Gene expression and T cell receptor profiling of *in vitro* expanded regulatory T cells in murine acute graft versus host disease models

DISSERTATION ZUR ERLANGUNG DES
DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT.)
DER FAKULTÄT FÜR BIOLOGIE UND VORKLINISCHE MEDIZIN
DER UNIVERSITÄT REGENSBURG

vorgelegt von

David Jonathan Dittmar

aus

Mainburg

im Jahr 2022

Gene expression and T cell receptor profiling of *in vitro* expanded regulatory T cells in murine acute graft versus host disease models

DISSERTATION ZUR ERLANGUNG DES
DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT.)
DER FAKULTÄT FÜR BIOLOGIE UND VORKLINISCHE MEDIZIN
DER UNIVERSITÄT REGENSBURG

vorgelegt von

David Jonathan Dittmar

aus

Mainburg

im Jahr 2022

Das Promotionsgesuch wurde eingereicht am:
7. Juli 2022

Die Arbeit wurde angeleitet von:
Prof. Dr. Michael Rehli

Unterschrift:

David Jonathan Dittmar

Contents

List of Figures

List of Tables

Summary

Acute graft versus host disease (aGvHD) poses the main cause of morbidity and mortality in patients having undergone allogeneic hematopoietic stem cell transplantation (allo HSCT), the only curative option for many hematologic malignancies. Especially for steroid-refractory forms of GvHD, clinicians are in dire need of novel and drug-independent treatment options. Constitutively *FOXP3*-expressing $CD4^+CD25^+$ regulatory T cells (Treg) are a successfully emerging therapeutic avenue, however, the prerequisites for their clinical efficacy in controlling aGvHD are not fully understood. Thus, sophisticated murine models and flow cytometry, facilitated by the work group around PD Dr. rer. nat. Petra Hoffmann and Prof. Dr. med. Matthias Edinger (WG Edinger/Hoffmann), were combined with several next generation sequencing (NGS) approaches, in an effort to shed light on the complex interplay between migration patterns, site-specific transcription profiles, and T cell receptor (TCR) repertoire dynamics of Treg during aGvHD emergence. In the present work, the conventional polyclonal *in vitro* expansion ("poly") of murine Treg was compared to an alloantigen-specific approach ("allo"). While allo Treg had a markedly less diverse T cell receptor (TCR) repertoire, both cell products retained their core Treg expression profiles, hence ameliorated GvHD symptoms in recipients following MHC-mismatched (C57BL/6→BALB/c) bone marrow transplantation (BMT). After seven days, Treg were re-isolated from recipient colon, liver and spleen compartments for gene expression and TCR repertoire profiling on the mRNA level. The directive role of the colon for the rapid, organ-specific adaption of Treg gene expression was most striking. Additionally, in the presence of GvHD-inducing conventional T cells (Tcon), a suppressive gene signature was induced, encoding anti-inflammatory cytokines, such as Il10, Fgl2 and Tgfb1, inhibitory surface molecules, such as PD-L1, PD-L2 or Ctla4 as well as directly cytotoxic granzymes

and perforin. Besides three *Tcrb* variable gene segments (*Trbv12-1*, *Trbv12-2*, *Trbv26*) found enriched in the colon of BMT mice, dominant donor Treg clonotypes could be observed across analyzed organs and were also largely shared between organs of individuals. A dominant allo-response became apparent across all tissues with little evidence for tissue-restricted clonal expansion of Treg.

Zusammenfassung

Die akute *Graft versus Host Disease* (aGvHD) stellt die Hauptursache für die hohe Morbidität und Mortalität bei Patienten nach allogener hämatopoetischer Stammzelltransplantation (allo-HSZT) dar. Bis heute ist die allo-HSZT der einzige kurative Therapieansatz für viele maligne hämatologische Erkrankungen. Insbesondere für die Behandlung von Kortikosteroid-refraktären Formen der aGvHD sind neue, Medikamenten-unabhängige Möglichkeiten dringend erforderlich. Ein neues, erfolgsversprechendes Therapeutikum bieten konstitutiv *FOXP3*-exprimierende, $CD4^+CD25^+$ regulatorische T-Zellen (Treg), deren klinische Wirksamkeit jedoch nicht hinreichend erforscht ist. Diese Arbeit hatte daher zum Ziel, deren Funktion im GvHD-Kontext besser zu verstehen. In Kooperation mit der Arbeitsgruppe von PD Dr. rer. nat. Petra Hoffmann und Prof. Dr. med. Matthias Edinger wurden komplexe Mausmodelle sowie Durchflusszytometrie mit verschiedenen *Next-Generation-Sequencing* (NGS)-Methoden verbunden, um das komplexe Zusammenspiel zwischen dem Migrationsverhalten, Organ-spezifischen Genexpressionsprofilen und der Dynamik des T-Zell-Rezeptor (TZR)-Repertoires von Treg während akuter GvHD zu beleuchten. Der Fokus lag dabei auf dem Vergleich der konventionellen, polyklonalen *in vitro*-Expansion muriner Treg ("poly-Treg") mit dem Ansatz einer Allo-Antigen-spezifischen Expansion ("allo-Treg"). Während das T-Zell-Rezeptor (TZR)-Repertoire der allo-Treg deutlich weniger Diversität aufwies, konnte gezeigt werden, dass die Expression von Treg-typischen Genen in beiden Zelltypen beibehalten wurde. Dies führte zu einer Verbesserung der GvHD-Symptome in Empfängern nach einer MHC-disparaten (C57BL/6→BALB/c) Knochenmarkstransplantation (KMT). Nach sieben Tagen wurden die transplantierten Treg aus den Organen Leber, Milz und Dickdarm der Empfänger isoliert, um deren Genexpression sowie die Zusammensetzung des TZR-Repertoires auf mRNA-Ebene zu untersuchen. Dabei

war besonders die Bedeutung des Dickdarms im Hinblick auf eine schnelle, organspezifische Anpassung der Genexpression auffällig. Des Weiteren konnte gezeigt werden, dass die Kotransplantation alloreaktiver, konventioneller T Zellen in Treg die Expression suppressiver Gene zur Folge hatte. Diese umfassten Transkripte antiinflammatorischer Zytokine (IL10, Fgl2, Tgfb1), inhibitorischer Oberflächenproteine (PD-L1, PD-L2, Ctla4) aber auch Granzyme und Perforin mit direkter zytotoxischer Wirkung. Neben einer Anreicherung von drei variablen Gen-Segmenten (*Trbv12-1*, *Trbv12-2*, *Trbv26*) im Colon von KMT-Mäusen, wurde beobachtet, dass dominante Klonotypen transplantierter Treg in allen untersuchten Organen auftraten und zwischen den Organen der Individuen weitestgehend gleichmäßig verteilt waren. Darüber hinaus konnten eine Allo-Antigen-spezifische Antwort in allen Geweben sowie wenig Evidenz für die Gewebs-spezifische klonale Expansion von Treg angeführt werden.

CHAPTER 1

Introduction

1.1 Hematopoietic stem cell transplantation

To date, hematopoietic stem cell transplantation (HSCT) is the only curative treatment option for most patients suffering from leukemia or other hematologic malignancies. While the stem cell source in the syngeneic or autologous HSCT is equal to the recipient, the transplanted cells are derived from a different individual in case of the allogeneic HSCT (allo HSCT). Numerous indications have been gathered and categorized by the European Society for Blood and Marrow Transplantation (EBMT) to provide a guideline for HSCT in children and adults, describing acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL) as the most common rationale for allo HSCT in grown-ups. In contrast, for childhood AML, allo HSCT is only considered standard of care in high-risk patients after their first complete remission, and especially in children, primary immunodeficiencies are frequent indications for allo HSCT (Duarte et al., 2019).

Generally, in the procedure also referred to as bone marrow transplantation (BMT), hematopoietic stem cells are derived from suitable donor bone marrow (BM), peripheral blood (PB) or cord blood (CB), and are subsequently transfused to a recipient for the repopulation of the vacant bone marrow niche. Usually, the recipient has to be conditioned by radiotherapy, chemotherapy, or a combination of both, with the aim of depleting the hematopoietic system, before the stem cells can be engrafted. Ideally, after the successful immune reconstitution, patients possess a healthy hematopoiesis with

a complete donor bone marrow chimerism and a minimal risk of relapse (Duarte et al., 2019). The latter is owed to a further beneficial effect of the allo HSCT, known as graft versus leukemia (GvL) effect which describes the elimination of residual leukemic cells by alloreactive donor T cells (Appelbaum, 2001; D. W. Barnes et al., 1956). Finding a suitable donor, however, can be challenging. Depending on the indication, the degree of tolerable mismatch on the human leukocyte antigen (HLA) level, and the preferred origin of the stem cell graft have to be carefully considered. As a general rule, HLA-matched sibling donors (MSDs) are favored over unrelated matched donors (MUDs), or unrelated individuals with permissive major histocompatibility complex (MHC) disparities, termed mismatched unrelated donors (MMUDs) (Duarte et al., 2019). Moreover, by virtue of the absence of a suitable matched donor or particular clinical indications, HLA haploidentical HSCT (haplo HSCT) is an additional treatment option. However, to limit the extent of alloreactivity, the graft is frequently depleted of T cells prior to transplant, a valid strategy for optimized transplant outcomes (X. Zheng & Tian, 2021). Still, a comparative study based on data gathered from 2011 trough 2018 suggests MUDs to be favorable over haploidentical donors (Gooptu et al., 2021). The main steps of the allogeneic transplantation procedure are illustrated in **Figure 1.1** in a simplified manner.

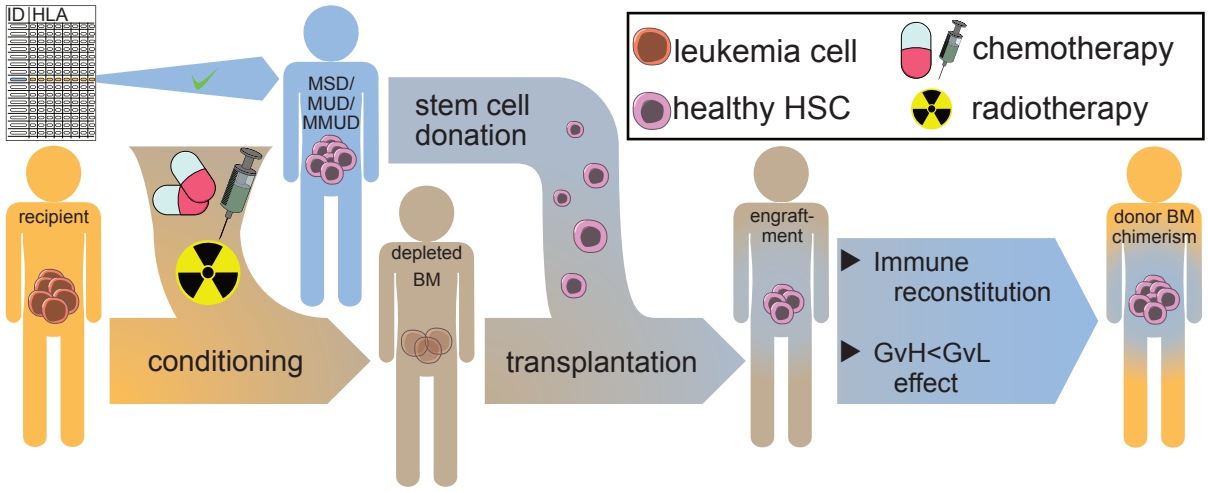

Figure 1.1: **Schematic overview of the allo HSCT procedure.** A healthy donor (MSD, MUD, MMUD) provides stem cells, either derived from BM, PB or CB. The patient is conditioned by radiation and/or chemotherapy to deplete the bone marrow compartment. Next, the recipient is infused with the donor HSCs for their engraftment. After immune reconstitution, the patient is capable of a healthy hematopoiesis. MSD: HLA-matched sibling donor, MUD: HLA-matched unrelated donor, MMUD: HLA-mismatched unrelated donor.

1.2 Acute graft versus host disease

1.2.1 Clinical appearance and epidemiology

Allo HSCT has been introduced as the only curative therapeutic avenue for many leukemia patients, and involves the beneficial GvL effect that is conveyed by alloreactive T cells. However, the frequently entailed complication known as graft versus host disease (GvHD) is brought about by the very same donor cell fraction and is the leading cause of morbidity and mortality among individuals having undergone allo BMT (Ferrara et al., 2009; Teshima et al., 2016; Zeiser & Blazar, 2017). Typically, acute GvHD (aGvHD) manifests within 100 days after transplant, but also delayed manifestations have been observed (Jagasia et al., 2015; Williams et al., 2021). The concomitant clinical appearance can be of aid in distinguishing acute from chronic GvHD (cGvHD), as mainly the skin, the intestine and the liver are the most frequently affected organs in the former (Holler et al., 2019; Justiz Vaillant et al., 2022). Hence, patients present with a maculopapular rash if the skin is involved. In the upper gastrointestinal tract (GIT), clinical manifestations may include nausea, anorexia or both while aGvHD affecting the lower GIT characteristically results in watery or bloody diarrhea and severe abdominal pain. When the liver is damaged, cholestasis can cause patients elevated serum levels of bilirubin. In contrast to aGvHD, cGvHD can affect a multitude of organs, such as the lungs, kidneys or the heart, besides the skin, the GIT and the liver, which are typically damaged by mechanisms different from aGvHD (Ferrara et al., 2009). Nevertheless, an overlap of acute and chronic GvHD symptoms is not uncommon (Jagasia et al., 2015; Williams et al., 2021).

As a severe and frequent complication, aGvHD occurs in 35%-45% of patients, even if a MSD is available (Ratanatharathorn et al., 1998; Verneris et al., 2015). When patients have to rely on single antigen-mismatched unrelated donors, the incidence is even higher, and when the respective data was stratified by minor *versus* major HLA-mismatch, no differences in the outcome could be detected (Wade et al., 2007). Depending on the origin of the graft, the five-year overall survival can range from ~30%-50% as observed in a cohort with GvHD being the main cause of death (Al-Kadhimi et al., 2014). In order to approximate the likelihood of transplant-related complications, risk prediction models may provide an evidence-based guideline (MacMillan et al., 2015). The key biological processes underlying the considerably complex pathogenesis of aGvHD will be discussed, hereinafter.

1.2.2 Pathophysiology

In advance, it seems worth emphasizing the primary role of the GIT as a target tissue of aGvHD, based on which the understanding of the pathogenesis will elaborated in the following (Hill & Ferrara, 2000). In general, the early events after allo HSCT causing aGvHD can be roughly divided into three phases. First, the host's tissues are damaged by the conditioning regimen, consisting of the single or combined treatment with radiation and chemotherapeutic agents. As a consequence, the tissue barrier integrity is compromised, and thus, microbiota can translocate from the colon's lumen to the epithelial layer and the lamina propria. Tissue-resident antigen-presenting cells (APCs), dislocated bacteria, epithelial cells and components of the extracellular matrix are disrupted by host conditioning. Consequently, the affected microenvironment becomes rich in danger-associated molecular patterns (DAMPs), such as adenosine triphosphate (ATP), pathogen-associated molecular patterns (PAMPs), such as lipopolysaccharide (LPS), tumor necrosis factor alpha (TNF-α) and various alarmins, such as interleukin (IL) 1 and IL-33. Supported by LPS-stimulated macrophages (MΦ), host APCs are activated, release further proinflammatory cytokines (e.g. IL-6, 12, 23) and accumulate MHC-II, costimulatory molecules, such as CD80/86 or CD40, as well as adhesion molecules (e.g. ICAM-1), on their surface (Ferrara et al., 2009; Hill & Koyama, 2020; Nassereddine et al., 2017).

Moreover, APCs abundantly present damage-associated host and microbial antigens *via* MHC class II to conventional $CD4^+$ donor T cells (Tcon) that recognize such as alloantigens, forming an immunological synapse between their TCR and costimulatory proteins (e.g. CD28) interacting with MHC-II and CD80/86 proteins on the APC cell surface, respectively. Consequently, in this second phase, Tcon are stimulated to proliferate, produce various cytokines (e.g. IL-2, IL-6) and differentiate into Th1 or Th17 cells, among other T helper cell subsets. Mainly driven by IL-2 and APC interaction, also $CD4^+CD25^+FOXP3^+$ regulatory T cells (Treg) are activated, proliferate and exert immunoregulatory functions by releasing antiinflammatory cytokines, such as IL-10 and TGF-β, and inhibitory surface molecules, such as the cytotoxic T-lymphocyte-associated protein 4 (CTLA-4), among other mechanisms (Ferrara et al., 2009; Nassereddine et al., 2017). The uptake of IL-33 *via* its receptor ST2 (ST2 with IL-1RACP) is another activation signal for Treg that is followed by the release of amphiregulin (AREG), a growth factor promoting tissue regeneration (Hill & Koyama, 2020). Of note, Treg can also drive Tcon to acquire a phenotype with suppressive functions, and thus become type 1 regulatory T (Tr1) cells (Hill & Koyama, 2020; Yamaguchi et al., 2011).

Despite the presence of Treg, the maleficent composition of donor $CD4^+$ T cells, releas-

ing pro-inflammatory cytokines, and attracted $CD8^+$ cytotoxic T-lymphocytes (CTLs) that directly target host tissue cells *via* Fas ligand (FASL) further drive inflammation and host cell apoptosis, respectively. The involvement of stimulated donor natural killer (NK) cells further aggravates host tissue damage. These events are involved in the third phase of aGvHD, the inflammatory effector phase, and it has long been established that the dysregulation of cytokines, i.e. the "cyotkine storm" is its key feature (Antin & Ferrara, 1992). The most important early events of aGvHD pathophysiology are summarized in **Figure 1.2**.

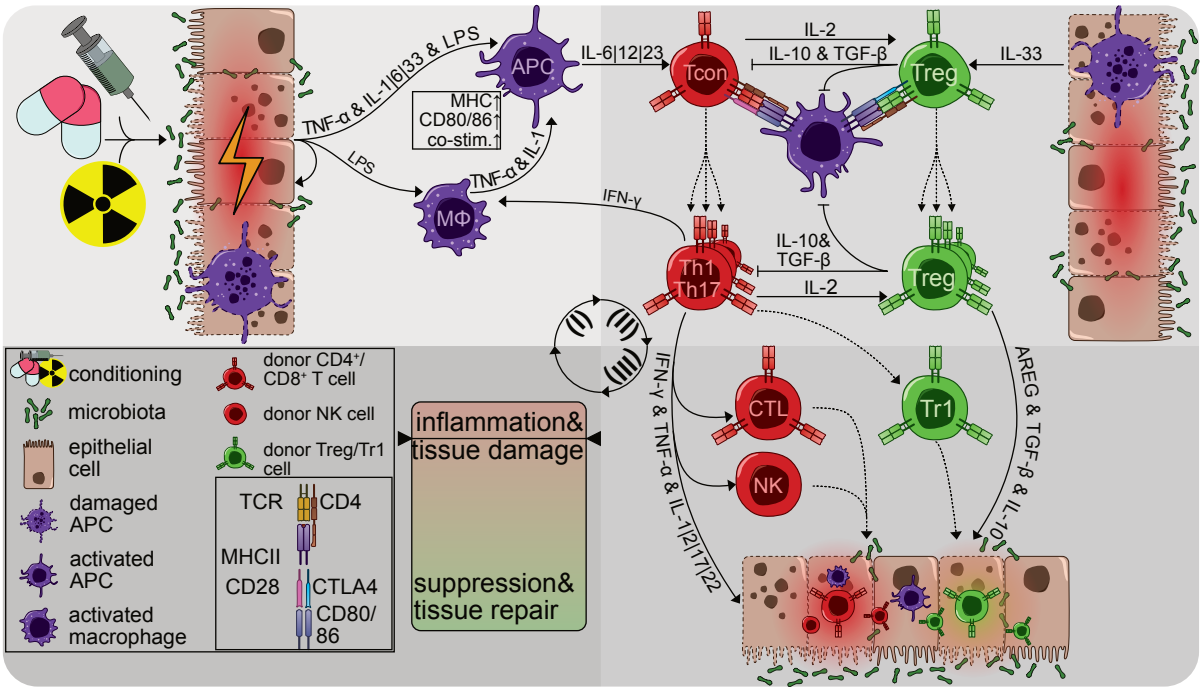

Figure 1.2: **Graphical summary of key processes in aGvHD pathophysiology.** (I) First phase: Host tissue damage by conditioning leads to host APC activation. Various danger signals, such as PAMPs, DAMPs and alarmins are released. (II) Second phase: Tcon are activated and proliferate, differentiate (e.g. Th1, Th17 cells) and release pro-inflammatory cytokines. Donor-type regulatory T cells are activated, i.a. by IL-2 or IL-33, and can exert suppressive function *via* inhibitory surface molecules (e.g. CTLA-4) or anti-inflammatory cytokines (e.g. TGF-β, IL-10), that also promote the conversion of Tcon to Tr1 cells. (III) Third phase: Tissue damage is further promoted by donor Tcon, CTLs and NK cells, while Treg are supporting tissue regeneration (e.g. *via* AREG) and dampen inflammation by various suppressive mechanisms. The imbalance between pro- and anti-inflammation results in aGvHD. The schematic is inspired by concepts from illustrations by Ferrara et al. (2009) and Hill & Koyama (2020) .

The comprehensive understanding of the pro- and anti-inflammatory events in aGvHD pathogenesis, with its numerous cell types, cell-contact-dependent and independent stimuli involved, is a challenging goal and has prompted experienced scientists in the field to embrace its complexity from various perspectives over the past few years (Fer-

rara et al., 2009; Hill & Koyama, 2020; Nassereddine et al., 2017; Teshima et al., 2016; Zeiser, 2019; Zeiser & Blazar, 2017). For the scope of this thesis, however, the involvement of Treg is of primary interest, hence, their most important immunoregulatory features will be adduced, hereinafter.

1.3 Regulatory T cells

1.3.1 Canonical immunoregulatory features

Treg are a specialized subset of $CD4^+$ T cells carrying the hallmark of a constitutively high expression of forkhead box protein P3 (FOXP3). As a master transcription factor (TF), the stable expression of FOXP3 is pivotal for peripheral tolerance conveyed by naturally arising Treg. Hence, if FOXP3 is deficient, affected individuals suffer from severe autoimmunity related complications which applies to humans and mice alike (Hori et al., 2003; Sakaguchi, 2004; Sharma & Rudra, 2018; Y. Zheng & Rudensky, 2007). Usually, Treg arise in the thymus and acquire FOXP3 expression upon high-affinity recognition of thymic self-antigens (Izcue et al., 2009; Picca et al., 2006). Especially in the intestine, regulatory lymphocytes exert exclusive functions in maintaining barrier integrity and tissue homeostasis (M. J. Barnes & Powrie, 2009; Coombes et al., 2005; Izcue et al., 2009). Of note, FOXP3 is not solely responsible for the differentiation and suppressor function of Treg. As demonstrated by Wu et al. (2006), the cooperation of NFAT and Treg plays a central role for Treg identity. Although typically expressed in Th1 T cells, T-bet expression in Treg enables them to specifically suppress Th17 and Th1 responses (Hoeppli et al., 2019; Koch et al., 2009). Another crucial transcription factor is FOXO1, in the absence of which Treg have shown hampered development with limited function (Kerdiles et al., 2010). Despite the known heterogeneity of Treg in mice and men, certain modes of action in immunosuppression are widely accepted.

In general, Treg employ multiple mechanisms to suppress inflammatory responses which are schematically depicted in **Figure 1.3**. One cell-contact-independent effector function relies on the secretion of effector molecules such as the antiinflammatory cytokines IL-10 and transforming growth factor beta 1 (TGF-β) by which Treg exert influence on the activation of effector T and B lymphocytes. In addition, IL-35, composed of EBI3 and the p35 subunit of IL-12, has been described as an antiinflammatory cytokine (Shevyrev & Tereshchenko, 2020). The growth factor AREG can be released by Treg to promote tissue repair in response to the alarmin IL-33 associating with its receptor suppression of tumorigenicity 2 (ST-2), and thus partly counteract tissue damage (Schiering et al., 2014; Zaiss et al., 2013). Moreover, the directly cytolytic granzyme (GZMB, GZMA) and perforin (PRF1) pathway is a means of targeting $CD4^+$ and $CD8^+$

effector T cells, among other cell types (Grossman et al., 2004).

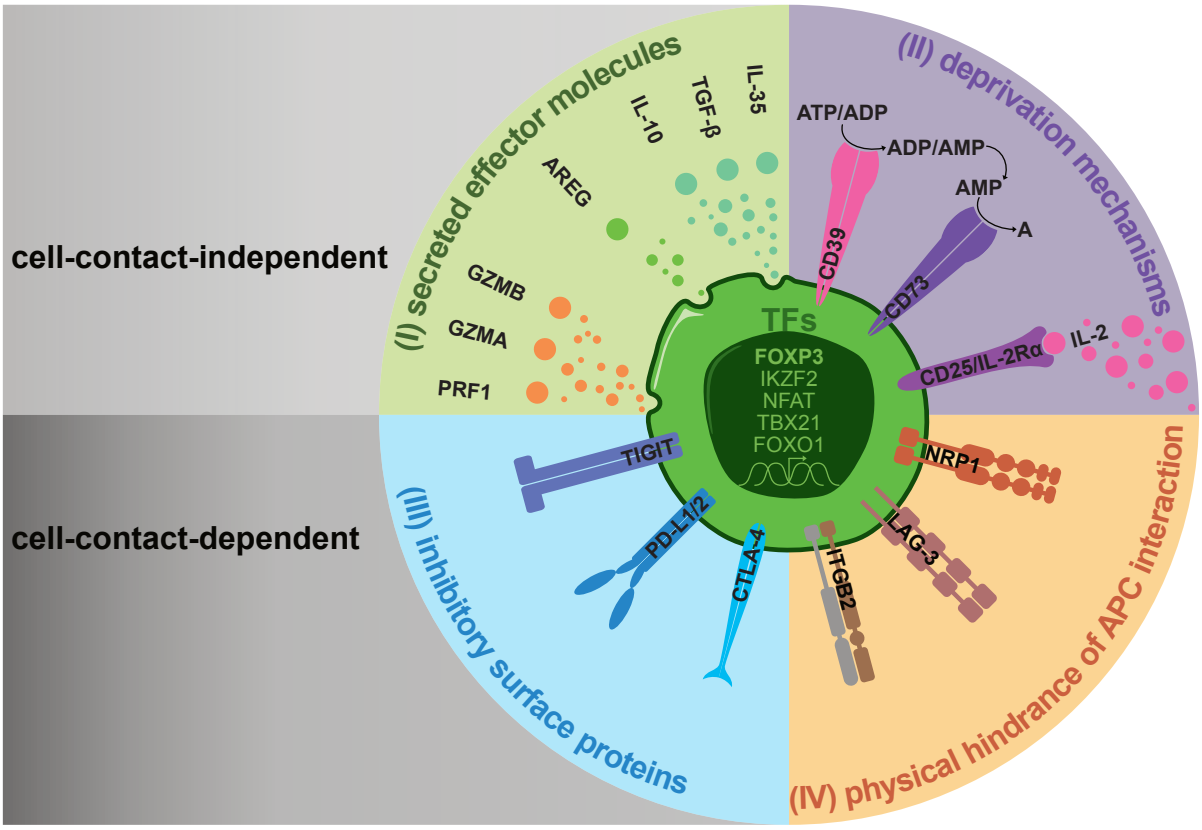

Figure 1.3: **Graphical summary of major suppressive Treg mechanisms.** (I) Secreted effector molecules and (II) deprivation mechanisms represent cell-contact-independent immunosuppressive functions. (III) Inhibitory surface proteins and (IV) physical deterrence of APC interaction by means of certain surface proteins comprise the mechanisms that are based on intercellular interaction. Essential TFs are listed in the schematic Treg nucleus.

Another immunosuppressive mechanism is the deprivation of factors that are key to the function of conventional $CD4^+$ or cytotoxic $CD8^+$ T cells at inflamed sites. Hence, the consumption of IL-2 *via* an interleukin receptor that contains the high affinity IL-2 receptor α (IL-2Rα/CD25) sub unit is a well-described principle (Yamaguchi et al., 2011). The catabolically active ectoenzymes CD39 and CD73 sequentially facilitate the degradation of pro-inflammatory Adenosine triphosphate (ATP), limiting the metabolic potential that can be harnessed by effector T cells (Antonioli et al., 2013).

A third layer of immunoregulation is facilitated by inhibitory surface molecules that act in a cell-contact-dependent manner, such as CTLA-4 which has a higher affinity towards APC costimulatory CD80/86 than the CD28 molecule, and thus, interferes with the APC-Tcon-interaction axis. Its inhibitory role has also been implicated in the direct uptake and degradation of its targets (Qureshi et al., 2011). Additionally, TIGIT typically promotes IL-10 production while reducing IL-12 secretion upon binding of CD155 on

DCs (Yu et al., 2009). The respective gene is a direct target of the FOXO1 transcription factor which is essential for Treg function *in vivo* (Kerdiles et al., 2010). By the interaction of programmed cell death ligand 1/2 (PD-L1/L2) on the surface of Treg with their respective targets PD-1/2 on activated Tcon, the latter can be rendered anergic or apoptotic (Deng et al., 2015; Shevyrev & Tereshchenko, 2020). Furthermore, PD-L1 interaction with host APC CD80 assures natural donor Treg expansion (Cassady et al., 2018). The upregulation of PD-1 on effector T cell surfaces in the aGvHD microenvironment stresses the importance of Treg PD-L1 conveyed anergy which is, however, insufficient to prevent aGvHD development (Hossain et al., 2017).

Lastly, Treg can directly deter the physical contact of naïve $CD4^+$ T cells with APCs by outcompeting them. To this end, several adhesion surface proteins can be highly expressed. The CD4 homolg LAG-3, for instance, has a high affinity towards MHC-II on APCs, thus allowing them to prevail over Tcon (Graydon et al., 2021). Besides LAG-3, Neuropilin 1 (NRP1) as well as various heterodimeric integrins, containing ITGB2 (e.g. LFA-1), can render the surface of APCs inaccessible for other T cell subsets. For a prolonged interaction between Treg and DCs during antigen recognition, NRP1 has been shown to be crucial (Mizui & Kikutani, 2008; Sarris et al., 2008). Furthermore, the LFA-1 surface antigen promotes the establishment and maintenance of the Treg interaction with APCs *via* ICAM-1 in which Treg can form aggregates around and shield APCs from cross-talk with Tcon (Onishi et al., 2008). Besides the function of Treg, their development also relies on the presence of LFA-1 (Marski et al., 2005).

1.3.2 Treg heterogeneity in homeostasis and inflammation

While the core features of Treg are well-known, the heterogeneity of Treg is also commonly accepted. More recent studies have made efforts to resolve the multitude of Treg phenotypes into the understanding of defined subpopulations, differing in their state of activation, TCR repertoire, microenvironment and concomitant gene expression patterns in unperturbed murine models (Miragaia et al., 2019; Zemmour et al., 2018). Depending on their gene expression modules, Treg can be roughly subdivided into natural, activated, central memory and effector Treg (Shevyrev & Tereshchenko, 2020). Although most essential genes and pathways are conserved between humans and mice, markers are not generally interchangeable which is why Treg heterogeneity in human peripheral blood has to be addressed separately (Wing et al., 2019). Especially FOXP3 positivity is not as reliable in humans as it is in murine Treg for lineage identification (J. Wang et al., 2007).

It is known that the TCR repertoires of Tcon and Treg usually show no overlap and thus, thymic and peripherally induced Treg (iTreg) differ in their TCR repertoire, with

the former displaying high affinity towards self-antigens, while the latter recognize foreign antigens with high affinity (Pacholczyk et al., 2006). In line with that, iTreg high affinity recognition of foreign antigens determines their function, which is distinct from that of thymic Treg (Wyss et al., 2016). While high affinity TCR harboring Treg have been described to mainly suppress inflammation by secretion of IL-10 and interaction with APCs by means of CTLA-4, low affinity TCR Treg are more prone towards releasing AREG to promote tissue homeostasis and regeneration (Sharma & Rudra, 2018; Shevyrev & Tereshchenko, 2020). Additionally, Delacher et al. (2017) described a special subset of Treg important for tissue homeostasis, characterized by high expression levels of *ST-2*, killer cell lectin like receptor G1 (*KLRG-1*) and *AREG*. A granular and unbiased approach by Miragaia et al. (2019) revealed non-lymphoid tissue-like (NLT) and lymphoid tissue-like (LT) Treg subpopulations present in the colon of unperturbed mice. The aforementioned signatures also served in characterizing colon-homed Treg in the present work. Furthermore, it has been shown that some gene expression modules can be shared between Treg and Tcon which, depending on their respective state of activation, can make a clear distinction challenging (Shevyrev & Tereshchenko, 2020; Zemmour et al., 2018).

Due to the abundance of activating stimuli in acutely inflamed sites, the usage of suppressive functions can undergo a shift from the deprivation of co-stimuli under homeostatic conditions towards the utilization of cytolysis and abundant release of antiinflammatory cytokines. By means of IL-10, IL-35 or TGF-β secretion, Treg can skew naïve Tcon towards differentiation to Tr1, Tr35 or iTreg, closing an amplified antiinflammatory positive feedback loop (Yamaguchi et al., 2011). Importantly, at the same time Tr1 cells lack key Treg features, such as FOXP3 expression (Song et al., 2021). Treg are essential in the modulation of aGvHD pathogenesis, with their diversified suppressive mechanisms and their potential to impose suppressive activity upon naïve Tcon or APCs in inflamed microenvironments. However, their efficacy is very often not sufficient to secure patient survival.

1.4 aGvHD - clinical challenge in need of novel treatment options

Obviously, aGvHD necessitates an adjuvant treatment to improve allo HSCT outcomes, as physiological mechanisms of immune regulation frequently fail. Several pharmacological lines of treatment are available and comprise strategies for the prevention as well as the amelioration of ongoing GvHD. A long-established GvHD prophylaxis approach is based on the combination of the calcineurin inhibitor cyclosporine with

methotrexate (Storb et al., 1986). For higher-grade GvHD, tacrolimus could be established as the preferred calcineurin inhibitor in MSD and unrelated donor transplants (Nash et al., 2000; Ratanatharathorn et al., 1998). And also less cytotoxic immunosuppressants, such as mycophenolate mofetil have proven a valid alternative to methotrexate (Bolwell et al., 2004). As far as the treatment of aGvHD is concerned, high-dose corticosteroid therapy most often represents the first line of treatment. In recent years, advances have expanded therapeutic options dramatically. Regimens include strategies of pre-treatment for risk minimization and various combinations of therapeutic agents. Currently, the EBMT recommends the guideline provided by Penack et al. (2020) for the management of aGvHD patients. Of note, the treatment of children requires additional considerations, and hence, e.g. in childhood ALL patients are cared for accordingly (Wölfl et al., 2022).

However, very often limiting toxicities outweigh the expected therapeutic benefit for the patient or pharmacological treatment is contraindicated. Especially for steroid refractory forms of aGvHD, newly emerging therapeutic avenues, such as immunological or cell-based therapies, have gained in urgency (Malard et al., 2020; Shapiro & Antin, 2020). The extracorporeal photochemotherapy (ECP) is a common immunological treatment, for instance, that results in the increase in peripheral and functional Treg in GvHD patients, and thus, indirectly harnesses Treg as therapeutic agent (Biagi et al., 2007). Among the cellular approaches, BM-derived myeloid suppressor cells (MSCs) can be listed as the very first approved therapy (Daly, 2012). In addition, preclinical studies have shown that donor-type $CD4^+CD25^+$ donor Treg are able to prevent aGvHD without impairing the GvL effect of alloreactive $CD4^+CD25^-$ donor Tcon (Edinger et al., 2003; Hoffmann et al., 2002). In a murine model of haplo HSCT, Treg were also efficacious in the treatment of ongoing GvHD (Riegel et al., 2020). In high-risk malignancy patients the co-infusion of Tcon has to be considered to minimize the risk of relapse. The concomitant elevated risk of developing aGvHD has been successfully counteracted by the administration of donor Treg (Di Ianni et al., 2011). The tissue homeostasis capacity of Treg in combination with their suppressive potential render them a promising therapy option to minimize transplant-related mortality (Ikegawa & Matsuoka, 2021). Recent advances in genome editing have, moreover, enabled the enhancement of Treg cell products for autoimmune diseases and transplant-related inflammation (Ferreira et al., 2019; Guo et al., 2021). By now, a substantial number of clinical trials in phases I, II or III that involve Treg as therapeutic agent are ongoing and, thereby demonstrate the enormous potential of Treg in aGvHD prevention and therapy (Hefazi et al., 2021).

1.4.1 Prerequisites for aGvHD treatment efficacy with Treg

Due to their potential to exert a multilayered set of functions for the regulation of inflammatory processes, Treg have proven effective for the treatment and prophylaxis of experimental aGvHD. However, the prerequisites for their clinical efficacy are not fully understood. The capacity to infiltrate target organs is one self-evident requirement. Also, the acquisition of a site-specific and suppressive gene expression program is likely one of the major determinants for Treg potency. Especially the role of the TCR repertoire is a pressing unanswered question. Most likely, the foundation of the clinical efficacy of Treg in preventing or ameliorating aGvHD is a complex interplay of their migratory potential, environmentally induced gene expression and clonal evolution.

The initial homing to secondary lymphoid organs, followed by proliferation and subsequent dissemination in GvHD target organs, such as colon, liver and skin characterize the spatiotemporal dynamics of $CD4^+$ allogeneic T cells (Beilhack et al., 2005). For the protective effect in the context of transplant-related alloreactivity, lymph node homing capacity has been shown to be essential for Treg clinical efficacy, as only $CD62L^+$ (selectin L, SELL) are able to protect mice from lethal GvHD (Ermann et al., 2005; Taylor et al., 2004). A variety of chemokine receptors and integrins, such as CCR9 and ITGAE (associating with ITGB7), capacitate Treg to migrate to the mucosa of barrier tissues (Agace et al., n.d.; Shevyrev & Tereshchenko, 2020). More specifically, CCR4 and CCR7 allow the infiltration of the inflamed intestine (Smigiel et al., 2013). The increase in CXCR3 levels governed by T-bet, enabled Treg to accumulate at foci of type 1 inflammation (Koch et al., 2009). Adapted from that, it has been shown that the very same homing capacity of Treg can be modified to enable the selective suppression of Th-1-inflamed sites (Hoeppli et al., 2019). Also, the sustained survival of Treg in sites of inflammation depends on the CCR8-mediated interaction with $CD11c^+$ donor APCs, as shown in a murine model of GvHD (Coghill et al., 2013).

Besides their migratory properties, certain functional pathways have emerged as fundamental for the therapeutic potential of Treg. For instance, enhancing Treg by conjugating high amounts of FasL on the cell surface enabled the resulting “killer” Treg to alleviate GvHD by direct killing of CTLs (Yolcu et al., 2013). The identification of relative GZMA deficiency in Treg stemming from GvHD patients has led to a haplo HSCT $Gzma^{-/-}$ mouse study, which identified GZMA as a key player for GvHD prevention (Ukena et al., 2011; Velaga et al., 2015).

As a general rule, a broad and unbiased TCR repertoire is considered beneficial (Wing & Sakaguchi, 2011). It has been shown that a high diversity in the Treg population assures their adequate function (Föhse et al., 2011). However, as pointed out in section 1.3.2, besides their target organ, the phenotype of Treg can be influenced by their TCR

(Zemmour et al., 2018). In line with this, low affinity-TCR Treg have been found to display enhanced suppressive potential (Sprouse et al., 2018). The issue, however, whether a Treg pre-selection skewed towards radiation-induced host antigens might pose an advantage in proliferation over naïve Tcon with an unbiased TCR repertoire, is an issue that so far has not been addressed by an adequate comparative study.

1.5 Objective - unraveling Treg efficacy in aGvHD

Treg are important regulators of the immune system and can present with a variety of phenotypes, some of which are associated with tissue homeostasis while other subtypes are specific for inflamed microenvironments. The requirements that have to be met for Treg to be clinically efficacious in the prevention or amelioration of acute GvHD, however, are still incompletely understood.

Hence, one goal of the present work was the detailed characterization of alloantigen-specifically ("allo") and polyclonally ("poly") *in vitro* expanded Treg prior to transplantation. Furthermore, Treg homing to colon, liver and spleen were investigated to gain more insight on the biological processes Treg undergo after organ infiltration. The adaption of gene expression patterns and the dynamics of TCR clonality under homeostatic as well as pathological conditions were of primary interest. Inquiring about the potential causality between clonotype and organ infiltration was aimed to be addressed by assessing Treg TCR chain pairing on the sing-cell level as an additional layer of evidence. The ultimate goal was the comprehensive understanding of Treg-specific mechanisms that could be harnessed to generate engineered Treg with enhanced potential for aGvHD prevention and therapy.

CHAPTER 2

Materials and Methods

2.1 Materials

All materials used to generate the presented results are listed and categorized by cusotm oligonucleotides, molecular biology kits, chemicals and reagents, technical devices and consumables. Furthermore, all relevant mouse strains are described.

2.1.1 Mice

Female BALB/c ($H2^d$) and C57BL/6 ($H2^b$) wild type mice were purchased from Charles River Laboratories (Sulzbach, Germany), congenic B6.SJL-Ptprca Pepcb/BoyJ (CD45.1; $H2^b$; here named 'CD45.1$^+$') and transgenic FoxP3EGFP mice (B6; CD45.2; $H2^b$; here named 'Foxp3^{gfp}' kindly provided by B. Malissen; Ref: Y. Wang et al. (2008); PubMed AccNr: 18209052) were bred in-house. Donors were 8–12 weeks, recipients 11–16 weeks old at the time of BMT. Mice were held under specific pathogen-free conditions and studies were approved by the Committee on Ethics of Animal Experiments at the Bavarian Government (Ref-No: 55.2-2532-2-430). All mouse-related work was facilitated by members of WG E/H.

2.1.2 Custom Oligonucleotides

All custom oligonucleotides, such as indexed primers and the template-switching oligonucleotide (TSO) are described in Table 2.1 and were ordered from Integrated DNA

Techonologies (IDT, Coralville, Iowa, USA).

Table 2.1: List of custom oligonucleotides with 5'→3' sequence and Illumina index designation. 'Bio-': 5' biotinylated; 'U': deoxyuridine; 'rG': ribo-Guanosin.

Name	Species	Chain	Type	i5/i7 index	Sequence 5'→3'
5'CDS Primer	universal	universal	cDNA synthesis	-	TTTTTTTTTTTTTTTTTTTTTTTTTVN
TCR_UMI_TSO	universal	universal	cDNA synthesis	-	Bio-GUCUCGUGGGCUCGGAGAUGUGUAUAAGAGACAGUNNNNUNNNNUNNNNUrGrGrG
TCR_1.PCR_for	universal	universal	1st PCR	-	GTCTCGTGGGCTCGGAGATGTG
GSP_TRA_rev	Mm	α	1st PCR	-	GTCGGTGAACAGGCAGAGGGT
GSP_TRB_rev	Mm	β	1st PCR	-	GAAGCCCCTGGCCAAGCACACGAG
Primer_i7 (for)	universal	universal	2nd PCR	701-729	CAAGCAGAAGACGGCATACGAGATXXXXXXXXGTCTCGTGGGCTCGG
mTCRA_rev_501	Mm	α	2nd PCR	501	AATGATACGGCGACCACCGAGATCTACACTAGATCGCTCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGATCTTTTAACTGGTACACAGCA
mTCRA_rev_502	Mm	α	2nd PCR	502	AATGATACGGCGACCACCGAGATCTACACCTCTCTATTCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGATCTTTTAACTGGTACACAGCA
mTCRA_rev_503	Mm	α	2nd PCR	503	AATGATACGGCGACCACCGAGATCTACACTATCCTCTTCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGATCTTTTAACTGGTACACAGCA
mTCRA_rev_508	Mm	α	2nd PCR	508	AATGATACGGCGACCACCGAGATCTACACCTAAGCCTTCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGATCTTTTAACTGGTACACAGCA
mTCRB_rev_501	Mm	β	2nd PCR	501	AATGATACGGCGACCACCGAGATCTACACTAGATCGCTCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTTGATGGCTCAAACAAGGAG
mTCRB_rev_502	Mm	β	2nd PCR	502	AATGATACGGCGACCACCGAGATCTACACCTCTCTATTCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTTGATGGCTCAAACAAGGAG
mTCRB_rev_517	Mm	β	2nd PCR	517	AATGATACGGCGACCACCGAGATCTACACGCGTAAGATCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTTGATGGCTCAAACAAGGAG
mTCRB_rev_520	Mm	β	2nd PCR	520	AATGATACGGCGACCACCGAGATCTACACAAGGCTATTCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTTGATGGCTCAAACAAGGAG

2.1.3 Molecular Biology Kits

Commercially available molecular biology kits that were used to conduct the presented experiments are listed in Table 2.2.

