A Cluster of Evolutionarily Recent KRAB Zinc Finger Proteins Protects Cancer Cells from Replicative Stress–Induced Inflammation

Filipe Martins[1,2], Olga Rosspopoff[1], Joana Carlevaro-Fita[1], Romain Forey[1], Sandra Offner[1], Evarist Planet[1], Cyril Pulver[1], HuiSong Pak[3,4,5], Florian Huber[3,4,5], Justine Michaux[3,4,5], Michal Bassani-Sternberg[3,4,5], Priscilla Turelli[1], and Didier Trono[1]

ABSTRACT

Heterochromatin loss and genetic instability enhance cancer progression by favoring clonal diversity, yet uncontrolled replicative stress leads to mitotic catastrophe and inflammatory responses that promote immune rejection. KRAB domain-containing zinc finger proteins (KZFP) contribute to heterochromatin maintenance at transposable elements (TE). Here, we identified an association of upregulation of a cluster of primate-specific KZFPs with poor prognosis, increased copy-number alterations, and changes in the tumor microenvironment in diffuse large B-cell lymphoma (DLBCL). Depleting two of these KZFPs targeting evolutionarily recent TEs, ZNF587 and ZNF417, impaired the proliferation of cells derived from DLBCL and several other tumor types. ZNF587 and ZNF417 depletion led to heterochromatin redistribution, replicative stress, and cGAS–STING-mediated induction of an interferon/inflammatory response, which enhanced susceptibility to macrophage-mediated phagocytosis and increased surface expression of HLA-I, together with presentation of a neoimmunopeptidome. Thus, cancer cells can exploit KZFPs to dampen TE-originating surveillance mechanisms, which likely facilitates clonal expansion, diversification, and immune evasion.

Significance: Upregulation of a cluster of primate-specific KRAB zinc finger proteins in cancer cells prevents replicative stress and inflammation by regulating heterochromatin maintenance, which could facilitate the development of improved biomarkers and treatments.

Introduction

Diffuse large B-cell lymphoma (DLBCL), the most common type of lymphoid malignancy, is plagued by relapse or refractoriness to first-line therapy in nearly 50% of cases (1). DLBCL arises from the neoplastic transformation of mature B cells during the final steps of their differentiation into antibody-secreting plasma cells or memory B cells. Three major molecular subtypes of DLBCL with prognostic relevance have been delineated based on similarities between the transcriptome of these tumors and their putative cell of origin (COO; ref. 2). Germinal center B-cell (GCB) and activated B-cell (ABC) DLBCLs are thought to derive from B cells blocked at the centroblast and plasmablast stages, respectively, whereas the transcriptome of so-called "unclassifiable" DLBCL does not evoke any of these two differentiation stages.

Genomic instability is a primary driver of the clonal evolution and diversification of cancer cells. It frequently results from errors occurring during genome replication, which starts with the assembly of pre-replicative complexes (pre-RC) during the G_1 phase of the cell cycle, an event known as origin licensing (3). During the early S phase, pre-RCs recruit the DNA polymerase and elongation factors to form the replication fork machinery. Subsets of licensed replication origins then become activated in a sequential and spatially regulated manner (4), allowing the bidirectional progression of replication forks across the genome. Conducting a proper replication program in the presence of short cell cycling, intercurrent mutagenic events, and exacerbated oncogene-induced transcriptional activity is challenging, and accordingly, DLBCL cells are prone to replicative stress (RS; ref. 5), characterized by the stalling and potential collapse of replication forks leading to genomic instability.

Increasing evidence suggests a central role for chromatin compaction in the selection and timing of activation of DNA replication origins (6). Heterochromatin regions are known to be late-replicating, a phenomenon attributed to their high content in repetitive DNA sequences. These regions are hotspots for homologous recombination (7), secondary DNA structures (8), and the formation of RNA–DNA hybrids that are major sources of RS (9). Pericentromeric regions and telomeres, which are enriched for satellite and simple repeats, are the most heterochromatin-dense regions of the genome. In addition, some 4.5 million transposable elements (TE) sprinkled over the rest of the human genome contribute significantly to its heterochromatin landscape. Most human TEs are retroelements belonging to the LTR (long terminal repeat)-containing endogenous retroviruses (ERV), long and short interspersed nuclear elements (LINE and SINE, respectively) or SVA (SINE-VNTR-Alu-like) families, which replicate by a "copy-and-paste" mechanism through reverse transcription of an RNA intermediate and integration of its DNA copy elsewhere in the genome. TEs are motors of genome evolution and a major source of *cis*-acting regulatory sequences that shape gene regulatory networks at play in broad aspects of human biology (10). The disruptive potential of TEs has led to the evolutionary selection of many protective

[1]School of Life Sciences, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland. [2]Clinics of Medical Oncology, Cantonal Hospital of Fribourg (HFR), Fribourg, Switzerland. [3]Department of Oncology, Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland. [4]Agora Cancer Research Centre, Lausanne, Switzerland. [5]Ludwig Institute for Cancer Research, University of Lausanne, Lausanne, Switzerland.

Corresponding Author: Didier Trono, EPFL SV LVG Station 19, 1015 Lausanne, Switzerland. E-mail: didier.trono@epfl.ch

Cancer Res 2024;XX:XX–XX

doi: 10.1158/0008-5472.CAN-23-1237

mechanisms, whether to prevent transposition during genome reprogramming in the germline and early embryogenesis or to control the transcriptional impact of TE-embedded regulatory sequences. Repression of TE loci via DNA methylation and heterochromatin formation through methylation of lysine 9 of histone 3 (H3K9me3), stand prominently among these mechanisms, and Krüppel-associated box (KRAB) zinc finger proteins (KZFP) among their mediators.

KZFPs constitute the largest family of transcription factors (TF) encoded by higher vertebrates, with nearly 400 members in humans alone, more than a third of them primate-restricted and the products of recent gene duplication events (11). KZFPs bind DNA in a sequence-specific manner through their C-terminal arrays of zinc fingers, and a large majority of human KZFPs have TE-embedded sequences as primary genomic targets. Most act as transcriptional repressors via the KRAB domain-mediated recruitment of KAP1 (KRAB-associated protein 1, also known as tripartite motif protein 28 or TRIM28; ref. 12), which serves as a scaffold for the assembly of a heterochromatin-inducing complex comprising nucleosome remodeling, histone deacetylase, and histone and DNA methyltransferase activities (13). Rather than just another silencing mechanism aimed at preventing TE mobilization, KZFPs are increasingly recognized as facilitating the domestication of TEs' regulatory potential, as the vast majority of TE integrants are no longer retrotransposition-competent but still harbor KZFP binding sites. Indeed, KZFPs not only repress TEs immediately after zygotic genome activation but also corral their transcriptional influences at various stages of development and in adult tissues (14). Thus, KZFPs both minimize the genotoxic potential of TEs and partner up with their genomic targets to shape lineage- and species-specific transcription regulatory networks.

KZFPs have been proposed to exert collective tumor-suppressive functions across many tumor types (15). However, individual family members have been proposed to promote disease progression in lymphocytic leukemia (16), lung (17), colorectal (18), and ovarian cancers (19). In addition, their main cofactor KAP1 functions as an E3-ubiquitin ligase capable of inactivating p53 (20), and the KAP1-recruited SETDB1 methyltransferase was recently found to foster immune evasion in melanoma and lung cancer by repressing TE transcription (21), a source of dsDNA/dsRNA species sensed by intracellular innate immunity pathways and of endogenous viral peptides with immunogenic properties (22).

Here, we took advantage of two large data sets of DLBCL patients with extensive clinical and molecular documentation (23) to explore the potential role of KZFPs in this hematologic malignancy. This led us to discover that a group of primate-specific KZFPs targeting evolutionarily recent TEs protect cancer cells from excessive RS associated with genome-wide alterations of the heterochromatin landscape, thus safeguarding these cells from cell-intrinsic inflammatory responses.

Materials and Methods

Cell culture

OCI-Ly7 (RRID:CVCL_1881) and U2932 (RRID:CVCL_1896) human lymphoma cell lines were obtained from DSMZ—German Collection of Microorganisms and Cell Cultures (www.dsmz.de). The other DLBCL cell lines, namely, SUDHL4 (RRID:CVCL_0539) and HBL1 (RRID:CVCL_4213), were kindly supplied by Elisa Oricchio (EPFL, Swiss Institute for Experimental Cancer Research) and typed using short tandem repeat profiling by Microsynth cell line authentication service. HL60 (RRID:CVCL_0002), K562 (RRID:CVCL_0004), LS1034 (RRID:CVCL_1382), SNU16 (RRID:CVCL_0076), and SW480 (RRID:CVCL_0546) cell lines were obtained from ATCC. SW480 were cultured in L15 medium (Sigma), whereas HL60, K562, U2932, HBL1, SUDHL4, SNU16, and LS1034 cells were grown in RPMI-1640 (Gibco), with 10% fetal bovine serum and 1% penicillin–streptomycin. OCI-Ly7 cells were grown in IMDM (Gibco) with 20% fetal bovine serum (FBS) and 1% penicillin–streptomycin. HEK293T cells were grown in DMEM (Gibco) supplemented with 10% FBS, 1% penicillin–streptomycin, and 2 mmol/L of L-glutamine. All cell lines were grown at 37°C and 5% CO_2 and used within 10–15 passages. Primary normal dermal fibroblasts were cultured in fibroblast basal medium (ATCC) supplemented with fibroblast growth kit-low serum (ATCC). Human M1-polarized macrophages and M-CSF were purchased from StemCell. After thawing, M1 macrophages were resuspended in a complete RPMI-1640 (Gibco) medium with M-CSF (StemCell) at a final concentration of 50 ng/mL. Human peripheral blood CD4 T cells were obtained from StemCell. After thawing, CD4 T cells were cultured in Xvivo 15 medium (Lonza), supplemented with 10% FCS, 1% penicillin–streptomycin, and 2 mmol/L of L-glutamine. Activation was performed by the addition of Dynabeads Human T Activator CD3/CD28 (25 μL/0.5 million cells, Life Technologies), along with IL2 (30 U/mL, Sigma).

All cell lines were tested for the absence of *Mycoplasma* contamination by Eurofins Genomics Europe using a qPCR method in August 2020 for U2932 and OCI-Ly7, in October 2020 for K562, in November 2020 for HL60, February 2021 for HBL1 and SUDHL4, and April 2023 for SW480. LS1034 was tested using a Mycoplasma Detection Kit (InvivoGen) based on isothermal PCR detection of the 16S rRNA gene.

KZFP knockdown experiments

The short hairpin sequences targeting ZNF587/417 were designed as described (24). ZNF587/417 shRNAs are targeting the following sequences: 5′-GCAGCATATTGGAGAGAAATT-3′ (shRNA.1) and 5′-AGTCGAAAGAGCAGCCTTATT-3′ (shRNA.2). The other shRNAs were purchased from Sigma-Aldrich, with shRNAs against ZNF587B/814 and ZNF586 targeting the following sequences: 5′-CCTTCTAAGCAGAGTATTTAT-3′ (shZNF587/814), 5′- GCTTATACATCTAGTCTCATT-3′ (shZNF586). An shRNA with the nontargeting sequence 5′-CAACAAGATGAAGAGCACCAAG-3′ (shScramble, shScr) was used as a negative control in all experiments. Lentivectors (LV) were produced as described at http://tronolab.epfl.ch. Cancer cell lines and primary cells were transduced with a multiplicity of infection (MOI) of 10. 48 hours after transduction, cells were selected for 3–4 days using 3 μg/mL of puromycin. For experiments performed between days 2 and 4, cells were used without selection while doubling the MOI. CD4 T cells were transduced 24 hours after the addition of activation beads.

Cell proliferation assays

The tetrazolium bromide (MTT) assay was used to determine cell proliferation for 4–7 days. Cells were plated after 5 days of LV transduction and 3 days of puromycin selection, except for SUDHL4 cells, which were plated 3 days after LV transduction. Briefly, 100 μL of cells were seeded in 96-well plates at a concentration of 5k/well for primary dermal fibroblasts, 20k/well for HBL1, HL60, K562; 40k/well for U2932, OCI-Ly7, and SUDHL4; 3k/well for SNU16, SW480, and LS1034. On each consecutive day, cells were incubated with 5 mg/mL MTT for 3 hours at 37°C. The reaction was stopped by the addition of 100 μL of 10% Ultrapure SDS (Invitrogen) to dissolve precipitated crystals. After overnight incubation at 37°C, the absorbance was measured on a microplate spectrophotometer at 570 nm while subtracting the background absorbance measured at 690 nm.

Flow cytometry

Cell death was determined using flow cytometry through the quantification of cell-surface Annexin V–APC (BioLegend, RRID: AB_2561515) and propidium iodide (PI; BioLegend) costaining. Briefly, 1 million cells were harvested 3 days after transduction and washed twice with cell staining buffer (BioLegend). Cells were resuspended in 100 μL of Cell Staining Buffer (BioLegend) and incubated in the dark for 15 minutes at room temperature following the addition of 5 μL of Annexin V–APC antibody and 10 μL of PI solution. The measurement of ecto-CALR (PE Anti-Calreticulin antibody [FMC 75], ab83220, RRID:AB_1859755) and HLA-I (APC anti-human HLA-A, B, C Antibody, RRID:AB_314878) surface expression was conducted after incubation with respective antibodies for 20 minutes at 4°C before being washed and resuspended in the same buffer, following the manufacturer's instructions. Cell-cycle distribution was analyzed by flow cytometry measurement of cellular DNA content using PI staining. Three million cells were collected, washed, and resuspended in 1 volume of ice-cold 1× Dulbecco's Phosphate Buffer Saline (PBS; BioConcept) before being fixed by the addition of 2 volumes of ice-cold 100% ethanol during slow vortexing. After 45 minutes of incubation at 4°C, cells were washed again in ice-cold PBS and resuspended in 1.5 volume of ice-cold PI staining solution (0.1% triton, 200 μg/mL RNAse A, and 50 μg/mL PI) for 30 minutes. Samples were then analyzed on a BD LSR II (Becton Dickinson) flow cytometer, using BD FACSDiva software, and quantified using FlowJo single-cell analysis software (FlowJo, RRID:SCR_008520).

CXCL10 measurement in cell culture supernatant

CXCL10 concentration was determined using the Human Pro-Quantum Immunoassay Kit according to the manufacturer's instructions. This method relies on antigen recognition by two antibodies, each conjugated to a specific DNA oligo. The presence of the antigen will allow the proximity ligation of the DNA oligo from one antibody to the other, and further quantification of the ligation product by quantitative PCR, as a readout of CXCL10 concentration. Fresh supernatants from U2932 KD and control cells were harvested at day 6 after transduction. Supernatants were mixed with the two oligonucleotide-conjugated antibodies, diluted 1:10 with assay dilution buffer, and incubated for 1 hour at room temperature. Oligonucleotide ligation was performed by the addition of Master mix and ligase and incubation for 1 hour at room temperature. Quantification of the ligation products was determined by quantitative PCR on an Applied Biosystems 7900HT Fast Real-Time PCT system (Thermo Scientific) using the recommended instrument settings. The Ct values were exported to the ProQuantum Cloud application (apps.thermofisher.com/apps/proquantum) to determine CXCL10 concentrations. Finally, supernatant protein concentrations were normalized by cell concentration measured at day 6.

Analysis of published data sets

The BAM files from 775 exon-enriched DLBCL transcriptomes were downloaded from the European Genome-Phenome Archive (EGA) at the European Bioinformatics Institute under the accession number EGAS00001002606 (23), upon data access agreement from Dave lab, Duke University. To ensure better analysis quality, we excluded 142 samples with fewer than 1 million counts in total. After filtering lowly expressed genes with an average expression <1 CPM, 23,560 genes were subjected to a univariate Cox regression analysis using the R-package "survival." For this purpose, each gene expression in normalized count units was compared with the overall survival (OS) followed by the application of an independent Cox model for each gene. 5,220 genes were found to be significantly associated with survival (Wald test P value <0.05) and 1,530 remained after correction for multiple testing using the Benjamini–Hochberg method (false discovery rate (FDR) <0.05). KZFP genes were identified using a *de novo* generated whole-genome annotation (Supplementary Table S1) using hidden Markov model (HMM) profile generated from seed sequences downloaded from Pfam (25) and created using hmmbuild from HMMER2 (hmmer.org). KZFP genes were further identified as protein-coding using the Ensembl BioMart database (GRCh38.p13, v103; ref. 26). Other human TF genes were identified using a previously published list (27). The categorization of the KZFPs was based on their evolutionary age as previously described (11).

