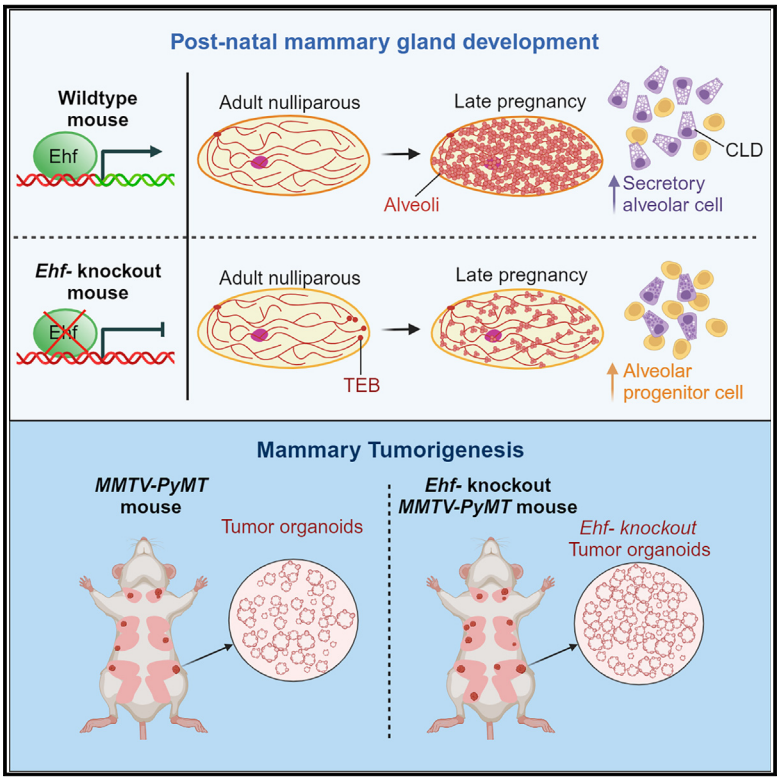
Article

Developmental Cell

Ehf controls mammary alveolar lineage differentiation and is a putative suppressor of breast tumorigenesis

Graphical abstract



Authors

Rebecca Nightingale, Camilla M. Reehorst, Natalia Vukelic, ..., Jane E. Visvader, Bhupinder Pal, John M. Mariadason

Correspondence

bhupinder.pal@onjcri.org.au (B.P.), john.mariadason@onjcri.org.au (J.M.M.)

In brief

Nightingale and Reehorst et al. identify a key role for the transcription factor EHF in mouse mammary gland development and in controlling the differentiation of mammary epithelial cells into milk-producing cells during pregnancy. They also show that Ehf is an important suppressor of tumor development in this tissue.

Highlights

- The transcription factor EHF is essential for mammary development in pregnancy
- Ehf drives differentiation of progenitor cells into milk-producing alveolar cells
- Ehf deletion increased tumor development in the mouse mammary gland

Nightingale et al., 2024, Developmental Cell *59*, 1988–2004
August 5, 2024
https://doi.org/10.1016/j.devcel.2024.04.022

CellPress

CellPress
OPEN ACCESS

Developmental Cell

Article

Ehf controls mammary alveolar lineage differentiation and is a putative suppressor of breast tumorigenesis

Rebecca Nightingale,[1,2,12] Camilla M. Reehorst,[1,2,12] Natalia Vukelic,[1,2] Nikolaos Papadopoulos,[1,2] Yang Liao,[1,2] Shalini Guleria,[1,2] Caroline Bell,[1,2] François Vaillant,[3,4] Sudip Paul,[5,6] Ian Y. Luk,[1,2] Amardeep S. Dhillon,[7] Laura J. Jenkins,[1,2] Riley J. Morrow,[1,2] Felicity C. Jackling,[3] Ashwini L. Chand,[1,2] David Chisanga,[1,2] Yunshun Chen,[3,4,8] David S. Williams,[1,2,11] Robin L. Anderson,[1,2,9] Sarah Ellis,[1,2] Peter J. Meikle,[5,6] Wei Shi,[1,2] Jane E. Visvader,[3,4] Bhupinder Pal,[1,2,*] and John M. Mariadason[1,2,10,13,*]

[1]Olivia Newton-John Cancer Research Institute, Heidelberg, VIC 3084, Australia
[2]School of Cancer Medicine, La Trobe University, Bundoora, VIC 3086, Australia
[3]Cancer Biology and Stem Cells Division, The Walter and Eliza Hall Institute, Parkville, VIC 3052, Australia
[4]Department of Medical Biology, University of Melbourne, Parkville, VIC 3052, Australia
[5]Metabolomics Laboratory, Baker Heart and Diabetes Institute, Melbourne, VIC 3004, Australia
[6]Baker Department of Cardiovascular Research Translation and Implementation, La Trobe University, Bundoora, VIC 3086, Australia
[7]The Institute for Mental and Physical Health and Clinical Translation, School of Medicine, Deakin University, Geelong, VIC 3220, Australia
[8]Bioinformatics Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, VIC 3052, Australia
[9]The Sir Peter MacCallum Department of Oncology, University of Melbourne, Parkville, VIC 3052, Australia
[10]Department of Medicine, University of Melbourne, Parkville, VIC 3052, Australia
[11]Department of Pathology, Austin Health, Heidelberg, VIC 3084, Australia
[12]These authors contributed equally
[13]Lead contact
*Correspondence: bhupinder.pal@onjcri.org.au (B.P.), john.mariadason@onjcri.org.au (J.M.M.)
https://doi.org/10.1016/j.devcel.2024.04.022

SUMMARY

The transcription factor EHF is highly expressed in the lactating mammary gland, but its role in mammary development and tumorigenesis is not fully understood. Utilizing a mouse model of *Ehf* deletion, herein, we demonstrate that loss of *Ehf* impairs mammary lobuloalveolar differentiation at late pregnancy, indicated by significantly reduced levels of milk genes and milk lipids, fewer differentiated alveolar cells, and an accumulation of alveolar progenitor cells. Further, deletion of *Ehf* increased proliferative capacity and attenuated prolactin-induced alveolar differentiation in mammary organoids. *Ehf* deletion also increased tumor incidence in the *MMTV-PyMT* mammary tumor model and increased the proliferative capacity of mammary tumor organoids, while low EHF expression was associated with higher tumor grade and poorer outcome in luminal A and basal human breast cancers. Collectively, these findings establish EHF as a non-redundant regulator of mammary alveolar differentiation and a putative suppressor of mammary tumorigenesis.

INTRODUCTION

The primary function of the mammary gland is to synthesize and secrete milk that sustains the growth and development of progeny.[1] The mammary gland is distinct from most other organs, as it undergoes major developmental differentiation in post-natal life.[2] The adult mammary gland comprises a complex, bilayered ductal epithelium,[3] consisting of apically oriented luminal epithelial cells surrounding a central lumen and an outer layer of basal/myoepithelial cells in contact with the basement membrane.[3,4] Upon pregnancy, progesterone and prolactin induce alveologenesis, an extensive tissue remodeling process that involves the co-operation of basal cells, hormone-sensing estrogen- and progesterone-receptor-positive (ER^+PR^+) luminal cells, and ER^-PR^- alveolar luminal cells.[5] Alveologenesis is initiated with a proliferative phase, involving proliferation of the epithelial compartment to expand the ductal system[6] and the formation of lobuloalveolar structures, which are comprised of a similar bilayered structure to that to the ductal system.[5] In mid pregnancy, the secretory initiation phase sees luminal alveolar progenitors differentiate into mature milk-secreting alveolar cells,[5] resulting in the upregulation of milk protein genes and the formation of intracellular lipid droplets.[7] Finally, by late pregnancy, secretory activation triggers milk proteins and lipid droplets to migrate into the alveolar lumen in preparation for active post-partum lactation.[6,8]

While a number of signaling pathways have been implicated as key molecular regulators of the luminal lineage, including but not

limited to Notch,[9,10] Prlr/Jak2/Stat5,[11–14] RANK,[15] and Erbb4[16] signaling, together with the transcription factors (TFs) ELF5,[10,15,17–19] GATA3 and its target FOXA1,[3,20,21] C/EBPβ,[22,23] and BLIMP1,[24] there remain significant gaps in our knowledge of the molecular mechanisms that drive alveologenesis.

The epithelium-specific Ets TFs (ESEs) are a subfamily of Ets TFs comprising four members, *ELF3*, *ELF5*, *EHF*, and *SPDEF*, so named due to their predominantly epithelium-specific expression profiles.[25–27] Among these TFs, the least characterized is EHF, which consists of a pointed domain (PNT) with the potential for homodimerization and co-operative binding with other TFs and gene transactivation, and an evolutionarily conserved winged helix-turn-helix ETS DNA-binding domain that binds the DNA consensus sequence A(T/C)(C/G)AGGAAGT.[28]

Previous mapping of cell-specific active enhancers and associated TF networks in human mammary epithelial cells (MECs) identified a transcriptional regulatory network driven by *EHF* and *ELF5* in luminal progenitor (LP) cells.[29] Consistent with this finding, upregulation of EHF is observed in the mouse and human LP population[30] and EHF was identified as a putative transcriptional regulator in LP cells in an independent single nucleus ATAC sequencing dataset.[31] Single-cell RNA sequencing (scRNA-seq) of mouse MECs determined *Elf5* and *Ehf* as among the most enriched TFs during differentiation toward the secretory lineage in gestation and lactation,[32] and a supporting study found that *Ehf* was induced during the initial phase of lactogenic differentiation in mouse MECs.[33] Conversely, the *Ehf* regulon was repressed in the basal compartment of the adult gland.[34] Interestingly however, high *Ehf* expression is observed in a rare mixed-lineage cluster identified among basal cells expressing a subset of luminal genes in adult mouse MECs, potentially representing a population of transit or "lineage-primed" cells[35] and supporting the notion that EHF may be involved in commitment to a luminal cell fate in the post-natal mammary gland. Nevertheless, while this evidence collectively suggests a critical role for EHF in luminal lineage differentiation of the mammary gland, direct functional studies to validate these observations are lacking and the role of EHF in post-natal mammary development and physiology has not previously been reported. Moreover, delineating regulators of normal mammary development has the potential to identify drivers of mammary tumorigenesis, as many pathways and TFs involved in mammary remodeling are frequently perturbed during oncogenesis.[36] Interestingly, recent studies in human and mouse breast cancer cells suggest a potential tumor-suppressive role for EHF,[37,38] though this remains to be tested directly in an *in vivo* setting.

We recently generated a mouse model that enables systemic or tissue-specific deletion of the DNA-binding Ets domain of the *Ehf* gene.[39] Herein, we utilize these Ehf^{KO} mice to investigate the role of Ehf in post-natal mammary development and breast tumorigenesis. In mice with systemic or mammary-specific *Ehf* deletion, we identify a defect in mammary alveolar differentiation, establishing EHF as a non-redundant TF required for the differentiation of alveolar cells in pregnancy. In addition, we demonstrate a significant increase in mammary tumor incidence in Ehf^{KO}*MMTV-PyMT* mice. This was consistent with low EHF expression in human breast cancers being associated with higher tumor grade and poorer outcome, thus identifying *Ehf* as a potential breast tumor suppressor.

RESULTS

Ehf deletion causes a lactation defect

To determine the impact of whole-body *Ehf* gene deletion on fertility, mammary gland development, and lactation, Ehf^{KO}, Ehf^{HET}, and Ehf^{WT} females were mated with Ehf^{WT} males. All Ehf^{KO} females became pregnant and produced viable and phenotypically normal offspring, born at the expected Mendelian ratio and litter sizes. However, the neonatal mortality rate from first litters born to Ehf^{KO} females was significantly higher (85.7%) compared with litters born to Ehf^{HET} (22.1%) and Ehf^{WT} (7.8%) females ($p < 0.001$, chi-squared test) (Figure 1A). Neonate lethality occurred within 48 h post parturition despite Ehf^{KO} mothers displaying typical nursing behavior. Necropsy examination of 1-day-old pups revealed an absence of milk in their stomachs. The 8 surviving pups from a total of 56 pups born from Ehf^{KO} females were fostered onto surrogate mothers. Lethality of neonates was independent of their genotype (Figure 1B), further indicating that the neonatal lethality was due to lactation insufficiency of Ehf^{KO} mothers. Second litters born to Ehf^{KO} females had a reduced but still significant neonatal mortality rate of 60% (Figure 1A), indicating that the lactation defect in Ehf^{KO} females is partially overcome by exposure to a subsequent pregnancy.

Expression of Ehf in the post-natal mammary gland

The inability of Ehf^{KO} females to support their pups suggested a potential defect in mammary gland development. As the role of *Ehf* has not previously been studied in this context, we first evaluated *Ehf* expression in whole mammary tissue during post-natal development. Relative to the pubertal gland, *Ehf* expression was 1.4-fold higher in the adult gland, then further induced in gestation, increasing 3.2-fold in early pregnancy (6.5 days pregnant [dP]), 8.1-fold at mid-pregnancy (10.5 dP), and peaking at 13.6-fold in late pregnancy (18.5 dP). Subsequently, *Ehf* levels rapidly declined during lactation (2–3 and 8 [days lactating]), returning to pre-pregnancy levels at 4 days post-involution (4 dPI) (Figure 1C).

To identify the specific MEC type(s) expressing *Ehf*, we interrogated published mouse scRNA-seq data of 15 annotated mouse MECs at four key stages of mammary gland development: nulliparous (NP) adult, mid-pregnancy (14.5 dP), during active lactation (6 dL), and in post-involution (11 dPI) (Figure 1D).[32] *Ehf* expression was largely restricted to MECs of the luminal lineage (Figure 1E) and was most highly expressed in alveolar progenitors in gestation (Avp-G) and lactation (Avp-L), and in differentiated alveolar cells in gestation (Avd-G) (Figure 1F). Similarly, publicly available scRNA-data from normal human mammary cells[40] (Figure 1G) confirmed that *EHF* expression was confined to epithelial cells of the luminal lineage (LP) and mature luminal (ML) (Figures 1H and 1I).

Loss of Ehf impairs ductal elongation in puberty and moderately reduces ductal area in the adult mammary gland

To eliminate the possibility that the observed lactation defect of Ehf^{KO} mothers was being driven by an abnormality in mammary tissue architecture occurring in early post-natal mammary development, we examined the mammary gland structure at puberty, during the critical stages of ductal elongation and branching

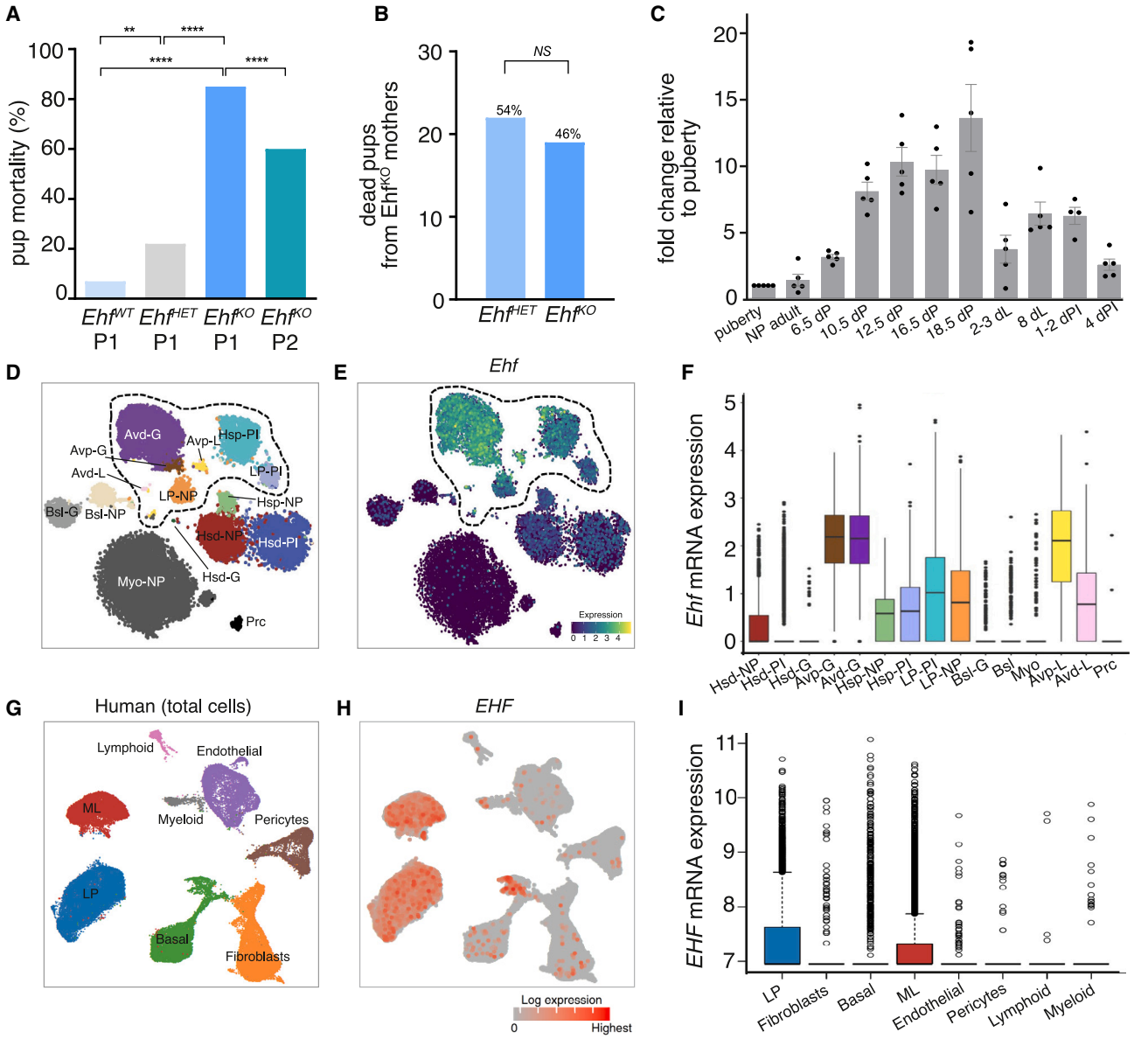


Figure 1. *Ehf* is highly expressed in luminal mammary epithelial cells and its deletion in mice causes a high pup mortality rate

(A) Mortality rate (%) of pups born to uniparous (P1) Ehf^{WT} (n = 17 mothers, 114 pups), Ehf^{HET} (n = 16 mothers, 95 pups), Ehf^{KO} (n = 10 mothers, 56 pups) mothers, and biparous (P2) Ehf^{KO} mothers (n = 6 mothers, 40 pups). $^{****}p \leq 0.0001$, $^{**}p \leq 0.01$. Chi-squared test.

(B) Distribution of genotypes of first-born litters from Ehf^{KO} mothers mated to Ehf^{HET} males. n = 41 out of the 48 dead pups genotyped. NS, not significant. Chi-squared test.

(C) *Ehf* mRNA expression relative to *β-actin* (*ACTB*) in whole mammary tissue from Ehf^{WT} females at puberty (6 weeks old), nulliparous adult stage (NP adult, 10 weeks old), and various stages of pregnancy and post involution (dP, days pregnant; dL, days lactation; dPI, days post involution). n = 5 mice per stage.

(D) 15 annotated clusters of mammary epithelial cells from *C57Bl/6N* female mice at 8 weeks old (NP adult), 14.5 dP, 6 dL, and 11 dPI identified as hormone-sensing differentiated (HsD), hormone-sensing progenitor (HsP), luminal progenitor (LP), luminal differentiated (LD), alveolar progenitor (Avp), alveolar differentiated (Avd), myoepithelial (Myo), asbal (Bsl), and PRC (Procr+ basal cells, enriched for mammary repopulating units) from Bach et al.[32] (STAR Methods).

(E) Identification of MEC subsets with expression of *Ehf* in the 15 annotated MEC clusters.

(F) *Ehf* mRNA expression in the 15 annotated mouse MEC clusters.

