Cohesin Releasing Factor WAPL Regulates Genome Structure and Function of Mature T Cells

Yaping Sun[1,¶], Gabrielle A. Dotson[2,¶], Lindsey A. Muir[2], Scott Ronquist[2], Katherine Oravecz-Wilson[1], Daniel Peltier[1], Keisuke Seike[1], Lu Li[1], Walter Meixner[2], Indika Rajapakse[2,3,*], Pavan Reddy[1,*]

[1]Department of Internal Medicine, Division of Hematology and Oncology, University of Michigan Rogel Cancer Center, Ann Arbor, MI, USA
[2]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA
[3]Department of Mathematics, University of Michigan, Ann Arbor, MI, USA.

[¶]These authors contributed equally to this work.
[*]Co-senior authors

[*]**Address correspondence to:** Indika Rajapakse **(**indikar@umich.edu)

or Pavan Reddy (reddypr@umich.edu)

Tel: +1-734-936-5280, Fax: +1-734-647-9271

Keywords: T cells, Hi-C, chromatin architecture, 3D nucleome, transplantation, GVHD

Conflict of Interest: We have no conflicts of interest to disclose.

ABSTRACT

The cohesin complex modulates gene expression and cellular functions by shaping three-dimensional (3D) organization of chromatin. WAPL, cohesin's DNA releasing factor, regulates 3D chromatin architecture. The 3D genome structure and its relevance to mature T cell functions is not well understood. We show that *in vivo* lymphopenic expansion, and allo-antigen driven proliferation, alters the 3D structure and function of the genome in mature T cells. Conditional deletion of *Wapl* in T cells reduced long-range genomic interactions, altered chromatin A/B compartments and the topologically associating domains (TAD) of the chromatin in T cells at baseline. Comparison of chromatin structure in normal and WAPL-deficient T cells after lymphopenic and allo-antigen driven stimulation revealed reduced loop extensions with changes in cell cycling genes. WAPL-mediated changes in 3D architecture of chromatin regulated activation, cycling and proliferation of T cells *in vitro* and *in vivo*. Finally, WAPL-deficient T cells caused reduced severity of graft-versus-host disease following experimental allogeneic hematopoietic cell transplantation. These data collectively characterize 3D genomic architecture of T cells *in vivo* and demonstrate biological and clinical implications for its disruption by cohesin releasing factor WAPL.

INTRODUCTION

The three-dimensional (3D) architecture of the genome includes coiling of genomic DNA around histone proteins to form the chromatin fiber, which folds into higher-order structures such as loops, domains, compartments, and chromosomes (Bonev and Cavalli, 2016; Finn and Misteli, 2019; Misteli, 2020; Rajapakse and Groudine, 2011; Rajapakse et al., 2011). High resolution chromatin conformation capture experiments reveal that the 3D spatial architecture of the chromatin at various scales is conserved, reproducible at the cellular level, and regulates gene expression (Cremer and Cremer, 2019; Cremer and Cremer, 2001; Dekker and Mirny, 2016; Dekker et al., 2002). It is increasingly appreciated that higher spatial organization can have specific alterations during mammalian development and in some pathologies including cancers and infections (Finn and Misteli, 2019; Waldman, 2020). However, whether *a priori* disruption of this 3D organization alters *in vivo* cellular functions and disease processes remains poorly understood.

The multi-unit cohesin ring complex plays a critical role in 3D genomic organization and in cell division. It consists of SMC1/3, SCC1 (RAD21), and STAG subunits that are loaded by the SCC2/SCC4 complex onto genomic DNA and establish the cohesin ring structure (Cuadrado and Losada, 2020; Nasmyth and Haering, 2009; Piche et al., 2019; Remeseiro et al., 2013). Cohesin dependency has been demonstrated by depletion of various cohesin units (Piche' et al., 2019). Cohesin release from chromatin is driven by WAPL, which opens an exit site at the interface of the SMC3/SCC1 subunits of the cohesin ring (Haarhuis et al., 2013; Haarhuis et al., 2017; Silva et al., 2020). Prior studies have elegantly demonstrated that the absence of WAPL reduces cohesin turnover, alters chromatin loop extensions, and leads to defects in interphase chromosome organization (Busslinger et al., 2017; Haarhuis et al., 2013; Haarhuis et al., 2017; Hill et al., 2020; Silva et al., 2020; Tedeschi et al., 2013).

WAPL is essential during mammalian embryonic development (Tedeschi et al., 2013) and in regulation of B cells (Hill et al., 2020). However, the role of WAPL-dependent chromatin alterations on *in vivo* functions after embryonal development or in mature T cells is not known. The

3D chromatin landscape has been recently described in T cell development and in T cell lines (Hu et al., 2018; Isoda et al., 2017; Misteli, 2020; Rawlings et al., 2011; Robson et al., 2017). It remains unclear however whether the specific changes in the 3D chromatin landscape influence or emerge from T cell development. Furthermore, the functional chromatin landscape in mature T cells and their alterations following *in vivo* activation remain unknown. The role of cohesin ring formation in mature T cell development, its function *in vivo*, and its disruption by the absence of WAPL in mature T cell immunity are also not fully understood.

Mature T cells can cause graft-versus-host disease (GVHD) following allogeneic hematopoietic cell transplantation (HCT), a potentially curative therapy against many hematological malignant and non-malignant diseases (Blazar et al., 2020; Wu and Reddy, 2017; Zeiser and Blazar, 2018). GVHD has precluded widespread utilization of this effective therapy. The chromatin landscape of allo-antigen stimulated T cells, and whether the disruption of this landscape can regulate the severity of T cell mediated GVHD remains unexplored.

Herein, we utilized Hi-C and RNA-seq to describe the 3D chromatin architecture of mature T cells *in vivo*, specifically at baseline (unstimulated, naive, pre-transplant) and following non-antigen stimulated lymphopenia induced proliferation (syngeneic) and allo-antigen driven (allogeneic) stimulation (Maeda et al., 2007). We generated T cell-specific WAPL-deficient mice and demonstrate that WAPL regulates the 3D chromatin structure, function, and *in vivo* biological response of T cells such as GVHD. These data collectively demonstrate that *a priori* disruption of chromatin structure regulates mature T cell function.

RESULTS

Characterization of mature naïve T cell genome architecture following *in vivo* stimulation

We first determined the genome architecture of mature naïve T cells at baseline and following *in vivo* lymphopenic and antigen-driven stimulation. To mimic clinically relevant *in vivo* stimulation, we compared naïve T cells before and after experimental syngeneic and allogeneic

transplantation. To this end, we utilized MHC-disparate B6 into a BALB/c model of transplantation (Reddy et al., 2008). CD62L[+] naive donor T cells were harvested from the splenocytes of B6 donors and transplanted into congenic syngeneic B6 or allogeneic BALB/c recipients (see Methods). Recipient animals were sacrificed 7 days after transplantation and their splenic T cells were isolated using congenic marker and analyzed for genomic architecture. We generated Hi-C contact maps to profile genome-wide chromatin interactions (Rao et al., 2014) for the harvested T cells. A higher incidence of long-range, inter-chromosomal interactions were observed in unstimulated T cells whereas syngeneic and allogeneic T cells had more pronounced intrachromosomal contact indicating mid- and short-range interactions (Figure 1A).

In Hi-C data, fewer interactions were captured as genomic distance between loci increased, a trend known as distance-dependent contact decay (Chen et al., 2015; Chen et al., 2016; Fudenberg et al., 2016; Fudenberg and Mirny, 2012; Lieberman-Aiden et al., 2009). However, strong long-range interactions were captured, and contact decay was used in evaluating their relative strength across the different groups. The magnitude of distance-dependent contact decay across the genome was distinct in the unstimulated naïve and lymphopenic/allo-antigen stimulated T cells. This was demonstrated by the patterns observed at the bottom left and top right corners of the intra-chromosomal regions along the diagonal of the Hi-C contact maps as shown in Figure 1A. Notably, these corner regions were significantly sparse in the allo-antigen stimulated T cells, suggesting that this may represent the strength of antigenic stimulation. The distal ends of the chromosomes in allo-antigen stimulated T cells did not interact frequently, in contrast to what was observed in unstimulated naïve and syngeneic (lymphopenic stimulated) T cells (Figure 1A).

We next performed element-wise subtraction between the Hi-C contact matrices to quantify the magnitude of dissimilarity between the measured contact frequencies in each pair of T cells at homeostasis (unstimulated), lymphopenic (syngeneic), or allo-antigen (allogeneic)

driven stimulation. The resulting subtraction matrices provided a comprehensive map of the quantitative interactional changes throughout the genome (Figure 1B) that provided greater clarity trends in magnitude and direction of change. Matrix subtraction between each genome-wide contact matrix confirmed the qualitative differences shown in Figure 1A. Matrix subtraction also demonstrated that unstimulated naïve T cells had more pairwise interactions than syngeneic and allogeneic cells T cells (Figure 1B). We found that about 94% of loci interactions had a higher contact frequency in unstimulated naïve T cells when compared with syngeneic and allogeneic T cells. Furthermore, this analysis supported that distal *intra*-chromosomal interactions were most highly enriched in unstimulated naïve T cells. By contrast, while distal *intra*-chromosomal interactions in syngeneic (lymphopenic) T cells were less enriched than in unstimulated naïveT cells, they were significantly more enriched than in the allo-antigen stimulated T cells (Figure 1B). Thus, unstimulated naïve, syngeneic, and allogeneic T cells demonstrate distinct and specific genome architectures.

Among all chromosomes, Chromosome 7 appeared to have the most significant dissimilarity between settings (Figure S3). Therefore, we took a closer look at the contact decay phenomenon for Chromosome 7. To this end, we extracted interacting loci from either half of the chromosome, represented by the red shaded area in Figure 1C. We characterized the dominant trends in interactions across the chromosome, highlighting with gray arcs only the highest contact frequency interactions (> 99.5th percentile) for clarity (Figure 1C). In the unstimulated naïve T cells, these interactions were mostly end-to-end. By contrast, interactions were mostly end-to-middle in syngeneic, and middle-to-middle in allogeneic T cells (Figure 1C). These data suggest that the varying degree of genomic distance-dependent interactions in T cells influence the observed chromatin folding patterns.

Generation of T cell conditional *WAPL* deficient mice

Because T cell genomic architecture changed following stimulation, we next assessed whether the changes are critical for T cell function. Cohesin promotes chromatin looping while WAPL is important for the release of cohesin from chromatin. Wapl is essential for embryonal development and its deficiency has been shown to cause defects in chromatin structure (Tedeschi et al., 2013). Specifically, WAPL has been shown to restrict chromosome loop extension (Busslinger et al., 2017; Haarhuis and Rowland, 2017; Haarhuis et al., 2017; Hill et al., 2020). Furthermore, previous studies demonstrated that WAPL is expressed in T cells and suggested that it may regulate T cell activation (Sun et al., 2013). Therefore, we determined whether WAPL regulates 3D chromatin architecture and the function of mature T cells. Because WAPL is critical for embryonal development we generated a T cell conditional WAPL knock-out (KO) mouse using CRISPR-Cas 9 and CD4-CRE systems (Ran et al., 2015; Tabebordbar et al., 2016). The *Wapl* locus in chromosome 14 qB has 18 exons and one non-coding exon (Figure S1A). We designed two sgRNAs specific to exon 2 of the *Wapl* gene to generate a double strand break in exon 2 (Figure 2A and Figure S1B). Our first line of mice carried an insert of two sgRNAs targeting exon 2 in *Wapl* (Figure 2A and Figure S1C). The second and third lines carried Rosa26-floxed STOP-Cas9 knockin on B6J (The Jackson Laboratory, Stock No:026175) or were CD4-CRE transgenic mice (The Jackson Laboratory, Stock No: 017336) (Figure S1C). Triple crosses were screened for sgRNA insert showing the positive 457 bp band (sgRNA-Wapl), (Figure 2B), CRE positive as a 100 bp band and loxP-SV40pA x3-loxP-CAS9-GFP (LSL) cleavage activity demonstrating 1123 bp for WT LSL, and 285 bp band for cleaved LSL after CRE recombination processing. The conditional KO mice developed normally. The WAPL KO T cells were verified for GFP expression (Figure 2B). The higher levels of LSL cleavage and GFP expression assured successful depletion of WAPL (Figure 2B, lane 4). We further confirmed WAPL protein depletion through Western blotting (Figure 2C) in T cells isolated from multiple knockout pups. Finally, to further confirm

efficient deletion we performed RNA-seq on the T cells sorted from these mice and compared them with littermate WT T cells, which demonstrated efficient loss of exon 2 in the *Wapl* gene (Figure 2D).

