Published in final edited form as:
*Immunity*. 2012 December 14; 37(6): 971–985. doi:10.1016/j.immuni.2012.10.007.

# The DNA damage- and transcription- associated protein Paxip1 controls thymocyte development and emigration

**Elsa Callen**[1], **Robert B. Faryabi**[1], **Megan Luckey**[2,*], **Bingtao Hao**[3,*], **Jeremy A. Daniel**[1,7], **Wenjing Yang**[4], **Hong-Wei Sun**[5], **Greg Dressler**[6], **Weiqun Peng**[4], **Hongbo Chi**[8], **Kai Ge**[9], **Michael S. Krangel**[3,**], **Jung-Hyun Park**[2,**], and **André Nussenzweig**[1,†]

[1]Laboratory of Genome Integrity, National Cancer Institute, NIH, Bethesda MD 20892

[2]Experimental Immunology Branch, National Cancer Institute, NIH, Bethesda MD 20892

[3]Department of Immunology, Campus Box 3010, Duke University Medical Center, Durham, NC 27710

[4]Department of Physics, The George Washington University, Washington, DC 20052

[5]Biodata Mining and Discovery Section, Office of Science and Technology, NIH, Bethesda MD 20892

[6]Department of Pathology, University of Michigan, Ann Arbor, MI 48109, USA. 262 Danny Thomas Place, Room E-7013, Memphis, TN 38105-2794

[7]The Novo Nordisk Foundation Center for Protein Research, Faculty of Health Sciences, University of Copenhagen, Copenhagen, Denmark

[8]Department of Immunology, St. Jude Children's Research Hospital, Memphis TN 38105

[9]Laboratory of Endocrinology and Receptor Biology, National Institute of Diabetes and Digestive and Kidney Diseases, NIH, Bethesda MD 20892

## Summary

Histone 3 lysine 4 trimethylation (H3K4me3) is associated with promoters of active genes and found at hot spots for DNA recombination. Here we have shown that PAXIP1, a protein associated with MLL3 and MLL4 methyltransferase and the DNA damage response, regulates RAG-mediated cleavage and repair during V(D)J recombination in $CD4^+$ $CD8^+$ DP thymocytes. Loss of PAXIP1 in developing thymocytes diminished Jα H3K4me3 and germline transcription, suppressed double strand break formation at 3' Jα segments but resulted in accumulation of unresolved T cell receptor a-chain gene (*Tcra)* breaks. Moreover, PAXIP1 was essential for release of mature single positive (SP) αβT cells from the thymus through transcriptional activation of sphingosine-1-phosphate receptor *S1pr1* as well as for natural killer T cell development. Thus, in addition to maintaining genome integrity during *Tcra* rearrangements, PAXIP1 controls distinct transcriptional programs during DP differentiation necessary for *Tcra* locus accessibility, licensing mature thymocytes for trafficking and natural killer T cell development.

†correspondence: andre_nussenzweig@nih.gov.
*These authors contributed equally to this work
**These authors contributed equally to this work

## Introduction

DNA-mediated processes including transcription and DNA repair are regulated by post-translational modifications of the histone components of chromatin (Taverna et al., 2007). Post-translational modifications are carried out by histone methyltransferases and demethylases, histone acetyltransferases, protein kinases, and ubiquitin-and SUMO-protein ligases. Specific histone modifications are recognized by so called “reader-effector” modules in non-histone proteins, Such reader proteins frequently carry enzymatic activity on the chromatin template which can render the DNA more accessible to the transcription and/or DNA repair machineries (Taverna et al., 2007).

The Recombination Activating Genes (*Rag1* and *Rag2*) are lymphoid specific proteins that are essential for V(D)J recombination, a process that generates diversity in the antigen receptor repertoire (Gellert, 2002; Schatz and Swanson, 2011). RAG1 and RAG2 (herein referred to as RAG) bind and cleave DNA at specific recombination signal sequences (RSSs) that flank each V, D and J gene segment. Thereafter the liberated DNA ends are recognized and repaired by the non-homologous end joining (NHEJ) pathway. The ability of RAG to initiate V(D)J recombination is dictated by the accessibility of RSSs within chromatin (Krangel, 2007; Osipovich and Oltz, 2010).

The first clue about the mechanisms regulating RAG activity was the discovery that germline transcripts of antigen receptor gene segments correlate with recombinationally active open chromatin configuration susceptible to double strand break (DSB) cleavage (Yancopoulos and Alt, 1985). The link between transcription, chromatin structure and recombination was strengthened by studies documenting the effects of transcriptional blockade on downstream histone modifications and RAG-dependent DSBs (Abarrategui and Krangel, 2006). A further link between germline transcription and increased accessibility was the discovery that RAG is targeted to histone 3 lysine 4 trimethylation (H3K4me3) (Liu et al., 2007; Matthews et al., 2007). The RAG2 protein contains a PHD finger at its C terminus, which interacts with H3K4me3. H3K4me3 is highly enriched in the 5’ end of all transcription units including chromatin encompassing recombinationally active genes in the immunoglobulin and T cell receptor loci. A recent study reported that RAG2 binds to virtually all genomic regions containing H3K4me3 (Ji et al., 2010). Moreover, RAG2 association with H3K4me3 is dependent on the RAG2 PHD domain (Ji et al., 2010), and lymphocyte development is partially impaired by PHD mutations that affect RAG2’s recognition of H3K4me3 (Matthews et al., 2007). Finally, biochemical studies revealed that H3K4me3 stimulates the catalytic function of RAG (Grundy et al., 2010; Shimazaki et al., 2009). Altogether, these studies suggest a model in which RAG2 association with H3K4me3 via its PHD-finger domain enhances DNA binding and cleavage. In this way beyond the recognition of the RSS in the DNA, the PHD chromatin reader in RAG recombinase increases the efficiency of V(D)J recombination.

Trimethylation of H3K4 is mediated by H3K4 methylases, which include at least 6 members (Set1A, Set1B, MLL1, MLL2, MLL3 and MLL4), each of which contains a conserved SET domain carrying the methyltransferase activity. Each H3K4 methylase exists in a large protein complex which shares four common subunits ASH2L, RBBP5, WDR5 and DPY30. Deletion of WDR5 results in a marked reduction of the recombination activity of RAG2 (Matthews et al., 2007), consistent with a role for H3K4me3 in enhancing V(D)J recombination. In addition to shared subunit composition, methylase complexes contain unique components which may provide target specificity to the complex. For example, PAXIP1 (Pax interaction with transcription-activation domain protein-1), is a subunit of the MLL3 and MLL4 complex (Cho et al., 2007; Issaeva et al., 2007; Patel et al., 2007) which regulates H3K4me3 and germline transcription initiation at IgG3 and IgG1 switch regions

during class switch recombination (Daniel et al., 2010; Schwab et al., 2011) (MLL3 is also known as KMT2C, and MLL4 is also known as ALR, MLL2, or KMT2D)(Cho et al., 2012). Transcription of switch regions is thought to render them in an accessible configuration that allows DNA cleavage by the enzyme, activation induced deaminase (AID) (Stavnezer et al., 2008). Nevertheless, loss of *Paxip1* does not affect amounts of H3K4me3 or transcription of Sμ and Sε, indicating that the PAXIP1-MLL3-MLL4 complex promotes accessibility of some but not all switch loci (Daniel et al., 2010).

RAG and AID-dependent DNA breaks at antigen receptor locus are marked by the phosphorylation of the histone variant H2AX (γ-H2AX) (Chen et al., 2000; Petersen et al., 2001). In addition to promoting germline transcription at the IgH locus, PAXIP1 exists in a separate complex that accumulates at sites of DNA damage in a manner dependent on the γ-H2AX DNA damage response pathway (Gong et al., 2009). Beyond promoting AID accessibility for IgG3 and IgG1 class switch recombination, PAXIP1 appears to function subsequently in the repair of a subset of AID dependent DSBs (Daniel et al., 2010).

Because H3K4me3 is thought to influence the accessibility of chromatin at V(D)J gene segments, it is of interest to determine the methylase activities that target the recombinase to specific loci. Here we have analyzed thymic development in mice deficient in PAXIP1. We have found that PAXIP1controls RAG mediated DSB formation and repair during *Tcra* recombination. Moreover, the development of natural killer T (NKT) cells, which require recombination and expression of a very limited and invariant TCRα chain, was ablated in the absence of PAXIP1. Surprisingly, mature *Paxip1*$^{-/-}$ thymocytes that successfully completed V(D)J recombination failed to exit the thymus due to impaired expression of the G protein-coupled sphingosine-1 phosphate receptor 1 (S1PR1). Thus, PAXIP1 regulates both recombinase accessibility and DNA repair in double positive (DP) T cells as well as the exit of mature single positive (SP) T cells from the thymus. We propose that PAXIP1 serves as a gatekeeper that ensures that cells with faulty rearrangements are not exported to the periphery.

## Results

### PAXIP1 regulates αβ T cell development

To determine the physiologic consequence of *Paxip1* loss during T cell development, we crossed *Paxip1*$^{f/f}$ mice with mice expressing Cre recombinase driven by the Lck proximal promoter (Hennet et al., 1995), which is active in immature thymocytes. Expression of *Paxip1* in mutant thymocytes was barely detectable by quantitative reverse transcriptase-polymerase chain reaction (RT-PCR), confirming the efficiency of deletion (Fig. S1). In *Paxip1*$^{f/f}$ × Lck Cre (referred to as *Paxip1*$^{-/-}$) mice, total thymocyte recovery was reduced to 50% of *Paxip1*$^{f/f}$ (referred to as WT) mice (Figures 1A and 1B). Consistent with this, staining for CD4 and CD8 co-receptors in 6- to 10- week old mice revealed a 25% decrease in the percentage and 50% decrease in the number of double positive (DP) thymocytes (Figures 1A and 1B), whereas the percentage of double negative (DN) thymocytes was similar (Figure S2A). The DN-DP thymocyte transition is associated with an approximately 100-fold expansion in thymus cellularity. We considered the possibility that reduced number of DP T cells might reflect inefficient proliferation. To test this, we examined the expression of CD25 which gets diluted during the DN-DP T cell transition (Crompton et al., 1994) (Figure S2B). Consistent with this, we observed a 3-fold increase in the percentage of *Paxip1*$^{-/-}$ DP T cells expressing CD25, indicative of decreased proliferation (Figure S2B).

Unlike the reduction in DP thymoctyes, we observed an unexpected increase in the number of CD4$^{+}$ and CD8$^{+}$ SP T cells in *Paxip1*$^{-/-}$ thymi (Figures 1A and 1B). SP thymocytes that accumulate in the absence of *Paxip1* were mature T cells, evidenced by a greater frequency

of *Paxip1*$^{-/-}$ SP T cells that express high amounts of TCRβ, CD69, and the maturation marker Qa2, and low amounts of heat stable antigen (HSA) in both CD4$^{+}$ and CD8$^{+}$ populations (Figures 1A and 1C; Figure S2C;). Associated with the accumulation of SP T cells in the thymus, there was a marked depletion in the number of mature *Paxip1* deficient T cells in the periphery (Figure 2). Whereas 78% of cells in WT lymph nodes were T cells, defined as TCRβ$^{+}$, only 42% of *Paxip1*$^{-/-}$ cells expressed TCRβ; moreover, there was 80% reduction in the number of CD4$^{+}$ and CD8$^{+}$ T cells in this compartment (n=16 mice analyzed) (Figure 2A). Finally, *Paxip1* deletion resulted in a 75% reduction in the number of splenic T cells (n=9) as well as a reduction of T cells in blood (Fig. 2B). The paucity of peripheral T cells (Figure 2) concomitant with the accumulation of mature T cells in thymus (Figure 1) suggested the possibility that PAXIP1 was required for T cell migration from the thymus to the periphery (see below).

