Tissue-specific features of the T cell repertoire following allogeneic hematopoietic cell transplantation in human and mouse

Susan DeWolf[†,1], **Yuval Elhanati**[†,2], **Katherine Nichols**[‡,3], **Nicholas R. Waters**[‡,3], **Chi L. Nguyen**[‡,3], **John B. Slingerland**[3], **Natasia Rodriguez**[4], **Olga Lyudovyk**[2], **Paul A. Giardina**[3], **Anastasia I. Kousa**[3], **Hana Andrlová**[3], **Nick Ceglia**[2], **Teng Fei**[5], **Rajya Kappagantula**[6,7], **Yanyun Li**[7], **Nathan Aleynick**[7], **Priscilla Baez**[4], **Rajmohan Murali**[7], **Akimasa Hayashi**[7,8], **Nicole Lee**[3], **Brianna Gipson**[3], **Madhumitha Rangesa**[4], **Zoe Katsamakis**[3], **Anqi Dai**[3], **Amanda G. Blouin**[7], **Maria Arcila**[7], **Ignas Masilionis**[9], **Ronan Chaligne**[9], **Doris M. Ponce**[4,10], **Heather J. Landau**[4,10], **Ioannis Politikos**[4,10], **Roni Tamari**[4,10], **Alan M. Hanash**[4,10,11], **Robert R. Jenq**[12], **Sergio A. Giralt**[4,10], **Kate A. Markey**[4,13,14], **Yanming Zhang**[7], **Miguel-Angel Perales**[4,10], **Nicholas D. Socci**[15], **Benjamin D. Greenbaum**[2,16], **Christine A. Iacobuzio-Donahue**[6], **Travis J. Hollmann**[7,17], **Marcel R.M. van den Brink**[§,3,4,10], **Jonathan U. Peled**[§,*,4,10]

[1]Leukemia Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

[2]Computational Oncology, Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

[3]Department of Immunology, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

[4]Adult Bone Marrow Transplantation Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

[5]Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

[6]David M. Rubenstein Center for Pancreatic Cancer Research, Memorial Sloan Kettering Cancer Center; New York, NY, USA.

[*]Correspondence: peledj@mskcc.org.
[†]Co-first authors
[‡]Co-second authors
[§]Co-senior authors

[7]Department of Pathology and Laboratory Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

[8]Department of Pathology, Kyorin University, Mitaka City, Tokyo, Japan

[9]Program for Computational and System Biology, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

[10]Weill Cornell Medical College, New York, NY, USA.

[11]Human Oncology & Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

[12]Departments of Genomic Medicine and Stem Cell Transplantation Cellular Therapy, Division of Cancer Medicine, University of Texas MD Anderson Cancer Center, Houston, TX, USA.

[13]Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA.

[14]Division of Medical Oncology, University of Washington; Seattle, WA, USA.

[15]Bioinformatics Core Facility, Memorial Sloan Kettering Cancer Center, New York, NY, USA.

[16]Physiology, Biophysics & Systems Biology, Weill Cornell Medicine, Weill Cornell Medical College, New York, NY, USA.

[17]Bristol Myers Squibb, Lawrenceville, NJ 08540.

Abstract

T cells are the central drivers of many inflammatory diseases, but the repertoire of tissue-resident T cells at sites of pathology in human organs remains poorly understood. We examined the site-specificity of T cell receptor (TCR) repertoires across tissues (5–18 tissues per patient) in prospectively collected autopsies of patients with and without graft-versus-host disease (GVHD), a potentially lethal tissue-targeting complication of allogeneic hematopoietic cell transplantation, as well as in mouse models of GVHD. Anatomic similarity between tissues was a key determinant of TCR repertoire composition within patients, independent of disease or transplant status. The T cells recovered from peripheral blood and spleen in patients and mice captured a limited portion of the TCR repertoire detected in tissues. Whereas few T cell clones were shared across patients, motif-based clustering revealed shared repertoire signatures across patients in a tissue-specific fashion. T cells at disease sites had a tissue-resident phenotype and were of donor origin based on single-cell chimerism analysis. These data demonstrate the complex composition of T cell populations that persist in human tissues at the end-stage of an inflammatory disorder following lymphocyte-directed therapy. These findings also underscore the importance of studying T cell in tissues rather than blood for tissue-based pathologies and suggest the tissue-specific nature of both the endogenous and post-transplant T cell landscape.

One Sentence Summary:

T cell receptor sequencing reveals tissue-specific features of the post-transplant T cell repertoire in mouse and human

INTRODUCTION

T cell immunity in barrier sites such as the skin and gastrointestinal (GI) tract is mediated by both circulating cells that traffic into tissues for specific responses as well as resident populations that persist in tissues for prolonged periods (1–4). The reactivity of T cells underpins not only their pathogen responses and the equilibrium they maintain with commensal organisms, but also homeostatic functions such as cancer surveillance (5). Although circulating T cells comprise a small minority of an individual's total T cell pool (6), organs are replete with tissue-resident T cells, to which various key protective and pathogenic functions have been ascribed (7). The T cell receptor (TCR) repertoire of tissue-resident cells has recently been studied in healthy tissues from organ donors (8), but those at sites of disease have not been well-characterized.

Allogeneic hematopoietic cell transplantation (allo-HCT) offers an opportunity to investigate the TCR repertoire within human tissues as the adaptive immune system is "reset" by a donor graft that rapidly repopulates lymphoid and peripheral tissues (9–11). The curative potential of allo-HCT can be attributed in part to the eradication of hematologic malignancies by alloreactive donor T cells. In 30–50% of recipients, however, alloreactive T cells mediate acute graft-versus-host disease (GVHD), an inflammatory process is triggered by the recognition of recipient antigens and primarily affects the skin, GI tract, and liver (12, 13). The reasons that a specific pattern of organ involvement is observed in any individual patient, among these typical target organs, are not known, although it has been proposed that the disease is initiated in a tissue-localized cascade of antigen-presentation, leading to broad, polyclonal donor T cell expansion (10). The local context in which immune activation occurs is important for GVHD pathogenesis; for example, relevant alloantigens, such as minor histocompatibility antigens, have distinct tissue distributions, (14) and certain cytokines can confer protective or pathogenic effects in different organs (15).

Treatments for GVHD target T cell function, including corticosteroids and continuation of prophylactic drugs such as calcineurin inhibitors (13). Among individuals with the highest-risk GVHD, which typically involves the lower GI tract and fails to respond to corticosteroids, mortality rates approach 50%, even with newer treatments such as the JAK/STAT inhibitor ruxolitinib (16, 17). Because of the inherent challenges in sampling tissue from patients, much of our understanding of GVHD T cell biology—and specifically of the involved T-cell receptor (TCR) repertoire—has been gleaned from animal models (18–21), circulating T cells from peripheral blood of patients (22), and diagnostic skin and GI biopsies (23–32). Recent analyses of human skin and of tissues from non-human primates suggest that donor T cells may acquire a tissue-resident phenotype, including early in the course of acute GVHD (33–35). Data are lacking about the T cells that persist in tissues over the course of GVHD, particularly in patients with steroid-refractory disease.

To address these gaps in understanding, we analyzed TCR repertoires in various organs from allo-HCT patients and mice with GVHD. We assembled an autopsy series of patients who had steroid-refractory GVHD and tracked T cell clones across their tissues, including esophagus, portions of the small and large intestine, liver, skin, and lymphoid structures including the spleen, in as many as 18 sites in a single patient. In parallel, we profiled TCR

repertoires in autopsy tissues from patients without GVHD, as well as from transplanted mice with GVHD as comparators.

RESULTS

Prospective collection of tissues from patients with GVHD by rapid autopsy

We prospectively enrolled adult recipients of allo-HCT whose course was complicated by steroid-refractory GVHD to a research-autopsy study over a three-year period (Table 1) (36). The patients (n = 7) received grafts of hematopoietic cells from fully matched donors (at ten HLA alleles); the grafts were infused either unmodified (n = 4) or following enrichment for CD34$^+$ cells by magnetic-bead sorting to deplete potentially alloreactive T cells (n = 3). Although CD34 selection was intended to decrease the risk of acute GVHD and recipients were not given pharmacologic prophylaxis (37), these three patients nevertheless developed GVHD. All the patients had GVHD involving the GI tract and received several lines of GVHD-directed therapy (Table 1, Tables S1–3); five also had skin involvement. All patients had acute GVHD, a T cell-driven process; two had an overlap syndrome, encompassing features of both acute and chronic GVHD, a process in which B cells are considered to play a key pathophysiological role (38). During autopsies, which were conducted following a median post-mortem interval of 9.75 hours (range 2–15.75), samples were collected from both GVHD-affected and non-GVHD-affected tissues, with an emphasis on GI tissues from the esophagus to the rectum, liver, skin, lymphoid organs including spleen (Fig. 1A). We were able to collect and viably preserve mononuclear cells from marrow, caval or cardiac blood from four GVHD patients.

Identification of unique T cell clones from tissues sampled during rapid autopsy

To characterize TCR repertoires, we sequenced the β-chain complementarity-determining region 3 (CDR3) from genomic DNA extracted from snap-frozen tissues from seven allo-HCT recipients (patients A-G; Fig. 1B–E, Methods) and from three patients enrolled in the same research-autopsy protocol who had non-hematologic malignancies and did not undergo HCT (patients H-J; Fig. 1B); in parallel, we sequenced tissues from a mouse model of GVHD (Figs. 1F–I, Figs. S1A,B). From 105 human samples, we recovered a median of 44,540 (standard deviation [SD] 14,887) total distinct clones per patient (Fig. 1B) and a median of 1597 (SD 4,263; range 86–29,290) unique clones per individual non-PBMC/marrow tissue (Fig. 1B, Table S4, sampling efficiency estimates in Methods). There was no correlation between the time to autopsy and number of unique clones recovered from GI-tract samples (Fig. S1C–D).

We tracked T cell clones across tissues, as defined by the nucleotide sequence of their TCR CDR3—the primary determinant of antigen specificity. This approach is similar to those employed in previous studies of blood and diagnostic biopsies of GVHD tissues (23–32). Several samples had a top clone (defined as the clone with the highest frequency in a given sample) that was highly abundant (frequency 12–63%; Fig. S1E, Table S4); this was particularly evident in those from patients D and E (both recipients of CD34-selected grafts from older donors, Table 1), which also had the highest clonality measures (Fig. 1E), suggesting expansion of clones in these tissues. Of note, the top clone in an individual tissue

was not the top sampled clone across all tissues from a given patient, which is explored in detail in subsequent figures. We interrogated the dataset for known public virus-reactive clones (39–41), and although some were present at low frequency, they did not account for dominant clones across most tissues (Table S5). In this heterogenous cohort, we did not observe associations between tissue repertoire diversity and transplant-related clinical variables that have previously been associated with diversity in the circulating T cell pool (p > 0.05 for graft source, CMV reactivation, patient age category, donor age category, and GVHD severity in linear mixed-effect models, Fig. S2) (42–47).

TCR repertoires in different tissues are distinct, while anatomically related tissues harbor greater TCR overlap

We next examined the similarity of the T cell repertoire across anatomy within each patient by quantifying repertoire overlap between tissue pairs via the Jensen-Shannon divergence (JSD) metric, a normalized index in which a value of 1 indicates complete repertoire divergence and 0 indicates identical clonal populations (Table S6). For example, pairwise comparisons between 15 samples from Patient D revealed the greatest repertoire overlap between small intestine and large intestine samples (dark-purple areas in Fig. 2A). The skin was divergent (yellow-to-green in color) from most GI tract samples. The large degree of overlap between different segments of the colon in patient D is reflected by the low JSD value of 0.14 (blue points in Fig. 2B, left panel clustered along the diagonal), with fewer clones shared between the descending colon and skin (JSD 0.62, Fig. 2B, center). An intermediate degree of sharing (JSD 0.42, Fig. 2B, right) was observed between samples from large intestine (descending colon) and small intestine (duodenum). A similar analysis was also performed with the Morisita index as an alternative divergence metric, which correlated strongly with the JSD (Fig. S3A). To contextualize the relationships between anatomic sites, we clustered normal human tissues by public gene-expression profiles (GTEx)(48) via dendrograms constructed from JSD distances, which confirmed a similar pattern of clustering based on known anatomic relationships (Fig. 2C).

Across all GVHD patients, we identified greater repertoire overlap across anatomically similar tissues, as revealed by dendrograms in which tissues with more similar TCR repertoires are clustered together (Fig. 2D, Fig. S3B–C). In nearly all cases, different segments of intestine clustered together at different levels of organization, for example segments of the large intestine (ascending, transverse, descending colon and rectum) clustered together, and segments of the small intestine (duodenum, jejunum, ileum) clustered together; spleen and PBMCs did not tend to cluster with GI tissues. This was also the case when we examined these higher-level tissue groups across patients. Repertoires sampled within broad anatomic regions were significantly more similar than those sampled between anatomic regions, specifically the large intestine (p < 0.0001, p < 0.01) and small intestine (p < 0.01, p < 0.001) for GVHD and non-GVHD subjects, respectively Fig. S3E), with a random labeling analysis supporting that this was not due to chance (Fig. S3F). Notably, some patients had much less repertoire overlap across all sampled tissues, for example patient C (Fig. S3B).

The pre-transplant circulating repertoires (PBMCs of recipient origin) were the farthest from the autopsy tissues: in the five patients with available pre-transplant PBMCs, the mean JSD value for pre-transplant blood was 0.998, n = 36 sample pairs. This is consistent with the achievement of complete donor chimerism, as assessed by clinical assays of whole blood or marrow, except for patient G who died with relapsed myeloid leukemia (Table S1). The peripheral T cell compartment was also 100% donor in the three GVHD patients in whom subset chimerism assays were performed. Taken together, these data suggest an anatomic specificity to the clones present in tissues, and this unsupervised analysis using T cell repertoire alone was able to cluster tissues in a pattern mirroring human anatomy.

To further investigate this finding, we asked whether this pattern of repertoire overlap in anatomically related tissues was specific to GVHD or a more general phenomenon reflective of tissue identity. We therefore profiled TCR repertoires of GI-tract tissues from three comparator rapid-autopsy participants (patients H-J, Figs. S3B–D). We again observed the greatest repertoire homology across anatomically similar structures in the GI tract. This parallel between GVHD patients and cancer patients who did not undergo transplantation is consistent with the notion that human tissue T cell repertoires may in part reflect a local phenomenon driven by tissue-specific properties.

