RESEARCH ARTICLE

Editorial Process: Submission:09/12/2025 Acceptance:03/27/2026 Published:04/07/2026

DNA Alterations in the Upstream Region of Exon 1 of *OSBPL10* in Northern Thai Patients with Diffuse Large B-Cell Lymphoma

Phuttirak Yimpak[1], Adisak Tantiworawit[2], Thanawat Rattanathammethee[2], Teerada Daroontum[3], Sirinda Aungsuchawan[1], Kanokkan Bumroongkit[1]*

Abstract

Background: Lymphoma is the most common hematologic malignancy in Thailand, with diffuse large B-cell lymphoma (DLBCL) being the predominant subtype. Early prognostic indicators are essential for guiding clinical decisions. Genetic alterations, particularly in regulatory regions, may serve as potential biomarkers. This study investigated sequence alterations upstream of exon 1 of the *OSBPL10* gene and their clinical relevance in a Northern Thai DLBCL population. **Methods:** Eighty-five DLBCL patients residing in Northern Thailand were genotyped to assess sequence alterations in the upstream regulatory region of *OSBPL10*. **Results:** Two previously reported single nucleotide polymorphisms (SNPs)—rs76150980 (G>C) and rs62244394 (C>G) and two additional alterations not listed in public SNP databases were identified. The heterozygous GC genotype of rs76150980 and the CG genotype of rs62244394 were significantly associated with a reduced risk of extranodal involvement (OR = 0.229 and 0.196, respectively). While no significant difference in overall survival (OS) was observed, individuals carrying the CG/GG genotypes of rs62244394 had significantly longer progression-free survival (PFS) compared to homozygous CC individuals (median PFS: 39.7 months, 95% CI: 30.4–49.0 vs. 17.3 months, 95% CI: 12.4–22.2; $p = 0.031$). Additionally, two upstream substitutions—C>T at chr3:31,981,252 and G>T at chr3:31,981,259—showed genotype frequencies of CC (87.1%), CT (3.5%), TT (9.4%) and GG (89.4%), GT (2.4%), TT (8.2%), respectively. At the allele level, the T allele at both positions was associated with increased International Prognostic Index risk (OR = 3.615 and 4.717), while the T allele at position 31,981,259 was associated with lower odds of the aggressive non-GCB subtype (OR = 0.324). No significant OS or PFS differences were observed for these variants. **Conclusion:** Variations upstream of *OSBPL10* were associated with extranodal involvement and selected clinical parameters in DLBCL, highlighting the prognostic potential of regulatory-region SNPs and the importance of population-specific genetic diversity in lymphoma research.

Keywords: *OSBPL10* gene- Single nucleotide polymorphism- Diffuse large B-cell lymphoma

Asian Pac J Cancer Prev, **27 (4)**, 1389-1397

Introduction

Lymphoma, the most common hematologic malignancy in Thailand, has an age-standardized incidence rate of about 6.3 per 100,000, according to national registry data [1]. One of the most aggressive forms of lymphoma is diffuse large B-cell lymphoma (DLBCL), which is characterized as a heterogenous disease [2]. Although lymphoma as a whole has a relatively high incidence, prognosis differs across subtypes. DLBCL can achieve high remission rates following appropriate treatment, whereas outcomes for other lymphoma subtypes are more variable. This underscores the importance of effective treatment planning from the time of diagnosis. Over the past decade, significant strides have been made in understanding and managing lymphoma. Advancements in molecular genetics have notably improved diagnostic accuracy, often contributing to better patient prognoses.

OSBPL10 (Oxysterol Binding Protein Like 10) gene encodes a member of the oxysterol-binding protein (OSBP) family [3]. It functions as an essential mediator of intracellular lipid transport and cholesterol homeostasis, and plays a regulatory role in cell signaling pathways. Dysregulation of *OSBPL10* protein and other OSBPL family members has been reported to contribute to cancer progression [4, 5]. In 2018, Dobashi and colleagues identified recurrent mutations in *OSBPL10*, primarily observed in exon 1, suggesting these as prognostic markers associated with favorable outcomes. Many of these mutations occur within RGYW/WRCY motifs, regions known to be susceptible to somatic hypermutation (SHM) [6]. However, the functional role of *OSBPL10* and its mutations in DLBCL remains unclear and requires further investigation.

*[1]Department of Anatomy, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand. [2]Division of Hematology, Department of Internal Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand. [3]Department of Pathology, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand. *For Correspondence: kanokkan.bumr@cmu.ac.th*

Non-coding regulatory mutations occur broadly across cancer types and can modulate gene expression through their effects on key regulatory elements [7]. Variants in upstream regulatory regions, including promoters and 5'untranslated regions (5' UTRs), are well-documented to affect gene expression and disease susceptibility. For example, single-nucleotide polymorphisms (SNPs) in promoter regions can alter transcription factor binding, chromatin accessibility, and epigenetic modifications, thereby affecting gene transcription and cancer risk [8]. Functional studies have also shown that 5' UTR variants can modulate mRNA stability or promote gene silencing, contributing to disease pathogenesis [9].

