

Next-Generation Sequencing Highlights of Diffuse Large B-cell Lymphoma in a Tertiary Care Hospital in North India

Review began 08/10/2022

Review ended 08/18/2022

Published 08/21/2022

© Copyright 2022

Mamgain et al. This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Garima Mamgain¹, Manisha Naithani², Priyanka Patra², Mukesh Mamgain¹, Sikha Morang³, Jhasketan Nayak¹, Karthik Kumar¹, Shashikant Singh¹, Anamika Bakliwal¹, Ashok Rajoreya¹, Sudeep Vaniyath¹, Debranjani Chattopadhyay¹, Rituparna Chetia¹, Arvind Gupta⁴, Gaurav Dhingra¹, Deepak Sundriyal¹, Uttam Kumar Nath¹

1. Department of Medical Oncology/Haematology, All India Institute of Medical Sciences, Rishikesh, Rishikesh, IND 2. Department of Biochemistry, All India Institute of Medical Sciences, Rishikesh, Rishikesh, IND 3. Department of Pharmacology, All India Institute of Medical Sciences, Rishikesh, Rishikesh, IND 4. Department of Pathology, All India Institute of Medical Sciences, Rishikesh, Rishikesh, IND

Corresponding author: Manisha Naithani, naithanimanisha@gmail.com

Abstract

Introduction: Next-generation sequencing (NGS) elucidates the diffuse large B-cell lymphoma (DLBCL) genetic characteristics by finding recurrent and novel somatic mutations. This observational study attempted to create an NGS panel with a focus on identifying novel somatic mutations which could have potential clinical and therapeutic implications. This panel was created to look for mutations in 133 genes chosen on basis of a literature review and it was used to sequence the tumor DNA of 20 DLBCL patients after a centralized histopathologic review.

Methods: The study included 20 patients having DLBCL. The quality and quantity of tumor cells were accessed by H&E staining and correlated with histopathology and Immunohistochemistry (IHC) status. Patients were grouped as ABC (activated B-cell), PMBL (primary mediastinal large B-cell lymphoma), and other or unclassified subtypes. The lymphoma panel of 133 was designed on targeted sequencing of multiple genes for the coding regions through NGS. The libraries were prepared and sequenced using the Illumina platform. The alignment of obtained sequences was performed using Burrows-Wheeler Aligner and identification of somatic mutations was done using LoFreq (version 2) variant caller. The mutations were annotated using an annotation pipeline (VariMAT). Previously published literature and databases were used for the annotation of clinically relevant mutations. The common variants were filtered for reporting based on the presence in various population databases (1000G, ExAC, EVS, 1000Japanese, dbSNP, UK10K, MedVarDb). A custom read-depth-based algorithm was used to determine CNV (Copy Number Variants) from targeted sequencing experiments. Rare CNVs were detected using a comparison of the test data read-depths with the matched reference dataset. Reportable mutations were prioritized and prepared based on AMP-ASCO-CAP (Association for Molecular Pathology-American Society of Clinical Oncology-College of American Pathologists), WHO guidelines, and also based on annotation metrics from OncoMD (a knowledge base of genomic alterations).

Results: The informativity of the panel was 95 percent. *NOTCH 1* was the most frequently mutated gene in 16.1% of patients followed by 12.9% who had *ARID1A* mutations. *MYD88* and *TP53* mutations were detected in 9.6% of the patient while 6.4% of patients had *CSF3R* mutations. *NOTCH 1* and *TP 53* are the most frequently reported gene in the middle age group (40-60). Mutation in *MYD88* is reported in every age group. *MYD88* (51%) is the most common mutation in ABC subtypes of DLBCL, followed by *NOTCH 1* (44%) and *SOC1 1* (33%) according to our findings. *NOTCH 1* mutations are frequent in ABC and PMBL subtypes. Closer investigation reveals missense mutation is the most frequent mutation observed in the total cohort targeting 68.4% followed by frameshift deletion reported in 26.3%. Six novel variants have been discovered in this study.

Conclusions: This study demonstrates the high yield of information in DLBCL using the NGS Lymphoma panel. Results also highlight the molecular heterogeneity of DLBCL subtypes which indicates the need for further studies to make the results of the NGS more clinically relevant.

Categories: Oncology, Healthcare Technology, Hematology

Keywords: informativity, lymphoma panel, genetic mutation, next-generation sequencing, diffuse large b cell lymphoma

Introduction

Among adult lymphoma, diffuse large B-cell lymphoma (DLBCL) is the commonest with 30-40% of global incidence [1]. DLBCL's molecular heterogeneity understanding has been much advanced due to gene expression profiling (GEP) studies allowing classification as germinal center B-cell-like (GCB), activated B-

How to cite this article

Mamgain G, Naithani M, Patra P, et al. (August 21, 2022) Next-Generation Sequencing Highlights of Diffuse Large B-cell Lymphoma in a Tertiary Care Hospital in North India. Cureus 14(8): e28241. DOI 10.7759/cureus.28241

cell-like (ABC), and primary mediastinal B-cell lymphoma (PMBL) also referred to as cell-of-origin (COO) classification [2,3]. The hallmark of the GCB subtype is the loss of PTEN (a tumor suppressor) and chromosomal translocation t(14;18)(q32;q21) juxtaposing the *BCL2* gene [4,5]. ABC subtype prominently reveals deletion of the INK4/ARF locus, *BCL2* amplification, and 3q27 rearrangements involving *BCL6* [6,7]. PMBL has classical Hodgkin lymphoma (CHL) like molecular hallmarks including *JAK2* amplifications and *SOCS1* deletions [8,9].

