Quantitative Determination of Estrogen Receptor, Progesterone Receptor, and HER2 mRNA in Formalin-fixed Paraffin-embedded Tissue—A New Option for Predictive Biomarker Assessment in Breast Cancer

Berit Maria Müller, MD, Ralf Kronenwett, MD,† ‡ Guido Hennig, PhD,† Heike Euting, PhD,† Karsten Weber, PhD,† Kerstin Bohmann, PhD,† Wilko Weichert, PhD,§ Gabriela Altmann,† Claudia Roth,† Klaus-Jürgen Winzer, MD,‖ Glen Kristiansen, MD,¶ Christoph Petry, PhD,† Manfred Dietel, MD,* and Carsten Denkert, MD**

Abstract: The development of optimized therapy strategies against malignant tumors is critically dependent on the assessment of tissue-based biomarkers in routine diagnostic tissue samples. We investigated a novel, fully automated, and xylene-free method for RNA isolation and biomarker determination using formalin-fixed paraffin-embedded (FFPE) tissue. The aim was to show that this approach is feasible and gives results that are comparable to the current gold standards. Expression of the breast cancer biomarkers ESR1, PGR, and HER2 was measured in a total of 501 FFPE tissue samples from 167 breast carcinomas, which had been stored for up to 21 years. Total RNA was extracted from tissue sections and biomarker expression was measured by kinetic RT-PCR (RT-kPCR). The results of the new method were compared with immunohistochemistry as the current gold standard. RNA was successfully isolated from all samples, with a mean yield of 1.4 μg/sample and fragment lengths of at least 150 bp in 99% of samples. RT-kPCR analysis of ESR1, PGR, and HER2 was possible in all samples. Comparing RT-kPCR results with standard IHC, we found a good concordance for ESR1 (agreement: 98.4%), PGR (84.4%), and HER2 (89.8%). We observed a low section-to-section variability of kPCR results for all 3 biomarkers (root of mean squared errors: 0.2 to 0.5 Ct values). The new approach is a reliable high-throughput instrument for standardized testing of biomarkers in clinical routine and for research studies on archived FFPE material up to 21 years old. For the assessment of ESR1, PGR, and HER2 the results are comparable to the current gold-standard.

Key Words: breast cancer, FFPE tissue, hormone receptor, RT-PCR, endocrine therapy

(*Diagn Mol Pathol* 2011;20:1–10)

The development of optimized therapy strategies against malignant tumors depends on a multidisciplinary approach, and the decision on the best therapeutic option is often based on biomarker assessment in formalin-fixed paraffin-embedded (FFPE) tumor tissue. For example, the state-of-the-art management of breast cancer includes the investigation of the predictive markers estrogen receptor (ESR1) and progesterone receptor (PGR), and HER2. These predictive markers are used to guide antiestrogenic therapy or anti-HER2-directed targeted therapy.[1–3]

For evaluation of predictive markers in FFPE tissue, immunohistochemistry (IHC) is currently the diagnostic gold standard.[4] IHC gives robust signals for many biomarkers, which could be correlated with morphology. However, the degree of expression of the marker can only be described in a semiquantitative way. Several studies have evaluated the interlaboratory variation in assessment of hormone receptor stainings.[5,6] Despite international efforts to standardize the cutoffs and the method,[7–9] a complete standardization of IHC for assessment of estrogen and progesterone receptors is difficult owing to the inherent semiquantitative nature of the technology. In clinical practice, substantial problems in biomarker diagnostics in breast cancer have been reported.[10,11] Therefore it is necessary to develop additional quantitative methods for biomarker assessment that are complementary to the current immunohistochemical techniques.

Owing to the current focus on targeted therapies, several additional predictive biomarkers are under development.

From the *Institute of Pathology; ‖Breast Cancer Center, Charité Hospital, Campus Mitte, Berlin; †Siemens Healthcare Diagnostics Products GmbH, Molecular Research Germany, Cologne; §Institute of Pathology, University Hospital, Heidelberg; ‡University of Düsseldorf, Universitätsstr. 1, Düsseldorf, Germany; and ¶Institute of Pathology, University Hospital, Zürich, Switzerland.

Berit Maria Müller and Ralf Kronenwett have contributed equally to this publication.

Ralf Kronenwett, Guido Hennig, Heike Euting, Karsten Weber, Kerstin Bohmann, Gabriela Altmann, Claudia Roth and Christoph Petry are employees at Siemens Healthcare Diagnostics Products GmbH, Molecular Research Germany, Nattermann Allee 12, 50829 Cologne, Germany.

Reprints: Berit Maria Müller, MD, Institute of Pathology, Charité Hospital, Campus Mitte, Charitéplatz 1, D-10117 Berlin, Germany (e-mail: berit.mueller@charite.de).

Many of the upcoming new biomarkers can be measured only on the mRNA level. For example, the RNA-based OncotypeDX 21-gene assay has recently been included in American treatment guidelines to identify patients with ESR1-positive, tamoxifen-treated breast cancer who do not need adjuvant chemotherapy.[12,13]

In addition, large collections of archival FFPE tissue samples exist in pathology institutions. In contrast to the usually limited collections of frozen tissue, these FFPE tissue samples could facilitate validation of biomarkers in adequately powered research studies.

Therefore, it is critical to implement quantitative, reliable, and standardized high-throughput methods for an investigation of mRNA expression in routine and archived FFPE tissue. However, the use of FFPE tissue for mRNA-based gene expression analyses have several limitations because formalin fixation leads to RNA fragmentation and cross-linkage with proteins and requires a manual deparaffinization step with harmful reagents.[14,15] Recently, we presented a novel standardized and fully automated method with an extraction-integrated xylene and ethanol-free deparaffinization step for fast isolation of total RNA from FFPE tissue sections.[16]

In this study, we evaluated the storage time-dependent performance and reproducibility of the new automated RNA isolation method using 501 samples from 167 FFPE breast tumors. Furthermore, we quantified the mRNA expression levels of the breast cancer biomarkers ESR1, PGR, and HER2 with kinetic reverse transcription polymerase chain reaction (RT-kPCR) and compared the results with state-of-the-art IHC.

