In-depth characterization of intratumoral heterogeneity in refractory B-cell non-Hodgkin lymphoma through the lens of a Research Autopsy Program

Keren Isaev,[1] Ting Liu,[1] Mehran Bakhtiari,[1] Kit Tong,[1] Rashmi Goswami,[2] Bernard Lam,[3] Ilinca Lungu,[3] Paul M. Krzyzanowski,[3] Amit Oza,[1] Neesha Dhani,[1] Anca Prica,[1] Michael Crump[1] and Robert Kridel[1]

[1]Princess Margaret Cancer Center - University Health Network; [2]Sunnybrook Health Sciences Center and [3]Ontario Institute for Cancer Research, Toronto, ON, Canada

Correspondence: R. Kridel
robert.kridel@uhn.ca

Received: February 23, 2022.
Accepted: June 15, 2022.
Prepublished: June 23, 2022.

https://doi.org/10.3324/haematol.2022.280900

Abstract

Intratumoral heterogeneity (ITH) provides the substrate for tumor evolution and treatment resistance, yet is remarkably understudied in lymphoma, due to the often limited amount of tissue that gets sampled during the routine diagnostic process, generally from a single nodal or extranodal site. Furthermore, the trajectory of how lymphoma, and especially non-Hodgkin lymphoma, spreads throughout the human body remains poorly understood. Here, we present a detailed characterization of ITH by applying whole-genome sequencing to spatially separated tumor samples harvested at the time of autopsy (n=24) and/or diagnosis (n=3) in three patients presenting with refractory B-cell non-Hodgkin lymphoma. Through deconvolution of bulk samples into clonal mixtures and inference of phylogenetic trees, we found evidence that polyclonal seeding underlies tumor dissemination in lymphoma. We identify mutation signatures associated with ancestral and descendant clones. In our series of patients with highly refractory lymphoma, the determinants of resistance were often harbored by founding clones, although there was also evidence of positive selection of driver mutations, likely under the influence of therapy. Lastly, we show that circulating tumor DNA is suitable for the detection of ancestral mutations but may miss a significant proportion of private mutations that can be detected in tissue. Our study clearly shows the existence of intricate patterns of regional and anatomical evolution that can only be disentangled through multi-regional tumor tissue profiling.

Introduction

Intratumoral heterogeneity (ITH) is a fundamental hallmark of cancer and is recognized at all levels of the tumor ecosystem, encompassing genetic, epigenetic and metabolic variation between individual cellular populations.[1] The progressive acquisition of somatic alterations across the genome is a major determinant of tumor evolution, leading to clonal expansions under selective pressures such as antineoplastic therapy.[2] A recent study from the Pan-Cancer Analysis of Whole Genomes (PCAWG) initiative identified almost universal evidence of ITH across 2,658 samples from 38 cancer types, with frequent evidence of both branching and linear evolution patterns.[3] Lymphomas arise from the malignant transformation of normal lymphocytes and are accordingly classified among malignancies of the hematopoietic system. While bone marrow and peripheral blood involvement can be observed, the most common lymphoma types often present without a noticeable circulatory component and form tissue masses in lymph nodes and/or extranodal sites.[4] Yet, unlike most solid malignancies where the site of primary disease is typically well defined and where modes of spread between the primary and metastatic sites have been elucidated using multi-regional sequencing approaches,[5,6] the knowledge about the modes of lymphoma dissemination remains limited. This observation stands in contrast to the other end of the evolutionary trajectory where understanding of the cellular origin of most lymphomas is increasing and can be traced back to specific stages of B- or T-cell differentiation.[7]

While surgical biopsies are critically important for accu-

rate diagnosis, larger surgical resections are not considered a part of the treatment strategy for lymphoma. This is in contrast to solid malignancies in which localized neoplasms are surgically removed, aiming to provide cure, with or without neoadjuvant or adjuvant therapies. As a result, the spatial distribution of subclones has been described in solid tumors to a much more granular degree than in hematological malignancies. To date, only a few studies have explored multisite heterogeneity in lymphoma, at a relatively small scale, typically comparing no more than two tumor sites.[8–11] One study reported the application of whole-exome sequencing to 12 samples from one patient with extranodal NK/T-cell lymphoma.[11] Overall, the full extent of spatial and temporal ITH largely remains to be uncovered in lymphoma. Herein, we propose to fill this knowledge gap by leveraging sample resources from an institutional Research Autopsy Program and applying state-of-the-art clonal deconvolution and phylogenetic reconstruction methods. Our goal was to provide an in-depth description of ITH in patients with highly refractory/resistant lymphoma and to enhance our understanding of clonal dissemination in lymphoma.

Methods

Research Autopsy Program

An institutional Research Autopsy Program was open at our institution between 2014 and 2017 and enrolled 115 patients after informed consent was obtained. The program, described in Bavi *et al.*, was built on the integration of multiple interfaces (patient-physician, autopsy-pathology, biospecimen-science), highlighting its integrated and collaborative nature.[12] Four patients with refractory non-Hodgkin lymphoma were enrolled: patient 1 with blastoid mantle cell lymphoma; patient 2 with primary mediastinal B-cell lymphoma; patient 3 with high-grade B-cell lymphoma with double hit (*MYC* and *BCL2* translocations); and patient 4 with diffuse large B-cell lymphoma. Patient 4 was excluded from this study as this patient presented with neurolymphomatosis and no autopsy samples with sufficient quality and/or lymphoma involvement were available. This study was approved by the University Health Network (UHN) Research Ethics Board (CAPCR 16-6342).

Samples

A total of 12, 39 and 25 tissue samples were harvested at the time of autopsy for patients 1, 2 and 3, respectively, and stored in the UHN biobank as frozen tissue. After pathology review of hematoxylin and eosin (H&E) sections cut from frozen tissue and exclusion of samples with low tumor content and/or necrosis, 3, 4 and 17 samples were available for DNA isolation. In addition, for patient 3, three diagnostic formalin-fixed and paraffin-embedded tissue (FFPET) biopsies from the time of diagnosis were available. Samples are listed in the *Online Supplementary Table S1*. DNA was extracted from frozen tissue samples using the Qiagen AllPrep DNA/RNA kit, from FFPET samples using the Qiagen AllPrep DNA/RNA FFPE kit and from peripheral blood cells using the QIAamp DNA Mini Kit. The germline control samples were determined to be free of tumor cell contamination through the lack of amplifiable, patient-specific immunoglobulin gene rearrangements.

Whole-genome sequencing and processing

Paired end sequencing (2x150 base pairs) was performed using an Illumina HiSeqX. Reads were mapped to the hg19 reference genome using bwa-mem[13] and duplicate reads were marked using Picard Tools, followed by local realignment and base quality score recalibration using GATK (4.0.5.1).[14] Somatic single nucleotide variants (SNV) and insertions and deletions (indels) were called using Mutect2[15] in tumor-normal mode with default parameters following the best practice workflow.[16] GnomAD population variants included in the GATK bundle were used as the germline resource. We also used Strelka (2.9.10)[17] to call SNV and indels and merged the two sets of mutation calls to include only their intersection. Mutations found in ENCODE blacklisted regions were removed.[18] Mutations were annotated by Annovar.[19] Additional filters were applied to remove mutations whose depth was lower than 60X (or 30X for FFPET samples), X and Y chromosome mutations, mutations with a population allele frequency higher than 0.001, and mutations annotated by GnomAD and not also annotated in COSMIC.

