Supplementary Information

Previous pilot study

In a pilot study, 7 pairs of additional samples were obtained in which a part of the same tumor had been both snap frozen and fixed in formalin. RNA was extracted from both, and expression data were generated using the DASL 502 Cancer Gene Chip. Correlation r^2 values between paired samples were reasonable, with a mean r^2 value for all seven paired replicates of 0.7 which is comparable with other studies comparing fresh or fixed cell line RNA or matched frozen and FFPE tissue samples (8, 24) using DASL technology.

Sample Preparation

To evaluate the typical tumor content of the tissue cores taken in this study, cores were embedded horizontally in wax blocks and these blocks were sectioned to obtain 5-micron sections throughout the whole TMA core. H&E slides were then prepared from these sections and viewed under light microscopy to determine the percentage tumor content of the core. The tumor content of tissue cores sampled in this fashion when assessed visually was estimated be at least 70% tumor cells. An example of one of these TMA core cross-sections is presented in Figure 2s.

An H&E stained slide was made subsequent to sampling of blocks, which was reviewed to ensure that the correct sample was taken (Figure 1s). Cores were stored at 4^0C until required in a sealed vial.

RNA extraction and quality control

Previous studies have demonstrated that the DASL assay produces reproducible gene expression results from degraded RNA such as that extracted from formalin fixed tissue (24, 41, 42) (8, 23). Gene expression profiles generated from the assay correlate well with profiles derived from intact RNA extracted from frozen tissue or cell lines (8, 23, 24) and yield biologically relevant results (24) (41). Many authors have suggested quality criteria to assess suitability of an RNA sample for use in the DASL assay. A commonly used quality control measure is quantitative real time PCR (qRT-PCR) of the housekeeping gene, *RPL13a*. Cycle threshold (Ct) values of ≤29 cycles (41, 42) or ≤28 cycles (8, 23, 42) have indicated adequate quality of RNA samples in previous work. Others have used spectrophotometry to measure RNA concentration and A_{260}/A_{280} ratios, RNA concentrations of <20ng/μL, <25ng/μL or total RNA input <50ng and A_{260}/A_{280} ratios <1.5 or <1.8 were reported to indicate that the sample is inadequate (8, 23, 42). The Agilent 2100 Bioanalyzer generates an RNA integrity number (RIN) for each RNA sample. It has been suggested that an RIN number >2.0 or ≤2.0 with a RNA fragment size >200nt indicates that RNA is suitable for use in the assay (23), others reported that RIN values do not correlate with the performance of RNA samples (42). RNA becomes increasingly degraded with storage time of tumor blocks as demonstrated by higher Ct values generated using qRT-PCR in older samples (8) but one study has reported successful profiling using blocks stored for over 24 years (24).

Concentrations of extracted RNA were determined in our samples using the ND-1000 spectrophotometer (NanoDrop, Wilmington, DE, USA). An aliquot of extracted RNA (5 µl) was removed for cDNA synthesis with the Superscript ™ First-strand synthesis system for RT-PCR (Invitrogen, UK), and the remaining was cryopreserved at -80°C until required for further analysis. Samples in Study 1 were analysed on the Agilent 2100 Bioanalyser using a RNA 6000 nano-chip and standard manufacturer's protocols to determine the average fragment length of the RNA. The electropherogram peaks were viewed for presence and size of RNA peaks in comparison to a RNA size marker and an RNA integrity (RIN) score (14) was determined for 21 samples in Study 2. Gel images were also produced for visual representation of RNA quantity and quality.

The RNA fragment length traces from all samples in Study 1 were visually evaluated and a fragment analysis trace score of 1-3 was created where 1 = peak visible and labeled with fragment size marker, 2 = peak visible but not labeled with fragment size marker and 3 = little or no peak visible (See Table 3s).

