Corresponding author(s): Soumya Raychaudhuri
Soumya@broadinstitute.org

Last updated by author(s): Feb 13, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	We used R version 3.5.1. The source code repository is located at https://github.com/immunogenomics/amp_phase1_ra.
Data analysis	We used R version 3.5.1. The source code repository is located at https://github.com/immunogenomics/amp_phase1_ra.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data presented in the manuscript are available through NIH IMMPORT (accession: SDY998 and SDY999) and and dbGAP (study accession: phs001457.v1.p1).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size was calculated. This study represents a "feasibility" study where single cell analyses were applied to a cohort of patients. Sample size was determined based on the total of number of patients recruited during the time period over phase 1 of this study. This study is a proof of principle, to demonstrate that single cell analyses can be applied to samples taken from a large cohort of patients from multiple research sites. Since the goal of the study was to test the feasibility of applying high-dimensional analysis, the total number of patients recruited here was considered sufficient for the sample size.
Data exclusions	Data were excluded from analyses based on specific quality control criteria as described in detail in the manuscript for each data sets. For synovial tissues that did not pass standard histologic QC (i.e. lack of identifiable lining structure) were excluded from main pipeline analysis. For single cell data, we discarded cells with fewer than 1,000 genes detected with at least one fragment. We also discarded cells that had more than 25% of molecules coming from mitochondrial genes. For bulk RNA-seq experiments, samples with low quality as determined by gene reads were excluded from subsequent analysis.
Replication	No experimental replication were performed in this study due to the nature of the study design
Randomization	No randomization was performed due to the cross-sectional nature of the study
Blinding	No blinding was performed in this study due to the cross-sectional nature of the study

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Antibodies used for flow cytometry and cell sorting:
antibody clone vendor catalog number Dilution
anti-CD45-FITC (Biolegend, HI30) H130 Biolegend 304006 1:400
anti-CD90-PE "5E10" Biolegend 328110 1:500
anti-Pdpn-PerCP eF710 NZ-1.3 eBioscience 46-9381-42 1:50
anti-CD3-PE-Cy7 UCHT1 Biolegend 300420 1:100
anti-CD19-BV421 HIB19 Biolegend 302233 1:20
anti-CD14-BV510 M5E2 Biolegend 301842 1:100
anti-CD34-BV605-A (eBioscence, 4H11) 581 Biolegend 343529 1:400
anti-CD4-BV650 (Biolegend, RPA-T4) RPA-T4 Biolegend 300536 1:50
anti-CD8a-BV711A RPA-T8 Biolegend 301044 1:100
anti-CD31-AF700 WM59 Biolegend 303134 1:100
CD27-APC M-T271 Biolegend 356410 1:100
anti-CD235a-APC-AF750 11E4B-7-6 Beckman Coulter A89314 1:10

Antibodies used for immunofluorescent microscopy studies:
antibody clone vendor catalog number Dilution
mouse anti-human CD8 C8/144B Genetex GTX72053 1:50 (3ug/ml)
rabbit anti-human IFNg polyclonal biorbyt orb214082 1:100 (10ug/ml)
Alexa Fluor 568 donkey anti-goat Ig G N/A Thermo Fisher Scientific Cat#A-11057 1:200 (10ug/ml)
Alexa Fluor 488 donkey anti-rabbit N/A Jackson ImmunoResearch Laboratories Cat#711-546-152 1:200 (6ug/ml)

Antibodies used for mass cytometry:
antibody clone metal dilution
CD45 HI30 141Pr 1:100
CD19 HIB19 142Nd 1:100

