Distinct biological subtypes and patterns of genome evolution in lymphoma revealed by circulating tumor DNA

Florian Scherer,¹* David M. Kurtz,¹,²,³* Aaron M. Newman,¹,⁴* Henning Stehr,⁵ Alexander F. M. Craig,¹ Mohammad Shahrokh Esfahani,¹ Alexander F. Lovejoy,⁴,⁵,⁶† Jacob J. Chabon,⁴ Daniel M. Klass,¹,⁴,⁵† Chih Long Liu,¹,⁴ Li Zhou,⁵ Cynthia Glover,¹ Brendan C. Visser,⁷ George A. Poultsides,⁷ Ranjana H. Advani,¹ Lauren S. Maeda,¹,³ Neel K. Gupta,¹,³ Ronald Levy,¹ Robert S. Ohgami,⁸ Christian A. Kunder,⁸ Maximilian Diehn,⁴,⁵,⁶‡ Ash A. Alizadeh¹,³,⁴,⁵‡

Patients with diffuse large B cell lymphoma (DLBCL) exhibit marked diversity in tumor behavior and outcomes, yet the identification of poor-risk groups remains challenging. In addition, the biology underlying these differences is incompletely understood. We hypothesized that characterization of mutational heterogeneity and genomic evolution using circulating tumor DNA (ctDNA) profiling could reveal molecular determinants of adverse outcomes. To address this hypothesis, we applied cancer personalized profiling by deep sequencing (CAPP-Seq) analysis to tumor biopsies and cell-free DNA samples from 92 lymphoma patients and 24 healthy subjects. At diagnosis, the amount of ctDNA was found to strongly correlate with clinical indices and was independently predictive of patient outcomes. We demonstrate that ctDNA genotyping can classify transcriptionally defined tumor subtypes, including DLBCL cell of origin, directly from plasma. By simultaneously tracking multiple somatic mutations in ctDNA, our approach outperformed immunoglobulin sequencing and radiographic imaging for the detection of minimal residual disease and facilitated noninvasive identification of emergent resistance mutations to targeted therapies. In addition, we identified distinct patterns of clonal evolution distinguishing indolent follicular lymphomas from those that transformed into DLBCL, allowing for potential noninvasive prediction of histological transformation. Collectively, our results demonstrate that ctDNA analysis reveals biological factors that underlie lymphoma clinical outcomes and could facilitate individualized therapy.

INTRODUCTION
Diffuse large B cell lymphoma (DLBCL), the most common type of non–Hodgkin’s lymphoma (NHL), displays remarkable clinical and biological heterogeneity (1). Although therapy is curative in most cases, 30 to 40% of patients ultimately relapse or become refractory to treatment (2, 3). Accurate prediction of patient outcomes would facilitate individualized treatments, yet conventional methods for risk stratification and personalized therapy selection are limited. For example, the International Prognostic Index (IPI) classifies patients into risk groups based on clinical parameters but has failed to demonstrate utility for directing therapy (4, 5). In addition, metabolic imaging with positron emission tomography/computed tomography (PET/CT) has failed to improve survival in patients who relapse after initial response to therapy, in part because of low specificity (6–8).

Biomarkers based on tumor molecular features hold great promise for risk stratification and therapeutic targeting but are currently difficult to measure in clinical settings. For example, most DLBCL tumors can be classified into two transcriptionally distinct molecular subtypes, each derived from a specific B cell differentiation state [cell of origin (COO)]: germinal center B cell–like (GCB) and activated B cell–like (ABC) DLBCL (9–11). These subtypes are prognostic and may also predict sensitivity to emerging targeted therapies (12–15). Although several methods for COO assessment have been developed, the current gold standard is based on microarray gene expression profiling, which is clinically impractical because of its reliance on fresh frozen tissues (10, 11). In contrast, immunohistochemistry is routinely used for COO classification on fixed clinical samples but suffers from low reproducibility and accuracy. Although newer methods can overcome some of these issues (16), all existing approaches rely on the availability of invasive tumor biopsies (16–19).

Separately, a subset of patients are diagnosed with DLBCL after histological transformation from an otherwise indolent and low-grade follicular lymphoma (FL); these patients represent another biologically defined risk group in need of improved biomarkers (20, 21). Although several genetic aberrations have been linked to this event, no single factor has been shown to accurately predict transformation. In addition, the molecular properties of transformation remain poorly understood (22–25).

