

The pioneer transcription factors Foxa1 and Foxa2 regulate alternative RNA splicing during thymocyte positive selection

Ching-In Lau¹, Jasmine Rowell¹, Diana C. Yanez¹, Anisha Solanki¹, Susan Ross¹, Masahiro Ono², and Tessa Crompton¹✉

¹UCL Great Ormond Street Institute of Child Health 30 Guilford Street, London WC1N 1EH, UK

²Department of Life Sciences, Imperial College London, London, UK

✉Address correspondence to Tessa Crompton, Phone: +44 (0)207 905 2893; Fax: +44 (0)207 813 8494; email: t.crompton@ucl.ac.uk

Key words: Foxa1, Foxa2, T-cell development, RNA splicing, Positive selection, Thymus

Abstract

During positive selection at the transition from CD4⁺CD8⁺ double-positive (DP) to single positive (SP) thymocyte, TCR signalling results in appropriate MHC-restriction and signals for survival and progression. We show that the pioneer transcription factors Foxa1 and Foxa2 are required to regulate RNA splicing during positive selection. Foxa1 and Foxa2 had overlapping/compensatory roles. Conditional deletion of both Foxa1 and Foxa2 from DP thymocytes reduced positive selection and development of CD4SP, CD8SP and peripheral naïve CD4⁺T-cells. Foxa1 and Foxa2 regulated expression of many genes encoding splicing-factors and regulators, including *Mbnl1*, *H1f0*, *Sf3b1*, *Hnrnpa1*, *Rnpc3*, *Prpf4b*, *Prf40b* and *Snrpd3*. Within the positively selecting CD69⁺DP cells, alternative RNA splicing was dysregulated in the double Foxa1/Foxa2 conditional knockout, leading to >850 differentially used exons (DEU). Many genes important for this stage of T-cell development (*Ikzf1-3*, *Ptprc*, *Stat5a*, *Stat5b*, *Cd28*, *Tcf7*) and splicing-factors (*Hnrnpab*, *Hnrnpa2b1*, *Hnrnpu*, *Hnrnpul1*, *Prpf8*) showed multiple DEU. Thus Foxa1 and Foxa2 are required during positive selection

to regulate alternative splicing of genes essential for T-cell development, and that by also regulating splicing of splicing-factors, they exert widespread control of alternative splicing.

Introduction

The production of $\alpha\beta$ T-cells in the thymus involves multiple stages of development during which haematopoietic precursors give rise to mature T-cells that can differentiate into functional effector T-cells. During this process, progenitor cells which do not express the coreceptor molecules CD4 and CD8 ($CD4^-CD8^-$ double negative (DN) cells) differentiate to become $CD4^+CD8^+$ double positive (DP) cells, which give rise to both CD4 single positive (SP) and CD8SP populations. Maturation from DP to SP follows successful rearrangement of the *Tcr α* locus, and requires TCR signalling: positive selection results in appropriate MHC restriction of SP cells, and is followed by negative selection of potentially self-reactive clones and selection of regulatory T-cells (Treg)(1-3). The strength and duration of the TCR signal that a developing cell receives broadly determine its fate, with the strongest signals leading to negative selection or CD4 Treg differentiation, usually at the SP stage in the medulla, intermediate signals leading to positive selection usually in the cortex, and weaker signals or lack of TCR signalling leading to death by neglect(4). For DP thymocytes undergoing positive selection, TCR signal strength and duration also influence CD4 and CD8 lineage choice. Those cells receiving stronger and longer TCR signals tend towards the CD4SP fate, while weaker/more transient signals favour CD8SP fate, and fate decisions are also influenced by the relative timing of cytokine and TCR signalling that a developing cell receives (2, 5, 6). Many models have been proposed to describe this process and to explain how positive selection ensures that CD4SP and CD8SP populations express TCR appropriately restricted by MHCII and MHCI, respectively(2, 3, 7). Currently, the consensus favours the kinetic signalling model(2, 4, 8), in which CD8 is downregulated first during positive selection, leading to a $CD4^+CD8^{lo}$ intermediate, with continued CD4 co-receptor expression allowing for prolonged stronger MHCII-TCR signalling, leading to differentiation to CD4SP; whereas cytokine signalling through the common gamma chain activates Stat5a and Stat5b, and rescues cells that have received an interrupted MHCI-TCR signal to induce differentiation to

CD8SP(9, 10). The CD4/CD8 lineage decision is also influenced by factors from the stroma, such as Notch and Hedgehog (Hh) signalling(11-14).

Many transcription factors contribute to the regulation of the CD4/CD8 lineage decision (2, 7, 15, 16). Additionally, epigenetic processes, such as DNA methylation and histone modifications, may be involved in locking-in the pattern of gene expression to generate stable CD4SP and CD8SP lineages(17), and potentially also in preparing for initiation of a particular programme of differentiation.

Here we investigate the role of the transcription factors Foxa1 and Foxa2 in T-cell development. The Foxa proteins are a highly conserved subfamily of forkhead box transcription factors, which contain unique wing-helix DNA-binding domains(18). The Foxa proteins can function as pioneer transcription factors, which by binding silent (condensed) chromatin early in a developmental programme prior to target gene activation, can act either to open up local chromatin, imparting competence to other transcriptional activators to initiate a developmental lineage, or to directly facilitate other factors binding to nucleosomal DNA(19, 20). Foxa2 has recently also been shown to demethylate tissue-specific regions of DNA, to generate stable lineage-specific DNA methylation patterns to enhance gene expression(21).

Foxa1 and Foxa2 proteins are closely related to each other and are widely co-expressed during embryogenesis and in several tissues postnatally including lung, liver, intestines, pancreas and thymus (12, 22-26). Genetic ablation of Foxa1 or Foxa2 in mice showed that they are both required for normal development during embryogenesis. Foxa2-deficient embryos display severe defects in notochord, floorplate and endoderm and die at E10-11, whereas Foxa1-null mice exhibit defects in the regulation of glucose homeostasis and die postnatally(27, 28). Foxa1 and Foxa2 play over-lapping and compensatory roles in the regulation of development of lung, liver and pancreas(22, 29-32).

Foxa2 is a well-recognized target gene of the Shh signalling pathway in the floorplate, and is required for the maintenance of Shh expression, but can also negatively regulate expression of components of the Shh signalling pathway(33-35). Expression of Foxa1 is also correlated to Hh signalling activities, and Foxa1-deficient mice exhibited elevated Shh and Gli2 expression in prostate, suggesting that Foxa1 can negatively regulate Shh signalling(36).

Little is known about the function of *Foxa1* and *Foxa2* in the immune system. In mouse models of autoimmune inflammation, ectopic *Foxa1* expression has been shown to drive differentiation and suppressive function of a novel subset of Treg(37). We recently found that expression of *Foxa1* and *Foxa2* in thymic epithelial cells (TEC) is required to maintain normal T-cell development and homeostasis of thymic and spleen Treg populations. Conditional deletion of *Foxa1* and *Foxa2* from TEC increased the proportion of medullary TEC, but reduced cell-surface MHCII expression on TEC, leading to a smaller thymus with a reduction in conventional CD4 T-cell differentiation but an increase in the CD4 Treg population(25). In the thymus, *Foxa1* and *Foxa2* are also expressed in developing T-cells, and *Foxa2* is a transcriptional target of *Shh* signalling after pre-TCR signal transduction(12, 26).

Here we investigate the function of *Foxa1* and *Foxa2* in developing $\alpha\beta$ T-cells. We show that *Foxa1* and *Foxa2* are required at the transition from DP to SP T-cell, where they regulate RNA splicing. Conditional deletion of both *Foxa1* and *Foxa2* led to a reduction in positive selection, differentiation and maturation of the SP populations, and a reduction in the peripheral CD4 T-cell pool.

Results

CD4SP and CD8SP development are impaired in the *Foxa1/2*CKO thymus

We examined *Foxa1* and *Foxa2* expression by QRT-PCR in FACS-sorted developing thymocytes from the DN3 stage onwards (Fig 1A). *Foxa1* and *Foxa2* were detected in all stages, and showed reciprocal patterns of expression, with *Foxa1* expression upregulated in DN4, following by decline in DP cells. In contrast, *Foxa2* was highly expressed in DN3 and DP stages, and expressed at lower levels in DN4 cells. Both *Foxa1* and *Foxa2* showed higher expression in CD4SP than in CD8SP cells.

To establish if *Foxa1* and *Foxa2* play a T-cell intrinsic role in thymocyte development, we conditionally deleted *Foxa1* and/or *Foxa2* from T-cells from the DP stage of development onwards, by crossing mice carrying a single or double loxP-flanked *Foxa1* and/or *Foxa2* allele to mice in which Cre is driven by the CD4-promotor. *Foxa1*^{fl/fl}CD4cre⁺ and *Foxa2*^{fl/fl}CD4cre⁺

mutant mice are referred to as Foxa1cKO and Foxa2cKO, respectively. Foxa1^{fl/fl}Foxa2^{fl/fl}CD4cre⁺ double mutant mice are referred to as Foxa1/2cKO. *Foxa1* and *Foxa2* were effectively deleted in Foxa1/2cKO CD4SP and CD8SP thymocytes, as expression of *Foxa1* and *Foxa2* were below detection by QRT-PCR, but detected in the control (Foxa1^{fl/fl}Foxa2^{fl/fl}CD4cre⁻) CD4SP and CD8SP populations (Fig 1B).

We compared thymocyte populations in Foxa1cKO, Foxa2cKO and Foxa1/2cKO thymus with their control littermates. The number of thymocytes were not significantly different between the conditional Foxa1 and/or Foxa2 mutants and their controls (Fig 1C). The Foxa1cKO thymus contained normal proportions of DP and SP populations (Fig 1D), but there was an increase in the proportion of DP cells in the Foxa2cKO compared to control (Fig 1E,G). In the Foxa1/2cKO thymus, the proportion and number of the DP population were increased and the proportion and number of CD4SP and CD8SP populations were reduced (Fig 1F-I). The fact that while the phenotype of the double Foxa1/2cKO thymus was more pronounced than that of the Foxa2cKO, the Foxa1cKO thymus appeared grossly normal, suggests that Foxa1 can compensate for Foxa2 at this developmental transition, and we therefore decided to use the double Foxa1/2cKO mice to investigate the impairment of differentiation from DP to SP cell.

The ratios of DP:SP, DP:CD4SP and DP:CD8SP were all increased in Foxa1/2cKO compared to control, while the ratio of CD4SP:CD8SP was decreased (Fig 1J). Thus, deficiency of both Foxa1 and Foxa2 in thymocytes led to less efficient development of SP populations and bias towards the generation of CD8 lineage cells over CD4SP. During differentiation from DP to SP stage, cell surface TCR expression is upregulated. The proportion of TCR β ^{hi} thymocytes was decreased in the absence of Foxa1 and Foxa2 (Fig 1K). When we gated on TCR β ^{hi} cells, and compared subset distribution, we found an increase in the proportion of DP and CD4⁺CD8^{lo} intermediate cells, and a decrease in the proportion of CD4SP population, indicating that fewer cells were completing positive selection and a partial arrest at the transition from DP and CD4⁺CD8^{lo} to CD4SP (Fig 1L).

Foxa1/2 deficiency influences maturation of SP cells and peripheral T-cell populations

We then investigated the maturation status of SP cells in the Foxa1/2cKO thymus by expression of HSA, CD69, Qa2 and CD62L. After positive selection, SP thymocytes retain high expression of heat-stable antigen (HSA (CD24)) and CD69, and as they mature HSA and CD69 are downregulated, Qa2 is upregulated, and expression of CD62L indicates that thymocytes are mature and ready to egress from the thymus. The proportion of HSA⁺CD69⁺ cells in both CD4SP and CD8SP populations was significantly decreased in Foxa1/2cKO compared to control thymus (Fig 2A). Furthermore, there was a significant decrease in the percentage of Qa2⁺ and CD69⁻CD62L⁺ cells in the CD4SP compartment in Foxa1/2cKO compared to control (Fig 2B-C), whereas we did not detect significant differences in Qa2 and CD62L expression in the CD8SP population (Supplementary Fig 2). Thus, in the absence of Foxa1 and Foxa2, not only were there fewer CD4SP and CD8SP cells, but their maturation was also affected, such that fewer mature CD62L⁺CD4SP were produced to egress from the thymus. In contrast, we found no significant difference in the proportion of CD4⁺CD25⁺Foxp3⁺ Tregs, NKT cells or NK cells in the Foxa1/2cKO thymus compared to control (Fig 2D,E).

In the spleen and lymph nodes changes in the CD4⁺ T-cell populations mirrored the thymus in the Foxa1/2cKO compared to control, with reductions in the proportion and number of conventional CD4⁺ T-cells overall and of naive CD44⁻CD62L⁺CD4⁺ T-cells in spleen (Fig 2F-H). In contrast, we did not detect significant differences in the number of peripheral CD8⁺ T-cells in the Foxa1/2cKO compared to control, suggesting that despite the reduction in CD8SP cells in the thymus, the peripheral CD8⁺ T-cell compartment is subject to homeostatic control and can expand to reach its normal size (Fig. 2F-G).

Foxa1/2 deficiency reduces TCR signalling

TCR signal strength is one factor which determines positive selection in the thymus and tonic TCR signalling is again required for maintenance and homeostasis of peripheral T-cell populations after egress from the thymus. Therefore, as a proxy to measure TCR signal strength in thymocyte subsets, we compared expression of proteins whose levels of expression are directly determined by TCR signalling between Foxa1/2cKO and control.