The clinical presentation, immunochemotherapy response, and the outcome of these subtypes differ. The ABC subtype has the worst prognosis whereas the GCB subtype has the comparatively best responses [2,7]. The availability of targeted medicines has made the identification of each molecular subtype essential so that each patient can receive the best possible therapy.

Despite recent breakthroughs in bench-to-bedside translation, the GEP-based classification has little influence on clinical practice due to difficulty in applying it for routine diagnosis due to molecular heterogeneity of disease [10-12]. The answer to this has emerged as high-throughput DNA sequencing in form of next-generation sequencing (NGS) allowing insights and detailed genomic characterization of DLBCL by identifying recurrent single nucleotide variants (SNV) prevalent in specific subtypes [13-22]. Despite the diversity, SNVs are particularly significant since their presence in actionable targets is associated with therapy response.

Moreover, notwithstanding pathophysiologic significance, the prediction of resistance to therapy or prognostic implications to immunochemotherapy treatment have not been thoroughly investigated in clinical trials. Additionally, activation-induced cytidine deaminase (AID)-driven somatic hypermutation (SHM) regulating the creation of recurrent SNVs, makes it difficult to distinguish between driver and passenger changes [23].

Hence to identify molecular subgroups and ultimately move towards precision treatment, requires a suitable panel of target genes NGS to uncover and describe the SNVs harbored by DLBCL patients. We devised an NGS lymphoma panel test after reviewing previous studies in DLBCL [13,14,20,23,24]. In the current study, the lymphoma panel was employed to sequence 20 patients diagnosed after a centralized histopathologic review. This study attempted to achieve information using the NGS lymphoma panel that can shed a light on the molecular heterogeneity of DLBCL subtypes.

Materials And Methods

This study included 25 DLBCL patients after a centralized histopathologic review. The quality and quantity of tumor cells were accessed by H&E staining and correlated with histopathology and immunohistochemistry (IHC) status. Patients were grouped as ABC, PMBL, and "other" or unclassified subtypes based on Hans algorithm. Five patients not having complete clinical information and immunohistochemistry data were excluded from this study. The study finally included 20 patients having DLBCL. Institutional Ethics Committee, All India Institute of Medical Sciences (AIIMS), Rishikesh approved (approval no. AIIMS/IEC/20/592) this study, and all participants gave written informed consent.

Pretreatment evaluation included complete physical examination, routine blood testing, and serum chemistry profile, including lactate dehydrogenase (LDH) levels and serum β 2-microglobulin. Additionally, computed tomography (CT) scans findings of the neck, chest, abdomen, and pelvis, and bone marrow examination were also noted. The Ann Arbor staging system was used. Immunohistochemically classified the enrolled patients into different subtypes [13].

The lymphoma panel of 133 was designed on targeted sequencing of multiple genes for the coding regions through NGS. The scope of this lymphoma panel included a panel of genes, wherein the prognostic significance of these genes and their mutations were intended to be studied. Previous publications were used to design it.

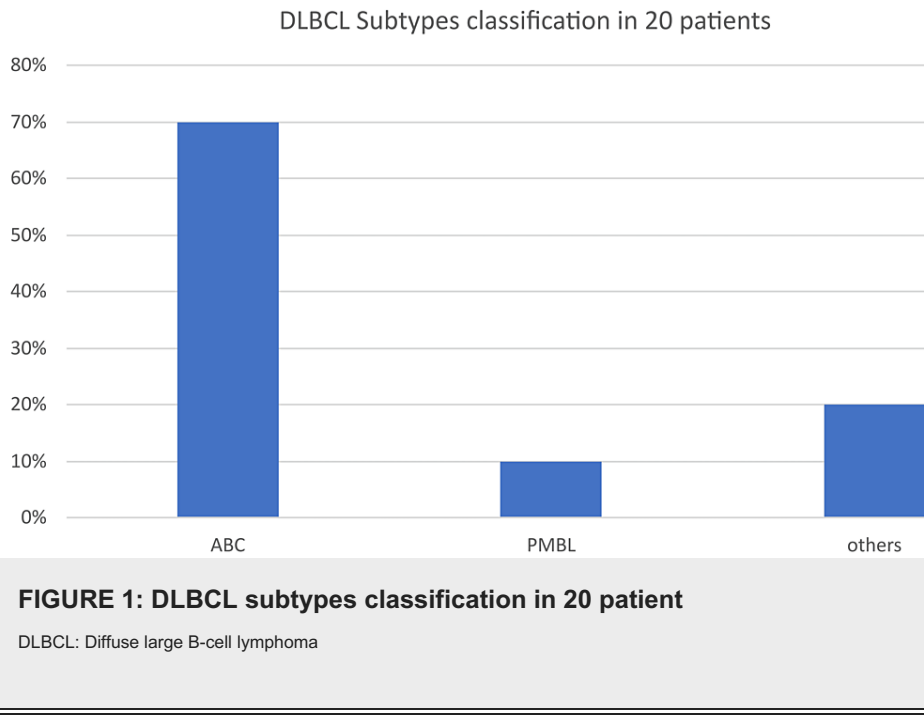
The cellularity of samples was checked before analysis. Targeted gene capture using a custom capture kit was done for genomic DNA. The libraries were prepared and sequenced using the Illumina platform. The human reference genome (GRCh38.p13/) alignment of obtained sequences was performed using Burrows-Wheeler Aligner (BWA) program [24,25]. Identification of somatic mutations was done using LoFreq (version 2) variant caller [26,27]. For clinical interpretation, only non-synonymous and splice site variants revealed in the coding regions were utilized.

