Article

Spatial and single cell mapping of castleman disease reveals key stromal cell types and cytokine pathways

David Smith [1], Anna Eichinger [2,3], Éanna Fennell [4], Zijun Y. Xu-Monette [5], Andrew Rech[6], Julia Wang [6], Eduardo Esteva [2], Arta Seyedian [7], Xiaoxu Yang [1], Mei Zhang[1], Dan Martinez[4], Kai Tan [1,7], Minjie Luo [4], Katherine J. Young [5], Paul G. Murray [4], Christopher Park [2], Boris Reizis [2] & Vinodh Pillai [6] ✉

To determine the cellular and molecular basis of Castleman Disease (CD), we analyze the spatial proteome and transcriptome from a discovery ($n = 9$ cases) and validation ($n = 13$ cases) cohort of Unicentric CD, idiopathic Multicentric CD, HHV8-associated MCD, and reactive lymph nodes. CD shows increased stromal cells that form unique microenvironments. Interaction of activated follicular dendritic cell (FDC) cytoplasmic meshworks with mantle-zone B cells is associated with B-cell activation and differentiation. CXCL13+ FDCs, PDGFRA + T-zone reticular cells (TRC), and ACTA2-positive perivascular reticular cells (PRC) were the predominant source of increased VEGF expression and IL-6 signaling. MCD is characterized by increased TRC while UCD shows increased B-reticular cells (BRC). VEGF expression by FDCs is associated with peri-follicular neovascularization. FDC, TRC and PRC of CD activates JAK-STAT, TGFβ, and MAPK pathways via specific ligand-receptor interactions. Here, we show that stromal-cell activation and associated B cell activation and differentiation, neovascularization and stromal remodeling underlie CD.

Castleman disease (CD) encompasses a group of disorders characterized by abnormal lymph node morphology[1–3]. Unicentric hyaline vascular CD (UCD) manifests as the enlargement of one or more lymph nodes in a single anatomical region[4], exhibits increased hyalinization and vascularity[5] and is typically managed through surgical resection[6]. A subset of UCD cases exhibit plasmacytosis, mediastinal involvement, or MCD-like inflammatory features, making diagnosis and treatment challenging[7]. Very low-level somatic mutational burden has been noted in a subset of UCD[8], but the significance and cell of origin in CD is unclear.

Multicentric CD (MCD) is marked by the enlargement of multiple lymph node regions throughout the body and is accompanied by laboratory abnormalities and systemic symptoms[9]. MCD exhibits a variable lymph node morphology, featuring abnormal germinal centers, increased interfollicular vascularity, and plasma cells. TAFRO (Thrombocytopenia, Anasarca, Fever, Reticulin fibrosis, and Organomegaly) is a recently identified variant of multicentric Castleman disease (MCD)[10]. Unlike other forms of idiopathic MCD (iMCD), TAFRO is distinguished by thrombocytopenia and hypogammaglobulinemia.

[1]Center for Single Cell Biology, Children's Hospital of Philadelphia Research Institute, Philadelphia, PA, USA. [2]Department of Pathology, New York University Grossman School of Medicine, New York City, NY, USA. [3]Department of Pediatrics, Dr. von Hauner Children's Hospital, University Hospital, Ludwig-Maximilians-Universität München, Munich, Germany. [4]School of Medicine, Bernal Institute, Limerick Digital Cancer Research Centre & Health Research Institute, University of Limerick, Limerick, Ireland. [5]Duke University Medical Center, Durham, NC, USA. [6]Department of Pathology and Laboratory Medicine, The Children's Hospital of Philadelphia and the University of Pennsylvania, Philadelphia, PA, USA. [7]Department of Pediatrics, The Children's Hospital of Philadelphia, Philadelphia, PA, USA. ✉e-mail: Pillaiv1@chop.edu

Histologically, TAFRO exhibits a mixed pattern with prominent stromal changes and minimal plasmacytosis. MCD is characterized by elevated levels of serum IL-6[11], VEGF[12], IL-1β, and CXCL13[13]. A subset of MCD cases are associated with HHV-8 infection[14] that drives the pathogenesis through viral IL-6[15]. In contrast, the etiology of HHV8-negative MCD (iMCD) is not known. IL-6 targeted therapy[16] and MTOR inhibitors[17] are used in the management of iMCD with variable efficacy.

Germline genetic aberrations have been sporadically reported in a subset of iMCD cases[8,18,19], but the cells initiating disease and contribution to pathogenesis is not clear. Bulk transcriptomics of lymphoid tissue from CD demonstrated that IL-6, VEGF, complement and vascular pathways were upregulated[20,21]. Prior immunohistochemical analysis suggested that VEGF and IL-6 expression might originate from lymph node cells[12,22], but conclusive identification of the exact cell type and their role in the pathogenesis of CD was lacking.

In this work, we identify the cells and pathways involved in UCD and MCD by utilizing single-cell spatial proteomic and transcriptomic approaches in combination with DNA sequencing and copy number analysis. We demonstrate that VEGF and IL6 produced by lymph node stromal cells such as FDC, TRC, and PRC are associated with specific clinicopathologic features of CD.

Results

Characteristics of Cases and Controls

The discovery cohort consisted of four MCD, three UCD and two RLN cases (Table 1). The lymph node resections were performed at first presentation of lymphadenopathy. RLN1 with reactive follicular hyperplasia and RLN2 with reactive interfollicular plasmacytosis were controls for the histological features of UCD and MCD lymph nodes. RLN and UCD cases did not show systemic symptoms or laboratory abnormalities except for UCD3. MCD cases comprised three HHV8-negative iMCD (MCD1,2 3), and one HHV8-positive MCD in a patient with HIV (MCD4). MCD4 developed in the setting of HIV infection. iMCD patients presented with fever, anasarca, renal dysfunction, anemia, thrombocytopenia and multicentric lymphadenopathy.
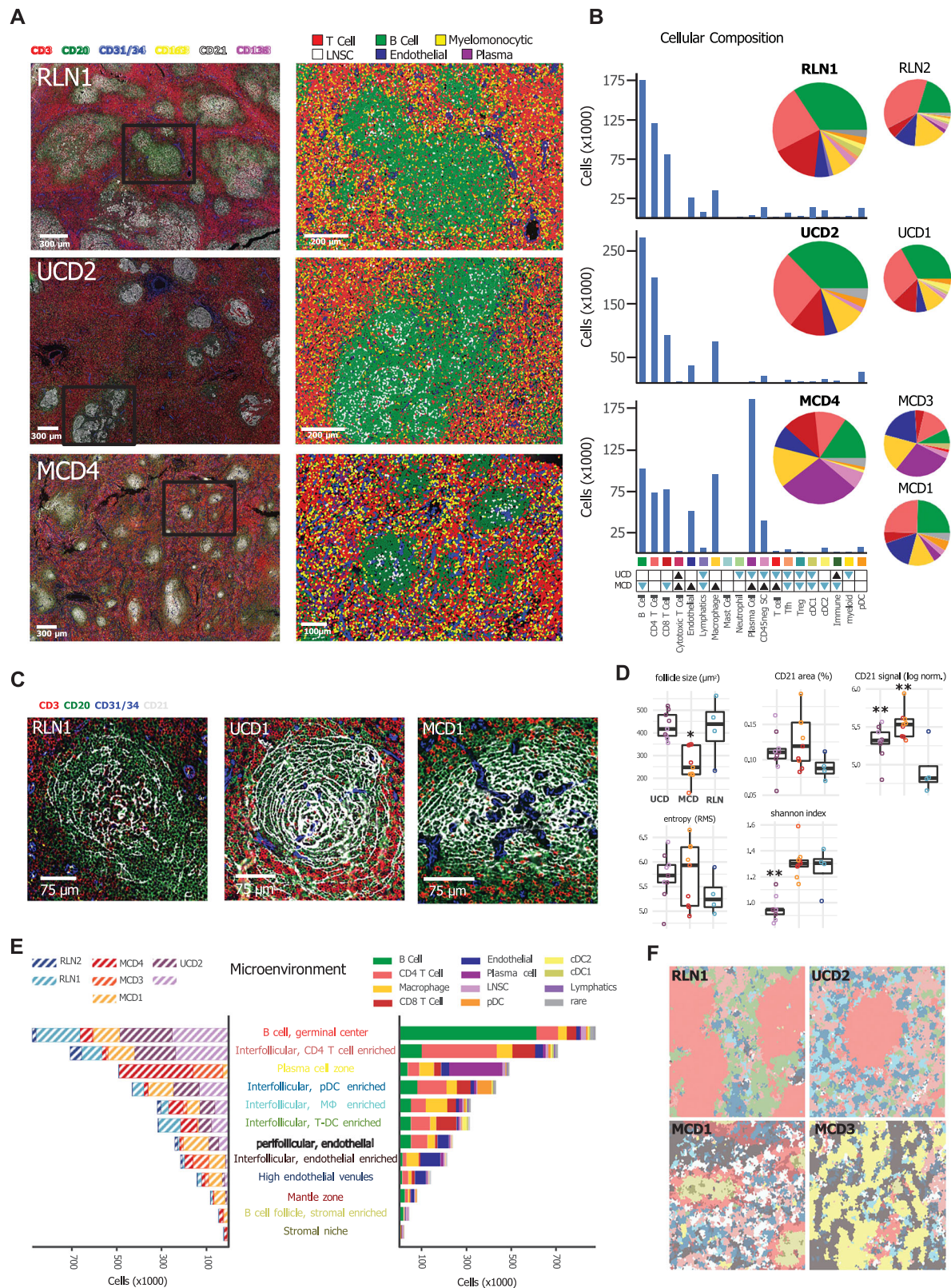
Cellular composition and spatial distribution

