AWARD NUMBER: W81XWH-17-1-0013

TITLE: Targeting Basal Breast Cancer

PRINCIPAL INVESTIGATOR: Dr. Pamela Cowin, Ph.D.

CONTRACTING ORGANIZATION: New York University School of Medicine, New York, NY

REPORT DATE: March 2021

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
March 2021	Annual Year	01Sep2020-28Feb2021

4. TITLE AND SUBTITLE
Targeting Basal Breast Cancer

5a. CONTRACT NUMBER
W81XWH-17-1-0013

5b. GRANT NUMBER
BC160959

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Pamela Cowin PhD

E-Mail: cowinp01@nyumc.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
New York University School of Medicine
550 First Ave
New York 10016

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT: Our goal was to determine 1) if Gpr identifies mammary stem and/or progenitors; 2) the significance of Gpr expression in murine and human breast tumors. We proposed to: a) identify and characterize Gpr+ cells, determine their potency by lineage tracing, assess the effect of Gpr cell ablation on mammary development; b) determine Gpr expression in human breast cancers and cell lines, and test if ablating Gpr+ cells affects mammary tumorigenesis in mouse models. In this six month grant period we have 1) identified genes that permit distinct Gpr-positive progenitor subpopulations to be identified thereby permitting their differential gene expression profiles to be compared; 2) transplanted Gpr cells to test their plasticity 3) demonstrated association of high Gpr expression with specific TNBC subsets and basal-type breast cancer cell lines; 4) re-established the cohort of compound tumor mice that were lost due to COVID shutdown to permit ablation studies to proceed; 5) established organoid cultures to permit parallel studies *in vitro.* Our results to date have shown that Gpr identifies mammary stem cells in the embryo and unipotent progenitors in perinatal and postnatal mammary gland located at invasive ductal tips. Our data in this period consolidate the finding that Gpr+ is a highly specific marker of multiple distinct mammary progenitors and that high Gpr expression is associated with early tumor onset and poor outcome in the most aggressive types of breast cancer.

15. SUBJECT TERMS
Cell adhesion, G-protein coupled receptor, stem cell marker, basal breast cancer

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U Unclassified	U Unclassified	U Unclassified	UU Unclassified	108	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Page

1. INTRODUCTION

One third of women with basal-type breast cancer have very poor outcome. The majority, however, have much lower risk yet receive similar treatment because they cannot be reliably distinguished from those with aggressive forms. Reproductive history is a significant risk factor and it has been proposed that this stems from proliferation occurring during ovulation and pregnancy exposing long-lived mammary stem cells (MaSCs) and progenitors to replicative error. Currently markers of breast stem cells and progenitors are neither specific nor comprehensive: those described to date label mutually exclusive subpopulations at distinct locations and their non-specificity for breast epithelial cells limits their utility as prognostic markers. We proposed to validate an adhesion G-protein coupled receptor (Gpr) as a novel and specific marker of mammary stem/progenitors and cancer stem cells.

2. KEYWORDS

Cell Adhesion, G-protein coupled receptors, Stem cell marker, Basal Breast Cancer

3. ACCOMPLISHMENTS

The major goals of the project were:

Task 1) Determine the role of Gpr+ cells in mammary development. (95% complete)

Task 2) Determine the significance of Gpr expression in human and mouse breast cancers. (75% complete)

- ***What was accomplished under these goals:***

Specific Aim 1. Determine the role of Gpr+ cells in mammary development.

Subtask 1. Validate expression of Gpr in Gpr-DTR-creER^{T2} and Gpr-lacZ mice by reporter expression, immunohistochemistry and FACS analysis. (1-12 months)

In Year 1 (Y1) we documented the generation of two mouse strains Gpr-lacZ and Gpr-DTR - creER^{T2} (Y1 Figs. 1 and 4). The Gpr-lacZ reporter (Y1 Fig 1) enabled spatial detection of Gpr expression by X-gal detection of beta-galactosidase fused to the Gpr ectodomain (Y1 Figs. 2, 3, 7). Colocalization with lineage markers (Y1 Fig. 6) showed that Gpr is expressed in a basal subpopulation at sites of predicted embryonic, pubertal and pregnancy-induced stem/progenitor activity. qPCR analysis of mRNA expression confirmed the dynamic developmental expression pattern (Y1 Fig. 7). FACS analysis indicated that Gpr+ cells expressed markers characteristic of regenerative mammary cells and lacked those of more committed cell types (Y1 Fig. 8). In year 2 (Y2) we investigated overlap of Gpr+ cells with other putative regenerative cell populations by FACS analysis (Y2 Fig.1) and demonstrated that Gpr+ cells clustered with cells expressing basal gene markers by mining scRNAseq datasets (Y2 Fig. 2, Fig 3). In year 3 (Y3) we found Gpr expression in embryonic cells expressing bilineage markers during early mammary specification (E15) (Y3 Fig. 1) and at later stages within keratin negative cell-types present at ductal tips

poised for ductal invasion of the fatpad (Y3 Fig. 1). In scRNAseq datasets we found Gpr in clusters of embryonic cells expressing bilineage markers (Y3 Fig. 2) but restricted to basal clusters in pubertal and adult mice (Y3 Fig. 3), together with s-Ship, Lgr5 and ProcR regenerative subpopulations. In summary, these analyses showed that Gpr appears to encompass and specifically mark multiple distinct subpopulations of mammary progenitors.

During the current reporting period (6 months 09.01.20-02.28.21) we have investigated further the relationship among these progenitor subpopulations. We have crossed Gpr-lacZ mice to s-Ship-EGFP and Lgr5-GFP strains and so far have colocalized Gpr-lacZ and the s-Ship-EGFP reporters in pubertal terminal end buds (TEB) **(Fig. 1).** We have crossed these bitransgenic mice to MMTV-Wnt1 and harvested the Gpr-lacZ;s-Ship-EGFP; MMTV-Wnt1 and Gpr-lacZ;Lgr5-EGFP;MMTV-Wnt1 glands and tumors **(Fig. 2).** Thus, we are now poised to test for reporter colocalization in embryos, pregnant glands, hyperplastic glands and tumors.

Fig. 1. Gpr and sShip are coexpressed in the outer cap cell layer of pubertal terminal end buds.
Gpr-lacZ expression detected by X-Gal staining (blue stain) and s-Ship-EGFP detected by immunolocalization with anti-EGFP (brown stain). Image on the left is counterstained with hematoxylin (purple). Note the colocalization to the outer flattened cap cell layer of these terminal end buds

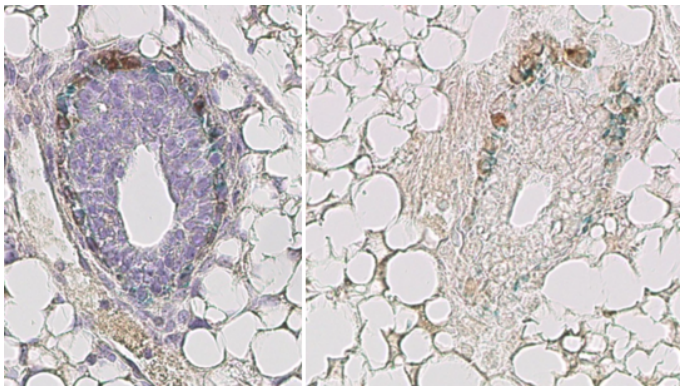


Fig. 2. Gpr-lacZ expression in MMTV-Wnt1 tumors.
Gpr-lacZ expression detected by X-Gal staining (blue stain) in a tumor whole mount derived from a Gpr-lacZ;s-Ship-EGFP;MMTV-Wnt1 mouse.

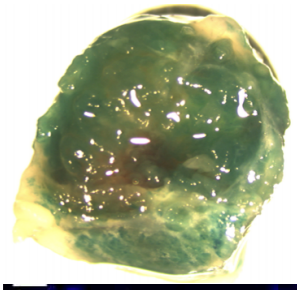
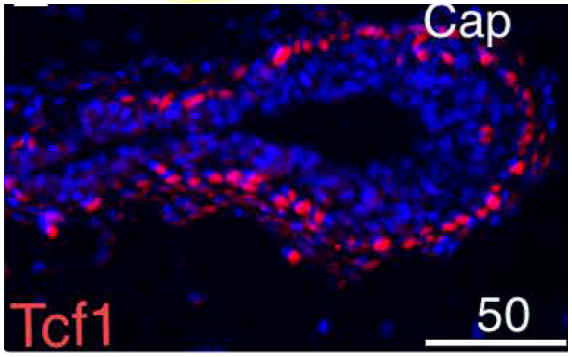


Fig. 3. Tcf1 encoded by the gene Tcf7 is specifically expressed in cap cell layer of TEB.
Immunolocalization of the Wnt-responsive transcription factor, Tcf1 (red stain), in a frozen section through a pubertal TEB, counterstained with DAPI (blue stain) to detect nuclei. Note Tcf1 is detected within nuclei of cells in the outer cap cell layer.



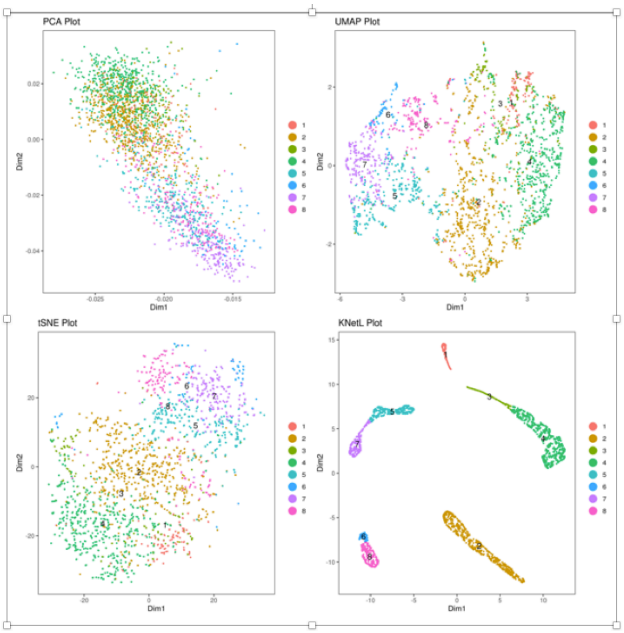
We have shown that Gpr-positive TEB cap cells specifically express the Wnt-sensitive transcription factor and tumor suppressor, Tcf1 (Fig. 3) encoded by the gene *Tcf7*. Given its

specificity for this cell type, we have used *Tcf7* expression to identify Gpr-positive cap cell populations in scRNAseq datasets. To compare the gene expression profiles of Gpr-positive subpopulations present in pubertal glands we have analyzed the scRNAseq dataset of Pal et al using *Tcf7* to identify cap cells and *Lgr5* to identify nipple proximal cells. Using KNetL plots we obtained the best segregation of cell clusters (Fig. 4).

Fig. 4. Cluster analysis of pubertal basal cell population scRNAseq dataset of Pal et al.

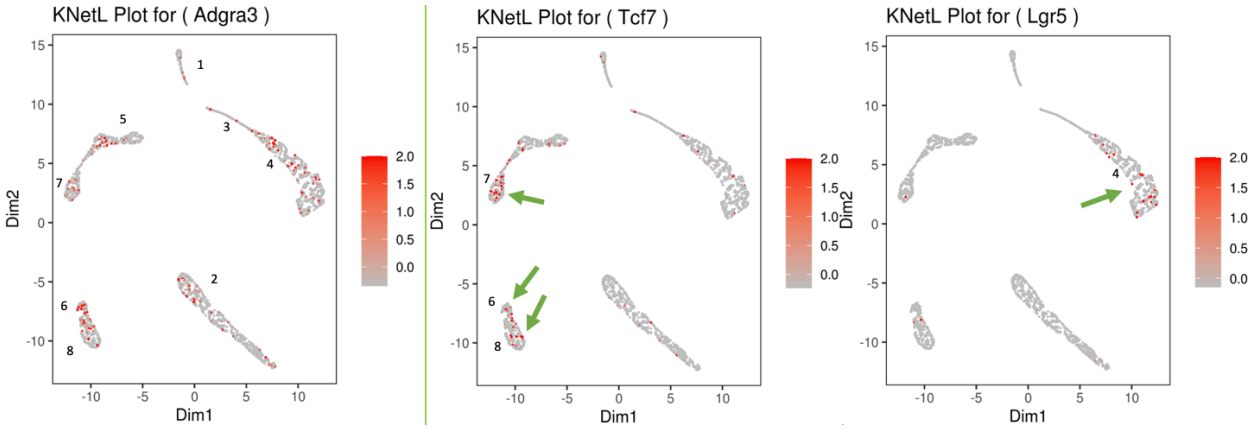
Top left: PCA
Top right: UMAP
Bottom left: tSNE
Bottom right: KNetL

Note: The KNetl algorithm provides the best cell cluster separation



As expected Gpr was present in many clusters, however using *Tcf7* expression as a marker, we identified the Gpr-positive cap cell population in clusters 6,7,8 and the Gpr-positive *Lgr5* expressing nipple proximal subpopulation in cluster 4. The remaining clusters contain sporadic Gpr-positive populations embedded along maturing ductal myoepithelium.

Fig. 5 Identification of basal subpopulations in scRNAseq clusters from pubertal glands
Left: Gpr (Adgra3) is expressed in most clusters. Center: *Tcf7* is specifically expressed in clusters 6,7,8 (green arrows) identifying these as cap cells. Right: In contrast *Lgr5* is exclusively found in cluster 4 (green arrow) identifying these cells as the nipple-proximal Gpr subpopulation.



Preliminary analysis of markers differentially expressed between these subpopulations (Fig. 5) confirms the low expression of genes encoding keratins and contractile proteins in the cap cells compared to the nipple proximal subpopulation. Moreover, cap cells are enriched in *Crispld2/lgl1,* first identified as a Drosophila tumor suppressor that function in basolateral polarity and cell migration and is regulated by the breast tumor suppressor *PTEN*.

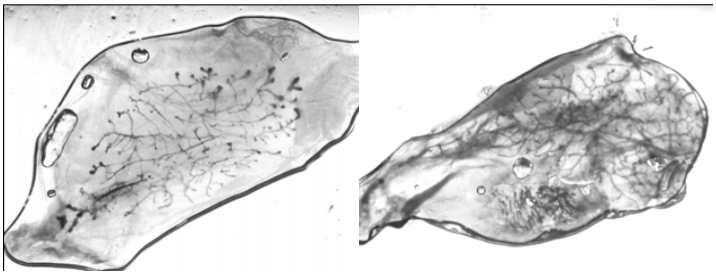
Subtask 2 and 3. Cross Gpr-DTR-creER^{T2} to Rosa26-STOP-reporter lines (1-18 months). Trace the Gpr lineage by inducing cre activity with TAM and detecting Rosa26R-reporter in Gpr descendants in conjunction with differentiation markers (18-30 months).

In this aim we carried out lineage tracing to address whether Gpr+ cells are stem cells or lineage-restricted progenitors during normal physiological development. In Y1 we crossed Gpr-DTR-creER^{T2} mice to Rosa26R-STOP-tdTomato (tdT) lines. When tamoxifen (Tam) is administered to these mice the tdT reporter is expressed in Gpr+ parental cells and in all cellular progeny. We established 3D confocal imaging of cleared tissue whole mounts (Y1 Fig 9 and 10) and showed tdT labeled cap cells in pubertal terminal end buds (Y1 Fig. 10) and cells in the basal layer of ducts of pubertal mice (Y1 Fig. 11). In Y2 we traced the tdT progeny of pubertal Gpr+ cells in mature ducts (Y2 Fig. 4), in alveoli and ducts during pregnancy (Y2 Fig. 5), in lactating glands (Y2 Fig 6.) and colocalized them with lineage-specific markers by immunofluorescence (Y2 Fig 7 and 8). We traced Gpr cell progeny after 6 months in virgin mice and after completing two cycles of pregnancy and involution (Y2 Fig. 9). These data demonstrated that pubertal Gpr+ cells are long-lived unipotent basal stem/progenitors. In Y3 we extended these findings by showing that cells expressing Gpr on leading tips of branches during pregnancy also function as unipotent basal progenitors (Y3 Fig. 7). In contrast, we demonstrated that early embryonic Gpr+ cells are bipotent at E14.5 and become restricted to unipotency around E19.5 (Y3 Fig. 6).

During the current reporting period we have begun to address the plasticity of Gpr+ progenitors by isolating MECS from Gpr-DTR-creER^{T2} mice and transplanting them into the cleared fat-pads of immunocompromised Nude mice. Both control wildtype and Gpr-DTR-creER^{T2} MECS generated mammary trees. However, the ductal trees produced by Gpr-DTR-creER^{T2} glands were disorganized and less directed in their outgrowth pattern than controls. (Fig. 6).

In this reporting period, to test if the preexisting unipotent pubertal Gpr+ progenitors remain unipotent or gain bipotency in a regenerative environment we have transplanted MECS from 6-week old pubertal donors that received Tam or vehicle one week beforehand. We are analyzing these glands by 3D confocal imaging to detect the fate of their cell progeny. To test if new progenitors generated under regenerative stimulation are unipotent or bipotent, we will repeat this experiment but giving Tam to host mice 1week post-transplantation. To investigate the effect of other cell types on the plasticity of progenitors we plan to compare transplantations of FACS purified Gpr+ MECs to those transplanted together with other mammary cell types.

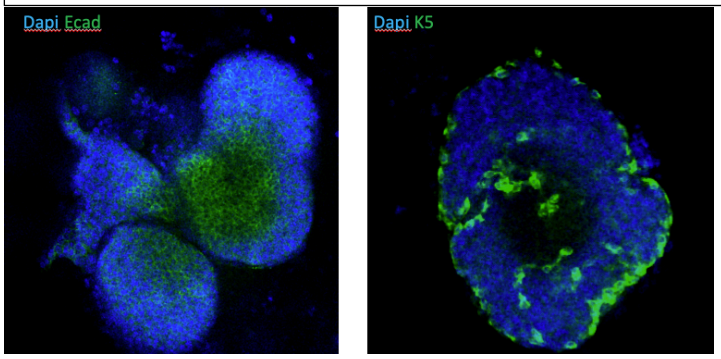
Fig. 6 Outgrowths from MECS transplanted into the cleared fat-pads of immunocompromised hosts
Outgrowths harvested 1 month after transplanting 100K MECS into the cleared fatpads of host Nude mice. Glands were fixed and stained with carmine to image the ducts and cleared of fat.
Left: Outgrowth from a control Gpr-wt MECS
Right: Outgrowth from Gpr-DTR-creER^{T2} MECS



Subtask 4 and 5. Ablate Gpr+ cells in Gpr-DTR:EGFP-creER^{T2} mice by administration of DTA, validate cell death by loss of Gpr expression and monitor the effect on mammary development.

Initially we proposed to examine the effects of killing Gpr-positive cells on mammary gland structure by administering Diptheria Toxin (DT) to 6-week old Gpr-DTR-creER^{T2}mice. However, this proved to be unfeasible because the mice deteriorated rapidly due to Gpr expression in vital organs (Y2 Fig. 10). To overcome this obstacle, we transplanted whole mammary glands from Gpr-DTR-creER^{T2} donors into immunocompromised Foxn1 nu mice but again encountered problems because suturing of the transplanted glands interfered with analysis of the mammary outgrowth. To circumvent this problem, we have transplanted mammary epithelial cells into a pilot cohort of mice. These have successfully generated outgrowths (Fig. 6) and we are now poised to examine the effects of administering DT on the integrity of these glands. Moreover, to permit faster analysis of the effects of DT on ductal morphogenesis, we have established mammary organoid cultures from Gpr-DTR-creER^{T2} mice (Fig. 7) that can be treated *in vitro*.

Fig. 7 Organoids derived from in vitro culture of MECS in Matrigel
Organoids established in Matrigel containing RSPO1, and BMP supplements, were cultured for 1 week then processed for tissue clearing, immunofluorescent detection and 3D confocal imaging. Left: luminal cell marker, E-cadherin (green) and Right: basal cell marker, Cytokeratin 5 (green); DAPI was used as a counterstain to visualize nuclei

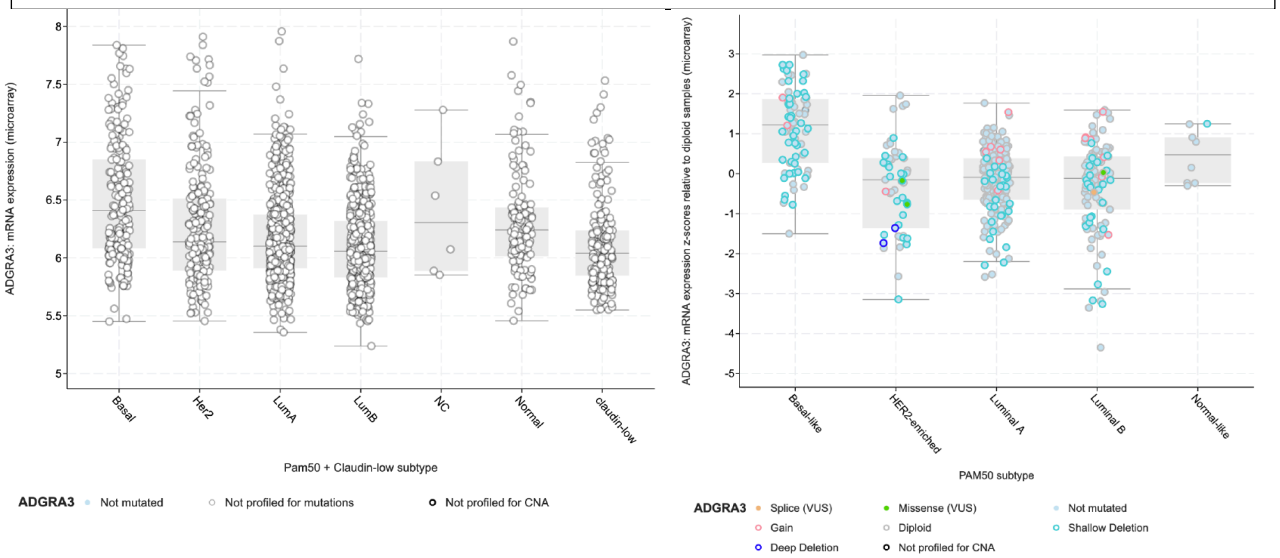


Specific Aim 2. To determine the significance of Gpr expression in breast cancers.
Task 1. Screen breast cancer cell lines for Gpr expression by qPCR and mine bioinformatics datasets of human breast cancers.

In Y1 we mined public datasets using the kmplotter algorithm to show that high expression of Gpr predicts poor prognosis in relapse free survival exclusively in the basal breast cancer subtype in human breast cancers (Y1 Fig 16). In Y3 we showed that high Gpr levels within this subtype correlate with worse patient outcome in terms of with disease-free survival (Y3 Fig. 10).

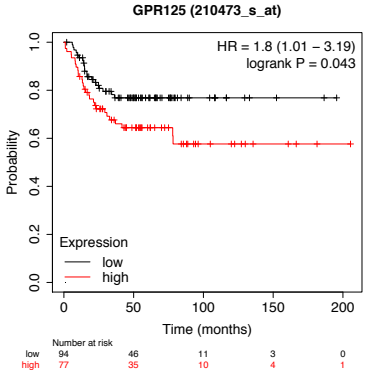
In the current reporting period we have interrogated the TCGA and METABRIC datasets and found Gpr is more highly expressed in ER-negative tumors compared to ER-positive tumors and in basal-type tumors compared to other PAM50 subgroups (Fig. 8)

Fig. 8. Gpr expression in human breast cancer subtypes
Highest Gpr expression is found in the basal breast cancer compared to other PAM50 subtypes in: Left) METABRIC dataset and Right) TCGA dataset



Utilizing Kmplotter to probe the Pietenpol classification of TNBC cancers, we have found high expression of Gpr is associated with poor outcome in terms of distant metastasis free survival specifically within the basal-like 1 (BL1) subtype of triple negative breast cancers (Fig. 9).

Fig. 9. Distant Metastasis-Free Survival
High expression of Gpr (red line) is associated with worse outcome in patients with BL1 subtype TNBC tumors (Pietenpol classification). Data derived using the kmplotter algorithm with best fit cutoff.



We have assessed Gpr expression in breast cancer cell lines by mining the DepMap Portal dataset and found Gpr is most highly expressed in TNBC cell lines compared to luminal cell lines and has the lowest expression in HER2-amplied cells.

Expression	Cell Line Name	Subtype Classification
5.714245518	MDAMB436	Basal B
5.219555769	HCC1143	Basal A
5.151371776	HCC1187	Basal A
5.092545742	CAL51	Basal B
4.810957159	MDAMB361	HER2 ER+
4.803227036	SUM159PT	Basal B
4.762348816	HCC70	Basal A
4.704871964	CAL148	LUM
4.602290585	HCC1569	Basal A
4.584361253	MDAMB231	Basal B
4.507160349	HMC18	Basal B
4.403949364	HCC1599	Basal A
4.311793718	HCC1954	Basal A
4.207892852	CAL851	Basal A
4.107687869	BT549	Basal B
4.032100843	HCC1500	Lum
4.027684877	BT20	Basal A
3.921245899	BT483	LUM
3.899473543	CAMA1	LUM
3.887525271	HS578T	Basal B
3.868884273	HCC1419	HER2amp
3.853995647	ZR751	LUM
3.846994687	CAL120	Basal B
3.821710215	EFM192A	HER2 amp
3.815575429	SUM229PE	Basal
3.791814071	SUM149PT	Basal B
3.772941338	BT474	HER2 amp
3.71259578	HCC1806	Basal A
3.683696454	MDAMB157	Basal A
3.673556424	HCC1395	Basal B
3.66108548	HDQP1	Basal A
3.578938713	SUM1315MO2	Basal
3.56315813	HCC2157	Basal A
3.513490746	MDAMB468	Basal A
3.379898164	MFM223	LUM
3.336283388	KPL1	LUM
3.264536431	UACC812	LUM
3.240314329	ZR7530	HER2 amp
3.163498732	T47D	LUM
3.152183419	MDAMB415	HER2 amp
3.116031993	HCC1428	LUM
3.051372102	MCF 7.00	LUM
2.987320866	EFM19	LUM
2.922197848	SUM52PE	LUM
2.912649885	HCC38	Basal A
2.887525271	HCC1937	Basal A
2.835924074	MDAMB175VII	HER2 amp
2.817623258	MDAMB453	HER2 amp
2.671293372	DU4475	LUM
2.508428653	SUM185PE	LUM
2.432959407	SUM44PE	LUM
2.301587647	HCC2218	HER2 amp
2.192194165	JIMT1	Basal A
2.014355293	SUM102PT	Basal
1.298658316	UACC893	HER2 amp
0.933572638	SKBR3	HER2 amp
0.50589093	AU565	HER2 amp
0.150559677	HCC202	HER2 amp

Fig. 9. Expression of Gpr in breast cancer cell lines in rank order of Gpr expression highest to lowest
Note that in general basal cell types (red) show higher levels of Gpr expression than luminal (green) or HER2 amplified (orange) cell types

Task 2 Test whether eradicating Gpr+ cells affect tumorigenesis and/or regression
Subtask 1: Generate bi-transgenic Gpr-DTR-creER^{T2}/MMTV-Wnt1 mice.
As reported previously the first experimental cohort of compound Gpr-DTR-creER^{T2}/Gpr-lacZ;MMTV-Wnt1 female mice began to produce tumors in March 2020 coinciding with a mandatory shutdown of all experiments due to the COVID-19 outbreak in New York and had to be euthanized. Dr. Spina, who volunteered for the very first pilot group that returned to work June 2020 set up the breeding again however the loss of this cohort of experimental animals set back our work by more than 8 months. The second cohort began developing tumors in February 2021. We have harvested and frozen organoids and cells from them. We now plan to transplant them into Foxn1 nu mice and expect the transplanted tumors to arise within 1-2 months.

Subtask 2: Ablate Gpr cells by DT administration and assess effects on tumor onset, progression, regression and histology.
In Y2 we reported that in contrast to wildtype mice, pubertal Gpr+ cells present in the MMTV-Wnt1 hyperplastic setting acquire bipotency (Y2 Fig. 12). We found Gpr cells were expanded in MMTV-Wnt1 tumors that develop with short latency and that these cells and retain embryonic features lacking keratin and SMA showing altered integrin levels and expressing transcription factors associated with embryonic and pubertal glands (Y3 Fig 11 and 12). Now that we have been able to regenerate tumors in donor mice we are poised to transplant MECs derived from these Gpr-DTR-creERT2 tumors and examine the effects of specifically ablating Gpr cells by injecting DT.

- **What opportunities for training and professional development has the project provided.**

  - **Dr. Cowin** acquired skills in tissue clearing and 3D-confocal imaging from colleagues in the Pathology Department, Cambridge who have pioneered this technique in the mammary gland (8).
  - **Dr. Spina** is a fully trained pharmacist and Ph.D. During the course of this project she has augmented her existing skills in breast cancer cell culture by acquiring new skills in a) mouse genetics b) histological analysis, c) tissue clearing and 3D-confocal imaging d) *in vivo* survival surgery and fat-pad clearing and transplantation. e) lineage tracing f) organoid isolation and mammosphere culture g) analysis of scRNAseq datasets
  - **Dr. Spina** took courses in ethics, animal welfare and professional career development. She attends the "works in progress" presentations of the stem cell and pharmacology and molecular oncology training programs as well as the breast cancer group within the NYU Cancer Center. She has presented in these venues as well as at the departmental retreat. She has met annually with her committee comprising: Dr. Mayumi Ito Ph.D. an expert in the lineage tracing of hair follicle and nail ectodermal appendages; Dr. Dimitris Placontonakis MD, an expert on adhesion-GPCR 133 in glioblastomas and Dr. Konstantin Itchenko, an expert in Adhesion-GPCR signal transduction.
  - **Dr. Spina and Cowin** have written two manuscripts together currently under revision and published a review on Embryonic Mammary development for Trends in Cell and Developmental Biology (see appendix).

- **How were the results disseminated to communities of interest?**
  - **Dr. Cowin** was scheduled to present these findings as an invited speaker at the Gordon Conference on Mammary Gland Biology in April 2020 (canceled due to COVID travel restrictions)
  - **Dr. Cowin** presented these findings to the Skirball Institute, NYUSOM Oct 9, 2019
  - **Dr. Cowin** presented these findings to the Dept of Opthalmology, NYUSOM on November 5, 2019
  - **Dr. Spina** was to have presented this work to the 10th Adhesion GPCR Workshop in Copenhagen (canceled due to COVID travel restrictions)
  - **Dr. Spina** presented this work to the Annual Skirball Institute Retreat in Oct 2019
  - **Dr. Spina** presented this work as a selected speaker for the Postdoctoral Association Research Day NYUSOM on September 23, 2019
  - **Dr. Spina** is scheduled to present this work by zoom at Princeton April 7 and MSKCC on April 8
  - **Drs. Spina and Cowin** have submitted two manuscripts and a review (see appendix)

- **What do you plan to accomplish during the next reporting period to accomplish the goals and objectives?**

- ❖ The attached manuscript was submitted to Nature Communications. In February we received reviewer's comments and were invited by the editor to submit a revised manuscript. Our current focus is therefore to carry out the requested experiments to bring the paper to publication. These entail providing greater insight into Gpr cell types through further scRNAseq and RNAseq analysis, exploring prognostic characteristics in TGCA and METABRIC datasets, determining plasticity of the Gpr cell type via transplantation and providing insight into potential role in cell migration via organoid experiments.
- ❖ Our second goal is, as originally planned, to treat glands, MECS and organoids with DT and examine the effects on ductal outgrowth and tumor propagation after transplantation into host mice

4. IMPACT

- ❖ **What was the impact on the development of the principal discipline of the project?**

  - ❖ Our expression and linage tracing studies in Y3 show that Gpr identifies mammary stem/bipotent progenitors in the embryonic mammary gland and unipotent basal progenitors in the perinatal pubertal and adult mammary gland.
  - ❖ Our results in Y1 showed that mice lacking this adhesion GPCR show delayed mammary ductal elongation and that the presence of the cytoplasmic signaling domain of this orphan receptor is essential. In Y2 we have shown that Gpr and its signaling domain are required for glandular development more generally and serve a common function as progenitors of the myoepithelium in multiple secretory organs.
  - ❖ Our results show that high levels of Gpr occur in TNBC breast cancer, and that patients with higher levels within these groups have particularly poor outcome. These support the concept that Gpr expression has value as a prognostic indicator of patient outcome in basal type breast cancer. This has pioneered a new field in breast cancer research since there are no studies besides our own on this Gpr. In Y2 and Y3 we addressed the mechanism by showing that Gpr+ cells acquire bipotency in the tumor situation. We have also shown that Gpr identifies a novel cell type that is expanded in tumors with reduced latency and displays mesenchymal undifferentiated features reminiscent of embryonic progenitors found at invasive tips in embryonic glands

- ❖ **What was the impact on other disciplines?**
  - ❖ The homozygous Gpr-DTR:EGFP-creER^{T2} mice, which have a pronounced eye phenotype display several features of "Dry Eye Syndrome" and could have utility as a model for this human inflammatory disease.
  - ❖ Our studies show that the pattern of Gpr expression is consistent with those affected by Mumps virus.
  - ❖ The expression of Gpr in several secretory glandular structures that share common ectodermal origin suggests it may play in ductal branching and secretory differentiation. Its presence in the stem cell compartments of several other ectodermal appendages including the bulge and secondary germ compartments of hair follicles (see Y1 report) suggest it may be an indicator of a more generalized stem cell function.

- **What was the impact on technology transfer?**
  - Nothing to report

- **What was the impact on society beyond science and technology?**
  - Linking breast development factor to breast cancer risk opens the door to preventative strategies linked to reproductive history.

5. CHANGES/PROBLEMS

- **Changes in approach and reasons for change**:
  - None since the last report

- **Actual or anticipated problems or delays and actions or plans to resolve them**
  - **Task 2 Subtask 1: Generate bi-transgenic Gpr-DTR -creER^{T2}/MMTV-Wnt1 mice**
- As reported previously the first experimental cohort of compound Gpr-DTR-creER^{T2}/Gpr-lacZ;MMTV-Wnt1 female mice began to produce tumors in March 2020 coinciding with a mandatory shutdown of all experiments due to the COVID-19 outbreak in New York and had to be euthanized. Dr. Spina, who volunteered for the very first pilot group that returned to work June 2020 set up the breeding again however the loss of this cohort of experimental animals set back our work by more than 8 months. The second cohort began developing tumors in February 2021. We have harvested and frozen organoids and cells from them. We now plan to transplant them into Foxn1 nu mice and expect the transplanted tumors to arise within 1-2 months.

- **Changes that had a significant impact on expenditures –**
  - Mouse costs during COVID lock down. We have incurred significant extra animal husbandry costs due to the loss of the mouse cohort involved in Task 2 as their tumors came during the lockdown period as detailed above. We have had to incur duplicate costs to regenerate these animals and to house a second set of progeny for over 7 months until they generate their tumors.

