ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΔΙΑΤΜΗΜΑΤΙΚΟ ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
«ΜΟΡΙΑΚΗ ΒΙΟΛΟΓΙΑ & ΒΙΟΪΑΤΡΙΚΗ»

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

Τελική Αναφορά
Μεταπτυχιακής Ερευνητικής Εργασίας

Distinct roles of the Special AT-rich sequence-Binding protein-1 during T-cell development

Παναγόπουλος Γεώργιος
Μεταπτυχιακός φοιτητής Διατμηματικού Προγράμματος
«Μοριακή Βιολογία & Βιοϊατρική»

Επιβλέπων Καθηγητής/Επιστημονικός Υπεύθυνος:
Καθ. Σπηλιανάκης Χαράλαμπος

Ηράκλειο, Σεπτέμβριος 2014

Θα ήθελα καταρχήν να ευχαριστήσω όλους όσους συνέβαλαν με οποιονδήποτε τρόπο στην επιτυχή εκπόνηση αυτής της διπλωματικής εργασίας.

Ευχαριστώ θερμά τον Επίκουρο καθηγητή κ. Χαράλαμπο Σπηλιανάκη για την επίβλεψη αυτής της διπλωματικής εργασίας και για την ευκαιρία που μου έδωσε να την εκπονήσω στο εργαστήριο γονιδιακής ρύθμισης στο ανοσοποιητικό σύστημα. Ήταν πάντα διαθέσιμος να μου προσφέρει τις γνώσεις και την εμπειρία του καθώς και να με καθοδηγήσει τόσο σε επιστημονικό όσο και σε προσωπικό επίπεδο. Θα του είμαι πάντα ευγνώμων για την εμπιστοσύνη που έδειξε στο πρόσωπό μου και την αμέριστη στήριξη που μου παρείχε καθ' όλη την παραμονή μου στο Ινστιτούτο. Ήταν τιμή μου που γνώρισα έναν τέτοιο άνθρωπο και επιστήμονα.

Ευχαριστώ επίσης, τον υποψήφιο Διδάκτορα κ. Πέτρο Τζέρπο για την απεριόριστη βοήθεια και τις συμβουλές του κατά τη διάρκεια των πειραματικών διαδικασιών. Τον ευχαριστώ θερμά για τις ιδέες που μου προσέφερε, για όλες τις ερευνητικές συζητήσεις που πραγματοποιήθηκαν καθ΄όλη τη διάρκεια εκπόνησης αυτής της εργασίας και ιδιαίτερα για την υπομονή του. Ήταν εκείνος που με έκανε να σκέφτομαι ως ερευνητής και να επιζητώ πάντα το καλύτερο. Αποτέλεσε και θα είναι πάντα πρότυπο για εμένα.

Θα ήθελα να πω ένα μεγάλο ευχαριστώ στην υποψήφια Διδάκτορα Χρύσα Σταθοπούλου για τις αμέτρητες ευχάριστες στιγμές που μοιραστήκαμε στο εργαστήριο, την καθημερινή της στήριξη εντός και εκτός εργασίας. Ήταν σαν μια δεύτερη αδερφή μου στην Κρήτη και νιώθω ευτυχής που τη γνώρισα.

Ευχαριστώ τα υπόλοιπα μέλη του εργαστηρίου Μανουέλα Καψετάκη, Καλλίνα Στρατήγη, Άντα Σαλατάι, Ευαγγελία Ταχματζίδη, Μαρία Στρατηγοπούλου, Νεφέλη Ντελεπιάνε και Δήμητρα Νικολέρη για την ευχάριστη συνεργασία και τις πολύτιμες συμβουλές τους κατά τη διάρκεια παραμονής μου στο εργαστήριο.

Επιπλέον, θα ήθελα να ευχαριστήσω τα υπόλοιπα μέλη της τριμελούς επιτροπής, τον καθηγητή κ. Ιωσήφ Παπαματθαιάκη και τον Επίκουρο καθηγητή κ. Χριστόφορο Νικολάου για τις συμβουλές και την καθοδήγησή τους ως προς την επιτυχή έκβαση των πειραματικών διαδικασιών. Ακόμα, ευχαριστώ τους Συνεργαζόμενους Ερευνητές του Ι.Τ.Ε. κ. Μιχαήλ Αιβαλιώτη και κ. Παύλο Παυλίδη για την τεχνική υποστήριξη και την ανάλυση των πειραματικών αποτελεσμάτων.

Τέλος, ευχαριστώ την οικογένεια και τους φίλους μου που είναι δίπλα μου σε κάθε βήμα της ζωής μου.

Master of Science Thesis

Summary

Special AT-rich sequence-binding protein 1 (SATB1) is a nuclear matrix protein with a central role in higher order chromatin organization and gene regulation. Elevated expression and/or mutation of SATB1 are associated with progression and poor prognosis in a plethora of malignancies. Although highly expressed in T cells, the molecular mechanism of SATB1 function in the thymus has been elusive. We identified protein interaction partners of SATB1 using a co-immunoprecipitation approach coupled with mass spectrometry analysis, using primary murine thymocyte extracts and created a comprehensive protein network map of SATB1 interactions in T cells. Among others, SATB1 interacts with several known transcription factors and kinases involved in long range chromatin interactions (YY1), thymic development, differentiation of naive $CD4^+$ precursorcells (Mi-2β) and in the establishment of T cell tolerance (AIRE). Moreover, several members of specific histone modification complexes involved in gene regulation (COMPASS and NSL complexes) were also found to co-precipitate with SATB1 protein, suggesting an active role in the establishment/maintenance of T cell chromatin structure and gene regulation. Moreover, we optimized the protocol for the chromatin immunoprecipitation of the SATB1 protein in naive $CD4^+$ T cells. Both the precipitated and the input DNA samples were sequenced in an Ion Proton™ sequencer. We established a pipeline for downstream bioinformatics analysis and quality tests of next generation sequencing data for nuclear matrix proteins. A combination of the protein interactors of SATB1 and its binding sites on chromatin, could shed more light on the mechanism underlying the function of SATB1 in T cell development, chromatin conformation and gene regulation via long-range chromatin interactions.

Introduction

In the past decade, research focused in the study of the three-dimensional (3D) organization of the eukaryotic genome as it determines cell fate and play important roles in the regulation of gene expression. Gene expression can be regulated "*in cis*" (intrachromosomally) by interactions between several regulatory elements (enhancers, promoters, locus control region etc.) on the same chromosome or "in trans" (interhromosomally) when the interacting regions localize on different chromosomes. During these genomic interactions, chromatin forms a pre-poised initial core conformation that gives the regulated genes the ability of rapid response to transcription activation signals[1 2].

The interchromosomal TH2-*Ifnγ* loci interactions take place in naïve $CD4^+$ cells and potentially keep the two loci poised for rapid expression upon T cell receptor (TCR) stimulation. During T helper type 1 (TH1) differentiation the *Ifnγ* gene is mainly regulated via intrachromosomal interactions with participating regulatory elements in the locus itself and in TH2 cells intrachromosomal interactions between the TH2-LCR and the cytokine gene promoters upon the action of TH2 cell lineage specific factors (Fig.1)[3].

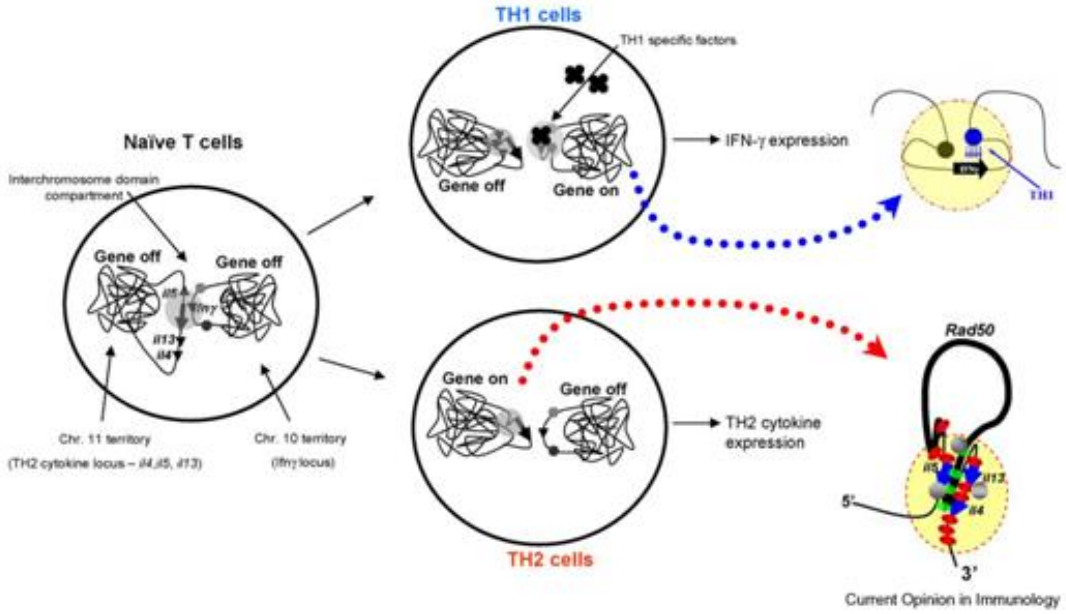


Figure 1. TH2 locus-Ifnγ long-range interactions during naive $CD4^+$ T cell differentiation to TH1 and TH2 cell lineages.