Importantly, the study of Thai population genetic structure by Wangkumhang and colleagues revealed strong subpopulation stratification within Thailand, identifying four genetically distinct Thai subpopulations. This population structure has key implications for genetic association studies, including in cancer, because regional ancestry may confound or interact with disease-associated variants [10]. Based on these findings, the present study aimed to investigate genetic alterations within the upstream regulatory region of *OSBPL10* (promoter and 5' UTR adjacent to exon 1) in DLBCL patients residing in Northern Thailand. The study also explored the correlation between these genetic variants and clinical parameters based on the International Prognostic Index (IPI), including age, Ann Arbor stage, lactate dehydrogenase (LDH) levels, Eastern Cooperative Oncology Group (ECOG) performance status, and extranodal involvement as well as overall survival (OS) and progression-free survival (PFS).

Given the significant genetic variation observed across different ethnic groups, studying *OSBPL10* variants in the Northern Thai population may yield localized insights that are often overlooked in broader genomic datasets. Therefore, this study aimed to characterize genetic variation in the upstream regulatory region of *OSBPL10* in Northern Thai DLBCL patients and examine the possible associations with clinical features. The findings provide preliminary, population-specific insights that may support future validation studies in larger and more diverse cohorts, and contribute toward developing more precise prognostic approaches and therapeutic strategies for DLBCL.

Materials and Methods

Sample collection

A total of 85 formalin-fixed, paraffin-embedded (FFPE) tissue samples were retrospectively collected from patients residing in Northern Thailand who were diagnosed with DLBCL between January 2018 and December 2019 at Maharaj Nakorn Chiang Mai Hospital in Chiang Mai Province, Thailand.

Collection of clinical data and biomarkers

Clinical information and DLBCL subtypes were retrieved from the lymphoma registry and electronic medical records at Maharaj Nakorn Chiang Mai Hospital (Chiang Mai, Thailand). Collected data included patient age, LDH levels, Ann Arbor staging, extranodal involvement, and ECOG performance status, which were used to calculate the IPI risk score. The data used for OS and PFS analysis included the following definitions: Overall survival (OS) was defined as the time from the date of specimen collection to the date of death from any cause or the date of the last follow-up. Progression-free survival (PFS) was defined as the time from specimen collection to tumor progression, death from any cause, or the last follow-up in cases where the disease remained at least in partial remission.

Samples were subsequently stratified based on multiple clinical and pathological parameters, including cell of origin (GCB vs. non-GCB), LDH levels (normal vs. elevated), extranodal involvement (nodal vs. extranodal), Ann Arbor stage (limited stage (I–II) vs. advanced stage (III–IV)), IPI risk score (low (0-3) vs. high (≥ 4)) to facilitate genotypic analysis in relation to clinical parameters.

DNA extraction and sequencing

The NucleoSpin® DNA FFPE XS kit (Macherey-Nagel, Germany) was used for DNA extraction from FFPE tissue samples in accordance with the manufacturer's protocol. Primers specific to the *OSBPL10* gene were designed using the Primer-BLAST tool provided by the NCBI, targeting the promoter region and 5' UTR of exon 1 (chr3: 31,981,074 to 31,981,302). The sequence of PCR primers was as follows: forward primer, CATTTCCCGGGGATTTGGAG; and reverse primer, TCGCCCTCCTGCTCTCTGG. DNA amplification was performed via standard polymerase chain reaction (PCR) using Platinum™ SuperFi™ DNA Polymerase (Invitrogen, USA). The thermal cycling conditions were as follows: initial denaturation at 98 °C for 30 seconds, followed by 35 cycles of denaturation at 98 °C for 10 seconds, annealing at 59–60 °C for 10 seconds, and extension at 72 °C for 30 seconds, with a final extension at 72 °C for 5 minutes. PCR products were assessed using 2% agarose gel electrophoresis. In cases where amplification yielded weak bands, re-amplification was performed under identical conditions to increase DNA yield. Amplified products were purified using the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel, Germany). The sequence was carried out either at Macrogen (Seoul, Korea) or the Medical Science Research Equipment Center (MSREC), Faculty of Medicine, Chiang Mai University. Sequence data were analyzed using SeqScape version 4.0 (Applied Biosystems, Foster City, CA, USA) by aligning electropherograms to reference sequences from GenBank (NC_000003.12, NM_017784.5). A mixed base detection threshold of 15% was applied, and heterozygous variants were automatically identified by the software.

Statistical analysis

The genotypic and allelic frequencies of the *OSBPL10* gene were determined and expressed as percentages. Associations between individual variants and clinical parameters were evaluated using binary logistic regression to determine odds ratios (ORs) with 95% confidence intervals (CIs). Kaplan–Meier survival analysis was

performed to assess differences in OS and PFS based on *OSBPL10* variant status. Survival curves for OS and PFS were compared using the log-rank test. Statistical analyses were conducted using IBM SPSS Statistics version 23, with a p-value of <0.05 considered indicative of statistical significance.

Results

Frequencies of OSBPL10 Variants in the Northern Thai Population and Their Global Comparisons

This study identified two previously reported SNPs: rs76150980 and rs62244394. In the Northern Thai population, the genotype frequencies for rs76150980 were 41.2% for GG, 27.0% for CG, and 31.8% for CC, with allele frequencies of 54.7% for G and 45.3% for C. For rs62244394, the genotype distribution was 38.8% for CC, 22.4% for CG, and 38.8% for GG, with allele frequencies of 50.0% for both C and G.