WAPL regulates T cell genome architecture

Because WAPL is known to regulate genome architecture and genome architecture entrains transcription (Kueng et al., 2006; Tedeschi et al., 2013; Wutz et al., 2017) we next evaluated the impact of WAPL deficiency on genome architecture and gene expression of unstimulated naïve T cells. To analyze whole genome architecture (structure) and gene expression (function), we integrated Hi-C and RNA-seq data (see Methods). We performed RNA-seq on T cells harvested from naïve B6 and those harvested on day+7 from transplanted syngeneic and allogeneic B6 recipients (see Methods). Integrating the RNA-seq data with the Hi-C data from above, we further analyzed qualitative and quantitative differences in the organization of the genome.

As shown in Figure 3A, unstimulated naïve T cells from the KO mice showed fewer interactions compared with WT T cells. Subtracting WT from the KO Hi-C contact matrix of the unstimulated naïve T cells yielded matrix differences, the overwhelming majority of which were negative (71%), indicating a decrease in contact frequencies and an overall weakening of long-range chromatin interactions in the absence of WAPL. Less than 25% of the shift in contact frequency was significant ($|\log_2 FC| \geq 1$ cutoff), however. Intriguingly, among the marked population of negative values (blue), Chromosome 11 had many positive values (red) in the difference matrix, indicating an increase in contact frequency in KO versus WT unstimulated T cells (Figure 3B, Table S1). These data suggest that the high frequency of interactions involving chromosome 11 may be a critical feature of the absence of WAPL in unstimulated naïve T cells.

We next explored the impact of WAPL deficiency on genome architecture following the stimulation of T cells in the context of lymphopenia (syngeneic T cells) and allo-antigen stimulation

(allogeneic T cells). Qualitatively, KO T cells had significantly greater interactions than WT T cells in both the syngeneic and allogeneic context (Figure 3A). These observations were quantified by matrix subtraction between KO and WT T cells in the context of syngeneic and allogeneic stimulation (Figure 3B). A significant proportion of the subtraction matrix elements were positive (27% in syngeneic and 45% in allogeneic) or zero (60% in syngeneic and 41% in allogeneic), demonstrating amplification of longer-range interactions and preservation of former interactions in the absence of WAPL, respectively (Figure 3B).

WAPL changes chromosome-level genome architecture

We next assessed changes at the chromosome level in the WT naïve unstimulated, stimulated lymphopenic (syngeneic), and allo-stimulated (allogeneic) T cells and compared the effect of WAPL deficiency in similar settings. The genome is hierarchically organized within the nucleus with the highest tier composed of chromosomes that occupy dedicated spaces around the nucleolus in an arrangement referred to as chromosome territories (Felsenfeld and Groudine, 2003; Boveri,T., 1909). These territories are important for transcriptional activity (Cremer and Cremer, 2019; Finn and Misteli, 2019; Rajapakse et al., 2009). Therefore, we next explored the structural and functional changes that occur within each chromosome following a loss of WAPL. To this end, we computed the Frobenius norm between pairs of matrices (Hi-C contact maps, denoting structure) and pairs of vectors (gene expression, denoting function) for each chromosome, gauging the level of variability between WT and KO data for each setting (Strang, G., 1993). We found that changes in genome architecture, as determined by the Frobenius norm of subtracted WT and KO Hi-C matrices, were minimal across each chromosome in unstimulated naïve T cells as compared with chromosomes in syngeneic and allogeneic T cells (Figure 3C). This conservation of structure is consistent with the notion that unstimulated naïve T cells are in a quiescent-like state and therefore less impacted by the loss of WAPL. However, in the context of lymphopenic or allo-antigen stimulation, changes in genome architecture were high across all

chromosomes, though gradually decreased with each chromosome in the allo-antigen stimulated setting (Figure 3C). Since we normalized for chromosome size, this decrease is not likely an artifact of our methodology. For changes in transcription, as determined by the Frobenius norm of subtracted WT and KO gene expression vectors, a few chromosomes exhibited notable changes. Specifically, Chromosome 13 in unstimulated naïve T cells, Chromosomes 4 and 14 in syngeneic T cells, and Chromosomes 4 and 11 in allogeneic T cells were most different in the absence of WAPL when compared to WT T cells (Figure 3C).

To further quantify differences between KO and WT T cells across the settings (naïve, syngeneic, and allogeneic), we utilized the Fiedler number. The Fiedler number is useful for measuring underlying connectivity in a genomic network, where a high Fiedler number indicates high topological stability (Chen et al., 2015). Here, we describe the Fiedler number on a per chromosome basis. We evaluated the correlation between structure and function in each chromosome by directly comparing connectivity (Fiedler number of chromosome-specific contact maps) and gene activity (chromosome-averaged gene expression) as shown in Figure 3D. In both WT and KO T cells, greater connectivity was demonstrated in the context of allo-antigen stimulation. By contrast, WT unstimulated naïve T cells demonstrated the lowest connectivity across all chromosomes yet exhibited the greatest increase in connectivity in the absence of WAPL. In both WT and KO T cells, across the three settings, Chromosome 13 was the most highly connected chromosome (Figure 3D).

Functional analysis of unstimulated T cells exhibited higher levels of gene expression only in chromosomes 3 and 13 whereas expression was notably higher in chromosomes 4, 7, 9, 12, 14 and 17 of syngeneic T cells and similar across the three settings in remaining chromosomes (Figure 3D). WT and KO T cells exhibited little difference in function except in unstimulated Chromosome 13, syngeneic chromosomes 4 and 14, and allogeneic chromosomes 4 and 11. Of note, unstimulated naïve, syngeneic, and allogeneic settings were most functionally dissimilar in

chromosomes 13 and 14 in WT and KO T cells (Figure 3D). Overall, these data show that there is a dissonance between the structural and functional trends at the chromosome level, as demonstrated by large changes in structure that do not yield comparable changes in function; or similarly, high connectivity (indicative of more active and accessible regions) accompanied by lower expression (Figure 3C and D).

Analysis of chromosome territory partitioning into A and B Compartments

We explored the next tier in the hierarchical organization of the genome - bi-partitioning of chromosome territories into individual stretches of accessible (active) and inaccessible (inactive) chromatin, termed A (euchromatin) and B (heterochromatin) compartments, respectively (Lieberman-Aiden et al., 2009). We demarcated these regions using the signed values of the Fiedler vector which measures underlying chromatin accessibility (Chen et al., 2015). The positive values of the Fiedler vector reflect compartment A and negative values reflect compartment B. We observed differential compartmentalization among all three settings of the WT T cells (Figure 3E). Specifically, we observed 274 genomic bins (100kb-length) with different chromatin states between unstimulated naïve and syngeneic WT T cells (Figure 3E). These bins include 295 genes that occupied distinct chromatin states between unstimulated naïve and syngeneic T cells. Some bins spanned multiple coding regions while others spanned only non-coding regions, explaining the inconsistency in the number of genes and corresponding genomic bins. By contrast, there were 645 bins containing 769 genes with different chromatin states between WT unstimulated naïve and allogeneic T cells (Figure 3E). However, there were 645 bins containing 714 genes with distinct chromatin states between WT syngeneic and allogeneic T cells (Figure 3E). Although the WT naïve/allogeneic and WT syngeneic/allogeneic comparisons had the same cumulative number of compartment differences genome-wide, the sets of differentially compartmentalized loci were not identical between the two pairs. Additionally, no single pairing had consistently more compartment differences per chromosome than the others (Figure S2A).

We also analyzed A/B compartment switch events mediated by the loss of WAPL by comparing WT and KO T cells (Figure 3F). We observed 184 genomic bins (100kb-length) containing 217 genes in the unstimulated naïve KO T cells that switched compartments when compared to WT T cells (Figure 3F). The switch events demonstrated a bias from compartment B to compartment A (70.7%) (Figure 3F). In the context of lymphopenic stimulation (syngeneic), KO T cells demonstrated 303 switch events involving 375 genes when compared to WT T cells. In the allogeneic context, KO T cells demonstrated 413 switch events involving 485 genes when compared to WT T cells (Figure 3F). The switch bias was once again in the direction of compartment B to compartment A in syngeneic T cells (60.4%) and in allogeneic T cells (53.8%). Additionally, allogeneic T cells tended to have more switch events per chromosome than unstimulated naïve and syngeneic T cells (Figure S2B). Interestingly, switched compartment loci tended to congregate towards the ends of chromosomes rather than in the middle or spread evenly throughout. No switch events were observed among contiguous regions of the genome in any of these settings. When quantified at the chromosome level, chromatin compartmentalization remained stable between unstimulated and stimulated T-cells as well as between WT and KO T cells (Tables S2 and S3).

WAPL impacts TADs and local gene transcription

Stability in global compartment organization throughout the genome did not preclude changes to local genome organization. In the next tier of genome organization, chromatin preferentially interacts within locally-distributed and insulated regions called topologically associating domains (TADs) that regulate transcription (Dixon et al., 2012). Thus, we next analyzed the impact of changes in TADs. One highly supported mechanism of TAD formation is loop extrusion (Fudenberg et al., 2016) mediated by the ring-shaped cohesin complex and WAPL, which enables TAD dynamics by promoting cohesin turnover (Haarhuis et al., 2017). Specifically, genes residing in the same TAD experience coordinated regulation and expression (Dixon et al., 2012; Nora et

al., 2012; Shen et al., 2012; Le Dily et al., 2014) while changes to TAD boundaries influence anomalous gene-enhancer interactions (Flavahan et al., 2016; Lupianez et al., 2015). We therefore utilized spectral graph theory to identify the positional boundaries that define TADs (Chen et al., 2016). We found that unstimulated T cells showed a 1.6% increase in the total number of TADs in KO T cells compared to WT T cells, a 3.2% increase for syngeneic T cells, and by contrast, a 0.5% decrease for allogeneic T cells (Figure S3A). Overall, there was surprisingly little fluctuation to the number of TADs following WAPL knockout, suggesting that the structural dissimilarity present between WT and KO T cells was likely a product of TAD rearrangement and shifting of domain boundaries rather than domain deletions or insertions. Further, TAD organization throughout the genome was not well conserved across settings. Out of 1,668 and 1,608 genome-wide TADs, respectively, WT unstimulated naïve and syngeneic T cells had only 778 boundaries in common. Out of 1,668 and 1,736 genome-wide TADs, respectively, WT unstimulated naïve and allogeneic T cells had 710 boundaries in common. Finally, out of 1,608 and 1,736 genome-wide TADs, respectively, WT syngeneic and allogeneic T cells had 609 boundaries in common.

We next utilized the variation of information (VI) metric to evaluate TAD boundary similarity (see Methods). Chromosome 7 demonstrated the most differential gene expression (163 in unstimulated naïve, 211 in syngeneic, and 123 in allogeneic) between WT and KO T cells amongst all chromosomes (Figure S3B). In contrast to differential gene expression, chromosome 11 contained the most dissimilar TAD boundaries between WT and KO in unstimulated and syngeneic subsets, with a 20% and 38% difference, respectively (Figure S3B). However, in the allogeneic setting, the TADs on chromosome 11 were not the most dissimilar between WT and KO T cells (Figure S3B). In the context of allogeneic T cells, maximal dissimilarity in TAD structure was found on Chromosome 17, with a 40% difference (Figure S3B). We therefore focused our

TAD analysis on Chromosome 7 (Figure 4) and chromosome 11 (Figure S4) because of the most prominent changes between the various groups.

