Arts & Sciences Electronic Theses and Dissertations

Arts & Sciences

Spring 5-15-2017

Epigenetic Activation of the Mouse T Cell Receptor Beta Recombination Center

Jiangyang Zhao
Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

Part of the Allergy and Immunology Commons, Genetics Commons, Immunology and Infectious Disease Commons, and the Medical Immunology Commons

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Molecular Genetics and Genomics

Dissertation Examination Committee:
Eugene Oltz, Chair
Shiming Chen
Takeshi Egawa
Jason Mills
Barry Sleckman
Ting Wang

Epigenetic Activation of the Mouse T Cell Receptor Beta Recombination Center
by
Jiangyang Zhao

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2017
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

List of Abbreviations

aa	amino acids
AD	Activation Domain
ACE	Accessibility Control Element
AML	Acute Myeloid Leukemia
APC	Antigen Presenting Cells
AR	Antigen Receptor
BM	Bone Marrow
CBF	Core Binding Factor
CD	Cluster of Differentiation
CLP	Common Lymphoid Progenitor
DBD	DNA Binding Domain
DN	Double Negative
DP	Double Positive
EG	ETS1 binding site to Gal4 binding site mutation
EGRm	EG with RUNX1 binding site null mutation
EmRG	RG with ETS1 binding site null mutation
ETP	Early T-cell Precursor
HR	Homologous Recombination
HSC	Hematopoietic Stem Cell
H3K4me3	histone H3 lysine 4 tri-methylation
H3K27ac	histone H3 lysine 27 acetylation
ID	Inhibition Domain
Kb	kilobase

kD   kilo Dalton

MAR   Matrix Attachment Region

MPP   MultiPotent Progenitor

NHEJ   Non-Homologous End Joining

NRDB   Negative Regulatory region of DNA Binding

PCR   Polymerase Chain Reaction

PHD   Plant HomeoDomain

RAG   Recombination-Activating Gene

RC   Recombination Center

RG   RUNX1 binding site to Gal4 binding site mutation

RSS   Recombination Signal Sequence

SE   Super Enhancer

SP   Single Positive

TAD   TransActivation Domain

TCR   T Cell Receptor

TF   Transcription Factor

WT   Wild Type

ZFN   Zinc Finger Nuclease

3C   Chromosome Conformation Capture

5C   Chromosome Conformation Capture Carbon Copy

Acknowledgements

First I would like to sincerely thank my mentor, Dr. Gene Oltz, for his continuous support and guidance through my whole Ph.D period. I have learnt a lot from him, not just in science and professionalism, but also in general way of thinking and problem solving. I cannot find a more wholehearted mentor than him.

Also, I am very grateful to my thesis committee, as well as Dr. Jim Skeath for their constructive and thoughtful suggestions in my experimental design and data analysis. I have met with amazing people who inspired me and set examples for me. I have benefited greatly from my training in DBBS at Washington University.

I would express my appreciation to all the past and current members of Oltz lab, especially Dr. Oleg Osipovich, who laid the foundation of this work. I would like to thank all other members for the encouragement, thoughtful discussion, and technical support. In addition, I will remember the help from our neighboring labs on the 7th floor of CSRB, where I constantly dropped by with questions and requests.

I am deeply indebted to my family and friends. My parents, Ming Zhao and Zhijun Jiang, are thousands of miles away, and yet I am always empowered by them regardless of distance. I truly cherish the wonderful time I spent with my wife Ran Hua in St. Louis. Hopefully I will reunite with her soon. I will never forget the help I got from my friends to overcome various hardships during this endeavor.

Jiangyang Zhao

Washington University in St. Louis

May 2017

ABSTRACT OF THE DISSERTATION

Epigenetic Activation of the Mouse T Cell Receptor Beta Recombination Center

by

Jiangyang Zhao

Doctor of Philosophy in Biology and Biomedical Sciences

Molecular Genetics and Genomics

Washington University in St. Louis, 2017

Dr. Eugene M. Oltz

Lymphocytes are the work horses of adaptive immunity. Compared to the B lymphocyte lineage, early stage progenitors of T lymphocytes maintain considerable potential for differentiation into other hematopoietic lineages. T lineage commitment requires the continuous coordination of transcription factors (TFs) by Notch1 signaling after multi-potent progenitors (MPPs) migrate to thymus. One of the first hall marks of T lineage commitment is expression of the T cell receptor β (TCRβ), which is encoded by the *Tcrb* locus following its assembly by V(D)J recombination, a somatic shuffling of the genome that joins one V, one D, and one J gene segment. *Tcrb* assembly is initiated at its recombination center (RC), composed of two DβJβ clusters. *Tcrb*-RC exhibits features of regulatory regions called super-enhancers (SEs), which are characterized by high level of active histone mark, H3K27ac, and by clusters of binding for TFs involved in cell fate decisions. A key *Tcrb*-RC enhancer, called Eβ, harbors two composite ETS1-RUNX1 binding motifs, which widely exist in regulatory elements for genes involved in T lymphopoiesis. ETS1 is sharply upregulated during T cell lineage commitment and recruits constitutively expressed RUNX1 to Eβ. However, the independent roles of these two TFs remain unclear, especially since

both are potent transactivators. In this study, I have shown that both ETS1 and RUNX1 are sufficient to independently activate Eβ in extrachromosomal reporter substrates. However, ETS1 by itself fails to activate Eβ in its native chromosomal context. By contrast, RUNX1 is sufficient to activate the endogenous Eβ element and its neighboring 25 kb region independently from ETS1. In addition, RUNX1 is sufficient to mediate long-range promoter-Eβ interactions, nucleosome clearance, and robust transcription throughout the *Tcrb* recombination center (RC). We also find that a RUNX1 domain, termed the negative regulatory domain for DNA binding (NRDB), can compensate for loss of ETS1 binding at adjacent sites. Thus, we have defined independent roles for RUNX1 in the activation of a T cell developmental enhancer, as well as its ability to mediate specific changes in chromatin landscapes that accompany long-range induction of RC promoters.

Chapter 1 : Introduction

1.1   TCRβ expression is a hallmark of T lineage commitment

Adaptive immunity is a crucial feature for the survival of all jawed vertebrates (1, 2). In this system, lymphocytes exist in large repertoires that contain different cell clones carrying antigen receptors (AR) with different binding capacities to fight potential pathogens (1–3). There are two types of lymphocytes for adaptive immunity, B and T cells, which are defined by their ARs, termed B-cell receptors (BCR) and T-cell receptors (TCR), respectively (4, 5). Compared to B cells, which secrete antibodies and recognize soluble antigens, T cells bind to peptide-MHC complexes that are processed by antigen presenting cells (APC) (6). The phenotypic differences between B and T cells result from choices of diverging transcriptional cascades from early stages of lymphocyte development, which are reinforced by external signaling (4, 5, 7–9).

Unlike B cells, T cells still maintain their potential to differentiate into other lymphoid and myeloid lineages during the early stages of their development (9, 10). Such plasticity has been investigated extensively using mouse models, which revealed that T lineage commitment is achieved as precursor cells gradually lose their potential for other lineages (9, 10). In mouse bone marrow, some hematopoietic stem cells (HSC) differentiate into multipotent progenitors (MPP), some of which further differentiate into common lymphoid progenitors (CLP) (11). Progenitor cells then migrate to the thymus, where they become early T-cell precursors (ETP).

ETPs are continuously stimulated by intrathymic epithelial cells through the Notch signaling pathways, which further induce expression of T-cell-specific transcription factors including TCF1, GATA3, and ETS1 (9, 10). These CD4$^-$CD8$^-$ double negative (DN) T cells will develop into CD4$^+$CD8$^+$ double positive (DP) T cells, which eventually differentiate into helper CD4$^+$ and cytotoxic CD8$^+$ single positive (SP) T cells (5).

DN cells can be further categorized into four developmental stages using two additional surface markers, CD44 and CD25 (12, 13). As precursor T cells progress from DN1 (CD44$^+$CD25$^-$) and DN2 (CD44$^+$CD25$^+$) to DN3 (CD44$^-$CD25$^+$) stages, they gradually lose their plasticity and eventually commit to the T lineage (10). At the DN3 stage, a small proportion (~2%) of cells will express TCRγ and TCRδ chains to become γδ T cells (14). Most cells; however, successfully express mature TCRβ chain, forming a surrogate TCR, termed pre-TCR, which is composed of TCRβ, an invariant surrogate pTα chain, and CD3 molecules (15). Once the pre-TCR is expressed the DN3 cells undergo β-selection to test TCR functionality (16). If they pass this checkpoint, they will further develop into DN4 (CD44$^-$CD25$^-$) and then DP cells (5). DP cells that successfully express a TCRα chain will have the chance to become mature αβ T cells (5). Therefore, TCRβ chain expression is the first critical event after precursor cells commit to the T lineage.

1.2 V(D)J recombination of *Tcrb* locus generates TCRβ coding sequence

Expression of TCRβ depends on successful rearrangement of the T cell receptor beta (*Tcrb*) locus. Like other AR loci, *Tcrb* is composed of linear arrays of coding segments named variable (V), diversity (D), and joining (J) (17). The somatic shuffling process to rearrange these segments is termed V(D)J recombination, in which one V, one D (if any), and one J segment are randomly excised and joined together to form a mature coding sequence (1).

V(D)J recombination is a multi-step biochemical process utilizing many enzymes. This process starts from recombination signal sequences (RSSs), which are located adjacent to all AR coding segments and facilitate joining to other segments (18, 19). RSSs are directional, each with a conserved heptamer, followed by a 12- or 23-nucleotide spacer, and an AT-rich nonamer (19). RSSs are recognized by RAG1/2 recombinase, which juxtaposes one 12-spacer RSS and one 23-spacer RSS (1). When coding segments are brought together by their RSSs, RAG1/2 will cleave them from RSSs and they may lose or gain nucleotides (2). They are then ligated through the non-homologous end joining (NHEJ) pathway (20).

The mouse *Tcrb* locus is located on chromosome 6 and spans over 660 kilobases (kb) (21). The upstream region contains thirty Vβ segments and spans over 400 kb, which is separated from the downstream DβJβ region by over 230 kb of trypsinogen genes (Fig. 1.1) (21, 22). Compared to Vβ segments, the downstream region is much smaller and spans 25 kb (21). This region harbors two Dβ-Jβ clusters, each with one Dβ, six usable Jβ, and a series of exons that encode the

constant region of TCRβ (Cβ) (21). At the very end of the *Tcrb* locus there is an isolated Vβ segment in the reverse orientation (23). All Vβ segments are followed by 23-spacer RSSs, whereas all Jβ segments are preceded by 12-spacer RSSs (1). The two Dβ segments are both flanked by 12-spacer RSSs in the upstream and 23-spacer RSSs in the downstream, allowing them to recombine with Vβ and Jβ segments, respectively (2). Although the RSS pairing principle may allow Vβ to Jβ joining, such a scenario is very rare (2). Instead, the order of recombination is tightly regulated to favor Dβ to Jβ joining before Vβ to DβJβ joining (24). It should be noted that the human *Tcrb* locus, especially its Dβ-Jβ region, is highly similar to the mouse gene, suggesting that conservation of the genetic structure and regulation across this region may be important for T cell development (21).
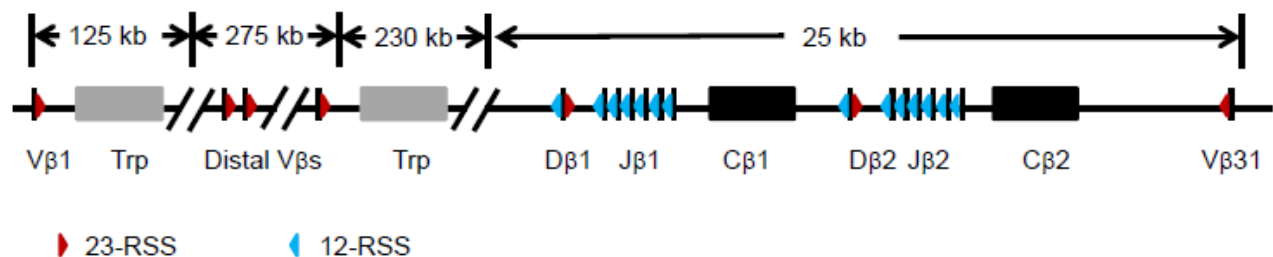
Figure 1.1 Genetic structure of the mouse *Tcrb* locus.

Schematic representation of the mouse *Tcrb* locus with an emphasis on the downstream DβJβ clusters. The V, D, and J segments are represented by black bars. Red and blue triangles represent 23-RSS and 12-RSS, respectively. The grey boxes are silent trypsinogen genes (Trp). The black boxes are constant regions.

1.3 *Tcrb*-RC is a super-enhancer driven by ETS1-RUNX1 co-binding

Despite lymphoid-specific expression, the RAG1/2 recombinase is allowed access to only a limited selection of RSSs at given developmental stages (2). Compared to other parts of *Tcrb*, the

4

DβJβ region is marked by strong binding of RAG1/2 recombinase (22). Regulation of RAG1/2 accessibility to this region is achieved by providing an open chromatin environment, which is marked by active histone modifications H3K27ac and H3K4me3 (25, 26). In addition, H3K4me3 directly recruits RAG2 by interacting with its PHD finger (26). Likewise, other AR loci initiate V(D)J recombination from similar regions with high accessibility to RAG1/2 recombinase, which are collectively termed recombination centers (RC) (22). The *Tcrb*-RC spans both DβJβ clusters (27, 28).

The profile of histone marks and open chromatin at *Tcrb*-RC coincide with features of actively transcribed genes (22, 29). Indeed, in DN cells with unrearranged *Tcrb* there is strong sterile transcription, termed germline transcription, going through both upstream Dβ1Jβ1Cβ1 and downstream Dβ2Jβ2Cβ2 clusters (30, 31). In the Dβ1Jβ1Cβ1 cluster, germline transcription is initiated by its promoter, pDβ1, located upstream of Dβ1 (30). With regard to the Dβ2Jβ2Cβ2 cluster, both regions upstream and downstream of Dβ2 have promoter activity (Fig. 1.2) (31, 32). Compared to the promoters, the enhancer Eβ, which is located between Cβ2 and Vβ31, exhibits broader control over the whole *Tcrb*-RC (Fig.1.2) (33). Without Eβ activity, germline transcription and histone acetylation throughout the *Tcrb*-RC are dramatically decreased (27, 29, 34). By contrast, loss of pDβ1 only affects its regional transcription; histone acetylation beyond its close vicinity is still retained (28, 29). In our previous research, we showed that Eβ physically interacts with pDβ1 to deliver the chromatin remodeling machinery SWI/SNF, which further allows accessibility to RAG1/2 recombinase at the *Tcrb*-RC (29). Thus, cis-elements within the *Tcrb*-RC are also termed accessibility control elements (ACEs) (35).

The pivotal role played by the *Tcrb*-RC in controlling a locus as large and complex as *Tcrb* makes it a phenotypical prototype for the recently developed concept super enhancer (SE) (36). A super enhancer is a cis-element cluster with high levels of H3K27ac, which functions as a cell fate switch in response to master regulators of transcription (37). Specifically, during T cell fate commitment, V(D)J recombination at *Tcrb* is initiated by activating Eβ within the *Tcrb*-RC (22, 24). Eβ responds to induced expression of ETS1, which brings constitutive RUNX1 to the two essential ETS1-RUNX1 composite motifs in Eβ (Fig. 1.2) (38). Loss of either ETS1 or RUNX1 binding completely abolishes Eβ activity (39). Notably, TF binding clusters harboring both ETS1 and RUNX1 sites exist widely in mouse and human enhancers (40–42). Of these clusters, composite ETS1-RUNX1 motifs with overlapping binding sites exist mainly in enhancers close to lymphocyte-specific genes, while other TF clusters harboring independent ETS1 and RUNX1 sites exist primarily in promoters (40). Therefore, a likely genetic activation network during T cell fate commitment is controlling many different loci at the same time through ETS1-RUNX1 composite binding. Precise regulation of such binding is crucial for T cell development, since aberrant genome-wide ETS1 and RUNX1 co-binding pattern is associated with autoimmune-related T cells (43).
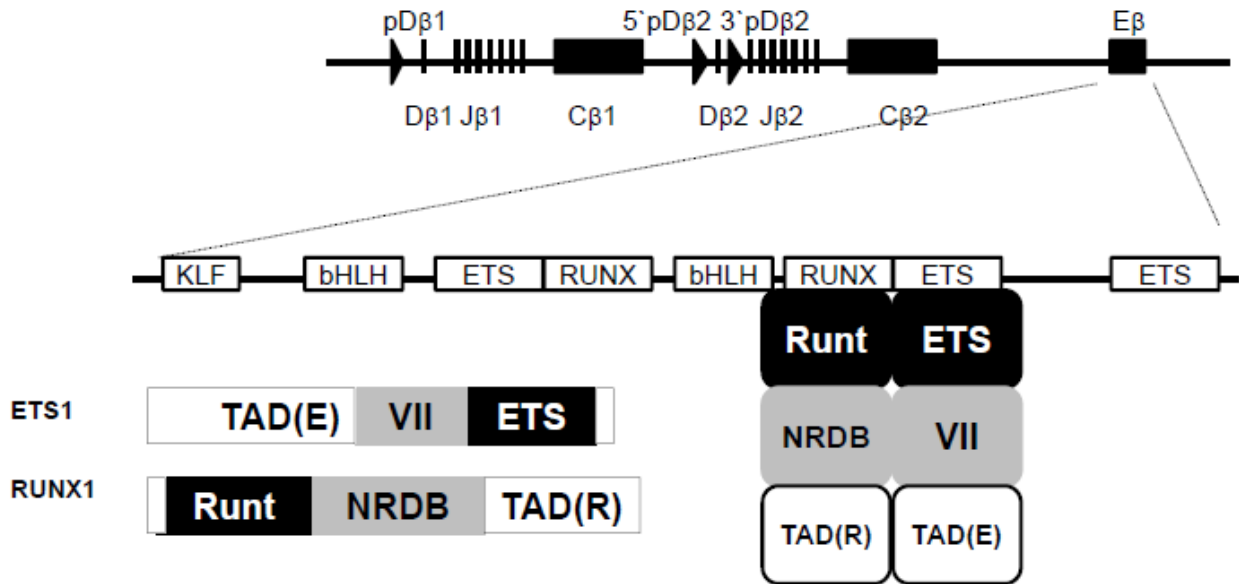
Figure 1.2 The mouse *Tcrb*-RC harbors two composite ETS1-RUNX1 binding sites.

The top panel is a schematic representation of the mouse *Tcrb*-RC with gene segments and cis-elements. The black triangles are promoters for Dβ1 (pDβ1), upstream of Dβ2 (5`pDβ2), and downstream of Dβ2 (3`pDβ2). The black square is enhancer Eβ. The middle panel is a zoom-in showing main Eβ transcription binding site. Note that an additional non-essential RUNX1 binding site immediately upstream of the second ETS1-RUNX1 motif is not shown. The bottom panel shows the domain structures of ETS1 and RUNX1 protein. ETS1 DBD and Runt DBD are shown with black boxes. Exon VII domain and NRDB domain are shown with grey boxes. The TAD for ETS1 and RUNX1 are labeled as TAD(E) and TAD(R) and shown with white boxes. The lower left panel shows co-binding of ETS1 and RUNX1 at one of their composite sites.

1.4   ETS1 and RUNX1 are essential for T cell development

While ETS1 and RUNX1 cooperatively function at their composite sites, both of them have

essential roles at certain stages during hematopoiesis. They may work independently or

cooperatively with other proteins, which is a common strategy to realize complex gene

regulation.

ETS1 is a representative member of the ETS protein family, all of which share similar ETS DNA binding domains (DBD) recognizing highly homologous binding sites (44, 45). There are more than 20 other ETS family members with different domain topologies and expression patterns (45). For example, PU.1 and SPIB, are primarily expressed in B lymphocytes (45). The closest family member to ETS1 is ETS2, which is nearly identical to ETS1 in amino acid (aa) sequence but is expressed primarily in placenta and hair follicles (46). By contrast, ETS1 is expressed primarily in hematopoietic cells, especially T lineage cells (47). ETS1 knock-out mice are viable, with aberrant thymic differentiation and greatly reduced T cell cellularity, as opposed to its minor disruption to B lineage cells (48).

The predominant isoform of mouse ETS1 is the 440-aa protein, termed p54 (Fig1.2), which accounts for around 90% of total endogenous ETS1 (49). The second abundant isoform is p42, which accounts for around 10% of total ETS1 (50). Other isoforms, p68 and p27, are much less expressed (44). Compared to p54, p42 lacks the exon VII domain, which inhibits ETS1 DNA binding after being phosphorylated by CaMKII or MLCK (51, 52). Thus, p42 has higher affinity to ETS1 binding sites than p54 (50). The DBD in ETS1 is a winged helix-turn-helix domain, whose DNA binding capacity is masked by two N-terminus and two C-terminus alpha helices (44). Therefore, although ETS1 is capable of DNA binding by itself, binding is greatly enhanced if cooperating factors partners counteract the auto-inhibition of ETS domain (44).

Compared to ETS1, RUNX1 belongs to a much smaller protein family with two additional members, RUNX2 and RUNX3 (53). This family is defined by their common DBD, termed a Runt domain (54). These proteins are also called core binding factor α (CBFα), since they usually form heterodimers with another non-DNA binding protein, CBFβ (55). CBFβ binding induces conformational changes of the Runt domain, which greatly enhance its DNA binding capacity (55). RUNX family members are similar to each other in terms of aa composition, but differ from each other in expression patterns (56). RUNX2 is primarily expressed in osteoblasts, whereas RUNX3 is essential for CD8$^+$ SP T cells and dorsal root ganglion neurons (56). Compared to RUNX2 and RUNX3, RUNX1 is essential from the early stage of hematopoiesis (57). Loss of RUNX1 leads to embryonic lethality due to a loss of hematopoiesis (57). Conditional knock-out of RUNX1 at specific stages of lymphocyte development abolishes affected lineages (57). There are three splicing isoforms of mouse RUNX1 (58). RUNX1a shares the same promoter with RUNX1b and is a truncated version without transactivation capacity (58). RUNX1c is transcribed from a different upstream promoter from that of RUNX1b, and shares the same downstream exons with RUNX1b (58). The 451-aa RUNX1b is the major isoform expressed in hematopoiesis lineages (Fig. 1.2) (58).

The common theme of individual ETS1 and RUNX1 DNA binding is auto-inhibition, which also is relevant to binding at composite sites. Either ETS1 or RUNX1 by themselves bind poorly to the composite site (39, 59, 60). By contrast, when the negative regulation of DNA binding (NRDB) domain of RUNX1 physically interacts with the exon VII domain of ETS1, conformational changes are induced in both the ETS and Runt domains, which greatly enhance their co-binding affinity (59, 60). While previous research has established the mechanism for

ETS1-RUNX1 cooperative binding, little has been done to elucidate how these two factors contribute to gene activation through their composite motifs. Both proteins may physically recruit other transactivators. The transactivation domain (TAD) of ETS1 interacts with ATF2 and the TAZ domain of CBP (41, 61), while RUNX1 interacts with p300 and the nuclear factor ALY (62, 63). Moreover, Runx1 is a context-dependent repressor for transcription. Its C-terminus VWRPY motif interacts with Groucho/TLE inhibitors at the CD4 silencer, which abrogates CD4 expression (64). Given the complexity of potential downstream effects at ETS1-RUNX1 composite motifs, it is necessary to experimentally determine specific roles of ETS1 and RUNX1 proteins on these motifs.

1.5   Conclusion and scope of thesis

ETS1 and RUNX1 are crucial co-regulators at the Eβ-driven *Tcrb*-RC and other T-lineage super enhancers harboring similar composite ETS1-RUNX1 motifs. Developmentally upregulated ETS1 expression, combined with constitutive hematopoietic RUNX1 expression, promotes cell fate commitment during T cell development via the genome-wide distribution of their co-binding motifs. However, at the molecular level, the specific functions performed by ETS1 and RUNX1 at these super enhancers are not clear due to a lack of experimental dissection.

Indeed, it was difficult to specifically recruit only ETS1 or RUNX1 to composite sites due to the cooperative nature of binding. In this study, I used *Tcrb*-RC as a model locus to convert ETS1 and RUNX1 sites to Gal4 binding sites, and examined the restoration of mutant Eβ activity by Gal4DBD fusion proteins. In Chapter 2, I test the feasibility of this approach using luciferase reporters and found that both Gal4-ETS1 and Gal4-RUNX1 fusion proteins are able to rescue their mutant enhancers. However, in cell models with knock-in alleles, only Gal4-RUNX1 activates its corresponding mutant Eβ. In Chapter 3, using a cell model with Eβ harboring double binding site mutations, I show that RUNX1 is sufficient to activate *Tcrb*-RC independently from ETS1. I further dissect RUNX1 domain functions using Gal4 fusion with truncated versions of the transcription factor. I find that the RUNX1-TAD is sufficient for proximal gene activation. To reach a distal promoter, the presence of either ETS1 or the NRDB domain is necessary. For the first time, this study also demonstrates spatial and temporal heterogeneity of different epigenetic and genetic processes during locus activation. In Chapter 2, I show that factor binding across *Tcrb*-RC saturates faster than restoration of Eβ activity. In Chapter 3, I show that while Eβ activity is fully restored at both DβJβ clusters, there is much less restoration of active histone marks in the upstream Dβ1Jβ1 cluster than in the downstream Dβ2Jβ2 cluster. Thus, the overall study provides mechanistic insight into super enhancer activation directed by ETS1-RUNX1 motifs. Moreover, the methodology used in this study will be applicable to dissect other complex mechanisms of gene regulation.

1.6   References

1. Jung, D., and F. W. Alt. 2004. Unraveling V(D)J Recombination: Insights into Gene Regulation. *Cell* 116: 299–311.

2. Gellert, M. 2002. V(D)J Recombination: RAG Proteins, Repair Factors, and Regulation. *Annu. Rev. Biochem.* 71: 101–132.

3. Sebzda, E., S. Mariathasan, T. Ohteki, R. Jones, M. F. Bachmann, and P. S. Ohashi. 1999. SELECTION OF THE T CELL REPERTOIRE. *Annu. Rev. Immunol.* 17: 829–874.

4. Pieper, K., B. Grimbacher, and H. Eibel. 2013. B-cell biology and development. *J. Allergy Clin. Immunol.* 131: 959–971.

5. Germain, R. N. 2002. T-cell development and the CD4–CD8 lineage decision. *Nat. Rev. Immunol.* 2: 309–322.

6. Blum, J. S., P. A. Wearsch, and P. Cresswell. 2013. Pathways of Antigen Processing. *Annu. Rev. Immunol.* 31: 443–473.

7. Weber, B. N., A. W.-S. Chi, A. Chavez, Y. Yashiro-Ohtani, Q. Yang, O. Shestova, and A. Bhandoola. 2011. A critical role for TCF-1 in T-lineage specification and differentiation. *Nature* 476: 63–8.

8. Schmitt, T. M., R. F. de Pooter, M. A. Gronski, S. K. Cho, P. S. Ohashi, and J. C. Zúñiga-Pflücker. 2004. Induction of T cell development and establishment of T cell competence from embryonic stem cells differentiated in vitro. *Nat. Immunol.* 5: 410–417.

9. Rothenberg, E. V., J. E. Moore, and M. A. Yui. 2008. Launching the T-cell-lineage developmental programme. *Nat. Rev. Immunol.* 8: 9–21.

10. Rothenberg, E. V, A. Champhekar, S. Damle, M. M. Del Real, H. Y. Kueh, L. Li, and M. A. Yui. 2013. Transcriptional establishment of cell-type identity: dynamics and causal mechanisms of T-cell lineage commitment. *Cold Spring Harb. Symp. Quant. Biol.* 78: 31–41.

11. Orkin, S. H., L. I. Zon, L. Ye, H. Huang, X. He, W. G. Tong, J. Ross, J. Haug, T. Johnson, J. Q. Feng, and et al. 2008. Hematopoiesis: An Evolving Paradigm for Stem Cell Biology. *Cell* 132: 631–644.

12. Rothenberg, E. V., and T. Taghon. 2005. MOLECULAR GENETICS OF T CELL DEVELOPMENT. *Annu. Rev. Immunol.* 23: 601–649.

13. Godfrey, D. I., J. Kennedy, T. Suda, and A. Zlotnik. 1993. A developmental pathway involving four phenotypically and functionally distinct subsets of CD3-CD4-CD8- triple-negative adult mouse thymocytes defined by CD44 and CD25 expression. *J. Immunol.* 150.

14. Raulet, D. H. 1989. The Structure, Function, and Molecular Genetics of the gamma/delta T Cell Receptor. *Annu. Rev. Immunol.* 7: 175–207.

15. Takahama, Y. 2006. Journey through the thymus: stromal guides for T-cell development and selection. *Nat. Rev. Immunol.* 6: 127–135.

16. von Boehmer, H., I. Aifantis, O. Azogui, J. Feinberg, C. Saint-Ruf, C. Zober, C. Garcia, and J. Buer. 1998. Crucial function of the pre-T-cell receptor (TCR) in TCR beta selection, TCR beta allelic exclusion and alpha beta versus gamma delta lineage commitment. *Immunol. Rev.* 165: 111–9.

17. Sleckman, B. P., J. R. Gorman, and F. W. Alt. 1996. ACCESSIBILITY CONTROL OF ANTIGEN-RECEPTOR VARIABLE-REGION GENE ASSEMBLY: Role of *cis* -Acting Elements. *Annu. Rev. Immunol.* 14: 459–481.

18. Kapitonov, V. V, and J. Jurka. 2005. RAG1 Core and V(D)J Recombination Signal Sequences Were Derived from Transib Transposons. *PLoS Biol.* 3: e181.

19. Sleckman, B. P., C. H. Bassing, F. W. Alt, M. M. Hughes, M. D'Auteuil, T. D. Wehrly, B. B. Woodman, F. Gärtner, J. M. White, and L. Davidson. 2000. Recombination signal sequences restrict chromosomal V(D)J recombinationbeyond the 12/23 rule. *Nature* 405: 583–586.

20. Rooney, S., J. Chaudhuri, and F. W. Alt. 2004. The role of the non-homologous end-joining pathway in lymphocyte development. *Immunol. Rev.* 200: 115–131.

21. Glusman, G., L. Rowen, I. Lee, C. Boysen, J. C. Roach, A. F. A. Smit, K. Wang, B. F. Koop, and L. Hood. 2001. Comparative Genomics of the Human and Mouse T Cell Receptor Loci. *Immunity* 15: 337–349.

22. Schatz, D. G., and Y. Ji. 2011. Recombination centres and the orchestration of V(D)J recombination. *Nat. Rev. Immunol.* 11: 251–263.

23. Acha-Orbea, H., A. N. Shakhov, L. Scarpellino, E. Kolb, V. Müller, A. Vessaz-Shaw, R. Fuchs, K. Blöchlinger, P. Rollini, J. Billotte, M. Sarafidou, H. R. MacDonald, and H. Diggelmann. 1991. Clonal deletion of V β814-bearing T cells in mice transgenic for mammary tumour virus. *Nature* 350: 207–211.

24. Jackson, A. M., and M. S. Krangel. 2006. Turning T-cell receptor ? recombination on and off: more questions than answers. *Immunol. Rev.* 209: 129–141.

25. McMurry, M. T., and M. S. Krangel. 2000. A Role for Histone Acetylation in the Developmental Regulation of V(D)J Recombination. *Science (80-.).* 287.

26. Matthews, A. G. W., A. J. Kuo, S. Ramón-Maiques, S. Han, K. S. Champagne, D. Ivanov, M. Gallardo, D. Carney, P. Cheung, D. N. Ciccone, K. L. Walter, P. J. Utz, Y. Shi, T. G. Kutateladze, W. Yang, O. Gozani, and M. A. Oettinger. 2007. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature* 450: 1106–1110.

27. Majumder, K., O. I. Koues, E. A. W. Chan, K. E. Kyle, J. E. Horowitz, K. Yang-Iott, C. H. Bassing, I. Taniuchi, M. S. Krangel, and E. M. Oltz. 2015. Lineage-specific compaction of Tcrb requires a chromatin barrier to protect the function of a long-range tethering element. *J. Exp. Med.* 212.

28. Majumder, K., L. J. Rupp, K. S. Yang-Iott, O. I. Koues, K. E. Kyle, C. H. Bassing, and E. M. Oltz. 2015. Domain-Specific and Stage-Intrinsic Changes in Tcrb Conformation during Thymocyte Development. *J. Immunol.* 195.

29. Osipovich, O., R. Milley Cobb, K. J. Oestreich, S. Pierce, P. Ferrier, and E. M. Oltz. 2007. Essential function for SWI-SNF chromatin-remodeling complexes in the promoter-directed assembly of Tcrb genes. *Nat. Immunol.* 8: 809–816.

30. Sikes, M. L., R. J. Gomez, J. Song, and E. M. Oltz. 1998. A developmental stage-specific promoter directs germline transcription of D beta J beta gene segments in precursor T lymphocytes. *J. Immunol.* 161: 1399–405.

31. McMillan, R. E., and M. L. Sikes. 2008. Differential activation of dual promoters alters Dbeta2 germline transcription during thymocyte development. *J. Immunol.* 180: 3218–28.

32. McMillan, R. E., and M. L. Sikes. 2009. Promoter activity 5′ of Dβ2 is coordinated by E47, Runx1, and GATA-3. *Mol. Immunol.* 46: 3009–3017.

33. Mathieu, N., W. M. Hempel, S. Spicuglia, C. Verthuy, and P. Ferrier. 2000. Chromatin Remodeling by the T Cell Receptor (Tcr)-β Gene Enhancer during Early T Cell Development. *J. Exp. Med.* 192.

34. Oestreich, K. J., R. M. Cobb, S. Pierce, J. Chen, P. Ferrier, and E. M. Oltz. 2006. Regulation of TCRβ Gene Assembly by a Promoter/Enhancer Holocomplex. *Immunity* 24: 381–391.

35. Cobb, R. M., K. J. Oestreich, O. A. Osipovich, and E. M. Oltz. 2006. Accessibility Control of V(D)J Recombination. *Adv. Immunol.* 91: 45–109.

36. Whyte, W. A., D. A. Orlando, D. Hnisz, B. J. Abraham, C. Y. Lin, M. H. Kagey, P. B. Rahl, T. I. Lee, and R. A. Young. 2013. Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. *Cell* 153: 307–319.

37. Niederriter, A. R., A. Varshney, S. C. J. Parker, and D. M. Martin. 2015. Super Enhancers in Cancers, Complex Disease, and Developmental Disorders. *Genes (Basel).* 6: 1183–200.

38. Carvajal, I. M., and R. Sen. 2000. Functional analysis of the murine TCR beta-chain gene enhancer. *J. Immunol.* 164: 6332–9.

39. Kim, W. Y., M. Sieweke, E. Ogawa, H. J. Wee, U. Englmeier, T. Graf, and Y. Ito. 1999. Mutual activation of Ets-1 and AML1 DNA binding by direct interaction of their autoinhibitory domains. *EMBO J.* 18: 1609–20.

40. Hollenhorst, P. C., K. J. Chandler, R. L. Poulsen, W. E. Johnson, N. A. Speck, and B. J. Graves. 2009. DNA Specificity Determinants Associate with Distinct Transcription Factor Functions. *PLoS Genet.* 5: e1000778.

41. Giese, K., C. Kingsley, J. R. Kirshner, and R. Grosschedl. 1995. Assembly and function of a TCR alpha enhancer complex is dependent on LEF-1-induced DNA bending and multiple protein-protein interactions. *Genes Dev.* 9: 995–1008.

42. Erman, B., M. Cortes, B. S. Nikolajczyk, N. A. Speck, and R. Sen. 1998. ETS-core binding factor: a common composite motif in antigen receptor gene enhancers. *Mol. Cell. Biol.* 18: 1322–30.

