J Mol Med (2009) 87:211–224
DOI 10.1007/s00109-008-0419-y

ORIGINAL ARTICLE

A novel approach for reliable microarray analysis of microdissected tumor cells from formalin-fixed and paraffin-embedded colorectal cancer resection specimens

Silke Lassmann · Clemens Kreutz · Anja Schoepflin · Ulrich Hopt · Jens Timmer · Martin Werner

Received: 3 June 2008 / Revised: 4 November 2008 / Accepted: 6 November 2008 / Published online: 6 December 2008

Abstract We present a novel approach for microarray analysis of RNA derived from microdissected cells of routinely formalin-fixed and paraffin-embedded (FFPE) cancer resection specimens. Subsequent to RNA sample preparation and hybridization to standard GeneChips (Affymetrix), RNA samples yielded 36.43±9.60% (FFPE), 49.90±4.43% (fresh-frozen), and 53.9% (cell line) present calls. Quality control parameters and Q-RT-PCR validation demonstrated reliability of results. Microarray datasets of FFPE samples were informative and comparable to those of fresh-frozen samples. A systematic measurement difference of differentially processed tissues was eliminated by a correction step for comparative unsupervised data analysis of fresh-frozen and FFPE samples. Within FFPE samples, unsupervised clustering analyses clearly distinguished between normal and malignant tissues as well as to further separate tumor samples according to histological World Health Organization (WHO) subtypes. In summary, our approach represents a major step towards integration of microarrays into retrospective studies and enables further investigation of the relevance of microarray analysis for clinico-pathological diagnostics.

Keywords mRNA expression profiling · Fixed tissue specimens · Cancer

Electronic supplementary material The online version of this article (doi:10.1007/s00109-008-0419-y) contains supplementary material, which is available to authorized users.

S. Lassmann (✉) · A. Schoepflin · M. Werner
Institute of Pathology, University Hospital,
Albert-Ludwigs University,
Breisacherstr. 115a,
79106 Freiburg, Germany
e-mail: silke.lassmann@uniklinik-freiburg.de

C. Kreutz · J. Timmer
Center for Biosystems Analysis (ZBSA),
Freiburg Center for Data Analysis and Modeling (FDM),
Albert-Ludwigs University,
Freiburg, Germany

U. Hopt
Department of Surgery, University Hospital,
Albert-Ludwigs University,
Freiburg, Germany

Introduction

Gene expression analysis by genome-wide mRNA microarray analysis is a widely applied technique in cancer research and is increasingly discussed as a supplementary tool for molecular diagnostics, especially for individualized cancer patient care [1–2]. With technical diversity and associated methodological advances, standardization of microarray analysis has become a pre-requisite for integrating this technique into a clinical setting [2–5].

One of the major factors influencing the performance and accuracy of microarray analysis is the source and processing of patient samples. So far, reliable microarray analysis is limited to fresh blood or fresh-frozen tissue samples. However, these samples may be unavailable from subsets of patients. The latter applies particularly to small, endoscopic or pretherapeutic tissue biopsies, which need to be entirely processed by formalin fixation and paraffin embedding (FFPE) for routine diagnostic histopathology. However, these FFPE tissue biopsies and/or cancer resection specimens, on which the primary histopathological diagnosis was made, represent the most precious tissue samples for molecular microarray analysis. In fact, the use

of such FFPE samples in large-scale, microarray-based retrospective studies from archival tissue specimens of patients with long-term clinical follow up will strengthen and expand the clinical relevance of gene expression profiles obtained in cell lines and/or fresh-frozen tissues.

The need to improve molecular analyses, especially microarray analysis from routinely FFPE-processed and histologically classified tissue specimens, is well recognized [6 and references therein] and has been fuelled by two main concepts: First, alter the entire workflow of diagnostic tissue collection and processing, for example, finding an optimal fixative for both histopathological and molecular analyses [7]. Second, modify "downstream" molecular techniques for analysis of FFPE samples, as successfully shown for quantitative RT-PCR [8, 9].

The main reasons why RNA extracts of routinely processed FFPE tissues are of poor quality for standard molecular analyses are (1) RNA fragmentation and (2) cross-linking to each other, to DNA and to proteins [6, 10, 11]. This problem is solved by, for example, choosing small fragments for detection by PCR-based methods [6, 8, 9]. RNA fragmentation and cross-linking, however, still prevent reliable microarray-based analysis, at least for those microarray platforms designed for (oligo-dT primed) detection of intact RNA sequences. Efforts to improve RNA extraction protocols and/or further RNA processing for microarray hybridization were partially successful [12–15]. Major improvements were however achieved by using a new microarray technique, which involves cDNA-mediated annealing, selection, extension, and ligation (DASL) as well as random priming for detection of degraded RNA [16–18].

Moreover, promising proof-of-concept evidence on the use of FFPE tissues for analysis of mRNA expression for conventional oligonucleotide microarrays has recently been provided [19–27], but the consistency of data with respect to routinely processed diagnostic FFPE tissues is still sparse. Above all, the use of specifically FFPE processed cell lines [25] and cell pellets [20] or tissues with optimal FFPE processing conditions [19, 21] may not reflect the FFPE samples of daily routine histopathological diagnostics. Recently, investigation of routinely processed fresh-frozen and FFPE tissues of diverse cancers by microarray was indeed successful, with the resulting mRNA expression profiles discriminating between cancer entities [22, 23]. Using different biological entities as tissue sources may however not have fully revealed the impact of tissue processing on mRNA expression profiles.

Together, valuable insights into successful handling of FFPE samples for microarray analyses have been obtained: In the case of FFPE samples specifically prepared [19–21, 25] and using "custom-made" microarrays [21, 22, 24, 25], microarray analyses were very successful in terms of the percent of genes detected (up to 54%). In slight contrast, investigation of FFPE tissue samples on a widely accepted research and discussed diagnostic platform (Affymetrix) yielded a less robust read-out, with present calls up to 30% [20, 23, 26]. Because the efficiency of probe set detection of 3′–5′ gene sequences is low in FFPE samples due to RNA fragmentation and the oligo-dT primed cDNA synthesis step, a recent study [27] proposed cut-off based analysis of probe sets located 3′ prime at the first 300–600 bp of the gene sequence only.

Here, we present a novel approach that allows efficient and reliable microarray analysis on Affymetrix GeneChips (HG-U133_Plus_2) from 100 ng total RNA extracted from microdissected (tumor) cells of archival, large, routine-processed FFPE colorectal cancer resection specimens. This novel approach (1) enables microarray analyses from the available diagnostic tissue specimen of an individual case, for example only a single small FFPE processed biopsy, (2) provides a tool for comparative microarray analyses of differently processed tissue samples, and (3) enables investigation of gene expression profiles in FFPE tissue samples in a biological context. Our study therefore paves the way for application of microarray analysis in large groups of archival tissue samples for retrospective studies and provides a basis of further investigation of the relevance of microarray analysis in a clinico-pathological diagnostic setting.

Materials and methods[1]

Tissue specimens

The study included fresh-frozen and formalin-fixed, paraffin-embedded (FFPE) resection specimens obtained from patients undergoing surgery for colorectal cancer at the Department of Surgery, University Hospital Freiburg, Germany (approval by local ethics committee; #251/04, Ethik-Kommission, Albert-Ludwigs Universität, Freiburg, Germany). Tissue specimens were subjected to routine pathological analysis, including "frozen section" diagnostics within 20 min from surgery and detailed macrocopic and histopathological diagnostics [28, 29] post fixation over night in neutral buffered formalin and paraffin embedding. Clinico-pathological parameters of the cases are given in Table 1. All patients had a curative (R0) resection. One patient (1/9, case 2) was suspected to have hereditary nonpolyposis colorectal cancer (HNPCC), as supported by routine testing for microsatellite instability. Other cases were sporadic colorectal cancers, with 4/9 cases (cases #6–9) also displaying microsatellite instability [30, 31]. The histotype of

[1] All case-specific histopathological and RNA data, the microarray datasets, and the results of the various clustering analyses not explicitly shown are available from the authors upon request.