RANKL MIH24 143Nd 1:50
CD64 10.1 144Nd 1:100
CD16 3G8 145Nd 1:100
CD8a RPA T8 146Nd 1:100
FAP Poly 147Sm 1:50
CD20 2H7 148Nd 1:100
CD45RO UCHL1 149Sm 1:100
CD38 HIT2 150Nd 1:100
CD279/PD-1 EH12.2H7 151Eu 1:100
CD14 M5E2 152Sm 1:100
CD69 FN50 153Eu 1:100
CD185/CXCR5 J252D4 154Sm 1:100
CD4 RPA T4 155Gd 1:100
Podoplanin NC-08 156Gd 1:100
CD3 UCHT1 158Gd 1:100
CD11c Bu15 159Tb 1:100
CD307d/FcRL4 413D12 160Gd 1:100
CD138 MI15 161Dy 1:100
CD90 5E10 162Dy 1:50
CCR2 K036C2 163Dy 1:100
Cadherin 11 3C10 164Dy 2:25
FoxP3 PCH101 165Ho 1:50
CD34 581 166Er 1:100
CD146/MCAM SHM-57 167Er 1:50
IgA 9H9H11 168Er 1:100
ICOS C398.4A 170Er 1:100
CD66b G10F5 171Yb 1:100
IgM MHM-88 172Yb 1:200
CD144/VE-Cadherin BV9 173Yb 1:100
HLA-DR L243 174Yb 1:100
IgD IA6-2 175Lu 1:100
CD106/VCAM-1 STA 176Yb 1:100
CD45 HI30 141Pr 1:100
CD19 HIB19 142Nd 1:100
RANKL MIH24 143Nd 1:50
CD64 10.1 144Nd 1:100
CD16 3G8 145Nd 1:100
CD8a RPA T8 146Nd 1:100
FAP Poly 147Sm 1:50
CD20 2H7 148Nd 1:100
CD45RO UCHL1 149Sm 1:100
CD38 HIT2 150Nd 1:100
CD279/PD-1 EH12.2H7 151Eu 1:100
CD14 M5E2 152Sm 1:100
CD69 FN50 153Eu 1:100
CD185/CXCR5 J252D4 154Sm 1:100
CD4 RPA T4 155Gd 1:100
Podoplanin NC-08 156Gd 1:100
CD3 UCHT1 158Gd 1:100
CD11c Bu15 159Tb 1:100
CD307d/FcRL4 413D12 160Gd 1:100
CD138 MI15 161Dy 1:100
CD90 5E10 162Dy 1:50
CCR2 K036C2 163Dy 1:100
Cadherin 11 3C10 164Dy 2:25
FoxP3 PCH101 165Ho 1:50
CD34 581 166Er 1:100
CD146/MCAM SHM-57 167Er 1:50
IgA 9H9H11 168Er 1:100
ICOS C398.4A 170Er 1:100
CD66b G10F5 171Yb 1:100
IgM MHM-88 172Yb 1:200
CD144/VE-Cadherin BV9 173Yb 1:100
HLA-DR L243 174Yb 1:100
IgD IA6-2 175Lu 1:100
CD106/VCAM-1 STA 176Yb 1:100

Validation

All commercial antibodies used for flow cytometry and cell sorting experiments were validated for flow cytometric analysis of human cells according to manufacturer's production information. Additional validation on synovial cells for cell type specificity were performed as described in Donlin and Rao et al., Methods for high-dimensonal analysis of cells dissociated from cyropreserved synovial tissue. Arthritis Res. Ther. 20, 139 (2018). For antibodies used in mass cytometry experiments, cell type specificity in synovial cells were tested and described in Donlin and Rao et al. For antibodies used in immunofluorescence microscopy experiments, all antibodies were tested for IF studies on human tissues and cells based on manufacturer's product