Expression of intracellular Nr4a1 is induced as an early consequence of TCR signal transduction and high expression requires relatively strong TCR signalling(38, 39). The proportion of intracellular Nr4a1^{hi} cells was significantly reduced in Foxa1/2cKO DP, CD4SP and CD8SP populations compared to control, suggesting that in the absence of Foxa1 and Foxa2 fewer cells had reached the threshold of TCR signal strength required to induce high levels of Nr4a1 (Fig 3A). Levels of cell-surface CD5 expression correlate with the strength of TCR signal transduction that a developing T-cell has received(40). As expected, mean fluorescence intensity (MFI) of anti-CD5 staining was lower in CD8SP and DP cells than in CD4SP cells (Fig 3B). MFI of CD5 was lower in Foxa1/2cKO DP, CD4SP and CD8SP populations than in their control counterparts, consistent with reduced TCR signal strength in the absence of Foxa1 and Foxa2.

Foxa1 and Foxa2 promote positive selection in DP thymocytes

Positive selection signals for survival and further maturation and requires interactions between self-peptide:MHC complexes in the thymic cortex, which may take place over a number of days, involving multiple or prolonged TCR-MHC interactions(41-43). TCR signalling for positive selection leads to cell-surface CD69 expression followed by upregulation of the cell-surface TCR complex. The proportions of DP cells that expressed cell surface TCR β ^{hi} and CD69 were decreased in the Foxa1/2cKO thymus compared to control (Fig 3C-D), consistent with a reduction in positive selection. To test this and to investigate mechanisms that might account for the reduced transition from DP to SP, we carried out transcriptome analysis on cells undergoing positive selection by RNA-sequencing FACS-sorted CD69⁺DP cells. RNA-sequencing identified only 176 differentially expressed genes (DEG) between Foxa1/2cKO and control datasets (adjusted (FDR) $p < 0.05$), of which 109 (62%) were more highly expressed in Foxa1/2cKO than control (Fig 3E and Table S1). Of the 176 DEG, 85 (~48%) had previously been identified as Foxa1/2 targets in genome-wide ChIPSeq analysis of Foxa1/2 binding sites in dopaminergic neuronal progenitors and 56 of those genes with verified Foxa1/2 binding sites (approximately two thirds) showed higher expression in the Foxa1/2cKO than control (Fig 3F and Table S1)(34, 44). DEG included genes involved in T-cell development and function, and approximately one quarter of DEG were

genes whose transcription has been shown to be regulated during positive selection(45) (Fig 3G and Table S1). Among these, DEG that were more highly expressed in *Foxa1/2cKO* than control datasets included genes known to reduce TCR signal strength: *Cbl*, a ubiquitin ligase, which negatively regulates TCR signalling(46) and *Themis*, which can attenuate TCR signalling during repertoire selection(47); and also genes associated with the CD8 lineage: *Cd8b1*; and *Lyst*, a lysosomal trafficking regulator required for CTL lytic granules(48), consistent with the CD4 lineage being more severely affected in the conditional knockouts. To test in an unbiased way, if *Foxa1* and *Foxa2* are required for the transcriptional response to TCR signalling for positive selection, we then used canonical correspondence analysis (CCA) to compare the overall pattern of gene expression in our datasets with transcriptome data from publically available datasets prepared from DP thymocytes that were receiving different strengths of TCR signals during selection (GSE38909)(49). The GSE38909 dataset contains TCR-transgenic AND DP thymocytes stimulated with a positively selecting peptide (gp250) or a non-selecting control peptide(49). We selected the 2000 most significantly differentially expressed genes from the GSE38909 dataset between DP thymocytes stimulated with the non-selecting control peptide and DP thymocytes stimulated with the positively selecting peptide and used these to generate a scale of unstimulated to TCR-signalling-for-positive-selection. We plotted our datasets against this scale. The CCA segregated the datasets by genotype: both control datasets were on the positive side of the axis, which corresponded to the transcriptional pattern induced by the positively selecting peptide, consistent with the fact that CD69⁺DP thymocytes have initiated positive selection by TCR signal transduction (Fig 3H). In contrast, both *Foxa1/2cKO* datasets fell on the negative side of the axis, thus showing an overall pattern of transcription that is closer to that of unstimulated DP cells than that of their control counterparts.

To confirm this impact on positive selection, we subdivided thymocytes by cell surface TCR β and CD69 expression into four different stages: TCR β ^{lo/neg}CD69⁻ (pre-selection thymocytes), TCR β ^{int}CD69⁺ (intermediate transition, undergoing positive selection), TCR β ^{hi}CD69⁺ (thymocytes after TCR signalling for initiation of positive selection) and TCR β ^{hi}CD69⁻ (more mature population). The proportion of pre-selection thymocytes was higher in the *Foxa1/2cKO* than control thymus, whereas there was no significant difference in the

proportion of the $\text{TCR}\beta^{\text{int}}\text{CD69}^+$ population, and both $\text{TCR}\beta^{\text{hi}}\text{CD69}^+$ and $\text{TCR}\beta^{\text{hi}}\text{CD69}^-$ populations were reduced, indicating that *Foxa1* and *Foxa2* expression in thymocytes promotes initiation of the process of positive selection, and also progression of cells during the differentiation process (Fig 3I). The loss of cells from the $\text{TCR}\beta^{\text{hi}}\text{CD69}^+$ population, but not its precursor $\text{TCR}\beta^{\text{int}}\text{CD69}^+$ population, suggests impairment at a late stage of positive selection and a failure to progress, and consistent with this, the proportion of cells undergoing cell death (Annexin V⁺) was increased in the $\text{CD69}^+\text{DP}$ population in the *Foxa1/2cKO* thymus compared to control (Fig 3J). When we gated on the most mature $\text{TCR}\beta^{\text{hi}}\text{CD69}^-$ thymocyte population, and compared CD4/CD8 subset distribution, we found that the proportion of DP cells increased by more than two-fold in the *Foxa1/2cKO* thymus compared to control, whereas the proportion of CD4SP cells was significantly reduced (Fig 3K), demonstrating a clear requirement for *Foxa1* and *Foxa2* in normal differentiation to CD4SP and positive selection, and suggesting that differentiation is dysregulated in the *Foxa1/2cKO*, so that more cells that had upregulated cell surface TCR and downregulated CD69 were unable to progress beyond the DP stage.

Foxa1 and Foxa2 regulate exon usage

To investigate how *Foxa1/2* might regulate positive selection we identified genes that encode verified transcription factors amongst the DEG between *Foxa1/2cKO* and control $\text{CD69}^+\text{DP}$ datasets to look for known regulators of positive selection or differentiation at this developmental stage which might function downstream of *Foxa1/2*. Twenty-four transcription factors were found, of which 13 had previously been verified as *Foxa1/2* targets in genome-wide ChIPSeq analysis in neuronal progenitors (Fig 4A and Table S1). Amongst the transcription factors whose expression were downregulated by *Foxa1/2* deletion there were no obvious down-stream candidates that might function to promote positive selection, although several showed changes in expression consistent with the developmental phenotype: for example *Ptma*, an anti-apoptotic gene (50); and *Notch3* which is transcriptionally regulated during positive selection but not required for positive selection(45, 51). Likewise, we found no obvious candidate genes which might function to inhibit positive selection or TCR signal strength amongst those genes which encode

transcription factors and whose expression were upregulated in the absence of Foxa1/2. *Ikzf2* was more highly expressed in Foxa1/2cKO than control, but is not required for differentiation at this stage of development; *Elk4* was also more highly expressed, but is required for positive selection and upregulated by TCR signal transduction, rather than functioning as a negative regulator(39, 52-54).

We then carried out Gene ontology (GO) term enrichment analysis of all DEG between Foxa1/2cKO and control CD69⁺DP datasets. This revealed over-representation of genes associated with terms connected with immunity and T-cell mediated immunity, including lymphocyte migration, cytokine signalling, cell killing and apoptotic process (Fig 4B and Table S2). Thus, as well as an overall enrichment in immunity related terms, as expected from the cell type, there was enrichment for terms related to processes required for differentiation to SP, such as lymphocyte migration, as differentiating cells must migrate to the medulla to complete their maturation and repertoire selection. Several terms related to cytokine signalling are also pertinent to the partial arrest at the DP and CD4⁺CD8^{lo} stage in the Foxa1/2 cKO, as cytokine signalling rescues cells whose MHCII-restricted TCR signal has been interrupted by down-regulation of CD8, allowing differentiation to CD8SP, whereas cytokine signalling is inhibited for differentiation to the CD4SP compartment.

Additionally, there was enrichment for the term 'chromatin silencing', related to the known function of Foxa transcription factors as pioneer factors. The GO analysis also showed enrichment for several terms related to RNA splicing, and for terms connected to cell matrix adhesion (Fig 4B). Amongst the top 20 most statistically significant DEG were genes associated with these processes (Fig 4C): for example, *Mmp14*, a membrane bound matrix metalloproteinase involved in breakdown of the extracellular matrix(55); *H1f0*, an H1 linker histone necessary for the condensation of nucleosome chains into higher-order chromatin structures, which is involved in the regulation of mRNA splice site recognition(56, 57); and *Mbnl1*, a regulator of alternative splicing, which functions in the control of T-cell development, and is over-expressed in mixed lineage leukemia(58-61).

Several other RNA splicing genes were differentially expressed between *Foxa1/2*cKO and control datasets (Fig 4D). Given this, and the fact that the GO enrichment analysis identified several terms associated with RNA splicing, we hypothesized that *Foxa1/2* might regulate the transition from DP to SP cell and positive selection by influencing RNA splicing in developing T-cells. To test this, we compared differential exon usage (DEU) between *Foxa1/2*cKO and control CD69⁺DP RNA-sequencing datasets. This analysis identified 852 events of differential exon usage (FDR $p < 0.05$) between the conditional knockout and control, which involved 628 different genes, 222 (~35%) of which had previously been shown to bind to *Foxa1/2* in genome-wide ChIPSeq analysis of *Foxa1/2* binding sites in dopaminergic neuronal progenitors (Fig 4E and Table S3). Conditional deletion of *Foxa1/2* therefore led to a greater number of changes in exon usage affecting more genes than the number of individual genes that were differentially expressed. Intersection between DEG and genes which contained differentially used exons revealed just 22 DEG (whose overall expression was differentially regulated and which also showed differential usage of individual exons) (Fig 4F), 18 of which (>80%) had previously verified *Foxa1/2* binding sites (Table S1). This intersection included several genes involved in the regulation of RNA splicing, with known splicing variants (Fig 4G): *Ranbp2*, a nucleoporin protein, controls alternative splicing patterns during nuclear speckle formation(62); *Sf3b1*, a well known splicing factor, mutations of which lead to myelodysplasia and anemia by globally disrupted splicing(63, 64); and *Smg1* which is mutated in Acute Myeloid Leukemia, its depletion resulting in disruption of alternative splicing(65, 66).

GO term enrichment analysis of genes which showed DEU revealed over-representation of genes associated with terms connected with the known functions of *Foxa1/2* in metabolic processes (metabolic process, cellular glucose homeostasis, response to insulin) and with their known functions in other tissues as epigenetic regulators and pioneer transcription factors (e.g. chromosome organization, chromatin remodeling, regulation of histone methylation, DNA methylation, DNA conformational change) (Fig 5A and Table S4). Many terms associated with mRNA splicing were also over-represented (Fig 5A). Genes involved in mRNA splicing showed multiple changes in exon usage, indicating that mRNA splicing factors are themselves subject to alternative splicing in developing T-cells, and that the regulatory

effects on splicing of the Foxa transcription factors may be amplified by regulation of splicing of components of the splicing machinery (Fig 5B-D). Additionally, the enrichment analysis highlighted positive regulation of NFK β signalling, a pathway which has been shown to regulate alternative splicing in T-cells(67).

Foxa1 and Foxa2 regulate alternative splicing of essential genes for T-cell development

GO terms associated with T-cell development and function showed over-representation of DEU genes (Fig 6A): for example, thymic T-cell selection; T-cell receptor signalling pathway; regulation of CD8-positive $\alpha\beta$ differentiation; lymphocyte differentiation; CD4-positive alpha-beta T-cell differentiation. Thus, Foxa1/2 control exon usage for genes involved in processes essential for the transition from DP to SP cell. Multiple differentially used exons were found in key genes for this transition, such as *Ikzf1*, *Ptprc*, *Stat5a*, *Stat5b*, *Cd28* and *Klf13* (Fig 6B)(9, 68-71), and many other essential genes for T-cell development showed differential usage of a single exon (for example: *Notch1*, *Rorc*, *Socs1*, *Pten*, *Orai1*, *Cbfb* and *Ets2*) (Fig 6B and Table S3). Many of the T-lineage genes which contained significantly differential exon usage have been previously described to be alternatively spliced during lymphocyte development or activation, such as the *Ikaros* family members, *Fyn*, *Ptprc* (CD45), *Stat5a*, *Stat5b* and *Cd28*(72, 73)(Fig 6B-E). However, the differential exon usage caused by absence of Foxa1 and Foxa2 did not always correspond to well-described alternative spliced variants of these genes. For example, alternative splicing of *Ptprc* exons 4, 5 and 6 is functionally important in T-cell differentiation(73), but in our datasets despite variation in expression of exon 5, only exons 2, 9, 13 were significantly different between control and conditional knockout after adjustment for false discovery ($p < 0.05$) (Fig 6B and E). Likewise, for *Ikzf1*, *Ikzf2* and *Ikzf3* we identified differentially expressed exons that were distinct from the alternative splice variants described in mouse thymocytes(74), indicating that absence of Foxa1 and Foxa2 led to wide dysregulation in the RNA splicing of genes that would normally display splice variants (Fig 6B-C).