The final quality control measure performed on extracted RNA samples was a quantitative real-time PCR (qRT-PCR) on a 90bp fragment of the control gene *RPL13a* performed in duplicate using SYBR Green and a 7500 real-time machine (Applied Biosystems) with standard protocols and PCR conditions. Primer sequences were Forward:GTACGCTGTGAAGGCATCAA and

Reverse:GTTGGTGTTCATCCGCTTG. Mean cycle threshold values were determined for all samples and those samples with a C_T value ≤29 cycles were used in the assay. If the sample had a Ct above this value it gave a reasonable reassurance of reproducible behavior in the DASL assay. Table 2s shows the results of replicates within the assay.

We used three RNA quality control measures, the ND-1000 spectrophotometer (Nanodrop, Wilmington, DE, USA), the Agilent 2100 Bioanalyser and qRT-PCR of *RPL13a*. Bioanalyzer traces were generated on all samples in Study 1. RNA concentration measure with spectrophotometry was not predictive of number of detected genes in Study 1 (Spearman's correlation, p=0.23) and only weakly predictive in Study 2 (Spearman's correlation, p=0.05). The Bioanalyzer traces for samples from Study 1 were not evaluable other than by subjective analysis of RNA fragment length traces and this did not correlate with the number of genes detected using the assay or the spectrophotometry derived RNA concentration (table 4s). In a limited number of samples in Study 2, the Bioanalyser RNA concentration and RIN number were generated. In these samples both increased RNA concentration and RIN number predicted greater number of detected genes, Spearman's correlation p=0.001 and Kruskal-Wallis test p=0.03 respectively. Quantitative RT-PCR was performed on all samples, but failed in Study 1. In Study 2, the C_T value were negatively correlated with number of genes detected (Spearman's correlation coefficient, p=0.003).

DASL expression arrays methodology

200ng total RNA was converted into cDNA using biotinylated random hexamers and oligo-deoxythymidine 18 primers, and Illumina-supplied reagents according to manufacturer's instructions. To monitor intra-assay variation, each DASL assay array plate included 6-12 technical replicates within and across plates and 2 replicate Stratagene Universal human reference RNAs (Agilent Technologies, UK). Pairs of query oligos were then annealed to complementary sequences (~50 bases) flanking the specific cDNA target site. The first oligo consists of two parts: the gene specific sequence and a universal PCR primer sequence at the 5' end. The second oligo consists of three parts: the gene specific sequence, a unique address sequence which is complementary to one of 1536 capture sequences on the array, and a universal PCR primer at the 3' end. The biotinylated cDNA was then bound to streptavidin-conjugated paramagnetic particles and mis-hybridised and non-hybridised oligos were washed away. Hybridised oligos were then extended and ligated to generate amplifiable templates. A PCR reaction was then performed with Cy3 and Cy5 labelled universal PCR primers and the double stranded PCR products were isolated and hybridised at 45°C for 18 hours to the Sentrix Universal-96 Array Matrix (SAM). The SAM is a fibre-optic assembly composed of 96 individual arrays assembled by loading pools of 3μm glass beads derivatised with oligonucleoides onto the etched ends of fibre optic bundles. About 50,000 optical fibres are hexagonally packed to form a 1.4mm diameter bundle (43). The fibre optic bundles are assembled into an array matrix, comprising 96 bundles arranged in an 8x12

matrix that matches the dimensions of standard micro-titer plates. With both platforms. The beads are positioned randomly on the array matrix so a decoding process is carried out to determine the location and identity of each bead in every location (8, 43). After hybridisation the arrays were scanned using the BeadArray™ scanner and analysed by BeadStudio v3 software (Illumina, USA). This software creates an intensity data file for each sample that is used in downstream data analysis.

