



Technische Universität München

Fakultät für Medizin

**T cell characterization in correlation with the intestinal microbiome
and GvHD after allogeneic stem cell transplantation**

Sebastian Jarosch

Vollständiger Abdruck der von der Fakultät für Medizin der Technischen Universität
München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz: Prof. Dr. Marc Schmidt-Supprian

Prüfer*innen der Dissertation: 1. Prof. Dr. Dirk H. Busch
2. Prof. Dr. Dirk Haller
3. Prof. Dr. Barbara Stecher-Letsch

Die Dissertation wurde am 09.09.2022 bei der Technischen Universität München ein-
gereicht und durch die Fakultät für Medizin am 21.02.2023 angenommen.

Parts of this thesis have previously been published or are submitted for publication:

Sebastian Jarosch*, Jan Köhler*, Sabrina Wagner, Elvira D'Ippolito, and Dirk H. Buscht†. ChipCytometry for multiplexed detection of protein and mRNA markers on human FFPE tissue samples. *STAR Protocols* (2022)

Sebastian Jarosch*, Jan Köhler*, Rim S.J. Sarker, Katja Steiger, Klaus-Peter Janssen, Arne Christians, Christian Hennig, Ernst Holler, Elvira D'Ippolito†, Dirk H. Buscht†. Multiplexed imaging and automated signal quantification in formalin-fixed paraffin-embedded tissues by ChipCytometry. *Cell Reports Methods* (2021)

Erik Thiele Orberg*, Elisabeth Meedt*, Andreas Hiergeist, Jinling Xue, Sakhila Ghimire, Melanie Tiefgraber, Sophia Gödel, Tina Eismann, Alix Schwarz, **Sebastian Jarosch**, Katja Steiger, Michael Gigl, Karin Kleigrew, Julius C. Fischer, Klaus-Peter Janssen, Michael Quante, Simon Heidegger, Peter Herhaus, Mareike Verbeek, Jürgen Ruland, Daniela Weber, Matthias Edinger, Daniel Wolff, Dirk H. Busch, Wolfgang Herr, Florian Bassermann, André Gessner, Li Deng, Ernst Holler, Hendrik Poeck. Identification of protective, metabolite-producing bacterial and viral consortia in allogeneic stem cell transplantation patients. *submitted and available at Research Square* (2021)

The methodological developments during this thesis led to additional publications:

Anna Ralser, Alisa Dietl, **Sebastian Jarosch**, Veronika Engelsberger, Andreas Wanisch, Klaus Peter Janssen, Michael Vieth, Michael Quante, Dirk Haller, Dirk H. Busch, Li Deng, Raquel Mejías-Luquet, Markus Gerhard†. *Helicobacter pylori* promotes colorectal carcinogenesis by deregulating intestinal immunity and inducing a mucus-degrading microbiota signature. *submitted and available at MedRxiv* (2022)

Carlson Tsui*, Lorenz Kretschmer*, Svenja Rapelius*, Sarah S. Gabriel, David Chisanga, Konrad Knöpper, Daniel T. Utzschneider, Simone Nüssing, Yang Liao, Teisha Mason, Santiago Valle Torres, Stephen A. Wilcox, Krystian Kanev, **Sebastian Jarosch**, Justin Leube, Stephen L. Nutt, Dietmar Zehn, Ian A. Parish, Wolfgang Kastenmüller, Wei Shi, Veit R. Buchholz†, Axel Kallies†. c-Myb orchestrates T cell exhaustion and response to checkpoint inhibition. *Nature* (2022)

Anna Purcarea*, **Sebastian Jarosch***, Jack Barton, Simon Grassmann, Ludwig Pachmayr, Elvira D'Ippolito, Monika Hammel, Anna Hochholzer, Karolin I. Wagner, Joost H. van den Berg, Veit R. Buchholz, John B. A. G. Haanen, Dirk H. Buscht†, and Kilian Schober†. Signatures of recent activation identify a circulating T cell compartment containing tumor-specific antigen receptors with high avidity. *Science Immunology* (2022)

Karolin I. Wagner*, Laura M. Mateyka*, **Sebastian Jarosch***, Vincent Grass, Simone Weber, Kilian Schober, Monika Hammel, Teresa Burrell, Behnam Kalali, Holger Poppert, Henriette Beyer, Sophia Schambeck, Stefan Holdener, Andrea Ströges-Achatz, Verena Haselmann, Michael Neumaier, Johanna Erber, Alina Priller, Sarah Yazici, Hedwig Roggendorf, Marcus Odendahl, Torsten Tonn, Andrea Dick, Klaus Witter, Hrvoje Mijojević, Ulrike Protzer, Percy A. Knolle, Andreas Pichlmair, Claudia S. Crowell, Markus Gerhard, Elvira D'Ippolito†, and Dirk H. Buscht†. Recruitment of highly cytotoxic CD8+ T cell receptors in mild SARS-CoV-2 infection. *Cell Reports* (2022)

David S. Fischer*, Meshal Ansari*, Karolin I. Wagner*, **Sebastian Jarosch**, Yiqi Huang, Christoph H. Mayr, Maximilian Strunz, Niklas J. Lang, Elvira D'Ippolito, Monika Hammel, Laura Mateyka, Simone Weber, Lisa S. Wolff, Klaus Witter, Isis E. Fernandez, Gabriela Leuschner, Katrin Milger, Marion Frankenberger, Lorenz Nowak, Katharina Heinig-Menhard, Ina Koch, Mircea G. Stoleriu, Anne Hilgendorff, Jürgen Behr, Andreas Pichlmair, Benjamin Schubert†, Fabian J. Theis†, Dirk H. Busch†, Herbert B. Schiller† & Kilian Schober†. Single-cell RNA sequencing reveals ex vivo signatures of SARS-CoV-2-reactive T cells through 'reverse phenotyping'. *Nature Communications* (2021)

Sophie Flommersfeld*, Jan P. Böttcher*, Jonatan Ersching, Michael Flossdorf, Philippa Meiser, Ludwig O. Pachmayr, Justin Leube, Inge Hensel, **Sebastian Jarosch**, Qin Zhang, M. Zeeshan Chaudhry, Immanuel Andrae, Matthias Schiemann, Dirk.H. Busch, Luka Cicin-Sain, Joseph C. Sun, Georg Gasteiger, Gabriel D. Victora, Thomas Höfer, Veit R. Buchholz† and Simon Grassmann†. Fate mapping of single NK cells identifies a type 1 innate lymphoid-like lineage that bridges innate and adaptive recognition of viral infection. *Immunity* (2021)

Thomas R. Müller, **Sebastian Jarosch**, Monika Hammel, Justin Leube, Simon Grassmann, Bettina Bernard, Manuel Effenberger, Immanuel Andrä, M. Zeeshan Chaudhry, Theresa Käuferle, Antje Malo, Luka Cicin-Sain, Peter Steinberger, Tobias Feuchtinger, Ulrike Protzer, Kathrin Schumann, Michael Neuenhahn, Kilian Schober† and Dirk H. Busch†. Targeted T cell receptor gene editing provides predictable T cell product function for immunotherapy. *Cell Reports Medicine* (2021)

Michael Hiltensperger, Eduardo Beltrán, Ravi Kant, Sofia Tyystjärvi, Gildas Lepennetier, Helena Domínguez Moreno, Isabel J. Bauer, Simon Grassmann, **Sebastian Jarosch**, Kilian Schober, Veit R. Buchholz, Selin Kenet, Christiane Gasperi, Rupert Öllinger, Roland Rad, Andreas Muschwackh, Christopher Sie, Lilian Aly, Benjamin Knier, Garima Garg, Ali M. Afzali, Lisa Ann Gerdes, Tania Kümpfel, Sören Franzenburg, Naoto Kawakami, Bernhard Hemmer, Dirk H. Busch, Thomas Misgeld, Klaus Dornmair and Thomas Korn†. Skin and gut imprinted helper T cell subsets exhibit distinct functional phenotypes in central nervous system autoimmunity. *Nature Immunology* (2021)

* These authors contributed equally

† Senior authors

Table of Contents

Table of Contents	I
List of Tables	V
List of Figures	VI
Aberrations	VIII
Abstract	X
1. Introduction	1
1.1 Allogeneic hematopoietic stem cell transplantation (aHSCT)	1
1.2 Graft-versus-host disease (GvHD) is a major complication after aHSCT	3
1.3 GvHD Pathophysiology	4
1.3.1 Activation of APCs	4
1.3.2 Donor T cell activation	5
1.3.3 Tissue damage and inflammation	6
1.4 Regulatory T cells in GvHD.....	7
1.5 Prevention and treatment of GvHD.....	9
1.6 The influence of the microbiome on GvHD	10
1.6.1 Associations from human studies	10
1.6.2 Mechanisms by which microbiota may influence GvHD	11
1.6.3 Microbiome modulation as a treatment option for GvHD	12
1.7 Multiplexed analysis methods for single-cell characterization	14
1.7.1 Single-cell RNA sequencing	14
1.7.2 Multiplexed imaging techniques.....	17
1.7.3 ChipCytometry	19
2. Aim of this thesis	20

3. Materials and Methods	21
3.1 Materials	21
3.1.1 Devices	21
3.1.2 Consumables	22
3.1.3 Chemicals and reagents	22
3.1.4 Reagent kits	23
3.1.5 Antibodies	24
3.1.5.1 ChipCytometry	24
3.1.5.2 Flow cytometry	25
3.1.5.3 Single cell RNA sequencing	25
3.1.6 Probes for fluorescence RNA <i>in-situ</i> hybridization	26
3.1.7 Buffers and media	26
3.1.8 Human samples	27
3.1.9 Gene Sets	27
3.1.10 Software	29
3.1.11 Code	30
3.2 Methods	31
3.2.1 Generation of single-cell suspensions	31
3.2.1.1 PBMC isolation from human blood samples	31
3.2.1.2 Generation of single-cell suspension from human gastrointestinal biopsies	31
3.2.1.3 Thawing single-cell suspensions	31
3.2.2 Establishment of ChipCytometry on FFPE tissues	31
3.2.2.1 Antigen retrieval testing	31
3.2.2.2 Autofluorescence treatment tests	31
3.2.2.3 Evaluation of tissue section adherence	32
3.2.2.4 Signal-to-Noise (SNR) ratio evaluation	32
3.2.2.5 Tissue integrity scoring	32
3.2.2.6 Manual counting	32
3.2.2.7 Computer-assisted image analysis of IHC staining	33
3.2.2.8 IHC staining within a ChipCytometry chip	33
3.2.2.9 Fluorescence RNA <i>in-situ</i> hybridization (FISH)	33
3.2.3 ChipCytometry on human FFPE tissues from gastrointestinal biopsies	33
3.2.3.1 Preparation of FFPE tissues	33
3.2.3.2 Staining, imaging, and bleaching	34
3.2.3.3 Data analysis	34
3.2.3.4 Distance measurements	34

3.2.4	Flow cytometry	35
3.2.5	Single-cell RNA sequencing	35
3.2.5.1	Cell staining and sorting	35
3.2.5.2	Preparation of single-cell RNA libraries	35
3.2.5.3	Sequencing	36
3.2.5.4	Sample demultiplexing using HLA information.....	36
3.2.5.5	Data analysis.....	37
3.2.6	16S rRNA sequencing	38
4.	Results	39
4.1	Patient cohort for the investigation of GvHD	39
4.2	Method development for ChipCytometry on FFPE samples.....	40
4.2.1	Implementation of ChipCytometry for FFPE sample handling	40
4.2.2	Marker establishment.....	43
4.2.3	Fluorescence RNA <i>in-situ</i> hybridization (FISH)	44
4.2.4	Automatic quantification of multiplexed image data	46
4.2.4.1	Stitching	47
4.2.4.2	Segmentation	47
4.2.4.3	Pre-processing and spatial spillover correction.....	49
4.2.4.4	Cell quantification from ChipCytometry imaging data	51
4.2.4.5	Quantification of phenotypically complex cell populations	53
4.2.4.6	Clustering of cell populations for unbiased cell type identification	54
4.3	Method development for scRNA sequencing from gastrointestinal biopsies	56
4.3.1	Cell extraction from gastrointestinal biopsies.....	56
4.3.2	Pilot scRNA sequencing experiment.....	57
4.3.2.1	Evaluation of sample demultiplexing.....	57
4.3.2.2	Transcriptomic comparison of fresh and frozen cells.....	58
4.4	Clinical parameters of the analyzed patient cohort	59
4.5	Multiparametric analysis of GvHD biopsies	60
4.6	GvHD is mainly linked to changes in T cell frequencies	62
4.7	Antibiotic treatment and immune reconstitution contribute to T cell variance ...	64
4.8	Antibiotic treatment reduces suppressive capacity of Tregs.....	65
4.9	Certain bacterial species are linked to Treg differentiation	68

4.10	Clonally expanded CD8 T cells are drivers for GvHD severity	70
4.11	Expanded T cell clones persist over time and are anatomically spread	72
4.12	Suppressive capacity of Tregs is directly linked to CD8 expansion	74
5.	Discussion	76
5.1	ChipCytometry allows deep <i>in-situ</i> phenotyping of single cells in gut biopsies	76
5.2	A titrated protocol allows scRNA seq from gastrointestinal punch biopsies	79
5.3	GvHD is associated with increased immune and T cell infiltration.....	80
5.4	The microbiome can influence Treg numbers and phenotype	81
5.5	Antigen-specific, systemically present CD8 T cells are linked to GvHD	82
5.6	Treg functionality and infiltration density are essential for GvHD suppression .	84
6.	Summary	85
7.	Acknowledgements	86
8.	Bibliography	87

List of Tables

Table 1. List of devices.....	21
Table 2. List of consumables.....	22
Table 3. List of chemicals and reagents.....	22
Table 4. List of reagent kits	23
Table 5. List of ChipCytometry antibodies.....	24
Table 6. List of flow cytometry antibodies	25
Table 7. List of single-cell RNA sequencing antibodies	25
Table 8. List of FISH probes.....	26
Table 9. Composition of buffers and media.....	26
Table 10. Origin of human samples	27
Table 11. Gene sets used in scRNA seq analysis	27
Table 12. List of software and python packages	29

List of Figures

Figure 1. Allogeneic hematopoietic stem cell transplantation for cancer treatment.	1
Figure 2. Clinical GvHD grading.	3
Figure 3. Three phases of GvHD.	4
Figure 4. Regulatory T cell function.	7
Figure 5. Mucosal immunity in healthy and GvHD situations.	12
Figure 6. Droplet-based 5' single-cell RNA sequencing with 10X genomics.	15
Figure 7. Computational analysis of scRNA sequencing experiments.	17
Figure 8. Comparison of antibody-based, FFPE-compatible HMTI methods.	18
Figure 9. ChipCytometry technology.	19
Figure 10. Filtering of cells in the scRNA sequencing experiment.	38
Figure 11. Human aHSCT patient cohort.	39
Figure 12. Antigen retrieval titration for the analyses of FFPE samples by ChipCytometry.	40
Figure 13. Enhancement of tissue integrity by the use pre-coated coverslips.	41
Figure 14. Titrated autofluorescence quenching and section thickness enhance staining intensity. ...	42
Figure 15. Cell types detectable with the established markers in human FFPE tissues.	43
Figure 16. High-quality, single cell resolution imaging of FFPE colon tissues with ChipCytometry.	44
Figure 17. Implementation of FISH mRNA staining into the ChipCytometry workflow.	45
Figure 18. Automated analysis pipeline for highly multiplexed imaging data.	46
Figure 19. Shading correction reduces bleaching-induced artefacts at adjacent positions.	47
Figure 20. Nuclei staining-based segmentation of single cells from tissue samples.	48

Figure 21. Pre-processing of ChipCytometry images for automated quantification.	49
Figure 22. Spatial spillover correction improves the reliability of automated image quantification.....	50
Figure 23. Multiparametric imaging and signal quantification of pancreatic and colon cancer tissues.	52
Figure 24. Robust detection of rare and phenotypically complex cell populations by ChipCytometry.	53
Figure 25. Neighbourhood embedding and clustering analysis for unbiased cell classification.....	54
Figure 26. Optimization of cell extraction, freezing and sorting of single cell suspensions.....	56
Figure 27. Pooling and demultiplexing of PBMC samples in scRNA sequencing.	58
Figure 28. scRNA sequencing data from fresh and frozen PBMCs.	58
Figure 29. Human patient samples analysed in this study.	59
Figure 30. Multiparametric analysis of GvHD biopsies by scRNA seq and ChipCytometry.	61
Figure 31. GvHD is mainly linked to changes in T cell frequencies.	62
Figure 32. Analysis of T cell infiltration by automated quantification of ChipCytometry data.	63
Figure 33. Analysis of gastrointestinal biopsies from GvHD samples via ChipCytometry.....	64
Figure 34. Treg heterogeneity in the ChipCytometry dataset is correlated to antibiotic treatment.....	65
Figure 35. The phenotype of regulatory T cell is influenced by antibiotic treatment.	66
Figure 36. Regulatory T cell differentiation is locked upon antibiotic treatment.	67
Figure 37. The abundance of certain bacterial species can influence Treg differentiation.	68
Figure 38. CD8 T cells are associated with GvHD severity.....	70
Figure 39. Conventional T cells show enhanced clonal expansion in severe GvHD patients.....	71
Figure 40. Expanded CD8 T cells are anatomically distributed and stable over time.	72
Figure 41. Direct connection between Tregs and CD8 T cell expansion in severe GvHD.	74

Aberrations

AB	Antibody	FISH	Fluorescence <i>in-situ</i> hybridization
ABX	Antibiotic treatment	FITC	Fluorescein isothiocyanate
aH SCT	Allogeneic hematopoietic stem cell transplantation	FMT	Fecal microbiota transplantation
AI	Artificial intelligence	FOXP3	Forkhead box P3
AF488	Alexa Fluor 488	GATA3	GATA binding protein 3
aGvHD	acute graft-versus-host disease	G-CSF	Granulocyte colony-stimulating factor
ANOVA	Analysis of variance	GEM	Gel bead in emulsion
APC	Antigen-presenting cell	GI	Gastrointestinal
AR	Antigen retrieval	GSEA	Geneset enrichment analysis
ASV	Amplicon sequence variant	GvHD	Graft-versus-host disease
BBKNN	Batch-balanced KNN	GvL	Graft-versus-Leukemia
BCoAT	butyryl CoA:acetate CoA transferase	GZMB	Granzyme B
BCR	B cell receptor	H&E	Hematoxylin Eosin
BSA	Bovine serum albumin	HBSS	Hanks' Balanced Salt Solution
BV421	Brilliant violet 421	HDAC	Histone deacetylase
BUV395	Brilliant ultraviolet 395	HDR	High dynamic range
cGvHD	chronic graft-versus-host disease	HEPES	4-(2-hydroxyethyl)-1-piperazineethane sulfonic acid
CD	Cluster of differentiation	HIER	Heat-induced antigen retrieval
cDNA	complementary DNA	HLA	Human leukocyte antigen
CDR3	Complementarity-determining region 3	HMTI	Highly multiplexed tissue imaging
CoA	Co-enzyme A	HRP	Horseradish peroxidase
cRPMI	complete RPMI	HSD	Honestly significant difference
CTLA4	Cytotoxic T-lymphocyte-associated protein 4	IBD	Inflammatory bowel disease
CM	Central memory	IEL	Intraepithelial leukocyte
DAMP	Danger-associated molecular pattern	IFN γ	Interferon gamma
DC	Dendritic cell	Ig	Immunoglobulin
DMSO	Dimethyl sulfoxide	IL	Interleukin
DNA	Deoxyribonucleic acid	ILC	Innate lymphoid cell
DPBS	Dulbecco's phosphate-buffered saline	IMC	Imaging mass cytometry
dsDNA	double-stranded DNA	ITGAE	Integrin alpha E
EDTA	Ethylenediaminetetraacetic acid	KEGG	Kyoto Encyclopedia of Genes and Genomes
eF570	eFluor 570	KNN	K-nearest neighbors
EM	Effector memory	MAIT	Mucosa-associated invariant T
EtOH	Ethanol	MEK	Methyl ethyl ketone
FCS	Fetal calf serum	MHC	Major histocompatibility complex
FFPE	Formalin-fixed paraffin embedded	mRNA	Messenger ribonucleic acid
		MUC2	Mucin 2
		n.s.	not significant

NFκB	Nuclear factor kappa-light-chain-enhancer of activated B cells	SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
NGS	Next-generation sequencing	SB	Sodium borohydride
NK	Natural killer	SBB	Sudan Black B
PAMP	Pathogen-associated molecular pattern	SCFA	Short-chain fatty acid
PB	Photobleaching	scRNA seq	Single cell RNA sequencing
PBMC	peripheral blood mononuclear cell	SCT	Stem cell transplantation
PBS	Phosphate buffered saline	SMA	Smooth muscle actin
PBST	PBS-Tween	SNP	Single-nucleotide polymorphism
PC	Principal component	SNR	Signal-to-noise ratio
PCA	Principal component analysis	STAT3	Signal transducer and activator of transcription 3
PCR	Polymerase chain reaction	TCR	T cell receptor
PD1	Programmed cell death protein 1	Teff	Effector T cell
PD-L1	Programmed death-ligand 1	TGFβ	Transforming growth factor beta
PE	Phycoerythrin	Th	T-helper
PerCP	Peridinin-chlorophyll-protein complex	TLR	Toll-like receptor
PI	Propidium iodide	TNFα	Tumor necrosis factor alpha
PLZF	promyelocytic leukemia zinc finger	TNFSF	Tumor necrosis factor superfamily
PRF1	Perforin-1	Treg	Regulatory T cell
pSTAT3	Phosphorylated STAT3	TRA	TCR alpha
pTregs	Peripherally induced Tregs	TRB	TCR beta
qPCR	Quantitative PCR	TRIS	tris(hydroxymethyl)aminomethane
Reg3a	Regenerating islet-derived protein 3 alpha	TRM	Tissue-resident memory
RNA	Ribonucleic acid	t-SNE	t-distributed stochastic neighbor embedding
ROI	Region of interest	TSO	Template switch oligo
RPMI	Roswell Park Memorial Institute Medium	tTregs	Thymus-derived Tregs
rRNA	Ribosomal ribonucleic acid	UMAP	Uniform manifold approximation and projection
RT	Room temperature	UMI	Unique molecular identifier
		UV	Ultraviolet

Abstract

Allogeneic hematopoietic stem cell transplantation (aHSCT) is a treatment option for several human diseases and shows promising curative results, especially in treating hematological malignancies. A major complication during this treatment is the development of a graft-versus-host disease (GvHD), an immune reaction against the recipient's organs that can be life-threatening for the patient and limits the success of aHSCT as a therapy. One of the main target organs of this T cell-mediated disease is the gastrointestinal tract, and first observational correlations to the patient's microbiome have been made. In order to systematically analyze infiltrating immune cells of GvHD patients and their relationship to the microbiome and GvHD outcome, existing studies need to be expanded towards multiplexed analysis methods with single-cell resolution.

In this project, we analyzed a cohort of aHSCT patients regarding their immune infiltrates in gastrointestinal biopsies. We further monitored the microbiome status via 16S rRNA sequencing of stool samples and the patient's clinical status via urinary and serum biomarkers. To retrieve the maximum information on the immune infiltration, we combined two emerging techniques for single-cell analytics, which are single-cell RNA sequencing and multiplexed imaging via ChipCytometry. A single-cell RNA sequencing workflow was adapted to fit the needs of small punch-biopsies regarding cell recovery after isolation. The ChipCytometry technology, developed for the analysis of cryopreserved tissue samples, was transferred to FFPE biopsies in order to be able to analyze an existing repository of patient samples. Furthermore, we developed an automated quantification pipeline to analyze infiltration differences based on multiplexed imaging data.

Utilizing the established methods, we found an increased density of immune infiltration in severe GvHD patients with a particular enrichment of CD8 T cells. This inflammation was counterbalanced by regulatory T cells, but these were tightly dependent on the microbiome of the respective patients. Regarding this connection, we found a reduction in the numbers and suppressive capacity of Tregs upon microbiome disruption introduced by broad-spectrum antibiotic treatment. Furthermore, we detected a dependency of Treg differentiation on a healthy microbiome and, more specifically, on the presence of SCFA-producing bacteria in the gut. CD8 T cells, strongly connected to GvHD severity in our dataset, were clonally expanded in GvHD patients. Clonotypes were stable over long periods and were found at different anatomic locations. This systemic antigen specificity of the disease-driving cells is an important finding regarding the mechanistic pathophysiology of GvHD, and identification of the specific target antigens might open up new therapeutic possibilities in the future.

Notably, we could show in a single patient of a clinical FMT trial that restoration of the microbiome can induce regulatory T cells, which we observed to be able to suppress the clonal expansion of disease-driving CD8 T cells. This suppression could finally lead to the mitigation of GvHD, further encouraging the idea of microbiome-targeting interventions as therapeutic but also prophylactic treatment options for aHSCT patients.

1. Introduction

1.1 Allogeneic hematopoietic stem cell transplantation (aHSCT)

Allogeneic hematopoietic stem cell transplantation (aHSCT) allows to replace the hematopoietic system of a patient with the one from a healthy donor and represents a promising treatment option for several human diseases. It can be utilized for the treatment of patients with anemias (Bolanos-Meade and Brodsky, 2009) or immune deficiencies (Dvorak and Cowan, 2008), and it allows the therapy of hematological malignancies with myeloablative doses of irradiation or chemotherapy (Jenq and van den Brink, 2010). Especially in cancer treatment, aHSCT has become one of the most potent treatment options for liquid tumors.

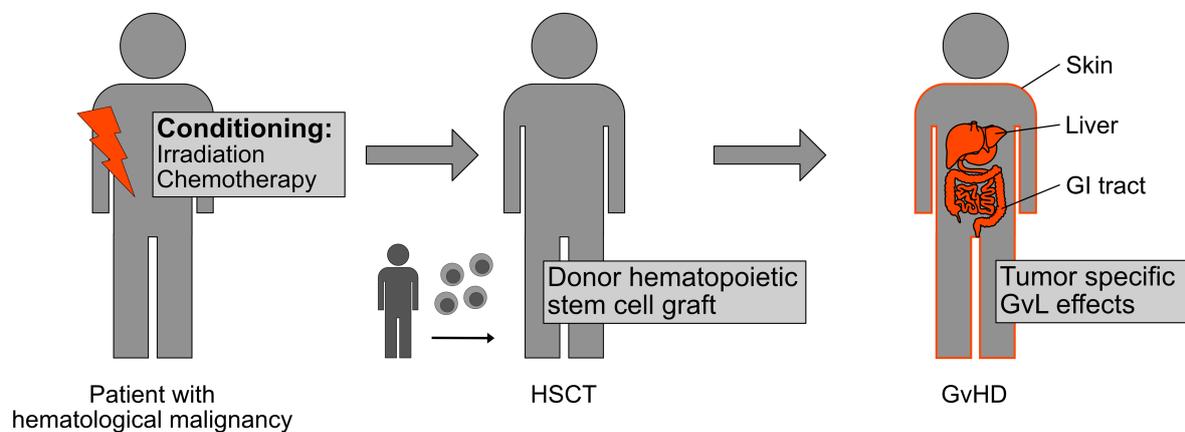


Figure 1. Allogeneic hematopoietic stem cell transplantation for cancer treatment.

Schematic depiction of aHSCT: The patient receives myeloablative conditioning to extinguish the malignant cells, which are then replaced by a hematopoietic stem cell graft. Donor immune cells then mediate the GvL effect and GvHD, where the recipient's organs are affected by a T cell-mediated inflammation. Mainly affected organs are the skin, liver, and gastrointestinal tract. Modified from (Shono and van den Brink, 2018).

After identifying the bone marrow as a source of hematopoietic stem cells (Lorenz et al., 1951), the first successful human transplantations were performed in 1959 for leukemia patients with identical twins as stem cell donors (Thomas et al., 1959). In the following years, aHSCT was further optimized and has become a routine treatment with more than 25.000 transplantations performed annually (Ferrara et al., 2009). Fully human leukocyte antigen (HLA)-matched siblings are the preferred donors for an aHSCT since, in this situation, the risk for complications during the procedure is minimized regarding graft rejections or graft-versus-host disease (GvHD, see 1.2). If such a sibling is unavailable, fully HLA-matched unrelated donors, partially HLA-matched relatives, or partially HLA-matched umbilical cord blood units can be used, broadening the applicability of stem cell transplantations (SCTs) at the expense of a higher risk for mismatch-associated complications. Fully haploid mismatched transplantations have recently gained popularity as they allow, in combination with post-transplant cyclophosphamide treatment, almost every patient to be transplanted from a haploidentical family member (Luznik et al., 2008). This is

especially important for countries with suboptimal donor registries and is nowadays more often applied than unrelated donor transplantations.

Hematopoietic stem cells are either derived from the bone marrow by aspiration from the iliac crest, mobilized from peripheral blood by granulocyte colony-stimulating factor (G-CSF), or extracted from umbilical cord blood (Giralt and Bishop, 2009). In most cases, stem cells from peripheral blood are used as the incidence of complications during stem cell extraction is very low, and the graft products have demonstrated high safety and efficiency (Stem Cell Trialists' Collaborative, 2005). Umbilical cord blood units can be a good alternative in fast-progressing malignancies since stem cells are faster available compared to the identification and processing of an unrelated, matched donor (Grewal et al., 2003). However, due to the low cell numbers available from umbilical cord blood, haploidentical transplantations are additional fast alternatives, more often used today.

Besides some treatments of benign hematological disorders like autoimmune diseases, primary immunodeficiencies or anemias and clinical trials for the treatment of solid tumors, the main field of application is still hematological malignancies (Storb et al., 2003). In this treatment strategy, the patient's hematopoietic cells are removed by myeloablative conditioning to extinguish the malignant cells. The compartment is subsequently replaced by the stem cell graft from an HLA-matched donor (Figure 1). The reason for its great potential in cancer treatment is that besides the higher curative rate of the myeloablative treatment, the immune-mediated graft-versus-leukemia (GvL) effect from the graft, representing an immune reaction against the tumor cells, leads to lower relapse rates and higher long-term survival (Kolb, 2008, Kolb et al., 1995). Unfortunately, allogeneic transplants are also associated with the risk of GvHD, an immune reaction against the recipient's organs, which still limits the success of this treatment strategy. Separating GvHD and GvL effects has been a central area of interest in aHSCT research in recent years. Both effects have been initially linked to an antigen-specific (either neo- or self-antigens) response of the T cell compartment (Korngold and Sprent, 1978), but also natural killer (NK) cells have recently been identified as substantial mediators for GvL responses (Ruggeri et al., 2002). Promising strategies for specific GvL induction after aHSCT are secondary adoptive NK cell transfers from the same donor (Miller et al., 2005) and adoptive T cell therapies with *in vitro* expanded or vaccine-primed autologous, tumor-antigen specific T cells (Rapoport et al., 2005).

Besides the tremendous therapeutic potential of aHSCT, it is still associated with several treatment-related toxicities, resulting in significant morbidity or even mortality after the transplantation. Apart from the risk of rejection (destruction of the graft by immunocompetent cells of donor origin), graft failure (loss of marrow function or failure to build up a hematopoietic system) or GvHD (see 1.2), infections are a severe risk in these immunocompromised patients (Giralt and Bishop, 2009). In many cases, a prophylaxis regimen via broad-spectrum antibiotics, antifungals, or antivirals is needed to prevent the patients from developing sepsis. These unavoidable treatments against infections alter the patient's microbiome, which plays an essential role in the maintenance of immune tolerance and in the development and severity of GvHD (see 1.6).

1.2 Graft-versus-host disease (GvHD) is a major complication after aHSCT

GvHD can be a life-threatening complication after aHSCT and limits the success of this potent, in many cases curative immunotherapy for hematological malignancies and other diseases. It is characterized by an immune reaction of donor T cells against the recipient's organs. The disease is subclassified into acute GvHD (aGvHD, within the first 100 days after transplantation) and chronic GvHD (cGvHD) at later stages (Billingham, 1966), with overlapping cases being reported (Schoemans et al., 2018). The incidence of acute GvHD among aHSCT patients is relatively high (35-50%), with 50% of the patients with aGvHD eventually developing a cGvHD (Jacobsohn and Vogelsang, 2007). HLA matching is used to avoid major histocompatibility differences between donor and recipient cells leading to GvHD reactions, and it has been shown that the incidence of GvHD is related to the number of mismatch alleles between the donor and recipient (Loiseau et al., 2007).

Grade	Organ-specific grading						Overall clinical grade
	Skin		Liver		GI tract		
0	No GvHD rash		<2 mg/dl total bilirubin		<500 ml/d diarrhea		No GvHD in any organ
1	Maculobapular rash <25 % BSA		2-3 mg/dl total bilirubin		500-999 ml/d diarrhea		1-2 skin <2 liver 0 GI tract
2	Maculobapular rash 25-50 % BSA		3.1-6 mg/dl total bilirubin		1000-1500 ml/d diarrhea		1-3 skin 1-2 liver 1-2 GI tract
3	Maculobapular rash >50 % BSA		6.1-15 mg/dl total bilirubin		>1500 ml/d diarrhea		2-4 skin 2-4 liver 2-4 GI tract
4	Maculobapular rash >50 % BSA bullous formation or desquamation >5 % BSA		>15 mg/dl total bilirubin		>2000 ml/d diarrhea		3 plus extreme constitutional symptoms

Figure 2. Clinical GvHD grading.

Overview of the clinical grading of GvHD for the individual organs and the resulting overall clinical severity grade. BSA=Body surface area. Visualized from (Glucksberg et al., 1974)

The three main affected organs are the skin, liver, and gastrointestinal tract (Vogelsang et al., 2003). The skin is the most common and early affected organ, and skin GvHD is marked by a maculopapular rash, able to spread throughout the body (Vogelsang et al., 2003). Gastrointestinal (GI) GvHD is characterized by diarrhea and can lead to vomiting, anorexia, and abdominal pain (Ferrara and Deeg, 1991). Liver GvHD can lead to hyperbilirubinemia and clinical jaundice, but it is difficult to differentiate from other reasons for liver dysfunction (Fujii et al., 2001). Clinical grading of GvHD is performed according to the Glucksberg criteria (Glucksberg et al., 1974) by scoring the involvement of the three mentioned target organs, and it is defined as 1 (mild), 2 (moderate), 3 (severe), and 4 (very severe) (Figure 2) with long-term survival rates as low as 5% for grade IV (Cahn et al., 2005). To minimize the risk of GvHD development, patients receive immunosuppressive therapeutics, which negatively affect susceptibility to infections due to slower immune reconstitution and can reduce the GvL effect. Therefore, it is essential to better understand the pathophysiology of GvHD and develop new treatment strategies to improve the effectiveness of aHSCT in treating hematological malignancies.

1.3 GvHD Pathophysiology

Already more than 50 years ago, Billingham formulated three required factors for GvHD development: (1) the graft needs to contain a sufficient number of immunocompetent cells, (2) the host needs to express antigens that are unknown to the graft, and (3) the host immune cells are not able to mount an immune response against the graft (Billingham, 1966).

All of these proposed concepts still fit the knowledge about GvHD nowadays as it is an inflammatory disease where activated donor T cells cause damage to the recipient's organs, especially the skin, the liver, and the GI tract. The preconditioning of the patients leads to tissue damage, serving as a starting point of an inflammatory cascade by the release of danger-associated molecular patterns (DAMPs) (Hill et al., 1997, Matzinger, 2002). The resulting inflammation can be subdivided into three main phases of GvHD: activation of antigen-presenting cells (APCs), donor T cell activation, and tissue destruction (Ferrara et al., 2009) (Figure 3).

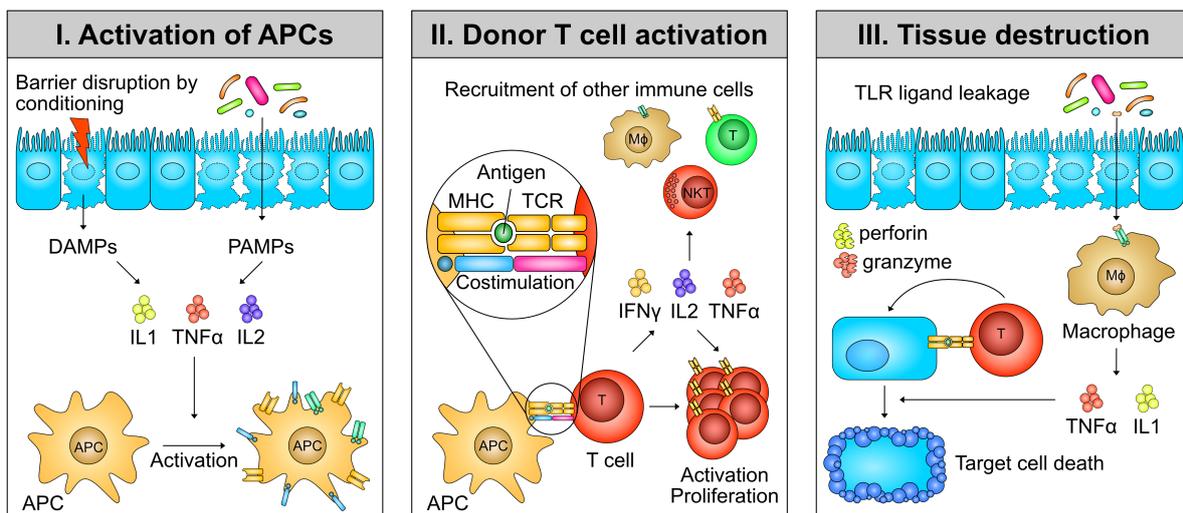


Figure 3. Three phases of GvHD.

In the first phase, conditioning-induced tissue damage leads to the production of cytokines via danger- and pathogen-associated molecular patterns (DAMPs/PAMPs). These signals lead to the activation of APCs by upregulation of costimulatory receptor- and MHC expression. In the second phase, donor T cells are activated by host APCs presenting minor histocompatibility antigens, which leads to T cell activation, proliferation, further enhanced cytokine secretion and recruitment of other immune cells. The primed and activated T cells can kill the target cells, finally leading to tissue destruction and severe GvHD symptoms. Visualized from (Ferrara et al., 2009)

1.3.1 Activation of APCs

The tissue damage in the recipient induced by the preconditioning regimen leads to the release of danger signals by dying cells, resulting in the production of cytokines, chemokines, and expression of costimulatory molecules by innate immune cells. Proinflammatory cytokine levels like TNFα (Choi et al., 2008) or IL12 (Mohty et al., 2005, Holler et al., 1990) have been correlated to the risk and severity of GvHD. Further, these signals lead to enhanced expression of MHC molecules and costimulatory receptors on the surface of host APCs, thereby activating those cells (Figure 3, left). Tissue damage in the GI

tract is exceptionally important in the development of GvHD because of the potential translocation of pathogen-associated molecular patterns (PAMPs) from the lumen into the tissue due to epithelial barrier disruption (Hill and Ferrara, 2000). These signals can be detected by innate immune receptors like toll-like receptors (TLRs), which sense microbial patterns and thereby further enhance the risk of immune cell and especially APC activation. The importance of activated APCs from the gastrointestinal tract has been underlined by findings in murine models that the Peyer's patches are an essential site for T cell priming in the setting of GvHD (Murai et al., 2003). Furthermore, a more direct association between PAMPs and GvHD in the gut has been demonstrated by reduced GvHD severity in mice treated with a lipopolysaccharide (LPS) antagonist, impairing the recognition of one of the most important PAMPs (Cooke et al., 2001). Because of these findings, the conditioning intensity is kept moderate in order to reduce toxicity and tissue damage, which eventually lead to a reduced risk of GvHD development.

1.3.2 Donor T cell activation

The second phase of the GvHD reaction is the activation of donor T cells by the before-activated APCs, and this step is crucial for determining the disease severity and tissue damage (Figure 3, middle). It has been shown in murine models that host APCs, still present in the tissues early after transplantation, are required for GvHD and GvL reactions (Zhang et al., 2002). After the loss of host APCs in the process of immune reconstitution, the donor cells can take over T cell stimulation by presenting host antigens, thus supporting the ongoing GvHD (Matte et al., 2004).

In HLA-mismatched transplantations, donor T cells can recognize the MHC molecules of host tissues (probably in the context of commonly expressed self-peptides), leading to T cell activation and proliferation. In HLA-matched situations, minor histocompatibility antigens (miHAs), which derive from polymorphic genes that differ between donor and recipient, are presented by the matched MHC molecule and eventually activate donor T cells. Some of these antigens have been identified, like the male-specific SMCY gene (Wang et al., 1995), the human HA-1/HA-2 (den Haan et al., 1995) alleles as well as specific genes that are only expressed in the donor and not in the recipient (Murata et al., 2003). It has been shown in mice that MHC II differences lead to CD4 T cell-mediated GvHD, whereas MHC class I differences mount alloreactive CD8 T cell responses (Csencsits and Bishop, 2003).

The activation of T cells leads to the production of cytokines (IL2, IFN γ , TNF α), causing further proliferation and activation of T cells as well as the recruitment of other immune cells. Regulatory T cells (see 1.4) and NKT cells can suppress these alloreactive responses and are, therefore, attractive candidates for cell therapy interventions of GvHD (Cohen and Boyer, 2006, Zeng et al., 1999, Hoffmann et al., 2002, Edinger et al., 2003).

1.3.3 Tissue damage and inflammation

The activated T cells ultimately mediate the destruction of host tissues (Figure 3, right). T cells using the Fas/FasL pathway of target cell killing are preferentially found in the liver, whereas T cells in the GI tract and the skin are mainly lysing target cells via the release of perforin or granzyme cytotoxic molecules (van den Brink and Burakoff, 2002). This target cell lysis can be detected histologically via the apoptosis of epithelial cells, known as a hallmark of GvHD (Azad et al., 2019, Epstein et al., 1980).

The migration of activated effector cells back into the circulation and to the target organs plays an essential role in disease development, and specific chemokines and homing receptors have been described to be associated with GvHD (Waldman et al., 2006, Wysocki et al., 2005). Stimulation of TLRs in the gut or the skin promotes innate immune cells to produce cytokines like TNF α , which can cause direct tissue damage by necrosis. Other cytokines indirectly mediate tissue damage by their immunostimulatory activity, leading to further T cell activation (Hill and Ferrara, 2000). The synergistic effect of cytokines and immune cells promotes inflammation, tissue injury and, finally, target organ damage. Eventually, extended tissue damage can lead to organ dysfunction and, therefore, life-threatening complications.

1.4 Regulatory T cells in GvHD

The gastrointestinal tract is a unique organ from an immunological point of view because of the frequent exposure to foreign antigens. This requires, on the one hand, tightly regulated tolerance mechanisms for food antigens and commensal bacteria to prevent chronic inflammation and tissue damage. On the other hand, immune responses against potentially harmful pathogens must be reliably mounted. Therefore, a controlled balance between potent immune cells fighting infections and regulatory mechanisms to dampen overshooting immune responses is inevitable in the gastrointestinal tract to keep a healthy environment.

Regulatory T cells (Tregs), specialized cells of the T helper cell lineage, are capable of suppressing immune responses and are the key players in maintaining peripheral tolerance, including immune homeostasis in the gut (Sakaguchi et al., 1995). Identified by the expression of the transcription factor forkhead box protein p3 (FOXP3) (Fontenot et al., 2003, Hori et al., 2003), they can suppress immune responses by direct cell-to-cell interaction, for example, via CTLA4 that, antagonistically binding to CD80/86 on APCs, hinders CD28 binding and therefore co-stimulation of T cells. Furthermore, the secretion of regulatory and immunosuppressive cytokines like IL-10 and TGF- β , directly acting on inflammatory cells, or consumption of IL-2, which is a stimulus for proliferation of activated T cells (Figure 4), are important modes of action for regulatory T cells (Josefowicz et al., 2012).

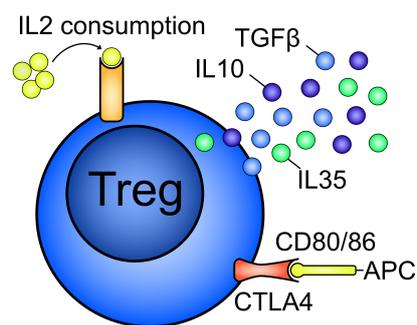


Figure 4. Regulatory T cell function.

Regulatory T cells can inhibit immune responses by cytokine release, cell contact dependent inhibition and IL2 consumption.

Based on their site of development, they are subdivided into thymic Tregs (tTregs) and peripherally-induced Tregs (pTregs). CD4 T cells that show a high affinity to self-antigens in the thymus are converted to tTregs, while pTregs develop in the periphery from mature conventional CD4 T cells upon antigen recognition in the presence of TGF- β . Tregs can be induced in solid cancer patients by the tumor microenvironment to avoid immunosurveillance and dampen antitumor immune responses (Togashi et al., 2019). On the contrary, deficiencies in the development and maintenance of Tregs are associated with autoimmune diseases and immune dysregulation, marking a high clinical potential for these cells.

In the case of GvHD, Tregs play an essential role in the control of the disease by suppressing inflammation. Many studies observed an inverse correlation between Treg frequencies in peripheral blood (Fujioka et al., 2013) or absolute counts in tissue samples (Rieger et al., 2006) and the development of acute GvHD. Preclinical studies also underlined the importance of Tregs for GvHD progression since their absence led to the aggravation of GvHD, and the simultaneous transfer of Tregs could inhibit the expansion of alloreactive T cells (Edinger et al., 2003, Cohen et al., 2002). The early availability of Tregs has thereby proven critical since GvHD was reduced compared to a later Treg administration (Nguyen

et al., 2007). Although therapeutic transfer of donor Tregs has substantial activity in preclinical models (Riegel et al., 2020), transferring sufficient numbers of regulatory T cells in human patients is challenging, and physiological frequencies have no protective capacity against GvHD. Tregs were administered in a 1:1 ratio to the conventional T cells (Hoffmann et al., 2002) in the mouse models of GvHD, but these numbers are very hard to achieve for human regulatory T cells.

Even the latest technological efforts, like the *in vitro* expansion or induction of regulatory T cells, have been unsuccessful so far due to the preferential proliferation of conventional CD4 T cells and the instability of the induced Tregs (Beres et al., 2011). However, data from our lab showed that the technology of Treg isolation might influence Treg recovery and that minimally manipulated Tregs, isolated label-free, have an improved recovery and capacity to suppress GvHD in preclinical models (Mohr et al., 2017). Such technologies would allow for the transfer of lower cell numbers and overcome the previously mentioned limitations in adoptive Treg transfer. However, unfortunately, the transfer of Tregs brings the risk of diminishing the GvL effect. For this reason, other prevention and treatment options for GvHD are currently used and are under further investigation.

1.5 Prevention and treatment of GvHD

Since T cells play a significant role in GvHD pathophysiology, T cell depletion was tested for GvHD prophylaxis in the early days of stem cell transplantations. The depletion methods included the negative selection of T cells by anti-TCR $\alpha\beta$ antibodies, positive selection of CD34+ stem cells, and the use of T cell neutralizing antibodies in the patient. These approaches showed a significant reduction of GvHD development but especially *ex vivo* T cell depletion is associated with reduced GvL effects, graft failure, and severe infections, therefore not improving the overall survival (Marmont et al., 1991, Platzbecker et al., 2004).

The most used strategy for GvHD prevention nowadays is the inhibition of T cell activation by administration of calcineurin inhibitors (e.g., cyclosporine, tacrolimus), often in combination with other immunosuppressants (e.g., methotrexate, mycophenolate mofetil) to improve the efficiency of immune cell suppression (Nash et al., 2000). In addition to the suppression of immune responses, the intensity of the conditioning regimen was reduced in order to mitigate tissue damage and GvHD initiation; however, in many cases, this therapeutic scheme leads only to a delayed onset of the disease (Mielcarek et al., 2003).

Since none of the GvHD prevention strategies applied so far were effective in extinguishing the risk of GvHD development, the incidence of acute GvHD today is still above 50% after aHSCT. The standard of care for acute GvHD is based on steroids, which have potent anti-inflammatory and immunosuppressive activity (MacMillan et al., 2002). However, more than 50% of severe GvHD patients develop a resistance to steroid treatment. Thereby, additional treatment options are needed. At the moment, extracorporeal photopheresis, an apheresis method inducing apoptosis of white blood cells by exposure to Methoxalen and UV light (Greinix et al., 2006), TNF α blockade (Levine et al., 2008) and Ruxolitinib as second-line treatment (Zeiser et al., 2020) are some promising intervention candidates as shown in clinical trials.

Still, GvHD prevention or risk adapted (e.g., biomarker guided) preemptive treatment would be of high interest since all the currently available treatment options have severe side effects and the patients need complex supportive care from specialized medical staff in order to prevent toxicities, severe infections, or irreversible tissue damage. A better understanding of GvHD pathophysiology and investigation of new treatment and prevention strategies is essential to reduce the risk of GvHD development and enhance the potential of aHSCT.

