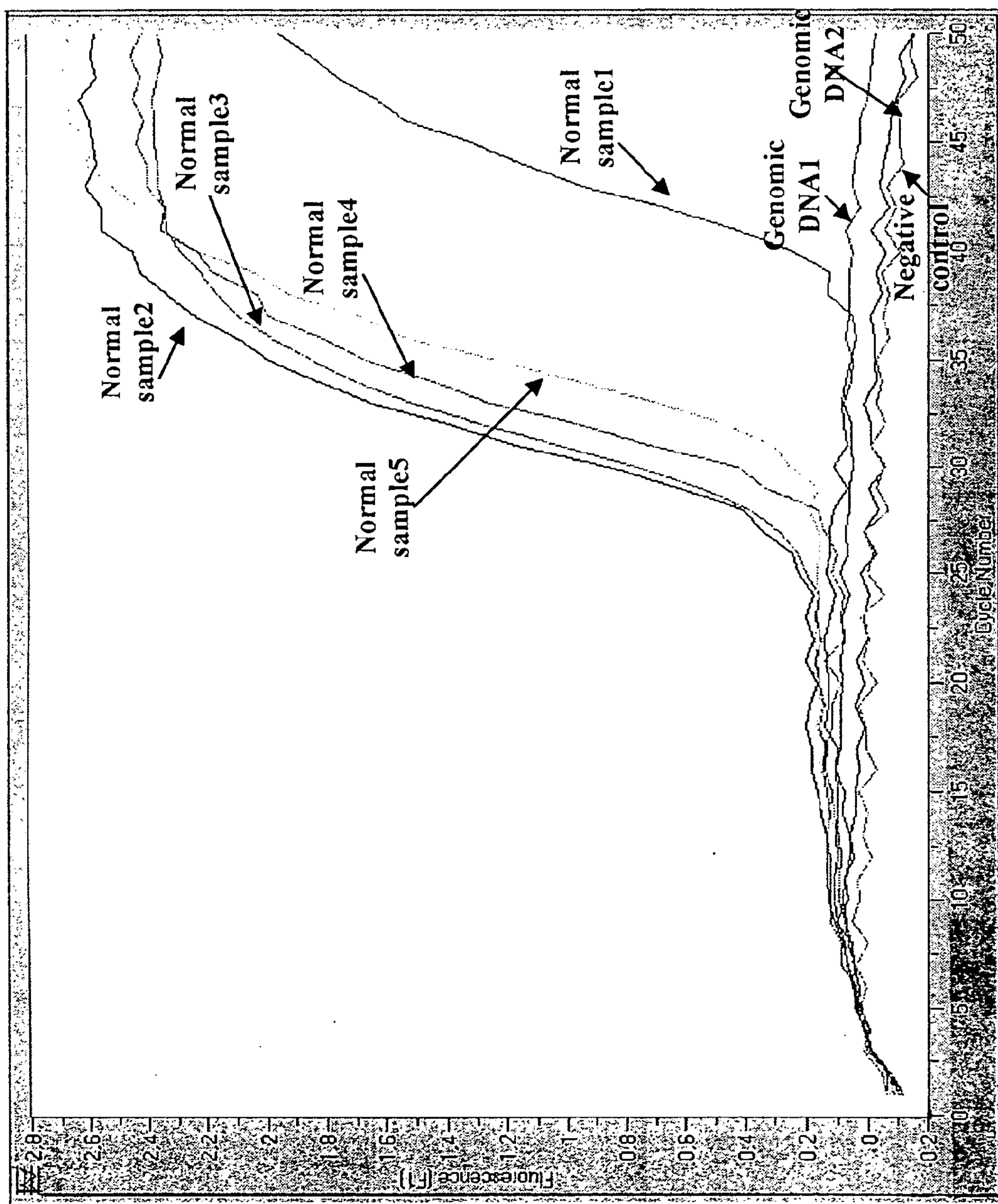


Figure 4
A typical real-time PCR graph for PBGD

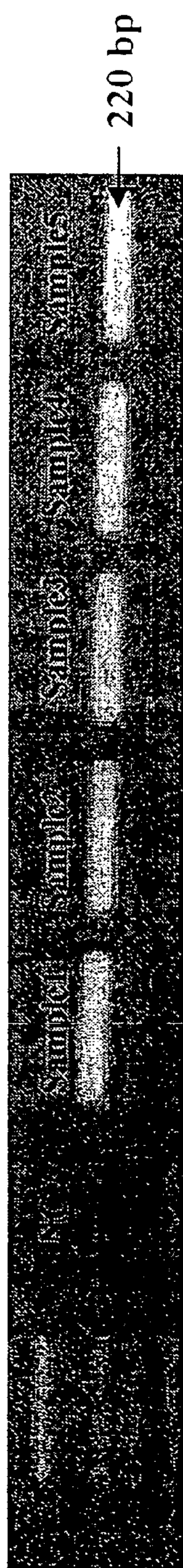


Figure 5.

