Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Matched cerebrospinal fluid and bone marrow single cell transcriptomics highlights the CXCR4-CXCL12 axis in ALL CNS disease

Signe Modvig

signe.modvig.stausboell@regionh.dk

Copenhagen University Hospital Rigshospitalet, Denmark https://orcid.org/0000-0001-9113-1097

Søren Sandgaard

University of Copenhagen

Bodil Als-Nielsen

University Hospital Rigshospitalet

Maria Thastrup

Rigshospitalet, University of Copenhagen

Ulrik Overgaard

Herlev University Hospital

Claus Christensen

Copenhagen University Hospital Rigshospitalet

Eva Ruud Kjær

Copenhagen University Hospital Rigshospitalet

Hanna Thorsson

Lund University

Henrik Lilljebjörn

Lund University https://orcid.org/0000-0001-8703-1173

Thoas Fioretos

Lund University https://orcid.org/0000-0002-3235-6154

Lars Olsen

University of Copenhagen https://orcid.org/0000-0002-6725-7850

Kjeld Schmiegelow

Copenhagen University Hospital https://orcid.org/0000-0002-0829-4993

Hanne Vibeke Marquart

University of Copenhagen https://orcid.org/0000-0001-9740-6522

Article

Keywords:

Posted Date: November 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11460878/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** conflict of interest to disclose.

Abstract

Central nervous system (CNS) relapse in acute lymphoblastic leukemia (ALL) poses a significant clinical challenge. Although experimental studies have highlighted candidate genes for CNS involvement, studies in humans are scarce. We used single-cell transcriptomic profiling of leukemic cells in matched cerebrospinal fluid (CSF), bone marrow and blood of patients with ALL to identify genes and cell states associated with CNS involvement and relapse. CSF leukemic cells showed a distinct expression profile with lipogenic metabolic rewiring and occupied transcriptional states more closely related to bone marrow than to blood. Trajectory analysis identified six genes with progressively increased expression toward CSF-associated leukemic states (SMAP2, CXCR4, SPRY1, PLD1, HSPB1 and ARPP21). Among 16 genes previously implicated in ALL CNS disease in experimental models, CXCR4, SCD, VEGFA and SPP1 were significantly increased in CSF, and CXCR4 was further increased in patients with relapse (log2FC 2.1, p = 0.00096). Supporting this, patients with CNS relapse had higher soluble CXCL12 in CSF (p = 0.044) and higher bone marrow CXCR4 surface expression at diagnosis (p = 0.034) than patients without relapse. In conclusion, single-cell transcriptomics identified molecular alterations associated with ALL CNS disease. The CXCR4–CXCL12 axis stands out as a marker of CNS relapse in ALL and a promising therapeutic target.

Introduction

Acute lymphoblastic leukemia (ALL) is the most common malignancy in childhood. It has a very high cure rate, with more than 90% of patients surviving long-term[1]. In recent years, immunotherapies have contributed to the improved patient outcome[2]. Especially the CD19-directed bispecific T-cell engager Blinatumomab has shown high efficacy and a favorable safety profile[3]. Blinatumomab is therefore being included as frontline therapy in current and upcoming protocols, with the goal to reduce toxicity while preserving relapse prevention [4,5].

Despite recent progress several critical barriers remain, preventing improvements in survival. One of them, central nervous system (CNS) disease, continues to pose a significant clinical challenge in ALL[6]. CNS involvement at time of diagnosis is present in 25–30% of patients, when including low-level leukemic blasts in the CSF detected by flow cytometry[7–9], more often in T-ALL than B-cell precursor (BCP-)ALL[9]. The presence of detectable leukemic cells in the cerebrospinal fluid (CSF) at diagnosis is associated with a higher risk of relapse involving CNS (CNS+ relapse)[7,8,10]. CNS+ relapses are common; they comprise roughly 20% of all relapses[1,10–12]. In high-risk patient groups such as adults and T-ALL, the cumulative 5-year incidence of CNS+ relapse is 6–10% [11,12]. It has recently been shown that blinatumomab does not decrease isolated CNS relapse rates to the same degree as other relapse types[13], which could be related to limited CNS activity and penetrance of the drug[13–15]. Thus, specific CNS-directed therapies remain central to CNS+ relapse prevention and treatment. However, the currently used drugs, such as intrathecal methotrexate and cytarabine, are associated with significant toxicities, prompting the need for less toxic alternatives.

In the attempt to identify new CNS-directed therapeutic targets, the mechanisms for leukemic cell CNS colonization have been intensively studied[16–19], and many candidate proteins have been suggested to play a role for CNS entry and/or persistence in experimental models of ALL[17,20–23]. However, xenograft models do not necessarily provide an accurate reflection of the CNS ALL-microenvironment interplay in humans[24]. Hence, studies of leukemic cells in the CSF of patients are needed.

Due to the scarcity and lability of the leukemic cells in patient-derived CSF, few studies have addressed their molecular characteristics. In this study, we applied single cell flow cytometric sorting and single cell RNA sequencing of leukemic cells in fresh CSF and matched bone marrow (BM) and peripheral blood (PB) sampled at diagnosis and CNS relapse to identify genes and pathways associated with CNS involvement and relapse. We then validated the leading candidate gene, CXCR4 (CD184), the receptor for the chemokine CXCL12, which directs migration and retention of hematopoietic and leukemic cells, in three independent cohorts using orthogonal assays.

Methods

Participants

Seven patients with ALL were included for single cell transcriptomic analyses of matched CSF and BM samples at diagnosis or CNS+ relapse, with PB analyzed in parallel when sufficient numbers of leukemic cells were present (Fig. 1A). Patients were treated according to the NOPHO ALL2008[1] and ALLTogether (NCT04307576, clinicaltrials.gov) protocols. CNS status was determined according to current guidelines[6,25]. Flow cytometric evaluation of leukemic cells in the CSF was performed on all patients as previously described[9,26] .

CXCR4 gene expression in BM leukemic cells in CNS+/- ALL patients was evaluated in single cell transcriptomics analysis of BM samples from time of diagnosis in 27 patients with ALL. Measurement of CSF CXCL12 levels by ELISA analysis was performed on CSF samples from 36 patients with ALL. Of the 36, 10 patients experienced CNS+ relapse (five sampled at diagnosis, five at time of CNS+ relapse), while 26 did not. Surface CXCR4 protein expression by flow cytometry on leukemic cells from BM at diagnosis was examined in 20 patients. Of these, 13 had subsequent relapse (eight CNS+ relapse, five isolated bone marrow relapse (CNS- relapse)). All patient cohorts are shown in Fig. 1A with analytical overlap shown in Fig. 1B.

All analyses were undertaken with the understanding and written consent of each subject including consent of a legal guardian for minors. The study methodologies conformed to the standards set by the Declaration of Helsinki, and the Capital Regional Ethics Committee approved the study (H-2-2010-002).

Sample preparation and single cell flow cytometric sorting

All samples were processed fresh on the day of sampling with CSF sample preparation initiated within 30 minutes of time of lumbar puncture. Samples were analyzed in 8/9 color flow cytometric panels on a FACSARIA III (Becton Dickinson) according to protocol-defined guidelines[26,27]. Leukemic cells were single-cell sorted into 384-well plates along with a small number of mature B- and T-cells. All available leukemic cells in the sample or up to a maximum of two plates per sample were sorted. Cells were index-sorted, allowing the immunophenotype of individual cells to be linked directly to their gene expression profile. Cells were annotated as leukemic cells based on the leukemia-associated immunophenotype during sorting, and all cell annotations were reviewed after sorting, with exclusion of any non-leukemic events. Immediately after sorting, plates were centrifuged, snap-frozen on dry ice, and stored at -80°C.

Single cell RNA sequencing and quality control

Library preparation was performed according to the CEL-seq2 protocol[28], adapted to FACS-sorted cells in 384-well plates[29]. Plates were sequenced on an Illumina NextSeq™ 500 at a read depth of 150,000 reads per cell. Raw reads were preprocessed and aligned to the human reference genome (GRCh38, Ensembl release 98) using STARsolo (STAR v2.7.11b)[30]. Quality control and downstream analysis was performed in R (version 4.5.0) with Seurat (version 5.5.0)[31]. Cells with < 1,200 UMIs, > 6,000 unique features or > 30% mitochondrial content were removed, and raw counts were normalized with SCTransform. Flow cytometry-based annotations were validated against unsupervised transcriptomic clusters. A detailed description of single cell RNA sequencing, preprocessing, quality control, and downstream analysis is available in the Supplementary Methods 1–3.