Given the many changes in exon usage observed in genes required for T-cell development and positive selection, it seemed likely that Foxa1/2's function during positive selection is to regulate RNA splicing of essential genes for developmental progression. Therefore, to investigate the impact of Foxa1/2 on the normal splicing of regulators of development, we identified transcription factors within the genes that showed DEU between Foxa1/2cKO and control. Of the 628 genes that showed DEU, 97 (15.5%), encoded transcription factors (Fig 7A and Table S2), many of which are important during T-cell development, and are specifically required for the transition from DP to SP cell (for example, *Yy1*, *Trim28*, *Tcf7*, *Tcf3*, *Tcf12*, *Stat5a*, *Stat5b*, *Sp3*, *Rorc*, *Klf13*, *Cbfb*) (Fig 7B).

Discussion

Here we show that the pioneer transcription factors Foxa1 and Foxa2 regulate alternative RNA splicing during T-cell development at the transition from DP to SP thymocyte. Conditional deletion of Foxa1 and Foxa2 from developing DP thymocytes led to reduced positive selection and a partial arrest at the transition from DP to SP thymocyte, with a reduction in SP cell maturation, and a reduced peripheral naïve CD4 T-cell population. Conditional deletion of Foxa1 and Foxa2 also led to significant changes in the expression of genes that regulate RNA splicing in cells undergoing positive selection, and concomitantly to >850 significantly differentially used exons.

Alternative RNA splicing is a mechanism that enables cells to generate many different proteins from a limited number of genes and is important in the regulation of development processes, including T-cell development(60, 67, 72, 75, 76). Our study indicates that Foxa1 and Foxa2 regulate the mRNA splicing of many genes that are important for progression at this developmental transition, and we therefore propose that aberrant RNA splicing of multiple genes may account at least in part for the reduction in positive selection and differentiation to SP cell observed when Foxa1 and Foxa2 were conditionally deleted from DP cells. In support of this, many of the genes that showed DEU, but were not DEG, between Foxa1/2cKO and control datasets are well-known to be required at this transition (eg. *Socs1*, *Ptpnc1*, *Rasgrp1*); and many encode transcriptional regulators of this stage of development (eg. *Ikzf1*, *Stat5a*, *Stat5b*, *Cbfb*, *Tcf7*, *Tcf3*, *Tcf12*, *Klf13*, *Sp3*, *Rorc*). GO term

enrichment analysis also highlighted terms associated with positive selection and differentiation to SP. In contrast, few DEG encoded known relevant transcription factors or regulators of TCR repertoire selection or differentiation from DP to SP, and *Themis*, a DEG with verified Foxa1/2 binding site, which was more highly expressed in the conditional knockout, and which can modulate TCR signalling in thymocytes, also showed differential exon usage. The way in which the transcriptional activity of Foxa1/2 relate to their regulation of splicing will require further investigation, as there is increasing evidence that splicing can occur co-transcriptionally as well as post-transcriptionally(77). Interestingly, approximately one third of genes which showed DEU but were not DEG had previously been shown to bind directly to Foxa1/2 in neuronal progenitors, suggesting that Foxa1/2 may act directly to regulate their splicing.

Overall, we identified only 176 DEG between Foxa1/2cKO and control CD69⁺DP cells, and of these more than 60% were more highly expressed in the conditional knockout than control, suggesting either that they were directly repressed by Foxa1/2, most likely by association of Foxa1/2 with a co-repressor, or that Foxa1/2 activate the transcription of an intermediate transcriptional repressor. Approximately 48% of DEG have been previously been shown to bind Foxa1/2 in whole genome ChipSeq screen of neuronal progenitors, suggesting that they were also likely to be direct targets (directly bound by Foxa1/2) in developing T-cells. These were distributed between DEG that were up- or down-regulated in the Foxa1/2cKO compared to control, and ~51% of up-regulated genes have previously been shown to bind Foxa1/2 by ChipSeq screening in other cell types(34, 44). We therefore think it is most likely that absence of Foxa1/2 in CD69⁺DP cells led to increased expression of these DEG because of a direct repressive impact of Foxa1/2 binding to these sites and recruitment of a co-repressor in control cells, rather than indirectly by transcriptional activation of an unknown downstream repressor of transcription. Although in cell transfection assays Foxa factors behave as transcriptional activators, Foxa1/2 have previously been shown to have repressor activity in other developing tissues, and Foxa2 has been shown to interact with the transcriptional co-repressor Tle family of proteins(34, 78, 79). Further investigation of Foxa2/Tle interactions in the regulation of differentiation from DP to SP cell will therefore

be important, given that Tle1, Tle3 and Tle4 are together required for commitment to the CD8 lineage at the transition from DP to SP(80).

Foxa1/2 conditional deletion led to increased transcription of several RNA splicing regulators, including *Sf3b1*, *Smg1*, and *Mbnl1*, mutation of which are associated with haematological malignancies and aberrant RNA splicing, but also significantly decreased transcription of splicing factors (*Prpf40b*, *Hnrnpa1*, *Snrpd3*) and the histone linker *H1f0*, which may be required for mRNA splice site recognition(56, 58-61, 63-66). Overall loss of Foxa1 and Foxa2 led to broad changes in splicing events, indicating that Foxa1 and Foxa2 are important regulators of mRNA processing during T-cell development. Foxa1/2 also regulated the mRNA splicing of many genes that encode splicing factors, suggesting that their influence on the splicing machinery is further amplified by regulation of alternative splicing of splicing components, as well as by a more direct effect on the transcription of splicing regulators, RNA binding and histone linker proteins.

Mbnl1 is highly expressed in the thymus and a recent study demonstrated that its constitutive knock out led to a hyperplastic thymus with retention of thymocytes and many mis-splicing events(60). *Mbnl1* was upregulated approximately two-fold in CD69⁺DP cells in the absence of Foxa1 and Foxa2 and was one of the most significantly differentially expressed genes in our datasets. In Foxa1/2cKO DP thymocytes we observed differential exon usage in some of the same genes that were affected when *Mbnl1* was constitutively deleted: for example, *Tcf7* exon(E)10 was affected by absence of Mbnl1, whereas our datasets showed changes in *Tcf7* E1 and E6; *Map4k4* E20 was affected in Mbnl1^{-/-} thymus, but Foxa1/2cKO led to changes in *Map4k4* E15 and E33; *Sptan1* E23 was affected in Mbnl1^{-/-} thymus, but Foxa1/2cKO changed usage of *Sptan1* E2, E11, E12 and E49(60).

Conditional deletion of Foxa1 alone from DP thymocytes did not grossly affect $\alpha\beta$ T-cell development, and the double Foxa1/2cKO had a stronger impact on differentiation from DP to SP than the single Foxa2cKO, suggesting overlapping or redundant functions, so that while Foxa1 may not be required at this developmental transition it can partially replace the requirement for Foxa2. Foxa1 and Foxa2 have compensatory and partially redundant roles in other tissues, including liver(22) and a recent study showed that conditional deletion of the third Foxa protein (Foxa3) in addition to Foxa1 and Foxa2 in adult liver abrogated liver

gene regulatory networks, destroying liver tissue homeostasis and function in the adult(30). It will therefore be important in the future to investigate the potential compensatory role of Foxa3 in T-cell development.

Materials and Methods

Mice

CD4-Cre transgenic mice were purchased from the Jackson Laboratory. Foxa1^{flox/flox} Foxa2^{flox/flox} mice were provided by Dr. Siew-Lan Ang(81). These mice were backcrossed to C57BL/6 strain for at least 6 generations. C57BL/6 mice were purchased from Envigo (UK). All mice were bred and maintained at University College London (UK) under UK Home Office regulations.

To selectively delete Foxa1 and/or Foxa2 in T-cells, we crossed Foxa1^{flox/flox} and Foxa2^{flox/flox} mice with CD4-cre⁺ mice. Subsequently, Foxa1^{flox/flox} Foxa2^{flox/flox} CD4-cre⁺ mice were bred. In this paper, we will refer these mice as Foxa1/2cKO. Similarly, the Foxa1^{flox/flox} CD4-cre⁺ is termed Foxa1cKO and the Foxa2^{flox/flox} CD4-cre⁺ is termed Foxa2cKO. For all experiments their control genotype is the CD4-cre⁻ littermate with the same floxed alleles as the experimental animals.

Genotyping

DNA from mice was extracted by lysis buffer containing 50mM KCL, 10mM Tris HCL (pH8.5), 1.5mM MgCL₂, 0.01% gelatin, 0.45% Nonidet P-40, 0.45% Tween 20, and 0.5µg/ml Protinease K in ultra pure water and incubating at 56°C overnight. PCR analysis was as described(82). Primer sequences for PCR to detect the *CD4-cre* transgene were forward 5'CGATGCAACGAGTGATGAGG3', reverse 5'GCATTGCTGTCACTTGGTCGT3'.

The PCR conditions for CD4-cre transgene were 1 min at 94°C, 1 min at 61°C, and 1 min at 72°C for 35 cycles. *Foxa1* and *Foxa2* WT and floxed gene primers and PCR conditions were as described (25).

Cell sorting

To obtain DN3 (CD4⁻CD8⁻CD25⁺CD44⁻), DN4 (CD4⁻CD8⁻CD25⁻CD44⁻), DP (CD4⁺CD8⁺), CD4SP (CD4⁺CD8⁻CD3⁺), CD8SP (CD4⁻CD8⁺CD3⁺) and CD69⁺DP (CD4⁺CD8⁺CD69⁺) thymocyte cell suspensions were sorted using Molecular Flow Cytometer (Cytomation, Fort Collins, CO).

Quantitative RT-PCR

RNA was extracted using Stratagene Strataprep Total RNA miniprep kit according to manufacturer's protocol and cDNA was synthesised using SuperScript III kit (Invitrogen) according to manufacturer's protocol. RT-PCR was carried out by analysis of cDNA sample in triplicate on an iCycler (Bio-rad Laboratories) using the iQTMSYBR Green Supermix according to manufacturer's instructions. The house keeping gene *Gapdh* was used to allow quantification of template and normalisation of each gene. *Gapdh*, *Foxa1* and *Foxa2* RT-PCR primers were purchased from Qiagen (Quantitect primer assay).

Antibodies and flow cytometry

Cells were stained using combinations of the following directly conjugated antibodies supplied by Biolegend (San Diego, US) or eBioscience (San Diego, US) : anti-CD3, anti-CD4, anti-CD5, anti-CD8, anti-CD24, anti-CD44, anti-CD62L, anti-CD69, anti-Foxp3, anti-NK1.1, anti-Qa2 anti-Nr4a1 and anti-TCR β (Supplementary Table 5). Cell suspensions were stained for 30 minutes on ice in phosphate buffered saline supplemented with 2% FCS and 0.01% sodium azide. For intracellular Foxp3 and Nr4a1 staining, cells were fixed and permeabilised using the ICfixation/permeabilisation kit (eBioscience). AnnexinV staining was carried out as described (83). Samples were analysed on an AccuriTM C6 flow cytometer (BD, Oxford, UK) and LSRII (BD Biosciences). Data were analyzed using FlowJo 10.4.2 (Tree Star).

RNA-sequencing and data analysis

CD69⁺CD4⁺CD8⁺ thymocytes from *Foxa1/2cKO* and control thymus were FACS sorted. Each sample was sorted independently from different mice, with 2 mice of each genotype per sort. RNA was prepared as described(84) and sequenced by UCL Genomics on the Illumina Next Seq 500. The sequenced data are publically available in the GEO database under

accession number GSE169602. The Basespace Sequence Hub was used for both FASTQ generation and RNA-Seq Alignment to the mouse reference genome UCSC mm10 (RefSeq gene annotation). Aligned reads were counted by the HTSeq python package with “union” overlap resolution mode(85). The Bioconductor package DESeq2 (1.30.0) was used to test for differential expression(86). P-values were plotted in a histogram resulting in a hill-shape, indicating an overestimation of the variance in the null distribution (Supplementary Fig 1). Therefore, the z-scores returned by DESeq2 were used as input to the CRAN package fdrtool (1.2.16) to re-estimate the null variance and subsequently the p-values(87). Adjusted p-values were then calculated by Benjamini-Hochberg false discovery correction (5%). Genes with adjusted p-values less than 0.05 were considered as DEGs.

The Bioconductor package DEXSeq (1.36.0) was used to test for differential exon usage(88, 89). The Python script `dexseq_prepare_annotation.py` was used to prepare the genome annotation with the following parameter: “-r no”. Subsequently, `dexseq_count.py` was used to generate counts of exons using the BAM alignment files as input with the following parameters “-p yes -s no -a 0”. The exon count files were then inputted into DEXSeq. Adjusted p-values were then calculated by Benjamini-Hochberg false discovery correction (5%).

To identify over-represented gene ontology terms in the set of DEG and genes containing DEU, we used the Bioconductor package GOseq (1.42.0)(90). For data visualisation, we used regularized logarithm (rlog) transformed counts generated by DESeq2 as input for heatmaps, which were generated using the CRAN package pheatmap (1.0.12): rows were centred; unit variance scaling was applied to rows; and both rows and columns were clustered using Pearson correlation distance and average linkage, where red represents higher expression and blue lower expression on a linear correlation scale. A value of 1 indicates a positive association, while a value of -1 indicates a negative association, and a value of 0 indicates no association. Venn diagrams were generated using the CRAN package VennDiagram (1.6.20).

To identify verified transcription factors amongst DEG and DEU genes we merged the most recent updated lists from the Riken mouse transcription factor database (91) and the mouse transcription factor list from The Animal Transcription Factor DataBase (92) and intersected our gene lists with this merged list.

Canonical correspondence analysis is a multivariate analysis that allows the comparison of experimental transcriptome data with publically available datasets from other laboratories(93). CCA was performed using the CCA function of CRAN package *vegan*, as previously described (83). The GSE38909 dataset was used as the environmental variable and our dataset was regressed onto it. The GSE38909 dataset contains TCR-transgenic AND DP thymocytes stimulated with a positively selecting peptide (gp250) or a non-selecting control peptide Hb (49). To represent environmental variables of interest, the 2000 most significant DEG (lowest p-values, calculated by moderated eBayes adjusted for false positives) between DP thymocytes stimulated with a non-selecting control peptide and DP thymocytes stimulated with the positively selecting peptide were used, to generate a scale of unstimulated to TCR-signalling-for-positive-selection, and we regressed our datasets onto this axis.