TABLE 1. Patient Characteristics With Relevant Clinicopathology Parameters

Characteristic	No. Patients	%
All cases	167	100
Median age at time of diagnosis: 60 y (range: 30-92)		
Year of diagnosis		
2007-2008	15	9.0
2003-2006	19	11.4
1995-1998	35	21.0
1992-1994	56	33.5
1987-1991	42	25.1
Histologic type		
Ductal carcinoma	143	85.6
Lobular carcinoma	17	10.2
Other carcinoma	7	4.2
Tumor size (mm)		
≤20	48	28.7
>20	96	57.5
Missing	23	13.8
Nodal status		
Positive	91	54.8
Negative	44	26.3
Missing	32	18.9
Histologic grade		
G1	23	13.8
G2	80	47.9
G3	64	38.3
Estrogen receptor expression (IHC)		
Positive (> 10%)	125	74.9
Negative (≤10%)	42	25.1
Progesterone receptor expression (IHC)		
Positive (> 10%)	76	45.5
Negative (≤10%)	91	54.5
HER2 score (IHC)		
Negative (0)	77	46.1
Negative (+)	51	30.5
Weakly positive (++)	15	9.0
Strongly positive (+++)	24	14.4

MATERIALS AND METHODS

Sample Preparation

One hundred sixty-seven tissue samples of invasive breast cancer diagnosed at the Charité Hospital, Berlin, Germany between 1987 and 2008 were included in this study; the use of tumor samples was approved by the institutional review board of the Charité University hospital. Tissue samples were fixed in neutrally buffered formalin. The surgical specimen were placed into formalin during the surgical procedure or—in some cases—after frozen section evaluation. Therefore, the time to fixation was usually between 5 and 20 minutes. The FFPE tissues were stored at room temperature. General handling procedures and storage conditions of the tissue samples have been similar in the study period. Table 1 gives a summarized overview about patient characteristics. The percentage of tumor tissue was evaluated on hematoxylin and eosin (H&E) stained sections.

From each paraffin block, 5 consecutive 10-μm sections were cut and placed into individual Sarstedt tubes (P/N 72.692.005, Sarstedt, Nümbrecht, Germany). Sections 1, 3, and 5 were used for RNA extraction. An adjacent section was used for IHC.

Immunhistochemical Staining

Immunohistochemical staining was carried out using the DISCOVERY XT autostainer (Ventana, Tucson, AZ), using a rabbit monoclonal antibody to ESR (clone SP1, Neomarkers, 1:50), a monoclonal mouse anti-human PGR antibody (clone PgR 636, Dako, 1:50) and a rabbit anti-human HER2 HercepTest polyclonal antibody (Dako, 1:300). Furthermore, we used the DAB MAP Kit (Ventana, Tucson, AZ) with the DISCOVERY Universal Secondary Antibody (Ventana, Tucson, AZ). For all staining runs, positive and negative controls were used. Immunohistochemical staining was evaluated by 2 pathologists (C.D. and B.M.) who were blinded to the results of RT-kPCR. For assessment of hormone receptor status, the percentage of positive tumor cells was evaluated.[17] The assessment of the HER2 reactivity was carried out according to the current ASCO/CAP guidelines.[18] IHC slides with more than 10% positive staining cells were considered to be ESR1 and PGR positive, and slides with a DAKO score of 3 were considered to be HER2 positive.

RNA Extraction

The fully automated isolation method of total RNA from FFPE tissue was done using iron oxide beads coated with a nanolayer of silica in combination with a liquid handling robot as described earlier.[16] In brief, FFPE sections were lysed with heat, proteinase K, and detergent. After lysis, residual tissue debris was removed from the lysis fluid with 40-μL silica-coated iron oxide beads. Following 15 minutes incubation at 65°C beads with surface-bound tissue debris were separated on a magnet and lysates were transferred to a 2-mL deep-well plate (96 wells). During magnetization, the melted paraffin separates and forms a ring around the tube wall through hydrophobic interactions. After deparaffinization, total RNA and DNA were bound to 40 μL of a fresh volume of beads with shaking at room temperature under chaotropic conditions. Then, beads were magnetically separated and supernatants were discarded. Next, surface-bound nucleic acids were washed 3 times followed by magnetization, aspiration and disposal of supernatants, respectively. Afterward, nucleic acids were eluted by incubation of the beads with 100-μL elution buffer for 10 minutes at 70°C with shaking. Finally, beads were separated and the remaining supernatant was incubated with 12-μL DNase I Mix to remove copurified DNA. After incubation for 30 minutes at 37°C, DNA-free total RNA solution was obtained. The method took a total of 4 hours and 20 minutes to extract 48 samples including hands-on time of 30 minutes. For each run, 2 positive controls [10-μm sections of the MaxArray Breast Receptor Control Cell Block (Invitrogen, Karlsruhe, Germany)] and 1 negative control (empty tube) were included.

Assessment of RNA Concentration and Quality

Concentration of RNA was assessed using the QUANT-iT RIBOGREEN assay (Invitrogen, Karlsruhe, Germany) according to the manufacturer's instruction. For assessment of mRNA yield, RNA expression of RPL37A was measured by RT-kPCR as described below.

For determination of the mRNA fragment length of the housekeeping gene glucose-6-phosphatdehydrogenase (G6PDH) qualitative RT-PCR analyses were done with 1 sense primer and 5 different antisense primers resulting in amplicons of 67, 151, 242, 379, or 453 bp as described by Liu et al.[19]

Gene Expression Analysis Using RT-kPCR

Expression of ESR1, PGR, HER2 and of the normalization genes RPL37A and CALM2 was assessed by one-step kinetic reverse transcription PCR using the SuperScript III PLATINUM One-Step Quantitative RT-PCR System with ROX (Invitrogen, Karlsruhe, Germany) according to manufacturer's instructions in an ABI PRISM 7900HT (Applied Biosystems, Darmstadt, Germany) with 30 minutes at 50°C, 2 minutes at 95°C followed by 40 cycles of 15 seconds at 95°C and 30 seconds at 60°C. Relative expression levels of genes of interest (GOI) were calculated as ΔCt value [$\Delta Ct = 20–(Ct_{GOI}–Ct_{RPL37A\ or\ CALM2})$]. ΔCt values positively correlate with relative gene expression. The optimal cut-off points for RPL37A-normalized data were defined for classification of ESR1, PGR, and HER2 status based on the column scatter plots. Positivity for ESR1 mRNA was defined as a ΔCt of more than 14.5, negativity for ESR1 was determined as a ΔCt less than 13.5 Those few samples that were between 13.5 and 14.5 were designated as equivocal for ESR1 mRNA expression. For PGR mRNA the upper and lower cutoff were 12 and 11, respectively, the samples between both cutoffs were regarded as equivocal. For HER2 mRNA the upper and lower cutoff were 19 and 18, respectively, the samples between both cutoffs were regarded as equivocal.

