The European Journal of Biomedical and Life Sciences 2024, No 1–2
ISSN 2310-5674

PREMIER Publishing
ppublishing.org

Section 2. Life sciences

DOI:10.29013/EJBLS-24-1.2-24-35

UNVEILING THE COMPLEX SNP LANDSCAPE IN LYMPHOMA THROUGH COMPREHENSIVE ANALYSIS FOR FUTURE FUNCTIONAL INTERPRETATION AND THERAPEUTIC APPLICATIONS

Ethan Zhang [1]

1

Cite: *Ethan Zhang. (2024). Unveiling the Complex SNP Landscape in Lymphoma Through Comprehensive Analysis for Future Functional Interpretation and Therapeutic Applications. The European Journal of Biomedical and Life Sciences 2024, No 1–2* https://doi.org/10.29013/EJBLS-24-1.2-24-35

Abstract:

The successful management of lymphoma increasingly hinges on a deep understanding of its genomic underpinnings, especially for early detection and targeted treatment strategies. The current study focused on the identification and functional categorization of single nucleotide polymorphisms (SNPs) specifically in lymphoma patients. Utilizing whole-genome sequencing data sourced from the Sequence Read Archive, we employed a rigorous analytical pipeline involving sequence alignment against the Homo sapiens chromosome 4 reference genome, followed by indexing, and variant calling. Our investigation led to the discovery of 1072 SNPs, a significant portion of which remain functionally uncharacterized. Among the categorized SNPs, variants were identified that have potential implications for gene regulation, splicing, and protein function. Statistical analyses revealed a significant association between these SNPs and the lymphoma cohort. Our findings offer nuanced insights into the complex genetic landscape of lymphoma and serve as a foundational platform for future research aimed at functional characterization and potential clinical applications. These identified SNPs could serve as potential biomarkers and contribute to the development of more effective diagnostic and therapeutic strategies for lymphoma.

Keywords: *Single Nucleotide Polymorphisms (SNPs), Lymphoma, TET2*

Introduction

Lymphoma

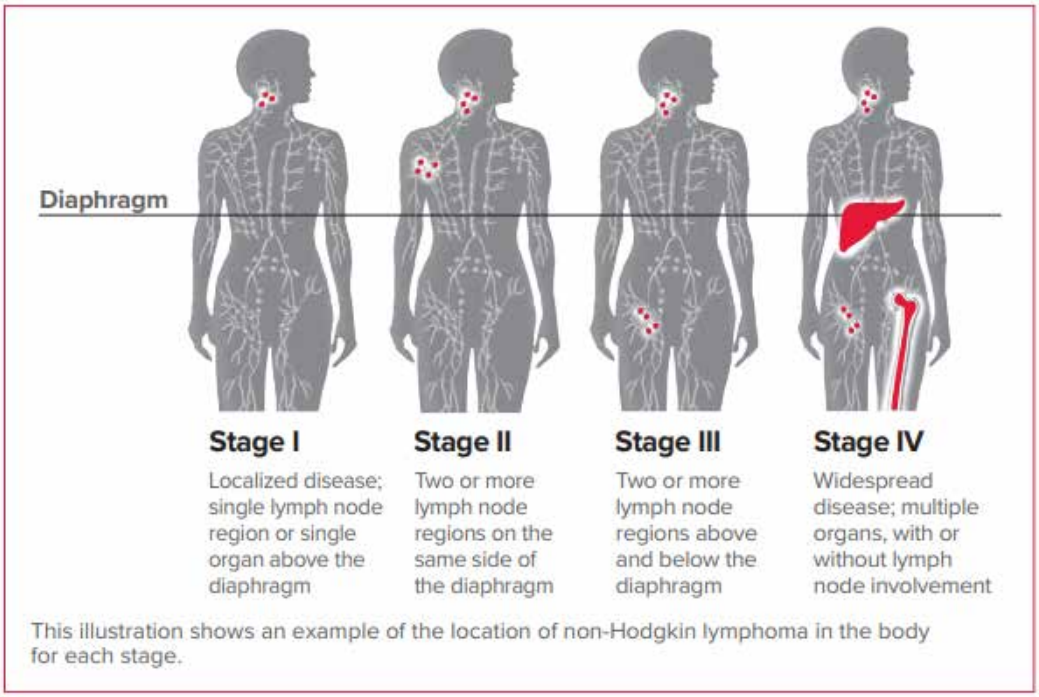
Lymphoma is a form of cancer that originates in the cells of the lymphatic system. The disease is categorized into two main types: Hodgkin lymphoma (HL) and non-Hodgkin lymphoma (NHL). Of the two, NHL is one of the most prevalent cancers in the United States, accounting for approximately 4% of all new cancer cases (“Key Statistics for Non-Hodgkin Lymphoma”, 2023). It is characterized by its

heterogeneous nature, comprising various malignancies that arise from the clonal proliferation of B-cell, T-cell, and natural killer cell subsets of lymphocytes at different stages of maturation. NHL represents around 4% of all malignancies and exhibits an overall survival rate of approximately 72% once diagnosed (Jamil, 2023). Likewise, the projected incidence of HL in the year 2023 is expected to encompass approximately 8,830 novel cases. Although constituting a relatively modest proportion of all newly diagnosed cancer cases, accounting for a mere 0.5%, HL bears a distinctive imprint on mortality rates. Projections indicate an estimated 900 deaths attributed to this condition every year, with 540 mortalities among males and 360 among females ("Cancer Stat Facts", 2023).

