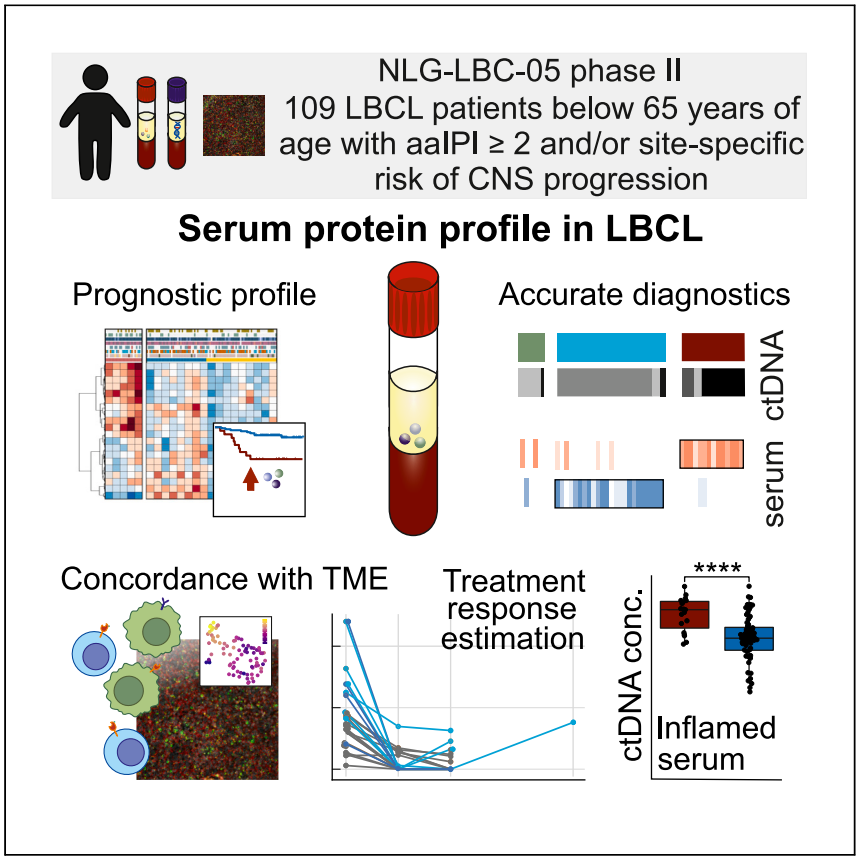
Med

CellPress
OPEN ACCESS

Article

Inflammatory and subtype-dependent serum protein signatures predict survival beyond the ctDNA in aggressive B cell lymphomas



Arffman et al. discovered novel associations between inflamed serum protein profile, exhausted tumor microenvironment, and poor survival in patients with primary high-risk large B cell lymphoma. They also reported findings of multiple serum proteins that complement several circulating tumor DNA assessments, such as subtype stratification and response evaluation.

Maare Arffman, Leo Meriranta, Matias Autio, ..., Mette Ølgod Pedersen, Suvi-Katri Leivonen, Sirpa Leppä

sirpa.leppa@helsinki.fi

Highlights

Landscape of 1,463 serum proteins in high-risk large B cell lymphoma (LBCL)

Inflamed serum protein profile predicts poor survival in primary LBCL

Several serum proteins complement circulating tumor DNA assessment

Lymphoma subtype marker proteins can be used to estimate response to therapy

Translation to Patients

Arffman et al., Med 5, 1–20
June 14, 2024
https://doi.org/10.1016/j.medj.2024.03.007

Please cite this article in press as: Arffman et al., Inflammatory and subtype-dependent serum protein signatures predict survival beyond the ctDNA in aggressive B cell lymphomas, Med (2024), https://doi.org/10.1016/j.medj.2024.03.007

Med

CellPress
OPEN ACCESS

Article

Inflammatory and subtype-dependent serum protein signatures predict survival beyond the ctDNA in aggressive B cell lymphomas

Maare Arffman,[1,2,3,15] Leo Meriranta,[1,2,3,15] Matias Autio,[1,2,3] Harald Holte,[4] Judit Jørgensen,[5] Peter Brown,[6] Sirkku Jyrkkiö,[7] Mats Jerkeman,[8] Kristina Drott,[8] Øystein Fluge,[9] Magnus Björkholm,[10] Marja-Liisa Karjalainen-Lindsberg,[11] Klaus Beiske,[12] Mette Ølgod Pedersen,[13,14] Suvi-Katri Leivonen,[1,2,3] and Sirpa Leppä[1,2,3,16,*]

SUMMARY

Background: Biological heterogeneity of large B cell lymphomas (LBCLs) is poorly captured by current prognostic tools, hampering optimal treatment decisions.

Methods: We dissected the levels of 1,463 serum proteins in a uniformly treated trial cohort of 109 patients with high-risk primary LBCL (ClinicalTrials.gov: NCT01325194) and correlated the profiles with molecular data from tumor tissue and circulating tumor DNA (ctDNA) together with clinical data.

Findings: We discovered clinically and biologically relevant associations beyond established clinical estimates and ctDNA. We identified an inflamed serum protein profile, which reflected host response to lymphoma, associated with inflamed and exhausted tumor microenvironment features and high ctDNA burden, and translated to poor outcome. We composed an inflammation score based on the identified inflammatory proteins and used the score to predict survival in an independent LBCL trial cohort (ClinicalTrials.gov: NCT03293173). Furthermore, joint analyses with ctDNA uncovered multiple serum proteins that correlate with tumor burden. We found that SERPINA9, TACI, and TARC complement minimally invasive subtype profiling and that TACI and TARC can be used to evaluate treatment response in a subtype-dependent manner in the liquid biopsy.

Conclusions: Altogether, we discovered distinct serum protein landscapes that dissect the heterogeneity of LBCLs and provide agile, minimally invasive tools for precision oncology.

Funding: This research was funded by grants from the Research Council of Finland, Finnish Cancer Organizations, Sigrid Juselius Foundation, University of Helsinki, iCAN Digital Precision Cancer Medicine Flagship, Orion Research Foundation sr, and Helsinki University Hospital.

CONTEXT AND SIGNIFICANCE

Clinical and biological heterogeneity remains a challenge for improving large B cell lymphoma (LBCL) survival. Here, by exploring the landscape of 1,463 serum proteins, Arffman et al. found an inflammatory profile that was associated with aggressive clinical characteristics, exhausted tumor microenvironment, and high circulating tumor DNA (ctDNA) levels. Furthermore, they show that multiple serum proteins have complementary potential together with ctDNA in subtype classification and response evaluation. The results suggest that serum protein profiling can enhance the detection of high-risk LBCL non-invasively and that it could be harnessed for accurate diagnostics and treatment decisions.

INTRODUCTION

Large B cell lymphomas (LBCL) comprise a heterogeneous group of aggressive lymphoid cancers.[1,2] The most common entity covering over 90% of all LBCLs is diffuse LBCL not otherwise specified (DLBCL NOS), which is further divided into germinal center B cell (GCB)- and activated B cell (ABC)-like subtypes. Other less common LBCL entities include, for example, high-grade B cell lymphoma

Please cite this article in press as: Arffman et al., Inflammatory and subtype-dependent serum protein signatures predict survival beyond the ctDNA in aggressive B cell lymphomas, Med (2024), https://doi.org/10.1016/j.medj.2024.03.007

with *MYC* and *BCL2* rearrangements, primary mediastinal LBCL (PMBCL), and T cell/histiocyte-rich LBCL (THRLBCL).[1,2] Although all LBCLs are treated with a curative intent using anthracycline-based immunochemotherapy, around 30% of patients have primary refractory disease or experience progression later during follow-up.[3] Inadequate dismantling of clinical and biological heterogeneity remains the main obstacle for improving outcomes, and there is a need for novel tools, which could be used to stratify patients with LBCL more accurately. Therefore, gene expression profiling,[4] mutational landscapes,[5,6] and, more recently, tumor microenvironment (TME) characteristics[7,8] have been utilized to subclassify LBCL. So far, however, targeted therapies based on molecular characterization have not been superior to standard of care. It may well be that characterization of host response on the systemic level may provide information beyond tumor tissue and thereby help to recognize patients with disturbed immunological response to lymphoma, high risk of death, and unmet therapeutic needs.

Accurate diagnosis of distinct LBCL entities relies on appropriate clinical settings, morphological and immunophenotypic assessment of tumor tissue, and, increasingly, also molecular and genomic profiling. The exact diagnosis may, however, be compromised due to inadequate or unrepresentative tissue biopsy caused by an anatomically challenging tumor location, urgency to initiate therapy, or spatial heterogeneity. Accordingly, there is a need for novel minimally invasive tools that can reduce the requirements for tumor tissue. Among them, circulating tumor DNA (ctDNA)-based liquid biopsies are anticipated to change the management of patients with aggressive LBCL.[9–11] However, genomic analysis and quantification of ctDNA are limited to tumor cell properties, overlooking the impact of TME and host response, which have emerging clinical impact in DLBCL.[7,8,12]

Analysis of serum proteins in routine blood samples guides treatment decisions and complements diagnostics of all patients with B cell lymphomas. Besides disclosing vital information of infectious diseases and organ function, serum proteins can also serve as another layer of liquid biopsy to characterize and monitor not only tumor cells but also host responses. For example, recent studies applying antibody-based immuno-oncology panels composed of 97 and 81 plasma and/or serum proteins were able to identify circulating inflammatory protein profiles, which were associated with outcome.[13,14] Nevertheless, systemic inflammation and other serum markers remain unlinked to molecular and clinical characteristics of LBCLs. Moreover, since exploration of single or a few soluble proteins have resulted in contradictory findings,[15–17] further research is warranted.

Here, we comprehensively characterized the serum proteome and correlated the findings with molecular and clinical characteristics of primary LBCL. We studied serum levels of 1,463 proteins in a discovery cohort of 109 high-risk patients treated in a Nordic lymphoma trial using a proximity extension assay (PEA). We identified an inflamed serum protein profile, which reflects host response to lymphoma, associates with exhausted TME features and high ctDNA burden, and translates to poor outcomes. We composed an inflammation score based on the inflammatory proteins and used it to predict outcome in an independent cohort of 122 patients with LBCL. Besides inflammation, we uncovered multiple serum proteins, which expand the liquid biopsy repertoire for accurate LBCL diagnosis beyond ctDNA. Taken together, this is the first large-scale study to characterize LBCL serum protein profiles and to integrate the findings with tumor and ctDNA characteristics in uniformly treated patients.

[1]Research Programs 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]Department of Oncology, Oslo University Hospital and KG Jebsen Centre for B Cell Malignancies, Oslo, Norway

[5]Department of Hematology, Aarhus University Hospital, Aarhus, Denmark

[6]Department of Hematology, Copenhagen University Hospital Rigshospitalet, Copenhagen, Denmark

[7]Department of Oncology, Turku University Hospital, Turku, Finland

[8]Department of Oncology, Skane University Hospital, Lund, Sweden

[9]Department of Oncology, Haukeland University Hospital, Bergen, Norway

[10]Department of Medicine, Karolinska University Hospital, Stockholm, Sweden

[11]Department of Pathology, Helsinki University Hospital, Helsinki, Finland

[12]Department of Pathology, Oslo University Hospital, Oslo, Norway

[13]Department of Pathology, Zealand University Hospital, Roskilde, Denmark

[14]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[15]These authors contributed equally

[16]Lead contact

*Correspondence: sirpa.leppa@helsinki.fi

https://doi.org/10.1016/j.medj.2024.03.007

Please cite this article in press as: Arffman et al., Inflammatory and subtype-dependent serum protein signatures predict survival beyond the ctDNA in aggressive B cell lymphomas, Med (2024), https://doi.org/10.1016/j.medj.2024.03.007

RESULTS

Clinical characteristics

Patient demographics of the discovery cohort (n = 109; NLG-LBC-05 trial[18]) are described in Table 1. All patients were treatment naive. Median age was 56 years (range 21–64 years). The majority of patients were males (n = 66, 60%) and had DLBCL NOS (n = 83, 76%), stage IV disease (n = 85, 78%), B symptoms (n = 71, 65%), and elevated lactate dehydrogenase (LDH) levels (n = 101, 93%). During the median follow-up time of 62 months (interquartile range 54–66 months), 21 (19%) patients relapsed and 17 (16%) died, translating to 81% (95% confidence interval [CI] 74%–88%) progression-free survival (PFS) and 84% (95% CI 78%–92%) overall survival (OS) at 5 years. The patients were treated with dose-dense immunochemotherapy combined with early systemic CNS prophylaxis, which overcame common adverse prognostic impacts of age-adjusted International Prognostic Index (aaIPI) score 3 and high-grade B cell lymphoma with BCL2 and/or BCL6 and MYC translocations.[18]

An inflamed serum protein profile is associated with poor outcome in high-risk LBCL

To characterize the circulating proteomic landscape in aggressive LBCL, we applied an antibody-based PEA to pretreatment serum samples collected from the patients and healthy donors. Distribution of protein profiles in relation to clinical variables is shown in Figure S1A. First, we explored relative serum concentrations of all studied proteins with consensus clustering and principal-component analysis (PCA), both of which suggested that the patients could be split into three groups according to their distinctive serum protein profiles (Figures 1A and S1B). As expected, allocation of the patients into three groups with k-means clustering reflected the PCA (clusters I–III). Serum profiles of healthy donors associated with clusters II and III, implying that the pretreatment serum protein profile of a subset of patients with LBCL was closely similar to the profile of healthy subjects.