Table 1 Summary of clinico-pathological parameters of the investigated cases with corresponding experimental data of RNA quality controls and microarray analysis

Case Id. # (year)	Sex	Age	Site	T	N	G	WHO	Sample type	% Tu.	RNA quantity/quality				Q-RT-PCR analysis (3′/5′ β-actin)			GeneChip HG-U133_ Plus_2 (GCOS, report file)						
										ng/μl	260/280	260/230	RIN	3′ Ct	5′ Ct	3′/5′	Chip Id. #	SF*	Bg	Raw Q	3′/5′ GAPDH	3′/5′ β-actin	% PC
Commercial reference RNA (32)										1,000	n.d.	n.d.	7.6	17.68	18.17	0.97	n.d.						
Cell line HCT116										1,183	2.07	2.22	9.2	18.04	17.61	1.02	1	2.1	35.16	1.25	1.28	29.64	53.9
1 (2007)	m	83	a	3	0	2	t	Fresh	>90	55.2	2.02	2.02	9.0	17.88	17.57	1.02	2	4.6	34.15	1.34	1.82	49.36	53.3
								FFPE	>90	100.5	2.02	2.07	2.3	20.24	22.66	0.89	3	12.9	34.28	1.27	1.82	10.08	47.5
			Normal epithelium					FFPE	0	126.7	2.05	2.14	2.3	21.45	23.56	0.91	17	24.6	33.59	1.23	2.11	5.85	44.2
2 (2008)	m	28	a	3	0	2	t	Fresh	>90	45.5	2.01	2.04	7.9	17.17	17.84	0.96	5	7.9	34.43	1.29	1.04	6.94	49.8
								FFPE	>90	186.4	2.01	2.05	2.3	19.73	22.98	0.86	6	20.3	32.80	1.22	0.69	5.49	42.1
3 (2007)	m	78	a	3	0	2	t	Fresh	>90	52.5	2.01	2.01	7.0	18.58	19.19	0.97	11	16.2	31.33	1.20	1.02	26.97	44.3
								FFPE	>90	193.1	2.00	2.00	2.1	22.83	26.33	0.87	4	32.0	34.16	1.23	0.96	8.09	33.3
4 (2008)	m	69	s	3	1	2	t	Fresh	>80	79.1	2.05	1.77	7.6	17.30	18.37	0.94	9	10.8	32.05	1.18	1.21	9.11	47.0
								FFPE	>80	247.0	2.01	2.00	2.3	19.52	24.09	0.81	10	21.1	34.06	1.20	0.50	5.42	39.6
			Normal epithelium					FFPE	0	215.9	2.04	2.18	2.2	20.72	23.08	0.90	16	24.3	35.18	1.32	1.93	5.06	42.9
5 (2007)	m	66	s	3	0	2	t	Fresh	>80	71.9	2.04	2.20	7.3	18.08	18.35	0.99	7	6.1	33.34	1.18	0.91	21.21	55.1
								FFPE	>80	288.7	1.99	1.90	2.2	18.31	22.02	0.83	8	18.5	36.71	1.35	0.91	9.21	42.0
			Normal epithelium					FFPE	0	228.6	2.07	2.18	2.1	24.13	27.11	0.89	18	159.4	32.79	1.17	0.79	10.19	12.7
6 (2003)	m	91	a	3	0	3	m	Fresh	>90	67.6	2.03	2.13	2.6	21.50	23.96	0.90	n.d.						
								FFPE	>90	1,108.2	2.08	2.07	2.6	19.54	24.02	0.81	12	26.8	39.12	1.49	1.20	7.15	36.1
7 (2001)	f	86	a	2	0	2	m	FFPE	>80	369.9	1.65	1.21	1.9	20.55	23.22	0.89	13	35.4	30.09	1.15	1.95	6.56	39.0
8 (2001)	m	67	a	3	0	3	m	FFPE	90	286.4	2.00	1.86	2.1	21.47	24.41	0.88	14	47.1	28.47	1.13	2.18	7.21	32.2
9 (2002)	f	90	a	3	0	2	m	FFPE	>80	169.1	2.00	2.10	2.2	22.65	25.68	0.88	15	58.4	32.11	1.21	2.28	17.58	25.6
Fresh-frozen mean (±std. dev.)										61.97	2.03	2.03	6.90	18.4	19.21	0.96	2,5,11,9,7	9.1	33.06	1.24	1.20	22.72	49.90
										(12.89)	(0.02)	(0.15)	(2.22)	(1.6)	(2.39)	(0.04)		(4.6)	(1.34)	(0.07)	(0.36)	(17.06)	(4.43)
FFPE mean (±std. dev.)										353.98	2.02	2.02	2.30	20.03	23.68	0.85	3, 6, 4, 10,	40.1	33.61	1.25	1.44	8.16	36.43
										(374.91)	(0.03)	(0.06)	(0.17)	(1.51)	(1.52)	(0.03)	8, 12–18	(39.7)	(2.79)	(0.15)	(0.66)	(3.46)	(9.60)

Refer to main text for discussion of parameters. Abbreviations are as follows: Site: *a* colon ascendens, *s* colon sigma, *T* pT category, *N* pN category, *G* grading [28], *WHO* histological subtype with *t* tubular and *m* mucinous [29], *% Tu* percent of tumor cells after microdissection, *Ct* cycle threshold, *SF** scaling factor. Note that this is mainly influenced by target signal values selected by the user. A major reduction of SFs is achieved by reducing the target signal value from the standard of 500 (shown) to, e.g., 200, which in our series resulted in more than 2-fold lower SFs (not shown). *Bg* background, *Raw Q* noise, *PC* present calls. Note: all tumors were curatively resected (R0)

tumors according to the World Health Organization (WHO) [29] was tubular in 5/9 cases (cases #1–5) and mucinous in 4/9 cases (cases #6–9).

For control purposes, a colorectal cancer cell line (HCT116, microsatellite instable; ATCC) as well as a commercially available universal reference RNA (Stratagene) [32] was included. Note that samples were processed at different time points (set 1=arrays 1–3; set 2=arrays 4–6; set 3=arrays 7–10; set 4=arrays 11, 12; set 5=arrays 13–18).

Microdissection and RNA extraction

First, relevant tissue areas, containing normal colorectal epithelial cells (resection margins at least 10 cm away from the tumor), invasive mucinous or tubular colorectal tumor cells were determined on hematoxylin and eosin (HE) stained sections (Fig. 1). Manual microdissection was then performed on serial sections stained for 15 s in instant hematoxylin (Shandon) to enrich normal epithelial cells or invasive tumor cells using fine needles (Fig. 1) [30, 31, 33, 34]. The cases displayed different tumor content and after microdissection tumor cell purity was >80% for all samples. All steps were performed under RNase-free conditions.

For fresh-frozen tissues, ten 10-µm serial sections were cut from each specimen onto sterile slide; microdissected cells were transferred into lysis buffer, mechanically disrupted by vigorous vortexing and Qiashredder columns. This was followed by RNA isolation and elution of RNA in 30 µl RNase-free water, according to the manufacturer's protocol (RNeasy Kit, Qiagen).

For FFPE tissues, four 5-µm serial sections were deparaffinized in xylene and graded ethanol [30, 33, 34] and RNA from microdissected cells was isolated using the RNeasy FFPE Kit with an additional DNase digestion step and elution of RNA in 15 µl RNase-free water (Qiagen).

All RNA samples were assessed for RNA quantity and quality using both the NanoDrop system (ND1000, Peqlab) as well as the Bioanalyzer 2100 system providing RNA integrity numbers [35] (RNA 6000 Nano LabChip, Agilent) (Table 1). The (tumor) cell content and RNA elution volumes were different for the processed samples, so

