

Spatial and molecular profiling of the classic Hodgkin lymphoma microenvironment reveals an immunosuppressive mononuclear phagocyte network

Benjamin Stewart

University of Cambridge

Martin Fergie

University of Manchester

Matthew Young

Wellcome Trust Sanger Institute <https://orcid.org/0000-0003-0937-5290>

Claire Jones

Newcastle upon Tyne NHS Foundation Trust

Ashwin Sachdeva

University of Manchester

Alex Blain

Newcastle University

Chris Bacon

Newcastle University

Vikki Rand

Teeside University

John Ferdinand

University of Cambridge

Kylie James

Garvan Institute of Medical Research

Krishnaa Mahbubani

University of Cambridge

Liz Hook

University of Cambridge

Nicolaas Jonas

University of Cambridge

Nicholas Coleman

University of Cambridge

Kourosh Saeb-Parsy

Cambridge University <https://orcid.org/0000-0002-0633-3696>

Matthew Collin

University of Newcastle <https://orcid.org/0000-0001-6585-9586>

Menna Clatworthy

University of Cambridge

Sam Behjati

Wellcome Sanger Institute <https://orcid.org/0000-0002-6600-7665>

Christopher Carey (✉ christopher.carey@newcastle.ac.uk)

Newcastle University <https://orcid.org/0000-0002-4293-3920>

Brief Communication**Keywords:**

Posted Date: January 7th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1233380/v1>

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

[Read Full License](#)

Abstract

Although a lymph node infiltrated by classic Hodgkin lymphoma is mostly composed of non-neoplastic immune cells, the malignant Hodgkin Reed-Sternberg cells (HRSC) successfully suppress an anti-tumor immune response, to create a cancer-permissive microenvironment. Accordingly, unleashing the dormant immune cells, for example by checkpoint inhibition, has been a central focus of recent therapeutic advances for this disease. Here, we profiled the global immune cell composition of normal and diseased lymph nodes by single-cell RNA sequencing, as a basis for interrogating the immediate vicinity of HRSC, first regionally and then at cellular resolution. Our analyses revealed specific immune cells and functional states associated with HRSC. Most prominently, we discovered a non-random spatial association of immunoregulatory mononuclear phagocytes positioned around HRSC, which express the immune checkpoints PD-L1, TIM-3, and the tryptophan-catabolizing protein IDO1. These findings provide a basis for rational targeting and activation of the anti-tumor immune response in classic Hodgkin lymphoma.

Main Text

Classic Hodgkin lymphoma (cHL) accounts for 12-15% of lymphoma diagnoses and occurs most frequently in adolescents and young adults, with another peak of incidence in older adults. Multi-agent cytotoxic chemotherapy (including selective use of radiotherapy) is curative in 80-95% of patients; in adults the potential for cure is largely determined by the disease stage¹⁻³. However, for patients with relapsed or refractory disease, the prospect of cure or long-term disease control is significantly diminished.

A cHL tumor is characterized by a minority population of neoplastic CD30⁺ Hodgkin Reed-Sternberg Cells (HRSC), comprising 1-5% of cells, derived from germinal centre (GC) B cells⁴. HRSC survive within a tumor microenvironment (TME), primarily composed of immune cells, containing abundant T-cells and a network of mononuclear phagocytes (MNP), including macrophages (MΦ), monocytes, and dendritic cells (DC)⁴.

HRSC establish an immunosuppressive niche through multiple mechanisms, including downregulation of major histocompatibility complex (MHC) proteins, and enhanced expression of programmed cell death ligand-1 (PD-L1), resulting from 9p23-p24 copy number gain or amplification^{5,6}. PD-L1 diminishes T-cell activation by ligation of its cognate receptor PD-1, which has been targeted by anti-PD-1 immune checkpoint blockade in relapsed/refractory disease, with evidence of therapeutic efficacy in the majority of patients⁷⁻⁹. Indeed, PD-L1 expression by HRSC and immune cells, including CD68⁺ tumor-associated MΦ (TAM), is almost universal in cHL¹⁰. However, the phenotypes, roles, and intercellular communication networks of MNP within the TME remain largely unknown. We hypothesize that MNP maintain the immunosuppressive niche through multiple mechanisms. Here, we present a robust characterization of the critical intercellular communication networks in the tumor microenvironment, with the aim of identifying novel therapeutic targets.

To gain an overview of the cellular ecosystem of cHL, we compiled a census of single cell transcriptomes, of 243,753 cells from lymphoma affected and unaffected lymph nodes. We sourced data from Aoki *et al*¹¹ incorporating droplet encapsulation single cell transcriptomes (10X genomics platform) from reactive and cHL nodes (Extended Data Fig. 1a). Additionally, we performed scRNAseq (10X genomics platform) with cell suspensions from healthy lymph nodes acquired from deceased organ donors, and two lymphoma lymph nodes, one with nodular sclerosis cHL (NSCHL), and one with nodular lymphocyte predominant Hodgkin lymphoma (NLPHL), a biologically distinct subtype of Hodgkin lymphoma (Extended Data Fig. 1a). Together, the combined dataset comprises tissue from 13 non-lymphoma-affected donors (8 deceased donors and 5 donors with reactive lymph node hyperplasia), plus one NLPHL and 23 cHL lymph nodes (Extended Data Fig. 1a). After quality control (Methods, Extended Data Fig. 1b-c), we performed dataset integration and dimensionality reduction using single-cell variational inference¹² and annotated cell types on the basis of marker genes and external dataset validation (Fig. 1a, Extended Data Fig. 1d-f).

The cellular ecosystem of cHL encompassed subsets of CD4⁺ and CD8⁺ T-cells, including regulatory T-cells (Treg), T follicular helper cells (Tfh), and exhausted CD4⁺ (ThExh) and CD8⁺ T-cells (CD8 TExh) cells (Fig. 1a). ThExh matched the recently identified LAG3⁺ subset, and expressed checkpoint molecules and exhaustion markers including *CD27*, *TNFRSF18*, *LAG3*, and *ICOS*^{11,13} (Fig. 1a, Extended Data Fig. 1g). The B cell compartment split into two large clusters of memory and naive subsets, in addition to plasmablasts, and germinal center B cells (Fig. 1a, Extended Data Fig. 1d-f).

