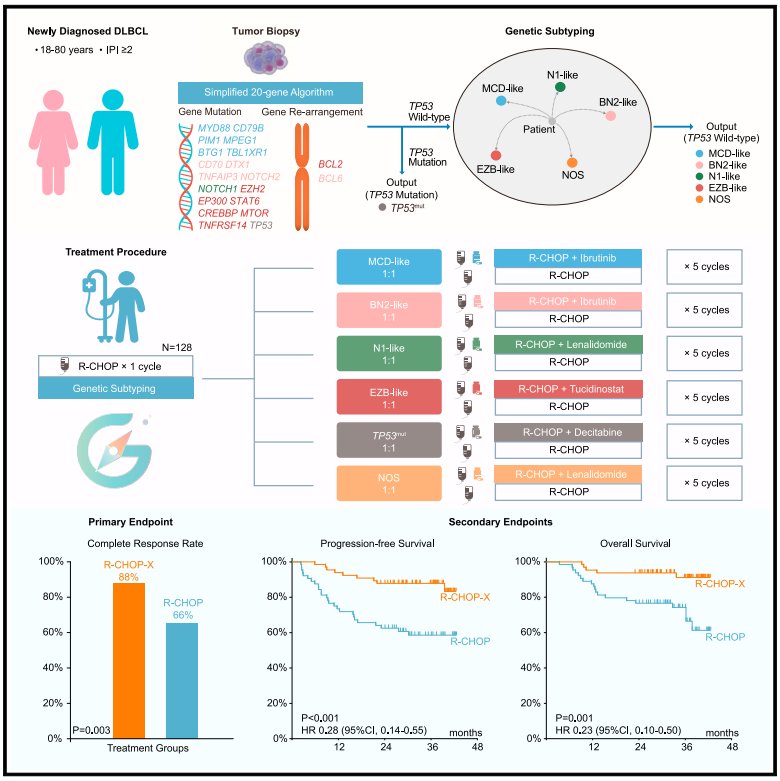
Cancer Cell

Article

# Genetic subtype-guided immunochemotherapy in diffuse large B cell lymphoma: The randomized GUIDANCE-01 trial

## Graphical abstract




## Highlights

- Targeted agents based on genetic subtyping improve R-CHOP efficacy in DLBCL
- Inflammatory and immune-depleted tumors have greater sensitivity to therapy
- Genetic subtype-guided immunochemotherapy is promising first-line therapy in DLBCL


## Authors

Mu-Chen Zhang, Shuang Tian, Di Fu, ..., Sai-Juan Chen, Peng-Peng Xu, Wei-Li Zhao

## Correspondence

pengpeng_xu@126.com (P.-P.X.), zhao.weili@yahoo.com (W.-L.Z.)



## In brief

Zhang et al. report findings from the GUIDANCE-01 trial showing genetic subtype-guided immunochemotherapy (R-CHOP-X) improves complete response rate, progression-free and overall survival in diffuse large B cell lymphoma (DLBCL). Among lymphoma microenvironment subtypes, inflammatory and immune-depleted tumors may benefit from R-CHOP-X. This mechanism-based tailored therapy dually targeting genetic and microenvironmental alterations is effective and safe in newly diagnosed DLBCL.


Zhang et al., 2023, Cancer Cell *41*, 1705–1716
October 9, 2023 
https://doi.org/10.1016/j.ccell.2023.09.004

CellPress

Cancer Cell

CellPress
OPEN ACCESS

Article

# Genetic subtype-guided immunochemotherapy in diffuse large B cell lymphoma: The randomized GUIDANCE-01 trial

Mu-Chen Zhang,[1,8] Shuang Tian,[1,8] Di Fu,[1,8] Li Wang,[1,2] Shu Cheng,[1] Hong-Mei Yi,[3] Xu-Feng Jiang,[4] Qi Song,[5] Yan Zhao,[1] Yang He,[1] Jian-Feng Li,[1] Rong-Ji Mu,[6] Hai Fang,[1] Hao Yu,[7] Hui Xiong,[7] Biao Li,[4] Sai-Juan Chen,[1] Peng-Peng Xu,[1,9,*] and Wei-Li Zhao[1,2,9,10,*]

[1]Shanghai Institute of Hematology, State Key Laboratory of Medical Genomics, National Research Center for Translational Medicine at Shanghai, Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China
[2]Pôle de Recherches Sino-Français en Science du Vivant et Génomique, Laboratory of Molecular Pathology, Shanghai, China
[3]Department of Pathology, Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China
[4]Department of Nuclear Medicine, Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China
[5]Department of Radiology, Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China
[6]Clinical Research Institute, Shanghai Jiao Tong University School of Medicine, Shanghai, China
[7]Department of Research and Development, Shanghai Righton Biotechnology Co. Ltd, Shanghai, China
[8]These authors contributed equally
[9]These authors contributed equally
[10]Lead contact
*Correspondence: pengpeng_xu@126.com (P.-P.X.), zhao.weili@yahoo.com (W.-L.Z.)
https://doi.org/10.1016/j.ccell.2023.09.004

## SUMMARY

We report the results of GUIDANCE-01 (NCT04025593), a randomized, phase II trial of R-CHOP alone or combined with targeted agents (R-CHOP-X) guided by genetic subtyping of newly diagnosed, intermediate-risk, or high-risk diffuse large B cell lymphoma (DLBCL). A total of 128 patients were randomized 1:1 to receive R-CHOP-X or R-CHOP. The study achieved the primary endpoint, showing significantly higher complete response rate with R-CHOP-X than R-CHOP (88% vs. 66%, p = 0.003), with overall response rate of 92% vs. 73% (p = 0.005). Two-year progression-free survival rates were 88% vs. 63% (p < 0.001), and 2-year overall survival rates were 94% vs. 77% (p = 0.001). Meanwhile, post hoc RNA-sequencing validated our simplified genetic subtyping algorithm and previously established lymphoma microenvironment subtypes. Our findings highlight the efficacy and safety of R-CHOP-X, a mechanism-based tailored therapy, which dually targeted genetic and microenvironmental alterations in patients with newly diagnosed DLBCL.

## INTRODUCTION

Diffuse large B cell lymphoma (DLBCL) is the most common subtype of non-Hodgkin lymphomas.[1] Although anti-CD20 antibody rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP) immunochemotherapy significantly improve clinical outcomes of DLBCL, only approximately 60% of patients with intermediate risk or high risk (International Prognostic Index [IPI] $\geq$ 2) achieve complete response (CR) upon R-CHOP treatment.[2,3] With the development of novel targeted agents, clinical trials have focused on combination with R-CHOP in these subsets of DLBCL patients. In the PHOENIX trial, the B cell receptor (BCR) signaling pathway inhibitor ibrutinib plus R-CHOP provided survival benefit among non-germinal center B cell (non-GCB) lymphoma patients under 60 years old with IPI 2–5.[4] In the ROBUST trial, lenalidomide plus R-CHOP improved prognosis of activated B cell (ABC) lymphoma patients with IPI 3–5.[5] Meanwhile, in the ECOG-ACRIN E1412 trial, lenalidomide plus R-CHOP reduced the risk of progression or death in DLBCL regardless of GCB and non-GCB subtype.[6] Recently, in the POLARIX trial, polatuzumab vedotin targeting surface antigen CD79b, rituximab, cyclophosphamide, doxorubicin, and prednisone prolonged survival of all patients with IPI 2–5.[7] These results indicate that novel targeted agents plus immunochemotherapy are promising therapeutic approaches in DLBCL.

Molecular heterogeneity may influence clinical outcomes in DLBCL.[8] Two whole-exome sequencing (WES) studies proposed genetic subtypes with comparable genetic features,[9,10] including MCD/C5 cluster, BN2/C1 cluster, and EZB/C3 cluster, while subtypes such as N1,[9] *TP53*-associated/C2 cluster, and *SGK1*-associated/C4 cluster were slightly different.[10] Recent studies have also proven the feasibility of simplified genetic subtyping with targeted sequencing of selected genes.[11,12] In addition to genetic alterations of tumor cells, the tumor microenvironment plays an essential role in DLBCL progression. Four major lymphoma microenvironment (LME) categories have been

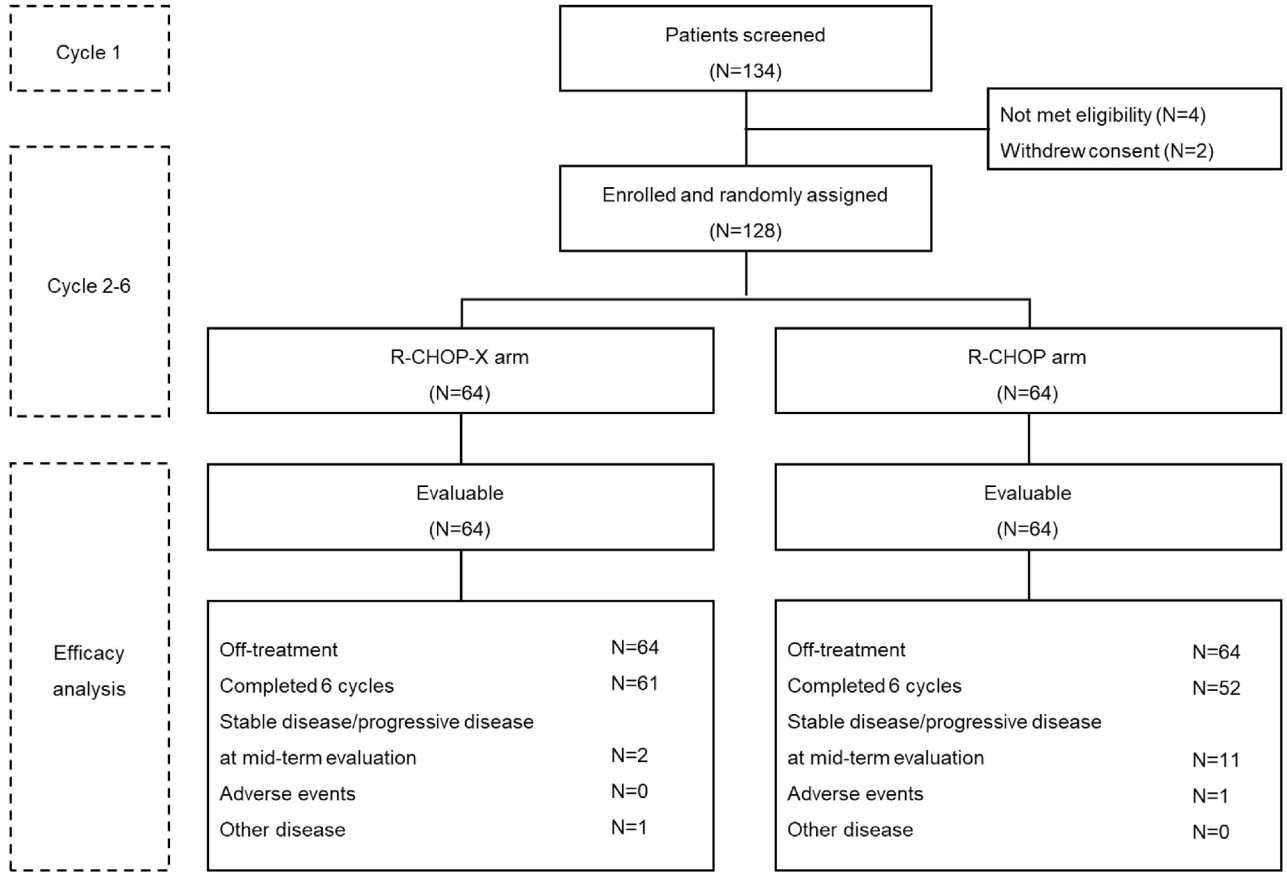



**Figure 1. GUIDANCE-01 trial consort table**
See also Figure S1.

defined, including germinal center-like (GC), with cell types commonly found in germinal centers, mesenchymal (MS) for the abundance of stromal cells and extracellular matrix pathways, inflammatory (IN) for the presence of inflammatory cells and pathways, and a depleted (DP) form by an overall lower presence of microenvironment-derived gene expression.[13] However, a therapeutic approach that selects different targeted agents based on multiple genetic subtypes has not been investigated.