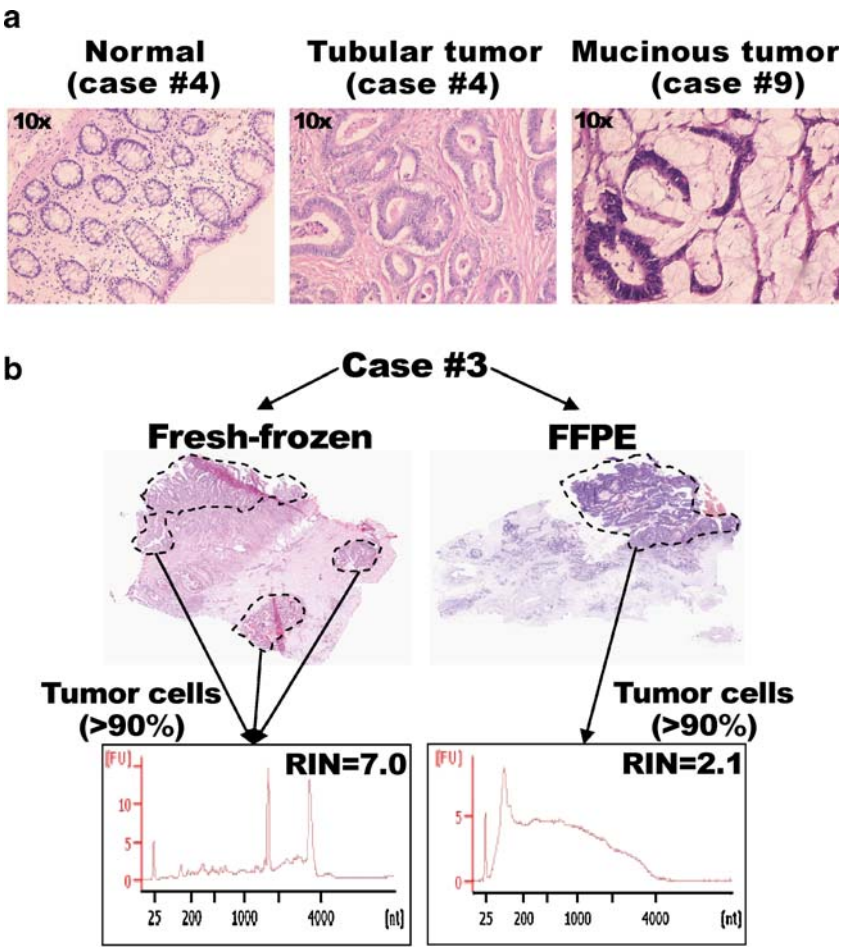


Fig. 1 Representative examples of tissue samples processed for microarray analysis. Panel (**a**) shows HE-stained FFPE sections of normal colorectal epithelium (*left*, case 4) as well as invasive tumor cells of a tubular (*middle*, case 4) and mucinous (*right*, case 9) colorectal tumor. Panel (**b**) shows hematoxylin and eosin (HE)-stained tissue sections of a matched fresh-frozen (*left*) and FFPE (*right*) tumor (case 3). Invasive tumor cells were selected for microdissection (*black dashed lines*), RNA isolation, and microarray analysis ("Materials and methods" section). Bioanalyzer2100 analysis of the corresponding tumor cell RNA extracts are given below, with high RNA quality of fresh-frozen tumor cells (RIN=7.0, good detection of 18S and 28S rRNA peaks) and poor RNA quality of FFPE tumor cells (RIN=2.1; major fragmentation, no detection of 18S and 28S rRNA). Refer also to Table 1

RNA concentrations differed between cases as well as between matched fresh-frozen and FFPE tissues. The different quantity of RNA was not a limiting factor, as only 100 ng of total RNA was required for further microarray processing.

In addition, after first strand cDNA synthesis for microarray hybridization (below), a 2-µl aliquot was taken for additional analysis of 3′ and 5′ located sequences of the β-actin gene (primers, probes, and reaction conditions according to Appendix B of the "Paradise Reagent System" protocol, Arcturus). The mean of the threshold cycles (Ct) of duplicate Q-RT-PCR analysis were used for calculation of the 3′ to 5′ β-actin ratio for each sample (Table 1).

RNA processing for microarray analysis

For hybridization to the HG-U133_Plus_2 GeneChip (Affymetrix), 100 ng of total RNA (as measured by Nanodrop, ND1000, Peqlab) were processed by the WT-Ovation™ FFPE RNA Amplification System (NuGEN) with slight modifications: After first strand cDNA synthesis, 2 µl of cDNA was saved for Q-RT-PCR analysis of 3′ and 5′ sequences of the β-actin gene (see above). The rest of 8 µl cDNA was processed for the second strand cDNA synthesis and Agencourt RNAClean bead purification. Purified cDNA was further processed by linear SPIA amplification, purification by the Qiaquick Purification Kit (Qiagen) and subsequent spectrophotometric measurement (Nanodrop, ND1000, Peqlab). For each sample, 5 µg of purified cDNA was then subjected to fragmentation and labelling using the FL-OvationTM cDNA Biotin Module V2 (NuGEN) according to the manufacturer's protocols. Fragmented and labelled cDNAs were stored at −20°C until use. Quality of all processed cDNA samples was assessed before and after fragmentation and labelling using the Bioanalyzer2100 system (RNA 6000 Nano LabChips, Agilent).

Microarray analysis

The entire fragmented and labelled cDNA (5 µg) of each sample was hybridized to the HG-U133_Plus_2 GeneChips (Affymetrix) by mixing 5 µl control oligonucleotide B2, 15 µl of 20× Eukaryotic Hybridization controls, 150 µl of 2× Hybridization buffer, 30 µl DMSO, 50 µl water (Affymetrix), and the fragmented and labeled cDNA. Hybridization was for 18 h at 45°C with rotation (60 rpm; GeneChip Hybridization Oven 640, Affymetrix). Subsequently, microarrays were washed and stained using the "GeneChip Hybridization, Wash and Stain Kit" (Affymetrix) with the GeneChip Fluidics work station 450 and wash protocol "450_0001". The scans were performed on the GeneChip Scanner 3000.

Microarray data evaluation and statistical testing

After microarray scanning, .cel files were generated by GCOS software (Affymetrix) and raw data was further processed using the sequence dependent GeneChip robust multi-array average expression measure (gcrma) [36] implemented in the corresponding BioConductor R statistical programming software package. Weakly regulated and therefore noisy and less informative probe sets were eliminated unsupervised by filtering according to standard deviation across all samples. For comparisons of measurements obtained from FFPE and fresh-frozen samples, an additional quantile normalization step of each array was applied to ensure that the observed effects are not due to global changes in the intensity levels. Average linkage hierarchical cluster analysis was performed using the R multi-dimensional analysis package "amap". Pearson correlation was used as distance measure. For clustering, the intensities for each probe set were standardized to mean zero and to a standard deviation of one.

For quality control, all standard assessment criteria implemented in the R package "simpleaffy" [37] were evaluated. To determine the effect of RNA fragmentation on microarray hybridization, the decrease in intensities between 5′ and 3′ ends for all the probe sets with 11 probes (99%) was calculated. To evaluate the reproducibility of individual gene detection, the correlation between probe sets belonging to the same gene (i.e., replicates) were calculated. Again, informative genes were selected unsupervised according to measured regulation. For comparison, a reference distribution was generated by a random selection of probe set pairs from the same set of the mostly regulated genes.

Supervised clustering analysis was performed after selection of class separating genes (here different cases), which were obtained by ranking of p values from standard analysis of variance (ANOVA). This ranking corresponds to the order of probe sets obtained after estimation of the false discovery rate according to Benjamini and Hochberg [38]. The design of our experiments was chosen to ensure that the grouping of the samples according to the five cases is orthogonal to the grouping according to the two types of tissue processing.

Unsupervised clustering analyses were performed after selection of the most variable probe sets between samples analyzed (see "Results" section). Elimination of the measurement difference observed between fresh-frozen and FFPE samples was done by centering the log intensities of each probe set independently to mean of zero in both groups. Because this step is independent on the case pair information, the results are still unsupervised with respect to the cases. Unsupervised clustering was done post correction as above for the most variable probe sets.

Validation of mRNA expression profiles by Q-RT-PCR and immunohistochemistry

For validation of mRNA expression profiles obtained, established Q-RT-PCR analysis of five genes [8, 30, 33, 34, 39, 40] was performed on the RNA extracts also used for microarray analysis. In brief, 600 ng of RNA were transcribed into 60 µl cDNA using MMLV. For each gene, 3 µl of cDNA, 3 µl of water, and each 3 µl forward and reverse primer, 3 µl TaqMan probe, 15 µl universal master mix were run for 40 cycles under standard conditions (Applied Biosystems). The resulting cycle threshold (Ct) values were used for comparison with microarray data and the $2^{(-ddCt)}$ formula for comparing gene expression patterns between different tissue samples with an universal reference RNA for normalization [32].

For immunohistochemistry of TOPOIIa, TYMS (TS), and STK6 (Aurora-A, STK15) [40], serial sections of FFPE tissues were deparaffinized, pretreated in a steamer (TOPOIIa=pH 9; TYMS=pH 6; STK6=pH 9) and stained (TOPOIIa=1:50 mouse mAb, DCS/Abcam, Cambridge, UK; TYMS=1:75 mouse mAB, clone TS106, DAKOCytomation, Glostrup, Denmark; STK6=1:50 Aurora Kinase 2, clone JLM28, Loxo/Novocastra, Dossenheim, Germany) according to standard protocols on an autostainer (DAKOCytomation).

Results

Comparison of RNA extracts of fresh-frozen and fixed tissue specimens

RNA extracts of microdissected FFPE tissues were of poorer quality (range of RINs [35]=1.9–2.6) as those obtained after microdissection of matching fresh-frozen tissues (range of RINs=7.0–9.0), the cancer cell line HCT116 (RIN=9.2), and the universal reference RNA (uRNA [32], RIN=7.6) ("Materials and methods" section; Table 1; Fig. 1).

The higher extent of RNA fragmentation in FFPE samples was also reflected by Q-RT-PCR analysis for 3′ and 5′ located sequences of the β-actin gene ("Materials and methods" section, Table 1). Fresh-frozen samples yielded RNA with a 3′/5′ ratio close to a perfect correlation of 1 (0.96±0.04), whilst this was reduced in FFPE samples (0.85±0.03). The HCT116 and uRNA RNA samples had 3′/5′ β-actin ratios of 0.97 and 1.02, respectively.

