Toward a New Molecular Taxonomy of Diffuse Large B-cell Lymphoma

Daisuke Ennishi, Eric D. Hsi, Christian Steidl, and David W. Scott

ABSTRACT

Diffuse large B-cell lymphoma (DLBCL) represents a grouping of clinically and biologically heterogeneous tumors. Application of advanced molecular technology has significantly expanded our knowledge of DLBCL pathobiology, allowing identification of subgroups with common, potentially targetable, biological themes. Here, we review the recent molecular analyses that could provide a paradigm shift to a new taxonomy, foundational to the rational transition to precision medicine. We discuss how classification systems may be synthesized into a common taxonomy, drawing strength from the relationships between genetic alterations, gene expression, and tumor microenvironment. Finally, challenges to translating such a taxonomy to the clinic will be outlined.

INTRODUCTION

Collectively, lymphoid cancers are the fourth most common cancers in both men and women and thus represent a significant healthcare issue. Diffuse large B-cell lymphoma, not otherwise specified (DLBCL, NOS) is the most common lymphoma subtype worldwide, accounting for 40% of all non-Hodgkin lymphomas (NHL; ref. 1). An aggressive disease, DLBCL requires immediate treatment but is one of only a very few cancers that can be cured even when disseminated at the time of initial diagnosis. The current standard of care for patients treated with curative intent is multiagent (CHOP) chemotherapy in combination with rituximab, an anti-CD20 antibody (R-CHOP; refs. 2, 3). Despite the significant improvement in outcomes with the addition of rituximab, approximately 40% of patients experience relapse or refractory disease. Further improvement in treatment outcome will rely on elucidating the molecular determinants related to treatment response. However, the pathologic and genetic heterogeneity driving DLBCL development has not been fully elucidated. Consequently, to date, targeted therapies have not significantly improved survival in patients with DLBCL, especially patients who have relapsed, and immunochemotherapy (R-CHOP) remains the standard of care.

For a more cogent understanding of the pathobiology of cancers, investigation of normal cellular counterparts is needed, especially in hematologic malignancies. DLBCL is postulated to largely arise from the malignant transformation of mature B cells that have experienced the germinal center (GC) reaction (Fig. 1; refs. 4, 5). The GC B cell is at particularly high risk for undergoing malignant transformation, due to processes essential to immunoglobulin affinity maturation including attenuation of certain DNA damage and cell proliferation checkpoints (6, 7). GCs are transient and dynamic structures with distinct microanatomic compartments: the dark zone (DZ) and the light zone (LZ; ref. 8–10). In the DZ, antigen-activated GC B cells undergo rapid proliferation and somatic hypermutation (SHM) of their immunoglobulin variable genes to generate high-affinity B-cell receptors (BCR). Although class switch recombination (CSR) was previously widely considered to occur in the GC, a recent study has strongly suggested that CSR mainly takes place after interaction with cognate T cells and thus prior to establishment of GCs (11). SHM and CSR are mediated by activation-induced cytidine deaminase, which introduces genomic instability as part of the processes responsible for affinity maturation and immunoglobulin class switching (12). Postreplicative GC B cells migrate to the more heterogeneous milieu of the LZ, where they interact with T follicular helper (Tfh) and follicular dendritic cells (FDC). Of note, although high-affinity GC B cells further differentiate into plasmablasts or memory B cells, a small subset of high-affinity LZ GC B cells cycles back to the DZ for additional rounds of SHM and proliferation, whereas the low-affinity GC B cells undergo apoptosis. These unique and complex events pose a significant risk to the genome of B cells, which have to endure high replication stress while undergoing multiple DNA mutation, breakage, and recombination events (13–16). When critical regulatory checkpoints fail, the physiologic GC reaction itself is thus proposed as the genesis of most mature B-cell lymphomas, including DLBCL. As a corollary, we propose in this review that a deeper understanding of the malignant transformation processes of B cells can provide
conceptual frameworks for a better classification system with therapeutic implications.

The current standard classification of aggressive B-cell lymphomas, the 2017 revision of the WHO classification, is the result of continued evolution of systems dating back to the 1950s (recently summarized by Swerdlow and Cook; ref. 17). The current state-of-the-art taxonomy incorporates chromosomal rearrangements and gene expression alongside morphology, immunophenotype, and disease site. Recent advances in characterizing genomic, epigenomic, transcriptomic, proteomic, and microenvironmental alterations have further significantly expanded our knowledge of DLBCL pathobiology. Insights gained from applying this new knowledge will likely lead to the identification of predictive molecular biomarkers and new drug targets, paving the way for clinical trials focused on disease subsets aiming to personalize therapy. Here, we will review the most recent innovations toward a new DLBCL molecular taxonomy, and discuss the apparent tension between existing and new concepts along with the potential to integrate these systems into a new taxonomy with an emphasis on actionability and clinical utility.

**CURRENT STATE-OF-THE-ART CLASSIFICATION: PATHOLOGY WORKUP**

The disease spectrum of aggressive B-cell lymphoma harbors significant clinical and biological heterogeneity. Classification into the neoplastic entities of the revised 2017 WHO classification requires the examination of tumor cell morphology and immunophenotype, molecular testing for recurrent chromosomal rearrangements, and integration with clinical and radiologic information, including sites of
Toward a New Molecular Taxonomy of DLBCL

Two-year overall survival (OS) was universally reported as less than 50%, providing a rationale to attempt intensive initial treatments, such as DA-EPOCH-R, R-hyper-CVAD, and R-CODOX-M/IVAC. It should be noted that conclusive data from well-controlled clinical studies do not currently exist that unequivocally demonstrate improved outcomes for patients with HGBL-DH/TH treated with these regimens (24, 28). It should be also noted that recent studies in which FISH testing was performed for all cases have shown the outcomes of HGBL-DH/TH treated with R-CHOP are superior to those previously reported, with the largest study showing a 5-year progression-free survival of 60% (23, 29–31). This likely reflects a historic selection bias in FISH testing of patients with high-risk features at diagnosis that was mitigated in the most recent studies by applying FISH to all tumors. Examples of this historic bias include selective testing of tumors with perceived “high-grade” morphologic features such as the “starry sky” patterns that are often seen in Burkitt lymphoma or other highly proliferative lymphomas, high proliferation rate, or poor response to up-front treatment.

