

High-resolution architecture and partner genes of *MYC* rearrangements in lymphoma with DLBCL morphology

Lauren C. Chong,¹ Susana Ben-Neriah,¹ Graham W. Slack,¹ Ciara Freeman,¹ Daisuke Ennishi,¹ Anja Mottok,² Brett Collinge,¹ Pau Abrisqueta,³ Pedro Farinha,¹ Merrill Boyle,¹ Barbara Meissner,¹ Robert Kridel,⁴ Alina S. Gerrie,¹ Diego Villa,¹ Kerry J. Savage,¹ Laurie H. Sehn,¹ Reiner Siebert,² Ryan D. Morin,^{5,6} Randy D. Gascoyne,¹ Marco A. Marra,⁶ Joseph M. Connors,¹ Andrew J. Mungall,⁶ Christian Steidl,^{1,*} and David W. Scott^{1,*}

¹Centre for Lymphoid Cancer, BC Cancer, Vancouver, BC, Canada; ²Institute of Human Genetics, University of Ulm and Ulm University Medical Center, Ulm, Germany; ³Department of Hematology, Vall d'Hebron Institute of Oncology (VHIO), University Hospital Vall d'Hebron, Barcelona, Spain; ⁴Division of Medical Oncology and Hematology, Princess Margaret Cancer Centre, Toronto, ON, Canada; ⁵Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, BC, Canada; and ⁶Canada's Michael Smith Genome Sciences Centre, Vancouver, BC, Canada

Key Points

- One third of *MYC* rearrangements in DLBCL occur in a cluster upstream of the *MYC* coding sequence and are enriched for *IGH* partners.
- Most *MYC* rearrangements in HGBL-DH patients have non-*IG* partners.

Genomic rearrangements in the *MYC* locus occur in ~12% of lymphomas with diffuse large B-cell lymphoma (DLBCL) morphology and are associated with inferior outcome. Previous studies exploring *MYC* rearrangements have primarily used fluorescence in situ hybridization (FISH) assays to characterize break-apart status but have rarely examined breakpoint location, and in some cases have not examined partner identity. We performed targeted sequencing of *MYC*, *BCL2*, *BCL6*, and the immunoglobulin (*IG*) loci in 112 tumors with DLBCL morphology harboring *MYC* rearrangement. We characterized the location of the *MYC* rearrangement at base pair resolution and identified the partner in 88 cases. We observed a cluster of breakpoints upstream of the *MYC* coding region and in intron 1 (the “genic cluster”). Genic cluster rearrangements were enriched for translocations involving *IGH* (80%), whereas nongenic rearrangements occurred mostly downstream of the *MYC* gene with a variety of partners, including *IGL* and *IGK*. Other recurrent partners included *BCL6*, *ZCCHC7*, and *RFTN1*, which has not previously been described as a *MYC* partner. We compared 2 commercially available FISH break-apart assays for the *MYC* locus and observed discordant results in 32% of cases examined, including some with *MYC-IGL* and *MYC-IGK* rearrangements. In cases of high-grade B-cell lymphoma with *MYC* and *BCL2* and/or *BCL6* rearrangement (HGBL-DH), so-called “double-hit” lymphomas, the majority of *MYC* rearrangements had non-*IG* partners (65%), with breakpoints outside the genic cluster (72%). In patients with de novo HGBL-DH of DLBCL morphology, *MYC-IG* rearrangements showed a trend toward inferior time to progression and overall survival compared with *MYC*-non-*IG* rearrangements. Our data reveal clinically relevant architecture of *MYC* rearrangements in lymphomas with DLBCL morphology.

Introduction

Genomic rearrangements in the *MYC* gene locus were first described through the characterization of the t(8;14)(q24;q32) translocation in Burkitt lymphoma 40 years ago.¹⁻³ *MYC* acts as a transcription factor regulating pathways including cell proliferation and growth, DNA replication, apoptosis, and differentiation, making it a powerful oncogene upon dysregulation.⁴ Similar *MYC* rearrangements, with a wider array of partner regions, have also been shown to occur in ~12% of lymphomas with diffuse large

B-cell lymphoma (DLBCL) morphology, and their presence is associated with inferior outcomes.⁵⁻⁹ In DLBCL-morphology tumors, rearrangements of the *MYC* locus are reported to occur concurrently with *BCL2* and/or *BCL6* rearrangements in 8% of cases, almost all of which are of the germinal center B-cell-like (GCB) subtype.^{7,9-14} These lymphomas have recently been classified as a distinct World Health Organization entity, "high-grade B-cell lymphoma with *MYC* and *BCL2* and/or *BCL6* rearrangements" (HGBL-DH),¹⁵ and have sufficiently poor prognosis for many centers to recommend dose-intensive chemotherapy regimens.^{16,17}

Although *MYC* rearrangements are known to have prognostic impact in DLBCL, previous large-scale studies examining their incidence and impact have mostly used break-apart fluorescence in situ hybridization (FISH) assays. These are able to detect the presence or absence of rearrangements but lack the resolution to identify the precise location of *MYC* breakpoints and require additional assays to identify rearrangement partners. Previous work in Burkitt lymphoma has classified *MYC* translocations according to recurrent breakpoint patterns, with breaks located in the *MYC* gene (class I), immediately upstream (class II), distally upstream (class III), or downstream.¹⁸ The patterns and relative incidences of these classes in lymphomas with DLBCL morphology have not been examined on a large scale. Currently, it remains unclear whether the location of the *MYC* break has any significance and whether breakpoint architecture differs between patient subsets (eg, HGBL-DH). Published studies have produced conflicting evidence about whether *MYC* rearrangement partner impacts patient outcome, with some demonstrating that *MYC*-immunoglobulin (*IG*) rearrangements are associated with inferior survival compared with *MYC*-non-*IG* rearrangements and other studies showing no association.^{12-14,19,20} However, the relationship among breakpoint location, *MYC* partner, and outcome has never been examined. Furthermore, some of these studies have used FISH break-apart assays with narrow gap sizes that have been demonstrated to lack the ability to detect a proportion of *MYC* locus rearrangements.^{21,22}

In this study, we used a targeted hybrid capture-based assay²³ to perform sequencing of the *MYC*, *BCL2*, *BCL6*, and *IG* (*IGH*, *IgL*, and *IGK*) loci in 112 tumors with DLBCL morphology that harbor *MYC* rearrangements. Using DNA extracted from routinely available formalin-fixed paraffin-embedded (FFPE) tissue, we characterized *MYC* rearrangements at base pair (bp) resolution and associated each rearrangement with its partner locus. By performing 2 FISH break-apart assays at the *MYC* locus, representing a wide-gap design and a narrower *MYC*-centric design, we were able to compare their sensitivity to detect breaks in the vicinity of *MYC*. Our data allowed us to identify recurrent patterns of rearrangement in the *MYC* locus and evaluate the impact of breakpoint architecture and partner identity on patient outcome.

Methods

Cohort selection and DNA extraction

To select cases for inclusion, 1048 biopsies from the BC Cancer lymphoma database were examined that were diagnosed with de novo DLBCL or transformed follicular lymphoma (tFL) with DLBCL morphology and had been assessed for *MYC* rearrangement status using the Vysis FISH break-apart assay (wide-gap design). Of these,

559 were also assessed with the Dako *MYC* break-apart assay (narrower *MYC*-centric design). Biopsies were selected if they were *MYC* break-apart positive by either assay and had available FFPE tumor tissue for analysis. This yielded a total of 112 cases: 95 with de novo DLBCL and 17 with tFL (Figure 1; Table 1; supplemental Table 1). Although the majority of biopsies emanated from population registry-based cohorts, 7 of the de novo DLBCL biopsies came from tissue microarrays consisting of selected patient cohorts, including 5 from patients who experienced central nervous system relapse. Break-apart assays for *BCL2* and *BCL6* (Vysis) were also performed for all biopsies. DNA was extracted from the FFPE tumors and used to create libraries for next-generation sequencing (supplemental Methods). Cell-of-origin subtype was assigned using the Lymph2Cx assay²⁴ or the Hans algorithm.²⁵ All patient specimens were collected as part of research projects approved by the University of British Columbia–British Columbia Cancer Agency Research Ethics Board (file H14-02304). This study abided by the Declaration of Helsinki.

