Enhanced stability of microRNA expression facilitates classification of FFPE tumour samples exhibiting near total mRNA degradation

JS Hall1,4, J Taylor1,2,4, HR Valentine1, JJ Irlam1, A Eustace1, PJ Hoskin1, CJ Miller2 and CML West*,1
1Translational Radiobiology Group, School of Cancer and Enabling Sciences, Manchester Academic Health Science Centre, The University of Manchester, Christie Hospital NHS Trust, Wilmslow Road, Manchester M20 4BX, UK; 2Applied Computational Biology and Bioinformatics Group, Paterson Institute for Cancer Research, Wilmslow Road, Manchester M20 4BX, UK; 3Cancer Centre, Mount Vernon Hospital, Rickmansworth Road, Northwood HA6 2RN, UK

BACKGROUND: As degradation of formalin-fixed paraffin-embedded (FFPE) samples limits the ability to profile mRNA expression, we explored factors predicting the success of mRNA expression profiling of FFPE material and investigated an approach to overcome the limitation.

METHODS: Bladder (n = 140, stored 3–8 years) and cervix (n = 160, stored 8–23 years) carcinoma FFPE samples were hybridised to Affymetrix Exon 1.0ST arrays. Percentage detection above background (%DABG) measured technical success. Biological signal was assessed by distinguishing cervix squamous cell carcinoma (SCC) and adenocarcinoma (AC) using a gene signature. As miR-205 had been identified as a marker of SCC, precursormiR-205 was measured by Exon array and mature miR-205 by qRT–PCR. Genome-wide microRNA (miRNA) expression (Affymetrix miRNA v2.0 arrays) was compared in eight newer FFPE samples with biological signal and eight older samples without.

RESULTS: RNA quality controls (QCs) (e.g., RNA integrity (RIN) number) failed to predict profiling success, but sample age correlated with %DABG in bladder (R = −0.30, P < 0.01) and cervix (R = −0.69, P < 0.01). Biological signal was lost in older samples and neither a signature nor precursor-miR-205 separated samples by histology, miR-205 qRT–PCR discriminated SCC from AC, validated by miRNA profiling (26-fold higher in SCC; P = 1.10 × 10−10). Genome-wide miRNA (R = 0.95) and small nucleolar RNA (R = 0.97) expression correlated well in the eight newer vs older FFPE samples and better than mRNA expression (R = 0.72).

CONCLUSION: Sample age is the best predictor of successful mRNA profiling of FFPE material, and miRNA profiling overcomes the limitation of age and copes well with older samples.

Keywords: microRNA; FFPE; degradation; microarray; snoRNA; profiling

Worldwide, there are an estimated one billion archived tissue samples, most of which are formalin-fixed and paraffin-embedded (FFPE) (Blow, 2007). Many of these samples are associated with long-term clinical follow-up data. This, coupled with the sequencing of the human genome and the subsequent abundance of molecular biology techniques to support both research and, increasingly, diagnosis, highlights the need to develop approaches for exploiting FFPE archives.

Although tissue and protein are well preserved in FFPE blocks, nucleic acids are degraded and chemically modified (Lee et al., 2005). These deleterious processes, which affect DNA and RNA, occur not only during sample embedding and processing (von Ahlfen et al., 2007), but also during long-term storage (Cronin et al., 2004). The RNA quality of FFPE samples stored for prolonged periods is known to reduce (Cronin et al., 2004), but there is little research on this phenomenon. Given that data quality is an important consideration for bioinformatics analyses, the impact of RNA degradation and chemical modification can be significant, ultimately overwhelming biological variation with noise. Little is known about the factors that determine RNA quality from FFPE tissues. Variables including tissue size, fixation time and storage temperature can negatively affect both RNA quality and the success of PCR assays in controlled laboratory conditions (von Ahlfen et al., 2007), but less is known about the processes that occur during prolonged storage. Studies have shown that older FFPE samples have higher Ct values (i.e., lower signal) than those stored for <3 years (Cronin et al., 2004; von Ahlfen et al., 2007). These data suggest that RNA degradation is not a single event, but continues beyond processing and throughout storage. In line with this, most RNA expression studies use FFPE samples <10 years old (von Ahlfen et al., 2007; Abdueva et al., 2010; Abramovitz et al., 2011), with very few, mostly recent, studies considering older samples (Cronin et al., 2004; Hall et al., 2011; Kennedy et al., 2011). This 10-year limit excludes the majority of FFPE material from analysis. In oncology, this is particularly pertinent for rare cancers, historical studies where shifts in aetiology may be relevant (Chaturvedi et al., 2011) and unique sample cohorts (West et al., 1997).

Recently, we showed that Exon array profiling and a specialised pipeline that exploits the redundant nature of the arrays can be used to derive a gene signature from FFPE without the requirement...
for matched fresh-frozen samples (Hall et al, 2011). We consider this to be the cutting edge of FFPE sample expression profiling (Linton et al, 2009) and, unlike earlier techniques, we captured biological information from 100% of samples aged 10–16 years, a higher success rate than other studies (Penland et al, 2007; Linton et al, 2008). We reasoned that exon array profiling coped with the RNA damage in FFPE samples.

Here, we show that while recent methods cope with RNA degradation in FFPE samples, they are less tolerant of the subsequent mRNA deterioration associated with longer-term storage. To compensate for the long-term deterioration of mRNA signals, we investigated microRNA (miRNA) profiling in older FFPE samples. MicroRNAs are small non-coding sequences of RNA, ~20–23 nucleotides long, which are involved in the regulation of countless genes. Similar to gene expression, miRNA expression appears to be tightly controlled and provides highly specific biomarkers for cancer (Lu et al, 2005; Lebanony et al, 2009). MicroRNAs appear to have enhanced stability in both plasma (Mitchell et al, 2008) and FFPE samples (Li et al, 2007; Hui et al, 2010). Importantly, we show that miRNAs are not subjected to the same deterioration seen in other RNA types, have robust expression regardless of sample age and can stratify patient samples with negligible miRNA signal.

