Whole-exome analysis reveals novel somatic genomic alterations associated with cell of origin in diffuse large B-cell lymphoma

Diffuse large B-cell lymphoma (DLBCL) is the most common lymphoma, with an estimated 27,650 cases diagnosed in 2016. Gene expression profiling has identified three subtypes of DLBCL: germinal center B-Cell like (GCB), activated B-cell like (ABC) and primary mediastinal B-cell lymphoma (PMBL). This molecular heterogeneity is indicative of a unique cell of origin (COO) giving rise to each subtype and is associated with clinical outcome; with treatment, leading to remission in ~80% of GCB patients but only ~50% in ABC patients. The poor prognosis of ABC-DLBCL is due, at least in part, to a high risk of early relapse in COO and therapeutic resistance. Immunochemotherapy-treate DLBCL patients.7 Building on this observation, we used the whole-exome sequencing (WES) method in DLBCL. In an effort to use genomics as a clinical predictor of therapeutic response, we used the whole-exome sequencing in 58 newly diagnosed frozen DLBCL tumors and paired germline DNA was performed at the Broad Institute and somatic mutations and exon-level copy number alterations (CNAs) were identified as previously described.7,10 A CNA was called for each gene: 10q11.21-10q24.23 loss (PTEN), 4q12-4q35.2 loss (IGL), 7q11.1-7q36.3 gain (CDK14), 3q12.1-3q29 loss (TP53), 2p13-2p12 gain (REL), 6q21 loss (PRDM1), 9p21 loss (CDKN2B), 18q21.33 gain (BCL2) and 9p24.1 gain (CD274). We estimated measures of association using odds ratios and report the association of genomic variants with COO using a χ² test. For this exploratory study, we used a nominal level of statistical significance (P < 0.05), and we did not adjust for multiple testing. While all mutations and CNAs identified in the 58 cases were analyzed for their association with COO, the variants reported in this study include (1) all mutations and CNAs that had an association with COO (P < 0.05), and (2) those previously identified as drivers of DLBCL.10-14 Patient characteristics are described in Table 1.

A total of 27 patients were classified as GCB, 26 as ABC, and five were unclassified (Figure 1a). In total, 37 genomic abnormalities were reported for their association with either GCB (Figure 1a, red boxes) or ABC (green boxes), or neither (blue boxes). We find that mutations in MYC, BCL2, TNFRSF14, GNA13 and FAT3 significantly associate (P < 0.05) with the GCB subtype, largely agree with previous reports.15,16 Mutations in P2RYB, EZH2 and FOXO1 have also been reported as GCB driver mutations and we find that mutations in these genes are restricted to GCB. MYC double-hits (MYC-DH, gene rearrangements of MYC with BCL2 and/or BCL6) were present only in GCB patients. Mutations in MYD88 associated (P = 0.03) with ABC in our dataset, while mutations in CD79B and TNFAIP3, both known to associate with ABC,17,16 trend towards association. The remaining mutations reported did not strongly associate with either ABC or GCB, suggesting that there is common biology underlying both subgroups.

In addition to mutational patterns, we identified several CNAs across both groups. Chromosomal losses (P < 0.05) associated with GCB DLBCL were found at chromosomes 10q11.21-10q24.23 and 4q12-4q35.2, with a gain at 7q11.1-1q36.3 towards GCB (P = 0.07). No copy-number variation was observed to directly associate with ABC patients however, a loss at 9p21 and gains at 18q21.33 trended towards the ABC subtype, supporting previous reports.18 A gain in 2p13-2p12 has been reported as being specific for GCB,18 but our data identify it occurring in both subtypes.

---

**Table 1. DLBCL patient characteristics**

| Characteristic | ABC (N = 26) | GCB (N = 27) | Unclassified (N = 5) | P-value |
|---------------|-------------|-------------|---------------------|---------|
| Diagnosis age | 60 (28-84) | 66 (26-88) | 68 (62-77); 0.2061 |
| Median (range): IQR | 55–71 | 60–75 | 66–70 |
| Age > 60 | 13 (50.0%) | 7 (26.9%) | 5 (100.0%); 0.0681 |
| Male | 17 (65.4%) | 15 (55.6%) | 3 (60.0%); 0.4646 |
| PS 2+ | 3 (11.5%) | 6 (22.2%) | 0 (0.0%); 0.3064 |
| Ann arbor stage III-IV | 23 (88.5%) | 16 (59.3%) | 2 (40.0%); 0.0159 |
| LDH > ULN | 17 (65.4%) | 16 (61.5%) | 3 (60.0%); 0.7734 |
| IPI | 0–1 | 4 (15.4%) | 8 (29.6%) | 2 (40.0%); 0.4829 |
| 2 | 10 (38.5%) | 6 (22.2%) | 1 (20.0%) |
| 3 | 9 (34.6%) | 9 (33.3%) | 1 (20.0%) |
| 4 or 5 | 3 (11.5%) | 4 (14.8%) | 1 (20.0%) |

