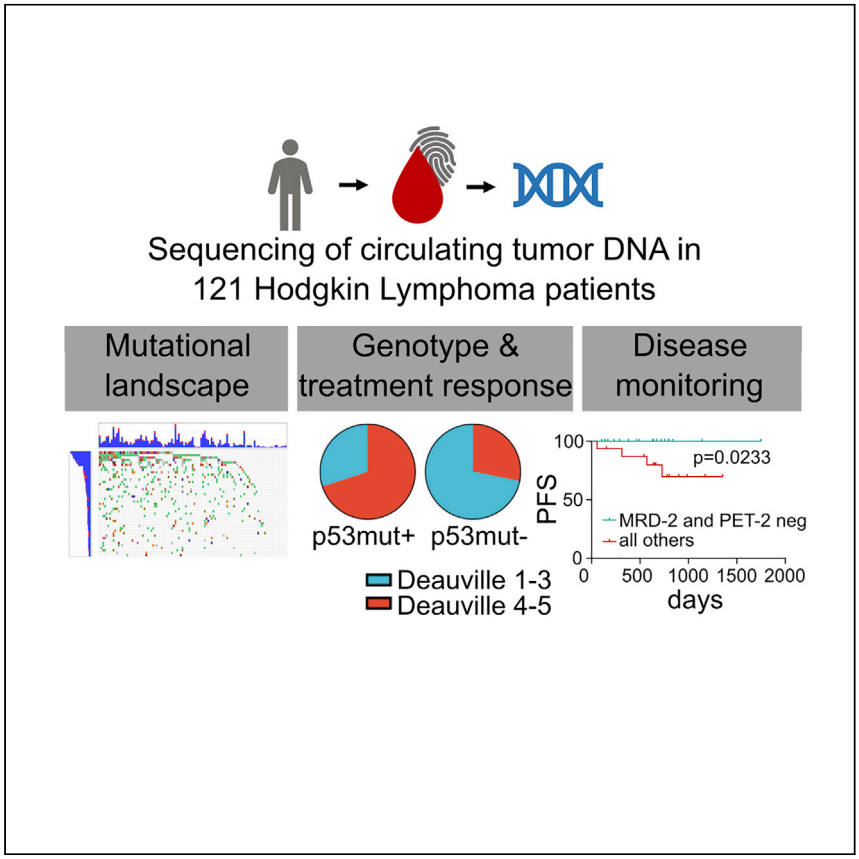
Med

CellPress

Clinical and Translational Article

In-depth cell-free DNA sequencing reveals genomic landscape of Hodgkin's lymphoma and facilitates ultrasensitive residual disease detection



Sobesky et al. show that circulating tumor DNA sequencing can be used to decipher the genomic landscape of Hodgkin's lymphoma and perform ultrasensitive minimal residual disease monitoring during and after treatment. Furthermore, they link various genotypes to tumor and patient phenotypes and outcomes.

Translation to Patients

Sophia Sobesky, Laman Mammadova, Melita Cirillo, ..., Peter Borchmann, Bastian von Tresckow, Sven Borchmann

sven.borchmann@uk-koeln.de

Highlights

ctDNA sequencing can comprehensively genotype most patients with Hodgkin's lymphoma

ctDNA sequencing reveals new associations of genotypes with phenotypes and outcome

Ultrasensitive minimal residual disease monitoring is possible with ctDNA sequencing

Sobesky et al., Med 2, 1171–1193
October 8, 2021
https://doi.org/10.1016/j.medj.2021.09.002

Med

CellPress

Clinical and Translational Article

In-depth cell-free DNA sequencing reveals genomic landscape of Hodgkin's lymphoma and facilitates ultrasensitive residual disease detection

Sophia Sobesky,[1,2,3] Laman Mammadova,[1,2,3] Melita Cirillo,[4] Esther E.E. Drees,[5,6] Julia Mattlener,[1,2,3] Helge Dörr,[1,2,3] Janine Altmüller,[7] Zhiyuan Shi,[1,2] Paul J. Bröckelmann,[1,2,3] Jonathan Weiss,[1,2] Stefanie Kreissl,[1,2,3] Stephanie Sasse,[8] Roland T. Ullrich,[1,2] Sarah Reinke,[9] Wolfram Klapper,[9] Elena Gerhard-Hartmann,[10,11] Andreas Rosenwald,[10,11] Margaretha G.M. Roemer,[6] Peter Nürnberg,[7] Anton Hagenbeek,[6] Josée M. Zijlstra,[6] Dirk Michiel Pegtel,[5] Andreas Engert,[1,2,3] Peter Borchmann,[1,2,3] Bastian von Tresckow,[3,12,13] and Sven Borchmann[1,2,3,14,*]

SUMMARY

Background: Individualization of treatment in Hodgkin's lymphoma is necessary to improve cure rates and reduce treatment side effects. Currently, it is hindered by a lack of genomic characterization and sensitive molecular response assessment. Sequencing of cell-free DNA is a powerful strategy to understand the cancer genome and can be used for extremely sensitive disease monitoring. In Hodgkin's lymphoma, a high proportion of cell-free DNA is tumor-derived, whereas traditional tumor biopsies only contain a little tumor-derived DNA.

Methods: We comprehensively genotype and assess minimal residual disease in 121 patients with baseline plasma as well as 77 follow-up samples from a subset of patients with our targeted cell-free DNA sequencing platform.

Findings: We present an integrated landscape of mutations and copy number variations in Hodgkin's lymphoma. In addition, we perform a deep analysis of mutational processes driving Hodgkin's lymphoma, investigate the clonal structure of Hodgkin's lymphoma, and link several genotypes to Hodgkin's lymphoma phenotypes and outcome. Finally, we show that minimal residual disease assessment by repeat cell-free DNA sequencing, as early as a week after treatment initiation, predicts treatment response and progression-free survival, allowing highly improved treatment guidance and relapse prediction.

Conclusions: Our targeted cell-free DNA sequencing platform reveals the genomic landscape of Hodgkin's lymphoma and facilitates ultrasensitive detection of minimal residual disease.

Funding: Mildred Scheel School of Oncology Aachen-Bonn-Cologne-Düsseldorf MD Research Stipend, Next Generation Sequencing Competence Network grant 423957469, Deutsche Krebshilfe grant 70112502, Deutsche Forschungsgemeinschaft (DFG) grant EN 179/13-1, the HL MRD consortium, and the Frau-Weiskam und Christel Ruranski-Stiftung.

Context and significance

Currently, most patients with Hodgkin's lymphoma are either treated too intensively and suffer from unnecessary side effects or are treated not intensively enough and experience relapse. Understanding the genetic drivers of each patient's disease before treatment and being able to very sensitively measure how much tumoral tissue is left during or after treatment could optimize treatment for each individual patient. Currently available technology for response assessment such as PET imaging is not very accurate. Here, we introduce a platform that analyzes the tumor genome of a Hodgkin's lymphoma patient and measures the remaining tumor burden using only blood samples. Our platform could improve Hodgkin's lymphoma treatment in the future by better defining risk groups of patients and assessing the remaining disease more sensitively than is currently possible.

INTRODUCTION

Classical Hodgkin's lymphoma (HL) is a B cell-derived malignant lymphoma that is now often curable with aggressive chemotherapy.[1,2] Nevertheless, many challenges

exist. Aggressive, multi-agent front-line chemo- and radiotherapy lead to early- and late toxicities such as secondary cancers, cardiovascular disease, infertility, fatigue, and osteonecrosis.[3–7] Furthermore, 10% to 30% of patients fail initial treatment, and many of these patients, who are often young, ultimately die.[8]

Individualizing treatment is key to improving outcomes and reducing long-term side effects in any cancer. Principally, two strategies for treatment individualization in a curable cancer exist: upfront treatment individualization by accurate risk assessment or accurate assessment of treatment response to stop treatment as soon as a cure is achieved.

In HL, both strategies are currently hampered by a lack of genomic characterization of the disease and sensitive molecular response assessment. Upfront risk assessment is currently purely based on clinical risk factors such as disease extent and has not improved much in recent decades.[9] Biological risk classification is impeded because a comprehensive genomic characterization of HL has so far been challenging due to the paucity of malignant Hodgkin Reed-Sternberg (HRS) cells in the typical HL tumor biopsy.[10] Past studies relied on cumbersome and technically challenging laser microdissection or flow sorting of biopsies for the enrichment of HL cells.[11–14] Response assessment in HL is currently based on positron emission tomography (PET), and strategies that include interim or end-of-chemotherapy PET-based response assessments to guide further treatment have recently been evaluated in clinical trials. In early-stage HL, PET-based response assessments have been mainly utilized to aim to spare patients with a negative PET scan from radiotherapy. Trials such as the RAPID trial,[15] the H10F/U trial,[16] and the HD16 trial[17] have shown that PET-based response assessments have predictive value in early-stage HL when it comes to progression-free survival (PFS); however, using it to guide treatment de-escalation by omitting radiotherapy resulted in a small but consistent loss in tumor control across trials. In advanced-stage HL, interim PET after two cycles is also prognostic, however, similar to early-stage HL, its ability to guide treatment is limited. Trials such as RATHL[18] or SWOG S0816[19] tested escalation of less intense ABVD (doxorubicin, bleomycin, vinblastine, and dacarbazine) chemotherapy in the case of a positive interim PET scan. These trials showed only limited improvement of outcomes compared to non-PET-guided strategies of comparable upfront treatment intensity. In contrast, PET-guided treatment was successful in the HD18 trial[20] that de-escalated treatment by reducing the number of more intense BEACOPPesc (bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone) treatment cycles in interim PET-negative patients without a loss of tumor control. As outlined above, sensitivity and specificity of PET-based response assessments is limited and highly divergent depending on details such as the order of different treatment regimens and the time point of PET assessment.[15–22] Overall, there is a high unmet clinical need for better biological risk classification and novel response assessment in HL.

A powerful strategy to unravel the cancer genome of a patient is the sequencing of cell-free DNA (cfDNA), which is released into the circulation from apoptotic cells and contains circulating tumor-derived DNA (ctDNA) with mutations, copy number alterations, gene fusions, or specific (clonal) immune-receptor (immunoglobulin H [IgH]) rearrangements representative of the tumor.[23,24] We and others have found that the plasma ctDNA content is much higher than one would expect in HL[24–28] considering the paucity of tumor cells, making it a highly relevant disease model to study ctDNA. Specifically, sequencing of ctDNA in HL allows easy and clinically feasible genotyping and biological risk assessment. Furthermore, carefully developed strategies to

[1]Department I of Internal Medicine, Center for Integrated Oncology Aachen Bonn Cologne Duesseldorf, University of Cologne, Köln, Germany

[2]Cancer Center Cologne Essen – Partner Site Cologne, CIO Cologne, University of Cologne, Köln, Germany

[3]German Hodgkin Study Group, Cologne, Germany

[4]University of Western Australia and Royal Perth Hospital, Perth, Australia

[5]Amsterdam UMC, Vrije Universiteit Amsterdam, Exosomes Research Group, Department of Pathology, Cancer Center Amsterdam, Amsterdam, the Netherlands

[6]Amsterdam UMC, Vrije Universiteit Amsterdam, Department of Hematology, Cancer Center Amsterdam, Amsterdam, the Netherlands

[7]Cologne Center for Genomics (CCG), University of Cologne, Cologne, Germany

[8]Department IV of Internal Medicine, Center for Integrated Oncology Aachen Bonn Cologne Duesseldorf, RWTH Aachen University, Aachen, Germany

[9]University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany

[10]Department of Pathology, University of Würzburg, Würzburg, Germany

[11]Comprehensive Cancer Center Mainfranken, Würzburg, Germany

[12]Department of Hematology and Stem Cell Transplantation, University Hospital Essen, University Duisburg-Essen, German Cancer Consortium (DKTK partner site Essen), Essen, Germany

[13]Cancer Center Cologne Essen - Partner Site Essen, West German Cancer Center, University Hospital Essen, University of Duisburg-Essen, Duisburg, Germany

[14]Lead contact

*Correspondence: sven.borchmann@uk-koeln.de

https://doi.org/10.1016/j.medj.2021.09.002

reduce sequencing errors enable highly sensitive and specific minimal residual disease (MRD) assessment by sequencing of ctDNA.[29,30]

Here, we comprehensively genotype and assess minimal residual disease in samples from 121 HL patients before, during, and after treatment. We present the largest integrated landscape analysis of mutations and copy number variations (CNVs) in HL to date and perform deep analysis of mutational processes driving HL, associations between HL and viruses and bacteria, and clonal structure of HL. We link several genotypes to HL phenotypes and treatment outcomes and show that MRD assessment is informative for overall treatment responses and PFS as early as one week after treatment initiation, suggesting high clinical utility.

RESULTS

Technical and clinical validation

We aimed to design a targeted ctDNA sequencing and bioinformatics platform for accurate baseline genotyping and MRD detection during follow up. To this end, cfDNA sequencing artifacts and stochastic sequencing errors need to be suppressed.[29] Combining existing knowledge,[25,26,29] our error suppression approach included tagging DNA molecules with unique molecular identifiers (UMIs) and a combination of two comparative error suppression (CES) methods: (1) deriving and excluding error-prone genomic regions in a identically processed set of cfDNA samples from a healthy control cohort and (2) judging each base call based on its modeled error-rate in an identically processed healthy control cohort considering the given tri-nucleotide context[31] (see STAR Methods for details).

For technical validation, we used spike-ins of well-characterized reference DNA (NA12878)[32,33] fragmented to cfDNA-size into healthy donor cfDNA. Observed error profiles depended on the use of UMI-based error suppression or additional CES (Figures 1A–1C). Combined UMI-based and CES error suppression resulted in improved sensitivity and specificity of mutation calling at 0.5% spiked-in mutated allele frequency (mAF) (Figures 1D and 1E). UMIs contributed most to sensitivity (Figure 1D), whereas CES was essential for high specificity (Figure 1E). For technical validation of MRD detection, we used lower spike-ins of 0.5%, 0.05%, 0.025%, 0.0125%, and 0.005% mAF and were able to show that combined error suppression is necessary to reliably distinguish MRD from the background below a mAF (corresponding to the MRD level) of 0.025% or 1 mutated allele in 4,000 DNA molecules (Figure 1F).

For clinical validation of mutation calling, we compared detected variants in 8 cfDNA samples with corresponding Formalin-Fixed Paraffin-Embedded (FFPE) tumor biopsies. We confirmed all mutations called in cfDNA with a mAF >4% (26/26) and 91.3% (63/69) of all mutations irrespective of mAF (Figures 1G and 1H). This is in line with our expectations, as mutation calling in artifact rich bulk FFPE samples with low HL cell content will fail to detect rare variants, underscoring the value of ctDNA sequencing in HL. For clinical validation of copy number assessment in cfDNA, we compared 9p24.1 (locus of PD-L1) gains assessed in cfDNA with gold standard fluorescence *in situ* hybridization (FISH). Average PD-L1 copies assessed by FISH in HL lymph node biopsies correlated strongly with copy number gain in cfDNA (Figure 1I). As an additional validation, we compared the baseline and remission sample detection of the 9p24.1 gain in patients with this gain and at least 3 remission samples (n = 5). We did not detect the 9p24.1 gain in any of the remission samples at the prespecified threshold (Figure 1J).

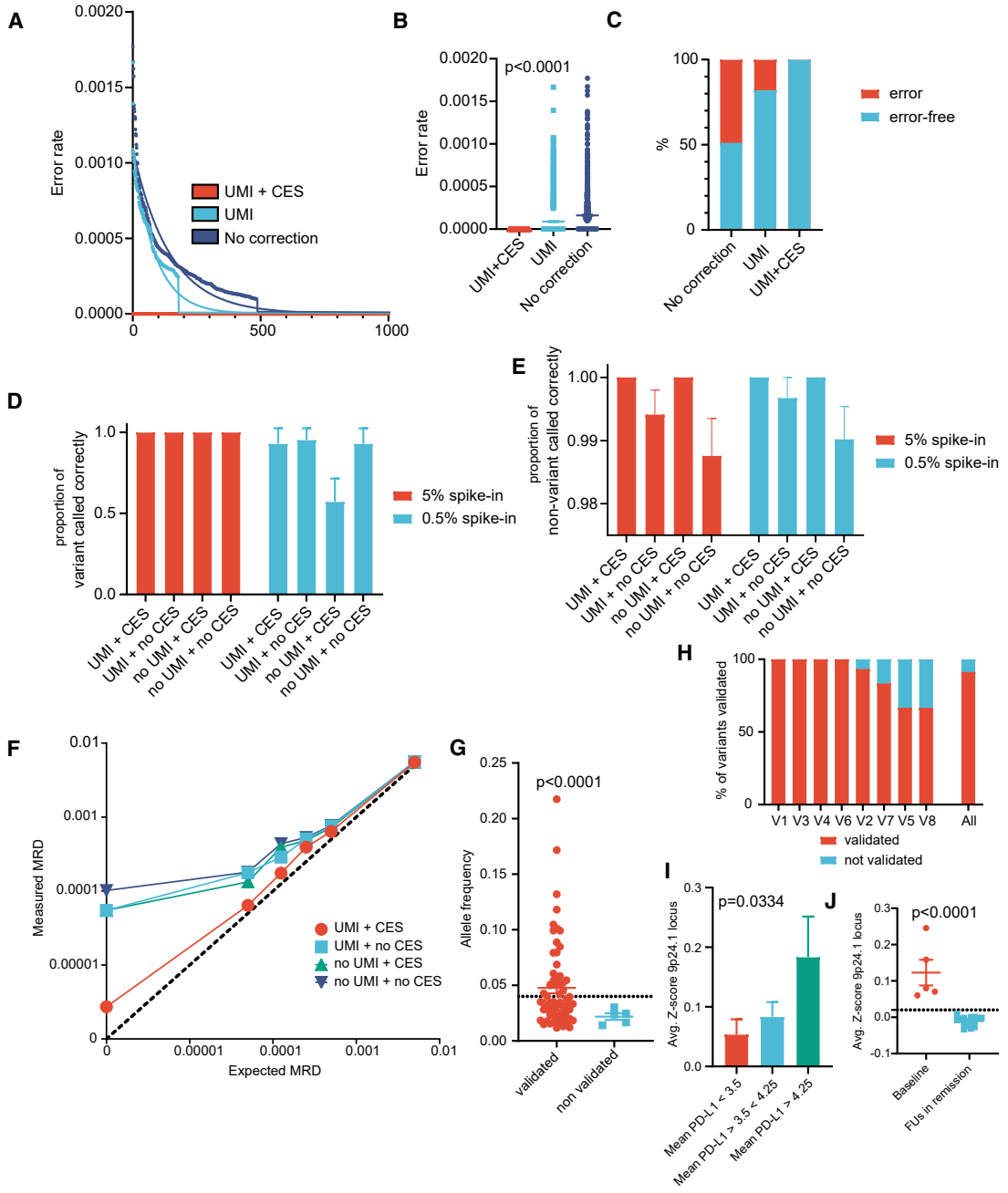


Figure 1. Technical and clinical validation of ctDNA sequencing platform

(A) Error profile of indicated error suppression methods across 1,000 random non-mutated bases. Dots indicate observed error rate at a base. Lines indicate a one-phase decay model of the respective error rate.