Table 2.2: List of commercially available molecular biology kits acquired from different manufacturers.

Kit	Supplier/Manufacturer
Chromium Next GEM Chip G Single Cell Kit	10x Genomics, Pleasanton, California, USA
Chromium Next GEM Single Cell 5′ Library and Gel Bead Kit v1.1	10x Genomics, Pleasanton, California, USA
Chromium Single Cell 5’ Library Construction Kit	10x Genomics, Pleasanton, California, USA
Chromium Single Cell V(D)J Enrichment Kit, Mouse T Cell	10x Genomics, Pleasanton, California, USA
Chromium i7 Multiplex Kit = Single Index Kit T Set A	10x Genomics, Pleasanton, California, USA
Agencourt® AMPure® XP- Kit	Beckman Coulter, Brea, USA
Nextera XT Index Kit v2 (i7)	Illumina, Inc., San Diego, California, USA
Uracil-DNA Glycosylase (UDG)	New England Biolabs (NEB), Ipswich, Massachusetts, USA
RNeasy® Mini Kit	Qiagen, Hilden, Germany
RNeasy® Micro Kit	Qiagen, Hilden, Germany
KAPA Library Quantification Kit, Complete kit (Universal)	Roche, Basel, Switzerland
SMART-Seq® Stranded Kit	Takara Bio Inc., Kusatsu, Shiga, Japan
Advantage® 2 PCR Kit	Takara Bio Inc., Kusatsu, Shiga, Japan
SMARTScribe™ Reverse Transcriptase	Takara Bio Inc., Kusatsu, Shiga, Japan
Qubit™ dsDNA HS Assay Kit	Thermo Fisher Scientific, Waltham, USA

2.1.4 Chemicals and Reagents

All chemicals and reagents were purchased from Sigma-Aldirch (St. Louis, USA) or Merck Millipore (Burlington, USA) if not indicated otherwise. See Table 2.3 for exceptions.

Table 2.3: List of chemicals and reagents acquired from different manufacturers.

Product	Supplier/Manufacturer
Aqua (1000 ml)	B. Braun Melsungen, Melsungen, Germany
Ethanol ≥99,5 %, Ph.Eur., reinst	Carl Roth, Karlsruhe, Germany
Monarch® DNA Elution Buffer	New England Biolabs (NEB), Ipswich, Massachusetts, USA
Desoxynucleosid-Triphosphat-Set	Roche, Basel, Switzerland
Gibco™ Distilled Water (RNase/DNase free)	Thermo Fisher Scientific, Waltham, USA
Dithioreitol (DTT), 0.1 M solution	Thermo Fisher Scientific, Waltham, USA

2.1.5 Technical Devices

A list of all technical devices that were directly operated for the generation of the present results is shown in Table 2.4.

Table 2.4: Technical devices are listed next to their supplier or manufacturer.

Device	Supplier/Manufacturer
Chromium Controller	10x Genomics, Pleasanton, California, USA
10x magnetic rack	10x Genomics, Pleasanton, California, USA
Agilent 2200 Tape Station	Agilent Technologies, Santa Clara, USA
CFX96™ qPCR system	Bio-Rad Laboratories, California, USA
Rotilabo-mini-centrifuge	Carl Roth, Karlsruhe, Germany
Intelli-Mixer RM-2L	Elmi-Tech, Riga, Latvia
Centrifuge 5804 R	Eppendorf, Hamburg, Germany
Eppendorf Research (plus) (2.5 µl/10 µl/100 µl/1000 µl)	Eppendorf, Hamburg, Germany
Heat sealer	Eppendorf, Hamburg, Germany
Mastercycler ep realplex	Eppendorf, Hamburg, Germany
Multipipette plus	Eppendorf, Hamburg, Germany
PCR Mastercycler nexus X2	Eppendorf, Hamburg, Germany
Thermomixer	Eppendorf, Hamburg, Germany

Table 2.4: Technical devices are listed next to their supplier or manufacturer. *(continued)*

Device	Supplier/Manufacturer
Mastercycler® nexus	Eppendorf, Hamburg, Germany
PIPETMAN Classic (10 µl/20 µl/200 µl/1000 µl)	Gilson, Midleton, USA
Thermo Cell	Hangzhou Bioer Co. Ltd., Hangzhou, China
SPROUT® MINI CENTRIFUGE 12V	Heathrow Scientific, IL, USA
Biofuge Fresco	Heraeus, Hanau, Germany
Biofuge Pico	Heraeus, Hanau, Germany
Herasafe Laminar air flow cabinet	Heraeus, Hanau, Germany
Megafuge 3.0R	Heraeus, Hanau, Germany
Multifuge 3S-R	Heraeus, Osterode Germany
IKA MS 3 Vortexer	IKA-Werke, Staufen Germany
MiSeq	Illumina, Inc., San Diego, California, USA
NextSeq 550	Illumina, Inc., San Diego, California, USA
10x magnetic rack replica	in-house production, 3D print
Pipetboy	Integra Biosciences, Zizers, Switzerland
UVP PCR HEPA Cabinet	Jena Bioscience, Jena, Germany
UVP PCR Cabinet	Jena Bioscience, Jena, Germany
Waterbath	Julabo, Seelstadt, Germany
Qubit 2.0 Fluorometer	Life Technologies, Carlsbad, USA
Pipet-Lite Multi Pipette L8-300XLS+	Mettler-Toledo, Giessen, Germany
Water purification system	Milipore, Eschborn, Germany
PCR Thermocycler PTC-200	MJ-Research; Biometra, Göttingen, Germany
NanoDrop 1000	PeqLab, Erlangen, Germany
Refrigerated SIGMA 2-Sartorius Centrifuge	Sartorius, Göttingen, Germany
Vortex-Genie 2	Scientific Industries, New York, USA
Dynabeads™ MPC™-S (Magnetic Particle Concentrator)	Thermo Fisher Scientific, Waltham, USA

Table 2.4: Technical devices are listed next to their supplier or manufacturer. *(continued)*

Device	Supplier/Manufacturer
ZX4 Advanced IR Vortex Mixer	VELP Scientifica, Usmate, Italy

2.1.6 Consumables

A list of all consumables that were used to conduct the presented experiments is provided in Table 2.5.

Table 2.5: List of consumable materials used for experiments are specified next to the supplier or manufacturer.

Product	Supplier/Manufacturer
(High Sensitivity) D1000 ScreenTape®	Agilent Technologies, Santa Clara, USA
(High Sensitivity) D1000 ScreenTape Ladder	Agilent Technologies, Santa Clara, USA
(High Sensitivity) D1000 ScreenTape Sample buffer	Agilent Technologies, Santa Clara, USA
High Sensitivity D5000 ScreenTape®	Agilent Technologies, Santa Clara, USA
High Sensitivity D5000 ScreenTape Ladder	Agilent Technologies, Santa Clara, USA
High Sensitivity D5000 ScreenTape Sample buffer	Agilent Technologies, Santa Clara, USA
Optical cap 8x strip	Agilent Technologies, Santa Clara, USA
Optical tube 8x strip	Agilent Technologies, Santa Clara, USA
(High Sensitivity) RNA ScreenTape Ladder	Agilent Technologies, Santa Clara, USA
(High Sensitivity) RNA ScreenTape Sample buffer	Agilent Technologies, Santa Clara, USA
(High Sensitivity) RNA ScreenTape®	Agilent Technologies, Santa Clara, USA
NeoTouch Premium Disposable Neoprene Gloves	Ansell, New Jersey, USA
TouchNTuff Disposable Gloves	Ansell, New Jersey, USA
VersaTouch Disposable Nitrile Gloves	Ansell, New Jersey, USA
AMPure XP Magnetic beads	Beckman Coulter, Brea, USA

Table 2.5: List of consumable materials used for experiments are specified next to the supplier or manufacturer. *(continued)*

Product	Supplier/Manufacturer
Hard-Shell® 96-Well PCR Plates, low profile, thin wall, skirted, clear	Bio-Rad Laboratories, California, USA
Microseal 'B' PCR Plate Sealing Film, adhesive, optical	Bio-Rad Laboratories, California, USA
SafeSeal SurPhob filter tips (10 µl/ 100 µl/ 1250 µl)	Biozym Scientific, Hessisch Oldendorf, Germany
Falcon tubes (200 ml/ 50 ml/ 15 ml)	Corning, New York, USA
Cell culture pipettes	Costar, Cambridge, USA / nerbe plus, Winsen (Luhe), Germany
Incidin OxyWipe S	Ecolab Deutschland, Monheim am Rhein, Germany
Combitips advanced (0.5 ml/ 2.5 ml/ 5.0 ml)	Eppendorf, Hamburg, Germany
DNA LoBind Tubes, DNA LoBind, 1.5 mL	Eppendorf, Hamburg, Germany
Eppendorf Safe-Lock Tubes, 0.5 mL	Eppendorf, Hamburg, Germany
Eppendorf Tubes 5.0 ml	Eppendorf, Hamburg, Germany
epT.I.P.S.® LoRetention Reloads, Eppendorf Quality™, 2 – 200 µl/ 50 – 1000 µL	Eppendorf, Hamburg, Germany
Heatsealing Film	Eppendorf, Hamburg, Germany
PCR Plate 96, semi-skirted	Eppendorf, Hamburg, Germany
Disposable scalpel	Feather, Osaka, Japan
MiSeq Reagent Kit v3 (600-cycle)	Illumina, Inc., San Diego, California, USA
NextSeq 500/550 High Output Kit v2 (75 Cycles)	Illumina, Inc., San Diego, California, USA
NextSeq 500/550 High Output Kit v2 (150 Cycles)	Illumina, Inc., San Diego, California, USA
Tips RT-LTS-A-300µL-/F/L-768/8	Mettler-Toledo, Giessen, Germany
Cryo tubes	Nunc, Wiesbaden, Germany
Parafilm	Pechiney Plastic Packaging, Chicago, USA

Table 2.5: List of consumable materials used for experiments are specified next to the supplier or manufacturer. *(continued)*

Product	Supplier/Manufacturer
Micro tube 1.5ml/ 0.5 ml DNA LowBind	Sarstedt, Nümrecht, Germany
Pipet tips 10 µl	Sarstedt, Nümrecht, Germany
Safe Seal Tube 2.0 ml/ 1.5 ml	Sarstedt, Nümrecht, Germany

2.2 Laboratory Methods

Hereafter, all wet-lab methods directly leading to the present results are thoroughly described. Methods facilitated by members of the co-operating work group are briefly summarized.

2.2.1 Treg *in vitro* expansion

Expansion of Treg and subsequent FACS analyses were facilitated by members of WG E/H according to established standard operating procedures (SOPs). In brief, freshly isolated C57BL/6 Foxp3^{gfp} (H2^{b}) donor Treg were cultured for 11-14 d in the presence of 2000 U/mL of recombinant human IL-2 (rh-IL2). Allo Treg cultures were supplemented with 30 Gy irradiated splenic DCs from BALB/c (H2^{d}) mice, and poly Treg were stimulated with conventional anti-CD3/CD28 beads.

2.2.2 Treg derived from *in vivo* models

All transplant experiments, Treg re-isolation from colon, liver and spleen together with successive FACS analyses were performed by staff of WG E/H. In short, homeostatic Treg were derived from pools of 2 to 5 healthy CD45.2^{+} C57BL/6 Foxp3^{gfp} (H2^{b}, Treg donors) mice separating CD62L^{+} from CD62L^{-} cells. Each wild type BALB/c (H2^{d}) recipient in the BMT control model was conditioned by total body irradiation (TBI, 8 Gy) and received 2.5x10^{6} T cell-depleted bone marrow (TCD BM) cells from congenic CD45.1^{+} C57BL/6 (H2^{b}) BM donors along with either 1.0x10^{6} *in vitro* expanded (11-13 d) allo or poly Treg at d0. Recipients in the prophylaxis model additionally received CD4^{+}CD25^{-} conventional T cells (Tcon) from BM donors at a Treg:Tcon ratio of 1:1 on d0. In contrast to the prophylaxis approach, recipients in the therapy model received either allo or poly Treg 11 d after TCD BM and Tcon administration at which the number of infused Treg was equal to the initially transplanted Tcon.

2.2.3 RNA preparation and QC

Total cellular RNA was isolated from Treg (freshly isolated, *in vitro* expanded, or re-isolated after BMT) obtained as homogenized RLT lysates at -20°C, using the RNeasy Mini Kit or the RNeasy Micro Kit from Qiagen (Hilden, Germany). Following the manufacturer's recommendations, samples originating from 500 to 500x10^3 cells were processed with the Micro Kit whereas the Mini Kit was applied to all samples exceeding 500x10^3 cells. RNA concentration and quality were simultaneously assessed using either RNA Screentape Analysis or High Sensitivity RNA Screentape Analysis from Agilent Technologies (Santa Clara, USA), depending on the expected yield. After quality control (QC), samples were thoroughly labeled and stored at -80°C for downstream methods. For samples passing the sensitivity threshold for the calculation of the RNA integrity index (RIN), the respective values are summarized in Tables B.1, B.2, B.3, B.4 and B.5.

2.2.4 Bulk RNA-seq library preparation, QC and sequencing

Total cellular RNA was obtained as described in section 2.2.3. Preparation of dsDNA libraries for Illumina sequencing was facilitated using the SMART-Seq Stranded Kit (Takara Bio Inc., Kusatsu, Shiga, Japan) according to the manufacturer's instructions. The composition of the 10x lysis buffer, shearing mix, first-strand synthesis mix, PCR1 mix, scZapR mix and PCR2 mix are provided in Tables 2.6, 2.7, 2.8, 2.9, 2.10 and 2.11, respectively. Depending on sample quality and abundance, fragmentation times and second PCR cycle numbers were adjusted, respectively. For some exceptionally scarce samples, first PCR products were pooled to minimize carry-over losses (Workflow "pp_ulow") as opposed to singularly processed samples (Workflow "sp_low"). After final cleanup using AMPure XP beads (Beckman Coulter, Brea, USA), the concentration of dsDNA libraries was measured with the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, USA). DNA fragment size distribution was assessed using the High Sensitivity D1000 ScreenTape Assay from Agilent Technologies (Santa Clara, USA). Libraries were stored at -20°C until pooling, and library pool concentration and fragment size distribution were assessed as in the quality control of individual libraries. Libraries with concentrations ranging from 0.5-75 ng/µL were permitted for sequencing on the NextSeq550 system (Illumina, Inc., San Diego, California, USA) following the manufacturer's guidelines for denaturing, dilution and loading of library pools. All bulk RNA-seq library metrics and metadata are summarized in Tables B.1, B.2, B.3, B.4 and B.5.

Table 2.6: 10x lysis mix according to manufacturer's instructions.

Component	Volume [µL]
10X Lysis Buffer	19
RNase Inhibitor	1
Total volume	20

Table 2.7: Shearing mix according to manufacturer's instructions.

Component	Volume/RXN [µL]
10X Lysis Mix	1
SMART scN6	1
scRT Buffer	4
Total volume	6

Table 2.8: First-strand synthesis mix according to manufacturer's instructions.

Component	Volume/RXN [µL]
SMART scTSO Mix	4.5
RNase Inhibitor	0.5
SMARTScribe RT	2.0
Total volume	7.0

Table 2.9: PCR1 reaction mix according to manufacturer's instructions.

Component	Volume/RXN [µL]
Nuclease-Free Water	2
SeqAmp CB PCR Buffer (2X)	25
SeqAmp DNA Polymerase	1
3' index (i5)	1
Total volume	29

Table 2.10: scZapR mix according to manufacturer's instructions.

Component	Volume/RXN [µL]
Nuclease-Free Water	16.8
10X ZapR Buffer	2.2
scZapR	1.5
scR-Probes	1.5
Total volume	22.0

Table 2.11: PCR2 reaction mix according to manufacturer's instructions.

Component	Volume/RXN [µL]
Nuclease-Free Water	26
SeqAmp CB PCR Buffer	50
PCR2 Primers	2
SeqAmp DNA Polymerase	2
Total volume	80

2.2.5 Bulk TCRrep-seq library preparation and QC

Total cellular RNA was obtained as described in section 2.2.3. TCR repertoire sequencing (TCRrep-seq) libraries for Illumina sequencing were prepared on the basis of (Mamedov et al., 2013) and the protocol was adapted, as follows. For the generation of 5'RACE-ready cDNA, total RNA ranging from 10 - 200 ng (low input) or 200 - 1000 ng (high input) in a total of 10 µl RNAse-free water was combined with 1 µL of 12 µM or 25 µM 5'-CDS Primer in a total volume of 11 µL, respectively (see Table 2.12). Reactions were incubated at 72°C for 3 min followed by 42°C for 2 min for poly-dT primer annealing. Next, 1 µL of 10 µM (low input) or 50 µM (high input) smarter template-switching oligonucleotide (TCR_UMI_TSO) was added at room temperature (RT). Reverse Transcription Buffer Mix (RTBM) consisting of 5x First-Strand-Buffer Mix (Takara Bio Inc., Kusatsu, Shiga, Japan), 20 mM dNTPs (Roche, Basel, Switzerland), 100 mM DTT (Thermo Fisher Scientific, Waltham, USA) was freshly prepared on ice (see Table 2.13) before it was supplemented with RNaseOUT Recombinant Ribonuclease Inhibitor (40 U/µL) (Thermo Fisher Scientific, Waltham, USA) and SMARTScribe Reverse Transcriptase (100 U/µL) (Takara Bio Inc., Kusatsu, Shiga, Japan) at RT (see Table 2.14. A volume of 8 µL of the resulting reverse transcription master mix was added to each reaction. After homogenization, the 20 µL reactions were incubated at 42°C for 90 min followed by 70°C for 10 min and put on hold at 10°C in a hot-lid thermal

cycler. Following reverse transcription, excess TSO was digested using Uracil-DNA Glycosylase (UDG) (NEB, Ipswich, Massachusetts, USA) by adding 4 µL of UDG master mix adjusted for TSO input amount (see Table 2.15) and successive incubation at 37°C for 60 min. The master mix for PCR1 was prepared on ice using the Advantage 2 PCR Kit (Takara Bio Inc., Kusatsu, Shiga, Japan) along with primers specific for the constant regions of the *Trac* and *Trbc* loci as well as a universal forward primer specific for a binding sequence introduced *via* the TSO. 24 µL of RACE-ready cDNA were combined with 26 µL of PCR1 master mix (see Table 2.16) and subjected to a two-step PCR program (95°C, 1min; 95°C, 30 s & 68°C, 70 s (x times); 68°C, 7 min; 8°C, hold). Depending on the RNA input amount, PCR cycle numbers ranged from 16 to 28. The 50 µL PCR product was purified with 27.5 µL of AMPure XP beads (Beckman Coulter, Brea, USA) according to the manufacturer's instructions. After elution in 23.5 µL of Monarch DNA Elution Buffer (NEB, Ipswich, Massachusetts, USA), reactions were split into 10 µL technical replicates for successive extension PCR (PCR2). PCR2 master mix (see Table 2.17) was assembled using the Advantage 2 PCR Kit (Takara Bio Inc., Kusatsu, Shiga, Japan) and gene-specific nested primers (*Trac*, *Trbc* loci) introducing Illumina Nextera XT i5 indices (see Table 2.1), distinguishing *Tcra* from *Tcrb* libraries within each reaction. After addition of 36 µl PCR2 master mix, 4 µL of i7 Nextera XT Index Kit v2 (Illumina, Inc., San Diego, California, USA) adapters were used for indexing libraries individually. Libraries were amplified using the same program as in PCR1 with 18 PCR cycles. Library purification was carried out using AMPure XP beads (Beckman Coulter, Brea, USA) at a volume ratio of 0.8 according to the manufacturer's instructions. After final elution in 17 µL of Monarch DNA Elution Buffer (NEB, Ipswich, Massachusetts, USA), concentration was measured with the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, USA) and fragment size profile was assessed using the High Sensitivity D1000 or D5000 ScreenTape Assay (Agilent Technologies, Santa Clara, USA). Libraries were stored at -20°C until pooling for sequencing. Oligonucleotide sequences are summarized in Table 2.1.

Table 2.12: Sample-wise annealing reaction composition for low or high input amounts of RNA; RXN: reaction.

Individual annealing RXN per sample	Volume/low-input RXN [µL]	Volume/high-input RXN [µL]
total RNA	1-10	1-10
5'-CDS Primer (12 µM)	1	-
5'-CDS Primer (25 µM)	-	1
RNase-free water ad total volume	11	11

Table 2.13: Reverse transcription buffer mix composition; RXN: reaction.

MM component	Volume/RXN [µL]
5X First-Strand Buffer	4.0
dNTP Mix (20 mM)	1.0
DTT (100 mM)	0.5
total volume	5.5

Table 2.14: Composition of reverse transcription mastermix for cDNA synthesis; RXN: reaction.

MM component	Volume/RXN [µL]
RTBM	5.5
RNAseOUT™ (40 U/µL)	0.5
SMARTScribe™ RT (100 U/µL)	2.0
total volume	8.0

Table 2.15: Composition of UDG digestion mastermix for removal of excess TSO; RXN: reaction.

MM component	Volume/low-input RXN [µL]	Volume/high-input RXN [µL]
UDG buffer (10X)	2.4	2.4
UDG (5000 U/mL)	1.0	1.4
Nuclease-free water	0.6	0.2
total volume	4.0	4.0

Table 2.16: Composition of 1st PCR mastermix for gene-specific enrichment; RXN: reaction.

MM component	Volume/RXN [µL]
PCR-grade water	15
Advantage 2 PCR buffer (10X)	5
equimolar dNTP Mix (10mM)	1
Primer 1 TCR_1.PCR_for (10 µM)	2
Primer 2 GSP_TRA_rev (10 µM)	1
Primer 3 GSP_TRB_rev (10 µM)	1
Advantage 2 Polymerase Mix (50X)	1
total volume	26

Table 2.17: Composition of 2nd PCR mastermix for index extension amplification of technical replicates per sample; RXN: reaction.

MM component	Volume/RXN [µL]
PCR-grade water	25
Advantage 2 PCR buffer (10X)	5
equimolar dNTP Mix (10 mM)	1
Primer 1 mTCRA_rev_5xx (10 µM)	2
Primer 2 mTCRB_rev_5xx (10 µM)	2
Advantage 2 Polymerase Mix (50X)	1
total volume	36

2.2.6 Bulk TCRrep-seq pooling, quantification and sequencing

Libraries with compatible barcodes were pooled assigning a sequencing quota corresponding to the estimated input cell number and library concentration of each sample. Pools of TCRrep-seq libraries were quantified by qPCR using the KAPA Library Quantification Kit (Roche, Basel, Switzerland) to accurately measure only fragments capable of cluster formation on an Illumina flow cell. Pools were diluted, denatured and loaded with a final target concentration of 9 pM and a PhiX (Illumina, Inc., San Diego, California, USA) percentage ranging from 8% to 10% to compensate for initially low library complexity owed to the nature of *Tcra* and *Tcrb* constant regions. Paired-end sequencing (300 bp) was facilitated using the Illumina MiSeq system (Illumina, Inc., San Diego, California, USA). All bulk TCRrep-seq library metrics and metadata are summarized in Tables B.1, B.2, B.3, B.4 and B.5.

2.2.7 scRNA-seq and scTCR-seq library preparation and sequencing

Single-cell gene expression (GEX) and TCR repertoire (VDJ) libraries were generated by laboratory staff of the LIT's (Leibniz-Institut für Immuntherapie, Regensburg, Germany) NGS core unit strictly following the protocols provided by 10x Genomics (Pleasanton, California, USA). Expanded or re-isolated Treg were FACS-purified and directly loaded on the Chromium Controller (10x Genomics, Pleasanton, California, USA) using the Chromium Next GEM Chip G Single Cell Kit in combination with the Chromium Next GEM Single Cell 5' Library and Gel Bead Kit v1.1 (10x Genomics, Pleasanton, California, USA). Following GEM transfer and GEM-RT incubation, cDNA was purified and subjected to amplification by PCR (12-14 cycles). Amplified cDNA was purified, and fragment size distribution and quantity were assessed using the High Sensitivity D5000 ScreenTape Assay (Agilent Technologies,Santa Clara, USA). ScTCR-seq (VDJ) libraries were generated by two rounds of target enrichment followed by enriched library construction using the Chromium Single Cell V(D)J Enrichment Kit, Mouse T Cell (10x Genomics, Pleasanton, California, USA). After purification with AMPure XP beads (Beckman Coulter, Brea, USA), 1:100 dilutions of VDJ libraries were subjected to QC as performed for cDNA QC, and single-read sequencing was facilitated using the NextSeq 550 system (Illumina, Inc., San Diego, California, USA) with 150 cycles. The same cDNA samples were subjected to 5' gene expression (GEX) library construction using the Chromium Single Cell 5' Library Construction Kit (10x Genomics, Pleasanton, California, USA) with 13-16 cycles of PCR. After double-sided size selection with AMPure XP beads (Beckman Coulter, Brea, USA), GEX library quality was controlled using the D1000 ScreenTape Assay (Agilent Technologies,Santa Clara, USA). GEX libraries were sequenced on a S1 flow cell in single-read mode (100 bp) on the NovaSeq 6000 system (Illumina, Inc., San Diego, California, USA). All relevant QC and library metrics are summarized for VDJ and GEX libraries in Tables B.6 and B.7.

2.3 Computational Methods

2.3.1 Bulk RNA-Seq data analysis

After sequencing on the NextSeq550 system (Illumina, Inc., San Diego, California, USA), the bcl2fastq Conversion Software (v1.8.4) provided by Illumina (Illumina, Inc., San Diego, California, USA) was used to execute base calling. Gene annotation from GENCODE Mouse release 16 was integrated in the mouse GRCm38 genome index to allow for spliced alignment. Reads were mapped to the mouse genome using STAR

(v2.5.3a) (Dobin et al., 2013) and data quality was simultaneously assessed using FastQC (v.0.11.3) (Andrews, 2015). Using the built-in —quantMode GeneCounts option of STAR aligner generated raw, uniquely mapped read counts per gene for each sequenced sample. Samples were excluded if the percentage of reads that could not be mapped due to insufficient fragment length exceeded 66%. Subsequent analyses were done in R (v4.1.0) (R Core Team, 2021) in conjunction with Bioconductor (v3.13) for managing and maintaining R packages with compatible versions (Gentleman et al., 2021; Morgan, 2022; Team, 2021). A comprehensive read count table (RCT) was assembled in R corresponding to a metadata file thoroughly describing all samples. To that end the fourth column of counts was extracted for each sequencing sample adding up reads for multiply sequenced biological samples. An annotation data frame was generated including EnsemblID, EntrezID, gene symbol, transcript length and gene type columns using the AnnotationDbi R package together with the Bioconductor annotation package org.Mm.eg.db (Carlson, 2021; Pagès et al., 2021). With the help of edgeR (Chen et al., 2021; Smyth et al., 2021), DGEList objects were generated from subsets of the comprehensive RCT and metadata as required for the indicated analyses. Next, rare transcripts were excluded *via* filterByExpression or a manual gene-wise threshold for the minimal number of samples required with counts per million (CPM) > 1, and library size normalization was performed. The calculation of group- or sample-wise $\log_2$-transformed as well as raw RPKM (reads per kilobase per million) and CPM values was done applying the cpm and rpkm or cpmByGroup and rpkmByGroup functions to DGEList objects, respectively, and data was scaled (Zscore) using R's scale function. If applicable, $\log_2$-transformed data was batch-corrected (PCR cycle No./ prep. batch) for visualized data by applying the removeBatchEffect function. Gene expression levels (RPKM) on the $\log_2$-scale of selected genes of interest (GOIs) were represented as bar plots generated using ggplot2 in conjunction with additional R packages (Kassambara, 2020; Slowikowski, 2021; Wickham, Chang, et al., 2022; Wickham & Seidel, 2022). Heatmaps representing $\log_2$RPKM or scaled expression data were generated using pheatmap (Kolde, 2019) or heatmap.2 from gplots (Warnes et al., 2022). For dimension reduction *via* uniform manifold approximation and projection (UMAP) the umap package (Konopka, 2022) was used and calculated embeddings were plotted with the help of ggplot2 along with additional R packages (Kassambara, 2020; Slowikowski, 2021; Wickham, Chang, et al., 2022; Wickham & Seidel, 2022). For each differential gene expression (DGE) analysis (Chen et al., 2021; Smyth et al., 2021), a design matrix consisting of the contrast of interest (COI) and feasible blocking factors was generated for dispersion estimation and fitting of a quasi-likelihood, negative binomial, generalized log-linear model to raw counts. The glmTreat function was used to conduct statistical tests per gene for the corresponding COI and linear model relative to a globally applied fold-change (FC) threshold of 1.5. DGE results were vi-

sualized as scatter plots (Kassambara, 2020; Slowikowski, 2021; Wickham, Chang, et al., 2022) representing gene-wise log_2CPM calculated based on the corresponding log_2FC and the average log_2CPM across samples of the indicated comparison. Data was filtered conditional on log_2CPM > 1 and log_2RPKM >1 values unless indicated otherwise. Genes of interest (GOIs) were highlighted as specified. The heatmap representation of genes DE between Treg residing in different tissues of baseline mice was heatmap.2 of the gplots R package (Warnes et al., 2022). Gene set enrichment analyses (GSEA) for reference gene sets (Hallmark gene sets for mouse: `http://bioinf.wehi.edu.au/MSigDB/v7.1/Mm.h.all.v7.1.entrez.rds`; Kegg gene sets for mouse: `http://bioinf.wehi.edu.au/MSigDB/v7.1/Mm.c2.cp.kegg.v7.1.entrez.rds`), were performed *via* competitive gene set tests using edgeR's camera.DGEList function. The enrichment was calculated based on the log_2FC-ranking of the the underlying COI. To contrast two individual DGE analyses, the fry function was used to carry out two-sided rotation gene set tests for determining the significance of the enrichment of a gene set from one comparison across the log_2FC-ranking of the respective other comparison. The results of both kinds of GSEA were visualized with the barcodeplot function of the limma package (Smyth et al., 2021). Barplots visualizing the significance levels of GSEA were generated with ggplot2 (Wickham, Chang, et al., 2022). The comprehensive log_2-transformed, PCR2-batch-corrected CPM matrix for the baseline, BMT control and prophylaxis models was transposed and annotated in R and exported for analysis with Graphia (Freeman et al., 2020). By means of graph-based clustering, Graphia identified co-expression networks of genes taking into account their Pearson's correlation (r≥0.845, k-NN using r with k>4, N(nodes)=3829, N(edges)=10.7k, Louvain cluster granularity = 0.6). The two-dimensional projection of the corresponding node network was exported for visualization. Metascape (`https://metascape.org`) was used for selected pathway enrichment analyses as indicated. The tool is uses the hypergeometric test and Benjamini-Hochberg *P* value correction to determine statistically significant over-representation of ontology terms within an input gene list (Zhou et al., 2019). Metascape was provided with lists of GOIs exported from R or Graphia along with the corresponding background gene lists to determine pathway enrichment using the default parameters for minimal overlap (3), *P* value cutoff (0.01) and minimal enrichment (1.5). The results were retrieved and the leading pathways per GroupID as well as the corresponding Benjamini-Hochberg corrected *P* values were extracted for visualization in R.

All steps in which data frame were joined, filtered, transposed or transformed were facilitated with the help of various R packages (K. Müller & Wickham, 2022; Shannon, 2022; Wickham, 2020; Wickham, François, et al., 2022; Wickham & Girlich, 2022). Manipulation and generation of spread sheets was carried out using additional R pack-

ages (Dragulescu & Arendt, 2020; Wickham, Hester, et al., 2022; Wickham & Bryan, 2022). For arranging multiple plots, additional R packages were used (Wickham, 2021; Wilke, 2020). A globally defined and consistently used color code for all samples was generated using the colorspace package (Ihaka et al., 2022).

2.3.2 Bulk TCRrep-seq data analysis

Paired-end sequencing data (300bp) was processed using the bcl2fastq conversion pipeline (v1.8.4) provided by Illumina (Illumina, Inc., San Diego, California, USA). Alignment of paired-end reads was facilitated using PEAR (v0.9.11) (Zhang et al., 2014). On the basis of the incorporated UMIs, MIGEC (v1.2.9) (Shugay et al., 2014) was used to assemble consensus sequences for molecular identifier groups (MIGs), representing unique *Tcra* or *Tcrb* transcripts. Using MiXCR (v3.0.1.8) (Bolotin et al., 2015) MIGs were mapped. With the help of the immunarch package (v0.6.6) (Nazarov et al., 2022) used under R (v4.0.3) (R Core Team, 2021) clonotype count tables, *Trbv* gene usage and repertoire diversity data frames were generated. First, data was parsed using the repLoad function, followed by repExplore to determine clonotype counts. Data was filtered based on a minimally required clonotype size and was downsampled accordingly. Afterwards, repDiversity was used to calculate the inverse Simpson Index (iSI), a surrogate for repertoire diversity. With the help of the geneUsage function provided with appropriate parameters (“musmus.trbv”, .type = “segment”, .ambig = “exc”) the *Trbv* gene usage was determined. Diversity barplots representing the iSI and selected *Trbv* gene usage frequency barplots were generated with the help of ggplot, ggsignif and additional R packages (Ahlmann-Eltze & Patil, 2021; Kassambara, 2020; Wickham, Chang, et al., 2022). The geneUsageAnalysis function was supplied with a distance matrix (method = “manhattan”) calculated based on *Trbv* gene usage data and the following parameters: .method = “tsne+kmeans”, .perp =4, .k=5. The resulting t-SNE was plotted with the help of the ggplot2 R package (Wickham, Chang, et al., 2022). For the visualization of overlapping fractions of *Tcrb* repertoires as circos plots, the circlize package was used (Gu et al., 2014). Barycentric distribution plots were plotted with ggplot 2 (Wickham, Chang, et al., 2022). All TCRrep-seq-related enrichment plots were generated using a modified version of limma’s (Smyth et al., 2021) barcodeplot function.

2.3.3 Single-cell RNA-seq analyses

Single-cell gene expression data was analyzed by bioinformatics staff of the LIT’s NGS core unit. The most important steps of analysis are described in the following. Demultiplexing of sequencing reads was facilitated with cellranger (v4.0.0) mkfastq (G. X. Y.

Zheng et al., 2017). Mapping of the genome was accomplished with cellranger count and the mouse index (Mouse reference, mm10 (GENCODE vM23/Ensembl 98)) supplied by 10x Genomics (Pleasanton, California, USA). Seurat (v4.0.0) (Hoffman, 2022; Satija et al., 2022) was used in connection with R (Gentleman et al., 2021; R Core Team, 2021) to perform quality control and analysis of single-cell data. The corresponding Seurat object was generated by importing cellranger output *via* Read10x, followed by CreateSeuratObject supplied with min.cells = 5 and min.features = 200. Appropriate thresholds for data filtering (nCount_RNAmin = 1000; nCount_RNAmax = 15000; nFeature_RNAmin = 200: nFeature_RNAmax = 2500; ribosome_pct > 10; percent.mt <5) as determined by QC were applied next. After normalization using Seurat's logNormalize function, data was scaled. Principal component analysis (PCA) was carried out based on the 2000 most variable features. Based on the number of loaded cells, the estimated doublet rate provided by 10x Genomics (Pleasanton, California, USA) was multiplied by the number of cells in the respective Seurat object to obtain the number of expected doublets (nExp). Without refinement, Doubletfinder's (McGinnis et al., 2019) doubletFinder_v3 function served the purpose of doublet removal and was run with the following parameters: pN = 0.25, pK = 0.09, nExp = nExp, PCs = 1:10. The count data per cell for each sample was joined by merging the respective Seurat objects, followed by a global normalization using logNormalize. Normalized counts of the comprehensive Seurat object were scaled, and a PCA was performed based on the 2000 most variable genes. Supplying the FindNeighbors, FindClusters and RunUMAP functions with suitable parameters (dim = 1:11, resolution = 0.5) resulted in clustering of the single-cells. The differential gene expression (DGE) between individual clusters as compared to all other clusters was assessed by means of a non-parametric Wilcoxon rank sum test. To that end, the FindAllMarkers function was executed with the default logFC = 0.25 and min.pct = 0.25 assuring only genes expressed by at least 25% of cells in either of the two compared groups were tested. Heatmaps were generated to represent the scaled gene expression of up to 10 cluster-defining genes for 2500 randomly sampled cells per cluster. Co-expressed gene signatures identified via bulk RNA-seq were used to generate scores by subtracting average gene expression of 100 control features from average expression levels per signature and cell *via* AddModuleScore. All analyzed features were binned based on average expression control features were randomly chosen from each bin. Analogously, scores for selected pathways were generated with 10 control features. Signature scores and expression levels of single genes of interest were projected on the pre-existing UMAP embedding *via* Seurat's FeaturePlot function. Cells isolated from the colon were processed individually by providing FindVariableFeatures with nfeatures = 2500 and dims = 1:30, and running FindNeighbors and RunUMAP. On this subset of cells, RNA velocity (La Manno et al., 2018) was used for pseudo-time-based trajectory inference. Filtering and normalization were

performed (min_shared_counts = 20, n_top_genes = 2000) before trajectories were inferred (n_pcs = 30, n_neighbors = 30) and overlaid with the pre-existing UMAP of the colon-subclustering.

2.3.4 Single-cell TCRrepseq data analysis

Single-cell TCR repertoire data was analyzed by bioinformatics staff of the LIT's NGS core unit. In brief, sequencing data from VDJ libraries was processed using cellranger (v5.0.0) together with the mm10 reference genome (refdata-cellranger-vdj-GRCm38-alts-ensembl-5.0.0) provided by 10x Genomics (Pleasanton, California, USA). Only cells with a single productive TCR and both gene expression and clonotype repertoire data coverage were kept for further analysis. Clonotypes of interest were highlighted in the UMAP embedding determined based on single-cell gene expression data as described above. Groups of Treg were designated based on categories of clonotypes of interest (containing/not containing any of *Trbv12-1*, *Trbv12-2*, *Trbv26*) and their distribution across the global and colon-restricted single-cell clusters was assessed using a two-sided Fisher's exact test in R (R Core Team, 2021). The $\log_2$-transformed enrichment for each cluster was visualized as lollipop plot generated with the R package ggplot2 (Wickham, Chang, et al., 2022). In analogy to the cluster-wise gene expression comparisons in the previous section, gene expression differences between colon-homing Treg with a TCR of interest and all other clonotypes were investigated. The resulting DEGs were visualized as volcano plot with the help of ggplot2 (Wickham, Chang, et al., 2022). Clonotype data was processed with the help of the R package immunarch (Nazarov et al., 2022) in analogy to the bulk TCRrep-seq data. Similarly, circos and barycentric distribution plots were generated. In addition, a heatmap representing the Morisita's Overlap Index was created using the repOverlap and vis functions provided by the immunarch R package (Nazarov et al., 2022).

2.3.5 Thesis development and compilation

This thesis was developed in the RStudio IDE (v2022.02.3+492 "Prairie Trillium" Release) using rmarkdown (Allaire et al., 2022), knitr (Xie, 2022b), and the R package bookdown (Xie, 2022a). In addition to knitr's kable function, kableExtra (Zhu, 2022) was used to improve representations of tables. LaTeX templates for the cover page and the overall structure of the dissertation were retrieved from the website of the Regensburg International Graduate School of Life Science (RIGeL: `https://rigel-regensburg.de/index.php?browse=119`) and were refined for the use with the bookdown package.

CHAPTER 3

Results

3.1 *In vitro* expansion of Treg - a dual approach

The potential of donor-type $CD4^+CD25^+$ regulatory T cells (Treg) to impair alloreactive $CD4^+CD25^-$ Tcon in a murine transplantation model of acute GvHD (C57BL/6 into BALB/c) and the resulting protection from lethal aGvHD was demonstrated previously (Hoffmann et al., 2002). However, besides their donor origin and ability to secrete interleukin 10, prerequisites for the clinical efficacy of Treg remain not fully understood. Therefore, two different expansion methods for cultivating Treg *in vitro* have been employed. The conventional polyclonal expansion ("poly") with anti-CD3/CD28 beads was performed in parallel to an alloantigen-specific expansion ("allo") in which spleen-derived, recipient-type $CD11^+$ dendritic cells (DCs) were irradiated prior to their addition to the culture as allogeneic stimulus. High amounts of recombinant human IL-2 (rhIL-2) were crucial for mimicking the typical micromilieu of activated Tconv in both expansion protocols. The rationale behind the allo branch was a potentially beneficial alloantigen-specific selection of Treg clonotypes prior to transplantation. A schematic representation of the dual approach facilitated by the work group of PD Dr. rer. nat. Petra Hoffmann and Prof. Dr. med. Matthias Edinger (WG E/H) is shown in **Figure 3.1 a**.

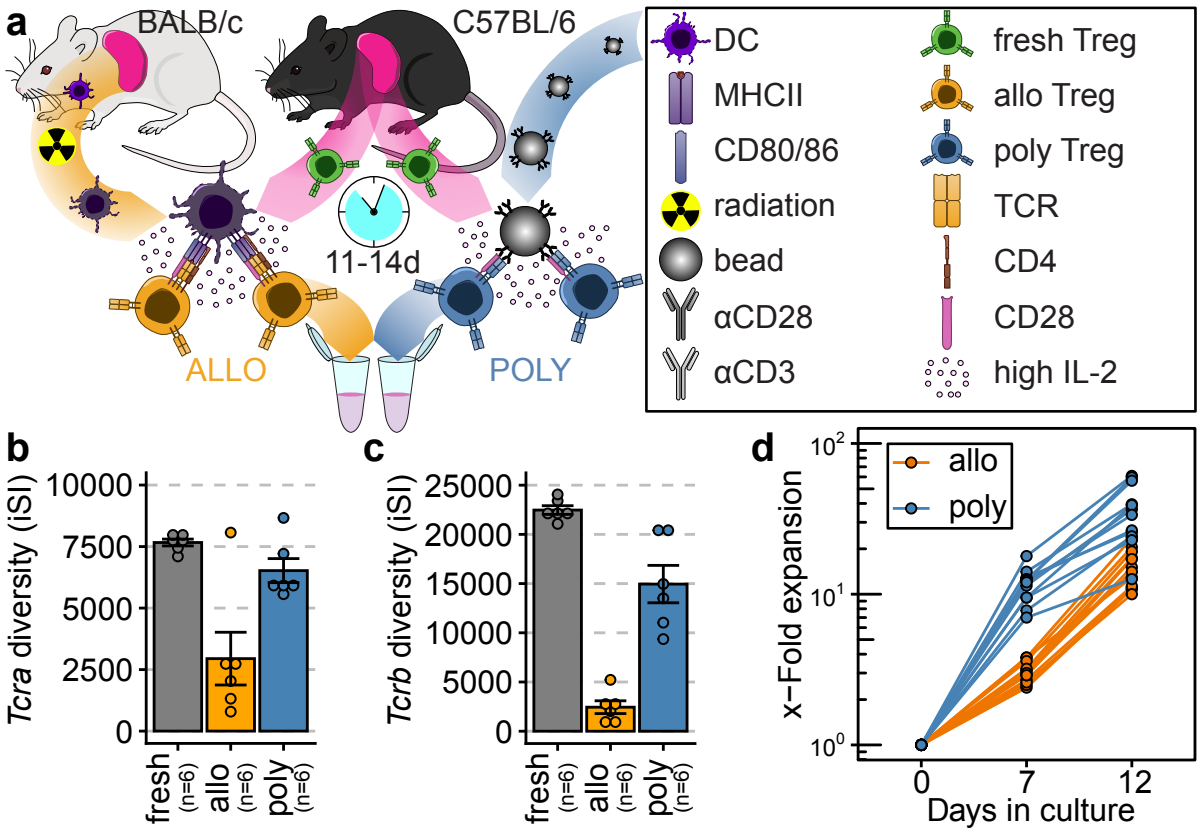

Figure 3.1: ***In vitro* expansion protocols determine Treg TCR diversity and expansion.** **a** Schematic depiction of the allo and poly Treg expansion protocols. Freshly isolated C57BL/6 Foxp3^{gfp} (H2^{b},Treg donors) Treg were cultured for 11-14 days in the presence of high concentrations of rhIL-2 (2000 U/mL) and either 30 Gy irradiated spleen-derived BALB/c (H2^{d}) DCs for allo or conventional anti-CD3/CD28 beads for poly Treg generation. **b,c** Barplots of TCR alpha (*Tcra*, b) and beta (*Tcrb*, c) chain repertoire diversities (inverse Simpson Index , iSI) are shown for spleen-derived fresh (d0) or expanded Treg (allo/poly, d11-14) after UMI-based sequencing of 5'RACE products of TRA/Tcrb mRNAs. Bars represent the mean±SE of n=5-6 independent experiments and individual data points are shown as dots. **d** Simplified expansion kinetics for allo and poly Treg are compared in a line plot (n=12 each, data kindly provided by WG E/H).

