

RESEARCH ARTICLE

Open Access



# Reliable gene expression profiling of formalin-fixed paraffin-embedded breast cancer tissue (FFPE) using cDNA-mediated annealing, extension, selection, and ligation whole-genome (DASL WG) assay

Mahesh Iddawela<sup>1,2,3,4,5,6\*</sup>, Oscar M. Rueda<sup>1</sup>, Marcus Klarqvist<sup>1</sup>, Stefan Graf<sup>1</sup>, Helena M. Earl<sup>2,3,4</sup> and Carlos Caldas<sup>1,2,3,4</sup>

## Abstract

**Background:** The difficulties in using formalin-fixed and paraffin-embedded (FFPE) tumour specimens for molecular marker studies have hampered progress in translational cancer research. The cDNA-mediated, annealing, selection, extension, and ligation (DASL) assay is a platform for gene expression profiling from FFPE tissue and hence could allow analysis of large collections of tissue with associated clinical data from existing archives, therefore facilitating the development of novel biomarkers.

**Method:** RNA isolated from matched fresh frozen (FF) and FFPE cancer specimens was profiled using both the DASL whole-genome (WG) platform, and Illumina BeadArray's, and results were compared. Samples utilized were obtained from the breast cancer tumour bank held at the Cambridge University Hospitals NHS Foundation Trust.

**Results:** The number of reliably detected probes was comparable between the DASL and BeadArray platforms, indicating that the source of RNA did not result in a significant difference in the detection rates (Mean probes- 17114 in FFPE & 17400 in FF). There was a significant degree of correlation between replicates within the FF and FFPE sample sets ( $r^2 = 0.96-0.98$ ) as well as between the two platforms (DASL vs. BeadArray  $r^2 = \text{range } 0.83-0.89$ ). Hierarchical clustering using the most informative probes showed that replicate and matched samples were grouped into the same sub-cluster, regardless of whether RNA was derived from FF or FFPE tissue.

**Conclusion:** Both FF and FFPE material generated reproducible gene expression profiles, although there was more noise in profiles from FFPE specimens. We have shown that the DASL WG platform is suitable for profiling formalin-fixed paraffin-embedded samples, but robust bioinformatics analysis is required.

**Keywords:** FFPE, WG DASL, Molecular Markers, Breast Cancer

\* Correspondence: mahesh.iddawela@monash.edu

<sup>1</sup>Cancer Research UK Cambridge Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, UK

<sup>2</sup>Department of Oncology, Addenbrooke's Hospital, University of Cambridge, Hills Road, Cambridge CB1 9RN, UK

Full list of author information is available at the end of the article



## Background

Formalin-fixed paraffin-embedded (FFPE) samples are a valuable source of clinical material, with great potential for predictive and prognostic biomarker discovery. RNA extracted from FFPE tissue is degraded, and therefore analysis is difficult since formalin fixing leads to cross-links with proteins and mono-methanol modification of nucleotide bases [1–3]. Improvements in technology has made analysis of this material feasible, with considerable potential for translational research using FFPE tissue.

The cDNA-mediated, annealing, selection, extension, and ligation assay (DASL<sup>®</sup>, Illumina, San Diego) is a platform that was developed for high-throughput gene expression analysis of partially degraded RNA. This assay initially assessed the expression of up to 1536 probes (a maximum of 512 genes, with three probes per gene) using 100–200 ng of degraded RNA [4]. A combination of random hexamers and oligo-dT priming is used for cDNA synthesis. The assay is designed to analyse 50 bp RNA fragments, and the sensitivity of the platform is enhanced by PCR amplification using common primers and by having 30 replicate beads per probe. These arrays are assembled randomly, and the position of each probe is determined following manufacture, which reduces spatial effects [5].

The initial DASL report showed that reproducible and disease-specific gene expression profiles can be generated using FFPE samples up to 10 years old [6]. They also compared gene expression in fresh frozen (FF) and FFPE specimens using eight matched cancer and normal samples, and derived a signature that could differentiate between cancerous and normal tissue in breast and colorectal cancers.

One of the limitations of the original DASL assay was the limited number of transcripts interrogated. Hence, the introduction of the whole-genome DASL assay (DASL WG), which profiles 24,000 transcripts, represents a marked improvement. In the initial report [7], the number and overlap of probes in FFPE samples compared to matched FF specimens were 88 % and 95 %, respectively. Furthermore, there was a 74 % overlap between differentially expressed genes in FF and FFPE samples. The assay was able to detect 1.5–2-fold changes in gene expression in FFPE specimens and produced high self-correlation between replicates ( $r^2 > 0.98$ ). Chien et al. showed that the DASL WG assay can be used to generate a signature typical of low-grade serous ovarian carcinoma and thereby distinguish it from borderline tumours [8]. This was subsequently used to differentiate serous and borderline tumours in three independent datasets with 80 % sensitivity. Another study investigated the effectiveness of DASL WG in predicting subtypes of familial breast cancer from FFPE samples; however the study did not include a direct comparison with matched FF material [9]. In another study, comparison of expression profiles

between FFPE and corresponding normal FF tissues found that only one quarter of genes identified in FF samples were also detected in FFPE samples [10]. Several studies using DASL have been published, but the majority have not used matched FF and FFPE samples, or have used different platforms to analyse FF samples [11, 12].