Targeted sequencing and structural variant detection

Prior to sequencing, a custom Agilent SureSelect hybrid capture assay was used to enrich for DNA in the regions of interest (supplemental Table 2). Regions for capture were chosen based on published literature and aimed to encompass all reported chromosomal breakpoints at the loci of interest.²⁶⁻³⁴ The total size of the target space was 7.8 Mb.

Libraries were pooled, and capture enrichment was performed, followed by sequencing on an Illumina HiSeq 2500 (v4 chemistry, 125 bp paired-end reads). Eleven cases had a second round of library construction and sequencing completed to further investigate breakpoints not yet identified in the *MYC* region, yielding 123 total libraries. To achieve ~200× average depth of coverage in the captured regions, 57 of the libraries were sequenced on 3 lanes, and the remaining 66 were sequenced on 1 lane (supplemental Figure 1).

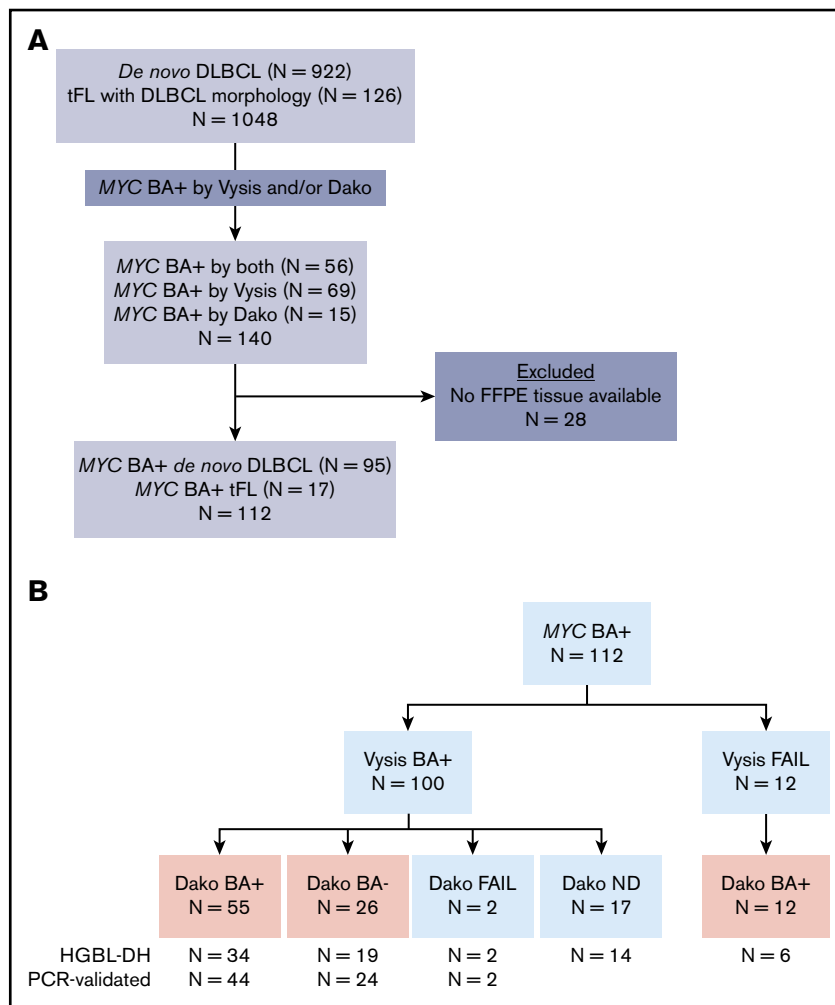
Reads were aligned to the GRCh37 reference genome using BWA-MEM (v0.7.5a),³⁵ and predicted structural variants (SVs) were identified using the deStruct (v0.2.1; <https://bitbucket.org/dranew/destruct>) and DELLY (v0.5.5)³⁶ tools in an ensemble approach (see supplemental Methods for details). Additional putative rearrangements were identified by manual inspection in IGV³⁷ of the *MYC* gene and surrounding region.

PCR-based SV validation

A subset of 70 cases with available remaining high-quality DNA was selected for SV validation. Validation was attempted for all predicted rearrangements in the *MYC*, *BCL2*, and *BCL6* regions in these cases. Each rearrangement had 3 sets of polymerase chain reaction (PCR) primers designed to amplify the fusion breakpoint and each of the 2 wild-type corresponding regions. The PCR-amplified fragments were generated using 35 cycles of PCR, followed by a second PCR reaction of 6 cycles to introduce Illumina sequencing adapters. Amplicons were pooled and sequenced on an Illumina MiSeq (250 bp paired-end reads). Sequenced reads were aligned to the putative fusion sequences, and reads mapping across a breakpoint were aligned to the hg19 reference with BLAT to verify exact breakpoint locations and identify any inserted sequence. Fusion sequences with no reads aligned to them were considered failed validations.

Figure 1. Cohort selected for capture sequencing. (A)

Biopsies from de novo DLBCL or tFL with DLBCL morphology were selected for inclusion if they were break-apart positive (BA+) at the *MYC* locus using the Vysis and/or Dako FISH assays. Cases with no available FFPE tissue were excluded, leaving a total of 112 cases. (B) The FISH break-apart (BA) status of the 112 cases are summarized. Cases in red boxes represent those successfully assessed with the Dako assay. ND, not done.



Whole-genome sequencing of DLBCL-derived cell lines

Genomic libraries were prepared for 19 DLBCL cell lines (supplemental Table 3) and sequenced on an Illumina HiSeq 2000 (paired-end 100 bp reads) or HiSeq X (paired-end 150-bp reads) platform (some previously described³⁸; see supplemental Methods). Reads were aligned to the GRCh37 reference using BWA-MEM (v0.7.5a),³⁵ and SVs in the *MYC*, *BCL2*, and *BCL6* regions were identified using DELLY (v0.5.5).³⁶

Outcome analysis

We used descriptive statistics to summarize the baseline characteristics of all patients (Table 1). Survival analysis was performed on (1) all patients with de novo DLBCL treated with R-CHOP chemotherapy (rituximab with cyclophosphamide, doxorubicin, vincristine, and prednisone) and (2) as above, but limited to patients representative of a population-based registry, excluding the patients from selected cohorts. Time to progression (TTP) was defined as the time from date of diagnosis to disease progression, relapse, or death related to lymphoma or its treatment. The Kaplan-Meier method was used to estimate overall survival (OS) and TTP, and log-rank tests were performed to compare risk between groups. Three-year estimates with 95% confidence intervals were calculated.

Statistical analysis was performed using STATA/IC statistical software (v13; StataCorp).

Results

FISH break-apart assays identify a cohort of 112 cases

We used 2 *MYC* break-apart FISH designs: the Vysis assay, which has probes that bind across a large gap (~1.6 Mb) that includes the *MYC* gene, and the Dako assay, which has a smaller gap (~90 kb) centered around *MYC*. In total, we examined 112 biopsies that harbored rearrangement of the *MYC* locus using Vysis and/or Dako assays (Figure 1). The Vysis assay was performed on all 112 biopsies and was evaluable in 100 cases. The Dako assay was performed on 95 of the biopsies and was evaluable in 93 cases. Of the 81 cases with interpretable results for both *MYC* break-apart assays, 55 were positive by both, whereas 26 (32%) were positive by Vysis and negative by Dako. *BCL2* and *BCL6* break-apart FISH assays were attempted on all biopsies and were successful in 106 and 109 cases, respectively (supplemental Figure 2). Of 107 biopsies in which double-hit status was evaluable by FISH (ie, break-apart positive for *BCL2* and/or *BCL6* or negative for both), 75 (70%) were classified as HGBL-DH, with 49 of these harboring a *BCL2* rearrangement (65%), 11 harboring a *BCL6*

Table 1. Baseline patient and first-line treatment characteristics of the entire cohort

Characteristic	de novo DLBCL (N = 95)			tFL (N = 17)		
	n	%	Missing	n	%	Missing
Male	66	69	0	10	59	0
Age >60 y	62	65	0	9	53	0
Ann Arbor stage III/IV	56	59	0	12	86	3
ECOG performance status >1	41	43	0	7	50	3
Elevated LDH	56	64	7	9	60	2
Extranodal sites (any)	60	63	0	10	77	4
Extranodal site >1	21	22	0	9	69	4
IPI 0-2	43	49	7	6	46	4
IPI 3	17	19	7	3	23	4
IPI 4-5	28	32	7	4	31	4
CHOP-containing regimen	86	91	0	14	82	0
Rituximab given	90	95	0	11	65	0
Palliation	2	2	0	2	12	0
Intensive regimen*	7	7	0	1	6	0

Median age for all patients is 64 years (range 28-87). For the de novo DLBCL patients, median duration of follow-up for living patients is 7.5 years (range 1.5-16).

