haematologica
Journal of the Ferrata Storti Foundation

Transcriptional landscape of nodular lymphocyte-predominant Hodgkin lymphoma and T-cell/histiocyte-rich large B-cell lymphoma: tumor microenvironment and checkpoint genes

by Ilja Kalashnikov, Veronica Russell, Rachel Kositsky, Leonardo P.A. Biral, Panu Kovanen, Johannes Dunkel, Marja-Liisa Karjalainen-Lindsberg, Annika Pasanen, Devang Thakkar, Jean L. Koff, David Jaye, Sarah Ondrejka, Eric D. Hsi, Emily F. Mason, Ridas Juskevicius, Cassandra Love, Jennifer Shingleton, Elizabeth Thacker, Lanie Happ, Clay Parker, Tushar Dave, Mette Ølgod Pedersen, Magdalena Czader, Jiehao Zhou, Dennis O'Malley, Mina Xu, Andrew Evans, Chee Leong Cheng, Matt McKinney, Chad M. McCall, Eric Tse, Rex Au Yeung, Yasodha Natkunam, Abner Louissaint, Jr., Jennifer Chapman, Catalina Amador, Amy Chadburn, Yuri Fedoriw, Amir Behdad, Carlos Galvez, Payal Sojitra, Andrew Evens, Kikkeri N. Naresh, Sandeep S. Dave and Sirpa Leppä

Received: August 2, 2026.
Accepted: November 5, 2026.

Citation: Ilja Kalashnikov, Veronica Russell, Rachel Kositsky, Leonardo P.A. Biral, Panu Kovanen, Johannes Dunkel, Marja-Liisa Karjalainen-Lindsberg, Annika Pasanen, Devang Thakkar, Jean L. Koff, David Jaye, Sarah Ondrejka, Eric D. Hsi, Emily F. Mason, Ridas Juskevicius, Cassandra Love, Jennifer Shingleton, Elizabeth Thacker, Lanie Happ, Clay Parker, Tushar Dave, Mette Ølgod Pedersen, Magdalena Czader, Jiehao Zhou, Dennis O'Malley, Mina Xu, Andrew Evans, Chee Leong Cheng, Matt McKinney, Chad M. McCall, Eric Tse, Rex Au Yeung, Yasodha Natkunam, Abner Louissaint, Jr., Jennifer Chapman, Catalina Amador, Amy Chadburn, Yuri Fedoriw, Amir Behdad, Carlos Galvez, Payal Sojitra, Andrew Evens, Kikkeri N. Naresh, Sandeep S. Dave and Sirpa Leppä. Transcriptional landscape of nodular lymphocyte-predominant Hodgkin lymphoma and T-cell/histiocyte-rich large B-cell lymphoma: tumor microenvironment and checkpoint genes. Haematologica. 2026 Nov 12. doi: 10.3324/haematol.2026.301728 [Epub ahead of print]

Publisher's Disclaimer.
E-publishing ahead of print is increasingly important for the rapid dissemination of science. Haematologica is, therefore, E-publishing PDF files of an early version of manuscripts that have completed a regular peer review and have been accepted for publication.
E-publishing of this PDF file has been approved by the authors.
After having E-published Ahead of Print, manuscripts will then undergo technical and English editing, typesetting, proof correction and be presented for the authors' final approval, the final version of the manuscript will then appear in a regular issue of the journal.
All legal disclaimers that apply to the journal also pertain to this production process.

Transcriptional landscape of nodular lymphocyte-predominant Hodgkin lymphoma and T-cell/histiocyte-rich large B-cell lymphoma: tumor microenvironment and checkpoint genes

Ilja Kalashnikov[1,2,3], Veronica Russell[4], Rachel Kositsky[4], Leonardo P.A. Biral[4], Panu Kovanen[5], Johannes Dunkel[5], Marja-Liisa Karjalainen-Lindsberg[5], Annika Pasanen[1,2], Devang Thakkar[6], Jean L. Koff[7], David Jaye[7], Sarah Ondrejka[8], Eric D. Hsi[9], Emily F. Mason[10], Ridas Juskevicius[10,11], Cassandra Love[4], Jennifer R. Shingleton[4], Elizabeth Thacker[6], Lanie Happ[6], Clay Parker[6], Tushar Dave[4], Mette Ølgod Pedersen[12,13], Magdalena Czader[14], Jiehao Zhou[14], Dennis O'Malley[15], Mina Xu[16], Andrew Evans[17], Chee Leong Cheng[18], Matthew McKinney[4], Chad M. McCall[4], Eric Tse[19], Rex Au Yeung[19], Yasodha Natkunam[20], Abner Louissaint, Jr.[21], Jennifer Chapman[22], Catalina Amador[23], Amy Chadburn[24], Yuri Fedoriw[25], Amir Behdad[26], Carlos Galvez[27], Payal Sojitra[28], Andrew Evens[29], Kikkeri N Naresh[30], Sandeep S. Dave[4], Sirpa Leppä[1,2,3]

[1]Research Program Unit, Applied Tumor Genomics, Faculty of Medicine, University of Helsinki, Helsinki, Finland.
[2]Department of Oncology, Helsinki University Hospital Comprehensive Cancer Center, Helsinki, Finland
[3]iCAN Digital Precision Cancer Medicine Flagship, Helsinki, Finland
[4]Duke University, Durham, North Carolina, United States
[5]Department of Pathology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
[6]Data Driven Bioscience, Durham, North Carolina, United States
[7]Winship Cancer Institute, Emory University School of Medicine, Atlanta, Georgia, United States
[8]Pathology and Laboratory Medicine Institute at Cleveland Clinic, Cleveland, Ohio, United States
[9]Mayo Clinic, Rochester, Minnesota, United States
[10]Vanderbilt University Medical Center, Nashville, Tennessee, United States
[11]PathGroup, Nashville, Tennessee, United States
[12]Department of Pathology, Zealand University Hospital, Copenhagen, Denmark
[13]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark
[14]Department of Pathology and Laboratory Medicine, Indiana University, Indianapolis, Indiana, United States
[15]Neogenomics, Fort Myers, Florida, United States
[16]Department of Pathology, Yale University, New Haven, Connecticut
[17]Department of Pathology and Laboratory Medicine, University of Rochester, Rochester, New York, United States
[18]Department of Anatomical Pathology, Singapore General Hospital, Singapore

[19]University of Hong Kong, Hong Kong
[20]Stanford University Medical Center, Stanford, California, United States
[21]Massachusetts General Hospital, Boston, Massachusetts, United States
[22]Department of Pathology, University of Miami and Sylvester Comprehensive Cancer Center, Miami, Florida, United States
[23]University of Nebraska, Omaha, Nebraska, United States
[24]Department of Pathology, Weill Cornell Medical College, New York, New York, United States
[25]University of North Carolina, Chapel Hill, North Carolina, United States
[26]Cleveland Clinic Florida, Weston, Florida, United States
[27]University of Illinois, Chicago, Illinois, United States
[28]Robert Wood Johnson Medical School, New Brunswick, New Jersey, United States
[29]Rutgers Cancer Institute, New Brunswick, New Jersey, United States
[30]Pathology Program, Translational Science & Therapeutics Division, Department of Pathology, Fred Hutchinson Cancer Center, Seattle, Washington, United States

Corresponding author:
Sirpa Leppä, MD, Professor
Department of Oncology and Research Programs Unit
University of Helsinki and Helsinki University Hospital Comprehensive Cancer Center
P.O. Box 180
FI-00029 Helsinki
Finland
Phone: +358 407558293
e-mail: sirpa.leppa@helsinki.fi

Running head: Transcriptional Landscape of NLPHL and THRLBCL

Key words: NLPHL, THRLBCL, transcriptome, tumor microenvironment, survival

Acknowledgments
We thank all participating sites for samples and clinical data.

Financial support
This research was supported by grants from the Research Council of Finland (S.L), Cancer Foundation Finland (S.L), Sigrid Juselius Foundation (S.L), University of Helsinki (S.L and I.K), Helsinki and Uusimaa Hospital District (S.L), Ida Montin Foundation (I.K), Orion Research Foundation (I.K), Biomedicum Helsinki Foundation

(I.K), Emil Aaltonen Foundation (I.K), Finnish Medical Foundation (I.K), and Finska Läkaresällskapet (I.K). Data generation was supported by Data Driven Bioscience in an unrestricted manner. S.D was supported by the Wellcome Trust Foundation and the National Institutes of Health.

Conflict of interest disclosure

R.K: Current employment: Astra Zeneca; E.D.H: Consultancy fees from Abcon Therapeutics; J.L.K: Consultancy fees from AbbVie, ADC Therapeutics, BeOne, Incyte, and Pierre Fabre; Research support from Viracta Therapeutics and Oncternal Therapeutics; S.L: Consultancy fees from AbbVie, Astra Zeneca, Genmab, Gilead, Incyte, and Roche, all outside the submitted work; honoraria from Gilead, Incyte, and Roche; research grants from BMS, Galapagos, Genmab, Novartis, and Roche, all outside the submitted work. C.G: Consultancy fees from Kite and AstraZeneca. S.D, C.L, J.R.S, E.T, L.H, C.P, and D.T own equity in Data Driven Bioscience. C.L, J.S, L.P.A.B: Consultancy fees from Data Driven Bioscience. Other authors do not have relevant conflicts of interest.

Prior presentation

Part of the work has been presented at the ASH Annual Meeting 2023.

Author contributions

I.K, V.R, R.K, S.D, and S.L conceived the study; I.K, V.R, R.K, and D.T analyzed the data; All other authors provided samples and/or clinical data; I.K and S.L wrote the manuscript; S.D and S.L supervised the study; and all authors read and approved the manuscript.

Data sharing

The RNA sequencing data will be publicly available through the European Genome Archive (EGA).

Artificial Intelligence

Claude Opus 5.5 (Anthropic, San Francisco, CA, USA), a commercial large language model, was used to check and correct the grammar and wording of author-written text, with prompts asking for grammar correction and shorter or clearer phrasing. It was not used to generate scientific content, analyze or interpret data, create figures, or search for or select references. All suggestions were reviewed by the authors, who take full responsibility for the content.

ABSTRACT

Nodular lymphocyte-predominant Hodgkin lymphoma (NLPHL) is a rare B-cell malignancy with an indolent course, yet some cases relapse or transform to T-cell/histiocyte-rich large B-cell lymphoma (THRLBCL). Biological factors underlying poor outcomes in NLPHL are largely unknown, and the relationship between THRLBCL and untransformed NLPHL has not been systematically defined. We analyzed 178 NLPHL and 22 THRLBCL cases using whole-transcriptome RNA sequencing and clinical data. NLPHL patients were younger, had limited-stage disease, and better outcomes than THRLBCL patients. Variant immuno-architectural patterns (IAP) in NLPHL showed THRLBCL-like transcriptional profiles, especially with splenic involvement and high Lymphocyte-Predominant International Prognostic Score (LP-IPS). In-silico deconvolution of the tumor microenvironment (TME) revealed B-cell enrichment in classic NLPHL, intermediate profiles in variant NLPHL, and increased activated CD4+ and CD8+ T-cells and macrophages in THRLBCL. Higher follicular helper T-cell content was associated with shorter event-free survival in an exploratory analysis. NLPHL cases showed strong B-cell and T-helper signatures, whereas THRLBCL cases were enriched for macrophage and cytotoxic T-cell pathways and expressed higher immune checkpoint genes, suggesting an immunosuppressive TME. Signatures defined a spectrum from classic to variant NLPHL and THRLBCL. External validation in independent cohorts reproduced this gradient, supporting a continuous microenvironmental spectrum spanning NLPHL and THRLBCL.

INTRODUCTION

Nodular lymphocyte-predominant Hodgkin lymphoma (NLPHL), also referred to as nodular lymphocyte-predominant B-cell lymphoma (NLPBL)[1], is a rare B-cell malignancy characterized by unique clinicopathological features, with a global incidence of approximately 1–3 cases per million annually[1,2]. Compared to classic Hodgkin lymphoma (cHL), NLPHL predominantly affects middle-aged males, usually presenting with isolated peripheral lymphadenopathy[3]. Although the clinical course of NLPHL is generally indolent and survival rates in young adults approach those of the general population, some patients experience recurrent disease and/or transformation to T-cell/histiocyte-rich large B-cell lymphoma (THRLBCL), an aggressive lymphoma associated with poor outcomes[4–7].

The histological complexity of NLPHL is captured by the Fan classification, which comprises six immuno-architectural patterns (IAP; patterns A–F)[8]. Classic patterns (A–B) exhibit malignant cells within a predominantly nodular B cell-rich tumor microenvironment (TME), while variant patterns (C–F) feature extrafollicular LP cells and/or a more diffuse and T-cell-rich TME[8]. In particular, variant pattern E closely resembles THRLBCL, complicating accurate diagnosis[1,2]. Additional NLPHL architectural patterns beyond the Fan classification have also been described[9].
The prognostic significance of the IAP classification in NLPHL is controversial. Early studies suggested that variant patterns correlate with worse clinical outcomes[10,11], and a pooled GHSG analysis of 583 patients reported significantly worse outcomes for patterns D-E versus A-C and F[12]. In contrast, a large international retrospective study conducted by the Global NLPHL One Working Group found no independent prognostic significance of individual patterns after adjusting for relevant clinical parameters[13]. Instead, the study introduced the Lymphocyte-Predominant International Prognostic Score (LP-IPS), which assigns one point each for the following criteria: age ≥45 years, stage III–IV disease, hemoglobin level <105 g/L, and splenic involvement[13]. Higher LP-IPS scores (≥2 points) were significantly associated with worse progression-free survival, overall survival, and an increased risk of transformation[13]. These conflicting data highlight an unmet need to define whether variant patterns add prognostic value beyond established clinical risk factors.