An agarose gel electrophoresis (2%) for the PBGD PRC products

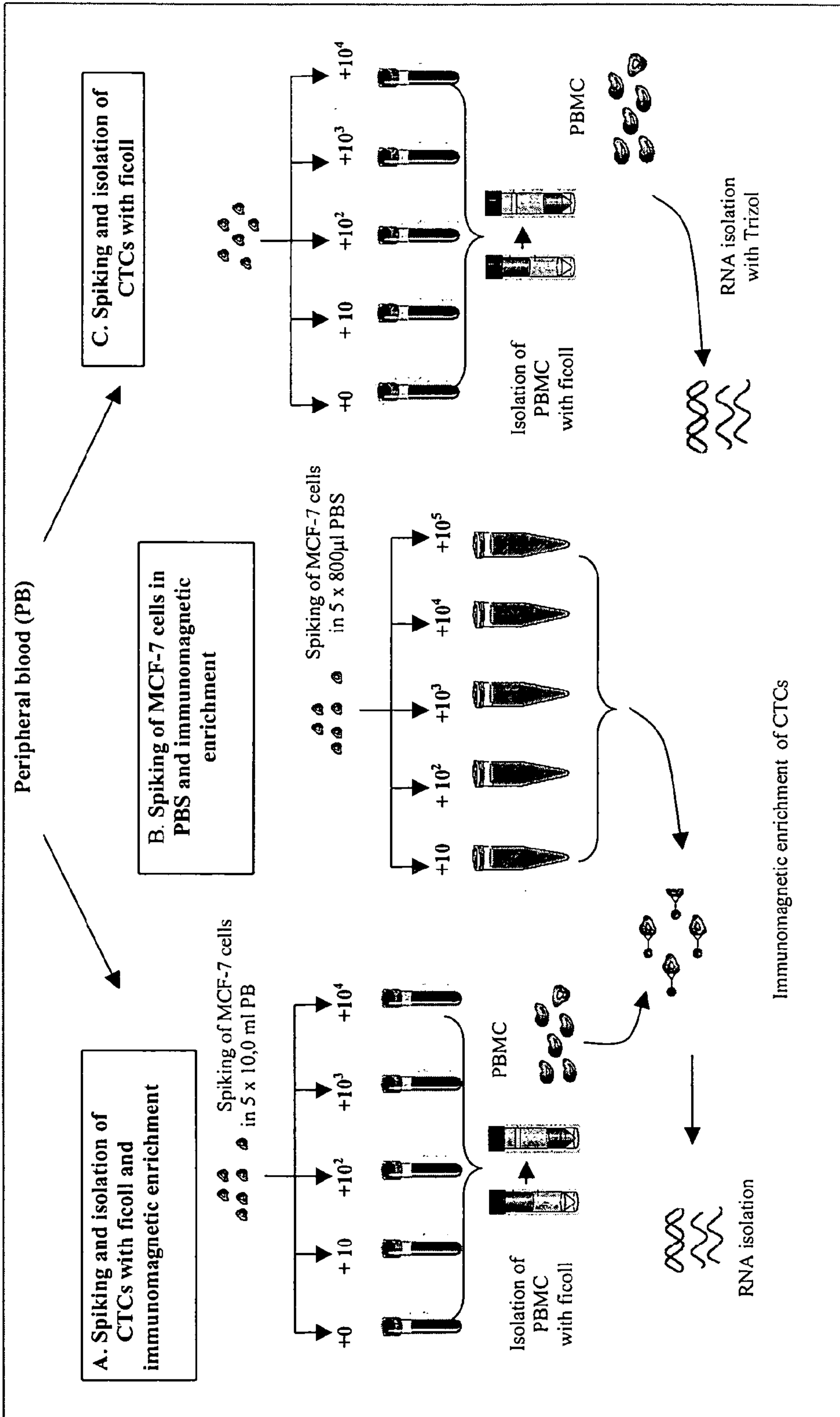


Figure 6

