



OPEN ACCESS

EDITED BY

Avinash Bhandoola,
National Institutes of Health (NIH),
United States

REVIEWED BY

Ferenc Livak,
National Institutes of Health (NIH),
United States

Ravi Krishnakant Patel,
University of California, San Francisco,
United States

*CORRESPONDENCE

Michael S. Krangel

✉ krang001@mc.duke.edu

RECEIVED 17 March 2023

ACCEPTED 21 June 2023

PUBLISHED 06 July 2023

CITATION

Mihai A, Roy S, Krangel MS
and Zhuang Y (2023) E protein binding
at the *Tcra* enhancer promotes *Tcra*
repertoire diversity.

Front. Immunol. 14:1188738.

doi: 10.3389/fimmu.2023.1188738

COPYRIGHT

© 2023 Mihai, Roy, Krangel and Zhuang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

E protein binding at the *Tcra* enhancer promotes *Tcra* repertoire diversity

Ariana Mihai, Sumedha Roy, Michael S. Krangel*
and Yuan Zhuang

Department of Immunology, Duke University School of Medicine, Durham, NC, United States

V(D)J recombination of antigen receptor loci is a highly developmentally regulated process. During T lymphocyte development, recombination of the *Tcra* gene occurs in CD4⁺CD8⁺ double positive (DP) thymocytes and requires the *Tcra* enhancer (E α). E proteins are known regulators of DP thymocyte development and have three identified binding sites in E α . To understand the contribution of E proteins to E α function, mutants lacking one or two of the respective binding sites were generated. The double-binding site mutant displayed a partial block at the positive selection stage of $\alpha\beta$ T cell development. Further investigation revealed loss of germline transcription within the *Tcra* locus at the J α array, along with dysregulated primary and impaired secondary V α -J α rearrangement. E α E protein binding increases *Tcra* locus accessibility and regulates TCR α recombination, thus directly promoting *Tcra* repertoire diversity.

KEYWORDS

E proteins, *Tcra* enhancer, V(D)J recombination, T cells, Thymus

1 Introduction

The adaptive immune system recognizes a wide variety of antigens by way of B cell receptors and T cell receptors (TCRs) on the surfaces of B and T cells, respectively. These highly diverse antigen receptors (AgRs) are generated by the process of V(D)J recombination during defined stages of B and T lymphocyte development (1). V(D)J recombination is catalyzed by the lymphoid-specific recombination activating gene (RAG) proteins, which act on AgR loci to create double-strand breaks between variable (V), diversity (D), and joining (J) gene segments and their respective recombination signal sequences. AgR diversity depends on combinatorial usage of V, D, and J segments, together with junctional heterogeneity introduced by non-homologous end joining DNA repair mechanisms (1, 2).

V(D)J recombination occurs in an ordered fashion during B and T cell development due to the activity of developmentally regulated enhancer and promoter elements (1). These elements drive germline transcription and chromatin accessibility at defined sites in

AgR loci, which are then permissive for RAG binding. RAG proteins generally assemble on highly accessible D and J gene segment recombination signal sequences to form a chromatin structure referred to as the recombination center (RC). RC-bound RAG is then able to capture V gene segments to complete the recombination process (1, 3).

Fetal liver or bone marrow origin precursors that migrate to the thymus can commit to the T cell fate and further differentiate into one of several mature T cell lineages. CD4⁺CD8⁻ double negative (DN) thymocytes rearrange the *Tcrγ*, *Tcrδ*, and *Tcrβ* genes, leading to TCRγ/TCRδ or TCRβ/pre-Tα (pre-TCR) pairings. This allows adoption of the γδ or αβ T cell lineage fates, respectively (4). Having passed the β-selection checkpoint, developing αβ T cells proceed to the CD4⁺CD8⁺ double positive (DP) stage and *Tcra* rearrangement. Uniquely, *Tcrδ* and *Tcra* gene segments are arranged in a single genetic locus. In the mouse, the *Tcra-Tcrδ* locus is organized as an array of about 100 V gene segments, followed by Dδ, Jδ, and Cδ gene segments, 60 Jα gene segments, and lastly Cα (5). This configuration leads to deletion of *Tcrδ* upon *Tcra* rearrangement. *Tcra* rearrangement requires the *Tcra* gene enhancer (Eα), located directly downstream of Cα (6). Lymphoid specific transcription factors (TFs) bind adjacent to or within four defined protein-binding elements (Tα1-Tα4) that make up Eα (7–9). These TFs include c-MYB, RUNX1, RUNX3, GATA3, ETS1/FLI1, SP1, CREB, LEF1/TCF1, CTCF, E2A/HEB, NFAT, AP1, and EGR1, with some having more than one identified binding site with the 275-bp element (10). TF occupancy and histone modification is observed as early as the DN stage, though Eα activation only occurs after pre-TCR signaling (11). Eα acts in *cis* to activate locus germline transcription, assemble an RC, and initiate Vα-Jα rearrangement. Initially, Eα activates the T early alpha (TEA) and *Traj49* promoters associated with the most Vα-proximal Jα gene segments (12–14). The assembled RC then directs an initial round of Vα-Jα rearrangement, referred to as a primary rearrangement, to nearby Jα segments. However, owing to the lack of D segments, *Tcra* is capable of undergoing multiple rounds of Vα-Jα rearrangement, and these secondary rearrangements will continue until a TCR is created that can mediate positive selection. Secondary rearrangements are thought to depend on RCs created by Eα and the promoter of the rearranged V gene segment (15). Following positive selection, Eα activity is downregulated in single-positive thymocytes and mature αβ T lymphocytes (16).

E proteins are a class of basic helix-loop-helix (bHLH) proteins that recognize a canonical CANNTG DNA sequence, also referred to as an E-box (17). In developing T cells, the E2A (*Tcf3*) and HEB (*Tcf12*) E proteins have wide-ranging targets including genes regulating cell survival and cell-cycle progression, control of developmental checkpoints, and stage-specific chemokine expression (17). E proteins are important regulators of V(D)J recombination. In both developing B and T cells, E proteins regulate the stage-specific expression of the RAG genes (18, 19). E proteins also regulate germline transcription at the *Igh*, *Igk*, *Igl*, *Tcrγ*, *Tcrδ*, and *Tcrβ* genes leading to recombination permissive chromatin environments (17, 20–22). Furthermore, E proteins are essential regulators of T cell development checkpoints, and have been previously shown to control DP cell development and

transition to the SP stage (23–26). Of the multiple TFs binding at Eα, E proteins are notable in that they occupy three identified binding sites during thymocyte development and show diminished binding in mature αβ T lymphocytes (11, 16, 27, 28). However, it is not known whether E proteins have direct effects on Eα activity and *Tcra* recombination, in part due to the broad impacts, noted above, on the DP thymocyte population in mouse E protein knockout models.

To assess the role of E proteins in *Tcra* rearrangement, E-boxes at Eα were deleted. Loss of the 5' E-box (E1) had no discernable effect on thymocyte development. However, the additional deletion of the 3' E-box (E3) impaired positive selection of DP thymocytes, with a concordant loss of CD4⁺ and CD8⁺ single positive (SP) thymocytes. The double E-box deletion reduced germline transcription across Jα segments, which is expected to result in reduced accessibility for Vα-Jα recombination. Consistent with this, alterations in Jα segment usage and invariant natural killer T (iNKT) cell development were detected, suggestive of dysregulated primary and impaired secondary Vα-Jα rearrangement. Therefore, E protein binding to Eα increases *Tcra* locus germline transcription, regulates Jα segment recombination, and promotes *Tcra* repertoire diversity.

2 Materials and methods

2.1 Mice

The ΔE1 mutation was generated in B6SJL/J (RRID : IMSR_JAX:100012). E1 was targeted with a two-guide and donor (Supplementary Table 1) CRISPR/Cas9 approach by pro-nuclear injection. Founders were screened by Sanger sequencing; identified mutants lacked the donor sequence. Founders were then crossed to lab-maintained C57BL/6. Eα^{ΔE1/ΔE1} pups from Eα^{ΔE1/+} by Eα^{ΔE1/+} breeding were screened by Sanger sequencing for C57BL/6 and SJL/J polymorphisms within *Tcra-Tcrδ*, all matching to the C57BL/6 reference sequence (Supplementary Table 1).

Single-guide (Supplementary Table 1) CRISPR/Cas9 electroporation of Eα^{ΔE1/ΔE1} was used to generate alleles with two mutated Eα E-boxes. Founders were screened by Sanger sequencing. The ΔE1ΔE3 and ΔE1ΔE3(1) alleles were detected in different founders and maintained separately by crossing to lab-maintained C57BL/6. The mutation was introduced onto a RAG-deficient background by crossing ΔE1ΔE3 to *Rag1*^{tm1Mom}/J (RRID : IMSR_JAX:002216).

Except where noted, all analyses were performed on mice 3–4 weeks of age obtained from heterozygous by heterozygous Eα allele breeding. All mice were of mixed C57BL/6 and SJL/J strain background.

