DDT Vol. 6, No. 20 October 2001

Human disease characterization: real-time quantitative PCR analysis of gene expression

James V. Snider, Mark A. Wechser and Izidore S. Lossos

Real-time quantitative PCR is the measurement of a fluorescent signal generated and measured during PCR as a consequence of amplicon synthesis. When used as reverse transcriptase-PCR (RT-PCR), real-time quantitative PCR has proved to be useful in accurately measuring expression levels of specific gene transcripts. When applied to questions of minimal residual disease, the technique has evolved from generically detecting the presence of disease cells in individuals, such as the AML1–ETO fusion transcript, to the identification of a specific gene, such as BCL-6, which is prognostic for determining the therapeutic outcome of patients with diffuse large B-cell lymphoma.

James V. Snider
***Mark A. Wechser**
Applied Biosystems
850 Lincoln Centre Drive
Foster City
CA 94404, USA
*tel: +1 800 874 9868
fax: +1 650 638 6045
e-mail: wechsema@applied-biosystems.com
Izidore S. Lossos
Division of Oncology
Rm 1100
Dept of Medicine
Stanford University
269 Campus Drive
CCSR Blvd
Palo Alto
CA 94305, USA

▼ The analysis of gene expression patterns is becoming increasingly important to our understanding of complex biological processes and their roles in disease pathology at the molecular level. The goal is to determine the roles that specific genes have in regulatory biology and pathogenesis by understanding how their function and alteration in expression contribute to disease. Identification of genes with pivotal roles in the disease process might provide novel targets for the design of specific diagnostic tests or gene-specific therapeutic interventions. To achieve this goal, it is necessary to gather gene-expression data on a wide variety of processes, such as cell- and tissue-specific expression, differentiation, response to stimuli, embryonic development and disease. Traditional methods for studying gene expression have significant limitations and, therefore, decreased use as the demand for rapid, quantitative and robust methods for the simultaneous evaluation of multiple genes increases.

Several techniques have been used over the years for the analysis of gene expression. Northern blots and RNase protection assays are the oldest and most well-established techniques. However, the value of these techniques is severely constrained by their requirements. For instance, northern blots require large amounts of RNA and are time consuming both in terms of personnel time and the time to obtain a result. In addition, northern blots use radioactive labels and enable the analysis of relatively few samples simultaneously. RNase protection assays require less RNA and are more sensitive than northern blots but are not amenable to analyzing large numbers of samples. This method is also time intensive and requires the use of radioactive labels. Most modern laboratories actively pursue technological options that do not require the use of radioactive labels.

DNA microarrays are a new tool used to measure the expression of tens of thousands of genes in parallel. In theory, this technique enables genome-wide evaluation of gene expression. DNA microarrays are useful for identifying genes that are coordinately regulated and whose products function either in a common pathway, such as a differentiation program, or in a physiological response or pathological process. These arrays have been used to identify disease subtypes with distinct clinical outcomes, as was recently reported for diffuse large B-cell lymphoma (DLBCL; Ref. 1). Microarray hybridization, however, has the following limitations: it requires large amounts of RNA, is time consuming, labor intensive, and can only be used as an investigational tool that cannot be applied to daily clinical usage. Thus, a quantitative method was needed that could translate laboratory discoveries from the 'bench to the patient's bedside', at

1359-6446/01/$ – see front matter PII: S1359-6446(01)01988-2

the same time being easy to perform, robust, reproducible and economical in its RNA requirements.

PCR techniques have become the preferred method for most nucleic acid analysis projects. When applied to questions of gene expression analysis, where RNA is used as the amplification template, the technique is termed reverse transcriptase PCR (RT-PCR). PCR techniques have several advantages over the northern blot and RNase protection assays: PCR is extremely sensitive, being performed successfully on samples as small as a single copy of template. The process is simple and is amenable to the rapid analysis of large numbers of samples.

