Comparative Quantitation of Lymphoma Gene Signatures in Parallel Formalin Fixed Paraffin Embedded and Frozen Tissue Lymph Nodes

Ebrahim Sakhinia1,2*, Jordan Goodchild3, Lia Patricia Menasce4, Joan Radford5, Mehrdad Asghari Estiar1, Masoud Sakhinia4 and Richard Byers6

1Department of Medical Genetic, Faculty of Medicine, Tabriz Molecular Analysis Centre (TGAC) and Tuberculosis and Lung Diseases research center, Tabriz university of Medical Science, Tabriz, Iran  
2Division of Regenerative Medicine, School of Medicine, Faculty of Medical and Human Sciences, the University of Manchester, UK  
3Department of Histopathology, Manchester Royal Infirmary, Manchester, UK  
4Department of Histopathology, Christie Hospital NHS Trust, Manchester, UK  
5Cancer Research UK Department of Medical Oncology, The University of Manchester and Christie Hospital NHS Trust, Manchester, UK  
6Faculty of Medicine, University of Liverpool, United Kingdom

Abstract

Microarray gene expression profiling of frozen samples has identified Indicator genes predictive outcome in diffuse large B-cell lymphoma (DLBCL) and follicular lymphoma (FL) and there is a now a need to translate these to use in routine clinical material. In order to evaluate the use of Indicator genes as a diagnostic tool in routinely processed formalin fixed paraffin embedded tissue (FFPET), we used a simple, practical polyA PCR based method for analysis of Indicator genes in DLBCL and FL, using parallel FFPET and frozen tissue lymph nodes. PolyA PCR was used to generate polyA cDNAs from 42 archival FFPET lymph nodes and 42 parallel frozen tissue lymph nodes. These polyA cDNAs were assayed by real-time PCR for 35 Indicator genes, identified from previous microarray studies, predictive of outcome in FL and DLBCL, together with housekeeping genes. Housekeeping genes were detectable at high levels in all cases, though the 35 Indicator genes were expressed at lower levels. There was statistically significant difference in gene expression between DLBCL and FL for one of the Indicator genes, HSP27 (p ≤ 0.02). There was no correlation between the Indicator gene expression levels in the FFPET and frozen lymph nodes. The results demonstrate that gene expression in FFPET is altered from that in frozen tissue and that gene expression data from frozen tissue is therefore not transferable for use in FFPET.

Keywords: Lymphoma; Indicator genes; Formalin fixed paraffin embedded tissue; PolyA PCR; Real-time PCR

Introduction

Recently it has become clear that the development of haematological malignancy may depend almost solely on disturbances in normal gene transcription within cells [1,2]. This breakdown in normal control can be quantified in the laboratory by measuring mRNA expression levels and a number of important studies have identified, through global gene expression profiling using microarray technology, small groups of genes for both FL [3-5] and DLBCL [6] predictive of outcome or response to treatment, so called Indicator genes. Although, the above microarray studies have suggested the possibility of more precise cancer diagnosis [7], verification of the data in larger or more diverse sample sets is hindered by the need for fresh or frozen tissue and for relatively large amounts of tissue. However, most archival and clinical tissue is in the form of formalin fixed paraffin embedded tissue (FFPET) which, with its attendant long-term follow-up and inclusively for all tumor types, is therefore becoming evermore important as a source of research material [8] and it would greatly facilitate and speed up research in this area [9]. Furthermore introducing gene expression profiling into routine oncology could revolutionise the way in which we manage cancers, improving patient outcomes.

There is therefore an urgent need to develop methods for measurement of mRNA expression in FFPET. Previous attempts have reported varying degrees of success but have used methods directed against a limited number of specific genes [10-12] or against abundant genes [13], which are of limited relevance for global transcriptional analyses.

In order to incorporate gene expression profiling of tumours into routine clinical practice, a robust, reproducible, reliable and cost effective procedure is required. PolyA reverse-transcriptase polymerase chain reaction (PolyA RT-PCR) enables globally amplification of cDNA transcripts representative of all poly adenylated mRNAs from a given sample whilst maintaining the relative ratios of each of the transcripts [14]. The technique is of sufficient sensitivity to amplify sub-picogram amounts of mRNA from single cells [15], and is therefore ideal for analysis of clinical samples, which given the increased use of needle core and fine needle aspiration samples, are of increasingly small size. We have previously successfully used this method to extract mRNA from FFPET [16] and for measurement of lymphoma associated Indicator genes in archived frozen samples of FL and DLBCL [17].