The Duke University Cancer Institute Transgenic and Knockout Mouse Shared Resource carried out the above CRISPR/Cas9-mediated mutagenesis. All mice were bred in a Duke University Division of Laboratory Animal Resources specific pathogen-free facility and handled in accordance with protocols approved by the Duke University Institutional Animal Care and Use Committee.

2.2 Antibodies

Fluorescently conjugated antibodies used in flow cytometry and cell sorting are commercially available and have been previously validated (Supplementary Table 2). CD1d tetramer was obtained from the National Institutes of Health Tetramer Core Facility.

2.3 Flow cytometry and cell sorting

Thymus was harvested and dissociated to single-cell suspension in FACS buffer (2.5% FBS and 2 mM EDTA supplemented PBS) and filtered using 70 nm nylon mesh. For analysis by flow cytometry, 3×10^6 cells were stained with fluorescently labeled antibodies and loaded CD1d-tetramer for 30 minutes at 4°C and then washed with excess FACS buffer. Samples were re-suspended in FACS buffer containing 7-Aminoactinomycin D (7-AAD) (ThermoFisher Scientific, Cat. A1310) or DAPI (Sigma-Aldrich, Cat. D9542) and analyzed on FACSCantoII or Fortessa X20 (BD Biosciences) cytometers available via the Duke University Cancer Institute Flow Cytometry Shared Resource. Analysis was performed using FlowJo (version 10.8.1) software. Gating scheme is shown in Supplementary Figure 1.

For isolation of preselection thymocytes, sorting was performed using an Astrios (Beckman-Coulter) cell sorter by the Duke University Flow Cytometry Shared Resource. Sample staining included a lineage dump (B220, CD11b, CD11c, CD19, GR-1, TER119, F4/80, and TCR δ antibodies). Gating scheme is shown in Supplementary Figure 2.

2.4 *Tcra* repertoire library preparation

Three to five million pre-selection DP thymocytes (CD4⁺CD8⁺CD3^{lo}) were sorted from E $\alpha^{+/+}$ and E $\alpha^{\Delta E1\Delta E3/\Delta E1\Delta E3}$ thymuses. Sorted cells were re-suspended in TRIzol and stored at -80°C for later RNA extraction. RNA was purified using Direct-Zol RNA Microprep (Zymo Research) kit with on-column DNase digestion. Sequencing libraries were prepared with modification of previously published methods (29, 30). Briefly, 5 ng of total RNA was reverse transcribed using SmartScribe Reverse Transcriptase (Takara Bio, Cat. 639538) with Trac-RT and SMARTnnna template switch oligo (Supplementary Table 1). Samples were then treated for 40 min at 37°C with 5 units uracil DNA glycosylase (NEB). cDNA purification was carried out using Ser-Mag Carboxylate-Modified Magnetic SpeedBeads (GE Healthcare Life Sciences). Q5 high-fidelity polymerase (NEB) was used to amplify cDNA (18 cycles), which was then purified using Ser-Mag Carboxylate-Modified Magnetic SpeedBeads. Q5 high-fidelity polymerase was used to perform dual-indexing of amplified cDNA, followed again by Ser-Mag Carboxylate-Modified Magnetic SpeedBeads purification. Sample quality control was performed by agarose gel electrophoresis. Replicate E $\alpha^{+/+}$ and E $\alpha^{\Delta E1\Delta E3/\Delta E1\Delta E3}$ indexed samples were pooled at equal concentration and final library preparation carried out using NEBNext Ultra II DNA

library preparation kit (NEB). Library was gel extracted using Zymoclean Gel DNA recovery kit (Zymo Research). Sequencing was performed on a MiSeq sequencer (Illumina) (300 x 300 bp) by the Duke University Cancer Institute Sequencing and Genomic Technologies Shared Resource.

2.5 *Tcra* repertoire analysis

Tcra repertoire analysis followed previously published protocols (29, 30). The following were performed using Migece v1.2.9 (31) (RRID : SCR_016337) (1): demultiplexing of raw fastq files (function: Checkout; parameters: -cute); (2) assessment of molecular identifier group (MIG) size distribution (function: Histogram; parameters: default); collapse of UMIs and filtering (function: AssembleBatch; parameters: -force-collision-filter -force-overseq 5). Collapsed reads were merged using MiTools v1.5 (<https://github.com/milaboratory/mitools>) (function: merge; parameters: -I -s 0.7). MiXCR v3.0.13 (32) (RRID : SCR_018725) was used to perform: read alignment (function: align; parameters: -s mmu -OvParameters.geneFeatureToAlign=VTranscript); assemble clonotypes (function: assemble; parameters: -OassemblingFeatures=[CDR2+FR3+CDR3] -OclusteringFilter.specificMutationProbability=1E-4); clones export (function: exportClones; parameters: default). VDJTools v1.2.1 (33) was used to convert files appropriately (function: Convert; parameters: default) and determine segment usage (function: CalcSegmentUsage; parameters: -u).

2.6 Generation of RAG-deficient DP thymocytes

Experimental mice were injected intraperitoneally with 150 μ g anti-CD3e (BioLegend, Cat. 145-2C11) at 3 weeks of age. At 10 days post-injection, thymus was harvested and dissociated to single-cell suspension in FACS buffer. Whole thymocytes were re-suspended in TRIzol and stored at -80°C for later RNA extraction.

2.7 Reverse transcription

RNA was purified using Direct-Zol RNA Microprep (Zymo Research) kit with on-column DNase digestion. cDNA synthesis was performed using SuperScript III Reverse Transcriptase (Invitrogen, Cat. 108080093) with random hexamers as per manufacturer instructions.

2.8 RT-qPCR

Reverse transcription quantitative polymerase chain reaction (RT-qPCR) was performed with PowerTrack SYBR Green Master Mix (ThermoFisher Scientific, Cat. A46109) as per manufacturer instructions, using 5 ng of cDNA per 20 μ L reaction. Reactions were

carried out in Axygen PCR microplate (PCR-96-LP-FLT-C) with MicroAmp Optical Adhesive Film (ThermoFisher Scientific, Cat. 4311971) using an Eppendorf MasterCycler qPCR machine. Values shown were normalized to those for *Actb*. Primers are shown in [Supplementary Table 1](#).

2.9 ChIP

Chromatin immunoprecipitation (ChIP) was adapted from a previously published protocol (34). Dynabeads M-280 Sheep anti-Mouse IgG (Invitrogen, Cat. 11202D) (100 μ L per sample) were washed four times using PBS-BSA (1x PBS, 5 mg/mL BSA) (ThermoFisher Scientific, Cat. 15260037). Dynabeads were resuspended in PBS-BSA (250 μ L per sample) with 8 μ g (per sample) anti-E2A (Santa Cruz Biotechnology, Cat. sc-416X; Yae), anti-HEB (Santa Cruz Biotechnology, Cat. sc-28364X; D-3), anti-GATA3 (Santa Cruz Biotechnology, Cat. sc-268X; HG3-31), or anti-RUNX1 (Santa Cruz Biotechnology, Cat. sc-365644X; A-2), and incubated at 4°C for 24 hours with rotation. Dynabeads were washed three times with PBS-BSA, and aliquoted to sample tubes for final wash. Whole thymocytes were resuspended at 10×10^6 cells/mL in 1% paraformaldehyde (ThermoFisher Scientific, Cat. 28906) in RPMI 1640 (ThermoFisher Scientific, Cat. 11875093) (supplemented with 10% FBS) for cross-linking. After 10 min at room temperature, cross-linking reaction was stopped by adding glycine to a final concentration of 0.125 M. Cross-linked cells were pelleted (at 4°C) and washed with 1x PBS (4°C). Cross-linked cells were resuspended at 20×10^6 cells/mL in ChIP Lysis Buffer (Santa Cruz Biotechnology, Cat. sc-45000) and incubated 5 min on ice. Cross-linked cells were pelleted and resuspended at 10×10^6 cells/mL in ChIP Lysis Buffer High Salt (Santa Cruz Biotechnology, Cat. sc-45001). Samples were sonicated for four cycles (30 s on, 30 s off) using a Bioruptor Pico (Diagenode). Samples were pelleted 15 min at 13200 rpm (4°C). From the supernatant, 1% was removed and stored at -20°C for use as input. Antibody-bound Dynabeads were resuspended in 500 μ L of sample supernatant (5×10^6 cells starting material) and incubated at 4°C for 16 hours with rotation. Beads were washed five times with ChIP Wash Buffer (Santa Cruz Biotechnology, Cat. sc-45002), with 3 min incubation at 4°C with rotation between washes. Beads were resuspended in 1 mL TE and rotated 1 min at room temperature. Supernatant was discarded and beads were resuspended in 200 μ L ChIP Elution Buffer (Santa Cruz Biotechnology, Cat. sc-45003). Samples were incubated 1 hour at 65°C at 1500 rpm. Supernatant was transferred to new tube and incubated 16 hours at 65°C at 1500 rpm. 150 μ L ChIP Elution Buffer was added to 50 μ L input sample and incubated 16 hours at 65°C at 1500 rpm. To each sample 200 μ L TE and 80 μ g RNase A (ThermoFisher Scientific, Cat. EN0531) were added, and then incubated 1 hour at 37°C at 1000 rpm. To each sample 80 μ g proteinase K (ThermoFisher Scientific, Cat. AM2546) was added, and then incubated 30 min at 56°C at 1000 rpm. DNA was purified by phenol/chloroform/ethanol extraction, and resuspended in 50 μ L TE. Samples were assayed by qPCR with PowerTrack SYBR Green Master Mix as per manufacturer instructions, using 2 μ L DNA per 20 μ L reaction. Reactions were

carried out in Axygen PCR microplate with MicroAmp Optical Adhesive Film using an Eppendorf MasterCycler qPCR machine. Values shown are relative to input. Primers are shown in [Supplementary Table 1](#). PBS-BSA, ChIP Lysis Buffer, ChIP Lysis Buffer High Salt, and ChIP Wash Buffer contained protease inhibitors (Roche, Cat. 11697498001).