Among other proteins, SATB1 (special AT-rich sequence-binding protein-1) participates in the maintenance of chromatin architecture by organizing higher-order chromatin loops in long-range interactions of the TH2, IFNγ and TNF loci[4 5]. It is a nuclear protein, identified as a global chromatin organizer tethering nuclear matrix attachment regions (MARS) that are 70% AT rich. In primary thymocytes SATB1 forms a functional nuclear architecture that has a "cage-like" pattern surrounding heterochromatin [6 7 8 9]. Apart from its architectural role, SATB1 acts as a regulatory

protein scaffold network. Its binding sites are hot spots for epigenetic modifications by various chromatin-modifying enzymes inducing the expression of several genes relevant to T cell maturation[10]. *Satb1* null mice arrest thymocyte development on the double positive stage (DP, $CD4^+CD8^+$) and the expression of cytokine genes is decreased highlighting its importance in T cell development[11]. Additionally, SATB1 is directly connected with the induction of X-inactivation and promoting aberrant growth and metastasis in different types of epithelial tumors [12 13 14].

A highly conserved DNA element located in the TH2 locus control region (TH2-LCR) showed DNAse I hypersensitivity both in naïve $CD4^+$ cells but also in activated lymphocytes and was involved in chromatin associations between interleukin gene loci residing on different chromosomes[15, 16, 17]. During initial experiments in our lab, this genomic region was used in a DNA affinity purification assay, coupled to mass-spectrometry analysis. This analysis resulted in the identification of two binding proteins, namely BACH1 and SATB1 (Fig. 2).

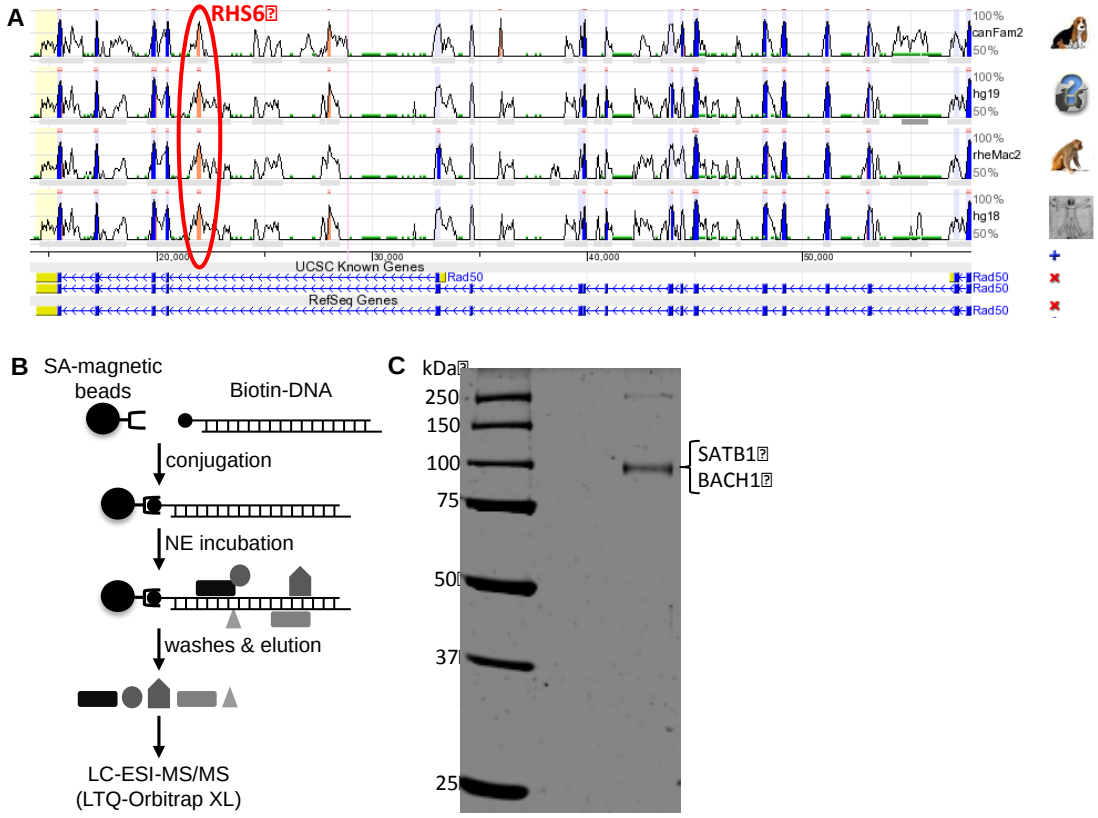


Figure 2: A) A region of 151 bp, part of the RHS6 DNase I hypersensitive site of the TH2-LCR shows 90% sequence conservation between mouse and human species. B) This conserved DNA region was biotinylated and utilized in a DNA affinity purification assay that upon mass spectrometry analysis resulted in the identification of SATB1 and BACH1 proteins (C).

Results

Co-Immunoprecipitation against SATB1 protein

In order to reveal the role of SATB1 during $CD4^+$ T cell differentiation we focused on determining its protein interactors in murine thymocytes. In line with this question we applied an anti-SATB1 co-Immunoprecipitation approach followed by Mass Spectrometry analysis (Fig. 3A) where 505 proteins were identified to interact with SATB1 (Table S1). In comparison with a negative control (IgG), 135 were highly enriched in the immunoprecipitated fraction with the anti-SATB1 antibody, 107 had low enrichment and 263 were enriched in the negative control (Fig. 3B). Among others, SATB1 was found to interact with several transcription factors and kinases known to be involved in tmymus development and establishment of T cell tolerance. In addition, several members of specific histone modification complexes were also coprecipitated with SATB1 suggesting an active role in the establishment/maintenance of the chromatin structure in $CD4^+$ T cells.

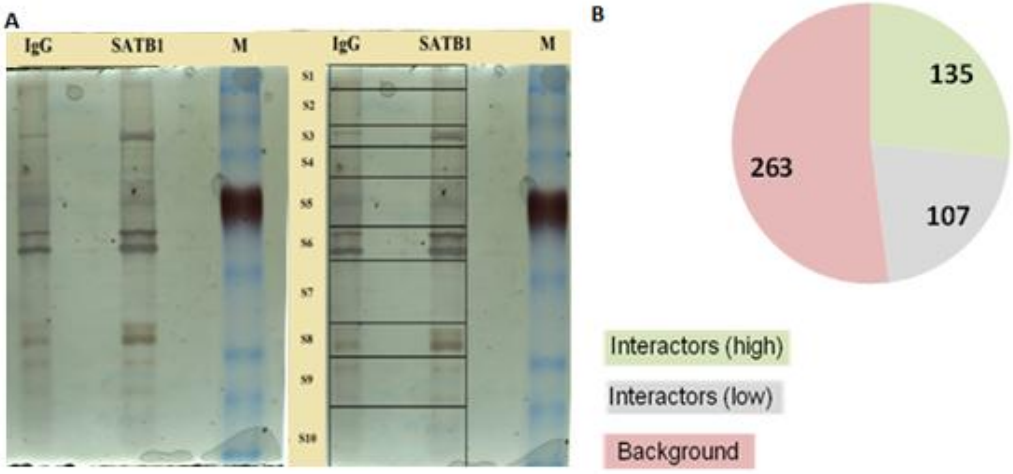


Figure 3. A) Silver stained SDS page (Left). Pairs of bands analysed with mass spectrometry (Right). B) Distribution of 505 proteins identified on a-SATB1 co-Immunoprecipitation assay.

SATB1 in the regulation of histone methylation

Trimethylation of Lysine 4 of Histone 3 (H3K4me3) is the hallmark of gene activation. Mono-, di-, and tri-methylation of H3K4 by a single histone methyltransferase (HMT) was firstly shown by Yeast SET-containing domain (Set1), component of the COMPASS complex (Complex Proteins Associated with Set1)[18 19]. In mammals, COMPASS is composed of SETD1A or SETD1B, WDR5, WDR82, RBBP5, ASH2L/ASH2, CXXC1/CFP1, HCFC1 and DPY30. In our experiments we found that 6 (SET1A, WDR5, RBBP5, ASH2L/ASH2, CXXC1/CFP1 and HCFC1) of these 8 components co-precipitated with the SATB1 protein. Furthermore, we identified MENIN, KAT8(MOF) and KMT2A(MLL1) as SATB1 co-interactors. These three proteins in combination with WDR5, RBBP5, ASH2L/ASH2, HCFC1 and DPY30 are the components of the mammalian COMPASS-like complex MLL1-2. Initially, MLL was

identified in an 11q23 chromosomal translocation that lead to myeloid and lymphoblastic leukemia[20].

MENIN, binds to the N-terminal domain of MLL and acts like a "molecular bridge" linking it to lens epithelium-derived growth factor (LEDGF)[21]. The N-terminal domain of MLL contains transcriptional co-activators, three AT hooks that bind AT-rich sequences, and a CXXC zinc-finger motif that binds non-methylated CG-rich sequences. In this way, MENIN/LEDGF, AT hook and the CXXC zinc-finger domains contribute to the localization of MLL–COMPASS-like complexes on chromatin [22, 23, 24].

SATB1 interacts with YY1 /Ino80 gene activation complex

Both YY1 (Yin-Yang 1) transcription factor and Ino80 DNA helicase co-precipitated with SATB1. YY1 binds to several regions of the TH2 locus, including RHS7, which is essential for intra-chromosomal interactions in this locus (Fig. 4A). For gene activation, YY1 needs to recruit Ino80 where it acts as a coactivator. Moreover, Ino80 is a component of some MLL1/MLL complexes, which, as previously mentioned, activates gene expression via H3K4me3. Knockdown of YY1 increases DNA methylation on the *IL-4* gene and RHS7 while it reduces H3K4-methylation, H3-acetylation in the TH2 locus and its regulatory elements, Th2 cytokine expression, intrachromosomal interactions and resistance to asthma allergy (Fig. 4B,C)[25]. These SATB1 interactors strengthen its role in recruitment of gene activating complexes.

