ARTICLE

Characterization of GECPAR, a noncoding RNA that regulates the transcriptional program of diffuse large B-cell lymphoma

Sara Napoli,[1] Luciano Cascione,[1,2] Andrea Rinaldi,[1] Filippo Spriano,[1] Francesca Guidetti,[1] Fangwen Zhang,[1] Maria Teresa Cacciapuoti,[3] Afua Adjeiwaa Mensah,[1] Giulio Sartori,[1] Nicolas Munz,[1] Mattia Forcato,[4] Silvio Bicciato[4], Annalisa Chiappella,[5] Paola Ghione,[6] Olivier Elemento,[7,8] Leandro Cerchietti, [6]Giorgio Inghirami[3] and Francesco Bertoni[1,9]

[1]Institute of Oncology Research, Faculty of Biomedical Sciences, USI, Bellinzona, Switzerland; [2]SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland; [3]Pathology and Laboratory Medicine Department, Weill Cornell Medicine, New York, NY, USA; [4]Center for Genome Research, Department of Life Sciences University of Modena and Reggio, Modena, Italy; [5]Ematologia, A.O.U. Città della Salute e della Scienza di Torino, Turin, Italy; [6]Department of Medicine, Division of Hematology and Medical Oncology, Weill Cornell Medicine, New York, NY, USA; [7]Institute for Computational Biomedicine, Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA; [8]Caryl and Israel Englander Institute for Precision Medicine, Weill Cornell Medicine, New York, NY, USA and [9]Oncology Institute of Southern Switzerland, Bellinzona, Switzerland.

Ferrata Storti Foundation

Haematologica 2022
Volume 107(5):1131-1143

ABSTRACT

Enhancers are regulatory regions of DNA, which play a key role in cell-type specific differentiation and development. Most active enhancers are transcribed into enhancer RNA (eRNA) that can regulate transcription of target genes by means of *in cis* as well as *in trans* action. eRNA stabilize contacts between distal genomic regions and mediate the interaction of DNA with master transcription factors. Here, we characterized an enhancer eRNA, GECPAR (germinal center proliferative adapter RNA), which is specifically transcribed in normal and neoplastic germinal center B cells from the super-enhancer of *POU2AF1*, a key regulatory gene of the germinal center reaction. Using diffuse large B-cell lymphoma cell line models, we demonstrated the tumor suppressor activity of GECPAR, which is mediated via its transcriptional regulation of proliferation and differentiation genes, particularly MYC and the Wnt pathway.

Introduction

Enhancers are regulatory DNA regions that positively drive gene transcription across neighboring genomic regions spanning many megabases and are characterized by distinct epigenetic features:[1,2] a high ratio of H3K4me1 to H3K4me3; enrichment of H3K27ac, which is deposited by the CREBBP/p300 complex;[3] high accessibility to chromatin readers such as bromodomain and extraterminal domain (BET) proteins and transcription factors (TF). Some enhancers are actively transcribed giving rise to noncoding RNA called enhancer RNA (eRNA).[4] Transcribed enhancers are more acetylated, more enriched of TF and co-activators, and are also more active in the transactivation of promoters, with which they interact inside 3D structures called enhancer-promoter loops.[5] Clusters of enhancers, called super-enhancers (SE), are strongly transcribed and produce several eRNA controlling key genes, which regulate cellular development and differentiation.[6,7] eRNA are crucial components of the regulatory chromatin machinery that controls the expression of key context-specific, protein-coding genes. They usually stabilize multiprotein complexes and constitute a scaffold for DNA loops by enforcing interactions between distant DNA regions, including those located on different chromosomes.[8-11] As they lack a poly A tail, their activity is restrained to the site of transcription and they undergo rapid decay. However, polyadenylated long intergenic non-coding RNA (lincRNA) also comprise enhancer-derived

Correspondence:

FRANCESCO BERTONI
francesco.bertoni@ior.usi.ch

SARA NAPOLI
sara.napoli@ior.usi.ch

Received: July 29, 2020.

Accepted: June 16, 2021.

Pre-published: June 24, 2021.

https://doi.org/10.3324/haematol.2020.267096

non-coding transcripts (e-lncRNA),[12] and the stabilization of these eRNA confers to them the capability to act *in trans*, regulating several distant targets.[13]

Individual eRNA are expressed in a tissue-specific manner. In normal B cells at various stages of differentiation, the expression of non-coding RNA can more precisely define cellular subsets than protein-coding transcripts.[14,15] In particular, eRNA are differentially expressed during B-cell development and they are associated with protein-coding genes that play an essential role in B-cell differentiation.

Diffuse large B-cell lymphoma (DLBCL) derives from germinal center (GC) B cells. DLBCL is typically divided into two main subtypes: GC B cell-like (GCB-DLBCL), whose transcriptional profile resembles that of light zone GC B cells, and activated B cell-like (ABC-DLBCL), whose transcriptome resembles that of plasmablasts.[16] However, DLBCL within each of these subgroups exhibit biological, genetic and transcriptional heterogeneity.[17-19] Lineage-specific and growth-dependent transcription factors like BCL6, Myc, NF-κB, p53, and E2F1 can activate specific genetic signatures, depending on the activation of unique subsets of enhancers[20,21] and contribute to disease heterogeneity. Here, we studied a unique eRNA associated with the *POU2AF1* gene, that we termed GECPAR, for germinal center proliferative adapter RNA. *POU2AF1* encodes the protein OCA-B, co-activator of OCT2, a B-cell specific transcription factor which plays a pivotal role in the regulation of normal and neoplastic GC B cells.[22,23] The SE proximal to *POU2AF1* is the most activated SE in GCB-DLBCL.[23] Loss of GECPAR correlated with reduced transcription of *TLE4*, which is a negative regulator of LEF1, a Wnt pathway effector protein that in turn regulates also NF-κB. GECPAR loss also increased MYC expression and proliferation of DLBCL cell lines. Conversely, its overexpression impaired cell proliferation. Collectively, our data provides evidence of the nodal role of GECPAR in the regulatory network modulating B-cell differentiation and proliferation.

Methods

Detailed descriptions of the experimental methods are included in the *Online Supplementary Appendix*.

Human samples, cell lines, small interfering RNA transfection

Established human DLBCL cell lines and patient-derived tumor xenograft cell lines (PDTX-CL) were grown as previously described.[24] All patients providing samples gave written informed consent. Molecular and clinical data acquisition and PDTX establishment were approved and carried out in accordance with Declaration of Helsinki and were approved by Institutional Review Boards of the New York Presbyterian Hospital, Weill Cornell Medicine (WCM) and the Ospedale San Giovanni Battista delle Molinette. Cell lines were checked for their identity.[24] Cells were transfected with small interfering RNA (siRNA) or locked-nucleic acid (LNA) using the 4D Nucleofector.

GECPAR cloning and infection into lymphoma cells, RNA sequencing

Cellular lysates were fractionated as previously described.[25] For strand-specific quantitative reverse transcription polymerase chain reaction (qRT-PCR), only the forward primer was used to amplify the antisense strand and only the reverse primer to amplify the sense strand. 5' and 3' rapid amplification of cDNA ends (RACE) was done using Invitrogen RACE System kits. GECPAR was cloned into the pGEM-T vector and subcloned in pCDH-CMV-MCS-EF1-copGFP. pCDH empty backbone or pCDH_GECPAR were transfected in HEK293T, and viral supernatant was then used to infect lymphoma cells. RNA sequencing (RNA-Seq) in cell lines was performed using the NEBNext Ultra II Directional RNA Library Prep.

Capture hybridization analysis of RNA targets (CHART) sequencing

CHART enrichment and ribonuclease H (RNAseH) mapping experiments were performed following previously reported protocols.[26,27] The enrichment of CHART signals was determined relative to the oligo controls. Conservative enrichment profiles were determined using the SPP package[28] and MACS,[29] as described by Vance and colleagues.[30]

Results

The super-enhancer associated with the *POU2AF1* gene locus is transcribed in normal B cells and diffuse large B-cell lymphoma cell lines

Analysis of publicly available RNA-Seq data on RNA polyA+ or polyA-[31] showed that CD20+ cells express a non-polyadenylated portion of the LOC100132078 transcript and also two isoforms of a more abundant antisense transcript (Figure 1A; *Online Supplementary Figure S1A*). Due to its proximity to the *POU2AF1* gene and its localization in a genomic region with characteristic SE features (highly acetylated, enriched in H3K4me1 but not H3K4me3, based on ENCODE ChIP-Seq data), we hypothesized that it could be an eRNA with particular relevance for GC B cells

In order to confirm the eRNA length reconstructed in CD20+ cells, we performed 5' and 3' RACE in the DLBCL cell line OCI-LY1. For the 3'-end detection we ran two reactions, with or without the addition of an artificial polyA tail. We identified a transcript lacking a polyA tail and another that was 400 bases longer and naturally polyadenylated. Similarly to the aforementioned polyA- transcript reported in CD20+ normal B cells, neither of the transcripts identified in DLBCL cells extended beyond the annotated first exon. The 5' RACE reaction reverse transcribed from exon 4 did not identify a specific 5'-end for exon 1, indicating that the long annotated transcript, LOC100132078, was likely not stable in our model. Conversely, reverse transcribing from exon 1, we identified a 5'-end located at nucleotide +366, mirroring our *in silico* observations for CD20+ normal B cells (Figure 1A and B). We renamed the stabilized portion of LOC100132078 we had sequenced in the OCI-LY1 model as "GECPAR".