The WHO classification, along with the recognition that many HGBL-DH/TH lack distinctive morphologic features, requires that FISH testing for MYC rearrangement be integrated into the diagnostic workup for all tumors with DLBCL morphology. Up-front FISH testing for BCL2 and BCL6 rearrangements or a sequential approach, where only MYC rearranged lymphomas are tested, are both rational approaches. Of biological importance, HGBL-DH/TH-BCL2 occurs almost exclusively in tumors with a GCB GEP (21, 31), allowing the screening of tumors for FISH testing by a COO test. However, this screening approach would reduce FISH testing only to approximately one-half of all tumors. Furthermore, this strategy would miss the rare population of HGBL-DH-MYC, which sometimes have an ABC gene-expression pattern (21). Therefore, the current recommendation is to perform FISH for MYC rearrangement in all tumors with DLBCL morphology. Data continue to emerge, with recent studies suggesting that the aggressive behavior of HG-DH/TH is largely due to cases with an immunoglobulin gene partner (heavy or light chain; ref. 23). This has implications on FISH strategies and expands the number of probes required to comprehensively evaluate a given case.

Once a diagnosis of DLBCL, NOS is made, a COO subtype is assigned. COO subtypes were originally described based on GEP by DNA microarray using RNA extracted from fresh tissue (32). Because of the technical difficulties in applying GEP in daily clinical practice, alternative classification methods based on IHC algorithms have been developed. Since the first published Hans algorithm in 2004 (33), which uses three markers (CD10, BCL6, and MUM1), several improved IHC algorithms have been suggested, such as “Choi,” “Tally,” and “Visco-Young” (34–36). Although the WHO recommends using GEP to assign COO, no such assays are currently widely available, so IHC-based algorithms are deemed acceptable. The WHO classification does not recommend which algorithm should be used but specifies that the method should be noted. Although these IHC-based COO tests have been widely available in routine clinical use, there is some degree of discrepancy with gold standard GEP-defined COO (15 to 50%), partly explained by their binary nature (not identifying 10%
to 15% of biopsies “unclassified” COO subtype by GEP; refs. 37, 38). In addition, as highlighted by several studies examining reproducibility among different laboratories (39), this lack of concordance is also in part due to sampling and technical issues as well as interobserver variation. These factors may have contributed to variable reports and conclusions regarding the prognostic value of IHC-based COO subtyping in the rituximab era (40, 41). The development of novel therapeutic agents with reported selective activity in ABC and GCB subtypes will depend on an accurate and reproducible assay for determining COO.

Determining the protein expression of MYC and BCL2 is also an important component of the current state of the art. Several groups have shown that the analysis of MYC and BCL2 protein expression in DLBCL identifies tumors that overexpress MYC and BCL2 (DEL), ranging from 25% to 40% of patients (42–44). Although there is ongoing debate about the threshold of MYC and BCL2 protein expression positivity by IHC, the thresholds recommended in the WHO classification are 40% (positive tumor cells) for MYC and 50% (positive tumor cells) for BCL2. Retrospective series have reported poor outcomes for this group of patients, with an OS of 30% to 40% at 5 years following R-CHOP therapy (42–44). In the context of COO subtypes, cases of DEL are more common in the ABC/non-GCB subtype compared with the GCB subtype, particularly when HGBL-DH/TH-BCL2 are excluded, with these tumors frequently being DEL (45, 46). Hu and colleagues suggested that coexpression of MYC and BCL2 may explain the adverse prognostic impact of ABC/non-GCB-DLBCL, whereas other studies showed that DEL status and COO subtype are independent prognostic factors (44, 45). Recognizing the poor outcome of patients and the potential for directed therapy, clinical trials are ongoing that investigate the efficacy of specifically targeting BCL2 in DEL DLBCL, NOS and HGBL-DH/TH lymphomas. It is important to note that DEL should not be considered a taxonomic class in and of itself, as lymphomas arrive at this expression phenotype from very distinct molecular mechanisms (47).

With an expanded knowledge base about pathobiology and new emerging therapies in the era of precision medicine, the translational lymphoma field has engaged in a deep and sometimes controversial discussion about appropriate diagnostic algorithms and standards for current pathology workups (Fig. 2) to provide the most relevant molecular and genetic information for outcome prediction and identification of potential therapeutic targets.

SUBTYPES OF DLBCL DEFINED BY TRANSCRIPTOMIC ANALYSIS

COO Classification

The evolution of technologies resulted in the exploration of cancer molecular subtypes first focusing on transcriptional profiles of tumor samples. Early transcriptional profiling of
primary DLBCL samples using microarrays classified tumors into various transcriptional subtypes with clinical correlates. In 2000, Alizadeh and colleagues first distinguished DLBCL into at least two major subgroups, unrecognizable on morphologic grounds, with a gene signature referred to as the ABC and GCB subtype reflecting the COO of tumors (32). Subsequent studies also demonstrated the biological and clinical distinction of COO subtypes and defined PMBCL as distinct from DLBCL, NOS (48–51).