In this study we tested the hypothesis that polyA PCR could produce globally amplified cDNA from RNA extracted from formalin fixed paraffin embedded samples. The specific aims were to: 1) To extract RNA from FFPET samples from patients with either FL or DLBCL, and for whom we have mRNA expression data from parallel frozen tumour samples. 2) To produce a polyA cDNA library using polyA RT-PCR. 3) To amplify cDNA using RTQ-PCR and subsequently measure the expression levels of 35 FL and DLBCL Indicator genes [1,5,6] in each sample.

*Corresponding author: Dr. Ebrahim Sakhinia, Department of Medical Genetic, Faculty of Medicine, Tabriz Molecular Analysis Centre (TGAC), Tuberculosis and Lung Diseases research center, Tabriz university of Medical Science, Tabriz, Iran, E-mail: esakhinia@yahoo.co.uk

Received August 24, 2012; Accepted September 24, 2012; Published September 26, 2012


Copyright: © 2012 Sakhinia E, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
sample. 4) To test the use of these gene expression profiles as classifiers of FL versus DLBCL. 5) To compare mRNA expression between archival frozen and FFPE samples from the same patients. The last two aims will test of the feasibility of switching from frozen to FFPE samples for routine gene expression analysis in FL and DLBCL.

**Materials and Methods**

**Tissue specimens**

Fifty archived FFPE lymph nodes (with at least 5 years follow up) with a diagnosis of either FL or DLBCL were obtained, with informed consent, from the archives of the Christie Hospital NHS Trust, Manchester, UK. The samples were selected from the archive on a sequential chronological basis, taking all samples for which consent was given, in the archive from 1995 backwards to 1989. Paraffin embedded sections from all cases were reviewed by two pathologists (RJB & LPM) and those unlikely to have sufficient material for analysis, with more than 10% necrosis or with uncertain or complex/mixed diagnoses were excluded. Parallel archived frozen tissue samples from the same patients had undergone mRNA expression analysis in a previous study [17]. A summary of the sample details is given below in table 1; total mRNA was extracted from 50 FFPE lymph node samples, 42 of which yielded RNA in enough quantity to perform gene expression analysis.

For each case the FFPE block was cooled on ice for twenty minutes, following which eight 10 μm thick sections were cut using a microtome. The eight sections from each block were then equally split between two 1.5 ml Eppendorf tubes (Starlab), such that each tube contained four rolled 10 μm sections. To allow for optimum extraction, the surface area of the tissues was ensured not to exceed 1cm.

**Total RNA isolation from samples FFPE**

The ‘Optimum™ FFPE RNA Isolation kit’ (Ambion ; Cat No. 47000) was used for RNA extraction, according to manufacturer’s instructions. Total RNA was extracted from 50 FFPE lymph node samples, 42 of which yielded RNA in enough quantity to perform gene expression analysis using RTQ-PCR. Total RNA from these 42 samples was visualised on 0.6% agarose gels.

**Poly Adenylated mRNAs (PolyA RT-PCR)**

Global amplification of cDNA corresponding to all expressed genes (polyA PCR), was carried out as previously reported [17]

**Real-time quantitative RT-PCR**

Taqman™PCR primers and probes were designed for each gene using Primer Express Software (Perkin Elmer/Applied Biosystems) and are listed in Table 2. For each gene, Taqman™PCR was applied to 1ng polyA cDNA from each sample and to 10 microlitre of all serial human genomic standards, using a Taqman™Gold kit as recommended by the manufacturer. Samples were analysed using an ABI Prism 7000 sequence detection system (PE Applied Biosystems) as recommended by the manufacturer and as described previously [18].

**Data Normalisation**

The expression levels of four housekeeping genes (β2-β, GAPDH, Beta actin and human ribosomal protein S9 mRNA) were measured by RT-PCR in each sample. Copy numbers obtained for the mean (Mhouse) of the four housekeeping genes in each sample were divided by the highest Mhouse in all samples resulting in a normalisation correction factor. Following real-time PCR amplification and quantification of the selected genes, this factor was then used for normalisation of expression levels of each of the 35 genes measured.

