ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΔΙΑΤΜΗΜΑΤΙΚΟ ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
«ΜΟΡΙΑΚΗ ΒΙΟΛΟΓΙΑ & ΒΙΟΪΑΤΡΙΚΗ»

ΤΜΗΜΑ ΒΙΟΛΟΓΙΑΣ
&
ΙΔΡΥΜΑ ΤΕΧΝΟΛΟΓΙΑΣ & ΕΡΕΥΝΑΣ
ΙΝΣΤΙΤΟΥΤΟ ΜΟΡΙΑΚΗΣ ΒΙΟΛΟΓΙΑΣ & ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ

SATB1-MEDIATED CHROMATIN LANDSCAPE IN MURINE T CELLS

Παπαματθαιάκης Διονύσιος-Αλέξανδρος
Μεταπτυχιακός φοιτητής Διατμηματικού Προγράμματος Μεταπτυχιακών Σπουδών
«Μοριακή Βιολογία & Βιοϊατρική»

Επιστημονικός Υπεύθυνος: Ιωάννης Ταλιανίδης
Επιβλέπων Καθηγητής: Χαράλαμπος Σπηλιανάκης

Ηράκλειο, Οκτώβρης 2020

ΕΥΧΑΡΙΣΤΙΕΣ

Καταρχάς θα ήθελα να ευχαριστήσω θερμά τον μέντορα μου, αναπληρωτή καθηγητή κ. Χαράλαμπο Σπηλιανάκη, για την αμέριστη εμπιστοσύνη που έδειξε στο πρόσωπο μου και την άριστη καθοδήγηση του τόσο σε επιστημονικό όσο και σε προσωπικό επίπεδο, η οποία έπαιξε καθοριστικό ρόλο στην έκβαση αυτής της μεταπτυχιακής εργασίας. Είμαι ευγνώμων που μου έδωσε τη δυνατότητα να ασχοληθώ με το συγκεκριμένο project *καθώς και που με εμπιστεύτηκε να διεκπεραιώσω άκρως απαιτητικές πειραματικές διεργασίες. Χωρίς την διαρκή επίβλεψη του, την διαθεσιμότητα του και την καίρια παρέμβαση του στην επίλυση προβλημάτων δεν θα είχε πραγματοποιηθεί τίποτα από όσα παρουσιάζονται παρακάτω. Δεν θα ξεχάσω ποτέ την πολύτιμη βοήθεια του στην διαμόρφωση των δυνατοτήτων μου καθώς χάρις σε αυτόν κατάφερα να ωριμάσω τόσο επιστημονικά όσο και προσωπικά, παίρνοντας όσα περισσότερα μπορούσα από την προσωπικότητα του κατά την συνεργασία μας.*

Θα ήθελα να ευχαριστήσω, την Μανουέλα Καψετάκη για τις πολύτιμες συμβουλές και την βοήθεια στην ανάλυση δεδομένων, την Ντέππυ Τσουκάτου για την τεχνική υποστήριξη και βοήθεια σε θέματα που αφορούν τον χειρισμό και τη χρήση πειραματόζωων.

Ευχαριστώ επίσης τα υπόλοιπα μέλη του εργαστηρίου, Τζώρτζη Παπαδόγκωνα, Αλέξανδρο Μινακάκη, Θανάση Φιλελέ, Γιοάνα Τρόκκα, Μαρία Λάσκου και Έλενα Στόκα για την άριστη συνεργασία, το ευχάριστο κλίμα που δημιουργούσαν στο εργαστήριο και τις πολύτιμες συμβουλές τους κατά τη διάρκεια αυτού του χρόνου.

Επιπλέον θα ήθελα να ευχαριστήσω την κ. Ειρήνη Στρατηδάκη για την αμέριστη βοήθεια και τεχνική υποστήριξη που μου παρείχε στην διεκπεραίωση των πειραμάτων αλληλούχισης.

Τέλος ευχαριστώ θερμά τους γονείς μου για την αγάπη και τη βοήθεια τους όλα αυτά τα χρόνια. Τη Σοφία για την αγάπη της και την καθημερινή στήριξη της στα εύκολα και στα δύσκολα. Τους φίλους μου που και είναι πάντα δίπλα μου δημιουργώντας ευχάριστες στιγμές.

Στους Γονείς μου

Στον Μέντορα μου

CONTENTS

ABSTRACT

Mammalian genomes are characterized by the higher order of their chromatin structure in the 3D nuclear space as their genome architecture epigenetically regulates gene expression patterns via the recruitment of many pioneer factors. Special AT-rich sequence Binding protein 1 (SATB1) is a chromatin organizer which plays a crucial role in T-cell development. The absence of SATB1 leads to arrested T-cell development and autoimmunity as it is a master regulator of the T-regulatory cell lineage, governing its early stage development. Although recent studies have uncovered SATB1's role in mediating short-range chromatin interactions in immune specific genes, the mechanism by which SATB1 mediates enhancer-promoter communication remains elusive. In the present study we focused on delineating the role of SATB1 in regulating the chromatin landscape of pioneer transcription factors and chromatin modifiers. We showed that SATB1 physically interacts with the p300 acetyltransferase and that they both share the same subnuclear localization and therefore they colocalize extensively. In order to investigate SATB1's role in enhancer patterning and formation we performed H3K27ac ChiP-seq experiments in wild type C57BL/6 (WT) and *Satb1*$^{fl/fl}$*CD4-Cre* (*Satb1* cKO) thymocytes. The overall levels of, the activating for transcription, H3K27ac remain the same between WT and *Satb1* cKO thymocytes, as also supported by immunofluorescence experiments, but specific immune-related genomic loci are deregulated in the absence of SATB1 and characterized by either increased or decreased H3K27ac deposition in their regulatory regions. We also observed the downregulation of immune checkpoint molecules as well as molecules responsible for the proper development, maturation and cell lineage commitment in combination with the upregulation of pro-inflammatory cytokines and transcription factors destined to be expressed in later time points. We conclude that SATB1 is necessary for the proper spatiotemporal regulation of gene expression of several T cell lineages and its absence may lead to an autoimmune disease-related phenotype.

ΠΕΡΙΛΗΨΗ

Το γονιδίωμα των ανώτερων θηλαστικών χαρακτηρίζεται από την πολύπλοκη και πολυδιάστατη οργάνωση της δομής της χρωματίνης στον τρισδιάστατο χώρο του πυρήνα. Η αρχιτεκτονική του ευκαρυωτικού γονιδιώματος επηρεάζει άμεσα την επιγενετική ρύθμιση της γονιδιακής έκφρασης μέσω της στρατολόγησης σημαντικών μεταγραφικών παραγόντων και αναδιοργανωτών της χρωματίνης. Η πρωτεΐνη Special AT-rich sequence Binding protein 1 (SATB1) είναι ένας οργανωτής της χρωματίνης διαδραματίζοντας σημαντικό ρόλο στη διαμόρφωση της χρωματινικής δομής των Τ λεμφοκυττάρων. Η απουσία της SATB1 οδηγεί

σε διακοπή της ανάπτυξης των Τ λεμφοκυττάρων προσδίδοντας στο ποντίκι φαινοτυπικά χαρακτηριστικά που προσομοιάζουν με αυτοανοσία. Τα τελευταία χρόνια, μελέτες για την SATB1 έχουν επιβεβαιώσει την ικανότητά της να δημιουργεί βρόχους σε μικρές γονιδιωματικές αποστάσεις καθώς και τον ρόλο της στην ορθή ανάπτυξη διαφόρων γενεαλογιών των Τ λεμφοκυττάρων. Παρόλα αυτά ο μηχανισμός μέσω του οποίου η SATB1 διαμεσολαβεί τη δημιουργία γειτνιάσεων μεταξύ ενισχυτών και υποκινητών γονιδίων παραμένει άγνωστος. Η παρούσα μελέτη εστιάστηκε στη μελέτη της αλληλεπίδρασης της πρωτεΐνης SATB1 με διάφορους μεταγραφικούς παράγοντες και στο ρόλο της να ρυθμίζει τη δράση τους. Παρατηρήθηκε αλληλεπίδραση της SATB1 με την ακετυλοτρασφεράση p300 και επιβεβαιώθηκε ότι έχουν παρόμοια υποκυτταρική κατανομή στον πυρήνα των θυμοκυττάρων. Επιπροσθέτως, πραγματοποιήθηκε ανοσοκατακρήμνιση χρωματίνης για την ενεργοποιητική της μεταγραφής τροποποίηση της ιστόνης H3K27ac σε αγρίου τύπου C57BL/6 (WT) και *Satb1*$^{fl/fl}$*CD4-Cre* (*Satb1* cKO) θυμοκύτταρα. Το συνολικό επίπεδο της ακετυλίωσης της ιστόνης 3 (H3) στη λυσίνη 27 (K27) δεν έδειξε σημαντικές διαφορές μεταξύ των δύο καταστάσεων, το οποίο επιβεβαιώθηκε και με πειράματα ανοσοφθορισμού για την H3K27ac και τη SATB1. Παρόλα αυτά παρατηρήθηκε απορρύθμιση στην εναπόθεση της ακετυλίωσης σε συγκεκριμένες γονιδιωματικές περιοχές όπου είτε μειώθηκαν είτε αυξήθηκαν τα επίπεδά της. Ενδιαφέρουσα ήταν η παρατήρηση μειωμένης εναπόθεσης και συνεπώς ενεργότητας σε ρυθμιστικές περιοχές γονιδίων που ευθύνονται για τον έλεγχο της ορθής ανάπτυξης, διαφοροποίησης και καθορισμού των διαφορετικών κυτταρικών γενεαλογιών των Τ κυττάρων. Αντίθετα, παρατηρήθηκε αυξημένη εναπόθεση σε ενεργοποιητικές προφλεγμονώδεις κυτοκίνες και μεταγραφικούς παράγοντες που καθορίζουν την ενεργοποίηση και τη δράση κυτταρικών γενεαλογιών οι οποίες αναπτύσσονται σε μετέπειτα στάδια. Συμπερασματικά, η SATB1 είναι ικανή και απαραίτητη για τη χωροχρονική ρύθμιση της γονιδιακής έκφρασης σε διάφορες γενεαλογίες των Τ κυττάρων και η απουσία της μπορεί να οδηγήσει σε φαινότυπο που προσομοιάζει αυτόν της αυτοανοσίας.

INTRODUCTION

1. Chromatin and Gene Expression

The nucleosome is the fundamental element of chromatin and is composed of an octamer of histones (two of each of H2A, H2B, H3 and H4) wrapped around 146bp of DNA[1,2]. Chromatin fibers can be observed in an electron microscope as beads on a string structure which are separated by linker DNA of 40bp. Such subnuclear organization favors the packaging of approximately 2-meter-long DNA in each nucleus but this is not the only reason of its existence. Chromosomal conformation and organization plays a crucial role in gene regulation and expression in order to construct a specific programme of spatio-temporal genome control. The level of chromatin compaction determines whether specific transcription factors are able to bind and regulate specific regions· nucleosomes act as a barrier depending on their post-translational modifications[3–5]. Specifically, covalent changes in the exposed histone tails of the nucleosomes can change the compaction levels of chromatin. For example, histone acetyltransferases (HATs) catalyze the acetylation of several residues and attract the binding of several transcription factors which use their bromodomains. This modification in reversed by histone deacetylaces (HDACs) which leads to transcriptional repression. Lysine methylation is catalysed by Histone Methyltransferases (HMTs) which is recognized by the chromo-domains and plant homeodomains (PHD) of transcription factors and it is reversed by histone demethylases (KDMs). So, the interplay of several histone-modifying enzymes with RNA Polymerase II regulates the transcriptional onset either by remodelling or stabilizing chromatin. Modification of specific lysine residues by methylation, such as H3K4 and H3K36 facilitate the active state of chromatin, whereas modifications of H3K9 define a repressive chromatin state. In general, methylation of CpG islands marks transcriptional inactive sites as this state of chromatin blocks the binding of several transcription factors. This phenomenon lies in the fact that this state promotes the binding of histone deacetylase enzymes as well as histone methyltransferases to establish this inactive chromatin state[6–8].