Within the myeloid compartment, we identified MΦ coexpressing *CD14*, *CD68*, and the M2 polarisation markers *FOLR2* and *MRC1* (Fig. 1b-c). These cells were transcriptionally distinct from classical monocytes, which expressed a characteristic signature including *S100A9*, *CD14*, *VCAN*, and *FCN1*¹⁴ (Fig. 1b-c). Amongst DCs, we identified cDC1 (key transcripts: *CLEC9A*, *CADM1*, and *IDO1*) and cDC2 (key transcripts: *CD1C*, *CLEC10A*, and *FCER1A*) (Fig. 1b-c). In addition to cDC1 and cDC2, we also identified a population of *LAMP3*⁺ DCs in both healthy and some lymphoma samples, expressing the chemokine receptor *CCR7* and the chemokines *CCL17* and *CCL19*, which we termed “activated DCs” (aDC) consistent with the nomenclature of transcriptionally similar cells described in human thymus and spleen^{15,16}, and in murine lung neoplasms¹⁷ (Fig. 1b-c, Extended data Fig. 1e). We found a population of plasmacytoid DCs (pDC) with a dominant contribution from lymphoma samples (Fig. 1d, Extended Data Fig. 2a), expressing *IL3RA* (CD123), *CLEC4C*, and *CXCR3* (Fig. 1b-c). The undiseased reference samples contributed a rare population of mast cells, expressing characteristic marker genes including *CPA3*, *KIT*, and *TPSAB1* (Fig. 1b-c).

A key challenge in comparing the cellular architecture of lymph nodes lies in quantifying the relative enrichment of different cell-types in health and disease. This has recently been addressed through a differential abundance testing method, based on partially overlapping neighborhoods of cells on a k-NN graph (MiloR tool¹⁸). Applying MiloR, we represented the integrated manifold as a k-NN graph partitioned into 22,573 overlapping neighborhoods, before statistical testing for differential cell abundance between

lymphoma affected and unaffected within each neighborhood (Methods). In lymphoma nodes, the T-cell compartment was polarised towards cytotoxic subsets including NK cells and effector memory CD8+ T-cells (CD8 Tem), in addition to CD4⁺ T-cells subsets expressing checkpoint apparatus including ThExh and Tfh cells (Fig. 1e, Extended Data Fig. 2b). The association of these subsets with disease subtypes diverged between cHL and NLPHL. cHL subtypes were enriched for exhausted T-cell subsets, which contrasted with the NLPHL sample, where the cellular landscape was dominated by cytotoxic subsets (NK cells, CD8 Tem, and CD8 Tcm) (Extended Data Fig 2c).

To fully appreciate the cHL microenvironment, it is crucial to assess how HRSC influence their immune neighbors. The scRNAseq dataset did not contain a population of HRSC, likely due to the rarity, size, and fragility of these cells. To establish the transcriptional program of HRSC, we leveraged a microarray dataset profiling microdissected HRSC and GCs¹⁹. Differential expression between HRSC and GC established an upregulated HRSC gene signature, including *TNFRSF8* (CD30) (Fig. 1f). Scoring of transcription factor regulons demonstrated activation of NF- κ B (*NFKB1* and its activatory heterodimer partner *RELA*) in HRSC (Fig. 1g), consistent with previous reports²⁰. We next performed *in silico* identification of molecular interactions between active transcription factors and potential targets within the HRSC geneset (Methods). This demonstrated an NF- κ B-centric network coordinating the upregulation of the chemokines *CCL5*, *CCL17*, and *CCL22* capable of the positioning and retention of ThExh via CCR5 and CCR4 ligation (Fig. 1h, Extended Data Fig 1g).

We used this lymph node-wide account of the immune landscape of Hodgkin lymphoma as a reference to interrogate the cellular composition of cHL tissues. We specified regions surrounding HRSCs, using targeted spatial transcriptomic profiling (Nanostring GeoMx Cancer Transcriptome Atlas), defining 300 μ m diameter regions of interest (ROI) in both PD-L1^{high} and PD-L1^{low} regions of 9 NSCHL and 1 Mixed cell cHL (MCCHL) lymph nodes, and follicular and interfollicular regions of one control reactive lymph node (Extended Data Fig. 3a-b). We represented these transcriptional profiles in a shared-nearest neighbor graph and identified 5 types (clusters) of microenvironment across cHL and control lymph nodes (Fig. 2a-b). We deconvoluted the cell composition of each cluster, using our scRNAseq atlas as reference, followed by cell-type count estimation and differential cell-type abundance analysis (Fig. 2c). The five types of microenvironments encompassed two 'neoplastic' HRSC enriched clusters (clusters 3 & 4), two 'non-neoplastic' (i.e. devoid of HRSC signatures; clusters 1 & 5), and one intermediate neighborhood (cluster 2). The two non-neoplastic neighborhoods were PD-L1^{low} and represented the normal lymph node follicular microenvironment (cluster 5) or fibrosing regions of diseased lymph nodes (cluster 1). Notably, visibly fibrotic regions are an adjunct diagnostic histological feature of NSCHL²¹. The neoplastic environments were PD-L1^{high} and exhibited divergent leukocyte enrichment. Cluster 3 was enriched for ThExh, Th cells, and NK cells, whereas cluster 4 exhibited myeloid cell infiltration, with enrichment of classical monocytes, macrophages, and cDC2 (Fig. 2c). Differentially expressed genes in cluster 4 suggested potent inflammatory signalling, with high expression of chemokines associated with Th2 responses and CCR3-dependent eosinophil recruitment (*CCL18*, *CCL13*, *CCL24*, *CCL26*, *CCL23*) and granulocyte attracting chemokines *CXCL1*, and *CXCL6* (Extended Data Fig. 3c). Cluster 1 contained

almost exclusively PD-L1^{low} regions and was enriched for stromal (fibroblast and endothelial) cells and classical monocytes (Fig. 2c). Differentially expressed genes in this cluster included genes encoding key signalling mediators of fibrosis *FGFR4*, *TGFB2*, and *PTCH1*, in addition to *TNN* encoding the matrix extracellular component tenascin (Extended Data Fig. 3c). Overall, these findings suggest heterogeneous immune infiltrates within the broader cHL microenvironment, some of which may progress to fibrosis.

Given the diversity of MNP enrichment in neoplastic regions, we sought to establish the spatial relationships of MNP subsets within the HRSC niche at single-cell resolution. Using MNP marker gene co-expression patterns in the scRNAseq data (Extended Data Fig 3d), we designed multiplexed immunofluorescence (IF) panels to identify MNPs in fixed cHL tissue sections. We identified CADM1⁺/CD11c⁺ cDC1, CD1c⁺/CD11c⁺ cDC2, LAMP3⁺ aDC with highly distinctive dendritic morphology, CD123⁺ pDC, CD11c⁺ monocytes and MΦ (CD11c⁺ ONLY), and CD30⁺ HRSC (Fig. 2d). We then phenotyped segmented cells and performed neighborhood analyses, taking a 25 μm-radius neighborhood around each CD30⁺ HRSC and measuring the relative enrichment of MNP subsets in these aggregated neighborhoods, compared to the remaining regions ('non-neighborhood') across the tissue section. This analysis revealed enrichment of cDC2 and CD11c⁺ monocytes in the immediate vicinity of HRSC across cHL samples. In contrast, both pDC and aDC were excluded from the HRSC niche and occupy regions with a low density of CD11c⁺ cells (Fig. 2e-g, Extended Data Fig. 4a-h). These findings indicate non-random, active organisation of MNP subsets in the immediate HRSC-neighborhood.