The COO designations likely represent transcriptional footprints of distinct patterns of genetic and epigenetic alterations superimposed on transcriptional patterns related to stage of differentiation of B cells. GCB-DLBCL lacks the expression of early post-GC differentiation markers and thus is more reminiscent of centroblasts and centrocytes in the DZ and LZ, respectively, whereas ABC-DLBCL is thought to correspond to GC-experienced B cells poised to undergo terminal differentiation (8, 32, 52). The pathogenesis of ABC-DLBCL is characterized by two main events: constitutive activation of NFκB and the blockade of terminal differentiation to plasma cells (53–56), achieved by distinct genetic alterations described later (57–61). On the other hand, underlying events of GCB-DLBCL appear more complex. Genetic studies have shown recurrent alterations in GNA13 (62), TNFRSF14 (63), and chromatin-modifying genes [including EZH2 (64) and CREBBP (65)], with only a few other biologically relevant lesions identified (13, 53, 66). An important example is chromosomal translocations of BCL2 detected in approximately 40% of GCB-DLBCL, and all HGBL-DH/TH-BCL2 tumors display the GCB gene-expression pattern (21, 31, 67, 68). It is important to note that COO does not necessarily reflect the time point in B-cell differentiation at which the critical driver mutations were accrued but rather the point at which differentiation is arrested. An example is that BCL2 rearrangements in tumors, which largely have a germinal center COO, occur during V(D)J recombination, a process that occurs earlier in differentiation in the bone marrow (69).

Clinically, the ABC subtype is associated with worse outcomes; population-based studies report 5-year OS rates of 35% for patients with ABC-DLBCL and 60% for patients with GCB-DLBCL in the pre-rituximab era (49), and 5-year OS rates of 56% for patients with ABC-DLBCL and 78% for patients with GCB-DLBCL following first-line R-CHOP treatment (45). Owing to the significant molecular and clinical differences between ABC- and GCB-DLBCL subtypes, COO identification offers opportunities to tailor treatment to the tumor biology. Of note, several targeted therapies have shown COO selective efficacy at the juncture of relapse/primary progression (70–74). Interestingly, at this time, this has not translated to improvement in outcomes when these agents have been combined with R-CHOP in the up-front treatment of DLBCL (75–79). It appears that additional granularity, beyond the groupings corresponding to a binary division based on B-cell differentiation stage, is required to adequately support precision medicine in DLBCL.

The initial requirement for the microarray technology and fresh-frozen tissue prevented penetration of GEP into routine clinical practice, where formalin-fixed paraffin-embedded (FFPE) biopsies are used. A range of technologies have now been described for GEP in FFPE, allowing the development of accurate and reproducible COO assignment. One example is the Lymph2Cx assay, measuring the expression of 20 genes (80). The concordance of this assay with traditional GEP methods from fresh-frozen samples was reported to be >95%, and survival outcomes were also similar to those obtained by the gold standard GEP technique. Importantly, concordance between two independent laboratories employing this technique was greater than 95% (80). Subsequent analyses in a population-based large cohort showed a significant prognostic effect independent of International Prognostic Index as well as DEL status (45). Although the Lymph2Cx has been used to assign COO within, and to select patients for, clinical trials, this approach is not currently available for routine pathology laboratories. It seems likely that this assay, or similarly developed assays that can be applied to FFPE, will be used increasingly in the future.

Other Gene Expression–Based Classifications

Subsequent to the description of the COO molecular subtypes, a number of additional transcriptionally defined systems for classification within DLBCL have been proposed. Those most recently described largely fit within the COO classification system. Einnishi and colleagues developed a gene-expression signature that distinguished HGBL-DH/TH-BCL2 from GCB-DLBCL—the so-called “DHTsig” (81). Meanwhile, Sha and colleagues identified a group of patients with gene expression intermediate between DLBCL and Burkitt lymphoma (82). Although independently derived, both signatures reveal a poor prognostic group within DLBCL tumors with a GCB gene expression profile with significant concordance of patients identified by both signatures. First, both signatures identified that almost all HGBL-DH/TH-BCL2 cases fell within this high-risk group along with an equally large group of non–HGBL-DH/TH-BCL2 DLBCL tumors. Interestingly, the HGBL-DH/TH-BCL2 and non-HGBL-DH/TH-BCL2 cases showed comparable outcomes after treatment with R-CHOP. These new aggressive signature subgroups roughly double the number of DLBCL tumors that would be classified as HGBL-DH/TH-BCL2 on the basis of FISH testing alone. Second, the gene-expression signatures were associated with centroblasts or intermediate zone GCB cells, with low expression of the centrocyte signature (52, 83). In addition, these tumors are enriched with signatures of high cell proliferation, as expected, and represent immunologically “cold” tumors with lower tumor-infiltrating lymphocytes (TIL) and frequent loss of MHC-I and/or MHC-II expressions. Third, the genetic backgrounds are very similar in the new high-risk subgroups: enrichment of gene mutations in epigenetic modifiers including EZH2 and CREBBP. Interestingly, some recurrent mutations in Burkitt lymphoma, such as DDX3X and BCL7A, were enriched in the new high-risk subgroups, whereas other Burkitt lymphoma–specific gene mutations (e.g., ID3 and TCF3, refs. 84, 85) were not observed, indicating the intermediate feature of this new subtype between DLBCL and Burkitt lymphoma. This identification of a biologically and clinically distinct group that encompasses tumors within both the current HGBL-DH/TH and GCB-DLBCL groups suggests that a redefinition of these boundaries may be warranted. Furthermore, the recent discovery of MYC and BCL2 rearrangements that are cryptic to FISH testing within the
DHTsig-positive group demonstrated that at least 19% of HGBL-DH/TH-BCL2 are being missed by FISH. This suggests that these expression signatures could be a more accurate means than FISH in identifying high-risk DLBCL with a GCB gene-expression pattern (86).