3.1.1 TCR repertoire reduction by allogeneic expansion

In order to investigate the TCR repertoire of expanded Treg, TRA/Tcrb-chain CDR3 region-specific PCR fragments were generated from total RNA. The employed 5' Rapid Amplification of cDNA Ends (RACE) system enabled unique molecular identifier (UMI) tagging of cDNA and TRA/Tcrb fragment enrichment was achieved by subsequent gene-specific PCR. The resulting TCR-chain fragments were subjected to massively parallel sequencing in which long paired-end reads assured sufficient sequence coverage. The work published by Mamedov et al. (2013) served as a guideline.

While poly Treg retained a broad TCR repertoire throughout *in vitro* culture, a gradual narrowing of the allo Treg TCR repertoire could be observed, as estimated by the inverse Simpson Index (iSI) of *Tcra* and *Tcrb* repertoires shown in **Figure 3.1 b** and **c**. The observed changes in TCR repertoire diversity were in line with the expected broad responsiveness to the beads in contrast to the restricted activation by the irradiated DCs of poly and allo Treg, respectively. Also, growth kinetics differed as the initial allo expansion was found delayed opposite to more rapidly expanding poly Treg which is shown in **Figure 3.1 d**. Both protocols allowed for the generation of Treg to an extent adequate for successive transplantation.

3.1.2 Preserved Treg identity in allo and poly Treg

A differential gene expression (DGE) analysis comparing allo and poly Treg revealed that the majority of genes defining Treg, as determined by a systematic meta analysis for Treg signature genes, was not significantly altered between the two cell products. Treg-signature genes defined by Aubert et al. (2020), expanded by two additional genes of interest (*Itgb7*, *Sell*) are highlighted (**Figure 3.2 a**). The heatmap in **Figure 3.2 b** compares core Treg gene expression levels of freshly isolated, allo and poly Treg and illustrates their overall uniformity in this regard.

For the functional pathway attribution of genes differentially expressed between allo and poly Treg, a gene set enrichment analysis (GSEA) was performed. In general, GSEA is a computational method to determine whether a predefined set of genes is significantly more abundant in one sort as compared to the other sort of samples in a pair-wise comparison. The results for the allo vs. poly comparison are represented as barcode plots in **Figure 3.2 c**. Here, several pathways, such as Glycolysis/Gluconeogenesis (Kegg), Hypoxia or IL-2/STAT-5-Signaling (Hallmark), were found enriched among genes upregulated in allo Treg, likely reflecting the divergent culture conditions of the expansion protocols.

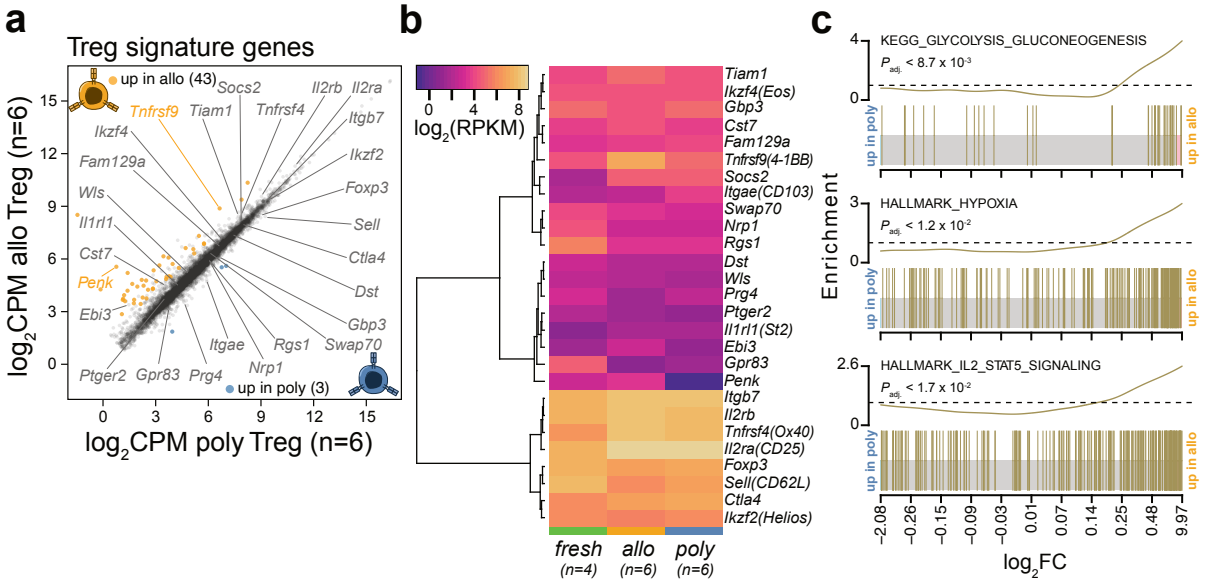

Figure 3.2: **Preserved Treg identity and divergent culture conditions are reflected by allo and poly Treg gene expression profiles.** **a** Scatter plot of global gene expression (as measured by RNAseq, n=6 independent experiments each, $\log_2$CPM>1, $\log_2$RPKM>1) contrasting expanded allo with poly Treg. Differentially expressed genes (DEGs) are colored as indicated by the description in the plot panel ($P_{\text{adj.}|\text{FC}|}$>1.5). Treg signature genes are labeled, as defined by Aubert et al. (2020) extended by *Itgb7* and *Sell*. **b** Heatmap showing the average $\log_2$-scaled reads per kilobase per million (RPKM) for Treg signature genes per group (freshly isolated, allo, poly). **c** GSEA barcode plots representing the enrichment (olive enrichment worm) of selected Kegg and Hallmark gene sets, stated for each panel individually. Detected genes (olive vertical bars) are ranked according to their $\log_2$FC in the allo vs. poly DGE analysis and Benjamini-Hochberg adjusted enrichment *P* values ($P_{\text{adj.}}$) of competitive gene set tests are provided.

3.2 MHC-disparate transplantation models

To track the fate of transplanted Treg after *in vitro* expansion, four different *in vivo* models were established and facilitated by WG E/H, of which three involved a complete MHC-mismatch. Samples derived thereof were kindly provided for further analysis. Per baseline sample, $CD4^+CD25^+$ Treg were re-isolated from colon, liver and spleen of 2 to 5 Treg donor strain animals ($CD45.2^+$ C57BL/6 $Foxp3^{gfp}$ mice, $H2^b$) in a homeostatic state, as shown in **Figure 3.3 a**. While no analyzable amounts of colon-homing naive $CD62L^+$ Treg could be retrieved, Treg originating from liver and spleen were additionally FACS sorted into $CD62L^+$ and $CD62L^-$ populations.

The BMT control model is illustrated in **Figure 3.3 b** and comprised the conditioning of wild type BALB/c recipients ($H2^d$) by total body irradiation with 8 Gy (TBI) followed by the transfer of T cell-depleted bone marrow (TCD BM) from MHC-mismatched con-

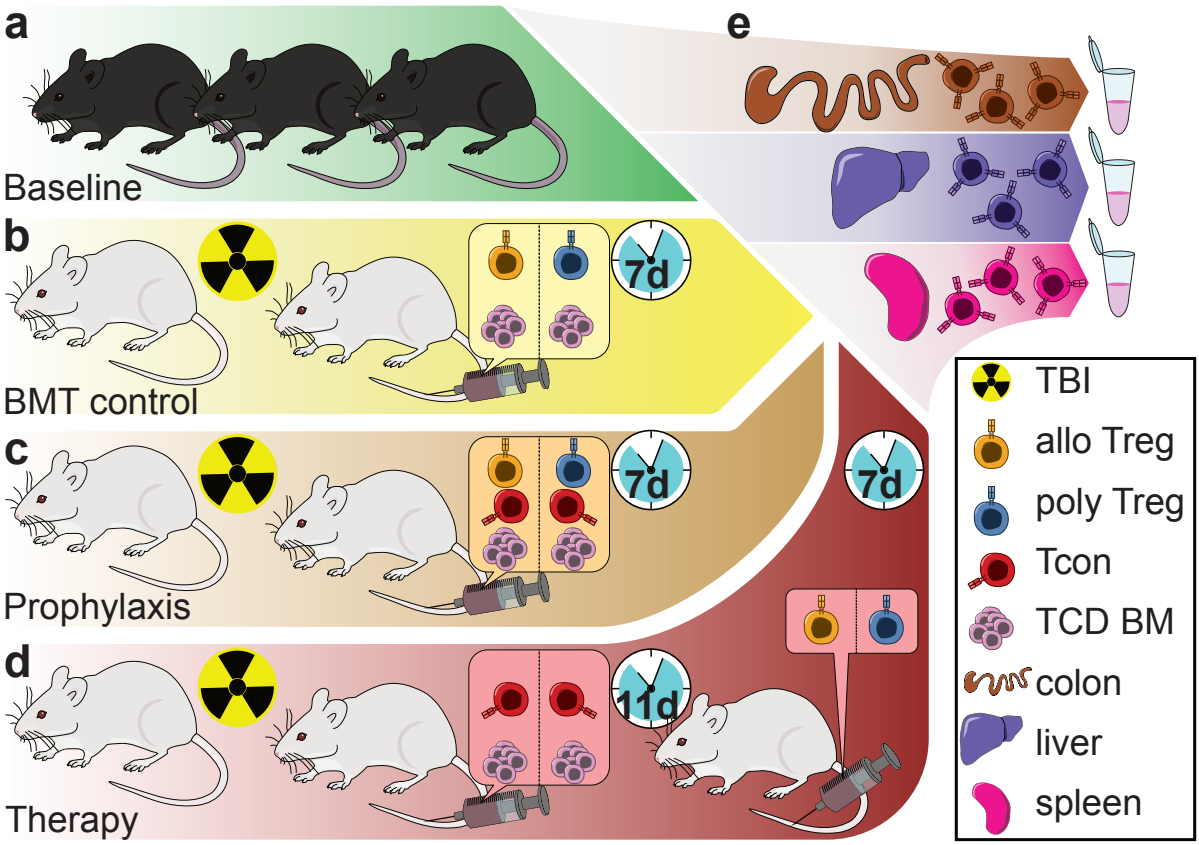

Figure 3.3: **Baseline and MHC-disparate *in vivo* models overview.** Schematic illustration of the baseline and BMT model and re-isolation procedures. **a** As a homeostatic control ("baseline"), Treg were derived from pools of 2 to 5 healthy CD45.2^{+} C57BL/6 Foxp3^{gfp} mice (H2^{b}, Treg donors). **b** In the BMT control model without GvHD ("no GvHD"), wild type BALB/c mice (H2^{d}) were conditioned by total body irradiation (TBI) with 8 Gy and received T cell-depleted bone marrow (TCD BM, 2.5x10^{6} cells/recipient) from congenic CD45.1^{+} C57BL/6 mice (H2^{b}, BM donors) along with *in vitro* expanded donor Treg (allo or poly, 1.0x10^{6} cells/recipient) at d0. **c** In the prophylaxis model, recipients additionally were infused with CD4^{+} CD25^{-} Tcon from BM donors at a Treg:Tcon ratio of 1:1 at d0. **d** In the therapy model, TCD BM and Tcon from BM donor strain mice were administered at d-11 and *in vitro* expanded donor Treg were given at d0. **e** In all transplant models, Treg were re-isolated from the indicated organs on d7 for further analyses.

genic CD45.1^{+}C57BL/6 mice (H2^{b}, BM donors) as well as *in vitro* expanded allo or poly Treg from Treg donor strain mice at day 0. Despite the radiation-induced colitis, this approach is referred to as the “no GvHD” model as no notable amounts of Tconv had arisen by day 7 after transplantation (information provided by WG E/H).

In contrast, the prophylaxis model (**Figure 3.3 c**) included the administration of BM donor strain Tconv in addition to donor Treg at a 1:1 ratio at day 0, resulting in GvHD as a direct consequence of co-transplanted alloreactive conventional T cells.

As shown in **Figure 3.3 d**, mice in the therapy model received TCD BM along with Tcon at day -11 and were only treated with allo or poly Treg at day 0. Consequently, the therapy model led to the most severe phenotype.

In all BMT models, Treg were re-isolated from recipient colon, liver and spleen compartments 7 days after initial transfer, as depicted in **Figure 3.3 e**. Animal handling, removal of the organs of interest, tissue digestion and fluorescence-activated cell sorting (FACS) were all performed by members of the WG E/H, before RNA was extracted for subsequent library preparation and RNA-sequencing.

3.2.1 Partial retention of allo-poly differences *in vivo*

As demonstrated in section 3.1.2, allo and poly Treg hardly differed in functionally important gene expression. Nevertheless, differences could be observed and the question arose whether the transcriptional imprint caused by the unequal culture conditions is still present 7 d after transplantation. Consequently, a comprehensive DGE analysis comparing allo to poly Treg of all BMT models (BMT control, prophylaxis and therapy) was performed. Despite the overall fewer significant DEGs between allo and poly Treg after BMT as illustrated by **Figure 3.4 a**, the GSEA with the set of genes upregulated in allo donor Treg revealed a highly significant enrichment across the $\log_2$FC ranking of the allo vs. poly comparison among transplanted Treg, as shown in **Figure 3.4 b**. However, when samples were stratified according to model and organ, only for the prophylaxis and BMT control models, individual allo vs. poly comparisons resulted in a significant enrichment of the same set of genes for all tissues, while in therapeutically administered Treg, only liver-homing cells retained the respective gene expression differences to a notable extent, as shown in **Figure 3.4 c**. Thus, the “memory” effect evoked by the *in vitro* expansion protocol faded with the conditioning dating back further.

Moreover, the top enriched Hallmark gene set (HALLMARK INTERFERON ALPHA RESPONSE) for the comprehensive allo vs. poly comparison of all transplanted Treg, was found enriched in Treg re-isolated from poly recipients, as illustrated in **Figure 3.4**

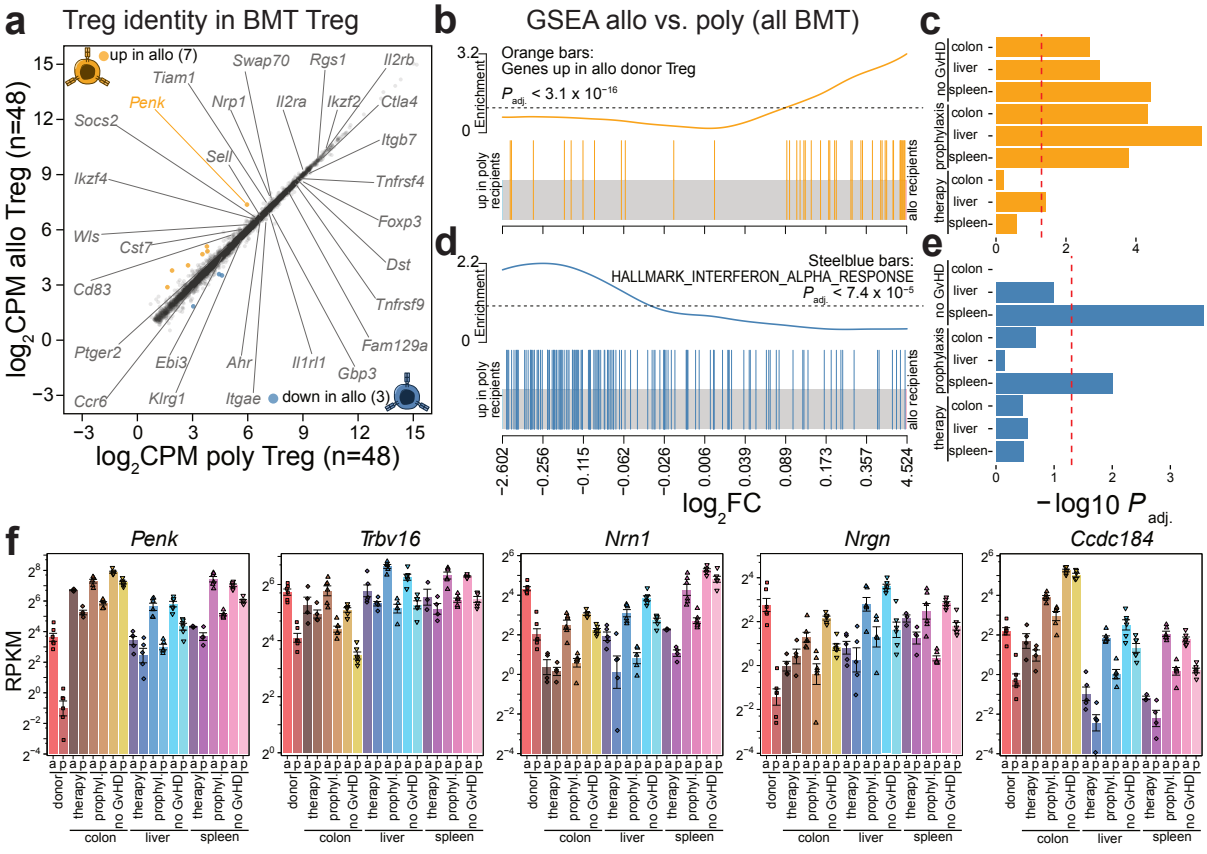

Figure 3.4: **Allo and poly Treg gene expression differences are partly preserved after BMT. a** Scatter plot representing global gene expression (log_2CPM>1, log_2RPKM>1) differences between allo and poly Treg across all BMT models after transplantation (n=48 each). Treg core signature genes are highlighted, as previously defined. **b** Barcode plot presenting GSEA results for allo-specifically upregulated genes in donor Treg ranked according to the log_2FC of the allo vs. poly comparison in BMT Treg. **c** Based on the gene set in **b**, individual enrichment *P* values are represented as -log_{10}($P_{adj.}$) bar plot for each organ per BMT model. In **b** and **c**, enrichment was determined by two-sided rotation gene set tests. **d** Barcode plot showing GSEA results for the top enriched Hallmark gene set across the same log_2FC ranking as in **b**. **e** For the same Hallmark gene set, individual enrichment *P* values are shown as -log_{10}($P_{adj.}$) bar plot for each organ per BMT model. In **d** and **e**, enrichment was determined by competitive gene set tests. All *P* values shown were Benjamini-Hochberg adjusted (**b**,**c**,**d**,**e**) and dashed red lines in **c** and **e** mark the FDR threshold of 0.05. **f** Bar plots of selected genes that were differentially expressed (DE) between allo and poly *in vitro* as well as the comprehensive *in vivo* analysis. Bars represent the mean±SE of RPKM values from n=4-6 independent experiments on a log_2-scaled y-axis. Individual data points are shown as dots.

d. However, when the individual allo vs. poly DGE analyses were interrogated in this regard, only spleen-derived Treg from BMT control or prophylaxis animals passed the FDR threshold of 0.05 (**Figure 3.4 e**), suggesting a more inflammatory micro environment in spleens of poly Treg recipients early after conditioning. Examples for genes remaining upregulated in allo Treg after BMT (*Penk*, *Tcrbv16*, *Nrn1*, *Nrgn*, *Ccdc184*) are shown in **Figure 3.4 f**. Apart from therapeutic Treg re-isolated from recipient colons, the selected genes are consistently more abundant in allo Treg as compared to their poly counterparts across organs and BMT models.

3.2.2 The Colon as major driver of Treg gene expression

Despite their differences in gene expression, allo as well as poly Treg maintained the phenotypic properties of canonical Treg 7 days after transplantation in all three BMT models. Next, the impact of the target tissue on Treg transcription was investigated and thus, Treg were characterized in terms of colon, liver and spleen-specific gene expression programs. To this end, a commonly used technique for dimension reduction, namely the uniform manifold approximation and projection (UMAP), was applied to the gene expression data of all Treg, in an effort to explore their interrelation. Results thereof were visualized as two-dimensional embedding (**Figure 3.5 a**). Although, distances along the axes of the UMAP plot cannot be directly interpreted, the formation of clusters can be translated to the transcriptional similitude of the respective samples. All Treg re-isolated from baseline as well as BMT model colons formed a dense cluster, clearly distinguishable from all other Treg groups. Similarly, donor Treg and liver or spleen-resident Treg were mapped to two further independent clusters. For transplanted, liver and spleen-derived Treg two adjacent clusters were observed, in which the separation by the corresponding BMT model was more pronounced for spleen than for liver homing cells. Altogether, the UMAP plot served as roadmap for biologically meaningful subsequent comparisons of groups of Treg.

3.2.2.1 Colon-specific gene expression in baseline Treg

As a barrier tissue harboring a multitude of bacteria strains, colon evoked the most prominent differences in Treg gene expression which were further dissected by first comparing colon to non-colon resident Treg of the baseline model, only taking $CD62L^-$ non-naive Treg into account. A considerable number of DEGs was found up- (N=1576) or downregulated (N=1150) in colon-resident Treg, as represented by **Figure 3.5 b** . The highlighting of non-lymphoid tissue (NLT) and lymphoid tissue (LT) gene signatures defined by Miragaia et al. (2019) revealed that for the majority of genes, the observed differences were in accordance with the expected direction of regulation. The

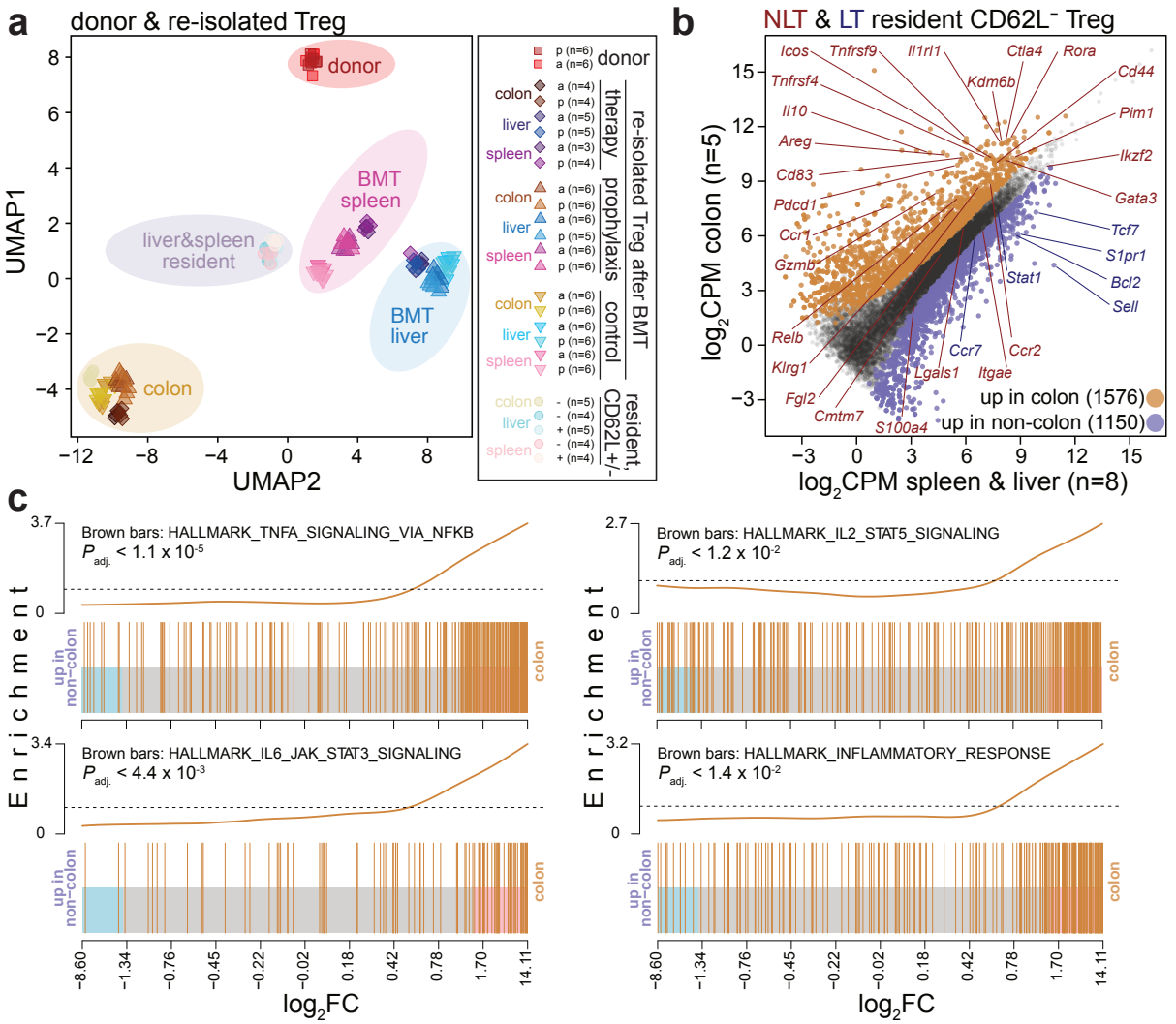

Figure 3.5: **CD62L⁻ baseline Treg gene expression is controlled by inflammatory stimuli of the healthy colon. a** UMAP embedding of donor, baseline and transplanted Treg (left panel). Symbols and colors discriminating allo/poly, CD62L+/-, the tissue and the model are described in the figure legend (right panel). **b** Scatter plot representing global average gene expression differences between colon (n=5) and liver/spleen (n=8) resident CD62L⁻ Treg. The number of significant DEGs ($P_{adj.|FC|>1.5}>0.05$) is provided in the embedded legend and data points are colored accordingly. NLT and LT signature genes as defined by Miragaia et al. (2019) are highlighted in the corresponding color of the plot title. **c** Barcode plots presenting GSEA results for the top 4 ranking Hallmark gene sets (stated per panel) that were found enriched when ranked according to the log_2FC of the colon vs. non-colon comparison in CD62L⁻ baseline Treg. Enrichment was determined by competitive gene set tests and Benjamini-Hochberg adjusted *P* values are provided for each panel.

consistency between these results and previously published work further corroborated the strong influence of the colon compartment on the gene expression of Treg. To elaborate on the function of colon-specific genes, GSEA were conducted and the results for the top 4 enriched Hallmark gene sets were summarized in **Figure 3.5 c**. All gene sets shown are associated with inflammatory processes, suggesting the presence of IL-2-producing Tcon and TNF-alpha secreting macrophages in the homeostatic colon.

To further characterize DEGs, non-colon-specific genes were simultaneously filtered for liver or spleen specificity. The resulting gene expression data was hierarchically clustered and visualized as a heatmap (**Figure 3.6 a**).

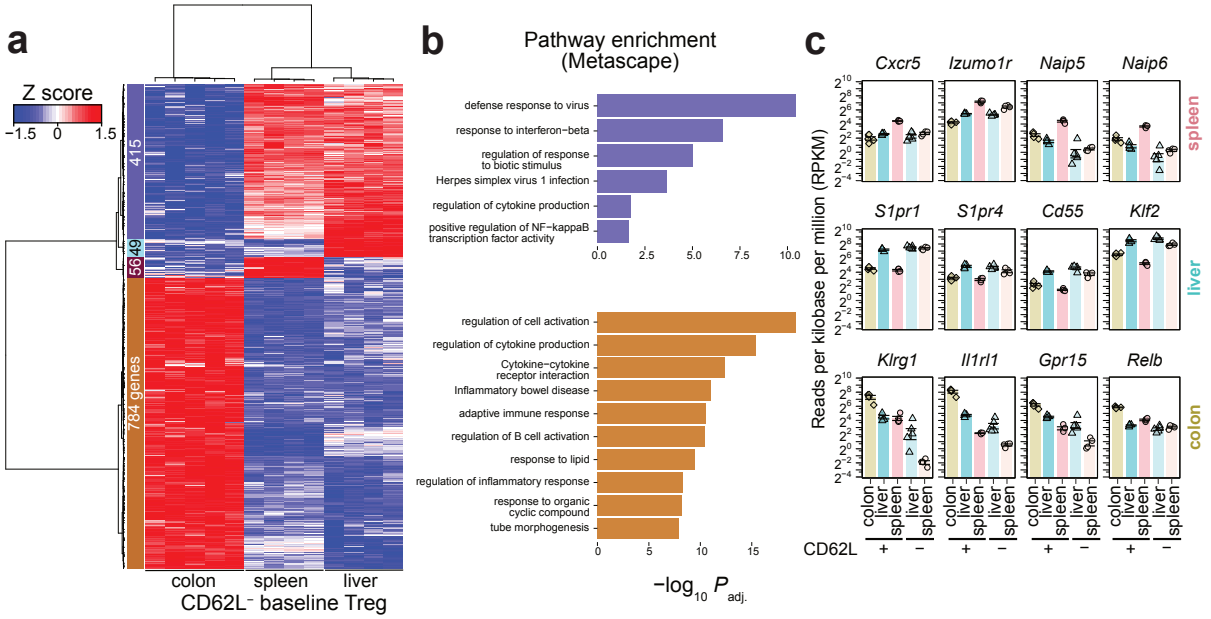

Figure 3.6: **Global DGE analysis defines functionally distinct colon and non-colon-specific Treg gene expression patterns of the homeostatic state. a** Heatmap representing hierarchically clustered and scaled expression data of genes that were DE between colon, liver and spleen resident CD62L$^-$ Treg. Clusters are color coded and sizes indicated by the side bar. **b** For the two bigger clusters defined in **a**, the top enriched pathways are ranked according to increasing Benjamini-Hochberg adjusted *P* values, as determined by a Metascape analysis with the appropriate background gene set, and visualized on the -log$_{10}$-scale. For spleen- or liver-specific clusters, no significantly enriched terms were observed. **c** Bar plots illustrating organ-specific gene expression levels for exemplary genes (top row: spleen, middle row: liver, bottom row: colon). Bars show mean±SE of RPKM values from n=4-5 independent experiments on a log$_2$-transformed y-axis. Individual data points are shown as color-coded distinct shapes.

For the two largest, easily distinguishable clusters, pathway enrichment analyses with Metascape were performed. The leading enriched pathways identified were clustered and ranked by decreasing significance, as shown in **Figure 3.6 b** (Zhou et al., 2019). According to that, Treg residing in the colon were in an activated state (e.g. regulation

of cell activation, regulation of cytokine production) specific to their environment (e.g. inflammatory bowel disease, response to lipid) while Treg located in liver or spleen had responded to different stimuli (e.g. defense response to virus, response to interferon-beta). Differences between liver and spleen were comparatively small, as represented by the two lesser clusters of genes. Selected tissue-specific genes were juxtaposed to illustrate exemplified gene expression differences (**Figure 3.6 c**).

3.2.2.2 Dominance of the colon over BMT Treg gene expression

The dominant influence of the colon on Treg gene expression, as shown in **Figure 3.5 a**, also determined the fate of transplanted cells. In analogy to the baseline model (section 3.2.2.1), the comparison of colon to non-colon migrating Treg yielded considerable numbers of up- (N=1056) and downregulated (N=866) genes (**Figure 3.7 a**).

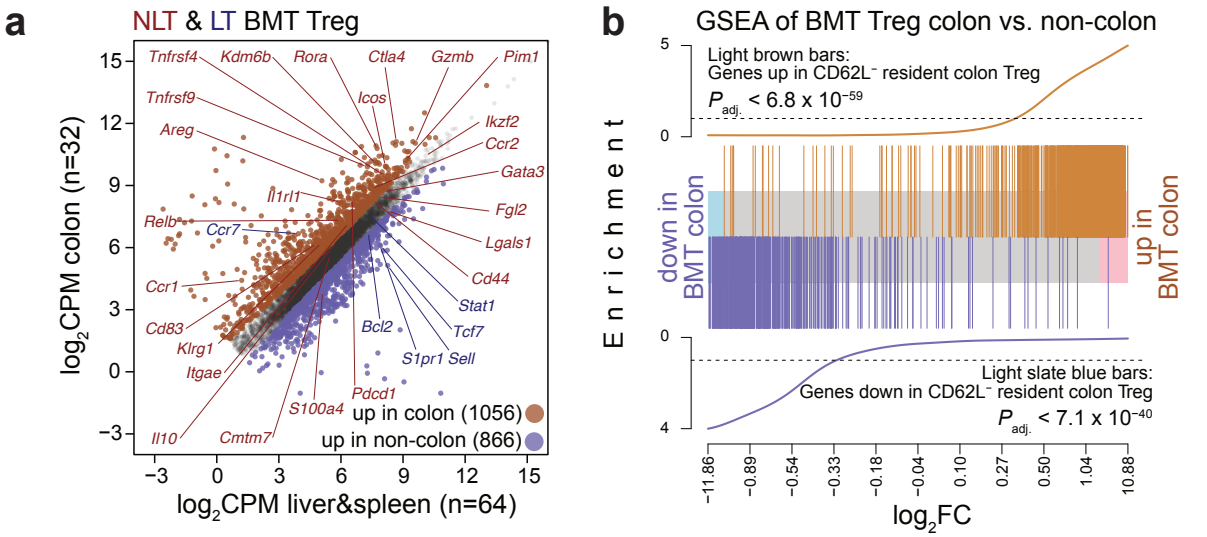

Figure 3.7: **Transplanted Treg rapidly acquire colon-specific gene expression profiles. a** Scatter plot contrasting average, filtered (log_2CPM>1, log_2RPKM>1) gene expression of colon vs. non-colon homing Treg of all BMT models (BMT control, prophylaxis, therapy). DEGs were determined under a linear model adjusting for the BMT model affiliation ($P_{adj.|FC|>1.5}$<0.05). The number of significantly upregulated (dark brown) and downregulated genes (slate blue) is indicated in the plot panel. NLT and LT signature genes as defined by Miragaia et al. (2019) are highlighted in the corresponding color of the plot title. **b** Barcode plot representing the GSEA results of the gene set upregulated in colon (light brown bars) or non-colon resident (light slate blue bars), ranked according to the log_2FC of the respective comparison in all BMT Treg. Benjamini-Hochberg adjusted *P* values are provided above or below the corresponding enrichment worm.

The accordance of the homeostatic with the transplantation models in this regard is underlined by the highly significant enrichment of colon-specific genes of the former among upregulated genes of the latter, as illustrated by the GSEA results represented

in **Figure 3.7 b**. Thus, transplanted Treg rapidly acquired the typical colon-specific gene expression imprint. Except very few induced (e.g. Ccr7) or quenched genes (e.g. Lgals1) due to transplantation, the NLT and LT gene signatures remained unaltered between baseline and BMT models regarding the colon vs. non-colon comparison axis.

3.2.3 Identification of functionally distinct co-expression networks by graph-based clustering

A more fine-grained approach for arranging genes into functional units was undertaken by a graph-based clustering of genes according to their gene-to-gene correlations across Treg groups. In anticipation of the subsequently intended projection of the graph-based clusters onto the single-cell allo Treg prophylaxis data, only bulk RNA-seq data from baseline, BMT control and prophylaxis models were considered for this. The analysis parameters were iteratively adjusted in such a way that only the main component of the graph remained, while small components with weak correlation edges were removed. As a result, a single graph with 9 well-distinguishable clusters could be determined and a two-dimensional representation thereof is depicted in **Figure 3.8 a**.

Ordering the scaled expression data of all bulk RNA-seq samples (donor, baseline, BMT control, prophylaxis, therapy) according to the graph-based clusters resulted in group, organ or model dependent arrangement of Treg when hierarchical clustering was applied (**Figure 3.8 b**).

Clusters 1, 2 and 3 were almost exclusively expressed in Treg re-isolated from the colon, with the therapy samples showing a slightly weakened gene expression (**Figure 3.8 b**). The corresponding Metascape analyses functionally linked the three gene sets to a state of activated (e.g. T cell activation, cluster 3) and differentiating Treg (e.g. lymphocyte differentiation, cluster 2) that actively produce cytokines (e.g. regulation of cytokine production, cluster 3), and at the same time employ catabolic processes (e.g. ubiquitin-dependent ERAD pathway, cluster 1), likely in response to that (**Figure 3.9 a, c, e**).

For liver and spleen-derived transplanted Treg, clusters 4 and 6 were most representative, in which a lower expression was observed in spleen homing BMT control Treg. Corresponding genes also showed a high expression in the majority of donor populations. Additionally, genes assigned to cluster 4 were moderately expressed in allo Treg homing to the colon of therapy animals, whereas transcripts from cluster 6 genes were found elevated in some prophylactic poly Treg that had migrated to the colon (**Figure 3.8 b**). The genes in these two clusters showed a strong association with processes

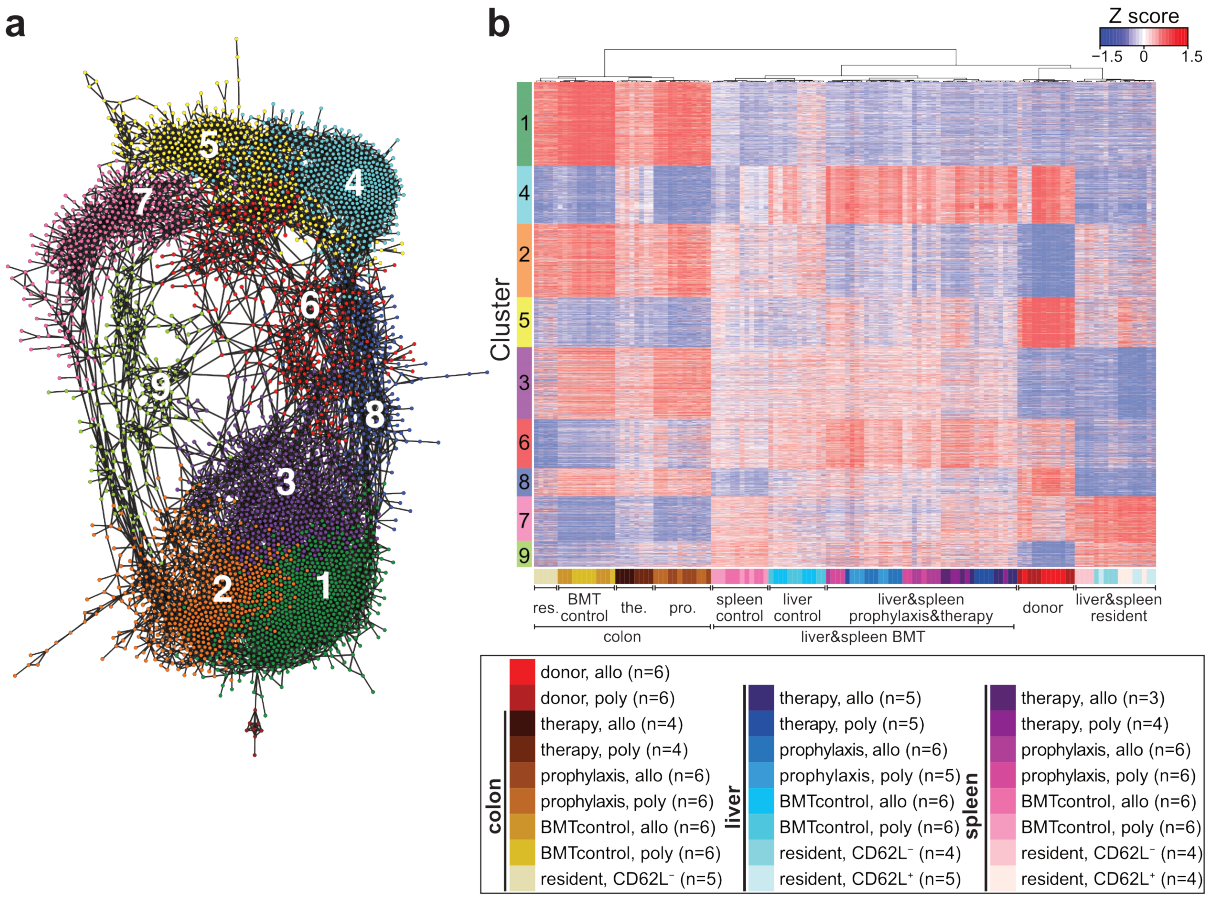

Figure 3.8: **Graph-based gene clusters separate Treg according to organ and/or model.** **a** Two-dimensional projection of the gene co-expression network generated *via* graph-based clustering (r≥0.845, k-NN using r with k>4, N(nodes)=3829, N(edges)=10.7k, Louvain cluster granularity = 0.6) of bulk RNA-seq data, excluding therapy samples. Clusters are color-coded and numbered. Nodes represent genes and edges correspond to their interrelation. **b** Heatmap representing scaled, $\log_2$-transformed expression data for clusters defined in **a**, as indicated by the row side-bar. Individual samples were hierarchically clustered as represented by the column dendrogram and the sample group is indicated by the column side bar (top panel; pro.: prophylaxis, res.: resident, the.: therapy), color-coded as described in the legend (bottom panel).

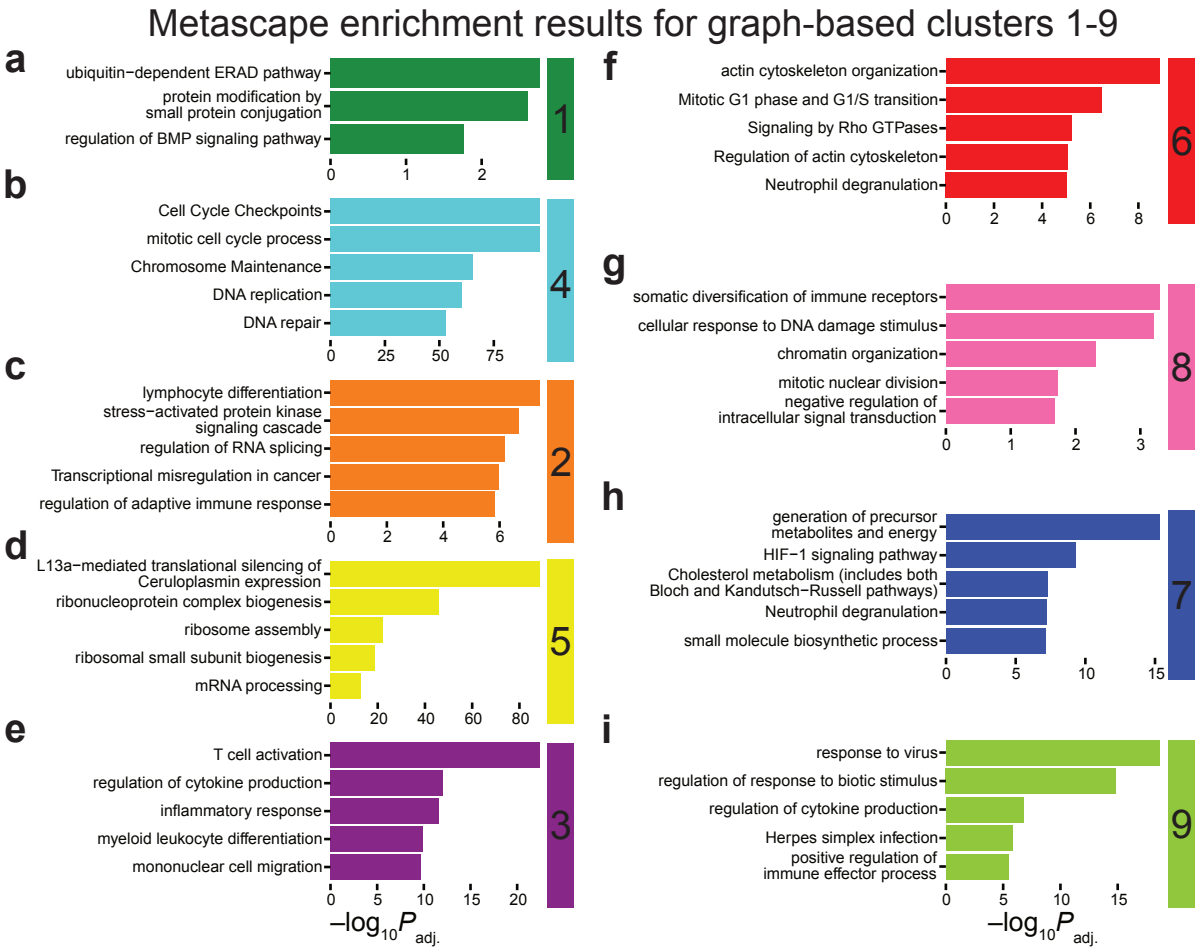

Figure 3.9: **Metascape enrichment analyses of graph-based clusters identify functional groups of pathways. a-i** For each co-expression cluster, the top five leading enriched pathways or terms are ranked according to increasing Benjamini-Hochberg adjusted *P* values ($P_{adj.}$), as determined by Metascape. Bars represent -$\log_{10}$-transformed $P_{adj.}$ values.

underlying cell proliferation (e.g. Cell Cycle Checkpoints, cluster 4; Mitotic G1 phase and G1/S transition, cluster 6), as shown in **Figure 3.9 b** and **f**. Donor Treg could further be distinguished by their characteristic expression of genes assigned to cluster 5 (**Figure 3.8 b**), which were associated with transcriptional activity (**Figure 3.9 d**).