Stages of Lymphoma

Lymphoma is systematically categorized into discrete stages according to the extent of its dissemination within the body. These progressive stages serve as pivotal navigational tools for the processes of diagnosis, therapeutic strategizing, and prognostic assessment. Stage I lymphoma, delineates a condition in which cancer is localized within a confined region, *in situ* (see Figure 1). This can manifest as either a solitary cancerous lymph node or as an occurrence within a specific lymphoid organ, such as the thymus, or even within a discrete segment of an extra lymphatic organ. The malignancy's reach is strictly constrained at this juncture. Transitioning to Stage II lymphoma, the scope of disease involvement slightly broadens. At this stage, the two lymph nodes which are located on the same side of the diaphragm may be affected ("What Are the Different Stages of Lymphoma", 2023).

Figure 1. *Stages of Lymphoma.* *This diagram illustrates the four stages of Lymphoma, detailing the progression and spread of the disease. Stage I shows localized involvement of a single lymph node region, while Stage II involves two or more lymph node regions on the same side of the diaphragm. In Stage III, cancer spreads to lymph node regions on both sides of the diaphragm, and Stage IV indicates the disseminated involvement of one or more extra lymphatic organs or tissues ("NHL Staging", n.d.)*



Alternatively, the disease may traverse from one lymph node to an adjoining organ. Though the propagation surpasses that of Stage I, it continues to be largely circumscribed within a single side of the diaphragm. Progressing to Stage III, an amplified dissemination pattern emerges. This disease often infiltrates multiple lymph nodes distrib-

uted both above and below the diaphragm, with potential engagement of the spleen. The spanning of both sides of the diaphragm reflects an escalated systemic presence of the disease. Ultimately, Stage IV denotes the utmost progression of lymphoma, signaling extensive metastasis. The malignancy breaches the confines of the lymphatic system, infiltrating distant organs scattered throughout the body. This may encompass two or more remote organs, such as the liver or lungs. The pervasive and systemic nature of the disease's dispersion in Stage IV underscores its substantial impact, frequently necessitating comprehensive and multifaceted therapeutic interventions ("What Are the Different Stages of Lymphoma", 2023).

Treatments

Treatment for lymphoma may include radiation, chemotherapy, or a combination of both. It may also include immunotherapy or other new treatments. The treatment that is best for patients will depend on many factors, such as the type of lymphoma the patients have. Hodgkin lymphoma is also called Hodgkin's disease. It may be treated with chemotherapy, a combination of chemotherapy and radiation, immunotherapy, stem cell transplantation, or other treatments. Many treatment options are now available for non-Hodgkin lymphoma. These include traditional chemotherapy, targeted drugs, and novel therapies that are available only through clinical trials. Patients with aggressive B cell non-Hodgkin lymphoma is often treated with R-CHOP chemotherapy, consisting of cyclophosphamide, doxorubicin, vincristine, prednisone, and rituximab, with alternatives considered for potential heart damage. T-Cell NHL, such as T-lymphoblastic lymphoma/leukemia, involves intensive chemotherapy and may include intrathecal chemotherapy, maintenance therapy, or stem cell transplant, with close monitoring for complications such as tumor lysis syndrome ("R-CHOP", 2022; "Treating Non-Hodgkin Lymphoma", 2023). For individuals with HIV infections, who face an elevated risk of aggressive NHL, the prognosis has improved significantly, particularly with the use of highly active antiretroviral therapy (HAART). HAART enables better tolerance to chemotherapy and immunotherapy and has somewhat alleviated challenges related to low blood cell counts, allowing for the administration of full doses of chemotherapy. Despite this progress, caution is exercised in chemotherapy for HIV-infected patients with lymphoma, and blood counts are closely monitored. The interaction between HIV infection and lymphoma prognosis remains complex, with modern anti-HIV therapy contributing to improved outcomes ("Treating Non-Hodgkin Lymphoma", 2023).

Immunotherapy has emerged as a vital treatment modality for both B-cell and T-cell NHL. In B-cell lymphomas, monoclonal antibodies targeting the CD20 antigen, such as rituximab (Rituxan), obinutuzumab (Gazyva), ofatumumab (Arzerra), and ibritumomab tiuxetan (Zevalin), have been proven effective. Their administration, typically via intravenous infusion, requires careful monitoring for side effects ranging from mild infusion-related reactions to more severe manifestations like skin conditions and blood cell count anomalies (Halim & Maher, 2020; "Treating Non-Hodgkin Lymphoma", 2023). Assessment for prior hepatitis B infection and vigilance for post-treatment infections are also imperative. For T-cell lymphomas, innovative immunotherapeutic agents such as T-cell engaging bispecific antibodies (e.g., mosunetuzumab (Lunsumio), epcoritamab (Epkinly), glofitamab (Columvi)) have been developed. These antibodies simultaneously target the CD3 protein on T cells and specific antigens on lymphoma cells, enhancing the immune response against the cancer. Monitoring for potential complications, including cytokine release syndrome and neurological issues, is essential, and the treatment approach must be tailored to the specific lymphoma type and the patient's overall health ("Treating Non-Hodgkin Lymphoma", 2023).

Chemotherapy is a cornerstone in lymphoma treatment, employing various drug categories such as alkylating agents to damage DNA, corticosteroids to reduce inflammation, platinum drugs to inhibit DNA replication, purine analogs to disrupt DNA and RNA, anti-metabolites to impede cell growth, anthracyclines to intercalate DNA, and other drugs to interfere with cell division. Standard regimens may be supplemented with immu-

notherapy drugs to target specific cancer cells, enhancing efficacy ("Treating Non-Hodgkin Lymphoma", 2023). Targeted drug therapies have also been developed, aiming to inhibit specific cellular processes. Proteasome inhibitors block enzyme complexes within cells to control cell division, while histone deacetylase (HDAC) inhibitors modulate gene activity. BTK inhibitors can target specific proteins involved in B-cell lymphoma cell growth and survival. These targeted therapies, administered through various routes such as intravenous infusion or oral capsules, present a sophisticated approach for treating specific NHL types, especially when standard treatments are insufficient ("Treating Non-Hodgkin Lymphoma", 2023).