(G) 8 annotated clusters of cells isolated from normal human breast.[40]

(H and I) Corresponding *EHF* expression in these clusters (values shown in I are logCPM). LP, luminal progenitor; ML, mature luminal.

morphogenesis, and in the adult NP setting when the mature ductal network has reached the margins of the mammary fat pad.[41]

Ductal elongation is led by highly proliferative terminal end buds (TEBs), consisting of an outer layer of pluripotent stem cells, termed cap cells,[42] and an inner multilayer of undifferentiated

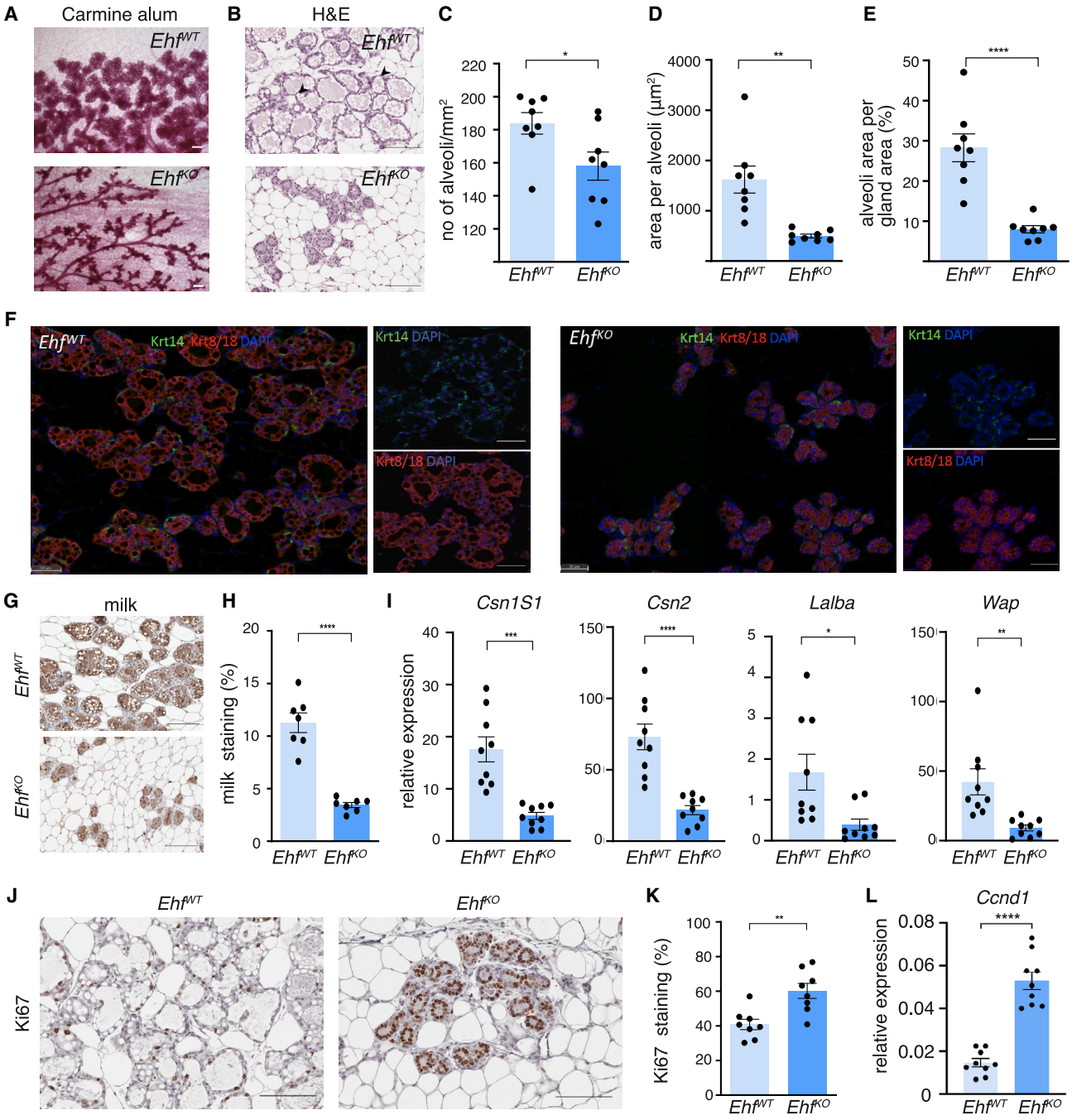


Figure 2. Ehf is indispensable for alveologenesis

(A) Representative whole mounts of inguinal mammary glands from *Ehf*^{WT} and *Ehf*^{KO} mice at 18.5 dP. Scale bars, 0.2 mm.

(B) Representative H&E-stained sections of thoracic mammary glands from *Ehf*^{WT} and *Ehf*^{KO} mice at 18.5 dP; arrows indicate cytoplasmic lipid droplets. Scale bars, 200 μM.

(C–E) Quantification of (C) number of alveoli per mm^2 of mammary gland tissue, (D) average area per alveoli, and (E) alveoli percentage area in mammary gland. Values shown are mean ± SEM of n = 8 mice per genotype.

(F) Representative Krt8/18 and Krt14 immunostained thoracic mammary glands from *Ehf*^{WT} and *Ehf*^{KO} mice at 18.5 dP. Red: Krt8/18; green: Krt14; blue: DAPI. Scale bars, 50 μM.

(G) Representative immunostained thoracic mammary glands from *Ehf*^{WT} and *Ehf*^{KO} mice at 18.5 dP for mouse-milk-specific protein. Scale bars, 200 μM.

(H) Percentage staining of milk protein in whole mammary gland. Values shown are mean ± SEM of n = 7 mice per genotype.

(I) mRNA expression in whole mammary gland tissue from *Ehf*^{WT} and *Ehf*^{KO} mice at 18.5 dP for milk proteins α-*casein* (*Csn1s1*), β-*casein* (*Csn2*), α-*lactalbumin* (*Lalba*), and *whey acidic protein* (*Wap*) expressed relative to β-*actin* (*ACTB*). Values shown are mean ± SEM of n = 9 mice per genotype.

(legend continued on next page)

body cells[43] that drive the filling of the mammary fat pad with an elaborate ductal system.[4] Histological examination of whole-mounted inguinal mammary glands in pubertal mice (6 weeks old) demonstrated that the directionality of TEBs was unaffected by *Ehf* deletion, with all TEBs oriented for forward growth toward the end of the mammary fat pad[44] (Figure S1A). However, the number of TEBs (Figure S1B), and the distance from TEBs to the end of the mammary fat pad, was significantly greater in Ehf^{KO} mice, indicating slower extension of ducts (Figure S1C).

In adult NP mice (10 weeks old), the ductal tree of inguinal glands of both Ehf^{WT} and Ehf^{KO} females was fully elongated (Figure S2A). However, while normal secondary and tertiary branching structures were observed (Figure S2B), a greater number of TEBs were still present in Ehf^{KO} glands (Figure S2C). The total number (Figures S2D and S2E) and average size of ducts (Figures S2D and S2F) was not significantly impacted, whereas the area of the mammary fat pad occupied by ducts was moderately reduced in Ehf^{KO} glands (Figures S2D and S2G). The bilayered ductal structure was not affected by *Ehf* loss, with a single layer of Krt8/18-positive cells enveloping a central lumen surrounded by an outer layer of Krt14-positive basal/myoepithelial cells (Figures S2H and S2I).

From these data, we concluded that as the ductal elongation defect in Ehf^{KO} gland at puberty is largely overcome by the adult stage, it is improbable that the lactation defect of Ehf^{KO} mothers is due to the subtle impact of *Ehf* deletion in the adult NP gland.

Ehf is indispensable for alveologenesis

To investigate the role of *Ehf* in mammary development during pregnancy, we studied the impact of *Ehf* loss at late gestation (18.5 dP), when *Ehf* expression peaks. There was no difference in the age of Ehf^{WT} and Ehf^{KO} females at the time of collection (Figure S3A), and absence of exon 8 of the *Ehf* transcript was confirmed in Ehf^{KO} whole mammary tissue at 18.5 dP (Figure S3B). Histological analyses of whole mounts of the inguinal mammary glands (Figure 2A) and H&E-stained corresponding thoracic mammary glands (Figure 2B) revealed that the mammary glands of Ehf^{WT} females at 18.5 dP consisted of well-differentiated lobuloalveolar units, with the mammary fat pad filled by expanded alveoli containing luminal secretory materials and cytoplasmic lipid droplets (CLDs) as well as lipids secreted into the luminal space.[18] Comparatively, Ehf^{KO} glands contained rudimentary low-density alveolar-like structures with an almost complete absence of CLD, indicative of failed secretory initiation[7] (Figures 2A and 2B). Ehf^{KO} glands also had fewer alveoli, with the mammary tissue largely comprised of adipocytes and stroma (Figures 2B and 2C), reduced area of individual alveoli, consistent with a lack of expansion from luminal secretion (Figure 2D), and reduced alveoli area to gland area ratio (Figure 2E). Consistent with the differences in abundance of CLDs and adipocytes, lipidomic profiling uncovered marked remodeling of the lipidome of the Ehf^{KO} mammary gland, including a significant decrease in monoalkyl-diacylglycerols (TG[O]), as well as alkenylphosphatidylethanolamine (PE[P]) and alkenylphosphatidylcholine (PC[P]), and a number of TG species containing polyunsaturated fatty acids (PUFAs), particularly 20:5, 22:5, 22:6, and 20:4 (Figure S3C).

To determine whether loss of *Ehf* also impacted the cellular composition and structure of alveoli, glands were co-stained with Krt8/18 and Krt14. Ehf^{KO} alveoli were comprised of the expected single layer of Krt8/18-positive cells enveloping a central lumen, surrounded by a discontinuous layer of Krt14-positive myoepithelial cells[42] (Figure 2F), demonstrating that the observed lactation deficiency is not due to perturbation of the bilayered alveoli anatomy.

To confirm that the loss of CLD reflected reduced milk production, Ehf^{WT} and Ehf^{KO} mammary glands were stained for mouse milk protein. Although some milk protein was present in the lumen of Ehf^{KO} alveoli, the percentage of milk within the gland was markedly reduced compared with that in Ehf^{WT} controls (Figures 2G and 2H). Furthermore, real-time quantitative PCR analysis determined significantly reduced mRNA expression of the dominant mouse milk proteins α-*casein* (*Csn1s1*), β-*casein* (*Csn2*), α-*lactalbumin* (*Lalba*), and *whey acidic protein* (*Wap*)[41] in Ehf^{KO} mammary tissue (Figure 2I).

A higher proportion of MECs in the mammary glands of *Ehf*-deficient mice stained positively for the cell proliferation marker, Ki67 (Figures 2J and 2K), which was further supported by an increased expression of the cell cycle regulator, *cyclinD1* (*Ccnd1*) (Figure 2L). Collectively, these findings demonstrate that Ehf^{KO} glands remain blocked in the proliferative phase of alveologenesis at 18.5 dP and identify *Ehf* as a non-redundant determinant of successful lactation.

Ehf deletion impacts an epithelial-cell-autonomous program

As the above findings were established in the context of whole-body *Ehf* deletion, we sought to confirm that the requirement of *Ehf* for alveologenesis is epithelial cell autonomous by deleting *Ehf* specifically in MECs. For this purpose, $Ehf^{fl/fl}$ mice were crossed to $MMTV^{Cre}$ mice to induce recombination in the mammary epithelium from early development.[45,46] As the $MMTV^{Cre}$ transgene may have off-target effects on lactation,[45–47] *Ehf*-mammary-specific knockout (Ehf^{MKO}) mice were studied alongside two control lines, $Ehf^{fl/fl}$ and $MMTV^{Cre}$, at late pregnancy (18.5 dP). There was no difference in the age of females at the time of collection (Figure S3D), and mammary-specific *Ehf* deletion was confirmed by real-time quantitative PCR (Figures S3E–S3G).

Examination of whole-mounted inguinal mammary glands and histological analysis of corresponding thoracic glands confirmed the phenotype observed in mice with constitutive *Ehf* deletion, with mammary glands from Ehf^{MKO} females also containing rudimentary low-density lobuloalveolar structures (Figure 3A) and a loss of secreted lipids and CLDs (Figure 3B). There were fewer alveoli in the mammary gland of Ehf^{MKO} mice compared with

(J) Representative immunostained thoracic mammary glands of Ehf^{WT} and Ehf^{KO} mice at 18.5 dP for the proliferation marker Ki67. Scale bars, 200 μM.
(K) Percentage positive Ki67 staining.
(L) mRNA expression in whole mammary gland tissue from Ehf^{WT} and Ehf^{KO} mice at 18.5 dP for *cyclinD1* (*Ccnd1*). Values shown are mean ± SEM of $n = 9$ mice per genotype. $^{****}p \leq 0.0001$, $^{***}p \leq 0.001$, $^{**}p \leq 0.01$, $^{*}p \leq 0.05$; NS, not significant. Unpaired t test.

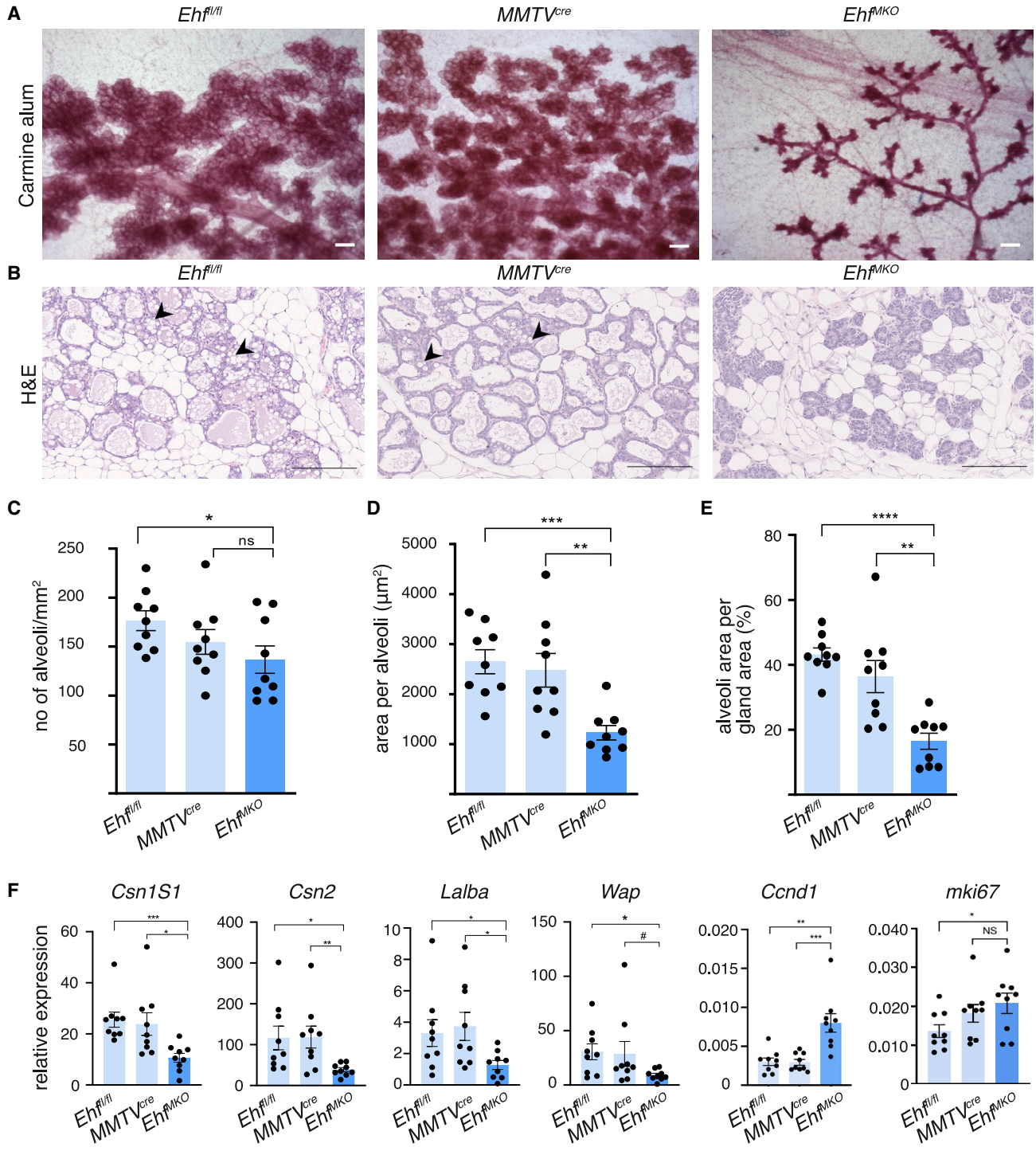


Figure 3. *Ehf* deletion impacts an epithelial-cell-autonomous program

(A) Representative whole mounts of inguinal mammary glands from $Ehf^{fl/fl}$, $MMTV^{Cre}$, and Ehf^{MKO} mice at 18.5 dP. Scale bars, 0.2 mm.

(B) Representative H&E-stained thoracic mammary glands from $Ehf^{fl/fl}$, $MMTV^{Cre}$, and Ehf^{MKO} females at 18.5 dP. Arrows indicate CLDs. Scale bars, 200 µM.

(C–E) Number of alveoli present per mm^2 of mammary gland tissue, (D) average area per alveoli, and (E) alveoli percentage area in mammary gland. Values shown are mean ± SEM of $n = 9$ mice per genotype.

(F) mRNA expression (relative to *β-actin*) of milk protein genes in whole mammary gland tissue from $Ehf^{fl/fl}$, $MMTV^{Cre}$, and Ehf^{MKO} mice at 18.5 dP: *α-casein* (*Csn1s1*), *β-casein* (*Csn2*), *α-lactalbumin* (*Lalba*), *whey acidic protein* (*Wap*), *cyclinD1* (*Ccnd1*), and *Ki67* (*Mki67*). Values shown are mean ± SEM of $n = 9$ mice per genotype. ****$p \leq 0.0001$, *** $p \leq 0.001$ **$p \leq 0.01$, *$p \leq 0.05$; NS, not significant. Unpaired t test. # $p \leq 0.05$, 1-tailed t test.

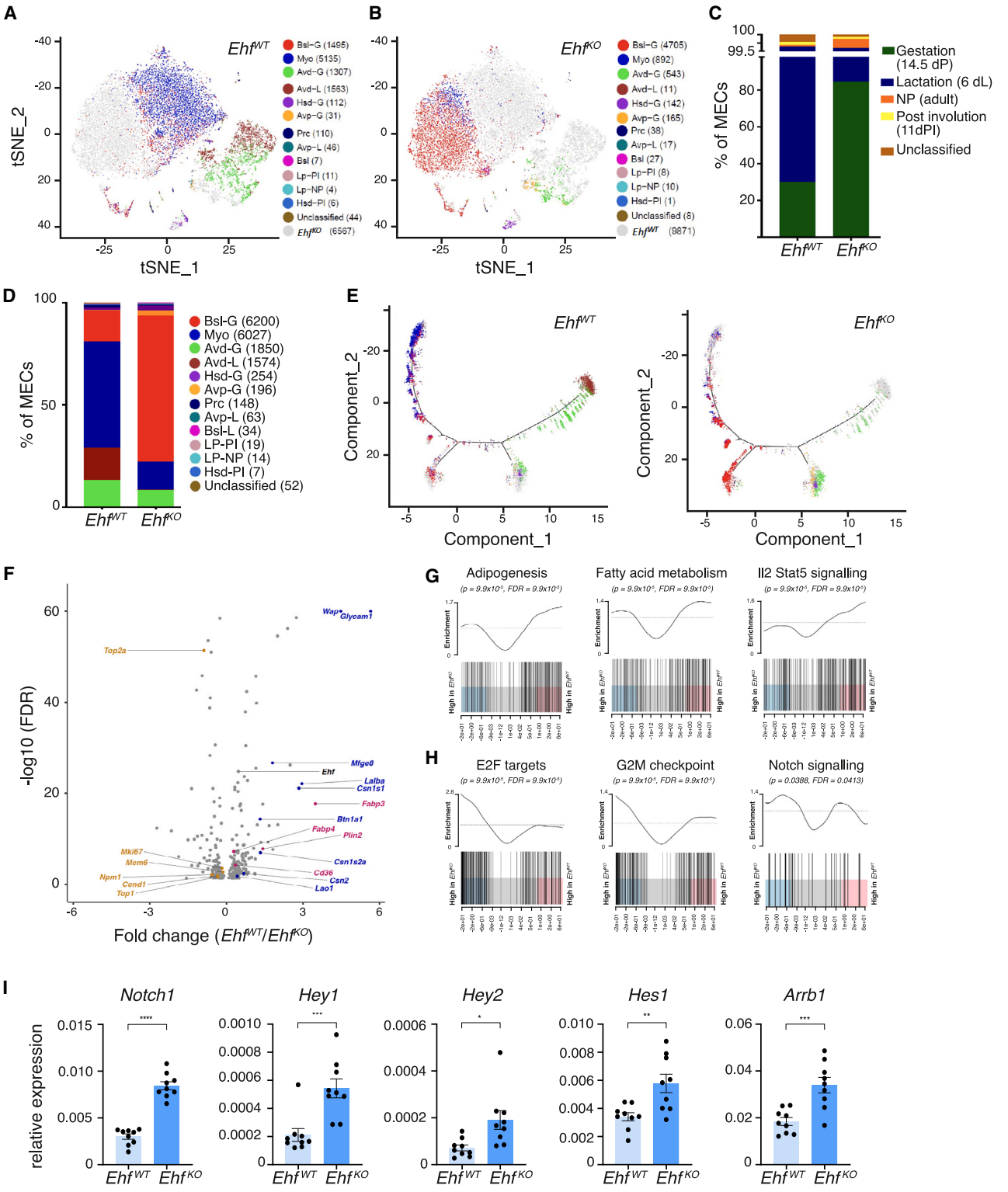


Figure 4. *Ehf* loss alters MEC subtype composition and modifies gene expression profiles at late pregnancy

(A and B) t-SNE plots of total mammary epithelial cells (MECs) from 18.5 dP Ehf^{WT} and Ehf^{KO} mammary gland; distribution of cells colored by identification of epithelial clusters, as defined by Bach et al.[32] Abbreviations defined in the methods section. Numbers in the legends indicate the number of cells identified for each group.