Tracking individual T cell clones across tissues to investigate tissue-specificity

Having observed that the pattern of global repertoire similarity between tissues (as measured by the JSD metric calculated for all clones) appeared to be driven in part by the identity of the underlying tissue, we asked if highly expanded or "top-frequency" clones are shared across tissues. We subsampled the dataset to the 10 most abundant clones in each tissue per patient to track individual clones across multiple sites (Fig. 3A) (e.g., for patient F, subsampling yielded 61 clones from 15 unique tissue samples). Tracking clone frequencies across samples revealed, for example, that the top clone in the ascending colon of Patient F was also the top clone in the ileum and was present in several other tissues, including in the skin, but it was not the most abundant clone in all structures (Fig. 3B). This pattern, in which some top clones were shared across tissues, was recapitulated in the other GVHD patients as well as in the non-GVHD comparators, especially in GI-tract samples, as illustrated for top 10 clones in small intestinal samples (Figs. 3C–D, Figs. S4A–H, and Fig. S5).

This visualization highlights that we did not identify the exact same top clones across all tissues or even GVHD-target tissues, nor were top clones unique to a given tissue. This suggests that some T cells are likely to be more tissue-specific, while others have a broader distribution. Notably, consistent with the above analysis of the global repertoire in these samples (Fig. 2), this analysis of the most abundant clones revealed comparable trends in both the GVHD and comparator patients. For the transplant patients, we were able to track T cell clones from the allograft itself (patients B, C, D, F, and G) or donor blood (patient E), revealing highly abundant donor clones in the graft persisting in the recipient tissues, specifically in recipients of unmodified grafts (i.e., not CD34-selected grafts) Figs. 3E–F, Figs S4L–0, Fig. S5).

We next sought to examine how the most abundant clones in circulation may reflect the top clones in the tissue. Focusing specifically on the top circulating clones in the four

patients with available post-mortem blood samples (Fig. 3G, Figs. S4I–K, Fig. S3E), we found that indeed some—but not all—dominant clones in the tissues were detected in the circulation (proportions of shared clones are tabulated in Table S8). That peripheral blood samples do not fully represent tissue-resident T cells underscores potentially important differences between T cell populations that are in circulation compared to those in GVHD-affected tissues. Though blood samples were not available from the comparator patients, we compared the degree of repertoire overlap between blood and GI tissues in publicly available datasets in which we observed a similar lack of repertoire overlap between blood and tissue in the non-inflammatory setting (Fig. S4P).

Evidence for common TCR motifs across patients within tissues

The detection of identical or public TCRs across individuals, particularly with common viral specificity, is well-established (40, 49). We therefore interrogated for TCRs shared across GVHD patients, particularly given the extensive sharing of HLA haplotypes in this cohort (for example, 6 of the GVHD patients carried the HLA allele A 02:01 and 4 carried allele A 01:01) (Table S9). Few clones were shared across GVHD patients (Fig. 4A, Table S10), irrespective of the degree of HLA allele sharing (Pearson's R= 0.11, p= 0.6)

Given the TCR clonal diversity across individuals, we asked whether there might be common features of TCRs in similar tissues across patients, perhaps reflecting similar reactivity. To evaluate this, we applied a published algorithm that identifies amino-acid patterns in different TCRs that are likely to share antigen specificity, thereby grouping lymphocyte interactions with paratope hotspots (GLIPH2) (50–52); this approach revealed motif-groups shared across patients (most shared motifs included in Table S11, including 1 group observed in the small intestine of all 10 patients).

To assess the extent of sharing of GLIPH groups across tissues and patients, we quantified overlap across all pairs of anatomic sites via the Jaccard index (Fig. 4B), a normalized measure of overlap that considers presence vs. absence of features, in contrast to the JSD which also considers frequency. As expected, the highest overlap indices—quantifying the likelihood that the same GLIPH groups are observed in different samples—were observed within each patient (Fig. S6A, dark-blue areas). However, we also identified moderate levels of sharing *across* patients within the GI tract (Fig. 4B zones of darker blue squares), particularly among pairs of small intestine samples between different patients (Fig. 4B red arrow), and among pairs of samples from small and large intestines (Fig. 4B, orange arrow) between patients. Strikingly, the overlap measured via the Jaccard index suggested broadly comparable degrees of sharing in both GVHD and non-GVHD patients (Fig. 4B; the rectangular zones of dark blue involve samples from GVHD patients A-G as well as from the non-GVHD comparator patients H, I, and J).

Since we recovered more clones per sample from the non-GVHD comparator patients than from GVHD patients (Fig. 1C), we assessed the role of sample size (number of unique clones) in determining GLIPH-group overlap (49). We found that the Jaccard index scales linearly with the geometric mean of the size of sample pairs being compared, regardless of whether the pairs were from GVHD patients or from comparator patients (Fig. S6B). To quantify GLIPH-group overlap across samples with considerably different sampling depths,

we defined a standardized tissue-overlap score (see Methods, schematized in Fig. 4C). In brief, after excluding intra-patient sample pairs, a set of raw Jaccard similarity values are plotted against their sample sizes and lines are fitted in log scale for any two sample subsets of interest. The Jaccard values expected at a standard mean repertoire size of 1000 clones are extrapolated as the standardized tissue-overlap score.

We applied the overlap score to estimate the degree of GLIPH-group overlap between tissues across patients. In unsupervised hierarchical clustering, GI sites clustered most closely together across patients (Fig. 4D, dark brown regions). Separately, liver and spleen also clustered together across patients, and skin was furthest away from the other tissues. As such, the landscape of GLIPH-group overlap across patients recapitulated the patterns we observed in individual patients analyzed via the underlying TCR nucleotide sequences (Fig. 2D). Taken together, these findings support the presence of common TCR motifs in specific tissues across patients, including both GVHD patients and non-GVHD comparators.

Infiltrating T cells in tissues have a tissue-resident phenotype and are primarily of donor-origin

Having identified the tissue-specificity of T cell clones present in tissues, we sought to characterize the phenotype of those T cells present in the tissues subjected to TCR sequencing. As freshly dissociated samples were not available, we characterized the phenotype of T cells in samples adjacent to those from which DNA was obtained using seven-color fluorescence microscopy (Figs. 5A–C). In light of recent reports describing a high abundance of tissue-resident T cells in human skin GVHD tissues (34) and gut GVHD in non-human primates (35), we hypothesized that an abundance of T cells in our tissue samples would bear a phenotype characteristic of tissue residence, which we evaluated via expression of CD69 (1). We also evaluated the proliferation marker Ki67, which is highly expressed in T cells in GVHD tissues (53) and FoxP3 as a marker of T regulatory cells. For this analysis, we focused on six tissues from patient D. Our analyses revealed that a mean of 51% and 42% of quantified CD8-negative (presumed CD4$^+$) and CD8$^+$ T cell in GI tract tissues, respectively, expressed CD69 (Fig. 5D), with a smaller fraction of CD69-positive cells identified in the skin. In this patient, < 15% of stromal T cells expressed Ki-67 or were FoxP3 positive (Fig. 5D, Fig. S7A). We specifically assessed how distance from parenchymal structures (crypt/villi, as defined by panCK positivity) was related to CD69 expression and found a trend toward higher frequencies of CD69$^+$ tissue resident T cells closer to the parenchymal cells in the GI tract tissues (Figs. S7B–C). Based on these histologic analyses of immediately adjacent tissue segments, we infer that most of the sequenced TCRs came from tissue-resident T cells as opposed to the alternative possibility that at the time of fixation these tissues contained primarily circulating T cells that happened to be traversing in them. This is consistent with the divergence in repertoires we observed when comparing post-mortem blood and tissues (Figs. 2 and 3).

Finally, in light of recent reports of long-term persistence of recipient T cells post-transplant in humans (33, 34, 54), we assessed tissue T cell chimerism in skin and GI tissues from the two GVHD patients who received sex-mismatched grafts by fluorescence in situ hybridization for loci on X and Y chromosomes. The majority of T cells were of donor

origin by FISH chimerism analysis (>90% of 641 CD3$^+$ T cells for Patient D and >90% of 24 cells in Patient A across images from the gut and skin); however, there were rare host T cells present, as illustrated in Fig. 5E and Fig. S7D. The donor origin of hematopoietic cells in Patient D was further corroborated by transcript abundances of X- and Y-chromosome genes generated by single-nucleus RNA-seq profiles (Fig. 5F, Fig. S11A–D).

TCR sequencing of mice with GVHD reveals greater repertoire overlap within similar anatomic regions

We next asked whether the TCR repertoire patterns we characterized across tissues in GVHD patients would be identified in a mouse model of GVHD. To focus on the site-specificity of the T cell infiltrate, we first examined a major-mismatch (MHC-disparate) model B6 (H2^b) → BALB/c (H2^d) so that the mismatched MHC antigens would be expressed across all recipient tissues. TCR repertoires were profiled in various tissues from GVHD mice at day 7 or 14 post-transplant in 3 separate experiments (132 samples from 16 unique GVHD recipient mice including donor T cell pool and two control animals, Fig. S1A), yielding a median of 15,754 clones per mouse with a SD 9,260 and a median of 3,371 clones per non-blood tissue, SD 3,388 (results including repertoire clonality summarized in Fig. 1F–I, Fig. S1B).

As in the human samples, we found greater repertoire overlap between tissues with greater anatomic similarity, both at day 7 and day 14 post-transplant (Fig. 6A, Fig. S8A, Table S13). We found significantly more similarity in the repertoire of gastrointestinal compared to non-gastrointestinal samples (p=1.7×10^{-27}, adjusted Wilcox), in keeping with the patient samples (Fig. S8B), with even greater overlap within GI tract compartments, specifically the small intestine (Fig. S8C). This finding was similar in a minor-mismatched (MHC-non-disparate) GVHD model (LP/J (H2^b) → B6 (H2^b)) (Fig. S8D). Furthermore, the most expanded clones were not identical across all tissues, including across different segments of the GI tract within a mouse, with some variability from animal to animal (Figs. 6B–C, Figs. S9A–B). Strikingly, the overall pattern of T cell repertoire overlap across tissues was very similar from mouse to mouse, as suggested by the small standard deviations in the JSD comparisons (Table S14).

Clone sharing between samples was only observed to any notable degree across samples collected from the same animal. Even the same tissues from replicate animals transplanted from aliquots of the same donor graft did not exhibit more sharing than disparate tissues from separate transplant experiments (Fig. S8E). This was the case even though the T cell pools in the donor grafts shared a considerable number of clones (Fig. 6D; 22% shared between any two grafts and 7% shared by all three). Critically, the top clones in the spleen and blood did not reflect the pattern of highly abundant TCRs in the tissues (Table S8), nor were the top tissue clones simply the top clones in the donor T cell population (Figs. 6B–C, Fig. S9. Furthermore, between day 7 and 14 post-transplant, we observed an increase in the cumulative frequency of the top 10 clones (p < 0.001 in a linear mixed-effect model) (Fig. S8F), consistent with ongoing clonal expansion within the tissues. To compare repertoires across mice, we again applied a motif-based analytic approach (see Methods) as we did in the human samples, utilizing the normalized overlap score (Fig. 4B). This unsupervised

approach revealed clustering of similar anatomic tissues (duodenum, colon, and ileum), with the dendrograms highlighting in particular the GI tract at day 14 post-transplant (Fig. 6E); tissues appeared more similar to each other at day 7 compared to day 14, consistent with recently published data (55).

DISCUSSION

While many prior studies have characterized circulating TCR repertoires (42–47, 56), our manuscript aligns with recent work on the analysis of T cells in tissues (8). Understanding T cell immunity in tissues is fundamental to elucidating the immune response in humans at the site of pathology (57) in both the endogenous (8), and transplant setting, the focus of this study. Mapping TCR clones across anatomic sites within an individual, we observed that each patient harbored a highly complex repertoire landscape in which different tissues were characterized by distinct clonal populations, disparate amounts of repertoire overlap across sites, and variable numbers of dominant clones shared across many tissues. Though the patient cohort was limited by size and heterogeneous with respect to clinical variables, we observed that the pattern of overlap between TCR repertoires allowed clustering of tissues by anatomic similarity, with the greatest homology observed between different sites within the GI tract. This demonstrates the technical feasibility of DNA-based TCR profiling in tissues via rapid-autopsy, despite the phenomenon of gut autolysis (58) and the classical description of acute GVHD as pauci-immune (59). While sampling remains a limitation of this work, the numbers of TCR sequences we recovered were comparable to a prior analysis of fresh-frozen GVHD biopsies (31), and approximately an order of magnitude higher than those obtained from formalin-fixed paraffin-embedded GVHD samples (32). Importantly, in the GVHD patients from whom we were able to collect blood at time of autopsy, the circulating T cell repertoire did not capture the complexity of TCRs identified in tissues, joining other reports in highlighting the limitations of peripheral blood studies in understanding pathophysiology in tissues. (25, 28, 29, 31, 32, 54, 60). While the majority of T cells were of donor origin as assessed by orthogonal assessments of chimerism, it remains possible that some of the clones were persisting host cells. Furthermore, our study does not address which TCRs in the tissues are indeed alloreactive (61) or otherwise antigen-specific.

Despite extensive HLA-sharing in the GVHD cohort, there was a paucity of public TCRs shared across patients, in keeping with prior analyses of diagnostic GVHD biopsies (31, 32). Moreover, in patients with multiple sites of GVHD, the dominant TCRs were not consistently observed across tissues within the same patient, raising the possibility that GVHD arose in each local tissue independently. Consistently, we observed only a minute sharing of TCR clones across mice with GVHD, even in inbred animals whose allografts were aliquots of the same suspension of splenic donor T cells, indicating that a similar stochastic process occurs in animals, consistent with prior analyses (21, 62). As in the patient samples, the TCR repertoire patterns in GVHD mice also exhibited clustering by anatomic region, particularly within segments of the GI tract, however there were also some dominant clones shared across sites, consistent with published work.(21). This is also notable because the mice did not receive GHVD prophylaxis or treatment. This complex picture in both humans and mice may represent a combination of some TCRs reactive to

tissue-specific antigens, some TCRs proliferating in response to MHC antigens expressed across tissues, and some potentially cross-reactive viral TCRs (63).