As shown previously in Figures 1 and 3, KO T cells in the unstimulated naïve setting demonstrated greater contact depletion towards the distal ends of the chromosomes and an overall decrease in *intra*-chromosomal interactions when compared to WT T cells. By contrast, KO T cells in the syngeneic and allogeneic settings maintained much of their sparsity in distal end interactions and demonstrated an increase in their *intra*-chromosomal interactions when compared to WT T cells. We further investigated enrichment in the inter-regional contacts on Chromosome 7 (Figure 4) and chromosome 11 (Figure S4). A unique feature emerged in the context of syngeneic and allogeneic T cells in the absence of WAPL, the appearance of 'corner peaks' (Figure 4A-C). Corner peaks are the enrichment in contacts seen at the bottom left and top right corners of TADs (Rao et al., 2014). We observed more abundant corner peaks in the KO syngeneic and allogeneic T cells (Figures 4B and C), consistent with the known activity of WAPL to limit loop extension (Haarhuis and Rowland, 2017; Haarhuis et al., 2017). Further, several studies have demonstrated that corner peaks are associated with a longer residence time of the cohesin complex at TAD borders (Haarhuis et al., 2017; Schwarzer et al., 2017; Szabo et al., 2019). We did not, however, observe this rise in corner peak abundance in the KO unstimulated naïve T cells (Figure 4A), suggesting that the development of corner peaks in the T cells might be a consequence of their proliferation following their activation.

Impact on the Cell Cycle Gene Network

WAPL is critical for sister chromatid cohesion and loop extrusion dynamics. These processes correlate with cellular proliferation and cell cycling, and because we observed changes in corner peaks of Chromosome 7 only in the context of syngeneic and allogeneic settings in the absence of WAPL, we next explored changes in cell cycle genes and the TADs that these genes reside in (Figures 4D-F). One particular gene-rich region in Chromosome 7, extending between positions

60Mb and 70Mb on the chromosome, exhibited considerable TAD reorganization in the KO T cells when compared to WT unstimulated (Figure 4D), syngeneic (Figure 4E), and allogeneic (Figure 4F) T cells. This region contains cell cycle genes (Dolatabadi et al., 2017; Inaba et al., 2018): *Fanci, Prc1,* and *Blm*. Unstimulated naïve T cells in the absence of WAPL gained two additional TAD borders in this region, eliminating the TAD co-occupancy of the cell cycle genes (Figure 4D). But these cell cycle genes (*Fanci, Prc1,* and *Blm*) were not differentially expressed, as they did not meet the cutoff of $|log_2FC| \geq 1$. However, there were seven non-cell cycle genes (*Mesp2, Arpin, Fes, Homer2, Saxo2, Cemip,* and *Arnt2*) directly upstream and downstream of the cell cycle genes in this region that were differentially expressed. These data suggest that in the unstimulated KO T cells, these three cell cycle genes on chromosome 7 were not differentially expressed despite the changes in TAD borders when compared to WT T cells. In the syngeneic context, both the WT and KO T cells merely demonstrated a minor shift in the TAD, maintaining the co-occupancy of the three cell cycle genes (Figure 4E). Similar to the unstimulated context, in the syngeneic T cells, there was a differential expression of non-cell cycle genes upstream and downstream of cell cycle genes in this region (*Agbl1, Isg20, Can, Hapln3, Ribp1, Mesp2, Anpep, Fes, Slc28a1, Homer2, Adamtsl3,* and *Tmc3*). In the allogeneic T cells, the TAD containing the three cell cycle genes gained an additional border in the absence of WAPL, splitting the genes into separate TADs but nonetheless still not affecting a change in the three cell cycle genes (Figure 4F). However, non-cell cycle genes (*Agbl1, Can, Rhcg,* and *Adamtsl3)* once again were differentially expressed. Thus, in the absence of WAPL, the highlighted region of Chromosome 7 containing the three cell cycle genes did not dynamically change but demonstrated significant changes in the expression and structural rearrangement of genes in their vicinity. No other cell cycle genes on Chromosome 7 demonstrated differential expression.

Because we did not see a change in the three cell cycle genes noted above in the analysis of Chromosome 7, yet WAPL is known to regulate cell cycling, we next explored the entire breadth

of cell cycle genes throughout the genome. To this end, we extracted and stitched together 141 Hi-C genomic bins that corresponded to a curated set of 170 cell cycle genes genome-wide (see Methods). We generated a Hi-C-derived 5C contact matrix to analyze the genome-wide cell cycle gene network. 5C, like Hi-C, is a derivative of the original chromosome conformation capture technique (Dekker et al., 2002), and is useful in identifying interactions among select genomic regions that bear relationship to one another (Dostie et al., 2006). In unstimulated naïve T cells, we observed that the connectivity of the cell cycle network decreased in the absence of WAPL, as determined by the element-wise Pearson correlations moving toward zero (Figure 5A). In syngeneic and allogeneic T cells, however, connectivity appeared to strengthen in the absence of WAPL, as demonstrated by the correlation tending towards ± 1 (Figure 5A). The connectivity (structure) of cell cycle genes changed between the WT and KO T cells in all settings, but the maximal change was noted in the syngeneic setting (Figure 5B). However, changes in the structure of the cell cycle network did not trend with the function (expression). The expression (function) of the cell cycle gene network changed the most between WT and KO T cells in the context of allogeneic setting (Figure 5B). The genome-wide structural analysis of the cell cycle gene network highlighted two subgroups of highly connected genes (shown in purple and green boxes in Figure 5A). One of these subgroups negatively correlated with most of the network (Figure 5A, purple box) while the other demonstrated positive correlation with most of the network (Figure 5A, green box) in all three settings. The measure of inter-connectivity within these two subgroups, in terms of degree centrality, is shown in Figure 5C. While gene expression of cell cycle genes *Fanci, Prc1,* and *Blm*, did not change in the absence of WAPL, other genes in the network did change significantly (Table S4). Five cell cycle genes were up-regulated in the absence of WAPL in unstimulated naive T cells. By contrast, 12 cell cycle genes were down-regulated and 4 up-regulated between WT and KO T cells in the syngeneic context. In the context of allo-antigen stimulation, only one gene was down-regulated.

WAPL induced changes in genome structure alter T cell gene expression and functions

We next determined whether the changes in the chromatin architecture caused by WAPL in T cells affected genome function. The 3D genome structural changes have been suggested to affect T cell development (Seitan et al., 2011). But whether *a priori* disruption of the cohesin ring affected T cell functions is not clear. The transgenic mice with conditional WAPL KO in T cells displayed normal birth and growth rates and generated enough mature naïve T cells. But to better analyze the developmental impact of WAPL deficiency on T cell development we next analyzed the thymii from 8-10 week old WAPL KO and WT littermates. The thymocytes from the WAPL KO animals showed significant changes when compared to the WT littermate controls. Specifically, they showed reduction in the total numbers of thymocytes, double positive (DP) and CD8 SP thymocytes (Figure 6A, Figure S5A). We next analyzed the secondary lymphoid organ (spleens) for peripheral T cell subsets. The WAPL KO demonstrated lower numbers of total T cells in the spleen (Figure 6B) and lower numbers of $CD3^+CD4^+$ and $CD3^+CD8^+$ T cells (Figure S5B and C) when compared to WT littermates. During the β-selection checkpoint in thymus, the β chain of the T cell receptor rearranged by the thymocytes must retain the structural properties allowing it to be presented on the surface of the thymocytes. We therefore assessed for $TCRβ^+$ $CD4^+$ and $CD8^+$T cells in the spleen and found that the WAPL KO mice demonstrated lower numbers of $TCRβ^+$ $CD4^+$ and $CD8^+$ T cells in the spleen when compared to WT littermates (Figure S5D and E). The splenocytes from the KO animals also demonstrated lower numbers of $CD3^+CD69^+$, $CD4^+CD69^+$, $CD8^+CD69^+$ and $CD4^+CD25^+$ T cells (Figure S6A-D).

Because WAPL deficiency altered chromatin architecture, we next analyzed whether these structural changes were associated with changes in gene expression and proliferation, at baseline and following *in vivo* stimulation. To this end, we first determined whether WAPL expression itself changed in T cells following T cells stimulation. Consistent with our previous observation, we observed significant upregulation of WAPL protein in T cells upon co-culture with allogeneic DCs

for 60 hours (Figure 6C) (Sun et al., 2019; Sun et al., 2015). To determine the impact of this structure-function relationship, we performed gene set enrichment analysis (GSEA) for the cell cycle gene network highlighted in our earlier Hi-C-derived 5C analyses (Figure 5). We found that cell cycle genes were differentially regulated between WT and WAPL-deficient T cells (Figure 6D). The differential expressions in genes with enrichment scores at top 2 and bottom 2 were verified by real-time quantitative PCR. *Gas2l1*, a gene induced upon growth arrest (Goriounov et al., 2003), and *Mcm3* that increases as cells progress from G0 into the G1/S phase and regulate cell cycle (Tsuruga et al., 1997), were significantly upregulated in WAPL-deficient T cells compared to WT T cells (Figure 6E and F). By contrast, *Sphk1*, a gene which showed the key role in TNF-a signaling and the canonical NF-kappa-B activation pathway important in inflammatory, apoptotic, and immune processes (Alvarez et al., 2010), and *Myh10*, a gene which was required for completion of cell division during cytokinesis (Kim et al., 2005; Straight et al., 2003) were downregulated in WAPL-deficient T cells (Figure 6G and -H). These observations demonstrated that changes in genome structure caused by WAPL deficiency in T cells promote differential gene expression that suppresses apoptosis, cell proliferation and cell cycle progress.

To further investigate the impact of WAPL deficiency on T cell cellular responses, we stimulated WAPL-deficient and WT T cells *in vitro*, either with allo-stimulation in mixed lymphocytes reaction (MLR) by co-culturing T cells with allogeneic DCs for 4 days or with CD3/CD28Ab for 2 or 3 days and pulsed with H^3-TdR. The WT T cells, demonstrated significant greater H^3-TdR incorporation when compared with WAPL KO T cells suggesting that WAPL deficiency caused reduced proliferation (Figure 6I and Figure S6E). T cell proliferation was directly assessed *in vitro* by dye dilution assay. Specifically, we utilized CellTrace™ FarRed dilution to avoid the interference with GFP fluorescence in WAPL KO T cells. The WAPL KO T cells demonstrated reduced dye dilutions when compared to WT T upon stimulation by either allogeneic DCs or CD3/CD28ab (Figure 6J and Figure S6F). It is possible that cell death from apoptosis could contribute to the reduction in

T cell expansion caused by WAPL deficiency. We therefore also determined apoptosis after stimulation of both WT and WAPL KO T cells. The WAPL deficient T cells also showed significantly decreased apoptosis when compared to littermate WT T cells (Figure 6K). These data demonstrate that absence of WAPL altered T cell proliferation and apoptosis following *in vitro* stimulation.

The interaction between cohesin and WAPL plays an essential role in maintaining chromosome structure and separation of sister chromatids during mitosis and cell proliferation (Peters and Nishiyama, 2012). Given the impact on cell cycling gene expression from the change in genomic architecture in the absence of WAPL, and the reduction in proliferation, we next examined cellular mechanisms underlying the reduced proliferative capacity of WAPL deficient T cells following *in vitro* and *in vivo* stimulation. Specifically, we explored the hypothesis that altered gene expression of cell cycling genes from the change in genomic architecture caused by WAPL deficiency leads to a reduction in cell cycling. To assess kinetic cell cycling we utilized flow cytometry analyses of DNA content with FxCycle™ Far Red Stain to avoid interference from the GFP fluorescence in WAPL KO T cells. Purified WT or WAPL KO T cells were stimulated with allogeneic DCs or CD3/CD28T cell activator dynabeads and analyzed for 2C and >2C populations. *WAPL* KO T cells demonstrated significantly lower 2C percentages but higher >2C percentages when compared with WT T cells at several time points after allo-DC or CD3/CD28 stimulation (Figure 6L and M and Figure S 7A and B). We next analyzed whether the effect on cell cycling was observed *in vivo*. To this end we once again utilized the allo-BMT model system, the host splenocytes were harvested and analyzed for WT or WAPL KO donor T cells as above. As shown in Figure 6N and Figure S7C, the *in vivo* stimulation of WT and WAPL KO T cells demonstrated similar differences as from in vitro stimulation. These data indicated that Wapl KO T cells had cell cycle deficiency, which impairs T cell proliferation and alters T cell function *in vitro* and *in*

vivo, consistent with cell cycle gene analyses by Hi-C derived 5C (Figure 5), GSEA and qPCR analyses (Figure 6D-H).

