INTERNATIONAL JOURNAL OF
MEDICAL BIOCHEMISTRY

DOI: 10.14744/ijmb.2025.47568
Int J Med Biochem 2025;8(3):165–177

Research Article

Genomic and structural analysis of genes involved in epigenetic regulations of diffuse large B cell lymphoma by computational approaches

Betul Akcesme[1,2], Abas Sezer[1]

[1]Department of Genetics and Bioengineering, International University of Sarajevo Faculty of Engineering and Natural Sciences, Sarajevo, Bosnia and Herzegovina
[2]Department of Medical Biology, University of Health Sciences, Hamidiye Faculty of Medicine, Istanbul, Türkiye

Abstract

Objectives: Despite the advances and huge efforts in determination of efficient therapy options for diffuse large B cell lymphoma (DLCBL), various issues, including resistance and toxicity, are among the main concerns. Additional contributor to the poor prognosis is genomic complexity of DLBCL, where high number of mutations are associated with the DLBCL pathogenesis. Epigenetic regulations expand the complexity of gene expression process by various mechanisms, including histone deacetylation and DNA methylation, without altering the DNA sequence.

Methods: Present study focused on analysis of genes involved in epigenetic regulations in DLBCL by using computational approaches, including sequence and structure-based predictions. C-bioportal database was used to find mutations, which are further analyzed by structural bioinformatic tools, including PredictSNP, SNPs&GO, AlphaMissense and DynaMut servers.

Results: Our results showed that mutations R1446H/C and Y1503F/D of CREBBP, L415P, H1451Y, Y1467D of EP300 and Y641N/F of EZH2 proteins are among most common mutations in DLBCL, with the mutations mostly being putative drivers and having a negative impact on the sequence and structure of the proteins.

Conclusion: The understanding of correlation of identified mutations in our study and DLBCL pathogenesis could contribute to the enhanced prediction, diagnosis and treatment of DLBCL. Identification of critical epigenetic mutations might enhance the efficiency and development of epigenetic drugs, further leading to better and more effective targeted therapies. However, further *in-vitro* and *in-vivo* investigations are required to verify our findings based on computational methods.

Keywords: Diffuse large B-cell lymphoma, epigenetic regulators, mutation analysis, structural bioinformatics

How to cite this article: Akcesme B, Sezer A. Genomic and structural analysis of genes involved in epigenetic regulations of diffuse large B cell lymphoma by computational approaches. Int J Med Biochem 2025;8(3):165–177.

Although recent advances in cancer treatment contributed to more efficient treatment options, the number of deaths caused by cancer is increasing. Recent report of the European Union (EU) discovers that the new cancer cases and cancer related deaths in EU states increased more than 2.3% and 2.4%, respectively, over the last two years, while the prognosis for 2040 is that there will be increase of 21% in new cancer cases compared to 2020 [1, 2].

Diffuse large B-cell lymphoma (DLBCL) is characterized as mature B-cell neoplasm and one of the most common lymphoma subtypes [3], where neoplastic cells are large in size and organized in a diffuse pattern when compared to the

Address for correspondence: Betul Akcesme, MD. Department of Genetics and Bioengineering, International University of Sarajevo Faculty of Engineering and Natural Sciences, Sarajevo, Bosnia and Herzegovina; Department of Medical Biology, University of Health Sciences, Hamidiye Faculty of Medicine, Istanbul, Türkiye
Phone: +387603111901 **E-mail:** betul.akcesme@sbu.edu.tr **ORCID:** 0000-0002-3730-9963
Submitted: January 22, 2025 **Revised:** March 11, 2025 **Accepted:** March 12, 2025 **Available Online:** June 17, 2025

non-cancerous tissue [3]. While median age for DLBCL diagnosis is between 60 and 70 years, younger patients can be diagnosed as well [4]. Enlargement of lymph nodes and highly aggressive growth of tumor mass in bone marrow, liver and other organs are some of the characteristic representations of DLBCL [3, 5]. During the cell development stages, malignant B-cell proliferation triggers DLBCL formation, which is classified into two main subgroups based on the cell of origin: Activated B-cell (ABC) DLBCL and germinal center B-cell (GCB) DLBCL [6]. Despite the advances and huge efforts in determination of efficient therapy options for DLCBL treatment, various issues, including resistance and toxicity, are among the main concerns. Additional contributor to the poor prognosis is genomic complexity of DLBCL, where high number of mutations are associated with the DLBCL pathogenesis [7]. Treatments of DLBCL commonly include groups of drugs such as CHOP (cyclophosphamide, doxorubicin, vincristine and prednisone) and R-CHOP (with addition of rituximab antibody) immunochemotherapy [8]. Additionally, chimeric antigen receptor-modified (CAR) T-cell therapy shows great potential for DLBCL treatment, especially for patients with resistance and relapsed DLBCL. This therapy is based on the genetic modification of the patient's own T-cells to target the cancer [9].

Epigenetic regulations expand the complexity of gene expression process by various epigenetic mechanisms, including histone deacetylation and DNA methylation, without altering the DNA sequence [10, 11]. Numerous groups of genes, including DNA methyltransferases (DNMTs) [12], ten eleven translocation genes (TET) [13], histone deacetylases (HDACs) [14], histone acetyltransferases (HATs) [15], polycomb group associated genes [16], switch/sucrose non-fermentable (SWI/SNF) genes [17], lysine methyltransferases (KMTs) [18], and lysine demethylases (KDMs) [19] are essential for epigenetic regulations. Mutations of these genes are leading to epigenetic dysregulations and alterations, further affecting gene expression and increasing the risk of cancer development. Various cancer types, including breast [20], prostate [21] and colorectal [22] cancer, acute myeloid leukemia [23], follicular lymphoma [24] and DLBCL [25] are associated with the mutations of epigenetic regulator genes. In particular, frequent mutations in the epigenetic regulatory genes CREBBP, KMT2D, EZH2 and TET2 were revealed upon genetic profiling of DLBCL patients, promoting the malignant transformation of B cells through altered epigenetic regulation [26]. Promising steps are being taken in relation to epigenetic therapy for DLBCL, where tazemetostat inhibitor of epigenetic regulator EZH2, might offer favorable treatment option for patients with EZH2 mutations. However, this approach still has certain risks, including selectivity of targets and oncogenic activation [27].

Additionally, important role in cancer research belongs to the computational analysis, where genomic data is analyzed by using various computational methods in order to correlate mutations with the cancer progression and pathogenesis. Computational analysis might be used to provide additional information and help in prognosis, early detection, as well as targeted therapy for cancer treatments [28]. Moreover, identification of critical epigenetic mutations might enhance the efficiency and development of epigenetic drugs, further leading to better and more effective targeted therapies [29].

The effort devoted to a better understanding of the factors and trigger points associated with cancer progression and development can have pivotal role in reducing the trend of new cancer cases. The scope of this study is to analyze genes categorized as epigenetic regulators involved in DLBCL pathogenesis, their mutations, along with the effects on protein sequence and structure by using multiple databases.