The allele frequencies for these SNPs in the Northern Thai population were compared with data from five populations in the 1000 Genomes Project Phase 3 (30X dataset): American (AMR), South Asian (SAS), European (EUR), African (AFR), and East Asian (EAS). The Northern Thai population exhibited allele frequencies that closely resembled the EAS group. In contrast, marked differences were observed when compared with the AMR, SAS, EUR, and AFR populations for both rs76150980 (Figure 1) and rs62244394 (Figure 2). In the AMR, SAS, EUR, and AFR populations, the G allele was the most frequent for rs76150980, whereas the EAS population had a nearly equal distribution of G and C alleles (Figure 1). For rs62244394, the EAS group showed a near-equal distribution of C and G alleles, while the AMR, SAS, EUR, and AFR predominantly carried the C allele (Figure 2).

Clinical relevance of rs76150980 and rs62244394 in Northern Thai individuals with DLBCL

We analyzed the association of the rs76150980 variant with clinical parameters in patients with DLBCL (Table 1). The heterozygous GC genotype was significantly associated with a reduced risk of extranodal involvement when compared to the GG homozygous genotype, which corresponds to the major allele and served as the reference

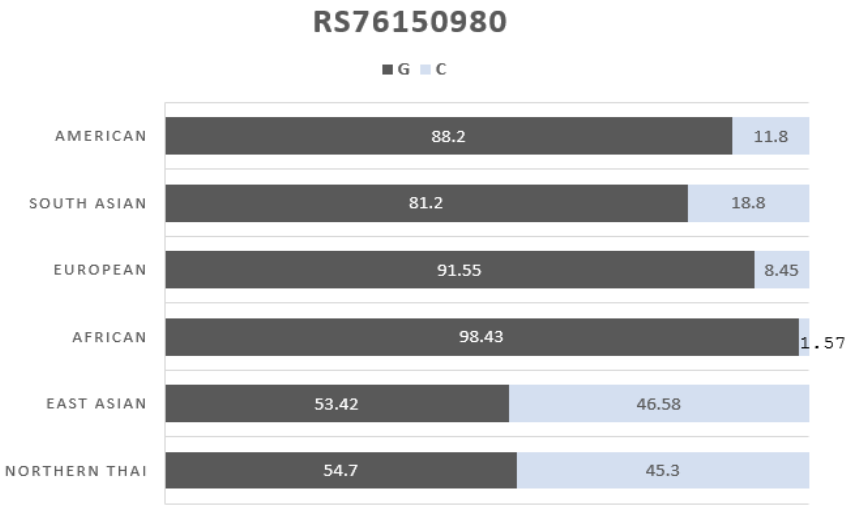

Figure 1. Allele Frequency of rs76150980 in Northern Thai Population Compared to Global Populations

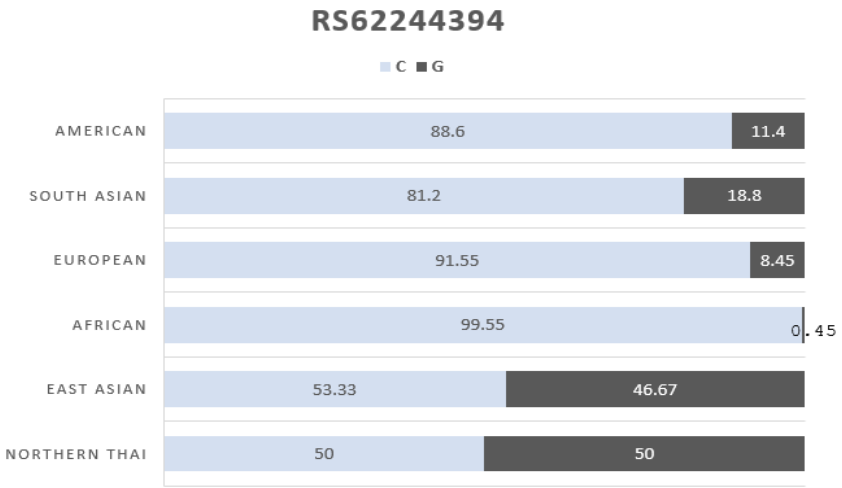

Figure 2. Allele frequency of rs62244394 in Northern Thai Population Compared to Global Populations

Table 1. Association of SNPs of rs76150980 and Clinical Data of DLBCL Patients

SNP	rs76150980 (n=85)					
	Genotypes				Alleles	
	GG	GC	CC	GC +CC	G	C
Cell of origin						
p-value	1.000	0.730	0.503	0.535	1.000	0.414
OR		1.246	1.523	1.385		1.358
95% CI		0.358 - 4.339	0.444 - 5.221	0.496 - 3.867		0.651 - 2.832
LDH						
p-value	1.000	0.969	0.698	0.825	1.000	0.655
OR		0.978	1.239	1.109		1.158
95% CI		0.324 - 2.957	0.420 - 3.654	0.443 - 2.773		0.608 - 2.207
Extranodal involvement						
p-value	1.000	0.013*	0.141	0.025*	1.000	0.064
OR		0.229	0.425	0.318		0.546
95% CI		0.072 - 0.734	0.136 - 1.327	0.117 - 0.864		0.288 - 1.036
Ann Arbor stage						
p-value	1.000	0.481	0.822	0.593	1.000	0.748
OR		0.678	0.887	0.783		0.903
95% CI		0.230 - 1.998	0.311 - 2.528	0.319 -1.922		0.484 - 1.683
IPI Risk Score						
p-value	1.000	0.117	0.874	0.324	1.000	0.748
OR		0.420	0.917	0.633		0.903
95% CI		0.142 - 1.244	0.314 - 2.678	0.255 - 1.569		0.484 - 1.683

* $p < 0.05$

(OR = 0.229; 95% CI: 0.072–0.734; p = 0.013).