2. T cell development

T lymphocytes originate from hematopoietic precursors of the lymphoid lineage. These multipotent cells migrate to the thymus upon stimulation by the appropriate signals and they are capable of proliferating and expanding their clonal population. A great spectrum of Chordota, from elasmobranchs to mammals, are capable of generating T cells as part of their immune system, making them and the thymus a unique, universal and irreplaceable feature of the immune system[9–13].

T cells are categorized in two classes based on whether they express αβ or γδ T cell receptors. Hematopoietic precursors enter the thymus and primarily reside in the cortex until their interaction with cortical epithelial cells. This communication leads to commitment and TCR expression which drives them to the medulla. In the medulla they are subjected to selection and further maturation in specific thymocyte subclasses[14–16]. Several key players are responsible for these developmental stages such as DLL4, which activates Notch signalling, a critical pathway for the initiation of the gene expression profile of T cells, Kit ligand which facilitates the proliferation of the precursors as they undergo T cell commitment and IL7 acting in a similar fashion to the Kit ligand [17–19]. Activation and proliferation initiated upon these signals leads to a distinct population of Double Negative T-cells (DN, $CD4^-CD8^-$). During these stages, DN T-cells are committing to αβ T cells which later on will mature to Double Positive (DP, $CD4^+CD8^+$) or γδ T cells. RORγt expression is responsible for the progression to the DP stage and those who successfully pass β selection are differentiating to CD4 or CD8 T cells according to the antagonistic expression of GATA3/ThPOK and Runx3 transcription factors respectively[20].

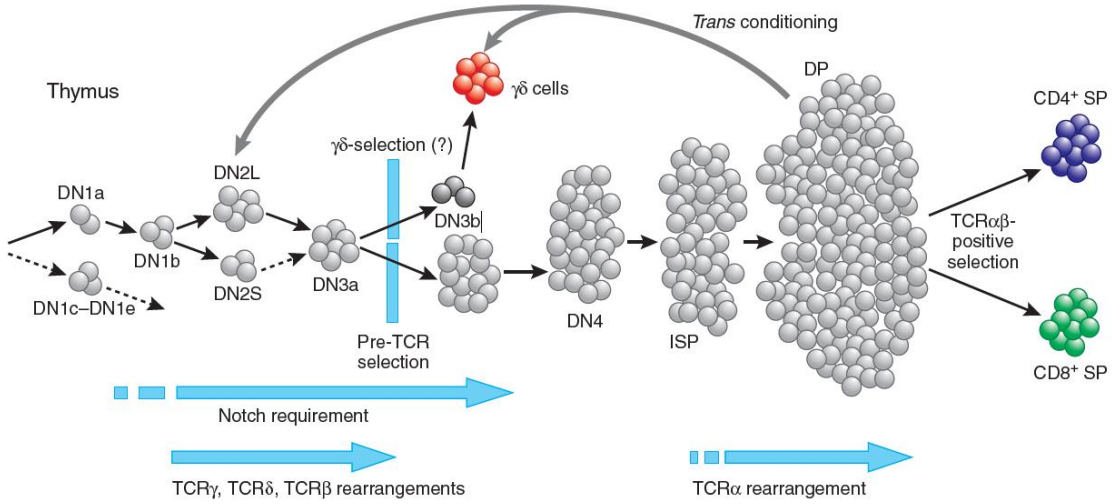

Figure 1 The stages of T cell development and differentiation in the thymus. The number of the cells indicate the relative number of each population at each stage.

3. 3D genome organization during T cell maturation

Major changes in the T cell 3D chromatin organization occur during the transition from the DN2 to DN3 stage of lineage commitment towards β selection and between DN4 to the DP stage where TCRβ is being rearranged and cells are being selected[21]. Specifically, the T cell lineage commitment transcription factor Bcl11b was shown to regulate 3D chromatin organization during T cell development. Several TADs with high Bcl11b binding, as well as

the TAD bearing the *Bcl11b* gene itself, tend to create more intra-TAD interactions during the transition from the DN2 to DN3 stage. The *Bcl11b* enhancer is being repositioned from the nuclear lamina to the interior of the cell nucleus during DN2 with the help from ThymoD, a cis-acting long non-coding RNA, which mediates the participation of its promoter and enhancer in the same loop. ThymoD null mice presented an increase in hypermethylated CpG islands within CTCF binding sites, thus disabling CTCF and cohesin binding in the *Bcl11b* enhancer. Coming to strengthen this notion, the ThymoD antisense transcript regulates the demethylation of CTCF binding sites[22]. T cell activation upon antigen stimulation detonates crucial rapid chromatin rearrangement. Specifically, splenocytes from C57BL/6 mice presented major chromatin reorganization during anti-CD3 treatment indicating transcriptional activation and rapid gene expression[23].

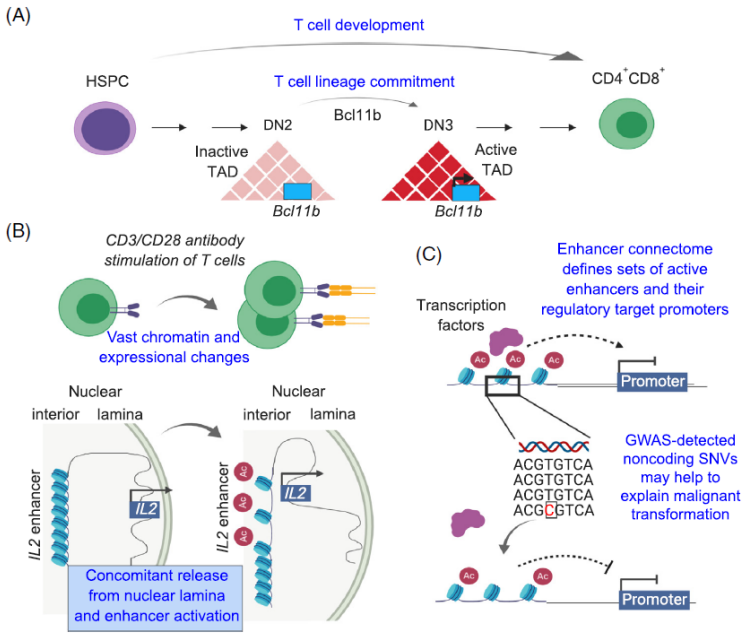

Figure 2 Genome organization changes in the 3D nucleus. (A) The topologically associated domain(TAD) where Bcl11b resides, is subjected to major conformation changes during DN2 and DN3 stages concomitantly with Bcl11b transcriptional activation. (B) The il2 locus of Jurkat T cells undergoes major reorganisation and decompaction in response to CD3/CD28 transmembrane stimulation. (C) Non-coding mutation in enhancer regions may disrupt the enhancer landscape and subsequently the enhancer activity, leading to gene expression deregulation.

4. Enhancer-promoter communication in the 3D cell nuclear space

The first models of gene organization in bacteria depicted promoters and other gene-proximal regulatory elements something that is highly complicated in higher eukaryotes in which regulation can occur hundreds of thousand bases far from the gene. So, in order for the DNA to be properly packed in the nucleus, it forms chromatin fibres which are being further compacted and organized in order to fit in the nucleus. This higher-order 3D organisation not

only facilitates the proper chromosomal segregation during mitosis and meiosis but also accounts for better and controllable gene regulation[24].

Through the last years it has been established that genome organisation plays a crucial role in transcriptional regulation. The nucleus is organised in such a fashion in order to spatiotemporally control gene expression by bringing in close proximity genomic regions capable of activating transcription. The genome is separated in several Topologically Associated Domains (TADs) which are regions of loops closely interacting in the 3D space in order to regulate gene activation or repression by bringing together or separating enhancers and promoters respectively[25,26]. TADs are quite specific insulated regions that tend to interact mainly within defined regions and the deletion of their boundaries leads to the abnormal or aberrant expression of nearby genes[21,27–32]. However, the disruption of an insulated or several insulated regions doesn't lead to exaggerated abnormalities in gene regulation, rather it causes aberrations only in nearby gene regions. This phenomenon leads to the assumption that either TAD boundaries are fine-tuning the cell's transcriptional plan or their function lies in the exclusive regulation of specific sets of genes[33–35]. It is believed that TADs facilitate the segregation of specific chemical activities and biochemical factories for the completion of specific biological functions in order to increase the efficiency by increasing the local concentration of key factors[36]. The existence of activated and repressed genomic regions strengthens the notion that transcriptional activation is tightly linked to chromosomal conformation. It is also possible that robust transcription factors are capable of bringing together these regions in order to initiate transcription[37,38].

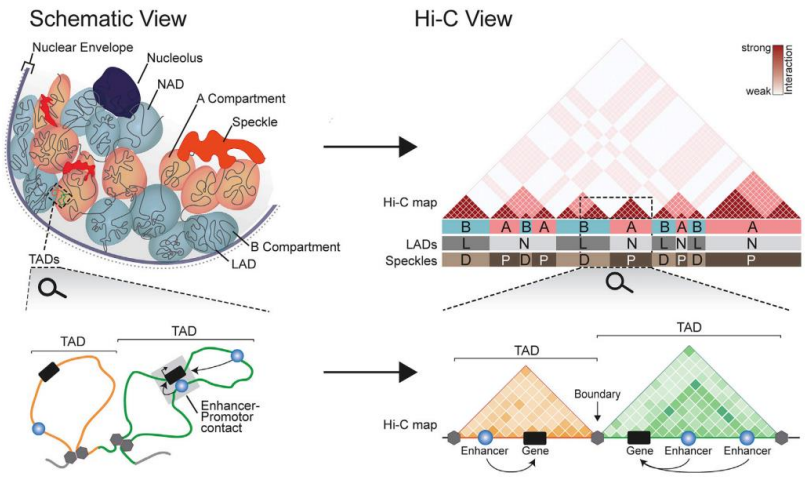

Figure 3 Schematic and contact matrix (from HI-C) representation of genome organisation. The transcriptionally active and inactive compartments are separated from each other to form A and B compartments, respectively. A compartments are closer to the nucleus centre and do not associate with Lamin-Associated Domains (LADs) in contrast to B compartments. The magnification shows that although these two distinct TADs are in a close proximity their interactions and gene regulation is strictly limited within their boundaries, which create discrete functional structures.

The quintessence of gene activation is the interaction of an enhancer to its cognate promoter mediated by several transcription factors and molecules. These long-range interactions must be tightly regulated in order to facilitate strict and specific activation of the correct genes in a spatiotemporal fashion, taking under acount that the number of enhancers are more than the number of genes[39–42]. In order to fine-tune the mode of action of enhancers to their cognate genes, several factors such as CTCF and YY1 sculpt the genomic landscape in order to organise the genome in district regulatory regions[34,35,43,44]. In addition, several other transcription factors such as p300 and the Mediator complex facilitate the generation of enhanceosomes capable to activate transcription[39–42]. It is obvious that an enhancer-promoter interaction is strictly dependent on gene expression, but the main question is whether gene expression is the cause or the consequence of 3D genome organisation. A revolutionary work has shown that even in the absence of GATA1, the contact of the mouse β-globin (Hbb) promoter to its locus control region could elicit strong transcriptional activation[45]. Other experiments have shown that even after the engagement of RNAPII and transcription onset the enhancer-promoter contact remains and probably travels along the transcribed regions with RNAPII or remains anchored at specific factories from where genes are reeled along[46].