WAPL deficiency in T cells improves survival after allogeneic BMT

Mature T cells in the allogeneic donor T cells are the principal mediators of GVHD, a major cause of mortality after allogeneic BMT. Because WAPL regulated mature T cell responses following *in vitro* and *in vivo* allo-stimulation we next determined whether this has a clinical impact on GVHD severity following experimental allo-BMT. To this end we once again utilized the MHC mismatched B6 (H2) → BALB/c (H2^d) mouse model, the congenic B6 animals served as the syngeneic controls. BALB/c recipient mice were lethally irradiated (800 cGy total body irradiation, split dose) and transplanted with T cell depleted WT BM from B6 donors along with purified mature T cells from the spleen of either WT or WAPL KO B6 animals (Sun et al., 2015). The recipient animals were monitored for survival and GVHD severity as described in Methods. We first determined whether WAPL expression changed after allo-BMT. Consistent with *in vitro* allo-stimulation as in Figure 6C, WT cells demonstrated higher expression of WAPL protein in donor T cells harvested from recipient spleens 7 days after allogeneic BMT when compared to syngeneic controls (Figure 7A). Survival analysis demonstrated that all the syngeneic animals survived, but the allogeneic animals that received WT T cells died with signs of severe clinical GVHD (Figure 7B). By contrast, allogeneic animals that received WAPL deficient T cells showed significantly improved survival (53% versus 17%; P < 0.01) (Figure 7B) and reduced clinical severity of GVHD (P < 0.01) (Figure 7C). We confirmed the reduction in GVHD with detailed histopathological analyses of GVHD target organs, the liver, GI tract and skin. As shown in Figure 7D, allogeneic animals that received T cells from WAPL deficient donors had significantly reduced histopathological GVHD in the liver (P<0.01), gastrointestinal tract (SI and LI) (P < 0.05) and skin (P < 0.05) on day +21 after BMT. Consistent with reduced mortality, the recipients of allogeneic WAPL KO T cells demonstrated

reduced serum levels of proinflammatory cytokines such as IFNγ and TNFα in sera when compared with WT T cell recipients (Figure 7E and F).

The allogeneic animals that received WT or KO T cells demonstrated >98% donor engraftment on day 21, ruling out mixed chimerism or engraftment as a cause for reduction in GVHD. Furthermore, consistent with above results, WAPL KO T cells showed significantly reduced expansion when compared with WT donor T cells (Figure 7G) in the recipient spleens harvested 7 days after BMT suggesting that the reduction in GVHD was consequence of reduced expansion of allo-T cells. Consistent with this, the allogeneic WAPL KO T cells demonstrated fewer absolute numbers of CD3$^+$CD69$^+$ cells (Figure 7H), and regulatory T cells (despite a higher percent) when compared to WT T cells (Figure 7I and J).

DISCUSSION

The cohesin complex and its regulators such as WAPL are critical determinants of 3D chromatin structure, which regulates replication, repair, and transcriptional processes (Cuadrado and Losada, 2020; Misteli, 2020; Nasmyth and Haering, 2009; Remeseiro et al., 2013). Given the paucity of data on genomic organization of mature T cells and its impact on T cell functions *in vivo,* we describe the genomic architecture of mature T cells following *in vivo* lymphopenic and allogeneic stimulation.

We find that mature naïve T cells show long range, interchromosomal interactions, throughout the genome at a baseline unstimulated state. Upon *in vivo* stimulation T cells showed pronounced medium and short range, *intra*-chromosomal interactions, with a change in distance dependent decay. The most pronounced changes in the chromatin architecture were observed in Chromosome 7, in the regions containing cell cycling genes, which correlated with the degree of proliferation and cell cycling following *in vivo* stimulation (Inaba et al., 2018). Thus, the changes in the 3D structure correlated with cellular functional states of mature T cells. Prior to availability

of Hi-C methods, explored structural features of the T cell genome following *in vitro* stimulation (Kim et al., 2004). Chromosome 6 was explored in its entirety, while the global spatial characterization was limited to chromosome territories. Here we define the entire genomic architecture on mature T cells following *in vivo* activation, including TADs and loops. We further connect those changes to genome wide functional changes in gene expression and T cell responses. Hu et al. demonstrated a key role for BCL11B in the development of T cells and associated 3D nucleome changes (Hu et al., 2018), but did not characterize the relevance of structural integrity to mature T cell functions. Our data focus on the genomic landscape of mature T cells. We define a mechanistic role for genomic architecture in regulation of gene expression, cellular function, biological responses *in vivo* that impact clinically relevant disease states such as GVHD. We determine that 3D architectural changes occur as a consequence of the cellular state of activation, and that its disruption by altering the function of cohesin complex, regulates mature T cell functions in response to allo-stimulation.

We further explored the mechanistic relevance of genomic structure to its function (gene expression) and to the cellular functions and biological impact of T mature cells by deleting WAPL, a key regulator of genomic structure. WAPL deletion led to reduction of long-range interactions in the baseline unstimulated state of naïve T cells. However, following lymphopenic (syngeneic) and antigen (allogeneic) activation there was a greater amplification of longer-range interactions following WAPL deletion. This is consistent with a previous report of on WAPL-deficient cell lines cells (Haarhuis et al., 2017) and with the notion of extruded DNA loops beyond CTCF barriers (Allahyar et al., 2018). These data suggest that WAPL alters genomic architecture, at baseline and after activation of T cells. The in vivo role for WAPL mediated regulation of chromatin structure is crucial for embryonal development (Tedeschi et al., 2013). We now extend these studies and demonstrate that WAPL also regulates *in vivo* immune responses mediated by T cells. T cell specific WAPL-deficient animals developed normally despite the T cells showing genomic

structural changes at homeostasis. Upon stimulation, WAPL deficient T cells demonstrated mitosis defects and more axial structures in the interphase suggesting that they exit mitosis with less intact cohesin. However, it is important to note that the absence of WAPL neither caused a complete loss of development of T cells nor a total shutdown of mature T cell proliferation. The T cells developed and proliferated in the absence of WAPL, albeit at a much lower level, the reasons for which remain unclear. One possible explanation is that separation of chromatids during mitosis in T cells may be independent, or only partially dependent, on WAPL, based on their strength and duration of stimulation/activation (Oliveira and Nasmyth, 2013; Srinivasan et al., 2019; Yuen and Gerton, 2018). This notion is consistent with the observation that several cohesinopathies in humans that are caused by mutations in various components of cohesin complex and yet do not appear to cause T cell defects (Piche et al., 2019; Remeseiro et al., 2013; Singh and Gerton, 2015). Future studies may determine the role of WAPL and disruption of cohesin in thymopoiesis and in regulation of T cells in secondary lymphoid organs.

The structural changes in our study reflect a polyclonal response from a combination of alloreactive/ lymphopenia induced proliferating cells, and the non-proliferating mature T cells. The T cells therefore might be in various stages of early/mid G1, S, G2, M phases. Thus, the genome contact frequency and associated structural changes are reflective of the T cells in these various stages after stimulation. T cells from naïve, syngeneic and allogeneic settings demonstrated significant changes in TAD and A/B compartmentalization, which were consistent despite the polyclonal nature of the T cell subsets. Nonetheless, while the genome architectures were significantly different, whether these are the direct cause or a consequence of proliferation differences cannot be definitively ascertained. Furthermore, whether antigen specific T cell responses vary based on the type of antigen cannot be determined from our study. However, our assessment of polyclonal responses is akin to biological and clinically relevant situations as allogeneic transplantation.

Our data collectively demonstrate for the first time to our knowledge that altering genomic structure *a priori*, at baseline, regulates T cell gene and cellular functions. Mechanistically the data show that WAPL alters chromatin architecture at cell cycling genes and thus links genomic structure and function. However, because centromeric and telomeric regions are highly repetitive and are typically unmappable, we filtered these regions in our Hi-C analysis (Lajoie et al., 2015; Ried and Rajapakse, 2017; Servant et al., 2015). In addition, features at the TAD and sub-TAD levels that are observed via bulk Hi-C are a population average, and could be variable and distinct between various T cell subsets that develop and differentiate after stimulation. Future studies that integrate single cell Hi-C and single-cell RNA-seq may refine these observations in specific T cell subsets. Regardless, our data demonstrate that disruption of 3D chromatin architecture by WAPL can directly regulate gene expression and cellular function of T cells in a physiologically and clinically relevant disease context.

There are no reported cases of isolated germline WAPL mutations in humans (Gard et al., 2009; Piche et al., 2019; Remeseiro et al., 2013; Singh and Gerton, 2015). This is likely because of its critical role during embryonic development. However, somatic mutations in WAPL have been linked to epithelial carcinogenesis (Waldman, 2020). Our study demonstrates that WAPL deficiency in T cells did not cause a profound defect in development of T cells, nor cause T cell malignancy. Thus, WAPL may play a nuanced role in different cell subtypes, depending on their developmental stage, context, and stimulation. Future studies will need to carefully assess the biological implications of WAPL deficiency on T cell subsets and other immune cells. Because WAPL deficient T cells caused reduction in GVHD, our data suggest that targeting it uniquely in T cells might be a novel strategy to mitigate immunopathologies such as GVHD, allograft rejection or autoimmunity. At a broader level, our data provide a proof of concept for the notion that targeting 3D genomic architecture may be a therapeutic strategy that can be potentially harnessed for directly modulating *in vivo* disease processes.

METHODS

Mice

B6 (H2^b, CD45.1 and CD45.2), BALB/c (H2^d, CD45.1 and CD45.2), Rosa26-floxed STOP-Cas9 knockin (B6J) and CD4CRE (B6) mice were purchased from The Jackson Laboratory and the National Cancer Institute. The ages of the mice used for experiments ranged between 8 and 12 weeks. Mice were housed in sterilized microisolator cages and received filtered water and normal chow.

Generating conditional WAPL KO T cells

We designed 2 sgRNAs targeting exon2 of *Wapl* locus which are localized in mouse Chromosome 14 qB (Figure S1) (Ran, et al.,2015). Three lines of mice with B6 background were used to generate conditional WAPL KO mice in T cells. The first line was generated to carry CRISPR-sgRNAs targeting insert to identify exon 2 in *Wapl* (Figure 2A). The second and third lines are Rosa26-floxed STOP-Cas9-GFP knockin on B6J (The Jackson Laboratory, Stock No:026175) and CD4-CRE transgenic mice (The Jackson Laboratory, Stock No: 017336). Tail DNA was screened for sgRNAs-Wapl insert positive, CRE positive and stop signal in loxP-SV40pA x3-loxP cleavage or deletion. The potential WAPL KO T cells were sorted for GFP positive (Figure 2B). To confirm conditional WAPL KO in T cells, T cells were processed for SDS-PAGE and detected with anti-Wapl antibody. And further, total RNAs were isolated from WAPL KO and WT T cells and processed for RNAseq.