The immunophenotype and spatial distribution of 645,285 single cells in RLN, 1,768,327 cells in UCD, and 2,071,397 cells in MCD lymph nodes were analyzed (Fig. 1). Cases and controls showed variably sized CD20 + B-cell follicles separated by interfollicular areas with CD3 + T cells, CD11b+ myelomonocytic cells, CD138+ plasma cells, CD31/34+ endothelial cells, and CD45-negative DAPI+ stromal cells (Fig. 1A). RLN showed predominantly lymphoid cells (B and T cells) with only minor proportions of the non-lymphoid cell types. Based on permutation tests of MCD vs. RLN, there were significantly ($q < 0.01$ and $|log2FC| > 0.53$) increased endothelial cells, plasma cells, macrophages, cytotoxic T cells, and stromal cells but decreased B cells compared to RLN (Fig. 1B and Supplemental Fig. 4). Both UCD and MCD showed decrease in Tfh, Treg, and cDC. Strikingly, non-lymphoid cells comprised more than 25% of lymph node cellularity in MCD1 and the majority of lymph node cellularity in MCD3. CD45-negative DAPI-positive stromal nuclei, consistent with follicular dendritic cells (FDCs), were more abundant in the germinal centers of UCD and MCD compared to RLN. However, enumeration of nuclei did not capture the extent and complexity of cytoplasmic projections emanating from FDC nuclei. Cytoplasmic meshworks of FDC play a crucial role in affinity maturation of B cells by presenting native antigen to B cells on complement receptors CD21 (CR2) and CD35 (CR1) (Fig. 1C). Hence, we focused on CD21 to visualize the extent of FDC meshworks. Interdigitation of FDC meshworks between concentric layers of B cells led to close interactions and was the basis of the well-described 'onion skin' appearance of CD mantle zones. Image-analysis techniques were used to quantify the extent and organization of CD21 meshworks.

Table 1 | Clinicopathologic Characteristics of Cases and Controls in the Discovery Cohort

Sample ID	Lymph node location	Age/Sex	Presentation	Diagnosis	Outcome	Treatment	Single nuclei transcriptomes	Single cell proteomes	Single nuclei immune repertoires
RLN1	Inguinal	4/M	Enlarged node	Reactive follicular hyperplasia	No recurrence	Resection	3785	527690	
RLN2	Hilar	17/M	Enlarged node	Reactive with plasma cells	No recurrence	Resection	2771	117595	119
UCD1	Paraspinal	14/F	Enlarged node (3.8 cm)	HVCD	No recurrence	Resection	4621	848405	
UCD2	Supraclavicular	15/F	Enlarged node	HVCD	No recurrence	Resection	6990	919922	116
UCD3	Mediastinum	15/F	Enlarged node (6 cm), anemia, increased ESR, CRP	HVCD with Plasma cells	No recurrence	Resection	7669		
MCD1	Axilla	13/M	Thrombocytopenia, fever, ascites, anemia, elevated ESR/CRP, elevated creatinine, splenomegaly, proteinuria and hematuria, hypertension	iMCD-TAFRO	CR	Sirolimus, IVIG, rituximab, siltuximab	4192	944503	
MCD2	Inguinal	14/M	Anemia, thrombocytopenia	iMCD- Mixed pattern- TAFRO	CR	Siltuximab, prednisone	1894		
MCD3	Cervical	15/F	Anemia, thrombocytopenia, anasarca, renal failure, rash	TAFRO syndrome	CR	prednisone, siltuximab, sirolimus, tocilizumab, eculizumab,	5246	359982	1963
MCD4	Axilla	15/F	HIV	HHV-8 associated MCD	LFU	Rituximab, steroids, anti-retrovirals	8333	770912	3264

Although the follicles of CD were smaller or similar in size to RLN, the area, brightness, and image entropy (measure of complexity) of CD21 signal was higher in UCD and MCD (Fig. 1D). UCD exhibited a lower Shannon diversity index of cells in follicles, consistent with the observed predominance of FDC nuclei. The findings suggested that aberrant activation and proliferation of FDC and other stromal cells may underlie CD.

Microenvironments and cellular interactions of CD

Next, we assessed the differences in the microenvironment between cases and controls. Twelve microenvironments were identified from the combined spatial scans of RLN, UCD, and MCD regions (Fig. 1E, F). B-cell germinal centers and interfollicular microenvironments enriched in CD4 T cells, macrophages or DC were noted, consistent with known lymph node structures. Many microenvironments were

Fig. 1 | Single-cell proteomic analysis of RLN, UCD and MCD lymph nodes. A The left column displays representative full scans of RLN1 ($n = 3$), UCD2 ($n = 5$), and MCD4 ($n = 4$), showcasing the expression of key lineage markers. In the right column, cell segmentations and annotations for the indicated germinal center region (black box) are provided. B-cells (green) form follicles, while T cells (red) and myeloid cells (yellow) populate the interfollicular areas. UCD and MCD exhibit increased intrafollicular stromal cell proliferation (white), with MCD additionally showing interfollicular endothelial cell (blue) and plasma cell (magenta) proliferation. RLN, reactive lymph node; UCD, unicentric Castleman disease; MCD, multicentric Castleman disease; CD45negSC, CD45-negative DAPI-positive stromal cells. **B** Bar plots illustrate the absolute abundance of cells for the representative region shown in **A**. Pie charts (inset) represent the relative abundance of cell types across all imaged regions for each sample. Differential abundance against RLN was tested using a permutation test. Significantly ($q < 0.01$ and $|\log2FC| > 0.53$) increased proportions are indicated with black triangles and significantly decreased proportions are indicated with blue triangles. **C** Increased CD21 + FDC meshworks (white) in the germinal center regions of UCD and MCD, representative of UCD

enriched in CD cases. An interfollicular macrophage-enriched microenvironment was noted in all CD suggestive of increased macrophage activity. Two endothelial-predominant microenvironments (perifollicular and interfollicular) and a plasma-cell region were identified in all MCD. UCD1/2, and MCD1 exhibited unique perifollicular-endothelial and plasmacytoid dendritic cell (pDC) regions, while MCD3 and MCD4 showed a predominance of interfollicular-endothelial and plasma cell microenvironments. MCD1 showed a unique mantle microenvironment and an stromal cell-rich B-cell follicle was present in MCD1 and MCD4. The microenvironments comprised multiple other cell types that likely interacted with each other akin to the functional interactions that result in normal lymph node structures. We hypothesized that interactions between locally proximal cells may drive the pathogenesis of CD. The distribution of distances between different cell types in each sample was calculated and tested against their RLN counterparts using a Kolmogorov-Smirnov test (Supplemental Table 2). Cell-cell pairings with an average distance <50μm were considered biologically relevant. The cell-cell interactions in UCD and MCD that were most different from RLN were those of plasma cells, macrophages, and endothelial cells with other cell types. These findings suggest that the distinctive microenvironments wherein non-lymphoid cells engage with lymphoid cells play a role in the pathogenesis of CD.

Functional analysis of expanded cell populations in CD

Single cell proteomics provided accurate enumeration of major cell types and microenvironments of CD but provided limited information on cell function. Hence, we performed single-nuclei RNA-seq of concurrent frozen tissue to characterize cell states and signaling pathways. Single-nuclei transcriptomes for 6,556 cells from RLN, 19,280 from UCD, and 24,281 from MCD lymph nodes were analyzed (Fig. 2). Distinct populations of B-cell, plasma-cell, T-cell, myelomonocytic, and stromal lineages were identified (Fig. 2A). RLN showed abundant germinal center B-cell nuclei that were decreased in UCD and MCD. Based on permutation tests against RLN, UCD showed significantly ($q < 0.01$ and $|\log2FC| > 0.53$) increased pDC, plasma cells, and naïve B cells. MCD showed increased nuclei of plasma cells, proliferating plasma cells (plasmablasts), endothelial cells, lymphatics, fibroblastic reticular cells (FRC), monocytes and macrophages, and cytotoxic T cells. Although non-lymphoid populations were underrepresented in transcriptomic data (likely due to dissociation limitations), the relative composition of expanded cell nuclei closely resembled the cell types observed in proteomic data.

We next sought to define functional differences of each cell type between cases and controls. We performed differential gene expression and pathway enrichment analyses of UCD and MCD cell types compared to those in RLN (Fig. 2B). In UCD and MCD, FDCs exhibited

($n = 5$), and MCD ($n = 4$). **D** Analysis of size, area, signal intensity, image entropy of CD21, and cell diversity (Shannon index) within each follicle reveals higher expression and organization of FDC meshworks across CD, and low cell diversity in UCD. Each plotted point represents the aggregate data from a single image region (MCD, $n = 11$; HVCD, $n = 9$, RLN = 4). Box plot bounds represent the interquartile range (p25-p75) with the horizontal bar representing the median. Whiskers extend to the full range of data. Significance versus RLN1 was calculated using Tukey's Honest Significance Difference test. *, $p < 0.05$; **, $p < 0.01$. **E** Twelve identified microenvironments are depicted. Sample composition (left graph) and cellular composition (right graph) of each microenvironment are shown. B cell germinal center and interfollicular CD4- and T-DC-enriched microenvironments are present in all samples. Plasma cell, macrophage, endothelial and stromal enriched microenvironments are increased in CD samples. **F** Voronoi diagrams of microenvironments in representative regions of cases and controls are shown. Color scheme as indicated by shading of microenvironment names in 1E. Source data are provided as a Source Data file in the supplementary dataset.

an upregulation of pathways linked to cytoplasmic projections and angiogenesis. Stromal populations displayed an elevated expression of MAPK, angiogenesis, and extracellular matrix-associated pathways. Notably, FRC of MCD showed a significant ($p < 0.05$) increase in VEGF, angiogenesis, MAPK, and JAK-STAT pathways. Macrophages were enriched in inflammatory pathways, while memory T cells were enriched in pathways associated with CCL19/21 binding, B-cell activation, and proliferation. IL-6 pathways were upregulated in activated B cells. Plasma cells were enriched in pathways indicative of increased antibody production. Collectively, these findings indicate that CD is marked by the involvement of various cell types and processes, including increased angiogenesis and extracellular matrix remodeling by stromal populations, inflammation by macrophages, T-cell activation, and IL-6-mediated differentiation of B cells into antibody-producing plasma cells. While these processes are part of normal immune responses, they appear to be hyperactive in CD.

