Short Technical Reports

Methods comparison for highresolution transcriptional analysis of archival material on Affymetrix Plus 2.0 and Exon 1.0 microarrays

Kim M. Linton^{1,5*}, Yvonne Hey^{2*}, Sian Dibben², Crispin J. Miller³, Anthony J. Freemont⁴, John A. Radford^{1,5}, and Stuart D. Pepper² ¹Cancer Research UK Department of Medical Oncology, The Christie NHS Foundation Trust, Manchester, UK, ²Molecular Biology Core Facility, Paterson Institute for Cancer Research, The University of Manchester, Manchester, UK, ³Applied Computational Biology and Bioinformatics Group, Paterson Institute for Cancer Research, The University of Manchester, Manchester, UK, ⁴School of Clinical & Laboratory Sciences, The University of Manchester, Manchester, UK, and ⁵School of Cancer and Imaging Sciences, The University of Manchester, Manchester, UK

BioTechniques 47:587-596 (July 2009) doi 10.2144/000113169 Keywords: archival; FFPE; expression microarray; Exon array; Plus 2.0

Supplementary material for this article is available at www.BioTechniques.com/article/113169. *K.L. and Y.H. contributed equally to this work

Microarray gene expression profiling of formalin-fixed paraffin-embedded (FFPE) tissues is a new and evolving technique. This report compares transcript detection rates on Affymetrix U133 Plus 2.0 and Human Exon 1.0 ST GeneChips across several RNA extraction and target labeling protocols, using routinely collected archival FFPE samples. All RNA extraction protocols tested (Ambion-Optimum, Ambion-RecoverAll, and Qiagen-RNeasy FFPE) provided extracts suitable for microarray hybridization. Compared with Affymetrix One-Cycle labeled extracts, NuGEN system protocols utilizing oligo(dT) and random hexamer primers, and cDNA target preparations instead of cRNA, achieved percent present rates up to 55% on Plus 2.0 arrays. Based on two paired-sample analyses, at 90% specificity this equalled an average 30 percentage-point increase (from 50% to 80%) in FFPE transcript sensitivity relative to fresh frozen tissues, which we have assumed to have 100% sensitivity and specificity. The high content of Exon arrays, with multiple probe sets per exon, improved FFPE sensitivity to 92% at 96% specificity, corresponding to an absolute increase of ~600 genes over Plus 2.0 arrays. While larger series are needed to confirm high correspondence between fresh-frozen and FFPE expression patterns, these data suggest that both Plus 2.0 and Exon arrays are suitable platforms for FFPE microarray expression analyses.

Introduction

It is now possible to obtain clinically useful gene expression profiles of formalin-fixed

and paraffin-embedded (FFPE) tumor tissues using a range of microarray platforms, including custom-made two-color spotted arrays and commercial platforms from Illumina (San Diego, CA, USA) (1,2), Affymetrix (Santa Clara, CA, USA) (3,4), and Agilent Technologies (Santa Clara, CA, USA) (5). FFPE is the worldwide standard tissue preservation method; however, the fixation process—as well as time in storage—compromises RNA extraction potential and quality and its consequent utility for molecular biology applications such as microarray gene expression profiling (6,7).

The development of FFPE microarray protocols, therefore, represents a longawaited opportunity to access vast FFPE archives with well-annotated clinical data for the molecular study of human cancers, especially rare tumors where a lack of freshfrozen tissue has prevented microarray studies using standard protocols.