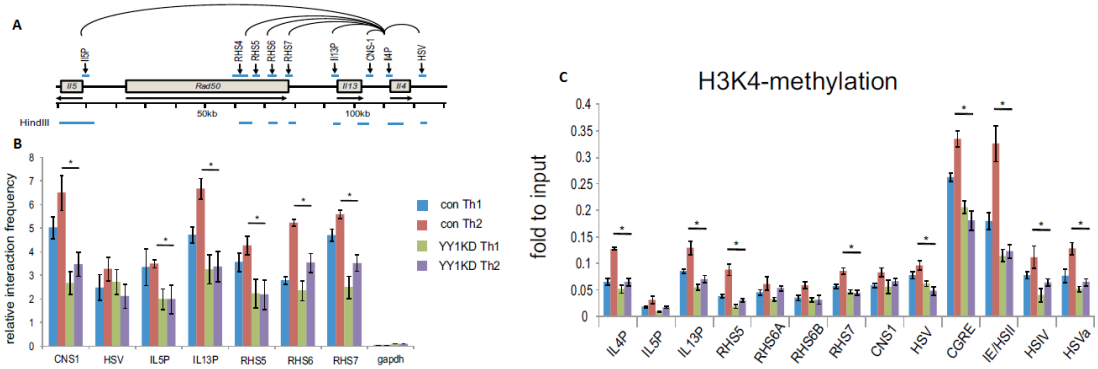


Figure 4. A) Schematic representation of the TH2 locus showing HindIII digestion fragments and their interactions with the *Il4* gene promoter. B) Long-range chromosomal interactions between the *Il4* gene promoter and other chromatin fragments of the Th2 cytokine locus measured by chromatin conformation capture (3C). C) H3K4methylation reduction in YY1 knockdown Th2 cells[25].

SATB1 in regulation of histone acetylation

Non-specific lethal (NSL) HAT complex is composed of MOF/KAT8, KANSL1, KANSL2, KANSL3, MCRS1, PHF20, OGT1/OGT, WDR5 and HCFC1. All proteins, with the exception of the PHF20, co-immunoprecipitated with the SATB1 protein. NSL complex induces acetylation of lysine 16 of Histone 3 (H3K16ac), which promotes gene activation. This complex shares components with other chromatin regulation complexes. The MCRS1 (Microspherule Protein1) is also one of the subunits of Ino80 chromatin remodeling complex, while WDR5 and HCFC1 are core components for the MLL/SET-containing histone methyltransferase complex. Interestingly, both Ino80 and MLL-SET complexes were also coimmunoprecipitated with SATB1. Based on genomic experiments, there is a "crosstalk" between these complexes that leads to a coordination of gene regulation in certain genes[26 27].

SATB1 mediates activation of CD4 expression

In our experiments, Mi-2β chromatin remodeler (CHD4) and the histone acetyltransferase p300 interacted with SATB1 as they co-immunoprecipitated in thymocyte protein extracts. Chromatin Immunoprecipitation experiments have revealed that Mi-2β is directly associated with the proximal CD4 enhancer in CD4-expressing thymocytes. As it was proposed, Mi-2β and p300 recruits HEB to E-box sites on the CD4 enhancer where Mi-2β promotes ATP-dependent chromatin fluidity. P300 with the presence of HEB or Mi-2β in the complex induces histone acetylation that converts chromatin to an open configuration. The activated enhancer interacts with the CD4 gene promoter leading to the stabilization of the transcription factors and the expression of the CD4 gene (Fig. 5). SATB1, could mediate Mi-2β binding to the CD4 enhancer or in forming the enhancer-promoter loop for induction of the CD4 gene expression during T cell development[28].

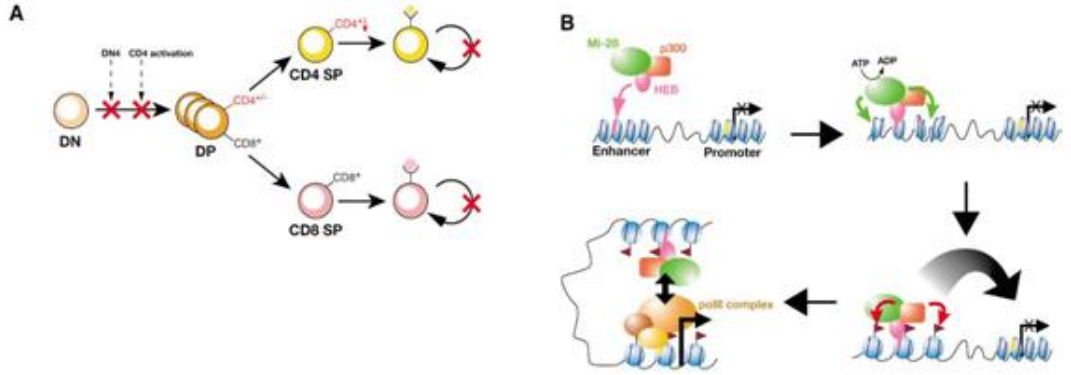


Figure 5. The role of Mi-2 in T cell development and a proposed model for CD4 gene regulation[28].

SATB1 is involved in immune tolerance

The autoimmune regulator protein (AIRE) was co-precipitated with the SATB1 protein in our experiments. AIRE is a transcription factor that controls the negative selection of self-reactive T cells, a procedure that is essential for the immune system to tolerate self-antigens and avoid autoimmunity. Mutations of human AIRE results in a rare multiorgan autoimmune disease, namely autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED)[29 30 31]. The *Aire* gene is expressed mainly in medullary epithelial cells (MECs) and cells of the monocyte-dendritic cell lineage, both cell types representing a population of antigen-presenting cells. It is also expressed in the spleen, fetal liver, and lymph nodes but at lower levels. AIRE's PHD1 domain binds to unmethylated Lysine 4 of Histone 3, a mark of relatively inactive chromatin. It is proposed that, with hypomethylayion of H3 tails, AIRE recognizes silenced genes and induce them to "wake up" transcriptionally, acting like a scaffold for transcription factors recruitment[32 33]. AIRE had previously shown to interact with the transcriptional activator CREB[34] (it is also precipitated with SATB1 protein in our experiments).

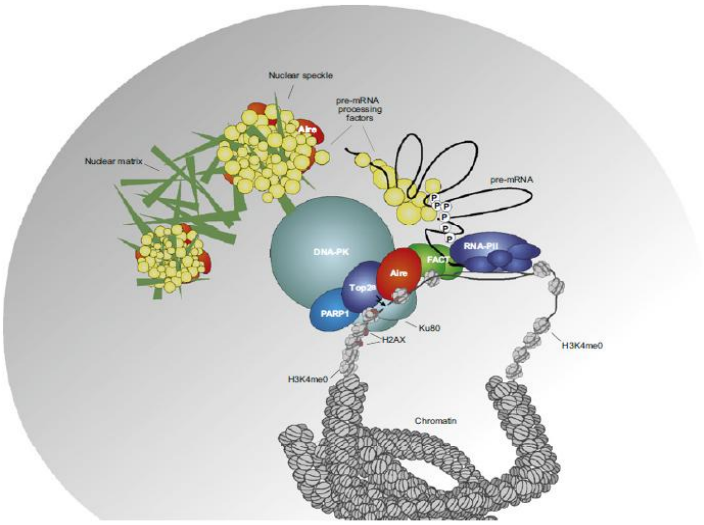


Figure 6. A model for AIRE-mediated regulation of gene expression upon entrance in the nucleus through the nucleopore complex, AIRE preferentially localizes to relatively transcriptionally inert chromatin regions, binding to hypo-methylated H3 tails[33].

Chromatin Immunoprecipitation against SATB1 protein

In order to reveal SATB1 binding sites on the mouse genome we performed chromatin immunoprecipitation assays against this protein followed by next generation sequencing analysis. Firstly, we applied this approach in thymocytes where chromatin is less compact and SATB1 is highly expressed. After 5 minutes of treatment on Covaris S2 sonicator, chromatin was fragmented in approximately 100-800bps (Fig. 7). Qubit™ fluorometer measurements gave 81.6ng of precipitated DNA fragments of the anti-SATB1 immunoprecipitation and 21.9ng of the negative control sample (IgG). Based on qPCR analysis, using primers for the RHS6 of the TH2 locus, the IP/IgG ratio was 21.56 indicating a considerable enrichment for the anti-SATB1 immunoprecipitated fraction..

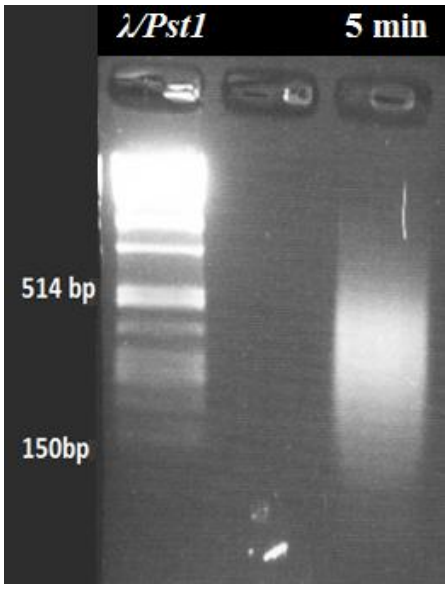


Figure 7. DNA fragmentation of the sonicated and pre-immunoprecipitated DNA from thymocytes, verified by agarose gel electrophoresis.

Consequently, we repeated this assay in naive $CD4^+$ cells. Since chromatin is more compact in this cell type, a 5-minute sonication sheared the chromatin in 100-1000bp. We verified this fragmentation both by agarose gel electrophoresis (Fig. 8A) and bioanalyzer measurements (Fig. 8B). Using Qubit™ fluorometer, we measured

151.2ng of precipitated DNA fragments for the anti-SATB1 immunoprecipitated fraction and 56.88ng of the negative control sample (IgG). Based on the subsequent qPCR analysis, using primers for the RHS6 region of the TH2 locus, the IP/IgG ration was 4.8.