The mutations were annotated using an annotation pipeline (VariMAT). Ensembl Variant Effect Predictor (VEP) program release 99 human gene Model was used for gene annotation of the variants [28,29]. Previously published literature and databases were used for the annotation of clinically relevant mutations. The common variants were filtered for reporting based on the presence in various population databases (1000G, ExAC, EVS, 1000japanese, dbSNP, UK10K, MedVarDb (in-house database) [30-34]. A custom read-depth-based algorithm was used to determine CNV (Copy Number Variants) from targeted sequencing experiments. Rare CNVs were detected using a comparison of the test data read-depths with the matched reference

dataset. Reportable mutations were prioritized and prepared based on AMP-ASCO-CAP (Association for Molecular Pathology-American Society of Clinical Oncology-College of American Pathologists), WHO guidelines [35,36], and also based on annotation metrics from OncoMD (a knowledge base of genomic alterations) [37].

Results

Lymphoma panel was performed on samples of only 20 DLBCL patients with a median age of 56 (range, 15-75) years. ABC accounts for 70% (14/20) of the patients, PMBL accounts for 10% (2/20), and other accounts for 20% (4/20) of the patients as depicted in Figure 1.



According to the lymphoma staging classification system, 40% of the patients included have AnnArbor stage 4. While 50% have stage 3. International Prognostic Index (IPI) scores categorize patients as having Low, Low-intermediate, and High-intermediate risk. In our cohort 52% of patients are at high intermediate risk, 33% are at low intermediate risk and only 15% of the patient are having the low-risk category. Total cohort 88% of the patients exhibited Clinical B symptoms.

The panel informativity was 95 percent. Nineteen out of 20 included subjects had at least one mutation. Variants per gene were found to be ranging from one to eight. Somatic mutational landscape examination after sequencing of all the genes yielded a result that among all the mutant genes 16.1% of patients had *NOTCH 1* mutations followed by 12.9% had *ARID1A* mutations. *MYD88* and *TP53* genes had a 9.6% mutation rate and *CSF3R* had a 6.4%. The rest of the genes like *FBXW7*, *PAX5*, *NF1*, *SOCS1*, *BRAF*, *INPPD*, *CUX1*, *JAK2*, *CBL*, and *CDKN1A* accounts for 3.2% of all mutations. Gene mutation frequencies in the various subtype are illustrated in Figure 2.

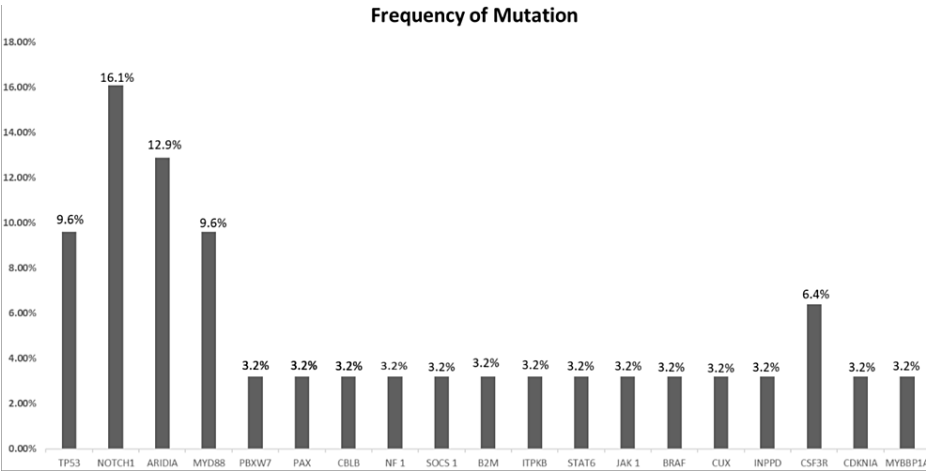


FIGURE 2: Mutation frequencies in the 20 patients.

The median overall sequencing depth for the lymphoma panel was 688X (120X-2049X), for individual genes sequencing depth is different as presented in Figure 3. The inner circle denotes the average depth for each patient and the outer ring shows the depth for each mutation for that particular patient.

