Journal of Mammary Gland Biology and Neoplasia (2019) 24:61–71
https://doi.org/10.1007/s10911-018-9418-y

STAT5-Driven Enhancers Tightly Control Temporal Expression of Mammary-Specific Genes

Ha Youn Shin[1] · Lothar Hennighausen[2,3] · Kyung Hyun Yoo[4]

Received: 19 June 2018 / Accepted: 12 October 2018 / Published online: 17 October 2018

Abstract

The de novo formation of milk-secreting mammary epithelium during pregnancy is regulated by prolactin through activation of the transcription factor STAT5, which stimulates the expression of several hundred mammary-specific genes. In addition to its key role in activating gene expression in mammary tissue, STAT5, which is ubiquitously expressed in most cell types, implements T cell-specific programs controlled by interleukins. However, the mechanisms by which STAT5 controls cell-specific genetic programs activated by distinct cytokines remain relatively unknown. Integration of data from genome-wide surveys of chromatin markers and transcription factor binding at regulatory elements may shed light on the mechanisms that drive cell-specific programs. Here, we have illustrated how STAT5 controls cell-specific gene expression through its concentration and an auto-regulatory enhancer supporting its high levels in mammary tissue. The unique genomic features of STAT5-driven enhancers or super-enhancers that regulate mammary-specific genes and their dynamic remodeling in response to pregnancy hormone levels are described. We have further provided biological evidence supporting the in vivo function of a STAT5-driven super-enhancer with the aid of CRISPR/Cas9 genome editing. Finally, we discuss how the functions of mammary-specific super-enhancers are confined by the zinc finger protein, CTCF, to allow exclusive activation of mammary-specific genes without affecting common neighboring genes. This review comprehensively summarizes the molecular pathways underlying differential control of cell-specific gene sets by STAT5 and provides novel insights into STAT5-dependent mammary physiology.

Keywords STAT5 · Enhancer · Mammary-specific gene expression

Introduction

The mammary gland is a secretory organ in mammals specialized in producing milk to feed young offspring. Uniquely, milk-secreting mammary epithelium only differentiates during gestation via a process that is tightly regulated by steroid and peptide hormones [1, 2]. Steroid hormones, such as estrogen and progesterone, are essential for stimulating ductal growth and branching in postnatal mammary tissue [3, 4] whereas a single peptide hormone, prolactin, controls the proliferation and differentiation of milk-secreting alveolar cells during pregnancy and lactation [5]. Upon prolactin binding to its receptor, the key downstream transcription factor, signal transducer and activator of transcription 5 (referred to as STAT5), is activated and gains entrance into the nucleus [6]. Nuclear-localized STAT5 binds a DNA recognition motif in the regulatory region of target genes and activates transcription. STAT5 exists as two isoforms, STAT5A and STAT5B, encoded by separate genes [7–9]. Although STAT5A and STAT5B display partially redundant functions in mammary development, loss-of-function studies in mice and human have additionally revealed distinct roles. Deletion of *Stat5a* in mice leads to impaired mammary alveolar development, although this is partially complemented by *Stat5b* [10–12]. Inactivation of *Stat5b* in mice triggers body growth defects and impaired immune cell proliferation [13, 14]. Ablation of both *Stat5a* and *Stat5b* in mice causes prenatal lethality due to fetal anemia [10, 15, 16]. STAT5 is widely expressed in the majority of tissues. Notably, STAT5A is predominantly

✉ Kyung Hyun Yoo
khryu@sookmyung.ac.kr

[1] Department of Biomedical Science and Engineering, Konkuk University, Seoul 05029, Republic of Korea

[2] Laboratory of Genetics and Physiology, National Institutes of Diabetes, Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA

[3] BK21 PLUS Project, Sookmyung Women's University, Seoul 04310, Republic of Korea

[4] Department of Biological Sciences, Sookmyung Women's University, Seoul 04310, Republic of Korea

detected in mammary tissue while STAT5B is more abundant in immune cells [17]. However, little is known about the mechanisms underlying the differential expression patterns and biological functions of STAT5.

Recent advances in genome-wide analysis have improved our fundamental understanding of the unique genomic structure of regulatory elements that drives cell-type specific gene expression [18, 19]. Genome-wide screening of histone modification markers and transcription factor binding is a commonly used technique to predict putative regulatory elements, such as enhancers and promoters. Active enhancers and promoters are frequently found in open chromatin regions devoid of nucleosomes and highly occupied by RNA polymerase II (RNA Pol II) and cofactors [20, 21]. In particular, active enhancers are generally modified by acetylation at lysine 27 of histone H3 (H3K27ac) whereas active promoters are enriched with trimethylation of H3K4 (H3K4me3). Advanced genome engineering technologies, such as CRISPR/Cas9, further facilitate determination of the biological functions of putative regulatory elements [22–24]. This review comprehensively summarizes the mechanisms by which STAT5 controls the transcription of mammary-specific gene sets, mechanisms involving STAT5 autoregulatory enhancers that respond to alterations in cytokine concentrations. We additionally describe the unique genomic features and biological functions of STAT5-driven mammary specific super-enhancers and their dynamic remodeling during mammary gland development.

Autoregulation of Mammary-Specific STAT5

Recent RNA-seq data revealed that *Stat5a* mRNA levels are higher in mammary tissue than in other cell types, suggesting mammary-specific transcriptional control of the *Stat5* locus [22]. To determine the regulatory elements that are exclusively active in mammary tissue, Metser et al. conducted DNase-seq analysis with the aim of identifying DNase I-hypersensitive sites (DHS) at the *Stat5* locus in eight different mouse tissues, including mammary tissue at day one of lactation (L1). The DNase-seq technique involves genome-wide sequencing of regions sensitive to DNase I cleavage and is commonly used to identify open chromatin regions that bind transcriptional machinery [20]. The *Stat5* locus encodes two distinct genes, *Stat5a* and *Stat5b*, positioned in a head-to-head orientation with a ~10 kb intergenic region between them [25]. While several DHS sites were identified in the promoter regions of both *Stat5a* and *Stat5b* in most tissues, including heart, liver, lung and muscle, a single site is present in the intergenic region of the *Stat5* locus exclusively in mammary tissue. This single open chromatin region is located closer to the transcription start site (TSS) of *Stat5a* than *Stat5b*. Chromatin immunoprecipitation-sequencing (ChIP-seq) experiments further revealed that this mammary-specific intergenic region is highly enriched with an active enhancer marker, H3K27ac, RNA Pol II and STAT5 binding (Fig. 1a). Additional ChIP-seq experiments demonstrated that mammary-enriched transcription factors, including the glucocorticoid receptor (GR), nuclear factor I/B (NFIB) and E74-like factor 5 (ELF5), co-localize with STAT5 in the putative enhancer region of the mouse genome. GR synergistically induces milk protein genes with STAT5 in response to prolactin stimulation [26]. NFIB acts as a co-regulator of STAT5 and activates mammary-specific genes [27], and ELF5 is predominantly expressed in mammary epithelium and shown to be essential for mammary development during pregnancy [28]. Collective data from DNase I hypersensitive and ChIP-seq studies have facilitated characterization of the putative enhancer regions exclusively activated in mammary tissue.