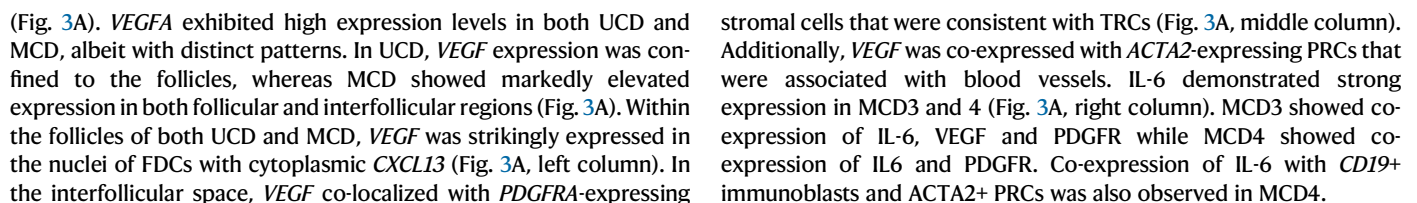## Stromal cells and origin of key cytokines

Since key pathways were enriched in stromal cell populations of CD, we analyzed them in greater detail (Fig. 2C–E). Single-nuclei transcriptomes from all non-hematopoietic cell types in all samples were extracted and annotated using stromal-specific markers (Fig. 2C). We identified distinct groups such as *CDH5 + ENG+* blood endothelial cells (BECs) and *PROX1+* lymphatic endothelial cells (LECs), which were particularly prevalent in MCD samples. We also noted a predominance of *ACTA2+* perivascular reticular cells (PRCs) and *PDGFRA/B+ CCL21*high *CCL19*low *CXCL12 +* T-zone reticular cells (TRCs) in MCD. In contrast, *CXCL13 +* FDC and *FDCSP + CLU +* B-zone reticular cells (BRCs) were mainly found in UCD samples (Fig. 2D). The findings suggest that distinct populations of stromal subsets are involved in the pathogenesis of UCD and MCD.

Considering the elevated levels of circulating VEGF and IL-6 typically seen in CD, we compared their expression across different samples and cell types (Fig. 2E). The highest expression of *VEGFA* and IL-6-associated genes (IL-6 module) were observed in stromal populations. Notably, high levels of *VEGFA* were uniquely seen in PRCs and TRCs across all MCD samples (Fig. 2E, top row). *VEGFA* was also expressed in FDCs across various CD samples, specifically UCD1, MCD1, and MCD3. Similarly, high expression levels of IL-6 associated genes were observed in PRC, CRC of MCD, and FDCs of UCD (Fig. 2E, bottom row). These results showed that stromal cells were the primary source of key cytokines in CD.

Spatial localization of VEGF and IL-6 expressing stromal cells

A sensitive multiplex nucleic acid in-situ hybridization (ISH) was used to co-localize key cytokines and stromal cells. RNAscope assays were performed for *VEGFA, IL-6, CD19, CXCL12, CXCL13, PDGFRA,* and *ACTA2*

(Fig. 3A). *VEGFA* exhibited high expression levels in both UCD and MCD, albeit with distinct patterns. In UCD, *VEGF* expression was confined to the follicles, whereas MCD showed markedly elevated expression in both follicular and interfollicular regions (Fig. 3A). Within the follicles of both UCD and MCD, *VEGF* was strikingly expressed in the nuclei of FDCs with cytoplasmic *CXCL13* (Fig. 3A, left column). In the interfollicular space, *VEGF* co-localized with *PDGFRA*-expressing stromal cells that were consistent with TRCs (Fig. 3A, middle column). Additionally, *VEGF* was co-expressed with *ACTA2*-expressing PRCs that were associated with blood vessels. IL-6 demonstrated strong expression in MCD3 and 4 (Fig. 3A, right column). MCD3 showed co-expression of IL-6, VEGF and PDGFR while MCD4 showed co-expression of IL6 and PDGFR. Co-expression of IL-6 with *CD19*+ immunoblasts and ACTA2+ PRCs was also observed in MCD4.

Fig. 2 | Comparative Analysis of Single Nuclei RNA-Sequencing and Pathway Enrichment in Cell Types of UCD and MCD versus RLN. A A UMAP plot of 50,117 single nuclei displays identified cell types and their relative abundance per sample, along with global abundance. MCD exhibits significantly increased proliferating plasma cells, cytotoxic and memory T cells, monocyte-macrophages, endothelial-lymphatics, and fibroblastic reticular cells (FRC). UCD demonstrates significantly increased plasmacytoid dendritic cells, plasma cells, and naive B cells. Differential abundance against RLN was tested using a permutation test. Significantly increased proportions are indicated with black triangles, and significantly decreased proportions are indicated with blue triangles. **B** The differentially expressed genes (DEGs) of UCD and MCD compared to RLN were tested for pathway enrichment using various databases, including Reactome, GO, KEGG, and TRANSFAC. The top significantly enriched pathways for MCD and UCD are displayed. *P*-value estimates were calculated from a cumulative hypergeometric test[40]. LNSC showed enrichment in pathways related to the extracellular matrix (ECM), cytoplasmic projections, actin cytoskeleton, angiogenesis, MAPK signaling, and VEGF signaling. Macrophages are enriched in inflammatory pathways, whereas plasma cells are enriched in pathways associated with antibody production. **C** Stromal subsets were annotated by expression of key lineage-defining markers. FDCs are characterized by strong expression of *CXCL13, CLU,* and *FDSP*. PRCs show high expression levels of *PDGFRB, CCL21, CCL19, ACTA2, TAGLN,* and *CXCL12*. TRCs show high *CCL21* and *CXCL12* expression. BECs show high *CDH5*, ENG, and *PECAM1*, while LECs show high *PROX1*. **D.** UMAPs of merged stromal subsets from cases and controls. TRCs, PRCs, BECs, and LECs are predominantly from MCD. BRCs and FDCs are predominantly from UCD. **E** SCT normalized IL-6 module score (*IL-6, OSMR, IL-6ST, LIFR*) expression is displayed in the top row while *VEGFA* categorized by stromal cell and sample type is shown in the bottom row. CRCs, PRCs, and FDCs of CD show high IL-6 module and VEGF expression. Gene expression significance versus RLN was calculated using Wilcoxon signed-rank test, and module score significance versus RLN was calculated using Tukey's Honest Significance Difference. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$. Source data are provided as a Source Data file in the supplementary dataset.

Integration of transcriptomic and proteomic data

Multiple stromal subsets were identified through SnRNA-seq, but accurate distinction on CODEX data was challenging due to the lack of specific surface markers. To address this, we integrated transcriptomics and proteomics data using MaxFuse[23] to determine the spatial localization of stromal subsets. This integration allowed us to identify the spatial distribution of key stromal subsets in and around germinal center follicles (Fig. 3B). FDC nuclei were accurately localized to the center of CD21-positive meshworks and were markedly increased in UCD and MCD. BRC were localized within follicles of RLN and UCD. PRC were detected in the blood vessel walls and were distinct from BEC and LEC.

Since key cells and pathways of CD were associated with germinal center follicles, we performed a detailed analysis of their cellular composition. Follicles identified by cell-microenvironment analysis were separated into concentric 100-pixel contours (Fig. 3C, left column). Stromal subtypes identified from snRNA-seq data were transferred to multiplexed image data and their abundance assessed at each contour layer. B cells and CD4 + T cells were increased in follicular and perifollicular areas, respectively. BEC showed a peak in the perifollicular region that was adjacent to peaks of stromal cells in MCD. The spatial findings are consistent with tropism of endothelial cells towards VEGF secreted by stromal cells and formed the basis of the diagnostic feature of penetrating blood vessels ('lollipop') in CD.

Ligand-receptor interactions of stromal cells

Ligand-receptor interactions that drive UCD and MCD (Fig. 4 and Supplemental Fig. 6) were identified from snRNA-seq. LIANA[24] in combination with Tensor-cell2cell[25] identified several factors that distinguish disease types (Fig. 4A). UCD was characterized by factors representing stromal cell interactions that activate the JAK-STAT, TGFβ, and MAPK pathways. The ligand-receptor pairs identified in UCD could be categorized as collagen-integrin extra cellular matrix (ECM) interactions, complement-mediated inflammatory interactions, and VEGF interactions between PRC and macrophages (Fig. 4B and Supplemental Fig. 5). Interestingly, UCD1 and 2 clustered and shared some factors with MCD1 and 2 while UCD3 clustered and shared factors with MCD3 and 4.