2.10 Statistical analysis

All reported data are from individual mice, with no repeated measurements from the same sample. Sample size was not predetermined by statistical methods. GraphPad Prism (version 9.5.0) software was used for all statistical analyses and generation of graphs.

3 Results

3.1 Normal $\alpha\beta$ T cell development upon deletion of a single E α E-box

To assess the impact of E protein binding at E α , the E-box upstream of T α 1 (E1) was targeted by CRISPR/Cas9. The resulting 13 base pair (bp) deletion (Δ E1) eliminates the 6 bp E-box; although the upstream 7 bp were also removed and base pair changes occur at the gRNA recognition sequence ([Figure 1A](#), [Supplementary Figure 3A](#)), no transcription factor binding was detected at these sites in previous dimethylsulfate genomic footprinting experiments (27). The generated allele is hereafter denoted as E α^{Δ E1.

E α deletion blocks thymocyte development at the DP stage, albeit with normal thymus cellularity (6). Using flow cytometry, no significant difference was detected in the numbers and proportions of total, DP, and SP thymocytes in the E α^{Δ E1/ Δ E1 mutants when compared to E $\alpha^{+/+}$ littermates ([Figures 1B–E](#), [Supplementary Figure 3B](#)). Further analysis of CD3 and CD69 expression did not show a significant difference in DP cells that are positively selected in E α^{Δ E1/ Δ E1 mice ([Figures 1F, G](#)).

3.2 Deletion of two E α E-boxes impairs positive selection

To further impact E protein binding at E α , the E3 E-box (located in T α 4) was targeted for deletion in the E α^{Δ E1 allele. The generated E α^{Δ E1 Δ E3} allele precisely eliminates the 6 bp E3 E-box ([Figure 2A](#)). A second allele, E α^{Δ E1 Δ E3(1)}, disrupts E3 by a 1 bp deletion ([Supplementary Figure 4A](#)). As in E α^{Δ E1/ Δ E1, thymic cellularity in E α^{Δ E1 Δ E3/ Δ E1 Δ E3} double mutants did not differ from that of wild-type littermates ([Supplementary Figure 4B](#)). However, analysis by flow cytometry revealed clear reductions of CD4⁺ and CD8⁺ thymocytes in the double mutants ([Figures 2B–E](#)). This suggests reduced positive selection, a supposition confirmed by the substantial reductions in both the frequencies and numbers of CD3⁺CD69⁺ DP thymocytes ([Figures 2F, G](#)). The developmental

block was similarly observed in $E\alpha^{\Delta E1\Delta E3(1)/\Delta E1\Delta E3(1)}$ double mutants (Supplementary Figures 4C–I).

To understand the molecular basis for this block, TF binding at $E\alpha$ was assessed by ChIP-qPCR. As per expectations, E2A and HEB binding was significantly reduced (Figures 3A–C). That reductions were only partial and were detected at all three E protein binding sites likely reflects residual E protein binding at the intact E2 site (in $T\alpha 3$), coupled with the inability of ChIP to resolve binding signals at E protein binding sites separate by no more than 150 bp. Interestingly, ablation of E-boxes in $T\alpha 1$ and $T\alpha 4$ provoked mildly reduced binding of GATA3 (whose defined binding site is in $T\alpha 3$), but not RUNX1 (whose defined binding sites are in $T\alpha 2$) (Figures 3A, D, E) (27). In contrast, no changes in transcription factor binding were detected at the RAG anti-silencer element (ASE), which contains binding sites for all of these factors. These results highlight that loss of E protein binding may mediate effects in part by destabilizing the binding of other components of the $E\alpha$ enhanceosome, with some transcription factors having greater dependence on E-protein binding than others.

Although impaired, $E\alpha$ likely retains considerable activity in double mutant mice given the continued progression of $\alpha\beta$ T cell development and generation of SP thymocytes. No impact was observed on the development of $\gamma\delta$ T cells (Supplementary Figures 5A–C).

3.3 $E\alpha$ E-box deletion reduces $Tcra$ - $Tcrd$ germline transcription

$E\alpha$ activity was evaluated by using RT-qPCR to assess transcription of known $E\alpha$ targets in DP thymocytes from $Rag1$ -deficient $E\alpha^{\Delta E1\Delta E3/\Delta E1\Delta E3}$ and $E\alpha^{+/+}$ littermates (Figure 4A). $E\alpha$ acts in *cis* to activate the T early- α (TEA) promoter immediately 5' of the $J\alpha$ array, and the $J\alpha 49$ promoter 15 kb downstream of TEA. These promoters target primary $V\alpha$ - $J\alpha$ rearrangements to the most proximal $J\alpha$ segments, after which they are excised (12, 13, 35, 36). $J\alpha$ accessibility in secondary rearrangements is driven by promoters of rearranged $V\alpha$ segments (15). The double E-box mutant showed