For the rs62244394 variant, similar associations were observed (Table 2). The heterozygous CG genotype was significantly associated with a lower risk of extranodal involvement compared to the CC homozygous genotype, which represents the major allele and served as the reference (OR = 0.196; 95% CI: 0.057–0.673; p = 0.010).

Survival analysis

For rs76150980, Kaplan–Meier survival analysis showed no significant difference in median OS between individuals with the reference genotype (GG; median OS: 29.9 months, 95% CI: 2.1–57.6) and those with the variant genotypes (GC or CC; median OS not reached), as determined by the log-rank test (Figure 3a). Similarly, rs76150980 was not significantly associated with median PFS (GG: 18.5 months, 95% CI: 9.2–27.9; GC/CC: 37.9 months, 95% CI: 27.6–48.3) (Figure 3b). For rs62244394, Kaplan–Meier analysis revealed no significant difference in median OS between individuals with the reference genotype (CC; median OS: 27.9 months, 95% CI: 5.8–49.9) and those with the variant genotypes (CG or GG;

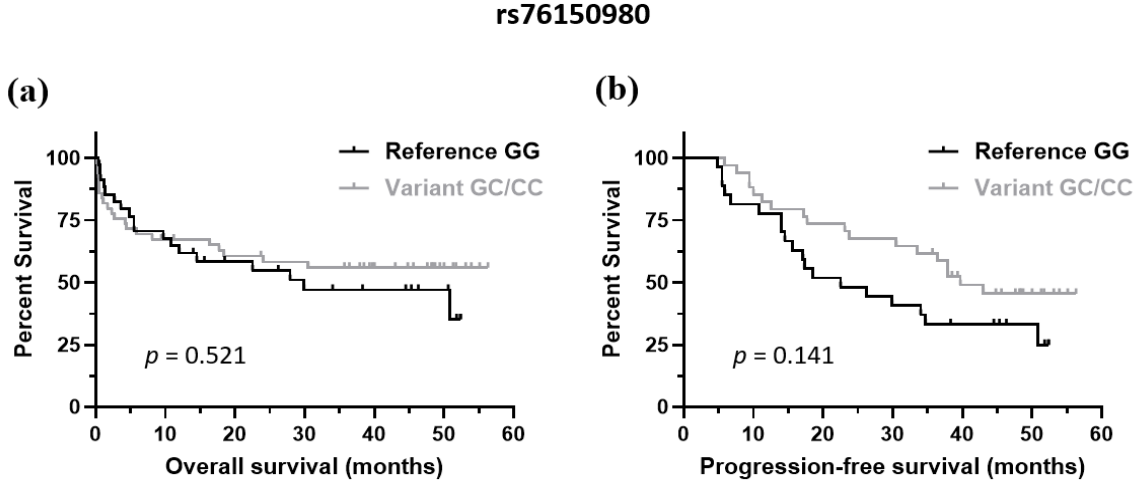

Figure 3. Survival Analysis of DLBCL Patients by rs76150980 SNPs Using Kaplan–Meier Plots

Table 2. Association of SNPs of rs62244394 and Clinical Data of DLCBL Patients

SNP	rs62244394 (n=85)					
	Genotypes				Alleles	
	CC	CG	GG	CG+GG	C	G
Cell of origin						
p-value	1.000	0.349	0.566	0.388	1.000	0.462
OR		2.000	1.393	1.575		1.312
95% CI		0.468 - 8.541	0.449 - 4.324	0.562 - 4.415		0.636 - 2.710
LDH						
p-value	1.000	0.897	1.000	0.951	1.000	1.000
OR		1.083	1.000	1.029		1.000
95% CI		0.324 - 3.626	0.359 - 2.783	0.407 - 2.602		0.527 - 1.896
Extranodal involvement						
p-value	1.000	0.010*	0.272	0.049*	1.000	0.197
OR		0.196	0.538	0.367		0.657
95% CI		0.057 - 0.673	0.178 - 1.625	0.135 - 0.998		0.347 - 1.244
Ann Arbor stage						
p-value	1.000	0.798	0.447	0.514	1.000	0.343
OR		0.857	0.679	0.738		0.740
95% CI		0.263 - 2.790	0.250 - 1.845	0.297 - 1.836		0.397 - 1.379
IPI Risk Score						
p-value	1.000	0.115	0.602	0.268	1.000	0.527
OR		0.391	0.761	0.593		0.818
95% CI		0.122 - 1.257	0.273 - 2.124	0.235 - 1.494		0.439 - 1.523

* $p < 0.05$

median OS not reached) (Figure 4a). However, individuals carrying the CG or GG genotypes had significantly longer median PFS compared to those with the homozygous CC genotype (CG/GG: 39.7 months, 95% CI: 30.4–49.0 vs. CC: 17.3 months, 95% CI: 12.4–22.2; p = 0.031, log-rank test) (Figure 4b).