Materials and Methods

Gene selection

Based on the literature analysis, genes selected for this study are mainly involved in epigenetic regulations and are potentially able to contribute to development of different cancer types. 69 genes belonging to the following groups: DNA methyltransferases, HDACs, HATs, ten-element translocation genes, polycomb group associated genes, SWI/SNF complex genes, KMTs, KDMs and sirtuins were included. C-bioportal database [30] was used for DLBCL studies, and the database with the largest sample size (1001 samples, Duke, cell 2017) was selected, from which genes involved in epigenetic regulations with certain mutations in DLBCL patients were selected.

Mutation analysis

Mutations in selected genes were further filtered based on the mutation type by using c-bioportal database [30]. Research focused on the mutations that are known to be putative drivers, including missense, truncating (nonsense, nonstop, frameshift deletion, frameshift insertion or splice site), inframe (deletion or insertion), splice and fusion mutations. Mutations that are categorized as mutations of unknown significance are excluded from the analysis. Following that, genes with often reported mutations in DLBCL patients were selected.

Sequence-based prediction

Impact of the mutations on the protein function was evaluated by using sequence-based computational tools, including PredictSNP [31], SNPs&GO [32], and AlphaMissense [33]; where PredictSNP is a consensus of six constituent tools, including MAPP, PhD-SNP, PolyPhen-1, PolyPhen-2, SIFT and SNAP for disease-related mutation prediction obtained from the Uniprot and Protein Mutant Database; SNPs&GO is a server for single point protein mutation prediction based on annotation of functional protein; and AlphaMissense is a tool used as a predictor of the pathogenicity of missense variants. Sequences of CREBBP (Q92793), EP300 (Q09472) and EZH2 (Q15910) were used for the analysis from the Uniprot database [34].

Structure-based prediction

Impact of the mutations on the protein stability was evaluated by using structure-based computational tool DynaMut [35], which is using normal mode analysis to evaluate changes

Table 4. Structure-based prediction results based on the DynaMut tool

Gene	Mutation	ΔΔG (kcal/mol)	$\Delta\Delta S_{Vib}$ (kcal.mol^{-1}.K^{-1})
CREBBP	R1446H	-1.135 (-)	0.244 (–)
	R1446C	-0.904 (-)	0.279 (–)
	Y1503F	-2.475 (-)	0.918 (–)
	Y1503D	-2.475 (-)	0.918 (–)
EP300	L415P	-1.341 (-)	0.409 (–)
	H1451Y	9.706 (+)	-0.167 (++)
	Y1467D	-2.399 (-)	0.882 (–)
EZH2	Y641N	0.578 (+)	0.053 (–)
	Y641F	0.578 (+)	0.053 (–)

ΔΔG change in the Gibbs free energy; $\Delta\Delta S_{Vib}$ change in vibrational entropy energy between wild-type and mutant; (+): Stabilizing; (-): Destabilizing; (++): Increased molecule flexibility; (--): Decreased molecule flexibility.

in protein stability. Protein Data Bank in Europe Knowledge Base (PDBe-KB) database [36] and Protein Data Bank in Europe (PDBe) database [37] were both used to select the PDB protein structures, chains containing the mutation positions, along with analysis of interaction interfaces.

Visualization of interactions in wild-type and mutant structures

Selected proteins were visualized by using BIOVIA Discovery Studio 2021. After selection of mutation position, interaction types, including hydrogen, electrostatic and hydrophobic bonds, along with the unfavorable bumps were selected from the interaction monitor. Additionally, molecular scope was set to any atom-to-atom interaction. Interactions with the surrounding residues were analyzed and evaluated in both, wild-type and mutant structures.

This study was conducted in accordance with the ethical principles of the Helsinki Declaration. As no human participants or patient data were directly involved, specific ethical approval was not required. All data used in this study are publicly available online at https://www.cbioportal.org/ [30].

Results

Based on the literature analysis, 69 genes that belong to the various groups of epigenetic regulators were selected (Appendix 1). By using the c-bioportal database, 14 genes, including DNMT3A, CREBBP, EP300, TAF1, EZH2, ARID1A, ARID1B, ARID5b, SETD1B, SETD2, SETD5, KMT2C and KMT2D are further filtered and selected with the reported mutations in DLBCL patients (Table 1).

Next, among above mentioned genes, genes with the highest number of reported mutations in various DLBCL patients from c-biportal database were selected, including CREBBP, EP300 and EZH2, with four (R1446H, R1446C, Y1503F, Y1503D), three (L415P, H1451Y, Y1467D) and two (Y641N, Y641F) often reported missense mutations in DLBCL patients, respectively. R1446H/C and Y1503F/D missense mutations of CREBBP were reported in 18 patients, L415P, H1451Y, and Y1467D missense mutations of EP300 were reported in 9 patients, while Y641N/F missense mutations of EZH2 were reported in 57 patients.

PredictSNP computational tool, together with the consensus tools, predicted that nearly all selected mutations are deleterious with the high expected accuracy percentages, except for Y1503F mutation with the expected accuracy of 58% for PhD-SNP and L415P mutation with the expected accuracy of 50% for SNAP tool (Table 2). Additionally, according to SNPs&GO computational tool, all mutations are categorized as disease related polymorphisms with the highest reliability index (RI) values being reported for EP300 gene, being 8, 10 and 9 for L415P, H1451Y and Y1467D mutations respectively, while the

Table 1. List of selected genes with the mutation percentages from C-bioportal database, 1001 DLBCL samples

No	Gene	Group	Mutation (%)	Source
1	DNMT3A	DNA methyltransferase	4.6	[38]
2	CREBBP	Histone acetyltransferase	11.5	[15]
3	EP300	Histone acetyltransferase	5.7	[15]
4	TAF1	Histone acetyltransferase	3.4	[39]
5	TET2	Ten element translocation	6.2	[40]
6	EZH2	Polycomb group	6.1	[41]
7	ARID1A	SWI/SNF complex	9.6	[42]
8	ARID1B	SWI/SNF complex	8.3	[42]
9	ARID5B	SWI/SNF complex	3.2	[43]
10	SETD1B	Lysine methyltransferase	8.3	[44]
11	SETD2	Lysine methyltransferase	6.2	[44]
12	SETD5	Lysine methyltransferase	3.1	[45]
13	KMT2C	Lysine methyltransferase	5.6	[46]
14	KMT2D	Lysine methyltransferase	24.7	[46]

DLBCL: Diffuse large B-cell lymphoma; SWI/SNF: Switch/sucrose non-fermentable.

Table 2. Sequence-based prediction results based on the PredictSNP and its consensus tools

Gene	Mutation	Predict SNP	MAPP	PhD-SNP	PolyPhen-1	PolypPhen-2	SIFT	SNAP
CREBBP	R1446H	87% (-)	51% (-)	88% (-)	74% (-)	81% (-)	79% (-)	85% (-)
	R1446C	87% (-)	76% (-)	88% (-)	74% (-)	81% (-)	79% (-)	85% (-)
	Y1503F	76% (-)	62% (-)	58% (+)	74% (-)	81% (-)	79% (-)	62% (-)
	Y1503D	87% (-)	77% (-)	82% (-)	74% (-)	81% (-)	79% (-)	81% (-)
EP300	L415P	76% (-)	57% (-)	82% (-)	59% (-)	81% (-)	79% (-)	50% (+)
	H1451Y	87% (-)	57% (-)	88% (-)	74% (-)	81% (-)	79% (-)	89% (-)
	Y1467D	87% (-)	86% (-)	86% (-)	74% (-)	81% (-)	79% (-)	89% (-)
EZH2	Y641N	87% (-)	82% (-)	77% (-)	59% (-)	60% (-)	53% (-)	85% (-)
	Y641F	87% (-)	86% (-)	82% (-)	74% (-)	65% (-)	79% (-)	89% (-)

%: Expected accuracy; (+): Neutral; (-): Deleterious.