MATERIALS AND METHODS

Patients and tissue

Samples of histologically proven transitional cell carcinoma of the bladder (stage T2, T3 or T4a) or high-grade non-muscle invasive carcinoma (T1 grade 3) from patients participating in the BCONF1 study (2000–2002) from the Christie Hospital. The samples were collected and processed at a single hospital site, using the same methods of fixation and embedding used today, that is, the same methods of fixation and embedding used today, although changes in the quality and composition of reagents cannot be excluded. Tumour biopsies were obtained under general anaesthetic and immersed in 4% neutral-buffered formalin. The FFPE blocks were stored at room temperature in a standard block storage unit. Unfortunately in the clinical setting other features associated with fixation were not recorded (e.g., duration of fixation, tumour volume, etc.); however, all samples were processed routinely and therefore are representative of the types of cohort available for FFPE research. All tumours contained ≥ 10% tumour material. Use of the material was approved by the local ethics committee (LREC 09/H1013/24).

Samples of histologically proven carcinoma of the cervix (FIGO stage Ib–IVa) from patients treated with radiotherapy alone with curative intent were selected for Exon array hybridisation (n = 141). Pre-treatment FFPE biopsies were obtained between November 2000 and April 2006 from 11 UK hospitals; sample processing was carried out according to local standard operating procedures. All tumours contained ≥ 10% tumour material. Use of the material was approved by the local ethics committee (LREC 09/H1013/24). Samples of histologically proven transitional cell carcinoma of the bladder (stage T2, T3 or T4a) or high-grade non-muscle invasive carcinoma (T1 grade 3) from patients participating in the BCONF1 study (2000–2002) from the Christie Hospital. The samples were collected and processed at a single hospital site, using the same standard operating procedure over the period they were obtained, that is, the same methods of fixation and embedding used today, although changes in the quality and composition of reagents cannot be excluded. Tumour biopsies were obtained under general anaesthetic and immersed in 4% neutral-buffered formalin. The FFPE blocks were stored at room temperature in a standard block storage unit. Unfortunately in the clinical setting other features associated with fixation were not recorded (e.g., duration of fixation, tumour volume, etc.); however, all samples were processed routinely and therefore are representative of the types of cohort available for FFPE research. All tumours contained ≥ 30% tumour material. Local ethical approval was obtained for using the human material (LREC: 08/H1011/63).

RNA extraction and quality control (QC)

RNA was extracted and DNase treated using RecoverAll Total Nucleic Acid Isolation Kit (Ambion, Austin, TX, USA), as per manufacturer’s instructions. We have previously shown that this kit was sufficiently optimised for sarcoma FFPE samples without the need for additional modification (Linton et al, 2009). RNA integrity and RNA quantification were measured using a Bioanlyser (Agilent Technologies Ltd, Santa Clara, CA, USA). Ratios of 260/230 and 260/280 were assessed using a Nanodrop 1000 Spectrophotometer (Thermo Scientific, Wilmington, DE, USA). Nanodrop-quantified cDNA yield was also recorded. Minimum requirements for hybridisation to Exon arrays were an input of 100 ng total RNA that amplifies to give a yield of ≥ 3.8 µg of cDNA. In two cervix carcinoma cases (V153 and V247), 50 ng of RNA was used as input, owing to a low RNA yield. As such, 100% of cases had usable amounts of RNA. Samples that were diluted (< 50 ng µl–1) were concentrated under vacuum (Eppendorf concentrator 5301; Eppendorf, Hamburg, Germany). The standard RNA quality parameters: RIN (a measure of the proportion of intact ribosomal RNA (rRNA)), 260/230 ratio and 260/280 ratio were recorded, but not used to screen samples. Full RNA QC data were available for 89% (125/141) of bladder and 87% (139/160) of cervix cancer samples hybridised to Exon arrays. The remaining samples had one or more parameters where data were not recorded.

Exon array hybridisation

One hundred nanograms of RNA was amplified using NuGen WT-Ovation FFPE v2 kit (NuGen Technologies, San Carlos, CA, USA). The WT-Ovation Exon Module V1.0 was used to generate ST-cDNA, and 3.8–4 µg was hybridised to Human Exon 1.0 ST arrays (Affymetrix, Santa Clara, CA, USA). Further details and raw data (CEL files) are available at http://bioinformatics.picr.man.ac.uk/vice (or GSE39067).

Exon array data analysis

Microarray data were normalised using RMA (Irizarry et al, 2003). R/BioConductor package anmap and anmap database (Yates et al, 2008) were used to filter non-exonic and multitranscripts probesets. Array performance was measured as the percentage of probesets flagged as ‘present’ with a conservative cutoff (%detected above background (%DABG) P < 0.01) and only those probesets ‘present’ in at least three samples were analysed. Gene level summaries were calculated by taking the median signal of filtered probesets that mapped to unique gene symbols. Affymetrix Exon 1.0ST array hybridisation reproducibility was assessed using standard Affymetrix Exon array QC measures (Supplementary Text and Supplementary Figures 1–3). Samples with more than two QC measures ± 10% of the cohort mean were considered outliers. Principle component analysis was also used to assess sample variation within cohorts. Supplementary Figures 1–3 show that no samples failed this QC check and so none were excluded from the analysis. An adenocarcinoma/squamous cell carcinoma (AC/SCC) ratio was derived by using a previously published signature comprising 2673 probesets (Hall et al, 2011). The median expression values of 2393 SCC probesets were divided by that of 278 AC probesets. Scores > 1 were classified as SCC and scores < 1 as AC. The same ratio was calculated for 100 sets of 2673 randomly sampled probesets for comparison.

TaqMan miRNA qRT–PCR

Ten nanograms of RNA was reverse transcribed using the TaqMan MicroRNA Reverse Transcription (RT) Kit (Applied Biosystems, Carlsbad, CA, USA) and pooled Taqman RT primers for hsa-miR-205 (000509), hsa-miR-26b (000407) and hsa-miR-16 (003911). Complementary DNA was amplified using the TaqMan PreAmp Master Mix, as per the manufacturer’s protocol, for 14 cycles (Applied Biosystems). Five microlitres of amplified cDNA was subjected to quantitative PCR using the same TaqMan primers and probes used in the RT stage. Standard Gene Expression Mastermix (Applied Biosystems) and thermocycling conditions were utilised, data were collected using an AB7900 and values exported from SDS2.1 for analysis in Excel. Relative quantification was performed

© 2012 Cancer Research UK

British Journal of Cancer (2012) 107(4), 684 – 694
using the 2−ΔΔCT method (Livak and Schmittgen, 2001), including normalisation of the target expression data (hsa-miR-205) to the mean of housekeeping miRNA expression (hsa-miR-26b and hsa-miR-16) (Pelotier and Latham, 2008).