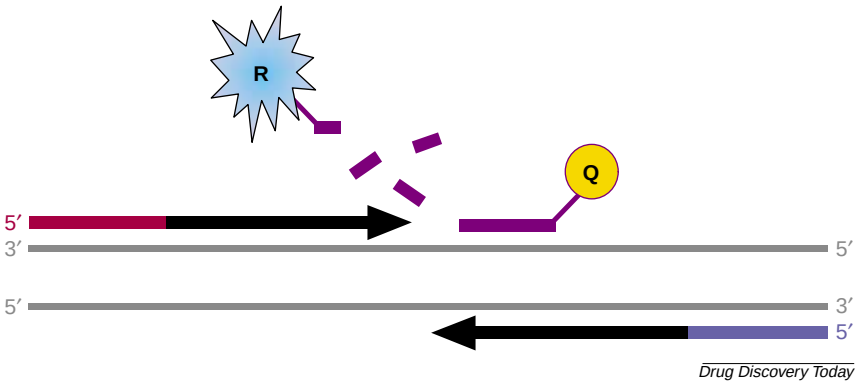

Figure 1. TaqMan® probe cleavage step. In a TaqMan® assay, *Taq* polymerase extends both PCR primers and one of the primer extension complexes will begin displacing and cleaving the labeled probe, resulting in a release of the reporter dye (R) and the generation of a fluorescent signal. This illustration shows a diagrammatic representation of this displacement and cleavage step. Abbreviation: Q, quencher dye.

However, PCR does have limitations: it is sensitive to inhibitors that can be present in the starting sample. Although PCR is regarded as a powerful qualitative detection technique, there are several inherent factors that have limited its value as a quantitative technique: (1) the plateau of product accumulation; and (2) variable reaction efficiency across different samples in the analysis. In an effort to overcome these two limitations, significant time and effort were invested to develop the technique of quantitative competitive PCR (Ref. 2). In quantitative competitive PCR a reference template is included in the reaction with the target sample. Typically, many duplicate assays will be performed where the reference template is included at varying concentrations. The objective being that at least one of the reference template concentrations will equal the concentration of the target template in the source sample and that the resulting PCR products will be generated in equal amounts.

Although this technique has been used successfully to quantify concentrations of target templates in samples, it is subject to significant limitations. The most substantial of these is that a reference template must be created and the assay must be optimized thoroughly for each target. Moreover, because many replicate assays must be performed it is difficult to analyze large numbers of samples. Consequently, quantitative competitive PCR has never been widely employed for gene-expression analysis projects.

Real-time PCR

A recent technological advancement, real-time PCR, is quickly becoming accepted for the routine quantitative analysis of RNA targets for gene-expression studies[3,4]. Real-time PCR, sometimes referred to as kinetic PCR, is a process by which a fluorescent signal is generated as the target sequence is amplified, and this fluorescent signal is measured in the reaction tube, during the PCR process[5–7]. There are two fundamental techniques used to generate this fluorescent signal. One is to include a dye, such as SYBR® Green, which binds double-stranded DNA. In this technique, any amplicon that is synthesized will bind the dye and generate a fluorescent signal that is detected by the instrument. The second, more widely used technique, uses a sequence-specific fluorescent probe that hybridizes to the target sequence during the annealing and extension phases of the reaction. The probe uses fluorescent resonance energy transfer (FRET) technology, whereby the emission spectrum of a fluorescent reporter dye at its 5′ end is effectively quenched by a second fluorescent dye at its 3′ end when the probe is intact[7]. During PCR the probe can hybridize to the target sequence after the denaturation step. The fluorescent signal is produced during primer extension. The 5′ nuclease assay method, first described by Holland and coworkers[4], then refined by Lee and coworkers[8] to include fluorescent labels, and subsequently commercialized by Applied Biosystems (Foster City, CA, USA) in 1996, is illustrated in Fig. 1.

During PCR, the forward and reverse primers hybridize to their cognate binding sites in the target nucleic acid and the TaqMan® probe hybridizes to its target sequence between the primers. As the PCR primers are extended during the extension phase of each amplification cycle, one of the strand extension complexes will run into the 5′ end of the labeled probe. *Thermus aquaticus (Taq)* DNA polymerase displaces the opposing probe sequence and, when the fork-like structure is created, the probe is nucleolytically cleaved by the 5′-nuclease activity of *Taq* DNA polymerase.

This cleavage causes the reporter dye and the quencher dye to separate, resulting in increased fluorescence of the reporter dye. This process continues until the entire probe is displaced and the amplicon is completed. This homogeneous assay is simple to perform and is becoming a *de facto* standard for nucleic-acid quantification procedures.

Instrumentation

Dedicated instrument systems are now available to measure real-time fluorescent signals based on FRET probe or SYBR® Green assays. The ABI PRISM® 7700 Sequence Detection System, the first such instrument, was introduced by Applied Biosystems in 1996. Applied Biosystems now has an offering of several instruments available to support these assays.