(B) Average error rate of indicated error suppression methods across 1,000 random non-mutated bases.

(C) Error-free rate of indicated error suppression methods across 1,000 random non-mutated bases.

(D and E) Technical validation of variant calling: shown is the proportion of correctly called mutations across 42 truly mutated bases (D) and the proportion of correctly called non-mutated bases across 1,527 random non-mutated bases (E) for the indicated error suppression methods.

(F) Technical validation of minimal residual disease (MRD) assessment with linearity test for indicated error suppression methods (n = 24). The dotted line indicates true linearity (the measured allele frequency [AF] in technical spike-in controls equals the expected (spiked-in) AF.

(G) Comparison between AF of variants identified in cfDNA that were validated in FFPE-tissue and those that were not (n = 8 clinical validation samples).

(H) Validation rate of variants identified in cfDNA in FFPE tissue samples shown separately across 8 sample pairs and summarized for all sample pairs.

Based on standard protocols for assay validation,[34] we defined a limit of blank of 0.00714% (approximately 1 in 14,000 DNA molecules) and a limit of detection of 0.00945% (approximately 1 in 10,500 DNA molecules) for MRD detection. Employing baseline variant calls from 6 HL patients, we tested 10 healthy donor samples as negative MRD controls. All healthy control samples showed no evidence of MRD (Data 1 in the accompanying Mendeley dataset[35]).

Sequencing and patient characteristics

Patients were sourced from 2 different cohorts: a primary cohort of 111 well-annotated patients from clinical trials used throughout the study unless indicated otherwise and a secondary cohort of 10 routine care patients, which was added to two parts of this study: investigating associations between genotypes and MRD with outcome (Supplemental methods Figure 1; please see STAR Methods for details). We produced a mean of 2.02×10^8 ($\pm 7.51 \times 10^6$ standard error of mean) reads per HL patient cfDNA sample with a mean of 6.84×10^7 ($\pm 3.65 \times 10^6$) uniquely mapped reads (Supplemental methods Figure 2A). The high duplication rate (mean: 3.67 $\pm$ 0.25) (Supplemental methods Figure 2B) was intentional and utilized for UMI-based error suppression. The mean post-UMI-deduplication per gene coverage was 1902× ($\pm$48), and the mean per patient coverage was 1670× ($\pm$90) (Supplemental methods Figures 2C and 2D). The median cfDNA concentration was 14.1 ng/mL plasma (range: 2.5–557.5 ng/mL) (Supplemental methods Figure 2E). Baseline cfDNA concentration was higher in males (median 19.7 versus 10.8 ng/mL; $p = 0.0587$) (Supplemental methods Figure 2F) and patients with a higher clinical stage ($p = 0.0001$) (Supplemental methods Figure 2G) or a higher International Prognostic Score[9] ($p < 0.0001$) (Supplemental methods Figure 2H). Patient characteristics were as expected with a representative stage, age, gender, and histological subtype distribution[1] (Table 1; Data 2 in the accompanying Mendeley dataset[35]).

Mutational landscape of HL

We detected 4,738 single base substitutions (SBSs) and small insertions or deletions (InDels) in our primary cohort with a median of 36 mutations per patient (range 1–175) (Supplemental methods Figure 2I; Data 3 in the accompanying Mendeley dataset[35]). Variants were identified in 109 out of 111 patients (98%). Non-synonymous mutations were predominant (n = 2,050), followed by non-coding mutations (n = 1,891), synonymous mutations (n = 614), mutations in untranslated regions (UTRs) (n = 92), and splicing mutations (n = 44) (Supplemental methods Figure 2J). This was as expected for our enrichment approach targeting mostly exons.

The genes with most non-synonymous, non-intronic, and non-intergenic mutations were TTN (n = 71), SOCS1 (n = 58), TNFAIP3 (n = 51), ITPKB (n = 50), STAT6 (n = 42), GNA13 (n = 36), B2M (n = 34), and CSF2RB (n = 28) (Data 4 in the accompanying Mendeley dataset[35]) (Figures S1A and S1B) after conservatively removing putative mutations detected in MUC4 or MUC16, as these could be artifacts resulting from paralogous alignment.[36] The number of mutations per patient was highly heterogeneous (Figures S1C and S1D).

Mutational frequency is not an indicator of relevance, for example, because large genes such as TTN, the most frequently mutated gene in this cohort, possibly only

(I) Comparison between HRS cell CD274/PD-L1 copies assessed by FISH in tissue and copy number *Z*-score in corresponding cfDNA samples (n = 51).
(J) Comparison between copy number *Z*-score of the 9p24.1 locus in corresponding baseline (n = 5) and follow-up cfDNA samples (n = 23).
Error bars show bootstrap confidence intervals for (D) and (E) and SEM for (B), (G), (I) and (J). One-way ANOVA (B), t test with Welch correction (G and J), and one-way ANOVA with trend test (I) were used.

Table 1. Patient characteristics of primary cohort

Patient characteristics primary cohort (n = 111)	Mean (n)	Range (%)
Study		
HD21	43	38.74
NIVAHL	68	61.26
Sex		
Male	54	48.65
Female	57	51.35
Age		
	32	[18-57]
IPS**		
0	13	11.71
1	39	35.14
2	29	26.13
3	20	18.02
4	8	7.21
5	2	1.80
Clinical stage		
1	2	1.80
2	75	67.57
3	17	15.32
4	17	15.32
Histology		
nodular sclerosis	75	67.57
classical Hodgkin's lymphoma without possible subtyping	20	18.02
mixed cellularity	14	12.61
lymphocyte rich	2	1.80
ECOG***		
0	74	66.67
1	36	32.43
2	1	0.90
Treatment		
2 cycles (4 doses) Nivolumab + 2 cycles Nivolumab-AVD + 2 cycles of AVD + 30 Gy IS-RT	34	30.63
4 cycles of Nivolumab-AVD + 30 Gy IS-RT	34	30.63
4 cycles of BRECADD or BEACOPPesc*	23	20.72
4 cycles of BRECADD or BEACOPPesc* + 30Gy RT on remaining PET-positive lesions	4	3.60
6 cycles of BRECADD or BEACOPPesc*	13	11.71
6 cycles of BRECADD or BEACOPPesc* + 30Gy RT on remaining PET-positive lesions	3	2.70

*Treatment allocation to BRECADD or BEACOPPesc in the HD21 trial is still blinded until the final analysis (expected in 2024).

**IPS: International Prognostic Score.

***ECOG: Eastern Cooperative Oncology Group Performance Status.

harbors many mutations because of its size.[36] To identify genes likely relevant in HL pathogenesis and not merely bystander mutations, we used a modified Vogelstein rule[37] (see STAR Methods for details). We identified 23 tumor suppressor genes (TSGs) and 20 oncogenes (OGs) (Figure 2A; Data 5 in the accompanying Mendeley dataset[35]). The number of OG and TSG mutations per patient was highly heterogeneous, with a median of 5 (Figures 2B and 2C). We identified recurrent mutations in several known oncogenic pathways in HL such as JAK-STAT signaling[11] (SOCS1, STAT6, CSF2RB, IL4R), class I major histocompatibility complex (MHC)-mediated antigen processing and presentation[12] (B2M), and NF-kappa B signaling[38]

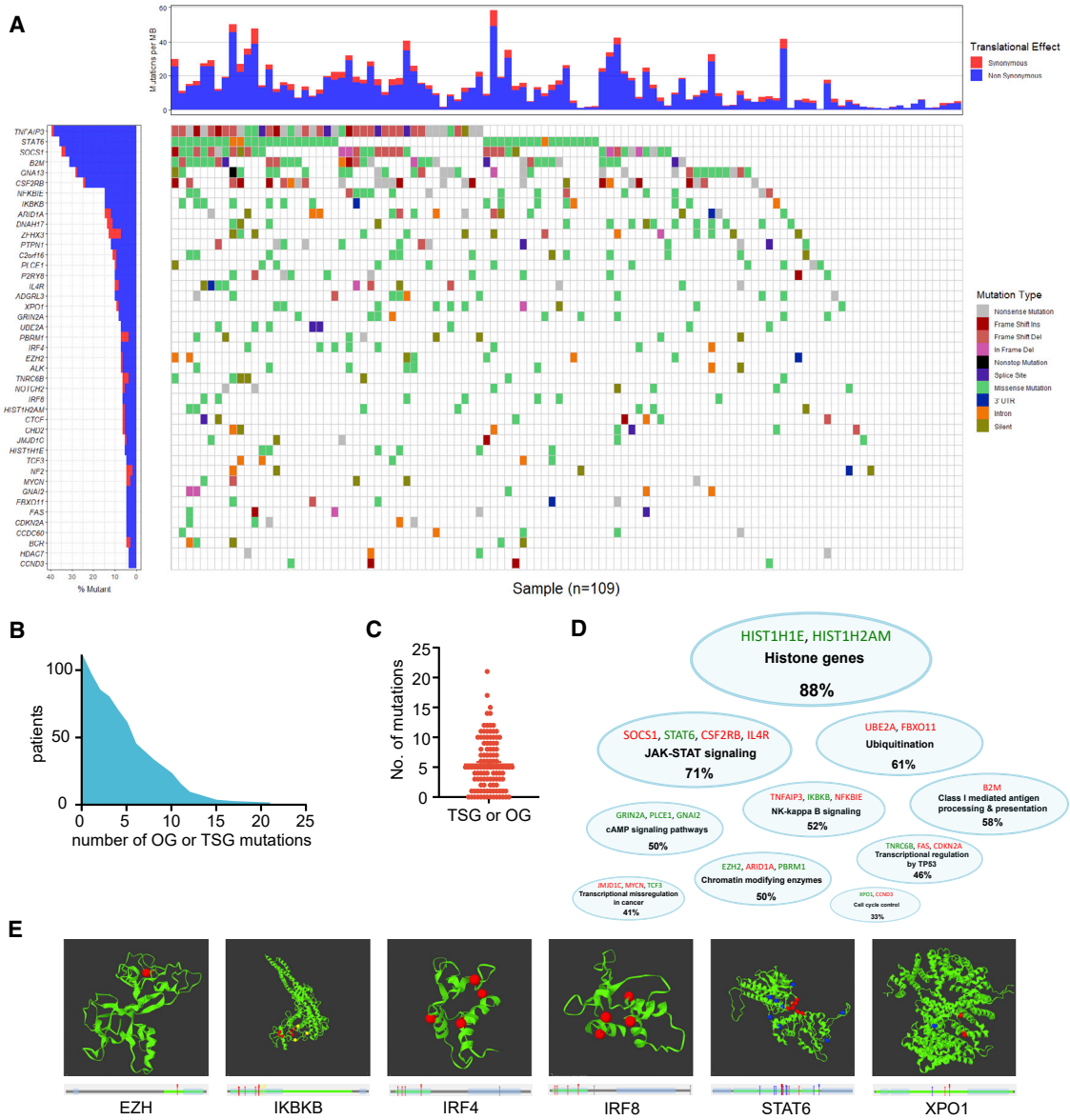


Figure 2. Mutational landscape of Hodgkin's lymphoma

(A) Waterfall plot of oncogenes and tumor suppressor genes identified in the patient cohort (n = 109; 2 patients without identified mutations not shown). The chart on the left shows the genes and proportions of patients with mutations. The chart on top shows the mutational burden per patient. Color codes indicate the type of mutation in a patient and proportion of synonymous versus non-synonymous mutations in genes (left chart) and patients (top chart), respectively.

(B) Cumulative plot of patients and number of mutations in identified oncogenes and tumor suppressor genes (n = 111).

(C) Number of mutations in identified oncogenes and tumor suppressor genes per patient (n = 111).

(D) Cloud figure of affected pathways by mutations in oncogenes and tumor suppressor genes. Percentages indicate proportion of patients affected by at least one mutation in indicated pathway. Shown identified oncogenes (green) and tumor suppressor genes (red) are color coded.

(E) 3D protein models of oncogenes with significant spatial clustering of mutations. Red, blue, and yellow dots indicate separate spatial clusters. The charts below the 3D models depict linear representations in which arrows with the same colors as in the 3D models indicate position of spatial clusters on this linear representation of the protein sequence.

Error bars show SEM.

(TNFAIP3, IKBKB, NFKBIE). In addition, we identified several recurrent TSGs and OGs involved in chromatin modification (EZH2, ARID1A, PBRM1); ubiquitination (UBE2A, FBXO11); cAMP signaling (GRIN2A, PLCE1, GNAI2); transcriptional cell

fate regulation by TP53 (TNRC6B, FAS, CDKN2A); cell-cycle regulation (XPO1, CCND3); transcriptional misregulation in cancer (JMJD1C, MYCN, TCF3); and interferon signaling (IRF4, IRF8, PTPN1), as well as histone genes (HIST1H1E, HIST1H2AM) (Figure 2D). Next, we expanded our analysis by allocating identified non-synonymous mutations to these pathways based on predefined gene sets (Data 6 in the accompanying Mendeley dataset[35]). Almost all patients had histone mutations (87.9%), while most patients had mutations in the JAK-STAT signaling pathway (71.0%), interferon signaling (60.7%), ubiquitination (60.7%), class I MHC-mediated antigen processing and presentation (57.9%), and nuclear factor (NF)-kappa B signaling (52.3%) (Figure 2D).

To understand their functional relevance, we analyzed spatial clustering of mutations in OGs within 3-dimensional protein structures using Mutation3D.[39] We identified significant spatial clustering in EZH2, IKBKB, IRF4, IRF8, STAT6, and XPO1 (Figure 2E). Mutations in EZH2 all occurred in residue 641 (NM_001203247), leading to a substitution of tyrosine by asparagine ($p_{cluster} = 6.67 \times 10^{-5}$). This mutation has recently been reported as recurrent in follicular lymphoma and diffuse large B cell lymphoma (DLBCL) and is associated with a gain-of-function of EZH2 with increased H3K27me activity.[40–42] Recurrent Y641 EZH2 mutations in HL have so far not been reported.[11]

For IKBKB, we identified two clusters ($p_{cluster} = 1.7 \times 10^{-3}$ each). Both clusters are located in the PKinase domain of IKBKB. One of the mutations results in a substitution of valine by isoleucine at residue 203 (NM_001556). This results in a gain-of-function of encoded IKK (IκB kinase) and increased NF-κB (nuclear factor 'kappa-light-chain-enhancer' of activated B cells) signaling, as has been shown by the profound immune deficiency caused by V203I IKBKB germline mutations.[43] Somatic V203I IKBKB mutations have also been identified in HL before.[14,25]

For IRF4, we identified a cluster in the IRF (DNA binding) domain ($p_{cluster} = 1.83 \times 10^{-4}$). IRF4 mutations clustering in the same domain have been observed in chronic lymphocytic leukemia (CLL)[44] and T cell leukemia,[45] while a subgroup of DLBCL harbors translocations affecting IFR4.[46] In HL, recurrent IRF4 mutations have not been described; however, it has been shown that IRF4 silencing inhibits HL proliferation in cell line models.[47] Functionally, mutations in the cluster identified by us are in close proximity to mutations that have been shown to increase IRF4 transcriptional activity.[45]

Similar to IRF4, we identified a cluster of mutations in IRF8 in the IRF (DNA binding) domain ($p_{cluster} = 2.52 \times 10^{-4}$). IRF8 mutations have not been described in HL but do occur recurrently in DLBCL,[48] and IRF8 knock down has been reported to lead to decreased DLBCL proliferation.[49]

For STAT6, we identified a cluster of mutations in the STAT (DNA binding) domain ($p_{cluster} = 5.56 \times 10^{-2}$). Recurrent mutations in this STAT6 domain have been identified in PMBCL[50] and HL before.[11] Within this cluster we identified multiple patients with the previously described N417Y, D419N, and N421S (NM_003153) mutations.[11] Functionally, STAT6 expression is important for HL survival,[51] with more dependency of the HL cell line L1236, which harbors the N417Y mutation, on constitutive STAT6 expression,[11] underscoring the oncogenicity of STAT6 in HL.