43. Peeters, J. G. C., S. J. Vervoort, S. C. Tan, G. Mijnheer, S. de Roock, S. J. Vastert, E. E. S. Nieuwenhuis, F. van Wijk, B. J. Prakken, M. P. Creyghton, P. J. Coffer, M. Mokry, and J. van Loosdregt. 2015. *Inhibition of Super-Enhancer Activity in Autoinflammatory Site-Derived T Cells Reduces Disease-Associated Gene Expression,. Cell reports* 12.12 (2015): 1986-1996.

44. Garrett-Sinha, L. A. 2013. Review of Ets1 structure, function, and roles in immunity. *Cell. Mol. Life Sci.* 70: 3375–3390.

45. Hollenhorst, P. C., L. P. McIntosh, and B. J. Graves. 2011. Genomic and Biochemical Insights into the Specificity of ETS Transcription Factors. *Annu. Rev. Biochem.* 80: 437–471.

46. Yamamoto, H., M. L. Flannery, S. Kupriyanov, J. Pearce, S. R. McKercher, G. W. Henkel, R. A. Maki, Z. Werb, and R. G. Oshima. 1998. Defective trophoblast function in mice with a targeted mutation of Ets2. *Genes Dev.* 12: 1315–26.

47. Muthusamy, N., K. Barton, and J. M. Leiden. 1995. Defective activation and survival of T cells lacking the Ets-1 transcription factor. *Nature* 377: 639–642.

48. Bories, J.-C., D. M. Willerford, D. Grévin, L. Davidson, A. Camus, P. Martin, D. Stéhelin, and F. W. Alt. 1995. Increased T-cell apoptosis and terminal B-cell differentiation induced by inactivation of the Ets-1 proto-oncogene. *Nature* 377: 635–638.

49. Dittmer, J. 2003. The Biology of the Ets1 Proto-Oncogene. *Mol. Cancer* 2: 29.

50. Leprivier, G., D. Baillat, A. Begue, B. Hartmann, and M. Aumercier. 2009. Ets-1 p51 and p42 isoforms differentially modulate Stromelysin-1 promoter according to induced DNA bend orientation. *Nucleic Acids Res.* 37: 4341–4352.

51. Cowley, D. O., and B. J. Graves. 2000. Phosphorylation represses Ets-1 DNA binding by reinforcing autoinhibition. *Genes Dev.* 14: 366–76.

52. Lindemann, R. K., M. Braig, P. Ballschmieter, T. A. Guise, A. Nordheim, and J. Dittmer. 2003. Protein kinase Calpha regulates Ets1 transcriptional activity in invasive breast cancer cells. *Int. J. Oncol.* 22: 799–805.

53. Cohen Jr., M. M. 2009. Perspectives on *RUNX* genes: An update. *Am. J. Med. Genet. Part A* 149A: 2629–2646.

54. Levanon, D., V. Negreanu, Y. Bernstein, I. Bar-Am, L. Avivi, and Y. Groner. 1994. AML1, AML2, and AML3, the Human Members of the runt domain Gene-Family: cDNA Structure, Expression, and Chromosomal Localization. *Genomics* 23: 425–432.

55. Huang, G., K. Shigesada, K. Ito, H. J. Wee, T. Yokomizo, and Y. Ito. 2001. Dimerization with PEBP2beta protects RUNX1/AML1 from ubiquitin-proteasome-mediated degradation. *EMBO J.* 20: 723–33.

56. Bae, S.-C., and Y. H. Lee. 2006. Phosphorylation, acetylation and ubiquitination: The molecular basis of RUNX regulation. *Gene* 366: 58–66.

57. Collins, A., D. R. Littman, and I. Taniuchi. 2009. RUNX proteins in transcription factor networks that regulate T-cell lineage choice. *Nat. Rev. Immunol.* 9: 106–115.

58. Challen, G. A., and M. A. Goodell. 2010. Runx1 isoforms show differential expression patterns during hematopoietic development but have similar functional effects in adult hematopoietic stem cells. *Exp. Hematol.* 38: 403–416.

59. Goetz, T. L., T. L. Gu, N. A. Speck, and B. J. Graves. 2000. Auto-inhibition of Ets-1 is counteracted by DNA binding cooperativity with core-binding factor alpha2. *Mol. Cell. Biol.* 20: 81–90.

60. Gu, T. L., T. L. Goetz, B. J. Graves, and N. A. Speck. 2000. Auto-inhibition and partner proteins, core-binding factor beta (CBFbeta) and Ets-1, modulate DNA binding by CBFalpha2 (AML1). *Mol. Cell. Biol.* 20: 91–103.

61. Nelson, M. L., H.-S. Kang, G. M. Lee, A. G. Blaszczak, D. K. W. Lau, L. P. McIntosh, and B. J. Graves. 2010. Ras signaling requires dynamic properties of Ets1 for phosphorylation-enhanced binding to coactivator CBP. *Proc. Natl. Acad. Sci. U. S. A.* 107: 10026–31.

62. Wee, H.-J., D. C.-C. Voon, S.-C. Bae, and Y. Ito. 2008. PEBP2-β/CBF-β–dependent phosphorylation of RUNX1 and p300 by HIPK2: implications for leukemogenesis. *Blood* 112.

63. Bruhn, L., A. Munnerlyn, and R. Grosschedl. 1997. ALY, a context-dependent coactivator of LEF-1 and AML-1, is required for TCRalpha enhancer function. *Genes Dev.* 11: 640–53.

64. Taniuchi, I., M. Osato, T. Egawa, M. J. Sunshine, S.-C. Bae, T. Komori, Y. Ito, and D. R. Littman. 2002. Differential Requirements for Runx Proteins in CD4 Repression and Epigenetic Silencing during T Lymphocyte Development. *Cell* 111: 621–633.

Chapter 2 : ETS1 and RUNX1 have distinct transactivation capacities on chromosomal and ectopic enhancers

2.1 Abstract

V(D)J recombination at the mouse T cell receptor β (*Tcrb*) is initiated from its recombination center (RC). Accessibility of the recombination machinery to the *Tcrb*-RC is tightly controlled by its enhancer, Eβ. Two transcription factors, ETS1 and RUNX1, bind cooperatively to the two composite motifs within Eβ, and are essential for Eβ function. Upregulated ETS1 expression during T cell lineage commitment recruits constitutive RUNX1 to Eβ. Since both proteins are important transactivators during certain stages of lymphocyte development, an important question is whether RUNX1 and ETS1 play independent roles during *Tcrb*-RC activation that can be dissected from their cooperative binding. To examine individual functions of ETS1 and RUNX1 at Eβ, I designed an inducible system and tested whether mutant versions of Eβ can be rescued by Gal4 fusion proteins. In this chapter, I first show that ETS1 or RUNX1 are sufficient to independently activate Eβ mutants on extrachromosomal luciferase substrates. However, in a cell model harboring knock-in Gal4 binding sites, ETS1 fails to activate its own enhancer mutant. By contrast, Gal4 fusions of RUNX1 are sufficient to activate Eβ harboring RUNX site mutations at the endogenous *Tcrb*-RC. I also observed delayed kinetics of germline transcription, a hallmark of Eβ activation, compared to fast transcription factor loading at the *Tcrb*-RC. I have also established a useful cell system to study *Tcrb*-RC activation and its epigenetic revision upon induction of its master enhancer element.

2.2   Introduction

Stepwise rearrangement of gene segments within antigen receptor loci, or V(D)J recombination, is tightly regulated by accessibility of RAG1/2 recombinase to each locus (1, 2). At the mouse T cell receptor beta (*Tcrb*) locus, the two downstream DβJβ clusters are highly enriched with RAG1/2 and undergo D-J recombination in DN thymocytes (2, 3). This region then maintains its accessibility to RAG proteins and allows further V-DJ recombination (3). Therefore, it is termed the *Tcrb* recombination center (*Tcrb*-RC) (2). During T lymphocyte development, highly active germline transcription throughout the *Tcrb*-RC loosens its chromatin structure, which permits accessibility by RAG proteins to the locus (4–7). To precisely control the tissue specificity of such transcription and, in turn, to restrict recombination of *Tcrb* to proper developmental stages, a short and potent enhancer, Eβ, is essential (8). Eβ is directly bound by T-lineage-specific transcription factors GATA3 and ETS1, both of which are upregulated with a network of other proteins induced by the master factor TCF1 (9, 10). Thus, *Tcrb* rearrangement controlled by Eβ is one of many processes essential for T cell development.

Of all the transcription factors that bind Eβ, ETS1 and RUNX1 are closely associated with each other since they form a cooperative complex sharing two composite motifs essential for Eβ function (11). The co-binding strategy is employed by many other loci with similar ETS1-RUNX1 composite motifs, which are widely distributed throughout the mouse genome (12). The

22

stringent requirement for both ETS1 and RUNX1 binding at these sites has raised interests in what individual functions they perform at the loci that they regulate together. When they do not binding cooperatively, both ETS1 and RUNX1 are independent transactivators (13–16). ETS1 deficiency in mice leads to severe defects and low T cellularity (17). RUNX1 deficiency leads to embryonic lethality, due to the complete loss of hematopoietic cell development (18). Conditional knock-out of RUNX1 at different T cell development stages severely blocks or alters lineage commitment (18, 19). Since RUNX1 expression is turned on at the very early stages of hematopoiesis and ETS1 is induced after T lineage commitment (20, 21), at least one possible physiological reason for ETS1-RUNX1 composite site is to regulate RUNX1 recruitment by newly expressed ETS1.

Still, this scenario does not explain the precise roles of ETS1 and RUNX1 in terms of enhancer and locus activation at *Tcrb*. Previous research regarding ETS1-RUNX1 co-binding mainly focused on their DNA binding capacity (22–24). Using *in vitro* assays with purified recombinant truncations of ETS1 and RUNX1, as well as synthetic oligonucleotides, these studies have identified DNA binding (DBDs) and interacting domains of both factors (22–24). ETS1 is equipped with an ETS DBD at its C terminus, which follows an exon VII domain with the capacity to interact with RUNX1 (23, 24). By contrast, the DBD of RUNX1, termed Runt, is situated at its N terminus, which is followed by the ETS1-interacting domain, called negative regulatory region of DNA binding (NRDB) (22, 24). When ETS1 and RUNX1 bind to each other through their interacting domains, the conformational changes in their DBDs counteracts "auto-inhibition" of DNA binding, and greatly increase their affinity to their composite sites (22–24). However, apart from DNA binding, the exact functions of ETS1 and RUNX1 transactivation

domains (TAD) are not well understood in the context of composite site binding. Although ETS1 and RUNX1 were examined for their activities at cis-elements using reporter assays (15, 25), there is no systematic dissection of their individual and cooperative functions at Eβ. In this study, I utilized truncations of ETS1 and RUNX1 proteins fused with Gal4DBD to test their activities on extrachromosomal luciferase substrates with corresponding binding site mutants of Eβ. I found that both ETS1 and RUNX1 independently activate mutant Eβ reporters.

Although reporter assays are informative to demonstrate the relative intensities of transactivation by different transcription factors, further investigation of epigenetic changes over a long distance requires an experimental systems in chromosomal settings. To construct cell and animal models with specific mutations, gene targeting techniques were developed to introduce exogenous vectors carrying the desired mutations into mammalian cells (26). However, traditional targeting is mostly restricted to embryonic stem cells, which permit high levels of homologous recombination (HR) (27). Even in these cells, the efficiency of targeting is still extremely low, so that thousands of clones must be screened to obtain the correct ones (28). To solve this problem, zinc finger nucleases (ZFN) were designed to generate site-specific double strand breaks (DSB), which dramatically increase the chance of HR between genomic DNA and targeting vectors (29). ZFNs are composed of a pair of FokI restriction enzymes fused with three to four zinc finger DBDs (29). Each zinc finger is capable of specifically binding to a nucleotide triplet. Thus, the combined sequence from three to four triplets constitutes a specific 9- to 12-nucloetide binding site for one ZFN subunit. A ZFN pair then binds specifically to an 18- to 24-nucleotide binding site. In this study, I successfully constructed cell models with Eβ mutation at *Tcrb*-RC using ZFN-assisted targeting.

The cell models revealed distinct behaviors by Gal4-ETS1 and –RUNX1 fusion proteins on their mutant enhancers. Unlike luciferase reporter assays, only Ga4-RUNX1 is able to activate Eβ with mutant RUNX1 binding sites. By contrast, Gal4-ETS1 does not activate Eβ with mutant ETS1 binding sites, even when the steady-state levels of the fusion proteins are boosted by inhibition of proteolysis. These results suggest that RUNX1 may be the major transactivator in the ETS1-RUNX1 duo. Also, the difference between results from reporters and knock-in alleles suggests the necessity to validate earlier research on reporter substrates using chromosomal settings. In addition, our inducible gene activation system provides a valuable tool to study kinetics of epigenetic changes at gene loci. Previously, kinetics studies of gene activation were limited to cytokine- and hormone-inducible genes, which respond at the end of long signaling cascades (30). Using similar strategies, it is possible to engineer inducible gene loci for many developmentally important genes, the activation processes of which would otherwise be very difficult to dissect.

2.3   Materials and Methods

Cell culture

The RAG1/P53-deficient thymoma cell line P5424 and its derivative cells were cultured at 37 °C and 5% CO_2 in RPMI 1640 medium supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 0.01% penicillin/streptomycin, and 50 μM β-mercaptoethanol. 293T cells were cultured under the same condition as P5424 cells only except that the medium was DMEM.

Cloning of the luciferase reporters

The promoter-less luciferase reporter plasmid (P-E-) was generated by removing a BglII-HindIII fragment from the pGL3-Promoter (Promega), which was then blunted and re-ligated. The reporter with only small Dβ1 promoter (PsE- or P+E-) was generated by replacing the BglII-HindIII fragment on the pGL3-Promoter plasmid with a 0.46kb PCR-amplified fragment from a previously described minilocus plasmid (31). A 0.57kb Eβ fragment amplified from a pBS vector harboring this enhancer was introduced to the BamHI site on PsE- to generate PsEs (or P+E+). Medium (2.39kb) and long (3.41kb) Eβ fragments were amplified directly from P5424 genomic DNA and ligated to the blunted BamHI ends from cut PsE- to generate PsEm and PsEl reporters. Medium (3.11kb) and long (3.65kb) Dβ1 promoters were amplified directly from P5424 genomic DNA and replaced the KpnI-HindIII fragment on the pGL3-Promoter plasmid to generate PmE- and PlE- reporters. The short, medium and long Eβ fragments were then cloned into the blunted BamHI sites on PmE- and PlE- reporters to generate PmEs, PmEm, PmEl, PlEs,

26

PlEm, and PlEl, respectively. For luciferase reporters with mutant EG, RG, EmRG, and EL enhancers, Gal4 or TF site mutations were first introduced into a 1.85 kb fragment spanning Eβ (HpaI-HindIII), using PCR-assisted mutagenesis with a high-fidelity polymerase according to the manufacturer's protocol (Phusion, NEB M0503). Oligonucleotide sequences used for site-directed mutagenesis and cloning are provided in Table 2.2. Mutagenesis was performed sequentially to alter both ETS1-RUNX1 sites in Eβ. Bacterial colonies from ligated plasmids were first screened by regular PCR. Positive candidates were further digested with restriction enzymes and confirmed by Sanger sequencing. Correct mutant Eβ fragments were then amplified by PCR and cloned into the P+E- reporter.

Construction and expression of Gal4 and LexA fusion proteins

Except the Gal4VP16 cDNA described in previous research (32), all other Gal4-ETS1 and Gal4-RUNX1 fusion proteins were prepared from an acceptor plasmid containing a cDNA for the Gal4 DNA-binding domain (DBD, aa 1-147). Full-length and truncated ETS1 and RUNX1 cDNAs were cloned in-frame to the Gal4DBD cDNA from either its upstream for C-terminus Gal4DBD fusions or its downstream for N-terminus Gal4DBD fusions. Each of these ETS1 and RUNX1 cDNAs was generated by PCR using primers provided in Table 2.2 (ETS1^{GTVE} aa 2-440, ETS1^{GT} aa 2-242, ETS1^{GTV} aa 2-330, ETS1^{TVEG} aa 1-439, ETS1^{TG} 1-242, ETS1^{TVG} aa 1-330, RUNX1^{GRNT} aa 1-451, RUNX1^{GNT} aa 184-451, RUNX1^{GN} aa 182-292, RUNX1^{GT} aa 291-451, RUNX1^{GNAD} aa 184-372, and RUNX1^{GID} aa 371-451). For LexA fusion proteins, ETS^{TVEL} and ETS1^{TVL} cDNAs were cloned by swapping the Gal4DBD cDNA in ETS^{TVEG} and ETS1^{TVG}

27

with PCR-amplified LexA cDNA (aa 1-202). The short Gal4DBD fusion cDNAs for ETS1[TVEG(65)] and ETS1[TVEG(93)] were cloned by swapping the complete Gal4DBD cDNA in ETS1[TVEG] with PCR-amplified cDNA for aa 1-65 and 1-93 of Gal4DBD. The final fused Gal4 and LexA fusion protein cDNAs were then excised and cloned into pcDNA3.1. Each version was validated by sequencing and protein expression in 293T cells. Expression vectors based on pcDNA3.1 were then used in luciferase assay. For constitutive Gal4-ETS1 fusion protein expression in Figs. 2.6, 2.7, and 2.8, fusion protein cDNAs were shuttled into a MSCV retroviral vector with IRES-hCD4 co-expression marker. For inducible expression of RUNX1[GNT], the cDNA was shuttled into the pFLRU:Thy1.1 tet-responsive lentiviral vector. Viruses containing each Gal4-ETS1 and Gal4-RUNX1 expression cassette were packaged as described previously (33) and used to spin-infect $5x10^6$ O3-73 derivative cells. Infected cells were cultured for 10 days, sorted for high human CD4 expression with an EasySep Human CD4+ T Cell Enrichment Kit (Stem Cell Technologies #19052), or for .high Thy1.1 expression by flow cytometry (Anti-Mouse/Rat CD90.1 APC, eBioscience 17-0900-82). Cells carrying Gal4-ETS1 expression cassettes were then ready for harvest or MG132 treatment as in Fig. 2.8A. Cells with inducible RUNX1[GNT] expression cassette were expanded, and induced with 1 μg/mL doxycycline (Dox) for 24 hours prior to molecular analysis.

Luciferase assay

Luciferase assays were perfumed with the Dual-Glo system (Promega Cat#2920). For each reaction, $5x10^6$ P5424 cells were transfected by electroporation (Biorad, 250 V/960 μF) with 5

µg of firefly luciferase reporter plasmid, 5 µg of transcription factor expression plasmid (if any), and 0.1 µg of Renilla luciferase control plasmid and cultured in 5 mL medium for overnight. Cells were harvested 24 hours post transfection, spun down, and resuspended in 20 µL culture medium at room temperature. The suspension was first mixed with 20 µL of the lysis buffer with firefly substrate and incubated at room temperature for 10 minutes before firefly luciferase reading. Then the whole reaction was mixed with 20 µL of the quench buffer with Renilla substrate and incubated at room temperature for another 10 minutes before Renilla luciferase reading.

ZFN cutting and detection of DSB by CEL-I (SURVEYOR) assay

Two Zinc Finger Nuclease pairs were customized for this study (ZFNs; Sigma-Aldrich). Each pair of ZFNs targets a DNA break to a site in *Tcrb*. ZFN1 targets sequence 5'-GGACTTGTCCTTAACTCCCTTTACCTAGCAAGATAGG-3' in the upstream of Dβ1. ZFN2 target is downstream of Eβ (5'-GTCTGTCTGTCCTGCATTCATTGGCTGGCTTTCGCT-3').

CEL-I assay was performed with the SURVEYOR Mutation Detection Kit (Transgenomic Cat# 706025) as the manufacturer's description to detect potential double strand break (DSB). Genomic DNA from samples was extracted with homemade lysis buffer and ethanol precipitation. Phusion PCR was performed with primers listed in Table 2.2 spanning the ZFN cutting site. Total PCR product was hybridized by heating and slow cooling on a thermocycler as

29

recommended by the manual. The hybridized product was then directly divided into halves and used for the non-cutting control and CEL-I cutting reactions at 42 °C for 1 hour. After cutting at the recommended condition in the manual, the samples were run on 2% TAE agarose gel to examine the digested product.

Construction of cell models

In O3-73 cloning, concurrent expression of 5×10^6 P5424 cells were transfected with 5 µg of each ZFN expression plasmid using electroporation (250 V/960 µF). After two days of recovery, cells were cloned by serial dilution and screened by PCR assays for the large *Tcrb*-RC deletion (Table 2.2). Clone O3-73 was selected as a parental cell line for future experiments (allele 1: intact *Tcrb*-RC, allele 2: deletion of Dβ1-Dβ2-Eβ).

To construct cell models within knock-in mutant enhancer, the same mutant fragments used in luciferase reporters are cloned into targeting vectors, which were generated in pBS by sequentially cloning a 3.2 kb 5' homology arm (BamHI-HpaI), a 0.7 kb mutant Eβ (HpaI-MluI), a 1.9 kb loxP:PGK-promoter:Puro:loxP cassette (MluI-NheI), and a 2.1 kb 3` homology arm (NheI-NotI).

For targeting, 5×10^6 O3-73 cells were co-transfected with 20 μg of a targeting vector and 5 μg each of expression plasmids for the ZFN2 pair, which introduce a DNA break ~200 bp downstream of Eβ. Two days post-electroporation, cells were selected with puromycin (1 μg/mL) and cultured for 5-6 d before sub-cloning by limiting dilution. Initially, single cell clones were screened by a PCR assay in which insertion of the targeting vector disrupts amplification (Table 2.2). Candidate recombinant clones were then screened by a long-range PCR assay for proper integration (Table 2.2), which was verified by sequencing. Successfully targeted clones were transfected with 1 μg of human CD4 expression plasmid and 10 μg of CRE expression plasmids and purified for CD4 expression. After one week, cultures were subcloned, screened for deletion of the drug marker cassette, and mutant alleles were verified by sequencing.

RT-qPCR

Total RNA was extracted by TRIzol (Roche #1697478) from O3-73 and its derivatives. In each reaction, 1 μg total RNA was digested by RQ1 RNase-free DNase (Promega M199A) in 20 μL total volume, as recommended by the manufacturer. The treated RNA was split equally for cDNA preparation with M-MuLV Reverse Transcriptase (NEB M0253) or used as a no RT control, with random hexamers. Real-time qPCR was performed with SYBR Green JumpStart Taq ReadyMix (Sigma S4438) using primers and conditions provided in Table 2.1.

Western blot

O3-73 cell lysates were prepared in RIPA buffer ($2x10^7$ cells in 300 μL) and sonicated for 40

cycles (30s, vortexed each 10 cycles, Diagenode Bioruptor Pico). Sonicated lysates were used

mixed with 2x Laemmli buffer, heated at 95 °C for 10 minutes, and chilled before loading to

12% PAGE gel (10 μL). Gels were transferred to PVDF membranes, incubated with detection

antibodies and visualized by ECL substrate (Thermo Fisher Scientific #34087). Primary

antibodies were used at a concentration of 1 μg/mL for protein detection and obtained from the

following sources: rabbit anti-Gal4DBD antibody (Santa Cruz sc-577x), rabbit anti-ETS1 (clone

N276; Santa Cruz sc-111), and mouse anti-GAPDH (clone G-9; Santa Cruz sc-365062).

Secondary antibodies, diluted $1:10^4$, were goat anti-rabbit IgG-HRP (Santa Cruz sc-2004) and

goat anti-mouse IgG-HRP (Promega W4021).

ChIP-qPCR

ChIP assays were performed as described previously (34). Sheared chromatin from O3-73 and its

derivatives were prepared after fixation with 1% formaldehyde and sonication with Diagenode

Bioruptor Pico for 20 cycles (30 s each, vortexed every 5 cycles). ChIP antibodies (all rabbit IgG

isotypes) were coupled to Protein A Dynabeads and were obtained from the following sources:

RUNX1 C terminus was provided by Dr. Takeshi Egawa (35). ETS-1 C20 (Santa Cruz sc-350x)

and normal rabbit IgG (Santa Cruz sc-2027x) were commercially available. Eluted ChIP and

input DNAs were purified with Qiagen QIAquick PCR purification kit and analyzed by RT-qPCR using primers and conditions provided in Table 2.1.

2.4 Results

Both ETS1 and RUNX1 activate luciferase reporters with mutant Eβ in precursor T cells

To functionally dissect the roles of ETS1 and RUNX1 in Eβ activation, I needed a cell model in which the Eβ-dependent activity of *Tcrb*-RC mimics that in primary mouse T cells. I chose the P53-/- Rag1-/- thymoma cell line, P5424, which maintains unrearranged *Tcrb*-RC and active germline transcription through this region (11, 36–38). To test the dependence of Eβ activity on its ETS1-RUNX1 binding motifs in P5424, I constructed a series of luciferase reporter sharing the same promoter for Dβ1 and transfected them into P5424 (Fig. 2.1). As expected, luciferase signal driven by wild type (WT) Eβ was ten-fold higher than the promoter only vector, which confirmed Eβ activity in P5424 cells. Since there are two ETS1-RUNX1 composite sites in Eβ, I converted either both of the RUNX1 or both of the ETS1 binding sites to Gal4 sites (RG or EG), which completely abolished the activity of Eβ (Fig. 2.1). In contrast, mutating only one ETS1 binding site (E1G and E2G) only partially crippled Eβ activity, suggesting that the two ETS1-RUNX1 composite sites work independently from each other (Fig. 2.1). These results are consistent with previous findings on Eβ composition (22).

To test whether ETS1 or RUNX1 proteins are sufficient to rescue their mutant enhancers, I co-transfected luciferase reporters with EG or RG mutant enhancers using Gal4 fusion protein expression vectors into P5424 cells (Fig. 2.2). As a positive control for rescue by Gal4 fusion proteins, Gal4VP16, a potent transactivator, induced both EG and RG mutant enhancers to at

least ten-fold higher than the physiological level of WT Eβ (Fig. 2.2). However, when the full-length RUNX1 or ETS1 proteins were fused to Gal4 DNA binding domain (DBD), the fusion proteins (RUNX1^{GRNT} or ETS1^{TVEG}/ETS1^{GTVE}) failed to fully restore RG or EG mutant enhancer activity (Fig. 2.2). By contrast, if the Runt or ETS DBDs were replaced by Gal4DBD, the fusion proteins (RUNX1^{GNT} or ETS1^{TVG}) were able to fully restore mutant RG or EG enhancer activity (Figs. 2.2 and 2.3). These results suggest that apart from inhibition of DNA binding (22–24), the Runt or ETS DBDs also inhibit the activity of transactivation domains (TAD) on reporter substrates. Notably, both ETS1^{GTV} and ETS1^{TVG} were equally or more competent than ETS1^{GTVE} and ETS1^{TVEG}, suggesting that activity of fusion proteins rely more on their composition than their domain topology (Fig. 2.2). For ETS1, its core competence of activity on Eβ is within its TAD, since both ETS1^{GT} and ETS1^{TG} restored EG activity to the WT level (Fig. 2.2). Similarly, RUNX1 TAD by itself (RUNX1^{GT}) is also sufficient to activate mutant RG enhancer (Fig. 2.4).

The sufficiency of Gal4 fusion proteins to rescue their corresponding EG and RG mutant enhancers does not exclude the potential involvement of endogenous RUNX1 or ETS1 binding at the intact halves of the composite sites. To further test the independent function of ETS1 and RUNX1 proteins, I engineered EGRm and EmRG mutant enhancers in which the core RUNX1 and ETS1 binding nucleotides were mutated in addition to existing EG and RG mutation, respectively (Fig. 2.3B). Surprisingly, ETS1^{TVG} failed to activate EGRm reporter as much as it did to EG reporter (Fig. 2.3B). Therefore, intact RUNX1 binding sites are required for full activity of ETS1^{TVG} at the mutant ETS1 sites. This observation led to a hypothesis that the presence of RUNX1 protein is required for activation of EG by Gal4-ETS1 fusion protein. However, it is challenged by the observation that the potent Gal4VP16 activated EmRG to a

similar high level to those of EG and RG, and only activated EGRm to a much lower level which was comparable to that of WT Eβ (Fig. 2.3B). On EmRG and EGRm mutant enhancers, neither Gal4VP16, nor the motifs recruit RUNX1. Thus, the difference between these two mutants is more likely due to changes in the positions of the two binding sites. Moreover, both RG and EmRG mutant enhancers were activated by ETS1^{TVG}, although RUNX1^{GNT} did activate EG, RG, and EmRG more strongly than ETS1^{TVG} (Fig. 2.3B). Taken together, I conclude that to activate Eβ on luciferase reporters, the positions hosting RUNX1 binding sites in the wild-type enhancer, whether mutant or not, must be accessible for transactivators.

In my system, the physiologically relevant scenarios for reporters with RG and EmRG mutant enhancers are activation using Gal4-RUNX1 fusion proteins. Given the importance of RUNX1 binding sites, I continued to investigate how RUNX1 activates Eβ by further dissecting RUNX1 functional domains. In previous *in vitro* research with purified proteins, ETS1-RUNX1 interactions were identified as important features of ETS1-RUNX1 co-binding to increase their affinity (22–24). Thus, it is possible that Gal4-RUNX1 fusion proteins recruit ETS1 via its ETS1-interacting domain NRDB (22–24) and ETS1 is still essential for the activity of Gal4-RUNX1 fusions. Indeed, losing ETS1 binding sites in mutant EmRG enhancer weakened its activation by Gal4-RUNX1 fusions, as opposed to the mutant RG enhancer (green bars vs. blue bars, Fig. 2.4). However, the TAD of RUNX1 by itself was sufficient to activate the mutant EmRG enhancer to WT Eβ level without ETS1 recruited by its DNA motif or NRDB domain (RUNX1^{GT}, Fig. 2.4). By contrast, the NRDB domain of RUNX1 alone did not activate mutant RG or EmRG enhancers, if there was any ETS1 recruited (RUNX1^{GN}, Fig. 2.4). These results suggest that RUNX1 is able to activate Eβ mutant reporters independently from ETS1, although

ETS1 may further assist such activation. Since RUNX1 C-terminus binds Groucho/TLE inhibitors (39, 40), its TAD may be further dissected into activation and inhibition sub-domains (AD and ID). As expected, AD, not ID, is responsible for RUNX1 activity on Eβ reporters (RUNX1^{GNAD} and RUNX1^{GID}, Fig. 2.4).

Collectively, the luciferase reporter assays with different versions of mutant Eβ and Gal4-ETS1/RUNX1 fusion proteins demonstrated that both ETS1 and RUNX1 may rescue their respective Eβ binding sites mutants once actively tethered by Gal4DBD. Meanwhile, RUNX1 behaves more independently than ETS1 due to the importance of its binding sites and its sufficiency without ETS1. To gain more physiological insights, it would be necessary to investigate further questions in chromosomal settings in P5424 cells (Chapter 3).

Generation of cell models harboring mutant versions of Eβ using Zinc Finger Nucleases

To construct P5424 cell models with mutant Eβ, the Oltz lab customized a commercially available Zinc Finger Nuclease (ZFN), which generates a double strand break (DSB) downstream of Eβ (ZFN2, Fig. 2.5A). DSBs greatly increase the efficiency of introducing site-specific homologous recombination (HR) between targeting vectors and chromosomes (41). For the convenience of targeting and later analysis, we first used another ZFN recognizing the upstream of pDβ1 (ZFN1, Fig. 2.5A) together with ZFN2 to delete the whole *Tcrb*-RC on one allele (Allele 1, Fig. 2.5A) while leaving the other allele intact (Allele 2, Fig. 2.5A). This

derivative of P5424 was termed O3-73 and was used as a wild type control in all experiments. In O3-73, when ZFN2 cuts the genomic DNA downstream of Eβ, targeting vectors harboring mutant Eβ and the drug selection cassette, loxP-Puro-loxP, flanked by long homologous arms may crossover with the endogenous locus and yield a new allele with the mutant Eβ and the puromycin-selectable marker (Fig. 2.5A). This allele, when treated with CRE recombinase, will lose the Puro cassette and leave only one loxP site (Fig. 2.5A).

Since the EG, RG, and EmRG mutant enhancers were shown to be activated by their respective Gal4-fusion proteins, I went on to make the corresponding mutant knock-in alleles (Figs. 2.5, 2.14, and 2.15). Cutting by ZFN2 was validated by CEL-I assay (Figs. 2.13A, 2.14A, and 2.15A). In the CEL-I assay, cut alleles repaired by non-homologous end joining (NHEJ) pathways harbored mismatches compared to the uncut alleles. Thus, PCR products spanning the cut sites were different for cut and uncut alleles. When they hybridized, they formed a "bubble" for CEL-I enzyme to recognize, which was cut into two smaller fragments. As expected, only samples with ZFN2 expression had CEL-I cutting products (Figs. 2.13A, 2.14A, and 2.15A).

During the cloning process, not all cells in culture undergo chromosomal modifications by HR or CRE recombination, thus the resulted cells had to be sub-cloned and screened for correct clones. The first step of screening was to check if the loxP-Puro-loxP cassette successfully replaced the ZFN2 cutting site (Figs. 2.5B, 2.14B, and 2.15B). In this assay, a conventional PCR spanning the ZFN2 cutting site lost its signal if the site was extended by the Puro cassette to much larger than

it could amplify. Indeed, I obtained clones with either no product or non-specific bands, which proceeded to further validation (Figs. 2.5B, 2.14B, and 2.15B).

Replacement of the ZFN2 site by the Puro cassette excluded clones without HR, but it could not distinguish wrong crossovers from correct ones. When crossover happened downstream of Eβ and in the upstream of ZFN2 cutting site, the new allele would have wild type Eβ and the Puro cassette (Wrong crossover, Fig. 2.5C). Thus it would be necessary to use PCRs specific to wild type and mutant Eβ to further screen clones. For example, in EG mutant cell cloning, two clones were identified to bear the Puro cassette (marked with * and #, Fig. 2.5B). However, when tested with wild type- and mutant-specific PCRs, only one clone (marked with *, Fig. 2.5C) carried the correct mutation and the other was wild type (marked with #, Fig. 2.5C). Mutation of Eβ was further validated by a functional readout, germline transcription through Jβ1.2 region (Fig. 2.5C). Alternatively, we used this functional readout as a screen for other mutant Eβ alleles (Fig. 2.14C) since it was difficult to find mutant-specific primers for genomic PCR. For the genomic PCR assay, a stronger and yet more difficult strategy was using a forward primer in the upstream of the left homologous arm together with wild type- or mutant-specific reverse primers, which could further ensure the intactness of the resulted alleles by long-range (3 kb) PCR products (Fig. 2.13C). During EmRG cell cloning, there were only two clones with the Puro cassette (Fig. 2.15B). Both of them passed the more stringent validation by long-range PCR (Fig. 2.15C). For all EG, RG, and EmRG clones passing validation of mutant Eβ, their exact genomic sequences were eventually confirmed by Sanger sequencing.

Cells with correct mutant Eβ sequences then went through another round of screening after CRE deletion of the Puro cassette (Figs. 2.5D, 2.14D, and 2.15D). The screening strategy using conventional PCR was similar to that for detection of replacement of the ZFN2 site by the loxP-Puro-loxP cassette. The difference was that positive clones would show an amplicon on gel instead of losing it. In addition, this new amplicon was slightly larger than that from O3-73 cells due to the extra sequence from the residual loxP site (Figs. 2.5D, 2.14D, and 2.15D). After screening, all correct EG, RG, and EmRG clones were further validated by Sanger sequencing. With these cells, I proceeded to test the chromosomal behaviors of Gal4 fusion proteins.