Table 3. Sequence-based prediction results based on the SNPs&GO computational tool and AI generated AlphaMissense server

Gene	Mutation	SNPs&GO		AlphaMissense	
		Effect	RI	Class	PS
CREBBP	R1446H	Disease related polymorphism	7	Likely pathogenic	0.995
	R1446C	Disease related polymorphism	7	Likely pathogenic	0.994
	Y1503F	Disease related polymorphism	4	Likely pathogenic	0.907
	Y1503D	Disease related polymorphism	5	Likely pathogenic	0.999
EP300	L415P	Disease related polymorphism	8	Likely pathogenic	0.999
	H1451Y	Disease related polymorphism	10	Likely pathogenic	0.985
	Y1467D	Disease related polymorphism	9	Likely pathogenic	0.999
EZH2	Y641N	Disease related polymorphism	3	Likely pathogenic	0.999
	Y641F	Disease related polymorphism	3	Likely pathogenic	0.965

RI: Reliability index.

lowest RI reported is 3 for both, Y641N/F mutations for EZH2 (Table 3). Moreover, all selected mutations are showed to be likely pathogenic based on the AlphaMissense tool with pathogenicity values higher than 0.9 (Table 3).

Pathogenicity Score: Pathogenicity values range from 0 to 1 representing the likelihood of pathogenicity as follows: 0.0 – 0.33 Likely benign; 0.34 – 0.66 ambiguous; 0.67 – 1.0 likely pathogenic.

Prior to structure-based prediction, protein structures, along with the chains containing the position of mutations were selected by using PDBe-KB and PDBe. For CREBBP mutations, R1446H/C and Y1503F/D, chain K of 8HAN (PDB ID) protein structure was selected, for EP300 L415P mutation, chain A of 7QGS (PDB ID) protein structure and for EP300 H1451Y and Y1467D mutations chain A of 5KJ2 (PDB ID) protein structures were selected. For EZH2 Y641N/F mutations, chain K of 6C23 (PDB ID) protein structure was selected. Structure-based prediction analysis revealed that most of the selected mutations are either destabilizing or/and they are increasing the molecule flexibility. R1446H/C and Y1503F/D mutations of CREBBP and L415P and Y1467D mutation of EP300 are both, destabilizing and increasing the molecule flexibility. On the other hand, only H1451Y mutation of EP300 decreases the molecule flexibility and stabilizes the structure (Table 4 and Fig. 1).

Furthermore, differences in the interactions in wild-type and mutant structures were observed. CREBBP wild-type Arg1446 has four conventional hydrogen bond interactions with Tyr1450, Val1449, Arg1443 and His1438 and two Pi-Alkyl interactions with Tyr1450 and Phe1440. CREBBP mutant His1446 has three conventional hydrogen bond interactions with Tyr1450, Vak1449 and Arg1443 and one Pi-Pi stacked interaction with Phe1440, while Cys1446 mutant of the same protein has three conventional hydrogen bond interactions with Tyr1450, Val1449 and Arg1443, Pi-Alkyl interaction with Phe1440 and Pi-donor hydrogen bond interaction with Tyr1450 (Fig. 2a-c). CREBBP wild-type Tyr1503 and mutant Phe1503 have four conventional hydrogen bond interactions with Leu1499, Gln1500, Met1506, and Leu1507, Pi-Alkyl interactions with Leu1499 and Ala1474, Pi-Pi T-shaped interactions with Trp1502 and Pi-Sigma bond interactions with Ile1471, while Asp1503 mutant of the same gene has same conventional hydrogen bond interactions without other interactions for the particular mutation position (Fig. 2d-f).