- **Significant changes in use or care of human subjects**
  - Nothing to report
- **Significant changes in use or care of vertebrate animals**
  - Due to COVID-19 lock down we lost a cohort of animals that developed tumors during this period and have had to repeat the breeding of significant numbers of mice
- **Significant changes in use of biohazard or select agants**
  - Nothing to report

6. PRODUCTS

- **Publications, conference papers and presentations**

- ***Manuscripts:*** see attached manuscripts that are currently under revision
  - **Spina E**, Simundza J, Incassati A, Chandramouli A, Watson CJ, Cowin P. Gpr125 Identifies Mammary Stem and Progenitor Cells and is Associated with Reduced Tumor Latency. Under revision for Nature Communications.
  - **Spina E**, Cowin P. Embryonic Mammary development. Review. Published in Seminar in Cell and Development Biology.
  - **Spina E**, Handlin R, Simundza J, Incassati A, Faiq M, Sainulabdeen A, Chan KC, Cowin P. Gpr125 Plays Critical Roles in Lacrimal Myoepithelia and Tear Film. bioRxiv. doi: https://doi.org/10.1101/2020.09.15.296749
- ***Books etc:*** Nothing to report

- ***Other publications, conference papers and presentations***
  Acknowledgement of Federal Support:

- **Website(s) or other internet site (s):** Nothing to report

- **Technologies or Techniques:** Nothing to report

- **Inventions, patent applications and/or licenses:** Nothing to report

- **Other products:**
  - **Research material:** Generation of Adgra3$^{cre/cre}$ knock out mouse model

7. PARTICIPANTS AND OTHER COLLABORATING ORGANIZATIONS

Individuals working on the project:

Name	Pamela Cowin
Project Role	P.I.
Research Identifier	
Nearest person month worked	12
Contribution to project	Directed research
Funding Support	DOD BC160959 40%

Name	Elena Spina
Project Role	Postdoctoral fellow
Research Identifier	
Nearest person month worked	12
Contribution to project	Performed work on Aim 1
Funding Support	DOD BC160959 100%

Has there been a change in the active support of the PI or senior key personnel since the last reporting period

P.I. Dr. Pamela Cowin – No change
Postdoc Dr. Elena Spina – No change

What other organizations were involved as partners?

Nothing to report

8. SPECIAL REPORTING REQUIREMENTS N/A

9. APPENDICES:

References cited.

1. Tabula Muris, C. *et al.* Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature* **562**, 367-372 (2018).
2. Wuidart, A. *et al.* Early lineage segregation of multipotent embryonic mammary gland progenitors. *Nat Cell Biol* **20**, 666-676 (2018).
3. Pal, B. *et al.* Construction of developmental lineage relationships in the mouse mammary gland by single-cell RNA profiling. *Nat Commun* **8**, 1627 (2017).
4. Bach, K. *et al.* Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing. *Nat Commun* **8**, 2128 (2017).

Manuscripts submitted:

- **Spina E**, Simundza J, Incassati A, Chandramouli A, Watson CJ, Cowin P. Adhesion GPCR Gpr125 is a Unifying Hallmark of Multiple Mammary Stem/Progenitors and Associated with Reduced Tumor Latency. Under revision for Nature Communications.
- **Spina E**, Cowin P. Embryonic Mammary development. Review. Seminar in Cell and Development Biology.
- **Spina E**, Handlin R, Simundza J, Incassati A,Faiq M, Sainulabdeen A, Chan KC, Cowin P. Gpr125 identifies myoepithelial progenitors at tips of lacrimal ducts and is essential for tear film. bioRxiv. doi: https://doi.org/10.1101/2020.09.15.296749

Adhesion GPCR Gpr125 is a Unifying Hallmark of Multiple Mammary Stem/Progenitors, and Associated with Tumor Latency

Departments of Cell Biology[1] and Dermatology[2] New York University School of Medicine. Department of Pathology, University of Cambridge[3].

Elena Spina[1], Julia Simundza[1], Angela Incassati[1], Anupama Chandramouli[1], Christine J Watson[3], Pamela Cowin[1,2]*

*Address for correspondence:

Dr. Pamela Cowin

New York University Medical Center

MSB 621, 550 1st Ave

New York, NY 10016

E-mail: pamela.cowin@nyulangone.org

Conflict of interest statement: The authors declare that no conflict of interest exists.

Abstract

Gpr125/*Adgra3*, is an orphan G-protein coupled receptor with homology to cell adhesion and axonal guidance factors. Loss of Gpr125 exacerbates cell movement defects in zebrafish planar polarity mutants however its role in mammals has remained understudied. Here we show that Gpr125 is a unifying hallmark of multiple mammary progenitors and is localized at sites of collective cell migration during development and at pushing margins of tumors. Gpr125 identifies bipotent mammary stem cells in the embryo and multiple long-lived unipotent basal progenitors in perinatal and postnatal glands. Bipotent Gpr125 progenitors retaining embryonic features (keratin18/14-negative) are greatly expanded in MMTV-Wnt1-early tumors that have reduced latency. Higher expression of Gpr125 within basal breast cancers, where there is a great unmet need for markers that can stratify risk, identifies patients that have a particularly poor outcome. This study highlights the utility of Gpr125 as a highly specific progenitor marker in multiple tissues and underscores the mechanistic parallels between intrinsic developmental properties and cancer.

Introduction

G-protein coupled receptors (GPCRs) characterized by a seven-pass transmembrane (7TM) domain fall into five categories: Glutamate, Rhodopsin, Adhesion, Frizzled and Secretin (GRAFS)[1]. Adhesion GPCRs form the second largest group, yet remain the least understood[2]. Gpr125 contains a domain resembling immunoglobulin-like cell adhesion molecules (Ig-CAM) and has leucine rich repeats (LRR) found in Slit1, an axonal guidance factor, and LRIG, a marker of hair follicle progenitors [3,4]. In keeping with functions suggested by these motifs, we recently discovered Gpr125 identifies myoepithelial progenitors at the migrating tips of embryonic lacrimal ducts, and others have reported its expression in spermatogonial progenitors [5,6].

The ligand for Gpr125 is unknown and it remains unclear whether it signals through G-proteins. Indeed, several studies suggest it acts via non-canonical routes [7-10]. Gpr125 interacts with several PDZ proteins typically involved in cell junctions, polarity, directional movement and morphogenesis. For example, in zebrafish Gpr125 clusters Disheveled (Dsh) into membrane subdomains and modulates planar cell polarity complexes directing convergent extension cell movements and facial motor neuron migration [7]. Gpr125 also binds Discs large (Dlg), a tumor suppressor member of the ZO-1 protein family [8,11]. This interaction implicates Gpr125 in cancer, where high expression has been correlated with good outcome in colon cancer and poor outcome in myeloid leukemia [10,12]. When introduced into cultured cells, Gpr125 is internalized constitutively to endosomes suggesting roles in receptor recycling[9]. Collectively, these studies lead us to hypothesize

that Gpr125 demarcates cells with stem/progenitor potency that participate in cell polarity and adhesive events linked to directed cell migration, Wnt signaling and cancer. To test this concept further we set out to study Gpr125 in the mammary gland and mammary cancers.

Mammary glands provide an ideal system to study developmental processes, such as cell growth, adhesion, migration, and differentiation *in vivo*. Between embryonic days 10-12 (E10-12) ectodermal cells rearrange into placodes that become committed to a mammary fate and sprout towards the mammary fat pad ~E15, then invade and branch to form a small bi-layered tree [13]. Mammary ducts comprise an internal luminal layer of hormone receptor positive and negative cells that express keratin (K) K8 and K18 encapsulated by a basal layer expressing smooth muscle actin (SMA), K14 and K5[14,15]. Development continues during puberty, when the permanent ductal system is extended by hormone-induced proliferation in terminal end buds (TEBs) [14,16-18]. During pregnancy, side-branches emerge followed by two waves of alveolar proliferation and differentiation to produce a functional lactating gland at birth [15,19-21]. Upon weaning, the gland involutes, to remove these temporary, and now redundant, branches and alveoli while retaining the permanent ductal system and regenerating the fat pad [22]. This cycle of development and destruction is repeated with each pregnancy. Thus, three functional types of stem/progenitor cells are required to support the natural life-cycle of the mammary gland: embryonic, pubertal and long-lived adult progenitors [18,23].

The existence of regenerative mammary cells dispersed throughout the mature ductal system was established by seminal experiments demonstrating that fragments from any part of the mammary epithelium could regenerate an entire gland when transplanted into a cleared mammary fat pad [24]. Serial passage of fragments and barcoded mammary cells provided the first evidence of a mammary hierarchy, with fully potent stem cells at the apex giving rise to more restricted ductal and alveolar progenitors [25,26]. These pioneering studies paved the way for analyses of cell subpopulations, defined by their surface marker expression and regenerative properties as mammary repopulating units (MRU) or mammary stem cells (MaSCs) [27,28]. In concert with lineage tracing and single cell RNA sequencing, these analyses produced a working model of the mammary hierarchy [14,23,27-36].

There is general agreement that the hierarchy begins with multipotent ectodermal stem cells, which give rise to bipotent embryonic mammary stem cells and long-lived unipotent luminal- and basal-restricted mammary progenitors. However, the developmental timing of potency restriction remains a source of debate [14,23,30-32,36-39]. A major gap in our knowledge concerns the location of putative stem and progenitor populations. Attempts to address this problem have identified mutually exclusive cell populations at disparate sites [23,40-45]. For example, cells that express Lgr5 are restricted to the nipple zone, whereas cells expressing s-SHIP are confined to ductal tips and branches, and cells expressing Procr and Bcl11b are dispersed along ductal borders [42-45]. Moreover, many of these proteins are expressed in additional mammary populations therein compromising

their specificity as stem/progenitor markers. To date no specific unifying progenitor hallmark has been identified. Here we show that Gpr125 identifies long-lived progenitors at multiple sites and stages of mammary development. Importantly, these cells are engaged in collective cell migration during normal development and their expansion is associated with reduced tumor latency and poor outcome in breast cancer.

Results

Gpr125 is expressed in nipple-proximal regions, terminal end buds and cells dispersed along mature ducts

As nothing is known about Gpr125 in the mammary gland we began by analyzing mRNA levels in whole mammary glands by qPCR over the course of mammary development (**Figure S1A**). Gpr125 mRNA was high during puberty, decreasing as mice reached maturity (12 weeks) and was higher during earlier stages of pregnancy than later. This temporal pattern indicates that Gpr125 mRNA levels are elevated during ductal elongation and branching, and decrease as alveoli expand and differentiate.

Next, we determined where Gpr125 protein is expressed by X-gal staining tissue from *Adgra3*$^{lz/+}$ mice in which β-galactosidase is fused to the first transmembrane region of Gpr125 (**Figure 1A, S1B, D**). *Adgra3*$^{lz/+}$ mice are viable, fertile and indistinguishable from wildtypes (**Figure S1E**). *Adgra3*$^{lz/lz}$ mice show an identical expression pattern but have delayed ductal elongation that rectifies by the end of puberty (**Figure S1F**). Gpr125-β-gal was expressed throughout the dormant pre-pubescent mammary tree (**Figure 1B, C**). As pubertal ductal elongation began, the X-gal staining pattern partitioned. Weak staining was retained in nipple-proximal ducts (**Figure 1D-F**) and persisted there throughout postnatal development. In contrast, robust Gpr125-β-gal expression appeared in proliferative TEBs (**Figure 1E-G**), and became reduced to an intense dot at ductal tips as the TEB reached the edges of the fat pad and regressed (arrows **Figure 1E**). In histological sections, Gpr125-β-gal was expressed in cap cells (**Figure 1G**) and also in

single cells dispersed among the basal layer of mature ducts (**Figure 1G**, inset). Gpr125-β-gal colocalized with SMA, p63 and low K14 (**Figure 1H-J**) but was absent from cells expressing E-cadherin (Ecad), estrogen receptor (ER) and progesterone receptor (PR) (**Figure 1K-M**). Co-expression of Gpr125-β-gal with PCNA and exclusion of p27 indicated their proliferative potential (**Figure 1N, O**).

Pubertal Gpr125+ cells display a marker profile of regenerative cell types

To characterize Gpr125+ cells further we mined the Tabula Muris scRNAseq dataset [46]. Gpr125 mRNA was detected within a subset of the "basal" epithelial cell cluster and some stromal cells (**Figure 2A**). We confirmed Gpr125 protein expression within a subset of basal cells by flow cytometry using a fluorogenic β-galactosidase (β-gal) substrate: Fluorescein di-β-D-galactopyranoside (FDG). Gating for Gpr125^{+}/FDG^{+} cells in suspensions of total mammary epithelial cells (MECs) from 6-week old pubertal *Adgra3*$^{lz/+}$ mice produced a clear enrichment for cells within the basal (CD24$^{med/low}$CD49f$^{+/Hi}$) population; FDG^{-} cells were concomitantly depleted within this gate (**Figure 2B**). Importantly, within the basal population, Gpr125^{+}/FDG^{+} cells displayed the highest level of integrins α6 and β1 (CD49f and CD29) (**Figure 2C**), which are hallmarks of regenerative MRU/MaSCs [27,28]. Conversely, Gpr125^{+}/FDG^{+} cells were low for the luminal progenitor marker CD61 (integrin β3) and negative for Sca-1 (**Figure 2C**) which is expressed on more committed cell types [27,47,48]. Gpr125-β-gal and s-SHIP-EGFP were both expressed in all cap cells surrounding the TEB (**Figure 2D**). By scRNAseq, cells expressing Gpr125 mRNA co-clustered with Lgr5 populations in datasets of embryonic,

pubertal and adult mammary epithelial cells (**Figures 2A, S2**)[34,35,37]. Collectively, these data show that Gpr125+ cells are found at multiple sites of predicted stem/progenitor activity, display a surface profile consistent with cell types high within the putative mammary hierarchy, encompass disparate cell populations with documented regenerative capacity and lack markers associated with lineage commitment.

Pubertal Gpr125+ cells are long-lived unipotent basal progenitors

To position pubertal Gpr125+ cells within the mammary hierarchy, we carried out lineage tracing to determine their physiological potency. We generated a mouse strain harboring a *creER*^{T2} module inserted after the endogenous *Adgra3* promoter (**Figures 3A and S1B,C**), crossed them to the Tomato (tdT) lineage reporter strain (B6.Cg-Gt(ROSA)26Sor$^{tm14(CAG-tdTomato)Hze/J}$) and initiated tracing in 5-week-old female progeny by delivering Tamoxifen (Tam) via IP (**Figure 3B**). 48 hours later few cells expressed tdT (data not shown). After two weeks we observed clusters of tdT+ cells within the basal layer of ducts (**Figure 3**) that were most abundant in the nipple-proximal region (**Figure 3C**). tdT+ cells displayed the characteristic bipolar shape of myoepithelial cells and by immunofluorescence, co-localized with the basal markers K5, K14, SMA and p63 (**Figure 3D-G**) and were devoid of the luminal markers Ecad and K8 (**Figure 3H,I**). tdT was also found in cap cells of TEB (**Figure 3J-L**) and in strips of cells extending down the basal surface of subtending ducts.

Next, we mated mice in which tracing had been initiated during puberty and analyzed their glands during pregnancy. Again, tdT+ cells were basally restricted. At p15.5. clusters of tdT+ cells were observed colocalizing with a subset of K5-expressing cells (**Figure 3M**) and surrounding immature Ecad+ alveoli (**Figure 3N**). At lactation day 6 (L6), tdT+ cells displayed the typical basket-like features of contractile myoepithelial cells, enmeshing fully differentiated Ecad+ alveoli (**Figure 3O, S3 movie**). We addressed the longevity of the Gpr125^{+} progenitors labeled during puberty by tracing their tdT+ progeny in both aged nulliparous mice and multiparous mice. In both, clusters and extensive strips of elongated tdT+ cells were found along basal ductal borders (**Figure 3P-R**). Collectively, these data show Gpr125 identifies long-lived unipotent basal progenitors.

Gpr125+ cells are located on the leading tips of side branches during pregnancy

During early pregnancy (p12-13.5) Gpr125-β-gal appeared focally where side-branches emerge from ducts (**Figure 4A-E**) and became concentrated at branch tips (**Figure 4F**). It persisted along the basal layer of permanent ducts during late pregnancy but was absent from differentiating alveoli (**Figure 4G-I**) with the exception of rare Gpr125-β-gal+ cells that likely represent the branch tip of each alveolar cluster (**Figure 4H arrowheads**).

To interrogate the potency of cells expressing Gpr125+ cells during pregnancy we initiated tracing in pregnant mice at p13.5 and analyzed their glands just prior to birth and during lactation (**Figure 4I**). Again, TdT+ cells exclusively displayed myoepithelial characteristics (**Figure 4J-N**).

Gpr125 identifies bipotent and unipotent progenitors during early embryogenesis

Extending our Gpr125 studies to embryonic mammary development, we found that early *Adgra3*$^{lz/+}$ embryos showed diffuse ectodermal Gpr125-β-gal expression that by E14.5 concentrated into ectodermal appendages, such as whisker follicles but was absent from the mammary line, placodes and buds (**arrows Figure 5A, A'**). However, at E15 Gpr125-β-gal appeared in the mammary sprout (**arrows Figure 5B, B'**), coincident with the onset of proliferation, indicated by nuclear PCNA and BrdU (**Figure 5C,D**), and was concentrated towards the leading tip together with K14 and p63 (**Figure 5E, F**). At this stage Gpr125-β-gal also became concentrated in the "bulge" stem cell compartment of hair follicles and whiskers (**Figure S4**). At E18.5 Gpr125-β-gal was strongly expressed in the rudimentary tree (**Figure 5G**) and concentrated in the lactiferous duct and the multilayered branch tips (**Figure 5H**). Although K14 expression was present in most cells (**Figure 5I**) and K18 was increased in central cells (**Figure 5J**), Gpr125+ branch tips lacked both keratins (**Figure 5I,J arrow and insets**) but expressed p63 (**Figure 5K**).

To test the potency of embryonic Gpr125+ cells we crossed *Adgra3*$^{cre/cre}$ to tdT mice and administered Tam to the pregnant dams to label embryos at E14.5 and at E19 (**Figure 5L**) then analyzed these pups at 8 weeks of age. Glands from mice labeled at E14.5 revealed tdT not only in basally-located bipolar cells co-expressing SMA (**Figure 5M**), but also in columnar cells situated above the SMA+ basal layer (**Figure 5N**) that expressed the luminal marker Ecad (**Figure 5O**). By contrast, when tracing was initiated at E19, tdT+ cells were exclusively basal, bipolar and SMA+ (**Figure 5P**).

These data show that Gpr125 appears at the onset of directed growth in a bipotent progenitor population expressing markers of both lineages (K14/K18). However, before birth Gpr125 cells become lineage-restricted and from thereon function as unipotent basal progenitors. Of note, Gpr125 expression concentrates at this stage in undifferentiated $p63^{+}K14-/K18^{-}$ cells at branch tips poised for ductal extension.

Gpr125+ cells retaining embryonic undifferentiated characteristics and bipotency are expanded in MMTV-Wnt1 glands

Next we investigated Gpr125 in mouse breast cancer models. Gpr125 showed the highest mRNA expression in MMTV-Wnt1 mice (**Table S5A**) which develop mixed lineage tumors enriched in cells with MaSCs profiles and show transcriptomic resemblance to basal-type breast cancer [49-54]. We therefore generated *MMTV-Wnt1*; *Adgra3*$^{lz/+}$ mice and analyzed glands over the course of tumor progression. In 8-week old mice, Gpr125-β-gal was expressed exclusively in basally-located cells within the nipple proximal hyperplasia and on hyperbranched tips (**Figure 6A-D**). Robust expression was seen in tumors, where Gpr+ cells formed large homogenous regions and concentrated at pushing margins (**Figure E-H**). Although Wnt-1 tumors display regions composed of K14+ and K18+ bilayers these populations rarely overlapped with Gpr125. Gpr125+ cells also lacked both SMA and K8 and although some expressed K14, the majority lacked both keratins (**Figure 6G-I)**. They did, however, express p63 as well as Tcf1 (**Figure 6J-K**). In normal gland the latter is found exclusively in a subset of embryonic cells and TEB cap cells

(**Figure S2 6L**). To determine the effect of Wnt expression on the potency of Gpr125 expressing cells, we generated *MMTV-Wnt1;Adgra3$^{cre/+}$;tdT* mice and performed lineage tracing by delivering Tam to 5-week old pubertal mice and harvesting hyperplastic glands from 12-week old mice (**Figure 6M**). tdT was found in both SMA+ and Ecad+ cells (**Figure 6N, O**). These data indicate that in the context of MMTV-Wnt1 transformation pubertal Gpr125+ cells retain the undifferentiated characteristics and bipotency of embryonic progenitors.

High Gpr125 expression is indicative of early tumor onset in mice and poor outcome in human basal-type breast cancer

We noted that Gpr125-β-gal expression was consistently more extensive in tumors arising with shorter latency (**Figure 7A**). MMTV-Wnt1 tumors have been separated into two subtypes with distinct gene expression (ex) profiles: Wnt1-early(ex) and Wnt1-late(ex), that correlate with early (average 6.5 weeks) and late (average 22.5 weeks) tumor onset respectively [49-51]. Mining these microarray data, we found Wnt1-early(ex) tumors have twice the level of Gpr125 mRNA as Wnt1-late(ex) tumors (**Table S5A**) and confirmed this by qPCR analysis ($p=0.0052$ $n=6$) (**Figure 7B**). To investigate the course of Gpr125+ cell expansion we carried out flow cytometry. In hyperplastic glands, the Gpr125$^+$/FDG$^+$ population localized within the traditional basal gate but in tumors it was expressed in a new population with intermediate CD49 levels (**Figure 7C**) that was more pronounced in the uninvolved glands and tumors that arose early. Collectively, these data show that expansion of the Gpr125 tumor population correlates with early tumor onset in mice. We

investigated GPR125 expression in human breast cancer and found it was highest within the basal-type (**Table S5B**). Interrogating publicly available datasets using the km plotter algorithm[55] showed high GPR125 levels within this subtype correlated with worse patient outcome in terms of relapse-free survival p=0.0054 and distant metastasis free survival (p=0.0043) (**Figure 7D,E red line**), and with disease-free survival using the breastmark algorithm[56] (**Figure 7F, blue line**).

Discussion

Here, we present the first report of Gpr125 expression and function in the developing mammary gland and breast cancer. Our results demonstrate the powerful ability of Gpr125 to specifically localize progenitors at multiple sites and stages of mammary development **(Figure 8).** We show that Gpr125 cells are concentrated at the tips of migrating structures during normal development and massed at pushing margins in tumors, where their expansion correlates with early tumor onset in mice and poor outcome in humans.

Gpr125 marks bipotent embryonic and unipotent perinatal mammary progenitors.

Our analyses show that in the embryo Gpr125 is expressed in stem cell compartments of other ectodermal appendages such as hair follicles and whiskers (**Figure S4**). This, together with its expression in lacrimal and spermatogonial progenitors, indicates that Gpr125 has important value in demarcating early progenitors in multiple tissues [6]. In the embryonic mammary gland, we found that E15 Gpr125+ cells are bipotent. Whether they are mammary stem cells or comprise a mixture of pre-committed unipotent basal and luminal progenitors will further require clonal analysis. In support of the former scenario, scRNAseq and molecular profiling studies could find no evidence of lineage bias among early embryonic progenitors [36,37,57]. However, by E19, Gpr125-expressing cells clearly become committed unipotent basal progenitors consistent with recent studies by Fre et al. that have indicated that lineage-restriction begins early and is completed before birth [32,38].

Gpr125 Identifies unipotent basal progenitors in multiple locations in postnatal glands.

During postnatal mammary development Gpr125 identifies long-lived unipotent basal progenitors at multiple locations. Gpr125 expression in the nipple proximal zone resembles that of Lgr5 [43,58,59]. Gpr125 and s-SHIP are both present in all cap cells of the TEB [42,60]. Gpr125 expression in progenitors dispersed throughout mature ductal system is consistent with findings from the earliest mammary transplantation studies and similar to that of Procr and Bcl11b [24-26] [44,45]. While the precise relationship of Gpr125+ cells to these diverse cell types requires further delineation, Gpr125 coexpression in Lgr5+ and s-SHIP+ cells shows it to be a common link between disparate progenitors. Thus, whereas mammary progenitor heterogeneity likely reflects the need to respond to niche and developmental stage-specific signals, we propose Gpr125 must serve a more universal role in progenitor biology. It is possible that the presence of Gpr125 in these progenitors relates to its involvement in Wnt signaling. Wnt signaling is critical at all stages of mammary development and has been shown to sustain progenitor potency in vitro [18,33,52,53,61,62]. Intriguingly, the closely related protein, Gpr124, selectively promotes canonical signaling by specific Wnt ligands[63]. In contrast, Gpr125, has been implicated in non-canonical Wnt signaling and detected in biochemical complexes with Ryk[7,12] [64].

Gpr125 marks sites of collective cell migration

It is striking that in several tissues Gpr125 is concentrated in a graded fashion at sites of directed migration [5]. These include the tips of the mammary sprout, rudimentary tree, pubertal TEB and side-branches. Its expression in these locations suggest an involvement of Gpr125 in directing progenitor migration. This concept is supported by previous reports showing Gpr125 levels influence facial motor neuronal precursor migration in zebrafish[7]. Indeed, regulation of directed cell movement may be a common function of the ADGRA family as Gpr124 is required for tip cell function in endothelia during angiogenesis [63].

High Gpr125 expression is indicative of early tumor onset in mice and poor outcome in human basal-type breast cancer.

A role in cell migration has considerable clinical significance in the setting of breast cancer, a disease where metastasis is the primary cause of death. Preliminary in silico analyses suggest Gpr125 could have prognostic utility in basal type breast cancer, where there is a great unmet need for markers that can stratify risk. In mice, Gpr125 mRNA is elevated and Gpr125+ cells, lacking lineage-specific keratins are greatly expanded in MMTV-Wnt1 tumors arising with short latency [49]. This Gpr125+ tumor cell-type shares features with Gpr125 progenitors found the multilayered tips of the embryonic mammary tree and pubertal TEB suggesting its pathological contribution to reduced latency may relate to its acquisition of intrinsic migratory and invasive properties of embryonic and pubertal progenitors.

Methods

Ethics statement. All animal protocols were approved by the Institutional Animal Care and Use Committee (IACUC) at NYU School of Medicine.

Mice. Mice were constructed by Ingenious Technologies, Ronkonkoma, NY as follows. A cassette containing *CreER^{T2}* followed by a 3' polyadenylation signal, harboring SV40-driven Neo flanked by FRT sites inserted in a central intron, was recombined into a bacterial artificial chromosome (BAC) to place *CreER^{T2}* under the control of the *Adgra3* promoter, excising 502 bp encompassing 221 bp of exon 1 and part of the following intron 1-2 of *Adgra3*. Mice generated from these ES cells were selected for germline transmission by PCR, verified by southern analysis and sequencing then bred to a Flp deleter strain to remove Neo. *Adgra3$^{lz/+}$* mice were generated by Regeneron using VelociGene methods [65] to modify a bacterial artificial chromosome (BAC) clone carrying the mouse *Adgra3* gene by replacement of sequence encompassing exons 16-19 with *lacZ* to produce expression of fusion protein comprising the N-terminal extracellular domain, the first transmembrane domain, and part of the first intracellular loop of Gpr125 fused to β-galactosidase (Figure 1A) [6].

Lineage Tracing: For lineage tracing experiments, *Adgra3-CreER^{T2}* mice were crossed to the fluorescent Rosa26R-lox.STOP.lox-tdTomato (tdT) lineage reporter strain (Stock No. 007909) Jackson laboratory. The transcriptional STOP was deleted by cre recombination by delivering tamoxifen. For tracing at mid-puberty tamoxifen was delivered ip at low dose: 1.5mg, and high dose: 5-15mg (delivered in 2.5mg aliquots every

other day). For tracing during pregnancy and embryogenesis 2 doses of 2.5mg Tam was given by oral gavage to *Adgar3*$^{cre/cre}$ pregnant dams over a 24hr time period. Pups were delivered at E19.5-E20.5 by caesarian section to avoid Tam-induced problems with delivery and fostered by SWR/J mice. Tissue was harvested at the indicated intervals over course of mammary gland development. For each experiment, two mammary glands of at least three mice were analysed. No fluorescence was observed in non-induced mice.

Tissue clearing and 3-D imaging: Mammary glands were excised and fixed overnight in 4% PFA then processed using a CUBIC protocol optimized for mammary gland [32,66]. Tissue was incubated in CUBIC Reagent 1A(10wt% Triton,5wt% N,N,N',N'-tetrakis (2-HP)ethylenediamine, 10wt% Urea, NaCl 25mM) clearing solution for 4 days, rinsed 3X in PBS, then incubated at 4C for 4 days in primary antibodies diluted in PBST containing 10% serum, rinsed again, then incubated in secondary antibody for 2 days, rinsed 3X, then cleared in CUBIC Reagent 2 (50w/v% Sucrose, 25w/v% Urea, 10w/v% Triethanolamine, 0.1w/v% Triton) for 24hrs. Primary antibodies: rabbit anti-K5 (Covance, PRB160P, 1:100); rat anti-K8 (Developmental Studies Hybridoma Bank, TROMA-I, 1:50); mouse anti-SMA (Dako, M0851, 1:100); rabbit anti-E-cadherin (Cell Signaling, 3195S, 1:100); rabbit anti-p63 (Abcam, ab124762,1:100); rabbit anti K14 (Abcam,Ab181595). Alexa Fluor-conjugated secondary antibodies Thermo Fisher Scientific, diluted 1:500: goat anti-mouse 647 (A21237); goat anti-rat 647 (A21247); goat anti-rabbit 647 (A21245). Cleared mammary tissues were imaged using a Zeiss 880 Laser Scanning inverted confocal microscope with 10X, 20X air Plan-Apochromat N.A. 0.8 M27 objective lenses.

X-gal staining: Embryos and mammary glands were fixed in 4% paraformaldehyde (PFA) (Sigma-Aldrich) at room temperature (RT) for 30-60 min, rinsed 3X in X-gal rinse buffer (2 mM MgCl2, 0.1% Sodium deoxycholate, and 0.2% NP-40 in PBS) at RT, then incubated in X-gal staining solution (50 mg/ml 5-bromo-4-chloro-3-indolyl-β Dgalactopyranoside in rinse buffer containing 5 mM potassium ferricyanide, 5 mM potassium ferrocyanide) (Applichem, Cheshire, CT) at RT overnight. After staining, glands were rinsed in PBS, post-fixed in 4% PFA overnight then prepared for whole mount analysis or processed for paraffin embedding, sectioning and histological analysis [67,68].

Mammary gland whole mounts. X-gal stained whole mounts were post-fixed in 4% PFA, washed twice with 1X PBS, dehydrated through an increasing ethanol gradient, cleared of lipids in Carnoy's Fixative (60% Ethanol, 30% Chloroform, 10% Glacial Acetic Acid) for two hours, and further cleared in Citrisolv (Fisher Scientific, Suwanne, GA) for two hours. Glands were pressed flat between the slide and coverslip under a heavy weight for 30 minutes, and imaged on a Leica dissecting microscope Model WILD M3Z (Leica Microsystems, Bannockburn, IL) with an Optronics digital camera Model 60800 (Goleta, CA). The glands were then re-hydrated through a decreasing ethanol gradient and counterstained with Carmine alum (500mL distilled water containing 1g Carmine and 2.5g aluminum potassium sulfate; Sigma Aldrich, St Louis, MO) diluted 1:4 in distilled water. Glands were once again dehydrated in ethanol, cleared in Carnoy's Fixative and Citrisolv,

and pressed flat before mounting under a coverslip with Cytoseal (VWR, West Chester PA) then re-photographed [69].

Immunohistochemistry: Mammary glands were removed from mice and fixed with either 10% neutral buffered formalin or 4% paraformaldehyde (PFA) and embedded in paraffin. Slides containing paraffin-embedded tissue sections were cleared of paraffin by incubation in a 60°C oven for one hour, and submerged in Citrisolv for 10 minutes. Sections were rehydrated through an ethanol gradient and rinsed in distilled water for 10 minutes. Antigen retrieval was performed by microwaving at 900 watts for 30 minutes in 10mM Citric Acid buffered to pH 6. From this point forward, the slides were washed thrice with 1X PBS between each step. Endogenous peroxidase activity was quenched by treating slides with 3% Hydrogen Peroxide (Sigma Aldrich) for 15 minutes at room temperature. Slides were blocked with 20% Normal Goat Serum for 30 minutes to reduce background signal. Primary antibodies were diluted in 2% bovine serum albumin (BSA, Sigma-Aldrich) in 1X PBS, and incubated at 4°C overnight. Primary antibodies: rabbit anti-K14 (Covance PRB-155P 1:500); mouse anti-K8 undiluted (Progen 65138); mouse anti-p63 (Neomarkers MS-1081-P1 1:300); mouse anti E-cadherin (BD 610182 1:100); mouse anti-PCNA (Dako M0987 1:500); rabbit anti TCF1 (Cell Signaling 22035 1:100);. Biotinlyated secondary antibodies were diluted in 2% BSA/PBS for 1 hour at room temperature, followed by HRP-conjugated Streptavidin (Vector Labs, Burlingame CA) for 30 minutes at room temperature. Colorimetric signal was developed using the DAB substrate (Vector Labs).

Flow cytometry. To detect Gpr125- β-gal expression, cells were labelled with fluorescein di-V-galactoside (FDG) according to manufacturer's protocol (Molecular Probes, Eugene, Oregon). Briefly, cells were resuspended at 107/mL in HF, and the samples were pre-warmed at 37°C for 10 minutes. FDG loading was performed by adding an equal volume of pre-warmed 2mM FDG (diluted in distilled water) to the cell suspension for exactly 1 minute at 37°C, then immediately quenched by adding 2mL ice-cold HF. The FDG-loaded cells were then centrifuged and stained with surface antibodies. The following antibodies were used to label cells for flow cytometry: biotinylated- TER119 (BD 553672,1:200), biotinylated-CD31 (BD 558737,1:200), biotinylated-CD45 (BD 553077,1:200), biotinylated-CD140a (eBioscience 12-1401-80,1:200), CD24-PE (BD 553262,1:400), CD49f-PerCP-Cy5.5 (Biolegend 313617,1:200,), CD24-FITC (BD553261,1:100), CD49f-PE (1:100), Streptavidin-AlexaFluor647 (Molecular Probes S21374,1:600), CD61-APC (Caltag,1:200), Sca1-PE-Cy7 (eBioscience 25-5981-81,1:600), CD29-Pacific Blue (Biolegend 102224,1:200). Cells were incubated with conjugated antibodies diluted in HF, for 30 minutes on ice in a dark container, washed with 2mL of HF, and resuspended in 250μL HF for analysis. Cell viability was assessed by adding 4',6-Diamidino-2-phenylindole (DAPI, Sigma-Aldrich) to the final suspension at a concentration of 1μg/mL. Data collection for flow cytometry was done on a Beckton-Dickinson (East Rutherford, NJ) LSRII analyser. Analyses were done using FlowJo software version 9.