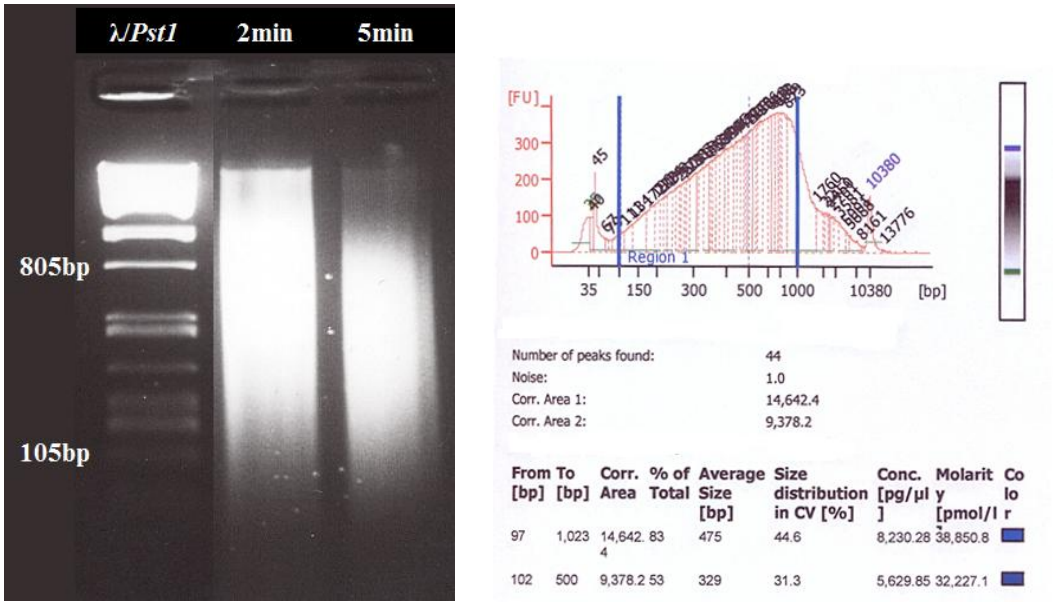


Figure 8. DNA fragmentation of the sonicated and pre-immunoprecipitated DNA from naive $CD4^+$ cells verified by A)agarose gel electrophoresis and B)bioanalyzer.

Sequencing of the precipitated DNA and peak calling

After resuspension of the immunoprecipitated DNA, both the SATB1 IP sample and the Input sample were further sonicated for a second time (4 minutes) to shear the DNA in the appropriate size for constructing the library used in the sequencing analysis (100-300bps). These final DNA fragments were used for library preparation and sequenced by the Ion Proton™ sequencer.

The output of this sequencing was a "fastq" file containing 65,4 millions of reads, free of the adaptors' sequence (Fig. 9). These data were mapped on the mouse genome database (mm9) and visualized in the IGV viewer (Fig. 10). Using a MACS peak caller and changing each time the p-value option, we identified a range of SATB1 binding sites (Table S2A, S2B). To check for statistical significance, we asked how many loci could be identified as SATB1 binding sites using the inverse ratio (Input/IP) in the MACS software (Table 1). Furthermore, based on the forward (IP/Input) peak calling, we designed primers for three SATB1 binding sites. The qPCR analysis using IP and Input samples before library construction could prove that SATB1 enrichment in those regions was indeed due to Immunoprecipitation, and not a library construction artifact. Finally though, none of the three regions was enriched in the qPCR analysis(IP/Input~1).

Matrix Attachment Regions and Chromosomal distribution

Furthermore, we wanted to check the distribution of SATB1 peaks in chromosomes and DNA elements (promoters, enhancers, genes, 5'UTRs, 3'UTRs etc.) in order to investigate if there is any preference for SATB1 binding using the "CEAS: Enrichment on chromosome and annotation" option of the Galaxy/Cistrome online

tool. Moreover, as SATB1 binds to Matrix attachment regions (MARS) we tried to identify MAR sites in SATB1 peaks and in the mouse genome in general. For MARS identification we used the emboss marscan online tool (http://emboss.bioinformatics.nl/cgi-bin/emboss/marscan). In addition, we used the CEAS online tool to find the chromosomal distribution of the MAR sites on SATB1 peak and in the whole mouse genome. Finally, in order to check for SATB1 preference for binding in specific MARs we compared the three distributions (SATB1 peak, MARs on SATB1 peak, MARs on mouse genome-mm9). We did not detect any difference between the three distributions (Fig. S1A,B,C).

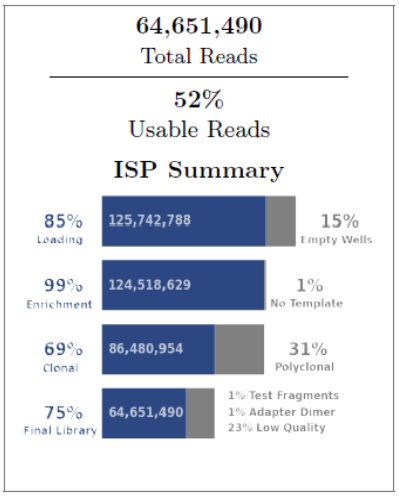


Figure 9. Run summary of the ChIP-sequencing.

p-value	IP/Input peaks	Input/IP peaks
10^{-5}	204.461	216.109
10^{-9}	34.932	54.735
10^{-13}	8.403	18.441

Table 1. Number of peaks in IP/Input comparison (2nd column) or Input/IP (3rd column) using the MACS peak caller with three different p-values

Binding motif and gene annotation

In order to detect any binding motif in SATB1 peaks we used the "SeqPos motif tool" option from the Galaxy/Cistrome online tool (http://cistrome.org/ap/). The binding motif for Jun protein (TGAGTCA) was the only motif that was identified, but its presence also in the reverse peaks (Input/IP) indicated that there is no significance in this identification (Fig. 11). Moreover, to define the genes (and their proximal regions) that include SATB1 binding sites, we used the "peak2gene: Peak Center Annotation" option of the Galaxy/Cistrome online tool. Using this tool we had both a list of genes that localize in a defined proximity to SATB1 peaks (Table S3) and a list of the peaks in each gene of the genome (Table S4).

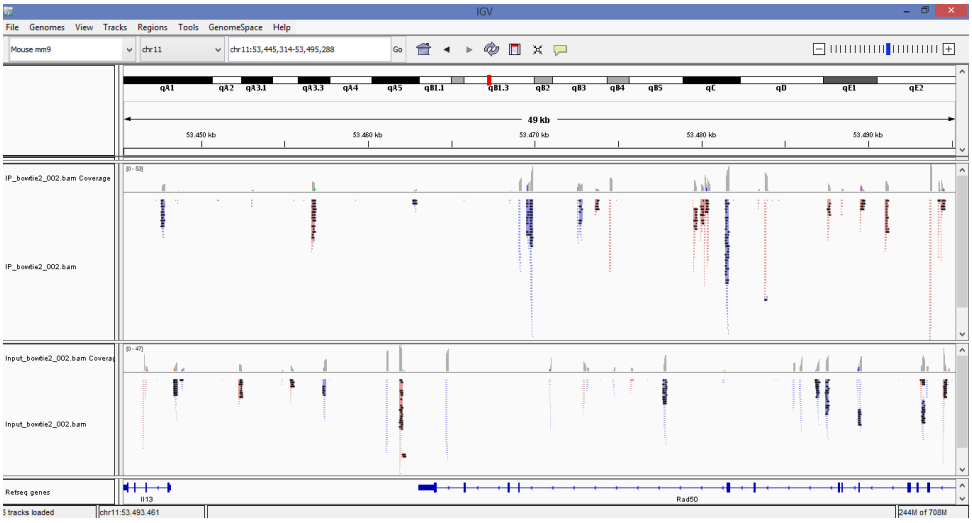


Figure 10. Visualization of the SATB1 IP and input DNA fragments (mapped in mm9) in IGV viewer.

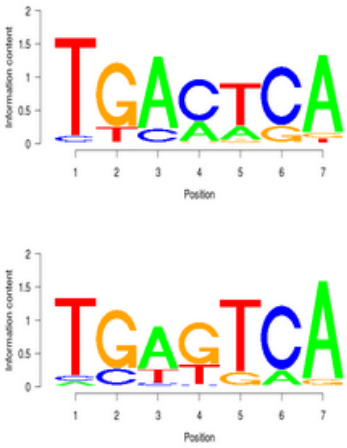


Figure 11. The JUN binding motif identified in SATB1 binding sites.

Discussion

Although several groups have focused on the genomic regions that interact during the formation of long-range chromatin interactions, less attention has been paid to the proteins that mediate these phenomena and the mechanisms through which they regulate and define chromatin conformation and genome locus localization. COMPASS complex, NSL complex, YY1 and AIRE proteins were all co-precipitated with the SATB1 protein and directly or indirectly associated with Histone 3 posttranslational modifications while, the third more enriched protein that precipitated with SATB1 is Histone 3. All these data leads to a clear association of SATB1 with histone modulator complexes and especially with H3K4 modifications.

Since SATB1 mediates long-range chromatin interactions, it would be of immense importance to investigate if during these phenomena SATB1 brings in close proximity distal genomic regions in order to activate genes in a faster and more accurate way via histone modification complexes. Moreover, the chromatin remodeler Mi-2β is the key protein in promoting CD4 expression in an enhancer-promoter looping event. It shall be investigated how SATB1 mediates Mi-2β binding to the CD4 enhancer or in forming the enhancer-promoter loop for induction of CD4 gene expression during T cell development.

Additionally, the Autoimmune Regulator (AIRE) is a transcription factor that controls the negative selection of self-reactive T cells, a procedure that is essential for the immune system to tolerate self-antigens and avoid autoimmunity. Creating an AIRE-dependent SATB1 conditional knockout mouse, in combination with molecular and imaging approaches will give us the chance to unravel mechanisms underlying AIRE regulation of T cell tolerance and the role of SATB1 in this process.