To further establish the in vivo function of STAT5 at the putative enhancers, CRISPR/Cas9 genome editing was employed to generate mice lacking two juxtaposed STAT5 binding motifs (GAS, TTCnnnGAA) within the intergenic enhancer region [22]. The GAS motif refers to an Interferon-Gamma Activated Sequence, a consensus DNA-recognition motif of all STAT family members, with the exception of STAT6. Once cytokine-stimulated STATs translocate into the nucleus, they bind to GAS motifs in the regulatory elements of target genes, inducing gene transcription [6]. Deletion of the two canonical STAT5 binding motifs in the *Stat5* locus led to the reduction of *Stat5a* expression by ~80% in mammary tissue whereas mRNA levels remained constant in T cells and liver, indicating that *Stat5a* gene expression is specifically controlled via autoregulation in mammary tissue. To further investigate whether the intergenic *Stat5* enhancer is involved in mammary-specific regulation of STAT5 target genes, mRNA levels of several target genes were examined in enhancer mutants. Notably, while the expression of mammary-specific genes, such as *Wap*, *Csn1s2b*, and *Csn2*, was significantly reduced, the common *Stat5* target gene, *Socs2,* appeared insensitive to STAT5 levels (Fig. 1b). *Wap* and casein (*Csn)* are milk protein genes activated by prolactin [29] while *Socs2* is a negative regulator of STAT5, which is induced by a subset of cytokines in most tissues [30]. Consistent with local gene expression data, RNA-seq experiments revealed global alterations in the expression of mammary-specific STAT5 target genes upon mutation of the *Stat5* enhancer. Approximately 155 genes were upregulated and ~200 genes downregulated in mutant mice. These findings suggest that the intergenic *Stat5* enhancer maintains high expression levels of STAT5 in mammary tissue via autoregulation and is further involved in regulation of mammary-specific STAT5 target genes.

Differential Cytokine Sensitivity of STAT5-Driven Enhancers

The mammary gland undergoes repeated cycles of proliferation and differentiation of epithelium, milk secretion and

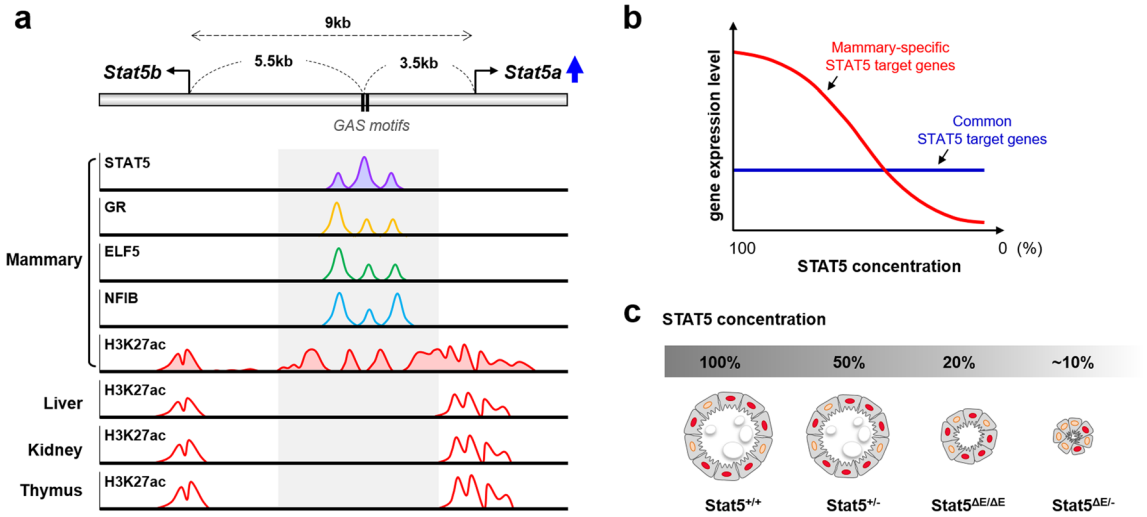

Fig. 1 Autoregulation of the *Stat5* locus in mammary tissue. **a** ChIP-seq profiles of an auto-regulatory enhancer at the *Stat5* locus. Mammary-specific STAT5 binding coincides with two adjacent GAS motifs, 3.5 kb upstream of *Stat5a*. Mammary-enriched transcription factors, including GR, ELF5, and NFIB, co-localize with STAT5. The active enhancer marker, H3K27ac, is highly enriched in the intergenic region of *Stat5* locus devoid of transcription factor binding in mammary tissue, but not liver or T cells. **b** Differential STAT5 sensitivity of mammary-specific genes and common STAT5 target genes. Mammary-specific STAT5 target genes are highly induced with increasing STAT5 concentrations, whereas common STAT5 target genes are insensitive. **c** Role of *Stat5* auto-regulatory enhancers in mammary alveolar differentiation. Mice carrying one intact *Stat5* allele (Stat5 $^{+/-}$) displayed normal alveolar mammary differentiation while those homozygous for the enhancer mutation (*Stat5*$^{\Delta E/\Delta E}$) showed less developed mammary epithelium with reductions in the numbers and sizes of MECs and those containing an enhancer mutation and null allele (*Stat5*$^{\Delta E/-}$) were underdeveloped

apoptosis during pregnancy [1, 2]. This cyclic developmental process is mainly controlled by fluctuation of prolactin levels. Increase in prolactin levels during pregnancy and lactation activates STAT5 and further induces genes essential for mammary alveolar cell proliferation and differentiation, and conversely, decline in prolactin levels during involution deactivates STAT5 and reduces the number of mammary epithelial cells (MECs) within alveoli in mice [31–33]. Willi and co-workers [34] further investigated the effects of cytokine levels on the structure of STAT5-driven enhancers at different mammary developmental stages. ChIP-seq experiments conducted with STAT5 and H3K27ac during lactation led to the identification of ~9200 putative STAT5-driven enhancers within the mammary genome at L1. Enhancers can be categorized into three groups, depending on the dissociation rates of active enhancer marks and mammary-enriched transcription factors upon termination of lactation. Enhancers are cis-acting regulatory elements comprised of nucleotide sequences in the genome. They are activated upon recruitment of transcription factors, followed by the establishment of a permissive histone modification landscape and subsequent induction of associated genes. Conversely, the detachment of these factors inactivates enhancers and leads to loss of gene transcription. These reversible chromatin features are often controlled by environmental cues, such as cytokines or hormones [20]. Among the ~9200 active enhancers established at L1, ~4370 lost their activity within 12 h of involution (I12) (designated 'class 1' enhancers), ~4520 'class 2' enhancers were inactivated after 24 h (I24) and 315 'class 3' enhancer activities were retained, even after 24 h of involution. Because ChIP-seq analyses were performed with mouse mammary tissue after the removal of lymph nodes and adipocytes, distinct features of classified enhancers may be specific to mammary epithelial cells. A search for STAT5 binding motifs within the three enhancer groups revealed the presence of a GAS motif in 38% class 1, 54% class 2 and 74% class 3 enhancers. Notably, enhancers containing the GAS motif were highly occupied by STAT5 in all three groups, with the most resistant enhancer group (class 3) being occupied by the highest levels of STAT5. Moreover, STAT5-sensitive enhancers (class 1 and 2) preferentially control mammary-specific genes whereas enhancers insensitive to STAT5 concentration (class 3) are prevalently linked to widely expressed genes. These results suggest that mammary-specific enhancers and universal enhancers have differential sensitivity to cytokine levels and STAT5 concentrations.

Genetic studies complementing genome-wide analyses have further confirmed that class 1 enhancers are sensitive to STAT5 concentrations. Mutant mice with four distinct genotypes were employed to investigate the effects of STAT5 integrity on enhancer function. Mice carrying two intact *Stat5* alleles (*Stat5*$^{+/+}$) expressed 100% STAT5 whereas those with a single *Stat5* allele (*Stat5*$^{+/-}$) expressed 50% STAT5, although their histological appearance was normal and STAT5 activation levels equivalent to wild-type mice (Fig. 1c). Mice homozygous for *Stat5* auto-regulatory enhancer mutation (*Stat5*$^{\Delta E/\Delta E}$) only expressed 25% STAT5 and showed reductions in the number and size of MECs in alveoli, while still displaying hallmarks of cell differentiation. Mice with one *Stat5*-null and

one enhancer mutant allele (*Stat5*$^{\Delta E/-}$) expressed ~10% STAT5 and were underdeveloped. Notably, the integrity of STAT5-sensitive class 1 enhancers was significantly reduced at L1, reminiscent of wild-type involution, in *Stat5* auto-regulatory enhancer mutants (*Stat5*$^{\Delta E/\Delta E}$). These mutational studies demonstrate that the *Stat5* auto-regulatory enhancer plays a critical role in maintaining STAT5 levels in mammary tissue and the integrity of STAT5-driven mammary-specific enhancers is significantly affected by the STAT5 concentration.