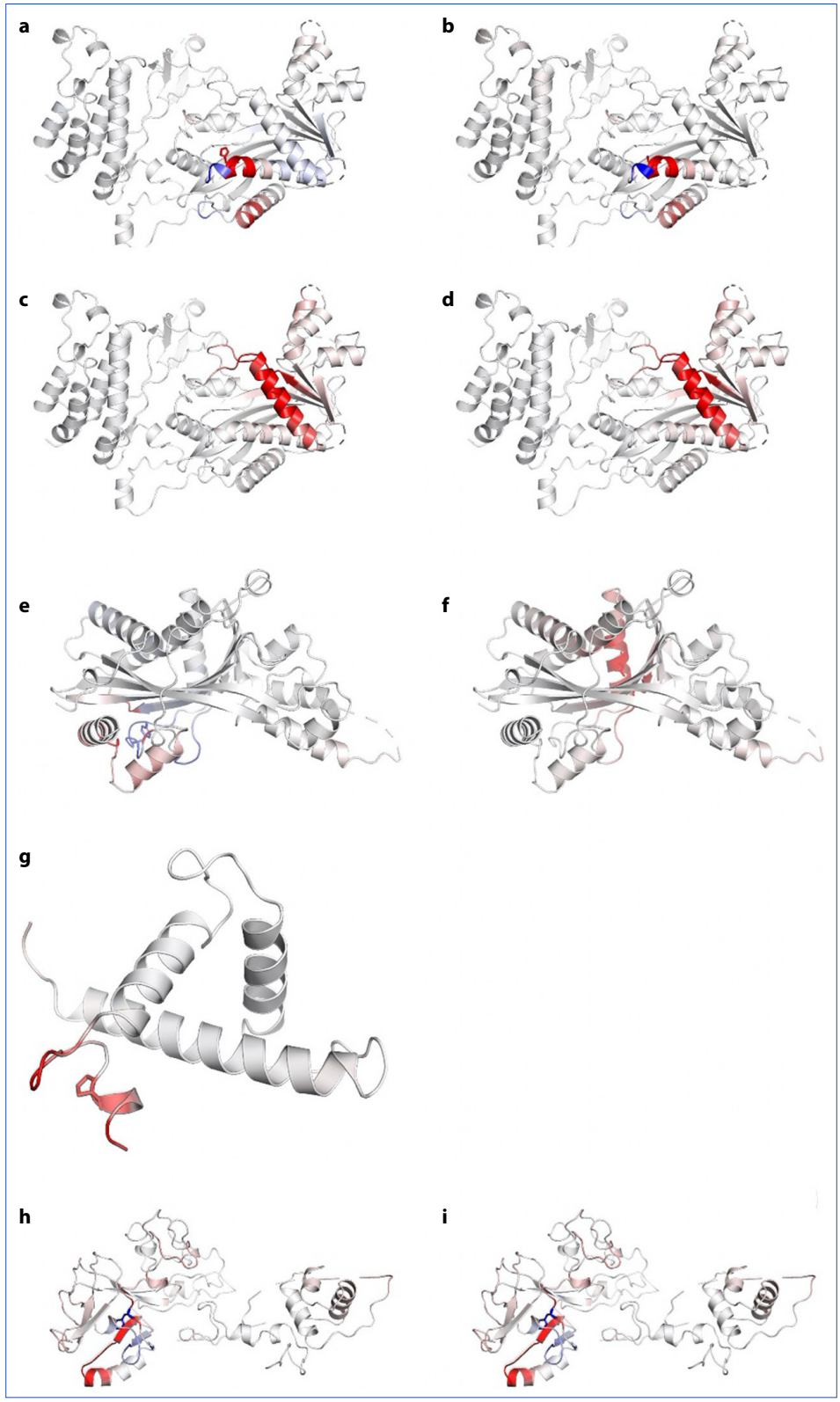

Figure 1. Flexibility visualization of mutations on CREBBP (a-d), EP300 (e-g) and EZH2 (h, i). (a) R1446H; (b) R1446C; (c) Y1503F; (d) Y1503D; (e) H1451Y; (f) Y1467D; (g) L415P; (h) Y641N; (i) Y641F.

Differences are also observed between wild-type and mutant structures on the positions 415, 1451 and 1467 of EP300. Leu415 wild-type structure has two conventional hydrogen bond interactions with Cys411 and Asn419, while Pro415 mutant structure has only one conventional hydrogen bond interaction with Asn419 (Fig. 3a, b). His1451 wild-type structure of

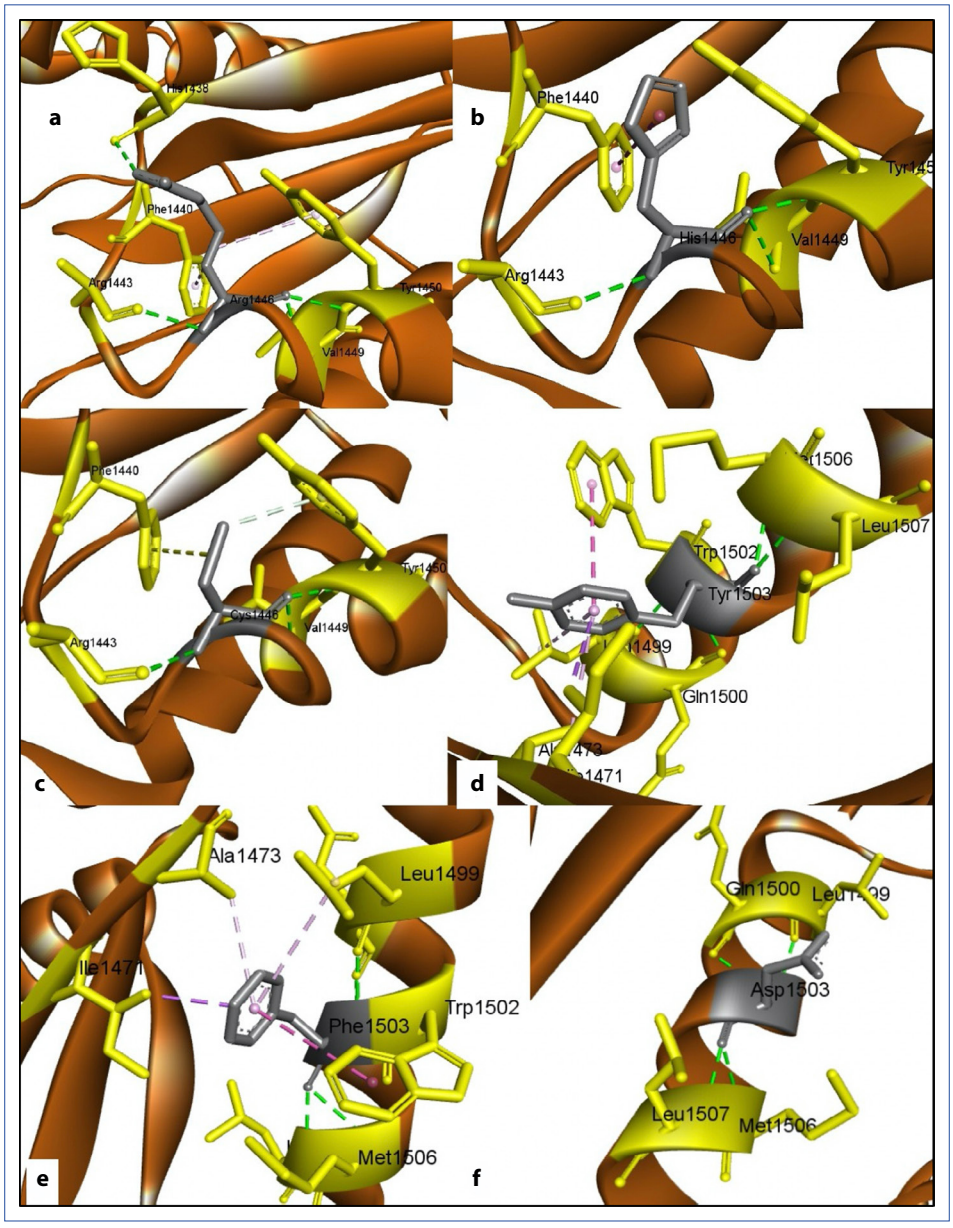

Figure 2. Visualization of interaction in wild-type and mutant structures on 1446 and 1503 positions of CREBBP protein. Brown – protein chain; yellow – interacting residues; grey – mutation position; conventional hydrogen bond interactions – green; Pi bond interactions – pink; (a) Arg1446 wild-type; (b) His1446 mutant; (c) Cys1446 mutant; (d) Tyr1503 wild-type; (e) Phe1503 mutant; (f) Asp1503 mutant (Created by BIOVIA Discovery Studio 2021).

the same protein has one carbon hydrogen bond interaction while Tyr1451 mutant has no interactions (Fig. 3c, d). Wild-type Tyr1467 of EP300 has four conventional hydrogen bond interactions with Leu1463, Gln1464, Met1470 and Leu147, three Pi-Alkyl interactions with Ile1435, Ala1437 and Leu1463 and one Pi-Pi T-shaped interaction with Trp1466 while Asp1467 mutant has same conventional hydrogen bond interactions without any Pi bond interactions (Fig. 3e, f).

Wild-type Tyr641 of EZH2 has two conventional hydrogen bond interactions with Phe665 and Ile708 and one Pi-Alkyl interaction with Arg685. Asn641 mutant has one conventional hydrogen bond interaction with Ile708 while Phe641 mutant has same conventional hydrogen bond interaction in addition to Arg685 Pi-Alkyl and Phe667 Pi-Pi stacked interaction (Fig. 4).