Cell cycle analysis, dimensionality reduction and compartment separation analysis

Cells were classified as S or G2/M phase using Seurat's CellCycleScoring on log-normalized counts, with cells scoring low for both modules classified as G1 (Supplementary Methods 4). The proliferative fraction for each compartment in each patient was calculated as the proportion of cells in S or G2/M phase. The first 30 principal components (PCs) were used for Uniform Manifold Approximation and Projection (UMAP) with cells colored by patient, tissue compartment and cell cycle phase. Transcriptional separation between compartments was assessed by partial least squares discriminant analysis (PLS-DA) and quantified using centroid distances and silhouette scores (Supplementary Methods 5).

Gene selection based on literature review

Since several proteins have been shown to be associated with CNS invasion in experimental models of ALL but most of these have not or only scarcely been studied in patients, we performed a thorough

litterature review of CNS ALL experimental model studies. Based on this, we selected 16 genes coding for proteins with experimental evidence for a role in CNS invasion and/or persistence (CCR7, CD79a, CD99, CXCR4, ICAM1, IL7R, ITGAL, ITGA4, ITGA6, JUN, LGMN, RAC2, SCD, SPP1, VEGFA, and ZAP70, Table S3 provides an overview of studies).

Differential gene expression and gene set enrichment analysis

The overall gene expression differences between CSF leukemic cells and leukemic cells in the BM and PB compartments, as well as the CXCR4 gene expression in BM between patients with and without CNS disease, were investigated with differential expression analysis using DESeq2 (version 1.50.2)[32] on per-sample pseudobulk aggregated raw counts (Supplementary Methods 6). For the tissue compartment comparisons, P-values were adjusted for multiple testing using the Benjamini-Hochberg method[33] and subsequent gene set enrichment analysis (GSEA) was carried out using fgsea (version 1.36.2)[34]with FDR correction. Within the CSF compartment, after per-sample pseudobulking, the expression levels of the 16 selected genes from the literature review were investigated between three primary and four relapse samples, with leukemia type as the contrast.

Trajectory inference and pseudotime analysis

For trajectory inference, data were integrated by canonical correlation analysis (CCA) and clustered using the Louvain algorithm (Supplementary Methods 7), and trajectory inference was carried out with Slingshot (version 2.18.0)[35], using the cluster with the highest proportion of CSF cells as the starting point (Supplementary Methods 8). The trajectory spanning all three compartments with no dependence on proliferating cells was selected for downstream analysis. Genes differentially expressed along this trajectory were identified, and Spearman correlations were calculated for the SCTransformed expression of all significant genes along the pseudotime axis. The resulting gene sets were used for overrepresentation analysis with clusterProfiler (version 4.18.4)[36] against the Gene Ontology Biological Processes database (Supplementary Methods 9)[37].

Retrospective analysis of CXCR4 expression on leukemic cells in bone marrow

CXCR4 was measured in cryopreserved bone marrow mononuclear cells by 9-color flow cytometry using the protocol-defined ALLTogether panel with the addition of CD184/CXCR4 in BV605 (clone 12G5, Catalogue#740418, BD Optibuild). CXCR4 median fluorescence intensities (MFI) were recorded for both leukemic cells and erythrocytes, which served as an internal negative control, being the only major CXCR4-negative cell population. MFI differences between leukemic cells and erythrocytes were calculated for each patient to account for inter-run variation.

Measurement of soluble CXCL12 levels in CSF

CXCL12 levels in CSF were measured in duplicate by ELISA (human CXCL12 Quantikine ELISA kit, Catalogue#DSA00, R&D Systems). Low, medium and high positive controls (Catalogue#QC20, R&D Systems) were added to each plate to measure inter-assay variation. Only samples with < 20% variation between duplicates were accepted for further analysis.

Results

Single cell transcriptomics was performed on leukemic cells and healthy lymphocytes from 19 samples of matched CSF (n = 7), BM (n = 7) and PB (n = 5) from seven patients with ALL (Fig. 1A), three from time of diagnosis and four from time of CNS+ relapse. Patient and sample characteristics are shown in Table S1.

Leukemic cells differ transcriptionally between the CSF, BM and PB compartments

Leukemic cells from individual patients clearly colocalized in UMAP space, while healthy T- and B-cells clustered together across patients and compartments (Fig. 2A–B), illustrating comparability of data across samples. Of note, leukemic cells from CSF clustered separately from BM and PB leukemic cells in most patients, indicating distinct gene expression profiles for leukemic cells in the CSF (Fig. 2B). Accordingly, centroid distances from PLS-DA showed minimal overlap between the three compartments, with leukemic cells in CSF consistently lying closer to the CSF centroid than to the BM and PB centroids (Fig. 2C) Similarity of leukemic cells across compartments was higher at diagnosis than at CNS relapse, with silhouette scores for leukemic cells in CSF lowest in diagnostic samples and highest at CNS relapse (Fig. 2D). Although substantial variation in proliferative fraction of leukemic cells was seen between patients (Fig. 2E and 2F), the proliferative fraction of CSF leukemic cells fell between those of the other two compartments, with BM leukemic cells having the highest proliferative fraction, and PB leukemic cells the lowest (median 45.8% in BM, 29.2% in CSF, and 25.1% in PB, Fig. 2F).

CXCR4 and VEGFA are upregulated in CSF leukemic cells at diagnosis and at CNS relapse

The CSF leukemic cell gene expression of 16 selected genes with experimental evidence for a role in ALL CNS disease (CCR7, CD79a, CD99, CXCR4, ICAM1, IL7R, ITGAL, ITGA4, ITGA6, JUN, LGMN, RAC2, SCD, SPP1, VEGFA, and ZAP70, Table S1) was compared with expression in combined BM and PB to investigate CSF-specific altered expressions in patients (Table S4). Of the 16, six showed significantly altered expression in CSF leukemic cells (Fig. 3A–B, Table S4). Four genes were increased (SCD: log2FC 2.48, p = 1,1x10e-5; VEGFA: log2FC 1.80, p = 0.011; CXCR4: log2FC 0.83, p = 0.0031; SPP1: log2FC 3.31, p = 0.0059), while two were decreased (CD99: log2FC -0.48, p = 0.025; JUN: log2FC -2.29, p = 7,1x10e-6).

Overall expression levels varied, with the lowest expression seen for SPP1 and VEGFA, while the highest expression was seen for CXCR4 and CD99.

Looking specifically at the expression profiles in CSF leukemic cells, we found that expression of CXCR4 (log2FC 2.13, p = 0.0020) and VEGFA (log2FC 3.94, p = 0.0039) were significantly upregulated in patients sampled at CNS relapse compared with patients sampled at primary diagnosis, while ITGA6 was significantly downregulated (log2FC 2.1, p = 0.017) (Fig. 3C).

CSF leukemic cells display gene expression profiles associated with lipogenic metabolic rewiring

Differential expression analysis of CSF leukemic cells vs pooled BM and PB leukemic cells revealed 42 upregulated and 47 downregulated genes after FDR correction (Table S4), including upregulation of genes associated to biosynthsis of cholesterol (SQLE, SC5D, CYP51A1) and fatty acids (and SCD and FADS1). GSEA demonstrated strong signals for cellular response to starvation and amino acid deficiency, cholesterol biosynthesis, translation initiation, and metabolism of steroids (Fig. 3D).

Trajectory analysis identifies upregulation of genes involved in response to oxidative stress and chemotaxis in CSF leukemic cells

After data integration across patients, leukemic cells from the three compartments separated in UMAP space with some overlap (Fig. 4A). Trajectory analysis showed that CSF leukemic cells occupied transcriptional states more closely related to BM leukemic cells than to PB leukemic cells (Fig. 4B). Although leukemic cells in proliferative states separated clearly in the projection of the integrated data, the trajectory was not influenced by cell cycle states, with all cell cycle states distributed evenly along it (Fig. 4C). The top 80 genes that were significantly altered along the pseudotime axis are shown in Fig. 4D).