Single Nucleotide Polymorphisms

Single Nucleotide Polymorphisms (SNPs) constitute a fundamental form of genetic variation among individuals, characterized by alterations at a singular nucleotide position within the genomic DNA sequence. With a prevalence that exceeds 1% across the human genome, SNPs occur approximately once in every 1000 base pairs and significantly contribute to the genetic diversity observed among distinct human populations ("Single Nucleotide Polymorphisms", n.d.; "What are single nucleotide polymorphisms", 2022). SNPs may manifest as either transitions (purine to purine or pyrimidine to pyrimidine) or transversions (purine to pyrimidine or vice versa), and their effects can range from being silent (synonymous), leading to no change in amino acid sequence, to causing changes in amino acid composition. The biological ramifications of SNPs extend to various key genetic processes, such as modulation of promoter activity, alteration of messenger RNA (mRNA) stability, and influence on subcellular localization. As a result, SNPs have the potential to precipitate various pathological conditions by affecting gene expression or function, thereby underscoring their significance in both normal physiological variation and disease susceptibility (Degtyareva *et al.*, 2021; "What are single nucleotide polymorphisms", 2022).

Investigating the intricate link between SNPs and cancer, particularly in the context of lymphoma, reveals that SNPs within promoter regions can prompt alterations in the number of methylation loci, consequently impacting gene expression and elevating the risk of cancer. Beyond methylation dynamics, promoter region SNPs exert their influence through promoter activity modulation, transcription-factor binding, and histone modifications, inclusive of DNA methylation (Deng *et al.,* 2017; Jiraskova *et al.*, 2019). Notably, SNPs can also exert effects on exonal regions, where they possess the capacity to suppress gene transcription and translation, thereby perturbing cancer susceptibility. SNPs occurring within genes associated with DNA mismatch repair, cell cycle regulation, metabolism, and immunity further contribute to genetic susceptibility to cancer. Understanding these multifaceted relationships between SNPs and cancer is pivotal for unraveling the intricacies of cancer etiology and pathogenesis (Deng *et al.,* 2017; Jiraskova *et al.*, 2019).

The utilization of SNPs as potential markers for early lymphoma diagnosis offers promising avenues in cancer diagnostics. By leveraging SNP markers and arrays, comprehensive genome screening at a global scale may enable the identification of chromosomal anomalies encompassing copy number variants, DNA amplifications, deletions, and loss of heterozygosity. In contrast to earlier technologies such as cytogenetics or gene candidate approaches, which demonstrated limited sensitivity, the utilization of SNP markers can present a robust alternative. Genetic aberrations, including copy number variations (CNVs) and translocations, have been established to exhibit distinct associations with specific human cancers, notably prevalent in hematological malignancies such as lymphomas (Shahrabi *et al.*, 2020; Etebari *et al.*, 2015). The potential to establish specific markers for such cases holds the key to identifying early manifestations of respective cancers. Notably, the expedience and cost-effectiveness of SNP microarrays facilitate rapid and comprehensive investigations, thereby enabling the unveiling of genetic anomalies implicated in oncogenesis. Recent applications of SNP arrays encompass genome-wide screening, fostering enhanced correlations between genetic patterns and phenotypes, encompassing disease etiology. This underscores the pivotal role of SNP arrays in unraveling the intricate landscape

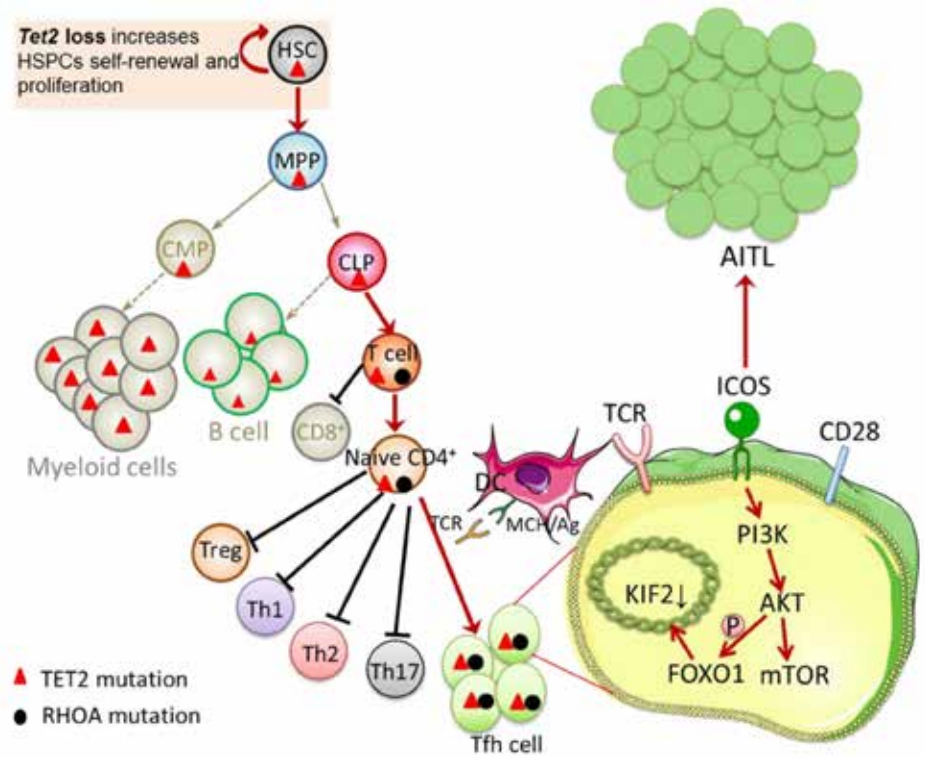
of cancer genetics and their potential contribution to early cancer diagnosis (Shahrabi *et al.*, 2020; Etebari *et al.*, 2015).