As for the rationale of the selection of potential targeted agents, DLBCL with MCD-associated mutations are often sensitive to Bruton's tyrosine kinase (BTK) inhibitor ibrutinib.[14] BN2 subtype is also notably enriched for BCR-NF-κB aberrations.[9,10] Histone acetyltransferases *CREBBP/EP300* mutations enriched in EZB subtype can be targeted by histone deacetylase (HDAC) inhibitors.[15,16] Demethylating agent decitabine is effective in acute myeloid leukemia (AML), myelodysplastic syndrome (MDS), and DLBCL patients with *TP53* mutations.[17,18] For N1 and unclassified subtypes, we applied lenalidomide for its potential clinical efficacy in both GCB- and ABC-DLBCL.[5,6,19] In this study, we investigated genetic subtype-guided targeted agents plus R-CHOP (R-CHOP-X) in a randomized, phase 2 clinical trial, to evaluate efficacy and safety and to explore genetic and LME subtypes that could benefit from R-CHOP-X in patients with newly diagnosed intermediate-risk or high-risk DLBCL.

## RESULTS

### Patient characteristics

Between July, 2019 and December, 2020, 134 patients were screened and treated with 1 cycle of standard R-CHOP. Four patients did not meet eligibility because of suspected central nervous system involvement by flow cytometry of cerebrospinal fluid (n = 1) and insufficient DNA concentration for targeted sequencing (n = 3), and 2 patients withdrew consent before randomization. Therefore, a total of 128 patients were randomly assigned 1:1 to the R-CHOP-X arm (n = 64) and the R-CHOP arm (n = 64), as shown in Figures 1 and S1. All 128 randomized patients were included in the efficacy and safety analyses. The median time from diagnosis to treatment was 18 days (19 days in the R-CHOP-X arm and 17 days in the R-CHOP arm). The median time from enrollment to response to classifying patients according to genetic subtypes was 14 days (range, 6–19). No patient had treatment delay in the second cycle due to timely results of genetic subtyping. Patient characteristics are summarized in Tables 1 and S1. Baseline clinical and pathological characteristics were comparable between the 2 groups with a median age of 64 years (range, 25–74). Most patients presented relatively high-risk disease: 77% with Ann Arbor stage III or IV, 80% with elevated serum LDH level, 52% with 2 or more extranodal involvement sites, and 65% with IPI 3–5. Among 29 patients

**Table 1. Patient characteristics**

| Characteristics | Total (n = 128) | R-CHOP-X (n = 64) | R-CHOP (n = 64) | p value |
|---|---|---|---|---|
| Age (year): median (range) | 64 (25–74) | 64 (29–74) | 64.5 (25–74) | 0.905 |
| Time from diagnosis to treatment (day): median (range) | 18 (2–84) | 19 (2–55) | 17 (2–84) | 0.547 |
| Sex-Male: n (%) | 67 (52%) | 32 (50%) | 35 (55%) | 0.595 |
| Stage III-IV: n (%) | 99 (77%) | 47 (73%) | 52 (81%) | 0.291 |
| Elevated LDH: n (%) | 102 (80%) | 53 (83%) | 49 (77%) | 0.380 |
| Extranodal sites ≥2: n (%) | 66 (52%) | 33 (52%) | 33 (52%) | >0.999 |
| ECOG performance status 0–1: n (%) | 106 (83%) | 54 (84%) | 52 (81%) | 0.639 |
| IPI 3–5: n (%) | 83 (65%) | 42 (66%) | 41 (64%) | 0.853 |
| Cell of origin: n (%) | | | | |
| GCB | 23/84 (27%) | 12/42 (28%) | 11/42 (26%) | 0.330 |
| ABC | 42/84 (50%) | 18/42 (43%) | 24/42 (57%) | |
| Unclassified | 19/84 (23%) | 12/42 (28%) | 7/42 (17%) | |
| MYC/BCL2 double expression: n (%) | | | | |
| Yes | 46 (36%) | 24 (38%) | 22 (34%) | 0.713 |
| No | 82 (64%) | 40 (63%) | 42 (66%) | |
| FISH: n (%) | | | | |
| MYC/BCL6 re-arrangement | 1 (1%) | 1 (2%) | 0 | >0.999 |
| MYC/BCL2 re-arrangement | 0 | 0 | 0 | – |
| 17p13 deletion | 19/117 (16%) | 10/57 (18%) | 9/60 (15%) | 0.709 |
| Genetic subtype: n (%) | | | | |
| MCD-like | 26 (20%) | 13 (20%) | 13 (20%) | >0.999 |
| BN2-like | 23 (18%) | 11 (17%) | 12 (19%) | |
| N1-like | 5 (4%) | 3 (5%) | 2 (3%) | |
| EZB-like | 3 (2%) | 1 (2%) | 2 (3%) | |
| *TP53*$^{mut}$ | 21 (16%) | 11 (17%) | 10 (16%) | |
| NOS | 50 (39%) | 25 (39%) | 25 (39%) | |

See also Table S1.

R-CHOP, rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone. R-CHOP-X, R-CHOP plus targeted agents. LDH, lactate dehydrogenase. ECOG, Eastern Cooperative Oncology Group. IPI, International Prognostic Index. GCB, germinal B cell. ABC, activated B cell. *TP53*$^{mut}$, *TP53* mutations.

with Ann Arbor stage II, 27 patients had elevated serum LDH level and 5 patients had 2 extranodal sites. Among 84 patients with available RNA-sequencing (RNA-seq) data, 71% of patients presented with non-GCB phenotype (ABC and unclassified). As assessed by immunohistochemistry, 36% of patients presented with MYC/BCL2 double expression. One patient presented with MYC/BCL6 re-arrangement lymphoma, and no patient had MYC/BCL2 re-arrangement. Simplified 20-gene algorithm was established (STAR methods) to categorize patients into MCD-like (26/128, 20%), BN2-like (23/128, 18%), N1-like (5/128, 4%), EZB-like (3/128, 2%), *TP53* mutations (*TP53*$^{mut}$) (21/128, 16%), and not otherwise specified (NOS) (50/128, 39%).

## Clinical efficacy

### *Primary endpoint*

The CR rate (CRR) at the end of the treatment was 88% (56/64, 95% CI, 79 to 96) in the R-CHOP-X arm and 66% (42/64, 95% CI, 54 to 78) in the R-CHOP arm (p = 0.003). The study met its pre-specified primary endpoint. In terms of genetic subtypes, CRRs of MCD-like, BN2-like, N1-like, EZB-like, *TP53*$^{mut}$, and NOS were 85% (95% CI, 62 to 100), 91% (95% CI, 71 to 100), 100%, 100%, 82% (95% CI, 55 to 100), and 88% (95% CI, 74 to 100) in the R-CHOP-X arm, and 54% (95% CI, 23 to 85), 67% (95% CI, 35 to 98), 50%, 0, 60% (95% CI, 23 to 97), and 80% (95% CI, 63 to 97) in the R-CHOP arm, respectively (Figure 2A). The difference observed within each genetic subtype was exploratory.

### *Secondary endpoints*

The overall response rate (ORR) at the end of the treatment was 92% (95% CI, 85 to 99) in the R-CHOP-X arm, and 73% (95% CI, 62 to 85) in the R-CHOP arm (p = 0.005). The ORRs were improved in most of the genetic subtypes (Figure 2B). With a median follow-up of 36 months, median progression-free survival (PFS) and overall survival (OS) were not reached. Two-year PFS rates were 88% (95% CI, 77 to 94) and 63% (95% CI, 49 to 73) (HR 0.28, 95% CI 0.14–0.55, p < 0.001) for R-CHOP-X and R-CHOP, respectively (Figure 2C). In addition, 2-year OS rates were also superior for R-CHOP-X, as compared to R-CHOP (94% [95% CI, 84 to 98] vs. 77% [95% CI, 64 to 85], HR 0.23, 95% CI 0.10–0.50, p = 0.001) (Figure 2D). The results of exploratory subgroup analysis of PFS and OS were shown

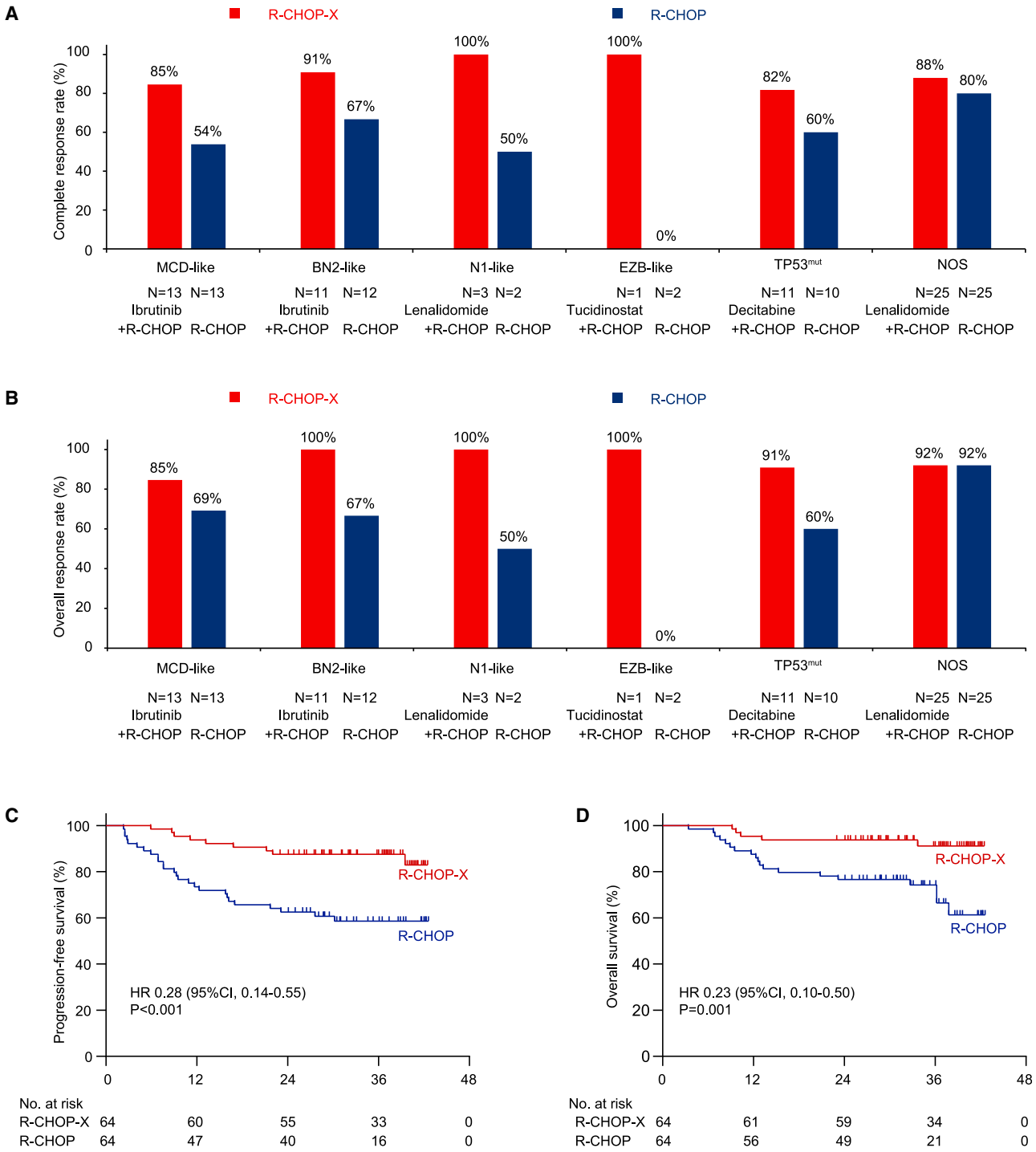



**Figure 2. Response and outcomes of GUIDANCE-01 cohort**

(A) Subgroup analysis of complete response rate according to genetic subtypes.

(B) Subgroup analysis of overall response rate according to genetic subtypes.