Gal4-ETS1 fusion protein fails to rescue mutant Eβ in *Tcrb*-RC

Rescuing of the EG mutant enhancer activity on luciferase reporters by Gal4-ETS1 fusion proteins inspired me to further test similar scenarios in a chromosomal setting. Unlike luciferase reporters, EG mutant cells have the whole *Tcrb*-RC, with promoters positioned further upstream from the EG mutant enhancer. To express the fusion transcription factors, I introduced their cDNAs into EG mutant cells using a retroviral transduction system. As expected, EG mutant enhancer completely abolished germline transcription through the Jβ1.2 and Jβ2.1 regions (Fig. 2.6). To our surprise, ETS1^{TVG}, which maintains the domain topology closest to WT ETS1 in all Gal4-ETS1 fusion proteins, failed to activate Jβ1.2 or Jβ2.1 germline transcription (Fig. 2.6). By contrast, the potent Gal4VP16 did activate EG mutant enhancer for both Jβ1.2 and Jβ2.1 regions, suggesting that this enhancer is accessible and activatable by Gal4 fusion proteins (Fig. 2.6). Meanwhile, the total ETS1 transcription, including both endogenous ETS1 and ectopic ETS1^{TVG}

40

signals, was at doubled in EG cells expression ETS1^{TVG} compared to untransduced cells (Fig. 2.6). This result suggests that, in terms of transcription levels, ETS1^{TVG} is comparable to endogenous ETS1.

However, transcription levels are not necessarily in proportion to steady-state protein levels, which may be affected by translation efficiency and protein degradation (42). Indeed, previous research observed that Gal4DBD fusion led to decreased expression of its fusion proteins (15). To test the influence of Gal4DBD on the steady-state expression levels of ETS1 fusion proteins, I needed a cell model in which ectopic wild type ETS1 protein and Gal4-ETS1 fusions are clearly comparable to each other. Since ETS1 is abundantly expressed in T lineage cells, I used a proB cell line 63-12, which expresses minimal endogenous ETS1, to compare different ectopic proteins (Fig. 2.7). I designed ETS1 fusion proteins with different lengths of additional domains, from the shortest HA tag (ETS1HA) to longer portions of Gal4DBD (ETS1$^{TVEG(65)}$, ETS1$^{TVEG(93)}$, and ETS1$^{TVEG(147)}$, Fig. 2.7A). The HA tag did not change protein expression levels compared to wild type ETS1 (Fig. 2.7B). By contrast, the 65 aa Gal4DBD portion slightly decreased its fusion protein expression (Fig. 2.7B). If the Gal4DBD was extended to 93 aa or 147 aa, Gal4 fusion protein expression was much weaker than wild type ETS1 (Fig. 2.7B). These results demonstrate that Gal4DBD decreases the steady-state levels of its fusion proteins.

Given that low steady-state level of Gal4-ETS1 fusion protein fails to activate the EG mutant enhancer, it is still possible that higher expression of this protein may pass a threshold to activate the mutant enhancer. To boost the expression of Gal4-ETS1 fusion proteins, I first tested if

41

Gal4DBD led to fusion protein degradation. Indeed, a low concentration (2 μM) of the protease inhibitor MG132 greatly increased ETS1^{TVG} expression in EG mutant cells, as detected by anti-Gal4 antibody (Fig. 2.8A). By contrast, endogenous ETS1 levels in cells without ectopic ETS1^{TVG} expression were not affected by titration of MG132 (Fig. 2.8A). Moreover, ETS1^{TVG} detected by anti-ETS1 antibody was comparable to the endogenous ETS1 when treated with 2μM MG132 (Fig. 2.8A). Meanwhile, MG132 did not alter total ETS1 transcription levels in EG mutant cells with or without ETS1^{TVG} expression (Fig. 2.8B). These results suggest that 2 μM MG132 increases ETS1^{TVG} steady-state level without affecting other essential cellular processes. However, increased ETS1^{TVG} expression failed to rescue Jβ1.2 germline transcription in EG mutant cells (Fig. 2.8B). Taken together, I conclude that expression of Gal4-ETS1 fusion protein does not activate knock-in EG mutant enhancer, as it does the luciferase reporters.

Fast binding of Gal4-RUNX1 fusion protein activates mutant Eβ in *Tcrb*-RC

The failure to rescue knock-in mutant Eβ activity by Gal4-ETS1 fusion naturally led to the question if the Gal4-RUNX1 fusion is competent in similar scenarios. Although in luciferase reporter assays, the Gal4-RUNX1 fusion protein RUNX1^{GNT} exhibited stronger potency than Gal4-ETS1 fusion protein ETS1^{TVG} in mutant Eβ activation (Fig. 2.3), it is possible that the *Tcrb*-RC allele harboring the RG mutant enhancer is not activatable. To test this possibility, I transduced RG mutant cells with Gal4DBD and RUNX1^{GNT} retroviral expression vectors (Fig. 2.9A). In contrast to EG mutant cells, Jβ1.2 germline transcription in RG mutant cells was

rescued by the corresponding RUNX1^{GNT} fusion protein to a level similar to that driven by WT Eβ, suggesting that the knock-in RG mutant enhancer activity was fully restored (Fig. 2.9A).

This fully restored mutant Eβ activity provides a useful tool to track the kinetics of *Tcrb*-RC activation process. In RG mutant cells bearing tet-inducible lentiviral expression cassettes for RUNX1^{GNT}, Jβ1.2 germline transcription was fully restored to WT level at 24 hours post induction by Dox (Fig. 2.9B). Before saturation, Jβ1.2 germline transcription gradually increased over the course of 24 hours (Fig. 2.9B). By contrast, Gal4 transcription exhibited a sharper increase from 0 to 4 hours post induction, after which its curve was much flatter (Fig. 2.9B). It should be noted that the Jβ1.2 primers amplify only unspliced germline transcripts, which are intermediate products quickly spliced into more stable transcripts. Thus, the level of unspliced Jβ1.2 germline transcription detected by this PCR is closely related to instant activity of Eβ. Similarly, since Gal4DBD actively causes its fusion proteins to degrade (Fig. 2.7), Gal4 transcription should reflect the instant levels of short-life Gal4 fusion proteins. Thus, different kinetics from these transcriptional results may suggest that induced RUNX1^{GNT} protein reaches its steady state of expression level much faster than induced mutant Eβ activity.

Indeed, the kinetics of RUNX1^{GNT} binding at Eβ measured by chromatin immuneprecipitation (ChIP) was very similar to that of Gal4 transcription, which increased sharply from 0 to 6 hours post induction and became flat afterwards, suggesting that available RUNX1^{GNT} at Eβ was quickly saturated post induction of its expression (Fig. 10A). Meanwhile, RUNX1^{GNT} binding at pDβ1 followed the same kinetics as its binding at Eβ (Fig. 10A). Since there is no RUNX1

43

binding site at pDβ1, detected RUNX1^{GNT} signal at pDβ1 was most likely from the same protein binding to Eβ, which was delivered via promoter-enhancer interaction between pDβ1 and Eβ (43). It is noticeable that ETS1 binding at both pDβ1 and Eβ also closely mimics RUNX1^{GNT} kinetics, which might result from a synergy between the intact ETS1 binding sites at RG mutant enhancer, the NRDB domain of RUNX1^{GNT}, and the ETS1 binding site at pDβ1 (13). Taken together, these results suggest that Eβ activation induces instant promoter-enhancer interaction between pDβ1 and Eβ, which leads to similar kinetics of transcription factor binding. Transcription factor binding at pDβ1 and Eβ was much faster than Jβ1.2 germline transcription as determined by two-way ANOVA test (Fig. 2.10B). By contrast, the differences between factor binding kinetic curves were small. Based on these observations, I conclude that fast transcription factor binding precedes gradual ascendance of Eβ activity during the process of *Tcrb*-RC activation.

2.5 Discussion

ETS1 and RUNX1 proteins are critical in lymphocyte development through their binding to cis-elements of lineage specific genes (18, 35, 44–47). Apart from functioning separately, they often work as co-binding partners on their composite sites (11, 22–24). Previous research focused mainly on their DNA binding capacity without detailed discussion of their individual functions. In my study, I utilized a representative model of ETS1-RUNX1 composite sites, the enhancer Eβ at mouse *Tcrb*-RC locus, to dissect ETS1 and RUNX1 functions. I found that in luciferase reporter assays, both ETS1 and RUNX1 are competent in independently rescuing ETS1 and

RUNX1 binding site mutations, although the RUNX1 binding sites seem more essential since they should be accessible to transactivators. By contrast, in cell models with knock-in ETS1 and RUNX1 binding site mutant alleles, only RUNX1 is able to activate the *Tcrb*-RC allele bearing its own mutant binding sites. The different behaviors of these two proteins demonstrate that results from luciferase substrate are not precise representation of actual chromosomal processes. The rich knowledge accumulated through luciferase and other reporter assays on genetic regulation in past research has to be validated using cell models bearing the same cis-element mutations. Moreover, such cell models will allow more versatile studies than reporter assays. For example, the kinetic comparison between facets of the RUNX1-induced gene activation process would otherwise be impossible without RUNX1 binding site mutants.

Of all questions answerable by mutant knock-in cell models, the ones regarding ETS1-RUNX1 cooperativity show the most advantages of these models over luciferase reporter substrates. ETS1 and RUNX1 interact with each other via the exon VII domain and the NRDB domain, respectively (22–24). To clearly test if one of the two proteins is essential for the enhancer activation mediated by the other, it is necessary to examine if the proteins physically bind to the enhancer. Although ChIP assays were feasible on transfected plasmid-based reporters (48), the binding patterns may be completely different between the same cis-elements on extrachromosomal reporters and in native chromatin (49). Thus, it is possible that different activities of ETS1 on reporters and EG mutant cells are due to the amounts of recruited RUNX1. RUNX1 may activate the reporter substrate with EmRG mutant enhancer independently from ETS1. This hypothesis will be more rigorously tested by ChIP assay to measure ETS1 binding on EmRG mutant alleles.

Additionally, the quantitative measurement by ChIP assay on mutant knock-in enhancers will provide insight into the actual levels of binding by Gal4 fusion proteins to mutant enhancers. Since Gal4DBD determines the low cellular steady-state level of Gal4-ETS1 fusion proteins (Fig. 2.7), it is likely that expression levels of other Gal4 fusion proteins are also affected. However, unlike Gal4-ETS1 fusions, which are unable to activate the chromosomal EG mutant enhancer, Gal4-RUNX1 fusions activate RG mutant enhancer (Fig. 2.9). Thus, it is possible that the local concentration of RUNX1^{GNT} near the RG enhancer is sufficiently above a threshold for Eβ activation. Indeed, RUNX1 signal at RG quickly saturated post induction of its expression (Fig. 2.10). If I titrate the induced steady-state levels of RUNX1^{GNT} with different Dox concentrations, I may be able to quantitatively determine the required threshold of the Runx TAD for Eβ activity using ChIP assay. In this experiment, an important difference between Gal4 fusion and wild type proteins has to be taken into consideration. Since Gal4DBD must dimerize to bind its binding site, its fusion proteins also have two copies of other domains recruited to Gal4 mutant sites, as opposed to their wild type counterparts. So at the molecular scale, RUNX1^{GNT} presents two TADs, which resemble ETS1-RUNX1 co-binding. Thus there are at least two possible scenarios for the required amount and type of TAD for Eβ activation. If both ETS1 and RUNX1 TADs are equally sufficient and necessary for Eβ, RUNX1^{GNT} may need to bind two fold as much as wild type RUNX1 to provide the same amount of TAD for Eβ activation. Alternatively, if only RUNX1 TAD is required, RUNX1^{GNT} may reach the same level of activity by binding at the same amount of wild type RUNX1.

Dimerization and low expression issues with the Gal4DBD may be overcome if we switch to other DNA binding approaches. To bind a single copy of fusion protein, it is possible to use zinc finger DBD (50). With three or four zinc finger domains, it will also increase the specificity of DNA binding throughout the genome (29). Another option will be using the recently developed CRISPR-Cas9 system with a mutant version of Cas9 protein without nuclease activity (51). However, both of these approaches may suffer from the bulky sizes of their DNA binding domains, which may affect the activity of TADs. For example, in our attempt to use LexA DBD (52) instead of Gal4 DBD to make fusion proteins, LexA-ETS1 fusions only modestly activate the luciferase reporters with LexA binding sites (Fig. 2.12). In this regard, Gal4DBD may still be the most potent option for the rescuing approach.

The rescuing approach benefits from using knock-in alleles to provide a chromosomal environment for the local interaction between Gal4 fusion proteins and mutant Eβ enhancers. Apart from this, knock-in alleles also provide the opportunity to investigate how Eβ controls a broad genomic region within the three-dimensional space of nucleus. Since the temporal tracings of ETS1 and RUNX1 binding at pDβ1 closely resemble those at Eβ (Fig. 2.10), the holocomplex between pDβ1 and Eβ must be formed quickly and instantly. Such scenarios are very difficult to investigate using luciferase reporter substrates, although one of my attempts implied the potential roles of genomic regions outside of known cis-elements in *Tcrb*-RC activation (Fig. 2.11). It has been established that CTCF-cohesin complexes mediate many long-range interactions throughout the human and mouse genomes (53). Since there are CTCF binding sites in the upstream of pDβ1 and in the downstream of Eβ, I went on to test if including these regions would increase Eβ luciferase reporter activity. When pDβ1 and Eβ fragments used in luciferase

assay were extended to include more genomic sequence, only extension of Eβ but not pDβ1 dramatically increased luciferase reporter activities (Fig. 2.11). This result suggests that regions other than the CTCF binding site in the downstream of Eβ further enhances its activity. It is possible that matrix attachment regions (MAR) and lamin-associated domains (LAD) in the extra sequences influenced the final activity of Eβ (54, 55). Nevertheless, all phenotypes resulting from CTCF sites, MARs and LADs in future analyses will be more relevant to physiological conditions on knock-in alleles. If I use Chromatin Conformation Capture (3C) assays in our cellular system, I will be able to quantitatively determine the levels of Eβ interaction with pDβ1 and other cis-elements during *Tcrb*-RC activation. If the same approach is applied to other genomic loci, it will eventually lead to a comprehensive and detailed understanding of the epigenetic regulation mechanisms of not only *Tcrb*-RC but also genomic domains in general.
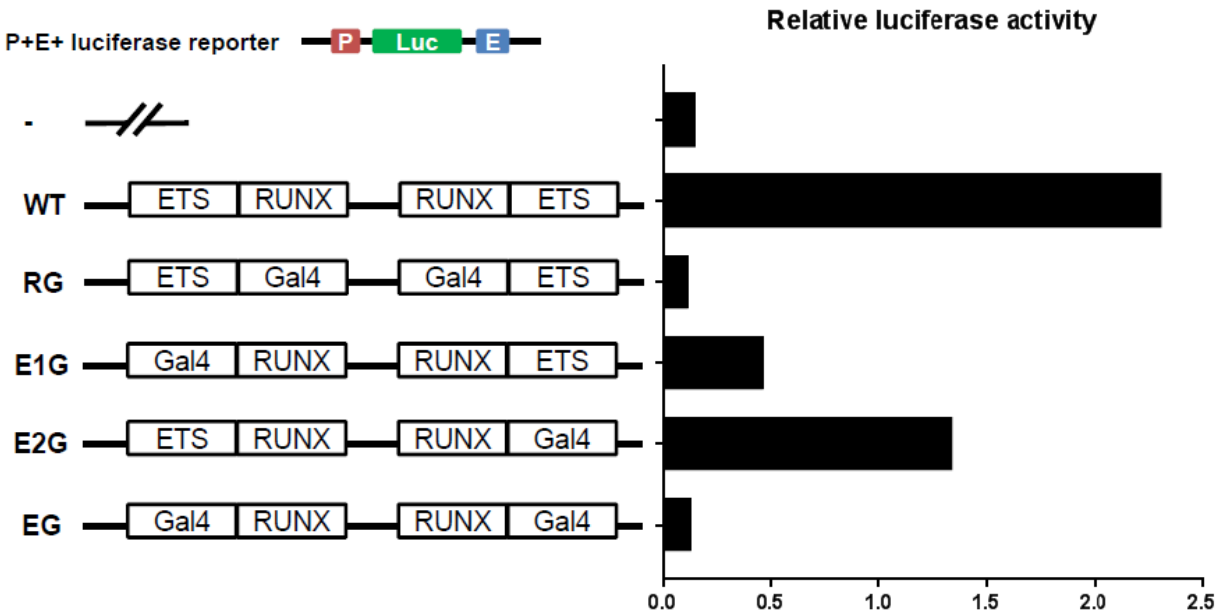
2.6   Figures

Figure 2.1 *Tcrb*-RC enhancer reporter activity is dependent on both ETS/RUNX binding motifs.
Schematic representation in the upper left panel shows a pGL3 luciferase reporter plasmid with a 0.49kb promoter upstream of Dβ1 (the red bar) and a 0.57kb wild-type enhancer Eβ of *Tcrb*-RC (the blue bar). The green bars labeled by "LUC" indicate firefly luciferase cDNA. This plasmid is termed P+E+. Reporter plasmids with the same promoter and different enhancer cassettes were used in this experiment. For simplicity, the rest of the left panel shows only the close-up depiction of the enhancer regions. "-" refers to no enhancer. WT refers to wild-type enhancer Eβ. RG refers to mutation of both RUNX1 binding sites to Gal4 binding sites. E1G refers to mutation of only the upstream ETS1 site to Gal4 site. E2G refers to mutation of only the downstream ETS1 site to Gal4 site. EG refers to mutation of both ETS1 sites to Gal4 sites. Firefly and control Renilla luciferase reporters were introduced to P5424 cells and measured as described in Materials and Methods. The right panel shows firefly luciferase signals normalized to Renilla signals from one experiment, which correspond to the enhancer variations on the left.
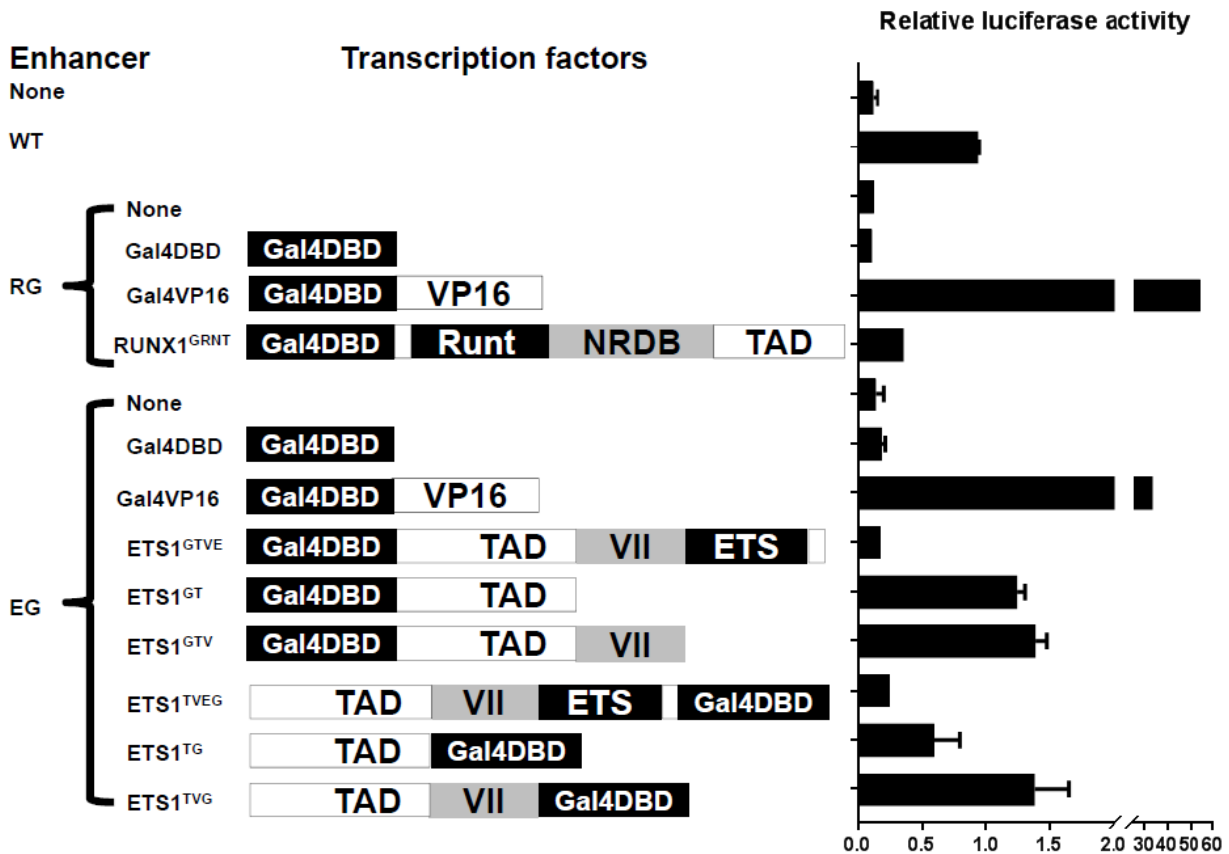
Figure 2.2 Gal4 fusion proteins restore mutant Eβ activity on reporters.

Schematic representation in the left panel shows the combination of different enhancers and transcription factors. All reporters share the same promoter for Dβ1 as described in Fig 2.1. For enhancers, "none" refers to no enhancer. WT, RG and EG enhancers are as described in Fig. 2.1. For transcription factors, "none" refers to no co-transfection. Gal4DBD refers to the 147 aa Gal4 DNA binding domain. Gal4VP16 refers to the fusion protein of Gal4DBD and VP16 activation domain. RUNX1 consists of an N terminus Runt DNA binding domain (Runt), a negative regulatory region of DNA binding domain (NRDB), and a transactivation domain (TAD). RUNX1^{GRNT} refers to the fusion of Gal4DBD and the full-length RUNX1. ETS1 consists of an N terminus transactivation domain (TAD), an exon VII domain (VII), and an ETS DNA binding domain (ETS). ETS1^{GTVE} refers to the fusion of the full-length ETS1 and Gal4DBD at the N terminus. ETS1^{GT} refers to the fusion of Gal4DBD at the N terminus and the ETS1 TAD. ETS1^{GTV} refers to the fusion of Gal4DBD at the N terminus, the ETS1 TAD and the ETS1 exon VII domain. ETS1^{TVEG} refers to the fusion of the full-length ETS1 and Gal4DBD at the C terminus. ETS1^{TG} refers to the fusion of the ETS1 TAD and Gal4DBD at the C terminus. ETS1^{TVG} refers to the fusion of the ETS1 TAD, the ETS1 exon VII domain and Gal4DBD at the C terminus. Firefly and control Renilla luciferase reporters were introduced to P5424 cells with transcription factor expression plasmids and measured as described in Materials and Methods.

50

The right panel shows firefly luciferase signals normalized to Renilla signals from two experiments, which correspond to the enhancer/TF combinations on the left.
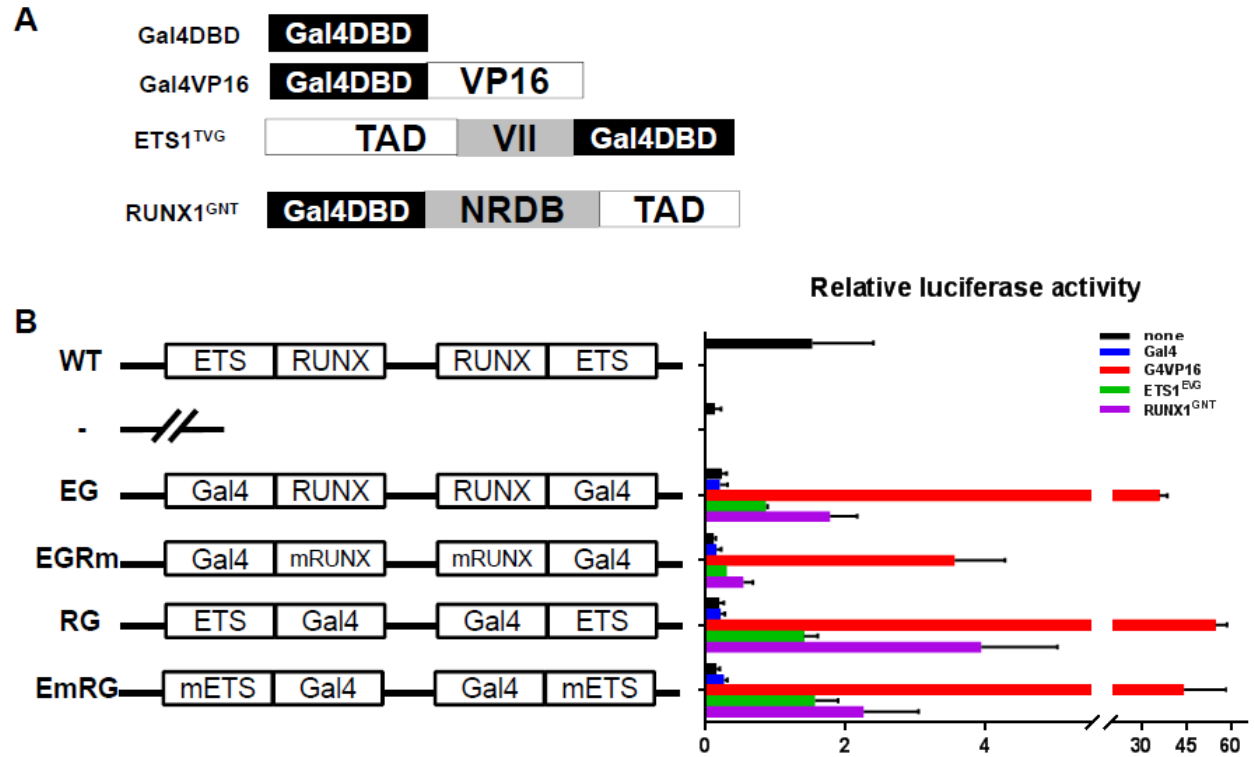
Figure 2.3 Gal4-ETS1 and –RUNX1 fusion proteins activate Eβ with ETS1 and/or RUNX1 binding site mutation in an independent manner.

(**A**) Cartoon depictions of fusion proteins used in this experiment. Gal4DBD, Gal4VP16 and ETS1^{TVG} are as described in Fig. 2.2. RUNX1^{GNT} refers to the fusion of Gal4DBD, the NRDB domain of RUNX1 and the RUNX1 TAD. (**B**) Schematic representation in the left panel shows different enhancers. All reporters share the same promoter for Dβ1 as described in Fig 2.1. WT, "-", RG and EG enhancers are as described in Fig. 2.1. EGRm refers to mutation of both ETS1 binding sites to Gal4 binding sites and null mutation of both RUNX1 sites. EmRG refers to mutation of both RUNX1 binding sites to Gal4 binding sites and null mutation of both ETS1 sites. In the right panel, different transcription factors are color-coded. Black refers to no ectopic protein. Blue refers to Gal4DBD. Red refers to Gal4VP16. Green refers to ETS1^{TVG}. Purple refers to RUNX1^{GNT}. All transcription factors are as described in Fig. 2.2. Data in the right panel are grouped by the enhancer types on reporters. WT and "-" enhancers were done without co-transfection of transcription factors. Firefly and control Renilla luciferase reporter plasmids were introduced to P5424 cells with transcription factor expression plasmids and measured as described in Materials and Methods. The right panel shows firefly luciferase signals normalized to Renilla signals from two experiments, which correspond to the enhancers on the left.
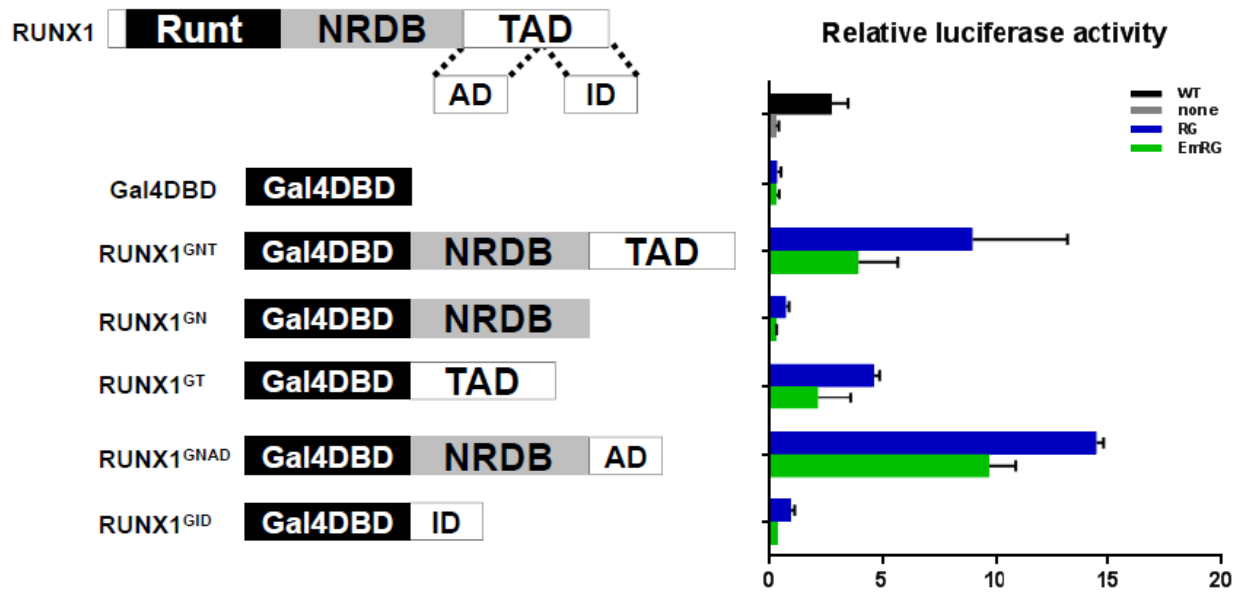
52

Figure 2.4 Gal4 fusion proteins with truncated RUNX1 identify functional domains responsible for Eβ reporter activity.

Schematic representation in the upper left panel shows dissection of RUNX1 domains as described in Fig. 2.2 and the further dissection of the TAD into Activation Domain (AD) and Inhibition Domain (ID). The lower left panel shows fusion proteins used in this experiment. Gal4DBD is as described in Fig. 2.2. RUNX1^{GNT} is as described in Fig. 2.3A RUNX1^{GT} refers to the fusion of Gal4DBD and the RUNX1 TAD. RUNX1^{GNAD} refers to the fusion of Gal4DBD, the NRDB domain of RUNX1 and the AD sub-domain of the RUNX1 TAD. RUNX1^{GID} refers to the fusion of Gal4DBD and the ID sub-domain of the RUNX1 TAD. In the right panel, enhancers on luciferase reporters are color-coded. WT, RG and EmRG enhancers are as described in Figs. 2.1 and 2.3. "none" refers to no enhancer. All reporters share the same promoter for Dβ1 as described in Fig 2.1. Black, grey, blue and green bars refer to WT, "none", RG and EmRG enhancers, respectively. Data in the right panel are grouped by fusion proteins. Firefly and control Renilla luciferase reporters were introduced to P5424 cells with transcription factor expression plasmids and measured as described in Materials and Methods. The right panel shows firefly luciferase signals normalized to Renilla signals from two experiments, which correspond to the fusions on the left.
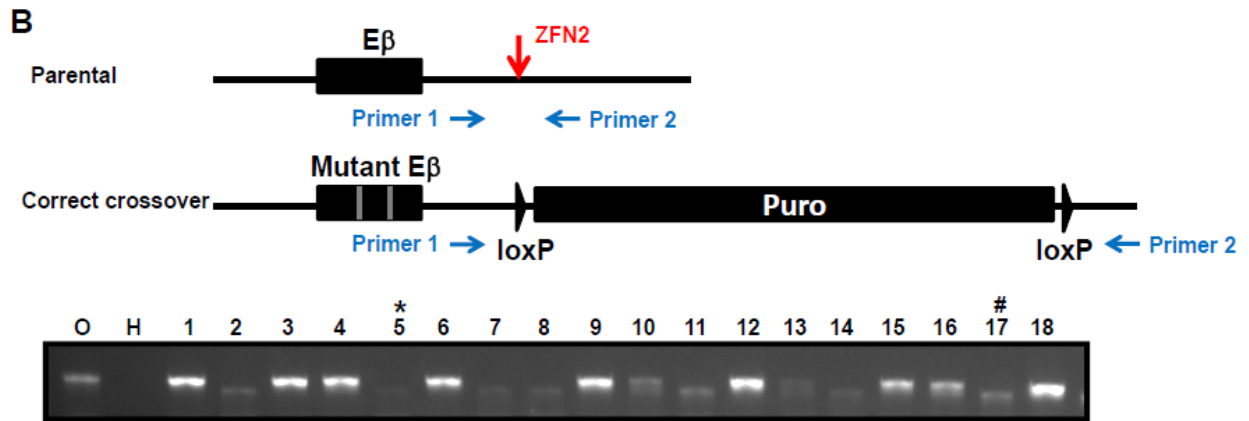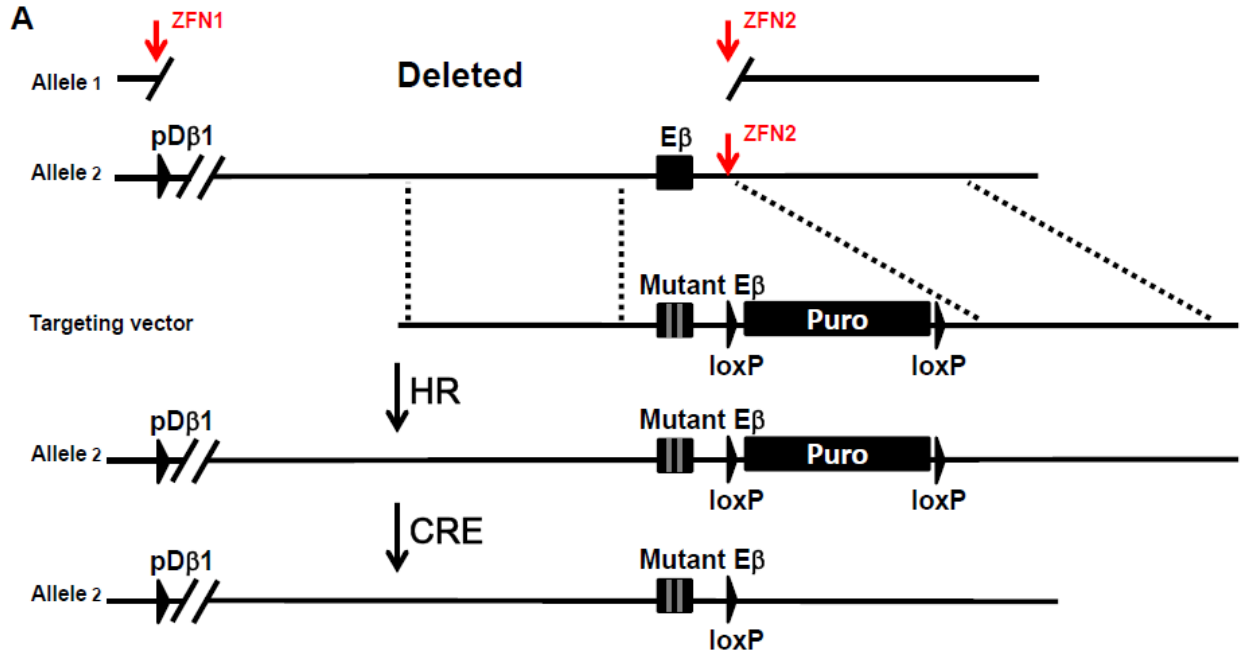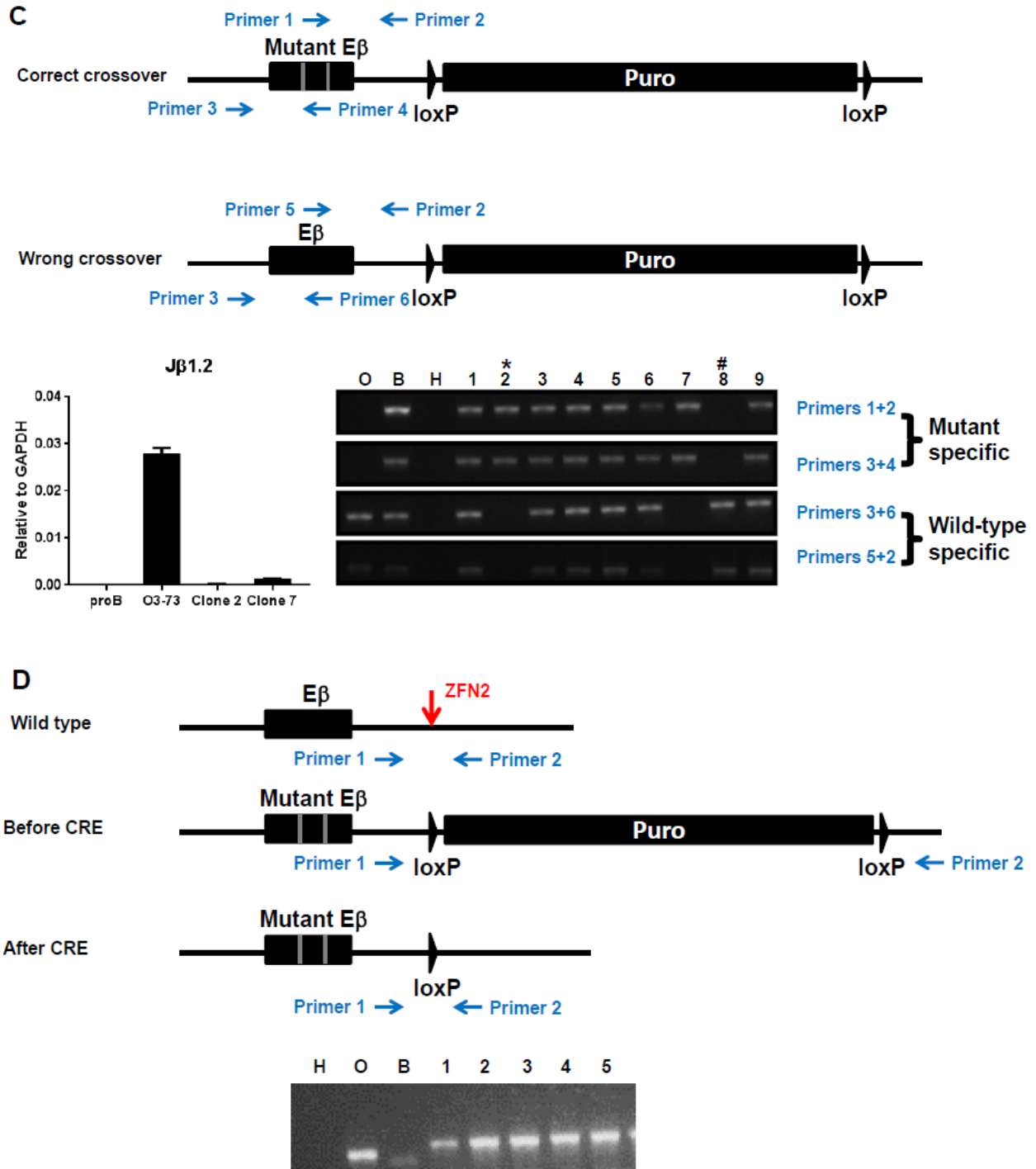
Figure 2.5 Zinc Finger Nuclease facilitates cloning of a knock-in cell model with the mutant EG enhancer.