Accordingly, we next asked which signals might coordinate the positioning of the MNP subsets and T-cells found in close association with HRSC. To interrogate these ligand-receptor interactions (LRI), we calculated the statistical enrichment of candidate LRI between MNPs and T-cells in our scRNAseq data using the CellPhoneDB tool²². This analysis predicted paracrine CCL3 and CCL4 signalling by MΦ and classical monocytes to *CCR5*- and *CCR1*-expressing cDC2 and ThExh (Fig. 3a). Furthermore, classical monocyte derived CXCL10 was predicted to signal to cDC1, ThExh, and Tfh via CXCR3. Nominated inhibitory interactions from ThExh included TIGIT signalling via NECTIN2 expressed by cDC2, classical monocytes, and MΦ (Fig. 3a).

Despite the close proximity of cDC2, macrophages, and monocytes to HRSC, the anti-tumor immune response remains inadequate in cHL. The failure of cDC2 to prime an effective T cell response may result from aberrant upregulation of inhibitory ligand and receptor molecules, amplifying immunosuppressive signalling in the HRSC microenvironment. To test this hypothesis, we designed a second IF panel to examine the expression of PD-L1, IDO1, and TIM-3 on CD11c⁺ and CD68⁺ cells, and their spatial relationships to CD30⁺ HRSC. HRSC exhibited extensive PD-L1 expression as expected. In addition, we found an enrichment of PD-L1⁺ CD11c⁺/CD68⁺ and CD11c⁺/CD68⁻ MNPs in CD30⁺ HRSC-centric neighborhoods (Fig. 3b-c). Similarly, we found variable coexpression of the co-inhibitory receptor TIM-3, and IDO1 - an immunomodulating tryptophan catabolizing enzyme - on CD11c⁺/CD68⁺ and CD11c⁺/CD68⁻ MNPs in close proximity to HRSC (Fig. 3c-d). The proportion of CD11c⁺/CD68⁻ and CD11c⁺/CD68⁺ MNPs expressing inhibitory molecules increased with age at diagnosis (p <0.05)

(Extended data Fig. 5a), suggesting immunosuppressive signalling in paediatric and young-adult cHL may differ compared with cHL in older adults. We noted no significant differences in the proportion of MNPs expressing inhibitory molecules according to sex, tumor type, or EBV status (data not shown).

HRSC have the remarkable ability to cohere a tumor mass, mainly composed of potential anti-tumor cells, by establishing an immunosuppressive microenvironment. Recently, unbiased scRNAseq profiling of cHL tumors has demonstrated enrichment of exhausted LAG3⁺ CD4⁺ T-cells around MHCII⁻ HRSC¹¹, and elaboration of CXCL13 by Tfh-like cells in the TME¹³. Our findings reveal a network of immunosuppressive MNPs closely associated with PD-L1⁺ HRSCs. These MNPs cooperate with HRSC-derived and NF-κB-directed chemokine expression, providing chemoattractant and inhibitory signals that recruit and instruct immunosuppressive T cells. Our analysis identifies classical monocytes as important signalling hubs, controlling retention of cDC2 and ThExh via CCR1-, CCR4-, CCR5-, and CXCR3-dependent signalling (Fig 3e).

Previous work has highlighted an association between myeloid cell infiltration and outcome in cHL; increased proportions of CD68⁺ TAM by immunohistochemistry, and an enhanced MΦ-associated gene expression signature correlates with inferior clinical outcomes in patients with advanced cHL^{23,24}. The positioning of these cells close to HRSC suggests that they have proximal roles in directing and amplifying the tolerogenic niche. Underlining the importance of the HRSC-associated MNP network, recent reports have demonstrated that eradication of HRSC, following first line anti-PD-1 immune checkpoint blockade, is associated with a reduction in PD-L1⁺ TAM, rather than a detectable cytotoxic response²⁵.

The substantial co-expression of inhibitory molecules and chemokines on HRSC-associated MNPs indicates functional redundancy in immunosuppressive signalling. This may partially explain resistance to targeting of individual components (e.g. PD-1), and offer opportunities for tailoring therapies on the basis of molecular profiling.

Methods

Tissue samples for scRNAseq

Healthy lymph nodes, unaffected by neoplastic disease, were acquired from donation after circulatory death (DCD) adult donors at the time of removal of organs for transplantation by the Cambridge Biorepository for Translational Medicine (CBTM) with ethical approval (reference 15/EE/0152, East of England—Cambridge South Research Ethics Committee) and consent from donor families. Lymph node samples were collected from inguinal, mesenteric, and thoracic regions at the end of the organ donation procedure, within 1-2 hours of cessation of circulation under cold ischaemic conditions. Lymph node specimens were maintained in ice-cold 0.9% saline for transport. Tissue was dissociated and processed as previously described²⁶. Lymph node data from (290b, 298c, 302c) are shared with James *et al.* and were processed with additional flow sorting as described²⁶. Patients who donated diseased tissue for

scRNAseq were enrolled in the 'Investigating how childhood tumors and congenital disease develop' study (NHS National Research Ethics Service reference 16/EE/0394).

Tissue samples for Nanostring and Multiplexed Immunofluorescence

The study was approved by a UK NHS Health Research Authority (HRA) Research Tissue Bank (CEPA Biobank, Newcastle upon Tyne Hospitals NHS Foundation Trust, reference number: 17/NE0070). The Research Tissue Bank released link-anonymized, formalin-fixed and paraffin embedded (FFPE) tissue, from patients with Classic Hodgkin lymphoma. All tissue samples were surplus to diagnostic requirements and were released in accordance with the terms of the ethical approval (ref: 17/NE0070).

Single cell RNA sequencing analysis

Single-cell RNA sequencing data from Hodgkin lymphoma specimens was acquired by direct data transfer from the Stiedl lab (University of British Columbia) as a raw counts matrix¹¹.

Mapping and quantification of scRNAseq generated for this study (Wellcome Sanger Institute (WSI) dataset) was performed using the kallisto-bustools toolkit (0.24.4)²⁷, against the GRCh38 human reference. The unfiltered count matrix was profiled using the *defaultDrops* function in the DropletUtils package (1.10.3)²⁸, in R (4.0.4).

Quality control of all cell-containing droplets was performed using the scanpy python package (version 1.7.1)²⁹. Across the concatenated dataset (WSI dataset and Aoki *et al.* data), cells with fewer than 500 distinct transcripts, fewer than 600 total counts, or greater than 10% counts originating from mitochondrial genes were filtered (Extended Data Fig. 1b). Genes with fewer than 10 counts across the dataset were filtered. The doublet detection tool scrublet³⁰ was run on a per-channel basis with default settings, before calculating the proportion of doublet calls per cluster in an initial high-resolution per-channel clustering solution (Extended Data Fig. 1c). Highly variable genes were called using the scanpy toolkit setting *n_top_genes* = 3000, *flavor*="seurat_v3". We performed data integration with batch correction (using donor as the batch key) and dimensionality reduction concurrently by training a single-cell variational inference model using the scvi-tools python package (version 0.9.0a0)¹². A 15-dimensional latent representation was used as input to nearest neighbor graph construction, Uniform Manifold and Approximation (UMAP), and Leiden clustering in scanpy.

