Journal of Molecular Biomarkers & Diagnosis

Hersey et al., J Mol Biomark Diagn 2019, 10:2

Research Article Open Access

A Novel Immunohistochemistry Method for Predicting Activated B Cell-Like (ABC) Subtype of Diffuse Large B-Cell Lymphoma (DLBCL)

Sarah Hersey[1], **Yixin Wang**[1*], **Alberto Risueno Perez**[2] **and Fadi Towfic**[2]

[1]*Precision Medicine and Companion Diagnostics, Celgene Corporation, 86 Morris Avenue, Summit, New Jersey, United States*

[2]*Research Analytics, Celgene Corporation, 86 Morris Avenue, Summit, New Jersey, United States*

[*]**Corresponding author:** Wang Y, Precision Medicine and Companion Diagnostics, Celgene Corporation, 86 Morris Avenue, Summit, New Jersey, United States, Tel: +1 908-897-6618; E-mail: yixwang@celgene.com

Rec Date: February 22, 2019, **Acc Date:** March 22, 2019, **Pub Date:** March 29, 2019

Abstract

Background: The ROBUST trial (CC-5013-DLC-002), is a phase 3 randomized, double-blind, placebo controlled, multicenter study to compare the efficacy and safety of Lenalidomide (CC-5013) plus R-CHOP (rituximab plus cyclophosphamide, doxorubicin, vincristine and prednisone) chemotherapy (R^2-CHOP) versus placebo plus R-CHOP chemotherapy in subjects with previously untreated activated B-cell (ABC) type Diffuse Large B-cell Lymphoma (DLBCL). The most commonly utilized immunohistochemistry (IHC) algorithm for DLBCL subtyping, developed by Hans, et al. has been shown to have approximately an overall percent agreement of 80% with the gene expression profiling (GEP) classification of DLBCL into the germinal center B-cell–like (GCB) and non-GCB subtypes. The non-GCB subtype includes both the ABC subtype as well as the unclassified or indeterminate subtype which is neither ABC nor GCB as defined by gene expression profiles. New antibodies and algorithms specific to ABC tumors have been proposed with an aim to improve the performance of the IHC algorithm for specifically detecting the ABC subtype. This article describes one such new IHC algorithm which could be deployed for ABC subtype determination in clinic.

Methods: We analyzed 100 cases of newly diagnosed DLBCL with CD20, CD10, Bcl-6, MUM1, FOXP1, Bcl-2, Ki-67 and CD5 IHC assays using laboratory developed tests (LDT) and compared different combinations of the IHC assay results to the GEP classification. Statistical analyses were applied to evaluate the possible effect of inter-laboratory and inter-observer variations for the IHC assays. Instead of using a decision tree algorithm approach to determine GCB and ABC, a novel approach was taken by using a weighted composite score algorithm. A new IHC algorithm using CD10, Bcl-6, MUM1, and FOXP1 was derived to identify ABC versus non-ABC tumors that closely approximated the GEP classification in this training set. The algorithm was assessed independently and also in conjunction with the Hans IHC algorithm to enhance testing performance. A separate set of 100 independent newly diagnosed DLBCL cases were used to validate the algorithms using LDTs developed independently from two different laboratories using different antibodies, different instrument systems and a total of four pathologists. It should be noted, the LDTs selected for use in this study were not pre-evaluated for performance.

Results: Statistical analyses indicated that the IHC assays and the algorithms for subtyping of DLBCL were robust and reproducible within the range of inter-laboratory and inter-observer variations. For the validation data set, comparing the GEP classification results and the IHC results which were derived using two independently developed LDTs per IHC marker and four pathologists, the new IHC algorithm using CD10, Bcl-6, MUM1, and FOXP1 achieved 81-91% concordance in identifying ABC tumors and 79%-86% concordance in overall classification between the individual pathologists' calls and the GEP classification. When used in conjunction with the Hans algorithm, the IHC results achieved 89-97% concordance in identifying ABC tumors and 84-89% concordance in overall classification to the GEP results in the validation data set, simulating the predictive power of the GEP classification.

Conclusion: The use of the new IHC algorithm alone and in combination with the Hans algorithm can accurately predict ABC tumors of DLBCL and facilitate subtyping of DLBCL using standard pathology materials and routinely validated IHC assays.

Keywords: DLBCL subtyping; Immunohistochemistry; Gene expression profiles; Companion diagnostics

Introduction

Diffuse large B-cell lymphoma (DLBCL) is a heterogeneous group of B-cell lymphomas [1-5]. Gene expression profiling (GEP) studies have shown that DLBCL can be divided into sub-groups of germinal center B-cell – like (GCB), activated B-cell – like (ABC), and unclassified tumors [6-9]. The GCB and ABC subtypes have different pathogenetic mechanisms that may impact the outcomes of DLBCL patients on targeted therapies [10]. For example, Lenalidomide (CC-5013) is being explored in combination with R-CHOP (rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone) to form R^2CHOP in treating ABC subtype DLBCL. Therefore, the assignment into cell of origin (COO) groups of DLBCL is becoming increasingly important with the emergence of novel therapies that have selective biological activity in sub-groups [10].