GECPAR is mainly chromatin associated and partially polyadenylated

In order to further characterize the physical characteristics of GECPAR, RNA was extracted from the cytoplasm, nucleoplasm and chromatin fractions. In GCB-DLBCL (OCI-LY1 and Karpas422) and ABC-DLBCL (HBL1, U2932) cell lines, GECPAR was transcribed but mostly retained on chromatin, in accordance with reported features of eRNA.[6,7] It was also clearly detected in the nucle-

oplasm and cytoplasm of OCI-LY1, a cell line with 5-fold higher levels of chromatin-associated GECPAR than the other cell lines (Figure 1C). Semi-quantitative directional RT-PCR showed that chromatin association was particular to GECPAR since its antisense transcript, when expressed, was more ubiquitously distributed (*Online Supplementary Figure S1B*). Quantification of *KCNQ1OT1*, *MALAT1* and β-actin mRNA served as a control for chromatin-associated, nuclear and cytosolic RNA, respectively (*Online Supplementary Figure S1C*).

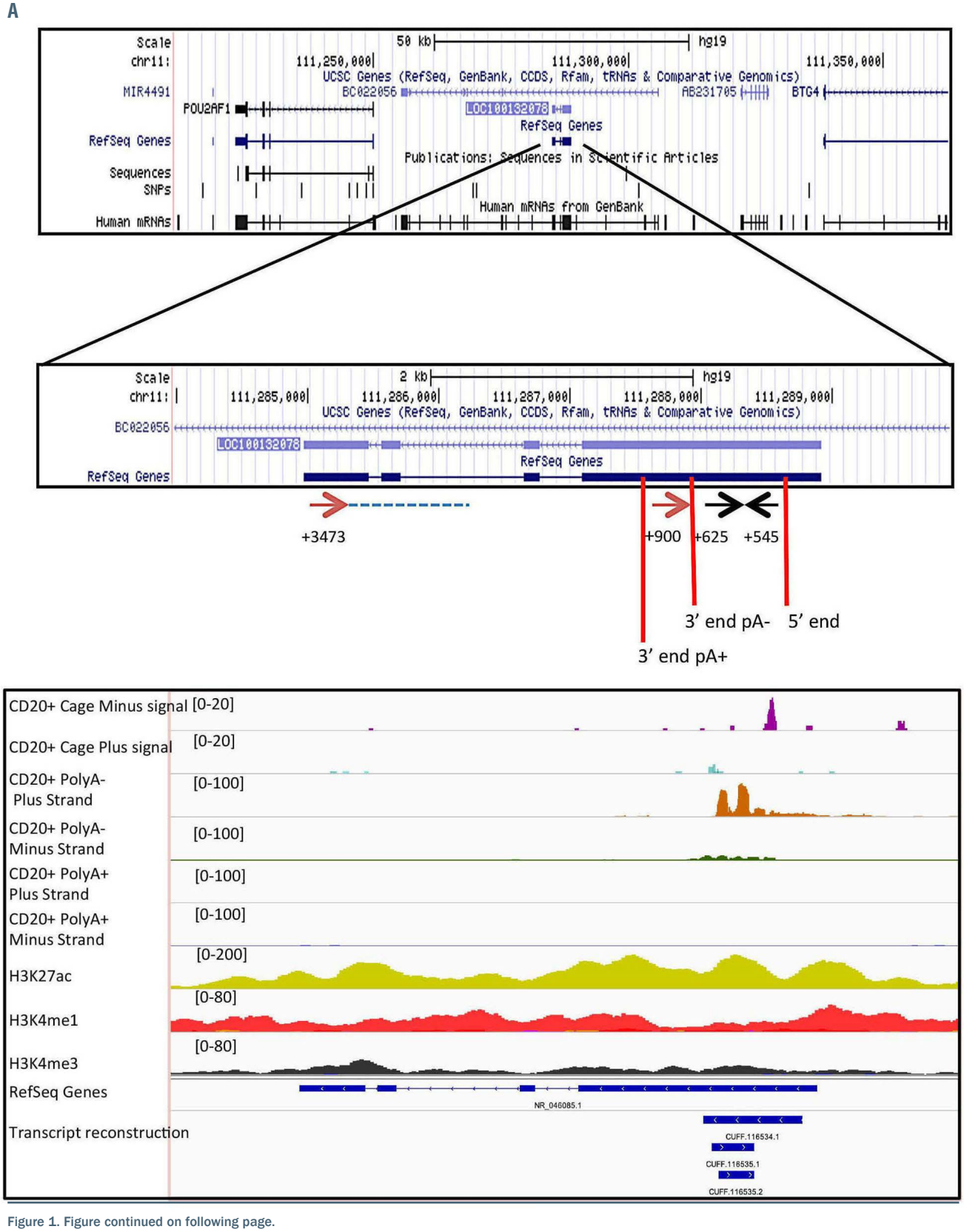

Figure 1. Figure continued on following page.

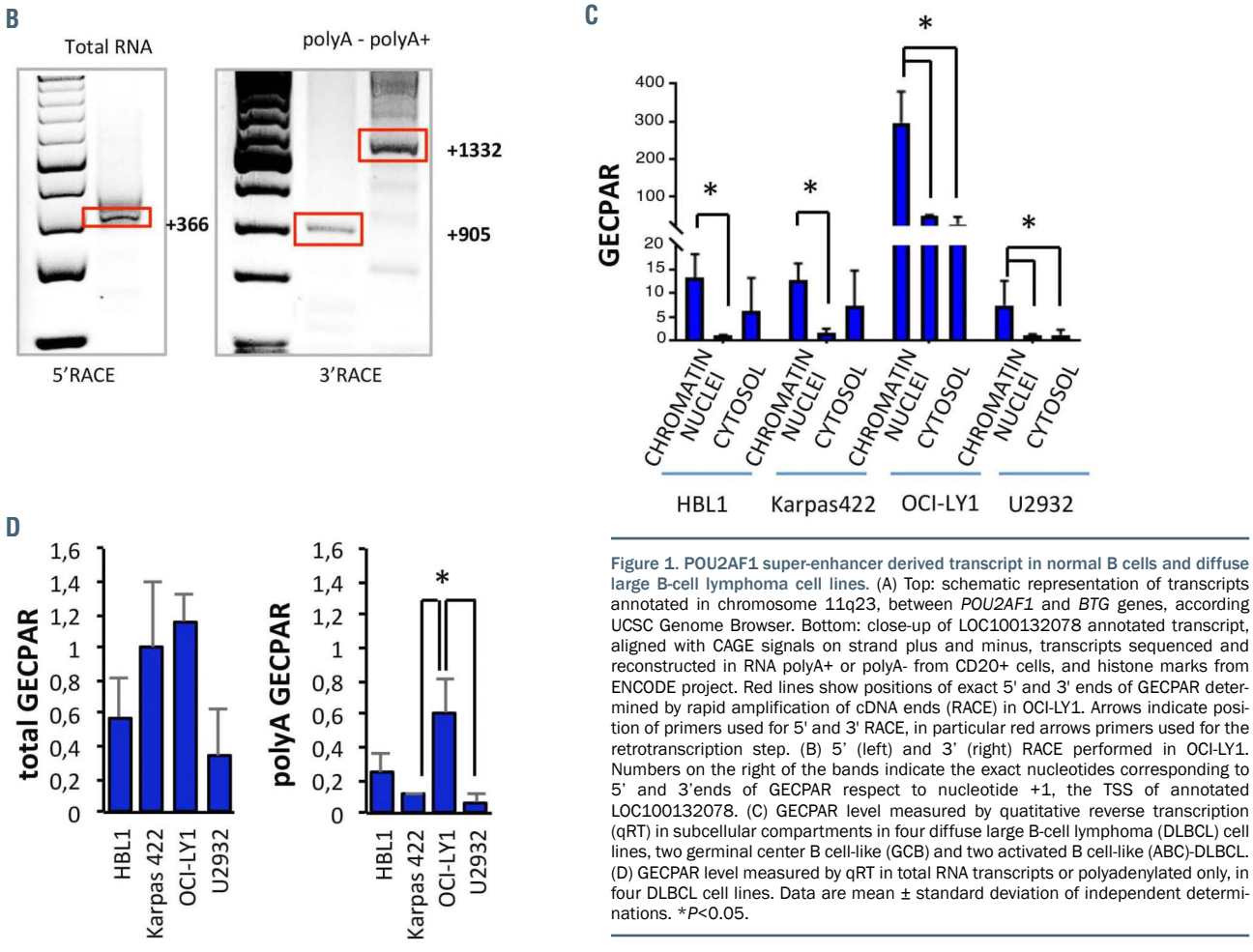

Figure 1. POU2AF1 super-enhancer derived transcript in normal B cells and diffuse large B-cell lymphoma cell lines. (A) Top: schematic representation of transcripts annotated in chromosome 11q23, between *POU2AF1* and *BTG* genes, according UCSC Genome Browser. Bottom: close-up of LOC100132078 annotated transcript, aligned with CAGE signals on strand plus and minus, transcripts sequenced and reconstructed in RNA polyA+ or polyA- from CD20+ cells, and histone marks from ENCODE project. Red lines show positions of exact 5' and 3' ends of GECPAR determined by rapid amplification of cDNA ends (RACE) in OCI-LY1. Arrows indicate position of primers used for 5' and 3' RACE, in particular red arrows primers used for the retrotranscription step. (B) 5' (left) and 3' (right) RACE performed in OCI-LY1. Numbers on the right of the bands indicate the exact nucleotides corresponding to 5' and 3'ends of GECPAR respect to nucleotide +1, the TSS of annotated LOC100132078. (C) GECPAR level measured by quatitative reverse transcription (qRT) in subcellular compartments in four diffuse large B-cell lymphoma (DLBCL) cell lines, two germinal center B cell-like (GCB) and two activated B cell-like (ABC)-DLBCL. (D) GECPAR level measured by qRT in total RNA transcripts or polyadenylated only, in four DLBCL cell lines. Data are mean ± standard deviation of independent determinations. **P*<0.05.