Results

Overview of patient and samples

In order to obtain detailed accounts of non-Hodgkin lymphoma (NHL) evolutionary trajectories, we collected samples from three patients who had been consented within the framework of an institutional Research Autopsy Program (Figures 1A, 3A and 5A).[12] All three patients had aggressive clinical presentations with death resulting from primary refractory or early relapsing disease. Patient 1 was a 50-60 year-old female patient with a diagnosis of stage IV blastoid mantle cell lymphoma (MCL), achieving a partial remission after six cycles of alternating R-CHOP/R-DHAP, followed by autologous stem cell transplantation and rituximab maintenance. She experienced early central nervous system relapse, 5 months after stem cell transplantation, treated with whole-brain radiation and intrathecal chemotherapy but ultimately died 1 year and 4 months after initial diagnosis. Patient 2 was a 20-30 year-

old male patient with a diagnosis of stage IIE primary refractory mediastinal B-cell lymphoma (PMBCL) with bulk, treated with six cycles of R-CHOP but found to have refractory disease on end-of-treatment positron emission tomography–computed tomography scan. He then received one cycle of GDP without response, followed by radiation, pembrolizumab, further palliative radiation and single agent chemotherapy, before dying from lymphoma 1 year and 2 months after diagnosis. Patient 3 was a 50-60-year old male patient with a diagnosis of stage IV high-grade B-cell lymphoma (HGBCL) with double hit rearrangement (*BCL2* and *MYC)* and large bulk in the retroperitoneum, initially treated with one cycle of R-CHOP followed by two cycles of DA-EPOCH-R without response, then treated with two cycles of GDP without response and followed by palliative radiation. This patient died within 6 months of diagnosis. All three patients had been treated at our institution before CAR T-cell therapy became widely available.

Whole-genome sequencing (WGS) was performed on a total of 27 tumor samples that included 24 cryopreserved autopsy samples from patients 1 (n=3), 2 (n=4) and 3 (n=17), as well as three spatially separated biopsies from the time of diagnosis, available as FFPET blocks, from patient 3 (*Online Supplementary Table S1*). The average coverage was 66X ± 3.1 standard deviations (SD) and 46X ± 11.3 for cryopreserved and FFPET samples, respectively. In addition, peripheral blood mononuclear cells were sequenced for all three patients (average coverage 36X ± 1.8) (*Online Supplementary Table S2*). Plasma was available for all three patients and was used for targeted sequencing using a hybrid capture panel spanning 158 kB. The mean on target coverage for ctDNA was 4,632X in the uncollapsed, de-duplicated bam files and 603X in the bam files with reads collapsed based on duplex consensus sequences and singleton correction.

Patient 1

Identification of somatic genetic alterations at the level of the whole genome

In this patient with mantle cell lymphoma, the average number of somatic SNV and indels was 8,776 +/- 2,249

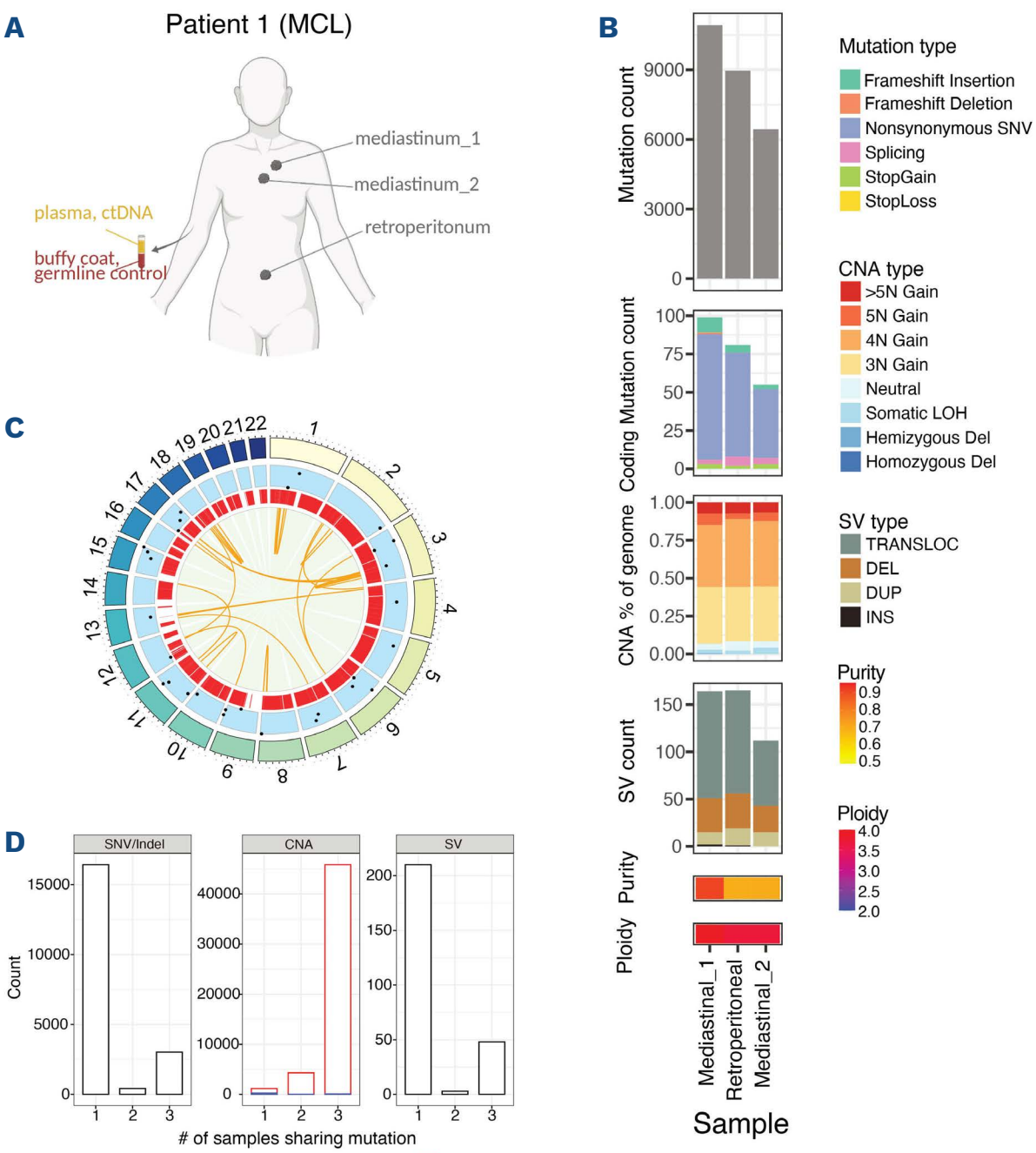


Figure 1. Overview of patient 1 (mantle cell lymphoma). (A) Spatially distinct tumor samples were collected from 3 tumor sites along with plasma and buffy coat. (B) High level overview of single nucleotide variants/insertions and deletions (SNV/indels), copy number aberration (CNA) and structural variant (SV) counts across samples. Purity and Ploidy estimates from Sequenza are shown. (C) Circos plot showing shared genetic aberrations. Chromosomes are shown in the outermost circle, followed by a track showing single nucleotide variants, a track showing copy number variation and finally structural variation shown in the innermost circle. (D) Plots showing the number of genetic events found in one or more samples. LOH: loss of heterozygosity; TRANSLOC: translocation; DEL: deletion; DUP: duplication; INS: insertion; MCL: mantle cell lymphoma. Panel (A) was created by BioRender.com.

SD per sample (Figure 1B). The average number of structural variants (SV) including larger insertions, deletions and translocations was 147 +/- 30 SD per sample. SNV, CNA and SV that were common to all samples (putatively ancestral) are shown in Figure 1C. The degree to which somatic alterations were shared across samples differed. The majority of mutations were private, i.e., observed in just one sample (n=16,417, 83%, Figure 1D), while copy number gains affected most genes, with evidence of whole-genome duplication in all samples (Figure 1B). We found a high confidence shared, i.e., putatively ancestral, splicing mutation in *TP53* affecting the splice acceptor site of intron 4 (*Online Supplementary Table S3*). This case also had a shared splicing mutation involving *FUBP1*, a gene that has previously been identified as a cancer-associated gene in chronic lymphocytic leukemia (CLL)[20] as well as in an MCL case arising from underlying CLL.[21] As expected, the t(11:14) chromosomal translocation was found in all samples. We compared SNV and indels detected in ctDNA with mutation calls within matching genomic regions from WGS of tissue biopsies. In this patient, in whom the number of mutations identified in the tumor samples was low, none of these mutations was found in ctDNA (0/2) (Figure 2A).