Despite the robustness of GEP in subclassifying DLBCL, such as Lymph2Cx assay (NanoString, WA), substantial time, cost, technical expertise, and resources are required, making it currently impractical for all regions of the world to perform GEP analysis on every patient with DLBCL. Various IHC algorithms have been developed to predict COO [1,11-14]. These algorithms use different combinations of antibodies to DLBCL tumors proteins to obtain a desirable COO classification. The results of the algorithms developed by Hans et al. and Choi et al. have correlated well with the corresponding GEP results and demonstrated survival differences between the GCB and non-GCB DLBCL groups [1,11]. The IHC assays use standard, formalin-fixed, paraffin-embedded (FFPE) tumor tissues. For example, the Hans algorithm uses three antibodies: CD10, Bcl-6 and MUM1 to classify DLBCL into GCB and non-GCB (including ABC and unclassified) subtypes. It resulted in a concordance of about 80% when compared with the GEP classification [1,12]. The Choi algorithm uses GCET1, CD10, Bcl-6, MUM1, and FOXP1 to determine COO of DLBCL and derived a greater than 85% concordance in a single study, with a single site performing the staining on tissue microarray and three pathologists interpreting to the GEP classification [11]. However, some antibodies used by the Choi algorithm have not been fully validated and are not routinely adapted in clinical labs. The other existing algorithms appear to have more potential variability and may not as robust to the use of different LDTs developed using different reagents and inter-pathologist interpretation [11-14]. Variability from inter-laboratory and inter-observer performance have been reported [15-20] with these methods when different LDTs, pathologists and laboratories have been used.

This study performed IHC analyses on 8 commonly used DLBCL markers to determine if a computational algorithm can be derived to identify ABC tumors from non-ABC tumors that would be comparable to GEP results. A total of 200 cases of DLBCL (including core needle biopsies and excisional biopsies) were obtained (Avaden Bio, WA). Tumor cellularity of the samples ranged from 25-95% with a mean of 77% and a median of 80%. Among them, 174 are excision biopsies and 26 are core needle biopsies. One-hundred samples were randomly selected as a training set for IHC testing of 8 laboratory developed tests (LDT) (Neogenomics, CA). The results of the 8 independent LDT tests were each evaluated by 3 independent pathologists. The IHC assay results were then compared to the Nanostring's Lymph2Cx assay results. New IHC algorithms were proposed to identify ABC tumors either alone or in combination with the Hans algorithm. The algorithms were locked and the validation set was performed under protocol. The proposed algorithms were validated using an independent set of 100 DLBCL samples in 2 independent laboratories and by a total of 4 independent pathologists. The performance of the IHC assays were compared to the Nanostring's Lymph2Cx assay results in the validation set [21].

Materials and Methods

DLBCL samples: A total of 200 cases of newly diagnosed DLBCL (including core needle biopsies and excisional biopsies) were obtained for the study (Avaden Biosciences, WA). Among them, 100 samples were selected as a training set. The remaining 100 samples were used as a validation set. The workflow for this study is summarized in Figure 1. Patient characteristics of the DLBCL samples between the training set and the validation set are summarized in Table 1.

Immunohistochemistry (IHC): For each sample, multiple 5 μm thick slides were prepared and a single slide was stained with hematoxylin and eosin and reviewed by the pathologist to define the tumor area. The remaining slides were then subjected to standard immunohistochemical procedures including antigen retrieval, incubation with antibodies, detection and counter-staining per each laboratory's standard operating procedures for the respective LDT tests. For each of the training samples, the following IHC tests were performed: CD20, CD10, Bcl-6, MUM1, FOXP1, Bcl-2, Ki-67 and CD5 by Neogenomics (Neogenomics, CA). For the validation set, the following IHC tests were performed CD20, CD10, Bcl-6, MUM1 and FOXP1. For the validation set, two laboratories (Neogenomics, CA, PPD, NC) were chosen to challenge the robustness of the algorithms by using the laboratories' respective independently LDTs. The antibody clones used by each of the laboratories for the IHC assays are listed in Table 2 and are different for each marker. Blocking, amplification and 3′-diaminobenzidine detection kits were used according to manufacturer's instructions. The immunostaining was performed per testing laboratory's protocols on either Benchmark XT (Ventana), Benchmark Ultra (Ventana), or Bond III (Leica) instruments. Each of the IHC assays in the study were validated with proper positive and negative controls in the testing laboratories.

For each IHC antibody stain, percentage of positive stained tumor cells in 5% increments (or <1%) and staining intensity (0, 1+, 2+ and 3+) in each of the DLBCL samples were scored.

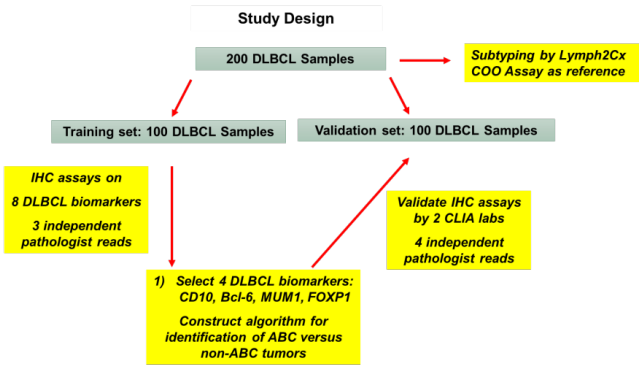

Figure 1: Workflow of the study. A total of 200 cases of DLBCL (including core needle biopsies and excisional biopsies) were analyzed by IHC and GEP for COO classification. One hundred samples were selected as a training set. The remaining 100 samples from were used as an independent validation set. For each of the training samples, the following IHC tests were performed: CD20, CD10, Bcl-6, MUM1, FOXP1, Bcl-2, Ki-67 and CD5. For the validation set, the following selected IHC tests were performed CD20, CD10, BCL6, MUM1 and FOXP1. For the validation set, two laboratories were chosen to challenge the robustness of the algorithms by using their respective independently laboratory developed tests. Each of the IHC assays in the study were validated with proper positive and negative controls in the testing labs. The scoring of the training set was provided by three independent pathologists. The scoring of the validation set was provided by two independent pathologists at each laboratory (n=4).

