Different expression of DNMT1, PCNA, MCM2, CDT1, EZH2, GMNN and EP300 genes in lymphomagenesis of low vs. high grade lymphoma

Pavlov, Katarina Horvat; Tadić, Vanja; Palković, Pamela Bašić; Sasi, Biljana; Magdić, Nives; Petranović, Matea Zajc; Klasić, Marija; Hančić, Suzana; Gršković, Paula; Matulić, Maja; ...

Source / Izvornik: **Pathology - Research and Practice, 2022, 239, 154170 - 154170**

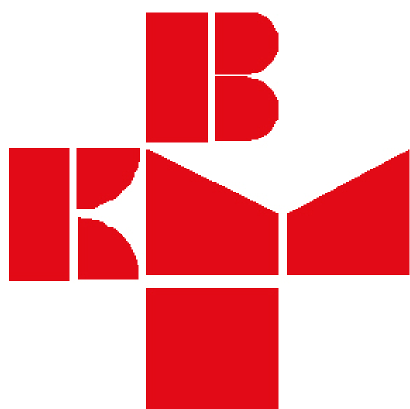
Journal article, Published version
Rad u časopisu, Objavljena verzija rada (izdavačev PDF)

https://doi.org/10.1016/j.prp.2022.154170

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:264:495235

Rights / Prava: Attribution-NonCommercial-NoDerivatives 4.0 International/Imenovanje-Nekomercijalno-Bez prerada 4.0 međunarodna

Download date / Datum preuzimanja: **2025-01-17**



Repository / Repozitorij:

Merkur University Hospital Repository

dabar
DIGITALNI AKADEMSKI ARHIVI I REPOZITORIJI

Pathology - Research and Practice 239 (2022) 154170

Contents lists available at ScienceDirect

Pathology - Research and Practice

journal homepage: www.elsevier.com/locate/prp

Different expression of *DNMT1, PCNA, MCM2, CDT1, EZH2, GMNN* and *EP300* genes in lymphomagenesis of low vs. high grade lymphoma

Katarina Horvat Pavlov [a], Vanja Tadić [b], Pamela Bašić Palković [b], Biljana Sasi [b], Nives Magdić [b], Matea Zajc Petranović [c], Marija Klasić [b], Suzana Hančić [a], Paula Gršković [b], Maja Matulić [b], Slavko Gašparov [a,d], Mara Dominis [a,d], Petra Korać [b,*]

[a] *Institute of Clinical Pathology and Cytology, Merkur University Hospital, Zagreb, Croatia*
[b] *Division of Molecular Biology, Department of Biology, Faculty of Science, University of Zagreb, Croatia*
[c] *Institute for Anthropological Research, Zagreb, Croatia*
[d] *Department of Pathology, Medical School Zagreb, University of Zagreb, Zagreb, Croatia*

ARTICLE INFO

Keywords:
Replication
Chromatin assembly
Germinal centre B lymphocytes
Follicular lymphoma
Diffuse large B cell lymphoma

ABSTRACT

Tumour cells develop by accumulating changes in the genome that result in changes of main cellular processes. Aberrations of basic processes such as replication and chromatin reassembly are particularly important for genomic (in)stability. The aim of this study was to analyse the expression of genes whose products are crucial for the regulation of replication and chromatin reassembly during lymphomagenesis (*DNMT1, PCNA, MCM2, CDT1, EZH2, GMNN, EP300*). Non-tumour B cells were used as a control, and follicular lymphoma (FL) and the two most common groups of diffuse large B cell lymphoma (DLBCL) samples were used as a model for tumour progression. The results showed that there are significant changes in the expression of the analysed genes in lymphomagenesis, but also that these changes do not display linearity when assessed in relation to the degree of tumour aggression. Additionally, an integrated bioinformatics analysis of the difference in the expression of selected genes between tumour and non-tumour samples, and between tumour samples (FL vs. DLBCL) in five GEO datasets, did not show a consistent pattern of difference among the datasets.

1. Introduction

During normal immune response, following exposure to antigens, naive B-cells migrate to the centre of the primary lymphatic follicles in the secondary lymphoid tissue and differentiate into centroblasts. Centroblasts proliferate, fill the follicular dendritic cell network and form a germinal centre. In the germinal centre, centroblasts with high affinity to antigen mature to centrocytes through somatic hypermutation and apoptosis. Cell maturity is achieved by proliferation, and therefore depends on the precise regulation and function of the genes that control cell division and replication. The accumulation of the alterations in the process of replication and the subsequent chromatin reassembly contributes to genomic instability and might contribute to the development of malignant tumours.

The most common tumours that arise from germinal centre B-cells are follicular lymphoma (FL) and diffuse large B-cell lymphoma (DLBCL), which together make up 60% of all lymphomas. FL is an indolent lymphoma whose clinical course is characterized by slow growth. Rarely it can transform into more aggressive type of lymphoma such as DLBCL. DLBCL is an aggressive lymphoma characterized by rapid growth. Most common subtype of DLBCL is the “not otherwise specified” (NOS) subtype which is divided based on gene expression profiling into two biologically different subgroups that also show different clinical course of the disease. The “germinal centre B-cell like” subgroup (GCB) has germinal centre lymphocytes as proposed cell of origin and is prognostically more favourable, while the “activated B-cell like” subgroup (ABC or non-GCB) most likely arises from activated B-lymphocytes and is prognostically less favourable [1]. Genes that might contribute to the onset and behaviour of these lymphomas are the ones involved in replication and chromatin reassembly, which have so far

* Corresponding author.
E-mail addresses: katarina.horvat@gmail.com (K.H. Pavlov), vanjatadic@me.com (V. Tadić), pamela.basic-palkovic@skole.hr (P.B. Palković), biljanasasi@gmail.com (B. Sasi), nivesmagdic@gmail.com (N. Magdić), matea@inantro.hr (M.Z. Petranović), marija.klasic@biol.pmf.hr (M. Klasić), suzanaparlov@gmail.com (S. Hančić), paula.grskovic@biol.pmf.hr (P. Gršković), mmatulic@biol.pmf.hr (M. Matulić), gasparovslavko@gmail.com (S. Gašparov), mara.dominis@gmail.com (M. Dominis), petra.korac@biol.pmf.hr (P. Korać).

https://doi.org/10.1016/j.prp.2022.154170
Received 5 May 2022; Received in revised form 7 October 2022; Accepted 11 October 2022
Available online 13 October 2022
0344-0338/© 2022 Elsevier GmbH. All rights reserved.

been analysed in other malignant tumours.

For example, CDT1, which marks the starting point of replication and helps regulate DNA replication, has so far been analysed as a potential predictor of biological behaviour in hepatocellular carcinoma. Increased *CDT1* expression was found to be associated with a higher tumour grade and tumour-node-metastasis stage. High *CDT1* expression correlated significantly with decreased overall survival and was identified as an independent predictor of poor overall survival [2–4]. GMNN, an inhibitor of CDT1, is involved in regulation that ensures only one replication of each part of the genome happens per cell cycle [5]. Changes in *GMNN* expression were found in some high-grade lymphomas, one of which was also DLBCL NOS, and an increase in *GMNN* expression was observed in other more aggressive neoplasms [6]. One of the minichromosome maintenance proteins that forms the hexameric complex around DNA is MCM2. When active in normal cells, it enables the start of replication [7]. Downregulation of *MCM2* gene was found in both high- and low-grade lymphomas [8]. Moreover, various studies have shown high prevalence of mutations in the *EP300* gene, encoding the p300 protein that regulates chromatin remodelling through histone acetylation, in both FL and in DLBCL patients [9,10]. PCNA is involved in the process of replication regulation and has been studied as a general biomarker for tumour development. It forms a ring that binds DNA polymerase to DNA molecule, and enables the elongation of a new DNA strand [11]. Changes in the *PCNA* gene expression have been studied as a potential prognostic factor in patients with chronic lymphocytic leukaemia (CLL), as well as a predictor of response to Fludarabine-based chemotherapy in that disease. The *PCNA* gene has also been studied as a potential target for cancer therapy, mainly for malignancies with a high proliferation index [11–17]. Mutation of *EZH2* has been shown to result in the overexpression of the gene in FL and DLBCL NOS. It has been shown that EZH2 contributes to tumorigenicity in breast cancer, non-small cell lung cancer, and also affects the immune microenvironment of the tumour. In the hormone-refractory metastatic prostate cancer, the *EZH2* gene is overexpressed. Its higher expression in localised prostate cancer has been shown to contribute to a poorer prognosis [18–21]. In recent years, several EZH2-specific inhibitors for the treatment of multiple malignancies, including B cell lymphomas, have entered clinical trials [22]. Deregulation of DNMT1, the enzyme primarily responsible for the accurate preservation of DNA methylation patterns after cell division, has been associated with the pathogenesis of various types of malignancies [23]. The upregulation of *DNMT1* has a role in the development of prostate cancer. *DNMT1* polymorphisms are associated with an increased risk of developing gastric cancer, breast cancer, and cervical cancer, while DNMT1 inhibitors are approved anti-cancer drugs used in the treatment of prostate cancer [24–28].

Taken together, existing data suggest that changes in the expression of genes responsible for replication and chromatin assembly contribute to malignant cell transformation, but data on their role in lymphomagenesis are still scarce. Therefore, the aim of this study was to examine changes in the expression of *GMNN, PCNA, EZH2, DNMT1, MCM2, CDT1* and *EP300* genes in non-tumour germinal centre B lymphocytes, FL cells and both subtypes of DLBCL NOS cells in order to assess their contribution to the lymphoma progression. The findings of our laboratory analyses were additionally validated by assessment of data on the RNA expression of FL, DLBCL and different types of healthy samples from a publicly available free NCBI Gene Expression Omnibus (GEO) database. It was expected that changes of *GMNN, PCNA, EZH2, DNMT1, MCM2, CDT1* and *EP300* expression will be shown in both indolent and aggressive lymphoma when compared with their non-tumour cell of origin, but with the higher intensity in more aggressive lymphoma type, e. g. that expression changes will display linearity when assessed in relation to the degree of tumour aggression.