Phylogenetic tree and tumor clonal evolution reconstruction using spatially distinct samples

In a first analysis of the evolutionary relationships between spatially separated tumor samples, using Treeomics,[22] we found clear evidence of branching evolution (*Online Supplementary Figure 1A*). The majority of analyzed SNV (82%) were exclusive to individual samples, rather than shared. In order to delineate clonal origins across samples within a given patient, we combined mutation (n=3,146, *Online Supplementary Table S4*), copy number and purity estimates using Pyclone-VI.[23] We observed one founding clone (population 1) harboring driver gene mutations including *TP53* (X126_splice) and *FUBP1* (X642_splice) (Figure 2B). Population 2 was prevalent in the Mediastinal_2 and Retroperitoneal samples while population 3 was only prevalent in Mediastinal_1 (Figure 2C; *Online Supplementary Figure 1B*) and contained an *EGFR* (A786V) mutation.

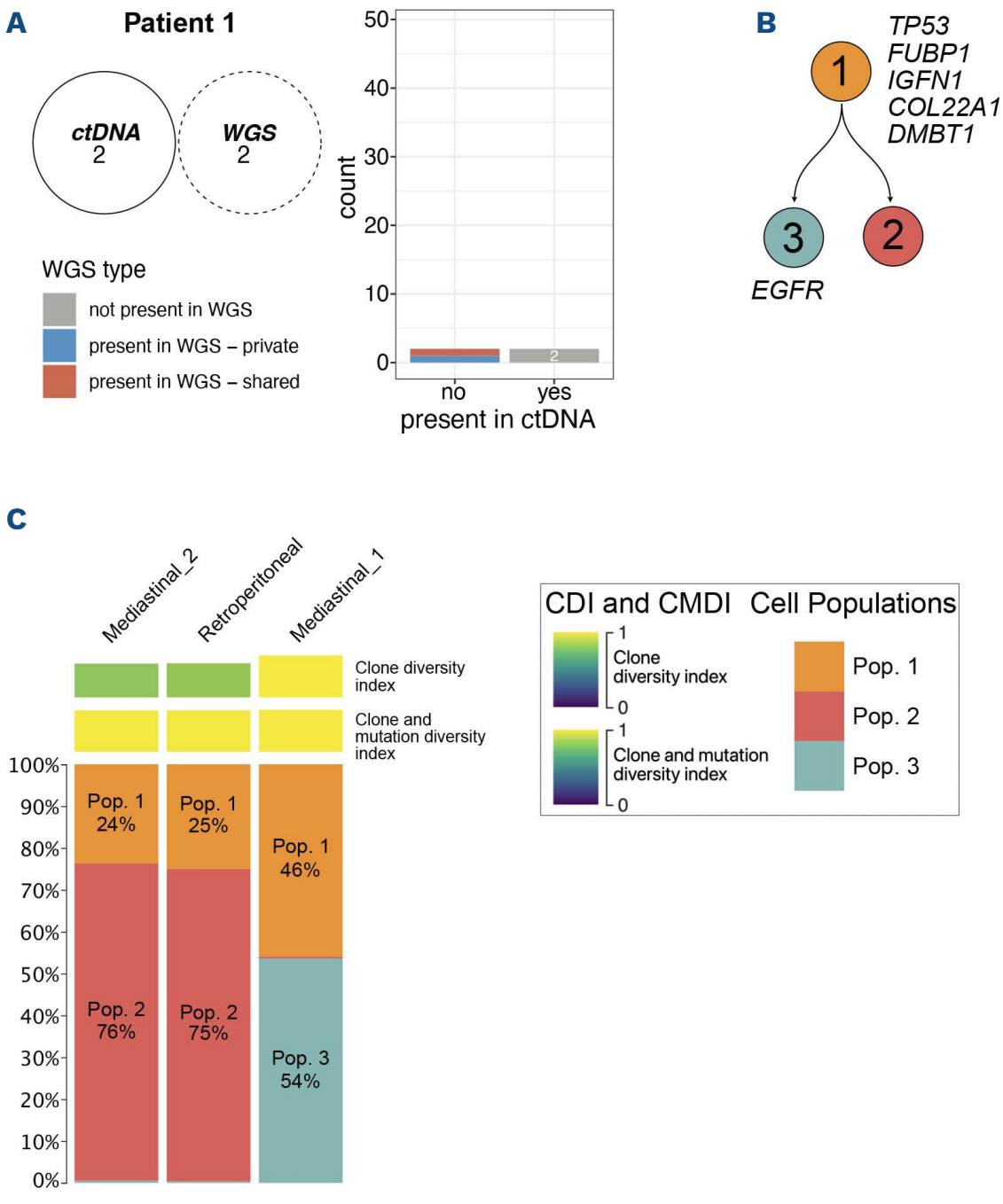


Figure 2. In depth analysis of patient 1 mutations (mantle cell lymphoma). (A) The Venn diagram shows the number of mutations identified in either circulating tumor DNA (ctDNA) or any of the whole-genome sequencing (WGS) samples, and their overlap. The barplots show the count of mutations, based on whether they are detected in ctDNA or not (x axis), and based on whether they are detected in WGS of tissue samples or not (coded by color). The analysis of WGS herein is restricted to the specific genomic regions covered by the ctDNA targeted sequencing panel. (B) Clonal phylogeny tree, as inferred by pyclone and pairtree. Driver genes with mutations in specific subclones are labeled. (C) Subclonal population frequencies across samples, with each sample annotated by its clone diversity index (CDI) and clone mutation diversity index (CMDI).

Distribution of driver gene mutations and mutation signatures across clonal lineage trees

We analyzed driver mutations defined as those coding SNV that affected previously described driver genes in the lymphoma context, as well as those specific positions that are reported in the COSMIC database. The number of such mutations was low and ranged from zero to three across the three clonal populations (*Online Supplementary Figure S1C*). We associated COSMIC mutation signatures with cell populations identified through our phylogenetic analysis. Several mutation signatures were found to account for mutations in the dominant founding clone (population 1) including a grouped set consisting of SBS9 (somatic hypermutation in lymphoid cells) and SBS85 (activation-induced cytidine deaminase [AID] in lymphoid cells) (*Online Supplementary Figure S1D* to *E*). While populations 2 and 3 contained far fewer mutations in this analysis, population 2 also gained the SBS17a signature (related to SBS17b and chemotherapy treatment). Meanwhile, mutations in population 3 were only present in the Mediastinal_1 sample and were related primarily to the SBS86 signature. Overall, mutation signatures associated with chemotherapy treatment and somatic hypermutation were observed across all samples.