We identified six genes that showed progressively increased expression along the inferred trajectory toward CSF-associated leukemic states (SMAP2, CXCR4, SPRY1, PLD1, HSPB1, and ARPP21), indicating stress adaptation and survival (HSPB1 and PLD1), chemotaxis (CXCR4) and slow-cycling states (SPRY1). Five genes were preferentially expressed in both CSF- and bone marrow-associated states relative to PB (CD24, WAKMAR2, IGLL1, RAG1, and MPP1), including markers suggesting immaturity (the pre-BCR gene IGLL, and the V(D)J recombination gene RAG1). Ribosomal and mitochondrial genes dominated the top genes expressed in BM-associated cell states, while a range of genes, including differentiation and maturity markers such as KLF2, TCF7, CD52, and MS4A1 (CD20), were predominantly expressed in PB-associated states.

Overrepresentation analysis of pseudotime-associated genes identified enrichment of pathways related to response to oxygen levels and oxidative stress in CSF-enriched leukemic cellular states (Fig. 4E). Conversely, genes associated with CSF-enriched leukemic cellular states showed reduced enrichment of

pathways including regulation of immune system- and apoptotic processes, lymphocyte activation, and cell motility (Fig. 4F).

CSF levels of the CXCR4-ligand CXCL12 are increased in patients who develop CNS relapse

CXCR4 was consistently found to be implicated across complementary analytical approaches, including pseudotime-associated expression dynamics, prior candidate-gene selection based on experimental CNS leukemia models, upregulation in CSF leukemic cells in pseudobulk differential expression analysis, and further upregulation in CNS+ relapse. We therefore those the CXCR4-CXCL12 axis for orthogonal validation.

We found higher CSF levels of soluble CXCL12 in diagnostic samples from 10 patients who later developed CNS+ relapse compared with 26 patients who did not experience relapse within the follow-up period (Mann-Whitney U test, $p = 0.044$, Fig. 5A).

There was no difference in CSF CXCL12 levels between patients with and without detectable leukemic cells in CSF by flow cytometry among patients without relapse (Fig. 5B). For a subset of the patients ($n = 12$), we investigated temporal dynamics of measured CXCL12 levels during induction therapy but found no trend (data not shown). Curiously, across all patients and timepoints, there was a slight inverse correlation between the number of leukemic cell and CXCL12 level in CSF ($r=-0.27$, $p = 0.032$, Fig. 5C).

CXCR4 surface levels are increased at diagnosis in BM of patients with later CNS+ relapse

Since CSF CXCL12 levels were increased in patients who developed CNS+ relapse but not in patients with CNS involvement at diagnosis, who did not develop relapse, we investigated if the same was the case for BM leukemic cell CXCR4 expression.

Measurement of surface CXCR4 expression by flow cytometry on BM leukemic cells from diagnosis in patients with later CNS+ relapse, CNS- relapse, and without relapse showed a significant difference in expression levels between the three groups ($p = 0.021$, Kruskall-Wallis test, Fig. 5E). Post hoc analyses showed significantly increased CXCR4 expression on BM leukemic cells from patients with later CNS+ relapse compared with patients without relapse in the follow-up period ($p = 0.034$, Dunn's test) and borderline significantly increased compared with patients with later CNS- relapse within the follow-up period ($p = 0.058$, Dunn's test).

The per-sample average CXCR4 gene expression levels in BM leukemic cells from diagnosis of 27 patients with and without CNS involvement, as measured by flow cytometry, but with no CNS+ relapse revealed no significant difference between the two groups, including in separate investigations of B- and T-ALL cases (Fig. 5F).

Discussion

In the present study, we have used single cell transcriptomics of matched CSF, bone marrow, and blood samples from patients with ALL to characterize the transcriptional landscape of leukemic blasts across compartments and stages of ALL CNS disease. This is the first transcriptomics study of CSF leukemic cells at single cell resolution in patients with ALL. We demonstrate direct upregulation of CXCR4 in leukemic cells in the CSF with further upregulation at time of CNS+ relapse. Our two independent corroborating analyses of the CXCR4-CXCL12 axis demonstrated consistent results. Thus, while cohort sizes were limited in this study, the consistency across independent analytical approaches strengthens the hypothesis of a role for the CXCR4-CXCL12 axis in CNS disease establishment and persistence in ALL.

The leukemic cells from the CNS displayed a distinct gene expression profile from bone marrow and peripheral blood. They showed upregulation of gene sets involved in response to starvation, including GCN2/EIF2AK4 response to amino acid deficiency, cholesterol biosynthesis and SREBP/SREBF target gene sets, as well as response to oxygen levels and oxidative stress. Also observed was enrichment of chaperone mediated autophagy and late endosomal microautophagy gene set. These signatures are all consistent with an integrated stress response activation of the leukemic cells in the nutrient-poor CSF microenvironment with a broad lipogenic metabolic rewiring, collectively indicating adaptation to the local milieu. Our findings in primary patient material thus extend the findings in cell lines and experimental models showing that CSF leukemic cells demonstrate metabolic rewiring with increased de novo synthesis of fatty acids and cholesterol[20,38], including the SREBP-target gene SCD[10 20], as well as higher levels of ROS[39].

Data integration and trajectory analysis revealed that bone marrow leukemic cells were closer to CSF leukemic cells in pseudotime than those of peripheral blood. Furthermore, genes associated with immaturity were preferentially expressed in both CSF- and bone marrow-associated cell states relative to PB, while maturity markers were associated with the peripheral blood cell state. Several leukemic cell CNS entry routes have been suggested in recent years[6,16,19,40,41], including a peripheral blood-independent, direct route from calvarian bone marrow along emissary vessels to the CNS, elegantly described by Yao et al[17]. The relative similarity of CSF and bone marrow leukemic cell gene expression profiles as opposed to peripheral blood could support such a direct route hypothesis, however, it could also simply reflect transcriptional adaptation of the CSF-derived leukemic cells to local adherence signals shared between the bone marrow and CNS meningeal niches. We did not find an upregulation of ITGA6 in the CSF associated cell states, a gene which was suggested by Yao et al to play a central role for CNS entry from calvarian bone marrow. However, this does not rule out a role for the integrin, since integrin functions are highly dependent on conformational changes in addition to mRNA/protein expression levels[42] .

We found upregulation of CXCR4 gene expression in CSF leukemic cells, in particular at time of CNS+ relapse, corroborating earlier findings in experimental models. Basile et al showed increased survival of ALL cells co-cultured with meningeal cells in CSF compared to ALL cells in CSF monoculture with CXCR4

inhibition leading to disruption of the leukemia-meningeal adhesion and decreased leukemic cell viability in the CSF[39]. Jonart et al found a similar effect of CXCR4 inhibition on leukemia-meningeal adhesion, as well as reversion of the meningeal-mediated leukemic cytarabine resistance and increased leukemic cell death upon Mavorixafor (AMD070) administration[43]. Jonart et al also demonstrated CXCL12 expression by the meningeal cells. We extend these preclinical studies with clinical evidence of elevated CXCR4 leukemic cell gene expression and elevated soluble CXCL12 levels in the CSF of patients with ALL CNS+ relapse, suggesting that this axis also plays a role in humans.

Our data implicate the CXCR4-CXCL12 axis as an early indicator of CNS+ relapse of ALL rather than simply associated with leukemic cell CNS presence: While CSF soluble CXCL12 protein levels as well as BM leukemic cell CXCR4 surface expression were elevated in patients who experienced CNS+ relapse, the soluble CXCL12 levels and CXCR4 gene expression levels were not increased in patients with CNS involvement at diagnosis but without later CNS relapse. Although the validation cohorts were assembled retrospectively, and residual confounding by baseline characteristics or risk group can therefore not be excluded, this observation is striking and new in the context of ALL. However, it shares interesting similarities with premetastatic niche formations described in solid cancers. Kaplan et al. demonstrated that lung metastasis by melanoma cells is preceded by increased fibronectin production by resident fibroblasts, which is then followed by integrin-dependent adhesion of VEGFR1 + hematopoietic bone marrow myeloid progenitors[44]. In the resulting premetastatic niche comprising VEGFR1 + myeloid progenitors, fibroblasts and fibronection, CXCL12 expression is increased, which leads to the homing of CXCR4 + melanoma cells and metastasis. It could be speculated that CXCL12 has a similar role in ALL and permits the homing of CXCR4 + cancer cells. However, the mechanism may be different, including the steps leading to increased CXCL12 expression. Meningeal cells are known to produce CXCL12 and to attract CXCR4 + cells also in non-leukemic settings, e.g. neuronal cells during embryogenesis[45], and they are also a known source of fibronectin[46]. Moreover, ALL cells in the CSF have previously been shown to upregulate VEGFA[22] and VEGFR1 as a supposed mechanism for hypoxic adaptation[47]. In accordance with this, we found upregulation of VEGFA in CSF leukemic cells as well as overrepresentation of genes involved in response to oxygen levels in CSF associated cell states. Thus, further studies of the role of secreted factors by meningeal cells and their interplay with leukemic cells in the early ALL CNS niche might provide important cues for relapse prevention.