Having failed to find common clones across patients, we clustered TCRs into motif groups using the published GLIPH2 algorithm (50–52). As expected, the greatest degree of motif-group sharing occurred in samples from the same patient, but we also found inter-patient overlap, particularly among anatomically similar regions. To the extent that TCRs assigned to the same GLIPH group are reflective of shared reactivity, this suggests that tissue-specific features sculpt their resident T cell repertoires. Moreover, the overall pattern of similar degrees of GLIPH-group sharing between non-transplant and transplant patients suggests a more general process, rather than solely the alloresponse. To account for the confounding effect of TCR sample size or depth on the measurement of motif-group sharing, we developed a tissue-overlap score, an important normalization that is often overlooked in TCR-sequencing studies.

One unanticipated finding was that the similarity in TCR profiles across anatomically similar tissues was observed in both GVHD patients and comparators. This observation is particularly notable considering the provenance of T cells in these two groups: endogenous T cells with decades of residency in the comparators versus recent arrival in the GVHD patients, either as mature T cells within the graft or *de novo* differentiation from hematopoietic stem cells. That the newly reconstituted immune system re-establishes tissue compartmentalization suggests this is a fundamental feature of the T cell repertoire in tissues and aligns with prior studies on the role of tissue-resident T cells in GVHD (35, 64, 65). While antigen presentation to alloreactive donor TCRs initiates GVHD (10), our study provides insights into the T cells that persist beyond diagnosis and through months of lymphocyte-directed therapy, which may be distinct in nature from those that trigger pathology acutely (60).

MATERIALS AND METHODS

Study design

Human specimens were studied as part of a prospective observational study whose objective was to characterize the T cell repertoire in autopsy tissues from allo-HCT patients with a history of GVHD. GVHD was also modeled in mice. The study design did not include blinding or randomization.

Patients and GVHD assessments

Clinical features of the seven GVHD patients and three non-GVHD comparators are tabulated in Table 1 and Table S1. The GVHD patients underwent allo-HCT at Memorial Sloan Kettering Cancer Center between 2014 and 2018. Structured retrospective assessments for GVHD through HCT day 100, 180 and 365 were conducted by an institutional consensus committee (Table S2) (66). In the cases where death occurred after day 365, GVHD symptoms in the 8 weeks prior to death were assessed following the same structured format. In cases where structured reviews did not assess GVHD outcomes for reasons of censoring

for competing risks (relapse, graft failure, subsequent transplant, stem cell boost, donor lymphocyte infusion, or death) GVHD symptoms were assessed for purposes of this study.

Human tissues

Between 2015 and 2018, seven allo-HCT recipients with GVHD were prospectively enrolled in the Last Wish Program (MSKCC IRB protocol #15–021) to donate tissues and other samples after their death (36). Consent was obtained antemortem. Antemortem PBMCs were viably preserved under biospecimen collection protocols (MSKCC IRB #06–107 or #09–141).

Candidates were identified through screening approaches that included referrals from primary physicians and computerized alerts to the study team when patients with a history of GVHD were admitted to the intensive care unit or when do-not-resuscitate orders were issued. Candidates were excluded only if they had infectious diseases that could pose a risk to autopsy personnel (e.g. hepatitis or HIV with detectable viral loads or concern for viral encephalitis). Research autopsies were deemed feasible if death was anticipated in the hospital or at a location within a 3-hr drive of the hospital. Comparator patients were selected retrospectively from a repository of other cancer patients who donated tissues to the Last Wish Program. Among 50 participants who had not received HCT and who had at least one available sample from the GI tract, we selected as comparators three patients who had 5 available GI tract samples and who were without known GI disease at time of death.

Cadavers were refrigerated as soon as possible after death and autopsies were performed promptly. During autopsy, tissue specimens were bisected and one portion immediately flash-frozen in liquid nitrogen and stored at −80°C while the remainder was formalin-fixed. For GI tissues, the outer muscle layers were removed prior to preservation. Post-mortem blood was aspirated from a cardiac chamber or from the inferior vena cava. Marrow was aspirated, when possible, from the anterior iliac spine or the posterior iliac crest with a Jamshidi needle. Mononuclear cells from postmortem blood and marrow aspirates were viably preserved either immediately or following refrigerated storage until the next business day.

As acute GVHD is classically a pauci-immune phenomenon without dense infiltrates of lymphocytes at the sites of disease (59), when multiple samples from the same tissue were available from the same patient, hematoxylin/eosin-stained sections of formalin-fixed tissues were reviewed by a pathologist (A.H. or C.I.D.) to prioritize for sequencing the samples with histologically apparent lymphocytes.

DNA extraction

Genomic DNA was extracted from frozen tissues with either the Mag-Bind Blood & Tissue DNA HDQ 96 Kit (Omega Bio-tek catalog # M6399) on an automated nucleic acid purification instrument (PerkinElmer Chemagic 360/MSM I) according to the manufacturer's protocol with up to 10 mg tissue input and elution into 150 μL elution buffer, or alternatively using the DNeasy Blood & Tissue Kit (QIAGEN catalog # 69504) according to the manufacturer's protocol with incubation at 55°C for digestion and elution in 0.5X Buffer AE. Isolated nucleic acids were further purified with (Genomic DNA Clean

& Concentrator-25, Zymo Research, catalog #D4065) due to concerns for red blood cell contamination. For viably preserved blood and bone marrow, pending cell number, CD3$^+$ T cells were flow-sorted using anti-human CD3 PE-Cy7 (clone UCHT1, eBioscience catalog # 25–0038-42) on a BD FACS Aria sorter, followed by extraction of DNA from CD3 T cells using the DNeasy Blood and Tissue kit; for smaller samples, DNA was extracted directly using DNeasy Blood and Tissue kit to minimize cell loss. For the donor samples, excess gDNA was utilized from clinical assays.

T cell receptor sequencing and analysis

TCRβ CDR3 sequencing was performed by Adaptive Biotechniques (Seattle, WA) to "survey" depth with reporting of CDR3 nucleotide and amino acid sequences (along with V- and J- allele/gene/family) and their associated sequencing data, including productive frequency, templates count, and the total amount of genomic DNA sequenced (ng) per sample. Clone size distributions were inspected for evidence of clonal skewing in the autopsy samples and none was found. One skin sample from Subject C was re-sequenced for technical reasons. Three samples from Subject D were excluded from some analyses: a marrow sample with only 10 clones recovered; a kidney and a heart sample were excluded because no other patients had samples from those organs. An estimation of the extent of sampling is included in Fig. S12A (67). The proportion of clones with TRBβ chains classically paired with MAITs and NKT cells (68) is included in Figs. S12B–C; however, we did not have paired TCRα chains for definitive evaluation of these populations. Clones were evaluated in multiple published databases of pathogen associated TCRs (39–41, 47, 69) (Table S5, Fig. S12D).

Repertoire diversity and overlap measures: T cell repertoire diversity of each sample was calculated at the nucleotide level, considering only productive clones, as previously described using clonality and slope (41), The Jensen Shannon divergence (JSD) was used to quantify repertoire overlap across tissues as previously described (41); Morisita was calculated using the horn_morisita function using the R package *abdiv*.

Dendrograms: The relationships in TCR repertoires between samples were summarized by calculating the pairwise Jensen-Shannon Divergences within each patient's samples. These were clustered using single-linkage hierarchical clustering and visualized as dendrograms in which the scale bars correspond to nodes as the heights of clustering hierarchy using R packages *ggtree*, *dendextend*, and *ape*.

GTEx data (version 8) were retrieved from https://www.gtexportal.org/home/datasets/on October 12, 2021. We filtered the expression data according to Ferreira et al. (70) as follows: protein-coding genes were identified from the Ensembl 104 by filtering the annotations to retain "protein_coding" gene_biotypes with "gene" types, according to Homo_sapiens.GRCh38.104.chr.gtf. Mitochondrial, hemoglobin ("HBB"), and MTATP6P1 (71) transcripts were removed. Genes with low expression (<2 transcripts per million (TPM)) were removed, and genes with high expression were capped at 10K TMP. Gene fold changes < 2 (relative to the sample with the lowest expression, per-gene) and an absolute difference of less than 10 TPM were removed. The same filtering was applied to

the median-aggregated and the per-individual expression data. JSD was used to quantify the between-tissue similarity and was plotted as dendrograms as described above. See supplementary code repository for a full list of all packages and versions used.

Alluvial plots & Heatmaps: To generate the heatmaps and alluvial plots, we generated a subsampled dataset for each patient with the top 10 clones by productive frequency per tissue (*top_n* function in R package *dplyr*). Fig. S12E presents clone frequency distribution for human and mouse samples, highlighting the cumulative abundance of the top 10 clones. For the heatmaps, each clone was ranked by frequency and plots were generated using the R package *pheatmap* without clustering such that the rows were organized by the mean frequency rank per row, excluding pre-transplant blood samples. The alluvial plots were generated using the R package *ggalluvial*. To facilitate plotting of stacked bars for clones with identical frequencies (tied for rank), random pseudocounts of negligible size ($\sim 10^{-7}$) were added to each frequency. Additionally, in cases where the top 10 clones involved ties, all tied clones were included.

GLIPH Motif Grouping: The GLIPH2 (51) algorithm was applied to a dataset comprised of all patient samples simultaneously (TCRs defined by the Vβ CDR3 combined with the V and J gene), with the following GLIPH2 parameters: simulation_depth=1000, kmer_min_depth=3, cdr3_length_cutoff=8, refer_file=refererence_files/ref_CD48.txt. HLA data were not included. This approach identified 135,479 GLIPH groups, including groups defined by either local ('motif') similarities and full-CD3-length 'global' similarities; filtering to include groups with > two unique CDR3s and yielded 49,979 groups, which were those included in subsequent analyses. For GLIPH analysis, tissue labels were simplified by grouping: blood and marrow were categorized as blood; the rectum and the ascending, descending, transverse, right, and left portions of the colon were categorized as large intestine; ileum, jejunum, and duodenum were categorized as small intestine. Jaccard distances were calculated for overlap between pairs of patient/tissue combinations as 1 minus the ratio of the number of groups that were observed in both combinations to the overall number of groups in both.

Tissue-overlap score for motif-based analysis: We noted that the Jaccard distance was correlated with sample size, specifically that the log of the distance between pairs of samples correlated linearly with the log of the geometric mean of the sample sizes in the pair (Fig. S6B). Moreover, the slopes of lines fitted to any set of pairs from different patients were comparable (Fig. S6B), indicating that Jaccard-based estimates of overlap scale universally with sample size. Of note, one patient (patient D) had a distinct distribution of GLIPH overlap (Fig. S6C). To compare the extent of GLIPH-group overlap between pairs of samples in a manner that incorporates varying sample size, we defined a standardized tissue-overlap score.

First, a linear fit (in log space) was constructed for all pairs of samples to derive a universal slope value (excluding intra-patient pairs and also excluding samples from the outlying patient D), such that the intercept estimates the average amount of sharing among all samples in the dataset. The underlying assumption is that the scaling of the Jaccard index with sample size is universal (the slope of the fit) but higher overlap for a group of samples

would be reflected by a higher overall Jaccard index across all sample sizes (the intercept of the fit). Next, sets of distances from sample groupings of interest (in this case, tissue of origin) were fitted anew with the universal slope value, and the resulting intercepts provide measures of the Jaccard-based overlap for the different groups. The intercept of the fit with log geometric mean 3 (1,000 TCRs) was defined as standard sharing and was used to compare groups of samples (schematized in Fig. 4C). Slopes for comparisons of pairs that include small-intestine samples are in Fig. S6D. This standard sharing was calculated for all possible pairs of tissue types and clustered using single-linkage hierarchical clustering (Fig. 4D). Code for calculating the tissue-overlap score is available at https://github.com/vdblab/Dewolf-TCR-landscape-2022.

GLIPH Motif Grouping for mouse samples: the GLIPH2 (50) motif analysis was performed on mouse samples with the same parameters as described above, except for the use of the mouse-specific reference file ref_CD48_m.txt. The analysis identified 198,472 unique GLIPH groups (102,825 after GLIPH groups were filtered to include only those that contained three or more unique CDR3s). Since fewer unique tissues were included in this analysis, we did not group the tissues as in the analysis of human samples. In addition, we visualized similarity between mouse samples in terms of presence or absence of TCRs representing GLIPH groups using t-Distributed Stochastic Neighbor Embedding (tSNE). Each sample was represented as a vector of values for each GLIPH group, where a value of 1 signifies that a GLIPH group was present in TCRs of a given sample and 0 signifies that it was absent. t-SNE plots in Fig. S10 thus represent in 2 dimensions how similar TCR repertoires are in different mouse samples in terms of GLIPH group usage and demonstrate that mouse samples cluster most closely to those from the same mouse, rather than by tissue region or transplant group; two of the three donor samples also clustered together.

Mice

Mice: Female B57BL6/J, BALB/cJ, and LP/J mouse strains were obtained from The Jackson Laboratory at 6–8 weeks of age. Mice were cohoused (3–5 mice/cage, randomly segregated prior to transplantation), provided food and water *ad libitum*, and maintained on a 12:12hr light-dark cycle. Experimental protocols were approved by the Institutional Animal Care & Use Committee at Memorial Sloan Kettering Cancer Center.

Transplantation: Mice were transplanted as previously described (72). Briefly, BALB/cJ recipients were conditioned with 950cGy split-dose total body irradiation from a cesium source (1100cGy for B57BL6/J recipients). C57BL6/J donor grafts, consisting of 10^7 T-cell-depleted bone marrow cells and 10^6 splenic $CD5^+$ T cells (4×10^6 for LP/J donors) purified using CD5 (Ly-1) MicroBeads with Miltenyi columns were administered to recipient mice via tail vein injection. When grafts were prepared from multiple donors, cells were pooled into a single tube from which aliquots were administered to individual recipients. Following transplant, mice were monitored for health twice daily and scored weekly for GVHD severity in accordance with protocols approved by the Institutional Animal Care and Use Committee.

Harvest: On day 7 or 14 post-transplant (43–44 days for the minor-MHC mismatch model), peripheral blood was collected from the retro-orbital plexus. Tissues were harvested using separate autoclaved tools for each animal and disposable forceps were used to handle each individual tissue. The liver was flushed of blood with buffered saline via the portal vein prior to harvest. Skin samples were taken from the ear most commonly, or when apparent, from sites of skin involvement by GVHD on the torso. Tissues were minced with razor blades in buffered saline prior to genomic DNA isolation via the QIAGEN DNeasy Blood & Tissue Kit.