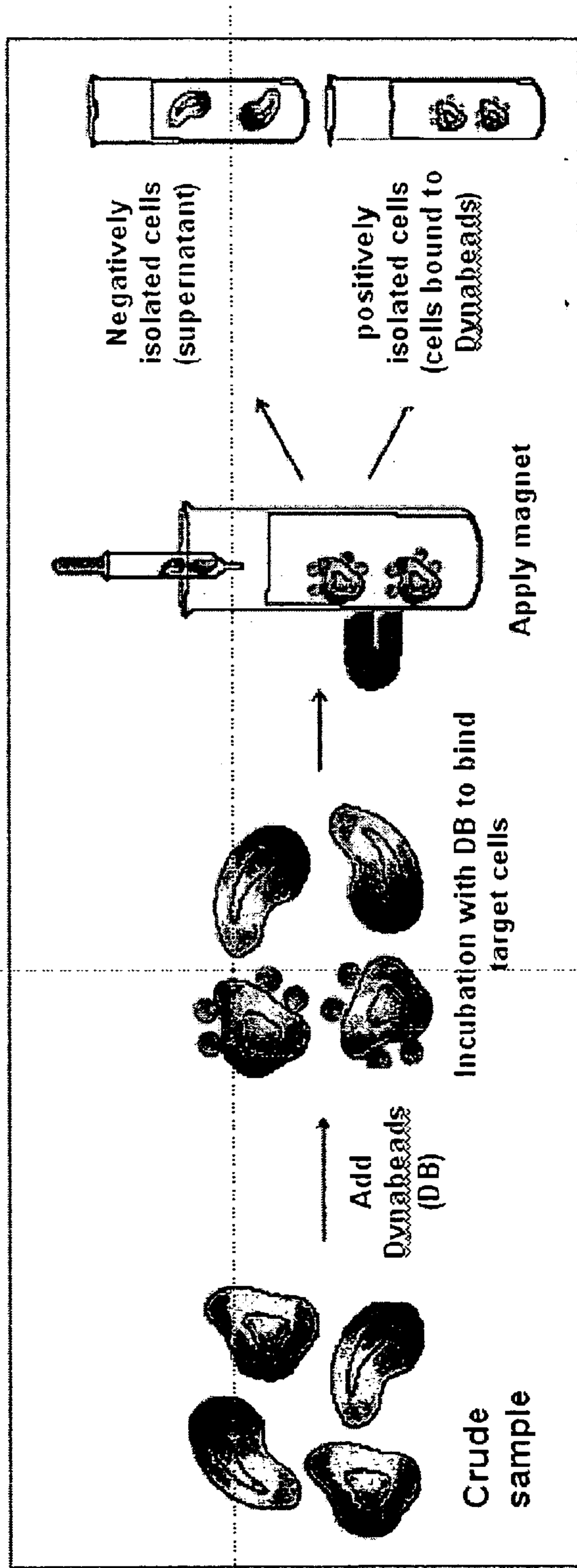


Figure 7

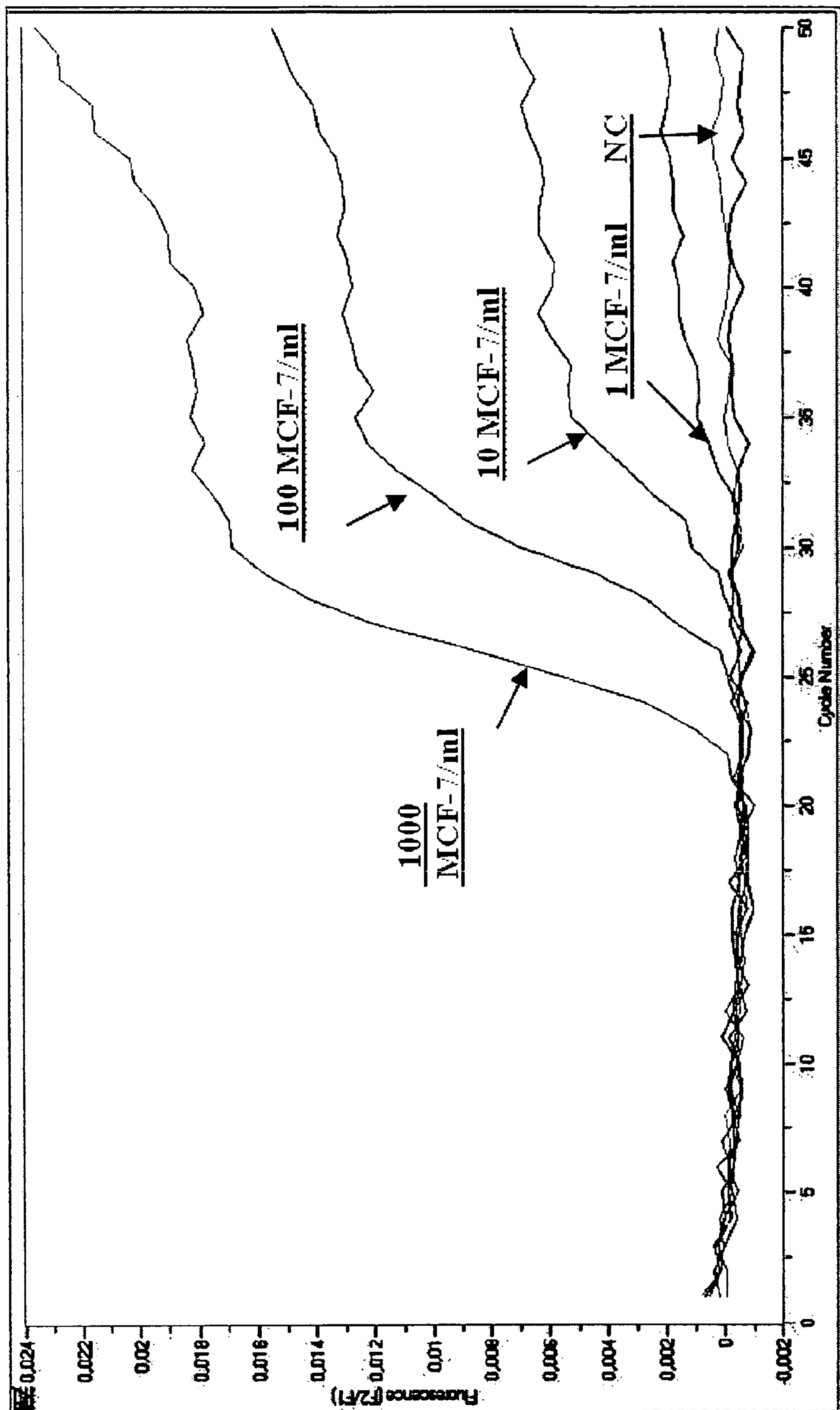


