Diffuse large B-cell lymphoma (DLBCL) displays marked clinical, pathological, and genetic heterogeneity. With currently frontline immunotherapy (RCHOP) only about 40% of patients are cured, with most relapses occurring within the first 2-3 years. Patients are currently risk-stratified primarily based on clinical features where the inclusion of molecular biomarkers into risk scores could impact the potential to identify those patients most likely to have refractory disease or have early relapse. Various cytogenetic studies have revealed the prognostic significance of genomic gains in DLBCL, but their lack of utility and reproducibility across clinical databases probably results from only different patient populations, but also the use of disparate platforms and analytical methods. The goal of the present study was to use a common and robust approach across different clinical databases, to identify common regions of copy number aberrations (CNAs) in DLBCL and genomic loci with robust prognostic value in DLBCL.

**Results**

**In-silico Identification of Overlapping CNAs in DLBCL**

**Figure 1**—After applying the GISTIC algorithm to three available in silico datasets, overlapping CNAs (MCR) were identified within either large chromosomal regions or peak regions in at least 2 of the 3 datasets.

**Aberrations (CNAs) and Calling Criteria**

<table>
<thead>
<tr>
<th>Chr.</th>
<th>Location</th>
<th>size (kb)</th>
<th>CN</th>
<th>MCR</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>22,622,538</td>
<td>10,867,785</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>31,536,224</td>
<td>10,083,152</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>41,200,100</td>
<td>10,083,152</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>51,944,706</td>
<td>10,083,152</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Integrated Analysis of CNA and Expression**

Raw expression data for 162 specimens of IS-172 were RMA-normalized and regional genes were correlated with scored CNAs by univariate t-tests. Genes exhibiting a 1.2 fold change and P < 0.05 after FDR correction were analyzed.

The expression of a total of 596 unique genes correlated with at least one aberration. Of these, significant gene expression correlation was:
- Not found for 17 abs
- Was found for 6 abs with peaks, but none mapped to the peak
- Was found for 11 abs with peaks, and mapped to the peak (see below)
- Was found for 6 abs with peaks, and mapped in MCR (see below)

**Pathway Analysis**

Ingenuity pathway analysis was performed on the total 596 unique genes which correlated with a CNA. Five canonical pathways were significantly enriched (FDR < 0.05).

**Method 1- Complexity based on total number of detectable CNAs within RCHOP and RCHOP-like treated samples**

- **Figure 2**—Concordance between the IS-172 and IS-169 datasets using GISTIC.
- **Figure 3**—Aberrations previously reported to be enriched in cell-of-origin (COO) subtypes. (A) were similarly tested for association after applying the newly established scoring criteria.

**Method 2- Complexity based on having an alteration within the CDKN2A-TP53-RB-E2F axis (Ref. 2)**

- **Figure 4**—Kaplan-Meier plot of dataset IS-172 and IS-169N=70) which were all RCHOP treated patients. Samples were scored as complex if any one of alterations at 2p, 3q, 5q, 10p, 11q, 14q, 17p, 18q, 19q, and 20q were present.
- **Figure 5**—Kaplan-Meier plot of dataset IS-172 and IS-169N=70) which were all RCHOP treated patients. Samples were scored as complex if any one of alterations at 2p, 3q, 5q, 10p, 11q, 14q, 17p, 18q, 19q, and 20q were present.

**Correlation of Clinically Relevant Aberrations**

Loss of 17p
- **Figure 6**—Loss of 17p, Gain of 16q

**References**