## PAXIP1 regulates *Tcra* gene recombination

The *Tcra* locus becomes accessible at the DP stage which correlates with Jα chromatin marked by H3K4me3 (Krangel, 2007). We used chromatin immunoprecipitation (ChIP) coupled with Illumina sequencing to examine genome wide H3K4me3 profiles of DP T cells (CD4$^{+}$CD8$^{+}$TCR$^{lo/med}$), CD4$^{+}$ SP T cells (CD4$^{+}$CD8$^{-}$TCR$^{hi}$) and CD8$^{+}$ SP T cells (CD4$^{-}$CD8$^{+}$TCR$^{hi}$) populations from sorted WT and *Paxip1*$^{-/-}$ thymocytes. Primary rearrangements utilize Jα segments at the 5' end of the 65-kb Jα array and secondary Vα-Jα rearrangements make use of additional Jα segments located 3' of the primary VαJα (Krangel, 2007). In WT DP T cells, we found that the sequence tag distribution of H3K4me3 was evenly spread along Jα locus as previously reported (Ji et al., 2010) (Figure 3A). In contrast, *Paxip1*$^{-/-}$ DP T cells displayed a skewed H3K4me3 profile with the highest H3K4me3 signals at the 5' Jα segments and a gradual decay toward the 3' end of the Jα array (Figure 3A). For example, in WT DP thymocytes, 41% of the total tag counts within the 65 kb Jα array were 3' of Jα30, whereas tag counts at the 3' end of the locus were reduced to 13% in *Paxip1*$^{-/-}$ DPs (Figure 3A). Consistent with the results in DP T cells, the distribution of H3K4me3 in *Paxip1*$^{-/-}$ CD4$^{+}$ SP and *Paxip1*$^{-/-}$ CD8$^{+}$ SP T cells also showed skewing towards 5' relative to 3' segments (Figure S3A). Thus, loss of *Paxip1* results in decreased H3K4me3 at 3' Jα segments in both DP and SP T cells.

To determine whether decreased H3K4me3 at 3' Jα segments correlated with impairment in Vα to Jα rearrangements, we examined the Jα repertoire in WT and *Paxip1*$^{-/-}$ thymocytes. TCRβ$^{lo/med}$ DP cells were sorted and TCR transcripts were reverse transcribed and PCR amplified with a TRAV12 (Vα8) family primer and a Cα primer, as previously described (Guo et al., 2002; Hawwari et al., 2005). Jα usage was then assessed by Southern blotting with a series of Jα-specific oligonucleotide probes. This analysis of cDNA revealed greater utilization of 5' segments and less usage of 3' Jα segments in *Paxip1*$^{-/-}$ thymocytes relative to WT (Figure 3B). Moreover, DP thymocytes from *Paxip1*$^{-/-}$ mice showed decreased TCR Cα mRNA as measured by quantitative RT-PCR (Figure 3C).

To determine whether skewed Jα usage is due to biased Vα to Jα recombination, we utilized DNA from sorted TCRβ$^{lo/med}$ DP thymocytes to amplify genomic fragments generated by TRAV12 (Vα8) to Jα rearrangements by real-time PCR. Recombination from TRAV12 to 5' Jα segments were detected at a similar amount in *Paxip1*$^{-/-}$ and WT thymocytes (Figure 3D). However, DNA rearrangements to more distal 3'Jα segments (eg. Jα37-Jα2) were reduced in *Paxip1* mutant cells relative to WT thymocytes (Figure 3D). Thus, decreased 3' Jα usage and transcription (Figures 3B and 3C) correlates with defective Vα to Jα rearrangements (Figure 3D) in *Paxip1*$^{-/-}$ thymocytes.

Altered 3' Jα usage and recombination in *Paxip1* deficient thymocytes could result from increased programmed cell death of DP thymocytes as seen in *Rorc*-deficient mice (Guo et

al., 2002), deregulation of *Rag* expression in DP thymocytes (Yannoutsos et al., 2001), or decreased cleavage of *Tcra* segments. We found that the viability of *Paxip1*$^{-/-}$ thymocytes (62%) was comparable to WT (67%) after overnight culture (Figure S3B). Moreover, *RAG* expression was not decreased in the absence of *Paxip1* (Figure S3C), indicating that lack of 3' Jα usage was not due to shortened DP T cells lifespan or deregulated recombinase expression.

Although *RAG* expression was intact in *Paxip1*$^{-/-}$ thymocytes (Figure S3C), impaired RAG-mediated cleavage of 3' Jα segments could account for altered Jα usage and recombination. To examine the efficiency of DNA double strand break formation, we performed ligation-mediated PCR to amplify Jα signal ends as described (McMurry et al., 1997). We found that DSBs formed efficiently in *Paxip1*$^{-/-}$ thymocytes at 5' Jα segments (eg. Jα61, Jα58 and Jα49) (Figure 3E). However, there was a dramatic decrease in DSB formation at 3' end of the Jα array (e.g. at Jα27, Jα13 and Jα6) (Figure 3E). Thus, deficiency in PAXIP1-dependent H3K4me3 at 3' Jα genes correlates with impaired cleavage of those segments.

Primary *Tcra* rearrangements are dependent on transcription mediated by T early α (TEA) and Jα49 promoters at the 5' end of the array, which limits recombination to the 5' Jα61-Jα52 gene segments. After initial Vα-Jα rearrangements, Vα promoters drive subsequent rounds of secondary rearrangements to Jα segments that are more 3' (Hawwari and Krangel, 2007). During primary and secondary rearrangements, the accessibility to limited sets of Jα segments correlates with activating histone marks such as H3K4 tri-methylation (Abarrategui and Krangel, 2006).

One possible explanation for the alteration in TCRα repertoire is that *Paxip1* deficiency reduces H3K4me3 and germline transcription throughout Jα rearrangements. If this were the case, the most profound recombination deficit might appear at the 3' end of the locus, as observed, because 3' recombination is dependent on prior 5' rearrangements (Krangel, 2007), and each step would be less efficient in the absence of *Paxip1*. To examine the effects of *Paxip1* on H3K4me3 and transcription at the 5'end of the locus without the confounding effects of ongoing rearrangements, we crossed *Paxip1*$^{-/-}$ mice with *Rag2*$^{-/-}$ mice expressing a rearranged *Tcrb* transgene. In the absence of *Rag2*, all recombination is eliminated. However, the presence of the *Tcrb* transgene allowed the *Tcra* locus to be transcriptionally active, and in a configuration whereby only 5' Jα segments were accessible. Comparison of H3K4me3 in the 5' end of the Jα array in *Rag2*$^{-/-}$ × TCRβ$_{\beta}$ × *Paxip1*$^{f/f}$ × *Lck Cre*$^{+}$ (*Rag2*$^{-/-}$ *Paxip1*$^{-/-}$) vs. *Rag2*$^{-/-}$ × TCRβ$_{\beta}$ × *Paxip1*$^{f/f}$ × *Lck Cre*$^{-}$ (*Rag2*$^{-/-}$) DP T cells revealed that PAXIP1 deficiency resulted in a 44% reduction in the number of sequence tags focused at Jα61-Jα52 (Figure 3F). In addition, there was a reduction in germline transcripts at all accessible Jα elements and an overall reduction in *Tcra* transcripts in *Rag2*$^{-/-}$ *Paxip1*$^{-/-}$ DP T cells relative to *Rag2*$^{-/-}$ (Figure 3G). These data suggest that PAXIP1 contributes to the accessibility of the *Tcra* locus by promoting H3K4me3 and germline transcription throughout the Jα cluster.

In addition to its role in transcription, PAXIP1 is recruited to sites of DNA damage where it promotes DSB repair by homologous recombination (Wang et al., 2010b). We observed that 5' Jα segments are cleaved in the absence of *Paxip1* (Figure 3E), and therefore we could assess whether DSBs are efficiently resolved. To ascertain whether *Tcra* and *Tcrd* locus integrity is compromised, we assayed for TCRα and δ associated chromosomal aberrations in metaphases spreads prepared from thymocytes. We performed FISH analysis with a *Tcra* and *Tcrd* locus probe which hybridizes to the middle of chromosome 14 together with a telomeric probe, which marks the ends of all chromosomes. Whereas TCRα and δ associated chromosomal aberrations were undetectable in WT thymocytes, 2.2% of

*Paxip1*$^{-/-}$ thymocytes showed a broken chromosome 14 evident by the *Tcra* probe hybridizing to the end of a shortened chromosome 14 (Figure 3H). This amount of TCRα and δ associated breaks is similar to that observed in mice lacking the NHEJ factor 53BP1 (Difilippantonio et al., 2008). In addition to TCRα and δ associated aberrations, there was a 4-fold increase in spontaneous DNA damage in *Paxip1*$^{-/-}$ thymoctyes (Figure 3H), consistent with a role for PAXIP1 in repairing replication-associated DSBs (Wang et al., 2010b), and with the observed decreased cellularity of *Paxip1*$^{-/-}$ thymocytes (Figure 1B and Figure S2B). We conclude that Jα rearrangements are less efficient in the absence of PAXIP1 because of impaired accessibility, cleavage, and DNA repair.

## PAXIP1 regulates NKT lineage development

Like conventional αβ T cell lineages, NK T cells arise from DP precursors in the thymus through secondary *Tcra* rearrangement, TCR-mediated signaling and selection (Godfrey and Berzins, 2007). NK T cells which express an invariant Vα14-Jα18 T cell receptor recognize lipid antigen presented by the molecule CD1d. Whereas *Paxip1* deficiency resulted in a 50% decrease in the total number of thymocytes, there was 90–95% decrease in the number of NK T cells stained by specific CD1d tetramers in the mutant thymus and spleen (n=4 mice analyzed) (Figures. 4A–C). To determine the stage of NK T cell development affected by the absence of *Paxip1* we analyzed CD1d tetramer positive thymocytes for their expression of HSA, CD44 and NK1.1. The few NKT cells in the *Paxip1*$^{-/-}$ thymus accumulated in the early immature populations (CD1d$^{tetramer+}$HSA$^{+}$CD44$^{lo}$NK1.1$^{-}$) (Figure 4A), suggesting that the block in NKT cell development is at early stages of positive selection.

To directly examine Vα14-Jα18 rearrangements, we performed quantitative RT-PCR with Vα and Jα specific primers (Gapin et al., 2001). Consistent with the fact that Jα18 is localized at the 3' end of the Jα cluster, which is under-utilized in the absence of PAXIP1 (Figures. 3A–E), we found that Vα14-Jα18 recombination was reduced 14 times in *Paxip1*$^{-/-}$ thymocytes (Figure 4D). Failure to generate the invariant Vα14-Jα18 T cell receptor provides an explanation for why NKT cell development is arrested at early stages in the absence of PAXIP1.

## PAXIP1 contributes to alterations in H3K4me3 during the DP-SP T cell transition

Next, we attempted to identify genes with deregulated H3K4me3 during the DP to SP T cell transition. To this end, we compared the ChIP-Seq of H3K4me3 in DP, CD4$^{+}$ SP, and CD8$^{+}$ SP T cell populations at gene promoters. In WT, the vast majority (27494; 98%) of H3K4me3 enriched genes were those that were shared amongst DP and SP T cells (Figure 5A, Table S1 and Figure S4A). Only 20 out of 27494 genes lost H3K4me3 during the DP to SP T cell transition (Figure 5A; inset, row S1 and Table S1) and 192 promoters gained H3K4me3 in both SP (CD4$^{+}$ and CD8$^{+}$) T cell populations (Figure 5A, inset, row S6 and Table S1). Nevertheless, a very small subset of H3K4me3 enriched genes were found uniquely in CD8$^{+}$ and CD4$^{+}$ T cells (121 and 166 genes respectively, Figure 5A, inset, row S4 and S5 and Table S1). For example *Runx3*, which is up-regulated during CD8$^{+}$ T cell differentiation (Wang et al., 2010a), showed H3K4me3 in CD8$^{+}$ but not in CD4$^{+}$ SP T cells; *Zbtb7b*, which is upregulated in CD4$^{+}$ SP T cells (He et al., 2010), exhibited H3K4me3 binding in CD4$^{+}$ but not in CD8$^{+}$ SP T cells. Nevertheless, very few genes (<2% of total) acquire H3K4me3 during the DP to SP T cell transition.