Here, we report a robust assessment of the DASL WG platform in gene expression profiling of matched FFPE and FF RNA samples from breast cancers. For orthogonal comparison, the same FF samples were also profiled with the Illumina BeadArray platform.

## Methods

Matched FF and FFPE samples from a cohort of breast cancer patients treated in 1999 at Cambridge University Hospitals NHS Foundation Trust were obtained from the tumour bank. This study has been approved by the Cambridge University Hospital Research Ethics committee.

### RNA extraction

Five to eight 5 µm-thick sections were cut from each block of FFPE tissue, and the samples were deparaffinised with xylene and digested with proteinase K for 14 h. Other steps, including purification and DNase treatment, were performed according to the manufacturer's protocol, and total RNA was stored at –80 °C after extraction (Roche high Pure Kit). The FF RNA was extracted using the Qiagen mirNEasy kit (Qiagen) according to the protocol.

### Whole-genome DASL

Total RNA extracted from FFPE and FF samples was converted to cDNA using oligo-dT<sub>18</sub> and random nanomer primers. Pairs of query oligonucleotides, including three unique pairs for each of 24,000 genes, were annealed to complementary sequences (~50 bp) flanking the specific cDNA target site. The biotinylated cDNA was then bound to streptavidin particles, and mis- and non-hybridised oligonucleotides were washed away. Each primer had a gene-specific sequence and a universal PCR primer site; a prolonged annealing time of 16 h was used, with a temperature gradient from 70 to 30 °C. PCR was performed using universal primers and a single colour (Cy3) assay. The amplified cDNA was then denatured with NaOH, precipitated, and washed with 75 % alcohol. The cDNA was then re-suspended and hybridised to Illumina Human Reference 8 BeadArrays for 16 h at 65 °C. The slides were washed and imaged using the Illumina BeadArray Reader (Illumina, San Diego, CA).

### Illumina expression BeadArrays

Total RNA (200 ng) from the same FF tumours assessed in the DASL WG assay was used to generate cRNA, using the Illumina TotalPrep RNA amplification kit (Ambion).

Reverse transcription with the T7 oligo (dT) primer was used to generate first strand cDNA, from which a second strand was produced using DNA polymerase. The RNA was subsequently degraded with RNase H and cleaned-up. *In vitro* transcription and biotin-labelled UTP were used to generate multiple copies of biotinylated cRNA, which was purified with a filter cartridge and quantified using a NanoDrop spectrophotometer.

The labelled cRNA target (1.5 µg) was hybridised to an array according to the Illumina HumanRef-6 BeadArray protocol. A maximum of 10 µl of cRNA was mixed with 20 µl of GEX-HYB hybridisation solution. The pre-heated 30 µl sample was dispensed into the large sample port of each array and incubated for 18 h at 58 °C at a rocker speed of five. The samples were hybridised to Illumina BeadArrays (Human Ref 6) and were washed according to the standard protocol and scanned with a BeadArray Reader (Illumina, San Diego, CA).

#### Analysis of gene expression data

Data from the Illumina microarray experiments were analysed using BeadStudio® software (Illumina, San Diego) and R (version 2.6.0) [13].

During this analysis, image files were loaded using the BeadArray package. The BASH algorithm was used to remove spatial artifacts [14]. The algorithm generates a grid of neighbours for each bead and identifies three types of defects: compact, diffuse (regions that contain more outliers than expected by chance), and extended. The average intensities of each probe were computed to summarise replicates and finally the data was quantile normalised.

Data for both experiments was combined to give a total of 24526 probes. Comparisons between FF (BeadArray) and FF (WG DASL) and between FF (WG DASL) and FFPE (WG DASL) were performed using identical probes with Student's *t*-tests. For carrying out comparisons between frozen samples used in DASL and BeadArray, a linear model per probe was fitted to detect those showing substantial differences. All probes with a *p*-value smaller than 0.05 were selected for clustering analysis.

For comparisons between paraffin and frozen samples in the DASL WG assay, a similar linear model was fitted and the probes with *p*-value < 0.05 were selected.

Correlations between these types of samples were computed using all probes and this selection of good probes. The final subset of probes that were selected in the platform and the tissue comparisons were used as input for the final cluster analysis. The data is deposited in a public data repository.

## Results

### Samples and quality control

As part of this study, five matched FFPE and FF samples were profiled in multiple technical replicates using the

DASL WG assay. The eight FF cancers were also profiled using Illumina BeadArray for inter-platform comparison, including five described above (Table 1). The size of the recovered RNA ranged between 100 and 200 base pairs in length. *RPLP13* QRT-PCR provided a quality control measure and only samples with a cycle threshold (Ct) value of 22–26 were deemed suitable for further analysis.