The 5′ nuclease assay and real-time PCR technology provide significant advances in the study of gene expression. The assay is homogeneous and there is no need for post-PCR processing. The measurement of a signal is achieved automatically during PCR, making this technique useful for studies requiring large numbers of samples. Because real-time PCR with TaqMan® probes can use dyes with different emission spectra for individual assays, the technique enables multiplexing of assays (i.e. more than one target can be measured in the same tube). This technique is highly reproducible and has a broad dynamic range. The large dynamic range and reproducibility are intrinsic to this technology because the signal is measured during the exponential phase of the amplification process. During the exponential phase, when the amplicon concentration is effectively doubling with each cycle, the PCR is robust. When a PCR enters plateau phase, the reaction becomes much less robust and the amplicon concentration might not continue to increase at all. Measuring the amount of amplicon present during the exponential phase of each cycle via the fluorescent product of probe cleavage or dye incorporation permits precise comparative measurements of different amplicons in the same or in different wells. Because the amount of product generated and quantified from each cycle is directly proportional to the initial copy number or concentration of the template, more abundant templates are detected earlier during the PCR than less abundant templates[5,9]. This fact has created a new paradigm for quantification using PCR. In this new paradigm, the cycle at which the product of the PCR crosses the detection threshold (Ct) is a direct measure of the target's abundance. Thus, we use the point in time when the target amplification is first detected as our metric for quantification rather than the amount of material at the end of the PCR.

Minimal residual disease

Before the availability of real-time analysis instruments, quantitative competitive PCR was commonly used for nucleic acid quantification assays. Wattjes and coworkers[10] performed a direct comparison between quantitative competitive PCR and the 5′ nuclease assay in which the sensitivity and accuracy of each technique were compared. This study measured expression of the AML1–ETO fusion transcript, generated by chromosomal translocation (8;21) that can be detected in acute myelogenous leukemia (AML) samples. In one experiment t(8;21) cells (Kasumi-1) were diluted in a negative cell-line (NB4) over a range of 1:0 to $1:10^5$. These results are summarized in Table 1.

The conclusion reached by Wattjes and colleagues[10] is that the 5′ nuclease assay used in real-time PCR and quantitative competitive PCR are qualitatively similar but that the 5′ nuclease assay is more accurate, more amenable to processing large numbers of samples and is easier to perform. This conclusion is supported by earlier work by Marcucci and colleagues[11], where the 5′ nuclease assay was used to detect AML1–ETO fusion transcripts as a marker of minimal residual disease (MRD). In this study, every patient who experienced AML relapse had AML1–ETO fusion transcripts >1000-fold higher than the β-actin reference. This level of AML1–ETO fusion transcript expression is 2–4 logs greater than that measured in patients in clinical remission (CR).

Kondo and coworkers developed a 5′ nuclease assay for this same AML1–ETO

Table 1. Comparison of quantitative competitive polymerase chain reaction (PCR) and the TaqMan® 5′ nuclease assay using Kasumi-1:NB4 cell-line dilution series

Kasumi-1:NB4	Quantitative competitive (equivalence point)	TaqMan standard curve (quantity)
1:0	10^6	7.3×10^5
1:10	10^5	1.2×10^5
1:100	10^4–10^3	3.8×10^4
$1:10^3$	10^3	3.0×10^3
$1:10^4$	10^2	5.7×10^2
$1:10^5$	10^2–10^1	2.0×10^2

The sensitivity of quantitative competitive PCR was determined by a serial-dilution of cells expressing cloned AML1/MTG8 sequence and competitor molecule, and this technique was compared with TaqMan® for sensitivity.

fusion transcript and applied it to the detection of MRD in AML patients harboring the fusion product. These studies showed that 5′ nuclease assay and real-time PCR provide a sensitive and accurate method for quantifying transcripts in AML patients, including those in CR. The authors did not infer any specific clinical interpretation from this data because their patient set was small – only 14 individuals[12]. Yet, this technique does provide simple, accurate quantification of a transcript with potentially high prognostic value. This work is one example of a rapidly growing body of research in which real-time PCR technology is being applied to the study of MRD. Although complete remission can be achieved in the majority of patients with acute lymphoblastic leukemia (ALL), a significant proportion – depending mainly on pre-therapeutic risk-factors – relapse and die of their disease[13]. Accurate quantification of MRD at the molecular level combined with the identification of patients at higher risk for ALL relapse could provide a significant opportunity to improve the therapeutic management of these patients.