Marker genes were calculated using the tf-idf metric and genes were ranked by the tf-idf metric, and p values calculated by a hypergeometric test, corrected for multiple testing (Benjamini-Hochberg method)³¹. These marker genes were used as a guide to manual annotation of cell types.

Cell types were annotated according to canonical marker expression (Extended Data Fig. 1d), supplemented with labelling using the celltypist python package (0.1.15)³². Concordance between annotated celltype labels and predicted reference annotations was assessed by calculating the Jaccard distance between cell label strings (Extended Data Fig. 1e-f).

Cell type predictions between the King *et al.* B cell atlas³³ was calculated by training a multilayer perceptron classifier on the reference data and calculating predictions on unseen test data, implemented in the scikit learn package (0.24.2)³⁴.

Relative abundance analysis was performed using the MiloR package (0.99.1)¹⁸. The annotated dataset was represented as a k=30 nearest neighbor graph calculated in scVI latent space. This graph representation was partitioned into 22,573 overlapping neighborhoods, and cell counts per 10X channel were calculated per neighborhood. Differential abundance between disease and health was calculated using the *testNhoods* function, setting the design as `~centre + disease`, and `norm.method="TMM"`.

Ligand receptor analysis was performed using CellPhoneDB (version 2.1.7)²². The count matrix was normalized to 1e4 total counts before input to the python implementation of CellPhoneDB which was run with `-iterations=1000`, `-threshold=0.1`, and `-result-precision=3`.

Microarray data analysis

The GEO accession GSE39133 was downloaded using the GEOquery R package (version 2.60.0)³⁵. We performed differential expression between GC and HRSC samples using limma (version 3.48.1)³⁶ and derived genesets by retaining genes with an absolute log fold change >3 and adjusted p value <0.01.

Transcription factor regulon and network analysis

Transcription factor regulons data were acquired using the dorothea R package (version 1.3.3)³⁷. Regulons with confidence levels "A", "B", and "C" were used to calculate transcription factor activities using the *run_viper* command.

We derived a network graph of transcription factors and targets using data from the OmnipathR R package (version 3.12)³⁸, plotting edges between transcription factors and targets, and between targets using the ggraph R package (version 2.0.5).

Nanostring spatial transcriptomics analysis

Ten of the FFPE cHL tissue biopsies used in the multiplexed IF analysis were selected for *in situ* transcriptomics (Extended Data Table 1), using the Nanostring GeoMx platform (Nanostring, Seattle, WA). Biopsies were selected on the basis of year of fixation (2014 onwards), high quality multiplex staining, and successful RNAscope quality control (QC) tests (Advanced Cell Diagnostics, Inc., Newark, CA). Five test slides were prepared, each arrayed with 2 unique, whole tissue biopsy sections; one slide also included reactive lymph node tissue. For each test slide, 4 consecutive 4-µm-thick sections were taken from each tissue block, prepared according to the manufacturer's instructions, and numbered 1-4: Slide 1, Hematoxylin & Eosin counterstain; Slide 2, CD45 and PAX5 immunostains (fluorescent IHC); Slide 3, *CD274* (PD-L1) and *PDCD1* (PD-1) - RNAscope probes, DNA counterstain; Slide 4, Oligonucleotide RNA probe hybridization.

Circular regions of interest (ROIs; uniform 300- μ m diameter) were labeled upon the *CD274/PDCD1* RNAscope slides, with reference to H&E and IHC slides for histomorphological context. PD-L1 was selected as the *CD274* probe fluorescent labelling was more robust than identifying HRSC cells by CD30 probe in initial experiments. In all selected tumors, the HRSC-dense areas of tissue were known to be PD-L1+ and HRSC were confirmed by DAPI signal and nuclear morphology. For each tumor, 3 ROIs were selected in 'PD-L1^{high}' areas and 3 ROIs in 'PD-L1^{low}' areas of tissue, unless otherwise stated. Assessments of PD-L1 intensity were made by visual inspection of normalized *CD274* probe fluorescent signal within each tumor. Sclerotic bands, germinal centers, and areas of low cellularity were avoided in cHL tissues, by referring to DNA counterstaining and paired IHC. In the control RN tissue, 3 ROIs were labelled in germinal centers and 3 ROIs in interfollicular areas.

RNA oligonucleotides from the 'Cancer Transcriptome Atlas' (n = 1812 probes), coupled to photocleavable oligonucleotide tags ('barcodes'), were hybridized overnight. The following day, hybridized barcodes were released from tissue by UV exposure from each ROI sequentially, before being counted and sequenced on the Illumina HiSeq 2500 platform.

Nanostring data analysis

Quality control steps were performed on exported data to exclude outlier probes and then data were normalized against the 75th percentile of signal from each tumor. Following data QC, 1372 (75.7%) probes were retained for analysis.

The nanostring expression matrix (ROI-by-gene) was analysed by principal component analysis (PCA) using the scanpy python package, before constructing a shared-nearest neighbor graph representation. Briefly, we identified k=10 nearest neighbors in PCA space using the FNN R package, before calculating the Jaccard distance between neighbor vectors after the approach of Levine *et al*³⁹. The resulting graph was clustered using the Leiden algorithm⁴⁰ with default settings.

Deconvolution was performed by first generating a matrix of one-hot-encodings of marker genes per cell type, including HRSC. We then used the *rlm* function in the MASS R package to perform robust linear regression. The resulting coefficients were treated as fractions to derive counts by multiplying by the number of nuclei in each ROI. These cell count estimates were then used to calculate differential abundance estimates for each cluster using DESeq2⁴¹.

Differential expression analysis between clusters was performed using limma (version 3.48.1)³⁶. The top 10 differentially expressed genes per cluster were selected for plotting (Extended Data Fig. 3c)

Tissue samples for multiplexed immunofluorescence

Formalin-fixed, paraffin embedded (FFPE) tumor biopsies (all excised lymph nodes) were selected from the pathology archives of Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK, by members of the Novopath (formerly CEPA) biobank, with appropriate ethical approval (reference

number: 17/NE0070). Limited, link-anonymized clinical data for study patients were collected by the Novopath biobank. Hematoxylin and eosin-stained (H&E) tissue sections and immunohistochemical tissue sections were reviewed by a Haematologist, together with the original pathology reports. Fifty-four cases were selected for the study, including biopsies from patients of all ages and histomorphological subtypes (Extended Data Table 1). Patient biopsies were selected on the basis of high-quality, whole lymph node resection biopsy tissue. Patients with coexisting malignant diagnoses were excluded from the study, even if these were not apparent in the tissue.