Patient 2

Identification of somatic genetic alterations at the level of the whole genome

In this patient with primary mediastinal B-cell lymphoma, the average number of all mutations was 373,288 +/- 6,482 and the average number of SV across samples was 89 +/- 18 per sample (Figure 3B). Unlike our observations for patient 1, the majority of mutations were shared across all four tumor samples analyzed (n=297,045, 59%) (Figure

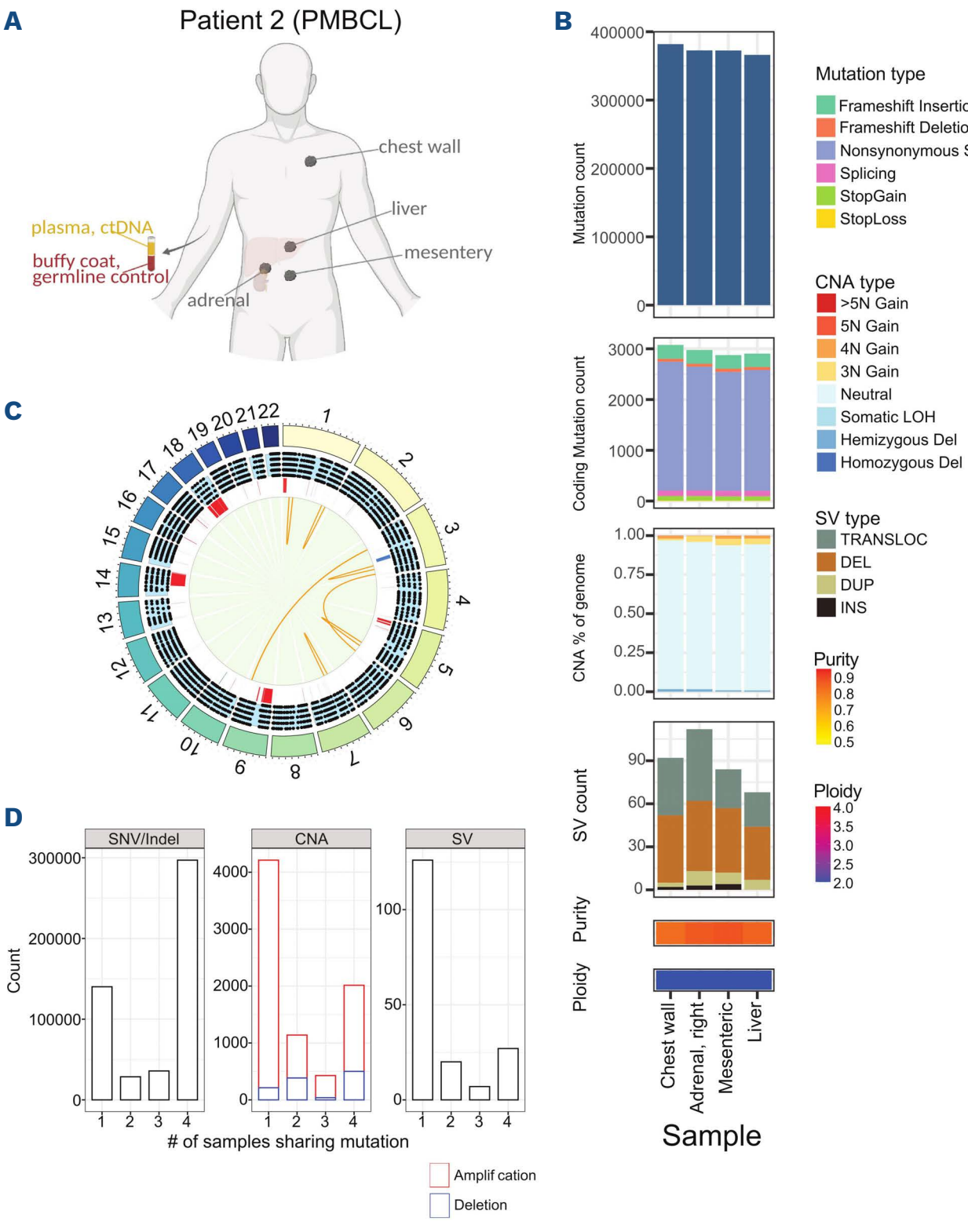


Figure 3. Overview of patient 2 (primary mediastinal B-cell lymphoma). (A) Spatially distinct tumor samples were collected from 4 tumor sites along with plasma and buffy coat. (B) High level overview of single nucleotide variants/insertions and deletions (SNV/indels), copy number aberration (CNA) and structural variant (SV) counts across samples. Purity and Ploidy estimates from Sequenza are shown. (C) Circos plot showing shared genetic aberrations. Chromosomes are shown in the outermost circle, followed by a track showing SNV, a track showing copy number variation and finally structural variation shown in the innermost circle. (D) Plots showing the number of genetic events found in one or more samples. CNA: copy number aberration; LOH: loss of heterozygosity; TRANSLOC: translocation; DEL: deletion; DUP: duplication; INS: insertion; PMBCL: primary mediastinal B-cell lymphoma. Panel (A) was created by BioRender.com.

3C and D). On the contrary, a substantial fraction of genes was affected by CNA in only one of the four samples. Samples from this patient harbored very high numbers of SNV which made it more challenging to assign pathogenic roles to individual gene mutations (*Online Supplementary Table S3*). However, certain genes that are commonly altered in B-cell lymphoma were mutated across all samples in this patient, including a nonsense mutation in *ARID1A* (Q74*), a recurrent hotspot mutation in *B2M* (X23_splice), a frameshift deletion in *BCL10* (S167Lfs*6) and a splicing mutation in *TNFAIP3* (X99_splice). The number of mutations that overlapped between WGS and ctDNA was 24, representing 71% of all mutations detected in circulating tumor DNA (ctDNA) and an average of 49% (45-47%) of all mutations found in the four tumor tissue biopsies (Figure 4A). Thus, a high percentage of ctDNA mutations was found across tumor samples. Further, the mutations present in ctDNA and in tissue based WGS were exclusively shared (24/24, 100%), compared to mutations present in WGS and not detected in ctDNA (17/46, 37%; chi-square *P*= 0.002).

Phylogenetic tree and tumor clonal evolution reconstruction using spatially distinct samples

Using results from Treeomics,[22] we observed a high percentage of coding SNV (69%) that were found in the trunk rather than in branches (*Online Supplementary Figure S2A*). A total of 22,402 mutations (*Online Supplementary Table S4*) were used for clonal phylogeny reconstruction using Pyclone-VI.[23] We observed an ancestral clone (population 1) with 13,385 mutations (60% of studied mutations) (Figure 4B). Population 2 emerged from the dominant clone with mutations including *PRDM9* (R842H) before forming two phylogenetic clades, founded by populations 3 and 4, respectively. Population 5 descended from population 3 and branched into populations 7 and 8. Population 7 was dominant in the adrenal gland and included mutations in *DNMT3A* (T835M) and *TFEB* (R13W), while population 8 was only found in the chest wall and contained *DDX21* (A688T) and *SOCS1* (S205N) mutations. Populations 6 and 9 descended from population 4, with population 9 further branching into population 10 found in the liver and