2. Materials and methods

For the laboratory analyses of non-tumour cells, FL cells and DLBCL cells, patients' tumour samples were collected. Fresh frozen tonsils and formalin fixed, paraffin embedded (FFPE) corresponding tissues (n = 10), fresh frozen and FFPE lymph nodes from patients with FL (n = 10), and FFPE lymph nodes from patients diagnosed with DLBCL NOS (10 DLBCL GCB cases and 10 DLBCL ABC cases) were included. Tonsils were collected from patients who underwent tonsillectomy due to recurring tonsillitis. Among the 30 patients diagnosed with lymphoma, men and women were equally represented with 15 men and 15 women. The age range was between 28 and 83 years of age.

All tissue samples were revised by three experienced hematopathologists (MD, SG, KHP), and the diagnoses were made using criteria of the World Health Organization (WHO) classification [1]. The study was approved by the Ethics committee of University Hospital Merkur and School of Medicine, University of Zagreb.

Tissue sections from tonsils and lymph nodes from patients with FL (4 μm tick) were mounted on membrane slides for laser capture microdissection (MembraneSlide 1.0 PEN, Carl Zeiss Microscopy), while tissue sections from FFPE lymph nodes from patients with DLBCL were mounted on standard adhesive histological slides. All slides were dried at 56 °C overnight and stained with hemalaun and eosin (HE) the next day. In brief, slides were deparaffinized using xylene substitute and rehydrated through ethanol series (absolute, 95%, and 70% ethanol, 2 min each). Afterwards, slides were washed with distilled water, treated with hematoxylin according to Harris for 5 min, rinsed with tap water for 2 min, and treated with eosin Y aqueous solution for 2 min. Dehydration with increasing ethanol series (70%, 95% and absolute ethanol, each 2 min) followed and slides were at the end air-dried and stored at 4 °C until the microdissection.

HE stained sections on membrane slides were analysed under the microscope. Germinal centres of tonsils and tumour areas of FL were selected with PALM RoboSoftware according to the manufacturer's instructions (Carl Zeiss Microscopy, Germany). Non-contact laser capture microdissection was then performed and selected tissue parts were collected in adhesive caps of collection tubes (AdhesiveCap opaque 500 μl, Carl Zeiss Microscopy, Germany). Microdissected samples were stored at − 80 °C until the RNA isolation. Whole section of DLBCL samples were scraped from histological slides and stored at − 80 °C until the RNA isolation.

RNA isolation of all samples was done according to the manufacturer's instructions of QIAGEN RNeasy Micro kit (Qiagen, Germany). After the RNA isolation, reverse transcription was done using PrimeScript Rtase (TaKaRa, Japan). Relative quantification of *DNMT1, PCNA, MCM2, CDT1, EZH2, GMNN* and *EP300* genes was analysed using *TBP1* gene as endogenous control. Following assays were used: Hs00945875_m1, Hs00696862_m1, Hs01091564_m1, Hs00368864_m1, Hs00544833_m1, Hs04276835_m1, Hs00914223_m1 and Hs00427620_m1 and all experiments were performed on 7500 Fast Real-Time PCR System (Applied Biosystems™, Foster City, CA, USA). Differences in gene expression between the germinal centre B-lymphocytes, FL cells and DLBCL cells were calculated using the ΔΔCt method.

In order to divide DLCBL into two subgroups, an immunohistochemical algorithm according to Hanc C et al. was used. [29] In brief: FFPE sections were deparaffinized using xylene substitution, rehydrated during heat-induced epitope retrieval (HIER) in Tris-EDTA buffer (pH 9.0), sections were then incubated with primary antibodies CD10 (IR648, Dako/Agilent), BCL6 (IR625, Dako/Agilent) and MUM1 (IR644, Dako/Agilent) for 20 min. Antigen visualization was done with EnVision Systems (K8002, Dako/Agilent), based on dextran polymer technology according to the manufacturer's recommendation. Tissues were mounted with The Tissue-Tek Glas Mounting Media (Sakura) and analysed using an Olympus BX51 microscope.

For assessment of protein presence in germinal centre cells or tumour cells, immunohistochemical staining was performed on FFPE sections with primary antibodies anti-EZH2 (SAB5500102, Sigma), anti-MCM2 (3619, Cell Signaling), anti-p300 (sc-585, Santa Cruz), anti-PCNA (2586, Cell Signaling), anti-CDT1 (sc-365305, Santa Cruz), anti-

DNMT1 (sc-271729, Santa Cruz) and anti-GMNN (sc-13015, Santa Cruz). Antigen visualization was also done with EnVision Systems (K8002, Dako/Agilent) according to the manufacturer's recommendation.

To compare differences between gene expression and protein presence in germinal centre B-lymphocytes and FL and/or DLBCL cells, Mann-Whitney U test was used. To assess the association between specific gene expression and protein presence, Spearman's rank-order correlation coefficient was used. Significance level was set at $p < 0.05$.

Additionally, integrated bioinformatics analyses were performed to validate our results on the analysed gene expression in the normal (non-tumour) B-cells, FL cells and DLBCL cells. Validation of the expression profiles of the seven investigated genes in tumour and non-tumour samples was performed using data from the Gene Expression Omnibus database (GEO, https://www.ncbi.nlm.nih.gov/geo/). A GEO browser was first searched for studies that measured gene expression in all samples of interest (non-tumour B-cells, FL cells and DLBCL cells) in the same dataset. After duplicate surveys were excluded (in which the samples partially or completely overlapped), the following datasets were selected for further analyses: GSE60, GSE12453, GSE12195, GSE32018 and GSE9327. Microarrays of these five datasets were performed on four different platforms: dataset GSE60 on spotted DNA/cDNA microarrays (non-commercial), datasets GSE12453 and GSE12195 on the Affymetrix U133 Plus 2.0 microarray, dataset GSE32018 on the Agilent-014850 Whole Human Genome Microarray 4x44K G4112F, and dataset GSE9327 on the CNIO Human Oncochip 1.0, 1.2 and 2.0. The number and types of non-tumour and tumour cells are shown in Tables 2–8.

After integrated bioinformatics GEO2R analysing, the results of the comparison of expression levels of the seven investigated genes between non-tumour B-cells and FL patients, non-tumour B-cells and DLBCL patients, and FL and DLBCL patients, were extracted from the datasets (Supplementary Tables 1–5). Because non-tumour B-cells types differed in five GEO datasets, both types of tumour samples were first compared separately with individual control cell types, and then with all healthy cell types together. Seven investigated differentially expressed genes (DEGs) with $|\log FC| > 2$ (FC = fold change) and p-value < 0.05 between non-tumour B-cells and tumour samples, and between two tumour samples (FL vs. DLBCL) were identified. DEGs with $\log FC < 0$ and significant p-values after Bonferroni`s correction for multiple testing were considered as down-regulated genes (highlighted blue in Supplementary Tables 1–5), while DEGs with $\log FC > 0$ were considered as up-regulated genes (highlighted red in Supplementary Tables 1–5).

We further constructed a protein-protein interaction (PPI) network using the Search Tool for the Retrieval of Interacting Genes (STRING, https://string-db.org/).

3. Results

3.1. Gene and protein expression in DLBCL subtypes

The level of gene expression and the average percentage of cells expressing the analysed proteins in the DLBCL GCB group and in the DLBCL ABC group were first determined, to evaluate whether the two should further be analysed as lower and higher-grade lymphoma or as unique DLBCL group.

In both DLBCL groups, GCB and ABC, the DNMT1 gene had the highest expression and the CDT1 and PCNA genes had the lowest expression. The expression levels of MCM2, P300, GMNN and EZH2 genes were also similar in both DLBCL groups.

Additionally, the average percentage of tumour cells in which the analysed proteins were expressed showed no significant difference between DLBCL GCB and ABC groups (Table 1).

As no significant difference in the level of gene expression and percentage of tumour cells that express analysed proteins was found between DLBCL GCB and ABC groups they were combined in unique DLBCL-NOS group for further analyses.

3.2. Gene and protein expression difference in germinative centre B-cells and lymphoma cells

In order to asses changes in selected gene/protein expression as possible contributors to low grade lymphoma development, samples of germinal centre B-cells were compared with FL cells. In order to evaluate those changes as contributors of high-grade lymphoma development, samples of germinal centre B-cells were compared with DLBCL NOS.

In germinal centre B-lymphocytes, the *DNMT1* and *MCM2* genes had the highest expression, the *CDT1* and *GMNN* genes had the lowest expression, and the expression of the remaining three analysed genes was between these two extremes. In the FL group, the *DNMT1* gene had the highest expression, the *GMNN*, *PCNA* and *CDT1* genes had the lowest expression, while the expression of the *EP300*, *MCM2* and *EZH2* genes was between these two extremes.

FL cells showed, in comparison to germinal centre B-cells, a significantly lower expression of *GMNN* ($p < 0.05$) and *MCM2* ($p < 0.01$), while DLBCL NOS cells showed a significantly lower expression of only *GMNN* ($p < 0.01$), but a significant upregulation of *CDT1* ($p < 0.05$) and *EP300* genes ($p < 0.01$) when compared with germinal centre B-cells (Fig. 1).

The average percentage of non-tumour B-cells in which the analysed proteins were expressed was (sorted in ascending order): CDT1 in 7.10% of cells, GMNN in 36.75% of cells, DNMT1 in 60.75%, MCM2 in 86.25%, EZH2 in 92.00%, PCNA in 94.00%, and p300 in 95.75% of cells.

A comparison of protein expression between non-tumour B-cells and FL cells showed a significantly lower percentage of cells expressing DNMT1 ($p < 0.05$), and GMNN, EZH2, MCM2, PCNA and CDT1 ($p < 0.01$) in FL than in non-tumour B-cells (Fig. 2a). Comparison of protein expression between DLBCL NOS and non-tumour B-cells showed lower expression of PCNA ($p < 0.01$) in DLBCL NOS samples (Fig. 2b).

3.3. Gene and protein expression difference between FL cells and DLBCL cells

Expression of *DNMT1, PCNA, MCM2, CDT1* and *EP300* showed a statistically significant higher expression in DLBCL NOS cells in comparison to FL cells, while the expression of GMNN was higher in FL cells ($p < 0.05$) (Fig. 3).

Average percentage of FL cells that expressed analysed proteins was as follows (sorted in ascending order): CDT1 was expressed in 0.80% of FL cells, DNMT1 in 2.00% of FL cells, GMNN in 4.80% of FL cells, MCM2 in 21.50% of FL cells, EZH2 in 41.50% of FL cells, PCNA in 65.50%of FL cells, and p300 in 93.50% of FL cells. The analysed proteins were present in DLBCL NOS cells as follows (sorted in ascending order): CDT1 11.89%, GMNN 43.29%, DNMT1 55.92%, MCM2 86.05%, EZH2 86.18%, PCNA 86.97%, and p300 95.13%. Comparison of protein presence in tumour tissues revealed higher percentage of cells

Table 1
The average percentage of DLBCL GCB and ABC tumour cells in which the analysed proteins were expressed.