1.6 The influence of the microbiome on GvHD

Since the main target organs of GvHD (gut, skin, liver) are exposed to bacteria, the investigation of potential interactions with the microbiome has been started early in the history of aHSCT and GvHD research. A particular focus of research has been the interplay between the microbiome and gastrointestinal GvHD as a primary indicator for total GvHD severity (Figure 2) and the main interaction site with commensal bacteria. Already in 1971, it was found that germ-free mice were protected from GvHD and had a prolonged survival (Jones et al., 1971). Since this protection may be derived from the less developed immunity in germ-free mice, van Bekkum et al. further showed that also antibiotic-treated (“decontaminated”) mice had prolonged survival after transplantation (van Bekkum et al., 1974). These early findings represent an auspicious connection between GvHD and the microbiome in murine experiments.

1.6.1 Associations from human studies

After these early observations from preclinical models, some clinical studies have been performed using broad-spectrum antibiotics for “gut decontamination” and isolation of the patients to mimic the germ-free situation of preclinical models in humans (Beelen et al., 1992). However, no clear trend of a benefit has been found for GvHD prevention (Passweg et al., 2016), and gut decontamination is, for this reason, not routinely recommended for treatment (Tombly et al., 2009). In more recent studies, technological advancements in microbiome research like 16S sequencing allowed to investigate the microbiome of GvHD patients in more detail. It was found that GvHD patients are associated with decreased stool microbial diversity (Jenq et al., 2012). 3-indoxyl sulfate, an indole metabolism product generated by gut bacteria, has been established as a clinical marker for microbial diversity, and low urinary levels are significantly correlated to GvHD severity (Weber et al., 2015). Furthermore, loss of microbial diversity has been linked to reduced survival and increased transplant-related mortality (Taur et al., 2014). Besides microbial diversity, increased abundance of *Lactobacilli* and reduced presence of *Clostridiales* has been associated with GvHD occurrence (Jenq et al., 2012). Other cohorts revealed *Blautia* species to reduce the risk of GvHD-related mortality (Jenq et al., 2015) and *Enterococci* to increase the risk of GvHD development (Holler et al., 2014). Lower abundance of *Bacteroides* and *Parabacteroides* and lower levels of the SCFA propionate, another fermentation product of the gut flora, were also linked to the risk of GvHD development (Biagi et al., 2015). These first findings of the interplay between microbiota and GvHD demonstrated the complexity of these interactions and need to be validated in multi-center studies to exclude site- and treatment-specific effects. One of these recent multi-center studies confirmed the association between the abundance of certain *Enterococci* species and the risk of GvHD (Stein-Thoeringer et al., 2019). Other observational studies have proven the positive effect of bacterial metabolites like short-chain fatty acids in more extensive, multi-center patient cohorts (Meedt et al., 2022, Markey et al., 2020). Taken together, there is increasing evidence that microbiota plays an essential role in GvHD development; however, mechanistic explanations and species-level associations are still under investigation, also drawing attention to inter-institutional differences (Peled et al., 2020).

1.6.2 Mechanisms by which microbiota may influence GvHD

At first view, the findings from human studies regarding the microbiome's influence on GvHD development seem paradoxical. On one side, there is evidence that germ-free mice and, in some clinical trials, also human individuals receiving gut decontamination are better protected from GvHD, suggesting that gastrointestinal bacteria drive GvHD. On the other side, high bacterial diversity is associated with a reduced risk of GvHD, hinting toward a beneficial role of microbiota in suppressing GvHD. One explanation for this paradox would be the need for proper balance between beneficial and potentially pathogenic bacteria in a healthy gut, which may be disrupted in an aHSCT setting by the preconditioning as well as the often-unavoidable antibiotic treatment of the immunocompromised patients. Furthermore, gut decontamination is only partially effective in human patients, leading to a dysbiosis with a high frequency of antibiotic-resistant, potentially pathogenic bacteria.

Detrimental effects of the microbiota on GvHD development are mainly dependent on bacterial translocation (Cooke et al., 2001), which in turn induces inflammation and recruitment of immune cells. The outgrowth of pathogenic bacteria in an imbalanced community can lead to a higher chance of PAMP translocation, for example by mucus-degrading bacteria, further reducing the intestinal barrier integrity. The translocation could be reduced with a decontamination. However, this would also affect the long-term advantageous effects of a diverse microbiome.

The beneficial effects of bacteria are linked to the production of bacterial metabolites like SCFA and riboflavin. The SCFA butyrate is an important energy source for intestinal epithelial cells and can improve wound healing and barrier integrity (Ma et al., 2012), favoring recovery after the damage induced by the preconditioning therapy. Furthermore, SCFAs can directly influence the immune response by induction of regulatory T cells (Arpaia et al., 2013). This induction can be mediated by binding to the G-protein-coupled receptor 43 (GPR43), an SCFA receptor expressed on regulatory T cells, or by inhibition of the histone deacetylase (HDAC), both leading to induction, proliferation, and accumulation of Tregs (Smith et al., 2013). The beneficial effects of microbiota in the setting of GvHD have been proven by the increased abundance of SCFA-producing bacteria, identified by copy numbers of the enzyme butyryl CoA:acetate CoA transferase (BCoAT) in stool samples of patients with low-grade GvHD (Meedt et al., 2022). This observation was correlated with the low abundance of *Clostridiales*, one of the main SCFA-producing classes, in severe GvHD patients and is thereby perfectly in line with other human studies assessing the link between the gut microbiome and GvHD (see 1.6.1). Riboflavin, synthesized by specific bacterial strains, can enhance the control of GvHD by recruiting MAIT cells, which can suppress immune responses as well (Varelias et al., 2018).

A "healthy" microbiome could largely reduce the risk of GvHD by maintaining regulatory mechanisms and barrier integrity (Figure 5, left). Most of these mechanisms are related to regulatory T cells, which directly interact with the microbiome via bacterial metabolites. Unfortunately, due to the preconditioning and antibiotic treatment, aHSCT patients have to deal with a disrupted epithelial barrier, an imbalanced

microbiome, and the risk of outgrowing pathogenic bacteria. All these factors together favor an inflammatory situation, which is finally responsible for T cell activation and tissue damage in the target organs (Figure 5, right). The strong associations between the microbiome and GvHD suppression via regulatory T cells is still an important therapeutic target for clinical interventions, and the systematic investigation of this interaction may help to develop new treatment strategies.

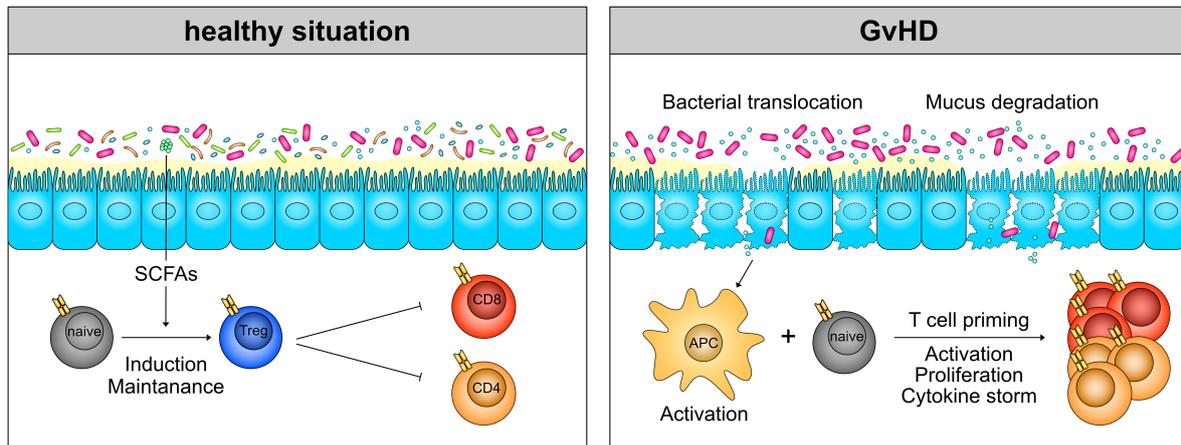


Figure 5. Mucosal immunity in healthy and GvHD situations.

In a healthy human gut, bacteria mount a regulatory milieu, allowing the induction and maintenance of regulatory T cells, which can dampen overshooting immune responses (left). The conditioning-induced damage in the epithelial barrier and the reduction of bacterial diversity leads to less Treg induction, mucus degradation, more danger signals, and finally, the inflammatory cascade of GvHD.

1.6.3 Microbiome modulation as a treatment option for GvHD

The previously described connection between microbiota richness and GvHD suggests microbiome modulation as a potential treatment strategy. Most patients show a loss of bacterial diversity due to the preconditioning in the weeks after transplantation, often leading to the outgrowth of pathogenic bacterial strains (Holler et al., 2014, Taur et al., 2012). This microbial imbalance is often further pronounced due to the treatment with broad-spectrum antibiotics, essential for many immunocompromised patients.

Fecal microbiota transplantation (FMT) is the most direct approach of microbiome modulation and has been applied to treat steroid-refractory GvHD patients. In the first patient reports, this strategy was curative in more than 75% of the patients, which otherwise have mortality rates of 70-80% (Kakihana et al., 2016, Spindelboeck et al., 2017, Malard et al., 2021). Furthermore, even prophylactic FMT treatment has been proven to be safe in aHSC T patients and restore microbial diversity (DeFilipp et al., 2018). Oral microbiota administration is achieved by capsules of frozen microbiota from healthy donors (Youngster et al., 2014), and it is still under investigation if the stem cell donor might be superior to unrelated donors. Besides transferring the total microbiome, single strain administration has been investigated and proven to modulate GvHD outcome, but with the risk of sepsis development in immunocompromised patients (Mehta et al., 2013). A safer approach would be the direct administration of

microbial metabolites like SCFAs (Mathewson et al., 2016) or indoles (Swimm et al., 2018). Similar effects can be achieved by sterile filtrates of fecal material from healthy donors, which would be a safer alternative to FMT by not containing living bacteria but only components and metabolites (Ott et al., 2017). Although there are promising and ongoing clinical trials, there is still a substantial lack of knowledge regarding the mechanistic understanding of the connection between the microbiome, infiltrating immune cells, and GvHD outcome.

Most studies conducted so far in the field of GvHD research have been performed with conventional methods from the pathology and molecular biology. These include immunohistochemistry staining of single markers in tissue samples and qPCR for gene expression analysis to quantify differences in tissue immune infiltrates concerning GvHD. However, since the disease is associated with a very complex phenotype and multiple factors like disease severity, timepoint, treatment strategy, and the microbiome can influence the immune infiltrates, novel and advanced technologies are needed for systematic studies of patient samples. These methods should, in particular, be able to achieve deep phenotyping of immune cells at single-cell and high spatial resolution to gain more information on the specific cell types involved, as well as their spatial location or interplay in the setting of the disease.

1.7 Multiplexed analysis methods for single-cell characterization

The development of sequencing techniques towards next-generation sequencing (NGS) allowed for the first time to perform gene expression analysis on cell populations and to compare the transcriptomes of certain groups. However, these bulk sequencing experiments provide only the average information of transcriptomic signatures in a population of cells. To overcome this limitation, transcriptomic sequencing on the level of single cells was developed, opening a new area of research. Several single-cell multi-omics techniques combine single-cell RNA sequencing with additional information like surface protein expression, T cell and B cell receptor (TCR/BCR) sequencing, chromatin accessibility, and many more. It has become a state-of-the-art technology, and the routinely developing tools for data analysis even enhance the power of existing scRNA seq datasets. In parallel to the advances in single-cell RNA sequencing, imaging methods were developed to analyze cells in their spatial context. Here, the resolution of imaging and the throughput for multiparameter imaging are continuously improved. Today, much effort has been made to bring those two fields together and develop methods that can interrogate the whole transcriptome in the spatial context of a tissue sample.

1.7.1 Single-cell RNA sequencing

After initial approaches using microarray analysis of amplified complementary DNA (cDNA) from single cells (Kurimoto et al., 2006), the first whole transcriptome analysis on single cells was performed in 2009 on a single mouse blastomere (Tang et al., 2009). From here on, the interest in scRNA sequencing was tremendous, leading to the development of several different techniques and a zoo of analysis packages. Single-cell isolation methods include the low-throughput methods of limited dilution (pipetting cells in 0.5 cells/aliquot) and micromanipulation (picking cells under the microscope). Higher throughput can be achieved by cytometry-based sorting of single cells (Picelli et al., 2013), combinatorial barcoding by splitting cell suspensions several times to a set of unique barcodes (Rosenberg et al., 2018), and droplet-based microfluidics, where single cells are encapsulated in aqueous spheres within an emulsion (Zheng et al., 2017, Macosko et al., 2015). scRNA seq utilizing droplet-based microfluidics is the most commonly used technique today. After singularization of cells, the mRNA is reverse transcribed to cDNA, which can be further amplified and modified for subsequent NGS.

In this project, we used the commercially available droplet-based microfluidic system from 10X genomics for single-cell RNA sequencing on flow cytometry sorted cells (Zheng et al., 2017). This method relies on a pool of uniquely barcoded gel beads and the encapsulation of single cells with single beads in an aqueous droplet in emulsion. The surface of each bead contains DNA tails composed of a capture sequence, a bead-specific barcode, and a tail-specific UMI to quantify the absolute abundance of mRNA molecules. After reverse transcription in emulsion, the molecules can be processed in bulk and traced back to individual cells and mRNA molecules thanks to the bead- and tail-specific barcodes (Figure 6A).

The methodology was initially developed for the 3' capture of mRNA molecules via the poly-A tail, common for all mRNA sequences. For polymorphic T cell receptor or B cell receptor sequences, however, most of the variability (and epitope specificity) is encoded in the CDR3 region at the 5' end of the transcript. Therefore, for profiling TCR/BCR repertoires, it is necessary to capture the mRNA on the 5' end to create constructs that contain the variable CDR3 region of the receptor and are short enough to be compatible with NGS-based sequencing. This is possible due to the use of a template switch oligonucleotide (TSO), which allows the addition of any sequence at the 3' end of a cDNA (Zhu et al., 2001). These technological optimizations allow now to simultaneously analyze the whole transcriptome and the immune cell receptor in the same single cell (Figure 6B). Immune cell receptor sequences can be amplified by using primers specific to the known gel bead sequence and the constant region of a TCR or BCR to perform immune profiling of single cells (Figure 6C).

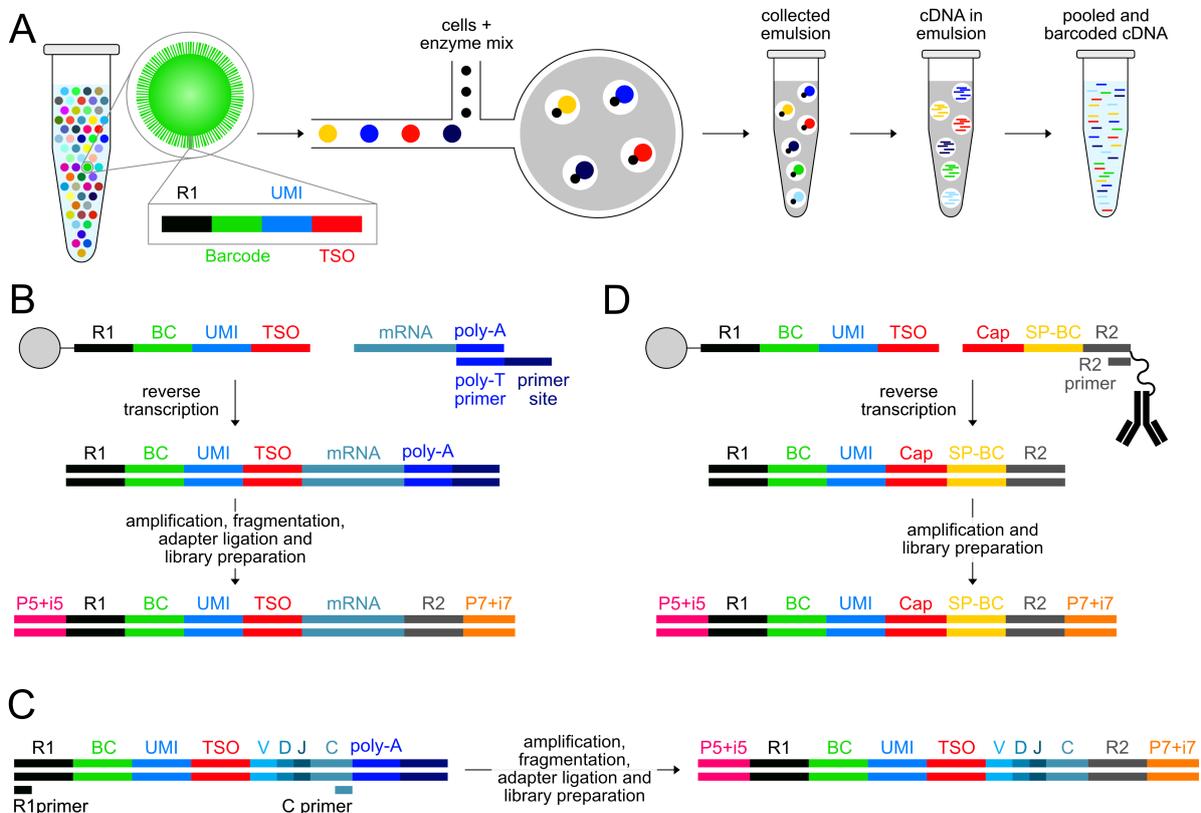


Figure 6. Droplet-based 5' single-cell RNA sequencing with 10X genomics.

(A) Schematic depiction of the 10X genomics 5' scRNA sequencing workflow: each unique gel bead is loaded with oligonucleotide tails, encoding Read 1 (R1), the bead-specific barcode (BC), the tail-specific unique molecular identifier (UMI), and a template switch oligo (TSO). By limited dilution, every cell is encapsulated with one gel bead and the reagents needed for reverse transcription in an emulsion within the microfluidic system. This allows amplifying and processing of the pooled cDNA, still capturing the cellular origin and the number of transcripts via BC and UMI. (B) The TSO captures the 5' end of the mRNA, and reverse transcription is initiated with a poly-T primer. After cDNA amplification, the DNA is fragmented, the Illumina Read 2 (R2) is added via ligation, and the indices are added via overhang PCR. (C) VDJ cDNA of TCR or BCR transcripts is amplified with an R1 primer and a specific primer binding to the constant region. Library preparation is performed similarly to the total mRNA (B). (D) Capturing antibody conjugated oligos works via a specific capture sequence (Cap) and reverse transcription initiated with an R2 primer. After cDNA amplification, the indices are added via index PCR. Figure adapted from (www.10xgenomics.com).

An additional advancement has been the implementation of antibodies conjugated to a barcoded, single-stranded DNA to capture surface marker expression. The DNA barcode can also be captured on the gel beads and amplified with specific primers (Figure 6D). Barcoded antibodies against broadly expressed markers (eg. CD45, MHC class I) can be used for the pooled analysis and demultiplexing of different samples within the same reaction via a cell “hashtag”, meaning one specific barcode identifying each subsample (Stoeckius et al., 2018). Furthermore, multiplexed surface expression profiling can be performed by combinatorial staining of several differently barcoded antibodies specific for a set of different epitopes (Stoeckius et al., 2017). Additional developments in the multi-omics direction include the integration of chromatin accessibility (Cao et al., 2018) and DNA methylation (Angermueller et al., 2016) to investigate epigenetic changes.

The rising interest in single-cell RNA sequencing datasets has led to a fast-developing field of bioinformatics, which provides researchers with analysis tools for datasets with high and increasing complexity. Various tools are used for quality filtering, alignment, and pre-processing of datasets, and guidelines have been formulated to guarantee a harmonized data analysis with high quality (Luecken and Theis, 2019). After filtering and pre-processing datasets, some general concepts of analysis help to understand and visualize high-dimensional datasets. One vital method to facilitate the analysis is dimensional reduction, which allows to visualize a gene expression landscape of several thousand genes in a two-dimensional space. The linear principal component analysis (PCA), as well as non-linear dimensional reduction methods like t-SNE (van der Maaten and Hinton, 2008) and UMAP (McInnes et al., 2018), are used to visualize the complex expression differences between single cells. Clustering algorithms on the neighborhood graph like Louvain or Leiden (den Haan et al., 1995) facilitate the detection of cells with similar expression profiles, and these clusters can finally be annotated to specific cell types within a single-cell RNA sequencing dataset (Figure 7A). For clonotype analysis of single-cell immune receptor data (TCR/BCR), additional tools have been developed, including algorithms for investigating receptor similarity based on gene usage and CDR3 sequences (Sturm et al., 2020).

All of the mentioned analyses generate new variables for every single cell or every single gene, with the need for special data formats suitable for scRNA sequencing datasets. One of these annotated data formats, specially developed for this purpose, is `anndata` (Virshup et al., 2021). Besides the count matrix (cells x genes), it contains a collection of arrays aligned to the cells (observables) and genes (variables) and additional annotation structures (Figure 7B). This data storage allows easy access and result storage for the different analyses of a scRNA sequencing experiment.

Besides annotation, pre-processing, and “classical” analysis of scRNA seq data based on the gene expression profile of single cells, computational methods allow the extraction of additional information without the need to change the experimental procedure. The possibilities are enormous and include the detection of copy number variants, allowing the investigation of inter-clonal differences in tumor samples (Fan et al., 2018) and analysis of genetic variation between donors (Kang et al., 2018). Furthermore, many tools have been developed in the field of transcriptional dynamics by RNA velocity analysis,

pioneered by La Manno et al. in 2018 (La Manno et al., 2018). These tools allow lineage tracing and pseudotime analysis and have shown to be able to unbiasedly detect the root cells of a trajectory (Bergen et al., 2020). RNA velocities are generally based on the ratio between spliced and un-spliced mRNA still containing intronic regions for each gene, which can be translated into gene activity per cell (Figure 7C). The ratio between spliced and unspliced mRNA can be used to differentiate cells into three central states for each gene: gene expression induction, repression, and steady-state (Figure 7D). By extracting dynamical genes, subpopulation kinetics can be identified for the whole dataset to infer each cell's developmental state in relation to the rest of the cells (Figure 7E).

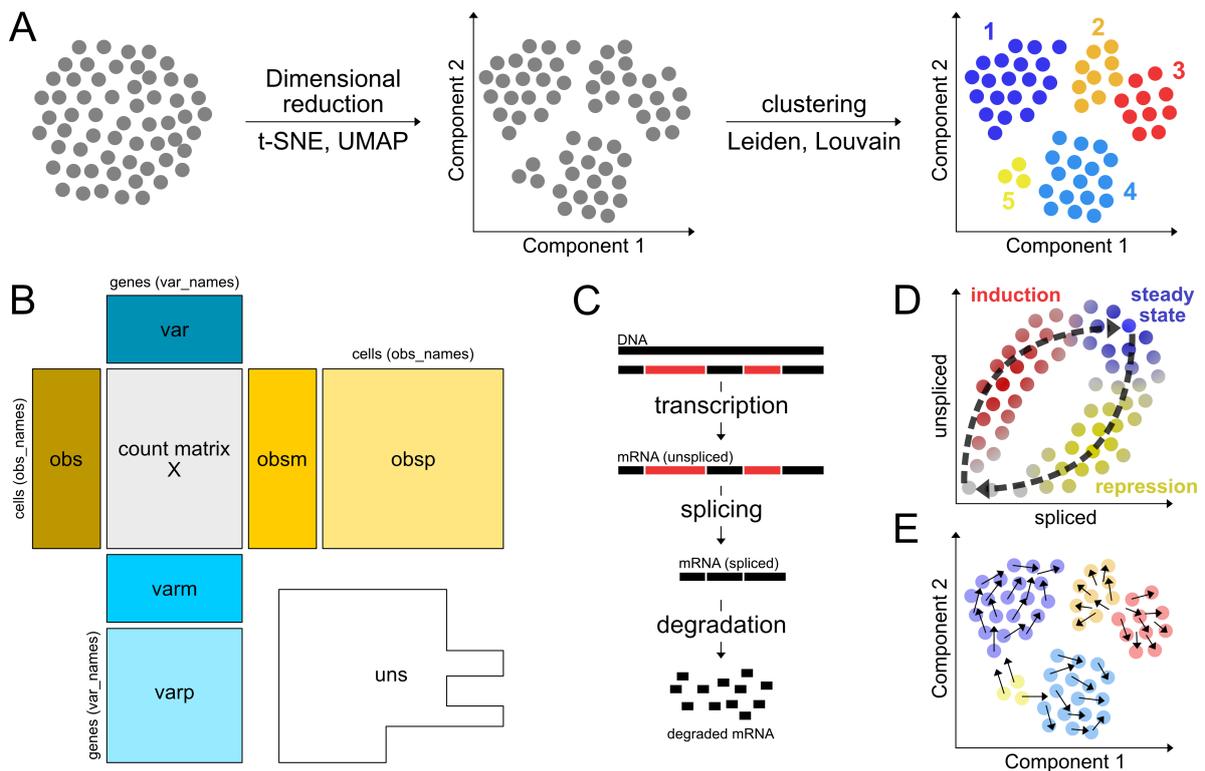


Figure 7. Computational analysis of scRNA sequencing experiments.

(A) After annotation and pre-processing, dimensional reduction is performed to visualize the gene expression landscape in two dimensions. Clustering cells according to their similarity is a good starting point for analyzing different cell populations. (B) Structure of anndata: in parallel to the count matrix X , one-dimensional annotations (obs/var), multi-dimensional annotations (obsm/varm), and pairwise relationships (obsp/varp) are stored for cells and genes. Unstructured annotations are separated (uns). Modified from (Virshup et al., 2021). (C) Processing steps of mRNA from transcription to degradation. (D) Transcriptional states of a cell for one hypothetical gene according to the frequency of spliced vs. unspliced mRNA. (E) Transcriptional dynamics over all genes can be visualized on the level of single cells as directionality of differentiation. (C-E) visualized from (Bergen et al., 2020, La Manno et al., 2018)

1.7.2 Multiplexed imaging techniques

Single-cell transcriptomics has expanded our knowledge about cell types and phenotypes in tissue samples, but it requires tissue dissociation and lacks the spatial context for cell-to-cell interactions. Tissues are characterized by defined ensembles of different cell types, interacting in complex networks that maintain organ health and tissue functions. Imbalances within these compositions can lead to tissue damage and pathological events. Therefore, the understanding of cellular phenotypes, together with

their spatial context, gains more and more importance in clinical research. There have been methods established which add the spatial component to RNA sequencing technologies by barcode-coated supplies that can track the spatial position of transcriptomic information (Stahl et al., 2016). However, the resolution is still limited and has not reached the single-cell level yet. Recent developments are increasing the resolution of spatial transcriptomics for fresh-frozen tissues (Chen et al., 2022). However, the whole transcriptome analysis still has some disadvantages compared to conventional imaging methods regarding spatial resolution.

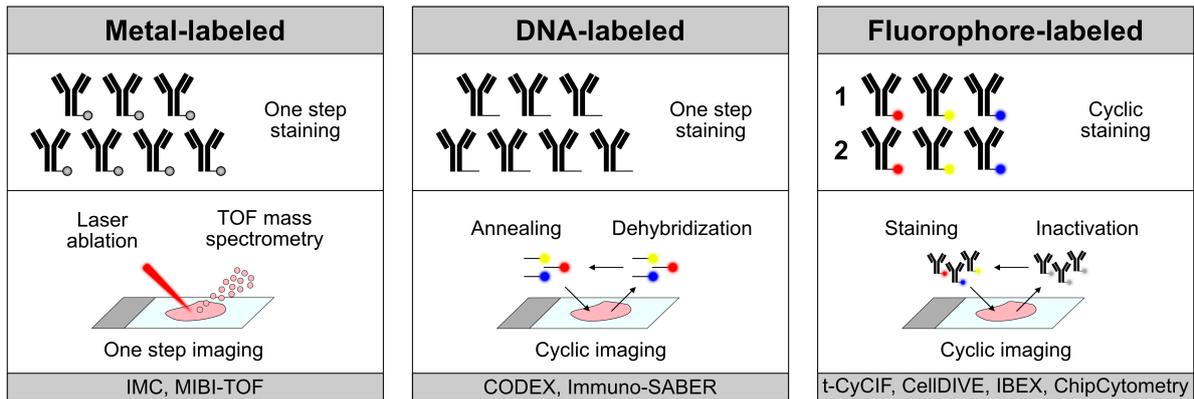


Figure 8. Comparison of antibody-based, FFPE-compatible HMTI methods.

Antibody-based HMTI method can be subclassified according to the antibody conjugation and the staining and detection method. Metal-labeled antibodies are stained in one round and imaged via laser ablation and absorption into a mass spectrometer (left). DNA-labeled antibodies are also stained in one round and imaged in iterative cycles with complementary, fluorescently labeled DNA oligos, which are de-hybridized by heating after each step (middle). Fluorophore-labeled antibodies are stained and imaged in iterative cycles, and the fluorophore is inactivated after each acquisition (right). Modified from (Hickey et al., 2022).

Highly multiplexed tissue imaging (HMTI) has gained much interest in the last few years. In parallel to the efforts to increase resolution in spatial transcriptomics, the number of markers to be analyzed on a tissue sample is increasing toward whole-transcriptome approaches. Most HMTI methods are based on antibodies conjugated to fluorophores, DNA, or metal tags, depending on the readout of the system (Figure 8). Tissues can be stained with up to 40 metal-tagged antibodies, micro-dissected using laser ablation, and particles analyzed by time-of-flight mass-spectrometry to finally reconstruct the tissue section (Giesen et al., 2014, Angelo et al., 2014). In addition to being a destructive method, imaging mass cytometry (IMC) requires expensive and complex instrumentation and user skills. Furthermore, the resolution of laser ablation might still be limited compared to fluorescent imaging. In these one-shot experiments, all epitopes of interest are stained and acquired simultaneously, thus preventing additional experimental steps. DNA-labeled antibodies expand the multiplexing capacity of isotope tags (restricted to 40 epitopes) and can be detected by cyclic imaging with fluorescently-labeled, complementary DNA oligos, which are removed after each imaging step (Goltsev et al., 2018, Kennedy-Darling et al., 2021). This methodology, called co-detection by imaging (CODEX), has further been developed using DNases for specific de-staining (SeqStain) (Rajagopalan et al., 2021), and signal amplification has been achieved via primer exchange (ImmunoSABER) (Saka et al., 2019). Currently, the number of analyzed markers is theoretically unlimited, and researchers have used the technology to investigate more than 100 epitopes on the same tissue sample. Besides its excellent multiplexing capacity, these methods

rely on custom-conjugated antibodies, which is why direct fluorescence-based methods are still key technologies in the field of HMTI. Tissue-based cyclic immunofluorescence (t-CyCIF), for example, uses iterative cycles of staining with fluorescently labeled primary or secondary antibodies, fluorescence acquisition with a slide scanning microscope, and fluorophore-inactivation by chemical bleaching (Lin et al., 2018). Additional technologies, e.g., cellDIVE (Gerdes et al., 2013), IBEX (Radtke et al., 2020), and ChipCytometry (Jarosch et al., 2021, Jarosch et al., 2022, Hennig et al., 2009) developed according to the same principle. All of these methods are compatible with off-the-shelf antibodies and have been proven to detect more than 60 epitopes.

1.7.3 ChipCytometry

ChipCytometry (Jarosch et al., 2021, Jarosch et al., 2022, Hennig et al., 2009) expands on the basic principles of t-CyCIF, adding two critical technological advantages. Firstly, 32-bit high dynamic range (HDR) imaging allows a better resolution between high and low signal intensities. Secondly, a proper HDR background is recorded before every staining and subtracted from the image to receive a net-fluorescence image. In addition, tissue sections are mounted into microfluidic chips, where they can be stored for later re-interrogation and staining, adding an essential value for precious sample material. After mounting a section on a chip, cyclic immunofluorescence is performed. The cycles include i) background acquisition before staining, ii) incubation of fluorophore coupled antibodies in up to five different channels, iii) imaging with an upright, inverted microscope, rediscovering each position in each cycle, and iv) photobleaching with the build-in HBO lamp (Figure 9). Due to an easier sample handling, this technology was tested and established for cryopreserved tissue samples only, which marks a disadvantage in regards to the applicability to retrospective clinical tissue repositories that contain almost exclusively formalin-fixed and paraffin-embedded (FFPE) tissues since this is still the preferred method for long-term preservation (Gaffney et al., 2018). Still, for the discussed advantages of the technology, we have chosen the ChipCytometry technology for multiplexed imaging, also aiming for a transfer of the technology to FFPE samples to decipher the spatial context of immune cells in GvHD patients and combine this information with the depth of scRNA sequencing for accurate phenotyping.

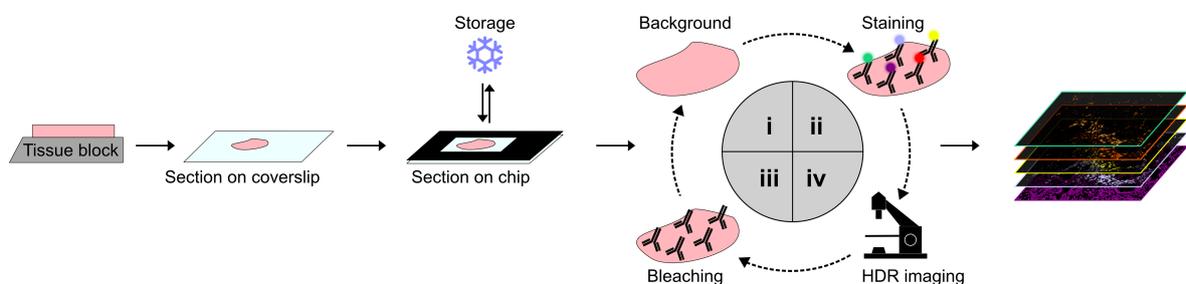


Figure 9. ChipCytometry technology.

Schematic overview of the ChipCytometry method. Sections are mounted on glass coverslips and later on chips containing a microfluidic chamber. Within the chip, the tissue can be stored, and cyclic immunofluorescence can be applied, meaning iterative cycles of background acquisition, antibody staining, imaging, and photobleaching for dye inactivation. Modified from (Jarosch et al., 2021).

2. Aim of this thesis

GvHD is a significant complication after aHSCT and limits the success of this otherwise promising treatment strategy for hematological malignancies. The disease is marked by donor cells attacking the recipient's organs, especially the skin, the gastrointestinal tract, and the liver, after T cell activation linked to the preconditioning of the patient. All primarily affected organs are seeded by commensal bacteria, and several preclinical as well as clinical studies have emphasized a connection between the microbiome and the severity of GvHD. However, the exact causal relationship between immune infiltrates, gastrointestinal GvHD, and microbiome composition remains unresolved.

In this study, we set out to characterize the immune infiltrates in GI biopsies of tumor patients who underwent aHSCT by the cutting-edge technologies of ChipCytometry multiplexed tissue imaging and scRNA seq. This experimental setting would allow obtaining the highest possible spatial and phenotypical resolution of immune infiltrates. The unique dataset should then be analyzed in relation to the patient-specific disease state and microbiome signatures. For this purpose, we specifically aimed at:

- establishing protocols for the application of ChipCytometry on FFPE samples
- adapting protocols of scRNA seq to small-sized GI biopsies
- characterizing the immune infiltrates of GI biopsies in relation to clinical parameters

ChipCytometry has some advantages compared to other multiplexed imaging approaches, but it had only been established for single cell suspensions and cryopreserved tissue samples. As large FFPE repositories of gastrointestinal biopsies were already available, it was essential to establish a protocol for ChipCytometry on FFPE tissue samples. For that, we aimed to test different methods for antigen retrieval, reduction of autofluorescent background and optimal tissue adherence. In parallel, we aimed to set up an analysis pipeline that allows the automated quantification of signal intensities and, ultimately, the analyses of immune cell infiltration differences between samples. For transcriptomic analysis of single cells from small punch biopsies, we wanted to optimize single cell extraction, freezing, and enrichment of immune cells to maximize cell recovery for the following analyses. Further, we needed to develop a good pooling strategy of multiple samples for a cost-efficient experimental setup.

After the technological establishments, we set out to generate an extensive transcriptomic and imaging dataset that would allow the investigation of GI-infiltrating immune cells in GvHD. This information was then supposed to be connected to clinically relevant GvHD biomarkers, treatment status of the patient and microbiome composition via 16 rRNA sequencing of stool samples. The output of these analyses should contribute to a better understanding of the mechanistic connections between immune cells, the microbiome, and GvHD severity, which will help to develop efficient treatment strategies against GvHD and improve the potential of HSCT for cancer treatment.

3. Materials and Methods

3.1 Materials

3.1.1 Devices

Table 1. List of devices

Device	Model	Supplier
10X Instrument	Chromium Controller	10x Genomics, Pleasanton, USA
Automated electrophoresis	2100 Bioanalyzer	Agilent, Santa Clara, USA
Automated staining system	Ventana Benchmark XT	Roche, Basel, Switzerland
Balance	EG 2200-2NM	Kern & Sohn GmbH, Balingen, Germany
	ACJ 320-4M	Kern & Sohn GmbH, Balingen, Germany
Benchtop fluorometer	Qubit 2.0	Life Technologies, Carlsbad, USA
ChipCytometry Instrument	Zellscanner ONE	Canopy Biosciences, St. Louis, USA
Counting Chamber	Neubauer improved	Paul Marienfeld & Co. KG, Lauda-Königshofen, Germany
Coverglass staining rack	-	Epredia, Michigan, USA
Flow Cytometer / cell sorter	CytoFLEX S	Beckman Coulter, Fullerton, USA
	FACS Aria II	Becton Dickinson, Heidelberg, Germany
	MoFlo Astrios	Beckman Coulter, Fullerton, USA
Incubator	HERAcell 240	Heraeus, Hanau, Germany
Laminar flow hood	HERAsafe	Heraeus, Hanau, Germany
Microscope	Axiovert 40C	Carl Zeiss, Jena, Germany
Microtome	RM2245	Leica, Wetzlar, Germany
Oven	Hybaid Shake `n` stack	Thermo Scientific, Waltham, USA
pH-meter	766	Knick, Berlin, Germany
Pressure cooker	6L Amazon Basics	Amazon, Seattle, USA
Slide scanner	BX61VS	Olympus, Tokio, Japan
	Aperio AT2	Leica, Wetzlar, Germany
Sequencer	HiSeq2500	Illumina, San Diego, USA
	NovaSeq6000	Illumina, San Diego, USA
Thermo shaker	Thermomixer comfort	Eppendorf, Hamburg, Germany
Waterbath	WNB10	Memmert, Schwabach, Germany

3.1.2 Consumables

Table 2. List of consumables

Name	Supplier	Identifier
Automat-Star glass coverslips (24 x 50 mm, No. 1)	Engelbrecht, Edermünde, Germany	K12460A1,0
Microscopy slides Superfrost® Plus	Thermo Scientific, Waltham, USA	#J1800AMNZ

3.1.3 Chemicals and reagents

Table 3. List of chemicals and reagents

Name	Supplier
Beta-mercaptoethanol	Life Technologies, Carlsbad, USA
Bovine serum albumin (BSA)	Sigma Aldrich, Taufkirchen, Germany
Collagenase IV	Sigma Aldrich, Taufkirchen, Germany
Dimethyl sulfoxide (DMSO)	Sigma Aldrich, Taufkirchen, Germany
Dulbecco's phosphate-buffered saline (DPBS)	PAN-Biotech, Aidenbach, Germany
Eosin	Morphisto, Offenbach, Germany
Ethanol 70%, 1 % MEK	Carl Roth, Karlsruhe, Germany
Ethanol absolute, 1 % MEK	Carl Roth, Karlsruhe, Germany
Ethylenediaminetetraacetic acid (EDTA)	Carl Roth, Karlsruhe, Germany
Fetal calf serum (FCS)	Sigma Aldrich, Taufkirchen, Germany
Ficoll	Sigma Aldrich, Taufkirchen, Germany
Gentamycin	Life Technologies, Carlsbad, USA
HBSS	Sigma Aldrich, Taufkirchen, Germany
Hematoxylin	Morphisto, Offenbach, Germany
HEPES	Carl Roth, Karlsruhe, Germany
Hoechst	Thermo Fisher, Darmstadt, Germany
Hydrogen peroxide (H ₂ O ₂)	Sigma Aldrich, Taufkirchen, Germany
L-Glutamine	Sigma Aldrich, Taufkirchen, Germany
Opal 520	Akoya Bioscience, Massachusetts, USA
Opal 570	Akoya Bioscience, Massachusetts, USA
Opal 650	Akoya Bioscience, Massachusetts, USA
Penicillin	Life Technologies, Carlsbad, USA
Propidium Iodide	Thermo Fisher, Darmstadt, Germany
Roticlear	Carl Roth, Karlsruhe, Germany

Materials

RPMI 1640	Life Technologies, Carlsbad, USA
Sodium Borohydrate	Sigma Aldrich, Taufkirchen, Germany
Sodium Chloride (NaCl)	Carl Roth, Karlsruhe, Germany
Sodium Citrate	Sigma Aldrich, Taufkirchen, Germany
Sodium Hydroxate (NaOH)	Sigma Aldrich, Taufkirchen, Germany
Sudan Back B	Sigma Aldrich, Taufkirchen, Germany
Tris(hydroxymethyl)aminomethane (TRIS)	Carl Roth, Karlsruhe, Germany
Tween20	Carl Roth, Karlsruhe, Germany

3.1.4 Reagent kits

Table 4. List of reagent kits

Kit	Supplier	Identifier
5' Feature Barcode Kit	10x Genomics, Pleasanton, USA	1000256
Chromium Next GEM Chip G Single Cell Kit	10x Genomics, Pleasanton, USA	1000120
Chromium Next GEM Chip K Single Cell Kit	10x Genomics, Pleasanton, USA	1000286
Chromium Next GEM Single Cell 5' Kit v2	10x Genomics, Pleasanton, USA	1000263
Chromium Next GEM Single Cell 5' Library and Gel Bead Kit v1.1	10x Genomics, Pleasanton, USA	1000165
Chromium Single Cell 5' Feature Barcode Library Kit,	10x Genomics, Pleasanton, USA	1000080
Chromium Single Cell 5' Library Construction Kit	10x Genomics, Pleasanton, USA	1000020
Chromium Single Cell Human TCR Amplification Kit	10x Genomics, Pleasanton, USA	1000252
Chromium Single Cell V(D)J Enrichment Kit, Human T Cell	10x Genomics, Pleasanton, USA	1000005
dsDNA hs assay kit	Life Technologies, Carlsbad, USA	Q32854
Dual Index Kit TN Set A	10x Genomics, Pleasanton, USA	1000250
Dual Index Kit TT Set A	10x Genomics, Pleasanton, USA	1000215
High sensitivity DNA Kit	Agilent, Santa Clara, USA	5067-4626
RNAscope Multiplex Fluorescent Reagent Kit v2	ACDBio, Newark, USA	323100
RNAscope Probe Diluent	ACDBio, Newark, USA	300041
Single Index Kit N Set A	10x Genomics, Pleasanton, USA	1000212
Single Index Kit T Set A	10x Genomics, Pleasanton, USA	1000213
UltraView Universal DAB Detection Kit	Roche, Basel, Switzerland	5269806001
ZellSafe Chip kit FFPE	Canopy Biosciences, St. Louis, USA	28050606/04

3.1.5 Antibodies

3.1.5.1 ChipCytometry**Table 5.** List of ChipCytometry antibodies

Epitope	Clone	Conjugate	Supplier	Identifier	Dilution
Annexin A1	EPR19342	PE	abcam	ab225512	1:100
anti-Rabbit	polyclonal	2nd FITC	BioLegend	406403	1:200
anti-Rabbit	polyclonal	2nd PE	BioLegend	406421	1:300
beta-Catenin	L54E2	PE	Cell Signaling	6898S	1:300
CD103	EPR4166(2)	AF488	abcam	ab225152	1:100
CD117	104D2	PE	Thermo Fisher	12-1178-41	1:100
CD123	6H6	PE	BioLegend	306006	1:150
CD133	clone 7	PE	BioLegend	372803	1:100
CD14	EPR3653	AF488	Abcam	ab133335	1:100
CD20	H1	PE	BD Bioscience	561174	1:200
CD20	H1	PerCP/Cy5.5	BD Bioscience	558021	1:25
CD3	SP7	-	Thermo Scientific	RM-9107-S1	1:150
CD4	polyclonal	AF488	R&D systems	FAB8165G	1:50
CD45	HI30	BUV395	BD Bioscience	563791	1:80
CD45	HI30	PerCP/Cy5.5	BioLegend	304028	1:50
CD45RA	HI100	BV421	BioLegend	304129	1:100
CD45RA	HI100	PE	BioLegend	304108	1:600
CD45RA	HI100	PerCP/Cy5.5	BioLegend	304122	1:100
CD45RO	UCHL1	PE	BioLegend	304206	1:150
CD56	123C3.D5	PerCP	Novus	33132PCP	1:100
CD57	HNK-1	PE	BioLegend	359611	1:100
CD68	KP1	FITC	Santa-Cruz	sc-20060	1:100
CD79a	HM47	PE	BioLegend	333503	1:100
CD8	C8/144B	PE	SantaCruz	sc53212 PE	1:50
Collagen IV	1042	AF488	Thermo Fisher	53-9871-80	1:100
Cytokeratin (Pan)	C11	AF488	BioLegend	628608	1:100
E-Cadherin	24E10	PE	Cell Signaling	7559S	1:100
Foxp3	236A/E7	PE	eBioscience	12-4777-42	1:30
GATA-3	L50-823	PE	BD Pharmingen	560074	1:50

Materials

Ki-67	B56	PE	BD Bioscience	556027	1:50
Ki-67	B56	PerCP/Cy5.5	BD Bioscience	561284	1:50
Mast Cell Chymase	CC1	PE	SantaCruz	sc-59586	1:100
Mast Cell Tryptase	G3	PE	SantaCruz	sc-33676	1:200
Muc2	SPM296	PE	Novus	34757PE	1:300
NF-kb	E379	AF488	abcam	ab190205	1:50
PD-1	NAT105	PE	BioLegend	367404	1:50
PD-L1	29E.2A3	PE	BioLegend	329706	1:200
pSTAT3	D3A7	PE	Cell Signaling	8119	1:150
SMA	1A4	eF570	eBioscience	41-9760-80	1:500
Vimentin	O91D3	AF488	BioLegend	677809	1:300
Vinculin	7F9	AF488	Invitrogen	53-9777-82	1:100

3.1.5.2 Flow cytometry

Table 6. List of flow cytometry antibodies

Epitope	Clone	Conjugate	Supplier	Identifier	Dilution
CD3	UCHT1	PE	Beckman Coulter	A07747	1:50
CD45	T29/33	Pacific Blue	Dako	PB986	1:50

3.1.5.3 Single cell RNA sequencing

Table 7. List of single-cell RNA sequencing antibodies

Epitope	Clone	Sequence	Supplier	Identifier	Dilution
TotalSeq-C0251	LNH-94/2M2	GTCAACTCTTTAGCG	BioLegend	394661	1:100
TotalSeq-C0252	LNH-94/2M2	TGATGGCCTATTGGG	BioLegend	394663	1:100
TotalSeq-C0253	LNH-94/2M2	TTCCGCCTCTCTTTG	BioLegend	394665	1:100
TotalSeq-C0254	LNH-94/2M2	AGTAAGTTCAGCGTA	BioLegend	394667	1:100
TotalSeq-C0255	LNH-94/2M2	AAGTATCGTTTCGCA	BioLegend	394669	1:100
TotalSeq-C Human Universal Cocktail V1.0	Several	Several	BioLegend	399905	-

3.1.6 Probes for fluorescence RNA *in-situ* hybridization**Table 8.** List of FISH probes

Gene	Target region	Opal dye	Supplier	Identifier
CD3E	26 - 1288	Opal 520	ACDBio, Newark, USA	553971
CTLA4	307 - 1224	Opal 650	ACDBio, Newark, USA	554341
IL10	122 - 1163	Opal 570	ACDBio, Newark, USA	602051
IFNG	80 - 1152	Opal 570	ACDBio, Newark, USA	310501

3.1.7 Buffers and media

Table 9. Composition of buffers and media

Buffer	Composition	
Antigen retrieval buffer, pH 8.5	10.00 mM	ddH ₂ O TRIS 1.00 mM EDTA
Blocking buffer	0.10 % (v/v) 5.00 % (v/v)	PBS Tween-20 Goat serum
Complete RPMI (cRPMI)	10.00 % (v/v) 0.12 % (w/v) 0.02 % (w/v) 1.00 % (v/v) 0.10 % (v/v) 0.10 % (v/v)	RPMI 1640 FCS HEPES L-Glutamine Penicillin/Streptomycin Gentamycin β-Mercaptoethanol
Digestion buffer	25.00 mM 0.10 % (w/v)	HBSS w/o Ca ²⁺ , Mg ²⁺ HEPES Collagenase IV
Digestion wash buffer	25.00 mM	cRPMI HEPES
EB buffer, pH 8.5	10.00 mM	ddH ₂ O Tris-HCl
FACS Buffer, pH 7.5	0.50 % (w/v)	PBS BSA
Freezing medium	10.00 % (v/v)	FCS DMSO
PBST	0.10 % (v/v)	PBS Tween-20
Quenching buffer	4.50 % (v/v) 24.00 mM	PBS Hydrogen Peroxide Sodium hydroxide

Materials

Sodium saline citrate buffer (20X), pH 7.0	3.00 mM	ddH ₂ O
	2.90 mM	Sodium Chloride (NaCl)
		Sodium citrate
Sudan Black B solution		70% EtOH
	0.10 % (w/v)	Sudan Black B

3.1.8 Human samples

All procedures were approved by the local ethics committee and, after informed, written consent of patients regarding the use of the tissue samples. Tissue samples were kindly provided from the repositories of Dr. Katja Steiger, Prof. Dr. Klaus-Peter Janssen, and Prof. Dr. Ernst Holler. Details can be found in the following table.