(C) Kaplan-Meier analysis of progression-free survival according to treatment arms. Survival curves were compared by log rank test. HRs and 95% CIs were estimated by Cox Proportional hazards model. A two-sided p value of <0.05 was considered statistically significant.

(D) Kaplan-Meier analysis of overall survival according to treatment arms. Survival curves were compared by log rank test. HRs and 95% CIs were estimated by Cox Proportional hazards model. A two-sided p value of <0.05 was considered statistically significant. See also Figure S2.

**Table 2. Common ≥ grade 3 treatment-related adverse events and adverse events of interest**

| | R-CHOP-X (n = 64) | | | R-CHOP (n = 64) | | |
|---|---|---|---|---|---|---|
| Adverse event | Grade 1-2 | Grade 3 | Grade 4 | Grade 1-2 | Grade 3 | Grade 4 |
| Hematological adverse events: n (%) | | | | | | |
| Neutropenia | 12 (19%) | 14 (22%) | 38 (59%) | 16 (25%) | 13 (20%) | 35 (55%) |
| Thrombocytopenia | 5 (8%) | 13 (20%) | 7 (11%) | 5 (8%) | 3 (5%) | 4 (6%) |
| Anemia | 13 (20%) | 16 (25%) | – | 14 (22%) | 13 (20%) | – |
| Febrile neutropenia | – | 13 (20%) | – | – | 7 (11%) | – |
| Non-hematological adverse events: n (%) | | | | | | |
| Infection | 10 (16%) | 4 (6%) | | 10 (16%) | 3 (5%) | |
| Nausea or vomiting | 13 (20%) | – | – | 14 (22%) | – | – |
| Increase in liver enzymes | 12 (19%) | 1 (2%) | – | 11 (17%) | 2 (3%) | – |
| Lung infection | 6 (9%) | 4 (6%) | – | 7 (11%) | 3 (5%) | – |
| Peripheral sensory neuropathy | 6 (9%) | – | | 7 (11%) | – | – |
| Rash | 2 (3%) | – | – | – | – | – |
| Interstitial lung disease | – | 2 (3%) | – | – | 1 (2%) | – |
| Gastrointestinal bleeding | – | 1 (2%) | – | – | 2 (3%) | – |
| Sepsis | – | 1 (2%) | – | – | – | – |
| Thromboembolic event | 1 (2%) | – | – | | – | – |

See also Table S2.
R-CHOP, rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone. R-CHOP-X, R-CHOP plus targeted agents.

in Figures S2A and S2B. Notable subgroups that did not show a clear benefit with R-CHOP-X included patient with 60 years of age or younger, patients with intermediate-low risk of IPI, patient with GCB subtype, and patients with NOS genetic subtype. Of 17 patients who died in this trial, 16 deaths were due to progressive disease, and 1 death due to heart failure and severe pulmonary infection 3 months after treatment.

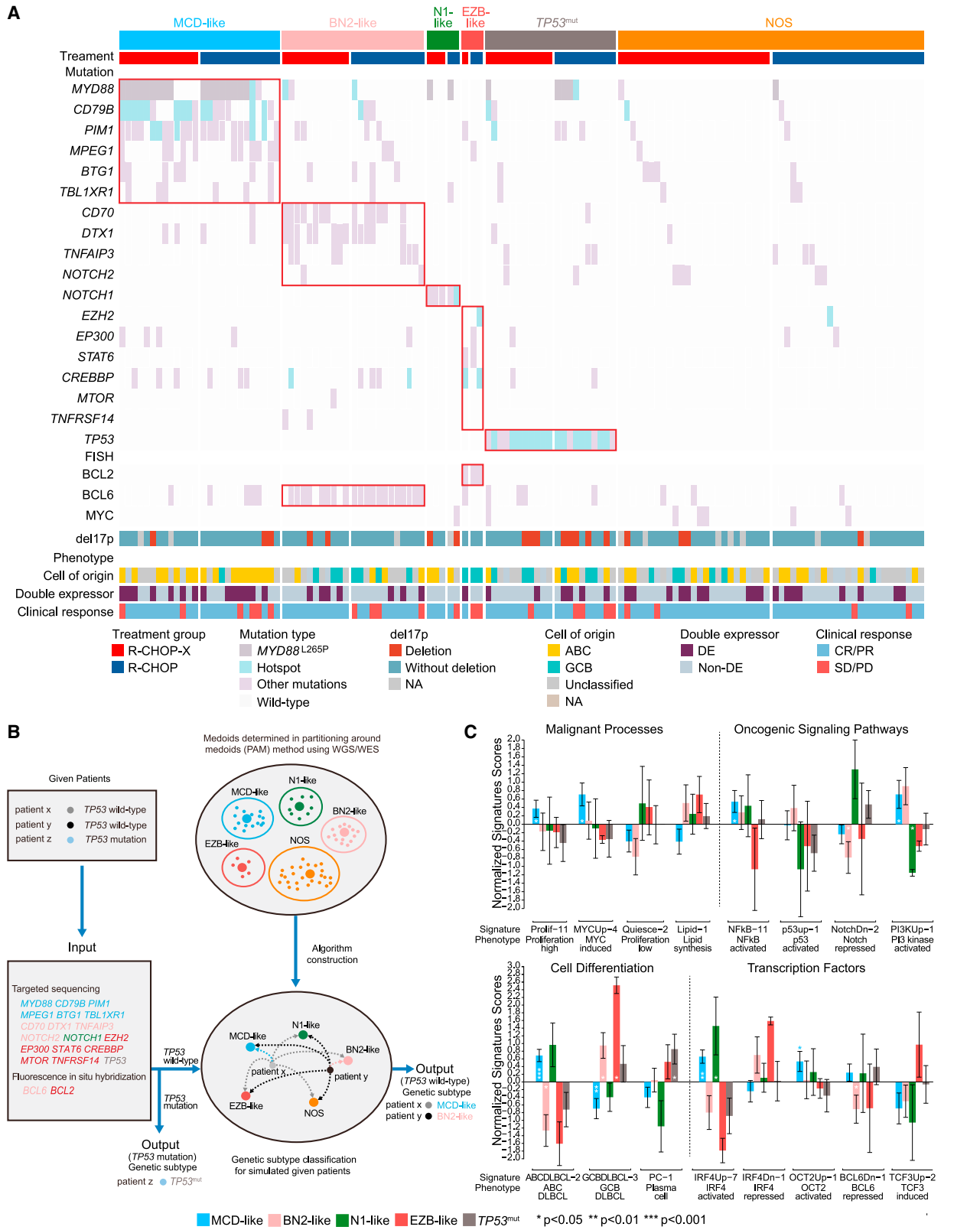
### *Subsequent treatment for lymphoma*

In the R-CHOP-X arm, among the 3 patients with partial response (PR), 2 received radiotherapy of the residual lesion revealed by final positron emission tomography-computed tomography evaluation, and 1 underwent splenectomy and was pathologically confirmed with DLBCL. Among the 3 patients with stable disease (SD), 2 received chimeric antigen receptor T cell (CAR T) therapy and achieved CR, and 1 died from disease progression. Two patients with progressive disease (PD) were salvaged with second-line chemotherapy and died from disease progression. In the R-CHOP arm, among the 5 patients with PR, 3 received radiotherapy and 2 received second-line treatment. Among the 7 patients with SD, 4 received second-line treatment followed by autologous hematopoietic stem cell transplantation and achieved CR, and 3 died from disease progression (including 1 received CAR T therapy). Among the 10 patients with PD, 3 received CAR T therapy and died from disease progression, and 7 received second-line therapy and 6 died from disease progression.

## Safety and dose intensity

Common hematological and non-hematological adverse events (AEs) are summarized in Table 2. For hematological toxicities, grade 3–4 neutropenia was the most common AEs in both groups (81% [52/64] in the R-CHOP-X arm and 75% [48/64] in the R-CHOP arm). Grade 3–4 thrombocytopenia was observed in 31% (20/64) of the R-CHOP-X arm and 11% (7/64) of the R-CHOP arm. Grade 3 anemia was observed in 25% (16/64) of the R-CHOP-X arm and 20% (13/64) of the R-CHOP arm. No grade 4 anemia was reported. Febrile neutropenia occurred in 20% (13/64) of the R-CHOP-X arm and 11% (7/64) of the R-CHOP arm and was of grade 3 in maximum. Despite the increased rates of cytopenia and thrombocytopenia, R-CHOP-X did not result in increased grade 3 pulmonary infection (6% [4/64] vs. 5% [3/64]) or gastrointestinal bleeding complications (2% [1/64] vs. 3% [2/64]), as compared to R-CHOP. Overall infection rates of all grades were 22% (14/64) and 20% (13/64) in R-CHOP-X and R-CHOP, respectively. No treatment-related mortality was observed.

Despite increased AE rates, similar dose intensity of chemotherapy was maintained between the R-CHOP-X and the R-CHOP arms (Table S2). In the R-CHOP-X arm, dose interruption of ibrutinib was reported in first 3 patients receiving ibrutinib 560 mg daily, 1 for grade 4 thrombocytopenia, 1 for grade 4 neutropenia, and another for grade 3 febrile neutropenia. After the protocol amendment in which ibrutinib was modified to 420 mg daily, these 3 patients were treated with ibrutinib 420 mg daily, and dose interruption of ibrutinib was reported in 24% (5/21) of the next 21 patients, 2 for grade 4 thrombocytopenia, 2 for grade 4 neutropenia, and 1 for grade 3 gastrointestinal bleeding, but no further dose reduction was reported. Dose interruption of lenalidomide was reported in 7% (2/28) of patients due to grade 4 neutropenia. Dose reduction of lenalidomide, tucidinostat, and decitabine was not reported. The proportion of patients who received ≥90% overall dose intensity was 79% (19/24) for ibrutinib, 93% for lenalidomide (26/28), 100% for tucidinostat (1/1), 100% for decitabine (11/11), 100% (64/64) for rituximab, vincristine, and prednisone, and 95% (61/64) for cyclophosphamide and doxorubicin. Rate of treatment discontinuation was 5% (3/64), of which 2 patients were due to stable disease at interim evaluation after 3 cycles, and 1 patient




(legend on next page)

due to heart failure after giant thymoma resection. In the R-CHOP arm, the proportion of patients who received ≥90% overall dose intensity was similar as that of the R-CHOP-X arm. Rate of R-CHOP discontinuation was 19% (12/64), of which 11 patients were due to stable or progressive disease at interim evaluation, and 1 patient because of pulmonary infection.

## Biomarker analysis

### *Simplified 20-gene algorithm for genetic subtyping*

To validate the simplified 20-gene algorithm, we analyzed the genetic features of the BC Cancer (BCC), Haematological Malignancy Research Network (HMRN), and National Cancer Institute (NCI) cohorts. As shown in Figure S3, excluding *TP53*$^{mut}$ cases from the simplified 20-gene algorithm, A53 cases and ST2 cases from the LymphGen algorithm, all the 4 subtypes (MCD-like, BN2-like, N1-like, and EZB-like subtypes in the simplified 20-gene algorithm; and MCD, BN2, N1, and EZB subtypes in the LymphGen algorithm) matched in 94% of patients in the BCC cohort, 93% of patients in the HMRN cohort, and 90% of patients in the NCI cohort. The main discordance occurred when patients assigned to the BN2 subtype by LymphGen were assigned to the MCD-like subtype by the simplified 20-gene algorithm (5 patients in the BCC cohort, 11 patients in the HMRN cohort, and 6 patients in the NCI cohort).

