Quantitative nuclease protection assay in paraffin-embedded tissue replicates prognostic microarray gene expression in diffuse large-B-cell lymphoma

Robin A Roberts[1], Constantine M Sabalos[2], Michael L LeBlanc[3], Ralph R Martel[2], Yvette M Frutiger[1], Joseph M Unger[3], Ihab W Botros[2], Matthew P Rounseville[2], Bruce E Seligmann[2], Thomas P Miller[4], Thomas M Grogan[1] and Lisa M Rimsza[1]

Gene expression profiling (GEP) has identified genes whose expression levels predict patient survival in diffuse large-B-cell lymphoma (DLBCL). Such discovery techniques generally require frozen samples unavailable for most patients. We developed a quantitative nuclease protection assay to measure expression levels of prognostic DLBCL genes using formalin-fixed, paraffin-embedded (FFPE) tissue. FFPE tissue was sectioned, permeabilized, denatured in the presence of specific probes, and hybridized to mRNA *in situ*. Nuclease subsequently destroyed non-hybridized probe. Alkaline hydrolysis freed mRNA-bound probes from tissue, which were transferred to ArrayPlates for probe capture and chemi-luminescent quantification. We validated assay performance using frozen, fresh, and FFPE DLBCL samples, then used 39 archived DLBCL, previously microarray analyzed, to correlate GEP and ArrayPlate results. We compared old (>18 years) with new (<2 months) paraffin blocks made from previously frozen tissue from the original biopsy. ArrayPlate gene expression results were confirmed with immunohistochemistry for BCL2, BCL6, and HLA-DR, showing agreement between mRNA species and the proteins they encode. Assay performance was linear to ~ 1 mg sample/well. RNase and DNase treatments demonstrated assay specificity for RNA detection, both fixed and soluble RNA detection. Comparisons were excellent for lysate *vs* snap-frozen *vs* FFPE ($R^2 > 0.98$ for all comparisons). Coefficients of variation for quadruplicates on FFPE were generally $<20\%$. Correlation between new and old paraffin blocks from the same biopsy was good ($R^2 = 0.71$). Comparison of ArrayPlate to Affymetrix and cDNA microarrays showed reasonable correlations. Insufficient power from small sample size prevented successfully correlating results with patient survival, although hazard ratios trended the expected directions. We developed an assay to quantify expression levels of survival prediction genes in DLBCL using FFPE, fresh, or frozen tissue. While this technique cannot replace GEP for discovery, it indicates that expression differences identified by GEP can be replicated on a platform applicable to archived FFPE samples.
Laboratory Investigation (2007) **87**, 979–997; doi:10.1038/labinvest.3700665; published online 13 August 2007

KEYWORDS: DLBCL; lymphoma; nuclease protection assay; paraffin-embedded tissue; prognosis; gene expression profiling

Diffuse large-B-cell-lymphoma (DLBCL) is the most common type of aggressive lymphoma, accounting for nearly 40% of lymphomas overall. The clinical International Prognostic Index (IPI) score for DLBCL helps to stratify patients into risk categories based on clinical features.[1] However, even within risk categories, patient outcome is variable. Because of the variable patient outcome, it has long been suspected that DLBCL may actually encompass more than one disease. In recent years, the results of extensive gene expression profiling (GEP) experiments have confirmed that DLBCL is not a uniform disease entity and that differences in GEP-defined subtypes can be related to patient survival.[2–4] GEP has yielded four different sets of prognostic genes based on the work of four different research groups.

The first group, known as the Leukemia and Lymphoma Molecular Profiling Project (LLMPP), studied 240 DLBCL using the LymphoChip. This chip was a spotted oligonucleotide microarray containing numerous elements known

to represent genes expressed by B- or T-lymphoid cells, genes involved in the immune response, and genes expressed by lymphoma and leukemia cell lines.[5] The investigators discovered four gene expression signatures highly correlated to patient survival called 'germinal center,' 'major histocompatibility (MHC) class II,' 'lymph node,' and 'proliferation.' Representative genes from these four signatures with the addition of the gene *BMP6* were used to create a 17-gene outcome predictor score, which provided additional prognostic value independent of the clinical IPI score.[2] Another group used the Affymetrix high-density oligonucleotide array platform and a supervised learning classification approach to develop a DLBCL outcome predictor score with best accuracy obtained using a set of 13 genes.[3] A third group of investigators performed a meta-analysis of the literature for genes that were previously reported to be significantly associated with survival in DLBCL. Using quantitative real-time PCR (qRT-PCR), these researchers assessed a series of 66 DLBCL and determined the six most predictive genes (*LM02, BCL6, FN1, CCND2, SCYA3,* and *BCL2*).[4] Only two genes were common to both the LymphoChip and qRT-PCR studies (*BCL6* and *FN1*). Thus, a total of 34 genes were previously identified in these studies as being significantly associated with patient outcome in DLBCL. Through our own analysis of the LymphoChip data set, we had previously identified several redox-related genes highly correlated with patient survival and had created a 'redox score,' including the representative genes, manganese superoxide dismutase (*SOD2*) and catalase (*CAT*).[6] All four of these studies used quantification of a small key set of genes that correlate with survival in DLBCL in their particular set of cases. All used techniques that relied on snap-frozen materials as the basis of analysis, making comparisons between the gene sets or expansion to wider groups of patients difficult. Thus, development of a method that is applicable to paraffin-embedded samples might be helpful in further exploring the prognostic utility of these and other, yet to be identified, gene sets.

Quantitative nuclease protection assays hold the promise of analyzing mRNA without the difficult steps of mRNA extraction followed by qRT-PCR and therefore may be applicable to formalin-fixed, paraffin-embedded (FFPE) tissues. We chose to use the ArrayPlate assay from High Throughput Genomics Inc. (HTG, Tucson, AZ, USA) that has been described previously for pharmaceutical discovery and development applications.[7] In this assay, thin tissue sections are treated with a permeabilizing/lysing buffer (HTG lysis buffer) and can be tested immediately or frozen for later analysis. Oligonucleotide probes designed for the genes of interest are incubated with the lysate to allow hybridization of specific DNA probe–mRNA heteroduplexes. Unhybridized probes and RNA are digested by a nuclease specific for single-stranded nucleic acids. Alkaline hydrolysis is then used to destroy the mRNA component from the heteroduplexes, leaving intact probes in concentrations proportional to the amounts of specific mRNA originally present. These remaining probes are then transferred to an ArrayPlate, which uses an oligonucleotide array and a sandwich hybridization linker to capture and label the probes for chemiluminescent detection and quantification. Originally designed for high-throughput screening, the assay has the potential to handle large sample numbers using small amounts of variously preserved tissue.

Using this technique, we developed an assay for all 36 genes of interest as well as housekeeping genes and genes aimed at determining the cellular composition of the sample (B cell, T cell, and macrophage genes). We performed several studies to test and validate the performance of the assay in both frozen and paraffin-embedded materials. We then turned to a subset of paraffin blocks from DLBCL cases for which matching frozen tissue had been previously analyzed by GEP, described in the study by Rosenwald *et al* to determine whether the ArrayPlate technique could similarly quantify the prognostically significant genes and whether this would correlate with the microarray data. The overall goal was to develop and validate an assay that would be useable for outcome prediction in all patients with DLBCL, not limited to the minority with available snap-frozen tissue. In this study, we describe our technique, compare fresh and frozen samples to paraffin samples, compare new paraffin blocks to nearly 20-year-old paraffin blocks, and compare the relationship between the results and microarray results as well as patient hazard ratio of death in a previously published set of 40 patients. These studies demonstrate the usefulness of this technique especially as applied to archived materials with potential consequent far-reaching applications for the field.

MATERIALS AND METHODS
Cell Line and Patient Materials, Preparation of FFPE and Frozen Blocks
To validate the design of the ArrayPlate, we began by using the RJ2.2.5 (RJ) cell line,[8] which is an MHC II-negative non-Hodgkin's lymphoma cell line. We also used a recently created RJ2.2.5 transfectant, which includes a CIITA expression vector that induces expression of MHC class II genes, including HLA-DRA, -DRB, -DPA, and -DQA.[9] The cells were grown under 5% CO_2 in RPMI media (with 10% fetal bovine serum, 1 mM sodium pyruvate, 2 mM L-glutamine, 100 U/ml penicillin, and 100 µg/ml streptomycin). Cells were initially seeded at 1×10^5 cells/ml, allowed to grow for 24 h, then harvested. Sixteen million cells were centrifuged to create cell pellets that were either (1) made into FFPE material by fixing in formalin for 4 h, routine overnight tissue processing on a Tissue-Tek (Sakura, McGaw Park, IL, USA) instrument, and paraffin embedding; (2) made into snap-frozen material by snap freezing in embedding media (optimal cutting temperature compound (OCT), Sakura Finetechnical Co., Torrance, CA, USA) in liquid nitrogen quenched in isopentane; or (3) put into HTG lysis buffer (HTG) at 4 million cells/ml for fresh samples.

Tissue used for this study came from retrospective samples of excess diagnostic tissue under an institutional review board-approved protocol. FFPE tissue used for testing the performance of the ArrayPlate included three benign lymph nodes with reactive follicular hyperplasia and two DLBCL chosen from our hospital files for having both multiple snap-frozen and paraffin blocks. FFPE blocks were from the routine surgical pathology archives and had been fixed in 10% neutral-buffered formalin and stored at room temperature. In addition, the two malignant samples consisted almost entirely of tumor. Samples were from incisional biopsies, which had been fixed in approximately 5 mm³ portions. A new paraffin block was created from a previously snap-frozen block of benign lymphoid tissue from 1989 by thawing to approximately −20°C, then immediately fixing in 10% formalin as described above and processing for Array-Plate analysis as described below.

To investigate whether the ArrayPlate results on FFPE samples were similar to GEP discovery results from snap-frozen tissue, we used FFPE samples, which had matching frozen tissue already analyzed by GEP on the LymphoChip and Affymetrix microarrays. We previously contributed 45 snap-frozen samples to the LLMPP study of DLBCL. Of these, 24 were included in the LLMPP study by Rosenwald *et al*, describing the clinical utility of an outcome predictor score in reflecting patient survival in *de novo* DLBCL, and paraffin blocks were available for 18 samples. We also had paraffin blocks on 22 cases profiled by the LLMPP but not included in the *de novo* DLBCL composed of 10 transformed and 12 relapsed DLBCL cases. These blocks ranged from 6 to 24 years of age, with an average of 14.8 years. Ten blocks were greater than 20 years.

Sample Preparation for the ArrayPlate Assay
Fresh cell pellets were put into HTG lysis buffer (4 million cells/ml), vortexed briefly, heated at 95°C for 10 min, and then frozen at −70°C until analysis. Three frozen or FFPE sections were cut at 5-μm thickness and immediately placed into 75 μl HTG lysis buffer (25 μl/section), vortexed briefly, heated at 95°C for 10 min, revortexed briefly, and then frozen at −70°C until analysis. One cut of the 5-μm-thick tissue was used per well on the ArrayPlate, but the same HTG lysis buffer containing multiple sections was used for the three ArrayPlate wells required to measure all the genes in the assay, or for multiple replicates of the same wells in the validation assays. It was important to cut the paraffin blocks into thin sections, as FFPE tissue not thinly cut gave poor results. Tissue quantities for the dilutional experiments were measured by comparing the weight of the microcentrifuge tubes with and without tissue sections from which most paraffin been removed.

ArrayPlate Assay
The ArrayPlate assay has been described previously.[7] After cells or tissue had been lysed, denatured, and permeabilized by heating in HTG lysis buffer as described above, the samples were frozen and sent to HTG for analysis. At HTG, the 50-mer probes specific for the genes of interest were incubated with the samples, forming specific probe–mRNA duplexes, and then unhybridized probes were digested by S1 nuclease. Next, alkaline hydrolysis destroyed the mRNA in the duplexes, leaving intact probes with stoichiometric concentrations proportional to the amounts of specific mRNA originally present. After neutralization, samples were transferred to ArrayPlates for probe detection. The Array-Plates contained a universal array of 16 unique, covalently bound, 25-mer 'anchor' oligonucleotides spotted in a 4 × 4 grid on the bottom of each well. This universal array was modified to bind 50-mer probes for the genes of interest at preselected positions by exposing the array to a mixture of 50-mer programming linker oligonucleotides that contained a 25-mer sequence to bind one of the probes at one end, and a 25-mer sequence to bind one of the anchor oligonucleotides on the other end. Three different mixtures of programming linker oligonucleotides distributed across three ArrayPlate wells were required to measure all the genes of interest in our assay.