Data Handling

The number of detectable genes in each sample was listed by Beadstudio software in sample control reports and represents the genes for which the target sequence signal is distinguishable from the negative controls, using a statistical procedure within Beadstudio. Control reports display internal standards from the Cancer Panel that measure assay performance (by detection of the control gene QARS), signal background and noise. The data were normalized within BeadStudio Gene Expression Module v3.4 (Illumina Inc). Fluorescence intensities from Cy3 and Cy5 dyes were averaged for each probe and the expression level of each gene was computed as an average of the intensities from three probes. Background correction and cubic B-spline smoothing methods were used and sample scaling was applied to remove the variation between plates. The quality of the normalization was assessed by testing the equality and the correlation between gene expression levels across sample replicates.

Background correction removes the gene expression from non-specific hybridization, estimated as the average expression for the negative control gene *OARS*. In BeadStudio, the outliers are removed using a method based on median absolute deviation. In the cubic spline method, 15 quantiles of the samples were scaled to have equal values as the corresponding quantiles in the reference sample (which is the average of all samples) and cubic B-spline smoothing was used to interpolate the values between quantiles. This method assumes that all transcripts have the same abundance and it is well suited for removing the effects that might cause non-linear transformation such as saturation. For points outside the interpolation interval, BeadStudio uses linear interpolation. In the sample scaling normalization, the samples that have technical replicates on different plates are used to calculate a scaling factor that is applied to all other samples to create equal average fluorescence intensity for each probe across the plates.

Statistical analysis of sample replicates

A number of RNA samples were replicated across the plates for DASL analysis. Biological replicates were also included which were RNA samples extracted from adjacent parts of the same tumor. Equality and correlation between gene expression levels in technical replicates was assessed using paired t-tests to compare replicate mean gene expression and Pearson's correlation coefficients (results presented in Table 2s).

Target validation with qRT-PCR

Small fragments (60-120bp) of *SPPI* exon1/2, *SPPI* exon 5/6 (location of DASL probes), based upon the NCBI reference sequence NM_000582.2, and the endogenous control *GAPDH* were amplified using Taqman® Gene Expression Assay probes (Hs00960942_m1 (Exons 1-2) and Hs00959010_m1 (Exons 5-6), Applied Biosystems, Warrington, UK) in duplicate from cDNA previously synthesised from RNA samples extracted from Study 1. cDNA was diluted 1:1 with nuclease free water and 1µl was used for each qRT-PCR assay. PCRs were performed with Gene Expression Master Mix (Applied Biosystems) in a 20µl final volume on the ABI 7900 (Applied Biosystems) using standard conditions. Automatic settings were used for baseline and threshold determination and Cycle threshold values for each target were exported to Excel for further analysis. Delta Ct values were determined as described previously (44), where delta Ct = Ct gene of interest-Ct internal control. Delta Ct values were then converted to $2^{-\text{delta Ct}}$ for the calculation of fold change between relapsers and non-relapsers where fold change = $2^{-\text{delta Ct}}$ relapsers/$2^{-\text{delta Ct}}$ non-relapsers.

Supplementary References

41. Bibikova M, Chudin E, Arsanjani A, et al. Expression signatures that correlated with Gleason score and relapse in prostate cancer. Genomics 2007; 89:666-72.
42. Abramovitz M, Ordanic-Kodani M, Wang Y, et al. Optimization of RNA extraction from FFPE tissues for expression profiling in the DASL assay. Biotechniques 2008; 44:417-23.
43. Galinsky VL. Automatic registration of microarray images. II. Hexagonal grid. Bioinformatics 2003; 19:1832-6.
44. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 2001; 29:e45.