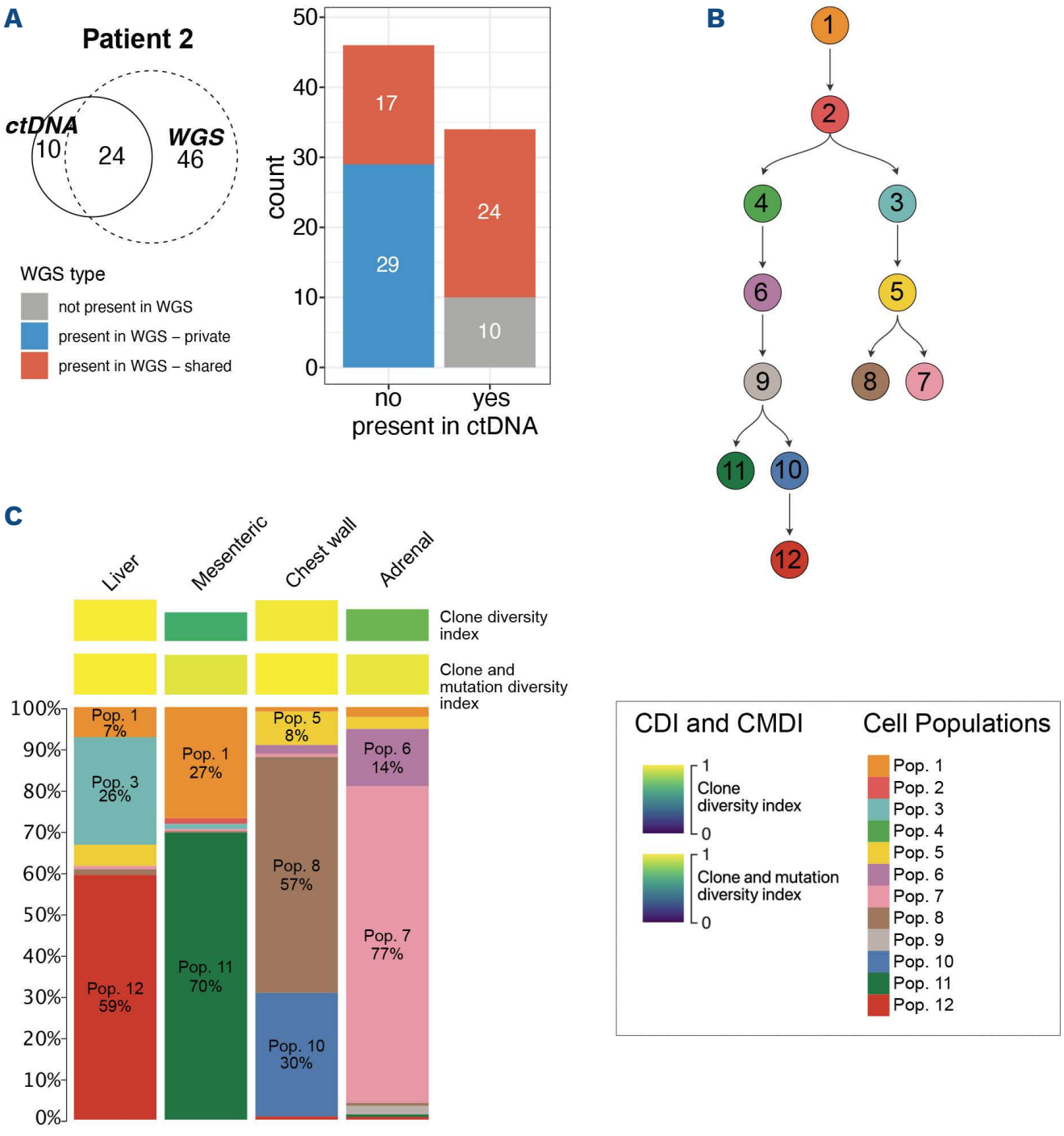


Figure 4. In depth analysis of patient 2 mutations (primary mediastinal B-cell lymphoma). (A) The Venn diagram shows the number of mutations identified in either circulating tumor DNA (ctDNA) whole-genome sequencing (WGS) samples, and their overlap. The barplots show the count of mutations, based on whether they are detected in ctDNA or not (x axis), and based on whether they are detected in WGS of tissue samples or not (coded by color). The analysis of WGS herein is restricted to the specific genomic regions covered by the ctDNA targeted sequencing panel. (B) Clonal phylogeny tree, as inferred by pyclone and pairtree. (C) Subclonal population frequencies across samples, with each sample annotated by its clone diversity index (CDI) and clone mutation diversity index (CMDI).

the chest wall and population 11 found only in the mesenteric lymph node. At least three of the metastatic sites were composed of clones from both phylogenetic clades, suggesting polyphyletic dissemination (Figure 4C; *Online Supplementary Figure S2B*).

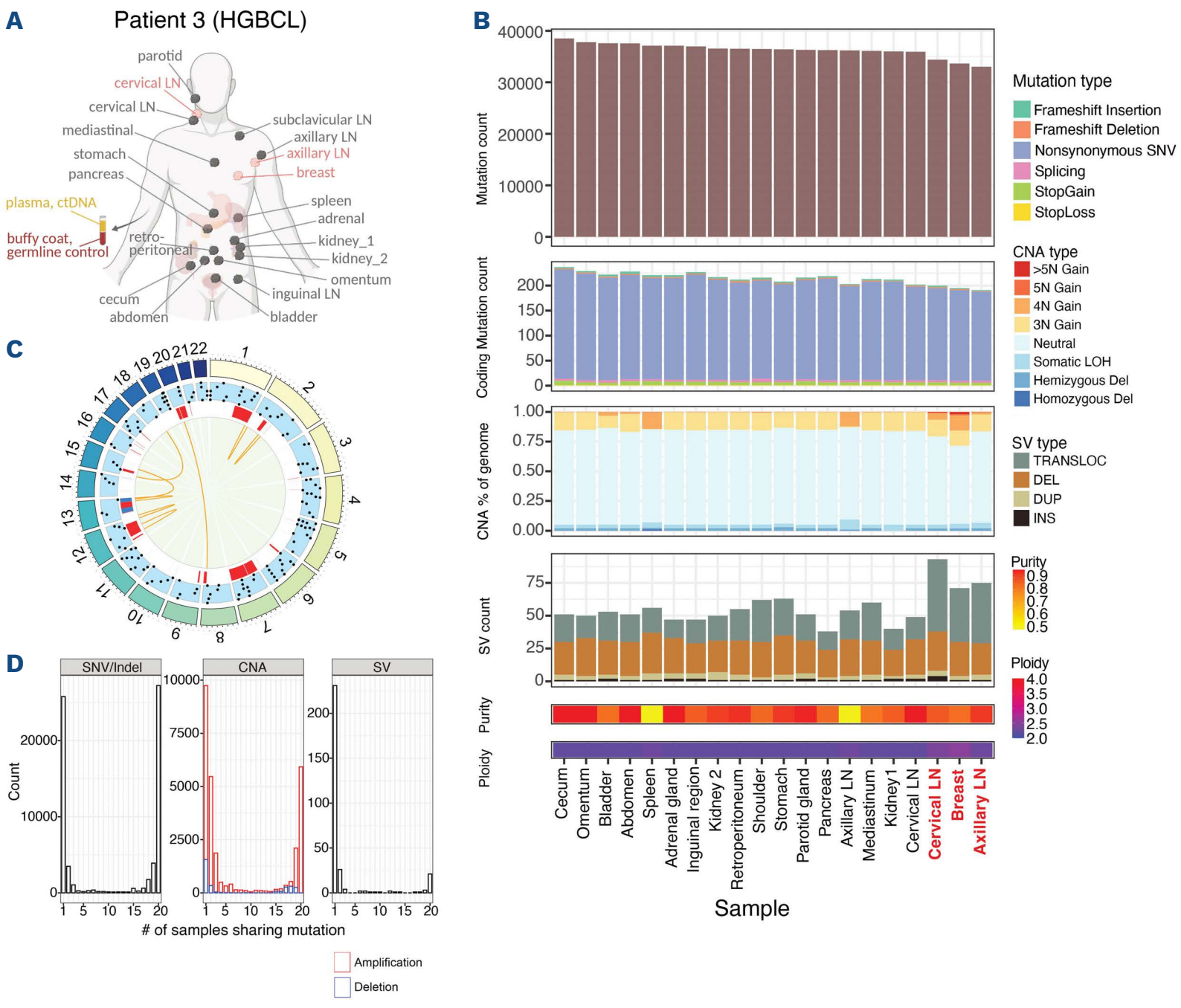
Distribution of driver gene mutations and mutation signatures across clonal lineage trees

The vast majority of driver mutations was found in the ancestral clone (n=227 or 10% of coding mutations in this clone), although up to 23 additional driver and COSMIC mutations were found in descending clones (population 7, 8, 11 and 12) (*Online Supplementary Figure S2C*). Interestingly, these cell populations were each found in only one sample: adrenal gland, chest wall, mesenteric lymph node and liver, respectively. Three sets of mutation signatures contributed to mutation patterns in the founding clone (population 1) and most descending populations (*Online Supplementary Figure S2D and E*). These mutation signatures included SBS36;SBS35 (defective base excision repair, including DNA damage due to reactive oxygen species), SBS9;SBS85 (somatic hypermutation and AID-induced mutagenesis) and SBS15 (defective DNA mismatch repair). Possibly, these signatures reflect the high number of mutations observed in this patient. Meanwhile, several signatures were only observed in some populations, including SBS11;SBS32 (alkylating agent treatment and prior treatment with azathioprine) which were only associated with population 10. Further, the SBS17b signature was found across multiple subclonal populations and not in the founding population and has been associated with chemotherapy treatment and damage through reactive oxygen elements.