Characteristics	Training cohort (n=100)	Validation cohort (n=100)	P value*
Age-median (range)	67 (18-98)	68 (23-91)	-
Gender-n (%)			
Male	58 (58%)	58 (58%)	1
Female	42 (42%)	42 (42%)	
Tissue Type–n (%)			
Nodal	36 (36%)	27 (27%)	0.04
Extra-nodal	64 (64%)	73 (73%)	
GEP result–n (%)			
ABC	27 (27%)	37 (37%)	0.19
GCB	60 (60%)	50 (50%)	
Unclassified	12 (12%)	12 (12%)	
No test	1 (1%)	1 (1%)	

*P values were derived from Chi-squared test to assess the variables between the training and the validation sets.

Table 1: Patient and tumor characteristics of training and validations sets.

For the training set, scoring was estimated visually by three independent pathologists trained on Neogenomics' LDTs who were blinded to the GEP results to allow for inter-reader reproducibility to be assessed. The scoring of the validation set was the same as the training set and provided by two independent pathologists at each laboratory (n=4).

Lymph2Cx COO Assay: GEP was used as a reference method in the study and performed by using Lymph2Cx COO Assay developed and analytically validated by NanoString (8). Lymph2Cx COO Assay includes a set of 15 COO genes and 5 "housekeeping" genes that had low variability across DLBCL samples. A weighted average expression of the 15 COO genes was used to generate a predictive score that assigned the sample into one of the 3 classes: ABC, GCB or Unclassified. Experimentally, 5 of 5 μm sections were prepared from each of the FFPE tissue blocks. A RNA input of 500 ng per sample was used to carry out Lymph2Cx COO Assay (NanoString, WA).

Construction and validation of new IHC algorithm: Percentage of positively stained cells for each of the IHC assays of the samples of the training set were tabulated on a spreadsheet, alongside with the corresponding GEP calls from Lymph2Cx COO Assay. Four of the 8 IHC stain results including CD10, Bcl-6, MUM1 and FOXP1 demonstrated abilities to discriminate ABC from non-ABC tumors when compared to the GEP calls (Figure 2). CD10 and Bcl-6 showed higher expression in GCB tumors than in ABC tumors, while MUM1 and FOXP1 showed higher expression in ABC tumors than in GCB tumors. The other 4 markers didn't show significant differential expression between the 2 groups. We then tested and chose combinations of these four IHC stains that could achieve high sensitivity and high specificity for predicting the ABC subtype versus non-ABC subtypes.

Two algorithms involving 4 IHC stains: CD10, Bcl-6, MUM1 and FOXP1 were proposed to discriminate ABC versus non-ABC tumors of DLBCL from the analysis of the training set (Figure 3). One algorithm has a composite score based on sum of weighted percentage of positively stained cells of the 4 markers (Figure 3A); the other first applied the Hans algorithm (1) and then the proposed algorithm to the non-GCB samples identified by the Hans algorithm (Figure 3B). The cutoff for the Hans algorithm positivity is 30% of the tumor cells positively stained.

Antibody	Clone	Manufacturer	Catalogue#
CD20	L26	Leica	NCL-L-CD20-L26
CD10	56C6	Leica	NCL-L-CD10-270
BCL6	LN22	Leica	NCL-L-Bcl-6-564
MUM1	MUM1p	Dako	M7259
FOXP1	JC12	Novus	NB100-65125
BCL2	124	Dako	M0887
Ki-67	MIB-1	BioCare	PM362DASS
CD5	4C7	Leica	NCL-L-CD5-4C7

Table 2A: Antibodies and instruments for IHC assays, Training set (Neogenomics).

The 4 IHC assays, the CD20 IHC assay as a diagnostic confirmation, and Lymph2Cx COO Assay were then performed on the 100 independent DLBCL samples of the validation set. The pre-determined algorithms including the cutoffs were computed based on the IHC stains of the samples of the validation set to determine their abilities to predict ABC and non-ABC tumors. The predicted IHC calls were compared to the Lymph2Cx COO Assay calls to evaluate the performance of the algorithms.

	Neogenomics Lab			PPD Lab		
Antibody	**Clone**	**Manufacturer**	**Catalogue#**	**Clone**	**Manufacturer**	**Catalogue#**
CD20	L26	Leica	NCL-L-CD20-L26	L26	Ventana	760-2531
CD10	56C6	Leica	NCL-L-CD10-270	SP67	Ventana	790-4506
BCL6	LN22	Leica	NCL-L-Bcl-6-564	GI91E	Ventana	760-4241
FOXP1	JC12	Novus	NB100-65125	SP133	Ventana	760-4611
MUM1	MUM1p	Dako	M7259	MUM1p	Dako	M7259

Table 2B: Antibodies and instruments for IHC assays, Validation set (Neogenomics and PPD).

