10796–10809 *Nucleic Acids Research, 2018, Vol. 46, No. 20* *Published online 4 October 2018*
doi: 10.1093/nar/gky891

Progressing super-enhancer landscape during mammary differentiation controls tissue-specific gene regulation

Hye Kyung Lee[1,*,†], Michaela Willi[1,†], Ha Youn Shin[2], Chengyu Liu[3] and Lothar Hennighausen[1,*]

[1]Laboratory of Genetics and Physiology, National Institute of Diabetes and Digestive and Kidney Diseases, US National Institutes of Health, Bethesda, MD 20892, USA, [2]Department of Biomedical Science and Engineering, Konkuk University, Seoul 05029, Republic of Korea and [3]Transgenic Core, National Heart, Lung, and Blood Institute, US National Institutes of Health, Bethesda, MD 20892, USA

Received July 17, 2018; Revised September 18, 2018; Editorial Decision September 19, 2018; Accepted September 20, 2018

ABSTRACT

The mammary luminal lineage relies on the common cytokine-sensing transcription factor STAT5 to establish super-enhancers during pregnancy and initiate a genetic program that activates milk production. As pups grow, the greatly increasing demand for milk requires progressive differentiation of mammary cells with advancing lactation. Here we investigate how persistent hormonal exposure during lactation shapes an evolving enhancer landscape and impacts the biology of mammary cells. Employing ChIP-seq, we uncover a changing transcription factor occupancy at mammary enhancers, suggesting that their activities evolve with advancing differentiation. Using mouse genetics, we demonstrate that the functions of individual enhancers within the *Wap* super-enhancer evolve as lactation progresses. Most profoundly, a seed enhancer, which is mandatory for the activation of the *Wap* super-enhancer during pregnancy, is not required during lactation, suggesting compensatory flexibility. Combinatorial deletions of structurally equivalent constituent enhancers demonstrated differentiation-specific compensatory activities during lactation. We also demonstrate that the *Wap* super-enhancer, which is built on STAT5 and other common transcription factors, retains its exquisite mammary specificity when placed into globally permissive chromatin, suggesting a limited role of chromatin in controlling cell specificity. Our studies unveil a previously unrecognized progressive enhancer landscape where structurally equivalent components serve unique and differentiation-specific functions.

INTRODUCTION

A major unresolved question is how the complexity of super-enhancers controls spatiotemporal regulation of lineage-specific programs in mammals. To gain insight into these biological systems, we continue to exploit the mammary gland, an organ characterized by extensive cell proliferation and differentiation during pregnancy. Mammary gland epithelium is composed of the basal myoepithelial lineage and milk-secreting luminal alveolar cells (1), which feature temporal regulatory dynamics during development. Differentiation of the luminal lineage is controlled by the cytokine prolactin (2,3) and advances during pregnancy and throughout lactation (4), resulting in the activation of mammary-specific genes (5). While differentiation of milk-secreting mammary epithelium is initiated during pregnancy, it is firmly established as lactation advances and accommodates changing needs of growing pups. However, it has not been investigated how daily milk production is regulated during progressing lactation.

It is established that cytokines control cell lineages through common transcription factors from the signal transducer and activator of transcription (STAT) family (6–10). Our studies (11,12) have demonstrated the importance of the transcription factors STAT5A and STAT5B (referred to as STAT5) in regulating mammary-specific genes during pregnancy and lactation (13,14). STAT5 is at the core of 440 mammary-specific super-enhancers, which are exclusively identified on the basis of chromatin immunoprecipitation sequencing (ChIP-seq) for STAT5, glucocorticoid receptor (GR), mediator complex subunit 1 (MED1) and acetylation of Histone 3 Lysine 27 (H3K27ac) (12), thereby activat-

*To whom correspondence should be addressed. Tel: +1 301 328 6183; Fax: +1 301 480 7312; Email: lotharh@mail.nih.gov
Correspondence may also be addressed to Hye Kyung Lee. Tel: +1 301 435 6635; Fax: +1 301 480 7312; Email: hyekyung.lee@nih.gov
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

ing mammary genes up to 1000-fold during pregnancy (11). Mammary enhancers have high affinities for STAT5, which could be an explanation for their unique tissue-specific activity. Genes such as *Csn1s1* and *Csn2* require low levels of STAT5 for their activation, while Whey acidic protein (*Wap*) and *Csn1s2b* require high levels (11). The mammary enhancer landscape has been investigated during pregnancy (12), but it is unknown whether it evolves further as mammary cells meet the increased demand of milk during progressing lactation.

Frequently, constituent enhancers within tissue-specific super-enhancers that are associated with developmentally important genes display redundant functions (15–18), suggesting mechanisms aimed at preventing deleterious consequences in case of enhancer mutations. Mutations of individual enhancers within super-enhancers in mice have demonstrated additive effects on gene expression (15–18). To gain insight into the biological significance of super-enhancers and their constituent enhancers controlling mammary-specific genes at different hormonal levels during lactation, we have focused on the well-characterized *Wap* super-enhancer locus. The *Wap* gene is activated more than 1000-fold by the tripartite super-enhancer (12), which is primarily built on the prolactin-sensing transcription factor STAT5 and other mammary-enriched transcription factors (11,12). Of the three individual enhancers with structurally similar chromatin features, the distal enhancer (E3) is the most influential one, while the other two play less prominent roles (12). However, only the deletion of all three enhancers completely ablates *Wap* expression. The more complex proximal enhancer (E1) contains independent binding sites for STAT5 and the transcription factors E74 like ETS transcription factor 5 (ELF5) and Nuclear factor 1B (NFIB). Disruption of all three binding sites results in the complete failure to establish the super-enhancer (12), demonstrating a functional hierarchy among hormone-sensing enhancers. However, it remains unknown to what extent individual enhancers control gene expression as mammary tissue progresses from pregnancy to fully established lactation.

Here, we asked whether progressing differentiation from pregnancy to advanced lactation leads to changes in the mammary enhancer landscape and a differential contribution of super-enhancer elements, which might support increased milk production. For this, we investigated the enhancer landscape at the peak of lactation and used mutant mice to explore the contribution of individual enhancers. Specifically, we focus on the *Wap* super-enhancer, which contains three constituent enhancers that are built on the master regulator STAT5A, GR and MED1 (12). The differential use of constituent enhancers within the *Wap* super-enhancer during pregnancy (12) makes it an attractive model to further explore whether progressive differentiation during lactation impacts their respective contributions.

MATERIALS AND METHODS

Mice

Mice carrying mutations in the *Wap* super-enhancer (ΔE1, ΔE3 and ΔE1a/2/3) have been described previously (12). All animals were housed and handled according to the guidelines of the National Institutes of Health (NIH) and all animal experiments were performed in accordance with the Animal Care and Use Committee (ACUC) of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). CRISPR/Cas9-targeted mice were generated with C57BL/6 mice (Charles River) by the transgenic core of the National Heart, Lung, and Blood Institute (NHLBI). Single-guide RNAs (sgRNAs) were designed using MIT's CRISPR Design tool (http://crispr.mit.edu/) and synthesized (OriGene and Thermo Fisher Scientific). Target-specific sgRNAs and *in vitro* transcribed *Cas9* messenger RNA (mRNA) were co-microinjected into the cytoplasm of fertilized eggs for founder mouse production. The ΔE1a/3 mutant mouse was generated by injecting a sgRNA for E1 GAS motif (Supplementary Table S8) into zygotes from mice carrying a deletion in E3 (ΔE3 mutant mice) (12). E12-R3 and E2-R3 mutant mice were generated by injecting sgRNAs in wild-type (WT) and ΔE1a/3 mutant mice, respectively (Supplementary Table S8). To screen for homozygous mice, all mice were genotyped by polymerase chain reaction (PCR) amplification and Sanger sequencing (Macrogen and Quintara Biosciences) with genomic DNA from mouse tails (Supplementary Table S9).

Chromatin immunoprecipitation sequencing (ChIP-seq) and data analysis

Mammary tissues from specific stages during lactation as well as kidney tissues were harvested, and stored at −80°C. The frozen-stored tissues were ground into powder in liquid nitrogen. Chromatin was fixed with formaldehyde (1% final concentration) for 15 min at room temperature and then quenched with glycine (0.125 M final concentration). Samples were processed as previously described (19). The following antibodies were used for ChIP-seq: STAT5A (Santa Cruz Biotechnology, sc-1081 and sc-271542), total STAT5 (Santa Cruz Biotechnology, sc-74442), STAT3 (Santa Cruz Biotechnology, sc-482), GR (Thermo Fisher Scientific, PA1-511A), H3K27ac (Abcam, ab4729), MED1 (Bethyl Laboratory, A300-793A) and RNA polymerase II (Abcam, ab5408). Libraries for next-generation sequencing were prepared and sequenced with an HiSeq 2500 instrument (Illumina) (19). Quality filtering was done using trimmomatic (20) (version 0.36) and Bowtie (21) (version 1.1.2) with the reference genome mm10 was applied to uniquely map the reads. Picard tools (Broad Institute. Picard, http://broadinstitute.github.io/picard/. 2016) was used to remove duplicates and subsequently, Homer (22) (version 4.8.2) software was applied to generate bedGraph files. Integrative Genomics Viewer (23) (version 2.3.81) was used for visualization. Each ChIP-seq experiment was conducted for two replicates.

Total RNA sequencing (total RNA-seq) and data analysis

Total RNA was extracted from frozen mammary tissue from WT mice at day 1 and 10 of lactation and purified with RNeasy Plus Mini Kit (Qiagen, 74134). Ribosomal RNA was removed from 1 μg of total RNAs and complementary DNA was synthesized using SuperScript III (Invitrogen). Libraries for sequencing were prepared according to

the manufacturer's instructions with TruSeq Stranded Total RNA Library Prep Kit with Ribo-Zero Gold (Illumina, RS-122-2301) and paired-end sequencing was done with an HiSeq 2500 instrument (Illumina). Quality control and alignment of total RNA-seq was done using trimmomatic (20) (version 0.36) and STAR RNA-seq (24) (version STAR 2.5.3a) using paired-end mode with the reference genome mm10. Subsequently, R (https://www.R-project.org/) and Bioconductor (25) were used. The R package Rsubread (26) was applied to generate the input files for DESeq2 (27) to conduct the RNA-seq analysis. The data were pre-filtered keeping only those genes that have at least 10 reads in total. Genes were categorized as significantly differentially expressed with an adjusted *P*-value below 0.05 and a fold change higher than 2 for upregulated genes and a fold change of smaller than −2 for downregulated ones. Normalized read counts are shown for each gene. The visualization was done using dplyr (https://CRAN.R-project.org/package=dplyr) and ggplot2 (28).