CHOP, cyclophosphamide, doxorubicin, vincristine, and prednisone; ECOG, Eastern Cooperative Oncology Group; IPI, International Prognostic Index; LDH, lactate dehydrogenase.

*Regimens included R-CODOX-M/IVAC (rituximab with cyclophosphamide, vincristine, doxorubicin, and high-dose methotrexate/ifosfamide, etoposide, and high-dose cytarabine), GDP (gemcitabine, dexamethasone, and cisplatin), and DA-EPOCH-R (dose-adjusted etoposide, prednisone, vincristine, cyclophosphamide, and doxorubicin with rituximab).

rearrangement (15%), and 15 (20%) harboring both *BCL2* and *BCL6* rearrangements in addition to the *MYC* rearrangement ("triple-hit" cases).

Capture sequencing results are highly concordant with FISH assays

The capture sequencing and SV detection pipeline was able to identify rearrangements closely matching the observed FISH break-apart results (supplemental Table 4). Of the 112 cases, 93 (83%) had a predicted *MYC* rearrangement (translocation or large intrachromosomal SV) that matched the observed FISH pattern. Predicted *BCL2* and *BCL6* rearrangements were found with capture in 95% (61/64) and 69% (18/26) of FISH break-apart-positive cases, respectively.

Prior to validation, predictions were classified as high, moderate, or low confidence based on the number of tools that identified them and their read support (supplemental Methods). Of the 70 cases for which SV validation was attempted, 61 had predicted *MYC* rearrangements. Fifty-four of these (89%) validated, and the exact breakpoint locations and the sequence of any inserted DNA were identified (supplemental Table 4). Of the 7 cases for which the *MYC* rearrangements were not validated, 4 did not produce conclusive results: 1 had very few sequenced reads in the validation assay (~3000 total vs median ~440 000) and likely failed to amplify the fusion as a result of the low amount of input DNA, and 3 showed no amplification of 1 or both "wild-type" alleles, indicating PCR priming failure or inadequate DNA quality. The remaining 3 cases did not validate for unknown reasons, and we consider these to be artifactual predictions (2 low confidence, 1 moderate confidence).

Thirty-seven validation cases had a predicted *BCL2* rearrangement, and 100% of these validated. Of 13 cases with *BCL6* rearrangements, 92% (12/13) were validated; the 1 failed case likely represented a PCR failure.

Of the cases with conclusive validation results, 49 had a moderate- or high-confidence *MYC* prediction, and 98% of these (48/49) successfully validated. In the cases in which *MYC* validation was inconclusive or not attempted, 94% (34/36) also had moderate- or high-confidence predictions, suggesting that the vast majority of *MYC* rearrangements in the nonvalidated cases are likely true positives. The 2 low-confidence predictions in these cases were omitted in downstream analyses.

Recurrent rearrangement breakpoint patterns in the *MYC* locus

In total, we identified 105 validated or high/moderate-confidence *MYC* rearrangements in 88 cases, including 93 translocations and 12 large (>2 Mb) intrachromosomal rearrangements (Figure 2). The large number of breakpoints identified with the capture sequencing approach allowed us to observe recurrent patterns in rearrangement anatomy. Many of the *MYC* breakpoints (35/105; 33%) occur in a cluster in close proximity to the *MYC* coding sequence, which we define here as the "genetic cluster," spanning from ~1.5 kb upstream of the transcription start site to the end of *MYC* intron 1 (~3.7 kb total). This cluster encompasses the class I and class II *MYC* translocations previously defined in Burkitt lymphoma.¹⁸ The remaining breakpoints are located outside the gene region, with 27% (19/70) occurring upstream of *MYC* in a space that includes previously defined class III *MYC* translocations.¹⁸ However, the majority of nongenic rearrangements occur telomeric to the genetic cluster (51/70; 73%) up to 0.6 Mb downstream of *MYC*.

Breakpoints in the genetic cluster are significantly enriched for *IGH* partners (28/35; 80%), whereas those outside of the genetic cluster are enriched for non-*IGH* rearrangements (63/70; 90%) (Table 2; $\chi^2 P < .0001$). Nongenic rearrangements are highly variable with respect to partners, with recurrent partners including the other *IG* genes (*IGL* and *IGK*) and previously described partner loci *BCL6* and *ZCCHC7/PAX5* (Figure 3).³² In addition, a novel recurrent partner, *RFTN1*, was identified in 4 cases. This gene encodes the RFTN1 protein, a lipid raft linker, and all breakpoints observed in the *RFTN1* gene occur in intron 1 or intron 2 (supplemental Figure 3). Rearrangements in the known single-hit cases were 50% *MYC-IG* (17/34), and 41% (14/34) occurred in the genetic cluster. In contrast, the majority of *MYC* rearrangements in HGBL-DH cases had non-*IG* partners (45/69; 65%), with the majority of breakpoints outside the genetic cluster (50/69; 72%). A similar trend was seen in patients with GCB-DLBCL; 65% (46/71) of *MYC* rearrangements had non-*IG* partners, and 70% (50/71) occurred outside the genetic cluster. The location of rearrangements with *IGH* and *IGL/IGK* partners showed a similar pattern to what has been reported in Burkitt lymphoma, suggesting a common mechanism and developmental stage at which these rearrangements arise (supplemental Figure 4).^{28,39,40}

FISH break-apart status at the *MYC* locus is variable between FISH assays

Break-apart *MYC* FISH assays were interpretable using Vysis and Dako probes in 81 of the 112 cases, allowing us to compare the patterns observed with the 2 assays. In total, 32% (26/81) of