Finally, we identified a cluster of mutations in XPO1 ($p_{cluster} = 9.71 \times 10^{-5}$) consisting mostly of the recurrent E571K (NM_003400) mutation substituting glutamate by

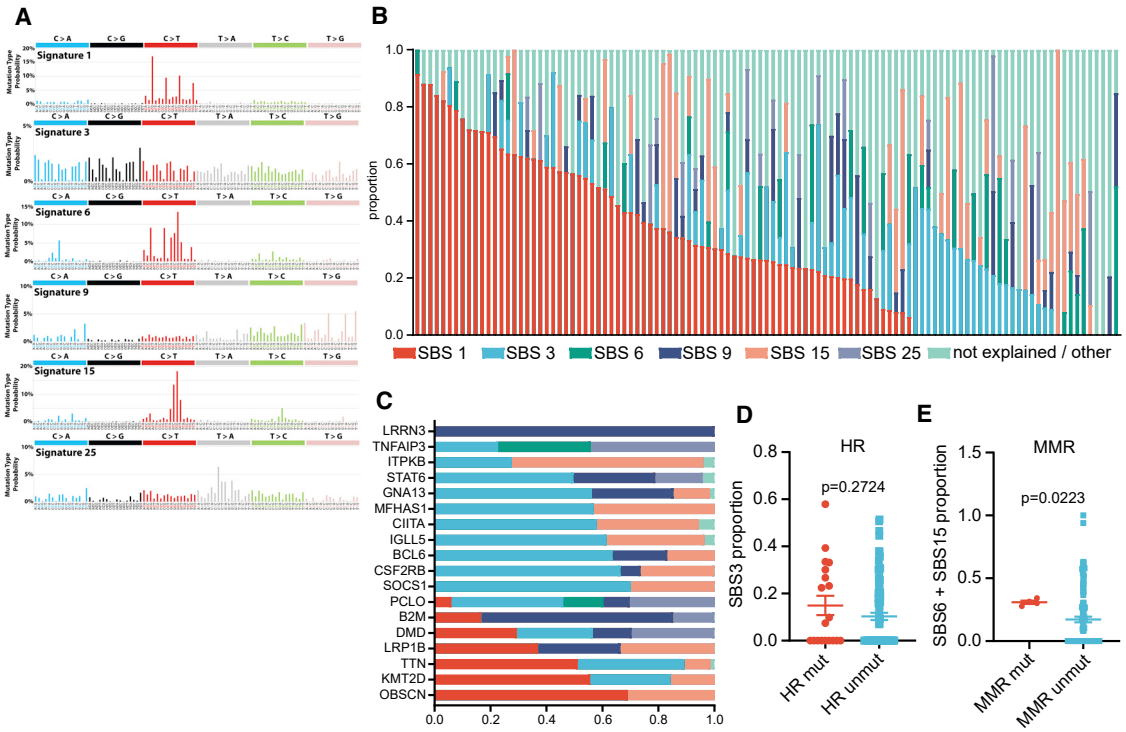


Figure 3. Mutational signatures in Hodgkin's lymphoma

(A) Mutational profile of mutational signatures identified in Hodgkin's lymphoma (figure adapted from https://cancer.sanger.ac.uk/signatures/sbs/).

(B) Contribution of identified mutational signatures to single patients sorted by contribution of SBS1 (n = 109).

(C) Contribution of identified mutational signatures to indicated genes.

(D) Comparison of SBS3 contribution between patients with and without mutations in genes involved in homologous recombination.

(E) Comparison of combined SBS6 + SBS15 contribution between patients with and without mutations in genes involved in mismatch repair.

Error bars show SEM. Mann-Whitney test was used for (G) and (H).

lysin. This mutation has been previously described in HL and used for disease response assessment in cfDNA by digital PCR.[52] Functionally, it has been shown that the XPO E571K mutation is a gain-of-function mutation leading to clonal B cell malignancy.[53]

Furthermore, we identified several co-occurring or mutually exclusive OGs and TSGs using DISCOVER[54] (Supplemental methods Figures 3A and 3B). However, co-occurrence or mutual exclusivity analysis did not reveal any clear relationships between pathways beyond the single gene relationships outlined in Supplemental methods Figure 3.

Mutational signatures

Mutational signatures have been identified as characteristic processes of somatic mutation acquisition in cancer, reflecting distinct mechanisms.[55] We identified 6 distinct COSMIC SBS signatures[55] (Figure 3A) using deconstructSigs:[56] SBS1 (deamination of 5-methylcytosine to thymine, associated with aging), SBS3 (defective homologous recombination [HR]-based DNA damage repair), SBS6 and SBS15 (defective DNA mismatch repair [MMR], associated with microsatellite instability), SBS9 (somatic hypermutation in lymphoid cells), and SBS25 (unknown cause, hitherto only identified in HL cell lines)[57] (Figure 3B).

Next, we examined the contribution of these mutational signatures to mutations in recurrently mutated genes in our primary cohort. We identified distinct mutational processes active in different genes. SBS1 was the biggest contributor in DMD, LRP1B, TTN, KMT2D, and OBSCN; SBS3 the biggest contributor in STAT6, GNA13, MFHAS1, CIITA, IGLL4, BCL6, CSF2RB, and SOCS1; SBS15 the biggest contributor in ITPKB; SBS9 the biggest contributor in B2M and LRRN3; and SBS25 the biggest contributor in TNFAIP3 (Figure 3C). Overall, SBS1—the aging signature—was the main contributor in genes that are not OGs or TSGs, suggesting a role of the underlying process in inducing bystander mutations. Interestingly, SBS3—the defective HR signature—majorly contributed to mutations in genes crucial for HL oncogenesis, such as STAT6, CSF2RB, or SOCS1,[11,12] highlighting a potential role for defective HR in HL pathogenesis. SBS9— the somatic hypermutation signature—was the sole contributor to LRRN3 mutations, which has not been empirically shown to be a target of somatic hypermutation but predicted to be one in one analysis.[58] SBS25—a potential HL specific signature—was the main contributor to mutations in TNFAIP3, one of the most frequently mutated genes in HL,[38] supporting the notion that SBS25 is indeed a HL signature and not an artifact.

Correlating mutational signatures with patient phenotypes, we observed that patients with a detection of SBS25 were older (38.1 versus 31.0 years of age; $p = 0.0208$) (Figure S2A), while detection of SBS15 was linked to lower lactate dehydrogenase (LDH) levels (217 versus 253 U/L; $p = 0.0441$) (Figure S2B) and less frequent erythrocyte sedimentation rate (ESR) elevation (OR = 0.32; $p = 0.0098$) (Figure S2C), suggesting that SBS15 could be associated with a less inflammatory and lower tumor burden or tumor cell turnover phenotype. However, uncertainty about the clinical and biological significance of these associations remains, and further studies are needed to confirm these associations.

Next, we examined if mutations in genes involved in HR or MMR (Data 7 in the accompanying Mendeley dataset[35]) were associated with SBS3 (defective HR) or SBS6 and SBS15 (defective MMR), respectively. The contribution of SBS3 was higher in patients with mutations in genes involved in HR, although this was not significant (Figure 3D). The combined contribution of SBS6 and SBS15 was significantly higher in patients with mutations in genes involved in MMR (Figure 3E). These results suggest that subsets of HL exist with defects in HR, MMR, or both and that the respective mutational process contributes fittingly to the mutations observed in these patient subsets.

Somatic CNVs

We used cnvkit[59] and Gistic 2.0[60] to identify focal and arm-level recurrent somatic CNVs (sCNVs). Overall, we identified 16 recurrent gains and 34 losses ($q < 0.25$) (Figures 4A and 4B; Figure S3A). The most common recurrent gains were 19p13.2 (65%); 2q31.2 (62%); 12q13.3 (STAT2, STAT6) (62%); 2p16.1 (REL) (61%); and 9p24.1 (JAK2, CD274) (55%). We also observed frequent gains of regions including NFKB1, MLL5, and PIK3CG (Figure 4C). The most common recurrent losses were 6q23.3 (TNFAIP3, ECT2L) (61%); 9q13 (59%); 13q32.3 (59%); 6q22.31 (58%); and 15q15.1 (57%). Of note, we also observed frequent losses of regions including B2M, HIST1H1E, HIST1H1B, and the HLA locus, underscoring the functional relevance of histone genes and immune escape in HL (Figure 4D). A separate, arm-level analysis using Gistic 2.0,[60] was consistent with these findings, identifying 2p, 2q, 5p, 9p, 10p, 12p, and 12q as recurrent arm-level gains and, among others, 4q, 6p, 6q, 7q, 11q, 17p, 18p, 18q, and 22q as recurrent arm-level gains. Both recurrent arm-level gains and losses were observed on 17q, 19p, and 19q. (Figure S3B).

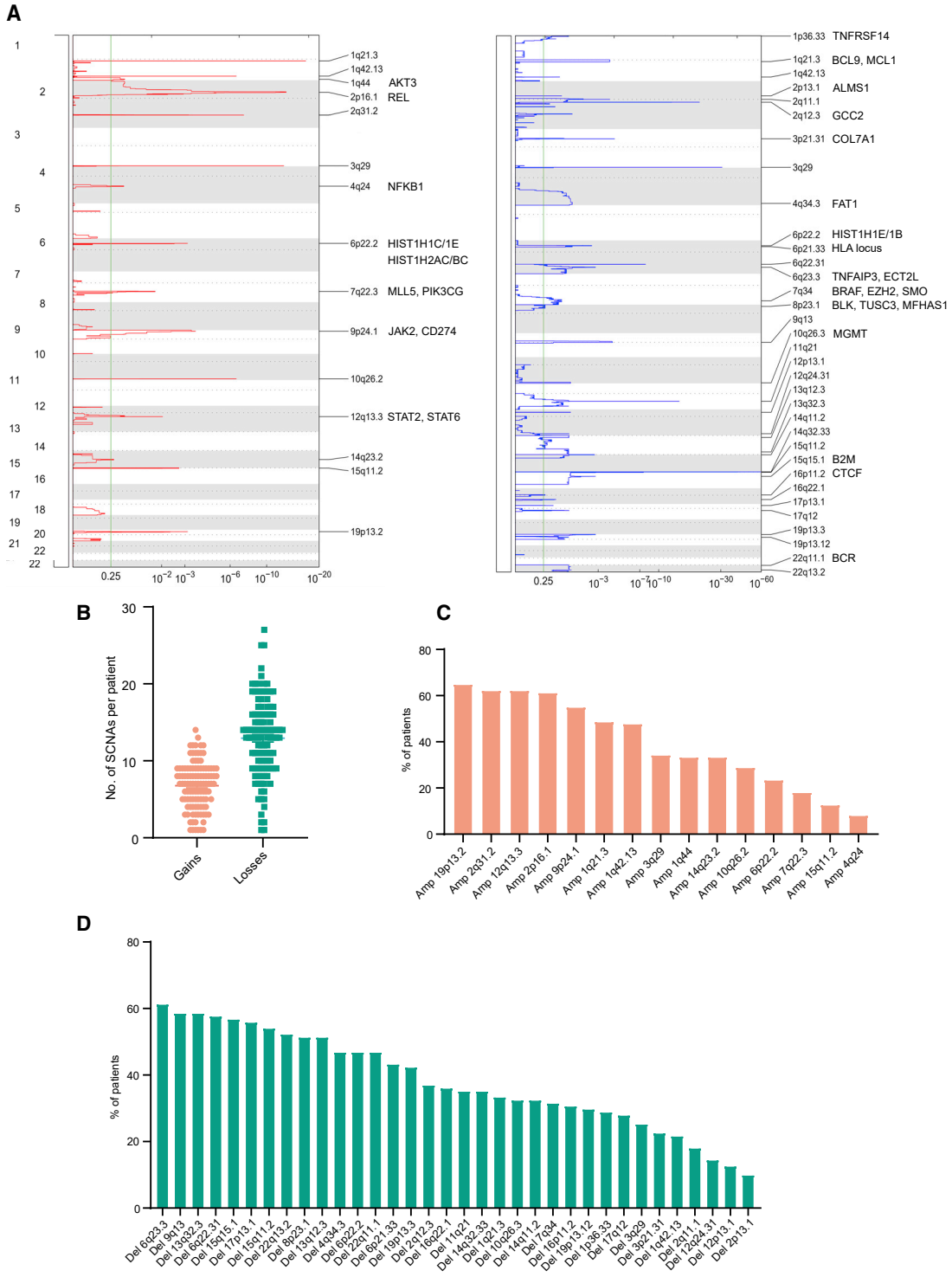


Figure 4. Copy number variations in Hodgkin's lymphoma

(A) Recurrent copy number gains (red) and losses (blue) in patient cohort (n = 111). The x axis shows the chromosomes, the y axis shows the GISTIC q values. Each peak indicates a recurrent copy number gain or loss with the width of the peak corresponding to the size of the recurrent region and the height to the q value. Selected genes within recurrent regions are annotated.

(B) Distribution of number of detected copy number gains and losses in patient cohort (n = 111).

(C) Most frequently detected recurrent copy number gains in patient cohort (n = 111).

(D) Most frequently detected copy number losses in patient cohort (n = 111).

Detection of viral and bacterial associations

We used a previously developed pipeline[61] built around KRAKEN[62] to identify viral and bacterial associations with HL by liquid biopsy. We detected Epstein-Barr virus (EBV) and human herpesvirus 6B (HHV6B), both known to be associated with HL,[63] in 17 (15.3%) and 7 (6.3%) patients, respectively (Figures S4A and S4B). Validating our findings, we found that EBV DNA in plasma as measured by our pipeline corresponded very well to EBV status of the tumor cells assessed by LMP1 staining of infiltrated HL lymph nodes (Figure S4C). Furthermore, we confirmed the known association of EBV+ HL with the mixed cellularity subtype,[64] as 41.7% (5/12) of patients with available histological subtype information and EBV association were of the mixed cellularity subtype compared to only 11.4% (9/79) of those who were not EBV-associated. HL cases associated with EBV were older (37.2 versus 31.1 years of age; $p = 0.0414$) (Figure S4D) and less likely to have mutations in genes involved in NF-kappa B signaling (OR = 0.35; $p = 0.0696$) (Figure S4E), in line with previous reports describing histological EBV detection associated with older HL patients[65] and fewer mutations in TNFAIP3.[38] Also, our observation may be related to the notion that EBV-infected B cells rely on NFkB for survival.[66] In contrast to previous reports,[11,14,] we detected no association between mutational burden (MB) and EBV in our primary cohort (6.0 versus 7.3 mutations/Megabase (Mb); EBV+ versus EBV–; $p = 0.3418$).

HL cases associated with HHV6B were older (42.2 versus 31.3 years of age; $p = 0.0144$) (Figure S4F), more likely to have extranodal involvement (OR = 5.17; $p = 0.0443$) (Figure S4G), and had a higher proportion of SBS6 detection (OR = 6.22; $p = 0.0105$) (Figure S4H). The identified associations of HHV6B with a distinct HL phenotype have not been described previously. Of note, EBV and HHV6B were not detected in any of the healthy control or matched germline samples (Data 8 in the accompanying Mendeley dataset[35]).

We also detected a number of bacterial taxa in the plasma of HL patients after extensive filtering to exclude false positives (see STAR Methods for details): *Streptococcus mitis*, *Burkholderia cepacia*, *Dolosigranulum pigrum*, *Mycobacteroides salmoniphilum*, *Ruminococcus gnavus*, and *Neisseria mucosa* (Data 9 in the accompanying Mendeley dataset;[35] Figure S4I). It is possible that the detection of some species, such as *Streptococcus mitis* as a common colonizer of humans, was increased in cfDNA due to HL-mediated immunosuppression.[67] Of note, we also detected *Moraxella catarrhalis* in a single plasma sample with no detection in the matched germline sample (Data 9 in the accompanying Mendeley dataset[35]). An uncontrolled IgD+ B cell response to *Moraxella catarrhalis* infection has recently been described as a possible early oncogenic driver in nodular lymphocyte-predominant HL.[68] However, the case in our series was a typical case of classical HL of the nodular sclerosis subtype. It was CD30+, CD15+, PAX5+, CD20–, and IgD–. Thus, closer examination of this particular case suggests either an accidental finding of *Moraxella catarrhalis* DNA or a different oncogenic mechanism at play.

HL phenotypes are characterized by distinct genotypes

Our primary cohort included 40 patients with a large mediastinal mass. These patients were characterized by a higher frequency of B2M (OR = 4.39; $p = 0.0009$) (Figure S5A) and IKBKB (OR = 3.78; $p = 0.0337$) (Figure S5B) mutations, as well as more mutations in genes involved in interferon signaling (OR = 3.05; $p = 0.0095$) (Figure S5C) and JAK-STAT signaling (OR = 2.43; $p = 0.0577$) (Figure S5D). These alterations have a high overlap with those recently reported in primary mediastinal B cell

lymphoma (PMBL)[69] and gray zone lymphoma with an anterior mediastinal mass.[70] Taken together with the notion that a disease continuum stretching from mediastinal HL via gray zone lymphoma to PMBL exists,[71] these findings are suggestive of a biological subgroup of HL with a proclivity to mediastinal growth biologically and phenotypically close to PMBL. Of note, out of 32 patients with a large mediastinal mass and histological subtyping available, 28 (87.5%) were of the nodular sclerosis subtype.