(A) Ehf^{KO} cells in gray.

(B) Ehf^{WT} cells in gray.

(legend continued on next page)

Ehf$^{fl/fl}$ but not *MMTV*^{Cre} glands (Figure 3C), while the size of individual alveoli (Figure 3D) and alveoli area to gland area ratio (Figure 3E) were significantly reduced in *Ehf*^{MKO} mice compared with both the *Ehf*$^{fl/fl}$ and *MMTV*^{Cre} controls. *Csn1s1*, *Csn2*, *Lalba*, and *Wap* mRNA levels were also significantly reduced (Figure 3F) while, conversely, expression of *Ccnd1* and *Mki67* were higher in the mammary tissue of *Ehf*^{MKO} mice (Figure 3F). These results establish that *Ehf* deficiency in the mammary gland during late pregnancy disrupts a mammary epithelial autonomous program.

Ehf^{KO} luminal cells have increased clonogenic potential

The effect of *Ehf* deletion on the clonogenic capacity of isolated luminal MECs (Lin$^-$ [Ter119$^-$CD31$^-$CD45$^-$] CD24$^+$CD29^{lo}) (Figure S4A) from *Ehf*^{WT} and *Ehf*^{KO} mammary glands at 18.5 dP was assessed. Consistent with the observed increase in Ki67 positivity, *Ehf*-deficient luminal cells in late pregnancy had a higher repopulating activity (Figures S4B and S4C), further supporting the existence of a differentiation block that causes cells to remain in the proliferative phase of alveologenesis.

Ehf loss alters MEC subtype composition and modifies gene expression profiles at late pregnancy

To investigate the impact of *Ehf* deletion on the mammary epithelial compartment during late pregnancy, we profiled the MEC population (Figure S4A) from *Ehf*^{WT} and *Ehf*^{KO} mice by scRNA-seq. Following gene and cell filtering, 9,871 *Ehf*^{WT} and 6,567 *Ehf*^{KO} cells were included in the analysis. Effective deletion of *Ehf* was confirmed by the absence of RNA-seq reads corresponding to *Ehf* exon 8 in *Ehf*^{KO} MECs (Figure S5A). Unsupervised cell clustering of the single-cell profiles of the *Ehf*^{WT} and *Ehf*^{KO} samples resulted in 18 clusters, visualized by t-distributed stochastic neighbor embedding (t-SNE) dimension reduction (Figure S5B). The lineage identity of these 18 transcriptional clusters was then interrogated by assessing expression of canonical mouse mammary epithelium lineage markers,[35,40] which identified clusters 1–4, 9, 10, 12, and 14–18 as basal (high expression of *Acta2*, *Krt5*, *Krt14*, *Sparc*, *Cxcl14*, and *Myl9*) (Figure S5C) and clusters 5–8, 11, and 13 as luminal (high expression of *Krt18*, *Krt19*, *Elf5*, *Kit*, *Ehf*, and *Prlr*) (Figure S5D).

The gene expression signature of each *Ehf*^{WT} and *Ehf*^{KO} MEC was further annotated according to one of the 15 MEC subtypes previously identified by Bach et al.[32] (Figures 4A and 4B). In concordance with their isolation during late pregnancy (18.5 dP), >99% of MECs from both *Ehf*^{WT} and *Ehf*^{KO} mammary gland were classified as gestation (G)-stage or lactation (L)-stage subtypes (Figure 4C). However, while the majority of *Ehf*^{WT} MECs were classified as L-stage subtypes (69.4%) (Bsl-L myo, Avd-L, Avp-L, and Prc-L), *Ehf*^{KO} MECs were primarily classified as the preceding G-stage subtypes (84.6%) (Bsl-G, Avd-G, Avp-G, and Hsd-G), indicating an overall delay in differentiation of *Ehf*^{KO} MECs (Figure 4C). More specifically, within the luminal lineage, the *Ehf*^{KO} gland contained a significantly reduced percentage of Avd-L and Avd-G cells and a corresponding increase in the precursory Avp-G population. Similarly, within the Bsl lineage, the proportion of Bsl-L Myo was markedly reduced, with a corresponding increase in the preceding Bsl-G subtype in the *Ehf*^{KO} gland (Figure 4D). To determine the impact of *Ehf* deletion on epithelial cell composition, we performed a pseudotime trajectory analysis along the mammary gland differentiation hierarchy. Overlaying the cell cluster identity with the pseudotiming states demonstrated that *Ehf*^{WT} MECs displayed a typical alveolar differentiation trajectory whereby the Bsl-G population exists at the bottom left axis, differentiating to produce the Bsl-L Myo (top left axis) and Avp-G compartments, followed by Avd-G and Avp-L, which represent intermediate populations in the trajectory toward production of terminally differentiated Avd-L cells (far right axis) (Figure 4E, left panel). In contrast, a high proportion of Bsl-G *Ehf*^{KO} MECs exist along the pseudotime trajectory, failing to terminally differentiate and produce Bsl-L Myo cells, there is also an over-accumulation of Avp-G cells, and lacking in production of Avd-L cells at the end-fate axis (right hand side), indicating a failure to differentiate into mature Avd-L cells (Figure 4E, right panel). These observations were further confirmed when trajectories of *Ehf*^{WT} and *Ehf*^{KO} MECs were plotted individually from unintegrated data (Figure S5E).

In addition to impacting MEC composition, *Ehf* deletion caused substantial transcriptional changes within each lactation and gestational subtype (Table S1). We elected to focus on the Avd-G compartment as a representative population to assess transcriptional changes resulting from *Ehf* deletion. Of the 352 differentially expressed genes (DEGs) in the Avd-G subset, 217 were downregulated and 135 upregulated in the *Ehf*^{KO} gland. Among the genes most significantly downregulated in *Ehf*^{KO}, Avd-G cells were those involved in the synthesis of milk proteins and lipids (*Wap*, *Lalba*, *Glycam1*, *Csn1s1*, *Csn1s2a*, *Csn2*, *Lao1*, *Mfge8*, and *Btn1a1*) and in fatty acid biosynthesis, metabolism, and transport (*Fabp3*, *Fabp4*, *Plin2*, and *Cd36*) (Figure 4F). Gene set enrichment analysis (GSEA) based on the "Hallmark gene sets" in the MSigDB database (Table S2) identified that the most downregulated pathways in *Ehf*^{KO} Avd-G cells included *protein secretion*, *apical junction*, and *apical surface*, signaling pathways critical for mammary gland development (including *IL2/STAT5 signaling*, *IL6/JAK/STAT3 signaling*, *estrogen response early*, *estrogen response late*, and *TGFbeta signaling*[43]), and several metabolic processes (*adipogenesis*, *fatty acid metabolism*, *glycolysis*, and *cholesterol homeostasis*) (Figures 4G and

(C) Percentage of cells classified as gestation stage (14.5 dP), lactation stage (6 dL), NP adult stage, post involution stage (11 dPI), and unclassified in *Ehf*^{WT} and *Ehf*^{KO} mammary gland at 18.5 dP.
(D) Percentage of each cell type. Numbers shown in the legend indicate the total number of cells identified for each group.
(E) Pseudotime analysis of *Ehf*^{WT} and *Ehf*^{KO} MECs along the trajectory of the mammary gland epithelial hierarchy. Left panel: *Ehf*^{KO} cells in gray. Right panel: *Ehf*^{WT} cells in gray. Refer to (A) and (B) for color code.
(F) Volcano plot indicating significantly differentially expressed genes in Avd-G MECs at late pregnancy. Yellow: genes involved in cell cycle progression; blue: synthesis of milk proteins and lipids; pink: fatty acid biosynthesis, metabolism, and transport.
(G and H) Gene set enrichment plots of Hallmarks significantly enriched in (G) *Ehf*^{WT} and (H) *Ehf*^{KO} Avd-G MECs at late pregnancy (18.5 dP).
(I) mRNA expression of the Notch signaling genes *Notch1*, *Hey1*, *Hey2*, *Hes1*, and *Arrb1* (relative to β-*actin*) in whole mammary gland tissue from *Ehf*^{WT} and *Ehf*^{KO} mice at 18.5 dP. Values shown are mean ± SEM of n = 9 mice per genotype. ****$p \leq 0.0001$, *** $p \leq 0.001$ **$p \leq 0.01$, *$p \leq 0.05$. Unpaired t test

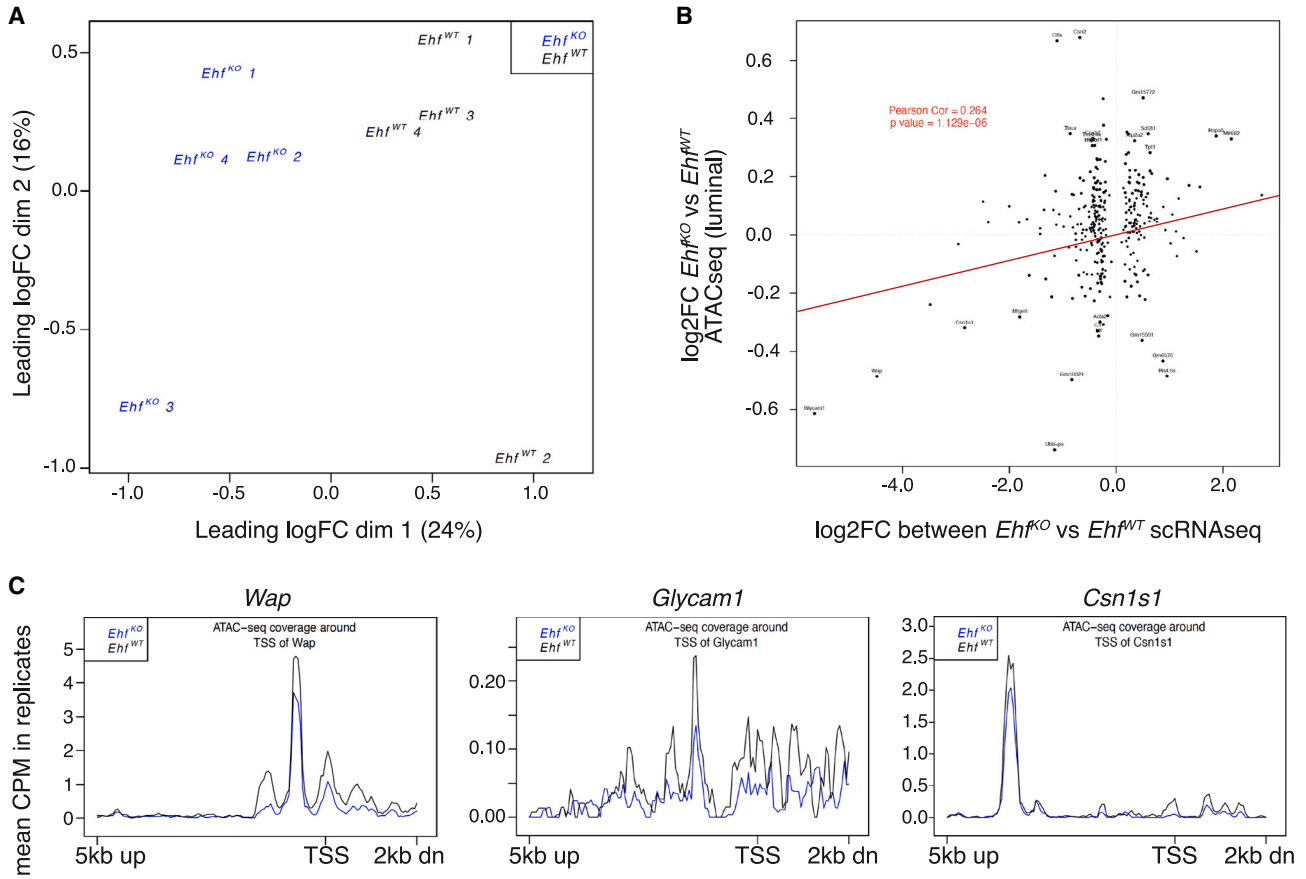


Figure 5. *Ehf* loss induces chromatin accessibility changes in DEGs in Avd-G cells

(A) Multidimensional scaling sample clustering of the ATAC-seq datasets (batch corrected) of luminal MEC samples isolated from 4 Ehf^{WT} and 4 Ehf^{KO} animals at 18.5 dP.

(B) Correlation of mRNA expression (scRNA-seq of Avd-G cells) and corresponding chromatin accessibility changes of the promoter regions (ATAC-seq of luminal cells) of the 331 genes differentially expressed in Avd-G cells between Ehf^{WT} and Ehf^{KO} samples at 18.5 dP.

(C) Representative ATAC-seq read coverage tracks of the regions 5 kb upstream to 2 kb downstream of the TSS of *Wap, Glycam1*, and *Csn1s1* in Ehf^{WT} and Ehf^{KO} luminal MECs at 18.5 dP.

S6A). Conversely, genes significantly upregulated in Ehf^{KO} Avd-G cells included those involved in cell cycle progression (*Mki67*, *Ccnd1*, *Npm1*, *Top1*, *Top2a*, and *Mcm6*) (Figure 4F). This was reflected in the GSEA, which revealed enrichment for *E2F targets*, *G2M checkpoint*, *MYC targets V1* and *V2*, *DNA repair*, *Notch signaling*, and *mitotic spindle* (Figures 4H and S6B). As hyperactive Notch signaling is known to impair alveologenesis,[48] we examined the expression of multiple Notch pathway components and targets, which confirmed increased expression of *Notch1*, *Hey1*, *Hey2*, *Hes1*, and *Arrb1* in Ehf^{KO} glands (Figure 4I).

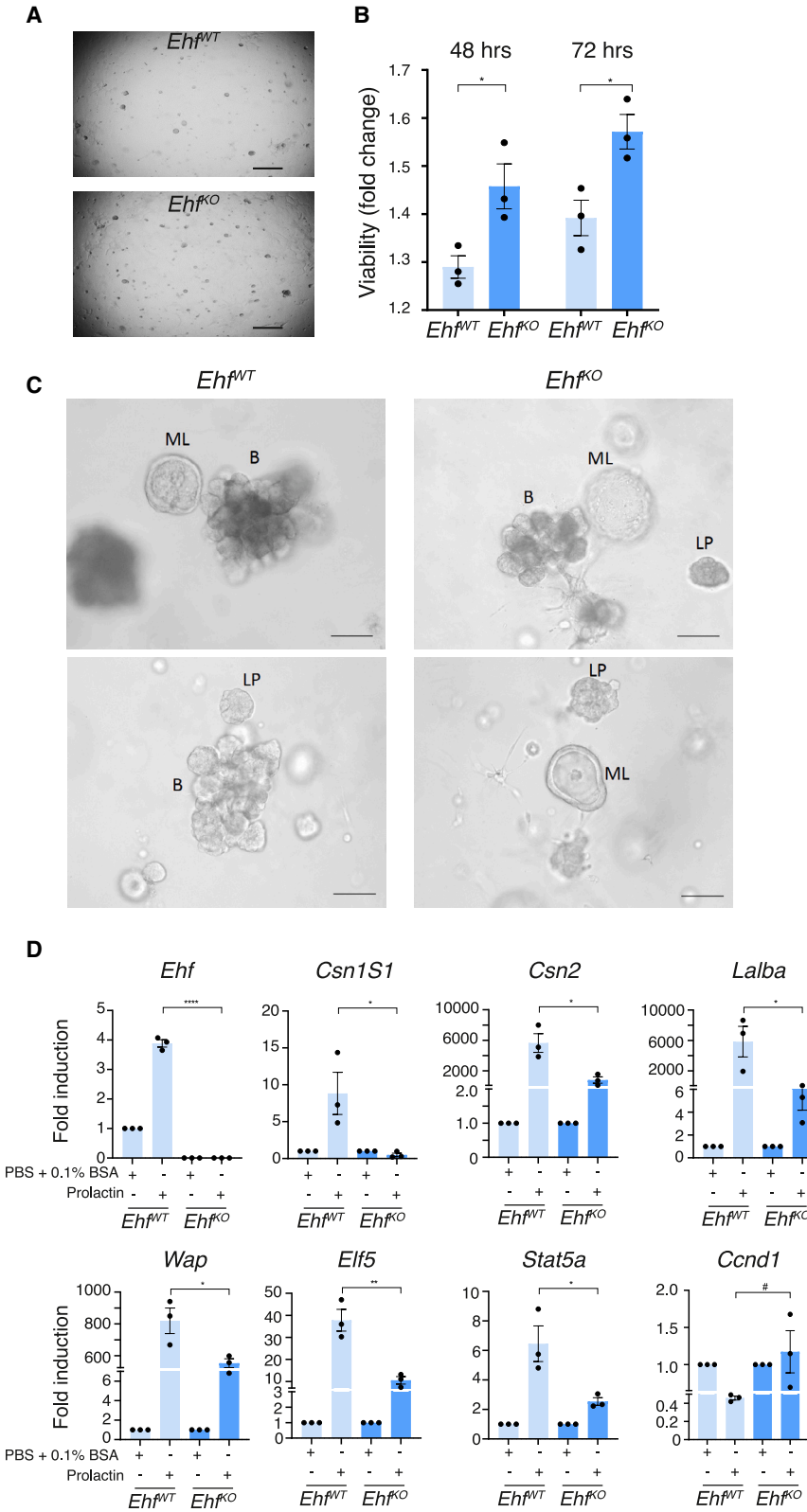
Collectively, these data demonstrate that in addition to reduced prevalence, Ehf^{KO} Avd-G cells display transcriptional profiles indicative of diminished alveolar differentiation and a greater progenitor-like state compared with their Ehf^{WT} counterparts.

Ehf loss induces corresponding chromatin accessibility changes in DEGs in Avd-G cells

To study the corresponding epigenetic changes induced by *Ehf* deletion in luminal MECs, we next performed ATAC-seq profiling of Ehf^{WT} and Ehf^{KO} luminal cells at 18.5 dP. Unsupervised clustering of the samples based on global ATAC-seq profiles resulted in clear segregation of the Ehf^{WT} and Ehf^{KO} luminal samples (Figure 5A), demonstrating that *Ehf* loss induces significant chromatin remodeling. To investigate chromatin accessibility changes in the genes differentially expressed between Ehf^{WT} and Ehf^{KO} Avd-G cells, we performed a correlation analysis between the ATAC-seq and corresponding scRNA-seq datasets. Of the 331 DEGs for which there was corresponding ATAC-seq data, a significant positive correlation (Pearson correlation 0.264, $p \leq 0.0001$) was observed between gene expression and chromatin accessibility (Figure 5B). ATAC-seq tracks showing reduced chromatin accessibility of 3 representative genes downregulated in Ehf^{KO} Avd-G cells are shown (Figure 5C). These data are consistent with *Ehf* loss impacting the alveolar differentiation program through transcriptional regulation.