Multiplex tissue staining, imaging, and analysis

Seven-color multiplex staining: Multiplex staining was performed after primary antibody concentration and stripping conditions were optimized by a pathologist (T.H.) according to the results of diaminobenzidine detection (Leica Bond Polymer Refine Detection DS9800) with Leica Bond RX stainer. Sections of formalin-fixed, paraffin-embedded tissues 4 μm thick were baked for 3 hours at 62°C upright with subsequent deparaffinization performed on Leica Bond RX, followed by 30 minutes of antigen retrieval with Leica Bond ER2, followed by 6 sequential cycles of staining, each round of which included a 40-minute combined blocking (Akoya antibody diluent/block ARD1001) and primary antibody incubation (Table S15). The primary antibody detection was performed using HRP conjugated species-specific secondary antibody polymer (Table S15) with a 10-minute incubation. The HRP-conjugated secondary antibody polymer was detected using fluorescent tyramide signal amplification using Opal dyes 520, 540, 570, 620, 650 and 690 accordingly (Table S15). After each staining cycle, a heat-induced stripping of the primary/secondary antibody complex was performed using Akoya AR9 buffer (AR900250ML) and Leica Bond ER2 (90% ER2 and 10% AR9) at 100°C for 20 minutes preceding the next cycle. After 6 sequential rounds of staining, sections were stained with Hoechst (Invitrogen 33342) to visualize nuclei and mounted with ProLong Gold antifade reagent mounting medium (Invitrogen P36930).

Multispectral imaging, spectral unmixing and cell segmentation: Seven-color multiplex-stained slides were imaged using the Vectra Multispectral Imaging System version 3 (Akoya). Scanning was performed at 20X (200X final magnification). Filter cubes used for multispectral imaging were DAPI, FITC, Cy3, Texas Red and Cy5. A spectral library containing the emitted spectral peaks of the fluorophores in this study was generated using the Vectra image analysis software (Akoya). Using multispectral images from single-stained slides for each marker, the spectral library was used to separate each multispectral cube into individual components (spectral unmixing) allowing for identification of the seven marker channels of interest using Inform 2.4 image analysis software. Images were exported to the Indica Labs Halo image analysis platform, followed by cell segmentation and signal thresholding that were performed separately on each image.

Cell and spatial analysis of Halo data: The output from the Halo segmentation and thresholding application was processed with a series of custom R scripts (available at https://github.com/vdblab/Dewolf-TCR-landscape-2022). The data were first loaded and parsed

into R data objects with cell phenotypes being assigned using a set of rules that mapped positive markers to cell phenotypes. This was done using a series of regular expressions.

Once phenotypes were assigned to all cells, cell fractions were computed for each type with specific functional markers (CD69, Ki67). These fractions were computed both globally and spatially. The second phase of the analysis was spatial. The distance from each cell to the closest panCK+ cell was measured as an approximation of each cell's distance to the nearest epithelial interface, that is villus/crypt structures as the parenchymal structures and lamina propria as the stromal structures. The cells were then binned into distance intervals from 0–25, 25–50, 50–100, 100–200, 200–300 microns. To visualize the spatial distribution of the various cell states, the cells were plotteds patially on a two-dimensional grid with colors indicating state and distance to panCK+ cells and shape indicating CD69 positivity.

Combined Immunophenotyping and FISH analysis: After immunophenotyping with mouse anti-CD3 antibody, the relevant tissue slides with coverslip were briefly reviewed and recorded under fluorescence microscopy for the intensity and quality of CD3 staining. After removing the coverslip, the slides were washed in 2X saline sodium citrate solution (SSC) at room temperature for 5 minutes, fixed with 10% neutral buffered formalin for 20 minutes, and dehydrated in a series of 70%, 85% and 100% alcohol for five minutes each. Slides were then pretreated for 10 minutes with 20 mM Citrate Buffer/1% NP-40 mixture, pH 6.0–6.5, followed by protease treatment (Abbott Molecular, Des Plaines, IL) at 40º C for 10–15 minutes and dehydration in a series of 70%, 85% and 100% of alcohol for 2 minutes each. FISH probes consisted of probes specific to the X chromosome centromere (labelled with spectrum orange) and to the Y chromosome specific probe (labelled with spectrum green), both from Abbott Molecular. After applying the FISH probes to the tissue areas, both tissue and probes were co-denatured at 94º C for 7 minutes, incubated at 37º C overnight, followed by post-hybridization washing in 2xSSC/0.3% NP-40 at 77º C for one minute. Tissue sections were counterstained with antifade medium without DAPI (Vector Laboratories, Burlingame, CA). The slides were evaluated under a fluorescence microscope coupled with appropriate filters for CD3 immunophenotype and for the X and Y chromosome probes. Signal analysis was performed in combination with tissue structure and cell morphology correlation, focusing on regions replete with CD3+ T cells. The sex chromosome complements, i.e., male (XY) or female (XX) were determined and recorded for all CD3 positive lymphocytes observed.

Single nucleus RNA sequencing

Sample preparation and sequencing: Nuclei extraction from snap frozen tissues was performed using Singulator™100 (S2 Genomics) system. Tissues between 20–50mg were processed using nuclei isolation cartridges following Singulator™100 standard nuclei isolation protocol. To reduce RNA degradation, DTT (Sigma-Aldrich Inc, cat no. 43816) and Protector RNAse inhibitor (Roche Diagnostics Deutschland GmbH, cat no. 0335399001) were added in the dissociation chamber (final concentration in 3.5mL final nuclei suspension: 1mM DTT and 0.5 U/uL RNAse inhibitor, https://www.protocols.io/view/nuclei-extraction-for-single-cell-rnaseq-from-froz-q26g74xzqgwz/v1). Nuclei were FACS sorted after staining with 7-aminoactinomycin D fluorescent intercalator (Life Technologies

Corporation, cat. no A1310). Nuclei yield and quality were evaluated on Countess II automatic cell counter by trypan blue and DAPI (Life Technologies, cat. no D3571) staining. Single Cell 5' Gene Expression and immune profiling was performed with a 10X genomics Chromium system using Chromium Next GEM Single Cell 5' GEM Kit v2 (PN-1000244) according to them Chromium Next GEM Single Cell 5' GEM Kit v2 (Dual Index) user guide. Briefly, 16,500 nuclei were loaded into the Chromium Next GEM CHIp K, targeting recovery of 10,000 nuclei after encapsulation and barcoding. Next generation sequencing libraries were constructed following the user guide and were sequenced on an Illumina NovaSeq 6000 system with the following read lengths: 1:28 cycle i7 Index: 10 cycles i5 Index: 10 cycles Read 2: 88 cycles.

RNA sequencing analysis: The Cell Ranger generated files (filtered_feature_bc_matrix.h5) for each patient were processed following the guidelines on the shunPykeR GitHub repository (73), an assembled pipeline of publicly available single cell analysis packages that allows for data analysis in a reproducible manner. Genes that were not expressed in any cell were removed from downstream analysis; cells were normalized to a total library size of 10,000 reads and gene counts were *log*-transformed (*log* = natural logarithm). Principal component analysis (PCA) was applied to reduce noise prior to data clustering. An optimal number of principal components was selected using the knee point (eigenvalues smaller than radius of curvature) of variance. Finally, Leiden clustering was used to identify clusters after PCA reduction. Quality control of the single cells using the total counts and number of genes did not identify any clusters requiring removal. After removal of doublets by Scrublet (74), PCA and unsupervised clustering analysis was reapplied to the filtered data, followed by batch effect correction across all samples using harmony (75). For the sex-linked transcript analysis (XIST and RPS4Y1 genes) in Fig. S11D, goblet cells were not shown due to low expression of both genes in the three patients; only one endothelial cell was detected for patient I and was not included.

Statistical analyses

To evaluate relationships between measures of TCR diversity (clonality, slope), clinical or mouse-experimental variables were constructed, with fixed-effects terms for tissue and a random-effect term for individuals using the *lmerTest::lmer* function in R. All p-values were calculated as two-sided. To assess how JSD values of similar tissues compare to those of less related tissues based on pooled JSD values across patients, p-values were calculated using the Kruskal-Wallis test, Benjamini Hochberg corrected for multiple comparisons.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

Author Manuscript

Competing interests:

IP has received research funding from Merck and serves as a DSMB member for ExcellThera. DMP has served as advisory board member for Evive Biotechnology (Shanghai) Ltd (formerly Generon [Shanghai] Corporation Ltd), Kadmon Corporation, CareDx, and Ceramedix. AMH holds intellectual property related to Interleukin-22 and GVHD. SAG receives research funding from Miltenyi Biotec, Takeda Pharmaceutical Co., Celgene Corp., Amgen Inc., Sanofi, Johnson and Johnson, Inc., Actinium Pharmaceuticals, Inc., and is on the Advisory Boards for: Kite Pharmaceuticals, Inc., Celgene Corp., Sanofi, Novartis, Johnson and Johnson, Inc., Amgen Inc., Takeda Pharmaceutical Co., Jazz Pharmaceuticals, Inc., Actinium Pharmaceuticals, Inc. RRR has served as advisory board member for MaaT Pharma, LisCure, Seres, Kaleido, and Protacta, is a consultant for Merck, Microbiome DX, and Karius, and reports a patent license fee and stock options for Seres. MAP reports honoraria from Abbvie, Astellas, Bristol-Myers Squibb, Celgene, Equilium, Incyte, Karyopharm, Kite/Gilead, Merck, Miltenyi Biotec, MorphoSys, Novartis, Nektar Therapeutics, Omeros, OrcaBio, Takeda, and VectivBio AG, Vor Biopharma. He serves on DSMBs for Cidara Therapeutics, Medigene, Sellas Life Sciences, and Servier, and the scientific advisory board of NexImmune. He has ownership interests in NexImmune and Omeros. He has received research support for clinical trials from Incyte, Kite/Gilead, Miltenyi Biotec, and Novartis. He serves in a volunteer capacity as a member of the Board of Directors of the American Society for Transplantation and Cellular Therapy (ASTCT) and Be The Match (National Marrow Donor Program, NMDP), as well as on the CIBMTR Cellular Immunotherapy Data Resource (CIDR) Executive Committee. BG has received honoraria for speaking engagements from Merck, Bristol Meyers Squibb, and Chugai Pharmaceuticals; has received research funding from Bristol Meyers Squibb and Merck; and has been a compensated consultant for Darwin Health, Merck, PMV Pharma and Rome Therapeutics of which he is a co-founder. CAID reports research support from BMS. TJH reports research support from BMS and Calico Labs. MVDB has received research support and stock options from Seres Therapeutics and stock options from Notch Therapeutics and Pluto Therapeutics; he has received royalties from Wolters Kluwer; has consulted, received honorarium from or participated in advisory boards for Seres Therapeutics, Vor Biopharma, Rheos Medicines, Frazier Healthcare Partners, Nektar Therapeutics, Notch Therapeutics, Ceramedix, Lygenesis, Pluto Therapeutics, GlaskoSmithKline, Da Volterra, Thymofox, Garuda, Novartis (Spouse), Synthekine (Spouse), Beigene (Spouse), Kite (Spouse); he has IP Licensing with Seres Therapeutics and Juno Therapeutics; and holds a fiduciary role on the Foundation Board of DKMS (a nonprofit organization). JUP reports research funding, intellectual property fees, and travel reimbursement from Seres Therapeutics, and consulting fees from DaVolterra, CSL Behring, and from MaaT Pharma. He serves on an Advisory board of and holds equity in Postbiotics Plus Research. He has filed intellectual property applications related to the microbiome (reference numbers #62/843,849, #62/977,908, and #15/756,845). Memorial Sloan Kettering Cancer Center (MSK) has financial interests relative to Seres Therapeutics.

Data and materials availability:

The datasets generated during in this study are available in the open-access ImmuneAccess repository https://doi.org/10.21417/SD2022GVHD. Analysis code is available at https://github.com/vdblab/Dewolf-TCR-landscape-2022.

References and Notes

1. Farber DL, Yudanin NA, Restifo NP, Human memory T cells: generation, compartmentalization and homeostasis. Nat Rev Immunol 14, 24–35 (2014). [PubMed: 24336101]

2. Wong MT, Ong DE, Lim FS, Teng KW, McGovern N, Narayanan S, Ho WQ, Cerny D, Tan HK, Anicete R, Tan BK, Lim TK, Chan CY, Cheow PC, Lee SY, Takano A, Tan EH, Tam JK,

Tan EY, Chan JK, Fink K, Bertoletti A, Ginhoux F, Curotto de Lafaille MA, Newell EW, A High-Dimensional Atlas of Human T Cell Diversity Reveals Tissue-Specific Trafficking and Cytokine Signatures. Immunity 45, 442–456 (2016). [PubMed: 27521270]

3. Rosato PC, Wijeyesinghe S, Stolley JM, Masopust D, Integrating resident memory into T cell differentiation models. Curr Opin Immunol 63, 35–42 (2020). [PubMed: 32018169]

4. Dominguez Conde C, Xu C, Jarvis LB, Rainbow DB, Wells SB, Gomes T, Howlett SK, Suchanek O, Polanski K, King HW, Mamanova L, Huang N, Szabo PA, Richardson L, Bolt L, Fasouli ES, Mahbubani KT, Prete M, Tuck L, Richoz N, Tuong ZK, Campos L, Mousa HS, Needham EJ, Pritchard S, Li T, Elmentaite R, Park J, Rahmani E, Chen D, Menon DK, Bayraktar OA, James LK, Meyer KB, Yosef N, Clatworthy MR, Sims PA, Farber DL, Saeb-Parsy K, Jones JL, Teichmann SA, Cross-tissue immune cell analysis reveals tissue-specific features in humans. Science 376, eabl5197 (2022).