In addition to highlighting its potential as an early indicator of CNS+ relapse, our findings draw attention to the CXCR4-CXCL12 axis as an emerging therapeutic target in ALL CNS disease. In ALL, high CXCR4 expression in BM was found to be associated with higher CIR and worse overall survival in two small cohorts[48,49]. CXCR4-inhibitors have been tested alone or in combination with chemotherapy in a large number of phase I/II trials in AML, and to a limited degree also in ALL[50]. These demonstrated a high degree of safety and tolerability but only a modest effect on bone marrow disease. Importantly, targeting of the CXCR4-CXCL12 axis specifically to prevent ALL CNS disease has not been studied in patients. This poses an intriguing option given the favorable safety profile of CXCR4 antagonists demonstrated in ALL. Although a preclinical model study showed no effect of Plerixafor (AMD3100) on ALL CNS

disease[16], other antagonists, in particular Mavorixafor, would be expected to have considerably higher CNS penetrance[51] and thus higher efficacy towards ALL CNS disease. Moreover, increased leukemic cell death has been observed in CSF when adhesion to meningeal cells is blocked[39], and Mavorixafor has demonstrated additional cytotoxic effects compared with other CXCR4 antagonists in vitro[43], making Mavorixafor a promising therapeutic candidate for CNS relapse prevention.

In conclusion, through transcriptomic characterization of leukemic cells in CSF and matched bone marrow and blood at single cell resolution, we demonstrate a distinct gene expression profile for CNS resident leukemic cells with broad metabolic rewiring and a profile more closely resembling bone marrow leukemic cells than those of peripheral blood. We show upregulation of CXCR4 expression, especially at time of CNS relapse, and demonstrate elevated CSF levels of soluble CXCL12 and CXCR4 surface expression of bone marrow leukemic cells in patients with later CNS+ relapse. These findings extend on the experimental studies suggesting a role for the CXCR4-CXCL12 axis in ALL CNS disease, and although larger studies are needed, they support that the CXCR4-CXCL12 axis could be a promising therapeutic target for prevention of ALL CNS+ relapse.

Declarations

Competing Interests

The authors declare no competing financial interests

Data availability statement

The datasets generated and analyzed during the current study are not publicly available, as deposition was not covered by the patient consent or the regulatory approvals obtained for this study, and the small cohort size could permit re-identification of individual patients. De-identified data may be made available to qualified researchers on reasonable request to the corresponding author, subject to ethics committee approval and a data transfer agreement.

Contributions

SM and HM conceived the work that led to the submission, KS, BAN, UO, MT, SM and HM acquired data, SS, LRO, and SM analyzed the data, CC, ERK, HT, TF, and HL played an important role in interpreting the results, SM and SS drafted the manuscript, all authors revised the manuscript and approved the final version.

References

1. Toft N, Birgens H, Abrahamsson J, Griškevic Ius L, Hallböök H, Heyman M *et al.* Results of NOPHO ALL2008 treatment for patients aged 1–45 years with acute lymphoblastic leukemia. *Leukemia* 2018; 32: 606–615.

2. Inaba H, Pui CH. Immunotherapy in pediatric acute lymphoblastic leukemia. Cancer and Metastasis Reviews. 2019; 38: 595–610.
3. Kantarjian H, Stein A, Gökbuget N, Fielding AK, Schuh AC, Ribera J-M *et al.* Blinatumomab versus Chemotherapy for Advanced Acute Lymphoblastic Leukemia. *New England Journal of Medicine* 2017; 376: 836–847.
4. Gupta S, Rau RE, Kairalla JA, Rabin KR, Wang C, Angiolillo AL *et al.* Blinatumomab in Standard-Risk B-Cell Acute Lymphoblastic Leukemia in Children. *New England Journal of Medicine* 2025; 392: 875–891.
5. Litzow MR, Sun Z, Mattison RJ, Paietta EM, Roberts KG, Zhang Y *et al.* Blinatumomab for MRD-Negative Acute Lymphoblastic Leukemia in Adults. *New England Journal of Medicine* 2024; 391: 320–333.
6. Thastrup M, Duguid A, Mirian C, Schmiegelow K, Halsey C. Central nervous system involvement in childhood acute lymphoblastic leukemia: challenges and solutions. Leukemia. 2022; 36: 2751–2768.
7. Thastrup M, Marquart HV, Schmiegelow K. Flow Cytometric Detection of Malignant Blasts in Cerebrospinal Fluid: A Biomarker of Central Nervous System Involvement in Childhood Acute Lymphoblastic Leukemia. Biomolecules. 2022; 12. doi:10.3390/biom12060813.
8. Perruso LL, Velloso E, Rocha V, Rego EM, Silva WF. Patterns and prognostic impact of CNS infiltration in adults with newly diagnosed acute lymphoblastic leukemia. *Ann Hematol* 2024; 103: 2033–2039.
9. Levinsen M, Marquart H V, Groth-Pedersen L, Abrahamsson J, Albertsen BK, Andersen MK *et al.* Leukemic blasts are present at low levels in spinal fluid in one-third of childhood acute lymphoblastic leukemia cases. *Pediatr Blood Cancer* 2016; 63: 1935–42.
10. van der Velden VHJ, de Launaij D, de Vries JF, de Haas V, Sonneveld E, Voerman JSA *et al.* New cellular markers at diagnosis are associated with isolated central nervous system relapse in paediatric B-cell precursor acute lymphoblastic leukaemia. *Br J Haematol* 2016; 172: 769–781.
11. Liang H, He B, Chen J, Luo B, Cai Z, Ding C *et al.* Paediatric-inspired regimen reduces central nervous system relapse in adult acute lymphoblastic leukaemia. *Br J Haematol* 2025. doi:10.1111/bjh.70067.
12. Cario G, Valsecchi MG, Conter V, Gotti G, Möricke A, Stanulla M *et al.* Results in pediatric T-ALL patients treated in trial AIEOP-BFM ALL 2009: Prognostic factors in the context of modern risk-adapted therapy. *Hemasphere* 2025; 9. doi:10.1002/hem3.70206.
13. Hogan LE, Brown PA, Ji L, Xu X, Devidas M, Bhatla T *et al.* Children's Oncology Group AALL1331: Phase III Trial of Blinatumomab in Children, Adolescents, and Young Adults with Low-Risk B-Cell ALL in First Relapse. *Journal of Clinical Oncology* 2023; 41: 4118–4129.
14. Ngo D, Otoukesh S, Stein AS, Forman S, Pullarkat V, Aldoss I. The safety of concurrent intrathecal chemotherapy during blinatumomab in adults with acute lymphoblastic leukemia. Leuk. Lymphoma. 2022; 63: 2754–2756.

