ARTICLE

Identification of a Zeb1 expressing basal stem cell subpopulation in the prostate

Xue Wang[1,2], Haibo Xu[3,4], Chaping Cheng[1,2], Zhongzhong Ji [1,2], Huifang Zhao[1], Yaru Sheng[1], Xiaoxia Li[1], Jinming Wang[1], Yu Shu[1], Yuman He[1], Liancheng Fan[5], Baijun Dong [5], Wei Xue[5], Chee Wai Chua [1,5], Dongdong Wu [3], Wei-Qiang Gao[1,2,6]* & Helen He Zhu [1,5,6]*

The basal cell compartment in many epithelial tissues is generally believed to serve as an important pool of stem cells. However, basal cells are heterogenous and the stem cell subpopulation within basal cells is not well elucidated. Here we uncover that the core epithelial-to-mesenchymal transition (EMT) inducer Zeb1 is expressed in a prostate basal cell subpopulation. The Zeb1+ prostate epithelial cells are multipotent prostate basal stem cells (PBSCs) that can self-renew and generate functional prostatic glandular structures at the single-cell level. Genetic ablation studies reveal an indispensable role for Zeb1 in prostate basal cell development. Utilizing unbiased single-cell transcriptomic analysis of over 9000 mouse prostate basal cells, we confirm the existence of the Zeb1+ basal cell subset. Moreover, Zeb1+ epithelial cells can be detected in mouse and human prostate tumors. Identification of the PBSC and its transcriptome profile is crucial to advance our understanding of prostate development and tumorigenesis.

[1] State Key Laboratory of Oncogenes and Related Genes, Renji-Med-X Stem Cell Research Center, Department of Urology, Ren Ji Hospital, School of Medicine and School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200127, China. [2] Med-X Research Institute, Shanghai Jiao Tong University, Shanghai 200030, China. [3] State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Sciences, Kunming 650223, China. [4] University of Chinese Academy of Sciences, Beijing 100049, China. [5] Department of Urology, Ren Ji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200127, China. [6] These authors contributed equally: Wei-Qiang Gao, Helen He Zhu. *email: gao.weiqiang@sjtu.edu.cn; zhuhecrane@shsmu.edu.cn

The prostatic epithelium, comprised of basal cells, secretory luminal cells and rare neuroendocrine cells, can serve as an excellent model to study stem cell biology due to its ability to regress and regenerate after repeated rounds of androgen deprivation and restoration in mice and rats[1,2]. Using in vivo subrenal capsule tissue regeneration, cell division mode analyses or in vitro organoid-forming assays, we and others have demonstrated the existence of prostate stem/progenitor cells in both basal and luminal prostate epithelia[3–10]. Importantly, recent in situ lineage tracing studies conducted in both mice and humans have provided further solid evidence for prostate stem/progenitors[11–16]. The basal cell compartment in many epithelial tissues including the epidermis[17,18], colorectum[19], bladder[20,21] and mammary gland[22,23] has been reported to enrich with stem cells. Using the K14rtTA/TetOCRE/RosaYFP or K5CREER/RosaYFP pan-basal cell lineage tracing mice, it has been shown that multipotent basal stem cells can contribute to the differentiation of prostate basal, luminal and neuroendocrine cell lineages especially during early postnatal development[13]. From a study of quantitative clonal mapping and three-dimensional reconstruction of adult human prostates, Moad et al. recently reported that multipotent basal stem cells are exclusively located in juxta-urethral niches and function in a directed migratory way to generate epithelial progenitors[24]. In contrast, using pan-basal or pan-luminal cell lineage tracing approaches, it is shown that basal and luminal cell lineages are independent self-sustained during prostate regeneration in adult mice[12]. On the other hand, despite the long-term postulation that basal cells contain more primitive prostate stem cells due to their resistance to castration, capability to differentiate into basal, luminal and neuroendocrine cell lineages of the prostate epithelium, and susceptibility to oncogenic transformation[4,5,13,25–29], the identity and nature of the prostate basal stem cells (PBSCs) within the heterogenous basal cell epithelia have not been elucidated. In addition, the existing markers for basal stem cells (CD117, CD133, CD44, Trop2, CD49f, Sca1, etc.)[4–6] are found in both the basal and luminal cell compartments.

In the present study, we identify a PBSC subpopulation that expresses Zeb1, an important EMT inducer[30,31], through both in vitro and in vivo functional, lineage tracing and genetic ablation analyses. Furthermore, we integrated single-cell RNA sequencing with computational data analysis, a powerful approach that is lately developed to map cellular diversity of a given tissue, i.e., cortical neurons[32] and cancer related immune cells[33], etc., to examine the prostate basal cell compartment. Our single-cell transcriptomics data provide additional supporting evidence for the existence of Zeb1-expressing PBSCs and uncover their gene expression profile.
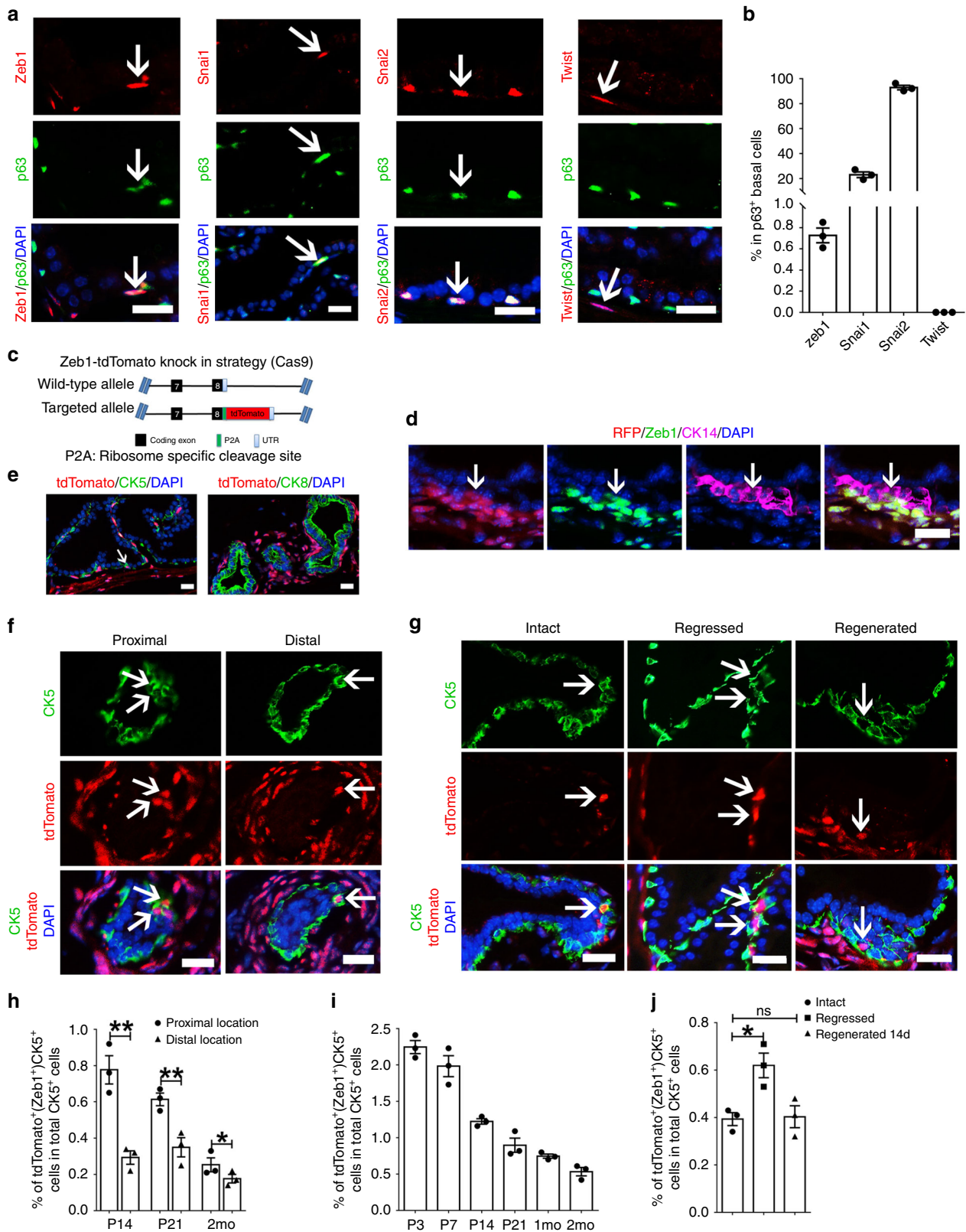
Results

Zeb1 is expressed in a prostate basal cell subpopulation. EMT has been previously considered to occur only in early embryonic development or pathological conditions, including tumor metastasis and wound healing[34]. Expression and function of core transcriptional factors such as Zeb1/2, Snai1/2, Twist1/2 which are required for EMT induction, have not been examined in normal prostate epithelia under physiological condition. Compared to prostate luminal cells, basal cells expressed low levels of epithelial-cell specific genes such as miRNA200 family[35] and E-cadherin[7] and exhibited a more mesenchymal-like phenotype. We therefore asked whether core EMT inducers are present in normal prostate basal cells. Immunofluorescent analysis of prostate sections showed that less than 1% of basal cells were positively labeled for Zeb1 immunostaining (Fig. 1a, b). Snai1 or Slug (Snai2) was expressed in 20% or 90% of p63+ basal cells,

respectively, while Twist1/2 positive staining was not detected in the basal layer (Fig. 1a, b). Considering the interesting expression pattern of Zeb1 in the prostate epithelium and the previously reported role of Zeb1 in stemness acquisition and maintenance of cancer stem cells (CSC) characteristics[31,36,37], we decided to investigate the biological relevance of Zeb1+ basal cells in the prostate.