Examination of standard and extended quality control parameters in microarray data sets

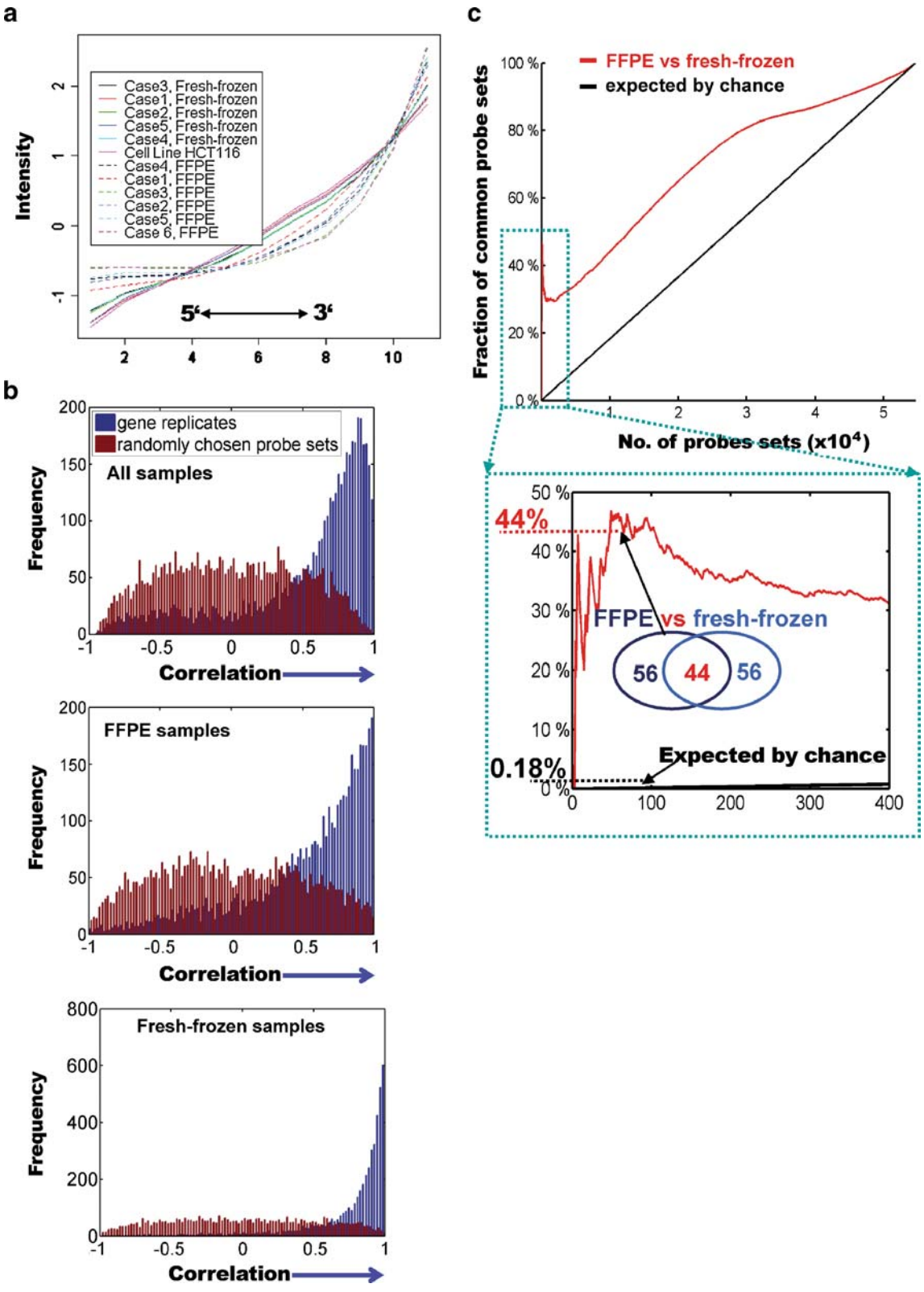
Irrespective of their performance in NanoDrop, Bioanalyzer2100, or Q-RT-PCR analysis, all RNA samples (except uRNA and case 6/fresh-frozen) were hybridized to the HG-U133_Plus_2 GeneChip ("Materials and methods" section). Successful hybridization was seen for all 18 samples (present calls=53.9% cell line HCT116; 49.90±4.43% fresh-frozen samples; 36.34±9.60% FFPE samples), with microarrays in an acceptable range of first line quality controls by "simpleaffy" software analysis [37] (Table 1, Supplementary Data Fig. 1).

Further to this, additional parameters of reliability, information content, and reproducibility were examined in detail in the matched fresh-frozen and FFPE samples ("Materials and methods" section): First, a slight reduction of signal intensities at the 5′ located probes was observed for FFPE as compared to fresh-frozen samples (Fig. 2a), independent of raw data processing in gcrma, rma, or mas5 [37, 41]. Second, correlation of gene-specific ("information") as compared to random ("no information") probe set replicates was shifted towards a perfect correlation of "1" (Fig. 2b), although less pronounced for FFPE samples. Third, the overlap, i.e., concordance of probe sets from fresh-frozen and FFPE samples, was high (Fig. 2c). Finally, correlation of all signals between microarray datasets (Supplementary Data Fig. 2) yielded a high reproducibility within fresh-frozen (r^2=0.949) and FFPE (r^2=0.923) samples and a high fidelity for matched fresh-frozen and FFPE samples (case 1—r^2=0.894; case 2—r^2=0.909; case 3—r^2=0.901; case 4—r^2=0.938; case 5—r^2=0.910). The mean correlation of microarrays of all samples was r^2=0.899.

Gene expression profiles in fresh-frozen and FFPE tissue samples of colorectal tumors

First, microarray datasets of matched fresh-frozen and FFPE samples were analyzed by supervised clustering after

Fig. 2 Summary of detailed microarray datasets analyses. Panel (**a**) shows the intensity of probe signals (*y*-axis) according to probe location within the gene sequence (*x*-axis) for the investigated samples. Note that for FFPE samples (*dashed lines*), loss of intensity is observed with increasing location of the probes towards the 5′ location of sequences. Panel (**b**) shows results of the correlation of intensities of gene-specific probe set replicates (*blue bars*) as compared to random probe sets (*red bars*) for all samples (*top*) and FFPE (*middle*) and fresh-frozen (*bottom*) samples separately. Note that, for all samples (*top*), the correlation (*y*-axis) is high for replicate probe sets (shift to 1 on *x*-axes) as compared to random probe sets (distribution from −1 to 1 on *x*-axes). Separate analysis of only fresh-frozen and FFPE showed essentially the same, but the effect was more pronounced for fresh-frozen samples. Panel (**c**) provides the results of concordance analysis of all probe sets from fresh-frozen and FFPE samples (*top*) and, as an example, of the top 100 regulated probe sets (*lower*). Note that, for the latter, an overlap of 44% of the probe sets is observed, as indicated in the *lower graph* on the *y*-axis and by the Venn diagram (*light line*=fresh-frozen, *dark line*=FFPE). By chance, not a single probe set would be expected (0.18%, indicated on *x*-axis)


a
Intensity
Case3, Fresh-frozen
Case1, Fresh-frozen
Case2, Fresh-frozen
Case5, Fresh-frozen
Case4, Fresh-frozen
Cell Line HCT116
Case4, FFPE
Case1, FFPE
Case3, FFPE
Case2, FFPE
Case5, FFPE
Case 6, FFPE
5'
3'
b
gene replicates
randomly chosen probe sets
All samples
FFPE samples
Fresh-frozen samples
Frequency
Correlation
c
FFPE vs fresh-frozen
expected by chance
Fraction of common probe sets
No. of probes sets (x10⁴)
44%
0.18%
FFPE vs fresh-frozen
56
44
56
Expected by chance

pre-selection of the most differentially expressed genes between cases ("Materials and methods" section). Not surprisingly, the resulting heatmaps and corresponding dendrograms showed clustering of fresh-frozen and FFPE samples strictly according to individual cases (Fig. 3a). This was not influenced by the number of pre-selected genes.

In view of the measurement difference in microarray datasets of fresh-frozen and FFPE samples (Fig. 2b,c), we next tested whether case-specific information is sustained if no presumptions are made. For this unsupervised clustering using the most variable probe sets ("Materials and methods" section), we also included the cell line sample (HCT116, an optimal RNA sample) and the FFPE sample of case #6 (derived from 2003 as compared to case #1–5 FFPE samples of 2007/2008) to obtain preliminary information if the source and/or age of the samples have an effect on microarray analysis.

This resulted in a clear-cut separation of samples according cell/tissue processing, whereby fresh-frozen samples and the cell line HCT116 formed one cluster and all FFPE samples were found in a second cluster (Fig. 3b). Within these two clusters, the cell line was separated from the fresh-frozen tissues and the case #6 FFPE sample from the more recent case #1–5 FFPE samples, suggesting that the source and/or age of the samples may impact mRNA expression profiles. As assessed by Affymetrix GCOS software, the percent of apparently "down-" or "up-" regulated (i.e., discordant) probe sets in matched fresh-frozen and FFPE samples was 7–12% and 6–11%, but was similar (8–12% and 6–12%, respectively) upon cross-examination of fresh-frozen and FFPE samples of two different cases. The list of differentially "regulated" (i.e., discordant) genes determined by unsupervised clustering of the 1,000 most variable probe sets, is given in Supplementary Table 1.