Table 10. Origin of human samples

Sample type	Source
Biopsies from aHSCT patients	University hospital of Regensburg (Prof. Ernst Holler)
Breast cancer resections	TUM Pathology Department (Dr. Katja Steiger)
Colorectal cancer resections	TUM Pathology Department (Dr. Katja Steiger)
Healthy adjacent tissues from colon resections	Molecular Tumor Biology (Prof. Dr. Klaus-Peter Janssen) at the Dept. of Surgery (TUM)
Pancreatic cancer resections	TUM Pathology Department (Dr. Katja Steiger)
PBMC samples	Voluntary, healthy donors

3.1.9 Gene Sets

Table 11. Gene sets used in scRNA seq analysis

Geneset	Source	Genes
G2M genes	(Tirosch et al., 2016)	BIRC5, TPX2, TOP2A, NDC80, CKS2, NUF2, CKS1B, MKI67, TMPO, CENPF, TACC3, FAM64A, SMC4, CCNB2, CKAP2L, CKAP2, AURKB, BUB1, KIF11, ANP32E, TUBB4B, GTSE1, KIF20B, HJURP, CDCA3, HN1, CDC20, TTK, CDC25C, KIF2C, RANGAP1, NCAPD2, DLGAP5, CDCA2, CDCA8, ECT2, KIF23, HMMR, AURKA, PSRC1, ANLN, LBR, CKAP5, CENPE, CTCF, NEK2, G2E3, GAS2L3, CBX5, CENPA
S genes	(Tirosch et al., 2016)	MCM5, PCNA, TYMS, FEN1, MCM2, MCM4, RRM1, UNG, GINS2, MCM6, CDCA7, DTL, PRIM1, UHRF1, MLF1IP, HELLS, RFC2, RPA2, NASP, RAD51AP1, GMNN, WDR76, SLBP, CCNE2, UBR7, POLD3, MSH2, ATAD2, RAD51, RRM2, CDC45, CDC6, EXO1, TIPIN, DSCC1, BLM, CASP8AP2, USP1, CLSPN, POLA1, CHAF1B, BRIP1, E2F8, HMGB2, CDK1, NUSAP1, UBE2C
Teff genes	(Höllbacher et al., 2020)	ACE, ACER1, ACSL6, ACSM1, ACVR1B, ACVR1C, ADAM23, ADAMTS10, ADAMTS17, ADCY9, ADGRE1, ADGRE5, AFF3, AGAP1, AGAP3, AGPAT9, AIF1, AJUBA, AK5, AKR1E2, ALOX5, ALPK1, AMIGO1, AMPD2, ANK3, ANKRD42, ANXA1, APBA2, APOL3, APP, AREG, ARHGAP32, ARHGAP39, ARHGAP40, ARL4A, ARMC12, ARMCX2, ARRD4, ASB13, ASTL, ATF3, ATHL1, AXIN2, AZIN2, B3GNT9, B4GALNT4, BACH2, BAIAP3, BAMBI, BCL2A1, BCL7A, BEGAIN, BFSP1, BHLHE40, BRSK1, BTBD11, C11orf74, C16orf54, C1orf115, C1orf162,

C1orf228, C1orf233, C2orf40, C3orf52, CACFD1, CACNA11, CAMK2N1, CAND2, CASS4, CBR3, CCDC112, CCDC184, CCND1, CCND3, CCR7, CCSAP, CD226, CD300A, CD40LG, CD69, CDC42EP4, CDH23, CDKN1A, CDS1, CERS6, CFH, CFL2, CHD7, CHN2, CHPF, CILP2, CKAP4, CLIC6, CLU, CMPK2, CNN3, COL1A1, COL27A1, COL6A1, CR1, CR2, CTSL, CXCL16, CYB561, CYP4F12, DCHS1, DDIT4, DENND5A, DHRS3, DISC1, DLG4, DLL1, DMXL2, DNAJC6, DOCK3, DPYSL4, DSE, DST, DTX1, DUSP1, DUSP2, DUSP5, DUSP8, DYNC2H1, DYNLT3, ECHDC3, EDA, EFHC2, EGR2, EHD4, ELMO2, ENPP2, EPHA1, EPHA4, EPS8, ETNK2, ETS2, EXD2, F2RL1, FADS3, FAM160A1, FAM19A1, FAM213A, FAM26F, FAM65A, FAM89A, FBP1, FBXL8, FGF9, FHL1, FKBP2, FKBP5, FLOT1, FOS, FOSB, FZD3, FZD8, GADD45G, GAS2L1, GAS7, GCLM, GCSAM, GIMAP7, GIMAP8, GLB1L2, GLUL, GNAO1, GNG7, GPR150, GPR160, GPR35, GRASP, GRB10, GREM2, GRIP1, GSAP, H1FO, HBEGF, HDGFL3, HID1, HIPK2, HOOK1, HS6ST1, HSF4, ID1, ID2, IER3, IER5, IFI44L, IFITM5, IGFBP3, IGFBP4, IL15, IL2, IL4I1, IL4R, IL5RA, IL7R, INPP4B, INSR, IQCK, IQSEC2, IRF2BP2, ISM1, ITGA1, JAKMIP3, JAM3, JUN, KCNQ1, KCTD3, KDM6B, KIAA2022, KIF13A, KIF19, KLF4, KRT2, LARGE, LATS2, LDOC1L, LMO4, LMTK3, LONRF1, LPGAT1, LPIN2, LRP1, LRP6, LRRC6, LYNX1, LYPD3, LYSMD2, MAGI3, MAML3, MAML213A, MAN1C1, MAP9, MAPRE3, MAPT, MARVELD1, MATK, MBNL3, MBOAT2, MCOLN2, MCOLN3, ME3, MEST, METTL1, MFAP3L, MID1IP1, MID2, MMP11, MMP15, MMP9, MPP2, MPP7, MTUS1, MTUS2, MYADM, MYO6, NAP1L3, NBEA, NBL1, NCOA7, NELL2, NEO1, NEU4, NFIA, NFKB1, NFKBIA, NINJ1, NINL, NKX3-1, NLGN2, NLRP3, NLRP6, NME4, NOL3, NPTXR, NR3C2, NR4A1, NR4A2, NT5C3B, NUDT11, OSBPL5, OSM, OTUD7A, P3H3, PARD3B, PARK2, PCSK5, PDE4D, PDZD4, PELI2, PFN2, PHLDB2, PIFO, PIK3C2B, PIM3, PLAC8, PLAG1, PLAUR, PLEKHA7, PLEKHG5, PLK2, PLXDC1, PLXNA4, PLXND1, PM20D2, PMPA1, PNP, POMGNT2, PPM1L, PPP4R4, PRKCA, PRKCE, PROSER2, PRRT1, PRRT4, PRSS22, PTGER4, PTK2, PTPRK, PVRL1, PVRL2, PVRL3, PWWP2B, RAB20, RAB27B, RALA, RALGPS2, RARA, RARG, RASD1, RASGEF1A, RASGEF1B, RASGRF2, RASSF2, RASSF6, RBM11, RBMS1, RELB, RGAG4, RGL1, RGS16, RHOB, RHOU, RIMKB, RIPK4, RNASE6, RND1, RNF122, RNF130, RNF144A, RNF207, ROBO3, RRAGD, RRAS2, RSAD2, RUFY3, S1PR1, SAV1, SBF2, SCML1, SCML4, SDK2, SEC14L2, SECTM1, SEMA4C, SERPINB6, SFMBT2, SFN, SGK1, SH2B3, SHF, SIAH3, SIPA1L2, SKIL, SLC22A17, SLC22A23, SLC22A31, SLC29A2, SLC30A4, SLC36A4, SLC39A8, SLC40A1, SMAD5, SMIM3, SNTB1, SOAT2, SORBS3, SOS2, SOX8, SPATA6, SPEF2, SPEG, SPG20, SPRY1, SREBF1, SSBP2, SSPO, ST6GALNAC1, ST6GALNAC2, STARD10, STOM, STXBP1, SUS4, SVIL, SYDE2, SYNLM, TAF4B, TANC1, TBC1D16, TCM2N, TCEA3, TCEAL2, TCF7, THEMIS, TIMP1, TLE1, TLR2, TMEM121, TMEM170B, TMEM184B, TMEM200A, TMEM30B, TMEM45B, TMEM63C, TMEM71, TMEM88, TMIGD2, TMOD2, TNFAIP3, TNFRSF10B, TNFRSF13C, TOP1MT, TPBGL, TPM2, TRABD2A, TRADD, TRAT1, TRIM2, TRIP10, TSPAN18, TSPAN6, TSPAN9, TXNRD3, UBE2E2, ULBP2, UNC13B, USP2, USP44, USP46, USP6NL, UST, VEGFB, VIPR1, WDR86, WNT1, WNT10B, WNT11, WNT7A, WWC2, XBP1, YES1, ZBED3, ZBTB10, ZBTB18, ZBTB42, ZBTB47, ZDHHC11, ZDHHC11B, ZFP36, ZFP36L2, ZFYVE9, ZMYND15, ZNF165, ZNF208, ZNF219, ZNF358, ZNF418, ZNF439, ZNF467, ZNF502, ZNF516, ZNF607, ZNF618, ZNF629, ZNF703, ZNF711, ZNF814, ZNF827, ZSCAN18, ZSWIM5

Treg genes

(Höllbacher et al., 2020)

ABHD11, ACTA2, ADAM28, ADAT2, ADRBK2, ADTRP, AKAP2, AKAP5, AKIP1, AKR1C3, ANKDD1B, ANKS1B, APOBEC2, AQP7, ARAP3, ARHGAP11A, ARHGAP11B, ARHGAP25, ARHGEF12, ASF1B, ASIC1, ASPM, ATP1B1, ATP6V0A1, BARD1, BFSP2, BLM, BLZF1, BUB1, C15orf53, C16orf74, CACNB2, CAMK1, CARD16, CARD17, CASC5, CASP1, CASP8, CBR4, CCDC141, CCM2, CCR3, CCRL2, CD27, CD79A, CD79B, CDC45, CDCA2, CDCA7L, CDHR3, CDK14, CDKL4, CDKN2A, CEACAM4, CENPE, CENPF, CENPJ, CEP128, CEP55, CHST2, CIITA, CIT, CLDN4, CPA5, CPNE2, CSF2RB, CTLA4, CTNNA1, CTTNBP2NL, CXCR6, DCLRE1A, DCP1B, DDIT2, DEPDC1, DOCK6, DSCC1, DUSP10, DUSP4, E2F7, ENTPD1, EPSTI1, ERI1, ETV7, F5, FAM110C, FAM174B, FAM179A, FAM184A, FAM19A2, FANK1, FBXO18, FCRL1, FDXR, FHL3, FITM2, FOXP3, GALNT3, GCNT1, GEN1, GINS2, GIPR, GJB6, GK, GLCE, GLRX, GNG8, GNRH1, GPLD1, GPR155, GPR19, GSN, GSTA4, GTSF1L, GXYLT1, HAVCR1, HDAC9, HECW1, HELLS, HES1, HIBCH, HIST4H4, HLA-DMA, HLA-DMB, HLA-DOA, HLA-DPA1, HLA-DPB1, HLA-DRA, HLA-DRB1, HLA-DRB5, HMCN1, HOXB2, HOXB3, HPGD, HS3ST3B1, HSPA1L, IKZF2, IKZF4, IL12A, IL1R1, IL1R2, IL2RA, IL32, INPP1, INPP5F, IRAK3, IRF4, ITGB8, JSRP1, KHDC1, KIAA1841, KIF14, KIF18A, KLHL2, LAMC3, LAYN, LGR4, LIMD1, LIN7A, LMCD1, LRG1, LRRC2, LRRC32, LSP1, MACC1, MARCH3, MELK, MEOX1, METTL7A, MFG8, MGST2, MPST, MYB, MYL6B, MYOF, NABP1, NCAPG, NDC80, NEBL, NEK2, NINJ2, NTHL1, NTNG2, NUF2, NUSAP1, OAS1, OGT, OR2A7, PAFAH2, PBK, PCED1B, PECCR, PGM2L1, PIK3R3, PITPNM2, PLGLB1, PLS3, PM20D1, PMAIP1, PMCH, POLE2, PRDM8, PRKAG1, PROB1, PRR11, PRR33, PSD3, PTGIR, PTC1, PYHIN1, RAB31, RAB37, RACGAP1, RAD51AP1, RASGRP4, RGD2, RTKN2, RTP4, RYR1, SASS6, SELP, SEMA3G, SFT2D1, SGMS1, SH3RF2, SHCBP1, SHMT2,

Materials

		SHTN1, SIM1, SIRPG, SKA1, SLC12A6, SLC16A10, SLC1A4, SLC2A8, SLC36A1, SLC43A1, SLC46A3, SLC9A7, SLFN14, SMC6, SMTN, SOX4, SPATS2L, SPC24, SPON2, SRGAP3, ST8SIA4, ST8SIA6, STAC, STAM, STRIP2, SUOX, SWAP70, TCF19, TCHP, TEP1, TIAM1, TICRR, TIGIT, TIMD4, TLDC2, TMEM169, TMEM244, TNFRSF13B, TNFRSF19, TNFRSF8, TNFRSF9, TNIP3, TOP2A, TOR4A, TOX, TOX2, TPP1, TPX2, TRIB1, TRPV6, TSHR, TTC22, TTN, TYMS, UBALD1, UTS2, VANGL1, VAV3, WBSCR27, WSCD2, XKRX, XYLB, YPEL2, ZBP1, ZBTB32, ZC3H12C, ZNF662, ZNF80
Treg suppression genes	This work	FOXP3, IL2RA, CTLA4, CD27, TNFRSF4, TNFRSF18, TNFRSF9, TIGIT, BATF, CCR4, CXCR3
Y-chromosome genes	ENSEMBL	AMELY, BCORP1, BPY2, BPY2B, BPY2C, CDY1, CDY1B, CDY2A, CDY2B, DAZ1, DAZ2, DAZ3, DAZ4, DDX3Y, EIF1AY, HSFY1, HSFY2, KDM5D, NLGN4Y, PCDH11Y, PRKY, PRORY, PRY, PRY2, PRYP3, PRYP4, RBMY1A1, RBMY1B, RBMY1C, RBMY1D, RBMY1E, RBMY1F, RBMY1J, RPS4Y1, RPS4Y2, SLC9B1P1, SRY, TBL1Y, TGIF2LY, TMSB4Y, TSPY1, TSPY10, TSPY2, TSPY3, TSPY4, TSPY8, TTTY10, TTTY12, TTTY13, TXLNGY, USP9Y, UTY, VCY, VCY1B, XKRY, XKRY2, ZFY
KEGG 2021 GvHD	KEGG	KIR3DL3, GZMB, HLA-A, HLA-B, HLA-C, HLA-DMA, HLA-DMB, HLA-DOA, HLA-DOB, HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQA2, HLA-DQB1, HLA-DRA, HLA-DRB1, HLA-DRB3, HLA-DRB4, HLA-DRB5, HLA-E, HLA-F, HLA-G, IFNG, FAS, IL1A, IL1B, IL2, FASLG, IL6, KIR2DL1, KIR2DL2, KIR2DL3, KIR3DL1, KIR3DL2, KLRC1, KLRD1, PRF1, KIR2DL5A, TNF, CD28, CD80, CD86

3.1.10 Software

Table 12. List of software and python packages

Software	Version	Developer/Publication	Website
Affinity Designer	1.9	Serif Europe Ltd., Nottingham, UK	https://affinity.serif.com/de/designer/
anndata	0.7.6	(Virshup et al., 2021)	https://github.com/scverse/anndata
Aperio ImageScope	12.4.0	Leica, Wetzlar, Germany	https://www.leicabiosystems.com/de-de/digitalpathologie/verwaltung/aperio-imagescope/
arcasHLA	0.4	(Orenbuch et al., 2020)	https://github.com/RabadanLab/arcasHLA
BaSiC	-	(Peng et al., 2017)	https://github.com/marrlab/BaSiC
Cellranger	6.0.2	10x Genomics, Pleasanton, USA	https://support.10xgenomics.com/single-cell-gene-expression/software/pipelines/latest/installation
diffxpy		David Fischer & Florian Hölzlwimmer	https://github.com/theislab/diffxpy
fcswrite	0.5.2	Zellmechanik Dresden, Dresden, Germany	https://github.com/ZELLMECHANIK-DRESDEN/fcswrite
Fiji	1.53	(Schindelin et al., 2012)	https://imagej.net/software/fiji/
FlowJo	10	FlowJo LLC, Ashland, USA	https://www.flowjo.com/
Freebayes	1.3.1	(Garrison and Marth, 2012)	https://github.com/freebayes/
leidenalg	0.8.3	(Traag et al., 2019)	https://github.com/vtraag/leidenalg
matplotlib	3.3.4	(Thomas et al., 2021)	https://matplotlib.org/

Microsoft Office	2019	Microsoft, Redmond, USA	-
numpy	1.19.5	(Harris et al., 2020)	https://numpy.org/
pandas	1.2.3	(Jeff et al., 2021)	https://pandas.pydata.org/
Prism	9	Graphpad, La Jolla, USA	https://www.graphpad.com/scientific-software/prism/
samtools	1.9	(Li et al., 2009)	http://www.htslib.org/
scanpy	1.8.1	(Wolf et al., 2018)	https://github.com/scverse/scanpy
Scikit-bio	0.5.6		https://github.com/biocore/scikit-bio
scikit-learn	0.24.1	(Pedregosa et al., 2011)	https://scikit-learn.org/stable/
scipy	1.6.1	(Virtanen et al., 2020)	https://scipy.org/
scirpy	0.10.0	(Sturm et al., 2020)	https://github.com/scverse/scirpy
scVelo	0.2.3	(Bergen et al., 2020)	https://github.com/theislab/scvelo
seaborn	0.11.1	(Waskom, 2021)	https://seaborn.pydata.org/
Souporcell	2.0	(Heaton et al., 2020)	https://github.com/wheaton5/souporcell
StarDist	0.3.0	(Weigert, 2020)	https://github.com/stardist/stardist-imagej
statsmodels	0.12.2	(Seabold and Perktold, 2010)	https://www.statsmodels.org
subset-bam	1.0	10x Genomics, Pleasanton, USA	https://github.com/10XGenomics/subset-bam
umap	0.4.6	(McInnes et al., 2018)	https://github.com/lmcinnes/umap
vcftools	0.1.16	(Danecek et al., 2011)	https://github.com/vcftools/vcftools
Velocity	0.17	(La Manno et al., 2018)	https://github.com/velocity-team/velocity.py

3.1.11 Code

All original code used and generated in this thesis has been deposited at the Github repository (<https://github.com/SebastianJarosch/ChipCytometry-Image-Processing>). The software release v1.1.0 available at Zenodo (SebastianJarosch, 2022) was used within this thesis, further updates on the pipeline will be made available in future releases.

3.2 Methods

3.2.1 Generation of single-cell suspensions

3.2.1.1 PBMC isolation from human blood samples

Fresh blood was diluted 1:5 in PBS before PBMC isolation. Leukosept tubes containing 15 ml Ficoll were centrifuged (1000 x g, RT, 2 min) before 35 ml of diluted fresh blood were added on top. After centrifugation (1000 x g, RT, 15 min), the leukocyte layer was collected and transferred to a new tube. Finally, the cells were washed with PBS and centrifuged (1500 rpm, RT, 7 min) before processing with either staining or freezing the generated cell suspensions. A maximum density of 2×10^7 cells/ml was frozen in freezing medium at -80°C (freezing speed of $1^\circ\text{C}/\text{min}$).

3.2.1.2 Generation of single-cell suspension from human gastrointestinal biopsies

Biopsies were transferred into pre-warmed (37°C) HBSS medium (Hank's Balanced Salt Solution w/o $\text{Mg}^{2+}/\text{Ca}^{2+}$) as soon as possible after resection. Tissue pieces were disrupted mechanically, transferred into 5 ml digestion buffer, and vortexed. Digestion takes place at the thermo shaker for twice 15 min (37°C , 220 rpm), vortexing the solution between the incubation steps. The resulting suspension was filtered using a $100 \mu\text{m}$ nylon mesh into a 50 ml Falcon tube, and the digestion was stopped by the addition of 40 ml digestion wash buffer. Centrifugation (1500 rpm, RT, 7 min) was followed by another washing step with wash buffer before the cells were resuspended in $250 \mu\text{l}$ freezing medium and frozen at -80°C (freezing speed of $-1^\circ\text{C}/\text{min}$).

3.2.1.3 Thawing single-cell suspensions

Each vial of frozen cell suspension was resuspended in 10 ml pre-warmed (37°C) cRPMI by pipetting back and forth the warm medium. Centrifugation (1500 rpm, 4°C , 7 min) was followed by washing with cold FACS buffer before staining the cells for flow cytometry.

3.2.2 Establishment of ChipCytometry on FFPE tissues

3.2.2.1 Antigen retrieval testing

Slides or coverslips mounted with FFPE tissue samples were either heated for 10, 20, or 30 min using a staining dish containing the antigen retrieval buffer and a water bath at a buffer temperature of 90°C or directly heated in a pressure cooker containing the retrieval buffer for 10 min at $\sim 120^\circ\text{C}$. Staining was performed overnight (see 3.2.3.2) for signal-to-noise ratio (SNR) evaluation (see 3.2.2.4).

3.2.2.2 Autofluorescence treatment tests

Autofluorescence quenching by Sudan Black B, sodium borohydride, and photobleaching was tested. For Sudan Black B treatment, sections were incubated for 5, 10, 20, or 30 min in Sudan Black B solution

before they were washed extensively, firstly with 70 % EtOH, ensuring to remove all remaining reagent from the tissue, and subsequently with PBST. For sodium borohydride, sections were incubated three times for 10 min in 0.1 % (m/v) solution in PBS and washed extensively with PBST. Photobleaching was conducted by incubating the sections in PBST underneath a white light source for 30 min. During this incubation time, the buffer was exchanged every 5 min to minimize the heating of the tissue during the process.

3.2.2.3 Evaluation of tissue section adherence

Tissue sections were mounted on either glass coverslips, polysilane-coated microscopy slides, or pre-treated coverslips (kindly provided by the Zellkraftwerk company) as tissue supplies, treated with different antigen retrieval intensities (no AR, 20 min 90°C water bath, 10 min 120°C pressure cooker) and processed for tissue integrity scoring via H&E staining (see 3.2.2.5)

3.2.2.4 Signal-to-Noise (SNR) ratio evaluation

Staining quality analysis was done by acquiring two areas of tissue, each covering five positions. The contrast and background of the HDR image were adjusted (same parameters for each condition), and images were exported as 16-bit grayscale .tiff images. The mean intensity measurements for SNR evaluation were performed with ImageJ with slightly different strategies for the areal, non-bleachable pan-cytokeratin staining, and the bleachable surface marker stainings: For pan-cytokeratin staining, crypts and background regions (lamina propria regions between the crypts) were segmented as freehand selections, and the SNR was calculated as crypt intensity divided by the mean of all background intensities. For bleachable surface markers, an additional background after photobleaching was recorded. The surface of cells was selected as a freehand line (3-pixel thickness) in ImageJ, and SNR was calculated as the staining intensity divided by the background intensity for each cell.

3.2.2.5 Tissue integrity scoring

Tissue integrity evaluation was performed by Dr. Sabrina Sarker and Dr. Katja Steiger on a scale from 0 (no tissue loss) to 3 (complete loss of tissue), based on Hematoxylin Eosin (H&E) staining. Slides or coverslips were first incubated in Mayer's Hematoxylin solution for 6 min, washed in tap water, and incubated in Eosin (1%) for another 6 min. After washing in tap water, dehydration of the slides or coverslips was achieved by incubation in 70% ethanol for 1 min, followed by 2 x 5 min absolute ethanol and 3 x 5 min Xylene incubation. Mounted sections were scanned with a slide scanner for subsequent transfer to the pathology department for integrity scoring.

3.2.2.6 Manual counting

All counting procedures were performed blinded by two scientists, and mean values were used for plotting and further analysis. Cells or positive staining events were evaluated only considering intact cells with an ordinary shape and complete surface marker staining.

3.2.2.7 Computer-assisted image analysis of IHC staining

Automated IHC and subsequent analysis were performed according to pathology standards by Dr. Sabrina Sarker and Prof. Dr. Katja Steiger from the pathology department. For more details on the methodology, refer to (Jarosch et al., 2021).

3.2.2.8 IHC staining within a ChipCytometry chip

For parallel ChipCytometry and IHC staining, IHC was performed within the chip after multiplexed immunofluorescence staining. After rinsing with tap water, 1 ml of Mayer's Hematoxylin solution was transferred to the chip and incubated for 6 min. The chip was washed with 10 ml tap water before incubation with 1 ml Eosin (1%) was performed for 6 min. The washing procedure with tap water was repeated, the chip was rinsed with 5 ml PBS and images were acquired on a slide scanner.

3.2.2.9 Fluorescence RNA *in-situ* hybridization (FISH)

RNA *in situ* hybridization in combination with ChipCytometry was performed as described before (Jarosch et al., 2022), using the RNAscope Multiplex Fluorescent Kit (ACD Bio). Briefly, after probe hybridization, the chip was loaded, a background was acquired, and the hybridized probes were finally detected with opal dyes. Opal 520, Opal 570, and Opal 650 were used for detection in the FITC, PE, and PerCP channels, respectively.

3.2.3 ChipCytometry on human FFPE tissues from gastrointestinal biopsies

ChipCytometry of human FFPE biopsies was performed according to the procedure described by us (Jarosch et al., 2021, Jarosch et al., 2022). Briefly, tissue sections were rehydrated on coverslips, and antigen retrieval was performed using TRIS-EDTA buffer (pH 8.5). Sections were then transferred to CellSafe chips, and cyclic immunofluorescence with photobleaching was performed within the chip. A detailed description of the methods can be found in the following section.

3.2.3.1 Preparation of FFPE tissues

Pre-cooled FFPE blocks (-10°C) were sectioned at a thickness of 4-5 μm , transferred to a water bath (40°C), and ultimately mounted on glass coverslips. The coverslips were dried for at least 10 h before paraffin melting was performed overnight at 60°C following an additional incubation for 30 min at 70°C to melt completely. After melting, coverslips were immediately immersed in Xylene for 10 min, followed by washing in two other Xylene-containing dishes for 10 min and 2 x 10 min absolute ethanol incubation. Coverslips were sequentially immersed in staining dishes containing 90% ethanol, 70% ethanol, 50% ethanol, and tap water for 5 min each to achieve rehydration. The antigen retrieval was performed in a dish with antigen retrieval buffer, placed within a water bath at 90 °C buffer temperature for 20 min. The coverslips were transferred into PBS at RT, before they were loaded onto the chip according to the manufacturer's recommendations. Chips were rinsed with PBST. Sections were blocked using 500 μl

blocking buffer for 1 h at RT and rinsed with 500 μ l Sudan Black B solution for autofluorescence quenching. After 10 min Sudan Black B incubation, chips were extensively washed with 70% EtOH and PBST. A first scan of the chip area containing the tissue sample was performed in the FITC channel (highest autofluorescence) to assess tissue quality and select the desired positions for multiplexed imaging.

3.2.3.2 Staining, imaging, and bleaching

After background acquisition to assess autofluorescence for each channel of the scheduled antibody mix, chips were rinsed with PBST to prepare them for subsequent staining. Up to five fluorophore-conjugated primary antibodies (Table 5) were pooled in PBST and centrifuged (14.000 rpm, 4 °C, 10 min), before staining to avoid the transfer of dye aggregates. Chips were incubated with 300 μ l of the antibody cocktail overnight at 4 °C, followed by extensive washing with 15 ml PBST. For staining with unconjugated primary antibodies followed by detection with a fluorophore-conjugated secondary antibody, the latter was incubated for 2 h at RT. Nuclei staining (Hoechst 33342) was incubated for 5 min at RT before acquisition. For more details on incubation times, dilutions, and tested antibodies, refer to section 3.1.5.1, Table 5 and (Jarosch et al., 2021). For final acquisition, exposure times were reduced compared to fresh frozen tissues due to the intrinsic higher autofluorescence of FFPE tissues: 300 ms for PE, 500 ms for FITC, 300 ms for PerCP-Cy5.5, 50 ms for BV421, and 1000 ms for BUV395. Photobleaching was performed directly after imaging by exposing each tissue position for 20 seconds to white light from the microscope build-in HBO lamp with a 364 nm long-pass filter to protect epitopes from UV light-induced damage.

3.2.3.3 Data analysis

Data analysis was performed based on an ImageJ script developed in this thesis for automatic signal quantification from ChipCytometry images. Briefly, images were first stitched and corrected for shading using the BaSiC plugin. Cells were segmented on the nuclei staining, and mean intensities were calculated for each cell and marker on spatial spillover-corrected images. The resulting matrix was then converted into an FCS file using the `fcswrite` function in python before populations were gated and quantified. For detailed information, refer to (Jarosch et al., 2021, Jarosch et al., 2022).

3.2.3.4 Distance measurements

Distances between pre-defined cell populations were calculated according to the coordinates of two individual cells using the following formula (Pythagorean theorem):

$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

The distance between a pair of cells is thereby defined by the square root of the summed squares of x and y coordinate differences between the cells.

3.2.4 Flow cytometry

A maximum of 5×10^6 cells from colon tissue single-cell suspensions (see 3.2.1.2) were stained per 50 μl master mix containing the respective antibodies (Table 6) per well of a V-bottom 96-well plate. Staining was performed for 30 minutes at 4 °C in the dark. Cells were washed twice with FACS buffer and either analyzed by flow cytometry using a CytoFLEX instrument or sorted using the MoFlo Astrios/FACS Aria II cell sorter. All centrifugation steps with cells were performed at 1500 rpm at 4 °C.

3.2.5 Single-cell RNA sequencing

3.2.5.1 Cell staining and sorting

A target 96-well V-bottom plate was coated with 250 μl FCS per well for at least 4 hours at 37 °C in the incubator. Before cell sorting, the FCS from the sorting plate was removed and replaced by 100 μl FACS buffer per well. A master mix containing fluorescently labeled and eventually DNA-barcoded antibodies was prepared and centrifuged (14.000 rpm, 4°C, 10 min) to remove antibody aggregates. After thawing single-cell suspensions (see 3.2.1.3), cells were resuspended in 100 μl of the master mix and stained for 30 minutes at 4 °C in the dark. Cells were washed two times with FACS buffer and filtered. PI (1:200) was added immediately before the sorting procedure, and living (PI⁻) CD45⁺ cells were finally sorted. Cells from individual samples were pooled during the sort resulting in one well containing pooled samples for one scRNA seq reaction. Cells from the same patient were always kept in separate pools.

3.2.5.2 Preparation of single-cell RNA libraries

After sorting the cells, they were centrifuged (1500 rpm, 4°C, 3 min), and the supernatant was carefully removed. Each cell pellet was resuspended in the master mix for reverse transcription (containing 37.8 μl of water to account for the volume of cell suspension) before 70 μl of the cell suspension was transferred to the Chromium Next GEM chip. The pellet integrity was checked under the microscope after each step to ensure that all cells were loaded onto the chip. Subsequently, 10x experiments have been performed according to the manufacturer's protocol. Briefly, after singularization of cells into Gel Beads-in-emulsion (GEMs), cells were lysed, and reverse transcription was completed within the GEMs. At this point, GEMs can be broken up and cDNA pooled since each molecule contains its specific combination of cell barcode and UMI. Following cDNA amplification, size selection was performed using SPRI beads to split cDNA molecules resembling antibody barcodes (short molecules, around 180 bp) and molecules derived from the mRNA (longer molecules, > 200 bp). Target enrichment for TCR sequences was accomplished by a nested PCR on the cDNA fraction using specific primers for the constant region of the TCR. Fragmentation, adapter ligation, and index PCR were performed on a fraction of the cDNA and the enriched cDNA to obtain the gene expression and VDJ libraries, respectively. The surface protein library was also generated via index PCR. All libraries were purified using SPRI beads to remove primers and molecules of irregular size. Quality control and fragment size analysis have been performed with a

High Sensitivity DNA Kit on a Bioanalyzer 2100 as recommended in the protocol, and library concentration was quantified with the Qubit dsDNA hs assay kit.

3.2.5.3 Sequencing

The different libraries (see 3.2.5.2) have been pooled according to their minimal required read counts (20.000 reads/cell for gene expression libraries, 5.000 reads/cell for cell surface libraries, and 5.000 reads/cell for VDJ libraries). Concentrations have been calculated using the following formula:

$$c_{(nM)} = \frac{c_{(ng/\mu l)}}{F \times size}$$

The concentration in ng/μl is determined from the Qubit DNA quantification, the size of the library is determined from the Bioanalyzer measurement, and the size factor (F) is set as 660 g/mol as DNA weight per base pair. Individual samples from one experiment have been pooled according to the cell number per sample, resulting in the following formula defining the dilution factor (DF) for each individual library:

$$DF = \frac{\left(\frac{c_{end}}{c_{(nM)}} \times \frac{r_{library}}{r_{total}} \times \frac{N_{library}}{N_{total}} \right)}{n_{libraries}}$$

c_{end}	= target concentration of the library pool
$r_{library}$	= required reads for the library
r_{total}	= sum of required reads for all libraries
$N_{library}$	= expected number of cells for the library
N_{total}	= sum of expected cell numbers for all libraries
$n_{libraries}$	= total number of libraries

Illumina paired-end sequencing was performed on the pooled libraries according to the total required reads with 150 cycles either on a Hiseq2500 or a Novaseq6000 instrument.

3.2.5.4 Sample demultiplexing using HLA information

Sample demultiplexing according to genetic variation was performed using souporecell (Heaton et al., 2020), resembling a pipeline of processing steps to finally achieve demultiplexing into n clusters. Briefly, the reads from the annotation file are first re-mapped to the reference before candidate variants (SNPs) are called using freebayes. After that, the variant alleles were quantified for each cell, and cells clustered according to this variant genotype. In the last step, doublets are inferred from the genotype information and the results (cell-cluster association) can be used for demultiplexing the transcriptomic data. For detailed information on the clustering, refer to (Heaton et al., 2020).

The more complicated part than clustering is the annotation of clusters back to individual patients by the HLA information. Unfortunately, HLA typing from scRNA seq data is not as reliable as from genomic data since many intronic regions (which often contain the minor differences between the HLA alleles) are missing in the NGS sequencing from mRNA molecules. We, therefore, developed a scoring system, allowing for donor reallocation using the patients' HLA information and the predicted HLA genotype

inferred with arcashHLA (Orenbuch et al., 2020). The scoring formula is represented by four sub-scores, calculated for each pair of patients and clusters of a pool: (1) a 4-digit HLA matching score, (2) a 2-digit HLA matching score, (3) a gender-related score, and (4) a cell number related score. The HLA matching scores are calculated by the number of allele matches between cluster and patient, subtracted by half of the number of matches with other patients to account for specificity in each allele. The 2-digit score is calculated as the 4-digit score but is weighted in half. The 2-digit score helps to gain associations between genotype and prediction due to the low resolution of HLA genotyping from scRNA seq data. The gender is accounted for by the male score of a cluster relative to the pool, multiplied by 1 for a male patient and -1 for a female patient. A male cluster would be expected to have a positive score, and a female cluster would be expected to have a negative score, resulting in a positive score in case of a gender match. The last information, which can be used for demultiplexing, is the number of cells that have been sorted in relation to the number of cells for a cluster. The whole score is divided by the positive difference between the cell number in the cluster and 10 % of the sorted cells but maximally divided by 1. These considerations lead to the following formula:

$$\left((N_{patient (4-digit)} - 0.5 \times N_{others (4-digit)}) + 0.5 \times (N_{patient (2-digit)} - 0.5 \times N_{others (2-digit)}) \right) + ((m_{cluster} - m_{pool}) * g) \times \frac{1}{\max\{(0.01 \times \sqrt{(c_{observed} - 0.1 \times c_{sorted})^2}), 1\}}$$

N = Number of allele matches

m = Male gene score

g = Gender (male = 1, female = -1)

c = Cell number

Having calculated the matrix of scores between donors and clusters, allocation can be performed. To solve this logical problem, scores are sorted in descending order, and allocation is done from the highest to the lowest score, stopping at the point where all clusters are annotated to a patient.

3.2.5.5 Data analysis

Annotation against the human reference genome (GRCh38/Ensembl 98) and the related VDJ reference was performed using Cell Ranger. All subsequent analysis has been performed using SCANPY (Wolf et al., 2018). The control dataset from the human gut cell atlas (James et al., 2020) was imported from the Cell Ranger output files, down-sampled to a maximum of 1000 cells per sample to fit the biopsy sizes, and included in all pre-processing steps in parallel to our samples. General data pre-processing was performed according to good practice in scRNA seq analysis (Luecken and Theis, 2019) with the following modifications: since the dataset contains plasma cells that are bigger in size compared to the other cell types, filtering was performed according to the factor between the number of counts and the number of genes per cells. This allowed excluding cells lying in between plasma cells and other leukocytes on the plot (Figure 10).

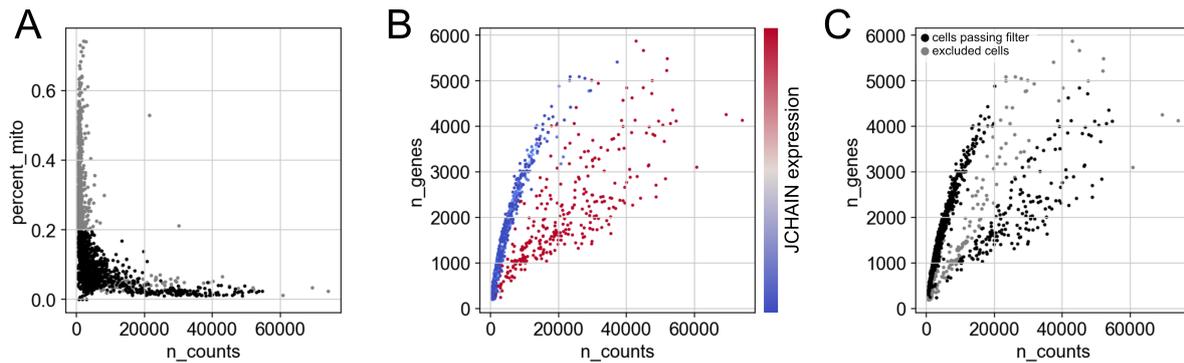


Figure 10. Filtering of cells in the scRNA sequencing experiment.

(A) Cells containing less than 20% mitochondrial genes were filtered. (B) Expression values of the JCHAIN gene, which is specific for plasma cells, show the separation of these cells on the number of counts vs. the number of genes plot. (C) Cells passing filter (black) and excluded (grey) cells are depicted. The following filters were applied: number of genes < 5,000, number of counts < 60,000, number of counts/number of genes < 5 OR number of counts / number of genes > 9, fraction mitochondrial genes < 0.2.

The number of counts, the fraction of mitochondrial genes, and cell cycle scores were regressed out, and batch effect correction was performed using batch-balanced k-nearest neighbors (bbknn). Donor reallocation was performed using souporecell and arcasHLA (see 3.2.5.4) or hashsolo (Bernstein et al., 2020) in combination with barcoded antibodies. Surface antibody reads were normalized by centered log-ratio transformation and scaled for each cell. Cell type annotation was based on Leiden clustering combined with sub-clustering of heterogeneous clusters and final annotation with known marker genes. Clonotype analysis was performed using scirpy (Sturm et al., 2020). Cells belonging to one clonotype were defined to have identical alpha and beta chain CDR3 nucleotide sequences. Both pairs of TRA/TRB sequences were considered in case additional chains were present. RNA velocities were calculated using velocityto (La Manno et al., 2018) and analyzed with scVelo (Bergen et al., 2020).

3.2.6 16S rRNA sequencing

16S sequencing was performed by Dr. Andreas Hiergeist from the Institute für Mikrobiologie und Hygiene in Regensburg according to core facility standards. Briefly, single-end sequencing was performed on the amplified V1-V3 region of the 16S gene. Normalization and absolute quantification were achieved by the spike in of specific amounts of *alicyclobacillus acidiphilus*, *agrobacterium radiobacter*, and *alcanivorax borkumensis* (Stammler et al., 2016). Amplicon sequence variants (ASV) were generated from demultiplexed and pre-processed sequencing reads using a dada2-based workflow (Callahan et al., 2016), and taxonomy was predicted using a SILVA classifier (Quast et al., 2013). For more details on the methodology of 16S rRNA sequencing and analysis, please refer to (Orberg et al., 2022)

For annotation of SCFA-producing species, the species producing Butyrate, Propionate, and Acetate were downloaded from the Virtual Metabolic Human database (Magnusdottir et al., 2017, Noronha et al., 2019), and bacterial species present in one or more of the lists were annotated with the respective metabolite.

4. Results

4.1 Patient cohort for the investigation of GvHD

To investigate the mechanistic connections between the gut microbiome, gastrointestinal immune cell infiltrates, and severity of acute GvHD in aHSCT patients, we analyzed a unique human patient cohort collected by the group of Prof. Dr. Ernst Holler at the University Hospital of Regensburg (Figure 11).

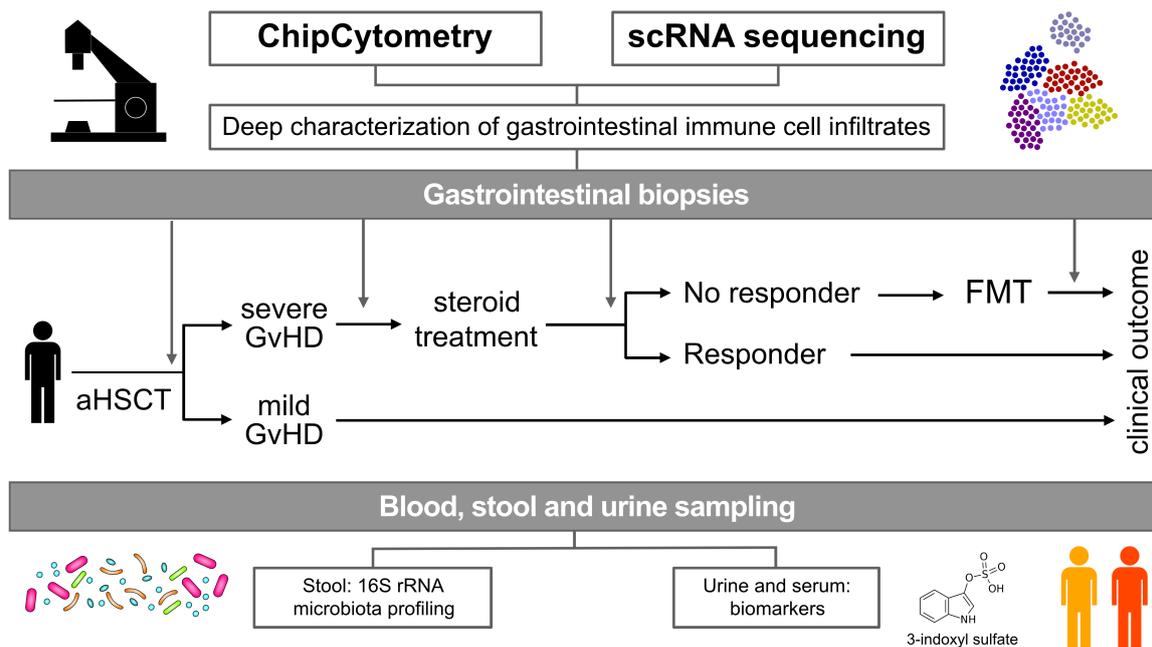


Figure 11. Human aHSCT patient cohort.

Graphical representation of the patient cohort. Stool and serum/urine samples are taken for the analysis of the microbiome and GvHD biomarkers respectively (bottom). Patients can take different paths of disease progression (middle) and gastrointestinal biopsies are taken according to the patient status (top)

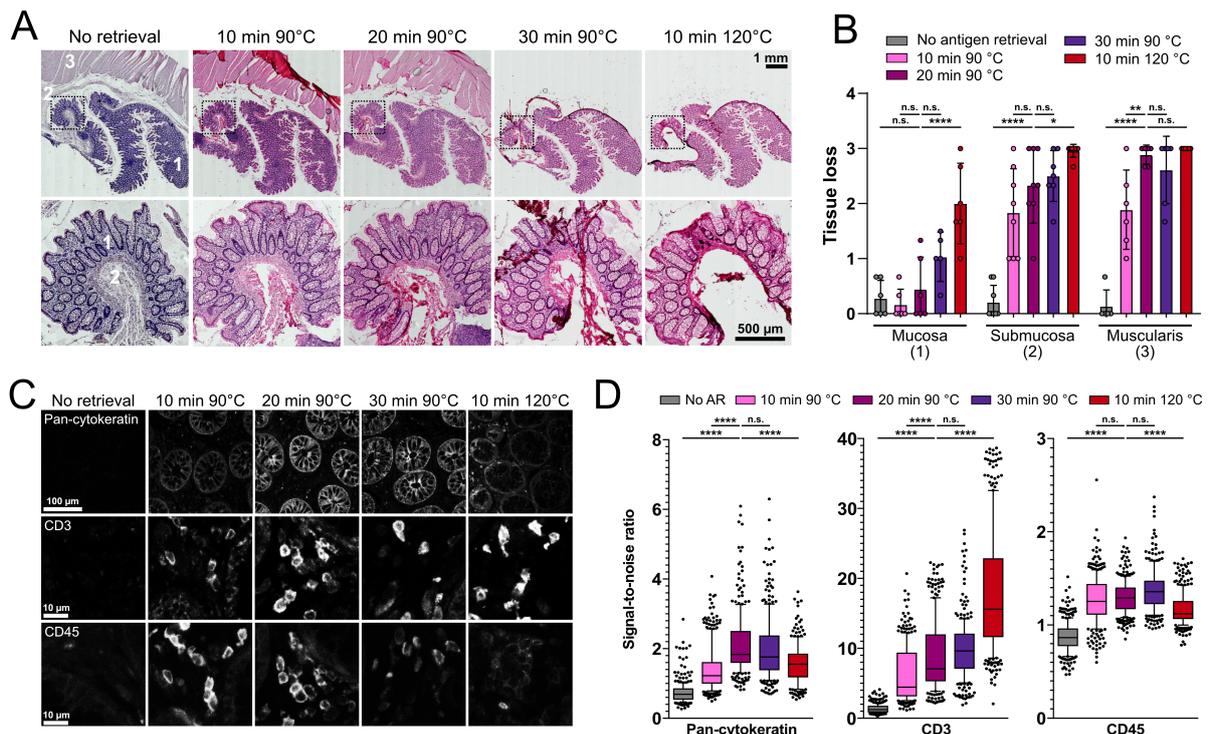
In this study, the gut microbiome is profiled by 16S rRNA sequencing of stool samples and GvHD severity is assessed by urinary or serum biomarkers. In addition, gastrointestinal biopsies are harvested at different time points during patient treatment, allowing in-depth analysis of infiltrating immune cells. To extract as much information as possible from these precious patient biosamples, we combined two emerging techniques for multiplex single-cell analysis: scRNA sequencing as the current state of the art for deep cell profiling and multiparametric imaging for spatially resolved phenotyping of cells within the tissue environment.

4.2 Method development for ChipCytometry on FFPE samples

At the beginning of this project, more than 300 gastrointestinal biopsies had already been collected as FFPE tissue samples, which is the most commonly used storage method for clinical samples (Gaffney et al., 2018). However, the ChipCytometry technology was available at the beginning of the project only for freshly frozen material, and methods for automated quantification of signal intensities were poorly developed. Therefore, as first steps, we had to adopt the technology for FFPE samples. This protocol development (Jarosch et al., 2021, Jarosch et al., 2022) will be further described within the next sections.