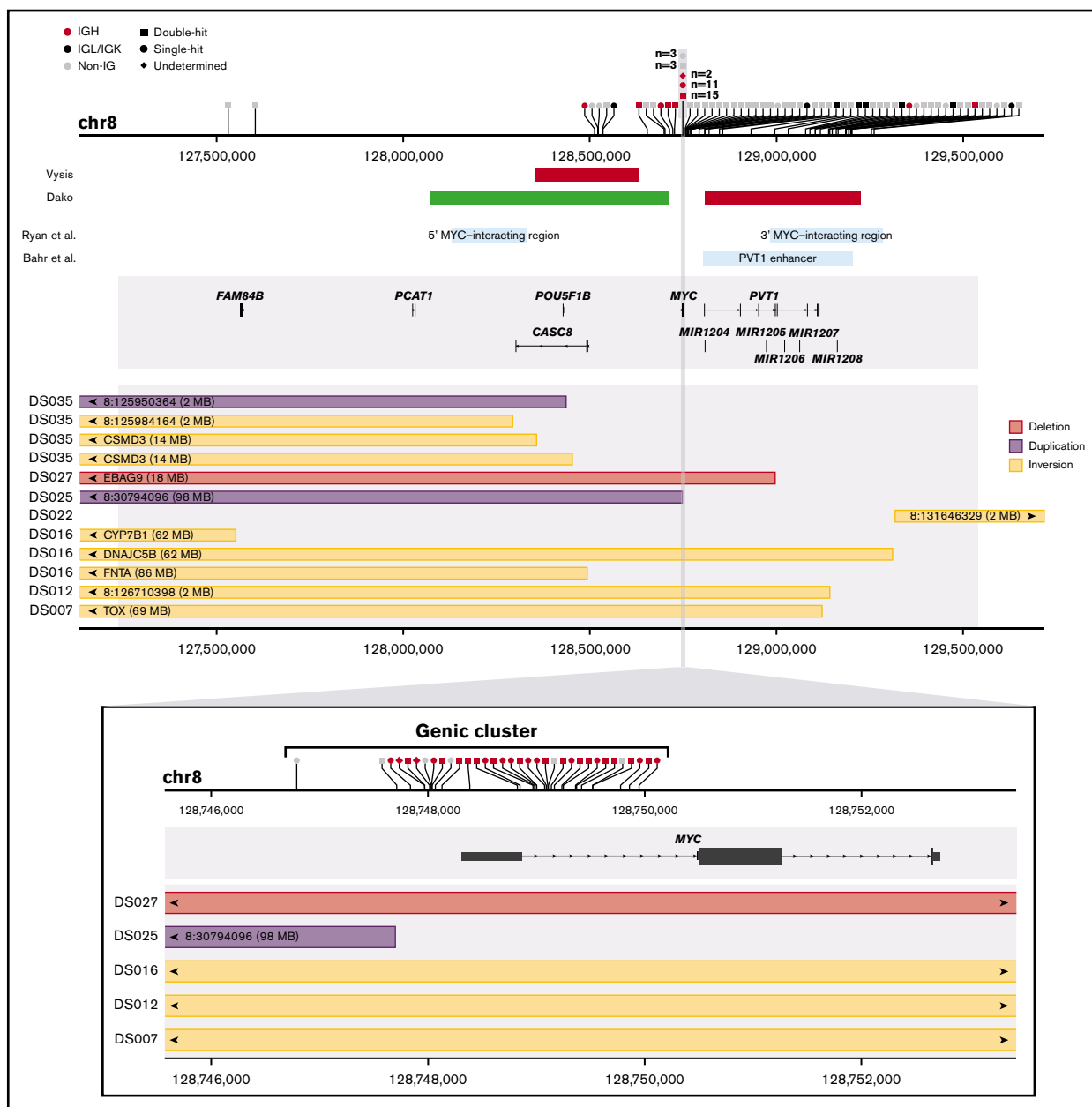


Figure 2. SVs in the *MYC* locus discovered by targeted capture sequencing. The horizontal axis represents genomic space on chromosome 8 (chr8). The upper panel summarizes the whole *MYC* capture region, and the bottom panel zooms in to the gene region. In each panel, the top track shows the location of translocation breakpoints (vertical ticks on the axis), and partner information is summarized with symbols above. The middle track shows gene models in the region. The bottom track shows the location of intrachromosomal rearrangements, with rectangles representing the span of the rearrangements. Arrows indicate that the end of the rearrangement is located outside of the plotting region at the labeled location. In the upper panel, rearrangements in the genic cluster have been binned for visualization. Additional middle tracks in the upper panel show the approximate location of reported enhancer regions (blue boxes) and the approximate binding location of FISH break-apart probes (green Vysis probe binds downstream of the plotted region). The shaded gray rectangle shows the captured region. Rearrangements that failed validation and low-confidence predictions in nonvalidated cases have been omitted. Intrachromosomal rearrangements smaller than 2 Mb have not been plotted, and rearrangements for which both reciprocal events were identified have only 1 displayed.

cases were negative when hybridized with the Dako probe (supplemental Table 5) and positive by Vysis. Because of the probe design, most of the cases found to have a genic breakpoint were Dako break-apart positive (26/27; 96%), whereas 43% (20/46) of cases with nongenic breakpoints did not produce a break-apart signal using the Dako assay. In some cases, we were able to further determine whether an *MYC-IGH* rearrangement was

present using a dual fusion FISH assay when capture did not identify a rearrangement (supplemental Methods; supplemental Table 1). With these data included, the Dako break-apart-negative cases had primarily *MYC*-non-*IGH* rearrangements (20/24; 83%), and 3 of these had an *IGL* or *IGK* partner. These findings are consistent with previous reports suggesting that the narrower *MYC*-centric Dako assay fails to detect a significant proportion of

Table 2. Summary of rearrangement partners identified in the *MYC* region using capture sequencing

	Total breakpoints	<i>IGH</i> partner	<i>IGK/IGL</i> partner	Non- <i>IG</i> partner
Genic cluster	35	28 (80)	0 (0)	7 (20)
Other	70	7 (10)	8 (11)	55 (79)
Single-hit	34	14 (41)	3 (9)	17 (50)
HGBL-DH	69	19 (28)	5 (7)	45 (65)
GCB	71	19 (27)	6 (8)	46 (65)
ABC/non-GCB	25	11 (44)	2 (8)	12 (48)

Total counts of rearrangements are shown with percentages in parentheses. Counts include translocations and large intrachromosomal rearrangements (>2 Mb). Rearrangements that failed validation and low-confidence predictions in nonvalidated cases have been omitted.

ABC, activated B-cell-like DLBCL.

rearrangements in the *MYC* locus.^{21,22} The large sample size and exact breakpoint locations elucidated in this study allowed us to more accurately describe the location where *MYC*-locus rearrangements were too far downstream from the coding space to be detectable with the Dako assay (Figure 4).

Recurrent rearrangement patterns in the *IGH*, *BCL2*, and *BCL6* loci

In addition to the *MYC* locus, other loci in the capture space showed recurrent patterns with respect to breakpoint anatomy. Most *BCL2* breakpoints with *IGH* partners were located in the 3' untranslated region and downstream of the *BCL2* gene, a region containing the previously designated major breakpoint region, minor cluster region, and intermediate cluster region

(supplemental Figure 5).³⁴ We identified additional breakpoints at the 5' end of *BCL2* (4 upstream, 1 in the 5' untranslated region, and 1 in intron 2) in the known variant cluster region with partners including *IGH*, *MYC*, and *IGK*.^{41,42} In the *BCL6* region, 52% of the breakpoints (12/23) occurred in the major breakpoint region comprising exon 1 and intron 1 (supplemental Figure 6).²⁷ However, the remaining breakpoints were not localized to the alternative breakpoint region (245-285 kb 5' of the gene) and were instead distributed throughout the region spanning from the 5' end of *BCL6* to ~225 kb upstream (telomeric), including some in reported enhancer elements.^{43,44} This suggests that additional breakpoints outside the major and alternative breakpoint regions are frequent in B-cell lymphomas with DLBCL morphology. Because many of these breakpoints occur in a space not captured with our design, it is possible that additional breakpoints exist that have not been identified. Recurrent patterns were also observed in the *IGH* locus (Figure 5). Rearrangements with *BCL2* occurred primarily in the D and J segments, whereas those with *MYC* occurred most often in switch regions upstream of constant segments.^{45,46} In *BCL2-IGH* rearrangements for which both reciprocal translocations were identified, the 2 *IGH* breakpoints always occurred in different *IGH* locations, with 1 in the D region and the other in the J region. Breakpoints with *BCL6* partners, similar to *MYC*, occur in *IGH* constant switch regions.