Reverse transcription and quantitative PCR. Total mRNA was generated from snap-frozen tissues (30-50mg) with a Qiagen RNA-Easy Mini column (Qiagen, Venio, Limburg). Tissues were homogenized in 1mL Trizol (Life Technologies) using a Polytron PT 1200C homogenizer (Kinematica, Bohemia, NY), mixed with 200µL Chloroform, and centrifuged at 12500rpm for 10 minutes in a 4°C minicentrifuge. The aqueous layer was decanted, mixed with an equal volume of 70% Ethanol, and added to the Qiagen spin column. The remaining preparation was performed according to manufacturer's protocol, and RNA was eluted in a final volume of 30-50µL RNase-free distilled water. RNA concentration and quality were assessed with a Nanodrop 2000 (Thermo Scientific, Waltham, MA) and stored at -80°C. cDNA was made from 100µg of RNA with a Verso cDNA Synthesis Kit 2000 (Thermo Scientific) and stored at -20°C. Quantitative PCR reactions were performed with three replicates of 10ng cDNA on a Bio-rad CFX96 Detection System (Bio-rad, Hercules, CA) using SYBR Green Real-Time PCR Master Mix (Life Technologies). Data was analyzed by the ΔΔCt method.

scRNA-seq analysis: scRNA-seq analysis were generated using available data from Tabula Muris (https://tabula-muris.ds.czbiohub.org/). We also conducted an analysis with two other single cell RNA-seq mammary gland datasets [34,35]. We processed the dataset using iCellR, Single (i) Cell R package, an interactive R package to work with high-throughput single cell sequencing technologies with the help of NYU Langone's Applied BioinformaticsLaboratories (https://www.biorxiv.org/content/10.1101/2020.03.31.019109v1).

Microarray analysis. Gpr125 mRNA expression in murine and human samples of breast cancer was obtained by using microarray data available at the Gene Expression Omnibus under the series GSE3165 (https://www.ncbi.nlm.nih.gov/geo/). Analysis was conducted using GEO2R. Version info: R 3.2.3, Biobase 2.30.0, GEOquery 2.40.0, limma 3.26.8. Analysis of Gpr125 in human breast cancer was carried out using kmplotter[55] https://kmplot.com/analysis/index.php?p=service&cancer=breast with criteria: Gpr125 affy ID 210473_s_at, Auto select best cutoff, excluding biased arrays and selecting for basal-type breast cancer RFS: N=618; DMSF: N=232 and using BreastMark[56]: http://glados.ucd.ie/BreastMark/mRNA_custom.html DFS, median cut off was selected for ssp2003 basal-type (N=318) [70] or ssp2006 basal-type (N=366) [71] datasets .

Statistics and reproducibility. Statistical significance was determined using GraphPad Prims software. Normal distribution of data was assessed using Shapiro-Wilk normality tests. Unpaired Student's t-test was performed. Data are always expressed as mean ± SEM. Each experiment was repeated independently at least 3 times. P values and N of repeat are indicated in the figure legends.

1 Fredriksson, R., Lagerstrom, M. C., Lundin, L. G. & Schioth, H. B. The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints. *Mol Pharmacol* **63**, 1256-1272, doi:10.1124/mol.63.6.1256 (2003).

2 Simundza, J. & Cowin, P. Adhesion G-protein-coupled receptors: elusive hybrids come of age. *Cell Commun Adhes* **20**, 213-226, doi:10.3109/15419061.2013.855727 (2013).

3 Fredriksson, R., Gloriam, D. E., Hoglund, P. J., Lagerstrom, M. C. & Schioth, H. B. There exist at least 30 human G-protein-coupled receptors with long Ser/Thr-rich N-termini. *Biochem Biophys Res Commun* **301**, 725-734 (2003).

4 Pickering, C. *et al.* The Adhesion GPCR GPR125 is specifically expressed in the choroid plexus and is upregulated following brain injury. *BMC Neurosci* **9**, 97, doi:10.1186/1471-2202-9-97 (2008).

5 Spina, E. *et al.* Gpr125 identifies myoepithelial progenitors at tips of lacrimal ducts and is essential for tear film. *bioRxiv 2020.09.15.296749*, doi:10.1101/2020.09.15.296749 (2020).

6 Seandel, M. *et al.* Generation of functional multipotent adult stem cells from GPR125+ germline progenitors. *Nature* **449**, 346-350, doi:10.1038/nature06129 (2007).

7 Li, X. *et al.* Gpr125 modulates Dishevelled distribution and planar cell polarity signaling. *Development* **140**, 3028-3039, doi:10.1242/dev.094839 (2013).

8 Yamamoto, Y. *et al.* Direct binding of the human homologue of the Drosophila disc large tumor suppressor gene to seven-pass transmembrane proteins, tumor endothelial marker 5 (TEM5), and a novel TEM5-like protein. *Oncogene* **23**, 3889-3897 (2004).

9 Spiess, K. *et al.* Arrestin-independent constitutive endocytosis of GPR125/ADGRA3. *Ann N Y Acad Sci* **1456**, 186-199, doi:10.1111/nyas.14263 (2019).

10 Wu, Y. *et al.* Elevated G-Protein Receptor 125 (GPR125) Expression Predicts Good Outcomes in Colorectal Cancer and Inhibits Wnt/beta-Catenin Signaling Pathway. *Med Sci Monit* **24**, 6608-6616, doi:10.12659/MSM.910105 (2018).

11 Woods, D. F. & Bryant, P. J. The discs-large tumor suppressor gene of Drosophila encodes a guanylate kinase homolog localized at septate junctions. *Cell* **66**, 451-464 (1991).

12 Fu, J. F. *et al.* Involvement of Gpr125 in the myeloid sarcoma formation induced by cooperating MLL/AF10(OM-LZ) and oncogenic KRAS in a mouse bone marrow transplantation model. *Int J Cancer* **133**, 1792-1802, doi:10.1002/ijc.28195 (2013).

13 Cowin, P. & Wysolmerski, J. Molecular mechanisms guiding embryonic mammary gland development. *Cold Spring Harb Perspect Biol* **2**, a003251, doi:10.1101/cshperspect.a003251 (2010).

14 Davis, F. M. *et al.* Single-cell lineage tracing in the mammary gland reveals stochastic clonal dispersion of stem/progenitor cell progeny. *Nat Commun* **7**, 13053, doi:10.1038/ncomms13053 (2016).

15 Watson, C. J. & Khaled, W. T. Mammary development in the embryo and adult: a journey of morphogenesis and commitment. *Development* **135**, 995-1003, doi:10.1242/dev.005439 (2008).

16 Rosen, J. M. Hormone receptor patterning plays a critical role in normal lobuloalveolar development and breast cancer progression. *Breast Dis* **18**, 3-9 (2003).
17 Daniel, C. W. & Silberstein, G. B. in *The Mammary Gland* (eds M.C. Neville & C.W. Daniel) 3-31 (Plenum, 1987).
18 Sreekumar, A. *et al.* WNT-Mediated Regulation of FOXO1 Constitutes a Critical Axis Maintaining Pubertal Mammary Stem Cell Homeostasis. *Dev Cell* **43**, 436-448 e436, doi:10.1016/j.devcel.2017.10.007 (2017).
19 Fata, J. E. *et al.* The osteoclast differentiation factor osteoprotegerin-ligand is essential for mammary gland development. *Cell* **103**, 41-50 (2000).
20 Brisken, C. *et al.* Essential function of Wnt-4 in mammary gland development downstream of progesterone signaling. *Genes and Development* **14**, 650-654 (2000).
21 Beleut, M. *et al.* Two distinct mechanisms underlie progesterone-induced proliferation in the mammary gland. *Proc Natl Acad Sci U S A* **107**, 2989-2994, doi:10.1073/pnas.0915148107 (2010).
22 Lloyd-Lewis, B. *et al.* Analysis of the Involuting Mouse Mammary Gland: An In Vivo Model for Cell Death. *Methods Mol Biol* **1501**, 165-186, doi:10.1007/978-1-4939-6475-8_7 (2017).
23 Lloyd-Lewis, B., Harris, O. B., Watson, C. J. & Davis, F. M. Mammary Stem Cells: Premise, Properties, and Perspectives. *Trends Cell Biol* **27**, 556-567, doi:10.1016/j.tcb.2017.04.001 (2017).
24 Smith, G. H. & Medina, D. A morphologically distinct candidate for an epithelial stem cell in mouse mammary gland. *J Cell Sci* **90 (Pt 1)**, 173-183 (1988).
25 Kordon, E. C. & Smith, G. H. An entire functional mammary gland may comprise the progeny from a single cell. *Development* **125**, 1921-1930 (1998).
26 Smith, G. H. Experimental mammary epithelial morphogenesis in an in vivo model: evidence for distinct cellular progenitors of the ductal and lobular phenotype. *Breast Cancer Res Treat* **39**, 21-31 (1996).
27 Shackleton, M. *et al.* Generation of a functional mammary gland from a single stem cell. *Nature* **439**, 84-88 (2006).
28 Stingl, J. *et al.* Purification and unique properties of mammary epithelial stem cells. *Nature* **439**, 993-997 (2006).
29 Van Keymeulen, A. *et al.* Lineage-Restricted Mammary Stem Cells Sustain the Development, Homeostasis, and Regeneration of the Estrogen Receptor Positive Lineage. *Cell Rep* **20**, 1525-1532, doi:10.1016/j.celrep.2017.07.066 (2017).
30 Van Keymeulen, A. *et al.* Distinct stem cells contribute to mammary gland development and maintenance. *Nature* **479**, 189-193, doi:10.1038/nature10573 (2011).
31 Rios, A. C., Fu, N. Y., Lindeman, G. J. & Visvader, J. E. In situ identification of bipotent stem cells in the mammary gland. *Nature* **506**, 322-327, doi:10.1038/nature12948 (2014).
32 Lloyd-Lewis, B., Davis, F. M., Harris, O. B., Hitchcock, J. R. & Watson, C. J. Neutral lineage tracing of proliferative embryonic and adult mammary stem/progenitor cells. *Development* **145**, doi:10.1242/dev.164079 (2018).
33 van Amerongen, R., Bowman, A. N. & Nusse, R. Developmental stage and time dictate the fate of Wnt/beta-catenin-responsive stem cells in the mammary gland. *Cell Stem Cell* **11**, 387-400, doi:10.1016/j.stem.2012.05.023 (2012).

34 Bach, K. *et al.* Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing. *Nat Commun* **8**, 2128, doi:10.1038/s41467-017-02001-5 (2017).
35 Pal, B. *et al.* Construction of developmental lineage relationships in the mouse mammary gland by single-cell RNA profiling. *Nat Commun* **8**, 1627, doi:10.1038/s41467-017-01560-x (2017).
36 Giraddi, R. R. *et al.* Single-Cell Transcriptomes Distinguish Stem Cell State Changes and Lineage Specification Programs in Early Mammary Gland Development. *Cell Rep* **24**, 1653-1666 e1657, doi:10.1016/j.celrep.2018.07.025 (2018).
37 Wuidart, A. *et al.* Early lineage segregation of multipotent embryonic mammary gland progenitors. *Nat Cell Biol* **20**, 666-676, doi:10.1038/s41556-018-0095-2 (2018).
38 Lilja, A. M. *et al.* Clonal analysis of Notch1-expressing cells reveals the existence of unipotent stem cells that retain long-term plasticity in the embryonic mammary gland. *Nat Cell Biol* **20**, 677-687, doi:10.1038/s41556-018-0108-1 (2018).
39 Ying, Z. & Beronja, S. Embryonic Barcoding of Equipotent Mammary Progenitors Functionally Identifies Breast Cancer Drivers. *Cell Stem Cell* **26**, 403-419 e404, doi:10.1016/j.stem.2020.01.009 (2020).
40 Fu, N. Y. *et al.* Identification of quiescent and spatially restricted mammary stem cells that are hormone responsive. *Nat Cell Biol* **19**, 164-176, doi:10.1038/ncb3471 (2017).
41 Huo, Y. & Macara, I. G. The Par3-like polarity protein Par3L is essential for mammary stem cell maintenance. *Nat Cell Biol* **16**, 529-537, doi:10.1038/ncb2969 (2014).
42 Bai, L. & Rohrschneider, L. R. s-SHIP promoter expression marks activated stem cells in developing mouse mammary tissue. *Genes Dev* **24**, 1882-1892, doi:10.1101/gad.1932810 (2010).
43 Plaks, V. *et al.* Lgr5-expressing cells are sufficient and necessary for postnatal mammary gland organogenesis. *Cell Rep* **3**, 70-78, doi:10.1016/j.celrep.2012.12.017 (2013).
44 Wang, D. *et al.* Identification of multipotent mammary stem cells by protein C receptor expression. *Nature* **517**, 81-84, doi:10.1038/nature13851 (2015).
45 Cai, S. *et al.* A Quiescent Bcl11b High Stem Cell Population Is Required for Maintenance of the Mammary Gland. *Cell Stem Cell* **20**, 247-260 e245, doi:10.1016/j.stem.2016.11.007 (2017).
46 Tabula Muris, C. *et al.* Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature* **562**, 367-372, doi:10.1038/s41586-018-0590-4 (2018).
47 Vaillant, F. *et al.* The mammary progenitor marker CD61/beta3 integrin identifies cancer stem cells in mouse models of mammary tumorigenesis. *Cancer Res* **68**, 7711-7717, doi:10.1158/0008-5472.CAN-08-1949 (2008).
48 Welm, B. E. *et al.* Sca-1(pos) cells in the mouse mammary gland represent an enriched progenitor cell population. *Dev Biol* **245**, 42-56 (2002).
49 Pfefferle, A. D. *et al.* The MMTV-Wnt1 murine model produces two phenotypically distinct subtypes of mammary tumors with unique therapeutic responses to an EGFR inhibitor. *Dis Model Mech* **12**, doi:10.1242/dmm.037192 (2019).
50 Pfefferle, A. D. *et al.* Transcriptomic classification of genetically engineered mouse models of breast cancer identifies human subtype counterparts. *Genome Biol* **14**, R125, doi:10.1186/gb-2013-14-11-r125 (2013).

51 Cleary, A. S., Leonard, T. L., Gestl, S. A. & Gunther, E. J. Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers. *Nature* **508**, 113-117, doi:10.1038/nature13187 (2014).
52 Roarty, K. & Rosen, J. M. Wnt and mammary stem cells: hormones cannot fly wingless. *Curr Opin Pharmacol* **10**, 643-649, doi:10.1016/j.coph.2010.07.004 (2010).
53 Hatsell, S., Rowlands, T., Hiremath, M. & Cowin, P. Beta-catenin and Tcfs in mammary development and cancer. *J Mammary Gland Biol Neoplasia* **8**, 145-158 (2003).
54 Herschkowitz, J. I. *et al.* Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol* **8**, R76 (2007).
55 Gyorffy, B., Surowiak, P., Budczies, J. & Lanczky, A. Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer. *PLoS One* **8**, e82241, doi:10.1371/journal.pone.0082241 (2013).
56 Madden, S. F. *et al.* BreastMark: an integrated approach to mining publicly available transcriptomic datasets relating to breast cancer outcome. *Breast Cancer Res* **15**, R52, doi:10.1186/bcr3444 (2013).
57 Spike, B. T. *et al.* A mammary stem cell population identified and characterized in late embryogenesis reveals similarities to human breast cancer. *Cell Stem Cell* **10**, 183-197, doi:10.1016/j.stem.2011.12.018 (2012).
58 Trejo, C. L., Luna, G., Dravis, C., Spike, B. T. & Wahl, G. M. Lgr5 is a marker for fetal mammary stem cells, but is not essential for stem cell activity or tumorigenesis. *NPJ Breast Cancer* **3**, 16, doi:10.1038/s41523-017-0018-6 (2017).
59 de Visser, K. E. *et al.* Developmental stage-specific contribution of LGR5(+) cells to basal and luminal epithelial lineages in the postnatal mammary gland. *J Pathol* **228**, 300-309, doi:10.1002/path.4096 (2012).
60 Rohrschneider, L. R., Custodio, J. M., Anderson, T. A., Miller, C. P. & Gu, H. The intron 5/6 promoter region of the ship1 gene regulates expression in stem/progenitor cells of the mouse embryo. *Dev Biol* **283**, 503-521, doi:10.1016/j.ydbio.2005.04.032 (2005).
61 Roarty, K. & Serra, R. Wnt5a is required for proper mammary gland development and TGF-beta-mediated inhibition of ductal growth. *Development* **134**, 3929-3939 (2007).
62 Zeng, Y. A. & Nusse, R. Wnt proteins are self-renewal factors for mammary stem cells and promote their long-term expansion in culture. *Cell Stem Cell* **6**, 568-577, doi:10.1016/j.stem.2010.03.020 (2010).
63 Vanhollebeke, B. *et al.* Tip cell-specific requirement for an atypical Gpr124- and Reck-dependent Wnt/beta-catenin pathway during brain angiogenesis. *Elife* **4**, doi:10.7554/eLife.06489 (2015).
64 Berndt, J. D. *et al.* Mindbomb 1, an E3 ubiquitin ligase, forms a complex with RYK to activate Wnt/beta-catenin signaling. *J Cell Biol* **194**, 737-750, doi:10.1083/jcb.201107021 (2011).
65 Valenzuela, D. M. *et al.* High-throughput engineering of the mouse genome coupled with high-resolution expression analysis. *Nat Biotechnol* **21**, 652-659, doi:10.1038/nbt822 (2003).

66 Lloyd-Lewis, B. *et al.* Imaging the mammary gland and mammary tumours in 3D: optical tissue clearing and immunofluorescence methods. *Breast Cancer Res* **18**, 127, doi:10.1186/s13058-016-0754-9 (2016).
67 Hatsell, S. J. & Cowin, P. Gli3-mediated repression of Hedgehog targets is required for normal mammary development. *Development* **133**, 3661-3670 (2006).
68 Chandramouli, A. *et al.* Ltbp1L is focally induced in embryonic mammary mesenchyme, demarcates the ductal luminal lineage and is upregulated during involution. *Breast Cancer Res* **15**, R111, doi:10.1186/bcr3578 (2013).
69 Imbert, A., Eelkema, R., Jordan, S., Feiner, H. & Cowin, P. ΔN89ß-catenin induces precocious development, differentiation, and neoplasia in mammary gland. *J. Cell Biology* **153**, 555-568 (2001).
70 Sorlie, T. *et al.* Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A* **100**, 8418-8423 (2003).
71 Hu, Z. *et al.* The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics* **7**, 96, doi:10.1186/1471-2164-7-96 (2006).

Acknowledgements:

This work was supported by Department of Defense W81XWH-17-1-0013 BC160959 (PC), NIH-R21-CA129905 and NYS Peter Rowley Award C02857 (PC), NIH training grants: T32GM066704 (JS), T32CA009161-44 (AI) and The Susan G Komen Foundation For The Cure (AI). We thank Dr. Aris Economides, Regeneron, Tarrytown NY for advice on *Adgra3*cre-ER^{T2} construction. We thank Drs. Alireza Khodadadi-Jamayran, Cynthia Loomis, Michael Cammer, Peter Lopez for assistance provided by the Applied Bioinformatics, Experimental Pathology, Microscope Laboratory and Flow Cytometry Research Cores: Grants P30CA016087, S10OD021747 and P30CA016087, and Drs. Alan Frey and David Frendewey for reading the manuscript.

Author Contributions: PC conceived the study, designed experiments and wrote the paper with input from ES, CJW ; ES JS AI AC conducted the experiments.

Figure 1

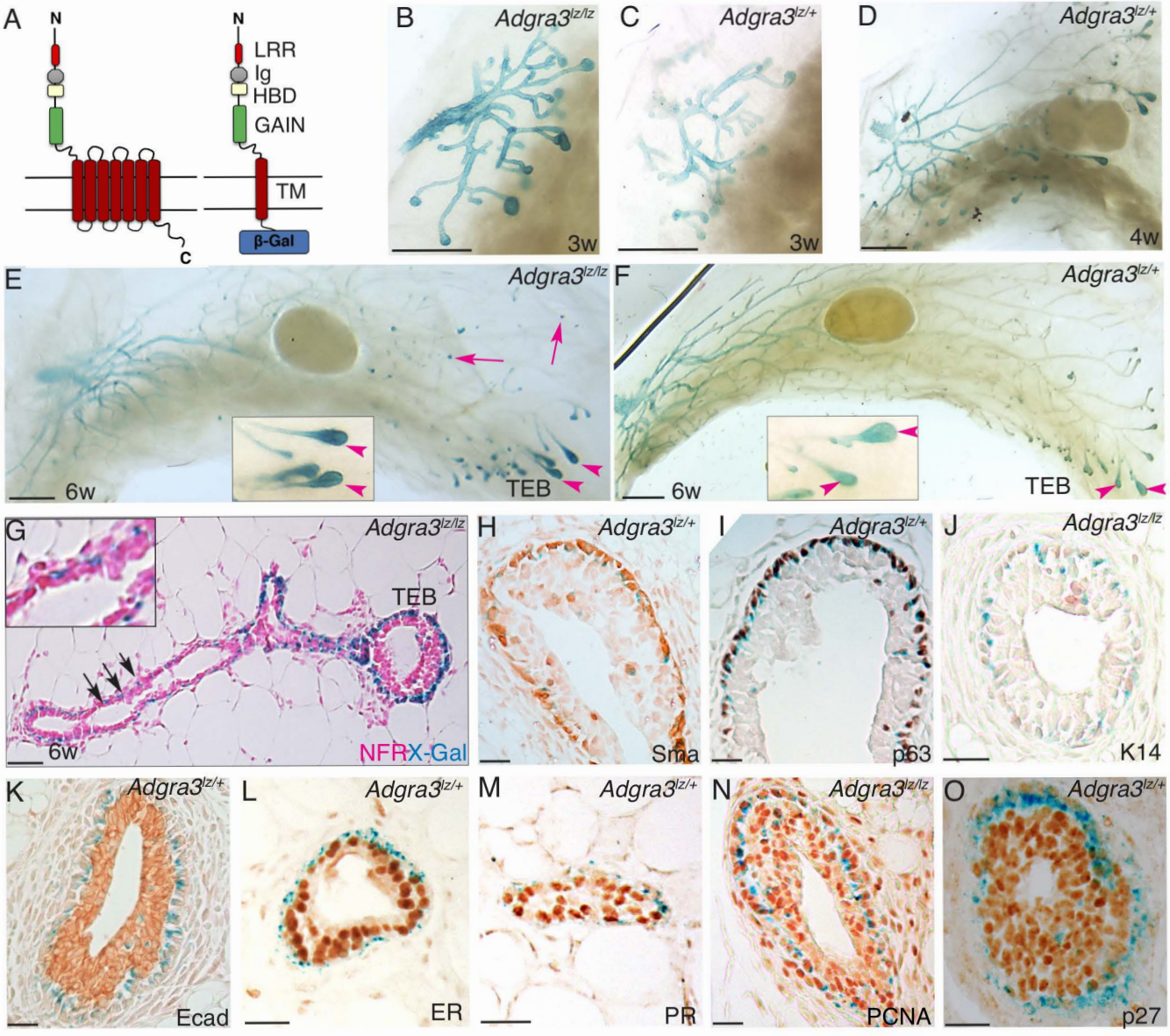
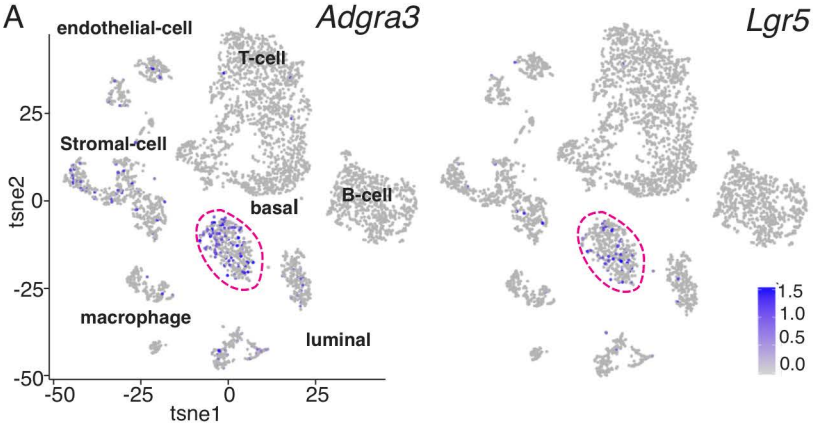
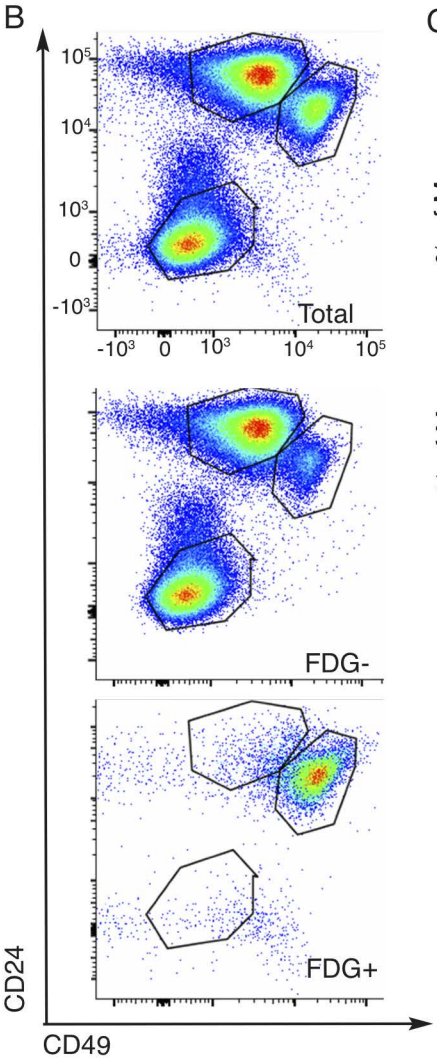

A
N
LRR
Ig
HBD
GAIN
TM
β-Gal
C
B
Adgra3lz/lz
3w
C
Adgra3lz/+
3w
D
Adgra3lz/+
4w
E
Adgra3lz/lz
6w
TEB
F
Adgra3lz/+
6w
TEB
G
Adgra3lz/lz
TEB
6w
NFRX-Gal
H
Adgra3lz/+
Sma
I
Adgra3lz/+
p63
J
Adgra3lz/lz
K14
K
Adgra3lz/+
Ecad
L
Adgra3lz/+
ER
M
Adgra3lz/+
PR
N
Adgra3lz/lz
PCNA
O
Adgra3lz/+
p27

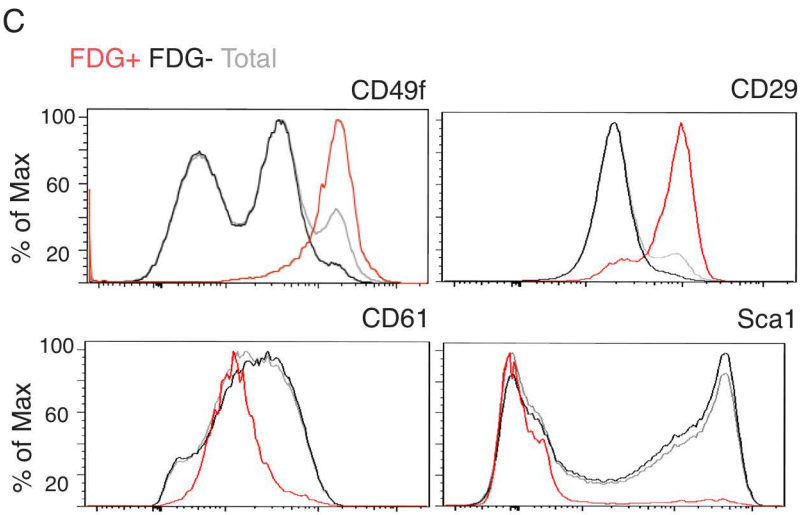
Figure 1. Gpr125 is expressed as predicted sites of stem/progenitor activity during pubertal mammary development.

A) Left: Schematic of Gpr125 protein, **Right**: Gpr125-β-gal fusion protein. N-terminus (N), leucine rich repeats (LRR), Immunoglobulin-like domain (Ig), hormone binding domain (HBD), GPCR autoproteolytic-inducing (GAIN) domain, transmembrane region (TM) and cytoplasmic region (C). **B-F)** Gpr125-β-gal expression in mammary whole mounts from pre-pubertal (3w) and pubertal (5w and 7w) nulliparous mice. Scale bar 2mm. **G)** section of X-gal treated whole mount counterstained with nuclear fast red (NFR), shows Gpr125 expression in the cap layer of terminal end buds (TEB) and in cells dispersed along the basal layer of subtending ducts. The inset box is a higher magnification of area indicated by arrows. **H-M)** X-gal stained sections of TEB with immunolocalization of (**H-J**) basal markers: smooth muscle actin (SMA), p63, Keratin (K14); Note the occasional cells expressing Gpr125 within the body layer all express basal cap cell markers. (**K-M**) luminal markers: E-cadherin (Ecad), estrogen receptor (ER), progesterone receptor (PR); (**N,O)** markers indicating proliferative status: proliferating nuclear cell antigen (PCNA) and p27. Scale bar in sections are 50μm.

Figure 2


A
Adgra3
Lgr5
endothelial-cell
T-cell
Stromal-cell
B-cell
basal
macrophage
luminal
tsne2
tsne1
25
0
-25
-50
-50
-25
0
25
1.5
1.0
0.5
0.0


B
Total
FDG-
FDG+
CD24
CD49
10^5
10^4
10^3
0
-10^3


C
FDG+ FDG- Total
CD49f
CD29
CD61
Sca1
% of Max
100
60
20

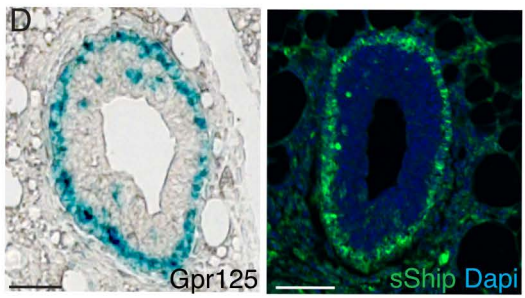

D
Gpr125
sShip Dapi

Figure 2. Gpr125+ cells have an MaSC/MRU profile and encompass distinct regenerative populations.

A) t-SNE plots of *Adgra3* and *Lgr5* in total mammary cells mined from Tabula Muris dataset [46]. **B)** Representative FACS dot plots of total MECs isolated from 6-week *Adgra3*$^{lz/+}$ pubertal mice stained with antibodies against CD24 and CD49f. Nb. Gpr125^{+}/FDG^{+} cells gated within the CD24$^{med/low}$/CD49f$^{+/hi}$ basal population and the basal population is depleted in FDG- populations. **C)** Histograms of expression of CD49f, CD29, Sca1 and CD61 in Gpr125^{+}/FDG^{+} (red line), Gpr125^{-}/FDG^{-}(black line) and total MECs (grey line). **D)** TEB sections show expression of Gpr125-β-gal (**left**) and sShip-EGFP cells in the cap cell layer of TEB (**right panel**). Scale bar = 50μm.

Figure 3

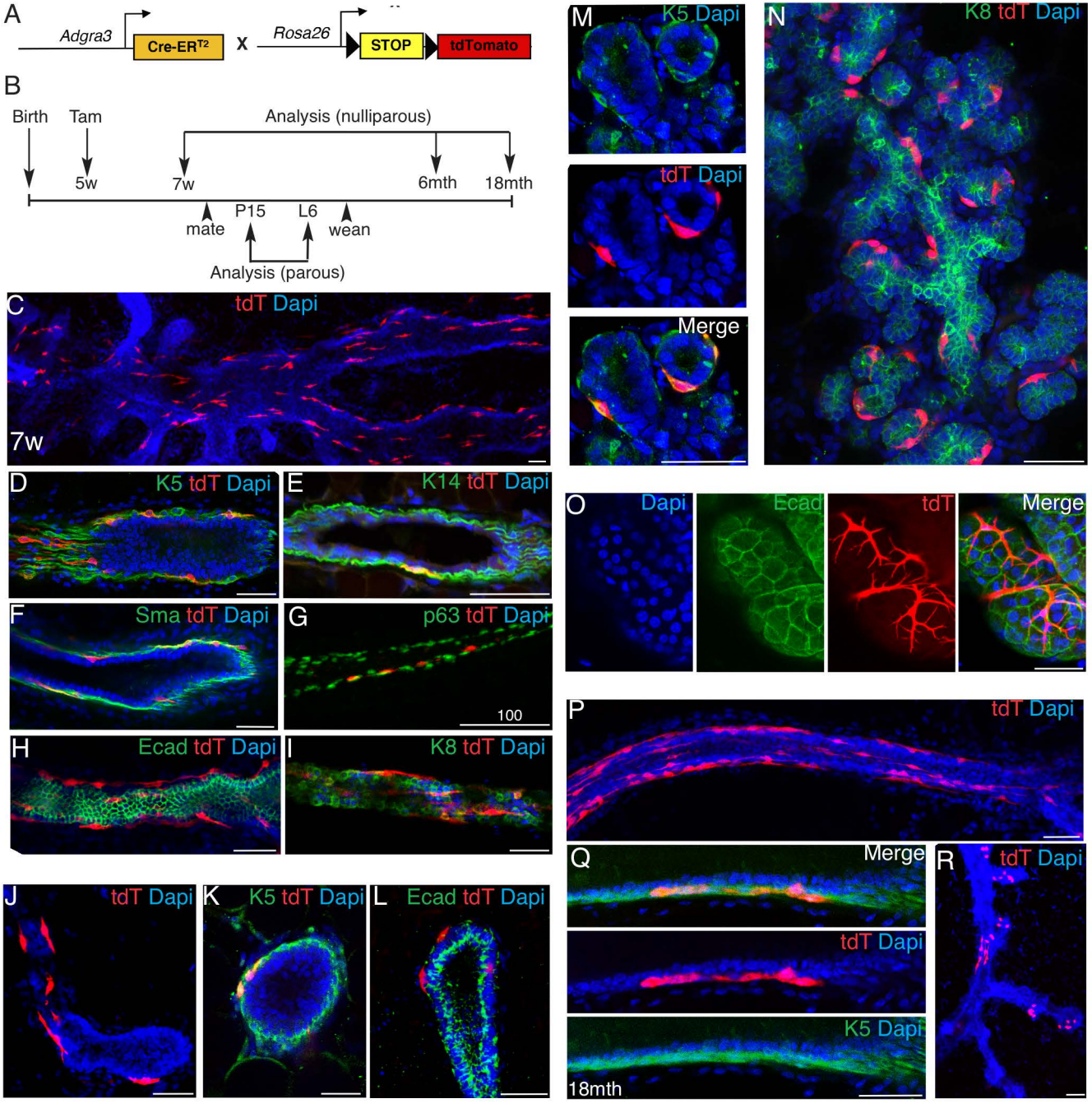

A
Adgra3
Cre-ER^{T2}
X
Rosa26
STOP
tdTomato
B
Birth
Tam
5w
Analysis (nulliparous)
7w
6mth
18mth
mate
P15
L6
wean
Analysis (parous)
C
tdT Dapi
7w
D
K5 tdT Dapi
E
K14 tdT Dapi
F
Sma tdT Dapi
G
p63 tdT Dapi
100
H
Ecad tdT Dapi
I
K8 tdT Dapi
J
tdT Dapi
K
K5 tdT Dapi
L
Ecad tdT Dapi
M
K5 Dapi
tdT Dapi
Merge
N
K8 tdT Dapi
O
Dapi
Ecad
tdT
Merge
P
tdT Dapi
Q
Merge
tdT Dapi
K5 Dapi
18mth
R
tdT Dapi

Figure 3. Pubertal Gpr125+ cells are long-lived unipotent basal progenitors.