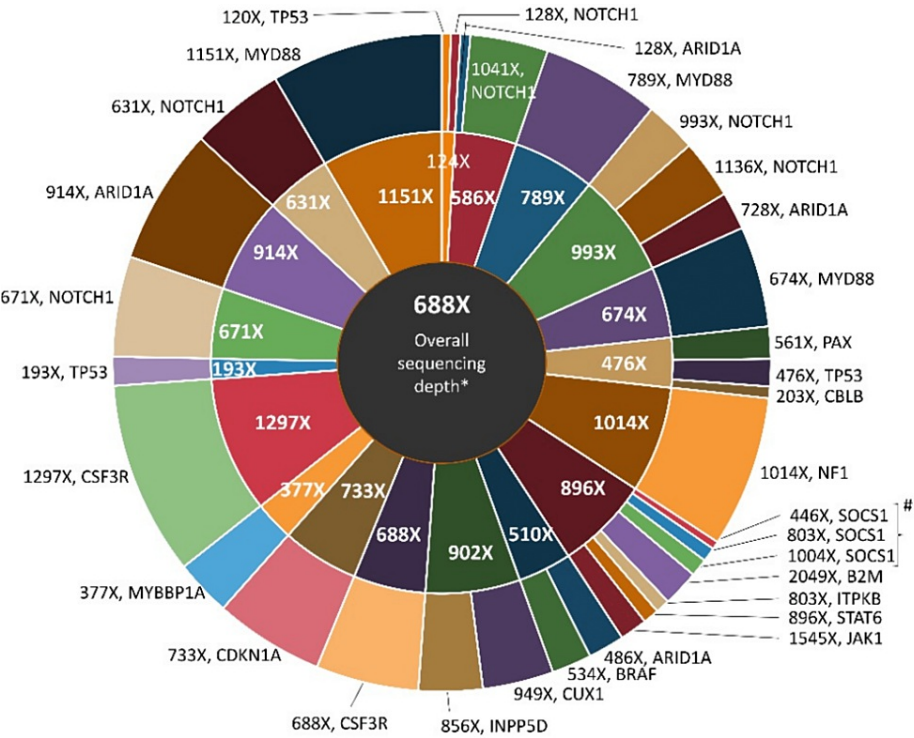


FIGURE 3: Overall Median Depth of Mutation

#one gene with different CDS (coding sequence) variants and amino acid change

Each finding in the lymphoma panel was correlated with clinical features like age, stage, IPI score, etc. The included cohort was divided into age groups of (n=2) 20 to 40 years, (n=9) 40 to 60 years, and those belonging to the (n=9) 60-80 age group. Only three mutations were observed in the younger age group (MYD88, CUX1, and INPP5D mutations). Genetic analysis of one patient of this age group revealed no mutation. Whereas NOTCH1 and TP53 are the most frequently reported gene in the middle age group (40-60). PAX, CBLB, BRAF, and CDKN1A frequency are reported only in the 60-80 years of age group. Mutation in MYD88 is reported in every age group. Of 20 patients, 15 patients are male and the most common mutations in male is in NOTCH1 and ARID1A (16% each), followed by SOCS1 (12%) and CSF3R (8%) whereas in females commonest mutations are MYD88 and TP53 (25% each). PAX, CBLB, and NF1 mutations are found only in the females included in the present study. According to IPI Score, patients having score 1 most

commonly have a mutation in *SOCS1* (43%). Those having a score 2 have *NOTCH 1* (22%) and enrolled patients with a score 3, both *NOTCH 1* and *ARID1A* (18%). As per Ann Arbor, most common mutations in stage 4 patients were in the *MYD88* gene followed by *CSF3R* (17%), whereas in stage 3 patients *SOCS 1* and *NOTCH 1* (23%) followed by *ARID1A* (15%) were commonly mutated and in stage 2 and stage 1 patients had mutated *TP53* gene. According to the Ki67 score, in the 60-80% score group, the most common mutations are in the genes of *NOTCH 1* (21%) followed by *ARID1A* and *TP53*. For those having scores of 40-60, the commonest mutations were found in *SOCS* (18%) followed by *MYD88* and *NOTCH 1* (12%). Overall *TP53* mutations were found in all stages. These are illustrated in Figure 4.

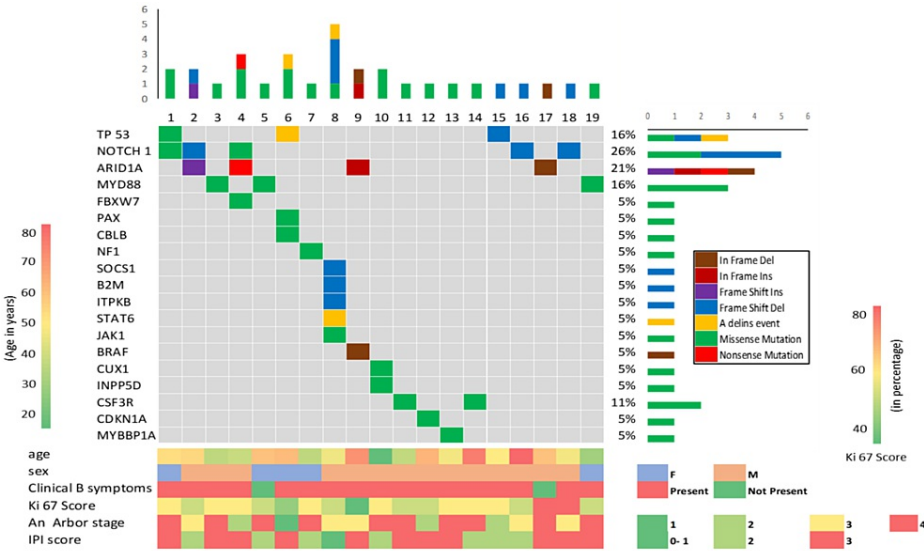


FIGURE 4: - Somatic Mutation landscape

In this landscape, a horizontal lane depicts a single gene and the vertical lines represent different mutations. One patient with no mutation is excluded in this figure. Different colors indicate mutation type. Schematic depiction of age, sex, clinical B symptoms, Ki 67 score, Ann arbor staging, and IPI scores with patients grouped by age (20-30 green, 30-40 light green, 40-50 yellow, 50-60 orange, and above 60 red color) and Clinical B symptoms is also shown (Red is Present and green is not present), Percentage of Ki 67 shown as (above 70 years old shown in red, between 50 to 60 year old shown in yellow and below 50 are shown in green), AnnArbor stage also indicated for each patient (Dark green 0-1, light green 2, yellow 3 and Red 4) and IPI score is indicated as 0-1 green color, 2 light green color, and 3 red colors).