To determine distinct features of the identified serum profiles, we examined the proteins and pathways characteristic for each of the clusters. Differences in the protein levels between the clusters were further visualized in a heatmap (Figure 1B). Accordingly, the serum protein profile in cluster I was distinguished from other clusters by elevated levels of proinflammatory cytokines, such as interferon gamma (IFNG) and interleukins 10 and 18 (IL-10 and IL-18); cytolytic effectors, including granzymes H and B (GZMH and GZMB); and checkpoint molecules, such as T cell immunoglobulin domain and mucin domain 3 (TIM3/HAVCR2), lymphocyte-activation gene 3 (LAG3), and programmed death ligand 1 (PD-L1; Figures 1C; Table S1). Accordingly, serum proteins with elevated levels in the cluster I patients were associated with inflammatory pathways and lymphocyte activity, reflecting clinically to B symptoms, and thus we named the cluster as the inflamed serum profile (Figure 1D). By contrast, serum proteins in clusters II and III, which encompassed the profiles of both patients and healthy donors, correlated with non-inflammatory pathways (Tables S2 and S3).

We found that the inflamed serum protein profile was enriched in the patients with THRLBCL ($p < 0.001$), non-GCB DLBCL ($p < 0.01$), bone marrow involvement ($p < 0.001$), or B symptoms ($p = 0.051$; Figures 1B; Table 1). Patients with the inflamed serum profile also had significantly worse OS and PFS compared to the other patients (Figures 1E and S1C). In multivariable analysis with aaIPI, bone marrow infiltration, and B symptoms, the inflamed serum profile retained a prognostic impact on OS (Figures 1F, S1D, and S1E). By using penalized logistic regression,

Please cite this article in press as: Arffman et al., Inflammatory and subtype-dependent serum protein signatures predict survival beyond the ctDNA in aggressive B cell lymphomas, Med (2024), https://doi.org/10.1016/j.medj.2024.03.007

Table 1. Patient characteristics

Characteristic	All patients (n = 109), n (%)	Inflamed profile (n = 24) , n (%)	Non-inflamed profile (n = 85), n (%)	Fisher's exact test, p
Age, median 45 years, range 21–64				p = 0.807
<60 years	73 (67)	17 (70)	56 (66)	–
≥60 years	36 (33)	7 (30)	29 (34)	–
Sex				p = 0.345
Male	66 (60)	17 (70)	49 (70)	–
Female	43 (40)	7 (30)	36 (30)	–
Histology				p < 0.001 (histology)
DLBCL NOS (Hans algorithm)	83 (76)	16 (67)	67 (79)	p < 0.008 (DLBCL COO)
GCB	44 (53)	4 (25)	40 (60)	–
Non-GCB	31 (37)	11 (69)	20 (30)	–
Not reviewed	8 (10)	1 (6)	7 (10)	–
PMBCL	8 (7)	0 (0)	8 (9)	–
THRLBCL	4 (4)	4 (17)	0 (0)	–
FL (grade 3B)	2 (2)	0 (0)	2 (3)	–
Intravascular	1 (1)	1 (1)	0 (0)	–
Not available	11 (10)	3 (13)	8 (9)	–
High-grade B cell lymphoma/examined	8/59	1/59	7/59	p = 1.000
Performance score (ECOG)				p = 0.246
0–1	72 (66)	12 (50)	60 (71)	–
2–3	37 (34)	12 (50)	25 (29)	–
Stage (Ann Arbor)				p = 0.574
I-II	4 (4)	0 (0)	4 (5)	–
III-IV	105 (96)	24 (100)	81 (95)	–
aaIPI				p = 0.153
0–2	69 (63)	12 (50)	57 (67)	–
3	40 (37)	12 (50)	28 (33)	–
CNS-IPI				p = 1.000
Low (1–3)	58 (53)	13 (54)	45 (53)	–
High (4–6)	51 (47)	11 (46)	40 (47)	–
B symptoms				p = 0.051
Yes	71 (65)	20 (83)	51(60)	–
No	38 (35)	4 (17)	34 (40)	–
Bulky disease				p = 0.064
Yes	46 (42)	6 (25)	40 (47)	–
No	63 (58)	18 (75)	45 (53)	–
LDH				p = 0.196
Elevated	101 (93)	24 (100)	77 (91)	–
Normal	8 (7)	0 (0)	8 (9)	–
Bone marrow infiltration				p < 0.001
No	79 (72)	10 (42)	69 (81)	–
Yes	30 (28)	14 (58)	16 (19)	–

DLBCL NOS, diffuse large B cell lymphoma not otherwise specified; GCB, germinal center B cell; ABC, activated B cell; PMBCL, primary mediastinal large B cell lymphoma; FL, follicular lymphoma; THRLBCL, T cell/histiocyte rich large B cell lymphoma; ECOG, Eastern Cooperative Oncology Group; aaIPI, age-adjusted International Prognostic Index; CNS-IPI, central nervous system international prognostic index; LDH, lactate dehydrogenase.

we established simplified models to detect the prognostic inflamed cluster (hazard ratio [HR] = 1.66, p < 0.01, accuracy rate 0.95; Figures S1F–SG) and patients with poor OS (HR = 5.73, p < 0.001; Figure S1H).

To aid the approximation of inflammation in different datasets, we composed an inflammation score of the most pertinent inflammatory proteins distinguishing

Please cite this article in press as: Arffman et al., Inflammatory and subtype-dependent serum protein signatures predict survival beyond the ctDNA in aggressive B cell lymphomas, Med (2024), https://doi.org/10.1016/j.medj.2024.03.007

CellPress
OPEN ACCESS

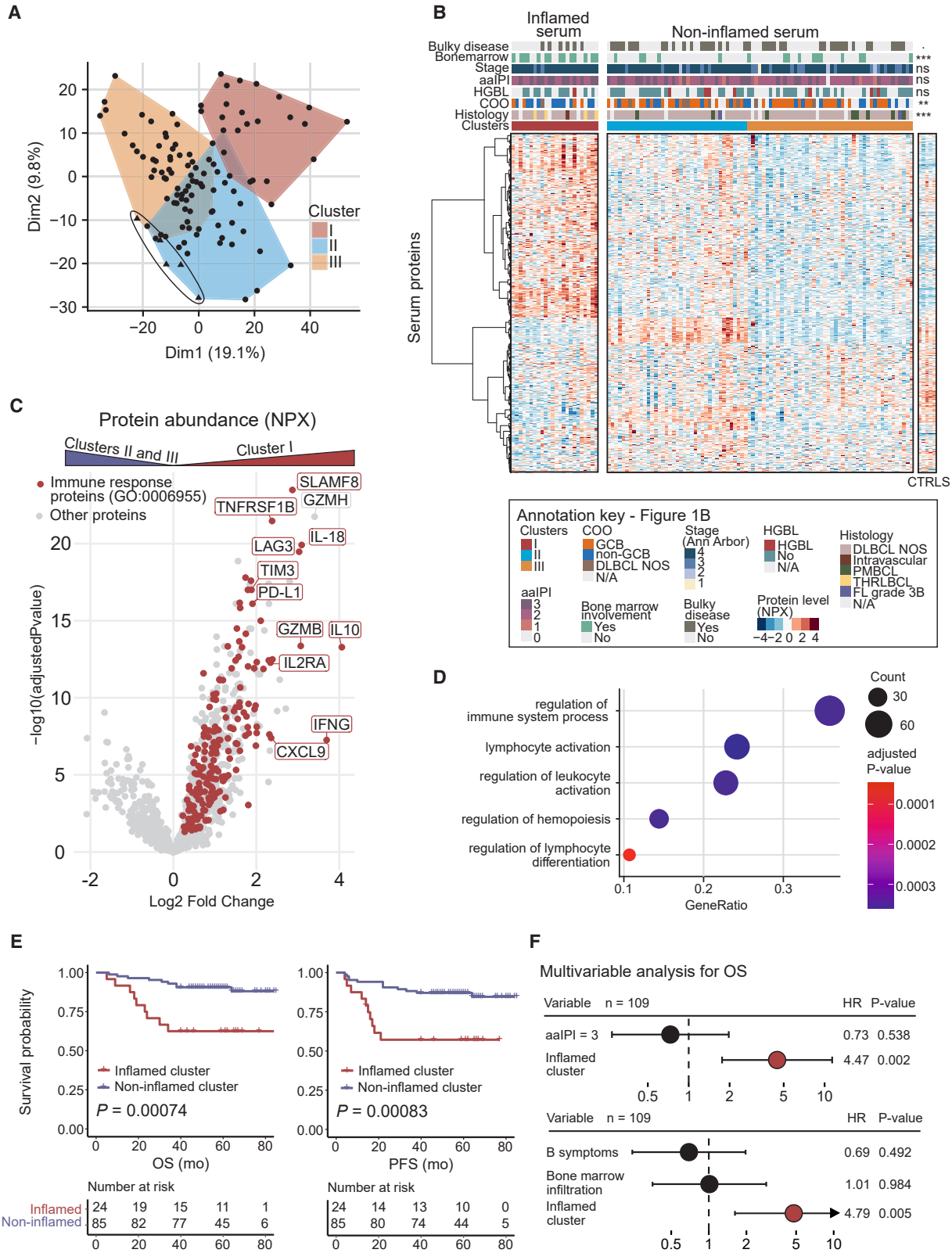


Figure 1. Serum protein profile and clinical characteristics of the patients with high-risk large B cell lymphoma

(A) Principal-component analysis (PCA) of serum samples, which are colored according to group assignment with k-means clustering (k = 3). Points represent samples from the patients, and encircled triangles represent serum samples from the healthy donors. See also Figure S1A.

(B) Heatmap of the serum protein levels (rows) in patient samples (columns) grouped according to their k-means clusters. Clinical variables are annotated on top of the heatmap, and cluster I is compared to clusters II and III with Fisher's exact test. High levels of proteins are colored red, whereas low levels are colored blue. Protein levels of serum samples from healthy donors (CTRLs) are shown on the right (n = 5).

(C) Volcano plot of differentially expressed proteins in cluster I compared to clusters II and III. Red dots represent immune response proteins (GO: 0006955), and gray dots represent proteins from other pathways. The most pertinent inflammatory proteins in GO: 0006955 and GZMH are annotated.

(D) Enrichment analysis of the pathways represented by serum sample proteins (log2 fold change > 0 and adjusted p value < 0.05) from the patients in the inflamed cluster compared to protein levels from the patients in the non-inflamed cluster (n = 221).

(E) Kaplan-Meier survival estimates for OS and PFS according to inflamed cluster (n = 24, red) and non-inflamed cluster (n = 85, purple). See also Figure S1C.

(F) Multivariable Cox regression models for OS of inflamed serum profile together with aaIPI = 3 and bone marrow infiltration and B symptoms. See also Figure S1D.

OS, overall survival; PFS, progression-free survival; aaIPI, age-adjusted IPI; dim, dimension.

the clusters (Figure S2A). Of these proteins, we aimed to reduce the score to mimic clinical purposes and hypothesized that secretory molecules, such as cytokines, would represent inflammation in situations where all inflammation score proteins would not be available. After ensuring positive correlation between cytokines measured with microbead-based immunoassay (Luminex) and PEA from our discovery cohort (Figure S2B), we saw a strong correlation (rho = 0.76) between the inflammation scores derived from the Luminex (Luminex inflammation score: IL-10, IL-18, and CXCL9) and Olink (Olink inflammation score: 12 proteins) data (n = 47; Figure 2A). As expected, the patients in the inflamed cluster had higher inflammation scores with both measurement assays compared to the patients in the non-inflamed cluster (Figure 2A). Furthermore, when we reduced the inflammation score proteins from the Olink inflammation score (n = 12) to those available in the Luminex (n = 3) in the Olink data, we found that a high inflammation score was still associated with worse OS and PFS (PFS: p = 0.011; Figures 2B and S2C). Finally, we sought to validate the results using serum samples from an independent cohort of patients with LBCL (n = 122, NLG-LBC-06 trial[19]). Patient demographics of the validation cohort are described in Table S4. Similarly to the discovery cohort, we detected variation between the patients in their serum cytokine levels (Figure S2D). In survival analysis, high pretreatment serum levels of inflammatory proteins (optimal cutoff; inflammation score >7.51, high n = 43, low n = 79) translated to worse survival (Figures 2C and S2E). Furthermore, patients with a high Luminex inflammation score more often had B symptoms, high LDH levels, and non-GCB DLBCL as compared to the other patients (Figures 2D; Table S4), validating our previous results. Lastly, the Luminex inflammation score associated with PFS independent of aaIPI (Figure 2E).

Altogether, by studying two independent LBCL cohorts, we uncovered the inflammatory response in the serum proteome, which translates to survival and associates with molecular and clinical characteristics of LBCL.