Statistical analysis: The $\chi 2$ test was used to compare patient characteristics between the training and the validation sets (Table 1). The COO calls made by the IHC algorithms for the validation set were compared to the subtypes as classified by Lymph2Cx COO Assay to determine accuracy of the IHC algorithms. Each pathologist's results were evaluated individually and then together. Overall percentage agreements (OPA) between the IHC calls from each pathologist to the Lymph2Cx COO Assay calls (ABCs and non-ABCs, excluding unclassified subtype) were calculated. Acceptance criteria was set at >80% concordance (ABC and non-ABC) between IHC and GEP for the validation of each of the algorithms. The protocol was written such that a >80% concordance between IHC and GEP for the validation set in both testing laboratories validates the algorithms across different laboratories and different IHC assays and platforms.

	GEP ABC	GEP GCB
IHC ABC	a	b
IHC non-ABC	c	d
Overall Percentage Agreement (OPA) = (a+d)/(a+b+c+d) x 100%		

Results

Patient characteristics of the training and validation sets

The training set included 58% males and 42% females with a median age of 67 years (range, 18-98 years). The validation set comprised 58% males and 42% females with a median age of 68 years (range, 23-91 years). The characteristics of the patients were not significantly different (Table 1). For the training set, 27 cases were classified into ABCs, 60 cases into GCBs and 12 cases into unclassified by GEP; and for the validation set, 37 cases were classified into ABCs, 50 cases into GCBs and 12 cases into unclassified by GEP. The classification of the patient cases with GEP-defined ABC, GCB and unclassified subtypes were also not significantly different between the training and the validation set (Table 1).

Reproducibility of inter-reader evaluation and inter-laboratory testing on IHC assays

One-hundred cases of DLBCL in the training set were tested in a single lab with the 8 IHC assays (Table 2), and the IHC stains were evaluated by 3 independent pathologists. The inter-pathologist variability was assessed by pair-wise comparison on assay positivity between the individual IHC assays, as well as between the Hans algorithm calls. Using 30% positively stained cells as the cutoff, the average pair-wise percentage agreement for the 4 selected IHC assay ranged from 83% to 93% (Table 3). The average concordance of the pathologists' calls for the Hans algorithm was 90% (Table 3). One-hundred cases of DLBCL in the validation set were tested in 2 independent labs for the 4 selected IHC markers. The laboratories used different antibody clones and instruments for some of the IHC assays (Table 2). In addition, The IHC stains of the validation set were evaluated by 4 independent pathologists (2 pathologists per laboratory). The average pair-wise percentage agreement between the pathologists of the individual IHC assays for the validation set ranged from 76% to 93% (Table 3). The average pair-wise concordance of the pathologists' calls for the Hans algorithm and the new IHC algorithm were 80% and 86%, respectively (Table 3).

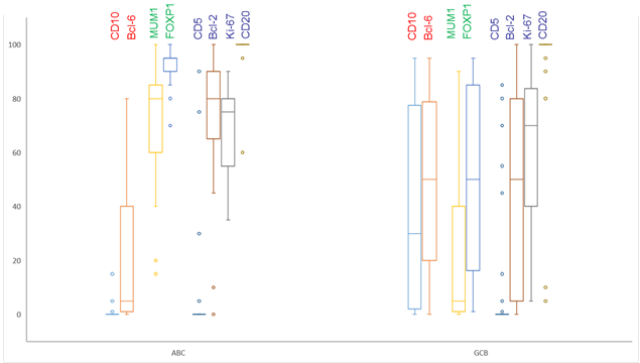

Figure 2: Percentage of positively stained cells in ABC and GCB tumors for individual IHC markers. The box plot illustrates the percentage of positively stain cells for each of the 8 IHC markers between the GEP defined ABC and GCB tumors. CD10 and Bcl-6 showed higher expression in GCB tumors than in ABC tumors, while MUM1 and FOXP1 showed higher expression in ABC tumors than in GCB tumors. The other 4 markers didn't show significant differential expression between the 2 groups.

A new IHC algorithm for classification of DLBCL

In the training set of 100 DLBCL cases, the use of GEP Lymph2Cx COO Assay identified 27 ABCs, 60 GCBs and 12 unclassified calls (1 no call). The Hans algorithm classified 47 GCBs and 53 non-GCBs. Seventy-three percent (73%) of GEP-defined GCB cases were classified

as GCBs and 100% GEP-defined ABC cases were classified as non-GCBs by the Hans' algorithm. The overall percentage agreement (71/87) between the Hans algorithm and the GEP result is 81%.

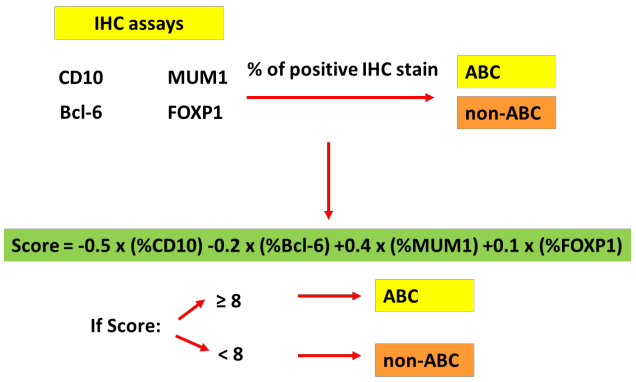

Figure 3A: IHC algorithms for ABC versus non-ABC classification of DLBCL. Two algorithms involving 4 IHC markers: CD10, Bcl-6, MUM1 and FOXP1 were proposed to discriminate ABC *vs.* non-ABC tumors of DLBCL from the analysis of the training set. One algorithm has a composite score based on sum of weighted percentage of positively stained cells of the 4 IHC markers.