MCD was enriched in factors that reflect the interactions of stromal cells, endothelial cells, macrophages, CD8 T cells, and Tfh, resulting in the activation of JAK-STAT, TGFβ, MAPK, and TNF pathways. The ligand-receptor pairs identified in MCD2 were interactions of TRCs with plasma cells through collagens and SDC1/CD138 (Fig. 4C and Supplemental Fig. 5). TRCs showed interactions of VEGF with its receptors on BEC. MCD4 exhibited similar interactions of PRC with BECs and LECs. These findings reveal key ligand-receptor interactions that could be targeted to alleviate MCD symptoms.

Spatial transcriptomics analysis of a validation cohort of CD

We performed 10x Xenium spatial transcriptomics analysis in a validation cohort of 13 CD cases and controls. The cohort consisted of 9 MCD (5 iMCD, 3 iMCD-TAFRO, and one HHV8 MCD), 2 UCD, and 2 reactive lymph node with Castleman-like features (RLN-CL) that ranged from 34-78 years in age and were distributed across sex and racial categories. UCD and MCD showed similar expansion of stromal cells and cell type compositions (Fig. 5A, B). Stromal subtype analysis revealed expansion of BEC in both UCD and MCD. TRC were relatively expanded in MCD while BRC were expanded in UCD (Fig. 5C), pattern that was also observed in SnSeq-data. We also observed a consistent IL-6 module and VEGF showed with high signals in TRC, FDC and PRC. We then validated predicted ligand-receptor interactions in the spatial transcriptomics data. For example, we identified "Factor 2" as a characteristic of UCD and MCD previously, in which LNSCs were senders and T cell populations were receivers. Each TRC-memory/effector T cell pair was scored based on the product of their respective DLL4 and NOTCH1 expression (Fig. 5E). This analysis confirmed the presence of genuine DLL4-NOTCH1 interactions, with a significant enrichment observed in the MCD mantle zone. The additional data provide confirmation of our findings in a validation cohort using a different methodology.
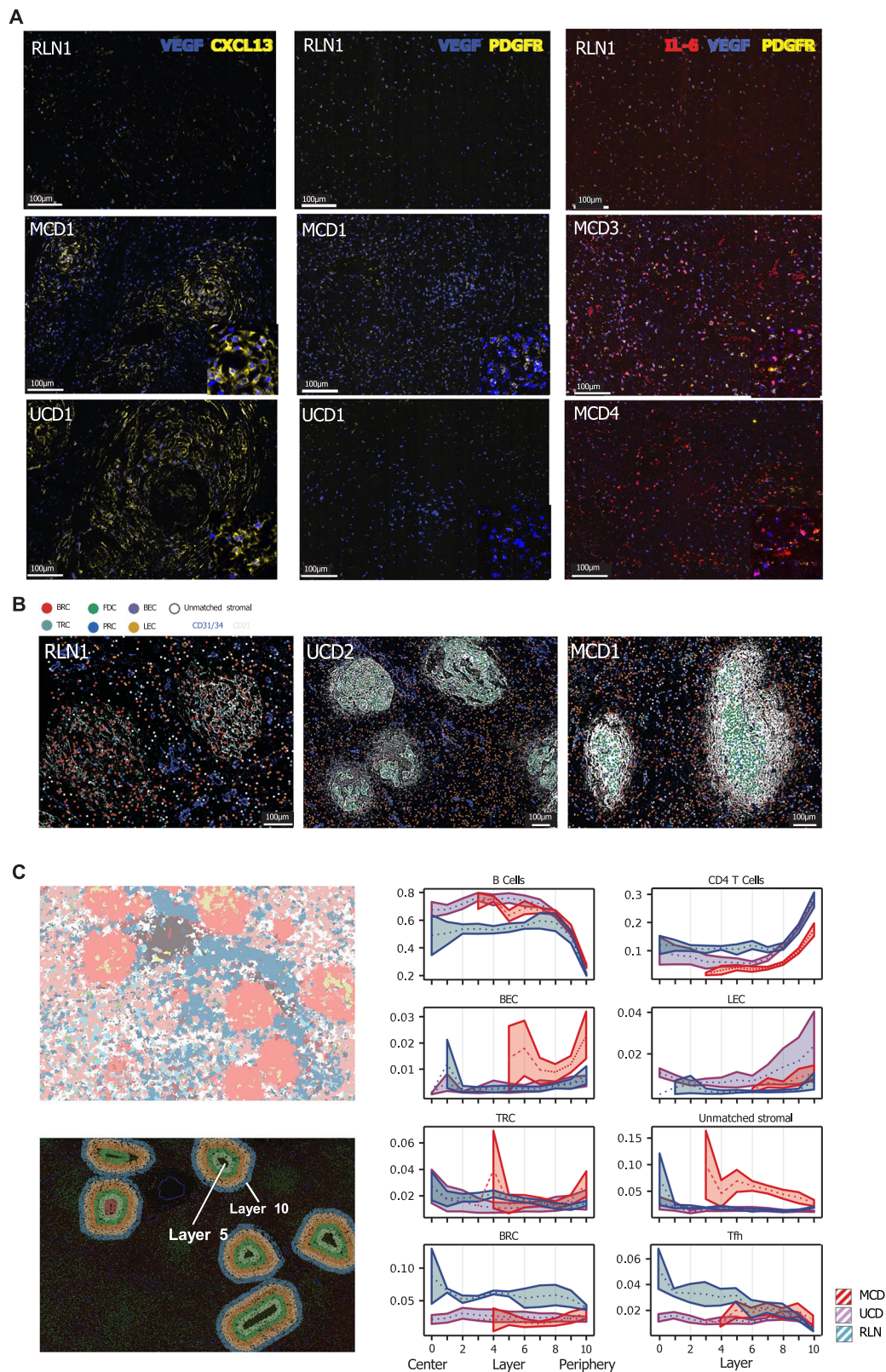
DNA sequencing, immune repertoire, and viral sequence analysis

Clinical NGS and SNP array DNA analysis did not reveal clinically significant Tier 1 or 2 single-nucleotide or copy number variants (Supplemental Table 3). Several Tier 3 variants of uncertain significance with high variant allele frequencies (VAFs) were noted in UCD1,3 and MCD1,4, suggesting possible germline variants. Reanalysis of data for variants previously reported in CD[8] revealed very low levels of *PDGFRB* c.1997 A > G, p.N666S (UCD2), *ALK* c.875 G > A, p.R292H (UCD2,3, MCD1,3) and *BCOR* c.3866 G > A, p.G1289D (UCD3) in some cases of UCD and MCD. These rare sequencing reads (2-5/1000) were close to the error rate of Illumina sequencing (1/1000) and thus are best considered as variants of uncertain significance.

5' sequencing was employed to analyze the T and B cell immune repertoire in a subset of cases. 1631 single-nuclei TCR and 6556 single-nuclei BCR sequences were identified from RLN, UCD, and MCD. Clonally restricted populations of B or T cells were not observed. VDJ and VJ analyses revealed that plasma cells of UCD3 and MCD3/4 showed high levels of somatic hypermutation and were predominantly class switched to IgG1 (Supplemental Fig. 6A, B) suggesting generation from a follicular reticular cell-driven germinal center process.

Given the association of MCD with HHV8 and HIV infection, we aligned sequencing reads to HHV8, HIV, and EBV reference genomes. HHV8 sequences were only identified in MCD4 with the highest proportion of viral sequences in plasma cells (Supplemental Fig. 6C). The

findings were consistent with the HHV8-driven etiology in MCD4. HHV-8, EB,V or HIV sequences were not detected in any other UCD, MCD, or RLN.

In conclusion, the findings indicate that the activation and proliferation of follicular reticular cells, resulting in VEGF and IL-6 secretion, are linked to B-cell activation, plasma cell differentiation, angiogenesis, and stromal remodeling across all subtypes of CD.

Discussion

Castleman disease (CD) encompasses a spectrum of lymphoproliferative disorders characterized by distinct histopathological and clinical features. Despite significant advances in understanding its pathogenesis, the cellular and molecular mechanisms driving CD remain incompletely understood. The absence of accurate cell culture or murine models limits functional studies of CD. Unlike malignant

Fig. 3 | Spatial Localization of Stromal Subtypes. A RNA ISH for *VEGFA/CXCL13*, *VEGF/PDGFRA* and *IL-6/VEGFA/PDGFRA* combinations in RLN (*n* = 1), UCD (*n* = 3) and MCD (*n* = 4) were evaluated. Markedly increased *VEGF* and *IL-6* in MCD that colocalized with *CXCL13* + FDC in follicles and *PDGFRA*+ stromal cells in interfollicular areas (2ⁿᵈ row). Colocalization of *CXCL13* + FDC and *VEGF* expression in follicles of UCD is noted (3ʳᵈ row). Co-expression of IL-6, VEGF and PDGFR+ stromal cells in MCD3 is shown (3ʳᵈ column, purple cells). Co-expression of IL-6 with PDGFR+ stromal cells in MCD4. Inset image shows high magnification. **B** Stromal subsets identified from snRNA-seq data are shown in CODEX images after integration by MaxFuse. BEC, LEC and PRC are associated with CD31/CD34+ blood vessels and are increased in UCD and MCD. Markedly increased FDC nuclei are closely associated with regions of dense CD21 meshworks. **C** Left column: Follicles were identified by cell microenvironment (Voronoi plot shown in top) and split into concentric 100-pixel contours (bottom). Stromal subtypes identified from snRNA-seq data were transferred to multiplexed image data and their abundance assessed at each contour layer (middle and right columns). Statistics were aggregated over each follicle per image region (MCD, *n* = 11; HVCD, *n* = 9, RLN, *n* = 3). Error bands represent a 90% confidence interval. Peaks of stromal cells are evident in layers before peaks of BEC and LEC of MCD. MCD is shown in red, UCD in purple and RLN in blue.

disorders, CD may not have a genetic basis but is instead an immune-driven disorder influenced by diverse antigenic stimuli, which are challenging to replicate in vitro without a clear understanding of their origins. Our study provides the most comprehensive spatial and single-cell characterization of CD to date, revealing key stromal cell subsets and cytokine pathways underlying the disease. This knowledge can serve as a foundation for developing more accurate in vitro and in vivo models of CD.