Discussion

Epigenetic regulators are potent proteins involved in various mechanisms related to gene expression regulations. Any alterations in their structure carries a high risk for tumor development [47]. Although there are numerous epigenetic regulators, focus of this study are epigenetic regulatory genes reported in DLCBL patients. Our results showed that CREBBP, EP300

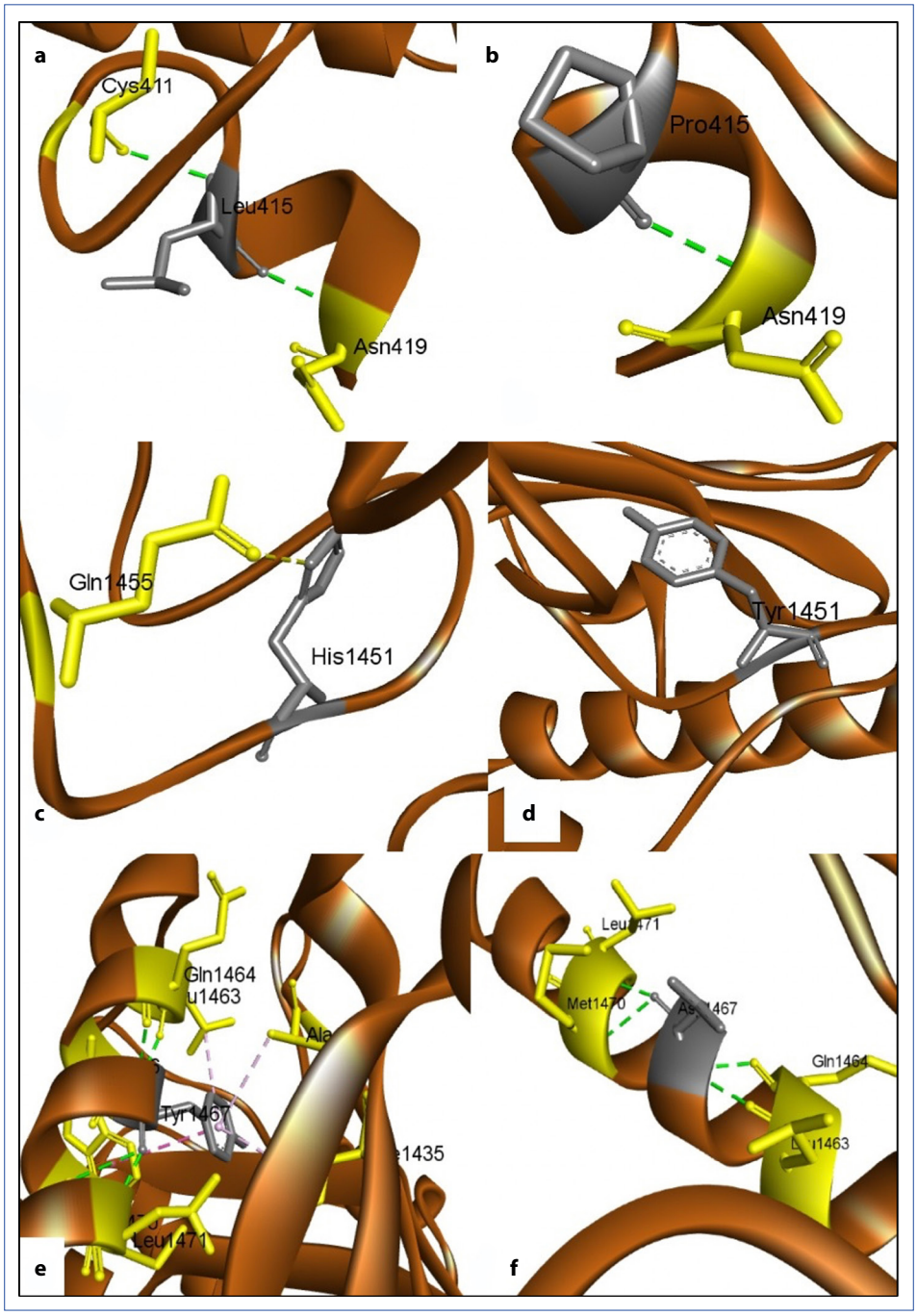

Figure 3. Visualization of interaction in wild-type and mutant structures on 415, 1451 and 1467 positions of EP300 protein. Brown – protein chain; yellow – interacting residues; grey – mutation position; conventional hydrogen bond interactions – green; Pi bond interactions – pink; (a) Leu415 wild-type; (b) Pro415 mutant; (c) His1451 wild-type; (d) Tyr1451 mutant; (e) Tyr1467 wild-type; (f) Asp1467 mutant (Created by BIOVIA Discovery Studio 2021).

and EZH2 are among most mutated epigenetic regulators in DLBCL. CREBBP and EP300 function as the histone acetyltransferases, while EZH2 has a role in polycomb repressive complex [48]. Various reports underline the importance of these genes in hematological malignancies, especially in DLBCL [49, 50]. CREBBP has crucial role in different processes, including regulation of gene expression, along with the involvement in cell growth, development and differentiation [50]. Similarly, EP300 has a pivotal role in cell proliferation, apoptosis and differentiation regulation through interactions with various transcription factors, including p53 and NF-κB [50]. Moreover, EZH2 acts as a transcriptional repressor, where EZH2 dysregulation is mainly associated with negative outcomes when it comes to the cancer progression and resistance [51].