Evidence indicated that after the completion of transcription, enhancer-promoter communication is lost concomitantly with gene repression[28]. TADs are submegabase-sized genomic domains which mainly present intradomain interactions[47,48]. This phenomenon restricts each domain and prevents cross-boundary communication events and aberrant or not-regulated gene activation. A reporter gene integration study showed that genes located in the same TAD function in a similar tissue-specific fashion[26], indicating the strong action of distal enhancers and that the overall structure is spatiotemporally regulated. The TAD-specificity case can't be completely confirmed because several data present differential expression (expression and repression) of several genes within the same TAD indicating that different genes can display different activation levels[49].

5. The Special AT-rich Binding Protein 1

The Special AT-rich Binding Protein 1 (SATB1) is a MAR-binding protein that serves as a chromatin remodeler and organizer mainly in the thymus and other organs as well (testis, brain, HSCs). Early research in the field has shown that despite the generally unchanged histology of SATB1 null versus wild type mice, there is a significant decrease in the size of the thymus, the spleen and the lymph nodes of the SATB1 KO animal. SATB1 null mice display a 2-fold decrease in the cell population in the thymus by two weeks of age and the maturation of T cells

is mostly arrested at the double positive stage with a striking 50% of the CD4 T cells undergoing apoptosis in the first 24 hours after activation with PMA and ionomycin andcontinued to die. DP cells, that usually don't exit the thymus, were found in the lymph nodes of the SATB1 null mice meaning that SATB1 is essential for the proper selection and maturation of T cells in the thymus[50].

Affinity chromatography experiments for SATB1 aiming in the purification of protein interactors showed that SATB1 was copurified with several essential factors, members of the NURD complex. Further co-immunoprecipitation experiments showed a clear interaction of SATB1 with Mi-2, mSin3A, MTA-2 and HDAC1. SATB1 also interacted with human recombinant hACF1 and hSNF2H, whereas truncated SATB1 failed to interact indicating that the 90-365 aa region is crucial for protein-protein interactions. These data indicate a significant role of SATB1 in chromatin remodelling and transcriptional regulation. ChIP experiments with SATB1-specific antibodies identified a 700bp fragment of the *IL-2Ra* gene locus as a high affinity binding region (SATB1 Binding Site700). Similar to SATB1, HDAC, ACF1 and ISWI were shown to interact with SBS700 in wt but not in SATB1 null mice. These data show that SATB1 regulates both the repression as well as the remodelling of the *IL-2Ra* locus. Furthermore, SATB1 loss lead to changed nucleosomal positioning in intronic sequences downstream of exon 1 and 7kb upstream of SBS700, delineating the role of SATB1 in regulating the epigenetic state of *IL-2Ra* locus as far as 7kb[51].

Immunofluorescence experiments in thymocytes presented a cage-like pattern of SATB1 in the nucleus and specifically in regions deprived of heterochromatin as deduced by the DAPI-dense staining of heterochromatin regions. Surprisingly, this pattern could not be distracted using high salt extraction or DNase I treatment, delineating the strong interaction of SATB1 with the bases of chromatin loops on the nuclear matrix. In further supporting these findings of SATB1's long-range regulatory potential, it was shown that SATB1 can regulate the chromatin structure of SBSs as long as 160kb from its binding site [52]. Following the previous findings, it was shown that SATB1 creates several chromatin loops and regulates the clustering of key cytokine genes within a 200kb region. These data show that SATB1 coordinately regulates the transcriptional activation of several cytokine genes (IL4, IL5, IL13, Rad50, Sept8, Kif-3a) by changing the chromatin structure upon Th2 cell activation[53].

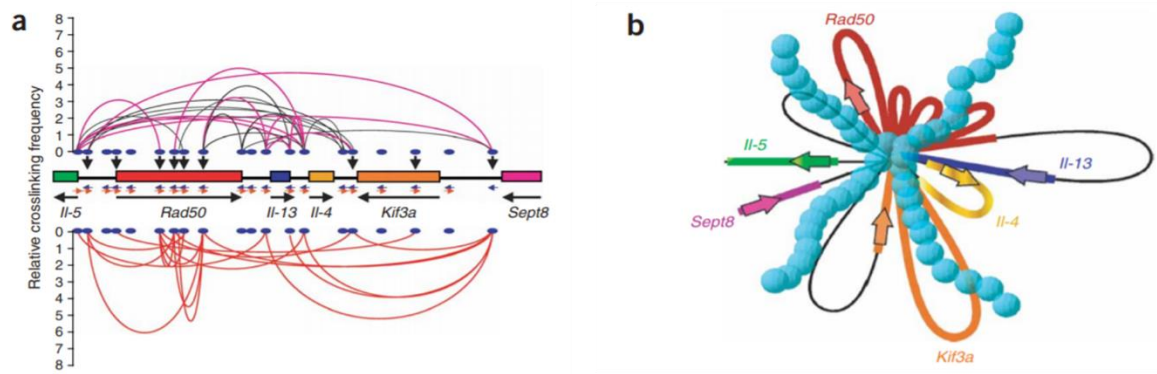

Figure 4 (a) The Th2 locus representation in which pink lines indicate increased crosslinking frequency after activation whereas the red lines indicate the proximity of the connecting regions (based on 3C and ChIP assays). The vertical arrows indicate the SATB1 binding sites. (b) Schematic model representation of Th2 locus conformation from interactions and events shown at (a). All loops are connected to the same central region decorated by SATB1.

Studying the post-translational modifications which modulate SATB1's function, Kumar et al showed that SATB1 can be phosphorylated and acetylated by PKC and PCAF respectively. They showed that when SATB1 is phosphorylated it can interact with HDAC1 and lead to transcriptional repression by directly binding to DNA. Dephosphorylated SATB1 is free to interact with PCAF and become acetylated leading to its disassociation from the DNA and the transcriptional de-repression of its target genes[54]. These data support a binary mode of action for SATB1 under several conditions.

Trying to identify the transcriptional program and the suppressive function of Tregs Bayer et al showed that the *Satb1* locus is controlled by FOXP3 through repressing its chromatin state. They showed that loss of FOXP3 releases SATB1 and leads to reprogramming of Tregs to effector T cells[55]. In another study, Kitagawa et al performed ChIP-seq experiments and showed that SATB1 binds many Treg-specific super-enhancers as well as common super-enhancers. What was more surprising was that ATAC-seq experiments revealed the binding of SATB1 in close chromatin regions presenting a pattern which opposes to the binding pattern of typical transcription factors[56].

Ghosh et al carried out a series of biophysical experiments and identified that the majority of SATB1's binding sites are found within the nucleosomal core of the chromatin fiber. These sites are composed of higher A/T content in their flanking sites (75%) than the unbound ones and showed that the distance between these dinucleotides which creates an atypical SATB1 binding consensus matters more that the AT content per se. More important was the finding that the CUT domain confers higher affinity whereas the Homeo Domain facilitates the specificity of binding[57].

MATERIALS AND METHODS

Animal models and cells.

The mice used for all the experiments were either C57BL/6 or *Satb1*$^{fl/fl}$*CD4-Cre* conditional knockout mice in the same genetic background. Mice were sacrificed and thymi were isolated. Tissues were smashed in ice cold 1X PBS and filtered through a 50µm pore mesh. Single cell suspensions were washed twice with cold 1X PBS supplemented with 1mM PMSF when needed. Cells were always quickly fixed and treated accordingly to each experimental procedure.

Co-Immunoprecipitation

Protein extracts were prepared from 4-9 weeks old female C57BL/6 mice. Thymocytes were incubated at room temperature for 30 min with rotation, in lysis buffer (150mM NaCl, 50mM Tris pH=7.5, 5% Glycerol, 1% Nonidet P-40, 1mM $MgCl_2$, 1X Proteinase Inhibitors and 1mM PMSF). 600µg of protein extract were incubated with 2.5-5µg anti-SATB1 antibody or rabbit IgG sera (#C2712 Santa Cruz Biotechnology) overnight at 4ºC. After 16 hours, 20µl of Dynabeads® Protein G (#10004D LifeTechnologies) were incubated with the immunocomplex at 4ºC for 2 hours. Beads with immunocomplexes were washed three times with buffer I (150mM NaCl, 50mM Tris pH=7.5, 5% Glycerol, 0.05% Nonidet P-40 and 1mM PMSF), twice with Nonidet P-40 free Buffer II (150mM NaCl, 50mM Tris pH=7.5, 5% Glycerol and 1mM PMSF) and finally resuspended in 2X SDS Loading Dye and boiled at 95ºC for 10min.

Chromatin Immunoprecipitation

Thymi were isolated from 4-9 weeks old female C57BL/6 or *Satb1*$^{fl/fl}$*CD4-Cre* mice, washed and filtered to make single cell suspensions. Cells ($15x10^6$) were cross-linked with 1% formaldehyde (0,1V fixation buffer: 11% Formaldehyde, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 50 mM Hepes pH 8.0 in 0.9V 1X PBS) at room temperature for 10min, quenched with 0.125 M glycine and lysed in SDS lysis buffer (1% SDS, 50 mM Tris (pH 8), 20 mM EDTA and 1X Protease Inhibitors - PIC). Nuclei were resuspended in 60µl TE Buffer (10mM Tris-Cl pH=8 and 1mM EDTA) and sonicated for 2-5 min using the Labsonic M – Tip sonicator for chromatin fragmentation in a range of 200-1000bp. Fragmentation was verified both by agarose gel electrophoresis and bioanalyzer analysis. 50µl of fragmented chromatin was used as Input DNA and to check shearing efficiency. For the immunoprecipitation, chromatin samples were treated with 1% Triton X-100 for 10 minutes at 37ºC and diluted with equal

volume of ChIP Binding Buffer (20 mM Tris pH 8.0, 2 mM EDTA, 30 mM NaCl, 0.2% Na-Deoxycholate, 2X PIC), in order to lower SDS concentration and were precleared using 10μl of prewashed and preblocked magnetic protein A/G beads. 30-50μl of precleared and preblocked with PBS/BSA (0.1%) magnetic protein A/G beads were incubated with 4-8μg of each antibody or IgG at 4^{o}C for 4-6 hours. The beads were washed with BSA/PBS buffer (1X PIC) three times, were resuspended in the precleared samples and incubated for 16 hours at 4°C with rotation. Beads were washed with fresh RIPA (50 mM Hepes (pH 8.0), 1% NP-40, 0.70% Na-Deoxycholate, 0.5 M LiCl, 1X PIC, 1 mM EDTA) buffer and were reverse cross-linked in ChIP Elution (10 mM Tris-HCl (pH 8.0), 5 mM EDTA, 300 mM NaCl and 1% SDS) buffer for 16 hours at 65^{o}C. Samples were diluted with ddH_2O and the supernatant was collected in order to be subjected in RNase and Proteinase K treatment.

Immunofluorescence

Freshly-prepared thymocytes were isolated and placed on Poly-D-Lysine coated glass coverslips. Cells were crosslinked with 4% PFA in 1X PBS for 10 minutes on ice and permeabilized with 0.5% TritonX-100 for 5 minutes on ice. Blocking was performed with 0.4% acetylated BSA for 30 minutes at room temperature and each antibody was diluted in detection buffer (4X SSC, 0.1% acetylated BSA and 0.01% Tween20). All secondary conjugated antibodies were used in a final dilution of 1:250.

Western Blot Analysis

Polyacrylamide gels were prepared at a 6 or 8% concentration. Nitrocellulose membranes were either blocked using 5% non-fat milk in Tris-Buffered Saline with 0.1% Tween-20 or 5% BSA (Fraction V) in TBST buffer and all antibodies were diluted in 1% of milk or BSA.