To better characterize and to more easily enrich the Zeb1+ prostate epithelial cells, we established a tdTomato knockin mouse model in which the coding sequence of tdTomato fluorescent reporter was linked to the last exon of Zeb1 by a P2A element (Fig. 1c). Using immunofluorescent triple-staining of RFP, Zeb1 and CK14 on mouse prostate sections, we confirmed that the tdTomato labeling faithfully reflected the endogenous expression of Zeb1 (Fig. 1d). TdTomato positive cells were only found in prostate basal cells (marked by CK5 immunostaining) but not from luminal cell compartment (labeled by CK8 immunostaining) (Fig. 1e). The Zeb1+/tdTomato+ prostatic basal cells were more frequently detected in the urethra-proximal region relative to the distal region, the location where prostate stem cells were suggested to reside[4,38,39] (Fig. 1f, h). Moreover, Zeb1+ basal cells located in the proximal region were more proliferative than those in the distal region (Supplementary Fig. 1a, b). In addition, we found that the percentage and absolute number of Zeb1+/tdTomato+ basal cells declined as the prostate development proceeded (Fig. 1i and Supplementary Fig. 1c). We then examined the dynamics of Zeb1+ basal cells during prostate regression and regeneration. We first checked the impact of castration and androgen replacement on bulk CK5+ basal cells and observed a moderate decrease of total basal cell number upon castration and a recovery of basal cell number after regeneration (Supplementary Fig. 2a). In contrast, as shown in Fig. 1g, j and Supplementary Fig. 2b–d, the proportion and absolute number of the Zeb1+ CK5+ population moderately augmented in regressed prostates, and then decreased to the intact prostate level after regeneration. We detected an increase of Ki67 positive cells in Zeb1+ basal cells from the proximal region of prostates in castrated mice, suggesting that the increase of Zeb1+ basal cell number may at least be partially contributed from cell proliferation (Supplementary Fig. 2e, f).

Zeb1+ basal cells are enriched for prostate basal stem cells. To test the role of Zeb1+ basal cells in prostate development, we performed a prostate organoid-forming assay in vitro using flow cytometry sorted Lineage− Sca-1+ CD49f^high (LSC)Zeb1+ and LSCZeb1− cells (Fig. 2a, b). While few and small organoids were produced from LSCZeb1− cells, significantly larger and more organoids were generated from LSCZeb1+ cells (Fig. 2c, e). Immunostaining analysis of frozen sections of organoids formed from sorted LSCZeb1+ cells showed generation of both basal (CK5, CK14 or p63 positive) and luminal (CK8 or AR positive) cells (Fig. 2d). Moreover, LSCZeb1+ cells possessed a serial organoid forming capacity, indicating a self-renewing characteristic (Fig. 2e). According to the gross appearance and the H&E staining of organoid sections, LSC cell-derived organoids can be divided into three types, the acinar, compact or lumen phenotypes (Fig. 2f, g). On the other hand, we performed organoid sectioning and immunostaining to further analyze the organoid phenotype and observed that both basal-only (p63+ CK8−) and multipotent (p63+ CK8+) organoids can be generated from prostate LSC cells (Fig. 2h, i). We could detect both tdTomato+ (Zeb1+) and tdTomato− (Zeb1−) cells in the organoids derived from LSCZeb1+ cells. The percentage of tdTomato+ (Zeb1+) cells remained stable along serial passages (Source data file), suggesting that LSCZeb1+ cells can undergo self-renewal and

differentiation. In addition, we performed organoid forming assays using LSCZeb1[+] and LSCZeb1[−] cells under androgen deprived condition. The number of organoids formed from Zeb1[+] cells (especially with the multipotent phenotype) was significantly more than organoids formed from Zeb1[−] cells following androgen deprivation (Supplementary Fig. 2g–i).

We then evaluated the potential of LSCZeb1[+] cells to generate prostates in vivo, a gold standard to assess the stem cell phenotype[4,40,41]. Using an in vivo kidney capsule transplantation assay, we grafted 1000 LSCZeb1[+] or LSCZeb1[−] mouse prostate cells in combination with rat embryonic urogenital sinus mesenchymal (UGM) stromal cells under the renal capsule of

Fig. 1 Zeb1+ basal cells are more frequently detected in the urethra-proximal region. a Co-immunostaining for Zeb1, Snai1, Snai2 (Slug), or Twist1/2 together with p63 on frozen sections of p21 wild-type mouse prostates. (Scale bars = 20 μm). **b** Quantification of the ratio of Zeb1+, Snai1+, Snai2+ or Twist1/2+ cells in p63+ prostatic basal epithelial cells. ($n = 3$ mice. Source data are provided as a Source Data file.) **c** Illustration of the strategy to generate the Zeb1-tdTomato reporter mouse model. **d** Triple staining of Zeb1, RFP and CK14 on a prostate section from Zeb1-tdTomato reporter mouse confirms that the td-Tomato labeling faithfully indicated the endogenous expression of Zeb1. (Scale bars = 20 μm). **e** Immunostaining of tdTomato and CK5 or CK8 on prostate frozen sections of 3-week-old wild type mice showing tdTomato expression is only found in prostate basal cells but not luminal cells. (Scale bars = 20 μm). **f** The tdTomato (Zeb1)+ prostate basal cells are more frequently found in the urethra-proximal region compared to the distal region. (Scale bars = 20 μm). **g** Section imaging of intact, regressed (21 days after castration) and regenerated (14 days after androgen replacement) prostate sections from Zeb1-tdTomato reporter mice. (Scale bars = 20 μm). **h** Quantification of tdTomato (Zeb1)+ prostate basal cells in urethra-proximal and distal regions at indicated developmental stages. ($n = 3$ mice. Data are analyzed by Student's t-test and are presented as mean + s.e.m. *$p < 0.05$, **$p < 0.01$. Source data are provided as a Source Data file.). **i** Quantification of tdTomato (Zeb1)+ prostate basal cells at indicated time points of prostate development. ($n = 3$ mice. Source data are provided as a Source Data file.). **j** Quantification of tdTomato (Zeb1)+ prostate basal cells at indicated time points of prostate regeneration. ($n = 3$ mice. Data are analyzed by Student's t-test and are presented as mean + s.e.m. *$p < 0.05$. Source data are provided as a Source Data file.) (In this figure, five fields per lobes (including the anterior, ventral and dorsolateral lobe) from 3 sections per mice of 3 mice were analyzed.).