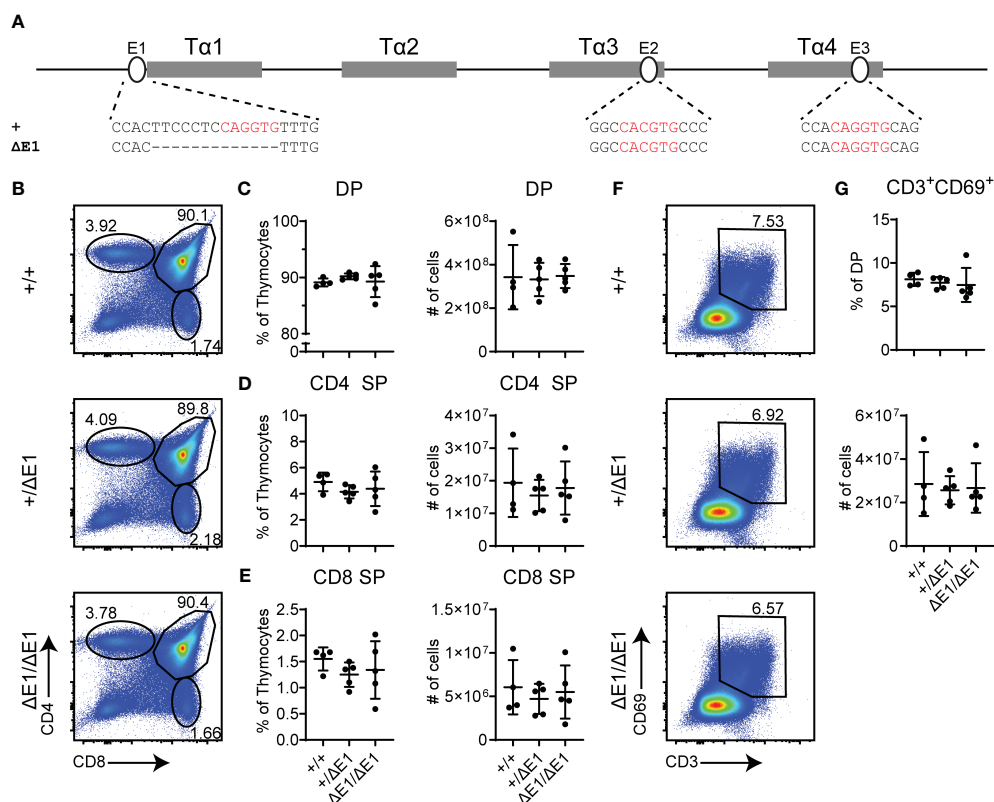


FIGURE 1

Loss of E1 does not impair T cell development. (A) Diagram of relative positions of $T\alpha 1$ - $T\alpha 4$ protein binding regions and E1-E3 E protein binding sites within $E\alpha$, with sequences of wild-type (+) and $\Delta E1$ mutation indicated below. E protein binding motifs are highlighted in red. (B) Representative flow cytometry plots, displayed as CD4 versus CD8, of live thymocytes from $E\alpha^{+/+}$, $E\alpha^{+/\Delta E1}$, and $E\alpha^{\Delta E1/\Delta E1}$ mice. Frequencies of gated populations are shown. (C–E) Frequencies and numbers of (C) DP thymocytes, (D) $CD4^+$ SP thymocytes, and (E) $CD8^+$ SP thymocytes, in $E\alpha^{+/+}$, $E\alpha^{+/\Delta E1}$, and $E\alpha^{\Delta E1/\Delta E1}$ mice, with gating as shown in Supplementary Figure 1. Note that $CD8^+$ SP are gated as $CD8^+TCR\beta^+CD24^-$ to exclude $CD8$ immature single positives. (F) Representative flow cytometry plots, displayed as $CD69$ versus $CD3$, of DP thymocytes (gated as shown in (B)) from $E\alpha^{+/+}$, $E\alpha^{+/\Delta E1}$, and $E\alpha^{\Delta E1/\Delta E1}$ mice. Frequencies of gated populations are shown. (G) Frequency and number of $CD3^+CD69^+$ DP cells in $E\alpha^{+/+}$, $E\alpha^{+/\Delta E1}$, and $E\alpha^{\Delta E1/\Delta E1}$, following gating as shown in (F). Data are pooled from 3 independent experiments and are plotted as mean \pm SD. $E\alpha^{+/+}$ (n = 4), $E\alpha^{+/\Delta E1}$ (n = 5), and $E\alpha^{\Delta E1/\Delta E1}$ (n = 5). Statistical analysis: one-way ANOVA with correction for multiple comparison using Tukey's *post hoc* testing. Significant differences were not detected.

a significant loss of germline transcription for the region spanning TEA to $C\alpha$ (*Trac*) (Figure 4B). Upstream of TEA, the $J\delta$ segments showed loss of transcription while proximal $V\alpha$ segments had no significant difference in expression (Figures 4C, D). Expression of genes downstream of *Tcra-Tcrd* likewise remained largely unperturbed (Figure 4E). Thus, in agreement with the partial development block, $\Delta E1\Delta E3$ causes reduced transcription of a subset of $E\alpha$ targets.

3.4 $E\alpha$ E-box deletion impairs TCR expression in thymic DP cells

Flow cytometry was also used to analyze surface TCR expression in recombinase-sufficient thymocytes. Notably, DP thymocytes of $\Delta E1\Delta E3$ mice displayed a substantial reduction in TCR β surface expression (Figure 5A), with a much smaller

reduction apparent on CD3⁺CD69⁺ DP thymocytes (Figure 5B). There was no difference in TCR β surface expression on CD4⁺ and CD8⁺ SP thymocytes (Figures 5C, D). Because TCR expression is upregulated during positive selection, the reduced frequency of CD3⁺CD69⁺ DP thymocytes (Figures 2F, G) could account in part for the overall reduction in TCR β expression in DP thymocytes. Alternatively, or in addition, the reduction in TCR β surface expression could reflect an increase of pre-TCR expression in $\Delta E1\Delta E3$ DP thymocytes due to impaired *Tcra* gene rearrangement and TCR $\alpha\beta$ surface assembly, or diminished TCR $\alpha\beta$ expression due to reduced transcription of rearranged *Tcra* genes. That there is no change in TCR β surface expression in more mature thymocyte populations may reflect selection for higher TCR expression during positive selection, or diminished effects of E proteins on $E\alpha$ in mature cells (16, 37).

$E\alpha$ activity is not restricted to $\alpha\beta$ T lymphocytes or their development, having been shown to contribute to normal

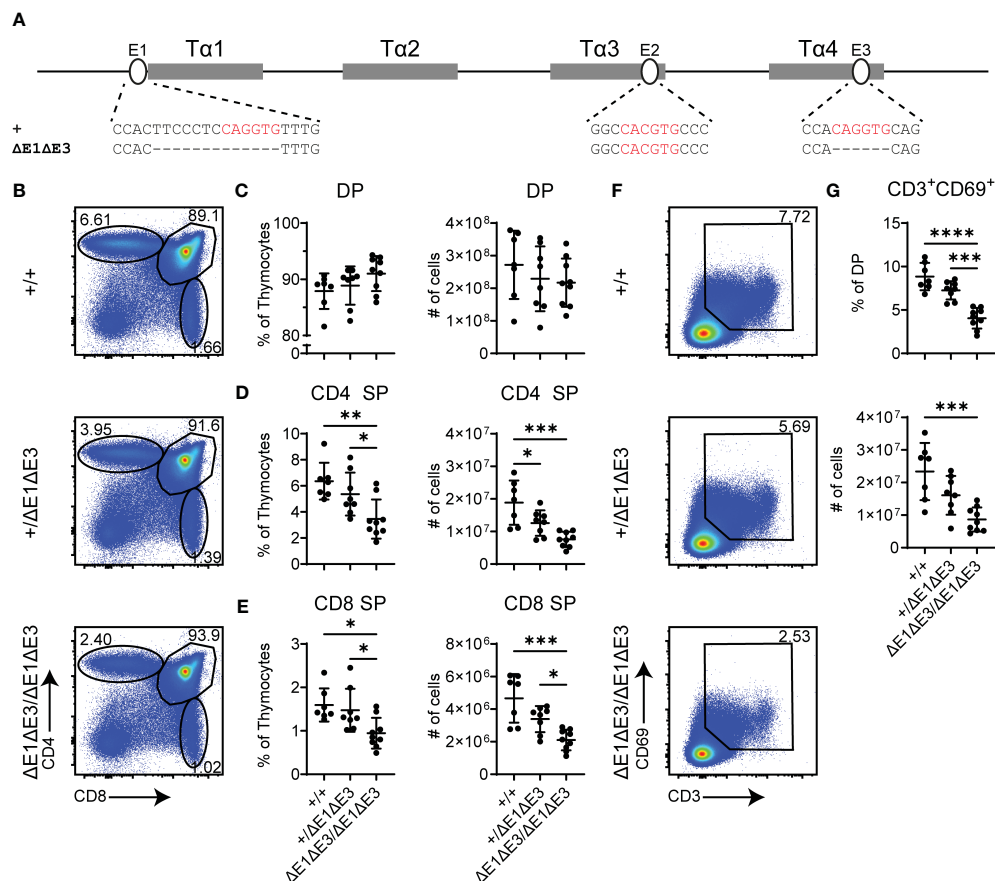


FIGURE 2

Loss of E1 and E3 impairs T cell development. (A) Diagram of relative positions of Tα1–Tα4 protein binding regions and E1–E3 E protein binding sites within $E\alpha$, with sequences of wild-type (+) and $\Delta E1\Delta E3$ mutation indicated below. E protein binding motifs are highlighted in red. (B) Representative flow cytometry plots, displayed as CD4 versus CD8, of live thymocytes from $E\alpha^{+/+}$, $E\alpha^{+/\Delta E1\Delta E3}$, and $E\alpha^{\Delta E1\Delta E3/\Delta E1\Delta E3}$ mice. Frequencies of gated populations are shown. (C–E) Frequencies and numbers of (C) DP thymocytes, (D) CD4⁺ SP thymocytes, and (E) CD8⁺ SP thymocytes, in $E\alpha^{+/+}$, $E\alpha^{+/\Delta E1\Delta E3}$, and $E\alpha^{\Delta E1\Delta E3/\Delta E1\Delta E3}$ mice. (F) Representative flow cytometry plots, displayed as CD69 versus CD3, of DP thymocytes [gated as shown in (B)] from $E\alpha^{+/+}$, $E\alpha^{+/\Delta E1\Delta E3}$, and $E\alpha^{\Delta E1\Delta E3/\Delta E1\Delta E3}$ mice. Frequencies of gated populations are shown. (G) Frequencies and numbers of CD3⁺CD69⁺ DP thymocytes in $E\alpha^{+/+}$, $E\alpha^{+/\Delta E1\Delta E3}$, and $E\alpha^{\Delta E1\Delta E3/\Delta E1\Delta E3}$, with gating as shown in (F). Data are pooled from 4 independent experiments and are plotted as mean \pm SD. $E\alpha^{+/+}$ (n = 7), $E\alpha^{+/\Delta E1\Delta E3}$ (n = 8), and $E\alpha^{\Delta E1\Delta E3/\Delta E1\Delta E3}$ (n = 9). Statistical analysis: one-way ANOVA with correction for multiple comparison using Tukey's *post hoc* testing. * p < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001.