DNA contamination in RNA preparations was assessed using a PAEP gene-specific kinetic PCR without reverse transcription as described above. Samples were considered to be substantially free of DNA when Ct values above 38 were detected. PCR assays were carried out in triplicates. The sequences of the primers and probes were as follows: ESR1: forward: 5′-GCCAAATT GTGTTTGATGGATTAA-3′, reverse: 5′-GACAAAA CCGAGTCACATCAGTAATAG-3′, probe: 5′-ATGC CCTTTTGCCGATGCA-3′ (amplicon length: 73 bp); PGR: forward: 5′-AGCTCATCAAGGCAATTGGT TT-3′, reverse: 5′-ACAAGATCATGCAAGTTATCAA GAAGTT-3′, probe: 5′-TTGATAGAAACGCTGTGAG CTCGA-3′ (100 bp); HER2: forward: 5′-CCAGCCTTC GACAACCTCTATT-3′, reverse: 5′-TGCCGTAGGT GTCCCTTTG-3′, probe: 5′-ACCAGGACCCACCA GAGCGGG-3′ (87 bp); RPL37A: forward: 5′-TGTGG TTCCTGCATGAAGACA-3′, reverse: 5′-GTGACAGC GGAAGTGGTATTGTAC-3′, probe: 5′-TGGCTGG CGGTGCCTGGA-3′ (65 bp); CALM2: forward: 5′-GA GCGAGCTGAGTGGTTGTG-3′, reverse: 5′-AGTCA GTTGGTCAGCCATGCT-3′, probe: 5′-TCGCGTCTC GGAAACCGGTAGC-3′ (72 bp); PAEP: forward: 5′-C ACAGAATGGACGCCATGAC-3′, reverse: 5′-AAACC AGAGAGGCCACCCTAA-3′, probe: 5′-AAGCCCTC AGCCCTGCTCTCCATC-3′ (72 bp).

Statistical Evaluation

The softwares PRISM 4 (GraphPad Software, Inc., La Jolla, CA) and MatLab (The MathWorks Inc., Version R2007b, Ismaning, Germany) were used. For comparison of multiple groups, the one-way ANOVA test was used. *P* values smaller than 0.05 were considered to be statistically significant. For correlation analyses, the Spearman correlation coefficient was calculated. Section-to-section and assay-to-assay variabilities were calculated as root of mean squared error.

Agreement between the RT-kPCR and IHC methods uncorrected for chance was calculated as the number of samples which agree divided by the total number of samples. Positive agreement was calculated as the number of samples positive in both methods divided by the number of samples positive in IHC. Negative agreement was determined accordingly. Agreement corrected for chance was determined by estimating Cohen's κ coefficient.

RESULTS

RNA Isolation From FFPE Tissue-RNA Yield and Fragment Size

The concentration of total RNA was assessed in the isolate from 1 section of each of the 167 tumors. The mean yield was 1.36 µg per section (range: 0.10 to 7.83 µg). We used the mRNA expression level of the housekeeping gene RPL37A as surrogate marker for the yield of mRNA obtained from each FFPE tissue section. Amplifiable mRNA of RPL37A could be detected in all 501 samples with Ct values below 28 in 498 (99%) of RNA preparations. In principle, RPL37A Ct values below 31 indicate sufficient RNA for quantification of biomarker expression. As assessed by DNA-specific kPCR for the PAEP gene in none of the 501 samples residual DNA was detected (data not shown).

To examine the fragment length of mRNA isolated from the FFPE tissue sections, we did RT-PCR analyses for the housekeeping gene G6PDH using primers that can amplify different fragment sizes. In all 167 samples, a 67-bp amplicon and in 166 of 167 (99%) samples, a 151-bp amplicon could be detected (Fig. 1). However, in only 3 of 167 samples (1.8%) a 242-bp fragment was successfully amplified, whereas in none of the samples a larger amplicon than 242 bp was observed. This shows that in more than 99% of samples mRNA fragments had a length of at least 150 bases.

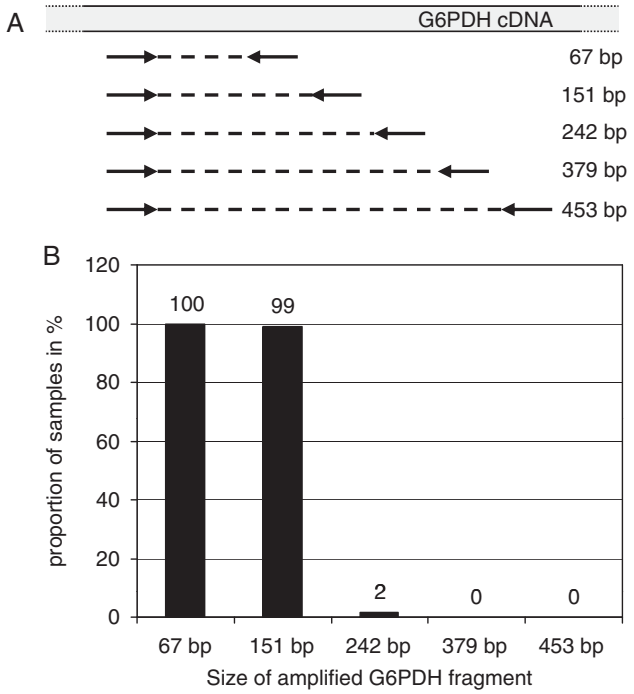

FIGURE 1. Assessment of fragment lengths of G6PDH mRNA in FFPE samples. A, Schematic depiction of RT-PCR-based assessment of 5 different fragment lengths (67, 151, 242,379, 453 nucleotides) of G6PDH mRNA using one common sense and 5 different antisense primers. B, Proportions of samples with detectable amplicons for each of the 5 different fragment sizes (n = 167). 99% of samples had a fragment size of at least 150 bp that is sufficient for most quantitative RT-PCR assays.

