Determining cell-of-origin subtypes of diffuse large B-cell lymphoma using gene expression in formalin-fixed paraffin embedded tissue

David W. Scott1*, George W. Wright2*, P. Mickey Williams3, Chih-Jian Lih3, William Walsh3, Elaine S. Jaffe4, Andreas Rosenwald5, Elias Campo6, Wing C. Chan7, Joseph M. Connors1, Erlend B. Smeland8, Anja Mottok1, Rita M. Braziel9, German Ott10, Jan Delabie11, Raymond R. Tubbs12, James R. Cook13, Dennis D Weisenburger14, Timothy C. Greiner7, Betty J. Glinsmann-Gibson15, Kai Fu7, Louis M. Staudt16#, Randy D. Gascoyne1,17#, Lisa M. Rimsza15#

1Centre for Lymphoid Cancer, British Columbia Cancer Agency, Vancouver, BC, Canada;  
2Biometric Research Branch, Division of Cancer Treatment and Diagnosis, National Cancer Institute, Bethesda, MD;  
3Molecular Characterization Laboratory, Leidos Medical Research Inc., Frederick National Laboratory for Cancer Research, Frederick, MD;  
4Laboratory of Pathology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD;  
5Institute of Pathology, University of Wuerzburg, Wuerzburg, Germany;  
6Department of Pathology, Hospital Clinic, Barcelona, Spain;  
7Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, NE;  
8Department of Immunology and Centre for Cancer Biomedicine, University of Oslo and Oslo University Hospital, Oslo, Norway;  
9Department of Pathology, Oregon Health and Sciences University, Portland, OR;  
10Department of Clinical Pathology, Robert-Bosch-Krankenhaus and Dr. Margarete Fischer-Bosch Institute of Clinical Pharmacology, Stuttgart, Germany;  
11Oslo University Hospital, Oslo, Norway;  
12Clinical Pathology, Cleveland Clinic Foundation, Cleveland, OH;  
13Department of Molecular Pathology and Laboratory Medicine, Cleveland Clinic, Cleveland, OH;  
14Dept. of Pathology, City of Hope, Duarte, CA;  
15Department of Pathology, University of Arizona, Tucson, AZ;  
16Metabolism Branch, Center for Cancer Research, National Cancer Institute, National Health Institutes, Bethesda, MD;  
17Department of Pathology, British Columbia Cancer Agency, Vancouver, BC, Canada

* These authors contributed equally to the work  
# These authors contributed equally to the work

Corresponding Author:  
Lisa M. Rimsza, M.D.  
1501 N. Campbell Avenue  
P.O. Box 245043  
Tucson, Arizona 85724-5043  
Telephone: 520-626-8396  
Fax: 520-626-6081  
Email: lrimsza@email.arizona.edu

Running Title: DLBCL cell-of-origin by gene expression in FFPET
Keypoint:

A 20-gene gene expression-based assay accurately and robustly assigns COO subtypes of DLBCL using formalin-fixed paraffin-embedded tissue.

Abstract:

The assignment of diffuse large B-cell lymphoma into cell-of-origin (COO) groups is becoming increasingly important with the emergence of novel therapies that have selective biological activity in germinal center B-cell-like (GCB) or activated B-cell-like (ABC) groups. The LLMPP’s Lymph2Cx assay is a parsimonious digital gene-expression (NanoString) based test for COO assignment in formalin-fixed paraffin-embedded tissue (FFPET) routinely produced in standard diagnostic processes. The 20-gene assay was trained using 51 FFPET biopsies; the locked assay was then validated using an independent cohort of 68 FFPET biopsies. Comparisons were made with COO assignment using the original COO model on matched frozen tissue. In the validation cohort the assay was accurate, with only one case with definitive COO being incorrectly assigned, and robust, with >95% concordance of COO assignment between 2 independent laboratories. These qualities, along with the rapid turn-around-time, make Lymph2Cx attractive for implementation in clinical trials and, ultimately, patient management.
Introduction:

Diffuse large B-cell lymphoma (DLBCL) is a heterogeneous group of cancers classified together on the basis of morphology, immunophenotype, genetic alterations and clinical behavior\(^1\). The distinction of DLBCL into cell-of-origin (COO) categories based on patterns of gene expression reminiscent of germinal center B-cell (the GCB group) and activated B-cell (the ABC group), with a small number of “unclassified” cases, as defined and characterized by the Lymphoma & Leukemia Molecular Profiling Project (LLMPP)\(^2,3\), has profound biological\(^4\), prognostic\(^5,6\) and potential therapeutic implications\(^7-9\). New therapeutic agents with selective activity in ABC and GCB DLBCL are under development. The original methods used to define these entities performed gene expression profiling (GEP) using microarrays on RNA derived from frozen tissue (FT). Subsequently, in an attempt to determine COO in standard practice using commonly available formalin-fixed paraffin-embedded tissue (FFPET) we and others used less precise\(^10\) but relatively inexpensive binary immunohistochemical (IHC) methods\(^11-13\). Recently, the feasibility of quantitating gene expression in FFPET in lymphoma has been demonstrated\(^14-18\). We sought to create a robust, highly accurate, validated molecular assay for COO distinction using GEP techniques applicable to FFPET\(^19\).

Methods:

Studies were performed on FFPET biopsies of *de novo* DLBCL cases that had been classified using the original GEP methods and published algorithm (the
“gold standard” method). Each case was centrally reviewed by a quorum of at least four LLMPP Pathologists and had GEP performed on matched FT biopsies using Affymetrix U133 plus 2.0 microarrays. Data are available at www.ncbi.nlm.nih.gov/geo/query/acc.cgi [accession number GSE53786]. The training cohort consisted of 51 cases (20 GCB, 19 ABC and 12 unclassified). The independent validation cohort of 68 cases, drawn from the validation cohort in Lenz et al., had proportions of GCB (28 cases), ABC (30 cases) and unclassified (10 cases) typically observed in DLBCL populations. Patient characteristics of the cohorts are shown in Table S1.

Tumor made up ≥60% of the surface area of the blocks. 10μm scrolls of FFPET were cut, to a surface area of 1cm², and tested in parallel at two independent laboratories – the Molecular Characterization Laboratory (MoCha), Frederick National Laboratory for Cancer Research in Frederick, Maryland and the Centre for Lymphoid Cancer (CLC), BC Cancer Agency in Vancouver, British Columbia. Nucleic acids were extracted using the Qiagen AllPrep FFPET kit and digital GEP was performed on 200ng RNA using NanoString technology (Seattle, WA). Details of study design are presented in the Supplement.

To assign COO by IHC, tissue microarrays were made using 0.6mm duplicate cores from 60/68 validation cohort cases, and stained for CD10, BCL6, IRF4/MUM1, FOXP1, GCET1 and LMO2. Two expert hematopathologists independently assessed the proportion of tumor cells stained, with consensus on discordant cases reached with a third hematopathologist. For the validation
studies, all individuals producing and analyzing the GEP and IHC data were blinded to the “gold standard” COO assignment.

Patients in the validation cohort received CHOP-type chemotherapy plus rituximab (R-CHOP (n=62); R-CNOP (n=3); R-CHEOP (n=1); R-ESHAP (n=1); R-CHOP/HSCT (n=1)). Progression free survival and overall survival, defined as per International Working Group response criteria\textsuperscript{20}, were estimated using the Kaplan-Meier method. Comparisons between groups were performed using the log-rank test.

This study was approved by the Institutional Review Boards of the participating centers in accordance with the Declaration of Helsinki.

**Results and Discussion:**

All FFPET biopsies yielded sufficient RNA at both laboratories for the NanoString technology analyses. A pilot study, using the training cohort, determined the FFPET gene expression by NanoString technology of 93 genes, identified in Lenz et al\textsuperscript{6} to differentiate ABC and GCB DLBCL subtypes. Fifteen genes, along with 5 “housekeeping” genes, were selected based on their ability to contribute to the accurate replication of the COO assignment model of Lenz et al\textsuperscript{6}. NanoString technology was then used to determine the expression of these 20 genes in FFPET-derived RNA from the training cohort, allowing a model to be built. As NanoString probes may vary in their hybridization efficiency lot-to-lot, a synthetic oligonucleotide reference was run alongside the patient samples, with the gene expression of the samples adjusted for the results in the reference. This
allows the model to be portable to new codeset lots and also reduces other sources of assay variability. Details of the model’s performance in the training cohort are presented in the Supplement. The locked model, including gene coefficients, thresholds and quality criteria, was then applied to the independent validation cohort. This assay has been named the Lymph2Cx.