Clinical data (OS, IPI details, and response to first-line therapy), driver mutation status, and COO classification of DLBCLs were extracted from the metadata provided in (23). The relationship between the 41 KZFP genes identified in Cox regression analysis was measured by Spearman rank correlation using the rcorr function of the Hmisc R-package and plotted using pheatmap. *KZFP* genes' genomic cluster affiliation and evolutionary ages were extracted as previously described (24). Hierarchical clustering analysis of lymphoma samples based on the expression of the 18 KZFPs located in the Chr19q coexpression module was computed using default parameters of the heatmap.2 function from gplots. The clusterProfiler R-package was used for enrichment analysis of DEGs using Hallmarks gene sets. CNAs were inferred from VCF files provided by Reddy and colleagues with gains and losses called as described in their publication (23).

The tables of read counts and metadata relative to the 553 samples sequenced by whole-genome RNA-seq in the GOYA trial were downloaded from GEO (accession number: GSE125966). Hierarchical clustering analysis of lymphoma samples based on the expression of the 18 KZFPs located in the Chr19q coexpression module was performed as previously detailed.

Processed RNA-seq data of tumor-derived cell lines (28) and IFNγ-treated monocytes were downloaded from the EMBL-EBI portal at E-MTAB-2706 and E-MTAB-7572, respectively.

To determine immune cell types from the tumor transcriptomes, CIBERSORTx analysis using the LM22 references was applied. The proportion of lymphoma cells was obtained by adding the relative proportions of cells from B-cell origin. CIBERSORTx analysis using LM4 as a reference was used to obtain the cell type-specific gene-expression levels of the 18 KZFPs of the Chr19q coexpression module across B-cell, myeloid-derived cell, CD4, and CD8 T-cell populations. The expression of these 18 KZFPs was not detected in CD8 T cells and was therefore not reported.

Data analysis

Unless otherwise specified, graphs were obtained by GraphPad Prism software version 9 (RRID:SCR_002798) or ggplot2 R-package.

RT-qPCR

Total RNA extraction was performed using the NucleoSpin RNA plus kit (Machery-Nagel) according to the manufacturer's recommendations. cDNA synthesis for qPCR was conducted using the Maxima H minus cDNA synthesis master mix (Thermo Scientific). Real-time quantitative PCR was performed using PowerUp SYBR Green Master Mix (Thermo Scientific) and run on a QuantStudioTM 6 Flex Real-Time PCR System.

RNA-seq libraries and downstream analyses

RNA-seq libraries were prepared using the Illumina Truseq Stranded mRNA kit. Libraries were sequenced in 75 or 100 bp paired-end formats on the Illumina Hiseq 4000 and NovaSeq 6000 sequencers,

respectively. RNA-seq reads were mapped to the hg19 and T2T-CHM13v2.0 human genome releases using hisat v2.1.0. Only uniquely mapped reads were used for counting over genes and repetitive sequence integrants (MAPQ > 10). Counts for genes and TEs were generated using featureCounts v2 and normalized for sequencing depth using the TMM method implemented in the limma package of Bioconductor (RRID:SCR_010943). Counts on genes were used as library size to correct both gene and TE expression. For repetitive DNA elements, an in-house curated version of the Repbase database (RRID:SCR_021169) was used. Differential gene expression analysis was performed using Voom as implemented in the Limma package of Bioconductor (29). *P* values were adjusted for multiple testing using the Benjamini–Hochberg method.

Cleavage under targets and tagmentation

Cleavage under targets and tagmentation (CUT&Tag) was performed as described (30) without modifications. For each mark, 150k cells were used per sample using the anti-H3K9me3 primary antibody (Active Motif, RRID:AB_2532132, RRID:AB_2532132), the anti-γH2AX primary antibody (Abcam, ab2893, RRID:AB_303388), and anti-rabbit IgG (Abcam, ab46540, RRID:AB_2614925) secondary antibody. A homemade purified pA-Tn5 protein (3XFlag-pA-Tn5-Fl, Addgene #124601, RRID:Addgene_124601) was produced and coupled with MEDS Oligos by the Protein Production and Purification of EPFL, as previously described (31). The purified recombinant protein was used at a final concentration of 700 ng/μL (1:250 dilution from homemade stock). Libraries were sequenced with 75 bp paired-end on the NextSeq 500 (Illumina). Reads were aligned to the T2T-CHM13v2.0 reference genome (32) using bowtie2 (RRID:SCR_016368). Only proper read pairs with MAPQ>10 were kept. CUT&Tag peaks were called using SEACR v1.3 on "stringent" mode and numeric threshold 0.01 for H3K9me3 and 0.001 for γH2AX, and merged with maximum distance allowed of 2,000 bp for H3K9me3 and 0 bp for γH2AX with bedtools 2.27.1 (RRID:SCR_006646). Encode blacklist for hg19 genome was lifted over to T2T-CHM13v2.0 genome and filtered out from the annotated peaks. Bedtools multicov was used to count mapped reads on the annotated and filtered peaks, and differential peak analysis was performed using Voom after library size correction (using the total number of aligned reads as a size factor) performed using the TMM method. Bigwig coverage tracks with the mean of replicate samples were generated using bedtools 2.27.1 and deeptools 3.3.1, and heat map representations of the coverage signal were performed using the computeMatrix function and plotHeatmap from deeptools 3.3.1.

EdU DNA synthesis monitoring flow cytometry

Lymphoma cells were pulse-labeled with 10 μmol/L 5-ethynyl-2′-deoxyuridine (EdU) for 20 minutes and subsequently fixed with 2% formaldehyde for 30 minutes at room temperature. EdU incorporation was detected using Click chemistry according to the manufacturer's instructions (Click-iT EdU Flow Cytometry Cell Proliferation Assay, Invitrogen). Cells were resuspended in 1× PBS (BioConcept) with 1% bovine serum albumin (BSA), 2 μg/mL DAPI, and 0.5 mg/mL RNase A for 30 minutes at room temperature and subsequently analyzed on a BD LSR II (Becton Dickinson) flow cytometer as previously described.

TrAEL-seq

In brief, 1 million cells transduced with ShScr and ShRNA.1 were collected 3 days after LV transduction. TrAEL-seq was performed as described (33). Preparation of TrAEL-seq adaptors: DNA oligonucleotides were synthesized and PAGE purified by Sigma-Genosys (Merck). The sequences were as follows:

1: [Phos]NNNNNNNNAGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTUGCGCAGGCCATTGGCC[BtndT]GCGCUACACTCTTTCCCTACACGACGCT.

2: [Phos]GATCGGAAGAGCACACGTCTGAACTCCAGTCUUUUGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T. Libraries were sequenced on an Illumina NextSeq 500 sequencer as 75 bp paired-end reads. All scripts used for the analyses of TrAEL-seq data can be found in the following git repository: https://github.com/fmartins/traelseq). UMI processed files were then aligned to the T2T-CHM13v2.0 genome using Bowtie2 (parameters: --local -D 15 -R 2 -N 1 -L 25 -i S, 1, 0.75; ref. 34). Alignment files (bams) were then deduplicated using umit_tools (parameters: --paired --extract-umi-method read_id --method unique). Reads were then summed in running windows using the bamCoverage function from deepTools (35). Count tables per region in the format of bedgraph files were then imported into R (36) for plotting and statistics. Bigwig files were computed using the bedGraphToBigWig software (37). For read polarity plots, the following formula was used: read polarity = (R − F)/(R + F), where F and R relate to the total forward and reverse read counts, respectively. To compute R and F, the following parameters were used in bamCoverage: for forward (--samFlagInclude 64 --samFlagExclude 16) and for reverse (--samFlagInclude 80).

Immunofluorescence

U2932 cells were fixed and permeabilized using True-Nuclear Transcription Factor Buffer Set (BioLegend) following manufacturer's instructions, incubated with primary anti-dsDNA antibody (1:100, Mybiosource, RRID:SCR_013493), anti-phospho-STING Ser365 (1:400, CST, RRID:AB_2827656), or anti-phospho-IRF3 Ser396 (1:100, CST, RRID:AB_2773013) for 30 minutes, washed, and incubated with secondary anti-Mouse IgG Alexa Fluor 488 (1:100, Invitrogen, RRID:AB_2534069) or anti-Rabbit IgG Alexa Fluor 568 (1:100, Invitrogen, RRID:AB_2534072) for another 30 minutes. Nuclei were stained with DAPI (4′, 6-diamidino-2-phenylindole, 1 μg/mL, Sigma-Aldrich). Images were acquired on a point-scanning confocal microscope SP8 Leica with an oil 40× objective 1.25 N.A. Images are Z-stacks of dimensions 256 × 256 in XY and spanning the entire cell in Z (voxel size 120 × 120 × 500 nm). Pinhole was set at 1 Airy Unit for an emission wavelength of 520 nm for all channels, acquired in sequential scan mode. DAPI was excited with a 405-nm laser (emission 420–470 nm), and its emission was collected between 420 and 470 nm. The excitation of the DAPI channel is used to make a transmission image. DsDNA (Alexa 488) was excited with a 488 nm laser, and its emission was collected between 500 and 530 nm. Phospho-STING and phospho-IRF3 (Alexa 568) were excited with a 552 nm laser, their emission was collected between 560 and 610 nm. DsDNA quantification estimation was performed in ImageJ/Fiji (RRID:SCR_002285). Each image is projected in Z using maximum intensity projection. The DAPI channel was subsequently median filtered (radius = 2 pixels), and the nuclear region was delineated by using a fixed intensity threshold, identical for all images. A band of 0.8 μm surrounding the nucleus is then defined by shrinking the nucleus region by 5 pixels, then expanding it by 0.8 micrometer, thus defining a cytosolic band region. Average intensities of DAPI and dsDNA channels were then measured within the shrank nuclear region and the cytosolic band. All experiments were reproduced at least once.

Phagocytosis assay

Human M1 macrophages (effector cells) were plated into 96-well flat clear bottom black walled culture plates (Thermo Scientific 165305) at a density of 1×10^4–2.5×10^4 cells/well in 50 μL complete

RPMI medium supplemented with M-CSF (50 ng/mL) overnight. U2932 cells (target cells) were transduced as described above, harvested on day 4, washed with pHrodo Wash Buffer (Sartorius), and labeled using pHrodo red (Sartorius) at a final concentration of 31.3 ng/mL for 45 minutes at 37°C in pHrodo Labeling Buffer (Sartorius), before being washed again and resuspended in complete RPMI medium. 3×10^4–7.5×10^4 target cells/well were then seeded over effector cells at a target:effector ratio of 3:1. 96-well plates were then placed into the IncuCyte S3 System (Sartorius) for time-lapse imaging. The pHrodo signal was measured by the acquisition of images from 3–5 fields/well every 1 to 1:30 hours for 24 hours.

Immunopeptidomics analysis

HLA-I peptides were immunopurified from U2932 cells. A first set of control and *ZNF587/417* KD cell pellets (30 million cells each) was used to generate a spectral library from data-dependent and data-independent acquisition methods (DDA and DIA, respectively). A second set of pellets (with at least 20 million cells) were acquired using the DIA method and analyzed by Spectronaut. The two sets of samples were processed as previously described to purify the HLA-I peptides (38). Briefly, anti-pan HLA-I monoclonal antibodies (W6/32) were purified from supernatant of HB95 hybridoma cells (ATCC HB-95) using protein-A sepharose 4B (Pro-A) beads (Invitrogen) and cross-linked to pro-A beads. Each sample was lysed in 1.5 mL PBS lysis buffer containing 0.25% sodium deoxycholate (Sigma-Aldrich), 0.2 mmol/L iodoacetamide (Sigma-Aldrich), 1 mmol/L EDTA, 1:200 Protease Inhibitors Cocktail (Sigma-Aldrich), 1 mmol/L phenylmethylsulfonylfluoride (Roche), 1% octyl-beta-D glucopyranoside (Sigma-Alrich), at 4°C for 1 hours and centrifuged at 4°C at 14,200 rpm for 50 minutes using a refrigerated table-top centrifuge (Eppendorf Centrifuge). The cleared lysates were loaded on wells of 96-well single-use filter microplates with 3-µm glass fibers and 25-µm polyethylene membranes (Agilent, 204495-100) containing the HB95 cross-linked beads. Once the lysates flew through, the plates were washed several times with solutions containing various concentrations of salts and two final washes of 2 mL of 20 mmol/L Tris-HCl pH8 using the Waters Positive Pressure-96 Processor (Waters). The HLA complexes bound to the beads were then directly eluted into preconditioned Sep-Pak tC18 100 mg Sorbent 96-well plates (Waters, 186002321) using 1% trifluoroacetic acid (TFA; Sigma Aldrich). After two washes of 1 mL 0.1% TFA, the HLA-I peptides were eluted using 25% acetonitrile (ACN; Sigma Aldrich) in 0.1% TFA. Peptides were dried using vacuum centrifugation (Concentrator plus, Eppendorf) and stored at −20 °C. Prior to LC-MS/MS analysis, iRT peptides (Biognosys) were spiked in each sample for retention time calibration through Spectronaut analysis. For each sample, dried peptides were resuspended in 12 µL, of which, 3 µL samples were injected for DDA and 2 µL for DIA. The Easy-nLC 1200 liquid chromatograph (Thermo Fisher Scientific) was coupled to a Q Exactive HF-X mass spectrometer (Thermo Fisher Scientific) for peptide tandem mass spectrometry data analysis. Peptides were separated on a homemade 50-cm analytical column with a tip size of ∼10 µm (OD = 365 µm, ID = 75 µm) and packed with ReproSil-Pur C18 (1.9-µm particles, 120 Å pore size, Dr Maisch GmbH). The peptide separation by chromatography was performed as previously described (39). For MS/MS data acquisition in DDA and DIA, ion sampling methods were used as described (40).

HLA typing and prediction of HLA binding affinity

HLA typing using HLA-HD v1.4.0 (41) on RNA sequencing data from U2932. For each sample, the consensus HLA haplotypes were determined as the consensus typing derived from all RNA-seq replicates. NetMHCpan-4.1 tool was used for predicting the binding affinity of the eluted peptides to the respective HLA alleles (42).

DDA and DIA data analysis

We used Spectronaut 15.7.220308.50606 (43) to build a spectral library from DDA and DIA MS/M data. The peptide identification search was done by Pulsar using an unspecific digestion type and a peptide length from 8–15 a.a. with a maximum variable modification of 5 per peptide. Acetylation of n-term and oxidation of methionine were used as variable modifications. The library was then generated using a peptide and PSM FDR ≤ 0.01 and protein FDR = 1 from a Pulsar search. The DIA MS/MS data were matched against the library as described (40), with the exception that this time iRT peptide calibration was used.

Data availability

The accession number for the RNA-seq, TrAEL-seq, and CUT&Tag data generated in this article is GEO: GSE229471. The data generated by others and analyzed in this study were obtained from EGA at EGAS00001002606, GEO at GSE125966, and EMBL-EBI portal at E-MTAB-2706, and E-MTAB-7572. All other raw data are available upon request from the corresponding author.

Results

A DLBCL transcriptome-wide association study reveals KZFPs predicting poor prognosis

To probe the potential involvement of KZFPs in DLBCL pathogenesis, we looked for a correlation between their levels of expression and disease outcome. For this, we used published RNA-seq data from 633 fresh-frozen paraffin-embedded DLBCL samples collected before the initiation of rituximab-based therapy (23). We conducted a univariate Cox proportional hazards regression analysis on the transcriptome of these samples to identify genes, the expression of which was either positively or negatively correlated with survival. Out of 23,560 expressed genes, 1,530 (6.5%) were found to be significantly associated with either a better ($n = 950$, 62.1%) or a worse ($n = 580$, 37.9%) outcome (**Fig. 1A**). KZFP genes were significantly enriched in prognosis-associated genes with 41 out of 395 (10.4%) expressed KZFPs associated with survival (1,530/23,560 vs. 41/395; Fisher exact test, $P = 0.0038$), 39 of them negatively. This stood in sharp contrast with the observed distribution of hazard ratios (HR) for all other TFs or remaining protein-coding genes, where good-prognosis hits dominated (**Fig. 1B**). Corroborating our results, a recent analysis of the same cohort found 12 of these 39 KZFP genes to be part of a KZFP-enriched poor prognosis gene expression signature (44).