Ehf-deficient mammary organoids display impaired lactogenic differentiation

To directly verify the lactogenic differentiation defect of Ehf^{KO} MECs, primary mammary organoids were generated from Ehf^{WT}



(legend on next page)

and Ehf^{KO} adult NP mice and their growth capacity assessed *in vitro*. Consistent with an increased proliferative capacity, the number of viable cells was higher in Ehf^{KO} compared with Ehf^{WT} organoids (Figures 6A and 6B). Interestingly, both Ehf^{WT} and Ehf^{KO} MECs were able to produce ML-type organoids (acinar structure and enlarged lumen), LP-type organoids (small size, reduced lumen with some exhibiting rudimentary budding), and basal or stem cell-like (B) organoids (large size, extensive branching morphology) (Figure 6C).[49,50]

We subsequently assessed the ability of Ehf^{WT} and Ehf^{KO} organoids to undergo prolactin-induced lactogenic differentiation. Notably, *Ehf* mRNA was induced 3.9-fold in Ehf^{WT} organoids 96 h post treatment (Figure 6D), indicating that its expression is induced by prolactin. Induction of the known prolactin-regulated genes, *Elf5*, *Stat5a*, *Csn1s1*, *Csn2*, *Lalba*, and *Wap*, and repression of the proliferation marker *Ccnd1*, was significantly attenuated in Ehf^{KO} compared with Ehf^{WT} organoids (Figure 6D), confirming the critical role of *Ehf* in lactogenic differentiation.

Ehf suppresses mammary tumor development in mice

Given its role in promoting differentiation and attenuating proliferation of luminal MECs, we explored whether Ehf could function as a suppressor of mammary tumorigenesis. We addressed this by crossing Ehf^{KO} mice to the *MMTV-PyMT* luminal breast tumor model[51] and assessing tumor incidence. The combined tumor weight at endpoint was not different between Ehf^{WT}*MMTV-PyMT* and Ehf^{KO}*MMTV-PyMT* mice, confirming that the ethical endpoint was accurately determined (Figure S7A), and efficient deletion of *Ehf* in Ehf^{KO}*MMTV-PyMT* mammary tumors was confirmed (Figure S7B). Notably, the delay in ductal elongation observed in Ehf^{KO} mice at puberty was also observed in Ehf^{KO}*MMTV-PyMT* mice (Figures S7C and S7D). Consistent with this delay, median tumor-free survival (82.5 days vs. 76.0 days for Ehf^{KO}*MMTV-PyMT* and Ehf^{WT}*MMTV-PyMT*, respectively, $p < 0.05$) (Figure S7E), and the time taken to reach the ethical endpoint (98.5 days vs. 83.0 days, $p < 0.01$), were increased in Ehf^{KO}*MMTV-PyMT* mice (Figure S7F).

Assessment of mammary tumor number at ethical endpoint demonstrated a significant increase in the number of macroscopic mammary tumors in Ehf^{KO}*MMTV-PyMT* mice (Figures 7A and 7B), suggesting a tumor-suppressive role for Ehf. Consistent with this effect, mammary tumor organoids derived from Ehf^{KO}*MMTV-PyMT* mice displayed increased cell viability compared with Ehf^{WT}*MMTV-PyMT* organoids (Figures 7C and 7D).

Histological assessment of primary mammary tumors resected from Ehf^{WT}*MMTV-PyMT* and Ehf^{KO}*MMTV-PyMT* mice revealed no major differences, with most tumors classified as invasive carcinomas of no special type (NST), with or without a ductal carcinoma *in situ* component (Figure S7G). Molecular characterization based on expression of the estrogen and progesterone receptors (*Esr1*/*Pgr*), *Erbb2*, and established markers of progression in the *MMTV-PyMT* model (*Ccnd1*, *Septin 9*, *Col1a1*, *Chad*, *Ezh2*, and *Gata3*)[51,55] also demonstrated no significant differences between genotypes (Figure S7H).

Low EHF expression is associated with poorer outcome and high tumor grade in human breast cancer

To extend these findings to human breast cancer, we examined associations between EHF expression and molecular subtype, outcome, and histological grade, using publicly available datasets.[52,53] A range of *EHF* expression was observed in each of the molecular subtypes of breast cancer (Figure 7E), with low *EHF* expression associated with poorer outcome in luminal A and basal-like subtypes (Figure 7F). Consistent with *Ehf* deletion impairing cell differentiation and maintaining cells in a proliferative state in the normal mouse mammary epithelium, *EHF* expression was significantly lower in high-grade breast cancers (Figure 7G) and inversely correlated with expression of the proliferation marker *MKI67* (Figure 7H). Finally, analysis of the TCGA breast cancer dataset[54] revealed a significant inverse correlation between *EHF* mRNA expression and methylation, indicating that loss of EHF expression may be epigenetically driven in human breast cancers (Figure 7I).

DISCUSSION

This study defines an essential and previously unknown role for the TF *Ehf* in post-natal mammary gland development and tumor suppression. In addition to its key role in alveolar differentiation during pregnancy, we found that *Ehf* was also important for development of the ductal network in puberty. This observation is consistent with previous data showing high *Ehf* expression in TEBs and mature ducts.[20] Furthermore, in line with the established function of estrogen in driving ductal elongation,[56] *Ehf* has been described as an estrogen-responsive TEB profile gene, with an induction trajectory increasing from pre-puberty to puberty and then declining in the adult gland.[57] However, the specific mechanism by which Ehf regulates ductal elongation, and whether this is aligned with or differs from its effects on alveolar differentiation during pregnancy, remains to be determined.

The impact of *Ehf* deletion on post-natal mammary development and differentiation was highly pronounced in pregnancy. Our findings indicate that Ehf^{KO} glands can undergo the initial phase of the alveologenesis program, as features of this stage, including tertiary branching and the formation of rudimentary alveolar buds, were evident in Ehf^{KO} glands at 18.5 dP.

Figure 6. *Ehf*-deficient mammary organoids display impaired lactogenic differentiation

(A) Representative images of primary mammary Ehf^{WT} and Ehf^{KO} organoids, t = 72 h post seeding. Scale bars, 500 μM.

(B) Viability of mammary organoids; fold-change normalized to t = 24 h post seeding. Values shown are mean ± SEM of n =3 organoid cultures generated from 3 independent animals per genotype.

(C) Representative images of primary mammary Ehf^{WT} and Ehf^{KO} organoids demonstrating structures indicative of mature luminal (ML), luminal progenitor (LP), and basal (B) cell type. Scale bars, 100 μM.

(D) mRNA expression of *Ehf*, milk protein and proliferation genes in primary mammary Ehf^{WT} and Ehf^{KO} organoids 96 h post prolactin treatment. Values shown are mean ± SEM of the fold-change relative to control-treated organoids, from n = 3 organoid cultures generated from 3 independent animals per genotype. $^{****}p \leq 0.0001$, $^{**}p \leq 0.01$, $^{*}p \leq 0.05$. Unpaired t test. $\#p \leq 0.05$, 1-tailed t test.

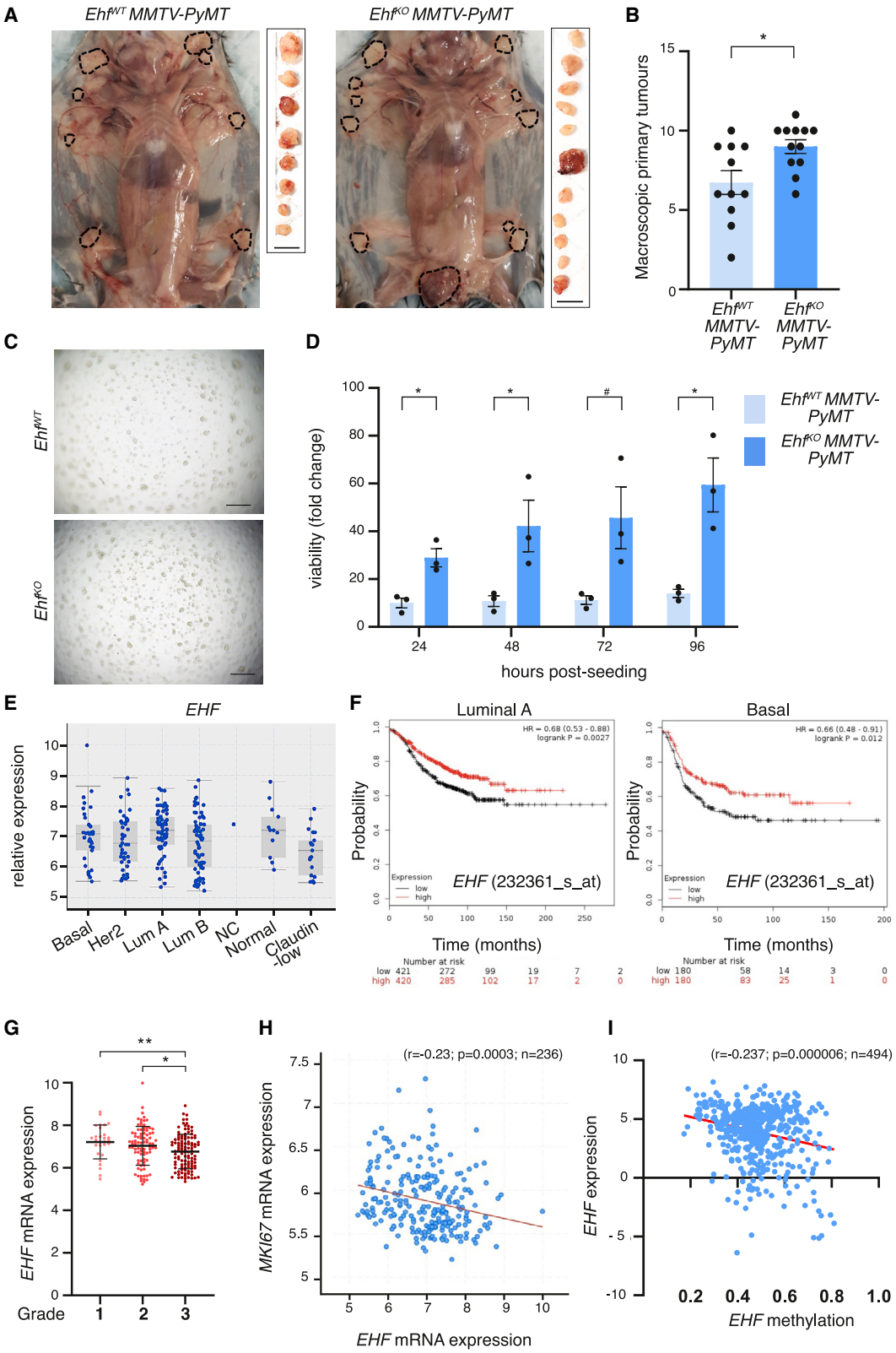


Figure 7. Ehf suppresses mammary tumor development in mice and its expression is associated with tumor grade and patient outcome

(A) Representative images of macroscopic primary mammary tumors *in situ* and following resection (boxed) in *Ehf*^{WT}*MMTV-PyMT* and *Ehf*^{KO}*MMTV-PyMT* mice at an ethically determined endpoint. Scale bars, 10 mm.

(legend continued on next page)

Conversely, the smaller alveoli and absence of CLDs in Ehf^{KO} mammary glands at 18.5 dP, which are classical features of mid-pregnancy,[7,19,41] indicate a failure of Ehf^{KO} MECs to undergo secretory initiation. Furthermore, the shift of MECs from the proliferative phase into the first alveolar differentiation transition is demarcated by a strong decline in the proliferation marker Ki67.[33] Our data indicating that 18.5 dP Ehf^{KO} MECs retain high Ki67 expression and proliferative capacity in clonogenic assays further support the notion that these cells exhibit a block in differentiation and remain in the proliferative phase of early pregnancy. Collectively, these findings demonstrate that *Ehf* is essential for functional alveolar differentiation during pregnancy and that, in the absence of *Ehf*, alveolar cells remain in a proliferative and undifferentiated state.

Interestingly, we observed that the basal cell population (Bsl-G) in Ehf^{KO} MECs expressed lower levels of milk protein and lipid genes at late pregnancy. Basal cells expressing milk protein gene transcripts have been proposed to represent cells transcriptionally primed toward a luminal cell fate.[35] These lineage-primed basal cells have been identified in the pubertal and adult mammary glands, with their prevalence increasing in pregnancy. The previous identification of *Ehf* enrichment in lineage-primed MECs,[35] paired with our observation that *Ehf* deletion causes overaccumulation of Bsl-G, suggests that Ehf may also contribute to priming of the basal compartment to a luminal cell fate in pregnancy.

One of the pathways predicted to be enriched in Ehf^{KO} MECs was Notch signaling. As sustained Notch signaling has been shown to impair alveologenesis through expansion of the LP population at the expense of differentiation,[48] *Ehf* deletion may impede alveolar differentiation, at least in part, through failed suppression of Notch signaling. Notably, constitutive Notch signaling also promotes the self-renewal and oncogenic transformation of LPs,[9] and may potentially also underpin the increased tumor incidence observed in Ehf^{KO}*MMTV-PyMT* mice. Whether sustained Notch signaling is the mechanism driving the disrupted alveologenesis and increased mammary tumorigenesis induced by *Ehf* deletion would therefore be worthy of further investigation.

Ehf shares a highly homologous DNA-binding domain with the related ESE TFs *Elf5* and *Elf3*.[27] Comparable with our findings, $Elf5^{HetKO}$ mice exhibit a lactation defect due to failed alveologenesis,[19] and *Elf5* has been shown to be an alveolar-lineage-specific regulator critical for both proliferation in early pregnancy and differentiation in late pregnancy.[17] Similar to *Ehf* deletion, heterozygous deletion of *Elf5* causes an overaccumulation of the MaSC-enriched and LP populations during pregnancy, which are blocked in differentiation, as well as an increase in mammary stem cell activity.[10] The functions of the remaining ESE-family TFs, *Elf3*, and *SPDEF* in post-natal mammary tissue have not been reported; however, the mammary developmental defect induced by *Ehf* deletion alone demonstrates that there is no functional redundancy between *Ehf* and the other ESE family members.

Nevertheless, it remains likely that EHF and ELF5 may cooperate to regulate a mammary differentiation program; EHF and ELF5 overlap in their expression pattern and biological function[27] and are also contiguously located on chromosome 11p13 in humans and chromosome 2 in mice.[58] Furthermore, *in silico* analyses identified EHF and ELF5 as the predominant TF regulatory network enriched in LP cells.[29] *Elf5* has also been shown to directly regulate *Ehf* in mammary ductal carcinoma cells.[59] Comparatively, in this study we observed that *Elf5* induction in response to prolactin treatment was markedly attenuated in Ehf^{KO} organoids, suggesting Ehf may also regulate *Elf5*. Conversely however, *Elf5* expression was increased in Ehf^{KO}-differentiated alveolar cells (Table S1), which may reflect an attempt of *Elf5* to compensate for *Ehf* loss. In this regard, while *Ehf* deletion resulted in complete neonate mortality of pups born to first litters nursed by Ehf^{KO} females, this phenotype was partially rescued by a subsequent parturition, a phenomenon that has been observed in other knockout models.[12,18] Whether this is due to increased expression of *Elf5* or other factors remains to be determined.

Consistent with its role in suppressing proliferation and promoting alveolar differentiation, this study also provides evidence that *Ehf* functions as a suppressor of mammary tumorigenesis. While whole-body Ehf^{KO} mice were used in this study, the known epithelium-specific expression of *Ehf*,[39] coupled with the increased growth capacity of Ehf^{KO}*MMTV-PyMT* tumor organoids, suggests that this effect is at least partially driven by disruption of the epithelial-cell-autonomous functions of *Ehf* in MECs. However, the possibility that the tumor microenvironment may contribute to tumor development and progression cannot be ruled out.

The potential tumor-suppressive role of EHF identified herein is also consistent with findings in the human breast cancer cell lines MDA-MB-231 and MDA-MB-453, and 4T1 mouse mammary tumor cells, where EHF overexpression attenuated cell proliferation, migration, invasion and metastasis, and epithelial

(B) Number of macroscopic primary mammary tumors in Ehf^{WT}*MMTV-PyMT* and Ehf^{KO}*MMTV-PyMT* mice at ethically determined endpoint. Values shown are mean ± SEM of $n = 11$ Ehf^{WT}*MMTV-PyMT* and 12 Ehf^{KO}*MMTV-PyMT* mice. $^{*}p < 0.05$. Unpaired t test.
(C) Representative primary mammary tumor organoids from Ehf^{WT}*MMTV-PyMT* and Ehf^{KO}*MMTV-PyMT* mice 96 h post seeding. Scale bars, 500 μM.
(D) Cell viability in mammary organoids from Ehf^{WT}*MMTV-PyMT* and Ehf^{KO}*MMTV-PyMT* mice. Values shown are mean ± SEM of the fold-change in cell viability normalized to t = 2 h post seeding from $n = 3$ organoid cultures generated from 3 independent animals per genotype. $^{*}p < 0.05$. Unpaired t test. $\#p \leq 0.05$, 1-tailed t test.
(E) *EHF* mRNA expression in METABRIC trial dataset[52] from intrinsic histological human breast cancer subtypes, basal-like (Basal), Her2 enriched (Her2), luminal A (Lum A), luminal B (Lum B), normal-like (Normal), claudin-low, and not classified (NC).
(F) Correlation of EHF expression with patient-relapse-free survival in luminal A and basal-like breast cancers, accessed from KM Plotter.[53]
(G) *EHF* mRNA expression according to histological grade of human breast cancers. Data shown are mean ± SEM mRNA expression values assessed by microarray in the METABRIC trial[52] of $n = 28$ grade 1, $n = 88$ grade 2, and $n = 111$ grade 3 breast cancers. $^{*}p < 0.05$, $^{**}p < 0.005$. Unpaired t test.
(H) Correlation of *EHF* mRNA expression with the proliferation marker *MKI67* across human breast cancer samples ($n = 236$, $r = -0.233$, $p = 0.0003$, Pearson's correlation).
(I) *EHF* mRNA expression vs. methylation in the TCGA dataset[54] (all cases, $n = 494$; $r = -0.237$, $p < 0.0001$, Pearson's correlation).

to mesenchymal transition, and promoted tumor cell apoptosis and cellular senescence.[37,38] A tumor-suppressive function for EHF has also been documented in cell-line and animal models of prostate,[60–62] pancreatic,[63,64] and colon carcinoma.[39,65]

Limitations of the study

While our findings establish a non-redundant role for *Ehf* in alveolar differentiation and tumor suppression in the mouse mammary gland, the study has limitations. Although the targeted deletion of the DNA-binding Ets domain of *Ehf* in this model enables us to conclude that the observed phenotype is due to loss of *Ehf* transcriptional activity,[39] the lack of an antibody to reliably detect mouse EHF protein precludes us from determining whether expression of the EHF protein is completely abolished. It therefore remains possible that other activities of EHF, such as protein-protein interactions mediated through the PNT domain, may be retained.

We further note that, in seeming contrast to the increased tumor incidence, tumor-free survival and the time to ethical endpoint were also increased in Ehf^{KO}*MMTV-PyMT* mice. However, we reason that this is likely due to the confounding effect of *Ehf* loss delaying development of the ductal epithelium and thereby delaying the onset of mammary tumors. Additional mouse models, whereby *Ehf* is deleted specifically in the adult gland, may therefore be required to further validate these findings.

In summary, this study demonstrates a previously unknown role for the *Ehf* TF in alveolar differentiation in the mammary gland, as well as acting as a putative suppressor of mammary tumorigenesis. These findings add a key element to our growing body of knowledge of the mechanism by which development of this critical tissue is regulated during pregnancy and the mechanisms that protect this tissue from malignant transformation.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.devcel.2024.04.022.

ACKNOWLEDGMENTS

This project was supported in part by an Austin Health Medical Research Foundation (AMRF) grant (2–1293) and the Operational Infrastructure Support Program, Victorian Government, Australia. C.M.R., N.V., S.G., L.J.J. and R.J.M. were supported by La Trobe University Australian Postgraduate Scholarships. The ONJCRI microscopy platform is supported by an equipment grant from the Australian Cancer Research Foundation, 2016. Y.C. was supported by an MRFF Investigator grant (1176199). P.J.M. was supported by an NHMRC L3 Investigator grant (2009965). J.E.V. was supported by an NHMRC Fellowship (1102742). B.P. was supported by a Victorian Cancer Agency (VCA) mid-career fellowship (MCRF21002). J.M.M. was supported by an NHMRC Senior Research Fellowship (1046092).