MicroRNA array hybridisation

One hundred nanograms of total RNA was tailed and ligated to FlashTag-Blotin-HSR as per the manufacturer’s protocol (Affymetrix). The ligated RNA was then hybridised to Affymetrix miRNA v2.0 arrays for 16 h at 48 °C. Further details and raw data (CEL files) are available at http://bioinformatics.picr.man.ac.uk/vice or GSE39067. Raw data were analysed using the Affymetrix ‘miRNA QC tool’ (v1.1.1.0), using standard parameters including quantile normalisation. The Affymetrix mirNA v2.0 array contains probesets for multiple species, and different small-non-coding RNA molecules such as stem loops (pre-miRNA precursors), and small nucleolar RNA (snoRNA) and small Cajal body-specific RNA (scaRNAs). Probesets were filtered to retain only those annotated for Homo sapiens and miRNA (mature, processed). SnoRNA (including CDBox and HACABox) were also considered, but independently to mature miRNAs. Pre-miRNA data from the miRNA v2.0 array were not considered owing to consistently low expression in FFPE and cell line RNA (Supplementary Figure 4A and B). Therefore, pre-miRNA data for miR205 were derived from the Exon array. Differential expression analysis of filtered probesets was performed using LIMMA (Smyth, 2004).

Cell lines and western data

Cervix cell lines were grown in DMEM + 10% foetal calf serum. Western blotting was performed using the following antibodies: p63 mouse monoclonal (BC4A4) (Abcam, Cambridge, UK) and alpha-tubulin mouse monoclonal (020M4753) (Sigma-Aldrich, Dorset, UK). Protein expression rather than pathological assignment was used to broadly classify lines as SCC (p63 + ) and AC (p63 − ).

p63 immunohistochemistry

Sections (4 μm) were dewaxed, rehydrated and the antigen retrieved by microwaving in Low pH Antigen Unmasking Solution (Vector Laboratories Inc., Peterborough, UK). After quenching endogenous peroxidase, nonspecific binding was blocked using 10% casein (Vector Laboratories Inc.). The primary antibody, mouse monoclonal (BC4A4) (Abcam), was applied at 10 μg ml−1, and the sections incubated at 4 °C in a humidified chamber overnight. The same concentration of mouse IgG2a control reagent (Dako Ltd, Ely, UK) was used as negative control. The antigen was detected with mouse EnVision Plus reagent (Dako Ltd) and visualised with 3,3’-diaminobenzidine (Dako Ltd). Sections were counterstained with haematoxylin, dehydrated and mounted. Batch-to-batch variation was assessed by running sections showing high and low p63 expression with each batch. Tumours exhibiting ≥5% positive nuclei were classified as p63 positive (Cho et al., 2003). Immunohistochemistry methods followed REMARK guidelines (McShane et al., 2005).

RESULTS

RNA quality affects the technical success of Exon array profiling

Optimisation of RNA extraction from FFPE is an important consideration and should take into account length of fixation and duration of sample storage (Ribeiro-Silva et al., 2007; Chung and Hewitt, 2010). To ensure that the RNA extraction protocol was optimised for our older cervix samples, we performed a pilot experiment on four samples stored 2–14 years (cohort median age: 13.9 years). Good yields (3–19 μg per 60 μm tissue) were obtained using the standard RecoverAll protocol and %DABG scores similar to cell line RNA (data not shown). Following the pilot experiments, remaining samples were hybridised. Complete RNA QC data were available for 125 bladder and 139 cervix samples. Figure 1A shows that neither RNA yield nor RIN correlated with %DABG, which measures array performance (Linton et al., 2008; Trabzuni et al., 2011). The average RIN of both the bladder and cervix samples was ~2.3 (Figure 1B), showing considerable rRNA degradation compared with cell line RNA (Figure 1C). The data show that rRNA integrity is a poor predictor of successful archival FFPE array profiling. The standard spectrophotometric determinants of RNA purity (260/230 and 260/280 ratio) also failed to correlate with %DABG (Figure 1A). Complementary DNA yield following NuGEN amplification correlated weakly with %DABG in both the bladder (R = 0.19) and cervix (R = 0.39) cohorts (P < 0.05 for both). Of note, yield might be a better predictor than suggested by these data, as samples yielding <3.8 μg cdNA cannot be hybridised to arrays. However, cdNA yield accounts for only a small amount of sample variation.

FFPE sample age is the predominant feature associated with poor array performance

The range of %DABG values was lower in the bladder (median 24.0% (range 12.6–34.0%)) than the cervix (median 19.2%, (range 4.4–40.9%)) samples. It is possible that tissue type contributes to the difference, but the bladder samples were younger (median age 6 years (6–8)) than the cervix cohort (median age 13 years (8–23)). Therefore, we examined whether FFPE block age at RNA extraction contributed to array performance. There was a statistically significant trend for %DABG to decrease with sample age in both the bladder (R = −0.30, P < 0.01; Figure 1D) and cervix (R = −0.69, P < 0.01; Figure 1E) series. This clearly demonstrates in two independent cohorts that the older an FFPE sample the poorer the array performance. The effect size was almost negligible in the younger bladder cancer cohort, but coupled with the cervix data it suggests that over time the %DABG might similarly decay. Neither RNA yield nor RIN correlated with the age of FFPE block (Supplementary Figures 5 and 6).