Among the 41 *KZFP* gene hits of our screen, 39 were protein-coding, 19 of them primate-specific and expressed at higher levels than their evolutionary older counterparts in all disease subtypes (**Fig. 1C**; ref. 11), suggesting a possible (COO-independent) protumoral role for these proteins despite their young evolutionary age. *ZNF662* and *ZNF354C*, the two better-prognosis *KZFP* genes identified by our analysis, reside on chromosomes 3 and 5, respectively. Interestingly, the long arm of chromosome 19 (Chr19q) harbors 32 of the 39 (82.1%) poor prognosis *KZFP* genes (Supplementary Fig. S1A), a significant overrepresentation of this genomic region as the entire chromosome 19 only encodes slightly more than half of all KZFPs (231/395, 58.5%; Fisher exact test, $P = 0.0034$). Noteworthy, this particular region also houses the cyclin E1 oncogene, another prognosis-related gene identified through our screen. Cyclin E1 has previously been associated with the disruption of the DNA replication program and the

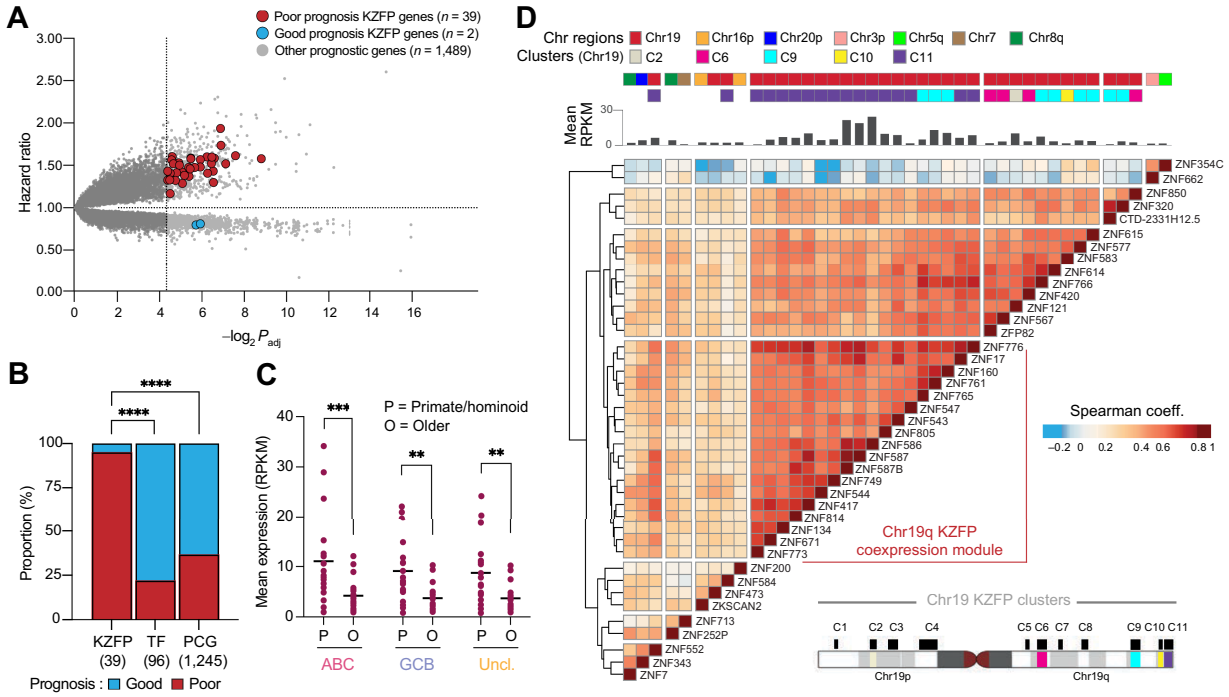


Figure 1.

A *KZFP* gene cluster as poor prognosis predictor in DLBCL. **A,** Volcano plot of HRs issued from univariate Cox regression analysis correlating individual gene expression with patient survival outcome are plotted against significance ($n = 612$ DLBCL RNA-seq with available survival data). The vertical dashed line indicates the threshold of significance (FDR < 0.05). The horizontal dashed line separates genes associated with an increased (>1) and reduced risk (<1) of death. **B,** Stacked bar plots showing the proportion of gene hits associated with good (blue) or poor prognosis (dark red). Genes were hierarchically categorized as protein-coding KZFPs, other TFs, and other protein-coding genes (PCG). Statistics, two-tailed Fisher exact test. **C,** Dot plot showing the mean expression in reads per kilobase of exon per million reads mapped (RPKM) of the 19 primate/hominoid-specific (P) and 20 evolutionary older (O) protein-coding KZFP gene hits correlated with survival across ABC ($n = 256$), GCB ($n = 281$), and unclassifiable ($n = 96$) DLBCL samples. Statistics, two-sided Mann–Whitney *U* test. **D,** Hierarchical clustering and heat map showing the Spearman correlation coefficients of pairwise comparison between the 41 KZFP gene hits based on their gene expression across 633 DLBCL samples—Euclidean distances and complete-linkage clustering method. The bar graph on top of the heat map shows the mean expression in RPKM and the upper annotation track shows the chromosomal location. The lower right ideogram displays KZFP gene hits positioning onto previously described Chr19q KZFP clusters (10). Statistical significance: **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$.

promotion of genomic instability when overexpressed in tumor cells (45). Additionally, it has been identified as a translocation partner with the IgH locus in specific cases of DLBCL (46). To evaluate if these *KZFP* genes were part of a single gene expression module, we asked which were systematically coexpressed in DLBCL samples. This led us to identify a subset of 18 *KZFPs* displaying a strong coexpression pattern. Fifteen of them were located within the C11 *KZFP* gene cluster located on the subtelomeric region of chromosome 19 (chr19q13.43) and the other 3 within the nearby C9 cluster (**Fig. 1D**; ref. 10). Strikingly, ZNF587, ZNF587B, and ZNF586, the 3 most highly expressed poor prognosis KZFPs from our screen were located over a 200 kb region encompassing 5 other KZFP hits, namely, ZNF814, ZNF417, ZNF671, ZNF776, and ZNF552 (Supplementary Fig. S1B). This prompted us to characterize the DLBCL subset overexpressing these coregulated unfavorable *KZFP* genes.

A *KZFP* gene cluster defines a DLBCL subset with cell-autonomous growth features

Through unsupervised hierarchical clustering of DLBCL samples based on the expression of these 18 coexpressed *KZFP* genes, we could delineate a subset encompassing about a quarter of all DLBCL samples ($n = 145$, 23%, **Fig. 2A**), to which we will refer to hereafter as $KZFP^{High}$ DLBCLs. To characterize differences between these samples and their $KZFP^{Low}$ counterparts, we conducted a gene set enrichment analysis (GSEA; ref. 47) for pathways cataloged in the Hallmark gene sets (48). $KZFP^{High}$ samples were characterized by a proliferative and highly metabolic state, as suggested by the upregulation of genes involved in cell-cycle regulation, DNA repair, mitotic spindle assembly, oxidative phosphorylation, and Myc response. Conversely, $KZFP^{Low}$ samples were characterized by the activation of IFN/inflammatory response and EMT (epithelial-to-mesenchymal transition) gene sets (**Fig. 2B**).

Using the metadata available for this patient cohort (23), we determined that individuals presenting with $KZFP^{High}$ DLBCL had survival rates twice shorter than those with a $KZFP^{Low}$ profile (**Fig. 2C**). $KZFP^{High}$ patients were slightly older (65 ± 13.2 vs. 59 ±16.2 years) and presented more frequently DLBCLs of the ABC subtype (47.2% vs. 34.7%; **Fig. 2D**), but were diagnosed at similar disease stages (**Fig. 2D**). To assess the independence of these KZFP risk groups and the outcome predicted by age, stage at diagnosis and COO, we conducted a multivariate Cox regression analysis, which confirmed that KZFP risk groups remained significantly associated with survival ($HR = 1.7$, $P = 0.001$). Intriguingly, $KZFP^{High}$ DLBCLs did not appear to present a lower response rate to first-line rituximab-based regimens (**Fig. 2D**), suggesting that patients in this group were instead more inclined to relapse after first achieving remission. To confirm the

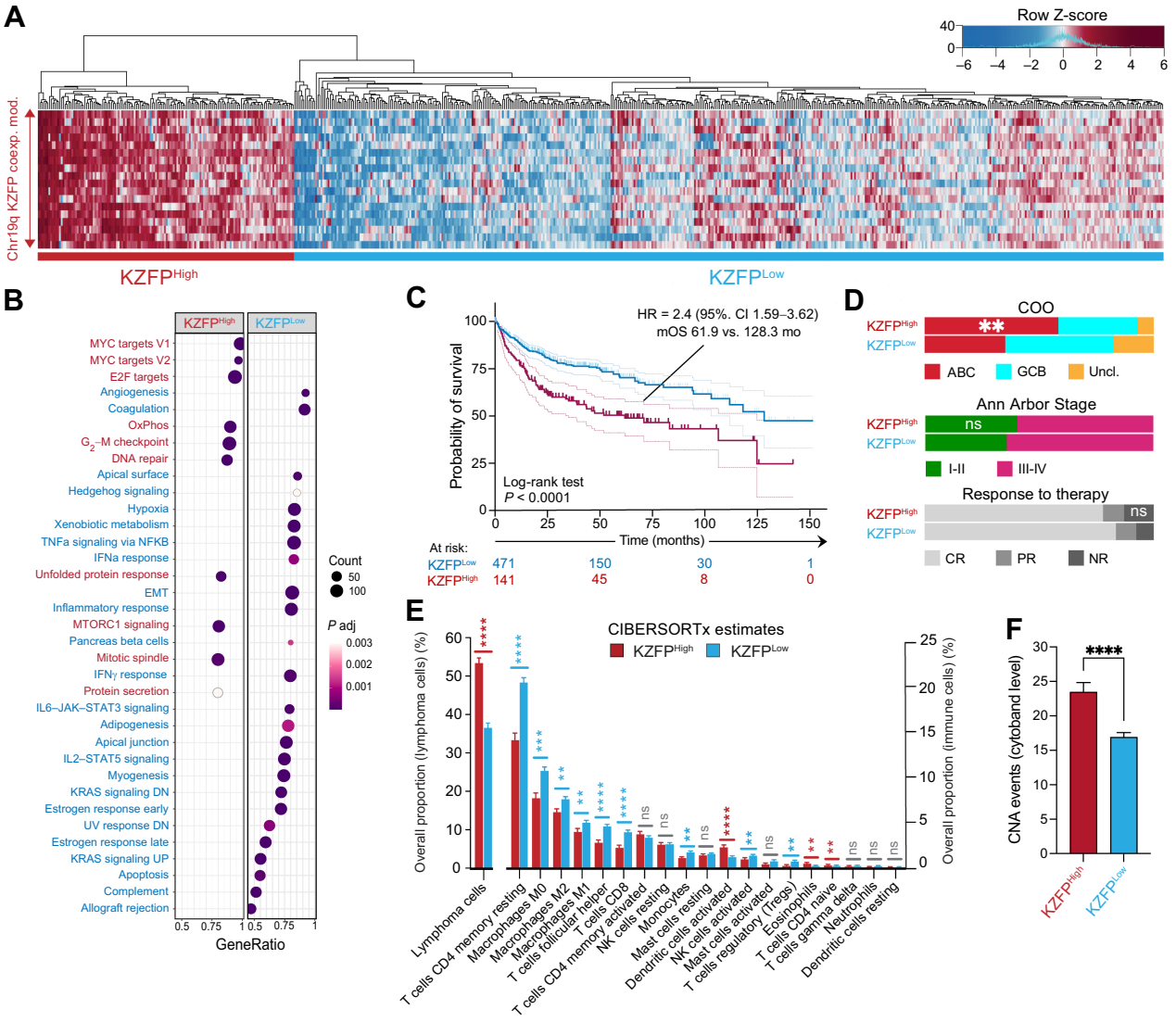
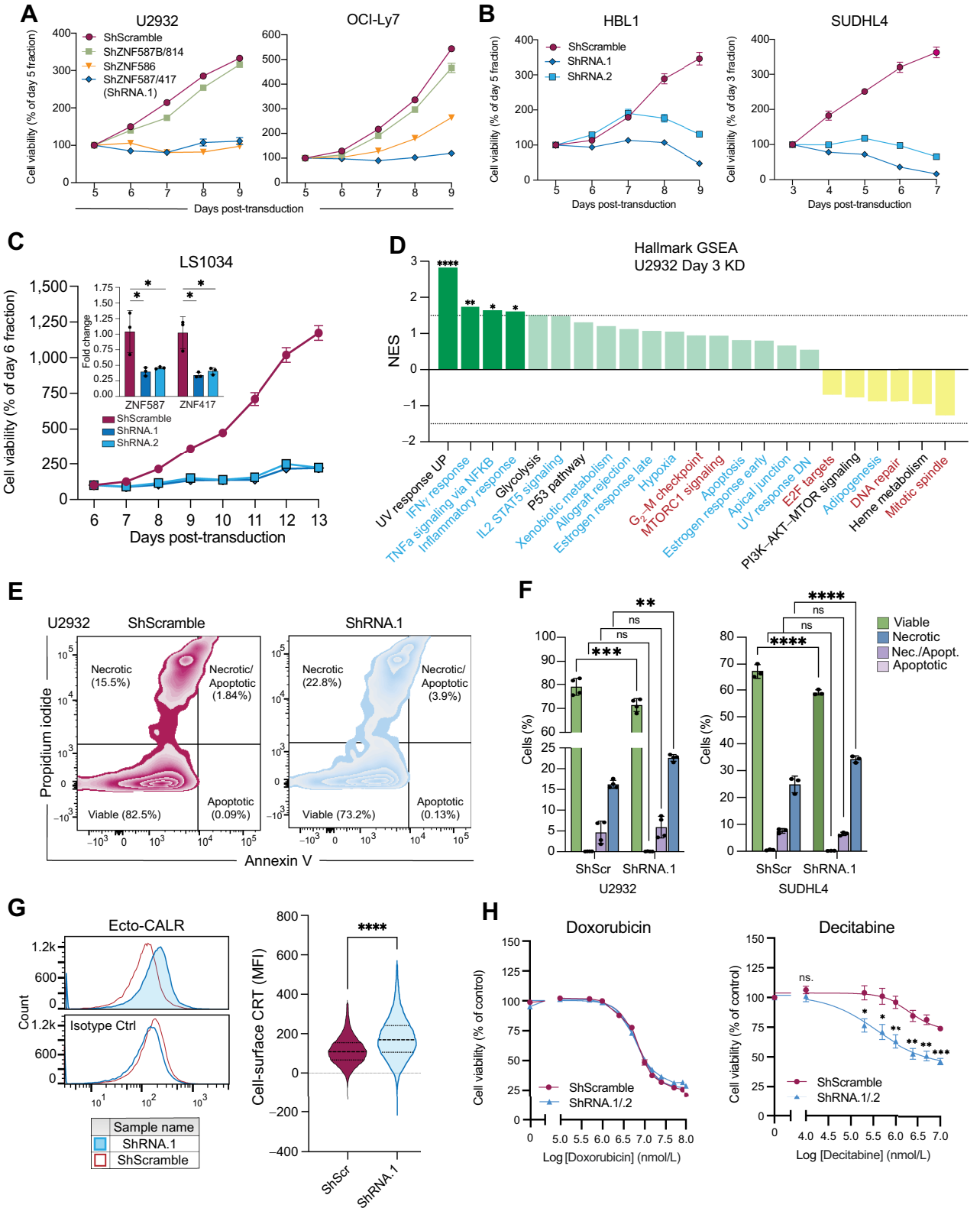


Figure 2.