Genetics

The TET2 gene, situated at chromosome 4 position 4q24 with accession number NM_001127208.3 (NCBI, 2023), has emerged as a central topic of investigation in contemporary molecular biology. Responsible for encoding a protein with complex and elusive functionality, this gene has piqued the scientific community's interest due to its potential involvement in transcription regulation, the initial phase of protein synthesis. Its role is particularly prominent in hematopoietic stem cell differentiation, a process that encompasses the formation of vital blood cell types, such as erythrocytes (red blood cells), leukocytes (white blood cells), and thrombocytes (platelets) ("Tet2 Gene: Medlineplus Genetics.", 2014; Shastry, 2009). In addition to these functions, the TET2 protein has been identified as possessing tumor-suppressive properties, acting to inhibit abnormal cellular proliferation, thus aligning it with the characteristics of established tumor suppressor proteins.

Within this context, Figure 2 serves to provide a comprehensive visual depiction of the proposed mechanism of action of TET2, offering a tangible and illustrative insight into its functional role within cellular processes. This figure illustrates the impact of TET2 and RHOA mutations on the body, showing a complex interaction that leads to enhanced self-renewal and proliferation of hematopoietic stem/progenitor cells. This cooperative effect suppresses $CD8^+$ T cell differentiation while favoring $CD4^+$ T cells and T follicular helper (Tfh) cells, leading to aberrant activation and transformation of Tfh cells. ICOS plays a critical role through specific pathways, including the ICOS-PI3K-mTOR signaling pathway and AKT activation, contributing to Tfh lineage differentiation and the transformation characteristics of angioimmunoblastic T-cell lymphoma (AITL) (Phillip, 2022; Hu, *et al.*, 2022).

Figure 2. ***TET2 and RHOA Mutation Impact on T Cell Development and AITL.*** *TET2 loss and RHOA G17V mutations synergistically enhance HSPC self-renewal and proliferation. They also suppress CD8+ T cell differentiation, favoring CD4+ T cells and leading to Tfh cell expansion. ICOS plays a co-stimulatory role via the ICOS-PI3K-mTOR and ICOS-PI3K-AKT-FOXO1 pathways, essential for Tfh cell differentiation and AITL progression*



Gene's Implication in Lymphoma

The Tet2 gene is characterized by a prevalent mutational occurrence across a diverse array of hematopoietic malignancies, including both myeloid and lymphoid cancers, as well as various solid tumor types ("Tet2", 2017; Jiang, 2020). Within the realm of cancer cells and specific immune cell subsets such as T cells, B cells, and macrophages, Tet2 assumes an essential role in the epigenetic modification of DNA. This process is a vital aspect governing gene regulatory mechanisms and can influence the activity of genes within malignant cells and certain immune cell populations. Notably, the epigenetic modifications known as 5hmC and 5mC. 5mC is the methylation of the fifth position of the pyrimidine ring of cytosine. 5hmC is a product of 5mC demethylation by the Ten-Eleven Translocation family proteins. 5hmC was found to regulate many cellular and developmental processes, including the pluripotency of embryonic stem cells, neuron development, and tumorigenesis in mammals (Shi *et al.*, 2017). 5hmC and 5mC are orchestrated by Tet2, and are consistently present under both physiological and pathological circumstances (Jiang, 2020).

In the context of hematological neoplasms, the impairment or abrogation of Tet2 function primarily leads to the abnormal production of 5mC, thus disrupting the epigenetic landscape. This perturbation is strongly associated with the progression of solid tumors, with frequent Tet2 mutations being a hallmark feature of hematopoietic malignancies (Cong, 2021; Jiang 2020) Experimental evidence from murine models has revealed that disruption of Tet2, or concurrent disruption of both Tet2 and Tet3, precipitates the spread of myeloid or lymphoid cell populations, culminating in the formation of aggressive tumor (Jiang, 2020). These insights emphasize the critical role of Tet2 in shaping the pathogenesis of both hematopoietic and solid cancers, thereby illuminating potential avenues for therapeutic interventions.

Through a synthesis of current research, this paper aims to provide a targeted examination of the TET2 gene, focusing specifically on the association of SNPs that may be found not only on TET2 but the entirety of chromosome 4 and their potential connection to lymphoma. Given the complexity and crucial nature of this gene, the present study underscores the need for an extensive analysis to decipher the underlying relationships between the TET2 gene's variations and the onset or progression of lymphoma. This research may pave the way for novel therapeutic interventions and contribute valuable insights into personalized medicine strategies.

Methods

In the current study, we undertook an analysis of sequence reads obtained from lymphoma patients and control subjects to identify genomic variants and meticulously examine their potential implications. The reference genome for *Homo sapiens* chromosome 4, accessed from Ensembl (Release 104) (Cunningham *et al.*, 2021), was utilized as a benchmark for accurate alignment and interpretation of genomic alterations detected in our study. Our analytical pipeline incorporated the selection of Sequence Read Archive (SRA) sequences based on library construction strategy and experimental design. SRA files were obtained using the fastq-dump utility of the SRA Toolkit v2.10.7 (NCBI, 2021), with subsequent analyses performed via the command-line interface. An initial assessment of sequence read quality was conducted using FastQC (v0.11.9) (Andrews, 2010), and sequence trimming was performed using Trimmomatic (v0.39) (Bolger *et al.*, 2014), focusing on regions with compromised base quality (Phred score < 33). This process improved the overall sequence data quality by removing low-quality bases.