Strong association of a transcript to chromatin usually correlates with its lack of polyadenylation consequent rapid degradation by the RNA exosome.[32] In order to determine if these features were applicable to GECPAR, we assessed its polyadenylation status. The latter was abundant in total transcripts reverse-transcribed using random hexamers, especially in the two GCB-DLBCL cell lines. Conversely, when oligo-dT was used for reverse transcription, GECPAR was clearly detectable in only OCI-LY1, in agreement with the higher abundance of GECPAR in this cell line. (Figure 1D).

GECPAR is predominantly transcribed in germinal center diffuse large B-cell lymphoma cell lines and patients

We measured GECPAR transcription by directional qRT-PCR in 22 DLBCL cell lines (GCB, n=16; ABC, n=8). The overlapping antisense transcript was evaluated in parallel as a control. GECPAR was more frequently expressed in GCB- than ABC-DLBCL cell lines (11/16 *vs*. 0/8; *P*=0.001). In particular, it was expressed at high levels in five (OCI-LY1, OCI-LY1b, OCI-LY8, OCI-LY18, VAL), and at lower levels in six (SU-DHL-4, SU-DHL-6, SU-DHL-16, SU-DHL-8, SU-DHL-10, TOLEDO) GCB-DLBCL cell lines. The transcript was barely detectable in the remaining five GCB and in all the eight ABC-DLBCL cell lines, while the antisense transcript was more broadly expressed in all cell lines (Figure 2A).

We also evaluated GECPAR level in a total RNA-Seq dataset[33] obtained from specimens derived from normal tonsil (n=31) and DLBCL patients (GCB, n=16; ABC, n=18). The transcript was significantly more expressed in normal cells compared to tumor cells, and, in accordance with our cell lines data, it was generally more abundant in GCB- than in ABC-DLBCL (Figure 2B). The higher GECPAR expression in GCB-DLBCL was confirmed in a validation cohort of 74 patients (GCB, n=31; non-GCB, n=43) (GSE145043) (*Online Supplementary Figure S2A*) and in a second one of 350 patients (GCB, n=183; ABC, n=167) (GSE10846). Variation of GECPAR expression in DLBCL cell lines and patients might be partially explained by its unstable genomic locus.[34-36] A focal deletion of the chromosomal region containing the eRNA was observed in three of 737 mature lymphoid tumors[37-41] (*Online Supplementary Figure S2B*).

The normal tonsil derived cells were then subdivided according to B-cell maturation stage.[42] GECPAR was most highly expressed by centroblasts while naïve B cells expressed the lowest levels. This observation further underlined the specific transcription of GECPAR in GC-derived cells. We also analyzed a catalog of murine lncRNA expressed in different developmental stages of B-cell maturation.[43] Similar to our observations in humans, the murine GECPAR orthologue was mainly expressed in GC B cells, confirming the specific and conserved association of GECPAR with the GC B-cell transcriptional program (*Online Supplementary Figure S2C*).

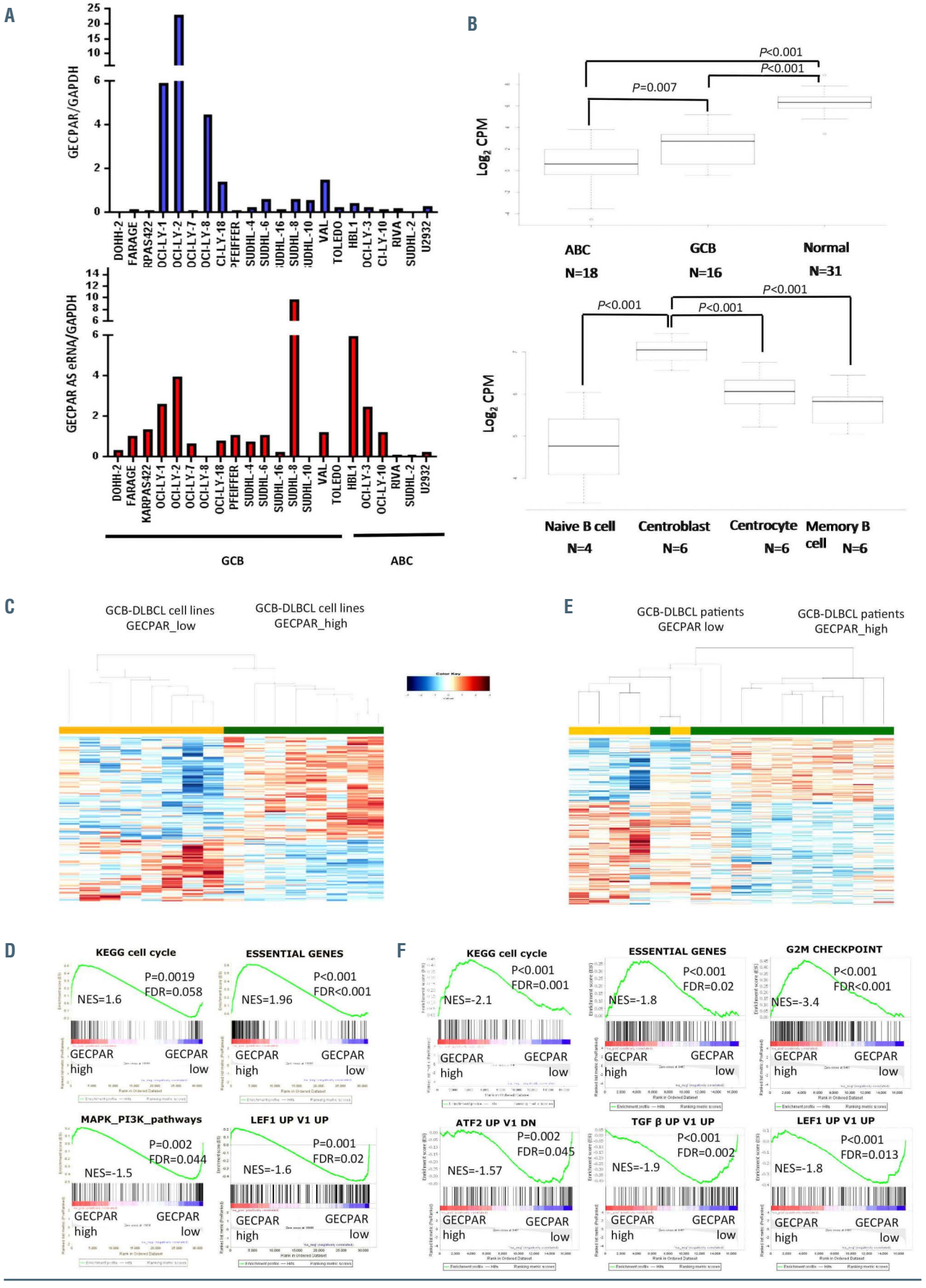

Figure 2. Legend on following page.

Figure 2. GECPAR specific expression in germinal center B cell-like cells and correlation with essential genes. (A) Top: GECPAR expression in a panel of 22 diffuse large B-cell lymphoma (DLBCL) cell lines, 16 germinal center B cell-like (GCB) and 8 activated B cell-like, bottom, expression level of GECPAR antisense transcript, measured as control. (B) Top: box plots of GECPAR expression quantified by total RNA sequencing in normal individuals or GCB- or ABC-DLBCL patients. Bottom: box plots of GECPAR expression in normal individuals stratified for cell of origin. (C) Heat map of differential gene expression, in GCB-DLBCL cell lines dichotomized for GECPAR expression. (D) Preranked gene set enrichment analysis, in GCB-DLBCL cell lines classified for GECPAR expression. (E) Heat map of differential gene expression, in 16 GCB-DLBCL patients, classified for GECPAR expression. (F) Preranked gene set enrichment analysis, in DLBCL patients classified for GECPAR expression