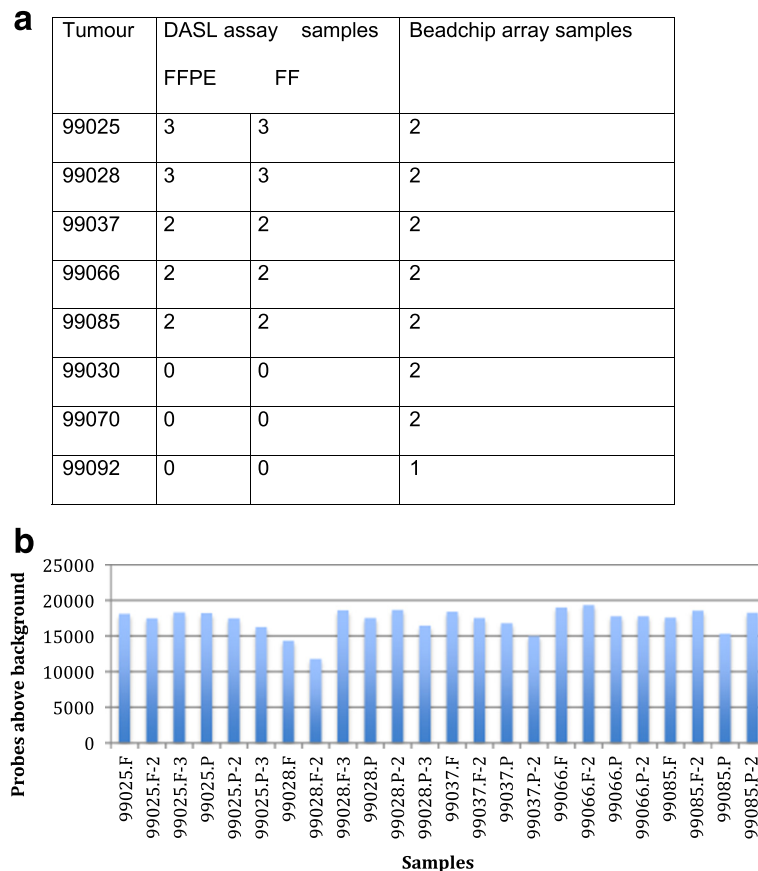
### Assay performance

To assess the DASL assay performance, the average number of genes detected above background were compared in the five matched FF and FFPE samples. The mean number of probes detected above background in FFPE samples were 17114 (range 15291–18662) and 17400 (range 11763–19320) in FF samples, and more probes were detected reliably in four of the FF samples used (Fig. 1b). The mean number of probes detected in replicates of each sample was 17598 (range 14872–19146) for FF and 17053 (16769–17773) for FFPE.

The probes were carefully re-annotated according to alignment characteristics (perfect match, good, bad or no match) and the signal was compared against these features to quantify their effect. While there was a significant low signal in the no match probes, the perfect matched probes had the greatest signal as expected. When the reliability of probe detection and the size of the genes were investigated, there was a significant difference. The genes that were shorter were more likely to be reliable than those which are longer. Further, the proportion of detected probes in FF and FFPE samples were compared, prior to normalisation, to assess the quality of the signal in the DASL assay. This showed that as expected, there is a greater variation in the signal intensity in the FFPE samples compared to FF samples.

**Table 1** Correlation of gene expression between matched FFPE and FF samples pre- and post-gene selection using the DASL WG assay and Illumina BeadArray

	Before gene selection	After gene selection
	$r^2$	$r^2$
DASL WG FFPE vs. FF		
99037	0.83	0.96
99085	0.86	0.97
99025	0.86	0.98
99028	0.85	0.96
99066	0.89	0.98
DASL WG FFPE vs. FF BeadArray		
99037	0.80	0.83
99085	0.82	0.86
99025	0.81	0.86
99028	0.81	0.85
99066	0.81	0.89



**Fig. 1 a** Samples and the number of technical replicates used for the DASL WG and Illumina BeadArray analysis & **(b)** The number of probes detected above background from Fresh (F) and FFPE (P) samples in the DASL WG assay

The reproducibility of the WG DASL assay for FFPE (using hierarchical clustering analysis), showed that all replicate FFPE samples clustered together (except for 99028P-3, which is a single outlier), suggesting that the assay has a high degree of reproducibility for FFPE analysis.

### FFPE vs. FF sample analysis

In order to assess the robustness of the DASL WG platform for analysing FF and FFPE samples, the correlation between different sample types was computed. It ranged between 0.83 and 0.89 for all the probes, and improved to 0.96–0.98 upon gene selection undertaken as described in the methods section (Figs. 1 and 2, Table 1).

Hierarchical clustering using the best performing probes showed significant homogeneity, with profiles from four of the samples (99085, 99028, 99066 and 99025) clustering together irrespectively of the origin (Fig. 3). There were outliers, but these were usually samples with a lower number of genes detected reliably.