RNA isolation and quantitative real-time PCR (qRT-PCR)

Total RNA was extracted from frozen mammary tissue of WT and mutant mice using a homogenizer and the PureLink RNA Mini kit according to the manufacturer's instructions (Invitrogen). Total RNA (1 μg) was reverse transcribed for 50 min at 50°C using 50 μM oligo dT and 2 μl of SuperScript III (Invitrogen) in a 20 μl reaction. Quantitative real-time PCR (qRT-PCR) was performed using TaqMan probes (*Wap*, Mm00839913_m1; *Ramp3,* Mm00840142_m1; mouse *Gapdh*, Mm99999915_g1, Thermo Fisher Scientific) on the CFX384 Real-Time PCR Detection System (Bio-Rad) according to the manufacturer's instructions. PCR conditions were 95°C for 30 s, 95°C for 15 s and 60°C for 30 s for 40 cycles. All reactions were done in triplicate and normalized to the housekeeping gene *Gapdh*. Relative differences in PCR results were calculated using the comparative cycle threshold (C_T) method.

Statistical analyses

For comparison of samples, data were presented as standard deviation in each group and were evaluated with a *t*-test and two-way ANOVA multiple comparisons using PRISM GraphPad. Statistical significance was obtained by comparing the measures from WT or control group, and each mutant group. A value of $*P < 0.05$, $**P < 0.001$, $***P < 0.0001$, $****P < 0.00001$ was considered statistically significant.

Enhancer and super-enhancer analysis

MACS2 (29) peak finding algorithm was used to identify regions of ChIP-seq enrichment over the background and promoter elements were subtracted to get enhancer elements. STAT5 enhancers were considered as true enhancer elements if they showed H3K27ac underneath. In order to receive only mammary-specific enhancer elements, binding sites also identified in liver or T cell were subtracted using bedtools (30). The R package DiffBind (31) was applied to categorize enhancer elements as L1 preferential, L10 preferential or L1/L10 shared. Super-enhancers were identified applying the ROSE algorithm (32,33). Therefore, each class of STAT5 enhancers was stitched using the default stitching size of 12.5 kb and combined with GR, MED1 and H3K27ac. The super-enhancers based on GR, MED1 and H3K27ac were overlapped and those which were identified at least in two of the three marks passed. Enhancers and super-enhancers were annotated to genes, taking into account the four closest up- and downstream genes and assigning the one with the highest expression level. L1 preferential super-enhancers were annotated using the mean of L1 gene expression, the L10 preferential super-enhancers were annotated using the mean of L10 gene expression and L1/L10 shared super-enhancers were annotated to both mean L1 and L10 gene expression. The used GTF file was filtered to remove predicted genes, comprising LOC, Riken and BC clones. Coverage plots (normalized to 10 million reads) were generated using Homer (22) software and R utilizing the packages dplyr and ggplot2 (28).

RESULTS

Evolving enhancer landscape and gene regulation during mammary differentiation

The average weight of a newborn pup is ∼1.4 g and increases more than 3.5-fold by day 10 of lactation (Supplementary Figure S1), a stage when daily milk production of dams is equivalent to their body weight (34). To begin to uncover how mammary differentiation progresses during lactation to meet the nutritional needs of the growing young, we focused on the impact of continued hormonal stimulation on the enhancer landscape and underlying changes in gene expression. Specifically, we investigated the progressive changes in the binding of the master transcription factor STAT5 at lactation days 1 (L1) and 10 (L10). Thereby, we uncovered the genome-wide presence of enhancer marks with preferred binding at L1 (L1 preferential), L10 (L10 preferential) and those being present throughout lactation (L1/L10 shared). We exclusively selected mammary-specific enhancer elements by removing STAT5 enhancer elements that were also identified in liver or T cells. A total of 1575 STAT5 enhancers were identified as L1 preferential, and 1573 as L10 preferential, whereas the majority of 11 466 STAT5 enhancers were shared between L1 and L10. The coverage plots showed that L1 preferential enhancers have strong STAT5 and GR binding at L1, but decreased binding at L10. Along with the transcription factor binding, acetylation of histone H3 at lysine 27 (H3K27ac) marks were strong at L1, but decreased at L10. In contrast, both the L10 preferential and the L1/L10 shared enhancers showed an increased STAT5 and GR binding between L1 and L10. The H3K27ac coverage increased between L1 and L10 in the L1/L10 shared enhancers, but an even stronger induction was observed at L10 preferential enhancers (Supplementary Figure S2).

In an earlier study, we identified 440 mammary-specific super-enhancers at L1 (12). Based on those findings, we used each of our three enhancer categories to identify super-enhancers specific for L1, L10 and those being active throughout lactation (L1/L10 shared). We calculated

super-enhancers in combination with GR, H3K27ac and MED1 for the STAT5 enhancer elements of each category separately (32,33). Using this approach, we unveiled a total of 591 super-enhancers, with 78 being L1 preferential (Supplementary Table S1), 90 L10 preferential (Supplementary Table S2) and the vast majority of 423 super-enhancers was shared between L1 and L10 (Supplementary Table S3). The STAT5 and GR coverage at L1 preferential constituent enhancers was decreased at L10, and H3K27ac marks are reduced even stronger, indicating the loss of enhancer activity (Figure 1A and Supplementary Figure S3A). In contrast, the L1/L10 shared enhancers as well as the L10 preferential ones gained STAT5 and GR binding, and elevated H3K27ac, confirming their increased activity at L10 (Figure 1B and C; Supplementary Figure S3B and C).

The changing transcription factor landscape and H3K27ac coverage suggested a biological impact on the associated genes, a question we addressed through RNA-seq analyses from mammary tissue at L1 and L10 (Supplementary Table S4). The annotation of the super-enhancers to RNA-seq expression data confirmed that genes annotated to shared L1/L10 enhancers were significantly induced from L1 to L10 (Figure 1D and Supplementary Tables S5–S7), including the mammary-specific genes *Wap* and *Csn1s2b*. Expression of genes assigned to L1 preferential super-enhancers is decreased, as exemplified by *Spp1*, which is known to be involved in mammary gland development and sensitive to STAT5 levels (11). Expression of genes that are associated to L10 preferential super-enhancers was induced significantly between L1 and L10, exemplified by *Csn1s1* (Figure 1E).

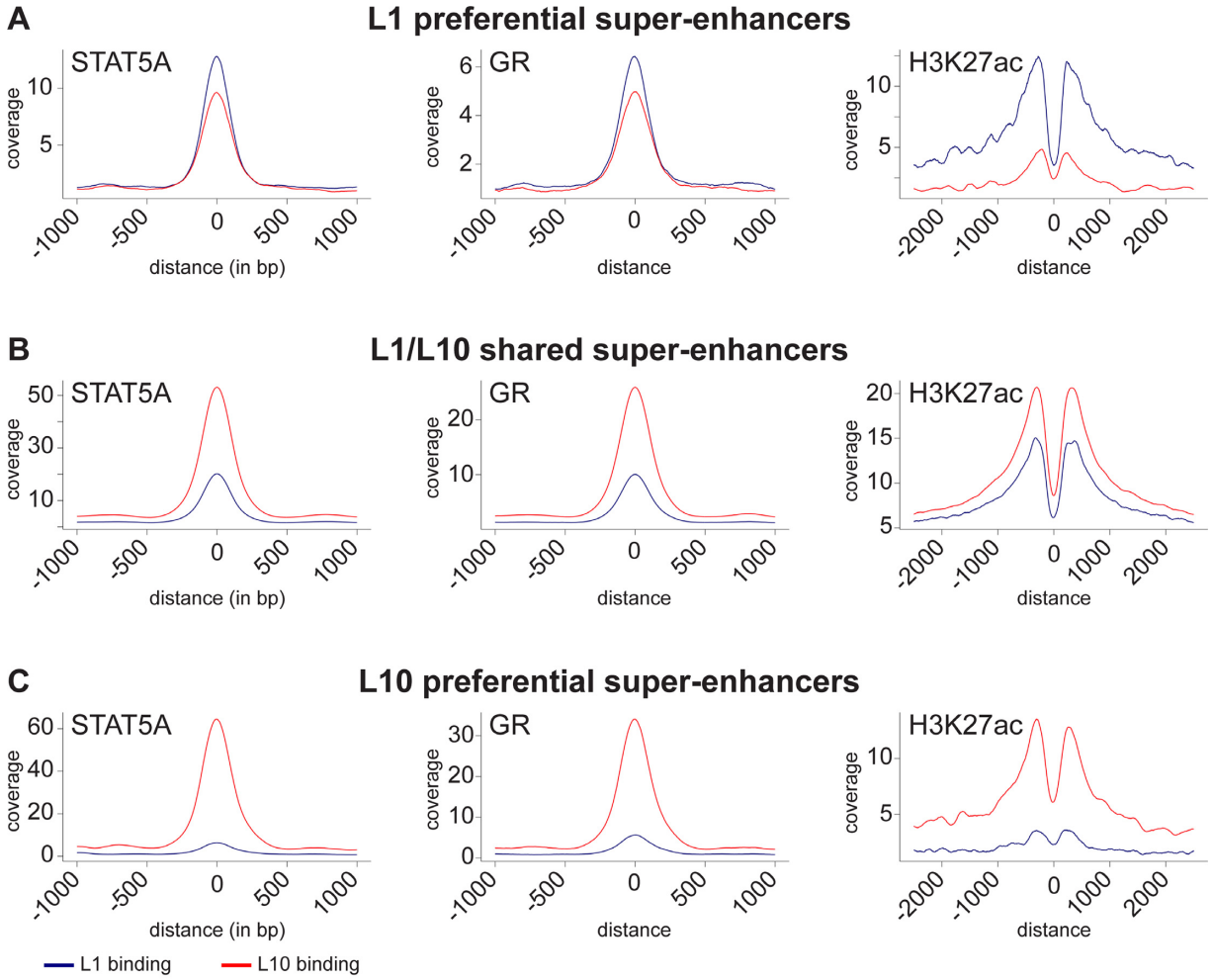
Among the 949 genes upregulated between L1 and L10 (increased more than 2-fold and an adjusted *P*-value smaller than 0.05), 124 were associated with super-enhancers (Figure 1E). Although overall the degree of elevated transcription factor coverage paralleled increased gene expression, it was not uniform among enhancers, suggesting that they responded uniquely to progressing differentiation (Figure 2). While STAT5 and GR binding at some enhancers increased >5-fold, others gained <2-fold. *Wap* and *Csn1s2b* expression increased up to 250-fold, which was accompanied by an ∼4–5-fold increase of STAT5 and GR coverage at its upstream enhancer (Figure 2A and B; Supplementary Table S4). In contrast, transcription factor loading at the three proximal enhancers of the *Csn2* locus increased between 2- and 3-fold and was accompanied by an ∼8-fold increased expression (Figure 2C). Concomitant increase of GR binding supports STAT5–GR interactions at tissue-specific enhancers and their potential cooperativity in the regulation of gene transcription (10,35–36). While at first glance it appeared that enhancers associated with a given locus display an equivalent response to lactation cues, our study provided new insight into unique responses. Notably, within the *Csn1s1* locus, transcription factor coverage of the distal enhancer increased 12-fold, while that of the proximal one <3-fold (Figure 2D). MED1 coverage at these enhancers paralleled that of transcription factors and Pol II loading paralleled expression levels (Supplementary Figure S4). While transcription factor occupancy increased at 1611 constituent enhancers of the L1/L10 shared and L10 preferential super-enhancers, a reduction was observed at 162 constituent enhancers of the L1 preferential ones, suggesting that they are partially decommissioned during differentiation. This is exemplified at the *Spp1* locus where transcription factor loading at the five enhancers was reduced by ∼60–80% (Figure 2E). This was paralleled by a loss of H3K27ac and a 93% reduction of *Spp1* mRNA levels (Figure 1E and Supplementary Table S4). Our results have revealed a dynamic transcription factor coverage of mammary enhancers along with gene expression changes, suggesting unique responses to a progressing differentiation program.