**Abbreviations:** ABC, activated B-Cell like; GCB, germinal center B-cell like

---

**LETTER TO THE EDITOR**

Diffuse large B-cell lymphoma (DLBCL) is the most common lymphoma, with an estimated 27,650 cases diagnosed in 2016. Gene expression profiling has identified three subtypes of DLBCL: germinal center B-Cell like (GCB), activated B-cell like (ABC) and primary mediastinal B-cell lymphoma (PMBL). This molecular heterogeneity is indicative of a unique cell of origin (COO) giving rise to each subtype and is associated with clinical outcome, with treatment, leading to remission in ~80% of GCB patients but only ~50% in ABC patients. The poor prognosis of ABC-DLBCL is due, at least in part, to a high risk of early relapse in COO and therapeutic resistance. Immunochemotherapy-treated DLBCL patients. Building on this observation, we used the whole-exome sequencing (WES) method in 58 newly diagnosed frozen DLBCL tumors and paired germline DNA was performed at the Broad Institute and somatic mutations and exon-level copy number alterations (CNAs) were identified as previously described.7,10 A CNA was called for each gene: 10q11.21-10q24.23 loss (PTEN), 4q12-4q35.2 loss (IGL), 7q11.1-7q36.3 gain (CDK14), 3q12.1-3q29 loss (TP53), 2p13-2p12 gain (REL), 6q21 loss (PRDM1), 9p21 loss (CDKN2B), 18q21.33 gain (BCL2) and 9p24.1 gain (CD274). We estimated measures of association using odds ratios and report the association of genomic variants with COO using a χ² test. For this exploratory study, we used a nominal level of statistical significance (P < 0.05), and we did not adjust for multiple testing. While all mutations and CNAs identified in the 58 cases were analyzed for their association with COO, the variants reported in this study include (1) all mutations and CNAs that had an association with COO (P < 0.05), and (2) those previously identified as drivers of DLBCL.10-14 Patient characteristics are described in Table 1.

A total of 27 patients were classified as GCB, 26 as ABC, and five were unclassified (Figure 1a). In total, 37 genomic abnormalities were reported for their association with either GCB (Figure 1a, red boxes) or ABC (green boxes), or neither (blue boxes). We find that mutations in MYC, BCL2, TNFRSF14, GNA13 and FAT3 significantly associate (P < 0.05) with the GCB subtype, largely agree with previous reports.15,16 Mutations in P2RYB, EZH2 and FOXO1 have also been reported as GCB driver mutations and we find that mutations in these genes are restricted to GCB. MYC double-hits (MYC-DH, gene rearrangements of MYC with BCL2 and/or BCL6) were present only in GCB patients. Mutations in MYD88 associated (P = 0.03) with ABC in our dataset, while mutations in CD79B and TNFAIP3, both known to associate with ABC,17,16 trend towards association. The remaining mutations reported did not strongly associate with either ABC or GCB, suggesting that there is common biology underlying both subgroups.

In addition to mutational patterns, we identified several CNAs across both groups. Chromosomal losses (P < 0.05) associated with GCB DLBCL were found at chromosomes 10q11.21-10q24.23 and 4q12-4q35.2, with a gain at 7q11.1-1q36.3 towards GCB (P = 0.07). No copy-number variation was observed to directly associate with ABC patients however, a loss at 9p21 and gains at 18q21.33 trended towards the ABC subtype, supporting previous reports.18 A gain in 2p13-2p12 has been reported as being specific for GCB,18 but our data identify it occurring in both subtypes.
Additionally, losses at 3q12.1-3q29 and 6q21 occurred in both subtypes.

In an effort to further understand genomic differences between DLBCL subtypes, we evaluated the relative percentage of each genomic instability (Figure 1b). Of the reported genomic alterations, 7/37 (18.9%) were only observed in GCB patients, whereas 2 out of 37 (5.4%) were specific to the ABC subtype. The majority (28 out of 37, 75.7%) overlapped between ABC and GCB, potentially indicative of similar biology between subtypes.

Patients diagnosed with ABC DLBCL have been reported to have a worse clinical prognosis than GCB patients, likely due to chronic B-cell receptor signaling and constitutive activation of NF-κB from acquired mutations in upstream genes such as CD79B, CARD11 and MYD88. Interestingly, we identified seven patients (Figure 1a, No. 47–53) within our ABC cluster that do not exhibit any of the 37 genomic alterations reported here that require further genomic study to better resolve the predictive survival analysis of DLBCL patients. Four additional cases (Figure 1a, No. 43–46) had only one genomic alteration. This may be indicative of other, yet to be identified genomic instabilities that may contribute to poor clinical outcome in these cases. Taken together, this analysis has further characterized the genetic profile of each COO subtype and has identified novel GCB CNAs that may contain candidate genes that provide insight on tumor biology and offer potential targets for therapy. Collectively, these data provide insight on the genetic heterogeneity of DLBCL and identify genomic variants that can inform subtype-specific therapy.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

ACKNOWLEDGEMENTS
This work was supported in part by the National Institutes of Health (P50 CA097274 to JRC and AJN R01 CA200703 to JRC), and the Predolin Foundation (to AJN).
thank the Mayo Clinic Center for Individualized Medicine for their support of the Bioinformatics Program and Biomarker Discovery Program.