Despite their distinct clinical presentations, NLPHL and THRLBCL share overlapping morphological and immunophenotypic characteristics, notably sparse malignant B cells embedded within an immune-rich TME[1,2]. Prior comparative studies analyzing gene expression profiles, copy number variations, and mutational landscapes underscore significant similarities between the malignant cells of both entities, suggesting a potential continuum between NLPHL and THRLBCL[14–19]. The main distinctions relate to differences in TME composition and cellular interactions. Earlier profiling of the non-malignant TME component suggested that THRLBCL exhibits a more tolerogenic immune profile than NLPHL[20]. Recently, Subramanian et al. conducted a large multimodal analysis of 171 NLPHL cases, defining 34 cell states across 14 lineages to organize NLPHL cases into three lymphocyte-predominant ecotypes (LPE1–3) with prognostic associations[21]. Notably, these associations were affected by whether patients received single-modality therapy or combined-modality therapy[21]. Another large study used NanoString-based profiling to examine immune checkpoint biomarkers in NLPHL and classic Hodgkin lymphoma and reported enrichment of *PDCD1* (PD-1) and *TIGIT* expression in NLPHL[22].

In addition, studies using high-plex immunofluorescence have recently advanced our understanding of the TME, highlighting its structured cellular composition and spatial organization, and further emphasizing the role of the TME in shaping clinical behavior[23,24]. Although the importance of the TME in NLPHL and THRLBCL is evident, genomic studies have historically focused predominantly on malignant cells, overlooking broader TME implications. To address this gap, we performed comprehensive gene expression profiling of tumors in a large, multicenter cohort of NLPHL and THRLBCL patients as part of the Atlas of Blood Cancer Genomes (ABCG) initiative. Importantly, the tumor samples comprised malignant cells embedded within the TME, allowing us to investigate the role of the TME in tumor dynamics and patient outcomes.

METHODS

Cohort and pathology review

The cohort included formalin-fixed paraffin-embedded (FFPE) tumor samples from 178 patients with NLPHL and 22 patients with THRLBCL, collected across 24 institutions as part of the Atlas of Blood Cancer Genomes (ABCG) initiative. Tissue biopsies and clinical data were collected with local institutional review board approval at each center in accordance with the Declaration of Helsinki. Written informed consent was waived because the study was retrospective and patient information was de-identified. A central panel of expert hematopathologists confirmed each diagnosis and classified NLPHL into immuno-architectural patterns (IAP) A to F using established histologic criteria. All THRLBCL cases were de novo, with no documented prior history of NLPHL. Sample timing relative to diagnosis is detailed in the Supplementary Methods.

RNA sequencing and processing

Nucleic acid extraction and library preparation used the Duoseq Research Kit (Data Driven Bioscience, Durham, NC) in RNA mode. Libraries were sequenced on the Illumina NovaSeq X platform. Two hundred cases met whole-transcriptome RNA sequencing (RNA-seq) quality standards. Reads were trimmed, aligned to the human genome (GRCh38), and quantified at the gene level.[25-31] Counts were restricted to protein-coding genes and normalized with DESeq2.[32] Immune-cell fractions were estimated with FARDEEP using the LM22 leukocyte signature matrix.[33,34] Full pipeline steps and software versions are provided in the Supplementary Methods.

Expression signature and external validation

Differential expression was assessed with DESeq2 across sequential and grouped IAP comparisons. A top-scoring pair (TSP) approach was used to derive a signature separating classic NLPHL, variant NLPHL, and THRLBCL.[35] Samples were split 75/25 into training and test sets, stratified by splenic involvement. Signature performance was assessed in the test set and across 1,000 bootstrap iterations. The signature was then applied without refitting to two independent public cohorts, Van

Loo et al.[20] and Subramanian et al.[21] Gene selection, scoring, and cohort handling are described in the Supplementary Methods.

Pathway and statistical analysis

Over-representation analysis was performed with clusterProfiler,[36] using xCell, CIBERSORT, and MSigDB Hallmark and canonical pathway (C2) gene sets as the universe.[37-39] All analyses used R version 4.5.0. Heatmaps and volcano plots were generated with pheatmap and EnhancedVolcano.[40,41] Principal component analysis and permutational multivariate analysis of variance (PERMANOVA) were used to compare the diagnoses. Survival analyses of Tfh content were exploratory and corrected for multiple testing. In the Subramanian cohort, the score and the Tfh fractions were tested against freedom from progression with LP-IPS-adjusted Cox models (Supplementary Methods). Categorical variables were compared with the Fisher exact test and continuous variables with the Wilcoxon rank-sum test. Discrimination was quantified by receiver operating characteristic area under the curve (ROC-AUC). Survival was analyzed by Kaplan-Meier estimation with the log-rank test and by univariable and multivariable Cox regression, reported as hazard ratios (HR) with 95% confidence intervals (CI).[42,43] Overall survival (OS) was measured from diagnosis to death from any cause. Event-free survival (EFS) was measured from diagnosis to progression, transformation, or death from any cause.

RESULTS

Patient characteristics and clinical outcomes

Table 1 summarizes the baseline clinical features. Patients with NLPHL were significantly younger at diagnosis compared to those with THRLBCL (median age: 42 vs 55 years). Both entities predominantly affected males (71% NLPHL, 84% THRLBCL). Limited-stage disease (68%) and a low LP-IPS score (0–1 point, 76%) were more common in NLPHL, whereas THRLBCL patients frequently presented with advanced-stage disease (89%) and a higher rate of splenic involvement (53%). All THRLBCL cases were de novo, with no prior history of NLPHL documented at diagnosis.

Of the reviewed NLPHL cases, IAP could be assessed in 140 samples (79%), among which classic patterns A–B were most frequent, accounting for 61%

(Supplementary Table S1). Among NLPHL patients, those with splenic involvement were older at diagnosis (median age: 49 vs. 40 years, p = 0.02).

Treatment modalities varied by disease stage and diagnosis. Radiotherapy was most commonly administered for limited-stage NLPHL (39%), while chemotherapy, mainly ABVD, was the most common regimen for advanced-stage disease (73%) (Supplementary Table S1). Rituximab was included in first-line therapy in approximately one-third of NLPHL cases. In contrast, nearly all THRLBCL patients received R-CHOP immunochemotherapy.

With a median follow-up of 6.4 years (range 0–39 years), NLPHL patients exhibited favorable outcomes, including a 5-year OS rate of 96% and an EFS rate of 78% (Figure 1A-B). However, 30 NLPHL patients experienced relapse, and transformation to aggressive lymphoma occurred in five cases. Patients with THRLBCL had significantly worse prognosis, with 5-year OS and EFS rates of 70% and 47%, respectively (Figure 1A-B).

Adverse prognostic factors identified in NLPHL included advanced stage (Figure 1C-D, OS, p = 0.03; EFS, p < 0.01), presence of splenic involvement (Figure 1E-F, OS and EFS, p < 0.01), and LP-IPS ≥2 points (Figure 2A-B, OS and EFS, p < 0.01). Variant patterns were associated with a trend toward poorer survival (Figure 2C-D). However, these differences did not reach statistical significance (OS, p = 0.1; EFS, p = 0.2).

Gene expression profiles reveal distinct immune landscape signatures in NLPHL and THRLBCL

Principal component analysis showed that NLPHL and THRLBCL occupy a largely shared transcriptional space, with the THRLBCL cases shifted along the first component rather than forming a separate cluster (Figure 3A). Diagnosis accounted for only 2.6% of the total variance in gene expression (PERMANOVA p = 0.0001). Accordingly, 19 of 22 THRLBCL cases fell inside the ellipse holding 95% of the NLPHL distribution, and the three outside had advanced-stage disease. The shift was toward the variant patterns. Over all genes, 21 of 22 THRLBCL cases lay closer to the average expression profile of variant than of classic NLPHL (p = 0.002).

NLPHL showed marked upregulation of canonical B-cell-associated and stromal genes (e.g. *MS4A1*/CD20, *CD79B*, and *COL19A1*) (Figure 3B, Supplementary Table S6). In contrast, THRLBCL exhibited elevated expression of genes associated with macrophage activation and monocyte differentiation (e.g., *MT1G*, *CD163*, *FCGR1A*) and higher expression of immune checkpoint-related genes (*HAVCR2*/TIM-3, *IDO1*), as well as T-cell-associated genes (*GZMH*), suggesting cytotoxic T-cell activity (Figure 3B, Supplementary Table S6).

Figure 3C shows a supervised heatmap in which genes were pre-selected by differential expression analysis and samples ordered by diagnostic class. Ordering NLPHL samples by IAP highlighted distinct expression patterns across IAP categories. NLPHL cases with classic patterns showed strong B-cell-associated signatures, whereas variant patterns, particularly IAP E, displayed intermediate gene-expression profiles that more closely resembled THRLBCL. NLPHL cases with splenic involvement demonstrated substantial transcriptional overlap with THRLBCL and were enriched for variant patterns, further supporting their biological proximity to THRLBCL. In addition, NLPHL cases with a high LP-IPS (≥2 points) or advanced stage showed gene-expression patterns that shifted toward a THRLBCL-like profile (Figure 3C). In contrast, we did not identify distinct transcriptional signatures associated with adverse clinical outcomes (relapse, transformation, or death) or with patient sex among NLPHL cases.

Figure 4A and Supplementary Table S2 illustrate a transcriptional continuum as NLPHL progresses through IAP categories and ultimately toward THRLBCL. In pairwise comparisons, expression of differentially expressed genes was associated with macrophage activation and increased inflammation. T-cell cytotoxicity increased gradually through the IAP categories, being highest in THRLBCL. Conversely, the genes in the B cell lineage were downregulated along the same trajectory. Overrepresentation analysis of upregulated genes (Figure 4B) highlighted enrichment of innate immunity and macrophage-associated pathways, whereas the corresponding analysis of downregulated genes (Figure 4C) emphasized loss of B-cell receptor signaling and B-cell maturation signatures.

Expression analysis captures transcriptional continuum from NLPHL to THRLBCL

To investigate the transcriptional continuum from NLPHL IAP A-E to THRLBCL, we developed robust gene signatures using the top scoring pair analysis (TSP) method (Supplementary Table S4). IAP F was excluded from this analysis due to the small number of available cases (n=4). This approach identified gene pairs that reliably differentiated between classic NLPHL, variant NLPHL, and THRLBCL across the entire cohort. The resulting heatmap illustrated progressive transcriptional shifts correlating with IAP progression from classic NLPHL (IAP A-B) to variant NLPHL (IAP C-E), culminating in THRLBCL (Figure 5A).

Gene pairs selected for the TSP signature consisted of one immune-response-related gene upregulated in THRLBCL or variant NLPHL, paired with one B-cell-related gene significantly upregulated in classic NLPHL. This approach yielded two major expression signatures: a B-cell signature and an immune signature. Aggregated signature scores showed significant differences across individual patterns and IAP categories in both training and test datasets as well as bootstrapped datasets (Supplementary Table S5). Classic NLPHL cases exhibited the lowest scores, variant NLPHL showed intermediate values, and THRLBCL displayed the highest scores, confirming a stepwise molecular transition along the disease spectrum (Figures 5B–E). An independent, cross-validated classifier reproduced this ordering, yielding out-of-fold THRLBCL probabilities that increased from classic NLPHL through variant NLPHL to THRLBCL (AUC 0.95; Spearman $\rho = 0.61$, $p < 0.001$; Figure S1).

External validation of the immune/B-cell signature score

We next tested whether the signature score generalizes to independent public datasets. In Van Loo et al.[20] (GSE7788; 10 NLPHL and 10 THRLBCL), the score robustly separated THRLBCL from NLPHL (ROC-AUC=1.00; Wilcoxon $p < 0.01$; Figure S2). In Subramanian et al.[21], restricting to cases with unambiguous IAP patterns and pooling the published training and validation cohorts (Classic n=71; Variant n=31) showed higher scores in variant compared with classic NLPHL (ROC-AUC=0.80; Wilcoxon $p < 0.01$; Figure S3; Supplementary Table S9). In that cohort the score was also higher in patients with high LP-IPS or advanced stage (Supplementary Table S11).

Tumor microenvironment deconvolution highlights cellular differences between NLPHL and THRLBCL

Given that gene expression in these cases predominantly stems from the TME due to the sparsity of malignant cells, we estimated relative immune cell proportions across NLPHL and THRLBCL by in silico cell-type deconvolution. Analysis revealed notable differences in immune cell composition among classic NLPHL, variant NLPHL, and THRLBCL (Figure 6A, Supplementary Table S3). Consistent with histopathological observations and gene expression data, classic NLPHL showed a higher relative abundance of B cells, including both memory B-cell (Figure S4A) and naïve B-cell subsets (Figure S4B). Macrophage composition varied notably among the entities, with THRLBCL showing significantly elevated proportions of both M1-like (Figure 6B) and M2-like macrophages (Figure 6C) compared to classic and variant NLPHL. However, classic and variant NLPHL did not differ significantly in M2-like macrophage proportions.

THRLBCL exhibited the highest proportion of activated CD4+ memory T cells, followed by variant and classic NLPHL (Figure S4C). In contrast, resting CD4+ memory T cells showed an inverse pattern, being significantly more abundant in classic NLPHL and decreased in variant NLPHL and THRLBCL (Figure S4D). Follicular helper T cells (Tfh; Figure 6D) were significantly more abundant in both classic and variant NLPHL compared to THRLBCL, with no significant differences between NLPHL subgroups. Finally, THRLBCL displayed the highest infiltration of CD8+ T cells (Figure S4E), accompanied by a significant reduction in naïve and memory B-cell populations.

Activated NK cells (Figure S4F) and resting NK cells (Figure S4G) were significantly enriched in THRLBCL compared to both classic and variant NLPHL, although no significant difference was found within NLPHL IAP subgroups. Regulatory T-cell fractions (Figure S4H) were comparable across all groups. Neutrophils, although rare, were most abundant in THRLBCL (Figure S4I).