Our spatial transcriptomic and proteomic analyses localized the origin of key CD cytokines IL-6, VEGF, and CXCL13. While prior studies showed the expression of these cytokines in bulk tissue[12,20,21], their cellular origin was not clear and were thought to be secreted from hematopoietic cells. A recent case report identified PRC as the origin of VEGF in a set of twins with familial iMCD[19]. However, wide applicability is limited by lack of controls, cases of UCD or other subtypes of MCD. Our cohort is the largest cohort of MCD and UCD cases and includes rare variants such as iMCD-TAFRO, HHV8 MCD, and UCD with plasmacytosis.

Our findings demonstrate that distinct stromal cell subsets predominate in different subtypes of CD. In UCD, BRCs and FDCs were significantly expanded. While direct visualization of the specific stromal subtype requires immunostaining, the downstream effects of their activation are evident morphologically. The histological feature known as "onion skin" refers to the concentric layering of mantle zone B cells. Our study suggests a mechanistic role for FDCs in forming this structure. Transcriptomics and spatial analysis show that FDC activation and close interaction with mantle zone B cells is the basis of 'onion skin' appearance. VEGF overexpression by FDCs correlated with significantly increased perifollicular neovascularization and penetrating blood vessels, forming the basis of the diagnostic "lollipop sign". Hypervascularization is necessary for stromal remodeling and lymph node expansion. In contrast, MCD was marked by the expansion of TRCs and PRCs. VEGF expression was highest in PRCs and TRCs of MCD cases. IL-6 expression was predominantly observed in PRCs and TRCs of MCD cases. These findings suggest that UCD and MCD are driven by distinct stromal-immune interactions, influencing disease presentation and progression.

Previously, prominent FDCs in CD were considered remnants of attenuated germinal centers. Using single-cell analysis, we show that these FDCs actively upregulate multiple activation pathways, highlighting their critical role in the pathogenesis of CD. The expansion of FDC meshworks in UCD facilitates close interactions with mantle zone B cells, potentially driving chronic antigen presentation and aberrant immune activation leading to B-cell activation and plasma cell differentiation. In MCD, stromal cell-mediated activation of T cells led to extrafollicular B-cell activation, resulting in plasmablast expansion. Notably, VDJ sequencing revealed polyclonal, class-switched, and somatically hypermutated IgG1+ plasma cells, supporting a stromal-driven germinal center response rather than non-specific antigen-independent activation.

Our findings suggest potential novel therapeutic strategies for CD. Given the central role of FDCs and TRCs in sustaining cytokine production and immune activation, targeted therapies disrupting these stromal-immune interactions could be effective. For instance, lymphotoxin-β receptor fusion proteins may prevent FDC hyper-activation, reducing CXCL13 and VEGF production. Additionally, VEGF inhibitors could mitigate neovascularization, particularly in endothelial-rich subtypes of MCD. Given the presence of CD20-expressing proliferating plasmablasts in MCD, a combination of anti-CD20 monoclonal antibodies with IL-6 or proteasome inhibitors may provide better disease control. Finally, our identification of DLL4-NOTCH1 signaling between TRCs and T cells as a critical interaction in CD suggests that NOTCH inhibitors could modulate the aberrant immune activation seen in MCD.

This study provides a framework for understanding CD pathogenesis, identifying distinct stromal and immune interactions that differentiate UCD from MCD. Our findings pave the way for new diagnostic and therapeutic strategies targeting key stromal subsets and cytokine networks in CD. A limitation of the study is the moderate sample size and its observational nature. Future research should focus on validating these findings in larger cohorts and developing preclinical models to test targeted interventions.
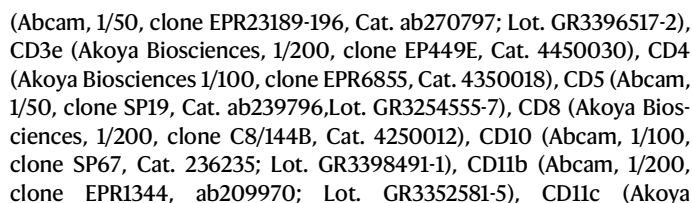
Methods
Specimen utilization workflow
The study was approved by the Children's Hospital of Philadelphia Institutional Review Board (IRB 16-013199) and performed in accordance with the Declaration of Helsinki with informed consent. Lymph node resections with clinicopathological diagnoses of UCD, MCD, and non-malignant non-infectious reactive lymph nodes (RLN) were identified from the institutional archives for the initial discover cohort (Table 1). We identified cases where rapid intraoperative frozen section diagnosis was conducted on whole lymph node resections. Upon removal from the patient, lymph node sections were embedded in Optimal Cutting Temperature (OCT) compound and frozen within a 10-minute window per clinical intraoperative diagnostic requirements that also ensured high-quality RNA reflecting the in vivo state. The lymph node was bisected along its long axis. One-half of the bisected lymph node was frozen, while the other half was fixed in formalin and subsequently embedded in paraffin (FFPE). The OCT-frozen half was used for single-nuclei transcriptomic analysis, while the corresponding FFPE half was used for single-cell proteomic imaging (Supplemental Fig. 1). Frozen tissue was used for DNA sequencing and SNP array analysis.

Multiplex spatial proteomic imaging and analysis
Image acquisition (CODEX/Phenocycler) and processing. Multiplexed spatial imaging of lymph node cross sections was performed using CODEX/Phenocycler from Akoya Biosciences. Clinically validated immunohistochemical markers were used to identify major immune and stromal cell types and subsets (Supplemental Table 1). After heat-induced epitope retrieval, a single 5 μm thick FFPE tissue was stained with a panel of forty-four DNA oligonucleotide-conjugated marker antibodies per the manufacturer's instructions. BCL-2 (Biolegend, 1/400 dilution; clone 100, Cat. 658702, Lot. B259980), BCL-6 (Biolegend, 1/200, clone IG191E/A8, Cat. 648301, Lot. B269747), CD1c

(Abcam, 1/50, clone EPR23189-196, Cat. ab270797; Lot. GR3396517-2), CD3e (Akoya Biosciences, 1/200, clone EP449E, Cat. 4450030), CD4 (Akoya Biosciences 1/100, clone EPR6855, Cat. 4350018), CD5 (Abcam, 1/50, clone SP19, Cat. ab239796,Lot. GR3254555-7), CD8 (Akoya Biosciences, 1/200, clone C8/144B, Cat. 4250012), CD10 (Abcam, 1/100, clone SP67, Cat. 236235; Lot. GR3398491-1), CD11b (Abcam, 1/200, clone EPR1344, ab209970; Lot. GR3352581-5), CD11c (Akoya

Biosciences, 1/50, clone 118/A5, At. 232177), CD15 (Biolegend, 1/200, clone HI98,Cat. 301902, Lot. B265372), CD20 (Akoya Biosciences, 1/400, clone L26, Cat. 4150018), CD21 (Biolegend, 1/400, clone Bu32, Cat. 354902, Lot. B267500), CD23 (Abcam, 1/100, clone SP23, Cat. ab242380; Lot. GR3285963-3), CD25 (Cell Marque, 1/100, clone 4C9, Cat. M210387575; Lot. 2022601 – line #5), CD30 (Cell Marque, 1/50, clone Ber-H2, M210387575, Lot. 1312708, #4), CD31 (Akoya

Fig. 4 | Driving Factors and Ligand-Receptor Interactions in CD. A Ligand-receptor interactions were assessed for each sample using the LIANA workflow. Inter-sample variation was analyzed with Tensor-cell2cell to identify distinct driver interactions ('factors') associated with disease states. Heatmaps display sample identity, source cell type, and receiver cell type for each factor. Enriched pathways for each factor are shown in the dot plot. Factor loadings and pathway weights were fit using a univariate linear model and p-values were estimated from a t-statistic as implemented in ref. [41]. UCD1 and 2 are characterized by factor 2. MCD1, 2 and 3 are characterized by factors 6, 10, and 11. MCD3 and 4 are additionally characterized by

factors 1, 8, 9, and 12. *denotes samples profiled using 5′ chemistry. **B** Significant (p < 0.05) LIANA-identified interactions of UCD and MCD cases compared to RLN are shown. Circos plots and dot plots with stromal subtypes as senders and all other cell types as receivers are shown. Dot plots illustrate individual ligand-receptor interactions clustered based on the Gene Ontology (GO) of the ligand. Interactions are grouped into categories representing ECM (integrin/collagen), inflammatory (complement components), and angiogenesis (VEGF) interactions. Source data are provided as a Source Data file in the supplementary dataset.