Sequence alignment to the reference human genome and indexing of the reads were achieved using Bowtie2 (v2.4.2) (Langmead & Salzberg, 2012). The resultant Sequence Alignment/Map (SAM) file was converted to a Binary Alignment/Map (BAM) file using SAMtools (v1.11) (Li *et al.*, 2009). Coverage depth at each genomic position was determined through the SAMtools 'depth' command, identifying regions likely to contain genetic variants. SNPs were identified using the binary variant call format (BCF) utilities from the SAMtools suite. The generated BCF file underwent stringent filtering via BCFTools (v1.11), retaining only high-confidence variants. These SNPs were further

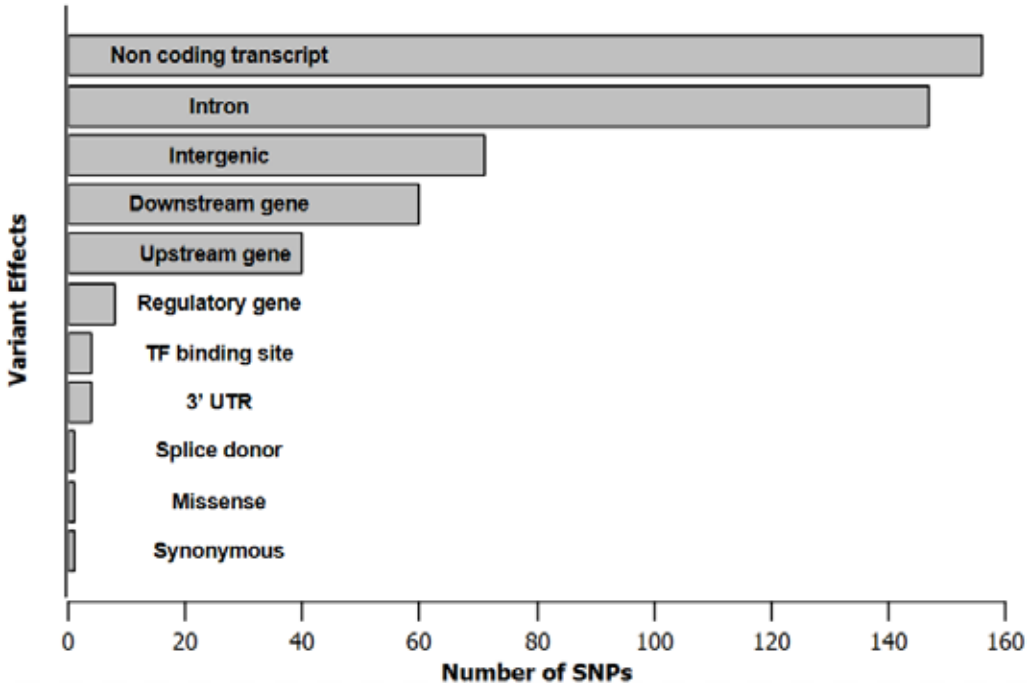
validated by comparing them to a reference dataset of SNPs from a healthy population using the Pandas library in Python. This approach allowed us to focus on lymphoma-specific genomic alterations.

Statistical analyses assessed the association between identified variants in the lymphoma and control cohorts, employing Chi-square tests for categorical variables evaluation. Contingency tables, constructed from the CSV file, exhibited subject counts with or without specific variants across both cohorts. The Python SciPy library was used to calculate the chi-square statistic and corresponding p-value, revealing a significant association ($p < 0.05$) between the identified variants and the lymphoma cohort. We further linked SNP accession numbers to their genomic consequences using a Python script we developed. SNP accession numbers were extracted from the SAMtools-generated BCF file, and the Ensembl REST API provided pertinent information regarding each SNP, including genomic location and allelic configuration. The parsed JSON responses facilitated data extraction, subsequently merged with genomic consequences and stored in a CSV file for further analysis. Our Python script allowed efficient information extraction and reduced manual error risk. All scripts used in this study are publicly available on our GitHub repository (https://github.com/crisprmax/SNP-identifier-Python), including comprehensive documentation and guidelines.

Results

In our endeavor to unravel the genetic heterogeneity and discern the functional roles of SNPs within the genome of a lymphoma clinical cohort, Figure 3 (referenced below) stands as a critical juncture in our investigative journey. This figure offers an overview, capturing both the distribution and the functional categories of the SNPs identified in our study. We analyzed a total of 1,072 SNPs, categorizing them based on their presumed functional roles. Intriguingly, a significant proportion of these variants have yet to be functionally characterized, posing an exciting avenue for future research. To maintain the integrity of the graphical representation, these uncharacterized variants were intentionally omitted, thereby enabling a more discernible visualization of the less abundant, but potentially functionally significant, variant categories.

