Article:
Ahmed, S orcid.org/0000-0003-4912-746X, Glover, P, Taylor, J et al. (8 more authors) (2021) Comparative analysis of gene expression platforms for cell-of-origin classification of diffuse large B-cell lymphoma shows high concordance. British Journal of Haematology, 192 (3). bjh.17246. pp. 599-604. ISSN 0007-1048

https://doi.org/10.1111/bjh.17246

Concordance of gene expression profiling approaches for cell-of-origin classification in diffuse large B-cell lymphoma

Sophia Ahmed[1,2], Paul Glover[2], Jan Taylor[2], Chulin Sha[1], Matthew A. Care[3], Reuben Tooze[2,3], Andrew Davies[4], David R. Westhead[1], Peter W.M. Johnson[4], Catherine Burton[2], Sharon Barrans[2]

[1]Faculty of Biological Sciences, University of Leeds, Leeds, UK.

[2]Haematological Malignancy Diagnostic Service, St James' University Hospital, Leeds, UK.

[3]Faculty of Medicine, University of Leeds, Leeds, UK.

[4]Cancer Research UK Centre and Southampton Clinical Trials Unit, University of Southampton, Southampton, UK.

Abstract

Cell-of-origin sub-classification of diffuse large B cell lymphoma (DLBCL) into activated B cell-like (ABC), germinal centre B cell-like (GCB) and unclassified or type III (UNC) by gene expression profiling is recommended in the latest update to the WHO classification of lymphoid neoplasms. There is however no accepted gold standard method or data set for this classification. Here we compare classification results using gene expression data for 68 formalin-fixed paraffin-embedded DLBCL samples measured on four different gene expression platforms (Illumina WG-DASL™ arrays, Affymetrix PrimeView arrays, Illumina TrueSeq RNA sequencing and the HTG EdgeSeq DLBCL Cell of Origin Assay EU) using an established platform agnostic classification algorithm (DAC) and the classifier native to the HTG platform, which is CE-IvD approved for *in vitro* diagnostic use. Classification methods and platforms show a high level of concordance, with agreement in at least 80% of cases and rising to much higher levels for classifications of high confidence. Our results demonstrate that COO classification by gene expression profiling on different platforms is robust, and that the use of the confidence value alongside the classification result is important in clinical applications.

Introduction

Diffuse large B cell lymphoma (DLBCL) is the most common type of non-Hodgkin lymphoma and has been revealed to consist of distinct subtypes on the basis of gene expression patterns that reflect the putative cell of origin (COO) [1]. The two main recognised COO subtypes within DLBCL are activated B-cell-like (ABC) and germinal centre B-cell-like (GCB) with a third category referred to as unclassified or Type III (UNC). GCB generally has better prognosis than ABC following standard R-CHOP chemotherapy, and this been consistently replicated in many studies using gene expression profiling (GEP) to assign COO groups [2].

Gene expression profiling can now be applied to routinely processed formalin fixed paraffin embedded (FFPE) diagnostic tissue biopsies, but despite this it has not been widely incorporated into routine clinical use, and the surrogate immunohistochemistry (IHC) based Hans test remains as standard practice. This uses just three markers (CD10, BCL6 and IRF4/MUM1) to classify patient samples as either GCB or non-GCB; however, reproducibility has proved difficult and this classification does not identify significant differences in overall survival [3]. COO classification has now been recognised in the latest 2016 update of the WHO classification of lymphoid neoplasms [4], which states that COO should be defined preferably by GEP, and recommends Hans IHC only where this is not possible.

We applied real-time COO classification, using the Illumina whole genome cDNA-mediated annealing, selection, extension and ligation (WG- DASL^{TM}) gene expression profiling assay and DAC classifier [5], to patients enrolled in the Randomised Evaluation of Molecular guided therapy for Diffuse Large B-cell Lymphoma with Bortezomib (REMoDL-B) study (NCT01324596) [6]. This aimed to evaluate the clinical efficacy of the combination of bortezomib with R-CHOP and to determine whether the COO subtypes respond differently. This study was the first large-scale study in DLBCL to use real-time molecular characterisation for prospective stratification and randomisation, and subsequent analysis of biologically distinct subgroups.

Since the initiation of this trial a number of commercially available platforms for COO assignment by GEP have emerged, including Lymph2Cx (Nanostring Technologies, WA) and HTG EdgeSeq DLBCL Cell of Origin Assay EU (HTG Molecular Diagnostics Inc., Tucson, AZ), and some authors have developed classifiers for use with RNA-seq data (e.g. [7]). It is important to appreciate that there is no gold standard for COO assignment, and all studies show a spectrum of gene expression patterns with a significant 'grey zone' of intermediate cases. Intermediate cases that fall close to a classification boundary have limited biological difference to cases falling just on the other side of the boundary, and as such, concordance

between different COO assignment methods may not be observed, nor should they necessarily be expected. However, to validate the technique for clinical practice it is important to understand the expected degree of agreement between different GEP platforms and classification algorithms, and how this relates to measures of classification confidence.