While the success of recently developed FFPE protocols has been met with enthusiasm from the scientific community, several limitations need to be overcome in order for this approach to reach its full potential. To maximally exploit existing FFPE supplies, protocols must be suitable for routinely processed archival FFPE tissues, where widely varying processing and storage conditions can significantly affect RNA degradation and chemical modification rates (8,9) and, in turn, priming for cDNA synthesis and array performance. The custom-developed Arcturus Paradise system (Molecular Devices, Sunnyvale, CA, USA) for FFPE and Agilent 3'-based custom arrays, for example, was useful for interrogating FFPE tissues prepared under optimal conditions [even after several years in storage (5)], but unsuccessful in up to 76% of routinely processed samples (10). One reason may be the use of relatively long 60-mer target probes, as experience with RT-PCR suggests that shorter probe sets are better at detecting extensively degraded RNAs typical of routinely processed FFPE (11). It is therefore likely that shorter microarray probes such as the 25-mer lengths employed in Affymetrix GeneChips will prove to be better for degraded RNAs (7). Furthermore, most custom arrays for FFPE templates have historically featured a preferential 3'-biased probe set design to accommodate degraded RNA and the standard use of oligo(dT) priming methods. However, even with short probe sets, the rationale for using oligo(dT)-only approaches has recently been questioned (3) since (i) this method will not label non-polyA-tailed fragments and (ii) adenine residues have higher rates



Figure 1. Improvements in transcript detection according to RNA extraction method, labeling protocol and microarray type. (A) Comparison of percentage present rates on Plus 2.0 arrays for two RNA extraction protocols (both NuGEN FFPE-labeled). In-house modified Ambion Optimum results were superior to Qiagen RNeasy ($45 \pm 9.4\%$ versus $39 \pm 10.4\%$, P = 0.01). (B) Comparison of percentage present rates on Plus 2.0 arrays for two target labeling protocols (both using modified-Optimumextracted RNAs). NuGEN WT-Ovation FFPE labeling and amplification was superior to Affymetrix One-Cycle labeling ($43 \pm 8.8\%$ versus $24 \pm 5.7\%$, P = 0.00). (C) Comparison of percentage present rates on Exon arrays for two target labeling protocols. NuGEN FFPE Exon (Qiagen RNeasy extracts) was superior to Affymetrix WT Exon (modified-Optimum extracts) (39 ± 3.5% versus $25 \pm 11.9\%$, P = 0.01). (D-F) Venn diagrams showing gene-level detection from paired formalinfixed paraffin-embedded (FFPE) and fresh-frozen tissue samples (numbers shown are the average from two paired-sample comparisons; P = present, A = absent). A 30% improvement in sensitivity [sensitivity = true positive ÷ (true positive + false negative)] is effected by using a combination of random and oligo(dT) primers to label modified-Optimum extracts, thus an average 81% of transcripts detectable in fresh-frozen tissues are also detected in FFPE. Average sensitivity and specificity for the same samples extracted with Qiagen RNeasy were 75% and 92%, respectively (data not shown). Exon array interrogation effects an average improvement in sensitivity to 92.5% in Qiagen RNeasy extracts labeled using the NuGEN FFPE Exon protocol (the same experiment was not performed on modified-Optimum extracts). Specificity [specificity = true negative ÷ (true negative + false positive) is satisfactory with all approaches; the false positive rate is only 5% with Exon arrays.

of formalin modification than other bases (12). Indeed, gene detection from FFPE has been found to be significantly better when random hexamer priming is used (13,14).

In this study, we examined routinely processed archival FFPE and the impact

of different RNA extraction and labeling protocols on final microarray data—using Affymetrix HG U133 Plus 2.0, X3P, and Exon 1.0 ST GeneChips—in order to identify optimal methods for interrogating archival fixed tissues.



Materials and methods

Sections of soft tissue sarcoma (STS) (including six leiomyosarcomas, three liposarcomas, three synovial sarcomas, and one unspecified pleomorphic spindle tumor) were cut from 1-8 year-old routine diagnostic FFPE tissue blocks and processed for RNA extraction as previously described (3). The number of tissue sections required depended on the labeling protocol and tissue cellularity; generally 10-12 sections each containing up to 1 cm² tissue were required to yield sufficient RNA for One-Cycle non-amplification reactions, while 2-3 sections were needed for labeling reactions incorporating an amplification step. In two sarcoma cases, paired unfixed (freshfrozen) tissues were taken from the same biopsy sent for routine FFPE processing, processed as previously described (3) and used as gold standards for measurement of FFPE array performance (sensitivity and specificity of gene detection).