Seed enhancer is required during pregnancy but not lactation

Given the increased transcription factor occupancy at mammary enhancers with progressing lactation, we turned to investigating the possibility of temporal-specific enhancer activities. The *Wap* tripartite super-enhancer (Figure 3A) is initially established in mammary epithelium during pregnancy and deletion of three transcription factor binding sites within the proximal seed enhancer (ΔE1) had uncovered its critical role in the establishment of the two distal enhancers during pregnancy (12) (Figure 3B; Supplementary Figure S5 as a control locus; L1 panel) and the activation of *Wap* (12). STAT5 and GR coverage at the three constituent *Wap* enhancers (Figure 3A) increased 3- to 4-fold as lactation progressed (Figure 3B and C) and coincided with 10-fold higher *Wap* mRNA levels (Figure 3D). To explore the possibility that elevated transcription factor loading at L10 might void the functional importance of the proximal seed enhancer E1, we investigated E1 mutant mice at L1 and L10. Notably, in the absence of E1 (ΔE1), *Wap* expression increased ∼220-fold during lactation (Figure 3D), suggesting that the seed enhancer is not essential for the activation of the distal enhancers at advanced differentiation. ChIP-seq experiments conducted on ΔE1 mutant mammary tissue at L10 revealed STAT5 and GR binding at the two distal enhancers (Figure 3B and E). As expected, no significant GR and STAT5 binding was observed at the mutant seed enhancer (ΔE1). Acquisition of H3K27ac marks at the most distal enhancer (E3) in ΔE1 mutants is indicative of its activity at L10, albeit to a lower extent than the WT (Figure 3B; Supplementary Figure S5 as control locus). These findings pinpoint differentiation-specific activities of individual cytokine-sensing enhancers.

Differential contribution of individual *Wap* enhancers during lactation

While at first glance, the three constituent enhancers within the *Wap* super-enhancer display equivalent transcription factor occupancy (Figure 3B), their individual contribution in the activation of *Wap* is far from understood (12). Although E3 by itself is sufficient to activate the *Wap* gene (12), the ability of E2 to activate gene expression during lactation remained unknown. Addressing the functional importance of E2, we generated mice in which STAT5-binding sites within enhancers E1 and E3 were disrupted using CRISPR/Cas9 genome editing, leaving E2 intact (ΔE1a/3) (Figure 3A). *Wap* expression in ΔE1a/3 mutant mice at L1 was reduced by 99.8% (Figure 4A), similar to



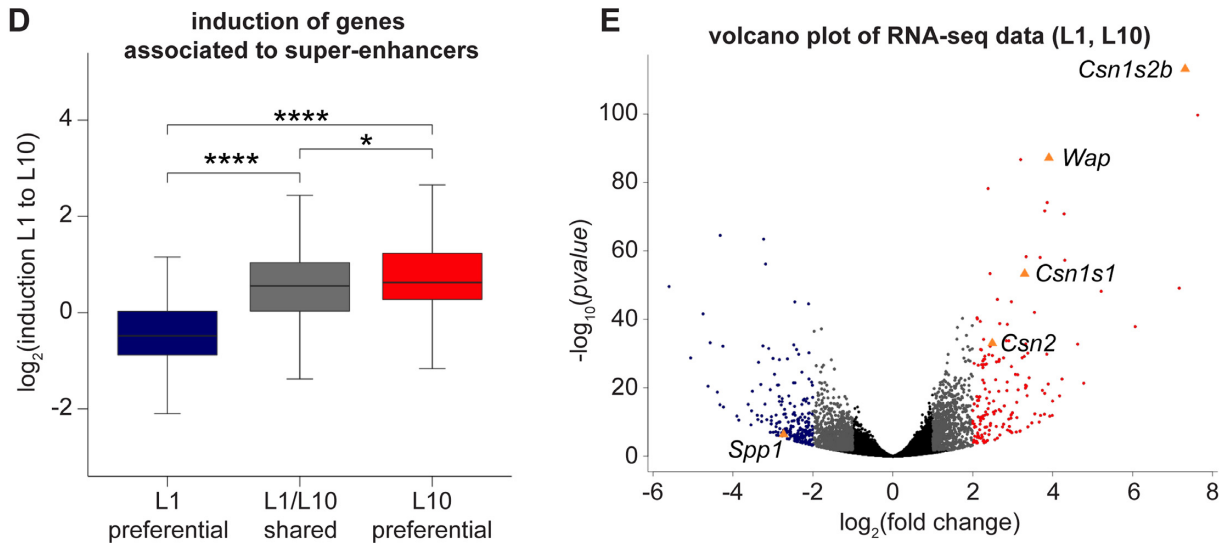


Figure 1. Genome-wide analyses of mammary-specific super-enhancers at L1 and L10. (**A**) Coverage plots of the mammary super-enhancers with preferential binding at L1. The coverage for STAT5, GR and H3K27ac decreased between L1 (blue) and L10 (red). (**B**) The binding of STAT5, GR and H3K27ac increased from L1 (blue) to L10 (red) for super-enhancers being active throughout lactation (L1/L10 shared). (**C**) L10 preferential super-enhancers showed increased binding of STAT5, GR and H3K27ac at L10 (red). (**D**) Genes associated to L1 preferential enhancers were not induced from L1 to L10, whereas genes that are assigned to L1/L10 shared and L10 preferential super-enhancers were significantly induced. Median, middle bar inside the box; IQR, 50% of the data; whiskers, 1.5 times the IQR. ANOVA with post-hoc Tukey HSD was used to evaluate statistical significance: $*P < 0.05$, $****P < 0.00001$. (**E**) Volcano plot of the L1 and L10 RNA-seq data (replicates L1 = 3; replicates L10 = 4).

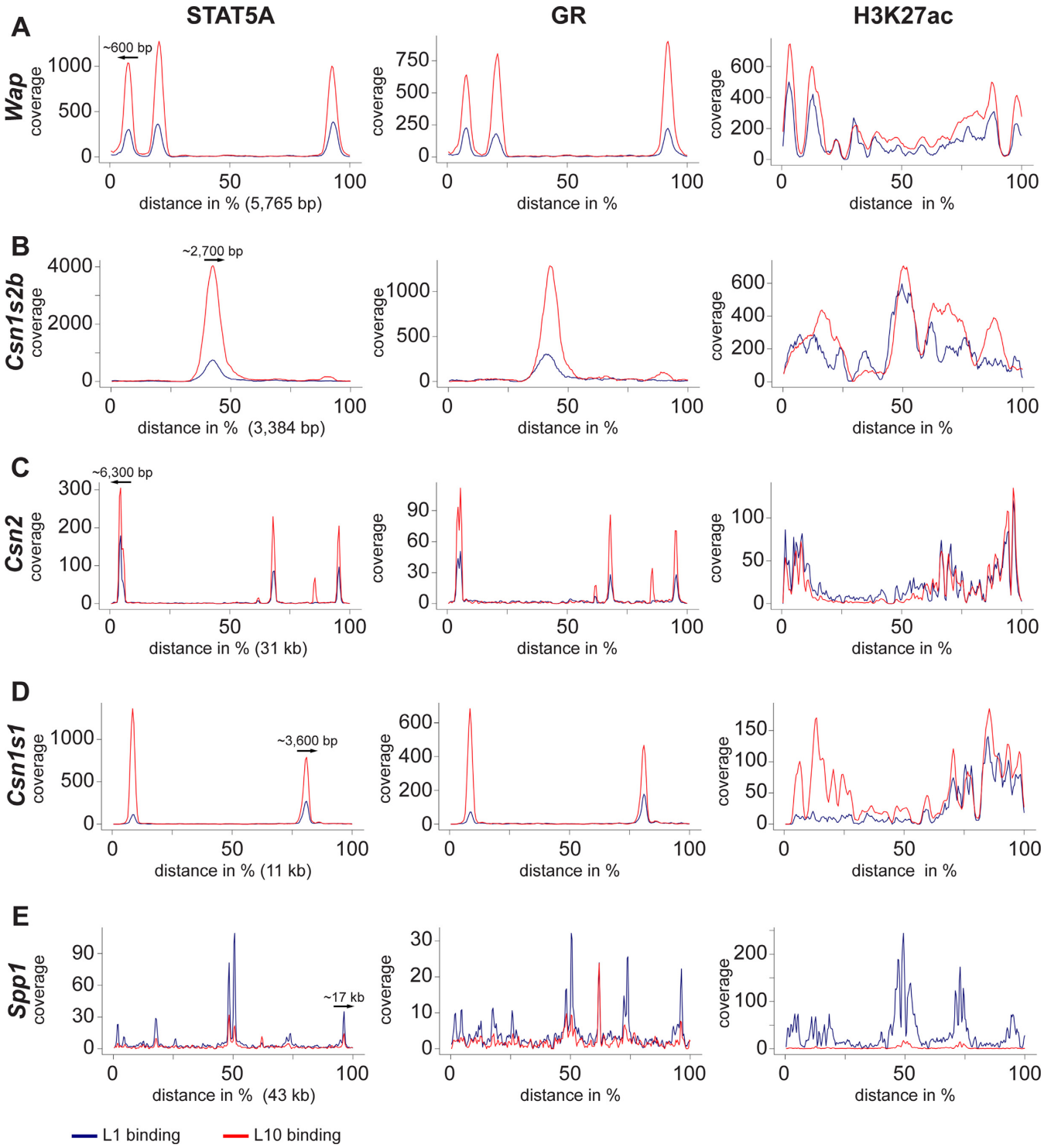


Figure 2. Genome-wide analyses of enhancer coverage in mammary tissue during lactation. (**A**) Coverage plots of the *Wap* locus (5765 bp) showing STAT5A, GR and H3K27ac at L1 (blue) and L10 (red). The distance of ∼600 bp to the *Wap* promoter is indicated by the arrow over the most proximal enhancer. (**B**) Coverage plots of the *Csn1s2b* locus (3384 bp) showing STAT5A, GR and H3K27ac at L1 (blue) and L10 (red). The distance of ∼2700 bp to the *Csn1s2b* promoter is indicated by the arrow over the most proximal enhancer. (**C**) Coverage plots of the *Csn2* locus (31 kb) showing STAT5A, GR and H3K27ac at L1 (blue) and L10 (red). The distance of ∼6300 bp to the *Csn2* promoter is indicated by the arrow over the most proximal enhancer. (**D**) Coverage plots of the *Csn1s1* locus (11 kb) showing STAT5A, GR and H3K27ac at L1 (blue) and L10 (red). The distance of ∼3600 bp to the *Csn1s1* promoter is indicated by the arrow over the most proximal enhancer. (**E**) Coverage plots of the *Spp1* locus (43 kb) showing STAT5A, GR and H3K27ac at L1 (blue) and L10 (red). The distance of ∼17 kb to the *Spp1* promoter is indicated by the arrow over the most proximal enhancer.