**AUTHOR CONTRIBUTIONS**

BAM, KW, JRC and AJN designed, analyzed and interpreted the data, and drafted the paper. YA, MJM, MM, ZY, performed the experiments, analyzed the data and edited the manuscript. SLS, GSN, SMA, TEW, AF, LR and BKL, collected patient specimens and edited the manuscript. SLS, GSN, SMA, TEW, AF, LR and BKL, collected patient specimens and edited the manuscript. SLS, GSN, SMA, TEW, AF, LR and BKL, collected patient specimens and edited the manuscript. SLS, GSN, SMA, TEW, AF, LR and BKL, collected patient specimens and edited the manuscript.

**REFERENCES**

1. Terada LR, DeSantis CE, Cerhan JR, Morton LM, Jemal A, Flowers CR. 2016 US lymphoid malignancy statistics by World Health Organization subtypes. CA Cancer J Clin 2016; 66: 443–459.

2. Lenz G, Wright GW, Emre NC, Kohlihammer H, Dave SS, Davis RE et al. Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. Proc Natl Acad Sci USA 2008; 105: 13520–13525.

3. Pon JR, Marra MA. Clinical impact of molecular features in diffuse large B-cell lymphoma and follicular lymphoma. Blood 2016; 127: 181–186.

4. Young RM, Shaffer AL 3rd, Phelan JD, Staudt LM. B-cell receptor signaling in diffuse large B-cell lymphoma. Semin Hematol 2015; 52: 77–85.

5. Turturro F. Constitutive NF-kB activation underlines major mechanism of drug resistance in relapsed refractory diffuse large B cell lymphoma. BioMed Res Int 2015; 2015: 484537.

6. Choi WW, Weisenburger DD, Greiner TC, Pietsch MA, Banham AH, Delabie J et al. A new immunostain algorithm classifies diffuse large B-cell lymphoma into molecular subtypes with high accuracy. Clin Cancer Res 2009; 15: 5494–5502.

7. Novak AJ, Asmann YW, Maurer M, Wang C, Slager SL, Hodge LS et al. Whole-exome analysis reveals novel somatic genomic alterations associated with outcome in immunotherapy-treated diffuse large B-cell lymphoma. Blood Cancer J 2015; 5: e346.

8. Wright G, Tan B, Rosenwald A, Hurt EH, Wiensner A, Staudt LM. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. Proc Natl Acad Sci 20032003; 100: 9991–9996.

9. Scott DW, Wright GW, Williams PM, Lih CJ, Walsh W, Jaffe ES et al. Determining cell-of-origin subtypes of diffuse large B-cell lymphoma using gene expression in formalin-fixed paraffin-embedded tissue. Blood 2014; 123: 1214–1217.

10. Lohr JG, Stojanov P, Lawrence MS, Auclair D, Chapuy B, Sougnez C et al. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. Proc Natl Acad Sci USA 2012; 109: 3879–3884.

11. de Miranda NF, Georgiou K, Chen L, Wu C, Gao Z, Zaravinos A et al. Exome sequencing reveals novel mutation targets in diffuse large B-cell lymphomas derived from Chinese patients. Blood 2014; 124: 2544–2553.

12. Monti S, Chapuy B, Takeyama K, Rodig SJ, Hao Y, Yeda KT et al. Integrative analysis reveals an outcome-associated and targetable pattern of p53 and cell cycle deregulation in diffuse large B cell lymphoma. Cancer Cell 2012; 22: 359–372.

13. Morin RD, Mendez-Lago M, Mungall AJ, Goya R, Mungall KL, Corbett RD et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. Nature 2011; 476: 298–303.

14. Pasqualucci L, Trifonov V, Fabbri G, Ma J, Rossi D, Chiarenza A et al. Analysis of the coding genome of diffuse large B-cell lymphoma. Nat Genet 2011; 43: 830–837.

15. Dubois S, Vially P-J, Mareschal S, Bohers E, Bertrand P, Ruminy P et al. Next generation sequencing in diffuse large B cell lymphoma highlights molecular divergence and therapeutic opportunities: a LYSAn Study. Clin Cancer Res 2016; 22: 2919–2928.

16. de Miranda NFCC, Georgiou K, Chen L, Wu C, Gao Z, Zaravinos A et al. Exome sequencing reveals novel mutation targets in diffuse large B-cell lymphomas derived from Chinese patients. Blood 2014; 124: 2544–2553.

17. Davis RE, Ngo VN, Lenz G, Tolar P, Young RM, Romesser PE et al. Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. Nature 463: 88–92.

18. Sebastián E, Alcoceba M, Martín-García D, Blanco Ó, Sanchez-Barba M, Balanzategui A et al. High-resolution copy number analysis of paired normal-tumor samples from diffuse large B cell lymphoma. Ann Hematol 2015; 95: 253–262.