Additional gene expression-based groups within DLBCL have been proposed on the basis of malignant cell autonomous (“intrinsic”) and microenvironment-related (“extrinsic”) properties. Oxidative phosphorylation (OxPhos) and non-OxPhos subtypes were developed according to the different fingerprints of cellular energy metabolism. Initially, DLBCL cases were clustered into three groups including the OxPhos cluster (OxPhos DLBCL), which is significantly enriched in genes involved in mitochondrial OxPhos (87). OxPhos DLBCLs do not have functional BCR signaling and are insensitive to BCR inhibition, suggesting that they are dependent on alternative survival mechanisms. Moreover, recent mitochondrial proteomics and gene-expression analysis have revealed that the mitochondrial proteome of OxPhos DLBCL cells was significantly enriched for enzymes involved in mitochondrial β-oxidation, the tricarboxylic acid (TCA) cycle, OxPhos, and detoxification of reactive oxygen species, compared with non-OxPhos DLBCL cells (88). Notably, OxPhos DLBCL cells were selectively sensitive to pharmacologic or genetic inhibition of fatty-acid oxidation, suggesting that the metabolic features of this subtype could be exploited therapeutically.

Finally, gene-expression studies of DLBCL have identified molecular signatures present in both GCB and ABC subtypes related to the tumor microenvironment (“extrinsic” factors) that correlated with outcome (89). The prognostically favorable “stromal-1” signature reflects reprogrammed stromal cells, extracellular matrix, and an active immune response. The less favorable “stromal-2” signature indicates elevated angiogenesis and blood vessel density. These studies suggest that metabolic and microenvironment features contribute to DLBCL pathogenesis. These signatures have only recently been translated onto tractable technology platforms allowing both the impact on patient outcomes to be confirmed as well as the possibility of patient selection for relevant therapeutic agents (90, 91).

**GENETIC ALTERATIONS AND MUTATIONAL PROFILING OF DLBCL**

With advances in genomic technology, the heterogeneity of DLBCL has been further dissected on the basis of genetic alterations—another layer of molecular definitions beyond GEP (92–96). Mutational landscape studies cumulatively examined well over 2,000 primary DLBCL samples that have revealed the presence of many recurrent alterations, with more than 150 putative lymphoma driver genes identified (92–100). Specifically, one of the large-scale whole-exome sequencing studies revealed a median of 17 genetic alterations per DLBCL case, which is a higher mutation rate compared with other hematologic malignancies (98), indicating that DLBCL are genetically characterized by significant intertumor and/or intratumor heterogeneity.

Up until very recently, the emerging landscape of genetic alterations in DLBCL was organized and functionally explored through the binary lens of COO. Here we will discuss the genetic alterations in the context of recent studies that proposed additional taxonomic groups. These studies aimed to improve the molecular classifications with integrated genetic and mathematical approaches, and identified genetic subtypes based on shared genomic abnormalities. In 2018, two seminal papers, by Chapuy and colleagues (98) and Schmitz and colleagues (99), put forward systems of molecular subgroups largely emerging from examining whole-exome sequencing, with a recent refinement of the Schmitz classification resulting in these groupings largely aligning (101). Recently, Lacy and colleagues applied targeted sequencing to more than 900 aggressive B-cell lymphomas drawn from a population-based registry (100). The consistency of the molecular subtypes identified by these three groups, despite the independent tumor cohorts and distinct mathematical approaches, provides confidence that they are truly biologically distinct. These subtypes are described below, noting that each group has used their own nomenclature and their full alignment has not been formally demonstrated.

The MCD subtype [aligning with C5 (98) and MYD88 (100)] is part of the ABC spectrum, and is strongly enriched for MYD88L265P mutations and CD79B mutations and amplifications, giving rise to activation of the NFκB pathway through signaling downstream of TLR9 and chronic active BCR signaling (58–61). Interestingly, although DLBCLs carrying the MYD88-mutant isoform did not respond to BTK inhibition in a recent clinical trial, exceptional responses were observed in tumors with concurrent MYD88 and CD79A/B mutations (72). Furthermore, an experimental study revealed that CD79B and surface IgM constitute a rate-limiting checkpoint against B-cell dysregulation by MYD88L265P, suggesting that these pathways may be functionally coupled (102, 103). Gain/amplifications of 18q, another frequent genetic alteration of the MCD group, encompass several driver genes such as BCL2, NFATC1, and MALT1 in a minimum common region at 18q21 and are associated with increased BCL2 expression and poor outcome (53, 104, 105). In addition, TCF4 gain/amplification at 18q21 has recently been shown to increase MYC expression by regulating the MYC enhancer (106). Other genetic alterations enriched in the MCD group include PRDM1, as well as copy-number gain/amplification of SPIB, a transcription factor contributing to a block of terminal differentiation by forming a complex with IFR4 as well as inactivating PRDM1 (73).