DLBCL subtype	DNMT1	PCNA	MCM2	CDT1	EZH2	GMNN	P300
DLBCL GCB	50.23%	87.5%	84.09%	12.36%	83.41%	42.05%	95%
DLBCL ABC	63.75%	86.88%	88.75%	11.25%	90%	45%	95.31%

DLBCL – diffuse large B cell lymphoma, GCB – germinal centre B-cell like, ABC – activated B-cell like

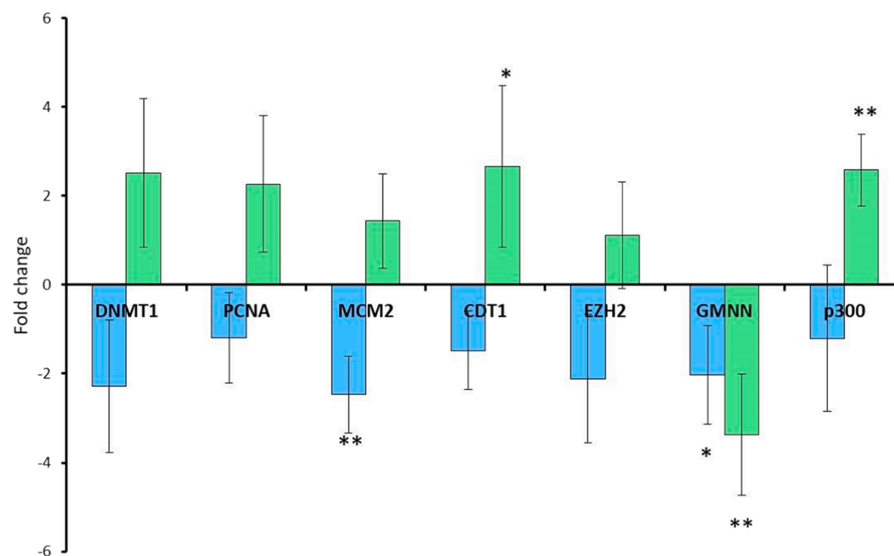


Fig. 1. Relative expression of the seven genes of interest in follicular lymphoma (FL) (blue) and diffuse large B-cell lymphoma not otherwise specified (DLBCL NOS) (green) compared to the non-tumour B cells. *MCM2* and *GMNN* gene expressions are significantly lower in FL cells than in non-tumour B-cells. *GMNN* expression is lower, and *CDT1* and *EP300* expressions are significantly higher in DLBCL NOS than in non-tumour B-cells. * $p < 0.05$; * * $p < 0.01$.

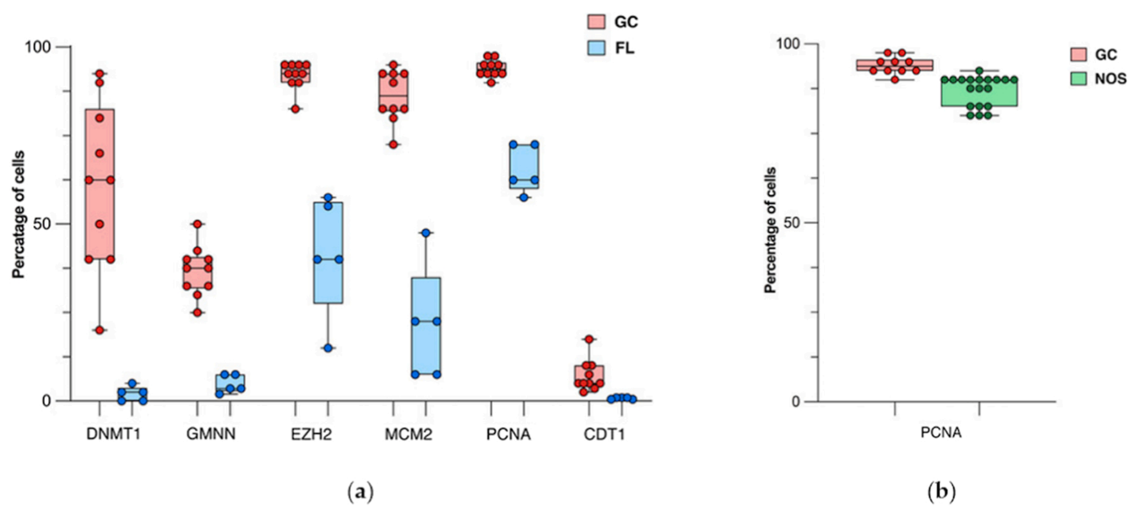


Fig. 2. a) Lower percentage of cells expressing DNMT1, GMNN, EZH2, MCM2, PCNA and CDT1 was found in follicular lymphoma (FL, blue) than in non-tumor B-cells from germinal centers (GC, red). b) Lower expression of PCNA was found in diffuse large B-cell lymphoma (DLBCL, green) than in non-tumor B-cells from germinal centers.

expressing DNMT1, PCNA, MCM2, CDT1, EZH2 and GMNN ($p < 0.01$) in DLBCL NOS than in FL (Fig. 4). The presence of protein p300 did not significantly differ between the two compared tumour cell types.

3.4. Validation of the results

3.4.1. Comparison of the expression of the investigated genes between DLBCL and FL samples

Comparison of gene expression between FL and DLBCL samples in five GEO datasets showed significant differences only in the GSE12195 dataset. Genes *DNMT1* (adjp = 2.12e-113, log2FC = −7.673), *CDT1*, *MCM2* (ajdp = 3.99e-104, log2FC = −7.332), *EP300*, *EZH2* (adjp = 5.92e-96, log2FC = −7.026), *PCNA* (adjp = 2.39e-106, log2FC = −8.261) and *GMNN* (adjp = 6.46e-94, log2FC = −6.979) (Tables 2–8, respectively) were significantly down-regulated in DLBCL compared to FL samples, while no difference in the expression was found in the remaining four GEO datasets. In the case of the *CDT1* and *EP300* genes, the microarray measured multiple probe sets representing the same gene, so in Tables 3 and 5 we just wrote that these genes were down-regulated in DLBCL when compared to FL samples (the individual values of adjusted p-value and log2FC multiple probes are shown in Supplementary Table 3).

3.4.2. Comparison of the expression of the studied genes between non tumour cells and FL

In the GSE60 dataset, a control sample consisted of activated B-lymphocytes, resting B-lymphocytes, tonsil germinal center (GC) B-cells, and resting/activated T-lymphocytes. The only significant difference in this dataset was found between FL samples and resting/activated T-lymphocytes in the *MCM2* gene (ajdp = 4.18e-3, log2FC = 2.161), which was upregulated in FL samples (Table 4).

The control sample of the GSE12195 dataset consisted of GC cells, memory and naive cells. All seven studied genes were significantly upregulated in FL samples compared to GC cells and to total control cells of this dataset, but not compared to memory and naive cells (Supplementary Table 3).

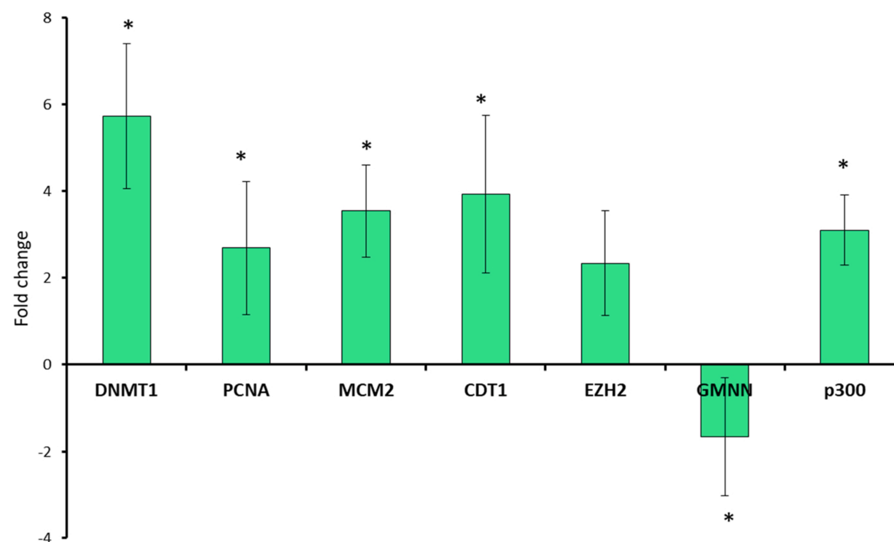


Fig. 3. Relative expression of the seven genes of interest in diffuse large B-cell lymphoma not otherwise specified (DLBCL NOS) compared to follicular lymphoma (FL). *DNMT1, PCNA, MCM2, CDT1* and *EP300* genes have significantly higher expression in DLBCL NOS cells than in FL cells, while *GMNN* is significantly less expressed in DLBCL compared to FL. * $p < 0.05$.

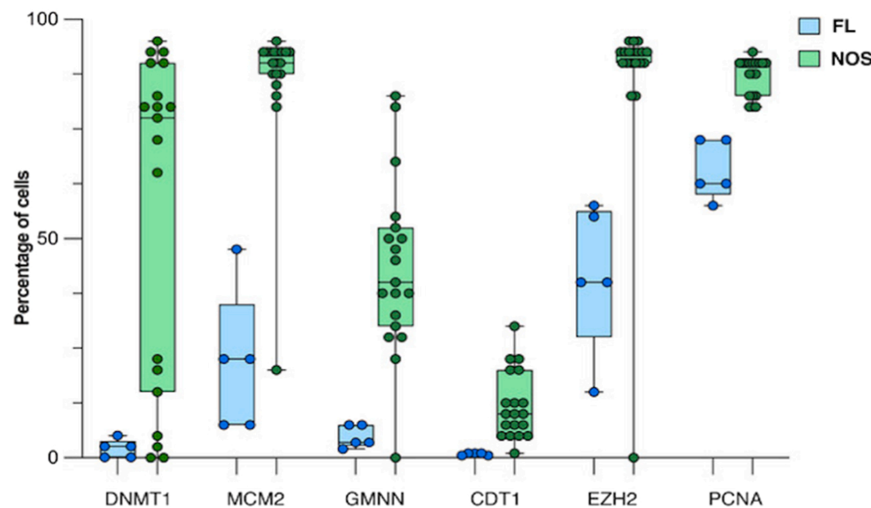


Fig. 4. Comparison of DNMT1, PCNA, MCM2, CDT1, EZH2 and GMNN protein presence in diffuse large B-cell lymphoma not otherwise specified (DLBCL NOS) and follicular lymphoma (FL) tumour samples showed significantly different percentage of two types of tumour cells expressing those proteins ($p < 0.01$).