Biological signal is lost in older samples

%DABG measures technical success, but does not reflect ‘biological signal’ success. To address this, we investigated whether a gene signature discriminated AC and SCC in older FFPE samples. To remove any tissue of origin concerns we focused on the cervix samples, which had the broadest age range. The cervix samples were divided into two series based on historical enrolment into two different studies. Cervix series 1 (CS1) comprised 112 FFPE samples with a median age of 12 years (range 8–18), and cervix series 2 (CS2) contained 48 FFPE samples with a median age of 18 years (range 16–23) (Table 1). We have previously published a gene signature that is capable of separating cervix into AC and SCC histology, based on the expression of 1062 SCC and 155 AC genes (Hall et al., 2011). This signature was derived on 28 cervix FFPE samples, but most importantly shown to independently validate in

The range of %DABG values was lower in the bladder (median 24.0% (range 12.6–34.0%)) than the cervix (median 19.2%, (range 4.4–40.9%)) samples. It is possible that tissue type contributes to the difference, but the bladder samples were younger (median age 6 years (6–8)) than the cervix cohort (median age 13 years (8–23)). Therefore, we examined whether FFPE block age at RNA extraction contributed to array performance. There was a statistically significant trend for %DABG to decrease with sample age in both the bladder (R = −0.30, P < 0.01; Figure 1D) and cervix (R = −0.69, P < 0.01; Figure 1E) series. This clearly demonstrates in two independent cohorts that the older an FFPE sample the poorer the array performance. The effect size was almost negligible in the younger bladder cancer cohort, but coupled with the cervix data it suggests that over time the %DABG might similarly decay. Neither RNA yield nor RIN correlated with the age of FFPE block (Supplementary Figures 5 and 6).

Biological signal is lost in older samples

%DABG measures technical success, but does not reflect ‘biological signal’ success. To address this, we investigated whether a gene signature discriminated AC and SCC in older FFPE samples. To remove any tissue of origin concerns we focused on the cervix samples, which had the broadest age range. The cervix samples were divided into two series based on historical enrolment into two different studies. Cervix series 1 (CS1) comprised 112 FFPE samples with a median age of 12 years (range 8–18), and cervix series 2 (CS2) contained 48 FFPE samples with a median age of 18 years (range 16–23) (Table 1). We have previously published a gene signature that is capable of separating cervix into AC and SCC histology, based on the expression of 1062 SCC and 155 AC genes (Hall et al., 2011). This signature was derived on 28 cervix FFPE samples, but most importantly shown to independently validate in
Figure 1 RNA quality affects the technical success of Affymetrix Exon arrays. (A) Table displaying the correlation coefficients (R) between RNA quantification and QCs against %DABG; RIN number, 260/230 ratio, 260/280 ratio and concentration. Complementary DNA yield following NuGen amplification was also considered as a surrogate for RNA quality. Bolded values show significant P-values: *P<0.05, **P<0.01. (B) Table displaying the median RNA quantification and QC values for the bladder cancer cohort (n = 125) and the combined cervix cancer cohort (n = 139). Brackets indicate range. (C) Bioanalyser traces showing intact cell line RNA (MCF10A) with an RIN number of 10 and a representative cervix FFPE sample (V554) with a RIN of 2.10. Fragment size is on the x axis (migration time in seconds), with abundance (fluorescent units) shown on the y axis. Formalin-fixed paraffin-embodied samples show loss of the 18S and 28S ribosomal subunit peaks seen in the cell line RNA. (D) x-y scatterplot showing %DABG against age of FFPE block (in years) for the bladder cancer cohort. The plotted line represents a line of best fit. (E) x-y scatterplot showing %DABG against age of FFPE block (in years) for the cervix cancer cohort. The plotted line represents a line of best fit.

Table I FFPE cohort information

| Cohort          | Tumour type | Number of patients | Median age of sample (years) | Median %DABG |
|-----------------|-------------|--------------------|-----------------------------|--------------|
| Bladder         | Bladder     | 141                | 6 (3–8)                     | 24.0 (12.6–34.5) |
| Cervix series 1 | Cervix      | 112                | 12 (8–18)                   | 22.6 (8.3–40.9) |
| Cervix series 2 | Cervix      | 48                 | 18 (16–23)                  | 13.9 (4.4–27.6) |
| Combined series | Cervix      | 160                | 13 (8–23)                   | 19.2 (4.4–40.9) |

Abbreviation: %DABG = percentage detection above background. Numbers in brackets represent the range.

The ratio in the AC samples was higher in the lung than cervix, possibly owing to differences in gene expression between the cancer types. Figure 2C shows the age range of the two cervix cohorts and Figure 2D the obvious biological signal in CS1 specimens with only a minority of SCC samples having ratios close to or below background. As with most classifiers, there are some histological misclassifications, but most of the younger samples were correctly classified as SCC. In contrast, the older SCC samples show a clear reduction in performance with most of the SCC samples having ratios close to or below background (Figure 2E). None of the FFPE samples had a ratio above 1.2, even though ~90% of cervix cancers are SCC. Taken together, these data demonstrate that the older FFPE samples have lost biological signal. We also saw a weak trend for decreasing housekeeper expression with sample age in our data; however, there was considerable sample variation (Supplementary Figures 7 and 8). Interestingly, there was no significant difference in housekeeping gene expression or distribution associated with %DABG filtering or between the older and newer samples.

mRNA transcript expression is progressively lost during long-term FFPE storage

Biological signal progressively decreased during storage as shown by the correlation between SCC/AC ratio and sample age (R = 0.76, Figure 3A). Furthermore, the strong correlation between SCC/AC...
ratio and %DABG (Figure 3B) demonstrates that this is a general feature of the data: FFPE sample storage is associated with a progressive drop in RNA quality, resulting in a general decline in the 'biological' signal to noise ratio. This is further demonstrated by the decrease in probeset signal for TP63, which encodes the tumour suppressor p63, a protein expressed predominantly in cells and tumours of squamous cell origin. Expression of the p63 protein, measured by immunohistochemistry, is currently considered to be the best single biomarker of cervical SCC (McCluggage, 2007). Like the SCC/AC ratio, Exon array-measured TP63 transcript abundance decreased with cervix sample age ($R = -0.63$, Figure 3C). As a positive control for RNA quality, and to confirm the relationship between p63 protein and transcript expression, p63 levels were measured in 16 cervix cancer cell lines by western blot (Figure 3D). The median gene expression level for TP63 was higher in p63-positive vs p63-negative cell lines (Figure 3E). While the background level of TP63 signal was relatively constant (green boxes) across cell lines and tumour cohorts, there is a decrease in the TP63 signal for SCC that is proportionate to RNA quality (hatched boxes). In the older cervix series this eliminated the difference between AC and SCC. These results are further supported by immunohistochemistry data for 152 samples, where median TP63 expression in the older samples was closer to background (p63 negative) levels (Figure 3F). As these data represent the median values for multiple probesets targeting the length of the gene, it is probable that the RNA degrades along the entire length of the TP63 transcript.
Mature miRNA expression can overcome mRNA degradation and accurately classify cervix tumours