KZFP^{High} DLBCLs display cell-autonomous growth features. **A,** Heat map of the unsupervised hierarchical clustering of 633 DLBCL samples based on the expression of the 18 KZFP genes of the Chr19q coexpression module—Euclidean distances and Ward's agglomeration method. **B,** Dot plots of significantly enriched Hallmark gene sets in KZFP^{High} vs. KZFP^{Low} DLBCLs. The *x*-axis represents the ratio of DEGs in these groups over the total number of genes in each set (gene ratio). Dot size represents the number of DEGs in each gene set (count) and the color scale represents the FDR adjusted *P* value. **C,** Kaplan–Meier survival curves of patients displaying a KZFP^{High} vs. KZFP^{Low} DLBCL signature as defined in **A**. Statistics, log-rank test. **D,** Stacked bar plots showing the proportion of the three different COO subtypes, the Ann Arbor stages, and the response to initial rituximab-based standard regimen among the patients belonging to each KZFP-defined group. CR, complete response; PR, partial response; NR, no response. Statistics, two-tailed Fisher exact test. **E,** Bar plot showing the differences in the mean proportion (± SEM) of lymphoma cells (left) and immune cell infiltrate (right) between KZFP^{Low} and KZFP^{High} DLBCLs using CIBERSORTx RNA-seq bulk deconvolution. Statistics, two-sided Mann–Whitney *U* test with Benjamini–Yekutieli correction. **F,** Bar plots showing the mean number (±SEM) of CNA detected by WES conducted on the same samples included in the transcriptomic analysis. **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$; ns, nonsignificant.

association of these 18 KZFPs with OS and disease progression (PFS), we used the transcriptomic data of an independent cohort of 553 DLBCL patients included in the GOYA trial, whose lymphomas were sequenced at diagnosis before rituximab/obinutuzumab—CHOP. Using the same unsupervised clustering methods, we found the same proportion of KZFP^{High} cases (28%) as in the Reddy and colleagues cohort (23%; Supplementary Fig. S2A). Among patients in the KZFP^{High} group, our analysis revealed a significant reduction in PFS, alongside the observed decline in OS already noted in the Reddy and colleagues cohort (Supplementary Fig. S2B and S2C). Notably, the GOYA cohort exhibited a higher proportion of GCB-DLBCLs, alongside a lower proportion of high-risk DLBCL cases, in contrast to the unselected Reddy and colleagues cohort (Supplementary Fig. S2E and S2F). These distinctions likely contributed to the observation that patients enrolled in the GOYA trial demonstrated a better OS compared with the Reddy and colleagues cohort (Supplementary Fig. S2D). These results confirm the robustness of this 18 KFZP gene signature as a promising prognostic biomarker in DLBCL.

This led us to hypothesize that *KZFP* expression may contribute to disease relapse not by causing primary chemoresistance of the bulk


A
U2932
OCI-Ly7
ShScramble
ShZNF587B/814
ShZNF586
ShZNF587/417 (ShRNA.1)
Cell viability (% of day 5 fraction)
Days post-transduction
B
HBL1
SUDHL4
ShScramble
ShRNA.1
ShRNA.2
Cell viability (% of day 5 fraction)
Cell viability (% of day 3 fraction)
Days post-transduction
C
LS1034
Cell viability (% of day 6 fraction)
Fold change
ZNF587
ZNF417
ShScramble
ShRNA.1
ShRNA.2
Days post-transduction
D
Hallmark GSEA
U2932 Day 3 KD
NES
UV response UP
IFNγ response
TNFa signaling via NFKB
Inflammatory response
Glycolysis
IL2 STAT5 signaling
P53 pathway
Xenobiotic metabolism
Allograft rejection
Estrogen response late
Hypoxia
G2-M checkpoint
MTORC1 signaling
Apoptosis
Estrogen response early
Apical junction
UV response DN
E2F targets
PI3K-AKT-MTOR signaling
Adipogenesis
DNA repair
Heme metabolism
Mitotic spindle
E
U2932
ShScramble
ShRNA.1
Necrotic (15.5%)
Necrotic/ Apoptotic (1.84%)
Apoptotic (0.09%)
Viable (82.5%)
Necrotic (22.8%)
Necrotic/ Apoptotic (3.9%)
Apoptotic (0.13%)
Viable (73.2%)
Propidium iodide
Annexin V
F
Viable
Necrotic
Nec./Apopt.
Apoptotic
Cells (%)
ShScr
ShRNA.1
U2932
SUDHL4
G
Ecto-CALR
Isotype Ctrl
Count
Sample name
ShRNA.1
ShScramble
Cell-surface CRT (MFI)
ShScr
ShRNA.1
H
Doxorubicin
Decitabine
Cell viability (% of control)
ShScramble
ShRNA.1/.2
Log [Doxorubicin] (nmol/L)
Log [Decitabine] (nmol/L)

population of lymphoma cells, but rather by contributing to their clonal diversity, thus facilitating the emergence of treatment-resistant subclones, and by fostering the immune evasion of residual disease. To explore this hypothesis, we applied a bulk RNA-seq deconvolution method (49) to DLBCL samples as a proxy for comparing the TME cell compositions of $KZFP^{High}$ and $KZFP^{Low}$ subtypes. This revealed a significant depletion of both innate and adaptative immune cells in $KZFP^{High}$ compared with $KZFP^{Low}$ DLBCL samples (**Fig. 2E**). This drop in immune gene signatures concerned T cells, activated natural killer (NK) cells, and monocyte-derived cells. To validate these observations, we applied the same bulk RNA-seq deconvolution method to the GOYA cohort, systematically comparing $KZFP^{High}$ and $KZFP^{Low}$ DLBCL groups. Our analysis consistently revealed marked differences in the immune landscape profile between these groups, except for activated dendritic cells and a few other minor cell types. Notably, this analysis confirmed the decreased proportion of macrophages and CD8 T cells, along with an increased proportion of lymphoma cells within the $KZFP^{High}$ group (Supplementary Fig. S2G). To ascertain whether the $KZFP^{High}$ signal might simply be a reflection of the higher proportion of lymphoma cells in this group, we conducted a detailed examination of the cell type–specific gene expression profile within lymphoma cells across both KZFP groups. This analysis revealed that despite the elevated proportion of lymphoma cells in the $KZFP^{High}$ group, they exhibited higher expression levels of the 18 KZFP genes, indicating a genuine transcriptional upregulation. Furthermore, no substantial modifications in the expression of these genes were observed within the myeloid and T-cell compartments (Supplementary Fig. S2H), except for notable changes involving ZNF586 in CD4 T cells, and minor fluctuations concerning ZNF671 and ZNF17/ZNF544 within CD4 T cells and myeloid cells, respectively.

These apparent differences in $KZFP^{High/Low}$ TME were further supported by the analysis of the mutational pattern of the KZFP-defined DLBCL groups, with $KZFP^{High}$ samples displaying significant enrichment for a recently defined high-risk molecular gene signature linked to *MYD88* and *CD79B* mutations, together with an increased prevalence of *BCL2* and *CDKN2A* alterations (Supplementary Fig. S2I; ref. 50). This mutational signature, known as the MCD subtype, has been shown to promote DLBCL immune evasion by reducing antigen presentation and NK cell activation among other mechanisms (51).

Furthermore, an examination of whole-exome sequencing (WES) data derived from the same sample cohort revealed a heightened occurrence of copy-number alterations (CNA) within $KZFP^{High}$ DLBCLs, indicative of increased clonal diversity (**Fig. 2F**). Notably, $KZFP^{High}$ samples exhibited a significant enrichment in copy-number gains affecting the 18 KZFP genes themselves, with 13 out of 144 (9%) cases, in contrast to 4 out of 489 (0.8%) cases observed in the $KZFP^{Low}$ group (Fisher exact test, $P \leq 0.0001$; Supplementary Fig. S2J). These combined results suggest that $KZFP^{High}$ DLBCLs exhibit a proliferative, genetically unstable, and cell-autonomous phenotype characterized by reduced immune infiltration and a heightened incidence of genetic alterations associated with more aggressive disease phenotypes.

ZNF587/417 depletion broadly impairs the growth and viability of highly proliferative cells

As a first step to investigate the impact of KZFPs on DLBCL homeostasis, we performed proliferation assays in U2932 and OCI-Ly7 cells, respectively representative of ABC- and GCB-DLBCL, transduced with lentiviral vectors expressing small hairpin RNAs (shRNA) against the 3 most highly expressed poor prognosis-associated KZFPs and their respective paralogs, or a scrambled sequence as a negative control (Supplementary Fig. S3A). *ZNF587B*/*ZNF814* and *ZNF417*/*ZNF587*, two pairs of paralogs with respectively 70% and 98% of amino acid homology, could be simultaneously repressed with a single shRNA targeting shared homologous regions in both paralog pairs. Although highly expressed in these cells (Supplementary Fig. S2B), *ZNF587B/814* could be knocked down without affecting the growth of the U2932 ABC or OCI-Ly7 GCB cell lines (**Fig. 3A**). In contrast, ZNF586 depletion stopped the proliferation of U2932 but was less efficient in decreasing OCI-Ly7 cell growth, whereas *ZNF587/417* knockdown (KD) triggered abrupt growth arrest in both settings (**Fig. 3A**). This result was confirmed with a second shRNA targeting *ZNF587/417* with similar KD efficiency (Supplementary Fig. S3A and S3C) and in two additional DLBCL cells lines, with growth arrest induced in both HBL1 (ABC) and SUDHL4 (GCB) cells, for the latter with a marked reduction in cell viability (**Fig. 3B**), in spite of the low expression level of these genes at baseline compared with the other tested DLBCL cell lines (Supplementary Fig. S3B). This effect was not restricted to lymphoma cells, because *ZNF587/417* depletion similarly impaired the growth of several other cancer cell lines derived from hematologic or solid tumors independently of their baseline expression level (Supplementary Fig. S3F), including HL-60 acute myeloid leukemia, K562 chronic myelogenous leukemia SNU16 gastric carcinoma, as well as SW480, HT29, and LS1034 colorectal cancer cells (**Fig. 3C**; Supplementary Fig. S3D). In contrast, it did not affect the growth of primary foreskin fibroblasts. However, it impaired the proliferation of activated human CD4 T cells (Supplementary Fig. S3E), suggesting an effect linked to the proliferation rate of cells and underlying cell viability rescue mechanisms, rather than their normal or transformed phenotype.

To understand the mechanisms underlying this observation, we performed RNA sequencing (RNA-seq) on U2932 cells harvested

Figure 3.

ZNF587/417 depletion impairs cell growth and viability. **A,** MTT proliferation assays of U2932 and OCI-Ly7 cells upon LV transduction with shRNAs targeting *ZNF587B/814*, *ZNF587/417*, and *ZNF586* or a control shRNA (shScr). **B,** MTT proliferation assays for HBL1 and SUDHL4 upon LV transduction with two different shRNAs targeting *ZNF587/417* (shRNA.1/.2) or a control shRNA. **C,** MTT proliferation assays of LS1034 cells upon LV transduction with shRNAs targeting *ZNF587/417* or a control shRNA (shScr). Top left corner, quantitative RT-PCR analysis of ZNF587 and ZNF417 mRNAs after LV transduction with two shRNAs targeting ZNF587/ZNF417 paralog pair in LS1034 cells. Statistics, Student *t* test. **D,** Waterfall plot of Hallmark GSEA signatures from RNA-seq data ranked by their normalized enrichment score (NES). Green, enriched signatures (NES > 0); yellow, depleted signatures (NES < 0). The *P* value cutoff <0.05 is indicated by a dotted line. Hallmark gene sets enriched in $KZFP^{High}$ or $KZFP^{Low}$ DLBCLs are highlighted in dark red and blue, respectively. **E,** Representative flow cytometry plot using Annexin V–APC/PI staining for cell death characterization in U2932 cells. **F,** Bar plots showing the percentage of cells undergoing early (PI^-/aV^+) vs. late apoptosis (PI^+/aV^+) and necrosis (PI^+/aV^-) in U2932 and SUDHL4 cells, 3 days after transduction with shRNA.1 targeting *ZNF587/417* or a control shRNA (mean ± SD of two independent replicates). Statistics, two-way ANOVA with Sidak multiple comparison test. **G,** Representative flow cytometry histogram (top) and violin plot quantification (bottom) of median fluorescence intensity (MFI) of calreticulin (CALR) expression on U2932 KD and control cells assessed 3 days after LV transduction. At least 35,000 cells were assessed per condition. Statistics, two-sided Mann–Whitney *U* test. **H,** Drug-sensitivity assay in U2932 cells assessed by MTT assay. After 3 days of LV transduction, KD and control U2932 cells were exposed for 72 hours to varying doses of doxorubicin (left) or decitabine (right). Values represent mean ± SEM cell viability as a percentage of DMSO-treated control samples (shScr). Experiments were performed in triplicate using ShRNA.1 and duplicate using ShRNA.2, with viability curves showing pooled results. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$; ns, nonsignificant.

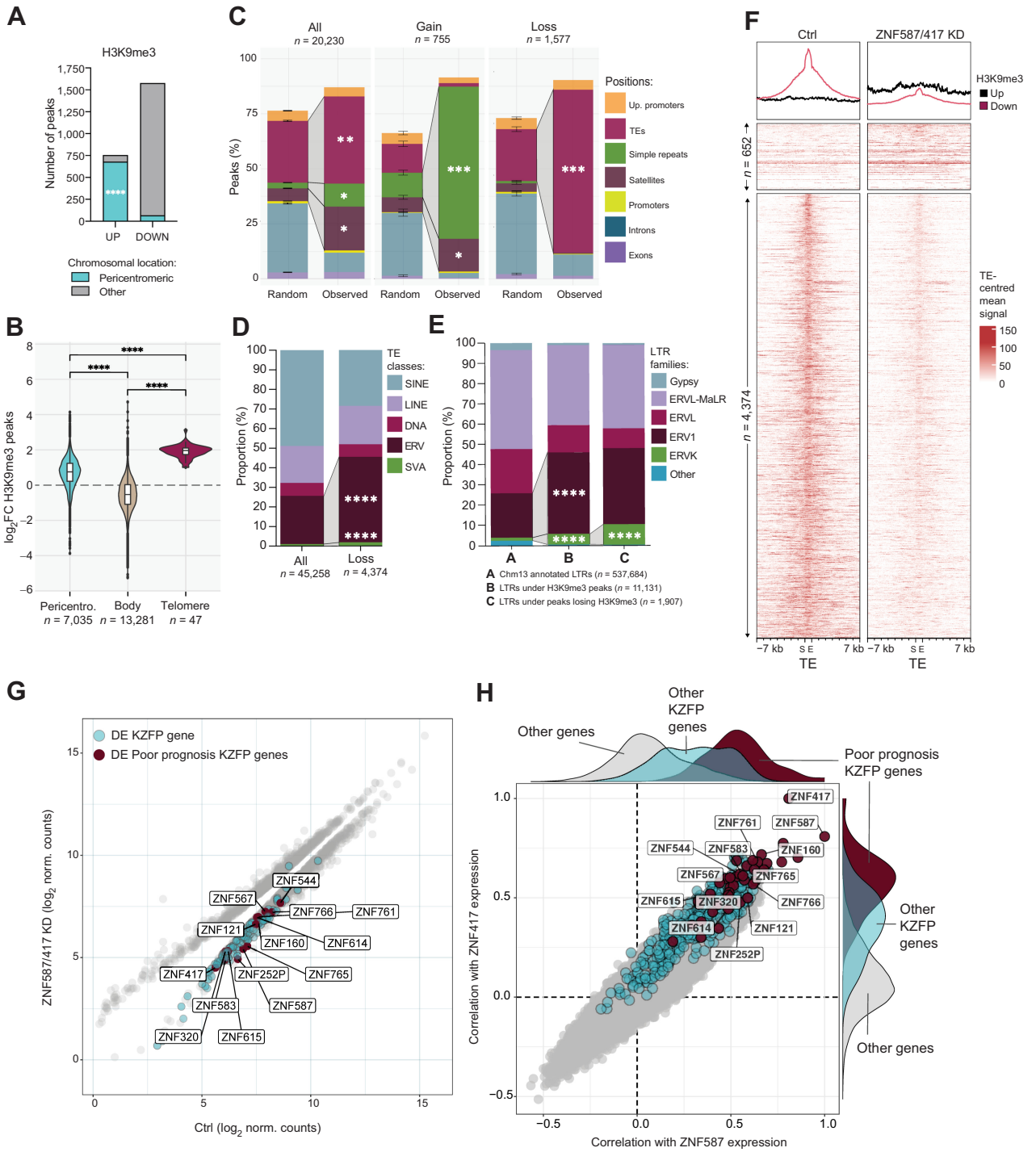


Figure 4.