The BN2 (aligning with C1 and NOTCH2) subtype is mostly assigned to ABC-DLBCL and unclassified DLBCL, and represents a novel group of tumors associated with favorable clinical outcomes. These genetic subtypes are characterized by BCL6 translocations and NOTCH2 activating mutations, and are proposed to arise from a marginal zone B-cell origin on the basis of the recurrent NOTCH2 mutations (107–110). NFκB signaling may be also activated in this subgroup by recurrent genetic alterations, such as deletion of TNFAIP3 (A20; ref. 57) and gain/amplification of BCL10 (60). Interestingly, recently our study discovered that TMEJ30A mutation, a genetic component of this subtype, is associated with favorable outcome due to increased accumulation of chemotherapy drugs and enhanced phagocytosis (111). This finding supports the clinical characteristics of
this group, and further provides potential subtype-specific therapeutic approaches. On the other hand, most EZB (aligning with C3 and BCL2) cases are assigned to GCB-DLBCL and are characterized by BCL2 translocations and EZH2 mutation. Other mutations of chromatin modifiers, such as CREBBP and KMT2D, as well as TNFRSF14 loss-of-function mutations and PTEN inactivation, were also components of these subtypes. Monoallelic and, less commonly, biallelic somatic mutations of KMT2D represent the single most frequent somatic mutations in DLBCL (~30%) as well as follicular lymphoma (~80%); refs. 92–96. Indeed, conditional inactivation of KMT2D in vivo leads to the expansion of GC B cells, and cooperates with BCL2 deregulation to increase the incidence of tumors recapitulating phenotypic and genetic features of human follicular lymphoma/DLBCL, thereby establishing KMT2D as a bona fide tumor suppressor gene (112, 113). Mutations in the histone acetyltransferases genes CREBBP and EP300 are also predominantly monoallelic, mutually exclusive, and accompanied by expression of the residual wild-type allele—a pattern consistent with a haploinsufficient tumor suppressor role. These mutations remove the histone acetyltransferase domain or introduce amino acid changes within this domain, which cause diminished affinity for Acetyl-CoA, resulting in the loss of function of CREBBP/EP300 proteins (65). EZH2, which encodes the catalytic component of the polycomb repressor complex 2, is one of the most frequently mutated genes in human lymphomas, especially GC-derived lymphomas, accounting for 25% to 30% of patients with follicular lymphoma and 30% of patients with GCB-DLBCL (92–96). EZH2 gain-of-function mutations drive lymphomagenesis by repressing target genes involved in proliferation checkpoints (e.g., CDKN1A) and B-cell terminal differentiation (e.g., IRF4 and PDRM1; refs. 114, 115). Of note, TNFRSF14, EZH2, and CREBBP mutations affect the immune microenvironment, as will be described later. The EZB group may be further divided into two groups using the DHIT-sig gene-expression signature, identifying groups with distinct mutational profiles and outcomes. The genetic profile, enriched with epigenetic modification genes and TNFRSF14, highlights similarity between EZB and follicular lymphoma. It is proposed that some of these tumors may arise from occult follicular lymphoma (i.e., transformation) or share a common progenitor cell (101, 116–118).

Chapuy and colleagues (98) describe the C4 cluster which also includes GCB cases and show better clinical outcome compared with the EZB cases. Genetic alterations involved in the C4 cluster include H1 linker histones and additional core histones including HIST1H1/2 genes, which were also recurrently mutated in follicular lymphoma (116). Additional alterations enriched in this cluster include several known driver genes, such as CD58, CARD11, and KLF11. Specifically, inactivating CD58 mutations are associated with immune recognition and modulate the cytolytic capacity of NK cells (119). Interestingly, CD58 was also reported to be an important marker of centrocytes in the LZ of GC (52, 120), and therefore this molecule may play a critical role for the distinction of EZB and C4. It is possible that the C4 cluster is composed of two subgroups. The first, labeled ST2 (101) or TET2/SGK1 (100), is characterized by mutations in SGK1 and TET2. The second, labeled SOCS1/SGK1, contains tumors with mutations in SGK1 and SOCS1. Interestingly, when PMBCLs were classified using the system put forward by Lacy and colleagues, they largely fell into the SOCS1/SGK1 group (100). This molecular subtype of DLBCL may explain the observation of tumors with a PMBCL gene expression signature that lack a mediastinal mass (121). It is suggested that further research is needed to firmly support these subtypes within C4.

Finally, the A53 (aligning with C2) subtype is defined by the presence of TP53 mutations, copy-number variations, and an increase in ploidy and is consistent with an earlier report describing complex copy-number states in a subgroup of DLBCL containing both ABC- and GCB-DLBCL (104). An additional group was identified by Schmitz and colleagues (99): the N1 subtype, characterized by mutations in NOTCH1, with this rare subtype comprising <2% of DLBCL.

These studies thus proposed novel molecular taxonomies with improved risk stratification beyond COO. Indeed, these studies might explain discrepant results from prior studies challenging the prognostic significance of COO assignment by GEP alone, because unknowing enrichment of study populations in favorable subtypes of ABC-DLBCL and/or unfavorable subtypes of GCB-DLBCL could affect results.

Future methods to determine the correct assignment to mutational groups will have to find the right balance between forcing classification of individual tumors and leaving tumors unclassified because of insufficient classification confidence. The clustering approach of Chapuy and colleagues (98) by definition forces tumors into a category, with only a small genetically undefined group (C0 – less than 4% of analyzed tumors), and may have produced molecular groups that were overinclusive of tumors. As clustering approaches are intractable for a tumor-by-tumor assignment assay, further development of their system is awaited. The refinement of the Schmitz and colleagues (99) system included an algorithm for assigning molecular subtype on a tumor-by-tumor basis (101). However, this algorithm results in 37% of tumors being unclassified. This may indicate that less prevalent genetic subtypes (such as the N1 subtype) may exist among these unassigned tumors or that whole-exome sequencing incompletely captures mechanisms by which tumors arrive at the biology of these newly defined entities. Finally, validation of these findings in larger sets of tumors with demographic diversity is key to providing a foundation for precision medicine in this disease.

**MICROENVIRONMENT BIOLOGY OF DLBCL**

**Biology of Immunologically Cold Tumors**

Traditionally, DLBCL had been recognized as less dependent on its microenvironment as compared to other lymphomas, in agreement with a near-complete disorganization and/or displacement of normal lymphoid architecture. In addition, given that tumors with very distinct genetic backgrounds share similar tumor microenvironment composition, this should not form the sole basis of a comprehensive taxonomy. However, there is increasing evidence that an immunologic niche and cross-talk with various immune cell types is critical for disease development and adds another
layer of complexity to genetic and molecular subtypes. In particular, it has been increasingly recognized that the disrupted cross-talk between lymphoma cells and the microenvironment contributes to the ability of lymphoma cells to escape the immune surveillance of the host in DLBCL (Fig. 3; ref. 122).