Ninety-nine percent of the independent validation cohort, with FFPET blocks ranging from 5 to 12 years of age, yielded gene expression of adequate quality (67/68 at both sites, with different single cases failing at each center). When considering the 58 cases designated as ABC or GCB by the “gold standard” method, the Lymph2Cx assay incorrectly assigned one case – an ABC by FT GEP assigned to GCB (Figure 1A). At 2%, this favorably compares with the 9%, 6% and 17% rates of misassignment by the “Hans”, “Tally” and “Choi” IHC-based algorithms, respectively (Figure 1A)\textsuperscript{11-13}. Of note, the accuracy of the IHC-based algorithms observed here is consistent with the original descriptions\textsuperscript{11-13} but superior to that seen in a recent report that also used FT-based COO designation as the comparator\textsuperscript{10}. In replication of the original COO model and distinct from the binary IHC models, the Lymph2Cx assay recognizes a group of unclassified cases, where confident assignment cannot be made to either the ABC or GCB subtype. Among the 58 cases, three (5%) at the MoCha site and four (7%) at the CLC site were designated as unclassified. Full assignment data, including that of the unclassified cases by the “gold standard method”, are shown in Table S2.
It has become increasingly apparent that pre-analytical, analytical and inter-observer variability contribute to poor reproducibility of COO assignment by IHC methods\textsuperscript{21}. The independent testing of the FFPET biopsies at 2 laboratories, beginning after the scrolling of the blocks, allowed determination of inter-laboratory concordance and the robustness and portability of the assay. The Lymph2Cx scores produced at the two sites show a very high degree of concordance (Figure 1B).

Outcomes in the validation cohort were used to determine whether the COO assignments made by the Lymph2Cx assay maintained the prognostic significance previously demonstrated for the original FT sample based method\textsuperscript{6}. The “gold standard method” and Lymph2Cx defined ABC groups both had significantly worse outcome than the GCB groups (Figure 2). Larger cohorts will need to be examined to provide the statistical power to determine whether this prognostic power is independent of other biomarkers, particularly the International Prognostic Index\textsuperscript{22}. Outcomes in the COO groups assigned by IHC were not significantly different in this cohort (Figure S1).

In summary, herein we have described a robust method for COO assignment, applicable to FFPET biopsies that are generated as part of routine diagnostic workflow. Tested against the “gold standard” FT Affymetrix-based method, the rates of correct assignment are very high and the prognostic significance is maintained. Thus, the Lymph2Cx assay brings to fruition the potential to use gene expression-based COO assignment, first described over a decade ago\textsuperscript{2}, in standard practice. With this demonstrated accuracy and
portability, combined with a rapid turn-around-time of less than 36 hours, the application of this assay will enable prospective selection of patients for therapeutic clinical trials and, ultimately, will guide appropriate patient management in clinical practice.

Acknowledgements: The project was funded by the National Cancer Institute, grant # 1U01CA157581-01. DWS is supported by a Canadian Institutes of Health Research Post-Doctoral Fellowship. AM is supported by a fellowship award by the Mildred Scheel cancer foundation. The authors would like to acknowledge the contributing efforts of past members of the LLMPP, our collaborators on previous DLBCL projects, and administrative support.

Author Contributions:

David W Scott: Designed and performed research, wrote manuscript
George Wright: Designed and performed research, wrote manuscript
P. Mickey Williams: Designed research, read and approved final manuscript
Chih-Jian Lih: Performed research, read and approved final manuscript
William Walsh: Performed research, read and approved final manuscript
Elaine S. Jaffe: Contributed vital reagents, performed research, read and approved final manuscript
Andreas Rosenwald: Contributed vital reagents, performed research, read and approved final manuscript
Elias Campo: Contributed vital reagents, performed research, read and approved final manuscript

Wing C. Chan: Contributed vital reagents, performed research, read and approved final manuscript

Joseph M Connors: Performed research, curated clinical and survival data and read and approved final manuscript

Erlend B. Smeland: Contributed vital reagents, read and approved final manuscript

Anja Mottok: Performed research, read and approved final manuscript

Rita M. Braziel: Contributed vital reagents, performed research, read and approved final manuscript

German Ott: Contributed vital reagents, performed research, read and approved final manuscript

Jan Delabie: Contributed vital reagents, performed research, read and approved final manuscript

Raymond R. Tubbs: Contributed vital reagents, performed research, read and approved final manuscript

James R. Cook: Contributed vital reagents, performed research, read and approved final manuscript

Dennis D. Weisenburger: Contributed vital reagents, performed research, read and approved final manuscript

Timothy C. Greiner: Contributed vital reagents, performed research, read and approved final manuscript
Betty J. Glinsmann-Gibson: Performed research, read and approved final manuscript

Kai Fu: Contributed vital reagents, performed research, read and approved final manuscript

Louis M. Staudt: Designed and performed research, read and approved final manuscript

Randy Gascoyne: Designed and performed research, edited and approved final manuscript

Lisa M. Rimsza: Designed and performed research, edited and approved final manuscript

Conflicts of Interest:
The authors declare they have no relevant conflicts of interest to disclose.