To further confirm our rationale in the selection of potential targeted agents in each subtype, we analyzed the genomic and transcriptomic data of tumor samples from newly diagnosed DLBCL patients previously reported,[20] including WES and whole-genome sequencing (WGS) of 325 patients and RNA-seq data available from 184 of the 325 patients. According to the simplified 20-gene algorithm, 184 patients with both WES/WGS and RNA-seq data available were categorized into MCD-like (n = 29), BN2-like (n = 24), N1-like (n = 9), EZB-like (n = 15), and *TP53*$^{mut}$ subtype (n = 23). Gene expression data showed that NF-κB activation (NFκB-11) was observed in MCD-like (p = 0.006) and BN2-like subtypes (p = 0.012), indicating the potential response to BTK inhibitor ibrutinib (Figure S4A). As revealed by gene set enrichment analysis, histone deacetylation signaling pathway was significantly upregulated in EZB-like subtype (p = 0.006, Figure S4B), indicating the potential response to HDAC inhibitor tucidinostat. Histone methyltransferase activity of histone H3 on lysine 9 (H3K9me3), which played an important epigenetic role on DLBCL progression through modulating endogenous retrovirus (ERV) expression, interferon-gamma (IFN-γ) production, and T cell activation,[21] was significantly upregulated in *TP53*$^{mut}$ subtype (p = 0.015), and related to downregulated IFN-γ production (p = 0.002), type I IFN production (p = 0.002), and T cell activation (p = 0.002) pathways (Figure S4C). Further detected by quantitative real-time PCR, ERV levels were significantly decreased in *TP53*$^{mut}$ subtype (ERVL-E $p < 0.001$, MER21C $p < 0.001$, HERVK11 p = 0.002, HERV16 p = 0.018, Figure S4D), as compared to those with wild-type *TP53*, suggesting the potential response to demethylating agent decitabine. Since no significant oncogenic signaling features were enriched in N1-like or NOS subtype, lenalidomide with broad anti-lymphoma activity was applied.

An overview of genomic signatures according to genetic subtypes for patients in GUIDANCE-01 cohort was provided in Figure 3A and Table S3. Twenty-one patients had *TP53*$^{mut}$ and referred as *TP53*$^{mut}$ subtype. Among patients with wild-type *TP53*, mutations in *MYD88*, *CD79B*, *PIM1*, *MPEG1*, *BTG1*, and *TBL1XR1* were enriched in patients with MCD-like subtype. Mutations of *CD70*, *DTX1*, *TNFAIP3*, and *NOTCH2* were enriched in patients with BN2-like subtype. N1-like subtype patients had typical *NOTCH1* mutations, and EZB-like subtype patients were detected with mutations of *EZH2*, *EP300*, *STAT6*, *CREBBP*, *MTOR*, and *TNFRSF14*. As detected by fluorescence *in situ* hybridization (FISH) analysis, 87% (20/23) of BN2-like patients presented with BCL6 re-arrangement. All of 3 EZB-like patients presented with BCL2 re-arrangement but without MYC re-arrangement. Schema diagram of our simplified 20-gene algorithm for genetic subtyping in this study is illustrated in Figure 3B.

Furthermore, we analyzed gene expression profile using RNA-seq data of 84 patients in this trial (Figure 3C and Table S4). With similar expression patterns observed in 184 patients used for establishing the simplified 20-gene algorithm, the subtypes exhibited distinct signatures in terms of malignant processes, B cell differentiation, and transcription factors. Briefly, MCD-like subtype expressed high level of cell proliferation (Prolif-11, p = 0.028), MYC oncoprotein (MYCUp-4, p = 0.010), ABC-DLBCL signatures (ABCDLBCL-2, $p < 0.001$), and genes induced by IRF4 (IRF4Up-7, p = 0.002), but lack of GCB-DLBCL signatures (GCBDLBCL-3, p = 0.004). EZB-like subtype expressed GCB-DLBCL signatures (GCBDLBCL-3, p = 0.018) and genes induced by TCF3 (TCF3Up-2). The subtypes also differed from oncogenic signaling pathways. NF-κB activation (NFκB-11) was highest in MCD-like (p = 0.042), and NOTCH repression (NotchDn-2) was lowest in BN2-like subtype (p = 0.045), while P53-targeted genes (p53up-1) were inhibited in *TP53*$^{mut}$ subtype. As detected by FISH analysis, deletion of 17p13 (del17p) was available in 117 of the 128 patients. Among these 117 patients, 19 were found to have del17p, including 8 patients with *TP53*$^{mut}$ (3 randomized to decitabine plus R-CHOP and 5 randomized to R-CHOP), 3 MCD-like (1 randomized to ibrutinib plus R-CHOP and 2 randomized to R-CHOP), 2 BN2-like (both randomized to ibrutinib plus R-CHOP), 2 N1-like (1 randomized to lenalidomide plus R-CHOP and 1 randomized to R-CHOP), and 4 NOS subtype (3 randomized to lenalidomide plus R-CHOP and 1 randomized to R-CHOP). No significant differences of PFS and OS were observed between R-CHOP-X arm and R-CHOP arm among these del17p patients.

### *Lymphoma microenvironment subtypes*

The LME categories were applied in 84 patients with RNA-seq data. Among all, 9 (11%), 27 (32%), 23 (27%), and 25 (30%)

**Figure 3. Genomic signatures according to genetic subtypes in GUIDANCE-01 cohort**

(A) Genomic alterations of patients with distinct genetic subtypes. Mutational status of 18 genes of the simplified 20-gene algorithm, 3 gene re-arrangements by FISH analysis, cell of origin classification by RNA-seq analysis, and MYC/BCL2 double expression status by immunohistochemistry are shown.

(B) Schema of the simplified 20-gene algorithm for DLBCL genetic subtype classification with simulated given patients.

(C) Gene expression signatures of patients with distinct genetic subtypes. Signature scores shown in bar plots were normalized with robust *Z* score method and error bars denote SEM. p values were derived from Wilcoxon rank-sum test. See also Figures S3, S4, Tables S3, and S4.

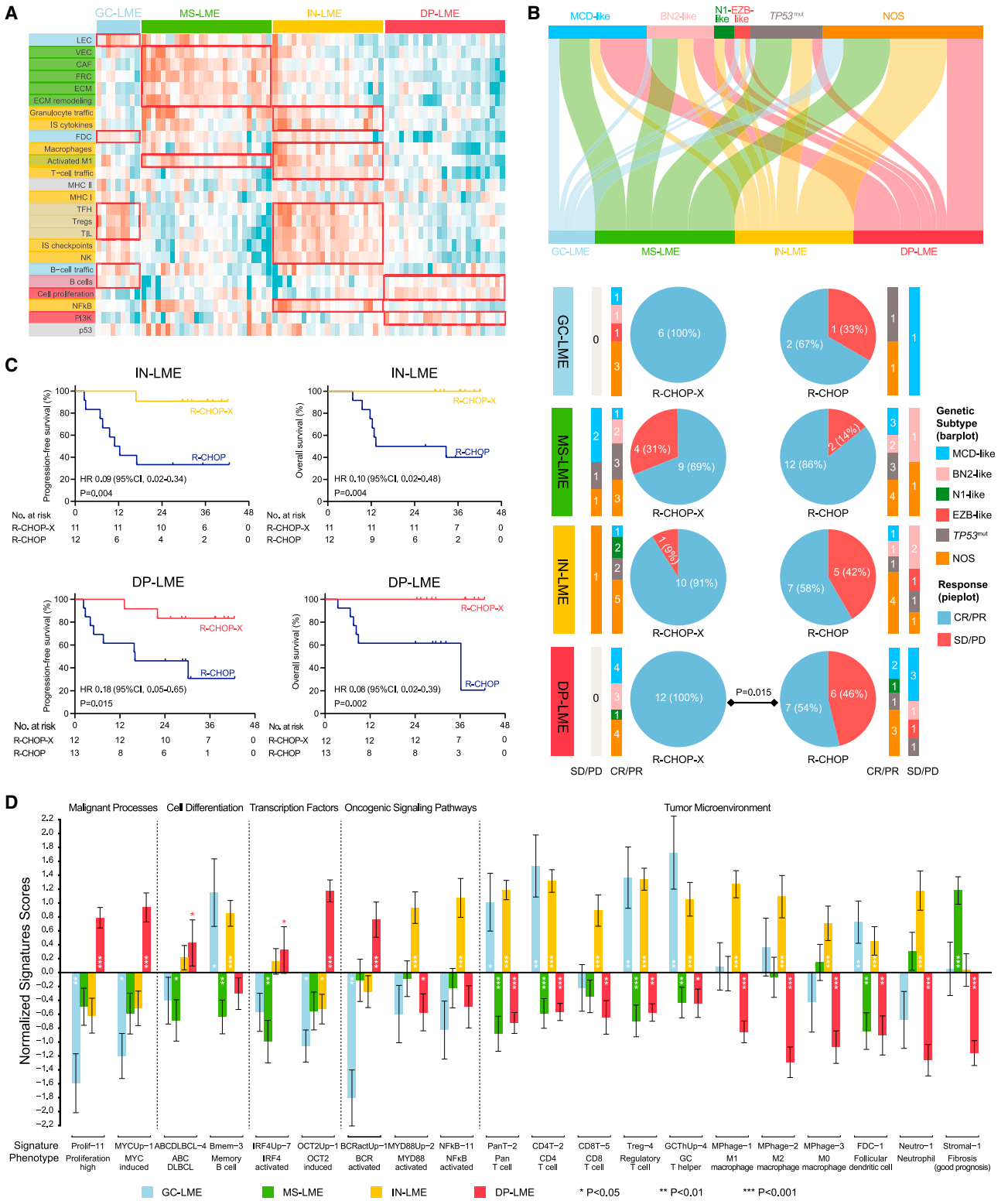



**Figure 4. Lymphoma microenvironment (LME) features according to genetic subtypes in GUIDANCE-01 cohort**

(A) Heatmap of 4 LME clusters of GUIDANCE-01 cohort.

(B) Association of LME categories and genetic subtypes, and related treatment response. Upper: Association of LME categories and genetic subtypes. No genetic subtype was significantly enriched in specific LME categories showed by Fisher's exact test ($p > 0.05$). Lower: The pie plot showed the number of patients with CR/PR or SD/PD in each LME subtype. The bar plots on its left/right side showed the corresponding genetic subtypes.

*(legend continued on next page)*

patients were categorized into GC-, MS-, IN-, and DP-LME, respectively (Figure 4A). No genetic subtype was significantly enriched in specific LME categories (Figure 4B). The ORRs were significantly higher in patients with DP-LME in the R-CHOP-X arm than those in the R-CHOP arm (100% vs. 54%, p = 0.015) (Figure 4B). PFS and OS of IN- and DP-LME patients in the R-CHOP-X arm were significantly superior to those in the R-CHOP arm (Figure 4C). We also identified signatures associated with LME subtypes (Figure 4D and Table S5). Proliferation signature (Prolif-11) and MYC activity (MYCUp-1) were increased in DP-LME ($p < 0.001$ and $p < 0.001$, respectively) and decreased in GC-LME (p = 0.004 and 0.013, respectively). Meanwhile, LME subtypes differed in cell differentiation, with DP-LME highly expressing activated B cells (ABCDLBCL-4, p = 0.020) and IN-LME expressing memory B cell signature (Bmem-3, $p < 0.001$). Accordingly, DP-LME highly expressed transcription factors IRF4 (IRF4Up-7, p = 0.019) and OCT2 (OCT2Up-1, $p < 0.001$). Among oncogenic signaling pathways, NF-κB (NFkB-11, $p < 0.001$) and MYD88-targeted genes (MYD88Up-2, $p < 0.001$) were upregulated in IN-LME, while BCR signature (BCRactUp-1, $p < 0.001$) was upregulated in DP-LME. In terms of tumor microenvironment, GC-LME expressed signature of follicular dendritic cells (FDC-1, p = 0.005), total T cells (PanT-2, p = 0.018), $CD4^+$ T cells (CD4T-2, p = 0.003), GC T helper cells (GCThUp-4, p = 0.004), and regulatory T cells (Treg-4, p = 0.007). MS-LME expressed stromal-related signature (Stromal-1, $p < 0.001$). IN-LME expressed signatures of $CD8^+$ T cells (CD8T-5, $p < 0.001$), neutrophils (Neutro-1, $p < 0.001$), and macrophages (MPhage-1, $p < 0.001$, M Phage-2, $p < 0.001$, and MPhage-3, $p < 0.001$). And DP-LME lacked all immune cell signatures.