Integration of clinical data showed that higher proportions of cytotoxic and activated T-cell subsets in NLPHL correlated with advanced stage, higher LP-IPS scores, and splenic involvement, consistent with their transcriptional profiles aligning more closely with THRLBCL. Interestingly, among the 22 deconvolved populations, a high

(above median) proportion of Tfh cells showed the strongest association with EFS (HR 2.67, 95% CI 1.21–5.87; log-rank p = 0.011), although no population remained significant after correction for multiple testing (q = 0.22; Supplementary Table S10). Notably, patients with high Tfh cell levels exhibited elevated expression of genes associated with cytotoxic T-cell immunity (*CD5*, *GZMK*) and higher expression of immune checkpoint-related genes (*TOX, PDCD1*/PD-1, *CTLA4*, *TIGIT*) (Supplementary Table S7). Within the high LP-IPS stratum (Figure 7A–B), baseline clinical characteristics, including age, sex, and stage, were equally distributed between patients with high versus low Tfh cell abundance. However, splenic involvement was more frequent in patients with high Tfh (70% vs 36%) (Supplementary Table S8). The Tfh association was not confirmed in the Subramanian cohort (HR 1.43, 95% CI 0.68–3.04, adjusted for LP-IPS). In that cohort, the LP-IPS itself was not prognostic either (Supplementary Table S11). All four transformations in our cohort occurred in Tfh-high patients.

These results indicate that NLPHL and THRLBCL represent a transcriptional and microenvironmental continuum, and that NLPHL cases with THRLBCL-like microenvironmental and clinical features (splenic involvement, high LP-IPS) are associated with poorer outcomes, whereas the remaining NLPHL cases generally follow a favorable course.

DISCUSSION

In this study, we systematically analyzed a cohort of 200 patients with NLPHL and THRLBCL, providing a comprehensive evaluation of gene expression profiles, immune microenvironment composition, and their associations with clinical and pathological characteristics. Centralized pathology review conducted by expert hematopathologists ensured consistent classification and accurate diagnosis across cases, emphasizing the critical role of international collaboration in studying rare hematological malignancies.

Comparative molecular profiling that jointly examined both NLPHL and THRLBCL has historically been limited to small cohorts. Van Loo et al. analyzed a microarray

dataset comprising 10 NLPHL and 10 THRLBCL cases and reported transcriptional features consistent with a more tolerogenic host immune response in THRLBCL[20]. A relationship between NLPHL and THRLBCL has been proposed repeatedly. The morphological overlap was recognized early, with the description of a diffuse T-cell/histiocyte-rich pattern within NLPHL[8] and of NLPHL with nodules resembling THRLBCL[45]. The LP cells are a clonal population of germinal-center-derived B cells with somatically mutated immunoglobulin genes[50]. Studies of the malignant cells have shown that microdissected LP cells resemble tumor cells of THRLBCL[14], that their expression profiles are highly similar[15], and that they share genomic imbalances [16]. The same genes are recurrently mutated in both entities[17,19], as whole-genome sequencing has recently confirmed[18]. An earlier comparative genomic hybridization study, however, reported distinguishable patterns[46]. Clonally related transformation from NLPHL to aggressive B-cell lymphoma has been demonstrated in a subset of patients[47]. Graded microenvironmental alterations across this spectrum have also been described[48].

Recently, Subramanian et al.[21] presented a large NLPHL-focused multimodal atlas that combined bulk RNA-seq with digital deconvolution and EcoTyper[44] to derive 34 cell states across 14 lineages, which organize into three lymphocyte-predominant ecotypes (LPE1–3). LPE3 was associated with worse freedom-from-progression after single-modality therapy but not after combined-modality therapy, and incorporating LPE3 into LP-IPS improved relapse prediction[21]. In contrast, our study was explicitly designed to quantify and model the transcriptional relationship between NLPHL (across IAP categories) and THRLBCL, and to link this molecular continuum to clinically actionable risk features, including splenic involvement and LP-IPS. The two studies are complementary, and the published Subramanian data allowed an external test of our signature.

Our cohort enabled us to investigate the transcriptional divergence between NLPHL and THRLBCL, highlighting distinct molecular signatures that define each entity. NLPHL was characterized by strong B-cell–related gene expression and immune pathways involving naïve and helper T cells. In contrast, THRLBCL displayed pronounced macrophage-, monocyte-, and cytotoxic T-cell–associated signatures,

consistent with its immunologically active and histiocytic-rich TME. Furthermore, variant NLPHL cases showed intermediate transcriptional profiles, reinforcing a transcriptional continuum from classic NLPHL to THRLBCL. To capture this spectrum quantitatively, we derived a robust immune/B-cell signature score that showed strong external validity in independent public datasets, maintaining its discriminative performance without refitting or reweighting. In the Van Loo et al. cohort[20], the score separated NLPHL from THRLBCL (ROC-AUC = 1.00), and in the Subramanian et al. cohort[21] it discriminated variant from classic NLPHL (ROC-AUC = 0.80). The score tracks clinical risk features but was not itself associated with outcome in either cohort (Supplementary Table S11). These transcriptional programs are therefore reproducible across platforms and cohorts and could aid classification in diagnostically challenging cases.

Our examination of the tumor microenvironment provided further insights into cellular compositions, with classic NLPHL showing a predominant B cell-rich TME, while variant NLPHL and THRLBCL demonstrated increasing infiltration of activated T cells and macrophages. Notably, NLPHL cases with splenic involvement exhibited significant transcriptional and TME overlap with THRLBCL, supporting the clinical observation that splenic involvement indicates a high-risk phenotype.

Clinically, our analyses confirmed the prognostic significance of the LP-IPS score and disease stage in NLPHL, and showed that these clinical factors closely align with the molecular features typical of THRLBCL. Notably, NLPHL patients with a higher proportion of Tfh cells had shorter EFS in an exploratory analysis. This is paradoxical, as Tfh cells are typically associated with preserved nodular architecture in NLPHL and are virtually absent in THRLBCL. One explanation is that these cells are functionally exhausted rather than supportive, since patients with high Tfh cell levels co-expressed immune checkpoint genes including *TOX, PDCD1, CTLA4,* and *TIGIT*. The association did not, however, survive correction for multiple testing and could not be separated from splenic involvement (Supplementary Table S10). It was not confirmed in the Subramanian cohort, which has a progression-only endpoint and shorter follow-up. Whether immune cell composition adds prognostic information

to the LP-IPS therefore remains open and needs testing in larger cohorts with harmonized endpoints.

We observed higher expression of checkpoint genes in THRLBCL and in high-risk NLPHL, including patients with elevated Tfh cell proportions and high LP-IPS. Notably, *HAVCR2* (TIM-3) and *LAG3* were progressively upregulated across NLPHL IAP categories toward THRLBCL. *IDO1* and *PDCD1LG2* were most prominently enriched in THRLBCL, supporting a gradient of increasing immune checkpoint activity along the biological spectrum. Gunawardana et al. similarly reported higher expression of immune checkpoint genes in NLPHL than in classic Hodgkin lymphoma using NanoString[22]. Taken together, these data support checkpoint-associated microenvironment states in NLPHL and motivate future studies to define their cellular sources and spatial organization.

While the strengths of our study include a large multicenter cohort, centralized pathology review, and standardized sequencing at a single center, we also recognize some limitations. Incomplete clinical data for some cases, treatment heterogeneity across institutions, and the inclusion of specimens collected at relapse or at an unconfirmed time point rather than exclusively at initial diagnosis limited our ability to comprehensively analyse clinical outcomes. Furthermore, the relatively small number of clinical events, particularly within NLPHL, restricted our capacity to robustly associate specific gene expression patterns with clinical outcomes such as death, relapse, or transformation. In addition, because our profiling is based on bulk transcriptomics and in silico deconvolution, we cannot directly resolve the spatial localization of checkpoint-expressing cells or infer functional immune-checkpoint activity. Because the malignant cells are sparse, the bulk transcriptome mostly reflects the microenvironment. Whether the malignant cells of NLPHL and THRLBCL are the same cannot be settled with our data. Thus, our results motivate subsequent spatially resolved and functional validation studies. Finally, the cross-sectional design precluded longitudinal profiling of individual patients across disease states.

Nonetheless, our findings deepen the biological understanding of NLPHL and THRLBCL, highlighting a transcriptional continuum between these entities and identifying splenic involvement as a potential indicator of aggressive disease biology. To date, THRLBCL has been considered a distinct entity[1,2]. However, our data show that the microenvironment draws no sharp line between the two. This is in keeping with the clinical picture, in which a THRLBCL-like pattern occurs within NLPHL and THRLBCL can arise after NLPHL. Whether THRLBCL should be regarded as a subset of NLPHL is for studies of the malignant cells[14–19,46] and for the classification bodies[49] to decide. Indeed, the International Consensus Classification has already renamed NLPHL as NLPBL, citing its close relationship to THRLBCL as one of the reasons[1].

Taken together, this study provides a systematic quantitative characterization of the transcriptional continuum across all NLPHL IAP categories and THRLBCL. It also derives a molecular signature of this spectrum that is reproducible across independent cohorts and sequencing platforms. Further studies, ideally prospective and incorporating standardized treatment protocols, are warranted to test whether immune cell composition adds prognostic information to the LP-IPS and to explore the clinical utility of the signature. Our work also highlights the importance of large-scale international collaborations in advancing research on rare lymphomas, laying a foundation for future translational and clinical studies.

REFERENCES

1. Campo E, Jaffe ES, Cook JR, et al. The international consensus classification of mature lymphoid neoplasms: a report from the clinical advisory committee. Blood. 2022;140(11):1229-1253.

2. Alaggio R, Amador C, Anagnostopoulos I, et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms. Leukemia. 2022;36(7):1720-1748.

3. Nogová L, Reineke T, Brillant C, et al. Lymphocyte-predominant and classical Hodgkin's lymphoma: a comprehensive analysis from the German Hodgkin Study Group. J Clin Oncol. 2008;26(3):434-439.

4. Posthuma HLA, Zijlstra JM, Visser O, Lugtenburg PJ, Kersten MJ, Dinmohamed AG. Primary therapy and survival among patients with nodular lymphocyte-predominant Hodgkin lymphoma: a population-based analysis in the Netherlands, 1993-2016. Br J Haematol. 2020;189(1):117-121.

5. Kalashnikov I, Tanskanen T, Pitkäniemi J, Malila N, Jyrkkiö S, Leppä S. Transformation and outcome of nodular lymphocyte predominant Hodgkin lymphoma: a Finnish Nationwide population-based study. Blood Cancer J. 2021;11(12):203.

6. Eichenauer DA, Plütschow A, Schröder L, et al. Relapsed and refractory nodular lymphocyte-predominant Hodgkin lymphoma: an analysis from the German Hodgkin Study Group. Blood. 2018;132(14):1519-1525.

7. Al-Mansour M, Connors JM, Gascoyne RD, Skinnider B, Savage KJ. Transformation to aggressive lymphoma in nodular lymphocyte-predominant Hodgkin's lymphoma. J Clin Oncol. 2010;28(5):793-799.

8. Fan Z, Natkunam Y, Bair E, Tibshirani R, Warnke RA. Characterization of variant patterns of nodular lymphocyte predominant hodgkin lymphoma with immunohistologic and clinical correlation. Am J Surg Pathol. 2003;27(10):1346-1356.

9. Carbone A, Gloghini A. "Intrafollicular neoplasia" of nodular lymphocyte predominant Hodgkin lymphoma: description of a hypothetic early step of the disease. Hum Pathol. 2012;43(5):619-628.

10. Hartmann S, Eichenauer DA, Plütschow A, et al. The prognostic impact of variant histology in nodular lymphocyte-predominant Hodgkin lymphoma: a report from the German Hodgkin Study Group (GHSG). Blood. 2013;122(26):4246-4252.

11. Hartmann S, Plütschow A, Mottok A, et al. The time to relapse correlates with the histopathological growth pattern in nodular lymphocyte predominant Hodgkin lymphoma. Am J Hematol. 2019;94(11):1208-1213.

12. Eichenauer DA, Basaran A, Bühnen I, et al. Refining histopathological growth pattern-based risk group discrimination in nodular lymphocyte-predominant Hodgkin lymphoma: an analysis from the German Hodgkin Study Group. Leukemia. 2025;39(7):1735-1743.

13. Binkley MS, Flerlage JE, Savage KJ, et al. International prognostic score for nodular lymphocyte-predominant hodgkin lymphoma. J Clin Oncol. 2024;42(19):2271-2280.

14. Brune V, Tiacci E, Pfeil I, et al. Origin and pathogenesis of nodular lymphocyte-predominant Hodgkin lymphoma as revealed by global gene expression analysis. J Exp Med. 2008;205(10):2251-2268.

15. Hartmann S, Döring C, Jakobus C, et al. Nodular lymphocyte predominant hodgkin lymphoma and T cell/histiocyte rich large B cell lymphoma--endpoints of a spectrum of one disease? PLoS One. 2013;8(11):e78812.

16. Hartmann S, Döring C, Vucic E, et al. Array comparative genomic hybridization reveals similarities between nodular lymphocyte predominant Hodgkin lymphoma and T cell/histiocyte rich large B cell lymphoma. Br J Haematol. 2015;169(3):415-422.

17. Schuhmacher B, Bein J, Rausch T, et al. JUNB, DUSP2, SGK1, SOCS1 and CREBBP are frequently mutated in T-cell/histiocyte-rich large B-cell lymphoma. Haematologica. 2019;104(2):330-337.

18. Rausch T, Schäfer H, Bein J, et al. Genetic lesions in nodular lymphocyte-predominant Hodgkin lymphoma and T cell/histiocyte-rich large B-cell lymphoma identified by whole genome sequencing. Leukemia. 2025;39(9):2215-2225.

19. Hartmann S, Schuhmacher B, Rausch T, et al. Highly recurrent mutations of SGK1, DUSP2 and JUNB in nodular lymphocyte predominant Hodgkin lymphoma. Leukemia. 2016;30(4):844-853.

20. Van Loo P, Tousseyn T, Vanhentenrijk V, et al. T-cell/histiocyte-rich large B-cell lymphoma shows transcriptional features suggestive of a tolerogenic host immune response. Haematologica. 2010;95(3):440-448.