MRD can be assessed using real-time PCR techniques with assays developed to detect specific markers. These markers include: (1) clonal tumor markers, such as rearranged immunoglobulin genes in multiple myeloma, chronic lymphocytic leukemia and B-cell ALL, T-cell receptor gene rearrangements in T-cell ALL; (2) chromosomal translocation markers such as t(14;18) in non-Hodgkin's lymphomas (NHL) or; (3) novel fusion products generated by chromosomal rearrangements, such as AML1–ETO in AML with translocation t(8;21), BCR–ABL in chronic myeloid leukemia and PML–RARA in acute promyelocytic leukemia[14–19]. These research reports have clearly shown that real-time PCR is a powerful tool for detecting clonal disease cells. Because patients in CR might have low but quantifiable levels of malignant clonal cells, the true prognostic value of this data is yet to be determined. It is necessary to define and establish quantitative criteria for each evaluated gene that will discriminate between patients with high or low risk of relapse.

Gene expression and minimal residual disease

The use of real-time PCR assays is certainly a valuable tool in characterizing MRD, but the technique is also being successfully applied to the study of other human disease states. For instance, telomerase activity is the most general molecular marker for the identification of human cancer and can be detected in 85% of all tumors, whereas most healthy tissues exhibit little or no telomerase expression[20]. De Kok and coworkers[21] developed a 5′ nuclease assay specific for the catalytic subunit of human telomerase (hTERT) and measured the expression level of this gene in both normal and tumor samples from several tissues. In some tissue types the results were distinct. For example, normal pancreatic tissue did not express hTERT above the detection limit, whereas all tumor samples did. Normal lung tissue had low levels of hTERT expression, whereas tumor tissues had hTERT expression levels that were at least five-fold higher. However, the ability to draw conclusions from these results was somewhat obscured by the heterogeneity of the tumor samples. For example, different extents of tissue infiltration by tumor cells could explain some of the observed results.

Finally, Lossos and colleagues[22] demonstrated how gene expression quantification by real-time PCR can complement genome-wide expression data from DNA arrays and use this approach to identify patients with DLBCL who have poor prognoses and who might require specifically tailored aggressive therapy. DLBCL represent the most frequent type of NHL, accounting for 30 to 40% of adult NHL. DLBCL is a diverse group of lymphoid neoplasms with heterogeneous clinical, histological, immunophenotypic, cytogenetic and molecular genetic features. Alizadeh and coworkers[1] evaluated DLBCL gene-expression profiling by DNA microarray techniques and identified two molecularly distinct forms of DLBCL: germinal center-like DLBCL, characterized by the expression of genes normally expressed in germinal center (GC) B cells, and activated B cell-like DLBCL characterized by expression of genes normally induced during *in vitro* activation of B cells.

Considerable gene expression heterogeneity was present within each DLBCL subgroup and no single gene, as assessed by microarray expression data, correlated absolutely in expression with either of these two DLBCL subgroups. Patients with the GC and activated-B-cell-like forms of DLBCL were found to have different prognoses: those with GC-like DLBCL had a significantly better overall survival (OS) than those with activated-B-cell-like DLBCL. These findings suggest that the cellular origin of DLBCL might determine lymphoma aggressiveness and tumor progression. However, the inherent limitations of DNA arrays prevented the application of this method to clinical use. Lossos and colleagues[22] designed a 5′ nuclease assay based real-time PCR method to quantify the expression of the BCL-6 gene, which is expressed in GC lymphocytes and germinal center-like DLBCL, and evaluated whether it might serve as a prognostic indicator for OS in DLBCL patients. To assess the validity of the method, BCL-6 gene expression was measured in K562 and HL60 cell lines that had been previously shown not to express BCL-6, and in GC lymphocytes and follicle center lymphoma (FCL) specimens that are known to express BCL-6 (Fig. 2). Barely detectable levels of the BCL-6 were found in the K562 and