Figure 1s: an H&E section of a block after TMA sampling

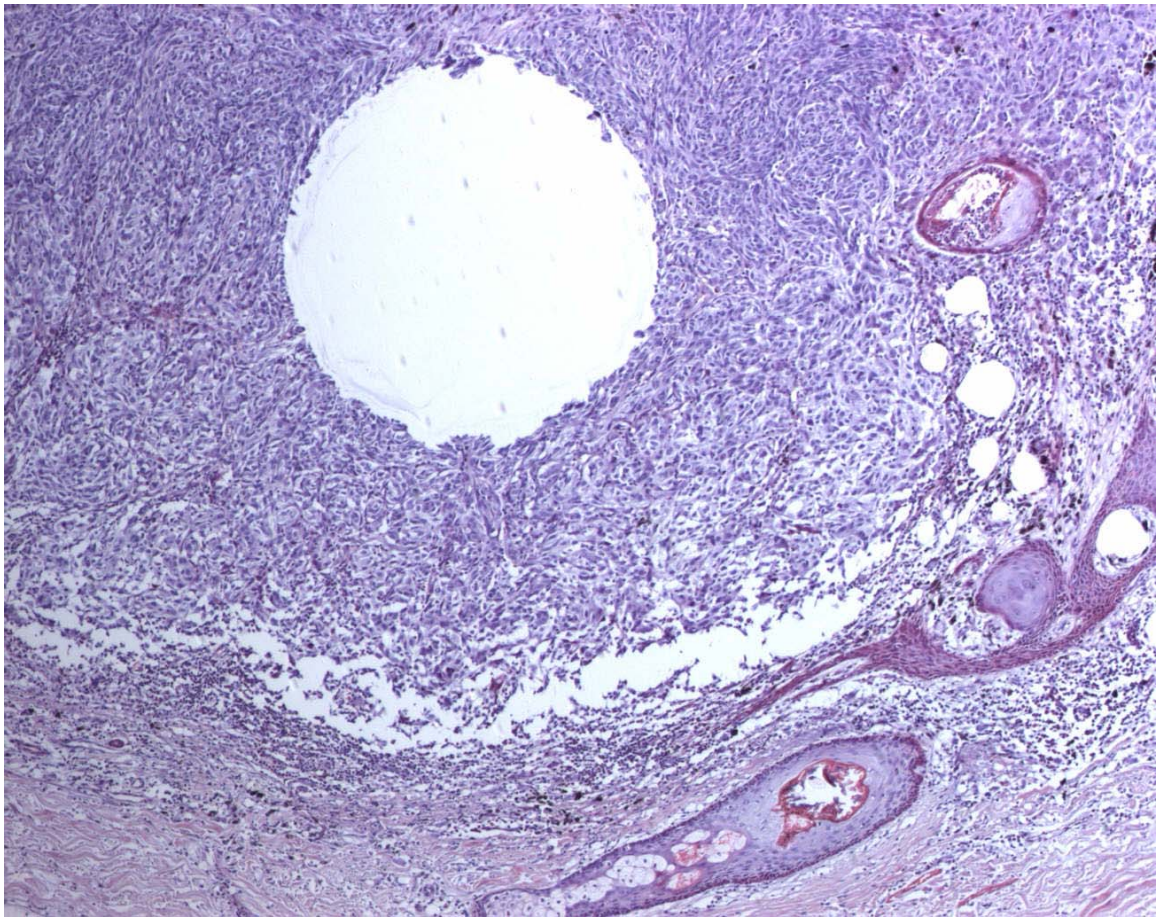

Figure 2s: A tumor core horizontally bisected to reveal the admixture of tumor and non-tumor cells.

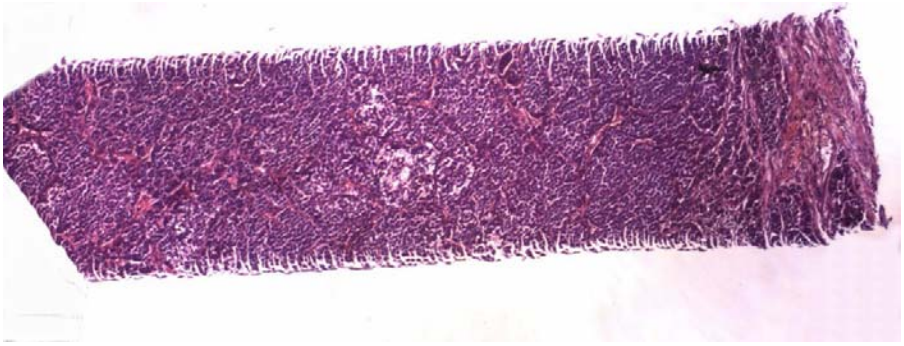

Table 1s: Quality control measures and their correlation with number of genes detected using DASL for Studies 1 and 2