Biosciences, 1/200, clone EP3095, Cat. 4150017), CD34 (Invitrogen, 1/100, Clone QBEND/10, Cat. MA1-10202; Lot. WE3277032A), CD40 (Abcam, 1/100, clone EPR20735, Cat. ab228818; Lot. GR3362546-5), CD45 (Akoya Biosciences, 1/200, clone 2D1, Cat. 4250099), CD45RA (Biolegend, 1/200, HI100, Cat. 304102, Lot. B255475), CD68 (Akoya Biosciences, 1/300, clone KP1, Cat. 4350001), CD69 (Abcam, 1/100, clone EPR21814, Cat. Ab234512; Lot. GR3402812-1), CD123 (Biolegend, 1/100, clone 6H6, Cat. 306002; Lot. B254974), CD134 (Biolegend, 1/100, clone Ber-ACT35, Cat. 350002, Lot. B247739), CD138 (Novus Biologicals, 1/200, clone 160, Cat. NBP2-89900; Lot. MF12AI/2315), CD163 (Novus Biologicals, 1/100, clone EDHu-1, Cat. NB110-40686; Lot. 149022B), CLEC9A (Abcam, 1/50, clone EPR22324, Cat. ab245121; Lot. GR3348002-8), Fascin (Millipore, 1/300, clone 55k-2, Cat. MAB3582; Lot 3601726), FOXP3 (Invitrogen, 1/100, clone 236 A/E7, Cat. 14-477-84; Lot. 2378013), Granzyme B (Abcam, 1/200, clone EPR20129-217, ab208586), HLA-DR (Akoya Biosciences, 1/300, clone EPR3692, Cat. 4450029), ICOS (Cell Signaling, 1/100, clone D1K2T, Cat. 89601BF; Lot. No. 9), IRF4/MUM1 (Abcam, 1/50, clone EPR5653, Cat. ab247962; Lot. GR3296798-1), IRF8 (Invitrogen, 1/100, clone ZI003, Cat. 39-8800; Lot UE 288651), Ki67 (Akoya Biosciences, 1/250, clone B56, Cat. 4250019), Mac2/Gal3 (Akoya Biosciences, 1/400, clone M3/38, Cat. 4450034), MPO (R&D Systems, 1/200, clone MAB3174, Cat. MAB3174;Lot. ZFZ0221042), PAX5 (Cell Signaling, 1/50, clone D7H5X, Item 12709BF; Lot. No. 2), PD-1 (Cell Signaling, 1/200, clone D4W2J, Item 86163BF; Lot. No. 7), Podoplanin (Akoya Biosciences, 1/200, clone NC-08, Cat. 4250004), TCF4 (Abcam, 1/100, clone NCI-R1-59-6, Cat. ab217668, GR3321095-5), Vimentin (Biolegend, 1/400, clone O91D3, Cat. 677802, Lot. B256050) Sections were then stained with three fluorophores conjugated to DNA oligonucleotides[26] and imaged with an inverted microscope (Keyence BZ-X700) using fluorophore-tagged (AF488, Atto550, Cy5) DNA oligonucleotide reporters. The process was iteratively repeated for 19 cycles to image all markers in the panel. Raw images were processed through the CODEX processing software (CODEX Processor 1.8.2.13). Different image tiles with similar z-planes were combined to create one QPTIFF file for analysis. Background subtraction, deconvolution, shading correction, and cycle alignment were performed.
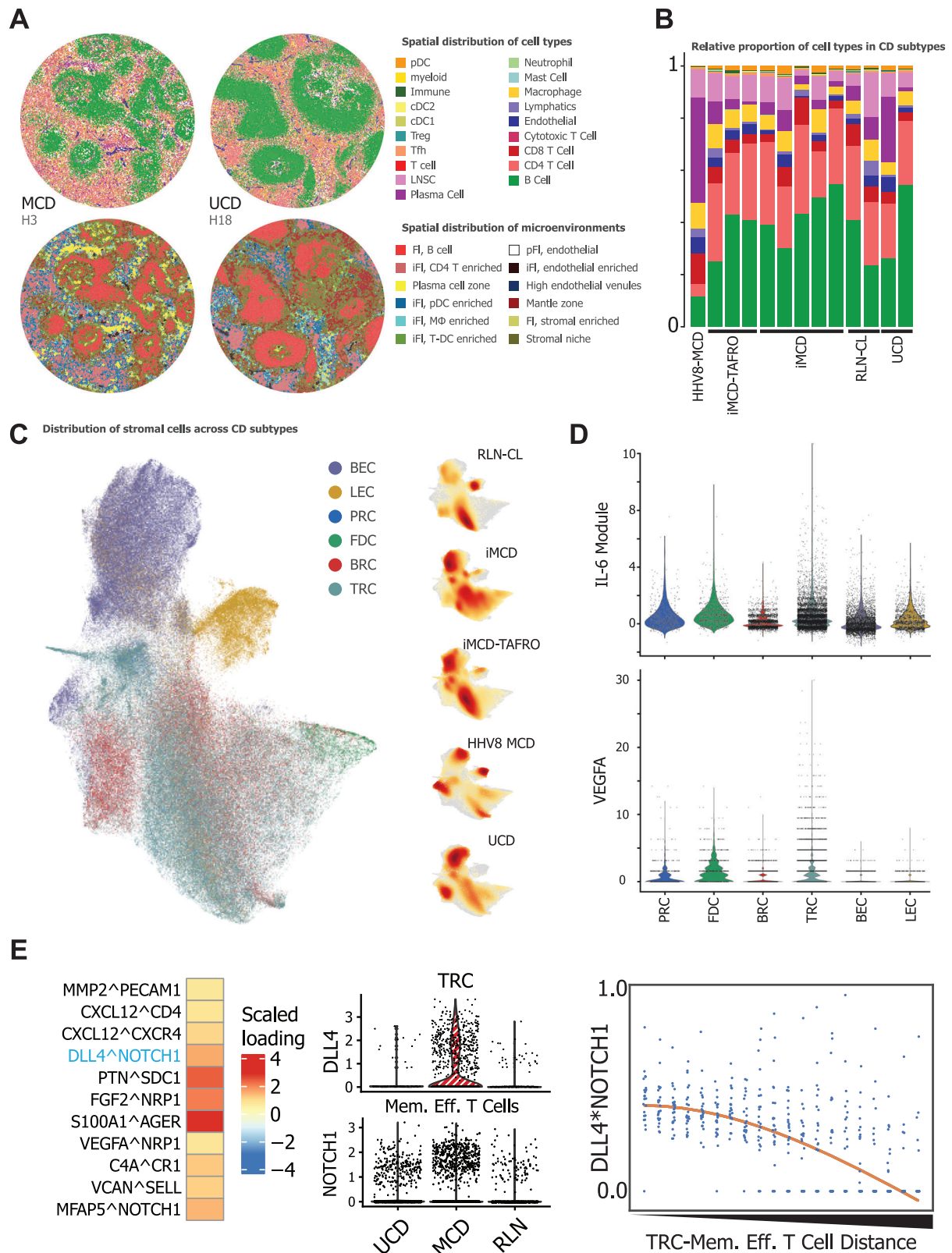
Lesional region selection and cell segmentation. Aggregated CODEX Processor images were split into component imaging regions. For each tissue, three to five representative regions were selected for a total of 35 regions. Within these regions, for each tile containing cells, the average channel intensity was collected. Using this sample population, each marker for each tile for all regions was evaluated using a two-tailed Grubb's test ($\alpha = 0.01$). If a tile contained more than 8 suboptimal markers/channels, then the tile was excluded from further analysis to ensure high-quality data. Over the entire experiment, this resulted in 211 of 2074 tiles being excluded. Cell segmentation was performed with CellSeg[27] using the default model weights. The GROWTH_METHOD and GROWTH_PIXELS parameters were determined through an iterative process, ultimately set as 'Sequential' and 5, respectively. CellSeg was also used for channel intensity quantitation with lateral bleed compensation enabled.

Cell phenotyping and identification. To assign cell type labels, data from representative regions of RLN, UCD and MCD (Supplemental Fig. 2A) were processed using Scanpy[28] v1.3.2. Briefly, marker intensities were log-normalized and scaled (Supplemental Fig. 2B). Cells were evaluated using Uniform Manifold Approximation and Projection (UMAP) and clustered using the Leiden algorithm[29]. Marker genes for each cluster were determined using Scanpy's rank_genes_groups function. Every cluster was assigned a cell type label after expert review of marker expression level and spatial localization of assigned cells (Supplemental Fig. 3). These representative data were used to construct a reference data set and predict cell labels for the remaining regions using Symphony[30] and Harmony[31], respectively. Resultant labels were reviewed and edited in batches through customized marker gating rules. This strategy was successful in annotating the majority of cells. Any remaining indeterminant cells were assigned labels using the phenotyping tools of Scimap (v0.22.9). The phenotyping matrix and gating values for these cells were iteratively determined by expert review. B cells, plasma cells, CD4 + T cells, CD8 + T cells, macrophages, regulatory T cells (Tregs), T follicular helper cells (Tfh), cytotoxic CD8 T cells, classic dendritic cells (cDC) 1 and 2, plasmacytoid dendritic cells (pDC), endothelial cells, lymphatic endothelial cells, neutrophils and mast cells were annotated using this methodology (Supplemental Fig. 1). Stromal cells were identified by the lack of pan-hematopoietic marker CD45 or other cell type specific markers in DAPI-positive nuclei. Minor populations of unclassified CD45+ cells were labeled as 'myeloid', and unclassified CD3+ cells were labeled 'immune'.

Cell distance analysis and spatial analysis. We used the suite of functionalities from Scimap (www.scimap.xyz) to determine the cellular and spatial architecture. Cell-cell distances and interactions (method = 'radius', radius=50) were calculated on a disease level. For each cell-cell pair combination, the distance distributions were scored using a Kolmogorov–Smirnov test against either the simulated distribution of complete spatial randomness or the empirical values measured in RLN tissue. P-values were adjusted using the Bonferroni method. The η_{50} distance was used to determine the closest cell pairings in CD compared to RLN.