## DISCUSSION

Recent molecular profiling studies lead to a novel and reproducible categorization of DLBCL genetic subgroups defined by specific oncogenic programs,[22] raising the possibility of designing a genetic subtype-guided targeted approach in treating DLBCL. Here, we reported the efficacy and safety of individualized immunochemotherapy guided by genetic subtypes (MCD-like, BN2-like, N1-like, EZB-like, *TP53*$^{mut}$, and NOS) in newly diagnosed DLBCL patients. This randomized phase 2 trial showed higher response rate and survival in the R-CHOP-X arm than in the R-CHOP arm. Meanwhile, the safety profile of R-CHOP-X is manageable, without new clinically significant or unexpected toxicity. The only protocol amendment was ibrutinib from 560 to 420 mg daily, which improved the intensity and the compliance of the immunochemotherapy. To prevent possible treatment delay caused by targeted sequencing analysis, the first cycle of R-CHOP was given to reduce the tumor burden, before a simplified genetic subtyping algorithm was applied to ensure the randomization at the second cycle of R-CHOP, thereby allowing rapid initiation of treatment and subsequent enrollment of high-risk patients. The percentage of intermediate-high or high-risk IPI patients was 65% in our GUIDANCE-01 trial, higher than recent major randomized trials in newly diagnosed DLBCL, such as GOYA, PHOENIX, REMoDL-B, and ROBUST trials.[2,4,5,23] Although this trial was not designed or powered to compare PFS or OS in patient subgroups, the exploratory subgroup analysis suggested a benefit with R-CHOP-X in patients older than 60 years of age, patients with intermediate-high or high risk of IPI, patients with non-GCB subtype as observed in POLARIX trial, and patients with classifiable genetic subtypes. Moreover, the simplified 20-gene algorithm we used can be output using targeted sequencing and FISH analysis instead of WES/WGS, which simplifies the process of data analysis and ensures its application in clinical trials and timely use of targeted agents based on genetic subtypes. Therefore, genetic subtype-guided targeted agents combined with R-CHOP are effective, safe, and clinically feasible in newly diagnosed DLBCL. With the encouraging results of the POLARIX trial,[7] this genetic subtype-guided treatment of targeted agents combined with Pola-R-CHP (Pola-R-CHP-X) may further improve the clinical outcomes for intermediate-risk or high-risk DLBCL patients.

Recent genomic studies illustrated molecular heterogeneity of DLBCL and delineated patients with distinct genetic subtypes.[9,24] We propose a simplified 20-gene algorithm to select targeted agents for individualized therapy. As revealed by post hoc RNA-seq, oncogenic signaling pathways varied from genetic subtypes were in accordance with the GenClass and the LymphGen models.[9,24] Interestingly, with the primary endpoint achieved, our study showed that genetic subtypes well corresponded to the sensitivity of targeted agents. Briefly, DLBCL with MCD-associated mutations are often sensitive to ibrutinib both in experimental models and in genomic analysis from the PHOENIX trial.[14,25] Ibrutinib is also potentially effective in N1 subtype.[25] BN2 subtype is related to the activation of NF-κB signaling pathway, and preclinical studies have shown that tumors belonging to this subtype are also suppressed by ibrutinib in murine xenografts.[24] However, we cannot rule out the discrepancy from the phenotypic heterogeneity of BN2, including ABC-, GCB-, and unclassified DLBCL.[9,10,24] Lenalidomide shows potential clinical effect in ABC-DLBCL and GCB-DLBCL, as reported in the ROBUST trial and the REMARC trial.[5,19] Lenalidomide is also effective in DLBCL with inflammatory gene alterations, including NOTCH pathway mutations.[26] These results may explain why we applied lenalidomide plus R-CHOP in N1-like and NOS subtypes. Moreover, EZB-like subtype is characterized by upregulated histone deacetylation signaling pathway and renders lymphoma cells sensitive to HDAC inhibitor tucidinostat.[27] The low proportion of EZB-like subtype was consistent with previous reports of relatively lower frequency of *EZH2* mutations in Chinese populations than in Western populations.[28,29] It is well established that decitabine is largely applied to treat MDS and AML patients who had *TP53* mutations.[17,30] In *TP53*$^{mut}$ AML, decitabine can induce IFN response

(C) Kaplan-Meier analysis of progression-free survival and overall survival according to LME categories with different treatment arms. HRs and 95% CIs were estimated by Cox Proportional hazards model. A two-sided p value of <0.05 was considered statistically significant.
(D) Gene expression signatures of patients with distinct LME subtypes. Signature scores shown in bar plots were normalized with robust *Z* score method and error bars denote SEM. p values were derived from Wilcoxon rank-sum test. See also Table S5.

via transcriptional activation of ERVs through the stimulator of IFN genes 1-dependent pathway.[31] In *TP53*$^{mut}$ DLBCL, we also found that the underlying mechanism of action of decitabine in combination with doxorubicin (a key cytotoxic agent in R-CHOP) was involved in an ERV-dependent manner, involving the inhibition of H3K9me3 occupancy on ERVs, subsequent enhancement of ERV activation, and unleashed IFN program.[21]

Lymphoma cells deploy genetic and epigenetic mechanisms to evade microenvironmental constraints of DLBCL. Decitabine also modulates the tumor microenvironment of *TP53*$^{mut}$ DLBCL through enhancing T cell-mediated antitumor response.[18] Patients with IN- and DP-LME, originally presented the worst prognosis upon R-CHOP treatment, benefited from R-CHOP-X (mainly treated with ibrutinib and lenalidomide). As expected, the IN-LME was composed of CD8$^+$ T cells, neutrophils, and macrophages, all of which could be targeted by ibrutinib and lenalidomide.[32–34] The DP-LME, contrasting with the other LMEs, was characterized by an overall lower presence of microenvironment-derived cells, but increased proliferation gene signatures.[13] Lenalidomide induces cell-cycle arrest and inhibits malignant B cell proliferation through regulation of cyclin-dependent kinases.[34] Meanwhile, our study showed that activation of BCR signaling pathway in DP-LME was counteracted by ibrutinib. Finally, it is worthy to point out that NOS patients could be further divided according to LME classification, providing an alternative option for targeted agents in this subset of DLBCL lack of genetic features.

### Limitations of the study

This study has limitations. First, the data were collected in a single institution, unblinded, and the relatively limited sample size might lead to possible selection bias. However, when we performed LME classification that was not prespecified, the distribution of LME subtypes remained balanced between R-CHOP-X and R-CHOP arms. Secondly, the individualization of targeted agents was not powered to detect a difference in each genetic subtype. A multicenter, randomized, phase 3 trial within each genetic subtype is currently ongoing (NCT05351346) to address this.

In conclusion, our findings demonstrate efficacy and safety of R-CHOP-X, a mechanism-based tailored therapy, which dually targets genetic and microenvironmental alterations in patients with newly diagnosed DLBCL.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Study design and participants
- METHOD DETAILS
  - Procedures
  - Outcomes
  - Sample processing
  - DNA-sequencing
  - RNA-sequencing
  - Sequence alignments and variant calling
  - Genetic subtyping
  - Gene expression quantification and preprocessing
  - Lymphoma microenvironment, gene expression signatures and cell of origin
  - Differential expression analysis and gene set enrichment analysis
  - Quantitative real-time PCR
  - Randomization and masking
- QUANTIFICATION AND STATISTICAL ANALYSIS

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.ccell.2023.09.004.

## ACKNOWLEDGMENTS

This study was supported, in part, by research funding from the National Key R&D Program of China (2022YFC2502600), National Natural Science Foundation of China (82130004, 81830007, 82170178, 82200201, and 82070204), Chang Jiang Scholars Program, Shanghai Municipal Education Commission Gaofeng Clinical Medicine Grant Support (20152206 and 20152208), Clinical Research Plan of Shanghai Hospital Development Center (SHDC2020CR1032B and SHDC2022CRD033), Multicenter Clinical Research Project by Shanghai Jiao Tong University School of Medicine (DLY201601), Collaborative Innovation Center of Systems Biomedicine, and the Samuel Waxman Cancer Research Foundation.

We truly thank Dr. Nikita Kotlov and Dr. Leandro Cerchietti for help with the LME classification, Dr. Chao-Fu Wang for help with pathological review, and Dr. Lei Dong for help with FISH analysis.

The computations in this paper were partially run on the Siyuan-1 cluster supported by the Center for High Performance Computing at Shanghai Jiao Tong University (SJTU-HPC), China. We thank Chen Li from SJTU-HPC for useful discussion of software and calculation.

## AUTHOR CONTRIBUTIONS

W.L.Z. conceived and designed the study as leading contact. M.C.Z., L.W., S.C., P.P.X., and W.L.Z. enrolled the patients and collected the data. M.C.Z. and S.T. performed bioinformatics investigation. D.F. and H.Y. were responsible for the simplified genetic subtyping algorithm. H.M.Y. was responsible for pathological review. X.F.J. and B.L. were responsible for PET-CT review. Q.S. was responsible for CT review. J.F.L., H.F., H.X., and S.J.C. gave technical support and advice. Y.Z. and Y.H. prepared biological samples. R.J.M. was responsible for supervision of study statistician. M.C.Z., S.T., P.P.X., and W.L.Z. interpreted the results and wrote the manuscript. All authors reviewed the manuscript critically and approved the content.

## DECLARATION OF INTERESTS

The authors declare that they have no competing interests.

Received: February 28, 2023
Revised: June 25, 2023
Accepted: September 5, 2023
Published: September 28, 2023

## REFERENCES

1. Siegel, R.L., Miller, K.D., and Jemal, A. (2019). Cancer statistics, 2019. CA A Cancer J. Clin. *69*, 7–34.

2. Vitolo, U., Trněný, M., Belada, D., Burke, J.M., Carella, A.M., Chua, N., Abrisqueta, P., Demeter, J., Flinn, I., Hong, X., et al. (2017). Obinutuzumab or Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone in Previously Untreated Diffuse Large B-Cell Lymphoma. J. Clin. Oncol. *35*, 3529–3537.

3. Xu, P.P., Fu, D., Li, J.Y., Hu, J.D., Wang, X., Zhou, J.F., Yu, H., Zhao, X., Huang, Y.H., Jiang, L., et al. (2019). Anthracycline dose optimisation in patients with diffuse large B-cell lymphoma: a multicentre, phase 3, randomised, controlled trial. Lancet. Haematol. *6*, e328–e337.

4. Younes, A., Sehn, L.H., Johnson, P., Zinzani, P.L., Hong, X., Zhu, J., Patti, C., Belada, D., Samoilova, O., Suh, C., et al. (2019). Randomized Phase III Trial of Ibrutinib and Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone in Non-Germinal Center B-Cell Diffuse Large B-Cell Lymphoma. J. Clin. Oncol. *37*, 1285–1295.

5. Nowakowski, G.S., Chiappella, A., Gascoyne, R.D., Scott, D.W., Zhang, Q., Jurczak, W., Özcan, M., Hong, X., Zhu, J., Jin, J., et al. (2021). ROBUST: A Phase III Study of Lenalidomide Plus R-CHOP Versus Placebo Plus R-CHOP in Previously Untreated Patients With ABC-Type Diffuse Large B-Cell Lymphoma. J. Clin. Oncol. *39*, 1317–1328.

6. Nowakowski, G.S., Hong, F., Scott, D.W., Macon, W.R., King, R.L., Habermann, T.M., Wagner-Johnston, N., Casulo, C., Wade, J.L., Nagargoje, G.G., et al. (2021). Addition of Lenalidomide to R-CHOP Improves Outcomes in Newly Diagnosed Diffuse Large B-Cell Lymphoma in a Randomized Phase II US Intergroup Study ECOG-ACRIN E1412. J. Clin. Oncol. *39*, 1329–1338.

7. Tilly, H., Morschhauser, F., Sehn, L.H., Friedberg, J.W., Trněný, M., Sharman, J.P., Herbaux, C., Burke, J.M., Matasar, M., Rai, S., et al. (2022). Polatuzumab Vedotin in Previously Untreated Diffuse Large B-Cell Lymphoma. N. Engl. J. Med. *386*, 351–363.