High-throughput sequencing (HTS) of circulating tumor DNA (ctDNA) in peripheral blood has recently emerged as a promising noninvasive approach for analyzing tumor genetic diversity and clonal evolution (26–32). Using cancer personalized profiling by deep sequencing (CAPP-Seq), an ultrasensitive capture-based targeted sequencing method, we performed deep molecular profiling...
of lymphoma tissue and cell-free DNA to define key biological features predictive of clinical outcomes (Fig. 1) (33, 34). Our findings reveal distinct patterns of genetic variation linked to adverse outcomes and emphasize the promise of noninvasive characterization of risk for managing patients with lymphoma.

RESULTS
Improved noninvasive profiling of tumor genetic heterogeneity in DLBCL
We and others previously showed that clonotypic immunoglobulin (Ig) V(D)J rearrangements can be detected and monitored in the peripheral blood of most DLBCL patients by HTS (IgHTS) (26, 27). However, IgHTS tracks a single tumor-specific genetic aberration and cannot capture the complex landscape of somatic variation in lymphoma. To overcome this shortcoming, we implemented a DLBCL-focused sequencing panel targeting recurrent single-nucleotide variants (SNVs), insertions/deletions, and breakpoints involving genes that participate in canonical fusions (BCL2, BCL6, MYC, and IGH). We also included Ig heavy-chain variable regions (IgVH) and the Ig heavy-chain joining cluster (IgLH) (table S1) (33–42). By profiling 92 human subjects at various disease milestones, we evaluated the technical performance of this targeted sequencing approach and the clinical utility of ctDNA for capturing DLBCL tumor genotypes.

We started by analyzing 76 diagnostic DLBCL tumor biopsies and 144 longitudinal plasma samples, 45 of which were obtained before treatment (figs. S1 to S6 and table S2). We identified somatic alterations in 100% of tumors with a median of 134 variants, including driver mutations in well-known DLBCL hotspot genes, IgH V(D)J rearrangements, and 89% of all chromosomal translocations previously identified by fluorescence in situ hybridization (FISH; fig. S1 and table S3). Applied to pretreatment plasma, our assay detected ctDNA in 100% of patients with 99.8% specificity when tumor genotypes were known (fig. S2). In addition, 91% of tumor-confirmed SNVs in driver genes could be noninvasively genotyped directly from pretreatment plasma, and this detection rate was directly correlated with ctDNA concentrations (fig. S3). At least one tumor-confirmed variant was identified by noninvasive genotyping in 87% of pretreatment plasma samples (39 of 45) and in all cases with ctDNA concentrations above 5 haploid genome equivalents (hGE)/ml (fig. S3B). Over this threshold, 95% of FISH-confirmed translocations in BCL2, BCL6, and MYC were detected by biopsy-free genotyping. This included a patient harboring a clinically important double hit lymphoma involving BCL2 and MYC, which is associated with poor prognosis (fig. S4A) (43–47). Because our panel targets multiple genomic regions and aberration types, we reasoned that it should have advantages over IgHTS for tumor genotyping and ctDNA assessment. In both historic studies of IgHTS and paired analyses in our own cohort, CAPP-Seq achieved higher sensitivity (Fig. 2, A and B) (26, 27). Thus, capture-based targeted sequencing can effectively detect somatic alterations in DLBCL tumors and plasma samples.

Because our approach can interrogate many mutations simultaneously, we next assessed whether the mutational architecture of DLBCL tumors is faithfully maintained in the plasma. We therefore determined and compared ctDNA burden serially over time, using mutations identified from either tumor biopsies or paired pretreatment plasma samples. Regardless of the source, the amount of ctDNA was highly concordant in serial plasma time points, both within individual patients and across all patients.
Emergence of Relapse PD

Scherer et al. suggest that, in most DLBCL patients, ctDNA is a robust surrogate for direct assessment of primary tumor genotypes.

Moreover, in nearly every patient, allele frequencies (AFs) of individual mutations found in both the primary tumor and the paired plasma were highly correlated (fig. S6). These data suggest that, in most DLBCL patients, ctDNA is a robust surrogate for direct assessment of primary tumor genotypes.