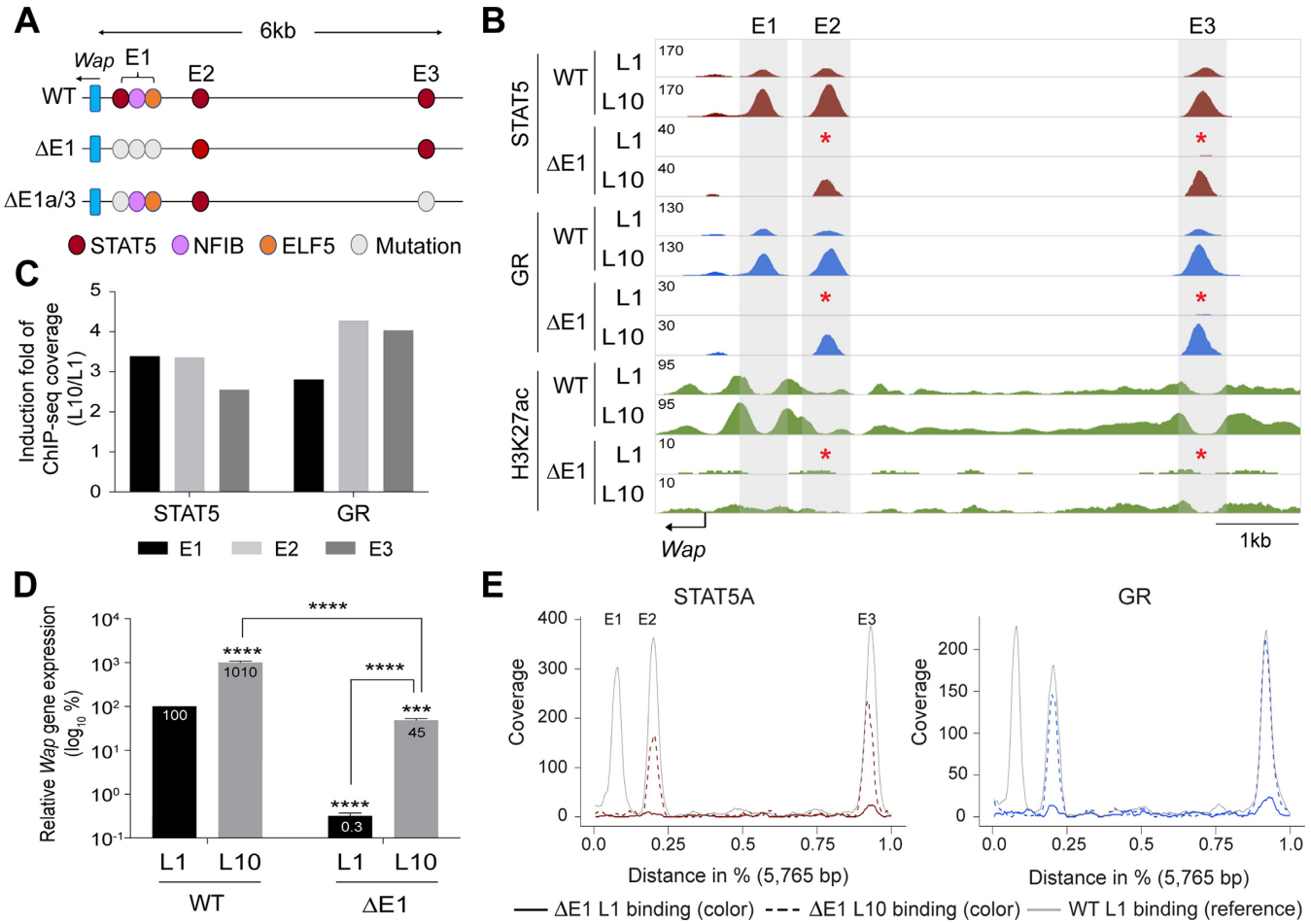


Figure 3. Requirement of the *Wap* seed enhancer during pregnancy but not lactation. (**A**) Diagram of the *Wap* enhancer region in WT and mutant mice. ΔE1 mice were generated by mutating the binding motifs for STAT5, NFIB and ELF5 (12). ΔE1a/3 mutant mice were generated by deleting STAT5-binding sites in E1 and E3 using CRISPR/Cas9 genome editing. The exact positions of the deletions are shown in Supplementary Table S9. (**B**) Genomic features of the *Wap* locus in lactating mammary tissues from WT and ΔE1 mutants. An asterisk indicates enhancer sites that showed prominent increase in STAT5A, GR and H3K27ac intensity in ΔE1 mutant tissue at L10. The *Lalba* locus was used as control (Supplementary Figure S5). (**C**) Induction of STAT5 and GR coverage at the three constituent *Wap* enhancers in WT tissue between L1 and L10. (**D**) *Wap* expression levels in mammary tissues from WT and ΔE1 mutant mice at L1 and L10 were measured by qRT-PCR and normalized to *Gapdh* levels. Results are shown as the means ± SEM of independent biological replicates (WT at L1, $n = 6$; WT at L10, $n = 12$; ΔE1 at L1 and L10, $n = 3$). ANOVA was used to evaluate the statistical significance of differences between L1 and L10 in WT and mutants. ***$P < 0.0001$, ****$P < 0.00001$. *Wap* expression levels between L1 and L10 were increased 10-fold in WT and 220-fold in ΔE1 mutants. (**E**) STAT5 and GR profiles at individual enhancers within the *Wap* super-enhancer in WT and ΔE1 mutant tissue based on the ChIP-seq data; red (STAT5), blue (GR), line (L1 of ΔE1), dotted line (L10 of ΔE1) and gray line (L1 of WT as a reference) (normalized to 10 million reads).

that observed in the absence of all three STAT5 enhancers (ΔE1a/2/3). Unexpectedly, E2 was unable to activate transcription despite being bound by transcription factors and located within active chromatin in the context of the intact super-enhancer. ChIP-seq analyses of ΔE1a/3 mutant tissue demonstrated greatly reduced binding of STAT5 and GR to E2 at L1 (Figure 4B; Supplementary Figure S6 as a control locus). This suggests that E2 by itself fails to adequately respond to hormonal cues during pregnancy. Having detected increased transcription factor binding at all three constituent enhancers within the *Wap* super-enhancer at L10, we asked whether E2 could be activated at L10 in the absence of E1 and E3 (ΔE1a/3). After 10 days of exposure to lactation hormones, *Wap* expression in ΔE1a/3 mutants increased 40-fold (Figure 4A), demonstrating activation of the otherwise silent E2 enhancer. Consistent with *Wap* activation, STAT5 and GR binding was now identified at E2 (Figure 4B and C). Although there was no significant STAT5 loading at the mutant E3 site, binding was observed at the mutant E1 site despite the deletion of the STAT5-binding motif (GAS). Since NFIB- and ELF5-binding sites are intact in the E1a mutant (12), it is possible that STAT5 is recruited through these transcription factors, or possibly the GR. H3K27ac was restored at E2 and the promoter region, suggesting that E2, possibly in conjunction with the promoter, responds to extended hormonal signaling and is sufficient to activate *Wap* expression, albeit not at normal levels. In contrast, *Wap* expression in the absence of all three enhancers (ΔE1a/2/3) was induced only 7-fold (Figure 4A), with little establishment of H3K27ac over the mutant regulatory regions (data not shown).