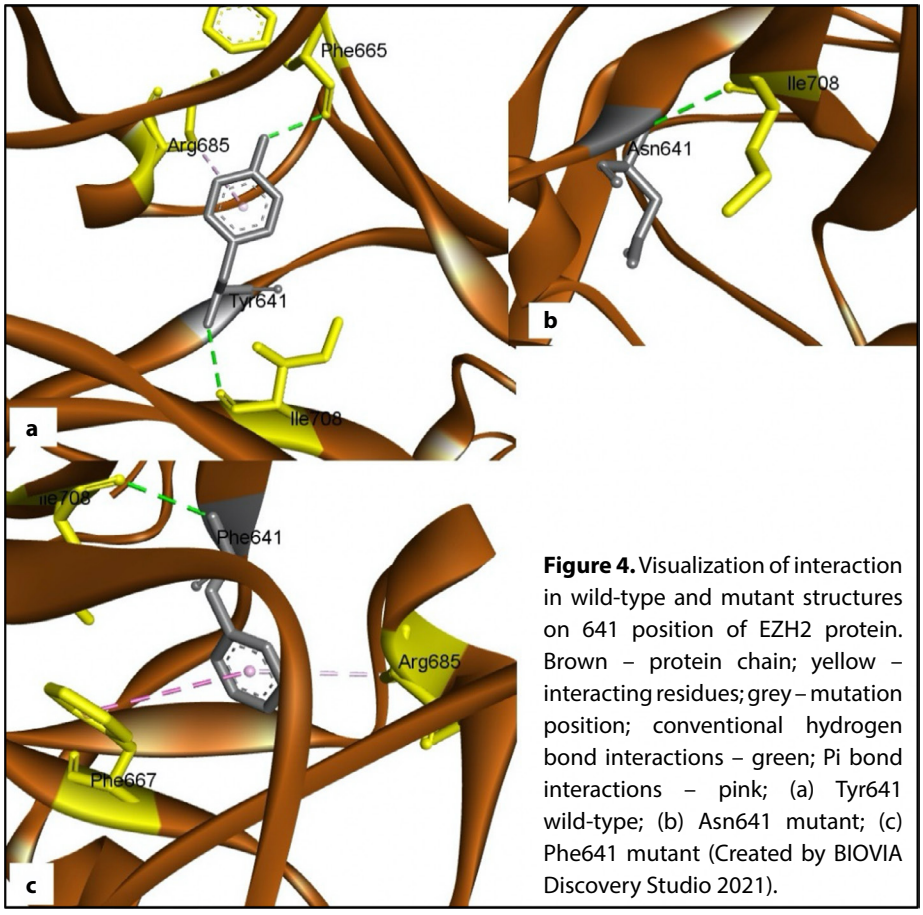

Figure 4. Visualization of interaction in wild-type and mutant structures on 641 position of EZH2 protein. Brown – protein chain; yellow – interacting residues; grey – mutation position; conventional hydrogen bond interactions – green; Pi bond interactions – pink; (a) Tyr641 wild-type; (b) Asn641 mutant; (c) Phe641 mutant (Created by BIOVIA Discovery Studio 2021).

Our results suggest that selected mutations, including R1446H/C and Y1503F/D for CREBBP, L415P, H1451Y and Y1467D for EP300, and Y641 N/F for EZH2 proteins have mostly negative impact on the sequence and structure of the proteins, being either deleterious, likely pathogenic, destabilizing and/or enhancing the molecule flexibility, along with the impact on the change of interactions between the residues.

Although the bonds in wild-type and mutant structures are not directly interfering with the interaction interfaces indicated in PDBe-KB database, they are in close proximity to mutation positions. In relation to that, selected mutations could possibly affect these interactions. Additionally, some of the mutations are positioned on ligand binding sites, which could impact the binding affinity and drug resistance [52]. Our results showed that most common CREBBP mutations are R1446 and Y1503. These mutations are also reported in relapse acute lymphoblastic leukemia as well [53]. Previous studies indicated the negative correlation between the R1446 and Y1503 mutations of CREBBP and cancer progression in DLBCL and B-cell lymphomas respectively [54]. Pasqualucci et al. [55], showed that the commonly shared pathogenic mechanisms of non-Hodgkin B-cell lymphoma are CREBBP and EP300 mutations. However, additional analysis is required to achieve a more comprehensive understanding.

Another common mutation selected in current study is H1451 of EP300. Duex et al. [56], discussed that H1451 residue mutations of EP300 lead to inactivation of HAT. This is in accordance with the findings of current study, as the most often reported mutation of EP300 gene is on H1451 in DLBCL patients. HAT domain has role in gene regulation by neutralizing the positive charge of histones, weakening their interaction with DNA and making chromatin more accessible for transcription. As the mutational hotspots occur in the HAT domain, it might alter the cellular regulations and development, leading to negative outcomes in relation to carcinogenesis. As Huang et al. [57] reported that CREBBP/EP300 gene mutations increased the rate of tumor progression in DLBCL with the lower progression-free survival and lover overall survival rate when compared to the patients without CREBBP/EP300 mutations. It is obvious that both CREBBP/EP300 mutations have a crucial effect on DLBCL pathogenesis.

Another commonly identified mutation is Y641 of EZH2. This particular mutation possesses the high importance, as the overexpression of EZH2 is closely related to tumor suppressor gene silencing [58]. It has been reported that Y641 and A677 heterozygous point mutations in EZH2 occur in 10–24% of all non-Hodgkin lymphoma cases [59]. In addition, Y641 mutation of EZH2 is associated with the immunodeficiency in lymphoid malignancies [60]. Morin et al. [61], also identified Y641 mutation of EZH2 and its correlation with the GCB subtype of DLBCL. According to their study, this particular mutation is related to the pathogenesis of GCB lymphomas.

Recent studies suggest that targeting the epigenetic regulators in various cancer types, including DLBCL, might poses beneficial effects regarding the treatment [27, 62, 63]. The understanding of correlation of analyzed mutation positions and DLBCL pathogenesis could contribute to the enhanced prediction, diagnosis, and treatment of DLBCL.

Conclusion

Any alterations in epigenetic mechanisms carry a high risk for cancer development, and understanding the factors and trigger points associated with cancer progression can play a pivotal role in reducing the incidence of new cancer cases. Since epigenetic regulators are reported to impact DLBCL progression, we focused our research on identifying and analyzing potential proteins and genes involved in the epigenetic regulation of DLBCL. Using various computational tools, we identified and analyzed the effects of commonly mutated positions in the CREBBP, EP300, and EZH2 proteins. Computational analysis, especially with recent advances in artificial intelligence and machine learning methods, plays an important role in identifying specific mutations and their correlation with cancer. However, it also has limitations, including the lack of experimental data, which necessitates additional *in vitro* and *in vivo* research. To the best of our knowledge, this is the first in silico study analyzing sequence and structure-based effects of R1446H/C, Y1503F/D, L415P, H1451Y, Y1467D and Y641N/F mutations.

Informed Consent: Informed consent was obtained from all participants.

Conflict of Interest Statement: The authors have no conflicts of interest to declare.

Funding: The authors declared that this study received no financial support.

Use of AI for Writing Assistance: No AI technologies utilized.

Authorship Contributions: Concept – B.A.; Design – B.A., A.S.; Supervision – B.A.; Data collection and/or processing – A.S.; Data analysis and/or interpretation – B.A., A.S.; Literature search – B.A., A.S.; Writing – A.S.; Critical review – B.A., A.S.

Peer-review: Externally peer-reviewed.