Next, we evaluated our method’s capability for biopsy-free detection of somatic alterations emerging during therapy or disease surveillance (Fig. 2, C and D, and figs. S7 and S8). We applied noninvasive genotyping to three patients with progressive disease receiving ibrutinib, an inhibitor of B cell receptor (BCR) signaling targeting Bruton tyrosine kinase (BTK). Resistance mutations in BTK have exclusively been described in tumor cells of patients with ibrutinib–refractory chronic lymphocytic leukemia and mantle cell lymphoma (48, 49). However, it remains unclear whether these mutations also occur in aggressive lymphomas, such as DLBCL, and whether they can be detected in plasma. By using ctDNA, we identified emergent resistance mutations in BTK that displayed distinct clonal dynamics in two of three patients (Fig. 2, C and D, and fig. S7, A and B).
patient, two adjacent BTK mutations encoding an identical amino acid substitution (BTK C481S) were found, but they were never observed within the same ctDNA molecule, demonstrating convergent evolution of independent resistant subclones (Fig. 2, C and D, and fig. S7A). These results suggest that tumor genotyping from plasma can facilitate monitoring of BTK-targeted therapy, regardless of histology. Thus, ctDNA profiling with CAPP-Seq has utility for real-time assessment of dynamic tumor processes, including clonal evolution and the acquisition of molecular resistance.

**Prognostic value of ctDNA in DLBCL**

Having demonstrated the technical performance of the assay, we next determined whether ctDNA analysis could facilitate early identification of clinically relevant risk groups in DLBCL. We started by comparing total ctDNA burden at diagnosis with standard clinical indices and risk of radiographic progression (Fig. 3 and fig. S9) (33). The amount of ctDNA was significantly correlated with serum lactate dehydrogenase (LDH; P < 1 × 10⁻⁴), the most commonly used biomarker for DLBCL (Fig. 3A and fig. S9A) (50).

---

**Fig. 3. Quantification of ctDNA in relation to DLBCL clinical indices and treatment response.** (A) Relationship between LDH and ctDNA concentration from pretreatment plasma time points. Correlations were determined separately above and below the upper limit of normal (ULN: 340 U/liter). (B) Correlation between MTV, measured from PET/CT imaging, and ctDNA concentrations from pretreatment plasma. Pretreatment LDH and MTV values in (A) and (B) were obtained as close in time as possible to blood draws used for plasma cell–free DNA sequencing (median, 6 days for LDH and 4 days for MTV). (C) Association between ctDNA concentration at diagnosis and Ann Arbor stage. Statistical comparison between early-stage (I + II) and late-stage (III + IV) patients was performed using Mann-Whitney U test. Means and SEMs are indicated. (D) Detection of ctDNA in relapsing patients as a function of time. Top: Cumulative fraction of patients with detectable ctDNA as a function of time before relapse. Bottom: Patient level data demonstrating ctDNA detection before relapse (n = 11). Clinical relapses were confirmed radiographically, and corresponding blood draws were taken within 30 days of diagnostic imaging, except for patients DLBCL088 (43 days) and DLBCL071 (78 days). All other blood draws were obtained between radiographic complete response and relapse (14 to 983 days before clinical relapse). Red circle, ctDNA detected; open circle, ctDNA not detected; black bars, imaging studies demonstrating complete response; red bars, imaging studies demonstrating detection of disease. Asterisks highlight patients with an isolated brain relapse. mo, months. (E) Kaplan-Meier analysis of PFS in patients with at least one ctDNA-positive plasma sample after the end of curative therapy compared to patients without detectable ctDNA after the end of curative therapy. Significance was assessed using the log-rank test.
Notably, whereas 100% of pretreatment samples had detectable ctDNA, only 37% of samples had abnormally high LDH, demonstrating superior sensitivity of ctDNA. Pretreatment ctDNA levels were strongly associated with metabolic tumor volumes (MTVs) measured using [18F]fluorodeoxyglucose PET/CT scans (Fig. 3B) (33). ctDNA concentrations at initial diagnosis were also significantly correlated with Ann Arbor stage (P = 3 × 10⁻⁴; Fig. 3C) and IPI (P < 1 × 10⁻⁴; fig. S9B) (5). Furthermore, we tested whether ctDNA concentrations at diagnosis were linked with the risk of future disease progression. In multivariate analyses incorporating key clinical parameters, higher ctDNA levels were continuously and independently correlated with inferior progression-free survival (PFS; table S4). Thus, pretreatment ctDNA in DLBCL can complement traditional clinical indices and serve as an independent prognostic biomarker.