After hybridization, probes from the sample were bound to array elements by the programming linker oligonucleotides. A mixture of detection linker oligonucleotides was added. The 50-mer detection linkers contained a 25-mer sequence that bound sample probe on the end not bound by the programming linker probe on one end, and a common 25-mer sequence to bind a detection probe on the other. Detection probe was added, which bound to all the detection linkers. The detection probe contained bound horseradish peroxidase. Upon the addition of chemiluminescent peroxidase substrate (Lumigen PS-atto, Lumigen Inc., Southfield, MI, USA), each array element gave off light proportional to the amount of sample probe bound at that position. Figure 1 shows a schematic of the assay.

The signals for all 1536 elements in an ArrayPlate were recorded simultaneously by imaging the plate from the bottom with a CCD-based Omix Imager (HTG). Images were analyzed using Vuescript software (HTG), which calculated average pixel intensity for each element to determine expression levels for each gene. Expression levels were normalized to the housekeeping gene *TBP* for which the signal level was arbitrarily set at 1000, except for the linearity experiments, where we used a weighted normalization of all genes to allow for titration.

For some array elements, signals were commonly too strong relative to other elements in the well. For these, signal was decreased by adding a 25-mer 'attenuation fragment' oligonucleotide that competed with one of the detection linkers for binding to sample probe, but that could not bind peroxidase-labeled detection probe. For rare mRNA transcripts, assay signal was increased by targeting the message with several non-overlapping probes and subsequently binding these non-overlapping probes at the same position
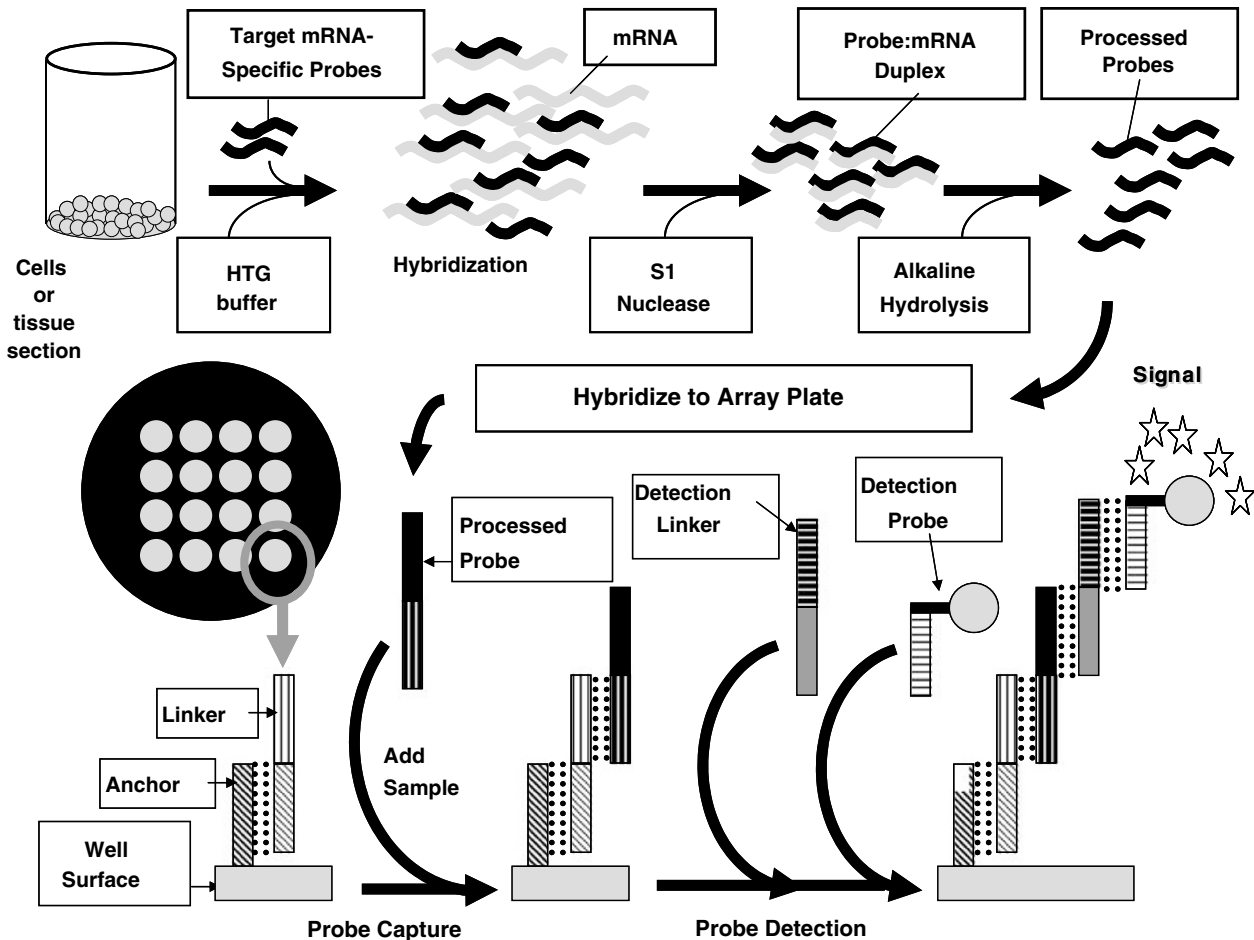
Figure 1 Diagram of ArrayPlate assay. HTG prints a Universal Array in every well of a 96-well plate with 16 different anchor oligonucleotides. A mixture of custom programming linkers is added, each designed to hybridize to a specific anchor and capture a gene-specific nuclease protection probe (shown for one linker). The array now captures probes for a custom set of genes. HTG lysis buffer lyses and/or permeabilizes sample, and then nuclease protection probes are added. The probes hybridize to all mRNA, soluble and cross-linked. S1 nuclease destroys nonspecific RNA producing a stoichiometric amount of target-mRNA/probe duplexes. Base hydrolysis releases probe from duplexes. Probes are transferred to a programmed ArrayPlate well, detection linker added, and both probes and detection linkers captured onto the array. The ArrayPlate is washed, HRP-labeled detection probe added, incubated, washed, then chemiluminescent substrate is added and the ArrayPlate imaged to measure the expression of every gene in all wells.

within the array by several distinct programming linkers. The same amounts of attenuating or amplifying oligos were always added, so expression levels were calculated for these array elements in the same way as for non-altered elements.

Selection of Genes and Probes for Programming of ArrayPlate Assay

As described at the beginning of this article, we used the key genes identified as prognostically important in four previous studies in DLBCL, which accounted for 36 genes of interest.[2–4,6] Because of the heterogeneity of cellular composition in human tumor samples, we also included probes designed to test the tumor composition for B cells (CD19, CD20), T cells (CD3), and histiocytes (CD68). Two housekeeping genes, TBP and PRKG1, were chosen based on previously published work assessing the utility of different endogenously expressed genes as housekeeping genes, which identified these two

genes as stably expressed at moderate or low levels in different types of lymphomas by qRT-PCR.[10] These two housekeeping genes were repeated at diagonal corners in each of the three wells used to create the assay. An oligo dT probe was added to assess the quantity of mRNA in the sample (since an oligo dT probe should detect all mRNA which has a poly-A tail). However, for technical reasons owing to the stringency of the assay, this probe was non-functional and not further utilized. A probe for cytochrome oxidase was also initially included because it is coded in mitochondrial DNA, and should be expressed at high levels. This turned out to bind both DNA and RNA, and so gave an extremely bright and generally oversaturated signal and was therefore not further considered, except that it could be used to distinguish whether there was insufficient material for the assay, or whether, if it had disappeared entirely, the sample was too degraded for use.

For each of the 44 genes of interest, four specific probes were designed, although not all were synthesized. Array-Builder 2.0 software (HTG) was used to design the oligo-nucleotides required for the assay to measure target transcripts in groups of 16. Briefly, with the user providing the accession numbers for the target genes and assigning their position in the array, the software retrieved each mRNA sequence from GenBank and ranked successive 50-mer stretches of the target gene sequences according to the melting temperature (T_m) of their 5′ and 3′ constituent 25-mers, giving preference to those 50-mers for which the T_m of each of the two 25-mer halves was nearest to 68°C. The four highest ranked and non-overlapping 50-mer sequences for each of the 16 target mRNA species were subjected to BLAST to identify homologous sequences. Sequences with homology to other genes were rejected and replaced with the next highest-ranking 50-mer sequence that was in turn submitted to BLAST. Sequences without significant homology were re-tained. Typically, 3–4 iterations were required to yield four suitable target 50-mer sequences within each of the 16 genes.

The software then created output files containing the sequences of the four oligonucleotides (programming linker, protection probe, detection linker, and attenuation fragment) required to measure a given 50-mer target in the assay.

For four genes, the signal had to be lowered with un-detectable competing oligonucleotide. For nine genes, the signal had to be increased, as shown in Figure 2. Between the two housekeeping genes, *PRKG1* and *TBP*, *TBP* had stronger and substantially more uniform luminescence (data not shown). *PKRG1* expression was variable and expression was low. Results were therefore normalized only to *TBP* set at 1000. Only one of the probes, for *HLA-DRB*, had to be redesigned, because the original single GenBank sequence did not identify the common polymorphisms in this highly polymorphic gene.

Table 1 lists the names of the genes of interest from the original papers, references, gene symbols for the plate, position at which probes begin for that gene, and the sequence of the designed probes. Figure 2 shows the DLBCL assay design, including which signals were reduced
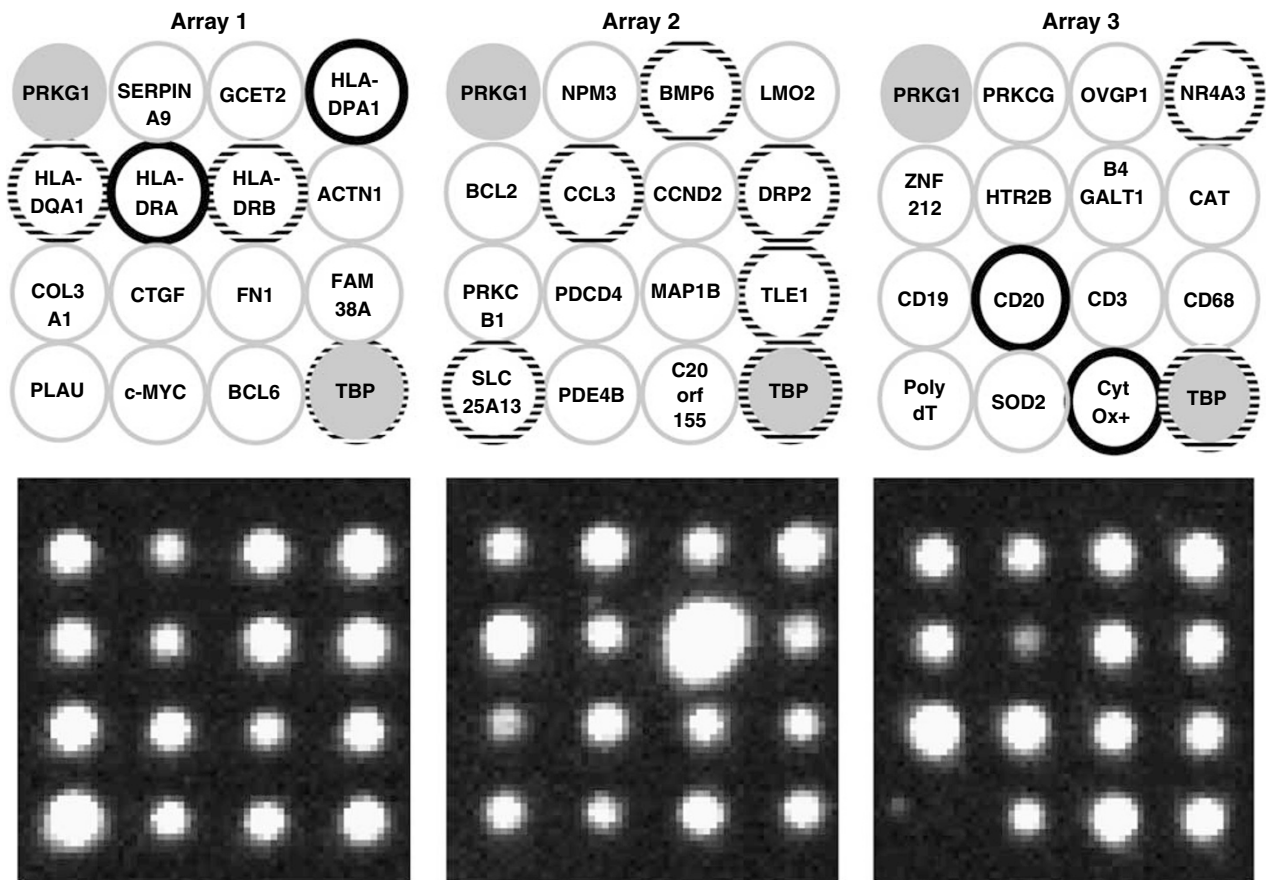
Figure 2 Setup of custom ArrayPlate for prognostic genes in DLBCL. The schematic of the arrays is above with a representative luminescence image from that array below. Genes circled in dashed lines had their signal attenuated with an undetectable competing oligonucleotide. Genes circled in black had their signal enhanced with multiple probes. Genes in gray circles were designed as housekeeping genes. Each array is one of three wells in 96-well plates with 16 locations for an anchor probe, which binds to a custom linker probe, which in turn binds to the specific probe (Figure 1). The putative housekeeping genes are repeated in each well.