We next determined whether PAXIP1 contributed to H3K4 trimethylation by comparing the distribution of H3K4me3 in WT and *Paxip1*$^{-/-}$ sorted populations. In DP and SP T cells, the vast majority of promoters (>99.95%) showed no dependence of H3K4me3 on *Paxip1* (Figure 5B). For example, when utilizing a 3-fold threshold, only 10 genes out of 21,902 unique RefSeq gene symbols showed altered H3K4me3 binding in *Paxip1*$^{-/-}$ DP cells

(Figure 5B), and only 93 genes showed at least a 2-fold change (Fig. S4B and Table S2). The majority of H3K4me3-bound genes (3 genes in DPs, 36 genes in $CD4^+$ and 32 genes in $CD8^+$ T cells) that were regulated by PAXIP1 showed an increase in H3K4me3 in WT relative to *Paxip1*$^{-/-}$ (Figure 5B, Tables S2–4 and Figure S4B). Notable among those genes that exhibited major deficits in H3K4me3 in *Paxip1*$^{-/-}$ thymocytes was *Paxip1* itself (likely because the construct deletes the *Paxip1* promoter, transcriptional start and exon 1 (Patel et al., 2007)) and the sphingosine-1 phosphate receptor 1 *S1pr1* (Figures 5A, 5C and 5D).

To determine whether there is a correlation between the genes that showed large deficits in H3K4me3 in the absence of *Paxip1* (Figure 5B) and genes showing large alterations in H3K4me3 in WT cells during the DP to SP T cell transition (Figure 5A), we compared changes of H3K4me3 in SP *Paxip1*$^{-/-}$ relative to SP WT vs. DP WT relative to SP WT T cells (Figures 5C and 5D). We found that 20 out of the 36 (56%) *Paxip1*-dependent genes that were deregulated in $CD4^+$ SP T cells (Figure 5B) altered their amounts of H3K4me3 by more than 3 times during the DP-SP T cell transition (Figure 5C). Similarly, 19 of 32 (60%) *Paxip1*-dependent genes in $CD8^+$ SP T cells (Figure 5B) exhibited a major increase in H3K4me3 during the DP-SP T cell transition (Figure 5D). Overall, there was a greater effect from *Paxip1* deficiency on H3K4me3 for those genes that gained H3K4me3, and thereby become transcriptionally activated during DP-SP T cell differentiation, relative to those genes that did not change H3K4me3 notably (Figures 5 C and 5D, genes in quadrants I and II compared to III and IV, P(Chi-square test) < 1e-16 for both $CD4^+$ and $CD8^+$ SP T cells). Thus, despite the fact that PAXIP1 was essential for H3K4 trimethylation in less than 0.16% of promoters in SP T cells (36 out of 21902) (Figure 5B), PAXIP1 contributed disproportionately to promoters in which H3K4me3 increases during the DP-SP T cell transition (27 out of 313 DP-SP promoters or 8.6%).

## PAXIP1-dependent H3K4me3 correlates with PAXIP1 binding to promoters

To determine why only a subset of genes is a target for PAXIP1-dependent H3K4me3, we performed genome-wide localization studies using ChIP-seq for PAXIP1 in WT and *Paxip1*$^{-/-}$ thymocytes. Because PAXIP1-binding showed some background signals in *Paxip1*$^{-/-}$ thymocytes, we considered only those differentially enriched regions at gene promoters (n=20,808) in which PAXIP1-binding in WT exceeded *Paxip1* deficient by greater than 3-fold with False Discovery Rates less than 0.01 (see methods). To more rigorously assess the PAXIP1-bound DNA sequences, we examined them for the enrichment of consensus motifs identified from publicly available ChIP-Seq databases (see methods). We identified 25 motifs with more than 2-fold enrichment and log likelihood ratio greater than 3 (Table S5; see extended methods). Using this method, we identified the ZBTB16 and PAX family motifs as among the highest scoring DNA sequences (Figure 6A and Table S5). Overall, 77% of the PAXIP1-binding sites scored positive for at least one ZBTB16 consensus motif and more than 56% were positive for at least one consensus paired domain or homeobox motif. It is known that PAXIP1 and PAX1 physically associate and colocalize in active chromatin (Lechner et al., 2000), and ZBTB16 has also been reported to interact with PAXIP1 (Rual et al., 2005). These observations support the specificity of our PAXIP1 binding sequences.

The list of identified PAXIP1 bindings was then used for Gene Ontology (GO) enrichment analysis, which enabled the identification of biological processes that are affected by PAXIP1. Interestingly, the top 11 enriched processes associated with the strongest PAXIP1 targets were those involved in T cell/lymphocyte differentiation, activation and V(D)J recombination (Figure 6B and Table S6), suggestive of T-cell specific roles for PAXIP1 in thymocytes .

PAXIP1 preferentially bound to promoters and overlapped with the closest nucleosome-free region to the transcription start site, as determined by the normalized distribution of PAXIP1 and H3K4me3 around transcription start sites (TSSs) (from +/−2 kb relative to TSS) (Figure 6C). On average, 91% of promoters occupied by H3K4me3 had PAXIP1 binding, and conversely 90% of PAXIP1-occupied promoters were also bound by H3K4me3. Given the strong positive effect of PAXIP1 on differential H3K4me3 and expression of several genes that were up-regulated during the DP-SP T cell transition (Figure 5), we asked whether we could discern the global effect of *Paxip1* on cognate sites of H3K4me3. By comparing the genome-wide localization of PAXIP1 to H3K4me3, we found that increased PAXIP1 binding intensity correlated with the extent of H3K4 trimethylation (Figure S5A and S5B). Specifically, those genes in SP lineages which showed the greatest deficits in H3K4me3 in the absence of PAXIP1 were associated with significant PAXIP1binding in WT T cells (Figures S5A and S5B).

To further determine the frequency of PAXIP1-dependent genes that exhibited direct PAXIP1 binding, we identified the number of PAXIP1-occupied promoters that had at least 2-fold increase in PAXIP1 dependent H3K4 tri-methylation (as determined in Figure S4B). We compared this to control groups consisting of the same number of genes taken from the next most relevant PAXIP1-dependent promoters with the highest increase in H3K4me3 not exceeding 2-fold. We observed 83%, 63% and 25% increase in direct PAXIP1 association with the most significant PAXIP1-dependent promoters in $CD8^+$ SP, $CD4^+$ SP, and DP T cell populations respectively relative to the controls (Figure 6D). Therefore, PAXIP1-dependent increases in H3K4me3 frequently result from direct binding of the PAXIP1 associated methyltransferase complex near the promoter.

## PAXIP1 affects thymocyte egress by regulating S1PR1 expression

Since $CD4^+$ and $CD8^+$ SP T cells accumulate in PAXIP1-deficient thymocytes (Figure 1), we hypothesized that genes involved in cell migration might be deregulated in both populations. We therefore compared the distribution of H3K4me3 in sorted $CD4^+$ SP and $CD8^+$ SP T cells from WT and *Paxip1*$^{-/-}$ thymocytes. Relative to *Paxip1*$^{-/-}$ $CD4^+$ T cells, 36 genes from WT $CD4^+$ T cells had increased expression of H3K4me3 near their promoters (Figure 5B and Figure 7A). 32 genes in WT $CD8^+$ T cells also had increased H3K4me3 deposition relative to *Paxip1*$^{-/-}$ $CD8^+$ thymocytes (Figure 5B and Figure 7A). Among these genes with deregulated H3K4me3 at their promoters, only 9 exhibited an H3K4me3 deficit in both *Paxip1*$^{-/-}$ $CD4^+$ and $CD8^+$ T lineages which accumulate in the mutant thymus (Figure 7A).

Examination of the list of common ($CD4^+$ and $CD8^+$ T cell) down-regulated genes in *Paxip1*$^{-/-}$ SP T cells revealed that the most important alteration, after *Paxip1* itself, was in the sphingosine-1-phosphate receptor *S1pr1* (Figure 7B and Table S7), which is essential for both αβ and γδ thymocyte and peripheral T cell emigration (Allende et al., 2004; Matloubian et al., 2004; Odumade et al., 2010). We observed abundant signal for H3K4me3 near *S1pr1* for both $CD4^+$ and $CD8^+$ T cell lineages, whereas H3K4me3 was barely detectable in the absence of *Paxip1* (Figure S6A). Another factor implicated in T cell homeostatis was IL6st (Figure 7B), which encodes the IL-6 receptor subunit, gp130. Surface staining for gp130 revealed that *Paxip1*$^{-/-}$ thymocytes were severely impaired in gp130 expression (Figure S6B).

In addition to S1PR1, the only other factors that have been implicated in thymocyte emigration are FOXO1 and KLF2 DNA binding proteins (Freitas and Rocha, 2009). Expression of all three factors (Figure 7C) and H3K4me3 association near their promoters (Figure S4A) is up-regulated during the DP-SP T cell transition; moreover, these gene products appeared to function in the same pathway, since FOXO1 is critical for KLF2

expression, which in turn drives S1PR1 expression (Freitas and Rocha, 2009; Kerdiles et al., 2009). To determine whether PAXIP1 promotes the expression of transcription factors critical for thymocyte emigration, we directly assayed *Foxo1*, *Klf2* and *S1pr1* expression in DP and SP cells by quantitative RT-PCR. As predicted, all three genes were up-regulated during the DP-SP transition in WT cells. Similarly, *Foxo1* and *Klf2* were also up-regulated in PAXIP1-deficient cells; however, S1PR1 RNA and protein failed to be induced in *Paxip1*$^{-/-}$ SP thymocytes (Figures 7C and 7D).

We conclude that amongst the factors known to drive thymocyte egress and the size of the peripheral pool, S1PR1 appears to be uniquely regulated by PAXIP1. Moreover, PAXIP1 bound directly to the S1PR1 promoter (Figure S5C). S1PR1 deficiency is characterized by a greater number of SP cells expressing β7 integrin and intermediate amount of CD69 in SP thymocytes (Matloubian et al., 2004), and *Paxip1*$^{-/-}$ SP thymocytes mirrored this phenotype (Figure 7E). Additionally, we found a profound accumulation of HSA$^{lo}$ mature γδ T cells in the thymus which also require PAXIP1 for thymic export (Odumade et al., 2010)(Figure S2D). Together, these data demonstrate that among several genes up-regulated in mature SP thymocytes, PAXIP1 is essential for induction of *S1pr1* during the DP to SP T cell transition.