(A) Schematic representation shows the overall cloning strategy. The top two alleles represent endogenous O3-73 alleles. Allele 1 shows a 22 kb deletion from upstream of the promoter of Dβ1 (pDβ1) to downstream of Eβ. Allele 2 shows pDβ1 and a close-up view of regions flanking

Eβ, with in-between regions omitted. The red arrows denote the Zinc Finger Nuclease (ZFN) cutting sites by ZFN1 and ZFN2. The regions on the targeting vector for recombination are aligned with Allele 2, with left and right homologous arms, mutant Eβ (two mutations marked by grey bars) and a loxP-Puro-loxP cassette for selection. The bottom two alleles represent the intermediate product after homologous recombination (HR) and the final product after CRE deletion of loxP cassette. **(B)** Schematic representation on top shows screening strategy for clones with successful crossover using primers (Primers 1 and 2) flanking the ZFN cutting site. The bottom shows representation of PCR products from different templates on 2% TBE agarose gel. O denotes O3-73 genomic DNA. H denotes water control. Numbers 1-18 are clones. * and # mark the same samples examined in later panels of this figure. Alignment of Primers 1 and 2 are shown on both parental and product alleles. The primer and clone numbers only apply to this panel. **(C)** The top schematic representation shows screening strategy for correct crossover products. Primers 1 and 4 align with mutant sequences of the two ETS1/RUNX1 motifs, respectively. Primers 2 and 3 align with consensus sequences of correct and wrong crossovers. Primers 5 and 6 align with wild type sequences of the two motifs. The lower right panel shows PCR product with different templates on 2% TBE agarose gel. O denotes O3-73 genomic DNA. B denotes genomic DNA from the bulk prior to sub-cloning. H denotes water control. Numbers 1-9 are clones. * and # mark the same samples examined in panel B of this figure. The primer and clone numbers only apply to this panel. The lower left panel shows RT-qPCR of unspliced germline transcription through Jβ1.2 to validate mutant Eβ activity. proB refers to a proB cell line R2K2. Clones 2 and 7 are the selected two correct clones. Data are presented as the average values of two PCR replicates (+S.D.) from one experiment. **(D)** Schematic representation on top shows the strategy to screen for clones with successful CRE deletion of the loxP-Puro-loxP cassette. See panel A of this figure for detailed description of alleles. The primer and clone numbers only apply to this panel. The lower panel shows PCR product with different templates on 2% TBE agarose gel. H denotes water control. O denotes O3-73 genomic DNA. B denotes genomic DNA from the bulk prior to sub-cloning. Numbers 1-5 are correct clones.
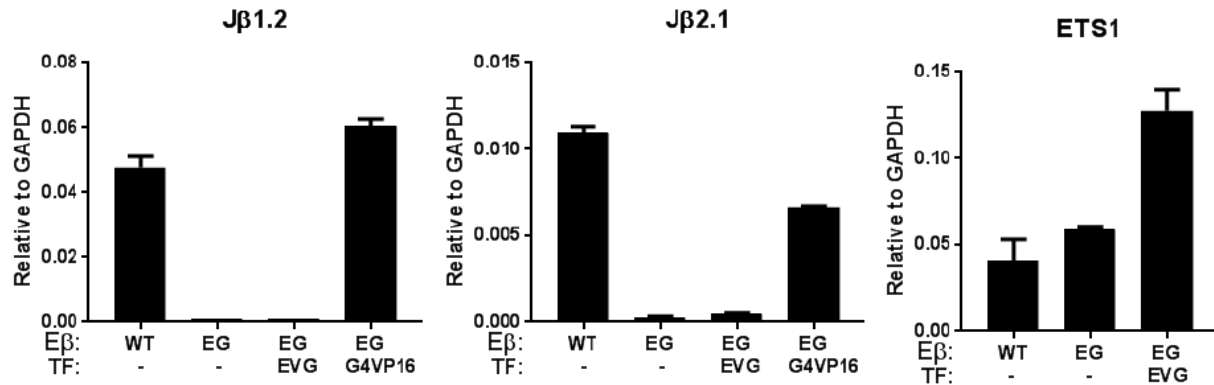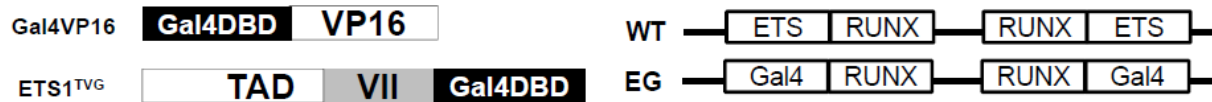
Figure 2.6 Chromosomal mutant Eβ lacking ETS1 binding site is activated by Gal4VP16 but not Gal4-ETS1 fusion proteins.

The top panel shows schematic representation of Gal4VP16 and ETS1^{TVG} fusion proteins as well as WT and EG enhancers (see also Fig. 2.2). The bottom panel shows RT-qPCR of unspliced germline transcription through Jβ1.2, Jβ2.1 regions and total level of ETS1 transcription for different versions of Eβ (WT and EG) induced by different transcription factors (ETS1^{TVG} and Gal4VP16, denoted by EVG and G4VP16, respectively). Data are presented as the average values of two PCR replicates (+S.D.) from one experiment.
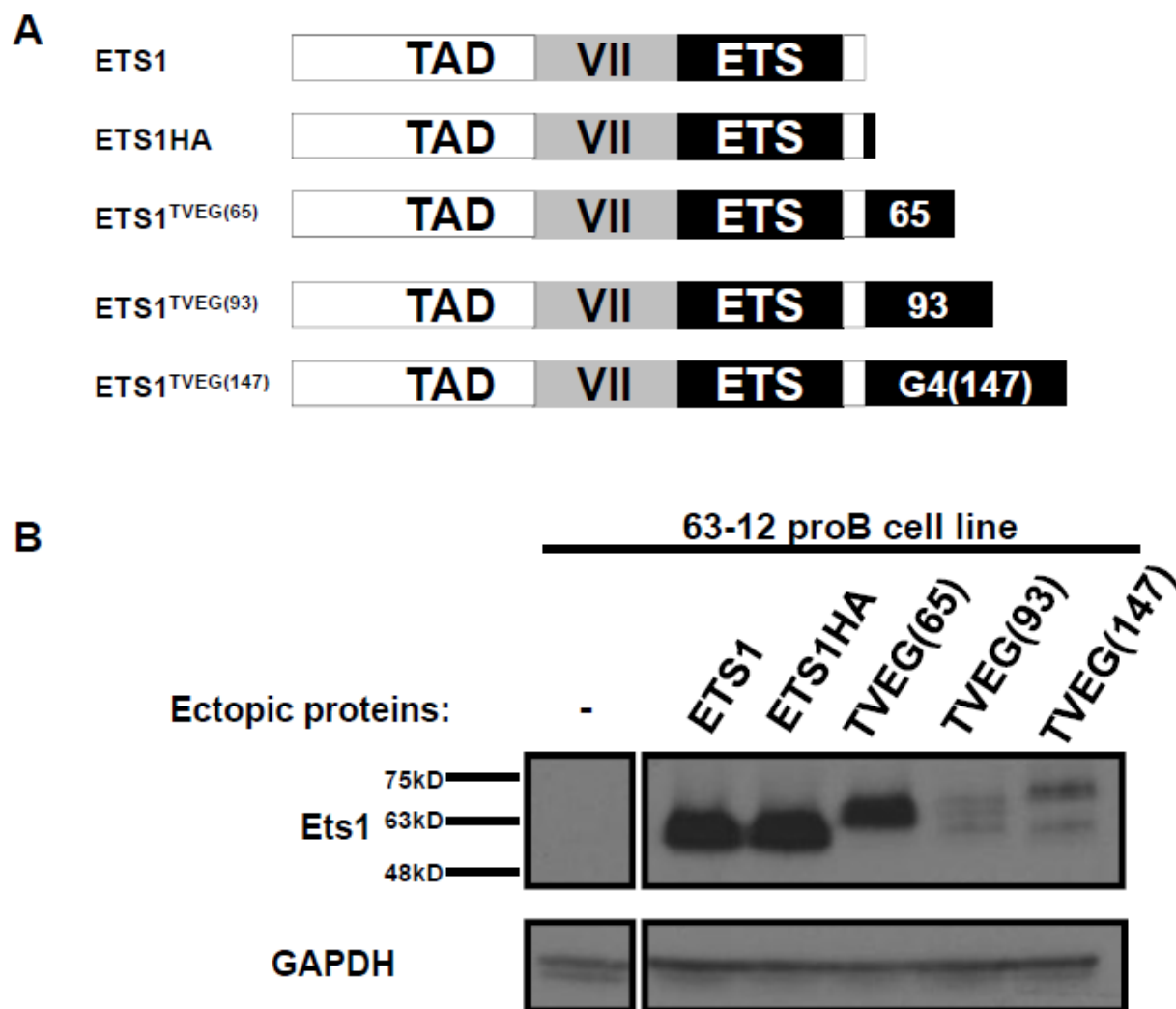
Figure 2.7 Gal4DBD decreases steady-state levels of fusion proteins in lymphoma cells.

(A) Cartoon depictions of wild-type ETS1 and its fusion proteins with different tags. ETS1HA is the fusion between full-length ETS1 and a C-terminus HA tag. ETS1$^{TVEG(65)}$, ETS1$^{TVEG(93)}$, ETS1$^{TVEG(147)}$ are full-length ETS1 fused with aa 1-65, aa 1-93 and aa 1-147 from Gal4 DBD, respectively. ETS1$^{TVEG(147)}$ is the same as ETS1^{TVEG} used in Fig. 2.2. (B) Western blotting analysis for expression of the proteins in panel A in the 63-12 proB cell line. Cells were transduced with retroviral vectors constitutively expression these proteins cells, enriched and processed as described in Materials and Methods.
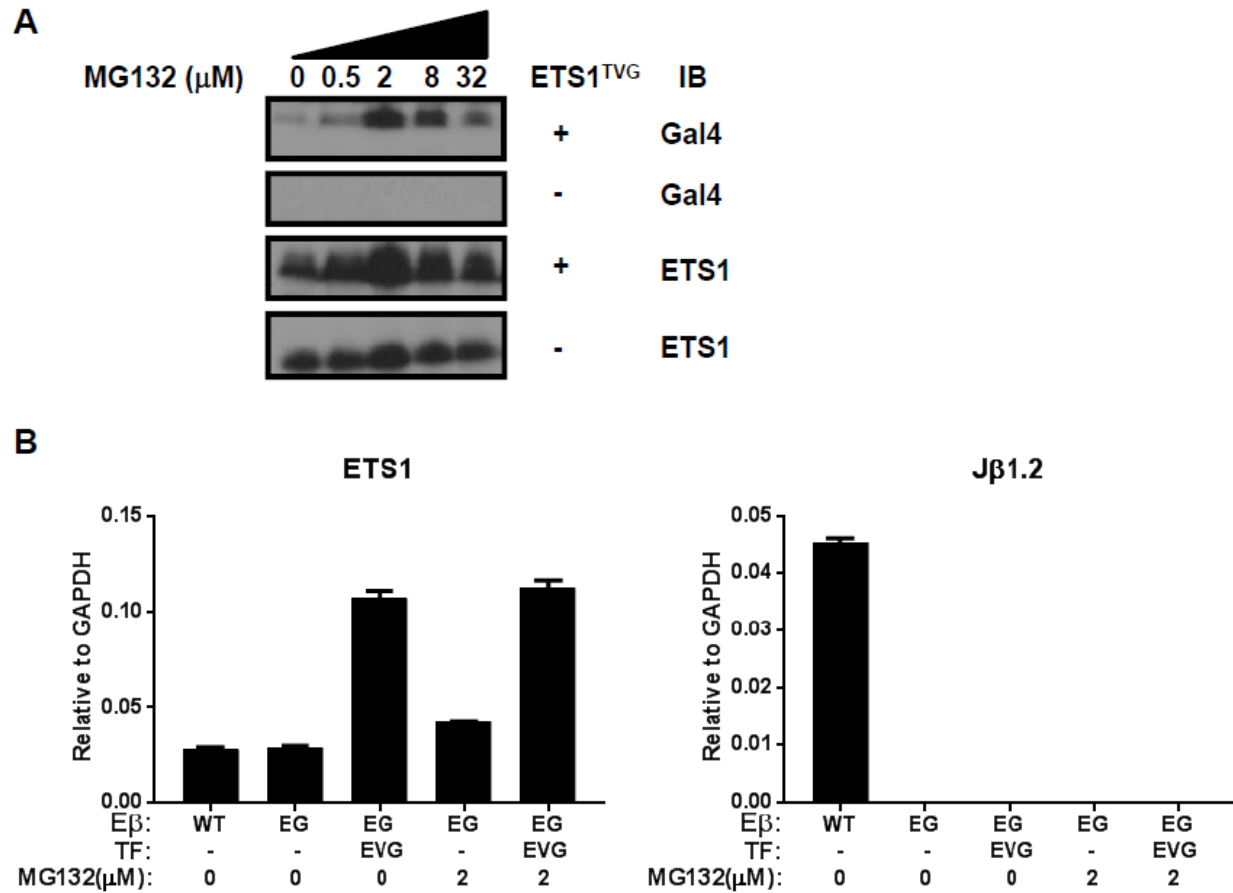
Figure 2.8 Stabilized Gal4-ETS1 fusion protein fails to activate chromosomal mutant Eβ lacking ETS1 binding site.

(A) The steady-state levels of ETS1[TVG] under the titration of the protease inhibitor MG132 concentration. The EG mutant cells were transduced with retroviral vectors constitutively expressing ETS1[TVG] protein and enriched as described in Materials and Methods. MG132 was added to the culture media of transduced and control samples at the labeled concentrations. Equal number of cells (2×10^6) were harvested after 24 hours and processed for conventional western blotting analysis for Gal4DBD or ETS1 epitopes. (B) RT-qPCR of total level of ETS1 transcription and unspliced germline transcription through Jβ1.2 region and for different versions of Eβ (WT and EG). Cells constitutively expressing transcription factors ETS1[TVG] and control cells were treated with or without 2μM MG132 for 24 hours prior to harvest. Data are presented as the average values of three PCR replicates (+S.D.) from one experiment.
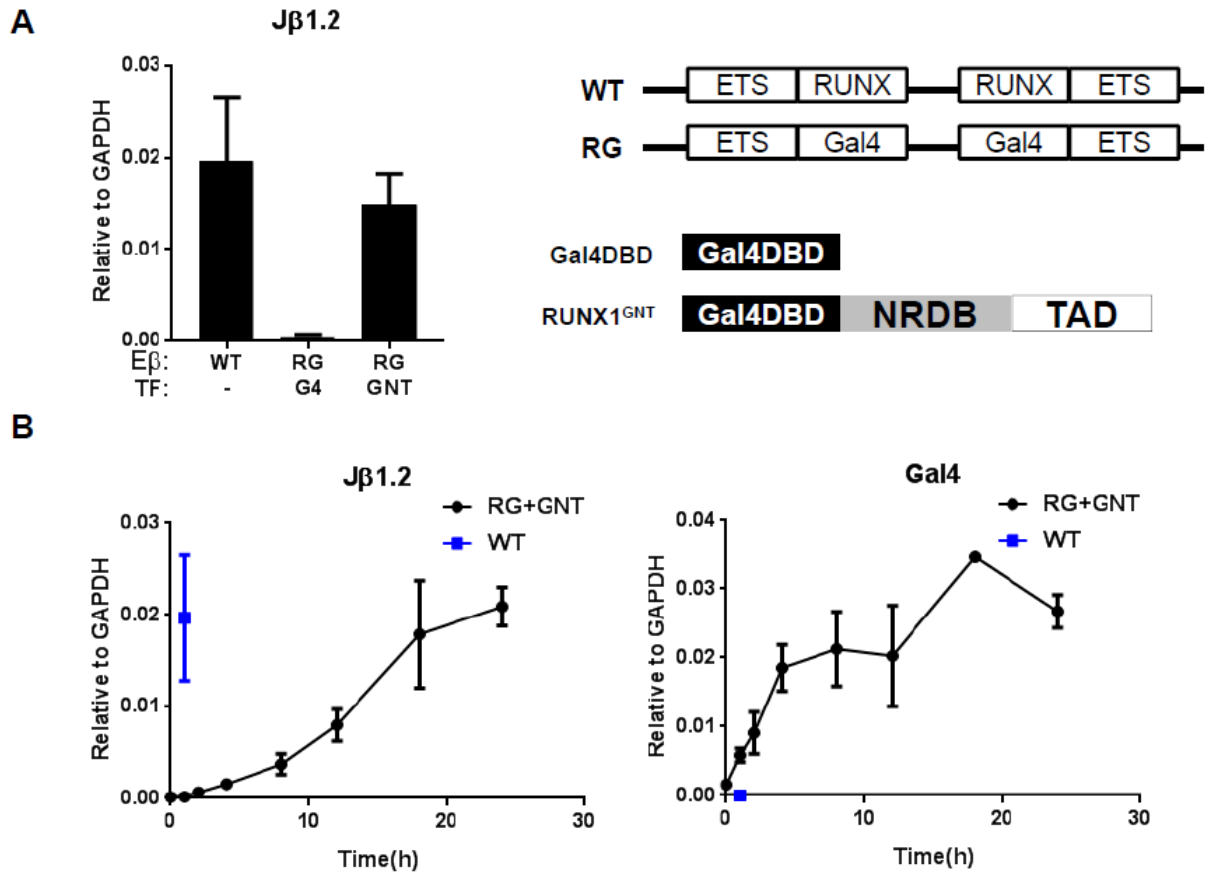
Figure 2.9 Chromosomal mutant Eβ lacking RUNX1 binding site is activated by Gal4-RUNX1 fusion protein.

(A) The right panel shows schematic representation of Gal4DBD and RUNX1[GNT] fusion proteins as well as WT and RG enhancers (see also Fig. 2.3). The left panel shows RT-qPCR of unspliced germline transcription through Jβ1.2 region for different versions of Eβ (WT and RG) induced by different transcription factors (Gal4DBD and RUNX1[GNT], denoted by G4 and GNT, respectively). EG cells were transduced with tet-inducible lentiviral vectors and enriched as in Materials and Methods. Transduced cells were harvested 24 hours after treated with 1 μg/mL Dox. Data are presented as the average values of three PCR replicates (+S.D.) from one experiment. **(B)** Different induction kinetics for unspliced germline transcription through Jβ1.2 region and Gal4 cDNA transcription. RG mutant cells with tet-inducible lentiviral vector expression RUNX1[GNT] (black dots and lines denoted by RG+GNT) were treated with 1 μg/mL Dox for 0, 1, 2, 4, 8, 12, 18, and 24 hours prior to harvest. Data from O3-73 cells are blue dots denoted by "WT". Data are presented as the average values of three PCR replicates (+S.D.) from one experiment.

A

B

Significant levels between curves

	Jβ1.2	ETS1 at Eβ	RUNX1 at Eβ	ETS1 at pDβ1
RUNX1 at pDβ1	****	n.s.	**	n.s.
ETS1 at pDβ1	***	*	*	
RUNX1 at Eβ	***	n.s.		
ETS1 at Eβ	**			
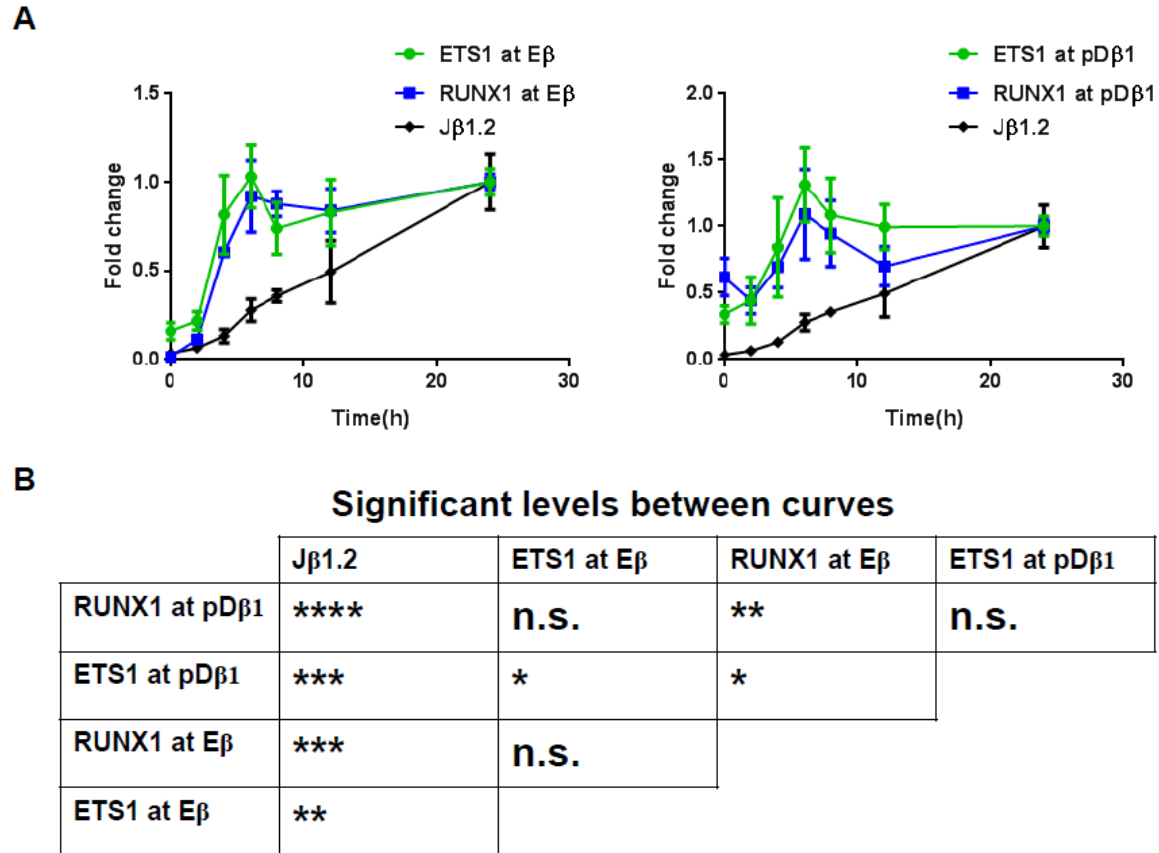
Figure 2.10 Fast transcription factor binding at *Tcrb*-RC cis-elements precedes Eβ activation.

(A) Different induction kinetics for unspliced germline transcription through Jβ1.2 region as well as ETS1 and RUNX1 binding at Eβ and the promoter for Dβ1 (pDβ1). RG mutant cells with tet-inducible lentiviral vector expression RUNX1^{GNT} were treated with 1 µg/mL Dox for 0, 2, 4, 6, 8, 12, and 24 hours prior to harvest. RT- and ChIP-qPCR were performed on RNA and chromatin collected for each time point, respectively, as described in Materials and Methods. The RUNX1 antibody recognizes the C-terminus of RUNX1 protein, which also exists in RUNX1^{GNT} fusion protein. For the convenience of comparison, data at 24 hours were set to 1 and other times were normalized to them. In the left panel, ETS1 and RUNX1 binding at Eβ are shown by green and blue dots and lines, respectively. The right panel shows ETS1 and RUNX1 binding at pDβ1. In both panels Jβ1 germline transcription is shown by black dots and lines. Data are presented as the average values of three PCR replicates (+S.D.) from one experiment. **(B)** Significant levels between all curves shown in panel A are determined by two-way ANOVA using time as one factor and curve as the other factor. n.s., *, **, ***, and **** represent non-significant (p>0.05), p≤0.05, p≤0.01, p≤0.001, and p≤0.0001, respectively.

61
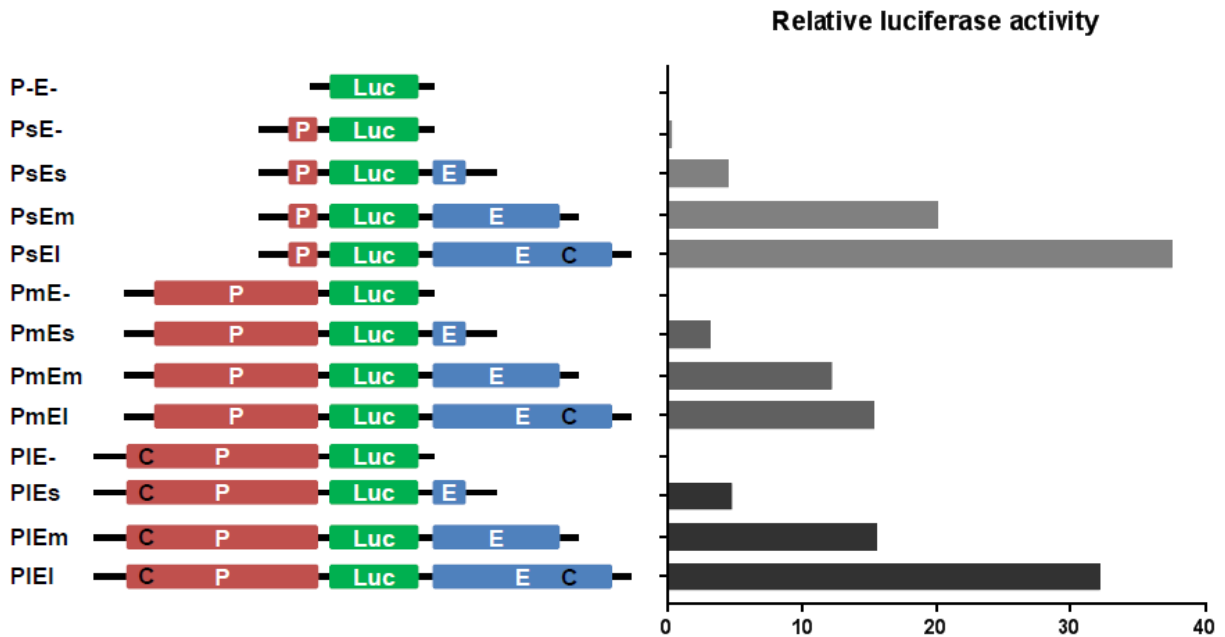
Figure 2.11 Longer *Tcrb* enhancer but not promoter increases luciferase reporter signal.

Schematic representation in the left panel shows different lengths of promoters and enhancers cloned to the pGL3 luciferase reporter plasmid. Short, medium, and long promoters sharing the same downstream end are 0.49kb, 3.11kb, and 3.65kb, respectively. They are labeled Ps, Pm, and Pl, respectively. Short, medium, and long enhancers sharing the same upstream end are 0.57kb, 2.39kb, and 3.41kb, respectively. They are labeled Es, Em, and El, respectively. The reporter without promoter or enhancer is labeled as P-E-. The short promoter and enhancer are also denoted in previous figures as P+ and E+, respectively. The red bars labeled by "P" are promoters. The green bars labeled by "LUC" indicate firefly luciferase cDNA. The blue bars labeled by "E" are enhancers. The black letter "C" indicates CTCF binding sites. All segments are drawn to scale. Firefly and control Renilla luciferase reporters were introduced to P5424 cells and measured as described in Materials and Methods. The right panel shows firefly luciferase signals normalized to Renilla signals from one experiment, which correspond to the promoter/enhancer variations on the left.
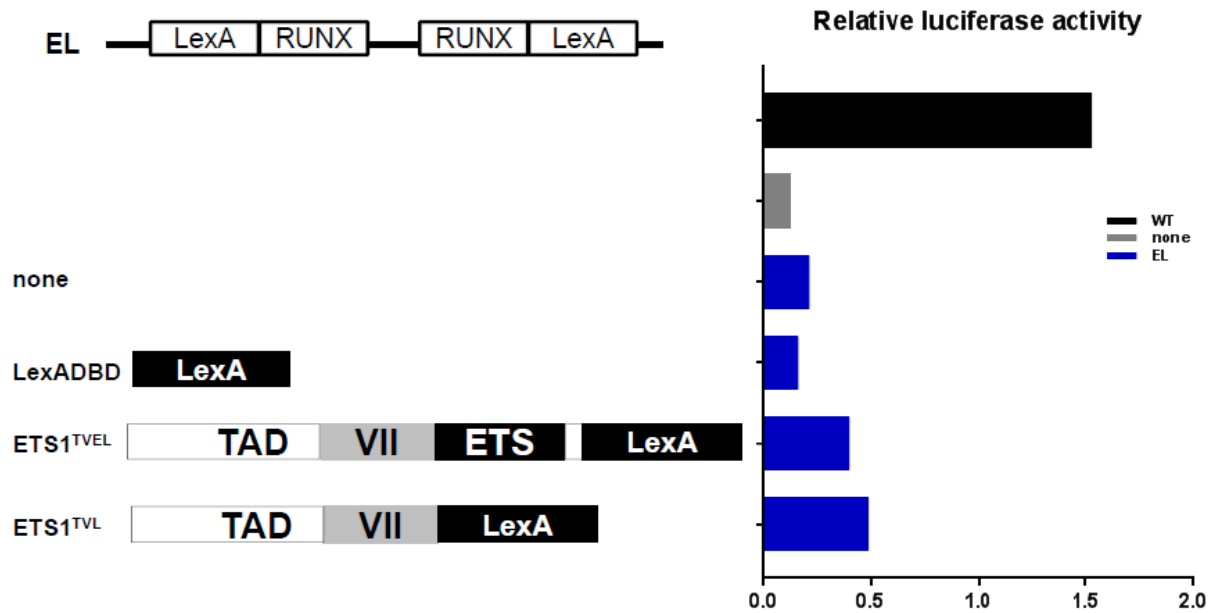
Figure 2.12 LexA-ETS1 fusion proteins fail to fully restore mutant Eβ activity on reporters.