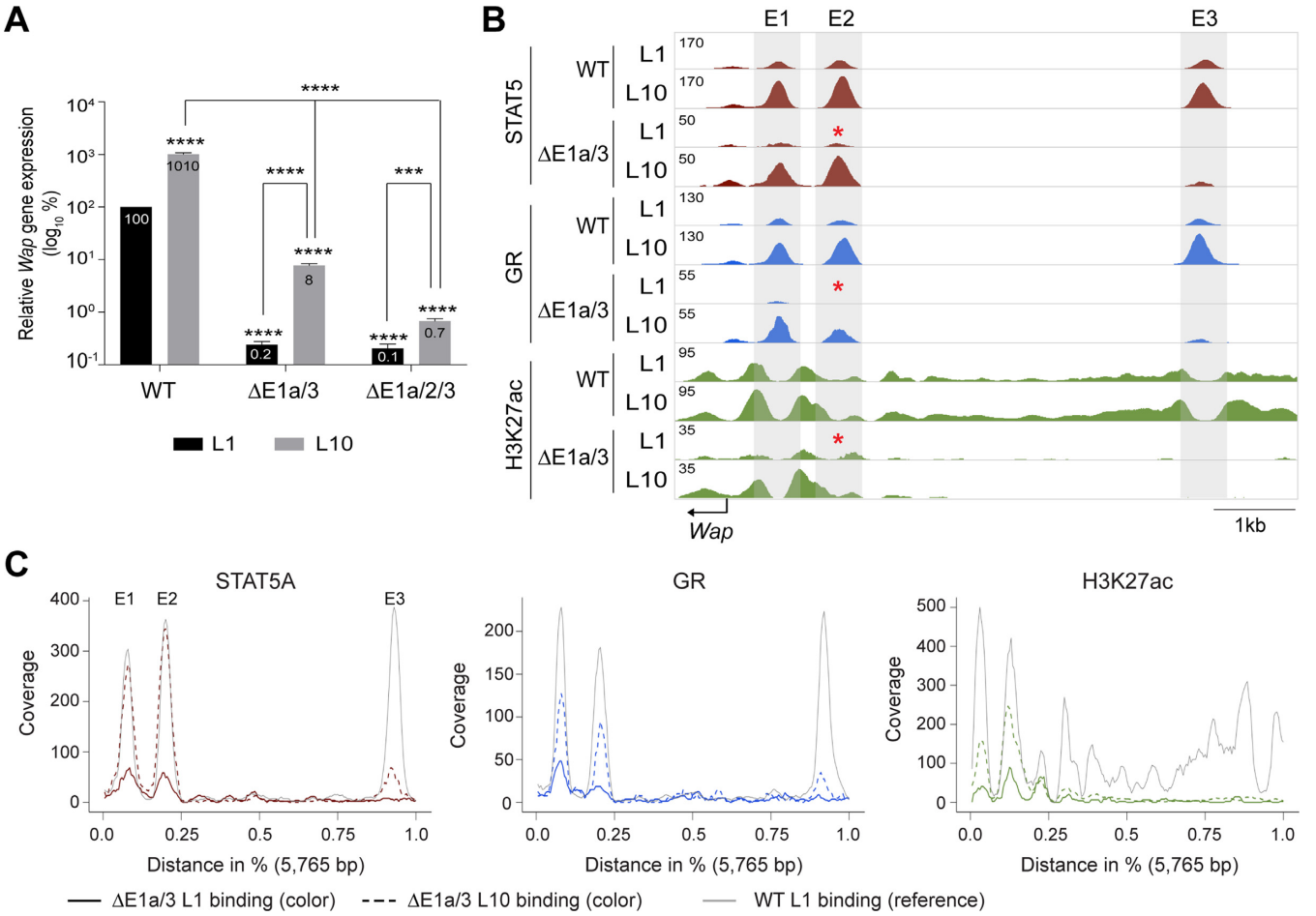


Figure 4. Differential contribution of constituent *Wap* enhancers during lactation. (**A**) *Wap* mRNA levels in mammary tissues of WT, ΔE1a/3 and ΔE1a/2/3 mutant mice at L1 and L10 were measured by qRT-PCR and normalized to *Gapdh* levels. Results are shown as the means ± SEM of independent biological replicates (WT and ΔS1a/3 at L1, $n = 6$; WT at L10, $n = 12$; ΔE1a/3 at L10, $n = 5$; ΔE1a/2/3 at L1 and L10, $n = 3$). ANOVA was used to evaluate the statistical significance of differences between WT and mutants. ***$P < 0.0001$, ****$P < 0.00001$. (**B**) ChIP-seq profile of the *Wap* locus in WT and ΔE1a/3 mutant mammary tissue at L1 and L10. Red asterisks indicate enhancer sites with a significant increase in STAT5A, GR and H3K27ac intensity at E2. The *Lalba* locus was used as a control (Supplementary Figure S6). (**C**) STAT5, GR and H3K27ac profiles at individual enhancers within the *Wap* super-enhancer in WT and ΔE1a/3 mutant tissue based on the ChIP-seq data; red (STAT5), blue (GR), H3K27ac (green), line (L1 of ΔE1a/3), dotted line (L10 of ΔE1a/3) and gray line (L1 of WT as a reference) (normalized to 10 million reads).

Mammary-restricted activity of the *Wap STAT5*-based enhancers

STAT5 is activated by a plethora of cytokines in non-mammary cells, including hepatocytes (37) and the immune system (38), and it remains an enigma why the activity of the three STAT5-based constituent *Wap* enhancers is restricted to mammary tissue. This conundrum might be explained by the need for additional transcription factors to establish mammary specificity. Alternatively, the establishment of a permissive genomic environment would be key to provide access of STAT5 to otherwise silent enhancers. Moreover, it is not clear to what extent the exceptionally high *Wap* mRNA levels during lactation reflect enhancer activity versus promoter activity or selective mRNA stabilization. We addressed these questions by moving *Wap* STAT5-based enhancers to the widely expressed *Ramp3* locus (Figure 5A). In contrast to *Ramp3*, genes further downstream (*Gm11981*, *Adcy1*, *Gm11985* and *Igfbp1*) are silent in mammary tissue and do not display H3K27ac active chromatin marks and Pol II binding. We used CRISPR/Cas9 genome editing and deleted ∼18 kb between the E2 enhancer and the *Ramp3* promoter. This region includes three CTCF sites (Figure 5B), which are known to moderate the influence of the *Wap* super-enhancer on the *Ramp3* gene (39). Moving the constituent enhancers E1 and E2 within 600 bp of the *Ramp3* promoter (E12-R3) (for details, see 'Materials and Methods' section and Supplementary Table S9) resulted in a 100-fold activation of *Ramp3* in L1 mammary tissue and more than 450-fold induction at L10 (Figure 5C and Supplementary Figure S7). This underscores that *Wap* enhancers E1 and E2 have the capacity to highly activate a common promoter in mammary tissue, probably the result of enhancer proximity and the absence of CTCF sites. STAT5 binding to E1 and E2 in mutant mammary tissue was accompanied by an expansion of H3K27ac coverage from the *Wap* enhancers into the *Ramp3* promoter (Figure 5B and C; Supplementary Figure S8 for control locus). This, together with Pol II binding to the *Ramp3* region, provides a

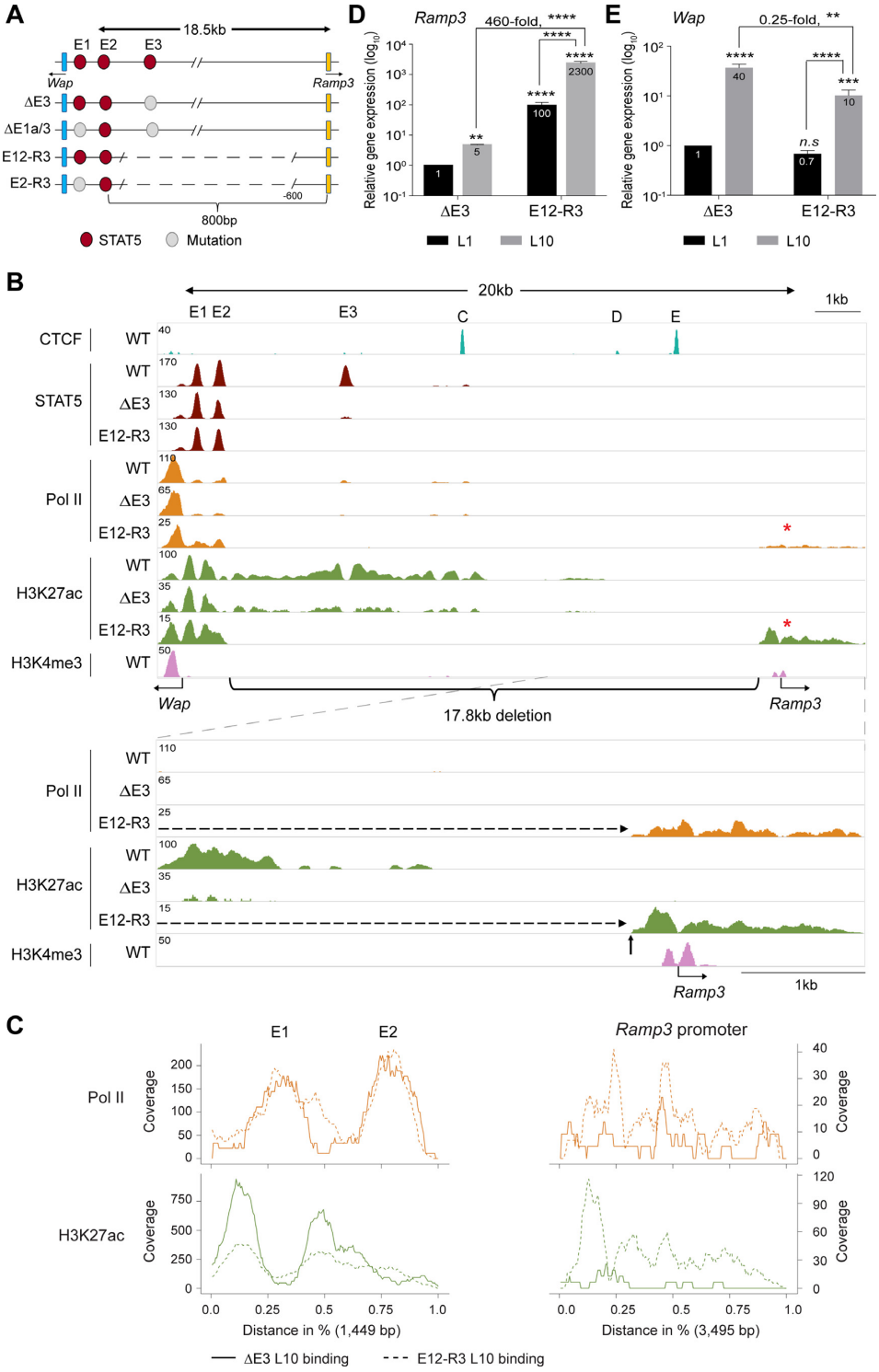


Figure 5. Capacity of *Wap* super-enhancer elements to activate the common *Ramp3* gene. (**A**) Diagram of mutants in which *Wap* enhancers were fused with the *Ramp3* promoter (E12-R3 and E2-R3). In E2-R3 mutants, the GAS site in E1 was deleted and was marked in gray. By deleting a 17.8 kb stretch of DNA, the *Wap* promoter and enhancers E1 and E2 were translocated within 600 bp of the *Ramp3* transcription start site. The exact positions of the deletions are shown in Supplementary Table S9. (**B**) STAT5A, RNA Pol II and H3K27ac landscapes at the *Wap-Ramp3* locus in WT, ΔE3 and E12-R3 mutant tissue at day 10 of lactation (L10) as well as the CTCF and H3K4me3 landscapes in WT mammary tissue at day 1 of lactation (L1). H3K27ac expansion and RNA Pol II loading occurred across the breakpoint into the *Ramp3* promoter and gene body (red asterisks). The *Lalba* locus was used as a control for ChIP-seq quality (Supplementary Figure S7). (**C**) RNA Pol II and H3K27ac coverage of the *Ramp3* promoter region in WT and mutant mammary tissue. (**D** and **E**) *Ramp3* (D) and *Wap* (E) mRNA levels in mammary tissues from ΔE3 and E12-R3 mutants at L1 and L10 were measured by qRT-PCR and normalized to *Gapdh* levels. Results are shown as the means ± SEM of independent biological replicates (ΔE3 at L1 and L10, E12-R3 at L1 and L10, $n = 3$). ANOVA was used to evaluate the statistical significance of differences between ΔE3 and E12-R3 mutant mice. $^{**}P < 0.001$, $^{***}P < 0.0001$, $^{****}P < 0.00001$. *n.s.* not significant.