DLBCL cell lines show patterns similar to patient samples

To further investigate rearrangement patterns, we also examined the *MYC*, *BCL2*, and *BCL6* loci for rearrangements in 19 DLBCL-derived cell lines (supplemental Tables 6 and 7). Using genome sequencing, we identified a *MYC* rearrangement in 12 of the cell lines (supplemental Figure 7). The patterns of rearrangement in the

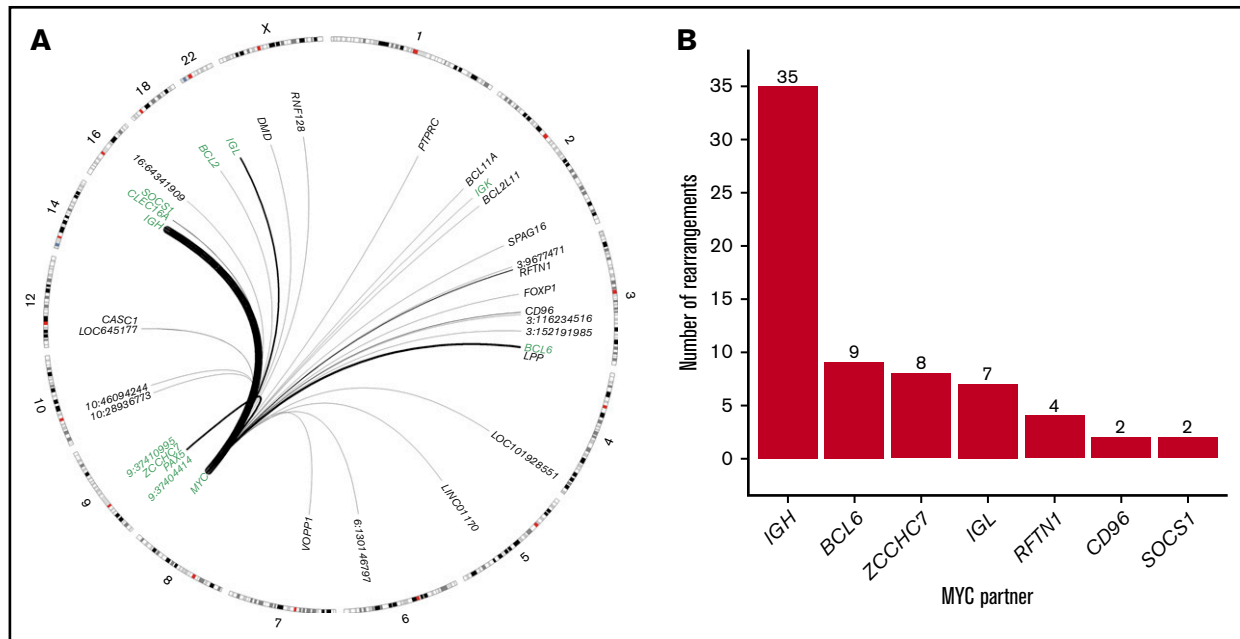


Figure 3. Identity and location of *MYC* rearrangement partners throughout the genome. (A) Circos plot showing identified rearrangement partners of *MYC*. Black arches represent translocations with thickness representing recurrence. Chromosomes with no identified partners are not drawn. Regions included in the capture space are labeled in green. (B) Absolute frequency of recurrent *MYC* rearrangement partners.

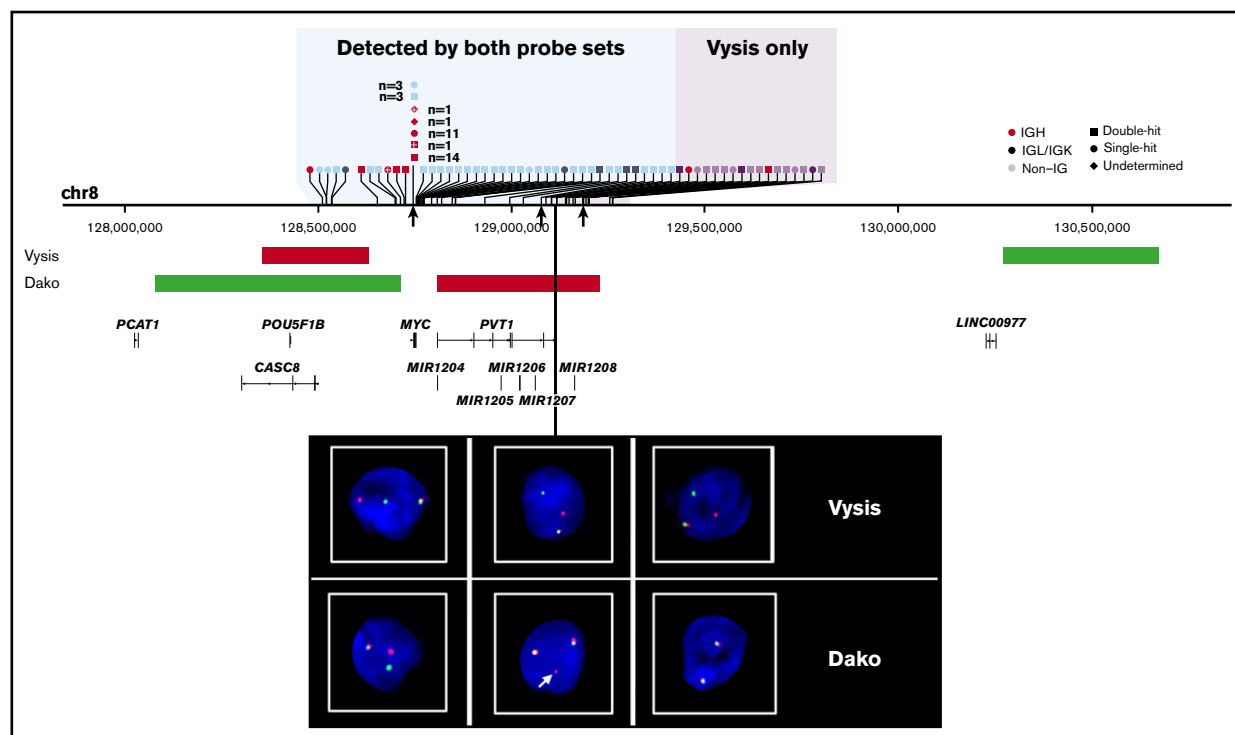


Figure 4. Sensitivity of rearrangement detection at the *MYC* locus with Vysis and Dako FISH probes. The horizontal axis represents genomic space on chromosome 8 (chr8). Translocations are represented by vertical ticks along the axis, with symbols describing the rearrangement partner labeled above. White crosses denote cases that were *MYC* break-apart positive by Dako only (ie, Vysis failed). The shaded blue region represents the space where rearrangements can be detected by both assays, and the shaded purple region shows where only Vysis can detect rearrangements. The approximate location of FISH binding probes is shown on the middle track. Representative FISH images show observed patterns: cases in the blue region produce a fusion signal and a break-apart signal in both assays (left); cases with breakpoints that lie within the Dako red probe binding region produce 2 fusion signals and a small additional red signal, as indicated by the white arrow (center); and cases in the purple region produce a break-apart signal with Vysis and 2 fusion signals with Dako (right). Breakpoints corresponding to the FISH images are indicated by black arrows under the chromosome track.

cell lines were consistent with those observed in primary patients, with a cluster of *MYC-IGH* rearrangements in the genic region and non-*IGH* partners distributed throughout the locus, predominantly downstream of the gene. The observed partner genes were also consistent and included *BCL6*, *ZCCHC7*, and *IGL*. Interestingly, NU-DHL-1 was confirmed to harbor an *MYC-RFTN1* rearrangement (previously published as *PVT1-RFTN1*³⁸). DLBCL cell line rearrangements in the *BCL2*, *BCL6*, and *IGH* loci showed patterns similar to primary patients with respect to breakpoint location and partner identity (supplemental Figures 8-10).

Identity of the *MYC* rearrangement partner impacts prognosis

To investigate whether the *MYC* rearrangement partner impacts patient prognosis, we used the Kaplan-Meier method to estimate TTP and OS in patients with de novo DLBCL treated with first-line R-CHOP. Genic breakpoint status did not confer inferior survival in comparison with nongenic breakpoints. Consistent with previous reports,^{12,13} we observed a nonsignificant trend toward inferior outcome in all patients harboring *MYC-IG* rearrangements compared with those with *MYC-non-IG* rearrangements (supplemental Figure 11). In the subgroup of HGBL-DH patients with *BCL2* rearrangement, a similar trend was observed, with *MYC-IG* cases showing inferior TTP ($P = .12$) and OS ($P = .09$). The same observations held true when considering only patients representative

of a population-based registry (ie, excluding patients from selected cohorts) (supplemental Figure 11).