Attenuated expression of MHC complexes plays a key role in immune escape of DLBCL (123, 124). The frequent deficiency of MHC-I expression on the surface of DLBCL cells was observed in DLBCL, based on genetic mechanisms such as inactivation of β2-microglobulin (B2M) and CD58 (119). On the other hand, B cells are themselves professional antigen-presenting cells; thus, MHC-II is normally expressed, and selection in the LZ of the GC involves antigen presentation via MHC-II to Tfh cells and FDCs regulated by the ubiquitin ligase MARCH1 and CD83 (52, 125). Within the LZ, the GC B cells receive prosurvival BCR signals by combining with Tfh cells through CD40, driving NFκB activation and subsequent IRF4-driven suppression of BCL6 (126, 127). In B-cell lymphomas, tumor antigens are also presented via MHC-II and recognized by CD4-positive T cells driving an antitumor immune response (128, 129). Thus, antigen presentation is concealed for these cells to escape killing in GC-driven B-cell lymphomas by reducing MHC-II expression. In keeping with this notion, MHC-II expression is often lost in GC-derived neoplasms, which was associated with an aggressive clinical course (130–133) and poor host tumor-infiltrating T-cell response (134, 135). In the COO-specific context, loss of MHC-II expression occurs more often in ABC-DLBCL compared with GCB-DLBCL, as the expression of transactivators of MHC-II is silenced when GC B cells transition to plasma cells (136). At the other end of the differentiation spectrum in the GC, a recent study demonstrated that loss of MHC-II expression also defined tumors that likely originated from the DZ of GC, with these tumors being associated with inferior treatment and an immune “cold” microenvironment (137).

**Molecular Basis of Immune Escape and Therapeutic Potential**

On the transcriptional level, although COO classification mainly reflects malignant B-cell features, two gene expression profiling studies have highlighted another level of DLBCL biological heterogeneity underlying the role of the microenvironment (87, 89). Although these studies suggest that the microenvironment component plays a major role in DLBCL pathogenesis, such descriptive studies do not provide any mechanistic insights into the interaction between lymphoma cells and the microenvironment. More recent studies have attempted to translate microenvironment biology into clinical assays of GEP using FFPE with prognostic properties in R-CHOP-treated patients (91, 138).

Mutational landscape studies have highlighted recurrent mutations with focal copy-number deletions involved in immune recognition of DLBCL cells. Inactivating mutations and deletions in the B2M gene, crucially impairing MHC-I assembly and cell-surface expression, occur in 30% of DLBCL cases. Copy-number loss of HLA-I loci at chromosome 6p21 is also a recurrent genetic event associated with reduced MHC-I expression of lymphoma cells (104). Genetic mechanisms of loss of MHC-II proteins in DLBCL are more complicated and diverse. In addition to deleted HLA-II loci, they inactivate CIITA, a critical transactivator of MHC-II expression, through inactivating somatic mutations. Interestingly, recent analyses also identified that MHC-II is downregulated by epigenetic aberrations, such as CREBBP (139, 140) and EZH2 (137) mutations. Of importance, HDAC3-specific and EZH2 inhibitors can rescue expression of MHC-II in CREBBP and EZH2-mutated lymphoma cells, respectively. These studies strongly suggest the potential of epigenetic reprogramming for priming the host immune system, providing attractive rationales for combination treatment strategies of epigenetic inhibitors with immune checkpoint inhibitors in a subset of DLBCLs.
Other recurrent genetic alterations associated with immune escape are inactivating mutations and focal copy-number deletions affecting CD58 (1p13.1), TNFRSF14 (1p36.32), and CD70 and TNFRSF14 (both 19p14.1). It should be noted that these genetic alterations were not consistently assigned to any genetic clusters as described above (98, 99), suggesting that immune microenvironment biology is a supportive layer rather than the basis of a molecular taxonomy. Functional models showed that CD58 surface expression on lymphoma cells modulated the cytolytic capacity of NK cells (119), and loss of TNFRSF14 can trigger both B cell–autonomous activation and B cell–extrinsic activation of the lymphoma microenvironment with B- and T-lymphocyte attenuator (BTLA) on Th cells (63, 92, 93, 141, 142). The advantage of losing CD70 expression remains to be investigated in DLBCL, but it is possible that loss of CD70–CD27 binding in the DLBCL context can alleviate the interaction of tumor B cells with potential antitumor T or NK effector cells. On the other hand, in line with the low expression of PD-L1 shown in DLBCL, the gain-of-function genetic changes of the PD-1/ PD-L1 axis is rare in patients with DLBCL (~10%; refs. 92, 104, 143). Of note, the C1 genetic subtype specifically harbors gains, amplifications, and translocations of the PD-L1/PD-L2 locus associated with increased expression, whereas its frequency is still low (20% of this subtype; ref. 98). These findings suggest that the PD-1/PD-L1 axis does not substantially contribute to the immune architecture of tumors, which explains the relatively low activity of immune-checkpoint inhibitors in patients with DLBCL compared with other cancers (144).

In addition to the genetic mechanism of immune escape, the heterogeneity of immune microenvironment subsets is being investigated by comprehensive deconvolution methods using transcriptome data of bulk tissues, and single-cell strategies including mass cytometry as well as single-cell RNA sequencing. With these advances, the expanded knowledge of tumor cell–microenvironment cross-talk will contribute to a more stable and nuanced taxonomy of DLBCL. This should facilitate subtype-specific therapeutic approaches as described in the following section.