ZNF587/417 depletion alters the heterochromatin landscape. **A,** Number of Cut&Tag peaks with significant (FDR <0.05) gain (UP) or loss (DOWN) of H3K9me3, 3 days after transduction with shRNA.1 targeting *ZNF587/417* or a control shRNA (from two independent replicates). **B,** Violin plots of the distribution of the $\log_2$-fold change (shRNA.1 vs. control) of H3K9me3 signal over called peaks in the listed genomic features. Statistics, two-sided Mann–Whitney *U* test. **C,** Observed genomic distribution of all H3K9me3 peaks (left) and peaks displaying a significant gain (center) or loss (right) of H3K9me3 signal vs. randomized peaks, over listed genomic features (mean value of duplicates ± SD from 1,000 permutations; FDR < 0.05). Genomic features overlapping with >50% an H3K9me3 peak were assigned in the following order: gene promoters (up to 1,000 bp upstream from the transcription start site), up.promoters (up to 5,000 bp upstream from the transcription start site), exons, TEs, satellites, simple repeats, and introns. Statistics, chi-squared test, alternative: greater. **D,** Proportion of TE classes overlapping all H3K9me3 called peaks (left) or peaks with a significant loss of H3K9me3 (FDR < 0.05; right). Asterisks indicate a significant enrichment compared with all H3K9me3 called peaks. Statistics, two-tailed Fisher exact test. (*Continued on the following page.*)

3 days after *ZNF587/417* KD. We identified 332 differentially expressed genes (DEG, fold change > 2 and FDR <0.05), 178 (53.6%) of which were downregulated (Supplementary Fig. S3G). GSEA revealed UV response UP as the most affected hallmark among DEGs (**Fig. 3D**). It encompassed several cyclins (*CCND3, CCNE1*), cyclin-dependent kinase inhibitors (*CDKN1A, CDKN1B, CDKN1C*), and TP53-responsive genes (*TP53I3*; Supplementary Fig. S2F). Reminiscent of the form of cell death induced by irradiation, Annexin V and PI staining revealed an increase in PI^+/AV^- cells, a pattern indicative of necrosis, upon *ZNF587/417* KD in U2932 and most prominently SUDHL4 cells (**Fig. 3E** and **F**). Moreover, ZNF587/417-depleted U2932 cells presented features of immunogenic cell death (ICD), with upregulation of cell-surface calreticulin (CALR), a potent "eat me signal" for mononuclear phagocytes (**Fig. 3G**).

We conducted drug-sensitivity assays to assess potential synergies between ZNF587/417 KD and doxorubicin, the main agent of the CHOP backbone or the hypomethylating agent decitabine. ZNF587/417 KD lymphoma cells did not reveal a higher sensitivity toward the first of these drugs (**Fig. 3H**), in line with our finding that all patients in this cohort responded equally well to first-line treatment (**Fig. 2D**). However, U2932 KD cells were more sensitive to decitabine than control cells (**Fig. 3H**), consistent with a cumulative impact of H3K9me3 and DNA methylation loss, as previously observed upon combining H3K9me3 methyltransferase G9A and DNA methyltransferase inhibitors in ovarian cancer cell lines (52).

ZNF587/417 depletion leads to a broad downregulation of KZFPs and alters the heterochromatin landscape of lymphoma cells

ZNF587 and ZNF417 are both KAP1-binding KZFPs that secondarily recruit the methyltransferase SETDB1 responsible for catalyzing the deposition of H3K9me3, a repressive histone mark promoting heterochromatin formation at targeted loci. The two paralogs recognize, almost exclusively, overlapping sets of primate-restricted TEs, including members of the ERVK and ERV1 classes of ERVs as well as many SVA integrants (24). To ask if the phenotype observed upon *ZNF587/417* KD cells correlated with changes in their chromatin status, we performed genome-wide profiling of H3K9me3 using CUT&Tag (30) in U2932 cells at day 3 of KD. Similar peak numbers ($n = 26{,}841$) and locations were detected in control and KD cells, but in ZNF587/417-depleted cells peak intensity significantly decreased at 1,577 and increased at 755 locations, most of the latter in pericentromeric regions (89.9% vs. 10.07%, $P < 0.0001$; **Fig. 4A**). Measurement of H3K9me3 across all called peaks confirmed a marked increase at pericentromeric regions and telomeres, contrasting with a global decrease across the rest of the genome following *ZNF587/417* KD (**Fig. 4B**).

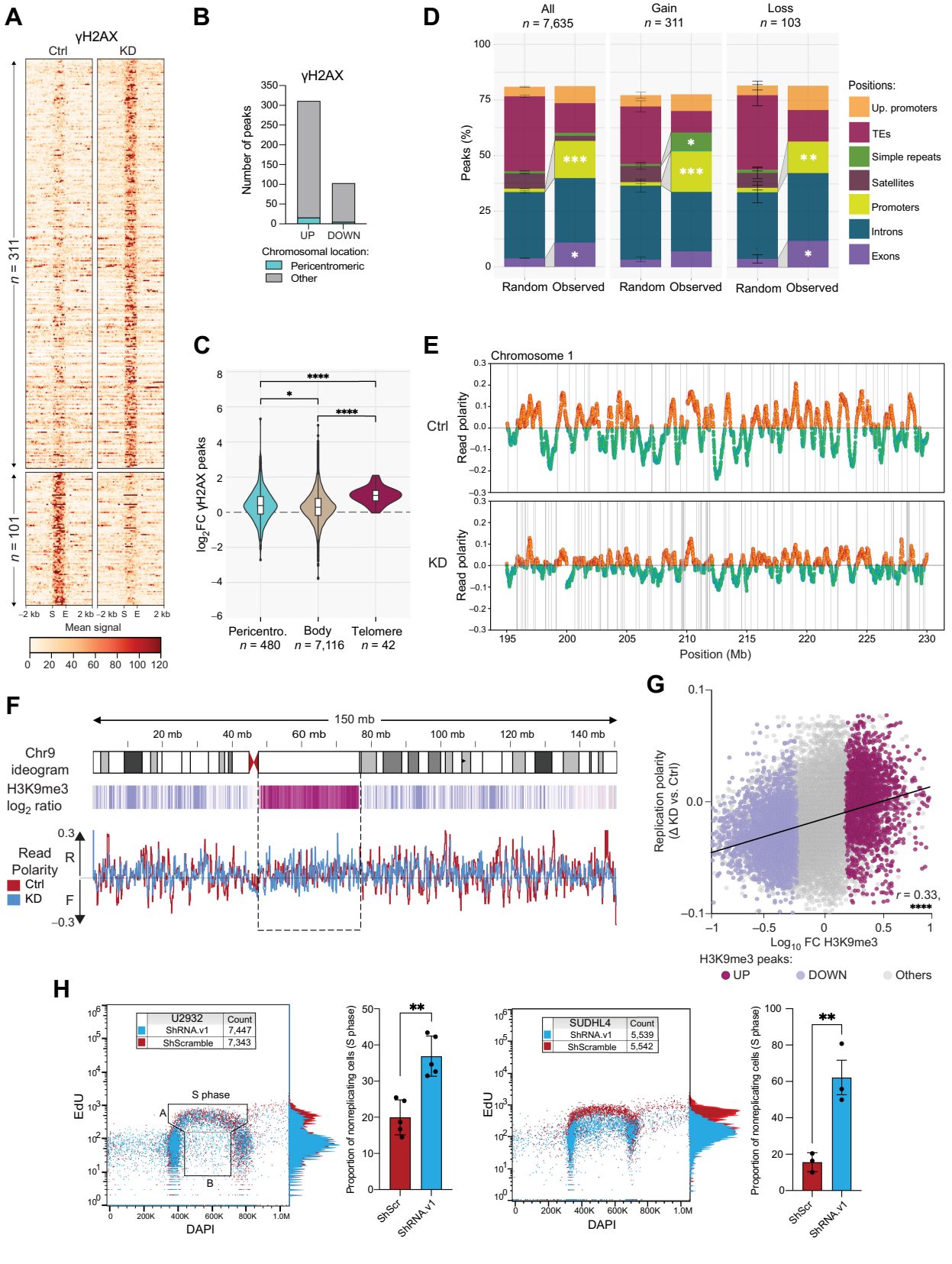
Although H3K9me3 peaks in control U2932 cells were enriched at TEs, simple repeats and satellite repeats, the KD triggered a preferential loss of this repressive mark at TEs and a gain at the other two categories of repeats (**Fig. 4C**). Among TEs, the most significant losers of H3K9me3 upon *ZNF587/417* KD belonged to the ERV and SVA subgroups (**Fig. 4D**), notably ERVK integrants previously identified as the preferential ERV targets of these KZFPs (**Fig. 4E**; ref. 24). When some increase in H3K9me3 was noted over TE inserts, it likely reflected spreading of this mark from surrounding satellite/simple repeats (**Fig. 4C** and **F**). Accordingly, peaks gaining H3K9me3 were on average substantially wider than TE-centered peaks losing this mark (Supplementary Fig. S4A). Still, the loss of H3K9me3 over a range of TEs extending beyond the sole known targets of ZNF587/417 suggested that additional factors contributed to this process. Confirming this suspicion, our RNA-seq analysis revealed the downregulation of another 100 *KZFP* genes (FDR <0.05) in U2932 cells 3 days after ZNF587/417 KD (**Fig. 4G**), including 12 poor prognosis *KZFP* gene hits. By contrast, only one *KZFP* gene was found upregulated in *ZNF587/417* KD cells at this time point. This downregulation was even broader at day 6 of KD, encompassing 184 and 133 *KZFP* genes (FDR <0.05) besides *ZNF587* and *ZNF417* in U2932 and OCI-Ly7 cells, respectively (Supplementary Fig. S4B). Although the reason for such broad downregulation is unknown, we observed that the expression of *ZNF587/417* was highly correlated with that of other *KZFP* genes in our DLBCL cohort (**Fig. 4H**), pointing to common regulatory mechanisms within this gene family. Supporting this hypothesis, a similar downregulation profile was found when analyzing the transcriptome of IFNγ-treated human macrophages (Supplementary Fig. S4C; ref. 53), suggesting a link with the IFN/inflammatory response.

ZNF587/417 depletion triggers RS in cancer cells

Our data pointed to a link between ZNF587/417 depletion and a UV irradiation-like transcriptional response. To probe this issue further, we measured the impact of *ZNF587/417* KD on Ser-139 H2AX phosphorylation (γH2AX) and p21 (*CDKN1A*) in U2932 and OCI-Ly7 cells 3 days after KD, which revealed a pronounced increase in these markers of DNA damage response (DDR; Supplementary Fig. S5A). We proceeded to the same analysis in selected solid tumor cell lines displaying a range of growth sensitivity to *ZNF587/417* KD. Depleting these KZFPs triggered an abrupt increase in p21 and γH2AX in LS1034 colorectal cancer cells (Supplementary Fig. S5A and S5B), the proliferation of which was most severely affected in this setting (**Fig. 3C**). In contrast, in SNU16 gastric and SW480 colorectal cancer, where growth impairment had been noted to be milder (see Supplementary Fig. S3D), baseline levels of γH2AX were high and did not further increase upon *ZNF587/417* knockdown (Supplementary Fig. S5B). It indicates that these cells tolerate higher levels of DNA damage than U2932, OCI-Ly7, or LS1034 cells, having likely evolved compensatory mechanisms that also dampen the impact of ZNF587/417 depletion.

To determine whether this DDR was concentrated in particular regions of the genome, we performed CUT&Tag for γH2AX in U2932 cells after 3 days of *ZNF587/417* KD. Peak numbers ($n = 7{,}635$) and locations were similar in control and KD cells, but γH2AX levels were significantly increased over 311 loci and decreased over 101 loci upon *ZNF587/417* KD (**Fig. 5A** and **B**), including 14 upregulated and 5 downregulated peaks located at pericentromeric regions (**Fig. 5B**). Although the number of peaks with a significant γH2AX increase was not enriched at those regions, the peaks located at pericentromeric regions ($n = 480$) and telomeres ($n = 42$) presented a higher global

(*Continued.*) **E,** Same as **D** over LTR families, including peaks annotated in the T2T-CHM13v2.0 genome release (left). Statistics, two-tailed Fisher exact test. **F,** Profile heat maps centered on TEs (± 7 kb) displaying significant change in H3K9me3 Cut&Tag signal (>50% overlap; mean signal coverage from duplicates). **G,** Scatter plot of RNA-seq data from ZNF587/417 KD versus control cells at day 3 of KD in U2932, outlining DEGs (gray dots, FDR < 0.05), poor prognosis KZFP genes (dark red dots), and other KZFP genes (turquoise dots). **H,** Scatter plot of Spearman correlation coefficients between ZNF417 (*y*-axis) or ZNF587 (*x*-axis) expression, and each gene expressed among the 633 DLBCLs analyzed transcriptomes ($n = 23{,}560$), outlining poor prognosis KZFP genes (dark red dots), other KZFP genes (turquoise dots), and other genes (gray dots), with density plots showing their distribution on each side. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$.


A
γH2AX
Ctrl
KD
n = 311
n = 101
−2 kb S E 2 kb
Mean signal
0 20 40 60 80 100 120
B
γH2AX
Number of peaks
UP
DOWN
Chromosomal location:
Pericentromeric
Other
C
log2FC γH2AX peaks
Pericentro. n = 480
Body n = 7,116
Telomere n = 42
D
All n = 7,635
Gain n = 311
Loss n = 103
Peaks (%)
Random Observed
Positions:
Up. promoters
TEs
Simple repeats
Satellites
Promoters
Introns
Exons
E
Chromosome 1
Ctrl
KD
Read polarity
Position (Mb)
F
150 mb
20 mb 40 mb 60 mb 80 mb 100 mb 120 mb 140 mb
Chr9 ideogram
H3K9me3 log2 ratio
Read Polarity
Ctrl
KD
R
F
G
Replication polarity (Δ KD vs. Ctrl)
Log10 FC H3K9me3
r = 0.33, ****
H3K9me3 peaks:
UP
DOWN
Others
H
U2932 Count
ShRNA.v1 7,447
ShScramble 7,343
S phase
EdU
DAPI
Proportion of nonreplicating cells (S phase)
ShScr
ShRNA.v1
SUDHL4 Count
ShRNA.v1 5,539
ShScramble 5,542