Table 1 Gene list for DLBCL ArrayPlate

Accession nos.	Name in original reference	ArrayPlate name	Reference[a]	Gene position	Target sequence (5′ start)
NM_006258	PRKG1	PRKG1	HK[b]	465	CGGTGGAGTATGGCAAGGACAGTTGCATCATCAAAGAAGGAGACGTGGGG
NM_175739	IMAGE 1334260	SERPINA9	R 2	934	TGCACCAGAAAGAGCAGTTCGCTTTTGGGGTGGATACAGAGCTGAACTGC
NM_152785	IMAGE 814622	GCET2	R 3	222	GCAAAGCCCCAAACAGAGAACATCCAGATGCTGGGATCACCATATCGCTG
NM_033554	HLA-DPα	HLA-DPA1	R 4	236	AAGAAGGAGACCGTCTGGCATCTGGAGGAGTTTGGCCAAGCCTTTTCCTT
NM_002122	HLA-DQα	HLA-DQA1	R 5	1391	GCAACAATGAAATTAATGGATACCGTCTGCCCTTGGCCCAGAATTGTTAT
NM_019111	HLA-DRα	HLA-DRA	R 6	335	TGGCCAACATAGCTGTGGACAAAGCCAACCTGGAAATCATGACAAAGCGC
NM_002124	HLA-DRβ	HLA-DRB	R 7	14	TGGAAACAGTTCCTCGGAGTGGAGAGGTTTACACCTGCCAAGTGGAGCAC
NM_001102	α-actinin	ACTN1	R 8	1922	AGACCTACCACGTCAATATGGCGGGCACCAACCCCTACACAACCATCACG
NM_000090	collagen type III α 1	COL3A1	R 9	4349	CAGTTCTGGAGGATGGTTGCACGAAACACACTGGGGAATGGAGCAAAACA
NM_001901	connective-tissue growth factor	CTGF	R 10	1698	TTCAGGAATCGGAATCCTGTCGATTAGACTGGACAGCTTGTGGCAAGTGA
NM_212482	fibronectin	FN1	R 11/L 5	7340	GGGAGAAAATGGCCAGATGATGAGCTGCACATGTCTTGGGAACGGAAAAG
NM_014745	KIAA0233	FAM38A	R 12	3947	GTGCTATGGCCTCTGGGACCATGAGGAGGACTCACCATCCAAGGAGCATG
NM_002658	urokinase plasminogen activator	PLAU	R 13	835	GGGTCGCTCAAGGCTTAACTCCAACACGCAAGGGGAGATGAAGTTTGAGG
NM_002467	c-myc	c-MYC	R 14	1477	CCACACATCAGCACAACTACGCAGCGCCTCCCTCCACTCGGAAGGACTAT
NM_138931	bcl-6	BCL6	R 1/L 6	1948	GATTCTAGCTGTGAGAACGGGGCCTTCTTCTGCAATGAGTGTGACTGCCG
M34960	TATA Box binding protein	TBP	HK	562	CGAAACGCCGAATATAATCCCAAGCGGTTTGCTGCGGTAATCATGAGGAT
M34960	TATA Box binding protein	TBP	HK	461	CAGCTTCGGAGAGTTCTGGGATTGTACCGCAGCTGCAAAATATTGTATCC
M34960	TATA Box binding protein	TBP	HK	774	GGTGGGGAGCTGTGATGTGAAGTTTCCTATAAGGTTAGAAGGCCTTGTGC
NM_006258	PRKG1	PRKG1	HK	465	CGGTGGAGTATGGCAAGGACAGTTGCATCATCAAAGAAGGAGACGTGGGG
NM_006993	NPM3	NPM3	R 16	418	GGCACCAGATTGTTACGATGAGCAATGATGTTTCTGAGGAGGAGAGCGAG
NM_001718	BMP6	BMP6	R 17	1566	ACCTTGGTTCACCTTATGAACCCCGAGTATGTCCCCAAACCGTGCTGTGC
NM_001718	BMP6	BMP6	R 17	1807	GGTGGGACGATGAGACTTTGAAACTATCTCATGCCAGTGCCTTATTACCC
NM_001718	BMP6	BMP6	R 17	1031	GCACAGAGACTCTGACCTGTTTTTGTTGGACACCCGTGTAGTATGGGCCT
NM_001718	BMP6	BMP6	R 17	2458	GCTCACCTCTTCTTTACCAGAACGGTTCTTTGACCAGCACATTAACTTCT
NM_005574	LM02	LMO2	L 1	2012	AAGGCCTTAAGCTTTGGACCCAAGGGAAAACTGCATGGAGACGCATTTCG
NM_000633	BCL2	BCL2	L 2	2165	CCTGCTTTTAGGAGACCGAAGTCCGCAGAACCTGCCTGTGTCCCAGCTTG
NM_002983.1	SCYA3	CCL3	L 3	715	ATGCTTTTGTTCAGGGCTGTGATCGGCCTGGGGAAATAATAAAGCACGCT
NM_002983.1	SCYA3	CCL3	L 3	30	CCTTTCTTGGCTCTGCTGACACTCGAGCCCACATTCCGTCACCTGCTCAG
NM_002983.1	SCYA3	CCL3	L 3	127	TGGCTCTCTGCAACCAGTTCTCTGCATCACTTGCTGCTGACACGCCGACC
NM_002983.1	SCYA3	CCL3	L 3	571	GTGTGTTTGTGATTGTTTGCTCTGAGAGTTCCCCTGTCCCCTCCCCCTTC
NM_001759.2	CCND2	CCND2	L 4	3666	GCGAGTAGATGAACCTGCAGCAAGCAGCGTTTATGGTGCTTCCTTCTCCC
NM_001939	DRP2 dystrophin related protein 2	DRP2	S 1	871	AGCAAAGATACCTCCCCGAAACAGCGGATCCAGAATCTCAGCCGCTTTGT
NM_001939	DRP2 dystrophin related protein 2	DRP2	S 1	3282	CACTGGCCCCACATTCCTCAACTAGTATTATTTGGGCTCTGGGCAGCAGC
NM_001939	DRP2 dystrophin related protein 2	DRP2	S 1	1030	GGGGCAATGGAGGAACTAAGCACTACTCTAAGCCAAGCTGAGGGAGTCCG
NM_001939	DRP2 dystrophin related protein 2	DRP2	S 1	3038	GACAGACCACTCCAGATACCGAGGCTGCAGATGATGTGGGGTCAAAGAGC
NM_002738	PRKACB protein kinase C-beta-1	PRKCB1	S 2	2787	AAAAGCACTTCAAGGGGTCAAAGGGCAACCAGCTTGGGTGCTACCTCAGT
NM_014456	H731 nuclear antigen	PDCD4	S 3	518	CAACCAGTCCAAAGGGAAGGTTGCTGGATAGGCGATCCAGATCTGGGAAA
NM_005909	3′ UTR of unknown protein	MAP1B	S 4	7037	CAAAACCAGCGGGCTTGAAAGAATCCTCGGATAAAGTGTCCAGGGTGGCT
NM_005077	Transducin-like enhancer protein 1	TLE1	S 5	3039	TTCTTTCTGGGTGATCTGGGGATCACGCCTTGCCCAAGTGTGAGATTACC

Table 1 Continued

Accession nos.	Name in original reference	ArrayPlate name	Reference[a]	Gene position	Target sequence (5′ start)
NM_005077	Transducin-like enhancer protein 1	TLE1	S 5	1703	TTGATCCTCCCCCTCACATGAGAGTACCTACCATTCCTCCAAACCTGGCA
NM_005077	Transducin-like enhancer protein 1	TLE1	S 5	1312	GCCTCCTCGGCAAGTTCCACTTCTTTGAAATCCAAAGAAATGAGCTTGCA
NM_005077	Transducin-like enhancer protein 1	TLE1	S 5	1255	GGAATCGACAAAAATCGCCTGCTAAAGAAGGATGCTTCTAGCAGTCCAGC
NM_014251	Uncharacterized	SLC25A13	S 6	1662	GCTTCCTTTGCAAATGAAGATGGGCAGGTTAGCCCAGGAAGCCTGCTCTT
NM_014251	Uncharacterized	SLC25A13	S 6	2037	CCTGATCACGTTGGGGGGCTACAAACTGGCAGTTGCTACATTTGCAGGGAT
NM_014251	Uncharacterized	SLC25A13	S 6	890	GGAGGAGTTTGTTCTGGCAGCTCAGAAATTTGGTCAGGTTACACCCATGG
NM_014251	Uncharacterized	SLC25A13	S 6	1536	CGAGTCAGTGCTCTGTCTGTCGTGCGGGACCTGGGGTTTTTTGGGATCTA
NM_002600	PDE4B Phosphodiesterase 4B, cAMP-specific	PDE4B	S 7	2128	CACCACCACTGGACGAGCAGAACAGGGACTGCCAGGGTCTGATGGAGAAG
NM_019095	E21G3 (Nucleostemin)	C20orf155	R 15	474	GCTCGAAACTGGGCCAATCAAAGATCAGCTTTGGGAAGTGCTCTTGATCC
M34960	TATA Box binding protein	TBP	HK	537	CCTAAAGACCATTGCACTTCGTCGCCGAAACGCCGAATATAATCCCAAGC
M34960	TATA Box binding protein	TBP	HK	461	CAGCTTCGGAGAGTTCTGGGATTGTACCGCAGCTGCAAAATATTGTATCC
M34960	TATA Box binding protein	TBP	HK	774	GGTGGGGAGCTGTGATGTGAAGTTTCCTATAAGGTTAGAAGGCCTTGTGC
NM_006258	PRKG1	PRKG1	HK	465	CGGTGGAGTATGGCAAGGACAGTTGCATCATCAAAGAAGGAGACGTGGGG
NM_002739	PRKCG Protein kinase C, gamma	PRKCG	S 9	901	CTGACGAAACAGAAGACCCGAACGGTGAAAGCCACGCTAAACCCTGTGTG
NM_002557	Oviductal glycoprotein	OVGP1	S 10	846	GGACGTACCTTTCGCCTCCTCAAAGCCTCTAAGAATGGGTTGCAGGCCAG
NM_173198	(MINOR) Mitogen induced nuclear orphan receptor	NR4A3	S 11	1055	CCAATGGCCTCTTTCCTCCCAAATAAACCACTGGCTTTCTCTTTGTCCCC
NM_173198	(MINOR) Mitogen induced nuclear orphan receptor	NR4A3	S 11	2957	TGTTCTGCAATGGACTTGTCCTGCATCGACTTCAGTGCCTTCGTGGATTT
NM_173198	(MINOR) Mitogen induced nuclear orphan receptor	NR4A3	S 11	2647	CCACCTTCTCCTCCAATCTGCATGATGAATGCCCTTGTCCGAGCTTTAAC
NM_173198	(MINOR) Mitogen induced nuclear orphan receptor	NR4A3	S 11	4095	CCCTGTCGATCCCTTCTGAGGTATGGCCCATCCAAGACTTTTAGGCCATT
NM_012256	Zinc-finger protein C2H2-150	ZNF212	S 12	518	GGTCACTGGAGAATGATGGCGTCTGTTTCACCGAGCAGGAATGGGAGAAT
NM_000867	5-Hydroxytryptamine 2B receptor	HTR2B	S 13	1809	CGAAATGGGATTAACCCTGCCATGTACCAGAGTCCAATGAGGCTCCGAAG
NM_001497	Uncharacterized	B4GALT1	S 8	1868	TCCAGGGCAACTCTAGCATCAGAGCAAAAGCCTTGGGTTTCTCGCATTCA
NM_001752	Catalase	CAT	T 1	1148	TTTTGCCTATCCTGACACTCACCGCCATCGCCTGGGACCCAATTATCTTC
NM_001770	CD19	CD19	B cell	128	GGAAGAGGGAGATAACGCTGTGCTGCAGTGCCTCAAGGGGACCTCAGATG
NM_152866	CD20	MS4A1	B cell	64	AACAAACTGCACCCACTGAACTCCGCAGCTAGCATCCAAATCAGCCCTTG
NM_000732	CD3-delta	CD3δ	T cell	410	GCCGACACACAAGCTCTGTTGAGGAATGACCAGGTCTATCAGCCCCTCCG
NM_001251	CD68	CD68	Histiocytes	667	TTCCCCTATGGACACCTCAGCTTTGGATTCATGCAGGACCTCCAGCAGAA
Unusable	poly dT	poly dT	MRNA	Poly-A tail	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
NM_000636	Manganese superoxide dismutase	SOD2	T 2	659	CCACTGCAAGGAACAACAGGCCTTATTCCACTGCTGGGGATTGATGTGTG
AY963585	Cytochrome oxidase	CytOx+	Loading control	524	CCCTGCCATAACCCAATACCAAACGCCCCTCTTCGTCTGATCCGTCCTAA
M34960	TATA Box binding protein	TBP	HK	537	CCTAAAGACCATTGCACTTCGTCGCCGAAACGCCGAATATAATCCCAAGC
M34960	TATA Box binding protein	TBP	HK	461	CAGCTTCGGAGAGTTCTGGGATTGTACCGCAGCTGCAAAATATTGTATCC
M34960	TATA Box binding protein	TBP	HK	774	GGTGGGGAGCTGTGATGTGAAGTTTCCTATAAGGTTAGAAGGCCTTGTGC

[a]Single letters with numbers refer to genes that are members of the four prognostic gene signatures and their position in the signature are as follows: R represents the paper by Rosenwald et al[2] with genes 1–17, L represents the paper by Lossos et al[4] with genes 1–6, S represents the paper by Shipp et al[3] with genes 1–13, T represents the paper by Tome et al[6] with 2 genes.