4.2.1 Implementation of ChipCytometry for FFPE sample handling

Before immunofluorescence staining can be applied on FFPE samples, epitopes need to be retrieved in order to make them accessible for staining antibodies. A commonly used retrieving strategy is heat-induced antigen retrieval (HIER), where the tissue is heated using a pressure cooker at 120 °C in a special buffer for antigen unmasking.



For ChipCytometry, tissue sections are mounted directly on coverslips instead of positively charged glass slides with polysilane-treated surfaces for improved tissue adhesion. This results in weaker tissue attachment and architecture preservation. Therefore, to overcome this limitation, we titrated the intensity of the antigen retrieval (Figure 12A). The lower temperatures achieved by tissue incubation in a water bath significantly reduced tissue loss (Figure 12B) while still preserving epitope accessibility, as indicated by the good SNR (Figure 12C-D). Eventually, we selected the 20-minute incubation at 90 °C in a water bath as optimal compromise between tissue preservation and staining intensity and as standard condition for our ChipCytometry experiments with FFPE sample material.

Although an increase in tissue integrity has been observed using a milder antigen retrieval, especially submucosa and muscularis tissue parts were still significantly affected by the retrieval treatment on coverslips. In collaboration with Canopy Biosciences, we were able to mimic the polysilane treatment commonly used on conventional microscopy slides on the coverslips for ChipCytometry. The use of these pre-treated coverslips could further reduce tissue detachment. Remarkably, we reached a comparable level of tissue integrity as observed with conventional microscopy slides (Figure 13A-B), potentially facilitating the use of the pressure cooker for hard-to-retrieve antigens again in the future. Quantification of differences between tissue supplies at our titrated retrieval condition of 90°C in the water bath showed significant reduction of tissue loss in submucosa and muscularis tissue parts on pre-treated compared to conventional coverslips and no significant differences to microscopy slides (Figure 13C).

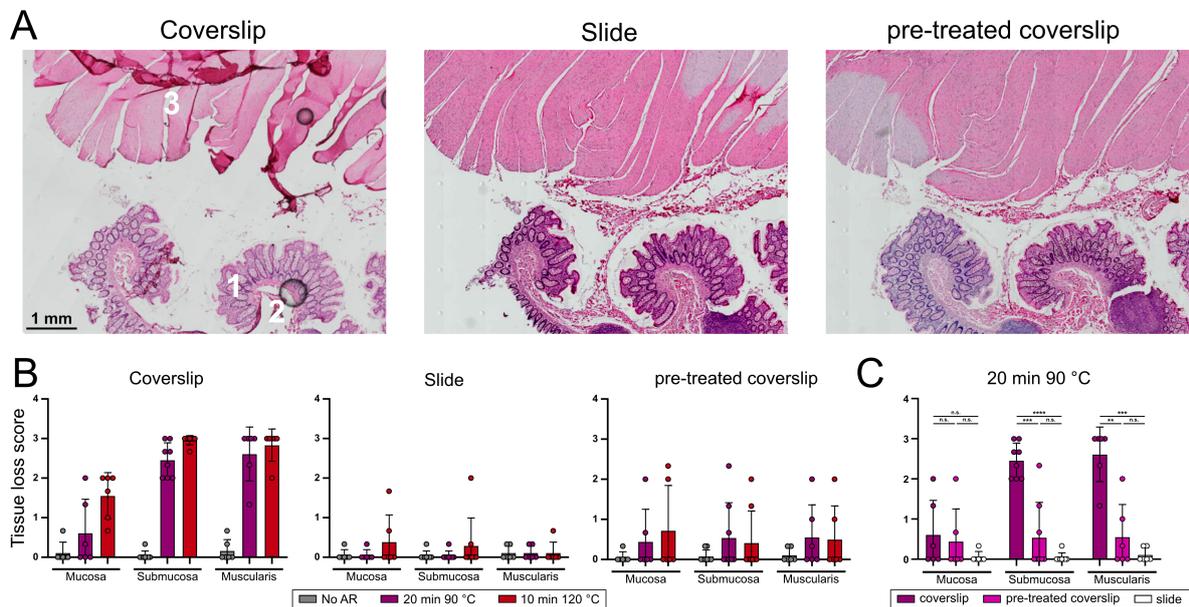


Figure 13. Enhancement of tissue integrity by the use pre-coated coverslips.

(A) Hematoxylin-eosin staining of consecutive tissue sections after antigen retrieval on different glass supplies. Numbers indicate regions of mucosa (1), submucosa (2), and muscularis (3). (B) Scoring of tissue loss for different retrieval conditions and supports (0 = no tissue integrity loss, 3 = complete loss of the tissue). (C) Scoring of tissue loss comparing different supports at the titrated retrieval condition of 20 min incubation at 90°C. Data in (B) and (C) are shown as mean ± standard deviation. Statistical testing in (C) was conducted by Tukey's test followed by Dunn's multiple comparison (**p<0.01; ***p<0.001; ****p<0.0001). A minimum of three donors in two independent experiments are shown. Modified from (Jarosch et al., 2021).

FFPE samples have been described to have an intrinsically high autofluorescence and we observed this tendency in the experiments described so far. ChipCytometry offers the unique opportunity to subtract the actual autofluorescence background from the staining images for each channel and each cell. However, high autofluorescence could dramatically reduce SNR eventually. We therefore wanted to reduce the autofluorescence of tissue samples to a minimum in order to guarantee a high staining intensity and image quality. For that purpose, we tested the three most commonly described autofluorescence quenching methods (Davis et al., 2014, Yang et al., 2017), which include treatments with photobleaching, Sudan Black B and Sodium Borohydride.

On slides pre-treated with antigen retrieval in TRIS-EDTA buffer at 90 °C in a water bath, we identified Sudan Black B to reduce autofluorescence most efficiently, leading to a higher SNR compared to the other tested treatments (Figure 14A-B). Chemical quenching seems to be more effective than light-induced quenching, since photobleaching performed the worst in this comparison. In a titration of Sudan Black B incubation times, we observed the highest gain in SNR at 10 min, with a decreasing staining intensity at longer incubation times (Figure 14C-D). The last parameter we optimized for tissue pre-treatment was the thickness of the sections, which can influence both the staining intensity and the autofluorescence. We cut consecutive slides of different tissue thickness, treated them for 10 min with Sudan Black B and analyzed the SNR for the different conditions. In these experiments, we identified 4-5 μm as the ideal thickness regarding SNR of FFPE samples (Figure 14E).

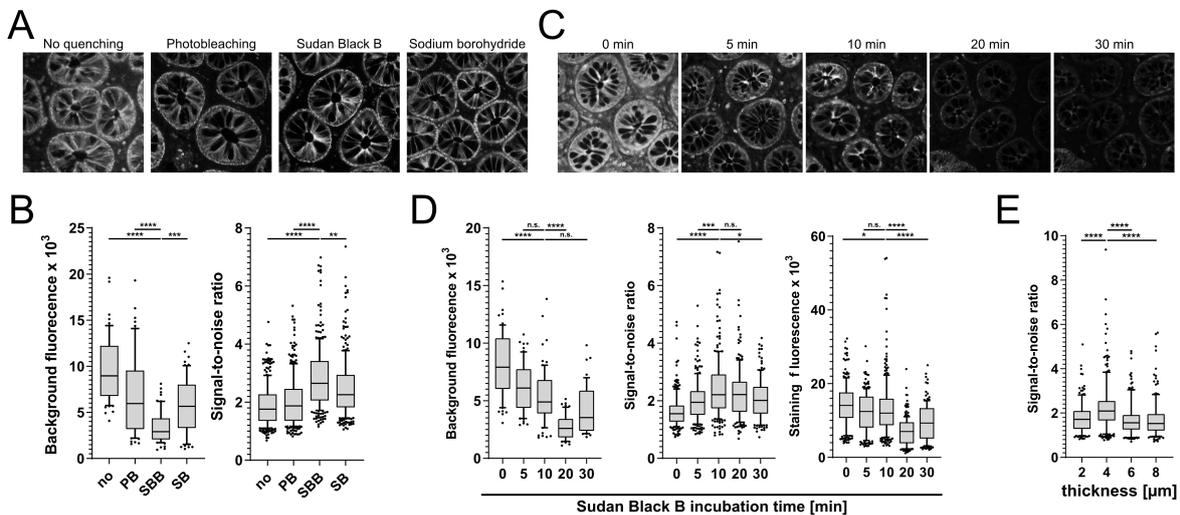


Figure 14. Titrated autofluorescence quenching and section thickness enhance staining intensity.

(A) Representative images for pan-cytokeratin staining in healthy colon sections pre-treated with different autofluorescence quenching methods. (B) Quantification of background area intensity and SNRs of pan-cytokeratin staining. (C) Representative images for pan-cytokeratin staining in healthy colon sections with titrated Sudan Black B incubation time. (D) Quantification of background intensity, SNRs and staining intensity is shown for conditions depicted in (C). (E) SNR values for different tissue thicknesses are shown regarding pan-cytokeratin staining. Data in figure (B), (D) and (E) are pooled from 3 donors and 2 independent experiments, depicted as interquartile ranges (whiskers extending to 10% and 90%, outliers plotted as dots). 8 positions per section were analyzed. Non-epithelial tissue parts were quantified as autofluorescence background. Statistical testing was conducted by Tukey's test followed by Dunn's multiple comparison (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$). PB=Photobleaching, SBB=Sudan Black B, SB=Sodium borohydride. Modified from (Jarosch et al., 2021).

4.2.2 Marker establishment

To fully exploit the potential of ChipCytometry in highly multiplexed spatial analysis, we needed to search for antibodies capable of recognizing epitopes in FFPE samples and that are ideally directly conjugated with a photo-bleachable fluorophore. Since the majority of commercially available antibodies for immunofluorescence are usually not conjugated but rather detected with fluorescently labeled secondary antibodies, or conjugated to photostable dyes, we had to rely on staining information for unconjugated clones and test the performance of the less validated corresponding conjugated batches for each antibody separately. After testing more than 100 different antibodies, we finally established a list of 32 validated clones (Table 5), which can be used for the staining of markers related to the general tissue architecture as well as of antigens for further deep characterization of especially immune cells, able to differentiate more than 30 different cell types (Figure 15). For the latter goal, immune cells are first identified by the expression of CD45 and then subclassified into different lineages using markers as CD3 (T cells), CD20 (B cells), CD56 (NK cells), CD14 (monocytes), CD68 (macrophages). T cells can be further characterized into cytotoxic (CD8) and helper (CD4) T cells, with resting/activated regulatory T cells being the most complex phenotype in this scenario.

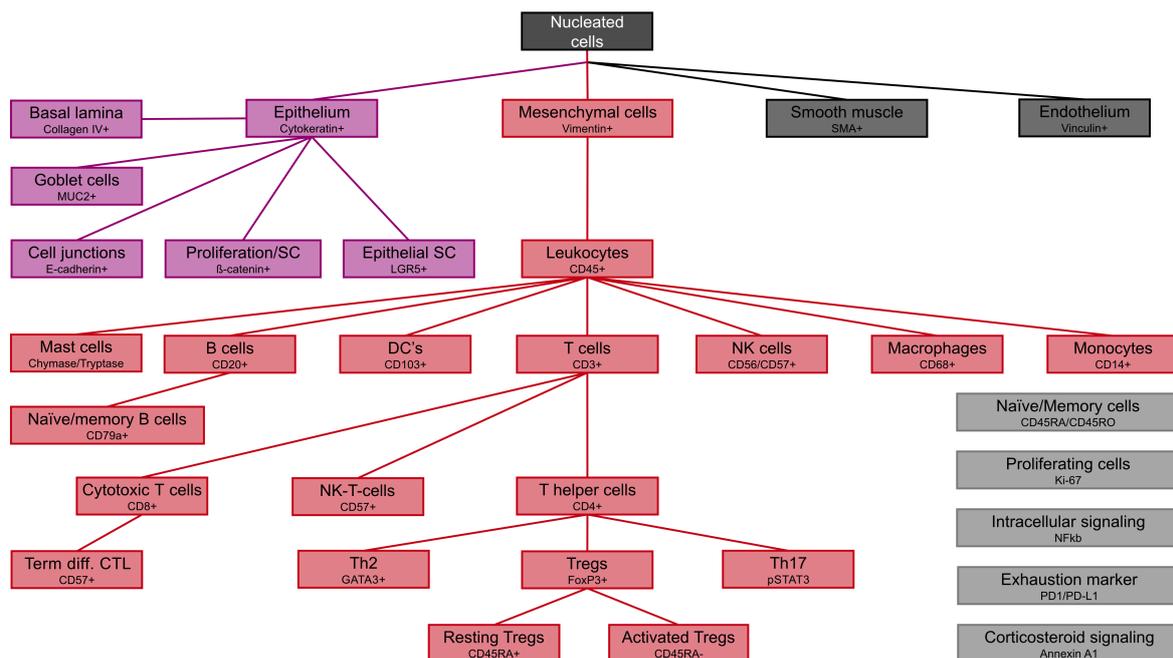


Figure 15. Cell types detectable with the established markers in human FFPE tissues.

All markers together can classify cells into the main compartments of a tissue (epithelium, mesenchymal cells, muscle and endothelium). Cells can be further subclassified into distinct functional subsets of epithelial cells (shown in magenta) and all different lineages of immune cells (shown in red). Finally, phenotypic subtyping can be achieved by several phenotypic markers (shown in light grey).

This set of markers allows a variety of analyses ranging from the qualitative detection of architectural changes to the deep characterization of cell populations. As an example, we have stained a colorectal cancer sample with the architectural markers pan-cytokeratin (epithelium), vimentin (lamina propria cells) and SMA (muscular cells). This three-marker combination already revealed obvious differences

in crypt architecture between tumor tissue and adjacent non-tumor tissue (Figure 16A). Deep phenotyping of immune cells was demonstrated in an inflamed colon tissue sample by visualizing the infiltration of T cells (CD3), B cells (CD20) and monocytes (CD14) as well as their relative position to the epithelium (Figure 16B). These images illustrate the potential of multiparameter imaging and showcase the resolution that allows to investigate each marker on each single cell. One example for such phenotypically complex event in this scenario is the activated regulatory T cell (nuclei+CD45+CD45RA-CD45RO+CD3+CD8-CD4+CD20-FoxP3+), which can be identified by the combined expression of several different markers (Figure 16C). Overlay of co-expressed markers as well as the absence of overlay between mutually exclusive markers served as quality control to show specificity of selected antibodies (Figure 16D).

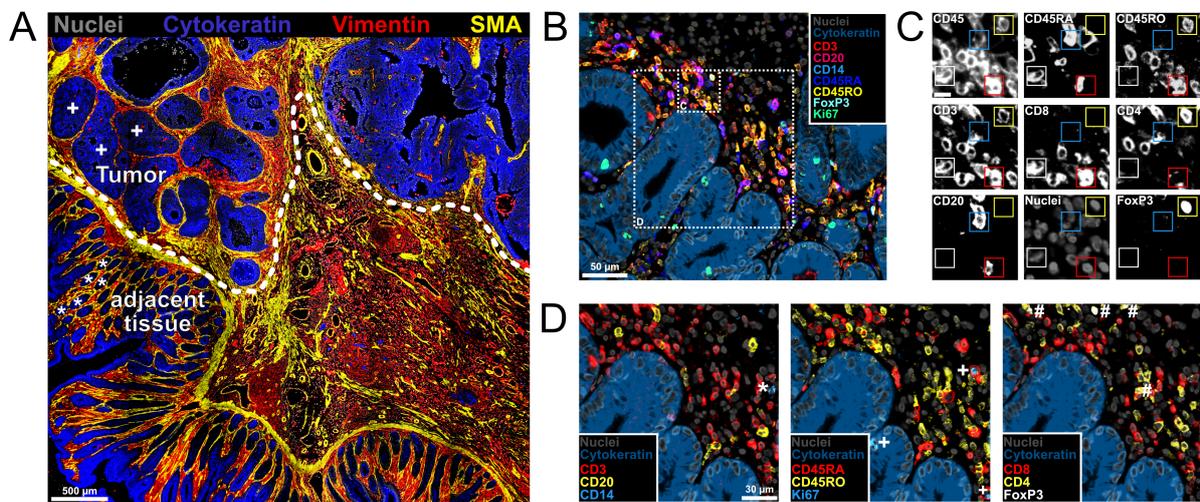


Figure 16. High-quality, single cell resolution imaging of FFPE colon tissues with ChipCytometry.

(A) ChipCytometry image of colon tumor and non-tumor adjacent tissue (above and below the dashed line, respectively). Asterisks (*) and crosses (+) indicate normal and degenerated crypts, respectively. (B) Representative images of multiplexed ChipCytometry in a high grade GvHD biopsy. Zoom areas shown in (C) and (D) are marked with dashed rectangles. (C) Single marker images are shown for a selected area from (B). High multiplexing enables to characterize for example an activated CD4+ Treg (yellow), a B cell (blue), an activated/memory (red) and a naïve (white) CD8+ T cell. Scale bar indicates 10 μm . (D) Multiplex ChipCytometry overlay of mutually exclusive markers (CD3/CD20/CD14, CD45RA/CD45RO, and CD4/CD8) and co-expressed markers (CD4/Foxp3). Exemplary cells are annotated from left to right as *:CD14+, +:Ki-67+, #:CD4+Foxp3+ cells. Modified from (Jarosch et al., 2021).

4.2.3 Fluorescence RNA *in-situ* hybridization (FISH)

Especially intranuclear markers and low-expressed proteins like cytokines are difficult to be detected via antibodies. In order to facilitate the detection of such targets, we aimed to combine the ChipCytometry technology with multiplexed fluorescence RNA *in-situ* hybridization (FISH) using the RNAscope technology (Wang et al., 2012). We established a protocol (Jarosch et al., 2022) that allows to stain up to three markers with multiplexed FISH, before continuing with the conventional ChipCytometry protocol (Figure 17A). Briefly, mRNA specific probes (Table 8) are hybridized after antigen retrieval but before chip loading and detected with an amplification system of complementary probes and dyes in a second step when the chip is loaded. We first checked the specificity of FISH by targeting CD3, a relatively abundant marker for which reliable antibody staining is available. Co-staining of CD3 on the mRNA and

protein level revealed a very good co-localization of staining signals as a sign for FISH specificity (Figure 17B).

As next step, we moved into the analyses of markers not detectable via antibody staining due to either low expression (CTLA4) or difficult preservation in tissues (cytokines like IFN- γ and IL10). FISH as methodological addition to ChipCytometry enabled us to perform even deeper characterization of cell types for example by measuring the expression of effector molecules like IFN γ for effector T cells or crucial effector molecules of regulatory T cells like IL10 and CTLA4 (Figure 17C). In this last case, co-localization of CTL4 and IL10 signals with the intranuclear antibody staining of Foxp3 served as quality validation of FISH. It is important to perform FISH staining before cyclic antibody staining due to the instability of the mRNA molecules. This allows to multiplex up to three mRNA targets in one cycle of FISH staining at the moment, which might be expanded to additional cycles in the future.

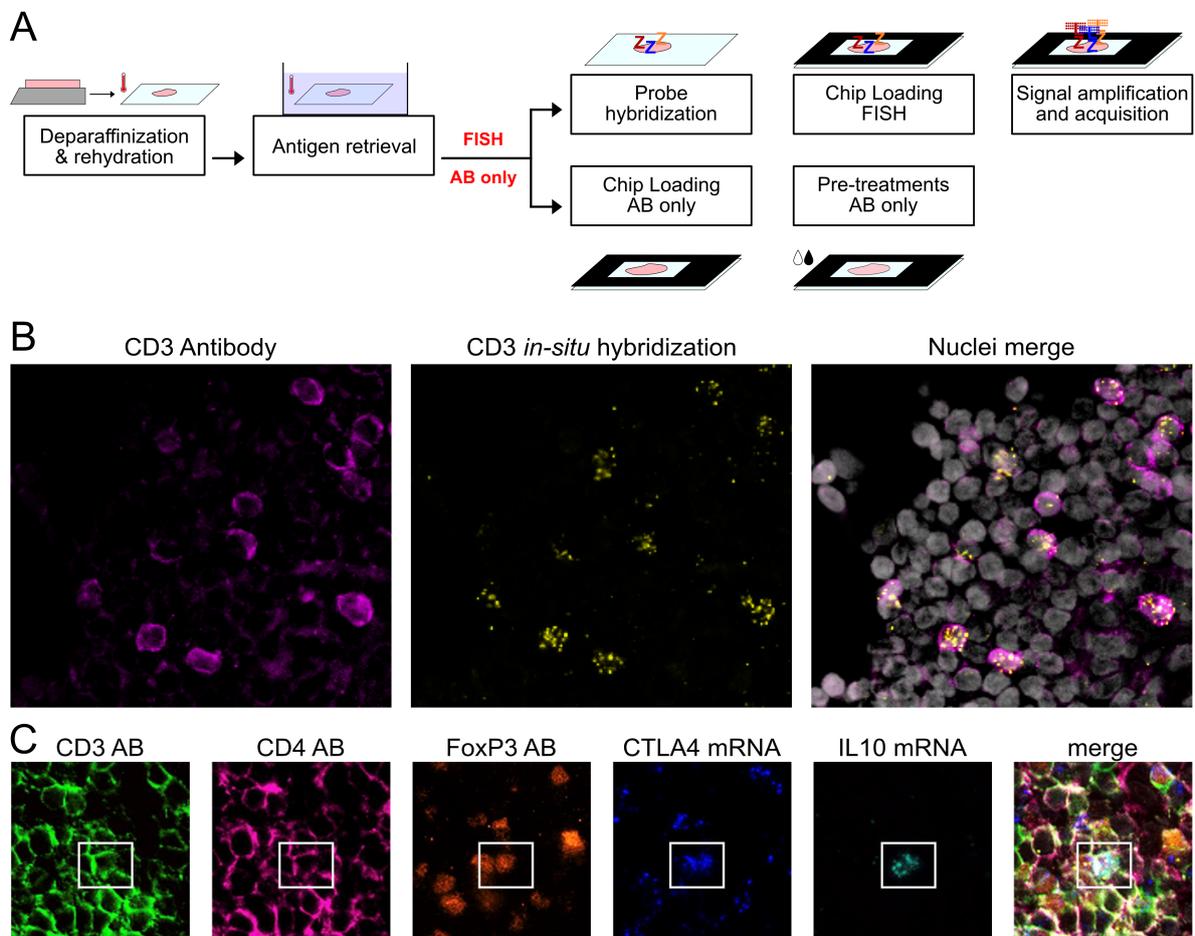


Figure 17. Implementation of FISH mRNA staining into the ChipCytometry workflow.

(A) Workflow of tissue pre-treatments before cyclic immunofluorescence is performed. Two options are depicted: either the conventional pre-treatment “AB only” without FISH (lower option), or the combination with FISH (upper option). (B) Exemplary staining of CD3 protein and mRNA on the same chip. (C) Multiplexed FISH combined with multiplexed antibody (AB) staining of a colon cancer sample. Modified from (Jarosch et al., 2022).

4.2.5 Automatic quantification of multiplexed image data

Multiparameter imaging generates a huge amount of data with a large number of markers for every cell down to subcellular resolution. Manual counting of cells is still a widely used procedure for cell quantification in routine diagnostics but becomes unrealistic to be applied to multiparameter imaging. For automated image analysis, the aim is to consider every cell within a tissue instead of analyzing representative views, and to capture complex phenotypes of individual cells consisting of co-expression of several markers. Indeed, the only option to achieve meaningful quantifications of multiplexed tissue staining is an automated approach for the quantification of all fluorescence intensities related to each segmented cell.

By adopting some preexisting tools (Lin et al., 2015, Lin et al., 2018) and implementing some new aspects of image correction, we developed an automated analysis pipeline for highly multiplexed imaging data. The pipeline uses ChipCytometry output files or stitched images from any multiplexed imaging method as input, and performs cell segmentation, image pre-processing and value calculation for each cell and each marker (Figure 18A). Calculated values are then converted into the FCS format, allowing for quantification similar to flow cytometry experiments by gating populations in two dimensional plots. The pipeline has been written as ImageJ plugin and can be downloaded together with its documentation from a continuously updated [GitHub repository](#). Running the pipeline on an imaging dataset does not require any programming knowledge and the user will be guided through the steps of parameter specification by simple click-and-select actions (Figure 18B). Methodological details of the individual pipeline steps will be described in the following section.

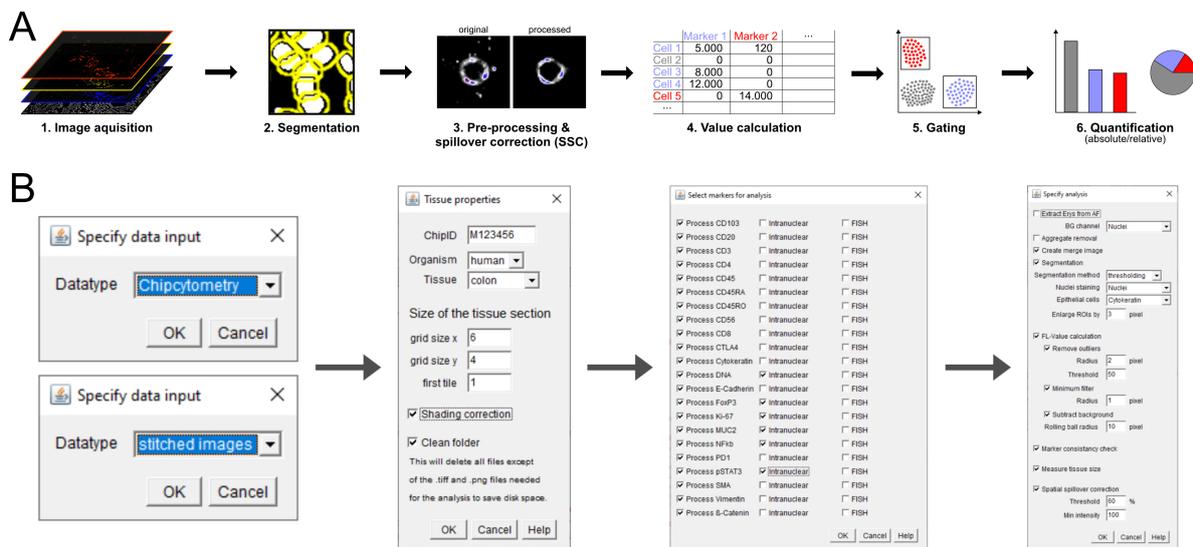


Figure 18. Automated analysis pipeline for highly multiplexed imaging data.

(A) Schematic overview of the individual steps of image quantification. (B) Screenshots from the actual ImageJ pipeline guiding the user through the steps of image analysis. Modified from (Jarosch et al., 2021, Jarosch et al., 2022).

4.2.5.1 Stitching

Before ChipCytometry images are processed for signal quantification, images of independently acquired positions must be stitched in order to generate one single image per marker containing the whole tissue area. Due to the photobleaching steps and the illumination characteristics of the instrument, areas of the tissue can be illuminated twice at the overlap of single positions, resulting in a grid of reduced signal intensity after image stitching (Figure 19, left image). This effect can be corrected using the BaSiC method for shading correction (Peng et al., 2017), which extrapolates a repetitive shading profile from a set of microscopy tiles. In order to apply this method on our images, we used the non-background subtracted fluorescent images to extract the shading profile for each staining and applied this profile on the individual tiles before stitching, resulting in a shading-corrected image (Figure 19, right image). It is important to perform this analysis not only on every channel, but on every staining, since the bleaching effect due to double illumination is exaggerated with the number of cycles, leading to stronger artefact in later acquisitions.

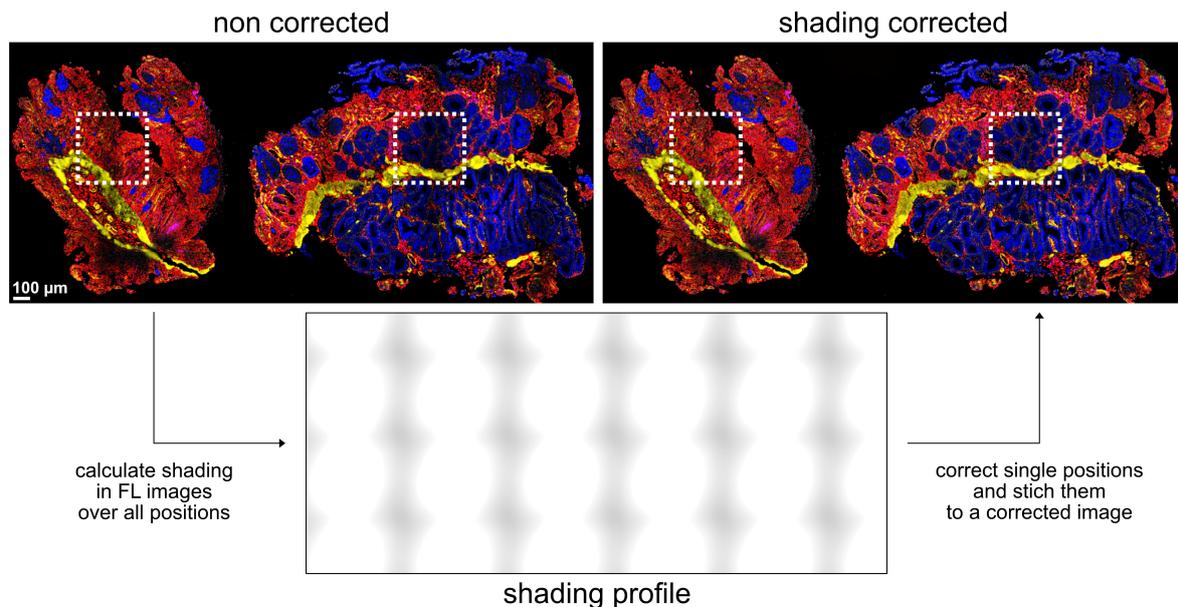


Figure 19. Shading correction reduces bleaching-induced artefacts at adjacent positions.

Differences between an uncorrected, stitched image (left) and a stitched image from tiles that have been corrected for repetitive shading beforehand (right). An exemplary shading profile corresponding to one of markers (pan-cytokeratin) in the composite image above is depicted in the lower part of the figure. Modified from (Jarosch et al., 2022).

4.2.5.2 Segmentation

Cell segmentation is a crucial part for automated analysis of imaging data. Several approaches have been exploited already, mainly differentiating into membrane-based and nuclei-based segmentations. Membrane-based segmentation is an advanced approach with the need of a good pan-membrane marker for all different cell types. It usually requires a seed, which is expanded towards the membrane staining and is essential for cells with a special shape like neurons or fibroblasts as well as for non-nucleated cells. For the aim of our project, we went with a nuclei-based approach, which has the

advantage of easy and robust stainability and showed a good performance in segmenting the cell types of our interest, having a convex shape with a low fraction of cytoplasm.

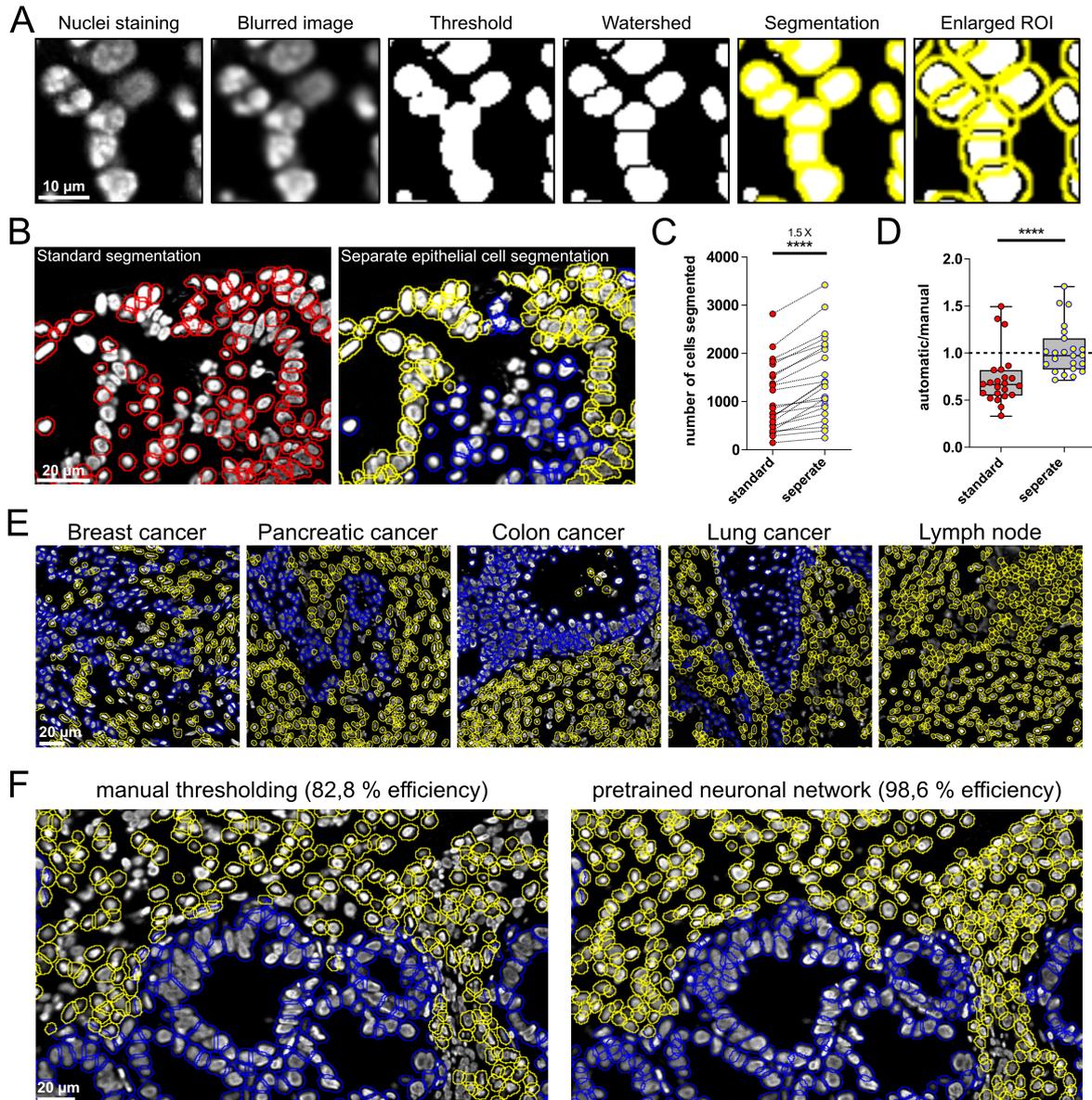


Figure 20. Nuclei staining-based segmentation of single cells from tissue samples.

(A) Different steps for cell segmentation are depicted. After blurring of the nuclei image, a threshold is applied, and shapes are refined by the watershed algorithm. Finally, particles can be detected, and boundaries enlarged for complete surface marker signal capture. Qualitative representation (B) and quantification (C) of segmented cells for either standard or separate epithelial/non-epithelial cell segmentation on a selected area of colon tissue samples from three different donors are depicted. Each dot represents one position (n per donor ≥ 6). (D) Ratio between the number of cells segmented either via standard or separate cell-type segmentation as shown in (C) and the number of cells quantified by manual counting. Data are depicted as interquartile range with whiskers extending from the minimum to the maximum of the dataset. (E) Representative segmentation images are depicted for different tissue types, each processed with separate segmentation of epithelial and non-epithelial cells. (F) Qualitative comparison between manual thresholding and neuronal network segmentation using StarDist is shown. Statistical testing in (C) and (D) was conducted by paired t test (**** $p < 0.0001$). Modified from (Jarosch et al., 2021, Jarosch et al., 2022).

We started out with a manual thresholding approach for nuclei detection, which allows the user to correct for differences in staining intensities between samples. Briefly, the nuclei image is blurred in order to reduce information and avoid over-segmentation. The blurred image is then binarized by applying a

manual threshold and the binary image is refined by watershed. Finally, particles are segmented, and regions of interest enlarged in order to capture the complete surface signal of a cell (Figure 20A). Since the particle detection is refined by some specific parameters regarding size and shape of the cells, we found that a separate segmentation of differently shaped cell types results in a better overall segmentation efficiency. For our case, we separated the elongated epithelial cells from the more roundly-shaped non-epithelial cells and obtained a higher number of segmentation events as well as a better accordance with cell numbers defined by manual counting, which served as a reference (Figure 20B-D). This approach is robustly working on different types of tissue; in lymphoid samples the expansion of ROIs is reduced in order to avoid spillover in these densely packed tissues (Figure 20E).

One major disadvantage of manual thresholding is the potential bias introduced by threshold selection. We therefore implemented the option of using a neuronal network model pre-trained on nuclei images from StarDist (Weigert, 2020). This algorithm is trained to predict distances to an object's boundaries in a star-shape and applies non-maximum suppression (removal of overlapping objects in machine learning) on candidate polygons to obtain the final segmentation. By still separating different cell types according to their shape, we can get an even improved and very robust segmentation of cells from more dense nuclei images, avoiding the risk of bias by manual thresholding, but more prone to over segmentation (Figure 20F).

4.2.5.3 Pre-processing and spatial spillover correction

As observed in all fluorescence imaging approaches, also in the ChipCytometry stainings, there is a chance of unspecific antibody binding, leading to a scattered background signal (Figure 21A).

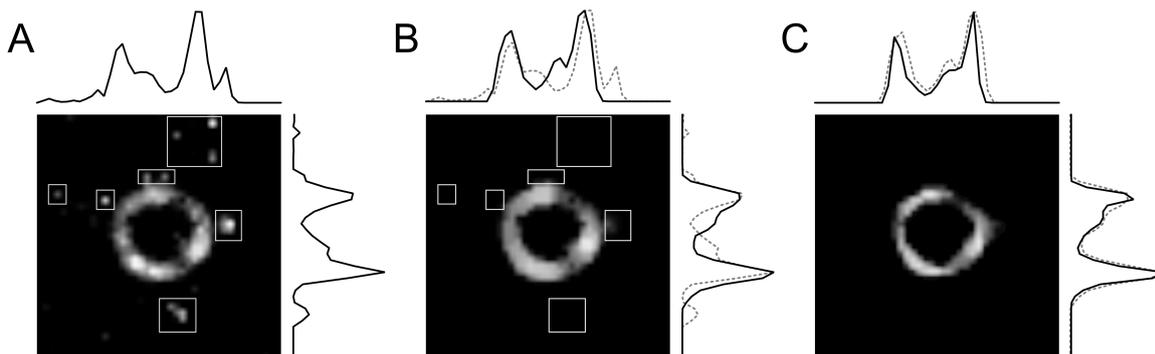


Figure 21. Pre-processing of ChipCytometry images for automated quantification.

(A) Fluorescence images often contain some scattered unspecific signal. A single cell image including noise is depicted together with the intensity profile of the intensity sums in x and y direction. (B) Same depiction as in (A), but processed with an outlier removal filter (replaces a pixel by the median of the pixels in the surrounding if it deviates from the median by more than a certain value). (C) Image from (B) was additionally processed with a minimum filter (grayscale erosion by replacing each pixel in the image with the smallest pixel value in that pixel's neighborhood).

In order to ensure a good and unbiased signal quantification for each segmented cells, we titrated and implemented some mild pre-processing steps in our quantification pipeline. The first filter that is applied, is the outlier removal. This filter replaces a pixel by the median of the pixels in the surrounding if it

deviates from the median by more than a certain value, leading to the removal of unspecific signal (Figure 21B). In a second step, the signal is sharpened by a minimum filter replacing each pixel in the image with the smallest pixel value in that pixel's neighborhood (Figure 21C), thereby minimizing the potential of signal spillover to neighboring cells.

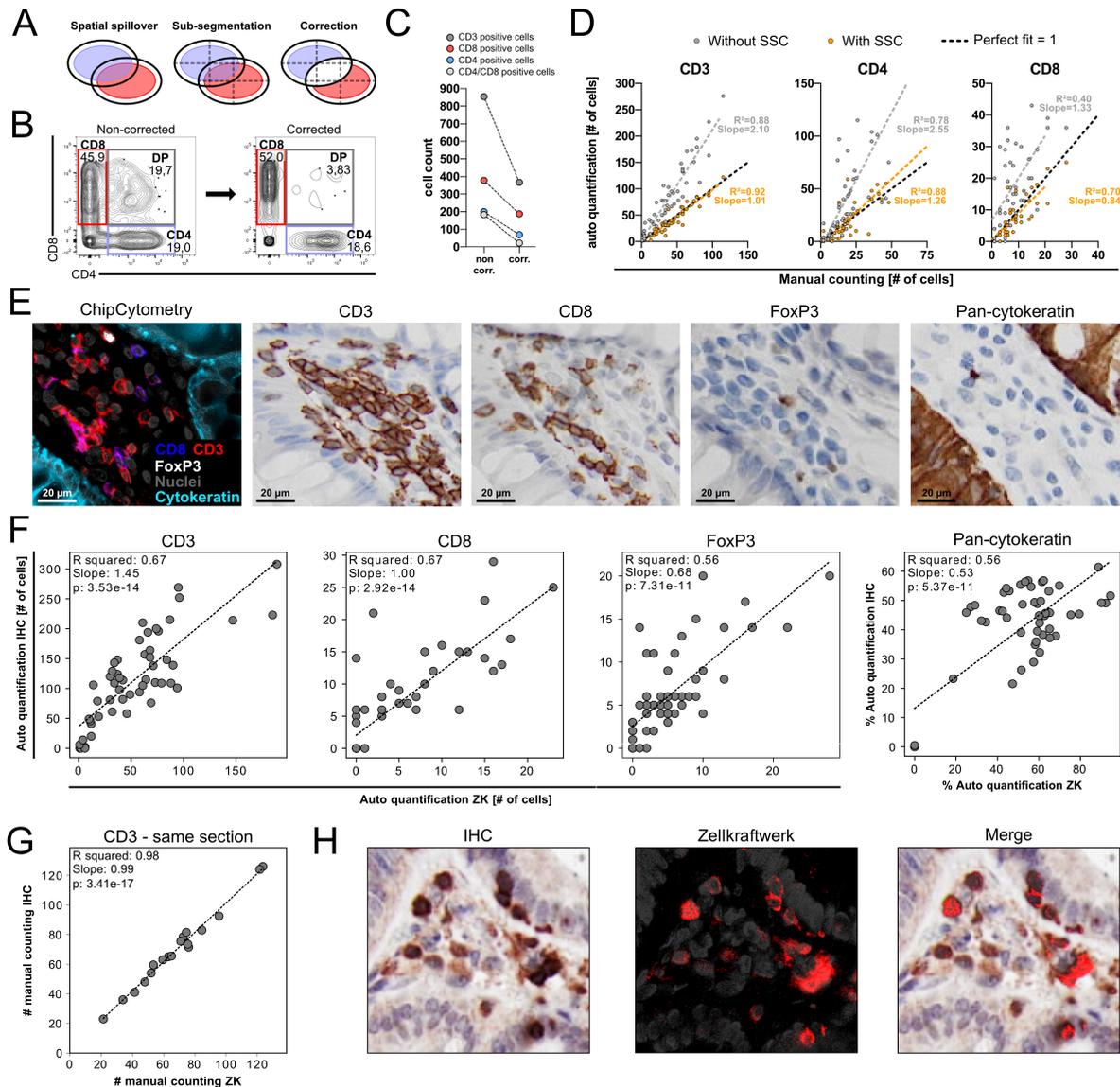


Figure 22. Spatial spillover correction improves the reliability of automated image quantification.

(A) Schematic depiction of SSC. Cells are sub-segmented into quadrants and the signal of a marker is deleted if the quadrant signal as fraction of total signal is exceeding a threshold value. (B) Representative data showing CD4 and CD8 staining on CD3 positive cells either with or without SSC on an inflamed colon tissue. (C) Quantification of cells positive or negative for CD8 and CD4 staining. Signal intensities were automatically quantified with or without SSC and cells were gated according to signal intensity. Cell count refers to the number of cells present in each corresponding gate. (D) Positive cells for CD3, CD4, and CD8 markers were quantified by either manual counting or automatic quantification, with or without SSC. Each dot represents one out of 10 representative positions for three donors. For automated quantification, number of positive cells was obtained by gating strategy using FlowJo software. Correlation analysis was done by Pearson's correlation. (E and F) Human healthy colon sections were stained with antibodies against pan-cytokeratin, CD3, CD8, and Foxp3. (E) Representative images of multiparameter ChipCytometry and individual marker IHC on consecutive slides. (F) Pearson's correlation between automated quantification of ChipCytometry staining and automated quantification of IHC staining. Each dot represents one out of 15 positions analyzed per donor ($n = 3$). Correlation analysis on CD3 staining with IHC and ChipCytometry on the same section (G) with corresponding images (H) is depicted to prove that consecutive slides lead to intrinsic variation in cell numbers. Modified from (Jarosch et al., 2021).

Whenever it comes to the signal quantification from cells within a tissue, spatial spillover is an important factor to consider. Spatial spillover describes the fact that the signal from a cell might also be detected within a neighboring cell due to their spatial organization. In our newly developed spatial spillover correction (SSC), each cell is sub-segmented into quadrants and surface marker signals are removed if the quadrant signal as fraction of total signal is exceeding a titrated threshold of 60% (Figure 22A). When we applied SSC to CD4 and CD8 staining data, known to be predominantly mutually exclusive, we found a strong reduction in the frequency of double positive cells (Figure 22B). Additionally, we also observed a substantial reduction of absolute cell numbers for single positive cells (Figure 22C), indicating the high extent at which spatial spillover could bias tissue imaging analyses.

To evaluate the accuracy of SSC, we next correlated cell numbers from corrected and non-corrected quantifications with manual counted cell numbers. Despite a good correlation in both scenarios, we found an overestimation of cell numbers without SSC, as indicated by correlation slopes above the value of 1.0 (Figure 22D). For corrected quantification, the correlation's slope values around 1.0 indicated that SSC could reduce most of the false-positive signals without influencing the true-positive events.

As a final validation, we compared the automatically generated data from ChipCytometry experiments with immunohistochemistry (IHC), still the gold standard in histological diagnostics. Consecutive sections were stained with single marker via IHC (CD3, CD8, FoxP3 and pan-cytokeratin); in parallel a single section was stained simultaneously for all markers using ChipCytometry (Figure 22E). Notably, we found a robust correlation between automated quantification for both methods (Figure 22F) with some variation being explained by the use consecutive slides. This hypothesis was validated by CD3 staining via ChipCytometry and IHC on the same section, leading to an increased correlation between those methods (Figure 22G-H).

4.2.5.4 Cell quantification from ChipCytometry imaging data

Once the images have been pre-processed and cells been segmented, the mean fluorescence intensity per cell for each marker can be used to identify and quantify cell populations in the dataset via gating (meaning thresholding marker expression to identify a cell population as either positive or negative for the specific marker). To facilitate this approach, the value table was converted into the FCS format and handled identically to flow cytometric data. This method allows to hierarchically navigate into the cell populations, starting from very basic discrimination of epithelial cells and leukocytes towards phenotyping of lineage populations like the assessment of PD1 expression within the T cell compartment (Figure 23A).

Besides qualitative analyses, the approach can be used to quantify the abundance of individual populations, which was done as a proof-of principle on tissue sections from colorectal and pancreatic adenocarcinoma containing also non-diseased adjacent tissue. Different populations were gated according to eleven markers (Figure 23A) and subsequently enumerated. As expected, we observed higher

abundance of epithelial cells in cancer tissues, most likely representing the malignant compartment, as well as higher infiltration of CD45+ immune cells both in the lamina propria and in the epithelial compartment of tumor regions, but to a lesser extent in pancreatic cancers (Figure 23B-C).