Multiplexed immunofluorescence

Multiplexed immunofluorescence (IF) was performed by sequentially immunostaining 4- μ m-thick tissue sections from selected FFPE tumor biopsies, with primary antibodies, secondary reagents, and unique fluorochromes. Each of the two multiplexes comprises 6 primary antibodies, followed by a nuclear counterstain, 4',6-diamidino-2-phenylindole, as per published protocols^{10,42,43}. All immunostaining was performed on the automated Ventana Benchmark platform (Roche Diagnostics, Rotkreuz, Switzerland). Slides were then air dried, mounted with Prolong Diamond Anti-fade mounting medium (#P36965; Life Technologies) and stored at 4°C in a light proof box until image acquisition. The target antigens, antibody clones, and dilutions for markers included in this report are listed in Extended Data Table 2.

Image acquisition

Test regions for multiplex IF analysis were initially identified for each tumor biopsy by reviewing the H&E and immunofluorescence tissue sections at 4x magnification. Three non-overlapping regions of interest were then acquired at 20x magnification for each multiplex panel. Each of these 3 test regions comprised 9 contiguous fields of view (FOVs), arranged as a 3x3 rectangular grid (measuring 2008 μ m x 1508 μ m in total). Therefore, there were a total of 27 FOVs per tumor, for each multiplex.

Regions were selected to exclude tissue artefacts, best represent the overall tissue, and to include CD30+ HRSCs. Regions were imaged and deconvoluted using the Vectra multispectral imaging platform (Vectra 3 and Inform 2.4.8, Akoya Biosciences, Marlborough, MA), using specific spectral libraries.

Computational image analysis

Cell detection and phenotyping: cell nuclei were segmented from the DAPI channel using inForm (version 2.4.8, Akoya Biosciences, Marlborough, MA; last accessed October 2021) with a fixed size cytoplasm region of $\leq 3\mu$ m grown around each cell nuclei. Cells were phenotyped based on the combinations of functional markers expressed using the phenotyping tool provided by inForm. Additional signals were manually thresholded for each case (PD-L1, TIM-3, IDO1, LAMP3).

Visualization of spatial density distributions: to visualize the spatial distribution of each phenotype, we applied a kernel density estimate to the spatial coordinates of each cell phenotype and rendered using contour plots. The kernel bandwidth set for each phenotype using Scott's method⁴⁴.

HRSC Neighborhood Analysis: To examine the immediate microenvironment surrounding HRSC cells, we determined a HRSC-niche neighborhood based on identifying cells whose centroid lay within a 25 μ m distance from a HRSC centroid to represent the microenvironmental niche (Extended data Fig. 4a). We apply a minimum distance of 10 μ m to minimise the effects of stain spillover between neighboring cells⁴⁵. We calculate the cell proportions of each cell phenotype within the HRSC-niche and for the entire image as a basis for comparison. To establish whether there are statistically significant differences across the dataset, Q-Q plots were used to identify the distributions as non-normal, and Wilcoxon signed-rank test applied to calculate statistical significance (Extended data Fig. 4g).

Nearest neighbor network analysis: We examined whether each cell phenotype tended to colocalize with or avoid each other cell phenotype. For a query phenotype, A, and a distant phenotype, B, we matched each cell of phenotype A to its nearest cell of type B and calculated the distance between each pair of cells (the nearest-neighbor distance). To establish a baseline distance which is independent of the density of an individual phenotype, we also randomly permuted the cell labels of type phenotype B so that we could estimate expected distance if the spatial local of cells of phenotype B were randomly distributed within the sample. We report the ratio of the observed nearest-neighbor distance to the expected nearest-neighbor distance for each cell to identify the distribution (Extended data Fig 4h). The median of this distribution is used as a summary statistic to visualise the pairwise relationships between all phenotypes, with a median nearest-neighbor distance ratio less than 1 indicating that cells tend to be placed closer together in the sample compared to a random distribution, and values greater than 1 indicating they tend to be further away (Extended Data Fig. 4h).

Declarations

Acknowledgements

BJS is supported by a Clinical Doctoral Fellowship from the Wellcome Trust (216366/Z/19/Z).

CDC is supported by grant funding from the JGW Patterson Foundation (<https://jgwpattersonfoundation.co.uk>), Bright Red (<https://brightred.org.uk>), the Lymphoma Research Trust, a Medical Research Council 'Confidence in Concepts' award, and a Wellcome Trust Small project grant award.

We thank Dr Christian Stiedl (The University of British Columbia) for generous sharing of scRNAseq data. We thank the Newcastle University / Newcastle upon Tyne NHS Foundation Trust Molecular diagnostics laboratory (NovoPath; <https://www.novopath.co.uk/>) for tissue handling and technical assistance.

Conflict of interest

None declared

Author contributions

BJS performed data analysis and wrote the manuscript, MF performed data analysis and contributed to the manuscript, MDY contributed to data curation, CJ, AS, AB, CB, VR, JRF performed scRNAseq experiments, KTM and KSP acquired deceased donor lymph node tissue, MRC and SB supervised the project, CDC performed microscopy experiments, analysed data, wrote the manuscript, and supervised the project.

Code availability

Code used to analyse data in the project is uploaded to Github at https://github.com/bjstewart1/hodgkin_lymphoma_analysis

Data availability

Data is made available to download on this website as an annotated data object (.h5ad file) and interactive cellxgene portal.

References

1. Sasse, S. *et al.* Long-Term Follow-Up of Contemporary Treatment in Early-Stage Hodgkin Lymphoma: Updated Analyses of the German Hodgkin Study Group HD7, HD8, HD10, and HD11 Trials. *J. Clin. Oncol.* **35**, 1999–2007 (2017).
2. Borchmann, P. *et al.* PET-guided treatment in patients with advanced-stage Hodgkin's lymphoma (HD18): final results of an open-label, international, randomised phase 3 trial by the German Hodgkin Study Group. *Lancet* **390**, 2790–2802 (2017).
3. Casasnovas, R.-O. *et al.* PET-adapted treatment for newly diagnosed advanced Hodgkin lymphoma (AHL2011): a randomised, multicentre, non-inferiority, phase 3 study. *Lancet Oncol.* **20**, 202–215 (2019).
4. Connors, J. M. *et al.* Hodgkin lymphoma. *Nat Rev Dis Primers* **6**, 61 (2020).
5. Roemer, M. G. M. *et al.* Major Histocompatibility Complex Class II and Programmed Death Ligand 1 Expression Predict Outcome After Programmed Death 1 Blockade in Classic Hodgkin Lymphoma. *J. Clin. Oncol.* **36**, 942–950 (2018).
6. Green, M. R. *et al.* Integrative analysis reveals selective 9p24.1 amplification, increased PD-1 ligand expression, and further induction via JAK2 in nodular sclerosing Hodgkin lymphoma and primary mediastinal large B-cell lymphoma. *Blood* **116**, 3268–3277 (2010).
7. Cader, F. Z. *et al.* A peripheral immune signature of responsiveness to PD-1 blockade in patients with classical Hodgkin lymphoma. *Nat. Med.* **26**, 1468–1479 (2020).
8. Armand, P. *et al.* Nivolumab for Relapsed/Refractory Classic Hodgkin Lymphoma After Failure of Autologous Hematopoietic Cell Transplantation: Extended Follow-Up of the Multicohort Single-Arm Phase II CheckMate 205 Trial. *J. Clin. Oncol.* **36**, 1428–1439 (2018).