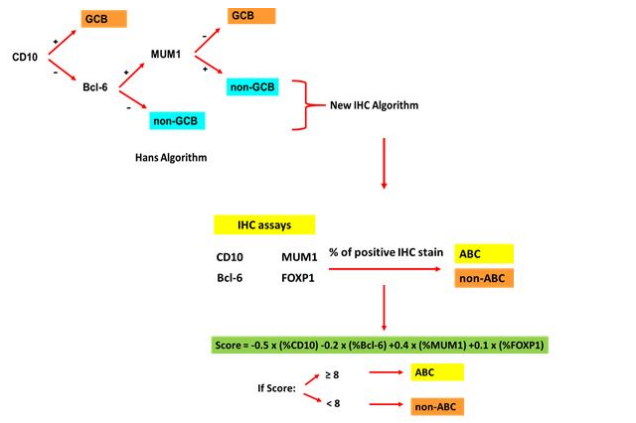

Figure 3B: IHC algorithms for ABC versus non-ABC classification of DLBCL. Two algorithms involving 4 IHC markers: CD10, Bcl-6, MUM1 and FOXP1 were proposed to discriminate ABC *vs.* non-ABC tumors of DLBCL from the analysis of the training set. The other algorithm applied Hans algorithm first and then the newly proposed algorithm to the non-GCB samples of the Hans result. The cutoff for Hans algorithm positivity is 30% of the tumor cells showing expression.

The box plot illustrates the percentage of positively stained cells for each of the 8 IHC markers between the GEP defined ABC and GCB tumors (Figure 2). CD10 and Bcl-6 showed higher expression in GCB tumors than in ABC tumors, while MUM1 and FOXP1 showed higher expression in ABC tumors than in GCB tumors. The other 4 markers didn't show significant differential expression between the 2 groups. Various combinations of the 4 selected IHC assays were tested to separate the cases into ABC and non-ABC subtypes according to the GEP results of the training set. Given the variation on evaluation of IHC stains, we tested the weighting of each of the 4 selected markers in the combinations. Also, inclusion of 2 markers for the ABC group (MUM1 and FOX1) and 2 markers for the GCB group (CD10 and Bcl-6) built redundancy and complementation for detection of the specific sub-group of DLBCL tumors. As the result, a new IHC algorithm calculated on 4 IHC stains of CD10, Bcl-6, MUMP1 and FOXP1 was chosen to achieve the highest possible agreement between the IHC calls and the GEP calls for the ABC cases (~90%) (Figure 1). Although the weights of the individual IHC markers in the new algorithm was developed based on the first pathologist's evaluation of the IHC stains of the training set, the remaining data from the other 2 pathologists were also examined to verify the algorithm. The classification results are highly concordant (For ABC call: average 98%, range: 96-100%; For overall call: average: 88%, range: 86-89%.) among the 3 pathologists' evaluation. However, because all 3 pathologists' evaluation was based on the IHC results of the same training set, this analysis was not considered as validation of the algorithm. Further testing with a set of independent samples was required for validation.

Assay/ algorithm	Average concordance from pair-wise comparisons (Training Set, 3 pathologists)	Average concordance from pair-wise comparisons (Validation Set, 2labs, 4 pathologists)
CD10	93%	93%
BCL6	83%	76%
MUM1	84%	85%
FOXP1	92%	92%
Hans Algorithm	90%	86%
New IHC Algorithm	--*	86%
*Not included for training set		

Table 3: Concordance of pathologist evaluation of IHC assays.

Validation of the new IHC algorithm for classification of DLBCL

The performance of the new IHC algorithm, either alone or in combination with the Hans algorithm, was evaluated in the 100 independent cases of the validation set. Each of the cases were assigned to GEP-defined ABC, GBC or unclassified subtype according to the Lymph2Cx COO Assay results. The new algorithms involve 4 IHC stains: CD10, Bcl-6, MUM1 and FOXP1 and derives a composite score based on sum of weighted percentage of positively stained cells of the 4 markers to discriminate ABC versus non-ABC tumors (Figure 3A). The concordance of the IHC calls by the new IHC algorithm to the GEP calls are shown in Table 4.

The percentage agreement of the ABC calls ranged from 81% to 91% among the 4 pathologists' evaluation. The overall percentage agreement ranged from 79% to 86% among the 4 pathologists' evaluation (Table 4). Further, the new IHC algorithm was applied in conjunction with the Hans algorithm (Figure 3B). The Hans algorithm was applied first to call out the GCB samples and then the proposed algorithm was applied to the non-GCB samples identified by the Hans algorithm in order to identify the ABC samples. The concordances of the IHC calls by the Hans and the new IHC algorithms to the GEP

calls are shown in Table 5. The percentage agreement of the ABC calls ranged from 89% to 97% among the 4 pathologists' evaluation using 2 independently derived sets of LDT tests (2 pathologists per set of LDT tests). The overall percentage agreement ranged from 84% to 89% among the 4 pathologists' evaluation (Table 5).