Genotype Distribution of Nucleotide Alterations in the Promoter Region of OSBPL10 Among Northern Thai DLBCL Patients

Additionally, we detected a C-to-T nucleotide substitution in the promoter region adjacent to exon 1 of the *OSBPL10* gene on chromosome 3, at position 31,981,252 (GRCh38.p14). The observed genotype frequencies at this site were 87.1% for CC, 3.5% for CT, and 9.4% for TT (Figure 5a). Another variant, a G-to-T substitution, was identified within the same promoter region at position 31,981,259 (GRCh38.p14). The genotype frequencies at this locus were 89.4% for GG, 2.4% for GT, and 8.2% for TT (Figure 5b).

We further analyzed the association of these variants with clinical parameters (Table 3 and 4). For the C-to-T

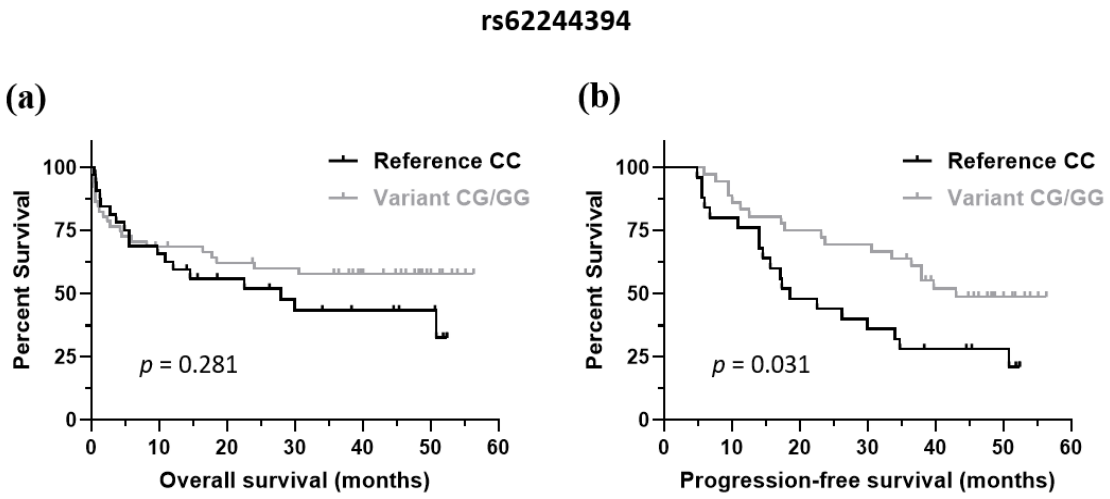

Figure 4. Survival Analysis of DLBCL Patients by rs62244394 SNPs Using Kaplan–Meier Plots

Table 3. Association of SNPs of C-to-T Substitution at Chromosome 3:31,981,252 and Clinical Data of DLCBL Patients

	C-to-T substitution at chr 3:31,981,252 (n=85)					
	Genotypes				Alleles	
	CC	CT	TT	CT+TT	C	T
Cell of origin						
p-value	1.000	0.591	0.274	0.214	1.000	0.115
OR		0.508	0.424	0.445		0.443
95% CI		0.043 – 5.990	0.091 – 1.975	0.115 – 1.721		0.161 – 1.129
LDH						
p-value	1.000	0.274	0.246	0.669	1.000	0.249
OR		0.255	3.571	1.361		1.970
95% CI		0.022 – 2.951	0.416 – 30.659	0.332 – 5.582		0.622 – 6.239
Extranodal involvement						
p-value	1.000	0.999	0.523	0.246	1.000	0.210
OR		-	1.723	2.585		2.088
95% CI		-	0.325 – 9.144	0.520 – 12.850		0.660 – 6.607
Ann Arbor stage						
p-value	1.000	0.839	0.438	0.451	1.000	0.285
OR		1.289	1.933	1.719		1.796
95% CI		0.112 – 14.868	0.365 – 10.239	0.421 – 7.015		0.615 – 5.246
IPI Risk Score						
p-value	1.000	0.804	0.153	0.170	1.000	0.048*
OR		1.364	4.773	3.068		3.615
95% CI		0.118 – 15.722	0.588 – 40.812	0.619 – 15.211		1.010 – 12.939

* $p < 0.05$

Table 4. Association of SNPs of G-to-T substitution at Chromosome 3:31,981,259 and Clinical Data of DLCBL Patients

	G-to-T substitution at chromosome 3:31,981,259 (n=85)					
	Genotypes				Alleles	
	GG	GT	TT	GT+TT	G	T
Cell of origin						
p-value	1.000	0.331	0.172	0.106	1.000	0.038*
OR		0.246	0.328	0.307		0.324
95% CI		0.015 – 4.162	0.066 – 1.624	0.073 – 1.286		0.112 – 0.938
LDH						
p-value	1.000	0.649	0.304	0.475	1.000	0.215
OR		0.520	3.120	1.820		2.274
95% CI		0.031 – 8.655	0.356 – 27.309	0.353 – 9.395		0.621 – 8.332
Extranodal involvement						
p-value	1.000	0.999	0.713	0.433	1.000	0.423
OR		-	1.378	1.929		1.620
95% CI		-	0.250 – 7.584	0.374 – 9.944		0.498 – 5.267
Ann Arbor stage						
p-value	1.000	0.999	0.574	0.323	1.000	0.279
OR		-	1.630	2.283		1.915
95% CI		-	0.297 – 8.953	0.444 – 11.737		0.590 – 6.213
IPI Risk Score						
p-value	1.000	0.999	0.199	0.116	1.000	0.045*
OR		-	4.133	5.511		4.717
95% CI		-	0.474 – 36.052	0.656 – 46.310		1.036 – 21.487