IPI: International Prognostic Index

All of the gene alterations in our population were categorized into three subtypes of DLBCL. ABC, PMBL, and Other. *MYD88* (51%) is the most common mutation in ABC subtypes of DLBCL, followed by *NOTCH 1* (44%) and *SOCS 1* (33%) according to our findings. *ARID1A*, *CBLB*, *CUX1*, *JAK 1*, *NF 1*, *NOTCH 1*, and *STAT 6* are the most commonly reported mutations in PMBCL subtypes. *TP 53* is the most prevalent mutation in other subtypes, while *NOTCH 1* is the most common mutation in DLBCL patients with the ABC and PMBL subtypes.

Again, upon closer investigation missense mutation is the most frequent mutation observed in the total cohort targeting 68.4% of the total cohort followed by Frameshift deletion reported in 26.3% of the total cohort. *NOTCH 1* involves many Frameshift deletion mutations while the *ARID1A*-only gene includes four different types of mutations (frameshift insertion, in-frame insertion, nonsense mutation, and in-frame deletion mutation). Notably, *MYD88* and *TP 53* include many missense mutations in 32% of the samples followed by some frameshift deletions reported in *TP 53* mutation reported in this cohort.

Novel mutations

The lymphoma panel identified many novel mutations which could potentially serve as means to stratify patients and possibly align with treatment options. These were mostly loss function mutations and these variants have not been reported in ExAC and 1000 genomes database.

In one patient having ABC subtype, a frameshift deletion (chr9:g.136496514del; c.7225del) resulted in prematurely truncated protein (downstream to codon 2409 at 13 amino acid, p.Gln2409SerfsTer13) was spotted in the *NOTCH 1* gene. Another variant detected in the sample of the same patient was a nonsense mutation (chr1:g.26731170C>T; c.1369C>T) in the *ARID1A* gene resulting in the introduction of a stop codon and truncation of the protein at codon 457 (p.Gln457Ter). A frameshift deletion

(chr17:g.7675181_7675197del; c.415_431del) leading to prematurely truncated protein (at codon 139, only 4 amino acids downstream of the said codon, p.Lys139AlafsTer4) was detected in the *TP53* gene of this same patient.

A sample of another patient having PMBL revealed frameshift deletion in the *SOCS1* gene (chr16:g.11255131del; c.348del) resulting in termination of the protein (2 amino acids downstream to codon 116, p.Ser116ArgfsTer2). Another frameshift deletion in the *SOCS1* gene (chr16:g.11255444_11255457del; c.22_35del) resulting in subsequent termination of the protein (at codon 8, 104 amino acids downstream, p.Ala8SerfsTer104) was detected in same patient. A frameshift deletion (chr16:g.11255019_11255029del; c.450_460del) resulting in termination of the protein, (downstream to codon 151, p.Leu151ArgfsTer98) was detected in the *SOCS1* gene of same patient.

Discussion

A lymphoma panel was performed on samples of only 20 DLBCL patients with a median age of 56 (range, 15-75) years. ABC accounts for 70% (14/20) of the patients, PMBL accounts for 10% (2/20), and other accounts for 20% (4/20) of the patients as depicted in similar results previously reported indicating molecular heterogeneity of DLBCL with a preponderance of ABC subtypes [38].

The lymphoma panel was designed for targeted sequencing of multiple genes for the coding regions through NGS. Our study included 133 genes and there are studies with multiple genes [38,39]. The informativity of our lymphoma panel was 95 percent similar to previous studies [38]. One patient sample did not yield genetic variation. The number of variants per gene was found to be ranging from one to eight.

On examination of the somatic mutational landscape, we found out that among all the mutant genes 16.1% of patients had *NOTCH 1* mutations followed by 12.9% *ARID1A* mutations. Mutation analysis in many studies has delineated that *NOTCH 1* mutations are one of the distinctive drivers of DLBCL [38,40]. Schmitz et al. developed a molecular classification based on 574 cases of DLBCL using various methods to identify genetic subtypes [41]. The four subtypes given by them include EZB, BN2, MCD and N1. The genetic hallmarks of BN2 subtypes include BCL6 fusions and variations of the pathway of NOTCH as found in our study. *ARID1A* mutations are common in ovarian clear cell carcinoma. Also detected in the 2-17 % range in myeloblastoma, neuroblastoma, Burkitt lymphoma, and Waldenström macroglobulinemia. Along with other genes (*HIST1H1C/D/E*, and *SMARCA4*), the *ARD1A* gene, one of the regulators of higher-order chromatin structure, has been reported to be mutated in $\geq 5\%$ of DLBCL or follicular lymphoma [42]. In our study, a much higher yield of 12.9% was seen, with *ARID1A* in all subtypes. This difference needs further studies since this has not been documented in Indian DLBCL patients previously, this finding is significant and further studies are required to give insight.