Genotyping primers

   1) CRISPR-sgRNAs targeting insert, 457 base pairs

Forward 5'-CCGTAAGATGAACCCTTACCC-3'

Reverse 5'-TCTGAGTAAGATCGGTGTTTCG-3'

   2)   loxP-SV40pA x3-loxP-CAS9-GFP (LSL) cleavage activity (1123 base pairs for WT LSL, 285 base pairs for cleaved LSL after CRE recombination processing)

Forward 5'-GCAACGTGCTGGTTATTGTG-3'

Reverse 5'-TAGTCTCCGTCGTGGTCCTT-3'

   3) CRE recombinase 100 base pairs

Forward 5'-GCG GTC TGG CAG TAA AAA CTA TC-3'

Reverse 5'-GTG AAA CAG CAT TGC TGT CAC TT-3'

DC isolation and purification

Dendritic cells (DC) were isolated from splenocytes either from WT B6 or BALB/c mice. Single cell suspensions were prepared according to the manufacturer's instruction then subjected to CD11c microbead (MACS) staining and positive selection using LS column (Miltenyi Biotec). The purity of enriched CD11c$^+$ DC preparation was 85.6~90%.

T cell isolation and purification, and mixed lymphocyte reaction (MLR)

WT and WAPL KO B6 T cells were isolated by negative selection (>95% purity) (Pan T Cell Isolation Kit II; Miltenyi Biotec). T cells were co-cultured with B6 WT or BALB/c DCs at a ratio of 40:1 (T cells versus DCs 3×10^5:7.5×10^3) for 96 hours using 96-well flat-bottomed plates (Falcon Labware), or stimulated with Dynabeads T cell activator CD3/CD28 (25 ul/10^6/ml) for 2 or 3 days respectively. Proliferation was determined by incubating the cells with H^3-thymidine

(1 Ci/well [0.037 MBq]) for the last 20 or 6 hours respectively. H^3-thymidine incorporation in T cells was counted on a 1205 Betaplate reader (Wallac, Turku, Finland).

BMT and systemic analyses of GVHD

BMTs were performed as described previously (Sun et al. 2015). The donor T cells (WT B6 or WAPL deficiency) were isolated from spleens and purified by negative selection (Pan T cell Isolation Kit II; Miltenyi Biotec). Bone marrows cells from tibia and fibula were harvested and TCD (T cell deletion) BM cells were isolated with positive deletion using anti-CD90.2 microbeads and LS column (Miltenyi Biotec). The recipient BALB/c mice received an 800-cGy total body irradiation on day -1 (split dose) and T cells (1×10^6, either B6 WT or WAPL deficiency T cells, and TCD BM cells (5×10^6, from WT B6 mice) were injected intravenously into the recipients on day 0. The syngeneic B6 control mice received a 1000-cGy total body irradiation on day -1. T cells (2×10^6,

isolated from B6 WT mice) and TCD BM cells (5×10^6 from B6 WT mice) were injected intravenously into the recipients on day 0. Mice were housed in sterilized microisolator cages and received normal chow and autoclaved hyperchlorinated drinking water for the first 3 weeks after BMT. Survival was monitored daily. The degree of systemic GVHD was assessed by a standard scoring system with five criteria scores: percentage of weight change, posture, activity, fur texture, and skin integrity, and graded from 0 to 2 for each criterion (maximum index =10) (Sun et al JCI 2015). Acute GVHD was also assessed by histopathologic analysis of the ileum and the ascending colon, liver, and ear skin. Specimens were harvested from animals on day 21 after BMT, then processed and stained with hematoxylin and eosin. Coded slides were examined systematically in a blinded manner by using a semiquantitative scoring system to assess the following abnormalities known to be associated with GVHD, small intestine: villous blunting, crypt regeneration, loss of enterocyte brush border, luminal sloughing of cellular debris, crypt cell apoptosis, outright crypt destruction, and lamina propria lymphocytic infiltrate; colon: crypt regeneration, surface colonocytes, colonocyte vacuolization, surface colonocyte attenuation, crypt cell apoptosis, outright crypt destruction, and lamina propria lymphocytic infiltrate; liver: portal tract expansion, neutrophil infiltrate, mononuclear cell infiltrate, nuclear pleomorphism, intraluminal epithelial cells, endothelialitis, hepatocellular damage, acidophilic bodies, mitotic figures, neutrophil accumulations, macrophage aggregates, macrocytosis; skin: apoptosis in epidermal basal layer or lower malpighian layer or outer root sheath of hair follicle or acrosyringium, lichenoid inflammation, vacuolar change, lymphocytic satellitosis. The scoring system denoted 0 as normal, 0.5 as focal and rare, 1.0 as focal and mild, 2.0 as diffuse and mild, 3.0 as diffuse and moderate, and 4.0 as diffuse and severe. Scores were added to provide a total score for each specimen (Sun et al, 2015).

Immunoblotting

T cell lysates were extracted as previous described (Sun et al, 2015), and 50 to 100 μg of protein extract was separated in SDS-PAGE and transferred onto a PVDF membrane (GE Healthcare).

The membrane was blocked with 5% nonfat milk for 30 minutes and then incubated overnight at 4°C with the following Abs in 5% nonfat milk: anti-Wapl rabbit polyclonal Ab (1:500 in nonfat milk, Proteintech Cat 16370-1-AP), anti–β-actin mouse mAb (1:1,000 in 5% nonfat milk; Abcam, catalog ab8226). After washing 3 times with TBST for 5 minutes, the blot was incubated with specific HRP-labeled secondary Ab, washed again with TBST, and signals generated and visualized using the Enhanced Chemiluminescence Kit (Thermo Scientific, Cat 32106).

Generation of Hi-C libraries for Sequencing

The *in situ* Hi-C protocols from Rao et al. (Rao et al., 2014) were adapted with slight modifications. For each Hi-C library, approximately 1 x 10^6 cells were incubated in 250 µl of ice-cold Hi-C lysis buffer (10mM Tris-HCl pH8.0, 10mM NaCl, 0.2% Igepal CA630) with 50µl of protease inhibitors (Sigma) on ice for 30 minutes and washed with 250 µl lysis buffer. The nuclei were pelleted by centrifugation at 2500x*g* for 5 minutes at 4°C, re-suspended in 50 µl of 0.5% sodium dodecyl sulfate (SDS) and incubated at 62°C for 10 minutes. Afterwards 145 µl of water and 25 µl of 10% Triton X-100 (Sigma) were added and incubated at 37°C for 15 minutes.

Chromatin was digested with 200 units of restriction enzyme MboI (NEB) overnight at 37°C with rotation. Chromatin end overhangs were filled in and marked with biotin-14-dATP (Thermo Fisher Scientific) by adding the following components to the reaction: 37.5 µl of 0.4mM biotin-14-dATP (Life Technologies), 1.5 µl of 10mM dCTP, 1.5 µl of 10mM dGTP, 1.5 µl of 10mM dTTP, and 8 µl of 5U/µl DNA Polymerase I, Large (Klenow) Fragment (NEB). The marked chromatin ends were ligated by adding 900 µl of ligation master mix consisting of 663 µl of water, 120 µl of 10X NEB T4 DNA ligase buffer (NEB), 100 µl of 10% Triton X-100, 12 µl of 10mg/ml BSA, 5 µl of 400 U/µl T4 DNA Ligase (NEB), and incubated at room temperature for 4 hours.

Chromatin de-crosslinking was performed by adding 50 µl of 20mg/ml proteinase K (NEB) and 120 µl of 10% SDS and incubated at 55°C for 30 minutes, adding 130 µl of 5M sodium chloride

and incubate at 68°C overnight. DNA was precipitated with ethanol, washed with 70% ethanol, and dissolved in 105 µl of 10 mM Tris-HCl, pH 8. DNA was sheared on a Covaris S2 sonicator. Biotinylated DNA fragments were pulled with the MyOne Streptavidin C1 beads (Life Technologies). To repair the ends of sheared DNA and remove biotin from unligated ends, DNA-bound beads were re-suspended in 100 µl of mix containing 82 µl of 1X NEB T4 DNA ligase buffer with 10mM ATP (NEB), 10 µl of 10 (2.5mM each) 25mM dNTP mix, 5 µl of 10U/µl NEB T4 PNK (NEB), 4 µl of 3U/µl NEB T4 DNA polymerase (NEB), and 1 µl of 5U/µl NEB DNA polymerase I, Large (Klenow) Fragment (NEB).

After end-repair, dATP attachment was carried out in 100 µl of reaction solution, consisting of 90 µl of 1X NEBuffer 2, 5 µl of 10mM dATP, and 5 µl of 5U/µl NEB Klenow exo minus (NEB). The reaction was incubated at 37°C for 30 minutes. The beads were then cleaned for Illumina sequencing adaptor ligation which was done in a mix containing 50 µl of 1X T4 ligase buffer, 3 µl T4 DNA ligases (NEB), and 2 µl of a 15 µM Illumina indexed adapter at room temperature for 1 hour. DNA was dissociated from the bead by heating at 98° C for 10 min, separated on a magnet, and transferred to a clean tube.

Final amplification of the library was carried out in multiple PCRs using Illumina PCR primers. The reactions were performed in 25 µl scale consisting of 25 ng of DNA, 2 µl of 2.5mM dNTPs, 0.35 µl of 10 µM each primer, 2.5 µl of 10X PfuUltra buffer, PfuUltra II Fusion DNA polymerase (Agilent). The PCR cycle conditions were set to 98° C for 30 seconds as the denaturing step, followed by 14 cycles of 98° C 10 seconds, 65° C for 30 seconds, 72° C for 30 seconds, then with an extension step at 72° C for 7 minutes.

After PCR amplification, the products from the same library were pooled and fragments ranging in sized of 300-500 bp were selected with AMPure XP beads (Beckman Coulter). The sized

selected libraries were sequenced to produce paired-end Hi-C reads on the Illumina HiSeq 2500 platform with the V4 of 125 cycles.

RNAseq

Naïve WT and WAPL KO T cells, or WT and WAPL KO T cells isolated from syngeneic or allogeneic BMT on day 7 were first purified as described previously (Sun et al, 2015). All dead cells were excluded by sorting for far red fluorescent reactive dye (Invitrogen, Cat. L10120). Then WT T cells were sorted for PE-CD3 and APC-CD45.2, WAPL KO T cells were sorted for APC-CD3, PE-CD45.2 and GFP positive. Each sample was pooled T cells from 3-4 mice, and samples in each group were biologically triplicated. RNAs were isolated using DNA/RNA mini Kit (Qiangen Cat 80204) by following the RNA isolation procedures. Sequencing was performed by the University of Michigan (UM) DNA Sequencing Core, using the Illumina Hi-Seq 4000 platform, paired end, 50 cycles and Ribosomal Reduction library prep.

At the UM Bioinformatics Core, we downloaded the reads files from the Sequencing Core's storage and concatenated those into a single FASTQ file for each sample. We checked the quality of the raw reads data for each sample using FastQC (version 0.11.3) to identify features of the data that may indicate quality problems (e.g., low quality scores, over-represented sequences, inappropriate GC content). We used the Tuxedo Suite software package for alignment, differential expression analysis, and post-analysis diagnostics. Briefly, we aligned reads to the reference genome (GRCm38) using TopHat (version 2.0.13) and Bowtie2 (version 2.2.1.). We used default parameter settings for alignment, with the exception of: "--b2-very-sensitive" telling the software to spend extra time searching for valid alignments. We used FastQC for a second round of quality control (post-alignment), to ensure that only high-quality data would be input to expression quantitation and differential expression analysis. We used Cufflinks/CuffDiff (version 2.2.1) for expression quantitation, normalization, and differential expression analysis, using GRCm38.fa as the reference genome sequence. For this analysis, we used parameter settings: "--multi-read-

correct" to adjust expression calculations for reads that map in more than one locus, as well as "--compatible-hits-norm" and "--upper-quartile–norm" for normalization of expression values. We generated diagnostic plots using the CummeRbund R package. We used locally developed scripts to format and annotate the differential expression data output from CuffDiff. Briefly, we identified genes and transcripts as being differentially expressed based on FDR ≤ 0.05, and fold change ≥ ± 1.5. We annotated genes with NCBI Entrez GeneIDs and text descriptions. RNA-seq reads in bam format were mapped to the most recent mouse genome (mm10) using IGV (Integrative Genomics viewer; https://software.broadinstitute.org/software/igv/ . The RNA-seq data reported here can be found in the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) with the series accession ID GSE134975.