**Early detection of DLBCL relapse**

Among the most promising clinical applications of ctDNA is its potential use for the detection of radiographically occult MRD (26, 27). We profiled plasma samples at times of radiographic complete response (n = 30) or recurrence (n = 8) from 11 patients, all of whom ultimately experienced disease progression despite therapy with curative intent. Whereas ctDNA was identified in all patients at the time of clinical relapse (Fig. 3D), it was also detectable as MRD before relapse in at least one plasma sample in 8 of 11 patients (73%), with ctDNA concentrations as low as 0.003% AF (0.11 hGE/ml). The mean elapsed time between the first ctDNA-positive time point and clinical relapse was 188 days, and all blood collections up to 3 months before relapse had ctDNA above the detection limit of our assay (Fig. 3D). When directly compared to IgHTS, our method detected MRD in twice as many patients with a mean lead time of >2 months, suggesting potential advantages in the surveillance setting (Fig. 3E and fig. S10) (26, 27). In contrast, ctDNA was undetectable in plasma samples from 10 patients who were disease-free for at least 24 months after therapy (51) and in 24 healthy adult subjects, demonstrating 100% specificity. Finally, we found that patients with ctDNA detected in plasma showed significantly inferior PFS compared to those with undetectable ctDNA (P = 3 × 10⁻⁴, log-rank test; Fig. 3F). This remained significant when controlling for “guarantee-time bias” (P = 8 × 10⁻⁵; likelihood ratio test), a potential confounding effect of comparing survival between groups when the classifying event (that is, ctDNA measurement) occurs during follow-up (52, 53). We observed a similar, though not significant, trend for overall survival (P = 0.056, log-rank test; fig. S11). Collectively, these results illustrate the promise of ctDNA profiling by targeted sequencing for improved MRD assessment and early relapse detection.

**COO classification**

COO classification of DLBCL is one of the strongest prognostic factors and a potential biomarker for future personalized therapies, yet accurate subtyping remains challenging in clinical settings (12–16, 19). We therefore used multiplexed somatic mutation profiling to develop a tool for COO classification from tumor or pretreatment plasma. By considering mutations enriched in GCB or non-GCB (ABC) DLBCL and targeted by our capture panel, we built a probabilistic classifier using a Bayesian approach (23, 54, 55). Patients in the training cohort were previously subtyped by microarray-based gene expression profiling of frozen tissues, currently considered the gold standard even if not clinically practical (fig. S12 and table S5) (23, 55). We then benchmarked the classifier performance using our cohort of 76 lymphoma tumor biopsies, predicting 44 patients as GCB and 32 as non-GCB (Fig. 4A). By comparing our results to a blinded, centralized immunohistochemical classification using the Hans algorithm (the current clinical standard), we observed a concordance rate of almost 80% (Fig. 4A) (17, 19). Patients identified as having GCB DLBCL by our classification approach had superior PFS over those identified as having non-GCB DLBCL (P = 0.02, log-rank test; Fig. 4B), consistent with previous descriptions of survival differences between COO subtypes (11). In addition, COO classifier scores were continuously associated with improved PFS (P = 3 × 10⁻³; Fig. 4C). Among patients analyzed by both immunohistochemistry and DNA genotyping, the Hans algorithm failed to stratify patient clinical outcomes, suggesting more accurate classification by our approach (Fig. 4D).

We next tested the COO classifier without knowledge of the tumor, using pretreatment plasma ctDNA (n = 41). The overall concordance between COO predictions from tumor tissue and biopsy-free plasma genotyping was 88% (Fig. 4E). Moreover, DLBCL molecular subtypes predicted directly from plasma were significantly associated with PFS in continuous models (P = 0.02; Fig. 4C). Thus, biopsy-free assessment of ctDNA has considerable potential for the classification of transcriptionally defined DLBCL subtypes.

**Patterns of genome evolution in patients with histological transformation**

Patients with aggressive DLBCL arising from histological transformation of an indolent FL represent another biologically defined risk group associated with poor prognosis (56, 57). We hypothesized that a comparative genomic analysis of paired tumor specimens might reveal biological features distinguishing histological transformation of FL (tFL), progression without transformation [nontransformed FL (ntFL)], and progression of DLBCL. Accordingly, we applied CAPP-Seq to three groups of paired tumor samples: (i) diagnostic FL versus tFL (n = 12), (ii) diagnostic FL versus ntFL (n = 12), and (iii) diagnostic de novo DLBCL versus relapsed/refractory DLBCL (rrDLBCL) (n = 7; Fig. 5 and figs. S13 and S14). We then compared the evolutionary history of these sequential tumor pairs by defining genetic alterations that were either common to both tumors or private to each (fig. S13A).

Among the three classes, we observed the greatest evolutionary distance among tumor pairs associated with histological transformation (Fig. 5, A and B, and figs. S13, B to D, and S14). This pattern was most pronounced when examining the fraction of mutations unique to the tumor biopsy at progression, which served to distinguish all three tumor subtypes (Fig. 5A and fig. S13D). Genomic divergence was independent of both the time to progression or transformation and the number of previous therapies, suggesting that this simple index could have utility as a biomarker of histological transformation (fig. S13E).

