From www.bloodjournal.org by guest on July 31, 2017. For personal use only.

Regular Article

blood

LYMPHOID NEOPLASIA

Diffuse large B-cell lymphoma genotyping on the liquid biopsy

Davide Rossi,[1,2,*] Fary Diop,[1,*] Elisa Spaccarotella,[1] Sara Monti,[1] Manuela Zanni,[1] Silvia Rasi,[1] Clara Deambrogi,[1] Valeria Spina,[1,2] Alessio Bruscaggin,[1,2] Chiara Favini,[1] Roberto Serra,[3] Antonio Ramponi,[4] Renzo Boldorini,[4] Robin Foà,[5] and Gianluca Gaidano[1]

[1]Division of Hematology, Department of Translational Medicine, University of Eastern Piedmont, Novara, Italy; [2]Division of Hematology, Institute of Oncology Research, Oncology Institute of Southern Switzerland, Bellinzona, Switzerland; [3]Laboratory of Informatics, University of Eastern Piedmont, Novara, Italy; [4]Division of Pathology, Department of Health Sciences, Maggiore della Carita' Hospital, University of Eastern Piedmont, Novara, Italy; and [5]Division of Hematology, Department of Cellular Biotechnologies and Hematology, Sapienza University, Rome, Italy

Key Points

- Plasma cfDNA genotyping is as accurate as genotyping of the diagnostic biopsy in detecting clonal somatic mutations in DLBCL.
- Plasma cfDNA genotyping is a real-time, noninvasive tool that can be used to track clonal evolution in DLBCL.

Accessible and real-time genotyping for diagnostic, prognostic, or treatment purposes is increasingly impelling in diffuse large B-cell lymphoma (DLBCL). Cell-free DNA (cfDNA) is shed into the blood by tumor cells undergoing apoptosis and can be used as source of tumor DNA for the identification of DLBCL mutations, clonal evolution, and genetic mechanisms of resistance. In this study, we aimed at tracking the basal DLBCL genetic profile and its modification upon treatment using plasma cfDNA. Ultra-deep targeted next generation sequencing of pretreatment plasma cfDNA from DLBCL patients correctly discovered DLBCL-associated mutations that were represented in >20% of the alleles of the tumor biopsy with >90% sensitivity and ~100% specificity. Plasma cfDNA genotyping also allowed for the recovery of mutations that were undetectable in the tissue biopsy, conceivably because, due to spatial tumor heterogeneity, they were restricted to clones that were anatomically distant from the biopsy site. Longitudinal analysis of plasma samples collected under rituximab-cyclophosphamide-doxorubicin-vincristine-prednisone (R-CHOP) chemotherapy showed a rapid clearance of DLBCL mutations from cfDNA among responding patients. Conversely, among patients who were resistant to R-CHOP, basal DLBCL mutations did not disappear from cfDNA. In addition, among treatment-resistant patients, new mutations were acquired in cfDNA that marked resistant clones selected during the clonal evolution. These results demonstrate that cfDNA genotyping of DLBCL is as accurate as genotyping of the diagnostic biopsy to detect clonally represented somatic tumor mutations and is a real-time and noninvasive approach to tracking clonal evolution and the emergence of treatment-resistant clones. (*Blood*. 2017;129(14):1947-1957)

Introduction

Diffuse large B-cell lymphoma (DLBCL) is a molecularly heterogeneous disease whose genetics may have clinical implications for patient stratification and treatment.[1,2] Mutations of clinical importance in DLBCL affect the *TP53* gene, whose variants are consistently associated with poor prognosis among patients treated with rituximab-cyclophosphamide-doxorubicin-vincristine-prednisone (R-CHOP),[3] and the *CARD11*, *CD79A*, *CD79B*, and *MYD88* genes, whose variants predict the benefit or no benefit from ibrutinib monotherapy.[4]

Limitations in accessing fresh tumor material from DLBCL tissue biopsies has prevented the rapid translation of DLBCL gene mutations into prognostic or predictive tools for clinical practice. Also, serial sampling of tumors to track the acquisition of drug-resistance mutations requires a rebiopsy of the DLBCL, which may not be routinely feasible in clinical practice. Therefore, alternative accessible sources of tumor DNA may help to complement the molecular diagnostic analyses that are routinely carried out on formalin-fixed paraffin-embedded (FFPE) tissue biopsies.

Cell-free fragments of DNA (cfDNA) are shed into the bloodstream by tumor cells undergoing apoptosis.[5] Accessing tumor cfDNA through the bloodstream has clear sampling advantages and allows serial monitoring of disease genetics in real time. cfDNA is also representative of the entire tumor heterogeneity, thus enabling to bypass the anatomical biases imposed by tissue biopsies in the reconstruction of the entire cancer clonal architecture and to identify resistant clones that are dormant in nonaccessible tumor sites.[5]

In DLBCL, cfDNA has been quantified or used to track the tumor clonotypic immunoglobulin gene rearrangement for minimal residual disease monitoring.[6-9] Conversely, the evidence that cfDNA mirrors the underlying tumor genetics in DLBCL and the proof that cfDNA can be used to track in real time clonal evolution-driven resistance upon treatment are currently limited to retrospective series.[10,11] In this prospective, training validation study, we show that cfDNA genotyping of DLBCL: (1) is as accurate as genotyping of the diagnostic biopsy to detect somatic mutations of allelic abundance >20%; and (2) is a noninvasive tool to track the emergence of treatment-resistant clones.

Submitted 31 May 2016; accepted 11 January 2017. Prepublished online as *Blood* First Edition paper, 17 January 2017; DOI 10.1182/blood-2016-05-719641.

*D.R. and F.D. contributed equally to this study.

The online version of this article contains a data supplement.

There is an Inside *Blood* Commentary on this article in this issue.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

© 2017 by The American Society of Hematology

From www.bloodjournal.org by guest on July 31, 2017. For personal use only.