Acknowledgements

We thank Ayad Eddaoudi for cell sorting. This work was funded by the MRC (MR/P000843/1; MR/5037764/1) and Great Ormond Street Hospital Children's Charity (GOSHCC). JR was supported by a studentship from the BBSRC London Interdisciplinary Biosciences Consortium (LiDO) and AS by a studentship from GOSHCC and the MRC. Research at UCL Great Ormond Street Institute of Child Health is supported by the NIHR Biomedical Research Centre at Great Ormond Street Hospital and UCL.

Authors' Contributions

C-IL and TC conceived the study and designed experiments. C-IL, JR and TC wrote the manuscript; JR, AS and MO analysed RNA-sequencing datasets. C-IL, DY and SR performed and analysed experiments. All authors critically reviewed the manuscript. The authors declare no competing financial interests.

References

1. Huynh A, Zhang R, Turka LA. Signals and pathways controlling regulatory T cells. *Immunol Rev.* 2014;258(1):117-31.
2. Littman DR. How Thymocytes Achieve Their Fate. *J Immunol.* 2016;196(5):1983-4.
3. Starr TK, Jameson SC, Hogquist KA. Positive and negative selection of T cells. *Annu Rev Immunol.* 2003;21:139-76.
4. Singer A, Adoro S, Park JH. Lineage fate and intense debate: myths, models and mechanisms of CD4- versus CD8-lineage choice. *Nat Rev Immunol.* 2008;8(10):788-801.
5. Klein L, Kyewski B, Allen PM, Hogquist KA. Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). *Nat Rev Immunol.* 2014;14(6):377-91.
6. Bosselut R. CD4/CD8-lineage differentiation in the thymus: from nuclear effectors to membrane signals. *Nat Rev Immunol.* 2004;4(7):529-40.
7. Carpenter AC, Bosselut R. Decision checkpoints in the thymus. *Nat Immunol.* 2010;11(8):666-73.
8. Egawa T. Regulation of CD4 and CD8 coreceptor expression and CD4 versus CD8 lineage decisions. *Adv Immunol.* 2015;125:1-40.
9. Park JH, Adoro S, Guinter T, Erman B, Alag AS, Catalfamo M, et al. Signaling by intrathymic cytokines, not T cell antigen receptors, specifies CD8 lineage choice and promotes the differentiation of cytotoxic-lineage T cells. *Nat Immunol.* 2010;11(3):257-64.
10. Brugnera E, Bhandoola A, Cibotti R, Yu Q, Guinter TI, Yamashita Y, et al. Coreceptor reversal in the thymus: signaled CD4+8+ thymocytes initially terminate CD8 transcription even when differentiating into CD8+ T cells. *Immunity.* 2000;13(1):59-71.
11. Laky K, Fowlkes BJ. Notch signaling in CD4 and CD8 T cell development. *Curr Opin Immunol.* 2008;20(2):197-202.
12. Solanki A, Yanez DC, Ross S, Lau CI, Papaioannou E, Li JW, et al. Gli3 in fetal thymic epithelial cells promotes thymocyte positive selection and differentiation by repression of Shh. *Development.* 2018;145(3).
13. Furmanski AL, Saldana JI, Rowbotham NJ, Ross SE, Crompton T. Role of Hedgehog signalling at the transition from double-positive to single-positive thymocyte. *European Journal of Immunology.* 2012;42(2):489-99.
14. Rowbotham NJ, Hager-Theodorides AL, Cebecauer M, Shah DK, Drakopoulou E, Dyson J, et al. Activation of the Hedgehog signaling pathway in T-lineage cells inhibits TCR repertoire selection in the thymus and peripheral T-cell activation. *Blood.* 2007;109(9):3757-66.

15. Taniuchi I. Views on helper/cytotoxic lineage choice from a bottom-up approach. *Immunol Rev.* 2016;271(1):98-113.
16. Naito T, Tanaka H, Naoe Y, Taniuchi I. Transcriptional control of T-cell development. *Int Immunol.* 2011;23(11):661-8.
17. Issuree PD, Ng CP, Littman DR. Heritable Gene Regulation in the CD4:CD8 T Cell Lineage Choice. *Front Immunol.* 2017;8:291.
18. Jackson BC, Carpenter C, Nebert DW, Vasiliou V. Update of human and mouse forkhead box (FOX) gene families. *Hum Genomics.* 2010;4(5):345-52.
19. Iwafuchi-Doi M, Donahue G, Kakumanu A, Watts JA, Mahony S, Pugh BF, et al. The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation. *Mol Cell.* 2016;62(1):79-91.
20. Zaret KS. Pioneer Transcription Factors Initiating Gene Network Changes. *Annu Rev Genet.* 2020;54:367-85.
21. Reizel Y, Morgan A, Gao L, Schug J, Mukherjee S, Garcia MF, et al. FoxA-dependent demethylation of DNA initiates epigenetic memory of cellular identity. *Dev Cell.* 2021;56(5):602-12 e4.
22. Kaestner KH. The FoxA factors in organogenesis and differentiation. *Curr Opin Genet Dev.* 2010;20(5):527-32.
23. Besnard V, Wert SE, Hull WM, Whitsett JA. Immunohistochemical localization of Foxa1 and Foxa2 in mouse embryos and adult tissues. *Gene Expr Patterns.* 2004;5(2):193-208.
24. Kaestner KH, Hiemisch H, Luckow B, Schutz G. The HNF-3 gene family of transcription factors in mice: gene structure, cDNA sequence, and mRNA distribution. *Genomics.* 1994;20(3):377-85.
25. Lau CI, Yanez DC, Solanki A, Papaioannou E, Saldana JI, Crompton T. Foxa1 and Foxa2 in thymic epithelial cells (TEC) regulate medullary TEC and regulatory T-cell maturation. *J Autoimmun.* 2018;93:131-8.
26. Rowbotham NJ, Hager-Theodorides AL, Furmanski AL, Ross SE, Outram SV, Dessens JT, et al. Sonic hedgehog negatively regulates pre-TCR-induced differentiation by a Gli2-dependent mechanism. *Blood.* 2009;113(21):5144-56.
27. Weinstein DC, Ruiz i Altaba A, Chen WS, Hoodless P, Prezioso VR, Jessell TM, et al. The winged-helix transcription factor HNF-3 beta is required for notochord development in the mouse embryo. *Cell.* 1994;78(4):575-88.
28. Shih DQ, Navas MA, Kuwajima S, Duncan SA, Stoffel M. Impaired glucose homeostasis and neonatal mortality in hepatocyte nuclear factor 3alpha-deficient mice. *Proc Natl Acad Sci U S A.* 1999;96(18):10152-7.
29. Gao N, LeLay J, Vatamaniuk MZ, Rieck S, Friedman JR, Kaestner KH. Dynamic regulation of Pdx1 enhancers by Foxa1 and Foxa2 is essential for pancreas development. *Genes Dev.* 2008;22(24):3435-48.
30. Reizel Y, Morgan A, Gao L, Lan Y, Manduchi E, Waite EL, et al. Collapse of the hepatic gene regulatory network in the absence of FoxA factors. *Genes Dev.* 2020;34(15-16):1039-50.
31. Lee CS, Friedman JR, Fulmer JT, Kaestner KH. The initiation of liver development is dependent on Foxa transcription factors. *Nature.* 2005;435(7044):944-7.
32. Wan H, Dingle S, Xu Y, Besnard V, Kaestner KH, Ang SL, et al. Compensatory roles of Foxa1 and Foxa2 during lung morphogenesis. *J Biol Chem.* 2005;280(14):13809-16.

33. Sasaki H, Hui C, Nakafuku M, Kondoh H. A binding site for Gli proteins is essential for HNF-3beta floor plate enhancer activity in transgenics and can respond to Shh in vitro. *Development*. 1997;124(7):1313-22.
34. Metzakopian E, Lin W, Salmon-Divon M, Dvinge H, Andersson E, Ericson J, et al. Genome-wide characterization of Foxa2 targets reveals upregulation of floor plate genes and repression of ventrolateral genes in midbrain dopaminergic progenitors. *Development*. 2012;139(14):2625-34.
35. Jeong Y, Epstein DJ. Distinct regulators of Shh transcription in the floor plate and notochord indicate separate origins for these tissues in the mouse node. *Development*. 2003;130(16):3891-902.
36. Gao N, Ishii K, Mirosevich J, Kuwajima S, Oppenheimer SR, Roberts RL, et al. Forkhead box A1 regulates prostate ductal morphogenesis and promotes epithelial cell maturation. *Development*. 2005;132(15):3431-43.
37. Liu Y, Carlsson R, Comabella M, Wang J, Kosicki M, Carrion B, et al. FoxA1 directs the lineage and immunosuppressive properties of a novel regulatory T cell population in EAE and MS. *Nat Med*. 2014;20(3):272-82.
38. Moran AE, Holzapfel KL, Xing Y, Cunningham NR, Maltzman JS, Punt J, et al. T cell receptor signal strength in Treg and iNKT cell development demonstrated by a novel fluorescent reporter mouse. *J Exp Med*. 2011;208(6):1279-89.
39. Bending D, Prieto Martin P, Paduraru A, Ducker C, Marzaganov E, Laviron M, et al. A timer for analyzing temporally dynamic changes in transcription during differentiation in vivo. *J Cell Biol*. 2018;217(8):2931-50.
40. Azzam HS, DeJarnette JB, Huang K, Emmons R, Park CS, Sommers CL, et al. Fine tuning of TCR signaling by CD5. *J Immunol*. 2001;166(9):5464-72.
41. Ross JO, Melichar HJ, Au-Yeung BB, Herzmark P, Weiss A, Robey EA. Distinct phases in the positive selection of CD8+ T cells distinguished by intrathymic migration and T-cell receptor signaling patterns. *Proc Natl Acad Sci U S A*. 2014;111(25):E2550-8.
42. Kisielow P, Miazek A. Positive selection of T cells: rescue from programmed cell death and differentiation require continual engagement of the T cell receptor. *J Exp Med*. 1995;181(6):1975-84.
43. Liu X, Bosselut R. Duration of TCR signaling controls CD4-CD8 lineage differentiation in vivo. *Nat Immunol*. 2004;5(3):280-8.
44. Metzakopian E, Bouhali K, Alvarez-Saavedra M, Whitsett JA, Picketts DJ, Ang SL. Genome-wide characterisation of Foxa1 binding sites reveals several mechanisms for regulating neuronal differentiation in midbrain dopamine cells. *Development*. 2015;142(7):1315-24.
45. Kasler HG, Young BD, Mottet D, Lim HW, Collins AM, Olson EN, et al. Histone deacetylase 7 regulates cell survival and TCR signaling in CD4/CD8 double-positive thymocytes. *J Immunol*. 2011;186(8):4782-93.
46. Huang F, Gu H. Negative regulation of lymphocyte development and function by the Cbl family of proteins. *Immunol Rev*. 2008;224:229-38.
47. Gascoigne NR, Acuto O. THEMIS: a critical TCR signal regulator for ligand discrimination. *Curr Opin Immunol*. 2015;33:86-92.
48. Sepulveda FE, Burgess A, Heiligenstein X, Goudin N, Menager MM, Romao M, et al. LYST controls the biogenesis of the endosomal compartment required for secretory lysosome function. *Traffic*. 2015;16(2):191-203.

49. Lo WL, Donermeyer DL, Allen PM. A voltage-gated sodium channel is essential for the positive selection of CD4(+) T cells. *Nat Immunol.* 2012;13(9):880-7.
50. Jiang X, Kim HE, Shu H, Zhao Y, Zhang H, Kofron J, et al. Distinctive roles of PHAP proteins and prothymosin-alpha in a death regulatory pathway. *Science.* 2003;299(5604):223-6.
51. Suliman S, Tan J, Xu K, Kousis PC, Kowalski PE, Chang G, et al. Notch3 is dispensable for thymocyte beta-selection and Notch1-induced T cell leukemogenesis. *PLoS One.* 2011;6(9):e24937.
52. Georgopoulos K. The making of a lymphocyte: the choice among disparate cell fates and the IKAROS enigma. *Genes Dev.* 2017;31(5):439-50.
53. Costello PS, Nicolas RH, Watanabe Y, Rosewell I, Treisman R. Ternary complex factor SAP-1 is required for Erk-mediated thymocyte positive selection. *Nat Immunol.* 2004;5(3):289-98.
54. Wang L, Xiong Y, Bosselut R. Tenuous paths in unexplored territory: From T cell receptor signaling to effector gene expression during thymocyte selection. *Semin Immunol.* 2010;22(5):294-302.
55. Kessenbrock K, Plaks V, Werb Z. Matrix metalloproteinases: regulators of the tumor microenvironment. *Cell.* 2010;141(1):52-67.
56. Kalashnikova AA, Winkler DD, McBryant SJ, Henderson RK, Herman JA, DeLuca JG, et al. Linker histone H1.0 interacts with an extensive network of proteins found in the nucleolus. *Nucleic Acids Res.* 2013;41(7):4026-35.
57. Di Liegro CM, Schiera G, Di Liegro I. H1.0 Linker Histone as an Epigenetic Regulator of Cell Proliferation and Differentiation. *Genes (Basel).* 2018;9(6): 310(6: 310).
58. Ho TH, Charlet BN, Poulos MG, Singh G, Swanson MS, Cooper TA. Muscleblind proteins regulate alternative splicing. *EMBO J.* 2004;23(15):3103-12.
59. Wang ET, Cody NA, Jog S, Biancolella M, Wang TT, Treacy DJ, et al. Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell.* 2012;150(4):710-24.
60. Sznajder LJ, Scotti MM, Shin J, Taylor K, Ivankovic F, Nutter CA, et al. Loss of MBNL1 induces RNA misprocessing in the thymus and peripheral blood. *Nat Commun.* 2020;11(1):2022.
61. Itskovich SS, Gurunathan A, Clark J, Burwinkel M, Wunderlich M, Berger MR, et al. MBNL1 regulates essential alternative RNA splicing patterns in MLL-rearranged leukemia. *Nat Commun.* 2020;11(1):2369.
62. Saitoh N, Sakamoto C, Hagiwara M, Agredano-Moreno LT, Jimenez-Garcia LF, Nakao M. The distribution of phosphorylated SR proteins and alternative splicing are regulated by RANBP2. *Mol Biol Cell.* 2012;23(6):1115-28.
63. Shiozawa Y, Malcovati L, Galli A, Sato-Otsubo A, Kataoka K, Sato Y, et al. Aberrant splicing and defective mRNA production induced by somatic spliceosome mutations in myelodysplasia. *Nat Commun.* 2018;9(1):3649.
64. Mupo A, Seiler M, Sathiaseelan V, Pance A, Yang Y, Agrawal AA, et al. Hemopoietic-specific Sf3b1-K700E knock-in mice display the splicing defect seen in human MDS but develop anemia without ring sideroblasts. *Leukemia.* 2017;31(3):720-7.
65. McIlwain DR, Pan Q, Reilly PT, Elia AJ, McCracken S, Wakeham AC, et al. Smg1 is required for embryogenesis and regulates diverse genes via alternative splicing coupled to nonsense-mediated mRNA decay. *Proc Natl Acad Sci U S A.* 2010;107(27):12186-91.