**Statistical analysis**

The data was not normally distributed and therefore non-parametric tests were used. Statistical analysis of the expression levels of the 35 genes was performed using the Mann Whitney test and Kruskal-Wallis with a p ≤ 0.05 for statistical significance. All the tests were performed using SPSS software v13 (Woking, Surrey, UK). As the expression levels for all the genes falls across a large range, the natural logarithm (LN) was used to plot all values to aid visual comparison. The ranking of the expression levels of Indicator genes between samples was presented as the mean rank statistical difference [18]. Spearman Correlation analysis was also performed.

**Results**

**RNA extraction and reverse transcription**

In order to validate the polyA PCR amplification step, the product was assessed by three means, as described previously [16], to ensure that the method was reliable and that the product could be used confidently in subsequent probing for Indicator genes. Firstly the polyA cDNA was visualised using gel electrophoresis (Figure 1A). Secondly, normal reverse-transcriptase PCR (RT-PCR) was performed on GAPDH and this was viewed using gel electrophoresis (Figure 1B,1C); the no RT control was negative (Figure 1B). Finally, as a quantitative measure, real time-quantitative PCR (RTQ-PCR) was performed for GAPDH.

**Quantification of Housekeeping Gene Expression from FFPE samples**

All samples showed strong expression of GAPDH, Ribosomal S9 and β-actin, although expression of IF2-β was less (Figure 2). However, there was no statistically significant difference in the value of Mhouse for the different diagnostic groups (p ≥ 0.60).

**Quantification and Comparison of FL or DLBCL Indicator Gene Expression in FFPE samples**

The expression levels of 35 Indicator genes for FL and DLBCL [1,5,6] were quantified and compared in FFPE samples (Table 2); these genes had been measured in parallel frozen samples in a previous study and were therefore chosen for purposes of comparison. The copy number for each gene in each sample was reproducible over at least duplicate tests demonstrating reliability of the method. Following normalisation of the data, only one gene, HSP27, showed significant difference in expression between FL and DLBCL (P ≤ 0.02), with higher expression in FL (Figure 3A). Comparison of the expression levels of each of the 35 Indicator genes measured in both FL and DLBCL is shown by means of the mean rank statistic in Figure 3B.

**Comparison of expression levels of FL and DLBCL indicator genes between parallel FFPE and Frozen Tissue samples**

Parallel archival frozen lymph node samples had previously been analysed for mRNA expression for the same genes as measured in the FFPE samples. The FFPE gene expression data for each sample was compared with that from the corresponding frozen samples in order to determine whether gene expression levels in FFPE was comparable with that of parallel frozen samples (Figure 4).

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Number</th>
<th>Average age of samples (yrs)</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>FL</td>
<td>16</td>
<td>13.4</td>
<td>All dead</td>
</tr>
<tr>
<td>DLBCL</td>
<td>24</td>
<td>13.6</td>
<td>All dead</td>
</tr>
<tr>
<td>Total</td>
<td>42</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Summary of FFPE sample details.
<table>
<thead>
<tr>
<th>Genes</th>
<th>Accession Number</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Beta Actin</td>
<td>NM_001101</td>
<td>Housekeeping gene</td>
</tr>
<tr>
<td>2 Ribosomoal S9</td>
<td>U14971</td>
<td>Housekeeping gene</td>
</tr>
<tr>
<td>3 IF2b</td>
<td>U23028</td>
<td></td>
</tr>
<tr>
<td>4 GAP</td>
<td>NM_002046</td>
<td></td>
</tr>
<tr>
<td>5 S-hT2B</td>
<td>X77307</td>
<td></td>
</tr>
<tr>
<td>6 h731</td>
<td>NM_145341</td>
<td></td>
</tr>
<tr>
<td>7 PKC-b-1</td>
<td>BC036472</td>
<td></td>
</tr>
<tr>
<td>8 OVGL</td>
<td>U09550</td>
<td></td>
</tr>
<tr>
<td>9 PKCG</td>
<td>BC047876</td>
<td></td>
</tr>
<tr>
<td>10 ZFPC150</td>
<td>BC022785</td>
<td></td>
</tr>
<tr>
<td>11 TLE1</td>
<td>M99435</td>
<td></td>
</tr>
<tr>
<td>12 MINOR</td>
<td>U12767</td>
<td></td>
</tr>
<tr>
<td>13 ACTA</td>
<td>J05192</td>
<td></td>
</tr>
<tr>
<td>14 NPM3</td>
<td>NM_006993</td>
<td></td>
</tr>
<tr>
<td>15 Bcl-6</td>
<td>U00115</td>
<td></td>
</tr>
<tr>
<td>16 c-myc</td>
<td>D89667</td>
<td></td>
</tr>
<tr>
<td>17 HLA-DQ-a</td>
<td>U77589</td>
<td></td>
</tr>
<tr>
<td>18 HLA-DR-a</td>
<td>K01171</td>
<td></td>
</tr>
<tr>
<td>19 KIAA0233</td>
<td>D87071</td>
<td></td>
</tr>
<tr>
<td>20 Urokinase</td>
<td>U08839</td>
<td></td>
</tr>
</tbody>
</table>