Discussion

In this study, we have performed the first large-scale characterization of *MYC* rearrangement architecture in lymphomas with DLBCL morphology. The nucleotide-level resolution of our sequencing assay allowed for the identification of recurrent rearrangement patterns that were not previously appreciated. Furthermore, we compared the sensitivity of 2 routinely available FISH assays and verified associations between *MYC* rearrangement partner and patient outcome.

We identified a genic cluster of *MYC* breakpoints located near the 5' end of the coding sequence, which is highly enriched for *MYC-IGH* rearrangements, and demonstrated that most nongenic *MYC* rearrangements have non-*IG* partners and occur downstream of the gene. All downstream breakpoints identified are located within a region ~565 kb telomeric of the coding space, which has been demonstrated to be enriched for enhancer elements.^{43,47} In conjunction with a previous study demonstrating that *MYC* rearrangements frequently occur within acetylated enhancer regions in B-cell lymphomas, the location of our mapped breakpoints may suggest that these rearrangements occur within enhancers in this region.⁴³ Future experiments investigating acetylation status in both the *MYC* region and the numerous partner regions

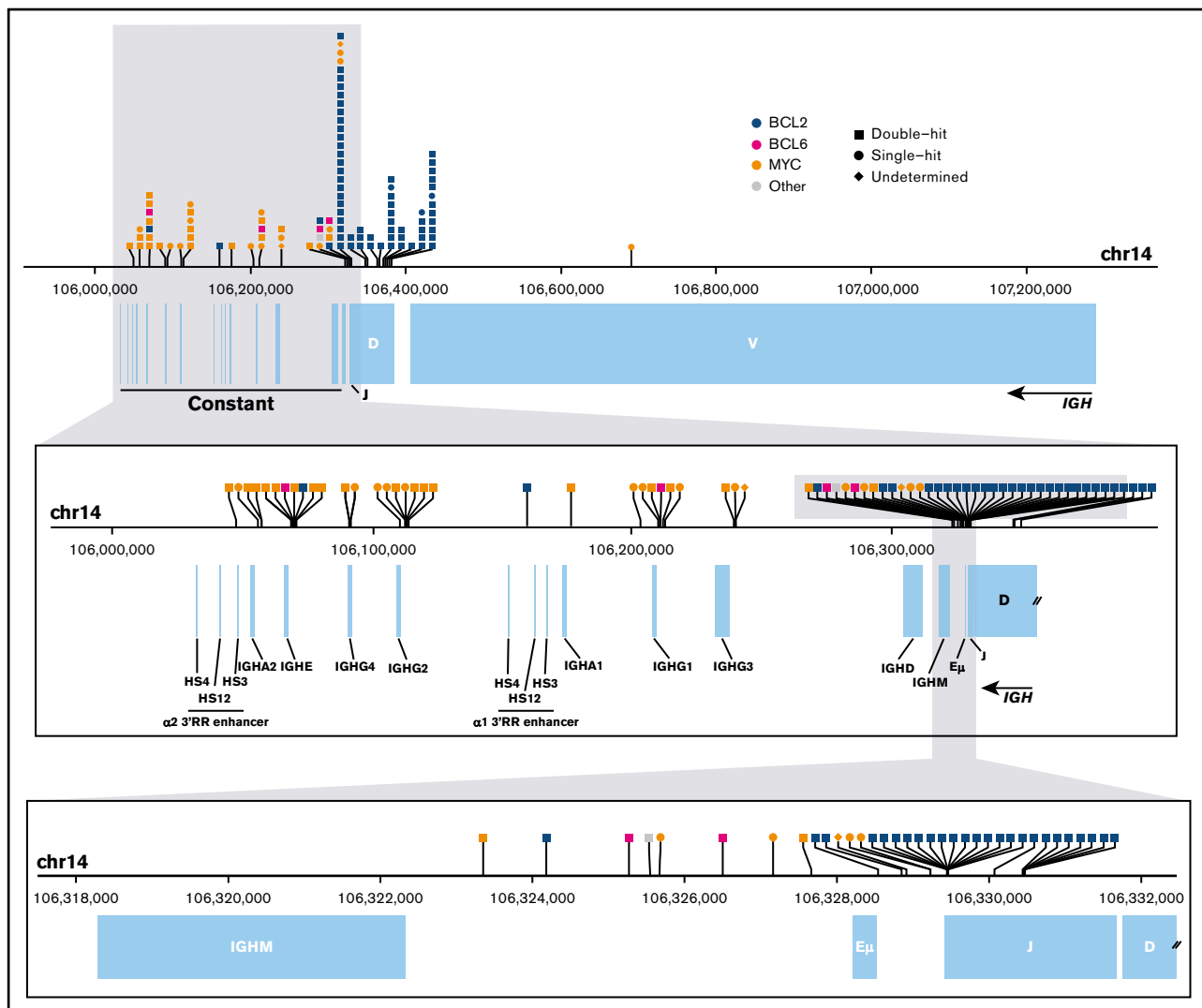


Figure 5. Translocation breakpoints in the *IGH* locus. The horizontal axis represents genomic space on chromosome 14 (chr14). The top panel shows the full *IGH* locus, the middle panel shows an enlargement of the constant region, and the bottom panel shows the smaller region surrounding the end of the variable diversity joining genes. In each panel, the top track indicates the location of translocation breakpoints (vertical tick marks), with symbols describing the identity of the partner region labeled above. Breakpoints shown on the uppermost panel have been binned for visualization. Annotations of the variable diversity joining segments, constant regions, and enhancer elements (E_{μ} and 3'RR) are depicted below.

discovered here may provide insight into this potential mechanism. Interestingly, another known superenhancer region 1.7 Mb downstream of the *MYC* gene is not captured in our targeted sequencing assay.⁴⁷ It is possible that some rearrangements have breakpoints in this region that are missed by our current design, and this may account for discrepancies between FISH and the capture results that can be improved in future experiments. We also identified numerous novel *MYC* rearrangement partners, including *RFTN1*, recurrently seen in 4 patient samples and the GCB-DLBCL cell line NU-DHL-1. This gene is known to be involved in B-cell receptor signal transduction.⁴⁸ All breakpoints observed occurred in intron 1 or 2 (supplemental Figure 3), and further investigation is required to determine whether these rearrangements impact *RFTN1* messenger RNA or protein expression, as well as how this may functionally impact tumor cells. Analysis of 19 DLBCL-derived cell line genomes found 12 cell lines harboring *MYC* rearrangements, including

9 HGBL-DHs (supplemental Table 7). Because the cell lines showed patterns similar to primary cases with respect to breakpoint location and partner identity, they appear to faithfully represent the spectrum of observed *MYC* rearrangements and provide opportunities to functionally study representative patterns in more detail. Interestingly, 9 of the 14 GCB cell lines profiled were HGBL-DHs, and only 1 had a sole *MYC* translocation, indicating a skew not representative of the general GCB patient population.

Breakpoints in the *IGH* locus showed highly recurrent patterns with respect to partner regions (Figure 5). The prevalence of *BCL2* rearrangements in the variable diversity joining (VDJ) segments of the *IGH* locus is consistent with their origin during RAG-mediated VDJ recombination in the bone marrow.⁴⁶ *MYC* and *BCL6* rearrangements occurred primarily in *IGH* switch regions upstream of constant gene segments, indicating their likely origin during AID-mediated class switch recombination in the

germinal center reaction. The mapping of breakpoints allowed us to observe the structure of the derivative chromosomes and location of regulatory enhancer elements resulting from characteristic *IGH* rearrangements with *MYC*, *BCL2*, and *BCL6* (supplemental Figure 12). In all 3 characteristic rearrangements, the breakpoint architecture places the partner in close proximity to a highly active *IGH* enhancer element (E_{μ} or 3'RR), suggesting a classical mechanism of oncogene overexpression.