15. Ngo D, Tinajero J, Song JY, Ma H, Quirk E, Koller P *et al.* The characterization and the impact of CSF pleocytosis during blinatumomab therapy for adult acute lymphoblastic leukemia. *Leuk Lymphoma* 2024; 65: 2016–2024.
16. Williams MT, Yousafzai YM, Elder A, Rehe K, Bomken S, Frishman-Levy L *et al.* The ability to cross the blood-cerebrospinal fluid barrier is a generic property of acute lymphoblastic leukemia blasts. *Blood* 2016; 127: 1998–2006.
17. Yao H, Price TT, Cantelli G, Ngo B, Warner MJ, Olivere L *et al.* Leukaemia hijacks a neural mechanism to invade the central nervous system. *Nature* 2018. doi:10.1038/s41586-018-0342-5.
18. Rajakumar SA, Grandal I, Minden MD, Hitzler JK, Guidos CJ, Danska JS. Targeted blockade of immune mechanisms inhibit B precursor acute lymphoblastic leukemia cell invasion of the central nervous system. *Cell Rep Med* 2021; 2. doi:10.1016/j.xcrm.2021.100470.
19. Modvig S, Jeyakumar J, Marquart HV, Christensen C. Integrins and the Metastasis-like Dissemination of Acute Lymphoblastic Leukemia to the Central Nervous System. Cancers (Basel). 2023; 15. doi:10.3390/cancers15092504.
20. Savino AM, Fernandes SI, Olivares O, Zemlyansky A, Cousins A, Markert EK *et al.* Metabolic adaptation of acute lymphoblastic leukemia to the central nervous system microenvironment depends on stearoyl-CoA desaturase. *Nat Cancer* 2020; 1: 998–1009.
21. Buonamici S, Trimarchi T, Ruocco MG, Reavie L, Cathelin S, Mar BG *et al.* CCR7 signalling as an essential regulator of CNS infiltration in T-cell leukaemia. *Nature* 2009; 459: 1000–1004.
22. Münch V, Trentin L, Herzig J, Demir S, Seyfried F, Kraus JM *et al.* Central nervous system involvement in acute lymphoblastic leukemia is mediated by vascular endothelial growth factor. *Blood* 2017; 130: 643–654.
23. Jost TR, Borga C, Radaelli E, Romagnani A, Perruzza L, Omodho L *et al.* Role of CXCR4-mediated bone marrow colonization in CNS infiltration by T cell acute lymphoblastic leukemia. *J Leukoc Biol* 2016; 99: 1077–1087.
24. Kohnken R, Porcu P, Mishra A. Overview of the use of murine models in leukemia and lymphoma research. *Front Oncol* 2017; 7. doi:10.3389/fonc.2017.00022.
25. Mahmoud HH, Rivera GK, Hancock ML, Krance RA, Kun LE, Behm FG *et al.* mahmoud_NEJM_1993_original_CNS123. *New England Journal of Medicine* 1993. doi:https://doi.org/10.1056/nejm199307293290504.
26. Thastrup M, Marquart HV, Levinsen M, Grell K, Abrahamsson J, Albertsen BK *et al.* Flow cytometric detection of leukemic blasts in cerebrospinal fluid predicts risk of relapse in childhood acute lymphoblastic leukemia: a Nordic Society of Pediatric Hematology and Oncology study. *Leukemia* 2020; 34: 336–346.
27. Modvig S Hallböök H, Siitonen S, Osnes LTN, Rosthøj S, Tierens A, Juvonen V, Vålerhaugen H, Hultdin M, Ehinger M, Porwit A, Matuzeviciene R, Stoskus M, Marincevic M, Lilleorg A, Taskinen M, Toft N, Jónsson OG, Pruunsild K, Vaitkeviciene G, Vettenranta K, MHO. Value of flow cytometry for MRD-

based relapse prediction in B-cell precursor acute lymphoblastic leukemia in a multi-center setting. Abstract submitted to ASH annual meeting 2019. 2019.

28. Hashimshony T, Senderovich N, Avital G, Klochendler A, de Leeuw Y, Anavy L *et al.* CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. *Genome Biol* 2016; 17: 77.
29. Muraro MJ, Dharmadhikari G, Grun D, Groen N, Dielen T, Jansen E *et al.* A Single-Cell Transcriptome Atlas of the Human Pancreas. *Cell Syst* 2016; 3: 385–394 e3.
30. Kaminow B, Yunusov D, Dobin A. STARsolo: accurate, fast and versatile mapping/quantification of single-cell and single-nucleus RNA-seq data. Preprint. 2021. doi:10.1101/2021.05.05.442755.
31. Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A *et al.* Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol* 2024; 42: 293–304.
32. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014; 15. doi:10.1186/s13059-014-0550-8.
33. Yoav Benjamini YH. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc B* 1995; 57: 289–300.
34. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. Fast gene set enrichment analysis. 2016. doi:10.1101/060012.
35. Street K, Risso D, Fletcher RB, Das D, Ngai J, Yosef N *et al.* Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* 2018; 19. doi:10.1186/s12864-018-4772-0.
36. Xu S, Hu E, Cai Y, Xie Z, Luo X, Zhan L *et al.* Using clusterProfiler to characterize multiomics data. *Nat Protoc* 2024; 19: 3292–3320.
37. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM *et al.* Gene ontology: Tool for the unification of biology. Nat. Genet. 2000; 25: 25–29.
38. Cousins A, Olivares O, Markert E, Manoharan A, Bubnova X, Bresolin S *et al.* Central nervous system involvement in childhood acute lymphoblastic leukemia is linked to upregulation of cholesterol biosynthetic pathways. *Leukemia* 2022; 36: 2903–2907.
39. Basile P, Jonart LM, Ebadi M, Johnson K, Kerfeld M, Gordon PM. The meninges enhance leukaemia survival in cerebral spinal fluid. *Br J Haematol* 2020; 189: 513–517.
40. Jost TR, Borga C, Radaelli E, Romagnani A, Perruzza L, Omodho L *et al.* Role of CXCR4-mediated bone marrow colonization in CNS infiltration by T cell acute lymphoblastic leukemia. *J Leukoc Biol* 2016; 99: 1077–1087.
41. März M, Meyer S, Erb U, Georgikou C, Horstmann MA, Hetjens S *et al.* Pediatric acute lymphoblastic leukemia—Conquering the CNS across the choroid plexus. *Leuk Res* 2018; 71: 47–54.
42. Scharff BFSS, Modvig S, Marquart HV, Christensen C. Integrin-Mediated Adhesion and Chemoresistance of Acute Lymphoblastic Leukemia Cells Residing in the Bone Marrow or the Central Nervous System. Front. Oncol. 2020; 10. doi:10.3389/fonc.2020.00775.
43. Jonart LM, Ostergaard J, Brooks A, Fitzpatrick G, Chen L, Gordon PM. CXCR4 antagonists disrupt leukaemia-meningeal cell adhesion and attenuate chemoresistance. *Br J Haematol* 2023; 201: 459–

469.

44. Kaplan RN, Riba RD, Zacharoulis S, Bramley AH, Vincent L, Costa C *et al.* VEGFR1-positive haematopoietic bone marrow progenitors initiate the pre-metastatic niche. *Nature* 2005; 438: 820–827.
45. Borrell V, Marín O. Meninges control tangential migration of hem-derived Cajal-Retzius cells via CXCL12/CXCR4 signaling. *Nat Neurosci* 2006; 9: 1284–1293.
46. DeSisto J, O'Rourke R, Jones HE, Pawlikowski B, Malek AD, Bonney S *et al.* Single-Cell Transcriptomic Analyses of the Developing Meninges Reveal Meningeal Fibroblast Diversity and Function. *Dev Cell* 2020; 54: 43–59.e4.
47. Kato I, Nishinaka Y, Nakamura M, Akarca AU, Niwa A, Ozawa H *et al.* Letters to Blood Hypoxic adaptation of leukemic cells infiltrating the CNS affords a therapeutic strategy targeting VEGFA. http://ashpublications.org/blood/article-pdf/129/23/3126/1400733/blood721712.pdf.
48. Ko SY, Park CJ, Park SH, Cho YU, Jang S, Seo E ju *et al.* High CXCR4 and low VLA-4 expression predicts poor survival in adults with acute lymphoblastic leukemia. *Leuk Res* 2014; 38: 65–70.
49. van den Berk LCJ, van der Veer A, Willemse ME, Theeuwes MJGA, Luijendijk MW, Tong WH *et al.* Disturbed CXCR4/CXCL12 axis in paediatric precursor B-cell acute lymphoblastic leukaemia. *Br J Haematol* 2014; 166: 240–249.
50. Cancilla D, Rettig MP, DiPersio JF. Targeting CXCR4 in AML and ALL. Front. Oncol. 2020; 10. doi:10.3389/fonc.2020.01672.
51. B. Fernandes T, C. F. Segretti M, C. Polli M, Parise-Filho R. Analysis of the Applicability and Use of Lipinski`s Rule for Central Nervous System Drugs. *Lett Drug Des Discov* 2016; 13: 999–1006.