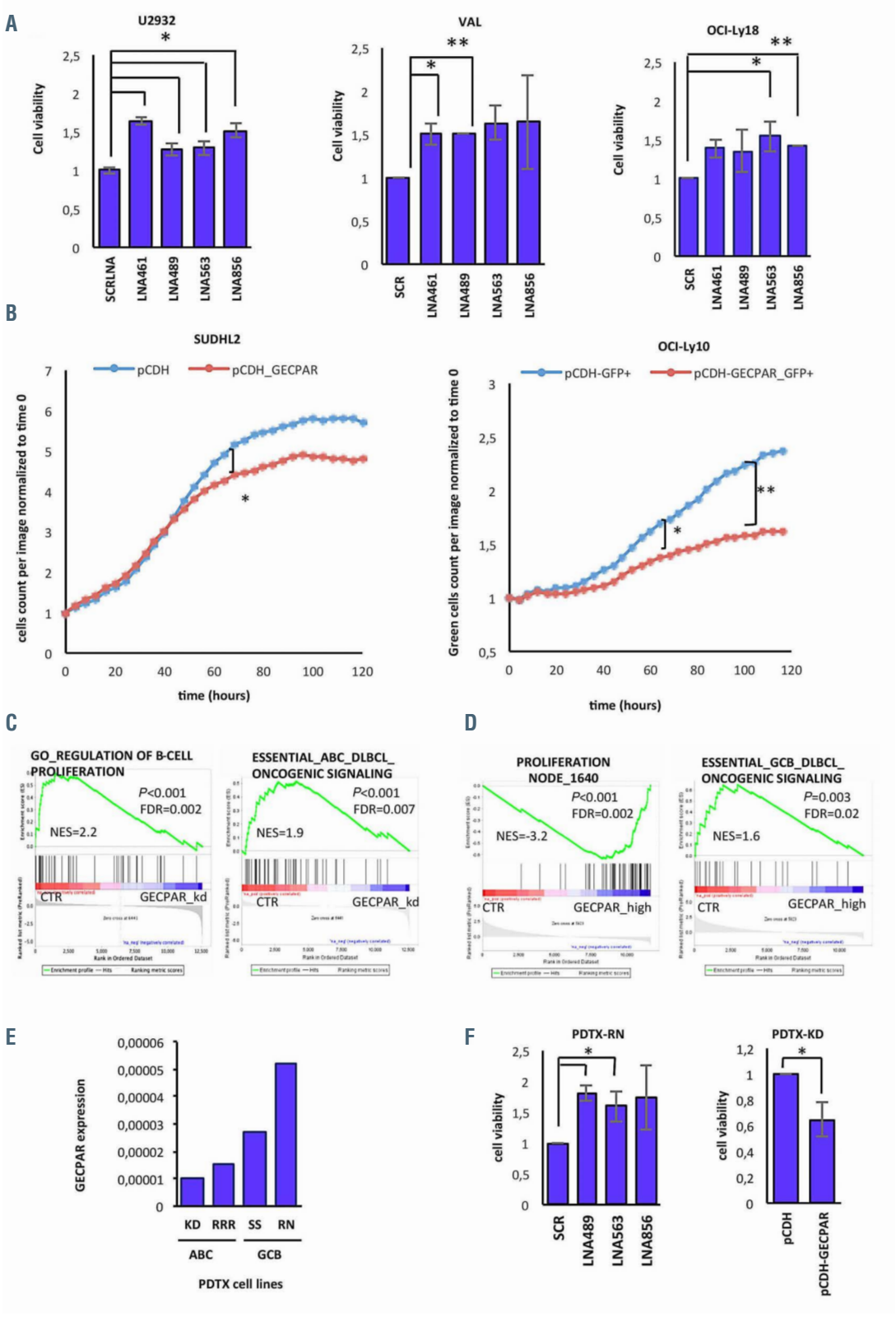

Figure 3. Legend on following page.

Figure 3. GECPAR antiproliferative activity and activation of germinal center B cell-like transcriptional program. (A) Proliferation assay after interference with GECPAR by 4 different LNA antisense oligonucleotides in U2932, VAL and OCI-LY18. Average of 3 independent experiments, *P<0.05, **P<0.01. (B) Growth curve of SUDHL2 GFP+ and SUDHL2 Gecpar- GFP+, left, or OCI-Ly10 GFP^{bright} and SUDHL2 Gecpar- GFP^{bright}, right, measured by Incucyte. Average of 3 independent experiments, *P<0.05, **P<0.01. (C) Preranked gene set enrichment analysis (GSEA) of RNA sequencing data after GECPAR knockdown (KD) in U2932. (D) Preranked GSEA of RNA sequencing data in GECPAR overexpressing SUDHL2 respect to control. (E) GECPAR expression in 4 patient-derived tumor xenograft models (PDTX) derived from 2 activated B cell-like diffuse large B-cell lymphoma (ABC-DLBCL) and 2 germinal center B cell-like (GCB)-DLBCL patients. (F) Left: proliferation assay in PDTX-RN 5 days after GECPAR KD. Right: proliferation assay in PDTX-KD 9 days after GECPAR infection.

GECPAR expression correlates with cell cycle genes and the germinal center diffuse large B-cell lymphoma oncogenic signature

In order to identify a gene expression signature associated with GECPAR, we focused on the 16 GCB-DLBCL cell lines with available expression profiling data[44] and split them in two groups based on the median GECPAR expression. We identified 122 significantly upregulated and 73 downregulated genes (absolute $\log_2$ fold change ≥0.59 and P≤0.05), that could divide GCB-DLBCL cell lines into high and low GECPAR expressers (Figure 2C; *Online Supplementary Table S1*). Transcripts that were more expressed in GECPAR- high than in GEPCAR-low expressers showed a significant enrichment of cell cycle genes and essential cell survival genes, while genes involved in MAPK and PI3K pathways, as well as LEF1 targets were comparatively less enriched (Figure 2D).

When we divided the 16 GCB-DLBCL patient specimens according to GECPAR expression GECPAR-high specimens showed an enrichment of cell cycle genes, particularly the G2M checkpoint as well as genes essential for cell survival (Figure 2E; *Online Supplementary Table S2*). Conversely, LEF1 targets and genes downstream of TGF-β and ATF2 were downregulated in DLBCL with high GECPAR expression (Figure 2F). Comparison of the genes associated with differential GECPAR expression in cell lines and clinical specimens (*Online Supplementary Table S3; Online Supplementary Figure S2D* and *E*) revealed that common genes were mainly involved in negative regulation of the cell cycle. Due to these observations, we hypothesized that GECPAR had an antiproliferative function.

GECPAR exhibits antiproliferative activity in diffuse large B-cell lymphoma cells

In order to investigate the putative antiproliferative role of GECPAR we induced degradation of GECPAR using LNA oligonucleotides in VAL, OCI-LY18 and OCI-LY1, three GCB DLBCL cell lines with high level of GECPAR and U2932, an ABC-DLBCL with moderate GECPAR expression (*Online Supplementary Figure S3A*). After 24 hours we measured *POU2AF1* mRNA and observed a negligible effect on its expression (*Online Supplementary Figure S3B*).Therefore, despite GECPAR transcription being dependent on activation of the same super-enhancer (*Online Supplementary Figure S3C* and *D*) needed for *POU2AF1* transcription (*Online Supplementary Figure S3E*), GECPAR itself was not essential for *POU2AF1* transcription. Degradation of GECPAR led to an increase in cell proliferation in all the tested cell lines, suggesting a tumor suppressor function of GECPAR (Figures 3A; *Online Supplementary Figure S3F to G*). In order to further confirm the antiproliferative activity of GECPAR, we then overexpressed GECPAR in SUDHL2 and OCI-Ly10, two ABC cell lines with low GECPAR levels. The growth of stable GFP-positive GECPAR-expressing cells (*Online Supplementary Figure S3H* and *I*) was followed by imaging in real time for 5 days. In both cell models, we measured a significant reduction in proliferation of cells overexpressing GECPAR compared to control infected cells (Figure 3B; *Online Supplementary Figure S3J)*. In particular, OCI-Ly10 expressed very intense GFP fluorescence (*Online Supplementary Figure S3I*) and could grow as a monolayer on L-poly-ornithin-coated surface allowing monitoring the growth of cells with specific green fluorescence intensity. On the contrary, SUDHL2 tended to form clusters, despite of the L-poly-ornithin coating, and the instrument could hardly discriminate fluorescence from single cells over time. In that case, we could measure the cell growth by phase contrast image analysis, more accurately. The number of total cells and of GFP expressing cells counted at time 0 are reported in the *Online Supplementary Figure S3K.*

As further confirmation, we analyzed GECPAR function also in two ABC- (PDTX-KD and PDTX-RRR) and two GCB- (PDTX-SS and PDTX-RN) DLBCL PDX models. We confirmed that GECPAR was higher in the two GCB than ABC cases (Figure 3E). Furthermore, we selected the PDX cells with the highest GECPAR expression (PDTX-RN) and we silenced GECPAR by LNA antisense oligonucleotides (*Online Supplementary Figure S4A*). GECPAR silencing increased the proliferation rate also in this model (Figure 3F; *Online Supplementary Figure S4B*). In addition, we overexpressed GECPAR in PDTX-KD cells, which had a very low amount of the transcript. We seeded the cells 24 hours (h) alter transduction and we monitored them (*Online Supplementary Figure S4C*). As for SUDHL2, although we could not monitor their growth along the whole experiment due to their tendency to form clusters, we measured GFP expression by fluorescence-activated cell sorting (FACS) (*Online Supplementary Figure S4D*), GECPAR expression by qRT-PCR (*Online Supplementary Figure S4E*) and cell viability by MTT assay (Figure 3F) after 9 days. As observed with ABC-DLBCL cell lines, also PDX cells, derived from an ABC-DBCL with low GECPAR expression, reduced their proliferation rate after GECPAR overexpression.