Figure 3. ***Different Variant Effects on Lymphoma Clinical Cohort.*** *This figure illustrates the distribution and functional categorization of SNPs found within a clinical cohort of lymphoma patients. A total of 1072 SNPs were analyzed, with the majority remaining uncharacterized*



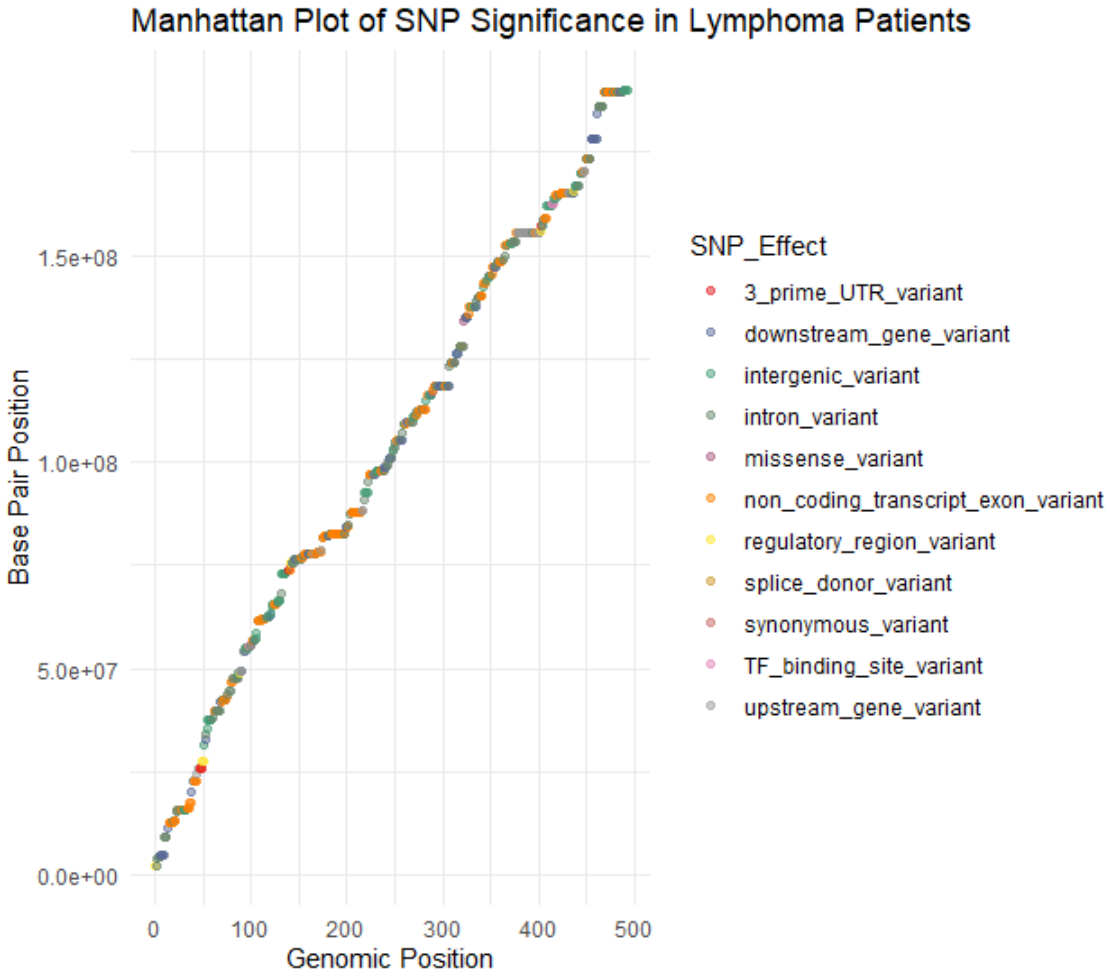
Within the dataset of characterized SNPs, multiple functional categories were represented, highlighting the diverse genetic landscape of the lymphoma clinical cohort (Table 1). Specifically, 156 SNPs were classified as non-coding transcript exon variants, indicating variations in regions that are transcribed but do not encode proteins. A further 147 SNPs were identified as intron variants, which could potentially influence splicing mechanisms. Additionally, 71 SNPs were categorized as intergenic variants, implicating regions between genes. The dataset also included 60 downstream and 40 upstream gene variants, suggesting a potential impact on the regulation of adjacent genes. Notably, 8 SNPs were found in regulatory regions and an additional 4 in transcription factor binding sites, emphasizing their likely role in gene regulation. Four SNPs were located within the 3' untranslated region (UTR), pointing to potential post-transcriptional regulatory effects. Lastly, the analysis unveiled a small number of functionally impactful SNPs: a single splice donor variant with potential implications for splicing processes, one missense variant that could alter protein function, and a lone synonymous variant with minimal impact on protein structure or function. These comprehensive insights underscore the diversity of SNPs present within the genome of the unknown organism, emphasizing potential functional repercussions that necessitate further exploration.

Table 1. ***Categorization and Descriptions of SNP Effects in Lymphoma Clinical Cohort****. This table enumerates the various types of SNPs identified within a clinical cohort of lymphoma patients, alongside brief descriptions outlining their potential functional implications*

SNP Effects	Description
Non-coding Transcript Exon Variants	Variations in regions that are transcribed but do not encode proteins.
Intron Variants	Potential influence on splicing mechanisms.
Intergenic Variants	Implicating regions between genes.
Downstream Gene Variants	Potential impact on the regulation of adjacent genes.
Upstream Gene Variants	Potential impact on the regulation of adjacent genes.
Regulatory Region Variants	Likely role in gene regulation.
Transcription Factor Binding Site Variants	Likely role in gene regulation.
3' Untranslated Region (UTR) Variants	Potential post-transcriptional regulatory effects.
Splice Donor Variant	Potential implications for splicing processes.
Missense Variants	Could alter protein function.
Synonymous Variant	Minimal impact on protein structure or function.