References:


Figure Legends:

Figure 1: Performance of the Lymph2Cx assay in the independent validation cohort.

(A) The Lymph2Cx model is shown in the form of a gene expression heatmap (upper panel) with 67 patient samples from the independent validation cohort arrayed left to right in ascending order of the assay score. The 20 genes that contribute to the model are shown at the left with the top 8 genes being over-expressed in activated B-cell-like DLBCL (ABC), the middle 5 genes being “housekeeping” genes and the lower 7 genes being over-expressed in germinal center B-cell-like DLBCL (GCB). In the lower panel, the cell-of-origin assignments are shown for the assay, the “gold standard method” using previously published algorithms on gene expression from frozen tissue and 3 immunohistochemistry based algorithms. The Lymph2Cx results shown are from the Molecular Characterization Laboratory (FNLCR, Frederick, MD), with 1 of the 68 cases in the independent validation cohort having failed. Results from the Centre for Lymphoid Cancer, BC Cancer Agency, Vancouver, Canada are shown in Figure S2.

(B) Comparison of the Lymph2Cx scores in the validation cohort from the 2 independent laboratories – the Molecular Characterization Laboratory (MoCha)
(FNLCR, Frederick, MD) and the Centre for Lymphoid Cancer (CLC), BC Cancer Agency (Vancouver, Canada). The dotted lines represent the thresholds between GCB, unclassified and ABC. The 66 of 68 cases where both laboratories generated results are shown. The 3 cases that gave discordant cell-of-origin assignments are shown in red. The concordance is 98%, when considering the ABC and GCB cases by the “gold standard” method, and 95% if the unclassified cases are included. The $R^2$ is 0.996 and the slope of the line of best-fit is 1.015. Comparisons in the training and total cohorts are shown in Figure S3.

**Figure 2: Patient outcomes according to cell-of-origin in the independent validation cohort.**

(A) Progression free survival in the cell-of-origin groups as determined by the Lymph2Cx assay. (B) Overall survival in the cell-of-origin groups as determined by the Lymph2Cx assay. (C) Progression free survival in the cell-of-origin groups determined by the “gold standard” method applying the previously described model$^6$ to gene expression on frozen tissue. (D) Overall survival in the cell-of-origin groups determined by the “gold standard” method. The $P$-values are from log-rank tests comparing the ABC and GCB groups. The log-rank tests are one-sided in the direction of greater Hazard for ABC. RR is the relative risk (with the 95% confidence interval in brackets) associated with the ABC group compared with the GCB group. The groupings in (A) and (B) are from the results at the Molecular Characterization Laboratory (FNLCR, Frederick, MD). Results from the Centre for Lymphoid Cancer, BC Cancer Agency, Vancouver, Canada are shown
in Figure S4.
Figure 2

A

Proportion FFS

P < 0.001  RR = 3.6 (1.6-8.4)

Time (years)

B

Proportion OS

P = 0.01  RR = 2.8 (1.1-7.3)

Time (years)

C

Proportion FFS

P = 0.01  RR = 2.6 (1.1-6.3)

Time (years)

D

Proportion OS

P = 0.04  RR = 2.3 (0.8-6.3)

Time (years)

Germinal-Center B-cell-like DLBCL  Unclassified DLBCL  Activated B-cell-like DLBCL
Determining cell-of-origin subtypes of diffuse large B-cell lymphoma using gene expression in formalin-fixed paraffin embedded tissue

David W. Scott, George W. Wright, P. Mickey Williams, Chih-Jian Lih, William Walsh, Elaine S. Jaffe, Andreas Rosenwald, Elias Campo, Wing C. Chan, Joseph M. Connors, Erlend B. Smeland, Anja Mottok, Rita M. Braziel, German Ott, Jan Delabie, Raymond R. Tubbs, James R. Cook, Dennis D. Weisenburger, Timothy C. Greiner, Betty J. Glinsmann-Gibson, Kai Fu, Louis M. Staudt, Randy D. Gascoyne and Lisa M. Rimsza

Information about reproducing this article in parts or in its entirety may be found online at:
http://www.bloodjournal.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at:
http://www.bloodjournal.org/site/misc/rights.xhtml#reprints

Information about subscriptions and ASH membership may be found online at:
http://www.bloodjournal.org/site/subscriptions/index.xhtml

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include digital object identifier (DOIs) and date of initial publication.