A control sample of the GSE12453 dataset consisted of germinal center (GC) B-cells, memory, naive and plasma cells, in the GSE32018 dataset of lymph nodes and reactive tonsils, while one in the GSE9327 consisted of reactive lymph nodes and spleen control. In those datasets (Supplementary Table 2, 4 and 5), no difference in expression between FL and different types of healthy control cells was found.

3.4.3. *Comparison of the expression of the studied genes between non-tumour cells and DLBCL*

In the GSE60 dataset the only significant difference was found between DLBCL samples and resting/activated T-lymphocytes in the *MCM2* gene, which was upregulated in DLBCL samples (Table 4). Again, there were multiple microarray probes representing the *MCM2* gene, so in Table 4 we only wrote that the *MCM2* gene was upregulated, while individual probe values can be found in Supplementary Table 1.

In the GSE12453 dataset the *CDT1* gene was upregulated in DLBCL samples in comparison with memory cells (adjp = 1.54e-5, log2FC = 2.576), naive cells (adjp = 3.11e-6, log2FC = 2.545) and plasma cells (adjp = 5.22e-5, log2FC = 2.183) (Table 3). In addition to the *CDT1* gene, when compared to naive cells, DLBCL samples also had significantly upregulated genes *EZH2* (adjp = 6.22e-6, log2FC = 2.229) (Table 6), *GMNN* (adjp = 5.73e-6, log2FC = 2.367) (Table 8) and *PCNA* (adjp = 8.76e-5, log2FC = 2.265) (Table 7).

All seven studied genes were significantly downregulated in DLBCL compared to memory cells, to naive cells and to the total control sample in the GSE12195 dataset (Supplementary Table 3).

In datasets GSE32018 and GSE9327 (Tables 2–8) no difference was found between DLBCL and healthy samples in the expression of the studied genes.

3.5. *Construction of protein-protein interactions (PPI) network (original and extended)*

In order to try to understand the additional molecular mechanisms that could be associated with FL and DLBCL progression, a PPI network was constructed using the STRING database. The initial PPI network contained all seven analysed proteins and PCNA and MCM2 were at the core of the network, each linked to the remaining six proteins (Fig. 5).

Table 2
The expression of DNMT1 gene between lymphomas and different types of control samples in five GEO datasets.

DNMT1						
GEO datasets		GSE60	GSE12453	GSE12195	GSE32018	GSE9327
Sample type	DLBCL	4	11	44	22	36
	FL	9	5	38	23	33
	control cells	B-lymphocytes (5), CD4-positive T-lymphocytes (5), germinal center B-lymphocyte (1), memory B-lymphocyte (1), umbilical cord blood B-lymphocyte (1), umbilical cord blood T-lymphocyte (1); total 14 cells	CD77 + and CD77 −GC B cells and plasma cells isolated from tonsils, and naive and memory B cells isolated from peripheral blood; 5 donors each, total 25 cells	germinal center B-lymphocytes (10), memory B-lymphocytes (5), naive pregerminal center B-lymphocytes (5); total 20 cells	6 reactive tonsils, 7 lymph nodes; total 13 cells	8 reactive lymph nodes, 5 splenic controls; total 13 cells
gene expression difference (log2 fold change, adj p-value)	DLBCL vs. FL	no significant difference	no significant difference	downregulated (p = 2.12e-113, log2 = −7.673)	no significant difference	no significant difference
	DLBCL vs. Activated B-lymphocytes	no significant difference				
	DLBCL vs. Resting B-lymphocytes	no significant difference				
	DLBCL vs. GC B-lymphocytes	no significant difference	no significant difference	no significant difference		
	DLBCL vs. Activated/resting T-lymphocytes	no significant difference				
	DLBCL vs. Lymph nodes				no significant difference	
	DLBCL vs. Memory cells		no significant difference	downregulated (p = 8.62e-92, log2 = −7.533)		
	DLBCL vs. Naive cells		no significant difference	downregulated (p = 2.41e-101, log2 = −7.157)		
	DLBCL vs. Plasma cells		no significant difference			
	DLBCL vs. Reactive lymph nodes					no significant difference
	DLBCL vs. Reactive tonsils				no significant difference	
	DLBCL vs. Spleen control					no significant difference
	DLBCL vs. Total control	no significant difference	no significant difference	downregulated (p = 3.95e-13, log2 = −3.716)		no significant difference
	FL vs. Activated B-lymphocytes	no significant difference				
	FL vs. Resting B-lymphocytes	no significant difference				
	FL vs. GC B-lymphocytes	no significant difference	no significant difference	upregulated (p = 5.70e-39, log2 =7.586)		
	FL vs. Activated/resting T-lymphocytes	no significant difference				
	FL vs. Lymph nodes				no significant difference	
	FL vs. Memory cells		no significant difference	no significant difference		
	FL vs. Naive cells		no significant difference	no significant difference		
	FL vs. Plasma cells		no significant difference			
	FL vs. Reactive lymph nodes					no significant difference
	FL vs. Reactive tonsils				no significant difference	
	FL vs. Spleen control					no significant difference
	FL vs. Total control	no significant difference	no significant difference	upregulated (p = 2.03e-8, log2 =3.957)		no significant difference
Platform		Technology type: spotted DNA/cDNA, non-commercial	Affymetrix U133 Plus 2.0 microarrays	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F	CNIO Human Oncochip 1.0, 1.2 and 2.0

adjp = after Bonferroni adjustement for multiple testing; logFC = Log2-fold change

Table 3
The expression of *CDT1* gene between lymphomas and different types of control samples in five GEO datasets.

CDT1						
GEO datasets		GSE60	GSE12453	GSE12195	GSE32018	GSE9327
Sample type	DLBCL	4	11	44	22	36
	FL	9	5	38	23	33
	control cells	B-lymphocytes (5), CD4-positive T-lymphocytes (5), germinal center B-lymphocyte (1), memory B-lymphocyte (1), umbilical cord blood B-lymphocyte (1), umbilical cord blood T-lymphocyte (1); total 14 cells	CD77 + and CD77 −GC B cells and plasma cells isolated from tonsils, and naive and memory B cells isolated from peripheral blood; 5 donors each, total 25 cells	germinal center B-lymphocytes (10), memory B-lymphocytes (5), naive pregerminal center B-lymphocytes (5); total 20 cells	6 reactive tonsils, 7 lymph nodes; total 13 cells	8 reactive lymph nodes, 5 splenic controls; total 13 cells
gene expression difference (log2 fold change, adj p-value)	DLBCL vs. FL	no significant difference	no significant difference	downregulated	no significant difference	no significant difference
	DLBCL vs. Activated B-lymphocytes	no significant difference				
	DLBCL vs. Resting B-lymphocytes	no significant difference				
	DLBCL vs. GC B-lymphocytes	no significant difference	no significant difference	no significant difference		
	DLBCL vs. Activated/ resting T-lymphocytes	no significant difference				
	DLBCL vs. Lymph nodes				no significant difference	
	DLBCL vs. Memory cells		upregulated (p = 1.54e-5, log2 = 2.576)	downregulated		
	DLBCL vs. Naive cells		upregulated (p = 3.11e-6, log2 = 2.545)	downregulated		
	DLBCL vs. Plasma cells		upregulated (p = 5.22e-5, log2 = 2.183)			
	DLBCL vs. Reactive lymph nodes					no significant difference
	DLBCL vs. Reactive tonsils				no significant difference	
	DLBCL vs. Spleen control					no significant difference
	DLBCL vs. Total control	no significant difference	no significant difference	downregulated	no significant difference	no significant difference
	FL vs. Activated B-lymphocytes	no significant difference				
	FL vs. Resting B-lymphocytes	no significant difference				
	FL vs. GC B-lymphocytes	no significant difference	no significant difference	upregulated		
	FL vs. Activated/ resting T-lymphocytes	no significant difference				
	FL vs. Lymph nodes				no significant difference	
	FL vs. Memory cells		no significant difference	no significant difference		
	FL vs. Naive cells		no significant difference	no significant difference		
	FL vs. Plasma cells		no significant difference			
	FL vs. Reactive lymph nodes					no significant difference
	FL vs. Reactive tonsils				no significant difference	
	FL vs. Spleen control					no significant difference
	FL vs. Total control	no significant difference	no significant difference	upregulated (p = 2.03e-7, log2 =2.726)	no significant difference	no significant difference
Platform		Technology type: spotted DNA/ cDNA, non-commercial	Affymetrix U133 Plus 2.0 microarrays	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F	CNIO Human Oncochip 1.0, 1.2 and 2.0

adjp = after Bonferroni adjustement for multiple testing; logFC = Log2-fold change

Table 4
The expression of *MCM2* gene between lymphomas and different types of control samples in five GEO datasets.