9. Ansell, S. M. *et al.* PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. *N. Engl. J. Med.* **372**, 311–319 (2015).
10. Carey, C. D. *et al.* Topological analysis reveals a PD-L1-associated microenvironmental niche for Reed-Sternberg cells in Hodgkin lymphoma. *Blood* **130**, 2420–2430 (2017).
11. Aoki, T. *et al.* Single-Cell Transcriptome Analysis Reveals Disease-Defining T-cell Subsets in the Tumor Microenvironment of Classic Hodgkin Lymphoma. *Cancer Discov.* **10**, 406–421 (2020).
12. Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. & Yosef, N. Deep generative modeling for single-cell transcriptomics. *Nat. Methods* **15**, 1053–1058 (2018).
13. Aoki, T. *et al.* Single-cell profiling reveals the importance of CXCL13/CXCR5 axis biology in lymphocyte-rich classic Hodgkin lymphoma. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).
14. Villani, A.-C. *et al.* Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science* **356**, (2017).
15. Park, J.-E. *et al.* A cell atlas of human thymic development defines T cell repertoire formation. *Science* **367**, (2020).
16. Madisson, E. *et al.* scRNA-seq assessment of the human lung, spleen, and esophagus tissue stability after cold preservation. *Genome Biol.* **21**, 1 (2019).
17. Maier, B. *et al.* A conserved dendritic-cell regulatory program limits antitumour immunity. *Nature* **580**, 257–262 (2020).
18. Dann, E., Henderson, N. C., Teichmann, S. A., Morgan, M. D. & Marioni, J. C. Differential abundance testing on single-cell data using k-nearest neighbor graphs. *Nat. Biotechnol.* (2021) doi:10.1038/s41587-021-01033-z.
19. Steidl, C. *et al.* Gene expression profiling of microdissected Hodgkin Reed-Sternberg cells correlates with treatment outcome in classical Hodgkin lymphoma. *Blood* **120**, 3530–3540 (2012).
20. Nagel, D., Vincendeau, M., Eitelhuber, A. C. & Krappmann, D. Mechanisms and consequences of constitutive NF- κ B activation in B-cell lymphoid malignancies. *Oncogene* **33**, 5655–5665 (2014).
21. International Agency for Research on Cancer & World Health Organization. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues.* (World Health Organization, 2008).
22. Efremova, M., Vento-Tormo, M., Teichmann, S. A. & Vento-Tormo, R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat. Protoc.* **15**, 1484–1506 (2020).
23. Steidl, C. *et al.* Tumor-associated macrophages and survival in classic Hodgkin's lymphoma. *N. Engl. J. Med.* **362**, 875–885 (2010).
24. Tan, K. L. *et al.* Tumor-associated macrophages predict inferior outcomes in classic Hodgkin lymphoma: a correlative study from the E2496 Intergroup trial. *Blood* **120**, 3280–3287 (2012).
25. Reinke, S. *et al.* Tumor and microenvironment response but no cytotoxic T-cell activation in classic Hodgkin lymphoma treated with anti-PD1. *Blood* vol. 136 2851–2863 (2020).

26. James, K. R. *et al.* Distinct microbial and immune niches of the human colon. *Nat. Immunol.* **21**, 343–353 (2020).
27. Melsted, P., Ntranos, V. & Pachter, L. The barcode, UMI, set format and BUStools. *Bioinformatics* **35**, 4472–4473 (2019).
28. Lun, A. T. L. *et al.* EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol.* **20**, 63 (2019).
29. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).
30. Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. *Cell Syst* **8**, 281–291.e9 (2019).
31. Young, M. D. *et al.* Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science* **361**, 594–599 (2018).
32. Domínguez, C. C. *et al.* Cross-tissue immune cell analysis reveals tissue-specific adaptations and clonal architecture across the human body. doi:10.1101/2021.04.28.441762.
33. King, H. W. *et al.* Single-cell analysis of human B cell maturation predicts how antibody class switching shapes selection dynamics. *Science Immunology* vol. 6 eabe6291 (2021).
34. Garreta, R. & Moncecchi, G. *Learning scikit-learn: Machine Learning in Python.* (Packt Publishing Ltd, 2013).
35. Davis, S. & Meltzer, P. S. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics* **23**, 1846–1847 (2007).
36. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
37. Garcia-Alonso, L., Holland, C. H., Ibrahim, M. M., Turei, D. & Saez-Rodriguez, J. Benchmark and integration of resources for the estimation of human transcription factor activities. *Genome Res.* **29**, 1363–1375 (2019).
38. Türei, D., Korcsmáros, T. & Saez-Rodriguez, J. OmniPath: guidelines and gateway for literature-curated signaling pathway resources. *Nat. Methods* **13**, 966–967 (2016).
39. Levine, J. H. *et al.* Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell* **162**, 184–197 (2015).
40. Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).
41. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
42. Feng, Z. *et al.* Multispectral imaging of formalin-fixed tissue predicts ability to generate tumor-infiltrating lymphocytes from melanoma. *J Immunother Cancer* **3**, 47 (2015).
43. Tóth, Z. E. & Mezey, E. Simultaneous visualization of multiple antigens with tyramide signal amplification using antibodies from the same species. *J. Histochem. Cytochem.* **55**, 545–554

(2007).

44. Scott, D. W. *Multivariate Density Estimation: Theory, Practice, and Visualization*. (John Wiley & Sons, 1992).
45. Bai, Y. *et al.* Adjacent Cell Marker Lateral Spillover Compensation and Reinforcement for Multiplexed Images. *Front. Immunol.* **12**, 652631 (2021).