Albeit less clearly confined, clusters 7 and 9 allotted genes were most abundant in liver and spleen-resident Treg, but were also moderately expressed in BMT Treg from same organs (**Figure 3.8 b**). In addition, donor Treg showed an overall high expression of genes in cluster 7 which were enriched for pathways associated with hypoxia (e.g. HIF−1 signaling pathway), among other terms (**Figure 3.9 h**). In contrast, cluster 9 comprised genes typically expressed upon viral infection (e.g. response to virus) (**Figure 3.9 i**).

Elevated levels of another set of genes belonging to cluster 8 were found in donor Treg,

and to a lesser extent in colon homing Treg from the BMT control and prophylaxis models (**Figure 3.8 b**). The corresponding representative pathways reflected processes deployed during immune receptor diversification (e.g. somatic diversification of immune receptors), potentially involving somatic recombination (e.g. cellular response to DNA damage stimulus, chromatin organization) (**Figure 3.9 g**).

The identified clusters were rich in TFs and prompted a more focused analysis. To this end, for the previously used order of clusters, several TFs were chosen and the corresponding gene expression data is represented as heatmap in **Figure 3.10 a**.

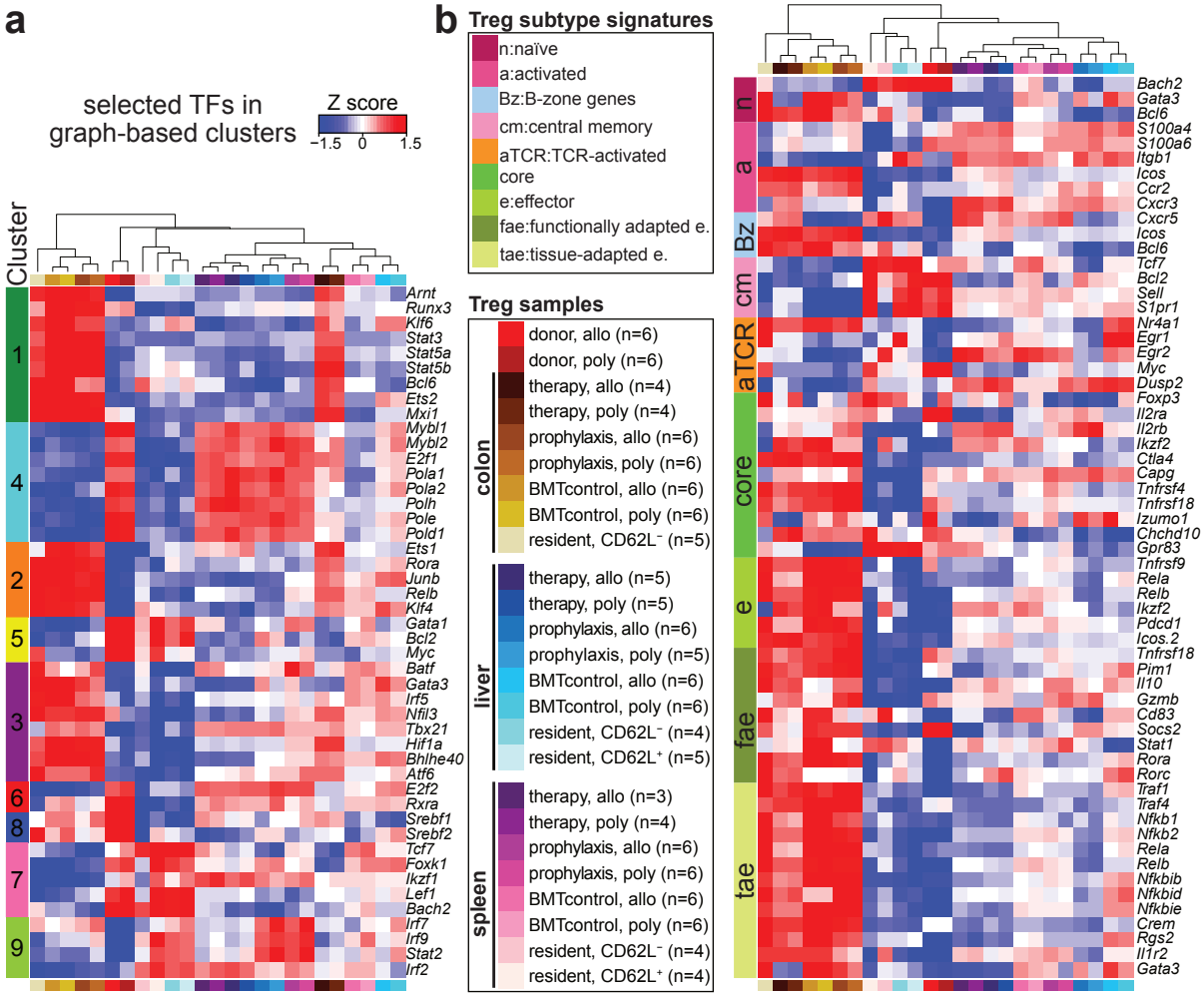

Figure 3.10: **Selected transcription factors of graph-based clusters and Treg subtype signatures. a,b** Scaled, group-average $\log_2$-transformed expression data represented as heatmaps according to the Zscore color key. Treg sample group affiliation is color-coded by bilateral column sidebars as shown in the Treg samples legend panel. Samples were hierarchically clustered. **a** Selected TFs of graph-based clusters 1-9, indicated by the row sidebar. **b** Treg subtype gene signatures as summarized by Shevyrev & Tereshchenko (2020), indicated by the row sidebar with color code according to the Treg subtype signature legend panel.

The hierarchical clustering of the samples based on the selected TFs clearly indicated

a close relationship between prophylaxis and therapy samples derived either from liver or spleen. Also, liver and spleen resident Treg displayed a homogeneous expression of the selected TFs. While resident and transplanted Treg from BMT control and prophylaxis models formed an expected cluster, therapeutically administered Treg also showed features akin to BMT control Treg from liver or spleen, in part owed to the expression of replication-associated TFs in cluster 4 (e.g. *Mybil2*, *E2f1*, *Pola1*) as well as the lack of *Gata3* and *Batf* expression, genes typically expressed in other colon Treg, along with further genes in cluster 3. The genes in cluster 4 were most distinctive for donor Treg and GvHD model-derived samples from liver and spleen. The T cell activation-associated TFs of clusters 1 (e.g. *Runx3*, *Stat3*, *Stat5a*, *Bcl6*), 2 (e.g. *Rora*, *Junb*, *Relb*) and 3 (e.g. *Irf5*, *Tbx21*, *Hif1a*) were well-represented in most colon-derived Treg, with cluster 3 showing a weaker and less uniform expression in therapy Treg.

As a complementary approach for functional characterization, the Treg subtype signatures summarized by Shevyrev & Tereshchenko (2020), empowered published single-cell analyses, were consulted, and consequently, previous data was also projected onto the according sets of genes (**Figure 3.10 b**). All colon-derived Treg showed a mostly homogeneous and strong expression of genes that are typical of effector Treg. The divergent activation stimuli either Treg in the colon or liver and spleen-homed BMT Treg were exposed to were reflected in their almost mutually exclusive activation signature. Genes indicating the spatiotemporal proximity to B cells, such as *Cxcr5*, *Icos* or *Bcl6* were most prominent in colon Treg. However, the chemokine receptor *Cxcr5* was selectively diminished in BMT control and prophylaxis Treg, likely as an early consequence of the irradiation. Donor cells and Treg residing in unperturbed livers and spleens could be associated with a central memory phenotype, however, the latter displayed decreased expression levels of core Treg genes which where on the other hand characteristic for most colon-derived Treg. While *Nr4a1* was also strongly expressed in the latter, Treg that had migrated to BMT control livers displayed a strong expression signature that typically follows TCR activation, indicative of a frequent antigen encounter.

In summary, the gene sets identified *via* graph-based clustering organized Treg according to distinct functional gene expression programs, mostly coinciding with their organ and/or model affiliation. Hence, Treg re-isolated from the colon were generally characterized by their immunological activity, whereas proliferation was induced in cells found in liver or spleen. In addition, the usage of single-cell based gene signatures allowed the approximation of the phenotypical composition in Treg samples derived from the *in vitro* and *in vivo* models.

3.3 Systematic dissection of *in vivo* models

3.3.1 Irradiation damage and early reconstitution phase

In order to delineate the impact of the disease models on Treg gene expression, systematic DGE analyses were individually performed for colon, liver and spleen. Thus, first of all, the comparison of the BMT control Treg with physiologically tissue-resident Treg was undertaken to cover the detrimental effects of the irradiation as well as the events in the early reconstitution phase after BMT. The resulting DEGs are visualized in **Figure 3.11**, with the previously defined suppression signature highlighted. The comparison of Treg re-isolated from the colon yielded a substantial number of up- (N=372) and downregulated (N=298) genes, and the GSEA for genes induced by BMT control treatment identified a pathway associated with glycolysis and gluconeogenesis (WP157) as significantly enriched. On the other hand, genes lost as compared to the homeostatic state were significantly enriched in a gene set that is indicative of cytokine-cytokine receptor interaction (mmu04060). The top-ranking differential genes for the aforementioned pathways were highlighted accordingly (**Figure 3.11 a**). The highest number of upregulated DEGs (N=1036) upon BMT control treatment were identified in Treg that could be isolated from the liver, and the most significantly associated functional gene set was involved in RNA polymerase I promoter opening (R-MMU-73728), suggesting the induction of proliferation. Genes implicated in interferon-beta response (GO: 0035458) were among the genes that were more abundant (N=622) in the baseline model (**Figure 3.11 b**). In analogy to the liver, Treg that had migrated to the spleen of BMT control animals upregulated (N=212) hallmarks of cellular senescence (R-MMU-2559583), a pathway known from a strong proliferative response. The downregulated (N=59) genes in Treg found in the spleen could not be significantly linked to a functional pathway (**Figure 3.11 c**). The majority of suppression signature genes was upregulated in liver-homing Treg, suggesting the reinforced exertion of cytolytic (*Gzma*,*Gzmb*,*Prf1*), immunosuppressive (e.g. *Il10*, *Tgfb1*, *Ebi3*) function along with cell-contact-dependent inhibitory mechanisms (*Ctla4*, *Fasl*, *Pdcd1lg1*).

3.3.2 Adapted Treg gene expression in the presence of Tcon

The systematic analysis covered the differences specifically evoked by the presence of co-transplanted Tcon in the prophylaxis and therapy models by virtue of comparing them to the BMT control model. First, Treg from the prophylaxis model were considered, and, the induced DEGs (N=189) in colon-homing Treg were significantly associated with the regulation of leukocyte activation (GO: 0002694), while genes involved in cytokine-cytokine receptor interaction (mmu04060) had reduced expression levels

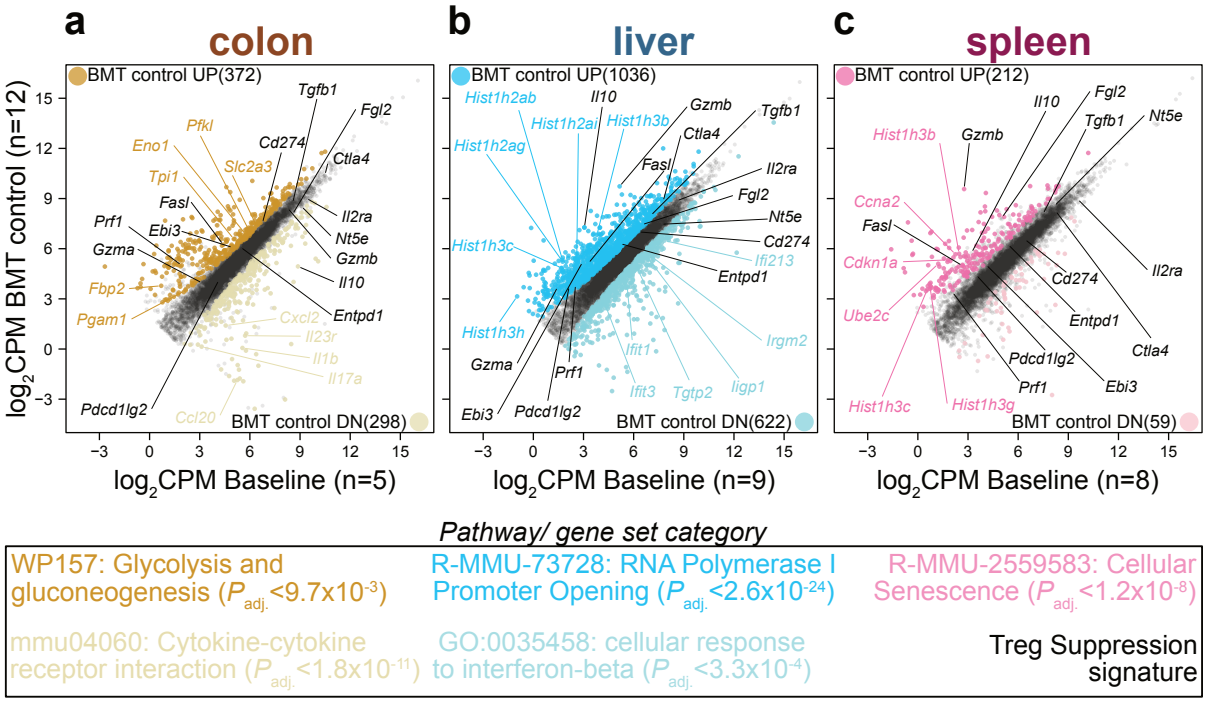

Figure 3.11: **Metabolic or proliferative activity of Treg is induced in early BM reconstitution.** Scatter plots contrasting average, filtered (log_2CPM>1, log_2RPKM>1) gene expression of Treg derived from BMT control animals with baseline data. Genes with $P_{adj.|FC|>1.5}$<0.05 were considered significant and were colored according to the embedded legends, indicating the corresponding number of DEGs. Selected, top-ranking GSEA-identified pathway genes were labelled, corresponding to the description in the bottom legend panel. Benjamini-Hochberg adjusted enrichment *P* values ($P_{adj.}$) are provided for each significant pathway, as determined by Metascape. Genes associated with suppression by Treg are highlighted in black. **a** Colon (n=5-12), **b** liver (n=9-12), and **c** spleen (n=8-12) data are shown separately.

(N=195) (**Figure 3.12 a**). This pathway was already decreased in Treg from BMT control animals as compared to the baseline model. The most prominently enriched pathway among upregulated genes in prophylactically transferred Treg homing to the liver (N=48) was interferon-beta response (GO: 0035456) and was indicative of a reconstituted homeostatic gene expression. In contrast, the IL-17 signaling pathway was significantly linked to the downregulated genes (N=107), as depicted in **Figure 3.12 b**. Prophylactic Treg derived from the spleen were highly associated with mitotic cell cycle (R-MMU-69278), showing a considerable number of significantly upregulated genes (N=505). Similar to the colon, BMT control Treg infiltrating the spleen featured a relative increase in genes required for cytokine-cytokine receptor interaction (mmu04060) (**Figure 3.12 c**). The suppression signature was almost completely upregulated in Treg that were challenged with Tcon in the colon, i.e. *Gzma*, *Gzmb*, *Prf1*, *Il10*, *Tgfb3*, *Entpd1*, *Ctla4* and *Fasl* showed a relative increase in expression level. Moreover, the cytolytic pathway along with *Fgl2* or *Il10* were induced in liver or spleen-homing Treg,

respectively.

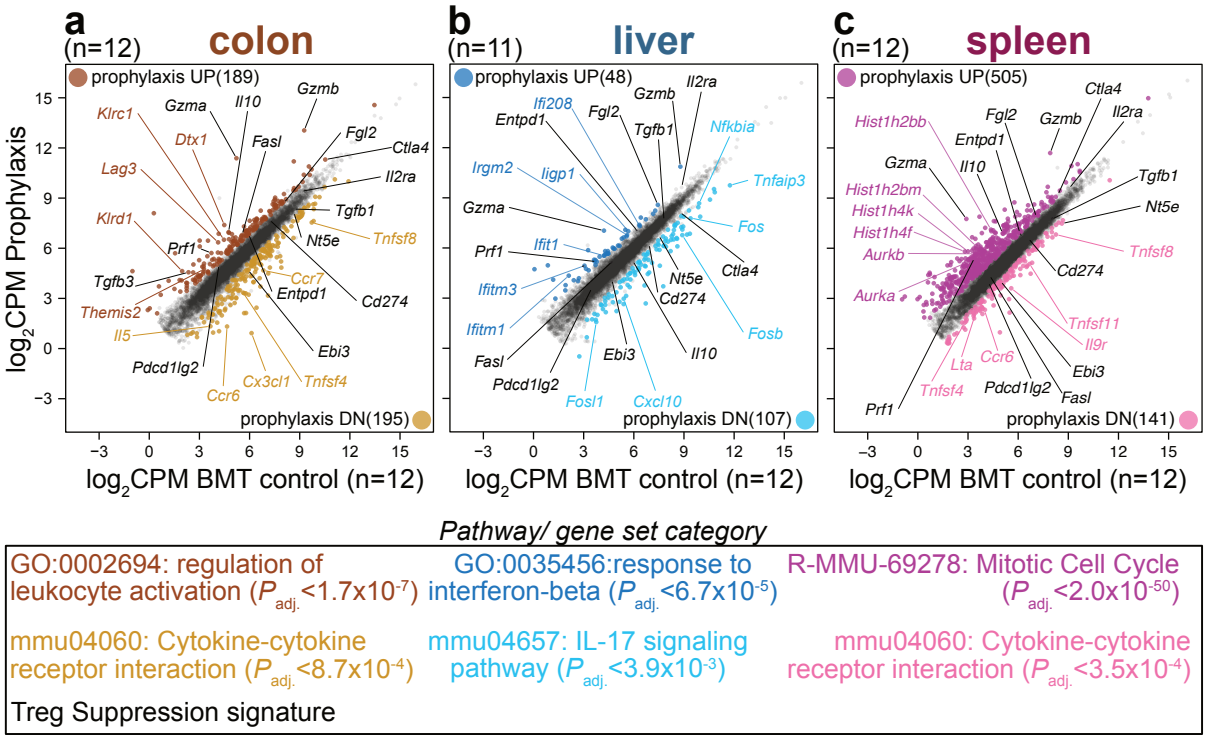

Figure 3.12: **Defense mechanisms and replication shape gene expression in Treg for prophylaxis of alloreactive T cell damage.** Scatter plots representing average, filtered (log_2CPM>1, log_2RPKM>1) gene expression of Treg derived from prophylaxis contrasted with the BMT control model. Genes with $P_{adj.|FC|>1.5}$<0.05 were considered significant and were colored according to the embedded legends, indicating the corresponding number of DEGs. Selected, top-ranking GSEA-identified pathway genes were labelled, corresponding to the description in the bottom legend panel. Benjamini-Hochberg adjusted enrichment *P* values ($P_{adj.}$) are provided for each significant pathway, as determined by Metascape. Genes associated with suppression by Treg are highlighted in black. **a** Colon (n=12), **b** liver (n=11-12), and **c** spleen (n=12) data are shown separately.

The analogous comparison of therapy to BMT control Treg yielded DEGs of a similar magnitude, with the largest number of up- (N=361) and downregulated genes (N=442) found in colon-infiltrating Treg. The latter were associated with the pathway of inflammatory response (GO: 0006954) (**Figure 3.13 a**).

The therapeutically administered Treg that had migrated to the liver showed a relative increase in catalytic activity (GO: 0043086) when compared to BMT control Treg, featuring a total of N=206 downregulated genes (**Figure 3.13 b**). The recurring cytokine-cytokine receptor interaction pathway (mmu04060) was enriched within the genes found downregulated (N=129) in spleen-homing therapy Treg, likely due to decaying inflammation, in analogy to the observations in Treg isolated from the colon early after BMT (**Figure 3.13 c**). Treg from all three analyzed organs paralleled each other in terms of a

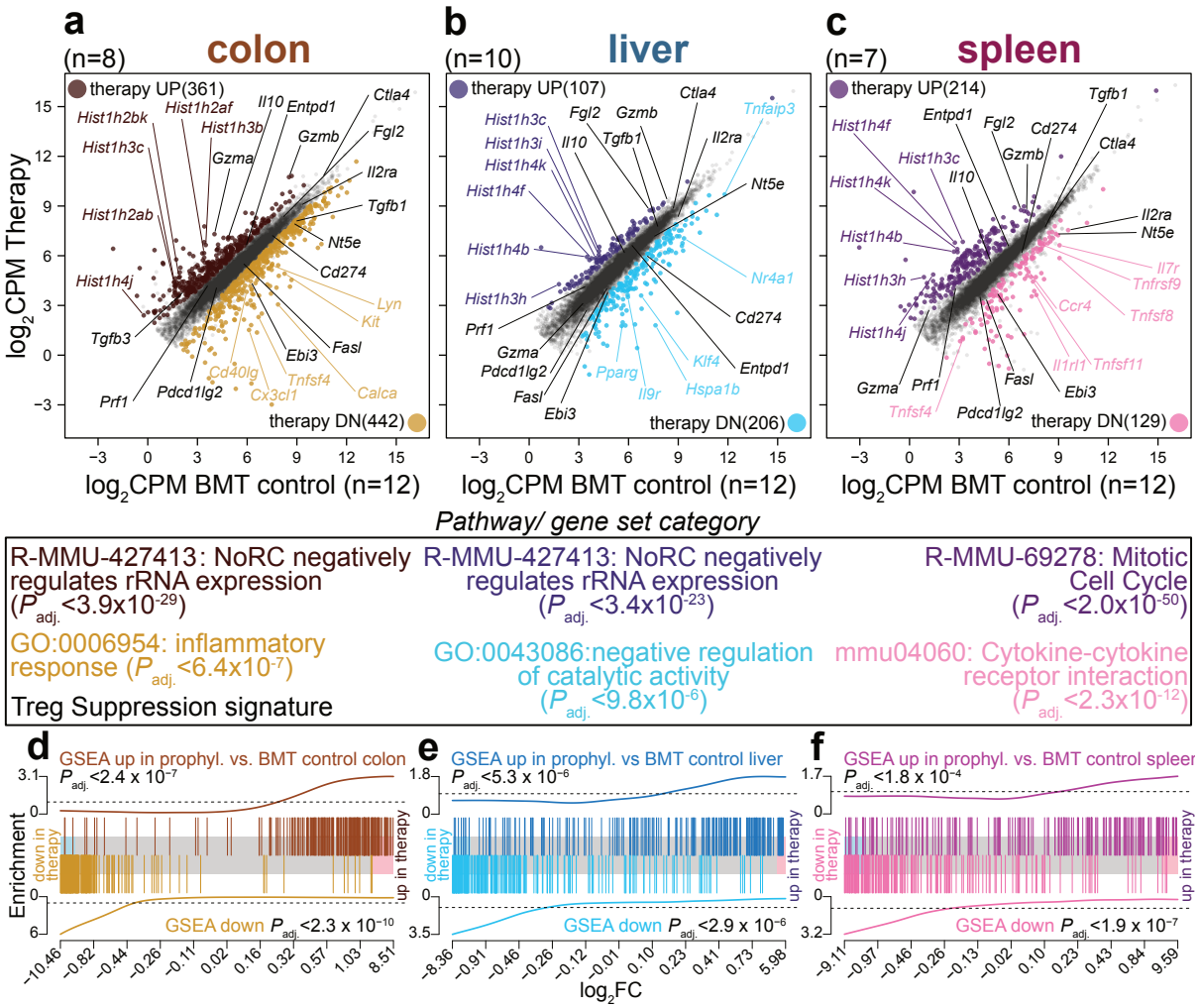

Figure 3.13: **Therapeutic transfer of Treg favors proliferation over immune responses. a-c** Scatter plots representing average, filtered (log_2CPM>1, log_2RPKM>1) gene expression of Treg derived from therapy contrasted with BMT control model. Genes with $P_{adj.|FC|>1.5}$<0.05 were considered significant and were colored according to the embedded legends, indicating the corresponding number of DEGs. Selected, top-ranking GSEA-identified pathway genes were labelled, corresponding to the description in the bottom legend panel. Benjamini-Hochberg adjusted enrichment *P* values ($P_{adj.}$) are provided for each significant pathway, as determined by Metascape. Genes associated with suppression by Treg are highlighted in black. **a** Colon (n=8-12), **b** liver (n=10-12), and **c** spleen (n=7-12) data are shown separately. **d-e** Organ-wise barcode plots representing the GSEA results of the gene sets upregulated (darker shades) and downregulated (lighter shades) in prophylaxis vs. BMT control, ranked according to the log_2FC of the therapy vs. BMT control comparison. Enrichment was determined by two-sided rotation gene set tests and each plot is provided with the corresponding Benjamini-Hochberg adjusted *P* value.

dominant replication signature, as genes upregulated in colon and liver (N=107) were involved in the remodelling complex-mediated regulation of rRNA expression (R-MMU-427413). Similarly, therapy Treg homed to the spleen featured upregulated genes (N=214) that were strongly associated with mitotic cell cycle (R-MMU-69278). The suppressive Treg signature was unequally induced in therapy Treg, and in the spleen, even a partial decrease was observed, i.e. *Il2ra*, *Nt5e* and *Ebi3* were among the downregulated genes, as opposed to the prophylaxis vs. BMT control comparison.

Despite apparent differences between the two Tcon-based models in contrast to the BMT control approach, parallels between the induced genes in prophylaxis and therapy were sought for. GSEA on the differentially regulated genes in prophylaxis was performed, ranking genes according to the differences found between therapy and BMT control Treg. For all three tissues, a significant enrichment of the organ-specifically up- and downregulated gene sets was observed, most prominent in the colon (**Figure 3.13 d, e** and **f**). Hence, a partial parallelism between the two model systems could be identified.

3.3.3 Treg are immunologically more active in prophylaxis than in therapy

The comparison between therapy and prophylaxis Treg was performed as a logical next step. Only in Treg from the colon, upregulated genes (N=32) allowed the identification of a significantly enriched pathway, namely negative regulation of rRNA expression by NoRC (R-MMU-427413) which is associated with proliferation. Corresponding downregulated genes were associated with immunoregulatory interactions between lymphoid and non-lymphoid cells (R-MMU-198933) (**Figure 3.14 a**).

Downregulated genes in liver (N=112) and spleen (N=87) were associated with gene ontology terms relating to the regulation of innate immune response (GO: 0045088) and the humoral immune response (GO: 0006959), respectively, indicating the relative preponderance of immune-related pathways in the prophylaxis Treg as compared to therapy. In line with that, therapy Treg derived from the liver showed significantly diminished expression levels of the suppressive genes *Il2ra*, *Gzma*, *Gzmb* and *Fasl*, with similar tendencies observed in Treg from spleens (**Figure 3.14 b** and **c**).

Besides the elevated immunological activity of prophylactic Treg, as determined by the above systematic approach, certain key molecules for homing, cytokine signalling and antimicrobial defense were more dominantly expressed in the therapy model, as summarized in **Figure 3.15**. After infiltration of the colon and liver of therapy animals, Treg expressed markedly higher levels of *Cxcr5*, required for homing to B cell zones, and

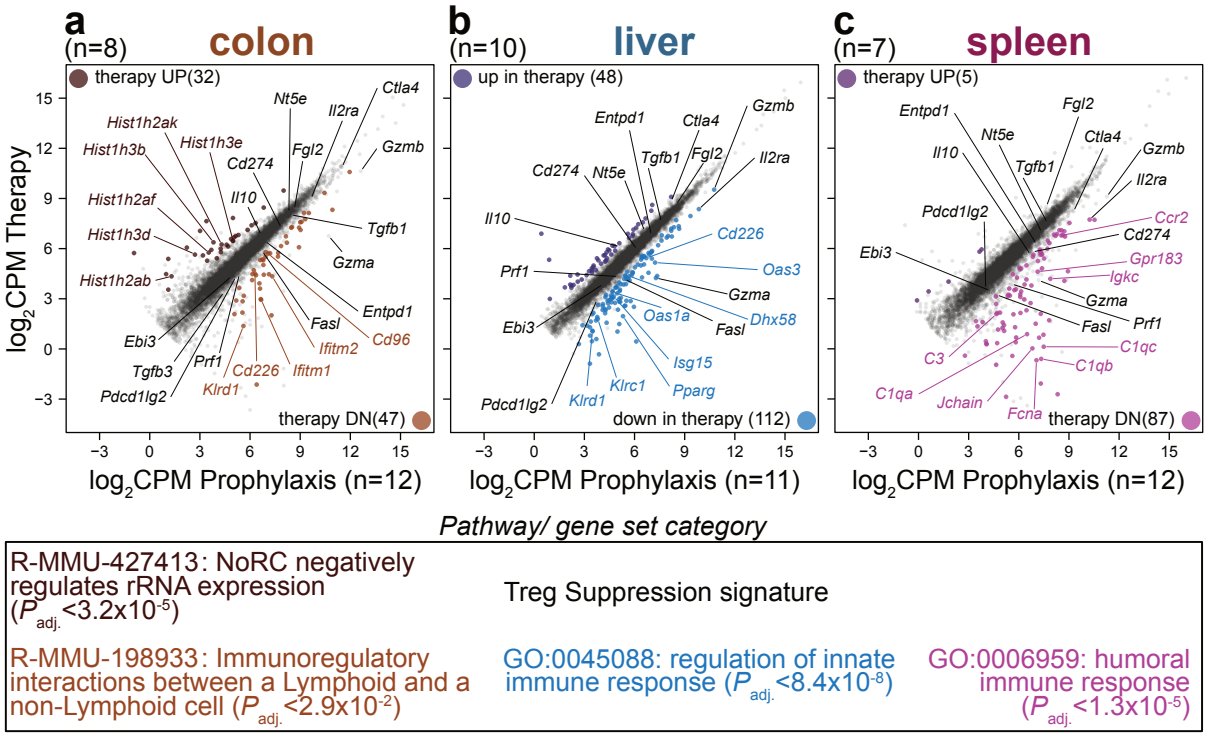

Figure 3.14: **Prophylactic Treg present with overall immunological superiority to therapy.** Scatter plots representing average, filtered (log_2CPM>1, log_2RPKM>1) gene expression of Treg derived from therapy contrasted with the prophylaxis model. Genes with $P_{adj.|FC|>1.5}$<0.05 were considered significant and were colored according to the embedded legends, indicating the corresponding number of DEGs. Selected, top-ranking GSEA-identified pathway genes were labelled, corresponding to the description in the bottom legend panel. Benjamini-Hochberg adjusted enrichment *P* values ($P_{adj.}$) are provided for each significant pathway, as determined by Metascape. Genes associated with suppression by Treg are highlighted in black. **a** Colon (n=8-12), **b** liver (n=10-11), and **c** spleen (n=7-12) data are shown separately.

the integrin *Itgb8*. Additionally, the cytokine receptor genes *Tnfrsf7* (*Cd27*) and *Il1r2*, a hallmark of tissue adaption, were more highly expressed in therapeutically administered Treg. Another observation was the dominance of the *Tcrb* variable segment *Trbv30*, a finding to be corroborated with the help of TCR repertoire data that will be analyzed later (**Figure 3.15 a**).

The integrin *Itgae*, involved in the homing to the intestinal mucosa, as well as *Ccr7* and *Sell* (CD62L), required for lymph node homing, showed a strong upregulation in therapeutic Treg that had seeded to the colon. The gene *Reg3b* was found elevated, which encodes an antimicrobial peptide that can also negatively regulate the translocation of bacteria from the intestinal lumen to epithelial layers. Of note, its expression levels were comparable to BMT control mice (**Figure 3.15 b**). The IL-6 signal transducer (*Il6st*) was representative for the reconstitution of a pathway that seems to play

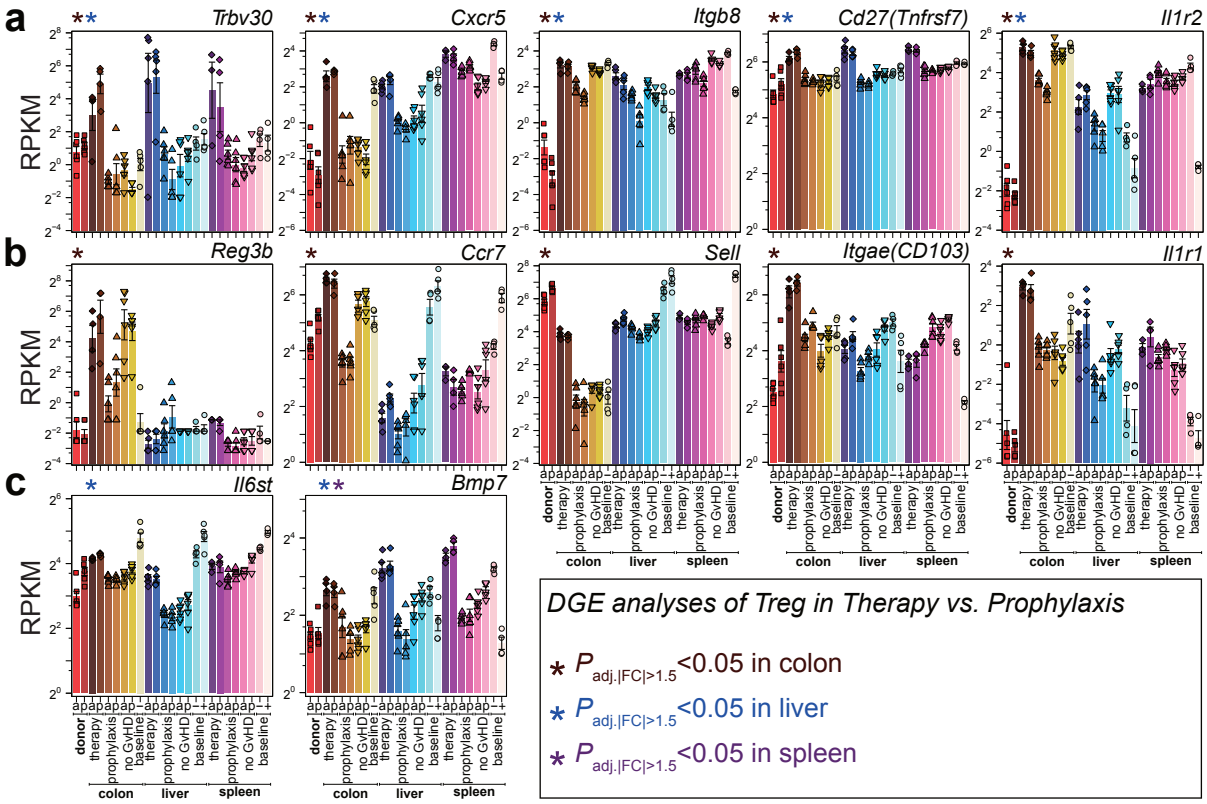

Figure 3.15: **Therapeutic Treg show specific homing and immunological signalling potential.** Gene expression levels of genes upregulated in therapeutic Treg when compared to prophylaxis, represented as bar plots. Asterisks indicate a $P_{adj.|FC|>1.5}$<0.05 and are colored according to the descriptive legend in the bottom right panel. Bars represent RPKM means±SE for n=4-6 independent experiments on a log2-scaled y-axis. Individual data points are shown as dots with shapes according to the corresponding *in vitro* or *in vivo* model.

an important role also in the homeostatic state, while the soluble TGF-β ligand *Bmp7* was significantly elevated in spleen and liver-homing Treg of the therapy approach, as compared to the prophylaxis, also showing a moderate induction in the BMT control model (**Figure 3.15 c**).

Despite the overall loss of immunologically relevant gene expression in therapy Treg, certain key homing genes were found induced as compared to the prophylaxis, indicating highly specific processes dependent on the time point of Treg transfer. In part, the gene expression of certain molecules was in line with Treg from the BMT control or baseline models, generally hinting at the reversal of the high extent of tissue damage in the prophylaxis model.

3.3.4 Suppressive potential of Treg across organs and models

In connection with the previously GSEA-identified differential functional gene sets, hereafter, the suppressive gene signature was extended and categorized according to section 1.3.1 for a more direct interpretability. As depicted in **Figure 3.16**, physical APC blockade *via Nrp1* or *Itgb2* was almost exclusive for liver and spleen, and moreover, *Nrp1* differed between allo and poly Treg.

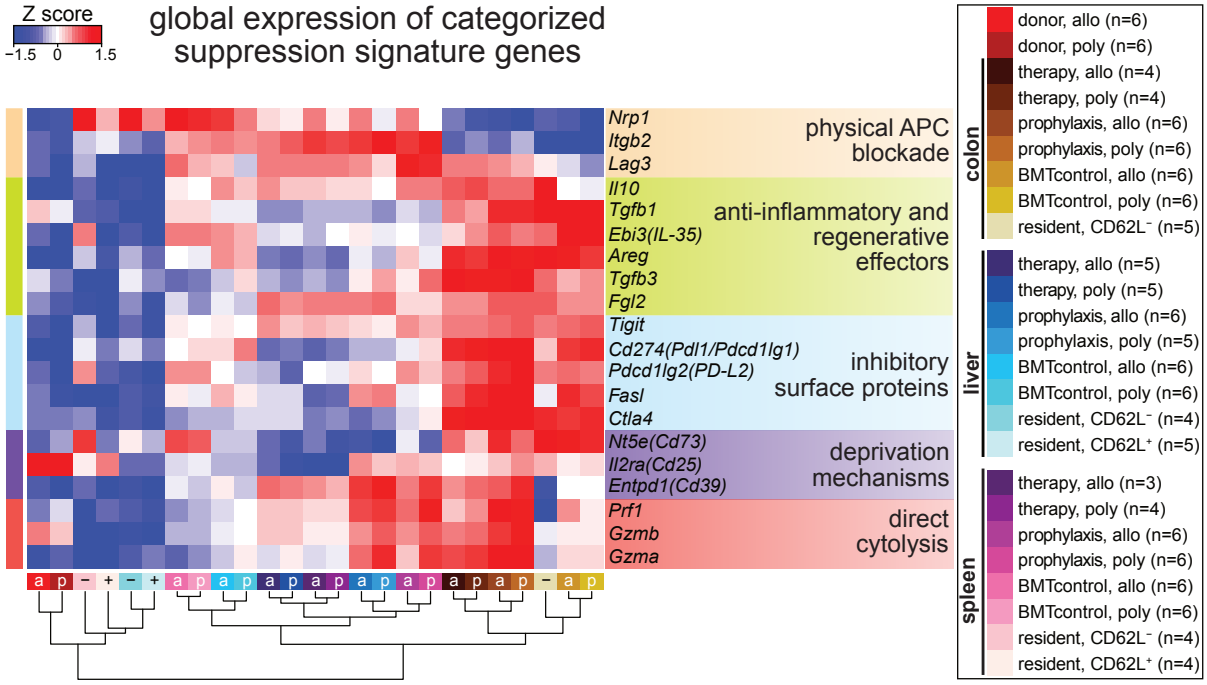

Figure 3.16: **Global categorized suppression signature gene expression elucidates differential usage of mechanism between organs and models.** Scaled, group-average $\log_2$-transformed expression data represented as heatmap according to the Zscore color key. Treg sample group affiliation is color-coded by the column sidebar as shown in the Treg samples legend panel. Samples were hierarchically clustered, as indicated by the dendrogram. The row sidebar indicates the category of suppressive genes, as described in the corresponding colored panels.

At first sight, the relative increase of anti-inflammatory, secreted effectors along with inhibitory surface proteins in Treg from the colon derived from all settings became apparent. In addition, the increased usage of *Fgl2* and *Tigit* by disease-burdened Treg from liver and spleen could be observed.

For an overview of the absolute expression levels of previously introduced gene categories of suppression-related genes, **Figure 3.17** compares all Treg sample categories with regard to that.

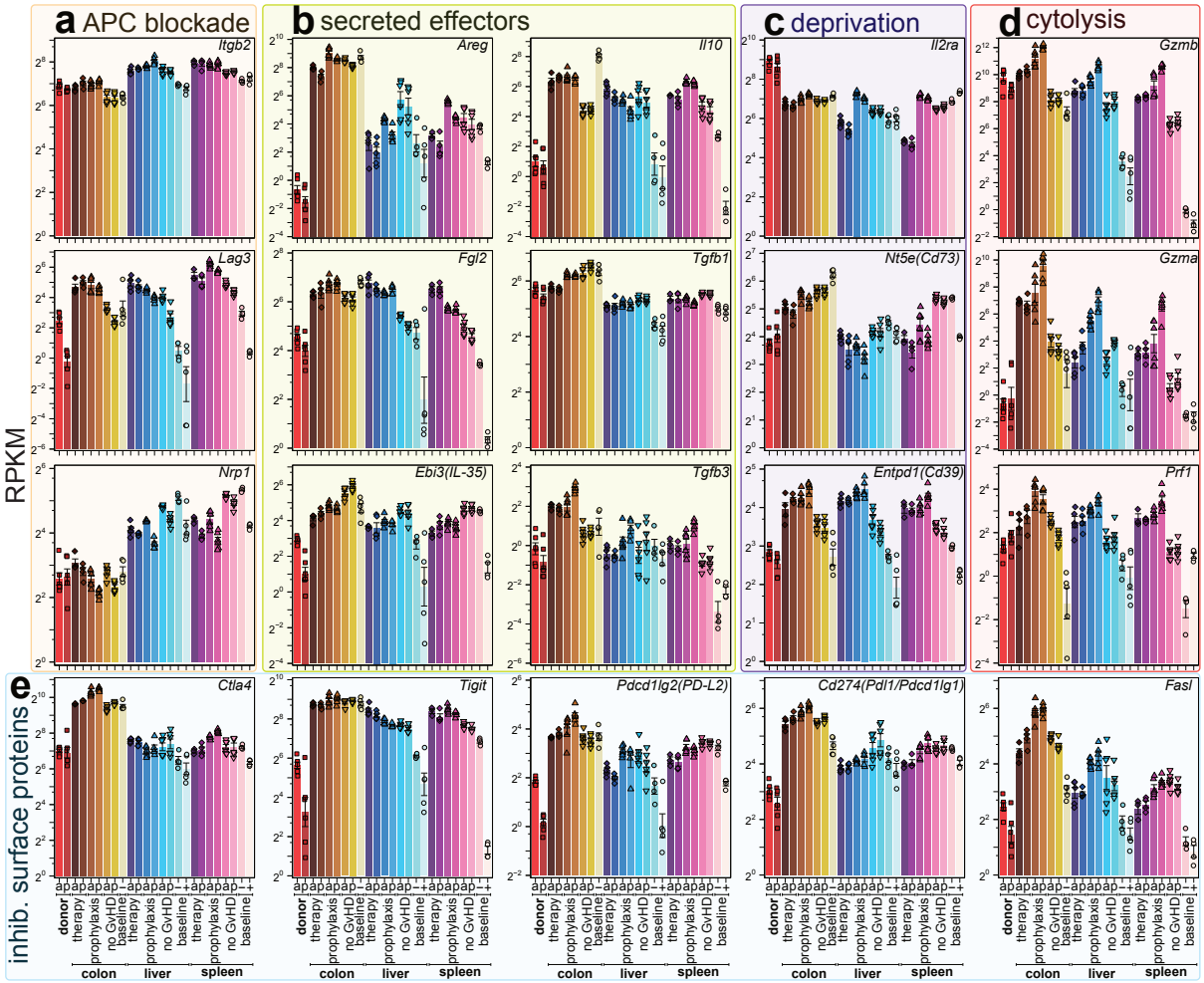

Figure 3.17: **Majority of genes encoding suppressive effector molecules are induced in Treg upon BMT.** Bar plots for Treg gene expression of functionally grouped suppression-related genes, enclosed by boxes colored according to the categories as defined in Figure 3.16. **a** Genes directly associated with physical hindrance of APC interaction. **b** Genes encoding secreted effector molecules for anti-inflammation and tissue homeostasis. **c** Deprivation mechanisms for IL-2 and ATP. **d** Granzyme-perforin-conveyed cytolysis. **e** Inhibitory surface molecules. Bars represent RPKM means±SE for n=4-6 independent experiments on a log2-scaled y-axis. Individual data points are shown as dots shaped corresponding to *in vitro* or *in vivo* model affiliation.

3.4 Investigation of Treg TCR repertoires *in vivo*

The gene expression of Treg in various settings of aGvHD has been thoroughly described. So far, however, only *in vitro* expanded allo and poly Treg were commented on with respect to their TCR repertoire. In the following, overlap and clonotype distribution analyses for all Treg categories derived from tissues are covered. Of note, also single chain information, either *Tcra* or *Tcrb*, will be referred to as clonotypes concerning all bulk TCRrepseq data despite no layer of information of that data type covers the pairing of alpha and beta chains. Moreover, only *Tcrb* data will be presented for the purpose of clarity.