MYD88 and *TP53* genes had a 9.6% mutation rate and *CSF3R* had a 6.4%. *MYD88* (40%) is the most common mutation in ABC subtypes of DLBCL gene alterations, according to our findings. Our findings have been mirrored by previous reports of *MYD88* mutation being reported in 30-40% of ABC type DLBCL [43], although lower frequencies have been reported by others [43,44]. *MYD88* has been selected as an essential mutation in the ABC [45,46]. The rate of TP53 variants present in our study was less than in previously reported publications [38]. Our study has 12% of samples showing *SOCS1* mutations while the previous study has reported higher (55.5%) PMBL patients [38].

In the present study, the cohort was divided into three age groups. *PAX*, *CBLB*, *BRAF*, and *CDKN1A* frequency are reported only in the 60-80 years of age group. *NOTCH 1* and *TP 53* are the most frequently reported gene in the middle age group (40-60). We observed only three mutations in the younger age group (*MYD88*, *CUX1*, and *INPP5D*). Mutation in *MYD88* is reported in every age group. in a previous study, authors found a correlation between age with *CD79B*, *KMT2D*, and *MYD88* mutations [38].

According to IPI Score in our study, in Score 1 most common mutation is found in *SOCS1* (43%) and score 2 *NOTCH1* (22%) and score 3, both *NOTCH1* and *ARID1A* (18%) followed by *MYD88* and *TP53* (12% each). The comparison of four independent DLBCL studies done by Mellert et al. suggests that *SOCS1* mutations may be used to forecast an extremely favorable overall survival rate of early-onset DLBCL patients (aged <60 years) [47]. Similarly, it is known that *TP53* alterations are more frequently associated with a lack of response to first-line chemotherapy [48] while *MYD88* mutations signal adverse prognostic impact [49]. Both these findings are mirrored in our study with lower score patients having higher *SOCS1* mutation and those having an IPI score of 3 having a high percentage of *TP53* and *MYD88* mutations.

The investigation also yielded missense mutation to be the most frequent mutation observed in the total cohort targeting 68.4% of the total cohort followed by frameshift deletion reported in 26.3% of the total cohort. *NOTCH 1* involves many frameshift deletion mutations while the *ARID1A*-only gene includes four different types of mutations (frameshift insertion, in-frame insertion, nonsense mutation, and frame deletion mutation). Notably, *MYD88* and *TP 53* include many missense mutations in 32% of the samples followed by some frameshift deletions reported in the *TP53* mutation reported in this cohort. A study by Fan et al. concluded that out of the total, somatic mutations detected in DLBCL, foremost were missense

type, followed by silent, splice-site, nonsense, and indels. Of the indels, maximum caused reading frameshifts [50].

In our present study, the lymphoma panel identified many novel mutations. These included a frameshift deletion in the *NOTCH1* gene, a nonsense mutation in the *ARID1A* gene, frameshift deletions in the *SOCS1* gene, and frameshift deletion in the *TP53* gene. These novel mutations can potentially be used for the stratification of patients in the future according to linking them to treatment options in coming studies.

Study limitations

The study's main limitations include its small sample size and lack of follow-up.

Conclusions

The importance of the information provided by the NGS lymphoma panel in patients with DLBCL is demonstrated in this study. Findings highlight the subtype molecular heterogeneity and identify novel mutations. This study's results point toward the importance of NGS, giving information about well-documented genes as well as novel mutations which can be potential targets in the future. The study adds to already existing knowledge significantly. Lymphoma panel has identified many novel mutations which could potentially have clinical applications in terms of stratification of patients according to treatment options. The major challenge which remains is pinpointing a clinical relevance. There clear need for an all-inclusive resource annotating associations of revealed genetic alterations with various pathophysiological pathways, possible drug targets, and their effect on therapy response. Further studies with NGS in DLBCL will make it possible to explore genetic variants and annotate and associate them with the phenotype for proper utilization in clinical settings.

Additional Information

Disclosures

Human subjects: Consent was obtained or waived by all participants in this study. Institutional Ethics Committee, All India Institute of Medical Sciences (AIIMS), Rishikesh issued approval AIIMS/IEC/20/592. The Institutional Ethics Committee at All India Institute of Medical Sciences, Rishikesh issued approval ECR/736/Inst/UK/2015/RR-18, dated 22/08/2020. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** This study was supported by All India Institute of Medical Sciences (AIIMS), Rishikesh. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgements

Study design and conceptualization is done by Garima Mamgain and Uttam Kumar Nath. Data analysis and interpretation are done by Garima Mamgain, Manisha Naithani & Priyanka Patra. Reviewing and revising the manuscript is done by Manisha Naithani. IHC Data interpretation is done by Arvind Gupta. Gaurav Dhingra and Deepak Sundriyal helped in preparing the draft of the manuscript. Sample acquisition was aided by Mukesh Mamgain, Sikha Morang, Jhasketan Nayak, Kartik Kumar, Shashikant Singh, Anamika Bakliwal, Ashok Rajoreya, Sudeep V, Debranjani Chattopadhyay and Rituparna Chetia.