To better elucidate the complexities of these data, we employed a Manhattan Plot (Figure 4). Our aim was to provide an overview of both the spatial arrangement and presumptive functional significance of SNPs in our clinical cohort. The x-axis of the plot arranges the SNPs according to their respective genomic coordinates on chromosome 4, offering a seamless view of their genomic distribution. Conversely, the y-axis demarcates the precise base pair positions for each individual SNP within this chromosome, thus enriching our comprehension of their genomic context. Each data point on the plot represents a unique SNP, with its y-axis placement signifying its specific base pair location on chromosome 4. To further enhance interpretability, we incorporated a color-coding scheme based on a scale that quantifies SNP Effect values. This chromatic stratification serves to spotlight SNPs that warrant particular attention, especially those with potential functional or regulatory roles.

Figure 4. ***Manhattan Plot of SNP Significance in Lymphoma Patients.*** *This plot provides an intricate overview of the spatial distribution and potential functional implications of SNPs specific to lymphoma patients. Each point populating the plot represents an individual SNP, with its vertical placement serving as a proxy for its base pair position on chromosome 4. When applicable, the coloration of each point is governed by a color scale that corresponds to its SNP Effect value*



Our study provides an analytical framework for the investigation of SNPs within a lymphoma clinical cohort. By using advanced visualization techniques, such as the Manhattan Plot, we have been able to delineate the spatial distribution and infer potential functional roles of these genetic variants. The color-coding system employed significantly aids in identifying SNPs with potential functional or regulatory implications, serving as a roadmap for targeted genomic studies in the future. Our findings not only enhance the current understanding of the genetic landscape in lymphoma but also pave the way for more in-depth explorations that could have clinical relevance.

Discussion:

The TET2 gene has been intricately linked to epigenetic modifications, cellular differentiation, and tumor suppression, signifying its critical role in hematopoiesis and cancer development. While we did not discover SNPs directly within TET2, the genetic variants that were identified in this study still offer insights into the broader genetic landscape associated with lymphoma patients. These variants can potentially serve as markers for early diagnosis and contribute to the development of personalized medicine strategies, which could have a significant impact on patient outcomes.

The prevalence of lymphoma, particularly NHL, stands as a significant contributor

to cancer-related morbidity and mortality, underscoring the urgent need for continued research efforts aimed at improving diagnostic tools, treatment modalities, and patient outcomes. The staging process of lymphoma, illustrated in Figure 1, further highlights the progressive nature of the disease and emphasizes the critical importance of early detection and intervention. Complementing this, the role of immunotherapy and targeted therapies in lymphoma treatment is elucidated. The potential of monoclonal antibodies and innovative immunotherapeutic agents is particularly noteworthy. These advancements in treatment strategies not only hold promise for enhancing patient responses but also for reducing adverse effects, thereby reinforcing the necessity of further research and clinical trials.

The systematic investigation of SNPs in our study offers a nuanced understanding of the genetic heterogeneity underlying lymphoma. Figure 3 serves as a cornerstone, providing an exhaustive overview of the potential functional implications of these genetic variants. The notable prevalence of «unknown» variants amplifies the imperative for continued research to elucidate their specific roles and consequences in the pathogenesis of lymphoma. Complementing this, the methodological rigor of our study stands as a testament to the robustness of our SNP analysis. Leveraging advanced bioinformatics tools and publicly available datasets, we have employed meticulous alignment, sequence trimming, and variant calling processes. This ensures not only the accuracy of our findings but also the reliability of the SNPs identified, thereby laying a solid foundation for future research in this critical area of lymphoma genetics.

Conclusion:

Our study aimed to identify and categorize SNPs in lymphoma patients. Utilizing a robust and comprehensive analytical pipeline that integrated multiple bioinformatics tools and statistical methods, we were able to identify several lymphoma-specific SNPs of particular interest. A notable finding is the identification of 1072 SNPs, the majority of which remain uncharacterized. These «unknown» SNPs present an intriguing area for future research, particularly in understanding their functional roles and potential involvement in lymphoma. Our study also revealed a substantial number of categorized SNPs, including 156 non-coding transcript exon variants, 147 intron variants, and 71 intergenic variants. These variants have potential implications for gene splicing, regulation, and even intergenic interactions.

Importantly, we also identified variants that are likely to have more immediate functional implications. These include 60 downstream and 40 upstream gene variants that may influence the regulation of neighboring genes, 8 regulatory region variants, and 4 transcription factor binding site variants that suggest potential regulatory roles. This investigation serves as an intricate tapestry of the SNP landscape in lymphoma, woven together through a rigorous analytical framework. Our findings not only offer a nuanced insight into the intricate genetic landscape associated with lymphoma but also serve as a foundational platform for future research endeavors focused on the functional interpretation and potential therapeutic applications of these genetic variants in the clinical management of the disease.

References:

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. Retrieved from URL: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*,– 30(15).– P. 2114–2120. URL: https://doi.org/10.1093/bioinformatics/btu170

Cong, B., Zhang, Q., & Cao, X. (2021). The function and regulation of TET2 in innate immunity and inflammation. *Protein & cell*,– 12(3).– P. 165–173. URL: https://doi.org/10.1007/s13238-020-00796-6

Cunningham, F., Achuthan, P., Akanni, W., Allen, J., Amode, M. R., Armean, I. M., & Flicek, P. (2021). Ensembl 2021. *Nucleic Acids Research*,– 49(D1).– D884-D891. URL: https://doi.org/10.1093/nar/gkaa942

Degtyareva, A. O., Antontseva, E. V., & Merkulova, T. I. (2021). Regulatory SNPs: Altered Transcription Factor Binding Sites Implicated in Complex Traits and Diseases. *International journal of molecular sciences*,– 22(12).– 6454 p. URL: https://doi.org/10.3390/ijms22126454