Correction of microarray datasets allows comparative analysis of fresh-frozen and FFPE tissue samples

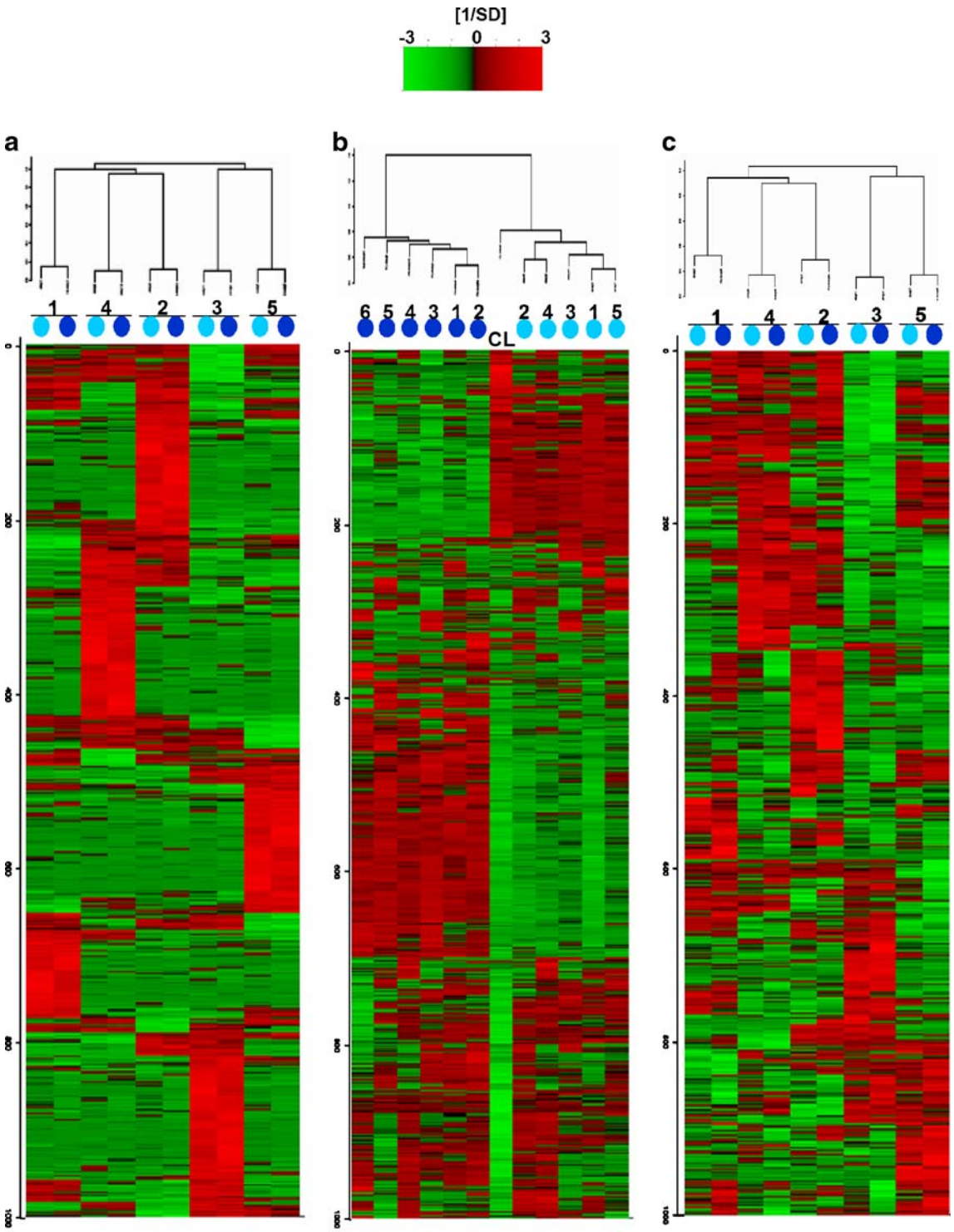
To account for the measurement difference in fresh-frozen and FFPE samples (Fig. 2a,b), we next tested whether a mathematical correction step allows comparative analysis of microarray datasets derived from differently processed clinical samples ("Materials and methods" section). For this, we subtracted the average log intensity for each probe set separately in the groups of fresh-frozen and FFPE samples. This correction step is the most obvious adjustment accounting for probe set dependent measurement differences, it is multi-directional and is independent on the intensities measured (Supplementary Data Fig. 3).

Indeed, unsupervised clustering analysis with corrected datasets of matched fresh-frozen and FFPE samples clearly separated samples according to the individual cases, irrespective of tissue processing (Fig. 3c, list of probe sets is given in Supplementary Table 1). The results were independent of the number of probe stets used, similar for raw data processing in gcrma (Fig. 3c), rma or mas5 and if only probe sets with positive present call were used. Moreover, upon unsupervised clustering with uncorrected or corrected microarray datasets including all 50,000 probe sets, samples clustered according to tissue processing or according to case, respectively (Supplementary Data Fig. 4). Therefore, the correction step preserves case-specific information in fresh-frozen and FFPE tissue samples, i.e., reflecting the molecular biology of individual tumors independent of tissue processing.

Validation of gene expression profiles in matched fresh-frozen and FFPE samples by Q-RT-PCR

Q-RT-PCR analysis was performed for detection of the housekeeping gene TATA-box binding protein (TBP) and the five genes thymidylate synthase (TYMS), survivin, signal transducer and activator of transcription STAT1 and STAT5 in the fresh-frozen and FFPE tumor tissues ("Materials and methods" section, Fig. 4a–f) [8, 30, 33, 34, 39, 40]. Correlation between microarray and Q-RT-PCR data in matched fresh-frozen and FFPE tumor samples was acceptable for 4/5 genes (range r^2=0.4867 to r^2=0.8044), but failed for 1/5 genes (STAT1, r^2=0.00029, Fig. 4e). Upon separate analysis of fresh-frozen and FFPE tumor samples, good correlations of microarray and Q-RT-PCR data (r^2=0.9866 and r^2=0.5109, respectively) were also obtained for STAT1 (Fig. 4f). Besides the possible technical bias introduced by the design of Q-RT-PCR primers, only detecting one 124 bp region within the STAT1 gene as

Fig. 3 Gene expression profiles of fresh-frozen and FFPE colorectal tumors. Panel (**a**) shows the heatmap and corresponding dendrogram of supervised clustering analysis of matched fresh-frozen and FFPE samples after pre-selection of the 1,000 most differentially expressed genes between cases, as determined by ANOVA ("Materials and methods" section). Panel (**b**) shows the heatmap and dendrogram of unsupervised clustering with the 1,000 most variable probe sets in matched fresh-frozen and FFPE samples (case #1–5) as well as one freshly harvested cell line (CL) and one older FFPE sample (case #6) as "controls" ("Materials and methods" section). Panel **(c)** shows the heatmap and dendrogram of unsupervised clustering analysis of matched fresh-frozen and FFPE samples with the 1,000 most variable probe sets after correction for the measurement difference of fresh-frozen and FFPE sample microarray datasets ("Materials and methods" section). *Numbers* and *filled circles* provide case identity and tissue processing (*light circles*=fresh-frozen; *dark blue circles*=FFPE); *CL* cell line HCT116. The *color bar* of expression levels is given at the *top* of the figure (*green*=low expressed genes; *red*=highly expressed genes). Note that after the correction step (panel **c**), the unsupervised clustering was strictly according to the case-specific biology and not tissue processing (panel **b**) and that the relatedness of cases was the same for as for supervised clustering (panel **a**). Refer also to main text and Supplementary Figure 4 (unsupervised clustering with or without correction with all 50,000 probe sets)


[1/SD]
-3
0
3
a
b
c
1 4 2 3 5
6 5 4 3 1 2 2 4 3 1 5
CL
1 4 2 3 5
0
200
400
600
800
1000