8. Sehn, L.H., and Salles, G. (2021). Diffuse Large B-Cell Lymphoma. N. Engl. J. Med. *384*, 842–858.

9. Schmitz, R., Wright, G.W., Huang, D.W., Johnson, C.A., Phelan, J.D., Wang, J.Q., Roulland, S., Kasbekar, M., Young, R.M., Shaffer, A.L., et al. (2018). Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. N. Engl. J. Med. *378*, 1396–1407.

10. Chapuy, B., Stewart, C., Dunford, A.J., Kim, J., Kamburov, A., Redd, R.A., Lawrence, M.S., Roemer, M.G.M., Li, A.J., Ziepert, M., et al. (2018). Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. Nat. Med. *24*, 679–690.

11. Lacy, S.E., Barrans, S.L., Beer, P.A., Painter, D., Smith, A.G., Roman, E., Cooke, S.L., Ruiz, C., Glover, P., Van Hoppe, S.J.L., et al. (2020). Targeted sequencing in DLBCL, molecular subtypes, and outcomes: a Haematological Malignancy Research Network report. Blood *135*, 1759–1771.

12. Mishina, T., Oshima-Hasegawa, N., Tsukamoto, S., Fukuyo, M., Kageyama, H., Muto, T., Mimura, N., Rahmutulla, B., Nagai, Y., Kayamori, K., et al. (2021). Genetic subtype classification using a simplified algorithm and mutational characteristics of diffuse large B-cell lymphoma in a Japanese cohort. Br. J. Haematol. *195*, 731–742.

13. Kotlov, N., Bagaev, A., Revuelta, M.V., Phillip, J.M., Cacciapuoti, M.T., Antysheva, Z., Svekolkin, V., Tikhonova, E., Miheecheva, N., Kuzkina, N., et al. (2021). Clinical and Biological Subtypes of B-cell Lymphoma Revealed by Microenvironmental Signatures. Cancer Discov. *11*, 1468–1489.

14. Wilson, W.H., Young, R.M., Schmitz, R., Yang, Y., Pittaluga, S., Wright, G., Lih, C.J., Williams, P.M., Shaffer, A.L., Gerecitano, J., et al. (2015). Targeting B cell receptor signaling with ibrutinib in diffuse large B cell lymphoma. Nat. Med. *21*, 922–926.

15. Andersen, C.L., Asmar, F., Klausen, T., Hasselbalch, H., and Grønbæk, K. (2012). Somatic mutations of the CREBBP and EP300 genes affect response to histone deacetylase inhibition in malignant DLBCL clones. Leuk. Res. Rep. *2*, 1–3.

16. Sun, Y., Gao, Y., Chen, J., Huang, L., Deng, P., Chen, J., Chai, K.X.Y., Hong, J.H., Chan, J.Y., He, H., et al. (2021). CREBBP cooperates with the cell cycle machinery to attenuate chidamide sensitivity in relapsed/refractory diffuse large B-cell lymphoma. Cancer Lett. *521*, 268–280.

17. Welch, J.S., Petti, A.A., Miller, C.A., Fronick, C.C., O'Laughlin, M., Fulton, R.S., Wilson, R.K., Baty, J.D., Duncavage, E.J., Tandon, B., et al. (2016). TP53 and Decitabine in Acute Myeloid Leukemia and Myelodysplastic Syndromes. N. Engl. J. Med. *375*, 2023–2036.

18. Zhang, M.C., Fang, Y., Xu, P.P., Dong, L., Shen, R., Huang, Y.H., Fu, D., Yan, Z.X., Cheng, S., Jiang, X.F., et al. (2021). Clinical efficacy and tumour microenvironment influence of decitabine plus R-CHOP in patients with newly diagnosed diffuse large B-Cell lymphoma: Phase 1/2 and biomarker study. Clin. Transl. Med. *11*, e584.

19. Thieblemont, C., Tilly, H., Gomes da Silva, M., Casasnovas, R.O., Fruchart, C., Morschhauser, F., Haioun, C., Lazarovici, J., Grosicka, A., Perrot, A., et al. (2017). Lenalidomide Maintenance Compared With Placebo in Responding Elderly Patients With Diffuse Large B-Cell Lymphoma Treated With First-Line Rituximab Plus Cyclophosphamide, Doxorubicin, Vincristine, and Prednisone. J. Clin. Oncol. *35*, 2473–2481.

20. Shen, R., Xu, P.P., Wang, N., Yi, H.M., Dong, L., Fu, D., Huang, J.Y., Huang, H.Y., Janin, A., Cheng, S., et al. (2020). Influence of oncogenic mutations and tumor microenvironment alterations on extranodal invasion in diffuse large B-cell lymphoma. Clin. Transl. Med. *10*, e221.

21. Fang, Y., M.C.Z., He, Y., Li, C., Fang, H., Xu, P.P., Cheng, S., Zhao, Y., Feng, Y., Liu, Q., et al. (2023). Human endogenous retroviruses as epigenetic therapeutic targets in TP53-mutated diffuse large B-cell lymphoma. Signal Transduct. Targeted Ther. https://doi.org/10.1038/s41392-023-01626-x.

22. Morin, R.D., Arthur, S.E., and Hodson, D.J. (2022). Molecular profiling in diffuse large B-cell lymphoma: why so many types of subtypes? Br. J. Haematol. *196*, 814–829.

23. Davies, A., Cummin, T.E., Barrans, S., Maishman, T., Mamot, C., Novak, U., Caddy, J., Stanton, L., Kazmi-Stokes, S., McMillan, A., et al. (2019). Gene-expression profiling of bortezomib added to standard chemoimmunotherapy for diffuse large B-cell lymphoma (REMoDL-B): an open-label, randomised, phase 3 trial. Lancet Oncol. *20*, 649–662.

24. Wright, G.W., Huang, D.W., Phelan, J.D., Coulibaly, Z.A., Roulland, S., Young, R.M., Wang, J.Q., Schmitz, R., Morin, R.D., Tang, J., et al. (2020). A Probabilistic Classification Tool for Genetic Subtypes of Diffuse Large B Cell Lymphoma with Therapeutic Implications. Cancer Cell *37*, 551–568.e14.

25. Wilson, W.H., Wright, G.W., Huang, D.W., Hodkinson, B., Balasubramanian, S., Fan, Y., Vermeulen, J., Shreeve, M., and Staudt, L.M. (2021). Effect of ibrutinib with R-CHOP chemotherapy in genetic subtypes of DLBCL. Cancer Cell *39*, 1643–1653.e3.

26. Hartert, K.T., Wenzl, K., Krull, J.E., Manske, M., Sarangi, V., Asmann, Y., Larson, M.C., Maurer, M.J., Slager, S., Macon, W.R., et al. (2021). Targeting of inflammatory pathways with R2CHOP in high-risk DLBCL. Leukemia *35*, 522–533.

27. Guan, X.W., Wang, H.Q., Ban, W.W., Chang, Z., Chen, H.Z., Jia, L., and Liu, F.T. (2020). Novel HDAC inhibitor Chidamide synergizes with Rituximab to inhibit diffuse large B-cell lymphoma tumour growth by upregulating CD20. Cell Death Dis. *11*, 20.

28. de Miranda, N.F.C.C., Georgiou, K., Chen, L., Wu, C., Gao, Z., Zaravinos, A., Lisboa, S., Enblad, G., Teixeira, M.R., Zeng, Y., et al. (2014). Exome sequencing reveals novel mutation targets in diffuse large B-cell lymphomas derived from Chinese patients. Blood *124*, 2544–2553.

29. Ren, W., Li, W., Ye, X., Liu, H., and Pan-Hammarström, Q. (2017). Distinct subtype distribution and somatic mutation spectrum of lymphomas in East Asia. Curr. Opin. Hematol. *24*, 367–376.

30. Greve, G., Schüler, J., Grüning, B.A., Berberich, B., Stomper, J., Zimmer, D., Gutenkunst, L., Bönisch, U., Meier, R., Blagitko-Dorfs, N., et al. (2021). Decitabine Induces Gene Derepression on Monosomic Chromosomes: In Vitro and In Vivo Effects in Adverse-Risk Cytogenetics AML. Cancer Res. *81*, 834–846.

31. Kogan, A.A., Topper, M.J., Dellomo, A.J., Stojanovic, L., McLaughlin, L.J., Creed, T.M., Eberly, C.L., Kingsbury, T.J., Baer, M.R., Kessler, M.D., et al.

(2022). Activating STING1-dependent immune signaling in TP53 mutant and wild-type acute myeloid leukemia. Proc. Natl. Acad. Sci. USA *119*, e2123227119.