66. Du Y, Lu F, Li P, Ye J, Ji M, Ma D, et al. SMG1 acts as a novel potential tumor suppressor with epigenetic inactivation in acute myeloid leukemia. *Int J Mol Sci*. 2014;15(9):17065-76.
67. Mallory MJ, Allon SJ, Qiu J, Gazzara MR, Tapescu I, Martinez NM, et al. Induced transcription and stability of CELF2 mRNA drives widespread alternative splicing during T-cell signaling. *Proc Natl Acad Sci U S A*. 2015;112(17):E2139-48.
68. Urban JA, Winandy S. Ikaros null mice display defects in T cell selection and CD4 versus CD8 lineage decisions. *J Immunol*. 2004;173(7):4470-8.
69. Mee PJ, Turner M, Basson MA, Costello PS, Zamoyska R, Tybulewicz VL. Greatly reduced efficiency of both positive and negative selection of thymocytes in CD45 tyrosine phosphatase-deficient mice. *Eur J Immunol*. 1999;29(9):2923-33.
70. Watanabe M, Lu Y, Breen M, Hodes RJ. B7-CD28 co-stimulation modulates central tolerance via thymic clonal deletion and Treg generation through distinct mechanisms. *Nat Commun*. 2020;11(1):6264.
71. Outram SV, Gordon AR, Hager-Theodorides AL, Metcalfe J, Crompton T, Kemp P. KLF13 influences multiple stages of both B and T cell development. *Cell Cycle*. 2008;7(13):2047-55.
72. Martinez NM, Lynch KW. Control of alternative splicing in immune responses: many regulators, many predictions, much still to learn. *Immunol Rev*. 2013;253(1):216-36.
73. Cho V, Mei Y, Sanny A, Chan S, Enders A, Bertram EM, et al. The RNA-binding protein hnRNPLL induces a T cell alternative splicing program delineated by differential intron retention in polyadenylated RNA. *Genome Biol*. 2014;15(1):R26.
74. Mitchell JL, Seng A, Yankee TM. Expression and splicing of Ikaros family members in murine and human thymocytes. *Mol Immunol*. 2017;87:1-11.
75. Baralle FE, Giudice J. Alternative splicing as a regulator of development and tissue identity. *Nat Rev Mol Cell Biol*. 2017;18(7):437-51.
76. Wu Z, Yates AL, Hoyne GF, Goodnow CC. Consequences of increased CD45RA and RC isoforms for TCR signaling and peripheral T cell deficiency resulting from heterogeneous nuclear ribonucleoprotein L-like mutation. *J Immunol*. 2010;185(1):231-8.
77. Tellier M, Maudlin I, Murphy S. Transcription and splicing: A two-way street. *Wiley Interdiscip Rev RNA*. 2020;11(5):e1593.
78. Sekiya T, Zaret KS. Repression by Groucho/TLE/Grg proteins: genomic site recruitment generates compacted chromatin in vitro and impairs activator binding in vivo. *Mol Cell*. 2007;28(2):291-303.
79. Wang JC, Waltner-Law M, Yamada K, Osawa H, Stifani S, Granner DK. Transducin-like enhancer of split proteins, the human homologs of *Drosophila* groucho, interact with hepatic nuclear factor 3beta. *J Biol Chem*. 2000;275(24):18418-23.
80. Xing S, Shao P, Li F, Zhao X, Seo W, Wheat JC, et al. Tle corepressors are differentially partitioned to instruct CD8(+) T cell lineage choice and identity. *J Exp Med*. 2018;215(8):2211-26.
81. Ferri AL, Lin W, Mavromatakis YE, Wang JC, Sasaki H, Whitsett JA, et al. Foxa1 and Foxa2 regulate multiple phases of midbrain dopaminergic neuron development in a dosage-dependent manner. *Development*. 2007;134(15):2761-9.
82. Lau CI, Outram SV, Saldana JI, Furmanski AL, Dessens JT, Crompton T. Regulation of murine normal and stress-induced erythropoiesis by Desert Hedgehog. *Blood*. 2012;119(20):4741-51.

83. Solanki A, Yanez DC, Lau CI, Rowell J, Barbarulo A, Ross S, et al. The transcriptional repressor Bcl6 promotes pre-TCR-induced thymocyte differentiation and attenuates Notch1 activation. *Development*. 2020;147(19).
84. Solanki A, Lau CI, Saldana JI, Ross S, Crompton T. The transcription factor Gli3 promotes B cell development in fetal liver through repression of Shh. *J Exp Med*. 2017;214(7):2041-58.
85. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166-9.
86. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
87. Strimmer K. fdrtool: a versatile R package for estimating local and tail area-based false discovery rates. *Bioinformatics*. 2008;24(12):1461-2.
88. Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA-seq data. *Genome Res*. 2012;22(10):2008-17.
89. Reyes A, Anders S, Weatheritt RJ, Gibson TJ, Steinmetz LM, Huber W. Drift and conservation of differential exon usage across tissues in primate species. *Proc Natl Acad Sci U S A*. 2013;110(38):15377-82.
90. Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol*. 2010;11(2):R14.
91. Kanamori M, Konno H, Osato N, Kawai J, Hayashizaki Y, Suzuki H. A genome-wide and nonredundant mouse transcription factor database. *Biochem Biophys Res Commun*. 2004;322(3):787-93.
92. Hu H, Miao YR, Jia LH, Yu QY, Zhang Q, Guo AY. AnimalTFDB 3.0: a comprehensive resource for annotation and prediction of animal transcription factors. *Nucleic Acids Res*. 2019;47(D1):D33-D8.
93. Ono M, Tanaka RJ, Kano M. Visualisation of the T cell differentiation programme by Canonical Correspondence Analysis of transcriptomes. *BMC Genomics*. 2014;15:1028.
94. Westendorf K, Durek P, Ayew S, Mashreghi MF, Radbruch A. Chromosomal localisation of the CD4cre transgene in B6.Cg-Tg(Cd4-cre)1Cwi mice. *J Immunol Methods*. 2016;436:54-7.

Figures

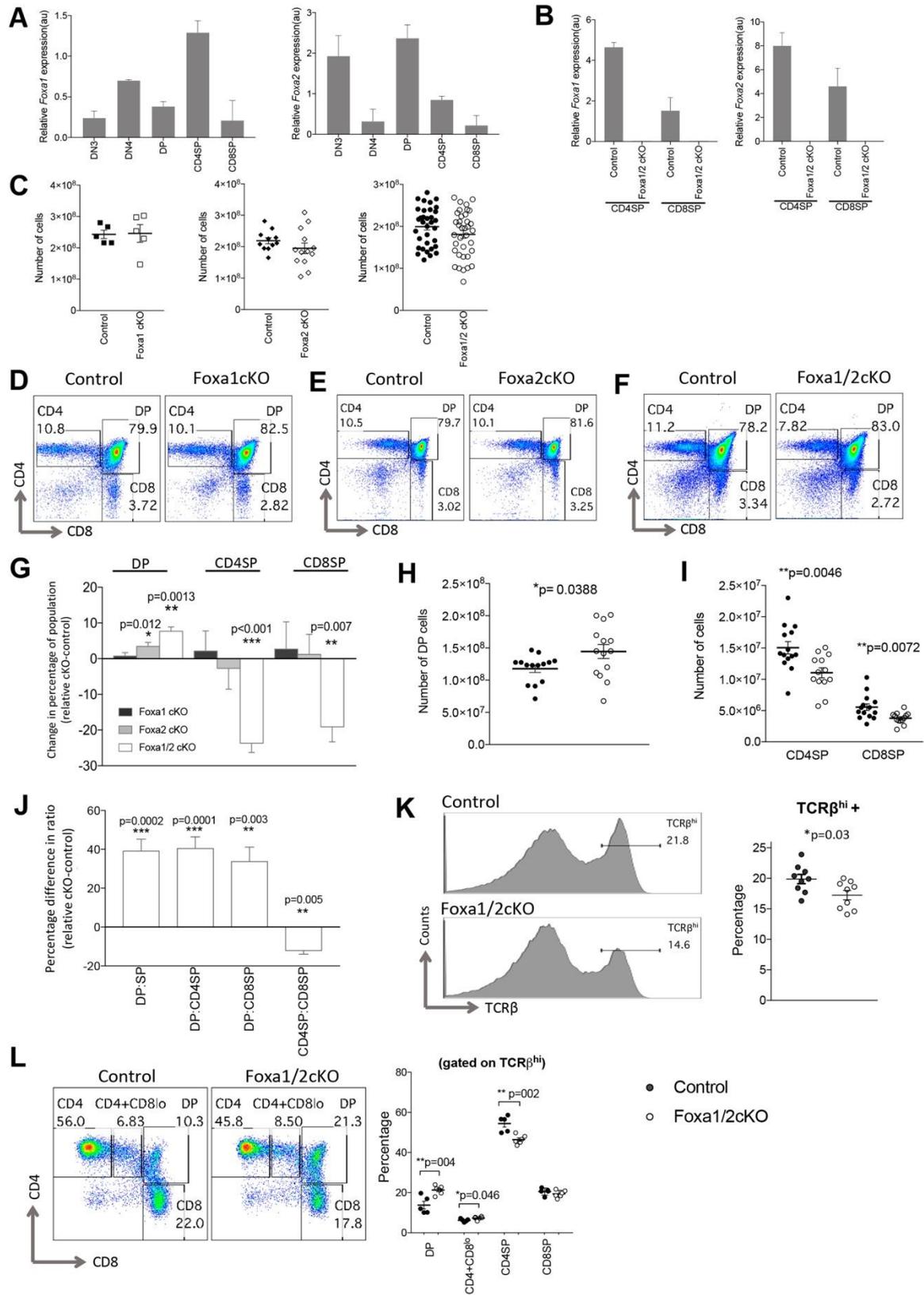


Figure 1. Foxa1 and Foxa2 are expressed in thymocytes and required in T-cell development

(A) *Foxa1* and *Foxa2* expression were measured by QRT-PCR in FACS-sorted thymocyte populations from wild type (C57BL/6) adult mice: DN3 (CD4⁻CD8⁻CD25⁺CD44⁻); DN4 (CD4⁻CD8⁻CD25⁻CD44⁻); DP (CD4⁺CD8⁺); CD4SP (CD4⁺CD8⁻CD3⁺); CD8SP (CD4⁻CD8⁺CD3⁺). Bar charts show relative expression of *Foxa1* (left panel) and *Foxa2* (right panel) in these FACS-sorted populations.

(B) Bar charts show relative expression of *Foxa1* (left panel) and *Foxa2* (right panel) measured by QRT-PCR in FACS-sorted CD4SP (CD4⁺CD8⁻CD3⁺) and CD8SP (CD4⁻CD8⁺CD3⁺) cells from *Foxa1/2cKO* and control thymus.

(C) Scatter plots show number of cells recovered from thymus of *Foxa1cKO*, *Foxa2cKO* and *Foxa1/2cKO* mice compared to control littermates.

(D-F) Flow cytometry profiles show anti-CD4 and anti-CD8 staining, giving percentage cells in regions shown for *Foxa1cKO* **(D)**, *Foxa2cKO* **(E)** and *Foxa1/2cKO* **(F)** thymus compared with control littermate thymus.

(G) Bar chart shows change in percentage of thymocyte population in cKO relative to control littermates (mean percentage of each population in control littermates subtracted from relative percentage of same population in conditional knockout) in *Foxa1cKO*, *Foxa2cKO* and *Foxa1/2cKO* thymus. For control vs *Foxa1cKO*: control n=5, *Foxa1cKO* n=5; for control vs *Foxa2cKO*: control n=9, *Foxa1cKO* n=9; for control vs *Foxa1/2cKO*: control n=15, *Foxa1/2cKO* n=16.

(H-I) Scatter plots show number of cells in DP **(H)** and CD4SP and CD8SP **(I)** populations in *Foxa1/2cKO* mice compared to control.

(J) Bar chart shows change in thymocyte subset ratio in *Foxa1/2cKO* relative to control littermates, calculated by mean ratio of control littermates subtracted from relative ratio from *Foxa1/2cKO*; control n=15, *Foxa1/2cKO* n=16.