GECPAR polarizes cells towards a germinal center B cell-like transcriptional program

We performed transcriptional analysis after GECPAR knockdown (KD) and overexpression in U2932 and SUDHL2 cells, respectively. Knockdown of GECPAR resulted in 1,099 significantly downregulated and 528 upregulated genes (*Online Supplementary Table S4*), while overexpression of GECPAR led to significant upregulation of 3,152 genes and downregulation of 787 genes (*Online Supplementary Table S5*). Genes upregulated after GECPAR silencing comprised proliferation genes, which were conversely downregulated in GECPAR-overexpressing cells. Further, while U2932, an ABC-DLBCL with moderate basal GECPAR expression still presented an enrichment of oncogenic genes typical of ABC-DLBCL after GECPAR knockdown (Figure 3D; *Online Supplementary*

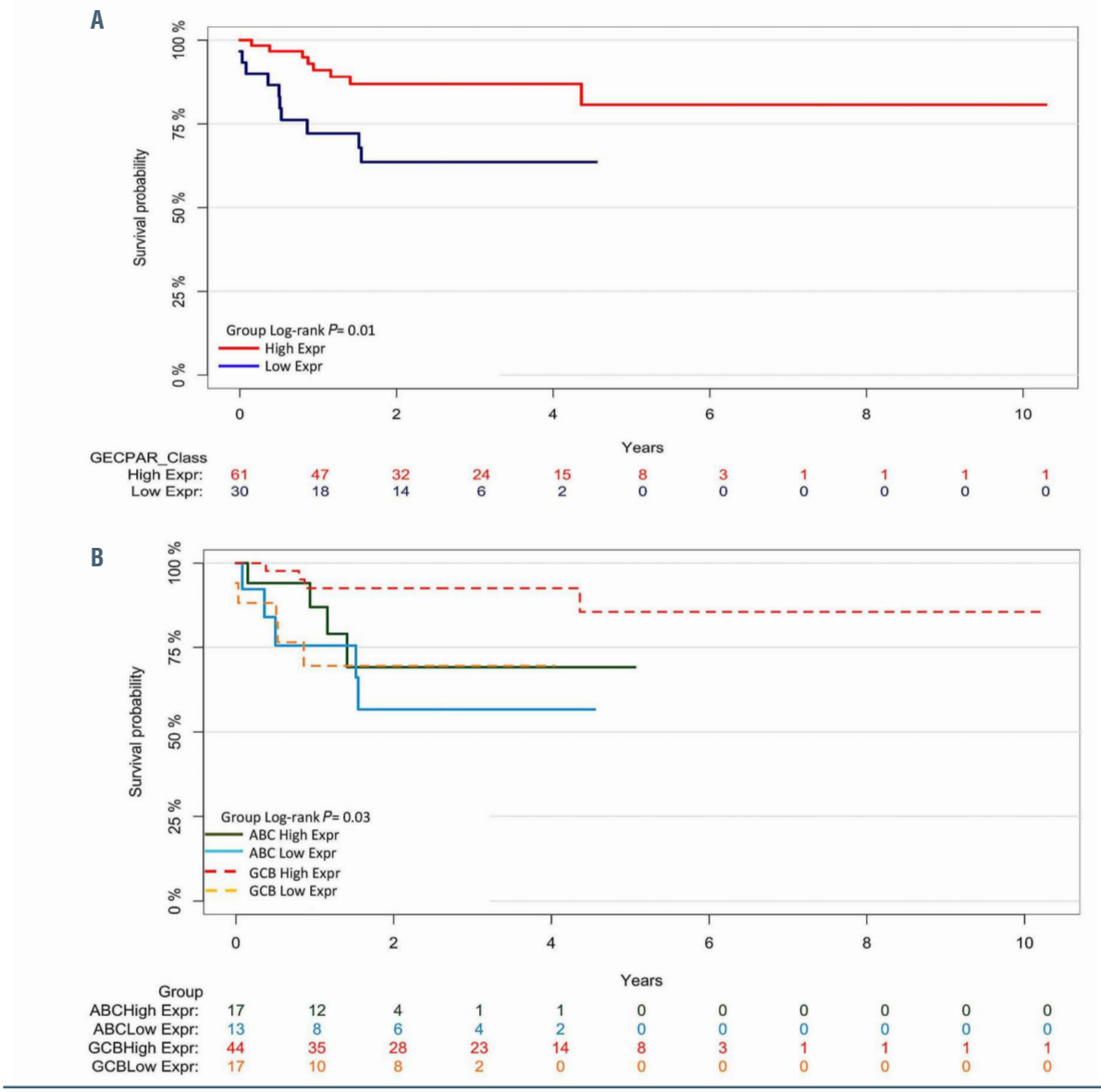

Figure 4. GECPAR has a favorable impact on the outcome of germinal center B cell-like diffuse large B-cell lymphoma patients. (A) Kaplan-Meier curves of diffuse large B-cell lymphoma (DLBCL) patients treated with R-CHOP (rituximab, cyclophosphamide, doxorubicin, vincristine, prednisone) and stratified for GECPAR expression. (B) Kaplan-Meier curves of germinal center B cell-like (GCB)- and activated B cell-like (ABC)-DLBCL patients treated with R-CHOP and stratified for GECPAR expression.

Table S6, left), the other ABC-DLBCL SUDHL2, showed an enrichment of GCB-DLBCL genes (Figure 3E; *Online Supplementary Table S6*, right), after GECPAR overexpression. Finally, GECPAR transcription was strongly induced by anti-IgM stimulation of the BCR (*Online Supplementary Figure S3L*). Together, these observations provided further support of GECPAR's role in maintaining the GC transcriptional program.

GECPAR expression has favorable prognostic impact in germinal center diffuse large B-cell lymphoma patients

We assessed the expression of GECPAR in 91 DLBCL patients treated with R-CHOP (rituximab, cyclophosphamide, doxorubicin, vincristine, prednisone) and its potential impact on the clinical outcome. We classified patients in three subgroups: low expressor (below the 15th percentile of the whole population), high expressor (over the 70th percentile), and neutral (in between). High expressor patients had a higher survival probability than low expressor (P=0.01) (Figure 4A). Then, we looked at the high and low expressors based on their cell of origin. As expected by previous analysis (*Online Supplementary Figure S2A*), the high expressors were mainly GCB patients. However, among GCB-DLBCL patients, cases with low GECPAR expression had the same risk of death as ABC patients, while the high expressors showed a better outcome (P=0.03). All together, these observations further sustain the tumor suppressor role we attributed to GECPAR based on our *in vitro* experiments.

GECPAR acts *in trans* regulating cell growth and differentiation by means of Wnt pathway

In order to identify the genes directly regulated by GECPAR, we performed CHART-Seq in OCI-LY1 and U2932. We identified 4,172 peaks in OCI-LY1 and 692 peaks in U2932 (Figure 5A; *Online Supplementary Tables S7* and S8).