References

1. WHO Classification of tumours of haematopoietic and lymphoid tissues. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J (ed): IARC Publications, Lyon, France; 2017. <https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/WHO-Classification-Of-Tumours-Of-Ha...>
2. Alizadeh AA, Eisen MB, Davis RE, et al.: Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*. 2000, 403:503-11. [10.1038/35000501](https://doi.org/10.1038/35000501)
3. Wright G, Tan B, Rosenwald A, Hurt EH, Wiestner A, Staudt LM: A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A*. 2003, 100:9991-6. [10.1073/pnas.1732008100](https://doi.org/10.1073/pnas.1732008100)
4. Huang JZ, Sanger WG, Greiner TC, et al.: The t(14;18) defines a unique subset of diffuse large B-cell lymphoma with a germinal center B-cell gene expression profile. *Blood*. 2002, 99:2285-90. [10.1182/blood.v99.7.2285](https://doi.org/10.1182/blood.v99.7.2285)
5. Pfeifer M, Grau M, Lenze D, et al.: PTEN loss defines a PI3K/AKT pathway-dependent germinal center subtype of diffuse large B-cell lymphoma. *Proc Natl Acad Sci U S A*. 2013, 110:12420-5. [10.1073/pnas.1305656110](https://doi.org/10.1073/pnas.1305656110)
6. Lenz G, Wright GW, Emre NC, et al.: Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. *Proc Natl Acad Sci U S A*. 2008, 105:13520-5. [10.1073/pnas.0804295105](https://doi.org/10.1073/pnas.0804295105)
7. Lenz G, Staudt LM: Aggressive lymphomas. *N Engl J Med*. 2010, 362:1417-29. [10.1056/NEJMra0807082](https://doi.org/10.1056/NEJMra0807082)

8. Savage KJ, Monti S, Kutok JL, et al.: The molecular signature of mediastinal large B-cell lymphoma differs from that of other diffuse large B-cell lymphomas and shares features with classical Hodgkin lymphoma. *Blood*. 2003, 102:3871-9. [10.1182/blood-2003-06-1841](#)
9. Reichel J, Chadburn A, Rubinstein PG, et al.: Flow sorting and exome sequencing reveal the oncogenome of primary Hodgkin and Reed-Sternberg cells. *Blood*. 2015, 125:1061-72. [10.1182/blood-2014-11-610436](#)
10. Salles G, de Jong D, Xie W, et al.: Prognostic significance of immunohistochemical biomarkers in diffuse large B-cell lymphoma: a study from the Lunenburg Lymphoma Biomarker Consortium. *Blood*. 2011, 117:7070-8. [10.1182/blood-2011-04-345256](#)
11. Mareschal S, Ruminy P, Bagacean C, et al.: Accurate classification of germinal center B-cell-like/activated B-cell-like diffuse large B-cell lymphoma using a simple and rapid reverse transcriptase-multiplex ligation-dependent probe amplification assay: a CALYM study. *J Mol Diagn*. 2015, [10.1016/j.jmoldx.2015.01.007](#)
12. Scott DW, Mottok A, Ennishi D, et al.: Prognostic significance of diffuse large B-cell lymphoma cell of origin determined by digital gene expression in formalin-fixed paraffin-embedded tissue biopsies. *J Clin Oncol*. 2015, 33:2848-56. [10.1200/JCO.2014.60.2383](#)
13. Morin RD, Mungall K, Pleasance E, et al.: Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. *Blood*. 2013, 122:1256-65. [10.1182/blood-2013-02-483727](#)
14. Mullighan CG: Genome sequencing of lymphoid malignancies. *Blood*. 2013, 122:3899-907. [10.1182/blood-2013-08-460311](#)
15. Morin RD, Johnson NA, Severson TM, et al.: Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat Genet*. 2010, 42:181-5. [10.1038/ng.518](#)
16. Morin RD, Mendez-Lago M, Mungall AJ, et al.: Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature*. 2011, 476:298-303. [10.1038/nature10351](#)
17. Pasqualucci L, Trifonov V, Fabbri G, et al.: Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat Genet*. 2011, 43:830-7. [10.1038/ng.892](#)
18. Zhang J, Grubor V, Love CL, et al.: Genetic heterogeneity of diffuse large B-cell lymphoma. *Proc Natl Acad Sci U S A*. 2013, 110:1398-403. [10.1073/pnas.1205299110](#)
19. Steidl C, Gascoyne RD: The molecular pathogenesis of primary mediastinal large B-cell lymphoma. *Blood*. 2011, 118:2659-69. [10.1182/blood-2011-05-326538](#)
20. Ritz O, Guiter C, Castellano F, et al.: Recurrent mutations of the STAT6 DNA binding domain in primary mediastinal B-cell lymphoma. *Blood*. 2009, 114:1236-42. [10.1182/blood-2009-03-209759](#)
21. Khodabakhshi AH, Morin RD, Fejes AP, et al.: Recurrent targets of aberrant somatic hypermutation in lymphoma. *Oncotarget*. 2012, 3:1308-19. [10.18632/oncotarget.653](#)
22. Bohers E, Mareschal S, Bertrand P, et al.: Activating somatic mutations in diffuse large B-cell lymphomas: lessons from next generation sequencing and key elements in the precision medicine era. *Leuk Lymphoma*. 2015, 56:1213-22. [10.3109/10428194.2014.941836](#)
23. Mareschal S, Dubois S, Vially PJ, et al.: Whole exome sequencing of relapsed/refractory patients expands the repertoire of somatic mutations in diffuse large B-cell lymphoma. *Genes Chromosomes Cancer*. 2016, 55:251-67. [10.1002/gcc.22328](#)
24. Li H, Durbin R: Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010, 26:589-95. [10.1093/bioinformatics/btp698](#)
25. Dubois S, Mareschal S, Picquenot JM, et al.: Immunohistochemical and genomic profiles of diffuse large B-cell lymphomas: implications for targeted EZH2 inhibitor therapy?. *Oncotarget*. 2015, 6:16712-24. [10.18632/oncotarget.3154](#)
26. Meyer LR, Zweig AS, Hinrichs AS, et al.: The UCSC Genome Browser database: extensions and updates 2013. *Nucleic Acids Res*. 2013, 41:D64-9. [10.1093/nar/gks1048](#)
27. Wilm A, Aw PP, Bertrand D, et al.: LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Res*. 2012, 40:11189-201. [10.1093/nar/gks918](#)
28. Wilson WH, Young RM, Schmitz R, et al.: Targeting B cell receptor signaling with ibrutinib in diffuse large B cell lymphoma. *Nat Med*. 2015, 21:922-6. [10.1038/nm.3884](#)
29. Li H, Handsaker B, Wysoker A, et al.: The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009, 25:2078-9. [10.1093/bioinformatics/btp352](#)
30. McLaren W, Gil L, Hunt SE, et al.: The ensembl variant effect predictor. *Genome Biol*. 2016, 17:122. [10.1186/s13059-016-0974-4](#)
31. Zerbino DR, Achuthan P, Akanni W, et al.: Ensembl 2018. *Nucleic Acids Res*. 2018, 46:D754-61. [10.1093/nar/gkx1098](#)
32. Auton A, Brooks LD, Durbin RM, et al.: A global reference for human genetic variation. *Nature*. 2015, 526:68-74. [10.1038/nature15393](#)
33. Lek M, Karczewski KJ, Minikel EV, et al.: Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016, 536:285-91. [10.1038/nature19057](#)
34. Nagasaki M, Yasuda J, Katsuoka F, et al.: Rare variant discovery by deep whole-genome sequencing of 1,070 Japanese individuals. *Nat Commun*. 2015, 6:8018. [10.1038/ncomms9018](#)
35. Moayyeri A, Hammond CJ, Hart DJ, Spector TD: The UK Adult Twin Registry (TwinsUK Resource). *Twin Res Hum Genet*. 2013, 16:144-9. [10.1017/thg.2012.89](#)
36. Li MM, Datto M, Duncavage EJ, et al.: Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J Mol Diagn*. 2017, 19:4-23. [10.1016/j.jmoldx.2016.10.002](#)
37. Momose S, Weißbach S, Pischmarov J, et al.: The diagnostic gray zone between Burkitt lymphoma and diffuse large B-cell lymphoma is also a gray zone of the mutational spectrum. *Leukemia*. 2015, 29:1789-91. [10.1038/leu.2015.34](#)
38. Roschewski M, Phelan JD, Wilson WH: Molecular Classification and Treatment of Diffuse Large B-Cell Lymphoma and Primary Mediastinal B-Cell Lymphoma. *Cancer J*. 2020, 26:195-205. [10.1097/PPO.0000000000000450](#)