**PERSPECTIVE AND CONCLUSION**

The discovery of recognizable COO subtypes approximately 20 years ago highlighted previously unrecognized heterogeneity. This provided the framework for further research into associated genetic and phenotypic features and clinical studies to test COO-specific novel targeted agents. However, recently, the emergence of multiple analytic approaches to comprehensively describe molecular features of DLBCL has added considerable texture and reignited the exciting challenge of developing a new widely accepted and clinically useful taxonomy of DLBCL. Ideally, this taxonomy would identify subgroups with homogeneous, potentially targetable, biology that could be used to form the basis of patient management in the emerging era of precision medicine. Furthermore, the system should be able to be widely and accurately deployed, assigning tumors in a suitable timeframe to avoid treatment delay. In the following paragraphs, we are offering a perspective into open scientific questions and potential solutions leveraging integrative approaches involving treating physicians, pathologists, and translational scientists.

**Avenues to a New Molecular Taxonomy**

Integrated data from several “omic” platforms have identified genetic clusters with substantial molecular similarities that stratify patient survival within each COO subtype, indicating that a new molecular taxonomy may have improved clinical utility (98, 99). To date, randomized clinical trials (REMOdL-B, PHOENIX, and ROBUST), specifically targeting ABC/non-GCB type, have failed to show additional effects of bortezomib, ibrutinib, and lenalidomide to R-CHOP in untreated DLBCL (75, 76, 145). These results suggest the limitation of isolated testing of COO subtypes before treatment, whereas the new genetic subtypes hold the, albeit yet to be tested, potential to provide groups enriched for patients more likely to benefit from these agents. Indeed, MCD includes tumors with both MYD88L265P and CD79B mutations, a genotype showing good response to ibrutinib in relapsed/refractory ABC-DLBCL (72).

We contend that the integration of genetic mutations and the transcriptome will provide a richer and more complete taxonomy, although the status quo still falls short of a truly integrative classification system, as current genetics-based systems do not incorporate epigenetics, the impact of the noncoding genome, or the differentiation state of the tumor cells. Regardless, an integration approach is attractive as it may robustly allow additional tumors to be assigned to the existing genetics groups, with these tumors arriving at a specific biology through mechanisms other than mutations of genes within these classifiers. In addition, this integration may uncover other biological subgroups that might currently reside in unclassified/“other” categories or that were inappropriately forced into a cluster designation (98, 99). Furthermore, subtypes defined by transcription phenotypes may provide further granularity within the genetic groups. This process has begun with the recently described subdivision of EZB using the DHIIts gene-expression signature, identifying a particularly poor-risk group of tumors within this group (101).

In addition, the relationship between genetics and immune microenvironment has been investigated, providing potential therapeutic interventions to restore the immune system’s antitumor activity, such as inhibitors of EZH2 and HDAC3, and its combination strategy with immune checkpoint inhibitors (137, 140). In the near future, opportunities to test different immunotherapies will increase with the success of CAR-T therapies (146, 147) and the efficacy of an anti-CD47 mAb (Hu5FG-G4) in patients with relapsed/refractory DLBCL (90). Thus, integrating a layer of microenvironmental insight into a new taxonomy will provide a more nuanced perspective on clinical utility and actionability by taking into account the intrinsic molecular profile and the host immune system together.

With regard to risk stratification, new molecular subtyping can detect a substantial number of favorable-prognosis patients who obtain sufficient benefit from the R-CHOP regimen, suggesting de-escalation of R-CHOP for such patients (e.g., EZB, DHIIts-). Furthermore, recent studies have
suggested the potential clinical value of circulating tumor DNA (ctDNA) for real-time response monitoring and early relapse detection during the disease course of DLBCL (148–150). ctDNA may also be useful at the time of diagnosis for patients without sufficient materials for genetic subtyping. Thus, further investigations are warranted for the integration of ctDNA into the management of DLBCL to facilitate precision medicine.

A draft of how transcriptionally, genetically, and microenvironmentally defined tumors relate to one another is shown in Fig. 4. However, a true integrated taxonomy will require careful examination of genomes, transcriptomes, and, ideally, epigenomes of a large representative group of DLBCL tumors. Related data resources are still being gathered in publicly available repositories. A final question is where the pathology boundaries of this new taxonomy should be placed. It should be noted that the genetics-based classification systems included HGBL-DH/TH with DLBCL morphology, without comprehensively identifying these tumors within their cohorts. The authors believe that it would be most beneficial to have a taxonomy that encompasses all aggressive B-cell lymphomas.

### Practical Diagnostic Considerations: Central Lab Structure, Future of Pathology Labs, Match Trials

It is anticipated that an integrative approach evaluating mutations, structural abnormalities, copy-number abnormalities, and gene expression will be required to dissect the complexity of DLBCL with sufficient granularity to truly inform personalized therapies. Although we are clearly still near the beginning of this journey, we now have the tools to begin to formulate a strategy to reach our destination. With well over 2,000 cases analyzed with publicly available data, we have a reasonable understanding of the landscape of the genetic alterations. Thus, we have sufficient data to make some decisions.

First, targeted sequencing strategies are sufficient for translating unbiased approaches used for discovery. With technical advances in library preparation and automation of analysis pipelines to perform assignments and identify actionable abnormalities, even integrative analysis involving the thousands of targets required will be possible on mid-level instrumentation within reach of large academic medical centers and regional reference laboratories. The simplified technical requirements will also enable reasonable turnaround times (measured in days, not weeks) required for the care of ill patients who need immediate therapy, because it is known that time to treatment is an important metric in achieving the best outcomes (151).