The aim of this study was to evaluate reproducibility of COO classification in DLBCL using different GEP platforms with the platform agnostic DAC algorithm [5] for potential implementation in routine use. We investigated an Affymetrix array based method and RNA-seq data, as well as the HTG EdgeSeq DLBCL Cell of Origin Assay EU with its native CE-IvD classifier and compared the results with those obtained using the now withdrawn WG-DASL™ GEP platform classified with the DAC (herein referred to as DASL_DAC).

Materials and Methods

Sample Selection

DASL_DAC classification had already been performed for the >1000 samples in the ReMoDLB trial [6]. From these, a subset of representative samples was selected based on the overall distribution of COO classes and classification confidences (n=119 in total (ABC n=33, Unc n=27, GCB n=59). Of these n=68 (ABC n=20, Unc n=10, GCB n=38) had adequate RNA and data QC from all platforms and were used in the final analysis (see Table S1).

RNA extraction

Total RNA was extracted from 5µm paraffin sections using the Ambion RecoverAll™ kit standard protocol, with an extended 16-hour protease digestion. RNA was assessed using either a NanoDrop Spectrophotometer or Qubit (Thermo Fisher Scientific) or TapeStation Genetic Analyser (Agilent), with 260:280 ratio or DV200 measurements for quality assessment, as per requirements of the expression platform manufacturer. Illumina WG-DASL™ and HTG EdgeSeq platforms did not require quality assessment of RNA, but Affymetrix Primeview arrays required verification of RNA purity by 260:280 ratio. The Illumina RNA Exome kit required quantitation performed using the Qubit Flourometer RNA quantitation high sensitivity kit (Thermo Fisher Scientific) and quality was assessed by Tapestation (Agilent) DV200 readings, which were required to be of 30 or above. Details of input requirements, represented genes, and data output are provided in Table S1.

Gene Expression Profiling (GEP)

Gene expression profiling methods tested in this study included four commercially available products; the now withdrawn Illumina WG-DASL™ Array (DASL), the Affymetrix Primeview Array, the HTG EdgeSeq DLBCL Cell of Origin Assay EU and an RNA sequencing approach, the Illumina TruSeq RNA Exome library (RNA-Seq) (see Supplemental Information for a

description of each method and Table S2 for the data acquisition and analysis pipeline). DASL was performed in real-time during REMoDL-B trial recruitment (Sept 2011-May 2015). GEP on Affymetrix and RNA-seq platforms was performed on the same stored RNA that was used for DASL. HTG EdgeSeq was carried out on either HTG recommended unstained sections (where available) or the same stored RNA (n=48 *vs* n=20 respectively). As part of the platform validation, paired RNA and FFPE sections from the same sample were processed on the HTG platform. Concordant results were generated in 13/14 pairs and only one sample was borderline and switched between class (RNA COO=GCB, FFPE section COO=ABC both classified with HTG EdgeSeq, compared with RNA COO DASL_DAC = GCB). This sample was not included in the final set of 68 samples. Prior to downstream COO classification, expression data from DASL, Affymetrix and RNAseq was quantile normalised using the limma package [8] implemented in R.4.2 [9].

Cell of Origin classification

Cell of Origin (COO) classification was performed using the DLBCL automated classifier (DAC) [5] with expression data from the DASL, Affymetrix and RNA-seq platforms. As well as providing a COO output, the classification is also associated with a classification probability (P)) for each class (P_{ABC}, P_{UNC} and P_{GCB} with $P_{ABC}+P_{UNC}+P_{GCB}=1$) and quality control metrics. In each case the assigned class is the one with highest probability, and here we define a log odds measure of classification confidence $L = \log_2(P_x/(1-P_x))$ where P_x is the probability of the assigned class. For example, a case with P_{ABC}=0.4 P_{UNC} =0.3 and P_{GCB} = 0.3 would be assigned as ABC with $L = \log_2(0.4/0.6) = -0.6$, and considered a 'weak' assignment (defined as $L < 0$) with the other two classes collectively more probable than the assigned class. Alternatively, P_{ABC}=0.8 P_{UNC} =0.1 and P_{GCB} = 0.1 has $L = \log_2(0.8/(0.2) = 2.0$ and is a 'confident' assignment (defined as $L \geq 1$: assigned class more than twice as probable as the other two collectively). Other moderate confidence assignments have $0 \leq L < 1$.