ChIP Library Preparation

The Illumina Nextera DNA library Prep kit was used for library preparation. ChIP DNA was tagmented using Tn5 enzyme which incorporates Illumina adaptor sequences in the ends of each fragment. The library was constructed upon amplification with PCR for 8-15 cycles. The index primers used are able to recognize the adaptor sequences and each one bears a specific barcode sequence. The final DNA library was purified by two-side size selection using Ampure Beads in order to maintain the intermediate fragments (150-300bp).

ChIP-seq Analysis

Each sample was destined to collect from 25 to 50 million reads in the Illumina Nextseq 550 instrument running in either a mid or high capacity flowcell. Peak calling was performed using the MACS2 software and each sample was normalized using its own input sample. For ChIP peak annotation we used the ChIP Seeker R package and for differential binding analysis we used the DiffBind R package. All analyses have been performed using the maximum possible significance threshold and anywhere changed is mentioned.

RESULTS

I. p300 acetyltransferase physically interacts with SATB1

Previous experiments of SATB1 immunoprecipitation coupled to Mass Spectrometry analysis in our lab revealed p300 as a potential SATB1 interactors, two crucial factors responsible for gene regulation and transcription. Therefore, in order to biochemically verify these protein interactions, we performed co-immunoprecipitation experiments. Whole-cell protein extracts were incubated with anti-SATB1 antibodies and then western blot was performed for p300. We observed specific bands in the SATB1 IP samples, although their molecular weight was slightly higher than the input sample, which are absent in the IgG control experiment. The membranes were striped and re-blotted for SATB1 in order to validate its precipitation.

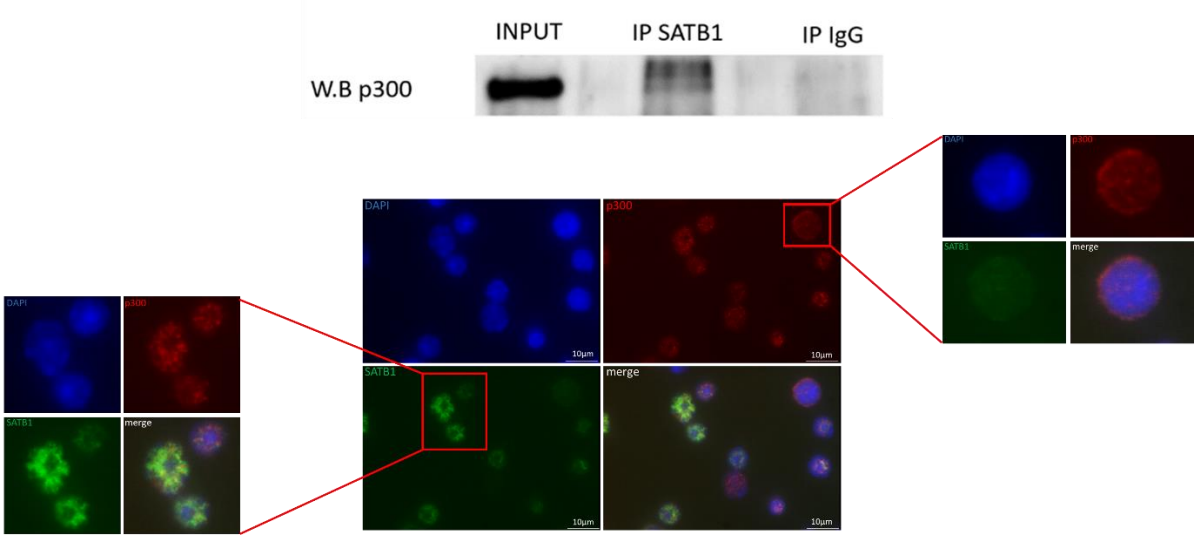

Figure 5 Top: Co-Immunoprecipitation experiment against SATB1 and rabbit IgG and Western Blot Analysis against p300. Bottom: Immunofluorescence experiment against SATB1 and p300. The two magnified cells present extensive colocalization of the two proteins. The SATB1 negative cell expresses p300 in a dispersed pattern possibly due to SATB1's absence.

Furthermore, immunofluorescence experiments were carried out in WT thymocytes in order to visualise the localization of these proteins in comparison to SATB1's cage-like pattern. Both p300 was detected in the cell nucleus of thymocytes and presented a cage-like pattern, quite similar to SATB1's subnuclear localization.

Interestingly, we observed many cells that lacked SATB1 but expressed p300 (probably thymic epithelial cells). In these cells, p300's pattern was dispersed and/or more perinuclear than in SATB1-positive thymocytes, where its localization resembles the cage-like pattern of SATB1. Thus SATB1's presence is indispensable for p300 localization in the thymocyte nucleus.

In order to verify if SATB1's presence is important for p300's expression we carried out Western Blot analysis in whole cell thymocyte extracts derived from WT and *Satb1* cKO animals.

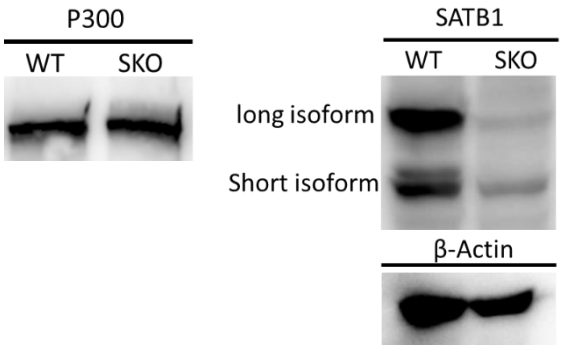

Figure 6 Left: Western Blot analysis of p300 expression between WT and Satb1 cKO protein extracts. Right: Western Blot analysis of SATB1 expression between WT and Satb1 Cko thymic protein extracts. SATB1 is still expressed in in a small percentage of cells of Satb1 cKO thymi, which are probably epithelial cells.

We observed that the p300 total protein levels in the wild type thymocytes and thymocytes that do not express the SATB1 protein were the same. Western blot analysis against SATB1 in protein extracts from WT and *Satb1* cKO thymocytes delineates the expression of Cre recombinase under the control of CD4 promoter and the efficacy to excise the 3rd exon of *Satb1*. Therefore, SATB1 doesn't regulate *EP300* expression rather than regulates p300 through their physical interaction and acts as a chaperone by recruiting/guiding p300 to its genomic binding sites.

II. SATB1-mediated enhancer landscape in T cells

SATB1 plays a profound role in T cell development and *Th2* locus gene expression[53] which is also supported by genomic experiments in our lab, as its loss leads to loss of H3K27ac-mediated chromatin loops. Therefore, it is clear that SATB1 plays an important role in regulating enhancer-promoter communication and therefore gene expression in T cells. Though, its role in generating and regulating specific enhancer regions remains elusive. In order to elucidate the role of SATB1 in enhancer formation and function we carried out a series of Chromatin Immunoprecipitation experiments against p300, H3K27ac which marks active enhancers, utilizing chromatin from both WT and *Satb1* cKO thymocytes.

In order to verify the efficiency of SATB1 chromatin immunoprecipitation we designed primers for specific genomic regions based on the top differentially bound enhancer regions as deduced from previous H3K27ac HI-ChIP experiments performed in the lab.

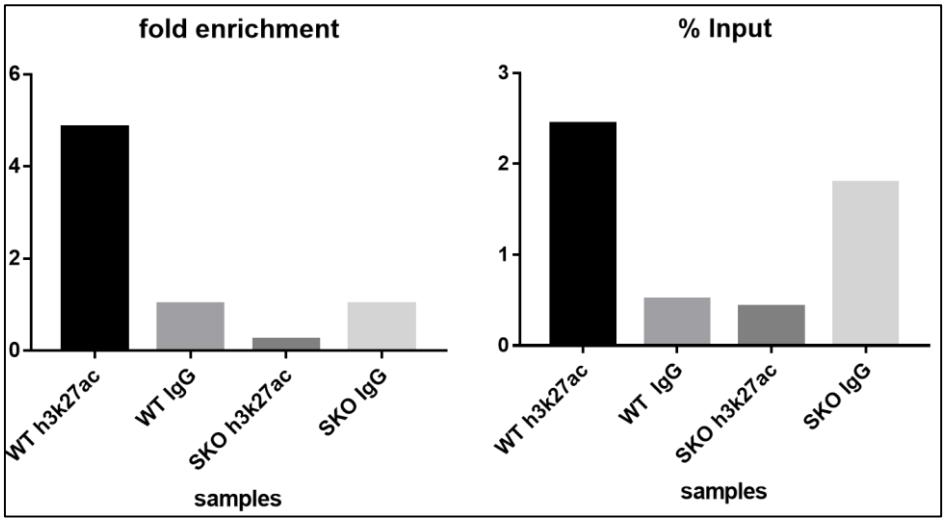

Figure 7 ChIP assay against SATB1 and rabbit IgG coupled to qPCR using primers that recognize a known enhancer region whose occupancy by H3K27ac is decrease in the absence of SATB1. The Satb1 Cko H3K27ac ChIP samples present a 4.5-fold decrease compared to the WT H3K27ac ChIP samples.

Based on the qPCR results we observed a 4.5-fold decrease of H3K27ac ChIP in the *Satb1* cKO compared to the WT. This finding indicates that indeed this enhancer is less bound and subsequently less active in the *Satb1* cKO thymocytes. Therefore, this region's DNA is underrepresented in our immunoprecipitated DNA fragment samples. We conclude that our protocol is valid and can be trusted in order to generate samples destined for sequencing.

So, we prepared thymocyte chromatin samples and created ChIP DNA libraries in order to perform sequencing. Each condition (WT and *Satb1* cKO) was performed from two biologically discrete samples. In general, our WT H3K27ac ChIP-seq samples generated 45534 and 19726 peaks, whereas the *Satb1* cKO H3K27ac ChIP-seq samples gave 11851 and 15340 thousand, as they were analysed equally with the same parameters by MACS2 peak caller.

Taking into account the H3K27ac peaks from each replicate of each condition we observe a mean 40% decrease in H327ac peaks in the absence of SATB1 but these calculations are not indicative of H3K27ac expression as WT1 sample generated 25808 more peaks than WT2.

In order to have a clearer image of H3K27ac distribution pattern in the nucleus we performed immunofluorescence experiments against SATB1 and the histone mark. H3K27ac has a similar cage like pattern like SATB1 and they present extensive co-localization. We observe that WT and *Satb1* cKO thymocytes have similar levels deposition of H3K27ac. In the image below we have a magnified area with 4 cells (2 of which are expressing and 2 of which do not express SATB1). It is clear that the intensity of the histone mark is higher in the SATB1 positive cells although present in high levels in the cKO.

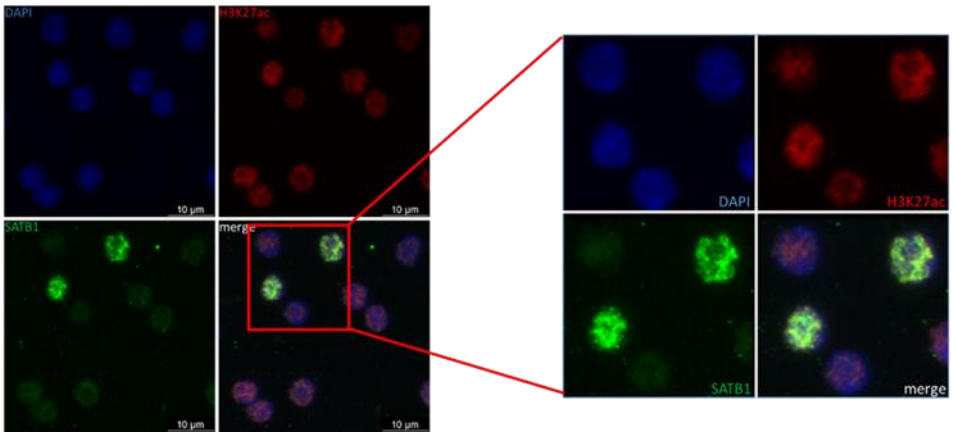

Figure 8 Immunofluorescence experiment against SATB1 and H3K27ac in Satb1 cKO thymocytes. In the magnified region of interest there are four cells, two that express SATB1 and two that don't.