**Genes Accession Number**

1=5'-TaqMan forward primer - 3'  
2= 5'-TaqMan revers primer - 3'  
3=5'-TaqMan(FAM)-probe-(TAMARA)-3'

**References**

Shipp et al. 2002 [1]

1=CCACCCCACTTCTCTCTAAAGGA
2=CATAATTTACAGAAAGCAATGCTATC
3=TTGCCCAGCTCCTCTCCCAAGTCC

1=GAGACGACGAGGAGGAGGAGGATTAA
2=GCAGGAAGAAACGAGACACATCCA
3=TCACCTGCTGCCCTGGGCTCTG

1=TGGAGTTGGGATGTGGAGTGGAAGTG
2=GGGACTCCCCAGACATGGA
3=ACCAGCCCCAAGAGACACAGAAG

1=ATGGGATTAACCTGCACTGTGA
2=GAGAAGCTATCTAGAGATGATG
3=TTCAATAGGGCTCCAGATTTCTCACC

1=CTTCTTCTTAAAGGACAGTGTTTTTGT
2=TCCTCAAAACGATCTCCTACAG
3=TAAGCTGCTAAAAACCCATGTGCTCT

1=GGGACTCCCCCTGCAAGAC
2=GAGAAGCTATCTAGAGATG
3=TTCAATAGGGCTCCAGATTTCTCACC

1=GCCATCAAGCTCCTGTCA
2=GAGATTGCGCTATCTGATTTG
3=TTCAATAGGGCTCCAGATTTCTCACC

1=GGCTCCCAGCCCTACAGTT
2=GCTGTTGGTCGACCCAGAA
3=AGGCCGCTTCGCCGTCACT

1=CCAAGCCCAAGGAGACAGA
2=GGGACTCCCCCTGCAAGAC
3=TTCAATAGGGCTCCAGATTTCTCACC

1=GGGACTCCCCCTGCAAGAC
2=GAGAAGCTATCTAGAGATG
3=TTCAATAGGGCTCCAGATTTCTCACC

1=CATGCCCACTGCAAGAC
2=GCTCCTGCTGCAGCTCTCAACG
3=TCACCTCCTCACCTCCAGAAAC

1=CCGGAGCGCAAATACTCTGT
2=CCCGGCTTCATCGTATTCCT
3=ATCGGTGGCTCCATCCTGGCCT

1=CCGAGGCCAAAGGACACAGA
2=CCAGATACCCCAAGAAAGAAGAA
3=AACGGGATCCTGCCGACTCATGACACA

1=GATGACTTCCAGCTCCAACCA
2=CACGCTCTCAGGCCTATCACG
3=AACGGGATCCTGCCGACTCATGACACA

1=CGCCATCTACAGGAGCAGAA
2=AGCAGCTTACAGGGCCTTAAAG
3=CCGAGGCCAAAGGACACAGA

1=GACAGCCAAACCCCTGCT
2=GCCCTCTGTCTTCGCCCTACCCCT
3=TTCAATAGGGCTCCAGATTTCTCACC

1=CGGAGACCAGCAAGAGAGT
2=CCGAGGCCAAAGGACACAGA
3=CCGAGGCCAAAGGACACAGA

1=CCGAGGCCAAAGGACACAGA
2=CCGAGGCCAAAGGACACAGA
3=CCGAGGCCAAAGGACACAGA

1=CAGGACGAGGACTGTTGACT
2=ATCGGTGGCTCCATCCTGGCCT
3=ATCGGTGGCTCCATCCTGGCCT

1=CCGAGGCCAAAGGACACAGA
2=CCGAGGCCAAAGGACACAGA
3=CCGAGGCCAAAGGACACAGA

1=GAGAAGCTATCTAGAGATG
2=TTCAATAGGGCTCCAGATTTCTCACC
3=TTCAATAGGGCTCCAGATTTCTCACC

1=GGCTCCCAGCCCTACAGTT
2=GCTGTTGGTCGACCCAGAA
3=AGGCCGCTTCGCCGTCACT
| 21 | BMP6 | NM_001718 | 1=AGTGTCGCCACCTCTCTGGCCCT 2=GTACTAGCTCACTCCTCTCTTT | Rosenwald et al. 2002 [5] |
| 22 | BCL-2 | M14745 | 1=TGAGTAAATCCATGCACTAACCTTT 2=CACCTGCACTCTTGCAATAATCTACCT | Husson et al. 2002 [6] |
| 23 | COL3A1 | NM_000090 | 1=CAAATCTGTCATGTCCTTCAATCAG 2=CAACATATCTCTGTAGTGTGACAA 3=TGCCCAAGTCTCTCTCCTCCTCCTACCT | Husson et al. 2002 [6] |
| 24 | CDC2 | NM_052987 | 1=CCTGGCAGACTCCAGATGA 2=GGAGCAGGGATCCAGCTCA 3=AGCCACAATTGAGTACCTGACGCT | Husson et al. 2002 [6] |

All accession numbers were taken from BLAST.