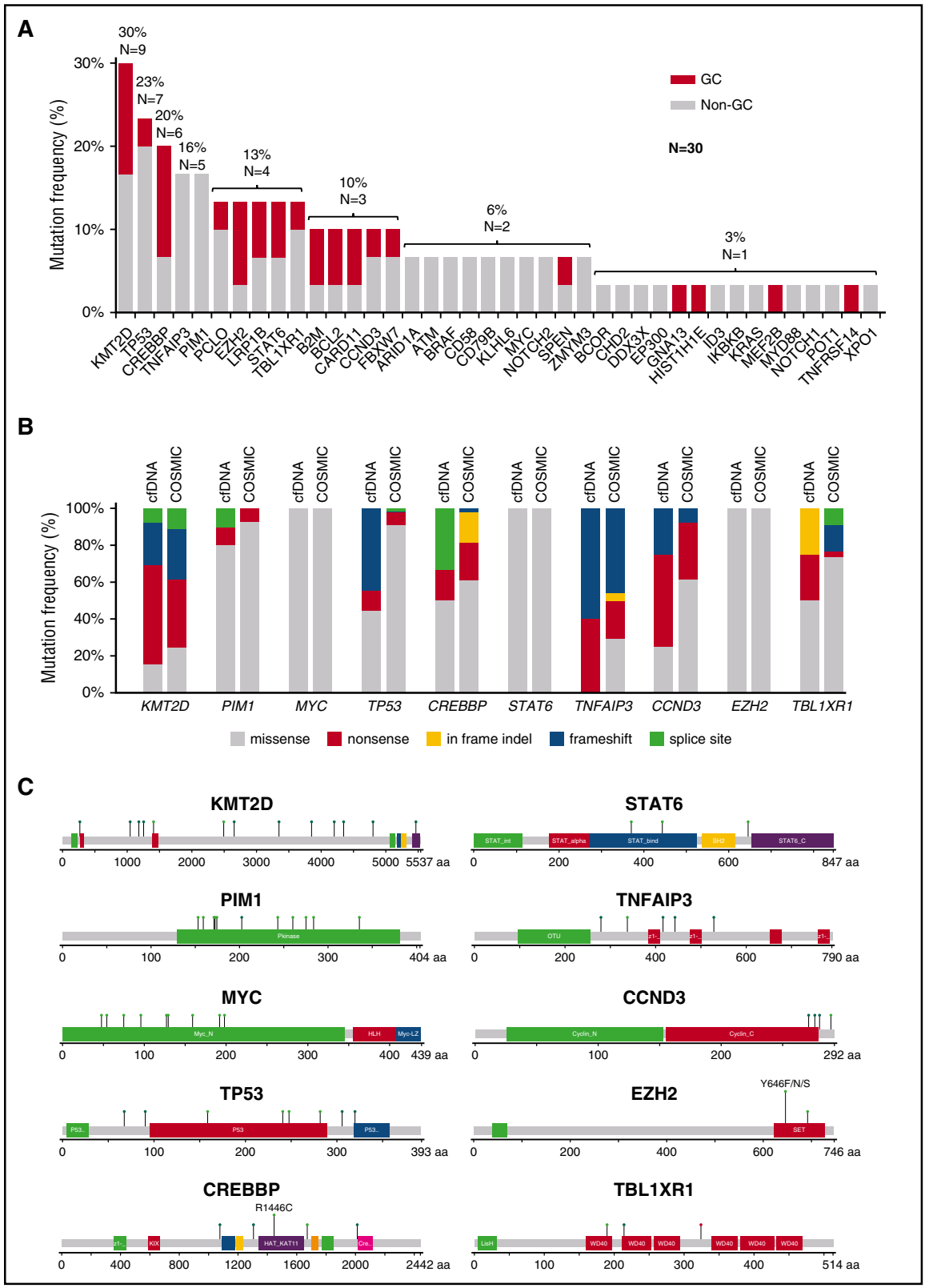

Figure 1. Prevalence and molecular spectrum of nonsynonymous somatic mutations discovered in plasma cfDNA. The percentage of training (A) and validation (D) DLBCL cases harboring nonsynonymous somatic mutations of genes investigated by targeted resequencing of plasma cfDNA. Color codes indicate the distribution of mutations among GC and non-GC DLBCL. The final number and prevalence of mutated cases is given for each gene. The molecular spectrum of nonsynonymous somatic mutations identified in plasma cfDNA of the training (B) and validation (E) DLBCL cases compared with the molecular spectrum of nonsynonymous somatic mutations that have been detected in the tumor gDNA of published DLBCL series and reported in the COSMIC database (version 76). Genes mutated in >10% cases of the DLBCL series are reported. The position and type of nonsynonymous somatic mutations that were identified in plasma cfDNA of the training (C) and validation (F) DLBCL cases. Genes mutated in >10% cases are reported. Mutation maps were obtained through MutationMapper version 1.0.1.

From www.bloodjournal.org by guest on July 31, 2017. For personal use only.

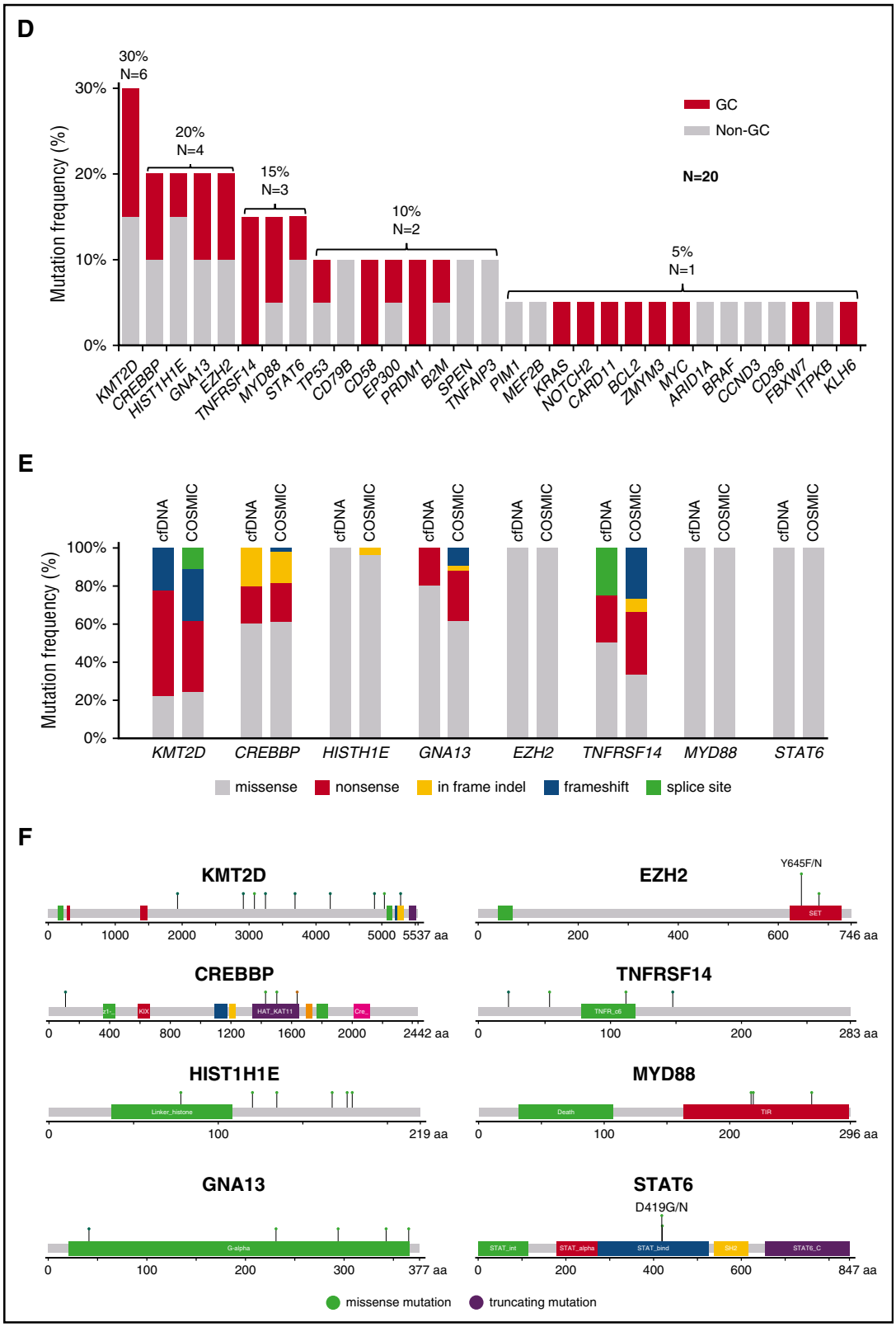

Figure 1. (Continued).

From www.bloodjournal.org by guest on July 31, 2017. For personal use only.

Methods

Patients

The study had a prospective, observational, nonintervention, unicentered design and consisted in the longitudinal collection of peripheral blood (PB) samples and clinical data from DLBCL patients treated with R-CHOP at the University of Eastern Piedmont. Inclusion criteria were: (1) male or female adults ≥18 years; (2) diagnosis of untreated DLBCL after pathological revision; (3) treatment with R-CHOP; and (4) evidence of signed informed consent. A total of 50 previously untreated DLBCL patients fulfilled the inclusion criteria and were recruited for the study from November 2013 to August 2015 as training series (n = 30), and from September 2015 to October 2016 as validation cohort (n = 20) (supplemental Table 1, available on the *Blood* Web site). All patients received R-CHOP. Recruited patients did not harbor either morphological/immunophenotypic evidence of PB involvement by DLBCL cells or bone marrow involvement by DLBCL, as documented by both the bone marrow biopsy and positron emission tomography/computed tomography (PET/CT).[12] The diagnosis and immunohistochemically defined cell of origin[13] were revised by 2 independent pathologists (R.B. and A.R.). The following biological material was collected: (1) cfDNA isolated from plasma at diagnosis before the start of treatment, during R-CHOP courses (on day 1 of each course before treatment infusion), at the end of treatment and at progression; and (2) normal germ line genomic DNA (gDNA) extracted from PB granulocytes after Ficoll separation. For comparative purposes, tumor gDNA from the paired DLBCL diagnostic tissue biopsy was available for 36 patients (extracted from fresh specimens in 25 cases and from FFPE specimens in 11 cases). In the remaining 14 cases, the leftover tissue from the FFPE diagnostic biopsy was not available or gave an insufficient amount of DNA for cancer personalized profiling by deep sequencing (CAPP-seq). In all instances, basal plasma samples were collected in close temporal proximity of the tumor tissue biopsy (7-14 days after diagnostic tissue biopsy) and before starting treatment. Clinical information was prospectively maintained in the University of Eastern Piedmont lymphoma database. Disease response was assessed by PET/CT.[13] Patients provided informed consent in accordance with local institutional review board requirements and the Declaration of Helsinki. Paired plasma cfDNA and normal gDNA from granulocytes collected from 6 healthy donors were used to set the experimental and biological background of the ultra-deep next generation sequencing (NGS) approach. The study was approved by the Ethical Committee of the Ospedale Maggiore della Carità di Novara, which is affiliated with the University of Eastern Piedmont (protocol code CE 112/15).

DNA extraction

In order to avoid preanalytic confounding effects, a standardized approach was used to extract cfDNA from plasma (supplemental Appendix).[14] PB granulocytes were separated by Ficoll gradient density centrifugation as a source of normal germ line gDNA. Tumor gDNA was isolated from fresh or FFPE diagnostic tissue biopsies containing >70% of tumor cells as estimated by morphology and immunohistochemistry. gDNA was extracted according to standard procedures.