description.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Clinical characteristics of 51 recruited patients. OA leukocyte-poor RA leukocyte-rich RA (n=15) (n=17) (n=19) Demographic variables Age, mean 71 64.2 57.3 (Range) (64-81) (42-79) (36-71) Females, n (%) 10 (66.7) 15 (82.4) 14 (73.7) RA-related variables Mean years of disease duration 15.7 5.5* (range) (<1-51) (<1-29) RF positive, n (%) 8 (47.1) 16* (84.2) CCP positive, n (%) 10 (58.8) 14 (73.7) DMARDs Prednisone, n (%) 10 (55.6) 4* (22.2) Methotrexate, n (%) 7 (41.2) 3 (15.8) TNFi, n (%) 4 (23.5) 2 (10.5) Rituximab, n (%) 0 (0) 1 (5.3) Abatacept, n (%) 1 (5.9) 1 (5.3) Tofacitinib, n (%) 2 (11.8) 1 (5.3) DMARDs = Disease-Modifying Antirheumatic Drugs. TNFi = TNF inhibitors (infliximab, etanercept, adalimumab, Golimumab). RhF = Rheumatoid Factor. CCP = Cyclic Citrullinated Peptide. *Significant p-value between leukocyte-poor RA and leukocyte-rich RA.
Recruitment	The study was performed in accordance with protocols approved by the institutional review board. A multicenter, cross-sectional study of individuals undergoing elective surgical procedures and a prospective observational study of synovial biopsy specimens from RA patients ≥ age 18, with at least one inflamed joint, recruited from 10 contributing sites in the network. Subjects in the biopsy portion were being asked to undergo a research procedure to obtain synovial tissue.
Ethics oversight	We have been approved by all relevant ethical regulations and the study protocol. Protocols were approved by University of Rochester Medical Center, Hospital for Special Surgery, University of Pittsburgh Medical Center, University of California San Diego, University of Colorado: Denver, Northwestern University, University of Birmingham UK, Queen Mary University of London, University of Alabama Birmingham, University of Massachusetts Medical Center

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Synovial T cells, B cells, monocytes, and fibroblasts were isolated from disaggregated synovial tissue. Briefly, disaggregated synovial cells were stained with antibodies against CD45 (HI30), CD90 (5E10), podoplanin (NZ1.3), CD3 (UCHT1), CD19 (HIB19), CD14 (M5E2), CD34 (4H11), CD4 (RPA-T4), CD8 (SK1), CD31 (WM59), CD27 (M-T271), CD235a (KC16), using human TruStain FcX in 1% BSA in Hepes-Buffered Saline (HBS,20 mM HEPES, 137 mM NaCl, 3mM Kcl, 1mM CaCl2) for 30 minutes. For validation experiments, RA and OA synovial tissue were disaggregated and synovial cells were stained with cell-type specific antibody panels. For each cell subset, up to 1000 cells were collected directly into buffer TCL (Qiagen). Antibody panels used to define cell subsets are fibroblasts: CD90 (5E10), podoplanin (NZ1.3), HLA-DR (G46-6); B cell subsets: HLA-DR (G46-6), CD11c (3.9), CD19 (SJ25C1), CD27 (M-T271), IgD (IA6-2), CD3 (UCHT1), CD14 (M5E2), CD38 (HIT2); Monocyte subsets: CD14-BV421 (M5E2), CD38-APC (HB-7), and CD11c-PECy7 (B-ly6). Immediately prior to sorting, DAPI or LIVE/DEAD viability dye was added to cell suspensions and cells were passed through a 100µm filter.
Instrument	T cells (CD45+, CD3+, CD14-), monocytes (CD45+, CD3–, CD14+), B cells (CD45+, CD3–, CD14–, CD19+), and synovial fibroblasts (CD45–, CD31–, PDPN+) were collected by fluorescence-activated cell sorting (BD FACSAria Fusion)
Software	Flowjo (version 10) was used for analysis
Cell population abundance	95% purity were achieved during sorting of synovial cells based on flow cytometry analysis during single cell sorting (second sort)

Gating strategy	Synovial cells were gated based on the following schemes: T cells (CD45+, CD3+, CD14-), monocytes (CD45+, CD3–, CD14+), B cells (CD45+, CD3–, CD14–, CD19+), and synovial fibroblasts (CD45–, CD31–, PDPN+)

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.