A) Genetic strategy used to target Tomato expression to Gpr125-expressing cells and their progeny. **B)** Schematic illustrating strategy to trace the lineage of cells that express Gpr125 during early puberty by delivering Tamoxifen (Tam) to (5w) nulliparous mice and harvesting glands during late puberty (7w), mid-pregnancy (P15); lactation(L6) and in aged mice. **C)** Representative 3D-imaging of mammary gland nipple-proximal region of *Adgra3-CreER^{T2};tdT* mice analyzed 2 weeks after labeling. **D-G)** tdT+ cells with immunolocalization of basal markers K5, K14, SMA, and p63, and exclusion of luminal markers Ecad, K8 **(H,I)**. **J)** tdT+ cells in the cap cell layer of TEB and subtending duct with immunolocalization of basal K5 **(K)** and exclusion of luminal Ecad **(L)**. **M)** p15 Immature alveoli show basally located tdT+ cells that express K5, but lack K8 expression **(N). O)** tdT+ cells form characteristic basket-like morphology of mature myoepithelial cells enmeshing an Ecad+ alveoli in lactating mammary gland. **P-R)** Extensive strips of tdT-positive cells along the outer basal layer of ducts in aged mice **(P,Q)** and after multiple pregnancies **(R).** Scale bar 50μm. Dapi = nuclear staining. Two glands from each of 5 mice were analyzed/stage.

Figure 4

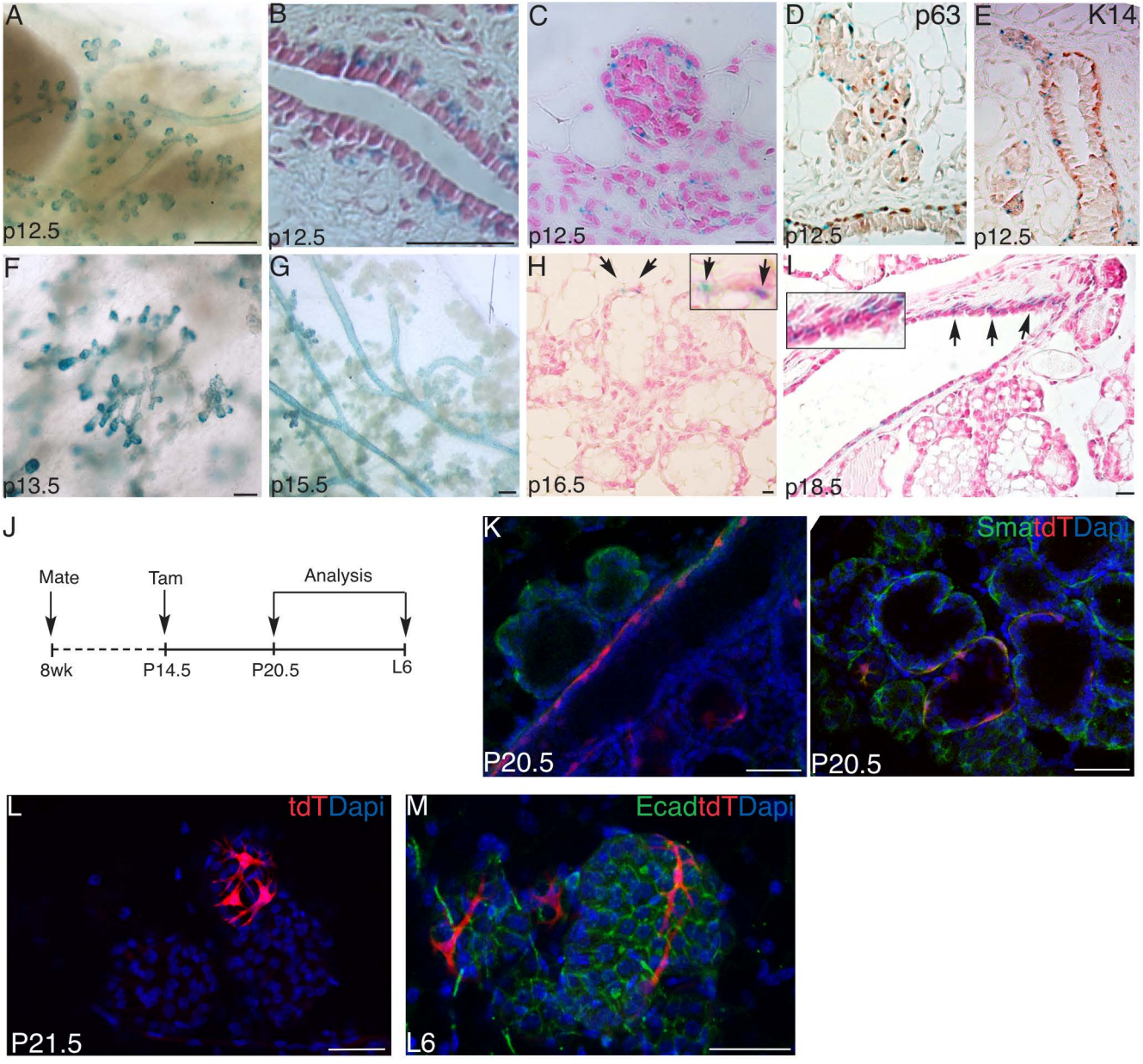

A
p12.5
B
p12.5
C
p12.5
D
p63
p12.5
E
K14
p12.5
F
p13.5
G
p15.5
H
p16.5
I
p18.5
J
Mate
Tam
Analysis
8wk
P14.5
P20.5
L6
K
P20.5
SmatdTDapi
P20.5
L
tdTDapi
P21.5
M
EcadtdTDapi
L6

Figure 4. Cells expressing Gpr125 on leading tips of branches during pregnancy are unipotent basal progenitors.

X-gal stained mammary whole mounts from pregnant mice showing Gpr125-β-gal expression at: **A)** sites of emerging side-branches at 12.5 days of pregnancy (p12.5), **F)** tips of elongating side-branches at p13.5 and **G)** in ducts but not alveoli at p15.5. Scale bar 500μm. **B, C, H, I)** X-gal stained sections counterstained with NFR. Boxed insets are higher magnification of regions indicated by arrows. **D, E)** X-gal stained sections of a permanent duct and side branch with immunolocalization of basal marker p63 and K14. Scale bar 20μm. **J)** Schematic illustrating strategy to trace the lineage of cells that express Gpr125 during early pregnancy by delivering Tam at p14.5 and harvesting glands at p20.5 and L6 mice. **K)** Tdt+ Sma+ cells were found in ducts and alveoli at p20.5 and at birth **(L)**. **M)** Fully differentiated tdt+ cells were devoid of luminal Ecad marker at L6. Scale bar 50μm. 1 gland from each of 5 mice were analyzed.

Figure 5

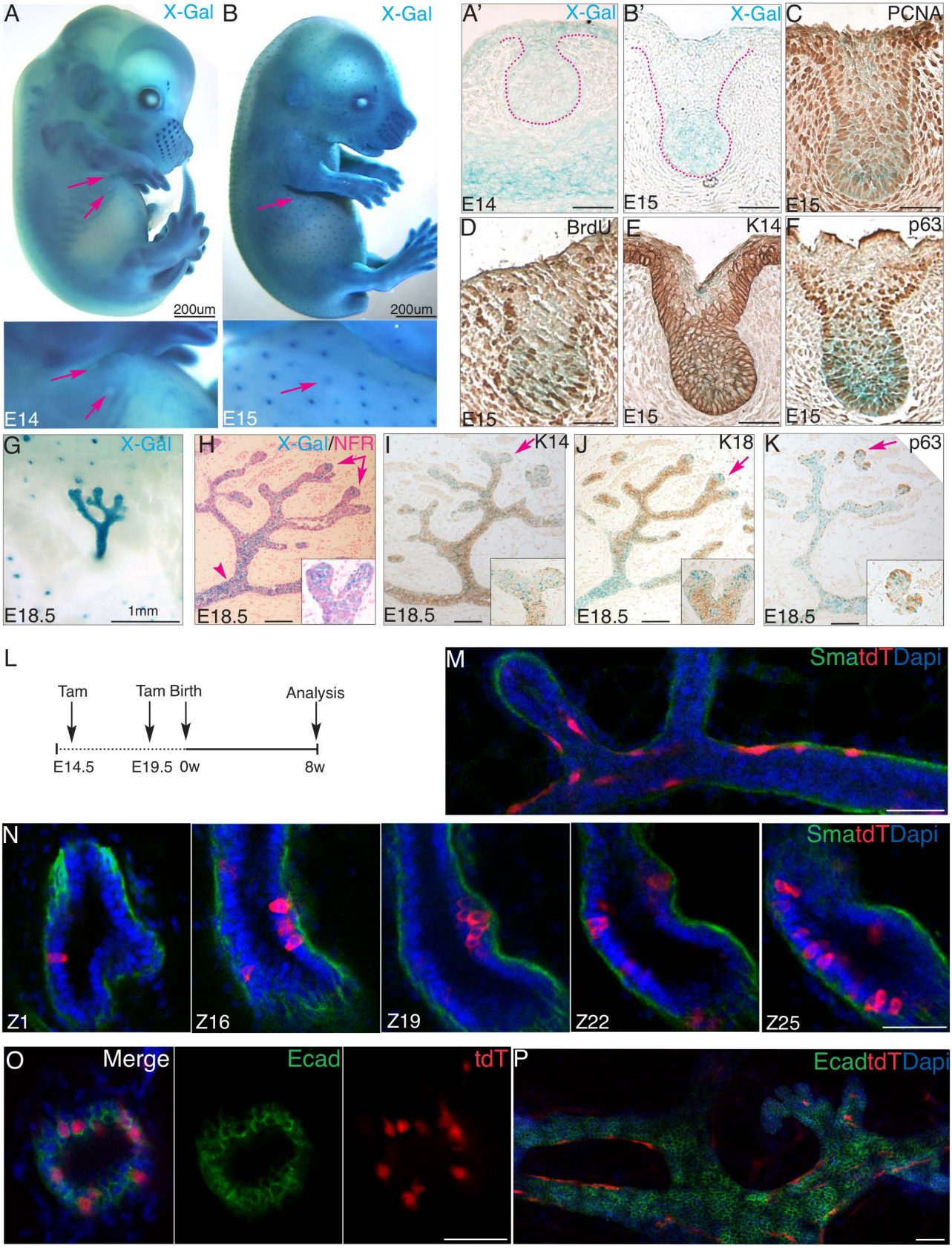

A
X-Gal
B
X-Gal
200um
200um
E14
E15
A'
X-Gal
B'
X-Gal
C
PCNA
E14
E15
E15
D
BrdU
E
K14
F
p63
E15
E15
E15
G
X-Gal
H
X-Gal/NFR
I
K14
J
K18
K
p63
E18.5
1mm
E18.5
E18.5
E18.5
E18.5
L
Tam
Tam
Birth
Analysis
E14.5
E19.5
0w
8w
M
SmatdTDapi
N
SmatdTDapi
Z1
Z16
Z19
Z22
Z25
O
Merge
Ecad
tdT
P
EcadtdTDapi

Figure 5. Gpr125 identifies an early bipotent and later unipotent basal progenitor population during embryogenesis.

A, B) X-gal stained *Adgra3*$^{lz/+}$ embryos at embryonic day (E)14 **(A)** and E15 **(B)**. Arrows indicate mammary buds and sprout respectively magnified in boxes below. Scale bar 200μm. **A'-F)** Gpr125-β-gal expression in sections of E14 buds **(A')** and E15 sprouts**(B'-F),** with immunolocalization for proliferative markers PCNA and BrdU **(C, D**), and K14 and p63 **(E,F)**. Scale bar 50μm. **G)** X-gal stained skin whole mount showing Gpr125-β-gal expression in the E18.5 tree and hair follicles encircling the developing nipple zone. Scale bar 1mm. **H-K)** Sections of the E18.5 rudiment tree stained with X-gal followed by NFR counterstain**(H)**; immunochemical detection of K14, K18, and p63 **(I-K)**. Boxed insets are higher magnification of branch tips regions indicated by arrows. Scale bar 50μm. **L)** Tracing of Gpr125+ cells in E14.5 or E19.5 embryos was initiated by delivering Tam in pregnant *Adgra3*$^{cre/cre}$ dams mated to Tdt mice. Mammary tissue from the progeny were analyzed at 8 weeks of age. **M-O)** 3-D images showing representative regions of pubertal ducts from E14.5 labelled embryo containing clusters of basally located tdT+ cells that co-express the basal marker SMA **(M)** as well as tdT+ columnar luminal cells lacking SMA **(N)** and expressing luminal marker Ecad **(O)**. **P)** Glands from progeny labeled at E19.5 show basally located tdT+ cells devoid of Ecad. Dapi=nuclear staining. Scale bar 50μm. Two glands from each of 3 mice were analyzed/stage.

Figure 6

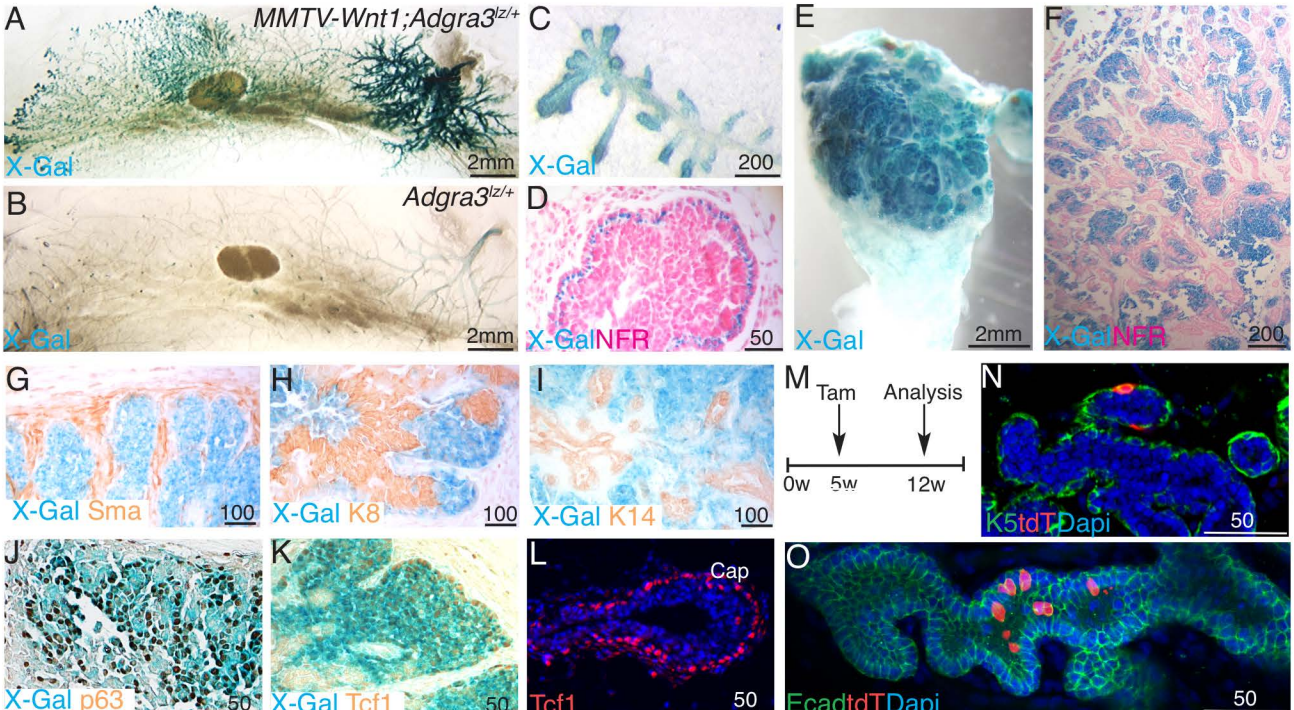
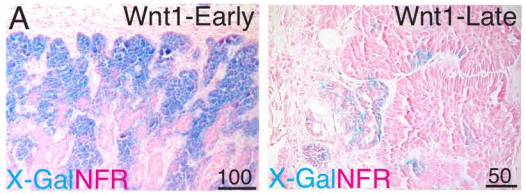
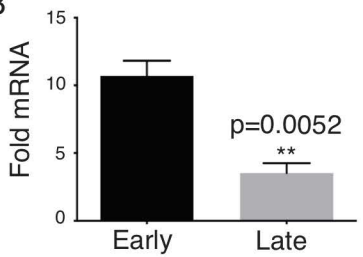

A
MMTV-Wnt1;Adgra3lz/+
X-Gal
2mm
B
Adgra3lz/+
X-Gal
2mm
C
X-Gal
200
D
X-GalNFR
50
E
X-Gal
2mm
F
X-GalNFR
200
G
X-Gal Sma
100
H
X-Gal K8
100
I
X-Gal K14
100
M
Tam
Analysis
0w
5w
12w
N
K5tdTDapi
50
J
X-Gal p63
50
K
X-Gal Tcf1
50
L
Cap
Tcf1
50
O
EcadtdTDapi
50

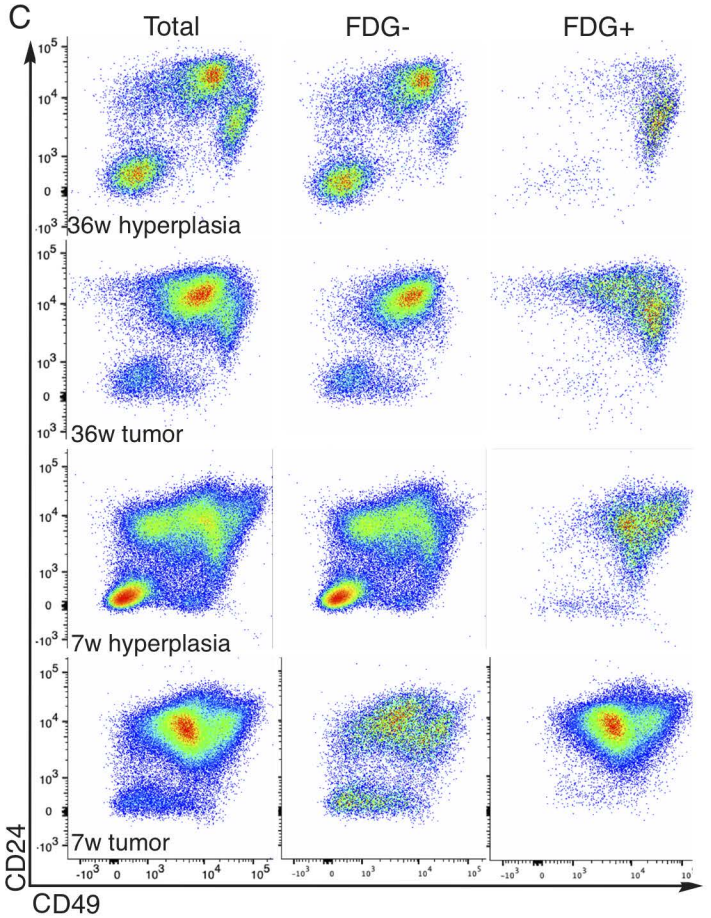
Figure 6. Gpr125+ progenitors are expanded in MMTV-Wnt1 tumors and retain embryonic features and bipotency.

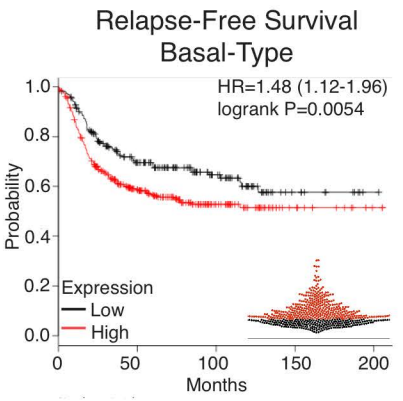
A-C) X-gal stained mammary whole mounts from 12-week old *MMTV-Wnt1;Adgra3$^{lz/+}$* mice show robust Gpr125-β-gal expression in hyperbranched ductal tips **(C)** and nipple-proximal zones compared to **(B)** control *Adgra3$^{lz/+}$* littermate. **D)** X-gal section of hyperbranched ductal tip counterstained with NFR presenting basal restriction of Gpr125-β-gal expression in 12-week *MMTV-Wnt1;Adgra3$^{lz/+}$* mice. **E-H)** X-gal stained whole mount and sections counterstained with NFR of *MMTV-Wnt1;Adgra3$^{lz/+}$* tumor (**G-K)** showing Gpr125 cells devoid of immunolocalization for SMA, K14 or K8, but expressing p63 and Tcf1. **L)** Tcf1 expression in the cap cells of normal TEB. **M-O)** Lineage tracing strategy in 5w *MMTV-Wnt1;Adgra3$^{lz/+}$* produced both **(N)** K5+ and **(O)** Ecad+ Tdt cells in hyperplastic glands. 1 gland from each of 5 mice were analyzed.

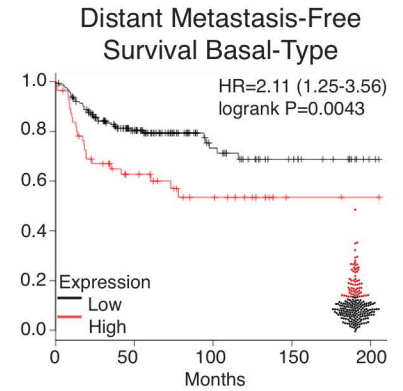
Figure 7


A
Wnt1-Early
Wnt1-Late
X-GalNFR
100
X-GalNFR
50


B
Fold mRNA
15
10
5
0
p=0.0052
**
Early
Late


C
Total
FDG-
FDG+
36w hyperplasia
36w tumor
7w hyperplasia
7w tumor
CD24
CD49


D
Relapse-Free Survival
Basal-Type
HR=1.48 (1.12-1.96)
logrank P=0.0054
Probability
Expression
Low
High
Months


E
Distant Metastasis-Free
Survival Basal-Type
HR=2.11 (1.25-3.56)
logrank P=0.0043
Expression
Low
High
Months

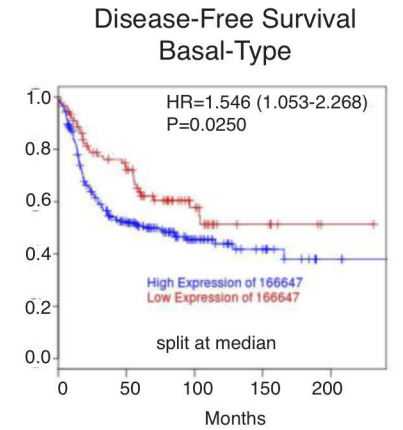
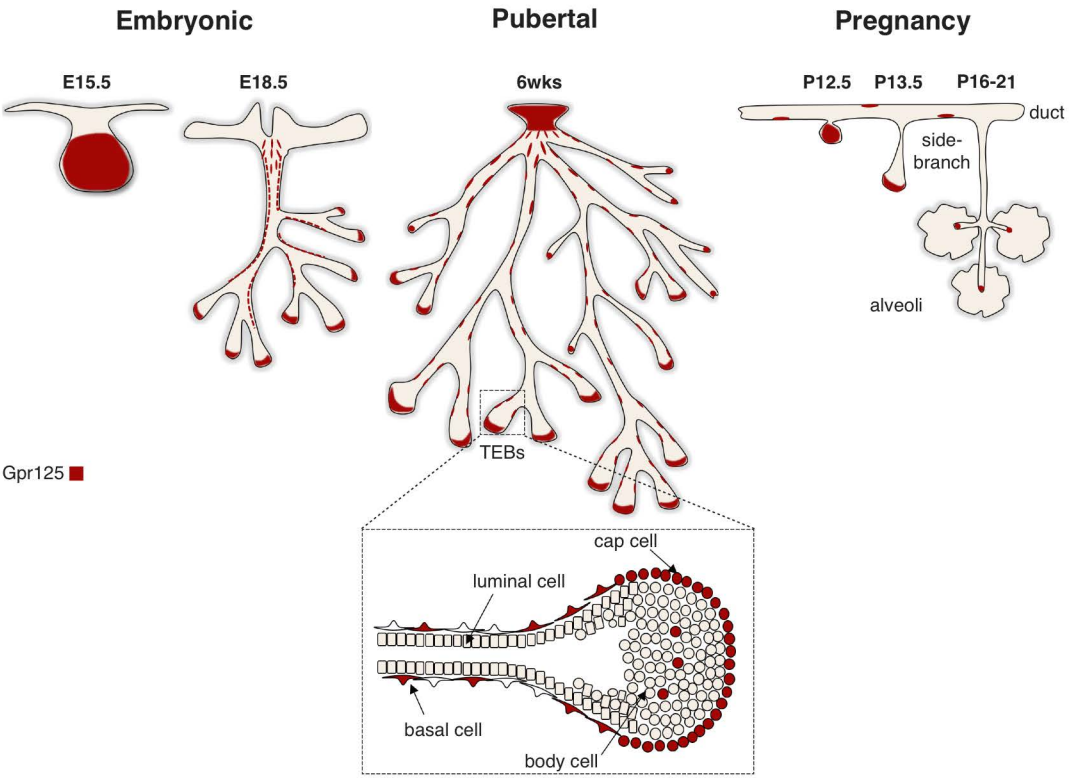

F
Disease-Free Survival
Basal-Type
HR=1.546 (1.053-2.268)
P=0.0250
High Expression of 166647
Low Expression of 166647
split at median
Months

Figure 7. High Gpr125 is predictive of poor outcome.

A) X-gal/NFR stained sections showing higher Gpr125-β-gal expression in a short (7w) latency tumor *MMTV-Wnt1;Adgra3*$^{lz/+}$ tumor (Wnt1-early) versus a long (36w) latency tumor (Wnt1-late) Scale bar 100μm and 50μm. **B)** Gpr125 mRNA levels in early and long latency tumors *MMTV-Wnt1;Adgra3*$^{lz/+}$ tumors (n=3 samples for each subtype, p=0.0052,**). **C)** Representative FACS dot plot of total MECs from *MMTV-Wnt1;Adgra3*$^{lz/+}$ hyperplastic uninvolved glands and associated short and long latency tumors stained with CD24 and CD49f. Plots are representative of 3 independent experiments. **D,E)** Relapse-free survival and Distant Metastasis Free Survival for high (**red lines**) and low Gpr125 mRNA expression (**black lines**) in human basal-like breast cancer subtype and autobestfit cutoff to divide patients into high and low expression (indicated in beehive plots) sourced from KMplotter (https://kmplot.com/analysis/index.php?p=service&cancer=breast)[55]. logrank P-value and Hazard Ratio (HR) indicate a significant association between high expression and poor prognosis. **F**) Disease-free survival utilizing BreastMark [56] ssp 2003/2006 datasets and splitting patients at the median level for expression (http://glados.ucd.ie/BreastMark/) into low (red line) and high (blue line) Gpr125 expression groups.

Figure 8

Gpr125 expression during mammary gland development



Gpr125 expression in MMTV-Wnt1 tumor

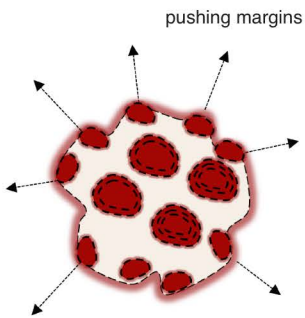


Figure 8. Schematic of location of Gpr125 progenitors over the course of mammary development

Gpr125 is expressed ~ E15 in a bipotent progenitor population concentrated towards the growing tip of the mammary sprout. Later, ~E18, as lineage segregation ensues, it becomes restricted to basal unipotent progenitors confined to K14+Lgr5+ cells in the nipple region and distal K14-/K18-negative cells at ductal tips. Gpr125 is retained in K14+/Lgr5+ cells in the nipple proximal zone throughout postnatal mammary development. During puberty Gpr125 is strongly expressed in sSHIP+ cap cells of terminal end buds during ductal elongation and in a population dispersed along the maturing ducts. During pregnancy Gpr125 expression increases in unipotent basal progenitors at tips of emerging side-branches but is absent from differentiated alveoli. In MMTV-Wnt1 early tumors Gpr125+ K14-/K18-negative bipotent progenitors are expanded and located in large islands and at pushing margins.

Figure S1

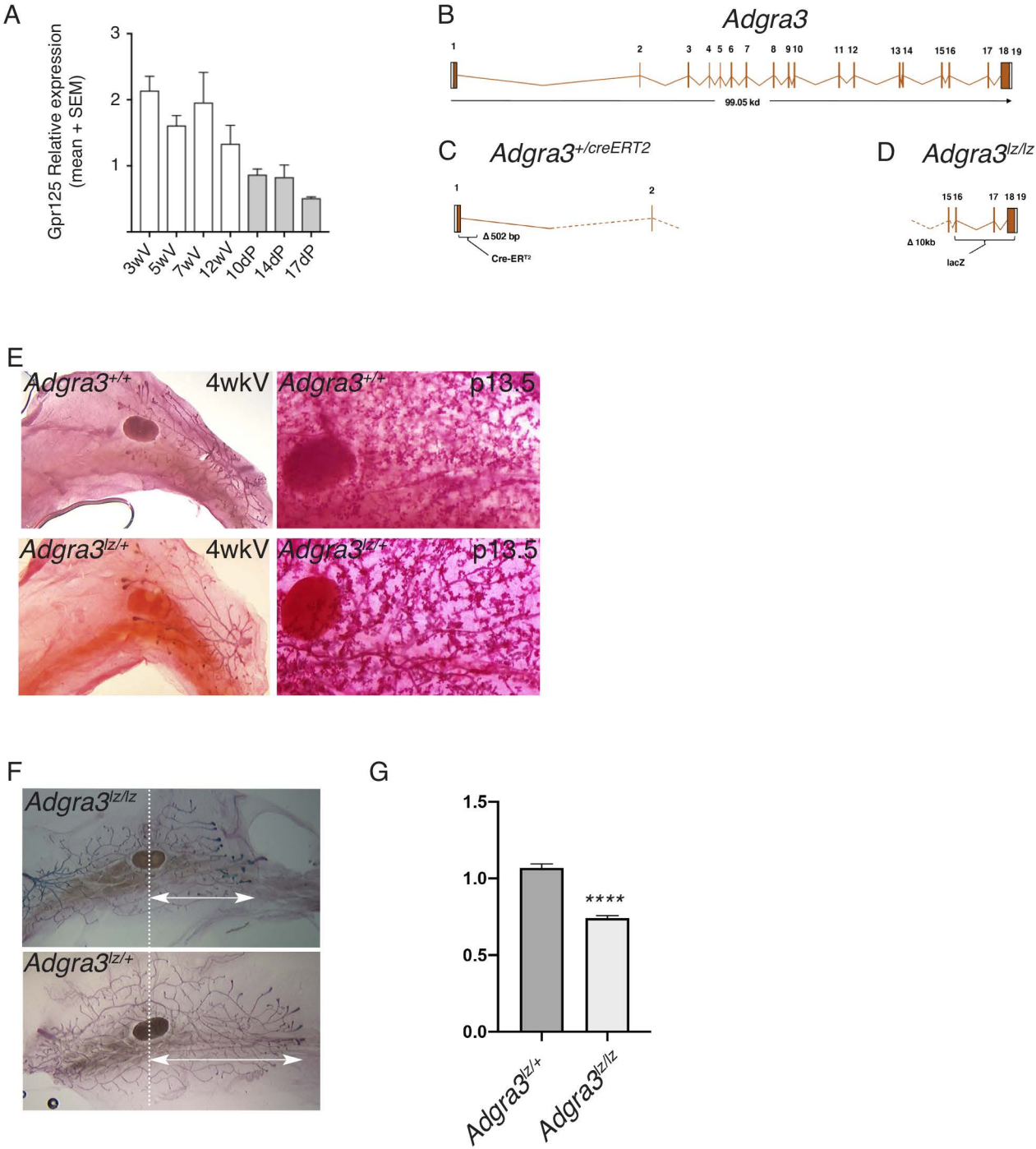
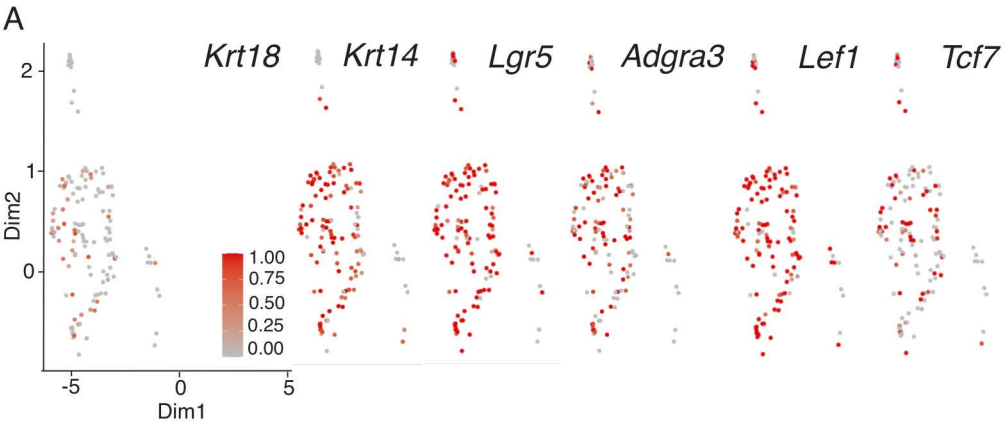
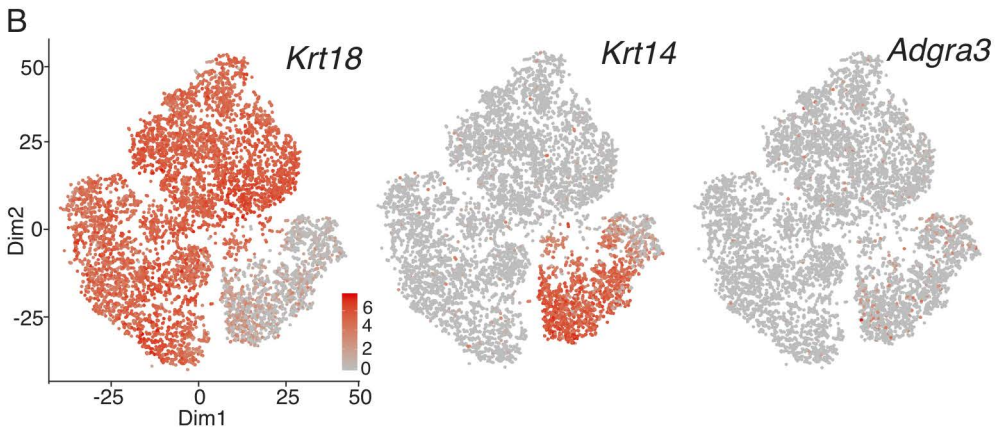

A
Gpr125 Relative expression (mean + SEM)
3
2
1
3wV
5wV
7wV
12wV
10dP
14dP
17dP
B
Adgra3
1
2
3
4 5 6
7
8
9 10
11 12
13 14
15 16
17
18 19
99.05 kd
C
Adgra3+/creERT2
1
2
Δ 502 bp
Cre-ERT2
D
Adgra3lz/lz
15 16
17
18 19
Δ 10kb
lacZ
E
Adgra3+/+
4wkV
Adgra3+/+
p13.5
Adgra3lz/+
4wkV
Adgra3lz/+
p13.5
F
Adgra3lz/lz
Adgra3lz/+
G
1.5
1.0
0.5
0.0
****
Adgra3lz/+
Adgra3lz/lz

Figure S1. *Adgra3*$^{lz/+}$ and *Adgra3*$^{cre/+}$ mouse strains

A) Bar plot of Gpr125 relative mRNA levels at different stages of mammary development normalized to male mammary glands; wV = weeks of age virgin nulliparous mice, dP = days of pregnancy. Each bar represents the mean ±SEM on 2-5 mice/stage with 4 technical replicates each**.** **B)** Schematic of *Adgra3* gene. **C)** *Adgra3*$^{cre/+}$ mice were generated by replacement of 502 bp after the first codon with a cassette containing c*reER*^{T2}. **D)** *Adgra3*$^{lz/+}$ mice were generated by deletion of 10 kb sequence downstream of the first TM and replacement by *lacZ* inserted in frame using Regeneron Velocigene technology [6]. **E)** Carmine stained mammary whole mounts from 4-week old and 13.5 days pregnant *Adgra3*$^{lz/+}$ and *Adgra3*$^{lz/lz}$ mice **F)** Mild impairment in ductal elongation in mammary wholemounts from pubertal *Adgra3*$^{lz/lz}$ females compared to littermates. Ducts from 7w old mice *Adgra3*$^{lz/+}$ mice have elongated further than those of *Adgra3*$^{lz/lz}$ littermates with respect to the midpoint of the inguinal lymph node marked by the dashed line. **G)** Quantitation of ductal elongation in glands from *Adgra3*$^{lz/lz}$ compared with pooled glands from *Adgra3*$^{lz/+}$ littermates ($p<0.0001$)****, n=6.