Neither mutation frequency of any of the identified OGs or TSGs, frequency of any of the identified recurrent copy number variations, oncogenic pathways, mutational signatures nor MB were associated with the nodular sclerosis or mixed cellularity subtype.

Probing age-genotype associations, we found EZH2 mutations to be associated with younger (21.8 versus 32.4 years; $p < 0.0001$) (Figure S5E) and IRF4 mutations to be associated with older patients (44.8 versus 31.3 years; $p = 0.0488$) (Figure S5F).

Furthermore, patients with any mutations in genes involved in transcriptional regulation by p53 (267 versus 226 U/L; $p = 0.0111$) (Figure S5G) had higher baseline LDH levels, suggesting higher tumor burden or tumor cell turnover in these patients.

Association between HL genotypes and treatment response

Biological predictors of response are currently not established in HL. Therefore, we examined associations of somatic mutations and sCNVs with early PET response after 2 treatment cycles as a surrogate endpoint for patients with ultimately worse outcomes.[19,21,72] For this analysis, we pooled patients from the primary and secondary cohort. Mutations in IRF8 (OR = 3.98; $p = 0.0523$) (Figure S6A); PIM1 (OR = 6.06; $p = 0.0198$) (Figure S6B); TP53 (OR = 5.95; $p = 0.0064$) (Figure S6C); and NOTCH1 (OR = 4.71; $p = 0.0586$) (Figure S6D) were associated with inferior response. Whereas copy number gains peaking at 12q13.3 (STAT2, STAT6) (OR = 0.40; $p = 0.0202$) (Figure S6E) and losses peaking at 7q34 (BRCA, EZH2, SMO) (OR = 0.34; $p = 0.0153$) (Figure S6F) and 10q26.3 (MGMT) (OR = 0.25; $p = 0.0061$) (Figure S6G) were associated with superior response. Both BRCA loss[73] and MGMT loss[74] might predispose to increased sensitivity to DNA-damaging agents, which were used as part of a patient's treatment.

We did not observe a relationship between baseline cfDNA concentration in plasma and treatment response assessed by PET (Deauville 1–3 versus Deauville 4–5: 13.4 versus 18.3 ng/mL; $p = 0.5470$) (Figure S6H).

Aiming to confirm the relevance of the identified effects of certain genotypes on PET response after 2 cycles of treatment, we evaluated if the identified high-risk mutations in IRF8, PIM1, TP53, or NOTCH1 were associated with worse or low risk copy number gains peaking at 12q13.3 or losses peaking at 7q34 and 10q26.3 with better PFS. We could only confirm a significant negative effect of TP53 mutations on PFS ($p = 0.0038$) (Figure S6I); however, with only 6 PFS events in the whole cohort, our ability to detect PFS differences was limited.

MB was highly heterogeneous (Figures S7A and S7B) and did not correlate with cfDNA concentration in plasma, excluding a technical bias where high ctDNA amounts allowed for the detection of more variants (Figure S7C). B2M mutations (10.0 versus 5.9 mutations/Mb; $p = 0.0002$) (Figure S7D), HLA locus losses (8.6

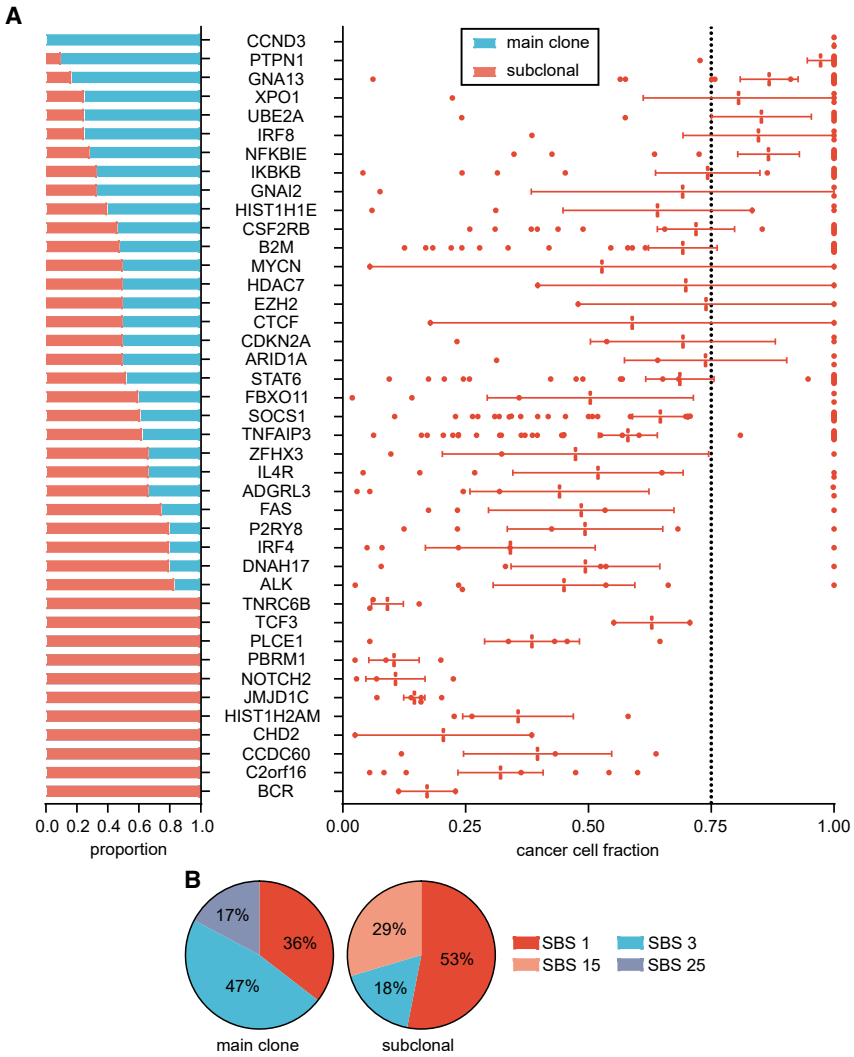


Figure 5. Clonal structure of Hodgkin's lymphoma

(A) Clonal structure is shown for all oncogenes and tumor suppressor genes (n = 42) including all patients where clonal structure could be derived (n = 51). The left chart shows percentage of patients with mutations in the indicated gene where at least one of the mutations in the indicated gene is in the main clone (cancer cell fraction [CCF] > 0.75). The right chart shows detailed CCFs. Each dot represents the CCF in the indicated gene in one patient where clonal structure could be assessed. The mean CCF for each gene is also shown.

(B) Contribution of single base substitution signatures to mutations in genes mainly mutated in main clones or subclones.

Error bars show SEM.

versus 5.9 mutations/Mb; p = 0.0103) (Figure S7E), and 9p24.1 gains (8.1 versus 5.9 mutations/Mb; p = 0.0343) (Figure S7F) were associated with higher MB, suggesting that high MB HL were selected for these immune escape variants during development. In line with this, any mutation in MHC class I antigen presentation genes was associated with higher MB (9.4 versus 4.1 mutations/Mb; p < 0.0001) (Figure S7G. These results suggest that a subset of HL exists that has a higher MB, and therefore, immune escape variants had a selective advantage in its evolution and became dominant.

Clonal structure of HL

ctDNA sequencing can be used to deconvolute the spatial and temporal clonal structure of cancer. We used PyClone[75] to differentiate between main clone mutations likely present in all HL cells and subclonal mutations likely only present in a fraction. Some OGs and TSGs mainly harbored main clone mutations (e.g., GNA13, XPO1, NFKBIE, IKBKB, CSF2RB, B2M) while some harbored mainly subclonal mutations (e.g., PRBM1, NOTCH2, CHD2, BCR) (Figure 5A). Some important HL OGs and TSGs (STAT6, SOCS1, TNFAIP3) harbored a mixture of main clone and subclonal mutations (Figure 5A). Clonal structure allows inferring timing of mutations with main clone mutations likely occurring earlier in oncogenesis. Thus, the diversity in observed clonal structure suggests that mutational processes in HL oncogenesis are likely ongoing without a clear temporal order of genes important in early versus late oncogenesis.

To disentangle the contribution of mutational processes to early versus late HL oncogenesis, we compared mutational signatures in subclonal versus main clone mutations. SBS3 (defective HR) contributed more to main clone mutations compared to subclonal ones, whereas SBS15 (defective MMR) contributed more to subclonal mutations (Figure 5B). This suggests that defective HR plays a bigger role in early HL oncogenesis, whereas defective DNA MMR contributes more to ongoing mutational processes after initial malignant transformation.

Minimal residual disease

Highly sensitive methods for response assessment that are predictive of treatment outcome are lacking in HL. Therefore, we used our ctDNA sequencing platform to assess MRD using samples collected after treatment initiation in a subset of patients from the primary and secondary cohorts (Figure 6A) by following individual mutational fingerprints over time.[19–21] First, we examined the relationship between MRD levels and PET-based response assessments at the same time point. MRD response was correlated with PET-based response assessment and grouped by Deauville score (Figure 6B). Next, we hypothesized that MRD, measured early during treatment, could predict further treatment response. Indeed, our data show that all patients with undetectable MRD after 2 cycles had a PET-negative response at the end of treatment, whereas all patients who did not achieve PET-negativity at the end of treatment had a less than complete MRD response after 2 cycles of treatment (Figure 6C). Examining an even earlier time point during treatment, we found that MRD assessed after only 1 week of treatment was highly predictive of PET-response after 2 cycles of treatment (Figure 6D). We had only a limited number of PFS events among patients with MRD measurement. Nevertheless, MRD after 2 cycles of treatment appeared predictive of PFS (Figure 6E). Combining PET-based and MRD response assessments after 2 cycles of treatment appeared to be of value, as no patient who was PET- and MRD-negative after 2 cycles of treatment had a PFS event (Figure 6F). To showcase clinical use cases for MRD monitoring with our assay, we finally highlight two patients illustrating that MRD assessment may function as an auxiliary tool complementing PET-based response assessment in more complex cases (Supplemental methods; Figure 4).

DISCUSSION

Here, we present the largest genomic landscape of HL to date and introduce our targeted ctDNA sequencing platform that allows for accurate genotyping using pretreatment samples and highly sensitive MRD detection in samples collected during and after treatment. While applied to HL, this study serves as a blueprint for other cancers with similar therapeutic challenges.

A

Cohort	1week	2cycles	EOT	FU	TOTAL
NIVAHL	6	18	11	-	**35**
HD21	-	7	1	-	**8**
Amsterdam	6	10	10	8	**34**
TOTAL	**12**	**35**	**22**	**8**	**77**

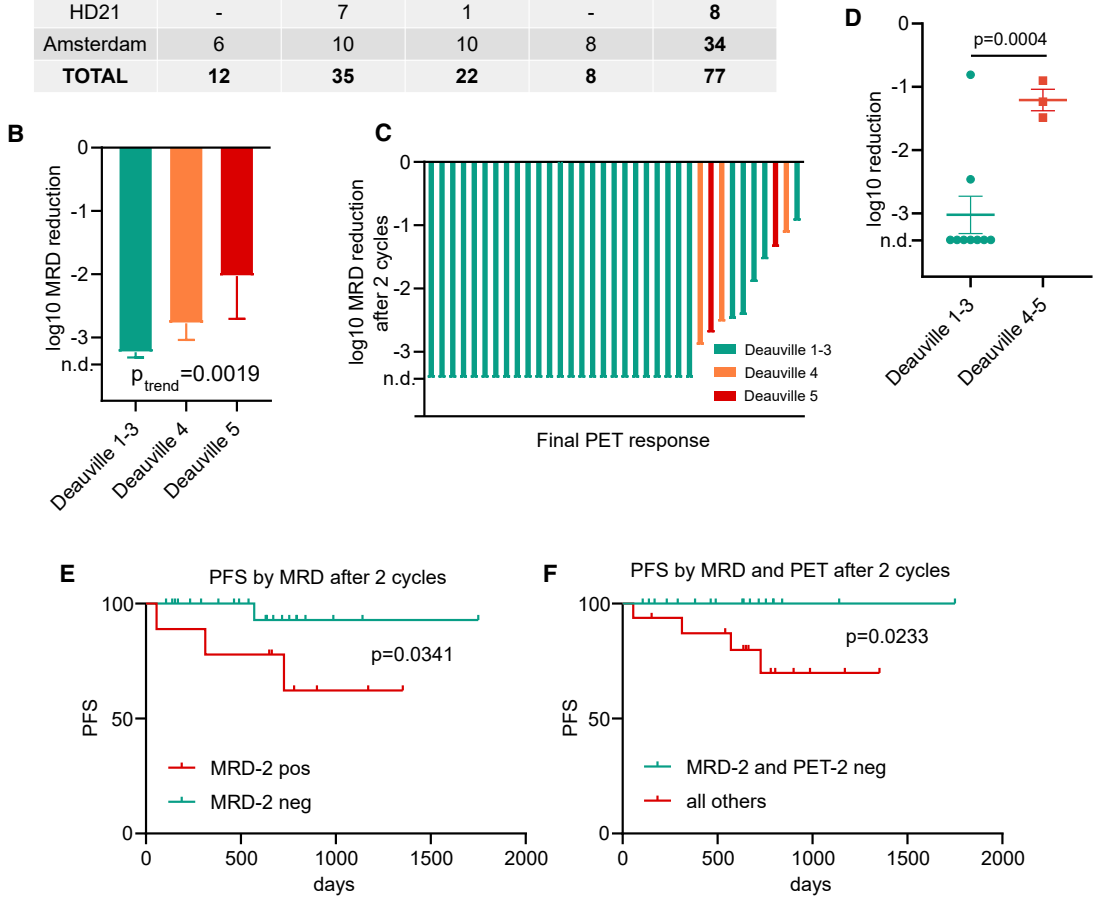


Figure 6. Detection of MRD

(A) Overview of during treatment and follow up samples used for MRD assessment.

(B) Correlation of MRD and corresponding PET response (n = 53 time points).

(C) Correlation of MRD after 2 cycles and PET-response at the end of treatment (n = 35 patients).

(D) Correlation of early MRD after 1 week and PET-response after 2 cycles (n = 12 patients).

(E) Kaplan-Meier analysis of progression-free survival (PFS) by MRD status after 2 cycles.

(F) Kaplan-Meier analysis of PFS by combined PET- and MRD-based response assessment after 2 cycles.

Error bars show SEM. ANOVA with trend test was used for (B), t test with Welch correction for (D), and log rank test for (E) and (F). EOT, end of chemotherapy treatment (before or after consolidative radiotherapy); FU, follow-up beyond last active treatment.

The low tumor DNA content in bulk HL tissue, due to the scarcity of HRS cells in primary lymph node biopsies,[13] has resulted in comparatively few genomic characterization studies in HL. Available knowledge is obtained from small studies with limited clinical data that relied on cumbersome sorting methods to enrich for HL-derived DNA. The three major studies in the field included only 10, 23, and 34 patients respectively.[11,12,] Reichel et al.[12] and Wienand et al.[14] used flow cytometry sorting to separate malignant cells from non-malignant cells in primary lymph node biopsies. Tiacci et al.[11] used laser microdissection to achieve the same. Recently, two studies introduced ctDNA sequencing in HL as a tool to genotype adult or pediatric HL.[25,26] Compared to both studies, our approach has several advantages. First, our target region is approximately 10× larger than the one used in both studies

while, at the same time, sequencing is performed at comparable depth. Second, our optimized bioinformatics process for error reduction and usage of UMIs results in highly reduced error rates, facilitating more accurate genotyping and an order of magnitude more sensitive MRD detection. Together, both advantages result in, on average, more mutations being detected per patient with high accuracy. Third, our bioinformatics platform allows for comprehensive genotyping including copy number variations and, if within the target region, structural variants. These advantages result in a higher proportion of patients with variants and a higher number of variants detected per patient. Therefore, we are convinced that our platform has the potential to become the state-of-the-art ctDNA pipeline in HL and, in addition, has wide applicability beyond HL, only requiring modifications of the target region. For example, it has recently been shown that improved sequencing protocols and bioinformatics can facilitate highly accurate genotyping in lung cancer[29] or diffuse large B cell lymphoma.[30] Importantly, our patient samples were collected in ancillary studies accompanying clinical trials across many treatment centers, showing that decentralized sample collection and centralized processing and analysis of liquid biopsies is feasible in hematological malignancies.

In our study, we made several novel findings with relevance to HL pathogenesis. First, we identify novel recurrent mutations in HL, such as Y641 EZH2 mutations and mutations in IRF4 and IRF8. Interestingly, Y641 EZH2 mutations have recently been shown to be able to modulate the germinal center niche toward lymphomagenesis by reducing survival dependency of germinal center B cells on T cell help and driving expansion of germinal center centrocytes.[76] Second, we disentangle the contribution of different mutational processes in HL pathogenesis at high resolution, suggesting an important role of HR and offering first evidence that COSMIC SBS25 might indeed be a HL signature observable in patients. Third, we link, for the first time, HL genotypes to phenotypes and show that mutations in some genes, such as TP53 or NOTCH1, are associated with a high-risk phenotype, while other alterations, such as BRCA1 or MGMT loss, are associated with a low-risk phenotype, possibly mediated by increased sensitivity to commonly used cytotoxic drugs. Fourth, we identify large mediastinal mass HL, which is mostly of the nodular sclerosis subtype, as a HL subgroup genetically more closely linked to PMBL than other classical HLs. Such a similarity of mediastinal HL of the nodular sclerosis subtype and PMBL has been proposed before based on histopathological similarity.[71,77] Fifth, we classify HL as a disease with high heterogeneity of MB and link high MB HL with an immune escape genotype. It is possible that these patients will respond differently to immune checkpoint inhibition.