We previously extracted RNA from STS using an in-house modification of the Optimum FFPE RNA Isolation Kit (Ambion Diagnostics, Austin, TX, USA) (3), which is no longer on the market. In this study, we re-extracted RNA from the same samples using the RNeasy FFPE Kit (Qiagen, Hilden, Germany), having previously obtained more consistent extract results with the latter kit than the RecoverAll Total Nucleic Acid Isolation Kit for FFPE (Ambion Diagnostics) (testing was carried out on cervical carcinoma sections; data not shown). RNA extracts were DNase-treated and technical replicates were labeled using Affymetrix protocols (One-Cycle, Two-Cycle, and Whole Transcript for Exon arrays) and NuGEN WT-Ovation systems (Pico, FFPE v2) for Plus 2.0 arrays and FFPE v2 for Exon arrays coupled with NuGEN FL-Ovation cDNA Biotin Module V2 (San Carlos, CA, USA). Samples were in turn hybridized to three Affymetrix microarray platforms for performance comparison: Human Genome U133 Plus 2.0, X3P, and Human Exon 1.0 ST (all from Affymetrix). Manufacturer's instructions were followed throughout and full details of extraction, labeling and hybridization protocols, and raw array data (.cel files) are available at http://bioinformatics.picr.man.ac.uk/vice.

Total RNA yield and purity was estimated by UV spectroscopy (Nanodrop ND-1000 Spectrophotometer; Nanodrop Technologies, Wilmington, DE, USA) and RNA quality was assessed on an Agilent 2100 Bioanalyzer.

Normalization and analysis of Plus 2.0 .cel files were performed using the

RNA extraction (sample)	TRIzol (fresh-frozen)				Ambion Optimum (FFPE)						Qiagen RNeasy (FFPE)	
Array	Plus 2.0		Exon		Plus 2.0				X-3P	Exon	Plus 2.0	Exon
Labeling protocol	One- Cycle	NuGEN FFPE	Affy WT Exon	NuGEN FFPE Exon	One- Cycle	Two- Cycle	NuGEN Pico	NuGEN FFPE	One- Cycle	Affy WT Exon	NuGEN FFPE	NuGEN FFPE Exon
Sample (name in MIAME VICEª)												
KL1 (RD3)					22.2	24		32.5		11.0	33.3	34.3
KL2 (FL4)					22.1	13.2		42.3	24.3		37.9	22.1
KL4 (FL9)					27.1	30.6		52.2			40.3	40.7
KL5 (RD+L11)					27.3		32.9					
KL9 (RD21)					27.6	35.5	37.9			19.7	42.6	37.9
KL14 (p29)					20.9		29.4	37.3				
KL16 (FL33)					32.3	33.8	39.2	50.6	35.9	32.8	42.2	40.9
KL20 (p24b)					22.9		38.2	37.1				
KL33 (RD52)					15.2		30.8					
KL37 (FL42)					17.6					27.5	27.8	
KL38 (p44c)					12		15.3	32.9			19	
KL35 ^b (RD56)					27.3			50.9		36.9	44.4	43.8
KL39 ^b (RD56F)	48.9	57	36.3	59.8								
KL36° (RD58)					27.5		37.2	54.9			52.1	46.9
KL40° (RD58F)	53.5	65.8		65.1								

Table 1. Comparison of Percent Present Results According to RNA Extraction, Array Type and Labeling Protocol