Schematic representation on the upper left panel shows the EL mutant enhancer with both ETS1 binding sites mutated to LexA binding sites. The lower left panel shows fusion proteins used in this experiment. "none" refers to no co-transfected protein. LexADBD refers to LexA DNA binding domain. ETS1^{TVEL} refers to the fusion protein with ETS1 full-length protein and LexADBD. ETS1^{TVL} refers to the fusion protein with ETS1 TAD, ETS1 exon VII domain and LexADBD. In the right panel, reporters with different enhancers are color-coded. The black bar refers to WT. The grey bar refers to "none" or no enhancer. The blue bars refer to reporters with EL mutant enhancer. Firefly and control Renilla luciferase reporters were introduced to P5424 cells with transcription factor expression plasmids and measured as described in Materials and Methods. The right panel shows firefly luciferase signals normalized to Renilla signals from one experiment, which correspond to the TF shown on the left.
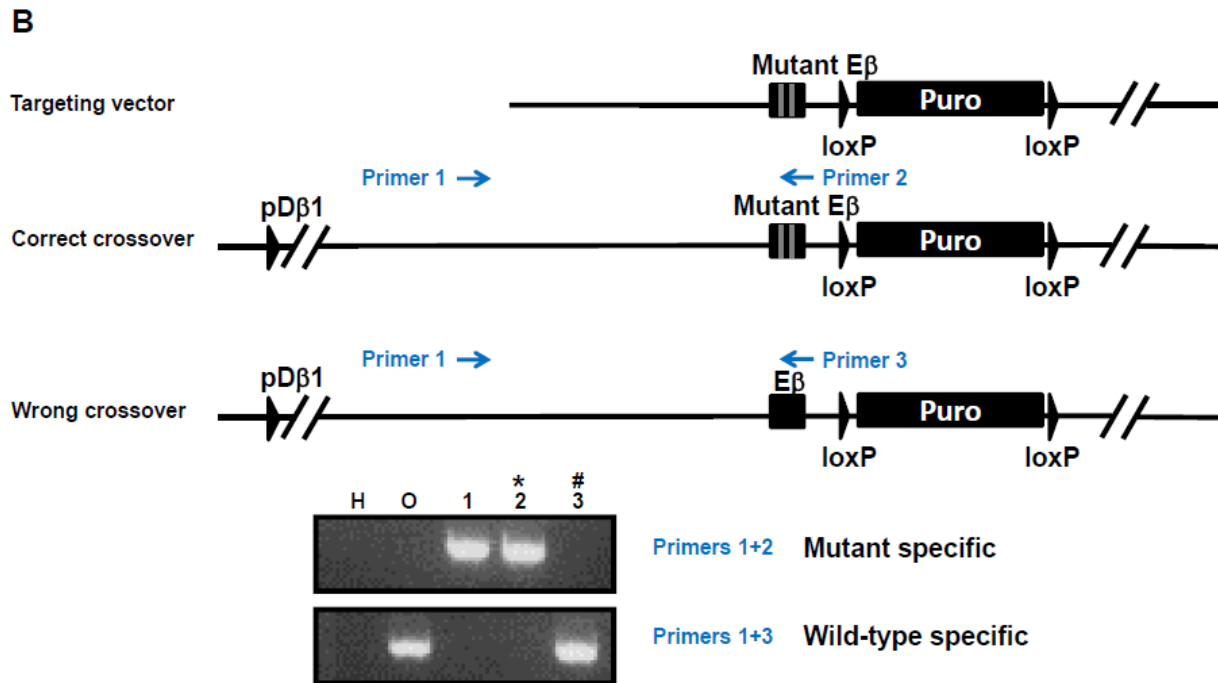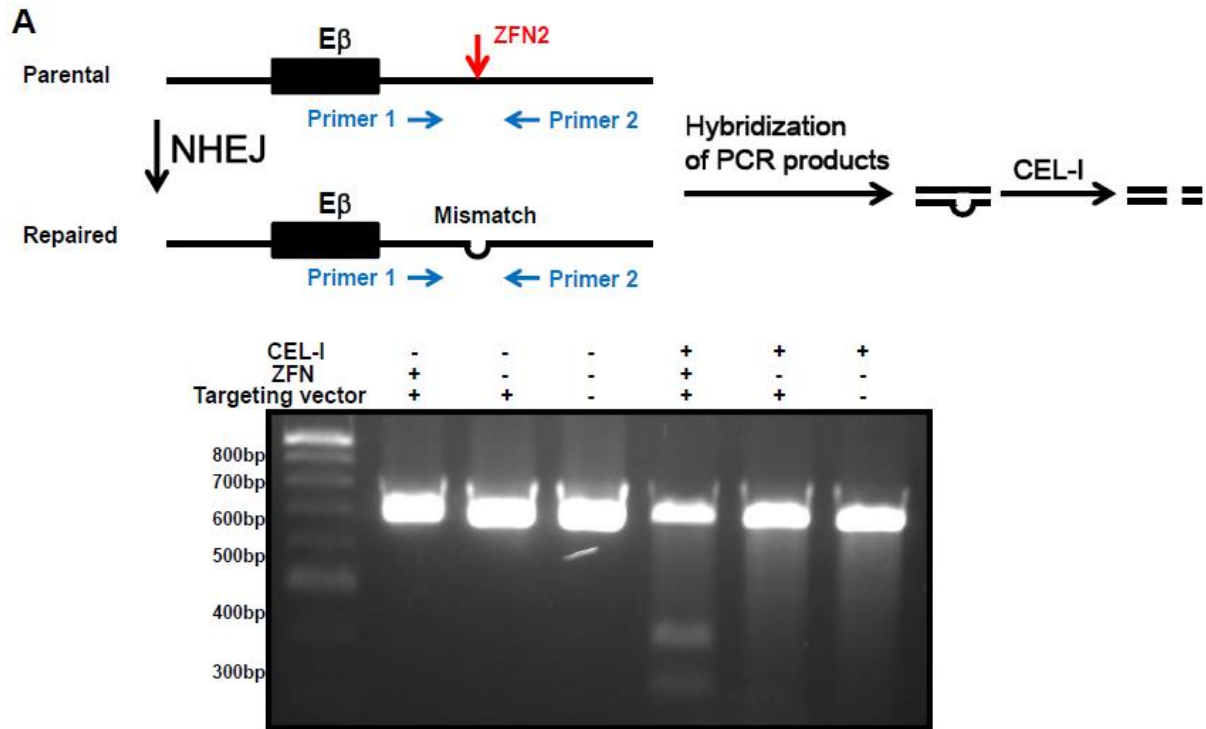
Figure 2.13 EG mutant knock-in allele cloning is quality-controlled by CEL-I assay and long-range PCR.

(A) Schematic representation on top shows the strategy and rationale for CEL-I assay. Double-strand DNA cut by ZFN and repaired by Non-Homologous End Joining (NHEJ) pathway results in gain or loss of nucleotides, which is shown as a mismatched bubble. The hybridized product with DNA strands from both parental and repaired templates is then cut by CEL-I enzyme. The bottom shows uncut and cut PCR products from different samples on 2% TAE agarose gel. **(B)** Schematic representation on top shows the strategy of long-range PCR. The vector and alleles show a close-up view of the left homologous arm, Eβ, and the loxP-Puro-loxP cassette, with other upstream and downstream regions omitted. Primer 1 aligns only to the endogenous alleles. Primers 2 and 3 align with mutant and wild-type sequences, respectively. The lower panel shows PCR product with different templates on 1% TBE agarose gel. H denotes water control. O denotes O3-73 genomic DNA. Numbers 1-3 are clones. * and # mark the same samples examined in Fig. 2.5. The primer and clone numbers only apply to this panel.

A

B

C

Jβ1.2
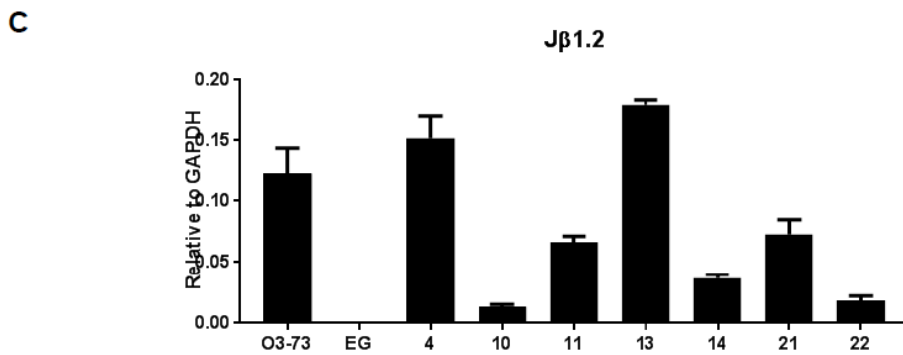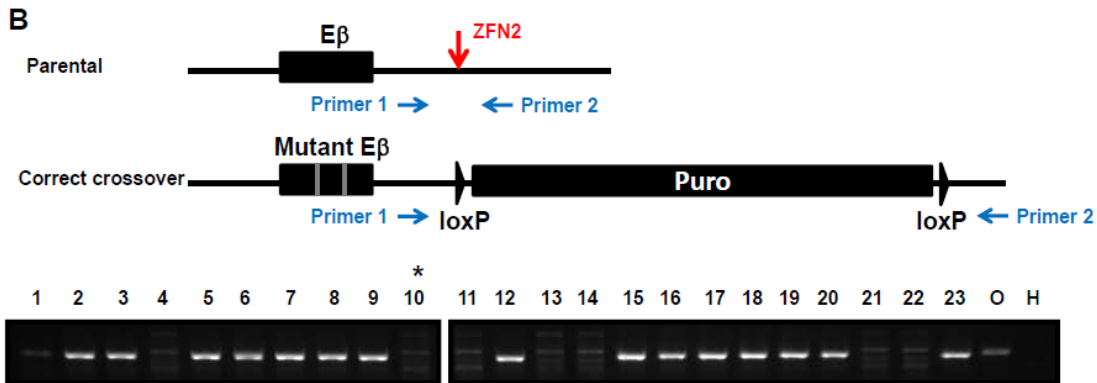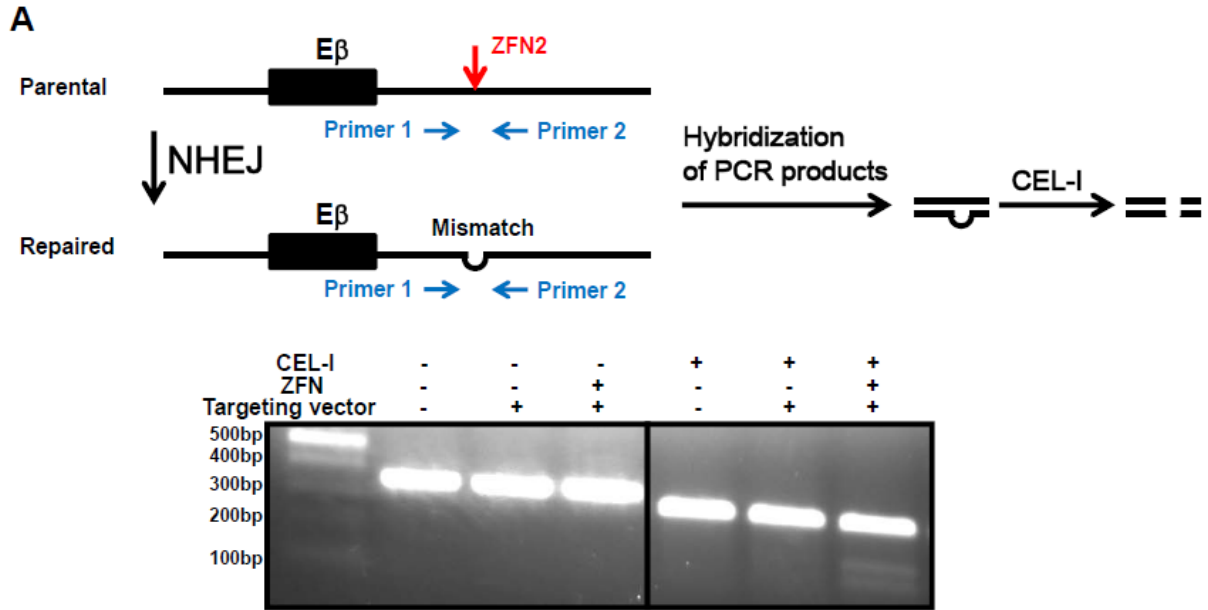
Figure 2.14 Screening and cloning of a knock-in cell model with the mutant RG enhancer.

(A) Schematic representation and experimental result of CEL-I assay used in RG cell cloning. See detailed description in Fig. 2.13A. **(B)** Schematic representation and experimental design are as described in Fig. 2.5B except that samples were from RG cell cloning. The bottom shows representation of PCR products from different templates on 2% TBE agarose gel. O denotes O3-73 genomic DNA. H denotes water control. Numbers 1-23 are clones. Sample marked by * was used in CRE deletion in panel D. The primer and clone numbers only apply to this panel. **(C)** Mutant Eβ activity for selected clones from panel B was validated of by RT-qPCR of unspliced germline transcription through Jβ1.2. O3-73 and EG denote O3-73 and EG cell RNA, respectively. Other numbered columns denote RNA from the clones with the same numbers from panel B. Data are presented as the average values of two PCR replicates (+S.D.) from one experiment. **(D)** Schematic representation and experimental design are as described in Fig. 2.5D except that samples were from RG cells cloning. The primer and clone numbers only apply to this panel. The lower panel shows PCR product with different templates on 2% TBE agarose gel. H denotes water control. O denotes O3-73 genomic DNA. B denotes genomic DNA from the bulk prior to sub-cloning. Numbers 1-4 are clones.
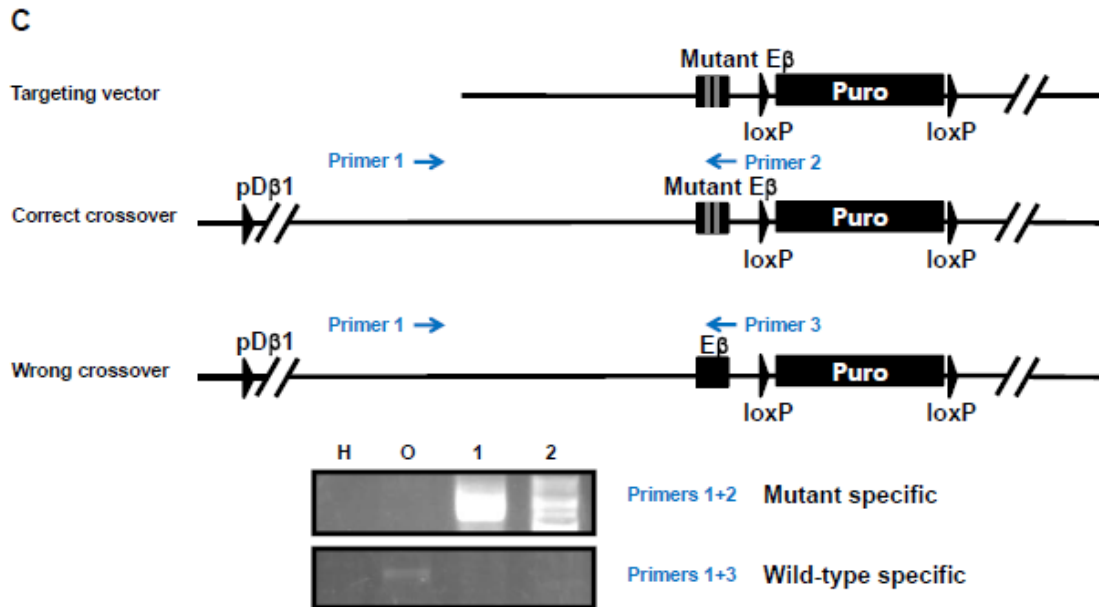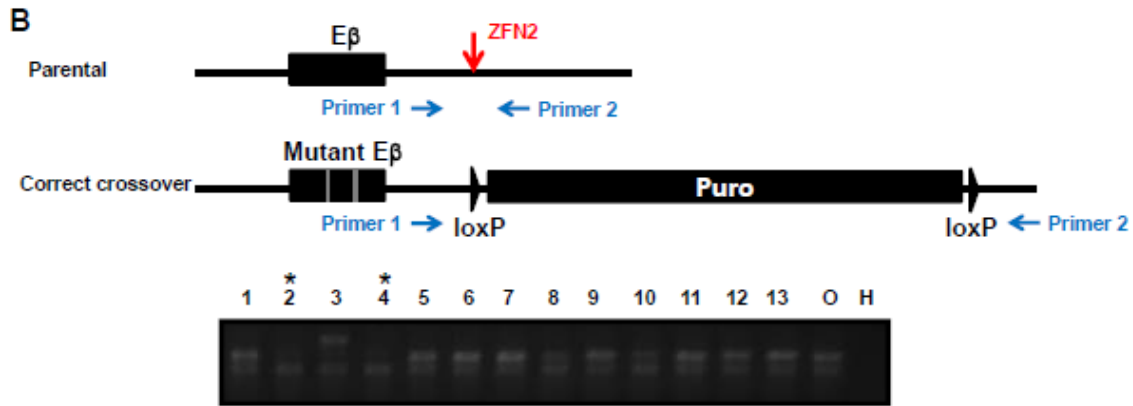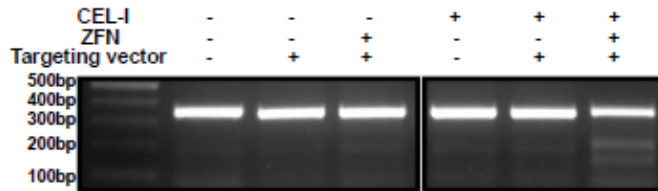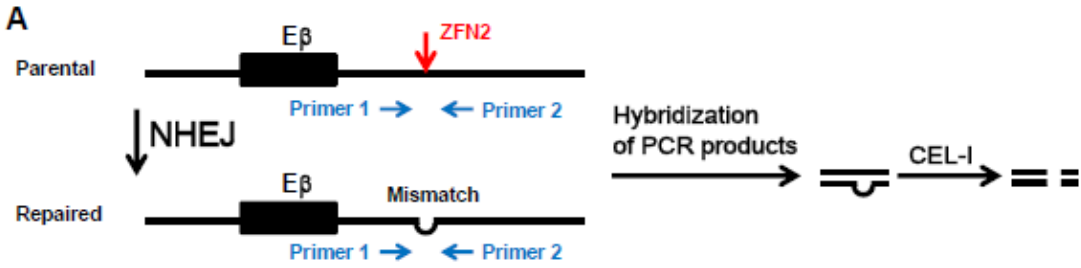
Figure 2.15 Screening and cloning of a knock-in cell model with the mutant EmRG enhancer.
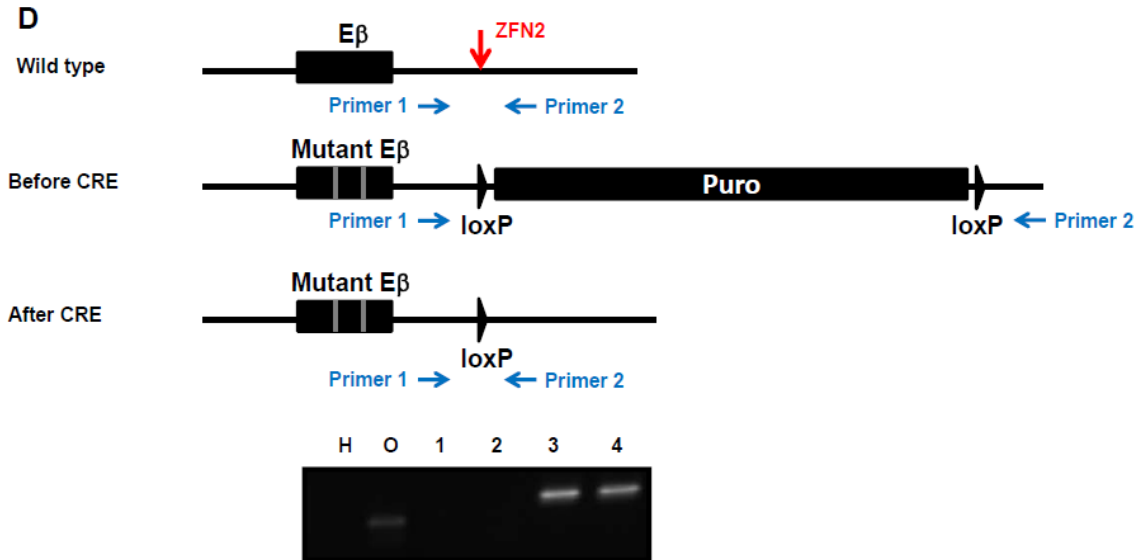
(**A**) Schematic representation and experimental result of CEL-I assay used in EmRG cell cloning. See detailed description in Fig. 2.13A. (**B**) Schematic representation and experimental design are as described in Fig. 2.5B except that samples were from EmRG cell cloning. The bottom shows representation of PCR products from different templates on 2% TBE agarose gel. O denotes O3-73 genomic DNA. H denotes water control. Numbers 1-13 are clones. Sample marked by * was used in long-range PCR in panel C. The primer and clone numbers only apply to this panel. (**C**) Schematic representation and experimental design are as described in Fig. 2.13B except that samples were from EmRG cells cloning. The lower panel shows PCR product with different templates on 1% TBE agarose gel. H denotes water control. O denotes O3-73 genomic DNA. Numbers 1 and 2 are clones. The primer and clone numbers only apply to this panel. (**D**) Schematic representation and experimental design are as described in Fig. 2.5D except that samples were from EmRG cells cloning. The primer and clone numbers only apply to this panel. The lower panel shows PCR product with different templates on 2% TBE agarose gel. H denotes water control. O denotes O3-73 genomic DNA. B denotes genomic DNA from the bulk prior to sub-cloning. Numbers 1-4 are clones.

2.7 Tables

Table 2.1 RT-qPCR and ChIP-qPCR primers

GAPDHF	GCCATCAACGACCCCTTCATTGACC
GAPDHR	GGTCTCGCTCCTGGAAGATGGTGAT
Jb1.2F	ATTGGTACCGTTAACCAGGCACAGTAGGACC
Jb1.2R	TTCCACCCGAGGTTCATCTTGACT
Jb2.1F	TATGCTGAGCAGTTCTTCGGACCA
Jb2.1R	AGTCCTGGAAATGCTGGCACAAAC
ETS1.F	GTCGATCTCAAGCCGACTCTCACCA
ETS1.R	TCCCAGTCGCTGCTGTTCTTTTGTG
Gal4.F	TCAAGGCTAGAAAGACTGGAACAGCT
Gal4.R	GGCATATCAGTCTCCACTGAAGCCA
EbFChIP	TTGACATTTACCAGGTCCTACA
EbRChIP	GTGGGGGAGCATTCACAGGCAC
pDb1FChIP	GCCCCTTCAGCAAAGATCATTTCAA
pDb1RChIP	TTGTGCAAGGTGGTGGTAAGATG
EaFChIP	CAGAAGCCACATCCTCTGGAAAGA
EaRChIP	GTGTAACGGCCCAGCCTACCTC

Note: Annealing at 60 °C

Table 2.2 Mutagenesis and cloning primers

ZFN1upF	ATCCCCAGCCAGATTCTTCT
ZFN2downR	CTCATGTGCTGTCTGCTCTGAATTC
ETS1Fmut	GAGATGGAGCTATGGCAGG
ETS1Rmut	CGGAGTACTGTCCTCCGTGTGGTTTGACATTTACCAGG
ETS2Fmut	CGGAGGACAGTACTCCGTGTGGCAAGTGTGGTTCC
ETS2Rmut	TCTCATGCAGCCCTTTCTAC
RUNX1Fmut	ACATCCTGTTGTGACTCCTG
RUNX1Rmut	CGGAGTACTGTCCTCCGTTGACATTTACCAGGTCCTACA
RUNX23Fmut	TCCCAAAATGCTCAGAATGTG
RUNX23Rmut	CGGAGTACTGTCCTCCGACATCCTGTCTTCAAACCCT
mutETS1Fmut	ACAGAAGGTTGTGACTCCTGGAGATGG
RUNX23Rmut.2	CGGAGGACAGTACTCCGCCCAAAATGCTCAGAATGTGA
mutETS2Rmut	ACAGAAGGTCTTCAAACCCTTCTCATGC
ETS1LexAmutF	ATATACAGTAGATGGAGCTATGGCAGG
ETS1LexAmutR	ATATACAGTTGTGGTTTGACATTTACCAGG
ETS2LexAmutF	ATATACAGTTGTGGCAAGTGTGGTTC
ETS2LexAmutR	ATATACAGTCTCATGCAGCCCTTTCTAC
5`armF	GCCAAGAAGATACTGGAGAATTATG
5`armR	AATCGGTCAGGGAAACAAGTATTAA
EbFclone	ATTGGTACCGTTAACCAGGCACAGTAGGACC
EbRclone	AATACGCGTGCAGGACAGACAGACCATCT
PuroF	AATACGCGTCCGATCATATTCAATAACCCTTAAT
PuroR	AATGCTAGCGAACCTCTTCGAGGGACCTA

HygroF	CCGGAGATGAGGAAGAGGAGAACA
HygroR	TCAGCAGCCTCTGTTCCACATACA
3`armF	AATGCTAGCATTCATTGGCTGGCTTTCGCT
3`armR	ATTGCGGCCGCAGGAATTTTTACTGCTGCTTGCGTC
ZFN2cutF	AGCCCTTTCTACCTCAGCCTCTA
ZFN2cutR	ACTGGGGCCCCTTGACTGCT
ETSwtGeno	ACATCCTGTTGTGACTCCTG
ETSmutGenoR	TGGTAAATGTCAAACCACACGGA
RUNXmutGenoR	TAGGACCTGGTAAATGTCAACGGA
WTRXGenoR	TGTAGGACCTGGTAAATGTCAAACC
mERXmutGenoR	AGGACAGTACTCCGACAGAAG
Gal4DBDF	ATTCTTAAGATGAAGCTACTGTCT
Gal4DBDR	ATTGATATCAGAACCACCTGAACCTCCTGATCCACCCGATACAGTCAACTG
RUNX1.184F	AATGATATCGATGATCAGACCAAGCCCGGGA
RUNX1.451R	ATTGCGGCCGCTCAGTAGGGCCGCCA
RUNX1.292R	ATTGCGGCCGCTCAACTGGAAAGTTCTGCAGAGAG
RUNX1.291F	AATGATATCTCCAGTCGACTCTCAACG
RUNX1.1.F	ATTGATATCATGCGTATCCCCGTAGATG
RUNX1.372.R	ATTGCGGCCGCTCACGAGCCGGTCTGGAAGG
RUNX1.371.F	AATGATATCGGCTCGCCCTCCTACCA
ETS1.2.F	ATTGGATCCAAGGCGGCCGTCGATCTC
ETS1.440.sR	ATTGCGGCCGCTAGTCAGCATCCGGCTTTACATCC
ETS1.242.sR	ATTGCGGCCGCTAACTGGCTCTCCCCAGGCA

ETS1.330.sR	ATTGCGGCCGCTATGTGTAGCCAGCCAGGGCA
ETS1.1.F	ATTGGTACCATGAAGGCGGCCGTCGATCTC
ETS1.440.nsR	ATTGGATCCGTCAGCATCCGGCTTTACATCC
ETS1.242.ns.R	ATTGGATCCACTGGCTCTCCCCAGGCA
ETS1.330.ns.R	ATTGGATCCTGTGTAGCCAGCCAGGGCA
LexA.F	ATTCCGCGGATGAAAGCGTTAACGGCCAG
LexA.R	ATTCCATGGTTACAGCCAGTCGCCGTT
Gal4.65.R	ATTCCATGGTCATTCCAGTCTTTCTAGCCTTGATTCCA
Ga.4.93.R	ATTCCATGGTCATAACAATGCTTTTATATCCTGTAAAGAATC

Note: Annealing at 62 °C with Phusion enzyme. mutETS1F pairs with RUNX1R to make upstream compsite site of EmRG.. 5`EbChIP in Table 2.2 and ZFN2cutR in this table are used to detect genomic Eb deletion.

2.8 References

1. Jung, D., and F. W. Alt. 2004. Unraveling V(D)J Recombination: Insights into Gene Regulation. *Cell* 116: 299–311.

2. Schatz, D. G., and Y. Ji. 2011. Recombination centres and the orchestration of V(D)J recombination. *Nat. Rev. Immunol.* 11: 251–263.

3. Jackson, A. M., and M. S. Krangel. 2006. Turning T-cell receptor ? recombination on and off: more questions than answers. *Immunol. Rev.* 209: 129–141.

4. Sikes, M. L., R. J. Gomez, J. Song, and E. M. Oltz. 1998. A developmental stage-specific promoter directs germline transcription of D beta J beta gene segments in precursor T lymphocytes. *J. Immunol.* 161: 1399–405.

5. McMillan, R. E., and M. L. Sikes. 2008. Differential activation of dual promoters alters Dbeta2 germline transcription during thymocyte development. *J. Immunol.* 180: 3218–28.

6. Osipovich, O., R. Milley Cobb, K. J. Oestreich, S. Pierce, P. Ferrier, and E. M. Oltz. 2007. Essential function for SWI-SNF chromatin-remodeling complexes in the promoter-directed assembly of Tcrb genes. *Nat. Immunol.* 8: 809–816.

7. Majumder, K., L. J. Rupp, K. S. Yang-Iott, O. I. Koues, K. E. Kyle, C. H. Bassing, and E. M. Oltz. 2015. Domain-Specific and Stage-Intrinsic Changes in Tcrb Conformation during Thymocyte Development. *J. Immunol.* 195.

8. Mathieu, N., W. M. Hempel, S. Spicuglia, C. Verthuy, and P. Ferrier. 2000. Chromatin Remodeling by the T Cell Receptor (Tcr)-β Gene Enhancer during Early T Cell Development. *J. Exp. Med.* 192.

9. Rothenberg, E. V, A. Champhekar, S. Damle, M. M. Del Real, H. Y. Kueh, L. Li, and M. A. Yui. 2013. Transcriptional establishment of cell-type identity: dynamics and causal mechanisms of T-cell lineage commitment. *Cold Spring Harb. Symp. Quant. Biol.* 78: 31–41.

10. Weber, B. N., A. W.-S. Chi, A. Chavez, Y. Yashiro-Ohtani, Q. Yang, O. Shestova, and A. Bhandoola. 2011. A critical role for TCF-1 in T-lineage specification and differentiation. *Nature* 476: 63–8.

11. Carvajal, I. M., and R. Sen. 2000. Functional analysis of the murine TCR beta-chain gene enhancer. *J. Immunol.* 164: 6332–9.

12. Hollenhorst, P. C., K. J. Chandler, R. L. Poulsen, W. E. Johnson, N. A. Speck, and B. J. Graves. 2009. DNA Specificity Determinants Associate with Distinct Transcription Factor Functions. *PLoS Genet.* 5: e1000778.

13. Giese, K., C. Kingsley, J. R. Kirshner, and R. Grosschedl. 1995. Assembly and function of a TCR alpha enhancer complex is dependent on LEF-1-induced DNA bending and multiple protein-protein interactions. *Genes Dev.* 9: 995–1008.

14. Nelson, M. L., H.-S. Kang, G. M. Lee, A. G. Blaszczak, D. K. W. Lau, L. P. McIntosh, and B. J. Graves. 2010. Ras signaling requires dynamic properties of Ets1 for phosphorylation-enhanced binding to coactivator CBP. *Proc. Natl. Acad. Sci. U. S. A.* 107: 10026–31.

15. Wee, H.-J., D. C.-C. Voon, S.-C. Bae, and Y. Ito. 2008. PEBP2-β/CBF-β–dependent phosphorylation of RUNX1 and p300 by HIPK2: implications for leukemogenesis. *Blood* 112.

16. Bruhn, L., A. Munnerlyn, and R. Grosschedl. 1997. ALY, a context-dependent coactivator of LEF-1 and AML-1, is required for TCRalpha enhancer function. *Genes Dev.* 11: 640–53.

17. Bories, J.-C., D. M. Willerford, D. Grévin, L. Davidson, A. Camus, P. Martin, D. Stéhelin, and F. W. Alt. 1995. Increased T-cell apoptosis and terminal B-cell differentiation induced by inactivation of the Ets-1 proto-oncogene. *Nature* 377: 635–638.

18. Collins, A., D. R. Littman, and I. Taniuchi. 2009. RUNX proteins in transcription factor networks that regulate T-cell lineage choice. *Nat. Rev. Immunol.* 9: 106–115.

19. Growney, J. D., H. Shigematsu, Z. Li, B. H. Lee, J. Adelsperger, R. Rowan, D. P. Curley, J. L. Kutok, K. Akashi, I. R. Williams, N. A. Speck, and D. G. Gilliland. 2005. Loss of Runx1 perturbs adult hematopoiesis and is associated with a myeloproliferative phenotype. *Blood* 106.

20. Rothenberg, E. V., J. E. Moore, and M. A. Yui. 2008. Launching the T-cell-lineage developmental programme. *Nat. Rev. Immunol.* 8: 9–21.

21. Rothenberg, E. V., and T. Taghon. 2005. MOLECULAR GENETICS OF T CELL DEVELOPMENT. *Annu. Rev. Immunol.* 23: 601–649.

22. Kim, W. Y., M. Sieweke, E. Ogawa, H. J. Wee, U. Englmeier, T. Graf, and Y. Ito. 1999. Mutual activation of Ets-1 and AML1 DNA binding by direct interaction of their autoinhibitory domains. *EMBO J.* 18: 1609–20.

23. Goetz, T. L., T. L. Gu, N. A. Speck, and B. J. Graves. 2000. Auto-inhibition of Ets-1 is counteracted by DNA binding cooperativity with core-binding factor alpha2. *Mol. Cell. Biol.* 20: 81–90.

24. Gu, T. L., T. L. Goetz, B. J. Graves, and N. A. Speck. 2000. Auto-inhibition and partner proteins, core-binding factor beta (CBFbeta) and Ets-1, modulate DNA binding by CBFalpha2 (AML1). *Mol. Cell. Biol.* 20: 91–103.

25. Yang, C., L. H. Shapiro, M. Rivera, A. Kumar, and P. K. Brindle. 1998. A role for CREB binding protein and p300 transcriptional coactivators in Ets-1 transactivation functions. *Mol. Cell. Biol.* 18: 2218–29.

26. Capecchi, M. R. 1989. The new mouse genetics: Altering the genome by gene targeting. *Trends Genet.* 5: 70–76.

27. Zwaka, T. P., and J. A. Thomson. 2003. Homologous recombination in human embryonic stem cells. *Nat. Biotechnol.* 21: 319–321.

28. te Riele, H., E. R. Maandag, and A. Berns. 1992. Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. *Proc. Natl. Acad. Sci. U. S. A.* 89: 5128–32.

29. Collin, J., and M. Lako. 2011. Concise Review: Putting a Finger on Stem Cell Biology: Zinc Finger Nuclease-Driven Targeted Genetic Editing in Human Pluripotent Stem Cells. *Stem Cells* 29: 1021–1033.

30. John, S., T. A. Johnson, M.-H. Sung, S. C. Biddie, S. Trump, C. A. Koch-Paiz, S. R. Davis, R. Walker, P. S. Meltzer, and G. L. Hager. 2009. Kinetic Complexity of the Global Response to Glucocorticoid Receptor Action. *Endocrinology* 150: 1766–1774.

31. Osipovich, O., R. Milley, A. Meade, M. Tachibana, Y. Shinkai, M. S. Krangel, and E. M. Oltz. 2004. Targeted inhibition of V(D)J recombination by a histone methyltransferase. *Nat. Immunol.* 5: 309–316.

32. Davison, J. M., C. M. Akitake, M. G. Goll, J. M. Rhee, N. Gosse, H. Baier, M. E. Halpern, S. D. Leach, and M. J. Parsons. 2007. Transactivation from Gal4-VP16 transgenic insertions for tissue-specific cell labeling and ablation in zebrafish. *Dev. Biol.* 304: 811–24.

33. Helmink, B. A., A. T. Tubbs, Y. Dorsett, J. J. Bednarski, L. M. Walker, Z. Feng, G. G. Sharma, P. J. McKinnon, J. Zhang, C. H. Bassing, and B. P. Sleckman. 2011. H2AX prevents CtIP-mediated DNA end resection and aberrant repair in G1-phase lymphocytes. *Nature* 469: 245–249.

34. Gopalakrishnan, S., K. Majumder, A. Predeus, Y. Huang, O. I. Koues, J. Verma-Gaur, S. Loguercio, A. I. Su, A. J. Feeney, M. N. Artyomov, and E. M. Oltz. 2013. Unifying model for molecular determinants of the preselection Vβ repertoire. *Proc. Natl. Acad. Sci. U. S. A.* 110: E3206-15.

35. Egawa, T., R. E. Tillman, Y. Naoe, I. Taniuchi, and D. R. Littman. 2007. The role of the Runx transcription factors in thymocyte differentiation and in homeostasis of naive T cells. *J. Exp. Med.* 204.

36. Tripathi, R. K., N. Mathieu, S. Spicuglia, D. Payet, C. Verthuy, G. Bouvier, D. Depetris, M. G. Mattei, W. M. HempeL, and P. Ferrier. 2000. Definition of a T-cell receptor beta gene core enhancer of V(D)J recombination by transgenic mapping. *Mol. Cell. Biol.* 20: 42–53.

37. Bonnet, M., F. Huang, T. Benoukraf, O. Cabaud, C. Verthuy, A. Boucher, S. Jaeger, P. Ferrier, and S. Spicuglia. 2009. Duality of Enhancer Functioning Mode Revealed in a Reduced TCR Gene Enhancer Knockin Mouse Model. *J. Immunol.* 183: 7939–7948.

38. Mombaerts, P., C. Terhorst, T. Jacks, S. Tonegawa, and J. Sancho. 1995. Characterization of immature thymocyte lines derived from T-cell receptor or recombination activating gene 1 and p53 double mutant mice. *Proc. Natl. Acad. Sci. U. S. A.* 92: 7420–4.

39. Taniuchi, I., M. Osato, T. Egawa, M. J. Sunshine, S.-C. Bae, T. Komori, Y. Ito, and D. R. Littman. 2002. Differential Requirements for Runx Proteins in CD4 Repression and Epigenetic Silencing during T Lymphocyte Development. *Cell* 111: 621–633.

40. Levanon, D., R. E. Goldstein, Y. Bernstein, H. Tang, D. Goldenberg, S. Stifani, Z. Paroush, and Y. Groner. 1998. Transcriptional repression by AML1 and LEF-1 is mediated by the TLE/Groucho corepressors. *Proc. Natl. Acad. Sci. U. S. A.* 95: 11590–5.