Figures

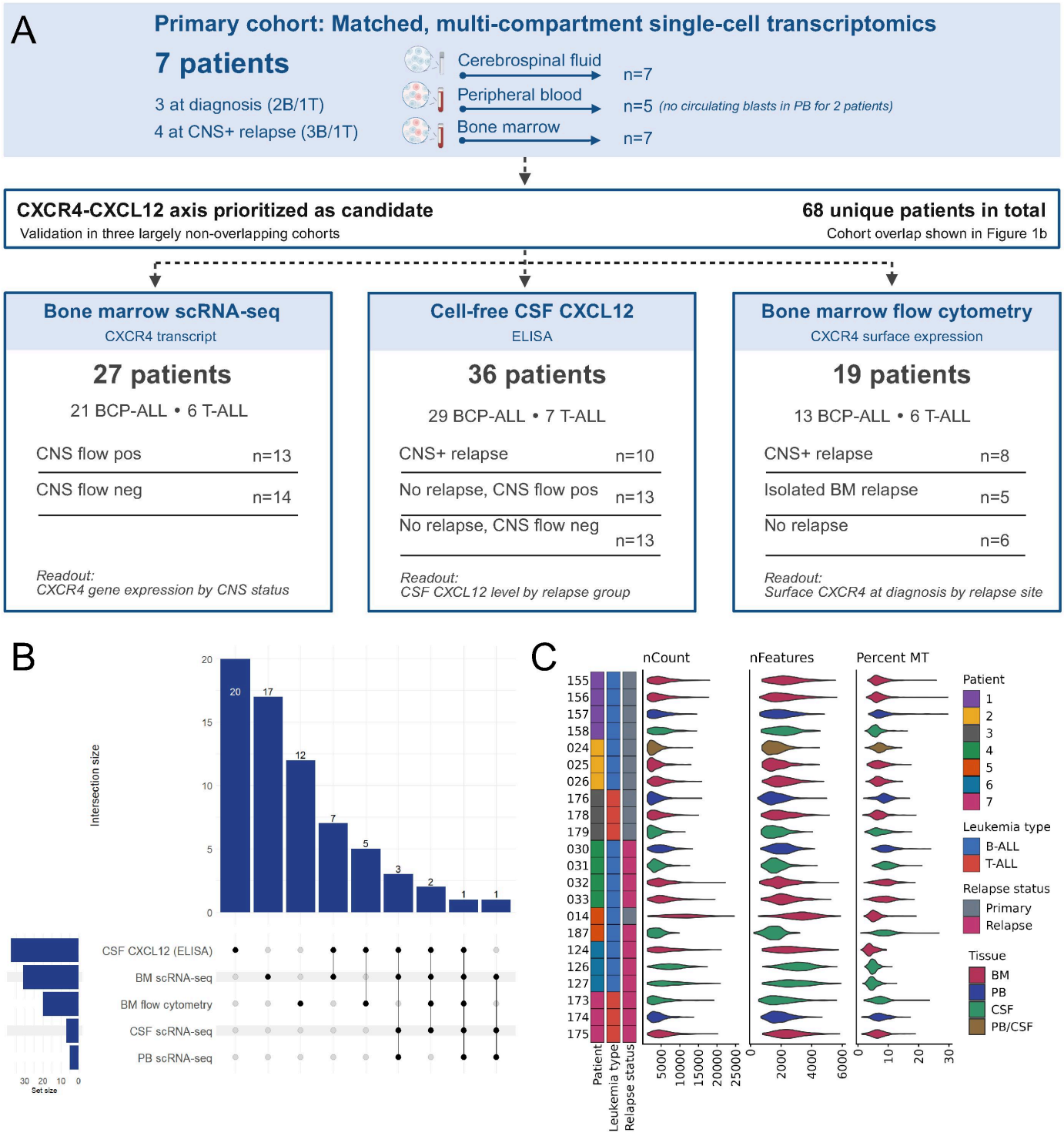


Figure 1

Cohort and sample overview. A. A discovery cohort underwent single cell RNA sequencing (scRNAseq) of matched cerebrospinal fluid, bone marrow and peripheral blood. Three additional cohorts were included for evaluation of CXCR4 gene expression in BM scRNAseq data, soluble CXCL12 quantification in CSF, and CXCR4 surface expression on bone marrow leukemic cells. Three patients from the primary cohort, all sampled at diagnosis, are also included in the bone marrow scRNA-seq cohort; the remaining

four were sampled at CNS relapse and are not. The CXCL12 ELISA and flow cytometry cohorts assess different analytes and compartments, so shared patients contribute independent measurements. B Upset plot showing overlap of analyses on patient level. C Overview of sample quality after filtering for the 22 sequenced plates from 19 samples in the matched, multicompartment single cell transcriptomics cohort. Number of transcripts (nCount), number of unique genes (nFeatures), and percentage of mitochondrial genes (percent MT) are shown for each plate, colored by patient, leukemia type, sample timepoint (relapse status) and tissue. Labels represent internal laboratory plate identifiers. BM: Bone marrow; CSF: cerebrospinal fluid; PB: Peripheral blood.