Numbers shown are percentage present rates: for Exon arrays these are computed as percentage of probe sets with DABG < 0.01 out of 1,411,406 probe sets represented per array; One-Cycle = Affymetrix One-Cycle Target Labeling Assay; Two-Cycle = Affymetrix Two-Cycle Target Labeling Assay; NuGEN Pico = WT-Ovation Pico RNA Amplification System; NuGEN FFPE = WT-Ovation FFPE RNA Amplification System V2; NuGEN FFPE Exon = WT-Ovation FFPE RNA Amplification System V2; NuGEN FFPE Exon = WT-Ovation FFPE RNA Amplification System V2; with the WT-Ovation Exon Module; Affy WT Exon = Affymetrix Whole Transcript Sense Target Labeling Protocol for Exon arrays.

ahttp://bioinformatics.picr.man.ac.uk/vice

^bPaired FFPE and fresh-frozen tissue samples for case 56

°Paired FFPE and fresh-frozen tissue samples for case 58

simpleaffy package in BioConductor (15), and mapping of probe sets was based on Affymetrix annotation downloaded from NetAffx (16). Exon array normalization and Detection Above Background (DABG) *P*-value calculations were carried out using Affymetrix Power Tools. XMap and Exonmap running Ensembl version 47_36i were used for annotation and mapping of exon probe sets, respectively (17–19).

Results and discussion

The purpose of this study was to compare a range of commercially available RNA extraction, labeling and Affymetrix microarray platforms to identify optimal combinations for gene expression profiling of routinely prepared archival FFPE tissues.

The performance of RNeasy FFPE and RecoverAllkits was similar; yields of $2-23 \,\mu g$ were obtained from two $8-\mu m$ sections of cervical carcinoma with $260/280 \, nm$ absorption ratios of 1.8-2.0. We compared 12 RNeasy FFPE STS extracts with paired extracts from the same samples prepared using an in-house modified Optimum FFPE protocol, where tissues were incubated for longer periods in higher concentrations of proteinase K and at higher temperatures. Average RNA integrity numbers were low regardless of extraction method (Optimum, 2.4 ± 0.42 ; RNeasy, 2.3 ± 0.13 ; P = 0.39). UV absorbance ratios were also similar for both methods: 260/280 Optimum 2.0 \pm 0.05, RNeasy 1.9 \pm 0.09; 260/230 Optimum 1.6 \pm 0.57, RNeasy 2.0 \pm 0.25) (see Supplementary Materials for QC results and Agilent Bioanalyzer plots, and refer to our online protocols for FFPE preparation for RNA extraction and the







Figure 3. Increased transcript detection with Exon arrays. Mapping was performed using the Affymetrix annotation of the probe sets. After removing duplications and ambiguous mapping, a total of 36576 Plus 2.0 probe sets were used. (A) These probe sets mapped to 5383 genes for which no probe sets were called present in any of the four replicates (two replicates each of KL9 and KL16). On Exon 1.0 arrays, there are a total of 66,828 exonic probe sets targeting this gene list, 11,072 of which are detected on all six arrays. (B) 3943 genes have at least 1 probe set present (shown as black dots). (C) 609 have at least 5 probe sets called present (shown as black dots). Gray dots in B and C identify genes for which no probe sets were called present in any of the four Plus 2.0 replicates.

modified Optimum protocol). Interestingly, despite these similarities, the modified-Optimum extracts outperformed the new RNeasy extracts (Figure 1A) for percentage present (PP) on Plus 2.0 arrays, presumably because the former employed mechanical disruption as well as enzymatic digestion, which may be helpful for hard-to-lyse tissues such as soft tissue sarcoma. This observation attests to the value of in-house optimization and, since the Optimum Kit is no longer on the market, presents a benchmarked opportunity for RNeasy FFPE Kit optimization.