To gain insight about the differential H3K27ac deposition between WT and cKO thymocytes we performed intersection analysis the WT versus the *Satb1* cKO H3K27ac peaks.

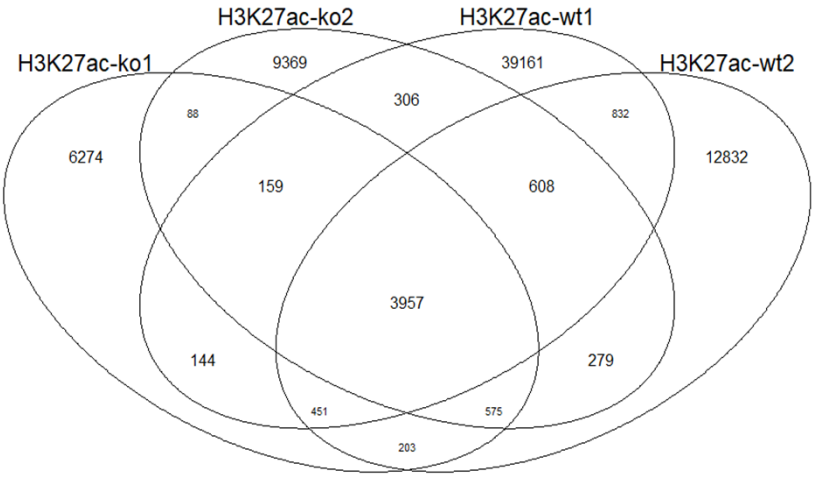

Figure 9 Venn Plot of the overlapping regions between all replicates in WT and Satb1 cKO H3K27ac samples

We found 4068 peaks intersecting in both conditions, 45455 peaks unique for the WT samples and 5487 peaks unique for the KO samples. Therefore 45455 peaks are retained in the absence of SATB1, and 5487 are created, meaning that these loci are decorated with activating H3K27ac marks because of the loss of SATB1, who previously led to their repression.

Additionally, using the ChipSeeker R package in order to visualize the H3K27ac distribution around TSS regions, we found that the distribution of WT and *Satb1* cKO thymocytes have a similar intensity with no significant differences.

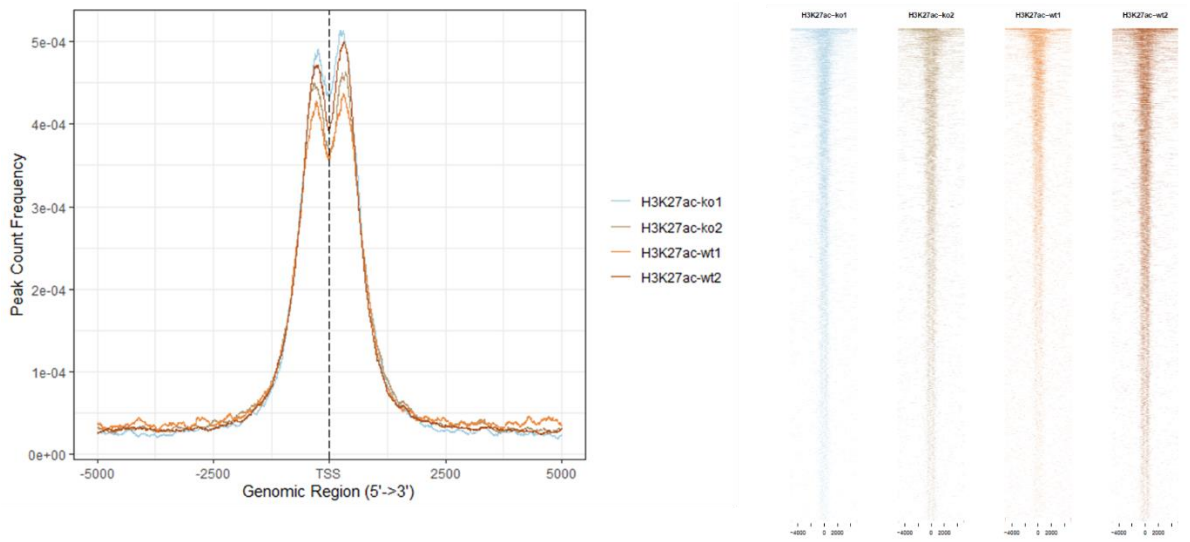

Figure 10 Average peak count frequency plot(left) and Heatmap(right) of H3K27ac distribution around TSS regions in WT and Satb1 cKO thymocytes.

However, these results are not conclusive and do not illustrate the H3K27ac intensity in specific regions of significantly decreased or increased H3K27ac levels in the absence of SATB1.

Differential binding analysis between WT and *Satb1* cKO T cells

In order to identify the significantly differential bound H3K27ac regions between WT and *Satb1* cKO thymocytes we performed differential binding analysis using the DiffBind R package. This tool uses both the aligned mapped single read files (BAM files) of all samples counts as well as the coordinates from the qualified peaks generated from MACS2 peak calling. All ChIP-seq libraries have been normalized over their sequencing depth and their control files (input chromatin sequencing samples). After counting the reads, the tools re-centers them in a range of 500bp and normalizes them using each sample's FrIP factor (Fraction of Reads in Peaks). 11904 intervals were common in all peaksets (WT and cKO). The analysis uses EdgeR and DESeq2 in order to perform differential analysis between the two conditions. Taking into account these parameters, the tool performed analysis with a FDR (False Discovery Rate) cut-off of 0.05 and generated 253(271 if FDR<=0.05) differentially bound regions (pvalue<1.04^{-3}). The results contained the normalized fold enrichment between the two conditions.

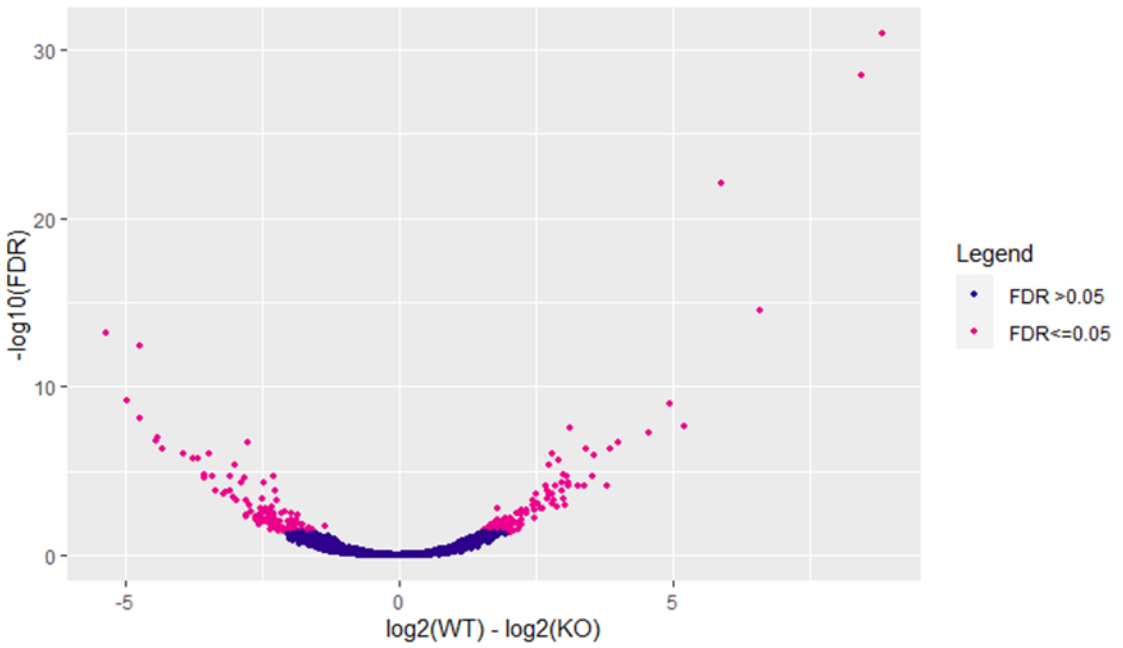

Figure 11 Volcano plot of significant differential bound H3K27ac regions. Red spots indicate the significantly differential bound regions. Those who are in the negative site of the x axis are upregulated whereas those in the positive site of the x axis are downregulated in the Satb1 cKO samples.

The common regions were plotted using a clustering analysis of differential binding affinity among the two conditions. The heatmap below indicates the two clusters of differential enrichment between the WT and *Satb1* cKO samples based on their binding affinity.

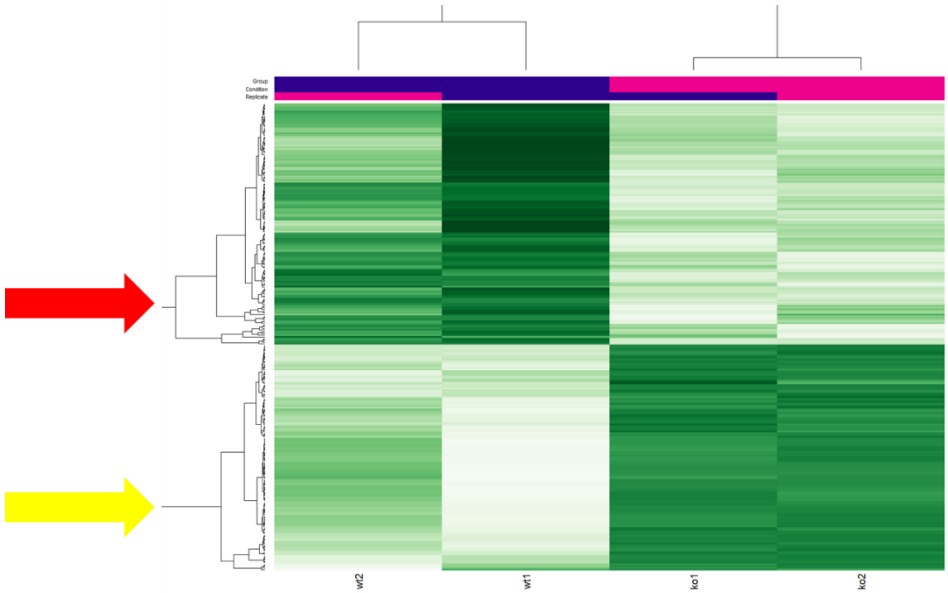

Figure 12 Binding affinity heatmap showing affinities for differentially bound sites. Sample clustering has been performed based on the condition. Two district clusters of binding sites show the enriched binding affinity of WT (red arrow) versus Satb1 cKO (yellow arrow) H3K27ac samples.

The red arrow indicates the regions that are enriched in the WT samples whereas the yellow those who are enriched in the KO samples. The fold-enriched peaks in each condition were separated in order to perform gene enrichment analysis with gProfiler.

The WT H3K27ac samples present enrichment in positive regulation of proteolysis and catabolic processes. What was more interesting was the enrichment in alpha-beta T cell differentiation and V(D)J recombination genes. These findings indicate that the absence of SATB1 leads to reduced deposition of H3K27ac and subsequently reduced activation of genomic loci that have a fundamental role in the proper early T cell development.