Library design for hybrid selection

A targeted resequencing gene panel, including coding exons and splice sites of 59 genes (target region: 207299bp) that are recurrently mutated in DLBCL and other mature B-cell tumors, was specifically designed for this project (supplemental Appendix; supplemental Table 2). The gene panel allowed a priori the recovery of at least 1 clonal mutation in 92.6% (95% confidence interval [CI]: 83.6-97.5) of DLBCL patients, as documented by in silico validation against public genomic datasets of DLBCL.[15-19]

CAPP-seq library preparation and ultra-deep NGS

In the training cohort, the gene panel was analyzed in plasma cfDNA collected at diagnosis, during R-CHOP courses, at the end of treatment, at progression (total cfDNA samples = 127), and, for comparative purposes to filter out polymorphisms, in germ line gDNA from the paired granulocytes. The tumor gDNA from the paired tissue biopsy was also investigated to test the accuracy of cfDNA genotyping. In the validation cohort, the gene panel was investigated in plasma cfDNA collected at diagnosis and in normal and tumor gDNA from paired granulocytes and the tissue biopsy, respectively. Tumor and germ line gDNA from tissues (median = 318 ng) were sheared through sonication before library construction to obtain 200-bp fragments. For plasma cfDNA, which is naturally fragmented, 2 to 717 ng (median = 17 ng) of DNA were used for library construction without additional fragmentation. The NGS libraries were constructed using the KAPA Library Preparation Kit (Kapa Biosystems), and hybrid selection was performed with the custom SeqCap EZ Choice Library (Roche NimbleGen). In the CAPP-seq of cfDNA, the manufacturer's protocols were modified as previously reported.[20] Multiplexed libraries (n = 6 per run) were sequenced using 300-bp paired-end runs on a MiSeq sequencer (Illumina).

Bioinformatic pipeline for variant calling

Details of the bioinformatic methods are supplied in the supplemental Methods. Nonsynonymous somatic mutation calling in plasma cfDNA was performed separately and was blinded from mutation calling in tumor gDNA. After CAPP-seq, FASTQ sequencing reads were initially deduped using FastUniq version 1.1. Then, the deduped FASTQ sequencing reads were locally aligned to the hg19 version of the human genome assembly using BWA version 0.6.2 and assembled into a mpileup file using SAMtools version 1. Single nucleotide variations and indels were called in plasma cfDNA vs germ line gDNA and tumor gDNA vs germ line gDNA, respectively, with the somatic function of VarScan2. Variants annotated as single nucleotide polymorphisms (SNPs) according to dbSNP 138 (with the exception of *TP53* variants that were manually curated and scored as SNPs according to the International Agency for Research on Cancer TP53 database), intronic variants mapping >2 bp before the start or after the end of coding exons, and synonymous variants were filtered out. Two independent statistical approaches (Fisher's exact test and *Z*-test) were then used to filter out variants below the base-pair resolution background frequencies in cfDNA across the target region. Only variants that had a significant call in both tests were retained (Bonferroni-adjusted $P < 6 \times 10^{-8}$) (supplemental Appendix).

Sanger sequencing

Sanger sequencing was also used to validate the most abundant plasma cfDNA mutations detected by CAPP-seq.

Statistical analysis

The sensitivity and specificity of plasma cfDNA genotyping were calculated in comparison with tumor gDNA genotyping as the gold standard. The analysis was performed with SPSS Statistics version 22.0

Results

Characteristics of the study cohorts

The study was based on a prospectively collected, consecutive series of 30 newly diagnosed DLBCL patients (training cohort) whose characteristics were consistent with an unselected cohort of DLBCL (supplemental Table 1). Upon R-CHOP treatment, 83.3% (95% CI: 65.9-93.1; n = 25/30) of patients achieved PET/CT-negative complete remission, whereas 16.7% (95% CI: 6.8-34.0; n = 5/30) failed to achieve complete remission. Of the patients that achieved complete remission (median follow-up: 6 months), 2 relapsed. The median number of cfDNA molecules per milliliter of plasma at disease presentation was 771.7 (range: 137.2-18 742.5). An independent validation series of 20 consecutive DLBCL patients was also assessed to confirm the accuracy of plasma cfDNA genotyping (supplemental Table 1).

From www.bloodjournal.org by guest on July 31, 2017. For personal use only.

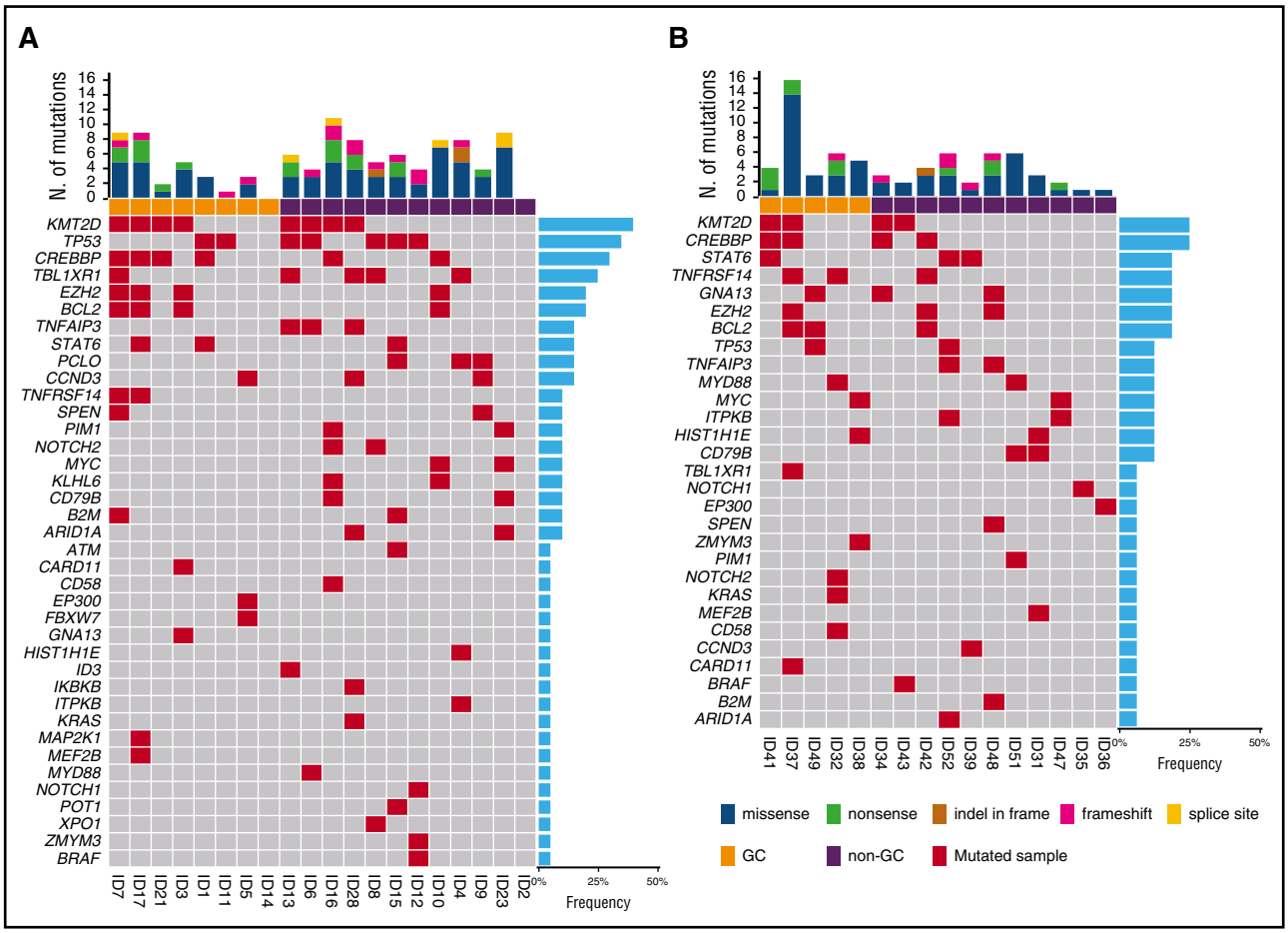

Figure 2. Overview of the DLBCL genotype discovered in tumor gDNA. Case-level mutational profiles of 20 training (A) and 16 validation (B) DLBCL tumors genotyped on gDNA from the tissue biopsy. Each column represents 1 tumor sample and each row represents 1 gene. The fraction of tumors with mutations in each gene is plotted (right). The number of mutations in a given tumor is plotted above the heat map. Cases are clustered according to the cell of origin.

Plasma cfDNA genotyping discloses somatic mutations in DLBCL-associated genes

To provide the proof of principle that plasma could function as a liquid biopsy for tracking recurrently mutated genes in DLBCL, plasma cfDNA collected at presentation from the training cohort was genotyped by using CAPP-seq, a targeted ultra-deep NGS approach for plasma cfDNA genotyping already validated in solid tumors (≥80% of the target region covered >1000× in all DLBCL; ≥80% of the target region covered >2000× in 17/30 DLBCL; supplemental Figure 1, supplemental Table 3).[20] Paired normal gDNA was also analyzed to confirm the somatic origin of mutations. Overall, within the interrogated genes, 66.6% (95% CI: 47.1-82.7%; n = 20/30) of patients harbored somatic mutations (total number: 129; range: 2-13 mutations per patient) that were detectable in plasma cfDNA (supplemental Table 4). In order to validate the NGS results, cfDNA was extracted from a second aliquot of plasma and then subjected to a second CAPP-seq and ultra-deep NGS sequencing. The robustness of the plasma cfDNA CAPP-seq approach and of the bioinformatic analysis was documented by the high concordance (R^2 = 0.916) of variant calling from the independent duplicate experiments (supplemental Figure 2), which consistently confirmed all the variants initially discovered in the plasma cfDNA, including those of low allele frequency, thus excluding their origin from a batch-specific experimental noise. Sanger sequencing consistently detected all plasma cfDNA mutations whose allele frequency was within the sensitivity range of this approach (allele frequency >10%) (supplemental Figure 2), thus validating the CAPP-seq results on a different experimental platform.