Isolation of RNA From Samples With Different Storage Times

The storage time of the FFPE samples used in our study ranged between 2 months and 21 years. Amplifiable mRNA of RPL37A could be detected in all samples, regardless of the different storage times, indicating an efficient isolation of mRNA even from 21 years old FFPE tissue (Fig. 2A). We found an age-dependent significant decrease of amplifiable mRNA during the first 10 years of storage of the samples. Similarly, yield of total RNA significantly decreased from a mean of 2.9-µg per section

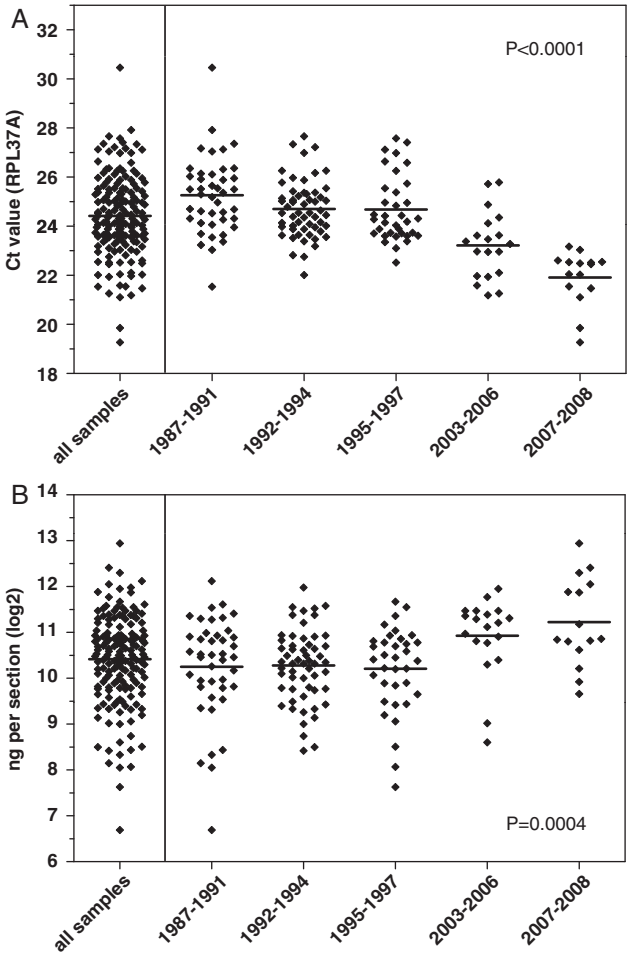

FIGURE 2. Age-dependent RNA yield in 167 FFPE tissue sections. Geometric means are indicated by bars. For comparison of multiple groups one-way ANOVA test was used to calculate the *P* values. A, Age-dependent yield of total RNA per section, assessed by the Ribogreen assay. B, Age-dependent yield of mRNA, assessed by RT-kPCR. RPL37A Ct value was used as surrogate marker for mRNA concentration. Small values indicate high concentrations and high Ct values small concentrations.

to 1.4 µg over the first 10 years of storage (Fig. 2B). No further decrease was observed in samples older than 10 years. There was no evidence for ongoing RNA fragmentation during long-term storage.

Comparison of Quantitative RT-PCR of ESR1, and PGR With IHC as the Current Gold-standard

In all 501 samples, the mRNA expression levels of ESR1, PGR, HER2, and the housekeeping genes RPL37A and CALM2 could be determined using RT-kPCR. The broad dynamic range gave a good basis for quantitative marker assessment. For ESR1 and PGR the RPL37A-normalized gene expression measured by RT-kPCR had a dynamic range of about 12 ΔCt values that corresponds to 3.5 log (base 10). Relative gene expression values of ESR1 showed a bimodal distribution whereas PGR was unimodally distributed (Figs. 3A, B).

Comparing the results from RT-kPCR with IHC a high concordance was found for both genes (Figs. 3A, B; Table 2). In the first analysis, all samples that were