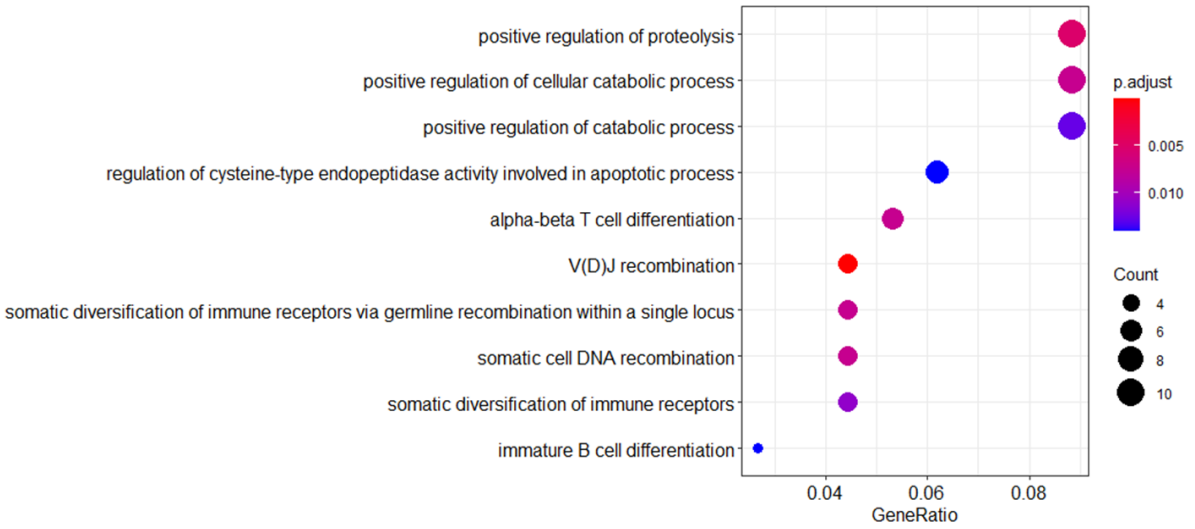

Figure 13 Dotplot of the downregulated gene sets in the absence of SATB1.

Defective alpha-beta T cell differentiation delineates the improper differentiation and maturation of double positive T cells toward the CD4 and CD8 T cell lineages. These findings possibly indicate problems in the development of αβ T cell population in the thymus of *Satb1* cKO mice. These results are consistent with previous FACS analysis experiments carried out in the host lab that depict the decreased expression of CD4 and CD8 T cells in the thymus.

alpha-beta T cell differentiation					
Tcf7	Lef1	Irf4	Il27	Cd83	Tgfbr2
Il4	Zbtb7b	Bcl6	Bcl11b	Gata3	Il18
Stat6	Rorc	Satb1	Bcl2	Irf1	Tox
Stat3	Foxp1	Il23a	Mir21a	Zap70	Runx3
Il12b	Il4ra	Il6	Zbtb16	Tnfsf4	Traj18
Ly9	Il2rg	Smad7	Tnfsf18	Il2	Ccl19

Table 1 Genes that belong to the alpha-beta T cell differentiation GO term

Additionally, the loss of SATB1 leads to decreased deposition of activating H3K27ac in genes involved in V(D)J recombination. This is consistent with previous experiments carried out in

the lab, which have shown the decreased expression of RAG1, RAG2 recombinases (Western Blot analysis) as well as the disturbed morphology of the thymus in *Satb1* cKO mice (immunohistochemistry). Problems in V(D)J recombination will lead to defective *Tcr* recombination and T cell receptor expression. These actions can cause the false positive selection of T cells in the thymus, which may lead to the release of these autoreactive T cells in the periphery.

V(D)J Recombination			
Tcf7	Rag1	Rag2	Dclre1c
Tcf3	Atm	Bcl11b	Lef1
Prkdc	Dcaf1	Lig4	Foxp1

Table 2 Genes that belong to the V(D)J recombination GO term

On the contrary, *Satb1* cKO thymocytes present enrichment in gene sets that are characterizing the Th1, Th2 and Th17 cell differentiation. This means that in the absence of SATB1, there is increased H3K27ac deposition and subsequently activation of genes governing these T cell lineages. These samples are prepared from whole thymus single cell suspensions, which are governed by the presence of Double Positive T cells (~90%).

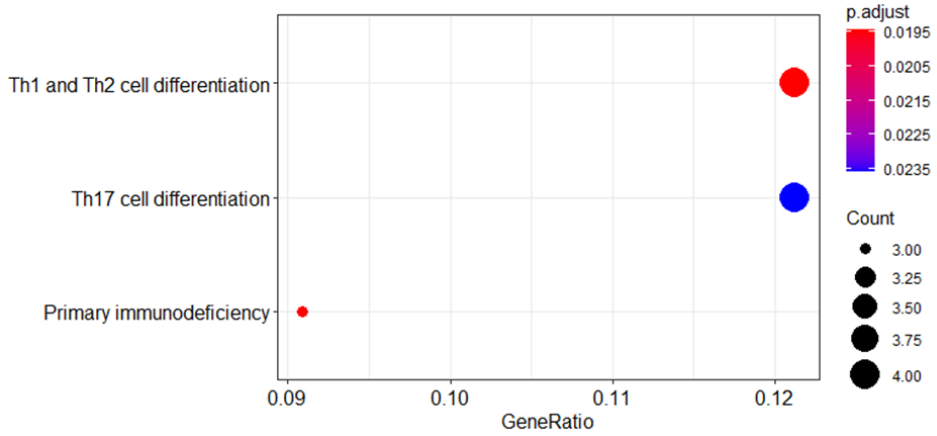

Figure 14 Dotplot of the upregulated gene sets in the absence of SATB1.

Naïve CD4 T cells differentiate to Th1 and Th2 cells by the stimulation of Il12 and Il4 respectively and when activated produce pro-inflammatory cytokines which recruit other cell population of the immune response against parasites and target cells[58].

Th1 Th2 cell differentiation								
Lck	Nfkbie	Jak2	Mapk13	Ifngr1	Il12a	Stat5a	Dll4	Fos
Il4	Tyk2	Rela	Cd247	Ifng	Notch2	Ikbkg	Nfatc2	Ppp3cb
Nfkbib	Nfatc1	Stat1	Dll1	Maf	Ppp3ca	Il12b	Il2	Mapk8
Il2rb	Ppp3r1	Zap70	Gata3	Stat4	Jun	Mapk12	Ikbkb	Ppp3cc
Stat6	Il5	Il2ra	Il13	Mapk3	Nfkb1	Notch3	Runx3	Lat
Jag2	H2-DMb2	Notch1	Stat5b	Mapk1	Ppp3r2	Rbpj	Jak3	Il4ra
Dll3	Ifngr2	Jag1	Nfkbia	H2-Eb2	Jak1	Il12rb2	Nfatc3	Il2rg

Table 3 Genes that belong to the Th1 Th2 cell differentiation GO term

Th17 cells are a Th1 subclass and play important role in the production of pro-inflammatory cytokines. Upon IL23 stimulation, Th17 cells are activated[59] and are capable of secrete (IL)-17, IL-21, IL-22 and IL-26. Most notably, IL-17 has been linked with many pathological conditions and autoimmune diseases[60]. Therefore, in the absence of SATB1 we observe the de-repression of genes that characterize pro-inflammatory T cell lineages and probably *Satb1* cKO thymocyte expression profile resembles that of activated and differentiated T cell lineages.

T17 cell differentiation							
Rela	Lck	Stat6	Il27ra	Stat5b	Il1rap	Il1r1	Jak1
Ikbkb	Il4	Tgfb1	Cd247	Nfkbia	Runx1	Stat1	Mtor
Il23a	Plcg1	Stat3	Tgfbr1	Hif1a	Ifngr2	Zap70	Nfkbib
Il6	Ahr	Stat5a	Gata3	Fos	Cd4	Il2ra	Il2rg
Il17a	Ifngr1	Ikbkg	Nfatc3	Hsp90aa1	Hsp90ab1	Prkcq	Il4ra
Jak2	Il6ra	Smad4	Nfkb1	Irf4	Nfkbie	Il1b	Il21r
Jak3	Rorc	Smad2	Il21	Il6st	Il2	Nfatc2	Lat

Table 4 Genes that belong to the Th17 cell differentiation GO term

- **Disrupted T cell lineage specification and activation control in *SATB1* cKO thymocytes**

  From selective enhancer region studies from our H3K27ac peaks, most notable was the 2.2 normalized fold decrease in the enhancer activity of *Cd8a/b* locus and the 1.85 normalized fold decrease in the enhancer activity of *Cd4* locus showing that possibly the depletion of SATB1 hurdles the proper development and commitment of DP cells toward the CD8 T cell lineage. This observation can be supported by previous FACS analysis and RNA-seq experiments performed in our lab, showing a similar decrease of *Cd8* and *Cd4* expression in the cell membrane and *Cd8, Cd4* mRNA expression respectively.

  CD4 1.85 normalized

  Surprisingly, in a similar context we observed that the enhancer activity of GATA3 was increased in the *Satb1* cKO thymocytes by normalized 2 fold. GATA-3 is a pioneer

transcription factor necessary for the Th2 locus cytokine genes in naïve and activated T cells as well as Th2 cell lineages[61]. Loss of GATA3 leads to defective expression of the Th2 locus cytokines[62]. Therefore, increased enhancer activity in the *Gata3* regulatory regions may lead to aberrant expression of activating cytokines and premature activation of T cells.

The cytokine receptor IL2Ra is another gene that displayed normalized 1.4-fold increased H3K27ac deposition in *Satb1* cKO thymocytes. SATB1 is known to regulate the *Il2Ra* locus through the creation of short range loops[51] and its absence leads to increased expression in DP T cells as shown by FACS analysis[50]. Our data highlight that this ectopic expression due to loss of SATB1 is caused by increased enhancer-promoter communication, due to possible defects in the recruitment of repressor complexes and delineate the important role of SATB1 in T cell activation control. Tiled CRISPR activation experiments have delineated the existence of an *Il2Ra* enhancer that contains an autoimmunity variant which leads to temporally altered expression of *Il2ra* expression without extracellular signals[63]. Possibly loss of SATB1 affects the regulation of *il2ra* enhancer locus in order to restrain its ectopic expression.

Additionally, we observed increased H3K27ac deposition in CD27 regulatory regions in *Satb1* cKO thymocytes by 1.5- normalized fold. CD27 is an essential molecule for CD4 and CD8 T cell function as its loss led to impaired response of these cells to viral infections[64].

Finally, *Ifnγ* receptor 1 had also increased H3K27ac deposition in the absence of SATB1 (2.46 normalized fold over WT cells). IFNγR1 is a pro-Th1 receptor that antagonizes the pro-Th2 receptor IL4R[65] promoting the Th1 lineage. Upon their activation, Th1 cells produce IFNγ and IL2 to activate macrophages and cytotoxic CD8 T cells in order to combat intracellular pathogens[66]. However, these actions can lead to inflammation and autoimmune disease-like features in combination with the ability of Th1 cells to promote IgG2a production from B cells[67].

- **<u>Disturbed gene regulation in *SATB1* cKO thymocytes</u>**

  An important remodeler, *Ino80d* presents a 1.2 normalized fold decrease in H3K27ac deposition in the *Satb1* cKO thymocytes. *Ino80d* belongs to the *Ino80* chromatin remodelling complex and is crucial for the transcriptional gene activation. Its recruitment on DNA is mediated by the chromatin remodeler YY1 and the decrease of its activation may lead to impaired remodelling of regulatory loci needed for Pol-II complex recruitment and transcriptional activation. Histone demethylase Phf2 presented a 1.22 normalized fold decrease of its H3K27ac deposition in the *Satb1* cKO thymocytes. This demethylase has

been found to erase the methylation pattern of H3K9 and H3K4 residues. So, the ablation of SATB1 leads to decreased expression and defective gene regulation, as the deposition of repressive histone marks controls the spatiotemporal gene expression.