[b]HK are putative hematopoietic housekeeping genes recommended in another paper by Lossos et al.[10]

(attenuated) or increased (amplified) (described above). Test samples were assayed in replicates, as indicated in figure legends, to provide measures of reproducibility, but because of the low variability found between replicates, one replicate per well of the assay was considered to be sufficient for our analysis of tumor samples.

Array Validation Experiments

Unlike fresh or frozen samples, when FFPE sections were treated with HTG lysis buffer, they tended not to dissolve completely. Even though there was remaining visible FFPE tissue in many cases, the assay performed similarly to when a similar weight of fresh or frozen sample that dissolved completely was used. To better understand this phenomenon, we treated a benign FFPE sample and a comparable frozen sample in HTG lysis buffer. The FFPE sample was centrifuged, and the pellet and supernatant were separated. The pellet was resuspended in an equal volume of HTG buffer, and all three samples assayed on the ArrayPlate.

Because the gene probes are not designed across exon boundaries, we wanted to be sure that the ArrayPlate measured concentrations of mRNA and not DNA. We therefore took benign tissue FFPE samples treated with HTG lysis buffer, and then the sample was either treated with RNase, DNase, or was left untreated and assayed on the ArrayPlate.

Assays were also performed with replicates and with dilutions to demonstrate the linearity and reproducibility of the assay with different sample types, including fresh, frozen, FFPE archived and recently made samples. Limited quantities of patient tissue forced us to limit some assays to only a single of the three arrays.

Immunohistochemistry

Immunohistochemistry (IHC) was performed on three samples to detect protein products for genes present on the array for which we had clinical IHC assays routinely performed in paraffin in our hospital. These included CD20, CD3, CD68 (data not shown for these first three), BCL2, BCL6, and HLA-DR. All staining was performed on the Ventana Benchmark XT instrument with Ventana I-View detection (Ventana Medical Systems Incorporated (VMSI), Tucson, AZ, USA). Our standard clinical laboratory staining procedures with on-instrument antigen retrieval were employed. Monoclonal antibodies were used against CD20 (VMSI, clone L26), CD3 (VMSI, clone PS1), CD68 (VMSI, clone KP1), BCL2 (VMSI, clone B4-2/100/D5), BCL6 (CNIO, Madrid, Spain, clone GI191E/A8), and HLA-DR (Biogenix, Bogota, Columbia, clone LN3). Photography was performed on a Labophot-2 microscope using a ×10 eyepiece and a ×40/0.65 objective lens (Nikon, Melville, NY, USA). A SPOT-RT 2.2.0 color camera and SPOT Advanced 4.0.9 software (Diagnostic Instruments, Sterling Heights, MI, USA) were used to capture and digitally acquire images, which were then inserted into PowerPoint 10 (Microsoft, Redmond, WA, USA) and cropped for publication.
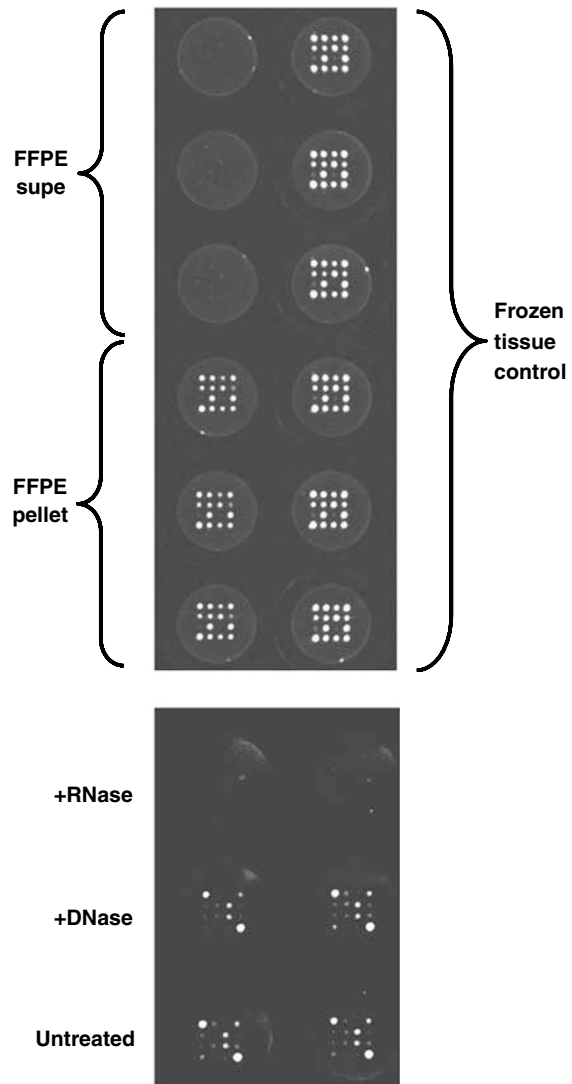

Figure 3 Array detection of RNA. At top is a demonstration that the nuclease protection assay detects both soluble and fixed RNA from FFPE samples. A highly crosslinked FFPE section and a comparable size frozen section from the same sample were treated with HTG lysis buffer. The FFPE sample was centrifuged, and the pellet and supernatant separated, then pellet resuspended in an equal volume of HTG lysis buffer, and all three samples assayed in an array of a somewhat different customization from the our ArrayPlate. Even in this case where almost all mRNA is fixed in the sample pellet, the assay measures similar amounts to the unfixed sample. At bottom, adding nucleases to samples demonstrates that the nuclease protection assay detects mRNA but not DNA. Sample lysates were either treated with DNase or RNase, or left untreated, before hybridization to an array of a somewhat different customization from the our ArrayPlate. DNase does not change the signal substantially, but RNase abrogates the signal.

Statistical Methods

Simple statistical calculations such as averages, s.d., linear correlations were performed in Excel 10 (Microsoft). Statistical analyses were performed on the 18 *de novo* untreated and 21 non-*de novo* relapsed or transformed DLBCL cases with results from both ArrayPlate and Affymetrix/

LymphoChip analysis of gene expression. Analyses of hazard ratios and survival were performed on all samples, because of the small sample size, unlike in the study by Rosenwald *et al*,[2] which used only *de novo* untreated samples. Spearman rank-order correlations for each pairwise comparison of three studied methods were derived for each gene. For a given comparison, analyses with fewer than 10 available observations were not performed due to power considerations. The median overall correlation between available genes for each pairwise comparison was computed; only genes with available data for all three pairwise comparisons were included to avoid bias. Univariate analysis results (hazard ratios, 95% confidence intervals, and *P*-values) of gene expression as measured by ArrayPlate on overall survival were obtained from the Cox regression model.[11] For each method, the genes were defined as either positively or negatively associated with outcome based on univariate Cox model estimates. Two-way tables, representing the level of agreement, were calculated between the LymphoChip *vs* the ArrayPlate, Affymetrix *vs* the

ArrayPlate, and the LymphoChip *vs* Affymetrix. The kappa statistic,[12] which adjusts for chance agreement, was used to statistically summarize the level of agreement between each method. Given that genes were not independent, confidence intervals for the kappa statistics were calculated using bootstrap resampling.[13]

RESULTS
ArrayPlate Measures All mRNA and Only mRNA in Samples

To characterize the performance of the assay when FFPE tissues were tested, sections were pretreated with RNase or DNase. As can be seen at the bottom of Figure 3, the assay chemiluminescence was abrogated after RNase but not DNase treatment. Therefore, RNA not DNA was measured in the assay. RNA is difficult to measure in fixed tissue by methods that require extraction, such as qRT-PCR, but relatively easy to measure *in situ* by methods such as *in situ* hybridization. RNA is believed to be crosslinked to other tissue molecules by
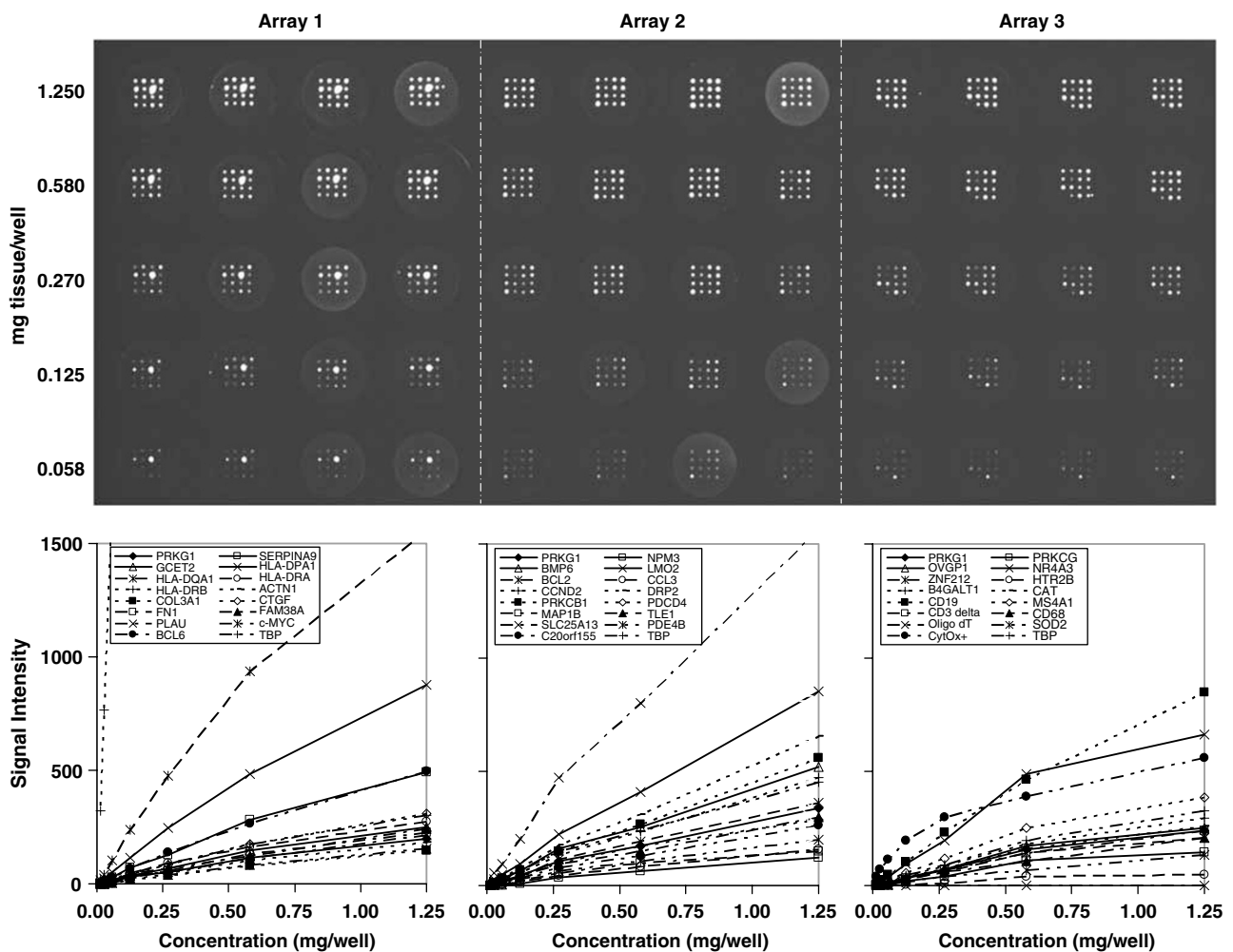