MCM2						
GEO datasets		GSE60	GSE12453	GSE12195	GSE32018	GSE9327
Sample type	DLBCL	4	11	44	22	36
	FL	9	5	38	23	33
	control cells	B-lymphocytes (5), CD4-positive T-lymphocytes (5), germinal center B-lymphocyte (1), memory B-lymphocyte (1), umbilical cord blood B-lymphocyte (1), umbilical cord blood T-lymphocyte (1); total 14 cells	CD77 + and CD77 −GC B cells and plasma cells isolated from tonsils, and naive and memory B cells isolated from peripheral blood; 5 donors each, total 25 cells	germinal center B-lymphocytes (10), memory B-lymphocytes (5), naive pregerminal center B-lymphocytes (5); total 20 cells	6 reactive tonsils, 7 lymph nodes; total 13 cells	8 reactive lymph nodes, 5 splenic controls; total 13 cells
gene expression difference (log2 fold change, adj *p*-value)	DLBCL vs. FL	no significant difference	no significant difference	downregulated (p = 3.99e-104, log2 = −7.332)	no significant difference	no significant difference
	DLBCL vs. Activated B-lymphocytes	no significant difference				
	DLBCL vs. Resting B-lymphocytes	no significant difference				
	DLBCL vs. GC B-lymphocytes	no significant difference	no significant difference	no significant difference		
	DLBCL vs. Activated/ resting T-lymphocytes	upregulated				
	DLBCL vs. Lymph nodes				no significant difference	
	DLBCL vs. Memory cells		no significant difference	downregulated (p = 3.54e-85, log2 = −7.800)		
	DLBCL vs. Naive cells		no significant difference	downregulated (p = 1.73e-88, log2 = −7.476)		
	DLBCL vs. Plasma cells		no significant difference			
	DLBCL vs. Reactive lymph nodes					no significant difference
	DLBCL vs. Reactive tonsils				no significant difference	
	DLBCL vs. Spleen control					no significant difference
	DLBCL vs. Total control	no significant difference	no significant difference	downregulated (p = 4.10e-13, log2 = −3.858)	no significant difference	no significant difference
	FL vs. Activated B-lymphocytes	no significant difference				
	FL vs. Resting B-lymphocytes	no significant difference				
	FL vs. GC B-lymphocytes	no significant difference	downregulated (p = 6.06e-3, log2 = −1.226)	upregulated (p = 2.29e-35, log2 =7.254)		
	FL vs. Activated/ resting T-lymphocytes	upregulated (p = 4.18e-3, log2 =2.161)				
	FL vs. Lymph nodes				no significant difference	
	FL vs. Memory cells		no significant difference	no significant difference		
	FL vs. Naive cells		no significant difference	no significant difference		
	FL vs. Plasma cells		no significant difference			
	FL vs. Reactive lymph nodes					no significant difference
	FL vs. Reactive tonsils				no significant difference	
	FL vs. Spleen control					no significant difference
	FL vs. Total control	no significant difference	no significant difference	upregulated (p = 6.50e-7, log2 =3.474)	no significant difference	no significant difference
Platform		Technology type: spotted DNA/ cDNA, non-commercial	Affymetrix U133 Plus 2.0 microarrays	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F	CNIO Human Oncochip 1.0, 1.2 and 2.0

adjp = after Bonferroni adjustement for multiple testing; logFC = Log2-fold change

Table 5
The expression of *EP300* gene between lymphomas and different types of control samples in five GEO datasets.

EP300						
GEO datasets		GSE60	GSE12453	GSE12195	GSE32018	GSE9327
Sample type	DLBCL	4	11	44	22	36
	FL	9	5	38	23	33
	control cells	B-lymphocytes (5), CD4-positive T-lymphocytes (5), germinal center B-lymphocyte (1), memory B-lymphocyte (1), umbilical cord blood B-lymphocyte (1), umbilical cord blood T-lymphocyte (1); total 14 cells	CD77 + and CD77 −GC B cells and plasma cells isolated from tonsils, and naive and memory B cells isolated from peripheral blood; 5 donors each, total 25 cells	germinal center B-lymphocytes (10), memory B-lymphocytes (5), naive pregerminal center B-lymphocytes (5); total 20 cells	6 reactive tonsils, 7 lymph nodes; total 13 cells	8 reactive lymph nodes, 5 splenic controls; total 13 cells
gene expression difference (log2 fold change, adj p-value)	DLBCL vs. FL	no significant difference	no significant difference	downregulated	no significant difference	no significant difference
	DLBCL vs. Activated B-lymphocytes	no significant difference				
	DLBCL vs. Resting B-lymphocytes	no significant difference				
	DLBCL vs. GC B-lymphocytes	no significant difference	no significant difference	no significant difference		
	DLBCL vs. Activated/ resting T-lymphocytes	no significant difference				
	DLBCL vs. Lymph nodes				no significant difference	
	DLBCL vs. Memory cells		no significant difference	downregulated		
	DLBCL vs. Naive cells		no significant difference	downregulated		
	DLBCL vs. Plasma cells		no significant difference			
	DLBCL vs. Reactive lymph nodes					no significant difference
	DLBCL vs. Reactive tonsils				no significant difference	
	DLBCL vs. Spleen control					no significant difference
	DLBCL vs. Total control	no significant difference	no significant difference	downregulated	no significant difference	no significant difference
	FL vs. Activated B-lymphocytes	no significant difference				
	FL vs. Resting B-lymphocytes	no significant difference				
	FL vs. GC B-lymphocytes	no significant difference	no significant difference	upregulated		
	FL vs. Activated/ resting T-lymphocytes	no significant difference				
	FL vs. Lymph nodes				no significant difference	
	FL vs. Memory cells		no significant difference	no significant difference		
	FL vs. Naive cells		no significant difference	no significant difference		
	FL vs. Plasma cells		no significant difference			
	FL vs. Reactive lymph nodes					no significant difference
	FL vs. Reactive tonsils				no significant difference	
	FL vs. Spleen control					no significant difference
	FL vs. Total control	no significant difference	no significant difference	upregulated	no significant difference	no significant difference
Platform		Technology type: spotted DNA/ cDNA, non-commercial	Affymetrix U133 Plus 2.0 microarrays	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F	CNIO Human Oncochip 1.0, 1.2 and 2.0

adjp = after Bonferroni adjustement for multiple testing; logFC = Log2-fold change

Table 6
The expression of *EZH2* gene between lymphomas and different types of control samples in five GEO datasets.

EZH2						
GEO datasets		GSE60	GSE12453	GSE12195	GSE32018	GSE9327
Sample type	DLBCL	4	11	44	22	36
	FL	9	5	38	23	33
	control cells	B-lymphocytes (5), CD4-positive T-lymphocytes (5), germinal center B-lymphocyte (1), memory B-lymphocyte (1), umbilical cord blood B-lymphocyte (1), umbilical cord blood T-lymphocyte (1); total 14 cells	CD77 + and CD77 −GC B cells and plasma cells isolated from tonsils, and naive and memory B cells isolated from peripheral blood; 5 donors each, total 25 cells	germinal center B-lymphocytes (10), memory B-lymphocytes (5), naive pregerminal center B-lymphocytes (5); total 20 cells	6 reactive tonsils, 7 lymph nodes; total 13 cells	8 reactive lymph nodes, 5 splenic controls; total 13 cells
gene expression difference (log2 fold change, adj p-value)	DLBCL vs. FL	no significant difference	no significant difference	downregulated (p = 5.92e-96, log2 = −7.026)	no significant difference	no significant difference
	DLBCL vs. Activated B-lymphocytes	no significant difference				
	DLBCL vs. Resting B-lymphocytes	no significant difference				
	DLBCL vs. GC B-lymphocytes	no significant difference	no significant difference	no significant difference		
	DLBCL vs. Activated/ resting T-lymphocytes	no significant difference				
	DLBCL vs. Lymph nodes				no significant difference	
	DLBCL vs. Memory cells		no significant difference	downregulated (p = 1.06e-86, log2 = −6.291)		
	DLBCL vs. Naive cells		upregulated (p = 6.22e-6, log2 = 2.229)	downregulated (p = 9.71e-84, log2 = −5.635)		
	DLBCL vs. Plasma cells		no significant difference			
	DLBCL vs. Reactive lymph nodes					no significant difference
	DLBCL vs. Reactive tonsils				no significant difference	
	DLBCL vs. Spleen control					no significant difference
	DLBCL vs. Total control	no significant difference	no significant difference	downregulated (p = 5.71e-13, log2 = −3.822)	no significant difference	no significant difference
	FL vs. Activated B-lymphocytes	no significant difference				
	FL vs. Resting B-lymphocytes	no significant difference				
	FL vs. GC B-lymphocytes	no significant difference	downregulated (p = 1.81e-3, log2 = −1.432)	upregulated (p = 1.24e-31, log2 = 6.944)		
	FL vs. Activated/ resting T-lymphocytes	no significant difference				
	FL vs. Lymph nodes				no significant difference	
	FL vs. Memory cells		no significant difference	no significant difference		
	FL vs. Naive cells		no significant difference	no significant difference		
	FL vs. Plasma cells		no significant difference			
	FL vs. Reactive lymph nodes					no significant difference
	FL vs. Reactive tonsils				no significant difference	
	FL vs. Spleen control					no significant difference
	FL vs. Total control	no significant difference	no significant difference	upregulated (p = 3.99e-10, log2 =4.003)	no significant difference	no significant difference
Platform		Technology type: spotted DNA/ cDNA,non-commercial	Affymetrix U133 Plus 2.0 microarrays	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F	CNIO Human Oncochip 1.0, 1.2 and 2.0

adjp = after Bonferroni adjustement for multiple testing; logFC = Log2-fold change

Table 7
The expression of *PCNA* gene between lymphomas and different types of control samples in five GEO datasets.