21. Subramanian A, Su S, Flerlage J, et al. Distinct cell state ecosystems for nodular lymphocyte-predominant Hodgkin lymphoma. Nat Commun. 2025;16(1):8473.

22. Gunawardana J, Law SC, Sabdia MB, et al. Intra-tumoral and peripheral blood TIGIT and PD-1 as immune biomarkers in nodular lymphocyte predominant Hodgkin lymphoma. Am J Hematol. 2024;99(11):2096-2107.

23. Weiss J, Gibbons K, Ehyaee V, et al. Specific polo-like kinase 1 expression in nodular lymphocyte-predominant hodgkin lymphoma suggests an intact immune surveillance program. Am J Pathol. 2024;194(1):165-178.

24. Younes S, Subramanian A, Khan A, Zhao S, Binkley M, Natkunam Y. Spatial phenotyping of nodular lymphocyte predominant Hodgkin lymphoma and T-cell/histiocyte-rich large B-cell lymphoma. Blood Cancer J. 2024;14(1):92.

25. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-2120.

26. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.

27. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25(16):2078-2079.

28. McKenna A, Hanna M, Banks E, et al. The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297-1303.

29. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14(4):417-419.

30. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res. 2015;4:1521.

31. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2025.

32. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

33. Hao Y, Yan M, Heath BR, Lei YL, Xie Y. Fast and robust deconvolution of tumor infiltrating lymphocyte from expression profiles using least trimmed squares. PLoS Comput Biol. 2019;15(5):e1006976.

34. Newman AM, Steen CB, Liu CL, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat Biotechnol. 2019;37(7):773-782.

35. Afsari B, Fertig EJ, Geman D, Marchionni L. switchBox: an R package for k-top scoring pairs classifier development. Bioinformatics. 2015;31(2):273-274.

36. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16(5):284-287.

37. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol. 2017;18(1):220.

38. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12(5):453-457.

39. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database (MSigDB) hallmark gene set collection. Cell Syst. 2015;1(6):417-425.

40. Kolde R. pheatmap: Pretty Heatmaps. R package; 2025.

41. Blighe K, Rana S, Lewis M. EnhancedVolcano: publication-ready volcano plots with enhanced coloring and labeling. R package; 2023.

42. Therneau TM. A package for survival analysis in R. R package version 3.7; 2024.

43. Kassambara A, Kosinski M, Biecek P, Fabian S. survminer: Drawing Survival Curves using 'ggplot2'. R package; 2025.

44. Steen CB, Luca BA, Esfahani MS, et al. The landscape of tumor cell states and ecosystems in diffuse large B cell lymphoma. Cancer Cell. 2021;39(10):1422-1437.e10.

45. Boudová L, Torlakovic E, Delabie J, et al. Nodular lymphocyte-predominant Hodgkin lymphoma with nodules resembling T-cell/histiocyte-rich B-cell lymphoma: differential diagnosis between nodular lymphocyte-predominant Hodgkin lymphoma and T-cell/histiocyte-rich B-cell lymphoma. Blood. 2003;102(10):3753-3758.

46. Franke S, Wlodarska I, Maes B, et al. Comparative genomic hybridization pattern distinguishes T-cell/histiocyte-rich B-cell lymphoma from nodular lymphocyte predominance Hodgkin's lymphoma. Am J Pathol. 2002;161(5):1861-1867.

47. Paschold L, Willscher E, Bein J, et al. Evolutionary clonal trajectories in nodular lymphocyte-predominant Hodgkin lymphoma with high risk of transformation. Haematologica. 2021;106(10):2654-2666.

48. Panayi C, Akarca AU, Ramsay AD, et al. Microenvironmental immune cell alterations across the spectrum of nodular lymphocyte predominant Hodgkin lymphoma and T-cell/histiocyte-rich large B-cell lymphoma. Front Oncol. 2023;13:1267604.

49. Salvaris RT, Allanson BM, Collins G, Cheah C. Nodular lymphocyte-predominant Hodgkin lymphoma: advances in disease biology, risk stratification, and treatment. Haematologica. 2024;109(11):3476-3487.

50. Braeuninger A, Küppers R, Strickler JG, Wacker HH, Rajewsky K, Hansmann ML. Hodgkin and Reed-Sternberg cells in lymphocyte predominant Hodgkin disease represent clonal populations of germinal center-derived tumor B cells. Proc Natl Acad Sci U S A. 1997;94(17):9337-9342.

TABLES

Table 1. Patient demographics

Characteristic	NLPHL (N = 178)	THRLBCL (N = 22)
Age at diagnosis, median (IQR), years	42.0 (27.1-54.9), N=165	55.4 (37.4-65.9), N=17
Missing, n	13/178	5/22
Sex		
Male	117/165 (70.9%)	16/19 (84.2%)
Female	48/165 (29.1%)	3/19 (15.8%)
Missing, n	13/178	3/22
Stage		
Limited	97/143 (67.8%)	2/18 (11.1%)
Advanced	46/143 (32.2%)	16/18 (88.9%)
Missing, n	35/178	4/22
Splenic involvement		
Yes	18/149 (12.1%)	8/15 (53.3%)
No	131/149 (87.9%)	7/15 (46.7%)
Missing, n	29/178	7/22
LP-IPS		
High (≥2)	32/132 (24.2%)	—†
Low (0-1)	100/132 (75.8%)	—†
Missing, n	46/178	—†
IAP		
A	58/140 (41.4%)	—†
B	28/140 (20.0%)	—†
C	21/140 (15.0%)	—†
D	16/140 (11.4%)	—†
E	13/140 (9.3%)	—†
F	4/140 (2.9%)	—†
Missing, n	38/178	—†
First-line therapy		
Watch and wait	16/135 (11.9%)	0/16 (0.0%)
Rituximab	9/135 (6.7%)	0/16 (0.0%)
Radiotherapy	35/135 (25.9%)	0/16 (0.0%)
Chemotherapy	48/135 (35.6%)	15/16 (93.8%)
Combined modality treatment	27/135 (20.0%)	1/16 (6.3%)
Missing, n	43/178	6/22
Patient status		
Alive	130/140 (92.9%)	8/15 (53.3%)
Dead	10/140 (7.1%)	7/15 (46.7%)
Missing, n	38/178	7/22

Values are n/N (%) with N the patients with available data, or median (Q1–Q3). †Defined for NLPHL only. IAP, immuno-architectural pattern. LP-IPS, Lymphocyte-Predominant International Prognostic Score.

FIGURE LEGENDS

Figure 1. Survival outcomes and covariates in NLPHL and THRLBCL according to stage and splenic involvement.

Overall survival (A) and event-free survival (B) for patients diagnosed with NLPHL and THRLBCL.

Overall survival (C) and event-free survival (D) stratified by limited versus advanced-stage disease in NLPHL patients.

Overall survival (E) and event-free survival (F) stratified by splenic involvement in NLPHL patients.

Figure 2. Survival outcomes and covariates in NLPHL and THRLBCL according to LP-IPS and IAP categories.

Overall survival (A) and event-free survival (B) stratified by LP-IPS category (low risk, 0–1 point versus high risk, ≥2 points) in NLPHL patients.

Overall survival (C) and event-free survival (D) stratified by IAP category in NLPHL patients.

Figure 3. Gene expression profiles differentiate between NLPHL and THRLBCL.

(A) Principal component analysis of NLPHL (green circles) and THRLBCL cases (red squares). The dashed line is the 95% concentration ellipse of the NLPHL cases.

(B) Volcano plot depicting differentially expressed genes between NLPHL and THRLBCL cases. Selected genes with absolute log2 fold change ≥1.5 and adjusted $P < 0.05$ are highlighted.

(C) Heatmap of the differentially expressed genes of NLPHL and THRLBCL cases. NLPHLs are separated by individual IAPs. Clinical annotations are displayed at the top.

Figure 4. Gene expression heterogeneity between NLPHL IAPs and THRLBCL.

(A) Heatmap displaying t-statistics from pairwise comparisons between successive IAP categories, between IAP groupings, and between NLPHL and THRLBCL. Each column represents a comparison contrast. Positive t-statistics indicate upregulation in the second group, and negative t-statistics indicate downregulation. Selected representative genes are highlighted. The t-statistics of all 185 genes are provided in Supplementary Table S2B.

(B-C) Overrepresentation analysis plot from gene set enrichment analysis for upregulated (B) and downregulated (C) genes.

Figure 5. Immune signatures in NLPHL and THRLBCL.

(A) Heatmap displaying expression of genes that reflect innate immunity ("Immune Signature") and B cell-mediated immunity ("B-cell Signature"). Genes included in each signature are indicated to the right. The individual IAPs and categories are indicated at the top.

(B-C) Box plots showing TSP score by individual IAP and THRLBCL on the training (C) and test sets (D).

(D-E) Box plots showing TSP score by IAP category and THRLBCL on the training (D) and test sets (E).

All pairwise comparisons are significant at $p<0.05$ unless otherwise indicated.

Figure 6. Tumor microenvironment of NLPHL and THRLBCL.

(A) Heatmap displaying tumor microenvironment composition as calculated from semi-supervised in silico cell-type deconvolution using the LM22 immune signature matrix. Cell types are shown on the right. Clinical annotations are displayed at the top.

(B-D) Box plots displaying the relative proportion of the indicated cell types in each IAP category and THRLBCL. All pairwise comparisons are significant at $p<0.05$ unless otherwise indicated. NS, not significant.

Figure 7. Survival of NLPHL patients according to LP-IPS category and follicular helper T-cell abundance.

(A-B) Overall survival (A) and event-free survival (B) curves for patients diagnosed with NLPHL stratified by LP-IPS category and the relative follicular T-helper cell abundance. p, log-rank test across the four groups. The Tfh contrast within each LP-IPS stratum is exploratory (Supplementary Table S10).

Figure 1

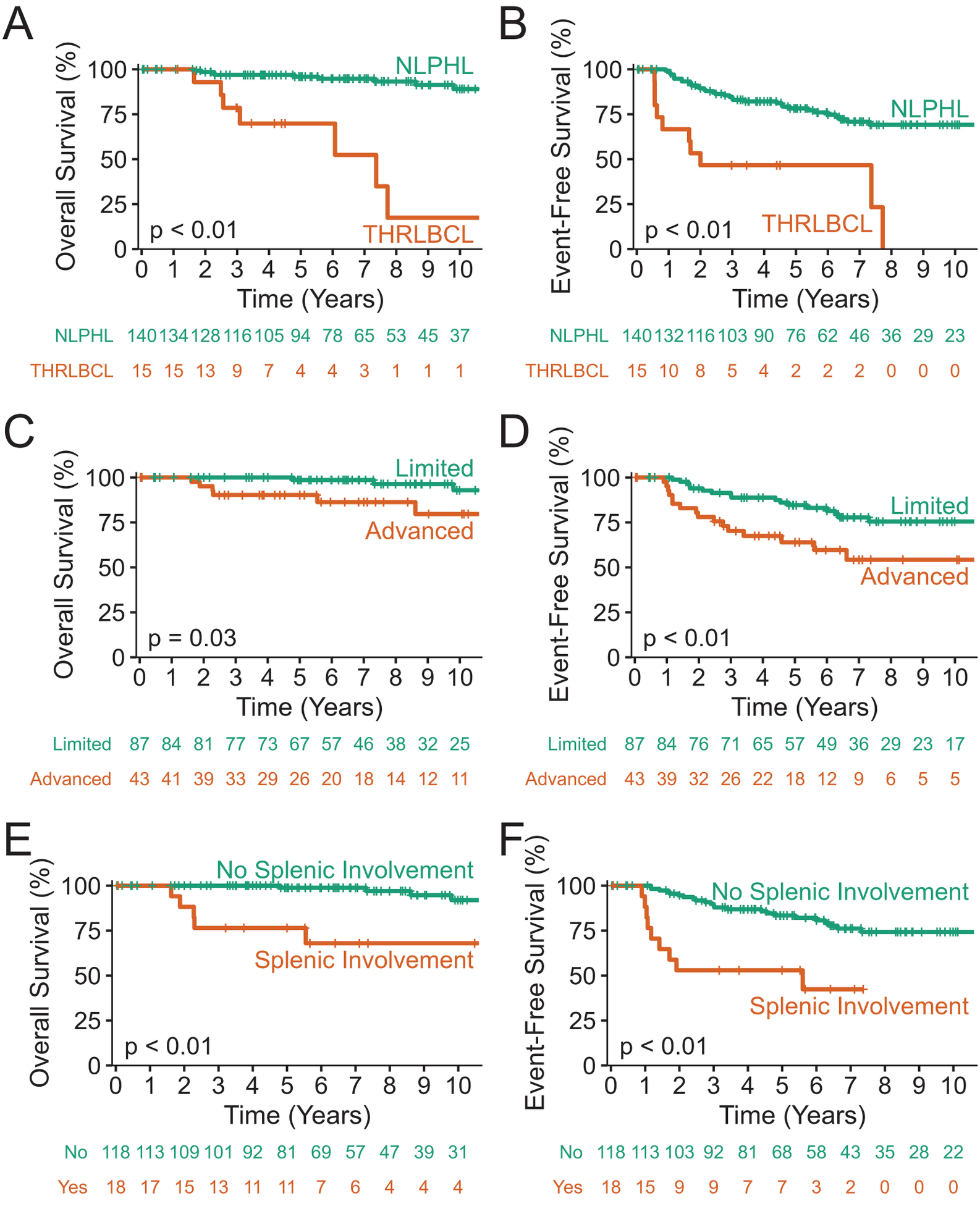

A
Overall Survival (%)
100
75
50
25
0
NLPHL
THRLBCL
p < 0.01
0 1 2 3 4 5 6 7 8 9 10
Time (Years)
NLPHL 140 134 128 116 105 94 78 65 53 45 37
THRLBCL 15 15 13 9 7 4 4 3 1 1 1
B
Event-Free Survival (%)
NLPHL
THRLBCL
p < 0.01
Time (Years)
NLPHL 140 132 116 103 90 76 62 46 36 29 23
THRLBCL 15 10 8 5 4 2 2 2 0 0 0
C
Overall Survival (%)
Limited
Advanced
p = 0.03
Time (Years)
Limited 87 84 81 77 73 67 57 46 38 32 25
Advanced 43 41 39 33 29 26 20 18 14 12 11
D
Event-Free Survival (%)
Limited
Advanced
p < 0.01
Time (Years)
Limited 87 84 76 71 65 57 49 36 29 23 17
Advanced 43 39 32 26 22 18 12 9 6 5 5
E
Overall Survival (%)
No Splenic Involvement
Splenic Involvement
p < 0.01
Time (Years)
No 118 113 109 101 92 81 69 57 47 39 31
Yes 18 17 15 13 11 11 7 6 4 4 4
F
Event-Free Survival (%)
No Splenic Involvement
Splenic Involvement
p < 0.01
Time (Years)
No 118 113 103 92 81 68 58 43 35 28 22
Yes 18 15 9 9 7 7 3 2 0 0 0