(K) Histograms show staining of anti-TCR β on thymocytes, giving the percentage of cells in the marker shown (TCR β ^{hi}). Scatter plot shows the percentage of TCR β ^{hi} cells in control and *Foxa1/2cKO* littermates.

(L) Flow cytometry profiles show anti-CD4 and anti-CD8 staining of thymus, gated on the TCR β^{hi} population, giving the percentage of cells in the regions shown. Dot plot shows the percentage of DP, CD4 $^+$ CD8 $^{\text{lo}}$, CD4SP and CD8SP, gated on TCR β^{hi} in Foxa1/2cKO and control thymus.

In scatter plots each symbol represents an individual mouse, control (dark circles) and Foxa1/2cKO (open circles). Bar charts and scatter plots show mean and SEM, giving significance by Student's t test: * $p < 0.05$; ** $p < 0.01$, *** $p < 0.001$.

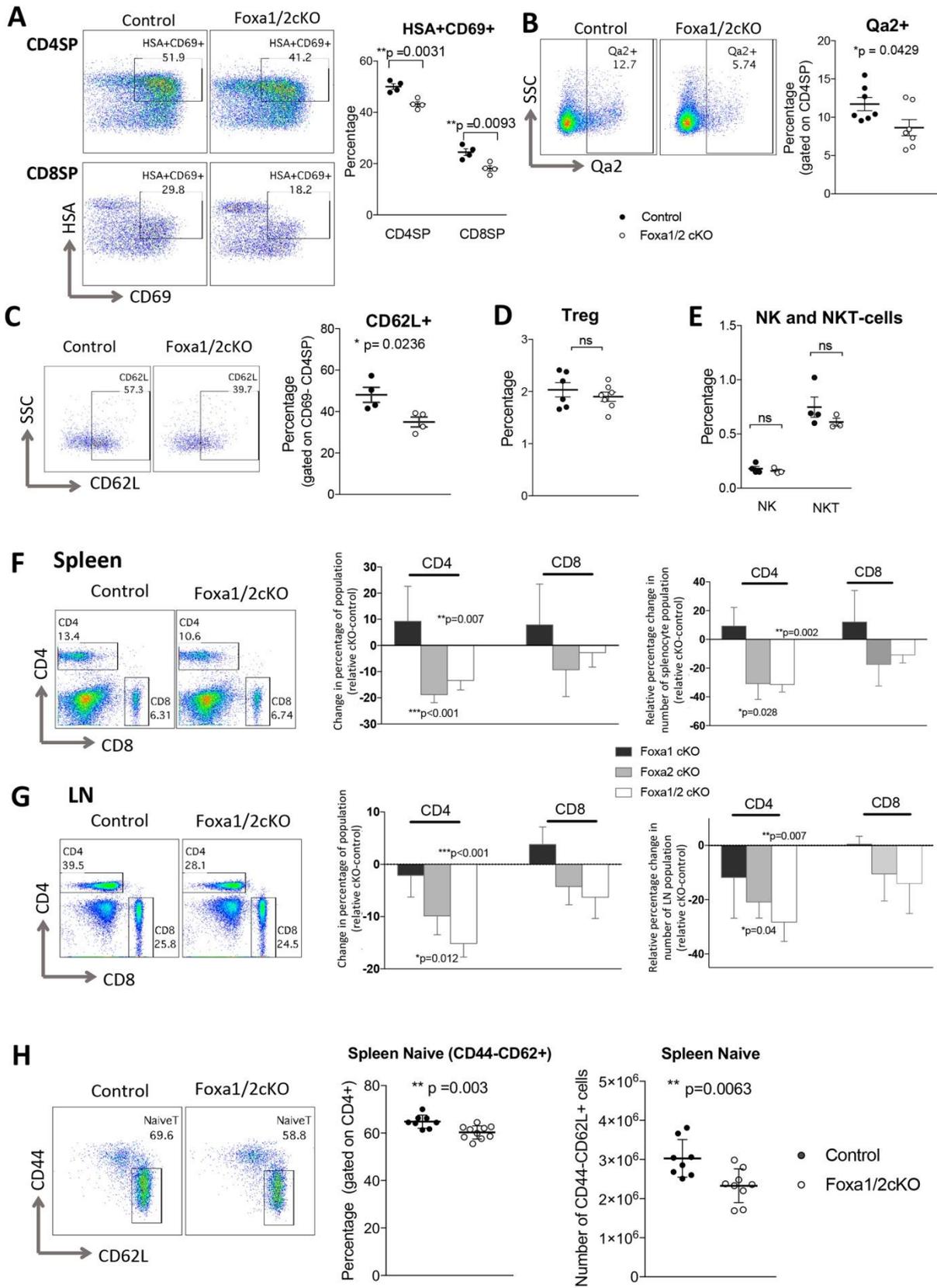


Figure 2. Maturation of SP and peripheral T-cells in Foxa1/2 cKO mice

(A) Flow cytometry profiles show anti-HSA and anti-CD69 staining on CD4SP and CD8SP cells, giving the percentage of cells in the region shown. Scatter plot shows the percentage of HSA⁺CD69⁺ cells in the CD4SP and CD8SP populations for control and Foxa1/2cKO.

(B) Flow cytometry profiles show anti-Qa2 staining, gated on CD4SP, giving the percentage of cells in the region shown. Scatter plot shows percentage of Qa2⁺ cells in the CD4SP population.

(C) Flow cytometry profiles show anti-CD62L staining, gated on the CD69⁻CD4SP population. Scatter plot shows the percentage of CD62L⁺ cells in the CD69⁻CD4SP population.

(D) Percentage of Tregs (CD8⁻CD4⁺CD25⁺iCFoxp3⁺) in the control and Foxa1/2cKO thymus.

(E) Percentage of NK (CD3⁻NK1.1⁺) and NKT-cells (CD3⁺NK1.1⁺) in the control and Foxa1/2cKO thymus.

(F-G) Flow cytometry profiles show anti-CD4 and anti-CD8 staining of spleen **(F)** and lymph node (LN) **(G)** in Foxa1/2cKO and control mice. Bar charts represent percentage change in T-cell composition and actual cell numbers in conditional knockout mice relative to control littermates in spleen **(F)** and LN **(G)**. Differences in percentage were calculated by subtraction of mean percentage or number of population in control littermates from the relative percentage or number of the population in cKO.

(H) Flow cytometry profiles show anti-CD44 and anti-CD62L staining, gated on CD4⁺ T-cells in Foxa1/2cKO and control spleen, giving the percentage of cells in the region shown. Scatter plots show percentage of CD44⁻CD62L⁺ cells (gated on CD4⁺) and number of CD44⁻CD62L⁺CD4⁺ T-cells (naive T-cells) in Foxa1/2cKO and control spleen.

In scatter plots each symbol represents an individual mouse, control (dark circles) and Foxa1/2cKO (open circles). Scatter plots and bar charts show mean and SEM, giving significance by Student's t test: * p<0.05; **p<0.01, ***p<0.001.

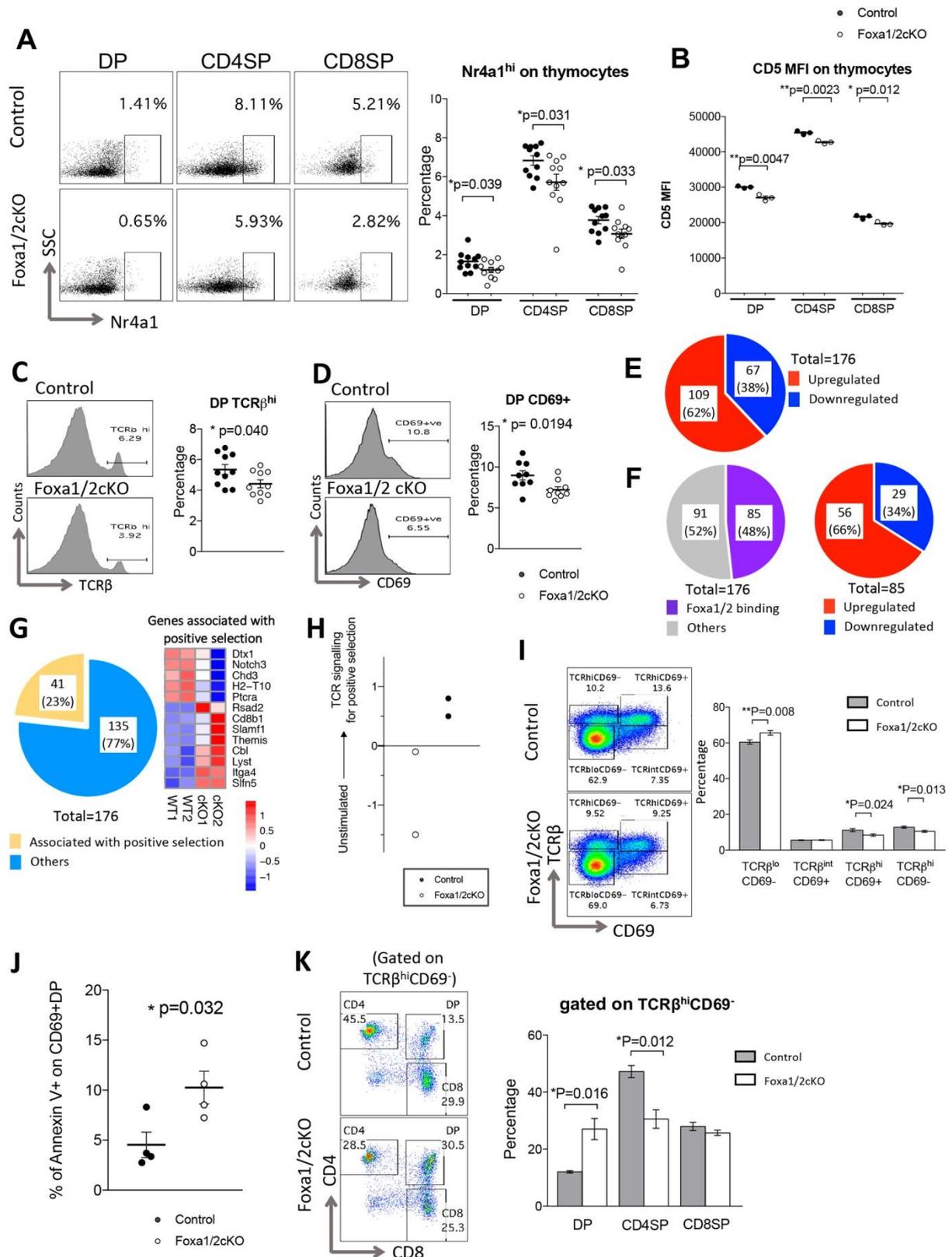


Figure 3. Foxa1/2 regulates TCR signal and promotes positive T-cell selection

(A) Flow cytometry profiles show intracellular anti-Nr4a1 staining, gated on DP, CD4SP and CD8SP in control and Foxa1/2cKO thymus. Scatter plot shows percentage of Nr4a1^{hi} in DP, CD4SP and CD8SP populations.

(B) Scatter plot shows MFI of anti-CD5 staining on DP, CD4SP and CD8SP populations.

(C-D) Histograms show staining of anti-TCR β **(C)** and anti-CD69**(D)** on the DP population in control and Foxa1/2cKO thymus. Scatter plots show percentage of TCR β ^{hi}**(C)** and CD69⁺**(D)** on DP populations.

(E) Pie chart represents the proportion of DEG that were upregulated (red) or downregulated (blue) in Foxa1/2cKO datasets compared to control.

(F) Left: Pie chart illustrates number of DEG that were identified as binding Foxa1/2 in genome-wide ChIPSeq analysis of Foxa1/2 binding sites in neuronal progenitors(34, 44) (purple) and DEG that were not (grey). Right: Pie chart represents those 85 DEG that were found to be Foxa1/2-binding, as defined in (F), that were upregulated (red) or down-regulated (blue) in our Foxa1/2cKO datasets.

(G) Pie chart represents number of DEG that are associated with positive selection(45) (yellow) and that are not associated with positive selection (blue). Pearson correlation clustering heatmap shows expression of selected positive selection associated DEG in control and Foxa1/2cKO, where red represents higher expression and blue lower expression on a linear correlation scale. A value of 1 indicates a positive association, while a value of -1 indicates a negative association, and a value of 0 indicates no association.

(H) Dot plot shows canonical correspondence analysis on a scale of unstimulated to TCR-signalling-for positive-selection(49).

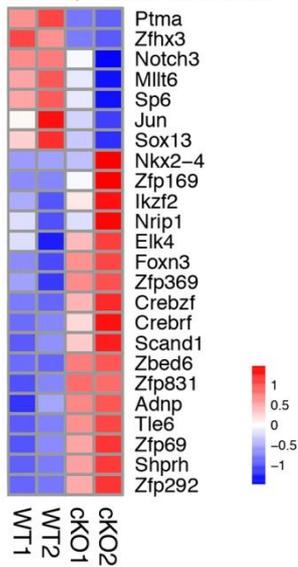
(I) Flow cytometry profiles show anti-TCR β and anti-CD69 staining, giving the percentage in regions shown for TCR β ^{lo}CD69⁻ (pre-selection), TCR β ^{int}CD69⁺ (selecting), TCR β ^{hi}CD69⁺ (post initiation of positive selection) and TCR β ^{hi}CD69⁻ (mature). Histogram shows mean percentage of cells in these four subsets, defined by staining against TCR β and CD69.

(J) Scatter plot shows percentage of Annexin V⁺ cells in the CD69⁺DP population.