MicroRNAs were explored for their ability to overcome the limitations of mRNA profiling. The miRNA miR-205 has previously been demonstrated as a marker of lung SCC (Lebanony et al., 2009). Two Exon array probesets target the full-length hsa-mir-205 (mir-205) precursor. Both probesets correlate with TP63 expression (Figure 4A) in our previously published cohort (Hall et al., 2011). We interrogated whether expression of these probesets correctly classified samples according to histology. The expression of both mir-205 probesets was similar to TP63 (Figure 4B). While the probe expression in SCC samples (hatched boxes) was higher in the cell lines and younger FFPE samples ($P<0.01$), there was no statistically significant difference between SCC and AC in the older samples. MicroRNA probesets on Exon arrays target precursor miRNAs (pre-miRNA) sequences, which are longer, in this case 110 nt, than the mature processed miRNAs (22 nt). The implication is that pre-miRNAs are subject to similar degradation effects as mRNAs. To address this, a subset of samples was randomly selected from the cervix cohort, to encompass all ages and RNA qualities (Supplementary Table 1). A qRT–PCR assay was designed to specifically detect the mature form of miR-205 and performed on the same RNA used to generate the array data. Both mir-205 probesets, and the median expression of TP63, failed to discriminate between AC and SCC in the older FFPE cohort (Figure 4C). However, signal for the mature miR-205 was significantly higher in SCC compared with AC across all RNA samples.
There was no significant age-related decay in signal between the CS1 and CS2 subsets. The CS1 samples showed a 7.6-fold difference in expression of miR-205 between AC and SCC, and the CS2 subset showed a 6.9-fold difference (Supplementary Figure S9). This confirms that mature miR-205 expression discriminates between AC and SCC in RNA with no discernable mRNA signal.

Global miRNA profiling confirms that miRNAs have enhanced stability in FFPE samples

To test whether this result generalised to global miRNA profiling, 16 FFPE samples and 2 cell lines were hybridised to Affymetrix miRNA v2.0 arrays. Clinical samples were randomly selected from either cohort of cervix samples. Figure 5A shows the probeset expression of hsa-miR-205 derived from the array data. There is clear separation of SCC and AC in the cervix samples, recapitulating the qRT–PCR data (Figure 4D). Taken together, these data show that miR-205 expression levels support histological discrimination, regardless of sample age or quality, even in FFPE samples where mRNA expression cannot. Figure 5B shows that the median probeset expression of all human miRNAs for eight young cervix samples correlates well with the probeset expression of eight old cervix samples ($R = 0.95$) with no apparent age-related bias. An equivalent plot for a random sampling of a similar sized subset of ($n = 1000$) mRNA (Exon) probesets was skewed towards the higher signal in younger (CS1) samples ($R = 0.88$) (Figure 5C). This shift is more pronounced for the 2395 probesets comprising the SCC component of the AC/SCC signature ($R = 0.72$). Together, these data suggest the miRNA data are more robust to the effects of
Gradient colouring is based on the size of the target molecule, between 48 nt (blue) and 250 nt (red). Pearson correlation coefficient ($R$) Probesets with positive or negative fold change less than twofold (log 2 $R$ median of eight CS2 samples. Pearson correlation coefficient ($y$) SCC-specific probesets derived from the AC/SCC signature (hollowed circles). $y$ CDbox and HAcaBox probes (miRNA 2.0 array).

in the literature that miR-205 is a suitable biomarker for SCC, adjustment ($P$ compared with AC) was statistically significant after FDR expression (Figure 5E), only miR-205 (26-fold higher in SCC vs FFPE processing and storage. We therefore compared miRNA profiles for SCC vs AC samples, irrespective of sample age (Figure 5D). While a number of probesets showed differential expression (Figure 5E), only miR-205 (26-fold higher in SCC compared with AC) was statistically significant after FDR adjustment ($P = 1.103 \times 10^{-5}$). This supports previous reports in the literature that miR-205 is a suitable biomarker for SCC, and for the first time we demonstrate its ability to discriminate cervix cancer histologies. Another small-non-coding class of RNA molecule represented on the Affymetrix miRNA v2.0 array, snoRNAs, shows stability across the two sample cohorts ($R = 0.97, P < 0.0001$) (Figure 5F). Figure 5F shows that this correlation is preserved in probesets detecting the entire range of target molecule sizes.

**Figure 5** Global microRNAome profiling confirms that microRNAs have enhanced stability in FFPE samples. (A) Affymetrix miRNA 2.0 array data for probeset hsa_miR-205_st. Each point represents an independent sample and the horizontal bar indicates the median. Eight samples were hybridised from the younger CS1 cohort (four AC and four SCC) and eight samples were hybridised from the older CS2 cohort (four AC and four SCC) along with a single example of SCC and AC cell line RNA. (B) $x$-$y$ scatterplot showing the miRNA (miRNA 2.0 array) median probeset expression values for all 1105 miRNA probes. $y$-axis values are the median of eight CS1 samples and $x$-axis values are the median of eight CS2 samples. Pearson correlation coefficient ($R$) is displayed. (C) $x$-$y$ scatterplot showing the mRNA (Exon array)-derived median probeset expression for 1000 randomly selected probesets (grey) and 2295 SCC-specific probesets derived from the AC/SCC signature (hollowed circles). $y$-axis values are the median of eight CS1 samples and $x$-axis values are the median of eight CS2 samples. Pearson correlation coefficient ($R$) is displayed. (D) Volcano plot showing the results of LIMMA differential expression analysis between eight SCC and eight AC samples. $x$ axis represents fold change (log2) and the $y$ axis details LIMMA odds ratio, a measure of statistical change. Probesets with positive or negative fold change less than twofold (log2 = 1) are coloured grey. Probesets with a fold change greater than twofold are coloured black. Star-shaped data points represent statistically significant probesets with an FDR-adjusted $P$-value < 0.05. (E) Table listing differentially expressed probesets with a log2 fold change > 1.5 fold. (F) $x$-$y$ scatterplot showing the snoRNA median probeset expression values for all 310 snoRNA, CDbox and HAcaBox probes (miRNA 2.0 array). $y$-axis values are the median of eight CS1 samples and $x$-axis values are the median of eight CS2 samples. Gradient colouring is based on the size of the target molecule, between 48 nt (blue) and 250 nt (red). Pearson correlation coefficient ($R$) is displayed.
DISCUSSION