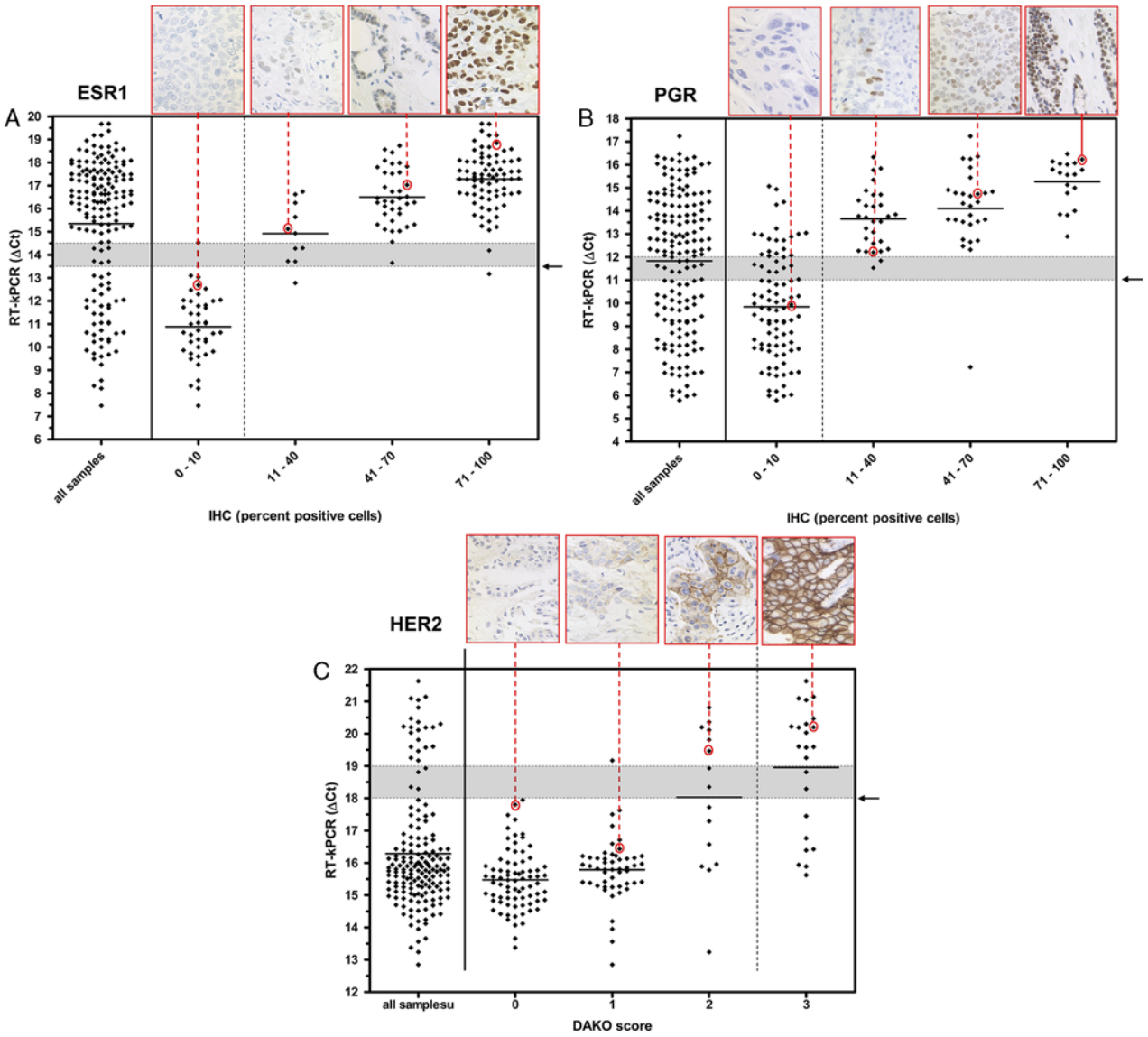

FIGURE 3. Comparison of ESR1 (A), PGR (B) and HER2 (C) protein expression analysis (determined by IHC) with mRNA expression (determined by RT-kPCR). Protein expression of 167 tumors is presented as percentage of positive tumor cells for ESR1 and PGR and as DAKO score for HER2. Cut-off levels of IHC positivity/negativity are indicated as vertical dotted lines. Ranges of mRNA expression values regarded as equivocal are marked by grey areas and horizontal dotted lines. The cut-off levels of RT-kPCR positivity/negativity including the equivocal mRNA values in the positive group are indicated by arrows. Corresponding immunohistochemical stainings of selected cases indicated by red circles are shown for each marker.

TABLE 2. Concordance of RT-kPCR and Central IHC (n = 167) as Well as Concordance of RT-kPCR Between Three Tissue Sections of a Tumor (n = 167) for ESR1, PGR and HER2 Status

	ESR1		PGR		HER2	
RT-kPCR	**Without eq**	**With eq**	**Without eq**	**With eq**	**Without eq**	**With eq**
Range equivocal Ct values/cutoffs	13.5–14.5	13.5	11–12	11	18–19	18
Proportion of equivocal samples	3.7%		6.7%		2.4%	
Concordance IHC and RT-kPCR						
Kappa	0.95	0.95	0.75	0.67	0.60	0.59
CI lower I	0.90	0.90	0.65	0.56	0.40	0.42
CI higher I	1.00	1.00	0.86	0.77	0.79	0.77
Overall Agreement	0.98	0.98	0.88	0.83	0.91	0.90
Positive agreement	0.98	0.98	0.99	0.99	0.64	0.67
Negative agreement	0.98	0.98	0.78	0.70	0.95	0.94
Agreement of RT-kPCR between three sections						
Overall agreement	1.00	0.99	1.00	0.98	1.00	0.99

Results are presented without samples that were regarded as equivocal (eq) on the mRNA level (left columns, range of Ct values designated as equivocal is given for each biomarker) and with the equivocal cases in the positive group (right columns, cutoff Ct value is indicated).

equivocal on the mRNA level (3.7% of samples for ESR1 assessment, 6.7% of samples for PGR) were not included in the comparison. In this analysis, an overall agreement of 98% between RT-PCR and IHC was found for ESR1. Positive and negative agreements were also 98%. For PGR, overall agreement of both methods was 88%, positive agreement was 99%. However, several samples with 10% or less positive staining cells were judged as positive in RT-kPCR resulting in a moderate negative agreement of 78%.

As second, more stringent analysis, all samples were evaluated and the samples that were equivocal on the mRNA level were included in the positive group (Table 2). Even in this analysis, almost perfect overall, positive and negative agreements of 98% were found for ESR1. For PGR, overall agreement of both methods was 83%, positive agreement was 99%. However, several samples with 10% or less positive staining cells were judged as positive in RT-kPCR resulting in a moderate negative agreement of 70%.

Comparison of Quantitative RT-PCR of Her2 With Immunohistochemistry

HER2 expression varied over a range of about 9 ΔCt values with a bimodal distribution of the relative gene expression values (Fig. 3C). In the first analysis, all samples that were equivocal on the HER2 mRNA level were excluded (2.4% of samples). In this first analysis, we observed an overall agreement of 91% between both methods comparing DAKO score 0 to 2+ with score 3+ (Fig. 3C, Table 2). Negative agreement was excellent (95%), whereas positive agreement was only 64%. In particular, the concordance was limited in samples with weak (score 2) or strong expression (score 3).

We conducted a second, more stringent analysis in which all cases were evaluated and the equivocal cases were included in the positive group (Table 2). Here we observed an overall agreement of 90% between both methods comparing DAKO score 0 to 2+ with score 3+ Negative agreement was excellent (94%), whereas positive agreement was only 67%. In particular, the concordance was limited in samples with weak (score 2) or strong expression (score 3).

Section-to-section and Assay-to-assay Variability of Biomarker Analysis

To assess the variability of the RT-kPCR results in different RNA eluates from the same tumor, we compared the results of 3 different tissue sections of each tumor. We found a highly significant correlation of the RPL37A Ct values between 2 sections of the 167 tumors (Spearman $R = 0.93$, 95% CI 0.90-0.95, $P < 0.0001$). The standard deviation of the RPL37A Ct values in 3 consecutive sections was in mean 0.36. For RPL37A-normalized levels of ESR1, PGR, and HER2, excellent correlations of gene expression between 2 tissue sections were observed (Figs. 4A–C). Variations above one ΔCt value were only observed in the low copy expression range in which RT-kPCR precision is the limiting factor. Comparing 3 sections, the root of mean squared errors of ESR1, PGR, and HER2 ΔCt values in 3 sections of a tumor were 0.23, 0.54, and 0.34, respectively (Fig. 4D). An age dependence of section-to-section variability was not observed. Finally, we examined the concordance of the RT-kPCR-based classification of the ESR1, PGR, and HER2 status as positive or negative in RNA isolated from 3 different tissue sections. The section-to-section agreement of RT-kPCR receptor status was excellent. For ESR1 in 99%, for PGR in 98%, and for HER2 in 99% of tumors an agreement of the classifications was found in all 3 sections including equivocal samples and even 100% without inclusion of samples that were regarded equivocal on the mRNA level (Table 2).