We therefore analyzed tumor biopsies obtained at diagnosis, along with follow-up plasma samples from patients with indolent lymphomas experiencing transformation (n = 8), progression without transformation (n = 7), or rrDLBCL (n = 11). In four patients, we additionally profiled follow-up plasma samples obtained before clinical evidence of transformation. Plasma genotyping results largely matched those from sequential tumors, with a higher fraction of emergent variants distinguishing tFL from other histologies (Fig. 5C). Separately, higher amounts of ctDNA were found to distinguish tFL and rrDLBCL from ntFL (Fig. 5D), suggesting that aggressive lymphomas display similar tumor cell proliferation.
light the potential of ctDNA as a noninvasive biomarker for early before clinical diagnosis. Together, these results demonstrate key ge-
in three of four plasma samples collected on an average of 66 days specificity (Fig. 5E). Moreover, our model successfully predicted tFL noninvasively classify tFL from ntFL with 83% sensitivity and 89%
work incorporating leave-one-out cross validation, we were able to
ering these discriminatory features within a logistic regression frame-
and turnover kinetics, despite their separate origins. When consid-
ering these discriminatory features within a logistic regression frame-
work incorporating leave-one-out cross validation, we were able to
noninvasively classify tFL from ntFL with 83% sensitivity and 89%
and our findings highlight the advantages of ctDNA as a noninvasive biomarker and provide a number of risk stratification strategies for clinical translation (Fig. 1).

For example, some patients with recurrent DLBCL undergo potentially curative subsequent therapies, including autologous stem cell transplantation (58). Although early detection of relapse has a potential for improving outcomes, surveillance imaging is considered to be largely ineffective for disease monitoring because of high false-
and our findings highlight the advantages of ctDNA as a noninvasive biomarker and provide a number of risk stratification strategies for clinical translation (Fig. 1).

For example, some patients with recurrent DLBCL undergo potentially curative subsequent therapies, including autologous stem cell transplantation (58). Although early detection of relapse has a potential for improving outcomes, surveillance imaging is considered to be largely ineffective for disease monitoring because of high false-
progression, with a mean lead time of more than 6 months. These results could inform clinical trial designs examining treatment paradigms based on early intervention directed by ctDNA detection.

DISCUSSION

Clinical and biological heterogeneity are key factors contributing to adverse risk and treatment failure in many can-
cers, including lymphomas. To address these challenges for patients with 
DLBCL, we applied CAPP-Seq, a highly sensitive targeted sequencing method, to analyze genetic profiles in 118 biop-
sies and 166 plasma samples from major disease milestones. In comparison to IgHTS, this approach achieved higher analytical and clinical sensitivity in capturing the mutational landscape of lymphoma and its clonal evolution. In addition, capture-based ctDNA analysis complemented cross-sectional imaging and facilitated the discovery of tumor molecular features and candidate biomarkers associated with high disease burden, relapse, non-GCB DLBCL, and histological transformation. Together,
In addition, accurate classification of GCB- and ABC-like molecular subtypes is important for determining prognosis in DLBCL patients. Here, we report a method for DLBCL classification based on integrating diverse somatic mutation profiles. This approach is both accurate and practical, allowing input material from either fixed tumor tissue or plasma samples, with high tumor-plasma concordance rates. Our noninvasive classification results were associated with clinical outcomes, suggesting a viable alternative to current methods that are limited by the requirement for invasive biopsies and suboptimal assay performance (11, 17, 61, 62). Moreover, the recent development of subtype-directed therapy has increased the importance of simultaneous disease classification and tumor genotyping (12–15). For example, patients classified as having ABC-like DLBCL by expression-based subtyping, and particularly those with ABC-like tumors that harbor gain-of-function mutations in BCR pathway genes (CD79B with or without MYD88), demonstrated a higher rate of ibrutinib efficacy (12).