Furthermore, when the mapped reads were visualized in the IGV viewer, we observed that they were not represented by both strands and that the input sample did not have coverage of the whole genome. These reads seem to be PCR duplicates,

since most of them were stacked up in columns wherein all reads start from the same nucleotide.

Moreover, since SATB1 peaks were not validated by qPCR analysis using the IP and Input samples before the library construction, it is possible that these peaks were artifacts produced during the library preparation. When we applied the reverse peakcalling (Input/IP) we detected similar number of peaks to the forward one (IP/Input), which proves that there was not statistical significance in the number of peaks that we detected. Downstream analysis for gene annotation, MARs identification, chromosomal distribution and binding motif did not show any preference for SATB1 binding. All the above, indicate a problem during the adaptor ligation step of the library preparation. The small fraction of fragments that was amplified during the PCR and mainly in a single-strand way, resulted in a biased sequencing for both the IP and the Input sample.

In conclusion, we have to repeat the chromatin immunoprecipitation for the SATB1 protein and sequence the IP and Input samples. IP (anti-SATB1) in SATB1 $CD4^+$ conditional knockout mice or rabbit IgG immunoprecipitation in wild type mice shall also be sequenced in order to get rid of false positive peaks. When we repeat the SATB1 ChIP-seq approach, the combination of the SATB1 binding sites with its protein interactors could reveal the SATB1 role in T cell development, chromatin conformation and gene regulation through long-range chromatin interactions.

Materials and Methods

Co-Immunoprecipitation

Thymocyte protein extracts were extracted from 8 male 4-6 weeks old C57BL/6 mice. Thymocytes incubated at room temperature for 30 min with rotation, in 1ml lysis buffer (150mM NaCl, 50mM Tris pH=7.5, 5% Glycerol, 1% Nonidet P-40, 1mM $MgCl_2$, 200U benzonase, 1X Proteinase Inhibitors, 1mM PMSF). 24mg of protein extract were precleared using 100ul of Dynabeads® Protein G (#10004D LifeTechnologies) at 4^oC for 1.5 h on an end-to-end rotator. 50ul of beads pre-blocked with 15ug rabbit anti-SATB1 or rabbit IgG sera (#C2712 Santa Cruz Biotechnology) at 4^oC for 4 h rotating and crosslinked at room temperature for 30 min with tilting/rotation using cross-linker Bis(sulfosuccinimidyl) suberate BS^3 (#21580 Thermo Fisher Scientific Inc.). Precleared lysates were subjected to immunoprecipitation by incubation with Dynabeads crosslinked with anti-SATB1 or IgG at 4^oC for 30min. Beads were washed 3 times with buffer I (150mM NaCl, 50mM Tris pH=7.5, 5% Glycerol, 0.05% Nonidet P-40, 1mM PMSF), 2 times with Nonidet P-40 free Buffer II (150mM NaCl, 50mM Tris pH=7.5, 5% Glycerol, 1mM PMSF) and boiled at 95^oC for 5min.

SDS page and Silver Staining

Proteins co-Immunoprecipitated with SATB1 were resolved via SDS-PAGE in a 10% polyacrylamide gels. The gel was stained with silver nitrate compatible with Mass Spectrometry Fixation: 50% MeOH, 12% HAc, 0.05% Formalin 2hrs, Washes: 35% EtOH 20 min 3 times, Sensitization: 0.02% $Na_2S_2O_3$ 2 min, Washes: H_2O 5min 3

times, Staining: 0.2% $AgNO_3$, 0.076% formalin 20 min, Washes: H_2O 1min 2 times, Developing: 6% Na_2CO_3, 0.05% formalin, 0.0004% $Na_2S_2O_3$, Stop Staining: 50% MeOH, 12% HAc 5min.

Mass Spectrometry

10 pairs of protein bands (anti-SATB1, IgG) were excised from silver stained polyacrylamide gels and cut into small pieces. Gel pieces were covered with destain solution (30mM potassium ferricyanide, 100mM sodium thiosulfate) and vortexed for 10 minutes in order to remove silver nitrate. For the reduction and alkylation of the cysteine residues, the gel pieces were covered with 100ul 10mM DTT and shaked 45 min at 56^{o}C followed by 45 min shaking at room temperature with 100ul 55mM IAA. Proteins were digested in 30ul of diluted Trypsin solution and incubated overnight at 37^{o}C. The next day, the supernatant was collected. For the extraction from the gel matrix of generated peptides, gel pieces were shaked for 20 min first in 50ul 50% acetonitril (ACN) and finally in 50ul 0.1% TFA/50% ACN. Peptide solutions were centrifuged by Speed Vac until dry power remained and then analyzed by means of liquid chromatography combined with mass spectrometry (LC/MS).

Chromatin Immunoprecipitation

Spleens and lymph nodes were removed from 6 male 6 weeks old C57BL/6 mice and were used to make single cell suspension. CD4 cells were isolated with CD4 (L3T4) MicroBeads (130-049-201 Miltenyi Biotec.). Cells (60x10^{6}) were crosslinked in 1% formaldehyde (0,1V fixation buffer: 11% formaldehyde, 100mM NaCl, 1mM EDTA, 0,5mM EGTA, 50mM Hepes pH=7,5) at room temperature for 1min, neutralized with 125mM glycine and lysed in a Nonidet P-40 buffer (5mM PIPES, 85mM KCl, 0.5% Nonidet P-40, 1mM PMSF). Nuclei were resuspended in 2ml sonication buffer (0,3% SDS, 10mM Tris pH=8, 1mM EDTA) and were sonicated for 5 min using Covaris S2 sonicator (Peak Power 140, Duty Factor 5.0, Cycles/Burst 200) for chromatin fragmentation in 200-800bp. Fragmentation was verified both by agarose gel electrophoresis and bioanalyzer processing. 10ul of fragmented DNA was used as Input DNA. For the immunoprecipitation, soluble chromatin solution (SCS) was diluted by adding 4ml of IP buffer (10mM Tris pH=8, 1mM EDTA, 100mM NaCl, 1% Triton X-100, 1X complete inhibitors, 1mM PMSF), in order to lower SDS concentration, and incubated with 10ug rabbit anti-SATB1 or rabbit IgG and 100ul Dynabeads® Protein G LifeTechnologies at 4^{o}C for 2 hours. The beads were washed with low salt buffer (20mM Tris pH=8, 150mM NaCl, 2mM EDTA, 0.1% SDS, 1% Triton X-100, 1mM PMSF), high salt buffer (20mM Tris pH=8, 500mM NaCl, 2mM EDTA, 0.1% SDS, 1% Triton X-100, 1mM PMSF), Lithium containing buffer (10mM Tris pH=8, 250mM LiCl, 2mM EDTA, 1% NP-40, 1mM PMSF), TE buffer (10mM Tris pH+8, 1mM EDTA, 50mM NaCl) to remove non-specifically bound chromatin. Crosslinks were reversed by overnight incubation in 250ul elution buffer (50mM Tris pH=8, 10mM EDTA, 1% SDS) at 65^{o}C. RNA contaminants and proteins were removed by digestion 30min at 37 ^{o}C adding 50ug RNAse and 1h at 55 ^{o}C with 20ug Proteinase K. DNA was purified using Phenol/Chloroform/Isoamyl alcohol extraction followed by ethanol

precipitation. DNA concentration was calculated using Qubit® 2.0 Fluorometer (Lifetechnologies).

Real-Time PCR

Quantitive PCR reactions were performed on a DNA Engine Opticon® 2 System (BioRad) using KAPA SYBR® FAST qPCR mix (KAPABIOSYSTEMS). The enrichment of the IP over IgG fractions in a targeted (positive) genomic region versus a negative control region was calculated using the formula:

Ct_{Input} was corrected by the dilution number $Ct'_{Input} = Ct_{Input} - \log_2 100$

$(2^{CtIp-Ct'Input} / 2^{CtIgG-Ct'Input})_{positive\ region} / (2^{CtIp-Ct'Input} / 2^{CtIgG-Ct'Input})_{negative\ region}$

As a positive region RHS6 was used (RHS6F 5'-CTCCAGGCTTACTGTTTCAGCA-3', RHS6R 5'-GCTGCCTCTGTACAATCATTCC-3') and as a negative control region we used HPRT1 (HPRTF 5'-CTGGGCCTAAATCTTGAGGAA-3', HPRTR 5'-TCCCAGATAATCACTCCGCA-3')

Primers for SATB1 peak validation:

CHR13F 5'- CCCAGCATGCAGTAGGGTAG-3', CHR13R 5'- AGGGACTTCGAGTGATTGGG-3', RHS4F 5'- GGAAACAAACCCCACGCAA-3', RHS4R 5'-AGCTGGTGAACAAGGACCTC-3', RHS6F 5'-TACACCCTCAGCCTTAGCAC-3', RHS6R 5'-ACAATGGAGAGGGATTGGGG-3'

95°C 5sec, 60°C 20sec, 72°C 1sec, plate read, 40 cycles, 72°C 5min, Melting Curve from 60°C 95°C, read every 0.5°C, hold 1sec, 15°C 1'sec

Library Preparation

Ion Plus Fragment Kit (Life Technologies) was used for library preparation. ChIP DNA was end-repaired and purified with two rounds of AMPure®XP bead capture to size-select fragments 100-250bp in length. The end-repaired DNA was ligated to Ion-compatible adapters, followed by nick repair to complete the linkage between adapters and DNA inserts. The library was amplified by PCR with 10 cycles for generating sufficient material. The final library was 170-220 bp in length and sequenced on an Ion Proton™ System (LIfeTechnologies).