3.4.1 *Tcrb* repertoire overlap and tissue distribution in BMT

First, the BMT control *Tcrb* repertoires of Treg re-isolated from colon, liver and spleen were compared to the repertoire of the respective transplanted donor population. The fraction of donor *Tcrb* repertoires overlapping with clonotypes observed in re-isolated Treg was larger when allo Treg were transferred to BMT control mice, in line with the observed narrowing of the repertoire by allogeneic *in vitro* expansion (section 3.1.1). In contrast, the fraction of *Tcrb* organ repertoires covered by the donor population was comparable between poly and allo Treg, and also inter-organ overlaps were relatively homogeneous (**Figure 3.18 a**, **b**).

When the most frequent clonotypes (top 500) were analyzed with respect to their distribution between the organs of all three recipient mice of independent experiments, a relatively even distribution became apparent, in which allo Treg presented more homogeneously distributed than poly Treg (**Figure 3.18 c**, **d,** leftmost column). Similarly, the organ-wise distribution of the top 250 *Tcrb* clonotypes across recipient mice pointed out that the most frequent beta chains were shared between individuals which was more obvious in allo Treg recipients (**Figure 3.18 c**, **d**, three rightmost columns).

An almost identical impression was elicited by the analogous visualization of Treg *Tcrb* data that were derived from prophylaxis animals. However, the extent of inter-organ overlaps tended to be bigger (**Figure 3.19 a**, **b**).

Furthermore, the most frequent *Tcrb* transcripts were shared between colon, liver and spleen as well as between individuals in the organ-stratified analysis, and in analogy to the BMT control model, the homogeneous nature of the clonotype distribution was more apparent in allo Treg recipients (**Figure 3.19 c**, **d**).

The observations regarding the TCR repertoires of BMT control and prophylaxis poly and allo Treg were widely recapitulated in therapeutically administered Treg (**Figure 3.20**). However, the availability of data was limited for experiment 2 because in each

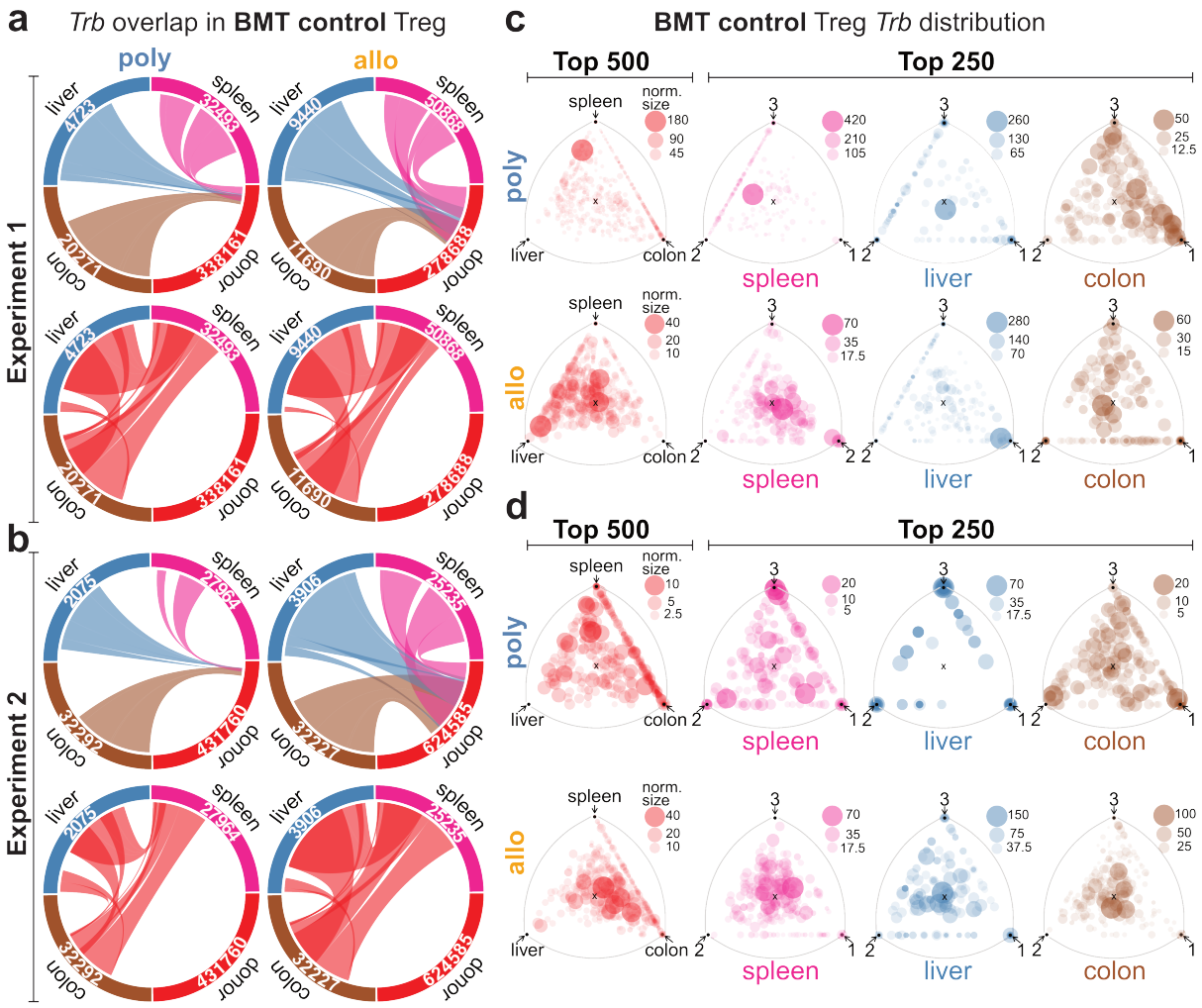

Figure 3.18: **Major *Tcrb* repertoire overlaps in allo Treg of BMT control mice. a**,**b** Poly and allo Treg *Tcrb* overlaps between the donor and re-isolated Treg populations (top panels) or between colon, liver and spleen (bottom panels) are represented as circos plots for BMT control experiment 1 (**a**) and 2 (**b**). The width of the bands corresponds to the fraction of the *Tcrb* repertoire involved in overlaps and the total repertoire sizes are indicated by the number of UMIs in each quadrant. **c**,**d** Barycentric triangle plots representing the most frequent poly and allo *Tcrb* clonotypes and their distribution between organs (top 500; red) or the organ-wise (top 250; brown, blue, purple) distribution across recipients (n=3) are shown for BMT control experiment 1 (**c**) and 2 (**c**). The bubble size is indicative of the normalized size of the *Tcrb* clonotype.

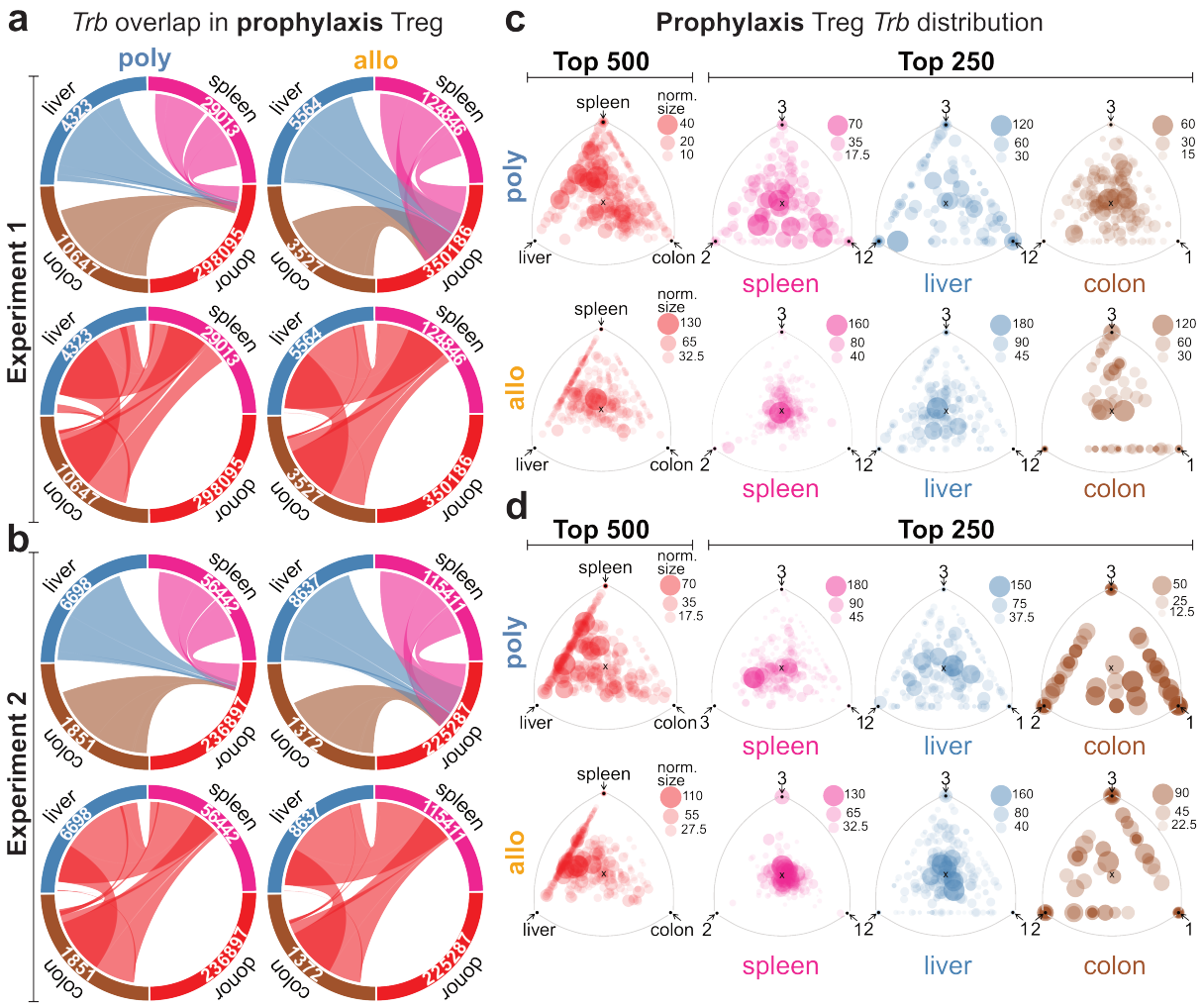

Figure 3.19: **Major *Tcrb* repertoire overlaps in allo Treg of prophylaxis mice. a**,**b** Poly and allo Treg *Tcrb* overlaps between the donor and re-isolated Treg populations (top panels) or between colon, liver and spleen (bottom panels) are represented as circos plots for prophylaxis experiment 1 (**a**) and 2 (**b**). The width of the bands corresponds to the fraction of the *Tcrb* repertoire involved in overlaps and the total repertoire sizes are indicated by the number of UMIs in each quadrant. **c**,**d** Barycentric triangle plots representing the most frequent poly and allo *Tcrb* clonotypes and their distribution between organs (top 500; red) or the organ-wise (top 250; brown, blue, purple) distribution across recipients (n=3) are shown for prophylaxis experiment 1 (**c**) and 2 (**c**). The bubble size is indicative of the normalized size of the *Tcrb* clonotype.

group (allo and poly), one animal predeceased their peers (**Figure 3.20 c**, **d**, three rightmost columns). Also, the number of transplanted Tconv was higher (500 k) than in experiment 1 (250 k), and hence, experiment 2 must be seen as an aggravated therapy model rather than a perfect replicate of experiment 1.

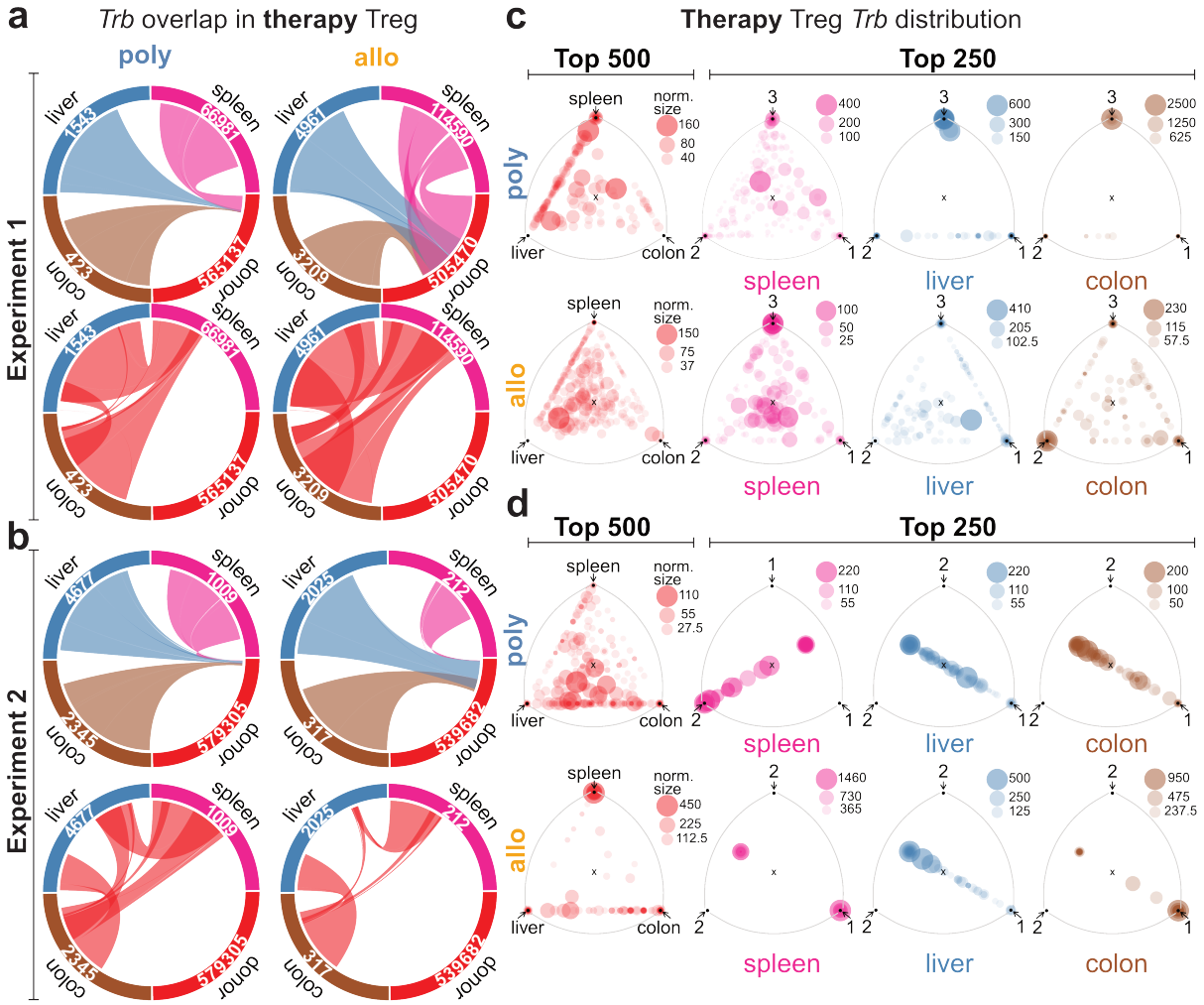

Figure 3.20: **Major *Tcrb* chain overlap in allo Treg of therapy mice.** **a**,**b** Poly and allo Treg *Tcrb* overlaps between the donor and re-isolated Treg populations (top panels) or between colon, liver and spleen (bottom panels) are represented as circos plots for therapy experiment 1 (**a**) and 2 (**b**). The width of the bands corresponds to the fraction of the *Tcrb* repertoire involved in overlaps and the total repertoire sizes are indicated by the number of UMIs in each quadrant. **c**,**d** Barycentric triangle plots representing the most frequent poly and allo *Tcrb* clonotypes and their distribution between organs (top 500; red) or the organ-wise (top 250; brown, blue, purple) distribution across recipients (n=2-3) are shown for therapy experiment 1 (**c**) and 2 (**c**). The bubble size is indicative of the normalized size of the *Tcrb* clonotype. For experiment 2 (**c**), data was only available for 2 individuals and was plotted twice in part.

In summary, the TCR repertoire data derived from Treg of all BMT models supported the observation that frequently detected clonotypes were shared between organs as

well as the organs of individual recipients which was more clearly indicated for the allo as compared to poly Treg groups.

Next, a TCR set enrichment analysis (TSEA) was performed based on the *Tcrb*s shared between the poly and allo donor Treg populations and the colon, liver and spleen-derived Treg from the respective recipients. A consistently stronger enrichment of allo Treg recipient-derived, donor-shared *Tcrb*s was observed as compared to the corresponding poly recipients, when clonotypes were ranked according the their frequency in the original graft. While a moderate enrichment characterized colon, liver and spleen infiltrating poly Treg in independent BMT control (**Figure 3.21 a**), prophylaxis (**Figure 3.21 c**) and therapy (**Figure 3.21 e**) experiments, the corresponding allo Treg were markedly more enriched for the *Tcrb* clonotypes expanded during *in vitro* culture (**Figure 3.21 b**, **d**, **f**).

These findings supported the assumption that organ infiltration is indeed mainly driven by the alloreactivity of transplanted Treg.

3.4.2 Differential usage of selected variable segments

In general, allo and poly Treg displayed a higher diversity prior to transplantation as compared to ones re-isolated from colon, liver and spleen. While BMT control and prophylactically transferred Treg were relatively comparable with respect to their *Tcrb* diversity, Treg from therapy animals tended to be characteristic of an even less divers repertoire (**Figure 3.22 a**).

Various clusters of data points corresponding to specific groups of Treg could be observed, when a t-distributed stochastic neighbor embedding (t-SNE), a means of reduced dimensionality, was generated based on *Trbv* usage data (**Figure 3.22 b**). One cluster, for instance, was occupied by poly Treg that had homed to the colon of BMT control and prophylaxis animals. Also, the donor populations were represented by independent clusters corresponding to their *in vitro* expansion. Another dense cluster was formed by liver and colon infiltrating poly Treg of the therapy model. Of note, the allo Treg re-isolated from therapy mice contributed to the cluster of allo BMT Treg homing to lymphoid tissues. In contrast, the allo Treg from BMT control and prophylaxis animals were observed to be in close proximity to poly Treg that were found in the liver and spleen. The group of spleen-directed poly Treg belonging to the therapy model did not show a specific cluster affiliation. The differential expression of the variable segment gene *Trbv16* between allo and poly Treg, as observed in section 3.2.1, was consistently more frequent in allo Treg across BMT and *in vitro* models when the respective repertoire data was consulted (**Figure 3.22 c**). Similarly, *Trbv30* was found upregulated in therapy Treg compared to prophylaxis (section 3.3.3). In line with that,

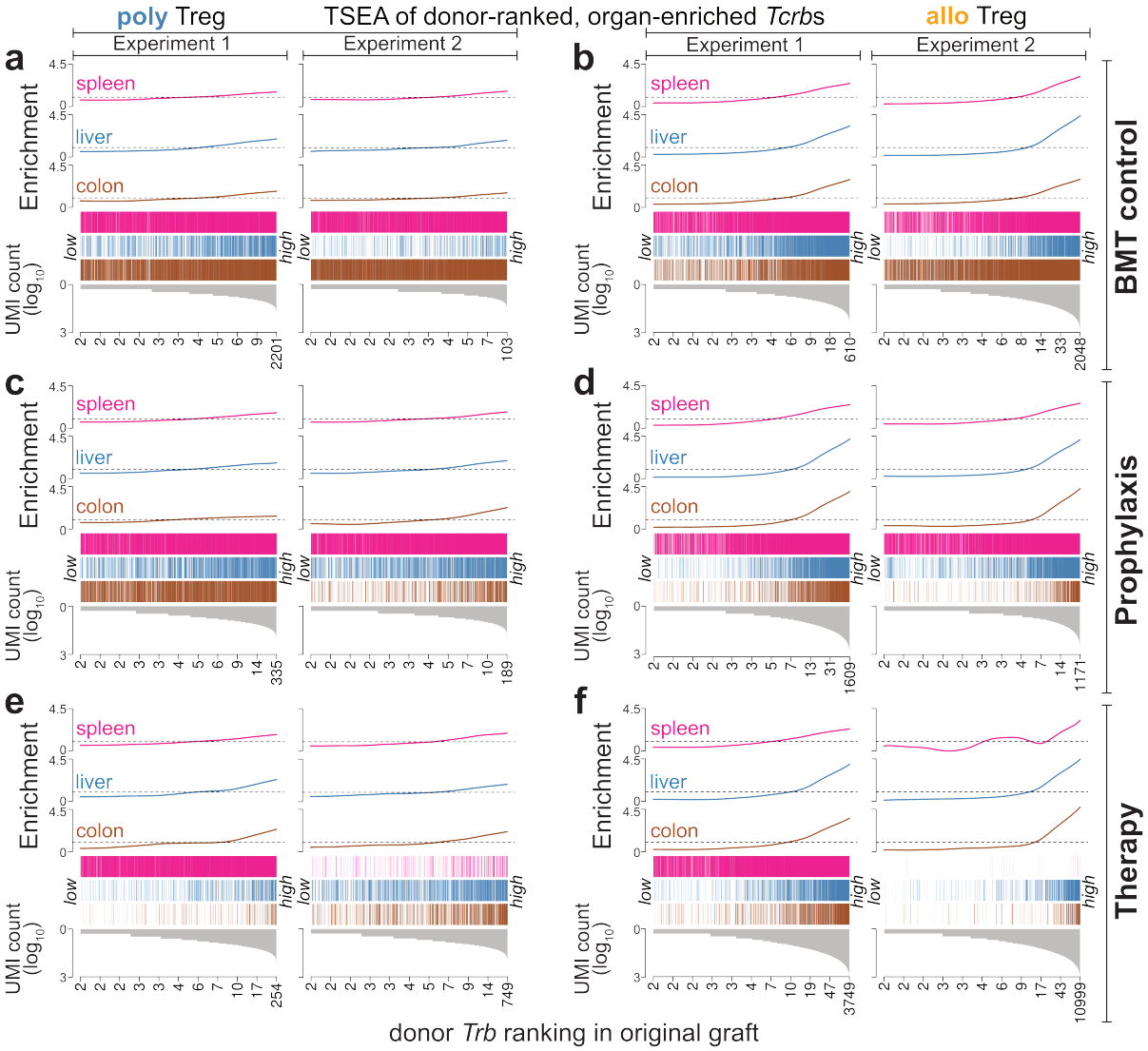

Figure 3.21: **TSEA of top-ranking, organ-shared donor Treg *Tcrb*s shows increased enrichment in allo BMT models. a-f** Barcode plots representing TSEA results of the organ-wise enrichment of organ-shared poly donor Treg (**a**,**c**,**e**) and allo donor Treg (**b**,**d**,**f**) TCR beta chain transcripts, ranked according to the frequency in the respective original graft (high/low). For the BMT control (**a**,**b**), the prophylaxis (**c**,**d**) and the therapy models (**e**,**f**), two independent experiments are shown. Bars represent *Tcrb* clonotypes and are colored according to the corresponding enrichment worms (brown: colon, blue: liver, purple: spleen). The frequency in the original graft is represented by by the $\log_{10}$-transformed UMI counts at the bottom of each panel.

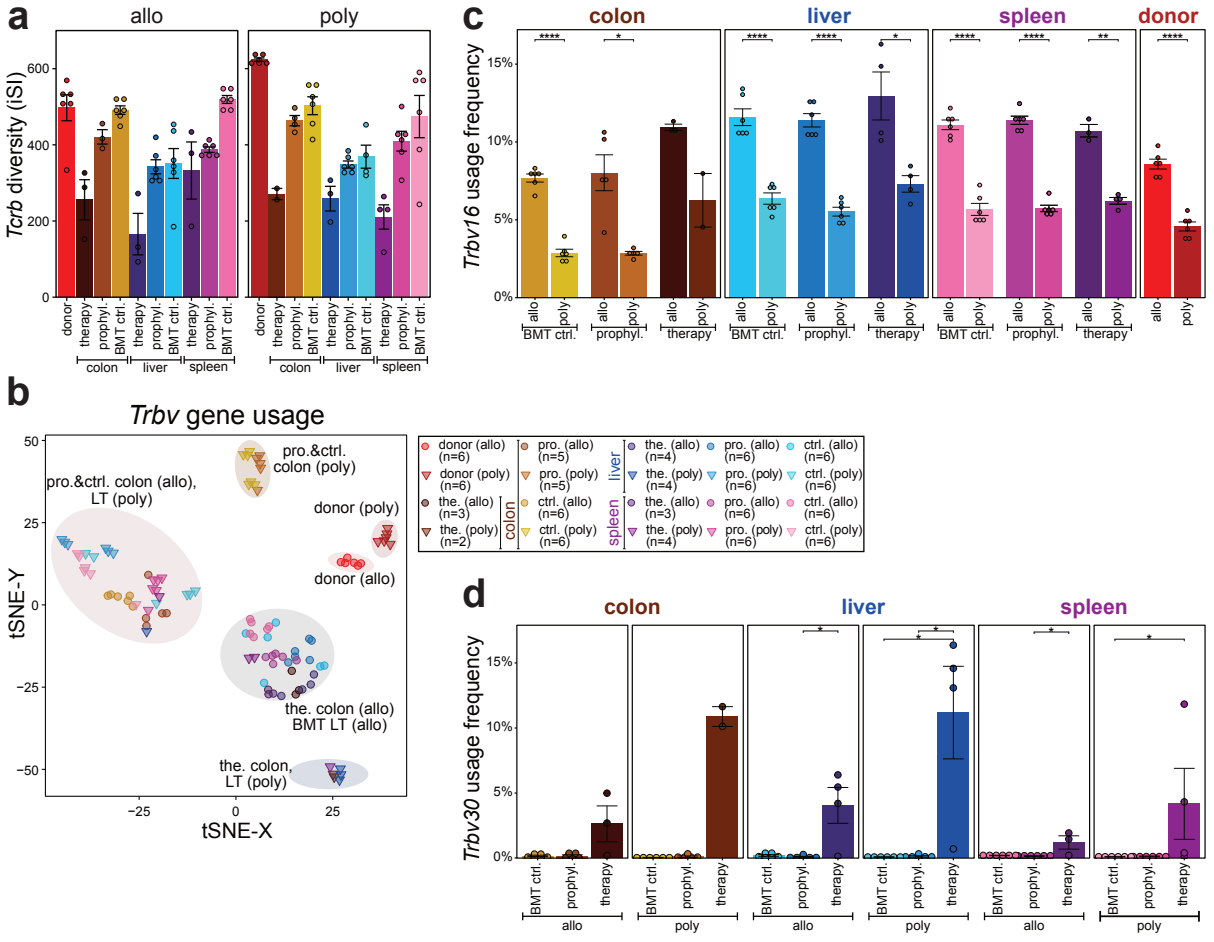

Figure 3.22: **Differential *Trbv* gene usage corroborates differences found in gene expression analyses. a** Barplots representing the mean *Tcrb* diversity (inverse Simpson Index, iSI) ±SE for n=2-6 replicates of allo and poly Treg prior to transplant (n=6) and after re-isolation from colon, liver and spleen (n=2-6). Individual data points are shown as dots. **b** T-SNE embedding of a distance matrix of *Trbv* segments across allo and poly Treg after *in vitro* expansion (donor) or after re-isolation from colon, liver and spleen of recipients. The shape of the individual data points corresponds to poly (triangle) or allo (circle) Treg, and colors correspond to the BMT model and organ, as shown in the legend panel. **c**, **d** Barplots representing the mean *Trbv16* (**c**) and *Trbv30* (**d**) usage frequencies ±SE for n=2-6 independent allo and poly Treg samples across BMT models. Individual data points are shown as dots, and significant differences are indicated by brackets annotated with asterisks corresponding to categorized *P* values (*: *P*<=0.05; **: *P*<=0.01; ***: *P*<=0.001; ****: *P*<=0.0001). **c** Organ- and model-wise comparisons of allo *versus* poly Treg were determined by a two-tailed t-test. **d** For each organ and expansion method**,** therapy Treg were compared to other BMT models by a Wilcoxon signed-rank test. BMT ctrl., ctrl.: BMT control model; prophyl., pro.: prophylaxis model; the.: therapy model; LT: lymphoid tissues.

this variable segment was almost exclusively observed in *Tcrb* repertoires of therapeutically administered Treg which was more apparent in poly Treg (**Figure 3.22 d**).

Another set of variable segments (*Trbv12-1*, *Trbv12-2*, *Trbv26)* stood out for its preponderance in BMT Treg that were re-isolated from BMT control and prophylaxis animals, and featured overall higher usage frequencies in poly Treg (**Figure 3.23 a**).

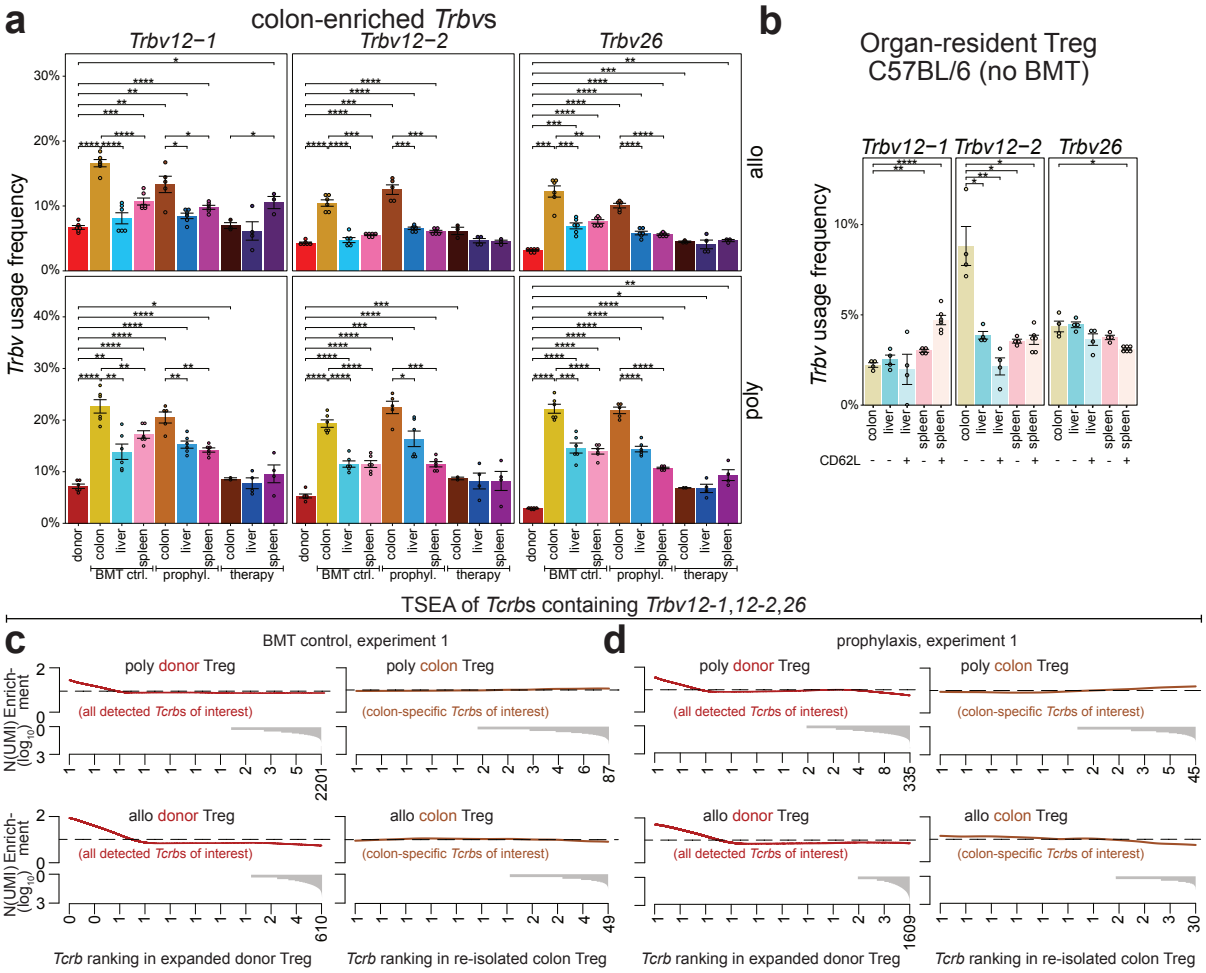

Figure 3.23: **Differential *Trbv* gene usage reveals colon-enriched variable segments.** **a**, **b** Barplots representing the mean *Trbv12-1*, *12-2*, *26* usage frequencies ±SE for n=2-6 independent allo and poly Treg samples across BMT models (**a**) and for n=4-6 homeostatic Treg samples (**b**). Individual data points are shown as dots, and significant differences as determined by a two-tailed t-test are indicated by brackets annotated with asterisks corresponding to the categorized *P* values (*: *P*<=0.05; **: *P*<=0.01; ***: *P*<=0.001; ****: *P*<=0.0001). **c**, **d** TSEA of *Tcrb*s with *Trbv12-1*, *12-2*, *26* for poly (upper panels) and allo (lower panels) Treg. One representative experiment is shown for the BMT control model (**c**) and prophylaxis (**d**). Enrichment results for the donor populations (red, left panels) applying the *Tcrb* ranking of all donor-detected, *Trbv* of interest-containing *Tcrb*s and the colon populations (brown, right panels) according to the *Tcrb* ranking of the tested set intersecting colon-detected *Tcrb*s of interest are shown above the respective $\log_{10}$-transformed UMI counts (N(UMI)).

Only one of the three variable segments (*Trbv12-2*) was characteristic for colon-resident Treg of the baseline model (**Figure 3.23 b**). The latter might be frequently incorporated in *Tcrb*s responding to antigens that are ubiquitous in the colon, while *Trbv12-1* and *Trbv26* could play a role in the recognition of antigens exposed early after the disruption of tissue by irradiation.

The TCR beta chain transcripts that contained any of the three previously introduced variable segments (*Tcrb*s of interest) and were observed in colon-homing Treg were relatively homogeneously distributed across the corresponding *Tcrb* ranking, as illustrated by the TSEA enrichment analyses in BMT control (**Figure 3.23 c**) and prophylaxis Treg (**Figure 3.23 d**). When all *Tcrb*s of interest were tested for an enrichment along the ranking of donor-detected *Tcrb*s of interest in the respective donor populations, a negative enrichment became apparent. Hence, such clonotypes were infrequently expanded *in vitro* and selectively accumulated in the colon.

3.5 aGvHD allo Treg prophylaxis on the single-cell level

The preceding bulk analyses provided insights into tissue-dependent gene expression and TCR beta chain repertoire changes under homeostatic and perturbed conditions. Next, a single-cell experiment of prophylactically transferred allo Treg was performed in order to explore the pairing of TCR alpha and beta chains into productive clonotype repertoires as well as to decipher the composition of sorted cells with regard to subpopulations and their quantitative contribution. Another layer of comparisons, i.e. differences between Treg populations with distinct specificities, was enabled by the integration of both data types in cellular resolution.

3.5.1 Recapitulation of rapidly tissue-adapted gene expression profiles

The rapid organ-specific adaption of Treg gene expression was the main finding in the single-cell allo prophylaxis approach, and thus corroborated previously presented observations. Consequently, colon-homing Treg formed an exclusive cluster after dimensionality reduction *via* UMAP. Treg re-isolated from spleen and liver occupied a large cluster with widely overlapping contributions from both organs. The *in vitro* expanded allo donor Treg were clearly separate from all *in vivo* samples in the same embedding. Global projections of the single-cell gene expression data distinguishing 13 variegated clusters (large panel), and representing Treg origins (smaller panels) are provided in **Figure 3.24 a**. Clusters 10 and 13 were excluded due to contaminating myeloid (cluster 10) and erythroid (cluster 13) gene signatures.

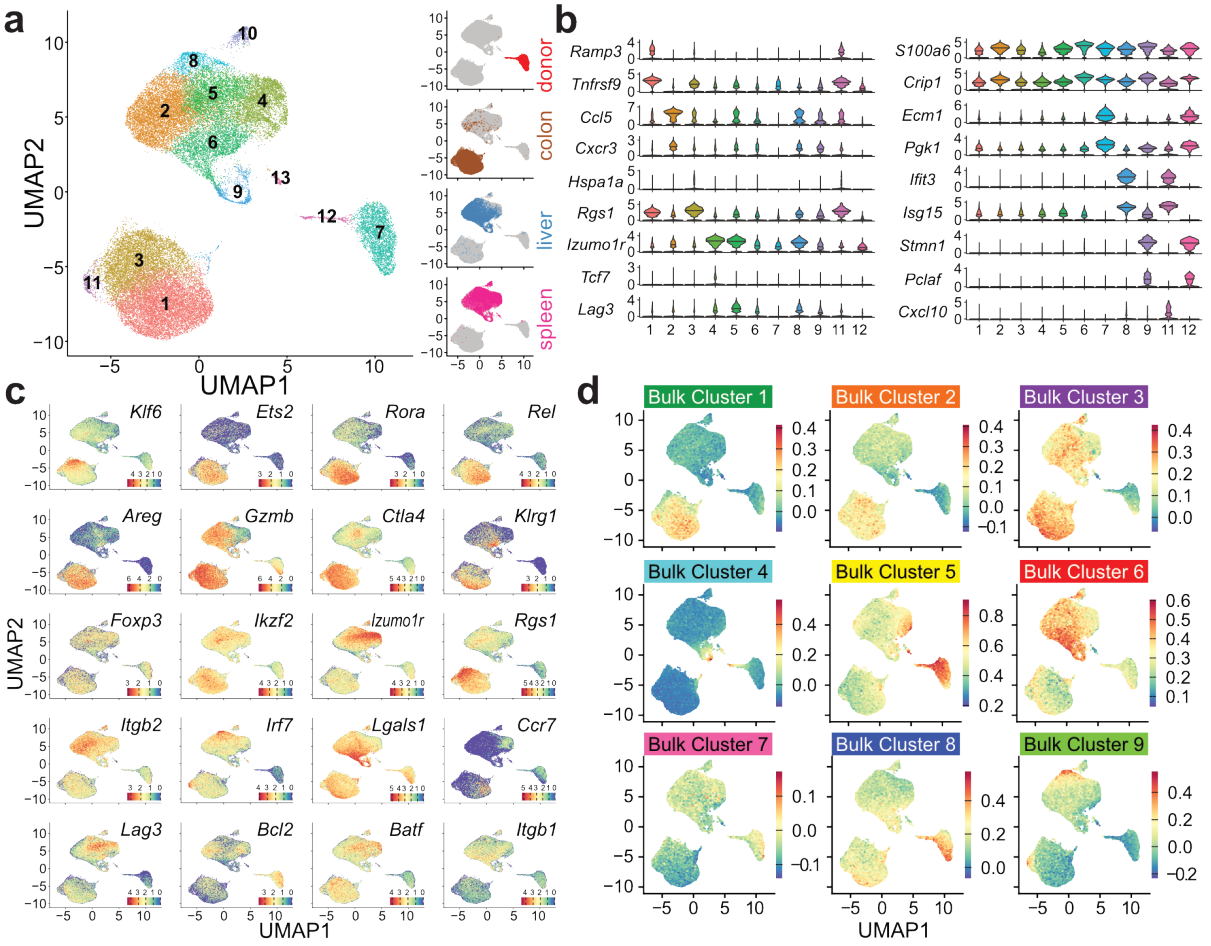

Figure 3.24: **Single-cell gene expression of prophylactic allo Treg recapitulates tissue adaption and reveals subpopulations.** **a** UMAP embedding of 37,457 allo donor Treg prior to prophylaxis (donor) and after re-isolation from colon, liver and spleen of n=3 animals, highlighted accordingly in the smaller representations of the plot. The bigger panel on the left shows the 13 differently colored clusters identified by unsupervised Louvain-clustering. The contaminating clusters 10 and 13, containing myeloid and erythroid signature genes, respectively, were excluded from further analyses. **b** Violin plots representing the $\log_2$-scaled expression distribution of cluster-defining genes for single-cell clusters 1-9, 11 and 12. **c,d** Log_2-scaled gene expression of selected marker genes (**c**) or bulk co-expression clusters (**d**) corresponding to the individual color scales within or next to plot panels and based on the same embedding as in **a**.

The gene expression levels for leading subpopulation-describing genes of clusters 1-9, 11 and 12 are provided as violin plots, colored according to the global UMAP plot (**Figure 3.24 b**). The effector Treg marker *Tnfrsf9* (4-1BB) was most common among colon-homing Treg (clusters 1,3 and 11). While the Treg activation marker *Cxcr3* was relatively limited to clusters 2, 8 and 9, occupied by liver and spleen infiltrating Treg, the expression of *S110a6* defined cluster 6, independent of the target organ.

The interrogation of selected marker genes allowed the functional distinction of the global clusters found in the single-cell gene expression data (**Figure 3.24 c**). The TFs *Klf6*, *Ets2*, *Rora* and *Rel*, for instance, showed the highest expression levels in either cluster 1 (*Klf6*), cluster 3 (*Rora*, *Rel*) or both colon clusters (*Ets2*), very much in line with this previously observed characteristic expression in colon Treg across all *in vivo* models analyzed *via* bulk sequencing. The genes *Ctla4*, *Areg* and *Gzmb*, typical of a NLT Treg signature, were found upregulated in colon-homing Treg of all *in vivo* models on the bulk level, arguing for tissue imprint strong enough not to be disrupted by irradiation or additional Tcon challenge. In accordance with that, the single cell clusters 1 and 3 comprised cells with a homogeneous high expression of all three markers, with Gzmb and Ctla4 also expressed in liver and spleen Treg, to a lower extent. In addition the colon-typical *Klrg1* was dominantly expressed in cluster 1, i.e. in roughly half of the Treg infiltrating the colon. A small subpopulation within cluster 6 was found expressing high levels of *Klrg1*, as well. While the core Treg markers *Foxp3* and *Ikzf2* were relatively uniformly expressed across single-cell clusters, *Izumo1r* was more prevalent in liver and spleen whereas *Rgs1* was found to be dominant in colon Treg.

The integrin encoding gene *Itgb2* was predominantly found in clusters 2,5 and 8, occupied by liver and spleen-derived Treg, and was found almost exclusively expressed in liver and spleen Treg of the BMT models on the bulk level. Analogously, *Irf7* was most characteristic of liver and spleen infiltrating Treg in prophylaxis but was also moderately expressed in colon samples, apart from the therapy model. *Irf7* was characteristic for clusters 8 (liver/spleen) and 11 (colon) while the expression in other clusters was low. The NLT signature gene *Lgals1* was more abundantly expressed in colon Treg of the baseline model, as previously described. However, under any of the pertubation models, it was found upregulated in liver and spleen infiltrating Treg. In line with that, *Lgals1* had the strongest signal in clusters 6 and 9, exclusively occupied by liver and spleen Treg, with a relatively high overall expression level. Also, in line with bulk data, the LT signature gene *Ccr7*, which was expressed the highest in liver and spleen resident Treg of baseline animals, but upregulated in colon homing Treg in BMT models, clearly stood out by means of the high signal in single-cell cluster1. Furthermore, *Lag3* was induced in the bulk of prophylactically administered Treg infiltrating spleen, and in line

with that, LT clusters 4 and 5 showed the strongest expression on the single-cell level. The central memory Treg marker *Bcl2* was predominantly expressed in donor as well as CD62L^{+} Treg residing in the liver and spleen of healthy mice, while still being moderately expressed in non-colon or liver homing Treg of the BMT control or prophylaxis models, respectively. Analogously, clusters 2, 4 and donor cluster 7 were characterized by the strongest *Bcl2* expression among analyzed single cells. In line with a fairly uniform expression in bulk Treg other than baseline, *Batf* was characteristic for single-cell colon clusters (clusters 1 and 3) as well as clusters 4 and 5, occupied by Treg re-isolated from liver and spleen. The integrin gene *Itgb1* was a identified as a feature of liver Treg as well as BMT model spleen Treg in the bulk approaches. In accordance with these findings, single-cell clusters 2,4 and 6 displayed the highest expression of *Itgb1*. A lot of important marker genes could be used to draw parallels between the gene expression analyses on the bulk level and the more detailed singe-cell approach.