Figure S2


A
Krt18
Krt14
Lgr5
Adgra3
Lef1
Tcf7
Dim2
Dim1
1.00
0.75
0.50
0.25
0.00


B
Krt18
Krt14
Adgra3
Dim2
Dim1

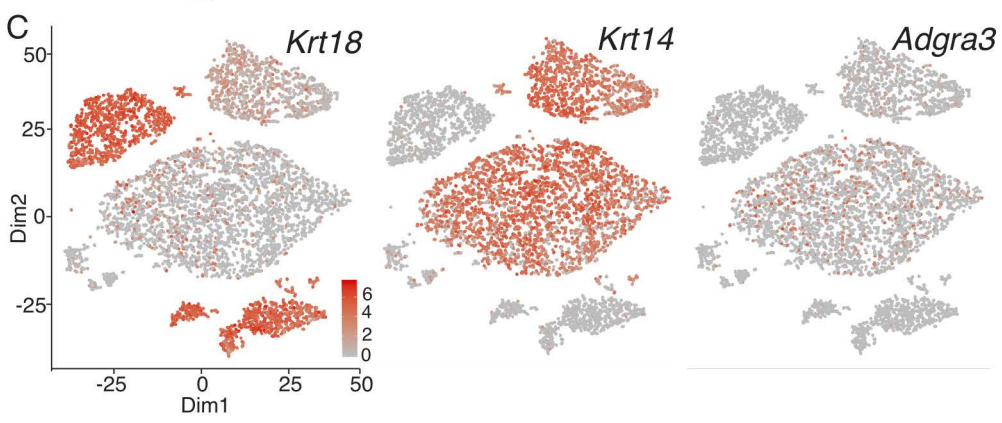
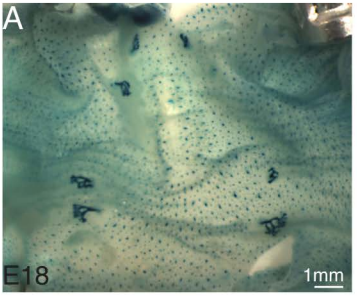

C
Krt18
Krt14
Adgra3
Dim2
Dim1

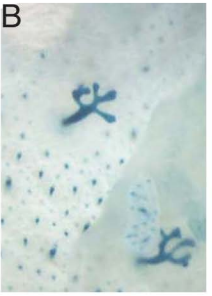
Figure S2. Gpr125 cells from pubertal and adult mice are confined to the basal cluster.

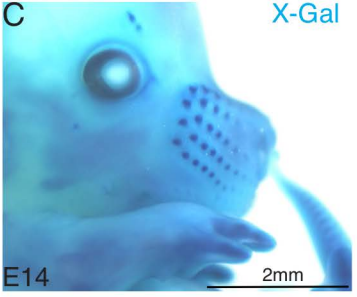
A) t-SNE plots of *Krt18, Krt14, Lgr5, Adgra3, Lef1 and Tcf7* in embryonic mammary cells mined from dataset generated by Wuidart et al [37].

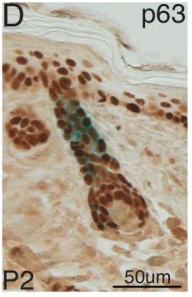
B,C) t-SNE plots of *Krt18*, *Krt14* and *Adgra3* in pubertal and adults mammary cells mined from dataset generated by Pal et al [35].

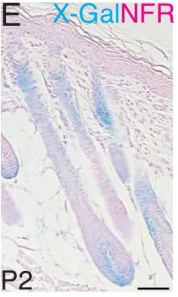
Figure S4


A
E18
1mm


B


C
X-Gal
E14
2mm


D
p63
P2
50um


E
X-GalNFR
P2

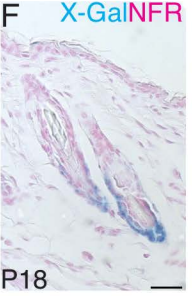

F
X-GalNFR
P18

Figure S4. Gpr125 expression in other ectodermal appendages

A,B) Skin from E18 embryo stained with X-gal shows Gpr125-β-gal expression in the five pairs of mammary trees and developing hair follicles. **C)** Whole mount of E14 embryo showing robust Gpr125-β-gal expression in the whisker-pad. **D-E)** Sections showing Gpr125-β-gal expression concentrated in the bulge and bulb compartments of P2 hair follicles during anagen, **(D)** immunolocalized with p63, or (**E)** counterstained with NFR. **(F)** Gpr125-β-gal expression in the secondary germ of P18 hair follicles during telogen

Figure S5

A

Mouse Tumor Model	Fold change
Wnt1-Early^{Ex}	3.17
p53null-Basal^{Ex}	1.76
Wnt1-Late^{Ex}	1.7
Class14^{Ex}	1.31
Class8^{Ex}	1.2
p53null-Luminal^{Ex}	0.95
Squamous-like^{Ex}	0.9
Neu^{Ex}	0.9
Erbb2-like^{Ex}	0.9
PyMT^{Ex}	0.8
Class3^{Ex}	0.78
Stat1^{Ex}	0.77
C3Tag^{Ex}	0.76
Myc^{Ex}	0.75
Normal^{Ex}	0.7
ClaudinLow^{Ex}	0.68
WapInt3^{Ex}	0.6

B

Human Breast Cancer subtype	Fold change
Human Basal-like	2.03
Human Normal-like	1.33
Human LumB	0.78
Human LumA	0.77
Human Her2-Enriched	0.75
Human Claudin-low	0.8

Figure S5. Gpr125 expression in murine and human breast cancer subtypes

A) Gpr125 mRNA fold change (from high to low mRNA level) in several mouse tumor models of breast cancers showing greater expression in Wnt1-Early^{Ex}, p53null-Basal^{Ex} and Wnt1-Late^{Ex} tumors. **B)** Gpr125 mRNA fold change higher in Human Basal like. Gpr125 mRNA fold change was obtained mining microarray data available at the Gene Expression Omnibus under the series GSE3165 [54].

Seminars in Cell and Developmental Biology xxx (xxxx) xxx

Contents lists available at ScienceDirect

Seminars in Cell and Developmental Biology

journal homepage: www.elsevier.com/locate/semcdb

Review

Embryonic mammary gland development

Elena Spina [a], Pamela Cowin [a,b,*]

[a] Department of Cell Biology, New York University School of Medicine, New York, NY, USA
[b] Department of Dermatology, New York University School of Medicine, New York, NY, USA

ARTICLE INFO

Keywords:
Embryonic mammary gland development
Breast cancer
Wnt
Hedgehog
Ltbp1

ABSTRACT

Embryonic mammary gland development involves the formation of mammary placodes, invagination of flask-shaped mammary buds and development of miniature bi-layered ductal trees. Currently there is a good understanding of the factors that contribute to ectodermal cell movements to create these appendages and of pathways that lead to mammary specification and commitment. Gene expression profiles of early bipotent mammary stem cells populations as well as cell surface proteins and transcription factors that promote the emergence of unipotent progenitors have been identified. Analyses of these populations has illuminated not only embryonic mammary development, but highlighted parallel processes in breast cancer. Here we provide an overview of the highly conserved pathways that shape the embryonic mammary gland. Understanding the dynamic signaling events that occur during normal mammary development holds considerable promise to advance attempts to eliminate cancer by restoring differentiative signals.

1. Introduction

1.1. Evolutionary importance and variety of mammary gland development

Mammary glands are the defining hallmark of mammals. Their emergence provided a significant advantage by enabling mothers to remain mobile while providing nutrition to their offspring. The importance of this evolutionary advance led Darwin to dedicate an entire chapter of his book "The Origin of Species" to their description.

Mammary glands arise as specializations of the embryonic ectoderm and comparative anatomy of living species suggests they evolved from ancestral hair-associated apocrine glands (for primary refs see [1,2]). For example, the duck-billed platypus produces milk from an abdominal patch comprising 100–200 compound mammolobular-pilosebaceous units. This secretion serves as an adhesive to bind the eggs to the mother and is wicked along specialized hairs to nourish the hatchlings. Koala bears form vestigial mammary hairs that regress as their nipples develop and similar structures that form in horses are retained. Squirrels develop bilateral sensory vibrissae and nipples from the same epidermal anlage. Molecular evidence suggests that cutaneous secretions that prevented desiccation of parchment eggs evolved to provide lactation, which enabled rapid growth of offspring by providing immediate inter-generational transfer of ~10% of maternal skeletal calcium [1,2]. With this fact in mind, it is not surprising that many proteins that regulate mammary development serve dual roles in bone metabolism [3].

2. Variety of mammary glands

There is a great variety in the position and number of mammary glands in different species (for primary refs see [1,2]). Manatees form axillary mammary glands, elephants and primates have a single pectoral pair, ungulates develop inguinal glands that are sometimes amalgamated to form an udder. Other species, for example pigs and the multi-mammate rat form two rows on either side of their body with a dozen glands in each. Marsupials have central abdominal glands that open into a pouch, and in the case of the kangaroo, secrete distinct types of milk from two types of nipple to feed offspring of different developmental stages. Whales and seals have retractable mammary glands to streamline their swimming, bats and possums make additional use their nipples as attachments sites to facilitate carrying their young. Setting aside such comparative studies and those driven by the motivation to maximize milk production and animal husbandry in farm animals [4], the majority of our knowledge about the embryonic stage of mammary development comes from studies on mice and rabbits. Genes associated with mammary development are highly conserved. Thus, although information on human embryonic mammary development is limited, a

* Correspondence to: Department of Cell Biology, New York University Medical Center, MSB 621, 550 1st Ave, New York, NY 10016, USA.
E-mail address: pamela.cowin@nyulangone.org (P. Cowin).

https://doi.org/10.1016/j.semcdb.2020.12.012
Received 5 September 2020; Received in revised form 3 December 2020; Accepted 29 December 2020

number of genes responsible for human developmental syndromes have been identified and studied further through the generation of genetically engineered mice with similar phenotypes (Table 1). It has also become apparent that the genes regulating directed collective cell migration, invasion, proliferation, cell potency and differentiation during murine embryonic mammary gland development are frequently reactivated in human breast cancer [5–11].

3. The basic developmental process

Murine mammary development begins mid-gestation ~ embryonic day 10.5 (E10.5) with the definition of bilateral milk lines. ~E11–E12 ectodermal cells coalesce into five pairs of mammary placodes. By E13.5 these cell movements form raised hillocks that quickly sink below the dermis to generate flask-shaped buds surrounded by a compacted highly specialized mammary mesenchyme. ~E15.5 the mammary primordium is destroyed in males but in females begins to proliferate and sprout towards a nearby fat pad. After invading the fat pad, it branches and forms a lumen, producing a hollow ductal tree by birth (Fig. 1). Mammary gland development, like that of all other ectodermal appendages, is regulated by highly conserved patterning pathways engaged in epithelial-mesenchymal cross-talk.

3.1. Definition of the mammary line by BMP4:TBX3 antagonism

Murine mammary glands form along arcs on either side of the body that roughly correspond to the dorsolateral-ventral (D/V) boundary of the underlying mesenchyme [12]. It has been proposed that these mammary lines are defined by mutual antagonism between ventral expression of the bone morphogenetic protein 4 (Bmp4) and dorsal expression of the transcription factor, Tbx3 (Table 2 and Fig. 2A) [12]. In rabbit embryos mammary lines are morphologically visible as a prominent ectodermal ridge [13]. Counterparts have been described in many species, however, in mice the corresponding structures are more nuanced and discernable only in histological sections, as a bilayered row of columnar cells within the presumptive region of the third mammary rudiment (MR) [14]. Molecularly, the field in which the mammary line will develop can be detected as early as ~E10.5 by a band of mesenchymal Tbx3 mRNA expression [15]. This is followed later ~E11.5 by an ectodermal line of Wnt10b mRNA between the fore and hind limbs and two streaks around the limbs where axillary and inguinal glands will form [16].

3.2. The essential role of early Wnt signaling along mammary line and limb streaks

Intriguingly, expression of TOP-gal, a transgenic reporter of canonical Wnt (Tcf/Lef) signaling, can be seen at E10.5 one day before the appearance of Wnt10b mRNA [17]. Its earlier activation suggests the involvement of additional Wnts or alternative regulators of Tcf/Lef genes. Wnt 3 and Wnt6, are present in the ectoderm and Wnt5a and Wnt11 in the mesenchyme at this time, and could contribute to boundary formation by antagonistic canonical and non-canonical Wnt signaling [17]. An absolute requirement for early Wnt signaling has been demonstrated by the finding that experimental expression of Dickkopf 1 (DKK1), a secreted Wnt inhibitor, abolishes the earliest molecular features (expression TOP-gal, Wnt10b and Tbx3) of mammary line formation [17]. In contrast, although deletion of Lef1, a downstream transcriptional mediator of canonical Wnt signaling, impairs all ectodermal appendages the effect on mammary rudiments is more attenuated that those produced by DKK1, suggesting partial redundancy with other members of the Tcf/Lef family [18]. Importantly, multiple studies (Table 2) have shown that Wnt signaling is essential not only for initiating mammary morphogenesis but remains critical at all subsequent stages of mammary formation and is active in both epithelial and mesenchymal compartments (Fig. 3).

Table 1
Table of human syndromes involving major regulators of embryonic mammary development.

Genes	Human defects/syndrome
TBX3	Ulnar mammary syndrome (UMS)
FGF10/FGFR2	Poland syndrome
EDA/EDAR	Hypohidrotic ectodermal dysplasia (HED)
PTHLH	Blomstrand chondrodysplasia

3.3. Migration of ectodermal cells into placodes

The mammary line and limb streaks are transient entities, and in less than a day, Tbx3, TOP-gal and Wnt10b expression become confined to the elliptical placodal thickenings of the ectoderm [15–17,19,20]. Placodes arise by ectodermal cell rearrangement rather than proliferation and form in a rather surprising temporal order [14]. Placode MR#3 forms first, followed swiftly by MR#4 through coalescence of cells at either end of the central mammary line. MR#1 emerges (sometimes preceding MR#4) together with MR#5 by movement of cells along the streaks encircling the fore and hind limbs respectively [21]. The last placode to develop, MR#2, is thought to receive cells from the forelimb streak as well as the mammary line. The suggestion that cells actively migrate along the central mammary line arose from early studies on rabbit embryos. Cells tagged with carbon were located over time and scanning electron micrographs showed cell processes polarized towards the placodes [13,21]. It has been proposed that Wnt expression may stimulate this migration in a manner similar to its actions in propelling cells out of intestinal crypts. The significance of the mammary line as a migratory path, however, has been challenged [14]. Placode MR#3 has been proposed to result instead from centripetal aggregation of ectodermal cells towards an Fgf10 attractant, emanating vertically from the tips of underlying somites (see Section 4.3) [22,23]. This model proposes that cells sustain Wnt signaling only when they enter the placode and Wnt signaling is extinguished elsewhere.

In this ongoing debate, movement of cells has, for the most part, focused on MR#3. Neuregulin 3 (Nrg3), for example, has emerged as a candidate that promotes migration of mammary progenitors at this site (Fig. 2B; see Section 4.4) [24,25]. In contrast, movement around the limbs has been relatively neglected. The earliest marker to appear ~E11 within the mesenchyme at these sites is latent Tgfβ-binding protein 1 (Ltbp1) (Fig. 4) [26]. Ltbp1 tethers and positions latent forms of Tgfβ, along fibronectin and elastin fibers and, importantly, plays an essential role in integrin-mediated stretch activation of this promigratory cytokine [27]. Thus, Ltbp1 is ideally placed in time and space to focally stimulate and guide movement of overlying ectodermal cells into MR# 1, 2, and 5. Ltbp1 is also strongly expressed in the mammary mesenchyme of all MR during invagination and sprouting (Fig. 4B) [26]. Functional testing of its role these processes, however, is precluded due to early lethality of Ltbp1 embryos [27]. In support of a promigratory role, LTBP1 appears in multiple metastasis signatures derived from human breast cancer cell lines [27].

4. Site-specific factors and regulatory genes

Analyses of mutant and genetically engineered mice has revealed the surprising fact that each placode pair is independently governed by its own unique complement of essential regulatory genes (Table 2). Attempts have been made to establish potential epistatic relationships by analyzing the effects of loss of one gene on the expression of others. As placode MR#3 has been studied the most intensively a working model is presented for this site in (Fig. 2). We propose that site specific genes act collectively either to direct cell migration and/or to augment and/or sustain canonical Wnt signaling, which must reach a critical threshold within a narrow temporal window for placodes to form and be

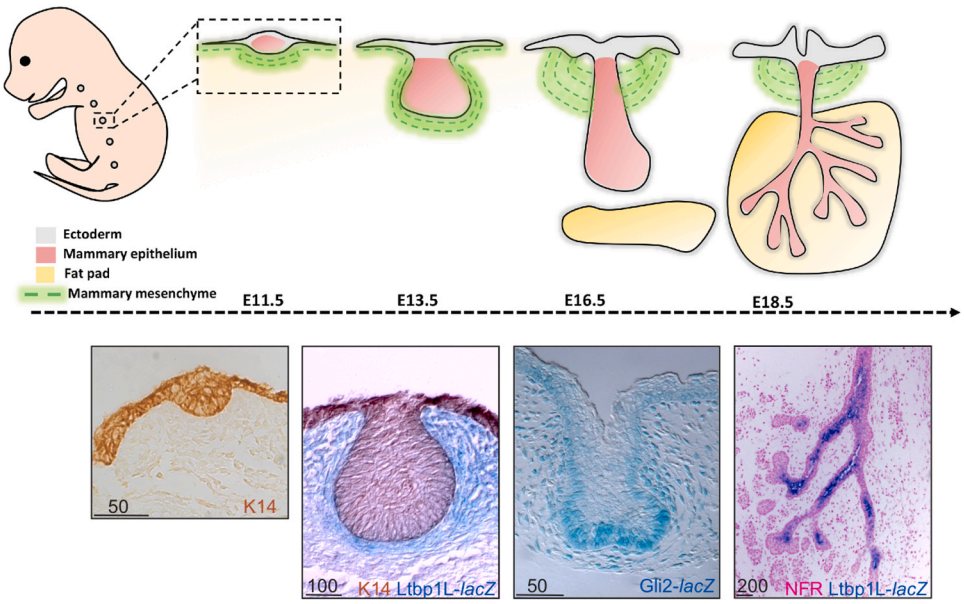


Fig. 1. Stages of murine embryonic mammary development. Left to right: Mammary development begins in the mouse embryo with thickening of the ectoderm to generate five pairs of raised elliptical placodes (pink). At E13.5 placodes invaginate to form flask-shaped buds surrounded by a specialized mammary mesenchyme (green). Once committed to mammary cell fate they proliferate and extend toward the underlying fat pad (yellow) (E16.5). At E18.5 a phase of branching morphogenesis ensues and microlumen develop to form a hollow bilayered rudimentary tree prior the birth. Histological sections show from left to right an E11.5 placode stained for keratin 14 (K14) (brown); E13.5 mammary bud stained for K14 (brown) and surrounded by mesenchymal Ltbp1L-*lacZ* expression (blue); E16.5 mammary sprout showing expression of Gli2-*lacZ* (blue); E18.5 rudimental tree expressing Ltbp1L-*lacZ* (blue) counterstained with Nuclear Fast Red (NFR) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

maintained.

4.1. *Gli3 repression of Hedgehog signaling is required for mammary fate of MR#2–5*

Gli3, is a transcription factor in the Hedgehog pathway and currently considered to be the most upstream regulator of MR#2–5. Hedgehog signaling operates by processing Gli transcription factors, tethered at the tip of the primary cilium, into transcriptional activators (A) or repressors (R). Briefly, binding of Hedgehog ligands to Patched receptors promotes entry of Smoothened, the Hedgehog signal transducer, into the primary cilium, which converts Gli2 to an activator (Gli2A) of gene transcription. Gli1 is induced by Hedgehog signaling, and hence used as a reporter of pathway activity, and serves as an amplifier. Gli3 generally represses the pathway but, depending on the cell context, can also function as an activator. Genetic manipulations designed to favor the Gli activator to repressor ratio have shown the surprising result that Hedgehog signaling must be repressed for most mammary rudiments to form properly (Fig. 5) [28]. This is in stark contrast to other ectodermal appendages, such as hair and teeth, which require positive Hedgehog signaling for appendage downgrowth (Fig. 5). Mice lacking Gli3 repressor function (*Gli3*$^{xt/xt}$) show abnormal patterning of the mammary field, displaying inappropriate dorsal expansion of Bmp4, constriction of the Tbx3 band and failure to concentrate Tbx3 within the presumptive MR#3 placodal region [29]. The targets of Gli3 repression at this early stage remain to be determined. However, Gli3 binding sites in the murine Bmp4 promoter, suggest that Bmp4 could be a target [30]. In this scenario, Gli3R functions to de-repress Tbx3 by sculpting the zone of Bmp4 expression [29].

At later stages of mammary development mice lacking Gli3 repressor function (*Gli3*$^{xt/xt}$) and those in which Hedgehog signaling is misactivated (*Gli2*$^{1ki/1ki}$) show identical phenotypes: failure to form MR#3 and MR#5 and severe impairment of MR#2, MR#4 invagination (Fig. 5) and loss of sexual dimorphism (Table 2) [28,29]. Intriguingly, in these strains, where the pathway is misactivated, hair follicles develop inappropriately within the mammary field (Fig. 5B) [29]. In a separate study, abrogation of Hedgehog signaling, through elimination of Smoothened, revealed the converse phenotype where mice displayed mammary features in appendages that would otherwise become hair follicles [31]. These studies support the concept that Gli3R specifies mammary fate by blocking Hedgehog-mediated hair follicle differentiation. Of note, the site of Gli3 activity is controversial. Veltmaat et al. focused attention on Gli3 mRNA expression in somites [23]. However, Gli3 protein was detected by immunohistochemistry in mammary buds and surrounding stroma [28]. Of note, Gli3 expression is not informative about the location and timing of its activity as this depends on post-translational processing. However, this can be determined by examining expression of Gli1-lacZ, a reporter of Hedgehog pathway activity on Gli2 and Gli3 mutant backgrounds. This approach provided compelling genetic evidence that the critical site of Gli3 repression lies within the mammary mesenchyme [28]. The direct gene targets of Gli3 repression at this site remain to be determined. However, molecular analysis of buds that remain evaginated in Gli2$^{iki/iki}$ and Gli3$^{xt/xt}$ (Fig. 5B) mice have demonstrated that the downstream consequences of Hedgehog pathway mis-activation involves failure to sustain markers of mammary mesenchyme specification [28,29].

4.2. *Tbx3 is required for placodal induction and maintenance of MR#1,3,4,5*

As discussed in Section 3.1, mesenchymal expression of Tbx3 mRNA precedes mammary line formation. Tbx3 mRNA is the earliest marker to become concentrated in the ectodermal placode ~E10.5 (Fig. 2B) [15,19,20]. In Tbx3 null mice, Wnt10b and Lef1 fail to accumulate, FGF signaling is abrogated, and placodes 1,3,4,5 do not form. Thus, Tbx3 acts downstream of Gli3R and upstream of Wnt and FGF signaling suggesting a key role in placodal induction. However, it is, in turn, augmented by these pathways in a positive feedback loop. Tbx3 levels are critical, as demonstrated by the finding that Tbx3 haploinsufficiency also leads to frequent aplasia of MR#1–3 ~E13.5 and reduced ductal branching in retained glands [15]. Thus, Tbx3 is required not only to position the mammary line and for placodal induction, but its level is critical for placodal maintenance. In other tissues, Tbx3 has been implicated in regulating cell proliferation, however as proliferation is not a prominent driver of placodal formation it is more likely that Tbx3′s ability to repress E-cadherin and promote migration are more relevant at this stage [15]. Of note, *TBX3* is highly relevant for human mammary development. C-terminal mutations in *TBX3,* which result in protein instability and impaired ability to repress target genes, are found in Ulnar-mammary-syndrome (UMS) (Table 1) [32]. Patients with this disorder show a pleiotropic phenotype resulting from defective patterning. Upper limb deficiencies range from missing digits and duplicated nail surfaces to complete absence of the forearm and hand. Their mammary phenotypes range from complete loss to hypoplasia of

Table 2
Effects of loss and gain of function of key regulators of embryonic mammary development in mice.

Genes	Mouse model	Phenotype	Ref.
Gli2A	*Gli3$^{xt/xt}$*	Loss MR3, MR5	[28]
Gli3R	*Gli2$^{1nki/1nki}$; Gli3$^{xt/+}$*	Impaired mm specification MR2, MR4: Bud evagination Inappropriate hair follicle formation Loss of sexual dimorphism	[29]
Tbx3	*Tbx3$^{-/-}$*	Loss MR1,3,4,5	[19]
	Tbx3$^{-/+}$	Loss MR1-3 hypoplastic branching in others	[20]
Nrg3	*Nrg3^{ska}/Nrg3^{ska}*	MR3 loss	[36]
	K14-Nrg3	MR4 duplication (more in adults)	
Fgfr2b	*Fgfr2b$^{-/-}$*	Loss of all MR	[22]
Fgf10	*Fgf10$^{-/-}$*	Loss of MR1,2,3,5	[22]
Eda	*K14-Eda*	Supernumerary between MR3-4 and in neck	[44, 48]
Pthlh	*Pthlh$^{-/-}$*	Arrest at late bud stage	[56,
Pth1r	*Pth1r$^{-/-}$*	Impaired mm specification Loss of sexual dimorphism	57]
Dkk1	*K14-PthrP*	Gain of mm differentiation and loss of hair follicles in ventral epidermis develops nipple characteristics	[17, 59]
	K5-rtTA; tetO-Dkk1	No placodes	[86]
Sostdc1	*Sostdc1$^{-/-}$*	Enlarged placodes, MR2 and 3 fused	[49]
	K14-rTA; tetO-WISE	Small MR	[87]
Lpr4	*Lrp4$^{mdit/mdit}$ Lpr4$^{mitt/mitt}$*	Fusion of MR2 and 3	[87]
	Lpr4$^{mdit/mdit}$; Lpr5$^{-/-}$	Rescued buds 2 and 3 fusion	
	Lpr4$^{mdit/mdit}$; Lpr6$^{-/-}$		
	Lpr4$^{mdit/mdit}$; Lpr5+/Lpr6$^{+/-}$		
	Lpr4$^{mdit/mitt}$; K14cre; β-catenin$^{fx/-}$	Small placodes, no bud downgrowth Inhibit ectopic Wnt signaling between MR2 and 3	
Lpr5	*Lpr5$^{-/-}$*	Small placodes-slow development	[68]
Lpr6	*Lpr6$^{-/-}$*	Impaired ductal outgrowth and adipogenesis	[67]
	MMTV-Wnt1; Lpr6$^{-/-}$	Rescue MR	
β-catenin	*K14cre; β-catenin$^{fx/-}$*	Small MR	[86]
	Dermo1-cre; β-catenin$^{lox/lox}$	Impaired mm specification Loss of sexual dimorphism	[64]
Lef1	*Lef1$^{-/-}$*	Loss MR2, MR3 Arrest MR1,4,5 at bud stage and demise	[18]
Pygo2	Pygo2$^{-/-}$	E18.5 MR loss or impaired elongation and branching	[69]

breasts and associated axillary hair, sweat glands. Nipple loss and presence of supernumerary and inverted nipples have been documented and defective lactation has been reported. Other features include dental, cardiac and genital abnormalities [32].

4.3. *Fgf10/Fgfr2b vertically augment placodal Wnt10b*

Fibroblast growth factor 10 (Fgf10) is essential for MR#1,2,3,5 and its receptor, Fgfr2b, is required to maintain all MR [22]. Fgf10 mRNA expression has been documented in the dermomyotome of the somites ~E10.5 onwards has been proposed to activate its receptor, Fgfr2b, which is expressed in the ectoderm [23]. Problems with this model include the facts that insolubility of Fgf10 make it unlikely to diffuse over such a distance, and classical transplantation studies have shown that mesenchyme alone can induce mammary fate within ectoderm independent of instruction from underlying tissues. Delamination and migration of cells expressing Fgf10 has been proposed as a solution [23]. Somitic Fgf10 expression is unaffected by Wnt inhibitors, whereas Fgf10 hypomorphs fail to express TOP-gal and Wnt10b, placing Fgf10 signals

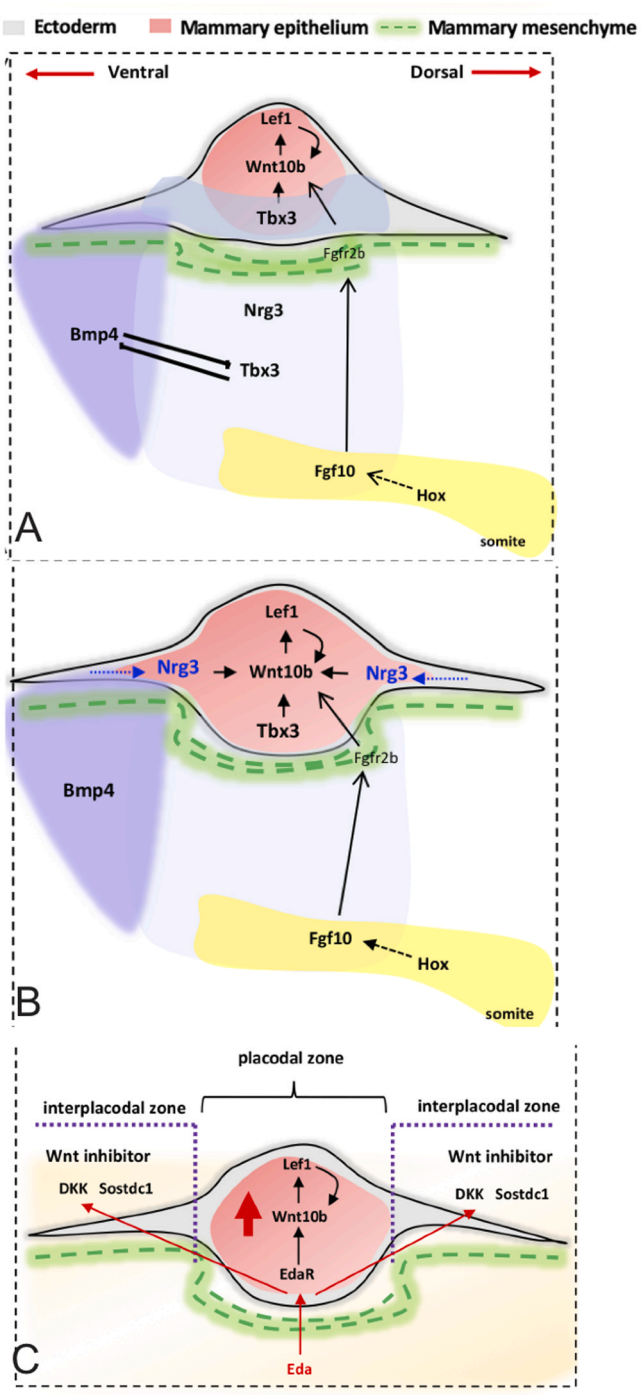


Fig. 2. Schematic of signaling pathways regulating mammary development. (A) The mammary line is specified by antagonism between dorsal Tbx3 and ventral Bmp4. Cells located at somite tips expressing Fgf10 are proposed to translocate towards the ectoderm where they activate their receptor, Fgfr2b. Tbx3, Wnt10b and Lef1 become focally upregulated along the mammary line. (B) Placode formation is induced by migration of cells. Nrg3 and its receptor restricted within the mammary epithelial compartments are proposed to induce migration of epithelial progenitor cells into the placode. Tbx3, Wnt10b and Fgf10 and FgfR are essential for the formation of most placodes. Tbx3 and Wnt10b must be maintained at a critical threshold during this period. (C) Eda/Edar signaling defines placodal and interplacodal regions by simultaneously stimulating expression of short-range activators (e.g. Wnt10b and PTHrP) and long-range inhibitors (e.g. DKK, Sostdc1) of mammary fate.