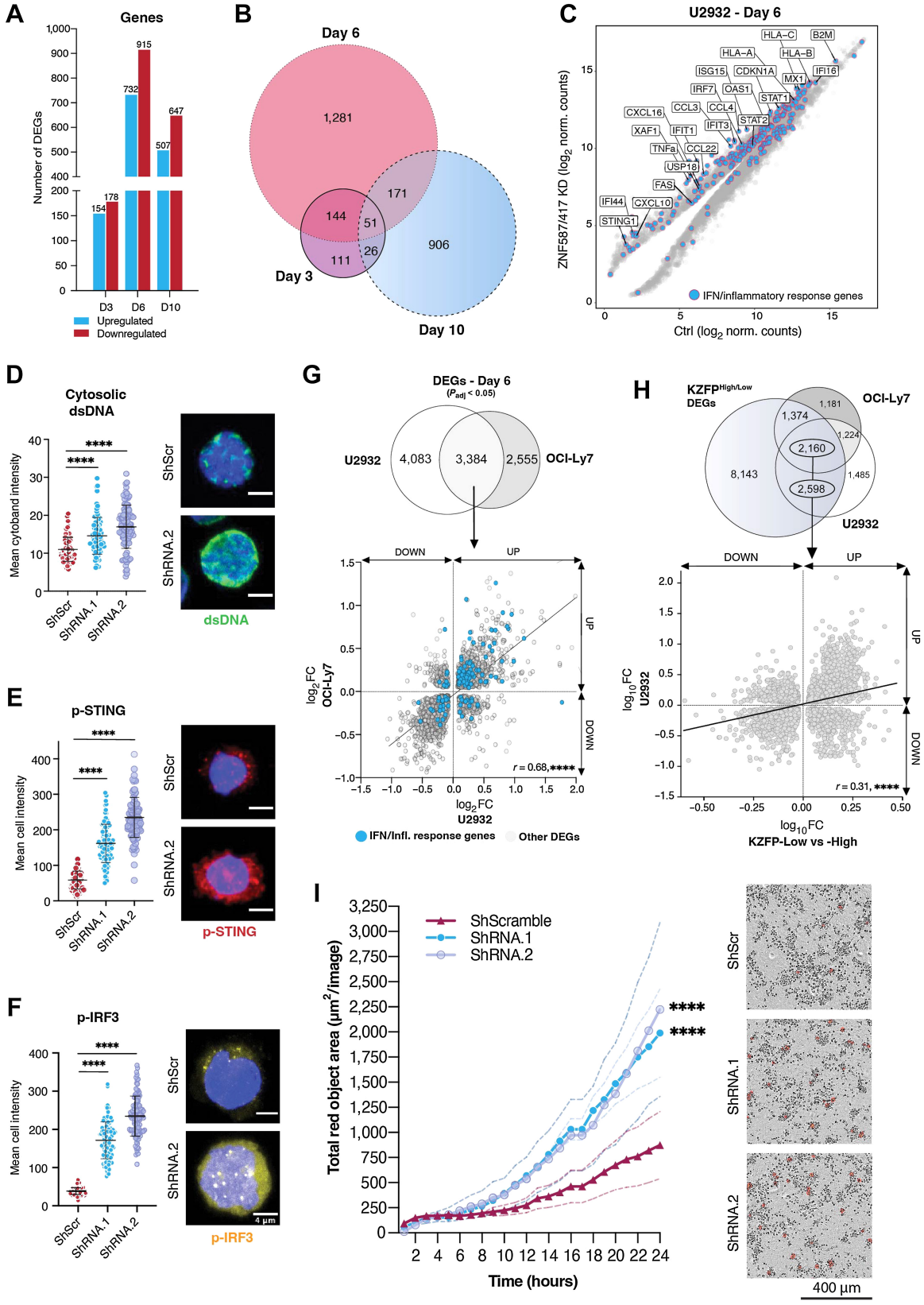
fold increase in γH2AX compared with other peaks upon *KZFP* KD (**Fig. 5C**). This reflected a significant enrichment in γH2AX-gaining peaks overlapping with simple repeats compared with other γH2AX-bearing loci (**Fig. 5D**). These results suggested that, when *ZNF587/417* are depleted, DNA damage was exacerbated at preexisting γH2AX hotspots rather than at TE loci targeted by these KZFPs. Because exacerbated DDR at preexisting γH2AX hotspots has been described with RS inducers such as hydroxyurea or aphidicolin (54), we characterized DNA replication fork progression at the genome-wide level using Transferase-Activated End Ligation sequencing (TrAEL-seq). We observed that *ZNF587/417* KD cells exhibited a global loss of replication polarity, indicating a lower frequency of diverging replication forks related to an increase in the number of replication origins (**Fig. 5E**). The excessive firing of replication origins is a known source of RS, notably due to the exhaustion of limiting factors necessary for replication fork progression (55) and collisions between replication and transcription (45), but it can also be a protective mechanism from RS triggered by primary stalling of replication forks (56). In support of the former hypothesis, we observed a strong correlation between H3K9me3 changes and replication fork polarity in *ZNF587/417* KD cells (**Fig. 5F** and **G**; Supplementary Fig. S5C and S5D), arguing in favor of a link between these two alterations. Strikingly, genomic regions with a loss in H3K9me3 exhibited a decrease in replication directionality, whereas the opposite was observed where H3K9me3 was gained (**Fig. 5F** and **G**), that is, predominantly at pericentromeric regions (**Fig. 5F**; Supplementary Fig. S5C). To probe if the cell growth impairment observed upon ZNF587/417 depletion was due to RS, we assessed DNA replication by measuring the incorporation of the thymidine analogue EdU in control and knockdown U2932 and SUDHL4 cells by flow cytometry. It revealed a decrease in EdU incorporation, along with a proportional increase of nonreplicating cells during the S phase of the cell cycle (**Fig. 5H**), thus confirming RS-related arrest.

ZNF587/417 depletion leads to cell-intrinsic inflammation

At day 6 of KD, 732 genes were upregulated and 915 downregulated (FC>2, FDR <0.05; **Fig. 6A**), including 195 of the 332 (58.7%) genes already noted as differentially expressed at day 3 (**Fig. 6B**). The transcriptome was then characterized by the upregulation of IFN-responsive and other inflammatory genes, and by an enrichment in Hallmark gene sets previously noted in $KZFP^{Low}$ DLBCL samples (Supplementary Fig. S6A). Interestingly, IFN-related DNA damage signature (IRDS) genes, such as *STAT1*, *IRF7*, *IFIT1/3*, *MX1*, *ISG15*, and *IFI44*, and *OASL* family members were upregulated at days 6 and 10, suggesting that the observed inflammatory response was at least partially related to genomic instability (**Fig. 6C**; ref. 57). TEs being well-known sources of cell-intrinsic inflammation, we monitored their transcription during these events. Even though some TE loci were detected as differentially expressed at day 6, frank TE upregulation became apparent only at day 10 after KD (Supplementary Fig. S6B). Furthermore, the TEs significantly upregulated at day 6 belonged to the SINE group, probably induced by the ongoing genotoxic stress (58). In contrast, most TEs significantly induced at day 10 were ERVs (Supplementary Fig. S6C), notably HERV9, HERV9N, HERV9NC, and LTR12C subfamily members of the primate-specific ERV1 family harboring ZNF587/417-binding motifs (Supplementary Fig. S6D). Of note, these ERV subsets already displayed an increased expression at day 6, albeit not reaching statistical significance at the TE family level (Supplementary Fig. S6C). Irrespectively, our transcriptome analyses revealed that IFN/inflammatory-related genes were already upregulated at day 3 of *ZNF587/417* knockdown, still in the complete absence of TE induction (**Fig. 3D**; Supplementary Fig. S6B). Therefore, TE transcripts and their products could not be the source of the inflammation initially observed upon KZFP depletion. The accumulation of cytosolic DNA molecules due to genotoxic stress is another known inducer of IFN response, triggered upon activation of the cGAS–STING DNA sensing pathway by micronuclei or free DNA fragments generated, respectively, in the context of lagging chromosomes or replication defects (59, 60). At day 3 of *KZFP* KD, U2932 cells presented a significant accumulation of cytosolic dsDNA (**Fig. 6D**) as well as of the dsDNA sensing intermediate mediator phospho-STING (**Fig. 6E**), together with the phosphorylation and nuclear translocation of IRF3, one of its main downstream effectors (**Fig. 6F**). Accordingly, upregulation or IRF3 target genes (61) was recorded at day 6 of *ZNF87/417* KD (Supplementary Fig. S6E). Furthermore, transcripts encoding CXCL10, a chemokine promoting the attraction of mononuclear phagocytes, T cells, and NK cells, were upregulated starting from day 3 up to day 10 following *ZNF587/417* KD, and we could confirm increased levels of this cytokine in the cell supernatant (Supplementary Fig. S6F). Induction of a cell-intrinsic inflammatory response was not restricted to ABC cells, as upregulation of IFN-related responsive genes similarly occurred in GCB-related OCI-Ly7 cells following *ZNF587/417* KD (**Fig. 6G**; Supplementary Fig. S6G), thus further highlighting that the effect of KZFP depletion

Figure 5.

ZNF587/417 depletion triggers RS. **A,** Profile heat maps of significantly changing γH2AX Cut&Tag peaks ($P < 0.05$) in U2932 control and KD conditions, 3 days after transduction (mean signal coverage from duplicates). **B,** Number peaks with a significant gain (UP) or loss (DOWN) in γH2AX in U2932 KD cells vs. control ($P < 0.05$). **C,** Violin plots of the distribution of the $\log_2$-fold change (shRNA.1 vs. control) of γH2AX signal over called peaks in the listed genomic features. Statistics, two-sided Mann–Whitney *U* test. **D,** Observed genomic distribution of all γH2AX peaks (left) and peaks displaying a significant gain (center) or loss (right) of γH2AX signal vs. randomized peaks, over listed genomic features (mean value of duplicates ± SD from 1,000 permutations; FDR < 0.05). Genomic features overlapping with a >50% γH2AX peak were assigned in the following order: gene promoters (up to 1,000 bp upstream from the transcription start site), up.promoters (up to 5,000 bp upstream from the transcription start site), exons, TEs, satellites, simple repeats, and introns. Statistics, chi-squared test, alternative: greater. **E,** TrAEL-seq read polarity plots over chromosome 1 in U2932 control and KD conditions, 3 days after transduction. Read polarity was calculated using a 250 kb sliding window spaced every 10 kb across the genome—see Materials and Methods. Orange and green points represent sliding windows with a reverse or forward strand bias, respectively. Vertical gray lines represent shifts from forward to reverse bias, highlighting strong replication origins. **F,** Top, heat map track of $\log_2$ ratio of Cut&Tag H3K9me3 signal between control and KD conditions surrounding the pericentromeric region of chromosome 9. Significant H3K9me3 gain is highlighted by the dashed rectangle. Bottom, overlapped read polarity plots of control (dark red) and KD (blue) conditions. **G,** Scatter plot of replication polarity vs. H3K9me3 Cut&Tag signal changes. *X*-axis, $\log_{10}$-fold change of all H3K9me3 peaks called in U2932 cells. *Y*-axis, replication polarity difference between KD and control conditions calculated from the absolute value of the percentage of reverse reads in a 2 Mb window surrounding the center of each H3K9me3 peak. Peaks with a significant gain in H3K9me3 are shown in red-violet, the one with a significant H3K9me3 loss in pastel purple, and others in gray. **H,** Representative flow cytometry analysis of the replication signal (EdU incorporation intensity; *y*-axis) and DNA content (DAPI staining; *x*-axis) in U2932 and SUDHL4 KD (blue) vs. control cells (dark red), 3 days after transduction. The *y*-axis side histogram shows changes in EdU intensity between KD (blue) and control (dark red) conditions. S phase cells are defined as events located within the A+B gates; A area, replicating cells; B area, nonreplicating cells. Right, barplots depicting the mean proportion (±SD) of nonreplicating S phase cells calculated as follows: (B/(A+B)) *100 from 5 and 3 biological replicates in U2932 and SUDHL4, respectively. Statistics, Student *t* test. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$.


A
Genes
Number of DEGs
1,000
900
800
700
600
500
400
200
150
100
50
0
915
732
647
507
178
154
D3
D6
D10
Upregulated
Downregulated
B
Day 6
1,281
171
144
51
906
111
26
Day 3
Day 10
C
U2932 - Day 6
ZNF587/417 KD (log2 norm. counts)
Ctrl (log2 norm. counts)
HLA-C
B2M
HLA-A
HLA-B
ISG15
CDKN1A
MX1
IFI16
IRF7
OAS1
STAT1
CXCL16
CCL3
CCL4
XAF1
IFIT1
STAT2
IFIT3
TNFa
CCL22
USP18
FAS
IFI44
CXCL10
STING1
IFN/inflammatory response genes
D
Cytosolic dsDNA
Mean cytoband intensity
ShScr
ShRNA.1
ShRNA.2
dsDNA
E
p-STING
Mean cell intensity
p-STING
F
p-IRF3
Mean cell intensity
4 µm
p-IRF3
G
DEGs - Day 6
(Padj < 0.05)
U2932
4,083
3,384
2,555
OCI-Ly7
DOWN
UP
log2FC OCI-Ly7
log2FC U2932
r = 0.68, ****
IFN/Infl. response genes
Other DEGs
H
KZFPHigh/Low DEGs
OCI-Ly7
1,181
1,374
1,224
2,160
8,143
1,485
2,598
U2932
log10FC U2932
r = 0.31, ****
log10FC
KZFP-Low vs -High
I
ShScramble
ShRNA.1
ShRNA.2
Total red object area (µm2/image)
Time (hours)
400 µm

was similar across different cellular contexts. Furthermore, upon intersecting genes differentially expressed at day 6 after KD in U2932 or OCI-Ly7 cells with the genes defining the $KZFP^{High}$ and $KZFP^{Low}$ groups of the DLBCL cohort, we found that *ZNF587/417* KD induced a shift of the cell lines toward a $KZFP^{Low}$-like transcriptome, supporting an important role for ZNF417 and ZNF587 in the gene-expression profile of poor prognosis DLBCL (**Fig. 6H**). Indicating that these changes would lead to an increased susceptibility to innate immunity clearance, we found that *ZNF587/417* KD cells displayed increased susceptibility to phagocytosis by human M1 polarized macrophages (**Fig. 6I**).

ZNF587/417 depletion modifies the antigenic profile of lymphoma cells

Our data indicate that the RS induced by *ZNF587/417* KD triggers an inflammatory response through the cytosolic release of DNA fragments and their sensing by the cGAS–STING pathway. This type of response normally leads to increased expression of HLA-I genes, which play a pivotal role in the immunogenicity of cancer cells. Accordingly, in U2932 cells, *HLA-C* gene expression started to be significantly upregulated at day 3 of *ZNF587/417* KD (Supplementary Fig. S3G), followed by *HLA-A*, *HLA-B*, and β2-microglobulin (*B2M*) at day 6 (**Fig. 6C**), whereas in OCI-Ly7, *HLA-B* and *HLA-C* transcripts were augmented only at this later time point (Supplementary Fig. S6G). Corroborating these transcriptomic changes, U2932 cells displayed a 2-fold increase in surface expression of HLA-A, -B, and -C at KD day 10, even though at that point their source transcripts were no longer upregulated (**Fig. 7A**). We went on to assess whether the upregulation of HLA-I observed in *ZNF587/417* KD cells was associated with an alteration of their HLA-bound peptidome. For this, HLA-bound peptides of KD day 10 and control U2932 cells were eluted and sequenced by mass spectrometry. We found 1,367 peptides derived from 1,050 different proteins (∼14.3% of the detected immunopeptidome and 23% of presented proteins) to be significantly more and 1,154 peptides from 876 proteins (∼12% of the detected immunopeptidome and 19.2% of presented proteins) to be less abundant in *ZNF587/417* KD cells ($P < 0.05$; **Fig. 7B**). In accordance with the transcriptome data, we noted an enrichment of peptides derived from inflammatory gene products in *ZNF587/417* KD cells, whereas KZFP-derived peptides were prevalent among depleted entities (**Fig. 7B**). Furthermore, the peptide repertoire of U2932 KD cells was altered, with noticeable changes in peptide length: 8 and 9-mers made up 90.5% of the peptides enriched in KD cells but only 65.9% of the peptides enriched in control cells, and conversely 10- and 11-mers constituted only 8.8% of the peptides enriched in U2932 KD versus 29.4% in control cells (**Fig. 7C**). We hypothesized these changes to be linked to differences in HLA allele-specific expression. By using previously described HLA binding prediction methods (**Fig. 7D**), we found that the percentage of peptides predicted to bind HLA-B (69.4% vs. 41.8%) and -C (7.7% vs. 3.5%) was increased in U2932 KD cells compared with control cells, where peptides were predicted to bind mostly HLA-A03:01 (50.1%), an allele recently associated with poor responses to immune-checkpoint blockade (62). The *ZNF587/417* KD also increased the proportion of peptides carrying C-terminal hydrophobic (leucine) at the expense of amphipathic (lysine) residues (**Fig. 7C**), suggesting underlying alterations of the proteasome degradation pathway and changes in HLA binding affinity. Therefore, we compared the immunogenicity of the eluted peptides using the Immune Epitope Database (IEDB) analytic tool (63), which predicted that immunopeptides eluted from U2932 KD cells were on average more immunogenic than that of control cells (**Fig. 7E**). Even though the contribution of KZFPs in DLBCL immune evasion needs to be further established, these results indicate that *ZNF587/417* KD reshuffles the immunopeptidome of lymphoma cells toward increased immunogenicity.

Discussion

Genomic instability confers cancer cells with evolutionary advantages by generating clonal diversity, which facilitates escape from immune surveillance and chemotherapeutic agents. Yet it has to be balanced to avoid mitotic catastrophes and antitumoral immune responses. Constitutive heterochromatin is a dynamic state reestablished after each cell division through epigenetic memory (64). The present work demonstrates that KZFPs, the TE-targeting partners of the KAP1–SETDB1 complex, play an important role in the maintenance of heterochromatin in cancer cells, protecting these from excessive RS, secondary inflammatory responses, and other immunologically detectable alterations. Our results further identify ZNF587 and ZNF417, two primate-specific paralogs, as leading protumoral KZFPs, and their upregulation, together with that of 16 other KZFP genes, as a strong negative prognosis indicator in DLBCL.