As expected, the NuGEN WT-Ovation amplification and labeling systems (Pico and FFPE) with their combined oligo(dT) and random hexamer priming strategy gave significantly better results than Affymetrix One-Cycle labeling; the average increase in PP rates on Plus 2.0 was 20 percentage points (range 10-27%) for modified-Optimum extracted samples labeled using NuGEN FFPE (Figure 1B). The NuGEN FFPE protocol gave slightly better PP rates than the Pico system and much higher rates than those obtained with Affymetrix Two-Cycle amplification (Table 1). Microarray hybridization (specificity) is increased with NuGEN systems through the use of amplified single-stranded cDNA (sscDNA) (20), which, in our experience, reduces background staining by minimizing variation in mismatch signal. The net effect of NuGEN FFPE labeling is a much greater reduction in mismatch probe signal intensities compared with Affymetrix 3'-driven reactions. Interestingly, we have found that Affymetrix One-Cycle labeling produces present calls that are not as



restricted to the 3' region as expected and, as a result, NuGEN labeled samples had a constant proportional increase in probe sets called present on Plus 2.0 arrays regardless of the distance from the 3' end of the transcript to the end of probe selection region/bases (Figure 2). One explanation is that oligo(dT) primers anneal to adenine-rich areas within the transcript as well as the poly(A)-tailed region itself. This may also explain why the increase in detection rates from extreme 3'-biased arrays such as Affymetrix X3P is relatively small in comparison to Plus 2.0 arrays when using One-Cycle labeling; the average benefit of the X3P array in two FFPE samples tested in this study was only 3-5% (Table 1).

Relative to Plus 2.0 expression profiles of paired 'gold-standard' fresh-frozen tissues (KL39 and KL40 are the fresh-frozen pairs for KL35 and KL36, respectively), where we have assumed 100% sensitivity and specificity of gene detection, the sensitivity from FFPE using NuGEN FFPE labeling (all other experimental conditions being equal) increased significantly from 50% on average using the One-Cycle protocol in our previous study (3) to 80% on average for the same samples (KL35:48 to 80%, KL36:51 to 81%) (Figure 1, D and E). A slight fall in specificity (from 96% to 92% and 95% to 88% for KL35 and 36, respectively) may be due to an inherent increase in variability associated with target amplification. These results compare favorably with recent published data: sensitivity and specificity results of 71% and 96%, respectively, were obtained in a similar comparative analysis by Almac Diagnostics using the NuGEN FFPE system and Affymetrix-based Disease Specific Microarrays (DSAs), which are deliberately enriched with 3' biased probe sets for degraded RNA/FFPE application (14). The 9–10% higher FFPE sensitivity achieved with our approach suggests that 3' probe set bias may no longer be necessary for FFPE templates when random primers are included in the reactions.

In order to test whether inter-gene expression relationships are faithfully preserved in FFPE, we compared fold changes generated between fresh-frozen tissue cases with those between paired FFPE, using transcripts present in both cases and differentially regulated with a fold change greater than log₂. Agreement for the direction of fold change was seen in 89% of differentially regulated genes. As a comparison, we carried out a similar experiment using data from cell lines (MCF7 and MCF10A) and found agreement in 83% of genes. Next, we compared scatter plots of signal intensity between fresh-frozen tissue and paired FFPE, applying an algorithm (21) to correct for differences introduced by different extraction protocols (in this case, TRIzol for fresh-frozen tissue and modified-Optimum for FFPE). This confirmed high data correlation, with R^2 values of 0.86 and 0.87 for KL36:KL40 and KL35:KL40 pairings (by comparison, $R^2 = 0.67$ for data generated for the same RNA extracted from one FFPE sample, KL36, using different methods, RNeasy and modified-Optimum). Comparably high correlations were obtained in another study similar to ours reported recently by Lassmann and colleagues (mean R^2 = 0.899) (4). Technical replicate data in our study (KL9 and KL16 Plus 2.0 triplicates) also confirmed excellent reproducibility of data (see Supplementary Materials for technical replicate plots).