Figures

Figure 1

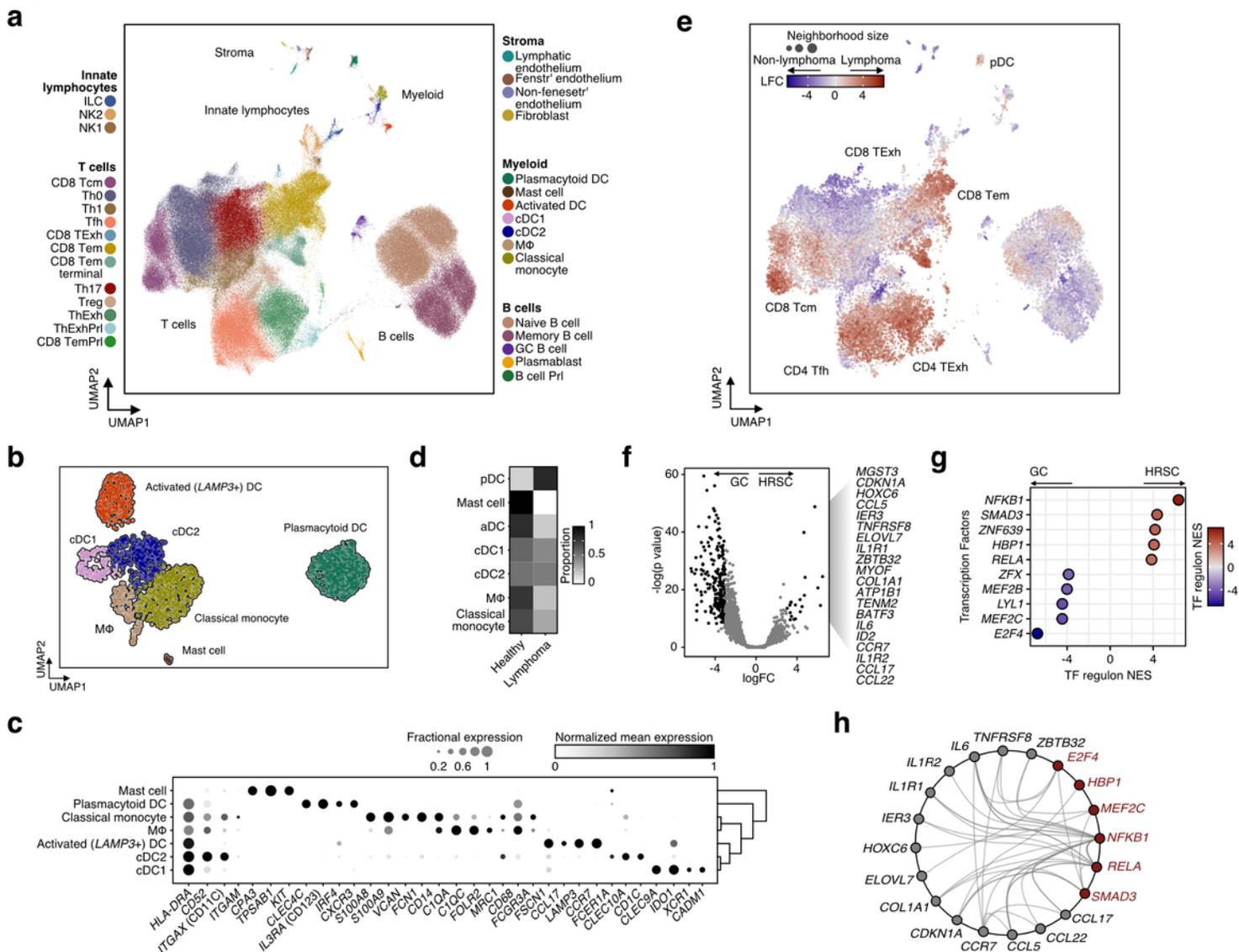


Figure 1

Transcriptional profiling of cell types in the cHL microenvironment. **a.** UMAP plot of 243,753 cells from an integrated scRNAseq dataset colored by cell type and organized by compartment (T cells, B cells, innate lymphocytes, myeloid, and stroma). **b.** UMAP plot of 2727 myeloid cells colored by cell type. **c.** Heatmap showing mean normalized expression levels (color) and fraction of cells expressing (dot size) markers of myeloid cell subsets. **d.** Heatmap showing the proportion of each cell type derived from healthy and

lymphoma samples. **e.** UMAP plot of 22,573 graph neighborhoods, colored by the differential abundance (log fold change) in lymphoma affected or non lymphoma affected (deceased-donor lymph nodes or reactive lymph node) samples. Dot size is proportional to neighborhood size (median neighborhood size = 50 cells). **f.** Volcano plot showing differentially expressed genes between microdissected GC and HRSC. Significant genes for HRSC are indicated ($|\text{LFC}| > 3$, adjusted p value < 0.01). **g.** Heatmap showing transcription factor (TF) regulon normalised enrichment scores for GC vs HRSC calculated with DoRothEA. **h.** Graph indicating interactions between genes (red, transcription factors; gray, other genes) differentially expressed in HRSC. Edges drawn where an interaction is documented by OmniPath.