5. Masopust D, Soerens AG, Tissue-Resident T Cells and Other Resident Leukocytes. Annu Rev Immunol 37, 521–546 (2019). [PubMed: 30726153]

6. Clark RA, Chong B, Mirchandani N, Brinster NK, Yamanaka K, Dowgiert RK, Kupper TS, The vast majority of CLA+ T cells are resident in normal skin. J Immunol 176, 4431–4439 (2006). [PubMed: 16547281]

7. Gray JI, Farber DL, Tissue-Resident Immune Cells in Humans. Annu Rev Immunol, (2022).

8. Poon MML, Caron DP, Wang Z, Wells SB, Chen D, Meng W, Szabo PA, Lam N, Kubota M, Matsumoto R, Rahman A, Luning Prak ET, Shen Y, Sims PA, Farber DL, Tissue adaptation and clonal segregation of human memory T cells in barrier sites. Nat Immunol 24, 309–319 (2023). [PubMed: 36658238]

9. O'Reilly RJ, Allogenic bone marrow transplantation: current status and future directions. Blood 62, 941–964 (1983). [PubMed: 6354307]

10. Koyama M, Hill GR, The primacy of gastrointestinal tract antigen-presenting cells in lethal graft-versus-host disease. Blood 134, 2139–2148 (2019). [PubMed: 31697827]

11. Beilhack A, Schulz S, Baker J, Beilhack GF, Wieland CB, Herman EI, Baker EM, Cao YA, Contag CH, Negrin RS, In vivo analyses of early events in acute graft-versus-host disease reveal sequential infiltration of T-cell subsets. Blood 106, 1113–1122 (2005). [PubMed: 15855275]

12. DeWolf S, Sykes M, Alloimmune T cells in transplantation. J Clin Invest 127, 2473–2481 (2017). [PubMed: 28628037]

13. Zeiser R, Blazar BR, Acute Graft-versus-Host Disease - Biologic Process, Prevention, and Therapy. N Engl J Med 377, 2167–2179 (2017). [PubMed: 29171820]

14. Griffioen M, van Bergen CA, Falkenburg JH, Autosomal Minor Histocompatibility Antigens: How Genetic Variants Create Diversity in Immune Targets. Front Immunol 7, 100 (2016). [PubMed: 27014279]

15. Burman AC, Banovic T, Kuns RD, Clouston AD, Stanley AC, Morris ES, Rowe V, Bofinger H, Skoczylas R, Raffelt N, Fahy O, McColl SR, Engwerda CR, McDonald KP, Hill GR, IFNgamma differentially controls the development of idiopathic pneumonia syndrome and GVHD of the gastrointestinal tract. Blood 110, 1064–1072 (2007). [PubMed: 17449800]

16. MacMillan ML, Robin M, Harris AC, DeFor TE, Martin PJ, Alousi A, Ho VT, Bolanos-Meade J, Ferrara JL, Jones R, Arora M, Blazar BR, Holtan SG, Jacobsohn D, Pasquini M, Socie G, Antin JH, Levine JE, Weisdorf DJ, A refined risk score for acute graft-versus-host disease that predicts response to initial therapy, survival, and transplant-related mortality. Biol Blood Marrow Transplant 21, 761–767 (2015). [PubMed: 25585275]

17. Zeiser R, von Bubnoff N, Butler J, Mohty M, Niederwieser D, Or R, Szer J, Wagner EM, Zuckerman T, Mahuzier B, Xu J, Wilke C, Gandhi KK, Socie G, Group RT, Ruxolitinib for Glucocorticoid-Refractory Acute Graft-versus-Host Disease. N Engl J Med 382, 1800–1810 (2020). [PubMed: 32320566]

18. Shlomchik WD, Graft-versus-host disease. Nat Rev Immunol 7, 340–352 (2007). [PubMed: 17438575]

19. Markey KA, MacDonald KP, Hill GR, The biology of graft-versus-host disease: experimental systems instructing clinical practice. Blood 124, 354–362 (2014). [PubMed: 24914137]

20. Zeiser R, Blazar BR, Preclinical models of acute and chronic graft-versus-host disease: how predictive are they for a successful clinical translation? Blood 127, 3117–3126 (2016). [PubMed: 26994149]

21. Wu Y, Fu J, Wang H, Yu XZ, Donor T-Cell Repertoire Profiling in Recipient Lymphoid and Parenchyma Organs Reveals GVHD Pathogenesis at Clonal Levels After Bone Marrow Transplantation in Mice. Front Immunol 12, 778996 (2021).

22. Latis E, Michonneau D, Leloup C, Varet H, Peffault de Latour R, Consortium C, Bianchi E, Socie G, Rogge L, Cellular and molecular profiling of T-cell subsets at the onset of human acute GVHD. Blood Adv 4, 3927–3942 (2020). [PubMed: 32818226]

23. Dietrich PY, Caignard A, Diu A, Genevee C, Pico JL, Henry-Amar M, Bosq J, Angevin E, Triebel F, Hercend T, Analysis of T-cell receptor variability in transplanted patients with acute graft-versus-host disease. Blood 80, 2419–2424 (1992). [PubMed: 1421414]

24. Dietrich PY, Caignard A, Lim A, Chung V, Pico JL, Pannetier C, Kourilsky P, Hercend T, Even J, Triebel F, In vivo T-cell clonal amplification at time of acute graft-versus-host disease. Blood 84, 2815–2820 (1994). [PubMed: 7919391]

25. Kubo K, Yamanaka K, Kiyoi H, Fukutani H, Ito M, Hayakawa R, Ohno R, Naoe T, Different T-cell receptor repertoires between lesions and peripheral blood in acute graft-versus-host disease after allogeneic bone marrow transplantation. Blood 87, 3019–3026 (1996). [PubMed: 8639925]

26. Liu X, Chesnokova V, Forman SJ, Diamond DJ, Molecular analysis of T-cell receptor repertoire in bone marrow transplant recipients: evidence for oligoclonal T-cell expansion in graft-versus-host disease lesions. Blood 87, 3032–3044 (1996). [PubMed: 8639927]

27. Koyama D, Murata M, Hanajiri R, Akashi T, Okuno S, Kamoshita S, Julamanee J, Takagi E, Miyao K, Sakemura R, Goto T, Terakura S, Nishida T, Kiyoi H, Quantitative Assessment of T Cell Clonotypes in Human Acute Graft-versus-Host Disease Tissues. Biol Blood Marrow Transplant 25, 417–423 (2019). [PubMed: 30359734]

28. Margolis DA, Casper JT, Segura AD, Janczak T, McOlash L, Fisher B, Miller K, Gorski J, Infiltrating T cells during liver graft-versus-host disease show a restricted T-cell repertoire. Biol Blood Marrow Transplant 6, 408–415 (2000). [PubMed: 10917576]

29. Hirokawa M, Matsutani T, Saitoh H, Ichikawa Y, Kawabata Y, Horiuchi T, Kitabayashi A, Yoshioka T, Tsuruta Y, Suzuki R, Miura AB, Sawada K, Distinct TCRAV and TCRBV repertoire and CDR3 sequence of T lymphocytes clonally expanded in blood and GVHD lesions after human allogeneic bone marrow transplantation. Bone Marrow Transplant 30, 915–923 (2002). [PubMed: 12476285]

30. Beck RC, Wlodarski M, Gondek L, Theil KS, Tuthill RJ, Sobeck R, Bolwell B, Maciejewski JP, Efficient identification of T-cell clones associated with graft-versus-host disease in target tissue allows for subsequent detection in peripheral blood. Br J Haematol 129, 411–419 (2005). [PubMed: 15842666]

31. Meyer EH, Hsu AR, Liliental J, Lohr A, Florek M, Zehnder JL, Strober S, Lavori P, Miklos DB, Johnson DS, Negrin RS, A distinct evolution of the T-cell repertoire categorizes treatment refractory gastrointestinal acute graft-versus-host disease. Blood 121, 4955–4962 (2013). [PubMed: 23652802]

32. Kanakry CG, Coffey DG, Towlerton AM, Vulic A, Storer BE, Chou J, Yeung CC, Gocke CD, Robins HS, O'Donnell PV, Luznik L, Warren EH, Origin and evolution of the T cell repertoire after posttransplantation cyclophosphamide. JCI Insight 1, (2016).

33. Divito SJ, Aasebo AT, Matos TR, Hsieh PC, Collin M, Elco CP, O'Malley JT, Baekkevold ES, Reims H, Gedde-Dahl T, Hagerstrom M, Hilaire J, Lian JW, Milford EL, Pinkus GS, Ho VT, Soiffer RJ, Kim HT, Mihm MC, Ritz J, Guleria I, Cutler CS, Clark RA, Jahnsen FL, Kupper TS, Peripheral host T cells survive hematopoietic stem cell transplantation and promote graft-versus-host disease. J Clin Invest 130, 4624–4636 (2020). [PubMed: 32516138]

34. Strobl J, Pandey RV, Krausgruber T, Bayer N, Kleissl L, Reininger B, Vieyra-Garcia P, Wolf P, Jentus MM, Mitterbauer M, Wohlfarth P, Rabitsch W, Stingl G, Bock C, Stary G, Long-term skin-resident memory T cells proliferate in situ and are involved in human graft-versus-host disease. Sci Transl Med 12, (2020).

35. Tkachev V, Kaminski J, Potter EL, Furlan SN, Yu A, Hunt DJ, McGuckin C, Zheng H, Colonna L, Gerdemann U, Carlson J, Hoffman M, Olvera J, English C, Baldessari A, Panoskaltsis-Mortari

A, Watkins B, Qayed M, Suessmuth Y, Betz K, Bratrude B, Langston A, Horan JT, Ordovas-Montanes J, Shalek AK, Blazar BR, Roederer M, Kean LS, Spatiotemporal single-cell profiling reveals that invasive and tissue-resident memory donor CD8(+) T cells drive gastrointestinal acute graft-versus-host disease. Sci Transl Med 13, (2021).

36. Iacobuzio-Donahue CA, Michael C, Baez P, Kappagantula R, Hooper JE, Hollman TJ, Cancer biology as revealed by the research autopsy. Nat Rev Cancer 19, 686–697 (2019). [PubMed: 31519982]

37. Bayraktar UD, de Lima M, Saliba RM, Maloy M, Castro-Malaspina HR, Chen J, Rondon G, Chiattone A, Jakubowski AA, Boulad F, Kernan NA, O'Reilly RJ, Champlin RE, Giralt S, Andersson BS, Papadopoulos EB, Ex vivo T cell-depleted versus unmodified allografts in patients with acute myeloid leukemia in first complete remission. Biol Blood Marrow Transplant 19, 898–903 (2013). [PubMed: 23467126]

38. Hill GR, Betts BC, Tkachev V, Kean LS, Blazar BR, Current Concepts and Advances in Graft-Versus-Host Disease Immunology. Annu Rev Immunol 39, 19–49 (2021). [PubMed: 33428454]

39. Emerson RO, DeWitt WS, Vignali M, Gravley J, Hu JK, Osborne EJ, Desmarais C, Klinger M, Carlson CS, Hansen JA, Rieder M, Robins HS, Immunosequencing identifies signatures of cytomegalovirus exposure history and HLA-mediated effects on the T cell repertoire. Nat Genet 49, 659–665 (2017). [PubMed: 28369038]

40. DeWitt WS 3rd, Smith A, Schoch G, Hansen JA, Matsen F. A. t., Bradley P, Human T cell receptor occurrence patterns encode immune history, genetic background, and receptor specificity. Elife 7, (2018).

41. DeWolf S, Grinshpun B, Savage T, Lau SP, Obradovic A, Shonts B, Yang S, Morris H, Zuber J, Winchester R, Sykes M, Shen Y, Quantifying size and diversity of the human T cell alloresponse. JCI Insight 3, (2018).

42. van Heijst JW, Ceberio I, Lipuma LB, Samilo DW, Wasilewski GD, Gonzales AM, Nieves JL, van den Brink MR, Perales MA, Pamer EG, Quantitative assessment of T cell repertoire recovery after hematopoietic stem cell transplantation. Nat Med 19, 372–377 (2013). [PubMed: 23435170]

43. Gkazi AS, Margetts BK, Attenborough T, Mhaldien L, Standing JF, Oakes T, Heather JM, Booth J, Pasquet M, Chiesa R, Veys P, Klein N, Chain B, Callard R, Adams SP, Clinical T Cell Receptor Repertoire Deep Sequencing and Analysis: An Application to Monitor Immune Reconstitution Following Cord Blood Transplantation. Front Immunol 9, 2547 (2018). [PubMed: 30455696]

44. Yew PY, Alachkar H, Yamaguchi R, Kiyotani K, Fang H, Yap KL, Liu HT, Wickrema A, Artz A, van Besien K, Imoto S, Miyano S, Bishop MR, Stock W, Nakamura Y, Quantitative characterization of T-cell repertoire in allogeneic hematopoietic stem cell transplant recipients. Bone Marrow Transplant 50, 1227–1234 (2015). [PubMed: 26052909]

45. Buhler S, Bettens F, Dantin C, Ferrari-Lacraz S, Ansari M, Mamez AC, Masouridi-Levrat S, Chalandon Y, Villard J, Genetic T-cell receptor diversity at 1 year following allogeneic hematopoietic stem cell transplantation. Leukemia 34, 1422–1432 (2020). [PubMed: 31772297]

46. Suessmuth Y, Mukherjee R, Watkins B, Koura DT, Finstermeier K, Desmarais C, Stempora L, Horan JT, Langston A, Qayed M, Khoury HJ, Grizzle A, Cheeseman JA, Conger JA, Robertson J, Garrett A, Kirk AD, Waller EK, Blazar BR, Mehta AK, Robins HS, Kean LS, CMV reactivation drives posttransplant T-cell reconstitution and results in defects in the underlying TCRbeta repertoire. Blood 125, 3835–3850 (2015). [PubMed: 25852054]

47. Pagliuca S, Gurnari C, Hong S, Zhao R, Kongkiatkamon S, Terkawi L, Zawit M, Guan Y, Awada H, Kishtagari A, Kerr CM, LaFramboise T, Patel BJ, Jha BK, Carraway HE, Visconte V, Majhail NS, Hamilton BK, Maciejewski JP, Clinical and basic implications of dynamic T cell receptor clonotyping in hematopoietic cell transplantation. JCI Insight 6, (2021).