32. Zhu, S., Gokhale, S., Jung, J., Spirollari, E., Tsai, J., Arceo, J., Wu, B.W., Victor, E., and Xie, P. (2021). Multifaceted Immunomodulatory Effects of the BTK Inhibitors Ibrutinib and Acalabrutinib on Different Immune Cell Subsets - Beyond B Lymphocytes. Front. Cell Dev. Biol. *9*, 727531.
33. Chiu, H., Trisal, P., Bjorklund, C., Carrancio, S., Toraño, E.G., Guarinos, C., Papazoglou, D., Hagner, P.R., Beldi-Ferchiou, A., Tarte, K., et al. (2019). Combination lenalidomide-rituximab immunotherapy activates anti-tumour immunity and induces tumour cell death by complementary mechanisms of action in follicular lymphoma. Br. J. Haematol. *185*, 240–253.
34. Gribben, J.G., Fowler, N., and Morschhauser, F. (2015). Mechanisms of Action of Lenalidomide in B-Cell Non-Hodgkin Lymphoma. J. Clin. Oncol. *33*, 2803–2811.
35. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.
36. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.
37. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. GigaScience *10*, giab008.
38. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.
39. Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. *38*, e164.
40. Fu, Y., Liu, Z., Lou, S., Bedford, J., Mu, X.J., Yip, K.Y., Khurana, E., and Gerstein, M. (2014). FunSeq2: a framework for prioritizing noncoding regulatory variants in cancer. Genome Biol. *15*, 480.
41. Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. *34*, 525–527.
42. Soneson, C., Love, M.I., and Robinson, M.D. (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res. *4*, 1521.
43. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47.
44. Swerdlow, S.H., Campo, E., Pileri, S.A., Harris, N.L., Stein, H., Siebert, R., Advani, R., Ghielmini, M., Salles, G.A., Zelenetz, A.D., and Jaffe, E.S. (2016). The 2016 revision of the World Health Organization classification of lymphoid neoplasms. Blood *127*, 2375–2390.
45. Scott, D.W., Wright, G.W., Williams, P.M., Lih, C.J., Walsh, W., Jaffe, E.S., Rosenwald, A., Campo, E., Chan, W.C., Connors, J.M., et al. (2014). Determining cell-of-origin subtypes of diffuse large B-cell lymphoma using gene expression in formalin-fixed paraffin-embedded tissue. Blood *123*, 1214–1217.
46. Nowakowski, G.S., LaPlant, B., Macon, W.R., Reeder, C.B., Foran, J.M., Nelson, G.D., Thompson, C.A., Rivera, C.E., Inwards, D.J., Micallef, I.N., et al. (2015). Lenalidomide combined with R-CHOP overcomes negative prognostic impact of non-germinal center B-cell phenotype in newly diagnosed diffuse large B-Cell lymphoma: a phase II study. J. Clin. Oncol. *33*, 251–257.
47. Morschhauser, F., Feugier, P., Flinn, I.W., Gasiorowski, R., Greil, R., Illés, Á., Johnson, N.A., Larouche, J.F., Lugtenburg, P.J., Patti, C., et al. (2021). A phase 2 study of venetoclax plus R-CHOP as first-line treatment for patients with diffuse large B-cell lymphoma. Blood *137*, 600–609.
48. Zhang, M.C., Fang, Y., Wang, L., Cheng, S., Fu, D., He, Y., Zhao, Y., Wang, C.F., Jiang, X.F., Song, Q., et al. (2020). Clinical efficacy and molecular biomarkers in a phase II study of tucidinostat plus R-CHOP in elderly patients with newly diagnosed diffuse large B-cell lymphoma. Clin. Epigenetics *12*, 160.
49. Cheson, B.D., Fisher, R.I., Barrington, S.F., Cavalli, F., Schwartz, L.H., Zucca, E., Lister, T.A., et al.; Alliance Australasian Leukaemia and Lymphoma Group; European Mantle Cell Lymphoma Consortium; Eastern Cooperative Oncology Group (2014). Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. J. Clin. Oncol. *32*, 3059–3068.
50. Huang, Y.H., Cai, K., Xu, P.P., Wang, L., Huang, C.X., Fang, Y., Cheng, S., Sun, X.J., Liu, F., Huang, J.Y., et al. (2021). CREBBP/EP300 mutations promoted tumor progression in diffuse large B-cell lymphoma through altering tumor-associated macrophage polarization via FBXW7-NOTCH-CCL2/CSF1 axis. Signal Transduct. Target. Ther. *6*, 10.
51. Qin, W., Fu, D., Shi, Q., Dong, L., Yi, H., Huang, H., Jiang, X., Song, Q., Liu, Z., Cheng, S., et al. (2021). Molecular Heterogeneity in Localized Diffuse Large B-Cell Lymphoma. Front. Oncol. *11*, 638757.
52. Li, M.M., Datto, M., Duncavage, E.J., Kulkarni, S., Lindeman, N.I., Roy, S., Tsimberidou, A.M., Vnencak-Jones, C.L., Wolff, D.J., Younes, A., and Nikiforova, M.N. (2017). 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. *19*, 4–23.
53. Ennishi, D., Jiang, A., Boyle, M., Collinge, B., Grande, B.M., Ben-Neriah, S., Rushton, C., Tang, J., Thomas, N., Slack, G.W., et al. (2019). Double-Hit Gene Expression Signature Defines a Distinct Subgroup of Germinal Center B-Cell-Like Diffuse Large B-Cell Lymphoma. J. Clin. Oncol. *37*, 190–201.
54. Phelan, J.D., Young, R.M., Webster, D.E., Roulland, S., Wright, G.W., Kasbekar, M., Shaffer, A.L., 3rd, Ceribelli, M., Wang, J.Q., Schmitz, R., et al. (2018). A multiprotein supercomplex controlling oncogenic signalling in lymphoma. Nature *560*, 387–391.
55. Bolstad, B.M., Irizarry, R.A., Astrand, M., and Speed, T.P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics *19*, 185–193.
56. Barbie, D.A., Tamayo, P., Boehm, J.S., Kim, S.Y., Moody, S.E., Dunn, I.F., Schinzel, A.C., Sandy, P., Meylan, E., Scholl, C., et al. (2009). Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature *462*, 108–112.
57. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation *2*, 100141.
58. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.
59. Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J.P. (2011). Molecular signatures database (MSigDB) 3.0. Bioinformatics *27*, 1739–1740.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Biological samples | | |
| Patient tumor biopsies from GUIDANCE-01 trial | This paper | NCT04025593 |
| Critical commercial assays | | |
| GeneRead DNA FFPE Tissue Kit | Qiagen | Cat# 56704 |
| QIAamp DNA Mini Kit | Qiagen | Cat# 51306 |
| RNeasy Mini Kit | Qiagen | Cat# 74104 |
| VAHTS mRNA-seq V3 Library Prep Kit | Vazyme | Cat# NA210-01 |
| FreeZol Reagent | Vazyme | Cat# R711-01 |
| HiScript III RT SuperMix for qPCR (+gDNA wiper) | Vazyme | Cat# R323-01 |
| ChamQ Universal SYBR qPCR Master Mix | Vazyme | Cat# Q711-03 |
| Dynabeads MyOne Streptavidin T1 | Thermo Fisher | Cat# 65601 |
| Qubit RNA HS Assay Kit | Thermo Fisher | Cat# Q32852 |
| Deposited data | | |
| Human genome reference GRCh37 | Genome Reference Consortium | http://hgdownload.soe.ucsc.edu/ |
| Human transcriptome reference GRCh38 | Ensembl | http://ftp.ensembl.org/pub/release-106/fasta/homo_sapiens/cdna/ |
| 1000 Genomes project data | The International Genome Sample Resource | RRID: SCR_006828 |
| COSMIC v77 | COSMIC | https://cancer.sanger.ac.uk/cosmic |
| Molecular Signatures Database | Liberzon et al.[28] | https://www.gsea-msigdb.org/gsea/msigdb |
| Raw RNA-seq data | This paper | GSA-Human database (HRA003285) |
| Software and algorithms | | |
| FastQC (version 1.11.4) | Babraham Institute | https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ |
| Burrows-Wheeler Aligner (version 0.5.9-tpx) | Li and Durbin[35] | https://bio-bwa.sourceforge.net/ |
| Trimmomatic (version 3.6) | Bolger et al.[36] | https://github.com/usadellab/Trimmomatic |
| Samtools (v0.1.18), | Danecek et al.[37] | https://github.com/samtools/samtools |
| Picard (v1.93) | Broad Institute | https://broadinstitute.github.io/picard/ |
| Genome Analysis Toolkit (v4.1.4.0) | McKenna et al.[38] | https://gatk.broadinstitute.org/hc/en-us |
| Annovar software (version 2017-07-17). | Wang et al.[39] | https://annovar.openbioinformatics.org/en/latest/ |
| Funseq 2.1.2 | Fu et al.[40] | https://github.com/gersteinlab/FunSeq2 |
| Kallisto (v0.46.0) | Bray et al.[41] | https://github.com/pachterlab/kallisto |
| R package tximport (v1.24.0) | Soneson et al.[42] | https://www.bioconductor.org/packages/release/bioc/html/tximport.html |
| R package preprocessCore (v1.58.0) | Bioconductor | https://www.bioconductor.org/packages/release/bioc/html/preprocessCore.html |
| R package limma (v3.48.3) | Ritchie et al.[43] | https://www.bioconductor.org/packages/release/bioc/html/limma.html |
| GSEA software | Subramanian et al.[27] | https://software.broadinstitute.org/cancer/software/gsea/wiki/index.php/Main_Page |
| SPSS software (v23.0) | SPSS Inc. | https://www.ibm.com/spss |
| GraphPad Prism software (v8.0.2) | GraphPad Prism | https://www.graphpad.com/ |
| LymphGen algorithm | Wright et al.[24] | Wright et al.[24] |
| LME | Kotlov et al.[13] | https://github.com/BostonGene/LME |
| Simplified 20-gene algorithm | This paper | https://doi.org/10.5281/zenodo.7648475 |

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Wei-Li Zhao (zhao.weili@yahoo.com).

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

- Raw RNA-seq data have been deposited at GSA-Human database (Accession number HRA003285). Accession numbers are listed in the key resources table.
- The original R code of the simplified 20-gene algorithm for genetic subtyping has been deposited at Zenodo. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Study design and participants

This was an open label, single-center, randomized, phase 2 trial (Figure S1). Patients aged 18–80 years with newly diagnosed, histologically confirmed CD20-positive DLBCL, Eastern Cooperative Oncology Group (ECOG) performance status of 0–2, IPI risk of intermediate or high (IPI $\geq$2), and adequate organ function were eligible for study participation. Exclusion criteria were primary central nervous system lymphoma, hepatitis B virus (HBV-DNA) or human immunodeficiency virus positivity, other active malignancy or serious medical conditions that could interfere with study treatment. The study protocol is provided as Data S1.

The study was approved by ethics committee and institutional review board of Shanghai Ruijin Hospital. Written informed consent was obtained from all participants in accordance with the Declaration of Helsinki. The study was registered on ClinicalTrial.gov (NCT04025593).

Pathological diagnosis was established according to the 2016 World Health Organization classification.[44] FISH of BCL2, BCL6, and MYC re-arrangements were performed for each patient. FISH of *TP53* gene (chromosome 17p13) deletion was performed for 117 of the 128 patients with adequate tumor tissues. The diagnosis of DLBCL was reviewed and confirmed by experienced pathologists (H.M.Y. and C.F.W.). Cell of origin subtype was assessed by RNA expression using Lymph2Cx assay for 84 of the 128 patients with RNA-seq data.[45]

## METHOD DETAILS

### Procedures

All patients were treated with 1 cycle of standard R-CHOP regimen (rituximab 375 mg/m$^2$ intravenously on Day 0, cyclophosphamide 750 mg/m$^2$, doxorubicin 50 mg/m$^2$, and vincristine 1.4 mg/m$^2$ [maximum 2.0 mg] intravenously on Day 1, prednisone 60 mg/m$^2$ [maximum 100 mg] orally on Days 1–5 of every 21-day cycle) after diagnosis, and were stratified by genetic subtypes and randomly assigned 1:1 to R-CHOP-X or R-CHOP for the remaining 5 cycles. In R-CHOP-X arm, patients with MCD-like and BN2-like subtypes received oral ibrutinib 560 mg on Days 1–21 of each cycle. Considering high frequency of severe adverse events of elderly patients in previously reported randomized phase 3 study[4] and safety profiles observed in the early enrolled 3 patients receiving ibrutinib, protocol amendment was made in August 2019, in which the dose of ibrutinib was changed from 560 mg to 420 mg daily. Patients with N1-like and NOS subtypes received oral lenalidomide 25 mg on Days 1–10 of each cycle, the dose and schedule of which were consistent with previous clinical trials in newly diagnosed DLBCL.[46,47] Patients with EZB-like subtype received oral tucidinostat 20 mg twice a week for 2 weeks of each cycle, the dose and schedule of which were consistent with our previous phase II trial in newly diagnosed DLBCL.[48] Patients with *TP53*$^{mut}$ subtype received intravenous decitabine 10 mg/m$^2$ Days $-5$ to $-1$ followed by standard R-CHOP of every 21-day cycle, the dose and schedule of which were consistent with our previous phase II trial in newly diagnosed DLBCL.[18]

G-CSF prophylaxis (pegfilgrastim 6 mg subcutaneously) was given from the second cycle of chemotherapy if grade $\geq$3 neutropenia was present in the first cycle. Primary prophylaxis of acyclovir and sulfamethoxazole were given to prevent cytomegalovirus and pneumocystis jiroveci infection. Lamivudine or entecavir was administered in occult carriers of HBV to prevent HBV reactivation. Tumor lysis prophylaxis and radiation therapy were performed for patients with bulky disease, or with residual disease at the end of treatment, at the discretion of physicians. Absolute neutrophil count >1,500/mL and platelet >75,000/mL at the time of retreatment were required, respectively. Targeted agents, including ibrutinib, lenalidomide, and tucidinostat, were not held or reduced midcycle regardless of grade 3 non-hematological AEs, grade 4 hematological AEs, or febrile neutropenia. Decitabine was reduced in case of treatment delay for more than 7 days because of grade 3 non-hematological AEs, grade 4 hematological AEs, or febrile neutropenia.

### Outcomes

The primary endpoint was CRR assessed by PET-CT at the end of the treatment. Secondary endpoints were PFS, OS, ORR, and safety. PET-CT scans were independently reviewed by nuclear medicine physicians (X.F.J. and B.L.) who were blinded to patient randomization. Treatment response was assessed according to Lugano 2014 criteria.[49] Interim efficacy was evaluated by PET-CT after 3 cycles. Patients who had achieved a complete or partial response received another 3 cycles. Patients who did not achieve CR or PR stopped receiving R-CHOP-X or R-CHOP at this point and were required with tumor biopsy. Final evaluation was performed by PET-CT at the end of the last cycle of treatment. CT of the neck, thorax, abdomen, and pelvis was repeated every 3 months thereafter to monitor disease progression until 1 year, then every 6 months until 2 years, and every year until 5 years. PFS was measured from the time of random assignment to date of progression, relapse, death from any cause, or date of last follow-up. OS was measured from the time of random assignment to death of any cause or date of last follow-up. Patients were monitored weekly for AEs, which were graded for severity with the National Cancer Institute Common Terminology Criteria for Adverse Events, Version 4.0 (NCI CTCAE v4.0).

### Sample processing

Formalin-fixed paraffin-embedded (FFPE) tumor biopsy samples were obtained from all patients before treatment for pathology review and targeted DNA sequencing. Fresh frozen samples were preserved for patients with residual tissue available from the baseline biopsies for post-hoc RNA-seq.