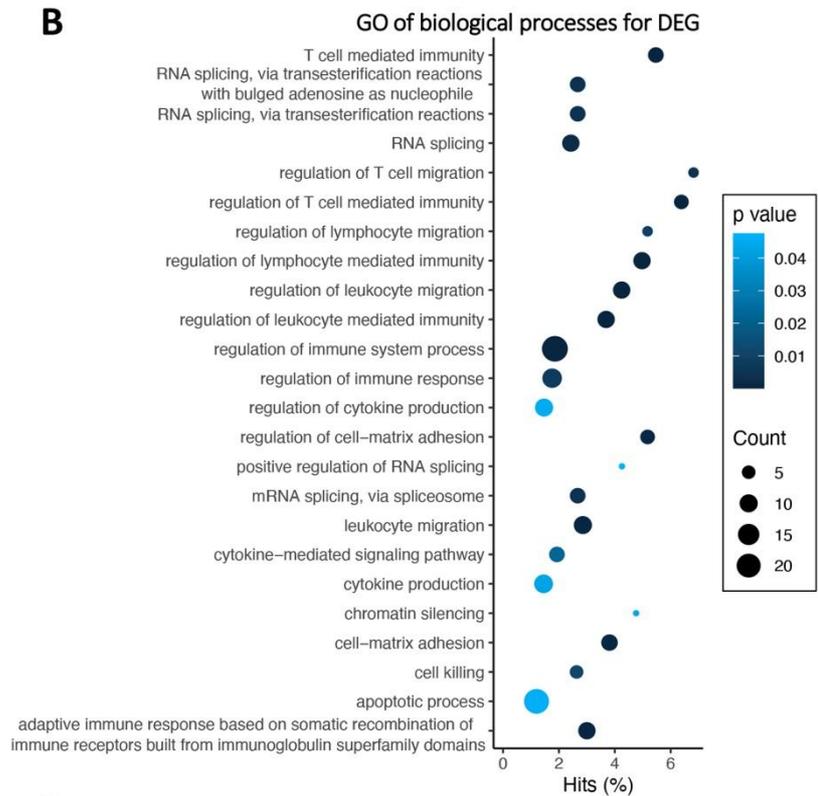
(K) Flow cytometry profiles show anti-CD4 and anti-CD8 staining gated on the TCR β^{hi} CD69 $^{-}$ population. Histogram shows the mean percentage of each thymocyte subpopulation, gated on TCR β^{hi} CD69 $^{-}$.

In scatter plots each symbol represents an individual mouse, control (dark circles) and Foxa1/2cKO (open circles). Scatter plots and bar charts show mean and SEM, giving significance by Student's t test: * $p < 0.05$; ** $p < 0.01$.

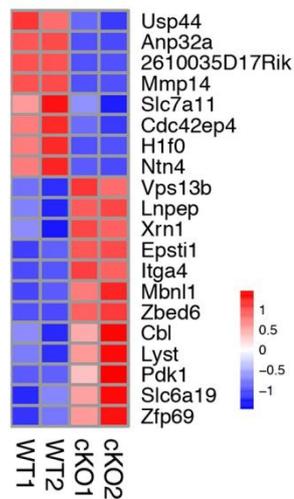
A Transcription factors in DEG



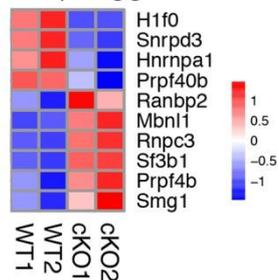
B



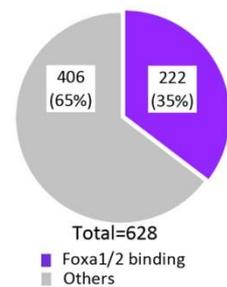
C Most significant DEG



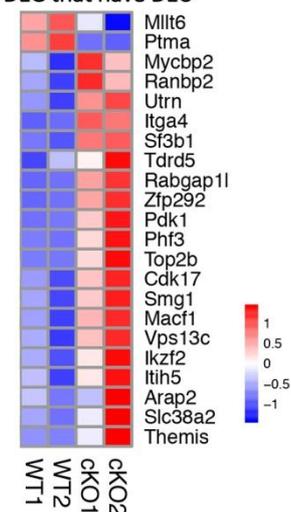
D RNA splicing genes in DEG



E



G DEG that have DEU



F

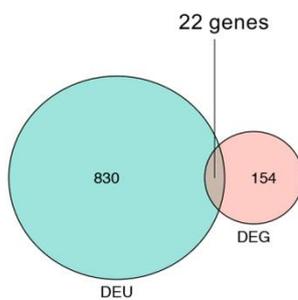


Figure 4. Foxa1/2 deficiency disturbs exon usage in developing T-cells

(A) Pearson correlation clustering heatmap shows expression of DEG that encode transcription factors in control and Foxa1/2cKO, where red represents higher expression and blue lower expression on a linear correlation scale. A value of 1 indicates a positive association, while a value of -1 indicates a negative association, and a value of 0 indicates no association.

(B) Dot plot of Gene Ontology (GO) analysis of biological processes associated with 176 DEG. The y-axis represents the biological process GO terms and the x-axis represents the percentage of genes found in that GO term category. Dot size represents the number of genes and the color indicates the p-value.

(C) Pearson correlation clustering heatmap shows expression of 20 most significant DEG in control and Foxa1/2cKO, where red represents higher expression and blue lower expression on a linear correlation scale. A value of 1 indicates a positive association, while a value of -1 indicates a negative association, and a value of 0 indicates no association.

(D) Pearson correlation clustering heatmap shows expression of DEG associated with RNA splicing in control and Foxa1/2cKO, where red represents higher expression and blue lower expression on a linear correlation scale. A value of 1 indicates a positive association, while a value of -1 indicates a negative association, and a value of 0 indicates no association. **(E)** Pie chart illustrates number of genes showing DEU: purple section represents number of genes that are identified as Foxa1/2 targets in genome-wide ChIPSeq analysis of Foxa1/2 binding sites in neuronal progenitors(34, 44).

(F) Proportional Venn diagram illustrates the intersection between DEG and genes showing DEU.

(G) Pearson correlation clustering heatmap shows expression of the 22 DEG which show DEU, where red represents higher expression and blue lower expression on a linear correlation scale. A value of 1 indicates a positive association, while a value of -1 indicates a negative association, and a value of 0 indicates no association.

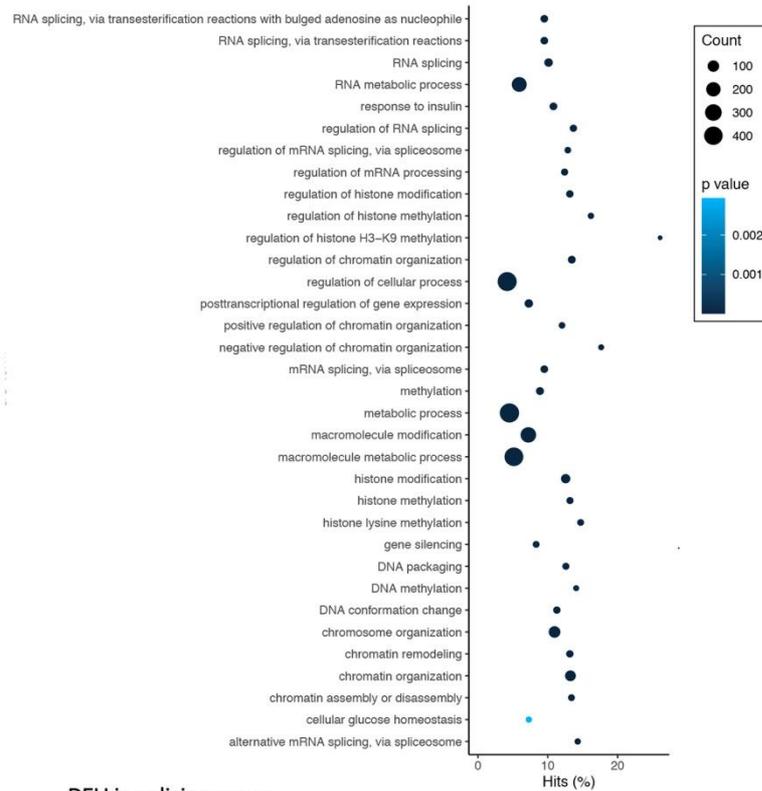
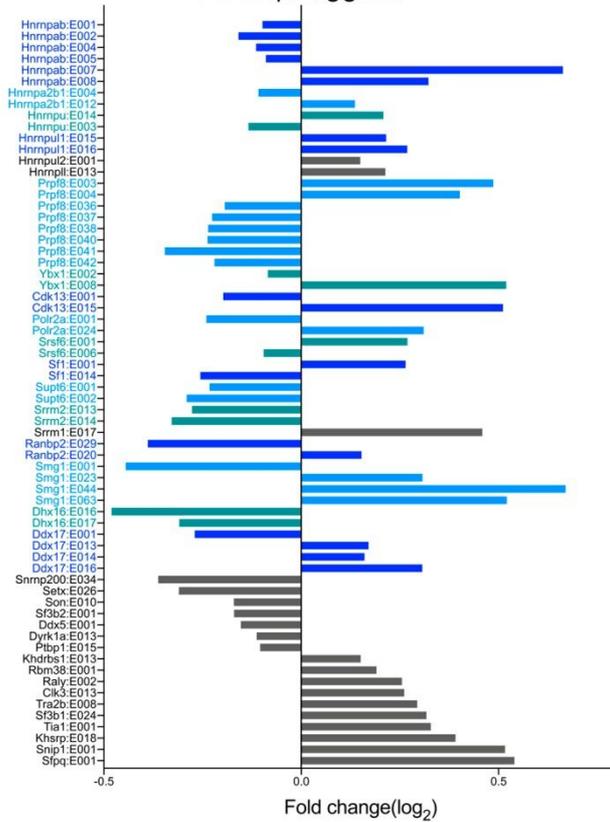
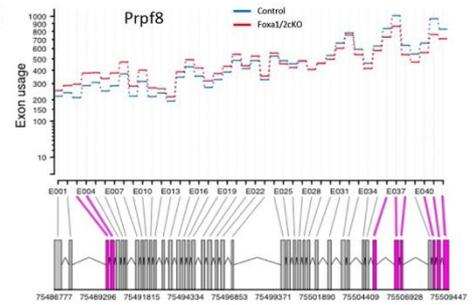
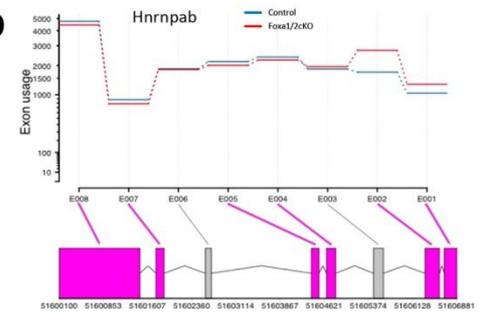
A**GO of biological processes for DEU****B****DEU in splicing genes****C****D**

Figure 5. Foxa1/2 regulate alternative splicing of splicing factors

(A) Dot plot of Gene Ontology (GO) analysis for genes showing DEU. Biological processes involving pioneer function, methylation, metabolic processes and RNA splicing are shown. The y-axis represents the biological process GO terms and the x-axis represents the percentage of genes found in that GO term category. Dot size represents the number of genes and the color indicates the p-value.

(B) Bar chart shows genes that contain DEU which are involved in mRNA splicing. Genes shown in colour show multiple DEU, genes in black show single differential exon changes.

(C-D) DEXSeq representation of *Prpf8* **(C)** and *Hnrnpab* **(D)** that show significant differential exon usage between Foxa1/2 cKO (red) in comparison to control (blue). The exons highlighted in pink have differential exon usage (FDR \leq 0.05).