Hi-C Analysis

We generated the Hi-C matrices using Juicer (Durand et al., 2016). Juicer uses BWA-mem to align each paired-end read separately. It then determines which reads can be mapped uniquely and keeps unambiguously mapped read pairs. Each read is assigned to a "fragment", determined by the restriction enzyme cut sites and paired reads that map to the same fragment are removed. The Juicer pipeline creates a binary data file (namely, the ".hic" file), which contains Hi-C contacts. The .hic file is imported into MATLAB-compatible variables via an in-house MATLAB script ('fdnLoadHic.m') (Ronquist et al. 2019). Then, Hi-C data were subsequently ICE normalized (Imakaev et al., 2012) and binned at 100kb and 1Mb resolution. Additionally, centromeric and telomeric regions were removed in the process. Hi-C data for this study are available through the Gene Expression Omnibus (GEO) database (accession number: PRJNA608895).

Frobenius Norm

The Frobenius norm describes the size or volume of a matrix. As such, the Frobenius norm can be used as a measure of variance or variability in data (Drineas and Mahoney, 2018). For matrix

A of dimension mn, that is $A \in R^{m \times n}$, the mathematical definition of the Frobenius norm is as follows:

$$||A||_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2} \, ,$$

where a_{ij} is an element of the data matrix A. Equivalently this expression can be written as,

$$||A||_F^2 = trace(A^T A) \, ,$$

where $trace$ is the sum of the diagonal elements of a matrix (Elden, 2007). When comparing two sets of data, the notation is modified as follows:

$$||A - B||_F^2 = trace((A - B)^T (A - B)) \, ,$$

where A and B are matrices of the same dimensions. The Frobenius norm of a vector is equivalent to the Euclidean norm.

A/B Compartmentalization

Chromatin can either take on a condensed heterochromatic form or a looser euchromatic form. We use the Fiedler vector – the eigenvector corresponding to the second smallest eigenvalue (Fiedler number) of the normalized Laplacian matrix – to describe this feature mathematically which allows us to bi-partition the data into signed compartments – "A" (positive values) corresponding to euchromatin and "B" (negative values) corresponding to heterochromatin. The Fiedler vector and first principal component (PC1) vector are mathematically equivalent. However, unlike the PC1 vector, the Fiedler vector is accompanied by its corresponding Fiedler number which allows us to assign a magnitude of connectivity to a network.

Identification of TADs

Topologically associating domains (TADs) were identified using spectral clustering as described in Chen et al. 2016. This technique calculates the Fiedler vector for a normalized Hi-C adjacency matrix and initially organizes neighboring regions whose Fiedler values have the same sign into shared domains. This initial TAD structure is repartitioned if for a given domain, the Fiedler number is less than a user-defined threshold (λ_{thr}) to ensure that the TADs are not too large. Resulting TADs are iteratively repartitioned until their respective Fiedler numbers are larger than λ_{thr} or until the smallest allowable TAD size (default is 3) is reached. We perform this procedure for chromosome-specific Hi-C contact maps at 100kb resolution. For our analysis, λ_{thr} was chosen to ensure a median TAD size of 900kb, as the expected median TAD size in mammalian genomes is 880kb (rounded to 900kb for our data since bins are in intervals of 100kb) (Bonev and Cavalli, 2016). A specific λ_{thr} was chosen for each chromosome and sample to ensure each TAD clustering set would have the same approximate median TAD size.

Quantifying TAD Structure Similarity

Variation of information (VI) is a metric used for comparing dataset clusterings (Meila et al. 2007) and has been used to quantify the similarity of TAD organization between samples in prior studies (Filippova et al. 2014; Sauerwald and Kingsford, 2018). VI treats the relationship between two clusterings as an information exchange, meaning it takes into account the loss and gain of TAD borders between samples. VI is equivalent to the sum of two conditional entropies, computed as:

$$VI(C,C') = H(C|C') + H(C'|C),$$

where C and C' are two clusterings and the two conditional entropies measure the information we lose about C and gain about C', respectively, going from cluster C to cluster C'. VI is bounded by 0 and log(n), where n is the number of genomic bins in each clustering.

Hi-C Derived 5C Contact Map Generation

We constructed a synthetic 5C contact map derived from a genome-wide 1Mb Hi-C contact map for genomic regions containing genes in the cell cycle gene networks. This was done by locating the genomic bins corresponding to cell cycle genes, extracting the interchromosomal and cell cycle loci-specific intrachromosomal interaction frequencies for those bins, and stitching them together in genomic order. Our working set of cell cycle genes was sourced from the KEGG database and included 170 genes distributed across all chromosomes in the genome. Some of the genes occupy the same 1Mb genomic bins resulting in a 141 by 141 dimensioned adjacency matrix at 1Mb-resolution.

Cell proliferation assay

CellTrace™ Far Red was utilized for cell proliferation assay. Purified T cells were labeled with Far Red at final concentration of 1 μmol/L according to manufacturer's instruction (Molecular Probes, C34564). Far Red-labeled T cells were cultured similarly to above MLR for allogeneic reaction or Dynabeads T cell activator CD3/CD28 stimulation for up to 7 days, then collected cells were examined with flow cytometry, gated for CD3 (APC) positive, and Far Red dilution was determined.

FxCycle™ Far Red Stain for DNA content measurement

For *in vitro* experiments, WT or WAPL KO T cells were treated as above MLR for allogeneic reaction or Dynabeads T cell activator CD3/CD28 stimulation for up to 5 days. For *in vivo* experiments, transferred WT and WAPL KO T cells were isolated and purified from spleens of BALB/c recipient mice on day 5 and 7 after allogeneic BMT. The T cells were fixed with 10% formaldehyde for 30 mins, washed 3 time with PBS and adjusted the sample cell concentration at 1×10^6 cells/mL. FxCycle™ Far Red stain (200 nM) (Molecular Probes, F10348) and 5 μL of

RNase A (20 mg/mL) (Roche Cat. 70294823) were added to flow cytometry samples and continued for incubation at room temperature for 30 minutes and protected from light. Samples were analyzed with flow cytometer (AttuneNxT) without washing using 633nm excitation and emission collected in a 660 bandpass. The DNA contents were determined as 2C and >2C by fluorescence intensities.

ELISA

Concentrations of TNF-α and IFN-γ in sera on day 21 after allogeneic BMT were measured with specific anti-mouse ELISA kits from BD Biosciences. Assays were performed per the manufacturer's protocol and read at 450 nm in a microplate reader (Bio-Rad). The concentrations were calculated from triplicate samples as mean ± SEM.

FACS

Single cell suspensions of spleens and thymii were prepared as previously described (Sun et al, 2015). Briefly, to analyze surface phenotype, purified T cells and thymocytes from B6 WT, WAPL KO deficiency mice, or transplanted animals, were washed with FACS wash buffer (2% bovine serum albumin [BSA] in phosphate-buffered saline [PBS]), pre-incubated with the rat anti-mouse FcR mAb 2.4G2 for 15 minutes at 4 ℃ to block nonspecific FcR binding of labeled antibodies, then resuspended in FACS wash buffer and stained with conjugated monoclonal antibodies purchased from BD Biosciences (San Jose, CA): allophycocyanin (APC)-conjugated monoclonal antibodies (MoAbs) to CD4, CD8, CD3, CD45.2, CD45.1, CD25 and CD69; phycoerythrin (PE)-conjugated MoAbs to CD3, CD4, CD8, CD25, and TCR□; allophycocyanin (APC)-conjugated MoAbs to CD3, CD4, and PerCP/Cy5.5-conjugated MoAbs to CD3, CD4 and CD8 were purchased from eBioscience (SanDiego, CA). Next, cells were analyzed using an AttuneNxT flow cytometer. For intracellular staining, cells were stained for CD4 and CD25 antibodies as above, then fixed with IC Fixation Buffer (Biolegend, Cat. No.420801) and incubated 20-60 minutes at

room temperature and continued by adding 2 mL of 1X Permeabilization Buffer (Biolegend Cat. No. 421002) and centrifuged at 400-600 x g for 5 minutes at room temperature. Cells were resuspended the cell pellet in 100 μL of 1X Permeabilization Buffer and PE-conjugated FoxP3 antibody (eBioscience, SanDiego, CA, Cat. 126403) was added at 0.5 ug/million cells/100 ul and incubated for 30 minutes at room temperature. Stained cells were resuspended in an appropriate volume of Flow Cytometry Staining Buffer for flow cytometry analyses. Apoptotic cells were detected by PE-Annexin V staining.

Statistical analyses:

Data were analyzed using Prism GraphPad (Version 7). Briefly, comparisons between 2 groups were calculated using t test, while comparisons between 2 groups at multiple time points were calculated utilizing Holm-Sidak method. Log-rank (Mantel-Cox) test was utilized to analyze all survival data. Mann-Whitney test was used for the statistical analysis of clinical scores (Sun et al, 2015).

Study approval

Study approval. All animal studies were reviewed and approved of by the University Committee on Use and Care of Animals of the University of Michigan, based on University Laboratory Animal Medicine guidelines.

REFERENCES

Allahyar, A., Vermeulen, C., Bouwman, B.A.M., Krijger, P.H.L., Verstegen, M., Geeven, G., van Kranenburg, M., Pieterse, M., Straver, R., Haarhuis, J.H.I.*, et al.* (2018). Enhancer hubs and loop collisions identified from single-allele topologies. Nat Genet *50*, 1151-1160.

Alvarez, S.E., Harikumar, K.B., Hait, N.C., Allegood, J., Strub, G.M., Kim, E.Y., Maceyka, M., Jiang, H., Luo, C., Kordula, T.*, et al.* (2010). Sphingosine-1-phosphate is a missing cofactor for the E3 ubiquitin ligase TRAF2. Nature *465*, 1084-1088.

Blazar, B.R., Hill, G.R., and Murphy, W.J. (2020). Dissecting the biology of allogeneic HSCT to enhance the GvT effect whilst minimizing GvHD. Nat Rev Clin Oncol *17*, 475-492.

Bonev, B., and Cavalli, G. (2016). Organization and function of the 3D genome. Nat Rev Genet *17*, 772.

Boveri, T., (1909). Die Blastomerenkerne von Ascaris megalocephala. Archiv fur Zellforschung, 3, p.181.

Busslinger, G.A., Stocsits, R.R., van der Lelij, P., Axelsson, E., Tedeschi, A., Galjart, N., and Peters, J.M. (2017). Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. Nature *544*, 503-507.

Chen, H., Chen, J., Muir, L.A., Ronquist, S., Meixner, W., Ljungman, M., Ried, T., Smale, S., and Rajapakse, I. (2015). Functional organization of the human 4D Nucleome. Proc Natl Acad Sci U S A *112*, 8002-8007.

Chen, J., Hero, A.O., 3rd, and Rajapakse, I. (2016). Spectral identification of topological domains. Bioinformatics *32*, 2151-2158.

Cremer, M., and Cremer, T. (2019). Nuclear compartmentalization, dynamics, and function of regulatory DNA sequences. Genes Chromosomes Cancer *58*, 427-436.

Cremer, T., and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet *2*, 292-301.

Cuadrado, A., and Losada, A. (2020). Specialized functions of cohesins STAG1 and STAG2 in 3D genome architecture. Curr Opin Genet Dev *61*, 9-16.

Dekker, J., and Mirny, L. (2016). The 3D Genome as Moderator of Chromosomal Communication. Cell *164*, 1110-1121.

Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002). Capturing chromosome conformation. Science *295*, 1306-1311.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.

Dolatabadi, S., Candia, J., Akrap, N., Vannas, C., Tesan Tomic, T., Losert, W., Landberg, G., Aman, P., and Stahlberg, A. (2017). Cell Cycle and Cell Size Dependent Gene Expression Reveals Distinct Subpopulations at Single-Cell Level. Front Genet *8*, 1.