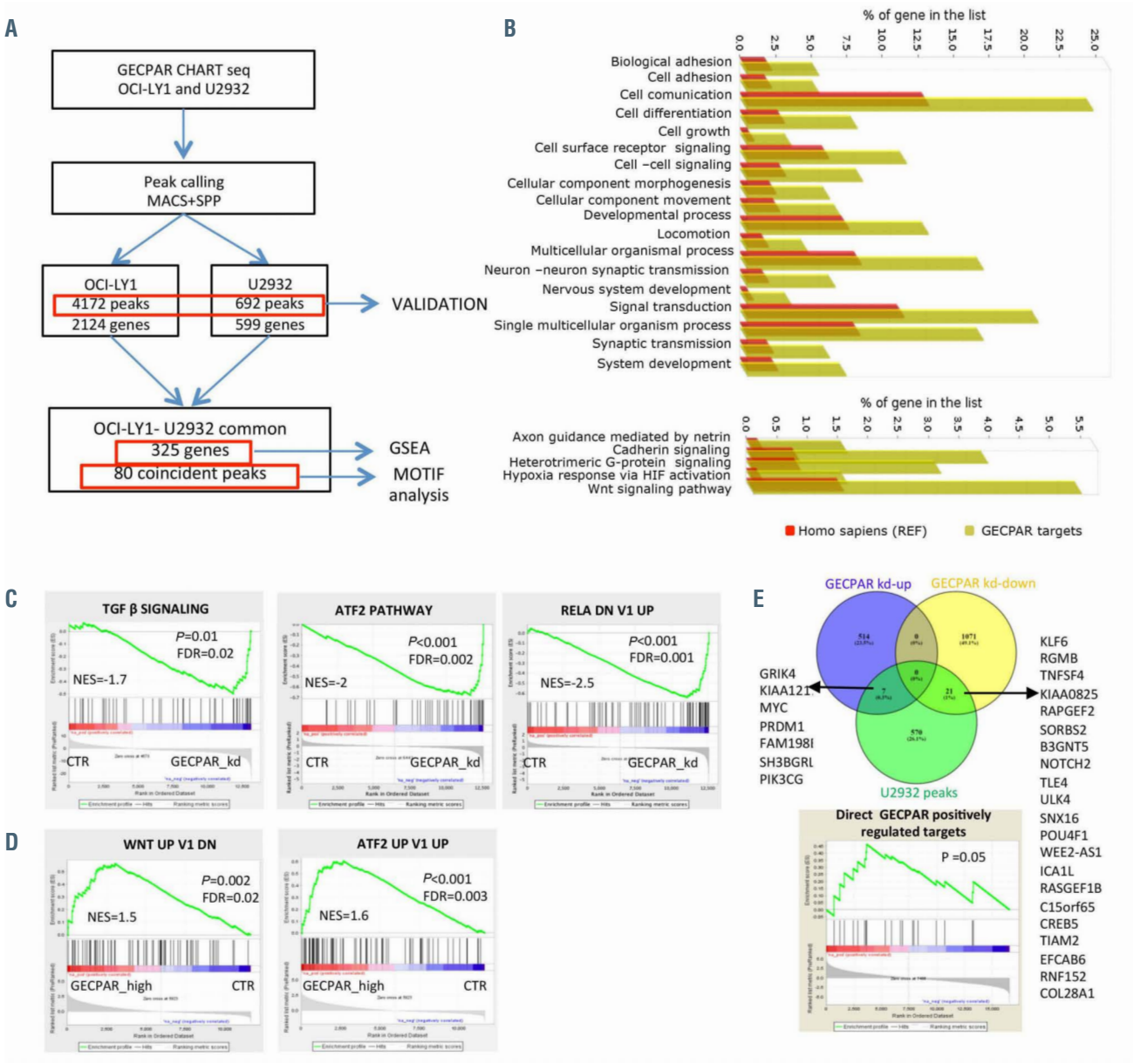

Figure 5 GECPAR *in trans* transcriptional regulatory function. (A) Pipeline of CHART experiment and analysis. (B) Panther gene ontology classification of 325 GECPAR target genes identified both in OCI-LY1 and U2932 by CHART. (C) Preranked gene set enrichment analysis (GSEA) of RNA sequencing (RNA-Seq) data after GECPAR knockdown (KD) in U2932. (D) Preranked GSEA of RNA-Seq data after GECPAR overexpression in SUDHL2. (E) Top: Venn diagram crossing genes with GECPAR binding detected by CHART sequencing and significant expression modulation after GECPAR KD in U2932. Direct downregulated (left) and upregulated (right) GECPAR targets are listed. Bottom, preranked GSEA of direct GECPAR positively regulated targets, in germinal center B cell-like diffuse large B-cell lymphoma patients dichotomized for GECPAR expression

The most prominent peaks were validated in an independent CHART experiment by qRT, confirming the robustness of both the enrichment experiment and downstream analysis (*Online Supplementary Figure S5B)*. As an additional control, we measured the levels of transcripts associated with GECPAR binding including *CREBBP*, *CREB5*, *TLE4* and *CYLD*. After 24 h of GECPAR silencing with LNA oligonucleotides in U2932, the levels of these transcripts were reduced by 50-80% (*Online Supplementary Figure S5C*) and after 72 h we noticed a reduction of 50% also in the level of CYLD and TLE4 proteins (*Online Supplementary Figure S5D*, top). We also measured the increase in protein levels in SUDHL2 and OCI-Ly10 stably overexpressing GECPAR, for TLE4 and CYLD, or CREBBP and CYLD, respectively (*Online Supplementary Figure S5D*, bottom).

GECPAR capture was done with a set of probes, selected after RNAseH sensitivity assay (*Online Supplementary Figure S5A*). Only peaks called by two different algorithms (MACS and SPP) were taken in account: 4,172 in OCI-LY1 and 692 in U2932 (Figure 5A, *Online Supplementary Tables S7* and *S8*). We identified a putative GECPAR binding motif. Among 78 CHARTseq peaks that fell within an interval of 10 kb in both cell lines there was a significant putative GECPAR binding motif (13 matches, *P*-values between 2.15×10^{-7} and 1.9×10^{-9}) (*Online Supplementary Figure S5E*),

In order to identify biological processes directly influenced by GECPAR independently of the cell of origin, we analysed 325 genes bound by the eRNA in both OCI-LY1 and U2932. The most significantly enriched classes of

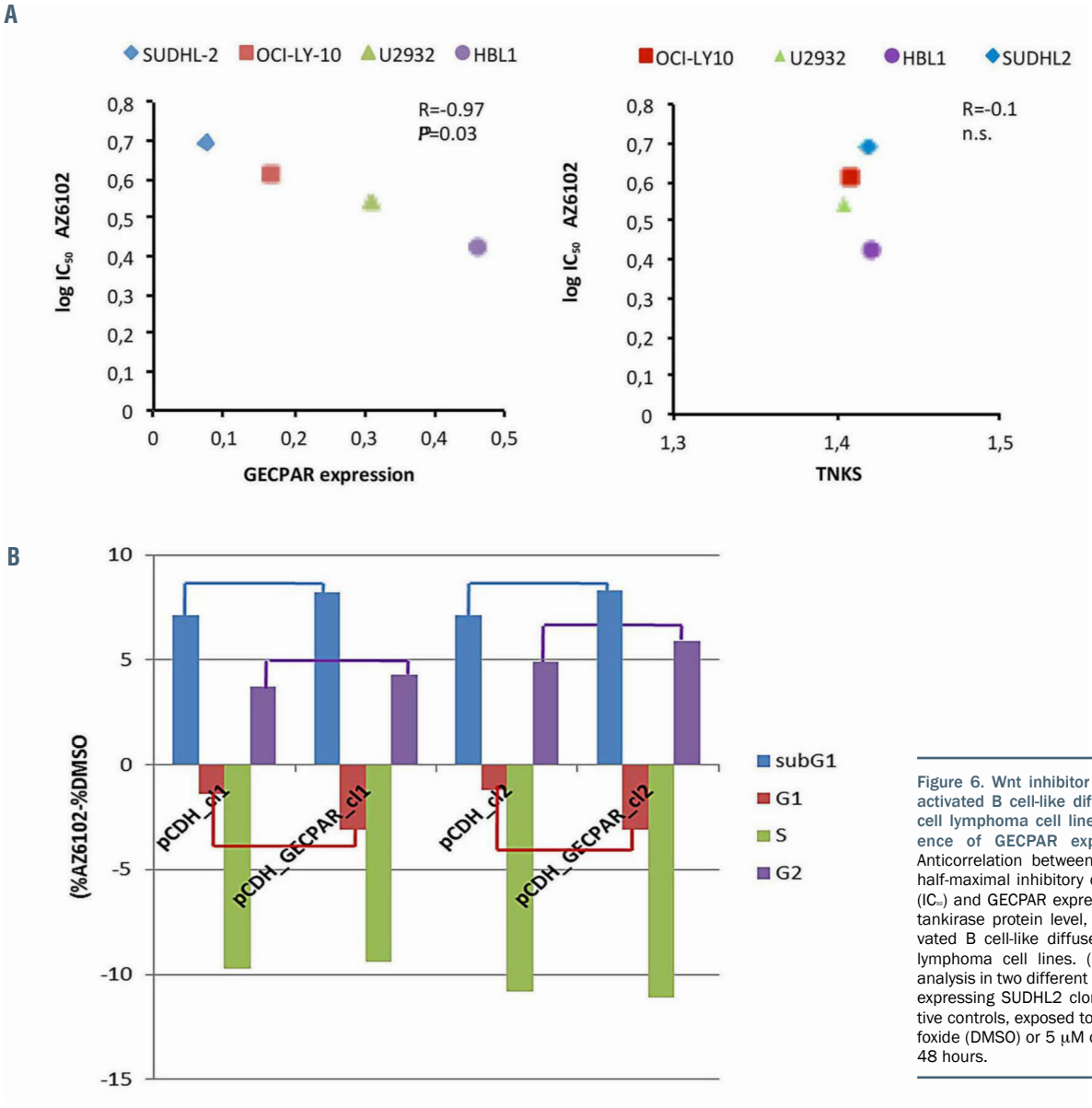

Figure 6. Wnt inhibitor sensitivity in activated B cell-like diffuse large B-cell lymphoma cell lines in dependence of GECPAR expression. (A) Anticorrelation between AZ6102 log half-maximal inhibitory concentration (IC_{50}) and GECPAR expression, left, or tankirase protein level, right, in activated B cell-like diffuse large B-cell lymphoma cell lines. (B) Cell cycle analysis in two different GECPAR overexpressing SUDHL2 clones and relative controls, exposed to dimethyl sulfoxide (DMSO) or 5 μM of AZ6102 for 48 hours.

genes belonged to the Wnt signaling pathway, cell growth and differentiation (Figure 5B). RNA-Seq data after GECPAR knockdown showed modulation of three pathways associated with development, differentiation and proliferation and known to cross-talk with the Wnt pathway, such as TGF β, NF-κB and MAPK (Figure 5C).[45-47] Negative regulators of TGF-β pathways including SMAD7, SMURF1 and SMURF2 (*Online Supplementary Figure S6A*) and negative regulators of MAPK signaling, DUSP1, DUSP8 and DUSP10 (*Online Supplementary Figure S6B*), were downregulated, after GECPAR silencing. Some of the downregulated genes belonging to the aforementioned pathways are also negatively regulated by NF-κB (*Online Supplementary Figure S6C*). Notably, WNT and MAPK pathways were also affected in SUDHL2 cells overexpressing GECPAR (Figure 5D).