Consistent with the typical spectrum of mutated genes in DLBCL, plasma cfDNA genotyping revealed somatic variants of *KMT2D* (also known as *MLL2*) in 30.0% (95% CI: 14.7-49.4; n = 9/30); *TP53* in 23.3% (95% CI: 9.9-42.2; n = 7/30); *CREBBP* in 20.0% (95% CI: 7.7-38.5; n = 6/30); *PIM1* and *TNFAIP3* in 16.6% (95% CI: 5.6-34.7; n = 5/30); *EZH2*, *STAT6*, and *TBL1XR1* in 13.3% (95% CI: 3.7-30.7; n = 4/30); *B2M*, *BCL2*, *CARD11*, *CCND3*, and *FBXW7* in 10.0% (95% CI: 2.1-26.5; n = 3/30); *CD58*, *CD79B*, and *MYC* in 6.6% (95% CI: 0.8-22.0; n = 2/30); and *EP300*, *GNA13*, *MEF2B*, *MYD88*, and *TNFRSF14* in 3.3% (95% CI: 0-17.2; n = 1/30) of cases (Figure 1A). Although the sample size was not powered to show significant enrichments of mutated genes within cell of origin categories of DLBCL, *EZH2* and *BCL2* mutations were, as expected, more frequent in germinal center (GC) DLBCL, whereas *TNFAIP3* and *PIM1* mutations were more frequent in non-GC DLBCL (Figure 1A).[1]

Among recurrently affected genes, the molecular spectrum of mutations identified in plasma cfDNA was highly consistent with that of variants that have been detected in tumor gDNA of published DLBCL series and reported in the COSMIC database (Figure 1B-C). *PIM1* and *MYC* were affected by multiple hotspot mutations that were suggestive of AID-related events (Figure 1C; supplemental Table 4).[21] Notably, patient identification code (ID) 23, who had the t(8;14) *MYC*

From www.bloodjournal.org by guest on July 31, 2017. For personal use only.

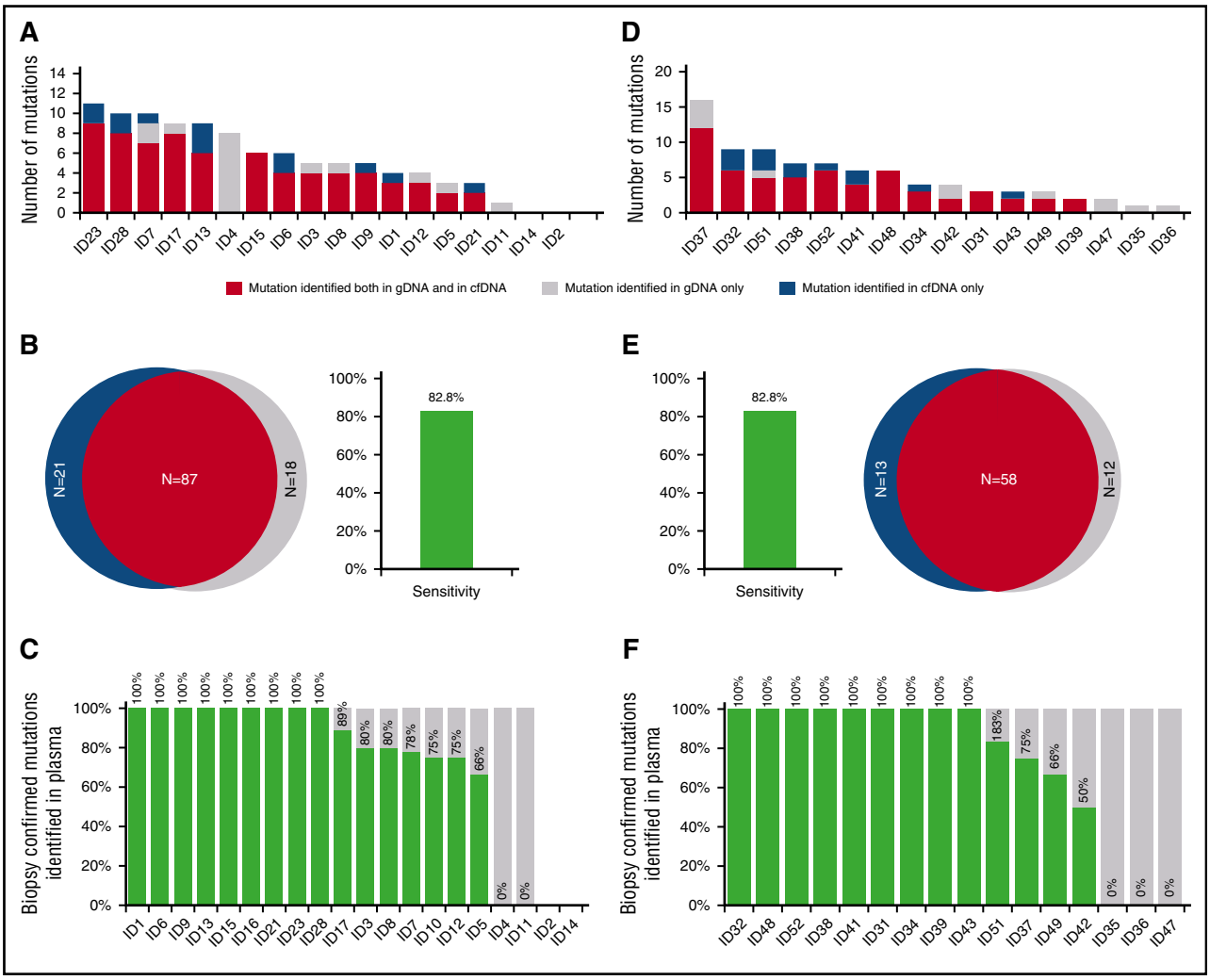

Figure 3. Concordance between plasma cfDNA and tumor gDNA genotyping. The number of mutations in a given tumor discovered in plasma cfDNA and/or tumor gDNA in the training (A) and validation (D) DLBCL series (cases provided with paired plasma cfDNA and tumor biopsy gDNA are represented). Mutations are color coded if they were identified in both plasma cfDNA and tumor biopsy gDNA (red), only in plasma cfDNA (blue), or only in tumor biopsy gDNA (gray). A Venn diagram summarizes the overall number of mutations discovered in both plasma cfDNA and tumor biopsy gDNA (red), only in plasma cfDNA (blue), or only in tumor biopsy gDNA (gray) from the training (B) and validation (E) DLBCL series. The corresponding overall sensitivity of plasma cfDNA genotyping in discovering biopsy-confirmed mutations is shown. The fraction of tumor biopsy–confirmed mutations that were detected in plasma cfDNA in the training (C) and validation (F) DLBCL series is shown. Patients are ordered by decreasing detection rates. The training cohort included 2 patients (ID2 and ID14) who were not informative because they were devoid of mutations within the target region. The green portion of the bars indicates the prevalence of tumor biopsy–confirmed mutations that were detected in plasma cfDNA. The gray portion of the bars indicates the prevalence of tumor biopsy–confirmed mutations that were not detected in plasma cfDNA.

translocation, also harbored multiple *MYC* mutations in cfDNA, consistent with the known accumulation of mutations in the translocated *MYC* gene due to the *IGH* enhancer-driven misfire of somatic hypermutation (supplemental Table 4).[22]

Analysis of 6 healthy donors by CAPP-seq did not disclose any somatic mutations in plasma cfDNA (supplemental Table 4), suggesting that the ultra-deep NGS and variant calling approaches used in this study did not pick up biological or analytical background noise in cfDNA.

Validation of biopsy-free DLBCL genotyping

The fresh or FFPE tissue biopsy of 20 DLBCL patients from the training cohort was genotyped by CAPP-seq, which was blinded from the mutational profile recovered in the paired plasma cfDNA. To validate our sequencing approach of FFPE samples, gDNA from paired frozen/FFPE specimens processed from the same DLBCL biopsy (n = 4) was subjected to CAPP-seq. Pairwise analysis of data showed high (96%) concordance in variant recovery from FFPE vs fresh samples (supplemental Figure 5).

In order to systematically derive the accuracy of cfDNA genotyping, the results of plasma cfDNA genotyping and tumor gDNA genotyping (gold standard; mutation spectrum shown in Figure 2A) were compared (Figures 3A-C) in 18 DLBCL training cases that were provided with the paired plasma/tissue samples and were mutated within the target region. Sequencing coverage of the tumor gDNA was comparable to that of plasma cfDNA (supplemental Figure 1).