Spatial neighborhoods were defined using Scimap's implementation of the Latent Dirichlet Allocation (LDA) method (method = 'radius', radius=50)[32]. The data were analyzed using 8, 10, 12, and 15 possible microenvironments. Twelve microenvironments provided the optimal balance between known lymph node structures and novel areas. Microenvironments were assigned labels based on unique cell type and spatial distribution.

Image and spatial statistics were calculated for each follicle, as defined by the "B cell, germinal center" and "B cell follicle, stromal enriched" annotations, and averaged for each region. The CD21 area was calculated as the thresholded CD21 fluorescence value and the CD21 signal was calculated as the normalized value in segmented cells. The concept of image entropy was used to quantify the degree of information and complexity in CD21 meshworks. Entropy was calculated as the local entropy within discs of 5-pixel radius. Across a follicle, this is simplified to *the root mean square.

Fig. 5 | Spatial transcriptomics of additional cohort of UCD and MCD cases.
Spatial transcriptomic profiling was performed on tissue microarrays of CD samples from 13 cases using the 10x Xenium platform. **A** Cell annotations and microenvironment analysis of an iMCD and UCD case are shown. Attenuated follicular (Fl) and expanded interfollicular (iFl) plasma cell and stromal microenvironments are noted in MCD. **B** Bar plots show relative cell proportions in subsets of RLN, MCD, and UCD. Expanded stromal and plasma cell subsets are noted. **C** UMAP visualization of stromal subtypes highlights distinct compositional differences between iMCD and UCD. Density projections illustrate differential expansion of BRC and TRC in UCD and MCD, respectively. BEC are expanded in both UCD and MCD (**D**). Elevated IL-6 module score (IL6, LIFR, OSMR) and VEGFA expression levels in iMCD cases. **E** DLL4^NOTCH1 interaction was evaluated in the Xenium data and showed a disease and distance dependency consistent with physical interaction. Source data are provided as a Source Data file in supplementary dataset.

snRNA-seq and analysis

Nuclei isolation and sequencing. Frozen tissue was sectioned into 40 μm sections, and nuclei were recovered using Chromium Nuclei Isolation Kit (PN-1000494) from 10x Genomics. Nuclei integrity and concentration were evaluated using a hemocytometer. Eight thousand nuclei were loaded onto a 10x Genomics Chromium Controller for a targeted recovery of 5000 nuclei. Single nuclei RNA was processed for sequencing by constructing gene expression (GEX) libraries (Chromium Next GEM Single Cell 3′ Kit v3.1, PN-1000268) or gene expression and V(D)J immune profiling libraries (Chromium Next GEM Single Cell 5′ Kit v2, PN-1000263; Chromium Single Cell Human TCR Amplification Kit, PN-1000252; Chromium Single Cell Human BCR Amplification Kit, PN-1000253). Briefly, the single nuclei suspension was mixed with RT Master Mix and loaded with barcoded single-cell 3′ or 5′ gel beads and partitioning oil onto microfluidic chips to encapsulate 5000 nuclei per sample using the Chromium Controller. After reverse transcription and cleanup, cDNA libraries were generated according to manufacturer's instructions with one additional cycle of PCR amplification to account for the relatively lower amount of RNA in nuclei than in whole cells. cDNA was fragmented and end-repaired, size-selected and PCR amplified to generate a 3′ or 5′ gene expression library. For VDJ library construction, full-length TCR or BCR transcripts were enriched from 4 μL of amplified cDNA, and 50 ng of enrichment product was used for library construction. Libraries were submitted for sequencing on an Illumina Novaseq S1-100 flow cell for a minimum sequencing depth of 25,000 reads/nuclei for gene expression and 5000 reads/nuclei for V(D)J profiling. Reads were processed using CellRanger Single-Cell Software Suite (v6.1.1 and v7.0.0) from 10x Genomics. Reads were aligned to the human reference genome GRCh38. Custom references were also prepared to include Epstein-Barr virus (NCBI:txid10376), human herpesvirus 8 (NCBI:txid37296), and human immunodeficiency virus 1 (NCBI:txid11676).

Data integration and annotation. Unique molecular identifier (UMI) counts were corrected for ambient RNA expression in R (v4.1.2) using SoupX[33] (v1.6.1). Further processing, visualization, and clustering was performed with Seurat[34] (v4.3.0). Briefly, counts were normalized using SCTransform, and cells were filtered on a library-specific basis of number of features, total counts, and mtRNA and rRNA content. Putative doublets were identified using DoubletFinder[35] (v2.0.3). Fully processed count data were integrated using Seurat's RPCA method. After integration and inspection of cell clusters, additional filtering was applied to remove likely cell debris based on feature count. Ultimately, the integrated data set consisted of 50117 nuclei. Ribosomal, mitochondrial, immunoglobulin, and HLA genes were filtered out. Cells were clustered with the Leiden algorithm and cell type labels were manually assigned based on the top marker genes. Subsequent data sets were annotated through label transfer using Symphony[31] (v0.1.0).

Differential expression and enrichment analyses. Differentially expressed genes (DEG) between RLN, UCD, and MCD cases were detected using the Wilcoxon Rank sum test implemented with Seurat. GEM libraries generated from 5′ and 3′ sequencing chemistries were batch corrected using ComBat[36] (SVA v3.46.0). Cell types were filtered on the criteria that at least 50 cells be represented among test groups. Certain genes were broadly upregulated across multiple cell types. To identify cell-type-specific DEG, the frequency of cluster significance was calculated for each gene, and genes in the top $\eta_{.96}$ were filtered out. Enrichment analysis was conducted with gProfiler2 (v0.2.1). As input to gProfiler, DEG were filtered by p-value < 0.1. AUCell[37] was used to score the signature of a gene set for each cell.

Ligand-receptor interaction analysis. Ligand-receptor interactions between cell-cell pairings were investigated using LIANA[24] (v0.1.8). For each case, top-scoring interactions were defined by an aggregate score<0.1. Unique interactions of CD compared to RLN were identified. Inter-sample variation was assessed using Tensor-cell2cell[25] to identify distinct driver interactions ('factors') associated with disease states. Cell types represented in less than 30% of the samples, based on a criterion of 30 observations, were excluded from the analysis. Individual ligand-receptor pair loadings were evaluated for enrichment against a weighted database of pathways provided by PROGENy[38].

Multiplex RNA in situ hybridization. RNA in situ hybridization (ISH) was performed on a Bond Rx (Leica Biosystems) automated staining platform using RNAscope LS Multiplex assay TSA Vivid dyes (ACD, 323275) according to the manufacturer's instructions. Probes for *VEGFA* (ACD 423168-C2), *IL-6* (ACD, 310378), *CXCL13* (ACD, 311328-C3), *CXCL12* (ACD, 422998-C3), *PDGFRA* (ACD, 604488-C3), *ACTA2* (ACD, 444778-C3) and *CD19* (ACD, 402718-C3) were obtained. Three RNA-scope probes were stained along with DAPI on sections from cases and controls. Stained slides were digitally scanned at 20X magnification on an Aperio FL slide scanner (Leica Biosystems, Germany). Appropriate positive and negative controls were used per the manufacturer's recommendation to ensure RNA integrity and exclude background signal.

VDJ sequencing and analysis. Immune repertoire profiling from V(D)J sequences was performed using the R package, Platypus (v3.5.0). Clonotypes were defined using the "double.and.single.chains" option. For quantitation of somatic hypermutation in B cells, full length sequences were assembled using MiXCR.

Targeted DNA sequencing and copy number variant analysis. A targeted next generation sequencing (NGS) panel which interrogates 238 malignancy-associated genes for sequence and copy number variants (CNVs) was performed on the extracted DNA. Briefly, DNA was fragmented and tagged for target enrichment using SureSelect[QXT] reagent kit (Agilent Technologies to generate adapter-tagged libraries, which were subjected to sequence analysis on the Illumina HiSeq platform for 150 bp paired-end reads (Illumina Inc.). All coding exons and the flanking intron sequences of targeted genes along with selected promoter and intronic regions were sequenced with a targeted average sequence depth of 1800x. NGS data were analyzed using clinical laboratory software ConcordS V2 and NextGENe V2 NGS Analysis Software (SoftGenetics). The annotated sequence variants and CNVs were classified per clinical somatic variant guidelines[39].

SNP array analysis. Genome-wide SNP array analysis was performed on genomic DNA extracted from lymph node tissue of the patients using the Illumina Infinium CytoSNP-850Kv1.2 BeadChip (Illumina Inc). The data were analyzed using vendor-provided analysis software (GenomeStudio). All genomic coordinates were based on the GRCh37/hg19 build of the human genome. The assay detects chromosomal gains, losses, and copy-neutral loss of heterozygosity (cnLOH) involving ≥10 SNP probes that are present in at least 10% of cells.

Validation Cohort of CD. An additional validation cohort of CD samples was identified (Supplemental Table 4). Tissues were collected in accordance with the Declaration of Helsinki, and data collection protocols were approved as being of minimal to no risk or exempt by the Institutional Review Board of Duke University (protocol ID Pro00104078, date of approval Oct 24, 2019). Twenty FFPE CD patient samples were constructed into a tissue microarray (TMA; 2 samples per patient). Upon H&E assessment, seven samples were excluded from downstream analysis due to lack of lesional tissue in the TMA. Histopathological diagnosis was confirmed after review of H&E and clinicopathological information.