Figure 2

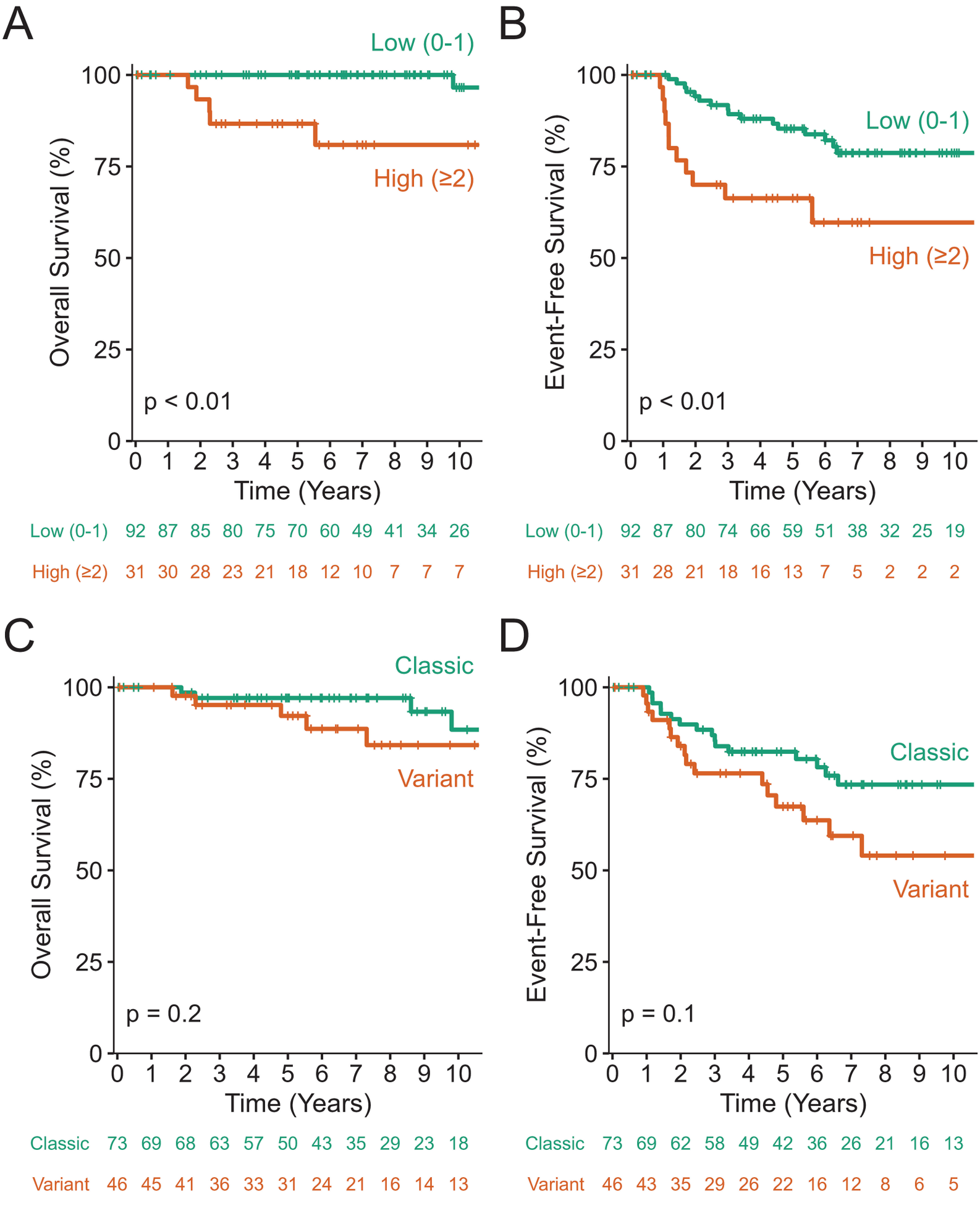

A
Low (0-1)
High (≥2)
Overall Survival (%)
100
75
50
25
0
p < 0.01
0 1 2 3 4 5 6 7 8 9 10
Time (Years)
Low (0-1) 92 87 85 80 75 70 60 49 41 34 26
High (≥2) 31 30 28 23 21 18 12 10 7 7 7
B
Low (0-1)
High (≥2)
Event-Free Survival (%)
100
75
50
25
0
p < 0.01
0 1 2 3 4 5 6 7 8 9 10
Time (Years)
Low (0-1) 92 87 80 74 66 59 51 38 32 25 19
High (≥2) 31 28 21 18 16 13 7 5 2 2 2
C
Classic
Variant
Overall Survival (%)
100
75
50
25
0
p = 0.2
0 1 2 3 4 5 6 7 8 9 10
Time (Years)
Classic 73 69 68 63 57 50 43 35 29 23 18
Variant 46 45 41 36 33 31 24 21 16 14 13
D
Classic
Variant
Event-Free Survival (%)
100
75
50
25
0
p = 0.1
0 1 2 3 4 5 6 7 8 9 10
Time (Years)
Classic 73 69 62 58 49 42 36 26 21 16 13
Variant 46 43 35 29 26 22 16 12 8 6 5

Figure 3

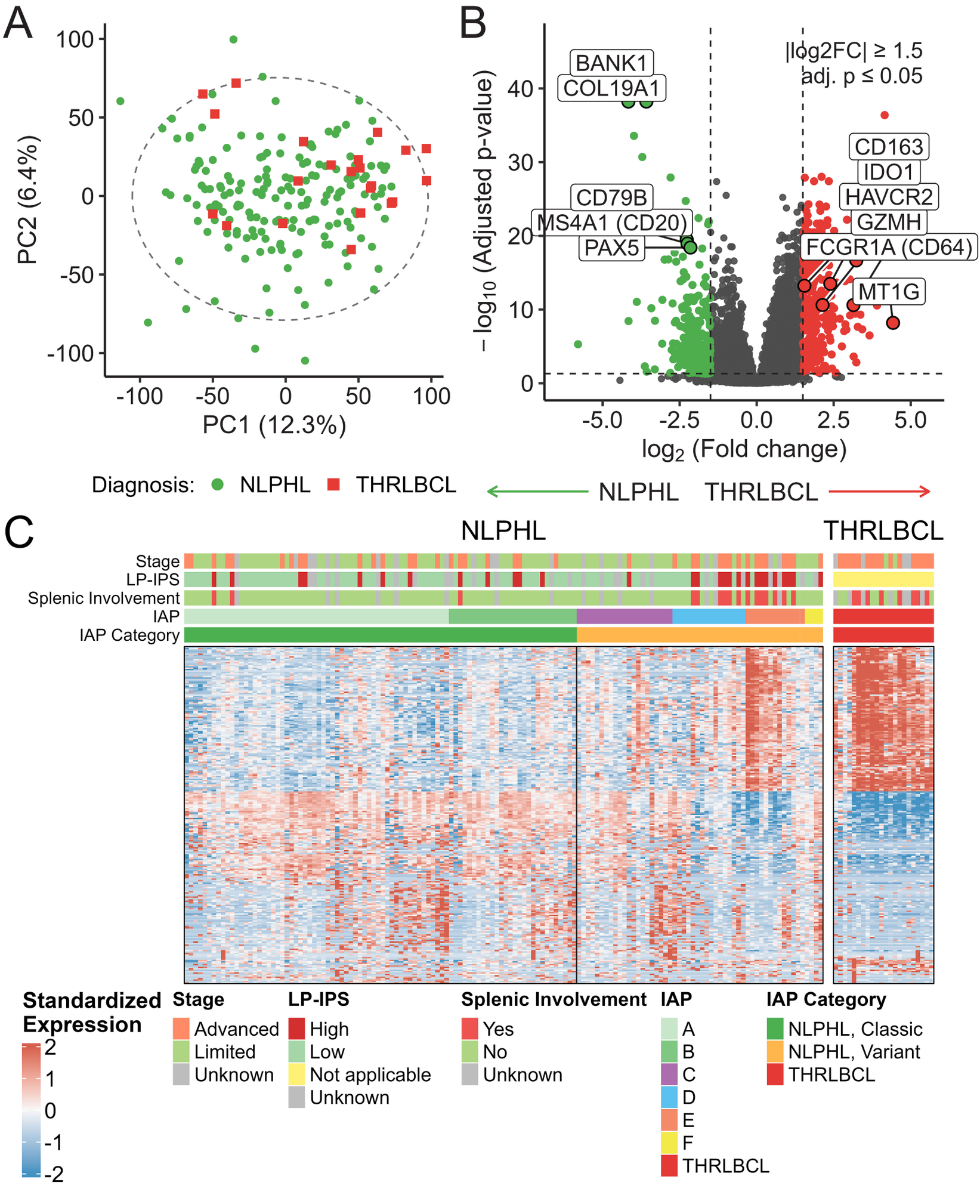

A
PC2 (6.4%)
PC1 (12.3%)
Diagnosis: NLPHL THRLBCL
B
BANK1
COL19A1
CD79B
MS4A1 (CD20)
PAX5
CD163
IDO1
HAVCR2
GZMH
FCGR1A (CD64)
MT1G
|log2FC| ≥ 1.5
adj. p ≤ 0.05
– log10 (Adjusted p-value)
log2 (Fold change)
NLPHL THRLBCL
C
NLPHL
THRLBCL
Stage
LP-IPS
Splenic Involvement
IAP
IAP Category
Standardized Expression
Stage
Advanced
Limited
Unknown
LP-IPS
High
Low
Not applicable
Unknown
Splenic Involvement
Yes
No
Unknown
IAP
A
B
C
D
E
F
THRLBCL
IAP Category
NLPHL, Classic
NLPHL, Variant
THRLBCL