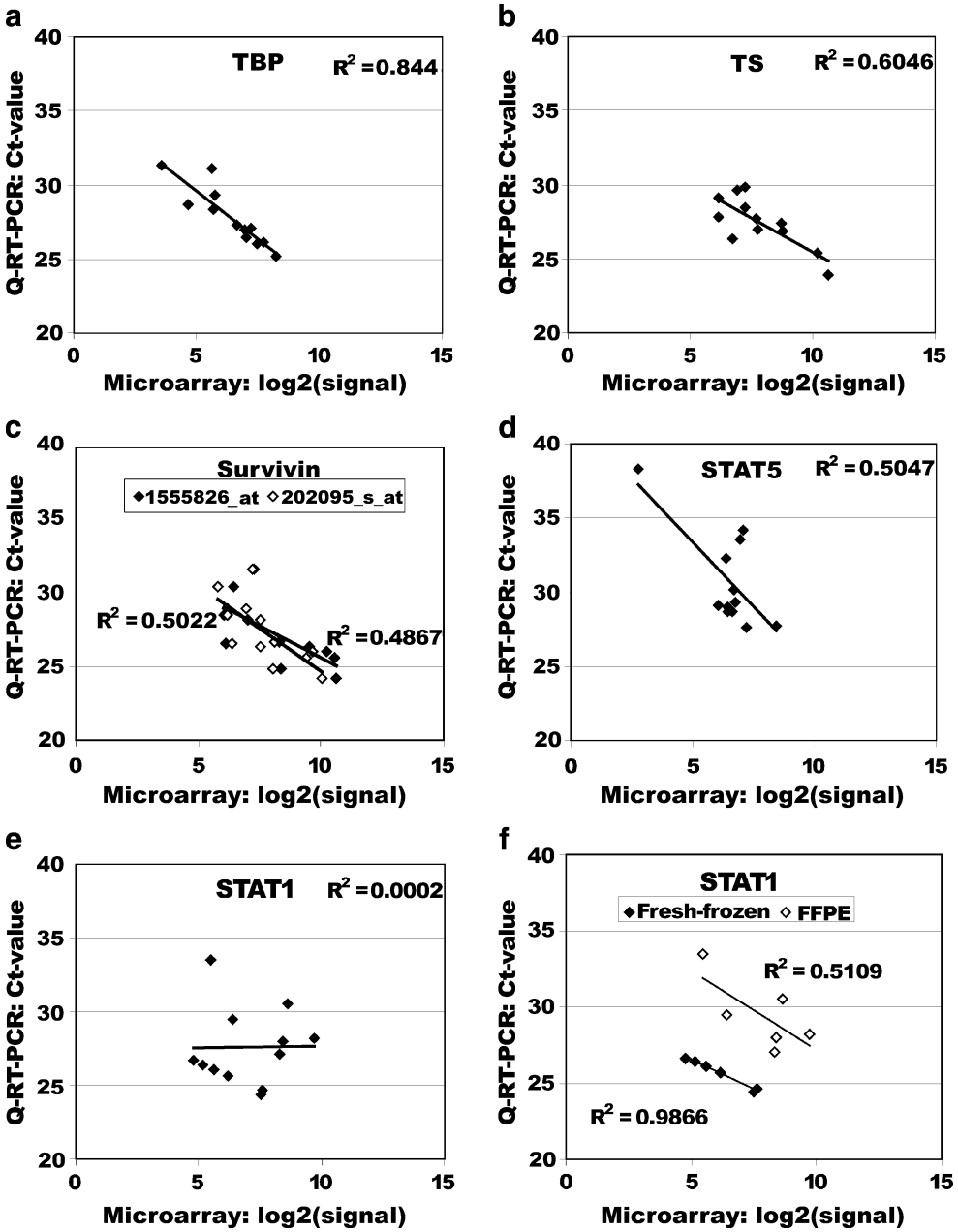


Fig. 4 Q-RT-PCR validation of microarray data from matched fresh-frozen and FFPE samples. The figure shows the correlations between microarray analysis (*x*-axis; log2(signal)) and Q-RT-PCR analysis (*y*-axis; Ct value) for expression of five genes (**a** TATA-box binding protein/TBP; **b** thymidylate synthase/TS; **c** survivin; **d** signal transducer and activator of transcription STAT5a; **e**, **f** STAT1) [14, 17, 39, 40, 41]. Note that, except for STAT1 (panel **e**, **f**), correlations were in a good range. The missing correlation for STAT1 was due to a marked difference in the performance of fresh-frozen and FFPE samples, as upon separate analysis of fresh-frozen and FFPE samples (**f**) correlations were r^2=0.9866 and r^2=0.5109, respectively

opposed to the microarray probe set spanning the entire STAT1 gene, this may also indicate an (additional) effect of the duration of tissue processing on rapid turnover genes.

Gene expression profiles in normal and malignant colorectal FFPE tissue samples

Having shown that comparison of microarray datasets from fresh-frozen and FFPE tissue samples is possible, we next examined whether microarray analysis of FFPE samples alone yields biologically relevant information. For this, we examined "biological replicates" of normal colorectal epithelium (n=3) and invasive tumors (n=9), with different tumor histotypes (5/9 tubular, 4/9 mucinous [29]).

Unsupervised clustering analysis by using the 1,000 most variable probe sets across all FFPE samples indeed yielded informative data on the biological context of the "replicates" (Fig. 5a). First, there was a clear distinction between normal epithelial and invasive tumor cell samples (list of probe sets is given in Supplementary Table 1). Genes down- and up-regulated in tumors included for example ANPEP, BMP2, MT1H (down) and CSPG2, KPNA2, COL1A2, SPARC, ALDOA, and SORD (up), as previously identified by us [42]. Other genes differentially expressed included for example the known genes TOPOIIa (topoisomerase IIa), TYMS (thymidylate synthase), and STK6 (Aurora-A), which were clearly up-regulated in tumors. Moreover, there was a clear distinction between

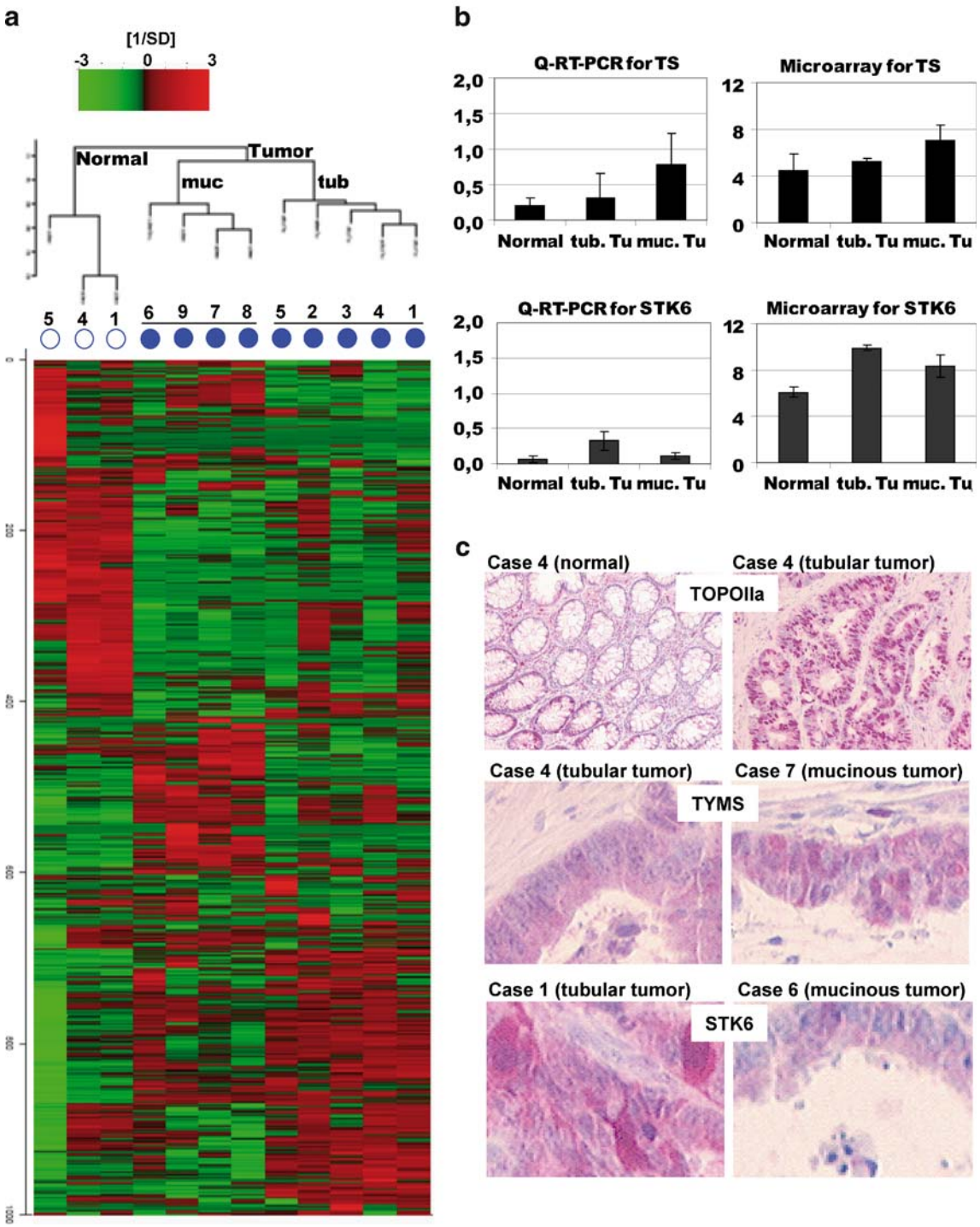


Fig. 5 Gene expression profiles of normal and malignant colorectal FFPE tissues. Panel (**a**) shows the heatmap and dendrogram of unsupervised clustering of all 12 FFPE samples with the 1,000 most variable probe sets ("Materials and methods" section). *Numbers* and *colored circles* provide case and tissue identity (*filled dark blue circles*=tumors; *lined dark blue circles*=normal colorectal epithelium). The *color bar* of expression levels is given at the top of the figure (*green*=low expressed genes; *red*=highly expressed genes). Note that (1) normal colorectal epithelium and invasive tumor samples clustered separately and that (2) within the tumor samples the two different histological subtypes were further separated in two corresponding clusters. Panel (**b**) shows correlation of Q-RT-PCR and microarray data (mean±std. dev) for TYMS (TS; *top row*) and STK6 (*bottom row*) expression in normal epithelial and invasive (tubular or mucinous) tumor cells. Panel (**c**) shows representative examples of immunohistochemistry for *top row* TOPOIIa (10×) in normal epithelium and invasive tumor (case #4, note the increased TOPOIIa staining in invasive tumor cells), for *middle row* TYMS/TS protein expression (40×) in a tubular (case #4, *left*) and a mucinous (case #7, *right*) tumor (note the higher rate of nuclear TYMS/TS staining in case #7), and for *bottom row* STK6 protein expression (40×) in a tubular (case #1, *left*) and a mucinous (case #6, *right*) tumor (note the higher intensity of STK6 staining in case #1)