41. Johnson, R. D., and M. Jasin. 2001. Double-strand-break-induced homologous recombination in mammalian cells. *Biochem. Soc. Trans.* 29.

42. Vogel, C., and E. M. Marcotte. 2012. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* 13: 227.

43. Oestreich, K. J., R. M. Cobb, S. Pierce, J. Chen, P. Ferrier, and E. M. Oltz. 2006. Regulation of TCRβ Gene Assembly by a Promoter/Enhancer Holocomplex. *Immunity* 24: 381–391.

44. Bae, S.-C., and Y. H. Lee. 2006. Phosphorylation, acetylation and ubiquitination: The molecular basis of RUNX regulation. *Gene* 366: 58–66.

45. Hollenhorst, P. C., L. P. McIntosh, and B. J. Graves. 2011. Genomic and Biochemical Insights into the Specificity of ETS Transcription Factors. *Annu. Rev. Biochem.* 80: 437–471.

46. Garrett-Sinha, L. A. 2013. Review of Ets1 structure, function, and roles in immunity. *Cell. Mol. Life Sci.* 70: 3375–3390.

47. Dittmer, J. 2003. The Biology of the Ets1 Proto-Oncogene. *Mol. Cancer* 2: 29.

48. Ishihara, K., M. Oshimura, and M. Nakao. 2006. CTCF-Dependent Chromatin Insulator Is Linked to Epigenetic Remodeling. *Mol. Cell* 23: 733–742.

49. Qiu, G.-H., C. H.-W. Leung, T. Yun, X. Xie, M. Laban, and S. C. Hooi. 2011. Recognition and suppression of transfected plasmids by protein ZNF511-PRAP1, a potential molecular barrier to transgene expression. *Mol. Ther.* 19: 1478–86.

50. Segal, D. J., and C. F. Barbas. 2000. Design of novel sequence-specific DNA-binding proteins. *Curr. Opin. Chem. Biol.* 4: 34–39.

51. Zetsche, B., S. E. Volz, and F. Zhang. 2015. A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nat. Biotechnol.* 33: 139–42.

52. Golemis, E. A., and R. Brent. 1992. Fused protein domains inhibit DNA binding by LexA. *Mol. Cell. Biol.* 12: 3006–14.

53. Merkenschlager, M., and D. T. Odom. 2013. CTCF and Cohesin: Linking Gene Regulatory Elements with Their Targets. *Cell* 152: 1285–1297.

54. Chattopadhyay, S., C. E. Whitehurst, and J. Chen. 1998. A Nuclear Matrix Attachment Region Upstream of the T Cell Receptor Gene Enhancer Binds Cux/CDP and SATB1 and Modulates Enhancer-dependent Reporter Gene Expression but Not Endogenous Gene Expression. *J. Biol. Chem.* 273: 29838–29846.

55. Berman, B. P., D. J. Weisenberger, J. F. Aman, T. Hinoue, Z. Ramjan, Y. Liu, H. Noushmehr, C. P. E. Lange, C. M. van Dijk, R. A. E. M. Tollenaar, D. Van Den Berg, and P. W. Laird. 2011. Regions of focal DNA hypermethylation and long-range hypomethylation in colorectal cancer coincide with nuclear lamina-associated domains. *Nat. Genet.* 44: 40–6.

Chapter 3 : Activation of Mouse *Tcrb*: Uncoupling RUNX1 Function from its Cooperative Binding with ETS1

Jiang-yang Zhao, Oleg Osipovich, Olivia I. Koues, Kinjal Majumder, Eugene M. Oltz. Activation of Mouse *Tcrb*: Uncoupling RUNX1 Function from its Cooperative Binding with ETS1.

3.1   Abstract

T lineage commitment requires the coordination of key transcription factors (TFs) in multi-potent progenitors (MPP) that transition them away from other lineages and cement T cell identity. Two important TFs for the MPP to T lineage transition are RUNX1 and ETS1, which bind cooperatively to composite sites throughout the genome, especially in regulatory elements for genes involved in T lymphopoiesis. Activation of the T cell receptor β (*Tcrb*) locus in committed thymocytes is a critical process for continued development of these cells, and is mediated by an enhancer, Eβ, which harbors two RUNX-ETS composite sites. An outstanding issue in understanding T cell gene expression programs is whether RUNX1 and ETS1 have independent functions in enhancer activation that can be dissected from cooperative binding. We now show that RUNX1 is sufficient to activate the endogenous mouse Eβ element and its neighboring 25 kb region by independently tethering this TF without coincidental ETS1 binding. Moreover, RUNX1 is sufficient for long-range promoter-Eβ looping, nucleosome clearance, and robust transcription throughout the *Tcrb* recombination center (RC), spanning both DβJβ

clusters. We also find that a RUNX1 domain, termed the negative regulatory domain for DNA binding (NRDB), can compensate for the loss of ETS1 binding at adjacent sites. Thus, we have defined independent roles for RUNX1 in the activation of a T cell developmental enhancer, as well as its ability to mediate specific changes in chromatin landscapes that accompany long-range induction of RC promoters.

3.2   Introduction

Cell fate commitment is coordinated largely by the temporal regulation of lineage-specifying transcription factors (TFs). During hematopoiesis in mammals, precursor cells in the thymus maintain developmental plasticity until they extinguish "master" TFs for the myeloid and B cell lineages, including SPI1 (1). Conversely, these precursor cells must induce or maintain expression of TFs that control key aspects of T-lineage expression programs, including BCL11b, TCF1, RUNX1, and ETS1 (1). Among the genomic targets for lineage-specifying factors are regulatory regions, dubbed super-enhancers (SEs), which are densely-spaced collections of enhancer elements associated with genes important for cell identity and function (2).

For T cell precursors, one such SE is associated with the locus encoding T cell receptor β (*Tcrb*) (3), which has a powerful enhancer (Eβ) situated near its 3' terminus. Eβ is essential for generating the SE, which spans both clusters of DβJβ gene segments and neighboring promoters, activating their expression as sterile transcripts in CD4⁻CD8⁻, double-negative (DN) thymocytes (4–6). Potent transcription and deposition of activating histone modifications over the DβJβ-Eβ

82

super-enhancer (7, 8), render it a prime target for RAG-1 and RAG-2, the V(D)J recombinase. Indeed, the *Tcrb*-SE has been designated as a "recombination center (RC)" due to its high load of RAG-1/2 in DN thymocytes that, in turn, triggers rapid recombination of DβJβ gene segments, an essential first step in the assembly of complete *Tcrb* genes (9, 10).

The critical regulatory element for formation of the *Tcrb*-RC is Eβ, which, like many enhancers that contribute to T-lineage SEs, has adjacent binding sites for the RUNX and ETS family of TFs (7, 11, 12). The biological significance of recurrent ETS/RUNX motifs is further underscored by their enrichment in T cell SEs harboring SNPs associated with auto-inflammatory conditions (13). The RUNX and ETS families are defined by their respective DNA binding regions, termed Runt and ETS domains, respectively (14, 15). Although these TFs can bind to isolated sites in the genome, compound ETS-RUNX motifs are commonly employed to attain tissue-specific cooperativity between family members in a number of cell lineages. For example, RUNX1 and the ETS member, SPI1, cooperatively regulate a gene encoding the M-CSF receptor in myeloid cells (16), while ETS1 and RUNX2 bind cooperatively to an element controlling the osteopontin gene in colorectal cancer cells (17). In T-lineage cells, RUNX1 and ETS1 target Eβ to generate the *Tcrb* super-enhancer and initiate its assembly (7).

At the biochemical level, both RUNX1 and ETS1 harbor domains that partially inhibit their binding to single, cognate sites, a feature that is overcome when the two bind cooperatively to composite sites (18–20). Specifically, regions flanking the ETS1 DNA binding domain undergo a significant conformational change when its exon-VII domain interacts with RUNX1, resulting

in a higher affinity of ETS1 binding to composite sites (19, 20). RUNX1 is stabilized by its interaction with a second protein, CBFβ, which protects RUNX1 from proteolysis and induces conformational changes in its Runt domain that enhance DNA binding (18, 21). However, the Runt domain of RUNX1 binds poorly to its cognate sites in chromatin until a region termed the "Negative Regulatory DNA-Binding" (NRDB) interacts with ETS1 (18, 20). Once bound to genomic sites, the transactivation domains (TADs) of ETS1 and RUNX1 recruit distinct sets of downstream factors that potentiate enhancer activities. Importantly, the TADs of both ETS1 and RUNX1 interact with the histone acetyltransferases p300/CBP, key enhancer-associated factors that alter epigenetic landscapes for gene activation (22, 23).

During T-lineage commitment, multi-potent progenitors in the thymus constitutively express RUNX1 but not ETS1 (24). Upon signaling through Notch and other receptor pathways, these progenitors induce expression of BCL11b and TCF1, which, in turn, activate the *Ets1* gene (1, 24, 25). Thus, ETS1 expression is a key development switch, pairing with RUNX1 to target cohorts of regulatory elements that control genes important for the T cell program, including the *Tcrb*-RC (26). Indeed, inactivation of either TF in mice leads to profound defects in T cell development and activation (27, 28). Because of its central role in T cell development, Eβ and its pair of ETS-RUNX composite sites have been studied extensively, especially with regard to activation of the *Tcrb* super-enhancer (7). Mutation of both ETS-RUNX composite sites within Eβ phenocopies a deletion of the entire enhancer (7, 20, 29, 30), including the loss of: (i) looping to promoters near the Dβ1 and Dβ2 gene segments (7, 31), (ii) transcription derived from these promoters (7, 30, 31), (iii) active histone modification over both DβJβ clusters (7, 31–33), and

(iv) DβJβ recombination (7, 30, 31). Thus, Eβ provides an excellent model to test mechanisms by which RUNX1 and ETS1 coordinate activation of a developmentally important super-enhancer, the *Tcrb*-RC, either through their individual or cooperative functions.

In this regard, we previously proposed a stepwise model for *Tcrb* activation in which promoter-enhancer holocomplex formation precedes nucleosome clearance, transcription, and activation (31). However, the unique roles for ETS1 and RUNX1 in each of these steps at *Tcrb* or any locus remains unclear due primarily to the requirement for cooperative binding of the TFs at their composite enhancer sites. Moreover, most cellular systems in which these mechanistic questions can be addressed either rely on extrachomosomal reporters or asynchronous activation of an endogenous gene (20, 34). Here, we report unique and cooperative functions of RUNX1 during activation of the endogenous *Tcrb*-RC using precursor T cells in which a Gal4 fusion of this TF is inducibly targeted to Eβ. We find that, when uncoupled from its need to bind cooperatively with ETS1, RUNX1 is sufficient to induce promoter-enhancer looping and transcription throughout the *Tcrb*-RC. Long-range transactivation by RUNX1 is uncoupled from extensive revisions to the chromatin landscape, features normally associated with promoter-driven transcription. Thus, although expression of RUNX1 is sufficient for transcriptional activation of the *Tcrb*-RC, ETS1 co-recruitment during T-lineage commitment may be required to boost *Tcrb*'s epigenetic status to that of a super-enhancer, ensuring its efficient assembly and persistent expression.

3.3   Materials and Methods

shRNA depletion of ETS1 and RUNX1

Previously reported shRNA target sequences for *Ets1* and *Runx1* (35, 36) were cloned into the pSIREN shuttle vector (Clontech), excised as BglII-MluI fragments, and inserted into pcDNA3.1 containing a truncated human CD4 cDNA (37). P5424 cells were cultured as described (38) and were transfected by electroporation (Biorad, 250 V/960 μF) with control (scrambled GFP 5'-AAGCTGGAGTACAACTACA-3'), ETS1-specific (target sequence 5'-GCAGACAGACTACTTTGCCAT-3'), or RUNX1-specific shRNA vectors (target sequence 5'-GCCCTCCTACCATCTATACTA-3'). Transfected cells were purified 48 hours post-transfection with an EasySep Human CD4+ T Cell Enrichment Kit (Stem Cell Technologies #19052).

Construction of cell models

P5424 cells were engineered to harbor only a single allele of the *Tcrb*-RC, termed O3-73 clone, by concurrent expression of two Zinc Finger Nucleases (ZFNs; Sigma). Each pair of ZFNs targets a DNA break to a site in *Tcrb*, one located upstream of Dβ1 (ZFN1 pair, target 5'-GGACTTGTCCTTAACTCCCTTTACCTAGCAAGATAGG-3') and a second, located downstream of Eβ (ZFN2 pair, target 5'-GTCTGTCTGTCCTGCATTCATTGGCTGGCTTTCGCT-3'). $5x10^6$ cells were transfected with 5 μg of each ZFN expression plasmid using electroporation (250 V/960 μF). After two days

86

of recovery, cells were cloned by serial dilution and screened by PCR assays for the large *Tcrb*-RC deletion (Table 3.2). Clone O3-73 was selected as a parental cell line for future experiments (allele 1: intact *Tcrb*-RC, allele 2: deletion of Dβ1-Dβ2-Eβ).

Targeting vectors for mutation of Eβ were generated as follows. Gal4 or TF site mutations were first introduced into a 1.85 kb fragment spanning Eβ (HpaI-HindIII), using PCR-assisted mutagenesis with a high-fidelity polymerase according to the manufacturer's protocol (Phusion, NEB M0503). Oligonucleotide sequences used for site-directed mutagenesis and cloning are provided in Table 3.2. All Eβ mutations were verified by sequencing. Mutagenesis was performed sequentially to alter both ETS-RUNX sites in Eβ. Targeting vectors were generated in pBS by sequentially cloning a 3.2 kb 5' homology arm (BamHI-HpaI), a 0.7 kb mutant Eβ (HpaI-MluI), a 1.9 kb loxP:PGK-promoter:Puro:loxP cassette (MluI-NheI), and a 2.1 kb 3` homology arm (NheI-NotI). For the ΔEβ targeting vector, a hygromycin-resistant cassette flanked by FRT sites was used for drug selection.

For targeting, $5x10^6$ O3-73 cells were co-transfected with 20 μg of a targeting vector and 5 μg each of expression plasmids for the ZFN2 pair, which introduce a DNA break ~200 bp downstream of Eβ. Two days post-electroporation, cells were selected with puromycin (1 μg/mL) or hygromycin (100 μg/mL) and cultured for 5-6 d before sub-cloning by limiting dilution. Initially, single cell clones were screened by a PCR assay in which insertion of the targeting vector disrupts amplification (Table 3.2). Candidate recombinant clones were then

screened by a long-range PCR assay for proper integration (Table 3.2), which was verified by sequencing. Successfully targeted clones were transfected with 1 µg of human CD4 expression plasmid and 10 µg of CRE or FLP expression plasmids and purified for CD4 expression. After one week, cultures were subcloned, screened for deletion of the drug marker cassette, and mutant alleles were verified by sequencing.

Gal4 fusion proteins

Expression vectors for all Gal4-RUNX1 fusion proteins were prepared from an acceptor plasmid containing a cDNA for the Gal4 DNA-binding domain (DBD, aa 1-147), into which was cloned, in-frame, cDNAs for portions of murine RUNX1. Each of these RUNX1 cDNAs was generated by PCR using primers provided in Table 3.2 (RUNX1^{GNT} aa 184-451, RUNX1^{GN} aa 182-292, and RUNX1^{GT} aa 291-451). The final fused Gal4-RUNX1 cDNAs were then excised and cloned into pcDNA3.1. Each version was validated by sequencing and protein expression in 293T cells. For inducible expression, each variant was shuttled into the pFLRU:Thy1.1 tet-responsive lentiviral vector. Viruses containing each Gal4-RUNX1 expression cassette were packaged as described previously (39) and used to spin-infect $5x10^6$ O3-73 derivative cells. Infected cells were cultured for 10 days, sorted for high Thy1.1 expression by flow cytometry (Anti-Mouse/Rat CD90.1 APC, eBioscience 17-0900-82), expanded, and induced with 1 µg/mL doxycycline (Dox) for 24 hours prior to molecular analysis.

RT-qPCR

Total RNA was extracted by TRIzol (Roche #1697478) from O3-73 and its derivatives. In each reaction, 1 μg total RNA was digested by RQ1 RNase-free DNase (Promega M199A) in 20 μL total volume, as recommended by the manufacturer. The treated RNA was split equally for cDNA preparation with M-MuLV Reverse Transcriptase (NEB M0253) or used as a no RT control, with random hexamers. Real-time qPCR was performed with SYBR Green JumpStart Taq ReadyMix (Sigma S4438) using primers and conditions provided in Table 3.1.

Chromatin Immunoprecipitation

ChIP assays were performed as described previously (40). Sheared chromatin from O3-73 and its derivatives were prepared after fixation with 1% formaldehyde and sonication with Diagenode Bioruptor Pico for 20 cycles (30 s each, vortexed every 5 cycles). ChIP antibodies (all rabbit IgG isotypes) were coupled to Protein A Dynabeads and were obtained from the following sources: RUNX1 C terminus was provided by Dr. Takeshi Egawa (41), ETS-1 C20 (Santa Cruz sc-350x), normal rabbit IgG (Santa Cruz sc-2027x), Gal4DBD (Santa Cruz sc-577x), total Histone H3 (Abcam ab1791), H3K4Me3 (Abcam ab8580) and H3K27Ac (Abcam ab4729). Eluted ChIP and input DNAs were purified with Qiagen QIAquick PCR purification kit and analyzed by qPCR using primers and conditions provided in Table S1.

Chromosome Conformation Capture assays

3C assays were performed on 10^7 O3-73 cells and its derivatives fixed with 2% formaldehyde. Chromatin was isolated and digested with HindIII, as described (42, 43). Bacterial artificial

chromosomes (BACs) spanning the entire *Tcrb* locus and a control *ERCC3* locus were digested with HindIII, ligated, and used as controls for Taqman qPCR efficiency in all 3C analyses (43). Primers and probes for these Taqman assays are listed in Table S1.

IP- and conventional western blotting

O3-73 cell lysates were prepared in RIPA buffer ($2x10^7$ cells in 300 μL) and sonicated for 40 cycles (30s, vortexed each 10 cycles, Diagenode Bioruptor Pico). For IP-westerns, each lysate was incubated with rabbit anti-Gal4DBD antibody (1 μg, Santa Cruz sc-577x) attached to Dynabeads M-280 sheep anti-rabbit IgG (50 μL) as described by the bead manufacturer. For conventional western blotting, sonicated lysates were used directly (10 μL). Whole cell or IP lysates were fractionated on 12% PAGE gels, transferred to PVDF membranes, incubated with detection antibodies and visualized by ECL substrate (Thermo Fisher Scientific #34087). Primary antibodies were used at a concentration of 1 μg/mL for protein detection and obtained from the following sources: mouse anti-Gal4DBD antibody (clone RK5C1; Santa Cruz sc-510), rabbit anti-ETS1 (clone N276; Santa Cruz sc-111), and mouse anti-GAPDH (clone G-9; Santa Cruz sc-365062). Secondary antibodies, diluted $1:10^4$, were goat anti-rabbit IgG-HRP (Santa Cruz sc-2004) and goat anti-mouse IgG-HRP (Promega W4021).

3.4 Results

Eβ function requires both ETS1 and RUNX1 in precursor T cells

The critical element for conferring SE status to *Tcrb* is its conventional enhancer, Eβ (8), which contains a pair of composite ETS-RUNX binding motifs as its core (Fig. 3.1A). To dissect the individual functions of these TFs in driving Eβ activation, we required a genetically tractable system for introducing mutations at the endogenous element. Such an approach in primary thymocytes would be unwieldy, since these cells fail to grow in long-term cultures. Thus, genetic and biochemical dissection of ETS-RUNX function at Eβ would require in vivo studies of multiple mutant mouse lines. To circumvent this obstacle, we selected a cell line that serves as an excellent model for precursor T cells based on transcriptome-wide data, as well as the transcriptional status of the *Tcrb* recombination center (44, 3). The cell line, P5424, is a thymoma derived from RAG1/P53-deficient mice, which retains a germline but transcriptionally active *Tcrb* locus (44). To serve as an appropriate model for dissecting ETS-RUNX cooperativity, Eβ function in P5424 must depend on these individual transcription factors. Indeed, depletion of either ETS1 or RUNX1 using shRNAs significant decreased germline Dβ1 or Dβ2 transcripts in P5424 (Figs. 3.1B and C).

These data; however, do not rule out the possibility that ETS1 or RUNX1 regulate *Tcrb* transcription independent of their effects on Eβ, including a potential role at the Dβ1 or Dβ2 promoters (5, 6). To definitively assess whether these factors are essential for Eβ function, we targeted the endogenous enhancer in P5424, introducing mutations at the ETS or RUNX motifs. For this purpose, we used a version of P5424, called O3-73, in which we had deleted the *Tcrb*-RC (both DβJβ clusters and Eβ) from one allele (Fig. 3.2A, see Materials and Methods). Initially, we made three mutant versions of Eβ in O3-73 using targeting vectors in combination with a zinc finger nuclease that introduces double-strand breaks at a site ~200 bp downstream of the core enhancer. The three mutant *Tcrb* alleles harbored (i) a complete Eβ deletion (ΔEβ), (ii) a replacement of both ETS1 sites in Eβ with Gal4 binding motifs (EG), or (iii) a replacement of both RUNX1 sites in the enhancer with Gal4 binding motifs (RG, Fig. 3.2A). All three of these targeted mutations were verified by sequencing. Replacement of either the ETS or RUNX motifs with Gal4 sites abolished binding of ETS1 and RUNX1 to the enhancer, respectively (Fig. 3.2B), as measured by chromatin immunoprecipitation (ChIP). Consistent with cooperative binding (11, 18, 20), the ETS mutations also abolished RUNX1 binding to the enhancer, while destruction of the RUNX motifs dramatically attenuated ETS1 binding to Eβ. Importantly, mutation of either the ETS or RUNX binding sites in Eβ crippled germline transcription throughout the recombination center, mimicking the ΔEβ allele, which completely lacks the enhancer (Fig. 3.2C). We conclude that Eβ function in this precursor T cell model depends completely on its two ETS-RUNX composite motifs.

Tethering of RUNX1 rescues mutant Eβ function

Replacement of the two RUNX motifs in Eβ with Gal4 sites allowed us to recruit mutant versions of RUNX1 to the enhancer for functional dissection. As a proof of concept, we introduced into O3-73 mutants a lentiviral vector that is inducible by doxycycline (Dox) for expression of a fusion protein in which the Runt domain of RUNX1 was replaced with the Gal4 DNA binding domain (RUNX1^{GNT}, Fig. 3.3A). The RUNX1^{GNT} fusion retains both the putative ETS1 binding domain (NRDB) and its transactivation domain. Induced levels of RUNX1^{GNT} protein were similar in all O3-73 variants, and did not impact the expression of endogenous ETS1 (Fig. 3.3B). Gal4-RUNX1 fusions efficiently targeted the mutant Eβ enhancer (RG), partially restoring ETS1 binding as measured by ChIP (Fig. 3.3C). Importantly, tethering of RUNX1^{GNT} was sufficient to rescue transcription of germline *Tcrb* alleles harboring the RG mutant enhancer, reaching levels comparable to those observed in the parental O3-73 line (Fig. 3.3D). As a control, the Gal4DBD failed to rescue ETS1 or RUNX1 binding at Eβ, nor did it rescue transcription of the *Tcrb*-RC. Moreover, germline Dβ1/Dβ2 transcription remained silent when RUNX1^{GNT} was expressed in O3-73 lines lacking the enhancer element (ΔEβ), excluding potential indirect effects of the fusion protein. These data demonstrate that recruitment of RUNX1 proteins containing only its NRDB and TAD portions is sufficient to restore Eβ function when binding of endogenous RUNX1 is precluded.

ETS1 binding is dispensable for restoration of enhancer function by RUNX1

One hypothetical model for ETS-RUNX cooperativity is that either transcription factor alone may be sufficient for enhancer activation, but stable binding of either to its target sites requires NRDB/exon VII interactions and co-occupancy of composite sites. Although tethering of RUNX1^{GNT} to the RG mutant enhancer fully rescued *Tcrb* germline transcription, the fusion also recruited significant amounts of ETS1 to Eβ. To test whether RUNX1 binding to Eβ is sufficient for enhancer activation in the absence of ETS1 occupancy, we generated a fourth version of O3-73, in which the RUNX motifs of Eβ were replaced with Gal4 sites and the two adjacent ETS motifs were destroyed (EmRG, see Fig. 3.2A). We then introduced RUNX1^{GNT} and Gal4-DBD expression vectors into the mutant O3-73 cells, which exhibited comparable steady-state levels of each protein (Fig. 3.4A). Additional mutation of the ETS motifs did not impair binding of RUNX1^{GNT} at Eβ (Fig. 3.4B, top panel, compare RG and EmRG). In contrast, destruction of the ETS motifs completely compromised ETS1 binding to the enhancer, despite elevated levels of RUNX1^{GNT} (Fig. 3.4B, bottom panel). These data indicate that the NRDB domain of RUNX1 cannot recruit ETS1 to Eβ independently of the ETS sites in its composite motifs. Although the EmRG enhancer lacked detectable ETS1, recruitment of RUNX1^{GNT} completely restored the expression of germline DβJβ transcripts (Fig. 3.4C). We conclude that ETS1 is dispensable for Eβ function if RUNX1 recruitment is uncoupled from its cooperative binding to composite ETS-RUNX motifs.

Dissection of RUNX1 functional domains for Eβ activation

The RUNX1 fusion that was sufficient to fully activate Eβ contained two previously defined

functional domains (20). The NRDB has been shown in vitro to interact with exon VII of ETS1,

presumably stabilizing the binding of both factors to composite sites (20). However, our data

suggest that ETS1 recruitment is dispensable for activation by the tethered RUNX1^{GNT} fusion

(Fig. 3.4). The transactivation domain (TAD) of RUNX1, therefore, may be sufficient to support

Eβ function if recruited to the enhancer independently of ETS1.

To further define functions of the two RUNX1 domains, we first expressed a Gal4 fusion of the

NRDB (RUNX1^{GN}, Fig. 3.3A) in O3-73 cells with mutant versions of Eβ (Fig. 3.5A). As shown

in Fig. 3.5B, the RUNX1^{GN} fusion was recruited to Eβ as efficiently as RUNX1^{GNT}, which has

both the NRDB and TAD. However, only the TAD-containing fusion permitted recruitment of

ETS1 to its neighboring sites in the RG mutant version of Eβ. These data indicate that the NRDB

of RUNX1 is insufficient for ETS1 recruitment, perhaps due to a lack of chromatin accessibility

at the mutant enhancer (see below). Consistent with these results, RUNX1^{GN} failed to activate

germline transcription from the distal Dβ1-Jβ1 cluster, while it induced only modest levels of

Dβ2-Jβ2 transcripts (Fig. 3.5C). Thus, the NRDB of RUNX1 is insufficient for ETS1

recruitment to Eβ and activation of this enhancer element.

To test whether the TAD of RUNX1 is sufficient for Eβ activation, we expressed a Gal4-TAD

fusion (RUNX1^{GT}, Fig. 3.3A) that was tethered to mutant enhancers with or without intact ETS

motifs (Fig. 3.5A). The RUNX1^{GT} fusion tethered efficiently to both the RG and EmRG

95

enhancer mutants (Fig. 3.5D). As expected, ETS1 was absent from the EmRG enhancer even after RUNX1^{GT} recruitment. In contrast, RUNX1^{GT} permitted binding of ETS1 to the RG enhancer, despite its lack of an NRDB domain, which is thought to be essential for cooperative binding (Fig. 3.5D). This finding may reflect the ability of RUNX1^{GT} to endow Eβ with sufficient chromatin accessibility, allowing for a substantial, albeit sub-wild type, levels of ETS1 binding to its exposed site (see below).

Importantly, the TAD of RUNX1 was sufficient to fully activate germline transcription of the *Tcrb*-RC when tethered to the RG enhancer, to which ETS1 also binds (Fig. 3.5E). In sharp contrast, when ETS1 binding was precluded at the EmRG enhancer, RUNX1^{GT} failed to activate transcription from the distal Dβ1 promoter. The Gal4-TAD fusion; however, supported modest expression levels from the more proximal Dβ2 promoter, when tethered to the EmRG enhancer (Fig. 3.5E). Together, these data provide several new insights into the function of RUNX1 domains at RUNX-ETS composite sites in Eβ, including (i) the TAD is sufficient for enhancer activation when modest levels of ETS1 are present, but is insufficient in the absence of ETS1 at the enhancer, (ii) the NRDB is dispensable for some ETS1 recruitment to these composite sites, (iii) the NRDB, in some manner, must compensate for an absence of ETS1 given that RUNX1^{GNT} fully activates the EmRG enhancer, and (iv) when activated solely by the TAD, enhancer function is limited spatially to the most proximal Dβ2 promoter. Thus, either ETS1 or efficient binding of RUNX1 by itself is required for maximal, long-range activation by Eβ.

Epigenetic mechanisms for short- versus long-range activation of the *Tcrb*-RC

Initiation and augmentation of transcription by distal enhancers is a multi-step process involving revisions to the regional epigenetic landscape and conformational changes that drive enhancer-promoter contacts (45, 46). Alterations in the epigenetic landscape germinate at the enhancer and may spread into neighboring regions, permitting access to nuclear factors (47). Enhancer-promoter looping is thought to deliver nucleosome remodelers and components of the basal transcription machinery to the promoter after initial recruitment to an activated/accessible enhancer (31, 48).

Given that the RUNX1 TAD can partially activate the proximal Dβ2, but not the more distal Dβ1 region, we tested whether specific aspects of multi-step promoter and SE activation are supported by this domain when compared with the complete RUNX1 fusion. As shown in Fig. 3.6A, the H3K27ac activation mark, a hallmark of SEs, is at least partially restored at the proximal Dβ2Jβ2 cluster when any TAD-containing version of RUNX1 fusions are recruited to Eβ, paralleling transcriptional activation of this region. Indeed, only modest levels of both H3K27ac and Dβ2Jβ2 transcription are observed when the RUNX1^{GT} is recruited to the EmRG enhancer lacking ETS1. By comparison, mutant versions of Eβ had little impact on levels of a promoter activation mark, H3K4me3 (Fig. 3.6A). Complete deletion of Eβ further reduced each of the activation marks, but at Dβ2Jβ2, H3K27ac and H3K4me3 remained above background levels observed in 3T3 fibroblasts (Figs. 3.7 E and F). These data indicate that mutant versions of

Eβ (RG and EmRG) retain some ability to modify chromatin and that the Dβ2 region has an inherent, Eβ-independent function to partially activate chromatin in T-lineage cells.

In contrast to the proximal Dβ2Jβ2 cluster, recruitment of RUNX1 fusion proteins to mutant versions of Eβ induced only modest increases in H3K27ac levels and partial restoration of H3K4me3 at the distal Dβ1Jβ1 region (Fig. 3.6A). Despite the minor impact on active chromatin marks, many of the RUNX1 fusions fully restored Dβ1Jβ1 germline transcription. We conclude that the long-range transactivation by RUNX1 is uncoupled from general revisions to the epigenetic landscape of Dβ1Jβ1 normally associated with SE formation and promoter-driven transcription.

A downstream effect of these epigenetic revisions may be changes in nucleosome density at promoters (32, 49), driving germline Dβ1 or Dβ2 transcription. As such, we performed ChIP assays for the H3 component of nucleosomes at the proximal Dβ2 and the distal Dβ1 promoters, as shown in Fig. 3.6B. Inactivating mutations of Eβ significantly increased nucleosome densities on both Dβ1 and Dβ2 promoters. Recruitment of most RUNX1 fusions partially or fully reversed the enhanced nucleosome occupancy at the promoters, compared with alleles harboring a fully active Eβ. One exception was the RUNX1^{GN} fusion, which lacks the TAD and failed to activate any transcription from the *Tcrb*-RC. Nucleosome density at Dβ1 and Dβ2 promoters was also elevated when the RUNX1^{GT} fusion was recruited to the mutant version of Eβ lacking its ETS1

binding sites (EmRG). These data suggest that full promoter activation by RUNX1 involves nucleosome clearance, a process requiring either the NRDB domain or cooperation with ETS1.

Our prior studies have shown that nucleosome clearance at Dβ1 involves Eβ looping to its promoter, which facilitates recruitment of the SWI/SNF chromatin remodeling complex (31, 32). Indeed, transcriptional activation of Dβ1 can be achieved in the absence of its promoter and Eβ looping if SWI/SNF is artificially tethered to this region (32). To examine the impact of RUNX1 fusions on promoter-enhancer loops in the *Tcrb*-RC, we performed chromosome conformation capture (3C) assays using Eβ as a viewpoint. As expected, inactivation of Eβ disrupts its communication with both Dβ promoter regions (Fig. 3.6C). Recruitment of RUNX1 to mutant enhancers with intact or disrupted ETS1 binding sites restored Eβ looping to both Dβ elements. However, tethering of only the RUNX1 TAD to the enhancer with ETS sites mutations failed to restore Eβ association with either Dβ1 or Dβ2. Collectively, chromatin and conformational data indicate that, like transcription, enhancer-promoter looping in the *Tcrb*-RC requires either RUNX1 with an intact NRDB, or cooperative binding of at least some ETS1. Moreover, these stringent requirements for Dβ1 activation are partly dispensable for Dβ2, which may rely on Eβ-independent chromatin modifications to offset any disruptions in looping or promoter accessibility. For long-range activation of Dβ1; however, both enhancer looping and subsequent changes in chromatin accessibility are required for promoter activation, even if histone modifications associated with SE status are not fully restored

3.5   Discussion

Cell fate decisions during development are initiated by TF combinations that coordinate lineage-specifying gene expression programs through their binding to conventional- and super-enhancers (2). Two such TFs, ETS1 and RUNX1, are particularly important for T lineage commitment, during which, they bind in a cooperative manner to large cohorts of enhancers controlling T cell expression programs. In this study, we probed the specific functions of RUNX1 in activating a super-enhancer region for T lineage commitment, *Tcrb*, independent of its requirement for cooperative binding with ETS1 to its target sites in Eβ. We found that RUNX1 was sufficient to initiate both promoter-Eβ looping and germline transcription within both DβJβ clusters when tethered directly to the enhancer. These data suggest that ETS1 and its transactivation domain (TAD) are dispensable for Eβ-mediated activation of the *Tcrb* super-enhancer in precursor T cells. Instead, the critical function of ETS1 during this lineage-commitment process may be to stabilize association of RUNX1 with the enhancer via cooperative binding. A contingent binding of RUNX1 to Eβ and other enhancers important for T lineage genes is consistent with the temporal expression of the TFs during thymocyte development (1, 25). Specifically, RUNX1 is constitutively expressed in many hematopoietic cells, including multi-potent progenitors. In these T lineage precursors, ETS1 is initially silent, precluding premature gene activation via the binding of RUNX1 to composite sites in key enhancers. When the multi-potent progenitors receive T lineage commitment cues, including Notch signaling (25), subsequent expression of ETS1 permits RUNX1 binding to composite sites, activating their target enhancers. In this way,

100

ETS1 serves as a developmental gatekeeper for RUNX1-dependent expression programs that guide T cell development.

The RUNX1-dominant model raised additional questions about specific roles for each of its domains during *Tcrb* activation. Prior studies of RUNX1 and its ETS1 binding partner had mainly focused on their DNA binding and interacting domains using purified proteins and reporter assays (19, 20). These *in vitro* studies implicated that the NRDB of RUNX1 was essential for its association with the ETS1 exon VII domain (20). We find that the NRDB was neither necessary nor sufficient for ETS1 binding to the endogenous Eβ element. However, both of these observations may relate primarily to effects on neighboring chromatin accessibility rather than direct protein-protein contacts. Indeed, the tethered NRDB failed to clear nucleosomes from Eβ, as measured by histone H3 occupancy, likely inhibiting the recruitment of ETS1 to an adjacent site. Conversely, tethering of the RUNX1-TAD generated accessible chromatin over Eβ, which presumably led to modest binding of ETS1 at neighboring sites. The chromatin opening function of the RUNX1-TAD may be mediated by its recruitment of the histone acetyltransferase, p300 (23), leading to SWI/SNF association via its binding to acetylated histones (49).