Pathologist 1: 98 samples: 37 ABC, 49 GCB, 12 unclassified by GEP. 2 no call.		
	GEP	
New IHC Algorithm	ABC (n=37)	GCB (n=49)
ABC (%)	30/37 (81%)	-
Non-ABC (%)	-	41/49 (84%)
Overall Accuracy (%)	71/86 (83%)	
Pathologist 2: samples: 36 ABC, 49 GCB, 12 unclassified by GEP. 3 no call.		
	GEP	
New IHC Algorithm	ABC (n=36)	GCB (n=49)
ABC (%)	29/36 (81%)	-
Non-ABC (%)	-	43/49 (88%)
Overall Accuracy (%)	72/85 (85%)	
Pathologist 3: 94 samples: 35 ABC, 48 GCB, 11 unclassified by GEP. 6 no call.		
	GEP	
New IHC Algorithm	ABC (n=35)	GCB (n=48)
ABC (%)	32/35 (91%)	-
Non-ABC (%)	-	39/48 (81%)
Overall Accuracy (%)	71/83 (86%)	
Pathologist 4: 92 samples: 32 ABC, 48 GCB, 12 unclassified by GEP. 8 no call.		
	GEP	
New IHC Algorithm	ABC (n=32)	GCB (n=48)
ABC (%)	26/32 (81%)	-
Non-ABC (%)	-	37/48 (77%)
Overall Accuracy (%)	63/80 (79%)	

Table 4: Comparison of classification results between new IHC algorithm versus GEP in validation set.

The results suggest that not only the performance of the new IHC algorithm is validated in the independent set of samples but also the combination of the Hans algorithm and the new IHC algorithm can accurately predict both ABC and GCB cases, simulating the performance of the GEP classification. The comparison between the new IHC algorithm and GEP is shown in Figure 4. The overall agreements between the two methods are excellent on both clear-cut ABC and GCB calls (Figure 4A). Most of the discordant calls between the two methods appeared in the middle range (e.g., 1,500 – 2,500 of the LST score) of the Lymph2Cx COO Assay score, where the new IHC algorithm tends to have more ABC calls (Figure 4A). In addition, 5 cases were identified that all 4 pathologists have the same classification but different from the GEP call (Figures 4A and 4B), suggesting a possibility that neither method is perfect in performance of COO classification. Experiments by using patient cohorts with outcome data is required to further evaluate the performance of these methods.

Pathologist 1: 98 samples: 37 ABC, 49 GCB, 12 unclassified by GEP. 2 no call.		
	GEP	
Hans+New IHC Algorithm	ABC (n=37)	GCB (n=49)
ABC (%)	34/37 (92%)	-
Non-ABC (%)	-	42/49 (86%)
Overall Accuracy (%)	76/86 (88%)	
Pathologist 2: samples: 36 ABC, 49 GCB, 12 unclassified by GEP. 3 no call.		
	GEP	
Hans+New IHC Algorithm	ABC (n=36)	GCB (n=49)
ABC (%)	32/36 (89%)	-
Non-ABC (%)	-	44/49 (90%)
Overall Accuracy (%)	76/85 (89%)	
Pathologist 3: 94 samples: 35 ABC, 48 GCB, 11 unclassified by GEP. 6 no call.		
	GEP	
Hans+New IHC Algorithm	ABC (n=35)	GCB (n=48)
ABC (%)	34/35 (97%)	-
Non-ABC (%)	-	39/48 (81%)
Overall Accuracy (%)	73/83 (88%)	
Pathologist 4: 92 samples: 32 ABC, 48 GCB, 12 unclassified by GEP. 8 no call.		
	GEP	
Hans+New IHC Algorithm	ABC (n=32)	GCB (n=48)
ABC (%)	30/32 (94%)	-
Non-ABC (%)	-	37/48 (77%)
Overall Accuracy (%)	67/80 (84%)	

Table 5: Comparison of classification results between Hans algorithm plus new IHC algorithm versus GEP in validation set.

Discussion

In the study, no single IHC marker has been sufficient in classifying DLBCL into subtypes. For this reason, combinations of antibodies and algorithms have been developed based on concordance with the classification results by GEP. The Hans algorithm is highly useful to determine GCB subtype of DLBCL. Apart from the Hans' algorithm, several other IHC stain algorithms have been proposed to classify subtypes of DLBCL [11-14].

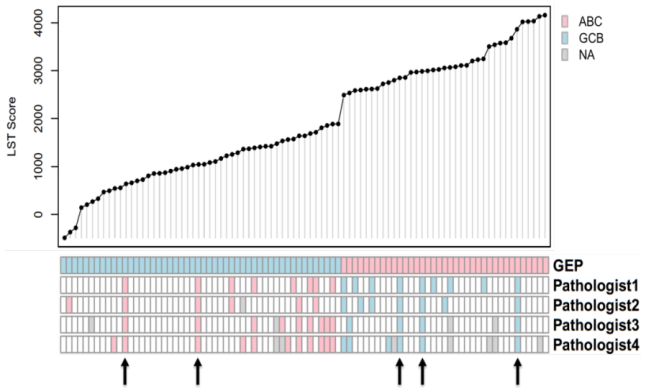

Figure 4A: Performance of the IHC algorithm compared to GEP in the validation set. The graph shows LST score from the Lymph2Cx COO Assay score (NanoString, WA) in ascending order of the 100 DLBCL samples of the validation set. The heatmap below the graph shows the ABC (pink) and GCB (blue) calls based on the GEP Lymph2Cx COO Assay. The pathologists' calls based on the new IHC algorithm in the heatmap only highlights the discordant results from the GEP calls. The overall percentage agreement ranged from 84% to 88% among the 4 pathologists' evaluation. The arrows below indicate 5 cases that all 4 pathologists have the same classification call but different from the GEP call.