expression profiles from tubular (CIN-type) and mucinous (MIN-type) sporadic cancers (list of probe sets is given in Supplementary Table 1): The cell-cycle-associated gene thymidylate synthase (TYMS) as well as the centrosome-associated gene STK6 was identified not only to be up-regulated in tumors but also to be differentially regulated in the two histotypes of colorectal tumors (TYMS/TS, highest expression in mucinous tumors; STK6, highest expression in tubular tumors).

Validation of gene expression profiles in normal and malignant FFPE samples by Q-RT-PCR and immunohistochemistry

Q-RT-PCR and immunohistochemistry were used to validate TYMS, STK6, and TOPOIIa expression within the different "biological replicates" of normal colorectal epithelium and colorectal cancer specimens ("Materials and methods" section, case #2 excluded due to familiar HNPCC-associated biology as opposed to sporadic colorectal cancers in cases #1 and #3–9).

Both Q-RT-PCR and microarray analyses revealed highest TYMS expression in mucinous tumors, followed by tubular tumors and normal epithelium (Fig. 5b). In addition, both Q-RT-PCR and microarray analyses showed highest STK6 expression in tubular tumors, followed by mucinous tumors and normal epithelium (Fig. 5b).

Irrespective of potential post-translational modifications, TOPOIIa, TYMS, and STK6 protein expression largely reflected mRNA levels of microarray and Q-RT-PCR analyses (Fig. 5c). Protein expression of (1) TOPOIIa was increased in tumors as compared to normal epithelial tissues, (2) nuclear TYMS appeared to be present at a higher rate in mucinous as compared to tubular tumors, and (3) STK6 was more highly expressed in tubular as compared to mucinous tumors.

Discussion

Routinely processed formalin-fixed and paraffin-embedded (FFPE) tissue samples represent an extensive and valuable source for large-scale, microarray-based retrospective studies. So far, reliable genome-wide microarray analyses of human tissue specimens are limited to the use of fresh-frozen tissue samples, in which the RNA molecules are still largely intact [1–6]. Here, we demonstrate that microarray analysis is possible, precise, and reliable from small amounts of RNA extracted after microdissection of tumor cells from FFPE tissue specimens processed for routine histopathological diagnostics. Furthermore, by mathematical correction of the different quality of fresh-frozen and FFPE sample microarray datasets, comparative analyses are possible irrespective of tissue processing. Most importantly, biologically relevant information can be obtained by microarray analysis of microdissected cell samples from different FFPE tissue specimens such as normal colorectal epithelium and invasive colorectal tumor cells or even different tumor histotypes of colorectal cancer [29].

The present study included fresh-frozen and FFPE samples of routinely processed colorectal cancer resection specimens and microarray analysis was successful in all samples, irrespective of the rather poor quality RNA in FFPE samples. Present calls (mean 36.4±9.6% for all normal and malignant colorectal FFPE tissues) were higher than in previous studies also using Affymetrix GeneChips [20, 23, 26]. This high performance in FFPE samples is most likely due to random priming of (degraded) RNA for cDNA synthesis, which is also an integral part of robust Q-RT-PCR approaches for FFPE samples [6, 8, 30, 33, 34, 40]. As optimal RNA integrity is difficult to guarantee entirely in a day-to-day clinico-pathological setting, random priming of cDNA synthesis may represent a major step forward to making (Q-RT-PCR and/or) microarray analysis more widely applicable to routine diagnostic tissue specimens.

Importantly, our approach works with small amounts (100 ng) of total RNA isolated from microdissected (tumor) cells of tissue sections derived from large diagnostic primary colorectal cancer resection specimens, i.e., distinguishing it from the previous reports on investigation of cell lines [25], cell pellets [20], and specifically prepared tissues [19, 21] and/or entire tissue "pieces" [18, 19, 25]. Our approach is feasible and reliable for a molecular (pathological) laboratory in a clinical environment, as demonstrated by performing the laboratory protocol at different time points for different samples and by applying a practicable microdissection step, which rules out gross contamination of "tumor"-specific microarray datasets by for example stromal cells [42]. Manual- [22, 26] and laser-assisted [21, 24] microdissection had been combined with microarray analysis of FFPE samples before, but required one or two steps of—(potentially biasing [43, 44])—in vitro transcription to obtain sufficient material for microarray hybridization. Still, subsequent studies should further strengthen the feasibility and performance of our approach in other laboratory settings and/or its application to other organ systems, for example those with higher RNase activities.

With reliable microarray datasets obtained from matched fresh-frozen and FFPE tissue samples of one cancer entity, our study also revealed the effect of tissue processing on microarray datasets. Previous studies [22, 23] also included matched fresh-frozen and FFPE samples, but as these were of different cancer entities the resulting clustering was primarily in favor of discriminating cancer entities. Interestingly, in these studies a clear separation of fresh-frozen

and FFPE samples occurred *within* the cancer clusters [22, 23]. This was similar to our findings of matched fresh-frozen and FFPE colorectal cancer samples and we further examined this phenomenon.

Although microarray datasets were informative, reproducible, and accurate within the groups of fresh-frozen and FFPE samples, we identified a so far neglected systematic measurement difference between fresh-frozen and FFPE samples. Not surprisingly, we found reduced signal intensities at probe sets corresponding to 5′ located gene sequences in FFPE samples. This was also exemplified by a poor correlation of microarray to Q-RT-PCR data for 1/5 tested genes (STAT1) between fresh-frozen and FFPE samples. Different localization of Q-RT-PCR primers and microarray probe sets within gene sequences, the efficiency of cDNA synthesis and the length of the Q-RT-PCR product may have influenced the correlation. The time of tissue processing may however be an additional variable, preferentially affecting specific genes, such as for example those induced by tissue hypoxia [45, 46]. Still, as all our microarray datasets were proven to be informative and reproducible, we examined whether patient-specific gene expression information is preserved in the fresh-frozen and FFPE samples of the same case. For this, we eliminated the measurement difference of fresh-frozen and FFPE samples at the level of data processing and demonstrated that this preserved the individual biology of the sample, i.e., patient information. Further, detailed studies are clearly warranted to fully investigate genes and biological pathways affected by tissue processing and to further confirm the correction step in independent sample sets, for example, by its application to existing microarray datasets [22, 23, 26].

Clearly, microarray analysis of FFPE samples from tissue specimens of different biological background by our approach yielded gene expression profiles that not only distinguished between normal and malignant colorectal cells, but also correctly recognized the different tumor histotypes of colorectal cancers described by the WHO [29].

Specifically, we detected ten genes differentially expressed in microdissected normal colorectal epithelial and invasive tumor cells, also previously identified by us in an unrelated set of fresh-frozen samples [42]. Validation of three additional genes by Q-RT-PCR and/or immunohistochemistry further supported our microarray results, with tumors expressing high levels of topoisomerase IIa (TOP-OIIa), thymidylate synthase (TYMS, TS), and the centrosome-associated kinase STK6 (Aurora-A, STK15, BTAK). Moreover, TYMS and STK6 were shown to be differentially expressed in sporadic tubular and mucinous tumors by microarray, Q-RT-PCR and immunohistochemical analyses. The sporadic tubular and mucinous tumors studied here were of the chromosomal- (CIN) and microsatellite (MIN)-instable type, respectively (Table 1, "Materials and methods" section). It is known that STK6 (STK15, AURKA) is amplified and overexpressed [30,31] preferentially in CIN tumors, whilst MSI tumors exhibit a higher proliferation rate [47] and associated need for TYMS. This is, at least partially, reflected in the present microarray, Q-RT-PCR, and immunohistochemical analyses.

In summary, the presented microarray approach is a highly valid tool to examine biological gene expression profiles from large-scale series of archival tissue specimens, allowing an extended insight into the molecular biology of cancers "*ex/in situ*". Its relevance for and potential integration into a clinico-pathological diagnostic setting however still awaits further investigation.

