Editorial note:

Diffuse large B-cell lymphoma (DLBCL) is the commonest subtype of non-Hodgkin lymphoma, accounting for about 30% to 40% of newly diagnosed non-Hodgkin lymphoma worldwide and in Hong Kong. In this topical update, Dr Choi Wai Lap reviews the genetic basis of DLBCL and its impact on classification, prognosis and even treatment. We welcome any feedback or suggestions. Please direct them to Dr. WS Wong (e-mail: sws_wong@yahoo.com.hk) of Education Committee, the Hong Kong College of Pathologists. Opinions expressed are those of the authors or named individuals, and are not necessarily those of the Hong Kong College of Pathologists.

Molecular Classification and Genetic Alterations of Diffuse Large B-cell Lymphoma

Dr CHOI Wai Lap Department of Clinical Pathology
Tuen Mun Hospital

Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLBCL</td>
<td>Diffuse Large B-cell Lymphoma</td>
</tr>
<tr>
<td>GEP</td>
<td>Gene expression profiling</td>
</tr>
<tr>
<td>COO</td>
<td>Cell of origin</td>
</tr>
<tr>
<td>IHC</td>
<td>Immunohistochemical</td>
</tr>
<tr>
<td>GCB</td>
<td>Germinal center B-cell like</td>
</tr>
<tr>
<td>ABC</td>
<td>Activated B-cell like</td>
</tr>
<tr>
<td>COO</td>
<td>Cell of origin</td>
</tr>
<tr>
<td>BCR</td>
<td>B-cell receptor</td>
</tr>
<tr>
<td>TLR</td>
<td>Toll-like receptor</td>
</tr>
</tbody>
</table>
Introduction

Diffuse large B-cell lymphoma (DLBCL) is the commonest subtype of non-Hodgkin lymphoma, accounting for about 30% to 40% of newly diagnosed non-Hodgkin lymphoma worldwide and in Hong Kong. DLBCL is heterogeneous in clinical presentation, morphology, immunophenotype, cytogenetics and prognosis. In the WHO Classification of Tumours of the Haematopoietic and Lymphoid Tissues published in 2008, several specific clinicopathological entities of DLBCL have been recognized, while leaving the rest to DLBCL, not otherwise specified, which is by far the most prevalent entity among the large B-cell lymphomas. In the following discussion, the term DLBCL will be used interchangeably with DLBCL, not otherwise specified.

Gene expression profiling and molecular classification of DLBCL

Gene expression profiling (GEP) is the simultaneous measurement of the transcription levels of thousands of genes to their corresponding messenger RNAs (mRNAs). GEP can be achieved by various technologies including DNA microarray, serial analysis of gene expression (SAGE) and most recently next generation sequencing (RNA-Seq).

Using DNA microarray technology on DLBCL, two distinct molecular subgroups were discovered based on the similarity of their gene expression pattern with a possible cell of origin (COO): the germinal centre B-cell-like (GCB-cell-like, or abbreviated as GCB) and the activated B-cell-like (ABC-like, or abbreviated as ABC). These molecular subgroups showed significantly different survival rates when treated with conventional cyclophosphamide, doxorubicin, vincristine and prednisolone (CHOP) chemotherapeutic regimen.

In order to facilitate the classification process by GEP, a statistical method based on Bayes' rule was then developed to estimate the probability of membership. Cases that had less than 90% probability of belonging to either molecular subgroup would be regarded as ‘unclassified’ molecularly. This statistical method laid the foundation of subsequent GEP studies using DNA microarray technology, which further confirmed the prognostication of the GCB and ABC molecular subgroups, even with the addition of rituximab to the CHOP (R-CHOP) regimen.

Immunohistochemical (IHC) algorithms to approximate the COO classification of DLBCL

Although the COO molecular classification of DLBCL is a powerful prognosticator, it is fairly difficult to apply in routine anatomical pathology practice. This can be attributed to the lack of expertise and facilities for DNA microarray in routine diagnostic laboratories, the relatively high costs of these DNA microarray chips, and the requirement of snap-frozen diagnostic materials.