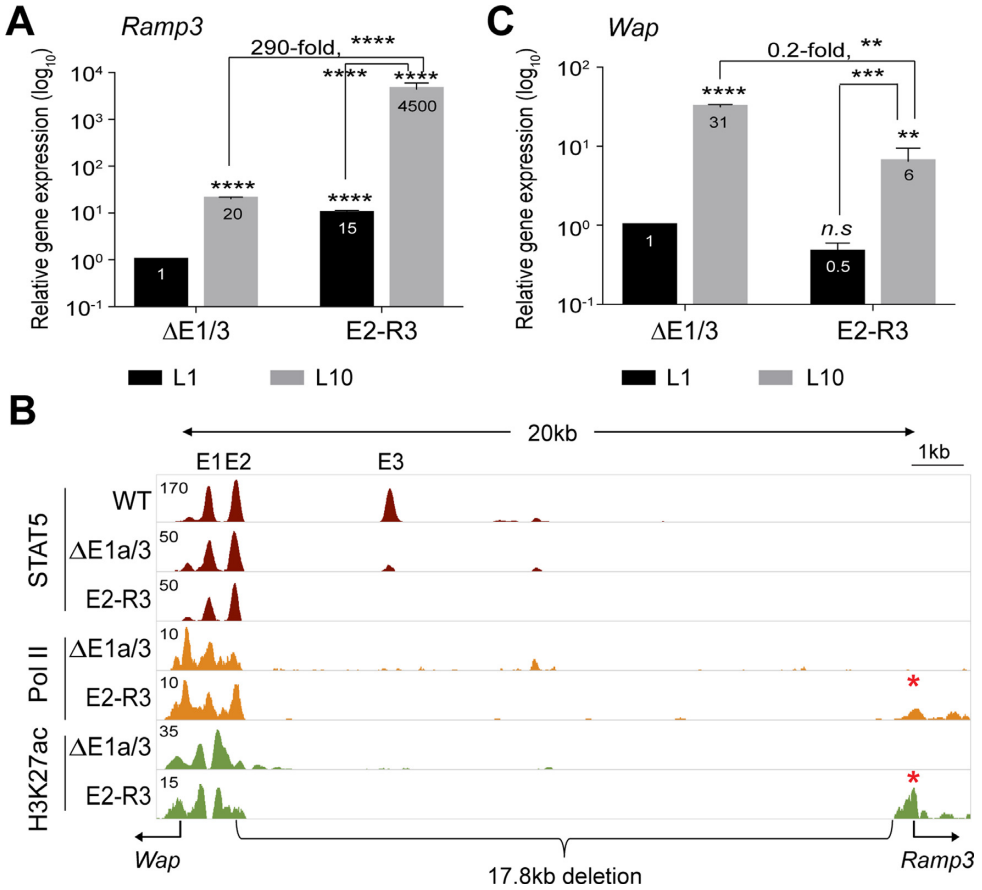


Figure 6. Capacity of constituent *Wap* enhancers to activate the common *Ramp3* gene. (**A** and **C**) *Ramp3* (A) and *Wap* (C) mRNA levels in mammary tissues from ΔE1a/3 and E2-R3 mutants at L1 and L10 were measured by qRT-PCR and normalized to *Gapdh* levels. Results are shown as the means ± SEM of independent biological replicates (ΔE1a/3 at L1 and L10, $n = 6$; E2-R3 at L1 and L10, $n = 3$). ANOVA was used to evaluate the statistical significance of differences between WT mice and mutant group. $^{**}P < 0.001$, $^{***}P < 0.0001$, $^{****}P < 0.00001$. (**B**) STAT5A, RNA Pol II and H3K27ac landscape of the *Wap*-*Ramp3* locus in WT, ΔE1a/3 and E2-R3 mutant tissue at L10. ChIP-seq analysis demonstrated an expansion of H3K27ac and RNA Pol II loading across the breakpoint into the *Ramp3* promoter and gene body. The *Lalba* locus was used as a control (Supplementary Figure S8).

mechanistic explanation for the greatly induced gene activation during pregnancy. While *Ramp3* levels in mutant mice increased ∼5-fold between L1 and L10 (Figure 5D), *Wap* levels increased 10-fold (Figure 5E), suggesting that *Wap* enhancers preferably activate their own promoter. Alternatively, *Wap* mRNAs could be specifically stabilized during lactation.

To demonstrate whether E2 by itself has the capacity to activate *Ramp3*, we generated mice in which the STAT5 site in E1 had been mutated and the region between E2 and the *Ramp3* promoter was deleted (Figure 5A). *Ramp3* levels in L1 mutant mammary tissue increased ∼15-fold (Figure 6A and Supplementary Figure S7), demonstrating that E2 can be activated during pregnancy when situated in a conducive environment. Similar to the E12-R3 fusion, we observed the expansion of H3K27ac marks and Pol II occupancy into the *Ramp3* promoter (Figure 6B; Supplementary Figure S9 for control locus). *Ramp3* expression in E2-R3 mutant mammary tissue increased almost 300-fold during lactation (Figure 6A), further highlighting the strength of this constituent enhancer with progressing differentiation. Of note, induction of the *Wap* gene was impaired in mutant mice (Figure 6C), further supporting the negative influence of the *Ramp3* locus on the inducibility of the *Wap* promoter.

Unlike *Wap*, the *Ramp3* gene is expressed across cell types (40) (Figure 7A), suggesting that the promoter is located in permissive chromatin accessible by the transcription machinery. In support of this, H3K27ac marks coincided with the *Ramp3* gene in kidney (Figure 7B). We tested the hypothesis that translocation of *Wap* enhancers into the permissive *Ramp3* locus would facilitate their activation in kidney and result in increased *Ramp3* expression. *Ramp3* expression in kidney was not altered in E12-R3 mice (Figure 7C), suggesting that the permissive chromatin did not facilitate the activation of the STAT5-based enhancers. Concordantly, no STAT5 binding was observed at E1 and E2 in WT and mutant kidney (Figure 7B and D for control locus, the *Pax8* locus (41)). Clearly, the presence of permissive chromatin does not necessarily facilitate context-independent activation of STAT5-based enhancers. Notably, fusion of *Ramp3* and *Wap* led to the establishment of H3K27ac marks over the *Wap* enhancers E1 and E2 in kidney (Figure 7B). However, this was not accompanied by a measurable activation of *Wap* and the binding of other transcription factors, including STAT3 and GR (Fig-

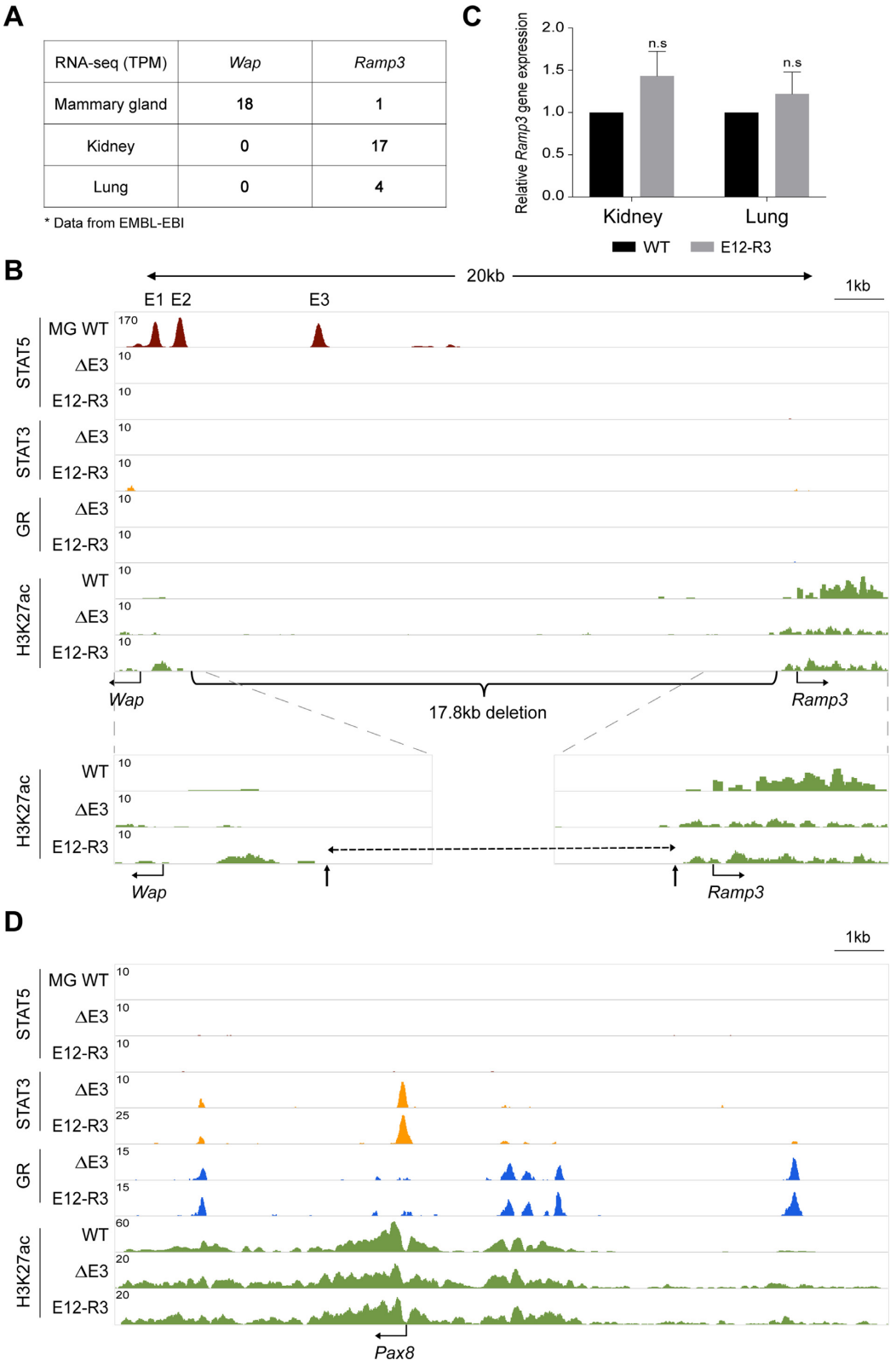


Figure 7. Analysis of *Wap* super-enhancer elements in kidney and lung. (**A**) TPM numbers from RNA-seq data of *Wap* and *Ramp3* in lactating mammary gland (DRR009919 from RIKEN FANTOM5 project), kidney and lung (GSE41637 from GEO) tissues were obtained from EMBL-EBI (https://www.ebi.ac.uk/). TPM, transcripts per kilobase million. (**B**) Genomic features of the 20 kb *Wap-Ramp3* locus in kidney tissue from ΔE3 and E12-R3 mutants. (**C**) *Wap* and *Ramp3* mRNA levels were measured by qRT-PCR in kidney and lung, from WT and E12-R3 female mutant mice (normalized to *Gapdh* levels). Results are shown as the means ± SEM of independent biological replicates (kidney and lung in ΔE3, $n = 3$; kidney in E12-R3, $n = 4$; lung in E12-R3, $n = 5$). A *t*-test was used to evaluate the statistical significance of differences in WT and mutants. (**D**) Kidney-specific *Pax8* locus (41) served as a ChIP-seq control for ΔE3 and E12-R3.

ure 7B). This study therefore demonstrates that the mere presence of H3K27ac marks is not sufficient to activate genes and establish enhancers *in vivo*.