In addition to *S1pr1* and *Paxip1*, there are 7 genes in which H3K4me3 is mis-regulated by more than 3-fold and 74 genes decreased by at least 50% in *Paxip1*$^{-/-}$ mature SP thymocytes (Figures 7A and 7B; Table S7). To determine whether failure to up-regulate *S1pr1* contributes to the accumulation of mature thymoctyes, we crossed *Paxip1*-deficient mice with mice expressing high amounts of an S1PR1 transgene under control of the human *Cd2* promoter (Liu et al., 2009). In contrast to *Paxip1* deficiency, SP thymocytes no longer accumulated in S1PR1Tg LckCre *Paxip1*$^{f/f}$ (*Paxip1*$^{-/-}$S1PR1Tg) mice (Figure 7F), and the frequency of CD4$^{+}$ and CD8$^{+}$ SP T cells in *Paxip1*$^{-/-}$S1PR1Tg was even lower than in WT. Moreover, S1PR1 overexpression normalized *Paxip1*$^{-/-}$ SP maturation, evidenced by a decrease in frequency of *Paxip1*$^{-/-}$S1PR1Tg cells expressing CD69 and HSA, and an increase in SP cells expressing β7-integrin relative to *Paxip1*$^{-/-}$ thymocytes (Figure 7F and Figure 1C). We conclude that among various deregulated genes in SP thymocytes, defective S1PR1 expression is a primary factor influencing the accumulation of T cells in the *Paxip1* deficient thymus.

## Discussion

Here we have shown that the MLL3 and 4 specific co-factor PAXIP1 regulates H3K4me3 in a subset of genes during thymocyte development. PAXIP1 controlled accessibility in DP thymocytes by promoting H3K4me3, transcription and cleavage within Jα chromatin. Based on these findings, it is likely that RAG2 binding to the Jα locus is also reduced in the absence of *Paxip1*; this could be the result of defective transcription and associated chromatin opening as well as from diminished interaction between the RAG2 PHD domain and H3K4me3. In any case, the PAXIP1-associated activity cannot be the only methyltransferase that targets RAG to the *Tcra* locus since the reduction in H3K4me3 and transcription was only 2-times lower in *Paxip1*$^{-/-}$ thymocytes relative to WT. A consequence of this redundancy is that Jα rearrangements are not eliminated in the absence of PAXIP1 but instead are focused to the 5' end of the Jα locus. Nevertheless, NK T cell development is blocked due to failure to recombine to the Jα18 segment which lies at the 3' end of the *Tcra* locus.

While the 5' end of the Jα locus is efficiently broken, it is inefficiently repaired, evidenced by the accumulation of TCRα and δ associated chromosomal aberrations in *Paxip1*$^{-/-}$ thymocytes. RAG mediated breaks are repaired by the NHEJ pathway. Interestingly,

PAXIP1 has been implicated in promoting homologous recombination rather than NHEJ (Wang et al., 2010b). Our ChIP analysis of genomic regions bound by PAXIP1 revealed that PAXIP1 associates near promoters (within 2kb of TSSs) of known NHEJ factors including XRCC4, DNA-PKcs and XRCC6 (Ku80). This suggests the possibility that PAXIP1's role in *Tcra* recombination might be linked to transcriptional regulation of NHEJ factors. However, we have found no difference in the expression of NHEJ factors in WT and *Paxip1*$^{-/-}$ thymocytes as measured by quantitative RT-PCR (not shown). Since PAXIP1 is recruited to DSBs in a MLL4-independent but γ-H2AX-dependent manner (Gong et al., 2009), it therefore seems more likely that PAXIP1's role in repair of RAG mediated DSBs is directly mediated by recruitment to DSBs rather than transcription of NHEJ proteins.

PAXIP1 regulates H3K4me3 in a small fraction of promoters in developing T cells. Among the genes that are positively regulated by PAXIP1 in CD4$^{+}$ or CD8$^{+}$ SP T cells, approximately 50% are those that changed H3K4me3 notably during the DP-SP T cell transition. Moreover, many of the genes that showed PAXIP1 dependence for H3K4me3 also exhibited PAXIP1 binding near their promoters. This indicates that PAXIP1 activity is most important for a subset of T cell specific genes whose expression is induced during DP to SP T cell differentiation. This includes *S1pr1*, but does not include other up-regulated genes such as *Bcl2*, *Sell* and *Il7r*.

It has been reported that PAX2 and PAX5 recruit PAXIP1 and the H3K4 methyltransferase complex to PAX-dependent genes (Patel et al., 2007; Schwab et al., 2011). If this mode of recruitment is generalizable, it suggests that complexes containing PAXIP1 and methyltransferases are targeted to DNA through interactions with sequence specific transcription factors. According to this hypothesis, transcription factors such as FOXO1 or KLF2 might facilitate association of PAXIP1 with the *S1pr1* promoter, which in turn recruits RNA polymerase II through ensuing H3K4me3. Other genes such as *Bcl2*, *Sell* and *Il7r* would be induced in a PAXIP1-independent manner because transcription factors that activate these promoters would not interact with PAXIP1.

Is there a connection between PAXIP1 functions in *Tcra* rearrangement and thymoctye export? One possibility is that PAXIP1 serves as a gatekeeper that ensures that developing thymocytes with faulty rearrangements are not exported to the periphery. Previously we have shown that in addition to its role in DNA double strand break repair, the ATM kinase prevents the transmission of DNA breaks to daughter cells (Callen et al., 2007). Impairment of both these repair and checkpoint functions may account for the high amount of antigen receptor associated breaks in mature peripheral *Atm*$^{-/-}$ lymphocytes, some of which are generated in earlier development as a result of failed V(D)J recombination (Callen et al., 2007). By promoting phosphorylation of H2AX, ATM is also required for the recruitment of PAXIP1 to DSB sites (Gong et al., 2009). PAXIP1 accumulation at DNA damage sites occurs independently of the MLL3 and 4 methyltransferase complex (Gong et al., 2009). Nevertheless, PAXIP1 association with sites of DNA damage during *Tcra* rearrangements might effectively deplete MLL3 and 4 of its transcription targeting co-factor PAXIP1. In this case, there would be less efficient induction of PAXIP1-dependent genes such as S1PR1 during the DNA damage response, which in turn would reduce thymic export. These two specialized roles of PAXIP1 in DNA repair and transcription might therefore have evolved as part of a checkpoint that prevents the propagation of potentially oncogenic DNA damage out of the thymus.

## Experimental Procedures

### Mice

*Paxip1*$^{f/f}$ (Kim et al., 2007), TCRβ transgenic (Shinkai et al., 1993), and Lck Cre transgenic (Hennet et al., 1995) and S1PR1 transgenic (Liu et al., 2009) mice were generated as previously described. *Rag2*$^{-/-}$ mice were obtained from Taconic Laboratories. All experiments were performed in compliance with NIH Intramural Animal Care and Use program.

### TCRα repertoire and recombination

Analysis of Jα usage in sorted TCRβ$^{lo/med}$ DP thymocytes was performed by Southern blot and probing of TRAV12 (Vα8)-Cα RT-PCR products with Jα-specific probes as described (Guo et al 2002; Hawwari et al. 2005). TCRα transcripts were analyzed by quantitative real-time PCR using a QuantiFast SYBR Green PCR kit (Qiagen). TCRα coding joints were analyzed in genomic DNA of sorted DP thymocytes by quantitative real-time PCR as above, with normalization to *B2m*. Primer sequences are provided in Table S8. Jα DSBs were quantified in genomic DNA of sorted DP thymocytes by ligation-mediated PCR (McMurry et al 1997) using primers and probes described (Seitan et al., 2011). Metaphases were obtained from thymocytes stimulated with anti-TCR (H57, 2 μg/ml; PharMingen) and anti-CD28 antibodies (5 μg/ml; PharMingen) and analyzed with BAC probes containing the TCRα locus (TCR Cα-232F18) and telomere-repeat specific peptide nucleic acid (PNA) probes (Applied Biosystems).

### Quantitative RT-PCR analysis of gene expression

Total RNA was extracted with TRIzol (Invitrogen) or RNeasy mini kit (Qiagen) and was reverse transcribed with Superscript III cDNA synth (Invitrogen). Quantitative RT-PCR was performed using SYBR Green (Perkin Elmer) with the 7900HT Fast Real-time PCR system (Applied Biosystems). Primers used are listed in Extended Experimental Procedures.

### Flow cytometry

Single cell suspensions from thymocytes, lymph nodes and spleen were isolated from 6–10 week old mice. Antibodies used for flow cytometric analysis are listed in Extended Experimental Procedures.

### Chromatin immunoprecipitation

At least 10 million sorted DP (CD4+CD8+TCR$^{lo/med}$), CD4$^{+}$ SP (CD4+CD8-TCR$^{hi}$) and CD8$^{+}$ SP (CD4$^{-}$CD8$^{+}$TCR$^{hi}$) populations was used to make cross-linked chromatin. Chromatin was processed for H3K4me3 or PAXIP1 ChIP-Seq with Illumina sequencing as described (Daniel et al., 2010).

### Computational and Statistical Analyses

Detailed description is available in the Extended Experimental Procedures.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We thank Gustavo Gutierrez-Cruz for technical assistance, Susan Sharrow, Larry Granger and Tony Adams for flow cytometry and Remy Bosselut for discussions. This work was supported by the Intramural Research Program

of the NIH, National Cancer Institute, and Center for Cancer Research to A.N. and J-H.P, and a Senior Scholar award from the Ellison Medical Foundation to A.N. M.S.K. was supported by NIH grant R37 GM41052.

## References

Abarrategui I, Krangel MS. Regulation of T cell receptor-alpha gene recombination by transcription. Nature immunology. 2006; 7:1109–1115. [PubMed: 16936730]

Allende ML, Dreier JL, Mandala S, Proia RL. Expression of the sphingosine 1-phosphate receptor, S1P1, on T-cells controls thymic emigration. The Journal of biological chemistry. 2004; 279:15396–15401. [PubMed: 14732704]

Callen E, Jankovic M, Difilippantonio S, Daniel JA, Chen HT, Celeste A, Pellegrini M, McBride K, Wangsa D, Bredemeyer AL, et al. ATM prevents the persistence and propagation of chromosome breaks in lymphocytes. Cell. 2007; 130:63–75. [PubMed: 17599403]

Chen HT, Bhandoola A, Difilippantonio MJ, Zhu J, Brown MJ, Tai X, Rogakou EP, Brotz TM, Bonner WM, Ried T, Nussenzweig A. Response to RAG-mediated VDJ cleavage by NBS1 and gamma-H2AX. Science. 2000; 290:1962–1965. [PubMed: 11110662]

Cho YW, Hong S, Ge K. Affinity purification of MLL3/MLL4 histone H3K4 methyltransferase complex. Methods Mol Biol. 2012; 809:465–472. [PubMed: 22113294]

Cho YW, Hong T, Hong S, Guo H, Yu H, Kim D, Guszczynski T, Dressler GR, Copeland TD, Kalkum M, Ge K. PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. The Journal of biological chemistry. 2007; 282:20395–20406. [PubMed: 17500065]

Crompton T, Moore M, MacDonald HR, Malissen B. Double-negative thymocyte subsets in CD3 zeta chain-deficient mice: absence of HSA+CD44-CD25− cells. European journal of immunology. 1994; 24:1903–1907. [PubMed: 7520000]

Daniel JA, Santos MA, Wang Z, Zang C, Schwab KR, Jankovic M, Filsuf D, Chen HT, Gazumyan A, Yamane A, et al. PTIP promotes chromatin changes critical for immunoglobulin class switch recombination. Science. 2010; 329:917–923. [PubMed: 20671152]

Difilippantonio S, Gapud E, Wong N, Huang CY, Mahowald G, Chen HT, Kruhlak MJ, Callen E, Livak F, Nussenzweig MC, et al. 53BP1 facilitates long-range DNA end-joining during V(D)J recombination. Nature. 2008; 456:529–533. [PubMed: 18931658]

Freitas AA, Rocha B. Homeostasis of naive T cells: the Foxo that fixes. Nature immunology. 2009; 10:133–134. [PubMed: 19148194]

Gapin L, Matsuda JL, Surh CD, Kronenberg M. NKT cells derive from double-positive thymocytes that are positively selected by CD1d. Nature immunology. 2001; 2:971–978. [PubMed: 11550008]

Gellert M. V(D)J recombination: RAG proteins, repair factors, and regulation. Annual review of biochemistry. 2002; 71:101–132.