host athymic nude mice. Compared to opaque and small grafts derived from LSCZeb1− cells, LSCZeb1+ cells generated semi-translucent and large grafts (Fig. 3a–c). Histological analysis indicated that Zeb1+ grafts underwent ductal morphogenesis with differentiation of basal (CK5+), luminal (CK8+) and rare neuroendocrine (Syp+) cell lineages (Fig. 3d, e). Pbsn expression further confirmed differentiation of functional secretory luminal cells (Fig. 3e). In addition, Zeb1+ grafts expressed the mouse β-integrin, which validated their mouse origin and indicated that they were not produced from the contamination of rat epithelial cells from the preparation of rat UGM cells (Fig. 3e). We further verified that the generated prostate tissues were originated from Zeb1+ donor cell grafts by positive staining of an MHC class I haplotype H-2k^b protein that is specifically expressed in C57BL/6 but not host athymic mouse cells (Fig. 3e). The unstructured tissue generated from LSCZeb1− cells was comprised of basal cells (CK5+), rare luminal cells (CK8+) and implanted rat urogenital sinus mesenchymal cells (mouse β-integrin− and Vimentin+) (Fig. 3f, g). These data suggested that LSCZeb1+ cells are multipotent prostate stem cells and possess the capacity of generating functional prostates.
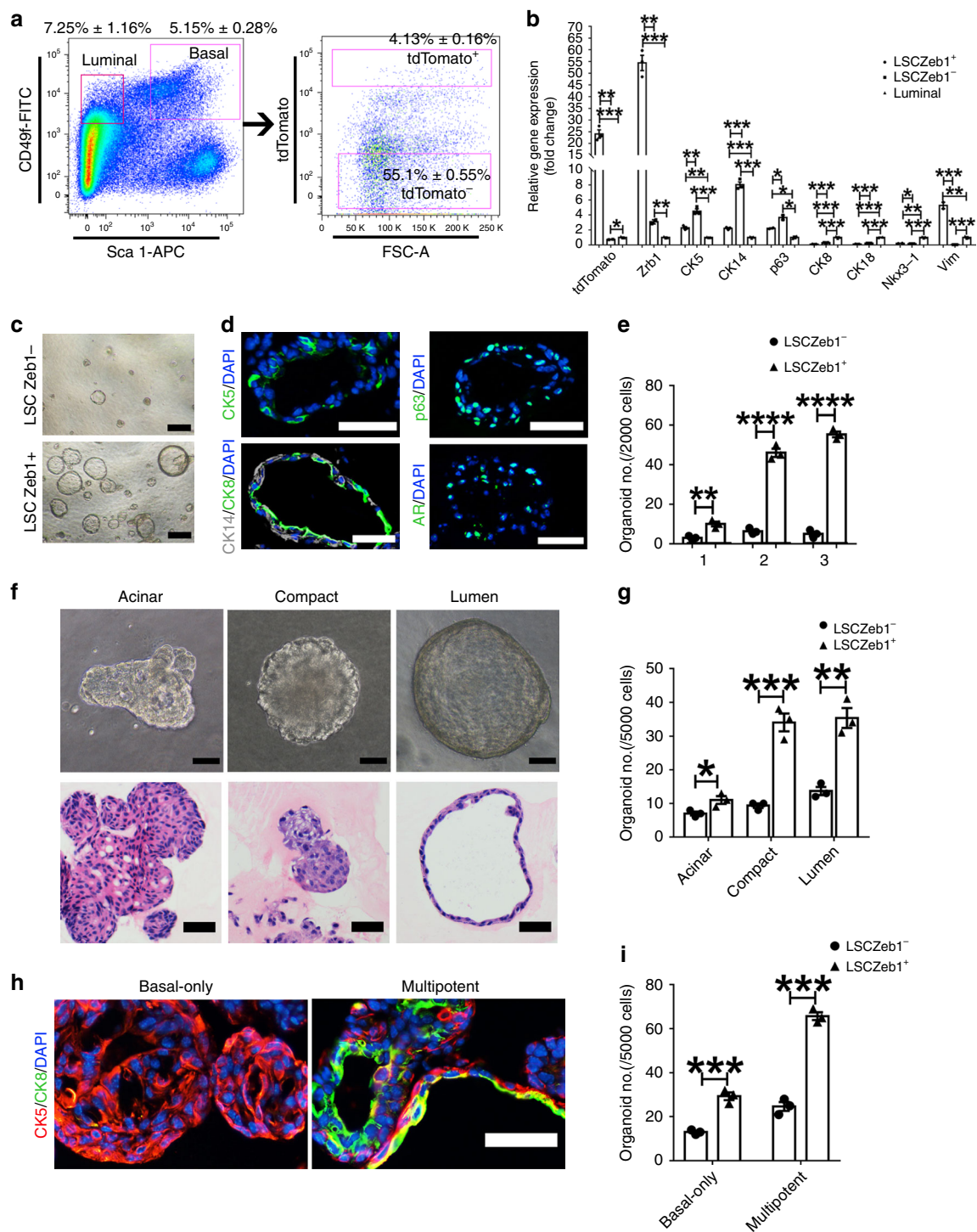
Single Zeb1+ basal cell can generate functional prostate. We then wondered whether a single Zeb1+ basal cell would generate a prostate in vivo. Single viable LSCZeb1+ or LSCZeb1− cell was FACS sorted into individual well of a 96-well plate, mixed with rat UGM stromal cells and transplanted into the renal capsules of nude mice. Each well was examined under the microscope to confirm the presence of a tdTomato+ or tdTomato− single cell (Fig. 3h). Fifteen prostates with well-differentiated prostate epithelial tubules containing all three basal, luminal and neuroendocrine cell lineages were generated from 36 single LSCZeb1+ cell transplants (Fig. 3i–l). In contrast, only 3 prostates were produced from 33 single LSCZeb1− cell xenografts (Fig. 3i–l). Furthermore, Zeb1+ Epcam+ cells harvested from the first generation of xenografted prostate tissues were able to form functional prostates in the secondary and tertiary sub-renal capsule transplantations (Fig. 3m, n). The percentage of tdTomato+ (Zeb1+) cells remained stable along serial implantation, suggesting that the LSCZeb1+ cell can undergo self-renewal and differentiation (Fig. 3o, p). Collectively, these data provided additional direct evidences that LSCZeb1+ cells represented a multipotent prostate stem cell population.

Lineage tracing supports multipotency of Zeb1+ basal cells. To further determine the role of Zeb1+ basal cells in prostate development in vivo, we generated a *Zeb1-CreERT2* mouse line and crossed it with *Rosa-CAG-LSL-tdTomato* mice to trace Zeb1+

basal cells during the postnatal prostate development (Fig. 4a, b). Tamoxifen administration to postnatal day 3 *Zeb1-CreERT2/ tdTomato* mice induced expression of tdTomato in 10% of Zeb1+ cells (Fig. 4c). Consistent with our immunostaining results and findings from the Zeb1-tdTomato knockin mouse model (Fig. 1), about 0.33% of CK5+ basal cells were labeled by tdTomato but none of the CK8+ luminal cells nor Syp+ neuroendocrine cells were labeled at 2 days after induction (Fig. 4d–f, i). Interestingly, 12 days after the tamoxifen administration, we found clusters of CK5+ CK8− tdTomato+ cells located in the outer basal cell layer as well as CK5− CK8+ tdTomato+ cells in the inner luminal compartment (Fig. 4g, i). Quantitative analysis exhibited that labeled basal cells expanded more than 16 folds and tdTomato expressing luminal cells increased from zero to around 4% in 12 days. Meanwhile, we also observed that about 3% of Syp+ neuroendocrine cells were marked with tdTomato at 12 days post induction (Fig. 4h). Those results further substantiated the notion that Zeb1+ basal cells were able to generate basal, luminal and neuroendocrine cell lineages.

To explore the role of Zeb1+ prostate basal cells during tissue regression and regeneration, we performed lineage tracing experiments on Zeb1 expressing epithelial cells in adult mice. Tamoxifen was given to adult Zeb1-CreERT2/tdTomato mice to induce the expression of tdTomato in Zeb1+ cells (Supplementary Fig. 3a). As shown in Supplementary Fig. 3b–d, about 0.34% of CK5+ basal cells were labeled by tdTomato but none of the CK8+ luminal cells nor Syp+ neuroendocrine cells were labeled at 2 days after induction. Mice were then castrated at 1 week after the injection of tamoxifen and reconstituted with testosterone pellets at 3 weeks post castration. Unlike the multilineage differentiation of Zeb1+ basal cells in early development, we found that Zeb1 + basal cells gave rise to mostly basal cells but not luminal or neuroendocrine cells based on lineage tracing analysis in prostate regeneration (Supplementary Fig. 3e, f).

In addition, we performed qRT-PCR assays for previously reported surface markers of stem/progenitor cells in the prostate or mammary gland[4,6,24,42–44]. We found that CD117, CD133 and Procr were expressed with an enrichment of more than 2 folds in the LSCZeb1+ versus the LSCZeb1− cells (Supplementary Fig. 4a). The expression levels of CD166, Trop2 and Dlk1 were slightly higher in the LSCZeb1+ than the LSCZeb1− cells (Supplementary Fig. 4a). Additionally, we conducted flow cytometric staining of LSCZeb1+ cells and found that CD117, CD133, Procr, Trop2, and Dlk1 were presented in 23.18%, 49.83%, 69.38%, 21.43%, 12.43% of the LSCZeb1+ cells respectively (Supplementary Fig. 4b). A combination of the top expressing markers in LSCZeb1+ cells, that are CD133 and Procr, cannot very well isolate the LSCZeb1+ cell population (Supplementary Fig. 4c).