Patient 3

Identification of somatic genetic alterations at the level of the whole genome

In this patient with high-grade B-cell lymphoma and double hit of *BCL2* and *MYC*, the average number of all mutations per sample was 36,316 +/- 1,337 (Figure 5B). The



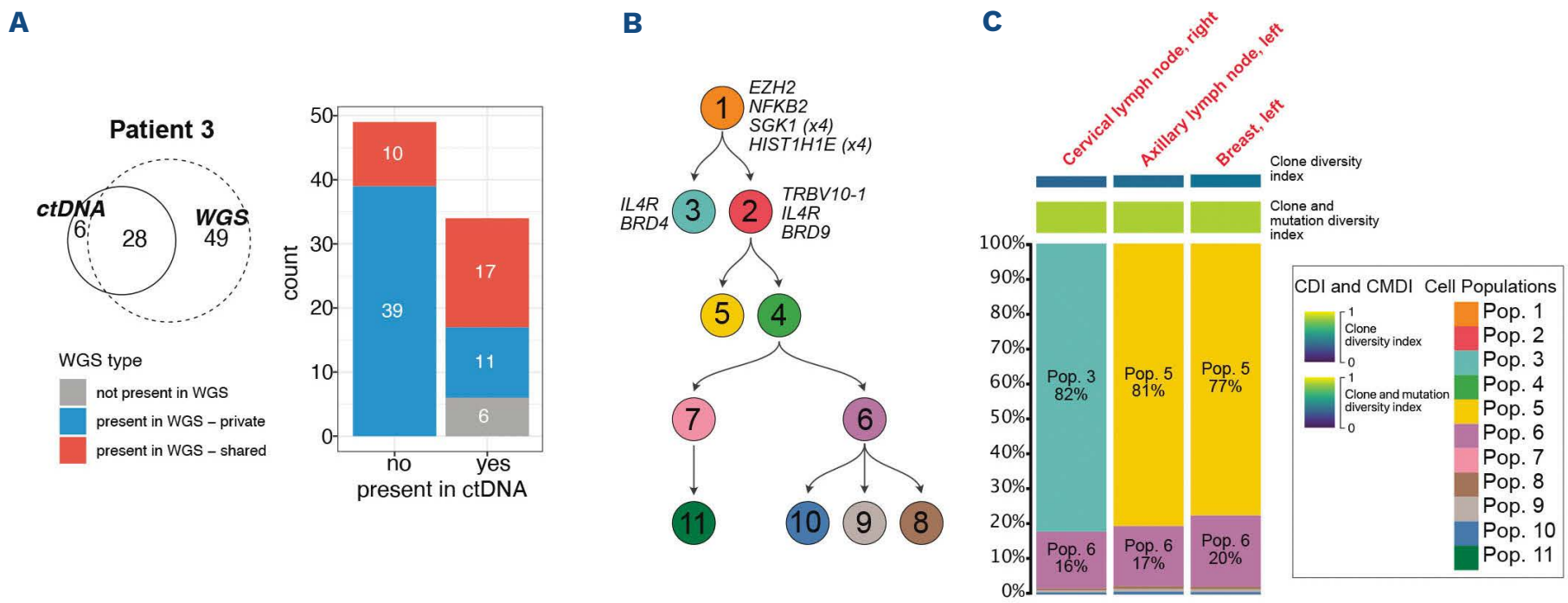
Continued on following page.

Figure 5. Overview of patient 3 (high-grade B-cell lymphoma). (A) Spatially distinct tumor samples were collected from 20 tumor sites (3 at diagnosis indicated in pink and 17 at autopsy) along with plasma and buffy coat. (B) High level overview of single nucleotide variants/insertions and deletions (SNV/indels), copy number aberration (CNA) and structural variant (SV) counts across samples. Purity and Ploidy estimates from Sequenza are shown. (C) Circos plot showing shared genetic aberrations. Chromosomes are shown in the outermost circle, followed by a track showing SNV, a track showing copy number variation and finally structural variation shown in the innermost circle. (D) Plots showing the number of genetic events found in one or more samples. LOH: loss of heterozygosity; TRANSLOC: translocation; DEL: deletion; DUP: duplication; INS: insertion; HGBCL: high-grade B-cell lymphoma. Panel (A) was created by BioRender.com.

average number of SV was 56 +/- 13. Overall, translocations and deletions were more common than insertions and duplications. Most mutations were either private (n=25,805, 39%) or found across all samples (n=27,243, 41%) with a uniform spread of mutations across shared samples (Figure 5C and D). Similarly, most genes were either affected by a CNA in only one sample or in all samples from the patient, while most SV were private. All samples harbored an *EZH2* mutation (A687V) and a *KMT2D* splice site mutation (*Online Supplementary Table S3*). Fifteen of 20 tumor samples (including the diagnostic FFPET specimens) harbored an additional indel in *KMT2D* (V2909fs). We identified a total of eight distinct gene mutations involving genes encoding linker histone H1 (*HIST1H1B, HIST1H1C, HIST1H1D*), found in eight to 20 of all tumor samples for this case. Additional mutations found in all samples included genes encoding proteins involved in signaling (*IL4R*: I242N; *JAK2*: V19A; *FOXO1*: W25G and *SGK1*: P85S, L149V, D167N and C824T). Moreover, ancestral mutations were identified in *TNFRSF14* (C34R) and *POU2F2* (T227A). Interestingly, five distinct *TP53* SNV or indel were found and were private with the exception of one mutation found in three autopsy samples. None of the *TP53* mutations was identified in the diagnostic FFPE samples. As expected, all samples from patient 3 harbored t(14:18) and *MYC* (non-immunoglobulin) translocations. The overlap of mutations between ctDNA and WGS consisted of a mean of 22 mutations (range, 20-25), with an average of 65% of these overlapping mutations detected in ctDNA (range, 59-74%) and an average of 52% (range, 47- 60%) of these mutations found in the tumor tissue samples (Figure 6A). Thus, the percentage of ctDNA mutations found in tissue samples was higher than the percentage of mutations from tissue samples found in ctDNA. Further, the mutations present in both ctDNA and WGS were more frequently shared than private, when compared to mutations detected only in WGS (17/28, 61% *vs.* 10/49, 20%; chi-square $P<0.001$). Hence, these results show that capture-based targeted sequencing of ctDNA provides complementary information to multi-regional WGS and has a higher likelihood of detecting clonal, rather than subclonal mutations.

Phylogenetic tree and tumor clonal evolution reconstruction using spatially distinct samples

Patient 3 allowed for detailed analysis of spatial relationships due to the number of available samples. Treeomics showed that the left breast and the left axilla diagnostic samples formed early diverging branches in the phylogeny, where the former had a private *PTEN* (V175E) mutation and the latter had a private *MYC* (S388R) mutation (*Online Supplementary Figure S3A*). Interestingly, the anatomically nearby axilla sample collected at autopsy also branched off earlier on with an acquired *TP53* (V172G) mutation. The third diagnostic sample from the right neck formed a branch from the main phylogeny along with the sample collected from the cervical lymph node upon autopsy. Additional diverging events led to one branch composed of the two kidney samples, adrenal gland and the spleen and one composed of the abdomen, cecum and inguinal region. Thus, the genetic profiles of the diagnostic samples were more similar to each other than to most autopsy samples. Moreover, these results reveal genetic relatedness between samples from spatially adjacent, but also between certain samples from distant sites and suggest a partial anatomical order in which lymphoma cells seeded throughout the body. For the Pyclone-VI analysis, we evaluated 12,024 mutations across 20 samples (Figure 6B; *Online Supplementary Table S4*), revealing a founding clone (population 1) containing 6,240 mutations (52% of those analyzed) including those in driver genes such as *EZH2* (A687V), *NFKB2* (A89V), *SGK1* (P95S, L159V, D177N, A285V), and *HIST1H1E* (S104F, A164P, P201S, G103D). This ancestral population branched into population 2 and population 3 where the latter was only prevalent in the cervical lymph node samples from diagnosis and autopsy, suggesting spatial localization, and contained mutations in *IL4R* (F229C) and *BRD4* (I146K). Population 4 and population 5 then diverged where the latter was prevalent only in the remaining two diagnostic samples and the axillary lymph node sample from autopsy. Population 4 further diverged into population 6 and population 7 with the former prevalent across all samples at varying levels and the latter found only in a subset of samples. Population 6 diverged into three subclones (populations 8, 9 and 10) present at mutually exclusive spatially distinct samples. The clone diversity index values were lowest in the three FFPE diagnostic samples as well as the autopsy samples from the axillary and cervical lymph nodes. These two autopsy sites had clonal compositions that resembled the composition found in the diagnostic biopsies, with a predominance of either populations 3 or 5 based on laterality (right-sided samples composed of populations 3 and 6, and left-sided samples composed of populations 5 and 6) (Figure 6C and D; *Online Supple-*