Inflamed serum protein profile correlates with tumor-infiltrating lymphocyte characteristics in the lymphoma microenvironment

We hypothesized that characterization of the tumor tissue could elucidate the molecular background of the circulating inflammatory proteins. Accordingly, we assessed gene expression levels of the Olink inflammation score proteins in our discovery population, in another Nordic cohort (NLG-LBC-04 trial[20]), and in a publicly available dataset[21] (Figures 3A and S3A). First, we examined gene expression in the lymphoma samples of the Nordic trial cohorts, including samples partially

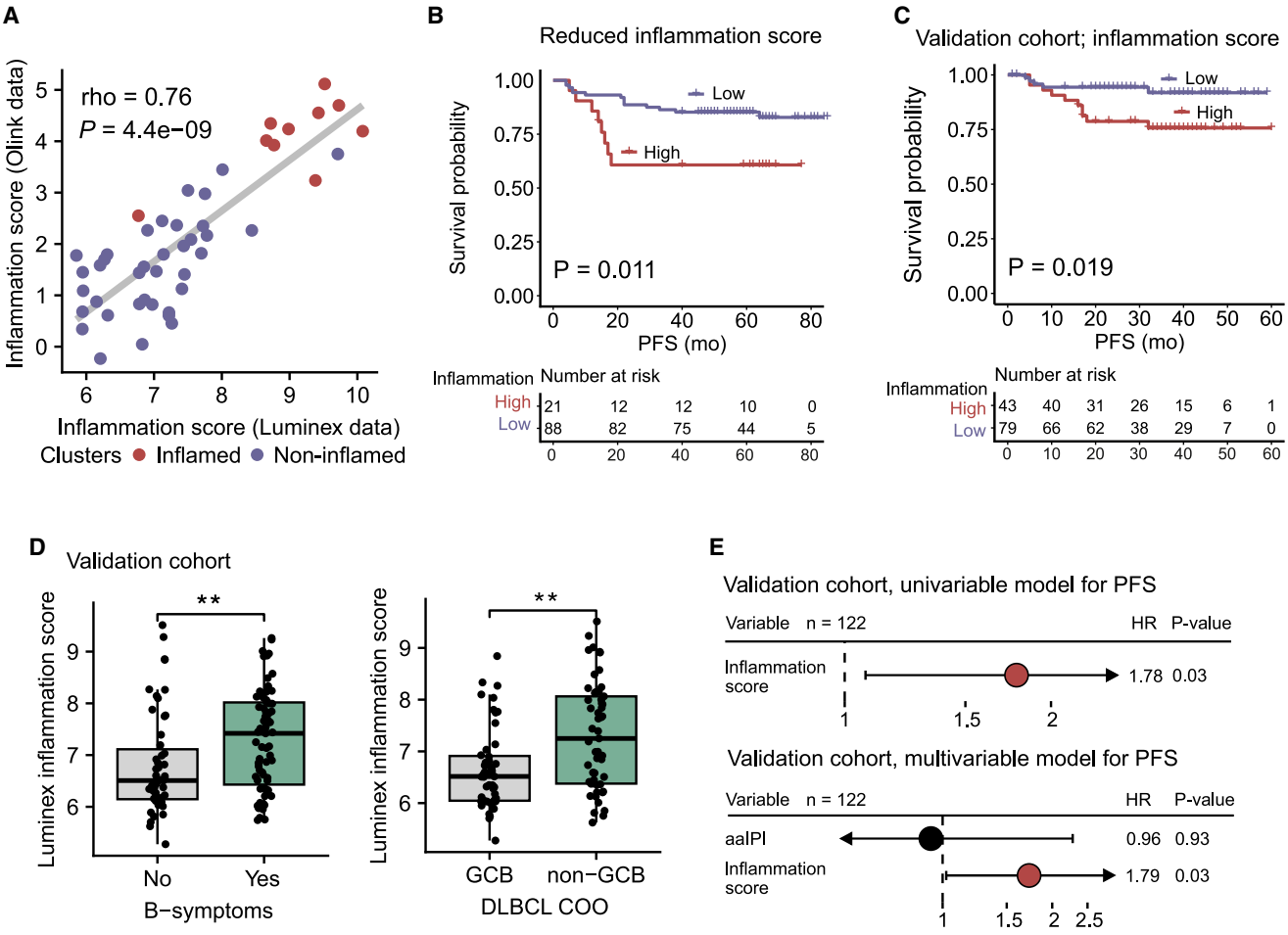


Figure 2. Inflammation score is validated in another LBCL cohort (NLG-LBC-06)

(A) Correlation scatterplot of the inflammation scores derived from the Olink data (IL-2RA, GZMB, IL-18, CXCL9, IFNG, PD-L1, IL-10, TIM3, LAG3, GZMH, TNFRSF1B, SLAMF8, y axis) and Luminex measurements (IL-10, IL-18, and CXCL9, x axis) in the discovery cohort (n = 47). The points represent patients and are colored according to the clusters (red: inflamed cluster, purple: non-inflamed cluster). See also Figure S2B.

(B) Kaplan-Meier survival estimate of PFS according to Olink inflammation score including only proteins in Luminex panel (IL-10, IL-18, and CXCL9). Optimal cutoff for high inflammation score (>4.81, n = 21) is used. See also Figure S2C.

(C) Kaplan-Meier survival estimates of PFS for the validation cohort according to Luminex inflammation score. Optimal cutoff for high inflammation score (>7.51, n = 43) is used. See also (E).

(D) Boxplots of Luminex inflammation score (y axis) in validation cohort according to B-symptom status (first image, x axis) and DLBCL COO (second image, x axis). The points represent patients. **p < 0.01.

(E) Univariable model and multivariable model for PFS of running inflammation score alone and together with aaIPI.

PFS, progression-free survival; aaIPI, age-adjusted IPI.

overlapping with the discovery serum cohort. We observed that a subgroup of LBCLs was characterized by high expression of genes encoding for the inflammation score proteins, which we had previously detected in the serum (*LAG3*, *GZMH*, *GZMB*, *IFNG*, *CD274*, *HAVCR2*, *IL10*, and *IL18*; Figures 3A and S3A). Likewise, we found a similar subgroup of patients with inflamed tumor tissue based on the gene expression of 624 DLBCLs (Figure S3B). In concordance with the lymphoma characteristics of the inflammatory cluster patients in our discovery cohort, these lymphomas were enriched for the ABC subtype, a subset characterized by low tumor cell fraction and high expression of T cell-exhaustion-related markers and IFNG-responsive genes, as well as cytolytic activity with dismal outcomes (data not shown).[7,22]

Please cite this article in press as: Arffman et al., Inflammatory and subtype-dependent serum protein signatures predict survival beyond the ctDNA in aggressive B cell lymphomas, Med (2024), https://doi.org/10.1016/j.medj.2024.03.007

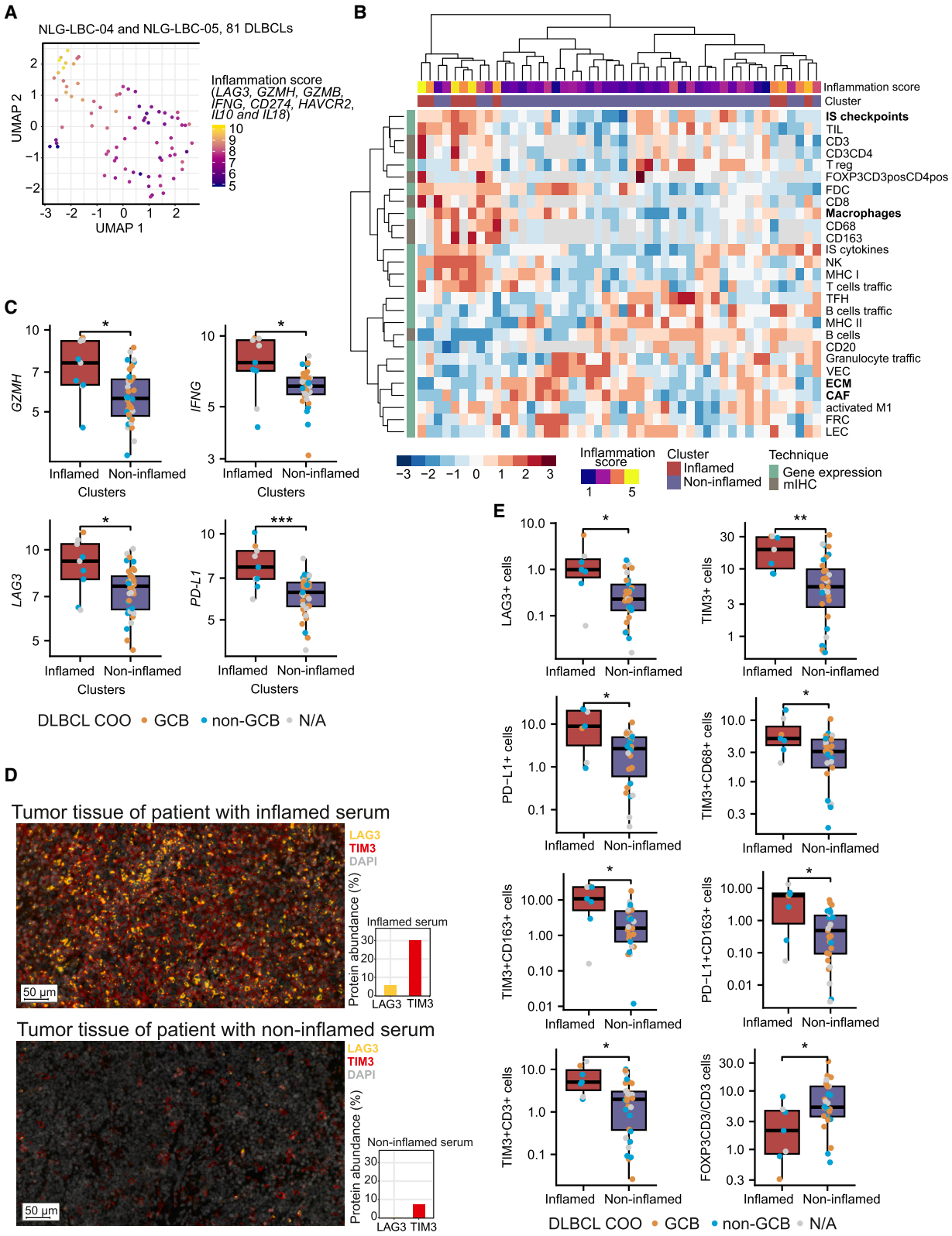


Figure 3. Patients with inflamed serum protein profile express inflammatory genes and tumor-infiltrating lymphocytes in their lymphoma tissue

(A) Dimensionality reduced uniform manifold approximation and projection (UMAP) plot according to gene expression profiling. Expression of inflammatory genes in the patients from two Nordic lymphoma trials profiled with NanoString nCounter Human PanCancer Immunoprofiling Panel. Total number of samples is 81, including samples from the study cohort (n = 65) and NLG-LBC-04 (n = 16) trial. Points represent patients and are colored according to inflammation score of available genes (*LAG3*, *GZMH*, *GZMB*, *IFNG*, *CD274*, *HAVCR2*, *IL10*, *IL18*). See also Figure S3A.

(B) Heatmap of predefined gene expression signatures[7] together with tumor cell phenotypes measured with multiplex immunohistochemistry (mIHC)[12,37] (rows). Columns represent patients. Patient clusters and inflammation scores are annotated on the top of the heatmap. High gene expression levels or proportion of cells in the TME are colored red, whereas low levels are colored blue. Significant differences in gene expression signatures between the clusters are bolded. See also Figures S3B and S3D.

(C) Boxplots depicting gene expression of the selected inflammatory genes (y axis) according to the serum protein clusters (x axis). The points represent individual patients (n = 48) and are colored according to DLBCL COO: orange, GCB DLBCL; blue, non-GCB DLBCL; gray, not available. *$p < 0.05$ and ***$p < 0.001$.

(D) Representative immunofluorescent images of diagnostic tumor tissues stained with selected protein markers. Top, a patient belonging to the inflamed cluster. Bottom, a patient belonging to the non-inflamed cluster. The bar graph illustrates the checkpoint marker proportions of LAG3 and TIM3 in each sample. Yellow color, LAG3; red color, TIM3; gray color, DAPI staining in nuclei.

(E) Boxplots showing the proportions of distinct cell phenotypes (y axis) in the diagnostic tumor biopsies of inflamed and non-inflamed cluster patients. The points represent individual patients (n = 39) and are colored according to the DLBCL COO: orange, GCB DLBCL; blue, non-GCB DLBCL; gray, not available. *$p < 0.05$ and **$p < 0.01$.

LEC, lymphatic endothelial cells; VEC, vascular endothelial cells; CAF, cancer-associated fibroblasts; FRC, fibroblastic reticular cells; IS, immune suppressive; ECM, extracellular matrix; FDC, follicular dendritic cells; MHC, major histocompatibility complex; TFH, T follicular helper cells; TIL, tumor-infiltrating lymphocytes; NK, natural killer cells.

Having found that the inflammatory signature characterized patients also at the gene expression level, we wanted to find out if the inflamed serum protein profile reflected the TME. Using predefined gene expression signatures describing TME statuses[7] and combining these with multiplex immunohistochemistry data on distinct tumor-infiltrating immune cells, we found that several inflammatory gene expression signatures associated with the inflamed cluster (Figures 3B, S3C, and S3D). For example, the signatures defined for macrophages and for immune-suppressive cytokines were upregulated in patients belonging to the inflamed serum cluster ($p < 0.01$ and $p < 0.05$, respectively; Figures 3B and S3D) together with several inflammatory genes, *GZMH*, *IFNG*, *LAG3*, and *PD-L1* (Figure 3C; Table S5). Nonetheless, the patients with inflamed serum were separated into distinct clusters according to the differences in their B cell and follicular dendritic cell gene expression ($p < 0.01$ in both, Kruskal-Wallis test; Figure 3B). Notably, in the non-inflamed patients, pathways concerning extracellular matrix and cancer-associated fibroblasts, which have been described to partly comprise the mesenchymal lymphoma microenvironment,[7] were upregulated ($p < 0.05$ and $p < 0.01$, respectively; Figure S3D).