The HTG EdgeSeq DLBCL panel measures the expression of 92 genes, but with only 13 of 20 DAC genes represented. While DAC can be used with a subset of its classifier genes, such an analysis is not presented here since it is not possible to quantify the effect of missing genes on the final assignments. Instead, for these data, we used the COO classification algorithm native to the HTG platform for comparison with the DAC assignments above. The HTG EdgeSeq DLBCL Cell of Origin Assay EU classifier is CE-IvD approved for *in vitro* diagnostic use. We note that the HTG COO classifier is trained to minimise the UNC class compared with other methods (https://www.freepatentsonline.co/y2018/0340231.html), preferring classification as either ABC or GCB: in this data set the HTG COO classifier calls 4 samples UNC, compared to an average of 9 in UNC for DAC.

Results

Full cell of origin classification results can be found in Figure 1 and Supplementary Table S3. The heatmap in Figure 1 shows that the same pattern of COO related gene expression over the gene set used by the DAC classifier is detected by each of the four gene expression measurement technologies (limited in the HTG data to those genes available on the platform). Figure 1 also reveals the broad trend that classifications that are discordant between two or more platforms tend to be ABC/UNC or GCB/UNC disagreements, rather than ABC/GCB, and tend to occur in cases with lower confidence on the DASL_DAC platform. The number of concordant and discordant assignments are further elucidated by the confusion metrices provided in Table S4.

First, we compare classifications from the DAC algorithm with the CE-IvD HTG EdgeSeq DLBCL COO Assay and classifier. Using DAC with gene expression data from the DASL platform (where DAC is most extensively validated, particularly in the REMoDL-B trial) we find 58/68 (85%) agreement on the COO class (Figure 1). However, 7 of the10 of the samples where classifications did not agree were classed as UNC by DASL_DAC and either ABC or GCB by HTG. We commented above that the HTG classifier is trained to minimize the UNC class, so these disagreements are not unexpected. In view of this, comparing only on cases not classed as UNC by DASL_DAC gives agreement of 55 of 58 (95% $\pm$ 5), and of the three discordant cases two are DASL_DAC weak confidence assignments ($L < 0$) and one is moderate confidence ($0 \leq L < 1$). Comparison of HTG with DAC derived from Affymetrix and RNA-seq gene expression platforms shows similar results, with, respectively, 82 $\pm$ 7 % and 79 $\pm$ 8% overall concordance, increasing to 88% $\pm$ 7 and 86% $\pm$ 8 when discounting cases assigned UNC by DAC. Overall therefore, there is a good level of concordance between HTG and the combination of DAC with any gene expression platform, with discordance generally associated with lower confidence cases and dominated by cases involving the UNC class.

Having confirmed the high level of agreement between DAC classifications and those from HTG, we next moved to a like-for-like comparison of the DAC algorithm applied to gene expression data from DASL arrays, Affymetrix arrays and RNA-seq. In this three-way comparison, overall 53 of 68 (78% $\pm$ 8) had the same class on all three platforms. The remaining 15 cases all agreed on two platforms, with the single discordant result approximately evenly spread between the platforms (five RNA-seq, seven DASL, three Affymetrix). The distributions of classification confidences are shown in Figure 2, where it is notable that the distribution of log-odds values is different depending on both platform and assigned class, and that GCB classifications are consistently of higher confidence. Of the 15 discordant cases (highlighted on Figure 2), 12 were ABC/UNC or GCB/UNC disagreements and only 3 were ABC/GCB disagreements. The average log-odds confidence of the discordant

assignments was 0.6 compared to 1.6 for assignments from DAC overall, showing that discordant calls are significantly lower in confidence ($P<0.01$, t-test). Viewed another way, compared to a 78% 3-way concordance overall, eliminating 'weaker' assignments and considering only moderately confident and confident assignments yields 3-way concordance of 39/48 (81%) and considering only confident assignments yields 25/25 (100%) 3-way concordance. These results are consistent with discordant calls being mostly associated with lower confidence on the discordant platform, as well as lying near the classification boundaries between UNC and the main ABC/GCB classes.

Discussion

The results reported above show a good degree of agreement between the DAC algorithm applied to DASL data and the native COO classifier with the HTG Edgeseq gene expression platform. Most disagreements are cases in the UNC class from DAC, a class which the training of the HTG classifier aims to minimise, and are preferentially classified as ABC or GCB by that algorithm. We prefer an approach that does not minimise the UNC class, since there is evidence that it is more than a group of intermediate cases, in particular containing cases with T cell dominated immune response [10]. Otherwise disagreements tend to have low classification confidence from the DAC algorithm, indicating that they may lie close to classification boundaries and reflecting biological heterogeneity and/or ongoing differentiation within the tumour.

The results for the DAC algorithm, when used with different gene expression measurement technologies, reveals complete concordance across all three platforms approaching 80%, with disagreements similarly dominated by those involving the UNC class and having low confidence on the discordant platform. Performance of the three gene expression platforms was similar, with no platform dominating the discordance statistics. Since these studies were carried out at significantly different times, it is possible that even this relatively small level of disagreement is influenced to some extent by RNA deterioration with time.