Differential Roles of STAT5-Driven Enhancers and Promoters

STAT5 activates both lineage-specific and ubiquitously expressed genes through distinct regulatory mechanisms [6, 35]. To elucidate the mechanisms by which STAT5 executes differential gene regulation, Zeng et al. [24] mapped genome-wide STAT5 binding sites in three different tissue types (lactating mammary epithelium, T helper cells, and liver) in mice via ChIP-seq analyses. In total, 400 STAT5 binding sites were commonly present in all three tissues, ~16,000 STAT5 binding sites exclusively in mammary tissues, 12,300 sites in liver, and 12,400 sites in T cells. Notably, shared STAT5 binding sites were significantly enriched in promoters and mammary-specific STAT5 binding sites predominantly detected in distal enhancers enriched with an active enhancer marker, H3K27ac. Mammary-enriched transcription factors co-localized with STAT5 on putative enhancers, but not promoters. A motif search revealed that DNA binding motifs of the mammary-enriched transcription factors ELF5, GR, and NFIB, are significantly enriched in STAT5-bound enhancers in mammary tissue whereas binding motifs for liver-specific HNF4A and FOXA1 are dominant in STAT5-driven enhancers in liver.

Differential cytokine sensitivity has also been reported in lineage-specific enhancers and STAT5-binding common promoters. Chromatin accessibility of lineage-specific enhancers is highly dependent on cytokine levels while common promoters are relatively insensitive. Although both *Wap* and *Socs2* expression are controlled by STAT5 [23, 30], chromatin accessibility of three enhancers upstream of mammary-specific *Wap* gene is highly sensitive to prolactin levels whereas that of the promoter found in widely expressed *Socs2* remains constant, independent of the growth hormone level (Fig. 2). Concordantly, genes in close proximity to lineage-specific enhancers are significantly induced in response to cytokines, compared to those near common promoters. The *Wap* gene is induced up to 1,000-fold during pregnancy while *Socs2* is only expressed at the basal level.

To further explore the role of STAT5 in the promoter regions of widely expressed genes, two adjacent STAT5-binding motifs within the *Socs2* promoter were mutated via CRISPR/Cas9 genome editing in mice. Animals with STAT5 mutant promoter regions showed impaired *Socs2* induction during pregnancy and lactation, along with precocious mammary gland development. However, basal activity of *Socs2* was maintained in mutants, suggesting that STAT5 binding is essential for cytokine-induced *Socs2* expression, but not basal activity. Taken together, these findings indicate that STAT5 differentially controls lineage-specific genes via enhancers and widely expressed genes through promoters, leading to the proposal that STAT5 is required for activation of cytokine-sensitive enhancers but only fine-tunes promoter activities upon cytokine stimulation.

STAT5-Driven Mammary-Specific Super-Enhancers

Recent progress in genome-wide technologies has led to the discovery of a hitherto unknown type of enhancer with unique genomic structures and biological functions. In 2013, Whyte et al. [36, 37] first introduced the novel concept of super-enhancers in embryonic stem cells, which drive cell-type specific gene control. In the same year, a similar concept of stretched enhancers was proposed by Parker et al. [38]. Further studies focused on characterization of the chromatin structure of super-enhancers in a wide range of cell types, including T cells [39], hair follicle cells [40] and chondrocytes [41]. Compared to conventional enhancers, super-enhancers are clusters of enhancers covering long regulatory regions (> 10 kb) that are highly occupied by H3K27ac, a subunit of the mediator complex MED1 [42], and lineage-specific or master transcription factors. Although super-enhancers are frequently located near genes specifically expressed in certain cell types, their bona fide biological function remains a subject of controversy due to the lack of in vivo studies [43]. The mouse genome contains 440 mammary-specific super-enhancers, and in vivo biological activities of super-enhancers were determined using experimental mouse genetics [23]. The chromatin structure of mammary-specific super-enhancers in cell nuclei obtained from frozen mammary tissue was assessed by ChIP-seq analyses with various mammary-enriched transcription factors and active enhancer markers. Similar to super-enhancers found in other tissues, mammary-specific super-enhancers are enriched with the H3K27ac marker, MED1, and RNA Pol II. Unique features of these super-enhancers include high occupancy of the master transcription factor, STAT5, with co-localization of mammary-enriched transcription factors, such as GR, NFIB, and ELF5 (Fig. 3a). For example, the mammary-specific *Wap* super-enhancer comprises three constituent enhancers, a proximal enhancer E1 (−650 bp for TSS), enhancer E2 (−1.4 kb from TSS), and the most distal enhancer E3 (−4.7 kb from TSS) that are highly occupied by mammary-enriched transcription factors.

The in vivo biological roles of mammary-specific super-enhancers were investigated by examining constituent enhancers of the *Wap* super-enhancer in mice. CRISPR/Cas9 genome editing and TALEN technology were utilized to

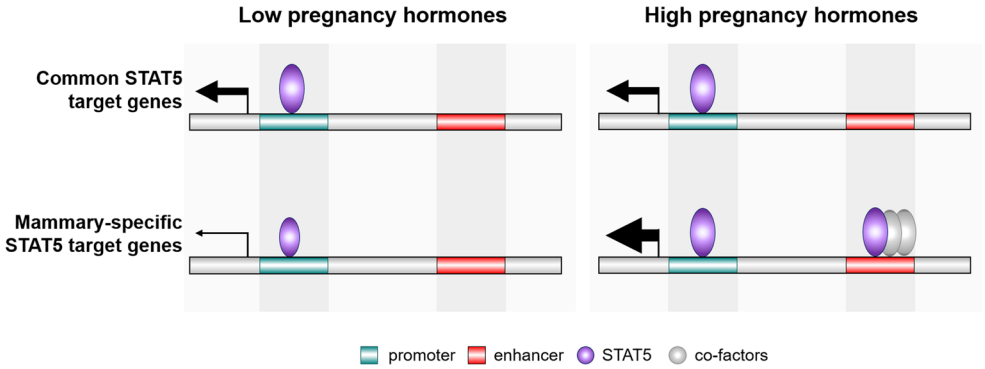

Fig. 2 Differential cytokine sensitivities of STAT5-driven enhancers and promoters. At low pregnancy hormone levels, STAT5 binds to the promoter of widely expressed *Socs2* or mammary-specific *Wap* and induces basal gene expression. In response to high levels of pregnancy hormones, enhancers upstream of mammary-specific *Wap* are occupied by STAT5 and mammary-enriched transcription factors and activate gene expression to a significant extent. Pregnancy hormones include prolactin, progesterone and estrogen, which are produced during gestation

generate mutations in the three individual enhancers and their combinations [44]. The STAT5 binding site at E1 was deleted via CRISPR/Cas9 and STAT5 binding sites at E2 and E3 with TALEN. Deletion of individual enhancers revealed differential roles within the *Wap* super-enhancer. Specifically, inactivation of E1 (ΔE1) reduced *Wap* gene induction at L1 by 62%, loss of E2 (ΔE2) resulted in 48% gene reduction, and deletion of distal E3 (ΔE3) led to 91% gene reduction. These data suggest that while the constituent enhancers have similar transcription factor occupancy and H3K27ac profiles, they play different roles within *Wap* super-enhancers, the most critical being the distal enhancer E3. These results are informative but do not elucidate the mechanisms underlying *Wap* gene induction up to 1,000-fold during lactation. Importantly, deletion of STAT5 binding sites at all three constituent enhancers (ΔE1/ΔE2/ΔE3) suppressed *Wap* gene expression at L1 to ~0.1% that of the wild-type control. Consistent with the functional hierarchy in individual mutants, combinatorial deletion of E1 and E2 (ΔE1/ΔE2) led to retention of 12.5% *Wap* mRNA whereas loss of E1 and E3 (ΔE1/ΔE3) reduced the mRNA level to 5.5%, validating the significance of the most distal enhancer, E3. These mouse genetic experiments clearly demonstrated the biological relevance of super-enhancers in cell type-specific gene control and the distinct roles of constituent enhancers.