Sequencing

After library preparation, IP and Input samples were single-end sequenced in an Ion Proton™ System. The output of this sequencing was a "fastq" file containing 65,4 millions of reads, free of the adaptors' sequence.

Adaptors' sequence: plus 3:2:15:2-CCATCTCATCCCTGCGTGTCTCCGACTCAG, minus 3:2:15:2-CCTCTCTATGGGCAGTCGGTGAT.

Mapping

ChIP-seq reads were aligned to the mouse genome database (mm9) using Bowtie2. For mapping, the "--very-sensitive-local" option was used same as "-D 20 -R 3 -N 0 -L 20 -i S,1,0.50" so that no mismatches were allowed.

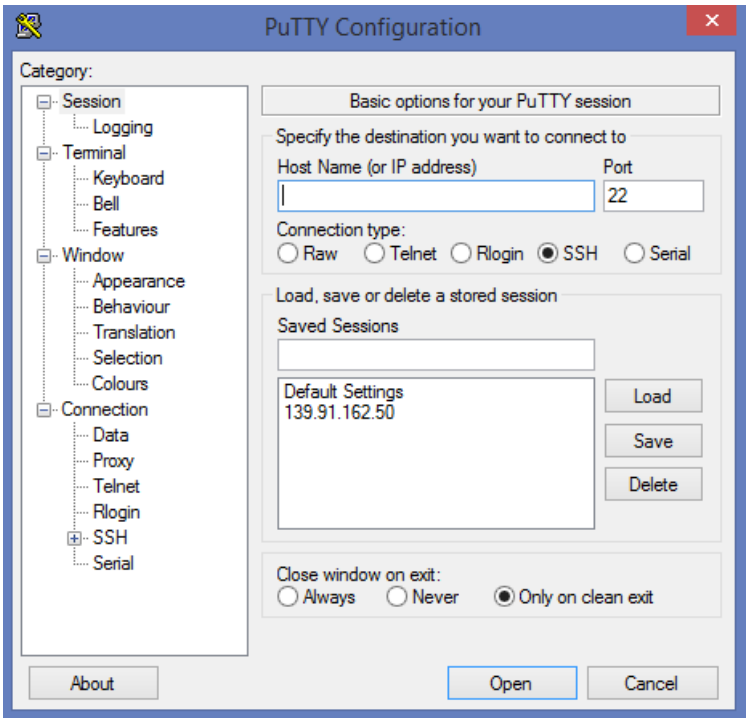
Firstly, we connected to the local server using the program PuTTy configuration

Host Name: 139.91.162.50

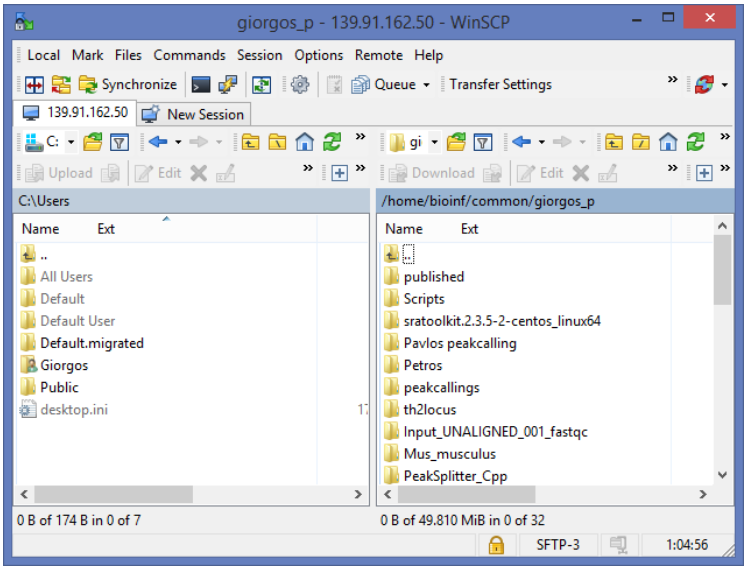
Port: 33333

username: giorgos

password: g10rg0s



Consequently, we uploaded our data to the local server using the WinSCP program (Host Name: 139.91.162.50
Port: 33333
username: giorgos
password: g10rg0s



Sequencing of the output could be either in .sra or .fastq form. For mapping these data to a genome we need a .fastq output. In order to make a conversion from .sra to .fastq we used the script

```
"fastq-dump filename.sra"
```

Since the data is in .fastq form we proceeded to mapping using the script:

in case of single-end sequencing

```
"bowtie2 --local --very-sensitive-local -p 8 --mm -x Mus_musculus/UCSC/mm9/Sequence/Bowtie2Index /genome -U filename.fastq | samtools view -uhS -F4 - | samtools sort - filename"
```

in case of pair-end sequencing

```
"bowtie2 -p12 --mm -x Mus_musculus/UCSC/mm9/Sequence/Bowtie2Index/genome -X 650 -1 pathway/filename1.fastq -2 pathway/filename2.fastq | samtools view -uhS -F4 - | samtools sort - filename"
```

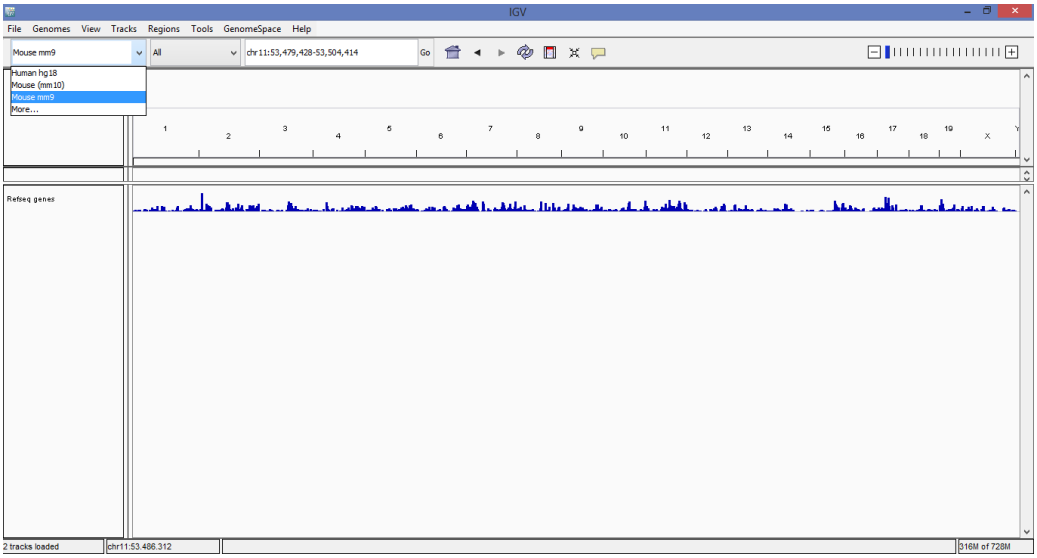
Create and index .bam.bai file for every .bam file (mapped file) using the following script:

```
"samtools index pathway/filename"
```

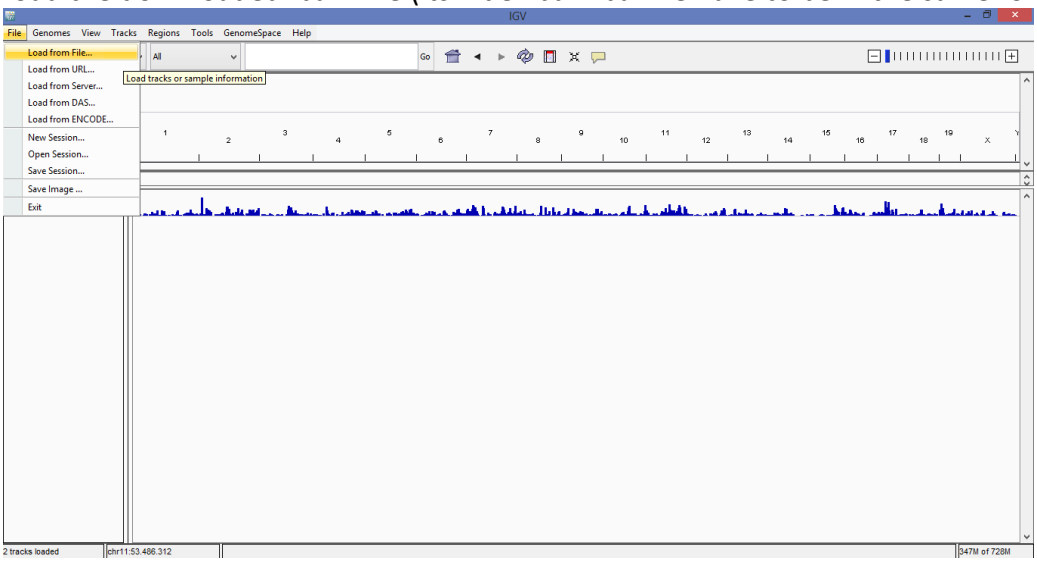
IGV viewer visualization

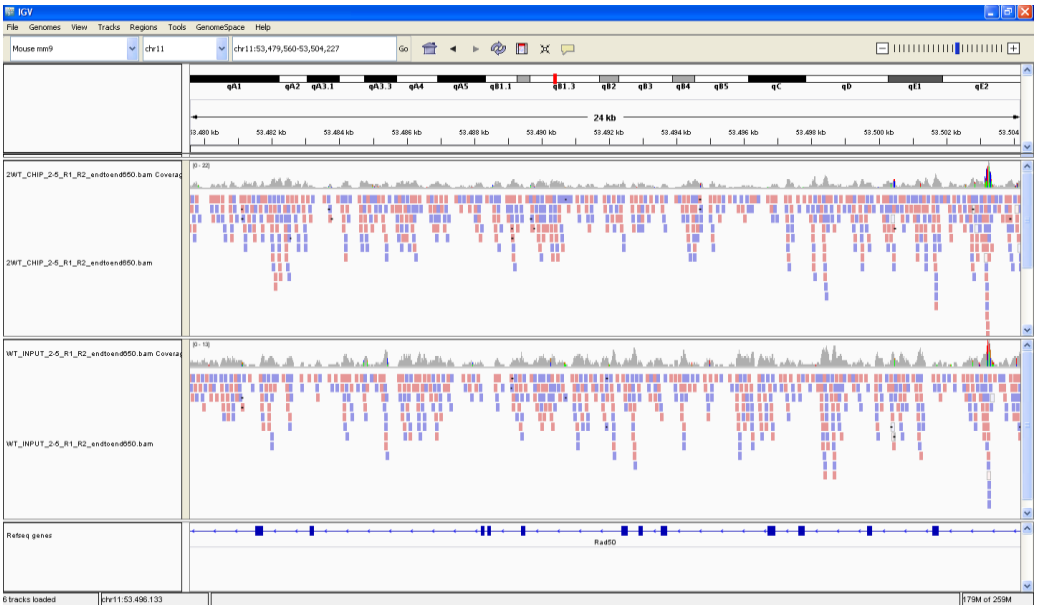
We were able to view our mapped data in the "IGV viewer" program.