Genotyping of plasma cfDNA collected at diagnosis identified a total of 108 somatic mutations, whereas genotyping of the gDNA from the diagnostic tissue biopsy identified 105 somatic mutations. Biopsy-confirmed tumor mutations were detectable with 82.8% (n = 87/105; 95% CI: 74.4-88.9) sensitivity in pretreatment plasma cfDNA samples (Figure 3B). Biopsy-confirmed tumor mutations not discovered in the cfDNA (n = 18/105) generally had low representation in the diagnostic biopsy tissue (median allele frequency in the tumor biopsy = 5.7%)

From www.bloodjournal.org by guest on July 31, 2017. For personal use only.

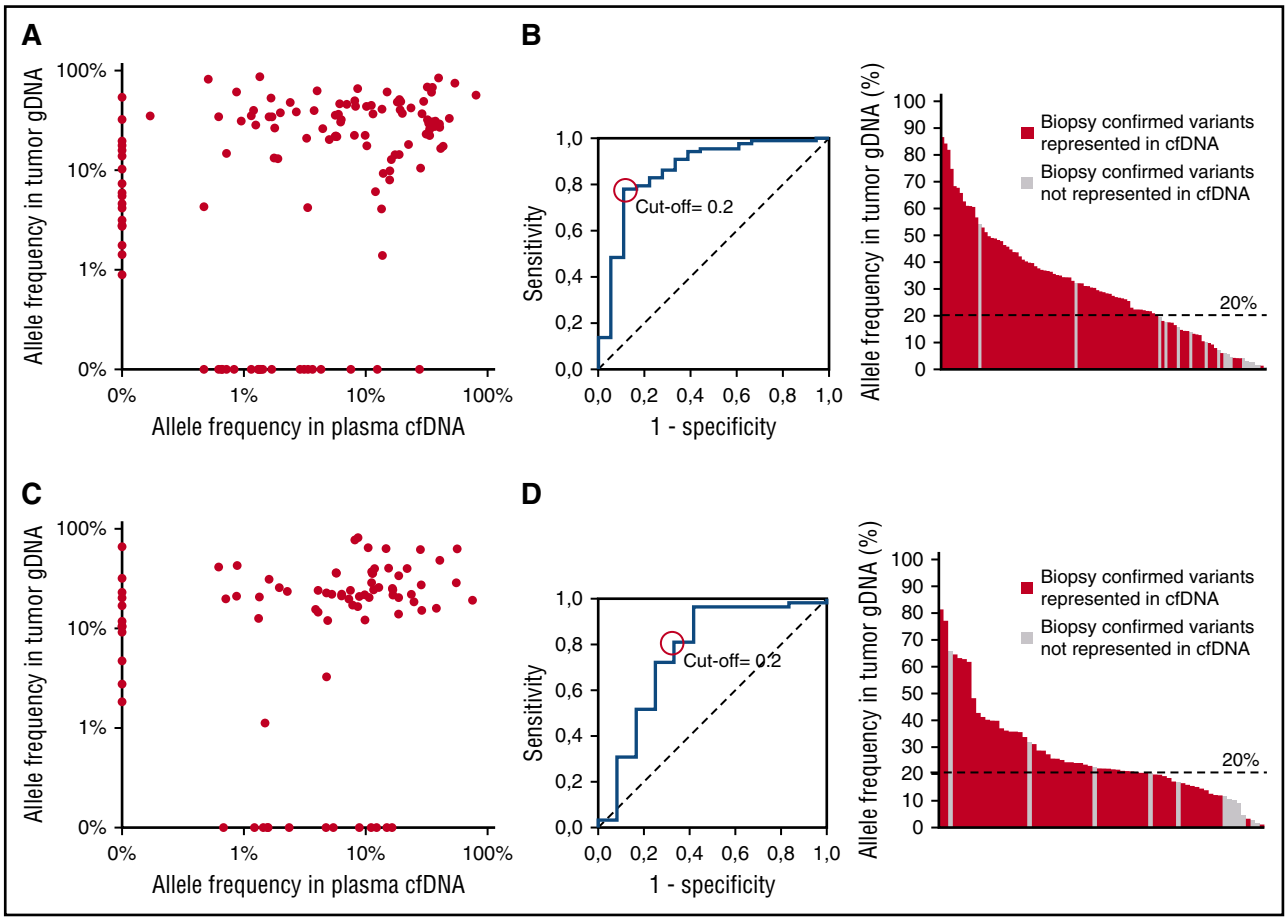

Figure 4. Accuracy of plasma cfDNA genotyping vs tumor gDNA genotyping. The mutation abundance in plasma cfDNA vs the mutation abundance in tumor gDNA is comparatively represented in the scatter plot for each variant identified in the training (A) and validation (C) DLBCL cases. ROC analysis illustrates the performance of plasma cfDNA genotyping in detecting biopsy-confirmed tumor variants according to the variant allele frequency of mutations in tumor gDNA in the training (B) and validation (D) DLBCL cases. The bar graph shows the allele frequency in tumor gDNA of the variants that were discovered in plasma cfDNA (red bars) or missed in plasma cfDNA (gray bars). The dash line tracks the 20% variant allele frequency threshold.

(Figure 4A-B). Consistently, using receiver operating characteristic (ROC) analysis, cfDNA genotyping showed the highest sensitivity (97.1%; 95% CI: 89.5-99.8; n = 68/70) in discovering mutations that were represented in ≥20% of the alleles of the tumor biopsy (Figure 4B), thus demonstrating that plasma cfDNA can accurately mirror the profiles of the most abundant clones found in tumor tissues. Plasma cfDNA was uninformative on tumor biopsy mutational status in 10% (95% CI: 1.5-31.3; n = 2/20) of DLBCL cases (Figure 3C). The representation of the variants in plasma correlated with the LDH levels at DLBCL diagnosis, but not with disease stage (supplemental Figure S6).

Plasma cfDNA genotyping disclosed additional 21 somatic mutations that were not detectable in the biopsy tissue, including 7 variants affecting *PIM1*, a known target of somatic hypermutation (Figures 2A and 3; supplemental Table 4).[21] Repeated ultra-deep NGS consistently confirmed these variants, thus excluding their origin from batch-specific experimental noise (supplemental Figure 2). Because the target region that has been sequenced in the compiled 20 DLBCL patients might support a total of 18 011 440 potential nonsynonymous variants if they were randomly distributed, and considering that only 21 nonsynonymous mutations were detected in plasma cfDNA but not in the tumor biopsy, the false positive rate of plasma cfDNA genotyping was 1.16×10^{-6} corresponding to a specificity >99.99% compared with tumor gDNA genotyping.

As previously reported in other cancer types,[23] nonsynonymous mutations occurring only in plasma cfDNA conceivably represented tumor mutations restricted to clones that were anatomically distant from the biopsy site, rather than false positive calls. Although the study lacked a systematic multiregional sequencing of tumor samples from multiple anatomical sites, the observations in support of the above interpretation are that: (1) the plasma cfDNA *FBXW7* p.S668G mutation of patient ID9, although undetectable in the diagnostic biopsy tissue, was identified in tumor cells from the cytospin of the cerebrospinal fluid collected at the time of isolated meningeal relapse (Figure 5); and (2) all mutations lacking in the diagnostic biopsy tissue disappeared from cfDNA upon achieving complete remission of DLBCL.

Plasma cfDNA was also genotyped by CAPP-seq in an independent cohort of 20 consecutive DLBCL patients. Overall, within the interrogated genes, 85.0% (95% CI: 63.1-95.6; n = 17/20) of validation patients harbored somatic mutations (total: 83 mutations; range: 1-12 mutations per patient) that were detectable in plasma cfDNA (Figure 1D; supplemental Table 4). The molecular spectrum of mutations identified in plasma cfDNA of patients in the validation series was highly consistent with that observed in the cfDNA of patients in the training series, and more in general with that of published DLBCL cohorts (Figure 1E-F). CAPP-seq of paired tumor (Figure 2B) and plasma samples collected from 16 DLBCL patients of the validation series confirmed the sensitivity of plasma cfDNA genotyping in recovering biopsy-confirmed mutations (Figure 3D-E). Biopsy-confirmed tumor mutations were detectable with 82.8% (n = 58/70,

From www.bloodjournal.org by guest on July 31, 2017. For personal use only.