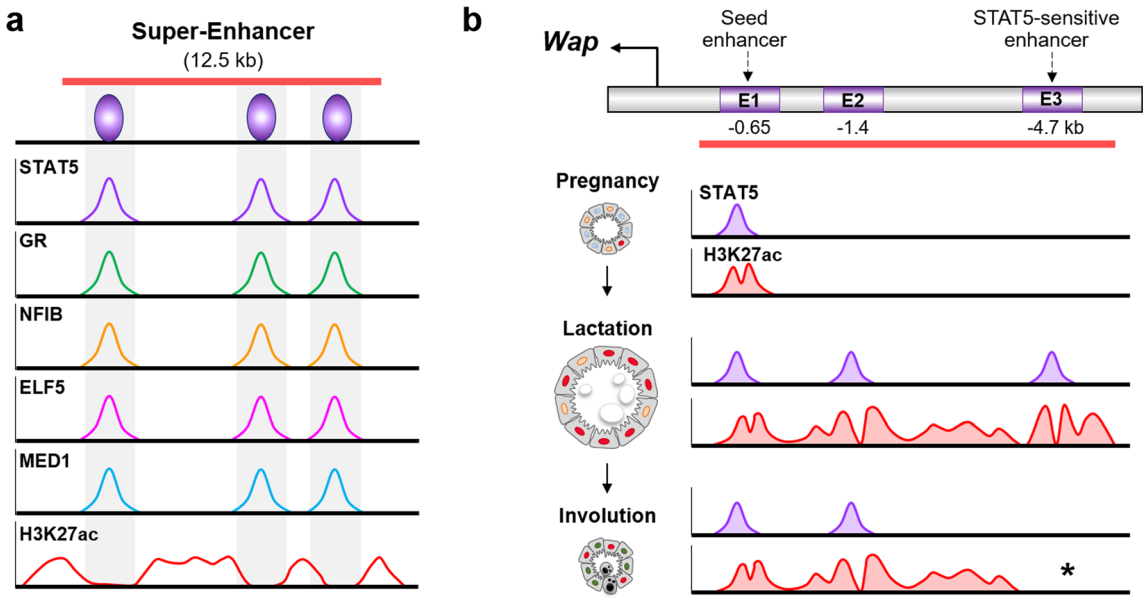

Fig. 3 *Stat5*-driven mammary-specific super-enhancers. **a** Genomic features of mammary-specific super-enhancers identified by ChIP-seq analyses. Super-enhancers are clusters of enhancers covering long regions of genomic DNA (> 12.5 kb). Constituent mammary-specific enhancers are co-occupied by mammary-enriched transcription factors (STAT5, GR, NFIB and ELF5), a mediator (MED1) and active enhancer marker, H3K27ac. **b** Dynamic remodeling of mammary-specific super-enhancers demonstrated by ChIP-seq analyses during mammary gland development. Mammary-specific *Wap* super-enhancer consists of three constituent enhancers, E1, E2 and E3. At the early pregnancy stages, only proximal E1 is occupied by STAT5 and H3K27ac. During lactation, all three enhancers are fully established. After termination of lactation, only E3 rapidly disappears, indicating that the most distal E3 is highly sensitive to STAT5 concentrations

Dynamic Remodeling of Mammary-Specific Super-Enhancers during Mammary Gland Development

Super enhancers are known to be sensitive to environmental changes. Previously, Adam et al. [40] showed that super-enhancers in hair follicle stem cells undergo remodeling during lineage progression. Dynamic remodeling has additionally been detected in mammary-specific super-enhancers. Chromatin structures of mammary-specific super-enhancers were investigated on day 13 of pregnancy (p13) and L1 in mice [34]. Importantly, >95% of super-enhancers were gradually established from p13 to L1. Approximately 7% super-enhancers were not established at p13 and only occupied at L1. In 56% super-enhancers, less than half the constituent enhancers were already occupied at p13. In 32% super-enhancers, more than half the constituents were occupied at p13 whereas 5% were fully established at p13. Only E1 was occupied at p13 whereas E2 and E3 were occupied at L1 in the case of the *Wap* super-enhancer (Fig. 3b). These genomic alterations suggest that super-enhancers are sensitive to extracellular signals, such as cytokine stimulation.

Concordantly, decommissioning of mammary-specific super-enhancers resembles the sequential process of enhancer establishment. ChIP-seq profiles of STAT5 and H3K27ac revealed rapid eradication of mammary-specific super-enhancers after termination of lactation [34]. Approximately 75% of the genes were associated with mammary-specific enhancers, with loss of at least one enhancer within 12 h of involution and 24% genes within 24 h, while 1% of the genes were resistant to involution. For instance, one constituent enhancer within the *Wap* super-enhancer was decommissioned within 12 h and two others within 24 h of involution (Fig. 3b). Importantly, the distal enhancer E3 first disappeared within 12 h in response to the decline in prolactin levels and E1 and E2 disappeared after 24 h of involution. Additional ChIP-seq experiments revealed that another STAT family member, STAT3, can effectively replace STAT5, since both STAT5 and STAT3 recognize identical DNA binding motifs [45, 46]. Distinct and reciprocal roles of STAT5 and STAT3 have been demonstrated during normal mammary gland development and in breast cancer cells [10, 11, 32, 47]. Whereas STAT5 is activated during late pregnancy and lactation, resulting in alveolar proliferation and differentiation [10, 11], STAT3 is activated during involution and leads to cell apoptosis [32]. In breast cancer cells, STAT5 and STAT3 compete to bind to the same regulatory site on the *Bcl6* gene, a transcriptional repressor involved in preventing the transcription of the β-casein gene and repressing the terminal differentiation of the mammary gland [47]. ChIP analyses have demonstrated that, although prolactin-stimulated STAT5 reduces *Bcl6* gene expression by inhibiting RNA Pol II binding, LIF-induced STAT3 increases *Bcl6* gene transcription by recruiting RNA Pol II in breast cancer cells. In the *Wap* super-enhancer, STAT5 binding at E1 and E2 was effectively replaced with STAT3 after 24 h of involution. Immunohistochemical analyses consistently revealed that loss of phosphorylated STAT5 (p-STAT5) within 12 h of involution could be rapidly replaced by p-STAT3 [34]. Thus, super-enhancers have the capacity to fine-tune the transcriptional activation of cell type-specific genes through sequential enhancer establishment and retrograde enhancer destruction in response to extracellular cytokine stimuli.