While COO classification remains relevant in clinical practice, further GEP subgroups, molecular high grade (MHG)[11] and DHitSig[12] have been recently defined by our group and others. These groups are largely consistent between studies, and add further subgroups beyond the COO, with both groups identifying a poor prognostic group within GCB. The results of this study provide confidence in classification across different gene expression technologies, both in the classification of COO, but also moving forward with more recently described classification schema (MHG/DHitSig).

Furthermore, we have previously shown that biologically relevant mutations differentiate appropriately between COO classes, both in REMoDL-B[6] and in our population dataset[13], providing further confidence that the COO calls are not only reproducible, but also biologically important. Large-scale sequencing studies in DLBCL have identified mutation clusters as an alternative approach to subclassification of DLBCL[14], [15]. There is significant overlap between COO and the mutational clusters, providing additional knowledge of the underlying biology of these disorders, and if used in combination the vision is towards providing a more personalised medicine approach for individual patients

Acknowledgements

This work was supported by a grant from Blood Cancer UK (formerly Bloodwise), grant number 15002.

Authorship

Contribution: The study was designed by SB, RT, DRW, AD, PWMJ and CB. Laboratory work was carried out by SA, SB. Data was analysed by SA, PG, JT, CS, MAC and DRW. The paper was written by SA, SB, JT and DRW.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Dr Sharon Barrans, Haematological Malignancy Diagnostic Service, St James' University Hospital, Leeds, LS9 7TF, UK. Email: sharon.barrans@nhs.net

Figures captions

Figure 1. Gene expression patterns and classification results for 68 diffuse large B cell lymphoma samples (columns). The heat map (bottom) shows gene expression levels (blue:low;red:high) for the 20 gene signature used by the DAC cell-of-origin classifier, as measured on four different platforms [Illumina DASL, Affymetrix, Illumina RNA-seq, HTG Molecular]. For each platform the genes are divided in two groups, those up-regulated in GCB (top, 8 genes) and those up-regulated in ABC (bottom, 12 genes). Only the 13 DAC classifier genes that are present on the HTG EdgeSeq DLBCL panel are shown. The classification results (top) are from the DAC classifier with expression data from the DASL, Affymetrix and RNA-seq platforms, compared to the results from the HTG COO classifier applied to HTG platform expression data. Samples are sorted by the class on the DASL platform (yellow, ABC;

green, unclassified (UNC); blue, GCB) and ordered by classification probability on the same platform.

Figure 2. Box and Whisker plots of log odds confidence measures from the DAC classifier applied to gene expression data from three different platforms (Illumina WG-DASL™, left; Affymetrix, middle; Illumina RNA-seq, right) split according to the assigned class (yellow, ABC; blue, GCB; green; UNC). Confidence values for samples that are discordant on the platform in question are shown as red circles.

References

1. Alizadeh, A.A., et al., *Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling.* Nature, 2000. **403**(6769): p. 503-11.
2. Lenz, G., et al., *Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways.* Proc Natl Acad Sci U S A, 2008. **105**(36): p. 13520-5.
3. Read, J.A., et al., *Evaluating cell-of-origin subtype methods for predicting diffuse large B-cell lymphoma survival: a meta-analysis of gene expression profiling and immunohistochemistry algorithms.* Clin Lymphoma Myeloma Leuk, 2014. **14**(6): p. 460-467 e2.
4. Swerdlow, S.H., et al., *The 2016 revision of the World Health Organization classification of lymphoid neoplasms.* Blood, 2016. **127**(20): p. 2375-90.
5. Care, M.A., et al., *A microarray platform-independent classification tool for cell of origin class allows comparative analysis of gene expression in diffuse large B-cell lymphoma.* PLoS One, 2013. **8**(2): p. e55895.
6. Davies, A., et al., *Gene-expression profiling of bortezomib added to standard chemoimmunotherapy for diffuse large B-cell lymphoma (REMoDL-B): an open-label, randomised, phase 3 trial.* Lancet Oncol, 2019. **20**(5): p. 649-662.
7. Reddy, A., et al., *Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma.* Cell, 2017. **171**(2): p. 481-494 e15.
8. Ritchie, M.E.P., B ;Wu, D; et al., *limma powers differential expression analyses for RNA-sequencing and microarray studies.* Nucleic Acids Research, 2015. **43**(7).
9. Team, R.C., *R: A language and environment for statistical computing*. 2013: R Foundation for Statistical Computing, Vienna, Austria.
10. Care, M.A., D.R. Westhead, and R.M. Tooze, *Gene expression meta-analysis reveals immune response convergence on the IFNgamma-STAT1-IRF1 axis and adaptive immune resistance mechanisms in lymphoma.* Genome Med, 2015. **7**: p. 96.
11. Sha, C.L., et al., *Molecular High-Grade B-Cell Lymphoma: Defining a Poor-Risk Group That Requires Different Approaches to Therapy.* Journal of Clinical Oncology, 2019. **37**(3): p. 202-+.
12. Ennishi, D., et al., *Double-Hit Gene Expression Signature Defines a Distinct Subgroup of Germinal Center B-Cell-Like Diffuse Large B-Cell Lymphoma.* J Clin Oncol, 2019. **37**(3): p. 190-201.
13. Lacy, S.E., et al., *Targeted sequencing in DLBCL, molecular subtypes, and outcomes: a Haematological Malignancy Research Network report.* Blood, 2020. **135**(20): p. 1759-1771.