	Study 1 **(188 samples)**	**Association between measure and detected genes. Test statistic and p-value.**	**Study 2** **(235 samples)**	**Associations between measure and detected genes. Test statistic and p-value.**
Number failed samples (%) (<250 genes)	4 (2.1)		2 (0.9)	
Number detected genes overall, mean (range)	434 (33-472)		457 (180-493)	
Age of block, years, mean (range) **Number genes detected, mean (range):** **Block <5 years old** **Block 5-10 years old** **Block >10 years old**	4.63 (1.48-7.84) 436 (33-468) 442 (317-472) N/A	Spearman correlation 0.11, p=0.13	6.26 (2.23-15.16) 466 (371-491) 456 (180-493) 441 (376-477)	Spearman correlation -0.24, p=0.0002
Melanin score: **0 (%)** **1 (%)** **2 (%)** **3 (%)**	In 102 samples: 5 (4.9) 30 (29.4) 42 (41.2) 25 (24.5)	Kruskal-Wallis chi squared 8.44, p=0.04	4 (1.7) 58 (24.9) 63 (27.0) 108 (46.4)	Kruskal-Wallis chi squared 4.02, p=0.26
Nanodrop RNA concentration, ng/μL, mean (range)	44.6 (0-87)	Spearman correlation 0.09, p=0.23	53.7 (0.4-209.9)	Spearman correlation 0.13, p=0.05
Bioanalyser RNA concentration, ng/μL, mean (range)	Not assessed		In 27 samples: 24.6 (0-56)	Spearman correlation 0.59, p=0.001
Bioanalyser RIN number, median (range)	Not assessed		In 21 samples: 1.5 (1.0 -2.3)	Kruskal-Wallis chi-squared 6.90, p=0.03
Ct values mean (range)	Failed		24.3 (19.9-27.3)	Spearmans correlation -0.20, p=0.003

Table 2s: Comparison of technical and biological replicates assessed in each study. Failed samples were classified as samples in which <250 genes were detected using the DASL assay. Pearson's correlation coefficients were calculated for each pair of replicates and mean gene expression was compared between the replicates using paired t-tests.

Study	No. technical replicates (no. of pairs with failed samples)	No. biological replicates (no. of pairs with failed samples)	Range of correlation coefficients (all pairs)	Range of correlation coefficients (pairs without failed samples)	Pairs with significantly different mean gene expression (t-test - p <0.05) (All samples)
Study 1	25 (2)	3 (1)	0.48-0.99	0.91-0.99	0
Study 2	25 (1)	11 (0)	0.38 -0.99	0.68-0.99 (technical) 0.41-0.97 (biological)	0

Table 3s: Descriptive characteristics of sample sets and larger study sets

Criterion	Study 1 overall data set	Study 1 sampled tumors	Study 2 overall data set	Study 2 sampled tumors
% Males	40	48	49	52
Mean age (range)	52.1(18-76)	54.9(19.9-78.5)	51.4(7.0-88.7)	52.0(14.4-88.0)
Median Breslow (range)	1.3 (0.2-15.0)	1.9 (0.9-12.0)	1.9 (0.75-24)	2.0(0.78-24.0)

Table 4s: Assessment of Bioanalyzer peak height and spectrophotometry derived RNA concentration and number of genes detected using the DASL assay for samples from Study 1.

	Spearman's correlation coefficient	**Significance value**
Spectrophotometry RNA concentration and Bioanalyzer peaks	-0.11	0.14
Genes detected and Bioanalyzer peaks	-0.07	0.37