CTCF as a Moderator of Mammary-Specific Gene Regulation

Super-enhancers regulate cell-type specific gene transcription and induce gene activation up to several thousand-fold. To avoid the spillover of transcriptional control to neighboring genes, the activities of super-enhancers need to be confined to their respective target genes. The genome is proposed to be compartmentalized into a topology-associated domain (TAD) that provides a shield from neighboring regulatory units [48]. The zinc-finger protein CCCTC binding factor (CTCF) is enriched in the border of TAD and serves an anchor for chromatin looping to perform its insulation function [49–51]. However, biological functions of CTCF sites have been mainly determined in cell lines and embryonic stem cells [52–55] and limited genetic evidence exists in the literature. Through ChIP-seq profiling of CTCF sites within mammary-specific *Wap* super-enhancers and examination of individual sites using CRISPR/Cas9 genome editing, the in vivo biological functions of CTCF sites were analyzed [56]. The *Wap* gene is located between two distinct genes, *Ramp3* and *Tbrg4*, expressed at low levels in many cell types (Fig. 4a). This locus contains five CTCF sites that are shared between different tissues. Site A coincides with a TAD border and site E presumed to serve as an anchorage for chromatin looping. Sites C and D preferentially respond to cytokines. To ascertain the specific role of each CTCF site in the *Wap* locus, sites were deleted individually and in combination via CRISPR/Cas9 genome editing. While none of the CTCF mutations affected *Wap* gene expression, loss of site E induced a five-fold increase in the neighboring *Ramp3* gene. Combined deletion of sites C, D and E further led to seven-fold induction of *Ramp3*, suggesting that C and D interact with site E. Moreover, chromosome conformation capture-on-chip (4C) experiments revealed that deletion of all three CTCF sites enhanced interactions between E3 of the *Wap* locus and the first intron of *Ramp3*. These results indicate that CTCF sites are not sufficient to block super-enhancer activity but are able to silence neighboring gene activation.

In vivo biological functions of CTCF sites were further investigated in the larger context of the mammary-specific *Csn* locus [57] consisting of five *Csn* genes that are exclusively expressed in mammary tissue and induced up to 1,000-fold

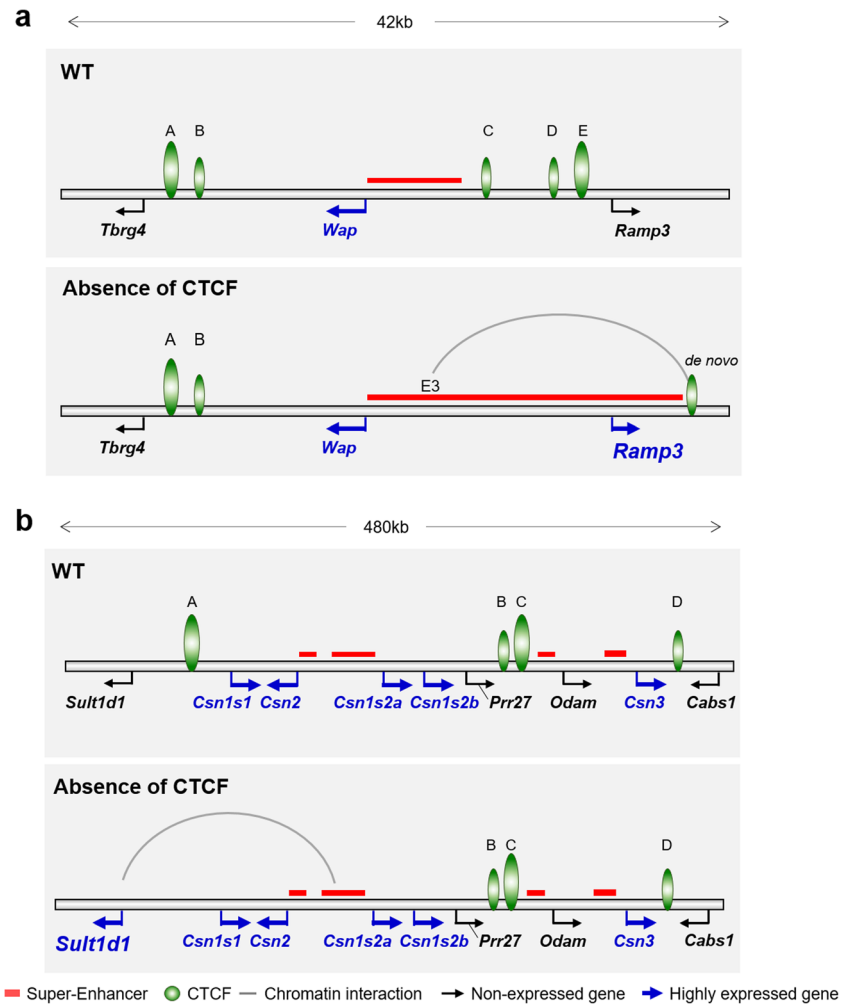

Fig. 4 The role of CTCF sites in mammary-specific *Wap* and *Csn* locus. **a** The porous nature of the CTCF site at the *Wap* locus. Mammary-specific *Wap* locus containing five CTCF binding sites (A to E) is located between *Tbrg4* and *Ramp3*. While deletion of sites C, D, and E did not affect *Wap* expression, the neighboring *Ramp3* gene was significantly induced, along with elevated chromatin interactions between E3 of the *Wap* super-enhancer and *Ramp3* locus. **b** Induction of neighboring gene activity and de novo establishment of chromatin interactions at the *Csn* locus in the absence of a functional CTCF site. The mammary-specific *Csn* locus contains four super-enhancers with four CTCF binding sites (A to D) within a 480 kb segment. Although all the CTCF sites examined did not function in restriction of the regulatory units, deletion of site A significantly induced neighboring *Sult1d1* activity through de novo chromatin interactions between the *Csn1s2a* super-enhancer and *Sult1d1* promoter

during pregnancy. Four mammary-specific super-enhancers are present within the extended *Csn* locus from *Csn1s1* to *Csn3* (Fig. 4b), two located between *Csn2* and *Csn1s2a*, one upstream of *Odam*, and one upstream of *Csn3*. The *Sult1d1* gene is positioned 95 kb upstream of *Csn1s1* and expressed at low levels in mammary tissue. In contrast, neither *Sutl1e1* positioned closer to *Csn1s1* nor *Cabs1* located 56 kb downstream of *Csn3* is expressed in mammary tissue. Lee and coworkers conducted CTCF ChIP-seq analyses on mouse mammary tissue during lactation and identified four CTCF sites (referred to as sites A, B, C, and D) in the *Csn* locus [57]. The individual CTCF sites were mutated via CRISPR/Cas9 genome editing in mice and their roles in establishing insulated neighborhoods investigated. While mutation of individual CTCF sites did not alter *Csn* gene levels in mammary tissue, loss of site A (but no other mutations) induced increased expression of neighboring *Sult1d1* up to 3-fold in mammary tissue. Although *Sult1d1* gene expression was induced in the absence of CTCF site A, the H3K27ac landscape and STAT5 binding profiles remained unchanged and only RNA Pol II binding levels were increased in the promoter region. These findings suggest that the only genes that are already expressed, even at low levels, are induced in response to loss of the CTCF site. To further ascertain whether promoter and enhancer interactions are affected in the absence of CTCF site A, chromosome conformation capture (3C) experiments were performed. Interactions between CTCF sites A and C were demonstrated in WT mammary tissue and showed a ~10-fold decrease upon removal of site A. De novo interactions between *Sultd1d* and the *Csn* locus were additionally established. Taken together, the data indicate that CTCF does not play a prominent role in demarcating distinct regulatory units but abrogation of functional CTCF sites triggers upregulation of neighboring genes that are already activated, further leading to establishment of de novo chromatin interactions.

Conclusions

Transcription factors convey extracellular signals to responsive genes in a wide range of cells. Limited studies to date have demonstrated how common transcription factors, such as STAT5, differentially regulate cell type-specific and widely expressed genes. This review provides an overview of the

differential regulatory mechanisms of STAT5 during mammary gland development.