When the single-cell data was highlighted according to the abundance of genes that were identified as co-expressed in the bulk analyses (section 3.2.3), clearly defined populations within the UMAP embedding could be demarcated (**Figure 3.24 d**). The bulk clusters 1 and 2, for instance, coincided with single-cell clusters 1,3 and 11, characterizing Treg re-isolated from the colon. Similarly, bulk cluster 5 described the single-cell donor population, and bulk cluster 6, typical of transplanted Treg infiltrating liver and spleen, exclusively covered the entirety of single-cell clusters belonging liver and spleen. While bulk cluster 9 was not as feasible for distinguishing bulk samples, the respective gene signature was specific for a small but immunologically highly active subpopulation of Treg that could be found in colon (cluster 11), liver and spleen (cluster 8) of prophylaxis animals.

A detailed overview over the single-cell cluster-typical gene expression patterns, including the description of the contaminating clusters 10 (macrophage) and 13 (erythrocyte), is provided in **Figure 3.25 a**. Of note, all three animals equally contributed to the entirety of the gene expression data, represented by recipient stratification of the data (**3.25 b**).

3.5.2 Common and colon-enriched clonotypes

Besides the detailed description of subpopulations enabled by single-cell gene expression analysis, the TCR repertoires of prophylactic allo Treg could be investigated at the clonotype resolution. As shown in **Figure 3.26 a**, a similar extent of repertoire overlap between the allo Treg donor and the re-isolated populations became apparent for each analyzed organ. Roughly half of the TCRs observed in colon, liver and spleen participated in an overlap with a combination of the respective other compartments.

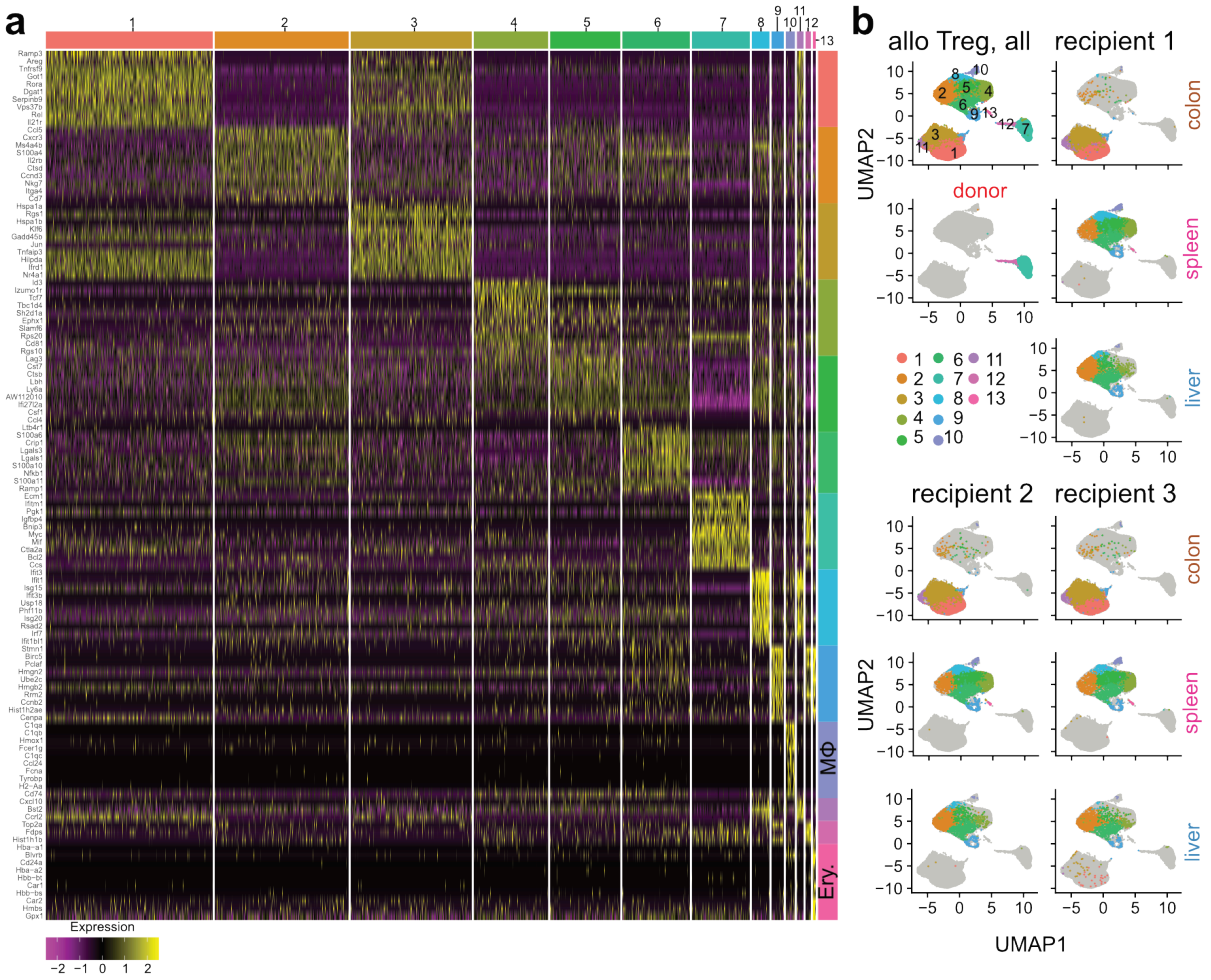

Figure 3.25: **Gene expression patterns of global single-cell clusters with equal recipient contributions.** **a** Log_2-normalized, scaled gene expression data for up to 10 top cluster-defining genes are represented as a heatmap for clusters 1-13, with the column and row sidebars indicating the single-cell cluster affiliation. MΦ: macrophage signature, Ery.: erythroid signature. **b** UMAP embeddings of global gene expression data, colored by clusters and stratified by recipients 1-3 or the pre-transplant allo donor Treg population.

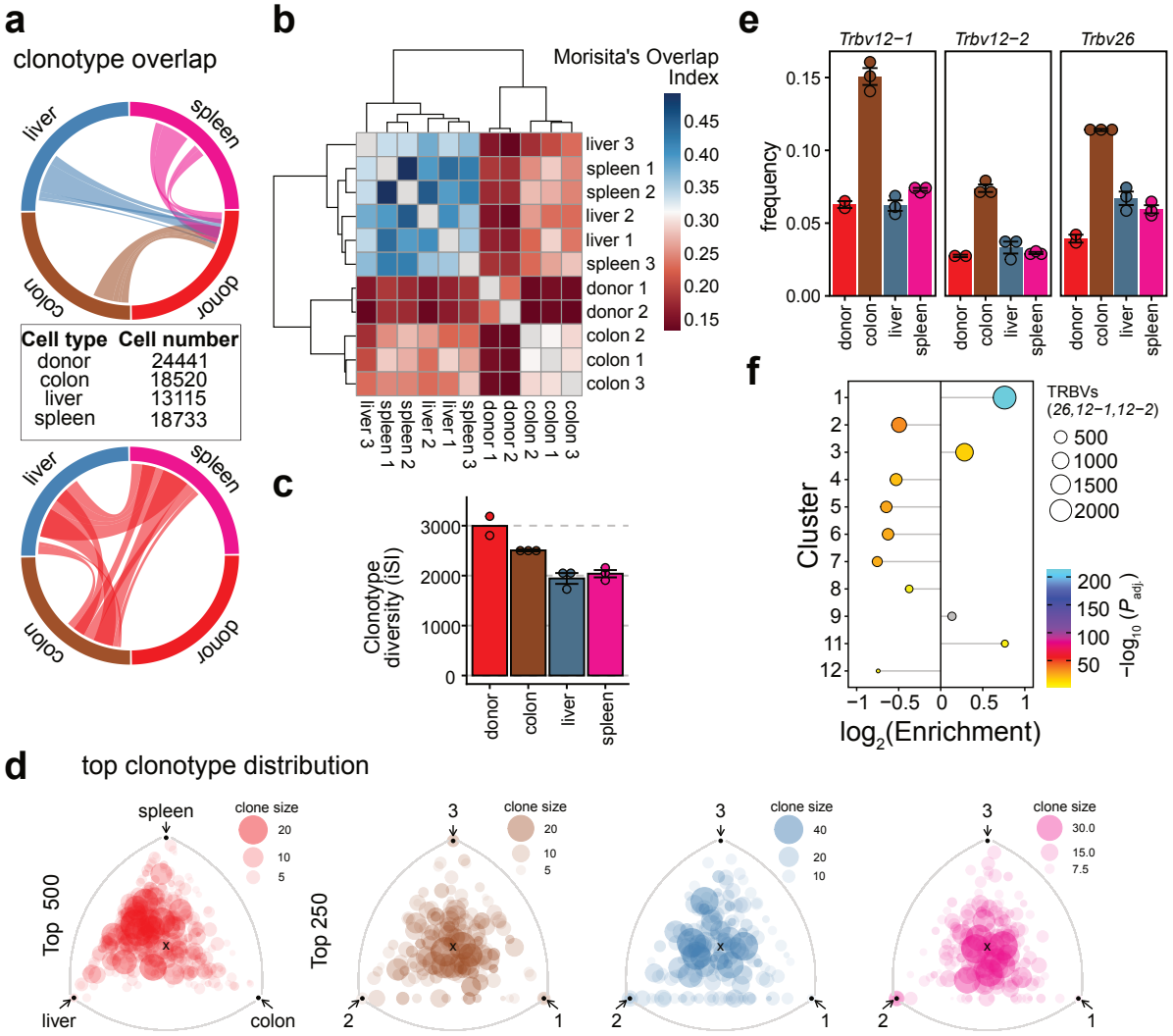

Figure 3.26: **Frequent clonotypes are shared between organs but colon-homing Treg are enriched for certain variable segments. a** Clonotype overlap between the donor and re-isolated Treg populations (top panel) or between colon, liver and spleen (bottom panel) are represented as CIRCOS plots. The width of the bands corresponds to the fraction of the TCR repertoire and the absolute numbers of cells with informative TCRs are provided (middle panel). **b** All samples' Morisita's Overlap Indices are represented as heatmap according to the color scale. Row and column dendrograms indicate the corresponding unsupervised clustering. **c** Clonotype barplots representing the diversity mean of the transplanted (n=2) or mean ±SE for n=3 re-isolated Treg populations, determined by the inverse Simpson Index (iSI). **d** Barycentric triangle plots representing the most frequent (top500/ top250) clonotypes and their distribution between organs (red) or the organ-wise (brown, blue, purple) distribution across recipients (n=3). The bubble size is indicative of the normalized size of the clonotype. **e** Barplots of selected TCR beta variable segment gene usage frequencies. Bars represent the mean (donor) or the mean ±SE for n=3 independent samples of re-isolated Treg. **f** Lollipop plot showing the $\log_2$-scaled enrichment of *Tcrbv12-1, Tcrbv12-2* and *Tcrbv26* across single-cell clusters 1-9, 11 and 12. The lollipop diameter corresponds to the clonotype size and the color encodes the -$\log_{10}$-transformed Benjamini-Hochberg adjusted *P* values.

Nevertheless, the unsupervised clustering of samples according to the Morisita's Overlap Index, distinguished the donor, colon and LT-homing Treg populations (**Figure 3.26 b**). The clonotype diversity was highest in the pre-transplant population and declined *in vivo*, as expected from the observations on the bulk level (**Figure 3.26 c**). Also, even more clearly than in the corresponding bulk data, frequent clonotypes were shared between all three organs as well as between all three recipients, when organs were analyzed separately (**Figure 3.26 d**). The prominent usage of the variable beta chain segments *Tcrbv12-1*, *Tcrbv12-2* and *Tcrbv26* by colon-homing Treg was observed in the single-cell data, in analogy to the previous findings of the BMT control and prophylaxis models (**Figure 3.26 e**). The single-cell, cluster-wise *Tcrbv* usage analysis identified the colon clusters (1,3 and 11) as significantly enriched for the three beta chain segments, as demonstrated in **Figure 3.26 f**, and thus paralleling the previously described observations of the TCRrepseq bulk analyses of the BMT control and prophylaxis models.

3.5.3 TCR-independently shaped Treg transcriptomes

Representative examples of clonotypes that were rather ubiquitous than organ-specific (top panels) or more frequent in either colon or liver (bottom panels) are shown in **Figure 3.27 a**. An example of a common clonotype (*Trav16D-DV11/Tcrbv3*) that could be traced across organs and also individuals is highlighted in the previously used global UMAP embedding (**Figure 3.27 b**). The organ-specific phenotype was acquired by Treg harboring that particular TCR, as exemplified by the gene expression of colon-typical genes (**Figure 3.27 c**).

3.5.4 Subclustering of colon homing Treg

In order to further elaborate on the observed heterogeneity in gene expression of Treg that had migrated to the colon, this subset was studied separately for the following analyses. In line with the preceding global analysis, a UMAP embedding was generated to visualize the distribution of colon Treg cells according to their transcriptomic likeness. Based on that, 8 colon subclusters were identified, as illustrated in **Figure 3.28 a**.

When gene expression differences between the colon subclusters were interrogated, various genes could be identified as cluster-typical. Most prominently, *Hspa1a* and *Jun* characterized cluster 3, for instance, while *Ramp3* showed the strongest expression in cluster 5. The colon subcluster 6 was rich in *Ifit1* and *Isg15* expressing Treg (**Figure 3.28 b**).

In order to shed light on functional differences between colon subclusters, various

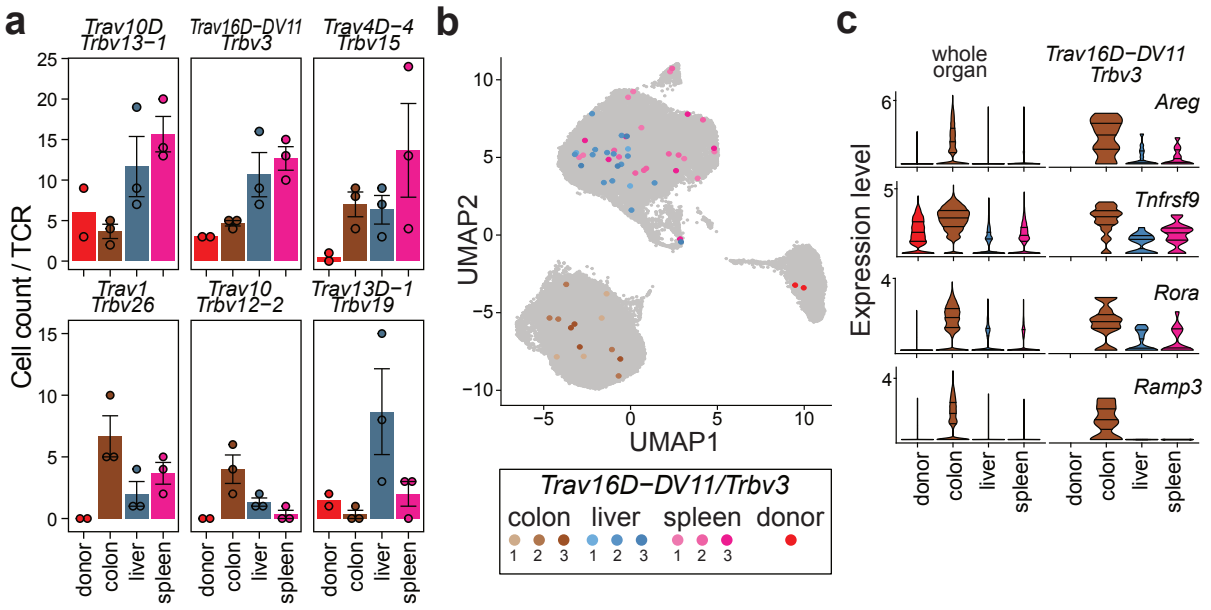

Figure 3.27: **Common clonotypes seed into all three organs and adopt site-specific gene expression. a** Cell counts for individual clonotypes that are ubiquitous (top) or unevenly distributed across organs (bottom) are shown as bar plots. Bars represent the mean of n=2 donor or mean±SE of n=3 recipients and individual data points are shown as dots. **b** Treg carrying one representative clonotype are highlighted within the previously used global UMAP embedding (top panel), distinguishing individual recipients as described in the legend (bottom panel). **c** Violin plots visualizing the organ-wise gene expression levels of four colon Treg signature genes, contrasting the entirety of Treg (left panels) with the subset expressing the indicated TCR (right panels).

marker genes or selected gene signatures were highlighted across the entire embedding (**Figure 3.28 c**). Cluster 6 contained Treg with a strong interferon response, paralleled by high levels of interferon regulatory factor 7 (Irf7), as compared to the other clusters. The relatively high abundance of *Myc* targets in cluster 7 Treg hinted at a small subpopulation that is more proliferative. Some previously identified colon-typical genes were rather uniformly expressed across colon subclusters (e.g. *Batf*, *Ctla4*, *Gzmb*). Furthermore, the majority of colon-homing allo Treg also expressed *Pdcd1lg2* (PD-L2) and *Fasl* as a means of cell contact-dependent regulatory function. Treg in a central zone with contributions from clusters 1-4 also expressed high levels of the killer cell lectin-like receptor subfamily members *Klrc1* (NKG2A) and *Klrd1* (CD94). In the right proportion of the embedding, Treg were found to be glycolytically more active than others (e.g. Pgk1, Dgat1) and typical tissue Treg markers were also found more abundantly expressed in Treg occupying the right pole of the embedding, with *Klrg1* demarcating cluster 2 and *Il1rl1* sharply highlighting cluster 5.

The important Treg effectors *Tnfrsf18* (GITR) and *Tnfrsf9* (4-1BB) highlighted clusters 1, 2, 5 and 7 along with elevated levels of the core Treg marker *Il2ra*. The previously in-

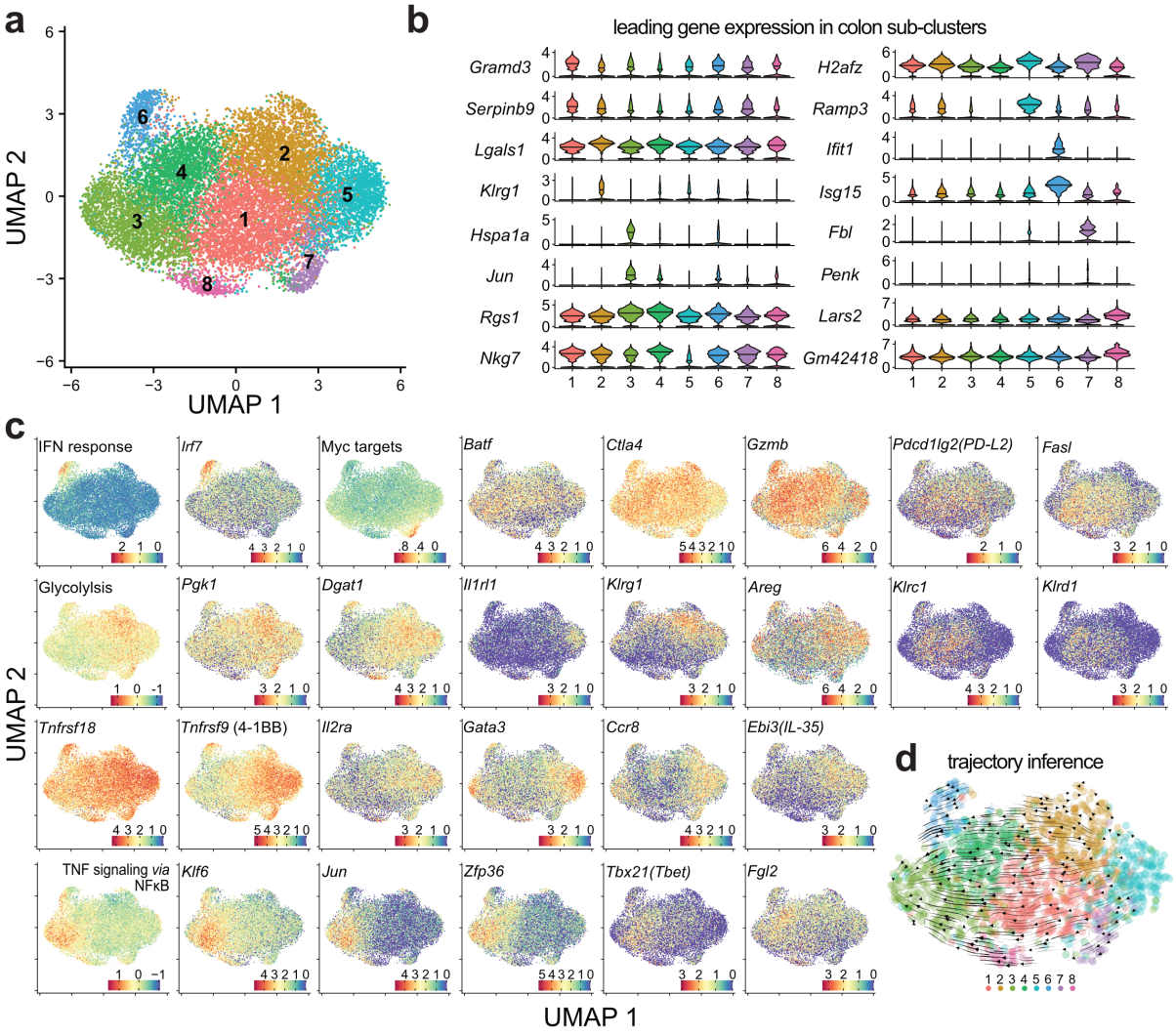

Figure 3.28: **Investigation of colon-homing allo Treg heterogeneity allows the inference of potential cell fate trajectories.** **a** UMAP embedding of 14,781 colon-homing allo Treg with the 8 identified Louvain-clusters highlighted. **b** Violin plots representing the $\log_2$-scaled gene expression levels of the leading colon cluster-defining genes across clusters 1-8. **c** Various representations of the UMAP embedding in **a** with color-coded $\log2_2$-transformed gene expression levels for the gene/pathway indicated in each plot panel. The expression level corresponds to the individually depicted color scale. **d** Visual representation of pseudotime-inferred cell fate trajectories across colon subclusters 1-8. The lines indicate potential differentiation paths and arrow heads indicate the corresponding directionality.

troduced effector Treg tissue adaption marker *Gata3* was also more strongly expressed in the right area of the UMAP embedding (e.g. cluster 5), but was also found equally high expressed in a small proportion of cluster 3. The gene expression distribution of the functionally important chemokine receptor *Ccr8* fairly coincided with *Gata3* expression. Along with the above mentioned genes, also the secreted effector *Ebi3* (IL-35) was found more highly expressed to the right end of the projection. In contrast, the antiinflammatory cytokine Fgl2 was more characteristic of Treg occupying clusters 3 and 4 on the left hand side of the UMAP, in which *Tbx21* (Tbet) was also more highly expressed. Lastly, genes involved in the TNF-induced signalling *via* NFκB, eg. *Klf6*, *Jun* and *Zfp36*, were also highly specific for clusters 3 and 4, demarcating Treg belonging to the left proportion of the UMAP as immunologically highly active.

After the functional dissection of the 8 colon subclusters, an established algorithm for the inference of proposed differentiation trajectories, namely RNA velocity, was applied to the pre-defined UMAP embedding in an effort to shed light on the fate of colon-homing allo Treg alongside a pseudotime-line (La Manno et al., 2018). Apart from very few exceptions, the the flow of the resulting streamlines originated on the right hand side (clusters 2 and 5), following an almond shape to the very left proportion of cluster 3.

The characteristic gene expression of colon-homing allo Treg in clusters 1-8 is provided in more detail as a heatmap in **Figure 3.29**.

3.5.5 Distribution of colon-enriched clonotypes and functional implications

Previously, the *Trbv* segments *12-1*, *12-2* and *26* were observed to be enriched in the colon-homing Treg of BMT control and prophylaxis animals, the latter being paralleled by the single-cell data. Consequently, the distribution of clonotypes containing any of the three variable segments was mapped to the colon-restricted UMAP embedding (**Figure 3.30 a**).

The distribution analysis of the three *Trbv* segments revealed a significant enrichment in clusters 2 and 7, whereas in clusters 1 and 3-5 were depleted in Treg with clonotypes incorporating any of the *Trv12-1*, *12-2* and *26* segments (**Figure 3.30 b**). When such Treg were tested for differentially expressed genes in comparison to Treg bearing a different TCR, a few genes came out as significantly regulated (**Figure 3.30 c**). The transcription co-factor homeodomain-only protein (*Hopx*) was found upregulated, a gene implicated in DC-mediated T cell unresponsiveness, so far known in iTregs. Furthermore, the two members of the TNF-alpha signaling via NFκB, *Rel* and *Bcl2a1d*,

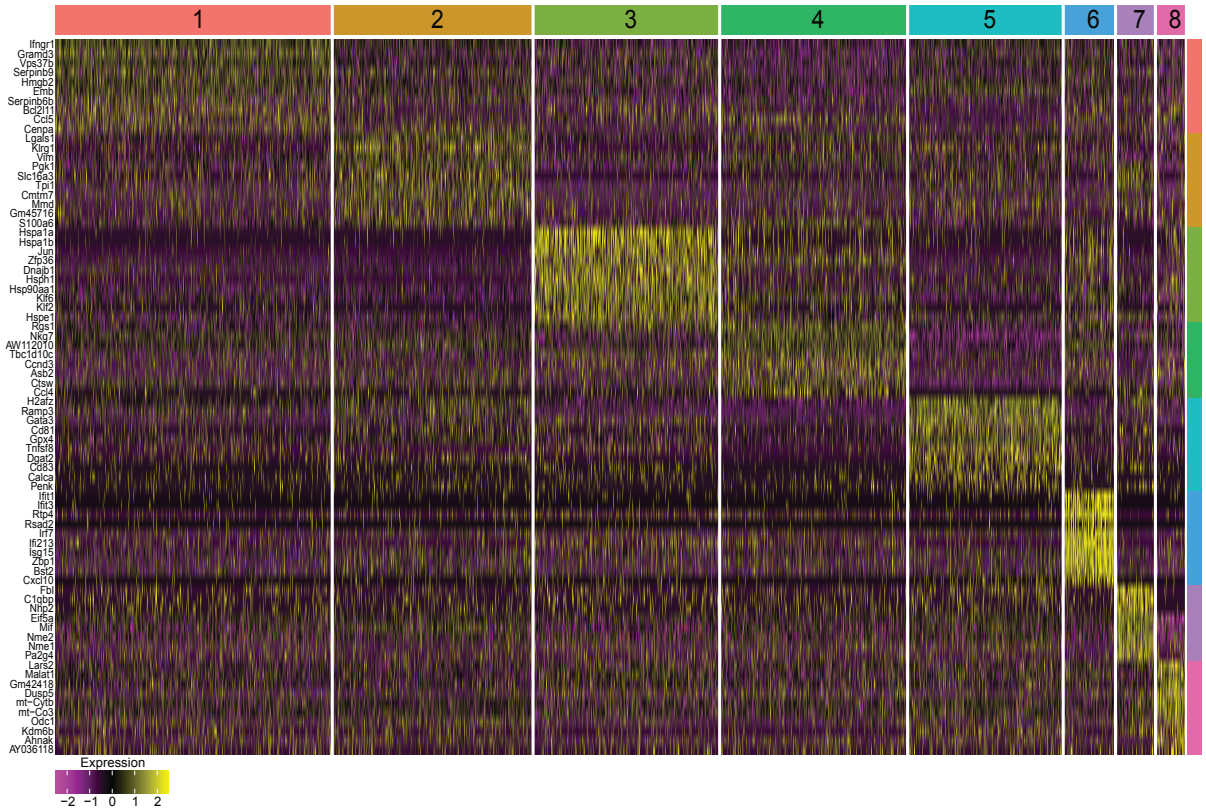

Figure 3.29: **Top ranking gene expression heatmap of colon clusters 1-8.** Log_2-normalized, scaled gene expression data for up to 10 top cluster-defining genes are represented as a heatmap for colon subclusters 1-8, with the column and row sidebars indicating the cluster affiliation.

were among the more highly expressed genes. The effector Treg marker *Rora* as well as the tissue Treg marker *Klrg1* were found upregulated in the Treg included in the set of interest. On the contrary, the core Treg gene *Rgs1* was elevated in Treg that did not belong to the singled-out subpopulation. Also, *Lag3* was more abundantly expressed, along with the effector memory T cell marker *Tnfsf8*. Higher levels of *Pdcd1* were observed in the non-set population, potentially rendering them more prone to PD-L1-mediated anergy.

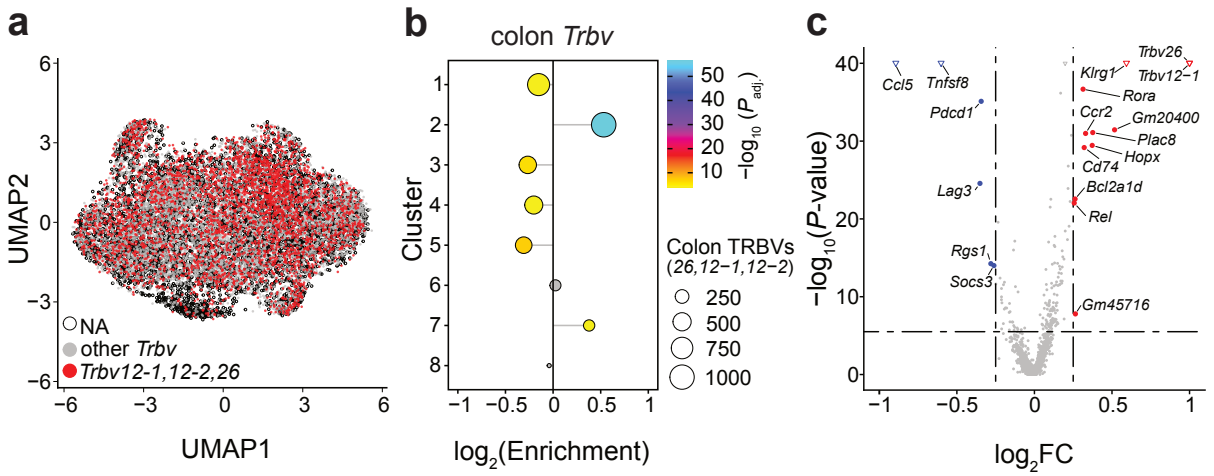

Figure 3.30: **Cluster 2 and 7 Treg are enriched for colon-enriched *Tcrb* variable segments and are functionally distinct.** **a** Previous UMAP embedding of colon-homing allo Treg, highlighting Treg harboring *Trbv12-1*, *12-2* or *26* ("set") variable segments (red dots), Treg other than that (grey dots) or Treg with no informative TCR data (NA, empty circles). **b** Lollipop plot representing the $\log_2$-transformed enrichment of indicated *Trbv* segments across colon subclusters 1-8. The color of the lollipops represents the the -$\log_{10}$-scaled Benjamini-Hochberg adjusted *P* value and the size corresponds to the clonotype size, as indicated in the legend. **c** Volcano plot depicting DEGs in the comparison of colon Treg belonging to the *Trbv* "set" *versus* Treg with TCRs other than that. Up- and down-regulated genes are colored in red or blue, respectively.

CHAPTER 4

Discussion

4.1 Gene expression and TCR repertoire dynamics in aGvHD

Undoubtedly, Treg hold a pivotal role in the regulation of the pathological processes underlying aGvHD. Their suppressive capacity is predicated on basal functions for tissue homeostasis as well as specific gene expression programs evoked by inflamed microenvironments. However, the prerequisites for their clinical efficacy in controlling aGvHD remain elusive. The processes that govern the distribution of Treg across target organs as well as the extent of tissue-specific gene expression imprints under the fulminant inflammation caused by major MHC-mismatch are insufficiently explored. It is not clear which clonotypes might be directive for the function of Treg in colon, liver and spleen in this context and whether the TCR repertoire evolution is regulated by tissue-specific antigens or alloantigens.

Therefore, the gene expression profiles and TCR repertoires of Treg in various settings of a MHC-mismatched murine aGvHD model (C57BL/6→BALB/c) were characterized in this thesis. A detailed analysis of alloantigen-specifically and polyclonally *in vitro* expanded Treg was carried out along with the thorough investigation of Treg that had migrated to the colon, liver and spleen of recipient mice to shed more light on the modalities of tissue adaption during transplantation in settings with increasing severity. A broad spectrum of samples derived from a homeostatic, a transplant control, a prophy-

laxis and a therapy model enabled the investigation of physiologically tissue-resident Treg as well as ones that had infiltrated the tissues of interest within 7 d after transfer. The unbiased approaches of RNA-sequencing and TCR repertoire-sequencing were employed to determine the impact of each setting and tissue on Treg gene expression and the concomitant TCR repertoire changes. Moreover, prophylactically administered allo Treg were explored by means of single-cell gene expression and TCR repertoire profiling, providing an even more detailed insight into functionally distinct subpopulations and their clonotypes after organ infiltration.

Both expansion methods yielded functional Treg with minor gene expression differences. As expected, allo Treg were characterized by a strongly narrowed TCR repertoire. Consequently, the donor Treg TCR repertoire fraction overlapping with that of the respective organ-infiltrating populations was markedly larger for allo Treg. Strikingly, the most frequent TCR beta chains were shared between target organs and the organs of individual recipients with a more even distribution of allo Treg as compared to poly Treg. The superior enrichment of organ-infiltrating allo Treg clonotypes among top-ranking donor Treg *Tcrb*s argued for an alloreactivity-driven migration pattern that may partially be ascribed to *Trbv16*, which was more frequently detected in allo Treg. Depending on their target organ, Treg differed in their *Trbv* usage. In the prophylaxis and control BMTs, the three *Tcrb* variable segment genes *Trbv12-1*, *Trbv12-2* and *Trbv26* were significantly enriched among Treg homing to the colon. One of the variable genes (*Trbv12-2*), was also characteristic of homeostatic, colon-resident Treg. The associated clonotypes may thus be involved in the recognition of flora-restricted antigens exposed early upon irradiation-induced damage (*Trbv12-1*, *Trbv26*) or ubiquitous, dietary or microbial antigens specific for the colon (*Trbv12-2*). In contrast, *Trbv30* almost exclusively occurred in *Tcrb*s of therapeutically given Treg, and might be indicative of translocated bacteria or fibrotic tissue as consequence of the initially undamped allo response.

Most notably, Treg that infiltrated the colon and the liver or spleen of recipients acquired the NLT and LT-typical gene expression signatures characteristic of the corresponding physiologically tissue-resident populations, respectively. Nevertheless, additional suppressive, metabolic or proliferation-associated gene expression modules were evoked by the perturbations underlying the BMT control, prophylaxis and therapy models, and differed between organs and models. Remarkably, the functionally distinct co-expression networks that distinguished various Treg populations across organs and experimental systems could be leveraged to identify single-cell subpopulations in the allo Treg prophylaxis approach. The latter also empowered the demonstration of how allo Treg gene expression was shaped by the target organ, independent of the clonotype. The granular dissection of colon-homing allo Treg revealed inferred differ-

entiation trajectories from homeostatic and metabolically more active states towards prominently induced immune responses. Furthermore, the colon subpopulation harboring TCRs consisting of globally colon-enriched *Trbv* segments was identified as functionally distinct.

4.2 Lessons from the transcriptional plasticity of Treg

4.2.1 Treg identity of allo and poly Treg

Both *in vitro* expansion approaches allowed the generation of cell products that featured the hallmarks of murine Treg, and were hence considered functional. The gene expression signature provided by Aubert et al. (2020) served as a meta-analysis-based reference for the assessment of the murine Treg identity. In the present analyses, the core Treg gene signature was extended by the lymph node homing marker *Sell* (CD62L), given its central role in the protective capacity of Treg in GvHD (Ermann et al., 2005; Taylor et al., 2004). Similarly, the heterodimerizing integrins encoded by the core Treg gene *Itgae* (CD103) and *Itgb7* have been shown to render Treg efficacious in the amelioration of ongoing GvHD (Zhao et al., 2008). Therefore, *Itgb7* was also appended to the core signature. Despite their preserved Treg identity, gene expression differences between allo and poly Treg could be identified, and interestingly, were in part preserved even after cells had seeded into their recipients' target organs. This maintained imprint hinted at the epigenetic re-wiring of Treg evoked by the partial, microenvironmental mimicry of tissue damage-associated antigens in the allo *in vitro* expansion. The expression of preproenkephalin (*Penk*) was consistently upregulated in allo Treg and has been suggested to be regulated by TNFR signaling by virtue of modulating the BATF-AP-1 complex (Aubert et al., 2020). In line with that, also the TNF receptor gene *Tnfrsf9* (4-1BB) was found upregulated in *ex vivo* allo Treg populations. The more obviously *in vivo* retained allo-poly differences in the BMT control and prophylaxis models in contrast to therapeutic Treg might be explained by the abundance of antigens exposed early after irradiation damage that are shared with the antigen landscape of irradiated DCs in the allo *in vitro* culture.

4.2.2 Tissue adaption and co-expression modules in transplanted Treg

One striking observation was the expression of NLT-typical genes evoked by both the healthy as well as the inflamed colon. Treg residing in or homing to that barrier tissue upregulated almost all of the characteristic genes identified by Miragaia et al. (2019).

Accordingly, homeostatic as well as transplanted Treg typically expressed LT signature genes. Of note, Treg of the BMT models were analyzed with respect to genes that were upregulated in colon Treg of the homeostatic state, and the gene signatures induced by this target organ or the two lymphoid tissues were largely paralleled by transplanted Treg, illustrating the consistently dominant tissue imprint across the employed models. The target tissue was hence clearly identified as the primary modulator of Treg gene expression profiles. On the single-cell level, the rapid target tissue adaption of prophylactically administered allo Treg was equally striking, as they formed a clearly separate cluster by means of their transcriptomic uniqueness. It should be highlighted that on the one hand, this demonstrated the directive influence of the tissue on Treg gene expression, and on the other hand, that Treg retained a striking gene expression plasticity even after the excessive stimulation *in vitro*.

The integration of all bulk-level gene expression data from the homeostatic, the BMT control and the prophylaxis models empowered the identification of co-expression networks employed by the analyzed Treg. Overall, the colonic Treg-identifying clusters could be broken down to T cell activation and differentiation processes, and LT-homing Treg featured replication-associated gene expression patterns. Some of the gene expression modules coincided with single-cell clusters defined by the target tissue-identity, while others overlapped with subpopulations that were not restricted to the donor or a single organ-homing population of Treg. Hence, there likely are distinct functional Treg species that play an important role in the allo Treg prophylaxis setting independent of the target tissue. Among the genes in in the co-expression clusters, TFs were particularly frequent. For instance, *Tbx21* (Tbet) was selectively upregulated in *in vivo* Treg of BMT animals, likely rendering them capable of migration to type 1 interferon-mediated inflammation as well as maintaining homeostasis (Koch et al., 2009). To decipher how the observed gene expression networks are governed by the implicated TFs, functional pooled Cas9 ribonucleotide-protein-complexes (RNPs) screens targeting the TFs in each of the graph-based clusters might be a complex but yet worthy approach to validate their function. Schumann et al. (2020) have identified how genes critical for adult human Treg are orchestrated by simultaneously targeting up to 40 TFs. The highlighting of different gene groups describing known Treg subtypes also stressed the dominant effector phenotype of Treg located in the colon featuring *Nr4a1* expression as a hallmark of TCR activation (Shevyrev & Tereshchenko, 2020).

4.2.3 Dynamic functions of Treg in different MHC-mismatch models

The challenge of co-transplanted Tcon induced a multilayered suppression-associated gene expression pattern in Treg that had infiltrated the colon, but also liver and spleen-directed Treg featured some key effectors. Increased *Tigit* expression in disease-burdened Treg from liver and spleen likely resulted in increased DC IL-10 production, promoting an anti-inflammatory milieu (Yu et al., 2009). The observed upregulation of Cd274 (PD-L1) might be essential in the maintenance of the Treg pool in the colon under perturbed conditions, considering the promoted proliferation of Treg by interaction with APC CD80 (Yi et al., 2011). Treg armed with FasL have proven efficacious in directly targeting activated Tcon, hence the induction of *FasL* expression in the prophylaxis model might be a major Treg effector mechanism for controlling aGvHD (Yolcu et al., 2013). The cytolytic effector *Gzma* was found upregulated In prophylactically and therapeutically administered Treg homing to the colon, in line with its pivotal role in the prevention of GvHD in the GIT (Velaga et al., 2015). The relevance of the ectoenzymes CD39 and CD73 in counteracting the pro-inflammatory environment rich in extracellular ATP is known from autoimmune diseases (Antonioli et al., 2013). At least, *Entpd1* (CD39) was found upregulated in prophylactic Treg after organ infiltration, highlighting ATP-deprivation as an essential mechanism early after BMT.

Strikingly, *Reg3b* was markedly higher expressed in therapy Treg of the colon, and contributes to antimicrobial defense against gram-negative bacteria by reducing the amount of mucosa-associated microbiota and the inhibition of translocation from the lumen to epithelial layers (Hrdý et al., 2020; Moyat et al., 2017; Ojo et al., 2019). Of note, equally high levels of *Reg3b* were observed in Treg transferred after conditioning only, suggesting that the translocation of bacteria from the lumen to intra-epithelial layers is a major damaging component in pathogenesis. In addition to *Sell* (CD62L), also the chemokine receptor gene0 *Ccr7* was more abundant in colonic Treg of the therapy setting as compared to prophylaxis, and has been been identified as a subpopulation-defining marker for homeostatic Treg that gain access to T cell zones rich in IL-2 (Smigiel et al., 2013). The present data indicate that this mechanism was re-established by therapeutic Treg migrating to the colon. The latter also expressed markedly higher levels of *Itgae* (CD103) encoding a critical integrin for mucosal homing that has been shown to potentiate Treg for the amelioration of ongoing GvHD (Zhao et al., 2008). The chemokine receptor *Cxcr5* proved efficacious in the amelioration of chronic GvHD by enabling Treg to home to germinal centers of secondary lymphoid organs (McDonald-Hyman et al., 2016). Similarly, therapeutically transferred Treg infiltration the colon and the liver featured markedly elevated levels of *Cxcr5* as compared to prophylaxis, emphasizing the importance of secondary lymphoid organ homing ca-

pacity in the context of aGvHD.

4.3 Clonal evolution of Treg *in vivo*

A feature in which allo and poly Treg differed was the markedly reduced TCR repertoire diversity of the latter, as expected from the stimulation of Treg with irradiated DCs instead of globally activating costimulatory CD3 and CD28. Due to their easy accessibility, dendritic cells were isolated from recipient-type spleens and it should be noted that their origin might already have influenced the Treg response after transplantation (Eisenbarth, 2019). Nevertheless, at least in the context of a complete MHC-mismatch, Treg with a less diverse TCR repertoire were not inferior to poly Treg in terms of core Treg marker levels, in contrast to a murine minor mismatch study that identified a broad repertoire as a requirement for proper Treg function (Föhse et al., 2011). The striking upregulation of the *Tcrb* variable segment *Trbv16* in allo Treg could be confirmed by means of comparing *Tcrb* repertoire data of allo and poly Treg across BMT models, and implicated associated TCRs in the recognition of allo-antigens.

In line with the less divers donor allo Treg populations, the fraction of overlapping *Tcrb*s with the populations recovered from recipients was markedly higher. The most frequent clonotypes were shared between the colon, the liver and the spleen of individual recipients and also between the three organs in all major mismatch models at which allo Treg featured a more even distribution as compared to poly Treg. The absence of tissue-restricted, dominant TCRs in the complete MHC-mismatch was clearly distinct from the the relevance of organ-specific clonality observed in the minor mismatch model presented by Föhse et al. (2011).

Of note, the focused clonotype repertoire of allo donor Treg favored organ infiltration, but did not interfere with the induction of specific genes imprinted by the target tissue. Clonotypes incorporating the allo-dominant *Trbv16* segment are likely involved in the migration properties of allo Treg. While a high affinity towards alloantigens can be presumed as causative for the efficient homing into colon, liver and spleen, an immunological superiority over other clonotypes is not implied.

The variable segment *Trbv30* was almost exclusively used in *Tcrb*s found in Treg that had infiltrated the organs of therapy animals which was more obvious for poly Treg. Hence, *Trbv30* might be essential for the recognition of antigens related to scar tissue or translocated bacteria that were absent in allo *in vitro* culture. As previously pointed out, therapeutic Treg located in the liver also featured prominent *Reg3b* expression as a surrogate for antimicrobial defense, supporting this assumption.

In spite of the widely organ-independent infiltration of ubiquitous Treg clones, the three

variable segments *Trbv12-1*, *Trbv12-2* and *Trbv26* were over-represented in colonic Treg from the BMT control and prophylaxis models. In addition, *Trbv12-2* was found frequently incorporated in physiologically colon-resident Treg. Thus, the corresponding TCRs are likely to be involved in the recognition of bacterial or dietary antigens exposed early after irradiation (*Trbv12-1*, *Trbv26*), or generally flora-associated antigens (*Trbv12-2*), that were all no longer present in the colon of therapy animals as a consequence of excessive tissue damage. Similarly, the single-cell study conducted by Muschaweck et al. (2021) revealed the directive role of the microbial composition in the gut for the Treg clonotype repertoire in a murine colitis model.