Recent studies have revealed that chromosomal instability (CIN) is a main driver of cancer progression and metastasis, and pointed to the induction of the cGAS–STING dsDNA sensing pathway as an important mediator of this process (65). CIN^{High} tumors are typically described as carrying a poor prognosis and displaying activation of genes related to IFN/inflammatory responses, EMT, and TNFα

Figure 6.

ZNF587/417 depletion leads to cell-intrinsic inflammation. **A,** Number of DEGs of RNA-seq analysis performed at day 3 ($n = 2$), day 6 ($n = 3$), and day 10 ($n = 2$) of ZNF587/417 KD vs. control U2932 cells. **B,** Euler diagrams of the overlap of DEGs upon ZNF587/417 KD in U2932 cells at each time point (days 3, 6, and 10). **C,** Scatter plot of $\log_2$ normalized counts of ZNF417/587 KD cells vs. control U2932 at day 6, outlining DEGs (gray dots), among which, genes belonging to type I/II IFN and inflammatory response Hallmark gene sets, Interferon Signaling Reactome, and cellular response to type I, II, and III IFN gene ontology terms are depicted in blue. **D,** Representative images and dot plot showing the mean intensity of cytosolic double-stranded DNA per cell, measured on z-stack immunofluorescence images ($n \geq 80$ cells per condition). Statistics, two-sided Mann–Whitney *U* test. **E** and **F,** Same as **D** for phosphorylated STING (**E**) and phosphorylated IRF3 (**F**) signal, respectively. **G,** Top, Venn diagrams showing the overlap of DEGs upon *ZNF587/417* KD in U2932 and OCI-Ly7 at day 6. Bottom, scatterplot of $\log_2$-fold changes of DEGs shared between U2932 (*x*-axis) and OCI-Ly7 (*y*-axis) KD cells. Blue dots highlight the genes belonging to IFN/inflammatory response terms detailed in **C**. The best-fit line (gray) results from the linear regression analysis of U2932 $\log_2$-fold changes onto OCI-Ly7 $\log_2$-fold changes. **H,** Top, Venn diagrams showing the overlap of DEGs upon *ZNF587/417* KD in U2932 and OCI-Ly7 cells at day 6 and genes discriminating $KZFP^{High}$ and $KZFP^{Low}$ DLBCLs. Bottom, scatterplot of $\log_2$-fold changes of DEGs shared between genes discriminating $KZFP^{High}$/$KZFP^{Low}$ DLBCLs (*x*-axis) and U2932 KD cells (*y*-axis). The best-fit line (black) results from the linear regression analysis of $KZFP^{High}$ vs. $KZFP^{Low}$ $\log_2$-fold changes onto U2932 KD $\log_2$-fold changes. For this panel, DEGs were defined with an FDR < 0.05 and a fold change >2. *r*, correlation coefficient. **I,** Phagocytosis assay of U2932 pHrodo-labeled cells cocultured with M1-polarized macrophages. Representative 24-hour course of pHrodo signal quantification using IncuCyte time-lapse imaging of *ZNF587/417* KD cells with two different shRNAs (shRNA.1, blue; shRNA.2, turquoise blue) and control cells (shScramble, dark red). Total pHrodo cell area per image acquired was calculated for each time point and plotted as a time course for each condition. Right, representative images of respective conditions. Statistics, two-way ANOVA followed by Bonferroni correction. Asterisks indicate significant differences at 24-hour time point compared with controls. ****, $P < 0.0001$.

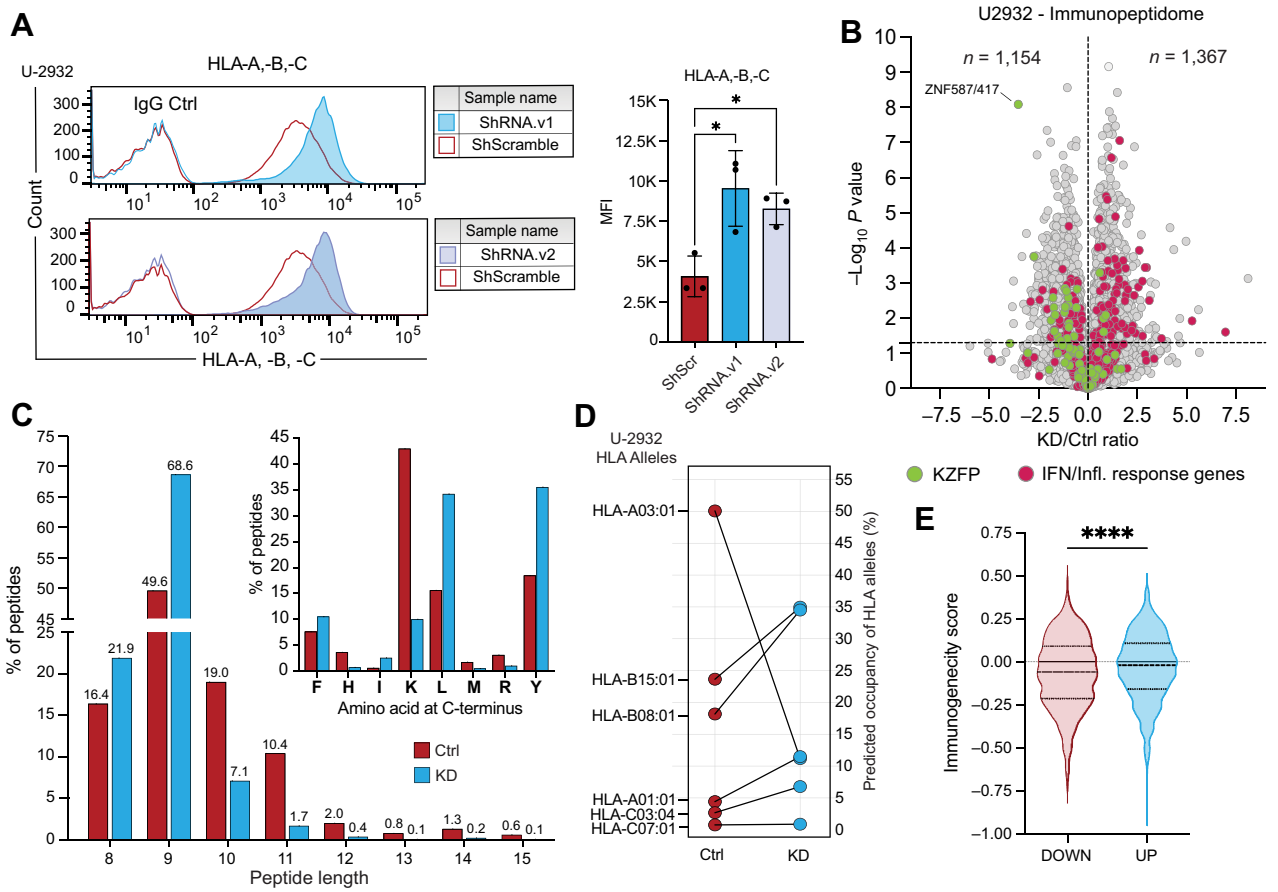


Figure 7.

ZNF587/417 depletion modifies the antigenic profile of lymphoma cells. **A,** Representative flow cytometry histograms (left) and bar plot quantification (right) of median fluorescence intensity (MFI) of HLA class I (HLA-A, -B, -C) expression on U2932 KD and control cells after 10 days with two different shRNAs targeting *ZNF587/417* (shRNA.1/.2) or a control shRNA (shScr). Statistics, Student *t* test (mean of MFI ± SD of independent triplicates). **B,** Volcano plot showing changes in the abundance of HLA-I–bound peptides between ZNF587/417 KD (shRNA.1/.2) and control cells from three independent experiments. The horizontal dashed line indicates a *P* value threshold of 0.05. *n*, number of enriched and depleted peptides. Peptides derived from gene products belonging to type I/II IFN and inflammatory response Hallmark gene sets, Interferon Signaling Reactome, and cellular response to type I, II, and III IFN gene ontology terms are highlighted in dark red. Peptides derived from KZFPs are highlighted in green. Statistics, Student *t* test. **C,** Bottom left, differences in the proportion of peptides of different lengths between control (dark red) and KD cells (blue). Top right, bar plot showing the differences in the percentage of peptides with listed C-terminal amino acid between control (dark red) and KD cells (blue). **D,** Line plot showing the changes in the proportion of predicted occupancy of HLA alleles present in U2932 cells between control (dark red) and KD cells (blue). **E,** Violin plots showing the predicted immunogenicity scores of eluted peptide ligands enriched in control (DOWN; red) and KD cells (UP; blue) from the IEDB analytic tool. *, $P < 0.05$; ****, $P < 0.0001$.

signaling (65). Here, we found KZFP^{High} DLBCLs to exhibit a higher average number of CNAs than their KZFP^{Low} counterparts, yet their transcriptome to be characterized by an upregulation of DNA-repair genes but CIN^{Low} expression signatures linked to cell cycle, Myc target, and OxPhos-related genes. Thus, KZFP^{High} DLBCL cells presumably combine higher degrees of genetic instability with an increased ability both to attenuate RS and to dampen inflammatory responses. As a consequence, the stabilization of their chromatin via upregulation of critical controllers of repetitive DNA results both in increasing the efficiency of their clonal expansion and diversification and in facilitating their immune evasion. ZNF587/417 recognize evolutionarily recent ERVs and SVAs, capable of robust transcriptional influences notably manifested during brain development (11, 24). In most tissues, *cis*-acting regulatory sequences embedded in these retroelements are maintained in a repressed state through a mix of histone and DNA methylation, the latter perpetuated across mitosis through the action of maintenance DNA methyltransferases, without the need for sequence-specific repressors. Our data suggest a model whereby, perhaps due to the state of DNA hypomethylation characteristic of cancer cells, maintenance of heterochromatin at these TE loci becomes strongly dependent on high levels of histone methylation-inducing KZFPs recognizing their sequence in order to prevent inflammatory responses. Chromatin accessibility is an important determinant of replication origins (66, 67). Here, modifications of H3K9me3 distribution upon ZNF587/417 depletion led to marked alterations of DNA replication with a multiplication of origins, increased directionality where this mark increased and loss of polarity where it dropped, accumulation of DSBs and cytoplasmic release of DNA fragments sensed by the cGAS/STING pathway, triggering the same type of immunogenic inflammation as observed in KZFP^{Low} DLBCL samples (**Fig. 8**).

The importance of KZFPs in DLBCL pathogenesis is supported by the recurrence of amplifications of the distal portion of Chr19q in this lymphoid malignancy (68). This phenomenon has so far been

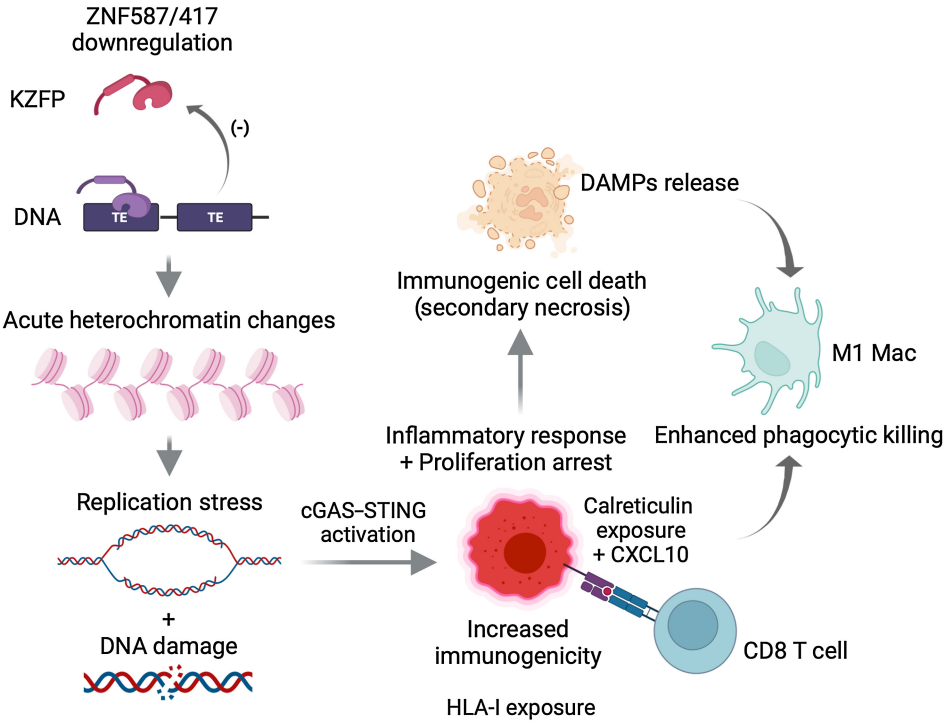


Figure 8.
Putative model of the cellular effects resulting from ZNF587/417 depletion in highly proliferative cells. Downregulation of ZNF587/417 leads to acute heterochromatin changes, followed by alterations of the DNA replication program leading to enhanced RS, DNA damage, and ICD. Cytosolic dsDNA from DNA damage (enhanced by secondary viral mimicry from TE upregulation) triggers cGAS–STING activation, IFN-like response, CXCL10 upregulation, enhanced phagocyting killing by M1 MAC, and enhanced MHC-I surface exposure in surviving cells (increased immunogenicity). (Created with BioRender.com.)

attributed to the presence in this chromosomal segment of *SPIB*, a gene encoding a repressor of the IFNβ gene activator IRF7 (69). However, SPIB acts in a COO-restricted manner (70), whereas Chr19q amplifications have been detected in both ABC and GCB-DLBCLs and demonstrated to correlate with poor prognosis independently from the lymphoma COO in prospective cohorts (68). Our results thus strongly suggest that this chromosomal amplification is driven by the protumoral potential of *ZNF587/417* and other components of *KZFP* gene cluster 11, as well as perhaps by the additional presence in the subtelomeric region of Chr19q of *KAP1*, which not only is an essential cofactor of these KZFPs but also itself a negative prognostic factor in non-Hodgkin B-cell lymphoma (71).

Our finding that γH2AX accumulates at hotspots of endogenous DNA damage upon *ZNF587/417* KD is reminiscent of chromatin alterations detected in lymphoblastoid cell lines exposed to chemical inducers of RS (54). Moreover, the reshuffling of the immunopeptidome observed in lymphoma cells surviving the acute genotoxic stress triggered by *ZNF587/417* KD closely resembles changes observed upon exposure of tumor cells to noncytotoxic doses of the nucleoside analogue gemcitabine (72). Together, these data indicate that methods aimed at inactivating ZNF587/417 could mimic the action of chemotherapeutic agents. Our finding that dependence on these KZFPs is not restricted to DLBCL cells but extends to a broad range of highly proliferative cancerous and noncancerous cells implies that these KZFPs and probably others engaged in maintaining heterochromatin at TE loci serve a core homeostatic role in safeguarding genome stability during DNA replication.

Authors' Disclosures

F. Martins reports grants from the Swiss National Science Foundation, the European Research Council, Personalized Health and Related Technologies (PHRT) strategic focus area of the Swiss Federal Institutes of Technology, the Swiss Personalized Health Network (SPHN) initiative of the Swiss Academy of Medical Sciences, Aclon Foundation, and Ludwig Institute during the conduct of the study. J. Carlevaro-Fita reports grants from the Swiss National Science Foundation, the European Research Council, the Aclon Foundation, and the Ludwig Institute during the conduct of the study. C. Pulver reports grants from the Swiss National Science Foundation, the European Research Council, PHRT strategic focus area of the Swiss Federal Institutes of Technology, SPHN initiative of the Swiss Academy of Medical Sciences, Aclon Foundation, and the Ludwig Institute during the conduct of the study. P. Turelli reports a patent for PCT/EP2022/06972 pending. D. Trono reports grants from ERC, the Swiss National Science Foundation, the Aclon Foundation, the Swiss Institutes of Technology, and the Ludwig Institute during the conduct of the study; in addition, D. Trono has a patent for KZFPs and cancer pending. No disclosures were reported by the other authors.