STAT5 expression levels in mammary epithelium are higher than in other cell types, a result of an auto-regulatory enhancer. High STAT5 concentrations are the underpinning of the activation of mammary-specific target genes during pregnancy and lactation. Mammary-specific gene regulation may be due to the unique chromatin features of the STAT5 autoregulatory enhancer and its target gene enhancers, such as the *Wap* and *Csn* enhancers. A search for motifs indicative of transcription factors demonstrated that these enhancer regions contain a group of DNA binding sites for mammary-enriched transcription factors, including STAT5, GR, NFIB, and ELF5 [22, 23]. ChIP-seq analyses further demonstrated that these enhancers are densely bound by those transcription factors and active histone markers. Similar genomic arrangements have been observed in the regulatory regions of the β-casein gene. For example, the proximal promoter and distal enhancer of the β-casein gene contain DNA binding motifs for transcription factors essential for mammary gland development (i.e., STAT5, GR, NF-1, ETS and C/EBPβ), a group called "composite response elements" [58]. CEBP/β is a member of the CCAAT/enhancer binding protein family of bZIP transcription factors essential for mammary gland development and expression of genes encoding milk proteins. Cooperation of STAT5, GR, and C/EBPβ synergized to induce the transcription of the β-casein gene in HC11 mammary epithelial cells. ChIP assays revealed that this synergistic effect is achieved by both protein-DNA binding and protein-protein interactions of these transcription factors [59]. Indeed, the amino-terminal domain of GR can interact with STAT5, whereas the DNA-binding domain of GR can interact with the amino-terminal transactivation domain of C/EBPβ. More importantly, because the DNA binding motif of GR is only present in the promoter, not in the enhancer, and STAT5 can only interact with C/EBPβ in the presence of full length GR, GR may bridge STAT5 and C/EBP β [59]. Further 3C experiments demonstrated the long-range chromatin interaction between the promoter and enhancer of the β-casein gene. Chromatin looping occurred only after the stimulation with both prolactin and hydrocortisone in HC11 cells, 3D mammary acini cultures and primary MECs isolated from lactating mice, but did not occur following treatment with a single hormone, in untreated cells and in MECs of virgin mice [60, 61].

Boosting expression of cell type-specific genes via an autoregulatory mechanism has also been observed in other cell types. GATA1 is a transcription factor essential for erythropoiesis and its autoregulation ensures the maintenance of high expression levels during erythroid lineage differentiation [62, 63]. STAT5 also differentially controls lineage-specific and widely expressed genes through enhancers and promoters, respectively. This distinct engagement of a transcription factor in gene regulation has been identified in chondrocytes with SOX9, a principal transcription factor of chondrocyte specification and differentiation [64]. SOX9 binding motifs as well as binding intensities are highly enriched in enhancers compared to promoters. Moreover, enhancer-associated genes are highly expressed in lineage-specific cells whereas promoter-associated genes are generally expressed in a wide range of cells.

Extensive cell type-specific gene control is driven by super-enhancers. Super-enhancers have been identified in a variety of cell types and their unique chromatin compositions, including lineage-specific transcription factor binding, are well characterized. This review focused primarily on the mammary-specific regulation of the *Wap* and *Csn* genes. Additional mammary-specific genes, such as *Wdnm1* and *Lalba*, the latter of which encodes alpha-lactalbumin, display gene expression patterns similar to those of *Csn* and *Wap*, respectively [65]. Indeed, the expression pattern of *Wdnm1* is more similar to *Csn2*, in that both genes are expressed from mid-pregnancy (p14) and increase during lactation [66]. In contrast, expression of both *Lalba* and *Wap* occurs in late pregnancy(p18 to L1). Our recent RNA-seq data also supports these findings [23]. Further studies are needed to determine whether the regulatory mechanisms of *Wdnm1* and *Lalba* are similar to those of *Csn2* and *Wap*, respectively.

In response to the levels of extracellular signals, the structures of super-enhancers evolve, leading to fine-tuning of specific gene activation. More detailed studies may be necessary to confirm whether super-enhancers are sufficient to determine specific cell types. The issue of whether introduction of a mammary-specific super-enhancer into a non-mammary locus can activate a silent gene in mammary tissue is of significant research interest. Super-enhancers that are co-occupied by mammary-enriched transcription factors (i.e., STAT5, GR, NFIB, ELF5) can control the transcription of mammary-specific gene sets involved in mammary tissue differentiation. In contrast, super-enhancers bound by embryonic stem cell markers (i.e., Oct4, Sox2, and Nanog) specifically regulate genes that determine the identity of embryonic stem cells. Hence, the induction of groups of genes involved in cell identity are dependent on the type of lineage-specific transcription factors bound to super-enhancers, with the expression of these genes possibly further driving specific cell types.

Extensive investigation of the potential role of CTCF in insulating neighborhoods has revealed that not all CTCFs within mammary-specific super-enhancers function as boundary molecules. Functional CTCF sites silence but do not block neighboring gene activation. Moreover, deletion of CTCF sites only promotes activation of genes that are already expressed. Integration of genome-wide analyses and CRISPR/Cas9 genome engineering technology should shed further light on the veiled genomic structures essential for cell type-specific control and their biological functions.

The recent development of single cell sequencing technology has provided new opportunities to determine the

transcriptome and chromatin features. This, in turn, may facilitate understanding of the mechanisms underlying cell-type specific gene controls. Indeed, single cell RNA-seq analyses have been exploited to understand the distinct gene expression profiles of different lineages of mammary epithelial cells, including luminal cells and basal cells [67–70]. However, determining the chromatin features of single cells is technically more challenging. In our experience, transcription factors could be easily detached from chromatin during isolation of single mammary epithelial cells, resulting in a failure of ChIP-seq analyses. Further technical advances are essential for deepening our understanding of the complex genetic circuits controlling mammary physiology.

Acknowledgements This work was supported by grants from Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (S2017R1C1B2002806, S201803S00103) and the IRP of NIDDK/NIH.