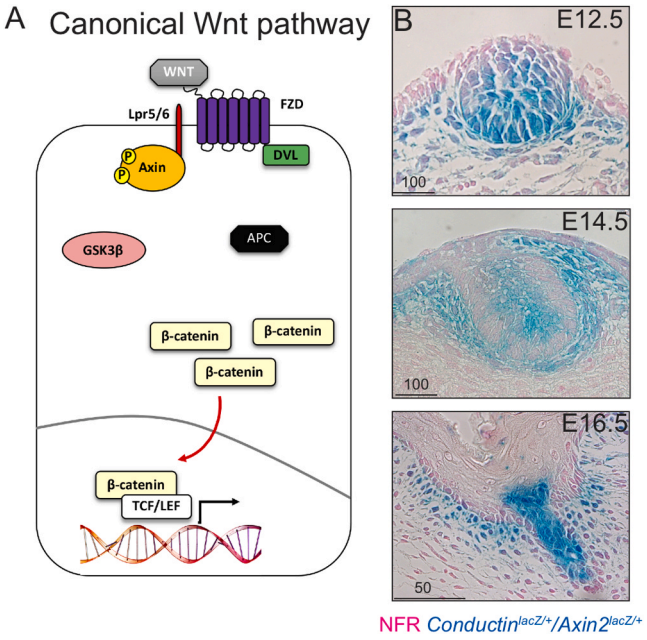


Fig. 3. Wnt signaling is important at every stage of embryonic mammary development (A) Schematic showing the key elements involved in activating canonical Wnt signaling, which plays critical roles at every stage of embryonic mammary development (see Table 2 for mammary phenotypes associated with this pathway). (B–D) Sections of embryos counterstained with NFR (pink) and stained with X-gal to locate expression of the Wnt signaling reporter, *Conductin*$^{lacZ/+}$/*Axin2*$^{lacZ/+}$ (blue). Note that Wnt signaling is activated both within the epithelial compartments of the placode ~E12.5, the mammary bud ~E14.5, and the proliferating mammary sprout ~E16.5 as well as in the surrounding mammary mesenchyme at all stages. Wnt signaling sustains placodal identity, plays a critical role in specifying the mammary mesenchyme, is important for ductal proliferation and elongation and has been shown to regulate stem cell potency *in vivo* and *in vitro* (see Table 2 for mammary phenotypes associated with Wnt signaling) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

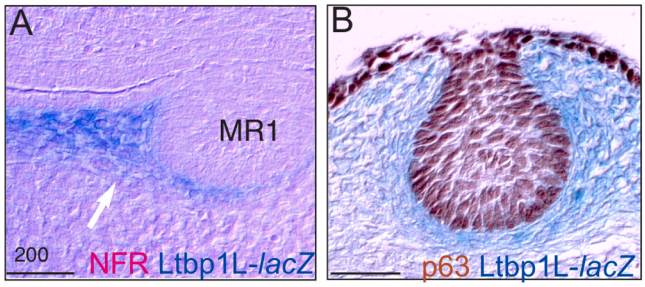


Fig. 4. Ltbp1 is an early marker of mammary mesenchyme. Sections of embryos and X-gal stained to locate Ltbp1L-*lacZ* reporter expression (blue) in (A) an axillary streak of mesenchymal cells (white arrow) oriented towards mammary rudiment 1 (MR1) counterstained with NFR (pink) ~E12. Scale bars 200 μm; and in (B) the mammary mesenchyme surrounding ~E14.5 bud stained with antibodies to p63 (brown) Scale bars, 100 μm (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).
From Chandramouli et al. [26,29].

upstream of Wnt10b and mammary line stratification. Of note, TOP-gal and Wnt10b expression aligns with the segmentation pattern of the thoracic somites and Pax3 mutants, which fail to extend the thoracic somites show dorsalization of the mammary line [23]. Moreover, humans with Poland syndrome, which is characterized by hypoplasia of structures derived from thoracic somites also affects breasts (Table 1)

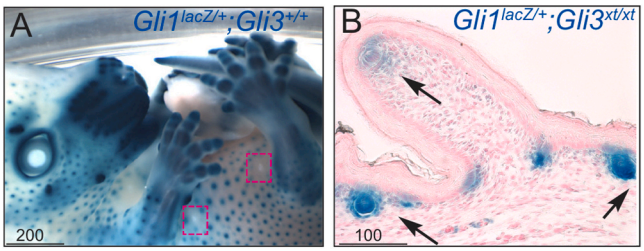


Fig. 5. Differential activation of Hedgehog signaling in ectodermal appendages.
(A) Hedgehog signaling detected by expression of the Gli1 $^{lacZ/+}$ reporter (blue) is activated in hair follicle placodes but remains repressed in mammary buds (lack of blue staining in pink boxed area). (B) In addition to bud loss, (*Gli3*$^{xt/xt}$) mutants that lack Gli repressor activity display evagination of the remaining buds and inappropriate formation of hair follicles within the mammary field (arrows) that express the Gli1-lacZ reporter of pathway activity. This data suggests that Gli3 repression of Hedgehog signaling is required for mammary fate by repressing hair follicle fate (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).
Chandramouli et al. [28,29].

[33]. Homeobox genes have recently been proposed to act upstream of Fgf10. Hoxc8, for example, although normally expressed in the E10.5 ectoderm, when mis-expressed induces somitic Fgf10 and ectopic placodes [34].

4.4. *Ngr3 stimulates aggregation of mammary progenitors into MR#3*

Neuregulin 3 (Nrg3) presents another example of a growth factor required for embryonic mammary development [24]. Nrg3 is a ligand of receptor tyrosine-protein-kinase ErbB4, implicated in chemoattraction, directed cell migration, adhesion, proliferation, and stratification [24]. Although Nrg3 appears ~E11 within the dermal mesenchyme underneath MR#3, by ~E12 is restricted to the epithelial placode together with its receptor, suggesting that it may operate in an autocrine manner [35]. Its requirement for embryonic mammary development became evident from studies of *Nrg3*^{ska} (scaramanga) mutants, which express reduced levels of Nrg3 and lack MR#3 (Table 2) [36]. By contrast, subsequent experiments showed that mice overexpressing K14-Nrg3 or those exposed to beads coated in recombinant Nrg3 produced a supernumerary MR between MR#3 and MR#4 [37]. Collectively these experiments show Nrg3 promotes mammary cell fate. Howard and colleagues elegantly visualized the effect of Nrg3 on progenitor recruitment into placodes by crossing *Nrg3*^{ska} mice to mice expressing s-SHIP-GFP, a marker of mammary progenitors [25]. They found that s-SHIP-GFP positive progenitors remained dorsally dispersed in *Nrg3*^{ska} mutants and the few cells that were present in *Nrg3*^{ska} hypoplastic placodes failed to adopt placodal cell arrangement and shape characteristics. These data provide strong evidence that Nrg3 regulates progenitor aggregation into, and rearrangement within, placodes (Fig. 2). Intriguingly K14-Nrg3 mice develop additional ectopic glands as adults suggesting that Nrg3 may continue to prime and/or recruit cells and is sufficient to induce de novo appendage development [37]. Of note, *Nrg3* mutants show changes in Tenascin C, a matrix protein associated with stem cell niches and a ligand for integrins found on mammary progenitors.

4.5. *Eda/Edar patterns placodal and inter-placodal domains by sculpting Wnt signaling*

A further route that serves to regulate placodal definition and spacing is the ectodysplasin (Eda) signaling pathway. Eda, a tumor necrosis factor-like ligand, binds to its receptor, Edar, and signals through an adaptor protein, Edaradd, to activate NF-κB [38,39]. This pathway is an important regulator of mammary gland development in humans

where mutations in *EDA*, *EDAR* or *EDARADD,* as well as in *IKK, the* inhibitor of NF-κB, cause hypohidrotic ectodermal dysplasia (HED) (Table 1) [40–44]. This condition involves defects in hair, teeth, sweat, salivary and mammary glands. The mammary phenotypes include loss and gain (polythelia) of nipples, compromised ability to breast feed and occasional loss of breasts altogether. Although Eda is dispensable in mice, for endogenous MR formation, K14-Eda expression, induces 2–3 extra MR along the central mammary line and others in the neck [45–48]. This suggests that Eda acts downstream of mammary line specification to regulate mammary placodal fate along this extended arc. Microarray analysis revealed that both Wnt signaling activators (Wnt10a,10b) and inhibitors (DKK1, Sostdc1) are transcriptional targets of Eda signaling and suppression of canonical Wnt signaling was found to reduce supernumerary placode induction by K14-Eda in a dose-dependent manner [45,46,49,50]. Eda is proposed to pattern placodes by stimulating expression of a short-range insoluble activator (e.g. Wnt ligands) and defines interplacodal regions, by simultaneously stimulating more diffusible Wnt inhibitors (e.g. DKK) (Fig. 2C) [45]. A similar reaction-diffusion mechanism has been proposed to pattern intestinal crypts [51,52]. This model provides an elegant explanation for how varying Eda expression can produce both loss and gain of MR. Changes in Eda/Edar signaling have been linked to adaptive changes in other skin appendages, for example sweat gland density in modern human populations, and it has been hypothesized that, in addition to causing HED, Eda/Edar may also be responsible for the prevalence of polythelia in humans [53]. Moreover, in addition to its importance for human appendage development Mikkola and colleagues have proposed that Eda patterning could provide the molecular mechanism responsible for evolutionary diversity in the number and position of mammary glands in different species [54]. In addition to modulating Wnt signaling, Eda has been shown to modulate expression of several other key factors including Fgf20 and parathyroid hormone-related peptide (PTHrP) and thus may have a wide repertoire of ways to modulate mammary development [46,55].

4.6. PTHrP: the master regulator of mammary mesenchyme differentiation

Once placodes have been defined the next step involves their rearrangement into hillocks ~E12 and invagination to form a flask shaped bud (Fig. 6). The mature bud comprises a sphere of concentrically oriented epithelial cells connected to the skin surface by a stalk surrounded by layers of condensed, elongated mesenchymal cells. In males, androgens cause the mesenchyme to enlarge, constricting the stalk and leading to severance and apoptosis [56]. In females, buds are quiescent at E14, become specified ~E15 then begin to proliferate and sprout towards the fat pad ~E16. PTHrP was discovered as a cause of humoral hypercalcemia of malignancy that likely arises due to misactivation of its physiological role in mobilizing skeletal calcium stores during lactation. During embryogenesis, however, PTHrP functions locally, and is considered to be the master epithelial inducer of mesenchymal differentiation. As such, it plays a critical role in mammary fate commitment. PTHrP is expressed within the mammary line and placodes but is not required until ~E15 when PTHrP it is secreted by the epithelial bud and acts in a paracrine fashion to activate the G-protein-coupled receptor, parathyroid hormone 1 receptor (PTHR1), within the surrounding mesenchyme [57,58]. Eda-Edar-NFκB signaling upregulates PTHrP mRNA within embryonic mammary buds [46,55]. Disruption of PTHrP or its receptor impairs mammary development in multiple ways. Buds form, but the mesenchymal cells that condense around them lack all hallmarks of mammary mesenchyme specification (Androgen Receptor (AR), Estrogen Receptor (ER), Lef1, Cadherin 11 and Tenascin C) [56,57,59]. The absence of AR prevents the mesenchyme from responding to androgens and so male buds are retained [56]. The undifferentiated immature mesenchyme fails to send reciprocal differentiative signals and, as a result, female buds lose their mammary

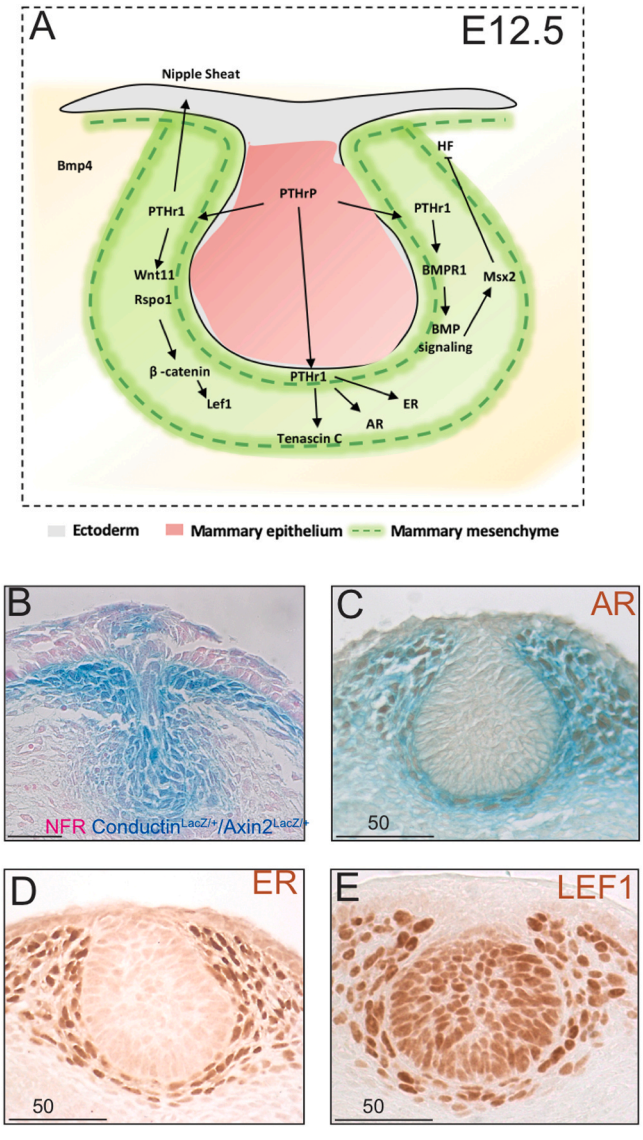


Fig. 6. PTHrP specifies the mammary mesenchyme. (A) PTHrP secreted from the mammary epithelium (pink) activates its receptor PTHR1 within the surrounding mammary mesenchyme (green). This acts through Wnt and BmpR1 pathways to induce androgen receptor (AR), which in males activates apoptosis and mesenchymal constriction as seen in (B). (C–E) Expression of AR, (D) estrogen receptor (ER), (E) Lef1 detected by immunohistochemistry. Mammary mesenchymal factors AR, ER, Lef1 and Tenascin C signal to the mammary bud and overlying epithelium to stimulate nipple differentiation and ductal sprouting in females (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

identity and fail to sprout, and the presumptive nipple reverts to an epidermal fate [59,60]. A converse phenotype occurs in response to PTHrP overexpression: ectopic mammary mesenchymal markers become expressed resulting in suppression of hair follicles and conversion of ventral epidermis to nipple-sheath. The importance of this pathway for human mammary development is demonstrated by human fetuses with Blomstrand chondrodysplasia (Table 1) that have homozygous null mutations in the gene for PTHR1 and lack breast duct development similar to *PTHrP*$^{-/-}$ mice [61].

PTHrP mediated epithelial-mesenchymal cross talk regulates the expression of many different molecules. Among these, BmpR1, Lef1, β-catenin and Msx2 have received the most attention [62,63]. β-catenin plays an essential role downstream of PTHrP signaling in specifying the mammary mesenchyme [64]. Loss of PTHrP inhibits mesenchymal Lef1

and β-catenin expression whereas overexpression of PTHrP upregulates them ectopically in the ventral mesenchyme [56,59]. Deletion of mesenchymal β-catenin phenocopies loss of PTHrP and blocks the ectopic induction of mammary mesenchymal markers by K14-PTHrP [64]. Further experiments demonstrated that PTHrP requires Wnt/β-catenin within the mesenchyme to induce mammary mesenchyme specification. While the mechanism by which PTHrP activates mesenchymal Wnt signaling remain obscure it has been proposed to likely result from induction of Wnt11 and R-Spondin 1 (Rspo1) [62,64]. Wnt signaling is active in both epithelial and mesenchymal compartments at this stage as demonstrated by the *Conductin*$^{lacZ/+}$/*Axin2*$^{lacZ/+}$ reporter (Fig. 3). Intriguingly, as discussed above, very similar phenotypes involving male bud retention, failure to maintain mammary mesenchyme markers, and loss of ductal downgrowth, are produced by misactivation of Hedgehog signaling (*Gli3*$^{xt/xt}$ and *Gli2*$^{1ki/1ki}$) [29]. In this case, however the phenotype is accompanied by loss of epithelial Wnt signaling [29]. While further work is required to decipher the interconnections, analyses on Gli3R, Eda and PTHrP suggest that Wnt signaling lies upstream and downstream of PTHrP in different compartments. Additional downstream effectors of PTHrP signaling include BMPs and Msx2 [65]. Addition of Bmp4 to cultured *PTHrP*$^{-/-}$ mammary buds has been shown to rescue their arrested sprouting. PTHrP induction of Msx2 is proposed to suppress hair follicles formation in the nipple region [65]. Supporting this concept, loss of hair follicles seen in ventral skin in K14-PTHrP mice is rescued by Msx2 deletion [65,66].

5. Ductal morphogenesis, branching and lumen formation: Wnts, Fgf10, Eda and Tgfβ pathways

Once the mammary rudiment has been specified, the last stages involve sprouting, branching and lumen formation. The onset of proliferation ~E15.5 coincides with resumption of Wnt reporter activity within the mammary sprout [17]. A requirement for Wnt activity at this stage is demonstrated by the fact that ablation of Wnt co-receptor, Lrp6, or pygopus, a canonical Wnt modulator delays and stunts outgrowth (Table 2) [67–69]. Branching begins ~E16. Deficiency in Eda, Edar, or NF-κB result in smaller ductal trees with less branches, and the converse is seen in mice overexpressing Eda and Edar [35,39]. Fgf10 is also implicated in ductal branching, Fgf10$^{-/-}$ glands fail to ramify and remain as sprouts, however FGF also promotes adipogenesis in the fat pad via CCAAT-enhancer-binding proteins beta (CEBPβ) [70]. Lumen formation in the embryonic mammary gland is thought to involve both apoptosis of cells at the center of the solid ductal chords as well as changes in cell adhesion. Of note Ltbp1 is one of the earliest reliable markers of luminal cell specification being expressed in cells surrounding micro lumen in the embryonic mammary tree (Fig. 1) and is a highly specific marker of ductal luminal cells in the adult gland [26]. This expression pattern suggests involvement of Tgfβ pathways in generating polarity and cavitation.

6. Embryonic mammary stem cells and progenitors

Mammary stem cells (MaSCs) are specified during embryonic development. Classical transplantation studies demonstrated that whole buds from as early as E13 could successfully regenerate a mammary tree in adult fat pads [71]. In contrast, mammary epithelial cells isolated from embryonic mammary rudiments did not acquire this capacity until E15.5 [7]. These experiments suggest that the mammary identity of epithelial cells is dependent upon signals from the mammary mesenchyme until E15.5. This emphasizes the critical importance of the early mammary mesenchyme not only as an inducer of mammary morphogenesis but also as a stem cell niche. By performing transplantations at limiting dilution Spike et al. found the frequency of mammary repopulating units (MRU) increased significantly from E15.5 onwards reaching a peak ~E18.5 and further showed that repopulating capacity was restricted to a cell population, which they termed fetal MaSCs (fMaSCs), that displayed CD24 and CD49f levels intermediate to those of adult luminal and basal populations [7]. A caveat to the use of the term fMaSC and to the interpretation of these results is that regeneration assays measure cell plasticity and acquisition of stem cell potency in an artificial situation. As such, MRU and fMaSCs are considered facultative stem cells [72] and lineage tracing studies in many tissues have shown that facultative stem cells differ in their behavior and potency from physiological stem cells [73,74]. Several other markers of adult cell populations with enhanced repopulating capacity are expressed in embryonic buds. For example, s-SHIP is expressed strongly at E13.5 and, as discussed above, has been utilized to follow cells migrating into mammary placodes [25,75,76]. As might be predicted from the requirement for Wnt signaling at all stages of mammary development, several Wnt target genes, such as Lgr5 and Axin2, are expressed at this time and have been used to isolate and characterize embryonic mammary progenitor populations [5,77,78]. Gene expression profiling of (Lgr5+CD49f^{hi}) embryonic progenitor cells derived from E14 buds uncovered enrichment for Wnt, Edar, Pth, Tgfβ, and Notch pathways, consistent with their critical roles *in vivo* as described throughout this review [5].

Lineage tracing studies have shown that K14-expressing cells labeled at E13 and E17 give rise to both basal and luminal lineages in adult mice, and thus at the population level are bipotent, whereas after birth K14 tracing is restricted to the basal lineage [5,73]. The Beronja group provided support for bipotency at a single cell level by showing that early embryonic ectodermal cells barcoded at E9 by intra-amniotic lentiviral injection give rise to both basal and luminal mammary lineages [79]. However, several lineage tracing experiments have demonstrated that unipotent lineage primed progenitors emerge prior to birth [5,77,80]. Using a multicolor lineage tracing approach with a Notch1-creERT2 mouse line, Lilja et al., showed that early embryonic Notch1-positive cells, which express both luminal and basal markers, give rise to both basal and luminal cells. However, by mathematical modeling they were able to demonstrate that, despite their undifferentiated phenotype, individual Notch1 cells are already lineage-primed unipotent progenitors at E12.5 [80]. Lineage restriction in later stages of embryogenesis has also been assessed by neutral lineage tracing approach using a R26-CreERT2;R26-Confetti mouse line, which avoids the bias inherent to the use of gene promoters. By this approach Lloyd-Lewis et al. found evidence that embryonic mammary progenitors had already become lineage restricted at E16.5-E17.5 [81].

scRNAseq studies have complemented findings derived from lineage tracing experiments [5]. Wuidart et al. also showed that at E14.5 Lgr5/CD49f^{hi} embryonic progenitors express a hybrid signature of both basal and luminal genes and similar results were found analyzing EpCAM+ populations [9]. However, by E17, as microlumen appear and branching occurs, evidence of lineage segregation appears [5,9]. Cells in the inner layer upregulate K8 expression [5] and Ltbp1 appears strictly in cells surrounding microlumen (Fig. 1) [26]. In contrast cells in the outer layer express higher p63 levels [5]. Three studies have addressed the mechanism by which cells become lineage primed. Lilja et al. showed that when Notch1 was expressed ectopically in basal cells it could switch them to luminal ER alpha negative cell fate [80]. Focusing on transcriptional regulons that might be responsible for directing lineage restriction Wuidart et al. demonstrated that ectopic expression of p63 in adult luminal cells could promote a basal-like state [5]. Collectively, these studies show that although very early embryonic mammary populations are multipotent, lineage priming at the individual cell level occurs early during embryonic mammary development and identify key roles for Notch1 and p63 in driving luminal or basal fate [5,80]. Moreover, they suggest a plasticity and dynamic behavior for these embryonic stem cells. Recent studies have investigated the epigenetic landscape over the course of late embryonic mammary development found that basal-like and luminal-like chromatin changes similar to those of adult cell types emerge ~E18 cells concluding that perinatal cells are primed and poised for specific routes of differentiation [82].

7. Conclusion

In summary, repression of Hedgehog signaling plays a central role in appendage fate and is the critical point of divergence between hair and mammary development. Mesenchymal patterning by Bmp4 and Tbx3 designates the mammary zone, which is also influenced by Hox gene expression in somites and limbs. Nrg3 promotes cell migration into placodes and Eda defines their radius and spacing. All pathways intersect and reinforce one another in complex site-specific ways that remain obscure, to promote critical thresholds of placodal Tbx3 and Wnts, which are required to sustain mammary fate of undifferentiated embryonic mammary cells expressing hybrid bilineage markers. The mammary mesenchyme is a critical stem cell niche. PTHrP induces mammary mesenchymal differentiation, which acts via Wnt and BMP signaling to unleash a suite of factors that instruct the bud and epithelium to commit to mammary and nipple identity. Subsequent actions of Eda, FGF and Tgfβ, promote lumen formation and ductal branching and are linked temporally to progressive restriction in progenitor potency driven by Notch, p63 and epigenetic modification. Thus, embryonic patterning pathways play two key roles regulating mammary progenitor potency and cell migration. These features frequently become revived in the context of breast cancer as evidenced by the preponderance of embryonic stem/progenitor signatures in breast tumors and reactivation of cell movement leading to metastasis [5–11]. Greater understanding of how these pathways collaborate to regulate embryonic mammary development is necessary to advance in our ability to thwart their collusion in the pathological setting. Classical studies on appendage development pointed to the significant role of epithelial and mesenchymal reciprocity [83]. Recent studies are now revealing a glimmer of hope that it may be possible to control some types of cancer by rebalancing signals between these compartments to restore physiological differentiation [84,85].

Conflict of interest statement

The authors declare that no conflict of interest exists.

Acknowledgments

This article was supported by Department of Defense USAMRMC W81XWH-17-1-0013 BC160959 and W81XWH-13-1-0290 BC123572 (PC).

References

[1] O.T. Oftedal, The mammary gland and its origin during synapsid evolution, J. Mammary Gland Biol. Neoplasia 7 (3) (2002) 225–252.
[2] O.T. Oftedal, D. Dhouailly, Evo-devo of the mammary gland, J. Mammary Gland Biol. Neoplasia 18 (2) (2013) 105–120.
[3] J.J. Wysolmerski, The evolutionary origins of maternal calcium and bone metabolism during lactation, J. Mammary Gland Biol. Neoplasia 7 (3) (2002) 267–276.
[4] K. Hughes, Comparative mammary gland postnatal development and tumourigenesis in the sheep, cow, cat and rabbit: Exploring the menagerie, Semin. Cell Dev. Biol. (2020).
[5] A. Wuidart, A. Sifrim, M. Fioramonti, S. Matsumura, A. Brisebarre, D. Brown, A. Centonze, A. Dannau, C. Dubois, A. Van Keymeulen, T. Voet, C. Blanpain, Early lineage segregation of multipotent embryonic mammary gland progenitors, Nat. Cell Biol. 20 (6) (2018) 666–676.
[6] A. Van Keymeulen, M.Y. Lee, M. Ousset, S. Brohée, S. Rorive, R.R. Giraddi, A. Wuidart, G. Bouvencourt, C. Dubois, I. Salmon, C. Sotiriou, W.A. Phillips, C. Blanpain, Reactivation of multipotency by oncogenic PIK3CA induces breast tumour heterogeneity, Nature 525 (7567) (2015) 119–123.
[7] B.T. Spike, D.D. Engle, J.C. Lin, S.K. Cheung, J. La, G.M. Wahl, A mammary stem cell population identified and characterized in late embryogenesis reveals similarities to human breast cancer, Cell Stem Cell 10 (2) (2012) 183–197.
[8] R. Eyre, D.G. Alférez, A. Santiago-Gómez, K. Spence, J.C. McConnell, C. Hart, B. M. Simões, D. Lefley, C. Tulotta, J. Storer, A. Gurney, N. Clarke, M. Brown, S. J. Howell, A.H. Sims, G. Farnie, P.D. Ottewell, R.B. Clarke, Microenvironmental IL1β promotes breast cancer metastatic colonisation in the bone via activation of Wnt signalling! Abstract, Nat. Commun. 10 (1) (2019) 5016.
[9] R.R. Giraddi, C.Y. Chung, R.E. Heinz, O. Balcioglu, M. Novotny, C.L. Trejo, C. Dravis, B.M. Hagos, E.M. Mehrabad, L.W. Rodewald, J.Y. Hwang, C. Fan, R. Lasken, K.E. Varley, C.M. Perou, G.M. Wahl, B.T. Spike, Single-cell transcriptomes distinguish stem cell state changes and lineage specification programs in early mammary gland development, Cell Rep. 24 (6) (2018) 1653–1666.e7.
[10] C. Dravis, C.Y. Chung, N.K. Lytle, J. Herrera-Valdez, G. Luna, C.L. Trejo, T. Reya, G. M. Wahl, Epigenetic and transcriptomic profiling of mammary gland development and tumor models disclose regulators of cell state plasticity, Cancer Cell 34 (3) (2018) 466–482 e6.
[11] C. Dravis, B.T. Spike, J.C. Harrell, C. Johns, C.L. Trejo, E.M. Southard-Smith, C. M. Perou, G.M. Wahl, Sox10 regulates stem/progenitor and mesenchymal cell states in mammary epithelial cells, Cell Rep. 12 (12) (2015) 2035–2048.
[12] K.W. Cho, J.Y. Kim, S.J. Song, E. Farrell, M.C. Eblaghie, H.J. Kim, C. Tickle, H. S. Jung, Molecular interactions between Tbx3 and Bmp4 and a model for dorsoventral positioning of mammary gland development, Proc. Natl. Acad. Sci. U. S. A. 103 (45) (2006) 16788–16793.
[13] A.Y. Propper, Wandering epithelial cells in the rabbit embryo milk line: a preliminary scanning electron microscope study, Dev. Biol. 67 (1) (1978) 225–231.
[14] M.Y. Lee, V. Racine, P. Jagadpramana, L. Sun, W. Yu, T. Du, B. Spencer-Dene, N. Rubin, L. Le, D. Ndiaye, S. Bellusci, K. Kratochwil, J.M. Veltmaat, Ectodermal influx and cell hypertrophy provide early growth for all murine mammary rudiments, and are differentially regulated among them by Gli3, PLoS One 6 (10) (2011), e26242.
[15] L.A. Jerome-Majewska, G.P. Jenkins, E. Ernstoff, F. Zindy, C.J. Sherr, V. E. Papaioannou, Tbx3, the ulnar-mammary syndrome gene, and Tbx2 interact in mammary gland development through a p19Arf/p53-independent pathway, Dev. Dyn. 234 (4) (2005) 922–933.
[16] J.M. Veltmaat, W. Van Veelen, J.P. Thiery, S. Bellusci, Identification of the mammary line in mouse by Wnt10b expression, Dev. Dyn. 229 (2) (2004) 349–356.
[17] E.Y. Chu, Canonical WNT signaling promotes mammary placode development and is essential for initiation of mammary gland morphogenesis, Development 131 (19) (2004) 4819–4829.
[18] C. van Genderen, R.M. Okamura, I. Farinas, R.G. Quo, T.G. Parslow, L. Bruhn, R. Grosschedl, Development of several organs that require inductive epithelial-mesenchymal interactions is impaired in Lef-1 deficient mice, Genes Dev. 8 (1994) 2691–2704.
[19] T.G. Davenport, L.A. Jerome-Majewska, V.E. Papaioannou, Mammary gland, limb and yolk sac defects in mice lacking Tbx3, the gene mutated in human ulnar mammary syndrome, Development 130 (10) (2003) 2263–2273.
[20] N.C. Douglas, V.E. Papaioannou, The T-box transcription factors TBX2 and TBX3 in mammary gland development and breast cancer, J. Mammary Gland Biol. Neoplasia 18 (2) (2013) 143–147.
[21] A.Y. Propper, B.A. Howard, J.M. Veltmaat, Prenatal morphogenesis of mammary glands in mouse and rabbit, J. Mammary Gland Biol. Neoplasia 18 (2) (2013) 93–104.
[22] A.A. Mailleux, B. Spencer-Dene, C. Dillon, D. Ndiaye, C. Savona-Baron, N. Itoh, S. Kato, C. Dickson, J.P. Thiery, S. Bellusci, Role of FGF10/FGFR2b signaling during mammary gland development in the mouse embryo, Development 129 (1) (2002) 53–60.
[23] J.M. Veltmaat, F. Relaix, L.T. Le, K. Kratochwil, F.G. Sala, W. van Veelen, R. Rice, B. Spencer-Dene, A.A. Mailleux, D.P. Rice, J.P. Thiery, S. Bellusci, Gli3-mediated somitic Fgf10 expression gradients are required for the induction and patterning of mammary epithelium along the embryonic axes, Development 133 (12) (2006) 2325–2335.
[24] N. Kogata, M. Zvelebil, B.A. Howard, Neuregulin 3 and erbb signalling networks in embryonic mammary gland development, J. Mammary Gland Biol. Neoplasia 18 (2) (2013) 149–154.
[25] N. Kogata, E. Oliemuller, O. Wansbury, B.A. Howard, Neuregulin-3 regulates epithelial progenitor cell positioning and specifies mammary phenotype, Stem Cells Dev. 23 (22) (2014) 2758–2770.
[26] A. Chandramouli, J. Simundza, A. Pinderhughes, M. Hiremath, G. Droguett, D. Frendewey, P. Cowin, Ltbp1L is focally induced in embryonic mammary mesenchyme, demarcates the ductal luminal lineage and is upregulated during involution, Breast Cancer Res. 15 (6) (2013) R111.
[27] A. Chandramouli, J. Simundza, A. Pinderhughes, P. Cowin, Choreographing metastasis to the tune of LTBP, J. Mammary Gland Biol. Neoplasia 16 (2) (2011) 67–80.
[28] S.J. Hatsell, P. Cowin, Gli3-mediated repression of Hedgehog targets is required for normal mammary development, Development 133 (18) (2006) 3661–3670.
[29] A. Chandramouli, S.J. Hatsell, A. Pinderhughes, L. Koetz, P. Cowin, Gli activity is critical at multiple stages of embryonic mammary and nipple development, PLoS One 8 (11) (2013), e79845.
[30] S. Kawai, T. Sugiura, Characterization of human bone morphogenetic protein (BMP)-4 and -7 gene promoters: activation of BMP promoters by Gli, a sonic Hedgehog mediator, Bone 29 (1) (2001) 54–61.
[31] A. Gritli-Linde, K. Hallberg, B.D. Harfe, A. Reyahi, M. Kannius-Janson, J. Nilsson, M.T. Cobourne, P.T. Sharpe, A.P. McMahon, A. Linde, Abnormal hair development and apparent follicular transformation to mammary gland in the absence of hedgehog signaling, Dev. Cell 12 (1) (2007) 99–112.
[32] M. Bamshad, R.C. Lin, D.J. Law, W.S. Watkins, P.A. Krakowiak, M.E. Moore, P. Franceschini, R. Lala, L.B. Holmes, T.C. Gebuhr, B.G. Bruneau, A. Schinzel, J. G. Seidman, C.E. Seidman, L.B. Jorde, Mutations in human TBX3 alter limb, apocrine and genital development in ulnar-mammary syndrome, Nat. Genet. 16 (3) (1997) 311–315.