In this cohort, we detected nine such patients by deep sequencing (tags for potential tumor transformation or progression) in patients with three distinct NHL types: tFL, ntFL, and rrDLBCL. The fraction of SNVs specific to tumor 2 (x axis) is compared with the proportion of SNVs shared between both tumors (y axis). Each dot represents a single patient. Shaded ovals highlight patients with different histologies, excluding outliers. The central node represents tumor 1, and the distance between tumor 1 and each patient’s tumor 2 (edge) is expressed as the fraction of unique mutations to both tumor 1 and tumor 2. Bar graph: percentage of SNVs unique to both tumor 1 and tumor 2 (nonshared mutations) for the median patient in each histological group. (C) Evolution of different types of NHL as determined by comparing diagnostic tumor samples from (A) (tumor 1) with follow-up plasma samples. The percentage of SNVs found uniquely in follow-up plasma compared to tumor 1 is shown for the three histologies. (D) Comparison of ctDNA concentrations in follow-up plasma samples from (C). Statistical comparisons in (C) and (D) were performed using the Mann-Whitney U test. Medians and ranges are indicated. (E) Performance metrics for the prediction of histological transformation from plasma. Sn, sensitivity; Sp, specificity; PPV, positive predictive value; NPV, negative predictive value. (F) Biopsy-free detection of an occult aggressive lymphoma subclone (tFL) in a patient histologically diagnosed with ntFL from a left inguinal lymph node biopsy (left, blue solid circle). The tumor site harboring the aggressive subclone (tFL, green dashed circle) was later identified in a retroperitoneal lymph node biopsy (right, green solid circle). Bottom: Venn diagram analysis of mutations found in tumor/plasma pairs at FL and tFL diagnosis. Mutations in key driver genes, such as CARD11 or PIM1, are indicated.

Fig. 5. Patterns of genome evolution in patients with histological transformation. (A) Comparison of mutation profiles from diagnostic tumor samples (“tumor 1”; FL or DLBCL) and follow-up tumor samples (“tumor 2”; transformation or progression) in patients with three distinct NHL types: tFL, ntFL, and rrDLBCL. The fraction of SNVs specific to tumor 2 (x axis) is compared with the proportion of SNVs shared between both tumors (y axis). Each dot represents a single patient. Shaded ovals highlight patients with different histologies, excluding outliers. (B) Network depiction of the mutational divergence between each tumor 1 and tumor 2 pair analyzed in (A). The central node represents tumor 1, and the distance between tumor 1 and each patient’s tumor 2 (edge) is expressed as the fraction of unique mutations to both tumor 1 and tumor 2. Bar graph: percentage of SNVs unique to both tumor 1 and tumor 2 (nonshared mutations) for the median patient in each histological group. (C) Evolution of different types of NHL as determined by comparing diagnostic tumor samples from (A) (“tumor 1”) with follow-up plasma samples. The percentage of SNVs found uniquely in follow-up plasma compared to tumor 1 is shown for the three histologies. (D) Comparison of ctDNA concentrations in follow-up plasma samples from (C). Statistical comparisons in (C) and (D) were performed using the Mann-Whitney U test. Medians and ranges are indicated. (E) Performance metrics for the prediction of histological transformation from plasma. Sn, sensitivity; Sp, specificity; PPV, positive predictive value; NPV, negative predictive value. (F) Biopsy-free detection of an occult aggressive lymphoma subclone (tFL) in a patient histologically diagnosed with ntFL from a left inguinal lymph node biopsy (left, blue solid circle). The tumor site harboring the aggressive subclone (tFL, green dashed circle) was later identified in a retroperitoneal lymph node biopsy (right, green solid circle). Bottom: Venn diagram analysis of mutations found in tumor/plasma pairs at FL and tFL diagnosis. Mutations in key driver genes, such as CARD11 or PIM1, are indicated.
could also guide therapy selection and improve treatment decisions by combining COO subtyping and assessment of favorable mutational patterns in a single assay (12). Separately, the framework we describe for disease classification using somatic alterations could extend to the noninvasive classification of many tumor types.

Finally, histological transformation of FL to DLBCL is characterized by a change from indolent to aggressive clinical behavior, associated with an unfavorable prognosis (56). We demonstrate that different NHL types, including tFL, exhibit distinct patterns of genome evolution. Among the subtypes that we evaluated, paired FL and tFL tumors showed the greatest evolutionary distance, on average, from their last common clonal progenitor, a finding that mirrors the marked shift in clinical presentation that accompanies transformation. By incorporating these genomic differences within a model, we found that FL transformation could be predicted with high sensitivity and specificity from ctDNA.

Given the clinical relevance of the reported results, further development and validation of our findings in larger patient cohorts will be needed. Such studies could lead to prospective clinical trials predicted with high sensitivity and specificity from ctDNA.

In summary, noninvasive genotyping and serial ctDNA monitoring are promising approaches for uncovering biology and improving patient management. We anticipate that ctDNA will have broad utility for dissecting tumor heterogeneity within and between patients with lymphomas and other cancer types, with applications for the identification of adverse risk groups, the discovery of resistance mechanisms to diverse therapies, and the development of risk-adapted therapeutics.

### MATERIALS AND METHODS

For detailed Materials and Methods, please see the Supplementary Materials.