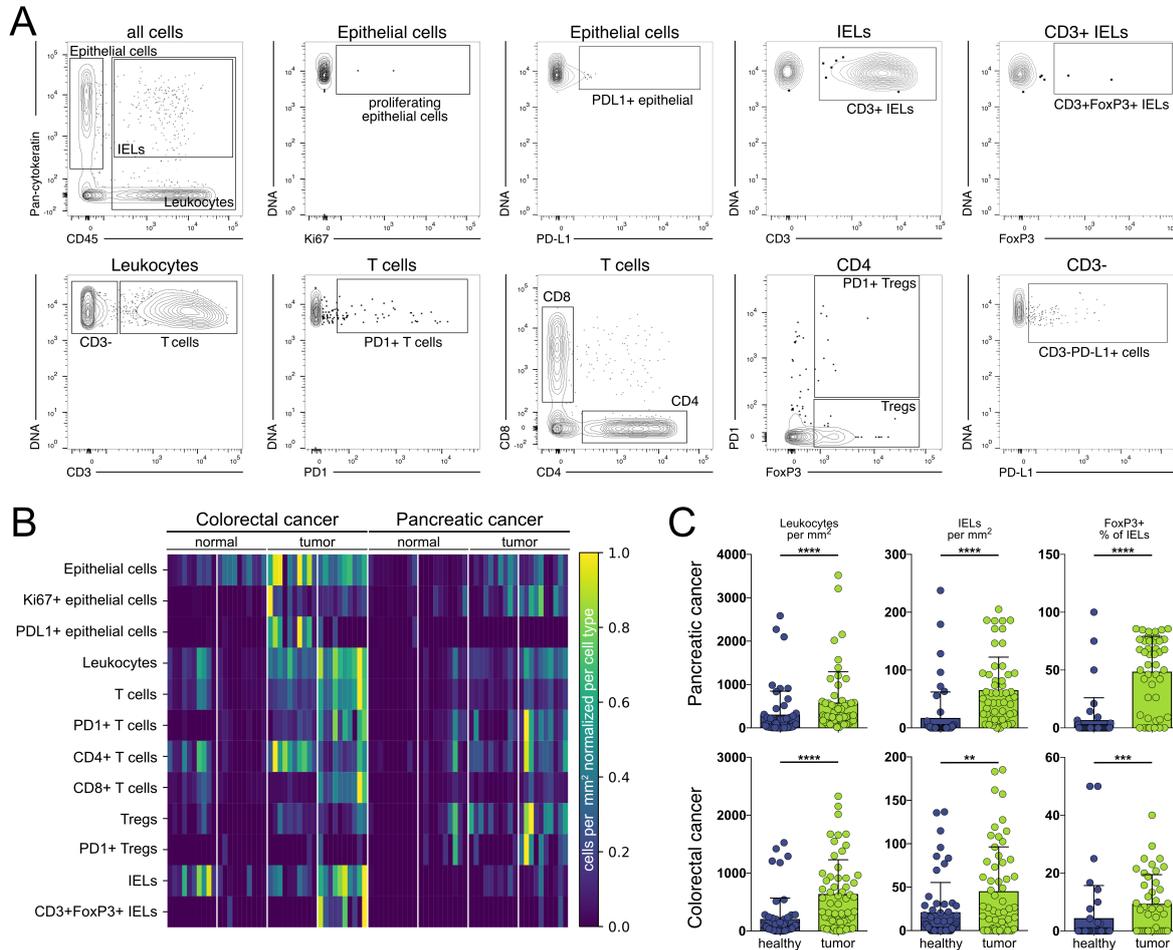


Figure 23. Multiparametric imaging and signal quantification of pancreatic and colon cancer tissues.

(A) Gating strategy applied for quantification of each sub-position of cancer and adjacent non-tumor tissues from colorectal and pancreatic tissue samples (B) Comparison of cancer and adjacent non-tumor tissues (n=2). Each column represents one of 10 representative positions of each tissue type and each row refers to a gated cell population. The color depicts the number of cells per mm², normalized per row (cell population). (C) Quantification of immune infiltrates in cancer and adjacent non-tumor tissues (n=4). Each dot indicates one representative position. Depicted mean +SD. Statistical testing was conducted by Mann-Whitney-test (**p < 0.01, ***p < 0.001, ****p < 0.0001).

These observations are in good accordance with earlier findings that describe a strong prognostic impact for tumor-infiltrating leukocytes for colorectal cancer (Angell et al., 2020, Sperlich et al., 2018) and a lower degree of immune infiltration in pancreatic cancers, which could explain unresponsiveness to checkpoint inhibitors (Skelton et al., 2017). The high proliferation state, measured by expression of Ki67, and pronounced PD-L1 expression marked the epithelial tumor cells (Figure 23B). T cells were also more abundant in tumor areas, including potentially exhausted PD-1+ T cells (Yaghoubi et al., 2019). Interestingly, lower infiltration of cytotoxic CD8+ T cells but higher presence of Tregs was observed in pancreatic tumors compared to colon tumors. The ability to detect populations such as PD1+Tregs and Foxp3+ IELs (Figure 23B-C) is clinically important, since these populations can be crucial for cancer

therapy outcome (Kumagai et al., 2020). Anti-PD-1 therapy was shown in some cases to amplify the suppressive activity of PD-1+ Tregs, thus leading to cancer progression (Kamada et al., 2019). Alternatively, the ratio of intraepithelial effector cell to Tregs has been suggested as a predictor for clinical outcome (Sinicrope et al., 2009).

4.2.5.5 Quantification of phenotypically complex cell populations

Having set all parameters for signal quantification, we started developing gating strategies to define rare and complex cell populations within a tissue by the expression of several markers. Given the importance of regulatory T cells in GvHD pathophysiology, we firstly focused on the quantification as well as sub-characterization of these cells in the gut biopsies of our patient cohort.

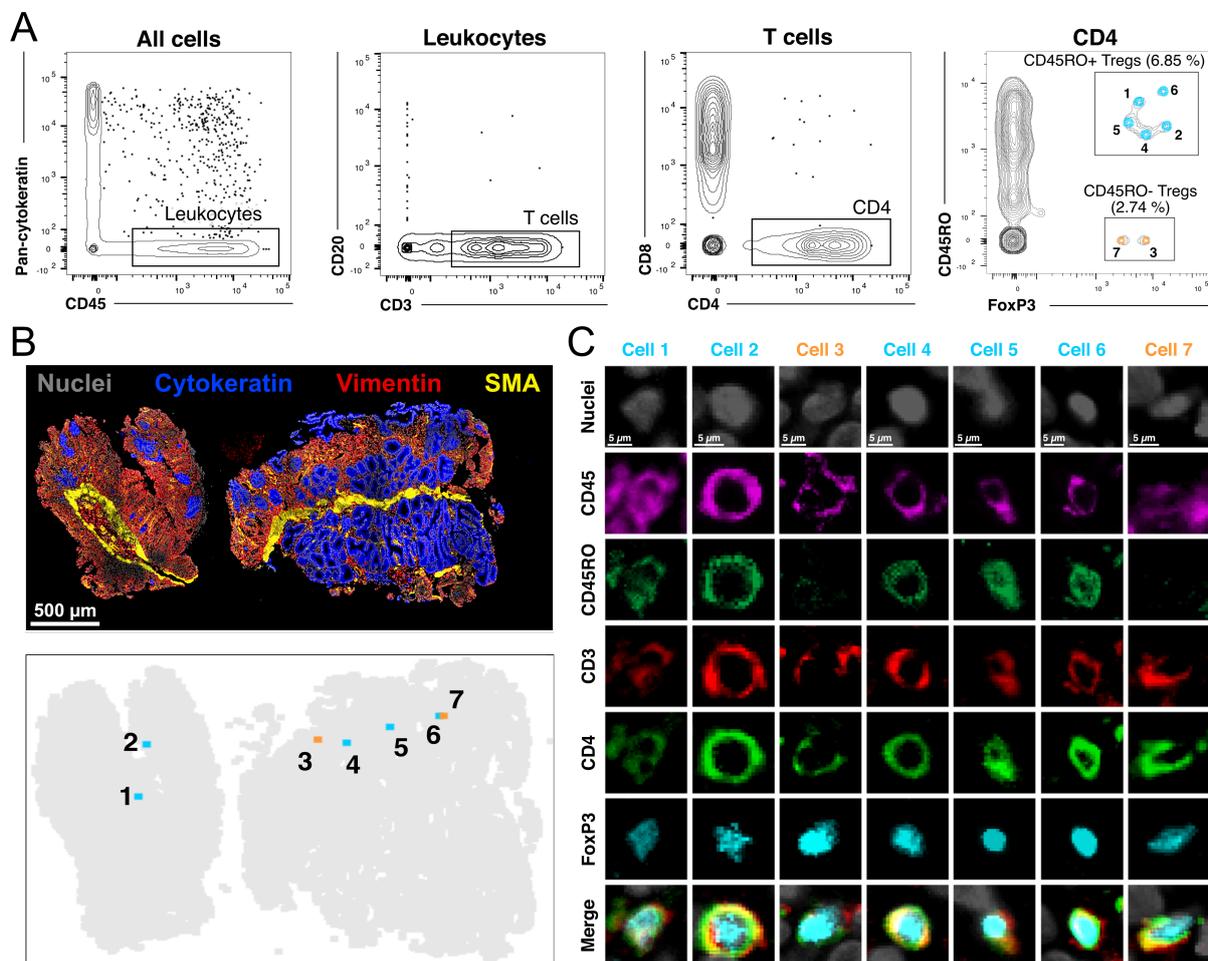


Figure 24. Robust detection of rare and phenotypically complex cell populations by ChipCytometry.

Tissue section from GvHD gut biopsy was stained with 18 markers (CD45, CD45RA, CD45RO, CD3, CD4, CD8, CD14, CD68, CD20, CD25, Foxp3, Gata3, pan-cytokeratin, Ki-67, PD-1, PD-L1, vimentin, SMA). (A) Representative gating strategy used to define tissue-infiltrating immune cells. Tregs were identified by sequential gating according to CD45+CD3+CD4+Foxp3+ expression and finally discriminated according to CD45RO expression. (B) Replotting of the gated Treg populations from (A) into the original stained tissue, to assess the spatial distribution/location. (C) Original staining images of the seven Tregs identified in (A and B). Modified from (Jarosch et al., 2021).

Regulatory T cells are often extremely rare events with complex phenotype. Here, we identified Tregs by the expression of CD45, CD3, CD4 and FoxP3, ending up with a population representing only 0.04% of the total quantified cells. This already rare population was further characterized by the expression of CD45RO as marker for activated regulatory T cells, leading to five activated Tregs (0.03% of all cells) and two resting Tregs (0.01% of all cells) as final characterized events (Figure 24A). We next wanted to check if the detected events represent real staining or are attributed to imaging artefacts. Since the position of each single cell is recorded in combination with the intensity values, we were able to track the gated cells back to their tissue localization (Figure 24B). This procedure allowed us to analyze the quality of the original images and to prove the reliability of the seven identified regulatory T cells by the original staining data (Figure 24C).

4.2.5.6 Clustering of cell populations for unbiased cell type identification

Subjective gating as qualitative or quantitative analysis approach is still knowledge-driven and therefore prone to biased analyses towards previously defined populations and phenotypes. This problem can be overcome by an unbiased, unsupervised clustering of highly multiplexed tissue staining data as it is routinely performed for complex datasets like for example from scRNA sequencing experiments.

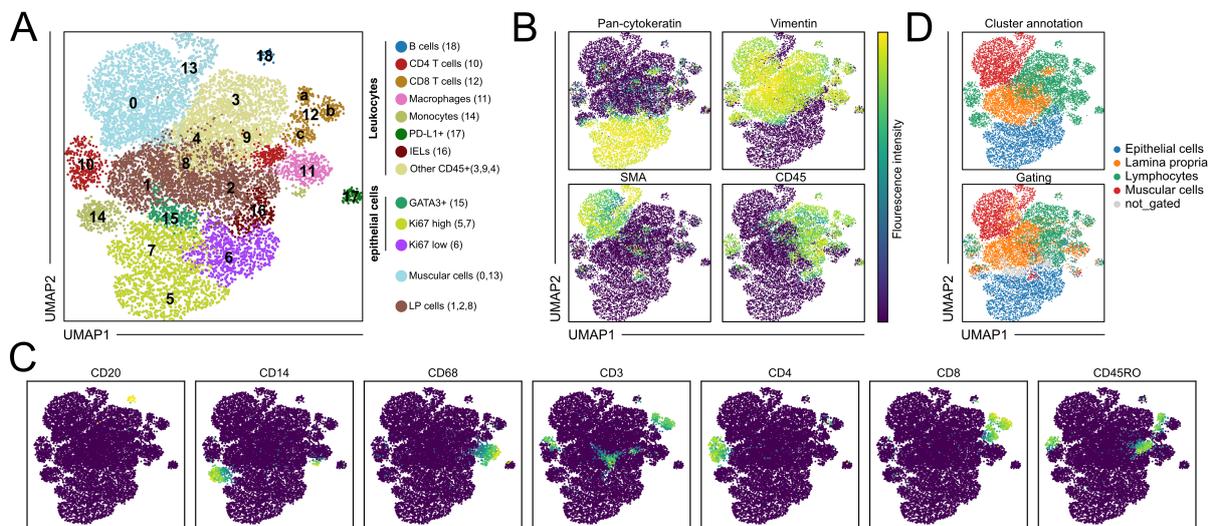


Figure 25. Neighbourhood embedding and clustering analysis for unbiased cell classification.

(A) Protein expression of 18 markers, was used as input to perform neighbourhood embedding and leiden clustering. Detected clusters were then annotated according to the fluorescence intensity per cluster. Subclusters of cluster 12 are annotated as a, b and c. (B) UMAPs showing distribution of pan-cytokeratin, vimentin, SMA and CD45 fluorescence intensities. (C) Fluorescence intensity distribution of markers that were used for annotation of clusters as depicted in (A). (D) Cell type annotation based on clustering together with the intensity data shown in (B) (top) is compared with gated populations is shown (bottom).

We performed an embedding based on expression of 18 proteins in inflamed colon biopsies via uniform manifold approximation and projection (UMAP) (Becht et al., 2018), using tools from the field of scRNA sequencing analysis. Leiden clustering (Traag et al., 2019) of the neighborhood-graph revealed a total

of 19 clusters, annotated to 13 cell types (Figure 25A) according to their marker expression (Figure 25B and C). The cluster annotation led to similar results compared to manual gating (Figure 25D).

Besides the cell type classification, this approach allows to explore additional variation in the individual cell types, as it can be detected in the CD8 T cell population/cluster 12 (Figure 25A). Here, we see three subclusters: CD3+CD8+CD45RO+ effector/memory CD8 T cells (cluster 12a), CD3+CD8+CD45RO-naïve CD8 T cells (cluster 12b) and CD3-CD8+CD45RO± cells (cluster 12c). The latter one could either be CD8-expressing NK cells (McKinney et al., 2021), which could be checked by additional staining for CD56. Or they could represent intestinal intraepithelial lymphocytes (Lin et al., 1994) that do not express CD3, which would perfectly fit in this scenario and would be in line with the additional pan-cytokeratin expression that can be detected in the cluster (Figure 25B). This population would have been missed by knowledge-driven gating approaches, since usually only CD3 positive cells would have been checked for CD8 expression, underlining the advantage of unbiased clustering approaches.

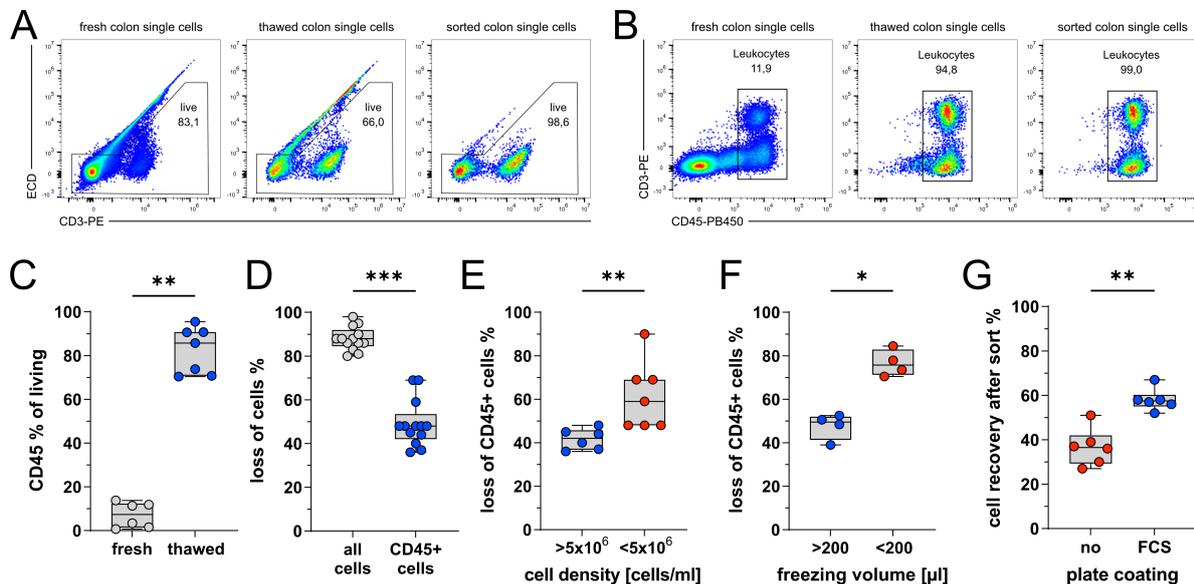
Taken together, we established with ChipCytometry a robust method for multiparametric staining of FFPE tissue samples. The high quality of staining allowed the further development of methods for automatic quantification and rare cell detection, which will facilitate the analysis of the immune infiltrates in GvHD patient samples.

4.3 Method development for scRNA sequencing from gastrointestinal biopsies

Tissue staining has the great advantage of preserving the *in situ* spatial information of the analyzed cells but, despite high multiplexing, the number of markers that can be simultaneously analyzed is still limited. For this reason, we decided to complement the information obtained by tissue staining with single-cell RNA sequencing, the state-of-the-art method for deep phenotyping of single cells. The whole transcriptome of each single cell can be investigated using this method, and further supplementation with paired TCR data or surface protein quantification is possible. Gastrointestinal biopsies are limited in size and the number of infiltrating immune cells might be very low after aHSCT conditioning, depending on the timepoint after aHSCT and the GvHD status of the patient. We therefore carefully optimized single-cell extraction procedures from gastrointestinal biopsies in order to analyze as many single cells per biopsy as possible with scRNA sequencing.

4.3.1 Cell extraction from gastrointestinal biopsies

The generation of single cell suspensions from tissues is a procedure that should be completed right after the collection of the tissue of interest in order to preserve high viability. In this project, cell extraction was performed at the University of Regensburg using mechanical disruption and collagenase digestion of tissue samples.



Exemplary FACS plots for single cell suspensions from colon tissue are shown for living cells (A) and leukocytes (B). Plots represent from left to right: freshly isolated cells, thawed cells and CD45+ sorted cells from the same material. (C) Percentage of CD45 positive cells is plotted for fresh and thawed material. (D) Cell loss in percent is plotted for all cells and CD45+ cells. (E-F) Loss of CD45 positive cells is plotted according to cell density in 1 ml freezing volume (E) and freezing volume for 5x10⁶ cells/ml (F). (G) Cell recovery after FACS sorting is depicted with and without FCS coating of the plate. Statistical testing in (C-G) was conducted by Mann-Whitney U test. (*p<0.05; **p<0.01; ***p<0.001; ****p<0.0001).

Single cell suspensions from multiple patients needed to be collected at different days before scRNA sequencing could be performed on a valuable number of cells, making it unavoidable to freeze the cells immediately after extraction. In this way, several samples can be thawed, cell-sorted and processed for scRNA sequencing in a pooled manner, guaranteeing sufficient cell numbers for a scRNA sequencing sample. Cell sorting in this procedure will ensure a pure population and defined cell number of living, CD45 positive cells for further processing. In order to minimize the impact of this procedure on our samples, we analyzed in detail the effect of freezing/thawing cycles on the quality of the generated single cell suspension, as cell viability is an important prerequisite for successful scRNA sequencing.

As expected, freezing and thawing reduced the total number of cells and the percentage of living cells, which could be selected by flow cytometry-based sorting prior further processing for sequencing (Figure 26A). Importantly, most of the cells suffering from the freezing procedure were non-immune cells as the percentage of CD45+ cells significantly increased after thawing (Figure 26B-C). This effect resulted in a cell-recovery of around 50% for CD45+ cells after thawing, which lies in the expected range for frozen cells (Figure 26D). We further checked on the influence of cell density on cell viability during freezing and found higher densities leading to improved cell survival (Figure 26E). For low cell numbers, which are expected from the gut biopsies used in this study, freezing volume therefore would need to be reduced. However, we found an increased cell loss when reducing the volume beyond 200 μ l (Figure 26F). Altogether, as optimal compromise, we decided to freeze the cells in 250 μ l independent of the cell counts. The cell loss appearing during the sort due to cell recovery afterwards could be further reduced by coating the sort plate with FCS (Figure 26G).

4.3.2 Pilot scRNA sequencing experiment

Before we started analyzing the precious gut biopsies from aHSCT patients, we performed a pilot experiment in order to find out if the pooling of patient samples for scRNAseq and the following demultiplexing with only HLA information would suffice for efficient donor separation. Additionally, we tested if freezing/thawing cycles might introduce a bias in gene expression compared to a fresh sample.

4.3.2.1 Evaluation of sample demultiplexing

PBMCs from four donors with known HLA status were pooled before scRNA sequencing in order to test demultiplexing performance in such an experiment (Figure 27A). SoupORcell (Heaton et al., 2020) was used to cluster cells according to single nucleotide polymorphisms (SNPs), which are unique for each individual, resulting in four distinct clusters and some detected doublets (Figure 27B). To allocate each cluster to one of the initial donors, HLA genotyping was performed on each cluster and a comparative score including the matching of HLA alleles together with gender allocation (see 3.2.5.4) was used for final re-allocation (Figure 27C). With this pilot experiment we were able to set up an analysis pipeline for demultiplexing of unlabeled cells based on SNPs and re-allocation via HLA haplotype.

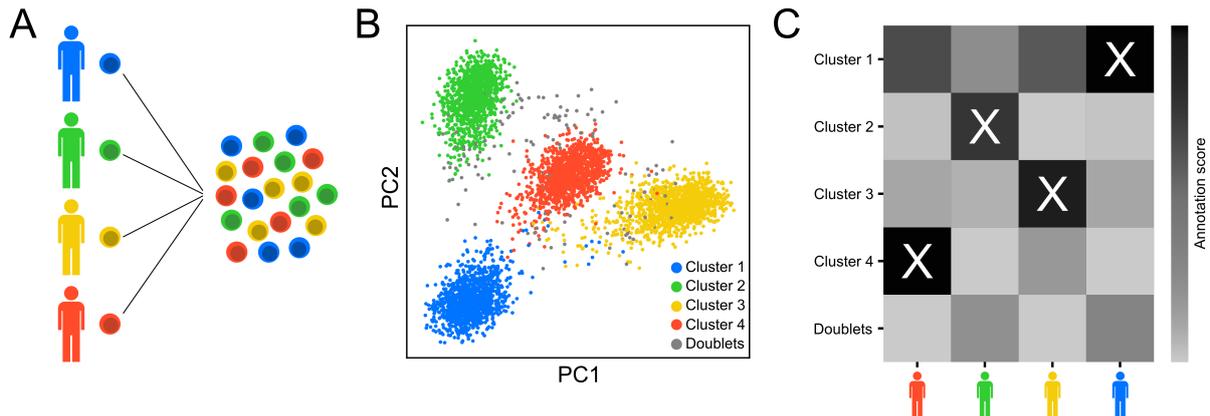


Figure 27. Pooling and demultiplexing of PBMC samples in scRNA sequencing.

(A) Schematic of the experimental setup: 7500 PBMCs from four individuals were sorted and processed as sample pool with scRNA sequencing. (B) PCA analysis on the single nucleotide polymorphisms of the individual donors resulting in four clusters annotated by soupcluster. (C) An annotation score matrix was used to allocate each cluster to one of the donors according to the HLA genotype. The X marks the matching combination between donor and cluster.

4.3.2.2 Transcriptomic comparison of fresh and frozen cells

The second point we wanted to address with the pilot experiment was to evaluate the transcriptomic bias induced by freezing and thawing of cells. Therefore, we compared the transcriptome of fresh and frozen PBMCs from the same patient with scRNA sequencing (Figure 28A). We found a good correlation between the mean gene expression in the fresh and the frozen sample except of a few genes (IFITIM, FOS, JUN, JUNB, DUSP1) related to apoptosis (Ameyar et al., 2003, Gil-Araujo et al., 2014, Park et al., 2016), that were upregulated in the frozen samples (Figure 28B). A similar good correlation was detected when comparing the cell type frequencies of the two conditions (always keeping in mind that blood was drawn at different timepoints) for both high and low abundant cell types (Figure 28C). Others have later confirmed these findings, identifying DMSO cryopreservation as the optimal storage method for samples processed with droplet-based scRNA sequencing methods (Wohnhaas et al., 2019).

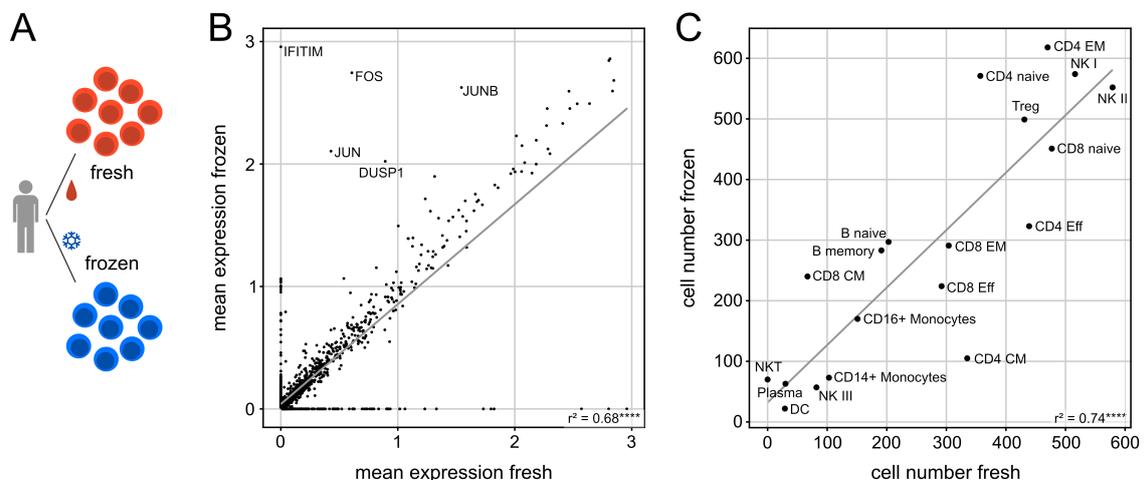


Figure 28. scRNA sequencing data from fresh and frozen PBMCs.

(A) Schematic of the experimental setup: fresh and frozen PBMCs from the same donor were processed with scRNA sequencing. (B) Gene expression correlation between the fresh and frozen sample. Plotted is the mean expression over all cells. (C) Cell number per annotated leiden cluster are plotted for fresh vs frozen PBMCs. Correlation analysis in (B) and (C) was done by Pearson's correlation (**** $p < 0.0001$).

4.4 Clinical parameters of the analyzed patient cohort

After establishing the methodologies for FFPE tissue staining via ChipCytometry (see 4.2) and single cell RNA sequencing (see 4.3), we applied these methods to the available cohort of human gut biopsies from GvHD patients (Figure 29A). Fresh biopsies were used for single cell extraction and subsequent single cell RNA sequencing whereas FFPE biopsies were used for multiplexed tissue imaging via ChipCytometry. The main parameters analyzed were GvHD severity, antibiotic treatment (ABX) and microbial diversity. Time points of sample collection were balanced between early and late biopsies (Figure 29B-C).

GvHD severity was clinically scored (see 1.2 and Figure 2) into mild (grade 1-2) and severe (grade 2-4), and correlated with increased levels of GvHD biomarkers like tumorigenicity 2 (ST2) (Vander Lugt et al., 2013), regenerating islet-derived 3-alpha (Reg3a) (Ferrara et al., 2011), and the MAGIC algorithm probability (MAP) score, which combines these two markers (Figure 29D). Microbial richness and alpha diversity were evaluated both via 16S RNA sequencing of stool material and urinary 3-Indoxyl-sulfate, which is an indirect parameter for microbial diversity (Weber et al., 2015). Remarkably, these clinical parameters regarding the microbiome were significantly reduced in antibiotic treated patients, revealing the causal link between the use of broad-spectrum antibiotics and microbiome disruption (Figure 29E).

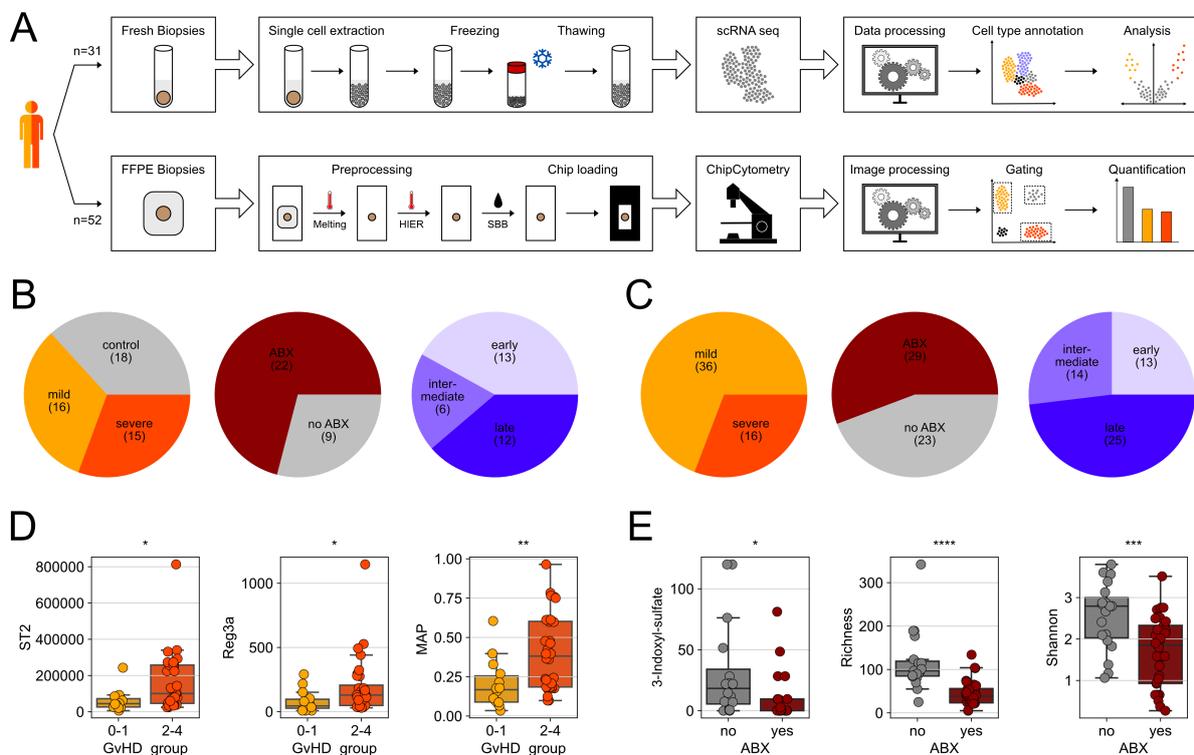


Figure 29. Human patient samples analysed in this study.

(A) Sample numbers and the workflow of the two analysis methods is depicted. (B) Patient characteristics of samples included in the scRNA dataset. (C) Patient characteristics of samples included in the ChipCytometry dataset. (D) Clinical parameters at the timepoint of the biopsy are correlated with GvHD severity. (E) Clinical parameters correlated with antibiotic treatment. Statistical testing in (D) and (E) was conducted by Kruskal-Wallis H test (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

4.5 Multiparametric analysis of GvHD biopsies

For the characterization of infiltrating immune cells, we extracted single cell suspensions from 31 aHSCT patient biopsies, sorted CD45 positive cells and performed single cell transcriptome, VDJ region and surface expression analysis by sequencing (Figure 30A). As controls, we used healthy adjacent tissue from colon resections and we integrated the scRNA sequencing data from the human gut atlas (James et al., 2020) containing data from healthy GI tissue samples.

After filtering and data processing, more than 22.000 cells from distinct immune cell lineages were left, representing tissue samples with different GvHD severity, from different time points after transplantation as well as from different anatomical locations within the gastrointestinal tract (Figure 30B). The distribution of cells revealed already some differences regarding GvHD severity, interestingly independent from the anatomical site. For example, an enrichment of CD8 T cell in the severe GvHD compartment can be observed from the UMAP, including cells from different anatomic locations and timepoints after SCT (Figure 30B).

Cell type annotation was performed via leiden clustering and annotation according to known cell type specific genes, for example CD19 and MS4A1/CD20 for the B cell compartment or CD3 and TCR constant chain expression for T cells (Figure 30C). Subclassification was further performed on differentiation genes, activation signatures or subset specific genes like FoxP3 and CTLA4 for Tregs (Figure 30D). The expression of surface markers further confirmed the annotation of cells via separating the annotated lineage populations independent from the leiden clustering (Figure 30E).

Using the ChipCytometry technology, we were able to analyze immune infiltrates within their spatial environment in GvHD biopsies, allowing us to quantify infiltration densities instead of absolute/relative cell numbers. Gastrointestinal biopsies of GvHD patients were thereby processed with the cyclic immunofluorescence approach, followed by automated quantification of the stained immune cell populations (Figure 30F). Hereby, we focused on the T cell compartment as significant immune population in the course of GvHD. The tissue architecture was visualized by markers like Hoechst (nuclei), pan-Cytokeratin (epithelium) and smooth muscular actin (muscular layer), whereas immune cells were stained by specific markers for each cell type. The visualization of CD4 and CD8 T cells showed already a difference in the CD4/CD8 ratio between the mild and the severe GvHD sample in line with our first observations from the scRNA seq dataset (Figure 30G).

Results - Multiparametric analysis of GvHD biopsies

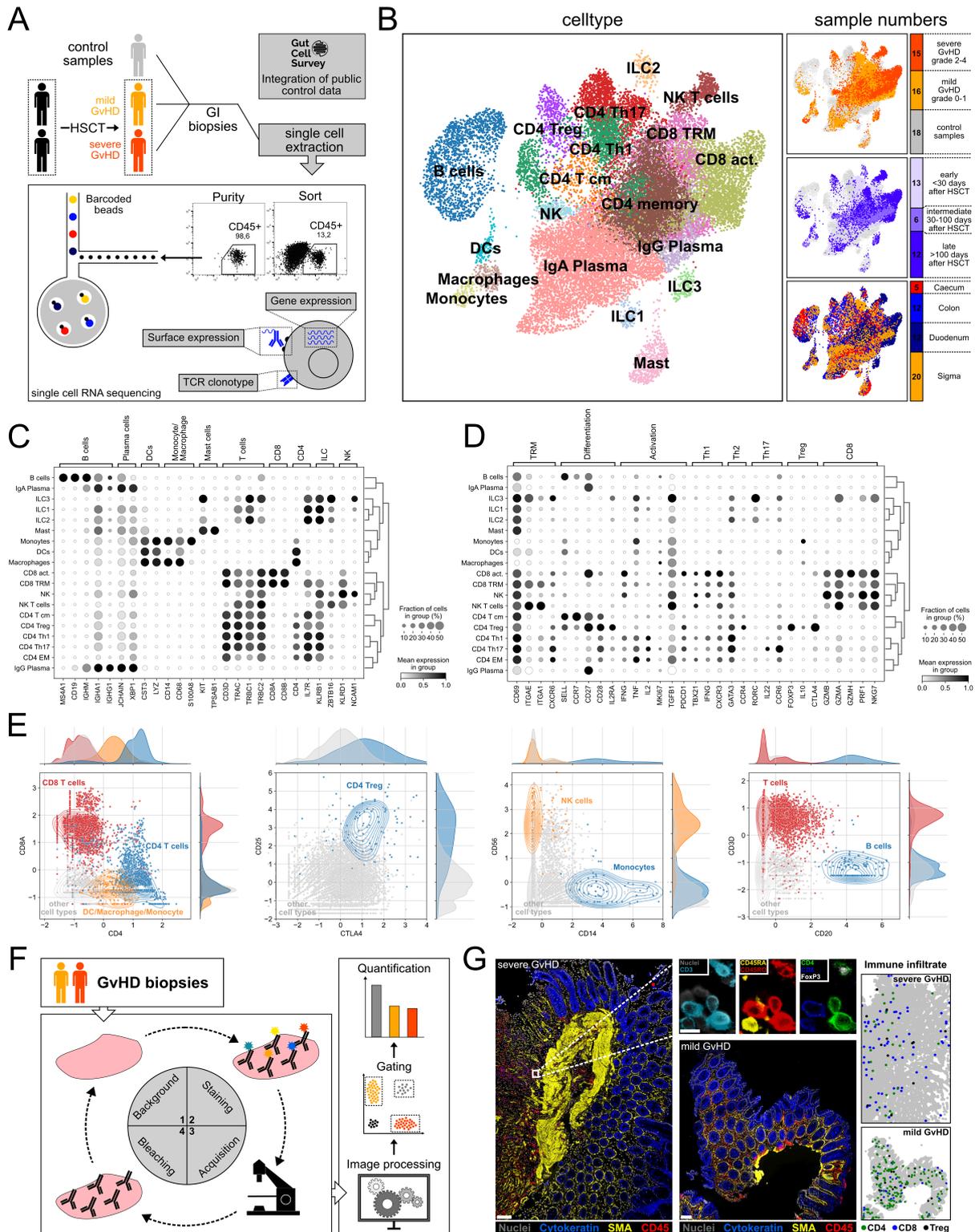


Figure 30. Multiparametric analysis of GvHD biopsies by scRNA seq and ChipCytometry.

(A) Schematic workflow of the scRNA sequencing experiments: Single cells were first isolated from fresh biopsies, sorted on living, CD45+ cells and processed with the scRNA seq workflow. (B) Annotation of cell types (left) as well as sample distribution according to GvHD severity, time point of collection and anatomical origin (right) are depicted on the UMAP. Cell type annotation was performed according to cell type specific marker genes (C) and sub-differentiation was according to phenotypic markers (D). (E) Validation of the annotated cell types by surface marker expression plots for lineage-specific surface markers. (F) Schematic representation of the ChipCytometry experiments: Cyclic immunofluorescence of gastrointestinal biopsies was followed by automated quantification of infiltrates. (G) Exemply image of a severe (left) and mild (right) GvHD sample. The zoom-in demonstrates multiplexed single cell resolution of the applied method. Plots on the right show the distribution of CD4, CD8 and Treg cells within the tissue samples. Scale bars indicate 100 μm in the overviews and 10 μm in the zoom-in.

4.6 GvHD is mainly linked to changes in T cell frequencies

When we compared cell type frequency distributions for the individual patients between control, mild and severe GvHD samples, we observed some group-specific differences like the higher frequencies of plasma cells in control samples and the high frequency of activated CD8 T cells and CD4 memory T cells in severe GvHD samples (Figure 31A).

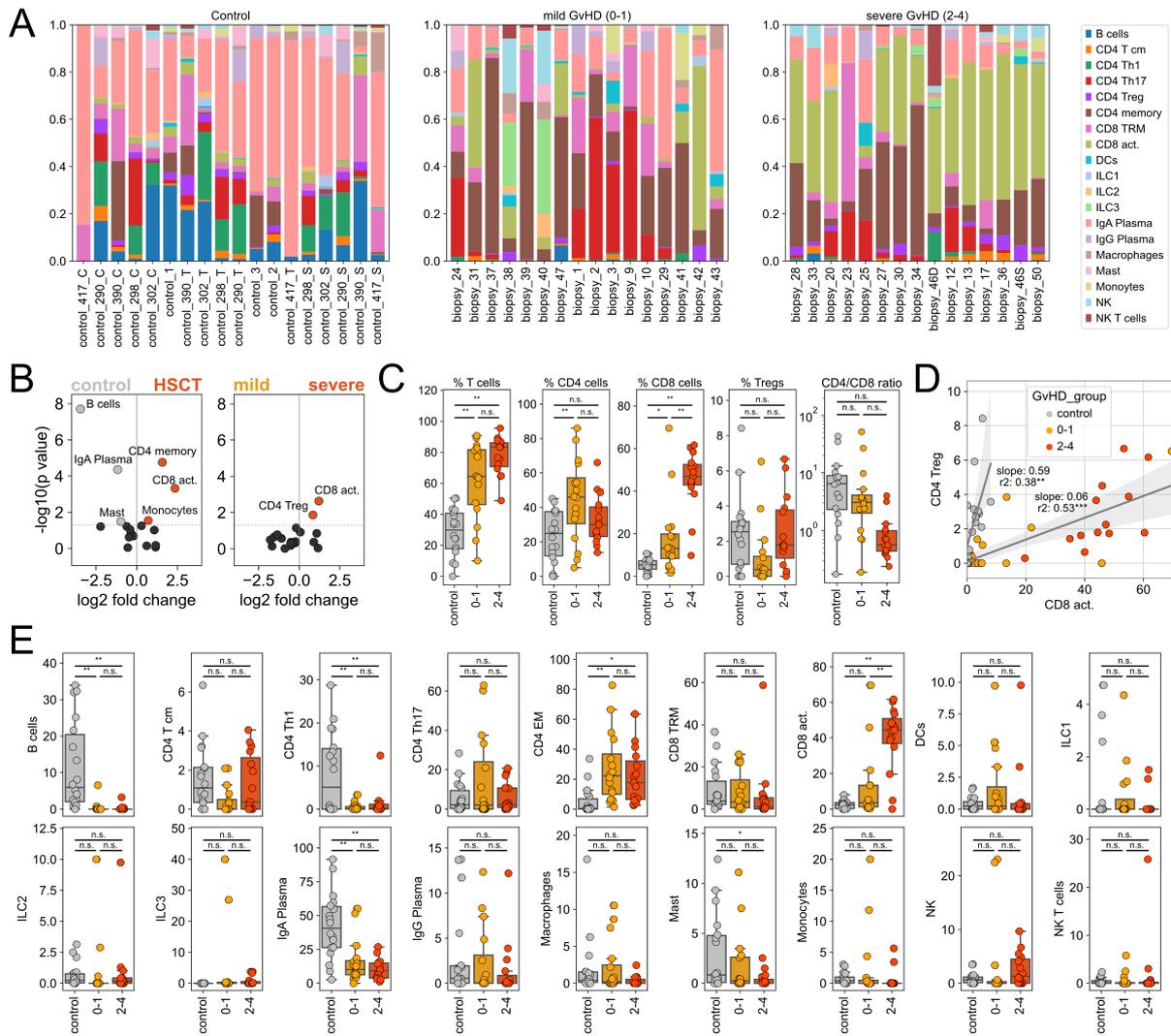


Figure 31. GvHD is mainly linked to changes in T cell frequencies.

(A) Cell type frequencies are depicted as bar plots for each individual sample, grouped into control, mild and severe GvHD samples. (B) Volcano plots show differential cell type frequencies between control and aHSCT samples (left) or between mild and severe GvHD (right). (C) Frequencies of infiltrating T cell subsets grouped by GvHD severity are depicted. Each dot represents an individual sample. (D) Linear regression analysis between activated CD8 T cell and regulatory CD4 T cell frequencies in aHSCT vs. control samples. (E) Frequencies of all other annotated cell populations grouped by GvHD severity are depicted. Each dot represents an individual sample. Statistical testing in (B) was conducted by Kruskal-Wallis H test. Statistical testing in (C) and (E) was conducted by one-way ANOVA followed by Tukey's HSD multiple comparison. Correlation analysis in (D) was done by Pearson's correlation. (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

These differences were quantified by calculating fold change differences between control and aHSCT patients and between mild and severe GvHD samples (Figure 31B). Besides the fact that the control

samples had significantly higher frequencies of B cells and plasma cells, which is the compartment of slowest reconstitution after aHSCT (Fiorenza and Turtle, 2021), we found especially the T cell compartment to be modified in association to GvHD severity. Indeed, GvHD severity was significantly linked to an increase in CD8 T cell frequencies, resulting in a drop of the CD4/CD8 ratio (Figure 31B-C). Furthermore, we observed a significant increase in regulatory T cell frequencies in severe compared to mild GvHD patients. The increased Treg frequencies strongly correlated with frequencies of activated CD8 T cells, suggesting a counter-regulation of inflammation (Figure 31D). Intriguingly, the slope of this correlation was reduced in aHSCT patients compared to healthy control samples. This speaks in favor of an insufficient regulation of inflammation by regulatory T cells, which is obviously not able to suppress GvHD.

Although not significant, we detected other interesting features of cell frequency differences between the three sample groups. Those are for example the gradual decrease of mast cells as well as the association of NK cells to the severe GvHD patients (Figure 31E). NK cells are known for a bilateral association in GvHD development, and can participate in GvHD reduction by repressing alloreactive T cells as well as promote GvHD by contribution to an inflammatory environment (Simonetta et al., 2017). This study will now focus on the significant differences in the T cell subsets, leaving space for follow-up studies investigating other aspects within the valuable datasets produced in this thesis.

The quantification of cell densities per square millimeter in the 52 analyzed biopsies showed significantly higher leukocyte and T cell infiltration in severe cases of GvHD. We could again detect a skewed CD4/CD8 ratio primarily dependent on a higher infiltration of CD8 T cells in severe GvHD, in contrast to the otherwise more stable CD4 T cell numbers (Figure 33A). The Treg infiltration also showed a trend of increase with GvHD severity, which was correlated with CD8 T cell infiltration as observed before in the single-cell RNA sequencing dataset (Figure 33B).

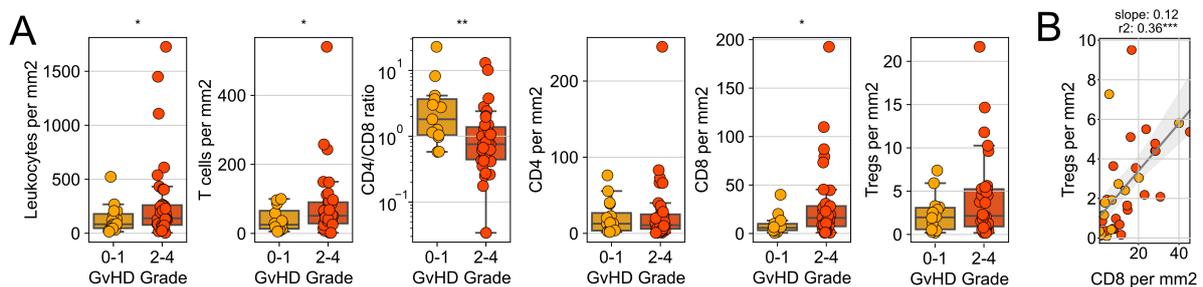


Figure 32. Analysis of T cell infiltration by automated quantification of ChipCytometry data.

(A) Quantified infiltration differences between mild and severe GvHD. (B) Correlation between CD8 T cell infiltrates and regulatory T cells. Statistical testing was conducted by Mann-Whitney U test. Correlation analysis in (B) was done by Pearson's correlation. (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$).

4.7 Antibiotic treatment and immune reconstitution contribute to T cell variance

We performed principal component analysis (PCA) on the data in order to find the clinical parameters explaining the main variance of the dataset. As input data for the PCA analysis, we used information generated by ChipCytometry such as the density of leukocytes, T cells, T helper cells, cytotoxic T cells and Tregs, as well as the CD4/CD8 ratio and the frequency of activated Tregs within the Treg compartment. Importantly, we found that samples do not cluster depending on GvHD severity (Figure 33A). Additionally, the principal components of the data were composed of parameters of general T cell infiltration on PC1 and parameters connected to regulatory T cells on PC2 (Figure 33B). To predict the clinical variables causing the highest variance within the dataset, we calculated the fold change in the single PCs for every clinical parameter, meaning that all parameters were divided in two groups (true vs. false, high vs. low, mild vs. severe) and the fold change between the mean value of the principal component was calculated between the two groups (Figure 33C-D). On the one hand, we found the time, in terms of days after SCT, to be significantly connected to PC1 and thereby with T cell infiltration. On the other hand, antibiotic treatment (ABX) as induction of dysbiosis, and urinary 3-indoxyl sulfate levels as indirect marker for microbial diversity were correlated with PC2 and the Treg infiltration (Figure 33C-D). Antibiotic treatment in our cohort is defined as systemic broad-spectrum antibiotic treatment up to seven days before the biopsy was taken, leading to a strong impact on the microbiome, which is why we see the inverse correlation between 3-indoxyl sulfate levels and ABX. According to these findings, we evaluated the absolute counts of infiltrating Treg according to the antibiotic treatment and GvHD severity. In this scenario, we detected an increase of Treg infiltration for severe GvHD cases in patients without antibiotic treatment as observed before. However, if the patient received systemic antibiotic treatment, this effect was suppressed (Figure 33E). Taken together, the analysis of the factors that explain the variance in the cohort showed that the time of immune reconstitution should be considered when analyzing biopsies from different time points and underlines the importance of the microbiome, especially in regard of regulatory T cells.