14. Chapuy, B., et al., *Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes.* Nature Medicine, 2018. **24**(5): p. 679-+.
15. Schmitz, R., et al., *Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma.* New England Journal of Medicine, 2018. **378**(15): p. 1396-1407.

Figure 1.

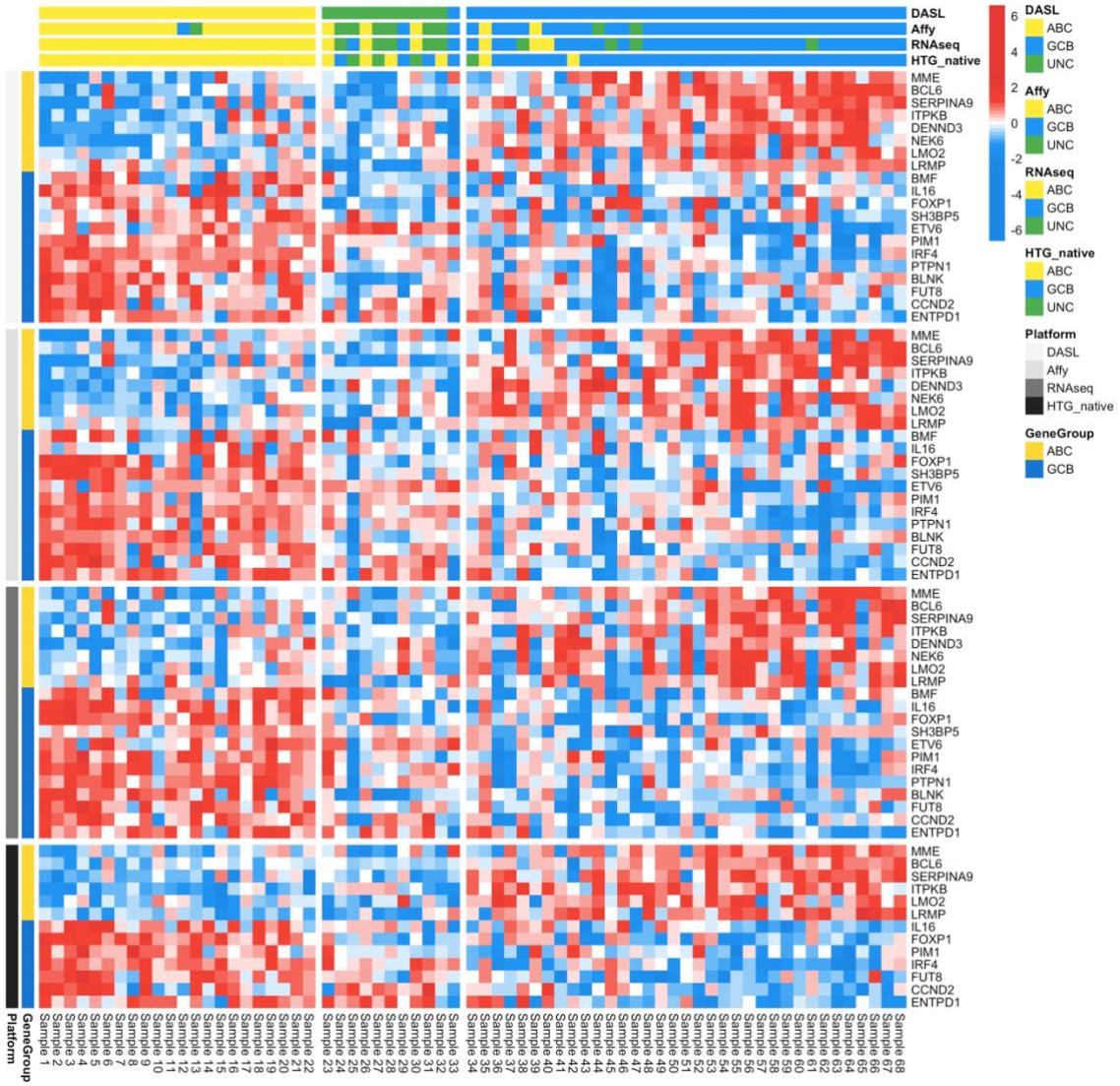
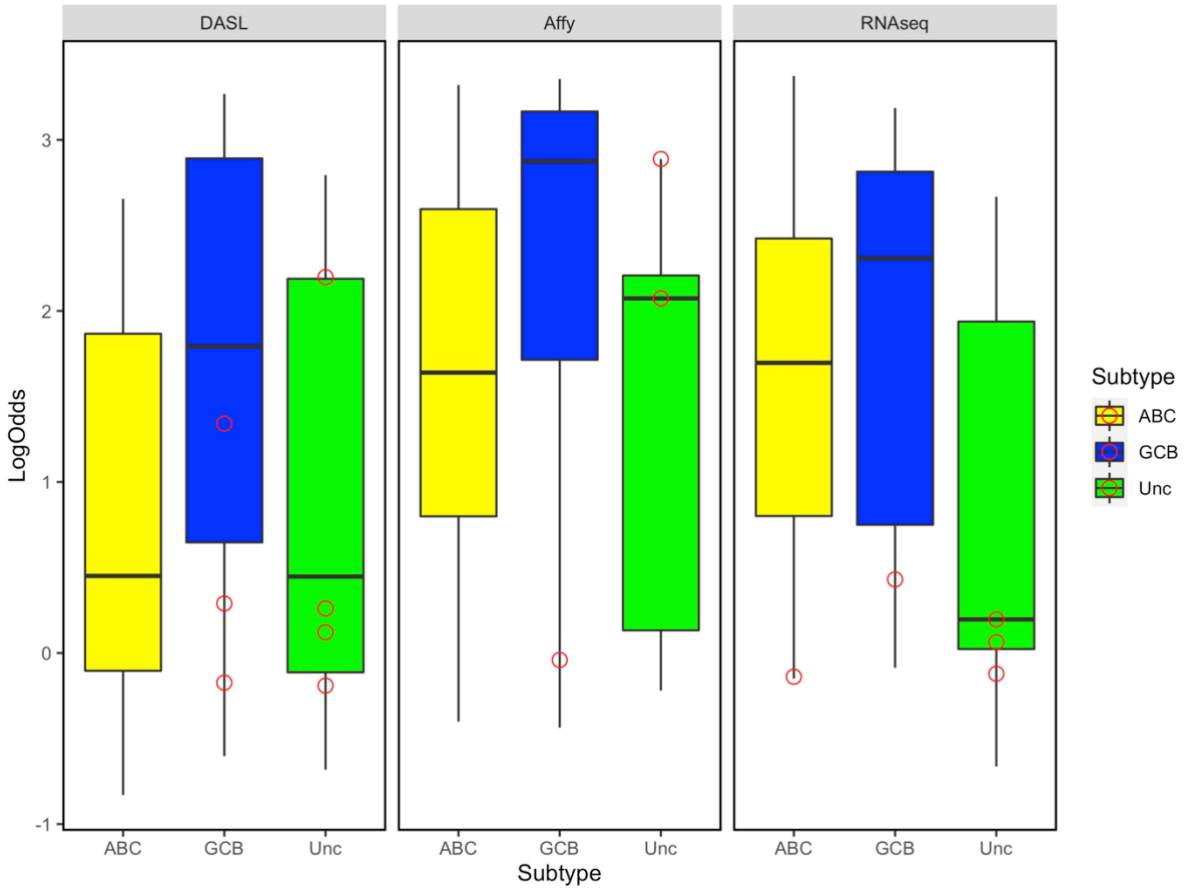