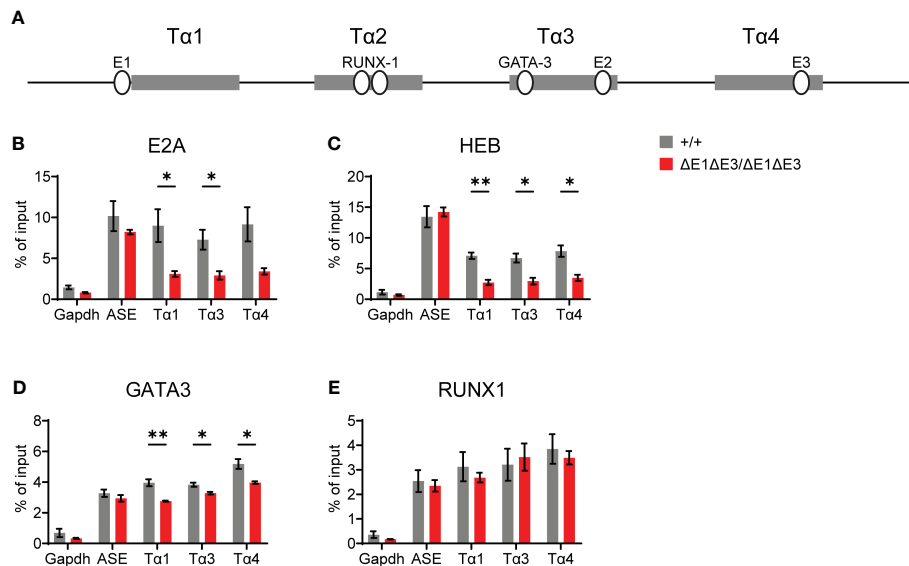


FIGURE 3

Loss of E1 and E3 reduces E protein and GATA3 binding at E α . (A) Diagram of relative positions of T α 1–T α 4 protein binding regions and GATA3, RUNX1, and E1–E3 E protein binding sites within E α . ChIP–qPCR of E2A (B), HEB (C), GATA3 (D), and RUNX1 (E) binding in E α ^{+/+} (n = 3) and E α ^{ΔE1ΔE3/ΔE1ΔE3} (n = 3) whole thymocytes. The RAG ASE and *Gapdh* served as positive and negative controls, respectively. Data were pooled from 3 independent experiments, one of which included an age-matched C57BL/6 as E α ^{+/+}. Data are plotted as mean \pm SEM. Statistical analysis: Student's t test. **p* < 0.05, ***p* < 0.01.

expression of *Tcrd* in $\gamma\delta$ T lymphocytes (6). TCR δ ⁺ thymocytes of $\Delta E1\Delta E3$ mice displayed a small but significant reduction in TCR δ surface expression (Figure 5E). This indicates that E proteins contribute in part to E α activity in $\gamma\delta$ T lymphocytes.

3.5 E α E-box deletion impairs TCR α rearrangement

As noted previously, TEA and J α 49 promoter-driven transcription normally target primary rearrangements to J α segments proximal to these promoters. Once these promoters are deleted by primary rearrangement, secondary rearrangement is thought to be directed by the promoter of the rearranged V α gene segment. However, in mice with genetic deletion of the TEA (or TEA and J α 49) promoter(s), the activation of cryptic downstream promoters causes dysregulated primary rearrangement directed more broadly across the central and distal J α segments. As such, changes in J α transcriptional activity are reflected by changes to the TCR α repertoire (12, 13, 15, 35, 36).

To assess how E protein binding at E α impacts TCR α rearrangement, J α segment usage was evaluated by performing 5' rapid amplification of cDNA ends (5'RACE) on preselection DP thymocytes (CD4⁺CD8⁺CD3^{lo}) sorted from E α ^{ΔE1ΔE3/ΔE1ΔE3} and E α ^{+/+} littermates. The double E-box mutants showed a significantly altered J α repertoire. While proximal and distal J α segments were underrepresented, most segments from J α 49 to J α 31 were significantly overrepresented (Figure 6, Supplementary Figure 6). This suggests a modest defect in primary rearrangements that preferentially affects TEA-dependent J α segments, coupled with a

substantial defect in secondary rearrangements required for usage of distal J α segments.

3.6 iNKT cells are reduced in double E α E-box mutant

DP thymocytes of *Rorc*^{-/-} mice are short-lived and the TCR α repertoire is consequently limited to primary V α -J α rearrangements (38). Invariant natural killer T cells (iNKT) are absent in *Rorc*^{-/-} mice and thus their characteristic V α 14-J α 18 TCR α chain is considered to be the product of secondary TCR α recombination (39, 40). iNKT cellularity and development was therefore assessed to provide additional evidence that E α E-box mutations impact secondary V α -J α rearrangement.

While there was no observed change in E α ^{ΔE1/ΔE1} (Supplementary Figures 7A, B), flow cytometry analysis determined iNKTs to be significantly reduced in both frequency and number in double E-box mutants (both E α ^{ΔE1ΔE3/ΔE1ΔE3} and E α ^{ΔE1ΔE3(1)/ΔE1ΔE3(1)}) (Figures 7A, B, Supplementary Figure 7C). Because all iNKT developmental stages were reduced numerically and proportionally (Figure 7C, Supplementary Figure 7D), iNKT development is obstructed prior to TCR surface expression and subsequent lineage commitment. This suggests diminished rearrangement of the V α 14-J α 18 invariant *Tcra* chain in double E-box mutants, and is in accord with reduced usage of J α 18 observed in repertoire analysis. In conjunction with the broadly reduced use of distal J α segments, this result indicates an impairment of secondary TCR α recombination upon E α E-box deletion.

Prior work demonstrated a deficiency in iNKT cells in mice lacking E protein HEB (26). Moreover, mice deficient in E protein inhibitors (Id2-Id3 double deficiency) were shown to have an increase in iNKT cells as well as elevated $V\alpha 14$ - $J\alpha 18$ rearrangement in preselection DP thymocytes (41). However, in neither case was a mechanism clearly established. The present work makes clear that one mechanism by which E proteins control iNKT cell development is by direct effects on the *Tcra* locus.

4 Discussion

TCR α rearrangement is a highly regulated and ordered process that takes place during the DP stage of thymocyte development (5). $E\alpha$ is a required *cis*-acting element which plays crucial roles in promoting locus accessibility and establishing the RC (14, 42). Among the TFs binding at $E\alpha$, E proteins have three identified binding sites and show recruitment as early as the DN stage (11, 16, 27). E2A and HEB carry out a myriad of roles crucial to lymphocyte development, including mediating germline transcription and recombination accessibility at other antigen receptor loci (17, 22, 43). The findings shown here indicate that $E\alpha$ E protein binding contributes to the regulation of TCR α recombination and repertoire diversity.

Although there was no effect of single E-box deletion ($\Delta E1$), double E-box disruption [$\Delta E1\Delta E3$ and $\Delta E1\Delta E3(1)$] produced a partial block in $\alpha\beta$ T cell development at the DP stage, and thus impaired generation of SP thymocytes. The present data cannot

distinguish whether the effects of E protein binding to E1 and E3 are distinct or redundant. Generation of $\Delta E3$ and $\Delta E2$ single mutants, as well as combinatorial E-box deletions, would be needed to more fully understand the contributions of the three E protein binding sites. We did not disrupt E2 in the present study because mutation of two E-protein binding sites proved sufficient to test the central hypothesis of the study, namely that E protein binding to $E\alpha$ is important for $E\alpha$ activity and *Tcra* recombination. The finding that $E\alpha$ GATA3 binding is at least partially dependent on E protein binding suggests that one function of E proteins is to stabilize the binding of other transcription factors on $E\alpha$, and that reduced binding of other transcription factors may contribute to impaired $E\alpha$ activity when E protein binding is prevented.

Intriguingly, germline transcription in $\Delta E1\Delta E3$ was primarily reduced across the $J\alpha$ array, even though $E\alpha$ is known to regulate chromatin structure and transcription over hundreds of kb upstream and downstream. Interaction with and accessibility at these other regions may be regulated or compensated by other $E\alpha$ binding TFs. Furthermore, while transcript abundance is reduced at both TEA and $J\alpha 49$, the overrepresentation of central $J\alpha$ segments in the TCR α repertoire indicates that recombination is being preferentially directed by the latter promoter (12). Consistent with this, $J\alpha 49$ is the most proximal $J\alpha$ segment to be present at a higher proportion in the repertoire. Collectively, these data suggest that E proteins binding at $E\alpha$ has a prominent role in regulating TEA functionality and thus the accessibility of TEA-dependent $J\alpha$ segments.