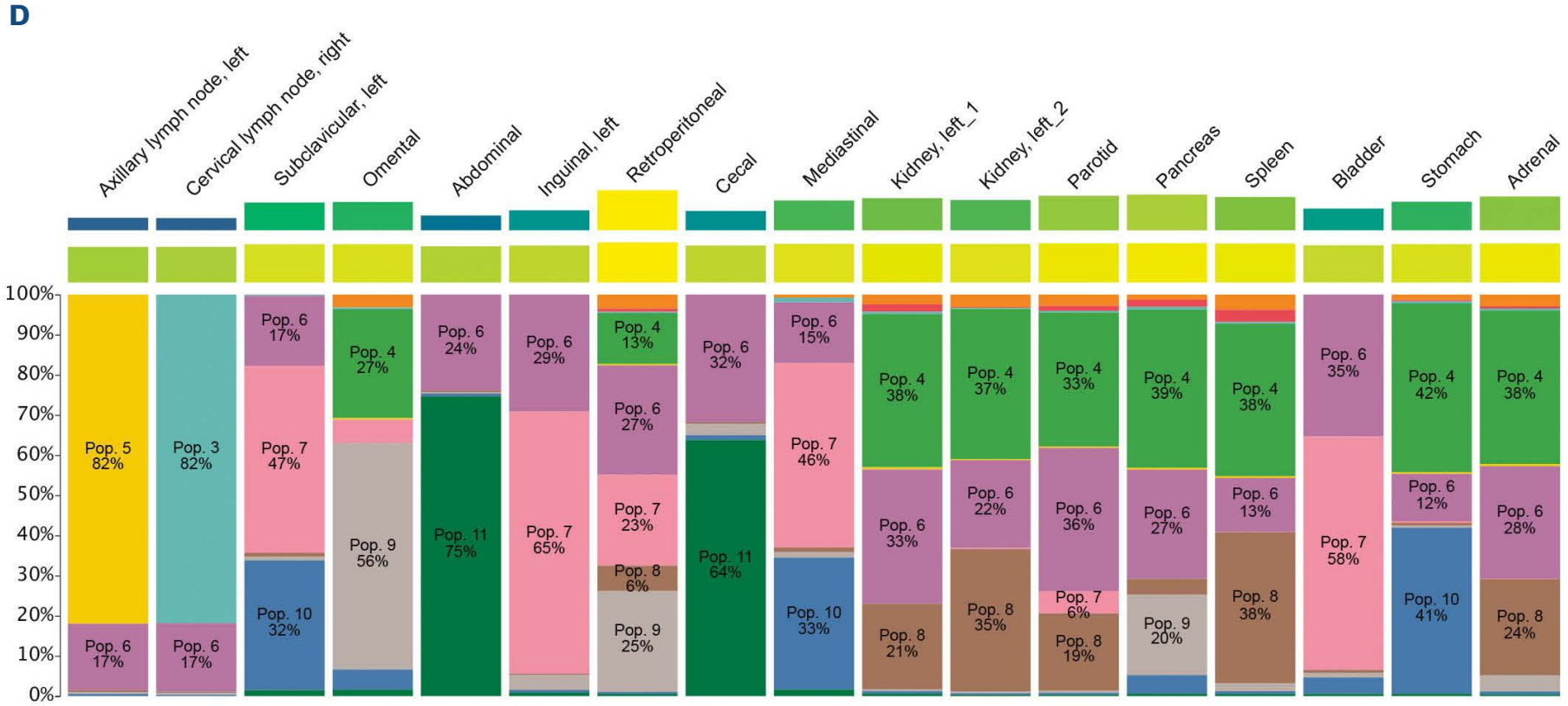


Figure 6. In depth analysis of patient 3 mutations (high-grade B-cell lymphoma). (A) The Venn diagram shows the number of mutations identified in either circulating tumor DNA (ctDNA) whole-genome sequencing (WGS) samples, and their overlap. The barplots show the count of mutations, based on whether they are detected in ctDNA or not (x axis), and based on whether they are detected in WGS of tissue samples or not (coded by color). The analysis of WGS herein is restricted to the specific genomic regions covered by the ctDNA targeted sequencing panel. (B) Clonal phylogeny tree, as inferred by pyclone and pairtree. Driver genes with mutations in specific subclones are labeled. (C) Subclonal population frequencies across samples, with each sample annotated by its clone diversity index (CDI) and clone mutation diversity index (CMDI).

mentary Figure S3B). The phylogenetic tree in this patient suggests that his lymphoma started in either the right cervical or left axillary regions and subsequently spread to sites situated below the diaphragm. The clones that metastasized to these infra-diaphragmatic sites were composed of population 4 or its descendants, and clonal mixtures were universally observed such that some sites were the result of polyphyletic seeding (e.g., left inguinal, bladder) and others arose from monophyletic dissemination (e.g., stomach, kidney). Our results thus illustrate that lymphoma can arise in a specific anatomical site, while metastatic dissemination occurs through the outgrowth of descendant populations, with eventually every single tumor site being multiclonal.

Distribution of driver gene mutations and mutation signatures across clonal lineage trees

In this patient, we observed that the highest number of driver and COSMIC mutations (n=24, 13% of coding mutations) were enriched in the ancestral clone (population 1) with, however, simultaneously several derivative clones (e.g., populations 8, 9, 10 and 11) also harboring up to eight

driver mutations (*Online Supplementary Figure S3C*). We observed several mutation signatures contributing to the dominant founding clone (population 1) including SBS9;SBS85 (somatic hypermutation and AID in lymphoid cells) which was also observed in the founding clones in the other two patients (*Online Supplementary Figure 3D* and *E*). This signature was observed in all other descendant populations in patient 3. Meanwhile several mutation signatures were only observed in specific populations such as SBS86;SBS39;SBS13 (chemotherapy treatment and error-prone polymerases) which was associated with populations 7 and 11 and SBS2 (activity of AID/APOBEC) found in population 7, prevalent only in a subset of patient samples including the abdomen and cecum. Overall, while phylogenetically distinct populations exhibited unique sets of driver gene mutations, mutation signatures contributing to observed mutations were often similar across samples and related to somatic hypermutation and prior chemotherapy exposure.

Discussion

Herein, we reconstructed the evolutionary histories of refractory lymphoma in three patients enrolled into an institutional Research Autopsy Program. We deciphered the underlying clonal phylogenies and their relationships to individual genetic alterations and mutation signatures. There were differences between patients, particularly related to the number and type of genetic alterations, as well as mutation signatures. However, while the size of our patient cohort was small, we observed commonalities across all three patients. We identified complex patterns of clonal spread, with all tumor sites containing multiple cohabitating clones and frequent occurrence of polyphyletic seeding underlying tumor cell dissemination. The three patients considered in this study were on the extreme end of treatment-resistant lymphoma and WGS revealed genetic alterations that are typically associated with poor outcome, such as *TP53* mutations (patient 1), aberrations in DNA repair pathways (patient 2) and double-hit *MYC* and *BCL2* translocations (patient 3). While most disease-driving genetic events were ancestral, subclonal alterations represented potential substrates for positive selection by chemotherapy, including for example *TP53* mutations in independent samples in patient 3. The majority of driver mutations were present at diagnosis in these cases with very poor responses to therapy, while also providing evidence of dynamic evolution, possibly shaped by selective pressures exerted, for example, by therapy. The assessment of mutations in cell-free DNA identified a large proportion of ancestral mutations but also revealed that subclonal mutations can readily be missed when relying exclusively on ctDNA analysis, as opposed to tissue-based sampling.