Zeb1 is required for development of prostatic basal cells. We then asked whether Zeb1 was functionally required for normal prostate development. To that end, we first utilized the clustered regularly interspaced short palindromic repeats (CRISPR)/ CRISPR-associated protein 9 (Cas9) system to achieve *Zeb1* gene knockout in prostate organoids in vitro (Fig. 5a). Efficient *Zeb1* deletion in mouse primary prostate epithelial cells was validated by immunoblotting (Fig. 5b). Zeb1 sgRNA transfected prostate organoids contained CK8+ luminal cells, however, Zeb1 knock-out organoids displayed a marked decrease of p63 expressing basal cells compared to organoids formed from mock-transfected prostate epithelial cells (Fig. 5c, d). We next investigated the role

of Zeb1 in prostate development in vivo. Due to the fact that Zeb1 knockout mice died shortly after birth[45], UGS from E16 Zeb1−/− embryos were dissected and transplanted beneath the renal capsule of male athymic mice to allow us to assess prostate epithelial development with Zeb1 deletion (Fig. 5e). Although Zeb1−/− UGS grafts developed into prostate and contained prostatic ductal structures, Zeb1 null grafts were smaller (Fig. 5e, f). Detailed histological examination revealed that approximately 50% of the Zeb1−/− UGS derived prostate epithelia virtually displayed all luminal cells that are positive of CK8 and AR without basal cells. In the other prostate epithelia, basal cell numbers were greatly reduced (Fig. 5f–h). Further immunofluorescent co-staining with

Fig. 2 LSCZeb1⁺ cells are enriched for prostate stem cells. a, b FACS and qRT-PCR quantification of Zeb1, canonical basal and luminal markers expression in Lineage⁻ Sca-1⁺ CD49f^hi tdTomato⁺, Lineage⁻ Sca-1⁺ CD49f^hi tdTomato⁻ prostate basal cell-enriched cells and Lineage⁻ Sca-1⁻ CD49f^lo luminal cells. RNA expression levels were normalized to Lineage⁻ Sca-1⁻ CD49f^lo luminal cells. The absolute number: basal cells, 14822 ± 966; luminal cells, 21312 ± 1215; tdTomato⁺ (Zeb1⁺) basal cells, 923 ± 67; tdTomato⁻ (Zeb1⁻) basal cells, 157265 ± 3434. ($n = 3$ independent experiments. Data are analyzed by Student's t-test and are presented as mean + s.e.m. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.). **c–e** In vitro serial organoid-forming assay showing more and larger organoids were produced from LSCZeb1⁺ cells compared to LSCZeb1⁻ cells. (Scale bars = 200 μm for bright-field images and 50 μm for the immunofluorescent images. The number 1,2,3 stands for the first, secondary and tertiary passaging. $n = 3$ replicates. Data are analyzed by Student's t-test and are presented as mean + s.e.m. **$p < 0.01$, ****$p < 0.0001$. Source data are provided as a Source Data file.). **f** Phase contrast and H & E staining of sections from LSC cell-derived organoids showing three different phenotypes, the acinar, compact and lumen type. (Scale bars = 200 μm for bright-field images and 20 μm for the immunofluorescent images.). **g** Quantification of the number of organoids with different phenotypes formed from LSCZeb1⁺ or LSCZeb1⁻ prostate cells. ($n = 3$ replicates. Data are analyzed by Student's t-test and are presented as mean + s.e.m. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Source data are provided as a Source Data file.). **h** Immunostaining of LSC cell-derived organoids (Scale bars = 50 μm). **i** Quantification of organoids with basal-only or multipotent phenotypes generated from LSCZeb1⁺ or LSCZeb1⁻ prostate cells. ($n = 3$ replicates. Data are analyzed by Student's t-test and are presented as mean + s.e.m. ***$p < 0.001$. Source data are provided as a Source Data file.).
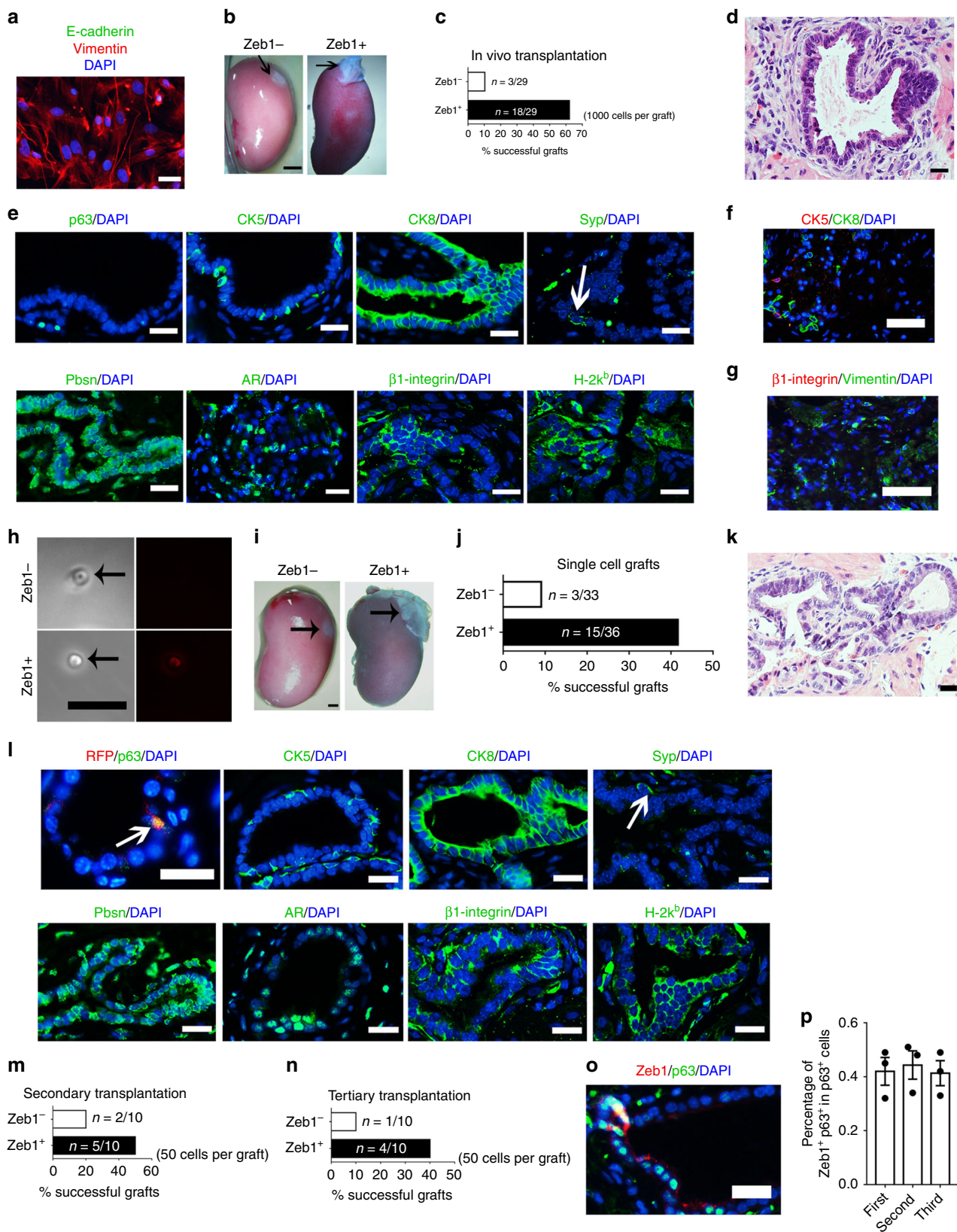
Zeb1 and p63 antibodies affirmed Zeb1 deletion and marked decrease of basal cells in Zeb1⁻/⁻ UGS grafts (Fig. 5g, h). These findings were further supported with additional staining of other basal cell markers CK5 and CK14 (Fig. 5g, h). Syp-expressing neuroendocrine cells could still be found in Zeb1⁻/⁻ UGS derived prostate epithelium (Fig. 5g). Collectively, our in vitro and in vivo data highlight an indispensable role for Zeb1 in prostate basal cell development.

Single cell RNA-seq confirms the Zeb1⁺ basal cell subset. To provide additional evidence for the existence of a Zeb1⁺ basal population, we performed unbiased single cell transcriptomic analysis of prostate basal cell-enriched LSC cell population. Using the 10x Genomics single cell system followed by Illumina sequencing, prostate LSC cells were purified from mouse prostates via FACS for single-cell RNA sequencing. Sorted LSC cells were confirmed by q-RT-PCR and immunostaining for expression of basal cell markers (Supplementary Fig. 5a–c). The flow chart for single-cell RNA sequencing data analysis was included in Supplementary Fig. 5d. To visualize cellular heterogeneity, 9833 single-cell transcriptome data were subjected to unsupervised Seurat clustering and projected onto two dimensions by t-distributed stochastic neighbor embedding (t-SNE). After removal of low-quality and contaminated non-epithelial cells (Supplementary Figs. 5e, f, 6), this led to a total of 9 prostate LSC cell clusters (Cluster 1–9 or termed as C1–C9) (Fig. 6a). Importantly, we further found that while all other LSC cell clusters displayed high levels of epithelial genes, the cell cluster C7, expressed both epithelial markers *Epcam*, *Cldn1*, *Ocln* and stromal cell markers including *Vim*, *Fn1*, and *Acta2*, stood out uniquely among all the cell clusters (Fig. 6a, b). Detailed analysis of highly variable genes in C7 showed that *Zeb1* was exclusively expressed in C7 (Fig. 6a, b). Other EMT transcriptional factors including *Zeb2, Snai1, Prrx1* and *Prrx2* were also highly transcribed in C7 (Fig. 6a–c). We performed unbiased GSEA using differential expressing genes in Cluster 7 versus all the other cell clusters. As shown in Supplementary Fig. 7a, b, GSEA analysis indicated that epithelial markers were downregulated and mesenchymal markers were positively enriched in the Cluster 7. Moreover, previously reported stem cell-related signatures were significantly upregulated in the cell cluster 7 (Supplementary Fig. 7c). Collectively, the unbiased examination of single-cell gene expression profiles of prostate LSC cells further substantiates a unique Zeb1 expressing LSC cell subpopulation with the mesenchymal and stem cell gene signature. Interestingly, when we used statistical and computational frameworks to delineate the differentiation states of the 9 cell clusters, both monocle and diffusion map scripts indicated that C7 was at one of the tips in the lineage tree (Supplementary Fig. 7d–g). We therefore chose

cluster 7 to serve as the beginning point to delineate the developmental trajectory using the Slingshot script. Three lineage trajectories originated from C7 toward committed basal cells or "intermediate cells" that expressed both basal and luminal cell signatures were revealed (described in more detail in Supplementary Fig. 8). Collectively, the unbiased examination of single-cell gene expression profiles of prostate basal cells identified a unique Zeb1 expressing LSC cell subpopulation with the mesenchymal gene signature and indicated that it positioned at the beginning of differentiation trajectories.