Sample ID	Lab	Pathologist	CD10+ %	Bcl-6+ %	MUM1+ %	FOXP1+%	IHC call	LPS Score	GEP Call
DLBCL620	Neo	1	1	10	40	90	ABC	637	GCB
DLBCL620	Neo	2	1	5	25	85	ABC	637	GCB
DLBCL620	PPD	3	0	80	50	100	ABC	637	GCB
DLBCL620	PPD	4	0	100	80	100	ABC	637	GCB
DLBCL624	Neo	1	5	70	70	95	ABC	1042	GCB
DLBCL624	Neo	2	5	70	70	95	ABC	1042	GCB
DLBCL624	PPD	3	0	100	90	100	ABC	1042	GCB
DLBCL624	PPD	4	20	100	100	100	ABC	1042	GCB
DLBCL600	Neo	1	90	1	30	90	non-ABC	2850	ABC
DLBCL600	Neo	2	60	1	50	85	non-ABC	2850	ABC
DLBCL600	PPD	3	60	40	90	100	non-ABC	2850	ABC
DLBCL600	PPD	4	50	50	80	100	non-ABC	2850	ABC
DLBCL625	Neo	1	5	65	5	90	non-ABC	2985	ABC
DLBCL625	Neo	2	5	65	1	70	non-ABC	2985	ABC
DLBCL625	PPD	3	0	80	2	60	non-ABC	2985	ABC
DLBCL625	PPD	4	0	100	20	70	non-ABC	2985	ABC
DLBCL595	Neo	1	0	85	10	100	non-ABC	3864	ABC
DLBCL595	Neo	2	1	70	5	95	non-ABC	3864	ABC
DLBCL595	PPD	3	0	100	40	100	non-ABC	3864	ABC
DLBCL595	PPD	4	0	100	40	100	non-ABC	3864	ABC

Figure 4B: Performance of the IHC algorithm compared to GEP in the validation set. The IHC data of individual markers and the GEP data for the 5 cases that all 4 pathologists have the same classification call but different from the GEP call.

Although IHC is routinely performed on archival FFPE tumor tissues in most pathology laboratories, making it practical for widespread clinical use, little data has been reported in identification of ABC tumors versus non-ABC tumors of DLBCL as compared to GEP classification. The existing algorithms appear to have more potential variability and may not as robust to the use of different LDTs developed using different reagents and inter-pathologist interpretation [11-14]. Our proposed IHC algorithm demonstrates a potential ability to accurately predict the ABC subtype of DLBCL tumors with only the addition of one IHC stain to the Hans algorithm. Further, the algorithm was locked prior to conducting the validation, which was conducted under a formal protocol. In conjunction with the Hans algorithm which accurately predicts GCB tumors, the use of the new IHC algorithm can facilitate subtyping of DLBCL using standard pathology materials and routinely validated IHC assays. The results of this study should allow laboratories with limited access to GEP to more accurately predict ABC tumors using common IHC methods.

The new algorithm places less weight on some GCB-specific markers and more weight to the ABC-specific markers such as FOXP1 as compared to other IHC methods. Because of its robust IHC staining, FOXP1 not only helps balance the weights between the 2 GCB and the 2 ABC markers but also off-set some of the inherent variability in readers' evaluation of the IHC stains. FOXP1 gene is on chromosome 3p14.1 and encodes a member of the FOX family of transcription factors [22]. FOXP1 mRNA has been shown to be highly expressed in the ABC subtype of DLBCL and other studies showed that high expression of FOXP1 protein is associated with an inferior survival in DLBCL patients [22]. In this study, we found that the addition of FOXP1 to the algorithm achieved a higher specificity for identification of the ABC tumors.

Although GEP defines about 10-15% of DLBCL cases into unclassified subtype, we only used GEP-defined ABC and GCB cases to derive the new algorithm in this study. Previous studies have shown that the GEP-defined unclassified DLBCL subtype behaved similarly to the ABC subtype in patient prognosis [5,6,15]. Whether or not the IHC algorithm can separate the unclassified cases into ABC and GCB subtypes or define the unclassified cased into a distinct entity with a clear prognostic implication requires further investigations.

Conclusion

In conclusion, we have identified and validated a new IHC algorithm for identification of ABC tumors of DLBCL. In conjunction with the Hans algorithm, it closely replicates the GEP based COO classification of DLBCL into the ABC and the non-ABC subtypes. It represents an improvement in using routine clinical IHC assays in DLBCL subtyping. The new method could facilitate future research and clinical development in DLBCL by using archival FFPE tumor materials and evaluation of patients with DLBCL for novel experimental therapies.

Acknowledgements

The authors would like to thank Anita Gandhi, Patrick Hagner and Greg Wessels for their critical reading of the manuscript.