Figure 2


DASL
Affy
RNAseq
LogOdds
3
2
1
0
-1
ABC
GCB
Unc
Subtype
Subtype
ABC
GCB
Unc

Comparative analysis of gene expression platforms for subtyping DLBCL shows high concordance.

Sophia Ahmed, Paul Glover, Jan Taylor, Chulin Sha, Matthew Care, Reuben Tooze, Andrew Davies, David Westhead, Peter Johnson, Cathy Burton, Sharon Barrans.

Supplementary Materials and Methods.

Affymetrix Transcriptome Arrays.

Transcriptome Arrays were generated using Affymetrix Primeview Human Gene Expression Assay following the manufacturer's protocol. They were stained and washed using the Affymetrix 450 Fluidics Station and arrays were scanned using the GeneChip® Scanner 3000 7 G (Affymetrix Inc.) to produce cel files for gene expression.

DASL

WG-DASL was performed according to Illumina protocols using 200 ng total RNA, and HumanRef-8 V3 arrays. Arrays were scanned on a BeadArray reader, data processed using GenomeStudio (Illumina United Kingdom), and exported as a text file.

HTG

HTG gene expression data was generated using the HTG EdgeSeq DLBCL Cell of Origin Assay. Starting material was either extracted RNA as described above or tissue sections from slides according to the manufactures protocol. Final libraries were pooled according to the HTG recommendations prepared for sequencing with the MiSeq Reagent Kit v2 300 cycles (Illumina).

RNAseq

Libraries were prepared using the Illumina RNA Exome kit following the manufacturer's protocol. Final library four-plex pools were quantitated using the Qubit Fluorimeter, and also checked for quality using the Agilent tapestation. Library pools were then further pooled into groups of 16 samples for sequencing with TG NextSeq® 500/550 High Output Kit v2 150 cycles (Illumina).