Load genome mm9

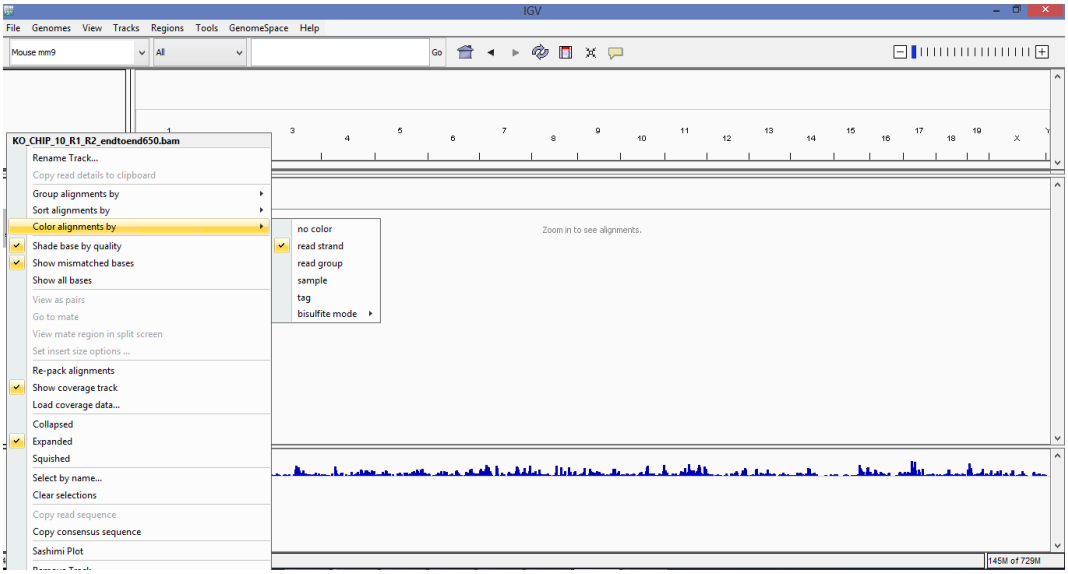


Load the downloaded .bam file (its index bam.bai file have to be in the same folder)





To add color to each strand of reading we (right) click on the track (left column) and chose option "Color alignments by/read strand"



Peak calling

Peak discovery was performed using MACS with the default parameters or changing the p-value option.

```
"macs14 -t treatmentfilename.bam -c controlfilename.bam -f BAM
-g mm -n outputfilename -w --bw 300 -p 1e-5 --call-subpeaks"
```

Options:
-g GSIZE, --gsize=GSIZE
Effective genome size. It can be 1.0e+9 or 1000000000,or shortcuts: 'hs' for human (2.7e9), 'mm' for mouse (1.87e9), 'ce' for C. elegans (9e7) and 'dm' for drosophila melanogaster (1.2e8),
DEFAULT:hs

-s TSIZE, --tsize=TSIZE
Tag size. This will overide the auto detected tag size.
DEFAULT: 25

--bw=BW
Band width. This value is only used while building the shifting model.
DEFAULT: 300

-p PVALUE, --pvalue=PVALUE
P-value cutoff for peak detection.
DEFAULT: 1e-5

-m MFOLD, --mfold=MFOLD

We selected the regions within MFOLD range of high-confidence enrichment ratio against background to build a model. The regions had to be lower than the upper limit, and higher than the lower limit.
DEFAULT:10,30

--nolambda
If True, MACS will use fixed background lambda as local lambda for every peak region. Normally, MACS calculates a dynamic local lambda to reflect the local bias due to potential chromatin structure.

--slocal=SMALLLOCAL
The small nearby region in basepairs to calculate dynamic lambda. This is used to capture the bias near the peak summit region. Invalid if there is no control data.
DEFAULT: 1000

--llocal=LARGELOCAL
The large nearby region in basepairs to calculate dynamic lambda. This is used to capture the surround bias.
DEFAULT: 10000.

--off-auto
Whether turn off the auto pair model process. If not set, when MACS failed to build paired model, it will use the nomodel settings, the '--shiftsize' parameter to shift and extend each tags.
DEFAULT: False

--nomodel
Whether or not to build the shifting model. If True, MACS will not build model. by default it means shifting size = 100, try to set shiftsize to change it.
DEFAULT: False

--shiftsize=SHIFTSIZE
The arbitrary shift size in bp. When nomodel is true, MACS will use this value as 1/2 of fragment size.
DEFAULT: 100

--keep-dup=KEEPDUPLICATES
It controls the MACS behavior towards duplicate tags at the exact same location -- the same coordination the same strand. The default 'auto' option makes MACS calculate the maximum tags at the exact same location based on binomal distribution using 1e-5 as pvalue cutoff; and the 'all' option keeps every tags. If an integer is given, at most this number of tags will be kept at the same location.
Default: auto

--to-small
When set, scale the larger dataset down to the smaller dataset, by default, the smaller dataset will be scaled towards the larger dataset.

DEFAULT: False

-w, --wig
Whether or not to save extended fragment pileup at every WIGEXTEND bps into a wiggle file. When --single- profile is on, only one file for the whole genome is saved. WARNING: this process is time/space consuming!!

-B, --bdg
Whether or not to save extended fragment pileup at every bp into a bedGraph file. When it's on, -w,--space and --call-subpeaks will be ignored. When --single-profile is on, only one file for the whole genome is saved. WARNING: this process is time/space consuming!!

-S, --single-profile
When set, a single wiggle file will be saved for treatment and input. Default: False

--space=SPACE
The resoluation for saving wiggle files, by default, MACS will save the raw tag count every 10 bps. Usable only with '--wig' option.

--call-subpeaks
If set, MACS will invoke Mali Salmon's PeakSplitter soft through system call. If PeakSplitter can't be found, an instruction will be shown for downloading and installing the PeakSplitter package. -w option needs to be on and -B should be off to let it work.
DEFAULT: False

--verbose=VERBOSE
Set verbose level. 0: only show critical message, 1:show additional warning message, 2: show process information, 3: show debug messages. DEFAULT:2

--diag
Whether or not to produce a diagnosis report. It's up to 9X time consuming. Please check 00README file for detail.
DEFAULT: False

--fe-min=FEMIN
For diagnostics, min fold enrichment to consider.
DEFAULT: 0

--fe-max=FEMAX
For diagnostics, max fold enrichment to consider.
DEFAULT: maximum fold enrichment

--fe-step=FESTEP
For diagnostics, fold enrichment step.
DEFAULT: 20

Marscan

For MARS identification we used the emboss marscan online tool (http://emboss.bioinformatics.nl/cgi-bin/emboss/marscan). Since there was a limit in the maximum length of the sequence that could be used by this tool (30.000bp), we had to split the chromosome sequences in fractions of 30.000bps, analyzed each one for MARS regions and merged the output lists in an tab delimited file (bed format). Split script "`split -a 2 -l 1300000 -d chr1.fa chr1.fa.`"

MARS on SATB1 peaks

For the identification of MARS in SATB1 peaks we used the "Intersect the intervals of two datasets" option from the Galaxy/Cistrome online tool (http://cistrome.org/ap/) in bed files containing MARS sites and SATB1 peaks. This option sets the overlap between elements of two datasets.

Binding Motif

In order to find a SATB1 binding motif we used the "SeqPos motif tool" option from the Galaxy/Cistrome online tool (http://cistrome.org/ap/) in the bed file containing the SATB1 peaks.

Chromosomal distribution

For chromosomal distribution we used the "CEAS: Enrichment on chromosome and annotation" option of the Galaxy/Cistrome online tool (http://cistrome.org/ap/). Peaks that identified with p-value=10^{-5} were used for the SATB1 peak distribution.

Representative sampling

In order to take a sample file from a tab delimited .bed file we used the script:
"`cat filename.bed |shuf | head -n numberoflines >filename.bed`"

Gene annotation

Loading SATB1's peaks to the "peak2gene:Peak Center Annotation" option of the Galaxy/Cistrome online tool (http://cistrome.org/ap/) we were able to detect the gene annotation of SATB1 binding sites (distance: 300bp, 1.000bp, 10.000bp/ genome: mm9).

Supplementary Materials

Table S1 Mass Spectrometry analysis
Table S2A Table S2B Peak calling p-value=10^{-9} (excel file A/ bed file B)
Table S3 Annotation for each peak (p-value=10^{-9}, distance:10.000bp)
Table S4A S4B S4C Annotation for each gene (p-value=10^{-9}, distance:300bp A/ 1.000bp B/ 10.000bp C)
Figure S1A S1B S1C Chromosomal distribution of SATB1 peaks (A), MARS mm9 (B) and MARS on SATB1 peaks (C).