AUTHOR CONTRIBUTIONS

Conceptualization, R.N., C.M.R., B.P., and J.M.M.; methodology, R.N., C.M.R., N.V., B.P., J.M.M., F.V., S.P., R.J.M., P.J.M., J.E.V., B.P., and J.M.M.; software, Y.L., D.C., Y.C., and W.S.; formal analysis, R.N., C.M.R., N.V., Y.L., F.V., S.P., D.C., Y.C., D.S.W., B.P., and J.M.M.; investigation, R.N., C.M.R., N.V., N.P., S.G., C.B., F.V., S.P., I.Y.L., L.J.J., A.L.C., D.C., Y.C., D.S.W., P.J.M., B.P., and J.M.M.; resources, R.N., C.M.R., Y.L., C.B., F.V., S.P., R.J.M., F.C.J., A.L.C., D.C., Y.C., D.S.W., S.E., P.J.M., W.S., J.E.V., B.P., and J.M.M.; data curation, Y.L., D.C., Y.C., and W.S.; writing – original draft, R.N., C.M.R., B.P., and J.M.M.; writing – review & editing, R.N., C.M.R., N.V., Y.L., F.V., S.P., A.S.D., R.L.A., S.E., P.J.M., W.S., J.E.V., B.P., and J.M.M.; visualization, R.N., C.M.R., N.V., N.P., Y.L., F.V., S.P., Y.C., P.J.M., W.S., B.P., and J.M.M.; supervision, R.N., A.S.D., R.L.A., P.J.M., W.S., J.E.V., B.P., and J.M.M.; funding acquisition, C.M.R., W.S., J.E.V., B.P., and J.M.M.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 17, 2022
Revised: March 3, 2024
Accepted: April 30, 2024
Published: May 22, 2024

REFERENCES

1. Rudolph, M.C., Russell, T.D., Webb, P., Neville, M.C., and Anderson, S.M. (2011). Prolactin-mediated regulation of lipid biosynthesis genes in vivo in the lactating mammary epithelial cell. Am. J. Physiol. Endocrinol. Metab. *300*, E1059–E1068. https://doi.org/10.1152/ajpendo.00083.2011.
2. Robinson, G.W., Karpf, A.B.C., and Kratochwil, K. (1999). Regulation of Mammary Gland Development by Tissue Interaction. J. Mammary Gland Biol. Neoplasia *4*, 9–19. https://doi.org/10.1023/A:1018748418447.
3. Fu, N.Y., Nolan, E., Lindeman, G.J., and Visvader, J.E. (2020). Stem Cells and the Differentiation Hierarchy in Mammary Gland Development. Physiol. Rev. *100*, 489–523. https://doi.org/10.1152/physrev.00040.2018.
4. Inman, J.L., Robertson, C., Mott, J.D., and Bissell, M.J. (2015). Mammary gland development: cell fate specification, stem cells and the microenvironment. Development *142*, 1028–1042. https://doi.org/10.1242/dev.087643.

5. Watson, C.J. (2022). Alveolar cells in the mammary gland: lineage commitment and cell death. Biochem. J. *479*, 995–1006. https://doi.org/10.1042/BCJ20210734.

6. Oakes, S.R., Hilton, H.N., and Ormandy, C.J. (2006). Key stages in mammary gland development - The alveolar switch: coordinating the proliferative cues and cell fate decisions that drive the formation of lobuloalveoli from ductal epithelium. Breast Cancer Res. *8*, 207. https://doi.org/10.1186/bcr1411.

7. McManaman, J.L. (2009). Formation of milk lipids: a molecular perspective. Clin. Lipidol. *4*, 391–401. https://doi.org/10.2217/clp.09.15.

8. Brisken, C., and Rajaram, R.D. (2006). Alveolar and lactogenic differentiation. J. Mammary Gland Biol. Neoplasia *11*, 239–248. https://doi.org/10.1007/s10911-006-9026-0.

9. Bouras, T., Pal, B., Vaillant, F., Harburg, G., Asselin-Labat, M.L., Oakes, S.R., Lindeman, G.J., and Visvader, J.E. (2008). Notch Signaling Regulates Mammary Stem Cell Function and Luminal Cell-Fate Commitment. Cell Stem Cell *3*, 429–441. https://doi.org/10.1016/j.stem.2008.08.001.

10. Chakrabarti, R., Wei, Y., Romano, R.A., DeCoste, C., Kang, Y., and Sinha, S. (2012). Elf5 regulates mammary gland stem/progenitor cell fate by influencing notch signaling. Stem Cells *30*, 1496–1508. https://doi.org/10.1002/stem.1112.

11. 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. https://doi.org/10.1101/gad.11.2.179.

12. Liu, X., Gallego, M.I., Smith, G.H., Robinson, G.W., and Hennighausen, L. (1998). Functional rescue of Stat5a-null mammary tissue through the activation of compensating signals including Stat5b. Cell Growth Differ. *9*, 795–803.

13. Ormandy, C.J., Camus, A., Barra, J., Damotte, D., Lucas, B., Buteau, H., Edery, M., Brousse, N., Babinet, C., Binart, N., and Kelly, P.A. (1997). Null mutation of the prolactin receptor gene produces multiple reproductive defects in the mouse. Genes Dev. *11*, 167–178. https://doi.org/10.1101/gad.11.2.167.

14. Ormandy, C.J., Naylor, M., Harris, J., Robertson, F., Horseman, N.D., Lindeman, G.J., Visvader, J., and Kelly, P.A. (2003). Investigation of the transcriptional changes underlying functional defects in the mammary glands of prolactin receptor knockout mice. Recent Prog. Horm. Res. *58*, 297–323. https://doi.org/10.1210/rp.58.1.297.

15. Lee, H.J., Gallego-Ortega, D., Ledger, A., Schramek, D., Joshi, P., Szwarc, M.M., Cho, C., Lydon, J.P., Khokha, R., Penninger, J.M., and Ormandy, C.J. (2013). Progesterone drives mammary secretory differentiation via RankL-mediated induction of Elf5 in luminal progenitor cells. Development *140*, 1397–1401. https://doi.org/10.1242/dev.088948.

16. Jones, F.E., Welte, T., Fu, X.Y., and Stern, D.F. (1999). Erbb4 Signaling in the Mammary Gland Is Required for Lobuloalveolar Development and Stat5 Activation during Lactation. J. Cell Biol. *147*, 77–88. https://doi.org/10.1083/jcb.147.1.77.

17. Oakes, S.R., Naylor, M.J., Asselin-Labat, M.L., Blazek, K.D., Gardiner-Garden, M., Hilton, H.N., Kazlauskas, M., Pritchard, M.A., Chodosh, L.A., Pfeffer, P.L., et al. (2008). The Ets transcription factor Elf5 specifies mammary alveolar cell fate. Genes Dev. *22*, 581–586. https://doi.org/10.1101/gad.1614608.

18. Choi, Y.S., Chakrabarti, R., Escamilla-Hernandez, R., and Sinha, S. (2009). Elf5 conditional knockout mice reveal its role as a master regulator in mammary alveolar development: Failure of Stat5 activation and functional differentiation in the absence of Elf5. Dev. Biol. *329*, 227–241. https://doi.org/10.1016/j.ydbio.2009.02.032.

19. Zhou, J., Chehab, R., Tkalcevic, J., Naylor, M.J., Harris, J., Wilson, T.J., Tsao, S., Tellis, I., Zavarsek, S., Xu, D., et al. (2005). Elf5 is essential for early embryogenesis and mammary gland development during pregnancy and lactation. EMBO J. *24*, 635–644. https://doi.org/10.1038/sj.emboj.7600538.

20. Kouros-Mehr, H., Slorach, E.M., Sternlicht, M.D., and Werb, Z. (2006). GATA-3 Maintains the Differentiation of the Luminal Cell Fate in the Mammary Gland. Cell *127*, 1041–1055. https://doi.org/10.1016/j.cell.2006.09.048.

21. Fu, N.Y., Pal, B., Chen, Y., Jackling, F.C., Milevskiy, M., Vaillant, F., Capaldo, B.D., Guo, F., Liu, K.H., Rios, A.C., et al. (2018). Foxp1 Is Indispensable for Ductal Morphogenesis and Controls the Exit of Mammary Stem Cells from Quiescence. Dev. Cell *47*, 629–644.e8. https://doi.org/10.1016/j.devcel.2018.10.001.

22. Seagroves, T.N., Krnacik, S., Raught, B., Gay, J., Burgess-Beusse, B., Darlington, G.J., and Rosen, J.M. (1998). C/EBPβ, but not C/EBPα, is essential for ductal morphogenesis, lobuloalveolar proliferation, and functional differentiation in the mouse mammary gland. Genes Dev. *12*, 1917–1928. https://doi.org/10.1101/gad.12.12.1917.

23. LaMarca, H.L., Visbal, A.P., Creighton, C.J., Liu, H., Zhang, Y., Behbod, F., and Rosen, J.M. (2010). CCAAT/enhancer binding protein beta regulates stem cell activity and specifies luminal cell fate in the mammary gland. Stem Cells *28*, 535–544. https://doi.org/10.1002/stem.297.

24. Elias, S., Morgan, M.A., Bikoff, E.K., and Robertson, E.J. (2017). Long-lived unipotent Blimp1-positive luminal stem cells drive mammary gland organogenesis throughout adult life. Nat. Commun. *8*, 1714. https://doi.org/10.1038/s41467-017-01971-w.

25. Sizemore, G.M., Pitarresi, J.R., Balakrishnan, S., and Ostrowski, M.C. (2017). The ETS family of oncogenic transcription factors in solid tumours. Nat. Rev. Cancer *17*, 337–351. https://doi.org/10.1038/nrc.2017.20.

26. Sharrocks, A.D. (2001). The ETS-domain transcription factor family. Nat. Rev. Mol. Cell Biol. *2*, 827–837. https://doi.org/10.1038/35099076.

27. Luk, I.Y., Reehorst, C.M., and Mariadason, J.M. (2018). ELF3, ELF5, EHF and SPDEF Transcription Factors in Tissue Homeostasis and Cancer. Molecules *23*, 2191. https://doi.org/10.3390/molecules23092191.

28. Madison, B.J., Clark, K.A., Bhachech, N., Hollenhorst, P.C., Graves, B.J., and Currie, S.L. (2018). Electrostatic repulsion causes anticooperative DNA binding between tumor suppressor ETS transcription factors and JUN–FOS at composite DNA sites. J. Biol. Chem. *293*, 18624–18635. https://doi.org/10.1074/jbc.RA118.003352.

29. Pellacani, D., Bilenky, M., Kannan, N., Heravi-Moussavi, A., Knapp, D.J.H.F., Gakkhar, S., Moksa, M., Carles, A., Moore, R., Mungall, A.J., et al. (2016). Analysis of Normal Human Mammary Epigenomes Reveals Cell-Specific Active Enhancer States and Associated Transcription Factor Networks. Cell Rep. *17*, 2060–2074. https://doi.org/10.1016/j.celrep.2016.10.058.

30. Lim, E., Wu, D., Pal, B., Bouras, T., Asselin-Labat, M.L., Vaillant, F., Yagita, H., Lindeman, G.J., Smyth, G.K., and Visvader, J.E. (2010). Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways. Breast Cancer Res. *12*, R21. https://doi.org/10.1186/bcr2560.

31. Chung, C.Y., Ma, Z., Dravis, C., Preissl, S., Poirion, O., Luna, G., Hou, X., Giraddi, R.R., Ren, B., and Wahl, G.M. (2019). Single-Cell Chromatin Analysis of Mammary Gland Development Reveals Cell-State Transcriptional Regulators and Lineage Relationships. Cell Rep. *29*, 495–510.e6. https://doi.org/10.1016/j.celrep.2019.08.089.

32. Bach, K., Pensa, S., Grzelak, M., Hadfield, J., Adams, D.J., Marioni, J.C., and Khaled, W.T. (2017). Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing. Nat. Commun. *8*, 2128. https://doi.org/10.1038/s41467-017-02001-5.

33. Williams, C., Helguero, L., Edvardsson, K., Haldosén, L.A., and Gustafsson, J.A. (2009). Gene expression in murine mammary epithelial stem cell-like cells shows similarities to human breast cancer gene expression. Breast Cancer Res. *11*, R26. https://doi.org/10.1186/bcr2256.

34. Giraddi, R.R., Chung, C.Y., Heinz, R.E., Balcioglu, O., Novotny, M., Trejo, C.L., Dravis, C., Hagos, B.M., Mehrabad, E.M., Rodewald, L.W., et al. (2018). Single-Cell Transcriptomes Distinguish Stem Cell State Changes and Lineage Specification Programs in Early Mammary Gland

Development. Cell Rep. *24*, 1653–1666.e7. https://doi.org/10.1016/j.celrep.2018.07.025.

35. Pal, B., Chen, Y., Vaillant, F., Jamieson, P., Gordon, L., Rios, A.C., Wilcox, S., Fu, N., Liu, K.H., Jackling, F.C., et al. (2017). Construction of developmental lineage relationships in the mouse mammary gland by single-cell RNA profiling. Nat. Commun. *8*, 1627. https://doi.org/10.1038/s41467-017-01560-x.
36. Valdés-Mora, F., Salomon, R., Gloss, B.S., Law, A.M.K., Venhuizen, J., Castillo, L., Murphy, K.J., Magenau, A., Papanicolaou, M., Rodriguez de la Fuente, L., et al. (2021). Single-cell transcriptomics reveals involution mimicry during the specification of the basal breast cancer subtype. Cell Rep. *35*, 108945. https://doi.org/10.1016/j.celrep.2021.108945.
37. Sakamoto, K., Endo, K., Sakamoto, K., Kayamori, K., Ehata, S., Ichikawa, J., Ando, T., Nakamura, R., Kimura, Y., Yoshizawa, K., et al. (2021). EHF suppresses cancer progression by inhibiting ETS1-mediated ZEB expression. Oncogenesis *10*, 26. https://doi.org/10.1038/s41389-021-00313-2.
38. Lim, S., Lim, J., Lee, A., Kim, K.I., and Lim, J.S. (2023). Anticancer Effect of E26 Transformation-Specific Homologous Factor through the Induction of Senescence and the Inhibition of Epithelial–Mesenchymal Transition in Triple-Negative Breast Cancer Cells. Cancers (Basel) *15*, 5270. https://doi.org/10.3390/cancers15215270.
39. Reehorst, C.M., Nightingale, R., Luk, I.Y., Jenkins, L., Koentgen, F., Williams, D.S., Darido, C., Tan, F., Anderton, H., Chopin, M., et al. (2021). EHF is essential for epidermal and colonic epithelial homeostasis, and suppresses Apc-initiated colonic tumorigenesis. Development *148*, dev199542. https://doi.org/10.1242/dev.199542.
40. Pal, B., Chen, Y., Milevskiy, M.J.G., Vaillant, F., Prokopuk, L., Dawson, C.A., Capaldo, B.D., Song, X., Jackling, F., Timpson, P., et al. (2021). Single cell transcriptome atlas of mouse mammary epithelial cells across development. Breast Cancer Res. *23*, 69. https://doi.org/10.1186/s13058-021-01445-4.
41. Anderson, S.M., Rudolph, M.C., McManaman, J.L., and Neville, M.C. (2007). Key stages in mammary gland development. Secretory activation in the mammary gland: it's not just about milk protein synthesis! Breast Cancer Res. *9*, 204. https://doi.org/10.1186/bcr1653.
42. Richert, M.M., Schwertfeger, K.L., Ryder, J.W., and Anderson, S.M. (2000). An atlas of mouse mammary gland development. J. Mammary Gland Biol. Neoplasia *5*, 227–241. https://doi.org/10.1023/a:1026499523505.
43. Slepicka, P.F., Somasundara, A.V.H., and Dos Santos, C.O. (2021). The molecular basis of mammary gland development and epithelial differentiation. Semin. Cell Dev. Biol. *114*, 93–112. https://doi.org/10.1016/j.semcdb.2020.09.014.
44. Paine, I.S., and Lewis, M.T. (2017). The Terminal End Bud: the Little Engine that Could. J. Mammary Gland Biol. Neoplasia *22*, 93–108. https://doi.org/10.1007/s10911-017-9372-0.
45. Wagner, K.U., Wall, R.J., St-Onge, L., Gruss, P., Wynshaw-Boris, A., Garrett, L., Li, M., Furth, P.A., and Hennighausen, L. (1997). Cre-mediated gene deletion in the mammary gland. Nucleic Acids Res. *25*, 4323–4330. https://doi.org/10.1093/nar/25.21.4323.
46. Wagner, K.U., McAllister, K., Ward, T., Davis, B., Wiseman, R., and Hennighausen, L. (2001). Spatial and temporal expression of the Cre gene under the control of the MMTV-LTR in different lines of transgenic mice. Transgenic Res. *10*, 545–553. https://doi.org/10.1023/a:1013063514007.
47. Robinson, G.W., and Hennighausen, L. (2011). MMTV-Cre transgenes can adversely affect lactation: Considerations for conditional gene deletion in mammary tissue. Anal. Biochem. *412*, 92–95. https://doi.org/10.1016/j.ab.2011.01.020.
48. Cazares, O., Chatterjee, S., Lee, P., Strietzel, C., Bubolz, J.W., Harburg, G., Howard, J., Katzman, S., Sanford, J., and Hinck, L. (2021). Alveolar progenitor differentiation and lactation depends on paracrine inhibition of Notch via ROBO1/CTNNB1/JAG1. Development *148*, dev199940. https://doi.org/10.1242/dev.199940.
49. Dekkers, J.F., van Vliet, E.J., Sachs, N., Rosenbluth, J.M., Kopper, O., Rebel, H.G., Wehrens, E.J., Piani, C., Visvader, J.E., Verissimo, C.S., et al. (2021). Long-term culture, genetic manipulation and xenotransplantation of human normal and breast cancer organoids. Nat. Protoc. *16*, 1936–1965. https://doi.org/10.1038/s41596-020-00474-1.
50. Rosenbluth, J.M., Schackmann, R.C.J., Gray, G.K., Selfors, L.M., Li, C.M.C., Boedicker, M., Kuiken, H.J., Richardson, A., Brock, J., Garber, J., et al. (2020). Organoid cultures from normal and cancer-prone human breast tissues preserve complex epithelial lineages. Nat Commun *11*, 1711. https://doi.org/10.1038/s41467-020-15548-7.
51. Attalla, S., Taifour, T., Bui, T., and Muller, W. (2021). Insights from transgenic mouse models of PyMT-induced breast cancer: recapitulating human breast cancer progression in vivo. Oncogene *40*, 475–491. https://doi.org/10.1038/s41388-020-01560-0.
52. Curtis, C., Shah, S.P., Chin, S.F., Turashvili, G., Rueda, O.M., Dunning, M.J., Speed, D., Lynch, A.G., Samarajiwa, S., Yuan, Y., et al. (2012). The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature *486*, 346–352. https://doi.org/10.1038/nature10983.
53. Győrffy, B. (2021). Survival analysis across the entire transcriptome identifies biomarkers with the highest prognostic power in breast cancer. Comput. Struct. Biotechnol. J. *19*, 4101–4109. https://doi.org/10.1016/j.csbj.2021.07.014.
54. Koboldt, D.C., Fulton, R.S., McLellan, M.D., Schmidt, H., Kalicki-Veizer, J., McMichael, J.F., Fulton, L.L., Dooling, D.J., Ding, L., Mardis, E.R., et al. (2012). Comprehensive molecular portraits of human breast tumours. Nature *490*, 61–70. https://doi.org/10.1038/nature11412.
55. Lin, E.Y., Jones, J.G., Li, P., Zhu, L., Whitney, K.D., Muller, W.J., and Pollard, J.W. (2003). Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases. Am. J. Pathol. *163*, 2113–2126. https://doi.org/10.1016/S0002-9440(10)63568-7.
56. Biswas, S.K., Banerjee, S., Baker, G.W., Kuo, C.-Y., and Chowdhury, I. (2022). The Mammary Gland: Basic Structure and Molecular Signaling during Development. Int. J. Mol. Sci. *23*, 3883. https://doi.org/10.3390/ijms23073883.
57. McBryan, J., Howlin, J., Kenny, P.A., Shioda, T., and Martin, F. (2007). ERα-CITED1 co-regulated genes expressed during pubertal mammary gland development: Implications for breast cancer prognosis. Oncogene *26*, 6406–6419. https://doi.org/10.1038/sj.onc.1210468.
58. Tugores, A., Le, J., Sorokina, I., Snijders, A.J., Duyao, M., Reddy, P.S., Carlée, L., Ronshaugen, M., Mushegian, A., Watanaskul, T., et al. (2001). The Epithelium-specific ETS Protein EHF/ESE-3 is a Context-dependent Transcriptional Repressor Downstream of MAPK Signaling Cascades. J. Biol. Chem. *276*, 20397–20406. https://doi.org/10.1074/jbc.M010930200.
59. Swahn, H., Sabith Ebron, J., Lamar, K.M., Yin, S., Kerschner, J.L., NandyMazumdar, M., Coppola, C., Mendenhall, E.M., Leir, S.H., and Harris, A. (2019). Coordinate regulation of ELF5 and EHF at the chr11p13 CF modifier region. J. Cell. Mol. Med. *23*, 7726–7740. https://doi.org/10.1111/jcmm.14646.
60. Cangemi, R., Mensah, A., Albertini, V., Jain, A., Mello-Grand, M., Chiorino, G., Catapano, C.V.v., and Carbone, G.M. (2008). Reduced expression and tumor suppressor function of the ETS transcription factor ESE-3 in prostate cancer. Oncogene *27*, 2877–2885. https://doi.org/10.1038/sj.onc.1210953.
61. Albino, D., Longoni, N., Curti, L., Mello-Grand, M., Pinton, S., Civenni, G., Thalmann, G., D'Ambrosio, G., Sarti, M., Sessa, F., et al. (2012). ESE3/EHF controls epithelial cell differentiation and its loss leads to prostate tumors with mesenchymal and stem-like features. Cancer Res. *72*, 2889–2900. https://doi.org/10.1158/0008-5472.CAN-12-0212.
62. Albino, D., Civenni, G., Rossi, S., Mitra, A., Catapano, C.V., and Carbone, G.M. (2016). The ETS factor ESE3/EHF represses IL-6 preventing STAT3 activation and expansion of the prostate cancer stem-like compartment. Oncotarget *7*, 76756–76768. https://doi.org/10.18632/oncotarget.12525.
63. Liu, J., Jiang, W., Zhao, K., Wang, H., Zhou, T., Bai, W., Wang, X., Zhao, T., Huang, C., Gao, S., et al. (2019). Tumoral EHF predicts the efficacy of anti-PD1 therapy in pancreatic ductal adenocarcinoma. J. Exp. Med. *216*, 656–673. https://doi.org/10.1084/jem.20180749.