Figure 4 ArrayPlate dilution study. At top is an image of replicates from serial dilutions of case 0, frozen benign hyperplastic lymph node tissue diluted from 1.25 to 0.058 mg/well. Graphs at the bottom show the assay linearity for each array after weighted normalization. Table 2 demonstrates the reproducibility of this experiment.

fixative, preserving the structure but making it difficult to remove intact. With extended archiving, the crosslinking may increase. In FFPE samples, the tissue is denatured and permeabilized, but not totally dissolved, as fresh or frozen

tissue is, by HTG lysis buffer. Nevertheless, the mRNA profiles measured are similar across samples, regardless of preservation method. In FFPE sections, some mRNA was not released from the tissue, but the probes bound to the mRNA

Table 2 Reproducibility of gene expression in frozen tissue in a dilution study

Array 1 (mg/well) elements	1.25 average	$n=4$ %CV	0.58 average	$n=4$ %CV	0.27 average	$n=4$ %CV	0.125 average	$n=4$ %CV
PRKG1	664	13	642	6	577	12	629	10
SERPINA9	1642	14	1605	5	1470	10	1417	4
GCET2	842	17	832	11	801	8	722	6
HLA-DPA1	2943	16	2766	10	2852	6	2385	11
HLA-DQA1	5220	14	5351	13	5522	10	4833	7
HLA-DRA	918	17	938	15	1077	7	933	12
HLA-DRB	46134	13	79053	10	94032	10	82515	9
ACTN1	608	8	595	7	550	6	834	55
COL3A1	483	10	440	5	420	16	345	6
CTGF	1023	10	986	8	954	17	869	8
FN1	752	5	745	6	639	13	632	6
FAM38A	809	6	712	13	797	8	879	19
PLAU	709	4	631	4	583	9	581	4
c-MYC	511	10	482	5	470	9	451	8
BCL6	1657	11	1504	8	1616	13	1378	9
TBP	1000		1000		1000		1000	
Average %CV		11		8		10		12

Array 2 (mg/well) elements	1.25 average	$n=4$ %CV	0.58 average	$n=4$ %CV	0.27 average	$n=4$ %CV	0.125 average	$n=4$ %CV
PRKG1	753	13	720	7	715	5	662	15
NPM3	274	34	265	36	231	36	143	77
BMP6	1157	35	1067	30	1093	27	983	34
LMO2	1896	21	1701	20	1530	20	1492	21
BCL2	807	12	803	10	787	10	718	7
CCL3	357	23	336	18	305	26	227	76
CCND2	1052	12	983	12	916	12	903	6
DRP2	1458	12	1296	5	1203	5	1186	4
PRKCB1	1236	13	1106	10	1021	18	1055	7
PDCD4	641	16	720	21	669	18	676	7
MAP1B	336	26	408	15	383	20	311	70
TLE1	672	23	621	25	634	18	603	30
SLC25A13	3469	3	3311	4	3199	6	3237	8
PDE4B	451	12	443	11	438	7	426	7
C20orf155	584	14	545	2	541	8	530	13
TBP	1000		1000		1000		1000	
Average %CV		18		15		16		26

Table 2 Continued

Array 3 (mg/well) elements	1.25 average	$n=4$ %CV	0.58 average	$n=4$ %CV	0.27 average	$n=4$ %CV	0.125 average	$n=2$[a]
PRKG1	736	15	803	10	732	14	738	
PRKCG	443	15	550	19	404	40	422	
OVGP1	782	11	880	11	784	9	842	
NR4A3	2033	8	2491	8	2151	17	2220	
ZNF212	646	12	725	10	645	16	662	
HTR2B	145	11	198	19	96	116	56	
B4GALT1	909	12	938	12	933	11	908	
CAT	803	13	833	8	806	4	813	
CD19	2629	19	2405	19	2583	23	2422	
MS4A1	1208	24	1306	25	1316	24	1345	
CD3 delta	750	23	732	27	759	32	723	
CD68	652	25	560	29	656	27	642	
Oligo dT								
SOD2	410	20	355	26	444	30	383	
CytOx+	1715	10	2025	19	3315	25	4638	
TBP	1000		1000		1000		1000	
Average %CV		16		17		28		
Average total %CV/assay		15		13		18		19

[a]Because there were only two replicates for this array, %CV was not calculated.

in situ and were protected from nuclease digestion there. The top of Figure 3 shows that the assay signal came mainly from the pelleted material, not the solubilized supernatant. The specimen shown in Figure 3 was an extreme case of fixation. We saw variability across other cases in the relative amounts of mRNA solubilized in HTG lysis buffer to mRNA remaining bound to the tissue section. This finding underscored the importance of measuring both soluble and fixed mRNA to compensate for variations in fixation from one case to another. For this reason, FFPE tissue must be analyzed in thin sections and extra paraffin removed from around the edges of the section, so that the probes are freely diffusible.

ArrayPlate Performs Similarly on Differently Preserved Samples

Figure 4 shows that assay performance on diluted frozen tissue sample replicates was reasonably linear from the detection limits up to about 1 mg of sample/well. Table 2 contains the same data, normalized to the TBP housekeeping gene, demonstrating that concentrations of starting material differing by up to 10-fold gave comparable results, with average %CVs (coefficients of variation) ranging from 15 to 19%.

We then studied the RJ2.2.5 cell line, to obtain sufficient quantities of homogeneous material, unlike biopsy material, which may differ from section to section, and is of limited

quantity. Replicate cell pellets, either fresh, frozen in OCT, or fixed and paraffinized, were run in the ArrayPlate assay. Figure 5 shows the linear correlations of assay signals between the differently treated samples, with $R^2 > 0.98$ for all three. Figure 6 plots the signal intensity for the three treatments ± s.d., showing that RNA expression as measured by the assay was substantially the same for fresh, frozen, or FFPE samples.

Because we wanted to do similar experiments with FFPE tissue samples, we first looked at the differences in gene expression in clinical FFPE tissue samples between sections. Since our sections are 5 μm thick, the three sets of patient sections from the same block are separated by about 10 sections. With the blocks being ∼ 5000 μm on a side, the sections from separate blocks can be quite distant. Figure 7 indicates that, while sections have some variability, the relatively small differences show that different sections or even blocks from the same patient have very similar results by ArrayPlate, while preserving the variability between patients.

Then, on a benign hyperplastic human lymph node biopsy for which we had sufficient material, we compared preservation methods. The new paraffin block was made from a portion of an 18-year-old snap-frozen biopsy, which was quickly defrosted, fixed in formalin, and embedded. The results were compared to a sister paraffin block made from the same biopsy at the time of the patient's surgery 18 years
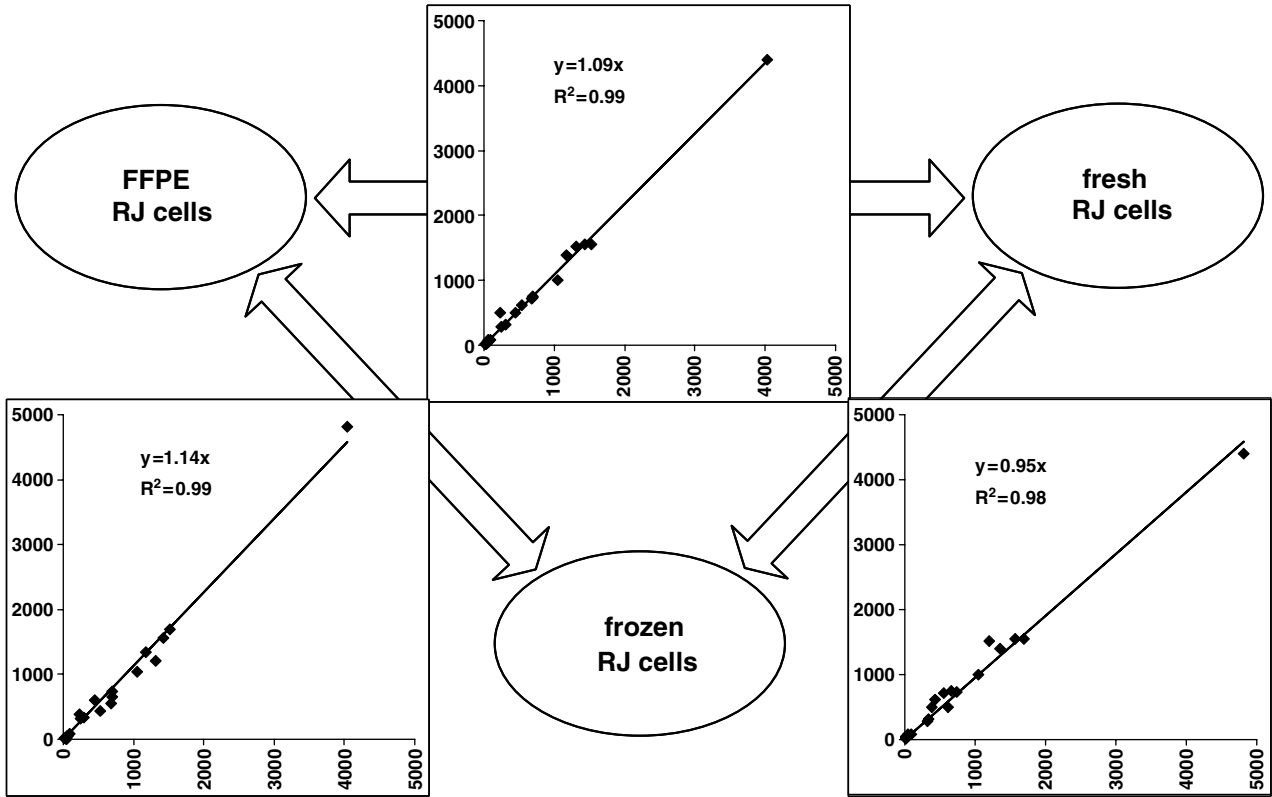
Figure 5 Comparison replicate cell line pellets after fresh preservation in HTG lysis buffer, snap freezing in OCT, or FFPE. Top graph has FFPE on the *x*-axis *vs* fresh on the *y*-axis, left graph FFPE *vs* frozen, and right graph fresh *vs* frozen. Only genes on the ArrayPlate expressed above background in the RJ2.2.5 cell line are shown. Results are nearly identical regardless of whether sample was fresh, frozen, or FFPE.
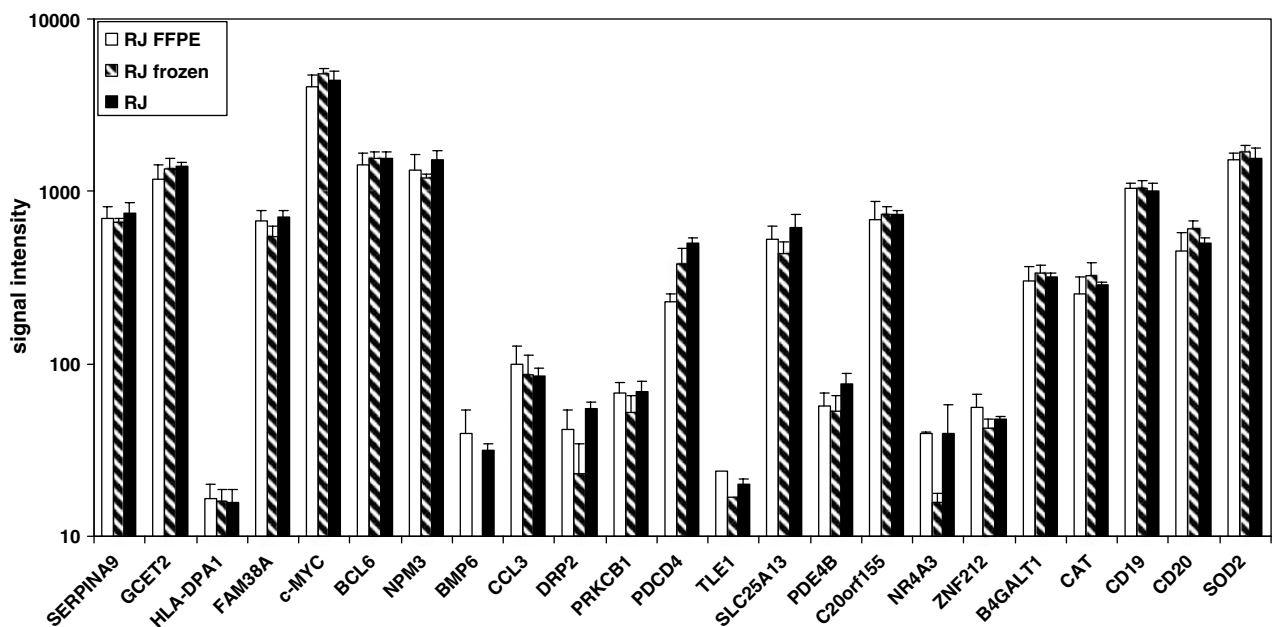