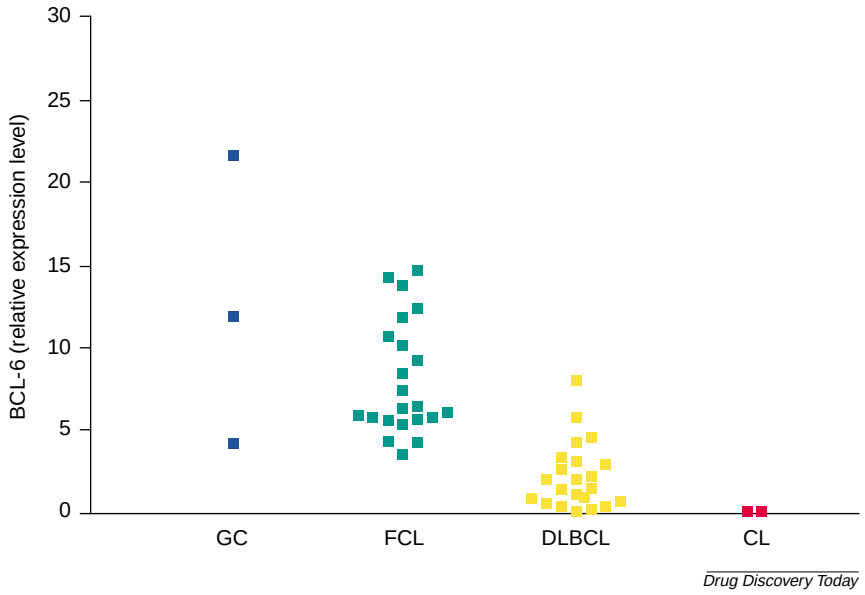

Figure 2. BCL-6 mRNA extension in unknown samples. To assess the validity of the technique, the expression of BCL-6 mRNA, relative to a glyceraldehyde-3-phosphate dehydrogenase (GADPH) control transcript, was evaluated in germinal center (GC) cells obtained from three tonsils from normal individuals (blue), 23 follicle-center lymphoma (FCL) specimens (green), 22 diffuse large-B-cell-lymphoma (DLBCL) specimens from the derivation set (yellow) and K562 and HL60 cell lines (CL; red). The CL cells are known to not express BCL-6, confirmed in this assay, whereas GC, FCL and DLBCL cells are known to express BCL-6 at different levels, also confirmed in this assay. Reproduced, with permission, from Ref. 1.

HL60 cell lines in agreement with previous evaluations of BCL-6 gene expression by northern blot analysis. BCL-6 gene expression was detected, as expected, in all the GC lymphocyte preparations and in all the FCL specimens. A similar range of BCL-6 gene expression was observed in GC cells and in FCL specimens.

Measurement of BCL-6 mRNA expression in DLBCL patients demonstrated that BCL-6 expression is an independent prognostic factor for the prediction of OS of DLBCL patients. BCL-6 mRNA expression was calculated as BCL-6 expression in tumor cells [normalized to glyceraldehyde-3-phosphate dehydrogenase (GAPDH)] relative to BCL-6–GAPDH expression in Raji (Burkitt's lymphoma) cell line [as established in the derivation set (expected to express a gene at a measurable level) and cross-validated in the validation set (known to express the gene of interest at a certain level) of DLBCL patients]. BCL-6 mRNA expression in DLBCL tumors (with a cut-off level of 1.3) predicted that the DLBCL patients with high (>1.3) BCL-6 gene expression have a significantly better OS than DLBCL patients with low (≤1.3) BCL-6 gene expression (median OS of 84 and 22 months, respectively, $p = 0.01$). In multivariate analysis, together with the elements of the current prognosis-prediction gold-standard – the International Prognostic Index (IPI) – BCL-6 gene expression was an independent survival-predicting factor. By contrast, the aggregate IPI score did not add further prognostic information to the patients' stratification by BCL-6 gene expression. BCL-6 mRNA expression assayed by real-time quantitative PCR correlated well with BCL-6 protein expression, assessed by immunohistochemistry using an anti-BCL-6 antibody. This study demonstrated that quantitative expression of a single gene, BCL-6, could predict the outcome of DLBCL patients.

As the genes involved in human disease are increasingly identified and characterized, it is clear that real-time PCR will be used as a powerful tool to quantify the expression level of these genes. As studies progress in determining patterns of gene expression, it is expected that microarray technology will be widely employed. Microarrays are powerful for measuring expression patterns and expression levels of large numbers of genes simultaneously but at a relatively low level of quantitative resolution. Real-time PCR can be synergistically employed with data generated from microarrays to accurately quantify the expression of select genes in large numbers of samples, as demonstrated by Lossos and coworkers[22]. Because many diseases are manifestations of the altered expression of specific genes, the widespread application of real-time PCR will continue to be a powerful tool in assisting our understanding of these pathological processes and providing better patient care.