The TEA and $J\alpha 49$ promoters are crucial to locus accessibility and directing primary recombination to proximal $J\alpha$ elements. In the

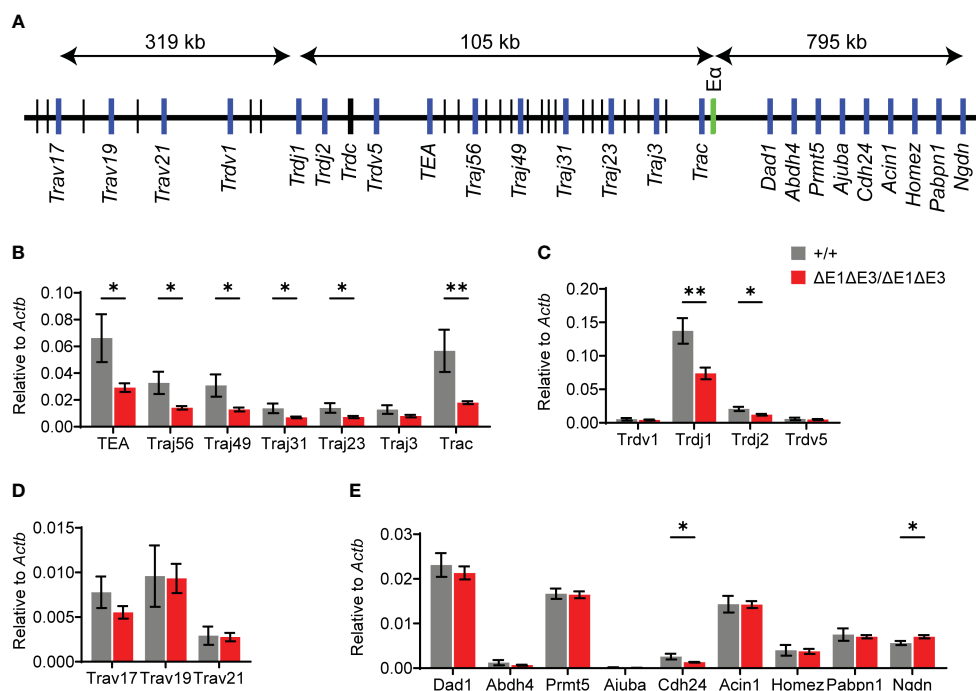


FIGURE 4

Loss of E protein binding at $E\alpha$ reduces *Tcra-Tcrd* locus transcription. (A) Diagram of murine *Tcra-Tcrd* locus region assayed by RT-qPCR. Diagram is not to scale; distances between features are indicated. Relative position of assayed genes and gene segments (blue), $E\alpha$ (green), and other gene segments (black) are shown. (B–E) RT-qPCR analysis of transcription in RAG-deficient $E\alpha^{+/+}$ (n = 5) and $E\alpha^{\Delta E1\Delta E3/\Delta E1\Delta E3}$ (n = 8) DP thymocytes. Data were pooled from 3 independent experiments. Data are plotted as mean \pm SEM. Statistical analysis: Student's t test. * $p < 0.05$, ** $p < 0.01$.

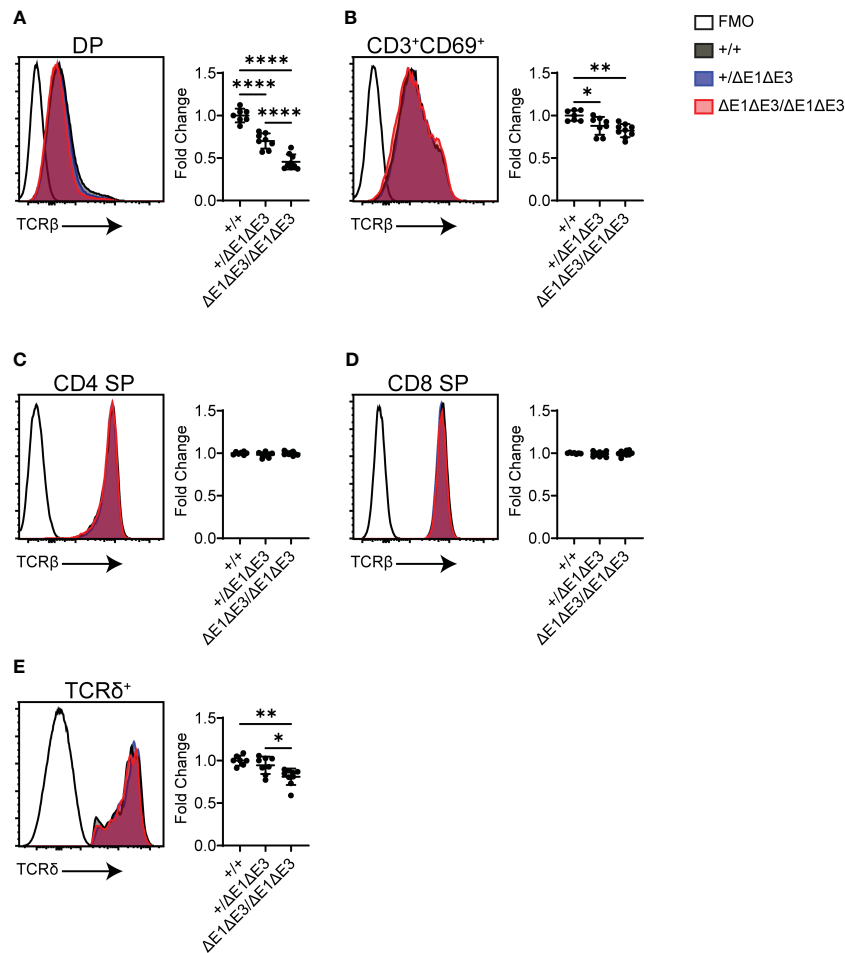


FIGURE 5

TCR surface expression is reduced in $\Delta E1\Delta E3$ DP thymocytes and thymic $\gamma\delta$ T lymphocytes. (A–D) Representative overlays of TCR β cell surface staining together with fluorescence-minus-one (FMO) controls, assayed by flow cytometry as shown in [Supplementary Figure 1](#). Graphs summarize TCR β median-MFI for $E\alpha^{+/+}$, $E\alpha^{+/\Delta E1\Delta E3}$, and $E\alpha^{\Delta E1\Delta E3/\Delta E1\Delta E3}$. (A) DP thymocytes; (B) positively selected (CD3 $^{+}$ CD69 $^{+}$) DP thymocytes; (C) CD4 $^{+}$ SP thymocytes; (D) CD8 $^{+}$ SP thymocytes. (E) Representative overlay of TCR δ surface staining together with FMO control, with graph summarizing TCR δ median-MFI of TCR δ^{+} thymocytes for $E\alpha^{+/+}$, $E\alpha^{+/\Delta E1\Delta E3}$, and $E\alpha^{\Delta E1\Delta E3/\Delta E1\Delta E3}$. Data pooled from 4 independent experiments. $E\alpha^{+/+}$ (n = 7), $E\alpha^{+/\Delta E1\Delta E3}$ (n = 8), and $E\alpha^{\Delta E1\Delta E3/\Delta E1\Delta E3}$ (n = 9). In all cases, summary graph data for all genotypes are presented with normalization to the average value for $E\alpha^{+/+}$ (set to 1) within individual experiments. Data are plotted as mean \pm SD. Statistical analysis: one-way ANOVA with correction for multiple comparison using Tukey's *post hoc* testing. *p < 0.05, **p < 0.01, ****p < 0.0001.

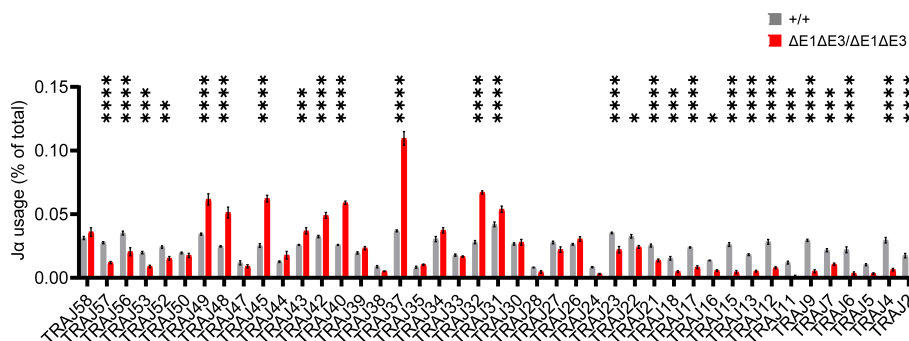


FIGURE 6

The J α repertoire is dysregulated in $\Delta E1\Delta E3$ mice. J α segment usage in pre-selection DP thymocytes from $E\alpha^{+/+}$ (n = 3) and $E\alpha^{\Delta E1\Delta E3/\Delta E1\Delta E3}$ (n = 3) assessed by 5' RACE. Values shown represent the frequencies with which individual J segments appear in the *TcrA* repertoire and are presented as mean \pm SEM. Statistical analysis: two-way ANOVA with correction for multiple comparison using Sidák *post hoc* testing. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