- **Loss of cell adhesion and proper T cell maturation in *SATB1* cKO thymocytes**

  Satb1 cKO ChIP-seq samples presented a 1.53 normalized fold decrease in H3K27ac deposition in the regulatory regions of several cell adhesion and cell to cell interaction molecules such as *thy1* (cd90), *pvr* (CD155), *ceacam20* (CEA Cell Adhesion Molecule 20), IGSF23 (immunoglobulin superfamily) whose function involves T cell positive selection, cell adhesion to the thymic epithelium and signal transduction. These findings indicate the abnormal development and maturation of T cells in the thymus upon SATB1 depletion and how its role is necessary for the proper specification, commitment and function of T cells. As intra-thymic interactions of thymocytes with thymic epithelial cells is crucial for their positive selection and maturation, decreased expression of these molecules may lead to false positive selection of autoreactive T cells in the thymus which can exit into the periphery and confer problems analogous to autoimmune disease-related features. These results can support thymus tissue section staining's presenting the loose communication between cell populations in the thymus in the absence of SATB1, where cells are more distant from each other delineating the loss of cell maturation integrity.

DISCUSSION

The past three decades have been crucial in unravelling the role of SATB1 in gene expression regulation with emphasis in immunity, cancer development and neuronal activity. Many research papers have focused on SATB1's role in mammalian physiology and development and how it affects the short- and long-range chromatin interactions of specific genomic loci. Although the effect of SATB1 on the gene expression profile of several cell types has been extensively studied it still remains elusive how SATB1 mechanistically shapes the expression pattern of T cells. Previous studies have supported its role in mediating short-range interactions in the Th2 gene locus as its loss leads to altered chromosomal conformation, but the mechanism by which it facilitates enhancer-promoter communication and whether SATB1 is crucial for the formation of T cell-specific enhancers remain elusive.

Throughout our investigation we tried to focus on the role of SATB1 in regulating the function of major transcription factors, as it is probably guiding them in their target sites and therefore how the absence of SATB1 can lead to impaired gene expression through disrupted recruitment of pioneer factors and affecting enhancer-promoter communication. Previous experiments performed in the host lab have indicated the possible interaction of SATB1 with major transcription factors and complexes (i.e p300, SET1A, the Compass Complex, YY1) and research carried out by other labs have delineated its interaction and co-operation with chromatin remodelling complexes and chromatin organizers[51]. Utilizing co-immunoprecipitation experiments we validated the interaction of SATB1 with p300, a coactivator responsible for regulation of gene expression. p300 protein seems to form a subnuclear cage-like pattern of localization, similar to that of the SATB1 protein with extensive colocalization between them. Interestingly, loss of SATB1 leads to a disturbed p300 pattern in the nucleus delineating its important role in its proper localization. We carried out ChIP experiments for p300 in WT and *Satb1* cKO thymocytes but we couldn't generate legitimate DNA libraries to provide valid sequencing results of its binding profile in the mouse thymocyte genome. Transcription factor ChIP experiments in thymocytes present difficulties due to the extremely compact state of their chromatin. By performing a 10-minute crosslinking with 1% PFA we weren't able to capture cofactor molecules. In the same setup H3K27ac ChIP samples were generated successfully as histones are tightly bound on DNA and can be easily detected and captured using specific antibodies. In order to overcome this obstacle in the future, we will perform Cut&Tag experiments against protein factors and histone marks between WT and *Satb1* cKO thymocytes. Cut&Tag is a novel technique which utilizes the power of a Protein A-

Transposase 5 fusion protein[68]. Protein A is capable to identify with high fidelity the Fc region of any used antibody in order to guide the fusion protein upon the target epitopes. Afterwards Tn5 will mark the flanking DNA regions of interest and will insert specific adapters recognized by Illumina sequencing machines. Using this approach, we aim to purify specific DNA fragments of transcription factors binding sites in order to perform sequencing and obtain the thymus-specific chromatin binding pattern of pioneer factors in the presence and absence of SATB1.

If our speculations are correct then possibly SATB1 recruits several pioneer protein factors upon their target sites in order for chromatin conformation to change, become more open for transcription factor complexes to bind and transcription to initiate. We believe that SATB1 plays a multitasking role in the regulation of T cell-specific gene expression patterns from the activation and aid of several transcription factors to the guidance of chromatin remodelers and organizers to their specific sites. Previous experiments in the lab have indicated the more compact state of chromatin in the absence of SATB1, meaning that its loss leads to disruption of the 3D chromatin structure in the nucleus. Loss of the integrity of the specific genomic interaction network leads to disrupted and even abnormal gene expression delineating the importance of SATB1 in maintaining the 3D nuclear structure of thymocytes.

In order to characterize T cell specific enhancers and elucidate the role of SATB1 in T cell-specific enhancers formation we performed H3K27ac ChIP-seq experiments. We observed that the overall H3K27ac levels remain the same in the absence of SATB1. H3K27ac presents a similar cage-like pattern as SATB1 does and they have extensive colocalization.

From our differential binding analysis, we observed the downregulation of gene sets important for alpha-beta T cell differentiation, V(D)J recombination and T cell receptor signalling pathway. These findings indicate the disturbed development and maturation of thymic populations in the thymus. *Satb1* cKO thymocytes have problems in committing to specific T cell lineages, have defective Tcr locus recombination and cannot receive the appropriate signals from their environment in order to be activated and multiply. On the other hand, *Satb1* cKO thymocytes have increased H3K27ac deposition in genes responsible for Th1-Th2 cell differentiation as well as Th17 cell differentiation. These observations highlight the pro-inflammatory tendency of these cells as these genes are destined to be expressed in later developmental stages from activated T cells. The upregulation of pro-inflammatory molecules will give false signals to the DP T cells which may undergo differentiation or activation without having received the appropriate signals from their environment in the appropriate timepoint.

The cells used for our assays, which were thymocytes, predominantly consisted of Double Positive T cells (~90% of all thymic population), so it is unusual to see pro-inflammatory molecules and transcription factors responsible for T cell activation, lineage specification and stimulation being upregulated in Satb1 cKO T cells.

Putting together the aforementioned results about how loss of SATB1 leads to improper T cell development and maturation and increase in expression of molecules destined to be activated in later T cell developmental stages, we assume that SATB1 controls the spatiotemporal regulation of the T cell genome in order for each developmental stage to be governed by the expression of the proper gene sets and restrain the expression of genes at their specific time point in order to avoid aberrant activation and deficient maturation. As we have seen, SATB1 has both activating and repressive roles in thymocytes as its absence leads to both up- and down-regulation in specific genes and regulatory elements activity.

Therefore, SATB1's divergent role might lie in the licensing of the expression of key T cell gene sets at the correct developmental stage via the regulation of several transcription factors and chromatin organizers. This has been partially shown in previous studies, in which SATB1 regulates the formation of the enhancer landscape in Tregs at an early developmental stage, before FOXP3 expression[69]. Surprisingly, they found that FOXP3 itself represses the genomic locus of *Satb1* in Tregs, as its ectopic expression leads to the conversion of Tregs to effector T cells. This result highlights the important role of SATB1 in controlling and guiding gene regulation of specific T cell lineages.

To summarize, a big amount of work has delineated SATB1's important role in gene regulation and chromatin organization in several tissues and biological processes. They have decoded many important aspects of its function and activity as well as how important SATB1 is for key developmental and regulatory processes in the thymus, neural tissue and ESCs, although there are still many aspects of its function that haven't been elucidated. Of great importance is its strong association with RNA molecules, as depletion of membranes and DNA do not lead to a disturbed subnuclear pattern of protein localization. SATB1's cage-like pattern is disrupted only in the presence of RNase A and this is a feature of great importance for its function.

In recent years the field of enhancer RNAs and enhancer long non-coding RNAs has attracted the attention of many research groups. Previous studies have identified enhancer RNAs (eRNAs) gene transcription as a crucial process for enhancer activation and mediation of transcriptional activation[70] which can be further supported with genome-wide analyses of RNA Pol II occupancy and active transcription in many animal models [71–76]. eRNAs are able to activate transcription by recruiting transcription factors and coactivators changing the

conformation of specific chromatin regions[77–84] Many long non-coding RNAs have been found to interact with either repressive complexes such as the PRC2, or activating complexes such as the COMPASS/SET1A complex and genome organizers such as the Cohesins and CTCF[85]. So, this knowledge creates a new question about the interplay of SATB1 with several eRNAs and enhancer long non-coding RNAs (elncRNAs) and how their cooperation facilitates gene regulation. Previous RNA-seq experiments carried out in the host lab have uncovered a set of elncRNAs being deregulated in the absence of SATB1 and their underlying enhancers controling the expression of immune-related genes of high importance for T cell development and immune response. Therefore, it is worth exploring how SATB1 and other pioneer factors interact with elncRNAs in the context of proper regulation of gene expression in the 3D nucleus.

As we mentioned in the results section, we observed a decrease in the enhancer activity of the *Cd8* locus in the absence of SATB1. Coming to support this observation, previous FACS analysis experiments, carried out in the host lab have shown that the CD8 T cell population in the thymus is decreased by 3-fold in the absence of SATB1 and increases in the periphery (spleen, lymph nodes). These facts strengthen the notion that SATB1 acts as a licensing factor during T cell development, maturation and activity of many T cell lineages. We have also identified an enhancer long non-coding RNA produced by the *Cd8* gene locus superenhancer, whose expression decreases dramatically in *Satb1* cKO thymocytes, as indicated by RNA-seq experiments. Therefore, it is worth investigating the role of SATB1 in the regulation of this elncRNA expression and how SATB1 shapes the chromatin ladscape and expression profile of CD8 T cells.

We believe that SATB1 controls a specific protein-RNA network whose role is to control proper gene regulation and chromatin conformation of thymocytes and other T cell lineages in order to maintain the proper cell maturation, development and function. The proposed role for SATB1 is the spatiotemporal regulation of immune-related gene sets in order to establish a controlled expression pattern that leads to the perfect development of many cell lineages and the proper function of the immune system.

BIBLIOGRAPHY

1. Khorasanizadeh, S. The Nucleosome: From Genomic Organization to Genomic Regulation. *Cell* **116**, 259–272 (2004).
2. Smith, C. L. & Peterson, C. L. ATP-Dependent Chromatin Remodeling. *Curr. Top. Dev. Biol.* **65**, 115–148 (2004).
3. Mellor, J. The dynamics of chromatin remodeling at promoters. *Molecular Cell* **19**, 147–157 (2005).
4. Workman, J. L. Nucleosome displacement in transcription. *Genes and Development* **20**, 2009–2017 (2006).
5. Li, B., Carey, M. & Workman, J. L. The Role of Chromatin during Transcription. *Cell* **128**, 707–719 (2007).
6. Fahrner, J. A. & Baylin, S. B. Heterochromatin: Stable and unstable invasions at home and abroad. *Genes and Development* **17**, 1805–1812 (2003).
7. Wallace, J. A. & Orr-Weaver, T. L. Replication of heterochromatin: Insights into mechanisms of epigenetic inheritance. *Chromosoma* **114**, 389–402 (2005).
8. Groth, A., Rocha, W., Verreault, A. & Almouzni, G. Chromatin Challenges during DNA Replication and Repair. *Cell* **128**, 721–733 (2007).
9. Bajoghli, B. *et al.* Evolution of Genetic Networks Underlying the Emergence of Thymopoiesis in Vertebrates. *Cell* **138**, 186–197 (2009).
10. Boehm, T. & Bleul, C. C. The evolutionary history of lymphoid organs. *Nat. Immunol.* **8**, 131–135 (2007).
11. Zapata, A. & Amemiya, C. T. Phylogeny of lower vertebrates and their immunological structures. *Current Topics in Microbiology and Immunology* **248**, 67–107 (2000).
12. Miracle, A. L. *et al.* Complex expression patterns of lymphocyte-specific gene during the development of cartilaginous fish implicate unique lymphoid tissues in generating an immune repertoire. *Int. Immunol.* **13**, 567–580 (2001).
13. Hirano, M. *et al.* Evolutionary implications of a third lymphocyte lineage in lampreys. *Nature* **501**, 435–438 (2013).
14. Takahama, Y. Journey through the thymus: Stromal guides for T-cell development and selection. *Nature Reviews Immunology* **6**, 127–135 (2006).
15. Petrie, H. T. & Zúñiga-Pflücker, J. C. Zoned out: functional mapping of stromal signaling microenvironments in the thymus. *Annual Review of Immunology* **25**, 649–679 (2007).