Figure 6 ArrayPlate gene expression measurements by method of preservation in cell lines. Another view of data from Figure 5 showing that gene expression in the RJ2.2.5 cell line does not differ with fresh, frozen, or FFPE cell pellets. Graph shows the average $\pm$ s.d. of four replicate wells. Only genes on the ArrayPlate expressed above background in the RJ2.2.5 cell line are shown. Differences between preservation methods are generally within assay variability for samples where sections contain essentially identical material.
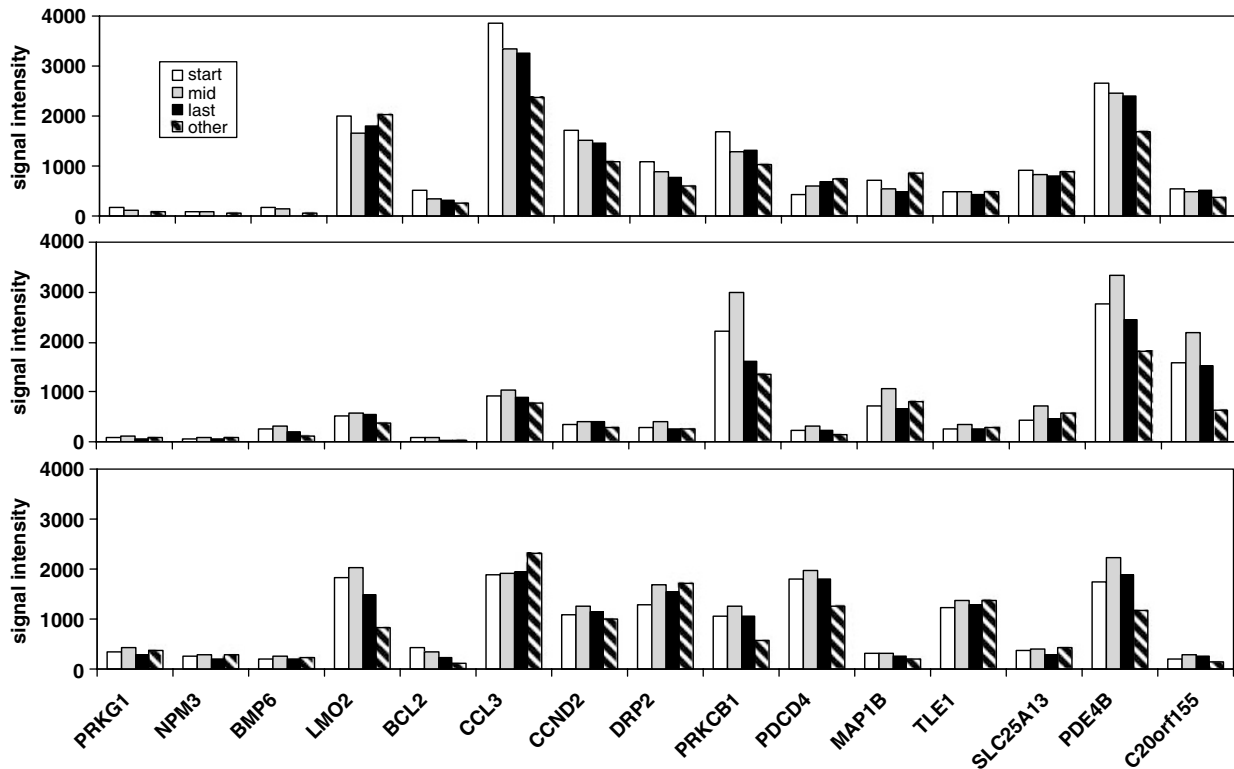
Figure 7 Similarities between non-adjacent FFPE sections from patient biopsies. One tissue section/well was run on array 1. Start section for each sample was the section that first contained significant tissue as the block was cut, mid section was 50–65 μm from first section, and last section was 115–150 μm from the first. Other section was from a different FFPE block of the same patient biopsy. Case 1 at top was an anaplastic DLBCL, case 2 in the middle an immunoblastic DLBCL, and case 4 at bottom a benign lymph node hyperplasia. While sections have some variability, the relatively small differences show that different sections or even blocks from the same patient have very similar GEP by ArrayPlate, while preserving the GEP variability between patients.

ago. As shown in Figure 8, the correlation between the new *vs* the old paraffin block from the same biopsy was good ($R^2 = 0.71$), with no more variability than what is seen in Figure 7, resulting from different sections of the same biopsy. This result indicated the applicability of the assay to very old archival material. ArrayPlate reproducibility with time of assay by method of preservation in tissue sections was also addressed. Frozen, recently FFPE, and archived FFPE tissue from case 0 above was assayed on array 2 on different days 1 week apart. The histogram in Figure 9 shows differences between samples, especially between the recent and archived FFPE samples, is largely within assay variability.

Comparison of ArrayPlate to Other Expression Measures
The ArrayPlate results accurately reflected that mRNA of *HLA-DRA* and *HLA-DQA* increased in the RJ2.2.5 cell line transfectant clone compared to the parent line as demonstrated with qRT-PCR (data not shown). ArrayPlate technology has been previously shown to have a good correlation with qRT-PCR in cell lines.[14] Because not all of the genes of interest are expected to be expressed in our B cell line or indeed in any single cell line, it would be difficult to make a uniform, large sample, which would express all of our genes of interest at concentrations relevant to our assay. The

samples expressing all the genes are clinical samples containing limited amounts of tissue. Therefore, qRT-PCR verification was not performed for all of our genes of interest. We have performed IHC for verification, and will in the long run validate the technology against the clinical 'gold standard' of patient outcome.

IHC on FFPE tissue sections was performed for CD20, CD3, CD68, BCL2, BCL6, and HLA-DR. Protein expression for these tissue sections correlated well with the gene expression results from FFPE sections on the ArrayPlate for all the genes tested. Results for BCL2, BCL6, and HLA-DR are shown in Figure 10.

Correlation between the ArrayPlate and Discovery GEP Techniques
We ran 40 cases of archived FFPE DLBCL on the ArrayPlate analysis. Thirty-nine cases were successfully analyzed. One case of transformed DLBCL had no signal at all. *In situ* hybridization with a poly-T probe demonstrated that the block from this case had no intact mRNA. Results on the remaining 39 cases were compared to previous Affymetrix and LymphoChip GEP results from matched frozen samples using Spearman rank statistics, as shown in Table 3. Since not all the gene results were available on all platforms, there were
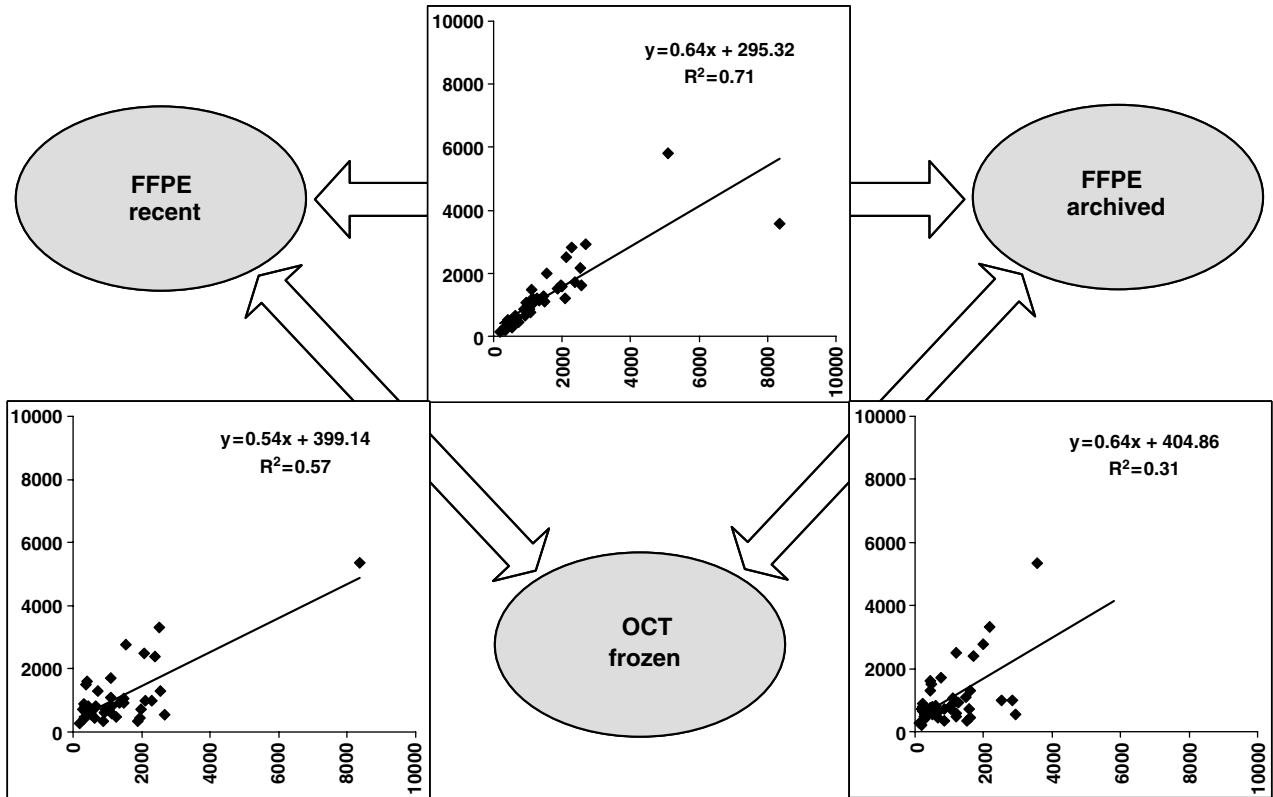
Figure 8 Comparison of tissue samples after snap freezing in OCT, archived or recent FFPE. Case 0, a large benign hyperplastic lymph node tissue biopsy from 18 years ago was preserved with both multiple frozen and archived FFPE blocks. A frozen block was recently FFPE. Top graph has recent FFPE on the *x*-axis *vs* archived FFPE on the *y*-axis, left graph recent FFPE *vs* frozen, and right graph archived FFPE *vs* frozen. The HLA-DRB gene was not included in the comparisons with frozen samples because the probe design for this gene changed between the frozen and FFPE experiments. Results are similar regardless of whether sample was FFPE either recently or 18 years ago, and reasonably similar if sample was frozen.
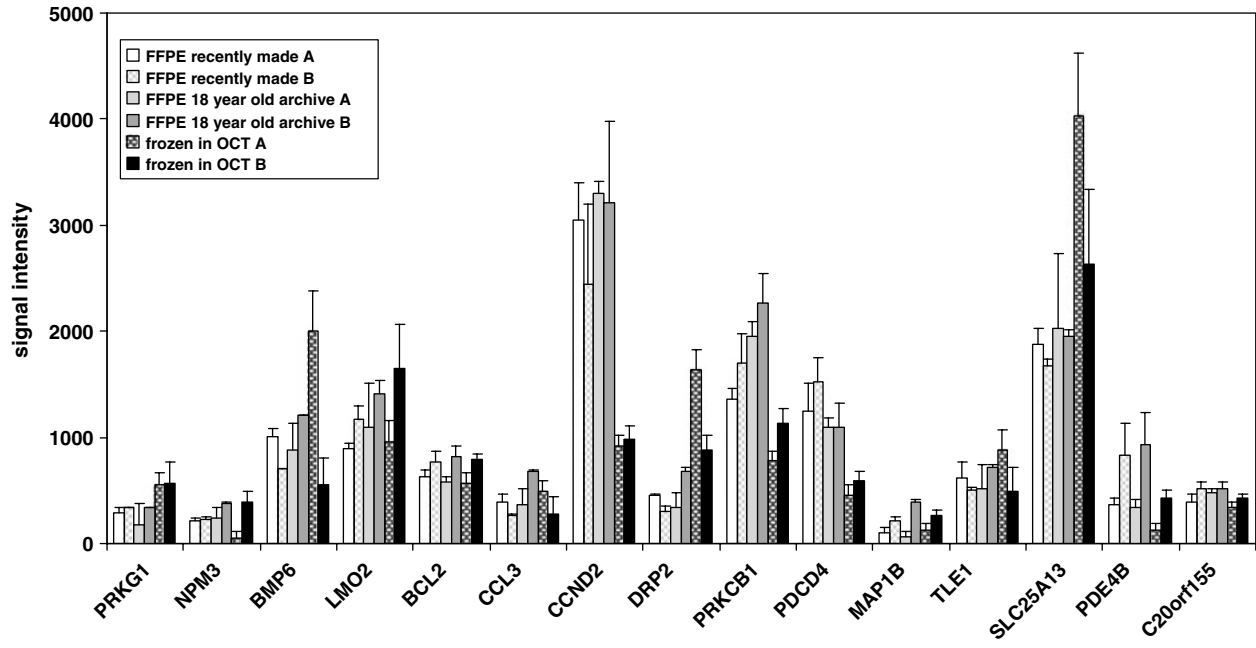