References

1 Alizadeh, A.A. *et al.* (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403, 503–511

2 Hayward, A. *et al.* (1999) Kinetics of competitive reverse transcriptase-PCR. In *PCR Applications: Protocols for Functional Genomics* (Innis, M. *et al.*, eds), p. 231, Academic Press

3 Higuchi, R. *et al.* (1992) Simultaneous amplification and detection of specific DNA sequences. *Biotechnology* 10, 413–417

4 Holland, P. (1991) Detection of specific polymerase chain reaction product by utilizing the 5′ to 3′ exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proc. Natl. Acad. Sci. U. S. A.* 88, 7276–7280

5 Heid, C. (1996) Real time quantitative PCR. *Genome Res.* 6, 986–994

6 Gibson, U. (1996) A novel method for real time quantitative RT-PCR. *Genome Res.* 6, 995–1001

7 Williams, P. *et al.* (1999) Real-time quantitative PCR: uses in discovery research. In *PCR Applications: Protocols for Functional Genomics* (Innis, M. *et al.*, eds) p. 365, Academic Press

8 Lee, L. (1993) Allelic discrimination by nick-translation PCR with fluorogenic probes. *Nucleic Acids Res.* 21, 3761–3766

9 Snider, J.V. (1999) Whole cell assays. In *PCR Applications: Protocols for Functional Genomics*, p. 329, Academic Press

10 Wattjes, M.P. *et al.* (2000) Comparison of nested competitive RT-PCR and real-time RT-PCR for the detection and quantification of AML1/MTG8 fusion transcripts in t(8;21) positive acute myelogenous leukemia. *Leukemia* 14, 329–335

11 Marcucci, G. *et al.* (1998) Detection of minimal residual disease in patients with AML1/ETO-associated acute myeloid leukemia using a novel quantitative reverse trancriptase polymerase chain reaction assay. *Leukemia* 12, 1482–1489

12 Kondo, M. *et al.* (2000) Real-time quantitative reverse transcription-polymerase chain reaction for the detection of AML1-MTG8 fusion transcripts in t(8;21)-positive acute myelogenous leukemia. *Leuk. Res.* 24, 951–956

13 Foroni, L. (1999) Investigation of minimal residual disease in childhood and adult lymphoblastic leukemia by molecular analysis. *Br. J. Haematol.* 105, 7–24

14 Rasmussen, T. *et al.* (2000) Quantitation of minimal residual disease in multiple myeloma using an allele-specific real-time PCR assay. *Exp. Hematol.* 28, 1039–1045

15 Ladetto, M. *et al.* (2000) Real-time polymerase chain reaction of immunoglobulin rearrangements for quantitative evaluation of minimal residual disease in multiple myeloma. *Biol. Blood Marrow Transplant.* 6, 241–253

16 Bruggemann, M. *et al.* (2000) Improved assessment of minimal residual disease in B cell malignancies using fluorogenic consensus probes for real-time quantitative PCR. *Leukemia* 2000, 14, 1419–1425

17 Pongers-Willemse, M.J. *et al.* (1998) Real-time quantitative PCR for the detection of minimal residual disease in acute lymphoblastic leukemia using junctional region specific TaqMan probes. *Leukemia* 12, 2006–2014

18 Dolken, L. *et al.* (1998) Quantitative detection of t(14;18)-positive cells by real-time quantitative PCR using fluorogenic probes. *BioTechniques* 25, 1058–1064

19 Estalilla, O.C. *et al.* (2000) 5′→3′ exonuclease-based real-time PCR assays for detecting the t(14;18)(q32;21): a survey of 162 malignant lymphomas and reactive specimens. *Mod. Pathol.* 13, 661–666

20 Shay, J. and Bacchetti, S. (1997) A survey of telomerase activity in human cancer. *Eur. J. Cancer* 33, 787–791

21 de Kok, J. *et al.* (2000) Real-time quantification of human telomerase reverse trancriptase mRNA in tumors and healthy tissues. *Clin. Chem.* 46, 313–318

22 Lossos, I. *et al.* (2001) The expression of a single gene, BCL-6, strongly predicts the survival in patients with diffuse large B-cell lymphoma. *Blood* 98, 945–951