Moreover, assay-to-assay variability was assessed by a second, independent measurement of the normalized expression values of ESR1, PGR, and HER2. Assay-to-assay variability (PCR variability) of the expression of

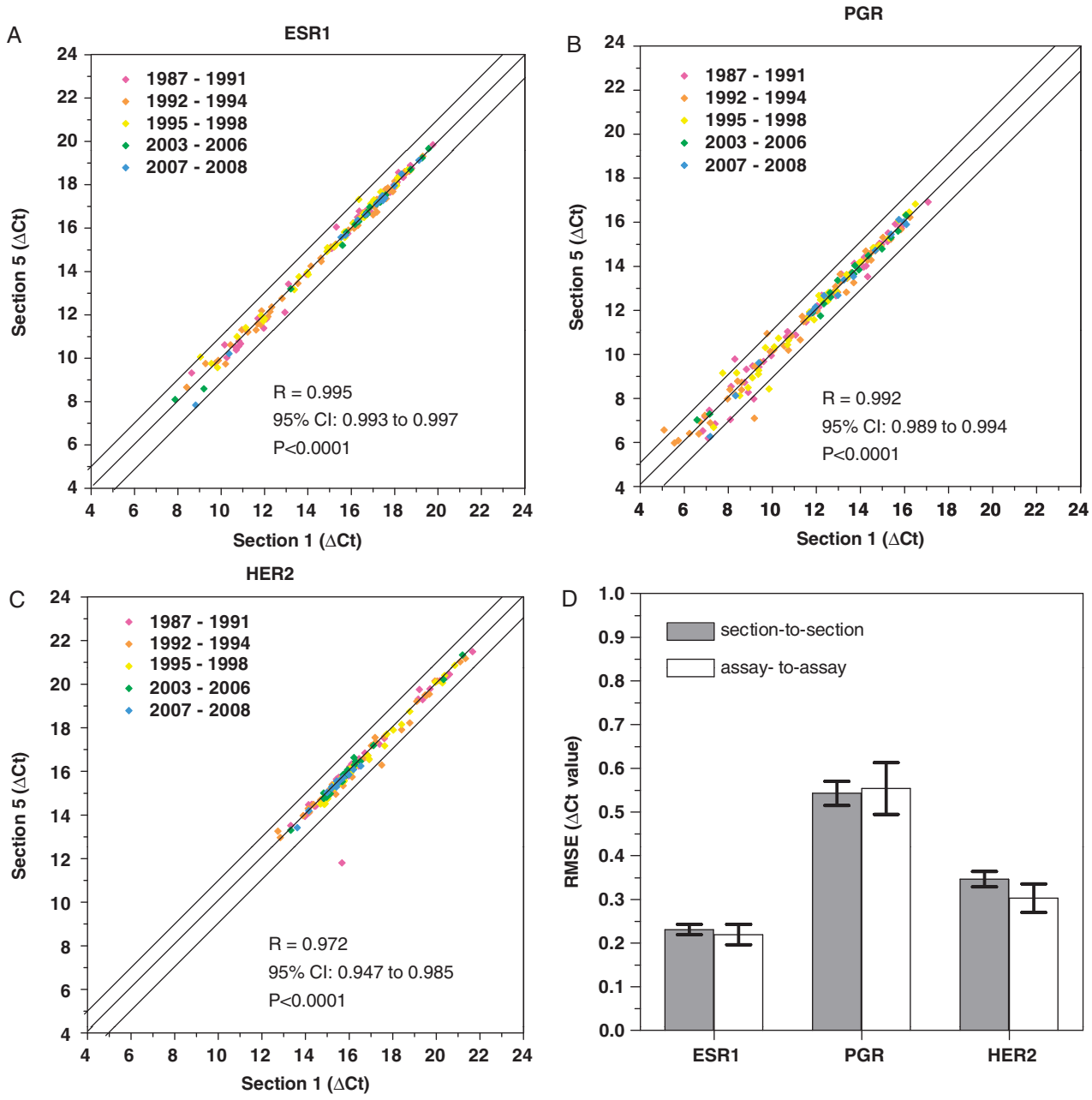

FIGURE 4. Section-to-section and assay-to-assay variability of ESR1, PGR, and HER2 mRNA expression. A to C, Comparison of RPL37A-normalized expression values of ESR1, PGR, and HER2 in 2 tissue sections of 167 breast cancer tumors. Different age groups are color-coded. The bisectrix and ±1 ΔCt are indicated by solid lines. Spearman coefficients, 95% confidence intervals and significance levels of the correlations are given. D, Section-to-section variability of normalized expression of ESR1, PGR, and HER2 in 3 tissue sections from 167 tumors, respectively, and assay-to-assay variability of the 3 normalized genes in RNA preparations from 167 tumor. Variability is indicated as root of mean squared errors. Bars represent 95% confidence intervals.

these genes was low and comparable to the section-to-section variability (Fig. 4D).

The results after normalization with the housekeeping gene CALM2 were similar to the results with RPL37A normalization (data not shown).

DISCUSSION

In this study, we evaluated the feasibility of quantitative biomarker assessment in FFPE tissue using a new fully automated system for RNA isolation linked to a kinetic RT-PCR. Using a cohort of 501 samples, we

compared the results with the current gold standards. The main result of our study was that the determination of biomarkers was possible in all 501 samples and showed a good to excellent correlation with the current diagnostic standard approaches. Samples that were stored for up to 21 years were suitable for the analysis, some older samples showed a slightly lower output of RNA, which could be easily compensated by normalization procedures. Our study did not show a correlation between the age of the samples and the amplifiable fragment sizes determined with RT-PCR of the G6PDH mRNA. Therefore, we interpret the reduced yield in older samples with reduced release of nucleic acids during lysis from surrounding cross-linked cellular components and proteins, but not with increased degradation of mRNA.[14,15] In general, there have been no significant changes in handling procedures and storage conditions in our institute within the last 21 years. So in our opinion, this cannot be the reason for the reduced yield in older samples. Results might be different owing to different fixation and storage conditions in other institutions, which might require additional testing.