Wnt signaling pathway is enriched in LSCZeb1⁺ cells. As an effort to uncover signaling pathways that are enriched in the cell cluster C7, we applied gene ontology (GO) enrichment analysis to the single-cell RNA-seq data. Coincidently, we found that the Wnt signaling pathway, which we have reported recently to promote self-renewal of prostate cancer stem cells[46], was one of the top hits preferentially presented in the cluster C7 (Fig. 6d). GSEA analysis further substantiated that Wnt signaling pathway was positively enriched in the cluster7 (Supplementary Fig. 9a, b). The Wnt signaling pathway was demonstrated previously to play an essential positive role in inducing EMT[47]. Our RT-PCR experiments confirmed a significant upregulation of key Wnt signaling components in LSCZeb1⁺ cells (Supplementary Fig. 9c). Moreover, we tested the impact of Wnt signaling pathway agonist Wnt3a on the organoid-forming capacity of LSCZeb1⁺ cells and observed that Wnt3a (100 ng/ml)) significantly increased the size and number of prostate organoids (Supplementary Fig. 9d–f). These data further suggested a positive role of Wnt signaling in Zeb1⁺ organoids outgrowth. It was recently reported that Procr, a Wnt pathway target, marks a novel multipotent mammary stem cell population[44]. Consistently, we also found a moderate (1.5 folds) but significant upregulation of Procr in LSCZeb1⁺ prostate cells (Supplementary Fig. 9c). Further flow cytometry examination revealed that Zeb1⁺ cells represented only a small subpopulation (8%) of Procr expressing prostate epithelial cells (Supplementary Fig. 9g), suggesting a more enrichment of PBSCs in Zeb1⁺ than Procr⁺ prostate epithelial cells. Furthermore, examination of prostate sections from APC^min mice with the aberrantly activated Wnt pathway revealed a significant increase of Zeb1⁺ basal cells (Fig. 6e, f). We conducted additional immunostaining analyses and found that Zeb1⁺ basal cells entered a more actively proliferating state in APC^min mice, further suggesting that the Wnt signaling promotes the expansion of Zeb1⁺ basal cells (Fig. 6g, h).

Zeb1⁺ cells are found in mouse and human prostate epithelia. We next asked the influence of oncogenic transformation on the

Zeb1 expression profile in the prostate epithelium. For that purpose, we crossed the Zeb1/tdTomato mice with the *Hi-Myc* prostate cancer mouse model. Immunofluorescent co-staining of Zeb1 or tdTomato together with basal or luminal cell markers revealed a 2-fold increase of Zeb1+ basal cells in *Zeb1/tdTomato; HiMyc* prostates (1.25% versus 0.73% in control mice at 3-month old, or 1.71% versus 0.71% in control mice at 5-month old) (Fig. 7a, b). Of note, we detected a small number of Zeb1+CK8+ luminal cells in *HiMyc* mice, which stood in contrast to the exclusive expression of Zeb1 in the basal cell compartment from wild type prostates (Fig. 7a, b). The pathology of Hi-Myc mice was evaluated by a pathologist and characterized in

Fig. 3 LSCZeb1$^+$ cells can self-renew and generate functional prostate in vivo. a Immunostaining of E-cadherin and Vimentin to validate the purity of UGM cells separated from UGS. (Scale bars = 20 μm.). **b** Prostate tissue generated from 1000 LSCZeb1$^+$ prostate cells at 2 months after renal capsule implantation. (implanted cell number: 1000, scale bars = 2 mm.). **c** Quantification of prostate tissue generation incidence from 1000-cell transplants. **d** H&E staining of 1000-cell implants shows well differentiated prostate epithelial tubules. (Scale bars = 20 μm.). **e** Immunostaining of p63, CK5, CK8, Syp, Pbsn, AR, mouse-specific β1-integrin and C57BL/6 donor-specific H-2k^b on sections of 1000-cell implants. (All scale bars = 20 μm.). **f, g** Immunostaining of frozen sections from LSCZeb1$^-$ cell transplants displaying unstructured tissue containing basal cells, rare luminal cells and mouse-β1-integrin$^-$ vimentin$^+$ rat mesenchymal cells. (Scale bars = 50 μm). **h** Phase contrast and fluorescent images of the single viable Lineage$^-$ Sca-1$^+$ CD49f^{hi} Zeb1$^+$ or Lineage$^-$ Sca-1$^+$ CD49f^{hi} Zeb1$^-$ prostate cell used for single cell renal capsule transplantation experiments in vivo. (Scale bars = 20 μm.). **i,** Prostate tissue generated from single Lineage$^-$ Sca-1$^+$ CD49f^{hi} Zeb1$^+$ prostate cell at 2 months after renal capsule implantation. (scale bars = 2 mm). **j** Quantification of prostate tissue generation incidence from single cell transplants. **k** H&E staining of single cell implants showing differentiated prostate epithelial tubules. (Scale bars = 20 μm.). **l,** Staining of RFP, p63, CK5, CK8, Synaptosphysin (Syp), Pbsn, Androgen receptor (AR), β1-integrin and H-2k^b on sections of single cell implants. (The white arrow points to a Syp$^+$ cell. Scale bars = 20 μm.). **m, n** Zeb1$^+$ EpCAM$^+$ cells harvested from first generation of renal capsule implantation were able to form prostates in secondary and tertiary transplantations. (50 prostate cells per graft). **o, p** Immunostaining and quantification of frozen sections from serial transplants showing a stable percentage of Zeb1$^+$ basal cells during the serial implantation. (n = 3 grafts. Source data are provided as a Source Data file.).
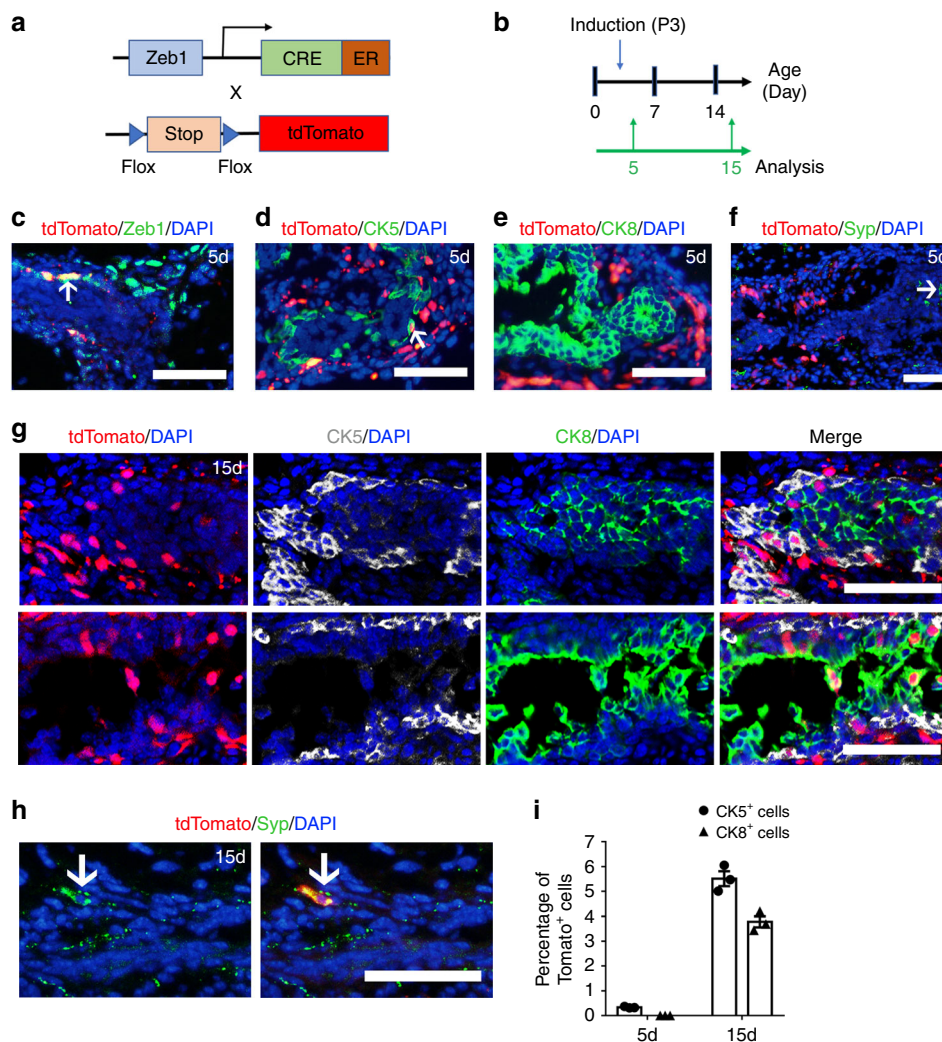