Similarly, the immune cell proportions in the TME associated with the inflamed and non-inflamed clusters (Figure 3B; Table S6). Furthermore, lymphomas of the patients with an inflamed serum protein profile were characterized by a higher abundance of checkpoint-molecule-positive immune cells, such as TIM3+ T lymphocytes and macrophages and PD-L1+ M2-like macrophages (Figures 3D and 3E). Conversely, relative proportions of the regulatory T cells in the whole T cell population were lower in the lymphomas of inflamed cluster patients (Figure 3E). Moreover, in the patients with GCB DLBCL and a non-inflamed serum protein profile, $CD8^+$ T cells were more often PD1+ (Figures S3E and S3F).

The findings demonstrate a link between the inflamed serum protein profile and TME enriched in checkpoint-positive tumor-infiltrating cells. However, even though we show significant associations, the connection between inflamed serum and inflamed TME remains incomplete, suggesting that other factors, such as tumor burden, spatial heterogeneity, and other biological characteristics, interfere with the origin of the serum protein profiles.

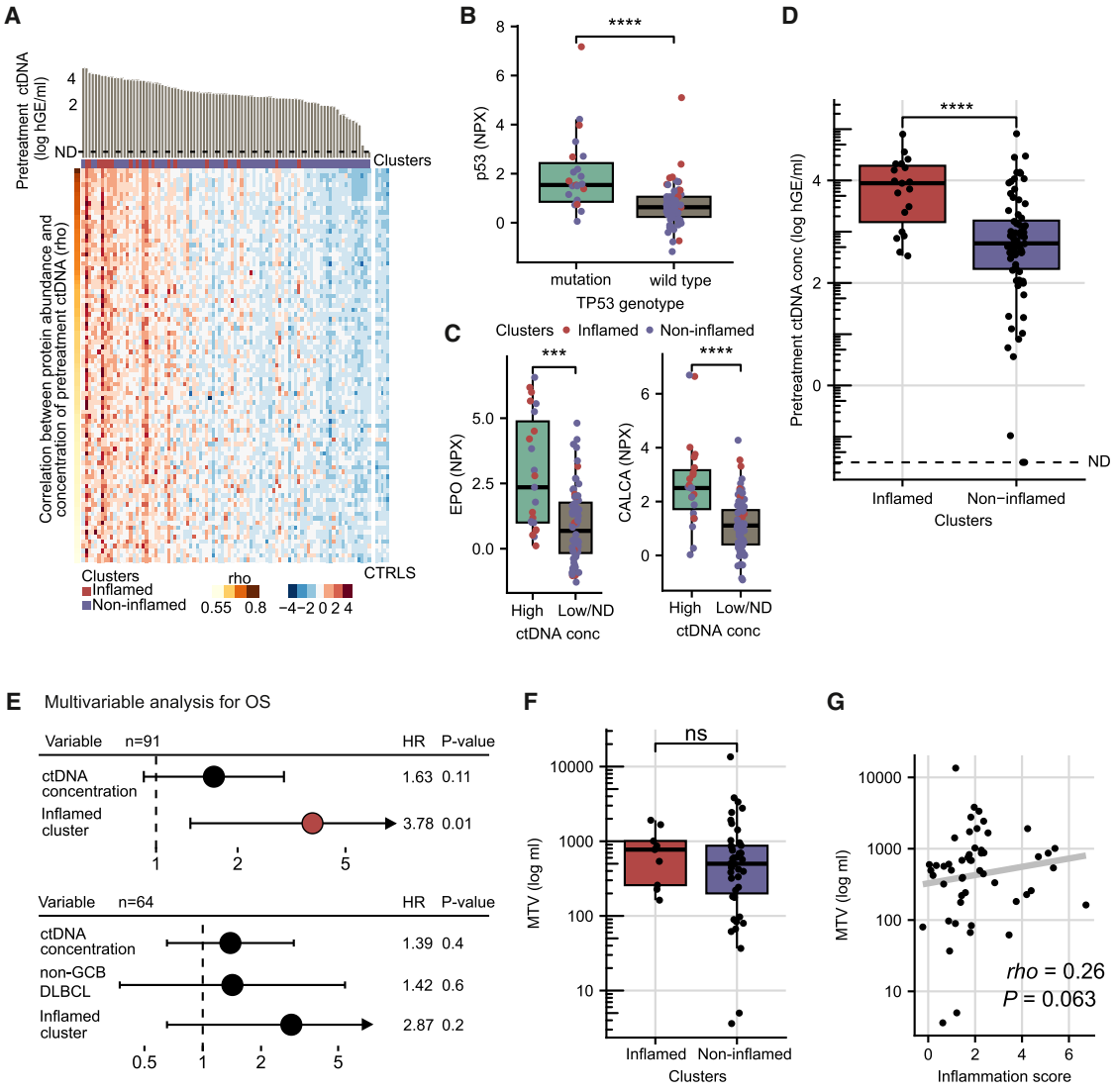


Figure 4. High ctDNA burden is reflected in several serum proteomic features

(A) Heatmap of serum protein levels that correlated with pretreatment ctDNA concentration (rho ≥ 0.5, n = 85). The columns represent patients, and the rows represent serum proteins. The serum protein clusters and the ctDNA concentration are annotated on the top of the heatmap. ctDNA levels below −1.5 log hGE/mL are not detectable. Protein levels of serum samples from healthy donors (CTRLs) are shown on the right (n = 5). See also Figure S4B.

(B) Boxplot of p53 protein concentration (y axis, normalized protein expression [NPX]) according to the TP53 mutation status from ctDNA (x axis). The dots represent patients and are colored according to the inflamed cluster (red) and non-inflamed cluster (purple). ****p < 0.0001. See also Figure S4C.

(C) Boxplots of erythropoietin (EPO; left) and calcitonin-related polypeptide alpha (CALCA; right) concentrations (NPX) according to patients with high (≥3.75 log hGE/mL) or low/not detectable ctDNA concentration. The dots represent patients and are colored according to the inflamed cluster (red) and non-inflamed cluster (purple). ***p < 0.001 and ****p < 0.0001.

(D) Boxplot of pretreatment ctDNA concentration (log hGE/mL, y axis) in the inflamed and non-inflamed cluster patients (x axis). ****p < 0.0001.

(E) Forest plot of Cox-proportional hazard model ratios for OS. Multivariable analyses of ctDNA concentration and inflamed serum protein profile and of ctDNA concentration, non-GCB DLBCL, and inflamed serum protein profile. ctDNA concentration is analyzed as a continuous variable, whereas non-GCB DLCBL and inflamed cluster are analyzed as categorized variables. See also Figure S4E.

(F) Boxplot of metabolic tumor volume (y axis) in the patients of the inflamed and non-inflamed clusters (x axis).

(G) Scatterplot of metabolic tumor volume (n = 52, y axis) and patients' serum inflammation score (x axis).

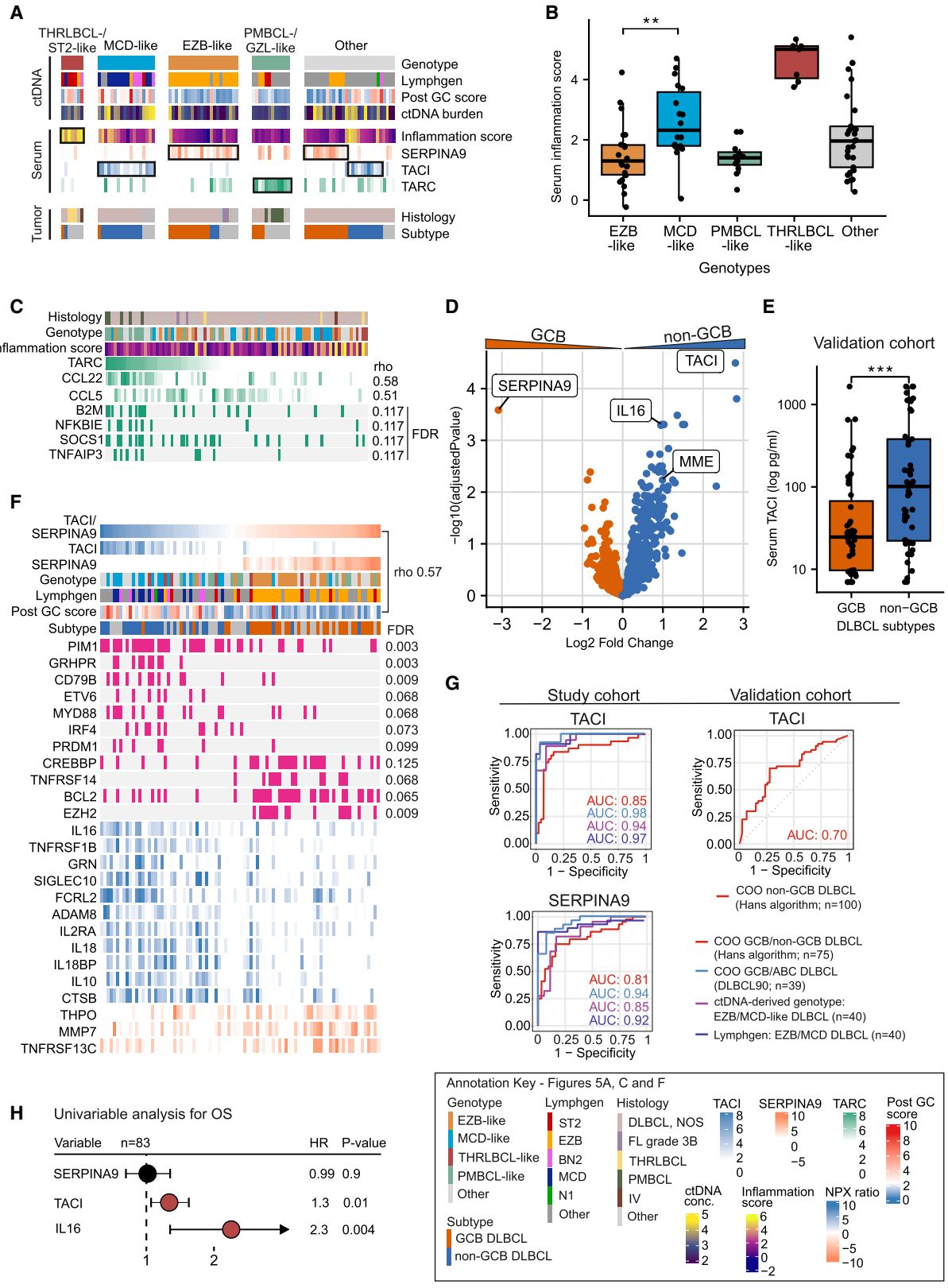
ND, not detectable; ns, not significant; FDR, false discovery rate; HR, hazard ratio; OS, overall survival; PFS, progression-free survival.

Inflamed serum protein profile co-occurs with a high pretreatment ctDNA concentration in plasma

To explore associations with serum protein characteristics beyond tumor biopsies and clinical variables, we expanded our analyses to another dimension of liquid biopsies

Please cite this article in press as: Arffman et al., Inflammatory and subtype-dependent serum protein signatures predict survival beyond the ctDNA in aggressive B cell lymphomas, Med (2024), https://doi.org/10.1016/j.medj.2024.03.007

CellPress
OPEN ACCESS



Please cite this article in press as: Arffman et al., Inflammatory and subtype-dependent serum protein signatures predict survival beyond the ctDNA in aggressive B cell lymphomas, Med (2024), https://doi.org/10.1016/j.medj.2024.03.007

Figure 5. Multiple serum proteins correlate with LBCL subtypes

(A) Oncoprint of ctDNA, serum, and tumor characteristics according to ctDNA-derived genetic subtypes. Columns represent individual patients, and rows different lymphoma characteristics. The associated serum characteristics per genotype are encircled.

(B) Boxplot of serum inflammation score (y axis) between different genotypic entities identified in the ctDNA data (x axis). **$p < 0.01$. See also Figure S4F.

(C) Oncoprint of the patients arranged according to serum TARC levels. Correlation between TARC, CCL22, and CCL5 is annotated as well as driver gene mutations, which associate with TARC concentration. Green bars represent mutations in the samples, and Wilcoxon rank-sum test p values corrected for FDR are shown per gene.

(D) Volcano plot of differential expression analysis comparing the serum protein abundances of GCB DLBCLs (orange) and non-GCB DLBCLs (blue). COO-associated proteins are annotated in the plot. See also Figures S4G and S4H.

(E) Boxplot of TACI levels (y axis) between COO subtypes (x axis) in the validation cohort (n = 100). ***$p < 0.001$.

(F) Oncoprint of patients arranged according to the ratio of TACI and SERPINA9 levels with driver genes associated with DLBCL subtypes. Pink bars represent mutations in the samples, and correlations with the TACI/SERPINA9 NPX ratio using Wilcoxon rank-sum test are annotated. Spearman's correlation coefficient between the TACI/SERPINA9 ratio and post-germinal center (GC) score is shown beside the annotations.

(G) Receiver operator curve analyses of protein abundances as DLBCL subtype predictors in the discovery cohort (TACI and SERPINA9) and validation cohort (TACI). Red lines denote analysis on GCB and non-GCB subtypes, light blue lines denote analysis on DLBCL90 data,[38] purple lines denote analysis on ctDNA derived EZB-like and MCD-like mutational subtypes, and blue lines denote EZB and MCD subtypes from the Lymphgen data, respectively.