Notwithstanding, we provide the first evidence, to our knowledge, that the NRDB has any function in transcriptional activation, since it can compensate for the complete absence of ETS1 at composite RUNX-ETS sites. Specifically, we found that the TAD of RUNX1, by itself, could not activate Eβ when the neighboring ETS1 sites were mutated. Enhancer activation; however,

was rescued in the absence of ETS1 when both the NRDB and TAD of RUNX1 were tethered to their native locations. Future studies using similar approaches may reveal the mechanisms by which the NRDB can compensate for enhancer-activating functions of ETS1. Taken together, our findings suggest that, in the native enhancer, the NRDB may contribute to a functional synergy with the TADs of RUNX1 and ETS1.

Like many developmentally controlled enhancers, Eβ-directed activation of the *Tcrb*-RC involves multiple revisions to its epigenetic and structural landscape, including histone modification, chromatin remodeling, and the formation of enhancer-promoter loops. The temporal and contingent relationships between these processes, leading ultimately to transcriptional activation, remains poorly understood at nearly all loci. Our study provided some key clues to the step-wise activation of a super-enhancer associated with the *Tcrb*-RC. First, we found that transcription could be largely uncoupled from acquisition of histone modifications over the super-enhancer region. This was especially true for long-range activation at the Dβ1Jβ1 cluster. Recruitment of various Gal4-RUNX1 fusions to Eβ completely restored transcription through this cluster with only modest, at best, restoration of activating histone marks, H3K27ac and H3K4me3. This finding is consistent with several reports showing that transcriptional activation of inducible genes may be uncoupled from H3K4me3 or other histone modifications at promoters (50, 51). In a broader sense, *Tcrb*, like many developmentally regulated genes, may be transcriptionally induced before it acquires its final epigenetic state as a super-enhancer, triggering key aspects of lineage commitment without full revisions of the epigenome (52).

In contrast to changes in histone modifications, *Tcrb*-RC transcription completely parallels the restoration of Eβ-promoter looping and nucleosome clearance from both types of regulatory elements, which are likely to be contingent processes. These new findings support a proposed step-wise model for activation of the *Tcrb*-RC (31), in which promoter-Eβ contacts generate a stable "holocomplex" to recruit additional factors, including chromatin remodelers. Indeed, recruitment of the nucleosome remodeling complex, SWI/SNF, to the *Tcrb*-RC, requires promoter-Eβ looping (31). In this study, we demonstrated that RUNX1 is sufficient to mediate promoter-Eβ looping, chromatin remodeling, and transcriptional activation. The RUNX1-TAD was again sufficient to mediate these processes, when modest levels of ETS1 occupied Eβ. Thus, ETS1 may support modest levels of promoter-Eβ looping in the absence of the RUNX1-NRDB, perhaps via its known interaction with ATF2 (53), a transcription factor that binds to Dβ promoters (54). In the absence of ETS1, tethering of the NRDB and RUNX1-TAD to Eβ is sufficient to drive loop formation with its distal promoters. In this case, we propose that association of RUNX1 with the nuclear factor, ALY (55), may be important to forge Eβ-promoter contacts. Since ALY is known to dimerize (55), it may form a bridge between RUNX1 molecules bound to the enhancer and Dβ promoters (38). Moreover, once the bridged promoters become even minimally active, ALY may further stabilize Eβ-promoter loops via its RNA-binding activity (56).

Together, our results demonstrate that RUNX1, when uncoupled from the requirement for cooperative ETS1 binding, is sufficient to drive long-range loop formation by Eβ, nucleosome clearance at its target promoters, and full transcriptional activation of the *Tcrb*-RC. Given the

103

broad distribution of ETS-RUNX motifs in the mammalian genome, our findings with the *Tcrb-RC* may be applicable to the activation of super-enhancers at other key hematopoietic genes. Future studies with inducible experimental platforms, such as the one described here, will be useful to address more detailed mechanisms for long-range enhancer activation of promoters for developmentally regulated genes.
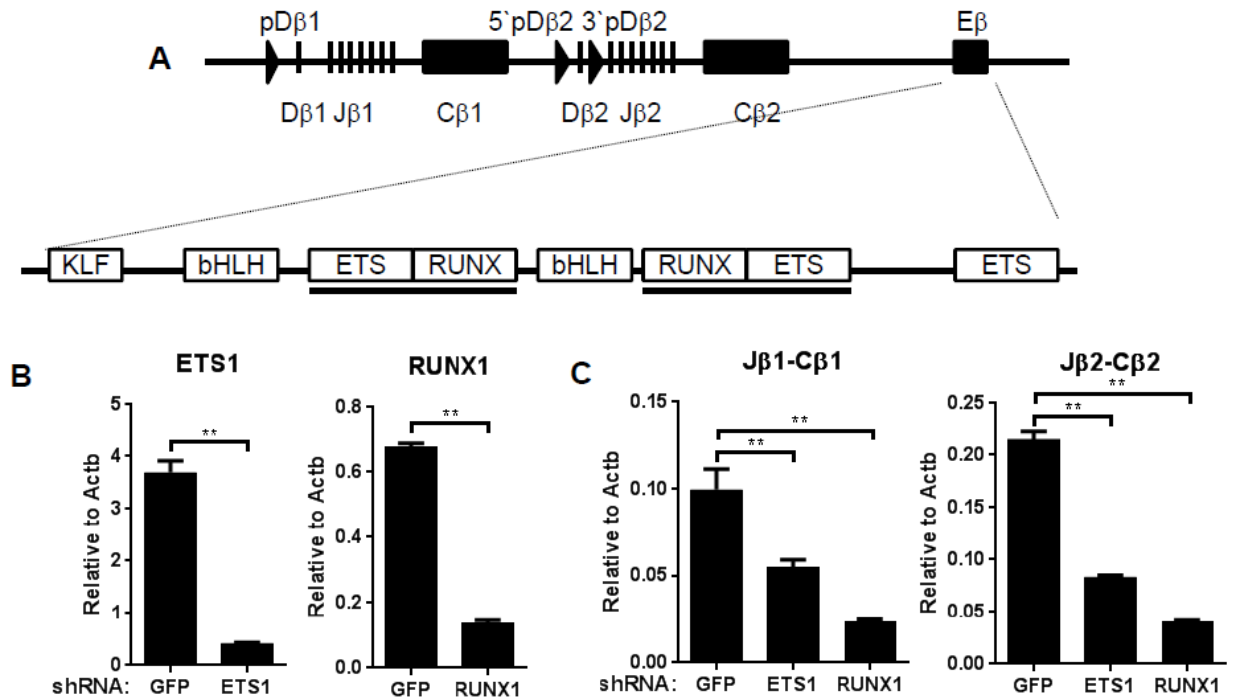
3.6 Figures

Figure 3.1 Eβ activity is dependent on ETS1 and RUNX1 in P5424 cells.

(A) Schematic representation of mouse *Tcrb*-RC showing promoters (pDβ1 and pDβ2), D elements (Dβ1 and Dβ2), J elements (Jβ1 and Jβ2 clusters), constant regions (Cβ1 and Cβ2) and the enhancer (Eβ). TF binding sites in Eβ identified by in vivo footprinting are highlighted below (33), including a pair of composite ETS-RUNX sites (black bars). **(B)** P5424 cells were transfected with shRNAs targeting ETS1, RUNX1, or control transcripts (GFP). Shown are ETS1 and RUNX1 mRNA levels measured by RT-qPCR and normalized to *Actb* transcripts. **(C)** Levels of spliced Jβ1-Cβ1 and Jβ2-Cβ2 germline transcripts in the indicated knock-downs as measured by RT-qPCR. In B and C, data are presented as the average values of three PCR replicates (±S.D.) and are derived from one of two biological replicates (see also Figs. 3.7 A and B), ** denotes statistical significance (p<0.01), as determined by Student's t-test.
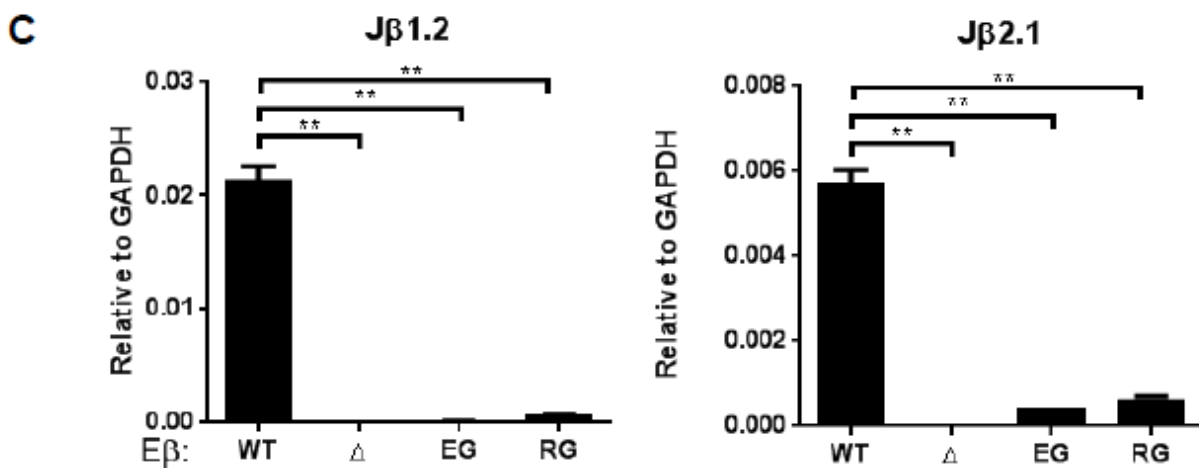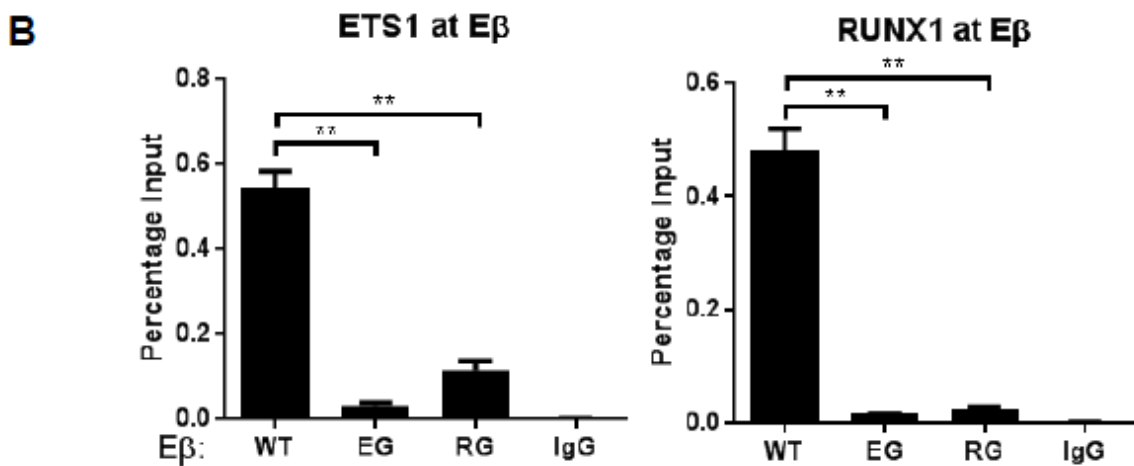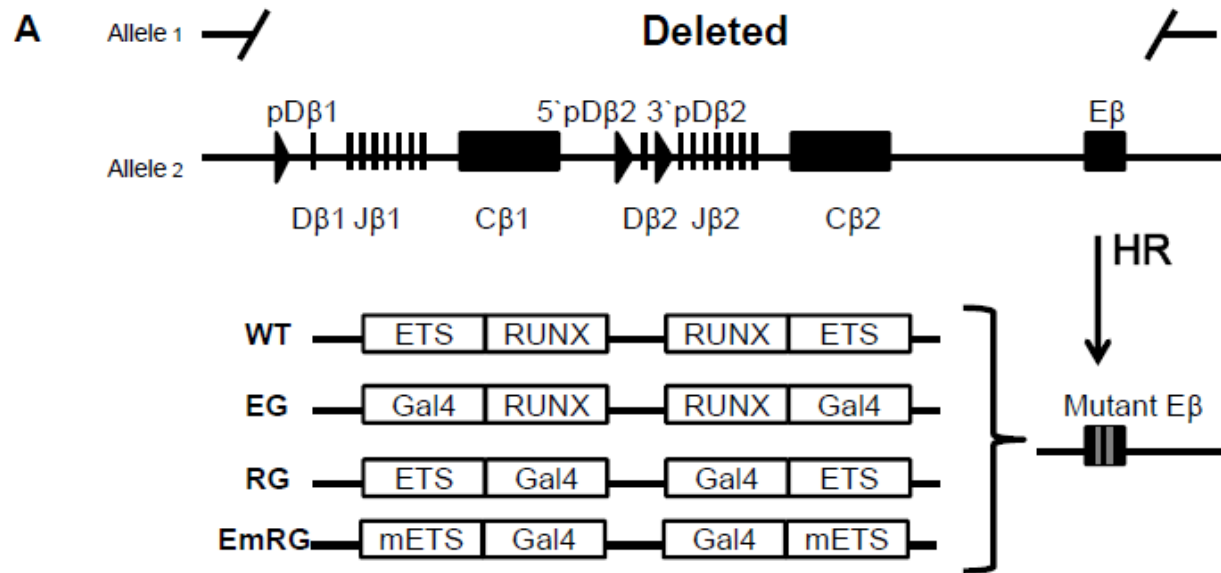
Figure 3.2 Mutation of ETS or RUNX motifs abolishes Eβ function.

(A) Schematic showing intact and mutant *Tcrb* alleles in O3-73 (top). As indicated, the entire RC, including Eβ, was deleted from the mutant allele. The intact allele was further targeted to generate a variety of mutant Eβ versions (bottom): EG (both ETS-RUNX sites converted to ETS1/Gal4), RG (ETS1/RUNX1 to Gal4/RUNX1 sites), EmRG (ETS1/RUNX1 to mutant ETS1/Gal4 sites). **(B)** ETS1 and RUNX1 binding at different versions of Eβ (WT, EG, and RG) measured by ChIP-qPCR. Control ChIPs with a non-specific IgG antibody are shown for WT cells. **(C)** Unspliced germline transcripts corresponding to Jβ1.2 and Jβ2.1 gene segments in cells with different versions of Eβ (WT, Δ, EG, and RG), as measured by RT-qPCR. For panels B and C, data are presented as the average values of three PCR replicates ($\pm$S.D.) and are derived one of two biological replicates (see also Figs. 3.7 C and D), ** denotes statistical significance ($p<0.01$), as determined by Student's t-test.

**A**

ETS1 — TAD | VII | ETS

RUNX1 — Runt | NRDB | TAD

RUNX1^GNT — Gal4DBD | NRDB | TAD

RUNX1^GN — Gal4DBD | NRDB

RUNX1^GT — Gal4DBD | TAD

B

Cell Fusion	WT	ΔEβ		RG	
	-	G4DBD	GNT	G4DBD	GNT

63kD
48kD
35kD
Gal4DBD 25kD
20kD
17kD
11kD

ETS1

GAPDH

C

ETS1 at Eβ

RUNX1 at Eβ

Eβ:	WT	RG	RG	IgG
TF:	-	G4	GNT	

D

Jβ1.2

Jβ2.1

Eβ:	WT	Δ	Δ	RG	RG
TF:	-	G4	GNT	G4	GNT
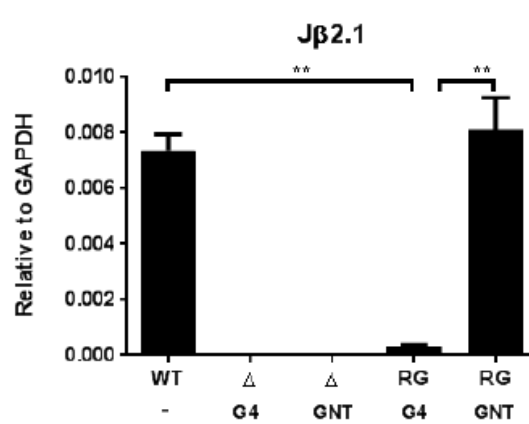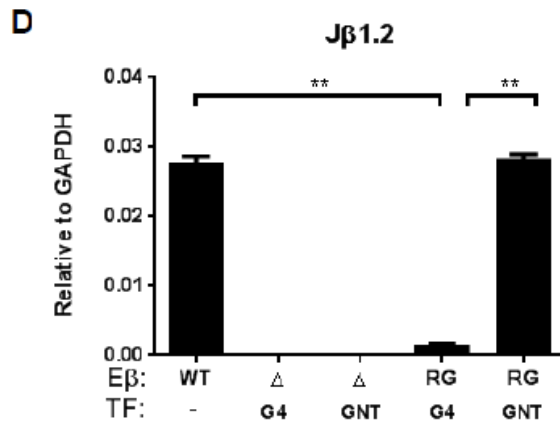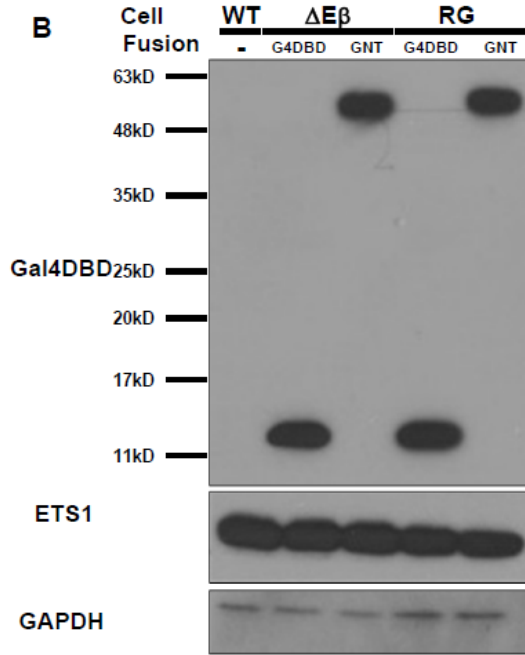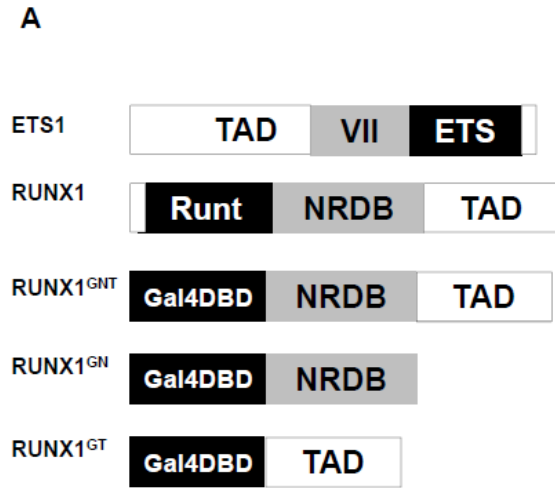
Figure 3.3 GAL4-RUNX tethering fully restores function to Eβ mutants lacking RUNX motifs.

A) Cartoon depictions of domains in full-length ETS1 and RUNX1 proteins, as well as in Gal4-RUNX1 fusions. ETS1 consists of an N terminus transactivation domain (TAD), exon VII domain (VII), and ETS DNA binding domain (ETS). RUNX1 consists of an N terminus Runt DNA binding domain (Runt), a negative regulatory region of DNA binding domain (NRDB), and a transactivation domain (TAD). **(B)** IP- and conventional western blotting analysis for expression of fusion proteins. Transduced cells were treated with 1 µg/mL Dox for 24 hours, and expression levels of Gal4 fusion proteins were detected by IP-western blotting. ETS1 and GAPDH expression were measured by conventional western blotting using the same cell lysates. **(C)** ChIP-qPCR assays for ETS1 and RUNX1 binding to the RG mutant enhancer upon expression of RUNX1^{GNT}. Control ChIPs with a non-specific IgG antibody are shown for WT cells. TF denotes Gal4DBD and Gal4-RUNX1 fusion proteins. **(D)** Unspliced germline transcripts corresponding to the Jβ1.2 and Jβ2.1 gene segments induced by RUNX1^{GNT} in cells with the RG mutant enhancer. For panels B and C, data are shown as averages ($\pm$ S.E.M.) from three biological replicates, ** denotes statistical significance (p<0.01), as determined by Student's t-test.
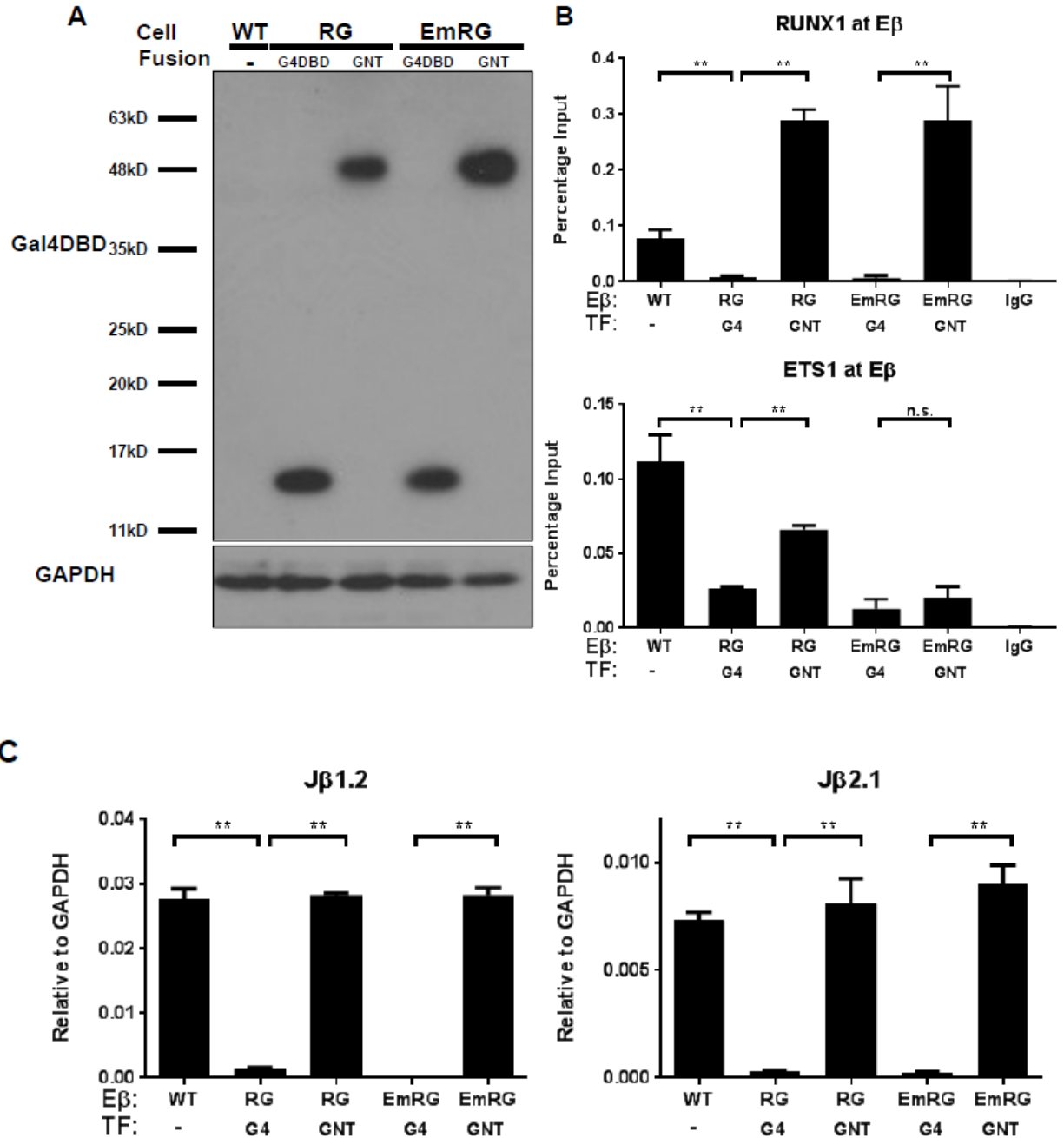
Figure 3.4 ETS1 is dispensable for Eβ activation by RUNX1.
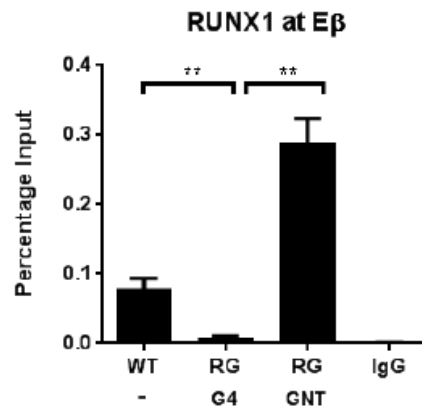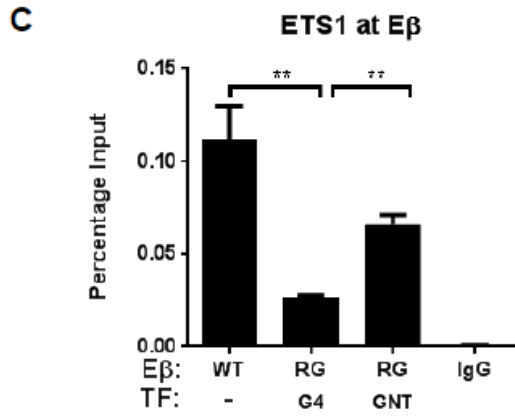
(**A**) Western blotting analysis for expression of Gal4-RUNX1 fusion proteins. Transduced cells were treated with 1 µg/mL Dox for 24 hours, and expression levels of Gal4 fusions or GAPDH were detected by IP- or conventional western blotting, respectively. (**B**) ChIP-qPCR assays for ETS1 and RUNX1 binding to the RG and EmRG mutant enhancers upon expression of RUNX1[GNT], as described in Fig. 3.3C. (**C**) Unspliced germline transcripts for Jβ1.2 and Jβ2.1 segments induced by RUNX1[GNT] in cells with RG and EmRG mutant enhancers, as described in

110

Fig. 3.3D. All cells for ChIP- and RT-qPCR assays were cultured in parallel with those shown in Fig. 3.3, the data from which are included again for direct comparisons.
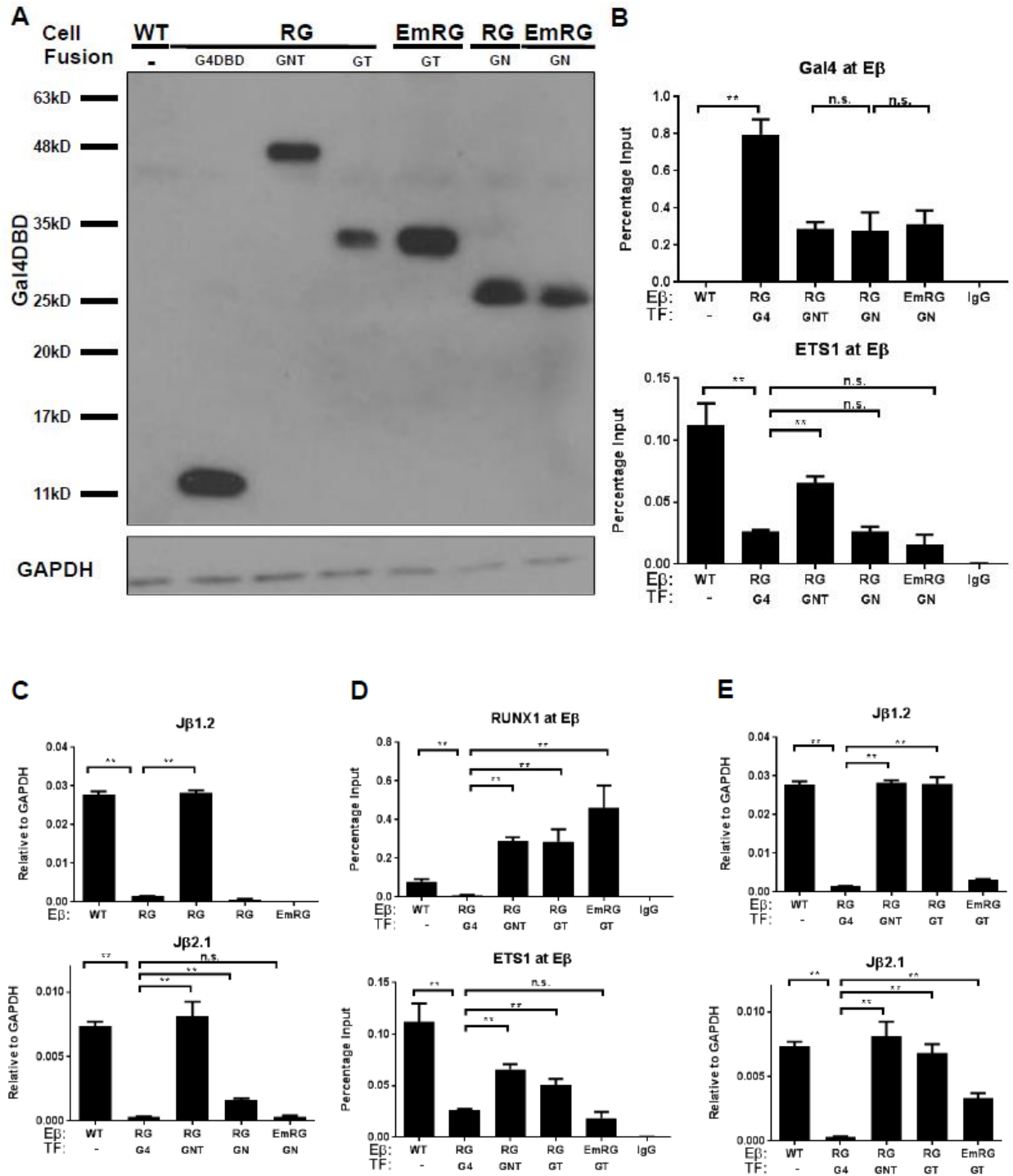
Figure 3.5 The TAD of RUNX1 is sufficient for short- but not long-range activation of the *Tcrb*-RC.

(A) Western blotting analysis for expression of indicated Gal4-RUNX1 fusion proteins, as described in Fig. 3.4A. **(B)** ChIP-qPCR assays for ETS1 and Gal4DBD binding to the RG and

EmRG mutant enhancers upon expression of RUNX1^{GN}, as described in Fig. 3.3C. **(C)** Unspliced germline transcripts for Jβ1.2 and Jβ2.1 segments induced by RUNX1^{GN} in cells with RG and EmRG mutant enhancers (see Fig. 3.3D). **(D)** ChIP-qPCR assays for ETS1 and Gal4DBD binding to the RG and EmRG mutant enhancers upon expression of RUNX1^{GT}. **(E)** Unspliced germline transcripts corresponding to Jβ1.2 and Jβ2.1 segments induced by RUNX1^{GT} in cells with RG and EmRG mutant enhancers. For panels B-E, data are shown averages ($\pm$ S.E.M.) of three biological replicates, n.s. (non-significant), ** ($p<0.01$), as determined by Student's t-test. All cells for ChIP- and RT-qPCR assays were cultured in parallel with those shown in Figs. 3.3&3.4, the data from which are included again for direct comparisons.