48. G. T. Consortium, The Genotype-Tissue Expression (GTEx) project. Nat Genet 45, 580–585 (2013). [PubMed: 23715323]

49. Elhanati Y, Sethna Z, Callan CG Jr., Mora T, Walczak AM, Predicting the spectrum of TCR repertoire sharing with a data-driven model of recombination. Immunol Rev 284, 167–179 (2018). [PubMed: 29944757]

50. Glanville J, Huang H, Nau A, Hatton O, Wagar LE, Rubelt F, Ji X, Han A, Krams SM, Pettus C, Haas N, Arleham CSL, Sette A, Boyd SD, Scriba TJ, Martinez OM, Davis MM,

Identifying specificity groups in the T cell receptor repertoire. Nature 547, 94–98 (2017). [PubMed: 28636589]

51. Huang H, Wang C, Rubelt F, Scriba TJ, Davis MM, Analyzing the Mycobacterium tuberculosis immune response by T-cell receptor clustering with GLIPH2 and genome-wide antigen screening. Nat Biotechnol 38, 1194–1202 (2020). [PubMed: 32341563]

52. Chiou SH, Tseng D, Reuben A, Mallajosyula V, Molina IS, Conley S, Wilhelmy J, McSween AM, Yang X, Nishimiya D, Sinha R, Nabet BY, Wang C, Shrager JB, Berry MF, Backhus L, Lui NS, Wakelee HA, Neal JW, Padda SK, Berry GJ, Delaidelli A, Sorensen PH, Sotillo E, Tran P, Benson JA, Richards R, Labanieh L, Klysz DD, Louis DM, Feldman SA, Diehn M, Weissman IL, Zhang J, Wistuba II, Futreal PA, Heymach JV, Garcia KC, Mackall CL, Davis MM, Global analysis of shared T cell specificities in human non-small cell lung cancer enables HLA inference and antigen discovery. Immunity 54, 586–602 e588 (2021). [PubMed: 33691136]

53. Miller WP, Srinivasan S, Panoskaltsis-Mortari A, Singh K, Sen S, Hamby K, Deane T, Stempora L, Beus J, Turner A, Wheeler C, Anderson DC, Sharma P, Garcia A, Strobert E, Elder E, Crocker I, Crenshaw T, Penedo MC, Ward T, Song M, Horan J, Larsen CP, Blazar BR, Kean LS, GVHD after haploidentical transplantation: a novel, MHC-defined rhesus macaque model identifies CD28-CD8+ T cells as a reservoir of breakthrough T-cell proliferation during costimulation blockade and sirolimus-based immunosuppression. Blood 116, 5403–5418 (2010). [PubMed: 20833977]

54. de Almeida GP, Lichtner P, Eckstein G, Brinkschmidt T, Chu CF, Sun S, Reinhard J, Madler SC, Kloeppel M, Verbeek M, Zielinski CE, Human skin-resident host T cells can persist long term after allogeneic stem cell transplantation and maintain recirculation potential. Sci Immunol 7, eabe2634 (2022).

55. Sacirbegovic F, Gunther M, Greco A, Zhao D, Wang X, Zhou M, Rosenberger S, Oberbarnscheidt MH, Held W, McNiff J, Jain D, Hofer T, Shlomchik WD, Graft-versus-host disease is locally maintained in target tissues by resident progenitor-like T cells. Immunity 56, 369–385 e366 (2023). [PubMed: 36720219]

56. Robins HS, Campregher PV, Srivastava SK, Wacher A, Turtle CJ, Kahsai O, Riddell SR, Warren EH, Carlson CS, Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. Blood 114, 4099–4107 (2009). [PubMed: 19706884]

57. Farber DL, Tissues, not blood, are where immune cells function. Nature 593, 506–509 (2021). [PubMed: 34035530]

58. Scheifele D, Bjornson G, Dimmick J, Rapid postmortem gut autolysis in infant rats: a potential problem for investigators. Can J Vet Res 51, 404–406 (1987). [PubMed: 3651898]

59. Salomao M, Dorritie K, Mapara MY, Sepulveda A, Histopathology of Graft-vs-Host Disease of Gastrointestinal Tract and Liver: An Update. Am J Clin Pathol 145, 591–603 (2016). [PubMed: 27247365]

60. Snyder ME, Moghbeli K, Bondonese A, Craig A, Popescu I, Fan L, Tabib T, Lafyatis R, Chen K, Trejo Bittar HE, Lendermon E, Pilewski J, Johnson B, Kilaru S, Zhang Y, Sanchez PG, Alder JK, Sims PA, McDyer JF, Modulation of tissue resident memory T cells by glucocorticoids after acute cellular rejection in lung transplantation. J Exp Med 219, (2022).

61. Morris H, DeWolf S, Robins H, Sprangers B, LoCascio SA, Shonts BA, Kawai T, Wong W, Yang S, Zuber J, Shen Y, Sykes M, Tracking donor-reactive T cells: Evidence for clonal deletion in tolerant kidney transplant patients. Sci Transl Med 7, 272ra210 (2015).

62. Zheng P, Tamaresis J, Thangavelu G, Xu L, You X, Blazar BR, Negrin RS, Zehnder JL, Iliopoulou BP, Meyer EH, Recipient-specific T-cell repertoire reconstitution in the gut following murine hematopoietic cell transplant. Blood Adv 4, 4232–4243 (2020). [PubMed: 32898248]

63. Poon MML, Byington E, Meng W, Kubota M, Matsumoto R, Grifoni A, Weiskopf D, Dogra P, Lam N, Szabo PA, Ural BB, Wells SB, Rosenfeld AM, Brusko MA, Brusko TM, Connors TJ, Sette A, Sims PA, Luning Prak ET, Shen Y, Farber DL, Heterogeneity of human anti-viral immunity shaped by virus, tissue, age, and sex. Cell Rep 37, 110071 (2021).

64. Morin-Zorman S, Wysocki C, Zhu J, Li H, Zorman S, Matte-Martone C, Kisanga E, McNiff J, Jain D, Gonzalez D, Rothstein DM, Lakkis FG, Haberman A, Shlomchik WD, In vivo dynamics of T cells and their interactions with dendritic cells in mouse cutaneous graft-versus-host disease. Blood Adv 3, 2082–2092 (2019). [PubMed: 31296496]

65. Abou-Daya KI, Tieu R, Zhao D, Rammal R, Sacirbegovic F, Williams AL, Shlomchik WD, Oberbarnscheidt MH, Lakkis FG, Resident memory T cells form during persistent antigen exposure leading to allograft rejection. Sci Immunol 6, (2021).

66. Dierov D, Webb N, Fatmi S, Nwanne C, Ciolino C, Mosesso K, Nieves J, Perales MA, Prockop SE, Ponce DM, Establishing a standardized system for review and adjudication of chronic graft-vs-host disease data in accordance with the National Institutes Consensus criteria. Adv Cell Gene Ther 2, (2019).

67. Deng C, Daley T, Calabrese P, Ren J, Smith AD, Predicting the Number of Bases to Attain Sufficient Coverage in High-Throughput Sequencing Experiments. J Comput Biol 27, 1130–1143 (2020). [PubMed: 31725321]

68. Godfrey DI, Uldrich AP, McCluskey J, Rossjohn J, Moody DB, The burgeoning family of unconventional T cells. Nat Immunol 16, 1114–1123 (2015). [PubMed: 26482978]

69. Werner L, Nunberg MY, Rechavi E, Lev A, Braun T, Haberman Y, Lahad A, Shteyer E, Schvimer M, Somech R, Weiss B, Lee YN, Shouval DS, Altered T cell receptor beta repertoire patterns in pediatric ulcerative colitis. Clin Exp Immunol 196, 1–11 (2019). [PubMed: 30556140]

70. Mele M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, Sammeth M, Young TR, Goldmann JM, Pervouchine DD, Sullivan TJ, Johnson R, Segre AV, Djebali S, Niarchou A, Consortium GT, Wright FA, Lappalainen T, Calvo M, Getz G, Dermitzakis ET, Ardlie KG, Guigo R, Human genomics. The human transcriptome across tissues and individuals. Science 348, 660–665 (2015). [PubMed: 25954002]

71. Zhao S, Zhang Y, Gamini R, Zhang B, von Schack D, Evaluation of two main RNA-seq approaches for gene quantification in clinical RNA sequencing: polyA+ selection versus rRNA depletion. Sci Rep 8, 4781 (2018). [PubMed: 29556074]

72. Tsai JJ, Velardi E, Shono Y, Argyropoulos KV, Holland AM, Smith OM, Yim NL, Rao UK, Kreines FM, Lieberman SR, Young LF, Lazrak A, Youssef S, Fu YY, Liu C, Lezcano C, Murphy GF, Na IK, Jenq RR, Hanash AM, Dudakov JA, van den Brink MRM, Nrf2 regulates CD4(+) T cell-induced acute graft-versus-host disease in mice. Blood 132, 2763–2774 (2018). [PubMed: 30381375]

73. Kousa A, Lemarquis AL, The shunPykeR's guide to single cell analysis (Version 1.0.0) [Computer software]. 10.5281/zenodo.7510613. (2023).

74. Wolock SL, Lopez R, Klein AM, Scrublet: Computational Identification of Cell Doublets in Single-Cell Transcriptomic Data. Cell Syst 8, 281–291 e289 (2019). [PubMed: 30954476]

75. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, Baglaenko Y, Brenner M, Loh PR, Raychaudhuri S, Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods 16, 1289–1296 (2019). [PubMed: 31740819]
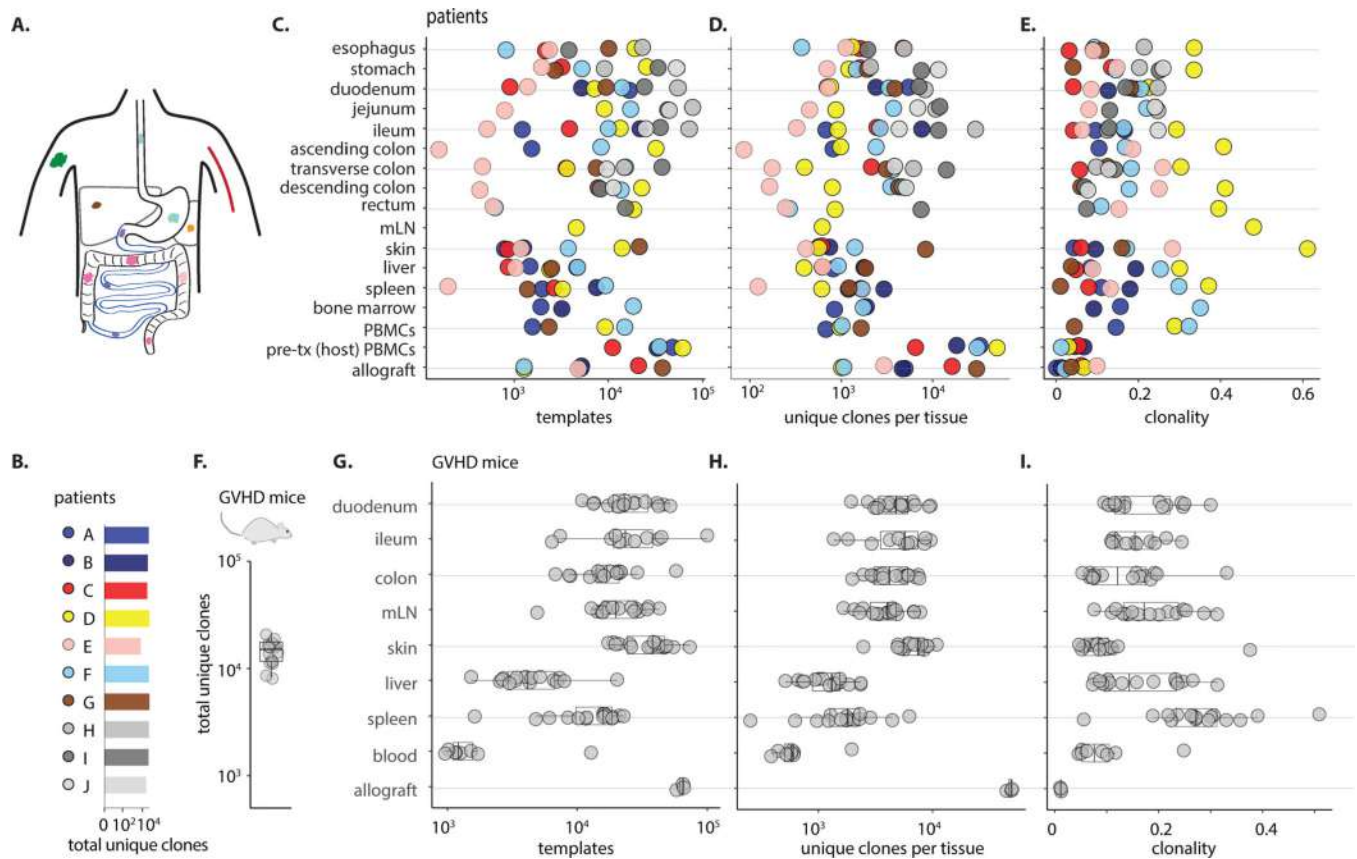
Fig. 1: Robust recovery TCR clones by sequencing genomic DNA from autopsy tissues.
(**A**) Schematic of tissues sampled from research-autopsy cohort: esophagus, stomach, small intestine (duodenum, jejunum, ileum), large intestine (ascending, transverse, descending), rectum, liver, skin, spleen, blood. (**B**) Number of unique clones recovered per patient, as defined by the nucleotide sequence of the CDR3 region (n = 102 total samples) from 7 GVHD patients (different colors) and 3 comparator patients (shades of grey) without GVHD. (**C**) Number of productive templates recovered from each sample. (**D**) Number of productive unique clones recovered from each sample. (**E**) Sample clonality. In (**C-E**), colors represent individual patients as in panel (**B**); samples with missing precise anatomic locations within the small intestine or large intestine were labeled as duodenum and ascending colon, respectively; one sample collected as 'mid-colon' was labeled as transverse colon. Post-mortem blood was collected by caval or cardiac puncture but is labeled here as "peripheral" blood mononuclear cells (PBMC) to use the conventional nomenclature for circulating blood cells; 'allograft' indicates material from the original graft (patients A, B, C, D, F, G) or donor PBMCs (patient E). (**F**) Number of unique clones recovered per GVHD mouse, as defined by the nucleotide sequence of the CDR3 region (n = 115 total samples from 16 mice, details Fig. S1A). (**G**) Number of productive templates (**H**) number of productive unique clones, and (**I**) clonality of each mouse sample. In (**G-I**), left and right edges of boxplot correspond to the first and third quartiles, respectively. Values are tabulated in Table S4.
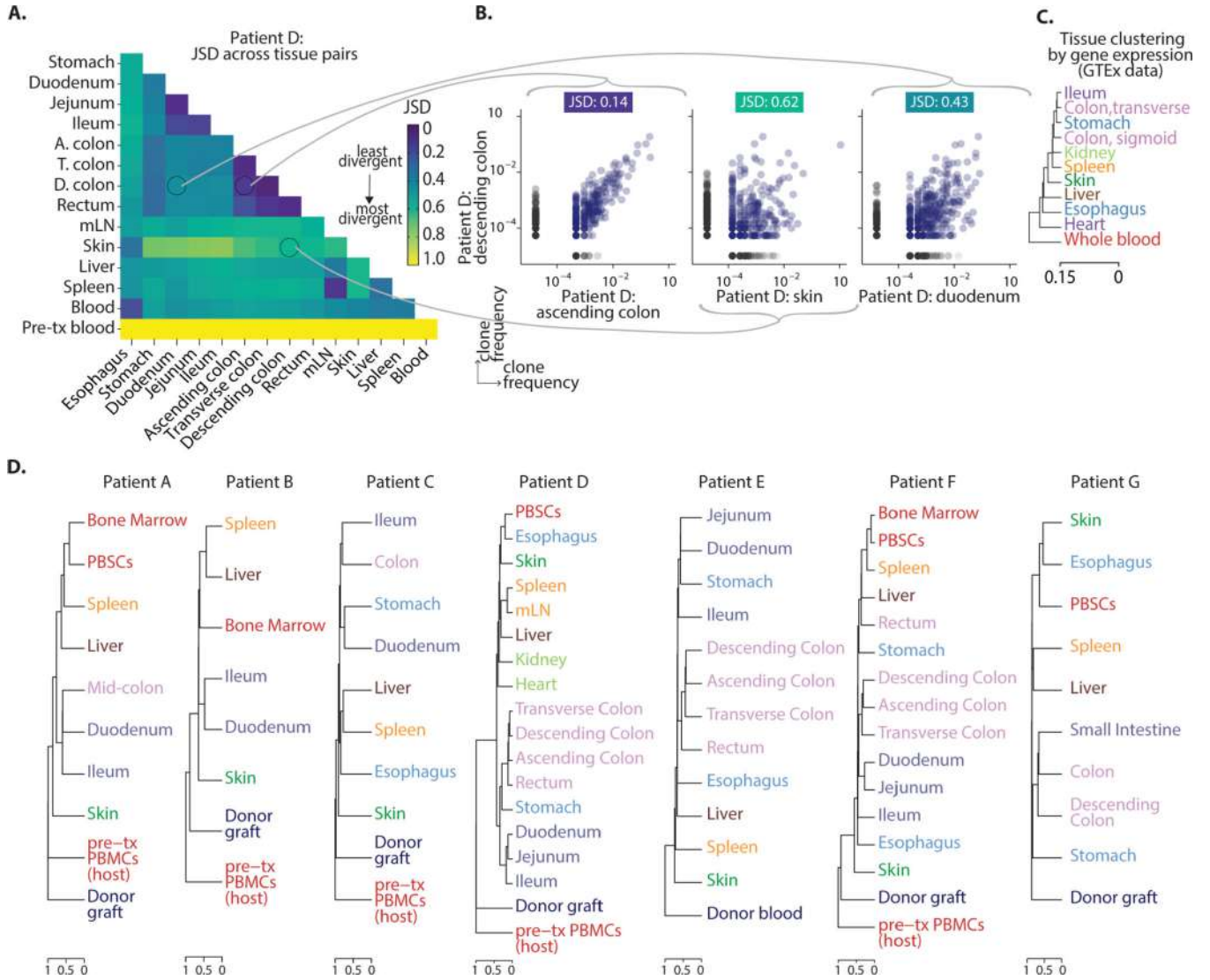
Fig. 2: TCR repertoires in different tissues are distinct, while anatomically related tissues harbor a greater degree of TCR overlap.