Genomic DNA was extracted from FFPE samples using a GeneRead DNA FFPE Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's descriptions. Total RNA was extracted from fresh frozen samples using Trizol and a RNeasy Mini Kit (Qiagen).

Before sequencing, genomic DNA was cut into fragments with ~200 bp using a focused ultrasonicator (No. M220, Covaris, Woburn, MA, USA). DNA quantity was determined using a Nanodrop 8,000 UV-Vis spectrometer (NanoDrop Technologies, Wilmington, DE, USA), Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA), and 2200 TapeStation Instrument (Agilent Technologies, Santa Clara, CA, USA). RNA quality was assessed by RNA 6000 NanoChip using 2100 Bioanalyzer (Agilent Technologies) and RNA concentration was measured using Qubit RNA HS Assay Kit by Qubit 2.0 Fluorometer (Life Technologies, Grand Island, NY, USA) according to the manufacturer's instructions.

### DNA-sequencing

Targeted sequencing was performed on FFPE tumor samples of all patients using NovaSeq (Illumina, San Diego, CA, USA), covering the CDS regions of 18 genes of the simplified 20-gene algorithm (*BTG1*, *CD70*, *CD79B*, *CREBBP*, *DTX1*, *EP300*, *EZH2*, *MPEG1*, *MTOR*, *MYD88*, *NOTCH1*, *NOTCH2*, *PIM1*, *STAT6*, *TBL1XR1*, *TNFAIP3*, *TNFRSF14*, and *TP53*). Library was constructed by lymphoma-associated gene mutation detection kit (Shanghai Rightongene Biotechnology Co., Ltd., Shanghai, China).

### RNA-sequencing

RNA-seq was performed on qualified frozen tumor samples of 84 patients. Double-stands cDNA was synthesized with the VAHTS mRNA-seq V3 Library Prep Kit for Illumina (Vazyme, Nanjing, China), purified and submitted to the pre-PCR reaction to establish the pre-library. Qualified DNA library was captured using Dynabeads MyOne Streptavidin T1 magnetic beads, followed by post-PCR reaction to obtain the final library. After quality control, High-throughput RNA-seq was performed using Illumina NovaSeq dual-end sequencing (2 × 150 bp) (Illumina, California, USA).

### Sequence alignments and variant calling

For all types of sequencing data, the FastQC (version 1.11.4, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) software was used to assess the quality of raw sequencing data.

The paired-end reads of targeted sequencing were aligned to the Human Genome Reference Consortium build 37 (GRCh37) using Burrows-Wheeler Aligner (BWA, version 0.5.9-tpx). The average depth of 69 capture and 59 amplicon targeted sequencing data were 335.5X (IQR 215.75X - 488.5X) and 1998X (IQR 1431X - 2736.5X), respectively.

Raw sequencing data were processed by using Trimmomatic (version 3.6) software to remove sequencing adapters and low-quality reads referring to the joint sequence fragments of the 3′ end and low-quality fragments with Q value <25 and fragments with <35 bp. Samtools (v0.1.18),[37] picard (v1.93), and Genome Analysis Toolkit (GATK, v4.1.4.0) were used for BAM file handling, local realignment, base recalibration and calling variants, respectively. Mutations in the coding region were annotated using the Annovar software (version 2017-07-17).[39]

Variants with depth >10 and VAF >0.05 were first filtered with variants detected by pooled results of 42 peripheral blood samples used as control samples in our previous study.[20] Then, the non-synonymous SNVs, indels and splicing sites were preserved with following criteria: (1) reported as somatic mutations in our previous studies[20,50,51]; (2) commonly considered as hotspot mutations like $MYD88^{L265P}$; (3) categorized into tier I and II variants according to Guideline for Evidence-Based Categorization of Somatic Variants[52]; (4) not observed in 1000 genome (1KG) project, observed in less than 0.01 in databases of 1KG project, or observed in less than 0.05 in databases of 1KG project associated with recurrently observation ($n > 5$) in hematological tumors in the COSMIC (v77) database. Variants were excluded with following criteria: (1) categorized into tier IV variants according to Guideline for Evidence-Based Categorization of Somatic Variants; (2) observed in paired peripheral blood samples by sanger sequencing. Sanger sequencing or digital PCR were conducted to confirm variants with critical VAF. Visual inspection was used to verify the final mutation annotation file (Table S3) and exclude potential false results.

### Genetic subtyping

The simplified 20-gene algorithm employed the Partitioning Around Medoids (PAM) clustering strategy to assign DLBCL samples into one of the pre-defined genetic subtypes: *TP53*$^{mut}$, MCD-like, BN2-like, N1-like, and EZB-like. This classification was based on mutation data of 18 genes and re-arrangement data of 2 genes assessed using FISH analysis (i.e., *BCL2* and *BCL6*). Candidate mutated genes were identified through WES and WGS of 325 patients, and their biological function were supported by RNA-seq data available from 184 patients. If any significant difference of relevant gene expression signature between the mutant and the wild-type patients was observed, the candidate mutated gene was considered functional (Table S6).

To generate the primary subtypes, clustering supervised by seed features (*MYD88*, *CD79B*, *BCL6* fusion, *NOTCH2*, *BCL2* fusion, *EZH2*, and *NOTCH1*) was first applied to the 325 patients, supported by 1000-times bootstrap. Subsequently, 18 genes with significantly different mutation patterns among these subtypes, along with re-arrangement data of *BCL2* and *BCL6*, were identified as clustering features. Patients with *TP53* mutations were assigned to a specific subtype with high priority. The set of features considered for possible association with each class was similar to those used in the GenClass algorithm: *TP53*$^{mut}$ (*TP53* mutations), MCD-like (*MYD88*, *CD79B*, *PIM1*, *MPEG1*, *BTG1*, *TBL1XR1* mutations), BN2-like (*BCL6* fusion, *NOTCH2*, *CD70*, *DTX1*, *TNFAIP3* mutations), N1-like (*NOTCH1* mutations), and EZB-like (*BCL2* fusion, *EZH2*, *TNFRSF14*, *CREBBP*, *EP300*, *MTOR*, *STAT6* mutations).

To validate the simplified 20-gene algorithm, we used 3 well-established datasets that were previously genetically categorized with LymphGen. These datasets included: 1) the BC Cancer (BCC) cohort, consisting of 320 newly diagnosed DLBCL patients with available clinical and genomic data from a targeted sequencing panel and FISH data,[53] 2) the UK population-based Haematological Malignancy Research Network (HMRN) cohort, consisting of 928 DLBCL patients from a real-world study with available clinical and panel-based DNA sequencing data,[11] and 3) the NCI cohort, consisting of 489 DLBCL patients from the dbGAP database.[9,54] In the NCI cohort, *BCL2* translocation and *BCL6* fusions results derived from RNA-seq were used as replacements due to the lack of FISH results during patient classification using the simplified 20-gene algorithm. All patients from the 3 validation cohorts, including the classification results of LymphGen algorithm and simplified 20-gene algorithm, were listed in Table S7.

### Gene expression quantification and preprocessing

The paired-end RNA-seq reads were pseudo-aligned to GRCh38 transcriptome (http://ftp.ensembl.org/pub/release-106/fasta/homo_sapiens/cdna/Homo_sapiens.GRCh38.cdna.all.fa.gz) using Kallisto (v0.46.0)[41] with default parameters and quantified as counts and transcripts per million (TPM). The original Kallisto outputs were integrated with R package "tximport" (v1.24.0)[42] and the counts/TPM of all alternative splicing transcripts of a gene were summed to represent the gene expression level. Protein coding genes, IG-C/V/D/J genes and TR-C/V/D/J genes were preserved, whereas histone- and mitochondrial-related genes were removed, resulting in ~20,000 genes for downstream analysis. Before LME classification and gene expression signatures calculation, log2 transformed TPM data (log2TPM) were processed with R package "preprocessCore" (v1.58.0), which implemented the quantile normalization described before[55] and removed batch effect caused by sequencing time, which was confirmed by PCA analysis, with R package "limma" (v3.48.3).[43]

### Lymphoma microenvironment, gene expression signatures and cell of origin

Patients were classified into 4 LME subtypes[13] with scores of 25 functional gene expression signatures and a KNN supervised model trained by a cohort of 4656 DLBCL patients (https://github.com/bostongene/lme). Gene expression signatures from SignatureDB (https://gdc.cancer.gov/about-data/publications/DLBCL-2018) were calculated with ssGSEA method[56] and compared between genetic or LME subtypes. Cell of origin was distinguished in patients with RNA-seq data using expression of signatures identified in Lymph2Cx assay.[45]

### Differential expression analysis and gene set enrichment analysis

Low expression genes with mean counts $\leq 5$ were removed. Then, quantile normalized and batch-adjusted logCPM (log2 transformed counts-per-million) were used for differential expression analysis (DEA) using R package "limma" (v3.48.3).[43] The gene list ranked by decreasing logFC (log2 fold-change) and ontology terms from gene ontology were used for gene set enrichment analysis (GSEA) that was implemented in R package "clusterprofiler" (v4.4.4).[57] Significantly enriched pathways were furtherly verified by GSEA software (Broad Institute, version 4.2.3).[58,59]

### Quantitative real-time PCR

Total RNA was extracted using FreeZol Reagent (Vazyme, Nanjing, China). Complementary DNA was synthesized using HiScript III RT SuperMix for qPCR with gDNA wiper (Vazyme, Nanjing, China). Quantitative real-time PCR (qRT-PCR) was performed by ChamQ Universal SYBR qPCR Master Mix (Vazyme, Nanjing, China) on LightCycler 480 System (Roche Diagnostic, Indianapolis, USA) with primers against ERVL-E, MER21C, HERVK11 and HERV16. Relative expressions were calculated using $_{\Delta\Delta}$CT method. Primer sequences were provided as follows: ERVL-E, forward: 5′-TCTCTATTGATGCCTTTATGAGT-3′, reverse: 5′-AGCATTGTGCTGTTTGTGAT-3'. MER21C, forward: 5′-GGAGCTTCCTGATTGGCAGA-3′, reverse: 5′-ATGTAGGGTGGCAAGCACTG-3'. HERVK11, forward: 5′-GGCTTGGTCCACAGATACAC-3′, reverse: 5′-TGCCTTGAATAGAGTGACCA-3'. HERV16, forward: 5′-CACCAGAAGGTCACCAGATA-3′, reverse: 5′- CTGTTGGGGAGTCCAGTTCT-3'.

### Randomization and masking

We used computer-assisted permuted-block randomization-with a block size of 4 and allocation ratio of 1:1 (R-CHOP-X, R-CHOP) for patients with different genetic subtypes and assigned patients in terms of the randomization results. The randomization sequence was generated by a statistician with no clinical involvement in the trial (R.J.M.). Investigators registered patients and assigned them based on the randomization sequence. Investigators and patients were not masked to treatment assignment.

## QUANTIFICATION AND STATISTICAL ANALYSIS

The objective of this study was to compare the efficacy and safety of R-CHOP-X with R-CHOP in newly diagnosed DLBCL patients. For sample size, we estimated that 65% of patients in the R-CHOP arm and 85% of patients in the R-CHOP-X arm would achieve CR. Sixty-four patients per group were required to show this difference with 5% significance (two-sided) and 76% power. Efficacy and safety analyses were by intention to treat. Genetic subgroup analysis was not prespecified in the sample size calculation. The difference in response or outcome observed within each genetic subgroup was exploratory. Statistical analyses were performed by Statistical Package for the Social Science (SPSS) 23.0 software (SPSS Inc., Chicago, IL, USA). Survival estimates were calculated by Kaplan-Meier method and survival curves were compared by Log rank test. Hazard ratios (HRs) and corresponding 95% confidence intervals (95%CIs) were estimated by Cox Proportional hazards model. A two-sided p value of $<0.05$ was considered statistically significant. Differences of normalized gene expression in 2 groups were analyzed using Mann-Withney test. Association of LME categories and genetic subtypes, and differences of treatment response between R-CHOP-X and R-CHOP arms within individual genetic or LME subgroup were analyzed using Fisher's exact test.