Intersection of CHARTseq and RNA-Seq data for U2932 cells with GECPAR knockdown identified *MYC* and *PRDM1* among seven genes negatively regulated by GECPAR, indicating that the eRNA influenced both the proliferative capability, reducing MYC, and the terminal differentiation to plasma cells, reducing PRDM1, the genes coding for BLIMP1. Interestingly, 21 direct GECPAR upregulated targets were positively correlated with GECPAR expression also in GCB-DLBCL specimens (Figure 5E). Among them there were *KLF6*, *NOTCH2*, components of BMP, cAMP and TNF-α pathways. Strikingly, we also identified *TLE4* (Groucho), which forms a corepressor complex with TCF/LEF1 and recruits HDAC to inhibit transactivation of TCF/LEF1 target genes.[48]

Our identification of GECPAR involvement in Wnt signaling prompted us to evaluate the activity of the tankyrase 1/2 (TNKS1/2) inhibitor, AZ6102, that prevents nuclear translocation of β-catenin.[49] For the four ABC-DLBCL cell lines we tested, GECPAR expression and sensitivity to AZ6102 were significantly anticorrelated (Figure 6A), suggesting that expression of GECPAR sensitized cells to Wnt pathway inhibition. All seven GCB-DLBCL cell lines tested where equally sensitive to Wnt pathway inhibition (*Online Supplementary Figure S7*). The differential sensitivity to AZ6102 in ABC-DLBCL was not related to tankyrase expression, since protein

levels were similar for the four cell lines (Figure 6A). Further, GECPAR overexpressing SUDHL2 cells were more sensitive to Wnt inhibition than the parental control, in terms of cell cycle perturbation. AZ6102 treatment more readily caused G2/M arrest, subG1 accumulation and decreased re-entry in G1 in GECPAR overexpressing cells (Figure 6B).

Discussion

eRNA have recently started to be recognized as potent modulators of coding gene transcription.[50,51] Here, we provide the first evidence of a lncRNA, transcribed in a SE specifically active during maturation of GC B cells, which plays an antiproliferative role in DLBCL models and is associated with favorable clinical outcome in GCB-DLBCL patients.

The lncRNA LOC100132078 was previously annotated as an unknown ncRNA, mainly expressed in lymph nodes and testis,[52] and reported among p53-induced eRNA in breast cancer.[53] Since it mapped inside a SE relevant for GC formation[3,23,43] and in a site of recurrent genomic instability in lymphoid tumors,[34-36] we elucidated its role in DLBCL, the neoplastic counterpart derived from GC B cells. We defined this lncRNA as eRNA according to the main features of this class of ncRNA: it was encoded within a SE; it was a non-polyA chromatin-associated transcript: its expression, highly cell type specific, was dependent on enhancer activation. We also identified a stabilized 970 nucleotide-long transcript, which, based on its expression pattern, we named GECPAR. It was less expressed in DLBCL samples than in normal tonsil B cells and *in vitro* experiments showed an inverse correlation with cell proliferation, suggesting an antitumoral function. The latter was further supported by the association between high GECPAR expression and favorable outcome in GCB DLBCL patients. GECPAR did not seem to act by *in cis* transactivation of the juxtaposed *POU2AF1* gene, which is strongly expressed in GC-derived malignancies.[22] Indeed, although GECPAR and *POU2AF1* transcript levels were correlated in cell lines and in clinical specimens, silencing of the eRNA did not strongly impair expression of the coding gene. This is not uncommon and might be due to redundant functions of multiple enhancers that target a given promoter.[54] On the contrary, GECPAR showed *in trans* activity and directly regulated the expression of several transcripts, mainly involved in cell growth and differentiation. These regulated genes were identified as common GECPAR targets in a GCB- and an ABC- DLBCL cell line, both of which had constitutively high GECPAR expression.

GECPAR expression was increased after BCR activation, an event that causes transcriptional reprograming of B cells. The exogenous overexpression of GECPAR in an ABC-DLBCL cell line confirmed its ability to switch the lymphoma cell towards the GCB-DLBCL transcriptional signature.

Nuclear enriched lncRNA regulating transcription *in trans* have been described and they often modulate cell development.[43,55] We propose that GECPAR is used by normal GC B cells to fine-tune the balance between proliferation and differentiation by directly repressing *MYC* and *PRDM1* expression. MYC has a stage-specific role in the GC, particularly in light zone B cells, namely centrocytes, from which GCB-DLBCL tumor cells derive. After antigen-driven selection, B cells that still need to improve their antigen affinity can re-enter in the dark zone where they undergo additional cycles of somatic hypermutation. This so-called "cyclic re-entry" is critical for maintaining the GC and is induced by the re-expression of MYC via BCR activation through NF-κB and FOXO1.[56,57] We propose GECPAR as a key surveillant of this process, as it directly represses MYC in that phase. Termination of the GC reaction is modulated by NF-κB activation downstream of the BCR. It induces IRF4, master regulator of terminal B-cell differentiation which in turn activates the plasma cell master regulator BLIMP1, encoded by *PRDM1*.[58] GECPAR itself directly represses *PRDM1*, impeding terminal differentiation into plasma blast. In conclusion, GECPAR, which is induced by BCR activation, would retain B cells in the GC light zone, reducing the tendency to re-enter in the dark zone or to exit and differentiate to plasma cells.

GECPAR also reduces B-cell proliferation rate and the tendency to differentiate, possibly by directly inducing TLE4, a negative repressor of TCF/LEF1. LEF1 is the key mediator of nuclear Wnt signaling and is important in lymphopoiesis. LEF1 is overexpressed in the nucleus of approximately 40% of DLBCL.[59] MYC and Wnt pathway are connected in a positive feedback-loop involving LEF1.[60] GECPAR, which directly inhibited MYC expression, indirectly enhanced its antiproliferative activity via TLE4 that contributed to the arrest of terminal differentiation induced by NF-κB. Indeed, GECPAR expression was inversely correlated with many LEF1 targets, in both DLBCL cell lines and specimens, and some of them were related to NF-κB regulation. Moreover, GECPAR silencing induced upregulation of important NF-κB genes, such as *CARD11*, *REL* and *IKBKB*, supporting the link between GECPAR and Wnt/NF-κB crosstalk. Several bidirectional connections between Wnt and NF-κB pathways[45] have been reported in cancer and in particular, in DLBCL.[61] We propose GECPAR as an additional layer of control of NF-κB activation in GC B cells, pausing terminal differentiation to plasma blasts.

The greater sensitivity of ABC-DLBCL with high GECPAR expression to pharmacological inhibition of Wnt further supports the relationship between GECPAR and Wnt pathway regulation and uncovers alternative therapeutic options for ABC-DLBCL patients.

In conclusion, our work describes a novel mechanism of regulation of GC differentiation, which might contribute to DLBCL pathogenesis, and could help in understanding the heterogeneity of this disease.

Disclosures

LC received a travel grant from HTG. FB received institutional research funds from Acerta, ADC Therapeutics, Bayer AG, Cellestia, CTI Life Sciences, EMD Serono, Helsinn, ImmunoGen, Menarini Ricerche, NEOMED Therapeutics 1, Nordic Nanovector ASA, Oncology Therapeutic Development and PIQUR Therapeutics AG; received consultancy fees from Helsinn and Menarini; provided expert statements to HTG; received travel grants from Amgen, Astra Zeneca, Jazz Pharmaceuticals and PIQUR Therapeutics. All the other authors have nothing to disclose.

Contributions

SN performed experiments, analyzed and interpreted data,

wrote the manuscript; SN and FB conceived and supervised the study, and extensively reviewed the manuscript; LC performed data mining; AR performed next generation sequencing; AAM reviewed the manuscript; FS, FG, FZ, GS performed experiments; MF, SB provided bioinformatic support; MTC, AC, PG and GI provided clinical samples data; OE and GI provided transcriptome data. All authors read and edited the manuscript.

Funding

Partially supported by funds from Gelu Foundation (to BF).