References

[1] Spilianakis C.G., Flavell RA. Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat Immunol*., 5, 1017 (2004)

[2] Nathan Harmston, Boris Lenhard Chromatin and epigenetic features of long-range gene regulation *Nucleic Acids Research* 41, 7185 (2013)

[3] Amsen D., Spilianakis C.G, Flavell R.A. How are T(H)1 and T(H)2 effector cells made? *Curr Opin Immunol.* 21, 153 (2009)

[4] Cai, S., Han, H.J. & Kohwi-Shigematsu, T. Tissue-specific nuclear architecture and gene expression regulated by SATB1. *Nat. Genet*. 34, 42–51 (2003)

[5] Pavan Kumar P. , Oliver Bischof , Prabhat Kumar Purbey , Dimple Notani , Henning Urlaub , Anne Dejean, Sanjeev Galande, Functional interaction between PML and SATB1 regulates chromatin-loop architecture and transcription of the MHC class I locus. *Nature Cell Biology*, 9, 45-56 (2007)

[6] Zheng Wang, Jiafu Long, The structural basis for the oligomerization of the N-terminal domain of SATB1. *Nucleic Acids Research*, 409, 4193–4202 (2012)

[7] Hui Nie, Philip W. Tucker, SATB1 is required for CD8 coreceptor reversal. *Mol. Immunol.* 46, 207–211 (2008)

[8] de Belle, I., Cai, S. & Kohwi-Shigematsu, T. The genomic sequences bound to special AT-rich sequence-binding protein 1 (SATB1) in vivo in Jurkat T cells are tightly associated with the nuclear matrix at the bases of the chromatin loops. J. *Cell Biol*. 141, 335–348 (1998)

[9] Prabhat Kumar Purbey, Sunita Singh, Dimple Notani, P. Pavan Kumar, Amita S. Limaye, Sanjeev Galande, Acetylation-Dependent Interaction of SATB1 and CtBP1 Mediates Transcriptional Repression by SATB1. *Mol & Cel Biol* 29,1321–1337 (2009)

[10] 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)

[11] Dimple Notani , Kamalvishnu P. Gottimukkala , Ranveer S. Jayani , Amita S. Limaye , Madhujit V. Damle , Sameet Mehta , Prabhat Kumar Purbey, Jomon Joseph , Sanjeev Galande, Global Regulator SATB1 Recruits b-Catenin and Regulates TH2 Differentiation in Wnt-Dependent Manner , *PLOS Biology*, 8 (2010)

[12] Ruben Agrelo, Abdallah Souabni, Maria Novatchkova, Christian Haslinger, Martin Leeb, Vukoslav Komnenovic, Hiroyuki Kishimoto, Lionel Gresh, Terumi Kohwi-Shigematsu, Lukas Kenner and Anton Wutz, SATB1 Defines the Developmental Context for Gene Silencing by Xist in Lymphoma and Embryonic Cells*. Developmental Cell* 16, 507–516, (2009)

[13] Britta Will, Thomas O. Vogler, Boris Bartholdy, Francine Garrett-Bakelman, Jillian Mayer, Laura Barreyro, Ashley Pandolfi, Tihomira I. Todorova, Ujunwa C. Okoye-Okafor, Robert F. Stanley, Tushar D. Bhagat, Amit Verma, Maria E. Figueroa, Ari Melnick, Michael Roth, Ulrich Steidl, Satb1 regulates hematopoietic stem cell self-renewal by promoting quiescence and repressing differentiation commitment. *Nat Immunol*. 14, 437–445 (2013)

[14] John D. Alvarez, Dag H. Yasui, Hiroyuki Niida, Tadashi Joh, Dennis Y. Loh,Terumi Kohwi-Shigematsu, The MAR-binding protein SATB1 orchestrates temporal and spatial expression of multiple genes during T-cell development. *GENES & DEVELOPMENT* 14, 521–535 (2000)

[15] Spilianakis C. et al. Interchromosomal associations between alternatively expressed loci. *Nature* 435:637 (2005)

[16] Lee G.R. et al. T helper cell differentiation: regulation by cis elements and epigenetics. *Immunity* 24:369 (2006)

[17] Kim L.K. et al. Oct-1 Regulates IL-17 Expression by Directing Interchromosomal Associations in Conjunction with CTCF in T Cells. *Mol Cell* 54:56 (2014)

[18] Bernd Schuettengruber, Anne-Marie Martinez, Nicola Iovino, Giacomo Cavalli, Trithorax group proteins: switching genes on and keeping them active. *Nat. Rev*. 12, 799-814

[19] Deqing Hu, Alexander S Garruss, Xin Gao, Marc A Morgan, Malcolm Cook, Edwin R Smith & Ali Shilatifard, The Mll2 branch of the COMPASS family regulates bivalent promoters in mouse embryonic stem cells. Nature Structural & Molecular Biology (2013)

[20] Trissa Miller, Nevan J. Krogan, Jim Dover, H. Erdjument-Bromage, Paul Tempst, Mark Johnston, Jack F. Greenblatt, Ali Shilatifard, COMPASS: A complex of proteins associated with a trithorax-related SET domain protein. *PNAS* 98, 12902–12907 (2001)

[21] Akihiko Yokoyama, Michael L. Cleary, Menin critically links MLL proteins with LEDGF on cancer-associated target genes. *Cancer Cell*. 14, 36–46 (2008)

[22] Man Mohan, Chengqi Lin, Erin Guest, Ali Shilatifard, Licensed to elongate: a molecular mechanism for MLL-based leukaemogenesis. *Nat. Rev.*10, 721-727 (2010)

[23] Edwin Smith, Chengqi Lin, Ali Shilatifard, The super elongation complex (SEC) and MLL in development and disease. *Genes Dev*. 25, 661-672 (2001)

[24] Janet L. Thornton, Gerwin H. Westfield, Yoh-hei Takahashi, et al. Context dependency of Set1/COMPASS-mediated histone H3 Lys4 Trimethylation. *Genes Dev*. 28 (2014)

[25] Soo Seok Hwang, Young Uk Kim, Sumin Lee, Sung Woong Jang, Min Kyung Kim, Byung Hee Koh, Wonyong Lee, Joomyeong Kim, Abdallah Souabni, Meinrad Busslingerc, Gap Ryol Lee, Transcription factor YY1 is essential for regulation of the Th2 cytokine locus and for Th2 cell differentiation. *PNAS* 110, 276–281 (2013)

[26] Xiaoming Zhao, Jiaming Su., Fei Wang, Da Liu, Jian Ding, Yang Yang, Joan W. Conaway, Ronald C. Conaway, Lingling Cao, Donglu Wu, Min Wu, Yong Cai, Jingji Jin, Crosstalk between NSL Histone Acetyltransferase and MLL/SET Complexes: NSL Complex Functions in Promoting Histone H3K4 Di-Methylation Activity by MLL/SET Complexes. *PLOS Genetics* 9, 1-11 (2013)

[27] Kin Chung Lam, Friederike Muhlpfordt, Juan M. Vaquerizas, Sunil Jayaramaiah Raja, Herbert Holz, Nicholas M. Luscombe, Thomas Manke, Asifa Akhtar, The NSL Complex Regulates Housekeeping Genes in Drosophila. *PLOS Genetics* 8, 1-11 (2012)

[28] Christine J. Williams, Taku Naito, Pablo Gomez-del Arco, John R. Seavitt,Susan M. Cashman,Beverly De Souza, Xiaoqing Qi, Piper Keables, Ulrich H. Von Andrian, Katia Georgopoulos, The Chromatin Remodeler Mi-2β is required for CD4 Expression and T cell Development. *Immunity* 20, 719–733, (2004)

[29] Andrew S. Koh, Alex J. Kuo, Sang Youn Park, Peggie Cheung, Jakub Abramson, Dennis Bua, Dylan Carney, Steven E. Shoelson, Or Gozani, Robert E. Kingston, Christophe Benoist, Diane Mathis, Aire employs a histone-binding module to mediate immunological tolerance, linking chromatin regulation with organ-specific autoimmunity. PNAS 105, 15878–15883 (2008)

[30] Rina Aharoni, Revital Aricha, Raya Eilam, Ido From, Keren Mizrahi, Ruth Arnon, Miriam C. Souroujon, Sara Fuchs, Age dependent course of EAE in Aire−/− mice. *Journal of Neuroimmunology* 262, 27–34 (2013)

[31] Anthony J. Bonito, Costica Aloman, M. Isabel Fiel, Nichole M. Danzl,Sungwon Cha, Erica G. Weinstein, Seihwan Jeong, Yongwon Choi, Matthew C. Walsh, Konstantina Alexandropoulos, Medullary thymic epithelial cell depletion leads to autoimmune hepatitis. *J Clin Invest*. 123:3510–3524 (2013)

[32] Bruno Kyewski, Pärt Peterson, Aire, Master of Many Trades. *Cell* 140, (2010)

[33] Jakub Abramson, Matthieu Giraud, Christophe Benoist, Diane Mathis, Aire's Partners in the Molecular Control of Immunological Tolerance. *Cell*, 123–135, (2010)

[34] Jukka Pitkanen, Vassilis Douca Thomas Sternsdor, Toshihiro Nakajima, Satoko Aratani, Kirsten Jensen, Hans Will, Perttu Vahamurto, Juha Ollila, Mauno Vihinen, Hamish S. Scott, Stylianos E. Antonarakis, Jun Kudoh, Nobuyoshi Shimizu, Kai Krohn, and Part Petersona, The Autoimmune Regulator Protein Has Transcriptional Transactivating Properties and Interacts with the Common Coactivator CREB-binding Protein. *The Journal of Biological Chemistry* 275, 16802–16809 (2000)