Godfrey DI, Berzins SP. Control points in NKT-cell development. Nature reviews. Immunology. 2007; 7:505–518. [PubMed: 17589542]

Gong Z, Cho YW, Kim JE, Ge K, Chen J. Accumulation of Pax2 transactivation domain interaction protein (PTIP) at sites of DNA breaks via RNF8-dependent pathway is required for cell survival after DNA damage. The Journal of biological chemistry. 2009; 284:7284–7293. [PubMed: 19124460]

Grundy GJ, Yang W, Gellert M. Autoinhibition of DNA cleavage mediated by RAG1 and RAG2 is overcome by an epigenetic signal in V(D)J recombination. Proceedings of the National Academy of Sciences of the United States of America. 2010; 107:22487–22492. [PubMed: 21149691]

Guo J, Hawwari A, Li H, Sun Z, Mahanta SK, Littman DR, Krangel MS, He YW. Regulation of the TCRalpha repertoire by the survival window of CD4(+)CD8(+) thymocytes. Nature immunology. 2002; 3:469–476. [PubMed: 11967541]

Hawwari A, Bock C, Krangel MS. Regulation of T cell receptor alpha gene assembly by a complex hierarchy of germline Jalpha promoters. Nature immunology. 2005; 6:481–489. [PubMed: 15806105]

Hawwari A, Krangel MS. Role for rearranged variable gene segments in directing secondary T cell receptor alpha recombination. Proceedings of the National Academy of Sciences of the United States of America. 2007; 104:903–907. [PubMed: 17210914]

He X, Park K, Kappes DJ. The role of ThPOK in control of CD4/CD8 lineage commitment. Annual review of immunology. 2010; 28:295–320.

Hennet T, Hagen FK, Tabak LA, Marth JD. T-cell-specific deletion of a polypeptide N-acetylgalactosaminyl-transferase gene by site-directed recombination. Proceedings of the National Academy of Sciences of the United States of America. 1995; 92:12070–12074. [PubMed: 8618846]

Issaeva I, Zonis Y, Rozovskaia T, Orlovsky K, Croce CM, Nakamura T, Mazo A, Eisenbach L, Canaani E. Knockdown of ALR (MLL2) reveals ALR target genes and leads to alterations in cell adhesion and growth. Molecular and cellular biology. 2007; 27:1889–1903. [PubMed: 17178841]

Ji Y, Resch W, Corbett E, Yamane A, Casellas R, Schatz DG. The in vivo pattern of binding of RAG1 and RAG2 to antigen receptor loci. Cell. 2010; 141:419–431. [PubMed: 20398922]

Kerdiles YM, Beisner DR, Tinoco R, Dejean AS, Castrillon DH, DePinho RA, Hedrick SM. Foxo1 links homing and survival of naive T cells by regulating L-selectin, CCR7 and interleukin 7 receptor. Nature immunology. 2009; 10:176–184. [PubMed: 19136962]

Kim D, Wang M, Cai Q, Brooks H, Dressler GR. Pax transactivation-domain interacting protein is required for urine concentration and osmotolerance in collecting duct epithelia. Journal of the American Society of Nephrology : JASN. 2007; 18:1458–1465. [PubMed: 17429055]

Krangel MS. T cell development: better living through chromatin. Nature immunology. 2007; 8:687–694. [PubMed: 17579647]

Lechner MS, Levitan I, Dressler GR. PTIP, a novel BRCT domain-containing protein interacts with Pax2 and is associated with active chromatin. Nucleic acids research. 2000; 28:2741–2751. [PubMed: 10908331]

Liu G, Burns S, Huang G, Boyd K, Proia RL, Flavell RA, Chi H. The receptor S1P1 overrides regulatory T cell-mediated immune suppression through Akt-mTOR. Nature immunology. 2009; 10:769–777. [PubMed: 19483717]

Liu Y, Subrahmanyam R, Chakraborty T, Sen R, Desiderio S. A plant homeodomain in RAG-2 that binds Hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. Immunity. 2007; 27:561–571. [PubMed: 17936034]

Matloubian M, Lo CG, Cinamon G, Lesneski MJ, Xu Y, Brinkmann V, Allende ML, Proia RL, Cyster JG. Lymphocyte egress from thymus and peripheral lymphoid organs is dependent on S1P receptor 1. Nature. 2004; 427:355–360. [PubMed: 14737169]

Matthews AG, Kuo AJ, Ramon-Maiques S, Han S, Champagne KS, Ivanov D, Gallardo M, Carney D, Cheung P, Ciccone DN, et al. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. Nature. 2007; 450:1106–1110. [PubMed: 18033247]

Matys V, Fricke E, Geffers R, Gossling E, Haubrock M, Hehl R, Hornischer K, Karas D, Kel AE, Kel-Margoulis OV, et al. TRANSFAC: transcriptional regulation, from patterns to profiles. Nucleic acids research. 2003; 31:374–378. [PubMed: 12520026]

McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. GREAT improves functional interpretation of cisregulatory regions. Nature biotechnology. 2010; 28:495–501.

McMurry MT, Hernandez-Munain C, Lauzurica P, Krangel MS. Enhancer control of local accessibility to V(D)J recombinase. Molecular and cellular biology. 1997; 17:4553–4561. [PubMed: 9234713]

Odumade OA, Weinreich MA, Jameson SC, Hogquist KA. Kruppel-like factor 2 regulates trafficking and homeostasis of gammadelta T cells. J Immunol. 2010; 184:6060–6066. [PubMed: 20427763]

Osipovich O, Oltz EM. Regulation of antigen receptor gene assembly by genetic-epigenetic crosstalk. Seminars in immunology. 2010; 22:313–322. [PubMed: 20829065]

Patel SR, Kim D, Levitan I, Dressler GR. The BRCT-domain containing protein PTIP links PAX2 to a histone H3, lysine 4 methyltransferase complex. Developmental cell. 2007; 13:580–592. [PubMed: 17925232]

Petersen S, Casellas R, Reina-San-Martin B, Chen HT, Difilippantonio MJ, Wilson PC, Hanitsch L, Celeste A, Muramatsu M, Pilch DR, et al. AID is required to initiate Nbs1/gamma-H2AX focus formation and mutations at sites of class switching. Nature. 2001; 414:660–665. [PubMed: 11740565]

Portales-Casamar E, Kirov S, Lim J, Lithwick S, Swanson MI, Ticoll A, Snoddy J, Wasserman WW. PAZAR: a framework for collection and dissemination of cis-regulatory sequence annotation. Genome biology. 2007; 8:R207. [PubMed: 17916232]

Rual JF, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, Li N, Berriz GF, Gibbons FD, Dreze M, Ayivi-Guedehoussou N, et al. Towards a proteome-scale map of the human protein-protein interaction network. Nature. 2005; 437:1173–1178. [PubMed: 16189514]

Schatz DG, Swanson PC. V(D)J recombination: mechanisms of initiation. Annual review of genetics. 2011; 45:167–202.

Schwab KR, Patel SR, Dressler GR. Role of PTIP in class switch recombination and long-range chromatin interactions at the immunoglobulin heavy chain locus. Molecular and cellular biology. 2011; 31:1503–1511. [PubMed: 21282469]

Seitan VC, Hao B, Tachibana-Konwalski K, Lavagnolli T, Mira-Bontenbal H, Brown KE, Teng G, Carroll T, Terry A, Horan K, et al. A role for cohesin in T-cell-receptor rearrangement and thymocyte differentiation. Nature. 2011; 476:467–471. [PubMed: 21832993]

Shimazaki N, Tsai AG, Lieber MR. H3K4me3 stimulates the V(D)J RAG complex for both nicking and hairpinning in trans in addition to tethering in cis: implications for translocations. Molecular cell. 2009; 34:535–544. [PubMed: 19524534]

Shinkai Y, Koyasu S, Nakayama K, Murphy KM, Loh DY, Reinherz EL, Alt FW. Restoration of T cell development in RAG-2-deficient mice by functional TCR transgenes. Science. 1993; 259:822–825. [PubMed: 8430336]

Stavnezer J, Guikema JE, Schrader CE. Mechanism and regulation of class switch recombination. Annual review of immunology. 2008; 26:261–292.

Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nature structural & molecular biology. 2007; 14:1025–1040.

Wang L, Xiong Y, Bosselut R. Tenuous paths in unexplored territory: From T cell receptor signaling to effector gene expression during thymocyte selection. Seminars in immunology. 2010a; 22:294–302. [PubMed: 20537906]

Wang X, Takenaka K, Takeda S. PTIP promotes DNA double-strand break repair through homologous recombination. Genes to cells : devoted to molecular & cellular mechanisms. 2010b

Yancopoulos GD, Alt FW. Developmentally controlled and tissue-specific expression of unrearranged VH gene segments. Cell. 1985; 40:271–281. [PubMed: 2578321]

Yannoutsos N, Wilson P, Yu W, Chen HT, Nussenzweig A, Petrie H, Nussenzweig MC. The role of recombination activating gene (RAG) reinduction in thymocyte development in vivo. The Journal of experimental medicine. 2001; 194:471–480. [PubMed: 11514603]