In a different use of our ctDNA sequencing platform, we show that it can be used for extremely sensitive MRD detection. MRD levels correlate with PET-based response assessments in our study and can predict treatment response as early as 1 week after treatment initiation, as well as predict relapse in PET-negative cases. Furthermore, MRD increases also precede relapse detected by imaging. Interestingly, MRD detection after 2 cycles of treatment is associated with worse PFS. Additionally, patients who are MRD- and PET-negative after 2 cycles have an excellent prognosis with no relapses in our group. However, it is important to consider that we only had very few PFS events in our study; therefore, bigger studies are needed to confirm these findings. Linking MRD detection and overall survival (OS) was not possible in our study because of a lack of OS events. Other disease biomarkers, such as thymus and activation regulated chemokine (TARC)[78] or exosomal miRNA,[79] have been evaluated in HL with some success, however patient numbers in these studies are low and preclude definitive conclusions. TARC has also been evaluated

in conjunction with PET-CT imaging, suggesting synergy.[80] ctDNA has been evaluated as a response biomarker in HL as well, for example, in a study that looked at normalization of sCNV signatures in plasma.[24] However, it is likely that our approach is more sensitive as shown in the technical validation presented here. It is unlikely that other available methods for MRD detection in HL can detect a remaining plasma tumor burden of 0.00714% (less than 1 in 14,000 DNA molecules), although a head-to-head comparison is currently missing. In addition to synergy between MRD evaluation based on ctDNA sequencing and PET-CT, there might also be synergy between MRD evaluation based on ctDNA sequencing and protein-based biomarkers such as TARC in HL. The idea of combining protein-based and DNA-based biomarkers has been tested in other cancers with remarkable success,[81] and similar studies in HL are needed. Future enhancements of our assay could be the inclusion of *real* or *in silico* cfDNA size selection,[82] phased variant enrichment,[83] or duplex sequencing.[84]

In summary, we present the largest comprehensive genomic profiling study of HL to date by utilizing our novel ctDNA sequencing platform and show that it can be used to accurately genotype HL at baseline and measure MRD with extremely high sensitivity. In an improvement on previous efforts, we use a well-annotated cohort of patients mainly treated within clinical trials to present an integrated platform that allows straightforward and comprehensive genotyping of HL at diagnosis and highly sensitive detection of MRD in almost all HL patients (>98%) by just requiring a blood sample. Our results pave the way for prospective clinical trials to evaluate prognostic genomic profiling by liquid biopsy and MRD-guided treatment in HL and how it can enhance current PET-guided treatment strategies. Such trials have to be carefully designed, considering various settings such as early- or advanced-stage HL or the given chemotherapy regimen among other factors, to answer how exactly such novel risk and response biomarkers can be integrated into current treatment concepts.

Limitations

Our study has several limitations. One is that we only had 6 PFS events in all patients so far, limiting our ability to detect PFS differences between genotypes. More follow up and larger cohorts will be needed in future studies to evaluate if the observed genotype outcome associations based on the PET-2 surrogate endpoint will be confirmed to be relevant for PFS. Another limitation is that absolute copy number inference is not reliably possible using ctDNA. Therefore, our analysis of clonal structure is also somewhat limited. Furthermore, while most patient characteristics such as age, gender, stage, or histological subtype were distributed as expected for a HL cohort and relationships between genotypes and patient characteristics could be analyzed, we were not able to account for differences in ethnicity, as most patients in our cohort were either white or ethnicity was not recorded.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - This paper does not report original code.
  - Data and code availability

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.medj.2021.09.002.

ACKNOWLEDGMENTS

We would like to thank Elisabeth Kirst for excellent technical assistance, the staff at the trial sites of the included clinical studies for supporting this ancillary study and finally all patients, without whom this study would not have been possible. S.S. was supported by a Mildred Scheel School of Oncology Aachen-Bonn-Cologne-Düsseldorf MD Research Stipend. Sequencing infrastructure was supported by DFG Research Infrastructure as part of the Next Generation Sequencing Competence Network [423957469], and next-generation sequencing (NGS) analyses were carried out at the production site WGGC Cologne. The research was partly funded by grant 70112502 from Deutsche Krebshilfe, grant EN 179/13-1 from the Deutsche Forschungsgemeinschaft (DFG), and a grant by the Frau-Weiskam und Christel Ruranski-Stiftung. It was also supported by the Hodgkin Lymphoma MRD consortium.

AUTHOR CONTRIBUTIONS

Conceptualization, S.B. and B.v.T..; methodology, S.S., L.M., M.C., J.A., E.G.H., A.R., D.M.P., and A.H.; investigation, all authors.; formal analysis, S.S., L.M., and S.B.; writing – original draft, S.S., L.M., and S.B.; writing – review & editing, all authors; funding acquisition, P.J.B., S.S., W.K., D.M.P., A.H., J.M.Z., P.N., A.E., B.v.T., and S.B.; supervision, S.B.

DECLARATION OF INTERESTS

P.J.B. reports research grants from BeiGene, Bristol Myers Squibb, Merck Sharpe & Dohme, and Takeda and personal fees and non-financial support from Bristol-Myers Squibb, Celgene, and Takeda, all outside the submitted work. S.S. received travel grants from GSK. D.M.P. reports being founder and CSO of Exbiome and an

occasional advisor for Takeda. B.v.T. reports personal fees and nonfinancial support from Bristol-Myers Squibb; personal fees from Amgen, Pfizer, Gilead Sciences, Pentixapharm, and Roche; grants, personal fees, and nonfinancial support from MSD and Takeda; and grants, personal fees, and nonfinancial support from Novartis. S.B. reports being founder, CEO, and shareholder of Liqomics and personal fees and non-financial support from Bristol-Myers Squibb and Takeda outside the submitted work.

Received: April 14, 2021
Revised: July 12, 2021
Accepted: September 15, 2021
Published: October 8, 2021

REFERENCES

1. Borchmann, S., von Tresckow, B., and Engert, A. (2016). Current developments in the treatment of early-stage classical Hodgkin lymphoma. Curr. Opin. Oncol. *28*, 377–383.

2. Küppers, R., Engert, A., and Hansmann, M.L. (2012). Hodgkin lymphoma. J. Clin. Invest. *122*, 3439–3447.

3. Kreissl, S., Mueller, H., Goergen, H., Mayer, A., Brillant, C., Behringer, K., Halbsguth, T.V., Hitz, F., Soekler, M., Shonukan, O., et al.; German Hodgkin Study Group (2016). Cancer-related fatigue in patients with and survivors of Hodgkin's lymphoma: a longitudinal study of the German Hodgkin Study Group. Lancet Oncol. *17*, 1453–1462.

4. Eichenauer, D.A., Thielen, I., Haverkamp, H., Franklin, J., Behringer, K., Halbsguth, T., Klimm, B., Diehl, V., Sasse, S., Rothe, A., et al. (2014). Therapy-related acute myeloid leukemia and myelodysplastic syndromes in patients with Hodgkin lymphoma: a report from the German Hodgkin Study Group. Blood *123*, 1658–1664.

5. Borchmann, S., Müller, H., Haverkamp, H., Baues, C., Marková, J., Hüttmann, A., Glunz, A., Fuchs, M., Borchmann, P., and Engert, A. (2019). Symptomatic osteonecrosis as a treatment complication in Hodgkin lymphoma: an analysis of the German Hodgkin Study Group (GHSG). Leukemia *33*, 439–446.

6. Borchmann, S., Müller, H., Hude, I., Fuchs, M., Borchmann, P., and Engert, A. (2019). Thrombosis as a treatment complication in Hodgkin lymphoma patients: a comprehensive analysis of three prospective randomized German Hodgkin Study Group (GHSG) trials. Ann. Oncol. *30*, 1329–1334.

7. Behringer, K., Mueller, H., Goergen, H., Thielen, I., Eibl, A.D., Stumpf, V., Wessels, C., Wiehlpütz, M., Rosenbrock, J., Halbsguth, T., et al. (2013). Gonadal function and fertility in survivors after Hodgkin lymphoma treatment within the German Hodgkin Study Group HD13 to HD15 trials. J. Clin. Oncol. *31*, 231–239.

8. Borchmann, P., Eichenauer, D.A., and Engert, A. (2012). State of the art in the treatment of Hodgkin lymphoma. Nat. Rev. Clin. Oncol. *9*, 450–459.

9. Hasenclever, D., and Diehl, V. (1998). A prognostic score for advanced Hodgkin's disease. International Prognostic Factors Project on Advanced Hodgkin's Disease. N. Engl. J. Med. *339*, 1506–1514.

10. Cirillo, M., Reinke, S., Klapper, W., and Borchmann, S. (2019). The translational science of hodgkin lymphoma. Br. J. Haematol. *184*, 30–44.

11. Tiacci, E., Ladewig, E., Schiavoni, G., Penson, A., Fortini, E., Pettirossi, V., Wang, Y., Rosseto, A., Venanzi, A., Vlasevska, S., et al. (2018). Pervasive mutations of JAK-STAT pathway genes in classical Hodgkin lymphoma. Blood *131*, 2454–2465.

12. Reichel, J., Chadburn, A., Rubinstein, P.G., Giulino-Roth, L., Tam, W., Liu, Y., Gaiolla, R., Eng, K., Brody, J., Inghirami, G., et al. (2015). Flow sorting and exome sequencing reveal the oncogenome of primary Hodgkin and Reed-Sternberg cells. Blood *125*, 1061–1072.

13. Borchmann, S., and Engert, A. (2017). The genetics of Hodgkin lymphoma: an overview and clinical implications. Curr. Opin. Oncol. *29*, 307–314.

14. Wienand, K., Chapuy, B., Stewart, C., Dunford, A.J., Wu, D., Kim, J., Kamburov, A., Wood, T.R., Cader, F.Z., Ducar, M.D., et al. (2019). Genomic analyses of flow-sorted Hodgkin Reed-Sternberg cells reveal complementary mechanisms of immune evasion. Blood Adv. *3*, 4065–4080.

15. Radford, J., Illidge, T., Counsell, N., Hancock, B., Pettengell, R., Johnson, P., Wimperis, J., Culligan, D., Popova, B., Smith, P., et al. (2015). Results of a trial of PET-directed therapy for early-stage Hodgkin's lymphoma. N. Engl. J. Med. *372*, 1598–1607.

16. Raemaekers, J.M.M., André, M.P., Federico, M., Girinsky, T., Oumedaly, R., Brusamolino, E., Brice, P., Fermé, C., van der Maazen, R., Gotti, M., et al. (2014). Omitting radiotherapy in early positron emission tomography-negative stage I/II Hodgkin lymphoma is associated with an increased risk of early relapse: Clinical results of the preplanned interim analysis of the randomized EORTC/LYSA/FIL H10 trial. J. Clin. Oncol. *32*, 1188–1194.

17. Fuchs, M., Goergen, H., Kobe, C., Kuhnert, G., Lohri, A., Greil, R., Sasse, S., Topp, M.S., Schäfer, E., Hertenstein, B., et al. (2019). Positron emission tomography-guided treatment in early-stage favorable Hodgkin lymphoma: Final results of the international, randomized phase III HD16 trial by the German Hodgkin Study Group. J. Clin. Oncol. *37*, 2835–2845.

18. Johnson, P., Federico, M., Kirkwood, A., Fosså, A., Berkahn, L., Carella, A., d'Amore, F., Enblad, G., Franceschetto, A., Fulham, M., et al. (2016). Adapted Treatment Guided by Interim PET-CT Scan in Advanced Hodgkin's Lymphoma. N. Engl. J. Med. *374*, 2419–2429.

19. Stephens, D.M., Li, H., Schöder, H., Straus, D.J., Moskowitz, C.H., LeBlanc, M., Rimsza, L.M., Bartlett, N.L., Evens, A.M., LaCasce, A.S., et al. (2019). Five-year follow-up of SWOG S0816: limitations and values of a PET-adapted approach with stage III/IV Hodgkin lymphoma. Blood *134*, 1238–1246.

20. Borchmann, P., Goergen, H., Kobe, C., Lohri, A., Greil, R., Eichenauer, D.A., Zijlstra, J.M., Markova, J., Meissner, J., Feuring-Buske, M., et al. (2017). PET-guided treatment in patients with advanced-stage Hodgkin's lymphoma (HD18): final results of an open-label, international, randomised phase 3 trial by the German Hodgkin Study Group. Lancet *390*, 2790–2802.

21. Gallamini, A., Tarella, C., Viviani, S., Rossi, A., Patti, C., Mulé, A., Picardi, M., Romano, A., Cantonetti, M., La Nasa, G., et al. (2018). Early chemotherapy intensification with escalated beacopp in patients with advanced-stage hodgkin lymphoma with a positive interim positron emission tomography/computed tomography scan after two abvd cycles: Long-term results of the GITIL/FIL HD 0607 trial. J. Clin. Oncol. *36*, 454–462.

22. Borchmann, P., Plütschow, A., Kobe, C., Greil, R., Meissner, J., Topp, M.S., Ostermann, H., Dierlamm, J., Mohm, J., Thiemer, J., et al. (2021). PET-guided omission of radiotherapy in early-stage unfavourable Hodgkin lymphoma (GHSG HD17): a multicentre, open-label, randomised, phase 3 trial. Lancet Oncol. *22*, 223–234.

23. Newman, A.M., Bratman, S.V., To, J., Wynne, J.F., Eclov, N.C., Modlin, L.A., Liu, C.L., Neal, J.W., Wakelee, H.A., Merritt, R.E., et al. (2014). An ultrasensitive method for quantitating

circulating tumor DNA with broad patient coverage. Nat. Med. *20*, 548–554.

24. Vandenberghe, P., Wlodarska, I., Tousseyn, T., Dehaspe, L., Dierickx, D., Verheecke, M., Uyttebroeck, A., Bechter, O., Delforge, M., Vandecaveye, V., et al. (2015). Non-invasive detection of genomic imbalances in Hodgkin/Reed-Sternberg cells in early and advanced stage Hodgkin's lymphoma by sequencing of circulating cell-free DNA: a technical proof-of-principle study. Lancet Haematol. *2*, e55–e65.

25. Spina, V., Bruscaggin, A., Cuccaro, A., Martini, M., Di Trani, M., Forestieri, G., Manzoni, M., Condoluci, A., Arribas, A., Terzi-Di-Bergamo, L., et al. (2018). Circulating tumor DNA reveals genetics, clonal evolution and residual disease in classical Hodgkin lymphoma. Blood *131*, 2413–2425.

26. Desch, A.K., Hartung, K., Botzen, A., Brobeil, A., Rummel, M., Kurch, L., Georgi, T., Jox, T., Bielack, S., Burdach, S., et al. (2020). Genotyping circulating tumor DNA of pediatric Hodgkin lymphoma. Leukemia *34*, 151–166.

27. Hohaus, S., Giachelia, M., Massini, G., Mansueto, G., Vannata, B., Bozzoli, V., Criscuolo, M., D'Alò, F., Martini, M., Larocca, L.M., et al. (2009). Cell-free circulating DNA in Hodgkin's and non-Hodgkin's lymphomas. Ann. Oncol. *20*, 1408–1413.

28. Camus, V., Viennot, M., Lequesne, J., Viailly, P.J., Bohers, E., Bessi, L., Marcq, B., Etancelin, P., Dubois, S., Picquenot, J.M., et al. (2021). Targeted genotyping of circulating tumor DNA for classical Hodgkin lymphoma monitoring: a prospective study. Haematologica *106*, 154–162.

29. Newman, A.M., Lovejoy, A.F., Klass, D.M., Kurtz, D.M., Chabon, J.J., Scherer, F., Stehr, H., Liu, C.L., Bratman, S.V., Say, C., et al. (2016). Integrated digital error suppression for improved detection of circulating tumor DNA. Nat. Biotechnol. *34*, 547–555.

30. Kurtz, D.M., Scherer, F., Jin, M.C., Soo, J., Craig, A.F.M., Esfahani, M.S., Chabon, J.J., Stehr, H., Liu, C.L., Tibshirani, R., et al. (2018). Circulating Tumor DNA Measurements As Early Outcome Predictors in Diffuse Large B-Cell Lymphoma. J. Clin. Oncol. *36*, 2845–2853.

31. Deng, S., Lira, M., Huang, D., Wang, K., Valdez, C., Kinong, J., Rejto, P.A., Bienkowska, J., Hardwick, J., and Xie, T. (2018). TNER: a novel background error suppression method for mutation detection in circulating tumor DNA. BMC Bioinformatics *19*, 387.

32. Zook, J.M., McDaniel, J., Olson, N.D., Wagner, J., Parikh, H., Heaton, H., Irvine, S.A., Trigg, L., Truty, R., McLean, C.Y., et al. (2019). An open resource for accurately benchmarking small variant and reference calls. Nat. Biotechnol. *37*, 561–566.

33. Krusche, P., Trigg, L., Boutros, P.C., Mason, C.E., De La Vega, F.M., Moore, B.L., Gonzalez-Porta, M., Eberle, M.A., Tezak, Z., Lababidi, S., et al.; Global Alliance for Genomics and Health Benchmarking Team (2019). Best practices for benchmarking germline small-variant calls in human genomes. Nat. Biotechnol. *37*, 555–560.