### DASL WG assay and Illumina BeadArray

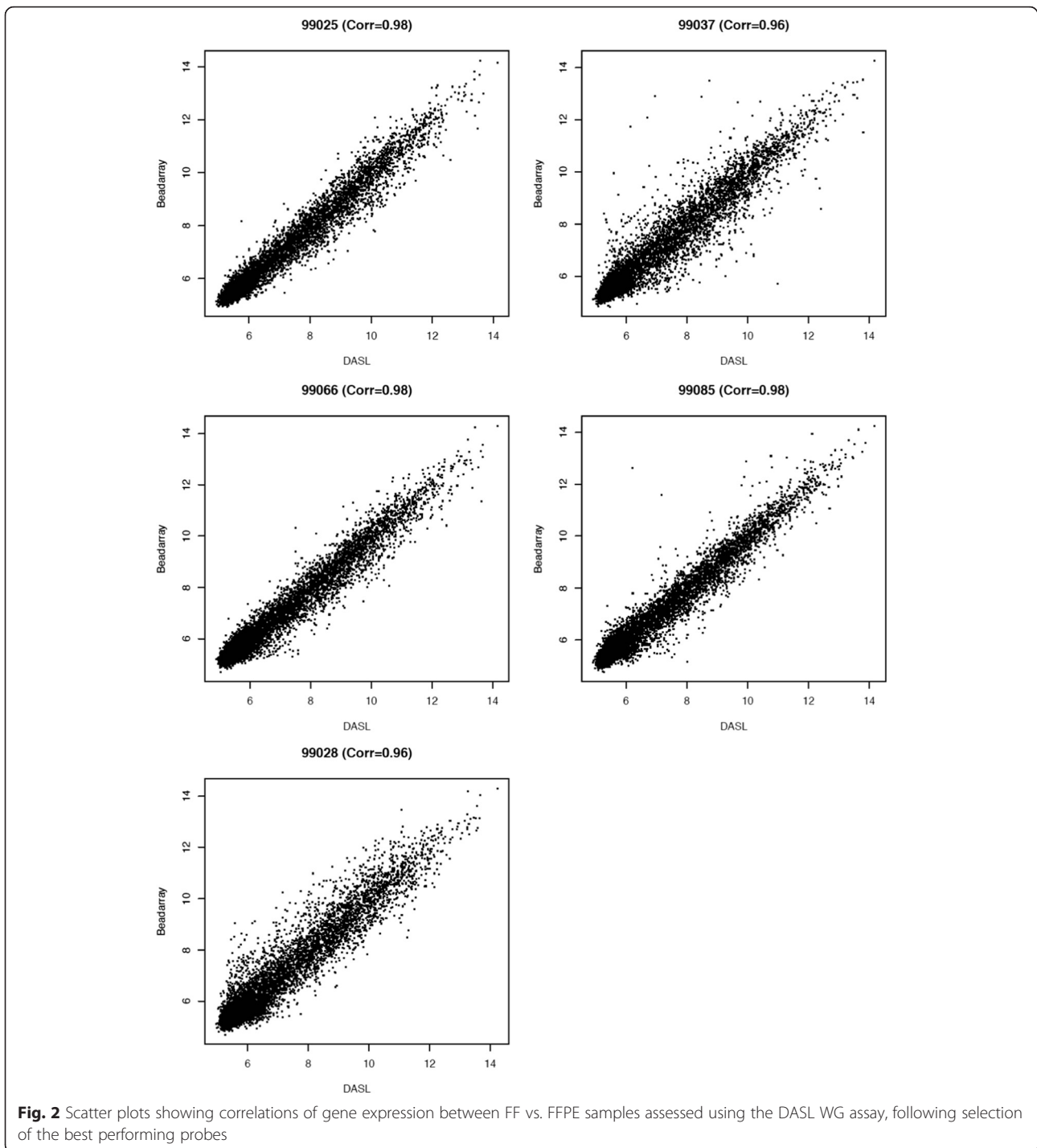
The correlation between all the probes in the DASL WG and BeadArray was assessed to compare the two platforms.

The correlation ranged between 0.80 and 0.82, and improved to 0.83–0.89 when only the best performing probes were used for the analysis (Table 1).

Although the DASL WG assay data (using FFPE and FF samples) and the BeadArray (using only FF specimens) are two independent platforms, hierarchical clustering using all probes showed that samples grouped according to the RNA sample source and the array platform used. In fact, all FFPE samples were closer together in the hierarchical clustering tree compared to FF specimens assessed with the BeadArray. Interestingly, the FF samples assessed by Illumina BeadArray and DASL WG did not cluster next to each other, but rather in two distinct groups.

When the analysis was repeated using only the best performing probes (as described in Methods) matched FF and FFPE samples clustered together (Fig. 4). This is one of the first studies to show robustly this significant overlap using DASL and BeadArray. Four out of the five tumours in which gene expression was assessed using both illumina BeadArray and DASL WG clustered together.

Gene Ontology analysis was undertaken to assess the probes that were discordant in the FF and FFPE samples