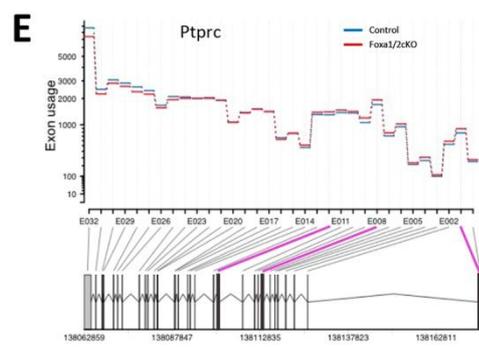
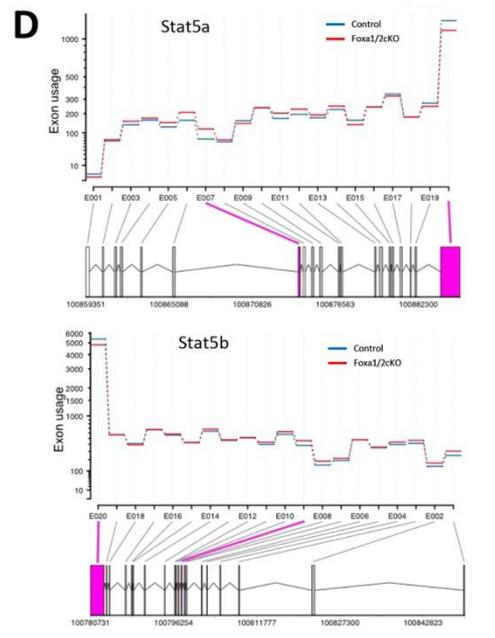
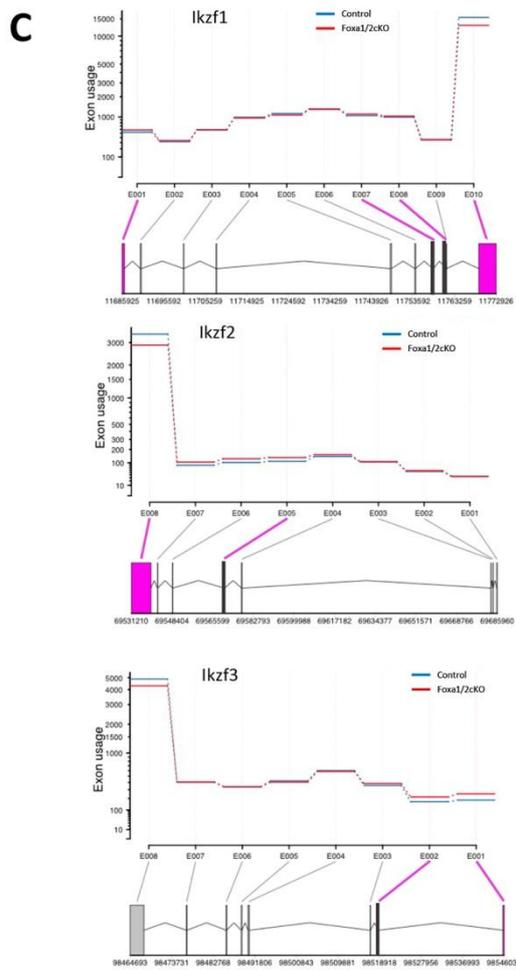
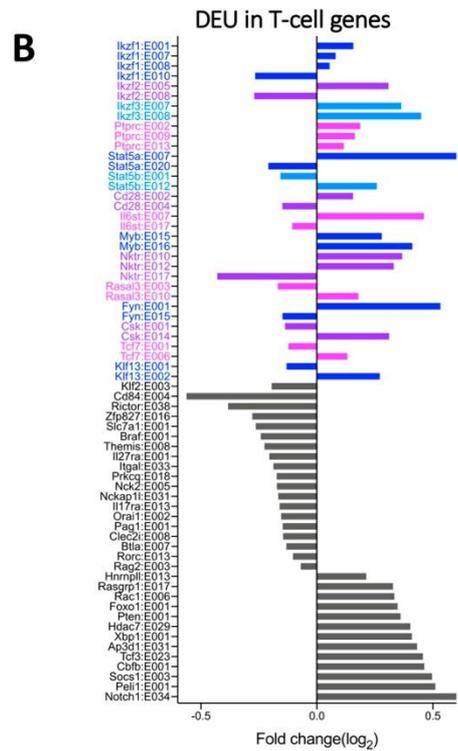
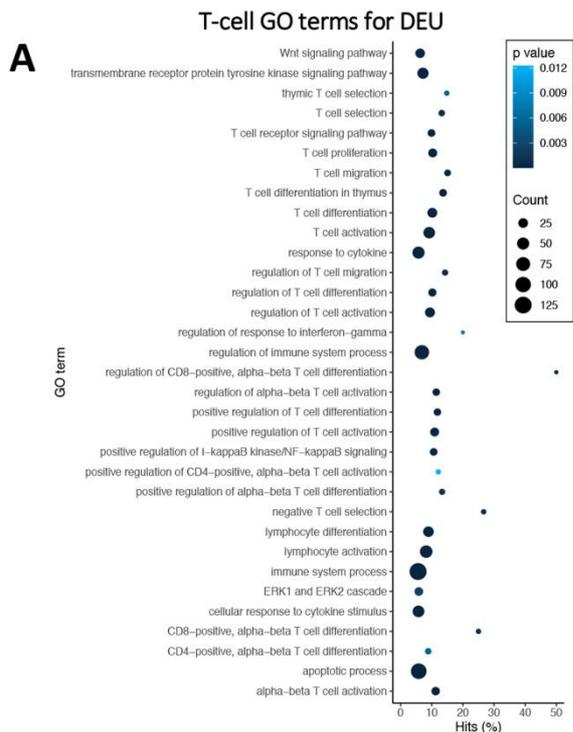


Figure 6. Foxa1/2 regulates alternative splicing of genes required for T-cell development

(A) Dot plot of Gene Ontology (GO) analysis for genes showing DEU. Biological processes associated with T-cell development and function are shown. The y-axis represents the biological process GO terms and the x-axis represents the percentage of genes found in that GO term category. Dot size represents the number of genes and the color indicates the p-value.

(B) Bar chart shows genes with DEU that are involved in T-cell development. Genes in colour represent those with multiple differential exon changes; genes in black have single differential exon changes.

(C-E) DEXSeq representation of *Ikzf1*, *Ikzf2*, *Ikzf3* **(C)**, *Stat5a*, *Stat5b* **(D)** and *Ptprc* **(E)** that show significant differential exon usage between Foxa1/2 cKO (red) in comparison to control (blue). The exons highlighted in pink have differential exon usage (FDR \leq 0.05).

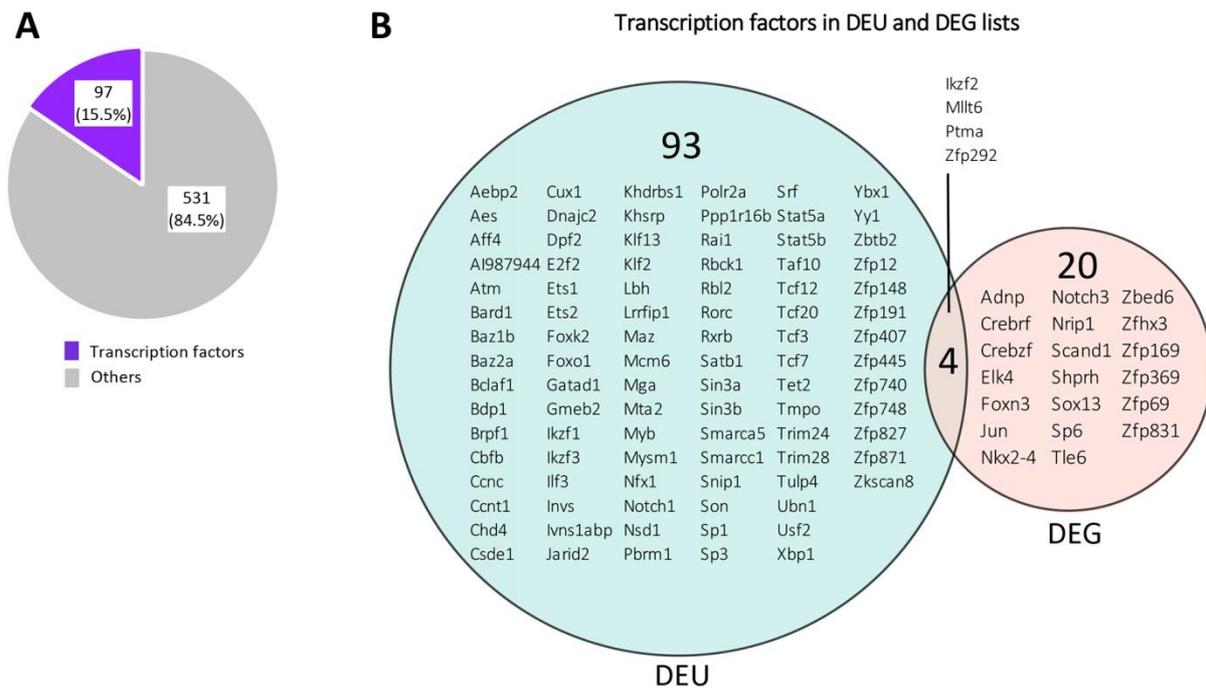


Figure 7. Foxa1/2 regulates alternative splicing of genes that encode transcription factors required for the transition from DP to SP thymocyte

(A) Pie chart illustrates the number of genes showing DEU. Purple section represents number of genes that encode transcription factors, and grey section represents number of genes that do not encode transcription factors.

(B) Proportional Venn diagram illustrates the intersection between DEG that encode transcription factors (pink) and and genes showing DEU that encode transcription factors (green), listing the genes in each category.

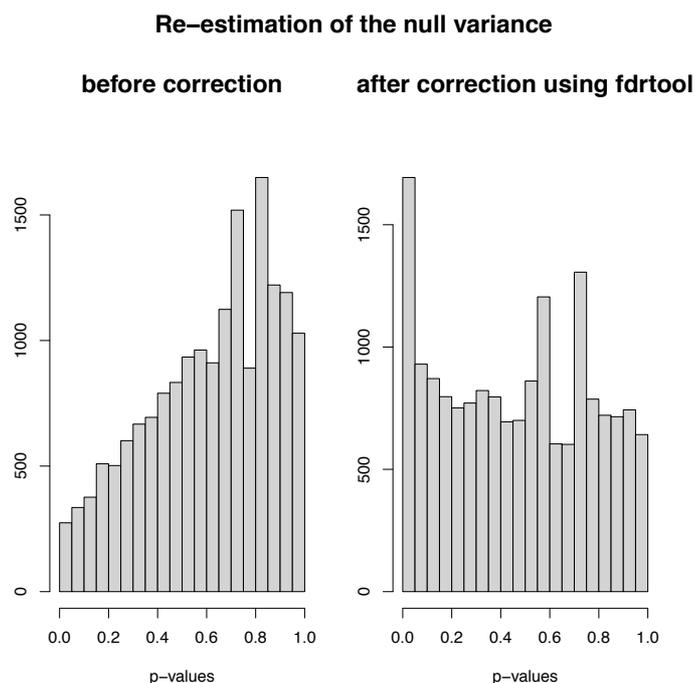


Fig. S1. P-value histogram showing before and after correction of overestimation of variance. Left, histogram of P-values obtained using DeSeq2 showing a hill-shape, indicating an overestimation of the variance in the null distribution. Right, histogram of P-values after z-scores returned by DESeq2 were used as input to fdrtool to re-estimate the null variance.

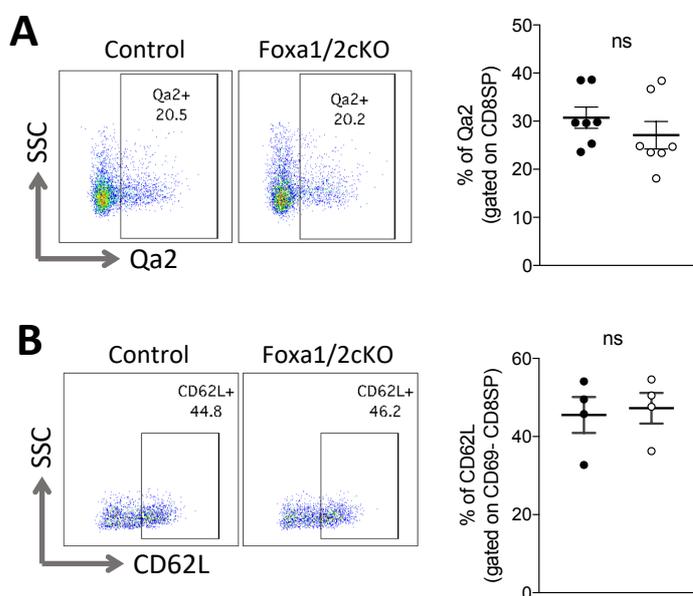


Fig. S2. (A) Flow cytometry profiles show anti-Qa2 staining, gated on CD8SP, giving the percentage of cells in the region shown. Scatter plot shows percentage of Qa2⁺ cells in the CD8SP population. **(B)** Flow cytometry profiles show anti-CD62L staining, gated on the CD69⁻ CD8SP population. Scatter plot shows the percentage of CD62L⁺ cells in the CD69⁻ CD8SP population.

Table S1. RNA-Seq datasets from CD4+CD8+CD69+ thymocytes between control and Foxa1/2cKO.

List of DEG (adjusted p-value < 0.05).

Gene names highlighted in blue indicate genes found in Foxa2 targets in (34).

Gene names highlighted in pink indicate genes found in Foxa1 targets in (44).

Gene names highlighted in purple indicate genes found in both Foxa1 and Foxa2 targets in (34, 44).

Gene names in green indicate genes transcriptionally regulated during positive selection (45).

[Click here to download Table S1](#)

Table S2. Biological Processes(BP) GO term enrichment analysis of DEG GO analysis of biological processes associated with 176 DEG.

[Click here to download Table S2](#)

Table S3. Differential exon usage analysis by DEXSeq from CD4+CD8+CD69+ thymocytes between control and Foxa1/2cKO.

List of the DEU (adjusted p-value < 0.05).

*CD4-Cre transgene contains first 2 exons of CD4 (Exon9 and 10 on reverse strand)(94)

Gene names highlighted in blue indicate genes found in Foxa2 targets in (34).

Gene names highlighted in pink indicate genes found in Foxa1 targets in (44).

Gene names highlighted in purple indicate genes found in both Foxa1 and Foxa2 targets in (34, 44).

[Click here to download Table S3](#)

Table S4. Biological Processes(BP) GO term enrichment analysis of DEU

GO analysis of biological processes associated with genes containing DEU.

padj = p-value after Benjamini-Hochberg false discovery correction

[Click here to download Table S4](#)

Table S5. List of antibodies used in the study

Antibody	Fluorochrome	Catalogue number	Company
CD3	Brilliant Violet 421	100227	Biolegend
CD3	Alexa Fluor® 700	100215	Biolegend
CD4	APC/Cyanine7	100413	Biolegend
CD4	PerCP/Cyanine5.5	100433	Biolegend
CD5	PerCP/Cyanine5.5	100623	Biolegend
CD8	Brilliant Violet 421	100737	Biolegend
CD8	PE	100707	Biolegend
CD24	Brilliant Violet 605	101827	Biolegend
CD44	Brilliant Violet 421	103039	Biolegend
CD62L	APC	104411	Biolegend
CD69	APC/Cyanine7	104525	Biolegend
Foxp3	Alexa Fluor® 647	126407	Biolegend
NK1.1	PE/Cyanine7	156513	Biolegend
Qa2	FITC	121709	Biolegend
Nr4a1	PE	12-5965-82	eBioscience
TCRb	PE/Cyanine7	109221	Biolegend