Consistent with previous reports, we observed underdetection of *MYC* break-apart-positive cases when using the Dako FISH assay (supplemental Table 5).^{21,22} *MYC*-centric FISH assays like Dako are effective at identifying genic breakpoints and, therefore, are biased toward finding *MYC-IG* rearrangements and missing *MYC-non-IG* rearrangements. The majority of cases with *MYC-non-IG* rearrangements are HGBL-DH (83%; 38/46), suggesting that the wider-gap Vysis probe has higher sensitivity for identifying these cases. This is particularly important because these cases now constitute a distinct World Health Organization entity, and routine *MYC* break-apart testing will require an assay that can reliably classify these patients. In our comparison of HGBL-DH patients with R-CHOP-treated de novo DLBCL, we confirmed a trend toward inferior survival in cases harboring *MYC-IG* rearrangements compared with those with *MYC-non-IG* rearrangements.^{12,13} However, the small size of this subset resulted in the trend being nonsignificant. Because our study used the Vysis probe design, we identified additional cases with both *MYC-IG* and *MYC-non-IG* rearrangements that would have likely been omitted in previous studies that used the Dako design.^{12,13} A potential limitation to our outcome analysis is that our cohort contains some cases identified in an era in which FISH assays were performed on the basis of defined patient or tumor characteristics and others for which FISH was applied to all tumors. Our findings will need to be further confirmed in large population-based studies in which FISH is applied to all biopsies.

Although our capture sequencing assay provides resolution not possible with FISH, it still has limitations. In 24 cases (21%), no *MYC* rearrangement was identified with sequencing that explained the observed FISH pattern. These discrepancies may be the result of technical issues (eg, degradation of FFPE DNA, breakpoints in low-coverage repetitive regions) or biological variability (eg, tumor heterogeneity). It is also likely that our design does not encompass all breakpoints that occur in the loci of interest. This is most evident at the *BCL6* locus: the design used here, which covers the major and alternative breakpoint regions, did not include 35% (8/23) of the identified *BCL6* breakpoints (supplemental Figure 6). In future experiments, an improved capture design may allow us to identify a higher proportion of rearrangements. Targeted sequencing may also be unable to characterize complex genetic rearrangements as a result of the inability of short reads to sequence across large regions. Some cases investigated had >50 predicted high-confidence rearrangements, including overlapping inversions, duplications, and deletions, and these numerous predictions may be the result of mapping difficulty in highly altered regions. Other cases had multiple rearrangements in the same locus (eg, *MYC-IGH* and *BCL2-IGH*), but our current method is unable to determine whether these occur on the same or alternate alleles, making it difficult to discern the sequence and potential mechanisms that give rise to these translocations. For these applications, long- or linked-read sequencing may improve upon the rearrangement detection pipeline.

This study provides a large-scale description of *MYC* rearrangements in B-cell lymphomas with DLBCL morphology at a resolution not previously attainable with conventional FISH methods. Our findings suggest that, given the heterogeneity of the rearrangements we characterized, future research and clinical trials may benefit from the use of both a wide-gap FISH assay and capture sequencing to stratify patients into subsets based on breakpoint location and partner. We anticipate that unresolved questions about the relationship between *MYC* rearrangement architecture and outcome can only be answered when larger numbers of patients are characterized. Further, examination in cohorts that have been treated with different therapeutic regimens is required to determine the clinical utility of assessing *MYC* rearrangement status and architecture. This work represents an important step toward understanding *MYC* rearrangements and their impact on patient outcomes, as well as designing effective strategies for detection in patient populations.

Acknowledgments

The authors thank Doug Horsman for his contributions to cytogenetics at BC Cancer. They also thank Felix Breden (Simon Fraser University) and Corey Watson (University of Louisville) for sharing their knowledge regarding the *IGH* locus. The authors thank and acknowledge staff at Canada's Michael Smith Genome Sciences Centre for the generation of discovery and validation sequencing datasets and acknowledge the team at the molecular cytogenetics laboratory, Institute of Human Genetics, Ulm University.

This work was supported by the Canadian Cancer Society Research Institute (704848, 705288), Genome Canada (4108), Genome British Columbia (171LYM), the Canadian Institutes of Health Research (GPH-129347), the Terry Fox Research Institute (1061), and the BC Cancer Foundation. A.M. was supported by fellowship awards from the Mildred Scheel Cancer Foundation (German Cancer Aid), the Michael Smith Foundation for Health Research, and Lymphoma Canada. R.S. is supported by the e:Med programme (Bundesministerium für Bildung und Forschung, FKZ 031A428H). C.S. is supported by a Michael Smith Foundation for Health Research career investigator award (5120).

Authorship

Contribution: L.C.C., S.B.-N., R.D.G., J.M.C., A.J.M., C.S., and D.W.S. designed the project; C.F., A.S.G., D.V., K.J.S., L.H.S., R.D.G., J.M.C., and D.W.S. collected patient data; G.W.S. and P.F. performed pathology; S.B.-N., D.E., A.M., B.C., P.A., R.K., and R.S. performed FISH analyses; S.B.-N., G.W.S., C.S., and D.W.S. performed cohort selection; S.B.-N., A.M., M.B., and B.M. performed nucleic acid extraction; M.A.M. and A.J.M. designed and performed the capture sequencing, whole-genome sequencing, and validation protocols; L.C.C. and R.D.M. performed sequencing and bioinformatics analysis; C.F. performed outcome analysis; L.C.C. wrote the manuscript; and all authors edited and approved the manuscript.

Conflict-of-interest disclosure: C.F. receives research funding from Roche/Genentech and honoraria from AbbVie, Seattle Genetics, and Roche/Genentech. The remaining authors declare no competing financial interests.

Correspondence: David W. Scott, Centre for Lymphoid Cancer, BC Cancer, 675 West 10th Ave, Vancouver, BC V5Z 1L3, Canada; e-mail: dscott8@bccancer.bc.ca; and Christian Steidl, Centre for Lymphoid Cancer, BC Cancer, 675 West 10th Ave, Vancouver, BC V5Z 1L3, Canada; e-mail: csteidl@bccancer.bc.ca.