DISCUSSION

By identifying progressive enhancer complexes in the mammary luminal lineage during lactation and the genetic dissection of a cytokine-sensing super-enhancer, we revealed hitherto unrecognized distinct functions of individual enhancer elements characterized by structurally equivalent enhancers with comparable transcription factor binding and H3K27ac coverage. We learned that while enhancer complexes originate with the initiation of luminal cell differentiation during pregnancy, their full establishment occurs with progressing differentiation and lactation driven by the master regulator STAT5. We also gained major new insight into the regulatory forces that ensure the exceptional expression of mammary-specific genes. The relative importance of constituent enhancers within a super-enhancer shifts as differentiation progresses, with the pregnancy-dependent seed enhancer not being required upon terminal differentiation. Through genetic experiments we also gained knowledge on the significance of permissive loci to respond to cytokine-sensing enhancers and the retention of cell specificity.

The concept that distinct signaling components, in conjunction with master regulators, converge to establish specific lineages has first been elegantly demonstrated in the hematopoietic system (42) and hair follicle lineages (43). STAT5 is not only central to hematopoietic lineages and mammary luminal cells that depend on cytokines (44,45), but also modulates gene expression in other cell types. The precise and cell-specific roles of STAT5-based enhancers might depend on context, possibly their relative position within super-enhancers, relative distance to promoters or accessibility in permissive chromatin (10,39,46,47). Chromatin accessibility is considered as a key to transcription factor binding and activating gene expression (48). Based on biochemical and genetic studies, a complex seed enhancer, which is bound by several mammary-enriched transcription factors, controls the *Wap* super-enhancer during pregnancy (12). It is likely that a pioneer factor, possibly STAT5 in the *Wap* super-enhancer, samples the genome and cooperative binding with other transcription factor results in stable occupancy as shown in other systems (49,50). Our study demonstrates that the seed enhancer concept is not absolute, but is void under signaling conditions that permit the activation of secondary enhancers.

Super-enhancers, also known as stretch enhancers (51), have originally been characterized as structural units, regardless of their biological activity. They are composed of several densely occupied transcription factor platforms that coincide with MED1 or H3K27ac marks (32–33,52). Although the transcription factor coverage between super-enhancers is frequently equivalent, expression levels of their associated genes can vary widely. In mammary tissue, expression of genes associated with equivalent super-enhancers can differ by more than 1000-fold (12). It is possible that promoter activity and preferential mRNA stabilization are contributing factors.

As shown in this study, enhancer-promoter proximity and the absence of CTCF sites (39) permit mammary super-enhancer elements to activate a receptive foreign promoter, albeit less than the own promoter. Genes regulated by tissue-specific super-enhancers are frequently located within chromatin loops that are anchored by CTCFs to possibly confine enhancer activity (53–57). In a few cases, disruption of CTCF-binding sites in mice led to ectopic expression of genes outside topologically associating domain (TAD) or sub-TAD (58,59). Deletion of CTCF sites between *Wap* and *Ramp3* results in *Ramp3* activation through an interaction of the *Wap*-E3 enhancer and the first intron of *Ramp3* gene in mutant mice (39). Thus, increased *Ramp3* expression in mutant mammary tissue is likely a combination of enhancer-promoter proximity and the absence of CTCF sites.

It can be predicted that permissive chromatin provides an accommodating environment for the binding of common and cell-enriched transcription factors and the activation of associated genes (48). Thus, our finding that STAT5-based *Wap* enhancers placed in open chromatin were unable to further induce an already active neighboring gene in kidney tissue was not expected. Notably, the *de novo* establishment of H3K27ac marks at the *Wap* promoter in kidney tissue failed to activate an otherwise silent promoter and enhancer. In contrast, in tissue culture cells the forced introduction of H3K27ac marks into enhancer regions leads to their activation (52), highlighting fundamental mechanistic differences between genuine *in vivo* and cell culture biology. Our study leads to the realization that permissive chromatin is not sufficient to activate STAT5-based enhancers, adding further intrigue on yet unknown regulatory elements embedded in common enhancers that exclusively respond to cell-specific cues.

DATA AVAILABILITY

ChIP-seq data of wild-type and mutant mammary tissue at L1 as well as wild-type kidney tissue were obtained under GSE74826 and GSM1000092 in the Gene Expression Omnibus (GEO). Liver and T cell data are available under accessions GSE31578, GSE27158 and GSE31039. RNA-seq data of mouse lactating mammary gland as well as kidney and lung tissues are available under DRR009919 from RIKEN FANTOM5 project and GSE41637 from GEO, respectively. ChIP-seq (two replicates under GSE115368) and RNA-seq (three or four replicates under GEO115369) data sets of wild-type and mutant mammary tissues at L1 and L10 have been deposited under GSE115370.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

This work utilized the computational resources of the NIH HPC Biowulf cluster (http://hpc.nih.gov).
Author contributions: H.K.L., M.W., C.L. and L.H contributed to study design. C.L. contributed to the generation of mutant mice. H.K.L. and H.Y.S. contributed to the establishment of mutant mouse lines. H.K.L. contributed to

experiments and data analysis. M.W. contributed to computational analysis. L.H. contributed to the supervision of the study. H.K.L., M.W. and L.H. wrote the manuscript and all authors approved the final version.

FUNDING

Intramural Research Programs (IRPs) of National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) and National Heart, Lung, and Blood Institute (NHLBI). Funding for open access charge: IRP of the NIH.
Conflict of interest statement. None declared.

REFERENCES

1. Hennighausen,L. and Robinson,G.W. (2005) Information networks in the mammary gland. *Nat. Rev. Mol. Cell Biol.*, **6**, 715–725.
2. Ormandy,C.J., Camus,A., Barra,J., Damotte,D., Lucas,B., Buteau,H., Edery,M., Brousse,N., Babinet,C., Binart,N. *et al.* (1997) Null mutation of the prolactin receptor gene produces multiple reproductive defects in the mouse. *Genes Dev.*, **11**, 167–178.
3. Horseman,N.D., Zhao,W., Montecino-Rodriguez,E., Tanaka,M., Nakashima,K., Engle,S.J., Smith,F., Markoff,E. and Dorshkind,K. (1997) Defective mammopoiesis, but normal hematopoiesis, in mice with a targeted disruption of the prolactin gene. *EMBO J.*, **16**, 6926–6935.
4. Robinson,G.W., McKnight,R.A., Smith,G.H. and Hennighausen,L. (1995) Mammary epithelial cells undergo secretory differentiation in cycling virgins but require pregnancy for the establishment of terminal differentiation. *Development*, **121**, 2079–2090.
5. Neville,M.C., McFadden,T.B. and Forsyth,I. (2002) Hormonal regulation of mammary differentiation and milk secretion. *J. Mammary Gland Biol. Neoplasia*, **7**, 49–66.
6. O'Shea,J.J., Lahesmaa,R., Vahedi,G., Laurence,A. and Kanno,Y. (2011) Genomic views of STAT function in CD4+ T helper cell differentiation. *Nat. Rev. Immunol.*, **11**, 239–250.
7. Wei,L., Vahedi,G., Sun,H.W., Watford,W.T., Takatori,H., Ramos,H.L., Takahashi,H., Liang,J., Gutierrez-Cruz,G., Zang,C. *et al.* (2010) Discrete roles of STAT4 and STAT6 transcription factors in tuning epigenetic modifications and transcription during T helper cell differentiation. *Immunity*, **32**, 840–851.
8. Vahedi,G., Takahashi,H., Nakayamada,S., Sun,H.W., Sartorelli,V., Kanno,Y. and O'Shea,J.J. (2012) STATs shape the active enhancer landscape of T cell populations. *Cell*, **151**, 981–993.
9. Yamaji,D., Na,R., Feuermann,Y., Pechhold,S., Chen,W., Robinson,G.W. and Hennighausen,L. (2009) Development of mammary luminal progenitor cells is controlled by the transcription factor STAT5A. *Genes Dev.*, **23**, 2382–2387.
10. Willi,M., Yoo,K.H., Wang,C., Trajanoski,Z. and Hennighausen,L. (2016) Differential cytokine sensitivities of STAT5-dependent enhancers rely on Stat5 autoregulation. *Nucleic Acids Res.*, **44**, 10277–10291.
11. Yamaji,D., Kang,K., Robinson,G.W. and Hennighausen,L. (2013) Sequential activation of genetic programs in mouse mammary epithelium during pregnancy depends on STAT5A/B concentration. *Nucleic Acids Res.*, **41**, 1622–1636.
12. Shin,H.Y., Willi,M., HyunYoo,K., Zeng,X., Wang,C., Metser,G. and Hennighausen,L. (2016) Hierarchy within the mammary STAT5-driven Wap super-enhancer. *Nat. Genet.*, **48**, 904–911.
13. Liu,X., Robinson,G.W., Wagner,K.U., Garrett,L., Wynshaw-Boris,A. and Hennighausen,L. (1997) Stat5a is mandatory for adult mammary gland development and lactogenesis. *Genes Dev.*, **11**, 179–186.
14. Cui,Y., Riedlinger,G., Miyoshi,K., Tang,W., Li,C., Deng,C.X., Robinson,G.W. and Hennighausen,L. (2004) Inactivation of Stat5 in mouse mammary epithelium during pregnancy reveals distinct functions in cell proliferation, survival, and differentiation. *Mol. Cell. Biol.*, **24**, 8037–8047.
15. Osterwalder,M., Barozzi,I., Tissieres,V., Fukuda-Yuzawa,Y., Mannion,B.J., Afzal,S.Y., Lee,E.A., Zhu,Y., Plajzer-Frick,I., Pickle,C.S. *et al.* (2018) Enhancer redundancy provides phenotypic robustness in mammalian development. *Nature*, **554**, 239–243.
16. Mouri,K., Sagai,T., Maeno,A., Amano,T., Toyoda,A. and Shiroishi,T. (2018) Enhancer adoption caused by genomic insertion elicits interdigital Shh expression and syndactyly in mouse. *Proc. Natl. Acad. Sci. U.S.A.*, **115**, 1021–1026.
17. Will,A.J., Cova,G., Osterwalder,M., Chan,W.L., Wittler,L., Brieske,N., Heinrich,V., de Villartay,J.P., Vingron,M., Klopocki,E. *et al.* (2017) Composition and dosage of a multipartite enhancer cluster control developmental expression of Ihh (Indian hedgehog). *Nat. Genet.*, **49**, 1539–1545.
18. Bahr,C., von Paleske,L., Uslu,V.V., Remeseiro,S., Takayama,N., Ng,S.W., Murison,A., Langenfeld,K., Petretich,M., Scognamiglio,R. *et al.* (2018) A Myc enhancer cluster regulates normal and leukaemic haematopoietic stem cell hierarchies. *Nature*, **553**, 515–520.
19. Metser,G., Shin,H.Y., Wang,C., Yoo,K.H., Oh,S., Villarino,A.V., O'Shea,J.J., Kang,K. and Hennighausen,L. (2016) An autoregulatory enhancer controls mammary-specific STAT5 functions. *Nucleic Acids Res.*, **44**, 1052–1063.
20. Bolger,A.M., Lohse,M. and Usadel,B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.
21. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.
22. Heinz,S., Benner,C., Spann,N., Bertolino,E., Lin,Y.C., Laslo,P., Cheng,J.X., Murre,C., Singh,H. and Glass,C.K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell*, **38**, 576–589.
23. Thorvaldsdottir,H., Robinson,J.T. and Mesirov,J.P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.*, **14**, 178–192.
24. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
25. Huber,W., Carey,V.J., Gentleman,R., Anders,S., Carlson,M., Carvalho,B.S., Bravo,H.C., Davis,S., Gatto,L., Girke,T. *et al.* (2015) Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods*, **12**, 115–121.
26. Liao,Y., Smyth,G.K. and Shi,W. (2013) The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.*, **41**, e108.
27. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
28. Wickham,H. (2009) *Ggplot2 : Elegant Graphics for Data Analysis*. Springer, NY.
29. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
30. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
31. Ross-Innes,C.S., Stark,R., Teschendorff,A.E., Holmes,K.A., Ali,H.R., Dunning,M.J., Brown,G.D., Gojis,O., Ellis,I.O., Green,A.R. *et al.* (2012) Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature*, **481**, 389–393.
32. Whyte,W.A., Orlando,D.A., Hnisz,D., Abraham,B.J., Lin,C.Y., Kagey,M.H., Rahl,P.B., Lee,T.I. and Young,R.A. (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*, **153**, 307–319.
33. Loven,J., Hoke,H.A., Lin,C.Y., Lau,A., Orlando,D.A., Vakoc,C.R., Bradner,J.E., Lee,T.I. and Young,R.A. (2013) Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell*, **153**, 320–334.
34. Knight,C.H., Maltz,E. and Docherty,A.H. (1986) Milk yield and composition in mice: effects of litter size and lactation number. *Comp. Biochem. Physiol. A Comp. Physiol.*, **84**, 127–133.
35. Stocklin,E., Wissler,M., Gouilleux,F. and Groner,B. (1996) Functional interactions between Stat5 and the glucocorticoid receptor. *Nature*, **383**, 726–728.
36. Stoecklin,E., Wissler,M., Moriggl,R. and Groner,B. (1997) Specific DNA binding of Stat5, but not of glucocorticoid receptor, is required for their functional cooperation in the regulation of gene transcription. *Mol. Cell. Biol.*, **17**, 6708–6716.