PCNA						
GEO datasets		GSE60	GSE12453	GSE12195	GSE32018	GSE9327
Sample type	DLBCL	4	11	44	22	36
	FL	9	5	38	23	33
	control cells	B-lymphocytes (5), CD4-positive T-lymphocytes (5), germinal center B-lymphocyte (1), memory B-lymphocyte (1), umbilical cord blood B-lymphocyte (1), umbilical cord blood T-lymphocyte (1); total 14 cells	CD77 + and CD77 −GC B cells and plasma cells isolated from tonsils, and naive and memory B cells isolated from peripheral blood; 5 donors each, total 25 cells	germinal center B-lymphocytes (10), memory B-lymphocytes (5), naive pregerminal center B-lymphocytes (5); total 20 cells	6 reactive tonsils, 7 lymph nodes; total 13 cells	8 reactive lymph nodes, 5 splenic controls; total 13 cells
gene expression difference (log2 fold change, adj *p*-value)	DLBCL vs. FL	no significant difference	no significant difference	downregulated (p = 2.39e-106, log2 = −8.261)	no significant difference	no significant difference
	DLBCL vs. Activated B-lymphocytes	no significant difference				
	DLBCL vs. Resting B-lymphocytes	no significant difference				
	DLBCL vs. GC B-lymphocytes	no significant difference	no significant difference	no significant difference		
	DLBCL vs. Activated/ resting T-lymphocytes	no significant difference				
	DLBCL vs. Lymph nodes				no significant difference	
	DLBCL vs. Memory cells		no significant difference	downregulated (p = 4.85e-83, log2 = −7.840)		
	DLBCL vs. Naive cells		upregulated (p = 8.76e-5, log2 = 2.265)	downregulated (p = 1.02e-95, log2 = −7.464)		
	DLBCL vs. Plasma cells		no significant difference			
	DLBCL vs. Reactive lymph nodes					no significant difference
	DLBCL vs. Reactive tonsils				no significant difference	
	DLBCL vs. Spleen control					no significant difference
	DLBCL vs. Total control	no significant difference	no significant difference	downregulated (p = 5.71e-13, log2 = −3.822)	no significant difference	no significant difference
	FL vs. Activated B-lymphocytes	no significant difference				
	FL vs. Resting B-lymphocytes	no significant difference				
	FL vs. GC B-lymphocytes	no significant difference	downregulated (p = 4.40e-2, log2 = −1.045)	upregulated (p = 2.22e-36, log2 = 8.268)		
	FL vs. Activated/ resting T-lymphocytes	no significant difference				
	FL vs. Lymph nodes				no significant difference	
	FL vs. Memory cells		no significant difference	no significant difference		
	FL vs. Naive cells		no significant difference	no significant difference		
	FL vs. Plasma cells		no significant difference			
	FL vs. Reactive lymph nodes					no significant difference
	FL vs. Reactive tonsils				no significant difference	
	FL vs. Spleen control					no significant difference
	FL vs. Total control	no significant difference	no significant difference	upregulated (p = 5.49e-9, log2 =4.439)	no significant difference	no significant difference
Platform		Technology type: spotted DNA/ cDNA, non-commercial	Affymetrix U133 Plus 2.0 microarrays	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F	CNIO Human Oncochip 1.0, 1.2 and 2.0

adjp = after Bonferroni adjustement for multiple testing; logFC = Log2-fold change

Table 8
The expression of *GMNN* gene between lymphomas and different types of control samples in five GEO datasets.

GMNN						
GEO datasets		GSE60	GSE12453	GSE12195	GSE32018	GSE9327
Sample type	DLBCL	4	11	44	22	36
	FL	9	5	38	23	33
	control cells	B-lymphocytes (5), CD4-positive T-lymphocytes (5), germinal center B-lymphocyte (1), memory B-lymphocyte (1), umbilical cord blood B-lymphocyte (1), umbilical cord blood T-lymphocyte (1); total 14 cells	CD77 + and CD77 −GC B cells and plasma cells isolated from tonsils, and naive and memory B cells isolated from peripheral blood; 5 donors each, total 25 cells	germinal center B-lymphocytes (10), memory B-lymphocytes (5), naive pregerminal center B-lymphocytes (5); total 20 cells	6 reactive tonsils, 7 lymph nodes; total 13 cells	8 reactive lymph nodes, 5 splenic controls; total 13 cells
gene expression difference (log2 fold change, adj p-value)	DLBCL vs. FL	no significant difference	no significant difference	downregulated (p = 6.46e-94, log2 = −6.979)	no significant difference	no significant difference
	DLBCL vs. Activated B-lymphocytes	no significant difference				
	DLBCL vs. Resting B-lymphocytes	no significant difference				
	DLBCL vs. GC B-lymphocytes	no significant difference	no significant difference	no significant difference		
	DLBCL vs. Activated/ resting T-lymphocytes	no significant difference				
	DLBCL vs. Lymph nodes				no significant difference	
	DLBCL vs. Memory cells		no significant difference	downregulated (p = 3.72e-75, log2 = −6.842)		
	DLBCL vs. Naive cells		upregulated (p = 5.73e-6, log2 = 2.367)	downregulated (p = 2.66e-79, log2 = −6.280)		
	DLBCL vs. Plasma cells		no significant difference			
	DLBCL vs. Reactive lymph nodes					no significant difference
	DLBCL vs. Reactive tonsils				no significant difference	
	DLBCL vs. Spleen control					no significant difference
	DLBCL vs. Total control	no significant difference	no significant difference	downregulated (p = 6.05e-13, log2 = −3.292)	no significant difference	no significant difference
	FL vs. Activated B-lymphocytes	no significant difference				
	FL vs. Resting B-lymphocytes	no significant difference				
	FL vs. GC B-lymphocytes	no significant difference	downregulated (p = 5.69e-3, log2 = −1.338)	upregulated (p = 5.22e-31, log2 = 6.956)		
	FL vs. Activated/ resting T-lymphocytes	no significant difference				
	FL vs. Lymph nodes				no significant difference	
	FL vs. Memory cells		no significant difference	no significant difference		
	FL vs. Naive cells		no significant difference	no significant difference		
	FL vs. Plasma cells		no significant difference			
	FL vs. Reactive lymph nodes					no significant difference
	FL vs. Reactive tonsils				no significant difference	
	FL vs. Spleen control					no significant difference
	FL vs. Total control	no significant difference	no significant difference	upregulated (p = 2.60e-8, log2 = 3.687)	no significant difference	no significant difference
Platform		Technology type: spotted DNA/ cDNA,non-commercial	Affymetrix U133 Plus 2.0 microarrays	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F	CNIO Human Oncochip 1.0, 1.2 and 2.0

adjp = after Bonferroni adjustement for multiple testing; logFC = Log2-fold change

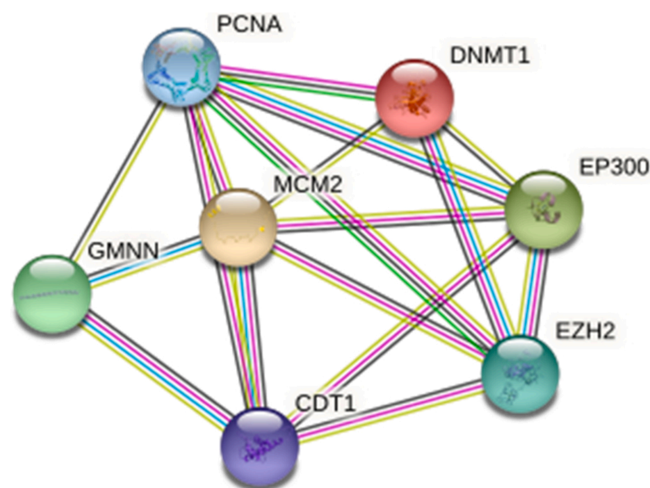


Fig. 5. Functional interactions between seven studied proteins. Figure was prepared using STRING Protein-Protein Interactions Network.

The average node degree was 4.86, average local clustering coefficient was 0.867, and the PPI enrichment p-value was 2.61e-08.

We then selected STRING`s option to build a larger PPI by adding additional proteins, in order to find out which other proteins may be involved in biological pathways and therefore relevant for lymphomagenesis. The new PPI network contained seven initial and five new proteins: SIRT1, MCM5, ORC6, CDC6 and RFC2. In this new PPI network, not only were PCNA and MCM2 linked to six proteins from the initial step, but also with four new proteins; MCM5, ORC6, CDC6 and RFC2. Furthermore, the protein CDT1, which was in the initial step linked with five proteins (all but DNMT1), was now linked not only to them but also to same four new proteins (MCM5, ORC6, CDC6 and RFC2). In addition, the newly added MCM5 protein was linked to nine proteins. The new PPI had 45 edges, average node degree was 7.5, average local clustering coefficient was 0.775, and the PPI enrichment p-value was 3.01e-12 (Fig. 6).

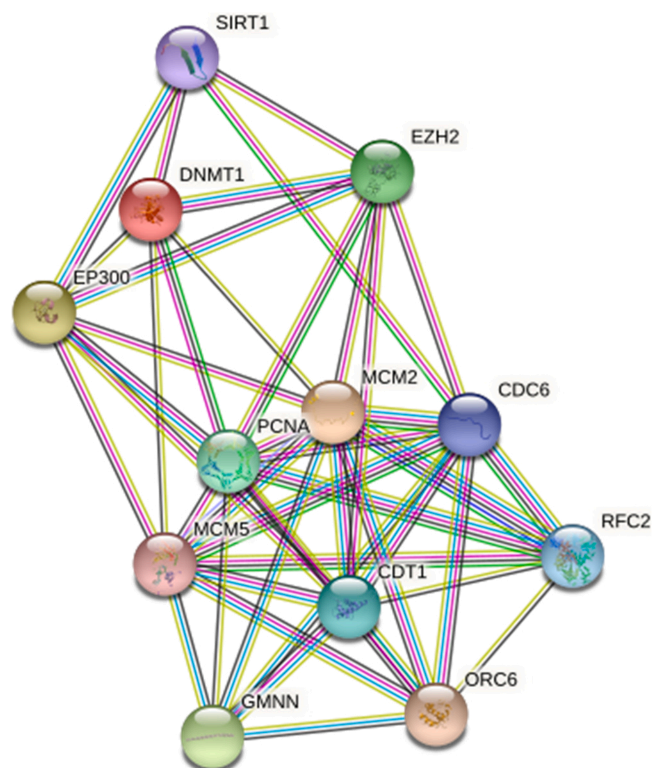


Fig. 6. Functional interactions between twelve proteins relevant to FL and DLBCL progression. In addition to seven initial proteins, five new proteins are MCM5 = Minichromosome Maintenance Complex Component 5; ORC6 = Origin Recognition Complex Subunit 6; CDC6 = Cell Division control protein 6 homolog; SIRT1 = NAD-dependent deacetylase sirtuin-1; and RFC2 = Replication factor C subunit 2. Figure was prepared using STRING Protein-Protein Interactions Network.

4. Discussion

In this study we aimed to evaluate changes in crucial genes that govern replication and chromatin assembly regulation in progression of lymphomagenesis. We used germinal centre B-lymphocytes, FL cells and DLBCL cells as a model. Non-tumour germinal centre B-cells were used as a control, an indolent lymphoma, FL, was used as an early step in lymphomagenesis, and DLBCL, as an aggressive lymphoma, represented lymphomagenesis progression. In this model, we expected to observe changes in basic cell processes that are becoming more pronounced with the level of tumour aggressiveness. Surprisingly, our results showed that there is no linearity in lymphoma progression.