References

1. Dalla-Favera R, Bregni M, Erikson J, Patterson D, Gallo RC, Croce CM. Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. *Proc Natl Acad Sci USA*. 1982;79(24):7824-7827.
2. Zech L, Haglund U, Nilsson K, Klein G. Characteristic chromosomal abnormalities in biopsies and lymphoid-cell lines from patients with Burkitt and non-Burkitt lymphomas. *Int J Cancer*. 1976;17(1):47-56.
3. Taub R, Kirsch I, Morton C, et al. Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. *Proc Natl Acad Sci USA*. 1982;79(24):7837-7841.
4. Ott G, Rosenwald A, Campo E. Understanding MYC-driven aggressive B-cell lymphomas: pathogenesis and classification. *Blood*. 2013;122(24):3884-3891.
5. Klapper W, Stoecklein H, Zeynalova S, et al; German High-Grade Non-Hodgkin's Lymphoma Study Group. Structural aberrations affecting the MYC locus indicate a poor prognosis independent of clinical risk factors in diffuse large B-cell lymphomas treated within randomized trials of the German High-Grade Non-Hodgkin's Lymphoma Study Group (DSHNHL). *Leukemia*. 2008;22(12):2226-2229.
6. Savage KJ, Johnson NA, Ben-Neriah S, et al. MYC gene rearrangements are associated with a poor prognosis in diffuse large B-cell lymphoma patients treated with R-CHOP chemotherapy. *Blood*. 2009;114(17):3533-3537.
7. Barrans S, Crouch S, Smith A, et al. Rearrangement of MYC is associated with poor prognosis in patients with diffuse large B-cell lymphoma treated in the era of rituximab. *J Clin Oncol*. 2010;28(20):3360-3365.
8. Hummel M, Bentink S, Berger H, et al; Molecular Mechanisms in Malignant Lymphomas Network Project of the Deutsche Krebshilfe. A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. *N Engl J Med*. 2006;354(23):2419-2430.
9. Scott DW, King RL, Staiger AM, et al. High-grade B-cell lymphoma with MYC and BCL2 and/or BCL6 rearrangements with diffuse large B-cell lymphoma morphology. *Blood*. 2018;131(18):2060-2064.
10. Obermann EC, Csato M, Dirnhofer S, Tzankov A. Aberrations of the MYC gene in unselected cases of diffuse large B-cell lymphoma are rare and unpredictable by morphological or immunohistochemical assessment. *J Clin Pathol*. 2009;62(8):754-756.
11. Ye Q, Xu-Monette ZY, Tzankov A, et al. Prognostic impact of concurrent MYC and BCL6 rearrangements and expression in de novo diffuse large B-cell lymphoma. *Oncotarget*. 2016;7(3):2401-2416.
12. Pedersen MØ, Gang AO, Poulsen TS, et al. MYC translocation partner gene determines survival of patients with large B-cell lymphoma with MYC- or double-hit MYC/BCL2 translocations. *Eur J Haematol*. 2014;92(1):42-48.
13. Copie-Bergman C, Cuillière-Dartigues P, Baia M, et al. MYC-IG rearrangements are negative predictors of survival in DLBCL patients treated with immunochemotherapy: a GELA/LYSA study. *Blood*. 2015;126(22):2466-2474.
14. Aukema SM, Kreuz M, Kohler CW, et al; Molecular Mechanisms in Malignant Lymphomas Network Project. Biological characterization of adult MYC-translocation-positive mature B-cell lymphomas other than molecular Burkitt lymphoma. *Haematologica*. 2014;99(4):726-735.
15. Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016;127(20):2375-2390.
16. Petrich AM, Gandhi M, Jovanovic B, et al. Impact of induction regimen and stem cell transplantation on outcomes in double-hit lymphoma: a multicenter retrospective analysis. *Blood*. 2014;124(15):2354-2361.
17. Friedberg JW. How I treat double-hit lymphoma. *Blood*. 2017;130(5):590-596.
18. Joos S, Falk MH, Lichter P, et al. Variable breakpoints in Burkitt lymphoma cells with chromosomal t(8;14) translocation separate c-myc and the IgH locus up to several hundred kb. *Hum Mol Genet*. 1992;1(8):625-632.
19. Johnson NA, Savage KJ, Ludkovski O, et al. Lymphomas with concurrent BCL2 and MYC translocations: the critical factors associated with survival. *Blood*. 2009;114(11):2273-2279.
20. Pedersen MØ, Gang AO, Poulsen TS, et al. Double-hit BCL2/MYC translocations in a consecutive cohort of patients with large B-cell lymphoma - a single centre's experience. *Eur J Haematol*. 2012;89(1):63-71.
21. Muñoz-Mármol AM, Sanz C, Tapia G, Marginet R, Ariza A, Mate JL. MYC status determination in aggressive B-cell lymphoma: the impact of FISH probe selection. *Histopathology*. 2013;63(3):418-424.
22. May PC, Foot N, Dunn R, Geoghegan H, Neat MJ. Detection of cryptic and variant IGH-MYC rearrangements in high-grade non-Hodgkin's lymphoma by fluorescence in situ hybridization: implications for cytogenetic testing. *Cancer Genet Cytogenet*. 2010;198(1):71-75.
23. Chong LC, Twa DDW, Mottok A, et al. Comprehensive characterization of programmed death ligand structural rearrangements in B-cell non-Hodgkin lymphomas. *Blood*. 2016;128(9):1206-1213.
24. Scott DW, Wright GW, Williams PM, et al. Determining cell-of-origin subtypes of diffuse large B-cell lymphoma using gene expression in formalin-fixed paraffin-embedded tissue. *Blood*. 2014;123(8):1214-1217.
25. Hans CP, Weisenburger DD, Greiner TC, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood*. 2004;103(1):275-282.
26. Bouamar H, Abbas S, Lin A-P, et al. A capture-sequencing strategy identifies IRF8, EBF1, and APRIL as novel IGH fusion partners in B-cell lymphoma. *Blood*. 2013;122(5):726-733.
27. Bosga-Bouwer AG, Haralambieva E, Booman M, et al. BCL6 alternative translocation breakpoint cluster region associated with follicular lymphoma grade 3B. *Genes Chromosomes Cancer*. 2005;44(3):301-304.

28. Einerson RR, Law ME, Blair HE, et al. Novel FISH probes designed to detect IGK-MYC and IGL-MYC rearrangements in B-cell lineage malignancy identify a new breakpoint cluster region designated BVR2. *Leukemia*. 2006;20(10):1790-1799.
29. Fabris S, Storlazzi CT, Baldini L, et al. Heterogeneous pattern of chromosomal breakpoints involving the MYC locus in multiple myeloma. *Genes Chromosomes Cancer*. 2003;37(3):261-269.
30. Walker BA, Wardell CP, Brioli A, et al. Translocations at 8q24 juxtapose MYC with genes that harbor superenhancers resulting in overexpression and poor prognosis in myeloma patients. *Blood Cancer J*. 2014;4(3):e191.
31. Kroenlein H, Schwartz S, Reinhardt R, et al. Molecular analysis of the t(2;8)/MYC-IGK translocation in high-grade lymphoma/leukemia by long-distance inverse PCR. *Genes Chromosomes Cancer*. 2012;51(3):290-299.
32. Huret J-L, Ahmad M, Arsaban M, et al. Atlas of genetics and cytogenetics in oncology and haematology in 2013. *Nucleic Acids Res*. 2013;41(Database Issue):D920-D924.
33. Bertrand P, Bastard C, Maingonnat C, et al. Mapping of MYC breakpoints in 8q24 rearrangements involving non-immunoglobulin partners in B-cell lymphomas. *Leukemia*. 2007;21(3):515-523.
34. Weinberg OK, Ai WZ, Mariappan MR, Shum C, Levy R, Arber DA. "Minor" BCL2 breakpoints in follicular lymphoma: frequency and correlation with grade and disease presentation in 236 cases. *J Mol Diagn*. 2007;9(4):530-537.
35. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26(5):589-595.
36. Rausch T, Zichner T, Schlattl A, Stütz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics*. 2012;28(18):i333-i339.
37. Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29(1):24-26.
38. 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(7):1256-1265.
39. Busch K, Keller T, Fuchs U, et al. Identification of two distinct MYC breakpoint clusters and their association with various IGH breakpoint regions in the t(8;14) translocations in sporadic Burkitt-lymphoma. *Leukemia*. 2007;21(8):1739-1751.
40. Richter J, Schlesner M, Hoffmann S, et al; ICGC MMML-Seq Project. Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. *Nat Genet*. 2012;44(12):1316-1320.
41. Hillion J, Mecucci C, Avenirin A, et al. A variant translocation t(2;18) in follicular lymphoma involves the 5' end of bcl-2 and Ig kappa light chain gene. *Oncogene*. 1991;6(1):169-172.
42. Dyer MJ, Zani VJ, Lu WZ, et al. BCL2 translocations in leukemias of mature B cells. *Blood*. 1994;83(12):3682-3688.
43. Ryan RJH, Drier Y, Whitton H, et al. Detection of enhancer-associated rearrangements reveals mechanisms of oncogene dysregulation in B-cell lymphoma. *Cancer Discov*. 2015;5(10):1058-1071.
44. Bunting KL, Soong TD, Singh R, et al. Multi-tiered reorganization of the genome during B cell affinity maturation anchored by a germinal center-specific locus control region. *Immunity*. 2016;45(3):497-512.
45. Lefranc M-P, Giudicelli V, Duroux P, et al. IMGT®, the international ImMunoGeneTics information system® 25 years on. *Nucleic Acids Res*. 2015; 43(Database Issue):D413-D422.
46. Lieber MR. Mechanisms of human lymphoid chromosomal translocations. *Nat Rev Cancer*. 2016;16(6):387-398.
47. Bahr C, von Paleske L, Uslu VV, et al. A Myc enhancer cluster regulates normal and leukaemic haematopoietic stem cell hierarchies [published correction appears in *Nature* 2018;558(7711):E4]. *Nature*. 2018;553(7689):515-520.
48. Saeki K, Miura Y, Aki D, Kurosaki T, Yoshimura A. The B cell-specific major raft protein, Raftlin, is necessary for the integrity of lipid raft and BCR signal transduction. *EMBO J*. 2003;22(12):3015-3026.