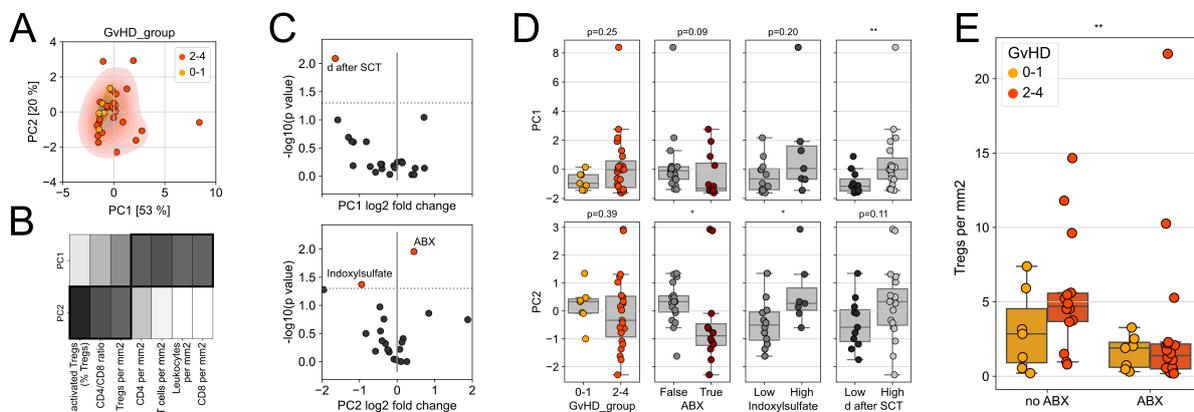


Figure 33. Analysis of gastrointestinal biopsies from GvHD samples via ChipCytometry.

(A) PCA analysis of the ChipCytometry dataset. The color indicates the GvHD severity of the corresponding patient. (B) PCA loadings are depicted as heatmap for PCA analysis in (A). (C) Volcano plot shows the fold change of clinical parameters along components 1 and 2. (D) Scatterplot for significantly associated clinical parameters. (E) Treg frequencies are plotted for the individual patients grouped on Antibiotic treatment and GvHD severity. Statistical testing was conducted by Mann-Whitney U test. (* $p < 0.05$; ** $p < 0.01$).

4.8 Antibiotic treatment reduces suppressive capacity of Tregs

After observing a connection between antibiotic treatment and abundance of the regulatory T cell infiltrate, we tried to investigate the mechanism behind this phenotype. We observed that antibiotic treatment induced a significant reduction in microbial diversity (Figure 29E), as also shown by others. In addition, the microbiome plays a crucial role in inducing and sustaining Treg phenotypes in the gut (Omenetti and Pizarro, 2015). Thereby, it is reasonable to speculate that the antibiotic-induced dysbiosis could impair Treg functions.

As first step, we analyzed the phenotype of *in situ* infiltrating Tregs. Tregs could be identified in the ChipCytometry dataset by the expression of CD3, CD4 and FoxP3, and were further sub-characterized into activated and resting Tregs by the expression of CD45RA/CD45RO (Miyara et al., 2009) (Figure 34A). Notably, we observed a significant reduction in the frequency of activated Tregs in patients that had received a systemic antibiotic treatment (Figure 34B), again linking antibiotics-induced dysbiosis to the Treg infiltrate, but this time in regard of functionality.

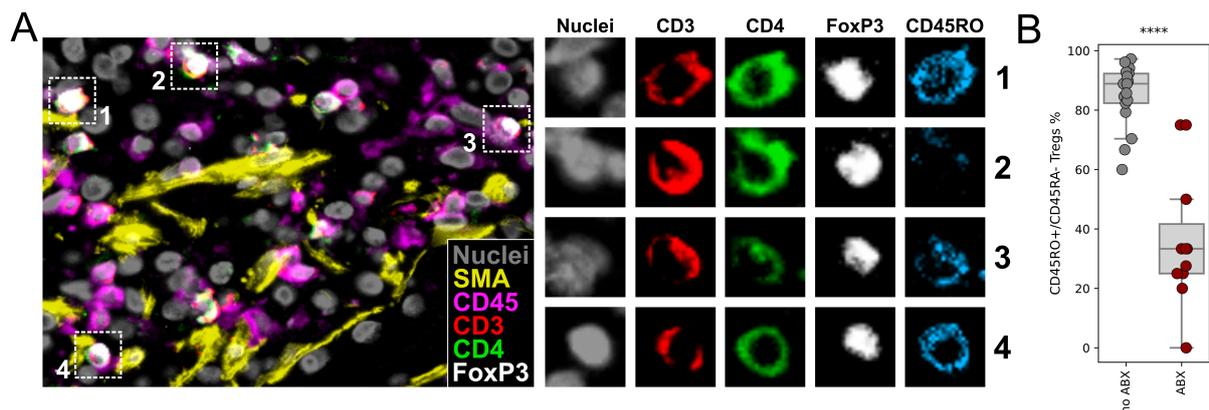


Figure 34. Treg heterogeneity in the ChipCytometry dataset is correlated to antibiotic treatment.

(A) ChipCytometry image of 4 regulatory T cells defined by the expression of markers shown on the right for each individual cell. (B) activated Treg frequency is shown for the ChipCytometry dataset. Statistical testing in (B) was conducted by Mann-Whitney U test (**** $p < 0.0001$).

To get more insights, we next analyzed the transcriptomic profiles of regulatory T cells. For the purpose, we extracted 207 regulatory T cells for in-depth analysis and re-clustered them according to their transcriptome. Leiden clustering separated the cells into three distinct clusters that showed slight differences in FOXP3 expression (Figure 35A). Tregs from antibiotic-treated patients were enriched in clusters 1 and 2, which showed reduced expression of FOXP3. GvHD groups were instead more equally distributed among the clusters (Figure 35A). Differential gene/surface antigen expression analysis between patients with or without systemic antibiotic treatment showed a suppression of genes or proteins associated with Treg phenotype or functionality (e.g., FOXP3, TNFSF4, CTLA4 gene expression and CD134, CD25, CD39, CCR4 and CXCR3 in surface markers) in antibiotic-treated patients (Figure 35B). On the contrary, this group showed increased expression of genes associated with conventional T cells (CD69, CD38, CD40) or with Treg instability (IRF1). Therefore, we scored genes associated with Treg

functionality (FOXP3 (Hori et al., 2003), IL2RA (Chinen et al., 2016), CTLA4 (Wing et al., 2008), CD27 (Muth et al., 2022), TNFRSF members (Ono et al., 2006), TIGIT (Yu et al., 2009), BATF (Hayatsu et al., 2017), CCR4 (Sugiyama et al., 2013), CXCR3 (Hoerning et al., 2011)) as well as gene sets for Treg and conventional effector CD4 T cells (Höllbacher et al., 2020). Notably, we found a significantly lower Treg suppression score in patients that had received systemic antibiotic treatment (Figure 35C). Moreover, we found an inverse correlation of Teff and Treg gene scores (Figure 35D), with enriched gene signatures of effector T cells in the antibiotic-treated group. The three Treg clusters thereby significantly differed in the described scores: Clusters 1 and 2 shared an effector-like signature whereas cluster 3 had a significantly higher Treg score and Treg suppression score signatures (Figure 35E-F). Altogether, these data indicate that systemic antibiotic treatment can negatively influence the proper reconstitution of the gut Treg niche after aHSCT.

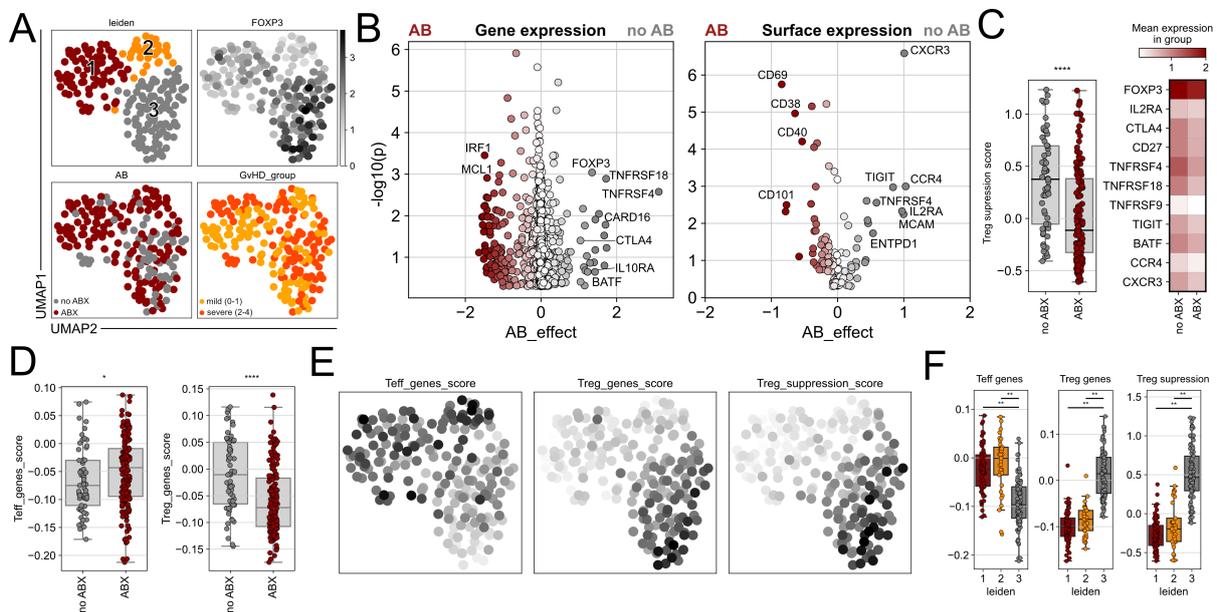


Figure 35. The phenotype of regulatory T cell is influenced by antibiotic treatment.

(A) 207 regulatory T cells have been extracted from the total dataset and were clustered into 3 distinct groups. (B) Differential gene expression analysis (left) and surface expression analysis (right) is shown according to antibiotic treatment (ABX). (C) Literature-defined genes were scored as Treg suppression signature. (D) Teff and Treg gene scores (Höllbacher et al., 2020) are depicted grouped according to antibiotic treatment. (E) UMAPs depict the distribution of the individual scores. (F) Scores are plotted for each individual cell grouped by Leiden clustering. Differential gene expression analysis in (B) was performed using Welch's t-test in diffxpy. Statistical testing in (C) and (D) was conducted by Kruskal-Wallis H test. Statistical testing in (F) was conducted by one-way ANOVA followed by Tukey's HSD multiple comparison (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

We next asked the question whether the observed less suppressive Tregs resulted from a de-differentiation of Tregs due to Treg instability or from an incomplete peripheral differentiation from conventional CD4 T cells. RNA velocity analysis of transcriptional dynamics allowed getting insights into the developmental direction of Tregs in the dataset. The suggested path derived from latent time analysis moved from Leiden cluster 1 via cluster 2 to cluster 3, which was the cluster with the highest Treg suppression score (Figure 36A-B). In line, Treg suppression score and Treg gene score increased with latent time whereas the Teff gene score decreased (Figure 36C). Pseudo-temporal gene expression analysis of

the top 100 velocity-associated genes showed overlap with the Teff gene signature in the early stage and with the Treg signatures in the later stage (Figure 36D), and the directionality of differentiation has even been detected within a single clonotype (Figure 36E). This evidence further confirmed the directionality of differentiation by cells derived from the same progenitor. We further observed reduced differentiation of Tregs in antibiotic-treated patients as well as a higher differentiation in severe GvHD cases, potentially related to counter-regulation of inflammation (Figure 36F). However, when we discriminated severe GvHD patients according to AB treatment, we again found the correlation between antibiotic treatment and Treg differentiation (Figure 36G). Altogether, our data indicate a GvHD-independent effect of reduced suppressive capacity in Tregs upon antibiotic treatment.

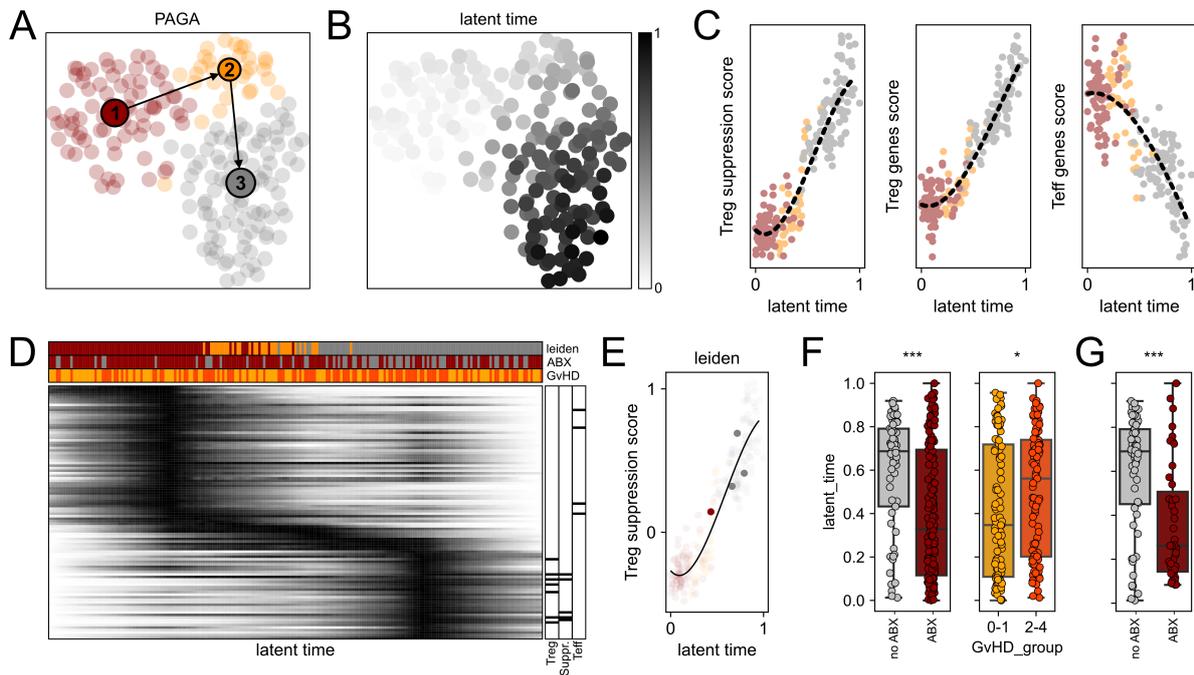


Figure 36. Regulatory T cell differentiation is locked upon antibiotic treatment.

(A) Dynamic modelling of transcriptional dynamics revealed the directionality of regulatory T cell differentiation from the Teff-like cluster 1 to the most suppressive cluster 3. Sown is partition-based graph abstraction (PAGA). (B) Latent time plotted as heatmap on the UMAP space. (C) Latent time is plotted against the gene signatures described before (see Figure 35). Each dot represents one cell and the color indicates the leiden clusters. (D) Heatmap of pseudo-temporal gene expression of the top 100 velocity-defining genes. Each column represents one cell (ordered by latent time) and each row represent one gene. Leiden cluster, antibiotic treatment and GvHD severity are indicated on the top of the graph and overlapping genes with the gene signatures are indicated on the right. (E) One expanded Treg clonotype is highlighted on the latent time vs. Treg suppression score plot also depicted in (C). (F-G) Latent time is plotted grouped by ABX or GvHD severity for all samples (F) and for severe GvHD patients only (G). Statistical testing in (C) and (D) was conducted by Kruskal-Wallis H test. Statistical testing in (F) was conducted by one-way ANOVA followed by Tukey's HSD multiple comparison (* $p < 0.05$; *** $p < 0.001$).

4.9 Certain bacterial species are linked to Treg differentiation

Since we found a significant connection between microbiome disruption via antibiotic treatment and Treg functionality, we wanted to further characterize this interaction on the level of bacterial species from 16S rRNA sequencing of matched stool samples. Therefore, we calculated the mean fold change of each of the three scores according to species presence in a patient for each bacterial species.

According to these fold-changes, species were clustered into five distinct groups: cluster 1, 2 and to a less extend cluster 4 seemed to support Treg development, whereas cluster 3 and 5 were more in favor for the less differentiated Tregs (Figure 37A). Within the clusters supporting Treg differentiation, we found an enrichment of *Clostridia* and *Bacteroidia* species, which have reduced frequencies in the clusters 3 and 5 (Figure 37B). These classes have been described to play an important role in regulatory T cell induction (Atarashi et al., 2011, Pandiyan et al., 2019). The clusters 3 and 5 were in contrast marked by a higher abundance of *Gammaproteobacteria*, which have previously been identified as a GvHD risk factor (Han et al., 2018).

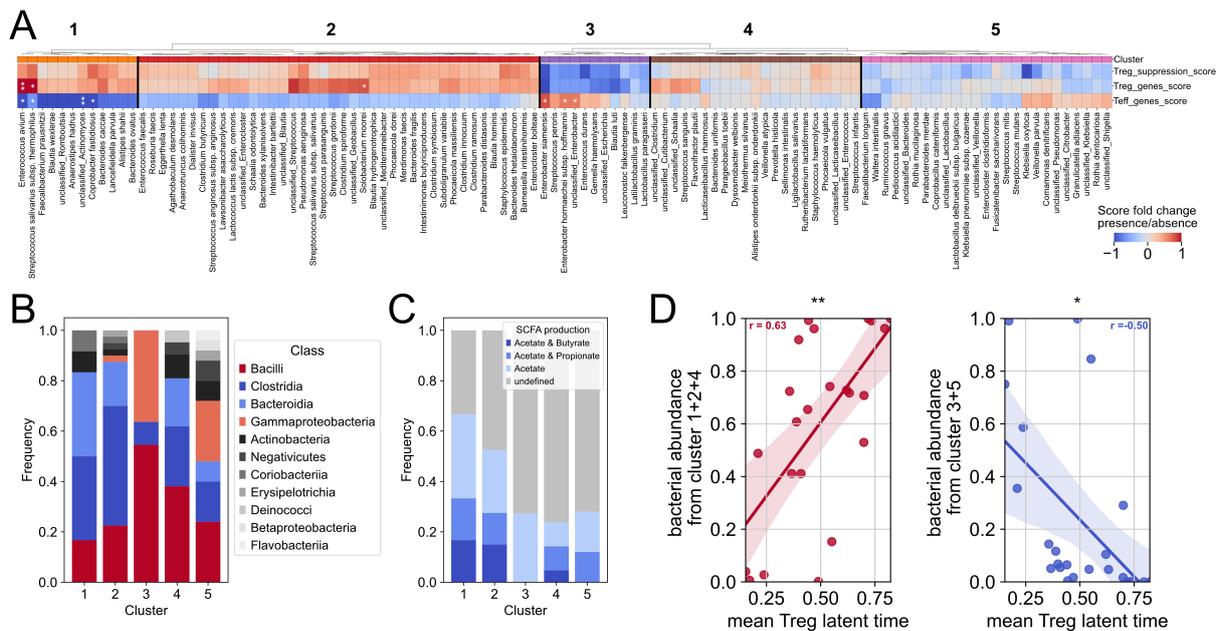


Figure 37. The abundance of certain bacterial species can influence Treg differentiation.

(A) Presence of bacterial species has been associated with changes in the gene expression scores defined in Figure 35. Species were subdivided by hierarchical clustering according to their association pattern into 5 distinct subgroups. (B) bacterial class abundance within the individual clusters is shown. (C) Ability of SCFA production for the individual species was retrieved from the Virtual Metabolic Human database and the frequency of metabolite producing species is depicted per cluster. (D) Linear regression analysis of the bacterial abundance of species from the clusters defined in (A) plotted against the mean latent time of Tregs as measure of differentiation. Statistical testing in (A) was conducted by Kruskal-Wallis H test. Correlation analysis in (D) was done by Pearson’s correlation. (* $p < 0.05$; ** $p < 0.01$).

An often-described pathway of Treg induction via specific microbiota is through bacterial metabolites. We therefore checked at the species level for known metabolites in the gut microbiome resource of the Virtual Metabolic Human database (Magnusdottir et al., 2017, Noronha et al., 2019), which contains

more than 800 semi-automatically curated strain-specific metabolic reconstructions for more than 600 species. By searching for all clustered strains in the database, we found enrichment in butyrate and propionate producing species in clusters 1, 2 and 4, which is in line with the supportive role of SCFAs to Treg development (Arpaia et al., 2013) (Figure 4C). Finally, the abundance of species from the Treg supporting clusters (cluster 1, 2 and 4) showed a positive correlation to the differentiation state of Tregs within a sample (Figure 37D).

In summary, we detected specific groups of bacteria that are positively associated with Treg differentiation in the gastrointestinal tract. These groups are enriched for SCFA producing species, suggesting this axis as potential mechanism of Treg regulation.

4.10 Clonally expanded CD8 T cells are drivers for GvHD severity

Our data indicate that the suppression of GvHD by Tregs could be strongly dampened by the antibiotic-induced gut bacteria disruption, and is therefore a potentially important finding for clinical interventions. On the pro-inflammatory side, we have seen that conventional T cells are significantly higher abundant in severe GvHD, and we therefore wanted to take a deeper look into this compartment. After extracting 7218 T cells from our scRNA seq cohort (Figure 38A), we performed differential gene expression analysis according to GvHD severity. We found cytotoxic effector gene signatures (NKG7, GZMA, GZMB, GZMK, HAVCR2) to be associated with T cells from severe GvHD patients (Figure 38B). Mild GvHD patients in contrast showed enrichment of memory-associated genes (IL7R, KLRB1), in line with the generally lower inflammation.

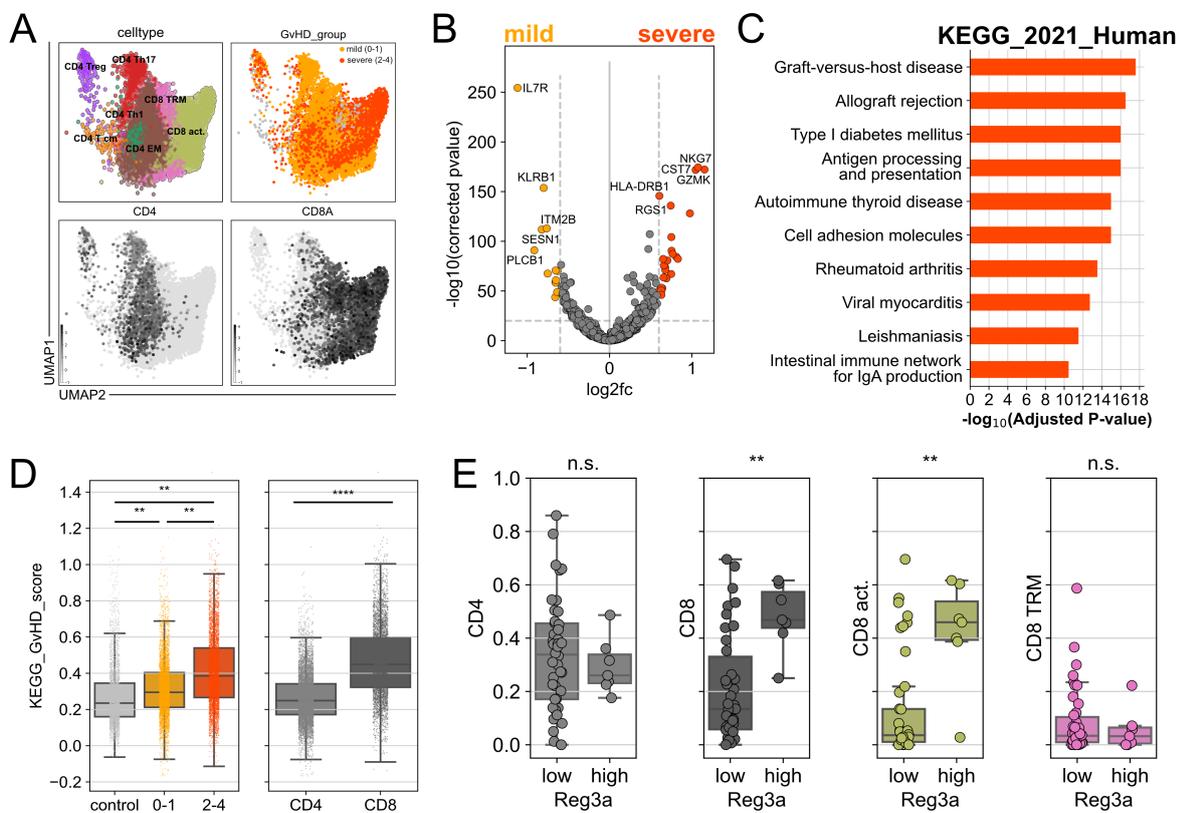


Figure 38. CD8 T cells are associated with GvHD severity.

(A) 7218 conventional T cells have been extracted from the total dataset and the individual cell types as well as marker expression are depicted. (B) Differential gene expression analysis is shown according to GvHD severity. (C) Gene set enrichment analysis for the KEGG 2021 human database was performed on genes that were higher expressed in severe patients as shown in (b). (D) Graft-versus-host disease score from (C) is depicted grouped by GvHD severity (left) or T cell subset (right). (E) Box Blots for cell type frequencies of CD4, CD8, CD8 activated T cells and CD8 TRM T cells (from left to right) are depicted. Patients are grouped according to low or high Reg3a levels. Differential gene expression analysis in (B) was performed using Welch's t-test in diffxpy. Statistical testing in (D) and (E) was conducted by Kruskal-Wallis H test or by one-way ANOVA followed by Tukey's HSD multiple comparison for more than two groups (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$).

Gene set enrichment analysis (GSEA) was next performed on the genes that were enriched in T cells from severe GvHD patients, and graft-versus-host disease was the top hit among other pathways associated with transplantation and autoimmunity (Figure 38C). When we scored the genes from this

pathway on our single cells, we found a positive correlation of GvHD severity with a higher score only in CD8 but not in CD4 T cells (Figure 38D). This is in line with our finding that especially CD8 T cells are associated with GvHD severity (Figure 31B-C and Figure 33B). Finally, the frequency of CD8 T cells were elevated in patients with high Reg3a levels and primarily depended on activated CD8 T cells (Figure 38E), underlining their importance in GvHD pathophysiology by connection to epithelial barrier disruption (Ferrara et al., 2011).

VDJ sequencing allowed us to investigate clonality within the T cell compartment and we found increased clonal expansion with GvHD severity (Figure 39A), including some highly expanded clonotypes with more than 15 cells per TCR. As expected, we could observe a connection between clonal expansion and TCR repertoire diversity for the individual patients, the latter being significantly reduced in severe GvHD samples for the whole T cell pool as well as for the CD8 compartment (Figure 39B). Regarding the T cell subset frequencies between control, mild and severe patients, the increased Treg frequency in control samples and the high proportion of activated CD8 T cells in severe GvHD patients are obviously detectable. These differences in frequencies are accompanied by clonal T cell expansion, which comprised some large clonotypes especially in the severe GvHD patients (Figure 39C). Although the absolute expansion was more obvious in the CD8 T cell compartment (Figure 39B-C), we found a significant relative expansion (clone size divided by number of T cells) in both, CD4 and CD8 T cells, with CD4 effector memory and activated CD8 T cells being the main drivers of expansion, respectively (Figure 39D). Altogether, we could show that activated CD8 T cells mainly drive severe GvHD. Furthermore, the clonal expansion associated with severity of the disease suggests for an antigen-specific inflammation rather than an unspecific activation of the overall T cell compartment.

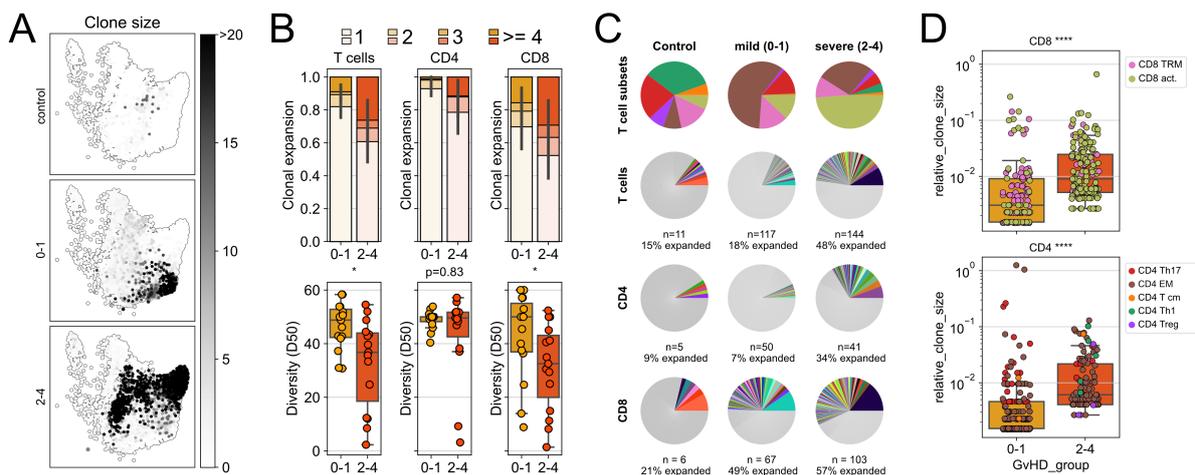


Figure 39. Conventional T cells show enhanced clonal expansion in severe GvHD patients.

(A) UMAPs show clonal expansion of T cells for the different patient groups. (B) Clonal expansion (top) and diversity (bottom) is shown for all T cells (left), CD4 T cells (middle) and CD8 T cells (right). (C) Pie charts depict the T cell type frequencies and clonal expansion in control, mild and severe GvHD patients. Clonal expansion analysis has been performed for all T cells and separated for CD4 and CD8 T cells. Clonotypes with a size smaller than 3 cells are depicted in gray. (D) relative clone size (clone size/number of T cells) is depicted for all individual clonotypes separated into CD8 (top) and CD4 T cells (bottom). Statistical testing in (B) and (D) was conducted by Kruskal-Wallis H test (* $p < 0.05$; **** $p < 0.0001$).

4.11 Expanded T cell clones persist over time and are anatomically spread

For three patients, multiple biopsies have been taken either longitudinally or from different sides of the gastrointestinal tract. This gave us a unique opportunity to investigate the persistence of expanded clonotypes over time and their anatomical spread within the gastrointestinal tract. When we calculated the repertoire overlap of these biopsies, intriguingly, we found some similarity in the TCR repertoire of biopsies from the same patients (Figure 40A).

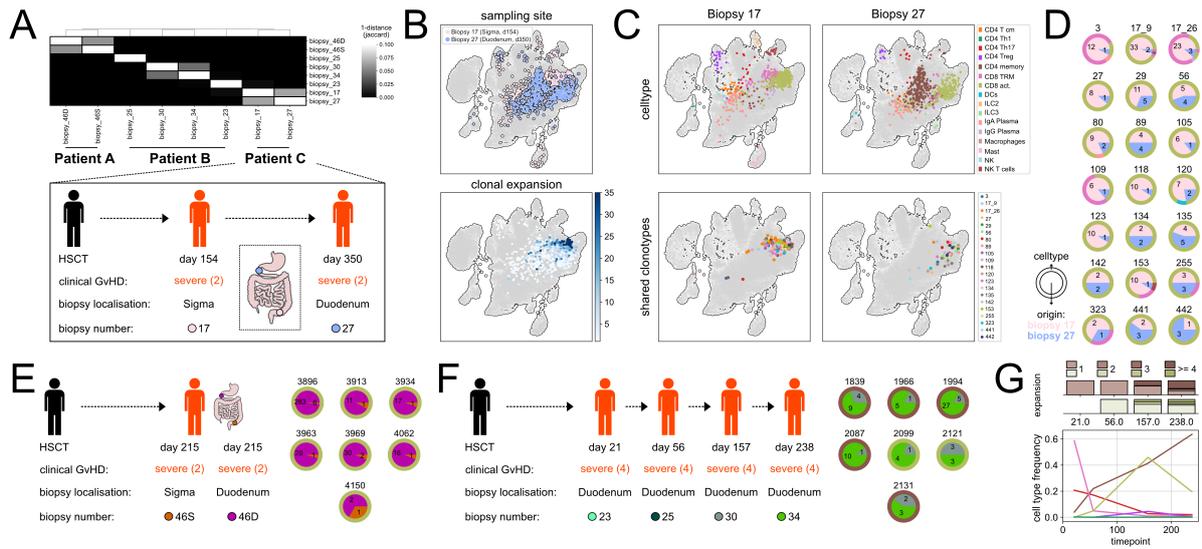


Figure 40. Expanded CD8 T cells are anatomically distributed and stable over time.

(A) The heatmap shows the repertoire overlap between all longitudinal samples (top). The schematic explains sample origins of biopsies 17 and 27 from the same patient. (B) UMAPs depict the distribution of the biopsies (top) and the clonal expansion (bottom) for patient C. (C) Cell type clusters (top) and location of shared clonotypes (bottom) are depicted for both biopsies. (D) All 21 shared clonotypes are depicted with the number of cells from the individual biopsies (inner circle) and the cell type each individual cell has been assigned to (outer ring). (E-F) Overlapping clonotypes and time course of GvHD with the respective biopsies for patient A (E) and patient B (F) are depicted. (G) T cell subset clonal expansion for CD4 effector memory T cells (top), CD8 activated T cells (middle) and T cell frequencies (bottom) are plotted over time for patient B.

For patient C, we in-depth analyzed two biopsies: biopsy 17 was collected at day 154 from the sigma, whereas biopsy 27 was taken at day 350 from the duodenum (Figure 40A). Although these biopsies were taken from anatomically distant sides and over the course of almost 200 days, we found similar immune infiltrates in both biopsies, except of the enriched CD4 effector memory cell frequency in the duodenal biopsy. Furthermore, and surprisingly, we observed 21 clonotypes overlapping between the biopsies (Figure 40B-C). These obviously systemically relevant clonotypes shared most of the time a rather activated phenotype (Figure 40C-D).

We could further validate this global appearance of expanded, activated CD8 T cells in another pair of biopsies from sigma and duodenum (Patient A) taken at the same time point (Figure 40E). For patient B, four longitudinal biopsies were available, and we found some clonal overlap only at the later time points between biopsies 30 and 34 (Figure 40F). Interestingly, here the overlapping clones predominantly derived from CD4 Tem T cells, and this cell type showed an increase in frequencies due to clonal

Results - Expanded T cell clones persist over time and are anatomically spread

expansion over all time points (Figure 40G). Still two out of seven overlapping clonotypes are from the CD8 activated T cell compartment with a similarly increasing clonal expansion over time (Figure 40G).

Taken together, we found clonal overlap between biopsies from the same patients independent from the anatomical location and stable over long time periods. This underlines clinical importance of antigen-specific T cell activation for GvHD pathophysiology, due to systemic relevance and stable antigen presentation, preserving the activated phenotype of the detected CD8 T cells.

4.12 Suppressive capacity of Tregs is directly linked to CD8 expansion

Results from our investigation of a sample cohort of GvHD patients revealed CD8 T cells with an activated signature as key players in GvHD progression. We furthermore found that regulatory T cells, which have the potential to suppress GvHD (Edinger et al., 2003), are dependent on the microbiome and can be suppressed in their numbers and functionality upon microbiome disruption.

In order to link these findings, we analyzed the relative clone size of T cells from severe GvHD patients according to the suppressive capacity of co-infiltrating Tregs in the same patient. Here we found a significantly lower clonal expansion for CD8 – but not CD4 – T cells in patients with high suppressive potential in the Treg compartment (Figure 41A).

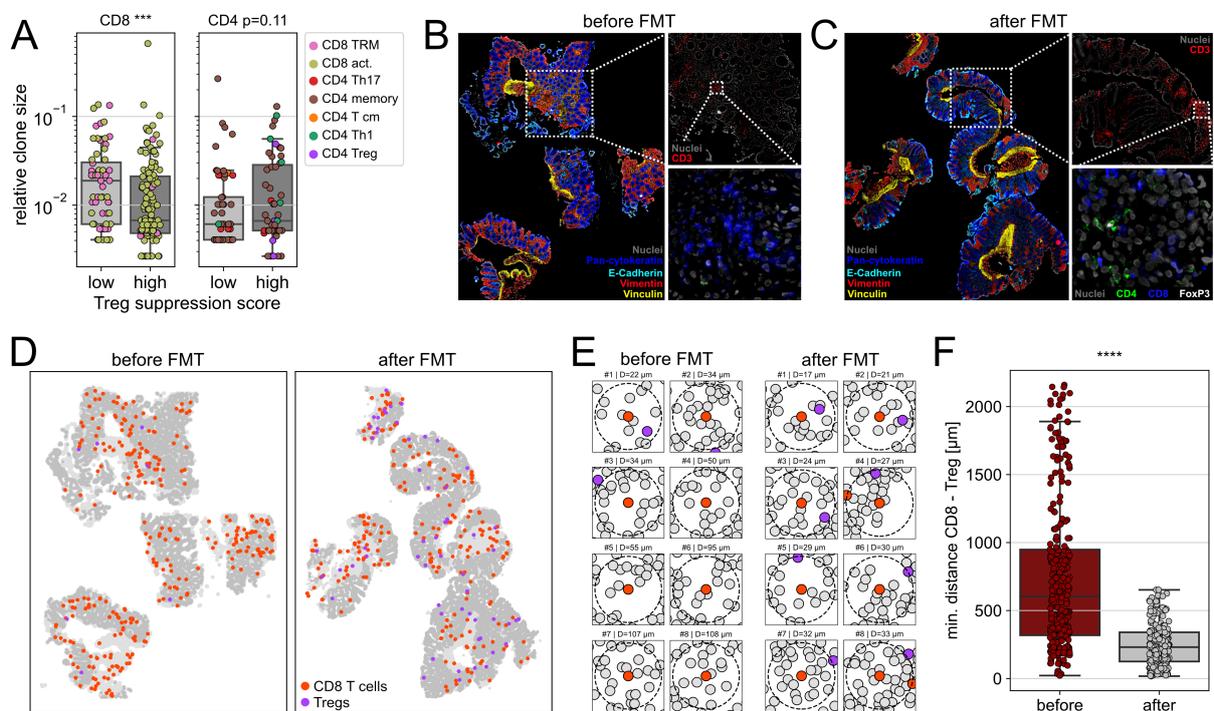


Figure 41. Direct connection between Tregs and CD8 T cell expansion in severe GvHD.

(A) Relative clone sizes are depicted for CD8 T cells (left) and CD4 T cells (right), grouped by the patients Treg suppression score being higher or lower than the mean. (B-D) Original ChipCytometry images of biopsies from a patient before (B) and after FMT (C) as well as the replotted, gated T cell populations (D) are depicted. (E) Eight CD8 T cells with the lowest distance to a Treg are depicted before and after FMT treatment. The circle indicates a radius of 30 μm . (F) The boxplot depicts the minimal distance to a Treg for each CD8 T cell before and after FMT treatment. Statistical testing in (A) and (F) was conducted by Kruskal-Wallis H test (** $p < 0.001$; **** $p < 0.0001$). Modified from (Orberg et al., 2022).

These findings highlight the importance of an intact microbiome to preserve a functional gastrointestinal Treg population, and support the potential of FMT or other probiotic approaches as therapeutic options for GvHD treatment. First pilot studies showed increasing abundance of protective bacterial species producing protective metabolites as well as an increase in Treg numbers after FMT treatment (Orberg et al., 2022, Kakahana et al., 2016). To assess whether FMT-induced restoration of gut Tregs could

ameliorate GvHD severity by suppressing CD8 T clones, we performed ChipCytometry on biopsies before and after FMT in a single patient and checked for the spatial organization of Tregs and CD8 T cells.

Firstly, we found a reduced density of CD8 T cell infiltration, accompanied by a higher density of regulatory T cells (Figure 41B-D). Furthermore, we calculated the minimal distance to the next regulatory T cells for each CD8 T cell in order to assess the interaction probability between regulatory and CD8 T cells. This analysis revealed more CD8 T cells having a Treg within a distance of 30 μm after FMT, presumably allowing interaction and suppression of CD8 T cells by spatial proximity (Figure 41E). Furthermore, we observed a significantly reduced distance between CD8 T cells and Tregs over the whole analyzed tissue (Figure 41F).

Taken together, with these analyses, we were able to link our two major findings by observing the suppression of clonal expansion of the systemically spread and disease driving CD8 T cells by differentiated regulatory T cells, dependent on an intact microbiome. We could further observe in a single patient, that FMT can restore microbial diversity and thereby lead to higher Treg frequencies. The higher Treg density upon FMT would then be able to favor interaction with CD8 T cells, facilitating a potential suppression of GvHD.

5. Discussion

5.1 ChipCytometry allows deep *in-situ* phenotyping of single cells in gut biopsies

At the start of this project, we had to bring together an existing repository of FFPE tissue samples from human aHSCT patients with ChipCytometry as a newly evolved technology to perform multiplexed imaging with spatial single-cell resolution and superior image quality. The technology advancements in detail are the subtraction of a true background, the high dynamic range imaging, and the ability to store samples for later re-interrogation. The methodology was developed for cryopreserved tissue samples, meaning that we had to generate a new treatment protocol to transfer the technology to FFPE tissue samples of our retrospective cohort. The main changes to the protocol were a tightly titrated step of antigen retrieval and the minimization of tissue autofluorescence to achieve the best possible signal-to-noise ratio.

After setting up a handling protocol for FFPE samples, we established antibodies compatible with the antigen-retrieved tissues. The current set of 32 markers allows the differentiation of more than 30 different cell types in tissue samples, focusing on immune cells. For broader applicability, this list of antibodies might be expanded for the staining and phenotyping of additional lineages, such as innate immune cells and cancer cells or apoptotic markers and checkpoint molecules to facilitate the phenotyping of tumor samples. The technology *per se* can be used with any antibody, as long as it can recognize its target within a fixed tissue sample and is conjugated to a suitable fluorophore. The platform, therefore, allows the study of particular cell types as it has been shown for MAIT cells, for example, identified by antibodies against Va7.2, CD161, PLZF, and CD3 (Leng et al., 2019). Existing repositories like the list of validated antibodies from the CyCIF-methodology (Lin et al., 2018) contain potential candidates for panel expansion.

For low-expressed markers that are difficult to detect via primary antibody staining, we combined the methodology with a fluorescence *in-situ* hybridization (FISH) approach (Wang et al., 2012). This advancement allows three additional targets to be stained, and the cyclic application could enable the expansion of the number of mRNA targets. Depending on mRNA stability, multiple cycles of fluorescent *in situ* hybridization could be performed with a chemical bleaching step to remove the signal and an additional HRP blocking in between each cycle to avoid the detection of probes from a previous channel. An alternative approach for up to twelve mRNA targets is the new RNAscope HiPlex Kit v2, which would enable one step of hybridization for 12 target probes followed by the specific, cyclic detection in rounds of three channels (488, 550, 650, all matched with the ChipCytometry filter sets). This procedure would be fully compatible with the proposed protocol for the combined detection of mRNA and protein targets.

Imaging experiments are often used for qualitative analysis of spatial cell type distribution and infiltration densities based on single markers. Quantitative studies are still mainly based on manual counting, which becomes unrealistic for the depth of phenotyping possible with the ChipCytometry technology.

Therefore, we developed an automated approach that enables us to perform robust quantification of phenotypically complex as well as very rare cell populations within tissue samples. Besides the advancement of superior data quality, we could implement an effective strategy of spatial spillover correction, leading to a more sophisticated quantification approach. This correction method for surface markers, based on validation of the signal distribution throughout a cell, mostly extinguished the detection of false positive events from both spatial spillover and unspecific staining artifacts and, therefore, highly increased the correlation between manual counts and automated quantification. Others have proposed methods for spatial spillover correction based on the spatial neighborhood of cells using adjacency matrices and correcting for the fraction of shared boundary between cells (Goltsev et al., 2018, Bai et al., 2021). However, these quantifications still lack the power to resolve positive and negative populations for each marker from imaging data. The quantification of staining intensities derived from ChipCytometry images can serve as a starting point for many different analyses, like the detection of phenotypically complex events by iterative gating, clustering of cells for unbiased cell type identification, and neighborhood analysis within a tissue sample.

Other imaging methods are specialized for higher throughput and increased multiplexing capacity, like the new PhenoCycler solution (Akoya Biosciences, former CODEX technology) or the GeoMx/CosMx system (Nanostring). However, the data quality of the ChipCytometry is still superior due to the special image acquisition (HDR, background subtraction), and it is fully compatible with third-party reagents as conventional immunofluorescent antibodies are used. A new instrument, the CellScape ChipCytometry device, was recently launched as a successor to the CellScanner used in this study. Besides the walk-away automation for up to four chips, this machine is equipped with a new optical core unit that allows higher resolution and up to 8x larger scan area per position for increased sample throughput. Since the general methodology has not changed, our handling protocol for FFPE samples is still applicable to this new generation platform, catching up on throughput with other methods. Furthermore, our work raised the company's interest in FFPE sample processing, and specific ChipCytometry kits for FFPE samples containing coated coverslips for better tissue adhesion are now on the market.

The developed pipeline for automated quantification can continuously be updated with new modules to be adapted to the current state-of-the-art image analysis. For example, we implemented the segmentation via a neuronal network approach and the correction of shading artifacts in the second version of the pipeline. The quantification of FISH signals is still in the testing phase, and there are additional ideas to be implemented, like the separate segmentation of the nucleus, cytoplasm and membrane to retrieve further information on the subcellular localization of the signal. This quantification of signal localization within the cell becomes particularly interesting in combination with the CellScape instrument and its higher resolution of 182 nm/pixel.

The technological achievements in multiplexed imaging led to the development of powerful image analysis tools, and ChipCytometry output data is primarily compatible with this software. Besides commercial software solutions like HALO (Indica labs), the Visiopharm software bundle, or Aperio ImageScope

(Leica), to name some of the commonly used ones, there are compelling open source tools like QuPath (Bankhead et al., 2017) as graphical user interface based analysis suite or the more advanced modular pipeline MCMICRO (Schapiro et al., 2022). Several AI-based image analysis algorithms are developed every month, and the combination of the best-performing models will ultimately lead to an optimal image analysis pipeline. For this reason, MCMICRO is an excellent framework for combining several toolsets developed by different groups, expanding on the basic principles that we proposed specifically for ChipCytometry data analysis.

The inter-sample variability is one significant challenge not resolved so far in image analysis. In our dataset, we overcame this problem by manually thresholding (gating) the marker intensity for each sample based on the marker distribution. However, pooled intensity data analysis would not be possible for multiple samples due to sample-specific differences in background intensity, signal intensity, or additional artifacts. Benchmarking and batch-correction, which has made substantial progress in the field of scRNA sequencing (Korsunsky et al., 2019, Haghverdi et al., 2018, Polanski et al., 2020), for different samples will be one of the main challenges in increasing applicability and automation in computer-assisted image analysis in the future.

5.2 A titrated protocol allows scRNA seq from gastrointestinal punch biopsies

Gastrointestinal punch biopsies are very limited in size (1-2 mm²), and we had to extract enough cells to perform transcriptomic phenotyping via scRNA sequencing. In addition to the limited size, immune reconstitution might not be completed depending on the time point after transplantation, yielding a further reduced density of immune cells within the biopsy as a specific challenge for aHSCT biopsies. We optimized cell extraction, cell suspension freezing, and immune cell enrichment and could extract sufficient numbers of cells to continue with scRNA sequencing experiments. The most important parameter here was keeping a high density of cells frozen without reaching too low volumes for sufficient cell recovery.

There have been advancements made in the meantime for the problems we were facing with our study: several patient samples collected on different days over a long period that should be analyzed in the same scRNA seq experiment. One solution for this problem could be the recently launched fixed-RNA scRNA sequencing kit (10X genomics) that allows to fix the sample on the day of the preparation and to generate scRNA sequencing libraries as soon as all samples are collected. This genome-wide, probe-based approach allows for whole transcriptome analysis, but since the original mRNA is never amplified, the CDR3 region, determining the specificity of immune cell receptors, cannot be sequenced with the technology. For this reason, the new approach is an excellent advancement for the field but would not have helped in our case for TCR profiling.

Before we performed scRNA sequencing with the aHSCT patient cohort, we checked for the correlation of gene expression between fresh and frozen samples. Besides a good correlation, in line with other studies suggesting DMSO cryopreservation for scRNA sequencing (Wohnhaas et al., 2019), we found a few differentially expressed genes associated with a stress response. Since these genes, however, were not specific for immune cell populations, they should not represent a significant problem as we processed all our samples the same way.

For sample pooling, we used SNP-based demultiplexing approaches which, in contrast to other methods published at this time like demuxlet (Kang et al., 2018), cluster cells solely on SNP differences, independent from the genotypes of the individual samples. After starting with scSplit (Xu et al., 2019), we moved on with Souporecell (Heaton et al., 2020), which in our hands showed a better capacity for sample demultiplexing. The better performance is most probably due to the modelling of ambient RNA in the sample as confounding factor for sample demultiplexing. Nowadays, most of these technologies have become less relevant due to the availability of cell tagging methods like cell hashing with barcoded antibodies (Stoeckius et al., 2018) or via lipid anchors (CellPlex technology, 10X genomics). However, they might still be interesting for limited sample material since additional staining steps, and corresponding sample loss, can be avoided.