Hans et al., therefore, developed an IHC algorithm based on GEP results using three IHC stains: CD10, BCL6 and MUM1/IRF4 in 2004. While the Hans algorithm is relatively simple and has been used most extensively in clinical trials, it has been criticized for basing on treatment results of CHOP instead of R-CHOP, which have become the standard regimen for DLBCL.

Other IHC algorithms using different IHC stain combinations and different cutoff levels of the various markers were thus developed in patients treated with R-CHOP, including the Nyman, Choi, Tally and Visco-Young algorithms. The first three of these and the Hans algorithm were compared head to head, and correlated with DNA microarray data of DLBCL samples from patients treated with R-CHOP. In that comparison, the Tally, Choi and Hans algorithms all showed high sensitivity, specificity and predictive values. The three algorithms also showed high concordance with DNA microarray results at 93%, 87% and 86%, respectively. However, the Nyman algorithm had a low sensitivity and negative predictive value despite showing a fairly high concordance (81%).

The most recent Visco-Young algorithm seemed to have a high concordance rate to DNA microarray data at 92.6%, but the sensitivity,
specificity and predictive values were uncertain.\textsuperscript{11} Besides, the hazard ratios of the COO molecular classification in univariate and multivariate analyses appeared to be less marked compared with the Nyman, Choi and Tally algorithms.\textsuperscript{10, 11}

The long list of IHC algorithms may seem puzzling and difficult for the anatomical pathologists to choose, but it seems that the Hans, Choi and Tally algorithms are among the most robust. Individual laboratory may choose among these three algorithms based on the availability of the IHC stains and familiarity to these algorithms.

**Other IHC markers useful for prognostication of DLBCL treated with R-CHOP**

In recent years, additional markers for further refinement of the prognostication within the GCB and ABC DLBCL subgroups emerged. Meyer et al, found that any SPARC expression in DLBCL microenvironment was associated with a significantly longer overall survival, and patients with high SPARC positivity in the microenvironment also had a significantly longer event-free survival. These survival differences were mainly due to the prognostic effect of SPARC positive cells in the ABC subgroup, but not in the GCB subgroup.\textsuperscript{12} Perry et al. further combined the COO classification, SPARC expression and microvessel density to formulate a new biologic prognostic model for all DLBCL, with the non-GCB phenotype, SPARC expression <5% and microvessel density quartile 4 being adverse factors for outcome.\textsuperscript{13}

BCL2 expression predicted for inferior survival in patients with GCB DLBCL but not ABC DLBCL when R-CHOP was used.\textsuperscript{14} Since the high BCL2 expression in GCB DLBCL is due to the t(14;18) translocation,\textsuperscript{13} it is not surprising that patients with GCB DLBCL harbouring BCL2 translocations would have poor outcomes.\textsuperscript{15} Similar to the double-hit B-cell non-Hodgkin lymphomas with both MYC and BCL2 translocations, recent studies also found that MYC and BCL2 protein co-expression in DLBCL predicted for poor outcomes.\textsuperscript{16-18} and it was observed that MYC/BCL2 co-expression was commoner in ABC DLBCL,\textsuperscript{17, 18} which appeared to contribute significantly to their gene expression signature and adverse outcome of ABC DLBCL.\textsuperscript{17} Another study also confirmed that high MYC expression, high BCL2 expression and low BCL6 expression were independent adverse factors for survival of DLBCL patients, and an IHC score based on the expression of these proteins and the MYC translocation status was also predictive for survival.\textsuperscript{19}

The expression of CD30 was recently observed to predict for superior 5-year overall and progression-free survival, with this favourable outcome maintained in both the GCB and ABC subgroups.\textsuperscript{20}

Among these prognostic markers discovered in the last few years, it seems that the combination of MYC and BCL2 expression is relatively easy to apply and has the potential of pinpointing about 30% of the DLBCL cases with incurable disease and poor survival outcomes when treated with the current standard R-CHOP regimen. Additionally, this segregation may be independent of the COO classification. Future clinical trials applying these newer markers or marker combinations are awaited.