(H) Univariable analysis for OS of the COO-associated serum proteins in patients with DLBCL.

based on ctDNA. We discovered that pretreatment ctDNA concentration, which surrogates tumor burden in DLBCL,[10,11] correlated positively with the levels of 85 proteins (6%) in the serum samples (rho $\geq$ 0.5; Figures 4A; Table S7). Furthermore, the majority of the correlating proteins associated with 18-fluorodeoxyglucose positron emission tomography-measured metabolic tumor volume (MTV; Figure S4A). The ctDNA-burden-associated proteins were enriched in pathways reflecting increased cell turnover rate, including cell-cycle regulation, DNA damage, and apoptosis (Figure S4B; Table S8). Among the proteins, high serum levels of p53 correlated with the mutations in *TP53* gene and p53 protein expression in the tumor tissue ($p < 0.0001$; Figures 4B and S4C). Additionally, patients with elevated ctDNA concentrations ($\geq$3.75 log haploid genome equivalents per milliliter plasma [hGE/mL]) were also characterized by high serum levels of erythropoietin ($p < 0.001$) and calcitonin-related polypeptide alpha ($p < 0.0001$; Figures 4C; Table S9), suggesting that the increase in certain serum protein levels may indicate systemic adaptation to the challenges imposed by the high tumor burden. In the analysis stratified according to the cell of origin (COO), we did not identify striking differences, suggesting that the levels of proteins reflecting tumor burden are mostly similar in the distinct COO subtypes (Figure S4D).

When we compared tumor burden between the clusters, we observed that patients with the inflamed serum profile had higher pretreatment plasma ctDNA concentrations compared to the non-inflamed patients ($p < 0.0001$; Figure 4D). In multivariable analysis with ctDNA, the inflamed protein profile retained a prognostic impact on survival (Figures 4E and S4E). However, when adjusted with non-GCB DLBCL, the prognostic impact of the inflamed protein profile was lost, supporting our previous results on the interaction of the inflamed serum profile and non-GCB DLBCL (Figures 4E and S4E). Moreover, MTV (n = 48) did not correlate significantly with the inflamed serum protein profile nor with the continuous inflammation score (rho = 0.26, p = 0.063; Figures 4F and 4G). Altogether, our findings reveal that the inflamed serum profile reflects high ctDNA burden and predicts poor survival independent of ctDNA concentration.

Multiple features in the circulating proteome reflect the characteristics of the ctDNA

Next, we analyzed the association of serum proteins with the phenotypic and genotypic features disclosed by ctDNA. We found that different ctDNA-based genotypic

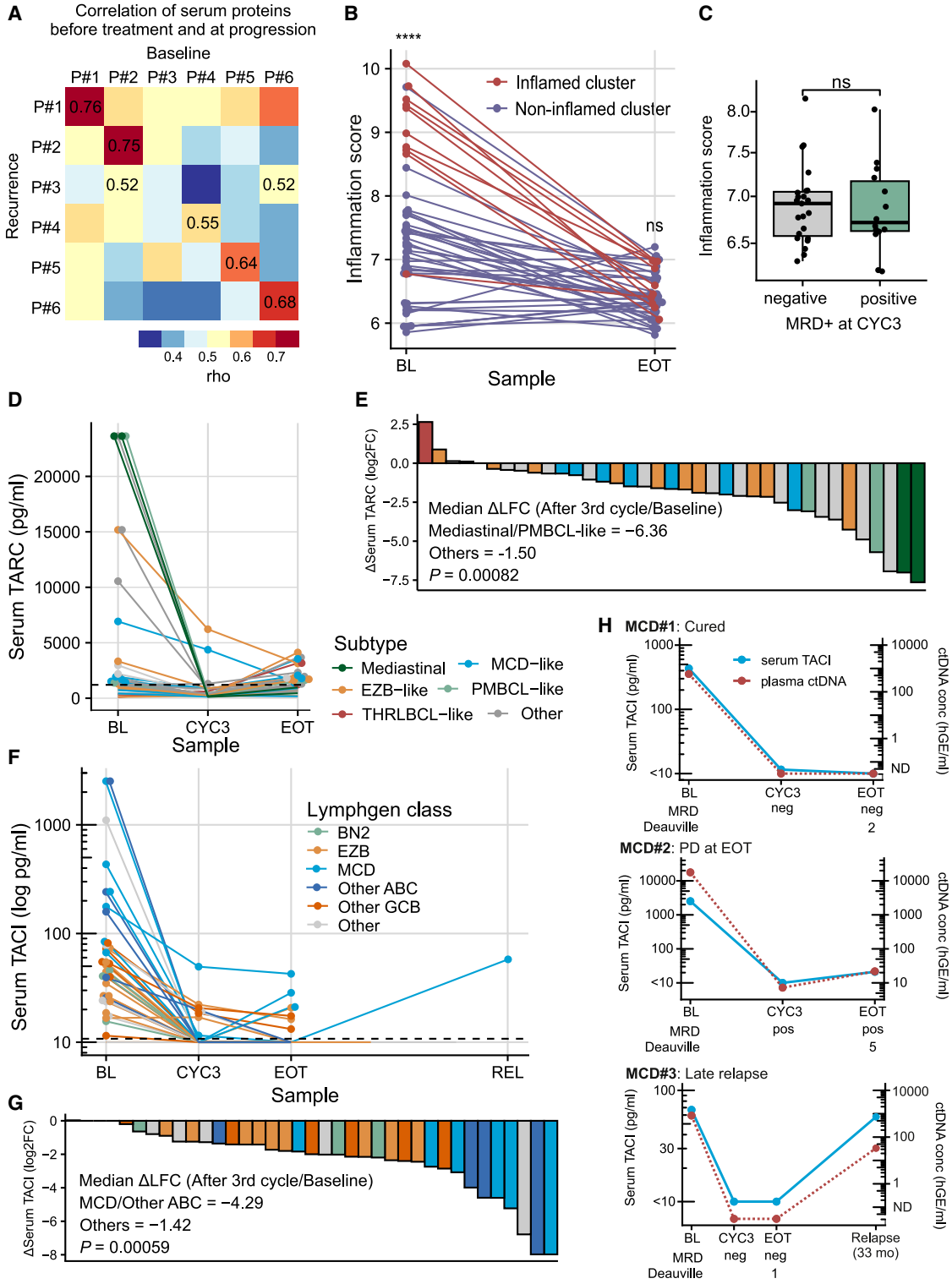


Figure 6. The levels of several serum proteins reflect treatment response

(A) Correlation plot of serum protein profiles from the patients with lymphoma progression (n = 6). Columns represent samples taken at BL, and rows represent samples taken at disease recurrence. Samples that correlate the most are annotated. Red color depicts high correlation and blue low correlation. See also Figure S5A.

Please cite this article in press as: Arffman et al., Inflammatory and subtype-dependent serum protein signatures predict survival beyond the ctDNA in aggressive B cell lymphomas, Med (2024), https://doi.org/10.1016/j.medj.2024.03.007

Figure 6. *Continued*

(B) Dot and line graph of serum inflammation scores before treatment and at EOT. Inflamed cluster patients are colored red, and non-inflamed cluster patients are colored purple. The difference of serum inflammation scores between inflamed and non-inflamed patients are marked at BL (****$p < 0.0001$) and EOT (ns, not significant).

(C) Inflammation scores (y axis) compared between MRD positivity and negativity after third treatment cycle (x axis). Gray box depicts MRD-negative patients, whereas green box depicts MRD-positive patients.

(D) Serum TARC kinetics during therapy. Dots and lines represent patients and are colored according to histological diagnosis (PMBCLs) or genotypes (other than PMBCLs). Black dashed line marks 1,206 pg/mL. See also Figure S5F.

(E) Waterfall plot describing the change in the serum TARC levels measured at BL and after three treatment cycles. Bars represent patients and are colored according to diagnosis (PMBCLs) or genotypes (other than PMBCLs). The change in TARC levels between mediastinal-/PMBCL-like genotypes and others are compared with Spearman correlation.

(F) Serum TACI levels during and after treatment and at disease relapse. Dots and lines represent patients, and colors represent Lymphgen/COO classification. Black dashed line represents the median of healthy controls' TACI levels.

(G) Waterfall plots describing the change in serum TACI levels measured at BL and after three cycles of therapy. Bars represent patients and are colored according to Lymphgen/COO classification. The change in TACI levels between MCD/other ABCs are compared with Spearman correlation.

(H) Patient examples of TACI (y axis, left) and ctDNA (y axis, right) dynamics during treatment, at the EOT, and at disease recurrence. Blue line represents TACI levels, and red line represents ctDNA concentration. MRD at CYC3 and EOT and Deauville score at EOT are marked below each patient's plot.

BL, baseline; EOT, end of therapy; MRD, minimal residual disease; ns, not significant.

subgroups from our previous work[11] correlated with distinct proteomic attributes in the serum (Figure 5A; STAR Methods). Notably, serum samples with a THRLBCL-like genotype, extending to the tumors histologically diagnosed as DLBCL NOS, associated with high serum inflammation scores (Figures 5A, 5B, and S4F). In contrast, PMBCLs and DLBCLs with PMBCL-/gray zone lymphoma-like genotypes[23,24] were associated with low serum inflammation scores but were characterized by elevated levels of thymic-activation-regulated chemokine TARC (CCL17) (Figures 5C and S4G) that promotes PMBCL cell growth[25] and represents a serum protein marker for Hodgkin's lymphoma.[26–28] Moreover, TARC levels correlated with serum levels of other cytokines, CCL22 (macrophage-derived chemokine [MDC]) and CCL5, which are involved in tumor-promoting CCR4+ Th2 T cell recruitment,[29,30] and with genomic drivers described in PMBCL (Figure 5C). Together, these findings show that serum levels of pathogenic markers can be explored and joined with ctDNA-based analysis.

Besides the separate LBCL entities and DLBCL NOS tumors with intermediate characteristics, we found that two serum proteins, SERPINA9 (GCET1) and TACI (TNFRSF13B), correlated with COO subtypes defined from the ctDNA or tissues (Figures 5D–5F). The patients with GCB DLBCL or those with the EZB/C3-like genotype had elevated serum levels of SERPINA9, whereas the patients with non-GCB/ABC DLBCLs or those with the MCD/C5-like genotype were characterized by higher serum levels of TACI and IL-16 (Figures 5D and S4H). Likewise, TACI serum levels were higher in patients with non-GCB DLBCL in our validation cohort ($p < 0.001$; Figure 5E). The ratio of these two proteins was associated with driver gene mutations characteristic for MCD and EZB subtypes and correlated with hypermutation signature-based post-germinal center score[11] in the ctDNA (Figure 5F). Serum levels of TACI and SERPINA9 were predictive for the molecular subtypes with excellent sensitivity and specificity, along with TACI levels predicting non-GCB DLBCL also in the validation cohort (Figure 5G). TACI and IL-16 levels were also associated with poor outcome (Figure 5H). Notably, although SERPINA9 and TACI are both established tissue markers for GCB and non-GCB/-ABC DLBCLs,[31,32] serum levels of some definitive GCB markers, such as membrane metallo-endopeptidase (MME) (CD10), did not reflect the subtype nor CD10 positivity (Figure 5D). Collectively, our results suggest that multiple serum proteins reflect ctDNA characteristics beyond inflammation, such as subtype-related molecular features, which could improve molecular characterization of LBCLs.

Please cite this article in press as: Arffman et al., Inflammatory and subtype-dependent serum protein signatures predict survival beyond the ctDNA in aggressive B cell lymphomas, Med (2024), https://doi.org/10.1016/j.medj.2024.03.007

Distinct serum protein levels reflect variable treatment responses and associate with ctDNA kinetics

After discovering clinical, inflammatory, and molecular disease correlates in the pretreatment serum of the patients with LBCL, we sought to understand how these patterns change during therapy and reflect different clinical courses.

Serum profiles prior to treatment and at the time of lymphoma progression resemble each other

To get an overview of the serum proteome profiles at different disease landmarks, we compared protein profiles of six available serum samples drawn at the time of lymphoma progression with the matched pretreatment samples (Figure 6A). Remarkably, most of the protein profiles were highly concordant between the two time points (rho 0.55–0.75; Figures 6A and S5A), which prompted us to examine the clinical utility of distinct circulating proteins in longitudinal samples using identified markers and to compare the results with ctDNA analysis.

Inflammatory serum profile declines during therapy

During treatment, serum protein estimates of the inflammation score (Luminex inflammation score: IL-10, IL-18, and CXCL9) and tumor burden (levels of serine protease HTRA2 and TNFRSF10B) declined in the patients who had elevated levels at baseline (Figures 6B, S5B, and S5C). In more detail, inflammation scores of the patients with inflamed serum at baseline were, at the end of therapy, similar to the inflammation scores of the other patients (Figure 6B). While the decline in inflammation associated with a pertinent decrease in ctDNA burden (rho = 0.43, $p < 0.0001$; Figure S5D), inflammation score did not associate with minimal residual disease positivity (MRD+) after three treatment cycles (Figure 6C), suggesting limited specificity of these markers to detect residual lymphoma during therapy. Moreover, inflammation score at the end of therapy did not associate with survival (Figure S5E), indicating that the highest clinical relevance for evaluating the inflammatory serum profile is at baseline.

Distinct subtype markers provide indicators of response in PMBCL-like LBCLs and ABC DLBCLs

Finally, we examined dynamic changes in the serum levels of proteins associated with distinct molecular disease subtypes (Figures 6D and 6F). After three treatment cycles, serum levels of both TARC and TACI declined in a subtype-dependent manner, with TARC levels declining below the previously determined cutoff value used for TARC positivity in Hodgkin's lymphoma[28] (1,206 pg/mL; Figures 6E and S5F) and most TACI levels below the accurate level of quantification (Figure 6G). Notably, in two relapsing patients with MCD/C5 ABC DLBCLs, TACI serum levels rebounded at the time of progression together with concurrent ctDNA levels, suggesting a complementary potential of TACI in the evaluation of response and recurrence with ctDNA (Figure 6H). Moreover, TACI serum levels in these patients were concordant with MRD status and Deauville score at the end of therapy (Figure 6H).