5.3 GvHD is associated with increased immune and T cell infiltration

In our two datasets of immune infiltration from aHSCT biopsies, we found GvHD to be associated with increased immune infiltration, especially on the site of T cells, which have been previously described as cellular GvHD mediators (van den Brink and Burakoff, 2002). In addition, we detected some exciting trends of differential composition regarding other immune cell populations in the single-cell sequencing data. For example, the frequency of B cells and plasma cells was significantly reduced in aHSCT samples, which is somehow expected due to the slow reconstitution of these cell types in the tissue (Fiorenza and Turtle, 2021). Since current interest is raised on the role of B cells in aHSCT as their depletion showed a reduction in GvHD (Shimabukuro-Vornhagen et al., 2009), our data might serve as a starting point for further investigations, and B cell receptor data might be generated from the existing cDNA libraries. In addition to B cells, the frequency of mast cells was significantly increased in the control samples and sequentially decreased in mild and severe GvHD. This effect of reduced mast cell numbers in severe GvHD patients has been described before in combination with the immunoregulatory function of these cells (Ustun et al., 2020). Although our data contains many valuable information on several immune cell types, we mainly focused with this thesis work on the T cell compartment as it showed the most significant difference in both datasets. Nevertheless, our comprehensive dataset should become a starting point for many follow-up studies also deeper looking at other immune cells.

Besides some haploidentical transplantations, most of our patients were transplanted with a graft either fully HLA-matched (12/12 alleles, 32,1% of the patients) or with one mismatch in an HLA class II allele (DPB mismatch: 11/12 or 10/12 alleles, 39,3% of the patients). The enrichment of T cell frequencies, especially within the MHC class I-restricted CD8 T cell compartment, hint towards the recognition of minor mismatches in case of alloreactivity. Importantly, besides the enrichment of T cells among all immune cells, we could also observe an increased T cell density as an absolute measure of infiltration in the biopsies of severe GvHD patients, speaking for T cell recruitment to the side of inflammation. The enrichment of CD8 T cells translated into a decreasing CD4/CD8 ratio from controls to mild GvHD to severe GvHD. In response to such inflammation, we detected an increase in regulatory T cells, as was described before for FoxP3 expression in the tissue (Ghimire et al., 2021). In addition, we could further correlate the Treg with the CD8 T cell infiltration. Therefore, we could directly show for the first time that the counterintuitive higher Treg infiltration in severe GvHD is a response to the higher inflammation and conventional T cell infiltration. Furthermore, we showed that the increase in Treg frequencies is still not high enough to fully counterbalance the effects mediated by the abundance of conventional T cells.

Interestingly, GvHD was not the main driver of the observed differences in cell composition, besides it associated with an increased CD8 T cell infiltrate. In contrast, we found two other clinical parameters most likely explaining the main variance of the dataset. The time point of the biopsy relative to the transplantation seemed to explain the extent of immune infiltration in general, in line with immune reconstitution after transplantation in the tissue. Microbiome disruption, in contrast, explained variance in Treg density and phenotype within the tissue, which led us to investigate this correlation further.

5.4 The microbiome can influence Treg numbers and phenotype

After observing several indications regarding the interplay between Tregs and the microbiome, like the reduced frequency of Tregs and the lower proportion of activated CD45RO+ Tregs in antibiotic-treated patients, we investigated the regulatory T cell compartment in-depth with a specific focus on the interplay with microbiome disruption. By doing so, we found two transcriptomically distinct subsets of regulatory T cells, one of them showing a higher gene signature of effector T cell genes with a low expression of Treg effector markers. In contrast, the other subset showed a more stable Treg phenotype characterized by a high score of suppression-associated genes and a stronger Treg signature. By analyzing transcriptional dynamics, we observed a differentiation pathway starting from the cluster with the lower Treg signature and developing into the more suppressive Tregs upon induction. Altogether, the phenotypic signature of the clusters and the directionality of differentiation would refer to the generation of peripherally induced Tregs. Interestingly, and further supporting the hypothesis of pTregs induction, the differentiation was dependent on an intact microbiome, since we observed less matured Tregs in antibiotic-treated patients. pTregs have been previously described to play an essential role in mucosal tolerance induction, as pTregs were induced in the colon by local microbial antigens and TCR repertoires did not overlap with thymic-derived Tregs (Lathrop et al., 2011).

Further investigation of the involved species showed positive correlation between Treg differentiation and SCFA-producing bacteria from the *Clostridia* and *Bacteroidia* classes. SCFA-producing bacteria have been linked to Treg induction (Arpaia et al., 2013, Smith et al., 2013) and were described to play a beneficial role in the context of GvHD (Meedt et al., 2022). Surprisingly, the species most significantly associated with Treg differentiation belongs to the *Enterococcus* genus, which was earlier described as promoting GvHD (Stein-Thoeringer et al., 2019). *Enterococcus avium* is a vancomycin-resistant species, and human infections are rare, meaning that further investigation is needed to confirm its association with Treg induction. Its ability to produce SCFAs (Noronha et al., 2019) might, however, explain the observed connection with Treg differentiation and underline the complexity regarding opportunistic infection and positive influence on Treg induction. It also underscored the importance of analysis on the species level since *Enterococcus faecium*, the dominant species in the association studies with GvHD (Stein-Thoeringer et al., 2019), cannot produce the protective SCFAs.

The number of patients included in our microbiome dataset is limited, and further investigation in larger cohorts on the role of different bacterial species on GvHD is required to confirm our findings. Importantly, as an internal control, we found that the detected signatures from our data have, for the most part, been previously described in the context of anti-inflammatory effects in the gut, supporting the directionality of our data. From the Treg-supporting cluster 1, for example, *Streptococcus salivarius* (Laws et al., 2021), *Faecalibacterium prausnitzii* (Qiu et al., 2013), *Blautia wexlerae* (Benitez-Paez et al., 2020), and *Bacteroides ovatus* (Tan et al., 2019) have been described to reduce inflammation and to induce tolerance or regulatory mechanisms.

5.5 Antigen-specific, systemically present CD8 T cells are linked to GvHD

T cells in severe GvHD patients showed an activated phenotype, and the upregulated genes were further associated with GvHD via GSEA. The connection to GvHD-related genes was mainly driven by CD8 T cells and less by CD4. The pathway gene set from the Kyoto Encyclopedia of genes and genomes (KEGG, entry hsa05332) contains curated genes from literature associated with GvHD pathophysiology. These genes are related to the three steps of GvHD development (see 1.3): IL1, IL6, and TNF are related to the generation of an inflammatory environment due to tissue damage (step 1). MHC class I and class II molecules, as well as other APC activation markers, are associated with antigen presentation to adaptive immune cells (step 2). Cytotoxicity markers like GZMB, PRF1, and T cell activation markers like IL2 or CD28 are linked to the tissue damage of target organs (step 3). The correlation between Reg3a levels in the circulation of patients and the frequency of CD8 activated T cells created an additional link to GvHD pathophysiology. Reg3a is a marker for tissue damage in GvHD (Ferrara et al., 2011), suggesting that CD8 T cells directly mediate the target organ damage by their cytotoxic function.

The analysis of the TCR repertoires revealed that the observed enrichment of CD8 T cells is associated with clonal expansion. The expanded clonotypes mostly showed an activated CD8 T cell signature, defined by the expression of granzymes, perforin, IFNG, and CD27, and low expression of ITGAE as a marker for tissue residency. Some patients in parallel showed a clonal expansion in CD4 effector memory T cells. The observation of increasing oligoclonality with GvHD severity support the hypothesis that the observed GvHD might be the result of an antigen-specific process rather than an unspecific activation of T cells.

It has been known for decades that T cells play an essential role and are required for GvHD development. However, the oligoclonality of these T cells has been investigated only partially in mouse models so far (Zheng et al., 2020, Wu et al., 2021). These studies found trends of expansion in the allogeneic transplantation setting, without phenotypic information on the T cells. Our findings confirmed previous data on clonal expansion, but shed, with a previously unseen depth of information, new light on the field of TCR repertoires in human GvHD patients. The observation that the clonotypes are stable over time in longitudinal samples and systemically distributed within the gastrointestinal tract, as indicated by matched samples from different anatomic regions, further corroborated the importance of antigen-specificity in the setting of GvHD. In a patient with four longitudinal samples between day 21 and 238 after stem cell transplantation, we could further observe an increase in clonal expansion over time with a more polyclonal repertoire at the early timepoints and increasing oligoclonality and shared clonotypes at the later time points. Overall, the data revealed that a restricted set of antigens seems to be responsible for the disease development. The removal of either antigen or specifically expanded T cells might be considered a therapeutic approach in the future.

One crucial step of future investigation will be the identification of the target epitopes recognized by the T cell receptors of the expanded T cells. Minor histocompatibility antigens have been described in the setting of GvHD, and TCRs were identified specific against the HA-1 and HA-2 peptides derived by point mutations of the HMHA1 gene (Spierings et al., 2009, Nicholls et al., 2009, den Haan et al., 1998), ACC-1 and ACC-2 peptides, derived from the BCL2A1 gene (Akatsuka et al., 2003), and LRH1 peptide generated by a frameshift mutation in the P2X5 gene (de Rijke et al., 2005). Bacteria play an essential role in modulating GvHD, and the question remains if, in addition to allogenic epitopes, bacterial epitopes could play a role in the observed CD8 T cell activation and clonal expansion.

To test this hypothesis and to identify potential antigen targets, we re-expressed the TCRs identified during this work in reporter cell lines for antigen screening techniques. In principle, there are two basic approaches for TCR epitope screening: screening of a fully randomized peptide library or working with a known peptidome, for example based on the human proteome. The entirely random screening is unbiased and not species-specific, with the drawback of an overwhelming number of targets to be tested. For a nonamer peptide pool, $\sim 8 \times 10^{11}$ possible peptides need to be considered, making an unbiased screening almost unrealistic. Therefore, we are working on a semi-randomized peptide library approach, where all possible peptide combinations for a given length are computationally restricted to a given HLA molecule, calculating probabilities of anchor residues from peptide databases (in collaboration with Kilian Schober and Benjamin Schubert). This approach would allow us to check the presence of candidate peptides in the human and bacterial genomes to further specify the source of target antigens for GvHD.

5.6 Treg functionality and infiltration density are essential for GvHD suppression

We identified and in-depth characterized two main mechanisms involved in the development of GvHD in aHSCT patients. These are the clonal expansion of CD8 T cells, leading to organ destruction and thereby driving gastrointestinal GvHD, and the dependence of Treg differentiation on an intact microbiome. We could finally show that these two findings are connected since the Treg differentiation status was inversely correlated with CD8 T cell expansion in patients. This connection, together with our findings regarding the interplay between microbiome and Tregs, strongly supports the idea of microbiome intervention for GvHD treatment or even prophylaxis. After some more experimental FMT approaches, there is increasing evidence that this type of microbiome intervention can dramatically increase the survival rates of severe GvHD patients (Zhao et al., 2021, Goeser et al., 2021).

We were able to analyze a first patient who developed steroid-refractory grade IV GvHD at day 290 after transplantation and was treated with FMT. For GvHD therapy, the patient received two FMT treatments (on days 306 and 314) from the same screened donor. This resulted in a restoration of microbial diversity and community towards the donor's microbiome, accompanied by an increase in protective metabolites like SCFAs (Orberg et al., 2022). ChipCytometry analysis of biopsies before and after FMT revealed regeneration of crypt structure, increased infiltration of regulatory T cells, and decrease in CD8 T cell density after FMT, again underlining the importance of the crosstalk between the microbiome and immune cells (Orberg et al., 2022). We could further observe a decreased distance between CD8 and Treg cells in the patient after FMT, potentially allowing spatial interaction and Treg-mediated suppression of inflammation. The patient remained GvHD-free until the last follow-up nine months after FMT. Despite being a single case, this evidence again demonstrates the curative potential of FMT for patients with otherwise very high mortality rates, reported to be greater than 90% (Westin et al., 2011).

Regarding GvHD prophylaxis, FMT might not be the best option for immunocompromised aHSCT patients due to the high risk of bacteremia or sepsis development. Since our data, as well as other studies (Arpaia et al., 2013, Smith et al., 2013), strongly support SCFAs as the mode of action for Treg induction by bacteria, the treatment with SCFAs might be a good alternative for a safer microbiome-related intervention in case of GvHD prophylaxis. Furthermore, butyrate can improve intestinal barrier integrity (Ma et al., 2012), which would be an intervention in the early development of GvHD. The treatment of inflammatory bowel diseases (IBD), which has been connected to SCFAs as protective metabolites as well, was unfortunately only effective in a subset of patients in the first clinical trials (Breuer et al., 1997). Pre- and probiotic approaches to increase the abundance of SCFA bacteria showed diverse clinical outcomes (Parada Venegas et al., 2019), and more systematic trials need to be performed to understand the reasons for treatment failure. However, although not effective in all patients, there are encouraging results of remission cases in IBD, and further mechanistic understanding might help to improve therapies in the future. For GvHD, preclinical studies showed GvHD protection through treatment with either butyrate or SCFA-producing bacterial species (Mathewson et al., 2016). Clinical translation is still pending, but there is increasing evidence that combinations of the elaborated treatment options might help to increase survival rates after aHSCT in the future.

6. Summary

GvHD is a significant threat after aHSCT, and a detailed understanding of the underlying pathogenic mechanisms is crucial for developing new therapies to treat the disease and thereby improve the potential of aHSCT for treating hematological malignancies. Several factors influencing disease severity have been known for a while, like the microbiome diversity and preconditioning intensity. However, these observations did mainly come from preclinical models. Direct mechanistic connections in human GvHD patients are lacking due to the often-low availability of human sample material and missing methods for in-depth immune cell characterization. This study aimed to systematically analyze the microbiome, clinical status, and immune infiltrates from a cohort of human aHSCT patients.

According to the repository of gastrointestinal biopsies from our patient cohort, we adapted and developed methods for multiplexed imaging and single-cell RNA sequencing of gastrointestinal immune cells. We developed an optimized cell extraction method to apply scRNA seq on the immune cell infiltrates of gastrointestinal punch biopsies and transferred the ChipCytometry technology for multiplexed imaging to FFPE samples. We complemented ChipCytometry as an HMTI method with an analysis pipeline for the automated quantification of imaging data to facilitate the investigation and phenotyping of immune populations within tissue samples.

We created a multi-omics scRNA sequencing dataset from 31 aHSCT biopsies containing transcriptomic, proteomic, and T cell receptor information. ChipCytometry imaging data from 52 biopsies containing spatial and phenotypic information on T cell infiltrates completed our dataset. In combination with the microbiome and clinical data, we found that GvHD is associated with an increased infiltration and clonal expansion of CD8 T cells, which were stable over time and systemically distributed throughout the gastrointestinal tract. Regulatory T cells are essential regulators for a balance between tolerance and protection from infections and have been described as critical players in suppressing GvHD. Regarding this regulatory pillar, we found a strong dependence on an intact microbiome, and particularly on SCFA-producing bacterial species, for the induction and differentiation of regulatory T cells. The differentiation status of Tregs was inversely correlated with the clonal expansion of CD8 T cells, pronouncing the suppressive role of Tregs and their ability to diminish GvHD progression.

Our data underline the importance of an intact microbiome in reducing GvHD severity. Recent clinical trials have shown that FMT can be curative for GvHD patients. We detected an increased Treg infiltration after FMT, associated with a decreased infiltration of CD8 T cells and a reduced distance between Tregs and CD8 T cells, allowing direct interaction. Besides shedding new light on the pathophysiology of GvHD, our data strongly support microbiome interventions for GvHD treatment. The strong phenotype of clonal expansion might be another target for therapies, and identifying the recognized antigens will be an essential follow-up of the project.

7. Acknowledgements

Many people were involved in the project and completing this thesis would not have been possible without the enormous support I had faced during the last years. First, I would like to thank **Dirk Busch** for supervising the project. His enthusiasm and scientific input were always important to me, but I'm also very grateful that I had the opportunity to influence the direction of the project myself. I also appreciate the supervision of my mentor **Elvira D'Ippolito**, who always supported me during the time. I wish to thank my second supervisor **Dirk Haller** for the scientific input. Besides the committee, I'm thankful for the additional support I received from **Simon Graßmann**, **Kilian Schober**, and **Thomas Müller**.

This project is a tandem project within the CRC1371 "Microbiome Signatures" with the group of **Ernst Holler** at the University of Regensburg. The group's clinical expertise was essential for the project's success, and I'm thankful for the input. This also involved **Elisabeth Meedt**, **Sakhila Ghimire**, **Doris Gaag**, and **Andreas Hiergeist**. The integration in the CRC provided an ideal environment to discuss the project with experts in the field. It allowed learning the required methodology and theory needed for the experimental setup. Many aspects of the ChipCytometry method have been developed in a team-work with **Jan Köhlen**, and I'm very thankful for the time we pushed the project together. I also always enjoyed working with **Sabrina Wagner**, who joined the team later, as well as discussing future directions with the people from Zellkraftwerk. The establishment of human tissue imaging approaches always needs human sample material. I'm thankful for the samples provided from the repositories of **Ernst Holler**, **Katja Steiger**, and **Klaus-Peter Janssen**, as well as for their expertise and tips in imaging technologies.

Besides the main project, I got involved in several collaborations, and I would like to thank **Anna Purcarea**, **Karolin Wagner**, and **Laura Mateyka** for being such great teammates. We had a fantastic time during collaborative projects, which were successfully published as a team effort. All single-cell RNA sequencing experiments have been performed together with **Monika Hammel**, and I thank her for going with me through more than 40 experiments and for always supporting my work in every way imaginable. In addition to the people from the institute named above, I would like to thank my other colleagues and friends **Sarah Dötsch**, **Justin Leube**, **Immanuel Andrä**, **Linda Warmuth**, **Adrian Straub**, **Julia Ritter**, **Inge Hensel**, **Anna Hochholzer**, **Anton Mühlbauer**, **Ludwig Pachmayr**, **Lorenz Kretschmer**, **Veronika Engelsberger**, **Saskia Kolb**, and **Anna Sichler** for sharing many hours of work, fun, and discussions and for making the time at the institute unforgettable. Here, I would especially like to highlight my gratefulness to **Anna Ralser** for our stimulating scientific discussions. But also, for being such a wonderful friend, and for numerous joint activities from coffee breaks and "pizza & wine" evenings to exciting tennis matches and other shared adventures I could always draw some motivation from.

And last but not least, I am very grateful for the support from my family. I thank my parents, **Uschi** and **Stephan Jarosch**, and my sister **Valentina Jarosch** for their enormous emotional support. The same is true for my girlfriend **Sophie Flommersfeld**, whom I'm very thankful for being my rock, supporting me in every situation and for re-motivating me after every failure.

8. Bibliography

Akatsuka, Y., Nishida, T., Kondo, E., Miyazaki, M., Taji, H., Iida, H., Tsujimura, K., Yazaki, M., Naoe, T., Morishima, Y., Kodera, Y., Kuzushima, K. & Takahashi, T. 2003. Identification of a polymorphic gene, BCL2A1, encoding two novel hematopoietic lineage-specific minor histocompatibility antigens. *J Exp Med*, 197, 1489-500.

Ameyar, M., Wisniewska, M. & Weitzman, J. B. 2003. A role for AP-1 in apoptosis: the case for and against. *Biochimie*, 85, 747-52.

Angell, H. K., Bruni, D., Barrett, J. C., Herbst, R. & Galon, J. 2020. The Immunoscore: Colon Cancer and Beyond. *Clin Cancer Res*, 26, 332-339.

Angelo, M., Bendall, S. C., Finck, R., Hale, M. B., Hitzman, C., Borowsky, A. D., Levenson, R. M., Lowe, J. B., Liu, S. D., Zhao, S., Natkunam, Y. & Nolan, G. P. 2014. Multiplexed ion beam imaging of human breast tumors. *Nat Med*, 20, 436-42.

Angermueller, C., Clark, S. J., Lee, H. J., Macaulay, I. C., Teng, M. J., Hu, T. X., Krueger, F., Smallwood, S., Ponting, C. P., Voet, T., Kelsey, G., Stegle, O. & Reik, W. 2016. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat Methods*, 13, 229-232.

Arpaia, N., Campbell, C., Fan, X., Dikiy, S., Van Der Veecken, J., Deroos, P., Liu, H., Cross, J. R., Pfeffer, K., Coffey, P. J. & Rudensky, A. Y. 2013. Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. *Nature*, 504, 451-5.

Atarashi, K., Tanoue, T., Shima, T., Imaoka, A., Kuwahara, T., Momose, Y., Cheng, G., Yamasaki, S., Saito, T., Ohba, Y., Taniguchi, T., Takeda, K., Hori, S., Ivanov, I., Umesaki, Y., Itoh, K. & Honda, K. 2011. Induction of colonic regulatory T cells by indigenous *Clostridium* species. *Science*, 331, 337-41.

Azad, S., Malhotra, V., Kirtani, P. & Choudhary, D. 2019. Gastrointestinal Biopsies for Evaluation of Acute Graft-Versus-Host Disease in Allogeneic Hematopoietic Stem Cell Transplant Patients. *Indian J Hematol Blood Transfus*, 35, 83-88.

Bai, Y., Zhu, B., Rovira-Clave, X., Chen, H., Markovic, M., Chan, C. N., Su, T. H., McIlwain, D. R., Estes, J. D., Keren, L., Nolan, G. P. & Jiang, S. 2021. Adjacent Cell Marker Lateral Spillover Compensation and Reinforcement for Multiplexed Images. *Front Immunol*, 12, 652631.

Bankhead, P., Loughrey, M. B., Fernandez, J. A., Dombrowski, Y., McArt, D. G., Dunne, P. D., McQuaid, S., Gray, R. T., Murray, L. J., Coleman, H. G., James, J. A., Salto-Tellez, M. & Hamilton, P. W. 2017. QuPath: Open source software for digital pathology image analysis. *Sci Rep*, 7, 16878.

Becht, E., McInnes, L., Healy, J., Dutertre, C. A., Kwok, I. W. H., Ng, L. G., Ginhoux, F. & Newell, E. W. 2018. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol*.

Beelen, D. W., Haralambie, E., Brandt, H., Linzenmeier, G., Muller, K. D., Quabeck, K., Sayer, H. G., Graeven, U., Mahmoud, H. K. & Schaefer, U. W. 1992. Evidence that sustained growth suppression of intestinal anaerobic bacteria reduces the risk of acute graft-versus-host disease after sibling marrow transplantation. *Blood*, 80, 2668-76.

Benitez-Paez, A., Gomez Del Pugar, E. M., Lopez-Almela, I., Moya-Perez, A., Codoner-Franch, P. & Sanz, Y. 2020. Depletion of *Blautia* Species in the Microbiota of Obese Children Relates to Intestinal Inflammation and Metabolic Phenotype Worsening. *mSystems*, 5.

Bibliography

- Beres, A., Komorowski, R., Mihara, M. & Drobyski, W. R. 2011. Instability of Foxp3 expression limits the ability of induced regulatory T cells to mitigate graft versus host disease. *Clin Cancer Res*, 17, 3969-83.
- Bergen, V., Lange, M., Peidli, S., Wolf, F. A. & Theis, F. J. 2020. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol*, 38, 1408-1414.
- Bernstein, N. J., Fong, N. L., Lam, I., Roy, M. A., Hendrickson, D. G. & Kelley, D. R. 2020. Solo: Doublet Identification in Single-Cell RNA-Seq via Semi-Supervised Deep Learning. *Cell Syst*, 11, 95-101 e5.
- Biagi, E., Zama, D., Nastasi, C., Consolandi, C., Fiori, J., Rampelli, S., Turrone, S., Centanni, M., Severgnini, M., Peano, C., De Bellis, G., Basaglia, G., Gotti, R., Masetti, R., Pession, A., Brigidi, P. & Candela, M. 2015. Gut microbiota trajectory in pediatric patients undergoing hematopoietic SCT. *Bone Marrow Transplant*, 50, 992-8.
- Billingham, R. E. 1966. The biology of graft-versus-host reactions. *Harvey Lect*, 62, 21-78.
- Bolanos-Meade, J. & Brodsky, R. A. 2009. Blood and marrow transplantation for sickle cell disease: overcoming barriers to success. *Curr Opin Oncol*, 21, 158-61.
- Breuer, R. I., Soergel, K. H., Lashner, B. A., Christ, M. L., Hanauer, S. B., Vanagunas, A., Harig, J. M., Keshavarzian, A., Robinson, M., Sellin, J. H., Weinberg, D., Vidican, D. E., Flemal, K. L. & Rademaker, A. W. 1997. Short chain fatty acid rectal irrigation for left-sided ulcerative colitis: a randomised, placebo controlled trial. *Gut*, 40, 485-91.
- Cahn, J. Y., Klein, J. P., Lee, S. J., Milpied, N., Blaise, D., Antin, J. H., Leblond, V., Ifrah, N., Jouet, J. P., Loberiza, F., Ringden, O., Barrett, A. J., Horowitz, M. M., Socie, G., Societe Francaise De Greffe De Moelle Et Therapie, C., Dana Farber Cancer, I. & International Bone Marrow Transplant, R. 2005. Prospective evaluation of 2 acute graft-versus-host (GVHD) grading systems: a joint Societe Francaise de Greffe de Moelle et Therapie Cellulaire (SFGM-TC), Dana Farber Cancer Institute (DFCI), and International Bone Marrow Transplant Registry (IBMTR) prospective study. *Blood*, 106, 1495-500.
- Callahan, B. J., Sankaran, K., Fukuyama, J. A., Mcmurdie, P. J. & Holmes, S. P. 2016. Bioconductor Workflow for Microbiome Data Analysis: from raw reads to community analyses. *F1000Res*, 5, 1492.
- Cao, J., Cusanovich, D. A., Ramani, V., Aghamirzaie, D., Pliner, H. A., Hill, A. J., Daza, R. M., Mcfaline-Figueroa, J. L., Packer, J. S., Christiansen, L., Steemers, F. J., Adey, A. C., Trapnell, C. & Shendure, J. 2018. Joint profiling of chromatin accessibility and gene expression in thousands of single cells. *Science*, 361, 1380-1385.
- Chen, A., Liao, S., Cheng, M., Ma, K., Wu, L., Lai, Y., Qiu, X., Yang, J., Xu, J., Hao, S., Wang, X., Lu, H., Chen, X., Liu, X., Huang, X., Li, Z., Hong, Y., Jiang, Y., Peng, J., Liu, S., Shen, M., Liu, C., Li, Q., Yuan, Y., Wei, X., Zheng, H., Feng, W., Wang, Z., Liu, Y., Wang, Z., Yang, Y., Xiang, H., Han, L., Qin, B., Guo, P., Lai, G., Munoz-Canoves, P., Maxwell, P. H., Thiery, J. P., Wu, Q. F., Zhao, F., Chen, B., Li, M., Dai, X., Wang, S., Kuang, H., Hui, J., Wang, L., Fei, J. F., Wang, O., Wei, X., Lu, H., Wang, B., Liu, S., Gu, Y., Ni, M., Zhang, W., Mu, F., Yin, Y., Yang, H., Lisby, M., Cornall, R. J., Mulder, J., Uhlen, M., Esteban, M. A., Li, Y., Liu, L., Xu, X. & Wang, J. 2022. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell*, 185, 1777-1792 e21.
- Chinen, T., Kannan, A. K., Levine, A. G., Fan, X., Klein, U., Zheng, Y., Gasteiger, G., Feng, Y., Fontenot, J. D. & Rudensky, A. Y. 2016. An essential role for the IL-2 receptor in Treg cell function. *Nat Immunol*, 17, 1322-1333.

Bibliography

Choi, S. W., Kitko, C. L., Braun, T., Paczesny, S., Yanik, G., Mineishi, S., Krijanovski, O., Jones, D., Whitfield, J., Cooke, K., Hutchinson, R. J., Ferrara, J. L. & Levine, J. E. 2008. Change in plasma tumor necrosis factor receptor 1 levels in the first week after myeloablative allogeneic transplantation correlates with severity and incidence of GVHD and survival. *Blood*, 112, 1539-42.

Cohen, J. L. & Boyer, O. 2006. The role of CD4+CD25hi regulatory T cells in the physiopathogeny of graft-versus-host disease. *Curr Opin Immunol*, 18, 580-5.

Cohen, J. L., Trenado, A., Vasey, D., Klatzmann, D. & Salomon, B. L. 2002. CD4(+)CD25(+) immunoregulatory T Cells: new therapeutics for graft-versus-host disease. *J Exp Med*, 196, 401-6.

Cooke, K. R., Gerbitz, A., Crawford, J. M., Teshima, T., Hill, G. R., Tesolin, A., Rossignol, D. P. & Ferrara, J. L. 2001. LPS antagonism reduces graft-versus-host disease and preserves graft-versus-leukemia activity after experimental bone marrow transplantation. *J Clin Invest*, 107, 1581-9.

Csencsits, K. L. & Bishop, D. K. 2003. Contrasting alloreactive CD4+ and CD8+ T cells: there's more to it than MHC restriction. *Am J Transplant*, 3, 107-15.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., Depristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., Mcvean, G., Durbin, R. & Genomes Project Analysis, G. 2011. The variant call format and VCFtools. *Bioinformatics*, 27, 2156-8.

Davis, A. S., Richter, A., Becker, S., Moyer, J. E., Sandouk, A., Skinner, J. & Taubenberger, J. K. 2014. Characterizing and Diminishing Autofluorescence in Formalin-fixed Paraffin-embedded Human Respiratory Tissue. *J Histochem Cytochem*, 62, 405-423.

De Rijke, B., Van Horssen-Zoetbrood, A., Beekman, J. M., Otterud, B., Maas, F., Woestenenk, R., Kester, M., Leppert, M., Schattenberg, A. V., De Witte, T., Van De Wiel-Van Kemenade, E. & Dolstra, H. 2005. A frameshift polymorphism in P2X5 elicits an allogeneic cytotoxic T lymphocyte response associated with remission of chronic myeloid leukemia. *J Clin Invest*, 115, 3506-16.

Defilipp, Z., Peled, J. U., Li, S., Mahabamunuge, J., Dagher, Z., Slingerland, A. E., Del Rio, C., Valles, B., Kempner, M. E., Smith, M., Brown, J., Dey, B. R., El-Jawahri, A., McAfee, S. L., Spitzer, T. R., Ballen, K. K., Sung, A. D., Dalton, T. E., Messina, J. A., Dettmer, K., Liebisch, G., Oefner, P., Taur, Y., Pamer, E. G., Holler, E., Mansour, M. K., Van Den Brink, M. R. M., Hohmann, E., Jenq, R. R. & Chen, Y. B. 2018. Third-party fecal microbiota transplantation following allo-HCT reconstitutes microbiome diversity. *Blood Adv*, 2, 745-753.

Den Haan, J. M., Meadows, L. M., Wang, W., Pool, J., Blokland, E., Bishop, T. L., Reinhardus, C., Shabanowitz, J., Offringa, R., Hunt, D. F., Engelhard, V. H. & Goulmy, E. 1998. The minor histocompatibility antigen HA-1: a diallelic gene with a single amino acid polymorphism. *Science*, 279, 1054-7.

Den Haan, J. M., Sherman, N. E., Blokland, E., Huczko, E., Koning, F., Drijfhout, J. W., Skipper, J., Shabanowitz, J., Hunt, D. F., Engelhard, V. H. & Et Al. 1995. Identification of a graft versus host disease-associated human minor histocompatibility antigen. *Science*, 268, 1476-80.

Dvorak, C. C. & Cowan, M. J. 2008. Hematopoietic stem cell transplantation for primary immunodeficiency disease. *Bone Marrow Transplant*, 41, 119-26.

Edinger, M., Hoffmann, P., Ermann, J., Drago, K., Fathman, C. G., Strober, S. & Negrin, R. S. 2003. CD4+CD25+ regulatory T cells preserve graft-versus-tumor activity while inhibiting graft-versus-host disease after bone marrow transplantation. *Nat Med*, 9, 1144-50.

Bibliography

- Epstein, R. J., McDonald, G. B., Sale, G. E., Shulman, H. M. & Thomas, E. D. 1980. The diagnostic accuracy of the rectal biopsy in acute graft-versus-host disease: a prospective study of thirteen patients. *Gastroenterology*, 78, 764-71.
- Fan, J., Lee, H. O., Lee, S., Ryu, D. E., Lee, S., Xue, C., Kim, S. J., Kim, K., Barkas, N., Park, P. J., Park, W. Y. & Kharchenko, P. V. 2018. Linking transcriptional and genetic tumor heterogeneity through allele analysis of single-cell RNA-seq data. *Genome Res*, 28, 1217-1227.
- Ferrara, J. L. & Deeg, H. J. 1991. Graft-versus-host disease. *N Engl J Med*, 324, 667-74.
- Ferrara, J. L., Harris, A. C., Greenson, J. K., Braun, T. M., Holler, E., Teshima, T., Levine, J. E., Choi, S. W., Huber, E., Landfried, K., Akashi, K., Vander Lugt, M., Reddy, P., Chin, A., Zhang, Q., Hanash, S. & Paczesny, S. 2011. Regenerating islet-derived 3-alpha is a biomarker of gastrointestinal graft-versus-host disease. *Blood*, 118, 6702-8.
- Ferrara, J. L., Levine, J. E., Reddy, P. & Holler, E. 2009. Graft-versus-host disease. *Lancet*, 373, 1550-61.
- Fiorenza, S. & Turtle, C. J. 2021. Associations between the Gut Microbiota, Immune Reconstitution, and Outcomes of Allogeneic Hematopoietic Stem Cell Transplantation. *Immunometabolism*, 3.
- Fontenot, J. D., Gavin, M. A. & Rudensky, A. Y. 2003. Foxp3 programs the development and function of CD4+CD25+ regulatory T cells. *Nat Immunol*, 4, 330-6.
- Fujii, N., Takenaka, K., Shinagawa, K., Ikeda, K., Maeda, Y., Sunami, K., Hiramatsu, Y., Matsuo, K., Ishimaru, F., Niiya, K., Yoshino, T., Hirabayashi, N. & Harada, M. 2001. Hepatic graft-versus-host disease presenting as an acute hepatitis after allogeneic peripheral blood stem cell transplantation. *Bone Marrow Transplant*, 27, 1007-10.
- Fujioka, T., Tamaki, H., Ikegame, K., Yoshihara, S., Taniguchi, K., Kaida, K., Kato, R., Inoue, T., Nakata, J., Ishii, S., Soma, T., Okada, M. & Ogawa, H. 2013. Frequency of CD4(+)FOXP3(+) regulatory T-cells at early stages after HLA-mismatched allogeneic hematopoietic SCT predicts the incidence of acute GVHD. *Bone Marrow Transplant*, 48, 859-64.
- Gaffney, E. F., Riegman, P. H., Grizzle, W. E. & Watson, P. H. 2018. Factors that drive the increasing use of FFPE tissue in basic and translational cancer research. *Biotech Histochem*, 93, 373-386.
- Garrison, E. & Marth, G. 2012. Haplotype-based variant detection from short-read sequencing.
- Gerdes, M. J., Sevinsky, C. J., Sood, A., Adak, S., Bello, M. O., Bordwell, A., Can, A., Corwin, A., Dinn, S., Filkins, R. J., Hollman, D., Kamath, V., Kaanumalle, S., Kenny, K., Larsen, M., Lazare, M., Li, Q., Lowes, C., McCulloch, C. C., McDonough, E., Montalto, M. C., Pang, Z., Rittscher, J., Santamaria-Pang, A., Sarachan, B. D., Seel, M. L., Seppo, A., Shaikh, K., Sui, Y., Zhang, J. & Ginty, F. 2013. Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. *Proc Natl Acad Sci U S A*, 110, 11982-7.
- Ghimire, S., Weber, D., Hippe, K., Meedt, E., Hoepting, M., Kattner, A. S., Hiergeist, A., Gessner, A., Matos, C., Ghimire, S., Wolff, D., Edinger, M., Hoffmann, P., Poeck, H., Herr, W. & Holler, E. 2021. GPR Expression in Intestinal Biopsies From SCT Patients Is Upregulated in GvHD and Is Suppressed by Broad-Spectrum Antibiotics. *Front Immunol*, 12, 753287.
- Giesen, C., Wang, H. A., Schapiro, D., Zivanovic, N., Jacobs, A., Hattendorf, B., Schuffler, P. J., Grolimund, D., Buhmann, J. M., Brandt, S., Varga, Z., Wild, P. J., Gunther, D. & Bodenmiller, B. 2014.

Bibliography

Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat Methods*, 11, 417-22.

Gil-Araujo, B., Toledo Lobo, M. V., Gutierrez-Salmeron, M., Gutierrez-Pitalua, J., Ropero, S., Angulo, J. C., Chiloeches, A. & Lasa, M. 2014. Dual specificity phosphatase 1 expression inversely correlates with NF-kappaB activity and expression in prostate cancer and promotes apoptosis through a p38 MAPK dependent mechanism. *Mol Oncol*, 8, 27-38.

Giralt, S. & Bishop, M. R. 2009. Principles and overview of allogeneic hematopoietic stem cell transplantation. *Cancer Treat Res*, 144, 1-21.

Glucksberg, H., Storb, R., Fefer, A., Buckner, C. D., Neiman, P. E., Clift, R. A., Lerner, K. G. & Thomas, E. D. 1974. Clinical manifestations of graft-versus-host disease in human recipients of marrow from HL-A-matched sibling donors. *Transplantation*, 18, 295-304.

Goeser, F., Sift, B., Stein-Thoeringer, C., Farowski, F., Strassburg, C. P., Brossart, P., Higgins, P. G., Scheid, C., Wolf, D., Holderried, T. a. W., Vehreschild, M. & Cruz Aguilar, M. R. 2021. Fecal microbiota transfer for refractory intestinal graft-versus-host disease - Experience from two German tertiary centers. *Eur J Haematol*, 107, 229-245.

Goltsev, Y., Samusik, N., Kennedy-Darling, J., Bhate, S., Hale, M., Vazquez, G., Black, S. & Nolan, G. P. 2018. Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell*, 174, 968-981 e15.

Greinix, H. T., Knobler, R. M., Worel, N., Schneider, B., Schneeberger, A., Hoecker, P., Mitterbauer, M., Rabitsch, W., Schulenburg, A. & Kalhs, P. 2006. The effect of intensified extracorporeal photochemotherapy on long-term survival in patients with severe acute graft-versus-host disease. *Haematologica*, 91, 405-8.

Grewal, S. S., Barker, J. N., Davies, S. M. & Wagner, J. E. 2003. Unrelated donor hematopoietic cell transplantation: marrow or umbilical cord blood? *Blood*, 101, 4233-44.

Haghverdi, L., Lun, A. T. L., Morgan, M. D. & Marioni, J. C. 2018. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat Biotechnol*, 36, 421-427.

Han, L., Jin, H., Zhou, L., Zhang, X., Fan, Z., Dai, M., Lin, Q., Huang, F., Xuan, L., Zhang, H. & Liu, Q. 2018. Intestinal Microbiota at Engraftment Influence Acute Graft-Versus-Host Disease via the Treg/Th17 Balance in Allo-HSCT Recipients. *Front Immunol*, 9, 669.

Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Van Kerkwijk, M. H., Brett, M., Haldane, A., Del Rio, J. F., Wiebe, M., Peterson, P., Gerard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C. & Oliphant, T. E. 2020. Array programming with NumPy. *Nature*, 585, 357-362.

Hayatsu, N., Miyao, T., Tachibana, M., Murakami, R., Kimura, A., Kato, T., Kawakami, E., Endo, T. A., Setoguchi, R., Watarai, H., Nishikawa, T., Yasuda, T., Yoshida, H. & Hori, S. 2017. Analyses of a Mutant Foxp3 Allele Reveal BATF as a Critical Transcription Factor in the Differentiation and Accumulation of Tissue Regulatory T Cells. *Immunity*, 47, 268-283 e9.

Heaton, H., Talman, A. M., Knights, A., Imaz, M., Gaffney, D. J., Durbin, R., Hemberg, M. & Lawnczak, M. K. N. 2020. Souporecell: robust clustering of single-cell RNA-seq data by genotype without reference genotypes. *Nat Methods*, 17, 615-620.

Bibliography

Hennig, C., Adams, N. & Hansen, G. 2009. A versatile platform for comprehensive chip-based explorative cytometry. *Cytometry A*, 75, 362-70.

Hickey, J. W., Neumann, E. K., Radtke, A. J., Camarillo, J. M., Beuschel, R. T., Albanese, A., Mcdonough, E., Hatler, J., Wiblin, A. E., Fisher, J., Croteau, J., Small, E. C., Sood, A., Caprioli, R. M., Angelo, R. M., Nolan, G. P., Chung, K., Hewitt, S. M., Germain, R. N., Spraggins, J. M., Lundberg, E., Snyder, M. P., Kelleher, N. L. & Saka, S. K. 2022. Spatial mapping of protein composition and tissue organization: a primer for multiplexed antibody-based imaging. *Nat Methods*, 19, 284-295.

Hill, G. R., Crawford, J. M., Cooke, K. R., Brinson, Y. S., Pan, L. & Ferrara, J. L. 1997. Total body irradiation and acute graft-versus-host disease: the role of gastrointestinal damage and inflammatory cytokines. *Blood*, 90, 3204-13.

Hill, G. R. & Ferrara, J. L. 2000. The primacy of the gastrointestinal tract as a target organ of acute graft-versus-host disease: rationale for the use of cytokine shields in allogeneic bone marrow transplantation. *Blood*, 95, 2754-9.

Hoerning, A., Koss, K., Datta, D., Boneschanski, L., Jones, C. N., Wong, I. Y., Irimia, D., Calzadilla, K., Benitez, F., Hoyer, P. F., Harmon, W. E. & Briscoe, D. M. 2011. Subsets of human CD4(+) regulatory T cells express the peripheral homing receptor CXCR3. *Eur J Immunol*, 41, 2291-302.

Hoffmann, P., Ermann, J., Edinger, M., Fathman, C. G. & Strober, S. 2002. Donor-type CD4(+)CD25(+) regulatory T cells suppress lethal acute graft-versus-host disease after allogeneic bone marrow transplantation. *J Exp Med*, 196, 389-99.

Höllbacher, B., Duhon, T., Motley, S., Klicznik, M. M., Gratz, I. K. & Campbell, D. J. 2020. Transcriptomic Profiling of Human Effector and Regulatory T Cell Subsets Identifies Predictive Population Signatures. *ImmunoHorizons*, 4, 585-596.

Holler, E., Butzhammer, P., Schmid, K., Hundsrucker, C., Koestler, J., Peter, K., Zhu, W., Sporrer, D., Hehlhans, T., Kreutz, M., Holler, B., Wolff, D., Edinger, M., Andreesen, R., Levine, J. E., Ferrara, J. L., Gessner, A., Spang, R. & Oefner, P. J. 2014. Metagenomic analysis of the stool microbiome in patients receiving allogeneic stem cell transplantation: loss of diversity is associated with use of systemic antibiotics and more pronounced in gastrointestinal graft-versus-host disease. *Biol Blood Marrow Transplant*, 20, 640-5.

Holler, E., Kolb, H. J., Moller, A., Kempeni, J., Liesenfeld, S., Pechumer, H., Lehmacher, W., Ruckdeschel, G., Gleixner, B., Riedner, C. & Et Al. 1990. Increased serum levels of tumor necrosis factor alpha precede major complications of bone marrow transplantation. *Blood*, 75, 1011-6.

Hori, S., Nomura, T. & Sakaguchi, S. 2003. Control of regulatory T cell development by the transcription factor Foxp3. *Science*, 299, 1057-61.

Jacobsohn, D. A. & Vogelsang, G. B. 2007. Acute graft versus host disease. *Orphanet J Rare Dis*, 2, 35.

James, K. R., Gomes, T., Elmentaite, R., Kumar, N., Gulliver, E. L., King, H. W., Stares, M. D., Bareham, B. R., Ferdinand, J. R., Petrova, V. N., Polanski, K., Forster, S. C., Jarvis, L. B., Suchanek, O., Howlett, S., James, L. K., Jones, J. L., Meyer, K. B., Clatworthy, M. R., Saeb-Parsy, K., Lawley, T. D. & Teichmann, S. A. 2020. Distinct microbial and immune niches of the human colon. *Nat Immunol*, 21, 343-353.

Jarosch, S., Köhlen, J., Sarker, R. S. J., Steiger, K., Janssen, K.-P., Christians, A., Hennig, C., Holler, E., D'ippolito, E. & Busch, D. H. 2021. Multiplexed imaging and automated signal quantification in formalin-fixed paraffin-embedded tissues by ChipCytometry. *Cell Reports Methods*, 1, 100104.

Bibliography

- Jarosch, S., Köhlen, J., Wagner, S., D'ippolito, E. & Busch, D. H. 2022. ChipCytometry for multiplexed detection of protein and mRNA markers on human FFPE tissue samples. *STAR Protocols*, 3, 101374.
- Jeff, R., Wes, M., Jbrockmendel, Joris Van Den, B., Tom, A., Phillip, C., Gfyoung, Sinhrks, Adam, K. & Simon, H. 2021. pandas-dev/pandas: Pandas 1.3.5.
- Jenq, R. R., Taur, Y., Devlin, S. M., Ponce, D. M., Goldberg, J. D., Ahr, K. F., Littmann, E. R., Ling, L., Gobourne, A. C., Miller, L. C., Docampo, M. D., Peled, J. U., Arpaia, N., Cross, J. R., Peets, T. K., Lumish, M. A., Shono, Y., Dudakov, J. A., Poeck, H., Hanash, A. M., Barker, J. N., Perales, M. A., Giral, S. A., Pamer, E. G. & Van Den Brink, M. R. 2015. Intestinal *Blautia* Is Associated with Reduced Death from Graft-versus-Host Disease. *Biol Blood Marrow Transplant*, 21, 1373-83.
- Jenq, R. R., Ubeda, C., Taur, Y., Menezes, C. C., Khanin, R., Dudakov, J. A., Liu, C., West, M. L., Singer, N. V., Equinda, M. J., Gobourne, A., Lipuma, L., Young, L. F., Smith, O. M., Ghosh, A., Hanash, A. M., Goldberg, J. D., Aoyama, K., Blazar, B. R., Pamer, E. G. & Van Den Brink, M. R. 2012. Regulation of intestinal inflammation by microbiota following allogeneic bone marrow transplantation. *J Exp Med*, 209, 903-11.
- Jenq, R. R. & Van Den Brink, M. R. 2010. Allogeneic haematopoietic stem cell transplantation: individualized stem cell and immune therapy of cancer. *Nat Rev Cancer*, 10, 213-21.
- Jones, J. M., Wilson, R. & Bealmeary, P. M. 1971. Mortality and gross pathology of secondary disease in germfree mouse radiation chimeras. *Radiat Res*, 45, 577-88.
- Josefowicz, S. Z., Lu, L. F. & Rudensky, A. Y. 2012. Regulatory T cells: mechanisms of differentiation and function. *Annu Rev Immunol*, 30, 531-64.
- Kakahana, K., Fujioka, Y., Suda, W., Najima, Y., Kuwata, G., Sasajima, S., Mimura, I., Morita, H., Sugiyama, D., Nishikawa, H., Hattori, M., Hino, Y., Ikegawa, S., Yamamoto, K., Toya, T., Doki, N., Koizumi, K., Honda, K. & Ohashi, K. 2016. Fecal microbiota transplantation for patients with steroid-resistant acute graft-versus-host disease of the gut. *Blood*, 128, 2083-2088.
- Kamada, T., Togashi, Y., Tay, C., Ha, D., Sasaki, A., Nakamura, Y., Sato, E., Fukuoka, S., Tada, Y., Tanaka, A., Morikawa, H., Kawazoe, A., Kinoshita, T., Shitara, K., Sakaguchi, S. & Nishikawa, H. 2019. PD-1(+) regulatory T cells amplified by PD-1 blockade promote hyperprogression of cancer. *Proc Natl Acad Sci U S A*, 116, 9999-10008.
- Kang, H. M., Subramaniam, M., Targ, S., Nguyen, M., Maliskova, L., Mccarthy, E., Wan, E., Wong, S., Byrnes, L., Lanata, C. M., Gate, R. E., Mostafavi, S., Marson, A., Zaitlen, N., Criswell, L. A. & Ye, C. J. 2018. Multiplexed droplet single-cell RNA-sequencing using natural genetic variation. *Nat Biotechnol*, 36, 89-94.
- Kennedy-Darling, J., Bhate, S. S., Hickey, J. W., Black, S., Barlow, G. L., Vazquez, G., Venkataraman, V. G., Samusik, N., Goltsev, Y., Schurch, C. M. & Nolan, G. P. 2021. Highly multiplexed tissue imaging using repeated oligonucleotide exchange reaction. *Eur J Immunol*, 51, 1262-1277.
- Kolb, H. J. 2008. Graft-versus-leukemia effects of transplantation and donor lymphocytes. *Blood*, 112, 4371-83.
- Kolb, H. J., Schattenberg, A., Goldman, J. M., Hertenstein, B., Jacobsen, N., Arcese, W., Ljungman, P., Ferrant, A., Verdonck, L., Niederwieser, D., Van Rhee, F., Mittermueller, J., De Witte, T., Holler, E., Ansari, H., European Group For, B. & Marrow Transplantation Working Party Chronic, L. 1995. Graft-versus-leukemia effect of donor lymphocyte transfusions in marrow grafted patients. *Blood*, 86, 2041-50.