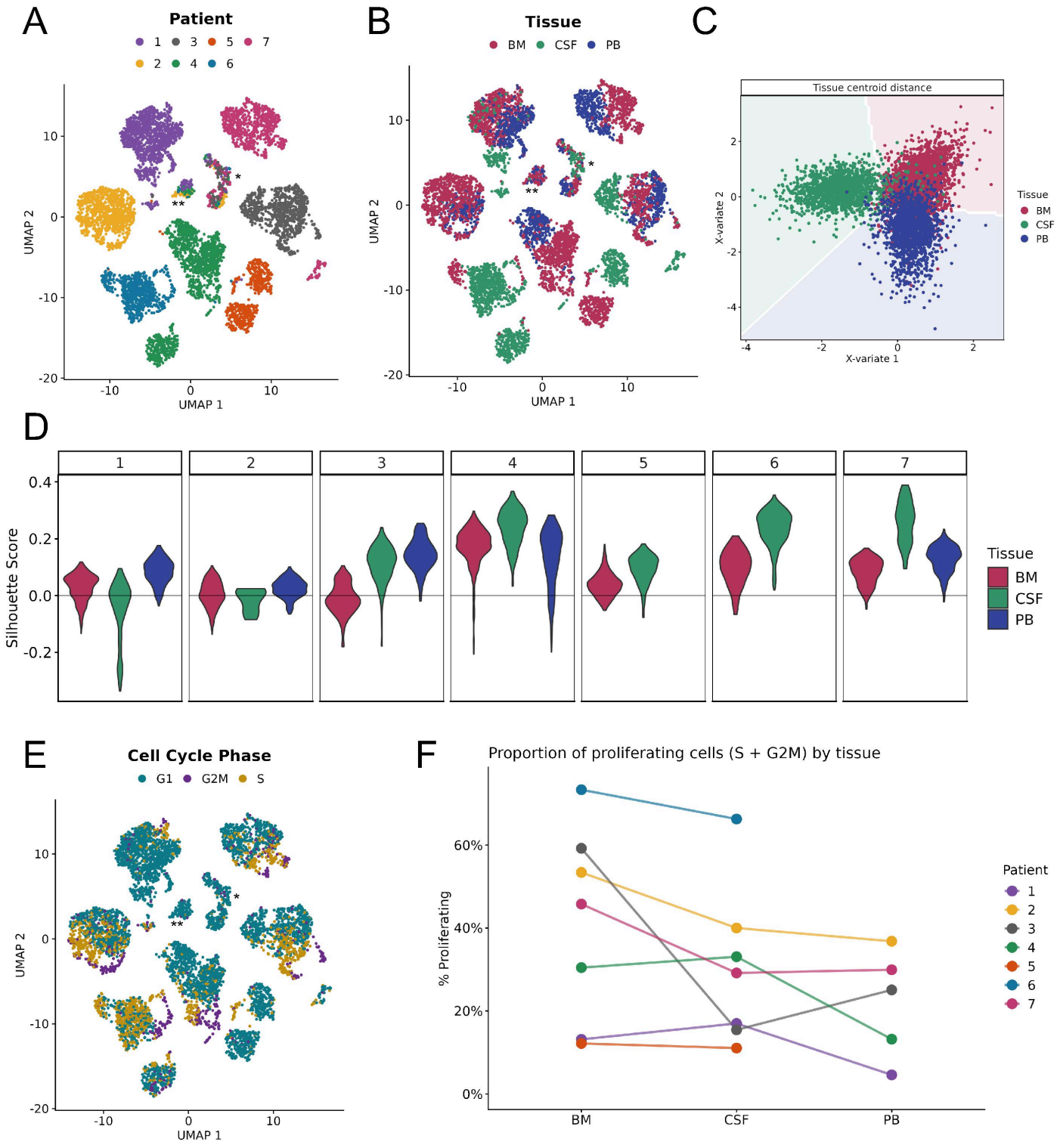


Figure 2

Influence of tissue on cell cycle phase and centroid distance. A-B. Uniform Manifold Approximation and Projection (UMAP) showing clustering of cells according to patient (A), and tissue (B). * = mature T-cells, ** = mature B-cells. C. Tissue centroid distance plot using Partial Least Squares Discriminant Analysis on the first 30 principal components in the PCA with tissue as response variable showing separation of leukemic cells from CSF (green) from bone marrow (red) and peripheral blood (blue) with very little

overlap. White lines illustrate boundaries for areas associated with tissue-specific centroids. D. Silhouette plot illustrating the normalized difference between the between-tissue separation and the within-tissue cohesion. Patients 1-3 are sampled at time of diagnosis, patients 4, 6, and 7 at time of relapse. Patient 5 CSF is sampled at time of isolated CNS relapse, bone marrow at diagnosis with no CNS involvement. E. Uniform Manifold Approximation and Projection (UMAP) showing clustering of cells according to cell cycle phase. F. Proportion of cells in S/G2M phase relative to cells in G1 phase in CSF, bone marrow and peripheral blood.

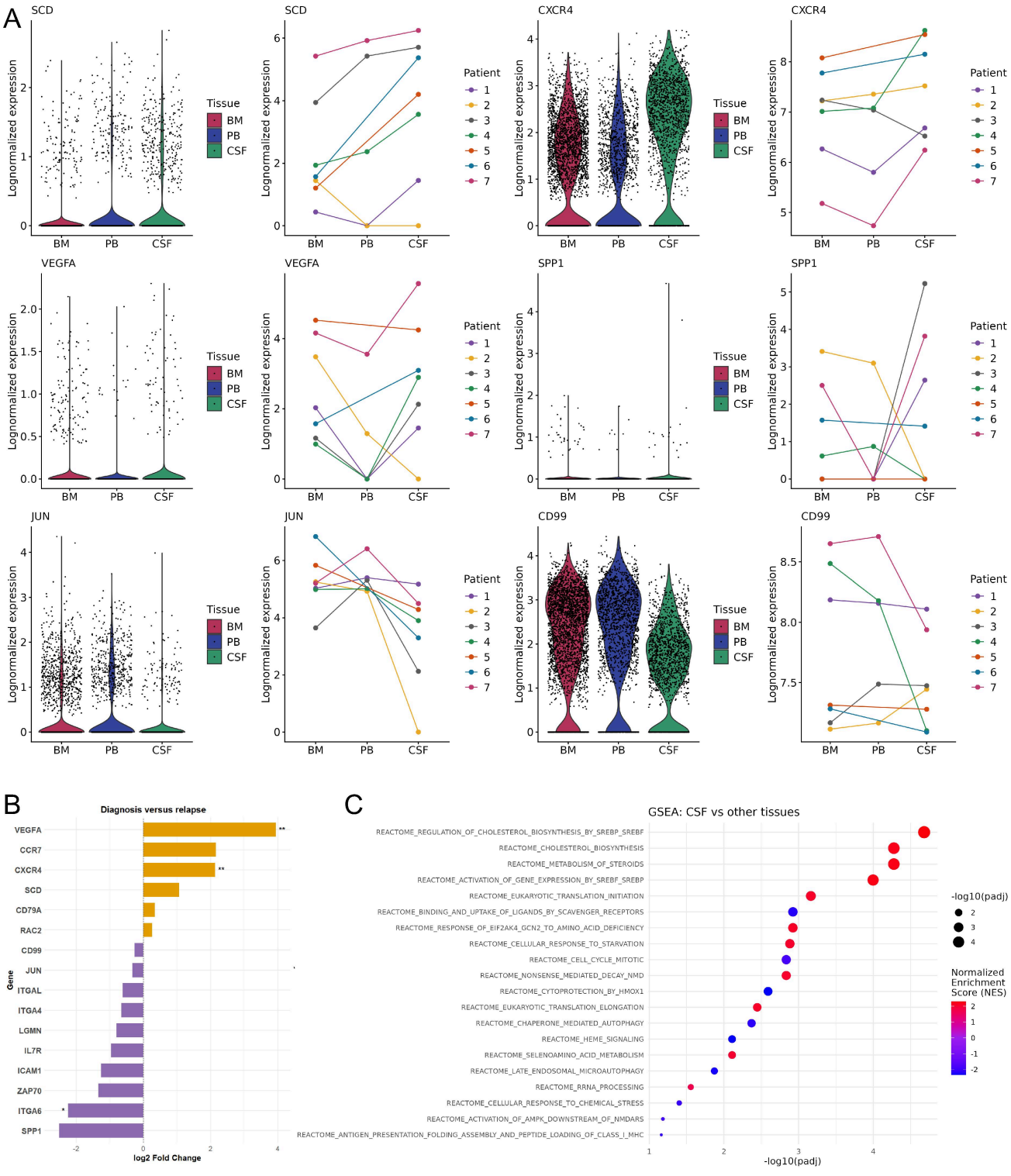


Figure 3

Selected genes and top pathways significantly altered in CSF. A-B. Based on a review of experimental model studies, 16 genes were identified for investigation. Of these, four were found significantly upregulated (A) and two downregulated (B) in leukemic cells of the CSF compared with those of combined bone marrow (BM) and peripheral blood (PB) by differential gene expression (DGE) analysis after sample-level pseudobulked aggregation. Violin plots show lognormalized gene expression for all cells, while spaghetti plots show lognormalized expression on pseudobulk sample level. C. Gene expression of the selected 16 genes in CSF leukemic cells of patients sampled at diagnosis compared with (unmatched) patients at time of relapse by DGE of pseudobulk samples. D. Gene set enrichment analysis showing pathways enriched in leukemic cells of CSF compared with BM/PB by DGE of pseudobulk samples.