References

1. Loven J, Hoke HA, Lin CY, et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell. 2013;153(2):320-334.
2. Whyte WA, Orlando DA, Hnisz D, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell. 2013;153(2):307-319.
3. Zhang J, Vlasevska S, Wells VA, et al. The CREBBP acetyltransferase is a haploinsufficient tumor suppressor in B-cell lymphoma. Cancer Discov. 2017;7(3):322-337.
4. Kim TK, Hemberg M, Gray JM, et al. Widespread transcription at neuronal activity-regulated enhancers. Nature. 2010;465 (7295):182-187.
5. Arnold PR, Wells AD, Li XC. Diversity and emerging roles of enhancer RNA in regulation of gene expression and cell fate. Front Cell Dev Biol. 2019;7:377.
6. Li W, Notani D, Rosenfeld MG. Enhancers as non-coding RNA transcription units: recent insights and future perspectives. Nat Rev Genet. 2016;17(4):207-223.
7. Liu F. Enhancer-derived RNA: a primer. Genomics Proteomics Bioinformatics. 2017;15(3):196-200.
8. Jiao W, Chen Y, Song H, et al. HPSE enhancer RNA promotes cancer progression through driving chromatin looping and regulating hnRNPU/p300/EGR1/HPSE axis. Oncogene. 2018;37(20):2728-2745.
9. Li W, Notani D, Ma Q, et al. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. Nature. 2013;498 (7455):516-520.
10. Fanucchi S, Shibayama Y, Burd S, et al. Chromosomal contact permits transcription between coregulated genes. Cell. 2013;155 (3):606-620.
11. Tan SH, Leong WZ, Ngoc PCT, et al. The enhancer RNA ARIEL activates the oncogenic transcriptional program in T-cell acute lymphoblastic leukemia. Blood. 2019;134 (3):239-251.
12. Marques AC, Hughes J, Graham B, et al. Chromatin signatures at transcriptional start sites separate two equally populated yet distinct classes of intergenic long noncoding RNAs. Genome Biol. 2013;14(11):R131.
13. Alvarez-Dominguez JR, Knoll M, Gromatzky AA, et al. The super-enhancer-derived alncRNA-EC7/Bloodlinc potentiates red blood cell development in trans. Cell Rep. 2017;19(12):2503-2514.
14. Agirre X, Meydan C, Jiang Y, et al. Long non-coding RNAs discriminate the stages and gene regulatory states of human humoral immune response. Nat Commun. 2019;10(1):821.
15. Verma A, Jiang Y, Du W, et al. Transcriptome sequencing reveals thousands of novel long non-coding RNAs in B cell lymphoma. Genome Med. 2015;7:110.
16. Mlynarczyk C, Fontan L, Melnick A. Germinal center-derived lymphomas: the darkest side of humoral immunity. Immunol Rev. 2019;288(1):214-239.
17. Pasqualucci L, Dalla-Favera R. Genetics of diffuse large B-cell lymphoma. Blood. 2018;131(21):2307-2319.
18. Wright GW, Huang DW, Phelan JD, et al. A probabilistic classification tool for genetic subtypes of diffuse large B cell lymphoma with therapeutic implications. Cancer Cell. 2020;37(4):551-568.
19. Chapuy B, Stewart C, Dunford AJ, et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. Nat Med. 2018;24(5):679-690.
20. Ceribelli M, Kelly PN, Shaffer AL, et al. Blockade of oncogenic IkappaB kinase activity in diffuse large B-cell lymphoma by bromodomain and extraterminal domain protein inhibitors. Proc Natl Acad Sci U S A. 2014;111(31):11365-11370.
21. Reddy A, Zhang J, Davis NS, et al. Genetic and functional drivers of diffuse large B cell lymphoma. Cell. 2017;171(2):481-494.
22. Teitell MA. OCA-B regulation of B-cell development and function. Trends Immunol. 2003;24(10):546-553.
23. Chapuy B, McKeown MR, Lin CY, et al. Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. Cancer Cell. 2013;24(6):777-790.
24. Gaudio E, Tarantelli C, Spriano F, et al. Targeting CD205 with the antibody drug conjugate MEN1309/OBT076 is an active new therapeutic strategy in lymphoma models. Haematologica. 2020;105(11):2584-2591.
25. Napoli S, Piccinelli V, Mapelli SN, et al. Natural antisense transcripts drive a regulatory cascade controlling c-MYC transcription. RNA Biol. 2017;14(12):1742-1755.
26. Vance KW. Mapping long noncoding RNA chromatin occupancy using capture hybridization analysis of RNA targets (CHART). Methods Mol Biol. 2017;1468:39-50.
27. Sexton AN, Machyna M, Simon MD. Capture hybridization analysis of DNA targets. Methods Mol Biol. 2016;1480:87-97.
28. Kharchenko PV, Tolstorukov MY, Park PJ. Design and analysis of ChIP-seq experiments for DNA-binding proteins. Nat Biotechnol. 2008;26(12):1351-1359.
29. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9(9):R137.
30. Vance KW, Sansom SN, Lee S, et al. The long non-coding RNA Paupar regulates the expression of both local and distal genes. EMBO J. 2014;33(4):296-311.
31. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57-74.
32. Pefanis E, Wang J, Rothschild G, et al. RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. Cell. 2015;161(4):774-789.
33. Teater M, Dominguez PM, Redmond D, et al. AICDA drives epigenetic heterogeneity and accelerates germinal center-derived lymphomagenesis. Nat Commun. 2018;9(1): 222.
34. Poretti G, Kwee I, Bernasconi B, et al. Chromosome 11q23.1 is an unstable region in B-cell tumor cell lines [Research Support, Non-U.S. Gov't]. Leuk Res. 2011;35(6):808-813.
35. Auer RL, Starczynski J, McElwaine S, et al. Identification of a potential role for POU2AF1 and BTG4 in the deletion of 11q23 in chronic lymphocytic leukemia. Genes Chromosomes Cancer. 2005;43(1):1-10.
36. Zhao C, Inoue J, Imoto I, et al. POU2AF1, an amplification target at 11q23, promotes growth of multiple myeloma cells by directly regulating expression of a B-cell maturation factor, TNFRSF17. Oncogene. 2008;27(1):63-75.
37. Chigrinova E, Rinaldi A, Kwee I, et al. Two main genetic pathways lead to the transformation of chronic lymphocytic leukemia to Richter syndrome. Blood. 2013;122(15): 2673-2682.
38. Boi M, Rinaldi A, Kwee I, et al. PRDM1/BLIMP1 is commonly inactivated in anaplastic large T-cell lymphoma. Blood. 2013;122(15):2683-2693.
39. Martinez N, Almaraz C, Vaque JP, et al. Whole-exome sequencing in splenic marginal zone lymphoma reveals mutations in genes involved in marginal zone differentiation. Leukemia. 2014;28(6):1334-1340.
40. Rinaldi A, Kwee I, Young KH, et al. Genome-wide high resolution DNA profiling of hairy cell leukaemia. Br J Haematol. 2013;162(4):566-569.
41. Rossi D, Trifonov V, Fangazio M, et al. The coding genome of splenic marginal zone lymphoma: activation of NOTCH2 and other pathways regulating marginal zone development. J Exp Med. 2012;209(9):1537-1551.
42. De Silva NS, Klein U. Dynamics of B cells in germinal centres. Nat Rev Immunol. 2015;15(3):137-148.
43. Brazao TF, Johnson JS, Muller J, et al. Long noncoding RNAs in B-cell development and activation. Blood. 2016;128(7):e10-19.
44. Tarantelli C, Gaudio E, Arribas AJ, et al. PQR309 is a novel dual PI3K/mTOR inhibitor with preclinical antitumor activity in lymphomas as a single agent and in combination therapy. Clin Cancer Res. 2018;24 (1):120-129.
45. Ma B, Hottiger MO. Crosstalk between Wnt/beta-Catenin and NF-kappaB signaling pathway during inflammation. Front Immunol. 2016;7:378.
46. Guo X, Wang XF. Signaling cross-talk between TGF-beta/BMP and other pathways. Cell Res. 2009;19(1):71-88.
47. Grumolato L, Liu G, Haremaki T, et al. beta-Catenin-independent activation of TCF1/LEF1 in human hematopoietic tumor cells through interaction with ATF2 transcription factors. PLoS Genet. 2013;9(8): e1003603.
48. Daniels DL, Weis WI. Beta-catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. Nat Struct Mol Biol. 2005;12(4):364-371.

49. Johannes JW, Almeida L, Barlaam B, et al. Pyrimidinone nicotinamide mimetics as selective tankyrase and wnt pathway inhibitors suitable for in vivo pharmacology. ACS Med Chem Lett. 2015;6(3):254-259.
50. Franco HL, Nagari A, Malladi VS, et al. Enhancer transcription reveals subtype-specific gene expression programs controlling breast cancer pathogenesis. Genome Res. 2018;28(2):159-170.
51. Zhang Z, Lee JH, Ruan H, et al. Transcriptional landscape and clinical utility of enhancer RNAs for eRNA-targeted therapy in cancer. Nat Commun. 2019;10(1):4562.
52. Fagerberg L, Hallstrom BM, Oksvold P, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol Cell Proteomics. 2014;13(2):397-406.
53. Leveille N, Melo CA, Rooijers K, et al. Genome-wide profiling of p53-regulated enhancer RNAs uncovers a subset of enhancers controlled by a lncRNA. Nat Commun. 2015;6:6520.
54. Li X, Fu XD. Chromatin-associated RNAs as facilitators of functional genomic interactions. Nat Rev Genet. 2019;20(9):503-519.
55. Guttman M, Donaghey J, Carey BW, et al. lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature. 2011;477(7364):295-300.
56. Pasqualucci L. Molecular pathogenesis of germinal center-derived B cell lymphomas. Immunol Rev. 2019;288(1):240-261.
57. Luo W, Weisel F, Shlomchik MJ. B cell receptor and CD40 signaling are rewired for synergistic induction of the c-Myc transcription factor in germinal center B cells. Immunity. 2018;48(2):313-326.
58. Boi M, Zucca E, Inghirami G, et al. PRDM1/BLIMP1: a tumor suppressor gene in B and T cell lymphomas. Leuk Lymphoma. 2015;56(5):1223-1228.
59. Tandon B, Peterson L, Gao J, et al. Nuclear overexpression of lymphoid-enhancer-binding factor 1 identifies chronic lymphocytic leukemia/small lymphocytic lymphoma in small B-cell lymphomas. Mod Pathol. 2011;24(11):1433-1443.
60. Hao YH, Lafita-Navarro MC, Zacharias L, et al. Induction of LEF1 by MYC activates the WNT pathway and maintains cell proliferation. Cell Commun Signal. 2019;17(1):129.
61. Bognar MK, Vincendeau M, Erdmann T, et al. Oncogenic CARMA1 couples NF-kappaB and beta-catenin signaling in diffuse large B-cell lymphomas. Oncogene. 2016;35(32):4269-4281.