37. Cui,Y., Hosui,A., Sun,R., Shen,K., Gavrilova,O., Chen,W., Cam,M.C., Gao,B., Robinson,G.W. and Hennighausen,L. (2007) Loss of signal transducer and activator of transcription 5 leads to hepatosteatosis and impaired liver regeneration. *Hepatology*, **46**, 504–513.
38. Yao,Z., Cui,Y., Watford,W.T., Bream,J.H., Yamaoka,K., Hissong,B.D., Li,D., Durum,S.K., Jiang,Q., Bhandoola,A. *et al.* (2006) Stat5a/b are essential for normal lymphoid development and differentiation. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 1000–1005.
39. Willi,M., Yoo,K.H., Reinisch,F., Kuhns,T.M., Lee,H.K., Wang,C. and Hennighausen,L. (2017) Facultative CTCF sites moderate mammary super-enhancer activity and regulate juxtaposed gene in non-mammary cells. *Nat. Commun.*, **8**, 16069.
40. Yue,F., Cheng,Y., Breschi,A., Vierstra,J., Wu,W., Ryba,T., Sandstrom,R., Ma,Z., Davis,C., Pope,B.D. *et al.* (2014) A comparative encyclopedia of DNA elements in the mouse genome. *Nature*, **515**, 355–364.
41. Tong,G.X., Yu,W.M., Beaubier,N.T., Weeden,E.M., Hamele-Bena,D., Mansukhani,M.M. and O'Toole,K.M. (2009) Expression of PAX8 in normal and neoplastic renal tissues: an immunohistochemical study. *Mod. Pathol.*, **22**, 1218–1227.
42. Trompouki,E., Bowman,T.V., Lawton,L.N., Fan,Z.P., Wu,D.C., DiBiase,A., Martin,C.S., Cech,J.N., Sessa,A.K., Leblanc,J.L. *et al.* (2011) Lineage regulators direct BMP and Wnt pathways to cell-specific programs during differentiation and regeneration. *Cell*, **147**, 577–589.
43. Lien,W.H., Guo,X., Polak,L., Lawton,L.N., Young,R.A., Zheng,D. and Fuchs,E. (2011) Genome-wide maps of histone modifications unwind *in vivo* chromatin states of the hair follicle lineage. *Cell Stem Cell*, **9**, 219–232.
44. Nosaka,T. and Kitamura,T. (2000) Janus kinases (JAKs) and signal transducers and activators of transcription (STATs) in hematopoietic cells. *Int. J. Hematol.*, **71**, 309–319.
45. Hennighausen,L. and Robinson,G.W. (2008) Interpretation of cytokine signaling through the transcription factors STAT5A and STAT5B. *Genes Dev.*, **22**, 711–721.
46. Vahedi,G., Kanno,Y., Furumoto,Y., Jiang,K., Parker,S.C., Erdos,M.R., Davis,S.R., Roychoudhuri,R., Restifo,N.P., Gadina,M. *et al.* (2015) Super-enhancers delineate disease-associated regulatory nodes in T cells. *Nature*, **520**, 558–562.
47. Li,P., Mitra,S., Spolski,R., Oh,J., Liao,W., Tang,Z., Mo,F., Li,X., West,E.E., Gromer,D. *et al.* (2017) STAT5-mediated chromatin interactions in superenhancers activate IL-2 highly inducible genes: Functional dissection of the *Il2ra* gene locus.. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 12111–12119.
48. Uyehara,C.M., Nystrom,S.L., Niederhuber,M.J., Leatham-Jensen,M., Ma,Y., Buttitta,L.A. and McKay,D.J. (2017) Hormone-dependent control of developmental timing through regulation of chromatin accessibility. *Genes Dev.*, **31**, 862–875.
49. Donaghey,J., Thakurela,S., Charlton,J., Chen,J.S., Smith,Z.D., Gu,H., Pop,R., Clement,K., Stamenova,E.K., Karnik,R. *et al.* (2018) Genetic determinants and epigenetic effects of pioneer-factor occupancy. *Nat. Genet.*, **50**, 250–258.
50. Mayran,A., Khetchoumian,K., Hariri,F., Pastinen,T., Gauthier,Y., Balsalobre,A. and Drouin,J. (2018) Pioneer factor Pax7 deploys a stable enhancer repertoire for specification of cell fate. *Nat. Genet.*, **50**, 259–269.
51. Quang,D.X., Erdos,M.R., Parker,S.C.J. and Collins,F.S. (2015) Motif signatures in stretch enhancers are enriched for disease-associated genetic variants. *Epigenetics Chromatin*, **8**, 23.
52. Hilton,I.B., D'Ippolito,A.M., Vockley,C.M., Thakore,P.I., Crawford,G.E., Reddy,T.E. and Gersbach,C.A. (2015) Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.*, **33**, 510–517.
53. Dowen,J.M., Fan,Z.P., Hnisz,D., Ren,G., Abraham,B.J., Zhang,L.N., Weintraub,A.S., Schuijers,J., Lee,T.I., Zhao,K. *et al.* (2014) Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell*, **159**, 374–387.
54. Hnisz,D., Day,D.S. and Young,R.A. (2016) Insulated Neighborhoods: Structural and Functional Units of Mammalian Gene Control. *Cell*, **167**, 1188–1200.
55. Hay,D., Hughes,J.R., Babbs,C., Davies,J.O.J., Graham,B.J., Hanssen,L., Kassouf,M.T., Marieke Oudelaar,A.M., Sharpe,J.A., Suciu,M.C. *et al.* (2016) Genetic dissection of the alpha-globin super-enhancer *in vivo*. *Nat. Genet.*, **48**, 895–903.
56. Schuijers,J., Manteiga,J.C., Weintraub,A.S., Day,D.S., Zamudio,A.V., Hnisz,D., Lee,T.I. and Young,R.A. (2018) Transcriptional dysregulation of MYC reveals common enhancer-docking mechanism. *Cell Rep.*, **23**, 349–360.
57. Huang,J., Li,K., Cai,W., Liu,X., Zhang,Y., Orkin,S.H., Xu,J. and Yuan,G.C. (2018) Dissecting super-enhancer hierarchy based on chromatin interactions. *Nat. Commun.*, **9**, 943.
58. Dowen,J.M., Fan,Z.P., Hnisz,D., Ren,G., Abraham,B.J., Zhang,L.N., Weintraub,A.S., Schujiers,J., Lee,T.I., Zhao,K. *et al.* (2014) Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell*, **159**, 374–387.
59. Hanssen,L.L.P., Kassouf,M.T., Oudelaar,A.M., Biggs,D., Preece,C., Downes,D.J., Gosden,M., Sharpe,J.A., Sloane-Stanley,J.A., Hughes,J.R. *et al.* (2017) Tissue-specific CTCF-cohesin-mediated chromatin architecture delimits enhancer interactions and function *in vivo*. *Nat. Cell Biol.*, **19**, 952–961.