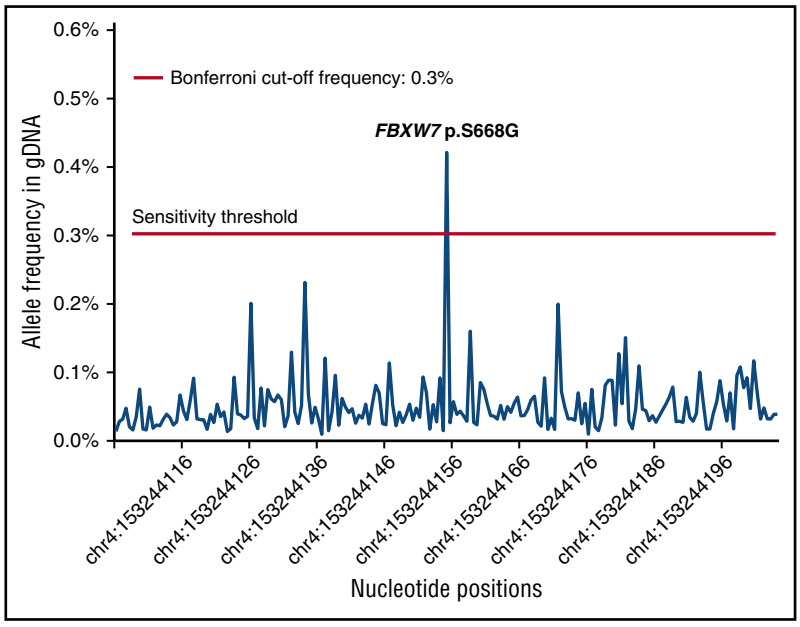

Figure 5. Detection of a *FBXW7* mutation in genomic tumor DNA from DLBCL cells of the cerebrospinal fluid. The p.S688G *FBXW7* somatic mutation was detected in the tumor gDNA from DLBCL cells of the cerebrospinal fluid at 0.42% frequency above the background, passing the identification cutoffs established by Bonferroni-corrected probabilities.

95% CI: 72.2-90.0) sensitivity in pretreatment plasma cfDNA samples. Most of the biopsy-confirmed tumor mutations not discovered in the cfDNA had low representation in the diagnostic biopsy tissue (<20%; Figure 4C). Consistently, using ROC analysis, cfDNA genotyping showed the highest sensitivity (91.3%; 95% CI: 79.1-97.1%; n = 42/46) in discovering mutations that were represented in ≥20% of the alleles of the tumor biopsy (Figure 4D).

Longitudinal monitoring of DLBCL genotype by using plasma cfDNA

A total of 127 plasma cfDNA samples were sequentially evaluated to assess the dynamics of mutations in plasma upon treatment with R-CHOP. Longitudinal analysis showed a rapid clearance of DLBCL mutations in the cfDNA of responding patients (Figure 6A). Among patients who were primary resistant to R-CHOP (patient ID12 and ID13), DLBCL mutations did not disappear from cfDNA (Figure 6B). In 1 of the 2 patients who responded to R-CHOP but ultimately had an early relapse in the central nervous system (patient ID9), mutations were still detectable in the plasma cfDNA sample collected while the patient was in complete remission (1 month before relapse). In addition, among patients who were primary refractory to R-CHOP or relapsed after treatment, new mutations appeared in the cfDNA that conceivably marked resistant clones that were selected during the clonal evolution process taking place under the selective pressure of treatment (Figure 7).

Discussion

By applying a training validation approach, this prospective study shows that plasma cfDNA genotyping: (1) is as accurate as genotyping of the diagnostic biopsy to detect somatic mutations of allelic abundance >20% in DLBCL; (2) allows the identification of mutations that are undetectable in the biopsy tissue conceivably because they are restricted to clones that are anatomically distant from the biopsy site; and (3) is a real-time, noninvasive tool to track clonal evolution and the emergence of treatment-resistant clones.

The identification of genomic alterations with clinical relevance in hematologic malignancies is increasing the need for assays that can routinely identify tumor mutational profiles. Among hematologic malignancies with a leukemic component, the accessibility of tumor cells in the PB has allowed the fast incorporation of gene mutations into genetic prognostic and predictive algorithms.[24-26] At variance with other hematologic malignancies, DLBCL typically lacks leukemic involvement, and bone marrow dissemination is infrequent.[27] Limited access to the tumor material has therefore hampered the development and validation of molecular prognostic models in DLBCL, whose molecular stratification represents an unmet medical need.[28] On these bases, among hematologic malignancies, DLBCL is an ideal model in which the liquid biopsy may allow a step forward in the translation of disease genetics into clinically useful markers, especially in the era of novel agents that are active in the molecular subgroup of the disease.[4]

Because of its >90% sensitivity and ~100% specificity, plasma cfDNA is an easily accessible source of tumor DNA that allows accurate profiling of DLBCL patients for cancer gene mutations represented in >20% of the tumor alleles, which is the sensitivity threshold of conventional Sanger sequencing methods that are broadly used to characterize tumor tissue specimens. On these bases, plasma cfDNA is an effective surrogate for direct tumor genotyping by conventional sequencing for the detection of clonally abundant mutations.

Plasma cfDNA genotyping also identifies variants that are subclonally represented in the tumor biopsy, although a proportion (~50%) of low-abundance mutations (ie, allele frequency <20% in the tumor biopsy) has been missed by our ultra-deep NGS approach, conceivably because of its chemistry-dependent sensitivity limit (~10^{-3}). Indeed, the allelic fractions of mutations in tumor biopsies and plasma samples are generally correlated, indicating that low abundant mutations in the tumor are scantily represented in plasma. The increasing evidence that small subclones have a clinical impact on treatment resistance and outcome on B-cell tumors prompts the development of sensitive approaches for cfDNA that are capable of exactly mirroring both the clonal and subclonal composition of the tumor.[29] For example,

From www.bloodjournal.org by guest on July 31, 2017. For personal use only.

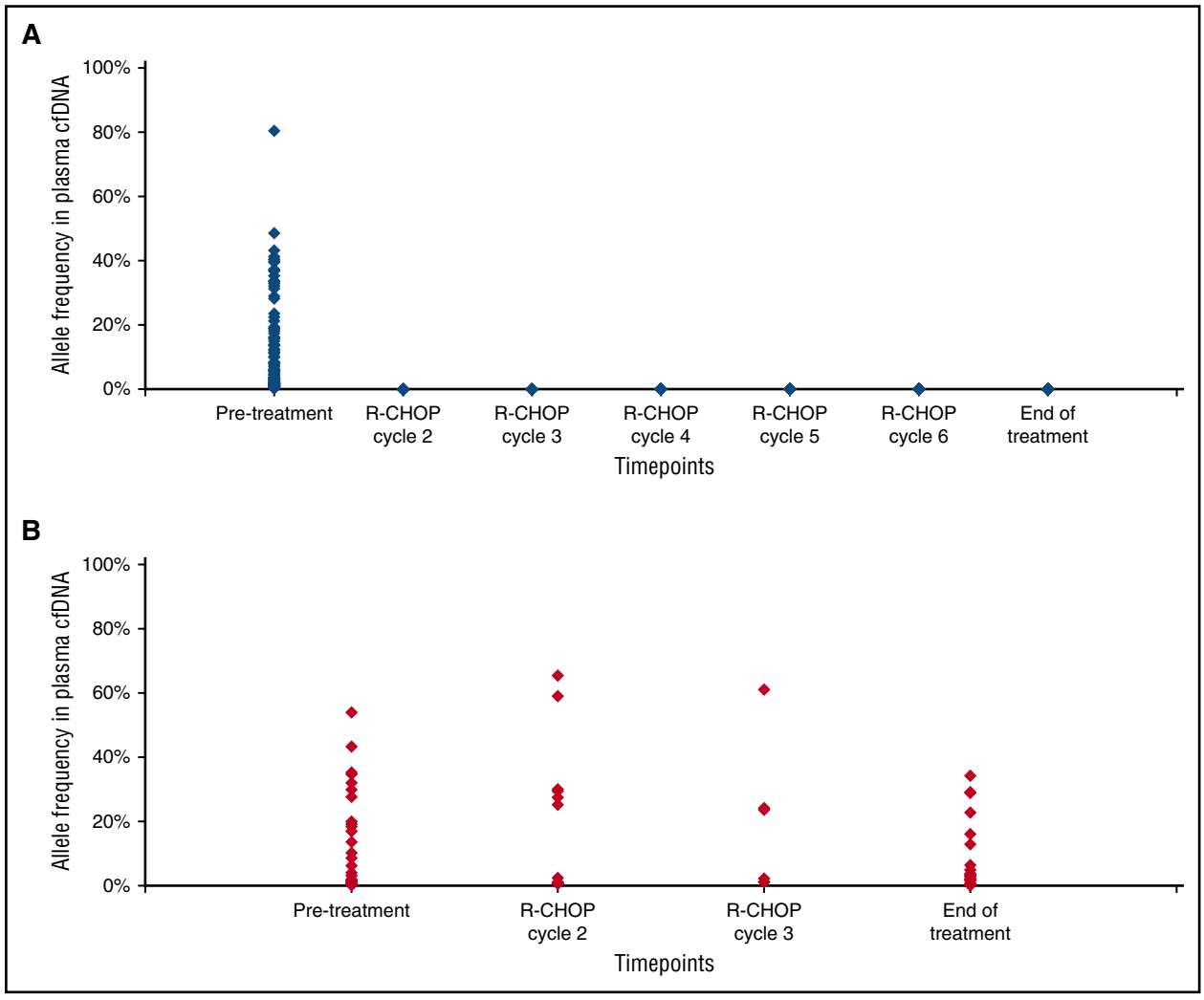

Figure 6. Longitudinal assessment of mutation abundance in plasma cfDNA upon R-CHOP treatment. The graphs represent the variant allele frequency of nonsynonymous somatic mutations in plasma cfDNA at baseline before the start of treatment, before each R-CHOP course, and at the end of treatment. Responding (A) and nonresponding (B) patients are pooled in the graphs. Each dot represents 1 single individual mutation. Among responding patients (A), baseline mutations (n = 102; mean allele frequency: 12.6%, range: 0.17%-80.4%) disappeared from plasma cfDNA. Among nonresponding patients (B), baseline mutations (n = 27; mean allele frequency: 14.3%, range: 0.14%-53.9%) remain detectable in plasma cfDNA before R-CHOP cycle 2 (mean allele frequency: 20.2%, range: 0.65%-65.3%), before R-CHOP cycle 3 (mean allele frequency: 18.8%, range: 1.1-61.1%), and at the end of treatment (mean allele frequency: 8.2%, range: 0.1%-34.2%).