[33] A. Baban, M. Torre, S. Bianca, A. Buluggiu, M.I. Rossello, M.G. Calevo, M. Valle, R. Ravazzolo, V. Jasonni, M. Lerone, Poland syndrome with bilateral features: case description with review of the literature, Am. J. Med Genet A 149A (7) (2009) 1597–1602.
[34] L.S. Carroll, M.R. Capecchi, Hoxc8 initiates an ectopic mammary program by regulating Fgf10 and Tbx3 expression and Wnt/beta-catenin signaling, Development 142 (23) (2015) 4056–4067.
[35] O. Wansbury, H. Panchal, M. James, S. Parry, A. Ashworth, B. Howard, Dynamic expression of Erbb pathway members during early mammary gland morphogenesis, J. Investig. Dermatol. 128 (4) (2008) 1009–1021.
[36] B. Howard, Identification of the scaramanga gene implicates Neuregulin3 in mammary gland specification, Genes Dev. 19 (17) (2005) 2078–2090.
[37] H. Panchal, O. Wansbury, S. Parry, A. Ashworth, B. Howard, Neuregulin3 alters cell fate in the epidermis and mammary gland, BMC Dev. Biol. 7 (2007) 105.
[38] P.H. Lindfors, M. Voutilainen, M.L. Mikkola, Ectodysplasin/NF-kappaB signaling in embryonic mammary gland development, J. Mammary Gland Biol. Neoplasia 18 (2) (2013) 165–169.
[39] S.M. Myllymaki, M.L. Mikkola, Inductive signals in branching morphogenesis – lessons from mammary and salivary glands, Curr. Opin. Cell Biol. 61 (2019) 72–78.
[40] A. Clarke, D.I. Phillips, R. Brown, P.S. Harper, Clinical aspects of X-linked hypohidrotic ectodermal dysplasia, Arch. Dis. Child 62 (10) (1987) 989–996.
[41] H. Mégarbané, C. Cluzeau, C. Bodemer, S. Fraïtag, M. Chababi-Atallah, A. Mégarbané, A. Smahi, Unusual presentation of a severe autosomal recessive anhydrotic ectodermal dysplasia with a novel mutation in the EDAR gene, Am. J. Med. Genet. A 146A (20) (2008) 2657–2662.
[42] C.Y. Cui, J.A. Smith, D. Schlessinger, C.C. Chan, X-linked anhidrotic ectodermal dysplasia disruption yields a mouse model for ocular surface disease and resultant blindness, Am. J. Pathol. 167 (1) (2005) 89–95.
[43] A. Haghighi, P. Nikuei, H. Haghighi-Kakhki, N. Saleh-Gohari, S. Baghestani, P. M. Krawitz, J. Hecht, S. Mundlos, Whole-exome sequencing identifies a novel missense mutation in EDAR causing autosomal recessive hypohidrotic ectodermal dysplasia with bilateral amastia and palmoplantar hyperkeratosis, Br. J. Dermatol. 168 (6) (2013) 1353–1356.
[44] M.L. Mikkola, Molecular aspects of hypohidrotic ectodermal dysplasia, Am. J. Med. Genet. A 149A (9) (2009) 2031–2036.
[45] M. Voutilainen, P.H. Lindfors, E. Trela, D. Lönnblad, V. Shirokova, T. Elo, E. Rysti, R. Schmidt-Ullrich, P. Schneider, M.L. Mikkola, Ectodysplasin/NF-κB promotes mammary cell fate via Wnt/β-catenin pathway, PLoS Genet. 11 (11) (2015), e1005676.
[46] M. Voutilainen, P.H. Lindfors, S. Lefebvre, L. Ahtiainen, I. Fliniaux, E. Rysti, M. Murtoniemi, P. Schneider, R. Schmidt-Ullrich, M.L. Mikkola, Ectodysplasin regulates hormone-independent mammary ductal morphogenesis via NF-kappaB, Proc. Natl. Acad. Sci. U. S. A. 109 (15) (2012) 5744–5749.
[47] J. Pispa, M.L. Mikkola, T. Mustonen, I. Thesleff, Ectodysplasin, Edar and TNFRSF19 are expressed in complementary and overlapping patterns during mouse embryogenesis, Gene Expr. Patterns 3 (5) (2003) 675–679.
[48] T. Mustonen, Ectodysplasin A1 promotes placodal cell fate during early morphogenesis of ectodermal appendages, Development 131 (20) (2004) 4907–4919.
[49] K. Närhi, M. Tummers, L. Ahtiainen, N. Itoh, I. Thesleff, M.L. Mikkola, Sostdc1 defines the size and number of skin appendage placodes, Dev. Biol. 364 (2) (2012) 149–161.
[50] I. Fliniaux, M.L. Mikkola, S. Lefebvre, I. Thesleff, Identification of dkk4 as a target of Eda-A1/Edar pathway reveals an unexpected role of ectodysplasin as inhibitor of Wnt signalling in ectodermal placodes, Dev. Biol. 320 (1) (2008) 60–71.
[51] H. Meinhardt, A. Gierer, Pattern formation by local self-activation and lateral inhibition, Bioessays 22 (8) (2000) 753–760.
[52] A. Gierer, H. Meinhardt, A theory of biological pattern formation, Kybernetik 12 (1) (1972) 30–39.
[53] M. Loukas, P. Clarke, R.S. Tubbs, Accessory breasts: a historical and current perspective, Am. Surg. 73 (5) (2007) 525–528.
[54] Y.G. Kamberov, S. Wang, J. Tan, P. Gerbault, A. Wark, L. Tan, Y. Yang, S. Li, K. Tang, H. Chen, A. Powell, Y. Itan, D. Fuller, J. Lohmueller, J. Mao, A. Schachar, M. Paymer, E. Hostetter, E. Byrne, M. Burnett, A.P. McMahon, M.G. Thomas, D. E. Lieberman, L. Jin, C.J. Tabin, B.A. Morgan, P.C. Sabeti, Modeling recent human evolution in mice by expression of a selected EDAR variant, Cell 152 (4) (2013) 691–702.
[55] T. Elo, P.H. Lindfors, Q. Lan, M. Voutilainen, E. Trela, C. Ohlsson, S.H. Huh, D. M. Ornitz, M. Poutanen, B.A. Howard, M.L. Mikkola, Ectodysplasin target gene Fgf20 regulates mammary bud growth and ductal invasion and branching during puberty, Sci. Rep. 7 (1) (2017) 5049.
[56] M.E. Dunbar, P.R. Dann, G.W. Robinson, L. Hennighausen, J.P. Zhang, J. J. Wysolmerski, Parathyroid hormone-related protein signaling is necessary for sexual dimorphism during embryonic mammary development, Development 126 (16) (1999) 3485–3493.
[57] J.J. Wysolmerski, W.M. Philbrick, M.E. Dunbar, B. Lanske, H. Kronenberg, A. E. Broadus, Rescue of the parathyroid hormone-related protein knockout mouse demonstrates that parathyroid hormone-related protein is essential for mammary gland development, Development 125 (7) (1998) 1285–1294.
[58] M.E. Dunbar, P. Young, J.P. Zhang, J. McCaughern-Carucci, B. Lanske, J.J. Orloff, A. Karaplis, G. Cunha, J.J. Wysolmerski, Stromal cells are critical targets in the regulation of mammary ductal morphogenesis by parathyroid hormone-related protein, Dev. Biol. 203 (1) (1998) 75–89.
[59] J. Foley, P. Dann, J. Hong, J. Cosgrove, B. Dreyer, D. Rimm, M. Dunbar, W. Philbrick, J. Wysolmerski, Parathyroid hormone-related protein maintains mammary epithelial fate and triggers nipple skin differentiation during embryonic breast development, Development 128 (4) (2001) 513–525.
[60] M. Dunbar, P. Dann, C. Brown, J. Van Houton, B. Dreyer, W. Philbrick, J. Wysolmerski, Temporally regulated overexpression of parathyroid hormone-related protein in the mammary gland reveals distinct fetal and pubertal phenotypes, J. Endocrinol. 171 (3) (2001) 403–416.
[61] A.S. Jobert, P. Zhang, A. Couvineau, J. Bonaventure, J. Roume, M. Le Merrer, C. Silve, Absence of functional receptors for parathyroid hormone and parathyroid hormone-related peptide in Blomstrand chondrodysplasia, J. Clin. Investig. 102 (1) (1998) 34–40.
[62] J. Hens, P. Dann, M. Hiremath, T.C. Pan, L. Chodosh, J. Wysolmerski, Analysis of gene expression in PTHrP-/- mammary buds supports a role for BMP signaling and MMP2 in the initiation of ductal morphogenesis, Dev. Dyn. 238 (11) (2009) 2713–2724.
[63] M. Hiremath, J. Wysolmerski, Parathyroid hormone-related protein specifies the mammary mesenchyme and regulates embryonic mammary development, J. Mammary Gland Biol. Neoplasia 18 (2) (2013) 171–177.
[64] M. Hiremath, P. Dann, J. Fischer, D. Butterworth, K. Boras-Granic, J. Hens, J. Van Houten, W. Shi, J. Wysolmerski, Parathyroid hormone-related protein activates Wnt signaling to specify the embryonic mammary mesenchyme, Development 139 (22) (2012) 4239–4249.
[65] J.R. Hens, P. Dann, J.P. Zhang, S. Harris, G.W. Robinson, J. Wysolmerski, BMP4 and PTHrP interact to stimulate ductal outgrowth during embryonic mammary development and to inhibit hair follicle induction, Development 134 (6) (2007) 1221–1230.
[66] I. Satokata, L. Ma, H. Ohshima, M. Bei, I. Woo, K. Nishizawa, T. Maeda, Y. Takano, M. Uchiyama, S. Heaney, H. Peters, Z. Tang, R. Maxson, R. Maas, Msx2 deficiency in mice causes pleiotropic defects in bone growth and ectodermal organ formation, Nat. Genet. 24 (4) (2000) 391–395.
[67] C. Lindvall, C.R. Zylstra, N. Evans, R.A. West, K. Dykema, K.A. Furge, B. O. Williams, The Wnt co-receptor Lrp6 is required for normal mouse mammary gland development, PLoS One 4 (6) (2009), e5813.
[68] C. Lindvall, N.C. Evans, C.R. Zylstra, Y. Li, C.M. Alexander, B.O. Williams, The Wnt signaling receptor Lrp5 is required for mammary ductal stem cell activity and Wnt1-induced tumorigenesis, J. Biol. Chem. 281 (46) (2006) 35081–35087.
[69] B. Gu, P. Sun, Y. Yuan, R.C. Moraes, A. Li, A. Teng, A. Agrawal, C. Rhéaume, V. Bilanchone, J.M. Veltmaat, K.I. Takemaru, S. Millar, E.Y.H.P. Lee, M.T. Lewis, B. Li, X. Dai, Pygo2 expands mammary progenitor cells by facilitating histone H3 K4 methylation, J. Cell Biol. 185 (5) (2009) 811–826.
[70] S. Rivetti, C. Chen, C. Chen, S. Bellusci, Fgf10/Fgfr2b signaling in mammary gland development, homeostasis, and cancer, Front Cell Dev. Biol. 8 (2020) 415.
[71] T. Sakakura, Y. Nishizuka, C.J. Dawe, Capacity of mammary fat pads of adult C3H/HeMs mice to interact morphogenetically with fetal mammary epithelium, J. Natl. Cancer Inst. 63 (3) (1979) 733–736.
[72] B. Lloyd-Lewis, O.B. Harris, C.J. Watson, F.M. Davis, Mammary stem cells: premise, properties, and perspectives, Trends Cell Biol. 27 (8) (2017) 556–567.
[73] A. Van Keymeulen, A.S. Rocha, M. Ousset, B. Beck, G. Bouvencourt, J. Rock, N. Sharma, S. Dekoninck, C. Blanpain, Distinct stem cells contribute to mammary gland development and maintenance, Nature 479 (7372) (2011) 189–193.
[74] M. Ito, Z. Yang, T. Andl, C. Cui, N. Kim, S.E. Millar, G. Cotsarelis, Wnt-dependent de novo hair follicle regeneration in adult mouse skin after wounding, Nature 447 (7142) (2007) 316–320.
[75] L. Bai, L.R. Rohrschneider, s-SHIP promoter expression marks activated stem cells in developing mouse mammary tissue, Genes Dev. 24 (17) (2010) 1882–1892.
[76] L. Tian, M.J. Truong, C. Lagadec, E. Adriaenssens, E. Bouchaert, H. Bauderlique-Le Roy, M. Figeac, X. Le Bourhis, R.P. Bourette, s-SHIP promoter expression identifies mouse mammary cancer stem cells, Stem Cell Rep. 13 (1) (2019) 10–20.
[77] C.L. Trejo, G. Luna, C. Dravis, B.T. Spike, G.M. Wahl, Lgr5 is a marker for fetal mammary stem cells, but is not essential for stem cell activity or tumorigenesis, NPJ Breast Cancer 3 (2017) 16.
[78] R. van Amerongen, A.N. Bowman, R. Nusse, Developmental stage and time dictate the fate of Wnt/beta-catenin-responsive stem cells in the mammary gland, Cell Stem Cell 11 (3) (2012) 387–400.
[79] Z. Ying, S. Beronja, Embryonic barcoding of equipotent mammary progenitors functionally identifies breast cancer drivers, Cell Stem Cell 26 (3) (2020) 403–419 e4.
[80] A.M. Lilja, V. Rodilla, M. Huyghe, E. Hannezo, C. Landragin, O. Renaud, O. Leroy, S. Rulands, B.D. Simons, S. Fre, Clonal analysis of Notch1-expressing cells reveals the existence of unipotent stem cells that retain long-term plasticity in the embryonic mammary gland, Nat. Cell Biol. 20 (6) (2018) 677–687.
[81] B. Lloyd-Lewis, F.M. Davis, O.B. Harris, J.R. Hitchcock, C.J. Watson, Neutral lineage tracing of proliferative embryonic and adult mammary stem/progenitor cells, Development 145 (2018), dev164079.
[82] E. Shema, B.E. Bernstein, J.D. Buenrostro, Single-cell and single-molecule epigenomics to uncover genome regulation at unprecedented resolution, Nat. Genet. 51 (1) (2019) 19–25.
[83] T. Sakakura, Mammary embryogenesis, in: M.C. Neville, C.W. Daniel (Eds.), The Mammary Gland, Plenum, New York, NY, 1987, pp. 37–63.
[84] X. Sun, A. Are, K. Annusver, U. Sivan, T. Jacob, T. Dalessandri, S. Joost, A. Füllgrabe, M. Gerling, M. Kasper, Coordinated hedgehog signaling induces new hair follicles in adult skin, Elife 9 (2020) 9.

[85] L.E. Dow, K.P. O'Rourke, J. Simon, D.F. Tschaharganeh, J.H. van Es, H. Clevers, S. W. Lowe, Apc restoration promotes cellular differentiation and reestablishes crypt homeostasis in colorectal, Cancer Cell 161 (7) (2015) 1539–1552.

[86] T. Andl, S.T. Reddy, T. Gaddapara, S.E. Millar, WNT signals are required for the initiation of hair follicle development, Dev. Cell 2 (5) (2002) 643–653.

[87] Y. Ahn, C. Sims, J.M. Logue, S.D. Weatherbee, R. Krumlauf, Lrp4 and Wise interplay controls the formation and patterning of mammary and other skin appendage placodes by modulating Wnt signaling, Development 140 (3) (2013) 583–593.

Gpr125 identifies myoepithelial progenitors at tips of lacrimal ducts and is essential for tear film.

Departments of Cell Biology[1] Dermatology[2] and Ophthalmology[3]

New York University School of Medicine

Elena Spina[1], Rebecca Handlin[1], Julia Simundza[1], Angela Incassati[1], Muneeb Faiq[3], Anoop Sainulabdeen[3], Kevin C Chan[3], Pamela Cowin[1,2]*

*Address for correspondence:

Dr. Pamela Cowin

New York University Medical Center

MSB 621, 550 1st Ave

New York, NY 10016

E-mail: pamela.cowin@nyulangone.org

Abstract

Gpr125, encoded by *Adgra3*, is an orphan adhesion G-protein coupled receptor (aGPCR) implicated in modulating Wnt signaling and planar polarity. Here we establish both physiological and pathological roles for Gpr125. We show that mice lacking Gpr125 or its signaling domains display an ocular phenotype with many hallmarks of human dry eye syndrome. These include squinting, abnormal lacrimation, mucus accumulation, swollen eyelids and inflammatory infiltration of lacrimal and meibomian glands. Utilizing a Gpr125-β-gal reporter and scRNAseq, we identify Gpr125 expression in a discrete population of cells located at the tips of migrating embryonic lacrimal ducts. By lineage tracing we show these cells function as progenitors of the adult lacrimal myoepithelium. Beyond defining an essential role for Gpr125 in tear film and identifying its utility as a marker of lacrimal progenitors, this study implicates Gpr125 in the etiology of blepharitis and dry eye syndrome, and defines novel animal models of these common maladies.

Introduction

Gpr125 is an orphan adhesion G-protein coupled receptor (aGPCR) that was discovered through homology searches of the human genome database (Bjarnadottir et al., 2006). Like other members of this family, Gpr125 has a large extracellular domain with sequence similarity to cell adhesion molecules (**Figure 1A**) (Bjarnadottir et al., 2006; Simundza and Cowin, 2013). Previous studies have highlighted Gpr125 as a marker of undifferentiated murine spermatogonial progenitors (Seandel et al., 2007), documented its elevation in the choroid plexus following injury and correlated high Gpr125 expression with both good and poor outcome in cancer (Fu et al., 2013; Pickering et al., 2008; Wu et al., 2018). When introduced into cultured cells Gpr125 undergoes constitutive clathrin-mediated internalization to endosomes suggesting a role in receptor recycling (Spiess et al., 2019). When expressed ectopically in zebrafish, Gpr125 interacts with the cytoplasmic adaptor Disheveled (Dsh) and recruits Frz7 and Glypican4 (Gpc4) complexes. When reduced it has little effect in wildtype zebrafish but in Wnt/planar cell polarity (PCP) mutants exacerbates defects in convergent extension and directed migration of facial branchiomotor neurons (FBMN) (Li et al., 2013). Gpr125 shares significant homology with Gpr124, which has been shown to regulate angiogenic sprouting and control selective Wnt signaling by stabilizing specific ligand receptor interactions (Anderson et al., 2011; Vanhollebeke et al., 2015). To date, the physiological function of endogenous Gpr125 in higher vertebrates has remained elusive. Here we uncover an essential physiological role for Gpr125 in the lacrimal gland, and a pathological role in the etiology of blepharitis and dry eye disease (DED).

Combined Results & Discussion

Mice lacking Gpr125 display blepharitis, blepharedema and mucoid accumulation

To address the role of native Gpr125, we developed mice that permit Gpr125 expression to be ablated and the lineage of cells normally expressing it to be traced by inserting a creER^{T2} cassette downstream of the *Adgra3* promoter (**Figure 1B**). Mice lacking Gpr125 expression (*Adgra3*$^{cre/cre}$) display a prominent eye phenotype (**Figure 1C**). *Adgra3*$^{cre/cre}$ mice squint as soon as their eyes open; whereas, heterozygous *Adgra3*$^{cre/+}$ are indistinguishable from wild-type littermates (**Figure 1C**). As *Adgra3*$^{cre/cre}$ mice mature, this early blepharitis progresses to blepharedema (swollen balding eyelids) and mucus precipitation (**Figure 1D**). The phenotype is constant in males but in females oscillates with reproductive status, becoming pronounced during pregnancy and lactation. During these stages, mice develop proptosis (bulging eyes) that resolves during weaning. The eye phenotype in *Adgra3*$^{cre/cre}$ mice is 100% penetrant on all strain backgrounds examined (C57B6/CH3, FVBN, and mixed). To dissect the role of Gpr125's adhesion ectodomain from its internal signaling functions we examined a second strain, *Adgra3*$^{lz/lz}$, which expresses the Gpr125 ectodomain and 1st transmembrane domain fused in frame to β-galactosidase and lacks regions required for signaling/adaptor functions (**Figure 1E**) (Seandel et al., 2007). Homozygous *Adgra3*$^{lz/lz}$ mice recapitulate the *Adgra3*$^{cre/cre}$ null phenotype (**Figure 1F**), whereas, *Adgra3*$^{lz/+}$ mice are normal. Collectively, these data demonstrate that Gpr125 has an essential physiological role in normal eye development and indicate that signaling downstream of the receptor is required. These analyses also

reveal that loss of Gpr125 protein or Gpr125 signaling is sufficient to trigger several common eye pathologies, such as blepharitis, blepharedema and mucoid accumulation.

Gpr125 in ocular structures and correlated pathologies

We examined adult eye globes by X-gal staining and in *Adgra3*$^{lz/+}$ mice found strong Gpr125-β-gal expression in the inner layer of the iris and in the ciliary body, which secretes aqueous humor (**Figure 2A, B**). As abnormal aqueous humor dynamics can alter intraocular pressure (IOP), which is a major risk factor for glaucoma, we measured IOP, but found no significant difference between wildtype and mutant genotypes (**Figure 2C**).

Next, we submitted both strains of mice for evaluation by a veterinary ophthalmologist. Examination of the lens and retina by slit lamp revealed well-documented characteristics of control B6 and FVBN mice, but no abnormality specifically linked to the *Adgra3*$^{cre/cre}$ or *Adgra3*$^{lz/lz}$ genotypes. Fluorescein staining revealed no evidence for corneal abrasion, but highlighted the presence of large mucoid precipitates around the eyelids of homozygous mutants (**Figure 2D**). This feature pointed towards abnormal tear film composition. Tears are required to lubricate corneal and conjunctival surfaces and to prevent eyes from desiccation (Botelho, 1964). They also function to protect eyes from microbial infection and preserve visual acuity. Tear film is composed of three layers, each secreted from a different source (**Figure 2E**). Goblet cells, clustered along the conjunctival rim, provide the inner mucus layer that spreads tear film evenly over the ocular surface (Gipson, 2016; Rios et al., 2000). Meibomian glands, found between eyelash follicles on the inner surface of eyelids, produce the outer lipid layer that prevents evaporation (Bron and Tiffany, 1998;

Nien et al., 2010). Lacrimal glands secrete the central aqueous component that contains water-soluble immuno-active and antibacterial proteins, as well as glucose, urea, and salts (Makarenkova et al., 2000). Defects in the volume or composition of any layer destabilizes tear film and induces DED prompting us to investigate the tear glands in more detail (Pflugfelder and de Paiva, 2017; Schaumberg et al., 2003; Schaumberg et al., 2002). As *Adgra3* mutant mice had swollen eyelids we looked first for changes in goblet cells and meibomian glands. Histological sections of eyelids stained with Alcian blue revealed goblet cells in *Adgra3*$^{cre/cre}$ mice (**Figure 2F**) but with greater variation in number (average =70/eyelid; range 7-210 n=23) compared to controls (average of 65 goblet cells/eyelid; range 45-94; n=23): some showed epithelial and goblet cell desquamation next to swathes of mucus; others showed clusters of goblet cells. Meibomian glands displayed inflammatory infiltration by T-cells and macrophages (**Figure 2G,H**). Surprisingly, given these phenotypes, goblet cells and meibomian glands were devoid of Gpr125 expression whereas eyelash follicles were positive in X-gal stained eyelids (**Figure 2I**). These data show that changes in goblet and meibomian glands seen in human DED occur of *Adgra3* mutants (Pflugfelder and de Paiva, 2017; Schaumberg et al., 2003; Schaumberg et al., 2002) but as a secondary consequence of a tear film abnormality caused by loss of Gpr125 elsewhere. By a process of elimination this led us to focus on the lacrimal gland.

Adgra3$^{cre/cre}$ and *Adgra3*$^{lz/lz}$ mice have abnormal lacrimation

We tested whether Gpr125 loss affected lacrimal function by measuring tear volume. *Adgra3*$^{cre/cre}$ and *Adgra3*$^{lz/lz}$ mice produced two to three-fold more tears than heterozygous

or wild-type controls (**Figure 3A**). Tear volume was greater in female than male mice. *Adgra3*$^{cre/cre}$ and *Adgra3*$^{lz/lz}$ mice often presented with a mild phenotype in one eye (squint only) and a severe phenotype (blepharedema and or mucus) in the other. When we separated eyes into mild and severe categories according to photographic assignment taken prior to measurement and then reanalyzed the data, we found tear volume for the mild phenotypic category was similar, and sometimes lower, than those of wildtypes. In contrast, those in the severe category showed high values indicative of excessive tearing (**Figure 3A**). Thus, our mice recapitulated the paradoxical phenomenon documented in human patients with DED where individuals with tear film abnormality originating from initial mild ocular dryness respond to the consequent eye irritation with compensatory hyper-lacrimation (Pflugfelder and de Paiva, 2017).

Gpr125-expressing cells are located at the leading tips of ducts during lacrimal development and function as progenitors of the lacrimal myoepithelium

Given the significant effect of *Adgra3* loss on lacrimal function, we sought to identify cell types that express Gpr125 over the course of lacrimal development. We began by mining scRNAseq data (Farmer et al., 2017). Gpr125 mRNA was detected in a small cell population that co-expressed keratin 14 and Sox10 mRNAs (**Figure 3B**). This population was present during the early developmental stage of ductal elongation (E16) but diminished by P4 (data not shown) as acinar differentiation ensued. Next, we stained embryos with X-gal to locate cells expressing Gpr125 (**Figure 3C**). Murine lacrimal glands emerge around embryonic day 13 (E13) as a bulbous outgrowth of the conjunctival epithelium, which by E15 has elongated as a bi-layered hollow duct with 4-5 bulbous tips.

Subsequent branching produces a compact mass of secretory acini, enmeshed by contractile myoepithelial cells, which fully differentiate after birth (Dean et al., 2004; Dean et al., 2005; Farmer et al., 2017; Makarenkova et al., 2000). Gpr125-β-gal appeared in the lacrimal bud as it emerged from the conjunctival rim ~E14, but by E15.5 it was restricted to a discrete population of cells located at the leading tips of migrating lacrimal ducts and by P1 at the front of lacrimal branches **(Figure 3C).** A potential role in directional outgrowth and collective cell migration is suggested by the fact that Gpr125 contains LRIG motifs that are present in Slit/Robo guidance factors (Bjarnadottir et al., 2006; Simundza and Cowin, 2013). This concept is supported by studies in zebrafish have shown that Gpr125 expression levels impact upon the migration of facial branchiomotor neurons (Li et al., 2013). Gpr124, is required for tip cell function during angiogenic sprouting raising the possibility that this family of proteins may serve similar roles in distinct cell types (Anderson et al., 2011; Vanhollebeke et al., 2015).

Intriguingly, during the course of these experiments we noted that Gpr125-β-gal was also expressed within a well-characterized “bulge” stem cell compartment of hair follicles and whiskers (**Figure 3C, D**) (Cotsarelis et al., 1990). This prompted us to ask if the embryonic Gpr125-positive cells present at ductal tips functioned as lacrimal progenitors. To test this, we performed lineage tracing by using the *creER*^{T2} cassette present in *Adgra3*^{cre} mice to activate expression of a lineage reporter. We labeled embryos harboring the ROSA-lox-STOP-lox-tdTomato reporters at E13-E15 by delivering tamoxifen to *Adgra3*$^{cre/cre}$ dams during mid-pregnancy. Lacrimal glands were harvested at 7 weeks and 6 months of age and analyzed by 3-D immunofluorescence confocal microscopy (**Figure**

3E,F). At 7 weeks, we found tdTomato (tdT)-labeled cells with an elongated shape along the basal borders of ducts and with stellate morphology enmeshing acini. These characteristics, together with their expression of keratin 5 (K5), identified them as contractile myoepithelial cells. A similar pattern was seen in glands from mice harvested at 6 months, indicating significant longevity of the original progenitor population (**Figures 3F,G**). Collectively, these data show that Gpr125+ cells at tips of migrating embryonic lacrimal ducts function as long-lived unipotent progenitors of the ductal and acinar myoepithelium. Recent studies have traced K14, K5, Sma and Runx lineages in the lacrimal gland and revealed a complex hierarchy comprising multipotent Runx cells at the apex and lineage restricted progenitors arising before birth (Basova et al., 2020). Our study adds to these analyses by providing a highly specific cell surface marker of Sma+ Sox10+ unipotent myoepithelial progenitors, locating these progenitors at the distal tips of elongating embryonic ducts, and demonstrating that they are already lineage restricted between E13-E15 of embryonic lacrimal development. Our findings of Gpr125 in stem cell compartment of several tissues together with early documentation of its expression in spermatogonial progenitors, show Gpr125 has widespread utility as marker for the localization and isolation of early progenitors.

Adgra3$^{cre/cre}$ and *Adgra3*$^{lz/lz}$ mice show inflammatory infiltration of lacrimal glands

Given Gpr125 expression in myoepithelial progenitors and its homology to adhesion receptors, we investigated the effect of Gpr125 loss on myoepithelial integrity of lacrimal glands. In histological sections we noted the presence of foci, composed of small round cells where the lacrimal acinar organization was disrupted in *Adgra3*$^{cre/cre}$ (**Figure 4A,B**)

and *Adgra3*$^{lz/lz}$ mice. Immunostaining for K5 was conspicuously absent in these foci indicating that myoepithelial cells were lost or disrupted (**Figure 4C,D**). These foci were surrounded by F480-positive macrophages (**Figures 4C,E**), and filled with cells recognized by CD4 and CD8 antibodies (**Figures 4C, F,G**) indicating infiltration by T-helper cells and cytoxic T-cells. Histological analysis of females displaying facial swelling (**Figure 4H**) and proptosis during pregnancy and lactation revealed enlarged lacrimal glands (**Figure 4I,J**) with swathes of macrophages around large areas of acinar loss (**Figure 4K**). Thus, loss of Gpr125 is accompanied by impaired lacrimal myoepithelial integrity and lymphocytic infiltration. Inflammatory infiltration is a common feature in later stages of human DED and a central feature of Sjogren's syndrome, the third most common auto-immune disease (Pflugfelder et al., 2018; Pflugfelder and de Paiva, 2017; Schaumberg et al., 2003; Schaumberg et al., 2002).

Analysis of knockout (KO) mice has revealed a suite of critical regulators of tear film (Chen et al., 2014; Cui et al., 2005; Dean et al., 2004; Dean et al., 2005; Gipson, 2016; Kenchegowda et al., 2011; Makarenkova et al., 2000; Marko et al., 2013; McMahon et al., 2014; Plikus et al., 2004; Tong and Gupta, 2016; Tsau et al., 2011). However, the involvement of GPCRs in this process has not been studied. In sporadic DED, tear film abnormality exposes the eye to irritation and desiccation, prompting compensatory excessive tearing and immune response, which leads to further lacrimal destruction (Pflugfelder and de Paiva, 2017). Our *Adgra3* mutants reproduce this complex spectrum of symptoms, from early eye discomfort to blepharedema (**Figure 1**), mucus accumulation **(Figures 1, 2D),** goblet cells desquamation **(Figure 2F),** compensatory hyper-lacrimation,

and inflammatory infiltration of meibomian **(Figures 2H)** and lacrimal glands **(Figure 4B-G).** Moreover, *Adgra3* mutants show worsening of their eye phenotype during pregnancy and lactation, recapitulating the hormonal/gender epidemiology of DED, which is more prevalent in women and exacerbated by pregnancy and post-menopause (Schaumberg et al., 2003; Schaumberg et al., 2002). Many genetic mouse models of DED arise from immune dysregulation or defects in matrix inhibition of immune activation and thus recapitulate late stages of DED (Park et al., 2015; Tong and Gupta, 2016). In contrast, our mice identify an initiating event during lacrimal development that predisposes mice to the full pathophysiological progression of DED. We show that in the absence of Gpr125, focal areas of the lacrimal gland become devoid of myoepithelium **(Figure 4D)** and infiltrated by lymphocytes and macrophages (**Figure 4C-G**). These results reinforce the concept that myoepithelial cells play a critical role in DED and complement studies that have shown that myoepithelial differentiation and contractile function is altered in Sjogren's patients (Hawley et al., 2018; Makarenkova and Dartt, 2015). Going forward, it will be important to determine if Gpr125 is involved in blepharitis and DED in humans. DED is a significant health problem that affects ~5% of the population overall is particularly prevalent in the elderly and women (Pflugfelder and de Paiva, 2017; Schaumberg et al., 2003; Schaumberg et al., 2002). As GPCRs are currently the targets of approximately 34% of drugs approved by the US Food and Drug Administration then deciphering Gpr125 signaling pathways holds promise to uncover novel targets for therapeutic intervention in these common conditions (Bassilana et al., 2019).

Materials and Methods

Mice

Mice were constructed by Ingenious Technologies, Ronkonkoma, NY as follows. A cassette containing *CreER*^{T2} followed by a 3' polyadenylation signal, harboring SV40-driven Neo flanked by FRT sites inserted in a central intron, was recombined into a bacterial artificial chromosome (BAC) to place *CreER*^{T2} under the control of the *Adgra3* promoter, excising 502 bp encompassing 221 bp of exon 1 and part of the following intron 1-2 of *Adgra3*. Mice generated from these ES cells were selected for germline transmission by PCR, verified by southern analysis and sequencing then bred to a Flp deleter strain to remove Neo. *Adgra3*$^{lz/+}$ mice were generated by Regeneron using VelociGene methods to modify a bacterial artificial chromosome (BAC) clone carrying the mouse *Adgra3* gene by replacement of sequence encompassing exons 16-19 with *lacZ* to produce expression of fusion protein comprising the N-terminal extracellular domain, the first transmembrane domain, and part of the first intracellular loop of Gpr125 fused to β-galactosidase (Figure 1A) (Seandel et al., 2007; Valenzuela et al., 2003). Animal experiments were approved by NYUMC institutional animal care and use committee and conformed to American Association for Accreditation of Laboratory Animal Care guidelines.

Ophthalmologic examination

Standard ophthalmic examination was performed by a trained veterinary ophthalmology consultant (Dr. Michael Brown, Animal Eyes of New Jersey). Slit lamp biomicroscopy was

used to assess the cornea, anterior chamber, iris, lens, and vitreous humor. Mydriasis was induced with tropicamide and the retina was examined via indirect ophthalmoscopy. Corneal fluorescein staining was performed by applying sodium fluorescein (1%), for 3 minutes to the cornea of mice. Excess fluorescein was removed by flushing with sterile phosphate buffered saline (PBS) and corneal staining was evaluated and photographed with a slit lamp biomicroscope (Humphrey-Zeiss, Dublin, CA) using a cobalt blue light. Punctate staining was recorded using a standardized National Eye Institute grading system of 0 to 3 for each of the five areas of the cornea.

Schirmer Tear Test

Tear production was measured via a modified Schirmer Tear Test. Briefly, 35mm x 5mm wide commercial Schirmer Tear Test standardized sterile strips (Schirmer Tear Test; Merck Animal Health) were transected with into two 15mm x 2.5mm strips, with the top notch removed. Individual strips were placed under the lower eyelid using forceps and removed after 15 seconds. The length of dye migration and wetting of the strip was measured in millimeters under a dissecting microscope.

Intraocular pressure measurement

Mice were anesthetized and maintained on isoflurane through a nose cone. IOPs were measured using a TonoLab rebound tonometer (Icare, Finland) within 5 min after isoflurane gas anesthesia induction. For every 6 valid measurements, the highest and lowest IOP values were automatically excluded by the device, and the average of the

remaining 4 IOP values was displayed along with the deviation. For quality control, only averages with slight deviation of less than 2.5 mmHg were considered acceptable readings. This procedure was repeated at least 3 times for each eye, and the acceptable readings were averaged IOP was measured eighteen times for each eye, and the average value was used for final analysis.

Histological Analysis

The exorbital lacrimal gland, salivary and parotid glands, and whole globes were removed from mice and fixed with either 10% neutral buffered formalin or 4% paraformaldehyde (PFA) and embedded in paraffin. For general histological assessment, sections were stained with hematoxylin and eosin (H&E), or with periodic acid-Schiff (PAS) and Alcian Blue to visualize conjunctival goblet cells. Goblet cells in the bulbar and palpebral conjunctiva were quantified by two separate readers. Serial sections of tissues were stained with antibodies for anti-CD4, anti-CD8, anti-F480, anti-cytokeratin CK5 optimized by the Experimental Histology Core, NYUMC for analysis by Akoya/PerkinElmer Vectra® multispectral imaging system then counterstained with Dapi.

Lineage Tracing

For lineage tracing experiments, *Adgra3*^{creERT2} mice were crossed to the fluorescent Rosa26R-lox.STOP.lox-tdTomato lineage reporter strain (Stock No. 007909) Jackson laboratory. The transcriptional STOP was deleted by cre recombination during embryonic development (E14.5-E15.5) by delivering tamoxifen (5mg per mouse-2

doses of 2.5mg) by oral gavage to *Adgar3*$^{cre/cre}$ dams with during mid-pregnancy. Pups were delivered at E19.5-E20.5 by caesarian section to avoid problems with delivery caused by Tamoxifen and fostered by SWR/J mice. Their tissues were harvested at 7 weeks and at 6 months to test for progenitor potency and longevity.

Tissue clearing and 3-D imaging

Lacrimal glands were excised and fixed overnight in 4% PFA then processed using a modified CUBIC (Reagent 1A) protocol (Davis et al., 2016). Tissue was incubated in CUBIC Reagent 1A clearing solution for 4 days, rinsed 3X in PBS then immunostained for 4 days at 4C in PBST containing 10% rabbit serum and rabbit anti-K5 (Covance, PRB160P, 1:100), rinsed again then 2 days in goat anti-rabbit AlexaFluor (AF) 647 (Thermo Fisher Scientific, A21245, lot number 1805235, 1:500), rinsed 3X then cleared in CUBIC R2 for 24hrs. Cleared lacrimal tissues were imaged using a Zeiss 880 Laser Scanning inverted confocal microscope with 20X air Plan-Apochromat N.A. 0.8 M27 objective lenses.

X-gal staining

Embryos, Eyes and lacrimal glands were fixed in 4% paraformaldehyde (PFA) (Sigma-Aldrich) at room temperature (RT) for 30-60 min, rinsed 3X in X-gal rinse buffer (2 mM MgCl2, 0.1% Sodium deoxycholate, and 0.2% NP-40 in PBS) at RT, then incubated in X-gal staining solution (50 mg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside in rinse buffer containing 5 mM potassium ferricyanide, 5 mM potassium ferrocyanide)

(Applichem, Cheshire, CT) at RT overnight. After staining, glands were rinsed in PBS, post-fixed in 4% PFA overnight then prepared for whole mount analysis or processed for paraffin embedding and sectioned for histological analysis.

scRNA-seq analysis

scRNA-seq analysis were generated mining data from Farmer DT. et al. (Farmer et al., 2017) We processed the dataset using iCellR, Single (i) Cell R package, an interactive R package to work with high-throughput single cell sequencing technologies with the help of NYU Langone's Applied Bioinformatics Laboratories (https://www.biorxiv.org/content/10.1101/2020.03.31.019109v1).