Figure 8A

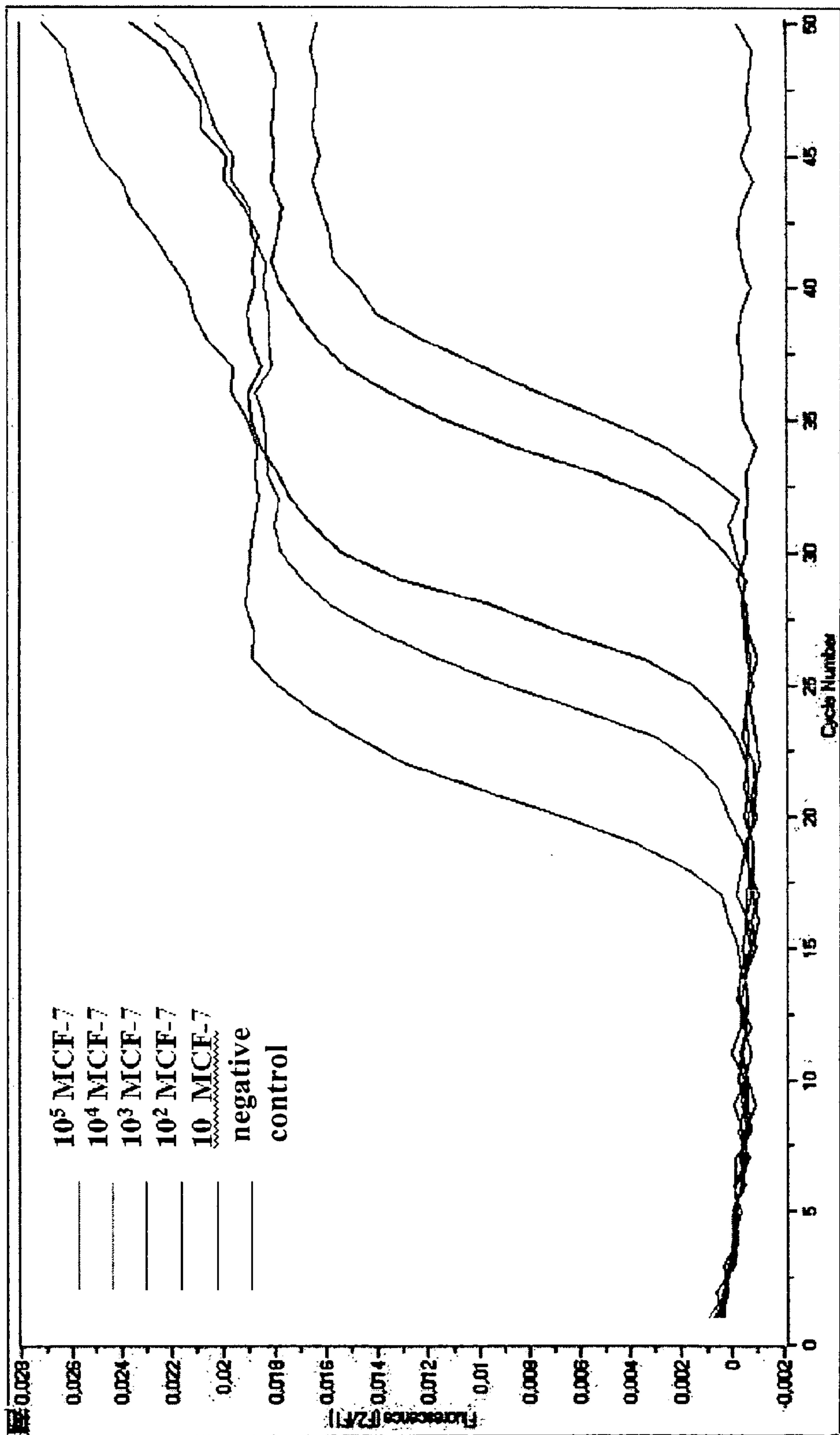