Specimen size, time to fixation and fixation duration affect the quality of RNA from FFPE archival samples (von Ahlfen et al., 2007). There is also a progressive deterioration of RNA quality associated with long-term storage. The lack of standardisation in tissue processing and storage in histopathology departments means that it is not possible to access samples handled uniformly. Given the vast archives of FFPE tumour material available, techniques are required that cope with a high degree of sample variation owing to processing and degradation. Approaches are also needed that predict whether samples can be successfully profiled with RNA expression techniques. This study investigated whether an approach that worked with samples stored for 10–16 years (Hall et al., 2011) could cope with the systematic degradation of RNA in even older FFPE samples (Cronin et al., 2004). We envisaged that the short 25-mer probes on the Exon arrays would deal with RNA fragmentation better than qRT–PCR and consequently we performed no pre-selection based on standard RNA quality metrics, or PCR expression of endogenous controls (Ribeiro-Silva et al., 2007; Reinholz et al., 2010; Waddell et al., 2010).

Standard RNA QC parameters did not predict microarray profiling success and there was only a very weak correlation for decreasing reference gene expression in older samples. The level of rRNA degradation, as measured by RIN, was not a predictor of FFPE microarray profiling success. The lack of correlation between RIN and %DABG agrees with previous studies demonstrating that 1-year-old FFPE expression, as measured by qRT–PCR, is comparable to that of fresh-frozen material, providing small amplicons are used (Cronin et al., 2004; von Ahlfen et al., 2007). This is further supported by a recent study of QC parameters in unfixed postmortem brain specimens, in which only 2.7% of the variation in array performance was explained by RIN (range 1–8.5) (Trabzuni et al., 2011). This suggests that the RNA deterioration that contributes to poor RIN can be overcome by the combined use of RNA amplification, that is not solely reliant on poly(A)-based priming, and multiple short reporters, such as the probesets found on Exon arrays.

In both the bladder and cervix cancer FFPE samples, a statistically significant association was seen between decreasing array performance (%DABG) and sample age, suggesting that during prolonged storage the RNA becomes increasingly incompatible with current methods. In the cervix cohort, we showed that samples with low %DABG also had low biological signal and it was not possible to distinguish between AC and SCC (AC/SCC ratio or TP63 transcript). As FFPE samples age, both total RNA quality and signal (%DABG) decreased and biological signal was lost. This finding supports the recommendation that FFPE samples should have RNA extracted as soon as possible after fixation, preferably within 1 year (von Ahlfen et al., 2007). With prospective translational studies or clinical tests this is sometimes achievable, but adhering to this recommendation would clearly remove many potentially valuable retrospective FFPE samples from further analysis. Our study is the first to show these trends in global mRNA profiles. Given that we have quantitatively validated Exon array FFPE data previously (Hall et al., 2011), the loss of biological signal in samples stored for 16–23 years should recapitulate with other methods for measuring RNA expression such as qRT–PCR or Quantigene. Other approaches were not explored owing to limited material available but loss of expression with age has been reported previously using qRT–PCR (Cronin et al., 2004). Our findings, therefore, are likely to generalise to other FFPE profiling methods given the similarity of approach (e.g., array designs featuring 3 reporters). While adjustments to data analysis and normalisation methods, such as normalising to invariant probesets/housekeeping genes (Cronin et al., 2004; Kennedy et al., 2011), might help correct for age-related effects, it is unreasonable to expect them to be able to reconstruct information that is missing in the original data. Formalin-fixed paraffin-embedded cohorts will inherently contain significant amounts of technical and biological variation, leading to the need for larger cohorts than an equivalent study using fresh-frozen material. There are currently no prospective QC metrics that can universally predict array performance. In this study, the strongest indicator of array success was sample age, therefore younger aged cohorts should be exploited where possible.

There is increasing evidence in the literature that miRNAs show enhanced stability in both plasma (Mitchell et al., 2008) and FFPE (Li et al., 2007; Hui et al., 2010). The miR-205 was investigated as a biomarker of SCC (Lebanony et al., 2009). The precursor transcript behaved in a similar way to miRNA and also failed to discriminate between AC and SCC in older samples. However, mature miR-205 was able to distinguish cervical samples by histological subtype. It was considered that this might be owing to the high expression of miR-205 (i.e., higher levels of transcript to degrade, compared with mRNA markers). As expected in the mRNA data, there was a significant decrease in signal in the older samples compared with the newer FFPE samples. This decrease was not observed when comparing the expression of human mature miRNAs (n = 1105) suggesting that the enhanced stability associated with miR-205 is a general feature of all detected miRNAs. Given the diverse roles of miRNAs and their enhanced stability independent of expression level, there is great potential for miRNA profiling to contribute to the development of clinical biomarkers (Lu et al., 2005; Lebanny et al., 2009) and the generation of signatures associated with outcome (Hui et al., 2010; De Preter et al., 2011) especially using FFPE samples.