(**A**) Heatmap of pairwise Jensen-Shannon divergence (JSD) values measure repertoire divergence, where 0 indicates identical repertoires and 1 indicates completely divergent repertoires, for all samples from patient D; values from other patients are tabulated in Table S6. (**B**) Scatter plots comparing repertoire overlap between illustrative sample pairs from patient D. Blue points are clones observed in both tissues, gray clones were observed in only one sample of the pair; ascending colon (n= 395 clones); descending colon (n= 799), overlap = 230; skin (n= 567), descending colon (n= 799), overlap = 130; duodenum (n= 764), descending colon (n= 799), overlap = 253. A pseudocount of 10^{-5} was added to frequencies to facilitate plotting on a logarithmic scale. (**C**) Clustering of tissues across anatomy by RNAseq gene expression profiling (agglomerative hierarchical clustering of JSD with single linkage) from the publicly available GTEx repository (see Methods). Scale bars correspond to nodes as the heights of clustering hierarchy. (**D**) Dendrograms constructed via agglomerative hierarchical clustering with single linkage of pairwise JSD comparisons

(Table S6) illustrate reproducible clustering of anatomically related tissues within each of 7 GVHD patients (patients A-G); see Fig. S3D for comparator patients. Branch heights show minimum pairwise distance between tissue pair JSD scores. Pre-transplant PBMCs are of host origin. Abbreviations: ascending colon (A. colon), transverse colon (T. colon), descending colon (D. colon).
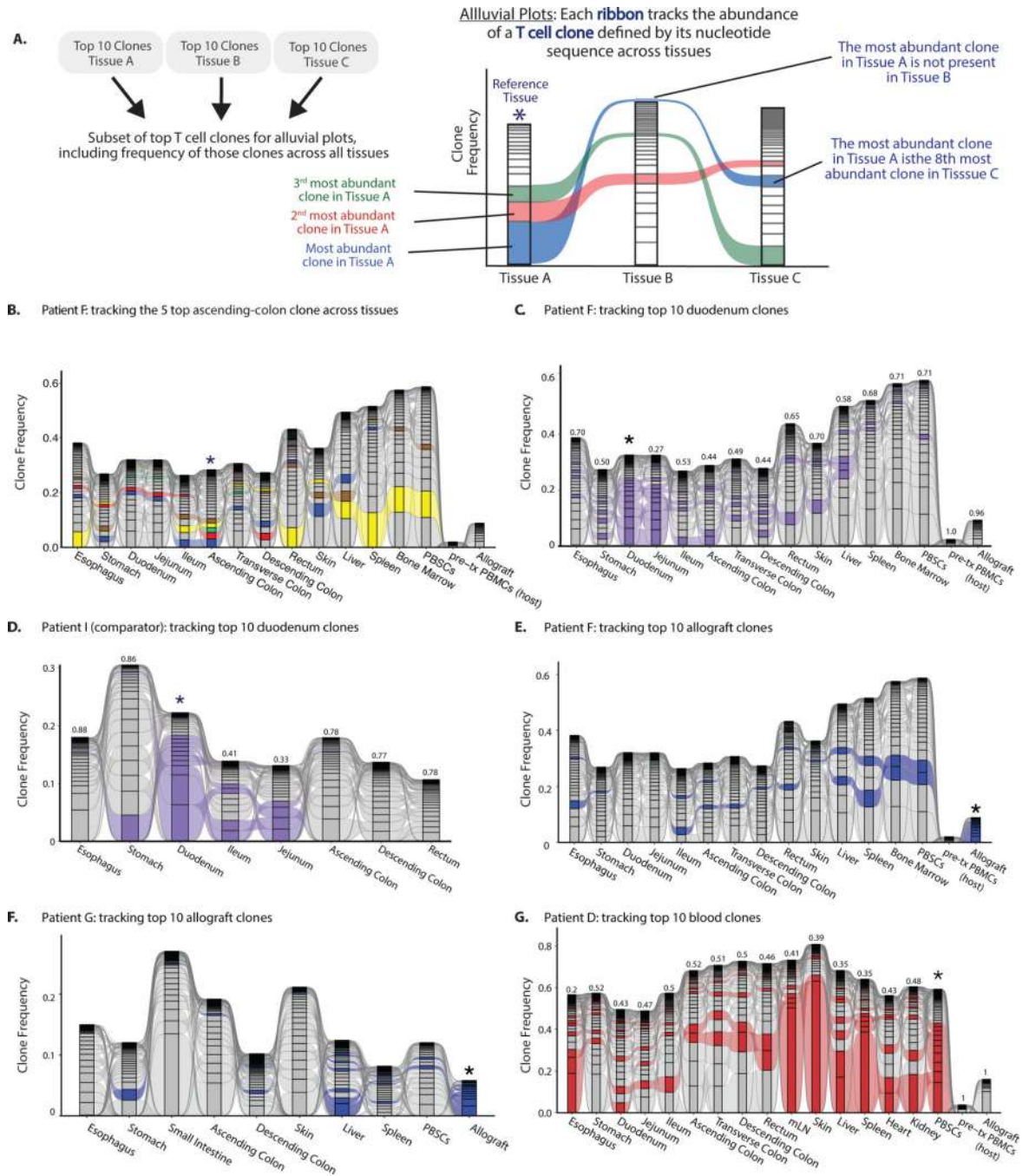
Fig. 3: Evidence for dominant clones shared across multiple sampled tissues.
Alluvial plots facilitate tracking of the frequency abundant TCR clones from a tissue across the other samples from the same patient. (**A**) Schematic illustrating that the most abundant clones from each tissue of a single individual were selected for visualization. Each tissue is represented by a bar of stacked units whose heights represent clone frequency, with the most abundant clones at the bottom. Colored ribbons between tissues trace the relative abundance of an individual TCR clone across the individual. Asterisk (*) marks the tissue whose top *n* clones are highlighted. (**B**) Alluvial plot tracking the frequency of the top

5 TCR clones from the ascending colon in patient F across all tissues from this patient. Asterisk indicates reference tissue. (**C-D**) Alluvial plots highlight top 10 duodenum TCRs in purple from a GVHD patient (**C**) and a comparator (**D**); see Fig. S4A–H for all patients. (**E-F**) Alluvial plots highlight top 10 donor allograft TCRs in dark blue from GVHD patients F and G. (**G**) Top 10 blood TCRs are highlighted in red from GVHD patient D. CDR3 of top clones are tabulated in Table S7; proportions of shared clones are tabulated in Table S8. Post-mortem blood was collected by caval or cardiac puncture but is labeled here as "peripheral" blood mononuclear cells (PBMC) to use the conventional nomenclature for circulating blood cells. Pre-tx indicates host pre-transplant PBMCs. For **C-D,G**, Jensen-Shannon divergence (JSD) values above each bar indicate the repertoire divergence of the tissue from the reference sample, where 0 indicates identical repertoires and 1 indicates completely divergent repertoires.
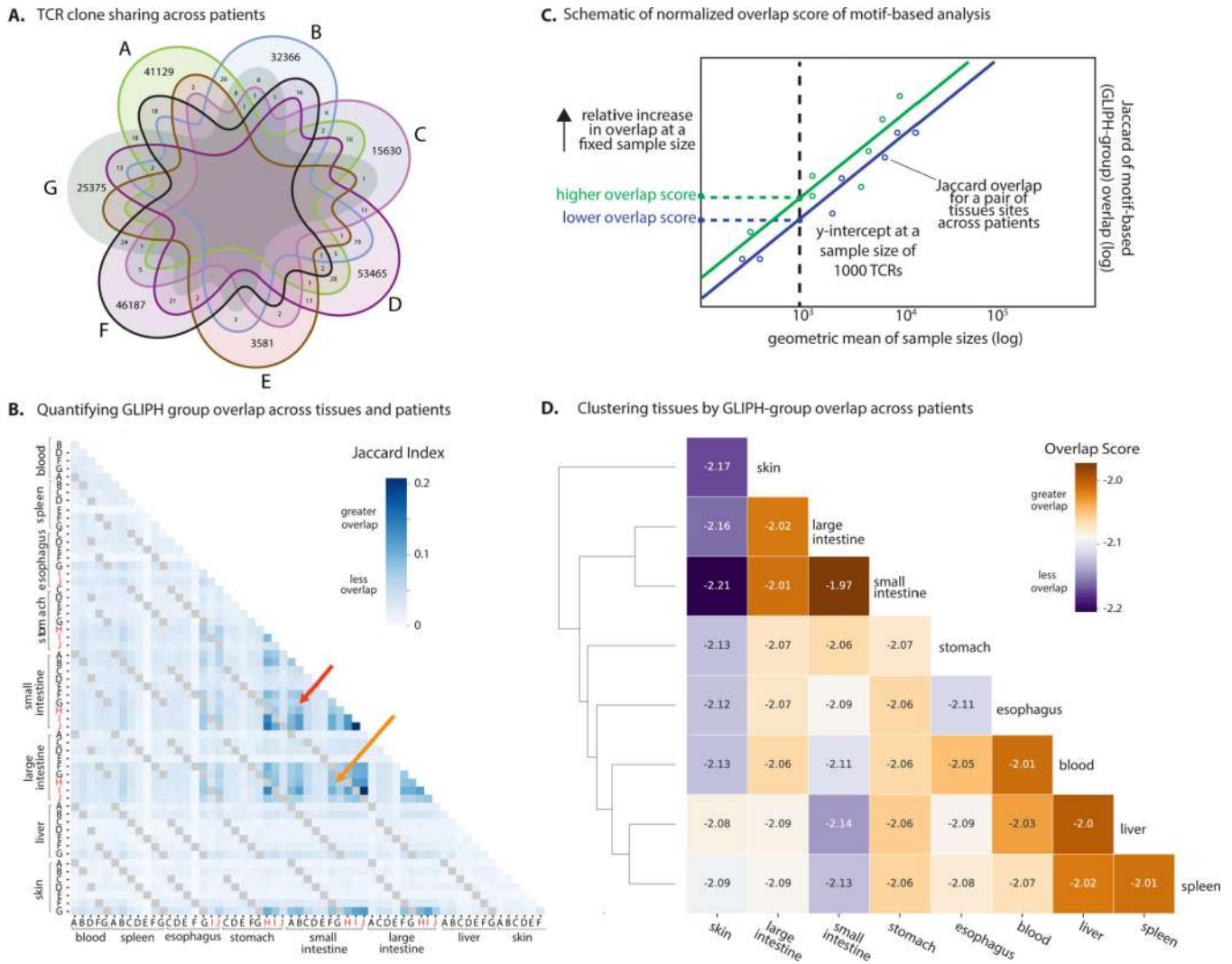
Fig. 4: Shared clonal motifs across patients in similar tissues.