64. Zhao, T., Jiang, W., Wang, X., Wang, H., Zheng, C., Li, Y., Sun, Y., Huang, C., Han, Z.B., Yang, S., et al. (2017). ESE3 inhibits pancreatic cancer metastasis by upregulating E-cadherin. Cancer Res. *77*, 874–885. https://doi.org/10.1158/0008-5472.CAN-16-2170.

65. Taniue, K., Oda, T., Hayashi, T., Okuno, M., and Akiyama, T. (2011). A member of the ETS family, EHF, and the ATPase RUVBL1 inhibit p53-mediated apoptosis. EMBO Rep. *12*, 682–689. https://doi.org/10.1038/embor.2011.81.

66. Guy, C.T., Cardiff, R.D., and Muller, W.J. (1992). Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. Mol. Cell. Biol. *12*, 954–961. https://doi.org/10.1128/mcb.12.3.954-961.1992.

67. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). <tt>edgeR</tt> : a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139–140. https://doi.org/10.1093/bioinformatics/btp616.

68. McCarthy, D.J., Chen, Y., and Smyth, G.K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res. *40*, 4288–4297. https://doi.org/10.1093/nar/gks042.

69. Chen, Y., Lun, A.T.L., and Smyth, G.K. (2016). From reads to genes to pathways: differential expression analysis of RNA-Seq experiments using Rsubread and the edgeR quasi-likelihood pipeline. F1000Res *5*, 1438. https://doi.org/10.12688/f1000research.8987.2.

70. Liao, Y., Smyth, G.K., and Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res. *47*, e47. https://doi.org/10.1093/nar/gkz114.

71. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. Cell *184*, 3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

72. Trapnell, C., Cacchiarelli, D., Grimsby, J., Pokharel, P., Li, S., Morse, M., Lennon, N.J., Livak, K.J., Mikkelsen, T.S., and Rinn, J.L. (2014). The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat. Biotechnol. *32*, 381–386. https://doi.org/10.1038/nbt.2859.

73. Qiu, X., Hill, A., Packer, J., Lin, D., Ma, Y.-A., and Trapnell, C. (2017). Single-cell mRNA quantification and differential analysis with Census. Nat. Methods *14*, 309–315. https://doi.org/10.1038/nmeth.4150.

74. Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H.A., and Trapnell, C. (2017). Reversed graph embedding resolves complex single-cell trajectories. Nat. Methods *14*, 979–982. https://doi.org/10.1038/nmeth.4402.

75. Law, C.W., Chen, Y., Shi, W., and Smyth, G.K. (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol. *15*, R29. https://doi.org/10.1186/gb-2014-15-2-r29.

76. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47. https://doi.org/10.1093/nar/gkv007.

77. 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. https://doi.org/10.1093/nar/gkt214.

78. 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. https://doi.org/10.1016/j.molcel.2010.05.004.

79. Liao, Y., Smyth, G.K., and Shi, W. (2014). FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.

80. Ma, J., Lanza, D.G., Guest, I., Uk-Lim, C., Glinskii, A., Glinsky, G., and Sell, S. (2012). Characterization of mammary cancer stem cells in the MMTV-PyMT mouse model. Tumour Biol. *33*, 1983–1996. https://doi.org/10.1007/s13277-012-0458-4.

81. Plante, I., Stewart, M.K.G., and Laird, D.W. (2011). Evaluation of mammary gland development and function in mouse models. J. Vis. Exp. *53*, 2828. https://doi.org/10.3791/2828.

82. Alshehry, Z.H., Barlow, C.K., Weir, J.M., Zhou, Y., McConville, M.J., and Meikle, P.J. (2015). An efficient single phase method for the extraction of plasma lipids. Metabolites *5*, 389–403. https://doi.org/10.3390/metabo5020389.

83. Huynh, K., Barlow, C.K., Jayawardana, K.S., Weir, J.M., Mellett, N.A., Cinel, M., Magliano, D.J., Shaw, J.E., Drew, B.G., and Meikle, P.J. (2019). High-Throughput Plasma Lipidomics: Detailed Mapping of the Associations with Cardiometabolic Risk Factors. Cell Chem. Biol. *26*, 71–84.e4. https://doi.org/10.1016/j.chembiol.2018.10.008.

84. Smyth, G.K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat. Appl. Genet. Mol. Biol. *3*, Article3. https://doi.org/10.2202/1544-6115.1027.

85. Buenrostro, J.D., Wu, B., Litzenburger, U.M., Ruff, D., Gonzales, M.L., Snyder, M.P., Chang, H.Y., and Greenleaf, W.J. (2015). Single-cell chromatin accessibility reveals principles of regulatory variation. Nature *523*, 486–490. https://doi.org/10.1038/nature14590.

86. Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: A method for assaying chromatin accessibility genome-wide. Curr. Protoc. Mol. Biol. *109*, 21.29.1–21.29.9. https://doi.org/10.1002/0471142727.mb2129s109.

87. Jamieson, P.R., Dekkers, J.F., Rios, A.C., Fu, N.Y., Lindeman, G.J., and Visvader, J.E. (2017). Derivation of a robust mouse mammary organoid system for studying tissue dynamics. Development *144*, 1065–1071. https://doi.org/10.1242/dev.145045.

88. Sachs, N., de Ligt, J., Kopper, O., Gogola, E., Bounova, G., Weeber, F., Balgobind, A.V., Wind, K., Gracanin, A., Begthel, H., et al. (2018). A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity. Cell *172*, 373–386.e10. https://doi.org/10.1016/j.cell.2017.11.010.

89. Sumbal, J., Chiche, A., Charifou, E., Koledova, Z., and Li, H. (2020). Primary mammary organoid model of lactation and involution. Front. Cell Dev. Biol. *8*, 68. https://doi.org/10.3389/fcell.2020.00068.

90. Morrow, R.J., Ernst, M., and Poh, A.R. (2023). Longitudinal quantification of mouse gastric tumor organoid viability and growth using luminescence and microscopy. Star Protoc. *4*, 102110. https://doi.org/10.1016/j.xpro.2023.102110.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit anti-mouse Milk Specific Protein	Accurate Chemical and Scientific Corporation	Cat# YNRMMSP
Rabbit anti-mouse Ki67 Monoclonal Antibody	Thermo Fisher Scientific	Cat# MA5-14520; RRID:AB_10979488
Dako envision anti-rabbit labelled polymer-HRP	Dako	Cat# K400311-2
Rat anti-mouse Krt8/18 TROMA-1	DSHB	Cat# AB-531826
ImmPRESS™ HRP Goat anti-Rat IgG Mouse adsorbed Polymer Detection Kit, Peroxidase	Vector Laboratories	Cat# MP-7444-15
Rabbit anti-mouse Keratin 14 antibody	Biolegend	Cat# 905304; RRID:AB_2616896
Jackson Immuno Research Labs Rat γ-globulin, unconjugated	Fisher Scientific	Cat# NC9948213
Purified rat anti-mouse CD16/CD32 (Mouse BD Fc Block™) (Clone 2.4G2)	BD Pharmingen	Cat# 553141; RRID:AB_394656
APC anti-mouse CD31 (rat, clone 390)	BioLegend	Cat# 102410; RRID:AB_312904
APC anti-mouse CD45 Antibody (rat, clone 30-F11)	BioLegend	Cat# 103112; RRID:AB_312977
APC anti-mouse TER-119/Erythroid Cells (rat, clone Ter119,)	BioLegend	Cat# 116212; RRID:AB_313713
Pacific Blue anti-mouse CD24 (rat, clone M1/69)	BioLegend	Cat# 101820; RRID:AB_572010
FITC-anti-mouse CD29 (mouse/rat clone HMβ1-1)	BioLegend	Cat# 102206; RRID:AB_312883
APC Rat IgG2b, κ Isotype Ctrl Antibody (clone RTK4530)	BioLegend	Cat# 400612; RRID:AB_326556
FITC Armenian Hamster IgG Isotype Ctrl Antibody (clone HTK888)	BioLegend	Cat# 400906; RRID:AB_2923258
Pacific Blue™ Rat IgG2b, κ Isotype Ctrl Antibody (clone RTK4530)	BioLegend	Cat# 400627; RRID:AB_493561
Biological samples		
Mammary glands dissected from female mice at various ages and stages of pregnancy	This paper	N/A
Chemicals, peptides, and recombinant proteins		
Quick Extract DNA Extraction	Astral Scientific	Cat# QE09050
MyTaq™ Red Mix	Bioline	Cat# BIO-25044
Agarose, Molecular Grade	Bioline	Cat# BIO-41025
1kb Plus DNA Ladder	Invitrogen	Cat# 10787018
Power SYBR Green PCR Master Mix	Thermo Fisher Scientific	Cat# 4368708
Carmine	Sigma Aldrich	Cat# C1022-5G
Aluminum potassium sulfate dodecahydrate	Sigma Aldrich	Cat# 237086-100G
Thymol	Sigma Aldrich	Cat# T0501-100G
Chloroform	Sigma Aldrich	Cat# C2432-500ML
Glacial acetic acid	Sigma Aldrich	Cat# A6283
Entellan™ mounting medium	Sigma Aldrich	Cat# 1079600500
Neutral Buffered Formalin 10% Clear 10L	Australian Biostain	Cat# 029-ANBFC.10L
Scott's tap water	Leica Biosystems	Cat# 3802900
Haematoxylin Mayer's (Modified formula)	Amber Scientific	Cat# MHM-1L
Eosin Y Solution, Aqueous	Sigma Aldrich	Cat# HT110232-1L
Xylene	Point of Care diagnostics	Cat# XYL2.5P
DPX Mountant for histology	Sigma Aldrich	Cat# 06522-500ML
Hydrogen peroxide 30%	Science Supply Australia	Cat# HA154/500ML
Sodium citrate tribasic dihydrate	Sigma Aldrich	Cat# S4641
UltraPure™ Ethylenediaminetetraacetic Acid, Disodium Salt, Dihydrate ($Na_2EDTA.2H_2O$)	Thermo Fisher Scientific	Cat# 15576028
Liquid 3, 3-diaminobenzidine (DAB)+ 2-component system	Agilent technologies Australia	Cat# K346811-2

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Normal goat serum	Thermo Fisher Scientific	Cat# 31872
Triton™ X-100	Sigma Aldrich	Cat# X100-100ML
Spectral- 40,6-Diamidino-2-phenylindole dihydrochloride (DAPI)	Akoya Biosciences	Cat# NC1526331
ProLong™ Diamond Antifade Mountant	Thermo Fisher Scientific	Cat# P36970
DMEM-F12	Thermo Fisher Scientific	Cat# 11320033
L-glutamine	Thermo Fisher Scientific	Cat# A2916801
Fetal Bovine Serum, heat inactivated	Thermo Fisher Scientific	Cat# 10100147
Insulin neutral (rys) 100IU/ml	Symbion	Cat# 272957
Hydrocortisone sodium succinate 100mg	Symbion	Cat# 328472
Recombinant mouse epidermal growth factor (EGF)	R&D Systems	Cat# RDS2028-EG-200
Gibco™ Collagenase, Type II, powder	Thermo Fisher Scientific	Cat# 17101015
Hyaluronate 4-glycanohydrolase (Hyaluronidase)	Sigma Aldrich	Cat# H3506
Deoxyribonuclease I (DNase I)	Worthington Biochemical corporation	Cat# LS002140
Trypsin (2.5%)	Thermo Fisher Scientific	Cat# 15090-046
Ethylene glycol-bis(2-aminoethylether)-N,N,N0, N0 -tetraacetic acid (EGTA)	Sigma Aldrich	Cat# E0396
Dispase II	Sigma Aldrich	Cat# 4942078001
Trypan Blue 0.4% (2,7-Naphthalenedisulfonic acid, 3,3-[(3,3-dimethyl[1,1-biphenyl]-4,4-di yl)bis(azo)] bis[5-amino-4-hydroxy-, tetrasodium salt)	Thermo Fisher Scientific	Cat# T10282
7-Aminoactinomycin D (7-AAD)	Sigma Aldrich	Cat# A9400
Chloroform (LC grade)	Merck	Cat# 1.02445.2500
Methanol (LC grade)	Merck	Cat# 1.06018.2500
Butanol	Merck	Cat# 1.01988.2500
Ammonium Formate	Sigma Aldrich	Cat# 516961-100G
Acetonitrile (LCMS grade)	Merck	Cat# 1.00029.2500
Isopropanol (LCMS grade)	Merck	Cat# 1.02781.4000
Water (LCMS grade)	Merck	Cat# 1.15333.2500
IGEPAL® CA-630	Sigma Aldrich	Cat# I8896
Digitonin, High Purity	Sigma Aldrich	Cat# 300410
TWEEN® 20	Sigma Aldrich	Cat# P1379
AMPure XP Reagent	Beckman Coulter	Cat# A63880
Cultrex Reduced Growth Factor Basement Membrane Extract, Type 2, Pathclear	R&D Systems	Cat# RDS3533-010-02
Advanced DMEM-F12	Thermo Fisher Scientific	Cat# 12634010
Cholera toxin	Sigma Aldrich	Cat# C8052
Penicillin-Streptomycin	Thermo Fisher Scientific	Cat# 15140122
HEPES	Thermo Fisher Scientific	Cat# 15630080
B-27 Supplement minus vitamin A	Thermo Fisher Scientific	Cat# 12587-010
N-Acetylcysteine	Sigma Aldrich	Cat# A7250
Recombinant mouse FGF2 (basic)	R&D Systems	Cat# 3139-FB-025
Recombinant mouse FGF-10	R&D Systems	Cat# 6224-FG-025
Recombinant mouse Wnt3a	R&D Systems	Cat# 1324-WN-002
Heparin sodium salt	R&D Systems	Cat# 2812/100
Recombinant murine R-Spondin-1	Peprotech	Cat# 315-32-100
Y-27632 2HCl	Selleck Chemicals	Cat# S1049
Bovine serum albumin (BSA)	Sigma Aldrich	Cat# A7030
Recombinant Mouse Prolactin Protein	R&D Systems	Cat# 1554-PL-050
Corning® Cell Recovery Solution	In vitro technologies	Cat# FAL354253

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical commercial assays		
ReliaPrep RNA Tissue Miniprep System - Fibrous Tissue protocol	Promega	Cat# Z6112
High-Capacity cDNA Reverse Transcription Kit	Applied Biosystems	Cat# 4368814
Chromium Next GEM Single Cell 3' Kit v3.1	10X Genomics	Cat# PN-1000268
Illumina Tagment DNA TDE1 Enzyme and Buffer Kit	Illumina	Cat# 20034197
MinElute Reaction Cleanup Kit	Qiagen	Cat# 28204
Opal 4-Color IHC manual kit	Akoya Biosciences	Cat# NEL810001KT
RealTime-Glo™ MT Cell Viability Assay	Promega	Cat# G9711
RNeasy® Micro Kit	Qiagen	Cat# 74004
Deposited data		
scRNA-seq of human breast tissue	Bach et al.[32]	GEO: GSE106273
scRNA-seq	This Manuscript	GEO: GSE221684
ATAC-seq	This Manuscript	GEO: GSE260878
Experimental models: Cell lines		
Ehf^{WT} and *Ehf*^{KO} mammary epithelial organoids	This paper	N/A
Ehf^{WT}*MMTV-PyMT* and *Ehf*^{KO}*MMTV-PyMT* mammary tumour organoids	This paper	N/A
Experimental models: Organisms/strains		
Mouse: C57bl/6 *Ehf*$^{fl/fl}$	Generated in house by Ozgene	N/A
Mouse: BALB/c-Tg(CMV-cre)1Cgn/J	Mouse Genetics Cologne (MGC) Foundation	Jackson Stock Number: 003465
Mouse: FVB/N (Tg(MMTV^{Cre})1Mam/J (MMTV-Cre/Line A)	Wagner et al.[45]	Jackson Stock Number: 003551
Mouse: *FVB/N-Tg(MMTV-PyVT)634Mul/J* (*MMTVPyMT*)	Guy et al.[66]	Jackson Stock Number: 002374
Oligonucleotides		
Primers for genotyping	See Table S3 for full sequence	N/A
Primers for real-time quantitative PCR	See Table S4 for full sequences	N/A
Software and algorithms		
edgeR	Robinson et al.,[67] McCarthy et al.,[68] Chen et al.,[69]	https://bioconductor.org/packages/release/bioc/html/edgeR.html
SPOT Basic Software	Spot Imaging	https://www.spotimaging.com/software/spot-basic/
Aperio ImageScope software v12.0.1.5027	Leica Biosystems	https://www.leicabiosystems.com/en-au/digital-pathology/manage/aperio-imagescope/
inForm® tissue analysis software version 2.2	Akoya Biosciences	https://www.akoyabio.com/phenoimager/inform-tissue-finder/
HALO software version 3.4	Indica Labs, USA	https://learn.indicalab.com/
FlowJo software v10.1r7	Tree Star	RRID:SCR_008520; URL: https://www.flowjo.com/solutions/flowjo
Rsubread v2.6.4	Liao et al.,[70]	https://bioconductor.org/packages/release/bioc/html/Rsubread.html
R	R Project for Statistical Computing	https://www.r-project.org/
Seurat v4	Hao*, Hao* et al.,[71]	https://satijalab.org/seurat/articles/install.html

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Monocle 2	Trapnell et al.,[72] Qiu et al.,[73] Qiu et al.,[74]	https://cole-trapnell-lab.github.io/monocle-release/docs/
Voom	Law et al.,[75]	N/A
Limma	Ritchie et al.,[76]	https://bioconductor.org/packages/release/bioc/html/limma.html
Subread	Liao et al.,[77]	https://subread.sourceforge.net/
HOMER v4.11	Heinz et al.,[78]	http://homer.ucsd.edu/homer/
FeatureCounts	Liao et al.,[79]	https://subread.sourceforge.net/featureCounts.html
Kaleido 3.0 software	Perkin Elmer	https://www.flinders.edu.au/content/dam/documents/research/health-medical-research-institute/ensight-user-manual.pdf
cellSens	Olympus	https://www.olympus-lifescience.com/en/software/cellsens
KMPlotter	Curtis et al.[52]	https://kmplot.com/analysis/
METABRIC dataset	Győrffy[53]	https://www.cbioportal.org/
TCGA dataset	Koboldt et al.,[54]	https://www.cancer.gov/tcga
GraphPad Prism v8.0	GraphPad Software	https://www.graphpad.com/
BioRender	BioRender	https://www.BioRender.com
Other		
MasterCycler EP Thermal Cycler	Eppendorf	https://www.gmi-inc.com/product/eppendorf-mastercycler-ep-thermal-cycler-range/
Omni Metal Beads- 2.4 mm diameter	Capella Science	Cat# 19-640
Nanodrop ND-2000	NanoDrop Technologies	Cat# ND-2000
ViiA™ 7 Real Time PCR system	Applied Biosystems	Cat# 4453534
Leica Biosystems RM2245	Leica Biosystems	https://www.leicabiosystems.com/en-au/histology-equipment/microtomes/leica-rm2245/
Epredia™ SuperFrost™ Plus Adhesion Control Microscope Slide	Fisher Scientific	Cat# 22-042-941
Zeiss Axioskop 2	Zeiss	N/A
Aperio AT2 Slide Scanner	Leica Biosystems	N/A
VECTRA 3 microscope	Akoya Biosciences	N/A
McIlwain tissue chopper, Model TC752	Mickle Laboratory Engineering	N/A
BD FACSAria Fusion	BD Biosciences	Flow Cytometry, WEHI URL: https://www.wehi.edu.au/research/research-fields/flow-cytometry
6495C QQQ mass spectrometer	Agilent	N/A
2200 Tapestation	Agilent	N/A
Illumina NextSeq 500	Illumina	N/A
24-well suspension plate	Greiner Bio-One	Cat# 662102
EnSight multimode plate reader	Perkin Elmer	N/A
Olympus CKX53 microscope	Olympus	https://www.olympus-lifescience.com/en/microscopes/inverted/ckx53

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, John Mariadason (john.mariadason@onjcri.org.au).