Having gained experience with 3' expression arrays, we sought to explore Exon arrays for expression profiling of archival material as these arrays differ markedly from other expression arrays in having multiple independent probe sets for every known and predicted exon in the genome (instead of a general one-one

mapping between probe set and gene). We postulated that this design feature may prove advantageous for FFPE samples where degradation between samples is not consistent. In a small number of samples tested, the NuGEN WT-Ovation FFPE coupled with the WT-Ovation Exon Module and FL-Ovation cDNA Biotin Module produced higher gene detection rates than the WT Sense Target Labeling Protocol recommended by Affymetrix for Exon arrays (Figure 1C). The latter includes an rRNA depletion step based on hybridization with a small number of oligonucleotide probes; the efficiency of this removal will be significantly affected by the degree of degradation in different samples. As a result, we used the NuGEN FFPE protocol to label all samples for Exon arrays (freshfrozen and FFPE). The paired samples allowed us to calculate specificity and sensitivity for FFPE Exon arrays, as we had done previously for the same RNA extracts on Plus 2.0. Given the large increase in content, it was possible that these arrays would not



perform as well as 3' expression arrays. At probe set/exon level, the sensitivity of FFPE Exon arrays was lower than the 80% obtained with Plus 2.0-66% (KL35) and 70% (KL36)— which is not unexpected since there are only four probes per probe set and their design may be constrained by the short length of most exons. However, at gene level, sensitivity improved dramatically to 93% for each sample (Figure 1F). Specificity was high at both exon and gene levels: for KL35 and KL36, respectively, exon level was 92% and 94% and gene level was 94% and 96%, suggesting, as before, that false-positive rates are low and therefore that detected genes are reliable. As with Plus 2.0 data, technical replicate data confirmed excellent reproducibility for Exon arrays; Pearson correlations for the three possible KL9 Exon array pairings gave R^2 values of 0.84, 0.82 and 0.82, and for KL16, 0.79, 0.78 and 0.77.

An Exon array probe set was considered present if the DABG score was <0.01. On Plus 2.0 arrays, KL9 had 18,427 probe sets called present on both technical replicates, which mapped to 7,936 genes (KL16 had 18,948 and 8027 respectively). When the same samples were analyzed on Exon arrays, a total of 407,039 probe sets were detected, representing an increase to 20,303 genes (KL16: 388,526 and 20,242, respectively). The intersection between the gene lists was 7515 for KL9 and 7561 for KL16, equating to almost 95% of genes detected on Plus 2.0 arrays. In some cases Exon-detected transcripts were simply not present on the Plus 2.0 array. In order to enumerate the absolute increase in gene detection on Exon arrays compared with Plus 2.0, we created a list of 5383 genes which had no probe sets called present on any of the four Plus 2.0 technical replicates, and mapped this list back to exon probe sets (Figure 3A). As shown in Figure 3B, 3943 of these genes had at least one probe set detected on the Exon array. By setting confidence limits to genes detected by at least 5 probe sets, Exon arrays detected an additional 609 genes compared with Plus 2.0 (Figure 2C). Examining the full content of the Exon arrays, we also found 110,568 'intergenic' probe sets detected on all Exon arrays representing many ESTs and Genscan predictions; 1566 of these show differential expression [fold change (FC)>2 and *P* < 0.01].

While the main objective of this study was to optimize the technical performance of Affymetrix arrays for FFPE substrates, we were interested in obtaining preliminary data on the biological relevance of developed protocols, fully recognizing that examining biological end points in a small and histologically heterogeneous series of sarcoma samples may lead to flawed results. A PubMed search of the literature identified several papers reporting gene signatures of diagnostic and prognostic relevance in soft tissue sarcoma, but only our recent paper (3) used the Affymetrix platform. In one of the largest studies, Francis and colleagues reported a ~500-gene diagnostic signature for leiomyosarcoma (LMS) using 27K spotted cDNA two-color arrays (22). Since six of the 13 cases in our study were LMS, we chose to test the diagnostic capacity of Francis's diagnostic genes in discriminating LMS from other subtypes in our series. Mapping was done using gene symbols (since their probe set ID and probe set sequence data are not publicly available) and included all probe sets reliably detected on our 13 samples. The 70-gene signature derived from this exercise clustered together 4 of the 6 LMS cases using an unsupervised approach (see Supplementary Material for dendrogram). Taking into account the cross-mapping difficulties of comparing data from different platforms, this result supports the biological validity of FFPE data, but larger studies specifically designed to address biological questions are needed to validate this observation.