Altogether, these findings reveal how multiple serum features change during therapy and provide ctDNA-independent means to assess treatment response.

DISCUSSION

Minimally invasive precision medicine tools of liquid biopsy, including ctDNA-based assays, are anticipated to revolutionize clinical decision-making in the treatment of patients with B cell lymphoma.[9–11] Thus far, however, most studies have examined only somatic mutations, while other circulating factors with synergistic profiling potential

Please cite this article in press as: Arffman et al., Inflammatory and subtype-dependent serum protein signatures predict survival beyond the ctDNA in aggressive B cell lymphomas, Med (2024), https://doi.org/10.1016/j.medj.2024.03.007

have remained poorly characterized and unintegrated with ctDNA. Here, by assessing the levels of 1,463 serum proteins, we identified an inflamed protein profile and other circulating proteins that expand minimally invasive prognostic and diagnostic immune-oncological profiling and estimation of treatment response in patients with LBCL. Moreover, we were able to validate the prognostic inflammatory signature and a subtype-specific profile in an independent validation cohort. Our findings are novel, and we anticipate growing interest in the fusion of ctDNA and serum proteomic features for minimally invasive and more cost-efficient molecular diagnostics.

Pretreatment ctDNA burden has previously been shown to associate with the levels of certain serum proteins, such as LDH at the time of lymphoma diagnosis.[33] We discovered that many other serum proteins, including those involved in apoptosis and thus cell-free DNA production, correlated with elevated ctDNA levels before treatment and with a concordant decline during therapy. As the majority of these cellular turnover-related proteins were not B cell specific, it is plausible to suggest that they could also be used as a surrogate to quantify tumor burden in other cancers.

In addition to serum proteins that reflect tumor burden, we uncovered several proteins, which correlated with molecular features of different lymphoma entities. For example, elevated levels of TARC, which is an established marker in Hodgkin's lymphoma,[26,27] were strongly associated with PMBCL-/GZ-like genotypes. As PMBCL-/GZ-like LBCLs can be difficult to biopsy due to their location in the mediastinum, new assessment methods are needed, and our results suggest that serum TARC levels could serve as an additional tool to diagnose these entities. Moreover, circulating TACI and SERPINA9 levels could be used to distinguish ABC and GCB DLBCLs. Given the specificity of these markers for B cells, we utilized TARC and TACI in complementing ctDNA-based molecular response evaluation and detected mutual potential between these liquid biopsies. These results suggest that joint analyses of serum proteins and ctDNA in baseline characterization could overcome the limitations of tissue biopsies and thereby improve molecular characterization and response evaluation. However, the clinical value of these markers warrants prospective validation. In addition, their feasibility in relapsed and refractory settings remains to be explored.

Although multiple dimensions of the tumor can be assessed with ctDNA, long turnaround times, special equipment, and limited resources impose contemporary challenges to routine ctDNA analyses. Furthermore, ctDNA-based assays do not capture the host's response to cancer. This warrants assessment of risk profiles and diagnostic characteristics by complementary methods such as serum or plasma protein profiling, which could provide results faster, as required in patient care. In this study, we identified a host response in a subset of patients with LBCL, which was molecularly characterized as systemic inflammation and associated with aggressive clinical characteristics and B symptoms. The hallmarks of systemic inflammation in response to cancer include increased production of proinflammatory cytokines and acute-phase proteins that enter the circulation,[34] resulting in many disabling symptoms, including fever and weight loss as examples of B symptoms. In patients with DLBCL, inflammatory characteristics have been described both in the tumor tissue and circulation.[7,12,13,16] Recently, Ask et al. described a protein profile in serum and plasma represented by a systemic protein deviation score, which correlated with the amount of PD1+ $CD8^+$ T cells in the peripheral blood and poor outcome.[13] Likewise, in our study, the patients with an inflamed serum protein profile had significantly worse survival compared to the patients with non-inflamed serum. In addition, their lymphomas were characterized by increased levels of gene expression pathways associating with the inflammatory TME. Similarly, they had higher abundances of checkpoint-protein-expressing

Please cite this article in press as: Arffman et al., Inflammatory and subtype-dependent serum protein signatures predict survival beyond the ctDNA in aggressive B cell lymphomas, Med (2024), https://doi.org/10.1016/j.medj.2024.03.007

immune cells, such as TIM3+ T lymphocytes and macrophages and PD-L1+ M2-type macrophages, whereas the proportion of regulatory T cells was lower. Together, the findings suggest that systemic inflammation combined with immune cell exhaustion creates a favorable environment for lymphoma development. Notably, inflammation reduced during therapy in patients with an inflamed serum protein profile, and besides TME characteristics, we believe that other biological factors such as tumor burden interfere with the onset of systemic inflammation profile.

We observed a novel association between inflamed serum protein profile and elevated ctDNA levels, which suggests that high tumor burden provokes inflammatory host response. On the other hand, a correlation between inflamed protein profile and biological features that are not associated with tumor burden, including ABC DLBCL, THRLBCL, and a high proportion of checkpoint-positive tumor-infiltrating immune cells, proposes that molecular tumor characteristics also play a part in the inflamed profile seen in the serum. Moreover, immunochemotherapy and granulocyte colony-stimulating factors contribute to the release of proteins into circulation by both healthy and malignant cells, thereby affecting the inflammatory protein profiles detected at the end of therapy. These aspects, along with the weak correlation with MTV, suggest that the inflamed serum protein profile is confounded by multiple factors in the MRD setting. Further understanding of the biological and clinical components contributing to the systemic inflammatory state is needed. For instance, a high proportion of checkpoint-positive T cells and macrophages in the lymphoma tissue of the patients with inflamed serum protein profiles implies that these patients could benefit from targeted immunotherapies. So far, however, the response rates to such therapies, namely PD1/PD-L1 blockage, have been low in DLBCL.[35,36]

The strengths of our study are prospectively collected and homogenously treated patient populations and the availability of longitudinally collected blood samples, tumor tissue, and comprehensive clinical data from the same patients that provide a possibility to correlate serum proteome data with ctDNA, tumor tissue, and patient demographics. In addition, we were able to validate our findings in an independent LBCL cohort with similar demographics, treatment, and follow-up data. We note that the patients included in the study consist of young, high-risk patients treated in clinical trials, and the results are reflected by the patient demographics, treatment, and survival. An additional strength of using circulating proteins is that the markers are also applicable to plasma samples,[13] which are usually collected in clinical routine.

Here, we combine two types of liquid biopsies, serum protein profiling and ctDNA, to characterize high-risk LBCL. We describe a systemic inflammatory response detected from the serum proteome, which associates with poor survival in two uniformly treated LBCL cohorts. We discover proteomic attributes that correlate with TME features and complement multiple ctDNA characteristics, which, with high clinical utility, could be harnessed both for non-invasive and accurate diagnostics as well as to guide treatment decisions and assess treatment response. We believe that the multidimensional potential of serum protein profiling extends the usage of liquid biopsy in the lymphoma field and hypothesize growing interest of serum protein profiling in oncology.

Limitations of the study

The limitations of our study include a small number of non-cancer control sera, and therefore the specificity of these findings to patients with lymphoma remains to be established. Moreover, we recognize that our study did not consider whether the measured proteins were soluble or incorporated in vesicles, such as exosomes,

which may have an impact on their detection. Nevertheless, the findings establish novel, clinically meaningful interactions between lymphoma and the host and expand liquid biopsy repertoire in patients with LBCL.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODELS AND SUBJECT DETAILS
  - Patients and samples
  - Discovery cohort and available data
  - Validation cohort
  - Additional samples, cohorts, and data
- METHOD DETAILS
  - Antibody-based proximity extension assay
  - Identification of clusters and serum protein associations
  - Bead-based immunoassay
  - Inflammation score
  - Gene expression profiling
  - Tissue protein expression analysis
  - Circulating tumor DNA analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Survival analyses

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.medj.2024.03.007.

ACKNOWLEDGMENTS

We thank Anne Aarnio for excellent technical assistance. This research was funded by grants from the Research Council of Finland (#357481; S.L.), Finnish Cancer Organizations (S.L.), Sigrid Jusélius Foundation (S.L.), University of Helsinki (S.L.), iCAN Digital Precision Cancer Medicine Flagship (S.L.), Orion Research Foundation sr (M. Arffman), and Helsinki University Hospital (S.L.). Part of the work was presented at the 64th ASH Annual Meeting.

AUTHOR CONTRIBUTIONS

Conception and design, M. Arffman, L.M., S.-K.L., and S.L.; provision of study materials or patients, H.H., J.J., P.B., S.J., M.J., O.F., M.B., and S.L.; collection and assembly of data, M. Arffman, L.M., M. Autio, H.H., S.-K.L., and S.L.; data analysis and interpretation, M. Arffman, L.M., M. Autio, S.-K.L., and S.L.; manuscript writing, M. Arffman, L.M., S.-K.L., and S.L.; final approval of manuscript, all authors; accountable for all aspects of the work including the manuscript, all authors.

DECLARATION OF INTERESTS

H.H. (all outside of the submitted work): Genmab: honoraria, safety committee; Gilead: honoraria, advisory board; Incyte: honoraria, advisory board; Nordic Nanovector: honoraria, safety committee; Novartis: honoraria, advisory board; Takeda: honoraria, advisory board. J.J. (all outside of the submitted work): BMS: consultancy;

Gilead: consultancy; Incyte: consultancy; Novartis: consultancy; Orion Pharma: consultancy; Roche: consultancy. M.B. (all outside of the submitted work): Astra Zeneca: consultancy; BMS/Celgene: consultancy; Incyte: consultancy; Janssen: consultancy; Mundipharma: consultancy; Nanexa: consultancy; Pfizer: consultancy; Roche: consultancy; Schain Research: consultancy; WntResearch: consultancy. M.J. (all outside of the submitted work): Abbvie: honoraria, research funding; Astra Zeneca: honoraria, research funding; BMS: honoraria, research funding; Genmab: honoraria; Incyte: honoraria; Janssen: honoraria, research funding; Kite/Gilead: consultancy, honoraria, research funding; Novartis: honoraria; Orion: honoraria; Roche: honoraria, research funding. S.L. (all outside of the submitted work): Genmab: consultancy, research funding; Gilead: consultancy; Incyte: consultancy; Nordic Nanovector: research funding; Novartis: consultancy, honoraria, research funding; Roche: consultancy, research funding; Merck: consultancy; Bayer: research funding; Celgene: consultancy, research funding; Orion: consultancy.

Received: January 6, 2024
Revised: February 24, 2024
Accepted: March 11, 2024
Published: April 4, 2024

REFERENCES

1. Khoury, J.D., Solary, E., Abla, O., Akkari, Y., Alaggio, R., Apperley, J.F., Bejar, R., Berti, E., Busque, L., Chan, J.K.C., et al. (2022). The 5th edition of the World Health Organization classification of haematolymphoid tumours: myeloid and histiocytic/dendritic neoplasms. Leukemia *36*, 1703–1719. https://doi.org/10.1038/s41375-022-01613-1.

2. Campo, E., Jaffe, E.S., Cook, J.R., Quintanilla-Martinez, L., Swerdlow, S.H., Anderson, K.C., Brousset, P., Cerroni, L., de Leval, L., Dirnhofer, S., et al. (2022). The international consensus classification of mature lymphoid neoplasms: a report from the clinical advisory committee. Blood *140*, 1229–1253. https://doi.org/10.1182/blood.2022015851.

3. Sehn, L.H., and Salles, G. (2021). Diffuse large B-cell lymphoma. N. Engl. J. Med. *384*, 842–858. https://doi.org/10.1056/NEJMra2027612.

4. Alizadeh, A.A., Eisen, M.B., Davis, R.E., Ma, C., Lossos, I.S., Rosenwald, A., Boldrick, J.C., Sabet, H., Tran, T., Yu, X., et al. (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature *403*, 503–511. https://doi.org/10.1038/35000501.

5. Schmitz, R., Wright, G.W., Huang, D.W., Johnson, C.A., Phelan, J.D., Wang, J.Q., Roulland, S., Kasbekar, M., Young, R.M., Shaffer, A.L., et al. (2018). Genetics and pathogenesis of diffuse large B-cell lymphoma. N. Engl. J. Med. *378*, 1396–1407. https://doi.org/10.1056/NEJMoa1801445.

6. Chapuy, B., Stewart, C., Dunford, A.J., Kim, J., Kamburov, A., Redd, R.A., Lawrence, M.S., Roemer, M.G.M., Li, A.J., Ziepert, M., et al. (2018). Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. Nat. Med. *24*, 679–690. https://doi.org/10.1038/s41591-018-0016-8.

7. Kotlov, N., Bagaev, A., Revuelta, M.V., Phillip, J.M., Cacciapuoti, M.T., Antysheva, Z., Svekolkin, V., Tikhonova, E., Miheecheva, N., Kuzkina, N., et al. (2021). Clinical and Biological Subtypes of B-cell Lymphoma Revealed by Microenvironmental Signatures in Lymphoma. Cancer Discov. *11*, 1468–1489. https://doi.org/10.1158/2159-8290.CD-20-0839.

8. Steen, C.B., Luca, B.A., Esfahani, M.S., Azizi, A., Sworder, B.J., Nabet, B.Y., Kurtz, D.M., Liu, C.L., Khameneh, F., Advani, R.H., et al. (2021). The landscape of tumor cell states and ecosystems in diffuse large B cell lymphoma. Cancer Cell *39*, 1422–1437.e10. https://doi.org/10.1016/j.ccell.2021.08.011.