There are multiple hypotheses concerning why miRNA is more stable than mRNA in FFPE. First, either the lack of structure, lack of a specific nucleotide target sequence, their small size or some other protective modification prevents the degradation that occurs with larger RNA molecules. Second, the miRNAs are in a protective environment such as the RNA-induced silencing complex, that by the virtue of formalin crosslinks are tethered to the miRNA or alternatively, because they are packaged into vesicles known as exosomes (Lee et al., 2009). It may also be that miRNAs provide a better template for amplification or have hybridisation kinetics that are less affected by the processes occurring in FFPE. It is interesting therefore that snoRNAs, a different class of small non-coding RNA, with their significantly larger size (48–250 nt) and different subcellular location also demonstrate enhanced stability in FFPE. SnoRNAs primarily guide chemical modifications of other RNA molecules (Mattaj et al., 1993). SnoRNAs have been associated with genetic conditions such as Prader–Willi syndrome (Cassidy et al., 2012) and dyskeratosis congenita (Mason and Bessler, 2011). It is becoming clear that snoRNAs also regulate other biological processes such as alternative splicing (Kishore and Stamm, 2006) and even function as non-classical miRNAs (Ender et al., 2008). Understanding the commonalities between snoRNAs and miRNAs will provide further insight into why some RNA molecules degrade less than others when formalin-fixed and paraffin-embedded.

Other studies have reported improvements in expression profiling by using miRNAs in conjunction with FFPE. This is, however, based on the best of our knowledge to show that archival FFPE RNA, that is severely degraded and empirically incompatible with mRNA expression methods, can yield meaningful biology when combined with miRNA expression profiling. Our study therefore provides strong support for the use of miRNA profiling in archival FFPE samples especially when an investigation requires the analysis of older, more recalcitrant material.

ACKNOWLEDGEMENTS

We would like to acknowledge Mark Wappett, Jodie Whittaker and Stuart Pepper from the CR-UK Molecular Biology Core Facility at the Paterson Institute for assistance with the project. We would
also like to thank Dr Kanwal Sikand for pathology advice. JSH is funded by a Cancer Research UK (CR-UK) project grant (C1094/ A9437) and the Manchester Cancer Research Centre (MRcC) (C1467/A7286). AE is funded by an MRC Biomarker grant (G0801525), and JT and CJM are funded by CR-UK (C147/ A6058). HRV is funded by UK Ecmc. the UK medical research council also contributed towards reagent costs (G0801525).

Conflict of interest

The authors declare no conflict of interest.

References

Abdueva D, Wing M, Schaub B, Triche T, Davicioni E (2010) Quantitative expression profiling in formalin-fixed paraffin-embedded samples by affymetrix microarrays. J Mol Diag 12: 409–417
Abramovitz M, Barwick BG, Willis S, Catzavelos C, Li Z, Kodani Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 57: 289–300
Blow N (2007) Tissue preparation: tissue issues. Nature 448: 959–963
Cassidy SB, Schwartz S, Miller JL, Driscoll DJ (2012) Prader-Willi syndrome. Genet Med 14: 10–26
Chaturvedi AK, Engels EA, Pfeiffer RM, Hernandez BY, Xiao W, Kim E, Cho NH, Kim YB, Park TK, Kim GE, Park K, Song KJ (2003) P63 and EGFR based platform. FASEB J 17: 35–42
Chung JY, Hewitt SM (2010) An optimized RNA extraction method from archival formalin-fixed paraffin-embedded tissue. Methods Mol Biol 611: 19–27
Cronin T, Pho M, Dutta D, Stephens JC, Shk S, Kiefer MC, Esteban JM, Baker JB (2004) 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 164: 35–42
De Preter K, Mestdagh P, Vermeulen J, Zeka F, Naranjo A, Bray I, Castel V, Ender C, Krek A, Friedlander MR, Beitzinger M, Weinmann L, Chen W, Mattaj IW, Tollervey D, Seraphin B (1993) Small nuclear RNAs in mammalian cells. Nature 360: 587–596
Ferrando AA, Downing JR, Jacks T, Horvitz HR, Golub TR (2005) A human snoRNA with a potential role in neurodegeneration. Nat Cell Biol 7: 635–645
Hyland J, Dhir R, Bathe OF, Winqvist O, Manne U, Shanmugam C, Vessella RL, Nelson PS, Martin DB, Tewari M (2008) Circulating microRNAs in serum of patients with prostate cancer. J Urol 180: 2030–2037
Lee J, Hever A, Willhite D, Zlotnik A, Hevezi P (2005) Effects of RNA degradation on gene expression analysis of human postmortem tissues. FASEB J 19: 1356–1368
Lee YS, Pressman S, Andrews AP, Kim K, White JI, Cassidy JY, Li X, Lubell K, Lim do H, Cho IS, Nakahara K, Preall JB, Bellare P, Sontheimer EJ, Carthew RW (2009) Silencing by small RNAs is linked to endosomal trafficking. Nat Cell Biol 11: 1150–1156
Linton KM, Hey Y, Dibben S, Miller C, Freemont A, Radford J, Pepper S (2009) Methods comparison for high-resolution transcriptional analysis of archival material on Affymetrix Plus 2.0 and exon 1.0 microarrays. Biotechniques 47: 587–596
Linton K, Hey Y, Dibben S, Miller C, Freemont A, Radford J, Pepper S (2009) Methods comparison for high-resolution transcriptional analysis of archival material on Affymetrix Plus 2.0 and exon 1.0 microarrays. Biotechniques 47: 587–596
Mason PJ, Bessler M (2011) The genetics of dyskeratosis congenita. Genet Med 13: 971–981
Mitchell PS, Parkin RK, Kroh EM, Black JM, Kay EW, Holt RJ, Proutski V, Ahdesmaki M, Farzidinov V, Goffard N, Hey F, McDuarry P, Mulligan K, Mussen J, O’Brien E, Oliver G, Walker SM, Mulligan JM, Wilson C, Winter A, O’Donoghue D, Mulcahy H, O’Sullivan J, Sheahan K, Hyland J, Dhir R, Bathe OF, Winqvist O, Manne U, Shannumugam C, Ramaswamy S, Leonard J, Smith JR, McDermott U, Wilson RH, Longley D, Marshall J, Cummings R, Sargent D, Johnston PG, Harkin DP (2011) Development and independent validation of a prognostic assay for stage II colon cancer using formalin-fixed paraffin-embedded tissue. J Clin Oncol 29: 4620–4626
Chajut A, Cohen D, Aharonov R, Mansukhani M (2009) Diagnostic assay based on hsa-miR-205 expression distinguishes squamous from non-squamous non-small-cell lung carcinoma. J Clin Oncol 27: 2030–2037
Lee J, Hever A, Willhite D, Zlotnik A, Hevezi P (2005) Effects of RNA degradation on gene expression analysis of human postmortem tissues. FASEB J 19: 1356–1358
Lebannon D, Benjamin H, Gilad S, Ezagouri M, Dow A, Ashkenazi K, Gfen N, Izraeli S, Rechavi G, Pass H, Nonaka D, Li J, Spector Y, Rosenfeld N, Chajut A, Cohen D, Abarorov R, Manukhanni M (2009) Diagnostic assay based on hsa-miR-205 expression distinguishes squamous from non-squamous non-small-cell lung carcinoma. J Clin Oncol 27: 2030–2037
Lee J, Hever A, Willhite D, Zlotnik A, Hevezi P (2005) Effects of RNA degradation on gene expression analysis of human postmortem tissues. FASEB J 19: 1356–1358
Lee YS, Pressman S, Andrews AP, Kim K, White JI, Cassidy JY, Li X, Lubell K, Lim do H, Cho IS, Nakahara K, Preall JB, Bellare P, Sontheimer EJ, Carthew RW (2009) Silencing by small RNAs is linked to endosomal trafficking. Nat Cell Biol 11: 1150–1156
Li J, Smyth P, Flavin R, Cahill S, Denning K, Aherne S, Guenther SM, O’Leary J, Shells O (2007) Comparison of miRNA expression patterns using total RNA extracted from matched samples of formalin-fixed paraffin-embedded (FFPE) cells and snap frozen cells. BMC Biotechnol 7: 36
Linton K, Hey Y, Dibben S, Miller C, Freemont A, Radford J, Pepper S (2009) Methods comparison for high-resolution transcriptional analysis of archival material on Affymetrix Plus 2.0 and Exon 1.0 microarrays. Biotechniques 47: 587–596
Linton KM, Hey Y, Saunders E, Jeziorowska M, Denton J, Wilson CL, Swindell R, Dibben S, Miller CJ, Pepper SD, Radford JD, Freemont AJ (2008) Acquisition of biologically relevant gene expression data by Affymetrix microarray analysis of archival formalin-fixed paraffin-embedded tumours. Br J Cancer 98: 1403–1414
Ferrando AA, Downing JR, Jacks T, Horvitz HR, Golub TR (2005) MicroRNA expression profiles classify human cancers. Nature 435: 843–848
Mason PJ, Bessler M (2011) The genetics of dyskeratosis congenita. Cancer Genet 204: 635–645
Mitchael J, Tollervey D, Seraphin B (1993) Small nuclear RNAs in messenger RNA and ribosomal RNA processing. FASEB J 7: 47–53
Clugcake WG (2007) Immunohistochemistry as a diagnostic aid in central nervous system tumors. Pathology 39: 97–111
McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM (2005) Reporting recommendations for tumour MARKer prognostic studies (REMARK). Eur J Cancer 41: 1690–1696
Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O’Briant KC, Allen A, Lin DW, Urban N, Drescher CW, Knudsen BS, Stirewal DT, Gentleman R, Vessella RL, Nelson PS, Martin DR, Tewari M (2008) Circulating microRNA signal from degraded clinical FFPE specimens