* $p < 0.05$

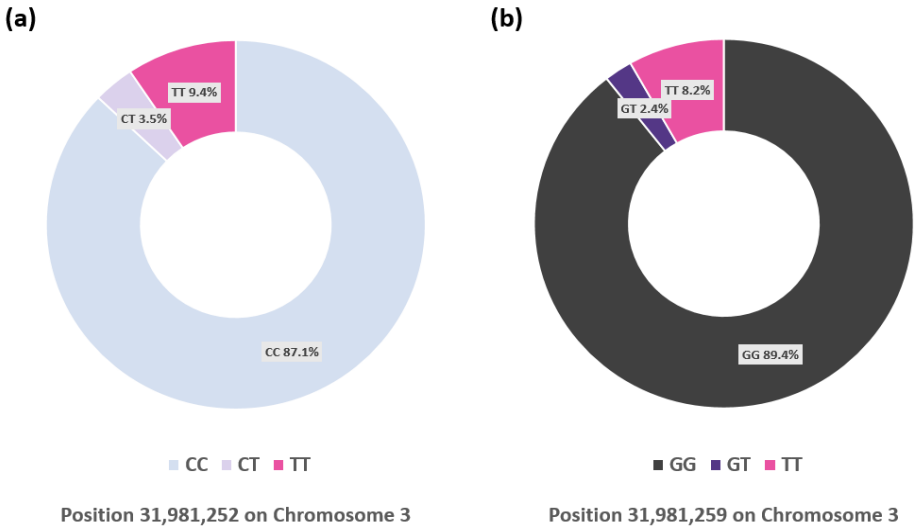

Figure 5. Genotype Frequency in the Section of Exon 1 at Position 31,981,252 (a) and Position 31,981,259 (b) on Chromosome 3 in these Northern Thai DLBCL patient

substitution at chromosome 3:31,981,252, genotype distributions were not significantly associated with IPI risk categories; however, allele-level analysis showed that carriers of the T allele had a higher likelihood of elevated IPI scores compared with those carrying the major allele (OR = 3.615; 95% CI: 1.010–12.939; p = 0.048). A similar pattern was observed for the G-to-T substitution at chromosome 3:31,981,259, where the T allele was also associated with increased IPI risk (OR = 4.717; 95% CI: 1.036–21.487; p = 0.045). For cell-of-origin classification (GCB vs. non-GCB), genotype-level comparisons were not significant; however, allele analysis at position 31,981,259 revealed that the T allele was associated with lower odds of the more aggressive non-GCB subtype (OR = 0.324; 95% CI: 0.112–0.938; p = 0.038). Survival analysis of these two SNPs, which has not been previously reported, revealed no significant differences in OS or PFS between individuals with the reference genotypes (CC at chr3:31,981,252 and GG at chr3:31,981,259) and those carrying the variant genotypes (CT, GT, or TT).

Discussion

This study focused on investigating genetic alterations within the upstream regulatory region of *OSBPL10*, with the objective of exploring their potential clinical relevance in Northern Thai patients diagnosed with DLBCL. Among the 85 cases analyzed, two reported SNPs, rs76150980 and rs62244394 and two previously unreported, were identified within this regulatory region.

The allele frequencies of rs76150980 and rs62244394 in the Northern Thai cohort closely resemble patterns observed in East Asian populations but differ from those in other global populations [11, 12]. East Asian populations include the major ethnic groups such as the Chinese, Japanese, and Korean [13, 14]. In Northern Thailand, Tai languages population constitute the predominant linguistic group [15]. Previous studies suggest that historical migration and admixture events, including the southward movement of Tai-Kadai–speaking groups from southern China, have influenced the genetic and demographic landscape of the region. Y-chromosome and mitochondrial DNA analyses indicate that Northern Thai populations share haplogroups common in East Asia, supporting a shared ancestry [16]. Genome-wide analyses further show that Northern Thai groups cluster more closely with southern Chinese Dai populations [17]. Together, these findings suggest that gene flow from southern China, combined with local admixture, has shaped the Northern Thai genetic landscape, producing a population closely linked to East Asian ancestries. This demographic history may explain why the *OSBPL10* alterations examined in this study are more frequent in Northern Thai and East Asian populations than in others.

We further conducted an in-depth analysis of genotypes in relation to clinical parameters, categorizing samples based on factors such as cell of origin, LDH levels, extranodal involvement, disease stage (Ann Arbor classification), and IPI risk score. Our results revealed that the heterozygous GC variant of rs76150980 and the CG variant of rs62244394 were significantly associated with a reduced risk of extranodal involvement compared to the homozygous wild-type genotypes. This observation supports the heterozygote advantage hypothesis, which suggests that having two different alleles may enable more efficient regulation of key biological processes.

Heterozygote advantage may help explain our findings regarding *OSBPL10* variants. In other disease settings, heterozygosity at some loci can be protective; for example, individuals who are human leukocyte antigen (HLA) heterozygous have been associated with a lower risk of colorectal cancer [18]. Similarly, a pan-cancer analysis from the UK Biobank found that HLA diversity, including heterozygosity, is associated with reduced risk of non-Hodgkin lymphoma, including DLBCL [19]. While the precise mechanism linking *OSBPL10* heterozygosity to lower extranodal involvement in DLBCL remains unclear, our findings suggest that genetic diversity at this locus may confer a slight protective effect, highlighting the need for additional functional studies to elucidate the consequences

of these variants.