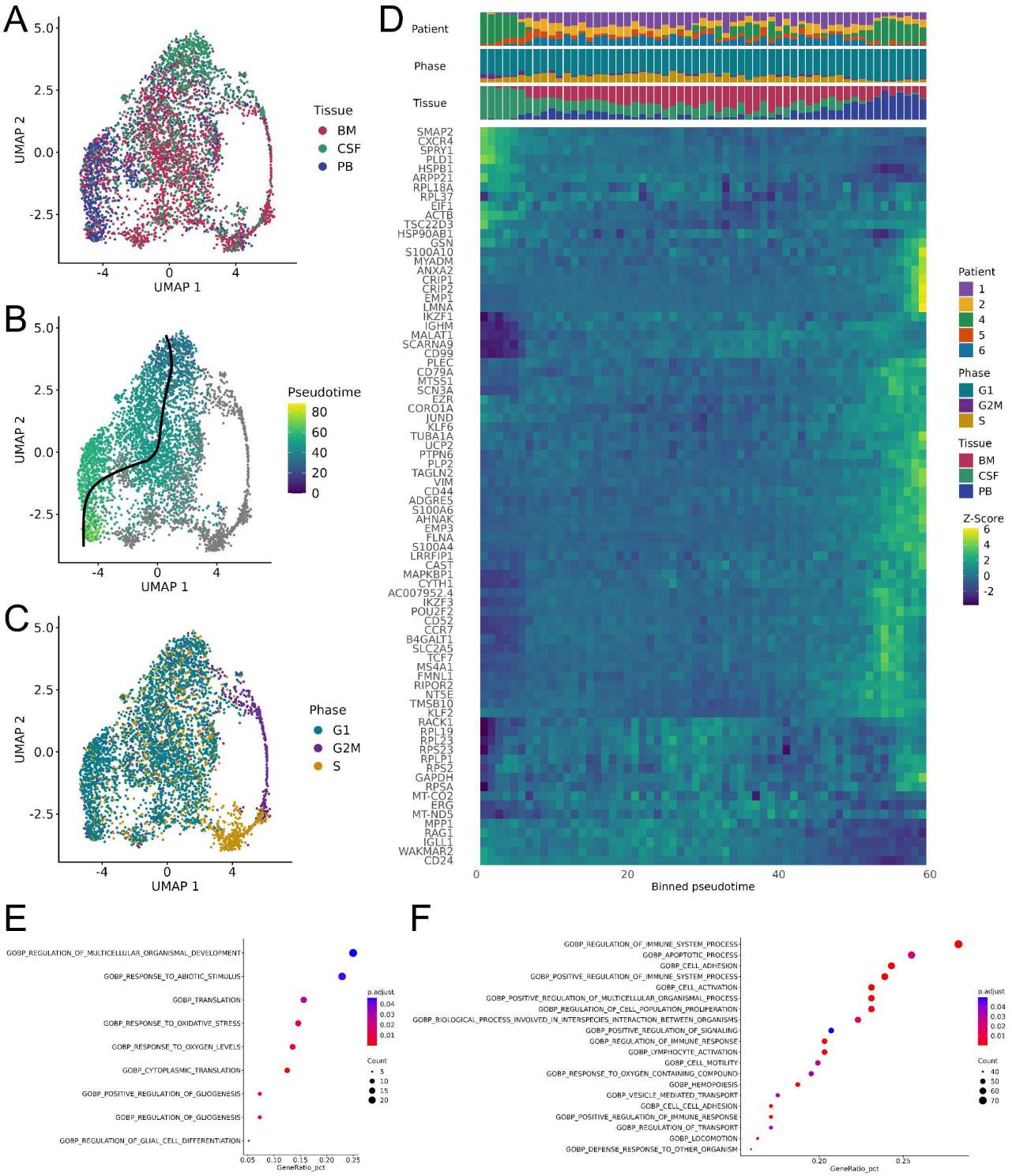


Figure 4

Trajectory analysis across tissues. A-C. Uniform Manifold Approximation and Projection (UMAP) showing integrated data from all samples of matched CSF, bone marrow (BM) and peripheral blood (PB) from the seven patients, colored by tissue (A), pseudotime (B), and cell cycle phase (C). The black line in panel B represents the trajectory spanning all three tissues with no dependence on proliferating cells. D. Heatmap illustrating the gene expression along a binned pseudotime axis for the top 80 genes

associated with pseudotime along the trajectory shown in panel B. E-F. Overrepresentation analysis based on Gene Ontology Biological Process of the genes significantly altered along the pseudotime axis with a negative (r<-0.1) (E) or positive (r>0.1) (F) Spearman correlation. A negative correlation corresponds to genes upregulated towards a CSF-associated cell-state, while a positive correlation corresponds to a peripheral blood-associated cell state.

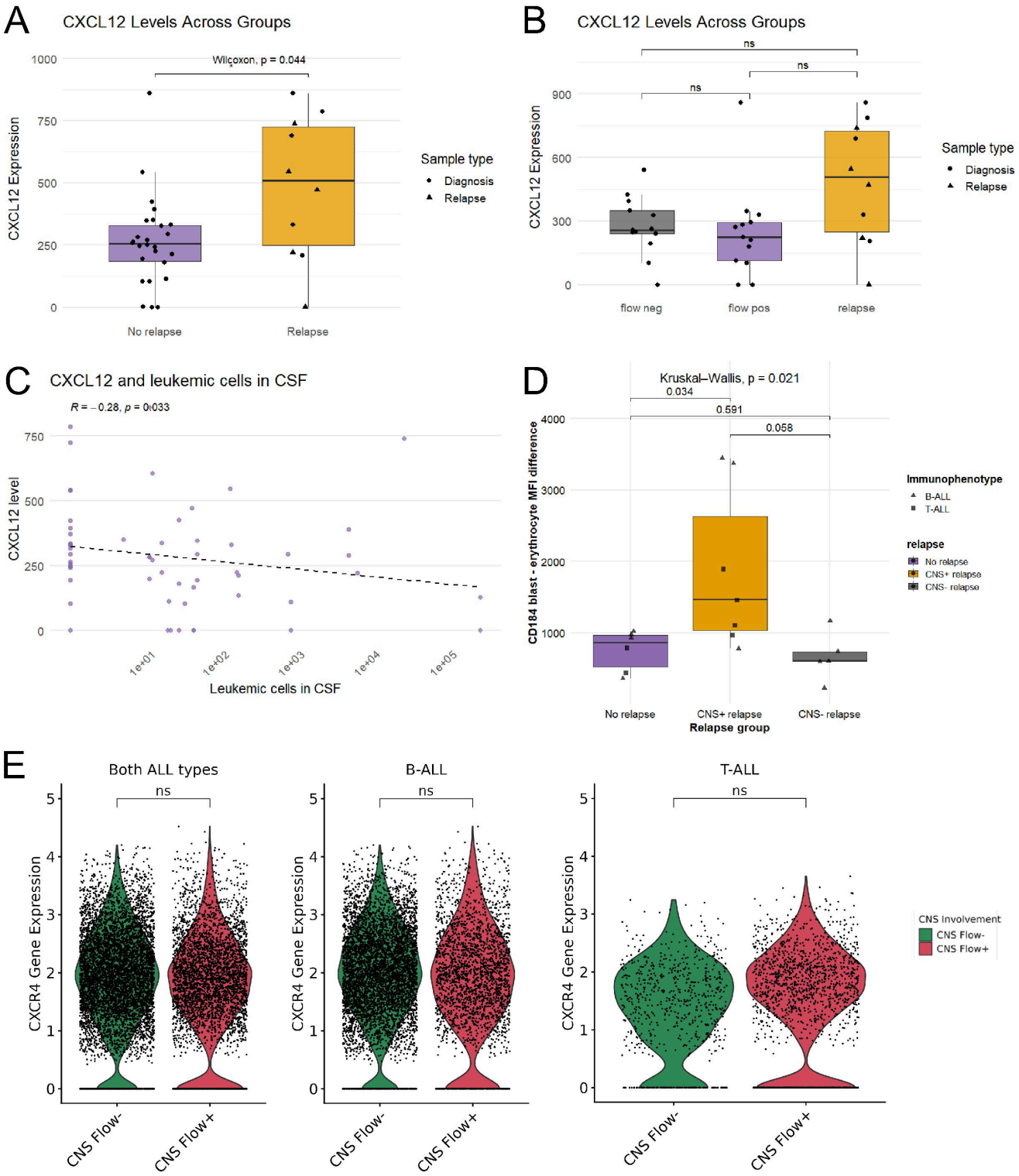


Figure 5

The CXCR4-CXCL12 axis in patients with and without CNS relapse. A. Soluble protein levels of CXCL12 in CSF of patients with and without CNS relapse. Patients with CNS relapse were sampled either at diagnosis or at time of relapse. B. Same data as A, but the patients without CNS relapse are divided according to whether they had detectable leukemic cells in the CSF at diagnosis as measured by flow cytometry (flow pos) or not (flow neg). C. CSF soluble CXCL12 protein levels across all samples and timepoints versus leukemic cells in the CSF measured by flow cytometry. Spearman correlation analysis shown in the top (R = Spearman's Rho). D. CXCR4 (CD184) surface marker expression measured by flow cytometry. The leukemic cell-erythrocyte median fluorescence intensity difference is displayed for patients with relapse involving the CNS (CNS+), isolated bone marrow relapse (CNS-) and patients without relapse. A three-group Kruskall Wallis test p-value is shown in the top, while pairwise comparison p-values are derived from Dunn's test for post hoc comparison. E. CXCR4 gene expression in BM leukemic cells of 27 flow pos or flow neg patients, as well as subdivided by immunophenotype (B-ALL (n=21, 13 flow neg vs. 9 flow pos) and T-ALL (n=6, 2 flow neg vs. 4 flow pos)).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarymaterial.pdf
- TableS4.xlsx