Acknowledgements The study would not have been possible without the long-term substantial research support to SL and MW by the Mushett Family Foundation, Chester, NJ, USA. The study has been funded in parts (in vitro experiments) by a grant to SL and MW from the German Research Foundation (DFG). SL and MW also acknowledge their collaborators Drs. Laura Tang and David Klimstra at the Memorial Sloan Kettering Cancer Center, New York. CK and JT acknowledge financial support by the BMBF-grant 0313921 FRIYSY.

References

1. Hoheisel JD (2006) Microarray technology: beyond transcript profiling and genotype analysis. Nat Rev Genet 7:200–210
2. Sotiriou C, Piccart MJ (2007) Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care. Nat Rev Cancer 7:545–553
3. Brazma A, Hingamp P, Quackenbush J et al (2001) Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet 29:365–371
4. MAQC ConsortiumShi L et al (2006) The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. Nat Biotechnol 24:1151–1161
5. Frueh FW (2006) Impact of microarray data quality on genomic data submissions to the FDA. Nat Biotechnol 24:1105–1107
6. Farragher SM, Tanney A, Kennedy RD et al (2008) RNA expression analysis from formalin fixed paraffin embedded tissues. Histochem Cell Biol 130:435–445
7. Vincek V, Nassiri M, Nadji M, Morales AR (2003) A tissue fixative that protects macromolecules (DNA, RNA, and protein) and histomorphology in clinical samples. Lab Invest 83:1427–1435
8. Specht K, Richter T, Müller U et al (2001) Quantitative gene expression analysis in microdissected archival formalin-fixed and paraffin-embedded tumor tissue. Am J Pathol 158:419–429
9. Antonov J, Goldstein DR, Oberli A et al (2005) Reliable gene expression measurements from degraded RNA by quantitative real-time PCR depend on short amplicons and a proper normalization. Lab Invest 85:1040–1050
10. Masuda N, Ohnishi T, Kawamoto S et al (1999) Analysis of chemical modification of RNA from formalin-fixed samples and optimization of molecular biology applications for such samples. Nucleic Acids Res 27:4436–4443
11. von Ahlfen S, Missel A, Bendrat K, Schlumpberger M (2007) Determinants of RNA quality from FFPE samples. PLoS ONE 2: e1261

12. Karsten SL, Van Deerlin VM, Sabatti C et al (2002) An evaluation of tyramide signal amplification and archived fixed and frozen tissue in microarray gene expression analysis. Nucleic Acids Res 30:E4
13. Xiang CC, Chen M, Ma L et al (2003) A new strategy to amplify degraded RNA from small tissue samples for microarray studies. Nucleic Acids Res 31:e53
14. Loudig O, Milova E, Brandwein-Gensler M et al (2007) Molecular restoration of archived transcriptional profiles by complementary-template reverse-transcription (CT-RT). Nucleic Acids Res 35:e94
15. Chen J, Byrne GE Jr, Lossos IS (2007) Optimization of RNA extraction from formalin-fixed, paraffin-embedded lymphoid tissues. Diagn Mol Pathol 16:61–72
16. Bibikova M, Talantov D, Chudin E et al (2004) Quantitative gene expression profiling in formalin-fixed, paraffin-embedded tissues using universal bead arrays. Am J Pathol 165:1799–1807
17. Bibikova M, Yeakley JM, Wang-Rodriguez J, Fan JB (2008) Quantitative expression profiling of RNA from formalin-fixed, paraffin-embedded tissues using randomly assembled bead arrays. Methods Mol Biol 439:159–177
18. Ravo M, Mutarelli M, Ferraro L et al (2008) Quantitative expression profiling of highly degraded RNA from formalin-fixed, paraffin-embedded breast tumor biopsies by oligonucleotide microarrays. Lab Invest 88:430–440
19. Lee J, Hever A, Willhite D, Zlotnik A, Hevezi P (2005) Effects of RNA degradation on gene expression analysis of human post-mortem tissues. FASEB J 19:1356–1358
20. Scicchitano MS, Dalmas DA, Bertiaux MA et al (2006) Preliminary comparison of quantity, quality, and microarray performance of RNA extracted from formalin-fixed, paraffin-embedded, and unfixed frozen tissue samples. J Histochem Cytochem 54:1229–1237
21. Coudry RA, Meireles SI, Stoyanova R et al (2007) Successful application of microarray technology to microdissected formalin-fixed, paraffin-embedded tissue. J Mol Diagn 9:70–79
22. Penland SK, Keku TO, Torrice C et al (2007) RNA expression analysis of formalin-fixed paraffin-embedded tumors. Lab Invest 87:383–391
23. Frank M, Döring C, Metzler D, Eckerle S, Hansmann ML (2007) Global gene expression profiling of formalin-fixed paraffin-embedded tumor samples: a comparison to snap-frozen material using oligonucleotide microarrays. Virchows Arch 450:699–711
24. Haque T, Faury D, Albrecht S et al (2007) Gene expression profiling from formalin-fixed paraffin-embedded tumors of pediatric glioblastoma. Clin Cancer Res 13:6284–6292
25. Lauss M, Vierlinger K, Weinhaeusel A et al (2007) Comparison of RNA amplification techniques meeting the demands for the expression profiling of clinical cancer samples. Virchows Arch 451:1019–1029
26. Linton KM, Hey Y, Saunders E et al (2008) Acquisition of biologically relevant gene expression data by Affymetrix micro-array analysis of archival formalin-fixed paraffin-embedded tumours. Br J Cancer 98:1403–1414
27. Srivastava PK, Küffer S, Brors B et al (2008) A cut-off based approach for gene expression analysis of formalin-fixed and paraffin-embedded tissue samples. Genomics. 91:522–529
28. Sobin LH, Wittekind C (2002) TNM classification of malignant tumours. UICC, 6th edn. Wiley Liss, Wilmington
29. Hamilton S, Aaltonen L (2000) Pathology and genetics of tumors of the digestive system. WHO classification of tumors. IARC, Lyon
30. Gerlach U, Kayser G, Walch A et al (2006) Centrosome-, chromosomal-passenger- and cell-cycle-associated mRNAs are differentially regulated in the development of sporadic colorectal cancer. J Pathol 208:462–472
31. Lassmann S, Weis R, Makowiec F (2007) Array CGH identifies distinct DNA copy number profiles of oncogenes and tumor suppressor genes in chromosomal- and microsatellite-unstable sporadic colorectal carcinomas. J Mol Med 85:293–304
32. Novoradovskaya N, Whitfield ML, Basehore LS et al (2004) Universal reference RNA as a standard for microarray experiments. BMC Genomics 5:20
33. Lassmann S, Schuster I, Walch A et al (2007) STAT3 mRNA and protein expression in colorectal cancer: effects on STAT3-inducible targets linked to cell survival and proliferation. J Clin Pathol 60:173–179
34. Lassmann S, Tang L, Capanu M et al (2007) Predictive molecular markers for colorectal cancer patients with resected liver metastasis and adjuvant chemotherapy. Gastroenterology 133:1831–1839
35. Schroeder A, Mueller O, Stocker S et al (2006) The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol 7:3
36. Wu Z, Irizarry RA, Gentleman R, Martinez-Murillo F, Spencer F (2004) A model-based background adjustment for oligonucleotide expression arrays. J Am Stat Assoc 99:909–917
37. Wilson CL, Miller CJ (2005) Simpleaffy: a BioConductor package for Affymetrix quality control and data analysis. Bioinformatics 21:3683–3685
38. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc B 57:289–300
39. de Kok JB, Roelofs RW, Giesendorf BA et al (2005) Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes. Lab Invest 85:154–159
40. Lassmann S, Shen Y, Jütting U et al (2007) Predictive value of Aurora-A/STK15 expression for late stage epithelial ovarian cancer patients treated by adjuvant chemotherapy. Clin Cancer Res 13:4083–4091
41. Irizarry RA, Bolstad BM, Collin F et al (2003) Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 31:e15
42. Wiese AH, Auer J, Lassmann S et al (2007) Identification of gene signatures for invasive colorectal tumor cells. Cancer Detect Prev 31:282–295
43. Kerkhoven RM, Sie D, Nieuwland M et al (2008) The T7-primer is a source of experimental bias and introduces variability between microarray platforms. PLoS ONE 3:e1980
44. Ma C, Lyons-Weiler M, Liang W et al (2006) In vitro transcription amplification and labeling methods contribute to the variability of gene expression profiling with DNA microarrays. J Mol Diagn 8:183–192
45. Chi JT, Wang Z, Nuyten DS et al (2006) Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers. PLoS Med 3(3):e47
46. Zhang H, Semenza GL (2008) The expanding universe of hypoxia. J Mol Med 86(7):739–746
47. Michael-Robinson JM, Reid LE, Purdie DM et al (2001) Proliferation, apoptosis, and survival in high-level microsatellite instability sporadic colorectal cancer. Clin Cancer Res 7:2347–2356