Second, for there to be any hope that we can advance the field with this new molecular taxonomy outlined above, a clinical-grade testing platform must be developed to support clinical trial strategies. National and international “match” style trials are planned in which all newly diagnosed DLBCL, NOS cases are enrolled and then placed in one of many treatment arms appropriate for the patient based on their molecular characterization. Such a test must be applicable in FFPE tissues, including small biopsies, because needle and endoscopic biopsies may represent an increasingly common specimen type. Knowing whether presence of normal accompanying tissues including non-neoplastic lymph node might affect the results must be part of the performance characteristics. It is not uncommon for a lymphoma diagnostic specimen to also contain other non-lymphomatous tissue in the same paraffin block when there is partial involvement or biopsy at an extranodal site.

This leads to the question: What is the clinical and/or academic incentive to develop such an assay for a pathologic

---

**Figure 4.** Biological links/integration between layers (gene-expression profiling, mutational profiling, tumor microenvironment biology). Schematic representation of the relationship between recently published transcriptomic [Ennishi and colleagues (81) and Sha and colleagues (82)] and genetic based [Chapuy and colleagues (98), Wright and colleagues (101), and Lacy and colleagues (100)] taxonomy, and tumor immune microenvironment. The suggested normal counterpart of the new taxonomic groups are described on the top. Potential therapeutic drug classes for genetic subtypes are represented (right). The molecular subtypes are labeled as per Wright and colleagues (101) where applicable for simplicity. The A53 (aka C2) molecular subtype is not shown. * align with C3 and BCL2; ** align with C1 and NOTCH2; *** align with C5 and MYD88.
entity? Of course, an academic laboratory might be interested in this task due to the institutional mission, challenge, and professional interests of the laboratory. However, most would not have the resources and/or regulatory expertise to rapidly develop such an assay platform that would, by necessity, be robust and transferable to several sites to have the redundancy needed to support a large-scale trial. A commercial laboratory or diagnostic company, perhaps in collaboration with an academic center, might be a reasonable option, but financial realities of a “return on investment” would add complexity. One strategy might be to build a targeted platform as noted above, but with one goal being to use a portion of the information from the output to replace a current standard-of-care assay. For example, if the integrative assay could replace FISH testing at a reasonable cost by identifying reliably HGBL-DH/TH within several days, then one could envision a path forward. Furthermore, it could lead to restricted use or removal of other ancillary tests such as IHC.

Assuming existence of a platform with adequate performance characteristics, a “match” style trial would be enabled. Central hematopathology laboratories would be able to confirm diagnosis of DLBCL, NOS by standard methods, ensure adequate tissue exists, and assign molecular taxonomy using the integrative testing (in the United States) as a Laboratory Developed Test or, if the hypothetical commercial entity strategy dictated, as an FDA-cleared in vitro diagnostic test.

In summary, the development of a universally accepted classification system would be of benefit to the field of DLBCL. In this review, we have outlined genomic, epigenetic, transcriptomic, proteomic, and microenvironment alterations that collectively could form the basis of a new taxonomy, with the potential to affect current treatment paradigms and provide a framework for future clinical trials. Ongoing studies are focusing on connecting these layers of alterations to form robust biological groups and on the implementation of clinically available assays with appropriate turnaround time to guide management. Such a molecular classification and additional pathobiological factors will help lay the groundwork for providing improved clinical benefit to patients with this disease.

Disclosure of Potential Conflicts of Interest

C. Steidl is a consultant at Seattle Genetics, Curis Inc., Roche, AbbVie, Juno Therapeutics, and Bayer and reports receiving commercial research grants from Bristol-Myers Squibb and Trillium Therapeutics Inc. D.W. Scott is a consultant at AbbVie, Janssen, and Celgene, reports receiving commercial research grants from Janssen, NanoString, and Roche/Genentech, and has ownership interest in patents for identifying molecular subtypes of lymphoma. One of these patents is licensed to NanoString. No potential conflicts of interest were disclosed by the other authors.

Acknowledgments

D. Ennishi, C. Steidl, and D.W. Scott acknowledge the support of the BC Cancer Foundation. C. Steidl is a recipient of a Michael Smith Foundation for Health Research Career Investigator award. D.W. Scott is a recipient of a Michael Smith Foundation for Health Research Health Professional-Investigator award.

Received February 14, 2020; revised May 4, 2020; accepted May 7, 2020; published first July 2, 2020.

REFERENCES


Downloaded from cancerdiscovery.aacrjournals.org on November 22, 2021. © 2020 American Association for Cancer Research.
Toward a New Molecular Taxonomy of DLBCL


Toward a New Molecular Taxonomy of Diffuse Large B-cell Lymphoma


<table>
<thead>
<tr>
<th>Updated version</th>
<th>Access the most recent version of this article at: doi:10.1158/2159-8290.CD-20-0174</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Cited articles</th>
<th>This article cites 149 articles, 56 of which you can access for free at: <a href="http://cancerdiscovery.aacrjournals.org/content/10/9/1267.full#ref-list-1">http://cancerdiscovery.aacrjournals.org/content/10/9/1267.full#ref-list-1</a></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>E-mail alerts</th>
<th>Sign up to receive free email-alerts related to this article or journal.</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Reprints and Subscriptions</th>
<th>To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at <a href="mailto:pubs@aacr.org">pubs@aacr.org</a>.</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Permissions</th>
<th>To request permission to re-use all or part of this article, use this link <a href="http://cancerdiscovery.aacrjournals.org/content/10/9/1267">http://cancerdiscovery.aacrjournals.org/content/10/9/1267</a>. Click on &quot;Request Permissions&quot; which will take you to the Copyright Clearance Center's (CCC) Rightslink site.</th>
</tr>
</thead>
</table>