Fig. 4 Lineage tracing supports the multipotency of Zeb1$^+$ basal cells. a Strategy to target tdTomato expression in Zeb1 expressing basal cells in vivo. **b** Illustration of protocols to trace the fate of Zeb1$^+$ basal cells during prostate postnatal development. **c** Zeb1 and tdTomato co-staining of prostate sections from *Zeb1-CreERT2/tdTomato* mice shows that tdTomato marks Zeb1 expressing cells. **d–f** tdTomato expression is found in CK5$^+$ basal cells, but not in CK8$^+$ luminal cells and Syp$^+$. neuroendocrine cells from *Zeb1-CreERT2/tdTomato* mouse prostates at 2 days after tamoxifen administration. **g** Clusters of CK5$^+$ CK8$^-$ tdTomato$^+$ cells can be detected in the outer basal cell layer and CK5$^-$CK8$^+$ tdTomato$^+$ cells can be found in the inner luminal compartment at 12 days after tamoxifen induction. **h** Syp$^+$ neuroendocrine cells are marked with tdTomato at 12 days post induction. **i** Percentage of tdTomato$^+$ cells in CK5$^+$ and CK8$^+$ cells at 2 days and 12 days after tamoxifen administration to postnatal day 3 mice. (n = 3 mice. Source data are provided as a Source Data file.). (In this figure, at least 20 sections each mouse prepared from 3 mouse prostates were analyzed. At least 60 fields for each immunostaining experiment were collected for analysis. Data are analyzed by Student's t-test and are presented as mean + s.e.m. All scale bars = 50 μm.).
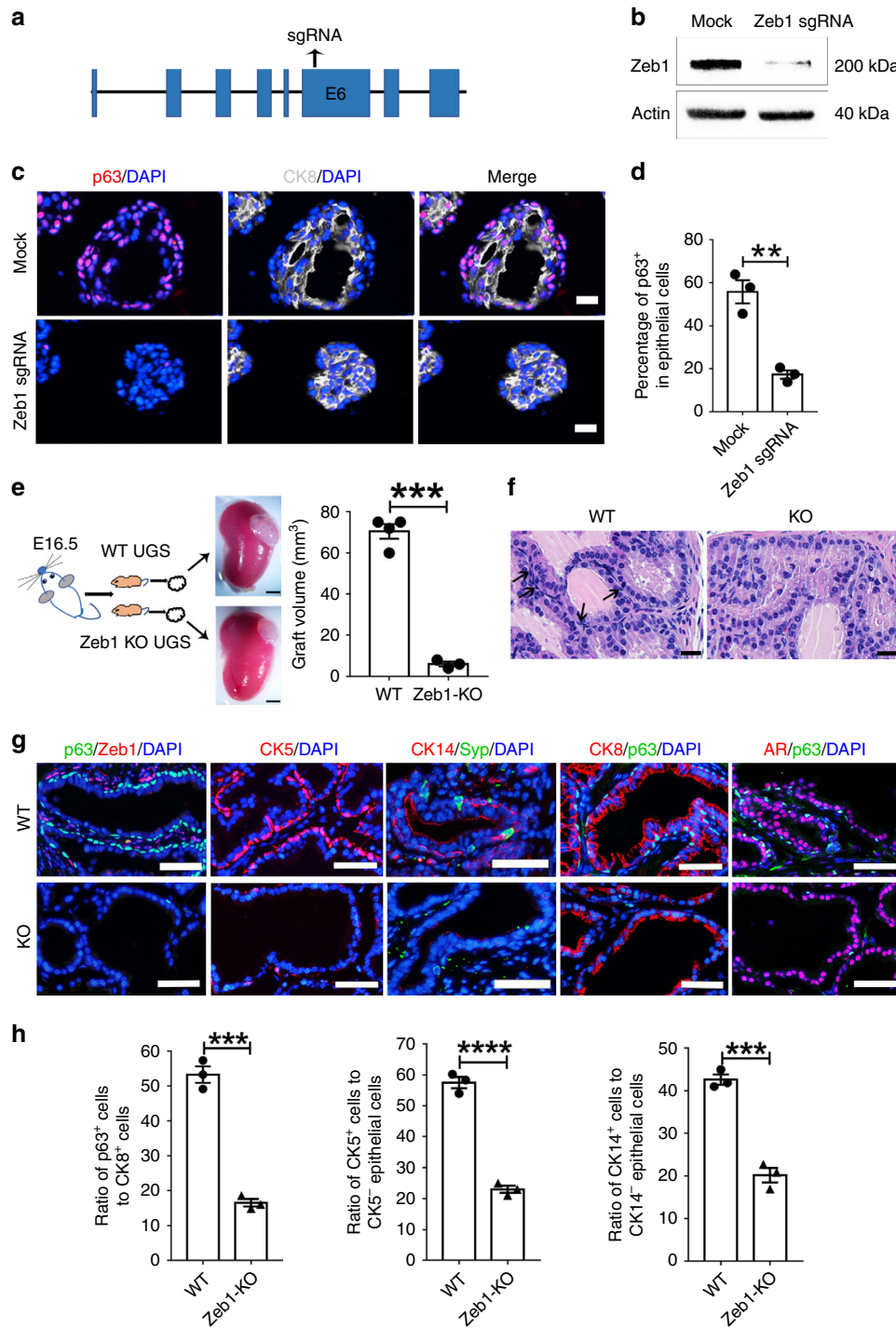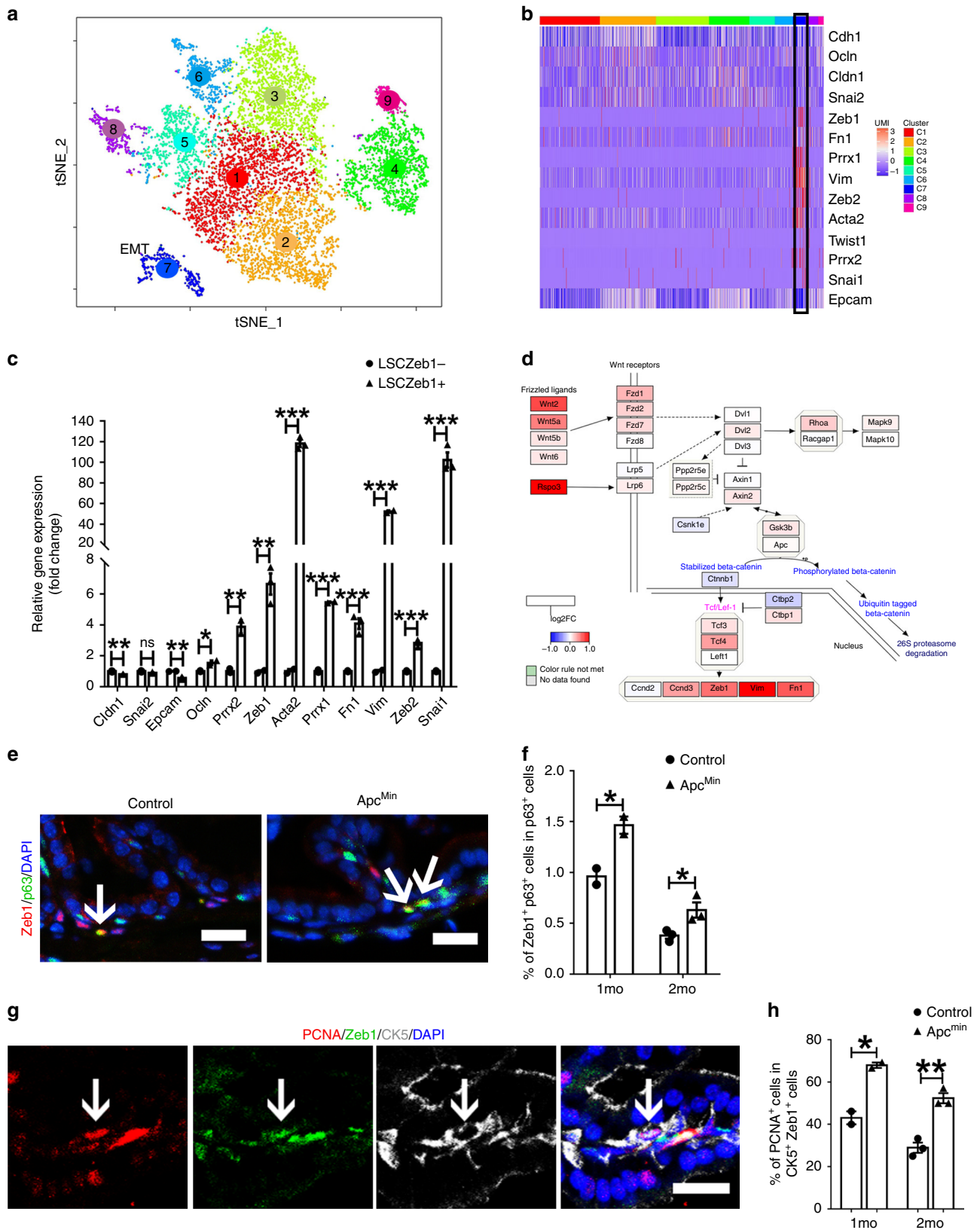
Fig. 5 Knockout of Zeb1 severely suppresses the basal cell development. a Illustration of sgRNA sequence for CRISPR/Cas9 system mediated Zeb1 knockout. **b** Immunoblotting confirms efficient Zeb1 ablation sgRNA transfected primary prostate epithelial cells. **c, d** Immunostaining of frozen sections from Zeb1 knockout organoids shows a severely impaired differentiation of basal cells. (Scale bars = 50 μm. $n = 3$ independent experiments. Data are analyzed by Student's t-test and are presented as mean + s.e.m. **$p < 0.01$.). **e** UGS from E16 wild-type or Zeb1$^{-/-}$ embryos are dissected and transplanted beneath the renal capsule of male athymic mice. Transplants are harvested and quantified the volume at 20 days later. ($n = 3$ mice. Data are analyzed by Student's t-test and are presented as mean + s.e.m. ***$p < 0.001$. Scale bars = 2 mm). **f** H&E staining of UGS implants showing differentiated prostate epithelial tubules but a remarkably decrease in basal cell number in Zeb1$^{-/-}$ UGS grafts. (Scale bars = 20 μm). **g** Staining of Zeb1, p63, CK5, CK14, Syp, CK8 and AR on sections of wild-type and Zeb1 knockout UGS implants. (Scale bars = 50 μm). **h** Quantification of basal cell development in Zeb1$^{-/-}$ UGS grafts. ($n = 3$ mice. At least 60 fields from 20 sections for each immunostaining experiment were analyzed. Data are analyzed by Student's t-test and are presented as mean + s.e.m. ***$p < 0.001$, ****$p < 0.0001$. Source data are provided as a Source Data file.).