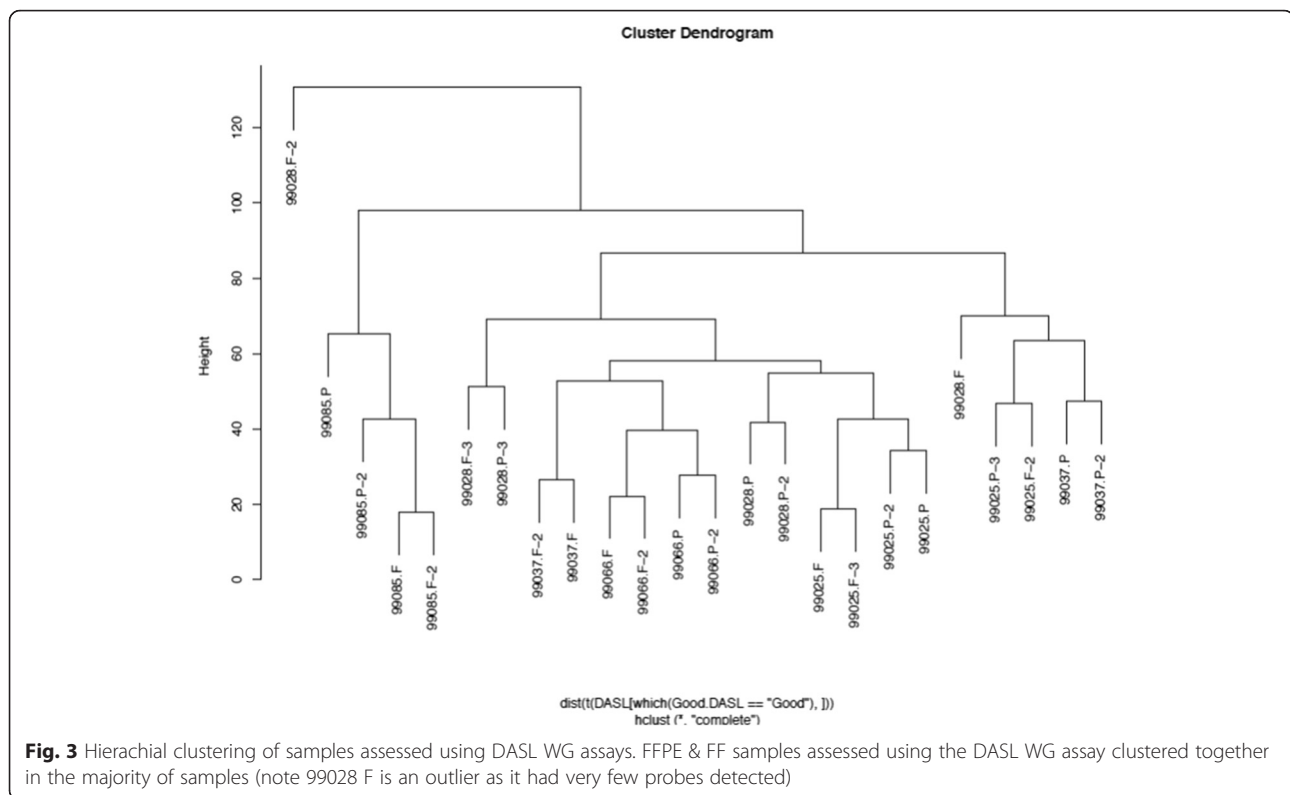


to get a better understanding of the biological processes involved (Supplementary File 1). This showed that as far as the biological processes are concerned: primary metabolic process and nucleic acid metabolic process were the two top in this category. While Poly (A) RNA binding and heterocyclic RNA binding were the top two molecular functions enriched in this group.

## Discussion

This study shows that reliable expression profiles could be generated from FFPE core biopsy tissue, and this supports the future investigation of this platform for analysis of multiple markers in breast cancer. The ability to use FFPE material makes large collections of currently unutilised resources available for molecular





marker discovery. Over the past few years, there has been considerable interest in using FFPE tissue for translational research, following the success of the *Recurrence score*<sup>®</sup> and the PAM50 score, which have demonstrated that reliable tests can be performed using FFPE tissue and used in routine clinical practice [15, 16]. This has led to improvements in the methods used for RNA extraction and in platforms used for FFPE analysis. Multiple tests have been developed for breast, lung, prostate, and colorectal cancer, which can potentially use FFPE samples for molecular marker analysis [17–19].

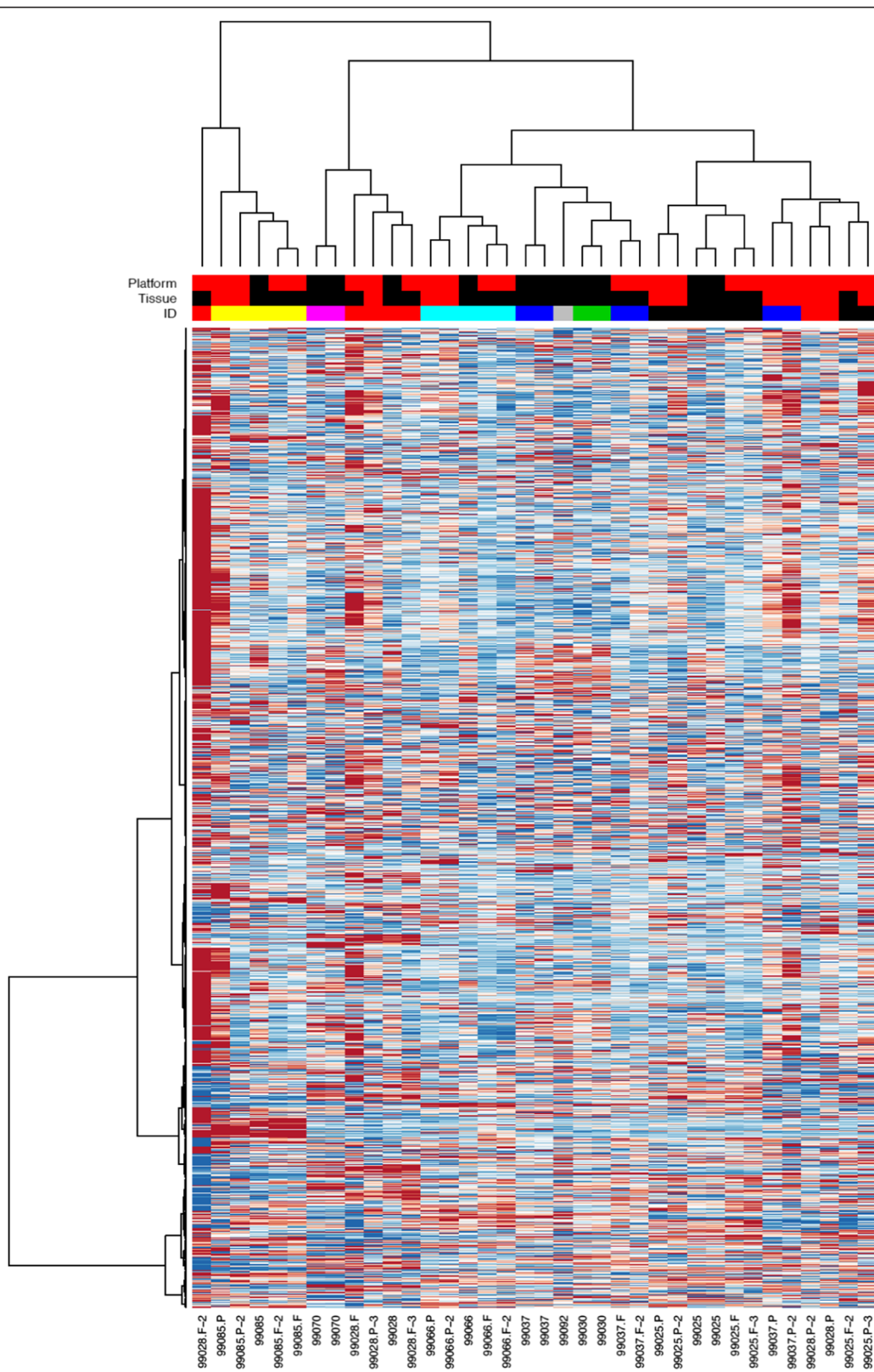
This is one of the first studies to assess matched FF and FFPE tissues with two different gene expression platforms, and to analyse the concordance of data generated. If progress is to be made with the analysis of FFPE tissue, it is imperative that the limitations of the platforms are investigated. Our study showed that comparable profiles and biological information could be derived.