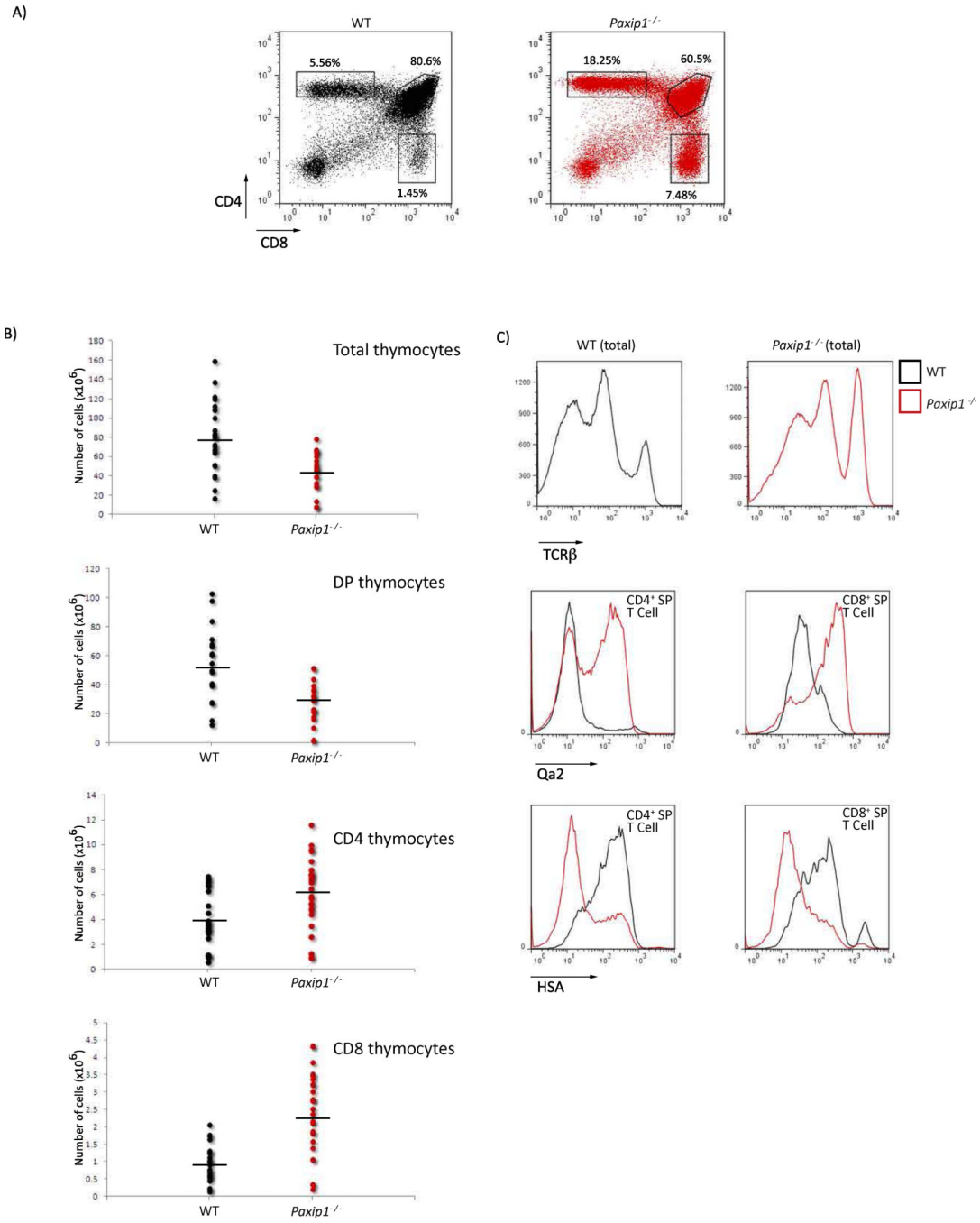



**Figure 1. PAXIP1 regulates thymocyte development**

(A) Flow cytometric analysis of WT and *Paxip1*$^{-/-}$ thymocytes stained with antibodies against CD4 and CD8. (B) Graphs showing numbers of total, DPs, CD4$^{+}$ and CD8$^{+}$ thymocytes from WT (black) and *Paxip1*$^{-/-}$ (red) mice. Average number in each population is indicated by the dashed line. (C) TCRβ, Qa2 and HSA expression in CD4$^{+}$ and CD8$^{+}$ cells from WT (black) and *Paxip1*$^{-/-}$ thymocytes. Same number of total thymocytes from WT and *Paxip1*$^{-/-}$ were collected for flow cytometry.

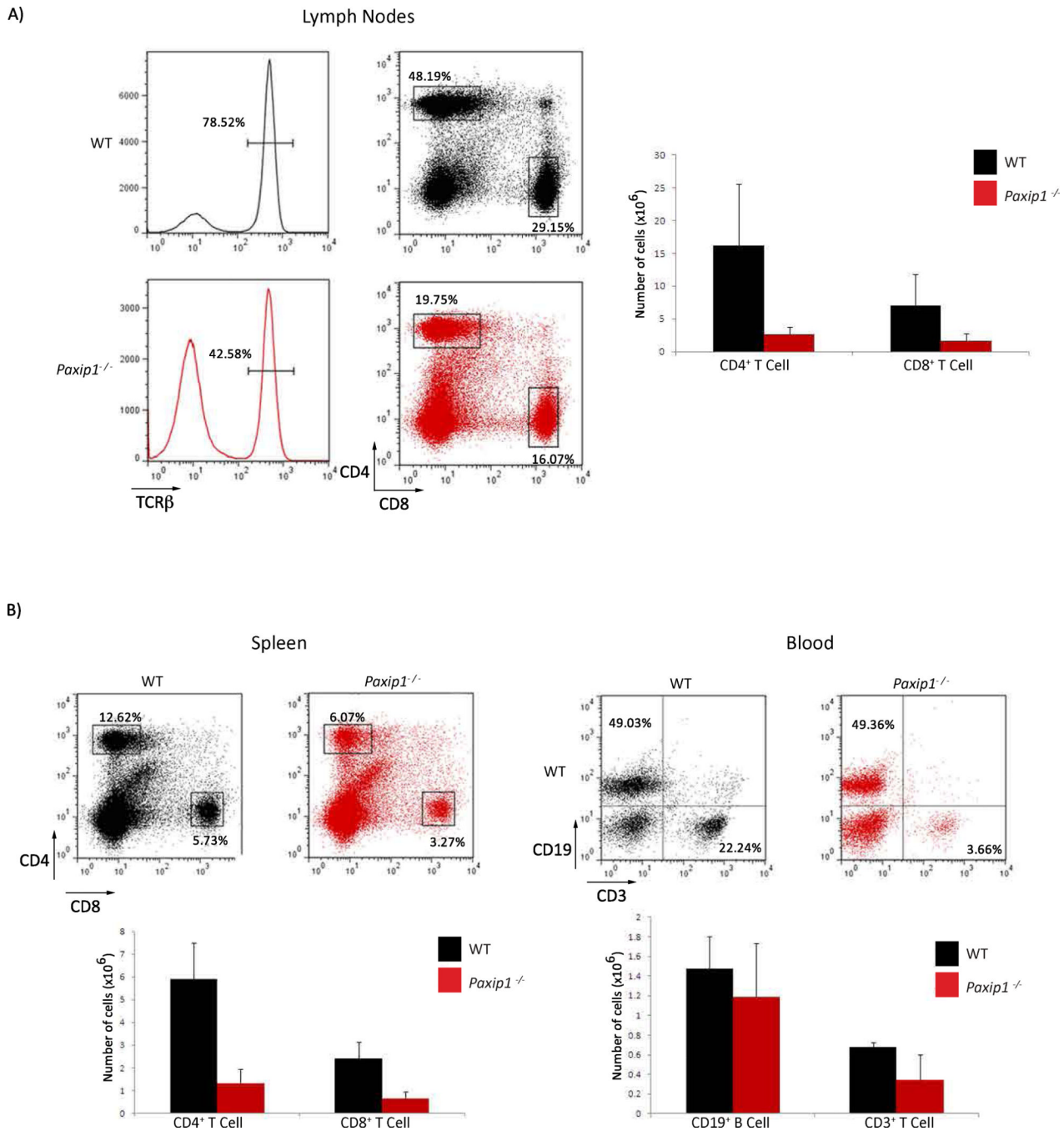



**Figure 2. Decreased number of peripheral T cells in the absence of PAXIP1**

(A) Left panel, flow cytometric analysis of WT and *Paxip1*$^{-/-}$ lymph node shows a decreased frequency of TCRβ$^{+}$, CD4$^{+}$ and CD8$^{+}$ SP cells in *Paxip1*$^{-/-}$ thymocytes. Right panel, bar graph shows the average number of T cells in WT and *Paxip1*$^{-/-}$ lymph node (n=16 mice). (B) Frequency of CD4$^{+}$ and CD8$^{+}$ T cells in spleen (left) and percentage of CD19$^{+}$ and CD3$^{+}$ lymphocytes in blood respectively was detected by flow cytometry. Representative examples are shown. Bar graphs (bottom) show the average number of CD4$^{+}$ and CD8$^{+}$ SP T cells in spleen (n=9 mice) and CD19$^{+}$ and CD3$^{+}$ cells in blood (n=5 mice) from WT and *Paxip1*$^{-/-}$ mice. Error bars represent standard deviation.

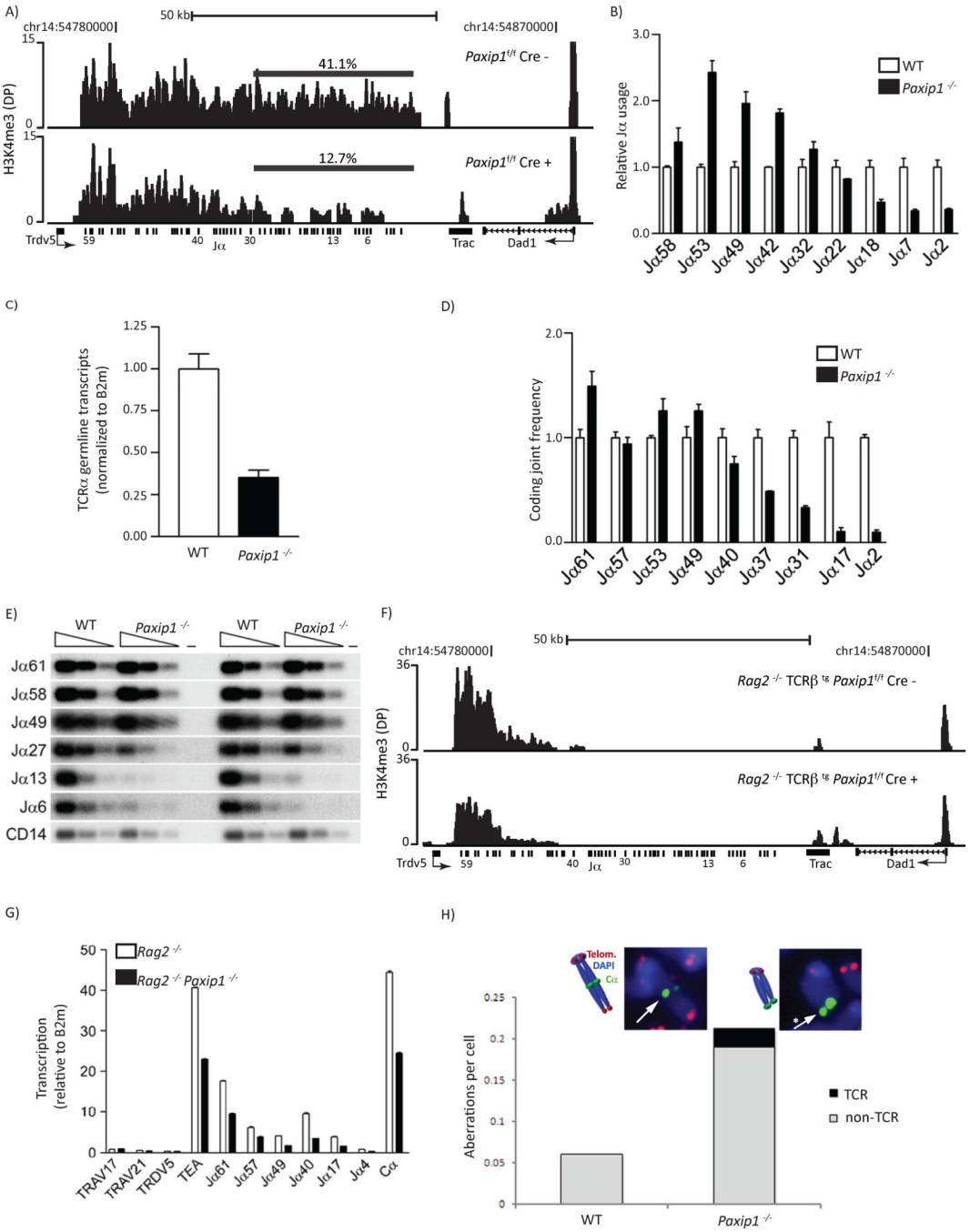



**Figure 3. PAXIP1 regulates *Tcra* recombination**

(A) Loss of PAXIP1 decreases H3K4me3 at 3' Jα segments in DP T cells. Gene tracks represent position of H3K4me3 across a region spanning Jα segments in WT (top) and *Paxip1*$^{-/-}$ (bottom). The x-axis represents the linear sequence of genomic DNA, and the y-axis represents the total number of mapped reads per million of mapped reads in 200 nucleotide windows. The genomic scale in kilobases (kb) is indicated above each track. (B) Jα usage in *Paxip1*$^{-/-}$ DP thymocytes relative to WT normalized to 1 for each Jα (mean ±SEM of two independent experiments). (C) Expression of TCR Cα transcripts in WT and *Paxip1*$^{-/-}$ DP T cells (mean±SEM of two independent experiments). (D) TRAV12 (Vα8)-Jα coding joints in *Paxip1*$^{-/-}$ DP thymocyte genomic DNA relative to WT normalized to 1

for each Jα (mean±SEM of two independent experiments). (E) DSBs were detected in serial 3-fold dilutions of genomic DNA isolated from WT and *Paxip1*$^{-/-}$ DP thymocytes and were detected by Southern blot. Results of two independent experiments are shown. (F) H3K4me3 ChIPseq profiles of the *Tcra* locus from *Rag2*$^{-/-}$ × TCRβ × *Paxip1*$^{f/f}$ × Lck Cre$^{+}$ and *Rag2*$^{-/-}$ × TCRβ × *Paxip1*$^{f/f}$ × Lck Cre$^{-}$ DP T cells show decreased tag counts in absence of PAXIP1. (G) Germline transcript abundance in *Rag2*$^{-/-}$ *Paxip1*$^{-/-}$ compared with *Rag2*$^{-/-}$ DPs as measured by real-time PCR. (H) Genomic instability analysis of metaphase spreads from thymocytes stimulated with anti-CD28 and anti-TCRβ for 2 days. Aberrations associated with the TCR Cα locus at chromosome 14 are represented in black. General instability outside the TCR Cα locus is represented in grey. Schematic and representative images of normal and broken chromosome 14 at the TCRαδ locus are shown. Arrow points to TCRαδ locus (green), telomere is shown in red, and DAPI counterstain is in blue.