References

1. Andrew, S., *FastQC: A Quality Control Tool for High Throughput Sequence Data* 2010: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
2. Martin, M., *Cutadapt removes adapter sequences from high-throughput sequencing reads.* EMBnet.Journal, 2011. **17**(1): p. 10-12.
3. Dobin, A.D., C.A.; Schlesinger, F; Drenkow,J; Zaleski, C; Jha, S; Batut, P; Chaison, M; Gingeras, T.R., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.
4. Liao , Y.S., G.K.; Shi W., *featureCounts: an efficient general purpose program for assigning sequence reads to genomic features.* Bioinformatics, 2014. **30**(7): p. 923-930.
5. Gautier, L.C., L; Bolstad, Benjamin, B.; Irizarry, R, *affy—analysis of Affymetrix GeneChip data at the probe level.* Bioinformatics, 2004. **20**(3): p. 307-315.

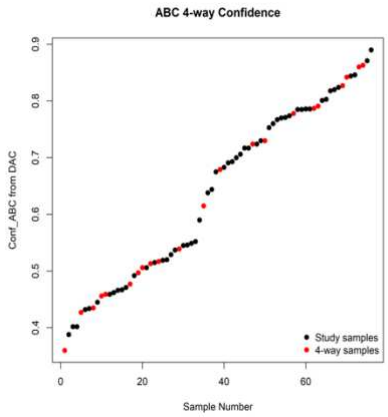
Supplemental Table S1: Samples grid for case selection.

	DASL	HTG	Affy	RNAseq
DASL (D)	286			
HTG (H)	237	248		
Affy (A)	137	115	137	
RNAseq (R)	85	74	69	90
D / H / A	115			
D / H / A / R	68			

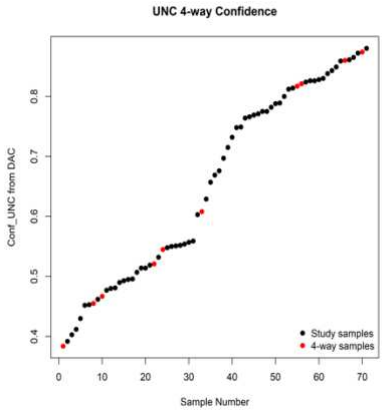
Supplemental Figure: Confidence values for the samples analysed by all 4 methods (red points, n=68) compared to the total number of samples in the study (black points, n=286). These plots show that the selected samples demonstrate a range of confidences as seen in the original dataset.

1a: ABC 4-way Confidence, **1b:** UNC 4-way Confidence, **1c:** GCB 4-way Confidence.

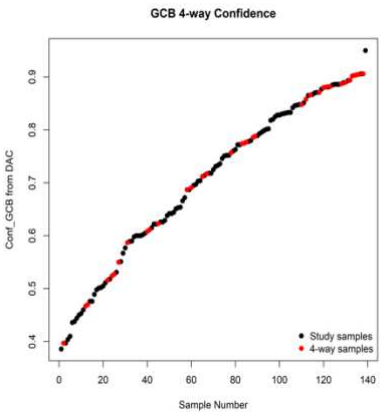
1a



1b



1c



Supplemental Table S2: Technical Comparison of Gene Expression Platforms

	Platform (cost per sample)	Input requirements			Pipeline for obtaining gene expression values and COO classification		
		Input type	Quantity	QC	Classifier genes represented		Software/pipeline for analysis
					genes	DAC	
Illumina	DASL (£118)	RNA	>50ng	NA	27,000	20/20	Expression analysis in Illumina Genomestudio; COO classification in DAC.
	RNA Exome (£390)	RNA	20–100ng	DV200 > 30%	21,415	20/20	Fastq files checked in fastqc [1] Quality filtering and trimming in Cutadapt [2]. Libraries aligned to human genome Gencode release 28 (GRCh38.p12) using STAR aligner [3]. Raw counts per gene were generated using the featureCounts [4] implemented in Rsubreads. COO classification in DAC.
Affymetrix	Primeview (£250)	RNA	50ng 20–200ng	260:280 (≥1.8).	>20,000	20/20	Expression data obtained from gene level intensities, extracted from the cel files using the Affy package [5] R v3.4.2. COO classification in DAC.
HTG	HTG-DLBCL panel (£235)	RNA or FFPE section	25ng RNA/ 0.5mm^2 section	NA	96	13/20	HTG_native COO classifier

Supplemental Table S3. COO classification of n=68 DLBCL samples on four GEP platforms. All gene expression data from the different platforms were analysed using the DAC (DLBCL Automatic Classifier) except HTG data (final column) was classified using the HTG native onboard analysis. DAC confidence for DASL classification is shown in column two. Mismatches are highlighted in red (ABC/GCB and mismatch involving unclassified respectively).