In a similar study to this, April et al. reported that > 88 % of probes were reliably detected using RNA from both sources, which is comparable to our results [7]. The quality and quantity of RNA that can be extracted from FFPE tissue varies [20]. There are many reasons for these variations, including RNase degradation and genomic DNA contamination [21]. Bibikova et al. [6] demonstrated that only 50 % of probes were detected in FFPE

compared to FF tumour specimens using DASL CP; however, there were differences between this and the DASL WG array which could account for these differences [22–24]. Linton et al. used Affymetrix Plus 2.0 arrays and found that only 50 % of the probes detected in FF samples were also identified using RNA from FFPE specimens [25]. This is not surprising as the probes in this array are not specifically designed for FFPE analysis and the assay only uses oligo-dT for reverse transcription [26].

Our study carried out a rigorous bioinformatics assessment of the WG DASL platform to validate its role in assessing FFPE samples. The confounding factors such as the GC content and the probe annealing position on the gene are some of the important factors that should be considered in the data analysis to select appropriate probes for comparison. Retrospectively, we also noted that some failed arrays had lower dye incorporation rates and lower total DNA concentration following the PCR reaction. This can potentially be used as a quality control step in the future to improve the data quality. It was also noted that genes that were shorter were more likely to be reliably detected than longer ones. This is going to have implication for selecting the optimal probes for future array design.

Ravo et al. analysed the gene expression of matched FF and FFPE samples on a DASL CP platform and also of FF samples on an Illumina BeadArray [27]. Their data suggest that DASL CP is more sensitive than the BeadArray,



**Fig. 4** Hierarchical clustering of FF and FFPE samples assessed using the DASL WG assay, and of FF samples assessed using the Illumina BeadArray. This shows significant clustering of FF and FFPE samples, regardless of whether the DASL WG or the BeadArray (indicated by number only) are used

especially in the identification of low abundance transcripts, due to the gene-specific RT-PCR process [27]. These two assays have different dynamic ranges, which also likely contributes to the differences observed. Similar to this study, where matched RNA was profiled using DASL WG, Mittempergher et al. also found that profiled samples clustered according to the RNA source [11].

One of the unique features of our study, compared to previous reports, is that we used matched FF and FFPE samples and assessed them using the same platform (DASL WG). This enabled us to compare the assay and its performance according to RNA source. Other studies, such as that of Ravo et al., used FFPE and matched cell line samples, rather than FF cancer specimens [27]. April et al. used cancer and matched normal tissue [7], and Wardell et al. analysed samples using different assays [9]. Others such as Sadi et al. compared limited number of probes and Abramovitz et al. compared custom panel of genes, limiting the value in assessing performance [1, 19].

Development of robust bioinformatics methods for analysis and interpretation of FFPE data is needed due to the variable quality of RNA [28]. The close clustering of two types of samples shows that reliable data can be generated using FFPE samples, with significant overlap with profiles from FF samples and thus attesting to the power of this technology to obtain reliable profiles from FFPE tissue. Our data confirm those of Mittempergher et al., which showed that samples profiled using DASL WG clustered together, irrespective of the source of RNA. They also observed that hierarchical clustering of the initial data led to FF and FFPE clustering in groups; however, following median centering of the  $\log_2$  transformed data and use of only the 5444 best performing probes, samples clustered according to type [11]. Our study confirms these results in and a second platform was also used to compare results from FF specimens, adding further evidence of the robustness of the assay.