In the last few years, expression profiling of archival material has moved from what was widely regarded as an unrealistic goal to a reality. This study builds on previous published work aimed at optimizing techniques for microarray analysis of archival FFPE. Our results suggest that several products are available for extracting microarray-suitable RNA from archival samples and that a 3' bias in array design may no longer be required for future FFPE applications if a combination of oligo(dT) and random hexamers is used. The latter strategy also significantly improves transcript detection compared with oligo(dT) priming alone. The poor quality of FFPE RNA leads to spurious high background signals in some probes, but this effect is significantly lessened using a hybridization protocol in which cDNA is used in place of cRNA. Together with the redundant design of Affymetrix arrays (the Plus 2.0 array has 11 pairs of probes per probe set) this approach generated reliable data from most archival samples tested in this study. We conclude that both Plus 2.0 and Exon arrays work well with routinelyprepared archival samples, producing highly sensitive and specific transcript detection. Exon arrays give higher detection rates and provide additional unique information. These encouraging results reflect significant progress in the search for an effective tool for microarray gene expression profiling of archival fixed tissues.

Acknowledgments

This work was partly funded by Cancer Research UK and charitable donations made to The Christie NHS Foundation Trust. We thank Catharine West for supplying cervical tissue samples for this study, Maria Jeziorska for identifying representative areas for macrodissection, and Ric Swindell for his help with statistical analyses.

The authors declare no competing interests.

References

- 1. Bibikova, M., E. Chudin, A. Arsanjani, L. Zhou, E.W. Garcia, J. Modder, M. Kostelec, D. Barker, et al. 2007. Expression signatures that correlated with Gleason score and relapse in prostate cancer. Genomics 89:666-672.
- 2. Hoshida, Y., A. Villanueva, M. Kobayashi, J. Peix, D.Y. Chiang, A. Camargo, S. Gupta, J. Moore, et al. 2008. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. N. Engl. J. Med. 359:1995-2004.
- 3. Linton, K.M., Y. Hey, E. Saunders, M. Jeziorska, J. Denton, C.L. Wilson, R. Swindell, S. Dibben, et al. 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.
- 4. Lassmann, S., C. Kreutz, A. Schoepflin, U. Hopt, J. Timmer, and M. Werner. 2009. A novel approach for reliable microarray analysis of microdissected tumor cells from formalinfixed and paraffin-embedded colorectal cancer resection specimens. J. Mol. Med. 87:211-224.
- Coudry, R.A., S.I. Meireles, R. Stoyanova, H.S. Cooper, A. Carpino, X. Wang, P.F. Engstrom, and M.L. Clapper. 2007. Successful application of microarray technology to microdissected formalin-fixed, paraffin-embedded tissue. J. Mol. Diagn. 9:70-79.
- 6. Godfrey, T.E., S.H. Kim, M. Chavira, D.W. Ruff, R.S. Warren, J.W. Gray, and R.H. Jensen. 2000. Quantitative mRNA expression analysis from formalin-fixed, paraffin-embedded tissues using 5' nuclease quantitative reverse transcription-polymerase chain reaction. J. Mol. Diagn. 2:84-91.
- Lee, J., A. Hever, D. Willhite, A. Zlotnik, and P. Hevezi. 2005. Effects of RNA degradation on gene expression analysis of human postmortem tissues. FASEB J. 19:1356-1358.
- 8. von Ahlfen, S., A. Missel, K. Bendrat, and M. Schlumpberger. 2007. Determinants of RNA quality from FFPE samples. PLoS One 2:e1261.
- Frank, M., C. Doring, D. Metzler, S. Eckerle, and M.L. Hansmann. 2007. Global gene expression profiling of formalin-fixed paraffinembedded tumor samples: a comparison to snap-frozen material using oligonucleotide microarrays. Virchows Arch. 450:699-711.
- Penland, S.K., T.O. Keku, C. Torrice, X. He, J. Krishnamurthy, K.A. Hoadley, J.T. Woosley, N.E. Thomas, et al. 2007. RNA expression analysis of formalin-fixed paraffinembedded tumors. Lab. Invest. 87:383-391.