Figure 4

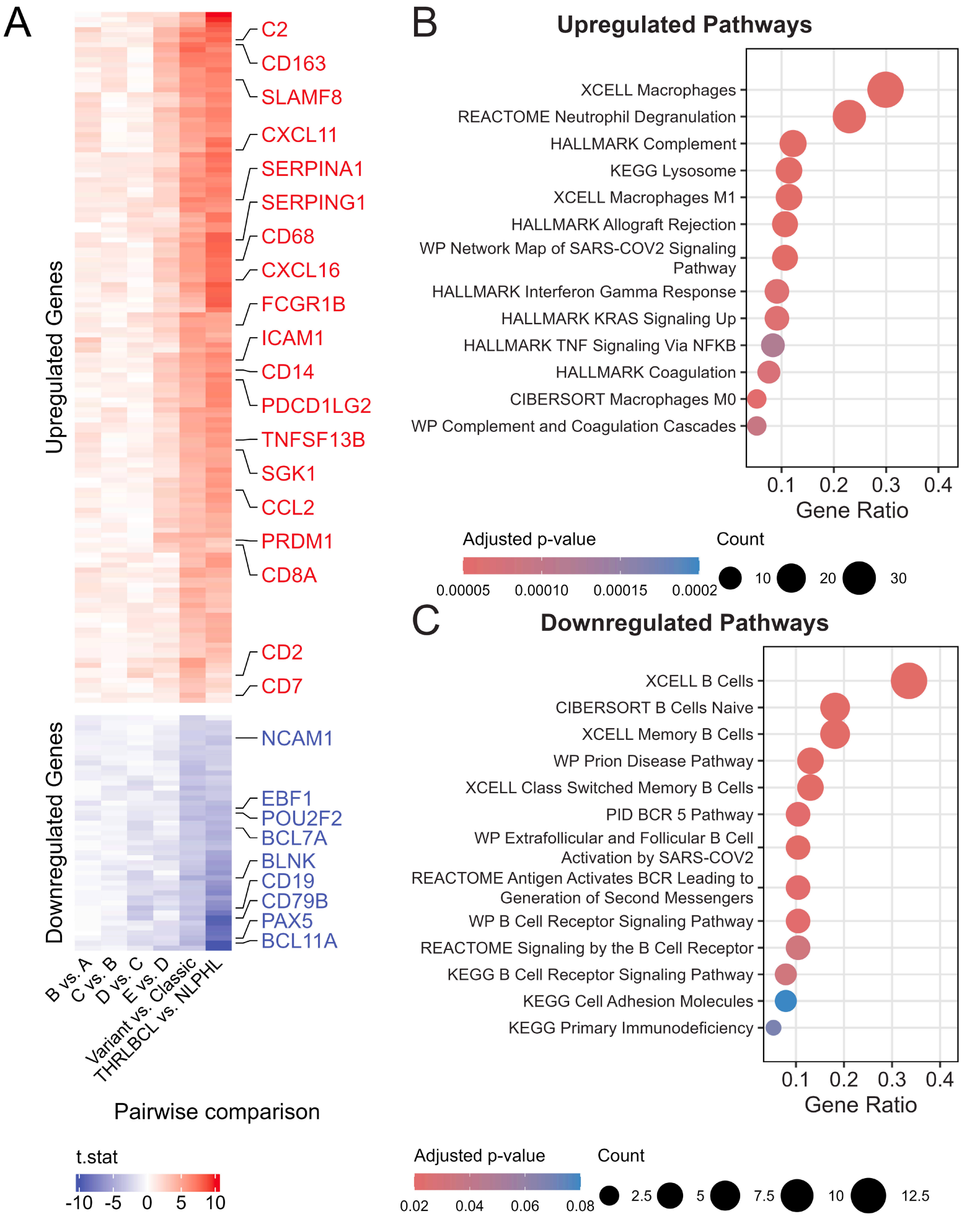

A
Upregulated Genes
C2
CD163
SLAMF8
CXCL11
SERPINA1
SERPING1
CD68
CXCL16
FCGR1B
ICAM1
CD14
PDCD1LG2
TNFSF13B
SGK1
CCL2
PRDM1
CD8A
CD2
CD7
Downregulated Genes
NCAM1
EBF1
POU2F2
BCL7A
BLNK
CD19
CD79B
PAX5
BCL11A
B vs. A
C vs. B
D vs. C
E vs. D
Classic vs. Variant
NLPHL vs. THRLBCL
Pairwise comparison
t.stat
-10 -5 0 5 10
B
Upregulated Pathways
XCELL Macrophages
REACTOME Neutrophil Degranulation
HALLMARK Complement
KEGG Lysosome
XCELL Macrophages M1
HALLMARK Allograft Rejection
WP Network Map of SARS-COV2 Signaling Pathway
HALLMARK Interferon Gamma Response
HALLMARK KRAS Signaling Up
HALLMARK TNF Signaling Via NFKB
HALLMARK Coagulation
CIBERSORT Macrophages M0
WP Complement and Coagulation Cascades
0.1 0.2 0.3 0.4
Gene Ratio
Adjusted p-value
0.00005 0.00010 0.00015 0.0002
Count
10 20 30
C
Downregulated Pathways
XCELL B Cells
CIBERSORT B Cells Naive
XCELL Memory B Cells
WP Prion Disease Pathway
XCELL Class Switched Memory B Cells
PID BCR 5 Pathway
WP Extrafollicular and Follicular B Cell Activation by SARS-COV2
REACTOME Antigen Activates BCR Leading to Generation of Second Messengers
WP B Cell Receptor Signaling Pathway
REACTOME Signaling by the B Cell Receptor
KEGG B Cell Receptor Signaling Pathway
KEGG Cell Adhesion Molecules
KEGG Primary Immunodeficiency
0.1 0.2 0.3 0.4
Gene Ratio
Adjusted p-value
0.02 0.04 0.06 0.08
Count
2.5 5 7.5 10 12.5

Figure 5

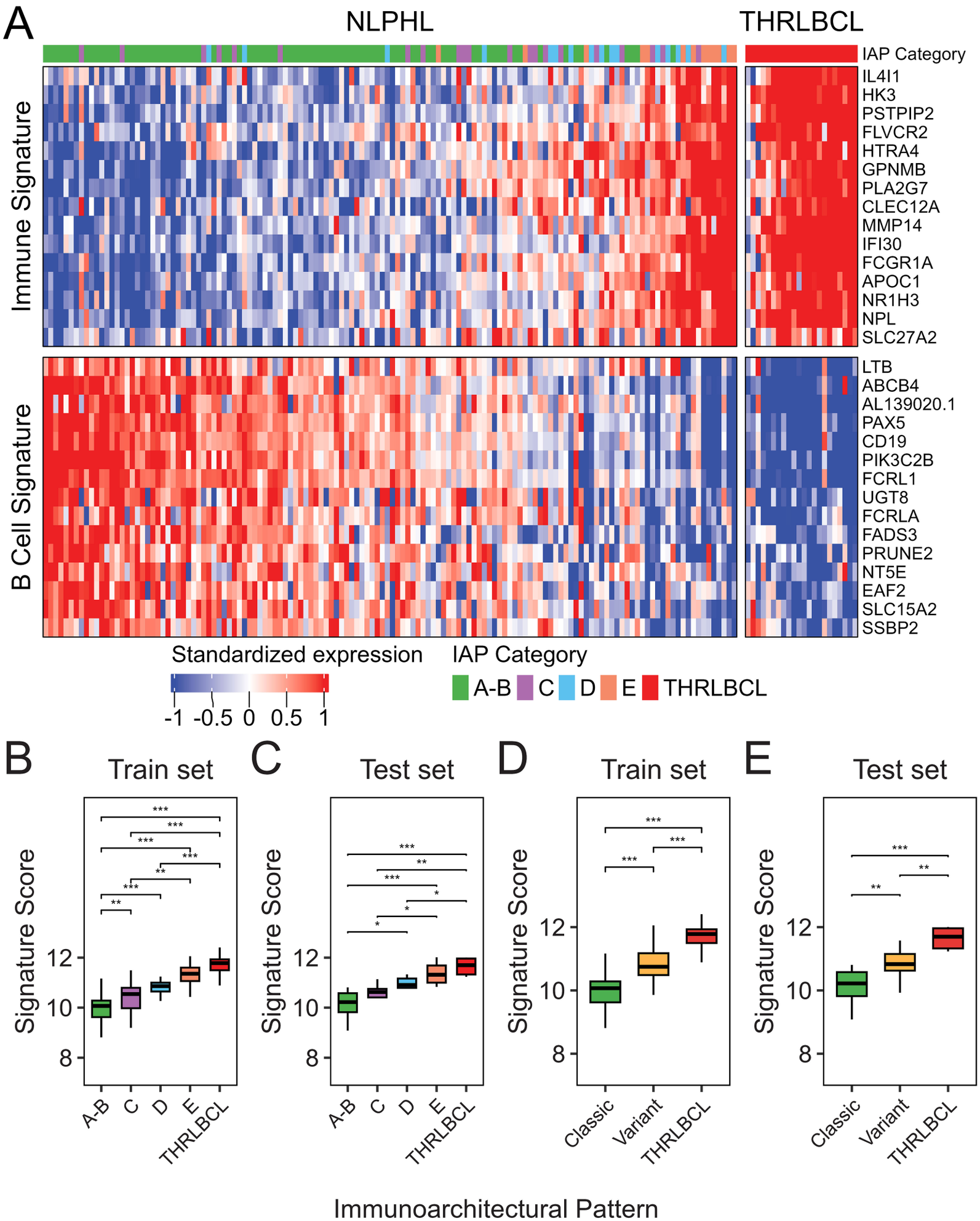

A
NLPHL
THRLBCL
IAP Category
Immune Signature
IL4I1
HK3
PSTPIP2
FLVCR2
HTRA4
GPNMB
PLA2G7
CLEC12A
MMP14
IFI30
FCGR1A
APOC1
NR1H3
NPL
SLC27A2
B Cell Signature
LTB
ABCB4
AL139020.1
PAX5
CD19
PIK3C2B
FCRL1
UGT8
FCRLA
FADS3
PRUNE2
NT5E
EAF2
SLC15A2
SSBP2
Standardized expression
-1 -0.5 0 0.5 1
IAP Category
A-B C D E THRLBCL
B
Train set
C
Test set
D
Train set
E
Test set
Signature Score
8 10 12
A-B C D E THRLBCL
Classic Variant THRLBCL
Immunoarchitectural Pattern

Figure 6

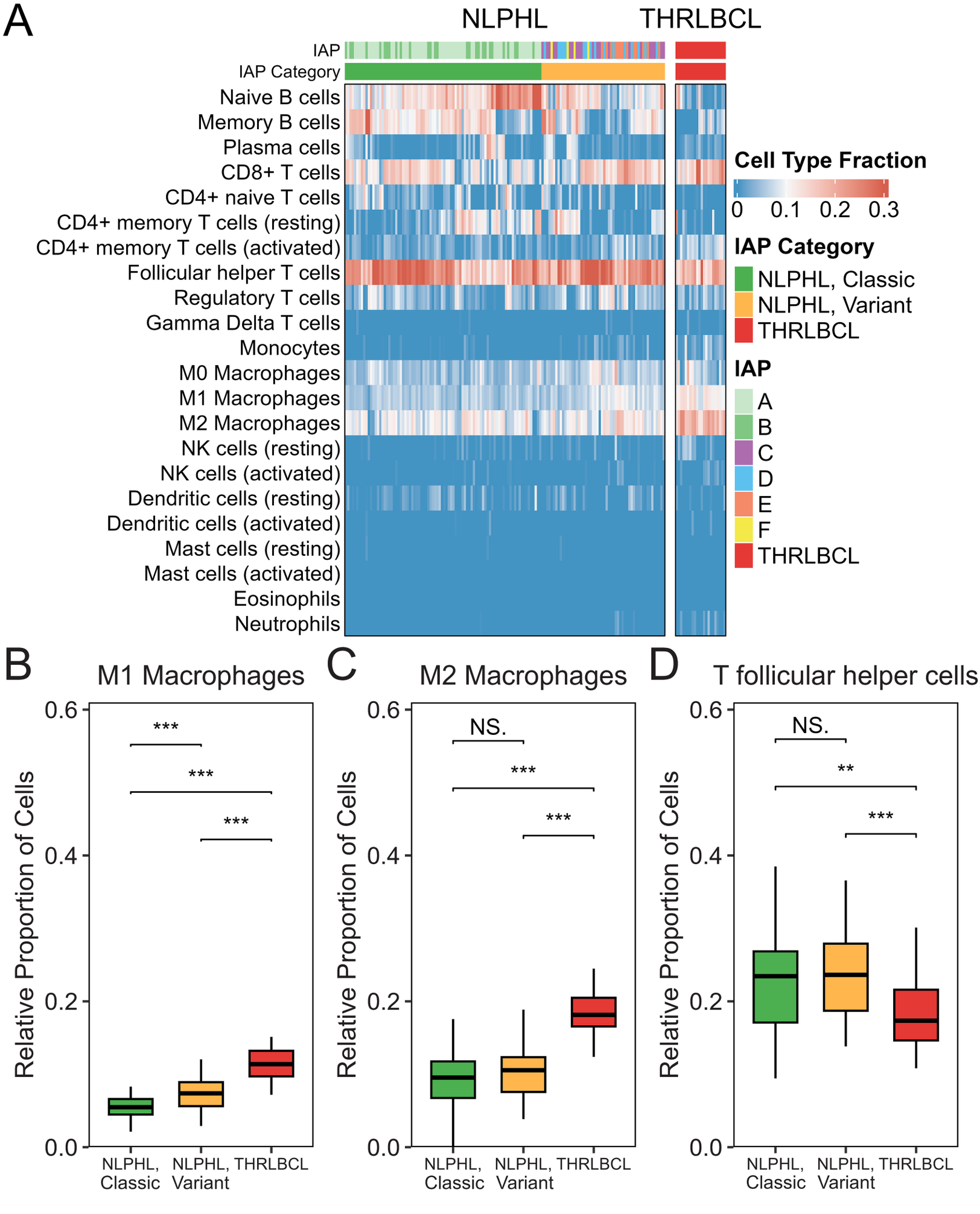
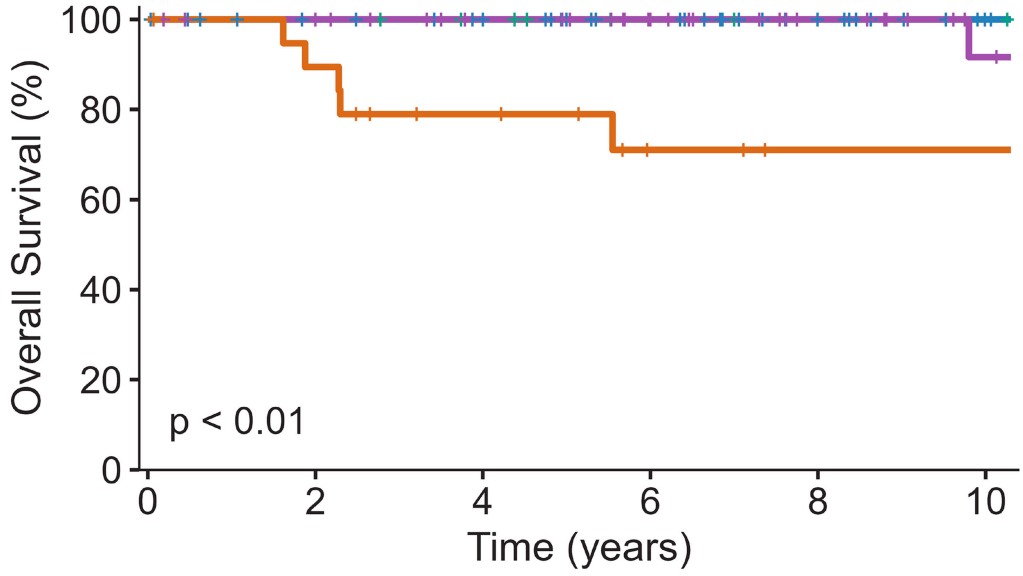

A
NLPHL
THRLBCL
IAP
IAP Category
Naive B cells
Memory B cells
Plasma cells
CD8+ T cells
CD4+ naive T cells
CD4+ memory T cells (resting)
CD4+ memory T cells (activated)
Follicular helper T cells
Regulatory T cells
Gamma Delta T cells
Monocytes
M0 Macrophages
M1 Macrophages
M2 Macrophages
NK cells (resting)
NK cells (activated)
Dendritic cells (resting)
Dendritic cells (activated)
Mast cells (resting)
Mast cells (activated)
Eosinophils
Neutrophils
Cell Type Fraction
0 0.1 0.2 0.3
IAP Category
NLPHL, Classic
NLPHL, Variant
THRLBCL
IAP
A
B
C
D
E
F
THRLBCL
B
M1 Macrophages
C
M2 Macrophages
D
T follicular helper cells
Relative Proportion of Cells
0.0
0.2
0.4
0.6
***
***
***
NS.
***
***
NS.
**
***
NLPHL, Classic
NLPHL, Variant
THRLBCL