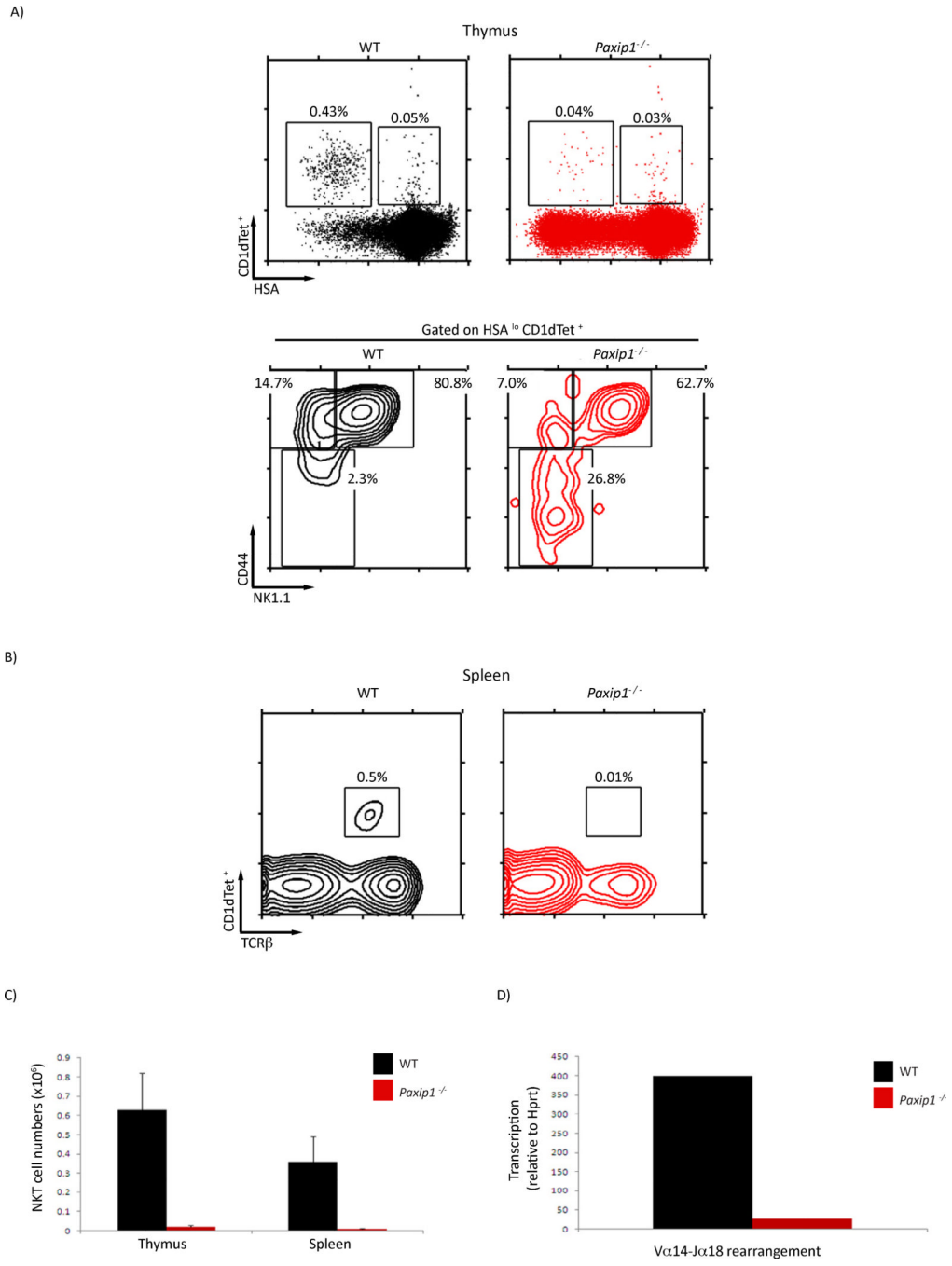



**Figure 4. PAXIP1 is essential for NK T cell development and differentiation**

(A) Top panel, flow cytometry of WT and *Paxip1*$^{-/-}$ thymocytes stained with PBS-57 loaded CD1d-tetramer (CD1dTet$^{+}$) and anti-CD24 (HSA). Bottom panel, CD44 and NK1.1 expression on gated HSA$^{lo}$CD1dTet$^{+}$ thymocytes. (B) Splenocytes stained with CD1dTet$^{+}$ and anti-TCRβ. (C) Quantitation of NK T cell numbers in thymus (n=4 mice analyzed for each genotype). (D) Q-RTPCR analysis of Vα14-Jα18 rearrangements in WT and *Paxip1*$^{-/-}$ thymocytes.

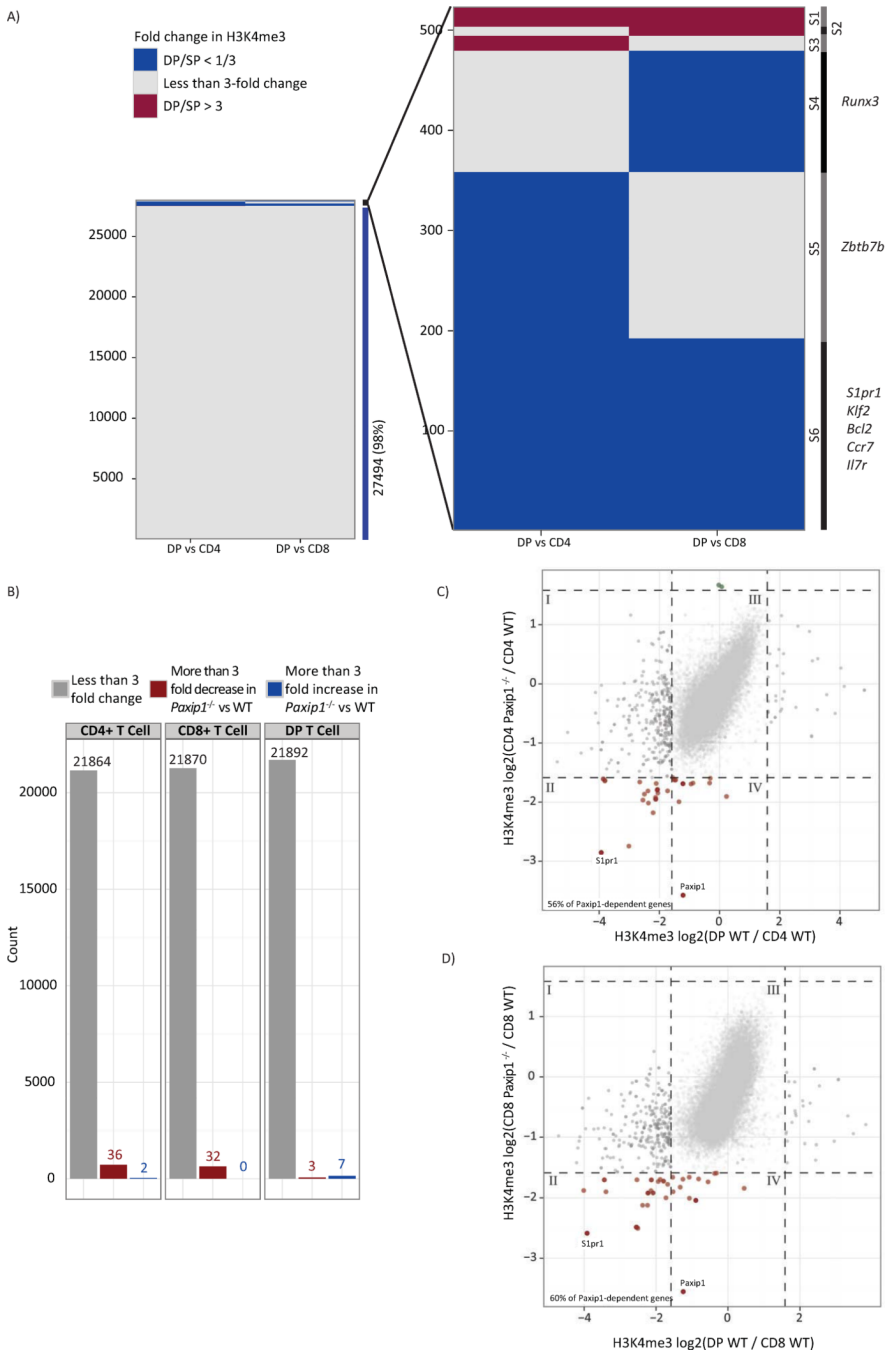



**Figure 5. PAXIP1 dependent H3K4me3 at promoters during thymocyte development**

(A) Differences in the distribution of H3K4me3 at promoters in DP and SP T cells. The fold change in H3K4me3 normalized tags per million reads (TPM) at the promoter of each gene in NCBI Build 37 RefSeq is indicated by color (left). The subset of genes with more than 3 fold change in H3K4me3 TPM is magnified (right), and representative genes for 6 identified categories are enumerated next to the color side bar (right). (B) PAXIP1 deficiency contributes to deregulation of H3K4me3 in less than 0.3% of genes. The count of gene symbols unaffected (grey), and positively (red) or negatively (blue) regulated by PAXIP1 is shown for $CD4^+$ SP, $CD8^+$ SP and DP T cells. Gene is designated as PAXIP1-dependent when H3K4me3 TPM at the promoter is changed by at least 3 fold. (C, D) Gain of

H3K4me3 at promoters during WT DP to SP T cell development correlates with deficit in H3K4me3 in the absence of PAXIP1 in SP T cells. Scatter plot of $CD4^+$ SP T cell and $CD8^+$ SP T cell(C and D respectively) show PAXIP1 - dependent H3K4me3 TMP fold enrichment (y-axis) versus difference in H3K4me3 TPM between DP WT and $CD4^+$ SP WT or $CD8^+$ SP WT T cells(C and D respectively) (x-axis). Each point depicts a gene's promoter in NCBI Build 37 RefSeq. Genes with more than 3 times reduction or increase in SP *Paxip1*$^{-/-}$ vs. SP WT T cells are depicted in red or green, respectively. The 3-fold enrichment boundary is marked by dashed line. Promoter is defined as +/− 5Kb from TSS.