Bibliography

- Korngold, R. & Sprent, J. 1978. Lethal graft-versus-host disease after bone marrow transplantation across minor histocompatibility barriers in mice. Prevention by removing mature T cells from marrow. *J Exp Med*, 148, 1687-98.
- Korsunsky, I., Millard, N., Fan, J., Slowikowski, K., Zhang, F., Wei, K., Baglaenko, Y., Brenner, M., Loh, P. R. & Raychaudhuri, S. 2019. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods*, 16, 1289-1296.
- Kumagai, S., Togashi, Y., Kamada, T., Sugiyama, E., Nishinakamura, H., Takeuchi, Y., Vitaly, K., Itahashi, K., Maeda, Y., Matsui, S., Shibahara, T., Yamashita, Y., Irie, T., Tsuge, A., Fukuoka, S., Kawazoe, A., Udagawa, H., Kirita, K., Aokage, K., Ishii, G., Kuwata, T., Nakama, K., Kawazu, M., Ueno, T., Yamazaki, N., Goto, K., Tsuboi, M., Mano, H., Doi, T., Shitara, K. & Nishikawa, H. 2020. The PD-1 expression balance between effector and regulatory T cells predicts the clinical efficacy of PD-1 blockade therapies. *Nat Immunol*, 21, 1346-1358.
- Kurimoto, K., Yabuta, Y., Ohinata, Y., Ono, Y., Uno, K. D., Yamada, R. G., Ueda, H. R. & Saitou, M. 2006. An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis. *Nucleic Acids Res*, 34, e42.
- La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastrioti, M. E., Lonnerberg, P., Furlan, A., Fan, J., Borm, L. E., Liu, Z., Van Bruggen, D., Guo, J., He, X., Barker, R., Sundstrom, E., Castelo-Branco, G., Cramer, P., Adameyko, I., Linnarsson, S. & Kharchenko, P. V. 2018. RNA velocity of single cells. *Nature*, 560, 494-498.
- Lathrop, S. K., Bloom, S. M., Rao, S. M., Nutsch, K., Lio, C. W., Santacruz, N., Peterson, D. A., Stappenbeck, T. S. & Hsieh, C. S. 2011. Peripheral education of the immune system by colonic commensal microbiota. *Nature*, 478, 250-4.
- Laws, G. L., Hale, J. D. F. & Kemp, R. A. 2021. Human Systemic Immune Response to Ingestion of the Oral Probiotic *Streptococcus salivarius* BLIS K12. *Probiotics Antimicrob Proteins*, 13, 1521-1529.
- Leng, T., Akther, H. D., Hackstein, C. P., Powell, K., King, T., Friedrich, M., Christoforidou, Z., Mccuaig, S., Neyazi, M., Arancibia-Carcamo, C. V., Hagel, J., Powrie, F., Oxford, I. B. D. I., Peres, R. S., Millar, V., Ebner, D., Lamichhane, R., Ussher, J., Hinks, T. S. C., Marchi, E., Willberg, C. & Klenerman, P. 2019. TCR and Inflammatory Signals Tune Human MAIT Cells to Exert Specific Tissue Repair and Effector Functions. *Cell Rep*, 28, 3077-3091 e5.
- Levine, J. E., Paczesny, S., Mineishi, S., Braun, T., Choi, S. W., Hutchinson, R. J., Jones, D., Khaled, Y., Kitko, C. L., Bickley, D., Krijanovski, O., Reddy, P., Yanik, G. & Ferrara, J. L. 2008. Etanercept plus methylprednisolone as initial therapy for acute graft-versus-host disease. *Blood*, 111, 2470-5.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. & Genome Project Data Processing, S. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25, 2078-9.
- Lin, J. R., Fallahi-Sichani, M. & Sorger, P. K. 2015. Highly multiplexed imaging of single cells using a high-throughput cyclic immunofluorescence method. *Nat Commun*, 6, 8390.
- Lin, J. R., Izar, B., Wang, S., Yapp, C., Mei, S., Shah, P. M., Santagata, S. & Sorger, P. K. 2018. Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes. *Elife*, 7.
- Lin, T., Matsuzaki, G., Yoshida, H., Kobayashi, N., Kenai, H., Omoto, K. & Nomoto, K. 1994. CD3-CD8⁺ intestinal intraepithelial lymphocytes (IEL) and the extrathymic development of IEL. *Eur J Immunol*, 24, 1080-7.

Bibliography

Loiseau, P., Busson, M., Balere, M. L., Dormoy, A., Bignon, J. D., Gagne, K., Gebuhrer, L., Dubois, V., Jollet, I., Bois, M., Perrier, P., Masson, D., Moine, A., Absi, L., Reviron, D., Lepage, V., Tamouza, R., Toubert, A., Marry, E., Chir, Z., Jouet, J. P., Blaise, D., Charron, D. & Raffoux, C. 2007. HLA Association with hematopoietic stem cell transplantation outcome: the number of mismatches at HLA-A, -B, -C, -DRB1, or -DQB1 is strongly associated with overall survival. *Biol Blood Marrow Transplant*, 13, 965-74.

Lorenz, E., Uphoff, D., Reid, T. R. & Shelton, E. 1951. Modification of irradiation injury in mice and guinea pigs by bone marrow injections. *J Natl Cancer Inst*, 12, 197-201.

Luecken, M. D. & Theis, F. J. 2019. Current best practices in single-cell RNA-seq analysis: a tutorial. *Mol Syst Biol*, 15, e8746.

Luznik, L., O'donnell, P. V., Symons, H. J., Chen, A. R., Leffell, M. S., Zahurak, M., Gooley, T. A., Piantadosi, S., Kaup, M., Ambinder, R. F., Huff, C. A., Matsui, W., Bolanos-Meade, J., Borrello, I., Powell, J. D., Harrington, E., Warnock, S., Flowers, M., Brodsky, R. A., Sandmaier, B. M., Storb, R. F., Jones, R. J. & Fuchs, E. J. 2008. HLA-haploidentical bone marrow transplantation for hematologic malignancies using nonmyeloablative conditioning and high-dose, posttransplantation cyclophosphamide. *Biol Blood Marrow Transplant*, 14, 641-50.

Ma, X., Fan, P. X., Li, L. S., Qiao, S. Y., Zhang, G. L. & Li, D. F. 2012. Butyrate promotes the recovering of intestinal wound healing through its positive effect on the tight junctions. *J Anim Sci*, 90 Suppl 4, 266-8.

Macmillan, M. L., Weisdorf, D. J., Wagner, J. E., Defor, T. E., Burns, L. J., Ramsay, N. K., Davies, S. M. & Blazar, B. R. 2002. Response of 443 patients to steroids as primary therapy for acute graft-versus-host disease: comparison of grading systems. *Biol Blood Marrow Transplant*, 8, 387-94.

Macosko, E. Z., Basu, A., Satija, R., Nemesh, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A. R., Kamitaki, N., Martersteck, E. M., Trombetta, J. J., Weitz, D. A., Sanes, J. R., Shalek, A. K., Regev, A. & Mccarroll, S. A. 2015. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*, 161, 1202-1214.

Magnusdottir, S., Heinken, A., Kutt, L., Ravcheev, D. A., Bauer, E., Noronha, A., Greenhalgh, K., Jager, C., Baginska, J., Wilmes, P., Fleming, R. M. & Thiele, I. 2017. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat Biotechnol*, 35, 81-89.

Malard, F., Loschi, M., Cluzeau, T., Huynh, A., Guenounou, S., Borel, C., Legrand, F., Granata, A., Devillier, R., Maisano, V., Magro, L., Orvain, C., Charbonnier, A., Panz-Klapuch, M., Markiewicz, M., Desmier, D., Moya, N., Mear, J.-B., Lhomme, F., Le Jeune, C., Cornillon, J., Robin, C., Holler, E., Daguindau, E., Bilger, K., Vehreschild, M., Le Bourgeois, A., Chevallier, P., Labussière-Wallet, H., Martineau, D., Couturier, M.-A., Bulabois, C.-E., Lanic, H., Camus, V., Chantepie, S., Ceballos, P., Gasc, C., Plantamura, E. & Mohty, M. 2021. Pooled Allogenic Fecal Microbiotherapy MaaT013 for the Treatment of Steroid-Refractory Gastrointestinal Acute Graft-Versus-Host Disease: Results from the Phase IIa Heracles Study and Expanded Access Program. *Blood*, 138, 262-262.

Markey, K. A., Schluter, J., Gomes, A. L. C., Littmann, E. R., Pickard, A. J., Taylor, B. P., Giardina, P. A., Weber, D., Dai, A., Docampo, M. D., Armijo, G. K., Slingerland, A. E., Slingerland, J. B., Nichols, K. B., Brereton, D. G., Clurman, A. G., Ramos, R. J., Rao, A., Bush, A., Bohannon, L., Covington, M., Lew, M. V., Rizzieri, D. A., Chao, N., Maloy, M., Cho, C., Politikos, I., Giralt, S., Taur, Y., Pamer, E. G., Holler, E., Perales, M. A., Ponce, D. M., Devlin, S. M., Xavier, J., Sung, A. D., Peled, J. U., Cross, J. R. & Van Den Brink, M. R. M. 2020. The microbe-derived short-chain fatty acids butyrate and propionate are associated with protection from chronic GVHD. *Blood*, 136, 130-136.

Bibliography

- Marmont, A. M., Horowitz, M. M., Gale, R. P., Sobocinski, K., Ash, R. C., Van Bekkum, D. W., Champlin, R. E., Dicke, K. A., Goldman, J. M., Good, R. A. & Et Al. 1991. T-cell depletion of HLA-identical transplants in leukemia. *Blood*, 78, 2120-30.
- Mathewson, N. D., Jenq, R., Mathew, A. V., Koenigsnecht, M., Hanash, A., Toubai, T., Oravec-Wilson, K., Wu, S. R., Sun, Y., Rossi, C., Fujiwara, H., Byun, J., Shono, Y., Lindemans, C., Calafiore, M., Schmidt, T. M., Honda, K., Young, V. B., Pennathur, S., Van Den Brink, M. & Reddy, P. 2016. Gut microbiome-derived metabolites modulate intestinal epithelial cell damage and mitigate graft-versus-host disease. *Nat Immunol*, 17, 505-513.
- Matte, C. C., Liu, J., Cormier, J., Anderson, B. E., Athanasiadis, I., Jain, D., Mcniff, J. & Shlomchik, W. D. 2004. Donor APCs are required for maximal GVHD but not for GVL. *Nat Med*, 10, 987-92.
- Matzinger, P. 2002. The danger model: a renewed sense of self. *Science*, 296, 301-5.
- McInnes, L., Healy, J. & Melville, J. 2018. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction.
- Mckinney, E. F., Cuthbertson, I., Harris, K. M., Smilek, D. E., Connor, C., Manferrari, G., Carr, E. J., Zamvil, S. S. & Smith, K. G. C. 2021. A CD8(+) NK cell transcriptomic signature associated with clinical outcome in relapsing remitting multiple sclerosis. *Nat Commun*, 12, 635.
- Meedt, E., Hiergeist, A., Gessner, A., Dettmer, K., Liebisch, G., Ghimire, S., Poeck, H., Edinger, M., Wolff, D., Herr, W., Holler, E. & Weber, D. 2022. Prolonged Suppression of Butyrate-Producing Bacteria Is Associated With Acute Gastrointestinal Graft-vs-Host Disease and Transplantation-Related Mortality After Allogeneic Stem Cell Transplantation. *Clin Infect Dis*, 74, 614-621.
- Mehta, A., Rangarajan, S. & Borate, U. 2013. A cautionary tale for probiotic use in hematopoietic SCT patients-Lactobacillus acidophilus sepsis in a patient with mantle cell lymphoma undergoing hematopoietic SCT. *Bone Marrow Transplant*, 48, 461-2.
- Mielcarek, M., Martin, P. J., Leisenring, W., Flowers, M. E., Maloney, D. G., Sandmaier, B. M., Maris, M. B. & Storb, R. 2003. Graft-versus-host disease after nonmyeloablative versus conventional hematopoietic stem cell transplantation. *Blood*, 102, 756-62.
- Miller, J. S., Soignier, Y., Panoskaltis-Mortari, A., Mcneaney, S. A., Yun, G. H., Fautsch, S. K., Mckenna, D., Le, C., Defor, T. E., Burns, L. J., Orchard, P. J., Blazar, B. R., Wagner, J. E., Slungaard, A., Weisdorf, D. J., Okazaki, I. J. & Mcglave, P. B. 2005. Successful adoptive transfer and in vivo expansion of human haploidentical NK cells in patients with cancer. *Blood*, 105, 3051-7.
- Miyara, M., Yoshioka, Y., Kitoh, A., Shima, T., Wing, K., Niwa, A., Parizot, C., Taflin, C., Heike, T., Valeyre, D., Mathian, A., Nakahata, T., Yamaguchi, T., Nomura, T., Ono, M., Amoura, Z., Gorocho, G. & Sakaguchi, S. 2009. Functional delineation and differentiation dynamics of human CD4+ T cells expressing the FoxP3 transcription factor. *Immunity*, 30, 899-911.
- Mohr, F., Fischer, J. C., Nikolaus, M., Stemberger, C., Dreher, S., Verschoor, A., Haas, T., Poeck, H. & Busch, D. H. 2017. Minimally manipulated murine regulatory T cells purified by reversible Fab Multimers are potent suppressors for adoptive T-cell therapy. *Eur J Immunol*, 47, 2153-2162.
- Mohty, M., Blaise, D., Faucher, C., Vey, N., Bouabdallah, R., Stoppa, A. M., Viret, F., Gravis, G., Olive, D. & Gaugler, B. 2005. Inflammatory cytokines and acute graft-versus-host disease after reduced-intensity conditioning allogeneic stem cell transplantation. *Blood*, 106, 4407-11.

Bibliography

- Murai, M., Yoneyama, H., Ezaki, T., Suematsu, M., Terashima, Y., Harada, A., Hamada, H., Asakura, H., Ishikawa, H. & Matsushima, K. 2003. Peyer's patch is the essential site in initiating murine acute and lethal graft-versus-host reaction. *Nat Immunol*, 4, 154-60.
- Murata, M., Warren, E. H. & Riddell, S. R. 2003. A human minor histocompatibility antigen resulting from differential expression due to a gene deletion. *J Exp Med*, 197, 1279-89.
- Muth, S., Klaric, A., Radsak, M., Schild, H. & Probst, H. C. 2022. CD27 expression on Treg cells limits immune responses against tumors. *J Mol Med (Berl)*, 100, 439-449.
- Nash, R. A., Antin, J. H., Karanes, C., Fay, J. W., Avalos, B. R., Yeager, A. M., Przepiorka, D., Davies, S., Petersen, F. B., Bartels, P., Buell, D., Fitzsimmons, W., Anasetti, C., Storb, R. & Ratanatharathorn, V. 2000. Phase 3 study comparing methotrexate and tacrolimus with methotrexate and cyclosporine for prophylaxis of acute graft-versus-host disease after marrow transplantation from unrelated donors. *Blood*, 96, 2062-8.
- Nguyen, V. H., Zeiser, R., Dasilva, D. L., Chang, D. S., Beilhack, A., Contag, C. H. & Negrin, R. S. 2007. In vivo dynamics of regulatory T-cell trafficking and survival predict effective strategies to control graft-versus-host disease following allogeneic transplantation. *Blood*, 109, 2649-56.
- Nicholls, S., Piper, K. P., Mohammed, F., Dafforn, T. R., Tenzer, S., Salim, M., Mahendra, P., Craddock, C., Van Endert, P., Schild, H., Cobbold, M., Engelhard, V. H., Moss, P. A. & Willcox, B. E. 2009. Secondary anchor polymorphism in the HA-1 minor histocompatibility antigen critically affects MHC stability and TCR recognition. *Proc Natl Acad Sci U S A*, 106, 3889-94.
- Noronha, A., Modamio, J., Jarosz, Y., Guerard, E., Sompairac, N., Preciat, G., Danielsdottir, A. D., Krecke, M., Merten, D., Haraldsdottir, H. S., Heinken, A., Heirendt, L., Magnusdottir, S., Ravcheev, D. A., Sahoo, S., Gawron, P., Friscioni, L., Garcia, B., Prendergast, M., Puente, A., Rodrigues, M., Roy, A., Rouquaya, M., Wiltgen, L., Zagare, A., John, E., Krueger, M., Kuperstein, I., Zinovyev, A., Schneider, R., Fleming, R. M. T. & Thiele, I. 2019. The Virtual Metabolic Human database: integrating human and gut microbiome metabolism with nutrition and disease. *Nucleic Acids Res*, 47, D614-D624.
- Omenetti, S. & Pizarro, T. T. 2015. The Treg/Th17 Axis: A Dynamic Balance Regulated by the Gut Microbiome. *Front Immunol*, 6, 639.
- Ono, M., Shimizu, J., Miyachi, Y. & Sakaguchi, S. 2006. Control of autoimmune myocarditis and multiorgan inflammation by glucocorticoid-induced TNF receptor family-related protein(high), Foxp3-expressing CD25+ and CD25- regulatory T cells. *J Immunol*, 176, 4748-56.
- Orberg, E. T., Meedt, E., Hiergeist, A., Xue, J., Ghimire, S., Tiefgraber, M., Gödel, S., Eismann, T., Schwarz, A., Jarosch, S., Steiger, K., Gigl, M., Kleigrewe, K., Fischer, J., Janssen, K.-P., Quante, M., Heidegger, S., Herhaus, P., Verbeek, M., Ruland, J., Weber, D., Wolff, D., Edinger, M., Busch, D., Herr, W., Bassermann, F., Gessner, A., Deng, L., Holler, E. & Poeck, H. 2022. Identification of protective, metabolite-producing bacterial and viral consortia in allogeneic stem cell transplantation patients. *Research Square*.
- Orenbuch, R., Filip, I., Comito, D., Shaman, J., Pe'er, I. & Rabadan, R. 2020. arcasHLA: high-resolution HLA typing from RNAseq. *Bioinformatics*, 36, 33-40.
- Ott, S. J., Waetzig, G. H., Rehman, A., Moltzau-Anderson, J., Bharti, R., Grasis, J. A., Cassidy, L., Tholey, A., Fickenscher, H., Seegert, D., Rosenstiel, P. & Schreiber, S. 2017. Efficacy of Sterile Fecal Filtrate Transfer for Treating Patients With *Clostridium difficile* Infection. *Gastroenterology*, 152, 799-811 e7.

Bibliography

- Pandiyan, P., Bhaskaran, N., Zou, M., Schneider, E., Jayaraman, S. & Huehn, J. 2019. Microbiome Dependent Regulation of Tregs and Th17 Cells in Mucosa. *Front Immunol*, 10, 426.
- Parada Venegas, D., De La Fuente, M. K., Landskron, G., Gonzalez, M. J., Quera, R., Dijkstra, G., Harmsen, H. J. M., Faber, K. N. & Hermoso, M. A. 2019. Short Chain Fatty Acids (SCFAs)-Mediated Gut Epithelial and Immune Regulation and Its Relevance for Inflammatory Bowel Diseases. *Front Immunol*, 10, 277.
- Park, H.-J., Lee, W.-Y., Jeong, H.-Y., Kang, H.-S., Kim, J.-B. & Song, H. 2016. Mitochondrial interferon-induced transmembrane protein-1 is a critical regulator of cell death in MPRO cells. *Biotechnology and Bioprocess Engineering*, 21, 561-566.
- Passweg, J. R., Baldomero, H., Bader, P., Bonini, C., Cesaro, S., Dreger, P., Duarte, R. F., Dufour, C., Kuball, J., Farge-Bancel, D., Gennery, A., Kroger, N., Lanza, F., Nagler, A., Sureda, A. & Mohty, M. 2016. Hematopoietic stem cell transplantation in Europe 2014: more than 40 000 transplants annually. *Bone Marrow Transplant*, 51, 786-92.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. & Duchesnay, E. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.
- Peled, J. U., Gomes, A. L. C., Devlin, S. M., Littmann, E. R., Taur, Y., Sung, A. D., Weber, D., Hashimoto, D., Slingerland, A. E., Slingerland, J. B., Maloy, M., Clurman, A. G., Stein-Thoeringer, C. K., Markey, K. A., Docampo, M. D., Burgos Da Silva, M., Khan, N., Gessner, A., Messina, J. A., Romero, K., Lew, M. V., Bush, A., Bohannon, L., Brereton, D. G., Fontana, E., Amoretti, L. A., Wright, R. J., Armijo, G. K., Shono, Y., Sanchez-Escamilla, M., Castillo Flores, N., Alarcon Tomas, A., Lin, R. J., Yanez San Segundo, L., Shah, G. L., Cho, C., Scordo, M., Politikos, I., Hayasaka, K., Hasegawa, Y., Gyurkocza, B., Ponce, D. M., Barker, J. N., Perales, M. A., Giral, S. A., Jenq, R. R., Teshima, T., Chao, N. J., Holler, E., Xavier, J. B., Pamer, E. G. & Van Den Brink, M. R. M. 2020. Microbiota as Predictor of Mortality in Allogeneic Hematopoietic-Cell Transplantation. *N Engl J Med*, 382, 822-834.
- Peng, T., Thorn, K., Schroeder, T., Wang, L., Theis, F. J., Marr, C. & Navab, N. 2017. A BaSiC tool for background and shading correction of optical microscopy images. *Nat Commun*, 8, 14836.
- Picelli, S., Bjorklund, A. K., Faridani, O. R., Sagasser, S., Winberg, G. & Sandberg, R. 2013. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods*, 10, 1096-8.
- Platzbecker, U., Ehninger, G. & Bornhauser, M. 2004. Allogeneic transplantation of CD34+ selected hematopoietic cells--clinical problems and current challenges. *Leuk Lymphoma*, 45, 447-53.
- Polanski, K., Young, M. D., Miao, Z., Meyer, K. B., Teichmann, S. A. & Park, J. E. 2020. BBKNN: fast batch alignment of single cell transcriptomes. *Bioinformatics*, 36, 964-965.
- Qiu, X., Zhang, M., Yang, X., Hong, N. & Yu, C. 2013. Faecalibacterium prausnitzii upregulates regulatory T cells and anti-inflammatory cytokines in treating TNBS-induced colitis. *J Crohns Colitis*, 7, e558-68.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. & Glockner, F. O. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*, 41, D590-6.
- Radtke, A. J., Kandov, E., Lowekamp, B., Speranza, E., Chu, C. J., Gola, A., Thakur, N., Shih, R., Yao, L., Yaniv, Z. R., Beuschel, R. T., Kabat, J., Croteau, J., Davis, J., Hernandez, J. M. & Germain,

Bibliography

R. N. 2020. IBEX: A versatile multiplex optical imaging approach for deep phenotyping and spatial analysis of cells in complex tissues. *Proc Natl Acad Sci U S A*, 117, 33455-33465.

Rajagopalan, A., Venkatesh, I., Aslam, R., Kirchenbuechler, D., Khanna, S., Cimbaluk, D., Kordower, J. H. & Gupta, V. 2021. SeqStain is an efficient method for multiplexed, spatialomic profiling of human and murine tissues. *Cell Rep Methods*, 1.

Rapoport, A. P., Stadtmauer, E. A., Aqui, N., Badros, A., Cotte, J., Chrisley, L., Veloso, E., Zheng, Z., Westphal, S., Mair, R., Chi, N., Ratterree, B., Pochran, M. F., Natt, S., Hinkle, J., Sickles, C., Sohal, A., Ruehle, K., Lynch, C., Zhang, L., Porter, D. L., Luger, S., Guo, C., Fang, H. B., Blackwelder, W., Hankey, K., Mann, D., Edelman, R., Frasch, C., Levine, B. L., Cross, A. & June, C. H. 2005. Restoration of immunity in lymphopenic individuals with cancer by vaccination and adoptive T-cell transfer. *Nat Med*, 11, 1230-7.

Riegel, C., Boeld, T. J., Doser, K., Huber, E., Hoffmann, P. & Edinger, M. 2020. Efficient treatment of murine acute GvHD by in vitro expanded donor regulatory T cells. *Leukemia*, 34, 895-908.

Rieger, K., Loddenkemper, C., Maul, J., Fietz, T., Wolff, D., Terpe, H., Steiner, B., Berg, E., Miehle, S., Bornhauser, M., Schneider, T., Zeitz, M., Stein, H., Thiel, E., Duchmann, R. & Uharek, L. 2006. Mucosal FOXP3⁺ regulatory T cells are numerically deficient in acute and chronic GvHD. *Blood*, 107, 1717-23.

Rosenberg, A. B., Roco, C. M., Muscat, R. A., Kuchina, A., Sample, P., Yao, Z., Graybuck, L. T., Peeler, D. J., Mukherjee, S., Chen, W., Pun, S. H., Sellers, D. L., Tasic, B. & Seelig, G. 2018. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science*, 360, 176-182.

Ruggeri, L., Capanni, M., Urbani, E., Perruccio, K., Shlomchik, W. D., Tosti, A., Posati, S., Rogaia, D., Frassoni, F., Aversa, F., Martelli, M. F. & Velardi, A. 2002. Effectiveness of donor natural killer cell alloreactivity in mismatched hematopoietic transplants. *Science*, 295, 2097-100.

Saka, S. K., Wang, Y., Kishi, J. Y., Zhu, A., Zeng, Y., Xie, W., Kirli, K., Yapp, C., Cicconet, M., Beliveau, B. J., Lapan, S. W., Yin, S., Lin, M., Boyden, E. S., Kaeser, P. S., Pihan, G., Church, G. M. & Yin, P. 2019. Immuno-SABER enables highly multiplexed and amplified protein imaging in tissues. *Nat Biotechnol*, 37, 1080-1090.

Sakaguchi, S., Sakaguchi, N., Asano, M., Itoh, M. & Toda, M. 1995. Immunologic self-tolerance maintained by activated T cells expressing IL-2 receptor alpha-chains (CD25). Breakdown of a single mechanism of self-tolerance causes various autoimmune diseases. *J Immunol*, 155, 1151-64.

Schapiro, D., Sokolov, A., Yapp, C., Chen, Y. A., Muhlich, J. L., Hess, J., Creason, A. L., Nirmal, A. J., Baker, G. J., Nariya, M. K., Lin, J. R., Maliga, Z., Jacobson, C. A., Hodgman, M. W., Ruokonen, J., Farhi, S. L., Abbondanza, D., Mckinley, E. T., Persson, D., Betts, C., Sivagnanam, S., Regev, A., Goecks, J., Coffey, R. J., Coussens, L. M., Santagata, S. & Sorger, P. K. 2022. MCMICRO: a scalable, modular image-processing pipeline for multiplexed tissue imaging. *Nat Methods*, 19, 311-315.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P. & Cardona, A. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods*, 9, 676-82.

Schoemans, H. M., Lee, S. J., Ferrara, J. L., Wolff, D., Levine, J. E., Schultz, K. R., Shaw, B. E., Flowers, M. E., Ruutu, T., Greinix, H., Holler, E., Basak, G., Duarte, R. F., Pavletic, S. Z., Party, E. T. C. W. & The, E.-N. I. H. C. G. T. F. 2018. EBMT-NIH-CIBMTR Task Force position statement on

Bibliography

standardized terminology & guidance for graft-versus-host disease assessment. *Bone Marrow Transplant*, 53, 1401-1415.

Seabold, S. & Perktold, J. 2010. statsmodels: Econometric and statistical modeling with python. *9th Python in Science Conference*.

Sebastianjarosch 2022. SebastianJarosch/ChipCytometry-Image-Processing: ChipCytometry image processing. v1.1.0 ed.: Zenodo.

Shimabukuro-Vornhagen, A., Hallek, M. J., Storb, R. F. & Von Bergwelt-Baildon, M. S. 2009. The role of B cells in the pathogenesis of graft-versus-host disease. *Blood*, 114, 4919-27.

Shono, Y. & Van Den Brink, M. R. M. 2018. Gut microbiota injury in allogeneic haematopoietic stem cell transplantation. *Nat Rev Cancer*, 18, 283-295.

Simonetta, F., Alvarez, M. & Negrin, R. S. 2017. Natural Killer Cells in Graft-versus-Host-Disease after Allogeneic Hematopoietic Cell Transplantation. *Front Immunol*, 8, 465.

Sinicrope, F. A., Rego, R. L., Ansell, S. M., Knutson, K. L., Foster, N. R. & Sargent, D. J. 2009. Intraepithelial effector (CD3+)/regulatory (FoxP3+) T-cell ratio predicts a clinical outcome of human colon carcinoma. *Gastroenterology*, 137, 1270-9.

Skelton, R. A., Javed, A., Zheng, L. & He, J. 2017. Overcoming the resistance of pancreatic cancer to immune checkpoint inhibitors. *Journal of Surgical Oncology*, 116, 55-62.

Smith, P. M., Howitt, M. R., Panikov, N., Michaud, M., Gallini, C. A., Bohlooly, Y. M., Glickman, J. N. & Garrett, W. S. 2013. The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science*, 341, 569-73.

Sperlich, A., Balmert, A., Doll, D., Bauer, S., Franke, F., Keller, G., Wilhelm, D., Mur, A., Respondek, M., Friess, H., Nitsche, U. & Janssen, K. P. 2018. Genetic and immunological biomarkers predict metastatic disease recurrence in stage III colon cancer. *BMC Cancer*, 18, 998.

Spierings, E., Gras, S., Reiser, J. B., Mommaas, B., Almekinders, M., Kester, M. G., Chouquet, A., Le Gorrec, M., Drijfhout, J. W., Ossendorp, F., Housset, D. & Goulmy, E. 2009. Steric hindrance and fast dissociation explain the lack of immunogenicity of the minor histocompatibility HA-1Arg Null allele. *J Immunol*, 182, 4809-16.

Spindelboeck, W., Schulz, E., Uhl, B., Kashofer, K., Aigelsreiter, A., Zinke-Cerwenka, W., Mulabecirovic, A., Kump, P. K., Halwachs, B., Gorkiewicz, G., Sill, H., Greinix, H., Hogenauer, C. & Neumeister, P. 2017. Repeated fecal microbiota transplantations attenuate diarrhea and lead to sustained changes in the fecal microbiota in acute, refractory gastrointestinal graft-versus-host-disease. *Haematologica*, 102, e210-e213.

Stahl, P. L., Salmen, F., Vickovic, S., Lundmark, A., Navarro, J. F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J. O., Huss, M., Mollbrink, A., Linnarsson, S., Codeluppi, S., Borg, A., Ponten, F., Costea, P. I., Sahlen, P., Mulder, J., Bergmann, O., Lundeberg, J. & Frisen, J. 2016. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, 353, 78-82.

Stammler, F., Glasner, J., Hiergeist, A., Holler, E., Weber, D., Oefner, P. J., Gessner, A. & Spang, R. 2016. Adjusting microbiome profiles for differences in microbial load by spike-in bacteria. *Microbiome*, 4, 28.

Bibliography

Stein-Thoeringer, C. K., Nichols, K. B., Lazrak, A., Docampo, M. D., Slingerland, A. E., Slingerland, J. B., Clurman, A. G., Armijo, G., Gomes, A. L. C., Shono, Y., Staffas, A., Burgos Da Silva, M., Devlin, S. M., Markey, K. A., Bajic, D., Pinedo, R., Tsakmaklis, A., Littmann, E. R., Pastore, A., Taur, Y., Monette, S., Arcila, M. E., Pickard, A. J., Maloy, M., Wright, R. J., Amoretti, L. A., Fontana, E., Pham, D., Jamal, M. A., Weber, D., Sung, A. D., Hashimoto, D., Scheid, C., Xavier, J. B., Messina, J. A., Romero, K., Lew, M., Bush, A., Bohannon, L., Hayasaka, K., Hasegawa, Y., Vehreschild, M., Cross, J. R., Ponce, D. M., Perales, M. A., Giralt, S. A., Jenq, R. R., Teshima, T., Holler, E., Chao, N. J., Pamer, E. G., Peled, J. U. & Van Den Brink, M. R. M. 2019. Lactose drives Enterococcus expansion to promote graft-versus-host disease. *Science*, 366, 1143-1149.

Stem Cell Trialists' Collaborative, G. 2005. Allogeneic peripheral blood stem-cell compared with bone marrow transplantation in the management of hematologic malignancies: an individual patient data meta-analysis of nine randomized trials. *J Clin Oncol*, 23, 5074-87.

Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P. K., Swerdlow, H., Satija, R. & Smibert, P. 2017. Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods*, 14, 865-868.

Stoeckius, M., Zheng, S., Houck-Loomis, B., Hao, S., Yeung, B. Z., Mauck, W. M., 3rd, Smibert, P. & Satija, R. 2018. Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol*, 19, 224.

Storb, R. F., Lucarelli, G., Mcsweeney, P. A. & Childs, R. W. 2003. Hematopoietic cell transplantation for benign hematological disorders and solid tumors. *Hematology Am Soc Hematol Educ Program*, 372-97.

Sturm, G., Szabo, T., Fotakis, G., Haider, M., Rieder, D., Trajanoski, Z. & Finotello, F. 2020. Scirpy: a Scanpy extension for analyzing single-cell T-cell receptor-sequencing data. *Bioinformatics*, 36, 4817-4818.

Sugiyama, D., Nishikawa, H., Maeda, Y., Nishioka, M., Tanemura, A., Katayama, I., Ezo, S., Kanakura, Y., Sato, E., Fukumori, Y., Karbach, J., Jager, E. & Sakaguchi, S. 2013. Anti-CCR4 mAb selectively depletes effector-type FoxP3+CD4+ regulatory T cells, evoking antitumor immune responses in humans. *Proc Natl Acad Sci U S A*, 110, 17945-50.

Swimm, A., Giver, C. R., Defilipp, Z., Rangaraju, S., Sharma, A., Ulezko Antonova, A., Sonowal, R., Capaldo, C., Powell, D., Qayed, M., Kalman, D. & Waller, E. K. 2018. Indoles derived from intestinal microbiota act via type I interferon signaling to limit graft-versus-host disease. *Blood*, 132, 2506-2519.

Tan, H., Zhao, J., Zhang, H., Zhai, Q. & Chen, W. 2019. Novel strains of *Bacteroides fragilis* and *Bacteroides ovatus* alleviate the LPS-induced inflammation in mice. *Appl Microbiol Biotechnol*, 103, 2353-2365.

Tang, F., Barbacioru, C., Wang, Y., Nordman, E., Lee, C., Xu, N., Wang, X., Bodeau, J., Tuch, B. B., Siddiqui, A., Lao, K. & Surani, M. A. 2009. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods*, 6, 377-82.

Taur, Y., Jenq, R. R., Perales, M. A., Littmann, E. R., Morjaria, S., Ling, L., No, D., Gobourne, A., Viale, A., Dahi, P. B., Ponce, D. M., Barker, J. N., Giralt, S., Van Den Brink, M. & Pamer, E. G. 2014. The effects of intestinal tract bacterial diversity on mortality following allogeneic hematopoietic stem cell transplantation. *Blood*, 124, 1174-82.

Taur, Y., Xavier, J. B., Lipuma, L., Ubeda, C., Goldberg, J., Gobourne, A., Lee, Y. J., Dubin, K. A., Socci, N. D., Viale, A., Perales, M. A., Jenq, R. R., Van Den Brink, M. R. & Pamer, E. G. 2012.

Bibliography

Intestinal domination and the risk of bacteremia in patients undergoing allogeneic hematopoietic stem cell transplantation. *Clin Infect Dis*, 55, 905-14.

Thomas, A. C., Michael, D., Antony, L., Elliott Sales De, A., John, H., Eric, F., Tim, H., Jody, K., David, S. & Nelle, V. 2021. matplotlib/matplotlib: REL: v3.3.4.

Thomas, E. D., Lochte, H. L., Jr., Cannon, J. H., Sahler, O. D. & Ferrebee, J. W. 1959. Supralethal whole body irradiation and isologous marrow transplantation in man. *J Clin Invest*, 38, 1709-16.

Tirosh, I., Izar, B., Prakadan, S. M., Wadsworth, M. H., 2nd, Treacy, D., Trombetta, J. J., Rotem, A., Rodman, C., Lian, C., Murphy, G., Fallahi-Sichani, M., Dutton-Regester, K., Lin, J. R., Cohen, O., Shah, P., Lu, D., Genshaft, A. S., Hughes, T. K., Ziegler, C. G., Kazer, S. W., Gaillard, A., Kolb, K. E., Villani, A. C., Johannessen, C. M., Andreev, A. Y., Van Allen, E. M., Bertagnolli, M., Sorger, P. K., Sullivan, R. J., Flaherty, K. T., Frederick, D. T., Jane-Valbuena, J., Yoon, C. H., Rozenblatt-Rosen, O., Shalek, A. K., Regev, A. & Garraway, L. A. 2016. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*, 352, 189-96.

Togashi, Y., Shitara, K. & Nishikawa, H. 2019. Regulatory T cells in cancer immunosuppression - implications for anticancer therapy. *Nat Rev Clin Oncol*, 16, 356-371.

Tomblyn, M., Chiller, T., Einsele, H., Gress, R., Sepkowitz, K., Storek, J., Wingard, J. R., Young, J. A., Boeckh, M. J., Center for International, B., Marrow, R., National Marrow Donor, P., European, B., Marrowtransplant, G., American Society Of, B., Marrow, T., Canadian, B., Marrow Transplant, G., Infectious Diseases Society Of, A., Society for Healthcare Epidemiology Of, A., Association of Medical, M., Infectious Disease, C., Centers for Disease, C. & Prevention 2009. Guidelines for preventing infectious complications among hematopoietic cell transplantation recipients: a global perspective. *Biol Blood Marrow Transplant*, 15, 1143-238.

Traag, V. A., Waltman, L. & Van Eck, N. J. 2019. From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep*, 9, 5233.

Ustun, C., Defor, T. E., Karadag, F. K., Don Yun, H., Nathan, S., Brunstein, C. G., Blazar, B. R., Weisdorf, D. J., Holtan, S. G. & Amin, K. 2020. Tissue mast cell counts may be associated with decreased severity of gastrointestinal acute GVHD and nonrelapse mortality. *Blood Adv*, 4, 2317-2324.

Van Bekkum, D. W., Roodenburg, J., Heidt, P. J. & Van Der Waaij, D. 1974. Mitigation of secondary disease of allogeneic mouse radiation chimeras by modification of the intestinal microflora. *J Natl Cancer Inst*, 52, 401-4.

Van Den Brink, M. R. & Burakoff, S. J. 2002. Cytolytic pathways in haematopoietic stem-cell transplantation. *Nat Rev Immunol*, 2, 273-81.

Van Der Maaten, L. & Hinton, G. 2008. Visualizing Data using t-SNE. *Journal of Machine Learning Research*, 9, 2579-2605.

Vander Lugt, M. T., Braun, T. M., Hanash, S., Ritz, J., Ho, V. T., Antin, J. H., Zhang, Q., Wong, C. H., Wang, H., Chin, A., Gomez, A., Harris, A. C., Levine, J. E., Choi, S. W., Couriel, D., Reddy, P., Ferrara, J. L. & Paczesny, S. 2013. ST2 as a marker for risk of therapy-resistant graft-versus-host disease and death. *N Engl J Med*, 369, 529-39.

Varelias, A., Bunting, M. D., Ormerod, K. L., Koyama, M., Olver, S. D., Straube, J., Kuns, R. D., Robb, R. J., Henden, A. S., Cooper, L., Lachner, N., Gartlan, K. H., Lantz, O., Kjer-Nielsen, L., Mak, J. Y., Fairlie, D. P., Clouston, A. D., McCluskey, J., Rossjohn, J., Lane, S. W., Hugenholtz, P. & Hill, G. R.

Bibliography

2018. Recipient mucosal-associated invariant T cells control GVHD within the colon. *J Clin Invest*, 128, 1919-1936.
- Virshup, I., Rybakov, S., Theis, F. J., Angerer, P. & Wolf, F. A. 2021. anndata: Annotated data. *bioRxiv*, 2021.12.16.473007.
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., Van Der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, I., Feng, Y., Moore, E. W., Vanderplas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., Van Mulbregt, P. & Scipy, C. 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*, 17, 261-272.
- Vogelsang, G. B., Lee, L. & Bensen-Kennedy, D. M. 2003. Pathogenesis and treatment of graft-versus-host disease after bone marrow transplant. *Annu Rev Med*, 54, 29-52.
- Waldman, E., Lu, S. X., Hubbard, V. M., Kochman, A. A., Eng, J. M., Terwey, T. H., Muriglan, S. J., Kim, T. D., Heller, G., Murphy, G. F., Liu, C., Alpdogan, O. & Van Den Brink, M. R. 2006. Absence of beta7 integrin results in less graft-versus-host disease because of decreased homing of alloreactive T cells to intestine. *Blood*, 107, 1703-11.
- Wang, F., Flanagan, J., Su, N., Wang, L. C., Bui, S., Nielson, A., Wu, X., Vo, H. T., Ma, X. J. & Luo, Y. 2012. RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J Mol Diagn*, 14, 22-9.
- Wang, W., Meadows, L. R., Den Haan, J. M., Sherman, N. E., Chen, Y., Blokland, E., Shabanowitz, J., Agulnik, A. I., Hendrickson, R. C., Bishop, C. E. & Et Al. 1995. Human H-Y: a male-specific histocompatibility antigen derived from the SMCY protein. *Science*, 269, 1588-90.
- Waskom, M. L. 2021. seaborn: statistical data visualization. *Journal of Open Source Software*, 6, 3021.
- Weber, D., Oefner, P. J., Hiergeist, A., Koestler, J., Gessner, A., Weber, M., Hahn, J., Wolff, D., Stammler, F., Spang, R., Herr, W., Dettmer, K. & Holler, E. 2015. Low urinary indoxyl sulfate levels early after transplantation reflect a disrupted microbiome and are associated with poor outcome. *Blood*, 126, 1723-8.
- Weigert, M. S., U.; Haase, R.; Sugawara, K.; Myers, G. 2020. Star-convex Polyhedra for 3D Object Detection and Segmentation in Microscopy. *EEE Winter Conference on Applications of Computer Vision (WACV)*, 2020, 3655-3662.
- Westin, J. R., Saliba, R. M., De Lima, M., Alousi, A., Hosing, C., Qazilbash, M. H., Khouri, I. F., Shpall, E. J., Anderlini, P., Rondon, G., Andersson, B. S., Champlin, R. & Couriel, D. R. 2011. Steroid-Refractory Acute GVHD: Predictors and Outcomes. *Adv Hematol*, 2011, 601953.
- Wing, K., Onishi, Y., Prieto-Martin, P., Yamaguchi, T., Miyara, M., Fehervari, Z., Nomura, T. & Sakaguchi, S. 2008. CTLA-4 control over Foxp3+ regulatory T cell function. *Science*, 322, 271-5.
- Wohnhaas, C. T., Lepar, G. G., Fernandez-Albert, F., Kind, D., Gantner, F., Violette, C., Hildebrandt, T. & Baum, P. 2019. DMSO cryopreservation is the method of choice to preserve cells for droplet-based single-cell RNA sequencing. *Sci Rep*, 9, 10699.
- Wolf, F. A., Angerer, P. & Theis, F. J. 2018. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol*, 19, 15.

Bibliography

- Wu, Y., Fu, J., Wang, H. & Yu, X. Z. 2021. Donor T-Cell Repertoire Profiling in Recipient Lymphoid and Parenchyma Organs Reveals GVHD Pathogenesis at Clonal Levels After Bone Marrow Transplantation in Mice. *Front Immunol*, 12, 778996.
- Wysocki, C. A., Panoskaltsis-Mortari, A., Blazar, B. R. & Serody, J. S. 2005. Leukocyte migration and graft-versus-host disease. *Blood*, 105, 4191-9.
- Xu, J., Falconer, C., Nguyen, Q., Crawford, J., Mckinnon, B. D., Mortlock, S., Senabouth, A., Andersen, S., Chiu, H. S., Jiang, L., Palpant, N. J., Yang, J., Mueller, M. D., Hewitt, A. W., Pebay, A., Montgomery, G. W., Powell, J. E. & Coin, L. J. M. 2019. Genotype-free demultiplexing of pooled single-cell RNA-seq. *Genome Biol*, 20, 290.
- Yaghoubi, N., Soltani, A., Ghazvini, K., Hassanian, S. M. & Hashemy, S. I. 2019. PD-1/ PD-L1 blockade as a novel treatment for colorectal cancer. *Biomed Pharmacother*, 110, 312-318.
- Yang, J., Yang, F., Campos, L. S., Mansfield, W., Skelton, H., Hooks, Y. & Liu, P. 2017. Quenching autofluorescence in tissue immunofluorescence. *Wellcome Open Research*, 2.
- Youngster, I., Russell, G. H., Pindar, C., Ziv-Baran, T., Sauk, J. & Hohmann, E. L. 2014. Oral, capsulized, frozen fecal microbiota transplantation for relapsing *Clostridium difficile* infection. *JAMA*, 312, 1772-8.
- Yu, X., Harden, K., Gonzalez, L. C., Francesco, M., Chiang, E., Irving, B., Tom, I., Ivelja, S., Refino, C. J., Clark, H., Eaton, D. & Grogan, J. L. 2009. The surface protein TIGIT suppresses T cell activation by promoting the generation of mature immunoregulatory dendritic cells. *Nat Immunol*, 10, 48-57.
- Zeiser, R., Von Bubnoff, N., Butler, J., Mohty, M., Niederwieser, D., Or, R., Szer, J., Wagner, E. M., Zuckerman, T., Mahuzier, B., Xu, J., Wilke, C., Gandhi, K. K., Socie, G. & Group, R. T. 2020. Ruxolitinib for Glucocorticoid-Refractory Acute Graft-versus-Host Disease. *N Engl J Med*, 382, 1800-1810.
- Zeng, D., Lewis, D., Dejbakhsh-Jones, S., Lan, F., Garcia-Ojeda, M., Sibley, R. & Strober, S. 1999. Bone marrow NK1.1(-) and NK1.1(+) T cells reciprocally regulate acute graft versus host disease. *J Exp Med*, 189, 1073-81.
- Zhang, Y., Louboutin, J. P., Zhu, J., Rivera, A. J. & Emerson, S. G. 2002. Preterminal host dendritic cells in irradiated mice prime CD8+ T cell-mediated acute graft-versus-host disease. *J Clin Invest*, 109, 1335-44.
- Zhao, Y., Li, X., Zhou, Y., Gao, J., Jiao, Y., Zhu, B., Wu, D. & Qi, X. 2021. Safety and Efficacy of Fecal Microbiota Transplantation for Grade IV Steroid Refractory GI-GvHD Patients: Interim Results From FMT2017002 Trial. *Front Immunol*, 12, 678476.
- Zheng, G. X., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B., Wheeler, T. D., Mcdermott, G. P., Zhu, J., Gregory, M. T., Shuga, J., Montesclaros, L., Underwood, J. G., Masquelier, D. A., Nishimura, S. Y., Schnall-Levin, M., Wyatt, P. W., Hindson, C. M., Bharadwaj, R., Wong, A., Ness, K. D., Beppu, L. W., Deeg, H. J., Mcfarland, C., Loeb, K. R., Valente, W. J., Ericson, N. G., Stevens, E. A., Radich, J. P., Mikkelsen, T. S., Hindson, B. J. & Bielas, J. H. 2017. Massively parallel digital transcriptional profiling of single cells. *Nat Commun*, 8, 14049.
- Zheng, P., Tamaresis, J., Thangavelu, G., Xu, L., You, X., Blazar, B. R., Negrin, R. S., Zehnder, J. L., Iliopoulou, B. P. & Meyer, E. H. 2020. Recipient-specific T-cell repertoire reconstitution in the gut following murine hematopoietic cell transplant. *Blood Adv*, 4, 4232-4243.

Bibliography

Zhu, Y. Y., Machleder, E. M., Chenchik, A., Li, R. & Siebert, P. D. 2001. Reverse transcriptase template switching: a SMART approach for full-length cDNA library construction. *Biotechniques*, 30, 892-7.