Author contributions

JSH managed the project, performed, analysed and interpreted experiments. JT and JSH analysed and interpreted the data. HRV performed immunohistochemistry, and JJI extracted RNA and recorded quality parameters. AE provided the bladder cancer dataset, generated from samples provided by PJH. CJM oversaw bioinformatics analysis and CMLW supervised the experimental studies. JSH wrote the paper with feedback from all authors.

Supplementary Information accompanies the paper on British Journal of Cancer website (http://www.nature.com/bjc)
microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci USA* 105: 10513–10518

Peltier HJ, Latham GJ (2008) Normalization of microRNA expression levels in quantitative RT-PCR assays: identification of suitable reference RNA targets in normal and cancerous human solid tissues. *RNA* 14: 844–852

Penland SK, Keku TO, Torrice C, He X, Krishnamurthy J, Hoadley KA, Woosley JT, Thomas NE, Perou CM, Sandler RS, Sharpless NE (2007) RNA expression analysis of formalin-fixed paraffin-embedded tumors. *Lab Invest* 87: 383–391

Reinholz MM, Eckel-Passow JE, Anderson SK, Asmann YW, Zschunke MA, Oberg AL, McCullough AE, Dueck AC, Chen B, April CS, Wickham-Garcia E, Jenkins RB, Cunningham JM, Jen J, Perez EA, Fan JB, Lingle WL (2010) Expression profiling of formalin-fixed paraffin-embedded primary breast tumors using cancer-specific and whole genome gene panels on the DASL(R) platform. *BMC Med Genomics* 3: 60

Ribeiro-Silva A, Zhang H, Jeffrey SS (2007) RNA extraction from ten year old formalin-fixed paraffin-embedded breast cancer samples: a comparison of column purification and magnetic bead-based technologies. *BMC Mol Biol* 8: 118

Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3: Article3

Trabzuni D, Ryten M, Walker R, Smith C, Imran S, Ramasamy A, Weale ME, Hardy J (2011) Quality control parameters on a large dataset of regionally dissected human control brains for whole genome expression studies. *J Neurochem* 119: 275–282

von Ahlen S, Missel A, Bendrat K, Schlumpberger M (2007) Determinants of RNA quality from FFPE samples. *PloS One* 2: e1261

Waddell N, Coccialesi S, Johnson J, Healey S, Marsh A, Riley J, da Silva L, Vargas AC, Reid L, Simpson PT, Lakhani SR, Chenevix-Trench G (2010) Gene expression profiling of formalin-fixed, paraffin-embedded familial breast tumours using the whole genome-DASL assay. *J Pathol* 221: 452–461

West CM, Davidson SE, Roberts SA, Hunter RD (1997) The independence of intrinsic radiosensitivity as a prognostic factor for patient response to radiotherapy of carcinoma of the cervix. *Br J Cancer* 76: 1184–1190

Yates T, Okoniewski MJ, Miller CJ (2008) XMap: annotation and visualization of genome structure for Affymetrix exon array analysis. *Nucleic Acids Res* 36: D780–D786

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License.