34. Armbruster, D.A., and Pry, T. (2008). Limit of blank, limit of detection and limit of quantitation. Clin. Biochem. Rev. *29* (*Suppl 1*), S49–S52.

35. Sobesky, S., Mammadova, L., Cirillo, M., Drees, E., Mattlener, J., Dörr, H., Altmüller, J., Shi, Z., Bröckelmann, P.J., Weiss, J., et al. (2021). Exhaustive circulating tumor DNA sequencing reveals genomic landscape of Hodgkin lymphoma and facilitates ultrasensitive residual disease detection. medRxiv. https://doi.org/10.1101/2021.03.16.21253679.

36. Lawrence, M.S., Stojanov, P., Polak, P., Kryukov, G.V., Cibulskis, K., Sivachenko, A., Carter, S.L., Stewart, C., Mermel, C.H., Roberts, S.A., et al. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature *499*, 214–218.

37. Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz Jr, L.A., and Kinzler, K.W. (2013). Cancer genome landscapes. Science *339*, 1546–1558.

38. Schmitz, R., Hansmann, M.L., Bohle, V., Martin-Subero, J.I., Hartmann, S., Mechtersheimer, G., Klapper, W., Vater, I., Giefing, M., Gesk, S., et al. (2009). TNFAIP3 (A20) is a tumor suppressor gene in Hodgkin lymphoma and primary mediastinal B cell lymphoma. J. Exp. Med. *206*, 981–989.

39. Meyer, M.J., Lapcevic, R., Romero, A.E., Yoon, M., Das, J., Beltrán, J.F., Mort, M., Stenson, P.D., Cooper, D.N., Paccanaro, A., and Yu, H. (2016). mutation3D: Cancer Gene Prediction Through Atomic Clustering of Coding Variants in the Structural Proteome. Hum. Mutat. *37*, 447–456.

40. Sahasrabuddhe, A.A., Chen, X., Chung, F., Velusamy, T., Lim, M.S., and Elenitoba-Johnson, K.S. (2015). Oncogenic Y641 mutations in EZH2 prevent Jak2/β-TrCP-mediated degradation. Oncogene *34*, 445–454.

41. Morin, R.D., Johnson, N.A., Severson, T.M., Mungall, A.J., An, J., Goya, R., Paul, J.E., Boyle, M., Woolcock, B.W., Kuchenbauer, F., et al. (2010). Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nat. Genet. *42*, 181–185.

42. Morin, R.D., Mendez-Lago, M., Mungall, A.J., Goya, R., Mungall, K.L., Corbett, R.D., Johnson, N.A., Severson, T.M., Chiu, R., Field, M., et al. (2011). Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. Nature *476*, 298–303.

43. Cardinez, C., Miraghazadeh, B., Tanita, K., da Silva, E., Hoshino, A., Okada, S., Chand, R., Asano, T., Tsumura, M., Yoshida, K., et al. (2018). Gain-of-function *IKBKB* mutation causes human combined immune deficiency. J. Exp. Med. *215*, 2715–2724.

44. Havelange, V., Pekarsky, Y., Nakamura, T., Palamarchuk, A., Alder, H., Rassenti, L., Kipps, T., and Croce, C.M. (2011). IRF4 mutations in chronic lymphocytic leukemia. Blood *118*, 2827–2829.

45. Cherian, M.A., Olson, S., Sundaramoorthi, H., Cates, K., Cheng, X., Harding, J., Martens, A., Challen, G.A., Tyagi, M., Ratner, L., and Rauch, D. (2018). An activating mutation of interferon regulatory factor 4 (IRF4) in adult T-cell leukemia. J. Biol. Chem. *293*, 6844–6858.

46. Salaverria, I., Philipp, C., Oschlies, I., Kohler, C.W., Kreuz, M., Szczepanowski, M., Burkhardt, B., Trautmann, H., Gesk, S., Andrusiewicz, M., et al.; Molecular Mechanisms in Malignant Lymphomas Network Project of the Deutsche Krebshilfe; German High-Grade Lymphoma Study Group; Berlin-Frankfurt-Münster-NHL trial group (2011). Translocations activating IRF4 identify a subtype of germinal center-derived B-cell lymphoma affecting predominantly children and young adults. Blood *118*, 139–147.

47. Aldinucci, D., Celegato, M., Borghese, C., Colombatti, A., and Carbone, A. (2011). IRF4 silencing inhibits Hodgkin lymphoma cell proliferation, survival and CCL5 secretion. Br. J. Haematol. *152*, 182–190.

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

49. Xu, Y., Jiang, L., Fang, J., Fang, R., Morse, H.C., 3rd, Ouyang, G., and Zhou, J.X. (2015). Loss of IRF8 inhibits the growth of diffuse large B-cell lymphoma. J. Cancer *6*, 953–961.

50. Ritz, O., Guiter, C., Castellano, F., Dorsch, K., Melzner, J., Jais, J.P., Dubois, G., Gaulard, P., Möller, P., and Leroy, K. (2009). Recurrent mutations of the STAT6 DNA binding domain in primary mediastinal B-cell lymphoma. Blood *114*, 1236–1242.

51. Baus, D., Nonnenmacher, F., Jankowski, S., Döring, C., Bräutigam, C., Frank, M., Hansmann, M.L., and Pfitzner, E. (2009). STAT6 and STAT1 are essential antagonistic regulators of cell survival in classical Hodgkin lymphoma cell line. Leukemia *23*, 1885–1893.

52. Camus, V., Stamatoullas, A., Mareschal, S., Viailly, P.J., Sarafan-Vasseur, N., Bohers, E., Dubois, S., Picquenot, J.M., Ruminy, P., Maingonnat, C., et al. (2016). Detection and prognostic value of recurrent exportin 1 mutations in tumor and cell-free circulating DNA of patients with classical Hodgkin lymphoma. Haematologica *101*, 1094–1101.

53. Taylor, J., Sendino, M., Gorelick, A.N., Pastore, A., Chang, M.T., Penson, A.V., Gavrila, E.I., Stewart, C., Melnik, E.M., Herrejon Chavez, F., et al. (2019). Altered nuclear export signal recognition as a driver of oncogenesis. Cancer Discov. *9*, 1452–1467.

54. Canisius, S., Martens, J.W.M., and Wessels, L.F.A. (2016). A novel independence test for somatic alterations in cancer shows that biology drives mutual exclusivity but chance explains most co-occurrence. Genome Biol. *17*, 261.

55. Alexandrov, L.B., Kim, J., Haradhvala, N.J., Huang, M.N., Tian Ng, A.W., Wu, Y., Boot, A., Covington, K.R., Gordenin, D.A., Bergstrom, E.N., et al.; PCAWG Mutational Signatures Working Group; PCAWG Consortium (2020).

The repertoire of mutational signatures in human cancer. Nature *578*, 94–101.

56. Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B.S., and Swanton, C. (2016). DeconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. Genome Biol. *17*, 31.

57. Petljak, M., Alexandrov, L.B., Brammeld, J.S., Price, S., Wedge, D.C., Grossmann, S., Dawson, K.J., Ju, Y.S., Iorio, F., Tubio, J.M.C., et al. (2019). Characterizing Mutational Signatures in Human Cancer Cell Lines Reveals Episodic APOBEC Mutagenesis. Cell *176*, 1282–1294.e20.

58. Duke, J.L., Liu, M., Yaari, G., Khalil, A.M., Tomayko, M.M., Shlomchik, M.J., Schatz, D.G., and Kleinstein, S.H. (2013). Multiple transcription factor binding sites predict AID targeting in non-Ig genes. J. Immunol. *190*, 3878–3888.

59. Talevich, E., Shain, A.H., Botton, T., and Bastian, B.C. (2016). CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. PLoS Comput. Biol. *12*, e1004873.

60. Mermel, C.H., Schumacher, S.E., Hill, B., Meyerson, M.L., Beroukhim, R., and Getz, G. (2011). GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol. *12*, R41.

61. Borchmann, S. (2019). An atlas of bacterial and viral associations in cancer. bioRxiv. https://doi.org/10.1101/773200.

62. Wood, D.E., and Salzberg, S.L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol. *15*, R46.

63. Jarrett, R.F., Hjalgrim, H., and Murray, P.G. (2020). The Role of Viruses in the Genesis of Hodgkin Lymphoma. In Hodgkin Lymphoma: A Comprehensive Overview, A. Engert and A. Younes, eds. (Springer International Publishing), pp. 25–45.

64. Boiocchi, M., De Re, V., Dolcetti, R., Carbone, A., Scarpa, A., and Menestrina, F. (1992). Association of Epstein-Barr virus genome with mixed cellularity and cellular phase nodular sclerosis Hodgkin's disease subtypes. Ann. Oncol. *3*, 307–310.

65. Jarrett, A.F., Armstrong, A.A., and Alexander, E. (1995). Epidemiology of EBV and Hodgkin's lymphoma. Ann. Oncol. *7*, 5–10.

66. Wei, W., Lin, Y., Song, Z., Xiao, W., Chen, L., Yin, J., Zhou, Y., Barta, S.K., Petrus, M., Waldmann, T.A., and Yang, Y. (2020). A20 and RBX1 Regulate Brentuximab Vedotin Sensitivity in Hodgkin Lymphoma Models. Clin. Cancer Res. *26*, 4093–4106.

67. Han, X.Y., Kamana, M., and Rolston, K.V.I. (2006). Viridans streptococci isolated by culture from blood of cancer patients: clinical and microbiologic analysis of 50 cases. J. Clin. Microbiol. *44*, 160–165.

68. Thurner, L., Hartmann, S., Fadle, N., Regitz, E., Kemele, M., Kim, Y.J., Bohle, R.M., Nimmesgern, A., von Müller, L., Kempf, V.A.J., et al. (2020). Lymphocyte predominant cells detect Moraxella catarrhalis-derived antigens in nodular lymphocyte-predominant Hodgkin lymphoma. Nat. Commun. *11*, 2465.

69. Chapuy, B., Stewart, C., Dunford, A.J., Kim, J., Wienand, K., Kamburov, A., Griffin, G.K., Chen, P.H., Lako, A., Redd, R.A., et al. (2019). Genomic analyses of PMBL reveal new drivers and mechanisms of sensitivity to PD-1 blockade. Blood *134*, 2369–2382.

70. Sarkozy, C., Hung, S.S., Chavez, E.A., Duns, G., Takata, K., Chong, L.C., Aoki, T., Jiang, A., Miyata-Takata, T., Telenius, A., et al. (2021). Mutational landscape of gray zone lymphoma. Blood *137*, 1765–1776.

71. Traverse-Glehen, A., Pittaluga, S., Gaulard, P., Sorbara, L., Alonso, M.A., Raffeld, M., and Jaffe, E.S. (2005). Mediastinal gray zone lymphoma: the missing link between classic Hodgkin's lymphoma and mediastinal large B-cell lymphoma. Am. J. Surg. Pathol. *29*, 1411–1421.

72. Adams, H.J.A., Nievelstein, R.A.J., and Kwee, T.C. (2015). Prognostic value of interim FDG-PET in Hodgkin lymphoma: systematic review and meta-analysis. Br. J. Haematol. *170*, 356–366.

73. Lord, C.J., and Ashworth, A. (2016). BRCAness revisited. Nat. Rev. Cancer *16*, 110–120.

74. Hegi, M.E., Diserens, A.C., Gorlia, T., Hamou, M.F., de Tribolet, N., Weller, M., Kros, J.M., Hainfellner, J.A., Mason, W., Mariani, L., et al. (2005). MGMT gene silencing and benefit from temozolomide in glioblastoma. N. Engl. J. Med. *352*, 997–1003.

75. Roth, A., Khattra, J., Yap, D., Wan, A., Laks, E., Biele, J., Ha, G., Aparicio, S., Bouchard-Côté, A., and Shah, S.P. (2014). PyClone: statistical inference of clonal population structure in cancer. Nat. Methods *11*, 396–398.

76. Béguelin, W., Teater, M., Meydan, C., Hoehn, K.B., Phillip, J.M., Soshnev, A.A., Venturutti, L., Rivas, M.A., Calvo-Fernández, M.T., Gutierrez, J., et al. (2020). Mutant EZH2 Induces a Pre-malignant Lymphoma Niche by Reprogramming the Immune Response. Cancer Cell *37*, 655–673.e11.

77. Perrone, T., Frizzera, G., and Rosai, J. (1986). Mediastinal diffuse large-cell lymphoma with sclerosis. A clinicopathologic study of 60 cases. Am. J. Surg. Pathol. *10*, 176–191.

78. Sauer, M., Plütschow, A., Jachimowicz, R.D., Kleefisch, D., Reiners, K.S., Ponader, S., Engert, A., and von Strandmann, E.P. (2013). Baseline serum TARC levels predict therapy outcome in patients with Hodgkin lymphoma. Am. J. Hematol. *88*, 113–115.

79. van Eijndhoven, M.A.J., Zijlstra, J.M., Groenewegen, N.J., Drees, E.E., van Niele, S., Baglio, S.R., Koppers-Lalic, D., van der Voorn, H., Libregts, S.F., Wauben, M.H., et al. (2016). Plasma vesicle miRNAs for therapy response monitoring in Hodgkin lymphoma patients. JCI Insight *1*, e89631.

80. Plattel, W.J., Visser, L., Diepstra, A., Glaudemans, A.W.J.M., Nijland, M., van Meerten, T., Kluin-Nelemans, H.C., van Imhoff, G.W., and van den Berg, A. (2020). Interim thymus and activation regulated chemokine versus interim ^{18}F-fluorodeoxyglucose positron-emission tomography in classical Hodgkin lymphoma response evaluation. Br. J. Haematol. *190*, 40–44.

81. Cohen, J.D., Li, L., Wang, Y., Thoburn, C., Afsari, B., Danilova, L., Douville, C., Javed, A.A., Wong, F., Mattox, A., et al. (2018). Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science *359*, 926–930.

82. Mouliere, F., Piskorz, A.M., Chandrananda, D., Moore, E., Morris, J., Smith, C.G., Goranova, T., Heider, K., Mair, R., Supernat, A., et al. (2017). Selecting Short DNA Fragments In Plasma Improves Detection Of Circulating Tumour DNA. bioRxiv. https://doi.org/10.1101/134437.

83. Kurtz, D.M., et al. (2019). Phased Variant Enrichment for Enhanced Minimal Residual Disease Detection from Cell-Free DNA. Blood *134*, 552.

84. Kennedy, S.R., Schmitt, M.W., Fox, E.J., Kohrn, B.F., Salk, J.J., Ahn, E.H., Prindle, M.J., Kuong, K.J., Shen, J.C., Risques, R.A., and Loeb, L.A. (2014). Detecting ultralow-frequency mutations by Duplex Sequencing. Nat. Protoc. *9*, 2586–2606.

85. Bröckelmann, P.J., Goergen, H., Keller, U., Meissner, J., Ordemann, R., Halbsguth, T.V., Sasse, S., Sökler, M., Kerkhoff, A., Mathas, S., et al. (2020). Efficacy of Nivolumab and AVD in Early-Stage Unfavorable Classic Hodgkin Lymphoma: The Randomized Phase 2 German Hodgkin Study Group NIVAHL Trial. JAMA Oncol. *6*, 872–880.

86. Liu, Y., Abdul Razak, F.R., Terpstra, M., Chan, F.C., Saber, A., Nijland, M., van Imhoff, G., Visser, L., Gascoyne, R., Steidl, C., et al. (2014). The mutational landscape of Hodgkin lymphoma cell lines determined by whole-exome sequencing. Leukemia *28*, 2248–2251.

87. Lohr, J.G., Stojanov, P., Lawrence, M.S., Auclair, D., Chapuy, B., Sougnez, C., Cruz-Gordillo, P., Knoechel, B., Asmann, Y.W., Slager, S.L., et al. (2012). Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. Proc. Natl. Acad. Sci. USA *109*, 3879–3884.

88. Reddy, A., Zhang, J., Davis, N.S., Moffitt, A.B., Love, C.L., Waldrop, A., Leppa, S., Pasanen, A., Meriranta, L., Karjalainen-Lindsberg, M.L., et al. (2017). Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma. Cell *171*, 481–494.e15.

89. Pasqualucci, L., Trifonov, V., Fabbri, G., Ma, J., Rossi, D., Chiarenza, A., Wells, V.A., Grunn, A., Messina, M., Elliot, O., et al. (2011). Analysis of the coding genome of diffuse large B-cell lymphoma. Nat. Genet. *43*, 830–837.

90. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

91. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

92. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T.,

O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al.; Exome Aggregation Consortium (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285–291.

93. Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M., and Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. *29*, 308–311.

94. Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., and Wilson, R.K. (2012). VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. *22*, 568–576.

95. Joos, S., Granzow, M., Holtgreve-Grez, H., Siebert, R., Harder, L., Martín-Subero, J.I., Wolf, J., Adamowicz, M., Barth, T.F., Lichter, P., and Jauch, A. (2003). Hodgkin's lymphoma cell lines are characterized by frequent aberrations on chromosomes 2p and 9p including *REL* and *JAK2*. Int. J. Cancer *103*, 489–495.

96. Roemer, M.G.M., Advani, R.H., Ligon, A.H., Natkunam, Y., Redd, R.A., Homer, H., Connelly, C.F., Sun, H.H., Daadi, S.E., Freeman, G.J., et al. (2016). PD-L1 and PD-L2 Genetic Alterations Define Classical Hodgkin Lymphoma and Predict Outcome. J. Clin. Oncol. *34*, 2690–2697.