a

b

c

d

e

f

g

h

Supplementary Fig. 10. In addition, we examined the Zeb1[+] epithelial cells in two other GEMM models of prostate cancer, the Probasin-Cre; Pten[fl/fl] (PKO) model and the TRAMP model. Consistently, we detected an increase in Zeb1[+] basal cells and emergence of Zeb1[+] p63[−] or Zeb1[+] CK5[−] cells in prostates of both tumor models (Supplementary Fig. 11). Similarly, positive

immunostaining of Zeb1 protein could be found in a small percent of CK14[+] basal cells in human prostate tissues (non-benign prostate hyperplasia (non-BPH) ($n = 3$) and BPH specimens ($n = 7$)) (Fig. 7c, d). During the examination of many sections of human prostate cancer samples, we infrequently found cancer tissues (Gleason score: $3 + 3$ or $3 + 4$, $n = 4$) with remaining

Fig. 6 Single-cell RNA-seq data reveal the gene transcriptional signatures of LSCZeb1+ cells. a A Seurat package and the first 12 principal components were applied to generate 9 different and stable clusters based on differential expressing genes among 9278 mouse prostate basal cell-enriched cells. C7 is marked with EMT due to its unique EMT expression profile. (Source data are provided as a Source Data file.). **b** A heatmap showing the scaled expression profile for EMT related genes. (The black-lined rectangle highlights the cluster 7). **c** qRT-PCR quantification of the mRNA expression of epithelial and stromal markers as well as EMT inducing transcriptional factors in Lineage− Sca-1+ CD49f^hi Zeb1+ and Lineage− Sca-1+ CD49f^hi Zeb1− prostate cells. ($n = 3$ independent experiments. Data are analyzed by Student's t-test and are presented as mean + s.e.m. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.). **d** Color-coded expression levels of Wnt signaling pathway genes in the Zeb1 expressing Cluster 7 were presented as were obtained according to log2 fold change (log2FC) between cluster 7 and the rest prostate basal cell clusters. The pathway drawing was modulated from the schematic for the mouse Wnt pathway at the pathviso website. **e**, **f** Co-immunostaining of Zeb1 and p63 on prostate sections reveals a significant increase of Zeb1+ basal cell number in APC^min mice compared to control animals. ($n = 3$ mice. Data are analyzed by Student's t-test and are presented as mean + s.e.m.. *$p < 0.05$. Scale bars = 20 μm. Source data are provided as a Source Data file.). **g** Images of PCNA, Zeb1 and CK5 triple staining on prostate sections from APC^min mice. (Scale bars = 20 μm.) **h** Quantification of the percentage of proliferating Zeb1+ prostate basal cells in Zeb1+ basal cells in both proximal and distal regions at indicated time points. ($n = 3$ mice. At least 20 fields per section of 3 sections each mouse prepared from 3 mouse prostates were analyzed. Data are analyzed by Student's t-test and are presented as mean + s.e.m. All scale bars = 20 μm. *$p < 0.05$, **$p < 0.01$.).

CK14+ basal cells, in which Zeb1 positive staining could be found (Fig. 7c, d). Intriguingly, Zeb1+ CK8+ luminal cells were barely detected in non-BPH but could be observed occasionally in BPH specimens and more frequently in prostate cancer specimens (Fig. 7e, f). Together, Zeb1 expressing epithelial cells can be detected in both basal and luminal layers of prostates from *Hi-Myc* mice and in human prostate samples.

Discussion

The prostate basal cell compartment is suggested to contain stem/progenitor cells[4,5,13,29], but the stem cell subpopulation within prostate basal cells and its transcriptional profile remain largely unknown. We here discover the existence of Zeb1-positive prostate epithelial cells in a small population of basal compartment but not in the luminal layer. Zeb1+ basal cells are more frequently found in the stem cell enriched urethra-proximal region. Importantly, utilizing functional methods and in vivo lineage tracing, we prove that the Zeb1+ prostate basal cells can self-renew and possess multipotency to generate all three prostatic epithelial cell lineages in vivo. Furthermore, in vitro and in vivo genetic ablation experiments showed a requirement of Zeb1 in basal cell differentiation. The unique Zeb1 expressing prostate basal cell cluster with mesenchymal cell features was further supported by single-cell RNA sequencing data. Our study underscores the importance of Zeb1 as a marker and an essential regulator for PBSCs.

EMT has been demonstrated to be utilized by tumor initiating cells to acquire stem cell attributes[31]. We find here that Zeb1 marks a population of PBSCs in prostates with mesenchymal features and higher expression levels of EMT inducers such as Snai1, Zeb2, Prrx1 and Prrx2. Support for our data comes from recent studies in mammary glands that multipotent Procr+ mammary stem cells display EMT characteristics and that EMT is required for the tissue-reconstitution activity of mammary basal cells[44,48–50]. Together, those observations demonstrate that like what was shown for tumor initiating cells, EMT is also closely associated with stemness in normal prostate and mammary stem cells. However, there is also tissue specific differences between stem cells in prostate versus mammary tissues. Firstly, we find that although most Zeb1+ prostate basal cells express Procr, Procr is present in a much larger population of prostate epithelial cells containing both basal and luminal cells (Supplementary Fig. 9g), suggesting that Zeb1 appears to be a better maker for basal stem cells in prostates than Procr. Secondly, Zeb1 and Snai1 are expressed in prostate basal cells but not in mammary epithelial cells[48–50], indicating distinct EMT inducing transcriptional factors or EMT programs are involved in respective stem cells.

Lineage tracing experiments conducted in early postnatal development in the current study demonstrate that Zeb1+ basal cells can self-renew and give rise to basal, luminal and neuroendocrine cell lineages. However, unlike the multilineage differentiation of Zeb1+ basal cells in early development, we found that Zeb1+ basal cells gave rise to mostly basal cells but not luminal or neuroendocrine cells in adult prostates regeneration. Although the absence of tdTomato tracing in regenerated luminal cells may be caused by a low Cre-ERT2 induction efficiency, these data suggested a stage dependent function of Zeb1+ basal cells, and were largely consistent with previous findings reported by professor Cédric Blanpain's group using pan-basal cell K5CRE-ER/RosaYFP and K14rtTA/TetO-CRE /RosaYFP lineage tracing mice. Loss of function experiments performed on prostate organoids and UGS transplants show that Zeb1 knockout results in a remarkably decrease in the basal cell compartment, but luminal and neuroendocrine cell differentiation are largely unaffected. Consistently, previous studies on renal capsule transplants from p63−/− UGS exhibit development of luminal and neuroendocrine cells independent of basal cells[51]. In addition, lineage tracing experiments on CK14, CK8, CK5 and NKX3.1 cells support the existence of self-sustained luminal cell progenitors[11–13,15]. Therefore, both multipotent Zeb1+ basal stem cells and stem/progenitor cells from the luminal compartment contribute to prostate development. In addition, compared to the sharp contrast between Zeb1+ and Zeb1− basal cells, the difference between Zeb1 wild-type and Zeb1 knockout cells in the renal capsule transplantation assay is less drastic, although the size of prostate tissue produced from Zeb1 knockout cells is significantly smaller and the basal cell differentiation is severely compromised. Our interpretation is that the self-renewal and differentiation of Zeb1+ prostate basal stem cells are regulated by complicate and intricate molecular mechanisms. Zeb1 itself is an important player in it but not the only one. Other signaling molecules, transcriptional factors, or epigenetic modifiers may also be required.