Materials availability

Mouse lines and organoids generated in this study are available upon request to the lead contact.

Data availability

- Sequencing data have been deposited in the GEO database under accession codes GEO: GSE221684 (scRNA-seq) and GEO: GSE260878 (ATAC-seq).
- This paper analyses existing publicly available datasets. Accession numbers are listed in the key resources table.
- This paper does not report any original code.
- Any additional information required to reanalyse the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Generation of conditional Ehf knockout mice

Mice carrying the *Ehf*$^{fl/fl}$ allele, whereby exon 8 and the coding region of exon 9 were flanked by *loxP* sites, were generated by oocyte injection techniques (Ozgene) in *C57bl/6* embryonic stem (ES) cells as previously described.[39] *Ehf*$^{fl/fl}$ mice were mated with *C57bl/6 Tg;CMV*^{Cre} mice (Original Source: Mouse Genetics Cologne (MGC) Foundation, provided by Warren Alexander, Walter and Eliza Hall Institute, Melbourne, Australia) to induce biallelic *Ehf* gene deletion in all tissues. The *CMV*^{Cre} allele was then bred out of the colony, and germline *Ehf*^{KO} and *Ehf*^{WT} control progeny were obtained by *Ehf* heterozygous mating and studied on a pure *C57bl/6* genetic background. *Ehf*$^{fl/fl}$ mice were also mated with mouse mammary tumor virus (MMTV) long terminal repeat (*Tg(MMTV*^{Cre}*)1Mam/J* (*MMTV-Cre/Line A*)) mice (Original source: Kay-Uwe Wagner,[45] Wayne State University, JAX stock #003551) to generate mammary epithelial specific *Ehf* deletion. *Ehf*$^{fl/fl}$ *MMTV*$^{Cre/+}$ (*Ehf*^{MKO}); *Ehf*$^{fl/fl}$ and *Ehf*$^{+/+}$*MMTV*^{Cre} (*MMTV*^{Cre}) genotypes for this study were obtained by mating *Ehf*$^{fl/+}$ females to *Ehf*$^{fl/+}$*MMTV*$^{Cre/+}$ males, and littermates were studied on a mixed *C57bl/6 FVB/N* genetic background.

Generation of Ehf knockout MMTV-PyMT mice

FVB/N-Tg(MMTV-PyVT)634Mul/J (*MMTV-PyMT*) males (Original source: William Muller[66] McGill University, JAX Stock#002374) expressing mouse mammary tumor virus (MMTV) long terminal repeat upstream of a cDNA sequence encoding the Polyoma Virus middle T antigen (PyVT) were used in mating due to failure of *MMTV-PyMT* females to lactate.[80] *Ehf*$^{+/-}$ females were used in mating due to failure of *Ehf* $^{-/-}$ females to lactate. Female mice heterozygous for germline *Ehf* deletion (*Ehf*$^{+/-}$) were then crossed to *MMTV-PyMT*$^{Tg/+}$ males; and their *Ehf* heterozygous deletion progeny were then interbred to generate *Ehf*^{KO}*MMTVPyMT* (*Ehf* $^{-/-}$*MMTV-PyMT*$^{Tg/+}$) and *Ehf*^{WT}*MMTVPyMT* (*Ehf* $^{+/+}$*MMTV-PyMT*$^{Tg/+}$) mice for this study. Female litter mates were studied on a mixed *C57bl/6 FVB/N* genetic background.

Mouse studies

All animal breeding and procedures were performed in accordance with specific approvals obtained from the La Trobe University (AEC16-58) and Austin Health (A2015-05274; A2018-05553; A2019-05648) Animal Ethics Committees (Melbourne) and conformed to the Australian Code of Practice for the Care and Use of Animals for Scientific Purposes. All animals in this study were housed in open-top boxes with basic enrichment and provided with standard mouse chow and drinking water *ad libitum*. For timed pregnancies, the Jackson laboratory guidelines were followed, whereby the soiled bedding from stud males was added to the females' boxes three nights prior to mating to induce synchronized estrous cycles. Adult female mice were mated overnight and assessed for the presence of vaginal plugs the following morning. The presence of a plug was counted as gestation day 0.5 (0.5 dP). The gestation stage of females was confirmed at the time of mammary gland collection by examination of the embryos developmental stage.

For the *MMTV-PyMT* tumour model, *Ehf*^{WT}*MMTV-PyMT* and *Ehf*^{KO}*MMTV-PyMT* female mice were examined and palpated twice weekly from 4 weeks of age onwards to assess for the onset of mammary tumours. Once palpable tumour/s were detected, their size was monitored by digital caliper measurements three times weekly until endpoint. Mice were euthanized by CO2 inhalation when tumours reached the ethically determined combined size limit of 1.5 cm^3 (Ethical endpoint cohort), or 6 weeks of age (6-week-old cohort). At ethical endpoint, macroscopic mammary tumours were excised and weighed. For animals of the 6-week-old cohort, inguinal mammary glands were whole mounted and carmine alum stained.

METHOD DETAILS

DNA genotyping

DNA extraction for genotyping was performed by incubating 2-3 mm of mouse tail biopsy samples in 30 μL Quick Extract DNA Extraction solution (Astral Scientific, Australia) at 65°C for 10 minutes followed by 98°C for 5 minutes. To amplify selected regions of DNA, 18 μL of a master mix consisting of gene-specific forward and reverse primers (6.25μL of each 100μM primer), 500 μL 2xMyTaq Red (Bioline, Australia) made up to 1 mL with nuclease free water was added to 2 μL of DNA. DNA was then amplified using a MasterCycler Thermal Cycler (Eppendorf, USA). Oligonucleotides used for genotyping are listed in Table S3. For DNA genotyping of *Ehf*, using the combination of primer A and C, the wild-type allele produced a 769 bp product, the floxed allele produced a 912 bp

product and the recombined allele produced a 252 bp product. To further confirm deletion of *Ehf*, the primer set B and C were utilized, whereby the wild-type allele produced a 501 bp product, the floxed allele produced a 598 bp product, and the recombined allele produced no PCR product due to primer B binding within the floxed region of *Ehf*. Genotyping for *MMTV*^{Cre} produced a 280 bp product in animals with the *MMTV*^{Cre} allele, with no product produced in wild-type mice. Genotyping for *MMTV-PyMT* produced a 600 bp product in animals with the *PyMT* transgene, with no product produced in wild-type mice.

Specific thermocycler conditions were as follows:

Genotyping for *Ehf*:

94°C	2:00 minutes	
94°C	0:30 minutes	x35 cycles
68°C	1:00 minutes	
72°C	1:20 minutes	
72°C	3:00 minutes	
4°C	∞	

Genotyping for *MMTV*^{Cre}:

94°C	2:00 minutes	
94°C	0:40 minutes	x35 cycles
60°C	1:00 minutes	
72°C	2:00 minutes	
72°C	5:00 minutes	
4°C	∞	

Genotyping for *MMTV-PyMT*:

94°C	2:00 minutes	
94°C	1:00 minutes	x35 cycles
55°C	1:00 minutes	
72°C	1:00 minutes	
72°C	7:00 minutes	
4°C	∞	

Gel electrophoresis

The products of DNA genotyping reactions were resolved and imaged by agarose gel electrophoresis to determine genotypes. Agarose (1.5% w/v, Bioline, Australia) in TAE buffer was heated in a microwave until homogenous. The agarose solution was subsequently poured into a gel tray to solidify. Once the gel had set it was suspended in TAE buffer in a gel tank and amplified DNA (10μL) applied to wells along the top of the gel. Where needed, a 1kb plus DNA ladder (2μL, Invitrogen, Australia) was included to accurately determine size of bands. DNA from mice with known genotypes were used as controls. The loaded gel was run at 90V for 60-75 minutes.

Expression of Ehf in MECs from publicly available datasets

Ehf expression in 15 annotated clusters of mouse mammary epithelial cells was derived using the online tool from the Marioni lab http://marionilab.cruk.cam.ac.uk/mammaryGland. Abbreviations: Basal in gestation (Bsl-G), Myoepithelial in lactation (Myo), Alveolar differentiated in gestation (Avd-G), Alveolar differentiated in lactation (Avd-L), Hormone sensing differentiated in gestation (Hsd-G), Alveolar progenitor in gestation (Avp-G), Procr^{+} basal cells (Prc), Alveolar progenitor in lactation (Avp-L), Basal in nulliparous gland (Bsl), Luminal progenitor in post involution (Lp-PI), Luminal progenitor in nulliparous gland (Lp-NP). Hormone sensing differentiated in post involution (Hsd-PI), and unclassified. *EHF* expression in human mammary cells was derived from scRNA-seq profiles of normal human breast tissues from 13 women with no family history of breast cancer, which were integrated and visualized in UMAP. Cell clusters of the integrated data were identified and annotated as described in Pal et al.[40] Log-count-per-million (log-CPM) values of *EHF* were computed using the *cpm* function of the *edgeR* package.[67–69]

Mammary gland RNA isolation, cDNA synthesis and quantitative real time polymerase chain reaction

RNA was extracted from whole thoracic mammary glands and primary mammary tumours using the ReliaPrep RNA Tissue Miniprep System Fibrous Tissue protocol (Promega, USA) using a metal bead homogenization step (Omni Metal Beads- 2.4 mm diameter, Capella Science, Australia), with the DNase I treatment step included. RNA concentration and quality was determined spectrophotometrically using the Nanodrop ND-2000 (NanoDrop Technologies, USA). RNA purity was determined by the ratio of absorbance at

260 nm and 280 nm, with ratios of >1.8 deemed of sufficient purity for cDNA synthesis. 200-500ng of RNA was reverse transcribed to cDNA using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, USA) as per manufacturer's instructions using random primers. Relative gene expression was determined by quantitative real time PCR using the ViiA 7 Real Time PCR system (Applied Biosystems, USA). The reaction mixture consisted of 0.75 μL 500nM forward primer, 0.75μL 500nM reverse primer, 1 μL of 1:10 diluted cDNA, 2.5μL 2x Power SYBR Green PCR Master Mix (Thermo Fisher Scientific, Australia). Samples were amplified using the following program: 95°C for 10 minutes followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. Expression levels of genes of interest were normalized to the house keeping gene β-actin (ACTB). Gene specific primers used for quantitative real time PCR are listed in Table S4.

Mammary gland whole mounts and histology

For mammary gland whole mounts, glands were excised and mounted onto SuperFrost slides (Thermo Fisher Scientific, Australia) then fixed in Carnoy's fixative (100% Ethanol, Chloroform, Glacial acetic acid; 6:3:1) overnight at 4°C. Glands were then gradually rehydrated in serial ethanol baths of 70%-, 35%- and 15%- ethanol and water for 5 minutes each, then immersed in distilled water for 5 minutes at room temperature. Glands were then immersed in Carmine alum (1g Carmine, 2.5g Aluminum potassium sulfate, Thymol) until the stain penetrated through the mammary epithelial tissue and then gradually dehydrated in serial water baths of 50%-, 35%-, 15%- ethanol and water for 5 minutes each then distilled water for 5 minutes at room temperature then cleared in xylene until only the epithelium remained.[81] Glands were mounted with Entellan™ (Sigma Aldrich). Formalin-fixed paraffin embedded (FFPE) thoracic and inguinal mammary gland sections and *MMTV-PyMT* primary mammary tumours were prepared by placing glands in cassettes in 10% neutral buffered formalin (Australian Biostain) for 48 hours before changing to 80% ethanol. Samples were processed and paraffin embedded (Department of Anatomical Pathology, Austin Hospital, Australia). 4 μm sections were cut using a Leica Biosystems RM2245 Semi-Automated Rotary Microtome (Leica Biosystems, Australia) onto SuperFrost slides (Fisher Scientific, Australia). Slides were de-paraffinised and rehydrated through serial xylene and ethanol washes, rinsed in RO water and stained in filtered Mayer's haematoxylin (Amber Scientific) for 60 seconds. After washing in warm tap water, slides were placed in Scott's tap water for 60 seconds then rinsed in water, counterstained in acidified aqueous eosin Y solution (Sigma Aldrich, USA) for 2-5 minutes before being dehydrated through serial ethanol and xylene washes and mounted with DPX (Sigma Aldrich, USA). Images from 18.5 dP carmine alum-stained whole mounts were captured using a Zeiss Axioskop 2 microscope using the ACHROPLAN 4x objective (NA 0.1) (Zeiss AxioCam HRc color camera and SPOT 5.6 software); Whole mounts from other stages and images from H&E- stained sections were captured using the Aperio AT2 Slide Scanner (Leica Biosystems, USA) using the Olympus 20x UPLanSApo (NA 0.75) objective.

Immunostaining

FFPE tissue sections (4 μm) were cut and placed onto electrostatic glass slides and processed by deparaffinizing and rehydrating through serial xylene and ethanol washes, then rinsed in tap water. Endogenous peroxidase was quenched by H_2O_2 (3% in RO water, Science Supply Associates, Australia) at room temperature for 10 minutes. Antigen retrieval was performed by incubation in 10mM sodium citrate pH 6.0 (sodium citrate tribasic dihydrate (Sigma-Aldrich, USA)), or 1mM EDTA pH 8.0 (Ultrapure EDTA, Invitrogen, Australia) at 100°C for 30 minutes in RO water. Slides were then cooled to below 65°C before washing in Tris-buffered saline with 0.05% Tween20 (TBST), then stained with the following primary antibodies: Rabbit anti-Mouse Milk Specific Protein (Accurate Chemical and Scientific Corporation, USA at 1:12,000 dilution), and Rabbit anti-mouse Ki67 Monoclonal Antibody (Thermo Fisher Scientific, UK at 1:150 dilution) overnight at 4°C in a humidified chamber. Slides were then washed in TBST and incubated with Dako envision anti-rabbit labelled polymer-HRP (Dako, USA) secondary antibody for 30 minutes at room temperature, followed by washing in TBST and visualization by addition of 3, 3-diaminobenzidine (DAB) (Dako, USA). Slides were counterstained with haematoxylin and dehydrated through serial ethanol and xylene washes then sealed with glass coverslips using DPX mounting solution (Sigma Aldrich, USA). Stained slides were scanned on an Aperio AT2 slide scanner (Leica Biosystems, USA) and images analysed using Aperio ImageScope software v12.0.1.5027 using a Nuclear v9.1 or Positive Pixel Count v9.1 algorithm. The algorithm was tested on three separate samples per stain before being applied to all samples.

Multicolour immunofluorescence staining was performed using the Opal 4-Color IHC manual kit (Akoya Biosciences). FFPE sections were deparaffinised and rehydrated through serial xylene and ethanol washes, then rinsed in tap water. Endogenous peroxidase was quenched by H_2O_2 (3% in RO water, Science Supply Associates, Australia) at room temperature for 10 minutes. Antigen retrieval was performed by incubation in AR6 buffer (Akoya Biosciences) in a microwave (LG). Sections were brought to boiling point at full power then simmered using 20% power for 15 minutes. Slides were cooled to room temperature for 20 minutes before washing in RO water followed by TBST, then blocked in 3% normal goat serum (Thermo Fisher Scientific), 0.3% Triton-X-100 (Sigma Aldrich) in TBS for 30 minutes, followed by Opal blocking buffer/diluent (Akoya Biosciences) for 10 minutes at room temperature. Slides were then stained with primary antibody rat anti-mouse Krt8/18 TROMA-1 (antibody deposited to the DSHB by Brulet, P. / Kemler, R. (DSHB Hybridoma Product TROMA-1)), at 1:100 dilution for 1 hour at room temperature. Slides were then washed in TBST and incubated with ImmPRESS™ HRP Goat anti-Rat IgG (Mouse Adsorbed-Detection kit, Vector Laboratories) secondary antibody for 10 minutes at room temperature, followed by washing in TBST. Slides were then incubated in Opal fluorophore solution Opal-690 at 1:100 dilution for 10 minutes at room temperature, followed by washing in TBST. A second antigen retrieval and blocking procedure was then performed as described above before staining with primary antibody rabbit anti-mouse Krt14 overnight at 4°C (Biolegend, at 1:800 dilution). Slides were then washed in TBST and incubated with Opal HRP Anti-Ms+Rb (Akoya Biosciences) for

10 minutes at RT, slides were then washed in TBST and incubated with Opal-570 at 1:150 dilution for 10 minutes at room temperature, followed by washing in TBST. A final antigen retrieval step was performed as described above, followed by washing in RO water and TBST. Slides were then stained with spectral DAPI (Akoya Biosciences) as per manufacturer's recommendation then washed in TBST and mounted with ProLong™ Diamond Antifade Mountant (Thermo Fisher Scientific). Stained slides were scanned on a VECTRA 3 microscope with a multispectral camera using the Olympus UPlanSApo 20x/N0.75 objective and analysis performed using Vectra 3.0 (Akoya Biosciences), inForm tissue analysis software version 2.2 (Akoya Biosciences) and HALO software version 3.4 (Indica Labs, USA).