9. Kurtz, D.M., Scherer, F., Jin, M.C., Soo, J., Craig, A.F.M., Esfahani, M.S., Chabon, J.J., Stehr, H., Liu, C.L., Tibshirani, R., et al. (2018). Circulating tumor DNA measurements as early outcome predictors in diffuse large B-cell lymphoma. J. Clin. Oncol. *36*, 2845–2853. https://doi.org/10.1200/JCO.2018.78.5246.

10. Scherer, F., Kurtz, D.M., Newman, A.M., Stehr, H., Craig, A.F.M., Esfahani, M.S., Lovejoy, A.F., Chabon, J.J., Klass, D.M., Liu, C.L., et al. (2016). Distinct biological subtypes and patterns of genome evolution in lymphoma revealed by circulating tumor DNA. Sci. Transl. Med. *8*, 364ra155. https://doi.org/10.1126/scitranslmed.aai8545.

11. Meriranta, L., Alkodsi, A., Pasanen, A., Lepistö, M., Mapar, P., Blaker, Y.N., Jørgensen, J., Karjalainen-Lindsberg, M.-L., Fiskvik, I., Mikalsen, L.T.G., et al. (2022). Molecular features encoded in the ctDNA reveal heterogeneity and predict outcome in high-risk aggressive B-cell lymphoma. Blood *139*, 1863–1877. https://doi.org/10.1182/blood.2021012852.

12. Autio, M., Leivonen, S.-K., Brück, O., Mustjoki, S., Mészáros Jørgensen, J., Karjalainen-Lindsberg, M.-L., Beiske, K., Holte, H., Pellinen, T., and Leppä, S. (2021). Immune cell constitution in the tumor microenvironment predicts the outcome in diffuse large B-cell lymphoma. Haematologica *106*, 718–729. https://doi.org/10.3324/haematol.2019.243626.

13. Ask, E.H., Tschan-Plessl, A., Gjerdingen, T.J., Sætersmoen, M.L., Hoel, H.J., Wiiger, M.T., Olweus, J., Wahlin, B.E., Lingjærde, O.C., Horowitz, A., et al. (2021). A Systemic Protein Deviation Score Linked to PD-1+ CD8+ T Cell Expansion That Predicts Overall Survival in Diffuse Large B Cell Lymphoma. Med *2*, 180–195.e5. https://doi.org/10.1016/j.medj.2020.10.006.

14. Pauly, F., Fjordén, K., Leppä, S., Holte, H., Björkholm, M., Fluge, Ø., Møller Pedersen, L., Eriksson, M., Isinger-Ekstrand, A., Borrebaeck, C.A.K., et al. (2016). Plasma immunoprofiling of patients with high-risk diffuse large B-cell lymphoma: a Nordic Lymphoma Group study. Blood Cancer J. *6*, e501. https://doi.org/10.1038/bcj.2016.113.

15. Vajavaara, H., Mortensen, J.B., Leivonen, S.-K., Hansen, I.M., Ludvigsen, M., Holte, H., Jørgensen, J., Bjerre, M., d'Amore, F., and Leppä, S. (2021). Soluble PD-1 but not PD-L1 levels predict poor outcome in patients with high-risk diffuse large B-cell lymphoma. Cancers *13*, 398. https://doi.org/10.3390/cancers13030398.

16. Rossille, D., Azzaoui, I., Feldman, A.L., Maurer, M.J., Labouré, G., Parrens, M., Pangault, C., Habermann, T.M., Ansell, S.M., Link, B.K., et al. (2017). Soluble programmed death-ligand 1 as a prognostic biomarker for overall survival in patients with diffuse large B-cell lymphoma: a replication study and combined analysis of 508 patients. Leukemia *31*, 988–991. https://doi.org/10.1038/leu.2016.385.

17. Rossille, D., Gressier, M., Damotte, D., Maucort-Boulch, D., Pangault, C., Semana, G.,

Le Gouill, S., Haioun, C., Tarte, K., Lamy, T., et al. (2014). High level of soluble programmed cell death ligand 1 in blood impacts overall survival in aggressive diffuse large B-Cell lymphoma: results from a French multicenter clinical trial. Leukemia *28*, 2367–2375. https://doi.org/10.1038/leu.2014.137.

18. Leppä, S., Jørgensen, J., Tierens, A., Meriranta, L., Østlie, I., de Nully Brown, P., Fagerli, U.-M., Larsen, T.S., Mannisto, S., Munksgaard, L., et al. (2020). Patients with high-risk DLBCL benefit from dose-dense immunochemotherapy combined with early systemic CNS prophylaxis. Blood Adv. *4*, 1906–1915. https://doi.org/10.1182/bloodadvances.2020001518.

19. Leppä, S., Jørgensen, J.M., Karjalainen-Lindsberg, M.-L., Beiske, K., Pedersen, M., Drott, K., Pasanen, A., Karihtala, K., Mannisto, S., Meriranta, L., et al. (2022). Biomarker-Driven Treatment Strategy in High-Risk Large B-Cell Lymphoma: Final Results of a Nordic Phase 2 Study. Blood *140*, 1768–1769. https://doi.org/10.1182/blood-2022-158183.

20. Holte, H., Leppä, S., Björkholm, M., Fluge, Ø., Jyrkkiö, S., Delabie, J., Sundström, C., Karjalainen-Lindsberg, M.-L., Erlanson, M., Kolstad, A., et al. (2013). Dose-densified chemoimmunotherapy followed by systemic central nervous system prophylaxis for younger high-risk diffuse large B-cell/follicular grade 3 lymphoma patients: results of a phase II Nordic Lymphoma Group study. Ann. Oncol. *24*, 1385–1392. https://doi.org/10.1093/annonc/mds621.

21. Reddy, A., Zhang, J., Davis, N.S., Moffitt, A.B., Love, C.L., Waldrop, A., Leppä, S., Pasanen, A., Meriranta, L., Karjalainen-Lindsberg, M.-L., et al. (2017). Genetic and functional drivers of diffuse large B cell lymphoma. Cell *171*, 481–494.e15. https://doi.org/10.1016/j.cell.2017.09.027.

22. Dufva, O., Pölönen, P., Brück, O., Keränen, M.A.I., Klievink, J., Mehtonen, J., Huuhtanen, J., Kumar, A., Malani, D., Siitonen, S., et al. (2020). Immunogenomic landscape of hematological malignancies. Cancer Cell *38*, 380–399.e13. https://doi.org/10.1016/j.ccell.2020.06.002.

23. Mottok, A., Hung, S.S., Chavez, E.A., Woolcock, B., Telenius, A., Chong, L.C., Meissner, B., Nakamura, H., Rushton, C., Viganò, E., et al. (2019). Integrative genomic analysis identifies key pathogenic mechanisms in primary mediastinal large B-cell lymphoma. Blood *134*, 802–813. https://doi.org/10.1182/blood.2019001126.

24. Sarkozy, C., Hung, S.S., Chavez, E.A., Duns, G., Takata, K., Chong, L.C., Aoki, T., Jiang, A., Miyata-Takata, T., Telenius, A., et al. (2021). Mutational landscape of gray zone lymphoma. Blood *137*, 1765–1776. https://doi.org/10.1182/blood.2020007507.

25. Viganò, E., Gunawardana, J., Mottok, A., Van Tol, T., Mak, K., Chan, F.C., Chong, L., Chavez, E., Woolcock, B., Takata, K., et al. (2018). Somatic IL4R mutations in primary mediastinal large B-cell lymphoma lead to constitutive JAK-STAT signaling activation. Blood *131*, 2036–2046. https://doi.org/10.1182/blood-2017-09-808907.

26. Plattel, W.J., van den Berg, A., Visser, L., van der Graaf, A.-M., Pruim, J., Vos, H., Hepkema, B., Diepstra, A., and van Imhoff, G.W. (2012). Plasma thymus and activation-regulated chemokine as an early response marker in classical Hodgkin's lymphoma. Haematologica *97*, 410–415. https://doi.org/10.3324/haematol.2011.053199.

27. Driessen, J., Kersten, M.J., Visser, L., van den Berg, A., Tonino, S.H., Zijlstra, J.M., Lugtenburg, P.J., Morschhauser, F., Hutchings, M., Amorim, S., et al. (2022). Prognostic value of TARC and quantitative PET parameters in relapsed or refractory Hodgkin lymphoma patients treated with brentuximab vedotin and DHAP. Leukemia *36*, 2853–2862. https://doi.org/10.1038/s41375-022-01717-8.

28. Diepstra, A., Nolte, I.M., van den Berg, A., Magpantay, L.I., Martínez-Maza, O., and Levin, L.I. (2023). Elevated serum TARC levels precede classic Hodgkin lymphoma diagnosis by several years. Blood *142*, 1928–1931. https://doi.org/10.1182/blood.2023020959.

29. Imai, T., Nagira, M., Takagi, S., Kakizaki, M., Nishimura, M., Wang, J., Gray, P.W., Matsushima, K., and Yoshie, O. (1999). Selective recruitment of CCR4-bearing Th2 cells toward antigen-presenting cells by the CC chemokines thymus and activation-regulated chemokine and macrophage-derived chemokine. Int. Immunol. *11*, 81–88. https://doi.org/10.1093/intimm/11.1.81.

30. Iellem, A., Mariani, M., Lang, R., Recalde, H., Panina-Bordignon, P., Sinigaglia, F., and D'Ambrosio, D. (2001). Unique chemotactic response profile and specific expression of chemokine receptors CCR4 and CCR8 by CD4+ CD25+ regulatory T cells. J. Exp. Med. *194*, 847–853. https://doi.org/10.1084/jem.194.6.847.

31. Montes-Moreno, S., Roncador, G., Maestre, L., Martínez, N., Sanchez-Verde, L., Camacho, F.I., Cannata, J., Martinez-Torrecuadrada, J.L., Shen, Y., Chan, W.C., and Piris, M.A. (2008). Gcet1 (centerin), a highly restricted marker for a subset of germinal center-derived lymphomas. Blood *111*, 351–358. https://doi.org/10.1182/blood-2007-06-094151.

32. Wada, K., Maeda, K., Tajima, K., Kato, T., Kobata, T., and Yamakawa, M. (2009). Expression of BAFF-R and TACI in reactive lymphoid tissues and B-cell lymphomas. Histopathology *54*, 221–232. https://doi.org/10.1111/j.1365-2559.2008.03203.x.

33. Kurtz, D.M., Green, M.R., Bratman, S.V., Scherer, F., Liu, C.L., Kunder, C.A., Takahashi, K., Glover, C., Keane, C., Kihira, S., et al. (2015). Noninvasive monitoring of diffuse large B-cell lymphoma by immunoglobulin high-throughput sequencing. Blood *125*, 3679–3687. https://doi.org/10.1182/blood-2015-03-635169.

34. Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674. https://doi.org/10.1016/j.cell.2011.02.013.

35. Juárez-Salcedo, L.M., Sandoval-Sus, J., Sokol, L., Chavez, J.C., and Dalia, S. (2017). The role of anti-PD-1 and anti-PD-L1 agents in the treatment of diffuse large B-cell lymphoma: the future is now. Crit. Rev. Oncol. Hematol. *113*, 52–62. https://doi.org/10.1016/j.critrevonc.2017.02.027.

36. Ansell, S.M., Minnema, M.C., Johnson, P., Timmerman, J.M., Armand, P., Shipp, M.A., Rodig, S.J., Ligon, A.H., Roemer, M.G.M., Reddy, N., et al. (2019). Nivolumab for relapsed/refractory diffuse large B-cell lymphoma in patients ineligible for or having failed autologous transplantation: a single-arm, phase II study. J. Clin. Oncol. *37*, 481–489. https://doi.org/10.1200/JCO.18.00766.

37. Autio, M., Leivonen, S.-K., Brück, O., Karjalainen-Lindsberg, M.-L., Pellinen, T., and Leppä, S. (2022). Clinical Impact of Immune Cells and Their Spatial Interactions in Diffuse Large B-Cell Lymphoma Microenvironment. Clin. Cancer Res. *28*, 781–792. https://doi.org/10.1158/1078-0432.CCR-21-3140.

38. Isaksen, K.T., Beiske, K., Smeland, E.B., Jørgensen, J., Brodtkorb, M., Myklebust, J.H., Jerkeman, M., Meriranta, L., Karjalainen-Lindsberg, M.-L., Leppä, S., et al. (2019). The DLBCL90 Double-Hit Gene Expression Signature Is Not Associated with Inferior Survival in Young High-Risk Patients with Diffuse Large B-Cell Lymphoma Treated with Dose-Intensive Immunochemotherapy. Blood *134*, 1485. https://doi.org/10.1002/jha2.109.

39. Wilkerson, M.D., and Hayes, D.N. (2010). ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics *26*, 1572–1573. https://doi.org/10.1093/bioinformatics/btq170.

40. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47. https://doi.org/10.1093/nar/gkv007.

41. Kassambara, A., Kosinski, M., and Biecek, P. (2021). In survminer: Drawing Survival Curves using 'ggplot2', R package version 0.4.9.

42. Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinf. *14*, 7–15. https://doi.org/10.1186/1471-2105-14-7.

43. Leppä, S., Meriranta, L., Arffman, M., Jørgensen, J., Karjalainen-Lindsberg, M., Beiske, K., Pedersen, M., Drott, K., Fluge, Ø., Jyrkkiö, S., et al. (2023). Biomarker-Driven Treatment Strategy in High-Risk Large B-Cell Lymphoma (NLG-LBC-06 Phase II Trial): Impact of ctDNA and TP53 Aberrations on Clinical Outcome. Hematol. Oncol. *41*, 51–52. https://doi.org/10.1002/hon.3163_21.