References

1. Hennighausen L, Robinson GW. Information networks in the mammary gland. Nat Rev Mol Cell Biol. 2005;6(9):715–25. https://doi.org/10.1038/nrm1714.
2. Hennighausen L, Robinson GW. Signaling pathways in mammary gland development. Dev Cell. 2001;1(4):467–75.
3. Bocchinfuso WP, Lindzey JK, Hewitt SC, Clark JA, Myers PH, Cooper R, et al. Induction of mammary gland development in estrogen receptor-alpha knockout mice. Endocrinology. 2000;141(8):2982–94. https://doi.org/10.1210/endo.141.8.7609.
4. Silberstein GB, Van Horn K, Shyamala G, Daniel CW. Progesterone receptors in the mouse mammary duct: distribution and developmental regulation. Cell Growth Differ. 1996;7(7):945–52.
5. Naylor MJ, Lockefeer JA, Horseman ND, Ormandy CJ. Prolactin regulates mammary epithelial cell proliferation via autocrine/paracrine mechanism. Endocrine. 2003;20(1–2):111–4. https://doi.org/10.1385/ENDO:20:1-2:111.
6. Hennighausen L, Robinson GW. Interpretation of cytokine signaling through the transcription factors STAT5A and STAT5B. Genes Dev. 2008;22(6):711–21. https://doi.org/10.1101/gad.1643908.
7. Azam M, Erdjument-Bromage H, Kreider BL, Xia M, Quelle F, Basu R, et al. Interleukin-3 signals through multiple isoforms of Stat5. EMBO J. 1995;14(7):1402–11.
8. Liu X, Robinson GW, Gouilleux F, Groner B, Hennighausen L. Cloning and expression of Stat5 and an additional homologue (Stat5b) involved in prolactin signal transduction in mouse mammary tissue. Proc Natl Acad Sci U S A. 1995;92(19):8831–5.
9. Mui AL, Wakao H, O'Farrell AM, Harada N, Miyajima A. Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin-5 transduce signals through two STAT5 homologs. EMBO J. 1995;14(6):1166–75.
10. Cui Y, Riedlinger G, Miyoshi K, Tang W, Li C, Deng CX, et al. Inactivation of Stat5 in mouse mammary epithelium during pregnancy reveals distinct functions in cell proliferation, survival, and differentiation. Mol Cell Biol. 2004;24(18):8037–47. https://doi.org/10.1128/MCB.24.18.8037-8047.2004.
11. Liu X, Robinson GW, Wagner KU, Garrett L, Wynshaw-Boris A, Hennighausen L. Stat5a is mandatory for adult mammary gland development and lactogenesis. Genes Dev. 1997;11(2):179–86.
12. Liu X, Gallego MI, Smith GH, Robinson GW, Hennighausen L. Functional rescue of Stat5a-null mammary tissue through the activation of compensating signals including Stat5b. Cell Growth Differ. 1998;9(9):795–803.
13. Udy GB, Towers RP, Snell RG, Wilkins RJ, Park SH, Ram PA, et al. Requirement of STAT5b for sexual dimorphism of body growth rates and liver gene expression. Proc Natl Acad Sci U S A. 1997;94(14):7239–44.
14. Imada K, Bloom ET, Nakajima H, Horvath-Arcidiacono JA, Udy GB, Davey HW, et al. Stat5b is essential for natural killer cell-mediated proliferation and cytolytic activity. J Exp Med. 1998;188(11):2067–74.
15. Yao Z, Cui Y, Watford WT, Bream JH, Yamaoka K, Hissong BD, et al. Stat5a/b are essential for normal lymphoid development and differentiation. Proc Natl Acad Sci U S A. 2006;103(4):1000–5. https://doi.org/10.1073/pnas.0507350103.
16. Miyoshi K, Shillingford JM, Smith GH, Grimm SL, Wagner KU, Oka T, et al. Signal transducer and activator of transcription (Stat) 5 controls the proliferation and differentiation of mammary alveolar epithelium. J Cell Biol. 2001;155(4):531–42. https://doi.org/10.1083/jcb.200107065.
17. Yamaji D, Kang K, Robinson GW, Hennighausen L. Sequential activation of genetic programs in mouse mammary epithelium during pregnancy depends on STAT5A/B concentration. Nucleic Acids Res. 2013;41(3):1622–36. https://doi.org/10.1093/nar/gks1310.
18. Ong CT, Corces VG. Enhancer function: new insights into the regulation of tissue-specific gene expression. Nat Rev Genet. 2011;12(4):283–93. https://doi.org/10.1038/nrg2957.
19. Ong CT, Corces VG. Enhancers: emerging roles in cell fate specification. EMBO Rep. 2012;13(5):423–30. https://doi.org/10.1038/embor.2012.52.
20. Shlyueva D, Stampfel G, Stark A. Transcriptional enhancers: from properties to genome-wide predictions. Nat Rev Genet. 2014;15(4):272–86. https://doi.org/10.1038/nrg3682.
21. Natoli G, Andrau JC. Noncoding transcription at enhancers: general principles and functional models. Annu Rev Genet. 2012;46:1–19. https://doi.org/10.1146/annurev-genet-110711-155459.
22. Metser G, Shin HY, Wang C, Yoo KH, Oh S, Villarino AV, et al. An autoregulatory enhancer controls mammary-specific STAT5 functions. Nucleic Acids Res. 2016;44(3):1052–63. https://doi.org/10.1093/nar/gkv999.
23. Shin HY, Willi M, HyunYoo K, Zeng X, Wang C, Metser G, et al. Hierarchy within the mammary STAT5-driven Wap super-enhancer. Nat Genet. 2016;48(8):904–11. https://doi.org/10.1038/ng.3606.
24. Zeng X, Willi M, Shin HY, Hennighausen L, Wang C. Lineage-specific and non-specific cytokine-sensing genes respond differentially to the master regulator STAT5. Cell Rep. 2016;17(12):3333–46. https://doi.org/10.1016/j.celrep.2016.11.079.
25. Miyoshi K, Cui Y, Riedlinger G, Robinson P, Lehoczky J, Zon L, et al. Structure of the mouse Stat 3/5 locus: evolution from Drosophila to zebrafish to mouse. Genomics. 2001;71(2):150–5. https://doi.org/10.1006/geno.2000.6433.
26. Pittius CW, Sankaran L, Topper YJ, Hennighausen L. Comparison of the regulation of the whey acidic protein gene with that of a hybrid gene containing the whey acidic protein gene promoter in transgenic mice. Mol Endocrinol. 1988;2(11):1027–32. https://doi.org/10.1210/mend-2-11-1027.
27. Robinson GW, Kang K, Yoo KH, Tang Y, Zhu BM, Yamaji D, et al. Coregulation of genetic programs by the transcription factors NFIB and STAT5. Mol Endocrinol. 2014;28(5):758–67. https://doi.org/10.1210/me.2012-1387.
28. Zhou J, Chehab R, Tkalcevic J, Naylor MJ, Harris J, Wilson TJ, et al. Elf5 is essential for early embryogenesis and mammary gland development during pregnancy and lactation. EMBO J. 2005;24(3):635–44. https://doi.org/10.1038/sj.emboj.7600538.