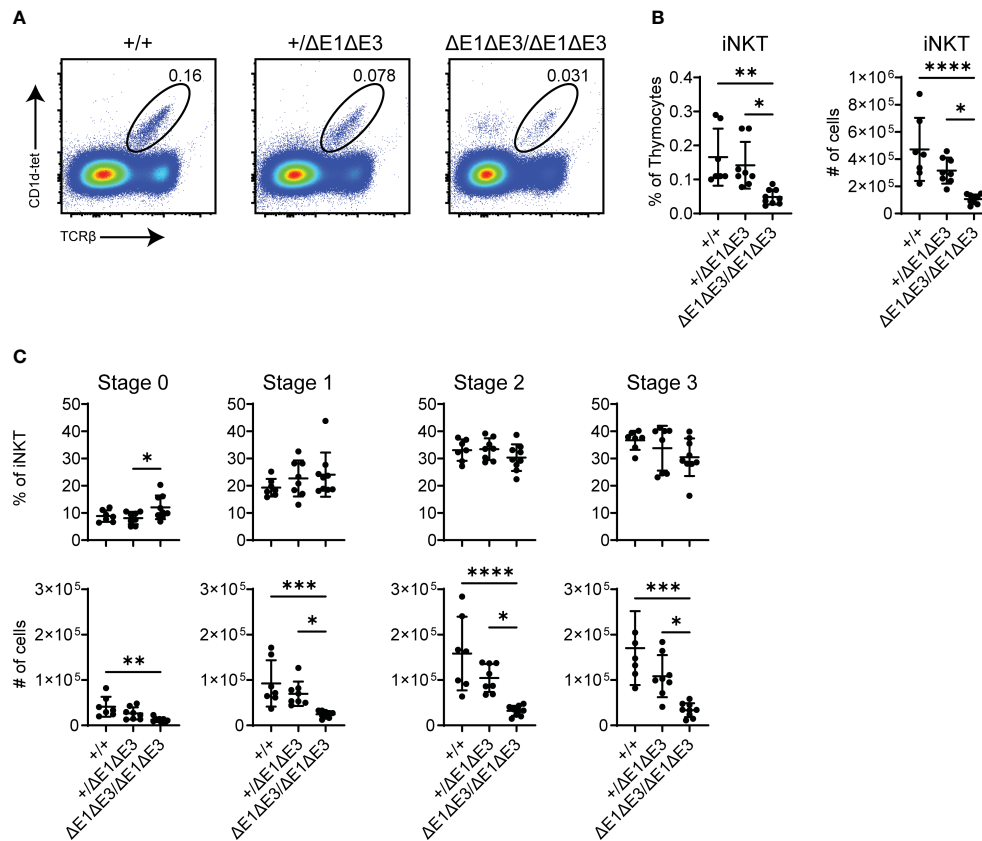


FIGURE 7

Fewer iNKT cells develop in $\Delta E1\Delta E3$ mice. (A) Representative flow cytometry plots, displayed as CD1d-tetramer versus TCR β , of live thymocytes from $E\alpha^{+/+}$, $E\alpha^{+/\Delta E1\Delta E3}$, and $E\alpha^{\Delta E1\Delta E3/\Delta E1\Delta E3}$ mice. Frequencies of gated populations are shown. (B) Frequencies and numbers of thymic iNKT cells in $E\alpha^{+/+}$ ($n = 15$), $E\alpha^{+/\Delta E1\Delta E3}$ ($n = 19$), and $E\alpha^{\Delta E1\Delta E3/\Delta E1\Delta E3}$ ($n = 15$) mice, with gating as shown in (A). Data pooled from 8 independent experiments. (C) Frequencies and numbers of thymic iNKT stage 0 (CD24⁺), stage 1 (CD24⁺CD44⁺NK1.1⁻), stage 2 (CD24⁺CD44⁺NK1.1⁺), and stage 3 (CD24⁺CD44⁺NK1.1⁺) cells in $E\alpha^{+/+}$ ($n = 7$), $E\alpha^{+/\Delta E1\Delta E3}$ ($n = 8$) and $E\alpha^{\Delta E1\Delta E3/\Delta E1\Delta E3}$ ($n = 9$) mice. Data were pooled from 4 independent experiments and are presented as mean \pm SD. Statistical analysis: one-way ANOVA with correction for multiple comparison using Tukey's *post hoc* testing. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

absence of the TEA promoter, or the TEA and J α 49 promoters, primary recombination is dysregulated and there is elevated usage of J α gene segments across the central and distal portions of the J α array (12, 15). However, relative usage of J α segments distal to J α 31 was reduced in $\Delta E1\Delta E3$ mice. This, together with the reduction in iNKT cells, indicates a substantial impairment of secondary V α -J α recombination in $\Delta E1\Delta E3$ mice. This suggests that E protein binding at E α may additionally regulate the promoters of rearranged V α gene segments, even though there is no effect on those promoters when they are more distant in the unrearranged locus.

The reported findings show that E protein binding at E α is important to but not necessary for $\alpha\beta$ T cell development. Ablation of all E α E-boxes may produce a more prominent developmental block at DP, but this may not be equivalent to $\Delta E\alpha$ given that other TFs may retain binding ability and thus maintain sufficient E α function to permit TCR α chain recombination, albeit at reduced levels. The importance of E protein binding is further emphasized by the broad conservation of E-boxes at E α in mammals (Supplementary Figure 8). The present results add to the

literature on E protein control of AgR gene recombination, with E proteins now shown to regulate all AgR loci. In contributing to locus accessibility and regulation of V α -J α rearrangement, E α -bound E proteins increase TCR α repertoire diversity and the potential for antigen recognition by $\alpha\beta$ T cells.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repository and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/geo/>, GSE227164.

Ethics statement

The animal study was reviewed and approved by Duke University Institutional Animal Care and Use Committee.

Author contributions

AM, SR, MK, and YZ contributed to conception and design of the study. Investigation and analysis was carried out by AM. AM and MK contributed to writing of manuscript original draft. AM, SR, MK, and YZ contributed to review, editing, and approval of submitted manuscript.

Funding

This study was funded by the NIH (R01-GM059638 and P01-AI102853 to YZ, and R35 GM136284 to MK).

Acknowledgments

We thank J. Mohammed, M. Cook, N. Martin, K. Parr, L. Martinek, and B. Li of the Duke University Cancer Institute Flow Cytometry Shared Resource for technical support and cell sorting. We thank G. Kucera, C. Bock, and M. Flowers of the Duke University Cancer Institute Transgenic and Knockout Mouse Shared Resource for assistance in generating mouse lines reported in this study. We thank J. Ratiu for advice on *Tcra* repertoire library and analysis methods. We thank the Duke University Cancer

Institute Sequencing and Genomic Technologies Shared Resource for assistance in sequencing TCR α repertoire library.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2023.1188738/full#supplementary-material>