So far, relative expression levels of *DNMT1* were assessed as a potential prognostic marker in oral squamous cell carcinoma and in non-small cell lung carcinoma. It was shown that in oral carcinoma *DNMT1* expression is the most discriminatory marker of poor outcome, while in lung carcinoma it can be regarded as an independent prognostic factor [26]. Furthermore, increased DNMT1 protein expression and overexpression of *DNMT1* gene were found in hepatocellular carcinoma tumour cells compared to normal hepatocytes [27]. Expression of DNMT1 protein in pancreatic ductal adenocarcinoma tissues was also observed, while there was no positive staining in the adjacent normal tissues [28]. We observed a decrease in the expression of both DNTM1 mRNA and protein in FL cells compared to GC lymphocytes, but the difference was significant only at the protein level. Our results can be explained by the different regulation of degradation of *DNMT1* mRNA and its protein product, as described by Leng et al., or by observation that only a certain number of transcribed mRNAs is translated into proteins [29,30] Our findings contradict the results by Agoston et al., who found that the longer half-life of DNMT1 protein is responsible for its dysregulation in breast cancer, rather than dysregulation on transcriptional level, suggesting that DNMT1 degradation is differently regulated in FL compared to breast cancer. [31] On the other hand, the expression of both DNMT1 mRNA and protein was significantly increased in DLBCL compared to FL (but not compared to GC B-cells), indicating an opposite role of DNMT1 in DLBCL compared to FL.

CDT1 overexpression in HCC was previously shown to promote hepatocellular carcinogenesis by causing genomic instability. Its expression was shown to be significantly increased with HCC grade [2]. Similarly to DNMT1, our results showed significant CDT1 protein expression decrease (but not significant mRNA expression decrease) in FL group compared to control group. CDT1 expression is low at both transcriptional and protein level in quiescent cells indicating expression regulation at transcriptional level, but its expression increases during G1 phase. [32] While its expression remains stable at mRNA level during cell cycle, the protein expression decreases upon entrance to S phase, suggesting that its expression in that phase is controlled by proteolysis. [33] Surprisingly, we observed the significant decrease in CDT1 expression in FL, an indolent lymphoma, but not in DLBCL, an aggressive lymphoma, suggesting that CDT1 degradation does not mark DLBCL development.

Furthermore, MCM2 protein and gene expression were decreased in FL group compared to control group, but both protein and gene expression were increased in DLBCL NOS group compared to FL group. This is consistent with studies that showed overexpression of MCM2 protein in DLBCL NOS with higher number of proliferating cells, as well as in tumours with poorer prognosis [6]. Additionally, while *EP300*

mutations have been identified as major pathogenetic mechanism in B-non-Hodgkin lymphomas (B-NHL) [7,8], little is known about the effect of the altered expression of non-mutated EP300 in DLBCL. Our results show *EP300* overexpression in DLBCL NOS compared both to control group as well as FL group, but no significant difference in the protein expression between DLBCL and the other two groups.

Changes of expression of *PCNA* have mostly been studied in carcinomas. Upregulation of *PCNA* has been found in gastric as well as breast carcinoma [34,35]. In chronic lymphocytic leukaemia (CLL) the *PCNA* gene expression has been studied as a marker of proliferation, as well as a potential predictive marker of response to chemotherapy. Its overexpression was found in CLL patients with poorer prognosis, and in patients that had poorer response to chemotherapy [14]. Surprisingly, in our study, DLBCL cells produced lower amounts of PCNA protein than non-tumour mature B-lymphocytes, despite higher mRNA expression, suggesting low translation effect or faster protein degradation. On the other hand, the expression of both mRNA and protein was higher in DLBCL NOS when compared to FL, which is consistent with expected higher proliferation rate in DLBCL NOS cells.

In DLBCL NOS compared to the FL group we also found elevated EZH2 expression, while its expression was lower in FL compared to the control group, but with no significant difference in mRNA expression between the groups, indicating different expression regulation at post-transcriptional and posttranslational levels or lower rate of protein degradation. Deregulation of EZH2 modulates pathogenesis of multiple lymphoid malignancies. It modulates oncogenesis by epigenetic repression of tumour suppressor genes, giving rise to potential malignant progression. However, our results suggest its opposite role in development of aggressive and indolent lymphoma.

GMNN has been studied in multiple tumours as marker of cell proliferation. Its absence has been found in CLL as well as in mantle cell lymphoma, neoplasms characterised as ones with low proliferation and mostly indolent behaviour. In contrast, GMNN has been expressed in lymphomas with high proliferation rate, such as DLBCL and Burkitt lymphoma. We found elevated protein amounts in the DLBCL NOS group when compared to the FL group, which supports findings of GMNN being connected to the level of proliferation of tumour cells [4]. Surprisingly, *GMNN* mRNA level was lower in DLBCL than in FL, indicating a longer half-life of its protein product compared to *GMNN* mRNA or different regulation of degradation between mRNA and protein in DLBCL, similar to DNMT1. [29] Also, lower expression of GMNN could have a direct effect on the expression of CDT1, even though GMNN was shown to have a dual role in regards to CDT1. [36,37] Additionally, we used integrated bioinformatics analysis to verify the results of our laboratory analyses of expression of seven genes involved in DLBCL and FL. Only one of five studied GEO datasets showed significant differences between DLBCL and FL samples; all seven studied genes were downregulated in DLBCL samples in the GSE12195 dataset. These results, as well as the absence of a difference in gene expression between the two lymphomas in the remaining four GEO datasets, are in contrast to our laboratory analysis.

The most similar comparison of gene expression levels between lymphomas and healthy cell samples between publicly available datasets and our samples was made for GSE60, GSE12453 and GSE12195 datasets, because these GEO datasets, just like our study, had GC B-cells as control cells (they also had other types of healthy control cells). While our laboratory analysis showed significantly lower expression of two genes (*GMNN* and *MCM2*) in FL than in GC B-cells, and a significant difference in the expression of three genes (*GMNN*, *CDT1* and *EP300*) between DLBCL NOS and GC B-cells, in two of three bioinformatics datasets no difference was found. However, in the GSE12195 dataset, all seven studied genes were significantly upregulated in FL compared to GC B-cells. In addition to GC B-cells, five datasets contained other types of healthy control cells, but a similar expression pattern of seven studied genes between DLBCL vs. control and FL vs. control was not detected in these other control cells either.

Perhaps the most surprising finding of the bioinformatics analysis used to validate the results of our laboratory analysis is that there are not only differences between our research and other studies, but that it is hard to draw firm conclusions from the results of bioinformatics studies about the expression of researched genes. The reason for this could be that a relatively small number of lymphoma and control samples were analysed, both in our laboratory analysis and in GEO datasets. In addition, these five datasets were analyzed using different platforms. The other reason might be population differences; in some populations FL is very rare, while in others it is common.

Moreover, the protein-protein interaction analysis showed that our network of seven analysed genes and their proteins had significantly more nodes than expected. That was not a surprise as it is known that they regulate replication and chromatin reassembly. The STRING software suggested additional five proteins (MCM5 = Minichromosome Maintenance Complex Component 5, ORC6 = Origin Recognition Complex Subunit 6, CDC6 = Cell Division control protein 6 homolog, SIRT1 = NAD-dependent deacetylase sirtuin-1, and RFC2 = Replication factor C subunit 2), which increased the statistical significance of the new PPI network.

Taken together, our results display the non-linearity of lymphoma development and suggest that the lower expression of proteins involved in expression regulation and chromatin assembly in FL could be responsible for the slower progression of indolent lymphoma compared to aggressive lymphomas like DLBCL.

CRediT authorship contribution statement

Conceptualization, **P.K., M.D, M.M. and S.G**; methodology, **K.H.P., V.T., P.B.P., B.S., N.M., S.H., M.K., M.Z.P., M. M. and P. K**. software, **M.Z.P. and P.K**; validation, **M.Z.P., K.H.P., V.T., P.B.P., B.S, N.M., M. D., S.G., and P. K.**; formal analysis, **V.T., P.B.P., B.S., N.M., S.H., M.Z. P., M.K., P.G. and P. K.**; resources, **M.D., S.G., and P.K.**; writing—original draft preparation, **K.H.P., M.K., P.G., M.Z.P., and P. K.**; writing—review and editing, **K.H.P., V.T., P.B.P., B.S., N.M., S.H., M. K., P.G. M.Z.P., M.D., S.G., M.M. and P. K.**; supervision, **P.K.**

All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Institutional review board statement

The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Ethics Committee of Merkur University Hospital (protocol code 03/1–8793/1 and 25. 9. 2019.).

Declaration of Competing Interest

None.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.prp.2022.154170.

References

[1] S.H. Swerdlow, E. Campo, N.L. Harris, E.S. Jaffe, S.A. Pileri, H. Stein, J. Thiele, W. H.O. Classification, Classification of tumours of haematopoietic and lymphoid tissues. Revised, 4th ed..,, IARC Press, Lyon, 2017.

[2] D. Karavias, I. Maroulis, H. Papadaki, C. Gogos, S. Kakkos, D. Karavias, V. Bravou, V, Overexpression of CDT1 is a predictor of poor survival in patients with

hepatocellular carcinoma, J. Gastrointest. Surg. . J. Soc. Surg. Aliment. Trac. 20 (2016) 568–579, https://doi.org/10.1007/s11605-015-2960-7.

[3] C. Caillat, A. Perrakis, Cdt1 and geminin in DNA replication initiation, Subcell. Biochem. 62 (2012) 71–87, https://doi.org/10.1007/978-94-007-4572-8_5.

[4] E.C. Obermann, K.L. Eward, A. Dogan, E.A. Paul, M. Loddo, P. Munson, G. H. Williams, K. Stoeber, DNA replication licensing in peripheral B-cell lymphoma, J. Pathol. 205 (2005) 318–328, https://doi.org/10.1002/path.1695.

[5] Y. Ishimi, Regulation of MCM2-7 function, Genes Genet. Syst. 93 (2018) 125–133, https://doi.org/10.1266/ggs.18-00026.

[6] E.C. Obermann, P. Went, A. Zimpfer, A. Tzankov, P.J. Wild, R. Stoehr, S.A. Pileri, S. Dirnhofer, Expression of minichromosome maintenance protein 2 as a marker for proliferation and prognosis in diffuse large B-cell lymphoma: a tissue microarray and clinico-pathological analysis, BMC Cancer 5 (2005) 162, https://doi.org/10.1186/1471-2407-5-162.

[7] L. Haery, J.G. Lugo-Picó, R.A. Henry, A.J. Andrews, T.D. Gilmore, Histone acetyltransferase-deficient p300 mutants in diffuse large B cell lymphoma have altered transcriptional regulatory activities and are required for optimal cell growth, Mol. Cancer 13 (2014) 29, https://doi.org/10.1186/1476-4598-13-29.