Figure 2

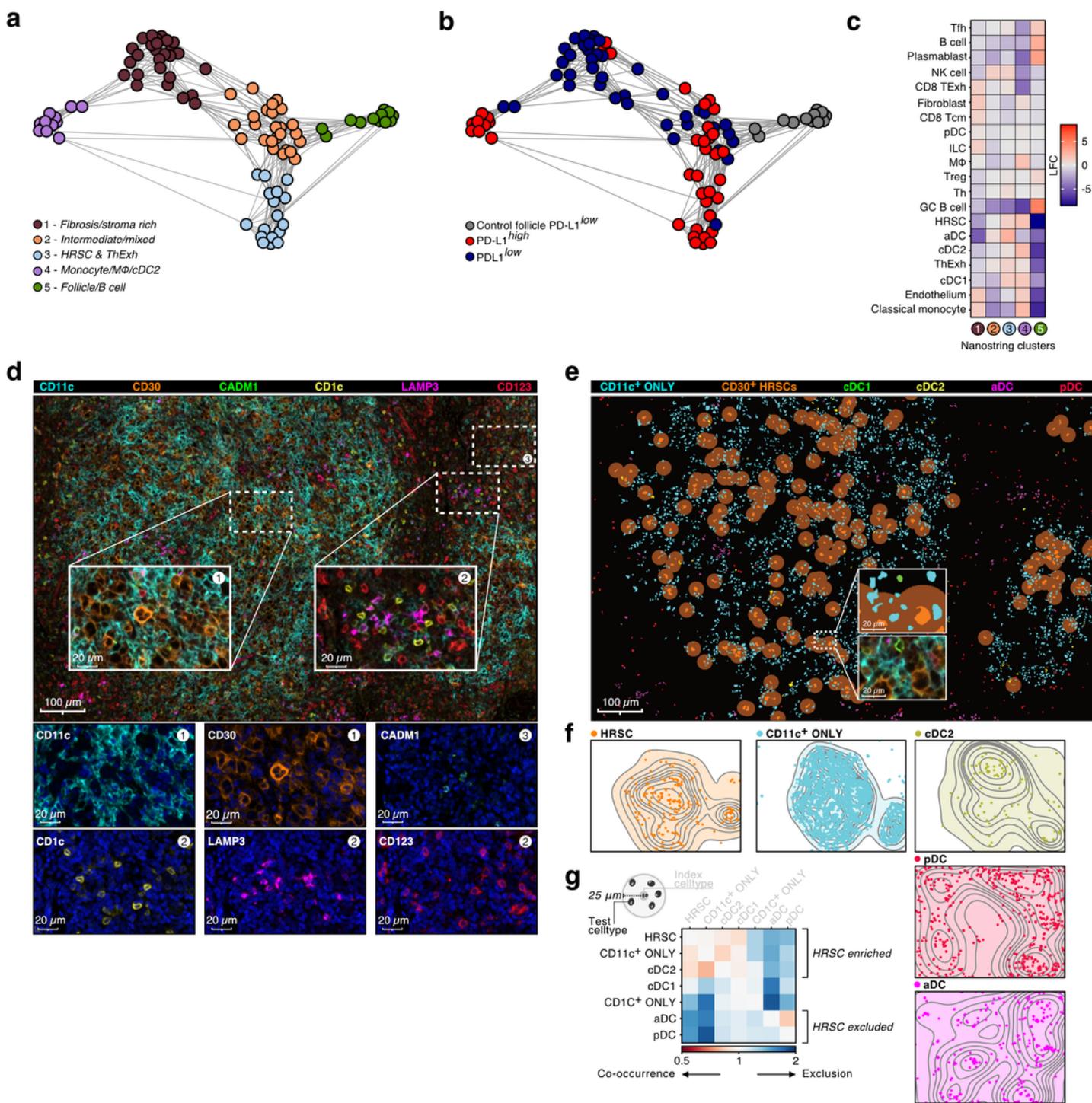


Figure 2

Mononuclear phagocyte enrichments in the Hodgkin Reed Sternberg cell microenvironment. a. Shared nearest-neighbor graph embedding of transcriptional profiles of nanostring ROI from reactive and cHL lymph nodes. Color indicates clusters identified using Leiden clustering with resolution = 1. **b.** Shared nearest-neighbor graph embedding of transcriptional profiles of nanostring ROI colored by PD-L1 expression status. **c.** Differential abundance estimates of cell-types deconvolved from nanostring ROI

between clusters shown in **(a)**, LFC estimates represent differential abundance of cell-types in each cluster vs all other clusters. **d.** Representative micrographs from multiplexed IF imaging of cHL samples. One of 3 test regions from tumor CHL-27 is shown, cropped to 1662 μm by 1076 μm ; CD11c (cyan), CD30 (orange), CADM1 (green), CD1c (yellow), LAMP3 (magenta), and CD123 (red) fluorescent signals are represented by unique pseudocolors. To improve clarity DAPI is not shown. Selected areas are highlighted, each at 100x magnification (160 μm x 108 μm); CD30+ HRSC-dense area (region 1, inset left), internodular area with no CD30+ HRSC (region 2, inset right), and area with CADM1+ cDC1 (region 3, inset top right). Individual pseudocolors are shown below the main image (second & third lines), with DAPI (blue) to identify cell nuclei. Each individual image refers to a multiplexed area above (inset), indicated by the corresponding number (top right). **e.** HRSC and DC map, corresponding to the same region from **(d)**. Phenotyped cells are identified by colored nuclei, with each cell type represented by a different color; 'CD11c+ ONLY' (CD11c+ CD1c-, cyan), CD30+ HRSC (light orange), cDC1 (green), cDC2 (yellow), LAMP3+ aDC (magenta), and pDC (red) are shown. Cells with no assigned phenotype ('null') and 'CD1c+ ONLY' cells are excluded from this visualization. The 'HRSC neighborhood' is shown as circles surrounding each CD30+ HRSC (dark orange). A selected area is highlighted, at 200x magnification (80 μm x 54 μm ; inset, center); **f.** Isobar plots show the location and density of each cell phenotype of interest, for the corresponding tumor region from **(d)** and **(e)** (note the uncropped 2008 μm by 1502 μm region is shown here). HRSCs (orange dots, left) and DC subsets (colors corresponding to each phenotype, right) are shown. **g.** Summarized interaction plot across all study tumors in aggregate, displaying the ratio of the nearest neighbor distance between phenotype pairs (index cell type (gray) to test cell type (black)) compared to the expected baseline distance. Red indicates cell type co-occurrence, blue indicates cell type exclusion.

expression of inhibitory molecules (IDO1, PD-L1, & TIM-3), cell-type defining markers for MNPs (CD11c & CD68), and HRSC (CD30). One of 3 test regions from CHL-7 is shown (2008 μm x 1502 μm), with CD11c (cyan), CD30 (orange), CD68 (magenta), PD-L1 (green), IDO1 (yellow), and TIM-3 (red) fluorescent signal represented by unique pseudocolors (to improve clarity DAPI is not shown). Inset 1 (100x magnification [160 μm x 108 μm]; left) shows CD30 and IDO1 (top), CD30 and PD-L1 (middle), and CD30 and TIM-3 (bottom), with DAPI (blue) to identify cell nuclei. Inset 2 (100x magnification [160 μm x 108 μm], right) shows representative CD30+ HRSC, surrounded by mononuclear phagocytes in closer detail. Individual pseudocolors from area 2, corresponding to each antibody, are shown below the main image (second & third lines). **c.** Proportion of mononuclear phagocytes expressing immune checkpoints and IDO1. Left panel: lower key represents the combinations of inhibitory molecules (black, expressed; white, unexpressed) corresponding to the proportions shown in boxplots, individual points correspond to cHL cases. Right upper panel: proportions of each cell type expressing, none, one, two or three inhibitory molecules in aggregate across all images. Right lower panel: Proportion of single or double positive cells expressing each indicated combination of inhibitory molecules. **d.** Expression of PD-L1, IDO1, and TIM-3 by MNPs. Representative images at 200x magnification (54 μm x 54 μm) show PD-L1 expression (green, top line) and corresponding CD11c (cyan) and CD68 (magenta, second line) for the same areas. Images are separated to improve clarity, as co-localized. IDO1 expression (yellow, cytoplasmic localization) with CD11c and CD68 (third line). TIM-3 expression (red, cytoplasmic localization) with CD11c and CD68 (bottom line). **e.** Model for intercellular interactions between MNP, dysfunctional T cells, and HRSC in the cHL TME. HRSC produce CCL22, CCL17, and CCL5 downstream of NF- κ B activity, orchestrating T-cell recruitment via indicated ligand receptor interactions. CCL3, CCL4, and CXCL10 produced by monocytes directs positioning of dysfunctional T-cells and cDC2. Widespread inhibitory molecule expression is seen on HRSC, MNP, and T cells.

Supplementary Files

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

- [ncfigS1.pdf](#)
- [ncfigS2.pdf](#)
- [ncfigS3.pdf](#)
- [ncfigS4.pdf](#)
- [ncfigS5.pdf](#)
- [ncfigS6.pdf](#)
- [ExtendedDataTables.docx](#)