97. Tokheim, C.J., Papadopoulos, N., Kinzler, K.W., Vogelstein, B., and Karchin, R. (2016). Evaluating the evaluation of cancer driver genes. Proc. Natl. Acad. Sci. USA *113*, 14330–14335.

98. Sondka, Z., Bamford, S., Cole, C.G., Ward, S.A., Dunham, I., and Forbes, S.A. (2018). The COSMIC Cancer Gene Census: describing genetic dysfunction across all human cancers. Nat. Rev. Cancer *18*, 696–705.

99. R Core Team (2020). R: A Language and environment for statistical computing. R Foundation for Statistical Computing (Vienna, Austria). https://www.R-project.org/.

100. Canty, A., and Ripley, B. (2020). boot: Bootstrap R (S-Plus) Functions. R package version 1.3-28.

101. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. *44* (D1), D457–D462.

102. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al.; The Gene Ontology Consortium (2000). Gene ontology: tool for the unification of biology. Nat. Genet. *25*, 25–29.

103. The Gene Ontology Consortium (2019). The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res. *47* (D1), D330–D338.

104. Fabregat, A., Sidiropoulos, K., Viteri, G., Forner, O., Marin-Garcia, P., Arnau, V., D'Eustachio, P., Stein, L., and Hermjakob, H. (2017). Reactome pathway analysis: a high-performance in-memory approach. BMC Bioinformatics *18*, 142.

105. Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. *47* (W1), W191–W198.

106. Stevens, S.J.C., Vervoort, M.B., van den Brule, A.J., Meenhorst, P.L., Meijer, C.J., and Middeldorp, J.M. (1999). Monitoring of epstein-barr virus DNA load in peripheral blood by quantitative competitive PCR. J. Clin. Microbiol. *37*, 2852–2857.

107. Jaiswal, S., and Ebert, B.L. (2019). Clonal hematopoiesis in human aging and disease. Science *366*, eaan4673.

108. Environmental Protection Agency (2000). Guidance for data quality assessment: Practical methods for data analysis (EPA). https://www.epa.gov/sites/default/files/2015-06/documents/g9-final.pdf.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological Samples		
Patient samples	This study	N/A
Chemicals, Peptides, and Recombinant Proteins		
green 496 dUTP	Enzo Lifesciences	Cat#42831
orange 552 dUTP	Enzo Lifesciences	Cat#42842
aqua 431 dUTP	Enzo Lifesicences	Cat#42853
Critical Commercial Assays		
Sureselect XT HS Target Enrichment Kit	Agilent	Cat#G9702C
Streck cell-free DNA tubes	Streck	Cat#218997
PAXgene Blood ccfDNA tubes	Qiagen	Cat#761115
DNA LoBind tube	Eppendorf	Cat#0030122208
QIAamp Circulating Nuclei Acid Kit	Qiagen	Cat#55114
QIAamp Blood Mini Kit	Qiagen	Cat#51104
GeneRead DNA FFPE Kit	Qiagen	Cat#180134
Deposited Data		
Supplementary Tables containing additional information and raw data	Mendeley Data	Mendeley Data: https://doi.org/10.17632/69my4hnj2y.1
Software and Algorithms		
fgbio (version 1.1.0)	N/A	http://fulcrumgenomics.github.io/fgbio/
Samtools (version 1.7)	N/A	http://www.htslib.org/
TNER (version 1.1)	N/A	https://github.com/ctDNA/TNER
ANNOVAR (no version)	N/A	https://annovar.openbioinformatics.org/en/latest/
Varscan (version 2.4.3)	N/A	http://varscan.sourceforge.net/
CNVKit (version 0.9.6)	N/A	https://github.com/etal/cnvkit
Gistic (version 2.0)	N/A	https://www.genepattern.org/modules/docs/GISTIC_2.0
Mutation 3D (no version)	N/A	http://www.mutation3d.org/
DISCOVER (version 0.9)	N/A	https://ccb.nki.nl/software/discover/
R (version 3.6.2)	N/A	https://www.r-project.org/
deconstructSigs (version 1.9)	N/A	https://github.com/raerose01/deconstructSigs
g:Profiler (no version)	N/A	https://biit.cs.ut.ee/gprofiler/gost
KRAKEN (version 2)	N/A	https://ccb.jhu.edu/software/kraken2/
PyClone (version 0.13.0)	N/A	https://github.com/Roth-Lab/pyclone
Bcl2fastq2 software (version 2.20)	Illumina	https://support.illumina.com/downloads/bcl2fastq-conversion-software-v2-20.html
BWA (version 0.7.17)	N/A	https://github.com/lh3/bwa
AddUMIsToBam (no version)	N/A	https://github.com/mbusby/AddUMIsToBam

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Sven Borchmann (sven.borchmann@uk-koeln.de).

Materials availability

This study did not generate new unique reagents.

This paper does not report original code.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Data and code availability

Primary sequencing data cannot be deposited publicly due to legal requirements and patient data protection. Upon entering a collaboration agreement and in consultation with the local ethics committee access to primary sequencing data might be granted. All secondary data derived from primary sequencing data is available within the article, supplementary information, or supplemental data files or available from the authors upon reasonable request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Patients

Patients were sourced from two different cohorts, a primary cohort of clinical trial patients with comprehensive data available and a secondary cohort of patients sourced from a single center (Amsterdam UMC), which was added to two parts of this study: investigating associations between genotypes and outcome and MRD (Supplemental methods Figure 1). Patients were de-identified. Samples for the primary cohort were collected within two German Hodgkin Study Group (GHSG) clinical trials: HD21 (n = 44, ClinicalTrials.gov Identifier: NCT02661503, approved by the local ethics committee of the University Hospital of Cologne No. 16-008) and NIVAHL (n = 69, (ClinicalTrials.gov Identifier: NCT03004833, approved by the local ethics committee of the University Hospital of Cologne No. 16-444). HD21 is a trial for newly diagnosed advanced stage HL patients. NIVAHL is a trial for newly diagnosed early-stage unfavorable HL patients.[85] Samples for the secondary cohort were routine care samples from Amsterdam University Medical Centers (UMC) and were collected within the BioLymph-study (2017-2019, VUmc Approved by the ethical committee of the Amsterdam UMC under METc registration number: 2017.008). The study is registered in the Dutch CCMO-register (toetsingonline.nl, NL60245.029.17). A part of this set of Amsterdam UMC samples (2014-2017, prior to the BioLymph study) has been collected through biobanking registered at the Biobank approval committee of VUmc, Amsterdam (2018.359).

Ethical considerations

All patients provided written informed consent to allow the collection of peripheral blood for research purposes. All human subject research was performed in accordance with approved protocols by the local ethics committees as detailed in the section *Patients* above and the Declaration of Helsinki.

METHOD DETAILS

Sample collection and processing

Blood samples were collected as part of an ancillary study to the main clinical trials (primary cohort) and routine care (secondary cohort). Samples were collected at predetermined time points within each trial in the primary cohort and at clinical visits during routine care (secondary cohort) We grouped samples into samples collected before treatment (all patients), approximately one week after treatment initiation, after two cycles of chemotherapy, at the end of treatment and follow up. Samples taken 6-10 days after initiation of treatment qualified as a "one week" sample. Samples collected during follow-up were only available for the secondary cohort. Samples were either collected in 6 or 9 mL EDTA tubes and processed within 2 hours, or in 10ml Streck cell-free DNA tubes (Streck, Omaha, NE) or 10ml PAXgene blood ccfDNA tubes (QIAGEN, Hilden, Germany) and processed within 3 days. In case Streck cell-free DNA or PAXgene Blood ccfDNA tubes were used, they were first centrifuged at 1900 g for 15 minutes at room temperature to separate the plasma. For further plasma purification, the plasma layer was pipetted into a clean 15ml

LoBind tube (Eppendorf, Hamburg, Germany) and centrifuged at 1900 g for another 10 minutes. Plasma was then aliquoted into 2-5ml cryotubes and stored at −80°C. The cell pellet fraction remaining after the first centrifugation step was also aliquoted into 2-5ml cryotubes and stored along with plasma samples at −80°C for use as a germline control. In case EDTA tubes were used, the same protocol was followed, but the first centrifugation step was performed at 900 g for 7 minutes at room temperature and the second centrifugation step was performed at 2500 g for 10 minutes at room temperature.

As samples were collected using two different specialized collection tubes (Streck and QIAGEN ccfDNA) we tested if the type of collection tube impacted results. Comparing the detection frequency of important mutations (Supplemental methods Figure 5A) and copy number variations (Supplemental method Figure 5B) in HL, we found no difference between collection tubes.

DNA extraction

To extract cell-free DNA from plasma samples, the QIAamp Circulating Nucleic-Acid Kit (QIAGEN) was used according to manufacturer's instructions but the Proteinase K incubation at 60°C was lengthened to 2 hours. All samples were extracted manually using a vacuum manifold (QIAGEN). 2-4ml of plasma were used as input. To extract germline DNA from cell pellets, the QIAGEN QIAamp DNA Blood Mini Kit (QIAGEN) was used according to manufacturer's instructions (spin protocol). 200µl of cell pellet was used as input. For DNA extraction from Formalin-Fixed Paraffin-Embedded (FFPE) lymph node samples 10 µm thick slides were cut. DNA extraction was done with the GeneRead DNA FFPE Kit (QIAGEN) according to manufacturer's instructions.

DNA processing, quality control and sequencing

For target enrichment we used customized RNA baits designed with the SureSelect platform (Agilent, Santa Clara, CA). Our target panel has evolved throughout the study, so we used 3 different versions of RNA bait designs targeting 3.847 Mb (version 1), 2.981 Mb (version 2) and 2.997 Mb (version 3). Version 1 of our target panel was designed heuristically by including recurrent locations of somatic mutations and CNVs based on previous studies in HL.[11,12,86] As these are very limited, we also reviewed genomic landscape studies in diffuse large B cell lymphoma and used them as an additional source for designing the target panel.[87–89] Version 2 has evolved from version 1 by omitting genes that were only affected in $\leq$ 1 patients in the first 35 patients analyzed using panel version 1. In version 3, somatically hypermutated intergenic regions were added to increase the average number of available mutations for MRD assessment in follow up samples as has been described by Kurtz et al.[30] Data 11 in the accompanying Mendeley dataset[35] shows which RNA bait design was used for which sample. Data 12 in the accompanying Mendeley dataset[35] shows the regions included in each design. For each sample, individual library preparations, hybridizations and captures were performed. For a first quantification of input DNA, the Qubit fluorometer (Invitrogen, Waltham, MA) was used. We used 25ng of DNA for library preparation where available but required a minimum amount of 15ng. For a more detailed quantification of input DNA considering the size distribution and as part of our quality control, the TapeStation 2200 System (Agilent) was used to quantify input DNA amounts in the size window of interest for cell-free DNA (50-700bp). We required a minimum of 25% of total input DNA to be in the correct size window. TapeStation DNA quantifications restricted to the size window of interest for cell-free DNA (50-700bp) were used for all further calculations requiring DNA amounts as input such as MRD quantification. Libraries were prepared using the

SureSelectXT HS Automated Target Enrichment (Agilent) for Illumina paired-end multiplexed sequencing protocol including a unique molecular identifier (UMI) sequenced with an additional index read and the Agilent Bravo automated liquid handling platform (Agilent). After quality control using the 2200 TapeStation System and quantification using the Qubit System (Thermo Fisher, Waltham, MA), pools of libraries were generated. These pools were then sequenced on an Illumina NovaSeq6000 sequencing system (Illumina, San Diego, CA) using S1, S2 or S4 flow cells and 2 × 100 bp read length or an Illumina HiSeq4000 sequencing system (Illumina) with 2 × 75 bp read length.

As different RNA bait designs were used for different patients, we tested if the version of our target panel impacted results. Comparing the detection frequency of important mutations (Supplemental methods Figure 6A) and copy number variations (Supplemental methods Figure 6B) in HL, we found no difference between collection tubes.

Basic data processing

Bcl2fastq2 software (v2.20, Illumina) was used to demultiplex bcl raw data. We used default settings to demultiplex the forward R1 and reverse R3 reads (–mask-short-adaptor-reads 22–minimum-trimmed-read-length 35–barcode-mismatches 1–adaptor-stringency 0.9). Fastq files were trimmed with the bcl2fasq2 software. The UMI R2 reads were demultiplexed in a separate run using the following settings:–mask-short-adaptor-reads 0 and–minimum-trimmed-read-length 0 without any adaptor trimming. Raw reads were mapped to the human genome reference-build hg19 using the Burrows Wheeler Aligner (BWA) alignment algorithm with a base quality threshold of 15 for read trimming (parameter: -q 15).[90] The resulting BAM files were further processed by a pipeline specifically designed for this project. Briefly, AddUMIsToBam (https://github.com/mbusby/AddUMIsToBam) was used to assign UMIs originating from every read to each aligned read in each BAM file. Fgbio's (http://fulcrumgenomics.github.io/fgbio/) GroupReadsByUmi function was then used to group UMIs using the adjacency strategy with a maximum of 1 edit allowed. Next, Fgbio's CallMolecularConsensusReads function was used to call a consensus read for each group of reads likely originating from the same molecule as identified by the same mapping and UMI with standard settings and accepting also one read of each unique input molecule to define that molecule's sequence. Finally, reads were realigned using BWA mem using standard settings.[90] MPILEUP files were generated using Samtools[91] ignoring overlaps (parameter: -x), adjusting mapping quality for reads containing excessive mismatches (parameter: -C 50), considering only reads with a minimum mapping quality of 40 (parameter: -q 40) and considering only bases with a minimum base quality of 30 (parameter: -Q 30). The MPILEUP file was limited to the genomic region captured during library preparation (parameter: -l).

Somatic single base substitution calling

To improve sensitivity and specificity, somatic single base substitution calling was aided by comparative error suppression (CES). For this procedure, cell-free DNA from a cohort of 10 healthy donor patients was prepared and sequenced as described above. Post UMI-processing consensus reads aligned to hg19 by BWA mem[90] were used as input. Using customized scripts, all positions with a non-reference (hg19) allele frequency > 5%, likely representing germline single nucleotide polymorphisms, were set to a non-reference allele frequency of 0 in the reference cohort. Next, TNER[31] was used to compile a database of position-specific average background error rates for each position in the target region using the healthy donor

controls as input. This database was used as input to TNER[31] to polish each cell-free DNA MPILEUP generated from post UMI-processing consensus reads aligned to hg19 by BWA mem[90] as described above with the polish strength parameter set to 0.05 (parameter: input.alpha = 0.05). Briefly, TNER[31] uses tri-nucleotide error rates in the healthy reference cohort to suppress likely sequencing errors, thus reducing tri-nucleotide error rate profiles to suppress stochastic sequencing errors. Additionally, stereotypical sequencing errors at positions with high error rates were suppressed using customized scripts. To this end, each potential variant position was compared against the position-specific average background error rate profile generated above. If the average background error at a specific position was > 0.5%, this position was blacklisted for variant calling as a position with a high rate of sequencing errors.

Next, this error-corrected dataset generated for each cell-free DNA sample was used as input for a customized variant calling pipeline. A variant was called at a position, when the following criteria were met: the minimum sequencing depth at a position was 50, at least 11 reads are supporting the variant, the variant has support on both the plus and the minus strand, the relative variant allele frequency of the variant is at least 0.1%, the variant has a relative variant allele frequency in the germline sample of less than 1%, the relative allele frequency in the cell-free DNA sample is at least 3 times larger than in the germline sample, a Fisher's exact test on a 2 × 2 contingency table including the variant versus reference and the cell-free DNA versus the germline reference sample has a p value < 0.05 and finally a one-tailed binominal test with the alternative hypothesis that the observed variant allele frequency is higher than the background error rate at that position has a p value < 0.05. As an additional filter, we checked each variant for its frequency in ExAC.[92] Variants with a frequency above 0.01 in ExAC version 0.3[92] were excluded. Furthermore, we excluded single base substitutions that were in close proximity (< 6 base pairs) to an identified InDel. Variants were annotated and classified using ANNOVAR and databases ExAC version 0.3 (parameter: protocol -exac03) dbNFSP version 3.0a (parameter: -protocol dbnsfp30a) and dbSNP version 147 (parameter: -protocol avsnp147).[93]

Somatic small InDel calling

We used Varscan2[94] to call small somatic insertions and deletions (InDels) and customized scripts for filtering identified variants. We used MPILEUP files generated as described above from post UMI-processing consensus reads aligned to hg19 by BWA mem[90] as input. First, a comprehensive, unfiltered list of all potential InDels was generated using Varscan2[94] with the following settings: minimum coverage of cell-free DNA sample file 1 (parameter:–min-coverage), minimum coverage of corresponding germline file 1 (parameter:–min-coverage-normal 1), minimum relative allele frequency 10^{-9} (parameter:–min-var-freq), minimum frequency to call a variant homozygous 0.75 (parameter:–min-freq-for-hom), minimum somatic p value 0.99 (parameter:–somatic-p value) and strand filter off (parameter:–strand-filter 0). Of note, the parameters were deliberately chosen to be very loose as more stringent filtering was supposed to take place in the next step.