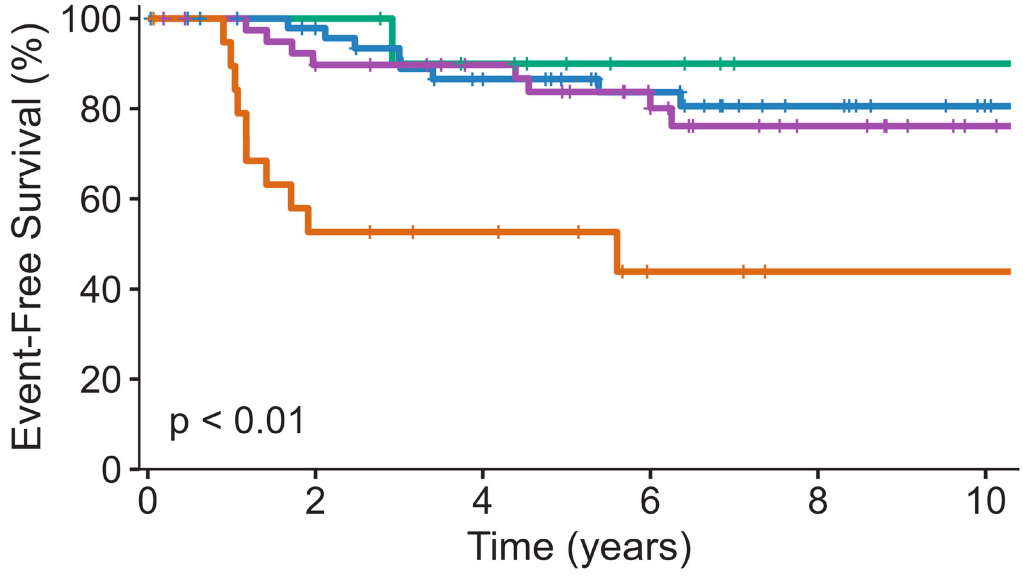
Figure 7

A



Tfh^{low} IPS^{high}	11	11	9	5	2	2
Tfh^{low} IPS^{low}	51	46	42	33	23	15
Tfh^{high} IPS^{low}	41	39	33	27	18	11
Tfh^{high} IPS^{high}	20	17	12	7	5	5

B



Tfh^{low} IPS^{high}	11	11	8	4	1	1
Tfh^{low} IPS^{low}	51	45	36	28	18	11
Tfh^{high} IPS^{low}	41	35	30	23	14	8
Tfh^{high} IPS^{high}	20	10	8	3	1	1

Transcriptional Landscape of NLPHL and THRLBCL: Tumor Microenvironment and Checkpoint Genes

Kalashnikov et al

Supplementary Material

Supplementary Methods

Sample collection, library preparation, and sequencing

The study cohort included formalin-fixed paraffin-embedded (FFPE) tumor samples from 178 patients with NLPHL and 22 patients with THRLBCL, collected across 24 institutions globally as part of the Atlas of Blood Cancer Genomes (ABCG) initiative. Tissue biopsies and clinical data were collected with approval from the local institutional review board at each center. A central panel of expert hematopathologists reviewed all cases to confirm the diagnosis of THRLBCL and to classify each NLPHL case into immuno-architectural patterns (IAP) A to F using established histologic criteria. Of the 178 NLPHL samples, 145 were obtained at initial diagnosis and 10 at relapse. Of the 22 THRLBCL samples, 14 were obtained at initial diagnosis and 6 at relapse. All THRLBCL cases were de novo, with no prior history of NLPHL documented in the clinical records. For the remaining samples, the timing of the biopsy relative to diagnosis could not be confirmed.

All nucleic acid extractions and sequencing library preparations were performed with the Duoseq Research Kit (EPXv3, Data Driven Bioscience, Durham, NC) in RNA mode, following the manufacturer's protocol. All libraries were sequenced on the Illumina NovaSeq X platform and analyzed according to manufacturer recommendations. Two hundred cases met the quality threshold and satisfied whole-transcriptome RNA sequencing (RNA-seq) quality standards.

RNA sequencing pipeline

FASTQ files containing RNA-seq reads were trimmed using Trimmomatic (v0.39) in paired-end mode to remove Illumina-specific adapter sequences and low-quality reads. The STAR aligner (v2.7.1) was used to map RNA reads to the human genome and transcriptome (GRCh38.p12, Ensembl release 98, with a PAR mask on chrY). Reads with at least one primary alignment to the human genome were extracted using samtools (v1.13) and further filtered based on manufacturer recommendations using GATK (v4.1.2.0). Transcript quantification was performed using salmon (v1.2.1), and the transcript-level data were summarized to the gene level using the tximport (v1.14.2) library in R. The result was a matrix representing the raw read counts for each gene per patient sample. This matrix was restricted to include only the expression of the protein-coding genes captured by exome sequencing. Genes with missing or uniformly low expression were excluded from further analysis, and the remaining expressed genes were normalized across the entire sample set in R using the DESeq2 (v1.40.2) library. Immune-cell fractions were estimated using FARDEEP with the LM22 leukocyte gene-signature matrix, as implemented for CIBERSORT and CIBERSORTx deconvolution.

Expression-based signature construction

NLPHL samples labeled with IAP A through E, along with THRLBCL samples, were each split into 75/25 train-test sets, stratified to maintain balanced spleen involvement across splits. Differential expression analysis was conducted using DESeq2 on the training set for the following sequential IAP comparisons: A versus B, B versus C, C versus D, D versus E, and E versus THRLBCL. Additional grouped comparisons were performed between patterns A-B and C-E, as well as between all NLPHL and THRLBCL training samples. A universe of 185 genes was defined by first selecting genes that were differentially expressed (adjusted P value <0.1 and |log2 fold change| >0.5) in both the A-B versus C-E and the NLPHL versus THRLBCL comparisons, or that were included in the LM22 gene list. This gene set was then further restricted to those with log2 fold changes that consistently moved in the same direction across successive comparisons spanning IAP A through E to THRLBCL. This universe was

used to select 15 top-scoring pairs (TSP), comprising 30 genes capable of distinguishing patterns A-B and C-E using the R library switchBox. For this step, samples were classified into high-risk and low-risk groups. Samples with spleen involvement, or IAP C through E, or THRLBCL were defined as high-risk, and IAP A and B with no spleen involvement were defined as low-risk. The 15 selected pairs (Supplementary Table S4) were fixed after training and applied unchanged to the test set and to the bootstrap resamples. For each sample, we computed (i) a pairwise vote score, the number of pairs in which the immune gene's expression exceeded that of its B-cell partner (range 0–15, ties counted as 0), and (ii) an aggregate score, the sum of the DESeq2-normalized counts of the immune genes minus that of the B-cell genes. Both are given in Supplementary Table S5. For visualization (Figure 5B–E), aggregate scores were transformed as log2(score + |min| + 1), where min is the lowest aggregate score across all 158 samples. Performance was assessed in the test set. To further evaluate signature generalizability, we performed 1,000 bootstrap iterations, assessing performance in distinguishing both grouped patterns (A-B versus C-E) and individual patterns or diagnostic classes. Significance of comparisons was assessed using the Wilcoxon rank-sum test.

External validation

External validation of the granular score was performed in two independent public cohorts without refitting or reweighting. First, we analyzed the Van Loo et al. microarray dataset (GSE7788), which includes 10 NLPHL and 10 THRLBCL cases. Affymetrix probe sets were mapped to gene symbols, and for each gene, one probe with the highest variance across samples was selected before scoring. Second, we analyzed the bulk RNA-seq transcripts-per-million (TPM) gene expression matrices and corresponding clinical tables from Subramanian et al., which provide separate NLPHL training (n=109) and independent validation (n=62) cohorts. For the comparison of primary classic versus variant NLPHL, the published training and validation subsets were pooled. To enable an unambiguous classic versus variant NLPHL comparison, we restricted analyses to cases confined to classic patterns A/B only (A, B, or A/B) or confined to variant patterns C-F only. Cases spanning both A/B and C-F patterns, or with missing or uncertain IAP annotation, were excluded. The granular score was computed for each sample using the same pre-specified approach defined in our training cohort. Group differences were assessed using the Wilcoxon rank-sum test, and discrimination was quantified using the receiver operating characteristic area under the curve (ROC-AUC). In every cohort, the score is the sum of the 15 immune genes minus the sum of the 15 B-cell genes (the pairings are given in Supplementary Table S4), computed on each platform's own expression values: DESeq2-normalized counts in our cohort, GC-RMA log2 intensities in the Van Loo data (one probe per gene, the probe with the highest variance) and TPM in the Subramanian data. Supplementary Table S5 lists the raw difference and Figure 5 shows its log2 after a constant shift. Figures S2 and S3 are on the external platforms' scales. The rank-based pair vote (number of pairs in which the immune gene exceeds the B-cell gene) was computed alongside and is reported in Supplementary Table S5 for our cohort and Supplementary Table S9 for the external cohorts. AL139020.1 has no probe on the GPL570 array and is absent from the Subramanian matrices, so 14 of the 15 pairs are used in the external cohorts.

Cross-validated confirmation of the continuum

To confirm that the classic to variant to THRLBCL ordering was not an artifact of the top-scoring-pair construction, we trained an independent classifier using fully nested leave-one-out cross-validation on the same DESeq2-normalized expression matrix. Within each outer fold, the 2,000 most variable genes were selected and standardized (z-scored). These features were used to fit an L2-penalized

logistic regression with balanced class weights discriminating NLPHL from THRLBCL. The regularization strength was tuned by five-fold inner cross-validation. Feature selection and standardization were performed inside each fold so that no sample contributed to its own prediction. The resulting out-of-fold THRLBCL probability was examined across NLPHL IAP categories and THRLBCL (Online Supplementary Figure S1).

Pathway analysis

To delineate the biological processes underlying the differential expression between groups, we performed an over-representation analysis of the model genes whose expression levels were associated with the transcriptional continuum between NLPHL and THRLBCL. The analysis was conducted using the clusterProfiler library in R. The overall pathway universe consisted of xCell and CIBERSORT gene sets, as well as Hallmark and canonical pathway (C2) gene sets from the Human Molecular Signatures Database (MSigDB).

Statistical analyses

All statistical analyses were performed with R (v4.5.0). Heatmaps were generated using the R package pheatmap and volcano plots using EnhancedVolcano. The Fisher exact test was used to assess differences in the frequency of dichotomous clinical variables. The Wilcoxon rank-sum test was used for non-parametric comparisons between two groups. Discrimination performance was quantified using ROC-AUC where indicated. For survival analysis, Kaplan-Meier estimates with the log-rank test, as well as univariable and multivariable Cox regression, were used, implemented with the R packages survival and survminer. Overall survival (OS) was defined as the time between the date of diagnosis and death from any cause. Event-free survival (EFS) was defined as the time from diagnosis to disease progression, transformation, or death from any cause.

Principal component analysis and PERMANOVA

Principal component analysis was run on the DESeq2-normalized, log2-transformed expression matrix (16,852 genes, 200 samples, mean-centered and not scaled). The variance explained by diagnosis was estimated with PERMANOVA on Euclidean distances (9,999 permutations). The 95% concentration ellipse of the NLPHL cases was drawn in the PC1–PC2 plane assuming a bivariate normal distribution. To place THRLBCL relative to the NLPHL patterns, the distance of each THRLBCL case to the average expression profile of classic (n = 86) and of variant (n = 54) NLPHL was computed over all genes (Euclidean distance on the same matrix, leaving out the 38 NLPHL cases without a pattern assignment), and the median difference between the two distances was compared with its distribution under 9,999 random reassignments of the classic and variant labels (one-sided p).

Follicular helper T-cell analyses

All 22 deconvolved populations were tested by median split for association with EFS (log-rank test) in the 123 NLPHL patients with survival data. Two populations that were zero in all samples could not be tested. P values were corrected across the 20 testable populations with the Benjamini–Hochberg procedure. Cox models for Tfh content are given unadjusted, adjusted for LP-IPS category or splenic involvement, and, in the 106 patients with a granular LP-IPS score, with splenic involvement entered separately from the rest of the score, and, in the 121 patients with known stage, with the score components (age ≥45 years, stage III–IV, splenic involvement) as separate covariates. Firth-penalized Cox regression was used where events were few (Supplementary Table S10).

External outcome test in the Subramanian cohort

Subramanian et al. (reference 21 of the main manuscript) published TPM expression matrices for their training (n = 109) and validation (n = 62) cohorts together with immunoarchitectural pattern, LP-IPS, stage, treatment and freedom from progression. Pattern was coded as classic (A or B only, n = 71), variant (C–F only, n = 31) or mixed. Freedom from progression was analysed in the 128 patients with follow-up, excluding 8 treated with rituximab alone as in the source study (30 progressions). The 30-gene signature was applied without refitting, and 29 of the 30 genes were present (AL139020.1 is absent), so 14 of the 15 gene pairs were usable. The score was computed as in the other cohorts, the sum of the immune genes minus the sum of the B-cell genes on TPM, and is shown and modelled as its log2 after a constant shift. Associations with pattern, LP-IPS and stage were quantified by ROC-AUC and Wilcoxon tests, and freedom from progression by Cox regression per standard deviation of the score, stratified by cohort and adjusted for LP-IPS. Tfh fractions were estimated with FARDEEP and the LM22 matrix and tested in the same way. As a reference, the LP-IPS and stage were tested in the same 128 patients and in our own survival set (Supplementary Table S11).

Software packages are cited in the reference list of the main manuscript, unless otherwise indicated.