Statistical Analysis

Experimental data are presented as mean ± SEM. P values for experiments comparing two groups were calculated using student's t test. For experiments comparing more than two groups, an Ordinary one-way ANOVA was used with multiple comparisons test. $P<0.05$ was considered statistically significant.

Online supplemental material

Fig. S1 (related to Fig. 2) shows immunostained meibomian gland in control *Adgra3*$^{+/+}$ *versus Adgra3*$^{cre/cre}$ mice.

Acknowledgements

This work was supported by Department of Defense W81XWH-17-1-0013 BC160959 (PC), NIH training grants: T32GM066704 (JS), T32CA009161-44 (AI) and The Susan G Komen Foundation For The Cure (AI). We thank Applied Bioinformatics, Experimental Pathology Research and Microscopy Laboratory Cores for the service provided: Grants P30CA016087, S10 OD021747 and P30CA016087. We thank Dr. Aris Economides, Regeneron, Tarrytown NY for advice on *Adgra3*-creER^{T2} construction and Dr. Michael Brown DVM (Animal Eyes of New Jersey, NJ) for ophthalmological evaluation.

The authors declare that no conflict of interest exists.

Author Contributions: PC conceived the study, designed experiments and wrote the paper with input from ES and RB; ES RB JS AI MF and AS conducted the experiments.

REFERENCES

Anderson, K.D., L. Pan, X.M. Yang, V.C. Hughes, J.R. Walls, M.G. Dominguez, M.V. Simmons, P. Burfeind, Y. Xue, Y. Wei, L.E. Macdonald, G. Thurston, C. Daly, H.C. Lin, A.N. Economides, D.M. Valenzuela, A.J. Murphy, G.D. Yancopoulos, and N.W. Gale. 2011. Angiogenic sprouting into neural tissue requires Gpr124, an orphan G protein-coupled receptor. *Proc Natl Acad Sci U S A*. 108:2807-2812.

Basova, L., G.J. Parfitt, A. Richardson, V. Delcroix, T. Umazume, D. Pelaez, D.T. Tse, I. Kalajzic, N. Di Girolamo, J.V. Jester, and H.P. Makarenkova. 2020. Origin and Lineage Plasticity of Endogenous Lacrimal Gland Epithelial Stem/Progenitor Cells. *iScience*. 23:101230.

Bassilana, F., M. Nash, and M.G. Ludwig. 2019. Adhesion G protein-coupled receptors: opportunities for drug discovery. *Nat Rev Drug Discov*. 18:869-884.

Bjarnadottir, T.K., D.E. Gloriam, S.H. Hellstrand, H. Kristiansson, R. Fredriksson, and H.B. Schioth. 2006. Comprehensive repertoire and phylogenetic analysis of the G protein-coupled receptors in human and mouse. *Genomics*. 88:263-273.

Botelho, S.Y. 1964. Tears and the Lacrimal Gland. *Sci Am*. 211:78-86.

Bron, A.J., and J.M. Tiffany. 1998. The meibomian glands and tear film lipids. Structure, function, and control. *Adv Exp Med Biol*. 438:281-295.

Chen, Z., J. Huang, Y. Liu, L.K. Dattilo, S.H. Huh, D. Ornitz, and D.C. Beebe. 2014. FGF signaling activates a Sox9-Sox10 pathway for the formation and branching morphogenesis of mouse ocular glands. *Development*. 141:2691-2701.

Cotsarelis, G., T.T. Sun, and R.M. Lavker. 1990. Label-retaining cells reside in the bulge of the pilosebaceous unit: implications for follicular stem cells, hair cycle, and skin carcinogenesis. *Cell*. 61:1329-1337.

Cui, C.Y., J.A. Smith, D. Schlessinger, and C.C. Chan. 2005. X-linked anhidrotic ectodermal dysplasia disruption yields a mouse model for ocular surface disease and resultant blindness. *Am J Pathol*. 167:89-95.

Davis, F.M., B. Lloyd-Lewis, O.B. Harris, S. Kozar, D.J. Winton, L. Muresan, and C.J. Watson. 2016. Single-cell lineage tracing in the mammary gland reveals stochastic clonal dispersion of stem/progenitor cell progeny. *Nat Commun*. 7:13053.

Dean, C., M. Ito, H.P. Makarenkova, S.C. Faber, and R.A. Lang. 2004. Bmp7 regulates branching morphogenesis of the lacrimal gland by promoting mesenchymal proliferation and condensation. *Development*. 131:4155-4165.

Dean, C.H., L.A. Miller, A.N. Smith, D. Dufort, R.A. Lang, and L.A. Niswander. 2005. Canonical Wnt signaling negatively regulates branching morphogenesis of the lung and lacrimal gland. *Dev Biol*. 286:270-286.

Farmer, D.T., S. Nathan, J.K. Finley, K. Shengyang Yu, E. Emmerson, L.E. Byrnes, J.B. Sneddon, M.T. McManus, A.D. Tward, and S.M. Knox. 2017. Defining epithelial cell dynamics and lineage relationships in the developing lacrimal gland. *Development*. 144:2517-2528.

Fu, J.F., T.H. Yen, Y. Chen, Y.J. Huang, C.L. Hsu, D.C. Liang, and L.Y. Shih. 2013. Involvement of Gpr125 in the myeloid sarcoma formation induced by cooperating MLL/AF10(OM-LZ)

and oncogenic KRAS in a mouse bone marrow transplantation model. *Int J Cancer*. 133:1792-1802.

Gipson, I.K. 2016. Goblet cells of the conjunctiva: A review of recent findings. *Prog Retin Eye Res*. 54:49-63.

Hawley, D., X. Tang, T. Zyrianova, M. Shah, S. Janga, A. Letourneau, M. Schicht, F. Paulsen, S. Hamm-Alvarez, H.P. Makarenkova, and D. Zoukhri. 2018. Myoepithelial cell-driven acini contraction in response to oxytocin receptor stimulation is impaired in lacrimal glands of Sjogren's syndrome animal models. *Sci Rep*. 8:9919.

Kenchegowda, D., S. Swamynathan, D. Gupta, H. Wan, J. Whitsett, and S.K. Swamynathan. 2011. Conditional disruption of mouse Klf5 results in defective eyelids with malformed meibomian glands, abnormal cornea and loss of conjunctival goblet cells. *Dev Biol*. 356:5-18.

Li, X., I. Roszko, D.S. Sepich, M. Ni, H.E. Hamm, F.L. Marlow, and L. Solnica-Krezel. 2013. Gpr125 modulates Dishevelled distribution and planar cell polarity signaling. *Development*. 140:3028-3039.

Makarenkova, H.P., and D.A. Dartt. 2015. Myoepithelial Cells: Their Origin and Function in Lacrimal Gland Morphogenesis, Homeostasis, and Repair. *Curr Mol Biol Rep*. 1:115-123.

Makarenkova, H.P., M. Ito, V. Govindarajan, S.C. Faber, L. Sun, G. McMahon, P.A. Overbeek, and R.A. Lang. 2000. FGF10 is an inducer and Pax6 a competence factor for lacrimal gland development. *Development*. 127:2563-2572.

Marko, C.K., B.B. Menon, G. Chen, J.A. Whitsett, H. Clevers, and I.K. Gipson. 2013. Spdef null mice lack conjunctival goblet cells and provide a model of dry eye. *Am J Pathol*. 183:35-48.

McMahon, A., H. Lu, and I.A. Butovich. 2014. A role for ELOVL4 in the mouse meibomian gland and sebocyte cell biology. *Invest Ophthalmol Vis Sci*. 55:2832-2840.

Nien, C.J., S. Massei, G. Lin, H. Liu, J.R. Paugh, C.Y. Liu, W.W. Kao, D.J. Brown, and J.V. Jester. 2010. The development of meibomian glands in mice. *Mol Vis*. 16:1132-1140.

Park, Y.S., A.E. Gauna, and S. Cha. 2015. Mouse Models of Primary Sjogren's Syndrome. *Curr Pharm Des*. 21:2350-2364.

Pflugfelder, S.C., F. Bian, K. Gumus, W. Farley, M.E. Stern, and C.S. De Paiva. 2018. Severity of Sjogren's Syndrome Keratoconjunctivitis Sicca Increases with Increased Percentage of Conjunctival Antigen-Presenting Cells. *Int J Mol Sci*. 19.

Pflugfelder, S.C., and C.S. de Paiva. 2017. The Pathophysiology of Dry Eye Disease: What We Know and Future Directions for Research. *Ophthalmology*. 124:S4-S13.

Pickering, C., M. Hagglund, J. Szmydynger-Chodobska, F. Marques, J.A. Palha, L. Waller, A. Chodobski, R. Fredriksson, M.C. Lagerstrom, and H.B. Schioth. 2008. The Adhesion GPCR GPR125 is specifically expressed in the choroid plexus and is upregulated following brain injury. *BMC Neurosci*. 9:97.

Plikus, M., W.P. Wang, J. Liu, X. Wang, T.X. Jiang, and C.M. Chuong. 2004. Morpho-regulation of ectodermal organs: integument pathology and phenotypic variations in K14-Noggin engineered mice through modulation of bone morphogenic protein pathway. *Am J Pathol*. 164:1099-1114.

Rios, J.D., K. Forde, Y. Diebold, J. Lightman, J.D. Zieske, and D.A. Dartt. 2000. Development of conjunctival goblet cells and their neuroreceptor subtype expression. *Invest Ophthalmol Vis Sci*. 41:2127-2137.

Schaumberg, D.A., D.A. Sullivan, J.E. Buring, and M.R. Dana. 2003. Prevalence of dry eye syndrome among US women. *Am J Ophthalmol*. 136:318-326.

Schaumberg, D.A., D.A. Sullivan, and M.R. Dana. 2002. Epidemiology of dry eye syndrome. *Adv Exp Med Biol*. 506:989-998.

Seandel, M., D. James, S.V. Shmelkov, I. Falciatori, J. Kim, S. Chavala, D.S. Scherr, F. Zhang, R. Torres, N.W. Gale, G.D. Yancopoulos, A. Murphy, D.M. Valenzuela, R.M. Hobbs, P.P. Pandolfi, and S. Rafii. 2007. Generation of functional multipotent adult stem cells from GPR125+ germline progenitors. *Nature*. 449:346-350.

Simundza, J., and P. Cowin. 2013. Adhesion G-protein-coupled receptors: elusive hybrids come of age. *Cell Commun Adhes*. 20:213-226.

Spiess, K., S.O. Bagger, L.J. Torz, K.H.R. Jensen, A.L. Walser, J.M. Kvam, A.K. Mogelmose, V. Daugvilaite, R.K. Junnila, G.M. Hjorto, and M.M. Rosenkilde. 2019. Arrestin-independent constitutive endocytosis of GPR125/ADGRA3. *Ann N Y Acad Sci*. 1456:186-199.

Tong, L., and P.K. Gupta. 2016. Need for Animal Models of Meibomian Gland Dysfunction. *Ophthalmol Ther*. 5:129-134.

Tsau, C., M. Ito, A. Gromova, M.P. Hoffman, R. Meech, and H.P. Makarenkova. 2011. Barx2 and Fgf10 regulate ocular glands branching morphogenesis by controlling extracellular matrix remodeling. *Development*. 138:3307-3317.

Valenzuela, D.M., A.J. Murphy, D. Frendewey, N.W. Gale, A.N. Economides, W. Auerbach, W.T. Poueymirou, N.C. Adams, J. Rojas, J. Yasenchak, R. Chernomorsky, M. Boucher, A.L. Elsasser, L. Esau, J. Zheng, J.A. Griffiths, X. Wang, H. Su, Y. Xue, M.G. Dominguez, I. Noguera, R. Torres, L.E. Macdonald, A.F. Stewart, T.M. DeChiara, and G.D. Yancopoulos. 2003. High-throughput engineering of the mouse genome coupled with high-resolution expression analysis. *Nat Biotechnol*. 21:652-659.

Vanhollebeke, B., O.A. Stone, N. Bostaille, C. Cho, Y. Zhou, E. Maquet, A. Gauquier, P. Cabochette, S. Fukuhara, N. Mochizuki, J. Nathans, and D.Y. Stainier. 2015. Tip cell-specific requirement for an atypical Gpr124- and Reck-dependent Wnt/beta-catenin pathway during brain angiogenesis. *Elife*. 4.

Wu, Y., W. Chen, L. Gong, C. Ke, H. Wang, and Y. Cai. 2018. Elevated G-Protein Receptor 125 (GPR125) Expression Predicts Good Outcomes in Colorectal Cancer and Inhibits Wnt/beta-Catenin Signaling Pathway. *Med Sci Monit*. 24:6608-6616.

Figure 1. Gpr125 loss induces blepharedema, blepharitis and mucus accumulation.

(A) Schematic of Gpr125 protein comprising N-terminus (N), leucine rich repeats (LRR), Immunoglobulin-like domain (Ig), hormone binding domain (HBD), GPCR autoproteolyis-inducing (GAIN) domain, 7-pass transmembrane region (TM) and cytoplasmic region (C). (**B)** Schematic of *Adgra3*. *Adgra3*$^{cre/cre}$ mice were generated by replacement of 502bp after the first codon with a creER^{T2} module. (**C)** Eye phenotype of *Adgra3*$^{cre/cre}$ compared to controls. **(D)** Examples of blepharedema and mucus accumulation in *Adgra3*$^{cre/cre}$ mice. **(E)** Schematic of the Gpr125-β-gal protein generated by deletion of 10 kb sequence downstream of the first TM and replacement by *lacZ*. **(F)** Eye phenotype of *Adgra3*$^{lz/lz}$ mice compared to controls.

Figure 2. Gpr125 is expressed in eyes and eyelids.

(A) Diagram of murine eye. (**B)** Section of X-gal stained *Adgra3*$^{lz/lz}$ eye shows Gpr125-β-gal expression in the ciliary body and iris. (**C)** Intraocular pressure (IOP) (mmHg) in male (blue) and female (pink) *Adgra3*$^{lz/lz}$ and *Adgra3*$^{cre/cre}$ mice compared to their respective FVBN and B6 controls. Each bar represents the mean ± SEM on 6-14 mice/group. ns, not significant. (**D)** Fluorescein stained corneas in *Adgra3*$^{cre/cre}$ and control mice. n=3 (**E)** Schematic of tear film. (**F)** Eyelid sections stained with alcian blue AB/PAS show goblet cells in *Adgra3*$^{cre/cre}$ and control mice. n=23. **(G-H)** Sections of eyelids showing meibomian glands stained with **(G)** H/E and **(H)** immunostained with antibodies: F480, CD4,CD8, CK5 and DAPI to detect macrophages, T-helper, cytotoxic T cells, cytokeratin 5 and nuclei respectively in *Adgra3*$^{cre/cre}$ mice. Control *Adgra3*$^{+/+}$ in S1A. **I)** X-gal stained whole

mounts of P10 eyelids from *Adgra3*$^{lz/lz}$ mice showing meibomian glands (black arrowheads) devoid of Gpr125 and strong expression in eyelash follicles (red arrows) Scale bar 100μm. n=3.

Figure 3. Gpr125 cells, located at ductal tips during development, function as lacrimal myoepithelial progenitors.

(A) Increased tear production observed in *Adgra3*$^{cre/cre}$ and *Adgra3*$^{lz/lz}$ male (blue) and female (pink) mice compared to controls. Each bar represents the mean ± SEM on 6-32 mice. **** $p<0.0001$, **, $p<0.05$ value significant; ns, not significant. **B)** t-SNE plot of cells clusters within E16 lacrimal glands (10). Zoomed images of E16 epithelial compartment (boxed region) show cells expressing Adgra3 mRNA also express myoepithelial markers, Keratin14 and Sox10 but not luminal markers Keratin19 or Aquaporin5. **C)** Gpr125-β-gal expression in embryos and in the bulge **(D)** of E14 whisker follicle stained with p63. **E)** Strategy for tracing the lineage of Gpr125-positive cells in E14.5-E15.5 embryos carrying the *Rosa26*.lox.STOP.lox.TdTomato reporter by tamoxifen injection of pregnant *Adgra3*$^{cre/cre}$ dams. 3D-confocal images of lacrimal glands from mice at **(F)** 7 weeks and **(G)** 6 months showing tdT expression in elongated myoepithelial cells along the basal border of ducts and stellate cells enmeshing acini colocalized with myoepithelial marker (K5). Scale bar 50μm.n=3.

Figure 4. Loss of Gpr125 leads to abnormal lacrimation and inflammatory infiltration of the lacrimal glands

(A) H/E section of lacrimal gland from *Adgra3*$^{cre/cre}$ mice shows foci of infiltration (arrows). **(B)** Control *Adgra3*$^{+/+}$. **(C-G)** Immunofluorescence of lacrimal gland co-stained for **(D)** K5, **(E)** macrophages, **(F)** T-helper, **(G)** cytotoxic T cells. **(H)** *Adgra3*$^{cre/cre}$ female with lacrimal mass sectioned and stained with H/E in **(I,J)**. **(K)** Immunofluorescence analysis as described above of boxed region in **I**. Scale bar 100μm.

Figure S1.

(A,B) Meibomian gland immunostained with antibodies: F480, CD4, CD8, CK5 and DAPI to detect macrophages, T-helper, cytotoxic T cells, cytokeratin 5 and nuclei respectively in control *Adgra3*$^{+/+}$ and *Adgra3*$^{cre/cre}$ mice. Scale bar 100μm.

Figure 1

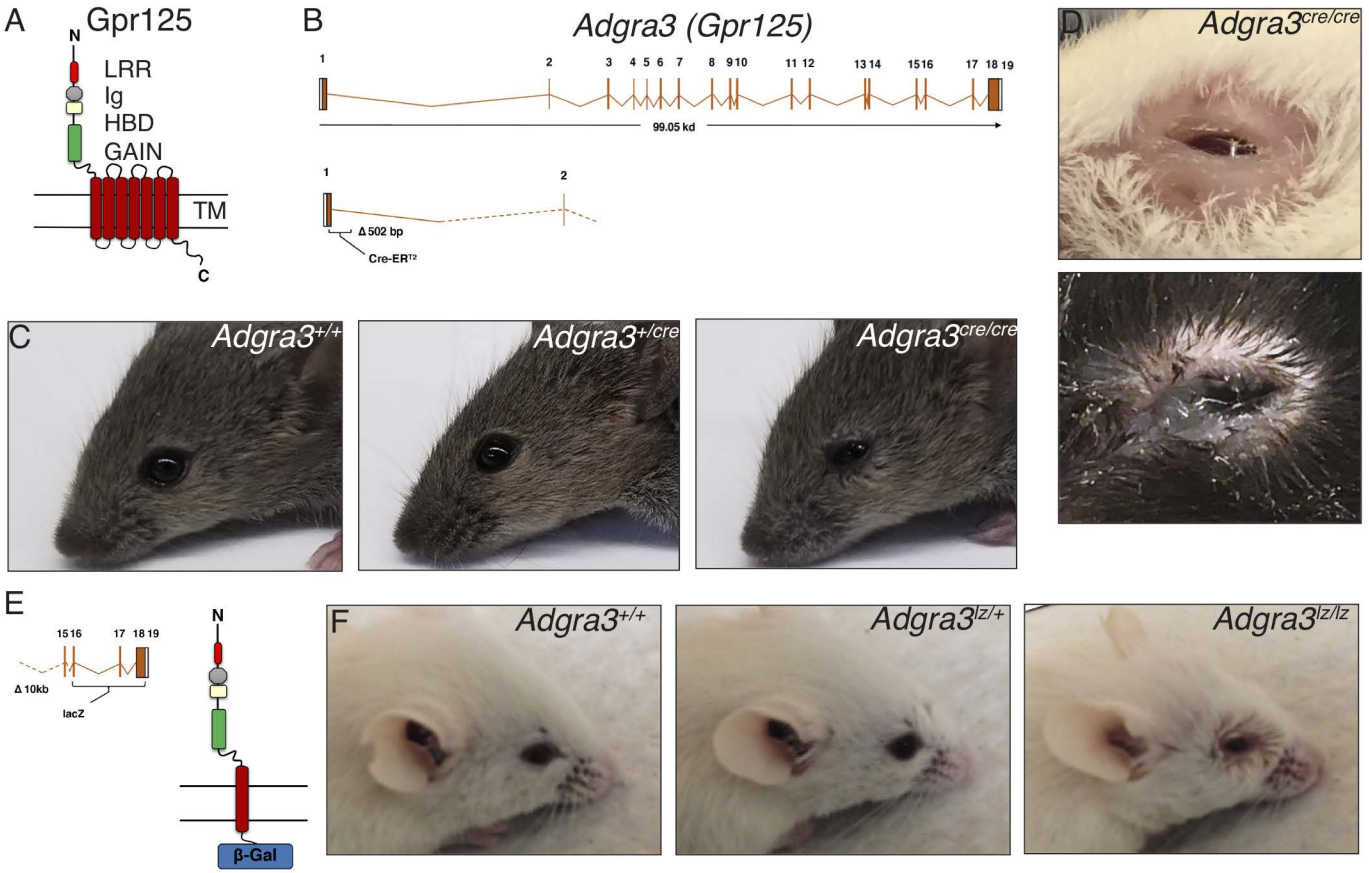

A
Gpr125
N
LRR
Ig
HBD
GAIN
TM
C
B
Adgra3 (Gpr125)
1
2
3
4 5 6
7
8 9 10
11 12
13 14
15 16
17 18 19
99.05 kd
1
2
Δ 502 bp
Cre-ER^{T2}
D
Adgra3$^{cre/cre}$
C
Adgra3$^{+/+}$
Adgra3$^{+/cre}$
Adgra3$^{cre/cre}$
E
15 16
17 18 19
Δ 10kb
lacZ
N
β-Gal
F
Adgra3$^{+/+}$
Adgra3$^{lz/+}$
Adgra3$^{lz/lz}$

Figure 2

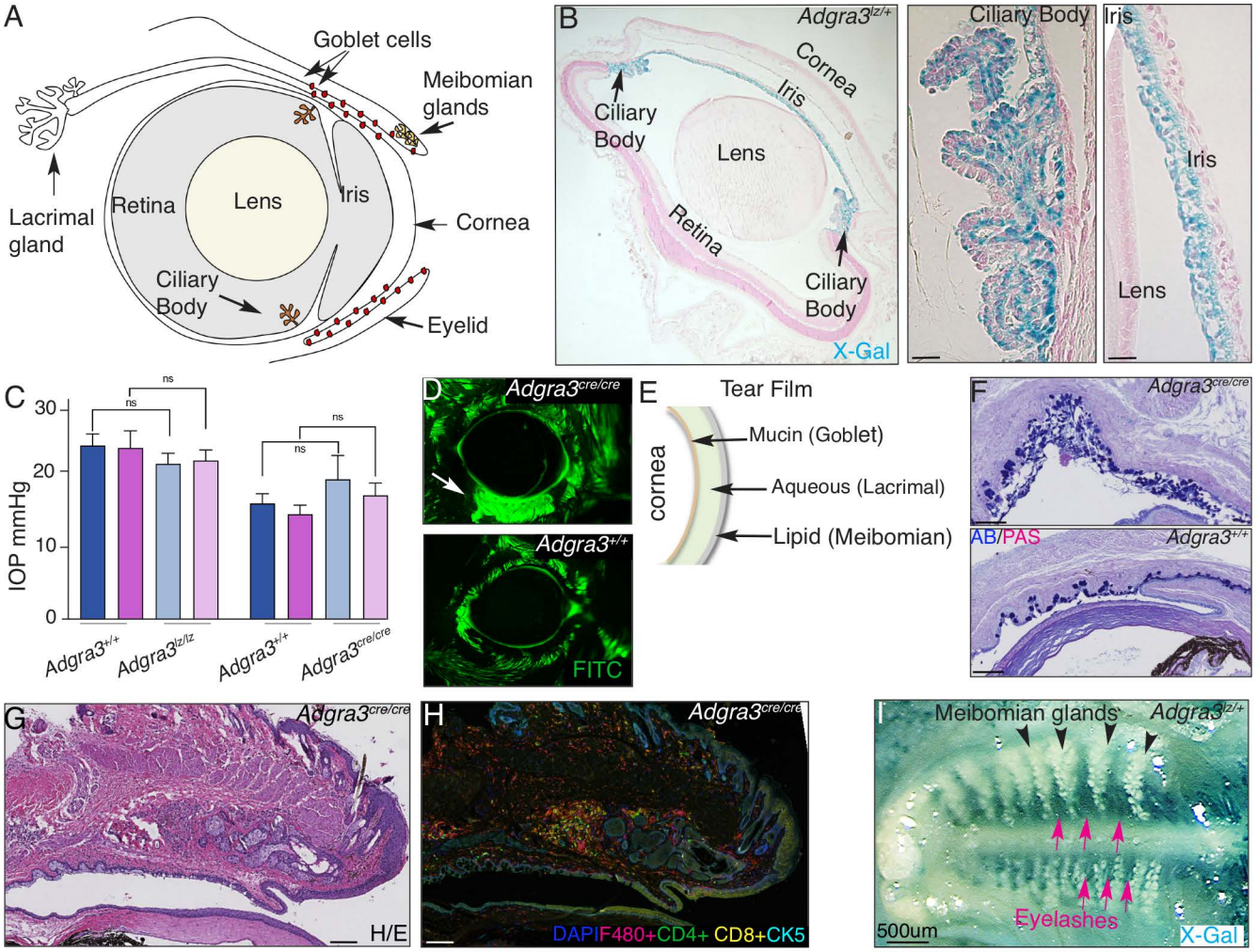

A
Goblet cells
Meibomian glands
Lacrimal gland
Retina
Lens
Iris
Cornea
Ciliary Body
Eyelid
B
Adgra3lz/+
Cornea
Iris
Ciliary Body
Lens
Retina
Ciliary Body
X-Gal
Ciliary Body
Iris
Iris
Lens
C
IOP mmHg
30
20
10
0
ns
ns
ns
ns
Adgra3+/+
Adgra3lz/lz
Adgra3+/+
Adgra3cre/cre
D
Adgra3cre/cre
Adgra3+/+
FITC
E
Tear Film
cornea
Mucin (Goblet)
Aqueous (Lacrimal)
Lipid (Meibomian)
F
Adgra3cre/cre
AB/PAS
Adgra3+/+
G
Adgra3cre/cre
H/E
H
Adgra3cre/cre
DAPIF480+CD4+ CD8+CK5
I
Meibomian glands
Adgra3lz/+
Eyelashes
500um
X-Gal

Figure 3

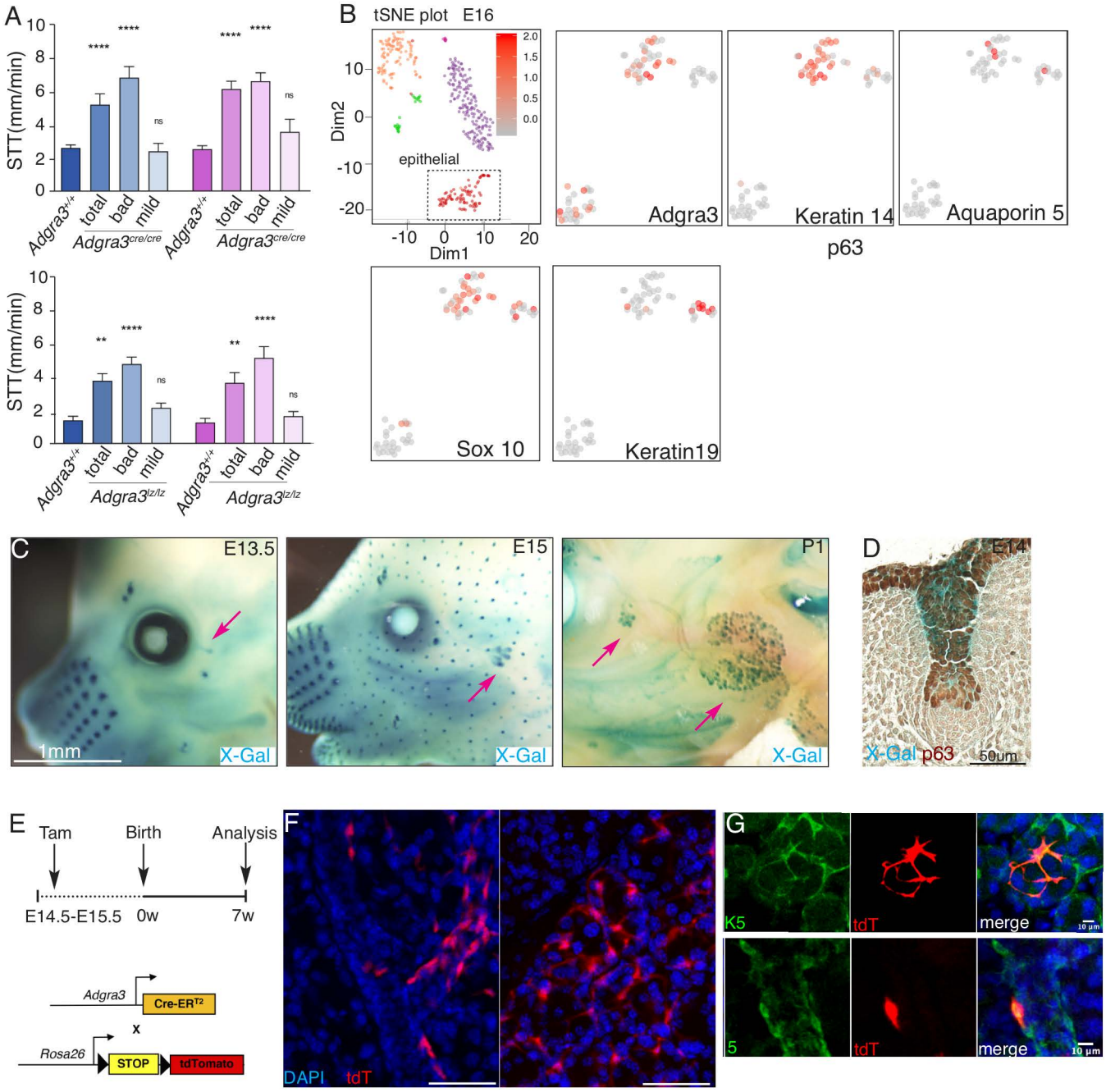

A
STT(mm/min)
Adgra3+/+
total
bad
mild
Adgra3cre/cre
Adgra3lz/lz
B
tSNE plot E16
Dim1
Dim2
epithelial
Adgra3
Keratin 14
Aquaporin 5
p63
Sox 10
Keratin19
C
E13.5
E15
P1
1mm
X-Gal
D
E14
X-Gal p63
50um
E
Tam
Birth
Analysis
E14.5-E15.5
0w
7w
Adgra3
Cre-ERT2
x
Rosa26
STOP
tdTomato
F
DAPI
tdT
G
K5
tdT
merge
5
10 μm

Figure 4

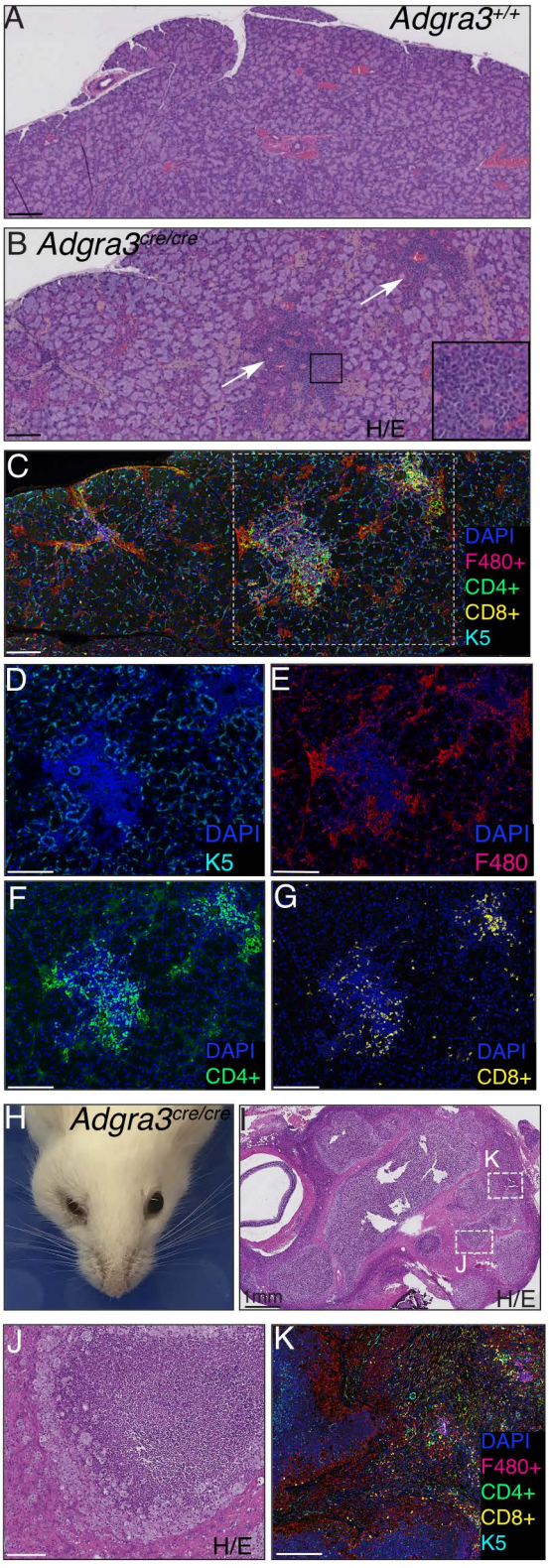
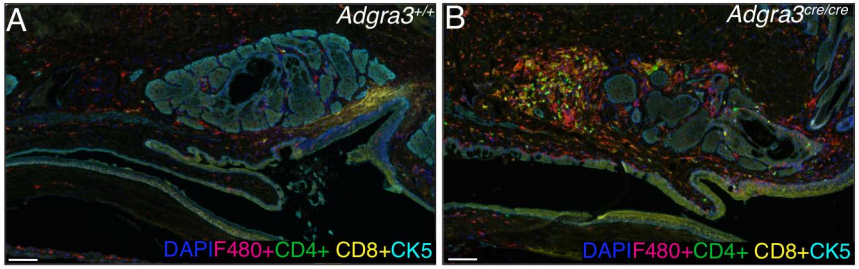

A
Adgra3+/+
B Adgra3cre/cre
H/E
C
DAPI
F480+
CD4+
CD8+
K5
D
DAPI
K5
E
DAPI
F480
F
DAPI
CD4+
G
DAPI
CD8+
H
Adgra3cre/cre
I
K
J
1mm
H/E
J
H/E
K
DAPI
F480+
CD4+
CD8+
K5

Figure S1


A
Adgra3+/+
DAPIF480+CD4+ CD8+CK5
B
Adgra3cre/cre
DAPIF480+CD4+ CD8+CK5