the incorporation of molecular barcoding in the library preparation chemistry and the application of in silico bioinformatic algorithms to suppress background NGS artifacts allow for an ~15-fold increase of CAPP-seq sensitivity.[30]

Plasma cfDNA represents a complementary source of tumor DNA for DLBCL genotyping compared with the tissue biopsy. On the one hand, the complete molecular heterogeneity of a tumor cannot be adequately assessed by a single or even multiple tissue biopsies, whereas cfDNA genotyping captures genetic information shed from all sites of the disease. Also, plasma is an accessible source of tumor DNA when DNA cannot be retrieved from the diagnostic biopsy tissue (ie, because of the limited size or poor conservation of the specimen). On the other hand, plasma cfDNA genotyping misses a proportion of small subclonal mutations. Also, in its current modality, plasma cfDNA genotyping cannot accurately differentiate de novo DLBCL vs transformed tumors, precisely define the DLBCL cell of origin, and fully detect tumor chromosomal translocations,[11] which can instead be all routinely scored by analyzing the biopsy tissue. These notions suggest that liquid biopsy is not a substitute for the tumor biopsy, but instead provides complementary information in DLBCL.

Treatment of DLBCL is currently undergoing a shift from chemotherapy toward regimens incorporating targeted agents.[4,31] Along with the clinical development of novel targeted agents in DLBCL, mutation-driven mechanisms of resistance to these drugs are increasingly emerging.[4] On these bases, cfDNA genotyping may be used as a strategy of molecular monitoring to inform the acquisition of targeted drug-resistance in DLBCL. cfDNA analysis can be easily repeated at multiple time points on plasma samples obtained before, during, and after treatment to provide a picture of the changes in tumor genetics, including dynamic changes in the mutation profile that occur during therapy.[22]

In addition to clonal evolution, longitudinal cfDNA genotyping can also provide information about residual disease in cancer.[11] Our CAPP-seq approach does not reach the sensitivity of typical minimal residual disease assays, and the number of informative cases (ie, cases that achieved a radiological remission but ultimately relapsed) is small in the present study cohort. Despite such limitations, the observation that,

From www.bloodjournal.org by guest on July 31, 2017. For personal use only.

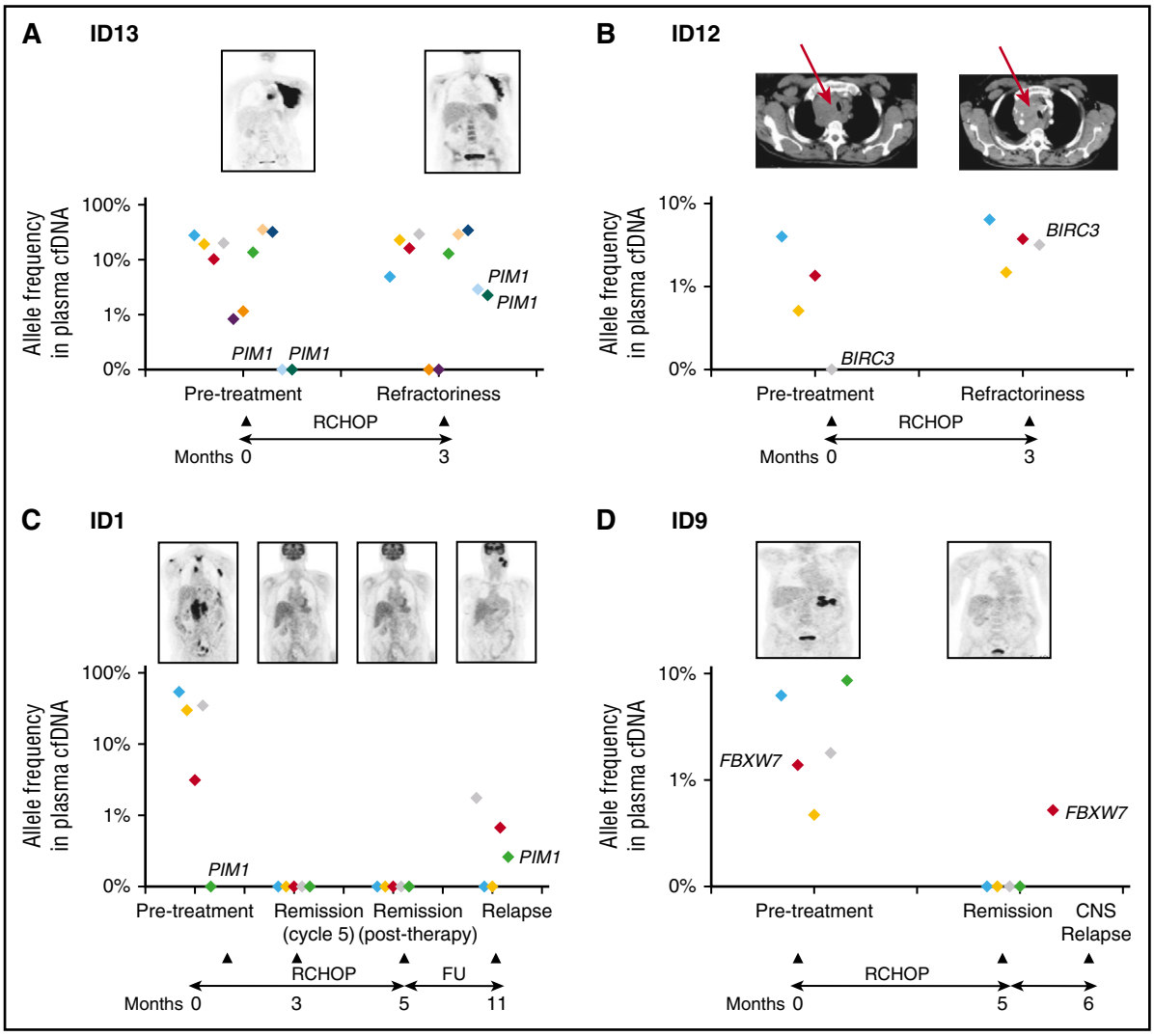

Figure 7. Noninvasive real-time monitoring of the DLBCL clonal evolution on plasma cfDNA. Sequential analysis of plasma cfDNA collected at presentation, upon treatment, and at the end of treatment. Disease response assessment by PET/CT is also shown at the corresponding time points. Representative patients are shown. Among nonresponding patients (ID13 [A] and ID12 [B]), baseline mutations remain detectable in plasma cfDNA. Among relapsing patients, baseline mutations cleared from plasma cfDNA and appeared again at the time of relapse (ID1 [C]) or persisted despite clinical remission (ID9 [D]). Color codes indicate different variants. Gene mutations selected by treatment are highlighted.

among early relapsing patients, low-level somatic mutations persisted in plasma while the disease was apparently in remission supports the development of plasma cfDNA mutational profiling as a tool for residual disease surveillance in DLBCL.[11]

There are relatively few studies investigating the importance of liquid biopsy in DLBCL that have shown very promising results with different approaches. In 2 studies, the *IGHV-D-J* gene segment of the rearranged immunoglobulin was used as a plasma cfDNA biomarker to identify patients at high risk of treatment failure, whereas in 2 additional studies, cfDNA was used to detect somatic mutations.[7,8,10,11] The retrospective nature of the previous studies that relied on archival material conceivably obtained and handled using different protocols might have precluded a tight control of the preanalytic factors that have a substantial effect on plasma cfDNA analysis, including the amount of "contaminating" DNA coming from circulating cells. The prospective design and the standardized protocols for sample processing used in our study allows for stringent control of preanalytic factors that might influence the results of plasma cfDNA genotyping. In particular, in less controlled situations, such as multicentered studies, the use of Cell-Free DNA BCT tubes that avoid storage-related preanalytic biases, can provide a broadly validated way to obtain stable cfDNA samples.[14]

An immediate clinical application of cfDNA genotyping in DLBCL patients is the incorporation of this assay in clinical trials to support post-hoc patient stratification according to baseline disease genetics, to develop treatment-specific prognostic, predictive, or actionable genomic biomarkers, and to track treatment resistance in a clinically relevant time frame.

Acknowledgments

The authors thank Carmela Ciardullo for her contribution to NGS.