Authors' Contributions

F. Martins: Conceptualization, data curation, formal analysis, validation, investigation, visualization, methodology, writing–original draft, project administration, writing–review and editing. **O. Rosspopoff:** Data curation, formal analysis, investigation. **J. Carlevaro-Fita:** Formal analysis, visualization. **R. Forey:** Investigation, methodology. **S. Offner:** Investigation. **E. Planet:** Formal analysis. **C. Pulver:** Investigation. **H. Pak:** Investigation. **F. Huber:** Formal analysis, investigation. **J. Michaux:** Investigation. **M. Bassani-Sternberg:** Investigation. **P. Turelli:** Investigation. **D. Trono:** Resources, supervision, funding acquisition, methodology, writing–original draft, project administration, writing–review and editing.

Acknowledgments

We thank C. Raclot, P.-Y. Helleboid, J. Duc, V. Glutz, R. Guiet, N. Chiaruttini, L. Simo Riudalbas, S. Sheppard, D. Grun, M. Begnis, F. Marzetta, E. Oricchio, and J. Pontis for their technical support and scientific advice, R. Guindon for help in the graphical output of the figures, the EPFL Flow Cytometry (FCCF), Protein Production and Purification (PTPSP), and Gene Expression (GECF) core facilities and the BioImaging and Optics Platform (BIOP) for the use of their systems and services. We thank A. Mayran and the laboratory of D. Duboule for help with the use of the InCucyte S3 Sartorius imaging system. We thank C. Cummings, R. Sudlow, H. Clevers, C. Bolen, L. Garraway, C. Fuchs, and M. Nowicka from Roche for their help in making the metadata relative to the GOYA trial publicly available for the research community and our publication. We also acknowledge all the members of the Trono laboratory for fruitful discussions and constant support. This study was supported by grants from the Swiss National Science Foundation and the European Research Council (KRABnKAP, no. 268721; Transpos-X, no. 694658), the Aclon Foundation, and the Ludwig Institute to D.Trono and the Personalized Health and

Related Technologies (PHRT) strategic focus area of the Swiss Federal Institutes of Technology, the Swiss Personalized Health Network (SPHN) initiative of the Swiss Academy of Medical Sciences (project no. 2017-407) to F. Martins.

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

Note

Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Received April 24, 2023; revised October 15, 2023; accepted January 19, 2024; published first February 12, 2024.

References

1. Sehn LH, Berry B, Chhanabhai M, Fitzgerald C, Gill K, Hoskins P, et al. The revised international prognostic index (R-IPI) is a better predictor of outcome than the standard IPI for patients with diffuse large B-cell lymphoma treated with R-CHOP. Blood 2007;109:1857–61.
2. Wright G, Tan B, Rosenwald A, Hurt EH, Wiestner A, Staudt LM. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large b cell lymphoma. Proc Natl Acad Sci 2003;100:9991–6.
3. Méchali M. Eukaryotic DNA replication origins: many choices for appropriate answers. Nat Rev Mol Cell Biol 2010;11:728–38.
4. Jackson DA, Pombo A. Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. J Cell Biol 1998;140:1285–95.
5. Murga M, Campaner S, Lopez-Contreras AJ, Toledo LI, Soria R, Montaña MF, et al. Exploiting oncogene-induced replicative stress for the selective killing of myc-driven tumors. Nat Struct Mol Biol 2011;18:1331–5.
6. Guilbaud G, Murat P, Wilkes HS, Lerner LK, Sale JE, Krude T. Determination of human DNA replication origin position and efficiency reveals principles of initiation zone organisation. Nucleic Acids Res 2022;50:7436–50.
7. Morales ME, White TB, Streva VA, DeFreece CB, Hedges DJ, Deininger PL. The contribution of alu elements to mutagenic DNA double-strand break repair. PLoS Genet 2015;11:e1005016.
8. De Magis A, Manzo SG, Russo M, Marinello J, Morigi R, Sordet O, et al. DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells. Proc. Natl. Acad. Sci. U.S.A. 2019;116:816–25.
9. Zeller P, Padeken J, van Schendel R, Kalck V, Tijsterman M, Gasser SM. Histone H3K9 methylation is dispensable for caenorhabditis elegans development but suppresses RNA:DNA Hybrid-associated repeat instability. Nat Genet 2016;48:1385–95.
10. Pontis J, Planet E, Offner S, Turelli P, Duc J, Coudray A, et al. Hominoid-specific transposable elements and KZFPs facilitate human embryonic genome activation and control transcription in naive human ESCs. Cell Stem Cell 2019;24:724–35.
11. Imbeault M, Helleboid P-Y, Trono D. KRAB Zinc-finger proteins contribute to the evolution of gene regulatory networks. Nature 2017;543:550–4.
12. Helleboid P, Heusel M, Duc J, Piot C, Thorball CW, Coluccio A, et al. The interactome of KRAB zinc finger proteins reveals the evolutionary history of their functional diversification. EMBO J 2019;38:e101220.
13. Friedman JR, Fredericks WJ, Jensen DE, Speicher DW, Huang XP, Neilson EG, et al. KAP-1, a novel corepressor for the highly conserved KRAB repression domain. Genes Dev 1996;10:2067–78.
14. Ecco G, Cassano M, Kauzlaric A, Duc J, Coluccio A, Offner S, et al. Transposable elements and their KRAB-ZFP controllers regulate gene expression in adult tissues. Dev Cell 2016;36:611–23.
15. Ito J, Kimura I, Soper A, Coudray A, Koyanagi Y, Nakaoka H, et al. Endogenous retroviruses drive KRAB zinc-finger family protein expression for tumor suppression. Sci Adv. 2020 Oct 21;6:eabc3020.
16. Busiello T, Ciano M, Romano S, Sodaro G, Garofalo O, Bruzzese D, et al. Role of ZNF224 in cell growth and chemoresistance of chronic lymphocitic leukemia. Hum Mol Genet 2017;26:344–53.
17. Yu S, Ao Z, Wu Y, Song L, Zhang P, Li X, et al. ZNF300 promotes chemoresistance and aggressive behaviour in non-small-cell lung cancer. Cell Prolif 2020;53:e12924.
18. Yang L, Hamilton SR, Sood A, Kuwai T, Ellis L, Sanguino A, et al. The previously undescribed ZKSCAN3 (ZNF306) is a novel "Driver" of colorectal cancer progression. Cancer Res 2008;68:4321–30.
19. Liu Y, Ouyang Q, Sun Z, Tan J, Huang W, Liu J, et al. The novel zinc finger protein 587B Gene, ZNF587B, regulates cell proliferation and metastasis in ovarian cancer cells in vivo and in vitro. CMAR 2020;12:5119–30.
20. Wang C, Ivanov A, Chen L, Fredericks WJ, Seto E, Rauscher FJ, et al. MDM2 interaction with nuclear corepressor KAP1 contributes to P53 inactivation. EMBO J 2005;24:3279–90.
21. Griffin GK, Wu J, Iracheta-Vellve A, Patti JC, Hsu J, Davis T, et al. Epigenetic silencing by SETDB1 suppresses tumour intrinsic immunogenicity. Nature 2021;595:309–14.
22. Tunbak H, Enriquez-Gasca R, Tie CHC, Gould PA, Mlcochova P, Gupta RK, et al. The HUSH complex is a gatekeeper of type I interferon through epigenetic regulation of LINE-1s. Nat Commun 2020;11:5387.
23. Reddy A, Zhang J, Davis NS, Moffitt AB, Love CL, Waldrop A, et al. Genetic and functional drivers of diffuse large B cell lymphoma. Cell 2017;171:481–94.
24. Turelli P, Playfoot C, Grun D, Raclot C, Pontis J, Coudray A, et al. Primate-restricted KRAB zinc finger proteins and target retrotransposons control gene expression in human neurons. Sci Adv 2020;6:eaba3200.
25. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res 2016;44:D279–285.
26. Cunningham F, Allen JE, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, et al. Ensembl 2022. Nucleic Acids Res 2022;50:D988–95.
27. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, et al. The human transcription factors. Cell 2018;172:650–65.
28. Klijn C, Durinck S, Stawiski EW, Haverty PM, Jiang Z, Liu H, et al. A comprehensive transcriptional portrait of human cancer cell lines. Nat Biotechnol 2015;33:306–12.
29. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.
30. Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun 2019;10:1930.
31. Bryson T, Henikoff S. 3XFlag-pATn5 protein purification and MEDS-loading (5x scale, 2L volume) V1; 2019. dx.doi.org/10.17504/protocols.io.8yrhxv6.
32. Nurk S, Koren S, Rhie A, Rautiainen M, Bzikadze AV, Mikheenko A, et al. The complete sequence of a human genome. Science 2022;376:44–53.
33. Kara N, Krueger F, Rugg-Gunn P, Houseley J. Genome-wide analysis of DNA replication and DNA double-strand breaks using TrAEL-seq. PLoS Biol 2021;19:e3000886.
34. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. Nat Methods 2012;9:357–9.
35. Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, et al.; Manke, T. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res 2016;44:W160–5.
36. R: A language and environment for statistical computing. 2013.
37. Kent WJ, Zweig AS, Barber G, Hinrichs AS, Karolchik D. BigWig and BigBed: enabling browsing of large distributed datasets. Bioinformatics 2010;26:2204–7.
38. Chong C, Marino F, Pak H, Racle J, Daniel RT, Müller M, et al. High-throughput and sensitive immunopeptidomics platform reveals profound interferonγ-mediated remodeling of the human leukocyte antigen (HLA) ligandome. Mol Cell Proteomics 2018;17:533–48.
39. Chong C, Müller M, Pak H, Harnett D, Huber F, Grun D, et al. Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes. Nat Commun 2020;11:1293.
40. Pak H, Michaux J, Huber F, Chong C, Stevenson BJ, Müller M, et al. Sensitive immunopeptidomics by leveraging available large-scale multi-HLA spectral libraries, data-independent acquisition, and MS/MS prediction. Mol Cell Proteomics 2021;20:100080.
41. Kawaguchi S, Higasa K, Shimizu M, Yamada R, Matsuda F. HLA-HD: an accurate HLA typing algorithm for next-generation sequencing data. Hum Mutat 2017;38:788–97.

42. Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. Nucleic Acids Res 2020;48:W449–54.
43. Bruderer R, Bernhardt OM, Gandhi T, Miladinović SM, Cheng L-Y, Messner S, et al. Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues. Mol Cell Proteomics 2015;14:1400–10.
44. Merdan S, Subramanian K, Ayer T, Van Weyenbergh J, Chang A, Koff JL, et al. Gene expression profiling-based risk prediction and profiles of immune infiltration in diffuse large B-cell lymphoma. Blood Cancer J 2021;11:2.
45. Macheret M, Halazonetis TD. Intragenic origins due to short G1 phases underlie oncogene-induced DNA replication stress. Nature 2018;555:112–6.
46. Nagel I, Akasaka T, Klapper W, Gesk S, Bottcher S, Ritgen M, et al. Identification of the gene encoding cyclin E1 (CCNE1) as a novel IGH translocation partner in t (14;19)(Q32;Q12) in diffuse large B-cell lymphoma. Haematologica 2009;94: 1020–3.
47. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci 2005;102: 15545–50.
48. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database hallmark gene set collection. Cell Syst 2015;1: 417–25.
49. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor-infiltrating immune cells with CIBERSORT. Methods Mol Biol 2018;1711:243–59.
50. Karube K, Campo E. MYC alterations in diffuse large B-cell lymphomas. Semin Hematol 2015;52:97–106.
51. Wright GW, Huang DW, Phelan JD, Coulibaly ZA, Roulland S, Young RM, et al. A probabilistic classification tool for genetic subtypes of diffuse large B cell lymphoma with therapeutic implications. Cancer Cell 2020;37:551–68.
52. Liu M, Thomas SL, DeWitt AK, Zhou W, Madaj ZB, Ohtani H, et al. Dual inhibition of DNA and histone methyltransferases increases viral mimicry in ovarian cancer cells. Cancer Res 2018;78:5754–66.
53. Chen H-J, Li Yim AYF, Griffith GR, De Jonge WJ, Mannens MMAM, Ferrero E, et al. Meta-analysis of in vitro-differentiated macrophages identifies transcriptomic signatures that classify disease macrophages in vivo. Front Immunol 2019; 10:2887.
54. Lyu X, Chastain M, Chai W. Genome-wide mapping and profiling of γH2AX binding hotspots in response to different replication stress inducers. BMC Genomics 2019;20:579.
55. Sørensen CS, Syljuåsen RG. Safeguarding genome integrity: the checkpoint kinases ATR, CHK1 and WEE1 restrain CDK activity during normal DNA replication. Nucleic Acids Res 2012;40:477–86.
56. Courtot L, Hoffmann J-S, Bergoglio V. The protective role of dormant origins in response to replicative stress. Int J Mol Sci 2018;19:3569.
57. Padariya M, Sznarkowska A, Kote S, Gómez-Herranz M, Mikac S, Pilch M, et al. Functional interfaces, biological pathways, and regulations of interferon-related DNA damage resistance signature (IRDS) genes. Biomolecules 2021;11:622.
58. Hagan CR, Rudin CM. DNA cleavage and Trp53 differentially affect SINE transcription. Genes Chromosom. Cancer 2007;46:248–60.
59. Mackenzie KJ, Carroll P, Martin C-A, Murina O, Fluteau A, Simpson DJ, et al. cGAS surveillance of micronuclei links genome instability to innate immunity. Nature 2017;548:461–5.
60. Shen YJ, Le Bert N, Chitre AA, Koo CX, Nga XH, Ho SSW, et al. Genome-derived cytosolic DNA mediates type I interferon-dependent rejection of B cell lymphoma cells. Cell Rep 2015;11:460–73.
61. Grandvaux N, Servant MJ, tenOever B, Sen GC, Balachandran S, Barber GN, et al. Transcriptional profiling of interferon regulatory factor 3 target genes: direct involvement in the regulation of interferon-stimulated genes. J Virol 2002;76: 5532–9.
62. Naranbhai V, Viard M, Dean M, Groha S, Braun DA, Labaki C, et al. HLA-A*03 and response to immune checkpoint blockade in cancer: an epidemiological biomarker study. Lancet Oncol 2022;23:172–84.
63. Calis JJA, Maybeno M, Greenbaum JA, Weiskopf D, De Silva AD, Sette A, et al. Properties of MHC Class I presented peptides that enhance immunogenicity. PLoS Comput Biol 2013;9:e1003266.
64. Stewart-Morgan KR, Petryk N, Groth A. Chromatin replication and epigenetic cell memory. Nat Cell Biol 2020;22:361–71.
65. Bakhoum SF, Landau DA. Chromosomal instability as a driver of tumor heterogeneity and evolution. Cold Spring Harb Perspect Med 2017;7:a029611.
66. Cadoret J-C, Meisch F, Hassan-Zadeh V, Luyten I, Guillet C, Duret L, et al. Genome-wide studies highlight indirect links between human replication origins and gene regulation. Proc. Natl. Acad. Sci. U.S.A. 2008;105:15837–42.
67. Cayrou C, Coulombe P, Vigneron A, Stanojcic S, Ganier O, Peiffer I, et al. Genome-scale analysis of metazoan replication origins reveals their organization in specific but flexible sites defined by conserved features. Genome Res 2011;21: 1438–49.
68. Dubois S, Tesson B, Mareschal S, Viailly P-J, Bohers E, Ruminy P, et al. Refining diffuse large B-cell lymphoma subgroups using integrated analysis of molecular profiles. EBioMedicine 2019;48:58–69.
69. Lenz G, Wright GW, Emre NCT, Kohlhammer H, Dave SS, Davis RE, et al. Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. Proc. Natl. Acad. Sci. U.S.A. 2008;105:13520–5.
70. Yang Y, Shaffer AL, Emre NCT, Ceribelli M, Zhang M, Wright G, et al. Exploiting synthetic lethality for the therapy of ABC diffuse large B cell lymphoma. Cancer Cell 2012;21:723–37.
71. Zhang P-P, Ding D-Z, Shi B, Zhang S-Q, Gu L-L, Wang Y-C, et al. Expression of TRIM28 correlates with proliferation and bortezomib-induced apoptosis in B-cell non-Hodgkin lymphoma. Leuk Lymphoma 2018;59:2639–49.
72. Gravett AM, Trautwein N, Stevanović S, Dalgleish AG, Copier J. Gemcitabine alters the proteasome composition and immunopeptidome of tumour cells. OncoImmunology 2018;7:e1438107.