Etebari, M., Navari, M., & Piccaluga, P. P. (2015). SNPs Array Karyotyping in Non-Hodgkin Lymphoma. *Microarrays (Basel, Switzerland)*,– 4(4).– P. 551–569. URL: https://doi.org/10.3390/microarrays4040551

Halim, L., & Maher, J. (2020). CAR T-cell immunotherapy of B-cell malignancy: the story so far. *Therapeutic advances in vaccines and immunotherapy*,– 8, 2515135520927164. URL: https://doi.org/10.1177/2515135520927164

Hu, L., Zhang, X., Li, H., Lin, S., & Zang, S. (2022, November 20). *Targeting TET2 as a therapeutic approach for angioimmunoblastic T cell lymphoma*. MDPI. URL: https://www.mdpi.com/2072-6694/14/22/5699

Jamil A., Mukkamalla S.K.R. Lymphoma. [Updated 2023 July 17]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan-. Available from: URL: https://www.ncbi.nlm.nih.gov/books/NBK560826/

Jiang S. (2020). Tet2 at the interface between cancer and immunity. *Communications biology*,– 3(1).– 667 p. URL: https://doi.org/10.1038/s42003–020–01391–5

Jiraskova K., Hughes D.J., Brezina S., Gumpenberger T., Veskrnova V., Buchler T., Schneiderova M., Levy M., Liska V., Vodenkova S., et al. Functional Polymorphisms in DNA Repair Genes Are Associated with Sporadic Colorectal Cancer Susceptibility and Clinical Outcome. *International Journal of Molecular Sciences*. 2019; 20(1):97. URL: https://doi.org/10.3390/ijms20010097

Key statistics for Non-Hodgkin Lymphoma. American Cancer Society. (2023). URL: http://www.cancer.org/cancer/types/non-hodgkin-lymphoma/about/key-statistics.html#:~:text=The%20American%20Cancer%20Society%27s%20estimates,11%2C780%20males%20and%208%2C400%20females).

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*,– 9(4).– P. 357–359. URL: https://doi.org/10.1038/nmeth.1923

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*,– 25(16). 2078–2079. URL: https://doi.org/10.1093/bioinformatics/btp352

Lymphoma-patient version. National Cancer Institute. (2023). URL: http://www.cancer.gov/types/lymphoma

National Center for Biotechnology Information (NCBI). (2021). SRA Toolkit v2.10.7. Retrieved from URL: https://github.com/ncbi/sra-tools

NHL staging. https://www.lls.org/lymphoma/non-hodgkin-lymphoma/diagnosis/nhl-staging. (2023). URL: https://www.lls.org/lymphoma/non-hodgkin-lymphoma/diagnosis/nhl-staging

Non-hodgkin lymphoma stages: Staging Non-Hodgkin Lymphoma. Staging Non-Hodgkin Lymphoma | American Cancer Society. (n.d.). URL: *https://www.cancer.org/cancer/types/non-hodgkin-lymphoma/detection-diagnosis-staging/staging.html*

Philip, M., Schietinger, A. *CD8+ T cell differentiation and dysfunction in cancer. Nat Rev Immunol* – 22. – P. 209–223 (2022). URL: *https://doi.org/10.1038/s41577-021-00574-3*

R-CHOP: *Combination chemotherapy cancer treatment*. City of Hope. (2022, June 14). URL: https://www.cancercenter.com/treatment-options/r-chop#:~:text=Doctors%20use%20R%2DCHOP%20to, other%20cancer%20drugs%20and%20treatments

Shahrabi, S., Ghanavat, M., Behzad, M. M., Purrahman, D., & Saki, N. (2020). *CD markers polymorphisms as prognostic biomarkers in hematological malignancies. Oncology reviews,*– 14(2).– 466 p. URL: *https://doi.org/10.4081/oncol.2020.466*

Shastry B.S. (2009). *SNPs: impact on gene function and phenotype. Methods in molecular biology (Clifton, N.J.),*– 578.– P. 3–22. URL: *https://doi.org/10.1007/978-1-60327-411-1_1*

Shi, D.-Q., Ali, I., Tang, J., & Yang, W.-C. (2017, July 5). *New insights into 5hmc DNA modification: Generation, distribution and function*. Frontiers. URL: https://www.frontiersin.org/articles/10.3389/fgene.2017.00100/full

Single nucleotide polymorphisms (snps). Genome.gov. (n.d.). URL: *https://www.genome.gov/genetics-glossary/Single-Nucleotide-Polymorphisms*

Stages of lymphoma. moffitt. (2023). URL: http://www.moffitt.org/cancers/lymphomas-hodgkin-and-non-hodgkin/diagnosis/stages

Tet2. Tet2 – My Cancer Genome. (2017). URL: https://www.mycancergenome.org/content/gene/tet2

Treating Non-Hodgkin Lymphoma: Non-Hodgkin Lymphoma Treatments. Non-Hodgkin Lymphoma Treatments | American Cancer Society. (2023). URL: https://www.cancer.org/cancer/types/non-hodgkin-lymphoma/treating.html

U. S. National Library of Medicine. (2023). Tet2 gene: Medlineplus Genetics. MedlinePlus. https://medlineplus.gov/genetics/gene/tet2/

U. S. National Library of Medicine. (2023). *Tet2 Tet methylcytosine dioxygenase 2 [homo sapiens (human)] – gene – NCBI*. National Center for Biotechnology Information. URL: https://www.ncbi.nlm.nih.gov/gene/54790

submitted 15.05.2024;
accepted for publication 29.05.2024;
published 22.07.2024

Contact: liliww2@gmail.com