(**A**) Venn diagram indicates the number of shared clones (nucleotide level) between sets of tissues illustrating lack of sharing across GVHD patients. Letters are patient identifiers. Diagram segments without any numerals have 0 shared clones (**B**) Pairwise comparison of GLIPH groups across anatomic tissues by each patient, quantified by the Jaccard similarity index in which higher values indicate more similarity. Gray cells are within-patient comparisons whose Jaccard values are not color-coded. Red (small intestine pair) and orange (small intestine compared to large intestine) arrows highlight tissues with relatively high overlap in GLIPH groups; GLIPH groups filtered for having > two unique CDR3s (**C**) Schematic of normalized overlap score of a motif-based analysis, such as GLIPH, enabling comparison of T cells in populations with disparate densities of T cell infiltrates. Overlap measure for a subgroup of tissue pairs (here one group in green and one in blue) defined by fitting the dependence using a common slope and assigning a standard Jaccard overlap for samples with geometric mean of 1000 TCRs. Standard overlap values highlighted by the dashed lines on the y-axis, with greater values on the y-axis indicating greater overlap within this sample group, accounting for sample size. (**D**) Calculated standard overlap for

all possible pairs (excluding patient D, see Methods) with heat map highlighting hierarchical clustering. Details regarding calculation of overlap score are in Fig. S6.
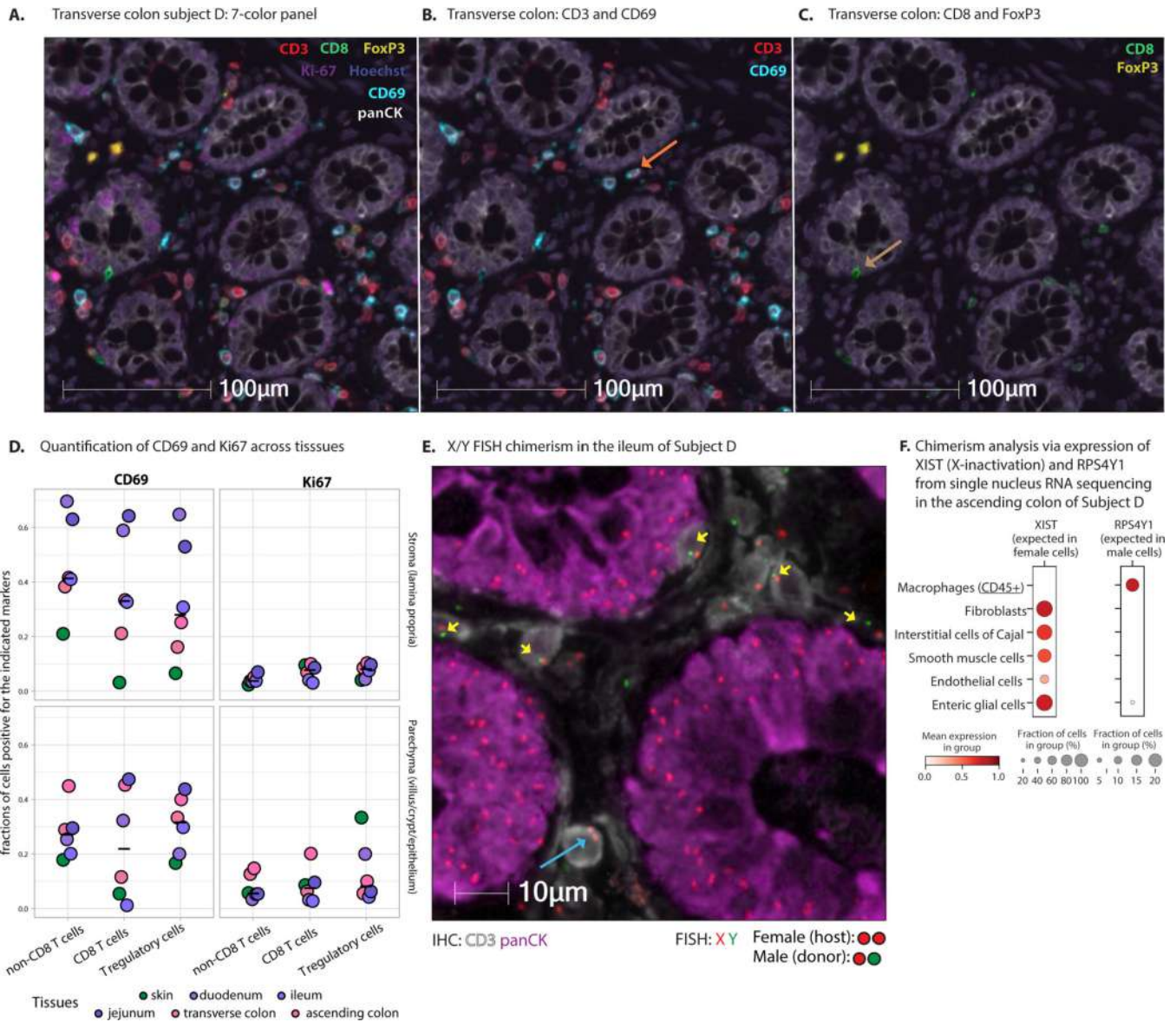
Fig. 5: T cells in GVHD tissues have a tissue-resident phenotype.
(**A-C**) Seven-color multiplex imaging of the transverse colon from patient D for characterization of CD3, CD8, regulatory T cell (Treg) marker FoxP3, tissue-resident marker CD69, proliferation marker Ki-67, epithelial marker panCK, and DAPI nuclear stain within a single tissue section. (**A**) Transverse colon specimen with all channels displayed; cross sections of intestinal crypts are defined by white panCK staining of epithelial cells. (**B**) red CD3 and blue CD69 co-expression in the same lamina propria cells (e.g. orange arrow) highlight tissue-resident T cells in the same section of transverse colon as in panel (**A**); (**C**) green CD8 stain identifies an intraepithelial CD8+ cell (brown arrow) and yellow Foxp3 stain identifies lamina propria Tregs in the same section of transverse colon in panels (**A**) and (**B**). (**D**) Automated algorithms were used both to segment images into individual cells and to annotate boundaries of two histological structures: panCK+ epithelial structures (i.e., villi/crypts) vs. panCK− stroma (i.e., lamina propria) in (6 tissues from

patient D). The fraction of all cells classified as having a tissue-resident phenotype (CD69[+]) or actively proliferating (Ki67[+]) within the non-CD8 T cell (CD3[+] CD8[−]), CD8 T cell (CD8[+]), and Treg (FoxP3[+]) compartments were enumerated (Table S12) and are plotted, stratified by histological zones. Each point is the fraction of cells of the population listed at the bottom that is positive for the marker listed; colors indicate different tissues. Black bars are median values across all tissues. Expanded imaging analysis included in Fig. S7E. (**E**) Representative image of combined immunophenotyping and fluorescence in situ hybridization (FISH) for X (red) and Y (green) chromosomes in the ileum of patient D (a female recipient of a male allograft) highlighting presence of predominantly male donor XY genotype (red/green dots) among CD3[+] T cells (white), in contrast to female recipient XX genotype (two red dots) in host parenchyma (magenta panCK[+]). Yellow arrows point to donor XY T cells; blue arrow points to a rare recipient T cell. (**F**) Dot plot indicating gene expression and coverage of XIST (X-chromosome inactivation gene) and RPS4Y1 (gene on Y-chromosome) in patient D from single nucleus RNA sequencing of the ascending colon; additional sequencing details and results from patient F and I included in Fig. S11.
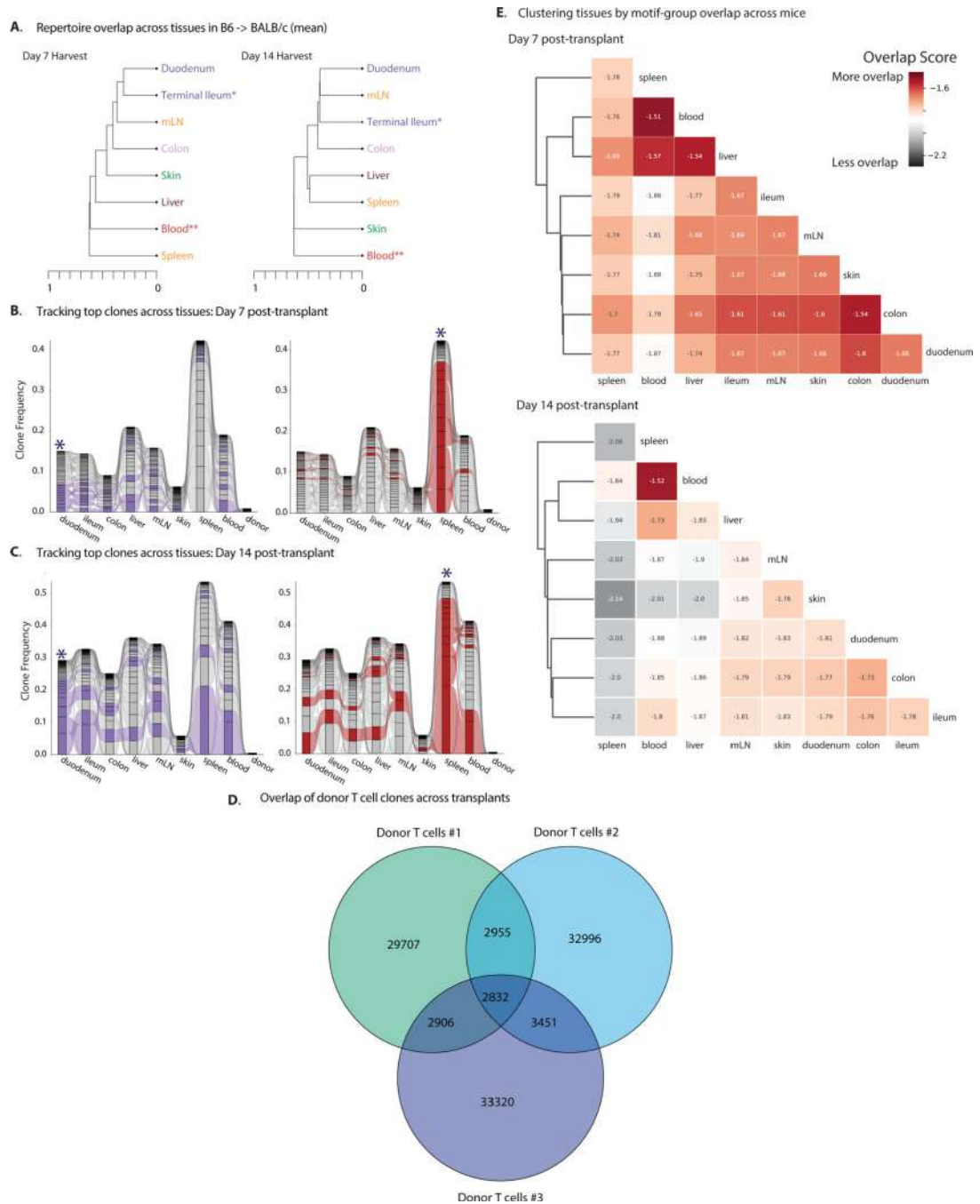
Fig. 6: In mouse GVHD, anatomically related tissues harbor greater repertoire similarity. (**A**) Dendrograms constructed via agglomerative hierarchical clustering with single linkage of mean JSD comparisons between tissue pairs at day 7 or 14 in a major MHC-mismatch mouse model of acute GVHD (nucleotide level). Branch heights show minimum pairwise distance between tissue pair JSD scores. Mean of 8 mice, pooled across two experiments, except where indicated: *n = 5–6 (two day 7 harvests, one day 14 harvest); ** n = 3 (one transplant); standard deviations included in Table S14. (**B, C**) Alluvial plots tracking top 10 clones identified in duodenum (purple) or spleen (red) across representative mice from Day

7 (**B**) or Day 14 (**C**) post-transplant. Asterisks indicate reference tissue. Additional alluvial plots in Fig. S9. Proportions of shared clones for spleen (all mice) are tabulated in Table S8. (**D**) Venn diagram of clonal overlap of donor T cell pool (1–2 mouse spleens per transplant) across three separate transplants; TCRs were defined at the amino-acid-sequence level. (**E**) Clustering tissues by motif-group overlap across mice. Motif-based analysis performed via GLIPH2 (see Methods and Fig. S10). Calculated standard overlap for all possible pairs of mouse GVHD samples stratified by day of harvest with heat map highlighting hierarchical clustering. Details regarding calculation of overlap score are in Fig. S6. mLN: mesenteric lymph nodes.

Table 1:

Rapid autopsy patient cohort

Pt ID	Disease	Age† (years)	Sex	GVHD subjects				
				Allograft	Donor type	Conditioning	GVHD Prophylaxis	GVHD sites*
A	AML	R: 20–44 D: 20–44	F	CD34-selected	Matched unrelated	Ablative	T-cell depleted	Lower & upper GI, liver, skin
B	MM	R: 45–64 D: 20–44	M	Unmodified PBSC	Matched unrelated	Reduced intensity	Calcineurin inhibitor/ methotrexate based	Lower & upper GI
C	MCL	R: 45–64 D: 20–44	M	Unmodified PBSC	Matched unrelated	Reduced intensity	Calcineurin inhibitor**/ methotrexate based	Lower & upper GI, mouth, skin, eye
D	MDS	R: 45–64 D:45–64	F	CD34-selected	Matched unrelated	Ablative	T-cell depleted	Lower & upper GI, skin
E	AML	R: 45–64 D: 45–64	F	CD34-selected	Matched related	Ablative	T-cell depleted	Lower & upper GI, skin, eye
F	AML	R: > 65 D: 20–44	M	Unmodified BM	Matched unrelated	Reduced intensity	Calcineurin inhibitor/ methotrexate based	Lower & upper GI, skin
G	AML	R: > 65 D: 20–44	M	Unmodified PBSC	Matched unrelated	Reduced intensity	Calcineurin inhibitor/ methotrexate based	Lower & upper GI
				Non-GVHD comparators				
				Sites of distant metastases				
H	Lung cancer	> 65	F	CNS, liver				
I	Breast cancer	20–44	F	CNS, liver				
J	Lung cancer	> 65	M	Bone, liver, pleura				

Abbreviations: Patient (Pt), graft-versus-host disease (GVHD), acute myeloid leukemia (AML), multiple myeloma (MM), mantle cell lymphoma (MCL), myelodysplastic syndrome (MDS), peripheral blood stem cell (PBSC), bone marrow (BM), central nervous syndrome (CNS).

*Throughout post-transplant time period, acute or chronic (GVHD details in Table S2)

**Tacrolimus and sirolimus

†To protect patient privacy, a range encompassing each patient's age at transplantation is listed. "R" denotes age of recipient and "D" denotes age of donor.