Using customized scripts, we filtered each raw InDel call file so only calls fulfilling the following criteria survived filtering: relative allele frequency of the matched germline < 0.5%, somatic p value of call < 0.01, absolute allele frequency of the matched germline < 4, not called as loss of heterozygosity by VarScan2,[94] at least 11 reads supporting the variant, at least one read on the plus strand and one read on the minus strand supporting the variant and minimum sequencing depth of 50 at the

variant location in the cell-free DNA sample. As an additional filter, we checked each variant for its frequency in ExAC.[92] Variants with a frequency above 0.01 in ExAC version 0.3[92] were excluded. Variants were annotated and classified using ANNOVAR and databases ExAC version 0.3 (parameter: protocol -exac03) dbNFSP version 3.0a (parameter: -protocol dbnsfp30a) and dbSNP version 147 (parameter: -protocol avsnp147).[93]

Somatic copy number calling

A pipeline was developed to identify somatic copy number changes based on CNVKit.[59] First, cell-free DNA from a cohort of 10 healthy donor patients was prepared and sequenced as described above. This healthy control cohort was used as a reference set to define an average technical background, against which somatic copy number changes of a specific sample were measured. The target region input for CNVKit[59] was defined as the region that was covered by the RNA baits for the respective panel design. CNVKit's batch function[59] was then run for each cell-free DNA sample, using post UMI-processing consensus reads aligned to hg19 by BWA mem[90] as input. Access data provided by CNVKit[59] for hg19 and the drop low coverage (parameter:–drop-low-coverage) option were used in all runs. Standard parameters were used for all other settings. Segmented copy number calls were generated and an amplitude threshold of 0.02 or −0.02 was used to define a segment as a somatic copy number gain or loss, respectively. We did deliberately not try to infer absolute somatic copy number changes, because (I) cell-free DNA is not purely tumor-derived and, sometimes heavily, diluted by background germline DNA reducing sensitivity, and (II) cell-free DNA-derived somatic copy number changes represent an average of individual tumor cells somatic copy number states in the body which are, especially in the case of HL, known to be very variable between individual tumor cells.[95,96]

To derive arm-level and focal recurrent and thus potentially biologically relevant somatic copy number changes, GISTIC 2.0.[60] was used. Segmented copy number calls were used as input and GISTIC 2.0. was run including the whole sample cohort with the following settings: cytoband and gene location information was taken from hg19, gain threshold 0.02 (parameter: -ta), loss threshold 0.02 (parameter: -td), calculate significance by gene (parameter: -genegistic 1), including broad-level analysis (parameter: -broad 1), threshold to distinguish broad from focal events 0.98 (parameter: -brlen), confidence level 0.75 (parameter: -conf), q-value threshold 0.25 (parameter: -qvt), performing arm peel off (parameter: -armpeel 1) and collapse method extreme (parameter: -gcm). Standard parameters were used for all other settings. Broad arm-level plots, gain and loss q-plots and copy number heatmap plots were generated from resulting output using GISTIC's[60] built-in plotting functions.

Additional variant filtering

For some analyses, mutations were only included if they occurred in the coding region of the genome. This is indicated in the respective analysis.

To identify recurrently mutated genes which are likely driver genes[37] and thus relevant for HL pathogenesis either as tumor suppressor genes or as oncogenes, we applied a slightly modified version of the Vogelstein rule,[37] which is a deliberate choice for a rather conservative approach tending to underestimate cancer driver genes.[37] Furthermore, the 20/20 rule originally described by Vogelstein et al.[37] and derived, ratiometric methods such as the 20/20+ algorithm have been shown to highly reliably predict functionally validated, known cancer driver genes and

have a high stability across datasets.[97] We used customized scripts to identify a gene as a tumor suppressor gene, if the ratio of the combined number of splice site and nonsense mutations as well as frameshift insertions and deletions to all coding mutations was $\geq$ 0.25. A gene was identified as an oncogene if it was not a tumor suppressor gene and the ratio of mutations leading to an amino acid change that is shared with at least one other mutation leading to an amino acid change to all mutations leading to an amino acid change was $\geq$ 0.20. Evaluation of genes as potential tumor suppressor genes or oncogenes was performed for all genes with at least 5 mutations in the cohort.

Technical and clinical validation and Quality control

For technical validation we spiked well-characterized reference DNA (NA12878)[32,33] fragmented to approximately cfDNA size into healthy donor cfDNA at proportions of 10%, 1%, 0.1%, 0.05%, 0.025%, 0.01% and 0%. Based on an allele frequency of 50% for heterozygous variants, these spike-in samples represented technical cfDNA validation samples with an expected mutated variant allele frequency of 5%, 0.5%, 0.05%, 0.025%, 0.0125% and 0.005%, respectively. Spike-ins of 5%, 0.5% and 0% were used to technically validate variant calling, whereas spike-ins of 0.1%, 0.05%, 0.025%, 0.01% and 0% were used to technically validate minimal residual disease detection. All spike-in samples were sequenced and bioinformatically processed as described above. Processing was identical to actual patient samples. For selected analyses UMI-based error suppression, CES or both were omitted. To calculate error profiles of sequencing with or without one or both error suppression methods, we used the 0% spike-in sample and randomly chose 1000 positions that were both not single nucleotide polymorphisms (SNPs) in NA12878 and the healthy donor sample – and should thus have zero non-reference bases detected at each position – and calculated each positions error rate as the combined variant allele frequency of all non-reference bases at the respective position. We modeled the error rate of different error suppression strategies as a decay process after sorting the error rates from high to low with Y denoting the modeled error rate, Y_0 denoting the highest observed error rate, K the estimated decay rate and X the position in the high-to-low sorted error rate profile:

$$Y = Y_0 * e^{-K*X}$$

For technical validation of variant calling, we compiled a positive truth set of mutations (n = 42) by identifying 22 cancer gene census genes[98] that were (I) covered by our panel, (II) had at least one consensus variant identified in NA12878 and (III) were called in the 10% spike-in sample and thus in principle covered by our sequencing panel. Similarly, we compiled a negative truth set of non-mutated positions (n = 1527) by randomly choosing one position in each genomic region covered by our panel of each of the 22 cancer gene census genes defined above that were both not single nucleotide polymorphisms (SNPs) in NA12878 and the healthy donor sample and including this position and both positions 3 bases up- and downstream, given they were also not single nucleotide polymorphisms (SNPs) in NA12878 and the healthy donor sample, into the negative truth set. The full positive and negative truth sets are available as Data 10 in the accompanying Mendeley dataset.[35] We then performed variant calling in the 0.5% spike-in sample with or without UMI-based error suppression, CES or both and calculated sensitivity and specificity to detect variants with an allele frequency of 0.5% based on these truth sets.

For technical validation of minimal residual disease detection, we slightly modified the above-described positive truth set by limiting it to variants detected with an absolute variant allele frequency > 29 resulting in 40 variants. We calculated minimal residual

disease as described above using this truth set as the technical validation baseline sample for the 0.1%, 0.05%, 0.025%, 0.01% and 0% spike-in samples with or without UMI-based error suppression, CES or both. To assess the contribution of UMI-based error suppression, CES or its combination to our assay's sensitivity for minimal residual disease detection, we compared measured minimal residual disease in our technical validation spike-in samples with expected minimal residual disease based on expected mutated allele frequency in the spike-ins. We calculated bootstrap confidence intervals using R[99] and the package boot[100] based on 1000 bootstrap replicates. We calculated the limit of blank of our assay for minimal residual disease detection based on standard protocols for assay validation[34] using the 0% spike-in sample as the blank sample and the 0.01% spike-in sample as a low-level sample. The limit of blank was defined as the mean blank minimal residual disease value + 1.645 x the 1000 replicate based bootstrap standard deviation of the blank sample's mean minimal residual disease value. The limit of detection was defined as the mean blank minimal residual disease value + 1.645 x the 1000 replicate based bootstrap standard deviation of the low-level sample's mean minimal residual disease value.

For clinical validation of variant calling, we subjected 8 corresponding Formalin-Fixed Paraffin-Embedded (FFPE) tumor biopsies to our sequencing and bioinformatics pipeline and compared called variants detected in the tumor biopsies with those detected in cfDNA.

For further clinical validation of the specificity of minimal residual disease detection, we tried to detect minimal residual disease in 10 healthy donor samples, which we used as minimal residual disease negative samples. We used variant call from 6 patients and subjected these 10 healthy donor samples to the standard minimal residual disease detection workflow using each patient's variant call set, resulting in 60 independent experiments.

To calculate mean per gene coverage we calculated the mean coverage across all locations of single base substitutions in a given gene across the whole cohort. To calculate mean per patient coverage, we calculated the mean coverage across all locations of single base substitutions in a given patient.

Clustering of mutations within 3-dimensional protein structures

Mutation3D[39] was used to cluster variants within 3-dimensional protein structures in all identified oncogenes with at least 8 variants in our cohort as this was heuristically identified as the minimum number of variants needed to perform a meaningful analysis for the respective oncogene. For each analyzed oncogene, all amino acid sequence changes caused by identified variants where used as input to Mutation3D,[39] which was run in advanced mode. Settings were chosen to include all ModBase homology models with a ModPipe quality score (MPQS) $\geq$ 1.1, requiring at least three mutations and at least one unique amino acid change per cluster and allowing a maximum intracluster distance between mutations of 25Å. We considered an identified cluster sufficiently likely to be truly present if its p value < 0.1.

Mutual exclusivity and Co-occurrence of mutations analysis

To test if oncogenes or tumor suppressor genes co-occur or are mutually exclusive, we used DISCOVER.[54] We calculated the background matrix using all mutations and small InDels but limited our analysis to amino acid sequence altering mutations in oncogenes and tumor suppressor genes. We used a p value of 0.05 as a threshold to detect mutual exclusivity or co-occurrence.

Mutational burden

Mutational burden was calculated by dividing the number of mutations occurring within the target region for a given patient by the target region size.

Mutational signatures

Mutational signatures were identified using R[99] and deconstructSigs.[56] We deconstructed the mutations in each patient and for each gene with > 30 mutations in our cohort into mutational signatures with the COSMIC single base substitution signatures as reference signatures.[55] Trinucleotide context normalization was based on trinucleotide context observations in the coding exome (parameter: tri.counts.method = 'exome'). We considered a mutational signature detected in a patient or gene if its associated weight was > 0.1. Mutational signatures detected in at least 10% of patients in our cohort were considered to be recurrent mutational signatures present in HL (COSMIC single base substitution signatures 1, 3, 6, 9, 15 and 25). For both patients and genes, the proportion of mutations not explained by any of these signatures was classified as other/not assigned.

Pathway associations

We assigned KEGG,[101] GO[102,103] or REACTOME[104] pathways to identified oncogenes and tumor suppressor genes using g:Profiler.[105]

Detection of viral and bacterial associations

To detect viral and bacterial associations of HL by liquid biopsy, we used a previously developed pipeline[61] built around KRAKEN.[62] In brief, we used post UMI-processing consensus reads aligned to hg19 by BWA mem[90] from all baseline cfDNA and germline control samples as input to KRAKEN[62] and generated output reports which we filtered with KRAKEN's built-in confidence scoring function with a confidence score of 0.25 (parameter:–confidence 0.25). We used the KRAKEN standard database including viral and bacterial reference genomes for our analysis. Next, we arranged and filtered the output reports in a way that (II) excluded all human-aligned reads, (II) included only species-level hits, (III) excluded all species with > 4 cumulative alignments in our healthy control cohort to account for misalignments and common sequencing contaminants and (IV) excluded all species that had > 2 germline samples with > 4 reads aligned to it. In general, a species was determined to be detected in a cfDNA sample, if > 4 reads aligned to it. However, for EBV (HHV4) we used 33 as a threshold, as low-level EBV DNA is commonly detected in blood of healthy humans in highly sensitive assays.[106] To further exclude likely false-positive taxa, we excluded common sequencing contaminants compiled within one of our previous studies,[61] taxa that are common sources of biotechnological products (e.g., enzymes), such as bacteria from the Thermaceae family and thus also more likely to be contaminants and finally taxa that do not colonize mammalian hosts (See Data 10 in the accompanying Mendeley dataset[35] for details).

Analysis of clonal structure of mutations

PyClone[75] was used to analyze clonal structure of mutations. All identified single base substitutions in a patient were used as input. Normal copy number was set to 2. As absolute copy number information was not available, minor copy number and major copy number were set to 0 and 2, respectively. Prior was set to total copy number (parameter:–prior total_copy_number). Burnin was set to 1000 (parameter:–burnin 1000). Default settings were used for all other parameters. As a cleaning step, starting from the cluster with the highest mean cellular prevalence, clusters were pooled with the next cluster in order of decreasing mean cellular prevalence until a cluster size of at least 3 mutations was reached. Subsequent analyses of clonal structure were only performed

for patients with at least 2 identified clones. To derive the cancer cell fraction (CCF) of each mutation, the cellular prevalence of the respective mutation was divided by the mean cellular prevalence of the cluster with highest mean cellular prevalence after cleaning as described above. The CCF of mutations in the main clone was set to 1. For subsequent dichotomized analyses, a mutation with a CCF > 0.75 was considered to be a main clone mutation and all other mutations were considered subclonal.

To analyze the relationship between mutational signatures and main clone versus subclonal mutations, deconstructSigs[56] was run separately on all genes in which > 70% of mutations were main clone mutations (main clone dataset) and all genes in which > 80% of mutations were subclonal (subclone dataset).

Minimal residual disease assessment

Minimal residual disease assessment was performed using customized R scripts. We used only single base substitution variant calls from baseline samples excluding non-coding RNA regions as input. Follow up samples were processed as described above including UMI-based and CES. Our process scans each follow up dataset at each variant position in the baseline dataset and identifies the sequencing depth and count of mutated alleles including only counts of the variant base identified in the baseline sample at each such position in the follow up sample. Then, it calculates *preliminary MRD,* defined as the sum of the count of all variant alleles defined as above in the follow up sample divided by the combined sequencing depth at these positions in the follow up sample. Next, we included a dusting step based on the assumption that mutated alleles should distribute evenly across all variants in the baseline file and outliers have a higher likelihood of being false tumor-derived somatic variant calls, for example CHIP variants[107] not captured in the sample's germline control leading to higher MRD detected than is truly present. We populated a Poisson distribution for each variant position with the expected number of events defined as: *preliminary MRD* $*$ *sequencing depth at variant position in follow up sample*. We then derived the Poisson-likelihood that the observed or a higher number of mutated alleles at this position occurs. If this likelihood was $< 10^{-5}$, we assumed this variant position had a higher than tolerable for MRD detection likelihood of being a false (e.g., non-tumor-derived) call and omitted it. After this dusting process we calculated final MRD as the sum of the remaining count of all variant alleles defined as above in the follow up sample divided by the combined remaining sequencing depth at these positions after the dusting process in the follow up sample.

To evaluate response in patients, we calculated the $\log_{10}$ reduction in MRD from baseline for a sample. We corrected for changes in cfDNA concentration between samples analogous to a method outlined before.[30] The $\log_{10}$ reduction was calculated according to the following formula:

$$MRD\ reduction = log_{10}\left(\frac{MRD_{follow\ up\ sample} * cfDNA\ concentration_{follow\ up\ sample}}{Average\ AF_{baseline\ sample} * cfDNA\ concentration_{baseline\ sample}}\right)$$

If MRD in a follow up sample was below the limit of blank, we denoted this as MRD not detected. For the purpose of statistical calculations within this manuscript we set MRD not detected to the midpoint between 0 and the MRD reduction of a fictitious *average* sample analogously to a previously described method[108] with:

$$MRD_{follow\ up\ sample} = limit\ of\ blank$$

$$cfDNA\ concentration_{follow\ up\ sample} = avg.\ cfDNA\ concentration\ of\ all\ follow\ up\ samples$$

$$Average\ AF_{baseline\ sample} = avg.\ AF\ of\ all\ baseline\ samples\ ,\ and$$

$$cfDNA\ concentration_{baseline\ sample} = avg.\ cfDNA\ concentration\ of\ all\ baseline\ samples.$$

Fluorescence *in situ* hybridization

FISH analyses were performed according to standard protocols. Briefly, bacterial artificial chromosome (BAC) clones were selected from the UCSC Genome Browser and ordered from Thermo Fisher. The BAC DNA was extracted from LB-Cultures according to standard protocols and labeled with Fluorescent Dye using the Nick Translation DNA Labeling System 2.0 (ENZO, New York, New York) according to manufacturer's recommendations. The following BAC clones and labeled dUTPs (Fluorescent Dyes) were used: PD-L1/upstream: RP11-963L3 & RP11-12D24 (green 496 dUTP, ENZ-42831); PD-L2/downstream: RP11-207C16 & RP11-845C2 (orange 552 dUTP, ENZ- 42842); centromere near control probe CTD-2024L1 & RP11-203L2 (aqua 431 dUTP, ENZ- 42853).

To analyze genetic alterations in 9p24.1 we used a combination of CD30 IHC and FISH assay. We produced serial slides for CD30 IHC and FISH. For evaluation, we used the Bioview System (Abbott Molecular, Chicago, Illinois). We produced scans of the IHC slide and marked regions of interest (ROIs) containing large amounts of CD30-positive HRSC. Likewise, a scan of the FISH slide with the DAPI filter was recorded. Both of the scanned images (IHC and FISH) were matched at equivalent points and an overlay was produced. 50 tumor cells per case were analyzed in several of the previously selected ROIs.

QUANTIFICATION AND STATISTICAL ANALYSIS

Discrete variables were compared using Fisher's exact test. Two continuous variables were compared using the Student's t test if normality of variables could be assumed or the Mann-Whitney test if not. Multiple continuous variables were compared using one-way ANOVA with Welsh correction if normality of variables could be assumed or the Kruskal-Wallis test if not. Normality of variables was assessed heuristically examining the distribution of variable values. Statistical tests used are indicated when performed.