4.4 Heterogeneity of prophylactic allo Treg homing to the colon

In the colon of prophylaxis animals, a pseudotime differentiation axis from a homeostatic state of basal metabolic and immunological activity towards a highly TNF-activated state was inferred.

Strikingly, a small *Irf7*-high population responding to type 1 interferons could be identified by means of single-cell resolution that was undetected on the bulk-level due to its low abundance relative to the entirety of the colon-homing population. Approximately half of the colon-located allo Treg featured considerable expression levels of *Ccr8*. The persistence of adoptively transferred Treg in a murine haplo-mismatch model of GvHD has been demonstrated to depend on this chemokine receptor (Coghill et al., 2013), and was thus identified relevant in the major mismatch context.

The natural killer cell receptors (NKRs) *Klrc1* (NKG2A) and *Klrd1* (CD94) are mainly known from $CD8^+$Treg or $CD8^+$T cells in general an little is known about the relevance and gene expression dynamics in $CD4^+$T cells (Ortega et al., 2004). In association with GvHD, the NKRs are mainly known from their relevance for leukemic clearance and little alloreactivity (Tanaka et al., 2004). In multiple sclerosis, CD94/NKG2A receptors have shown to play an important role in regulating T cell activity (Correale & Villa, 2008). Most notably, the two aforementioned NKRs were expressed in roughly a third of the colon-infiltrating allo Treg in the prophylaxis model, providing a novel role for *Klrc1* and *Klrd1* in the regulation of Tcon-mediated alloreactive responses in aGvHD.

Surprisingly, significant gene expression differences could be identified in prophylactic allo Treg within the colon when Treg harboring colon-enriched *Trbv* segment-associated TCRs were contrasted with all other clonotypes. The upregulation of *Klrg1* hinted at potential superiority in exerting tissue homeostasis (Delacher et al., 2017). The co-transcription factor Hopx has been implicated in the maintained function of iTreg and

was required for limiting inflammation in a murine multiple sclerosis model (Jones et al., 2015; Jones & Hawiger, 2017). The higher relative abundance of *Hopx* expression in Treg carrying a *Tcrb* of interest, together with the concomitantly lower *Pdcd1* (PD-1) levels hinted at a specialized Treg subpopulation standing out for its persistent functionality and longevity.

4.5 Strengths and limitations

While the Treg-Tcon ratio of 1 was consistent across prophylaxis and therapy experiments, the second therapy replicate (N=500x10^3) differed from the first one (N=250x10^3) in terms of the absolute numbers of transplanted T cells. Therefore, the former has to be repeated analogously. A consequence of the aggravated therapy model was the difficulty to isolate samples of sufficient quality, and thus not all library preparations were successful (e.g. spleen-derived allo Treg samples). In consequence of the inflammation caused by the conditioning and the dominant alloreactive immune response, the re-isolation of Treg proved challenging. As a result, samples were scarce and of limited quality for some conditions.

As a strength it should be pointed out that low-input RNA-seq was successfully applied for a consistently robust library preparation generated from scarce samples. The requirement of 6 independent samples per condition is considered good practice, and was achieved for all Treg groups throughout this thesis (Schurch et al., 2016). In addition, the high sensitivity of the in-house TCRrepseq library protocol allowed an appropriate representation of the TCR repertoire for most of the samples. The integration of UMIs and the unbiased cDNA priming improved the validity of the data.

4.6 Conclusion

Both allo and poly expansion yielded cell products with a uniform Treg identity with allo Treg featuring an allo-specifically narrowed clonal repertoire that promoted alloreactivity-driven organ infiltration. Frequent clonotypes seeded into all analyzed target organs, with the tissue emerging as the primary determinant for gene expression changes. Little evidence was in support of dominant organ-restricted clonal expansion, however some TCR families were enriched in colonic Treg. Characteristic NLT and LT signatures known from homeostasis were rapidly established in co-transplanted Treg which also induced additional suppressive genes after tissue homing. The organ-dependent imprint on gene expression was independent of the TCR, since Treg carrying the same TCR were be observed in all tissues. A colon-homing subpopulation expressing two

NKRs was newly identified.

Many of the observed genes evoked by the presence of Tcon are implicated in known functions of Treg and could be linked to aGvHD pathogenesis. It stands to reason that Treg naturally adopt gene expression modules that ameliorate GvHD. Furthermore, identified candidate genes pose targets for engineering Treg with enhanced clinical efficacy. Oligoclonal Treg cell products defined by variable segments found enriched in colonic Treg may prove advantageous for protection against GvHD. However, it is noteworthy that the unrestricted modification of suppressive capacity might negatively affect the GvL effect.

4.7 Outlook

This thesis focused on the detailed characterization of Treg regarding their gene expression and TCR repertoire changes in complete MHC-mismatch models of aGvHD. The analogous analysis of the co-transplanted, aGvHD-inducing Tcon of the prophylaxis and therapy settings would pose a valuable addition to the project's data sets. The corresponding samples have already been subjected to the isolation of RNA and are available for library preparation. Moreover, an exact replicate of the first therapy experiment has been conducted by WG E/H and samples derived thereof are pending for TCR repertoire and RNA-seq library preparation. The corresponding data will provide evidence for the validation of the therapy-related findings in this work. In analogy to the single-cell allo Treg prophylaxis experiment, the therapy setting was illuminated with cellular resolution, and the corresponding analyses will be accomplished in the near future. These results will be most crucial to validate and extend the observed differences between the two Tcon-based models. Establishing a FACS-staining for the newly identified NKRs NKG2A and CD94 on Treg will be crucial for both the validation on the protein-level as well as sorting. In order to address the pivotal role of the microbiome in the pathobiology of aGvHD, great efforts have been undertaken by WG E/H to generate and maintain germ-free mice. These animals are now available for the BMT model systems deployed in the present work. The downstream analysis of the transcriptional plasticity and clonotype evolution of adoptively transferred Treg in germ-free mice will be an invaluable addition to the observations gathered so far. Moreover, conducting serial transplant experiments is another long-term goal within the present project and will serve the aim to validate the enhanced efficacy of *in vivo* selected Treg clonotypes in aGvHD prevention and amelioration.

In addition to colon, liver and spleen, their vicinal lymph nodes, the bone marrow and the small intestine could provide insights into the involvement of additional Treg populations. The discrimination between colonic Treg located in the *lamina propria* and the

lamina intraepithelialis would allow a more detailed understanding of Treg interactions at the disrupted barrier. The challenging isolation of Treg infiltrating the skin of BM recipient mice is a long-term goal of WG E/H and will ultimately enable the incorporation of another major aGvHD target organ into the scope of this project.

Despite the fact that the investigation of Treg clonotype repertoires did provide little evidence for a tissue-restricted expansion, a few interesting findings could be derived from the bulk and single-cell TCR repertoire data. The *Tcrb* variable segments *Trbv12-1*, *Trbv12-2* and *Trbv26* could be linked to the potential recognition of damage-associated antigens in the colon, and thus pose potential targets for Treg with an engineered specificity that could prove superior in migrating to and persisting at sites of severe tissue damage. The procedure presented by Roth et al. (2018) represents the first approach for replacing the endogenous TCR without disrupting its transcriptional regulation. Later, more advanced methods enabling the orthotopic replacement of the TCR were developed and could prove valuable in the application of candidate TCRs in the presented aGvHD models (Moosmann et al., 2022; T. R. Müller et al., 2021; Schober et al., 2020).

Facing the extent of future analyses, the establishment of spatial transcriptomics in the project's context will prove worth pursuing. This tool will enable the histophathological scoring of tissue cryo sections and at the same time provide gene expression and positional identity data for all cell types involved at the single-cell resolution (Ståhl et al., 2016).

References

Literature

Agace, W. W., Higgins, J. M., Sadasivan, B., Brenner, M. B., & Parker, C. M. (n.d.). *T-lymphocyte–epithelial-cell interactions: integrin αE(CD103)β7, LEEP-CAM and chemokines*. 6.

Al-Kadhimi, Z., Gul, Z., Chen, W., Smith, D., Abidi, M., Deol, A., Ayash, L., Lum, L., Waller, E. K., Ratanatharathorn, V., & Uberti, J. (2014). High incidence of severe acute graft-versus-host disease with tacrolimus and mycophenolate mofetil in a large cohort of related and unrelated allogeneic transplantation patients. *Biology of Blood and Marrow Transplantation: Journal of the American Society for Blood and Marrow Transplantation*, *20*(7), 979–985. https://doi.org/10.1016/j.bbmt.2014.03.016

Antin, J. H., & Ferrara, J. L. (1992). Cytokine dysregulation and acute graft-versus-host disease. *Blood*, *80*(12), 2964–2968.

Antonioli, L., Pacher, P., Vizi, E. S., & Haskó, G. (2013). CD39 and CD73 in immunity and inflammation. *Trends in Molecular Medicine*, *19*(6), 355–367. https://doi.org/10.1016/j.molmed.2013.03.005

Appelbaum, F. R. (2001). Haematopoietic cell transplantation as immunotherapy. *Nature*, *411*(6835), 385–389. https://doi.org/10.1038/35077251

Aubert, N., Salomon, B. L., & Marodon, G. (2020). *Characterization of a regulatory T cells molecular meta-signature identifies the pro-enkephalin gene as a novel marker in mice*. 638072. https://doi.org/10.1101/638072

Barnes, D. W., Corp, M. J., Loutit, J. F., & Neal, F. E. (1956). Treatment of murine leukaemia with X rays and homologous bone marrow; preliminary communication. *British Medical Journal*, *2*(4993), 626–627. https://doi.org/10.1136/bmj.2.4993.626

Barnes, M. J., & Powrie, F. (2009). Regulatory T cells reinforce intestinal homeostasis. *Immunity*, *31*(3), 401–411. https://doi.org/10.1016/j.immuni.2009.08.011

Beilhack, A., Schulz, S., Baker, J., Beilhack, G. F., Wieland, C. B., Herman, E. I., Baker, E. M., Cao, Y.-A., Contag, C. H., & Negrin, R. S. (2005). In vivo analyses of early events in acute graft-versus-host disease reveal sequential infiltration of T-cell subsets. *Blood*, *106*(3), 1113–1122. https://doi.org/10.1182/blood-2005-02-0509

Biagi, E., Di Biaso, I., Leoni, V., Gaipa, G., Rossi, V., Bugarin, C., Renoldi, G., Parma, M., Balduzzi, A., Perseghin, P., & Biondi, A. (2007). Extracorporeal photochemotherapy is accompanied by increasing levels of circulating CD4+CD25+GITR+Foxp3+CD62L+ functional regulatory T-cells in patients with graft-versus-host disease. *Transplantation*, *84*(1), 31–39. https://doi.org/10.1097/01.tp.0000267785.52567.9c

Bolwell, B., Sobecks, R., Pohlman, B., Andresen, S., Rybicki, L., Kuczkowski, E., & Kalaycio, M. (2004). A prospective randomized trial comparing cyclosporine and short course methotrexate with cyclosporine and mycophenolate mofetil for GVHD prophylaxis in myeloablative allogeneic bone marrow transplantation. *Bone Marrow Transplantation*, *34*(7), 621–625. https://doi.org/10.1038/sj.bmt.1704647

Cassady, K., Martin, P. J., & Zeng, D. (2018). Regulation of GVHD and GVL activity via PD-L1 interaction with PD-1 and CD80. *Frontiers in Immunology*, *9*, 3061. https://doi.org/10.3389/fimmu.2018.03061

Coghill, J. M., Fowler, K. A., West, M. L., Fulton, L. M., Deventer, H. van, McKinnon, K. P., Vincent, B. G., Lin, K., Panoskaltsis-Mortari, A., Cook, D. N., Blazar, B. R., & Serody, J. S. (2013). CC chemokine receptor 8 potentiates donor Treg survival and is critical for the prevention of murine graft-versus-host disease. *Blood*, *122*(5), 825–836. https://doi.org/10.1182/blood-2012-06-435735

Coombes, J. L., Robinson, N. J., Maloy, K. J., Uhlig, H. H., & Powrie, F. (2005). Regulatory T cells and intestinal homeostasis. *Immunological Reviews*, *204*, 184–194. https://doi.org/10.1111/j.0105-2896.2005.00250.x

Correale, J., & Villa, A. (2008). Isolation and characterization of CD8+ regulatory T cells in multiple sclerosis. *Journal of Neuroimmunology*, *195*(1-2), 121–134. https://doi.org/10.1016/j.jneuroim.2007.12.004

Daly, A. (2012). Remestemcel-L, the first cellular therapy product for the treatment of

graft-versus-host disease. *Drugs of Today*, *48*(12), 773. https://doi.org/10.1358/dot.2012.48.12.1885866

Delacher, M., Imbusch, C. D., Weichenhan, D., Breiling, A., Hotz-Wagenblatt, A., Träger, U., Hofer, A.-C., Kägebein, D., Wang, Q., Frauhammer, F., Mallm, J.-P., Bauer, K., Herrmann, C., Lang, P. A., Brors, B., Plass, C., & Feuerer, M. (2017). Genome-wide DNA-methylation landscape defines specialization of regulatory T cells in tissues. *Nature Immunology*, *18*(10), 1160–1172. https://doi.org/10.1038/ni.3799

Deng, R., Cassady, K., Li, X., Yao, S., Zhang, M., Racine, J., Lin, J., Chen, L., & Zeng, D. (2015). B7H1/CD80 interaction augments PD-1-dependent T cell apoptosis and ameliorates graft-versus-host disease. *Journal of Immunology (Baltimore, Md.: 1950)*, *194*(2), 560–574. https://doi.org/10.4049/jimmunol.1402157

Di Ianni, M., Falzetti, F., Carotti, A., Terenzi, A., Castellino, F., Bonifacio, E., Del Papa, B., Zei, T., Ostini, R. I., Cecchini, D., Aloisi, T., Perruccio, K., Ruggeri, L., Balucani, C., Pierini, A., Sportoletti, P., Aristei, C., Falini, B., Reisner, Y., … Martelli, M. F. (2011). Tregs prevent GVHD and promote immune reconstitution in HLA-haploidentical transplantation. *Blood*, *117*(14), 3921–3928. https://doi.org/10.1182/blood-2010-10-311894

Duarte, R. F., Labopin, M., Bader, P., Basak, G. W., Bonini, C., Chabannon, C., Corbacioglu, S., Dreger, P., Dufour, C., Gennery, A. R., Kuball, J., Lankester, A. C., Lanza, F., Montoto, S., Nagler, A., Peffault de Latour, R., Snowden, J. A., Styczynski, J., Yakoub-Agha, I., … arrow Transplantation (EBMT). (2019). Indications for haematopoietic stem cell transplantation for haematological diseases, solid tumours and immune disorders: Current practice in Europe, 2019. *Bone Marrow Transplantation*, *54*(10), 1525–1552. https://doi.org/10.1038/s41409-019-0516-2

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. *Nature Medicine*, *9*(9), 1144–1150. https://doi.org/10.1038/nm915

Eisenbarth, S. C. (2019). Dendritic cell subsets in T cell programming: Location dictates function. *Nat Rev Immunol*, *19*(2, 2), 89–103. https://doi.org/10.1038/s41577-018-0088-1

Ermann, J., Hoffmann, P., Edinger, M., Dutt, S., Blankenberg, F. G., Higgins, J. P., Negrin, R. S., Fathman, C. G., & Strober, S. (2005). Only the CD62L+ subpopulation of CD4+CD25+ regulatory T cells protects from lethal acute GVHD. *Blood*, *105*(5), 2220–2226. https://doi.org/10.1182/blood-2004-05-2044

Ferrara, J. L., Levine, J. E., Reddy, P., & Holler, E. (2009). Graft-versus-host disease. *The Lancet*, *373*(9674), 1550–1561. https://doi.org/10.1016/S0140-6736(09)60237-3

Ferreira, L. M. R., Muller, Y. D., Bluestone, J. A., & Tang, Q. (2019). Next-generation regulatory t cell therapy. *Nature Reviews. Drug Discovery*, *18*(10), 749–769. https://doi.org/10.1038/s41573-019-0041-4

Föhse, L., Suffner, J., Suhre, K., Wahl, B., Lindner, C., Lee, C.-W., Schmitz, S., Haas, J. D., Lamprecht, S., Koenecke, C., Bleich, A., Hämmerling, G. J., Malissen, B., Suerbaum, S., Förster, R., & Prinz, I. (2011). High TCR diversity ensures optimal function and homeostasis of Foxp3+ regulatory T cells. *European Journal of Immunology*, *41*(11), 3101–3113. https://doi.org/10.1002/eji.201141986

Gooptu, M., Romee, R., St Martin, A., Arora, M., Al Malki, M., Antin, J. H., Bredeson, C. N., Brunstein, C. G., Chhabra, S., Fuchs, E. J., Ghosh, N., Grunwald, M. R., Kanakry, C. G., Kekre, N., McGuirk, J. P., McNiece, I. K., Mehta, R. S., Mielcarek, M., Milano, F., … Eapen, M. (2021). HLA-haploidentical vs matched unrelated donor transplants with posttransplant cyclophosphamide-based prophylaxis. *Blood*, *138*(3), 273–282. https://doi.org/10.1182/blood.2021011281

Graydon, C. G., Mohideen, S., & Fowke, K. R. (2021). LAG3's enigmatic mechanism of action. *Frontiers in Immunology*, *11*. https://www.frontiersin.org/article/10.3389/fimmu.2020.615317

Grossman, W. J., Verbsky, J. W., Barchet, W., Colonna, M., Atkinson, J. P., & Ley, T. J. (2004). Human t regulatory cells can use the perforin pathway to cause autologous target cell death. *Immunity*, *21*(4), 589–601. https://doi.org/10.1016/j.immuni.2004.09.002

Guo, W., Su, X., Wang, M., Han, M., Feng, X., & Jiang, E. (2021). Regulatory t cells in GVHD therapy. *Frontiers in Immunology*, *12*. https://www.frontiersin.org/article/10.3389/fimmu.2021.697854

Hefazi, M., Bolivar-Wagers, S., & Blazar, B. R. (2021). Regulatory t cell therapy of graft-versus-host disease: Advances and challenges. *International Journal of Molecular Sciences*, *22*(18), 9676. https://doi.org/10.3390/ijms22189676

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*(9), 2754–2759.

Hill, G. R., & Koyama, M. (2020). Cytokines and costimulation in acute graft-versus-host disease. *Blood*, *136*(4), 418–428. https://doi.org/10.1182/blood.201900

0952

Hoeppli, R. E., MacDonald, K. N., Leclair, P., Fung, V. C. W., Mojibian, M., Gillies, J., Rahavi, S. M. R., Campbell, A. I. M., Gandhi, S. K., Pesenacker, A. M., Reid, G., Lim, C. J., & Levings, M. K. (2019). Tailoring the homing capacity of human Tregs for directed migration to sites of Th1-inflammation or intestinal regions. *American Journal of Transplantation*, *19*(1), 62–76. https://doi.org/10.1111/ajt.14936

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. *Journal of Experimental Medicine*, *196*(3), 389–399. https://doi.org/10.1084/jem.20020399

Holler, E., Greinix, H., & Zeiser, R. (2019). *Acute Graft-Versus-Host Disease* (E. Carreras, C. Dufour, M. Mohty, & N. Kröger, Eds.; 7th ed.). Springer. http://www.ncbi.nlm.nih.gov/books/NBK553993/

Hori, S., Nomura, T., & Sakaguchi, S. (2003). Control of regulatory T cell development by the transcription factor Foxp3. *Science (New York, N.Y.)*, *299*(5609), 1057–1061. https://doi.org/10.1126/science.1079490

Hossain, M. S., Kunter, G. M., El-Najjar, V. F., Jaye, D. L., Al-Kadhimi, Z., Taofeek, O. K., Li, J.-M., & Waller, E. K. (2017). PD-1 and CTLA-4 up regulation on donor T cells is insufficient to prevent GvHD in allo-HSCT recipients. *PLOS ONE*, *12*(9), e0184254. https://doi.org/10.1371/journal.pone.0184254

Hrdý, J., Alard, J., Couturier-Maillard, A., Boulard, O., Boutillier, D., Delacre, M., Lapadatescu, C., Cesaro, A., Blanc, P., Pot, B., Ryffel, B., Chamaillard, M., & Grangette, C. (2020). Lactobacillus reuteri 5454 and Bifidobacterium animalis ssp. lactis 5764 improve colitis while differentially impacting dendritic cells maturation and antimicrobial responses. *Scientific Reports*, *10*(1), 5345. https://doi.org/10.1038/s41598-020-62161-1

Ikegawa, S., & Matsuoka, K. (2021). Harnessing treg homeostasis to optimize post-transplant immunity: Current concepts and future perspectives. *Frontiers in Immunology*, *12*. https://www.frontiersin.org/Article/10.3389/fimmu.2021.713358

Izcue, A., Coombes, J. L., & Powrie, F. (2009). Regulatory Lymphocytes and Intestinal Inflammation. *Annual Review of Immunology*, *27*(1), 313–338. https://doi.org/10.1146/annurev.immunol.021908.132657

Jagasia, M. H., Greinix, H. T., Arora, M., Williams, K. M., Wolff, D., Cowen, E. W., Palmer, J., Weisdorf, D., Treister, N. S., Cheng, G.-S., Kerr, H., Stratton, P., Duarte,

R. F., McDonald, G. B., Inamoto, Y., Vigorito, A., Arai, S., Datiles, M. B., Jacobsohn, D., … Flowers, M. E. D. (2015). National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: I. The 2014 Diagnosis and Staging Working Group report. *Biology of Blood and Marrow Transplantation: Journal of the American Society for Blood and Marrow Transplantation*, *21*(3), 389–401.e1. https://doi.org/10.1016/j.bbmt.2014.12.001

Jones, A., & Hawiger, D. (2017). Peripherally Induced Regulatory T Cells: Recruited Protectors of the Central Nervous System against Autoimmune Neuroinflammation. *Frontiers in Immunology*, *8*, 532. https://doi.org/10.3389/fimmu.2017.00532

Jones, A., Opejin, A., Henderson, J. G., Gross, C., Jain, R., Epstein, J. A., Flavell, R. A., & Hawiger, D. (2015). Peripherally Induced Tolerance Depends on Peripheral Regulatory T Cells That Require Hopx To Inhibit Intrinsic IL-2 Expression. *Journal of Immunology (Baltimore, Md.: 1950)*, *195*(4), 1489–1497. https://doi.org/10.4049/jimmunol.1500174

Justiz Vaillant, A. A., Modi, P., & Mohammadi, O. (2022). *Graft Versus Host Disease*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK538235/

Kerdiles, Y. M., Stone, E. L., Beisner, D. L., McGargill, M. A., Ch'en, I. L., Stockmann, C., Katayama, C. D., & Hedrick, S. M. (2010). Foxo Transcription Factors Control Regulatory T Cell Development and Function. *Immunity*, *33*(6), 890–904. https://doi.org/10.1016/j.immuni.2010.12.002

Koch, M. A., Tucker-Heard, G., Perdue, N. R., Killebrew, J. R., Urdahl, K. B., & Campbell, D. J. (2009). The transcription factor T-bet controls regulatory T cell homeostasis and function during type 1 inflammation. *Nature Immunology*, *10*(6), 595–602. https://doi.org/10.1038/ni.1731

MacMillan, M. L., Robin, M., Harris, A. C., DeFor, T. E., Martin, P. J., Alousi, A., Ho, V. T., Bolaños-Meade, J., Ferrara, J. L. M., Jones, R., Arora, M., Blazar, B. R., Holtan, S. G., Jacobsohn, D., Pasquini, M., Socie, G., Antin, J. H., Levine, J. E., & Weisdorf, D. J. (2015). A refined risk score for acute graft-versus-host disease that predicts response to initial therapy, survival, and transplant-related mortality. *Biology of Blood and Marrow Transplantation: Journal of the American Society for Blood and Marrow Transplantation*, *21*(4), 761–767. https://doi.org/10.1016/j.bbmt.2015.01.001

Malard, F., Huang, X.-J., & Sim, J. P. Y. (2020). Treatment and unmet needs in steroid-refractory acute graft-versus-host disease. *Leukemia*, *34*(5), 1229–1240. https://doi.org/10.1038/s41375-020-0804-2

Mamedov, I. Z., Britanova, O. V., Zvyagin, I. V., Turchaninova, M. A., Bolotin, D. A.,

Putintseva, E. V., Lebedev, Y. B., & Chudakov, D. M. (2013). Preparing unbiased T-cell receptor and antibody cDNA libraries for the deep next generation sequencing profiling. *Frontiers in Immunology*, *4*, 456. https://doi.org/10.3389/fimmu.2013.00456

Marski, M., Kandula, S., Turner, J. R., & Abraham, C. (2005). CD18 Is Required for Optimal Development and Function of CD4+CD25+ T Regulatory Cells. *The Journal of Immunology*, *175*(12), 7889–7897. https://doi.org/10.4049/jimmunol.175.12.7889

McDonald-Hyman, C., Flynn, R., Panoskaltsis-Mortari, A., Peterson, N., MacDonald, K. P. A., Hill, G. R., Luznik, L., Serody, J. S., Murphy, W. J., Maillard, I., Munn, D. H., Turka, L. A., Koreth, J., Cutler, C. S., Soiffer, R. J., Antin, J. H., Ritz, J., & Blazar, B. R. (2016). Therapeutic regulatory T-cell adoptive transfer ameliorates established murine chronic GVHD in a CXCR5-dependent manner. *Blood*, *128*(7), 1013–1017. https://doi.org/10.1182/blood-2016-05-715896

Miragaia, R. J., Gomes, T., Chomka, A., Jardine, L., Riedel, A., Hegazy, A. N., Whibley, N., Tucci, A., Chen, X., Lindeman, I., Emerton, G., Krausgruber, T., Shields, J., Haniffa, M., Powrie, F., & Teichmann, S. A. (2019). Single-Cell Transcriptomics of Regulatory T Cells Reveals Trajectories of Tissue Adaptation. *Immunity*, *50*(2), 493–504.e7. https://doi.org/10.1016/j.immuni.2019.01.001

Mizui, M., & Kikutani, H. (2008). Neuropilin-1: the glue between regulatory T cells and dendritic cells? *Immunity*, *28*(3), 302–303. https://doi.org/10.1016/j.immuni.2008.02.012

Moosmann, C., Müller, T. R., Busch, D. H., & Schober, K. (2022). Orthotopic T-cell receptor replacement in primary human T cells using CRISPR-Cas9-mediated homology-directed repair. *STAR Protocols*, *3*(1), 101031. https://doi.org/10.1016/j.xpro.2021.101031

Moyat, M., Bouzourene, H., Ouyang, W., Iovanna, J., Renauld, J.-C., & Velin, D. (2017). IL-22-induced antimicrobial peptides are key determinants of mucosal vaccine-induced protection against H. pylori in mice. *Mucosal Immunology*, *10*(1), 271–281. https://doi.org/10.1038/mi.2016.38

Müller, T. R., Jarosch, S., Hammel, M., Leube, J., Grassmann, S., Bernard, B., Effenberger, M., Andrä, I., Chaudhry, M. Z., Käuferle, T., Malo, A., Cicin-Sain, L., Steinberger, P., Feuchtinger, T., Protzer, U., Schumann, K., Neuenhahn, M., Schober, K., & Busch, D. H. (2021). Targeted t cell receptor gene editing provides predictable t cell product function for immunotherapy. *Cell Reports Medicine*, *2*(8). https://doi.org/10.1016/j.xcrm.2021.100374

Muschaweck, M., Kopplin, L., Ticconi, F., Schippers, A., Iljazovic, A., G'alvez, E. J. C., Abdallah, A. T., Wagner, N., Costa, I. G., Strowig, T., & Pabst, O. (2021). Cognate recognition of microbial antigens defines constricted CD4+ T cell receptor repertoires in the inflamed colon. *Immunity*, *54*(11), 2565–2577.e6. https://doi.org/10.1016/j.immuni.2021.08.014

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*(6), 2062–2068.

Nassereddine, S., Rafei, H., Elbahesh, E., & Tabbara, I. (2017). Acute Graft Versus Host Disease: A Comprehensive Review. *Anticancer Research*, *37*(4), 1547–1555. https://doi.org/10.21873/anticanres.11483

Ojo, B. A., O'Hara, C., Wu, L., El-Rassi, G. D., Ritchey, J. W., Chowanadisai, W., Lin, D., Smith, B. J., & Lucas, E. A. (2019). Wheat Germ Supplementation Increases Lactobacillaceae and Promotes an Anti-inflammatory Gut Milieu in C57BL/6 Mice Fed a High-Fat, High-Sucrose Diet. *The Journal of Nutrition*, *149*(7), 1107–1115. https://doi.org/10.1093/jn/nxz061

Onishi, Y., Fehervari, Z., Yamaguchi, T., & Sakaguchi, S. (2008). Foxp3+ natural regulatory T cells preferentially form aggregates on dendritic cells in vitro and actively inhibit their maturation. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(29), 10113–10118. https://doi.org/10.1073/pnas.0711106105

Ortega, C., Romero, P., Palma, A., Orta, T., Peña, J., García-Vinuesa, A., Molina, I. J., & Santamaría, M. (2004). Role for NKG2-A and NKG2-C surface receptors in chronic CD4+ T-cell responses. *Immunology and Cell Biology*, *82*(6), 587–595. https://doi.org/10.1111/j.0818-9641.2004.01284.x

Pacholczyk, R., Ignatowicz, H., Kraj, P., & Ignatowicz, L. (2006). Origin and t cell receptor diversity of Foxp3+CD4+CD25+ t cells. *Immunity*, *25*(2), 249–259. https://doi.org/10.1016/j.immuni.2006.05.016

Penack, O., Marchetti, M., Ruutu, T., Aljurf, M., Bacigalupo, A., Bonifazi, F., Ciceri, F., Cornelissen, J., Malladi, R., Duarte, R. F., Giebel, S., Greinix, H., Holler, E., Lawitschka, A., Mielke, S., Mohty, M., Arat, M., Nagler, A., Passweg, J., … Basak, G. W. (2020). Prophylaxis and management of graft versus host disease after stem-cell transplantation for haematological malignancies: Updated consensus recommen-

dations of the european society for blood and marrow transplantation. *The Lancet Haematology*, *7*(2), e157–e167. https://doi.org/10.1016/S2352-3026(19)30256-X

Picca, C. C., Larkin, J., Boesteanu, A., Lerman, M. A., Rankin, A. L., & Caton, A. J. (2006). Role of TCR specificity in CD4+ CD25+ regulatory T-cell selection. *Immunological Reviews*, *212*, 74–85. https://doi.org/10.1111/j.0105-2896.2006.00416.x

Qureshi, O. S., Zheng, Y., Nakamura, K., Attridge, K., Manzotti, C., Schmidt, E. M., Baker, J., Jeffery, L. E., Kaur, S., Briggs, Z., Hou, T. Z., Futter, C. E., Anderson, G., Walker, L. S. K., & Sansom, D. M. (2011). Trans-endocytosis of CD80 and CD86: a molecular basis for the cell-extrinsic function of CTLA-4. *Science (New York, N.Y.)*, *332*(6029), 600–603. https://doi.org/10.1126/science.1202947

Ratanatharathorn, V., Nash, R. A., Przepiorka, D., Devine, S. M., Klein, J. L., Weisdorf, D., Fay, J. W., Nademanee, A., Antin, J. H., Christiansen, N. P., Jagt, R. van der, Herzig, R. H., Litzow, M. R., Wolff, S. N., Longo, W. L., Petersen, F. B., Karanes, C., Avalos, B., Storb, R., … Wingard, J. R. (1998). Phase III Study Comparing Methotrexate and Tacrolimus (Prograf, FK506) With Methotrexate and Cyclosporine for Graft-Versus-Host Disease Prophylaxis After HLA-Identical Sibling Bone Marrow Transplantation. *Blood*, *92*(7), 2303–2314. https://doi.org/10.1182/blood.V92.7.2303

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*(3), 895–908. https://doi.org/10.1038/s41375-019-0625-3

Roth, T. L., Puig-Saus, C., Yu, R., Shifrut, E., Carnevale, J., Li, P. J., Hiatt, J., Saco, J., Krystofinski, P., Li, H., Tobin, V., Nguyen, D. N., Lee, M. R., Putnam, A. L., Ferris, A. L., Chen, J. W., Schickel, J.-N., Pellerin, L., Carmody, D., … Marson, A. (2018). Reprogramming human T cell function and specificity with non-viral genome targeting. *Nature*, *559*(7714), 405–409. https://doi.org/10.1038/s41586-018-0326-5

Sakaguchi, S. (2004). Naturally arising CD4+ regulatory t cells for immunologic self-tolerance and negative control of immune responses. *Annual Review of Immunology*, *22*, 531–562. https://doi.org/10.1146/annurev.immunol.21.120601.141122

Sarris, M., Andersen, K. G., Randow, F., Mayr, L., & Betz, A. G. (2008). Neuropilin-1 expression on regulatory T cells enhances their interactions with dendritic cells during antigen recognition. *Immunity*, *28*(3), 402–413. https://doi.org/10.101

6/j.immuni.2008.01.012

Schiering, C., Krausgruber, T., Chomka, A., Fröhlich, A., Adelmann, K., Wohlfert, E. A., Pott, J., Griseri, T., Bollrath, J., Hegazy, A. N., Harrison, O. J., Owens, B. M. J., Löhning, M., Belkaid, Y., Fallon, P. G., & Powrie, F. (2014). The alarmin IL-33 promotes regulatory T-cell function in the intestine. *Nature*, *513*(7519), 564–568. https://doi.org/10.1038/nature13577

Schober, K., Müller, T. R., & Busch, D. H. (2020). Orthotopic T-Cell Receptor Replacement-An "Enabler" for TCR-Based Therapies. *Cells*, *9*(6), E1367. https://doi.org/10.3390/cells9061367

Schumann, K., Raju, S. S., Lauber, M., Kolb, S., Shifrut, E., Cortez, J. T., Skartsis, N., Nguyen, V. Q., Woo, J. M., Roth, T. L., Yu, R., Nguyen, M. L. T., Simeonov, D. R., Nguyen, D. N., Targ, S., Gate, R. E., Tang, Q., Bluestone, J. A., Spitzer, M. H., … Marson, A. (2020). Functional CRISPR dissection of gene networks controlling human regulatory T cell identity. *Nature Immunology*, *21*(11), 1456–1466. https://doi.org/10.1038/s41590-020-0784-4

Schurch, N. J., Schofield, P., Gierliński, M., Cole, C., Sherstnev, A., Singh, V., Wrobel, N., Gharbi, K., Simpson, G. G., Owen-Hughes, T., Blaxter, M., & Barton, G. J. (2016). How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *RNA (New York, N.Y.)*, *22*(6), 839–851. https://doi.org/10.1261/rna.053959.115

Shapiro, R. M., & Antin, J. H. (2020). Therapeutic options for steroid-refractory acute and chronic GVHD: An evolving landscape. *Expert Review of Hematology*, *13*(5), 519–532. https://doi.org/10.1080/17474086.2020.1752175

Sharma, A., & Rudra, D. (2018). Emerging functions of regulatory t cells in tissue homeostasis. *Frontiers in Immunology*, *9*. https://www.frontiersin.org/article/10.3389/fimmu.2018.00883

Shevyrev, D., & Tereshchenko, V. (2020). Treg heterogeneity, function, and homeostasis. *Frontiers in Immunology*, *10*. https://www.frontiersin.org/Article/10.3389/fimmu.2019.03100

Smigiel, K. S., Richards, E., Srivastava, S., Thomas, K. R., Dudda, J. C., Klonowski, K. D., & Campbell, D. J. (2013). CCR7 provides localized access to IL-2 and defines homeostatically distinct regulatory t cell subsets. *Journal of Experimental Medicine*, *211*(1), 121–136. https://doi.org/10.1084/jem.20131142

Song, Y., Wang, N., Chen, L., & Fang, L. (2021). Tr1 cells as a key regulator for maintaining immune homeostasis in transplantation. *Frontiers in Immunology*, *12*.

https://www.frontiersin.org/article/10.3389/fimmu.2021.671579

Sprouse, M. L., Shevchenko, I., Scavuzzo, M. A., Joseph, F., Lee, T., Blum, S., Borowiak, M., Bettini, M. L., & Bettini, M. (2018). Cutting Edge: Low-Affinity TCRs Support Regulatory T Cell Function in Autoimmunity. *The Journal of Immunology*, *200*(3), 909–914. https://doi.org/10.4049/jimmunol.1700156

Ståhl, P. L., Salmén, 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, Å., Pontén, F., Costea, P. I., Sahlén, P., Mulder, J., Bergmann, O., … Frisén, J. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, *353*(6294), 78–82. https://doi.org/10.1126/science.aaf2403

Storb, R., Deeg, H. J., Whitehead, J., Appelbaum, F., Beatty, P., Bensinger, W., Buckner, C. D., Clift, R., Doney, K., & Farewell, V. (1986). Methotrexate and cyclosporine compared with cyclosporine alone for prophylaxis of acute graft versus host disease after marrow transplantation for leukemia. *The New England Journal of Medicine*, *314*(12), 729–735. https://doi.org/10.1056/NEJM198603203141201

Tanaka, J., Toubai, T., Tsutsumi, Y., Miura, Y., Kato, N., Umehara, S., Kahata, K., Mori, A., Toyoshima, N., Ota, S., Kobayashi, T., Kobayashi, M., Kasai, M., Asaka, M., & Imamura, M. (2004). Cytolytic activity and regulatory functions of inhibitory NK cell receptor-expressing T cells expanded from granulocyte colony-stimulating factor-mobilized peripheral blood mononuclear cells. *Blood*, *104*(3), 768–774. https://doi.org/10.1182/blood-2003-11-3870

Taylor, P. A., Panoskaltsis-Mortari, A., Swedin, J. M., Lucas, P. J., Gress, R. E., Levine, B. L., June, C. H., Serody, J. S., & Blazar, B. R. (2004). L-selectinhi but not the l-selectinlo CD4+25+ t-regulatory cells are potent inhibitors of GVHD and BM graft rejection. *Blood*, *104*(12), 3804–3812. https://doi.org/10.1182/blood-2004-05-1850

Teshima, T., Reddy, P., & Zeiser, R. (2016). Acute Graft-versus-Host Disease: Novel Biological Insights. *Biology of Blood and Marrow Transplantation: Journal of the American Society for Blood and Marrow Transplantation*, *22*(1), 11–16. https://doi.org/10.1016/j.bbmt.2015.10.001

Ukena, S. N., Velaga, S., Geffers, R., Grosse, J., Baron, U., Buchholz, S., Stadler, M., Bruder, D., Ganser, A., & Franzke, A. (2011). Human regulatory T cells in allogeneic stem cell transplantation. *Blood*, *118*(13), e82–92. https://doi.org/10.1182/blood-2011-05-352708

Velaga, S., Ukena, S. N., Dringenberg, U., Alter, C., Pardo, J., Kershaw, O., & Franzke,

A. (2015). Granzyme A Is Required for Regulatory T-Cell Mediated Prevention of Gastrointestinal Graft-versus-Host Disease. *PloS One*, *10*(4), e0124927. https://doi.org/10.1371/journal.pone.0124927

Verneris, M. R., Lee, S. J., Ahn, K. W., Wang, H.-L., Battiwalla, M., Inamoto, Y., Fernandez-Vina, M. A., Gajewski, J., Pidala, J., Munker, R., Aljurf, M., Saber, W., Spellman, S., & Koreth, J. (2015). HLA-mismatch is associated with worse outcomes after unrelated donor reduced intensity conditioning hematopoietic cell transplantation: An analysis from the CIBMTR. *Biology of Blood and Marrow Transplantation : Journal of the American Society for Blood and Marrow Transplantation*, *21*(10), 1783–1789. https://doi.org/10.1016/j.bbmt.2015.05.028

Wade, J. A., Katovich Hurley, C., Takemoto, S. K., Thompson, J., Davies, S. M., Fuller, T. C., Rodey, G., Confer, D. L., Noreen, H., Haagenson, M., Kan, F., Klein, J., Eapen, M., Spellman, S., & Kollman, C. (2007). HLA mismatching within or outside of cross-reactive groups (CREGs) is associated with similar outcomes after unrelated hematopoietic stem cell transplantation. *Blood*, *109*(9), 4064–4070. https://doi.org/10.1182/blood-2006-06-032193

Wang, J., Ioan-Facsinay, A., Voort, E. I. H. van der, Huizinga, T. W. J., & Toes, R. E. M. (2007). Transient expression of FOXP3 in human activated nonregulatory CD4+ T cells. *European Journal of Immunology*, *37*(1), 129–138. https://doi.org/10.1002/eji.200636435

Wang, Y., Kissenpfennig, A., Mingueneau, M., Richelme, S., Perrin, P., Chevrier, S., Genton, C., Lucas, B., DiSanto, J. P., Acha-Orbea, H., Malissen, B., & Malissen, M. (2008). Th2 lymphoproliferative disorder of LatY136F mutant mice unfolds independently of TCR-MHC engagement and is insensitive to the action of Foxp3+ regulatory T cells. *Journal of Immunology (Baltimore, Md.: 1950)*, *180*(3), 1565–1575. https://doi.org/10.4049/jimmunol.180.3.1565

Williams, K. M., Inamoto, Y., Im, A., Hamilton, B., Koreth, J., Arora, M., Pusic, I., Mays, J. W., Carpenter, P. A., Luznik, L., Reddy, P., Ritz, J., Greinix, H., Paczesny, S., Blazar, B. R., Pidala, J., Cutler, C., Wolff, D., Schultz, K. R., … Sarantopoulos, S. (2021). National Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host Disease: I. The 2020 Etiology and Prevention Working Group Report. *Transplantation and Cellular Therapy*, *27*(6), 452–466. https://doi.org/10.1016/j.jtct.2021.02.035

Wing, J. B., & Sakaguchi, S. (2011). TCR diversity and Treg cells, sometimes more is more. *European Journal of Immunology*, *41*(11), 3097–3100. https://doi.org/10.1002/eji.201142115

Wing, J. B., Tanaka, A., & Sakaguchi, S. (2019). Human FOXP3+ Regulatory T Cell Heterogeneity and Function in Autoimmunity and Cancer. *Immunity*, *50*(2), 302–316. https://doi.org/10.1016/j.immuni.2019.01.020

Wölfl, M., Qayed, M., Benitez Carabante, M. I., Sykora, T., Bonig, H., Lawitschka, A., & Diaz-de-Heredia, C. (2022). Current prophylaxis and treatment approaches for acute graft-versus-host disease in haematopoietic stem cell transplantation for children with acute lymphoblastic leukaemia. *Frontiers in Pediatrics*, *9*. https://www.frontiersin.org/article/10.3389/fped.2021.784377

Wu, Y., Borde, M., Heissmeyer, V., Feuerer, M., Lapan, A. D., Stroud, J. C., Bates, D. L., Guo, L., Han, A., Ziegler, S. F., Mathis, D., Benoist, C., Chen, L., & Rao, A. (2006). FOXP3 Controls Regulatory T Cell Function through Cooperation with NFAT. *Cell*, *126*(2), 375–387. https://doi.org/10.1016/j.cell.2006.05.042

Wyss, L., Stadinski, B. D., King, C. G., Schallenberg, S., McCarthy, N. I., Lee, J. Y., Kretschmer, K., Terracciano, L. M., Anderson, G., Surh, C. D., Huseby, E. S., & Palmer, E. (2016). Affinity for self antigen selects Treg cells with distinct functional properties. *Nature Immunology*, *17*(9), 1093–1101. https://doi.org/10.1038/ni.3522

Yamaguchi, T., Wing, J. B., & Sakaguchi, S. (2011). Two modes of immune suppression by Foxp3+ regulatory T cells under inflammatory or non-inflammatory conditions. *Seminars in Immunology*, *23*(6), 424–430. https://doi.org/10.1016/j.smim.2011.10.002

Yi, T., Li, X., Yao, S., Wang, L., Chen, Y., Zhao, D., Johnston, H. F., Young, J. S., Liu, H., Todorov, I., Forman, S. J., Chen, L., & Zeng, D. (2011). Host APCs augment in vivo expansion of donor natural regulatory T cells via B7H1/B7.1 in allogeneic recipients. *Journal of Immunology (Baltimore, Md.: 1950)*, *186*(5), 2739–2749. https://doi.org/10.4049/jimmunol.1002939

Yolcu, E. S., Kaminitz, A., Mizrahi, K., Ash, S., Yaniv, I., Stein, J., Shirwan, H., & Askenasy, N. (2013). Immunomodulation with donor regulatory T cells armed with Fas-ligand alleviates graft-versus-host disease. *Experimental Hematology*, *41*(10), 903–911. https://doi.org/10.1016/j.exphem.2013.04.016

Yu, X., Harden, K., C Gonzalez, L., 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. *Nature Immunology*, *10*(1), 48–57. https://doi.org/10.1038/ni.1674

Zaiss, D. M. W., van Loosdregt, J., Gorlani, A., Bekker, C. P. J., Gröne, A., Sibilia, M.,

van Bergen en Henegouwen, P. M. P., Roovers, R. C., Coffer, P. J., & Sijts, A. J. A. M. (2013). Amphiregulin enhances regulatory t cell-suppressive function via the epidermal growth factor receptor. *Immunity*, *38*(2), 275–284. https://doi.org/10.1016/j.immuni.2012.09.023

Zeiser, R. (2019). Advances in understanding the pathogenesis of graft-versus-host disease. *British Journal of Haematology*, *187*(5), 563–572. https://doi.org/10.1111/bjh.16190

Zeiser, R., & Blazar, B. R. (2017). Acute Graft-versus-Host Disease - Biologic Process, Prevention, and Therapy. *The New England Journal of Medicine*, *377*(22), 2167–2179. https://doi.org/10.1056/NEJMra1609337

Zemmour, D., Zilionis, R., Kiner, E., Klein, A. M., Mathis, D., & Benoist, C. (2018). Single-cell gene expression reveals a landscape of regulatory T cell phenotypes shaped by the TCR. *Nature Immunology*, *19*(3, 3), 291–301. https://doi.org/10.1038/s41590-018-0051-0

Zhao, D., Zhang, C., Yi, T., Lin, C.-L., Todorov, I., Kandeel, F., Forman, S., & Zeng, D. (2008). In vivo-activated CD103+CD4+ regulatory T cells ameliorate ongoing chronic graft-versus-host disease. *Blood*, *112*(5), 2129–2138. https://doi.org/10.1182/blood-2008-02-140277

Zheng, X., & Tian, Z. (2021). Which is better, HLA-matched sibling or haploidentical transplantation? *Cellular & Molecular Immunology*, *18*(5), 1347–1347. https://doi.org/10.1038/s41423-021-00640-9

Zheng, Y., & Rudensky, A. Y. (2007). Foxp3 in control of the regulatory T cell lineage. *Nature Immunology*, *8*(5), 457–462. https://doi.org/10.1038/ni1455

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., Benner, C., & Chanda, S. K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Communications*, *10*(1), 1523. https://doi.org/10.1038/s41467-019-09234-6

Software-related Literature

Andrews, S. (2015). *FastQC*. https://qubeshub.org/resources/fastqc

Bolotin, D. A., Poslavsky, S., Mitrophanov, I., Shugay, M., Mamedov, I. Z., Putintseva, E. V., & Chudakov, D. M. (2015). MiXCR: software for comprehensive adaptive immunity profiling. *Nature Methods*, *12*(5), 380–381. https://doi.org/10.1038/nmeth.3364

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635

Freeman, T. C., Horsewell, S., Patir, A., Harling-Lee, J., Regan, T., Shih, B. B., Prendergast, J., Hume, D. A., & Angus, T. (2020). *Graphia: A platform for the graph-based visualisation and analysis of complex data*. 2020.09.02.279349. https://doi.org/10.1101/2020.09.02.279349

Gu, Z., Gu, L., Eils, R., Schlesner, M., & Brors, B. (2014). circlize Implements and enhances circular visualization in R. *Bioinformatics (Oxford, England)*, *30*(19), 2811–2812. https://doi.org/10.1093/bioinformatics/btu393

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M. E., Lönnerberg, P., Furlan, A., Fan, J., Borm, L. E., Liu, Z., Bruggen, D. van, Guo, J., He, X., Barker, R., Sundström, E., Castelo-Branco, G., … Kharchenko, P. V. (2018). RNA velocity of single cells. *Nature*, *560*(7719), 494–498. https://doi.org/10.1038/s41586-018-0414-6

McGinnis, C. S., Murrow, L. M., & Gartner, Z. J. (2019). DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Systems*, *8*(4), 329–337.e4. https://doi.org/10.1016/j.cels.2019.03.003

Shugay, M., Britanova, O. V., Merzlyak, E. M., Turchaninova, M. A., Mamedov, I. Z., Tuganbaev, T. R., Bolotin, D. A., Staroverov, D. B., Putintseva, E. V., Plevova, K., Linnemann, C., Shagin, D., Pospisilova, S., Lukyanov, S., Schumacher, T. N., & Chudakov, D. M. (2014). Towards error-free profiling of immune repertoires. *Nature Methods*, *11*(6), 653–655. https://doi.org/10.1038/nmeth.2960

Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics (Oxford, England)*, *30*(5), 614–620. https://doi.org/10.1093/bioinformatics/btt593

Zheng, G. X. Y., 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., … Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nature Communications*, *8*(1), 14049. https://doi.org/10.1038/ncomms14049

R packages and documentation

Ahlmann-Eltze, C., & Patil, I. (2021). *Ggsignif: Significance brackets for ggplot2*. https://CRAN.R-project.org/package=ggsignif

Allaire, J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., Wickham, H., Cheng, J., Chang, W., & Iannone, R. (2022). *Rmarkdown: Dynamic documents for r*. https://CRAN.R-project.org/package=rmarkdown

Carlson, M. (2021). *Org.mm.eg.db: Genome wide annotation for mouse*.