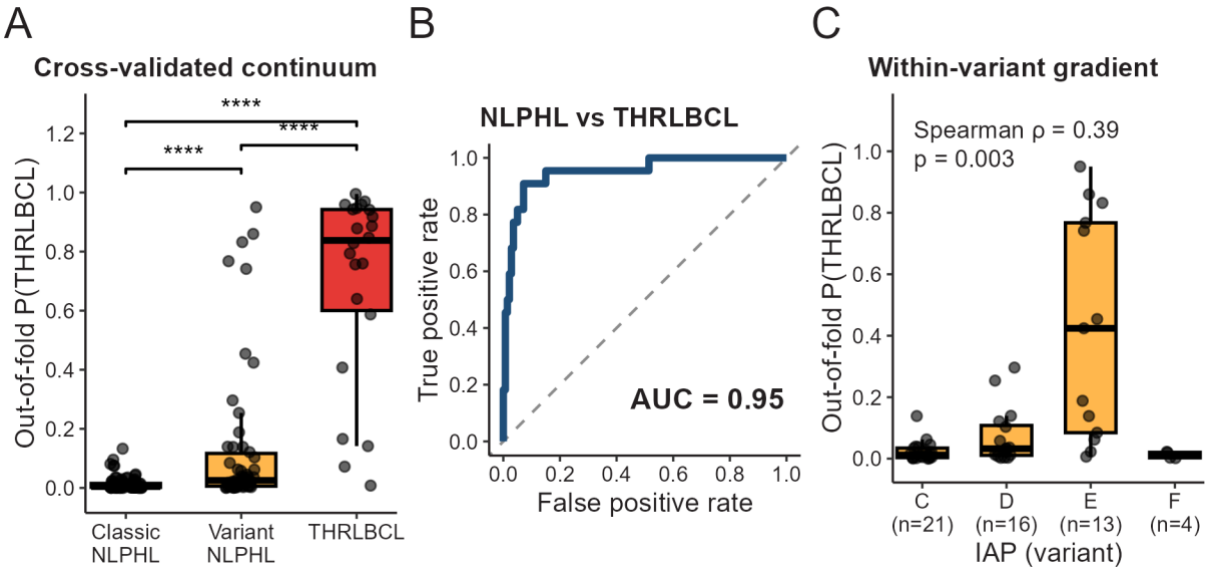


Figure S1 (A-C). Cross-validated confirmation of the NLPHL-to-THRLBCL continuum. Out-of-fold probabilities from an independent classifier of NLPHL versus THRLBCL. (A) Box plots of the THRLBCL probability across classic NLPHL, variant NLPHL, and THRLBCL. (B) Receiver operating characteristic (ROC) curve for the discrimination of THRLBCL versus NLPHL. (C) Box plots of the THRLBCL probability within variant NLPHL across IAP C–F. All pairwise comparisons are significant at $p<0.05$ unless otherwise indicated.

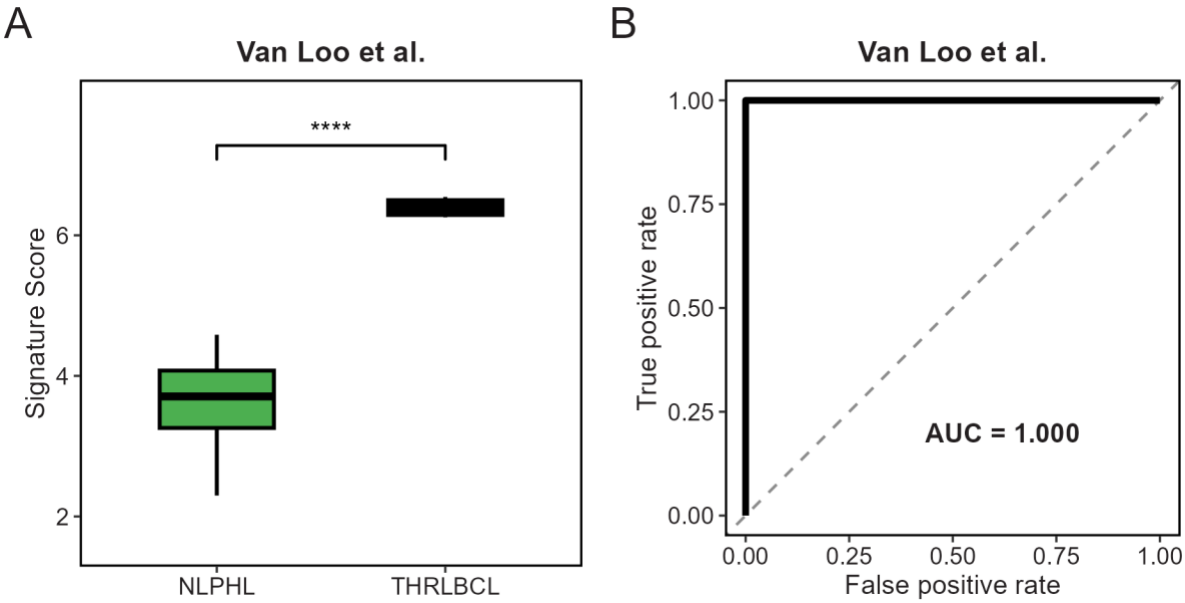


Figure S2 (A-B). External validation of the signature score in Van Loo et al. (A) Box plots showing the distribution of the signature score across NLPHL and THRLBCL samples. (B) Receiver operating characteristic (ROC) curve evaluating discrimination of THRLBCL versus NLPHL using the signature score. All pairwise comparisons are significant at $p<0.05$ unless otherwise indicated.

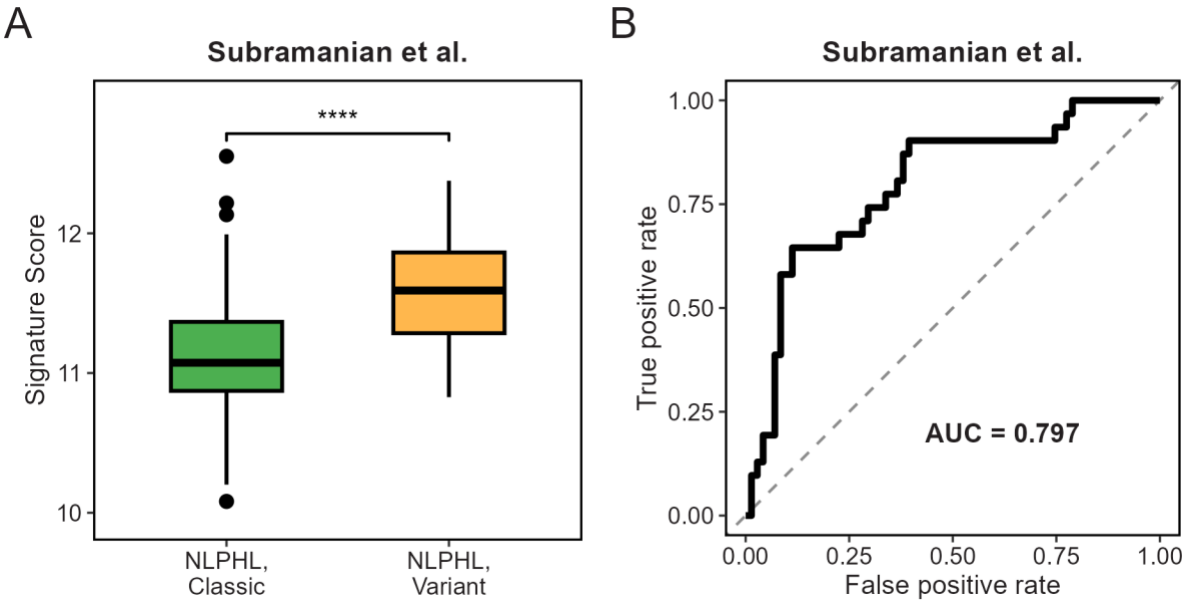


Figure S3 (A-B). External validation of the signature score in Subramanian et al. (A) Box plots showing the distribution of the signature score in classic versus variant NLPHL samples. (B) Receiver operating characteristic (ROC) curve evaluating discrimination of variant versus classic NLPHL using the signature score. All pairwise comparisons are significant at $p<0.05$ unless otherwise indicated.

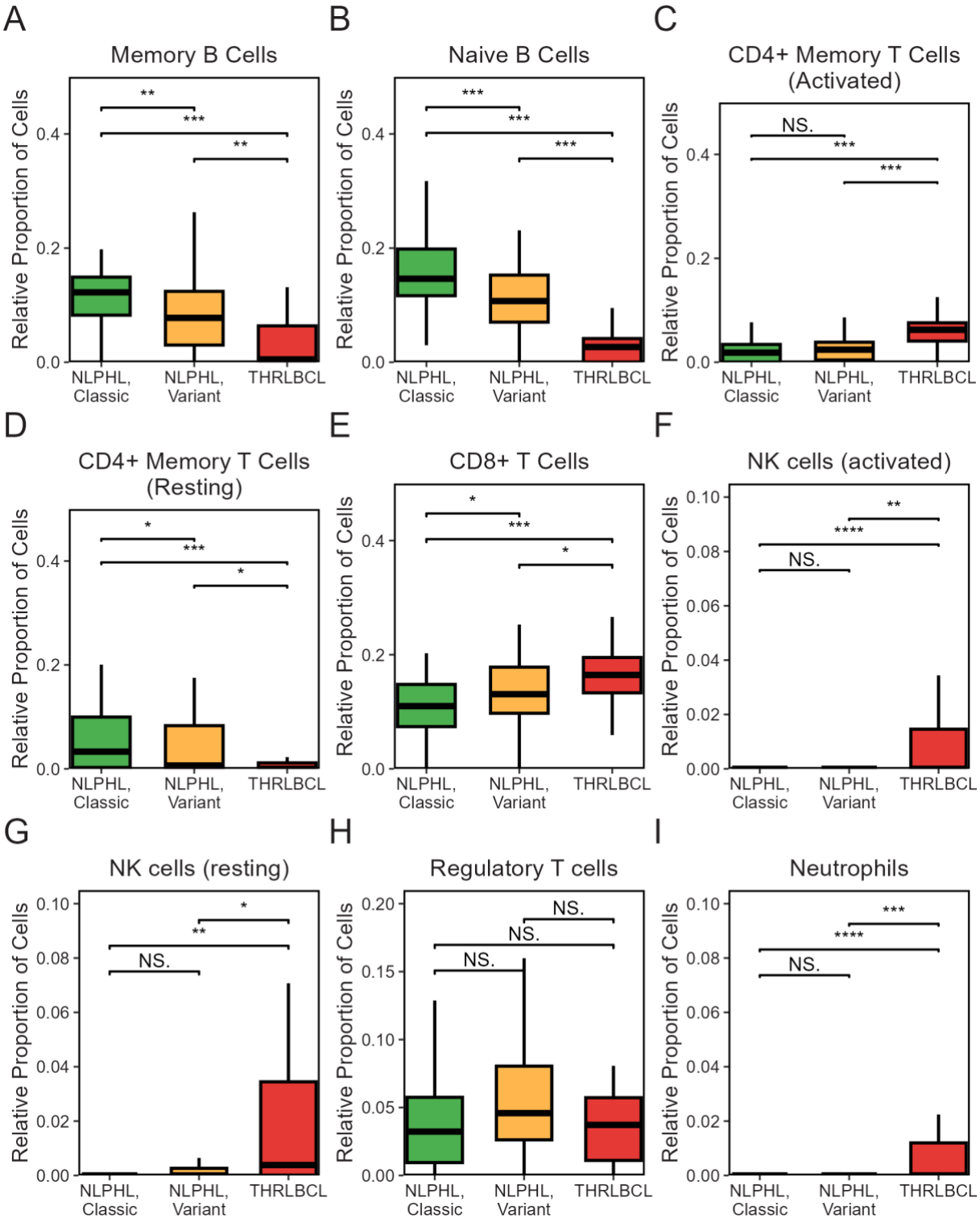


Figure S4 (A-I). Tumor microenvironment of NLPHL and THRLBCL. Box plots showing the relative proportion of the indicated cell types in each IAP category and THRLBCL. All pairwise comparisons are significant at $p<0.05$ unless otherwise indicated. NS, not significant.

Supplementary Tables

Supplementary Tables S1–S11 are provided as a separate Excel file, one table per sheet.

Supplementary Table S1. Clinical information for the 200 sequenced patients: diagnosis, age, sex, stage, splenic involvement, LP-IPS, IAP, first-line therapy, outcome, sequenced biopsy and biopsy site.

Supplementary Table S2. (A) Log2 fold changes for the sequential IAP comparisons and NLPHL versus THRLBCL (training set), with direction. (B) t-statistics of the 185 genes shown in Figure 4A for the six comparisons, with direction.

Supplementary Table S3. Deconvolved immune cell fractions (FARDEEP, LM22), 22 populations per sample, for the 200 samples.

Supplementary Table S4. The 15 top-scoring gene pairs of the signature with the log2-normalized expression of the 30 genes in all samples.

Supplementary Table S5. Signature score and pair vote per sample, with IAP group and training or test set (n = 158).

Supplementary Table S6. Differential expression between NLPHL and THRLBCL (DESeq2): log2 fold change relative to NLPHL, adjusted P value and direction for genes with $|\log2 \text{ fold change}| \geq 1.5$ and adjusted $P < 0.05$.

Supplementary Table S7. Genes differentially expressed between Tfh-high and Tfh-low NLPHL (DESeq2, adjusted $P < 0.05$; log2 fold change relative to Tfh-low).

Supplementary Table S8. Clinical characteristics of Tfh-high versus Tfh-low patients within the high LP-IPS stratum (n = 20 and 11). P values, Fisher's exact test.

Supplementary Table S9. External validation of the signature score in Van Loo et al. and Subramanian et al.: Wilcoxon P and ROC-AUC with 95% CI for the summed score and for the pair vote.

Supplementary Table S10. Tfh content and event-free survival in NLPHL (n = 123): log-rank tests of all deconvolved populations with Benjamini–Hochberg q values, Cox models for Tfh content, and the Tfh by LP-IPS interaction.

Supplementary Table S11. Signature score and Tfh content in the Subramanian et al. cohort: cohort description, score versus pattern and clinical features, Tfh content versus freedom from progression, and reference outcome models.