Dostie, J., Richmond, T.A., Arnaout, R.A., Selzer, R.R., Lee, W.L., Honan, T.A., Rubio, E.D., Krumm, A., Lamb, J., Nusbaum, C*., et al.* (2006). Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. Genome Res *16*, 1299-1309.

Drineas, Petros, and Michael W. Mahoney. "Lectures on randomized numerical linear algebra." *The Mathematics of Data* 25 (2018): 1.

Eldén, Lars. *Matrix methods in data mining and pattern recognition*. Society for Industrial and Applied Mathematics, 2007.

Felsenfeld, G., and Groudine, M. (2003). Controlling the double helix. Nature *421*, 448-453.

Finn, E.H., and Misteli, T. (2019). Molecular basis and biological function of variability in spatial genome organization. Science *365*.

Flavahan, W.A., Drier, Y., Liau, B.B., Gillespie, S.M., Venteicher, A.S., Stemmer-Rachamimov, A.O., Suva, M.L., and Bernstein, B.E. (2016). Insulator dysfunction and oncogene activation in IDH mutant gliomas. Nature *529*, 110-114.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A. (2016). Formation of Chromosomal Domains by Loop Extrusion. Cell Rep *15*, 2038-2049.

Fudenberg, G., and Mirny, L.A. (2012). Higher-order chromatin structure: bridging physics and biology. Curr Opin Genet Dev *22*, 115-124.

Gard, S., Light, W., Xiong, B., Bose, T., McNairn, A.J., Harris, B., Fleharty, B., Seidel, C., Brickner, J.H., and Gerton, J.L. (2009). Cohesinopathy mutations disrupt the subnuclear organization of chromatin. J Cell Biol *187*, 455-462.

Goriounov, D., Leung, C.L., and Liem, R.K. (2003). Protein products of human Gas2-related genes on chromosomes 17 and 22 (hGAR17 and hGAR22) associate with both microfilaments and microtubules. J Cell Sci *116*, 1045-1058.

Haarhuis, J.H., Elbatsh, A.M., van den Broek, B., Camps, D., Erkan, H., Jalink, K., Medema, R.H., and Rowland, B.D. (2013). WAPL-mediated removal of cohesin protects against segregation errors and aneuploidy. Curr Biol *23*, 2071-2077.

Haarhuis, J.H., and Rowland, B.D. (2017). Cohesin: building loops, but not compartments. EMBO J *36*, 3549-3551.

Haarhuis, J.H.I., van der Weide, R.H., Blomen, V.A., Yanez-Cuna, J.O., Amendola, M., van Ruiten, M.S., Krijger, P.H.L., Teunissen, H., Medema, R.H., van Steensel, B*., et al.* (2017). The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. Cell *169*, 693-707 e614.

Hill, L., Ebert, A., Jaritz, M., Wutz, G., Nagasaka, K., Tagoh, H., Kostanova-Poliakova, D., Schindler, K., Sun, Q., Bonelt, P.*, et al.* (2020). Wapl repression by Pax5 promotes V gene recombination by Igh loop extrusion. Nature *584*, 142-147.

Hu, G., Cui, K., Fang, D., Hirose, S., Wang, X., Wangsa, D., Jin, W., Ried, T., Liu, P., Zhu, J.*, et al.* (2018). Transformation of Accessible Chromatin and 3D Nucleome Underlies Lineage Commitment of Early T Cells. Immunity *48*, 227-242 e228.

Inaba, T., Honda, H., and Matsui, H. (2018). The enigma of monosomy 7. Blood *131*, 2891-2898.

Isoda, T., Moore, A.J., He, Z., Chandra, V., Aida, M., Denholtz, M., Piet van Hamburg, J., Fisch, K.M., Chang, A.N., Fahl, S.P.*, et al.* (2017). Non-coding Transcription Instructs Chromatin Folding and Compartmentalization to Dictate Enhancer-Promoter Communication and T Cell Fate. Cell *171*, 103-119 e118.

Kim, K.Y., Kovacs, M., Kawamoto, S., Sellers, J.R., and Adelstein, R.S. (2005). Disease-associated mutations and alternative splicing alter the enzymatic and motile activity of nonmuscle myosins II-B and II-C. J Biol Chem *280*, 22769-22775.

Kim, S.H., McQueen, P.G., Lichtman, M.K., Shevach, E.M., Parada, L.A., and Misteli, T. (2004). Spatial genome organization during T-cell differentiation. Cytogenet Genome Res *105*, 292-301.

Kueng, S., Hegemann, B., Peters, B.H., Lipp, J.J., Schleiffer, A., Mechtler, K., and Peters, J.M. (2006). Wapl controls the dynamic association of cohesin with chromatin. Cell *127*, 955-967.

Lajoie, B.R., Dekker, J., and Kaplan, N. (2015). The Hitchhiker's guide to Hi-C analysis: practical guidelines. Methods *72*, 65-75.

Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O.*, et al.* (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science *326*, 289-293.

Lupianez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R.*, et al.* (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. Cell *161*, 1012-1025.

Maeda, Y., Tawara, I., Teshima, T., Liu, C., Hashimoto, D., Matsuoka, K., Tanimoto, M., and Reddy, P. (2007). Lymphopenia-induced proliferation of donor T cells reduces their capacity for causing acute graft-versus-host disease. Exp Hematol *35*, 274-286.

Misteli, T. (2020). The Self-Organizing Genome: Principles of Genome Architecture and Function. Cell *183*, 28-45.

Nasmyth, K., and Haering, C.H. (2009). Cohesin: its roles and mechanisms. Annu Rev Genet *43*, 525-558.

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J.*, et al.* (2012). Spatial partitioning of the regulatory landscape of the X-inactivation centre. Nature *485*, 381-385.

Oliveira, R.A., and Nasmyth, K. (2013). Cohesin cleavage is insufficient for centriole disengagement in Drosophila. Curr Biol *23*, R601-603.

Peters, J.M., and Nishiyama, T. (2012). Sister chromatid cohesion. Cold Spring Harb Perspect Biol *4*.

Piche, J., Van Vliet, P.P., Puceat, M., and Andelfinger, G. (2019). The expanding phenotypes of cohesinopathies: one ring to rule them all! Cell Cycle *18*, 2828-2848.

Rajapakse, I., and Groudine, M. (2011). On emerging nuclear order. J Cell Biol *192*, 711-721.

Rajapakse, I., Groudine, M., and Mesbahi, M. (2011). Dynamics and control of state-dependent networks for probing genomic organization. Proc Natl Acad Sci U S A *108*, 17257-17262.

Rajapakse, I., Perlman, M.D., Scalzo, D., Kooperberg, C., Groudine, M., and Kosak, S.T. (2009). The emergence of lineage-specific chromosomal topologies from coordinate gene regulation. Proc Natl Acad Sci U S A *106*, 6679-6684.

Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S.*, et al.* (2015). In vivo genome editing using Staphylococcus aureus Cas9. Nature *520*, 186-191.

Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S.*, et al.* (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665-1680.

Rawlings, J.S., Gatzka, M., Thomas, P.G., and Ihle, J.N. (2011). Chromatin condensation via the condensin II complex is required for peripheral T-cell quiescence. EMBO J *30*, 263-276.

Reddy, P., Negrin, R., and Hill, G.R. (2008). Mouse models of bone marrow transplantation. Biol Blood Marrow Transplant *14*, 129-135.

Remeseiro, S., Cuadrado, A., and Losada, A. (2013). Cohesin in development and disease. Development *140*, 3715-3718.

Ried, T., and Rajapakse, I. (2017). The 4D Nucleome. Methods *123*, 1-2.

Robson, M.I., de Las Heras, J.I., Czapiewski, R., Sivakumar, A., Kerr, A.R.W., and Schirmer, E.C. (2017). Constrained release of lamina-associated enhancers and genes from the nuclear envelope during T-cell activation facilitates their association in chromosome compartments. Genome Res *27*, 1126-1138.

Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., Haering, C.H., Mirny, L.*, et al.* (2017). Two independent modes of chromatin organization revealed by cohesin removal. Nature *551*, 51-56.

Seitan, V.C., Hao, B., Tachibana-Konwalski, K., Lavagnolli, T., Mira-Bontenbal, H., Brown, K.E., Teng, G., Carroll, T., Terry, A., Horan, K.*, et al.* (2011). A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. Nature *476*, 467-471.

Servant, N., Varoquaux, N., Lajoie, B.R., Viara, E., Chen, C.J., Vert, J.P., Heard, E., Dekker, J., and Barillot, E. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. Genome Biol *16*, 259.

Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V.*, et al.* (2012). A map of the cis-regulatory sequences in the mouse genome. Nature *488*, 116-120.

Silva, M.C.C., Powell, S., Ladstatter, S., Gassler, J., Stocsits, R., Tedeschi, A., Peters, J.M., and Tachibana, K. (2020). Wapl releases Scc1-cohesin and regulates chromosome structure and segregation in mouse oocytes. J Cell Biol *219*.

Singh, V.P., and Gerton, J.L. (2015). Cohesin and human disease: lessons from mouse models. Curr Opin Cell Biol *37*, 9-17.

Srinivasan, M., Petela, N.J., Scheinost, J.C., Collier, J., Voulgaris, M., M, B.R., Beckouet, F., Hu, B., and Nasmyth, K.A. (2019). Scc2 counteracts a Wapl-independent mechanism that releases cohesin from chromosomes during G1. Elife *8*.

Straight, A.F., Cheung, A., Limouze, J., Chen, I., Westwood, N.J., Sellers, J.R., and Mitchison, T.J. (2003). Dissecting temporal and spatial control of cytokinesis with a myosin II Inhibitor. Science *299*, 1743-1747.

Strang, G., and Strang, G., (1993). Introduction to linear algebra (Vol. 3). Wellesley, MA: Wellesley-Cambridge Press.

Sun, Y., Oravecz-Wilson, K., Bridges, S., McEachin, R., Wu, J., Kim, S.H., Taylor, A., Zajac, C., Fujiwara, H., Peltier, D.C.*, et al.* (2019). miR-142 controls metabolic reprogramming that regulates dendritic cell activation. J Clin Invest *129*, 2029-2042.

Sun, Y., Oravecz-Wilson, K., Mathewson, N., Wang, Y., McEachin, R., Liu, C., Toubai, T., Wu, J., Rossi, C., Braun, T.*, et al.* (2015). Mature T cell responses are controlled by microRNA-142. J Clin Invest *125*, 2825-2840.

Sun, Y., Tawara, I., Zhao, M., Qin, Z.S., Toubai, T., Mathewson, N., Tamaki, H., Nieves, E., Chinnaiyan, A.M., and Reddy, P. (2013). Allogeneic T cell responses are regulated by a specific miRNA-mRNA network. J Clin Invest *123*, 4739-4754.

Szabo, Q., Bantignies, F., and Cavalli, G. (2019). Principles of genome folding into topologically associating domains. Sci Adv *5*, eaaw1668.

Tabebordbar, M., Zhu, K., Cheng, J.K.W., Chew, W.L., Widrick, J.J., Yan, W.X., Maesner, C., Wu, E.Y., Xiao, R., Ran, F.A.*, et al.* (2016). In vivo gene editing in dystrophic mouse muscle and muscle stem cells. Science *351*, 407-411.

Tedeschi, A., Wutz, G., Huet, S., Jaritz, M., Wuensche, A., Schirghuber, E., Davidson, I.F., Tang, W., Cisneros, D.A., Bhaskara, V.*, et al.* (2013). Wapl is an essential regulator of chromatin structure and chromosome segregation. Nature *501*, 564-568.

Tsuruga, H., Yabuta, N., Hashizume, K., Ikeda, M., Endo, Y., and Nojima, H. (1997). Expression, nuclear localization and interactions of human MCM/P1 proteins. Biochem Biophys Res Commun *236*, 118-125.

Waldman, T. (2020). Emerging themes in cohesin cancer biology. Nat Rev Cancer *20*, 504-515. Wu, S.R., and Reddy, P. (2017). Tissue tolerance: a distinct concept to control acute GVHD severity. Blood *129*, 1747-1752.