Overall, the method allowed successful isolation of total RNA with a mean yield of 1.36-μg per section. Assuming about 2 ng input RNA per one-step RT-kPCR reaction each sample would allow the quantitative analysis of at least hundred biomarker genes. We observed a low biologic section-to-section variability which did not exceed the PCR variability itself, suggesting that the variability is contributed by the PCR rather than by the RNA isolation technique or the heterogeneity of the tumor.

Different manual protocols for RNA extraction from FFPE tissue have been developed.[20–23] Most of the protocols use similar principles such as deparaffinization, cell disruption, and release of RNA from cross-linked proteins by proteinase K. RNA is subsequently extracted using silica columns or beads and chaotropic salts. Especially the deparaffinization step employs harmful reagents and needs ethanol/xylene treatment steps in combination with centrifugations impeding a complete automation. Furthermore, the complexity of FFPE samples from different tissue types compromises the standardization of RNA extraction protocols. Thus, the use of the existing methods in a routine laboratory for high-throughput analyses is comparably difficult and labor-intensive.

Compared with other methods for RNA extraction, one major new feature of the presented method is the extraction-integrated deparaffinization step by means of an efficient hydrophobic absorption of paraffin to the inner polypropylene tube wall of the tube. This replaces the harmful and labor-intensive ethanol/xylene steps and is one precondition for complete automation. Another important and innovative feature of the method is the introduction of a bead-based negative selection step to bind and eliminate any undigested tissue (debris) in the absence of chaotropic salts before RNA extraction. The separation of tissue debris is another key for complete automation; otherwise such material would interfere with accurate liquid handling resulting in clogging of pipette tips. In addition, this negative selection step is the prerequisite to define a standardized lysis time (here 1 h) and hence makes the method applicable to the complexity of FFPE samples with variable tissue input and cell/matrix composition. The complete automation results in a processing time of 4 hours and 20 minutes for 48 samples including hands-on time of 30 minutes, which supports the use for high-throughput analyses in a routine laboratory.

We evaluated the feasibility of quantitative RT-kPCR for assessment of expression of the 3 breast cancer-related biomarkers ESR1, PGR, and HER2 using the RNA isolated with our new method. The results from a series of 3 sections of each tumor showed that a highly reproducible and robust expression analysis of genes is possible over a large dynamic range of 3 to 4 logs in tissue samples from the same tumor. This reproducibility is a prerequisite for the use of the new method in clinical routine.

Comparing quantitative RT-kPCR with the standard method IHC for assessment of ESR1, PGR, and HER2 expression levels, we found an excellent concordance between both methods for ESR1 status with an agreement of 98%. Agreements of PGR and HER2 expression were slightly lower with values of 83% and 90%, respectively. This is in line with the results of other groups assessing receptor status in FFPE tissue using RT-kPCR.[24–28] Best agreement between PCR and IHC was always found for ESR1 (agreement 91% to 94%) whereas PGR and HER2 showed somewhat lower concordance in all published studies.

HER2 overexpression is currently being tested by a combination of IHC as a standard method and FISH or SISH to detect HER2 gene amplification.[29–31] Looking at concordance of HER2 status between IHC and RT-kPCR, we found an excellent negative agreement of 94%, whereas positive agreement was only 67%. In particular, the concordance was limited in samples with weak positive (score 2) or strong positive expression (score 3). A similar observation was made by Gong et al[32] in fresh-frozen tissue using microarrays and by other groups in FFPE tissue using quantitative RT-PCR for HER2 testing in breast cancer specimens.[26,28] Cases with Score 0 to 1 are similarly low in mRNA expression, whereas the group with Score 2+ to 3+ has similarly high mRNA levels. This suggests that further studies could define a clinically relevant cut-off for HER2 overexpression based on RT-kPCR. Also, concordance between genomic HER2 amplification and HER2 mRNA expression needs to be clarified in further studies.

For classification of ESR1, PGR, and HER2 status in RT-kPCR we used optimized cut-off values based on the bimodal distribution of expression values in column scatter plots. Furthermore, for each marker we defined an intermediate so called equivocal zone between the clearly positive and negative range. Similar to the situation for HER2, the equivocal zone might require the

repeated assessment of biomarkers using additional diagnostic tests by other methods. The cut-off values used in this investigation are a basis for retrospective and prospective validations in clinical studies. As a first prospective validation, the PREDICT substudy of the neoadjuvant GeparQuinto study of the German Breast Group is currently conducted, to validate biomarkers for response to neoadjuvant chemotherapy in breast cancer.

In conclusion, our feasibility study shows that the novel method enables a reliable and high-throughput purification of RNA from the valuable archives of stored tumor tissues. The achieved agreement between IHC and RT-PCR, and the large dynamic range and reproducibility of the new method opens a new route in the field of personalized medicine. As mRNA-based receptor determination can predict tamoxifen response and clinical outcome, our new method may improve the predictive value of ESR1 status.[33,34] As many upcoming biomarkers from gene expression studies are measured on the mRNA level, this method is a major technical improvement for implementation of reproducible and cost-efficient testing of such biomarkers in clinical routine and in research studies using archived FFPE material in molecular pathology diagnostic testing.

ACKNOWLEDGMENTS

The authors thank Britta Beyer, Petra Wachs, and Torsten Acht for their excellent technical assistance and Britta Dahl and Martina Eickmann for proofreading.