10x Xenium spatial transcriptomics. For high-resolution in-situ gene expression, 10x Xenium spatial transcriptomics was conducted. The pre-designed Xenium Human Multi-Tissue and Cancer Panel was used with 100 additional custom-designed probes (FFVG7Q), focussing on CD specific genes. These 100 additional genes included 8 HHV8 (KSHV) specific probes designed against LANA1 (ORF73), vIL6 (K2), vIRF1, PF8 (ORF59), ORF36 and ORF21. Human FFPE Castleman's disease tissue TMA sections were cut to 5 micron and mounted onto Xenium slides. Deparaffinization and target retrieval were performed, and the sections were then hybridized with gene probes overnight at 50 °C, which was then followed by post-hybridization washes. Subsequently, ligation was conducted at 37 °C for 2 h and amplification was conducted at 30 °C for 2 h. The tissues were then autofluorescence-quenched. Nuclei staining was performed before loading onto the Xenium Analyzer instrument for imaging. Images were then processed and segmented through the 10x Xenium processing pipeline and results were reviewed in the Xenium Explorer app. Cell segmentation was achieved by identifying nuclei from the DAPI signal and growing the membrane mask from this nuclear mask.

10x Xenium data analysis. Raw 10x Xenium data was processed through the Xenium Ranger v3.0.1 pipeline. Cell annotations were transferred from the snRNA-seq dataset using scArches v0.6.1. LDA neighborhood probabilities were calculated de novo and assigned with k-means clustering. As with the CODEX data, the number of motifs and clusters was 12. Based on cell type proportions, neighborhoods were aggregated and named to match the CODEX profiles. For ligand-receptor cell-cell interactions, the nearest paired cell by euclidean distance was selected for the ligand-receptor score.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Raw sequencing data for all snRNA-seq experiments have been deposited in the GEO database under accession code GSE296614. The snRNAseq, CODEX images and Xenium source datasets have been deposited in FigShare. (https://doi.org/10.6084/m9.figshare.28394915). All other data are available in the article and its Supplementary files or from the corresponding author upon request. Source data are provided with this paper.

Code availability

All custom code is provided at github.com/qu4drupole/Castleman_2025 (https://doi.org/10.6084/m9.figshare.28873094.v1).

References

1. Wang, H. W., Pittaluga, S. & Jaffe, E. S. Multicentric castleman disease: Where are we now?. *Semin. Diagnostic Pathol.* **33**, 294–306 (2016).
2. Nishimura, M. F., Nishimura, Y., Nishikori, A., Yoshino, T. & Sato, Y. Historical and pathological overview of Castleman disease. *J. Clin. Exp. Hematop* **62**, 60–72 (2022).
3. Dispenzieri, A. & Fajgenbaum, D. C. Overview of castleman disease. *Blood* **135**, 1353–1364 (2020).
4. Keller, A. R., Hochholzer, L. & Castleman, B. Hyaline-vascular and plasma-cell types of giant lymph node hyperplasia of the mediastinum and other locations. *Cancer* **29**, 670–683 (1972).
5. Wu, D., Lim, M. S. & Jaffe, E. S. Pathology of castleman disease. *Hematol. Oncol. Clin. North Am.* **32**, 37–52 (2018).
6. van Rhee, F. et al. International evidence-based consensus diagnostic and treatment guidelines for unicentric Castleman disease. *Blood Adv.* **4**, 6039–6050 (2020).
7. Zhang, M. Y. et al. UCD with MCD-like inflammatory state: surgical excision is highly effective. *Blood Adv.* **5**, 122–128 (2021).
8. Li, Z. et al. Recurrent PDGFRB mutations in unicentric castleman disease. *Leukemia* **33**, 1035–1038 (2019).
9. Frizzera, G., Banks, P. M., Massarelli, G. & Rosai, J. A systemic lymphoproliferative disorder with morphologic features of Castleman's disease. Pathological findings in 15 patients. *Am. J. Surg. Pathol.* **7**, 211–231 (1983).
10. Hawkins, J. M. & Pillai, V. TAFRO syndrome or castleman-kojima syndrome: a variant of multicentric Castleman disease. *Blood* **126**, 2163 (2015).
11. Yoshizaki, K. et al. Pathogenic Significance of Interleukin-6 (IL-6/BSF-2) in Castleman's Disease. *Blood* **74**, 1360–1367 (1989).
12. Nishi, J. et al. Expression of vascular endothelial growth factor in sera and lymph nodes of the plasma cell type of Castleman's disease. *Br. J. Haematol.* **104**, 482–485 (1999).
13. Pierson, S. K. et al. Plasma proteomics identifies a 'chemokine storm' in idiopathic multicentric castleman disease. *Am. J. Hematol.* **93**, 902–912 (2018).
14. Soulier, J. et al. Kaposi's sarcoma-associated herpesvirus-like DNA sequences in multicentric Castleman's disease. *Blood* **86**, 1276–1280 (1995).
15. Brousset, P., Cesarman, E., Meggetto, F., Lamant, L. & Delsol, G. Colocalization of the viral interleukin-6 with latent nuclear antigen-1 of human herpesvirus-8 in endothelial spindle cells of Kaposi's sarcoma and lymphoid cells of multicentric Castleman's disease. *Hum. Pathol.* **32**, 95–100 (2001).
16. Beck, J. T. et al. Brief report: alleviation of systemic manifestations of castleman's disease by monoclonal anti-interleukin-6 antibody. *N. Engl. J. Med* **330**, 602–605 (1994).
17. Arenas, D. J. et al. Increased mTOR activation in idiopathic multicentric castleman disease. *Blood* **135**, 1673–1684 (2020).
18. Nagy, A. et al. Next-generation sequencing of idiopathic multicentric and unicentric castleman disease and follicular dendritic cell sarcomas. *Blood Adv.* **2**, 481–491 (2018).
19. Chan, J. Y. et al. Single-cell landscape of idiopathic Multicentric castleman disease in identical twins. *Blood* **143**, 1837–1844 (2024).
20. Wing, A. et al. Transcriptome and unique cytokine microenvironment of castleman disease. *Mod. Pathol.* **35**, 451–461 (2021).
21. Horna, P., King, R. L., Jevremovic, D., Fajgenbaum, D. C. & Dispenzieri, A. The lymph node transcriptome of unicentric and idiopathic multicentric Castleman disease. *Haematologica* **108**, 207–218 (2023).
22. Lai, Y.-M. et al. Expression of interleukin-6 and its clinicopathological significance in Castleman's disease. *Chin. J. Hematol.* **34**, 404–408 (2013).
23. MaxFuse enables data integration across weakly linked spatial and single-cell modalities. *Nat Biotechnol.* **42**, 1096–1106 (2024).
24. Dimitrov, D. et al. Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data. *Nat. Commun.* **13**, 3224 (2022).
25. Armingol, E. et al. Context-aware deconvolution of cell-cell communication with Tensor-cell2cell. *Nat. Commun.* **13**, 3665 (2022).
26. Black, S. et al. CODEX multiplexed tissue imaging with DNA-conjugated antibodies. *Nat. Protoc.* **16**, 3802–3835 (2021).
27. Lee, M. Y. et al. CellSeg: a robust, pre-trained nucleus segmentation and pixel quantification software for highly multiplexed fluorescence images. *BMC Bioinforma.* **23**, 46 (2022).
28. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
29. Traag, V. A., Waltman, L. & van Eck, N. J. From louvain to leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).
30. Korsunsky, I. et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).

31. Kang, J. B. et al. Efficient and precise single-cell reference atlas mapping with Symphony. *Nat. Commun.* **12**, 5890 (2021).
32. Chen, Z., Soifer, I., Hilton, H., Keren, L. & Jojic, V. Modeling multiplexed images with. *J. Comput Biol.* **27**, 1204–1218 (2020).
33. Young, M. D. & Behjati, S. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *Gigascience* **9**, giaa151 (2020).
34. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 e3529 (2021).
35. McGinnis, C. S., Murrow, L. M. & Gartner, Z. J. Doubletfinder: doublet detection in single-cell rna sequencing data using artificial nearest neighbors. *Cell Syst.* **8**, 329–337 e324 (2019).
36. Zhang, Y., Parmigiani, G. & Johnson, W. E. ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genom. Bioinform* **2**, lqaa078 (2020).
37. Aibar, S. et al. SCENIC: single-cell regulatory network inference and clustering. *Nat. Methods* **14**, 1083–1086 (2017).
38. Schubert, M. et al. Perturbation-response genes reveal signaling footprints in cancer gene expression. *Nat. Commun.* **9**, 20 (2018).
39. Horak, P. et al. Standards for the classification of pathogenicity of somatic variants in cancer (oncogenicity): Joint recommendations of Clinical Genome Resource (ClinGen), Cancer Genomics Consortium (CGC), and Variant Interpretation for Cancer Consortium (VICC). *Genet Med* **24**, 986–998 (2022).
40. Raudvere, U. et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res* **47**, W191–W198 (2019).
41. Badia-I-Mompel, P. et al. decoupleR: ensemble of computational methods to infer biological activities from omics data. *Bioinform Adv.* **2**, vbac016 (2022).

Author contributions
V.P. performed study concept and design; V.P., D.S., A.E., E.F., J.W., A.R., K.T., M.Z., M.L., P.M., C.P., and B.R. performed writing, review, and revision of the paper; V.P., D.S., A.E., E.E., E.F., Z.X., A.S., X.Y., D.M., J.W., A.R., K.Y., and M.L. provided acquisition, analysis and interpretation of data, and statistical analysis.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-61214-1.

Correspondence and requests for materials should be addressed to Vinodh Pillai.

Peer review information *Nature Communications* thanks Antonino Carbone, Annunziata Gloghini and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.