This confirms the potential of the DASL WG platform for FFPE analysis. Many groups have reported that correlations are poorer for genes with low expression levels. This is of interest as the DASL assay relies on gene-specific amplification; therefore, the signal would be expected to be independent of the expression level. However, it is possible that cDNA synthesis and PCR are less efficient for genes expressed at low levels, especially when modification of the 3' UTR could cause marked effects.

## Conclusions

Together with appropriate histopathological criteria, molecular classification of cancers can contribute to better patient stratification. This is one of the first studies to assess the DASL WG assay, and importantly to include a comparison with Illumina BeadArray's. The data

demonstrate that reproducible, biologically relevant signatures can be generated using FFPE specimens and the DASL WG assay. This has marked implications for future research using FFPE materials and will impact studies aimed at molecular marker discovery, especially where only FFPE is available.

## Abbreviations

Ct, cycle threshold; DASL cDNA-mediated, annealing, selection, extension, and ligation; FF, fresh frozen; FFPE, formalin-fixed paraffin-embedded; WG, whole-genome

## Availability of data and materials

Supplementary file available with Gene Ontology data. All the other data has been deposited to a public repository and available to download (<https://figshare.com/s/e8161d813544575c283e>).

## Authors' contributions

MI designed and performed the experiments. MI, OR, MK, and SG performed data analysis. MI, OR, HME and CC wrote the manuscript, with input from all authors. HME and CC supervised the research. MI was funded by Addenbrooke's Charitable Trust and Cancer Research UK. All authors read and approved the final manuscript.

## Competing interests

The authors declare that they have no competing interests.

## Ethics approval and consent to participate

This study was approved by the Cambridge University Hospitals Research Ethics Committee.

## Author details

<sup>1</sup>Cancer Research UK Cambridge Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, UK. <sup>2</sup>Department of Oncology, Addenbrooke's Hospital, University of Cambridge, Hills Road, Cambridge CB1 9RN, UK. <sup>3</sup>Cambridge Breast Unit, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. <sup>4</sup>NIHR Cambridge Biomedical Research Centre and Cambridge Experimental Cancer Medicine Centre, Cambridge, UK. <sup>5</sup>Department of Anatomy and Developmental Biology, Monash University, Clayton, Victoria 3800, Australia. <sup>6</sup>School of Clinical Sciences, Monash University, Clayton, Victoria, Australia.

Received: 10 March 2016 Accepted: 5 August 2016

Published online: 20 August 2016

## References

- Abrahamsen HN, Steiniche T, Nexø E, Hamilton-Dutoit SJ, Sørensen BS. Towards quantitative mRNA analysis in paraffin-embedded tissues using real-time reverse transcriptase-polymerase chain reaction: a methodological study on lymph nodes from melanoma patients. *J Mol Diagn*. 2003;5(1):34–41.
- Abrahamsen HN, Nexø E, Steiniche T, Hamilton-Dutoit SJ, Sørensen BS. Quantification of Melanoma mRNA Markers in Sentinel Nodes: Pre-Clinical Evaluation of a Single-Step Real-Time Reverse Transcriptase-Polymerase Chain Reaction Assay. *J Mol Diagn*. 2004;6(3):253–9.
- Cronin M, Pho M, Dutta D, Stephans JC, Shak S, Kiefer MC, Esteban JM, Baker JB. Measurement of Gene Expression in Archival Paraffin-Embedded Tissues: Development and Performance of a 92-Gene Reverse Transcriptase-Polymerase Chain Reaction Assay. *Am J Pathol*. 2004;164(1):35–42.
- Gunderson KL, Kruglyak S, Graige MS, Garcia F, Kermani BG, Zhao C, Che D, Dickinson T, Wickham E, Bierle J, et al. Decoding Randomly Ordered DNA Arrays. *Genome Res*. 2004;14:870–7.
- Fan J-B, Yeakley JM, Bibikova M, Chudin E, Wickham E, Chen J, Doucet D, Rigault P, Zhang B, Shen R, et al. A Versatile Assay for High-Throughput Gene Expression Profiling on Universal Array Matrices. *Genome Res*. 2004;14:878–85.
- Bibikova M, Talantov D, Chudin E, Yeakley JM, Chen J, Doucet D, Wickham E, Atkins D, Barker D, Chee M, et al. Quantitative gene expression profiling in formalin-fixed, paraffin-embedded tissues using universal bead arrays. *Am J Pathol*. 2004;165(5):1799–807.