Figure 9 ArrayPlate reproducibility with time by method of preservation in tissue sections. Frozen, recently FFPE, and archived FFPE tissue from case 0 was assayed on array 2 on different days. Graph shows the average ±s.d. of 2–4 replicate wells. Differences between samples are largely within assay variability.
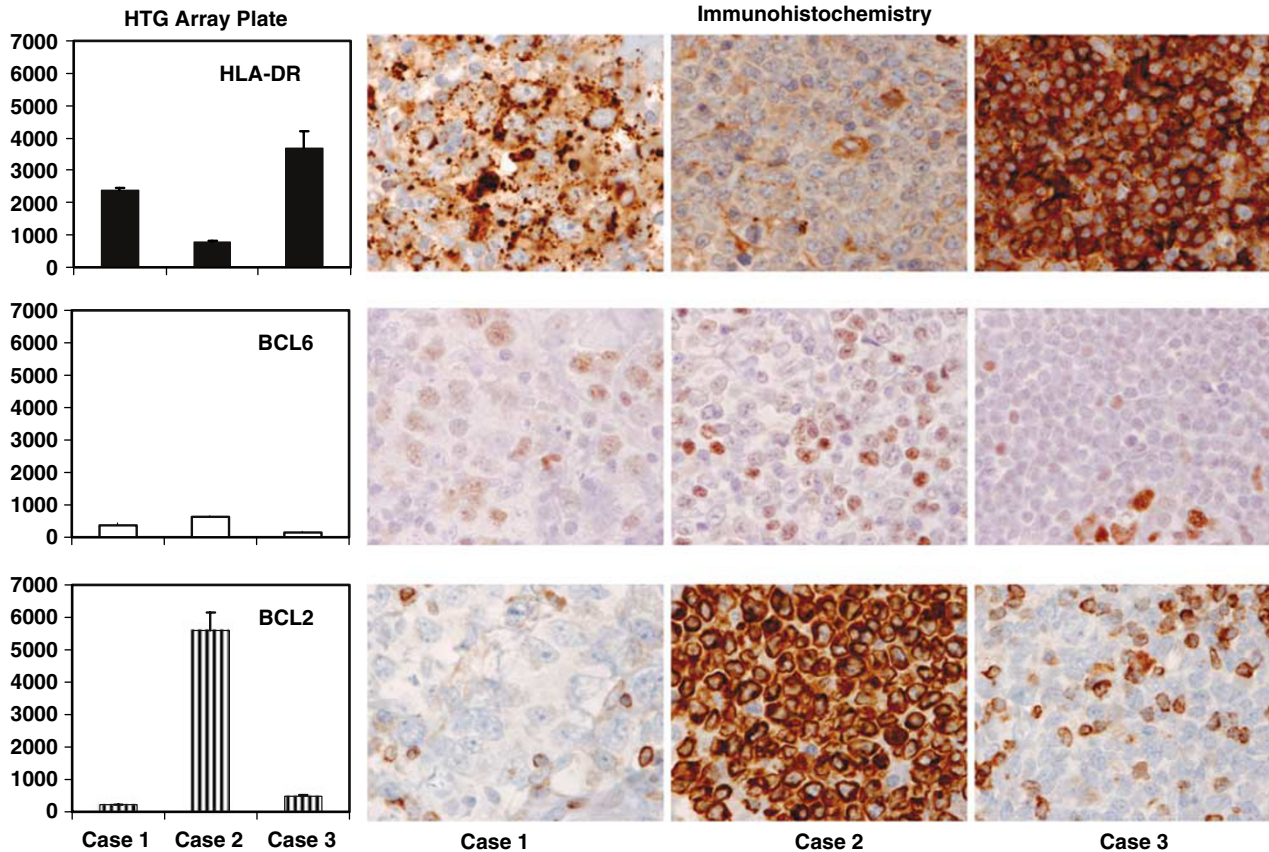
Figure 10 Comparison of results of ArrayPlate and IHC. ArrayPlate gene expression on the left and IHC on the right are shown for three FFPE patient tissue samples. Case 1 was an anaplastic DLBCL, case 2 an immunoblastic DLBCL, and case 3 a benign lymph node hyperplasia. ArrayPlate results correlate well with IHC.

frequent missing data as indicated. Of 33 genes comparable between the two microarray platforms, 27 (82%) were significantly correlated. Of 32 genes comparable between the ArrayPlate and the LymphoChip, 20 (61%) were significantly correlated, and for the ArrayPlate to Affymetrix correlation there were 37 comparable genes with 19 (51%) significantly correlated. Of the 32 genes comparable between all three technologies, 19 (59%) were significantly correlated between all three. The median correlation for ArrayPlate *vs* Affymetrix was 0.52, for ArrayPlate *vs* LymphoChip 0.55, and Affymetrix *vs* LymphoChip 0.78, as shown in Figure 11. The correlation between microarrays is understandably higher since those two analyses were performed on the same aliquot of RNA derived from a single frozen tissue block, while the ArrayPlate analysis was performed on a different FFPE block from the same specimen. Overall, these are excellent correlations for this type of technology, since correlations between Affymetrix and spotted arrays have been reported to be around 0.4,[15,16] and indicate that relationships between genes and survival that have been detected in discovery experiments could also be seen with the ArrayPlate. The genes were clustered into two groups as either being positively or negatively associated with survival outcome based on Cox regression. The

crude agreements between the methods were 79% for the LymphoChip *vs* the ArrayPlate, 75% for Affymetrix *vs* the ArrayPlate, and 89% for the LymphoChip *vs* Affymetrix. The kappa statistics were 0.76, 0.67, and 0.75, respectively. Landis and Koch[17] have characterized kappa statistics between 0.40 and 0.60 to represent 'moderate' agreement between measurements.

Hazard Ratios for Prognostic Genes with ArrayPlate Compared to Discovery GEP Techniques
The results for the 36 previously discovered prognostic genes were then compared to survival for the 39 cases successfully analyzed using the ArrayPlate. We first performed univariate analyses of gene expression levels *vs* patient survival; however, none of the genes was significantly correlated with survival, which we attributed to the low number of cases, specifically the few *de novo* previously untreated cases, in this study group (since all of these genes were associated with survival in larger groups of patients). The survival predictors for all the different groups were calculated from only *de novo* untreated samples, whereas some of our samples were previously treated relapsed or transformed DLBCL. We then calculated the hazard ratios of death for each gene and

Table 3 Spearman rank correlations by gene

Gene	Correlation (*P*-value, *n*)		
	ArrayPlate with Affymetrix	ArrayPlate with LymphoChip	Affymetrix with LymphoChip
PRKG1	NA, <10	NA, <10	−0.22 (0.38, 18)
SERPINA9	Missing	NA, <10	Missing
GCET2	0.57 (0.03, 14)	0.57 (0.03, 14)	0.79 (<0.0001, 18)
HLA-DPA1	0.54 (0.02, 18)	0.85 (<0.0001, 18)	0.63 (0.005, 18)
HLA-DQA1	−0.49 (0.07, 14)	0.15 (0.62, 14)	0.46 (0.05, 18)
HLA-DRA	0.37 (0.13, 18)	0.63 (0.005, 18)	0.82 (<0.0001, 18)
HLA-DRB	0.36 (0.14, 18)	0.62 (0.007, 18)	0.52 (0.03, 18)
ACTN1	0.24 (0.34, 18)	0.17 (0.51, 18)	0.68 (0.002, 18)
COL3A1	0.52 (0.03, 18)	0.54 (0.02, 18)	0.86 (<0.0001, 18)
CTGF	0.34 (0.18, 17)	0.55 (0.02, 17)	0.72 (0.0007, 18)
FN1	0.60 (0.009, 18)	0.56 (0.02, 18)	0.96 (<0.0001, 18)
FAM38A	0.42 (0.09, 17)	−0.02 (0.93, 16)	0.24 (0.35, 17)
PLAU	0.54 (0.03, 17)	0.73 (0.0008, 17)	0.92 (<0.0001, 18)
c MYC	0.68 (0.002, 18)	0.64 (0.004, 18)	0.96 (<0.0001, 18)
BCL6	0.69 (0.002, 18)	0.66 (0.003, 18)	0.89 (<0.0001, 18)
NPM3	0.65 (0.007, 16)	0.52 (0.04, 16)	0.70 (0.001, 18)
BMP6	0.34 (0.17, 18)	0.39 (0.11, 18)	−0.09 (0.71, 18)
LMO2	0.65 (0.005, 17)	0.74 (0.0006, 17)	0.98 (<0.0001, 18)
BCL2	0.61 (0.009, 17)	0.67 (0.003, 17)	0.79 (0.0001, 18)
CCL3	0.30 (0.22, 18)	0.38 (0.11, 18)	0.77 (0.0002, 18)
CCND2	0.76 (0.0003, 18)	0.70 (0.001, 18)	0.96 (<0.0001, 18)
DRP2	−0.26 (0.34, 15)	Missing	Missing
PRKCB1	0.26 (0.30, 18)	0.06 (0.82, 18)	0.83 (<0.0001, 18)
PDCD4	0.89 (<0.0001, 18)	0.86 (<0.0001, 18)	0.91 (<0.0001, 18)
MAP1B	0.38 (0.19, 14)	−0.03 (0.91, 13)	0.27 (0.29, 17)
TLE1	0.85 (<0.0001, 16)	Missing	Missing
SLC25A13	0.50 (0.03, 18)	Missing	Missing
PDE4B	0.59 (0.01, 18)	0.55 (0.02, 18)	0.64 (0.004, 18)
C20orf155	0.30 (0.23, 18)	0.37 (0.15, 17)	0.23 (0.38, 17)
PRKCG	0.18 (0.59, 11)	Missing	Missing
OVGP1	NA, <10	Missing	Missing
NR4A3	0.42 (0.08, 18)	0.17 (0.52, 17)	0.30 (0.25, 17)
ZNF212	0.22 (0.46, 14)	Missing	Missing
HTR2B	NA, <10	Missing	Missing
B4GALT1	0.33 (0.18, 18)	0.36 (0.14, 18)	0.64 (0.005, 18)
CAT	0.37 (0.13, 18)	0.45 (0.06, 18)	0.75 (0.0004, 18)
CD19	0.77 (0.0002, 18)	0.77 (0.0002, 18)	0.92 (<0.0001, 18)
MS4A1	0.53 (0.02, 18)	0.56 (0.01, 18)	0.97 (<0.0001, 18)
CD3δ	0.53 (0.02, 18)	0.63 (0.005, 18)	0.96 (<0.0001, 18)
CD68	0.32 (0.19, 18)	0.36 (0.14, 18)	0.47 (0.05, 18)
OLIGO DT	Missing	Missing	Missing

Table 3 Continued

Gene	Correlation (*P*-value, *n*)		
	ArrayPlate with Affymetrix	ArrayPlate with LymphoChip	Affymetrix with LymphoChip
SOD2	0.51 (0.03, 18)	0.50 (0.03, 18)	0.87 (<0.0001, 18)
CYTOX	Missing	Missing	Missing
TBP	Missing	Missing	Missing
Median[a]	0.52	0.55	0.78

Bold font shows those genes with all three comparisons present and significant.

[a]Medians include only those genes with all three comparisons available.

method. For hazard ratios >1, there is an increased risk of death and for ratios <1, there is a decreased risk of death. These hazard ratios usually trended in the same direction for each gene as they did in the larger data sets from which their significance was derived. Table 4 shows two-way comparisons of agreement (results for hazard ratios above or below 1) between each of the three different quantitative mRNA methods and the directions predicted by individual genes for outcome. For methods, the Affymetrix and Lymphochip microarrays agree 88% in this data set, ArrayPlate agrees with Affymetrix 83%, and with Lymphochip 87%. Agreement with previously predicted outcome is 79% for Affymetrix, 91% for Lymphochip, and 80% for ArrayPlate. The higher agreement of Lymphochip methods with the predictors is expected because the greatest number of prognostic genes (17) was derived from Lymphochip data. Altogether, there are 19 genes concordant and predictive by hazard ratio for all three methods, and four more concordant and predictive for the ArrayPlate and one of the other microarray methods, while not being present on the other microarray. These genes are divided into signatures as follows: 12/17 Rosenwald signature genes, 5/6 Lossos, 6/13 Shipp, 1 Tome, suggesting that the Rosenwald and Lossos signatures are more robust.