- 11. Cronin, M., M. Pho, D. Dutta, J.C. Stephans, S. Shak, M.C. Kiefer, J.M. Esteban, and J.B. Baker. 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.
- 12. Paik, S., C.Y. Kim, Y.K. Song, and W.S. Kim. 2005. Technology insight: application of molecular techniques to formalin-fixed paraffin-embedded tissues from breast cancer. Nat. Clin. Pract. Oncol. 2:246-254.
- 13. Xiang, C.C., M. Chen, L. Ma, Q.N. Phan, J.M. Inman, O.A. Kozhich, and M.J. Brownstein. 2003. A new strategy to amplify degraded RNA from small tissue samples for microarray studies. Nucleic Acids Res. *31*:e53.
- 14. Farragher, S.M., A. Tanney, R.D. Kennedy, and H.D. Paul. 2008. RNA expression analysis from formalin fixed paraffin embedded tissues. Histochem. Cell Biol. 130:435-445.
- 15. Wilson, C.L. and C.J. Miller. 2005. Simpleaffy: a BioConductor package for Affymetrix Quality Control and data analysis. Bioinformatics 21:3683-3685.
- 16. Liu, G., A.E. Loraine, R. Shigeta, M. Cline, J. Cheng, V. Valmeekam, S. Sun, D. Kulp, and M.A. Siani-Rose. 2003. NetAffx: Affymetrix probesets and annotations. Nucleic Acids Res. 31:82-86.
- Yates, T., M.J. Okoniewski, and C.J. Miller. 2008. X:Map: annotation and visualization of genome structure for Affymetrix exon array analysis. Nucleic Acids Res. 36:D780-D786.
- 18. Okoniewski, M.J., T. Yates, S. Dibben, and C.J. Miller. 2007. An annotation infrastructure for the analysis and interpretation of Affymetrix exon array data. Genome Biol. 8:R79.
- 19. Flicek, P., B.L. Aken, K. Beal, B. Ballester, M. Caccamo, Y. Chen, L. Clarke, G. Coates, et al. 2008. Ensembl 2008. Nucleic Acids Res. *36*:D707-D714.
- 20.Barker, C.S., C. Griffin, G.M. Dolganov, K. Hanspers, J.Y. Yang, and D.J. Erle. 2005. Increased DNA microarray hybridization specificity using sscDNA targets. BMC Genomics 6:57.
- 21. Sims, A.H., G.J. Smethurst, Y. Hey, M.J. Okoniewski, S.D. Pepper, A. Howell, C.J. Miller, and R.B. Clarke. 2008. The removal of multiplicative, systematic bias allows integration of breast cancer gene expression datasets - improving meta-analysis and prediction of prognosis. BMC Med. Genomics 1:42.
- 22. Francis, P., H.M. Namlos, C. Muller, P. Eden, J. Fernebro, J.M. Berner, B. Bjerkehagen, M. Akerman, et al. 2007. Diagnostic and prognostic gene expression signatures in 177 soft tissue sarcomas: hypoxia-induced transcription profile signifies metastatic potential. BMC Genomics 8:73.

Received 10 December 2008; accepted 8 April 2009.

Address correspondence to Kim Linton, Cancer Research UK Department of Medical Oncology, The Christie NHS Foundation Trust, Wilmslow Road, Withington M20 4BX, Manchester, UK. email: kim.linton@christie.nhs.uk