The main limitation of our study is evidently its small sample size, with only three patients included with different lymphoma diagnoses. It may thus be challenging to generalize upon our findings at present time. Nonetheless, to the best of our knowledge, our description of multi-regional WGS, dissection of clonal dissemination patterns and integration with state-of-the-art ctDNA analysis provides one of the most refined descriptions of ITH in lymphoma to date. Our findings shed light on the evolutionary pathways underlying refractory disease and underscore that the genetic determinants of treatment resistance appear to be largely present at the time of diagnosis in those patients who present with rapid disease progression and/or primary refractory disease. Recently, Makishima *et al.* reported one patient with extranodal NK/T-cell lymphoma in which they performed WES of 12 tumor samples, including two samples from the time of autopsy.[11] Similar to our observation, they found clear evidence of spatial ITH. This observation was made despite the disseminated nature of this patient's lymphoma, with evidence of intravascular invasion. However, most previous research aiming to describe ITH in lymphoma has been limited by scarce sampling, generally considering no more than two serial samples in a given patient.[8,10,24–28] On the other hand, research autopsy programs are challenging to implement but offer access to potentially unrivaled tissue resources that can expand our understanding of tumor evolution.[12,29] Notwithstanding the ethical and logistical challenges associated with research autopsy programs, systematic analysis of clonal trajectories in larger patient cohorts remains a compelling strategy to reveal generalized disease trends that could aid in improving clinical decisions.

Disclosures

No conflicts of interest to disclose.

Contributions

KI analyzed and interpreted WGS data, and wrote the manuscript. TL analyzed ctDNA data. MB and KT processed samples. RG performed pathology review. BL, IL and PML generated ctDNA data. AO, ND, AP and MC established the Rapid Autopsy Program and/or recruited patients. RK designed the study, interpreted data, wrote the manuscript and provided general oversight of the project.

Acknowledgments

The authors wish to thank The Center for Applied Genomics, The Hospital for Sick Children, Toronto, Canada for assistance with WGS.

Funding

This research was supported by a New Idea Award from the

Leukemia and Lymphoma Society (8013-19) and a Large-Scale Applied Research Project funded by Genome Canada and the Ontario Research Fund (271LYM). The Research Autopsy Program was generously funded and supported by the Princess Margaret Cancer Foundation.

Data-sharing statement

WGS data for bulk tumor samples and ctDNA data have been deposited to the European Genome-Phenome Archive (EGAS00001005996). Qualified investigators will be able to access these data through a data-sharing agreement.

References

1. Greaves M. Evolutionary determinants of cancer. Cancer Discov. 2015;5(8):806-820.
2. McGranahan N, Swanton C. Clonal heterogeneity and tumor evolution: past, Present, and the future. Cell. 2017;168(4):613-628.
3. Dentro SC, Leshchiner I, Haase K, et al. Characterizing genetic intra-tumor heterogeneity across 2,658 human cancer genomes. Cell. 2021;184(8):2239-2254.
4. 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.
5. Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med. 2012;366(10):883-892.
6. McPherson A, Roth A, Laks E, et al. Divergent modes of clonal spread and intraperi-toneal mixing in high-grade serous ovarian cancer. Nat Genet. 2016;48(7):758-767.
7. Seifert M, Scholtysik R, Küppers R. Origin and pathogenesis of B cell lymphomas. Methods Mol Biol. 2019;1956:1-33.
8. Araf S, Wang J, Korfi K, et al. Genomic profiling reveals spatial intra-tumor heterogeneity in follicular lymphoma. Leukemia. 2018;32(5):1261-1265.
9. Haebe S, Shree T, Sathe A, et al. Single-cell analysis can define distinct evolution of tumor sites in follicular lymphoma. Blood. 2021;137(21):2869-2880.
10. Magnes T, Wagner S, Thorner AR, et al. Spatial heterogeneity in large resected diffuse large B-cell lymphoma bulks analysed by massively parallel sequencing of multiple synchronous biopsies. Cancers. 2021;13(4):650.
11. Makishima K, Suehara Y, Abe Y, et al. Intratumor heterogeneity of lymphoma identified by multiregion sequencing of autopsy samples. Cancer Sci. 2022;113(1):362-364.
12. Bavi P, Siva M, Abi-Saab T, et al. Developing a pan-cancer research autopsy programme. J Clin Pathol. 2019;72(10):689-695.
13. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv. 2013 March. https://doi.org/10.48550/arXiv.1303.3997 [preprint, not peer-reviewed]
14. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43(5):491-498.
15. Benjamin D, Sato T, Cibulskis K, Getz G, Stewart C, Lichtenstein L. Calling somatic SNVs and indels with Mutect2. bioRxiv. 2019;861054. [preprint, not peer-reviewed]
16. Website. https://gatk.broadinstitute.org/hc/en-us/articles/360035889791?id=11136.
17. Kim S, Scheffler K, Halpern AL, et al. Strelka2: fast and accurate calling of germline and somatic variants. Nat Methods.2018;15(8):591-594.
18. Boyle-Lab. Blacklist/hg19-bla cklist.v2.bed.gz at master Boyle-Lab/Blacklist. https://github.com/Boyle-Lab/Blacklist (accessed November 27, 2021).
19. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38(16):e164.
20. Landau DA, Tausch E, Taylor-Weiner AN, et al. Mutations driving CLL and their evolution in progression and relapse. Nature. 2015;526(7574):525-530.
21. Klener P, Fronkova E, Berkova A, et al. Mantle cell lymphoma-variant Richter syndrome: Detailed molecular-cytogenetic and backtracking analysis reveals slow evolution of a pre-MCL clone in parallel with CLL over several years. Int J Cancer. 2016;139(10):2252-2260.
22. Reiter JG, Makohon-Moore AP, Gerold JM, et al. Reconstructing metastatic seeding patterns of human cancers. Nat Commun. 2017;8:14114.
23. Gillis S, Roth A. PyClone-VI: scalable inference of clonal population structures using whole genome data. BMC Bioinformatics. 2020;21(1):571.
24. Isaev K, Ennishi D, Hilton L, et al. Molecular attributes underlying central nervous system and systemic relapse in diffuse large B-cell lymphoma. Haematologica. 2021;106(5):1466-1471.
25. Magnes T, Wagner S, Thorner AR, et al. Clonal evolution in diffuse large B-cell lymphoma with central nervous system recurrence. ESMO Open. 2021;6(1):100012.
26. Greenawalt DM, Liang WS, Saif S, et al. Comparative analysis of primary relapse/refractory DLBCL identifies shifts in mutation spectrum. Oncotarget. 2017;8(59):99237-99244.
27. Juskevicius D, Lorber T, Gsponer J, et al. Distinct genetic evolution patterns of relapsing diffuse large B-cell lymphoma revealed by genome-wide copy number aberration and targeted sequencing analysis. Leukemia. 2016;30(12):2385-2395.
28. Kridel R, Chan FC, Mottok A, et al. Histological transformation and progression in follicular lymphoma: a clonal evolution study. PLoS Med. 2016;13(12):e1002197.
29. Krook MA, Chen H-Z, Bonneville R, Allenby P, Roychowdhury S. Rapid research autopsy: piecing the puzzle of tumor heterogeneity. Trends Cancer Res. 2019;5(1):1-5.