### SUPPLEMENTARY MATERIALS

[www.sciencetranslationalmedicine.org/cgi/content/full/8/364/364ra155/DC1](http://www.sciencetranslationalmedicine.org/cgi/content/full/8/364/364ra155/DC1)

Materials and Methods

Fig. S1. Overview of DLBCL tumor genotyping results.

Fig. S2. Sensitivity and specificity of ctDNA detection in DLBCL pretreatment plasma samples.

Fig. S3. Performance assessment of biopsy-free tumor genotyping from DLBCL plasma samples.

Fig. S4. Utility of biopsy-free genotyping for transplantation detection and ctDNA monitoring.

Fig. S5. Analysis of biopsy-free ctDNA monitoring in serial plasma samples.

Fig. S6. Correlation of mutant AF from pretreatment tumor/plasma pairs.

Fig. S7. Noninvasive detection of ibrutinib resistance mutations in lymphoma patients.

Fig. S8. Noninvasive detection of an emergent somatic alteration after targeted therapy in a patient with tFL.

Fig. S9. Relationship between pretreatment ctDNA concentration and key DLBCL clinical indices.

Fig. S10. Performance comparison of CAPP-Seq and IgHTS for DLBCL relapse detection.

Fig. S11. Association between ctDNA positivity after curative therapy and overall survival.

Fig. S12. Genomic features incorporated into the DLBCL COO classifier.

Fig. S13. Analysis of mutation evolution in serial lymphoma tumor biopsies.

Fig. S14. Evolutionary patterns distinguishing lymphoma histologies.

Table S1. DLBCL selector design with references and final coordinates.

Table S2. Overview of patients, samples, and clinical characteristics.

Table S3. Somatic mutations and V(D)J recombination sequences detected in tumor biopsies.

Table S4. Univariate and multivariate outcome analysis.

Table S5. Illustrative example of DLBCL subtype determination.

Table S6. Overview of patients, samples, and clinical characteristics.

Table S7. DLBCL selector design with references and final coordinates.

Table S8. Summary of results for noninvasive ctDNA monitoring.

Table S9. Summary of results for noninvasive ctDNA monitoring.

Table S10. Summary of results for noninvasive ctDNA monitoring.

Table S11. Summary of results for noninvasive ctDNA monitoring.

Table S12. Summary of results for noninvasive ctDNA monitoring.

Table S13. Summary of results for noninvasive ctDNA monitoring.

Table S14. Summary of results for noninvasive ctDNA monitoring.

Table S15. Summary of results for noninvasive ctDNA monitoring.

Table S16. Summary of results for noninvasive ctDNA monitoring.

Table S17. Summary of results for noninvasive ctDNA monitoring.

Table S18. Summary of results for noninvasive ctDNA monitoring.

Table S19. Summary of results for noninvasive ctDNA monitoring.

Table S20. Summary of results for noninvasive ctDNA monitoring.

Table S21. Summary of results for noninvasive ctDNA monitoring.

Table S22. Summary of results for noninvasive ctDNA monitoring.

Table S23. Summary of results for noninvasive ctDNA monitoring.

Table S24. Summary of results for noninvasive ctDNA monitoring.

Table S25. Summary of results for noninvasive ctDNA monitoring.

Table S26. Summary of results for noninvasive ctDNA monitoring.

Table S27. Summary of results for noninvasive ctDNA monitoring.

Table S28. Summary of results for noninvasive ctDNA monitoring.

Table S29. Summary of results for noninvasive ctDNA monitoring.

Table S30. Summary of results for noninvasive ctDNA monitoring.

Table S31. Summary of results for noninvasive ctDNA monitoring.

Table S32. Summary of results for noninvasive ctDNA monitoring.

Table S33. Summary of results for noninvasive ctDNA monitoring.

Table S34. Summary of results for noninvasive ctDNA monitoring.

Table S35. Summary of results for noninvasive ctDNA monitoring.

Table S36. Summary of results for noninvasive ctDNA monitoring.

Table S37. Summary of results for noninvasive ctDNA monitoring.

Table S38. Summary of results for noninvasive ctDNA monitoring.

Table S39. Summary of results for noninvasive ctDNA monitoring.

Table S40. Summary of results for noninvasive ctDNA monitoring.

Table S41. Summary of results for noninvasive ctDNA monitoring.

Table S42. Summary of results for noninvasive ctDNA monitoring.

Table S43. Summary of results for noninvasive ctDNA monitoring.

Table S44. Summary of results for noninvasive ctDNA monitoring.

Table S45. Summary of results for noninvasive ctDNA monitoring.

Table S46. Summary of results for noninvasive ctDNA monitoring.

Table S47. Summary of results for noninvasive ctDNA monitoring.