7. April C, Klotzle B, Royce T, Wickham-Garcia E, Boyaniwsky T, Izzo J, Cox D, Jones W, Rubio R, Holton K, et al. Whole-genome gene expression profiling of formalin-fixed, paraffin-embedded tissue samples. *PLoS One*. 2009;4(12):e8162.
8. Chien J, Fan JB, Bell DA, April C, Klotzle B, Ota T, Lingle WL, Gonzalez Bosquet J, Shridhar V, Hartmann LC. Analysis of gene expression in stage I serous tumors identifies critical pathways altered in ovarian cancer. *Gynecol Oncol*. 2009;114(1):3–11.
9. Waddell N, Cocciardi S, Johnson J, Healey S, Marsh A, Riley J, da Silva L, Vargas AC, Reid L, Simpson PT, et al. Gene expression profiling of formalin-fixed, paraffin-embedded familial breast tumours using the whole genome-DASL assay. *J Pathol*. 2010;221(4):452–61.
10. Kibriya MG, Jasmine F, Roy S, Paul-Brutus RM, Argos M, Ahsan H. Analyses and interpretation of whole-genome gene expression from formalin-fixed paraffin-embedded tissue: an illustration with breast cancer tissues. *BMC Genomics*. 2010;11:622.
11. Mittempergher L, de Ronde JJ, Nieuwland M, Kerkhoven RM, Simon I, Rutgers EJ, Wessels LF, Van't Veer LJ. Gene expression profiles from formalin fixed paraffin embedded breast cancer tissue are largely comparable to fresh frozen matched tissue. *PLoS One*. 2011;6(2):e17163.
12. Rentoft M, Laurell G, Coates PJ, Sjoström B, Nylander K. Gene expression profiling of archival tongue squamous cell carcinomas provides sub-classification based on DNA repair genes. *Int J Oncol*. 2009;35(6):1321–30.
13. Tarraga J, Medina I, Carbonell J, Huerta-Cepas J, Minguez P, Alloza E, Al-Shahrour F, Vegas-Azcarate S, Goetz S, Escobar P, et al. GEPAS, a web-based tool for microarray data analysis and interpretation. *Nucleic Acids Res*. 2008;36(Web Server issue):W308–314.
14. Cairns JM, Dunning MJ, Ritchie ME, Russell R, Lynch AG. BASH: a tool for managing BeadArray spatial artefacts. *Bioinformatics*. 2008;24:2921–2.
15. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, et al. A Multigene Assay to Predict Recurrence of Tamoxifen-Treated, Node-Negative Breast Cancer. *N Engl J Med*. 2004;351(27):2817–26.
16. Nielsen TO, Parker JS, Leung S, Voduc D, Ebbert M, Vickery T, Davies SR, Snider J, Stijleman IJ, Reed J, et al. A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor-positive breast cancer. *Clin Cancer Res*. 2010;16(21):5222–32.
17. Greytak SR, Engel KB, Bass BP, Moore HM. Accuracy of Molecular Data Generated with FFPE Biospecimens: Lessons from the Literature. *Cancer Res*. 2015;75(8):1541–7.
18. Duffy MJ. Predictive Markers in Breast and Other Cancers. *Clin Chem*. 2005;51(3):494–503.
19. Sadi AM, Wang DY, Youngson BJ, Miller N, Boerner S, Done SJ, Leong WL. Clinical relevance of DNA microarray analyses using archival formalin-fixed paraffin-embedded breast cancer specimens. *BMC Cancer*. 2011;11:253:251–13.
20. Li J, Smyth P, Cahill S, Denning K, Flavin R, Aherne S, Pirota M, Guenther SM, O'Leary JJ, Sheils O. Improved RNA quality and TaqMan Pre-amplification method (PreAmp) to enhance expression analysis from formalin fixed paraffin embedded (FFPE) materials. *BMC Biotechnol*. 2008;8:10.
21. Waldron L, Ogino S, Hoshida Y, Shima K, McCart Reed AE, Simpson PT, Baba Y, Noshko K, Segata N, Vargas AC, et al. Expression profiling of archival tumors for long-term health studies. *Clin Cancer Res*. 2012;18(22):6136–46.
22. Srivastava PK, Küffer S, Brors B, et al. A cut-off based approach for gene expression analysis of formalin-fixed and paraffin-embedded tissue samples. *Genomics*. 2008;91:522–529.
23. Gilbert MT, Haselkorn T, Bunce M, Sanchez JJ, Lucas SB, Jewell LD, Van Marck E, Worobey M. The isolation of nucleic acids from fixed, paraffin-embedded tissues-which methods are useful when? *PLoS One*. 2007;2(6):e537.
24. Specht K, Richter T, Müller U, Walch A, Werner M, Hofler H. Quantitative Gene Expression Analysis in Microdissected Archival Formalin-Fixed and Paraffin-Embedded Tumor Tissue. *Am J Pathol*. 2001;158(2):419–29.
25. Linton KM, Hey Y, Saunders E, Jeziorska M, Denton J, Wilson CL, Swindell R, Dibben S, Miller CJ, Pepper SD, et al. Acquisition of biologically relevant gene expression data by Affymetrix microarray analysis of archival formalin-fixed paraffin-embedded tumours. *Br J Cancer*. 2008;98(8):1403–14.
26. Bibikova M, Talantov D, Chudin E, Yeakley JM, Chen J, Doucet D, Wickham E, Atkins D, Barker D, Chee M, et al. Quantitative Gene Expression Profiling in Formalin-Fixed, Paraffin-Embedded Tissues Using Universal Bead Arrays. *Am J Pathol*. 2004;165(5):1799–1807.
27. Ravo M, Mutarelli M, Ferraro L, Grober OM, Paris O, Tarallo R, Vigilante A, Cimino D, De Bortoli M, Nola E, et al. Quantitative expression profiling of highly degraded RNA from formalin-fixed, paraffin-embedded breast tumor biopsies by oligonucleotide microarrays. *Lab Invest*. 2008;88(4):430–40.
28. Waldron L, Simpson P, Parmigiani G, Huttenhower C. Report on emerging technologies for translational bioinformatics: a symposium on gene expression profiling for archival tissues. *BMC Cancer*. 2012;12:124.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