Figure 8B

10/10

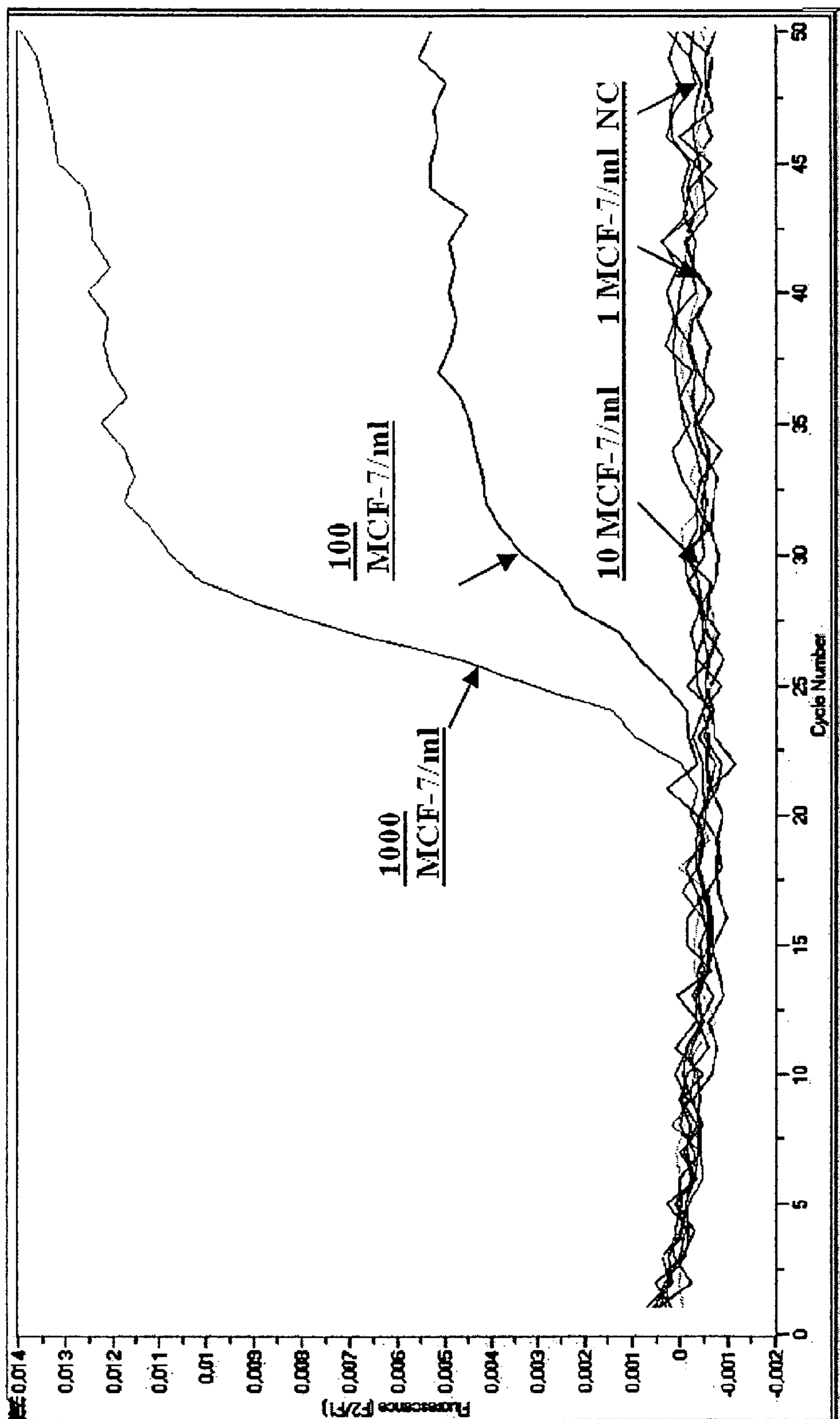


Figure 8C