References

1. Joint Research Centre. Cancer cases and deaths on the rise in the EU. European Commission. Available at: https://joint-research-centre.ec.europa.eu/jrc-news-and-updates/cancer-cases-and-deaths-rise-eu-2023-10-02_en. Accessed Jun 2, 2025.
2. Joint Research Centre. European Cancer Information System: 21% Increase in new cancer cases by 2040. European Commission. Available at: https://joint-research-centre.ec.europa.eu/jrc-news-and-updates/european-cancer-information-system-21-increase-new-cancer-cases-2040-2022-03-16_en. Accessed Jun 2, 2025.
3. Padala SA, Kallam A. Diffuse large B-cell lymphoma. Treasure Island (FL): StatPearls Publishing; 2023.
4. Kesavan M, Eyre TA, Collins GP. Front-line treatment of high-grade B-cell non-Hodgkin lymphoma. Curr Hematol Malig Rep 2019;14(4):207–18. [CrossRef]
5. Li S, Young KH, Medeiros LJ. Diffuse large B-cell lymphoma. Pathology (Phila) 2018;50(1):74–87. [CrossRef]
6. Cioroianu AI, Stinga PI, Sticlaru L, Cioplea MD, Nichita L, Popp C, et al. Tumor microenvironment in diffuse large B-cell lymphoma: Role and prognosis. Anal Cell Pathol 2019;2019:1–9. [CrossRef]
7. You H, Xu-Monette ZY, Wei L, Nunns H, Nagy ML, Bhagat G, et al. Genomic complexity is associated with epigenetic regulator mutations and poor prognosis in diffuse large B-cell lymphoma. Oncoimmunology 2021;10(1):1928365. [CrossRef]
8. Lewis WD, Lilly S, Jones KL. Lymphoma: Diagnosis and treatment. Am Fam Physician 2020;101(1):34–41.
9. Hopfinger G, Jäger U, Worel N. CAR-T cell therapy in diffuse large B-cell lymphoma: Hype and hope. Hemasphere 2019;3(2):e185. [CrossRef]
10. Perri F, Longo F, Giuliano M, Sabbatino F, Favia G, Ionna F, et al. Epigenetic control of gene expression: Potential implications for cancer treatment. Crit Rev Oncol Hematol 2017;111:166–72. [CrossRef]
11. Loscalzo J, Handy DE. Epigenetic modifications: Basic mechanisms and role in cardiovascular disease (2013 Grover Conference series). Pulm Circ 2014;4(2):169–74. [CrossRef]
12. Lyko F. The DNA methyltransferase family: A versatile toolkit for epigenetic regulation. Nat Rev Genet 2018;19(2):81–92. [CrossRef]
13. Rasmussen KD, Helin K. Role of TET enzymes in DNA methylation, development, and cancer. Genes Dev 2016;30(7):733–50. [CrossRef]
14. Hassell KN. Histone deacetylases and their inhibitors in cancer epigenetics. Dis 2019;7(4):57. [CrossRef]
15. Gujral P, Mahajan V, Lissaman AC, Ponnampalam AP. Histone acetylation and the role of histone deacetylases in normal cyclic endometrium. Reprod Biol Endocrinol 2020;18(1):84. [CrossRef]
16. Ma R, Zhang Y, Sun T, Cheng B. Epigenetic regulation by polycomb group complexes: Focus on roles of CBX proteins. J Zhejiang Univ Sci B 2014;15(5):412–28. [CrossRef]
17. Wanior M, Krämer A, Knapp S, Joerger AC. Exploiting vulnerabilities of SWI/SNF chromatin remodelling complexes for cancer therapy. Oncogene 2021;40(21):3637–54. [CrossRef]
18. Bhat KP, Ümit Kaniskan H, Jin J, Gozani O. Epigenetics and beyond: Targeting writers of protein lysine methylation to treat disease. Nat Rev Drug Discov 2021;20(4):265–86. [CrossRef]
19. Sterling J, Menezes SV, Abbassi RH, Munoz L. Histone lysine demethylases and their functions in cancer. Int J Cancer 2021;148(10):2375–88. [CrossRef]
20. Man X, Li Q, Wang B, Zhang H, Zhang S, Li Z. DNMT3A and DNMT3B in breast tumorigenesis and potential therapy. Front Cell Dev Biol 2022;10:916725. [CrossRef]
21. Shin HJ, Hua JT, Li H. Recent advances in understanding DNA methylation of prostate cancer. Front Oncol 2023;13:1182727. [CrossRef]

22. Dobre M, Salvi A, Pelisenco IA, Vasilescu F, De Petro G, Herlea V, et al. Crosstalk between DNA methylation and gene mutations in colorectal cancer. Front Oncol 2021;11:697409. [CrossRef]
23. Venney D, Mohd-Sarip A, Mills KI. The impact of epigenetic modifications in myeloid malignancies. Int J Mol Sci 2021;22(9):5013. [CrossRef]
24. Amin R, Braza MS. The follicular lymphoma epigenome regulates its microenvironment. J Exp Clin Cancer Res 2022;41(1):21. [CrossRef]
25. Jiang Y, Melnick A. The epigenetic basis of diffuse large B-cell lymphoma. Semin Hematol 2015;52(2):86–96. [CrossRef]
26. Isshiki Y, Melnick A. Epigenetic mechanisms of therapy resistance in diffuse large B-cell lymphoma (DLBCL). Curr Cancer Drug Targets 2021;21(4):274–82. [CrossRef]
27. Chung C. A promising future for precision epigenetic therapy for follicular and diffuse large B-cell lymphoma? Blood Lymphat Cancer 2022;12:99–106. [CrossRef]
28. Shi X, Teng H, Shi L, Bi W, Wei W, Mao F, et al. Comprehensive evaluation of computational methods for predicting cancer driver genes. Brief Bioinform 2022;23(2):bbab548. [CrossRef]
29. Ribeiro ML, Sánchez Vinces S, Mondragon L, Roué G. Epigenetic targets in B- and T-cell lymphomas: Latest developments. Ther Adv Hematol 2023;14:204062072311734. [CrossRef]
30. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012;2(5):401–4. [CrossRef]
31. Bendl J, Stourac J, Salanda O, Pavelka A, Wieben ED, Zendulka J, et al. PredictSNP: Robust and accurate consensus classifier for prediction of disease-related mutations. PLoS Comput Biol 2024;10(1):e1003440. [CrossRef]
32. Capriotti E, Calabrese R, Fariselli P, Martelli PL, Altman RB, Casadio R. WS-SNPs&GO: A web server for predicting the deleterious effect of human protein variants using functional annotation. BMC Genomics 2013;14:1–7. [CrossRef]
33. Tordai H, Torres O, Csepi M, Padányi R, Lukács GL, Hegedűs T. Analysis of AlphaMissense data in different protein groups and structural context. Sci 2024;11(1):495. [CrossRef]
34. The UniProt Consortium. UniProt: The universal protein knowledgebase. Nucleic Acids Res 2018;46(5):2699. [CrossRef]
35. Rodrigues CH, Pires DE, Ascher DB. DynaMut: Predicting the impact of mutations on protein conformation, flexibility and stability. Nucleic Acids Res 2018;46(W1):W350–5. [CrossRef]
36. The PDBe-KB Consortium. PDBe-KB: A community-driven resource for structural and functional annotations. Nucleic Acids Res 2020;48(D1):D344–53. [CrossRef]
37. Armstrong DR, Berrisford JM, Conroy MJ, Gutmanas A, Anyango S, Choudhary P, et al. PDBe: Improved findability of macromolecular structure data in the PDB. Nucleic Acids Res 2020;48(D1):D335–43. [CrossRef]
38. Jin B, Robertson KD. DNA methyltransferases, DNA damage repair, and cancer. Adv Exp Med Biol 2013;754:3–29. [CrossRef]
39. Hilton TL, Li Y, Dunphy EL, Wang EH. TAF1 histone acetyltransferase activity in Sp1 activation of the cyclin D1 promoter. Mol Cell Biol 2005;25(10):4321–32. [CrossRef]
40. Zhang X, Zhang Y, Wang C, Wang X. TET (Ten-eleven translocation) family proteins: Structure, biological functions and applications. Signal Transduct Target Ther 2023;8(1):297. [CrossRef]
41. Wang W, Qin JJ, Voruganti S, Nag S, Zhou J, Zhang R. Polycomb group (PcG) proteins and human cancers: Multifaceted functions and therapeutic implications. Med Res Rev 2015;35(6):1220–67. [CrossRef]
42. Alver BH, Kim KH, Lu P, Wang X, Manchester HE, Wang W, et al. The SWI/SNF chromatin remodelling complex is required for maintenance of lineage specific enhancers. Nat Commun 2017;8(1):14648. [CrossRef]
43. Lin C, Song W, Bi X, Zhao J, Huang Z, Li Z, et al. Recent advances in the ARID family: Focusing on roles in human cancer. OncoTargets Ther 2014;7:315–24. [CrossRef]
44. Antunes ETB, Ottersbach K. The MLL/SET family and haematopoiesis. Biochim Biophys Acta Gene Regul Mech 2020;1863(8):194579. [CrossRef]
45. Sessa A, Fagnocchi L, Mastrototaro G, Massimino L, Zaghi M, Indrigo M, et al. SETD5 regulates chromatin methylation state and preserves global transcriptional fidelity during brain development and neuronal wiring. Neuron 2019;104(2):271–89. e13. [CrossRef]
46. Tian X, Zhang S, Liu HM, Zhang YB, Blair CA, Mercola D, et al. Histone lysine-specific methyltransferases and demethylases in carcinogenesis: New targets for cancer therapy and prevention. Curr Cancer Drug Targets 2013;13(5):558–79. [CrossRef]
47. Chen JF, Yan Q. The roles of epigenetics in cancer progression and metastasis. Biochem J 2021;478(17):3373–93. [CrossRef]
48. Lue JK, Prabhu SA, Liu Y, Gonzalez Y, Verma A, Mundi PS, et al. Precision targeting with EZH2 and HDAC inhibitors in epigenetically dysregulated lymphomas. Clin Cancer Res 2019;25(17):5271–83. [CrossRef]
49. Mondello P, Ansell SM, Nowakowski GS. Immune epigenetic crosstalk between malignant B cells and the tumor microenvironment in B cell lymphoma. Front Genet 2022;13:826594. [CrossRef]
50. Zhu Y, Wang Z, Li Y, Peng H, Liu J, Zhang J, et al. The role of CREBBP/EP300 and its therapeutic implications in hematological malignancies. Cancers 2023;15(4):1219. [CrossRef]
51. Kim KH, Roberts CWM. Targeting EZH2 in cancer. Nat Med 2016;22(2):128–34. [CrossRef]
52. Kim P, Zhao J, Lu P, Zhao Z. mutLBSgeneDB: Mutated ligand binding site gene database. Nucleic Acids Res 2017;45(D1):D256–63. [CrossRef]
53. Mullighan CG, Zhang J, Kasper LH, Lerach S, Payne-Turner D, Phillips LA, et al. CREBBP mutations in relapsed acute lymphoblastic leukaemia. Nature 2011;471(7337):235–9. [CrossRef]
54. Mondello P, Tadros S, Teater M, Fontan L, Chang AY, Jain N, et al. Selective inhibition of HDAC3 targets synthetic vulnerabilities and activates immune surveillance in lymphoma. Cancer Discov 2020;10(3):440–59. [CrossRef]