Alterations in the promoter and 5' UTR regions, such as those identified in our study, can significantly influence gene expression by affecting the binding of transcription factors [20]. These changes may enhance or suppress transcriptional activity, impacting mRNA stability and translation efficiency [21], thus altering gene expression levels. We propose that the heterozygous variants of *OSBPL10* may have a balanced regulatory effect, reducing extranodal involvement risk.

In the survival analysis, neither rs76150980 nor rs62244394 had a substantial impact on overall survival in the studied population. However, the presence of the altered genotype (CG/GG) of rs62244394 SNP may be associated with improved prognosis in terms of disease progression. These findings suggest the potential clinical relevance of these variants and warrant further investigation in larger, more diverse cohorts.

Additionally, the alterations at positions 31,981,252 and 31,981,259 on chromosome 3, located within the promoter region of the *OSBPL10* gene, have not been reported previously in the 1000 Genomes Project Phase 3: 30X dataset. These alterations warrant further investigation from multiple perspectives. In our study, allele-level analyses suggested that the minor alleles at these two positions were associated with higher IPI risk and, in the case of position 31,981,259, with lower odds of the aggressive non-GCB subtype. These findings indicate that even subtle regulatory alterations may contribute to inter-individual variability in DLBCL clinical features. While the precise mechanisms remain unclear, it is possible that promoter-region alterations modulate *OSBPL10* expression in B cells. Functional validation and studies in larger, multi-ethnic cohorts will be necessary to confirm these effects and clarify the role of *OSBPL10* alterations in DLBCL pathogenesis.

In summary, our study investigated the clinical significance of the heterozygous CG variant of the *OSBPL10* gene in Northern Thai DLBCL patients, focusing on SNPs rs76150980 and rs62244394. These variants were associated with a lower risk of extranodal involvement. Moreover, alterations in the promoter region of *OSBPL10* at positions 31,981,252 and 31,981,259, previously unreported, may influence DLBCL clinical features, including IPI risk and subtype. These results suggest that even subtle regulatory alterations in *OSBPL10* may contribute to inter-individual variability in DLBCL clinical presentation. Our findings provide a basis for further research into the molecular mechanisms of DLBCL and may inform future studies exploring potential therapeutic targets. Moreover, these findings can serve as valuable information for the Northern Thai population, which has a distinct genetic background from other Thai subpopulations. Due to the small sample size in this study, further investigations in larger and more diverse cohorts are needed to assess the broader impact of these variants across populations.

Author Contribution Statement

Conceptualization: Yimpak P, Bumroongkit K, Tantiworawit A, Rattanathammethee T, Daroontum T, Aungsuchawan S. Data curation: Yimpak P, Bumroongkit K, Tantiworawit A, Rattanathammethee T, Daroontum T, Aungsuchawan S. Formal analysis: Yimpak P, Bumroongkit K, Tantiworawit A, Rattanathammethee T, Daroontum T. Funding acquisition: Yimpak P, Bumroongkit K. Investigation: Yimpak P, Bumroongkit K, Daroontum T, Rattanathammethee T, Tantiworawit A, Daroontum T. Methodology: Yimpak P, Bumroongkit K. Project administration: Yimpak P, Bumroongkit K, Daroontum T. Resources: Daroontum T, Rattanathammethee T, Tantiworawit A. Software: Yimpak P. Supervision: Bumroongkit K. Validation: Yimpak P, Bumroongkit K, Daroontum T, Tantiworawit A, Rattanathammethee T, Daroontum T. Visualization: Yimpak P. Writing – original draft: Yimpak P. Writing – review & editing: Yimpak P, Bumroongkit K, Tantiworawit A, Rattanathammethee T, Daroontum T, Aungsuchawan S.

Acknowledgements

General

We would like to express our heartfelt gratitude to Mrs. Rochana Phuackchantuck for her assistance with statistical consultations.

Funding Statement

This study was supported by the Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand [grant number ANA-2563-07765, No. 160/2564].

Scientific Body approved / student thesis

This study is part of a doctor of Philosophy (Ph.D.) dissertation.

Availability of data

The data are not publicly available due to patient privacy concerns but are available from the corresponding author upon reasonable request and with permission from the ethics committee.

Ethical Declaration

This research was granted an ethical exemption from the requirement for written consent by the Ethics and Research Committee of the Faculty of Medicine, Chiang Mai University [Certificate of approval No. 248/2021, Study code ANA-2563-07765].

Conflict of Interest

The authors declare no conflict of interest.