DISCUSSION
The medical importance of developing a prognostic model in DLBCL, which moves beyond the currently used clinical IPI staging system, is obvious given the broad range of outcomes for similarly staged patients. DLBCL prognosis has been the focus of recent intense research and publication in this field. Several groups have investigated and created their own methods of outcome prediction based on the expression of from 6 to 17 genes quantified from samples of snap-frozen tumor biopsies.[2–4,6] We did not attempt to select new genes of interest. Our goal was to take the best leads from the
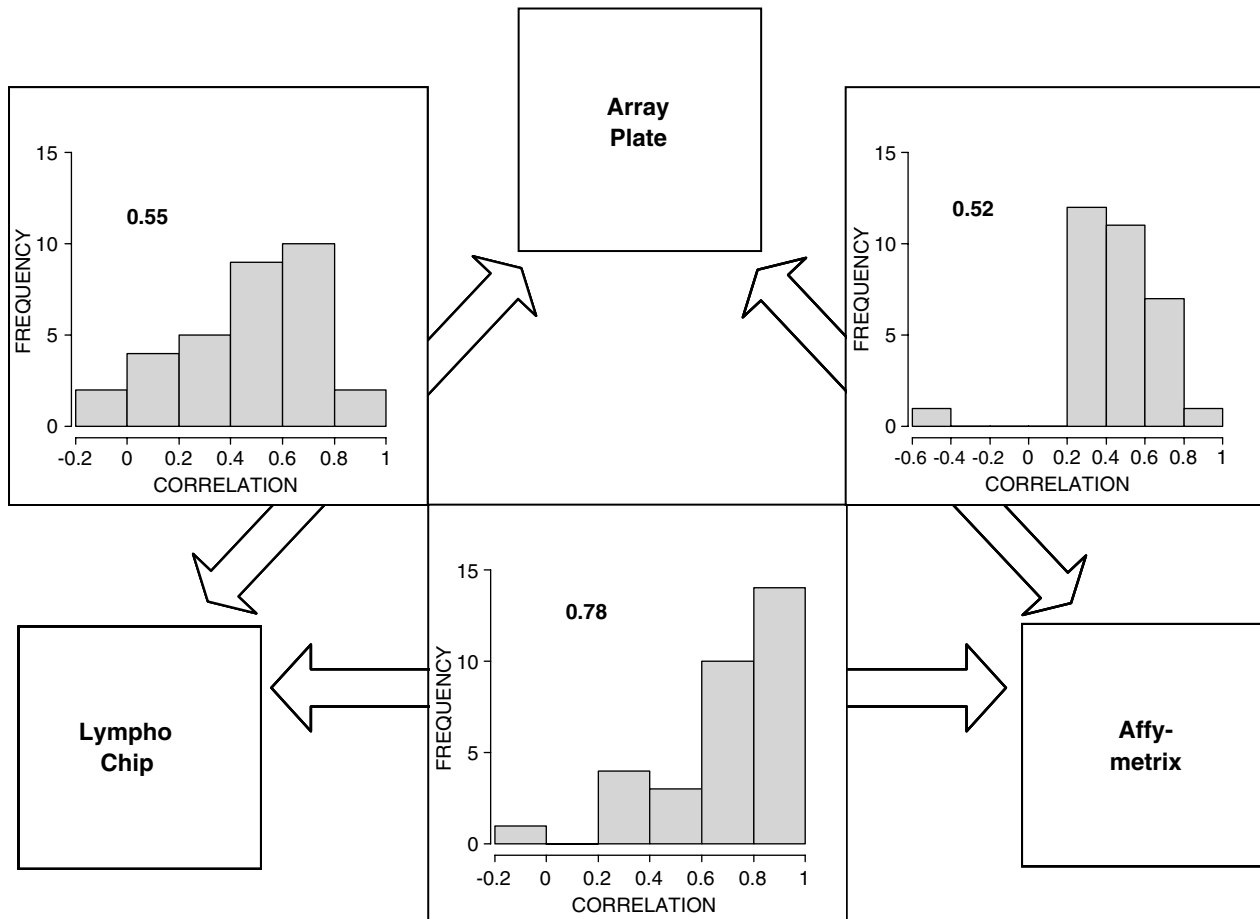
Figure 11 Correlations between the ArrayPlate assay and the LymphoChip and Affymetrix microarrays. Histograms show the distribution of correlations for each pair of comparisons across the gene sets, and the median correlation for each is indicated.

previously published comprehensive microarray studies and see if the key genes could be quantified in FFPE tissues. The significance of the current work is in the development of an assay that can be used for outcome prediction on routinely fixed, paraffin-embedded tissues, which is the only material available for analysis for many patients.

The ArrayPlate assay developed for this study was extremely reproducible with low %CVs on replicate and repeated assays. The correlation among the various types of preparations, including snap-frozen pellets in OCT, fresh samples, and old and new paraffin blocks, was very good. The ability to obtain similar results whether using frozen or paraffin-embedded materials is particularly notable and most likely results from measuring both soluble and cross-linked mRNA from the fixed tissues. All blocks analyzed in the current work had been fixed and handled at the University Medical Center, Tucson, which may have improved their quality; however, we have since analyzed an additional 120 FFPE blocks from a variety of hospitals with good success (data not shown). A typical qRT-PCR procedure using paraffin-embedded materials requires a long series of steps to first isolate RNA, which can include deparaffinization,

protease digestion, phenol:chloroform extraction, alcohol precipitation of the nucleic acids, followed by resuspension, DNase digestion, a second phenol:chloroform extraction, alcohol precipitation, and resuspension. The mRNA template is then reverse transcribed to create cDNA that undergoes PCR amplification. In the ArrayPlate assay, the mRNA is hybridized *in situ* therefore bypassing several steps in the qRT-PCR process, including extraction and amplification, which can both introduce bias.

The comparison of an old paraffin block (18 years) *vs* a newly created paraffin block is particularly important since it demonstrates the feasibility of using the ArrayPlate assay to test archival material from previous clinical trials or stored patient samples to validate gene expression correlations to survival from discovery experiments. This aspect may be critical to generating meaningful results from patients who may not otherwise have available material for techniques that require snap-frozen samples, such as patients who are biopsied in smaller hospitals or outlying clinics, those with minimal tissue available for biopsy, or those in whom the possibility of lymphoma was not anticipated before biopsy so that tissue was not snap-frozen.

Table 4 Agreement between hazard ratios for different technologies and predicted outcomes[a]

	LymphoChip <1	LymphoChip >1	
Affymetrix <1	**15**	1	33 comparisons
Affymetrix >1	3	**14**	29 agreement
	Affymetrix <1	Affymetrix >1	
ArrayPlate <1	**14**	3	36 comparisons
ArrayPlate >1	3	**16**	30 agreement
	LymphoChip <1	LymphoChip >1	
ArrayPlate <1	**14**	2	30 comparisons
ArrayPlate >1	2	**12**	26 agreement
	Predicted <1	Predicted >1	
Affymetrix <1	**17**	1	38 comparisons
Affymetrix >1	7	**13**	30 agreement
	Predicted <1	Predicted >1	
LymphoChip <1	**18**	0	32 comparisons
LymphoChip >1	3	**11**	29 agreement
	Predicted <1	Predicted >1	
ArrayPlate <1	**16**	1	35 comparisons
ArrayPlate >1	6	**12**	28 agreement

[a]This table shows two-way comparisons of agreement (results for hazard ratios above or below 1) between each of the three different methods used on the 39 samples analyzed and between the methods and the predicted values from the different gene signatures. Bold font shows those genes in agreement. Genes with missing data on one of two assessment techniques or for which the hazard ratio was 1 ($|HR-1| < 0.01$) were excluded from the agreement tables. For methods, the Affymetrix and LymphoChip microarrays agree 88% in this data set, ArrayPlate agrees with Affymetrix 83% and with LymphoChip 87%. Agreement with previously predicted outcome is 79% for Affymetrix, 91% for LymphoChip, and 80% for ArrayPlate.

The level of agreement between the ArrayPlate and both the LymphoChip and Affymetrix arrays, as shown by the kappa statistics, suggests that ArrayPlate technology may be suited to detect outcome associations. While confidence intervals are wide because of the limited number of *de novo* samples tested, we are collecting samples from a much larger cohort of patients for further testing.

While we were not able to demonstrate a direct relationship between the ArrayPlate and patient survival, attributed to the low number of *de novo* DLBCL samples in the study, we were able to demonstrate that the hazard ratios agreed with predictions from the larger data sets of Rosenwald, Shipp, Tome, and Lossos 80% of the time, similar to the microarray methods, indicating that it may be possible for

clinically significant information to be generated with this technique. Analysis of a larger data set is underway.

In summary, this is an early report on use of the Array-Plate, a quantitative gene expression protocol for testing paraffin-embedded materials. While this technique does not replace the strength of GEP, which can analyze the entire genome for discovery studies, it does indicate that the significant genes identified by GEP studies can be adapted into a useful platform applicable to nearly all patients. The platform is easily modifiable and can be adjusted as new genes of interest are discovered.

DISCLOSURES/DUALITY OF INTEREST
Thomas Grogan is the Chief Medical Officer of Ventana Medical Systems Inc. (Tucson, AZ, USA), which markets the immunohistochemistry stainer used in this study. Thomas Miller has a financial interest in Ventana Medical Systems Inc., which markets the immunohistochemistry stainer used in this study. Ralph Martel, Constantine Sabalos, Ihab Botros, Matthew Rounseville, and Bruce Seligman are employed by High Throughput Genomics (Tucson, AZ, USA), which developed and owns the ArrayPlate assay used in this study.

1. Shipp MA, Harrington DP, Anderson JR, et al. A predictive model for aggressive non-Hodgkin's lymphoma. N Engl J Med 1993;329:987–994.
2. Rosenwald A, Wright G, Chan WC, et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. N Engl J Med 2002;346:1937–1947.
3. Shipp MA, Ross KN, Tamayo P, et al. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. Nat Med 2002;8:68–74.
4. Lossos IS, Czerwinski DK, Alizadeh AA, et al. Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. N Engl J Med 2004;350:1828–1837.
5. Alizadeh A, Eisen M, Davis RE, et al. The lymphochip: a specialized cDNA microarray for the genomic-scale analysis of gene expression in normal and malignant lymphocytes. Cold Spring Harb Symp Quant Biol 1999;64:71–78.
6. Tome ME, Johnson DBF, Rimsza LM, et al. A redox signature score identifies diffuse large B-cell lymphoma patients with a poor prognosis. Blood 2005;106:3594–3601.
7. Martel RR, Botros IW, Rounseville MP, et al. Multiplexed screening assay for mRNA combining nuclease protection with luminescent array detection. Assay Drug Dev Technol 2002;1:61–71.
8. Accola RS. Human B cell variants immunoselected against a single Ia antigen subset have lost expression of several Ia antigen subsets. J Exp Med 1983;157:1053–1058.
9. Glinsmann-Gibson BJ, Cycon KA, Murphy SP, et al. Induced major histocompatibility class II (MHCII) expression does not alter chemosensitivity, radiosensitivity, redox potential, or proliferation rate in a CIITA-transfected diffuse large B cell lymphoma (DLBCL) cell line [abstract]. Blood 2006;108:578a.
10. Lossos IS, Czerwinski DK, Wechser MA, et al. Optimization of quantitative real-time RT-PCR parameters for the study of lymphoid malignancies. Leukemia 2003;17:789–795.
11. Cox D. Regression models and life tables. J R Stat Soc B 1972;B34:187–220.
12. Fleiss JL. The measurement of interrater agreement. Statistical Methods for Rates and Proportions. John Wiley & Sons Inc.: New York, NY, 1981, pp 212–236.

13. Efron B. The Jackknife, the Bootstrap, and Other Resampling Plans. [38]. CBMS-NSF Monographs. Society of Industrial and Applied Mathematics: Philadelphia, PA, 1982.

14. Sawada H, Taniguchi K, Takami K. Improved toxicogenomic screening for drug-induced phospholipidosis using a multiplexed quantitative gene expression ArrayPlate assay. Toxicol In vitro 2006;20:1506–1513.

15. Irizarry RA, Warren D, Spencer F, *et al*. Multiple-laboratory comparison of microarray platforms. Nat Methods 2005;2:345–350.

16. Kuo WP, Jenssen TK, Butte AJ, *et al*. Analysis of matched mRNA measurements from two different microarray technologies. Bioinformatics 2002;18:405–412.

17. Landis RC, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–174.