[8] L. Pasqualucci, D. Dominguez-Sola, A. Chiarenza, G. Fabbri, A. Grunn, V. Trifonov, L.H. Kasper, S. Lerach, H. Tang, J. Ma, et al., Inactivating mutations of acetyltransferase genes in B-cell lymphoma, Nature 471 (2011) 189–195, https://doi.org/10.1038/nature09730.

[9] A. González-Magaña, F.J. Blanco, Human PCNA structure, function and interactions, Biomolecules 10 (2020) 570, https://doi.org/10.3390/biom10040570.

[10] S.C. Wang, PCNA: a silent housekeeper or a potential therapeutic target? Trends Pharmacol. Sci. 35 (2014) 178–186, https://doi.org/10.1016/j.tips.2014.02.004.

[11] S. Lim, R. Khoo, K.M. Peh, J. Teo, S.C. Chang, S. Ng, G.L. Beilhartz, R.A. Melnyk, C. W. Johannes, C.J. Brown, et al., BioPROTACs as versatile modulators of intracellular therapeutic targets including proliferating cell nuclear antigen (PCNA, Proc. Natl. Acad. Sci. U. S. A 117 (2020) 5791–5800, https://doi.org/10.1073/pnas.1920251117.

[12] V. Ribrag, L. Massade, A.M. Faussat, F. Dreyfus, C. Bayle, A. Gouyette, J.P. Marie, Drug resistance mechanisms in chronic lymphocytic leukemia, Leukemia 10 (1996) 1944–1949.

[13] M. Cardano, C. Tribioli, E. Prosperi, Targeting proliferating cell nuclear antigen (PCNA) as an effective strategy to inhibit tumor cell proliferation, Curr. Cancer Drug Targets 20 (2020) 240–252, https://doi.org/10.2174/1568009620666200115162814.

[14] A. Del Giglio, S. O'Brien, R.J. Ford, J. Manning, H. Saya, M. Keating, D. Johnston, D.F. Chamone, A.B. Deisseroth, Proliferating cell nuclear antigen (PCNA) expression in chronic lymphocytic leukemia (CLL), Leuk. Lymphoma 10 (1993) 265–271, https://doi.org/10.3109/10428199309148548.

[15] C. Yang, Y. Zhang, Y. Chen, F. Ragaller, M. Liu, S. Corvigno, H. Dahlstrand, J. Carlson, Z. Chen, A. Näsman, et al., Nuclear IGF1R Interact with PCNA to Preserve DNA Replication after DNA-Damage in a Variety of Human Cancers, PloS One 15 (2020), e0236291, https://doi.org/10.1371/journal.pone.0236291.

[16] Y. Zhao, X.X. Wang, W. Wu, H. Long, J. Huang, Z. Wang, T. Li, S. Tang, B. Zhu, D. Chen, EZH2 Regulates PD-L1 Expression via HIF-1α in Non-Small Cell Lung Cancer Cells, Biochem. Biophys. Res. Commun. 517 (2019) 201–209, https://doi.org/10.1016/j.bbrc.2019.07.039.

[17] X. Wang, L.T. Brea, J. Yu, Immune modulatory functions of EZH2 in the tumor microenvironment: implications in cancer immunotherapy, Am. J. Clin. Exp. Urol. 7 (2019) 85–91.

[18] S. Varambally, S.M. Dhanasekaran, M. Zhou, T.R. Barrette, C. Kumar-Sinha, M. G. Sanda, D. Ghosh, K.J. Pienta, R.G.A.B. Sewalt, A.P. Otte, et al., The Polycomb Group Protein EZH2 Is Involved in Progression of Prostate Cancer, Nature 419 (2002) 624–629, https://doi.org/10.1038/nature01075.

[19] B. Li, W.-J. Chng, EZH2 abnormalities in lymphoid malignancies: underlying mechanisms and therapeutic implications, J. Hematol. Oncol. 12 (2019) 118, https://doi.org/10.1186/s13045-019-0814-6.

[20] S.K. Loo, E.S. Ch'ng, C.H. Lawrie, M.A. Muruzabal, A. Gaafar, M.P. Pomposo, A. Husin, M.S. Md Salleh, A.H. Banham, L.M. Pedersen, et al., DNMT1 Is Predictive of Survival and Associated with Ki-67 Expression in R-CHOP-Treated Diffuse Large B-Cell Lymphomas, Pathol. (Phila.) 49 (2017) 731–739, https://doi.org/10.1016/j.pathol.2017.08.009.

[21] S.K. Loo, S.S. Ab Hamid, M. Musa, K.K. Wong, DNMT1 is associated with cell cycle and DNA replication gene sets in diffuse large B-Cell Lymphoma, Pathol. Res. Pract. 214 (2018) 134–143, https://doi.org/10.1016/j.prp.2017.10.005.

[22] D.H. Kim, H.-M. Kim, P.T.T. Huong, H.-J. Han, J. Hwang, H. Cha-Molstad, K. H. Lee, I.-J. Ryoo, K.E. Kim, Y.H. Huh, et al., Enhanced anticancer effects of a methylation inhibitor by inhibiting a novel DNMT1 target, CEP 131, in cervical cancer, BMB Rep. 52 (2019) 342–347.

[23] I. Faleiro, R. Leão, A. Binnie, R.A. de Mello, A.-T. Maia, P. Castelo-Branco, Epigenetic therapy in urologic cancers: an update on clinical trials, Oncotarget 8 (2017) 12484–12500, https://doi.org/10.18632/oncotarget.14226.

[24] H. Li, J.-W. Liu, L.-P. Sun, Y. Yuan, A. Meta-Analysis, of the Association between DNMT1 Polymorphisms and Cancer Risk, BioMed. Res. Int. 2017 (2017) 3971259, https://doi.org/10.1155/2017/3971259.

[25] C.P. Hans, D.D. Weisenburger, T.C. Greiner, R.D. Gascoyne, J. Delabie, G. Ott, H. K. Müller-Hermelink, E. Campo, R.M. Braziel, E.S. Jaffe, et al., Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray, Blood 103 (2004) 275–282, https://doi.org/10.1182/blood-2003-05-1545.

[26] H. Kim, Y.M. Kwon, J.S. Kim, J. Han, Y.M. Shim, J. Park, D.-H. Kim, Elevated MRNA Levels of DNA Methyltransferase-1 as an Independent Prognostic Factor in Primary Nonsmall Cell Lung Cancer, Cancer 107 (2006) 1042–1049, https://doi.org/10.1002/cncr.22087.

[27] Y. Saito, Y. Kanai, T. Nakagawa, M. Sakamoto, H. Saito, H. Ishii, S. Hirohashi, Increased protein expression of DNA methyltransferase (DNMT) 1 Is significantly correlated with the malignant potential and poor prognosis of human hepatocellular carcinomas, Int. J. Cancer 105 (2003) 527–532, https://doi.org/10.1002/ijc.11127.

[28] J.-J. Zhang, Y. Zhu, Y. Zhu, J.-L. Wu, W.-B. Liang, R. Zhu, Z.-K. Xu, Q. Du, Y. Miao, Association of increased DNA methyltransferase expression with carcinogenesis and poor prognosis in pancreatic ductal adenocarcinoma, Clin. Transl. Oncol. . Publ. Fed. Span. Oncol. Soc. Natl. Cancer Inst. Mex. 14 (2012) 116–124, https://doi.org/10.1007/s12094-012-0770-x.

[29] F. Leng, J. Yu, C. Zhang, et al., Methylated DNMT1 and E2F1 are targeted for proteolysis by L3MBTL3 and CRL4DCAF5 ubiquitin ligase, Nat. Commun. 9 (2018) 1641, https://doi.org/10.1038/s41467-018-04019-9.

[30] A.B. Shyu, M.F. Wilkinson, A. van Hoof, Messenger RNA regulation: to translate or to degrade, EMBO J. 27 (3) (2008) 471–481, https://doi.org/10.1038/sj.emboj.7601977.

[31] A.T. Agoston, P. Argani, S. Yegnasubramanian, A. De Marzo, M.A. Ansari Lari, J. L. Hicks, N.E. Davidson, W.G. Nelson, Increased protein stability causes DNA methyltransferase 1 dysregulation in breast cancer, J. Biol. Chem. 280 (18) (2005) 18302–18310, https://doi.org/10.1074/jbc.M501675200.).

[32] G. Xouri, Z. Lygerou, H. Nishitani, V. Pachnis, P. Nurse, S. Taraviras, Cdt1 and geminin are down-regulated upon cell cycle exit and are over-expressed in cancer-derived cell lines, Eur. J. Biochem 271 (2004) 3368–3378, https://doi.org/10.1111/j.1432-1033.2004.04271.x.

[33] H. Nishitani, S. Taraviras, Z. Lygerou, T. Nishimoto, The human licensing factor for DNA replication Cdt1 accumulates in G1 and is destabilized after initiation of S-phase, J. Biol. Chem. 276 (2001) 44905–44911.

[34] M. Juríková, Ľ. Danihel, Š. Polák, I. Varga, Ki67, PCNA, and MCM proteins: markers of proliferation in the diagnosis of breast cancer, Acta Histochem 118 (2016) 544–552, https://doi.org/10.1016/j.acthis.2016.05.002.

[35] X.-Y. He, X.-M. Pan, M.-M. Jin, Y.-L. Yang, Z.Q. Yang, D. Yan, J.-X. Ma, Long non-coding RNA AK027294 promotes tumor growth by upregulating PCNA in gastric cancer, Eur. Rev. Med. Pharmacol. Sci. 23 (2019) 5762–5769, https://doi.org/10.26355/eurrev_201907_18314.

[36] K. Klotz-Noack, D. McIntosh, N. Schurch, N. Pratt, J.J. Blow, Re-replication induced by geminin depletion occurs from G2 and is enhanced by checkpoint activation, J. Cell. Sci. 125 (2012) 2436–2445.

[37] A. Ballabeni, R. Zamponi, J.K. Moore, K. Helin, M.W. Kirschner, Geminin deploys multiple mechanisms to regulate Cdt1 before cell division thus ensuring the proper execution of DNA replication, Proc. Natl. Acad. Sci. U.S.A. 110 (30) (2013) E2848–E2853, https://doi.org/10.1073/pnas.1310677110.