44. Wright, G.W., Phelan, J.D., Coulibaly, Z.A., Roulland, S., Young, R.M., Wang, J.Q., Schmitz, R., Morin, R.D., Tang, J., Jiang, A., et al. (2020). A probabilistic classification tool for genetic subtypes of diffuse large B cell lymphoma with therapeutic implications. Cancer Cell *37*, 551–568.e514. https://doi.org/10.1016/j.ccell.2020.03.015.

Please cite this article in press as: Arffman et al., Inflammatory and subtype-dependent serum protein signatures predict survival beyond the ctDNA in aggressive B cell lymphomas, Med (2024), https://doi.org/10.1016/j.medj.2024.03.007

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Serum (discovery cohort)	This paper and Leppä et al.[18]	www.ClinicalTrials.gov, NCT01325194
Serum (validation cohort)	This paper and Leppä et al.[19]	www.ClinicalTrials.gov, NCT03293173
Serum (healthy donors)	This paper	N/A
Critical commercial assays		
Olink Explore 1536 library	Olink Proteomics	www.olink.com
Human Magnetic Luminex Assay	R&D Systems	LXSAHM-08
Human Magnetic Luminex Assay	R&D Systems	LXSAHM-07
Deposited data		
Proximity Extension Assay data	This paper	N/A
Bead-based immunoassay (Luminex) data	This paper	N/A
Clinical patient data (discovery cohort)	Leppä et al.[18]	www.ClinicalTrials.gov, NCT01325194
Clinical patient data (validation cohort)	Leppä et al.[19]	www.ClinicalTrials.gov, NCT03293173
nCounter® PanCancer Immune Profiling Panel	Autio et al.[12]	N/A
DLBCL90 assay	Isaksen et al.[38]	N/A
RNAseq gene expression data	Reddy et al.[21]	N/A
multiplex IHC data	Autio et al.[12,37]	N/A
Targeted sequencing data	Meriranta et al.[11]	EGAS00001005835
Software and algorithms		
R (version ≥3.6.1)	R core team	www.R-project.org
limma (version 3.50.3)	Ritchie et al.[40]	https://bioconductor.org/packages/release/bioc/html/limma.html
survminer (version 0.4.9)	Kassambara et al.[41]	https://github.com/kassambara/survminer
GSVA (version 1.42.0)	Hänzelmann et al.[42]	https://github.com/rcastelo/GSVA
ConsensusClusterPlus (version 1.58.0)	Wilkerson et al.[39]	https://bioconductor.org/packages/release/bioc/html/ConsensusClusterPlus.html
Lymphgen 2.0	Wright et al.[44]	https://llmpp.nih.gov/lymphgen/lymphgendataportal.php?version=2.0

RESOURCE AVAILABILITY

Lead contact

Further information and requests for data should be directed to the lead contact, Sirpa Leppä (sirpa.leppa@helsinki.fi).

Materials availability

The protein data used in this study is available upon reasonable request from the corresponding author.

Data and code availability

The patient characteristics of the discovery and validation cohorts are presented in Tables 1 and S4. Identifier numbers are listed in the key resources table. This paper does not report any original codes. Any additional information of the code used in this study is available upon reasonable request from the lead contact.

EXPERIMENTAL MODELS AND SUBJECT DETAILS

Patients and samples

All patients in the discovery and validation cohorts signed informed consent before study participation. The Institutional Review Boards, National Medical Agencies,

and Ethics Committees in Finland, Norway, Denmark, and Sweden approved the protocols and sampling.

Discovery cohort and available data

Discovery cohort consisted of 109 patients aged 21–64 years with high-risk (age adjusted International Prognostic Index (aaIPI) $\geq$ 2 and/or site-specific risk factors for central nervous system (CNS) progression) LBCL, who were included in the Nordic LBC-05 phase II trial (registered at www.ClinicalTrials.gov, NCT01325194).[18] Treatment consisted of two courses of high-dose methotrexate (HD-Mtx) in combination with biweekly rituximab (R), cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP-14), followed by four courses of biweekly R-CHOP with etoposide (R-CHOEP-14) and one course of R-HD-cytarabine. The patient characteristics are described in Table 1. Pre-treatment serum samples from 109 patients and post-treatment serum samples collected at the time of lymphoma progression from six patients were available for analysis. In addition, 47 pre-treatment serum samples were available for technical validation in the Luminex analysis. Gene expression profiling (ncounter PanCancer Immune Profiling Panel [NanoString Technologies, Seattle, WA, United States], n = 48)[12] and mIHC (n = 39) data from tumor tissue of overlapping patients were previously produced.[12,37] In addition, metabolic tumor volume at diagnosis (n = 52) and ctDNA data (n = 91)[11] were available for correlative analyses.

Validation cohort

Independent validation cohort consisted of pre-treatment serum samples and clinical data from 122 patients included in the Nordic LBC-06 phase II trial (registered at www.ClinicalTrials.gov, number NCT03293173).[19] The cohort had similar inclusion criteria and treatment as the discovery population, with following differences: PMBCLs were excluded and depending on the biological risk factors patients were stratified to receive either R-CHOEP-14 (no biological risk factors) or dose-adjusted etoposide, doxorubicin, cyclophosphamide, vincristine, prednisone, and rituximab (DA-EPOCH-R; biological risk factors). Biological high-risk was defined as a presence of at least one of the following factors: C-MYC translocation, DH, 17p/TP53 deletion, co-expression of MYC and BCL2 (double protein expression; DPE), P53+ and/or $CD5^{+}$. Serum samples were collected prior to therapy of which 122 were included in the validation. Patient demograpics are presented in Table S4.

Additional samples, cohorts, and data

In addition, gene expression profiling data[12] from the Nordic LBC-04 trial (registered at www.ClinicalTrials.gov, NCT01502982)[20] and gene expression data by Reddy et al.[21] obtained from the corresponding authors were used.

METHOD DETAILS

Antibody-based proximity extension assay

Serum samples (45 μL) were analyzed for 1463 proteins using antibody-based proximity extension assay (PEA, Olink Proteomics AB, Uppsala, Sweden) with Olink Explore 1536 library (Product #91120). The quantified relative protein levels were analyzed as normalized protein expression (NPX) on a log 2 scale. After excluding technical control samples and proteins with concentrations below limit of detection in $\geq$75% samples according to manufacturer's instructions the data was available for 1,398 proteins in 115 patient samples and five donors with no known history of cancer that were used as controls.

Please cite this article in press as: Arffman et al., Inflammatory and subtype-dependent serum protein signatures predict survival beyond the ctDNA in aggressive B cell lymphomas, Med (2024), https://doi.org/10.1016/j.medj.2024.03.007

Identification of clusters and serum protein associations

We identified the serum clusters using k-means clustering and the optimal number of clusters was determined by manual examination of dimensionality reduced data (principal component analysis) and consensus clustering. We performed consensus clustering by using R package *ConsensusClusterPlus* (version 1.58.0[39]) evaluating maximum of six clusters, with 1000 bootstraps resampling 80% of samples with all the proteins. We used R package limma[40] to identify proteins with significantly different levels between different conditions. Significant associations were determined using a cut-off for significant differential expression was Benjamini-Hochberg corrected p value <0.05 and log2 fold change >1.

Bead-based immunoassay

We applied a bead-based immunoassay to measure absolute levels of selected marker proteins in serum samples from LBCL patients and healthy donors. We used two multiplexed Luminex panels provided by R&D Systems (Human Magnetic Luminex Assay, Minneapolis, MN, USA) according to manufacturer's Human Luminex Discovery Assay and protocol on the Bio-Plex 200 instrument (Hercules, CA, USA). The analytes used for inflammation score were IL10, IL18 and CXCL9, and analytes used in subtype characterization and kinetics assay was CCL17/TARC and TACI/TNFRSF13B. Protein levels were measured in duplicates using 50 µL serum as a starting material and the mean of duplicate measurements was used. Samples with protein concentrations below or over the standard range were assigned the concentration of the lowest or the highest standard dilution per protein, respectively.

Luminex-based measurements including IFNG, which correlated negatively with PEA-based NPX values in technical validation were excluded from further analyses. For the kinetic analysis of selected marker proteins in follow-up serum samples from the patients in the discovery cohort, 47 baseline, 41 mid treatment (after three cycles of therapy), 47 end of therapy and three progression samples were analyzed.

Inflammation score

Inflammation score was composed to evaluate the contribution of inflammation in tissues and serum as a continuous variable with reduced number of markers. Inflammation score was calculated as the geometric mean of marker protein serum concentrations. In PEA data (Olink), we used log2-transformed NPX values of twelve markers (IL2RA, GZMB, IL18, CXCL9, IFNG, PD-L1, IL10, TIM3, LAG3, GZMH, TNFRSF1B and SLAMF8) to compose the inflammation score. In bead-based immunoassay data (Luminex), the inflammation score was estimated with available cytokine markers (IL10, IL18 and CXCL9) and applied to their log10-transforemed absolute concentrations (pg/mL). The optimal cutoff for inflammation score in survival analyses was similar for OS and PFS and determined with R package *survminer* version 4.1.2.[41]

Gene expression profiling

Gene expression datasets. Gene expression profiling data of the diagnostic tumor biopsies from the patients in our discovery cohort (n = 48) was produced with ncounter PanCancer Immune Profiling Panel (NanoString Technologies, Seattle, WA, United States).[12] The gene expression levels of 252 genes encoding the proteins measured with PEA in serum were available. Additionally, gene expression profiling based molecular subtypes determined with DLBCL90 assay (NanoString) were available for 39 patients.[38] Further exploration and validation of the

inflammation score were performed using log2-transformed gene expression data (RNAseq) from 624 patients from Reddy et al.[21]

Gene expression data analysis. Differentially expressed genes were determined using the R package limma[40] with same criteria used in the serum protein analyzes above. Uniform manifold approximation and projection analysis was done with R package *umap* (arXiv 2018 1802.03426). Inflammation score was calculated as the geometric mean of the available gene expression levels of the immune markers *LAG3, GZMH, GZMB, IFNG, CD274, HAVCR2, IL10, IL18* (n = 8, discovery cohort) and *LAG3, GZMH, GZMB, IFNG, CD274, HAVCR2, IL10, IL18, SLAMF8* (n = 9, Reddy et al. data21). Gene expression data from the NanoString and RNAseq datasets were analyzed with UMAP dimensionality reduction and colored according to inflammation score per tumor to explore and visualize patterns of inflammatory component contribution in different subsets of LBCLs. The immune signatures reported by Kotlov et al.[7] were estimated for the discovery cohort using the Nanostring immune panel data of available genes of the functional gene expression signatures composed by the authors. Our gene expression profiling data included 59% of all genes and 79% of inflammatory genes used in the original work. The gene set variation analysis was performed as previously described[7,42] and visualized by hierarchical clustering.

Tissue protein expression analysis

Multiplex immunochemistry (mIHC) was performed for the diagnostic biopsies of 39 patients of the discovery cohort. Briefly, tissue micro-arrays were stained in multiple rounds with antibody panels of immune cell markers and the acquired immunofluorescent figures were analyzed based on pixel co-localization to quantify cell fractions with different immunophenotypic marker combinations.[12,37] Immune cell fractions per tumor sample were analyzed according to the inflamed serum profile and other circulating proteins.[12,37] In total, 17 tissue protein markers which yielded a total of 76 combinations were available for analysis.

Immunohistochemistry for p53 was performed on whole tissue sections and tissue micro-arrays of available diagnostic tissues according to clinical routines and the positivity was evaluated by expert hematopathologists.[43]

Circulating tumor DNA analysis

Pretreatment circulating tumor DNA (ctDNA) from concurrently drawn plasma samples was profiled for 91 patients of the discovery cohort using an in-house targeted high-throughput sequencing-based assay.[11] Quantified ctDNA levels expressed as log10-transformed haploid genome equivalents per milliliter of plasma (log hGE/ml), the somatic mutations in patients and minimal residual disease (MRD) status according to ctDNA were available.[11] In dynamic analyses, MRD status according to ctDNA after three cycles and at the end of therapy was available for 58 and 71 patients, respectively.

The genomic subtypes according to Wright et al.[44] were determined for all the cases with available driver mutations used in the LymphGen 2.0 tool. Identified lymphoma mutations from pretreatment ctDNA and/or diagnostic tumor tissue derived DNA with the targeted sequencing panel were used for the analysis. Copy number alterations were not available for consideration. *BCL2* and *BCL6* translocation status as determined by examination of break-apart fluorescent *in situ* hybridization (FISH) of the diagnostic tissues was utilized for cases with available data. Composite genotypes were assigned to the subgroup with the highest likelihood.

Please cite this article in press as: Arffman et al., Inflammatory and subtype-dependent serum protein signatures predict survival beyond the ctDNA in aggressive B cell lymphomas, Med (2024), https://doi.org/10.1016/j.medj.2024.03.007

QUANTIFICATION AND STATISTICAL ANALYSIS

All analyses were performed in R environment (version ≥3.6.1). The significance of clinical factors between patient groups was determined with Fisher's exact test. Statistical tests were in general non-parametric and two-sided, and statistical details accompany figures in pertinent captions. *p* values <0.05 were considered statistically significant and significance is marked in figures as follows: ns, $p \geq 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$.

Survival analyses

The overall median follow-up time for the discovery and validation cohorts were 62 and 37 months. The Kaplan-Meier method with log-rank test was used to estimate survival rates between different patient groups. Cox regression was used to estimate survival in univariable and multivariable analyses.