References

- Schatz DG, Ji Y. Recombination centres and the orchestration of V(D)J recombination. *Nat Rev Immunol* (2011) 11(4):251–63. doi: 10.1038/nri2941
- Helmink BA, Sleckman BP. The response to and repair of RAG-mediated DNA double-strand breaks. *Annu Rev Immunol* (2012) 30:175–202. doi: 10.1146/annurev-immunol-030409-101320
- Lin SG, Ba Z, Alt FW, Zhang Y. RAG chromatin scanning during V(D)J recombination and chromatin loop extrusion are related processes. *Adv Immunol* (2018) 139:93–135. doi: 10.1016/bs.ai.2018.07.001
- Ciofani M, Zúñiga-Pflücker JC. Determining $\gamma\delta$ versus $\alpha\beta$ T cell development. *Nat Rev Immunol* (2010) 10(9):657–63. doi: 10.1038/nri2820
- Carico Z, Krangel MS. Chromatin dynamics and the development of the TCR α and TCR δ repertoires. *Adv Immunol* (2015) 128:307–61. doi: 10.1016/bs.ai.2015.07.005
- Sleckman BP, Bardon CG, Ferrini R, Davidson L, Alt FW. Function of the TCR alpha enhancer in alphabeta and gammadelta T cells. *Immunity* (1997) 7(4):505–15. doi: 10.1016/S1074-7613(00)80372-6
- Ho JC, Yang LH, Morle G, Leiden JM. A T-cell-specific transcriptional enhancer element 3' of c alpha in the human T-cell receptor alpha locus. *Proc Natl Acad Sci U S A*. (1989) 86(17):6714–8. doi: 10.1073/pnas.86.17.6714
- Winoto A, Baltimore D. A novel, inducible and T cell-specific enhancer located at the 3' end of the T cell receptor alpha locus. *EMBO J* (1989) 8(3):729–33. doi: 10.1002/j.1460-2075.1989.tb03432.x
- Balmelle N, Zamarreño N, Krangel MS, Hernández-Munain C. Developmental activation of the TCR alpha enhancer requires functional collaboration among proteins bound inside and outside the core enhancer. *J Immunol* (2004) 173(8):5054–63. doi: 10.4049/jimmunol.173.8.5054
- Hernández-Munain C. Recent insights into the transcriptional control of the *Tcra/Tcrd* locus by distant enhancers during the development of T-lymphocytes. *Transcription* (2015) 6(4):65–73. doi: 10.1080/21541264.2015.1078429
- del Blanco B, García-Mariscal A, Wiest DL, Hernández-Munain C. Tcr alpha enhancer activation by inducible transcription factors downstream of pre-TCR signaling. *J Immunol* (2012) 188(7):3278–93. doi: 10.4049/jimmunol.1100271
- Hawwari A, Bock C, Krangel MS. Regulation of T cell receptor alpha gene assembly by a complex hierarchy of germline alpha promoters. *Nat Immunol* (2005) 6(5):481–9. doi: 10.1038/ni1189
- Villey I, Caillol D, Selz F, Ferrier P, de Villartay JP. Defect in rearrangement of the most 5' TCR-J alpha following targeted deletion of T early alpha (TEA): implications for TCR alpha locus accessibility. *Immunity* (1996) 5(4):331–42. doi: 10.1016/S1074-7613(00)80259-9
- Ji Y, Little AJ, Banerjee JK, Hao B, Oltz EM, Krangel MS, et al. Promoters, enhancers, and transcription target RAG1 binding during V(D)J recombination. *J Exp Med* (2010) 207(13):2809–16. doi: 10.1084/jem.20101136
- Hawwari A, Krangel MS. Role for rearranged variable gene segments in directing secondary T cell receptor alpha recombination. *Proc Natl Acad Sci U S A*. (2007) 104(3):903–7. doi: 10.1073/pnas.0608248104
- del Blanco B, Angulo Ú, Krangel MS, Hernández-Munain C. T-Cell receptor α enhancer is inactivated in $\alpha\beta$ T lymphocytes. *Proc Natl Acad Sci U S A*. (2015) 112(14):E1744–53. doi: 10.1073/pnas.1406551112
- Belle I, Zhuang Y. E proteins in lymphocyte development and lymphoid diseases. *Curr Top Dev Biol* (2014) 110:153–87. doi: 10.1016/B978-0-12-405943-6.00004-X
- Miyazaki K, Watanabe H, Yoshikawa G, Chen K, Hidaka R, Aitani Y, et al. The transcription factor E2A activates multiple enhancers that drive rag expression in developing T and B cells. *Sci Immunol* (2020) 5(51). doi: 10.1126/sciimmunol.abb1455
- Yoshikawa G, Miyazaki K, Ogata H, Miyazaki M. The evolution of rag gene enhancers and transcription factor e and id proteins in the adaptive immune system. *Int J Mol Sci* (2021) 22(11). doi: 10.3390/ijms22115888
- Murre C. Epigenetics of antigen-receptor gene assembly. *Curr Opin Genet Dev* (2007) 17(5):415–21. doi: 10.1016/j.gde.2007.08.006
- Inlay MA, Tian H, Lin T, Xu Y. Important roles for e protein binding sites within the immunoglobulin kappa chain intronic enhancer in activating ν kappa jkappa rearrangement. *J Exp Med* (2004) 200(9):1205–11. doi: 10.1084/jem.20041135
- Jones ME, Zhuang Y. Regulation of V(D)J recombination by e-protein transcription factors. *Adv Exp Med Biol* (2009) 650:148–56. doi: 10.1007/978-1-4419-0296-2_12
- Jones ME, Zhuang Y. Acquisition of a functional T cell receptor during T lymphocyte development is enforced by HEB and E2A transcription factors. *Immunity* (2007) 27(6):860–70. doi: 10.1016/j.immuni.2007.10.014

24. Jones-Mason ME, Zhao X, Kappes D, Lasorella A, Iavarone A, Zhuang Y. E protein transcription factors are required for the development of CD4(+) lineage T cells. *Immunity* (2012) 36(3):348–61. doi: 10.1016/j.immuni.2012.02.010
25. Kadakia T, Tai X, Kruhlak M, Wisniewski J, Hwang IY, Roy S, et al. E-protein-regulated expression of CXCR4 adheres preselection thymocytes to the thymic cortex. *J Exp Med* (2019) 216(8):1749–61. doi: 10.1084/jem.20182285
26. D'Cruz LM, Knell J, Fujimoto JK, Goldrath AW. An essential role for the transcription factor HEB in thymocyte survival, tcr rearrangement and the development of natural killer T cells. *Nat Immunol* (2010) 11(3):240–9. doi: 10.1038/ni.1845
27. Hernández-Munain C, Sleckman BP, Krangel MS. A developmental switch from TCR delta enhancer to TCR alpha enhancer function during thymocyte maturation. *Immunity* (1999) 10(6):723–33. doi: 10.1016/S1074-7613(00)80071-0
28. Dauphars DJ, Mihai A, Wang L, Zhuang Y, Krangel MS. Trav15-dv6 family tcrd rearrangements diversify the tcr repertoire. *J Exp Med* (2022) 219(2). doi: 10.1084/jem.20211581
29. Mamedov IZ, Britanova OV, Zvyagin IV, Turchaninova MA, Bolotin DA, Putintseva EV, et al. Preparing unbiased T-cell receptor and antibody cDNA libraries for the deep next generation sequencing profiling. *Front Immunol* (2013) 4:456. doi: 10.3389/fimmu.2013.00456
30. Turchaninova MA, Davydov A, Britanova OV, Shugay M, Bikos V, Egorov ES, et al. High-quality full-length immunoglobulin profiling with unique molecular barcoding. *Nat Protoc* (2016) 11(9):1599–616. doi: 10.1038/nprot.2016.093
31. Shugay M, Britanova OV, Merzlyak EM, Turchaninova MA, Mamedov IZ, Tuganbaev TR, et al. Towards error-free profiling of immune repertoires. *Nat Methods* (2014) 11(6):653–5. doi: 10.1038/nmeth.2960
32. Bolotin DA, Poslavsky S, Mitrophanov I, Shugay M, Mamedov IZ, Putintseva EV, et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat Methods* (2015) 12(5):380–1. doi: 10.1371/journal.pcbi.1004503
33. Shugay M, Bagaev DV, Turchaninova MA, Bolotin DA, Britanova OV, Putintseva EV, et al. VDJtools: unifying post-analysis of T cell receptor repertoires. *PLoS Comput Biol* (2015) 11(11):e1004503. doi: 10.1038/s41467-017-00380-3
34. Carr TM, Wheaton JD, Houtz GM, Ciofani M. JunB promotes Th17 cell identity and restrains alternative CD4(+) T-cell programs during inflammation. *Nat Commun* (2017) 8(1):301. doi: 10.1038/sj.emboj.7601866
35. Abarrategui I, Krangel MS. Noncoding transcription controls downstream promoters to regulate T-cell receptor alpha recombination. *EMBO J* (2007) 26(20):4380–90. doi: 10.1038/ni1379
36. Abarrategui I, Krangel MS. Regulation of T cell receptor-alpha gene recombination by transcription. *Nat Immunol* (2006) 7(10):1109–15. doi: 10.4049/jimmunol.2000864
37. Rodríguez-Caparrós A, Álvarez-Santiago J, López-Castellanos L, Ruiz-Rodríguez C, Valle-Pastor MJ, López-Ros J, et al. Differently regulated gene-specific activity of enhancers located at the boundary of subtopologically associated domains: TCR α enhancer. *J Immunol* (2022) 208(4):910–28. doi: 10.1038/ni791
38. Guo J, Hawwari A, Li H, Sun Z, Mahanta SK, Littman DR, et al. Regulation of the TCRalpha repertoire by the survival window of CD4(+)CD8(+) thymocytes. *Nat Immunol* (2002) 3(5):469–76. doi: 10.1016/j.immuni.2005.03.011
39. Egawa T, Eberl G, Taniuchi I, Benlagha K, Geissmann F, Hennighausen L, et al. Genetic evidence supporting selection of the Valpha14i NKT cell lineage from double-positive thymocyte precursors. *Immunity* (2005) 22(6):705–16. doi: 10.1073/pnas.0408449102
40. Bezbradica JS, Hill T, Stanic AK, Van Kaer L, Joyce S. Commitment toward the natural T (iNKT) cell lineage occurs at the CD4+8+ stage of thymic ontogeny. *Proc Natl Acad Sci U S A*. (2005) 102(14):5114–9. doi: 10.3389/fimmu.2018.00042
41. Roy S, Moore AJ, Love C, Reddy A, Rajagopalan D, Dave SS, et al. Id proteins suppress E2A-driven invariant natural killer T cell development prior to TCR selection. *Front Immunol* (2018) 9:42. doi: 10.3390/ijms21228478
42. Rodríguez-Caparrós A, Álvarez-Santiago J, Del Valle-Pastor MJ, Suñé C, López-Ros J, Hernández-Munain C. Regulation of T-cell receptor gene expression by three-dimensional locus conformation and enhancer function. *Int J Mol Sci* (2020) 21(22). doi: 10.1038/ni1260
43. Murre C. Helix-loop-helix proteins and lymphocyte development. *Nat Immunol* (2005) 6(11):1079–86. doi: 10.1038/nmeth.3364