55. Pasqualucci L, Dominguez-Sola D, Chiarenza A, Fabbri G, Grunn A, Trifonov V, et al. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. Nature 2011;471(7337):189–95. [CrossRef]
56. Duex JE, Swain KE, Dancik GM, Paucek RD, Owens C, Churchill MEA, et al. Functional impact of chromatin remodeling gene mutations and predictive signature for therapeutic response in bladder cancer. Mol Cancer Res 2018;16(1):69–77. [CrossRef]
57. Huang YH, Cai K, Xu PP, Wang L, Huang CX, Fang Y, et al. CREBBP/EP300 mutations promoted tumor progression in diffuse large B-cell lymphoma through altering tumor-associated macrophage polarization via FBXW7-NOTCH-CCL2/CSF1 axis. Signal Transduct Target Ther 2021;6(1):10. [CrossRef]
58. Huang J, Gou H, Yao J, Yi K, Jin Z, Matsuoka M, et al. The noncanonical role of EZH2 in cancer. Cancer Sci 2021;112(4):1376–82. [CrossRef]
59. Majer CR, Jin L, Scott MP, Knutson SK, Kuntz KW, Keilhack H, et al. A687V EZH2 is a gain-of-function mutation found in lymphoma patients. FEBS Lett 2012;586(19):3448–51. [CrossRef]
60. Li B, Chng WJ. EZH2 abnormalities in lymphoid malignancies: Underlying mechanisms and therapeutic implications. J Hematol Oncol 2019;12(1):118. [CrossRef]
61. Morin RD, Johnson NA, Severson TM, Mungall AJ, An J, Goya R, et al. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nat Genet 2010;42(2):181–5.
62. Ribeiro ML, Reyes-Garau D, Armengol M, Fernández-Serrano M, Roué G. Recent advances in the targeting of epigenetic regulators in B-cell non-Hodgkin lymphoma. Front Genet 2019;10:986. [CrossRef]
63. Wang N, Ma T, Yu B. Targeting epigenetic regulators to overcome drug resistance in cancers. Signal Transduct Target Ther 2023;8(1):69. [CrossRef]

Appendix 1. 68 genes selected for this study mainly involved in epigenetic regulations in different cancers

Gene	Name	Percentage	Group
DNMT1			DNA Methyltransferase
DNMT3A	DNA Methyltrasferase 3 Alpha	4.6	
DNMT3B			
DNMT3L			
HDAC1			HDAC
HDAC2			
HDAC3			
HDAC8			
HDAC4			
HDAC5			
HDAC7			
HDAC9			
HDAC6			
HDAC10			
HDAC11			
CREBBP	cAMP-responsible element binding protein	11.5	HAT
EP300	Histone acetyltransferase p300	5.7	
TAF1	Transcription initiation factor TFIID subunit 1	3.4	
TET1			TET
TET2	Ten element translocation 2	6.2	
TET3			
EZH2	Enhencer of zeste homolog 2	6.1	Polycomb group proteins
ARID1A	AT-rich interactive domain 1A gene	9.6	SWI/SNF complex
ARID1B	AT-rich interactive domain 1B gene	8.3	
PBRM1			
SMARCA4			
SMARCB1			
ARID1			
ARID2			
ARID3			
ARID4			
ARID5 (B)	AT-rich interactive domain 5B	3.2	
JARID1			
JARID2			
SETD1A			KMTs
SETD1B	Histone lysine methyltransferase 1B	8.3	
SETD2	Histone lysine methyltransferase 2	6.2	
SETD5	Histone lysine methyltransferase 5	3.1	
KMT2A			
KMT2C	Lysine methyltransferase 2C	5.6	
KMT2D	Lysine methyltransferase 2D	24.7	
KMT3A			
KMT3B			
KMT6A			
KMT8			
KMT2A			
KMT1C			
KMT1E			
KMT2E			
KMT3C			

Appendix 1. Cont.

Gene	Name	Percentage	Group
KMT3E			
KDM1A			KDMs
KDM3A			
KDM5A			
KDM2B			
KDM4B			
KDM5B			
KDM6B			
KDM3B			
KDM5C			
KDM6A			
SIRT1			SIRTUINS
SIRT2			
SIRT3			
SIRT4			
SIRT5			
SIRT6			
SIRT7			