39. Bueno R, Stawiski EW, Goldstein LD, et al.: Comprehensive genomic analysis of malignant pleural mesothelioma identifies recurrent mutations, gene fusions and splicing alterations. *Nat Genet.* 2016, 48:407-16. [10.1038/ng.3520](#)
40. Arruga F, Vaisitti T, Deaglio S: The NOTCH Pathway and Its Mutations in Mature B Cell Malignancies . *Front Oncol.* 2018, 8:550. [10.3389/fonc.2018.00550](#)
41. Schmitz R, Wright GW, Huang DW, et al.: Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma . *N Engl J Med.* 2018, 378:1396-407. [10.1056/NEJMoa1801445](#)
42. Koh Y: Genomics of diffuse large B cell lymphoma. *Blood Res.* 2021, 56:S75-9. [10.5045/br.2021.2021049](#)
43. Lee JH, Jeong H, Choi JW, Oh H, Kim YS: Clinicopathologic significance of MYD88 L265P mutation in diffuse large B-cell lymphoma: a meta-analysis. *Sci Rep.* 2017, 7:1785. [10.1038/s41598-017-01998-5](#)
44. Lunning MA, Green MR: Mutation of chromatin modifiers; an emerging hallmark of germinal center B-cell lymphomas. *Blood Cancer J.* 2015, 5:e361. [10.1038/bcj.2015.89](#)
45. Ngo VN, Young RM, Schmitz R, et al.: Oncogenically active MYD88 mutations in human lymphoma . *Nature.* 2011, 470:115-9. [10.1038/nature09671](#)
46. Lohr JG, Stojanov P, Lawrence MS, et al.: Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc Natl Acad Sci U S A.* 2012, 109:3879-84. [10.1073/pnas.1121343109](#)
47. Mellert K, Martin M, Lennerz JK, et al.: The impact of SOCS1 mutations in diffuse large B-cell lymphoma . *Br J Haematol.* 2019, 187:627-37. [10.1111/bjh.16147](#)
48. Reddy A, Zhang J, Davis NS, et al.: Genetic and functional drivers of diffuse large B cell lymphoma . *Cell.* 2017, 171:481-494.e15. [10.1016/j.cell.2017.09.027](#)
49. Choi JW, Kim Y, Lee JH, Kim YS: MYD88 expression and L265P mutation in diffuse large B-cell lymphoma . *Hum Pathol.* 2013, 44:1375-81. [10.1016/j.humpath.2012.10.026](#)
50. Fan Z, Pei R, Sha K, Chen L, Wang T, Lu Y: Comprehensive characterization of driver genes in diffuse large B cell lymphoma. *Oncol Lett.* 2020, 20:382-90. [10.3892/ol.2020.11552](#)