**Genetic alterations of GCB DLBCL**

In the recent few years, global screening of the genomes and transcriptomes of DLBCL using next generation sequencing (NGS) have rapidly broadened our understanding of the pathogenesis of the COO subgroup of DLBCL. These studies showed that certain genetic alterations are associated with individual molecular subgroups, shedding light on the oncogenic pathways that are essential for the GCB and ABC DLBCL.

As aforementioned, the t(14;18) is found in about 35% GCB DLBCL, leading to over-expression of BCL2, which is a key anti-apoptotic protein in GCB DLBCL.\textsuperscript{21} The over-expression is due to the juxtaposition of BCL2 to the potent regulatory elements of the immunoglobulin locus in the t(14;18), as well as by disrupting the suppression by BCL6.\textsuperscript{22, 23} Other mechanisms of BCL2 up-regulation found in GCB DLBCL without a t(14;18), include deregulation of Miz1, aberrant
somatic hypermutation of BCL2 promoter sequences, and mutations in the BCL2 coding sequence.\textsuperscript{22,24}

About 15\% of DLBCL harbour the t(8;14) which brings the MYC gene under the control of the immunoglobulin promoter, but this translocation is found in both COO subgroups.\textsuperscript{25,26} Amplifications of the microRNA (miR) 17-92 cluster on chromosome 13q are found in 12\% of GCB DLBCL.\textsuperscript{27} This cluster acts as a potential oncogene and accelerates MYC-induced lymphomagenesis and enhances oncogenesis by increasing proliferation and survival via inhibition of the tumor suppressor PTEN and thus activation of AKT.\textsuperscript{27-30} Alternative mechanisms of inhibition of PTEN include deletions of PTEN on chromosome 10q, which are found in around 11\% of DLBCL with a preference in GCB DLBCL, and are mutually exclusive to the miR-17-92 amplifications.\textsuperscript{27,30} Similarly, AKT can also be activated by phosphatidylinositol 3 kinase (PI3K) with activating mutations, which are found in about 8\% of DLBCL, and are mutually exclusive to the loss of PTEN.\textsuperscript{31} GNA13 mutations are present in up to 25\% of GCB DLBCL.\textsuperscript{32,33} GNA13 encodes Go13, a G-protein that increases Rho-mediated motility, and Rho-mediated increased PTEN activity and potent AKT inhibition.\textsuperscript{34,35} GNA13 mutations thus probably contribute to the spread of the neoplastic cells outside of the germinal centre niche and the deregulation of the AKT/mTOR pathway.

The chromatin modifying genes seem to play key roles in GCB DLBCL pathogenesis, as mutations of a number of these genes are preferentially found in GCB DLBCL. Recurrent activating mutations affecting residue Tyr641 of the polycomb-group histone methyltransferase EZH2 have been found in 21.7\% of GCB DLBCL.\textsuperscript{36} Mutated EZH2 trimethylates Lys27 of histone H3 and represses gene expression more vigorously by increased affinity to the substrate.\textsuperscript{36-38} Another histone methyltransferase MLL2 is mutated in around 30\% of DLBCL, with the majority of the mutations being inactivating ones.\textsuperscript{32,36} Myocyte enhancer factor 2B (MEF2B) is a DNA binding protein that cooperates with histone modifying enzymes to regulate gene expression, and is mutated in about 9\% of DLBCL.\textsuperscript{32,36} In one study, monoallelic deletions and inactivating mutations of CREBBP and EP300 were found in nearly 39\% of GCB DLBCL and only 17\% of ABC DLBCL.\textsuperscript{39} These two acetyltransferases are transcriptional co-activators in multiple signaling pathways. For instance, CREBBP and EP300 acetylate and inactivate BCL6 by disrupting the recruitment of histone deacetylases and thus hindering the ability of BCL6 to repress transcription.\textsuperscript{40,41}

Mutations of p53 are more frequent in GCB DLBCL than in ABC DLBCL (38\% vs 18\%). Most of the mutations are missense or nonsense inactivating mutations, and mutated p53 confer worse survivals within both the GCB and ABC subgroups of DLBCL.\textsuperscript{42}