Sample	DASL	DASL confidence	Affy	RNAseq	HTG
1	ABC	0.863	ABC	ABC	ABC
2	ABC	0.86	ABC	ABC	ABC
3	ABC	0.842	ABC	ABC	ABC
4	ABC	0.827	ABC	ABC	ABC
5	ABC	0.791	ABC	ABC	ABC
6	ABC	0.787	ABC	ABC	ABC
7	ABC	0.778	ABC	ABC	ABC
8	ABC	0.73	ABC	ABC	ABC
9	ABC	0.679	ABC	ABC	ABC
10	ABC	0.724	ABC	ABC	ABC
11	ABC	0.615	ABC	ABC	ABC
12	ABC	0.539	GCB	ABC	ABC
13	ABC	0.517	UNC	ABC	ABC
14	ABC	0.513	ABC	ABC	ABC
15	ABC	0.506	ABC	ABC	ABC
16	ABC	0.497	ABC	ABC	ABC
17	ABC	0.477	ABC	ABC	ABC
18	ABC	0.459	ABC	ABC	ABC
19	ABC	0.456	ABC	ABC	ABC
20	ABC	0.435	ABC	ABC	ABC
21	ABC	0.427	ABC	ABC	ABC
22	ABC	0.36	ABC	ABC	ABC
23	UNC	0.545	ABC	ABC	ABC
24	UNC	0.608	UNC	UNC	GCB
25	UNC	0.817	UNC	GCB	UNC
26	UNC	0.821	ABC	ABC	ABC
27	UNC	0.86	UNC	UNC	UNC
28	UNC	0.874	UNC	UNC	ABC
29	UNC	0.521	GCB	GCB	GCB
30	UNC	0.467	ABC	ABC	UNC
31	UNC	0.455	UNC	UNC	GCB
32	UNC	0.384	UNC	UNC	ABC
33	GCB	0.397	GCB	GCB	GCB
34	GCB	0.467	GCB	GCB	UNC
35	GCB	0.47	ABC	ABC	ABC
36	GCB	0.516	GCB	GCB	GCB
37	GCB	0.524	GCB	GCB	GCB
38	GCB	0.527	GCB	UNC	GCB
39	GCB	0.55	ABC	ABC	GCB
40	GCB	0.587	GCB	ABC	GCB
41	GCB	0.608	GCB	GCB	GCB
42	GCB	0.611	GCB	GCB	ABC
43	GCB	0.623	GCB	GCB	GCB

44	GCB	0.687	UNC	GCB	GCB
45	GCB	0.691	GCB	UNC	GCB
46	GCB	0.712	GCB	GCB	GCB
47	GCB	0.717	UNC	UNC	GCB
48	GCB	0.757	GCB	GCB	GCB
49	GCB	0.774	GCB	GCB	GCB
50	GCB	0.775	GCB	GCB	GCB
51	GCB	0.777	GCB	GCB	GCB
52	GCB	0.786	GCB	GCB	GCB
53	GCB	0.788	GCB	GCB	GCB
54	GCB	0.848	GCB	GCB	GCB
55	GCB	0.858	GCB	GCB	GCB
56	GCB	0.866	GCB	GCB	GCB
57	GCB	0.871	GCB	GCB	GCB
58	GCB	0.88	GCB	GCB	GCB
59	GCB	0.881	GCB	GCB	GCB
60	GCB	0.882	GCB	GCB	GCB
61	GCB	0.887	GCB	UNC	GCB
62	GCB	0.89	GCB	GCB	GCB
63	GCB	0.894	GCB	GCB	GCB
64	GCB	0.902	GCB	GCB	GCB
65	GCB	0.903	GCB	GCB	GCB
66	GCB	0.904	GCB	GCB	GCB
67	GCB	0.905	GCB	GCB	GCB
68	GCB	0.906	GCB	GCB	GCB

Table S4: Confusion Matrices for three COO classes across the four gene expression platforms. Concordant numbers of samples on the diagonal in bold, and discordant numbers in upper and lower halves.

HTG

	ABC	GCB	UC
ABC	**22**	2	4
GCB	0	**33**	3
UC	0	1	**3**

DASL

AFFY

	ABC	GCB	UC
ABC	**20**	2	3
GCB	1	**32**	1
UC	1	2	**6**

DASL

	ABC	GCB	UC
ABC	**23**	1	1
GCB	2	**31**	1
UC	3	4	**2**

HTG

RNAseq

	ABC	GCB	UC
ABC	**22**	3	3
GCB	0	**29**	2
UC	0	4	**5**

DASL

	ABC	GCB	UC
ABC	**25**	2	1
GCB	1	**28**	2
UC	2	6	**1**

HTG

	ABC	GCB	UC
ABC	**25**	2	1
GCB	0	**29**	2
UC	0	3	**6**

AFFY