16. Love, P. E. & Bhandoola, A. Signal integration and crosstalk during thymocyte migration and emigration. *Nature Reviews Immunology* **11**, 469–477 (2011).
17. Calderón, L. & Boehm, T. Synergistic, context-dependent, and hierarchical functions of epithelial components in thymic microenvironments. *Cell* **149**, 159–172 (2012).
18. Buono, M. *et al.* A dynamic niche provides Kit ligand in a stage-specific manner to the earliest thymocyte progenitors. *Nat. Cell Biol.* **18**, 157–167 (2016).
19. Zamisch, M. *et al.* Ontogeny and Regulation of IL-7-Expressing Thymic Epithelial Cells. *J. Immunol.* **174**, 60–67 (2005).
20. Rothenberg, E. V. Programming for T-lymphocyte fates: Modularity and mechanisms. *Genes Dev.* **33**, 1117–1135 (2019).
21. Hu, G. *et al.* Transformation of Accessible Chromatin and 3D Nucleome Underlies Lineage Commitment of Early T Cells. *Immunity* **48**, 227-242.e8 (2018).
22. Isoda, T. *et al.* Non-coding Transcription Instructs Chromatin Folding and Compartmentalization to Dictate Enhancer-Promoter Communication and T Cell Fate. *Cell* **171**, 103-119.e18 (2017).
23. Rawlings, J. S., Gatzka, M., Thomas, P. G. & Ihle, J. N. Chromatin condensation via the condensin II complex is required for peripheral T-cell quiescence. *EMBO J.* **30**, 263–276 (2011).
24. Felsenfeld, G. & Dekker, J. Genome architecture and expression. *Curr. Opin. Genet. Dev.* **22**, 59–61 (2012).
25. De Laat, W. & Duboule, D. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* **502**, 499–506 (2013).
26. Symmons, O. *et al.* Functional and topological characteristics of mammalian regulatory domains. *Genome Res.* **24**, 390–400 (2014).
27. Stadhouders, R. *et al.* Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. *Nat. Genet.* **50**, 238–249 (2018).
28. Bonev, B. *et al.* Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell* **171**, 557-572.e24 (2017).
29. Dixon, J. R. *et al.* Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**, 331–336 (2015).
30. Krijger, P. H. L. *et al.* Cell-of-origin-specific 3D genome structure acquired during somatic cell reprogramming. *Cell Stem Cell* **18**, 597–610 (2016).
31. Pękowska, A. *et al.* Gain of CTCF-Anchored Chromatin Loops Marks the Exit from

Naive Pluripotency. *Cell Syst.* **7**, 482-495.e10 (2018).

32. Johanson, T. M. *et al.* Transcription-factor-mediated supervision of global genome architecture maintains B cell identity. *Nat. Immunol.* **19**, 1257–1264 (2018).
33. Nora, E. P. *et al.* Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* **169**, 930-944.e22 (2017).
34. Schwarzer, W. *et al.* Two independent modes of chromatin organization revealed by cohesin removal. *Nature* **551**, 51–56 (2017).
35. Haarhuis, J. H. I. *et al.* The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell* **169**, 693-707.e14 (2017).
36. Wang, S. *et al.* Spatial organization of chromatin domains and compartments in single chromosomes. *Science (80-.).* **353**, 598–602 (2016).
37. Boija, A. *et al.* Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell* **175**, 1842-1855.e16 (2018).
38. Chong, S. *et al.* Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science (80-.).* **361**, (2018).
39. Plank, J. L. & Dean, A. Enhancer function: Mechanistic and genome-wide insights come together. *Molecular Cell* **55**, 5–14 (2014).
40. Pombo, A. & Dillon, N. Three-dimensional genome architecture: Players and mechanisms. *Nature Reviews Molecular Cell Biology* **16**, 245–257 (2015).
41. Long, H. K., Prescott, S. L. & Wysocka, J. Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. *Cell* **167**, 1170–1187 (2016).
42. Furlong, E. E. M. & Levine, M. Developmental enhancers and chromosome topology. *Science* **361**, 1341–1345 (2018).
43. Wutz, G. *et al.* Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J.* **36**, 3573–3599 (2017).
44. Weintraub, A. S. *et al.* YY1 Is a Structural Regulator of Enhancer-Promoter Loops. *Cell* **171**, 1573-1588.e28 (2017).
45. Deng, W. *et al.* Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* **149**, 1233–1244 (2012).
46. Papantonis, A. *et al.* Active RNA Polymerases: Mobile or Immobile Molecular Machines? *PLoS Biol.* **8**, e1000419 (2010).

47. Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
48. Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
49. Hnisz, D., Day, D. S. & Young, R. A. Insulated Neighborhoods: Structural and Functional Units of Mammalian Gene Control. *Cell* **167**, 1188–1200 (2016).
50. Alvarez, J. D. *et al.* The MAR-binding protein SATB1 orchestrates temporal and spatial expression of multiple genes during T-cell development. *Genes Dev.* **14**, 521–535 (2000).
51. Yasui, D., Miyano, M., Cai, S., Varga-Weisz, P. & Kohwi-Shigematsu, T. SATB1 targets chromatin remodelling to regulate genes over long distances. *Nature* **419**, 641–645 (2002).
52. Cai, S., Han, H. J. & Kohwi-Shigematsu, T. Tissue-specific nuclear architecture and gene expession regulated by SATB1. *Nat. Genet.* **34**, 42–51 (2003).
53. Cai, S., Lee, C. C. & Kohwi-Shigematsu, T. SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. *Nat. Genet.* **38**, 1278–1288 (2006).
54. Pavan Kumar, P. *et al.* Phosphorylation of SATB1, a Global Gene Regulator, Acts as a Molecular Switch Regulating Its Transcriptional Activity In Vivo. *Mol. Cell* **22**, 231–243 (2006).
55. Beyer, M. *et al.* Repression of the genome organizer SATB1 in regulatory T cells is required for suppressive function and inhibition of effector differentiation. *Nat. Immunol.* **12**, 898–907 (2011).
56. Kitagawa, Y. *et al.* Guidance of regulatory T cell development by Satb1-dependent super-enhancer establishment. *Nat. Immunol.* **18**, 173–183 (2017).
57. Ghosh, R. P. *et al.* Satb1 integrates DNA binding site geometry and torsional stress to differentially target nucleosome-dense regions. *Nat. Commun.* **10**, (2019).
58. Amsen, D., Spilianakis, C. G. & Flavell, R. A. How are TH1 and TH2 effector cells made? *Current Opinion in Immunology* **21**, 153–160 (2009).
59. Aggarwal, S., Ghilardi, N., Xie, M. H., De Sauvage, F. J. & Gurney, A. L. Interleukin-23 promotes a distinct CD4 T cell activation state characterized by the production of interleukin-17. *J. Biol. Chem.* **278**, 1910–1914 (2003).
60. Shuttleworth, S. *et al.* Progress in the development of small molecule therapeutics targeting Th17 cell function for the treatment of immune-inflammatory diseases. in

Progress in Medicinal Chemistry **50**, 109–133 (Elsevier B.V., 2011).

61. Zheng, W. P. & Flavell, R. A. The transcription factor GATA-3 is necessary and sufficient for Th2 cytokine gene expression in CD4 T cells. *Cell* **89**, 587–596 (1997).
62. Pai, S. Y., Truitt, M. L. & Ho, I. C. GATA-3 deficiency abrogates the development and maintenance of T helper type 2 cells. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 1993–1998 (2004).
63. Simeonov, D. R. *et al.* Discovery of stimulation-responsive immune enhancers with CRISPR activation. *Nature* **549**, 111–115 (2017).
64. Hendriks, J. *et al.* CD27 is required for generation and long-term maintenance of T cell immunity. *Nat. Immunol.* **1**, 433–440 (2000).
65. Mowen, K. A. & Glimcher, L. H. Signaling pathways in Th2 development. *Immunological Reviews* **202**, 203–222 (2004).
66. Liew, F. Y. TH1 and TH2 cells: A historical perspective. *Nature Reviews Immunology* **2**, 55–60 (2002).
67. Stevens, T. L. *et al.* Regulation of antibody isotype secretion by subsets of antigen-specific helper T cells. *Nature* **334**, 255–258 (1988).
68. Kaya-Okur, H. S. *et al.* CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.* **10**, 1–10 (2019).
69. Kitagawa, Y. *et al.* Guidance of regulatory T cell development by Satb1-dependent super-enhancer establishment. *Nat. Immunol.* **18**, 173–183 (2017).
70. Tuan, D., Kong, S. & Hu, K. Transcription of the hypersensitive site HS2 enhancer in erythroid cells. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 11219–11223 (1992).
71. Kim, Y. J., Xie, P., Cao, L., Zhang, M. Q. & Kim, T. H. Global transcriptional activity dynamics reveal functional enhancer RNAs. *Genome Res.* **28**, 1799–1811 (2018).
72. Djebali, S. *et al.* Landscape of transcription in human cells. *Nature* **489**, 101–108 (2012).
73. De Santa, F. *et al.* A Large Fraction of Extragenic RNA Pol II Transcription Sites Overlap Enhancers. *PLoS Biol.* **8**, e1000384 (2010).
74. Kharchenko, P. V. *et al.* Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. *Nature* **471**, 480–486 (2011).
75. Bonn, S. *et al.* Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat. Genet.* **44**, 148–156 (2012).
76. Chen, R. A. J. *et al.* The landscape of RNA polymerase II transcription initiation in C.

elegans reveals promoter and enhancer architectures. *Genome Res.* **23**, 1339–1347 (2013).

77. Sigova, A. A. *et al.* Transcription factor trapping by RNA in gene regulatory elements. *Science (80-.).* **350**, 978–991 (2015).
78. Kaikkonen, M. U. *et al.* Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. *Mol. Cell* **51**, 310–325 (2013).
79. Lai, F., Gardini, A., Zhang, A. & Shiekhattar, R. Integrator mediates the biogenesis of enhancer RNAs. *Nature* **525**, 399–403 (2015).
80. Gardini, A. *et al.* Integrator regulates transcriptional initiation and pause release following activation. *Mol. Cell* **56**, 128–139 (2014).
81. Bose, D. A. *et al.* RNA Binding to CBP Stimulates Histone Acetylation and Transcription. *Cell* **168**, 135-149.e22 (2017).
82. Li, W. *et al.* Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* **498**, 516–520 (2013).
83. Hsieh, T. H. S. *et al.* Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. *Cell* **162**, 108–119 (2015).
84. Isoda, T. *et al.* Non-coding Transcription Instructs Chromatin Folding and Compartmentalization to Dictate Enhancer-Promoter Communication and T Cell Fate. *Cell* **171**, 103-119.e18 (2017).
85. Gil, N. & Ulitsky, I. Regulation of gene expression by cis-acting long non-coding RNAs. *Nat. Rev. Genet.* **21**, 102–117 (2020).