Table S48. Summary of results for noninvasive ctDNA monitoring.

Table S49. Summary of results for noninvasive ctDNA monitoring.

Table S50. Summary of results for noninvasive ctDNA monitoring.

Table S51. Summary of results for noninvasive ctDNA monitoring.

Table S52. Summary of results for noninvasive ctDNA monitoring.

Table S53. Somatic mutations and V(D)J recombination sequences detected in tumor biopsies and a list of driver genes used in this work.

Table S54. Univariate and multivariate outcome analysis.

Table S55. Illustrative example of DLBCL subtype determination.


15. S. Batzoglou, R. Levy, A. A. Alizadeh, Mutations in early follicular lymphoma progenitors and predicts recurrence in patients with stage II colon cancer.


Acknowledgments: We thank the patients and their families who participated in this study. We would like to thank R. Tibshirani for the statistical advice related to the CAPP-Seq COO classifier. We also thank L. Pasquallucci for providing detailed information about genes informative for COO classification and J. Kress for assistance with the graphic design.

Funding: This work was supported by the Damon Runyon Cancer Research Foundation [DR-CL171-14 (to A.A.A.) and PST#09-16 (to D. M. Kurtz)], the American Society of Hematology Scholar Award (to A.A.A.), the V Foundation for Cancer Research Abeoloff Scholar Award (to A.A.A.), the German Research Foundation [SCH 1870/1-1 (to F.S.)], the Stanford TRAM (Translational Research and Applied Medicine) Pilot Grant (to A.A.A. and F.S.), the American Society of Clinical Oncology Young Investigator Award (to D. M. Kurtz), the National Cancer Institute (R01CA188298 and 1K99CA187192-01A1), the U.S. NIH Director’s New Innovator Award Program (1-DP2-CA186569), and the Ludwig Institute for Cancer Research. Author contributions: F.S., D. M. Kurtz, A.M.N., M.D., and A.A.A. developed the concept, designed the experiments, analyzed the data, and wrote the manuscript. F.S., D. M. Kurtz, A.M.N, H.S., M.S.E., and C.L.L. performed the bioinformatics analyses. F.S., D. M. Kurtz, A.F.M.C., A.F.L., J.J.C., D. M. Klass, and L.Z. performed the molecular biology experiments related to CAPP-Seq. C.A.K. performed all Hans immunohistochemistry analyses. C.G., B.C.V., G.A.P., R.H.A., L.S.M., N.K.G., R.L., and R.S.O. provided patient specimens and/or clinical data. M.D. and A.A.A. contributed equally as senior authors. All authors commented on the manuscript at all stages. Competing interests: A.M.N, D. M. Klass, and L.Z. are coinventors on patent applications related to CAPP-Seq. A.M.N, M.D., and A.A.A. are consultants for Roche Molecular Systems. Data and materials availability: Custom software used in this work was previously published and is available by request for nonprofit use (34).
Distinct biological subtypes and patterns of genome evolution in lymphoma revealed by circulating tumor DNA
Science Translational Medicine 8 (364), 364ra155. [doi: 10.1126/scitranslmed.aai8545]

Editor's Summary

The telltale DNA in lymphoma

Diffuse large B cell lymphoma is a relatively common type of tumor that can exhibit a wide range of behaviors, from indolent and curable cancers to ones that are very aggressive and difficult to treat. By analyzing DNA in tumor samples and blood of lymphoma patients, Scherer et al. have shown that specific genetic characteristics can determine each tumor's cell of origin and identify tumors that are going to transform into more aggressive subtypes and may require more intensive treatment. The authors also demonstrated that circulating tumor DNA in the patients' blood is suitable for this analysis, allowing for periodic monitoring of each patient without repeated invasive biopsies.

The following resources related to this article are available online at http://stm.sciencemag.org.
This information is current as of November 15, 2016.

---

Article Tools
Visit the online version of this article to access the personalization and article tools:
http://stm.sciencemag.org/content/8/364/364ra155

Supplemental Materials
"Supplementary Materials"
http://stm.sciencemag.org/content/suppl/2016/11/07/8.364.364ra155.DC1

Related Content
The editors suggest related resources on Science's sites:
http://stm.sciencemag.org/content/scitransmed/8/346/346ra92.full
http://stm.sciencemag.org/content/scitransmed/7/313/313ra182.full
http://stm.sciencemag.org/content/scitransmed/7/302/302ra133.full
http://stm.sciencemag.org/content/scitransmed/6/224/224ra24.full

Permissions
Obtain information about reproducing this article:
http://www.sciencemag.org/about/permissions.dtl