A

H3K27Ac

| | Jβ1 | Jβ2 |

TF: - Gal4 GNT GT Gal4 GNT GT - Gal4 GNT GT Gal4 GNT GT
Eβ: WT | RG | EmRG | WT | RG | EmRG

H3K4Me3

| | Jβ1 | Jβ2 |

TF: - Gal4 GNT GT Gal4 GNT GT - Gal4 GNT GT Gal4 GNT GT
Eβ: WT | RG | EmRG | WT | RG | EmRG

B

H3 at pDβ1

TF: - Gal4 GNT GT GN Gal4 GNT GT GN - -
Eβ: WT | RG | EmRG | Δ 3T3

H3 at pDβ2

TF: - Gal4 GNT GT GN Gal4 GNT GT GN - -
Eβ: WT | RG | EmRG | Δ 3T3

H3 at 5`Eβ

TF: - Gal4 GNT GT GN Gal4 GNT GT GN - -
Eβ: WT | RG | EmRG | Δ 3T3

C

pDβ1   5`pDβ2  3`pDβ2   Eβ

Dβ1 Jβ1   Cβ1   Dβ2 Jβ2   Cβ2

Dβ1-Eβ

Eβ: WT RG RG RG EmRG EmRG
TF: -   G4 GNT GT GNT   GT

Dβ2-Eβ

Eβ: WT RG RG RG EmRG EmRG
TF: -   G4 GNT GT GNT   GT

114

Figure 3.6 RUNX1 is sufficient to restore aspects of the *Tcrb*-RC epigenetic landscape.
(A) ChIP-qPCR assays for H3K27ac and H3K4me3 levels at Jβ1 or Jβ2 following induction of the indicated Gal4-RUNX1 fusions in clones containing either the RG or EmRG mutant versions of Eβ. **(B)** Total histone H3 ChIPs showing changes in nucleosome occupancy induced by Gal4-RUNX1 fusions. H3 levels were monitored at pDβ1, pDβ2, and immediately upstream of Eβ (5`Eβ) in RG and EmRG mutant cells expressing the indicated RUNX1 fusion proteins. Clones containing wild-type (WT) or ΔEβ versions of *Tcrb*, as well as 3T3 fibroblasts are included as controls. Values for IgG control ChIPs in both A and B were significantly lower than all samples shown. **(C)** Promoter-enhancer interactions induced by Gal4-RUNX1 fusions at RG and EmRG versions of Eβ, as measured by 3C-qPCR (42). Data are presented as the average values of three PCR replicates (±S.D.) and are derived one of two biological replicates (see also Fig. 3.8). For panel C, n.s. (non-significant), ** (p<0.01), as determined by Student's t-test.
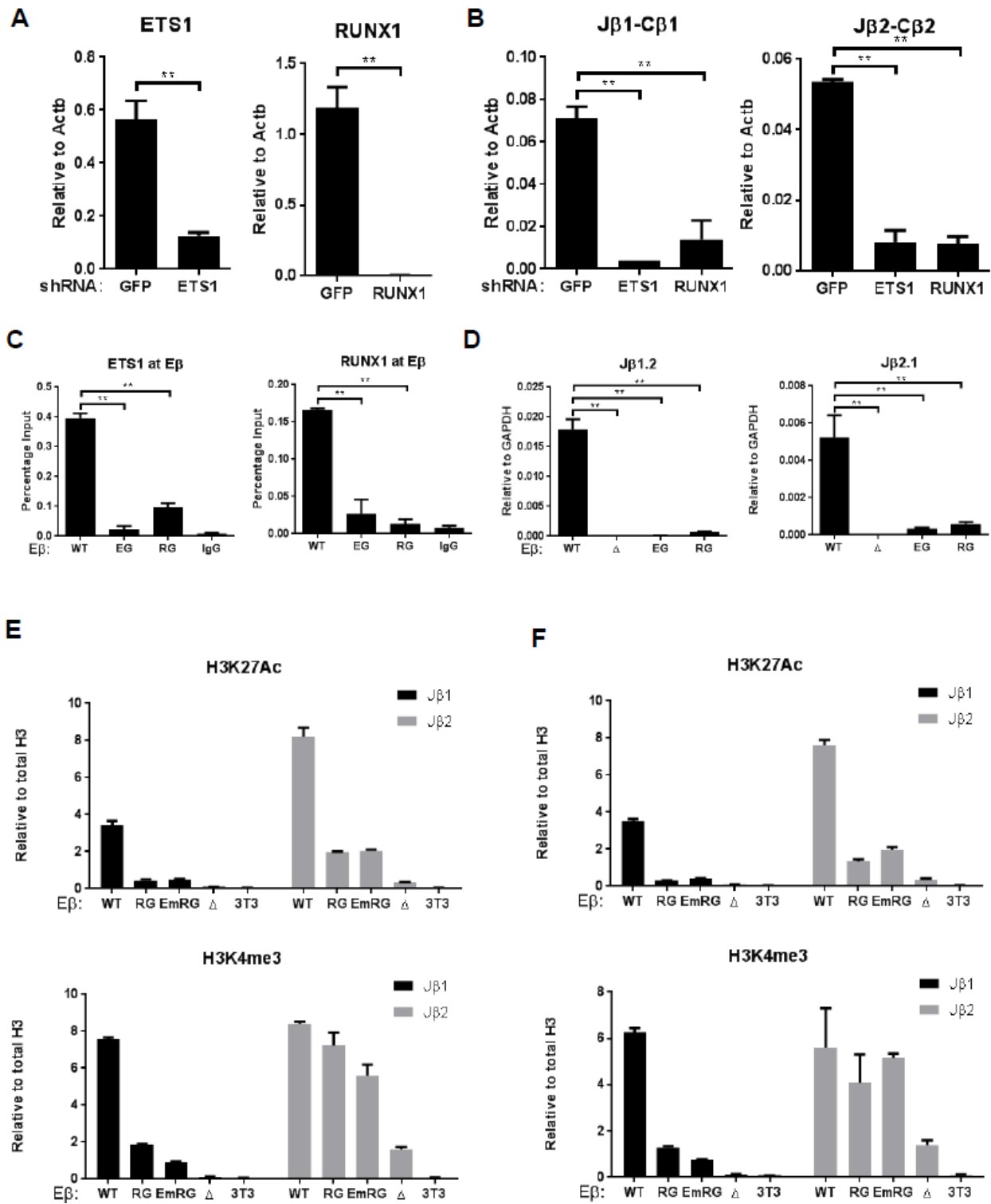
Figure 3.7 Biological replicates for Figs. 3.1&3.2 and basal H3K27ac and H3K4me3 levels at Jβ1 or Jβ2

(A) and **(B)** show data from the biological replicate corresponding to experiments presented in Figs. 3.1 B and C, respectively. **(C)** and **(D)** show data from the biological replicate corresponding to experiments presented in Figs. 3.2 B and C, respectively. **(E)** ChIP-qPCR assays for basal H3K27ac and H3K4me3 levels at Jβ1 or Jβ2 of the indicated cells (see Fig. 6A for details). **(F)** Biological replicate for experiments shown in panel E.
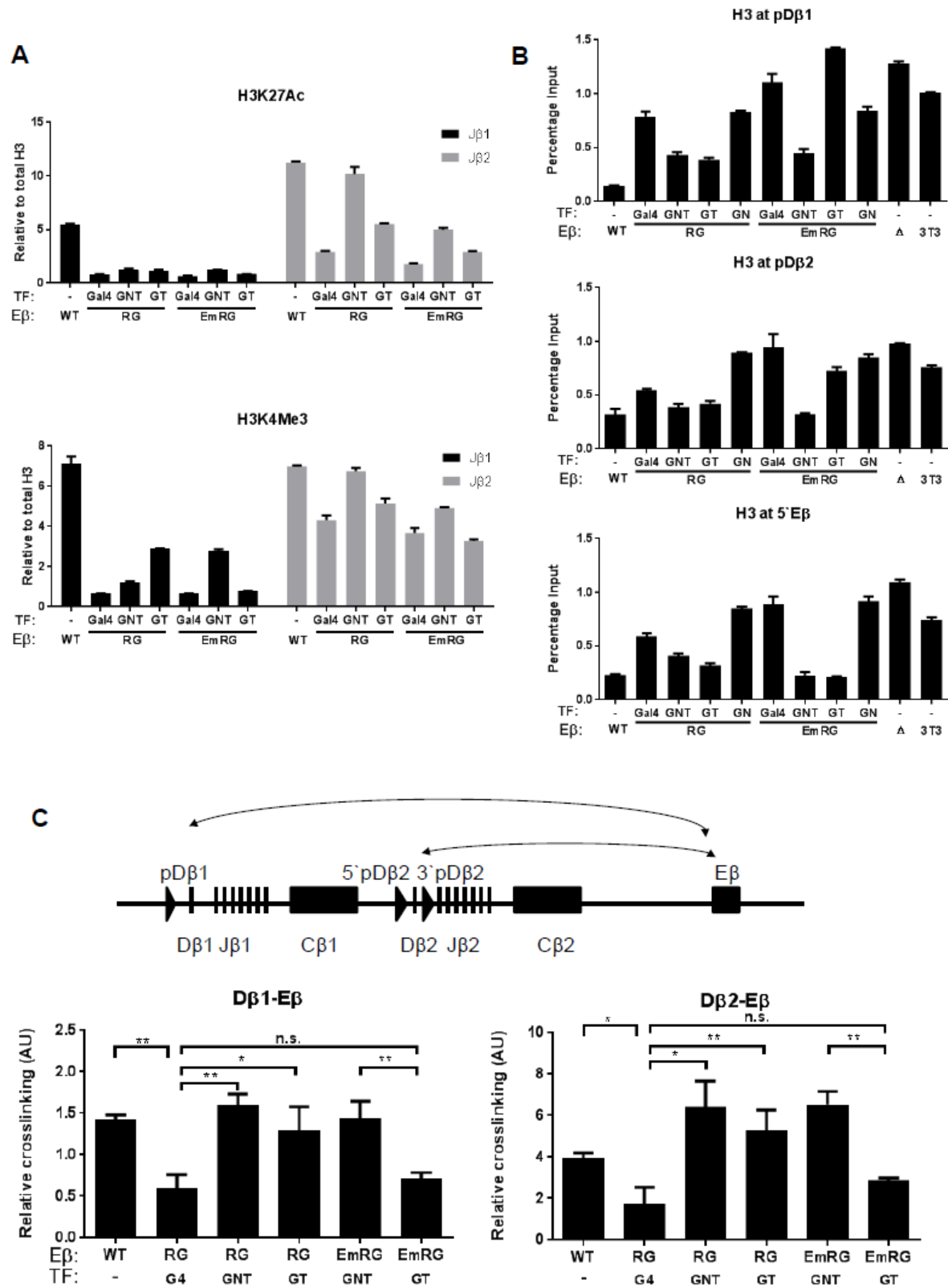
Figure 3.8 Biological replicates for Fig. 3.6

A, B and C show data from the biological replicates corresponding to experiments presented in Fig.6 A, B and C, respectively.

3.7   Tables

Table 3.1 RT-qPCR, ChIP-qPCR, and 3C assay primers and probes

Jb1 F	GAACCAGACTCACAGTTGTAGAGG
Cb1 R	GCTCTCCTTGTAGGCCTGAG
Jb2 F	ACGACTCACCGTCCTAGAGG
Cb2 R	CATTCACCCACCAGCTCAG
ETS1 F	ACAGCTTTGTTGTCCATCTG
ETS1 R	AGATCTGTCCATCCTTCCTG
Actb F	GGCTGTATTCCCCTCCATCG
Actb R	CCAGTTGGTAACAATGCCATGT
Runx F	GGCACTCTGGTCACCGTCAT
Runx R	CGTTGAATCTCGCTACCTGGTT
GAPDHF	GCCATCAACGACCCCTTCATTGACC
GAPDHR	GGTCTCGCTCCTGGAAGATGGTGAT
Jb1.2F	ATTGGTACCGTTAACCAGGCACAGTAGGACC
Jb1.2R	TTCCACCCGAGGTTCATCTTGACT
Jb2.1F	TATGCTGAGCAGTTCTTCGGACCA
Jb2.1R	AGTCCTGGAAATGCTGGCACAAAC
EbFChIP	TTGACATTTACCAGGTCCTACA
EbRChIP	GTGGGGGAGCATTCACAGGCAC
Jb1upChIP	ATGCAGGCAAGCTCAACTGGGAGTA
Jb1downChIP	TACATCTTCCCTGCCCAAGCCTGTC
Jb2FChIP	AGAAACGCTGTATTTTGGCTCAGGA

Jb2RChIP	CTGGCTATCCTTAGCCTGGGAAATG
pDb1FChIP	GCCCCTTCAGCAAAGATCATTTCAA
pDb1RChIP	TTGTGCAAGGTGGTGGTAAGATG
pDb2FChIP	TAGATGGATATCCGTTCCCAAGCCA
pDb2RChIP	TCTACCTCCAGAACTGGTAAGGCAA
5`EbFChIP	TGTAGCACTGACAGAATGGTAAGT
5`EbRChIP	CAGCACTGTCCCTTTGCAGGTCA
ERCC3F3C	GCCCTCCCTGAAAATAAGGA
ERCC3R3C	GACTTCTCACCTGGGCCTACA
ERCC3probe3C	AAAGCTTGCACCCTGCTTTAGTGGCC
EbF3C	TGGGACTAGTGAACTGAATATGTCTTC
Db1R3C	AAGGCATTGTTGCATGATCC
Db2R3C	TGGGGCCCTCACTTTTCTTA
Ebprobe3C	GGAGAGGCAGAGTGGTAGGAA

Note: Annealing at 60 ℃

Table 3.2 Mutagenesis and cloning primers

ZFN1upF	ATCCCCAGCCAGATTCTTCT
ZFN2downR	CTCATGTGCTGTCTGCTCTGAATTC
ETS1Fmut	GAGATGGAGCTATGGCAGG
ETS1Rmut	CGGAGTACTGTCCTCCGTGTGGTTTGACATTTACCAGG
ETS2Fmut	CGGAGGACAGTACTCCGTGTGGCAAGTGTGGTTCC
ETS2Rmut	TCTCATGCAGCCCTTTCTAC
RUNX1Fmut	ACATCCTGTTGTGACTCCTG
RUNX1Rmut	CGGAGTACTGTCCTCCGTTGACATTTACCAGGTCCTACA
RUNX23Fmut	TCCCAAAATGCTCAGAATGTG
RUNX23Rmut	CGGAGTACTGTCCTCCGACATCCTGTCTTCAAACCCT
mutETS1Fmut	ACAGAAGGTTGTGACTCCTGGAGATGG
RUNX23Rmut.2	CGGAGGACAGTACTCCGCCCAAAATGCTCAGAATGTGA
mutETS2Rmut	ACAGAAGGTCTTCAAACCCTTCTCATGC
5`armF	GCCAAGAAGATACTGGAGAATTATG
5`armR	AATCGGTCAGGGAAACAAGTATTAA
EbFclone	ATTGGTACCGTTAACCAGGCACAGTAGGACC
EbRclone	AATACGCGTGCAGGACAGACAGACCATCT
PuroF	AATACGCGTCCGATCATATTCAATAACCCTTAAT
PuroR	AATGCTAGCGAACCTCTTCGAGGGACCTA
HygroF	CCGGAGATGAGGAAGAGGAGAACA
HygroR	TCAGCAGCCTCTGTTCCACATACA
3`armF	AATGCTAGCATTCATTGGCTGGCTTTCGCT

3`armR	ATTGCGGCCGCAGGAATTTTTACTGCTGCTTGCGTC
ZFN2cutF	AGCCCTTTCTACCTCAGCCTCTA
ZFN2cutR	ACTGGGGCCCCTTGACTGCT
ETSwtGeno	ACATCCTGTTGTGACTCCTG
ETSmutGenoR	TGGTAAATGTCAAACCACACGGA
RUNXmutGenoR	TAGGACCTGGTAAATGTCAACGGA
WTRXGenoR	TGTAGGACCTGGTAAATGTCAAACC
mERXmutGenoR	AGGACAGTACTCCGACAGAAG
Gal4DBDF	ATTCTTAAGATGAAGCTACTGTCT
Gal4DBDR	ATTGATATCAGAACCACCTGAACCTCCTGATCCACCCGATACAGTCAACTG
RUNX1.184F	AATGATATCGATGATCAGACCAAGCCCGGGA
RUNX1.451R	ATTGCGGCCGCTCAGTAGGGCCGCCA
RUNX1.292R	ATTGCGGCCGCTCAACTGGAAAGTTCTGCAGAGAG
RUNX1.291F	AATGATATCTCCAGTCGACTCTCAACG

Note: Annealing at 62 °C with Phusion. mutETS1F pairs with RUNX1R to make upstream compsite site of EmRG.. 5`EbChIP in Table S1 and ZFN2cutR in this table are used to detect genomic Eβ deletion.

3.8   References

1. Rothenberg, E. V, A. Champhekar, S. Damle, M. M. Del Real, H. Y. Kueh, L. Li, and M. A. Yui. 2013. Transcriptional establishment of cell-type identity: dynamics and causal mechanisms of T-cell lineage commitment. *Cold Spring Harb. Symp. Quant. Biol.* 78: 31–41.

2. Whyte, W. A., D. A. Orlando, D. Hnisz, B. J. Abraham, C. Y. Lin, M. H. Kagey, P. B. Rahl, T. I. Lee, and R. A. Young. 2013. Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. *Cell* 153: 307–319.

3. Vanhille, L., A. Griffon, M. Maqbool, J. Zacarias-Cabeza, L. Dao, N. Fernandez, B. Ballester, J. Andrau, and S. Spicuglia. 2015. High-throughput and quantitative assessment of enhancer activity in mammals by CapStarr-seq. *Nat. Commun.* 6: 6905.

4. Tripathi, R. K., N. Mathieu, S. Spicuglia, D. Payet, C. Verthuy, G. Bouvier, D. Depetris, M. G. Mattei, W. M. HempeL, and P. Ferrier. 2000. Definition of a T-cell receptor beta gene core enhancer of V(D)J recombination by transgenic mapping. *Mol. Cell. Biol.* 20: 42–53.

5. Sikes, M. L., R. J. Gomez, J. Song, and E. M. Oltz. 1998. A developmental stage-specific promoter directs germline transcription of D beta J beta gene segments in precursor T lymphocytes. *J. Immunol.* 161: 1399–405.

6. McMillan, R. E., and M. L. Sikes. 2008. Differential activation of dual promoters alters Dbeta2 germline transcription during thymocyte development. *J. Immunol.* 180: 3218–28.

7. Majumder, K., O. I. Koues, E. A. W. Chan, K. E. Kyle, J. E. Horowitz, K. Yang-Iott, C. H. Bassing, I. Taniuchi, M. S. Krangel, and E. M. Oltz. 2015. Lineage-specific compaction of Tcrb

requires a chromatin barrier to protect the function of a long-range tethering element. *J. Exp. Med.* 212.

8. Mathieu, N., W. M. Hempel, S. Spicuglia, C. Verthuy, and P. Ferrier. 2000. Chromatin Remodeling by the T Cell Receptor (Tcr)-β Gene Enhancer during Early T Cell Development. *J. Exp. Med.* 192.

9. Schatz, D. G., and Y. Ji. 2011. Recombination centres and the orchestration of V(D)J recombination. *Nat. Rev. Immunol.* 11: 251–263.

10. Ji, Y., A. J. Little, J. K. Banerjee, B. Hao, E. M. Oltz, M. S. Krangel, and D. G. Schatz. 2010. Promoters, enhancers, and transcription target RAG1 binding during V(D)J recombination. *J. Exp. Med.* 207.

11. Hollenhorst, P. C., K. J. Chandler, R. L. Poulsen, W. E. Johnson, N. A. Speck, and B. J. Graves. 2009. DNA Specificity Determinants Associate with Distinct Transcription Factor Functions. *PLoS Genet.* 5: e1000778.

12. Erman, B., M. Cortes, B. S. Nikolajczyk, N. A. Speck, and R. Sen. 1998. ETS-core binding factor: a common composite motif in antigen receptor gene enhancers. *Mol. Cell. Biol.* 18: 1322–30.

13. Peeters, J. G. C., S. J. Vervoort, S. C. Tan, G. Mijnheer, S. de Roock, S. J. Vastert, E. E. S. Nieuwenhuis, F. van Wijk, B. J. Prakken, M. P. Creyghton, P. J. Coffer, M. Mokry, and J. van Loosdregt. 2015. *Inhibition of Super-Enhancer Activity in Autoinflammatory Site-Derived T Cells Reduces Disease-Associated Gene Expression,*. *Cell reports* 12.12 (2015): 1986-1996.

14. Chuang, L. S. H., K. Ito, and Y. Ito. 2013. RUNX family: Regulation and diversification of roles through interacting proteins. *Int. J. Cancer* 132: 1260–1271.

15. Hollenhorst, P. C., L. P. McIntosh, and B. J. Graves. 2011. Genomic and Biochemical Insights into the Specificity of ETS Transcription Factors. *Annu. Rev. Biochem.* 80: 437–471.

16. Petrovick, M. S., S. W. Hiebert, A. D. Friedman, C. J. Hetherington, D. G. Tenen, and D. E. Zhang. 1998. Multiple functional domains of AML1: PU.1 and C/EBPalpha synergize with different regions of AML1. *Mol. Cell. Biol.* 18: 3915–25.

17. Wai, P. Y., Z. Mi, C. Gao, H. Guo, C. Marroquin, and P. C. Kuo. 2006. Ets-1 and runx2 regulate transcription of a metastatic gene, osteopontin, in murine colorectal cancer cells. *J. Biol. Chem.* 281: 18973–82.

18. Gu, T. L., T. L. Goetz, B. J. Graves, and N. A. Speck. 2000. Auto-inhibition and partner proteins, core-binding factor beta (CBFbeta) and Ets-1, modulate DNA binding by CBFalpha2 (AML1). *Mol. Cell. Biol.* 20: 91–103.

19. Goetz, T. L., T. L. Gu, N. A. Speck, and B. J. Graves. 2000. Auto-inhibition of Ets-1 is counteracted by DNA binding cooperativity with core-binding factor alpha2. *Mol. Cell. Biol.* 20: 81–90.

20. Kim, W. Y., M. Sieweke, E. Ogawa, H. J. Wee, U. Englmeier, T. Graf, and Y. Ito. 1999. Mutual activation of Ets-1 and AML1 DNA binding by direct interaction of their autoinhibitory domains. *EMBO J.* 18: 1609–20.

21. Huang, G., K. Shigesada, K. Ito, H. J. Wee, T. Yokomizo, and Y. Ito. 2001. Dimerization with PEBP2beta protects RUNX1/AML1 from ubiquitin-proteasome-mediated degradation. *EMBO J.* 20: 723–33.

22. Yang, C., L. H. Shapiro, M. Rivera, A. Kumar, and P. K. Brindle. 1998. A role for CREB binding protein and p300 transcriptional coactivators in Ets-1 transactivation functions. *Mol. Cell. Biol.* 18: 2218–29.

23. Kitabayashi, I., A. Yokoyama, K. Shimizu, and M. Ohki. 1998. Interaction and functional cooperation of the leukemia-associated factors AML1 and p300 in myeloid cell differentiation. *EMBO J.* 17: 2994–3004.

24. Rothenberg, E. V., J. E. Moore, and M. A. Yui. 2008. Launching the T-cell-lineage developmental programme. *Nat. Rev. Immunol.* 8: 9–21.

25. Weber, B. N., A. W.-S. Chi, A. Chavez, Y. Yashiro-Ohtani, Q. Yang, O. Shestova, and A. Bhandoola. 2011. A critical role for TCF-1 in T-lineage specification and differentiation. *Nature* 476: 63–8.

26. Cauchy, P., M. A. Maqbool, J. Zacarias-Cabeza, L. Vanhille, F. Koch, R. Fenouil, M. Gut, I. Gut, M. A. Santana, A. Griffon, J. Imbert, C. Moraes-Cabé, J.-C. Bories, P. Ferrier, S. Spicuglia, and J.-C. Andrau. 2016. Dynamic recruitment of Ets1 to both nucleosome-occupied and -depleted enhancer regions mediates a transcriptional program switch during early T-cell differentiation. *Nucleic Acids Res.* 44: 3567–85.

27. Bories, J.-C., D. M. Willerford, D. Grévin, L. Davidson, A. Camus, P. Martin, D. Stéhelin, and F. W. Alt. 1995. Increased T-cell apoptosis and terminal B-cell differentiation induced by inactivation of the Ets-1 proto-oncogene. *Nature* 377: 635–638.

28. Okuda, T., J. van Deursen, S. W. Hiebert, G. Grosveld, and J. R. Downing. 1996. AML1, the Target of Multiple Chromosomal Translocations in Human Leukemia, Is Essential for Normal Fetal Liver Hematopoiesis. *Cell* 84: 321–330.

29. Sun, W., B. J. Graves, and N. A. Speck. 1995. Transactivation of the Moloney murine leukemia virus and T-cell receptor beta-chain enhancers by cbf and ets requires intact binding sites for both proteins. *J. Virol.* 69: 4941–9.

30. Bouvier, G., F. Watrin, M. Naspetti, C. Verthuy, P. Naquet, and P. Ferrier. 1996. Deletion of the mouse T-cell receptor beta gene enhancer blocks alphabeta T-cell development. *Proc. Natl. Acad. Sci. U. S. A.* 93: 7877–81.

31. Oestreich, K. J., R. M. Cobb, S. Pierce, J. Chen, P. Ferrier, and E. M. Oltz. 2006. Regulation of TCRβ Gene Assembly by a Promoter/Enhancer Holocomplex. *Immunity* 24: 381–391.

32. Osipovich, O., R. Milley Cobb, K. J. Oestreich, S. Pierce, P. Ferrier, and E. M. Oltz. 2007. Essential function for SWI-SNF chromatin-remodeling complexes in the promoter-directed assembly of Tcrb genes. *Nat. Immunol.* 8: 809–816.

33. Bonnet, M., F. Huang, T. Benoukraf, O. Cabaud, C. Verthuy, A. Boucher, S. Jaeger, P. Ferrier, and S. Spicuglia. 2009. Duality of Enhancer Functioning Mode Revealed in a Reduced TCR Gene Enhancer Knockin Mouse Model. *J. Immunol.* 183: 7939–7948.

34. Schmitt, T. M., R. F. de Pooter, M. A. Gronski, S. K. Cho, P. S. Ohashi, and J. C. Zúñiga-Pflücker. 2004. Induction of T cell development and establishment of T cell competence from embryonic stem cells differentiated in vitro. *Nat. Immunol.* 5: 410–417.

35. Dutta, D., S. Ray, J. L. Vivian, and S. Paul. 2008. Activation of the VEGFR1 chromatin domain: an angiogenic signal-ETS1/HIF-2alpha regulatory axis. *J. Biol. Chem.* 283: 25404–13.

36. Yang, L., W. Yu, J. DU, S. Chao, M. Chen, H. Zhao, and J. Guo. 2009. Overexpression or knock-down of runt-related transcription factor 1 affects BCR-ABL-induced proliferation and migration in vitro and leukemogenesis in vivo in mice. *Chin. Med. J. (Engl).* 122: 331–7.

37. David, R., M. Groebner, and W.-M. Franz. 2005. Magnetic Cell Sorting Purification of Differentiated Embryonic Stem Cells Stably Expressing Truncated Human CD4 as Surface Marker. *Stem Cells* 23: 477–482.

38. McMillan, R. E., and M. L. Sikes. 2009. Promoter activity 5′ of Dβ2 is coordinated by E47, Runx1, and GATA-3. *Mol. Immunol.* 46: 3009–3017.

39. Helmink, B. A., A. T. Tubbs, Y. Dorsett, J. J. Bednarski, L. M. Walker, Z. Feng, G. G. Sharma, P. J. McKinnon, J. Zhang, C. H. Bassing, and B. P. Sleckman. 2011. H2AX prevents CtIP-mediated DNA end resection and aberrant repair in G1-phase lymphocytes. *Nature* 469: 245–249.

40. Gopalakrishnan, S., K. Majumder, A. Predeus, Y. Huang, O. I. Koues, J. Verma-Gaur, S. Loguercio, A. I. Su, A. J. Feeney, M. N. Artyomov, and E. M. Oltz. 2013. Unifying model for molecular determinants of the preselection Vβ repertoire. *Proc. Natl. Acad. Sci. U. S. A.* 110: E3206-15.

41. Egawa, T., R. E. Tillman, Y. Naoe, I. Taniuchi, and D. R. Littman. 2007. The role of the Runx transcription factors in thymocyte differentiation and in homeostasis of naive T cells. *J. Exp. Med.* 204.

42. Hagège, H., P. Klous, C. Braem, E. Splinter, J. Dekker, G. Cathala, W. de Laat, and T. Forné. 2007. Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nat. Protoc.* 2: 1722–1733.

43. Majumder, K., L. J. Rupp, K. S. Yang-Iott, O. I. Koues, K. E. Kyle, C. H. Bassing, and E. M. Oltz. 2015. Domain-Specific and Stage-Intrinsic Changes in Tcrb Conformation during Thymocyte Development. *J. Immunol.* 195.

44. Mombaerts, P., C. Terhorst, T. Jacks, S. Tonegawa, and J. Sancho. 1995. Characterization of immature thymocyte lines derived from T-cell receptor or recombination activating gene 1 and p53 double mutant mice. *Proc. Natl. Acad. Sci. U. S. A.* 92: 7420–4.

45. Hatzis, P., and I. Talianidis. 2002. Dynamics of Enhancer-Promoter Communication during Differentiation-Induced Gene Activation. *Mol. Cell* 10: 1467–1477.

46. Ong, C.-T., and V. G. Corces. 2011. Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nat. Rev. Genet.* 12: 283–293.

47. Ho, Y., F. Elefant, N. Cooke, and S. Liebhaber. 2002. A Defined Locus Control Region Determinant Links Chromatin Domain Acetylation with Long-Range Gene Activation. *Mol. Cell* 9: 291–302.

48. Bulger, M., and M. Groudine. 2011. Functional and Mechanistic Diversity of Distal Transcription Enhancers. *Cell* 144: 327–339.

49. Peterson, C. L., and J. L. Workman. 2000. Promoter targeting and chromatin remodeling by the SWI/SNF complex. *Curr. Opin. Genet. Dev.* 10: 187–192.

50. Kiefer, H. L. B., T. M. Hanley, J. E. Marcello, A. G. Karthik, and G. A. Viglianti. 2004. Retinoic acid inhibition of chromatin remodeling at the human immunodeficiency virus type 1 promoter. Uncoupling of histone acetylation and chromatin remodeling. *J. Biol. Chem.* 279: 43604–13.

51. Zhang, H., L. Gao, J. Anandhakumar, and D. S. Gross. 2014. Uncoupling Transcription from Covalent Histone Modification. *PLoS Genet.* 10: e1004202.

52. Pérez-Lluch, S., E. Blanco, H. Tilgner, J. Curado, M. Ruiz-Romero, M. Corominas, and R. Guigó. 2015. Absence of canonical marks of active chromatin in developmentally regulated genes. *Nat. Genet.* 47: 1158–1167.

53. Giese, K., C. Kingsley, J. R. Kirshner, and R. Grosschedl. 1995. Assembly and function of a TCR alpha enhancer complex is dependent on LEF-1-induced DNA bending and multiple protein-protein interactions. *Genes Dev.* 9: 995–1008.

54. Spicuglia, S., S. Kumar, J.-H. Yeh, E. Vachez, L. Chasson, S. Gorbatch, J. Cautres, and P. Ferrier. 2002. Promoter activation by enhancer-dependent and -independent loading of activator and coactivator complexes. *Mol. Cell* 10: 1479–87.

55. Bruhn, L., A. Munnerlyn, and R. Grosschedl. 1997. ALY, a context-dependent coactivator of LEF-1 and AML-1, is required for TCRalpha enhancer function. *Genes Dev.* 11: 640–53.

56. Zhou, Z., M. Luo, K. Straesser, J. Katahira, E. Hurt, and R. Reed. 2000. The protein Aly links pre-messenger-RNA splicing to nuclear export inmetazoans. *Nature* 407: 401–405.

Chapter 4 : Conclusions and Future Directions

In my thesis, I first investigated ETS1-RUNX1 cooperativity at the developmentally important super enhancer *Tcrb*-RC using extrachromosomal reporter substrates. My results demonstrated independent competency of both ETS1 and RUNX1 proteins on Eβ in these substrates. Next, I constructed cell models with knock-in mutant versions of Eβ to examine ETS1 and RUNX1 functions in chromosomal settings. In contrast to results from extrachromosomal substrates, RUNX1, but not ETS1, was capable of activating chromosomal mutant Eβ. To further delineate the mechanisms by which RUNX1 activates the *Tcrb*-RC, I constructed an inducible cell system and dissected RUNX1 domain functions with or without the presence of ETS1. The results established the sufficiency of RUNX1 in Eβ activation, which is independent from ETS1. RUNX1 is able to tolerate loss of ETS1 activity by using its NRDB domain in long-distance promoter activation by Eβ, while the RUNX1 TAD by itself maintains basal activity to proximal cis-elements. In addition to analysis of RUNX1 domains in *Tcrb*-RC activation, my research also revealed the spatial and temporal heterogeneity of different genetic and epigenetic processes during locus activation. Specifically, simultaneous transcription factor loading to cis-elements across the whole *Tcrb*-RC precedes maximal induction of Eβ activity on both distal and proximal promoters. However, such activation, especially at the distal promoter, is uncoupled from full restoration of nucleosome depletion and active histone marks.

These results not only provided insights into mechanisms underlying Eβ function and *Tcrb*-RC activation, but also offered several starting points for future studies: (i) further investigation of ETS1 function at Eβ with recruited Gal4-ETS1 fusion proteins, (ii) the temporal and spatial analysis of *Tcrb*-RC activation by high-throughput approaches, (iii) V(D)J recombination and Vβ usage with mouse models, and (iv) cooperativity of similar transcription factors on other loci. These proposed studies will deepen our understanding above and beyond the *Tcrb* gene regulation.

Function of ETS1 in Eβ activation

In Chapter 2, I showed that the Gal4-ETS1 fusion protein failed to activate Dβ1 germline transcription, even when its total protein level was boosted by MG132 protease inhibitor (Figs. 2.7 & 2.8). While ETS1 may not be the direct executer of long-range transactivation, it has been suggested to assist RUNX1 TAD to reach upstream pDβ1 in Chapter 3 (Fig. 3.5). To obtain a more complete picture on ETS1-RUNX1 cooperativity, it is necessary to know what epigenetic changes ETS1 by itself induces. After overexpressing Gal4-ETS1 in EG mutant cells, ChIP assay to examine active histone mark changes and 3C assay to examine long- and short-range promoter-enhancer interactions could be performed. If the binding of Gal4-ETS1 is weak, it will be necessary to use MG132 to boost its expression. In addition, it is possible that ETS1 by itself, like the RUNX1 TAD, retains partial activity for proximal Dβ2 promoters. If this is true, it will suggest that there is synergy between ETS1 and RUNX1 TAD in reaching a distal cis-element. Together, they pass a certain threshold to reach beyond the proximal promoters and to activate pDβ1 from a long distance.

Kinetic analysis of *Tcrb*-RC

Kinetic studies of gene regulation involve two very different time-scales. To observe the synthesis of a certain transcript, it is necessary to use high-speed microscopy to take motion pictures of single cells over hundreds of seconds (1). Another observation focuses on the average signals from a population of cells over hours of time (1, 2). In our previous studies, we proposed a stepwise model for Eβ holocomplex formation and *Tcrb*-RC activation (3). To observe details of this model would a resolution to distinguish temporal differences between epigenetic events. For example, such resolution is needed to record the difference between factor loading on Eβ and those on pDβ1. The reason is that factor loading on Eβ and pDβ1 seems highly synchronous, suggesting that holocomplex formation is instant (Fig. 2.10). In this sense, we will need the high-speed microscope to observe the order of these events.

Meanwhile, the delayed restoration in germline transcription as opposed to factor binding suggests that different types of genetic and epigenetic events may be resolved over the course of hours and days (Fig. 2.10). Especially, in Chapter 3 I observed the uneven restoration of active histone marks over *Tcrb*-RC at 24 hours post induction of Gal4-RUNX1 expression (Fig. 3.5). If these marks are further restored over longer time, I may be able to observe the temporal difference between histone modification and Eβ activation. To investigate the *Tcrb*-RC from as many viewpoints as possible, one would use ChIP-chip or ChIP-Seq with H3K27ac and H3K4me3 antibodies on *Tcrb*-RC at different time points. With these results, a full dynamic depiction of how *Tcrb*-RC is ignited may emerge.

V(D)J recombination and Vβ usage in mouse models

The final product of *Tcrb* gene activation is the TCRβ chain. In our cell models, we would be able to test Dβ-Jβ recombination if we overexpress the RAG1 enzyme. This will further examine the extent of *Tcrb*-RC activation. However, V segments at *Tcrb* in P5424 cells are largely silent compared to primary cells (data unpublished). Therefore, it will be necessary to introduce the inducible system into mouse models. With ZFN technology or the recently developed CRISPR-Cas9 system (4, 5), it will be convenient to construct mouse models with one mutant Eβ allele in RAG1-deficient background. The mice will be bred to obtain homozygous Eβ mutants. These mice will be crossed to mice carrying the tet-inducible expression cassettes of different Gal4 fusion proteins. One would first perform chromosome conformation capture carbon copy (5C) assay to investigate potential loops between V segments and *Tcrb*-RC (6). Next, crosses between the inducible mouse models with RAG1-sufficient mice would be able to examine the influence of different *Tcrb*-RC activation processes on Vβ usage and on in Vβ to DβJβ recombination.

Inducible system for other loci and transcription factors

ETS1-RUNX1 composite binding motifs widely exist across the genome. For example, Eα, the enhancer for *Tcra* locus, is also dependent on its single ETS1-RUNX1 binding motif (7). If Eα is active in P5424, the same inducible system will be applicable to *Tcra*. The same principle may apply to other loci given suitable cell models. Apart from ETS1-RUNX1 composite sites, other transcription factor partnerships may also be studied using the inducible system. For example, in

B lymphocytes, PU.1, an ETS family member, cooperate with IRF4 to activate their target sites (8). Therefore, it is possible to construct cells with PU.1 and IRF4 binding site mutation and test the rescuing by Gal4-PU.1 and Gal4-IRF4 fusion proteins.

References

1. Coulon, A., C. C. Chow, R. H. Singer, and D. R. Larson. 2013. Eukaryotic transcriptional dynamics: from single molecules to cell populations. *Nat. Rev. Genet.* 14: 572–84.

2. John, S., T. A. Johnson, M.-H. Sung, S. C. Biddie, S. Trump, C. A. Koch-Paiz, S. R. Davis, R. Walker, P. S. Meltzer, and G. L. Hager. 2009. Kinetic Complexity of the Global Response to Glucocorticoid Receptor Action. *Endocrinology* 150: 1766–1774.

3. Oestreich, K. J., R. M. Cobb, S. Pierce, J. Chen, P. Ferrier, and E. M. Oltz. 2006. Regulation of TCRβ Gene Assembly by a Promoter/Enhancer Holocomplex. *Immunity* 24: 381–391.

4. Collin, J., and M. Lako. 2011. Concise Review: Putting a Finger on Stem Cell Biology: Zinc Finger Nuclease-Driven Targeted Genetic Editing in Human Pluripotent Stem Cells. *Stem Cells* 29: 1021–1033.

5. Mali, P., L. Yang, K. M. Esvelt, J. Aach, M. Guell, J. E. DiCarlo, J. E. Norville, and G. M. Church. 2013. RNA-Guided Human Genome Engineering via Cas9. *Science (80-.).* 339.

6. Dostie, J., T. A. Richmond, R. A. Arnaout, R. R. Selzer, W. L. Lee, T. A. Honan, E. D. Rubio, A. Krumm, J. Lamb, C. Nusbaum, R. D. Green, and J. Dekker. 2006. Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res.* 16: 1299–309.

7. Giese, K., C. Kingsley, J. R. Kirshner, and R. Grosschedl. 1995. Assembly and function of a TCR alpha enhancer complex is dependent on LEF-1-induced DNA bending and multiple protein-protein interactions. *Genes Dev.* 9: 995–1008.

8. Brass, A. L., A. Q. Zhu, and H. Singh. 1999. Assembly requirements of PU.1-Pip (IRF-4) activator complexes: inhibiting function in vivo using fused dimers. *EMBO J.* 18: 977–91.