References

1. Hans CP, Weisenburger DD, Greiner TC, Gascoyne RD, Delabie J, et al. (2004) Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. Blood 103: 275-282.
2. Schaffer M, Chaturvedi S, Alvarez JD, Frans S, Aquino R (2018) Comparison of Immunohistochemistry Assay Results with Gene Expression Profiling Methods for Diffuse Large B-Cell Lymphoma Subtype Identification in Matched Patient Samples. J Mol Biomark Diagn 9: 2.
3. Gifford GK, Gill AJ, Stevenson WS (2016) Molecular subtyping of diffuse large B-cell lymphoma: update on biology, diagnosis and emerging platforms for practising pathologists. Pathology 48: 5-16.

4. O'Malley DP, Auerbach A, Weiss LM (2015) Practical applications in immunohistochemistry: evaluation of diffuse large B-cell lymphoma and related large B-cell lymphomas. Arch Pathol Lab Med 139: 1094-1107.
5. Pon JR, Marra MA (2016) Clinical impact of molecular features in diffuse large B-cell lymphoma and follicular lymphoma. Blood 127: 181-186.
6. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, et al. (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 403: 503.
7. Wright G, Tan B, Rosenwald A, Hurt EH, Wiestner A, et al. (2003) A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. Proc Natl Acad Sci 100: 9991-9996.
8. Scott DW, Wright GW, Williams PM, Lih CJ, Walsh W, et al. (2014) Determining cell-of-origin subtypes of diffuse large B-cell lymphoma using gene expression in formalin-fixed paraffin-embedded tissue. Blood 123: 1214-1217.
9. Mareschal S, Ruminy P, Bagacean C, Marchand V, Cornic M, et al. (2015) Accurate Classification of Germinal Center B-Cell–Like/Activated B-Cell–Like Diffuse Large B-Cell Lymphoma Using a Simple and Rapid Reverse Transcriptase – Multiplex Ligation-Dependent Probe Amplification Assay: A CALYM Study. J Mol Diagn 17: 273-283.
10. Czuczman MS, Trněný M, Davies A, Rule S, Linton KM, et al. (2017) A phase 2/3 multicenter, randomized, open-label study to compare the efficacy and safety of lenalidomide versus investigator's choice in patients with relapsed or refractory diffuse large B-cell lymphoma. Clin Cancer Res 23: 4127-4137.
11. Choi WW, Weisenburger DD, Greiner TC, Piris MA, Banham AH, et al. (2009) A new immunostain algorithm classifies diffuse large B-cell lymphoma into molecular subtypes with high accuracy. Clin Cancer Res 15: 5494-5502.
12. Meyer PN, Fu K, Greiner TC, Smith LM, Delabie J, et al. (2011) Immunohistochemical methods for predicting cell of origin and survival in patients with diffuse large B-cell lymphoma treated with rituximab. J Clin Oncol 29: 200-207.
13. Natkunam Y, Farinha P, Hsi ED, Hans CP, Tibshirani R, et al. (2008) LMO2 protein expression predicts survival in patients with diffuse large B-cell lymphoma treated with anthracycline-based chemotherapy with and without rituximab. J Clin Oncol 26: 447-454.
14. Nyman H, Jerkeman M, Karjalainen-Lindsberg ML, Banham AH, Leppä S (2009) Prognostic impact of activated B-cell focused classification in diffuse large B-cell lymphoma patients treated with R-CHOP. Mod Pathol 22: 1094-1101.
15. De Jong D, Xie W, Rosenwald A, Chhanabhai M, Gaulard P, et al. (2007) Retracted: Immunohistochemical prognostic markers in diffuse large B-cell lymphoma: validation of tissue microarray as a prerequisite for broad clinical applications (a study from the Lunenburg Lymphoma Biomarker Consortium). J Clin Oncol 25: 805-812.
16. Batlle-López A, de Villambrosía SG, Francisco M, Malatxeberria S, Sáez A, et al. (2016) Stratifying diffuse large B-cell lymphoma patients treated with chemoimmunotherapy: GCB/non-GCB by immunohistochemistry is still a robust and feasible marker. Oncotarget 7: 18036.
17. Reber R, Banz Y, Garamvölgyi E, Perren A, Novak U (2013) Determination of the molecular subtypes of diffuse large B-cell lymphomas using immunohistochemistry: a case series from the Inselspital, Bern, and a critical appraisal of this determination in Switzerland. Swiss Med Wkly 143: 13748.
18. Bellas C, García D, Vicente Y, Kilany L, Abraira V, et al. (2014) Immunohistochemical and molecular characteristics with prognostic significance in diffuse large B-cell lymphoma. PloS one 9: e98169.
19. Coutinho R, Clear AJ, Owen A, Wilson A, Matthews J, et al. (2013) Poor concordance among nine immunohistochemistry classifiers of cell-of-origin for diffuse large B-cell lymphoma: implications for therapeutic strategies. Clin Cancer Res 19: 6686-6695.
20. Muris JJ, Meijer CJ, Vos W, van Krieken JH, Jiwa NM, et al. (2006) Immunohistochemical profiling based on Bcl - 2, CD10 and MUM1 expression improves risk stratification in patients with primary nodal diffuse large B cell lymphoma. J Pathol 208: 714-723.
21. Montes-Moreno S, Roncador G, Maestre L, Martinez N, Sanchez-Verde L, et al. (2008) Gcet1 (centerin), a highly restricted marker for a subset of germinal center-derived lymphomas. Blood 111: 351-358.
22. Banham AH, Beasley N, Campo E, Fernandez PL, Fidler C, et al. (2001) The FOXP1 winged helix transcription factor is a novel candidate tumor suppressor gene on chromosome 3p. Cancer Res 61: 8820-8829.