1=TaqMan primer 1 (5′–3′); 2=TaqMan primer 2 (5′–3′); 3=TaqMan [6-carboxyfluorescein (FAM)]-probe-[6-carboxyteramrthylrhodamine (TAMARA)] (5′–3′).

Table 2: List of primers and probes used.
Correlation analysis between parallel FFPET and frozen tissue samples demonstrated statistically significant correlation for HSF1 (correlation coefficients -0.35, p ≤ 0.036), P21cp1 (correlation coefficients -0.32, p ≤ 0.054) and GAPDH alone showed positive correlation (correlation coefficients 0.31, p ≤ 0.051) (Figure 5).

**Discussion**

The aims of the study were to determine the feasibility of mRNA extraction, polyA PCR global cDNA amplification and RTQ-PCR of gene expression profiling from routinely processed FFPET blocks from patients with FL and DLBCL and to compare the gene expression profiles obtained with those in parallel frozen samples. The results demonstrate that the methodology is suitable for extraction and the amplification of mRNA from FFPET samples but that the gene expression profile in the resultant cDNA differs from that in cDNA prepared from frozen samples.

The main obstacle to using FFPET samples in routine molecular diagnosis is degradation caused to nucleic acids within the sample by the process of fixation, which both cross links and fragments the mRNA present. Frequent breaks in the RNA results in a number of intact amplicons to be too low for amplification [19], whilst with increasing age, more breaks may appear in the RNA, leading to the number of intact amplicons to be too low for amplification of the gene of interest [19]. We have previously addressed this problem using polyA amplification to increase the amount of material for analysis, followed by real-time PCR using primers and probes for relatively short sequences within 300 bp of the 3’ end. Since formalin fixation results in mRNA fragments approximately 200 bp long, this strategy is capable of successful analysis [16].

The results demonstrate that all four housekeeping genes measured were expressed but that three of the four, namely GAPDH, β-Actin and Ribosomal S9 were expressed at higher levels than IF2-β. Since housekeeping genes, by definition, are constitutively expressed by all cells, the likely explanation for this is that the IF2-β gene is more vulnerable to disruption during formalin fixation than the other housekeeping genes. This could be due to a structural difference in the gene, or maybe because the gene is expressed at lower levels constitutively in cells.