Chen, Y., Lun, A. T., McCarthy, D. J., Ritchie, M. E., Phipson, B., Hu, Y., Zhou, X., Robinson, M. D., & Smyth, G. K. (2021). *edgeR: Empirical analysis of digital gene expression data in r*.

Dragulescu, A., & Arendt, C. (2020). *Xlsx: Read, write, format excel 2007 and excel 97/2000/XP/2003 files*. https://github.com/colearendt/xlsx

Gentleman, R., Carey, V., Morgan, M., & Falcon, S. (2021). *Biobase: Base functions for bioconductor*. https://bioconductor.org/packages/Biobase

Hoffman, P. (2022). *Seurat: Tools for single cell genomics*. https://CRAN.R-project.org/package=Seurat

Ihaka, R., Murrell, P., Hornik, K., Fisher, J. C., Stauffer, R., Wilke, C. O., McWhite, C. D., & Zeileis, A. (2022). *Colorspace: A toolbox for manipulating and assessing colors and palettes*. https://CRAN.R-project.org/package=colorspace

Kassambara, A. (2020). *Ggpubr: ggplot2 based publication ready plots*. https://rpkgs.datanovia.com/ggpubr/

Kolde, R. (2019). *Pheatmap: Pretty heatmaps*. https://CRAN.R-project.org/package=pheatmap

Konopka, T. (2022). *Umap: Uniform manifold approximation and projection*. https://github.com/tkonopka/umap

Morgan, M. (2022). *BiocManager: Access the bioconductor project package repository*. https://CRAN.R-project.org/package=BiocManager

Müller, K., & Wickham, H. (2022). *Tibble: Simple data frames*. https://CRAN.R-project.org/package=tibble

Nazarov, V. I., Tsvetkov, V. O., Rumynskiy, E., Popov, A. A., Balashov, I., & Volobueva, M. (2022). *Immunarch: Bioinformatics analysis of t-cell and b-cell immune repertoires*. https://CRAN.R-project.org/package=immunarch

Pagès, H., Carlson, M., Falcon, S., & Li, N. (2021). *AnnotationDbi: Manipulation of SQLite-based annotations in bioconductor*. https://bioconductor.org/packages/AnnotationDbi

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Satija, R., Butler, A., Hoffman, P., & Stuart, T. (2022). *SeuratObject: Data structures for single cell data*. https://CRAN.R-project.org/package=SeuratObject

Shannon, J. (2022). *Ecoflux: Functions for ecological flux studies including sap flux and soil/stem c efflux.*

Slowikowski, K. (2021). *Ggrepel: Automatically position non-overlapping text labels with ggplot2*. https://github.com/slowkow/ggrepel

Smyth, G., Hu, Y., Ritchie, M., Silver, J., Wettenhall, J., McCarthy, D., Wu, D., Shi, W., Phipson, B., Lun, A., Thorne, N., Oshlack, A., de Graaf, C., Chen, Y., Langaas, M., Ferkingstad, E., Davy, M., Pepin, F., & Choi, D. (2021). *Limma: Linear models for microarray data*. http://bioinf.wehi.edu.au/limma

Team, T. B. D. (2021). *BiocGenerics: S4 generic functions used in bioconductor*. https://bioconductor.org/packages/BiocGenerics

Warnes, G. R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., Lumley, T., Maechler, M., Magnusson, A., Moeller, S., Schwartz, M., & Venables, B. (2022). *Gplots: Various r programming tools for plotting data*. https://github.com/talgalili/gplots

Wickham, H. (2020). *reshape2: Flexibly reshape data: A reboot of the reshape package*. https://github.com/hadley/reshape

Wickham, H. (2021). *Pryr: Tools for computing on the language*. https://github.com/hadley/pryr

Wickham, H., & Bryan, J. (2022). *Readxl: Read excel files*. https://CRAN.R-project.org/package=readxl

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., & Dunnington, D. (2022). *ggplot2: Create elegant data visualisations using the grammar of graphics*. https://CRAN.R-project.org/package=ggplot2

Wickham, H., François, R., Henry, L., & Müller, K. (2022). *Dplyr: A grammar of data manipulation*. https://CRAN.R-project.org/package=dplyr

Wickham, H., & Girlich, M. (2022). *Tidyr: Tidy messy data*. https://CRAN.R-project.org/package=tidyr

Wickham, H., Hester, J., & Bryan, J. (2022). *Readr: Read rectangular text data*. https://CRAN.R-project.org/package=readr

Wickham, H., & Seidel, D. (2022). *Scales: Scale functions for visualization*. https://CRAN.R-project.org/package=scales

Wilke, C. O. (2020). *Cowplot: Streamlined plot theme and plot annotations for ggplot2*. https://wilkelab.org/cowplot/

Xie, Y. (2022a). *Bookdown: Authoring books and technical documents with r markdown*. https://CRAN.R-project.org/package=bookdown

Xie, Y. (2022b). *Knitr: A general-purpose package for dynamic report generation in r*. https://yihui.org/knitr/

Zhu, H. (2022). *kableExtra: Construct complex table with kable and pipe syntax*.

APPENDIX A

List of Abbreviations

In Table A.1, all abbreviations used within this thesis are described.

Table A.1: List abbreviations with descriptions.

Abbreviation/Unit	Description
°C	degree celcius
5'CDS Primer	5' cDNA synthesis primer
A	adenine
aGvHD	acute graft versus host disease
ALL	acute lymphoblastic leukemia
allo	allogeneic (eng.)/ allogen (dt.)
allo Treg	alloantigen-specifically expanded Treg
allo-Treg	Alloantigen-spezifisch expandierte Treg
AML	acute myeloid leukemia
Anml.	animal number
APC	antigen-presenting cell
AREG	amphiregulin
ATP	adenosine triphosphate
bas./res.	baseline, homeostatic model of resident Treg
BM	bone marrow
BMT	bone marrow transplantation

Table A.1: List abbreviations with descriptions. *(continued)*

Abbreviation/Unit	Description
C	cytosine
CB	cord blood
cDNA	complementary DNA
COI	contrast of interest
CPM	counts per million
CTL	cytotoxic T-lymphocyte
CTLA-4	cytotoxic T-lymphocyte-associated protein 4
d	day(s)
DAMP	danger-associated molecular pattern
DC(s)	dendritic cell(s)
DE	differentially expressed
DEG(s)	differentially exressed gene(s)
DGE	differential gene expression
DNA	deoxyribonucleic acid
dNTP(s)	deoxynucleoside triphosphate(s)
don.	in vitro expanded donor Treg
DTT	dithioreitol
EBMT	European Society for Blood and Marrow
ECP	extracorporeal photochemotherapy
EtOH	ethanol
Exp. No	experiment number
FACS	fluorescence-activated cell sorting
FASL	Fas ligand
FC	fold change
for	forward
FOXP3	forkhead box protein P3
Frag.	fragmentation timeused in SMART-Seq
g	gram
G	guanine
GEX	gene expression
GOI(s)	gene(s) of interest
GSEA	gene set enrichment analysis
GSP_TR[A/B]_rev	gene-specific Tcra/Tcrb reverse primer
GvHD	graft-versus-host disease

Table A.1: List abbreviations with descriptions. *(continued)*

Abbreviation/Unit	Description
Gy	Gray
GZMA/GZMB	granzyme A/B
h	hour(s)
haplo	haploidentical
HSC	hematopoietic stem cell
HSCT	hematopoietic stem cell transplantation
HSZT	hämatopoetische Stammzelltransplantation
IL	interleukin
iSI	inverse Simpson Index
iTreg	peripherally induced regulatory T cell(s)
IVTE	in vitro expansion type
kb	kilobase(s)
KLRG-1	killer cell lectin like receptor G1
KMT	Knochenmarkstransplantation
LPS	lipopolysaccharide
LT	lymphoid tissue-like
M. No	internal number for the mouse/ mouse culture/ pool of mice
MHC	major histocompatibility complex
MIG	molecular identifier group
min	minute(s)
mL	milliliter
Mm	mus musculus
MM	mastermix
MMUD	HLA-mismatched unrelated donor
ms	millisecond(s)
MSC	myeloid suppressor cell
MSD	HLA-matched sibling donor
MUD	HLA-matched unrelated donor
МФ	macrophage(s)
NaCl	sodium chloride
ng	nanogram
NGS	next generation sequencing

Table A.1: List abbreviations with descriptions. *(continued)*

Abbreviation/Unit	Description
NK cell	natural killer cell
NKR	natural killer cell receptor
NLT	non-lymphoid tissue-like
nt	nucleotide(s)
o/n	over night
Padj.	adjusted p value
PB	peripheral blood
PBS	phosphate-buffered saline
PCR2	number PCR2 cycles used in SMART-Seq protocol
poly Treg	polyclonally expanded Treg
poly-Treg	polyklonal expandierte Treg
prc./BMTctrl./ctrl.	BMT control model
PRF1	perforin
pro./prohyl.	GvHD model, Prophylaxis
QC	quality control
RACE	Rapid Amplification of cDNA Ends
RCT	read count table
rev	rev
RNA	ribonucleic acid
RNA-seq	total RNA-sequencing
RPKM	reads per kilobase per million
rpm	rotations per minute
RT	room temperature
RTBM	reverse transcription buffer mix
s	second(s)
S. No	internal sample number
Sample ID	internal sample ID
SE	standard error of the mean
SELL	selectin L
SOP	standard operating procedure
ST-2	suppression of tumorigenicity 2
T	thymine
t-SNE	t-distributed stochastic neighbor embedding

Table A.1: List abbreviations with descriptions. *(continued)*

Abbreviation/Unit	Description
TBI	total body irradiation
TCD BM	T cell-depleted bone marrow
Tcon	conventional T cell(s)
TCR	T cell receptor
TCR_UMI_TSO	smarter template-switching oligonucleotide
TCRrep-seq	TCR repertoire sequencing
TDG	G/T mismatch-specific thymine-dna glycosylase
TF	transcription factor
TGF-β	transforming growth factor beta 1
Th cell	T helper cell
the.	GvHD model, Therapy
TNF-α	tumor necrosis factor alpha
Tr1 cell	type 1 regulatory T cell
Treg	regulatory T cell(s)
TSEA	TCR set enrichment analysis
TSO	template-switching oligonucleotide
Tx	transplantation
Tx198 No	internal descriptor for transplant experiment
TZR	T-Zell-Rezeptor
UDG	Uracil-DNA Gylcosylase
UMAP	uniform manifold approximation and projection
UMI	unique molecular identifier
v/v	volume per volume
VDJ	TCR repertoire (10x Genomics libraries)
w/o	without
w/v	weight per volume
µg	microgram
µL	microliter

APPENDIX B

Sample and library metrics

In Tables B.2, B.3, B.4 and B.5, all relevant bulk library metrics and metadata for Treg from the baseline, BMT control (“prc”), prophylaxis (“pro”) and therapy (“the”) models are provided, respectively. The donor Treg populations are additionally summarized in a separate table (Table B.1). The metrics and metadata for single-cell libraries (GEX and VDJ) are shown in Tables B.6 and B.7, respectively.

The following abbreviations apply to all bulk library tables: S. No: internal sample number; M. No: Internal number for the mouse (re-isolated BMT Treg), the mouse culture (expanded Treg) or the pool of mice (baseline Treg); Sample ID: internal sample ID; Tx198 No: internal descriptor for transplant experiment; Exp. No: experiment number corresponding to this thesis; IVTE: type of *in vitro* expansion; MIGs: molecular identifier groups corresponding to individual UMIs; Workflow: applied workflow in SMART-Seq library preparation; Frag. [min]: fragmentation time applied in SMART-Seq library preparation; PCR2: number of PCR2 cycles used in SMART-Seq library preparation; Anml.: animal number in accordance with shown results.

Table B.1: Bulk library metrics and metadata for donor Treg samples.

								TCRrep-seq libraries		SMART-Seq libraries			
S. No	Sample ID	Culture time [d]	IVTE	BMT	Exp. No	Cell Count	RIN	Tra MIGs	Trb MIGs	Workflow	Frag. [min]	PCR2	Unique Reads
5	176_d14_rp	d14	NA	NA	NA	1000000	7.0	NA	NA	NA	NA	NA	NA
23	196_d14_ra	d14	NA	NA	NA	800000	10.0	NA	NA	NA	NA	NA	NA
26	198_d13_ra	d13	allo	pro	2	3000000	10.0	363756	225287	sp_low	6	12	11750488
27	198_d13_rp	d13	poly	pro	2	3000000	10.0	363945	236897	sp_low	6	12	11601442
64	199_d12_ra	d12	allo	pro	1	3000000	9.6	416994	350186	sp_low	6	12	11969967
65	199_d12_rp	d12	poly	pro	1	3000000	9.6	331739	298095	sp_low	6	12	11631357
102	208_d11_ra	d11	allo	the	2	12600000	7.9	225577	539682	sp_low	6	11	10817821
103	208_d11_rp	d11	poly	the	2	6000000	9.2	298946	579305	sp_low	6	11	14086614
137	210_d11_ra	d11	allo	the	1	3000000	9.5	220751	505470	sp_low	6	11	16118900
138	210_d11_rp	d11	poly	the	1	6000000	9.7	458893	565137	sp_low	6	11	13590714
217	241_d12_ra	d12	allo	prc	1	10000000	10.0	83357	278688	sp_low	6	12	15837908
218	241_d12_rp	d12	poly	prc	1	10000000	10.0	117378	338161	sp_low	6	12	15413920
237	245_d12_ra	d12	allo	prc	2	10000000	9.8	343000	624585	sp_low	6	12	14085733
238	245_d12_rp	d12	poly	prc	2	10000000	9.8	287403	431760	sp_low	6	12	14418397

Table B.2: Bulk library metrics and metadata for baseline samples.

								TCRrep-seq libraries		SMART-Seq libraries			
S. No	M. No	Sample ID	CD62L	Origin	Pool	Cell Count	RIN	Tra MIGs	Trb MIGs	Workflow	Frag. [min]	PCR2	Unique Reads
215	005	005_C_rn_neg	neg	colon	5	53000	NA	NA	NA	sp_low	6	14	16741604
261	007	007_C_rn_neg	neg	colon	7	8188	NA	268	319	sp_low	6	14	15378679
262	008	008_C_rn_neg	neg	colon	8	6000	NA	293	310	sp_low	6	14	10023908
263	009	009_C_rn_neg	neg	colon	9	8699	NA	151	236	sp_low	6	14	13492671
264	010	010_C_rn_neg	neg	colon	10	10700	4.2	270	266	sp_low	6	14	14732958
208	001	001_L_rn_neg	neg	liver	1	14000	6.8	737	1677	sp_low	6	14	14510373
210	002	002_L_rn_neg	neg	liver	2	31400	5.2	605	1497	sp_low	6	14	11203030
214	004	004_L_rn_neg	neg	liver	4	33000	6.7	597	1560	sp_low	6	14	12565270
260	010	010_L_rn_neg	neg	liver	10	53000	NA	3567	6570	sp_low	6	14	16342168
207	001	001_L_rn_pos	pos	liver	1	8200	3.9	125	338	sp_low	6	14	13052210
209	002	002_L_rn_pos	pos	liver	2	9500	3.2	25	101	sp_low	6	14	11412005
211	003	003_L_rn_pos	pos	liver	3	4300	7.9	NA	NA	sp_low	6	14	11560694
213	004	004_L_rn_pos	pos	liver	4	8400	NA	138	390	sp_low	6	14	11326806
259	010	010_L_rn_pos	pos	liver	10	3400	NA	84	238	sp_low	6	14	9922170
202	002	002_S_rn_neg	neg	spleen	2	1400000	9.4	170204	339958	sp_low	6	12	16483181
204	003	003_S_rn_neg	neg	spleen	3	960000	9.3	185448	346006	sp_low	6	12	11257781
206	004	004_S_rn_neg	neg	spleen	4	600000	8.5	125175	234498	sp_low	6	12	15653268
258	010	010_S_rn_neg	neg	spleen	10	497000	NA	28009	55252	sp_low	6	12	13206804
201	002	002_S_rn_pos	pos	spleen	2	1200000	9.4	203742	214328	sp_low	6	12	17823342
203	003	003_S_rn_pos	pos	spleen	3	900000	8.8	149086	183849	sp_low	6	12	14743361
205	004	004_S_rn_pos	pos	spleen	4	700000	8.1	142178	223493	sp_low	6	12	15173562
257	010	010_S_rn_pos	pos	spleen	10	779000	NA	30251	73244	sp_low	6	12	14210758
1	176	176_d0_rn_pos	pos	spleen	11	1000000	8.7	154132	33059	NA	NA	NA	NA
12	196	196_d0_rn_pos	pos	spleen	12	1000000	9.5	33033	91049	NA	NA	NA	NA

Table B.3: Bulk library metrics and metadata for BMT control samples.

											TCRrep-seq libraries		SMART-Seq libraries			
S. No	M. No	Sample ID	IVTE	Origin	BMT	Anml. No	Exp. No	Tx198 No	Cell Count	RIN	Tra MIGs	Trb MIGs	Workflow	Frag. [min]	PCR2	Unique Reads
217	241	241_d12_ra	allo	donor	prc	NA	1	10	10000000	10.0	83357	278688	sp_low	6	12	15837908
218	241	241_d12_rp	poly	donor	prc	NA	1	10	10000000	10.0	117378	338161	sp_low	6	12	15413920
219	4181	4181_S_ra	allo	spleen	prc	1	1	10	24600	8.6	5873	6622	sp_low	6	13	15452164
220	4181	4181_L_ra	allo	liver	prc	1	1	10	4600	NA	1123	1754	sp_low	6	14	14802196
221	4181	4181_C_ra	allo	colon	prc	1	1	10	17000	8.8	4795	5827	sp_low	6	13	14339946
222	4183	4183_S_ra	allo	spleen	prc	2	1	10	14600	9.3	12998	28901	sp_low	6	12	15351001
223	4183	4183_L_ra	allo	liver	prc	2	1	10	11600	7.3	2379	4212	sp_low	6	14	14959041
224	4183	4183_C_ra	allo	colon	prc	2	1	10	86000	8.2	4210	4291	sp_low	6	14	14939994
225	4190	4190_S_ra	allo	spleen	prc	3	1	10	6600	9.1	8260	15345	sp_low	6	13	14808714
226	4190	4190_L_ra	allo	liver	prc	3	1	10	7900	6.7	2022	3474	sp_low	6	14	15140592
227	4190	4190_C_ra	allo	colon	prc	3	1	10	62000	NA	1367	1572	sp_low	6	14	15090620
228	4192	4192_S_rp	poly	spleen	prc	1	1	10	22000	NA	2259	3555	sp_low	6	14	14771381
229	4192	4192_L_rp	poly	liver	prc	1	1	10	2000	NA	450	859	sp_low	6	14	15235095
230	4192	4192_C_rp	poly	colon	prc	1	1	10	4600	9.1	7883	9173	sp_low	6	13	14986120
231	4193	4193_S_rp	poly	spleen	prc	2	1	10	14000	9.4	7790	12551	sp_low	6	13	14415941
232	4193	4193_L_rp	poly	liver	prc	2	1	10	5700	NA	1089	2180	sp_low	6	14	15497421
233	4193	4193_C_rp	poly	colon	prc	2	1	10	38000	7.9	4780	5302	sp_low	6	13	15165460
234	4196	4196_S_rp	poly	spleen	prc	3	1	10	20000	8.9	10546	16387	sp_low	6	13	14908260
235	4196	4196_L_rp	poly	liver	prc	3	1	10	3900	NA	914	1684	sp_low	6	14	15369038
236	4196	4196_C_rp	poly	colon	prc	3	1	10	35000	8.7	4686	5796	sp_low	6	13	13790781
237	245	245_d12_ra	allo	donor	prc	NA	2	12	10000000	9.8	343000	624585	sp_low	6	12	14085733
238	245	245_d12_rp	poly	donor	prc	NA	2	12	10000000	9.8	287403	431760	sp_low	6	12	14418397
239	4208	4208_S_ra	allo	spleen	prc	2	2	12	28000	NA	4906	6743	sp_low	6	14	14156222
240	4208	4208_L_ra	allo	liver	prc	2	2	12	6400	NA	599	1066	sp_low	6	14	15130083
241	4208	4208_C_ra	allo	colon	prc	2	2	12	52000	9.5	10027	10158	sp_low	6	13	14476292
242	4209	4209_S_ra	allo	spleen	prc	3	2	12	9100	NA	3236	4397	sp_low	6	14	16120951
243	4209	4209_L_ra	allo	liver	prc	3	2	12	1800	NA	460	801	sp_low	6	14	14472322
244	4209	4209_C_ra	allo	colon	prc	3	2	12	55000	9.6	13355	12569	sp_low	6	12	14811502
245	400	400_S_ra	allo	spleen	prc	1	2	12	50000	9.7	11163	14095	sp_low	6	13	15079329
246	400	400_L_ra	allo	liver	prc	1	2	12	7000	NA	1320	2039	sp_low	6	14	15421066
247	400	400_C_ra	allo	colon	prc	1	2	12	55000	9.7	9126	9500	sp_low	6	12	15865549
248	4214	4214_S_rp	poly	spleen	prc	2	2	12	21000	NA	4072	6393	sp_low	6	14	13624907
249	4214	4214_L_rp	poly	liver	prc	2	2	12	3000	NA	322	556	sp_low	6	14	14439760
250	4214	4214_C_rp	poly	colon	prc	2	2	12	47000	9.6	6823	7060	sp_low	6	13	15359992
251	4216	4216_S_rp	poly	spleen	prc	3	2	12	40000	9.7	6022	8780	sp_low	6	14	15522755

Table B.3: Bulk library metrics and metadata for BMT control samples. *(continued)*

											TCRrep-seq libraries		SMART-Seq libraries			
S. No	M. No	Sample ID	IVTE	Origin	BMT	Anml. No	Exp. No	Tx198 No	Cell Count	RIN	Tra MIGs	Trb MIGs	Workflow	Frag. [min]	PCR2	Unique Reads
252	4216	4216_L_rp	poly	liver	prc	3	2	12	1800	NA	208	340	sp_low	6	14	14985141
253	4216	4216_C_rp	poly	colon	prc	3	2	12	60000	9.6	8978	10172	sp_low	6	13	13391519
254	4205	4205_S_rp	poly	spleen	prc	1	2	12	50000	9.7	8907	12791	sp_low	6	13	14706123
255	4205	4205_L_rp	poly	liver	prc	1	2	12	5300	NA	752	1179	sp_low	6	14	14282638
256	4205	4205_C_rp	poly	colon	prc	1	2	12	53000	9.6	13843	15060	sp_low	6	13	13523016

Table B.4: Bulk library metrics and metadata for prophylaxis samples.

											TCRrep-seq libraries		SMART-Seq libraries			
S. No	M. No	Sample ID	IVTE	Origin	BMT	Anml. No	Exp. No	Tx198 No	Cell Count	RIN	Tra MIGs	Trb MIGs	Workflow	Frag. [min]	PCR2	Unique Reads
64	199	199_d12_ra	allo	donor	pro	NA	1	2	3000000	9.6	416994	350186	sp_low	6	12	11969967
65	199	199_d12_rp	poly	donor	pro	NA	1	2	3000000	9.6	331739	298095	sp_low	6	12	11631357
66	341	341_L_ra	allo	liver	pro	1	1	2	14600	6.6	1457	1677	sp_low	4	13	12190226
67	341	341_C_ra	allo	colon	pro	1	1	2	20000	6.6	1224	959	sp_low	4	13	12175663
68	341	341_S_ra	allo	spleen	pro	1	1	2	100000	9.5	28422	21458	sp_low	6	12	13030172
69	346	346_L_ra	allo	liver	pro	2	1	2	16000	7.5	1451	1655	sp_low	6	13	17217024
70	346	346_C_ra	allo	colon	pro	2	1	2	13200	2.3	2211	2319	pp_ulow	3	11	19839713
71	346	346_S_ra	allo	spleen	pro	2	1	2	120000	9.4	69539	55169	sp_low	6	12	11316174
72	348	348_L_ra	allo	liver	pro	3	1	2	27000	9.6	2042	2232	sp_low	6	13	14932866
73	348	348_C_ra	allo	colon	pro	3	1	2	18000	8.9	289	249	sp_low	6	13	11893679
74	348	348_S_ra	allo	spleen	pro	3	1	2	111000	9.6	58713	48219	sp_low	6	13	12318346
75	343	343_L_rp	poly	liver	pro	1	1	2	12200	8.9	726	871	sp_low	6	14	NA
76	343	343_C_rp	poly	colon	pro	1	1	2	30400	9.2	4418	3384	sp_low	6	13	12078455
77	343	343_S_rp	poly	spleen	pro	1	1	2	50000	9.6	9903	10647	sp_low	6	13	11888225
78	347	347_L_rp	poly	liver	pro	2	1	2	13500	7.9	1232	1665	sp_low	6	14	11894634
79	347	347_C_rp	poly	colon	pro	2	1	2	37000	9.0	3813	3965	sp_low	6	13	12026336
80	347	347_S_rp	poly	spleen	pro	2	1	2	60000	9.7	10210	11277	sp_low	6	13	11980908
81	349	349_L_rp	poly	liver	pro	3	1	2	20600	9.3	1404	1787	sp_low	6	13	12259556
82	349	349_C_rp	poly	colon	pro	3	1	2	30000	8.2	3046	3298	sp_low	6	13	12019691
83	349	349_S_rp	poly	spleen	pro	3	1	2	50000	9.3	6283	7089	sp_low	6	13	11952544
26	198	198_d13_ra	allo	donor	pro	NA	2	1	3000000	10.0	363756	225287	sp_low	6	12	11750488
27	198	198_d13_rp	poly	donor	pro	NA	2	1	3000000	10.0	363945	236897	sp_low	6	12	11601442
28	293	293_L_ra	allo	liver	pro	1	2	1	20900	8.7	2729	3496	sp_low	6	14	10710661
29	293	293_C_ra	allo	colon	pro	1	2	1	4200	1.3	540	645	pp_ulow	3	11	17552173
30	293	293_S_ra	allo	spleen	pro	1	2	1	115700	9.6	42501	48124	sp_low	6	12	12246032
31	294	294_L_ra	allo	liver	pro	2	2	1	17300	6.6	2545	3083	sp_low	4	14	11883124
32	294	294_C_ra	allo	colon	pro	2	2	1	3580	NA	287	351	pp_ulow	3	11	13670174
33	294	294_S_ra	allo	spleen	pro	2	2	1	102000	9.4	25007	28421	sp_low	6	12	14277113
34	296	296_L_ra	allo	liver	pro	3	2	1	14400	9.5	1487	2058	sp_low	6	14	11791873
35	296	296_C_ra	allo	colon	pro	3	2	1	3100	NA	298	376	pp_ulow	3	11	7792444
36	296	296_S_ra	allo	spleen	pro	3	2	1	142000	9.3	41840	38866	sp_low	6	13	12470370
37	297	297_L_rp	poly	liver	pro	1	2	1	11400	9.4	1879	2398	sp_low	6	14	11920321
38	297	297_C_rp	poly	colon	pro	1	2	1	3400	NA	382	302	pp_ulow	3	11	11870266
39	297	297_S_rp	poly	spleen	pro	1	2	1	85700	9.7	20022	22752	sp_low	6	13	12093875
40	298	298_L_rp	poly	liver	pro	2	2	1	7200	3.2	1483	1653	pp_ulow	3	11	43519022

Table B.4: Bulk library metrics and metadata for prophylaxis samples. *(continued)*

											TCRrep-seq libraries		SMART-Seq libraries			
S. No	M. No	Sample ID	IVTE	Origin	BMT	Anml. No	Exp. No	Tx198 No	Cell Count	RIN	Tra MIGs	Trb MIGs	Workflow	Frag. [min]	PCR2	Unique Reads
41	298	298_C_rp	poly	colon	pro	2	2	1	3000	2.1	388	354	pp_ulow	3	11	8657807
42	298	298_S_rp	poly	spleen	pro	2	2	1	30000	9.7	6591	6922	sp_low	6	14	12017064
43	299	299_L_rp	poly	liver	pro	3	2	1	15500	8.1	2114	2647	sp_low	6	13	12012869
44	299	299_C_rp	poly	colon	pro	3	2	1	7100	3.3	896	1195	pp_ulow	3	11	11289454
45	299	299_S_rp	poly	spleen	pro	3	2	1	105000	9.6	19710	26768	sp_low	6	12	11973325

Table B.5: Bulk library metrics and metadata for therapy samples.

											TCRrep-seq libraries		SMART-Seq libraries			
S. No	M. No	Sample ID	IVTE	Origin	BMT	Anml. No	Exp. No	Tx198 No	Cell Count	RIN	Tra MIGs	Trb MIGs	Workflow	Frag. [min]	PCR2	Unique Reads
137	210	210_d11_ra	allo	donor	the	NA	1	5	3000000	9.5	220751	505470	sp_low	6	11	16118900
138	210	210_d11_rp	poly	donor	the	NA	1	5	6000000	9.7	458893	565137	sp_low	6	11	13590714
139	278	278_S_ra	allo	spleen	the	1	1	5	165000	9.3	26512	25407	sp_low	6	11	17932361
140	278	278_L_ra	allo	liver	the	1	1	5	20887	7.6	1751	1404	sp_low	6	14	11845549
141	278	278_C_ra	allo	colon	the	1	1	5	10000	6.9	683	645	sp_low	6	14	12269393
142	279	279_S_ra	allo	spleen	the	2	1	5	251000	8.7	52436	49465	sp_low	6	11	12872215
143	279	279_L_ra	allo	liver	the	2	1	5	10620	5.9	470	465	sp_low	4	16	10950258
144	279	279_C_ra	allo	colon	the	2	1	5	10000	7.9	1112	1031	sp_low	6	14	11499938
145	280	280_S_ra	allo	spleen	the	3	1	5	215000	8.9	34475	39718	sp_low	6	11	17703279
146	280	280_L_ra	allo	liver	the	3	1	5	15100	6.9	3033	3092	sp_low	6	13	12703432
147	280	280_C_ra	allo	colon	the	3	1	5	11000	8.7	1508	1533	sp_low	6	14	11456289
148	281	281_S_rp	poly	spleen	the	1	1	5	119800	9.2	13227	14504	sp_low	6	11	15009726
149	281	281_L_rp	poly	liver	the	1	1	5	6000	NA	355	366	sp_low	4	16	10649865
150	281	281_C_rp	poly	colon	the	1	1	5	6000	6.4	272	285	sp_low	4	16	10552630
151	282	282_S_rp	poly	spleen	the	2	1	5	166000	9.2	27551	27563	sp_low	6	11	14790705
152	282	282_L_rp	poly	liver	the	2	1	5	8800	7.4	1189	1119	sp_low	6	14	11812858
153	282	282_C_rp	poly	colon	the	2	1	5	5600	NA	132	126	pp_ulow	3	10	29059831
154	290	290_S_rp	poly	spleen	the	3	1	5	213000	8.3	21374	24914	sp_low	6	11	14630858
155	290	290_L_rp	poly	liver	the	3	1	5	7000	6.3	35	58	sp_low	4	16	8608969
156	290	290_C_rp	poly	colon	the	3	1	5	566	2.3	30	12	pp_ulow	3	10	NA
102	208	208_d11_ra	allo	donor	the	NA	2	4	12600000	7.9	225577	539682	sp_low	6	11	10817821
103	208	208_d11_rp	poly	donor	the	NA	2	4	6000000	9.2	298946	579305	sp_low	6	11	14086614
104	284	284_S_ra	allo	spleen	the	1	2	4	38000	2.2	23	160	pp_ulow	3	10	NA
105	284	284_L_ra	allo	liver	the	1	2	4	4800	NA	353	312	sp_low	4	16	8815559
106	284	284_C_ra	allo	colon	the	1	2	4	1600	NA	21	21	pp_ulow	3	10	NA
107	288	288_S_ra	allo	spleen	the	2	2	4	53500	3.3	32	52	pp_ulow	3	10	NA
108	288	288_L_ra	allo	liver	the	2	2	4	8100	9.8	1563	1713	sp_low	6	13	16779847
109	288	288_C_ra	allo	colon	the	2	2	4	5380	5.8	357	296	sp_low	4	14	12378813
110	x296	x296_S_rp	poly	spleen	the	1	2	4	78200	7.0	863	939	sp_low	6	14	14319032
111	x296	x296_L_rp	poly	liver	the	1	2	4	13100	9.1	1808	1682	sp_low	6	14	12027943
112	x296	x296_C_rp	poly	colon	the	1	2	4	12686	6.6	1045	950	sp_low	6	13	12748589
113	x297	x297_S_rp	poly	spleen	the	2	2	4	110250	2.9	94	70	pp_ulow	3	10	NA
114	x297	x297_L_rp	poly	liver	the	2	2	4	15994	9.1	2996	2995	sp_low	6	14	16885717
115	x297	x297_C_rp	poly	colon	the	2	2	4	12000	7.7	1607	1395	sp_low	6	14	14848183

Table B.6: Single-cell GEX library (10x) metrics and metadata.

Sample	Number of Reads	Estimated Number of Cells	Mean Reads per Cell	Median Genes per Cell	Valid Barcodes	Reads Mapped to Genome	Fraction Reads in Cells	Total Genes Detected	Median UMI Counts per Cell	Cells per cluster	Cells with TCR
Donor allo	376523583	5560	67720	2332	0.875	0.915	0.829	16567	9369	3066	2037
Colon recipient 1	442390212	7866	56240	1735	0.860	0.913	0.817	17297	4491	5253	3530
Liver recipient 1	345854859	5336	64815	1672	0.886	0.909	0.661	16328	3992	2708	1889
Spleen recipient 1	463924730	6676	69491	1826	0.862	0.912	0.766	16564	4740	3879	2795
Colon recipient 2	455731582	6832	66705	1671	0.882	0.905	0.864	17047	4353	4903	3345
Liver recipient 2	281835717	4633	60832	1659	0.895	0.916	0.723	16124	4064	2580	1841
Spleen recipient 2	406491956	7703	52770	1649	0.892	0.910	0.769	16683	4256	4650	3249
Colon recipient 3	437143148	6593	66304	1739	0.864	0.923	0.896	16721	4574	4948	3317
Liver recipient 3	260585726	3690	70619	1625	0.875	0.915	0.677	15809	3963	1860	1329
Spleen recipient 3	385355848	6265	61509	1813	0.853	0.918	0.784	17095	4755	3610	2557

Table B.7: Single-cell VDJ library (10x) metrics and metadata.

Sample Name	Number of Read Pairs	Estimated Number of Cells	Mean Read Pairs per Cell	Number of Cells With Productive V-J Spanning Pair	Valid Barcodes	Reads Mapped to Any V(D)J Gene	Reads Mapped to TRA	Reads Mapped to TRB	Fraction Reads in Cells
Donor allo*	132123300	19314	6840	18318	0.955	0.890	0.272	0.618	0.456
Donor allo	34278811	5124	6689	4501	0.968	0.934	0.231	0.702	0.655
Colon recipient 1	40069704	6876	5827	5915	0.971	0.935	0.360	0.574	0.745
Liver recipient 1	36121241	5294	6823	4376	0.975	0.948	0.255	0.692	0.589
Spleen recipient 1	54058522	6287	8598	5595	0.972	0.945	0.304	0.641	0.685
Colon recipient 2	55842398	5906	9455	5133	0.975	0.940	0.358	0.582	0.790
Liver recipient 2	33139445	4330	7653	3724	0.975	0.944	0.273	0.670	0.626
Spleen recipient 2	38079303	6918	5504	6039	0.975	0.948	0.311	0.637	0.675
Colon recipient 3	40687781	5734	7095	5039	0.972	0.940	0.354	0.586	0.761
Liver recipient 3	25524413	3491	7311	2921	0.974	0.946	0.268	0.677	0.582
Spleen recipient 3	44480090	5528	8046	4974	0.970	0.936	0.303	0.632	0.697

* Technical replicate of sample ”Donor allo” only a VDJ library was prepared from; a higher input cell number was chosen for a larger TCR repertoire coverage.

APPENDIX C

Supplementary Software

All supplementary software that was used throughout this thesis is listed in Table C.1. For Microsoft Office applications, the license (Volume License 2019) was provided by the University of Regensburg (Regensburg, Germany). All other software was used in compliance with the GNU GENERAL PUBLIC LICENSE.

Table C.1: List of additional software used with specifications and purpose.

Designation	Version	Purpose
Affinity Designer	Version 1.10.5	editing data graphics and schematic illustrations
BBEdit	Version 13.5.7 (415124, 64-bit Intel, sandboxed)	code preparation
bcl2fastq Conversion Software	Version 1.8.4	base calling software
cellranger	Version 4.0.0 / 5.0.0	scData demultiplexing
Doubletfinder	Version 2.0.3	scData double removal
FastQC	Version 0.11.3	sequencing raw data quality control
git	Version 2.32.1 (Apple Git 133)	versioning and backup of analysis and thesis R scripts

Table C.1: List of additional software used with specifications and purpose. *(continued)*

Designation	Version	Purpose
GNU bash	Version 3.2.57(1)-release (x86_64-apple-darwin21)	basic file operations, standardized primary analysis scripts
Graphia	Version 3.0	graph-based clustering of genes into co-expression networks
Illumina Experiment Manager	Version 1.19.1	sample sheet generation
Illumina Sequencing Analysis Viewer	Version 2.4.7	sequencing raw data quality control
Microsoft ® Excel for Mac	Version 16.57 (22011101)	table preparation
Microsoft® PowerPoint for Mac	Version 16.57 (22011101)	presentation preparation
Microsoft® Word for Mac	Version 16.57 (22011101)	text preparation
MIGEC	Version 1.2.9	consensus sequence assembly for MIGs
MiXCR	Version 3.0.1.8	TCRrep-seq data mapping to VDJ database
PEAR	Version 0.9.11	TCRrep-seq paired-end read alignment
R	R version 4.1.0 (2021-05-18)	data analysis and visualization; thesis compilation
RStudio	2022.02.3+492 ”Prairie Trillium” Release	development environment for R scripts and thesis
STAR	Version 2.5.3a	sequencing read alignment
Zotero	Version 5.0.96.3	citation management

APPENDIX D

Acknowledgement

First of all, I want to thank my research supervisor Prof. Dr. Michael Rehli for the continued support throughout the process of generating and analyzing the presented data over the past four years. In particular, his skill to visualize complex data in an easily comprehensible way has broadened my scientific horizon. Moreover, I want to extend my gratitude for the opportunity to participate in an exciting project within the SFB TR221 and, of course, the funding that made my work possible.

Hanna Stanewsky, Margit Nützel, Johanna Raithel and Ute Ackermann deserve my gratitude for their great job in the wet lab. I also want to thank Dr. Nicholas Strieder for his computational support, ideas and consultation. I thankfully look back to the fruitful discussions with Alexander Fischer.

I want to thank the members of the work group around Prof. Dr. Matthias Edinger and PD Dr. Petra Hoffmann for their seamless collaboration and the two group leaders for all the ideas based on their extensive experience in the field of GvHD. Especially Franziska Pielmeier has contributed a rich and continued effort to generate samples that made this work possible.

Furthermore, I want to thank Prof. Dr. Gunter Meister for providing access to the Illumina MiSeq system and Norbert Eichner for his continued assistance in operating the sequencer as well as data transfer.

My mentors Prof. Dr. Richard Warth and Prof. Dr. Thomas Winkler provided valuable feedback as part of my annual research reports and I am thankful for their time and

consideration.

It goes without saying that only due to the tireless support from my partner, Janina Herold, I was able to focus on writing the thesis. She also contributed to my progress by helping me prioritize in a productive manner.