References

1. Thanasitthichai S, Ingsirorat R, Chairat C, Chiawiriyabunya I, Wongsena M, Laowahutanont P, et al. Cancer in thailand vol.Xi 2019-2021 [internet]. Bangkok: National cancer institute; 2025 [cited 2025 nov 20].
2. Green TM, Young KH, Visco C, Xu-Monette ZY, Orazi A, Go RS, et al. Immunohistochemical double-hit score is a strong predictor of outcome in patients with diffuse large b-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and

prednisone. J Clin Oncol. 2012;30(28):3460-7. https://doi.org/10.1200/jco.2011.41.4342.
3. Genecards. The human gene database. *OSBPL10* gene - oxysterol binding protein like 10 [internet]. Israel: Weizmann institute of science; c1996-2025 [2025 nov 14; cited 2025 nov 19]. Available from: Https://www.Genecards.Org/cgi-bin/carddisp.Pl?Gene=osbpl10.
4. Du X, Turner N, Yang H. The role of oxysterol-binding protein and its related proteins in cancer. Semin Cell Dev Biol. 2018;81:149-53. https://doi.org/10.1016/j.semcdb.2017.07.017.
5. Tian K, Ying Y, Huang J, Wu H, Wei C, Li L, et al. The expression, immune infiltration, prognosis, and experimental validation of OSBPL family genes in liver cancer. BMC Cancer. 2023;23(1):244. https://doi.org/10.1186/s12885-023-10713-9.
6. Dobashi A, Togashi Y, Tanaka N, Yokoyama M, Tsuyama N, Baba S, et al. *TP53* and *OSBPL10* alterations in diffuse large b-cell lymphoma: Prognostic markers identified via exome analysis of cases with extreme prognosis. Oncotarget. 2018;9(28):19555-68. https://doi.org/10.18632/oncotarget.24656.
7. Weinhold N, Jacobsen A, Schultz N, Sander C, Lee W. Genome-wide analysis of noncoding regulatory mutations in cancer. Nat Genet. 2014;46(11):1160-5. https://doi.org/10.1038/ng.3101.
8. Deng N, Zhou H, Fan H, Yuan Y. Single nucleotide polymorphisms and cancer susceptibility. Oncotarget. 2017;8(66):110635-49. https://doi.org/10.18632/oncotarget.22372.
9. Evans DGR, van Veen EM, Byers HJ, Wallace AJ, Ellingford JM, Beaman G, et al. A dominantly inherited 5' UTR variant causing methylation-associated silencing of *BRCA1* as a cause of breast and ovarian cancer. Am J Hum Genet. 2018;103(2):213-20. https://doi.org/10.1016/j.ajhg.2018.07.002.
10. Wangkumhang P, Shaw PJ, Chaichoompu K, Ngamphiw C, Assawamakin A, Nuinoon M, et al. Insight into the peopling of mainland southeast asia from thai population genetic structure. PLoS One. 2013;8(11):e79522. https://doi.org/10.1371/journal.pone.0079522.
11. National library of medicine, national center for biotechnology information. dbSNP rs76150980 [internet]. Bethesda: National library of medicine; c2024 [updated 2024 sep 3; cited 2025 nov 20]. Available from: Https://www.Ncbi.Nlm.Nih.Gov/snp/rs76150980.
12. National library of medicine, national center for biotechnology information. dbSNP rs62244394 [internet]. Bethesda: National library of medicine; c2024 [updated 2024 sep 3; cited 2025 nov 20]. Available from: Https://www.Ncbi.Nlm.Nih.Gov/snp/rs62244394.
13. Wang Y, Lu D, Chung YJ, Xu S. Genetic structure, divergence and admixture of han chinese, japanese and korean populations. Hereditas. 2018;155:19. https://doi.org/10.1186/s41065-018-0057-5.
14. Pan Z, Xu S. Population genomics of east asian ethnic groups. Hereditas. 2020;157(1):49. https://doi.org/10.1186/s41065-020-00162-w.
15. Kutanan W, Kampuansai J, Fuselli S, Nakbunlung S, Seielstad M, Bertorelle G, et al. Genetic structure of the mon-khmer speaking groups and their affinity to the neighbouring tai populations in northern thailand. BMC Genet. 2011;12:56. https://doi.org/10.1186/1471-2156-12-56.
16. Brunelli A, Kampuansai J, Seielstad M, Lomthaisong K, Kangwanpong D, Ghirotto S, et al. Y chromosomal evidence on the origin of northern thai people. PLoS One. 2017;12(7):e0181935. https://doi.org/10.1371/journal.pone.0181935.
17. Kutanan W, Liu D, Kampuansai J, Srikummool M, Srithawong S, Shoocongdej R, et al. Reconstructing the human genetic history of mainland southeast asia: Insights from genome-wide data from thailand and laos. Mol Biol Evol. 2021;38(8):3459-77. https://doi.org/10.1093/molbev/msab124.
18. Tsai YY, Qu C, Bonner JD, Sanz-Pamplona R, Lindsey SS, Melas M, et al. Heterozygote advantage at HLA class i and ii loci and reduced risk of colorectal cancer. Front Immunol. 2023;14:1268117. https://doi.org/10.3389/fimmu.2023.1268117.
19. Wang QL, Wang TM, Deng CM, Zhang WL, He YQ, Xue WQ, et al. Association of HLA diversity with the risk of 25 cancers in the uk biobank. EBioMedicine. 2023;92:104588. https://doi.org/10.1016/j.ebiom.2023.104588.
20. Guo Y, Jamison DC. The distribution of SNPs in human gene regulatory regions. BMC Genomics. 2005;6:140. https://doi.org/10.1186/1471-2164-6-140.
21. Mustafa A, Shabbir M, Badshah Y, Khan K, Abid F, Trembley JH, et al. Genetic polymorphism in untranslated regions of PRKCZ influences mRNA structure, stability and binding sites. BMC Cancer. 2024;24(1):1147. https://doi.org/10.1186/s12885-024-12900-8.