REFERENCES

1. Payne SJL, Bowen RL, Jones JL, et al. Predictive markers in breast cancer—the present. *Histopathology*. 2008;52:82–90.
2. Turner NC, Jones AL. Management of breast cancer—part II. *BMJ*. 2008;337:107–110.
3. Geyer CE, Forster J, Lindquist D, et al. Lapatinib plus capecitabine for HER-2 positive advanced breast cancer. *N Engl J Med*. 2006; 355:2733–2743.
4. Harvey JM, Clark GM, Osborne CK, et al. Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. *J Clin Oncol*. 1999;17:1474–1481.
5. Rhodes A, Jasani B, Barnes DM, et al. Reliability of immunohistochemical demonstration of oestrogen receptors in routine practice: interlaboratory variance in the sensitivity of detection and evaluation of scoring systems. *J Clin Pathol*. 2000;53:125–130.
6. Rüdiger T, Höfler H, Kreipe HH, et al. Quality assurance in Immunohistochemistry: results of an interlaboratory trial involving 172 pathologists. *Am J Surg Pathol*. 2002;26:873–882.
7. Price KN, Goldhirsch A. Clinical trial update: International Breast Cancer Study Group. *Breast Cancer Res*. 2005;7:252–254.
8. Goldhirsch A, Ingle JN, Gelber RD, et al. Thresholds for therapies: highlights of the St Gallen International Expert Consensus on the primary therapy of early breast cancer 2009. *Ann Oncol*. 2009;20: 1319–1329.
9. Kurosumi M. Immunohistochemical assessment of hormone receptor status using a new scoring system (J-Score) in breast cancer. *Breast Cancer*. 2007;14:189–193.
10. Layfield LJ, Gupta D, Mooney EE. Assessment of tissue estrogen and progesteronereceptor levels: a survey of current practice, techniques, and quantitation methods. *Breast J*. 2000;6:189–196.
11. Allred DC. Hormone receptor testing in breast cancer: a distress signal from Canada. *Oncologist*. 2008;13:1134–1136.
12. Paik S, Shak S, Tang G, et al. A Multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med*. 2004;351:817–2826.
13. Harris L, Fritsche H, Mennel R, et al. American Society of Clinical Oncology. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *J Clin Oncol*. 2007;25:5287–5312.
14. Masuda N, Ohnishi T, Kawamoto S, et al. Analysis of chemical modification of RNA from formalin-fixed samples and optimization of molecular biology applications for such samples. *Nucleic Acid Res*. 1999;27:4436–4443.
15. Páska C, Bögi K, Szilák L, et al. Effect of formalin, acetona, and RNAlater fixatives on tissue preservation and different size amplicons by real-time PCR from paraffin-embedded tissue. *Diagn Mol Pathol*. 2004;13:234–240.
16. Bohmann K, Hennig G, Rogel U, et al. RNA extraction from archival formalin-fixed paraffin-embedded tissue: a comparison of manual, semiautomated, and fully automated purification methods. *Clin Chem*. 2009;55:1719–1727.
17. Kreienberg R. *Interdisziplinäre S3- Leitlinie für die Diagnostik, Therapie und Nachsorge des Mammakarzinoms*. Munich, Germany: Zuckerschwerdt GmbH; 2008.
18. Wolff AC, Hammond ME, Schwartz JN, et al. American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol*. 2007;25:118–145.
19. Liu H, Huang X, Zhang Y, et al. Archival fixed histologic and cytologic specimens including stained and unstained materials are amenable to RT-PCR. *Diagn Mol Pathol*. 2002;11: 222–227.
20. Abramovitz M, Ordanic-Kodani M, Wang Y, et al. Optimization of RNA extraction from FFPE tissue for expression profiling in the DASL assay. *Biotechniques*. 2008;44:417–423.
21. Ribeiro-Silva A, Zhang H, Jeffrey SS. RNA extraction from ten year old formalin-fixed paraffin embedded breast cancer samples: a comparison of column purification and magnetic bead-based technologies. *BMC Molecular Biology*. 2007;8:118.
22. Stanta G, Schneider C. RNA extracted from paraffin-embedded human tissues is amenable to analysis by PCR amplification. *Biotechniques*. 1991;11:304, 306, 308.
23. Rupp GM, Locker J. Purification and analysis of RNA from paraffin-embedded tissue. *Biotechniques*. 1988;6:56–60.
24. Ma XJ, Hilsenbeck SG, Wang W, et al. The HOXB13: IL17BR expression index is a prognostic factor in early-stage breast cancer. *J Clin Oncol*. 2006;24:4611–4619.
25. Cronin M, Pho M, Dutta D, et al. Measurement of gene expression in archival paraffin-embedded tissues: development and performance of a 92-gene reverse transcriptase-polymerase chain reaction assay. *Am J Pathol*. 2004;164:35–42.
26. Esteva FJ, Sahin AA, Cristofanilli M, et al. Prognostic role of a multigene reverse transcriptase-PCR assay in patients with node-negative breast cancer not receiving adjuvant systemic therapy. *Clin Cancer Res*. 2005;11:3315–3319.
27. Badve SS, Baehner FL, Gray RP, et al. Estrogen-and progesterone-receptor status in ECOG 2197: comparison of immunohistochemistry by local and central laboratories and quantitative reverse transcription polymerase chain reaction by central laboratory. *J Clin Oncol*. 2008;26:2473–2481.
28. Iverson AA, Gillett C, Cane P, et al. A single-tube quantitative assay for mRNA levels of hormonal and growth factor receptors in breast cancer specimens. *J Mol Diagn*. 2009;11:117–130.
29. Di Palma S, Collins N, Bilous M, et al. A quality assurance exercise to evaluate the accuracy and reproducibility of chromogenic in situ hybridisation for HER2 analysis in breast cancer. *J Clin Pathol*. 2008;61:757–760.
30. Yaziji H, Goldstein LC, Barry TS, et al. Her-2 testing in breast cancer using parallel tissue-based methods. *JAMA*. 2004;291: 1972–1977.

31. Dietel M, Ellis IO, Höfler H, et al. Comparison of automated silver enhanced in situ hybridisation (SISH) and fluorescence ISH (FISH) for the validation of HER2 gene status in breast carcinoma according to the guidelines of the American Society of Clinical Oncology and the College of American Pathologists. *Virchows Arch*. 2007;451:19–25.
32. Gong Y, Yan K, Lin F, et al. Determination of oestrogen-receptor status and ERBB2 status of breast carcinoma: a gene-expression profiling study. *Lancet Oncol*. 2007;8:203–211.
33. Lacroix M, Querton G, Hennebert P, et al. Estrogen receptor analysis in primary breast tumors by ligand-binding assay, immunocytochemical assay, and northern blot: a comparison. *Breast Cancer Res Treat*. 2001;67:263–271.
34. Baehner FL, Habel LA, Quesenberry CP, et al. Quantitative RT-PCR analysis of ER and PR by Oncotype DX indicates distinct and different associations with prognosis and prediction of tamoxifen benefit. *Presented at the 29th Annual San Antonio Breast Cancer Symposium, 2006,* Abstract #45 (2006).