This work was supported by Special Program Molecular Clinical Oncology 5 × 1000 No. 10007 grant, Associazione Italiana per la Ricerca sul Cancro Foundation (Milan, Italy); iCARE No. 17860 grant, Associazione Italiana per la Ricerca sul Cancro and Marie Curie Actions; Fondazione Cariplo grant 2012-0689; Progetto Ricerca

From www.bloodjournal.org by guest on July 31, 2017. For personal use only.

Finalizzata grants RF-2010-2307262 and RF-2011-02349712 from the Ministero della Salute (Rome, Italy); Futuro in Ricerca 2012 grant RBFR12D1CB from the Ministero dell'Istruzione, dell'Università e della Ricerca (Rome, Italy); and by grant KFS-3746-08-2015 from the Swiss Cancer League (Bern, Switzerland) and grant 320030_169670/1 from the Swiss National Science Foundation (Bern, Switzerland).

Authorship

Contribution: D.R. and G.G. designed the study, interpreted data, and wrote the manuscript; F.D. contributed to data interpretation and manuscript preparation; E.S., S.M., S.R., C.D., V.S., A.B., and C.F. performed molecular studies and contributed to manuscript revision; M.Z. and A.R. provided study material and contributed to manuscript revision; A.R. and R.B. performed pathological revision; R.S. contributed to bioinformatic analysis; and R.B. and R.F. contributed to data interpretation and manuscript revision.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Davide Rossi, Division of Hematology, Oncology Institute of Southern Switzerland and Institute of Oncology Research, Via Vincenzo Vela 6, 6500 Bellinzona, Switzerland; e-mail: davide.rossi@eoc.ch; and Gianluca Gaidano, Division of Hematology, Department of Translational Medicine, University of Eastern Piedmont, Via Solaroli 17, 28100 Novara, Italy; e-mail: gianluca.gaidano@med.uniupo.it.

References

1. Pasqualucci L, Dalla-Favera R. The genetic landscape of diffuse large B-cell lymphoma. *Semin Hematol*. 2015;52(2):67-76.
2. Sehn LH, Gascoyne RD. Diffuse large B-cell lymphoma: optimizing outcome in the context of clinical and biologic heterogeneity. *Blood*. 2015; 125(1):22-32.
3. Xu-Monette ZY, Wu L, Visco C, et al. Mutational profile and prognostic significance of TP53 in diffuse large B-cell lymphoma patients treated with R-CHOP: report from an International DLBCL Rituximab-CHOP Consortium Program Study. *Blood*. 2012;120(19):3986-3996.
4. Wilson WH, Young RM, Schmitz R, et al. Targeting B cell receptor signaling with ibrutinib in diffuse large B cell lymphoma. *Nat Med*. 2015; 21(8):922-926.
5. Diaz LA Jr, Bardelli A. Liquid biopsies: genotyping circulating tumor DNA. *J Clin Oncol*. 2014;32(6): 579-586.
6. Hohaus S, Giachelia M, Massini G, et al. Cell-free circulating DNA in Hodgkin's and non-Hodgkin's lymphomas. *Ann Oncol*. 2009;20(8):1408-1413.
7. Kurtz DM, Green MR, Bratman SV, et al. Noninvasive monitoring of diffuse large B-cell lymphoma by immunoglobulin high-throughput sequencing. *Blood*. 2015;125(24):3679-3687.
8. Roschewski M, Dunleavy K, Pittaluga S, et al. Circulating tumour DNA and CT monitoring in patients with untreated diffuse large B-cell lymphoma: a correlative biomarker study. *Lancet Oncol*. 2015;16(5):541-549.
9. Roschewski M, Staudt LM, Wilson WH. Dynamic monitoring of circulating tumor DNA in non-Hodgkin lymphoma. *Blood*. 2016;127(25): 3127-3132.
10. Bohers E, Viailly PJ, Dubois S, et al. Somatic mutations of cell-free circulating DNA detected by next-generation sequencing reflect the genetic changes in both germinal center B-cell-like and activated B-cell-like diffuse large B-cell lymphomas at the time of diagnosis. *Haematologica*. 2015;100(7):e280-e284.
11. Scherer F, Kurtz DM, Newman AM, et al. Distinct biological subtypes and patterns of genome evolution in lymphoma revealed by circulating tumor DNA. *Sci Transl Med*. 2016;8(364): 364ra155.
12. Cheson BD, Fisher RI, Barrington SF, et al; Alliance, Australasian Leukaemia and Lymphoma Group; Eastern Cooperative Oncology Group; European Mantle Cell Lymphoma Consortium; Italian Lymphoma Foundation; European Organisation for Research; Treatment of Cancer/ Dutch Hemato-Oncology Group; Grupo Español de Médula Ósea; German High-Grade Lymphoma Study Group; German Hodgkin's Study Group; Japanese Lymphorra Study Group; Lymphoma Study Association; NCIC Clinical Trials Group; Nordic Lymphoma Study Group; Southwest Oncology Group; United Kingdom National Cancer Research Institute. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J Clin Oncol*. 2014;32(27):3059-3067.
13. 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.
14. El Messaoudi S, Rolet F, Mouliere F, Thierry AR. Circulating cell free DNA: Preanalytical considerations. *Clin Chim Acta*. 2013;424: 222-230.
15. Norton SE, Lechner JM, Williams T, Fernando MR. A stabilizing reagent prevents cell-free DNA contamination by cellular DNA in plasma during blood sample storage and shipping as determined by digital PCR. *Clin Biochem*. 2013;46(15): 1561-1565.
16. Pasqualucci L, Trifonov V, Fabbri G, et al. Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat Genet*. 2011;43(9):830-837.
17. Lohr JG, Stojanov P, Lawrence MS, et al. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc Natl Acad Sci USA*. 2012;109(10):3879-3884.
18. Zhang J, Grubor V, Love CL, et al. Genetic heterogeneity of diffuse large B-cell lymphoma. *Proc Natl Acad Sci USA*. 2013;110(4):1398-1403.
19. 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.
20. Newman AM, Bratman SV, To J, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat Med*. 2014;20(5):548-554.
21. Pasqualucci L, Neumeister P, Goossens T, et al. Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature*. 2001; 412(6844):341-346.
22. Rabbitts TH, Hamlyn PH, Baer R. Altered nucleotide sequences of a translocated c-myc gene in Burkitt lymphoma. *Nature*. 1983; 306(5945):760-765.
23. Murtaza M, Dawson SJ, Tsui DW, et al. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature*. 2013;497(7447):108-112.
24. Döhner H, Weisdorf DJ, Bloomfield CD. Acute myeloid leukemia. *N Engl J Med*. 2015;373(12): 1136-1152.
25. Hunger SP, Mullighan CG. Acute lymphoblastic leukemia in children. *N Engl J Med*. 2015;373(16): 1541-1552.
26. Rossi D, Gaidano G. The clinical implications of gene mutations in chronic lymphocytic leukaemia. *Br J Cancer*. 2016;114(8):849-854.
27. Sehn LH, Scott DW, Chhanabhai M, et al. Impact of concordant and discordant bone marrow involvement on outcome in diffuse large B-cell lymphoma treated with R-CHOP. *J Clin Oncol*. 2011;29(11):1452-1457.
28. Weinstock DM, Dalla-Favera R, Gascoyne RD, et al. A roadmap for discovery and translation in lymphoma. *Blood*. 2015;125(13):2175-2177.
29. Rossi D, Khiabanian H, Spina V, et al. Clinical impact of small TP53 mutated subclones in chronic lymphocytic leukemia. *Blood*. 2014; 123(14):2139-2147.
30. Newman AM, Lovejoy AF, Klass DM, et al. Integrated digital error suppression for improved detection of circulating tumor DNA. *Nat Biotechnol*. 2016;34(5):547-555.
31. Younes A, Thieblemont C, Morschhauser F, et al. Combination of ibrutinib with rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP) for treatment-naive patients with CD20-positive B-cell non-Hodgkin lymphoma: a non-randomised, phase 1b study. *Lancet Oncol*. 2014;15(9):1019-1026.

From www.bloodjournal.org by guest on July 31, 2017. For personal use only.

2017 129: 1947-1957
doi:10.1182/blood-2016-05-719641 originally published online January 17, 2017

Diffuse large B-cell lymphoma genotyping on the liquid biopsy

Davide Rossi, Fary Diop, Elisa Spaccarotella, Sara Monti, Manuela Zanni, Silvia Rasi, Clara Deambrogi, Valeria Spina, Alessio Bruscaggin, Chiara Favini, Roberto Serra, Antonio Ramponi, Renzo Boldorini, Robin Foà and Gianluca Gaidano

Updated information and services can be found at:
http://www.bloodjournal.org/content/129/14/1947.full.html

Articles on similar topics can be found in the following Blood collections
Lymphoid Neoplasia (2595 articles)

Information about reproducing this article in parts or in its entirety may be found online at:
http://www.bloodjournal.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at:
http://www.bloodjournal.org/site/misc/rights.xhtml#reprints

Information about subscriptions and ASH membership may be found online at:
http://www.bloodjournal.org/site/subscriptions/index.xhtml

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.
Copyright 2011 by The American Society of Hematology; all rights reserved.