Mammary epithelial cell preparation and flow cytometry

Thoracic and inguinal mammary glands were resected, lymph nodes removed, and samples collected into cold MEC media (DMEF12 supplemented with 1 mM glutamine (Thermo Fisher Scientific), 5μg/ml insulin (Symbion), 500ng/mL hydrocortisone (Symbion), 10ng/mL Epidermal growth factor (R&D systems) and 5% Heat inactivated Fetal bovine serum (Thermo Fisher Scientific). For the scRNA-seq study, mammary glands of 18.5 dP females of the same genotype were pooled. Glands were mechanically dissociated with a McIlwain tissue chopper (Mickle Laboratory Engineering), then digested for 1 hour at 37°C in an orbital shaker with 300 U/mL collagenase (Sigma), 100 U/mL hyaluronidase (Sigma) and 0.1 mg/mL DNase I (Worthington) in MEC media containing 5% Fetal bovine serum. The resulting suspension was diluted in PBS with 2% fetal bovine serum and cells pelleted by centrifugation at 1050 rpm for 5 minutes at room temperature. The cell pellet was further digested in 0.25% trypsin-EGTA with 0.1 mg/mL DNase I for 3 minutes at 37°C, then diluted in PBS with 10% fetal bovine serum and cells pelleted by centrifugation at 1050 rpm for 5 minutes at room temperature. The cell pellet was then digested in 5 mg/mL Dispase (Sigma Aldrich) with 0.1mg/mL DNase I (Roche Diagnostics) at 37°C for 5 minutes, then diluted in PBS with 2% fetal bovine serum. Cells were passed through a 100 μm cell strainer (Miltenyi Biotec 130-110-917) to obtain a single cell suspension, then live cells counted using trypan blue. Single cells were pelleted by centrifugation at 1050 rpm for 5 minutes at room temperature, then blocked in rat γ-globulin (Jackson Laboratories) and anti-CD16/CD32 Fcγ III/II receptor antibody (BD Pharmingen) with 0.1 mg/mL DNase I for 10 minutes at room temperature. Antibody incubations were performed at 4°C for 25 minutes. The following antibodies were used for Fluorescence-activated cell sorting (FACS): APC anti-mouse CD31 (rat, clone 390, BioLegend), APC anti-mouse CD45 Antibody (rat, clone 30-F11, BioLegend), APC anti-mouse TER-119/Erythroid Cells (rat, clone Ter119, BioLegend), Pacific Blue anti-mouse CD24 (rat, clone M1/69, BioLegend) and FITC-anti-mouse CD29 (mouse/rat clone HMβ1-1, BioLegend). Isotype controls APC Rat IgG2b, κ Isotype Ctrl Antibody (clone RTK4530), FITC Armenian Hamster IgG Isotype Ctrl Antibody (clone HTK888), Pacific Blue™ Rat IgG2b, κ Isotype Ctrl Antibody (clone RTK4530) were used to set up the sorting gates. Cells were diluted in PBS with 2% fetal bovine serum then pelleted by centrifugation at 1050 rpm for 5 minutes, and 0.2mg/ml 7-Aminoactinomycin D (Sigma Aldrich) added to exclude dead cells. The cell suspension was then filtered through a 40 μm cell strainer and the luminal cell population ($Ter119^{-}CD31^{-}CD45^{-}$ $CD24^{+}CD29^{lo}$) sorted on a FACS Aria Fusion using a 100 μm nozzle (Becton Dickinson). Data analysis was performed on the single, live gate using FlowJo software (v 10.1r7, Tree Star).

Lipidomic profiling

Lipidomic profiling was performed by the Baker Metabolomics Facility. Lipids were extracted from mammary gland homogenates (50 μg of protein) using a single-phase chloroform:methanol extraction method as described previously.[82,83] Briefly, 10μl of tissue homogenates in PBS were mixed with 10 μl of internal stand mix as previously described[83] in 1:1 chloroform:methanol, and 200 μl of 2:1 chloroform:methanol, followed by sonication and centrifugation to pellet the protein precipitate. The supernatant was removed and dried under vacuum before the samples were reconstituted in equal volumes (50 μl) of water saturated butanol and 10 mM ammonium formate in methanol. Analysis of mammary gland extracts was performed on an Agilent 6495C QQQ mass spectrometer with an Agilent 1290 series HPLC system and a ZORBAX eclipse plus C18 column (2.1x100 mm, 1.8 μm, Agilent) with the thermostat set at 45°C as described previously.[83] Mass spectrometry analysis was performed in both positive and negative ion modes (negative ion mode was used for free fatty acid species) with dynamic scheduled multiple reaction monitoring (MRM) and polarity switching. The solvent system consisted of solvent A: 50% H_2O / 30% acetonitrile / 20% isopropanol (v/v/v) containing 10 mM ammonium formate and solvent B: 1% H_2O / 9% acetonitrile / 90% isopropanol (v/v/v) containing 10 mM ammonium formate. The gradient was as follows; starting with a flow rate of 0.4 ml/minute at 15% B and increasing to 50% B over 2.5 minutes, then to 57% B over 0.1 minutes, to 70% B over 6.4 minutes, to 93% B over 0.1 minute, to 96% B over 1.9 minutes and finally to 100% B over 0.1 minute. The solvent was then held at 100% B for 0.9 minutes. Equilibration was as follows; solvent was decreased from 100% B to 15% B over 0.2 minutes and held for an additional 3.8 minutes. The following mass spectrometer conditions were used; gas temperature, 150°C, gas flow rate 17 L/min, nebulizer 20psi, sheath gas temperature 200°C, capillary voltage 3500V and sheath gas flow 10 L/min. Lipid species within each class were analyzed by the same method. The levels of individual lipid species were calculated by taking a ratio of the area under the curve of the lipid of interest to the area under the curve of corresponding internal standard. The ratio was then multiplied by the amount of internal standard added into the sample. The levels for lipid classes were calculated from the sum of individual species within each class.

Single cell capture and library preparation for sequencing

To obtain sufficient cells for scRNA-seq analysis, mammary glands from two female mice per genotype were pooled. The 10x Chromium Single cell 3' V3 kit was used for freshly sorted mammary epithelial cell capture and cDNA preparation according to the 10x single cell 3' protocol.

scRNA-seq bioinformatic analysis

10X single-cell RNA sequencing data were processed using the cellCounts function included in Rsubread package (version 2.6.4).[70] Sequencing reads were mapped to the mouse reference genome GRCm38/mm10. Mapped reads were assigned to NCBI RefSeq mouse genes (build 38.1) and then collapsed to UMIs. Samples were demultiplexed, and cell barcodes were called based on known barcode sequences included in the assay kit and gene expression profile of each barcode. The called barcodes must have a gene expression profile distinct from those of ambient RNA barcodes detected in the data. Cells with $\geq$20% mitochondrial UMIs, $\geq$40% ribosomal UMIs or <200 or >5000 expressed genes were excluded from further analysis. Genes with no known symbols, mitochondrial and ribosomal genes and genes expressed in fewer than 3 cells in every sample were also removed from the analysis. 11258 MECs from Ehf^{WT} and 7275 from Ehf^{KO} mice were obtained, of which 9871 and 6567 respectively, were included in the final analysis after cell filtering. UMI counts were loaded into R and the Seurat pipeline[71] in R was used to select variable genes, cluster cells, and perform dimension reduction. Cell types were annotated according to one of the 15 subtypes identified in Bach et al.,[32] (GSE106273: GSM2834498: NP_1, GSM2834499: NP_2, GSM2834500: G_1, GSM2834501: G_2, GSM2834502: L_1, GSM2834503: L_2; GSM2834504: PI_1, GSM2834505: PI_2). The Bach et al. data were converted to $\log_2$-CPM, quantile normalized and precision weighted with the *voom* function of the *limma* package.[75,76] A reference gene expression vector for each of the 15 cell types was derived by fitting linear models to each gene using expression data of the gene in all cells belonging to a cell type. The reference vector contained representative gene expression levels for a cell type. We then performed Pearson correlation between the gene expression levels of each cell in the Ehf^{WT} and Ehf^{KO} sample and each of the 15 reference vectors. Each cell was assigned to a cell type where it had the highest Pearson correlation. Where the highest Pearson Correlation was lower than 0.3 the cells were annotated us 'Unclassified'. The UMI count matrix was provided to Monocle 2[72–74] for pseudotime trajectory reconstruction from the unintegrated Ehf^{WT} and Ehf^{KO} samples. Monocle 2 selects genes differentially expressed between clusters of cells to define biological process and employs a reversed graph embedding algorithm to reconstruct a pseudotime trajectory in an unsupervised manner. This algorithm maps the high dimensional gene expression space to a much lower dimensional space in which structure of the trajectory is determined. Monocle 2 detects the sequence of changes to global gene expression levels as a cell progresses through the biological process and identifies branch points that describe significant divergences in cellular state automatically. Significantly differentially expressed genes (DEGs) were identified between the Ehf^{WT} and Ehf^{KO} samples for each of the identified cell types. Cells that failed to achieve a Pearson coefficient of 0.3 in the correlation analysis were excluded. UMI counts of genes were converted to $\log_2$-CPM, quantile normalized and precision weighted with *voom*.[75] A linear model was fitted to each gene, and an empirical Bayes moderated *t* statistic were used to assess differences in expression.[84] DE genes were called with an FDR cut-off of 0.05, and a mean UMI value greater than 1 in at least one of the Ehf^{WT} or Ehf^{KO} samples. The enrichment of gene sets in the expression changes between Ehf^{WT} and Ehf^{KO} samples for Avd-G cells was performed using the *mroast* function in the *limma* package[75,76] to test the enrichment of Hallmark genesets in the Molecular Signatures Database (MSigDB) database. An FDR cut-off of 0.05 was used for identifying enriched Hallmark gene sets.

Library preparation for ATAC sequencing

Nuclei were extracted from 50,000 cells per sample by resuspending in 50 μl of ATAC-seq lysis buffer (10 mM Tris-HCl (pH 7.5), 10 mM NaCl, 3 mM MgCl2, 0.1% Tween-20, 0.1% IGEPAL CA-630 (Sigma Aldrich), 0.01% Digitonin (Sigma Aldrich)) and incubating on ice for 3 minutes followed by washing with 10 mM Tris-HCl (pH 7.5), 10 mM NaCl, 3 mM MgCl2, 0.1% Tween-20. Samples were then centrifuged at 500 *g* for 10 minutes at 4°C and the supernatant discarded. Transposition of the DNA and ligation of adapters was performed using the Tagment DNA TDE1 Enzyme and buffer kit (Illumina) as per the manufacturer's instructions. Briefly, DNA was digested by incubation with Tn5 Transposase enzyme at 37°C for 30 minutes. Transposed DNA was then purified using the MinElute reaction cleanup kit (Qiagen), and partially PCR-amplified using Nextera custom barcoded primers[85] for 5 cycles. To reduce GC-bias, the number of additional PCR cycles required was determined to be 9 cycles (14 cycles in total) by using a SYBR Green quantitative real time PCR reaction and calculating the number of cycles corresponding to 1/3 of the maximum fluorescent intensity.[86] Libraries were then purified by AMPure XP magnetic bead purification (Beckman Coulter). DNA quantity and quality was assessed using the Tapestation 2200 system (Agilent). All libraries were sequenced on the Illumina NextSeq500 with single end 130 bp and 1x8 base index reads.

ATAC-seq bioinformatic analysis

ATAC-seq Reads were aligned to the GRCm39/mm39 build of the mouse reference genome using the *Subread aligner*[77] in the Subread package (version 2.0.6). Only uniquely mapped reads were included. Peaks in each individual sample were called by using *HOMER*[78] (v4.11) with a false discovery rate (FDR) cutoff of <1e-5. The peaks from all the samples were then merged by taking unions of the overlapping peaks which produced 195818 merged regions. The mapped reads in the individual samples were then assigned to the merged regions using *featureCounts*.[79] Multi- assignment of a single read was allowed. Merged regions were required to have at least 10 CPMs in at least 4 samples to be included in the analysis. Read counts were converted to $\log_2$-CPM, quantile

normalized and precision weighted using the *voom* function of the *limma* package.[75,76] Batch effect was removed by using the *removeBatchEffect* function in *limma*. Sample clustering was performed based on multi-dimensional scaling. A correlation analysis was performed between the 352 DE genes identified in the Avd-G population from the scRNA-seq analysis and the regions of differentially accessible chromatin identified in the luminal cell population from Ehf^{WT} and Ehf^{KO} samples. Out of the 352 DE genes, 331 genes commonly found between the scRNA-seq data and the ATAC-seq data were included in the correlation analysis. Pearson correlation was performed by first calculating the $\log_2$ CPM values of promoter regions of all genes in the ATAC-seq data. The promoter regions were defined as the 5kb up and 2kb down region around the TSS of each gene in the mm39 RefSeq gene annotation. A scatter plot was generated to show the $\log_2$ fold changes in scRNA-seq and ATAC-seq data. Gene read coverage track figures show the coverage levels (CPM) of the ATAC-seq reads in the promoter regions of the selected genes. The selected genes have fold changes $\geq$1.2 in scRNA-seq Avd-G data and in ATAC-seq luminal data. The denominators for calculating CPM values are the number of mapped reads in each library. 50-bp bins in the promoter regions were used.

Clonogenic capacity assays of primary mammary epithelial cells

For colony forming assays, 1000 sorted mammary luminal epithelial cells from 18.5 dP Ehf^{WT} and Ehf^{KO} mice were embedded in Cultrex Reduced Growth Factor Basement Membrane Extract, Type 2, Pathclear (R&D systems). 20μl domes were then pipetted on a 24-well suspension plate (Greiner Bio-One) and cultured for 11 days in a low O_2 incubator in Advanced DMEM-F12 (Themo Fisher Scientific) supplemented with 5% heat inactivated fetal bovine serum, 1 mM L-glutamine (Thermo Fisher Scientific), 5 μg/ml insulin (Austin pharmacy), 10 ng/ml mouse recombinant epidermal growth factor (R&D systems), 0.5 μg/ml hydrocortisone (Austin Pharmacy) and 20 ng/ml cholera toxin (Sigma Aldrich). Colonies were scored manually, and images were captured on an Olympus CKX53 microscope using the LCACHN 20x objective (NA 0.4) (Olympus, Japan) and cellSens standard software (Olympus, Japan).

Primary mammary and tumour organoid culture

Primary mammary organoids and mouse mammary tumour organoids were generated using established protocols.[87–89] Specifically, the thoracic and inguinal mammary glands from 10-week-old nulliparous Ehf^{WT} and Ehf^{KO} mice were resected (lymph nodes removed). For *MMTV-PyMT* tumour organoid generation, the largest primary mammary tumour at ethical endpoint was resected. The resected tissue was collected into Advanced DMEM-F12 (Thermo Fisher Scientific) supplemented with 1 mM L-glutamine (Thermofisher Scientific), Penicillin-Streptomycin (Thermo Fisher Scientific) and 10 mM HEPES (Thermo Fisher Scientific) on ice. In a tissue culture hood, samples were mechanically dissociated by chopping into ~1 mm^3 pieces, then digested in Advanced DMEM-F12 containing 300U/mL Collagenase Type 2 (Thermo Fisher Scientific) supplemented with 1 mM glutamine (Thermo Fisher Scientific), Penicillin-Streptomycin (Thermo Fisher Scientific) and 10mM HEPES (Thermo Fisher Scientific), 5 μg/mL Insulin (Symbion), 500 ng/mL Hydrocortisone (Symbion), 10 ng/mL Mouse recombinant epidermal growth factor (R&D systems) and 5% Heat inactivated fetal bovine serum. Samples were dissociated by repeat pipetting then incubated on an orbital shaker at 37°C for 45-60 minutes, vortexing every 15 minutes. Following digestion, 10 mLs of cold wash medium was added, and the cell digestion was passed through a 100 μM MACS smart strainer (Miltenyi Biotec 130-110-917) and pelleted by centrifugation at 400*g* for 5 minutes at 4°C. The supernatant was removed, and the cell pellet was washed in 10 mL of cold washing medium, then again pelleted by centrifugation at 400*g* for 5 minutes at 4°C. The supernatant was again removed, and the cell pellet was resuspended in cold Cultrex Reduced Growth Factor Basement Membrane Extract, Type 2, Pathclear (R&D systems). 2x 25 μl domes were pipetted per well in 3 wells of a 24-well suspension plate (Greiner Bio-One, 662102), and domes set at 37°C for 30 minutes. For the growth of primary mammary organoids, Basal organoid medium was added to each well containing Advanced DMEM-F12 (Thermofisher Scientific) supplemented with 1 mM L-glutamine (Thermofisher Scientific), Penicillin-Streptomycin (Thermofisher Scientific) and 10 mM HEPES (Thermofisher Scientific), 5 μg/mL insulin, 100 ng/mL hydrocortisone, B-27 Supplement minus vitamin A (Thermofisher Scientific), 1.25 μM N-Acetylcysteine (Sigma Aldrich), 50 ng/mL Recombinant mouse epidermal growth factor (R&D systems), 5 ng/mL Recombinant Mouse FGF basic (R&D systems), 10 ng/mL Recombinant Mouse FGF-10 (R&D systems), 10 ng/mL Recombinant Mouse Wnt-3a (R&D systems), 4 μg/mL Heparin sodium salt (R&D systems), 50 ng/mL Recombinant Murine R-Spondin-1 (Peprotech) with 5 μM Y-27632 (SelleckChem) added for the first 72 hours of culturing. *MMTV-PyMT* tumour organoids were cultured in Basal organoid medium as described above, excluding Recombinant Mouse FGF-10 and Recombinant Mouse Wnt-3a.

RealTime-Glo™ MT Cell Viability Assay

Viability of primary mammary organoids and *MMTV-PyMT* tumour organoids were determined using a protocol established by Morrow et al.[90] For primary mammary organoids, organoids generated from mammary gland from 3 Ehf^{WT} and 3 Ehf^{KO} 10-week-old adult nulliparous mice were cultured for two passages before assessing viability. For *MMTV-PyMT* tumour organoids, the viability of organoids generated from the largest primary mammary tumour from 4 Ehf^{WT}*MMTV-PyMT* and 4 Ehf^{KO} *MMTV-PyMT* mice at ethical endpoint was assessed. Organoids were disrupted into single cells and plated in technical triplicates into 96 well Uclear white edged tissue culture plates (Interpath) at a seeding density of 20,000 cells per well in 50 μl Reduced Growth Factor Basement Membrane Extract, Type 2, Pathclear (R&D systems). RealTime-Glo™ MT Cell Viability Assay (Promega) reagents were prepared according to the manufacturer's instructions in basal organoid culture medium and 200 μl added per well. Organoids were cultured in a humidified 37°C low oxygen incubator and imaged at the same time daily for between three and four days. Luminescence was imaged on an EnSight multimode plate reader (Perkin Elmer) and quantified with Kaleido 3.0 software (Perkin Elmer). For primary mammary organoids the Day 0 reading was taken 24 hours post seeding; For *MMTV-PyMT* tumour organoids the Day 0 reading was taken at

2 hours post seeding. Wells containing 50 μl Type 2 Reduced Growth Factor Basement Membrane Extract, Culture medium and assay reagents were used as a reference for background signal. Luminescence was calculated as fold change to the Day 0 reading with background luminescence removed.

Imaging and Prolactin treatment of organoids

Images of organoids were captured on an Olympus CKX53 microscope using the UPLFLN IPC 4x objective (NA 0.13) (Olympus, Japan) and cellSens standard software (Olympus, Japan). Prolactin treatment experiments were performed by plating Ehf^{WT} and Ehf^{KO} organoids at a matched seeding density into 6x 25 μl domes and culturing until equivalent confluency (72-96 hours). Organoids were then washed in PBS and cultured in Basal organoid medium containing 0.1% BSA (vehicle control) or Lactation medium containing Advanced DMEM-F12 (Thermo Fisher Scientific) supplemented with 1 mM L-glutamine (Thermofisher Scientific), 100 U/ml Penicillin 100 μg/ml Streptomycin (Thermo Fisher Scientific), 10 mM HEPES (Thermo Fisher Scientific), 5 μg/mL insulin (Symbion), 100 ng/mL hydrocortisone (Symbion) and 5 μg/mL recombinant mouse Prolactin protein (R&D systems) for 96 hours, then harvested for RNA isolation. Mammary organoids were cultured in a humidified 37°C low oxygen incubator for the duration of the study.

Harvesting of primary mammary organoids and RNA isolation

Culture medium was removed, and organoids were washed with cold PBS, then collected by incubation for 1 hour at 4°C on an orbital shaker in 1 mL of cold Cell recovery solution per well (Invitro technologies). Pipette tips and tubes were precoated with 1% wt/vol PBS-BSA. Organoids were transferred with precoated pipette tips into precoated tubes and centrifuged at 200*g* for 5 minutes at 4°C to pellet, the supernatant was removed, and the organoid pellet was resuspended in 350 μl RLT lysis buffer (Qiagen), and RNA extraction performed using the RNeasy micro kit (Qiagen) according to the manufacturer's instructions, with the DNase I treatment step included. cDNA synthesis and quantitative real time polymerase chain reaction were performed as described for whole mammary gland tissue.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical details can be found in relevant figures, figure legends or results. All data are presented as mean ± standard error of mean (SEM) unless otherwise noted. Statistical analyses were performed using GraphPad Prism v8.0 software (GraphPad Software, La Jolla, CA, USA). Groups were typically compared using the student's T-test or Chi-square analysis. In all cases $p<0.05$ was considered statistically significant.