**Genetic alterations of ABC DLBCL**

Constitutive activation of NF-κB plays an important role in ABC DLBCL survival,\textsuperscript{43,44} which can be due to several distinct genetic alterations affecting both positive and negative regulators of the pathway. These genetic alterations are thus predominantly seen in ABC DLBCL. \textit{TNFAIP3} encodes for the negative NF-κB regulator A20 and is inactivated in about 30\% of cases of DLBCL by biallelic mutations and/or deletions,\textsuperscript{45} which are almost exclusively found in ABC DLBCL and rare in GCB DLBCL.\textsuperscript{45-48}

The B-cell receptor (BCR) induced activation of NF-κB requires CARD11, a scaffold protein that mediates the activation of IkB kinase β.\textsuperscript{49} Mutations of CARD11 are found in around 10\% of ABC DLBCL and a smaller subset of GCB DLBCL.\textsuperscript{45} These mutants probably activate the NF-κB pathway in the absence of BCR signaling and lead to constitutive NF-κB activation.\textsuperscript{50} CD79A and CD79B are proximal BCR subunits and were found to be mutated in around 20\% of ABC DLBCL.\textsuperscript{43} The mutations increase surface BCR expression and abrogate a feedback inhibitor of BCR signaling,\textsuperscript{43} leading to chronic active BCR signaling and consequent activation of the NF-κB pathway.

NF-κB is also activated after stimulation of Toll-like receptors (TLR) and interleukin (IL)-1 and 8
receptors. MYD88 functions as a signaling adaptor protein and assembles a protein complex that activates NF-κB following TLRs, and IL-1 and IL-8 receptor stimulations. MYD88 mutations are found in about 30% of ABC DLBCL. They all lead to the same amino acid substitution (L265P) in the TLR/IL-1 receptor domain. The MYD88 L265P mutant may promote cell survival by activating NF-κB signaling.

Normal GC B-cell development into plasma cells requires down-regulation of BCL6 by NF-κB, and expression of IRF4, BLIMP1 and XBP1. In ABC DLBCL, several mechanisms that can interrupt this GC B-cell to plasma cell transition occur. Chromosomal translocations of BCL6 on chromosome 3q27 are detected in about 35% of DLBCL cases, and are twice commoner in ABC DLBCL than in GCB DLBCL. The commonest translocations involve the immunoglobulin heavy-chain promoter, resulting in constitutive expression of BCL6. Deregulated expression of BCL6 is thought to result in differentiation blockage, reduced p53-mediated apoptotic response to DNA damage, and enhanced proliferation.

The PRDM1 gene on chromosome 6q21 encodes for BLIMP1, a zinc finger transcriptional repressor that represses genes involved in BCR signaling and proliferation, and acts as a tumour suppressor. Inactivating mutations and deletions of PRDM1 are found in up to 30% of ABC DLBCL. Additionally, PRDM1 can be inactivated by transcriptional repression through constitutively active BCL6, as it is the case in patients carrying BCL6 translocations. Indeed, chromosomal translocations of BCL6 and genetic alterations affecting PRDM1 are mutually exclusive.

Amplifications of chromosome 18q21 and gain of BCL2 gene are observed in 21% and 46% of ABC DLBCL, respectively. Together with NF-kB transactivation, they likely contribute to the BCL2 over-expression observed in 59% of ABC DLBCL. Amplifications of the telomere of chromosome 19q are found in about 25% of ABC DLBCL. SPIB, an ETS family transcription factor, may be a possible candidate gene on chromosome 19q essential for ABC DLBCL as down-regulation of SPIB was toxic to ABC DLBCL cell lines.

Homozygous or heterozygous deletions of the INK4a/ARF locus are observed in about 30% of ABC DLBCL. P16INK4a and p14ARF regulate the pRB and the p53 tumour suppressors, and inactivation of the p53 pathway via INK4a/ARF is found to inhibit apoptosis.

Conclusions

The emergence of technologies like DNA microarray and next generation sequencing has unraveled the molecular and genetic basis of DLBCL. The much better understanding of the pathogenetic pathways and the potential targets along these pathways has gradually opened the door for potential personalized medicine or tailored therapy for DLBCL patients in future. Pathologists should be prepared to provide such information of molecular classification and genetic alterations of DLBCL when such a day comes.

References