It should be stressed however that successful amplification of a housekeeping gene from cDNA does not ensure that Indicator gene expression will be unaffected by fixation in FFPET samples. By definition, housekeeping gene expression is uniformly high in all cells, and so even if extensive degradation had occurred, they may still be detectable after PCR. Conversely, formalin fixation may completely destroy genes that are expressed at only low levels in cells, such as some Indicator genes, and thus these would not be detectable. The FFPET samples used in this study were all over five years old. It is possible that RNA stored in FFPET does not in fact degrade during the fixation process itself, but degrades whilst in the paraffin over time [20]. If this were to be the case, then gene expression analysis from FFPET may be more feasible from FFPET of a younger age.

Despite the above concerns mRNA was successfully amplified and the gene expression profile of 35 Indicator genes measured in the resultant cDNA. However, only one of the 35 genes measured, Heat shock protein 27 (HSP27), demonstrated statistically significant (p ≤ 0.02) increased expression in FL compared to DLBCL, indicating that

![Figure 1A: Gel Electrophoresis of total RNA isolated from FFPET samples.](image1)

![Figure 1B: Gel Electrophoresis of PolyA cDNA Product.](image2)

![Figure 1C: Gel Electrophoresis of the GAPDH PCR products amplified from the PolyA cDNAs.](image3)

![Figure 2: Expression levels (LN) of each of the four housekeeping genes and for the mean of four (Mhouse). The expression levels of the four housekeeping genes are shown in natural log scale (LN). L= Natural Logarithm, Mhouse = Mean of the housekeeping genes.](image4)
**Figure 3A:** HSP27 expression levels in both FL and DLBCL, Mhouse expression showed as control, LN=Natural Logarithm, Mhouse=Mean of the housekeeping genes.

**Figure 3B:** Expression patterns of 35 indicator genes plus four housekeeping genes in FL and DLBCL.

**Figure 4:** Comparison of expression levels of (A) FL and (B) DLBCL indicator genes between parallel FFPET and Frozen Tissue samples.
A

Figure 5: Correlation analysis between parallel FFPET and archived frozen tissue samples.

this was the only gene capable of differentiating between the two types of lymphoma in this study.

Using archival data from parallel frozen samples, correlation analysis was performed to statistically compare gene expression of 35 Indicator genes in frozen versus FFPET tissue from identical patients (Figure 5). The results showed that gene expression in FFPET was altered from that of frozen tissue, with positive correlation between the level of gene expression in frozen samples and the level in the corresponding FFPET sample for GAPDH alone, and that therefore gene expression data from frozen tissue studies was not transferable to use in FFPET. This inability to translate between the two gene expression profiles means that if FFPET were to replace the use of frozen tissue in the future in routine use, then new Indicator genes will need to be identified from FFPET directly in future studies. These findings have reiterated conclusions from previous studies, which all indicate that gene expression in FFPET samples is not comparable to that of frozen samples. For example, Chen et al. [21] identified 63 genes whose expression levels were significantly different between hepatocellular adenoma and well differentiated hepatocellular carcinoma using microarray studies. When these genes were further analysed using RTQ-PCR from parallel FFPET samples, only eight of the original 63 genes were found to be expressed in concordance with the microarray results. More significantly, our previous work demonstrated that in the 42 parallel frozen samples for which Indicator gene expression was quantified, ten of the 35 genes showed significant differences in expression between FL and DLBCL [16]. However, in this study we found that in the corresponding FFPET samples, only one of the 35 genes was significantly expressed. This difference is due to the significantly increased degradation of RNA extracted from FFPET, causing the discriminative power of each gene to drop when using FFPET.

Overall this study has demonstrated the feasibility of extracting and globally amplifying RNA from routinely processed, human FFPET samples, using polyA RT-PCR. RTQ-PCR can then be successfully performed for both housekeeping genes and Indicator genes. However, the expression levels of these genes differs from that in parallel frozen samples and therefore, despite the ability to extract mRNA, carry out polyA PCR and measure gene expression in FFPET, Indicator genes identified in frozen samples cannot be used for prediction of outcome in FFPET.

References