29. Hennighausen LG, Sippel AE. Mouse whey acidic protein is a novel member of the family of 'four-disulfide core' proteins. Nucleic Acids Res. 1982;10(8):2677–84.
30. Harris J, Stanford PM, Sutherland K, Oakes SR, Naylor MJ, Robertson FG, et al. Socs2 and elf5 mediate prolactin-induced mammary gland development. Mol Endocrinol. 2006;20(5):1177–87. https://doi.org/10.1210/me.2005-0473.
31. Chapman RS, Lourenco PC, Tonner E, Flint DJ, Selbert S, Takeda K, et al. Suppression of epithelial apoptosis and delayed mammary gland involution in mice with a conditional knockout of Stat3. Genes Dev. 1999;13(19):2604–16.
32. Humphreys RC, Bierie B, Zhao L, Raz R, Levy D, Hennighausen L. Deletion of Stat3 blocks mammary gland involution and extends functional competence of the secretory epithelium in the absence of lactogenic stimuli. Endocrinology. 2002;143(9):3641–50. https://doi.org/10.1210/en.2002-220224.
33. Zhao L, Melenhorst JJ, Hennighausen L. Loss of interleukin 6 results in delayed mammary gland involution: a possible role for mitogen-activated protein kinase and not signal transducer and activator of transcription 3. Mol Endocrinol. 2002;16(12):2902–12. https://doi.org/10.1210/me.2001-0330.
34. Willi M, Yoo KH, Wang C, Trajanoski Z, Hennighausen L. Differential cytokine sensitivities of STAT5-dependent enhancers rely on Stat5 autoregulation. Nucleic Acids Res. 2016;44(21):10277–91. https://doi.org/10.1093/nar/gkw844.
35. Baik M, Yu JH, Hennighausen L. Growth hormone-STAT5 regulation of growth, hepatocellular carcinoma, and liver metabolism. Ann N Y Acad Sci. 2011;1229:29–37. https://doi.org/10.1111/j.1749-6632.2011.06100.x.
36. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell. 2013;153(2):307–19. https://doi.org/10.1016/j.cell.2013.03.035.
37. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, et al. Super-enhancers in the control of cell identity and disease. Cell. 2013;155(4):934–47. https://doi.org/10.1016/j.cell.2013.09.053.
38. Parker SC, Stitzel ML, Taylor DL, Orozco JM, Erdos MR, Akiyama JA, et al. Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. Proc Natl Acad Sci U S A. 2013;110(44):17921–6. https://doi.org/10.1073/pnas.1317023110.
39. Vahedi G, Kanno Y, Furumoto Y, Jiang K, Parker SC, Erdos MR, et al. Super-enhancers delineate disease-associated regulatory nodes in T cells. Nature. 2015;520(7548):558–62. https://doi.org/10.1038/nature14154.
40. Adam RC, Yang H, Rockowitz S, Larsen SB, Nikolova M, Oristian DS, et al. Pioneer factors govern super-enhancer dynamics in stem cell plasticity and lineage choice. Nature. 2015;521(7552):366–70. https://doi.org/10.1038/nature14289.
41. Liu CF, Lefebvre V. The transcription factors SOX9 and SOX5/SOX6 cooperate genome-wide through super-enhancers to drive chondrogenesis. Nucleic Acids Res. 2015;43(17):8183–203. https://doi.org/10.1093/nar/gkv688.
42. Yin JW, Wang G. The mediator complex: a master coordinator of transcription and cell lineage development. Development. 2014;141(5):977–87. https://doi.org/10.1242/dev.098392.
43. Pott S, Lieb JD. What are super-enhancers? Nat Genet. 2015;47(1):8–12. https://doi.org/10.1038/ng.3167.
44. Gaj T, Gersbach CA, Barbas CF 3rd. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol. 2013;31(7):397–405. https://doi.org/10.1016/j.tibtech.2013.04.004.
45. Li M, Liu X, Robinson G, Bar-Peled U, Wagner KU, Young WS, et al. Mammary-derived signals activate programmed cell death during the first stage of mammary gland involution. Proc Natl Acad Sci U S A. 1997;94(7):3425–30.
46. Watson CJ, Neoh K. The Stat family of transcription factors have diverse roles in mammary gland development. Semin Cell Dev Biol. 2008;19(4):401–6. https://doi.org/10.1016/j.semcdb.2008.07.021.
47. Walker SR, Nelson EA, Yeh JE, Pinello L, Yuan GC, Frank DA. STAT5 outcompetes STAT3 to regulate the expression of the oncogenic transcriptional modulator BCL6. Mol Cell Biol. 2013;33(15):2879–90. https://doi.org/10.1128/MCB.01620-12.
48. de Laat W, Duboule D. Topology of mammalian developmental enhancers and their regulatory landscapes. Nature. 2013;502(7472):499–506. https://doi.org/10.1038/nature12753.
49. Bell AC, West AG, Felsenfeld G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell. 1999;98(3):387–96.
50. Yusufzai TM, Tagami H, Nakatani Y, Felsenfeld G. CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. Mol Cell. 2004;13(2):291–8.
51. Gaszner M, Felsenfeld G. Insulators: exploiting transcriptional and epigenetic mechanisms. Nat Rev Genet. 2006;7(9):703–13. https://doi.org/10.1038/nrg1925.
52. Dowen JM, Fan ZP, Hnisz D, Ren G, Abraham BJ, Zhang LN, et al. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. Cell. 2014;159(2):374–87. https://doi.org/10.1016/j.cell.2014.09.030.
53. de Wit E, Vos ES, Holwerda SJ, Valdes-Quezada C, Verstegen MJ, Teunissen H, et al. CTCF binding polarity determines chromatin looping. Mol Cell. 2015;60(4):676–84. https://doi.org/10.1016/j.molcel.2015.09.023.
54. Blinka S, Reimer MH Jr, Pulakanti K, Rao S. Super-enhancers at the Nanog locus differentially regulate neighboring pluripotency-associated genes. Cell Rep. 2016;17(1):19–28. https://doi.org/10.1016/j.celrep.2016.09.002.
55. Narendra V, Rocha PP, An D, Raviram R, Skok JA, Mazzoni EO, et al. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. Science. 2015;347(6225):1017–21. https://doi.org/10.1126/science.1262088.
56. Willi M, Yoo KH, Reinisch F, Kuhns TM, Lee HK, Wang C, et al. Facultative CTCF sites moderate mammary super-enhancer activity and regulate juxtaposed gene in non-mammary cells. Nat Commun. 2017;8:16069. https://doi.org/10.1038/ncomms16069.
57. Lee HK, Willi M, Wang C, Yang CM, Smith HE, Liu C, et al. Functional assessment of CTCF sites at cytokine-sensing mammary enhancers using CRISPR/Cas9 gene editing in mice. Nucleic Acids Res. 2017;45(8):4606–18. https://doi.org/10.1093/nar/gkx185.
58. Wyszomierski SL, Rosen JM. Cooperative effects of STAT5 (signal transducer and activator of transcription 5) and C/EBPbeta (CCAAT/enhancer-binding protein-beta) on beta-casein gene transcription are mediated by the glucocorticoid receptor. Mol Endocrinol. 2001;15(2):228–40. https://doi.org/10.1210/mend.15.2.0597.
59. Kabotyanski EB, Huetter M, Xian W, Rijnkels M, Rosen JM. Integration of prolactin and glucocorticoid signaling at the beta-casein promoter and enhancer by ordered recruitment of specific transcription factors and chromatin modifiers. Mol Endocrinol. 2006;20(10):2355–68. https://doi.org/10.1210/me.2006-0160.
60. Kabotyanski EB, Rijnkels M, Freeman-Zadrowski C, Buser AC, Edwards DP, Rosen JM. Lactogenic hormonal induction of long distance interactions between beta-casein gene regulatory elements. J Biol Chem. 2009;284(34):22815–24. https://doi.org/10.1074/jbc.M109.032490.
61. Rijnkels M, Kabotyanski E, Shore A, Rosen JM. The chromatin landscape of the casein gene locus. Horm Mol Biol Clin Investig. 2012;10(1):201–5. https://doi.org/10.1515/hmbci-2012-0004.
62. Moriguchi T, Suzuki M, Yu L, Takai J, Ohneda K, Yamamoto M. Progenitor stage-specific activity of a cis-acting double GATA

motif for Gata1 gene expression. Mol Cell Biol. 2015;35(5):805–15. https://doi.org/10.1128/MCB.01011-14.
63. Leddin M, Perrod C, Hoogenkamp M, Ghani S, Assi S, Heinz S, et al. Two distinct auto-regulatory loops operate at the PU.1 locus in B cells and myeloid cells. Blood. 2011;117(10):2827–38. https://doi.org/10.1182/blood-2010-08-302976.
64. Ohba S, He X, Hojo H, McMahon AP. Distinct transcriptional programs underlie Sox9 regulation of the mammalian chondrocyte. Cell Rep. 2015;12(2):229–43. https://doi.org/10.1016/j.celrep.2015.06.013.
65. Neville MC, McFadden TB, Forsyth I. Hormonal regulation of mammary differentiation and milk secretion. J Mammary Gland Biol Neoplasia. 2002;7(1):49–66.
66. Robinson GW, McKnight RA, Smith GH, Hennighausen L. Mammary epithelial cells undergo secretory differentiation in cycling virgins but require pregnancy for the establishment of terminal differentiation. Development. 1995;121(7):2079–90.
67. Bach K, Pensa S, Grzelak M, Hadfield J, Adams DJ, Marioni JC, et al. Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing. Nat Commun. 2017;8(1):2128. https://doi.org/10.1038/s41467-017-02001-5.
68. Nguyen QH, Pervolarakis N, Blake K, Ma D, Davis RT, James N, et al. Profiling human breast epithelial cells using single cell RNA sequencing identifies cell diversity. Nat Commun. 2018;9(1):2028. https://doi.org/10.1038/s41467-018-04334-1.
69. Sun H, Miao Z, Zhang X, Chan UI, Su SM, Guo S, et al. Single-cell RNA-Seq reveals cell heterogeneity and hierarchy within mouse mammary epithelia. J Biol Chem. 2018;293(22):8315–29. https://doi.org/10.1074/jbc.RA118.002297.
70. Giraddi RR, Chung CY, Heinz RE, Balcioglu O, Novotny M, Trejo CL, et al. Single-cell transcriptomes distinguish stem cell state changes and lineage specification programs in early mammary gland development. Cell Rep. 2018;24(6):1653–66 e7. https://doi.org/10.1016/j.celrep.2018.07.025.