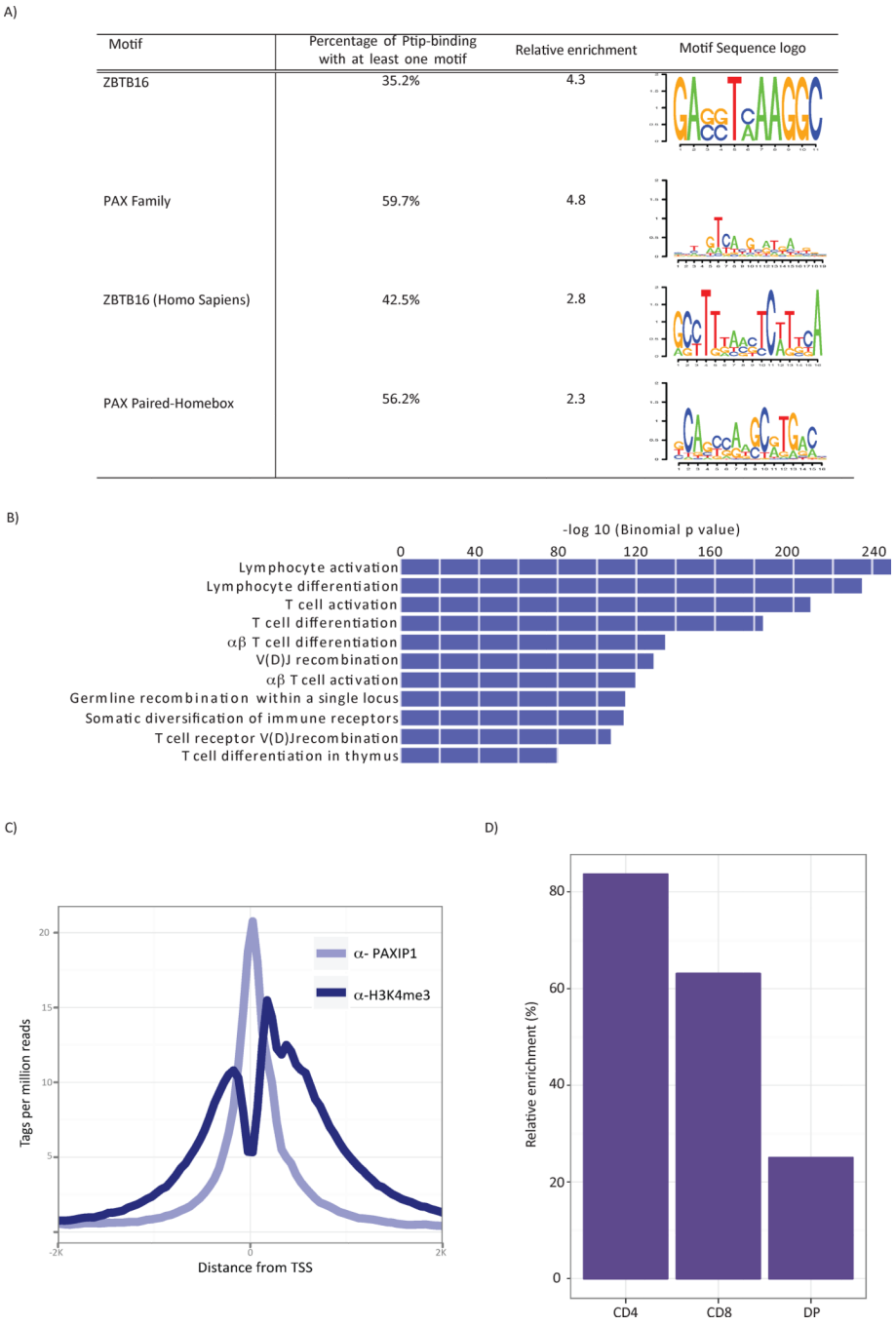
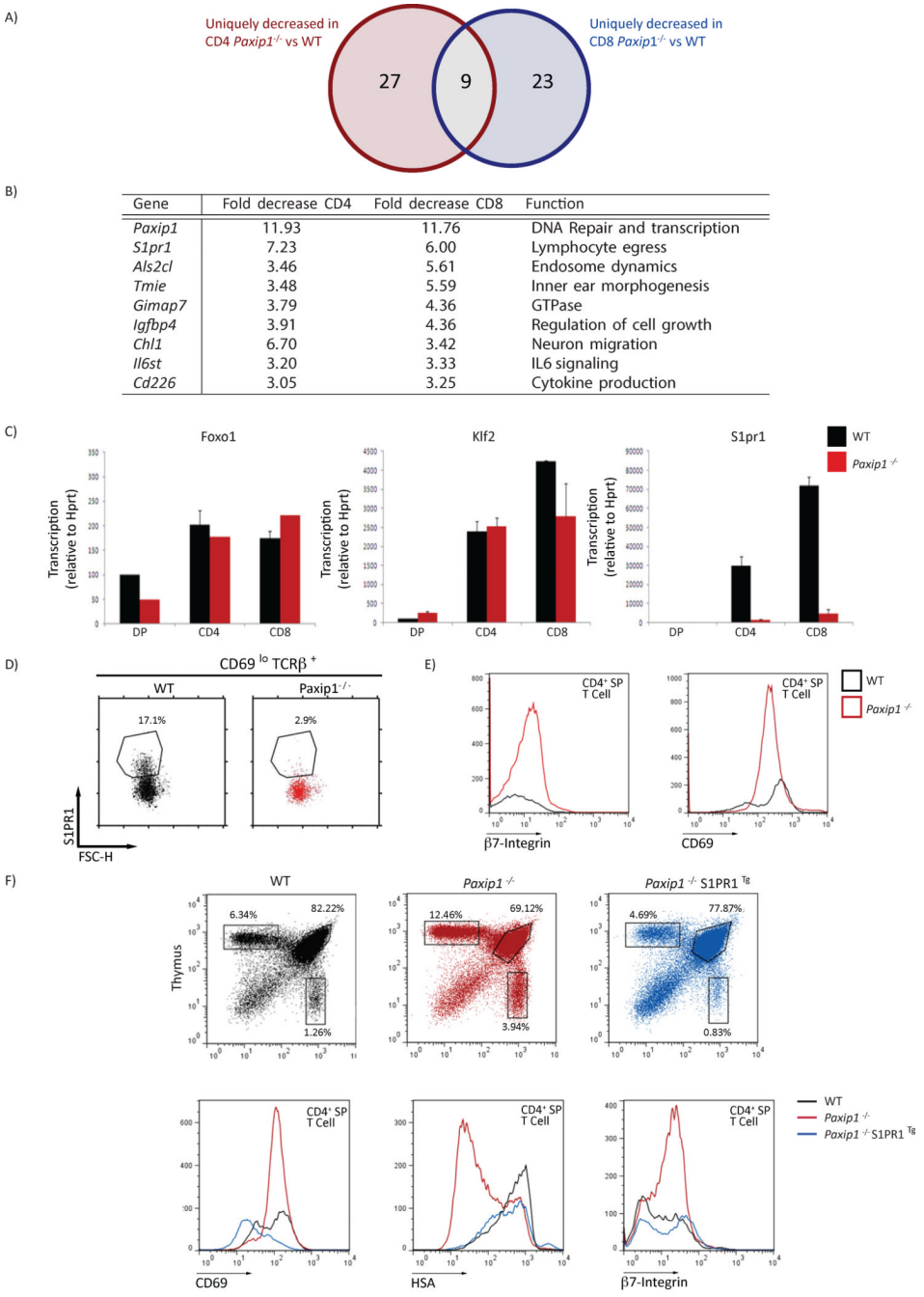



**Figure 6. ChIP-seq analysis of PAXIP1-binding sites**

(A) PAXIP1-binding sites are enriched for Pax and Zbtb16 motifs. Enriched motif models at PAXIP1-bound sites were identified and computed relative to two background models as described in the methods. The motif of PAX paired-omeodomains was computed based on *de novo* motif analysis using ChIP-seq data. Motif models for PAX 2,4,6,8 and ZBTB16 were obtained from TRANSFAC (Matys et al., 2003)and PAZAR (Portales-Casamar et al., 2007), respectively. All motif models have log likelihood ratio greater than 3 relative to the background models. (B) PAXIP1-binding sites are strongly associated with processes linked to T cell development. Top, terms in biological process Gene Ontology (GO) are enriched for T cell differentiation and activation. GREAT (McLean et al., 2010) was used for analysis

of functional gene annotations. Only terms with greater than 2 fold binomial enrichment are shown. (C) PAXIP1 is enriched at sites occupied by H3K4me3. The normalized tags per million reads of PAXIP1 and H3K4me3 in a 50bp window across a 4 kb region centered on the transcription starting sites (TSS) of NCBI Build 37 RefSeq genes is plotted. (D) Correlation between alteration in H3K4me3 and direct PAXIP1-binding. For each lineage, the percent increase in PAXIP1 -occupancy of promoters with more than 2-fold differential H3K4me3 (TSS +/− 5kb) was computed relative to a control set of the same size consisting of the next most relevant PAXIP1 -dependent promoters with the highest increase in H3K4me3 not exceeding 2-fold.




| Gene | Fold decrease CD4 | Fold decrease CD8 | Function |
|---|---|---|---|
| *Paxip1* | 11.93 | 11.76 | DNA Repair and transcription |
| *S1pr1* | 7.23 | 6.00 | Lymphocyte egress |
| *Als2cl* | 3.46 | 5.61 | Endosome dynamics |
| *Tmie* | 3.48 | 5.59 | Inner ear morphogenesis |
| *Gimap7* | 3.79 | 4.36 | GTPase |
| *Igfbp4* | 3.91 | 4.36 | Regulation of cell growth |
| *Chl1* | 6.70 | 3.42 | Neuron migration |
| *Il6st* | 3.20 | 3.33 | IL6 signaling |
| *Cd226* | 3.05 | 3.25 | Cytokine production |

**Figure 7. PAXIP1 is required for *S1pr1* expression**

(A, B) A very few genes exhibit H3K4me3 deficit in both *Paxip1*$^{-/-}$ CD4$^{+}$ SP and *Paxip1*$^{-/-}$ CD8$^{+}$ SP T cell lineages. The Venn diagram shows the overlap of genes with more than 3 times reduction in H3K4me3 TPM in both *Paxip1*$^{-/-}$ CD4$^{+}$ and CD8$^{+}$ SP T cells. The numbers represent the total number of unique gene symbol in each area. (B) List of genes with more than 3 times reduction in H3K4me3 TPM in both *Paxip1*$^{-/-}$ CD4$^{+}$ and CD8$^{+}$ SP T cells. The second and third columns show H3K4me3 reduction in *Paxip1*$^{-/-}$ CD4$^{+}$ and CD8$^{+}$ SP T cells with respect to WT CD4$^{+}$ SP and CD8$^{+}$ SP T cells, respectively. (C) Expression of *Foxo1*, *Klf2* and *S1pr1* in sorted DP, CD4$^{+}$ and CD8$^{+}$ WT (black) and *Paxip1*$^{-/-}$ (red) thymoctyes measured by qRT-PCR. (D) Surface S1PR1 expression on

TCRβ$^{hi}$CD69$^{lo}$ mature thymocytes from WT (black) and *Paxip1*$^{-/-}$ (red) mice. (E) β7-Integrin and CD69 expression in CD4$^{+}$ SP T cells was measured by flow cytometry using the same number of total WT (black line) and *Paxip1*$^{-/-}$ thymocytes (red line). (F) S1PR1 overexpression decreases accumulation of mature CD4$^{+}$ and CD8$^{+}$ SP T cells in *Paxip1* deficient thymocytes (top) and normalizes expression of maturation markers CD69, HSA and β7-integrin in SP T cells.