Wutz, G., Varnai, C., Nagasaka, K., Cisneros, D.A., Stocsits, R.R., Tang, W., Schoenfelder, S., Jessberger, G., Muhar, M., Hossain, M.J*., et al.* (2017). Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. EMBO J *36*, 3573-3599.

Yuen, K.C., and Gerton, J.L. (2018). Taking cohesin and condensin in context. PLoS Genet *14*, e1007118.

Zeiser, R., and Blazar, B.R. (2018). Acute Graft-versus-Host Disease. N Engl J Med *378*, 586.

Figure 1. Genome-wide structural effects on unstimulated naïve T cells after syngeneic/allogeneic transplantation. (A) Genome-wide Hi-C contact maps for WT unstimulated naïve, syngeneic, and allogeneic T cells. Balanced (KR-normalized) maps are shown at 1Mb resolution and log-scale. Numbers along the axes in the first contact map denote the chromosome at that respective region of the map and the number in the inset bottom left corner of the map indicates the highest value among elements in the maps. Contact maps contain only positive values, so the lowest value will always be zero. **(B)** Matrix subtraction between settings. Red matrix elements indicate higher contact frequencies (enrichment) in the setting in the numerator compared to the setting in the denominator. The opposite (depletion) is true for blue matrix elements. **(C)** (Top) Illustration of a single *intra*-chromosomal region of the genome-wide contact map. The upper right quadrant (red shaded area) represents all interactions between either half of the chromosome and denotes the area where we typically observe distance-dependent contact frequency decay. These interactions can be categorized as end-to-end, end-to-middle, or middle-to-middle. (Bottom) We evaluate this area of Chromosome 7 (1045 – 114 Mb by 976 – 1045 Mb) across all three settings. Gray arcs along the ideogram highlight the top 0.5% of pairwise interactions between either half of the chromosome. The break in the ideogram indicates where the centromere typically lies; we filter centromeric regions from our data.

A

B

100 bp---Cre
1123 bp---WT-LSL

285 bp ----CD4cre-LSL

457 bp---gRNAs-Wapl insert

} Genotyping

| 92 | 56 | | 1 | 2 | 62 |

Purified T cells were examined for GFP expression by FACS

142 kD----Wapl
42 kD-----β-actin

} T cells were purified and processed for detection of WAPL protein by WB

NC-WT
PC Cas9-GFP
1   2   3   4

Pups

C

D

Chromosome 14

Figure 2. Generating conditional WAPL knockout T cells utilizing CRISPR-CAS9 system.

(A) Map of 2 sgRNAs which are targeting exon 2 of *Wapl* gene on Chromosome 14. **(B)** Genotyping and confirmation of *Wapl* knocked out. Mice tail DNA was processed for PCR genotyping to confirm expressions of CRE and sgRNA-Wapl insert and Cas9-stop signal expression or cleavage. GFP expression was determined by flow cytometry and WAPL protein deletion was determined by Western Blotting. **(C)** Confirmation of WAPL KO in T cells on multiple pups according to the screen procedures in (B) by Western Blotting. (D) Confirmation of sgRNAs targeting segment in exon 2 of *Wapl* gene was deleted in biological triplicates by RNA-seq.

Figure 3. Genome-wide effects of WAPL knockout. (A) Genome-wide Hi-C contact maps for unstimulated naïve, syngeneic, and allogeneic T cells in the absence of WAPL. Balanced (KR-normalized) maps are shown at 1Mb resolution and log-scale. **(B)** Matrix subtraction between WT and KO T cells in each setting. Red matrix elements indicate increased pairwise interaction frequencies in the absence of WAPL while blue matrix elements indicate decreased frequencies in the absence of WAPL. Red arrows along the unstimulated naïve subtraction matrix point out a contrasting stripe of increased contact frequencies involving chromosome 11. **(C)** Structure and function differences between WT and KO T cells in each setting. The left y-axis (blue) reflects the Frobenius norm of the difference between WT and KO Hi-C matrices (measure of structural change) in each setting and the right y-axis (orange) reflects the Frobenius norm of the difference between WT and KO gene expression vectors (measure of functional change) in each setting. Chromosomes are indicated along the x-axis. **(D)** Montage of structure-function plots by chromosome. Fiedler numbers of Hi-C contact maps for each chromosome and cell type are plotted against the mean gene expression measured in transcripts per million (TPM). Gray shaded plots denote chromosomes with significant changes in structure and/or function. **(E)** Differential compartmentalization between each WT T cell setting. We observed 274 loci (100kb-length) with different chromatin states between unstimulated naïve and syngeneic T cells and 645 between both unstimulated naïve and allogeneic as well as syngeneic and allogeneic. Vertical bars below each heatmap indicate chromosomes in numerical order (1 to 19) from left to right. The heatmap color intensity reflects the magnitude of the Fiedler values for each locus. **(F)** A/B compartment switch events between WT and KO T cells. We observed 184 loci (100kb-length) with different chromatin states between WT and KO unstimulated naïve T cells, 303 between both WT and KO syngeneic T cells, and 413 between WT and KO allogeneic T cells.

Figure 4. Comparison of TAD organization between WT and KO T cells. (A-C) Chromosome

7 comparison of chromatin accessibility (top, track) and TAD clustering (bottom, map) between

WT and KO unstimulated naïve (A), syngeneic (B), and allogeneic (C) T cells. TADs are shown along a normalized (observed/expected) log-scale contact map at 100-kb resolution and are colored according to chromatin accessibility with TADs containing both accessible and inaccessible loci denoted 'mixed' and colored black. **(D-F)** The light blue boxed area on the contact maps highlights gene-rich region 60Mb – 70Mb in Chromosome 7. Zoomed in panels to the right offer a closer look at the region's interaction pattern and TADs, chromatin accessibility (bar plot directly right), and gene expression (rightmost bar plot). This region contains three cell cycle genes (*Fanci, Prc1, Blm*), whose chromatin accessibility is indicated by red bars in the chromatin accessibility plot and whose gene expression is indicated by the purple bars in the gene expression plot. There are more than three highlighted bars in both plots because the three genes each spanned two 100kb-length regions/bins. Green bars in the gene expression plot represent 100kb-length regions/bins containing one or more non-cell cycle genes that are differentially expressed ($|\log_2 FC| \geq 1$).

Figure 5. Cell cycle gene network across T cells. (A) Hi-C-derived 5C contact maps representing the cell cycle gene network. Maps are shown at 1-Mb resolution as a Pearson correlation of normalized (observed/expected) contacts. The purple and green boxes highlight subgroups of interest that are assessed in C. The purple subgroup is comprised of 10 loci (1Mb length) ranging non-contiguously from 18 Mb to 197 Mb and containing 10 cell cycle genes. The cell cycle gene loci in WT and KO T cells in each subgroup is comprised of 12 loci (1Mb

length) ranging non-contiguously from 1,489 Mb to 1,606 Mb and containing 16 cell cycle genes. **(B)** Structural and functional differences between cell cycle gene loci in WT and KO T cells in each setting. Difference is measured as the Frobenius norm of the difference between WT and KO data as previously described in Figure 2C. **(C)** Degree centrality (i.e. the row sum of a matrix) of subgroups of interest.

Figure 6. WAPL deficiency impairs T cell development, proliferation, and cell cycling.

(A) Total numbers of thymocytes isolated from WT and WAPL KO thymii were counted. Data were combined from WT (9) and WAPL KO (10) mice (mean ± SEM). **(B)** Total T cell numbers of

spleens from WT and WAPL KO mice were counted. Data were combined from WT (23) and

WAPL KO (26) mice (mean ± SEM). **(C)** WAPL protein expression was upregulated in T cells

when co-cultured with allogeneic DCs *in vitro*. A representative Western blot was shown. **(D)** The

mRNA expression of cell cycle genes was analyzed utilizing GSEA. Gene sets of WT and WAPL

KO T cells from naïve, syngeneic and allogeneic conditions were sorted by enrichment scores.

All groups were in biological triplicates. **(E-H)** According to ES in (D), the mRNA expressions of

top or bottom two genes were analyzed by quantitative real-time PCR. Data were combined from

three independent experiments (mean ± SEM). **(I)** H^3-TdR incorporation in WT and WAPL KO T

cells were analyzed by mixed lymphocyte reaction when stimulated by allogeneic DCs for 4 days.

Data were combined from three independent experiments (mean ± SEM). **(J)** Cell proliferation

was analyzed by Far Red dilution for either WT or WAPL KO T cells when stimulated with

allogeneic DCs for up to 7 days. Data were combined from two independent experiments (mean

± SEM). **(K)** Cell apoptosis was analyzed by FACS after Annexin V staining for WT and *Wapl* KO

T cells when stimulated with allogeneic DCs for up to 7 days. Data were combined from two

independent experiments (mean ± SEM). **(L)** After WT and WAPL KO T cells were stimulated

with allogeneic DCs for 3 and 5 days *in vitro*, cells were stained with FxCycle™ Far Red, DNA

contents were separated as 2C and >2C by flow cytometry. Combined data were obtained from

three independent experiments (mean ± SEM). **(M)** After WT and WAPL KO T cells were

stimulated with CD3/CD28Ab for 2-5 days in *vitro*, cells were stained with FxCycle™ Far Red,

DNA contents were separated as 2C and >2C by flow cytometry. Combined data were obtained

from three independent experiments (mean ± SEM). **(N)** WT and WAPL KO T cells were isolated

from recipient spleens on days 5 and 7 after allogeneic BMT *in vivo* and were stained with

FxCycle™ Far Red, DNA contents were separated as 2C and >2C by flow cytometry. Combined

data were obtained from two independent experiments (mean ± SEM).

Figure 7. WAPL-deficient T cells mitigate GVHD severity in allogeneic BMT.
(A) *Wapl* expression was upregulated in T cells on day 7 after allogeneic BMT. A representative Western blot was shown. **(B)** Allogeneic MHC mismatched B6 (H2^b) into BALB/c(H2^d) GVHD model. Survival percentages were recorded on daily base. P-values were obtained using the log-rank test based on comparisons between animals that received either TCD BM plus WT T cells (blue circles) or TCD BM plus WAPL KO T cells (red squares). Combined data were from 2

independent experiments. **(C)** GVHD score was assessed by a standard scoring system as before (Sun, et al, 2015). The Mann-Whitney U test was used for the statistical analysis of clinical scores between animals that received either TCD BM plus WT T cells (blue circles) or TCD BM plus WAPL KO T cells (red squares). Combined data were from 2 independent experiments. **(D)** Histopathological analysis of the B6 into BALB/c BMT model. Bowel (small and large intestine), liver, and skin were obtained after BMT on day 21. GVHD scores were from 2 independent experiments (mean ± SEM). **(E and F)** Sera were collected from recipient mice on day 21 after allogeneic BMT as in (B), the concentrations of IFNγ and TNFα were measured by ELISA. Significant lower concentrations of INFγ and TNFα were detected in the sera collected from mice transferred with WAPL KO T cells. Combined data were from 2 independent experiments (mean ± SEM). **(G)** *In vivo* T cell expansion was determined by isolating transferred WT and WAPL KO in BALB/c recipients on day 7 after allogeneic BMT. Data were combined from 3 independent experiments (mean ± SEM). **(H)** Transferred donor T cells (WT and WAPL KO) were isolated from recipient's spleens on day 7 after allogeneic BMT and stained for $CD3^+$ and $CD69^+$. The absolute numbers of activated WAPL KO T cells ($CD3^+CD69^+$) were significantly lower than WT T cells. Data were combined from 3 independent experiments (mean ± SEM). **(I and J)** Donor T cells were isolated as in (H) on day 7 after allogeneic BMT and stained for $CD4^+$ 25^+ $FoxP3^+$ for identifying T regulatory cells. The absolute numbers of WAPL KO Treg cells were significantly lower than WT Treg cells **(I),** although the percentages were the same between WT and WAPL KO T cells **(J).** Data were combined from 3 independent experiments (mean ± SEM). P-values were obtained using unpaired t-test.