At different developmental stages, castration-mediated regression and regeneration stages, we consistently observed that Zeb1+ basal cells are more enriched in the urethral-proximal region. In addition, we found that there was a gradient of cell proliferation activity in Zeb1+ basal cells from proximal to intermediate to distal region. This is somehow similar to a "directional flow" phenotype of human prostate basal stem cells as reported recently[24]. However, without real-time 3D imaging or 3D lineage tracing, it is not conclusive at this moment whether Zeb1+ basal stem cells function in a "direction flow" way in mouse prostates as reported in human prostates. In addition, we want to point out that Zeb1− basal cells can also form organoids, although with a much lower efficiency. Therefore, Zeb1+ basal cells were enriched for basal stem cells, but Zeb1− basal cells also have a degree of progenitor activity.
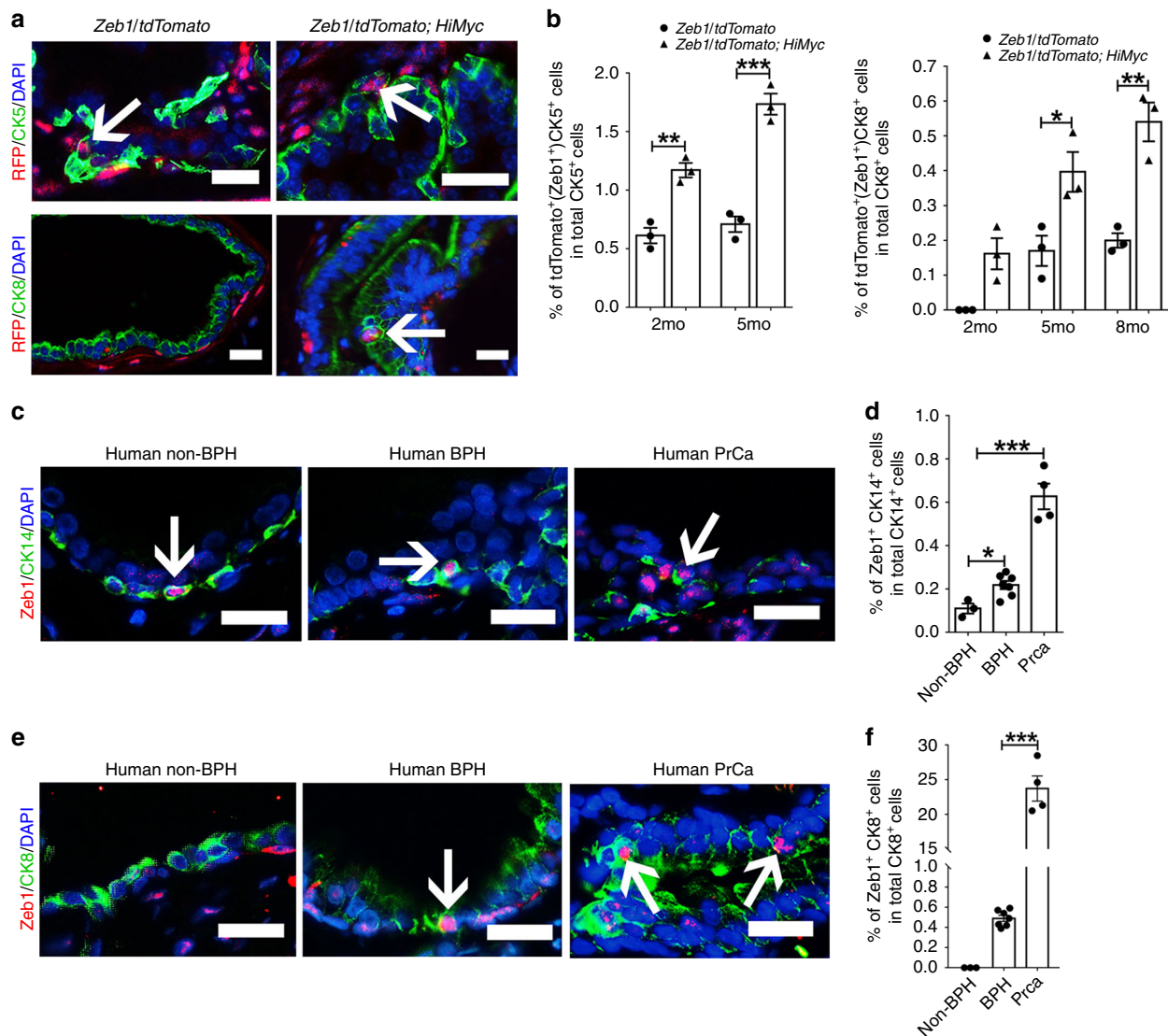
Fig. 7 Zeb1⁺ epithelial cells are found in both mouse and human prostates. a Images of tdTomato and CK5 or CK8 staining on prostate sections from *Zeb1/tdTomato* or *Zeb1/tdTomato; HiMyc* mice. (Scale bars = 20 μm). **b** Quantifications of tdTomato⁺ basal or tdTomato⁺ luminal cells in prostates from *Zeb1/tdTomato* or *Zeb1/tdTomato; HiMyc* mice. (n = 3 mice. In panel **a–d** at least 20 fields per section of 3 sections each mouse prepared from 3 mouse prostates were analyzed. Data are analyzed by Student's t-test and are presented as mean + s.e.m. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Source data are provided as a Source Data file.) **c**, **d** Images and quantifications of Zeb1⁺ basal epithelial cells in human non-BPH, BPH or prostate cancer specimens. (non-BPH (n = 3), BPH (n = 7). For cancer samples, Gleason score = 3 + 3 or 3 + 4, n = 4. At least 15 fields per section of 3 sections each human sample were analyzed. Data are analyzed by Student's t-test and are presented as mean + s.e.m. *$p < 0.05$, ***$p < 0.001$. Scale bars = 25 μm. Source data are provided as a Source Data file.). **e**, **f** Images and quantifications of Zeb1 and CK8 co-immunostaining on sections from human non-BPH, BPH or prostate cancer specimens. (non-BPH (n = 3), BPH (n = 7). For cancer samples, Gleason score = 3 + 3 or 3 + 4, n = 4. At least 15 fields per section of 3 sections each human sample were analyzed. Data are analyzed by Student's t-test and are presented as mean + s.e.m. ***$p < 0.001$. Scale bars = 25 μm. Source data are provided as a Source Data file.).

Intriguingly, we find that different from the exclusive basal expression of Zeb1 in wild-type prostate, Zeb1⁺ epithelial cells can be detected in the luminal compartment in Hi-Myc mice or in human prostate hyperplastic or tumor samples. There are two possible explanations for the appearance of Zeb1⁺ luminal cells: It could be derived from transformed Zeb1⁺ basal cells; or alternatively, the transformed luminal cells may acquire Zeb1 expression. Clarification of these two possibilities needs further investigations. Nevertheless, the appearance of Zeb1⁺ luminal cells at early stages of prostate tumorigenesis is of special interest, because most clinical prostate cancers display a dominant luminal phenotype by loss of basal cell phenotypes, and Zeb1 has been well shown in the acquisition of cancer stem cell properties[36,52].

This Zeb1⁺ luminal cell population might represent prostate tumor initiating cells, which can be also related to two different cellular origins for prostate cancers: basal versus luminal cells.

Single cell RNA analysis or RNA-seq has emerged as a very powerful tool to analyze heterogenous cell population of a tissue or organ. Previous studies have reported the composition of prostate epithelia in adult prostate from mouse and human prostate via single-cell q-PCR or RNA-seq. Utilizing a Fluidigm multiplex qPCR-based single-cell expression analysis platform, Joao D. Barros-Silva and et al examined the cell component of prostates from hormone naïve and castrated mice. They found that LY6D⁺ luminal cells are castration resistant prostate luminal progenitors[53]. Through single-cell RNA sequencing and flow

cytometry, Gervaise H. Henry and et al specified markers for human prostate epithelial and stromal cell compartments[54]. To our knowledge, this work is the first report to decipher the heterogeneity within the prostate basal epithelium using unbiased single-cell transcriptomics. Importantly, along with the single-cell RNA-seq data, our immunocytochemical staining, functional experiments, and cell lineage tracing analysis together support that Zeb1 marks a multipotent prostate basal stem cells and Zeb1 itself is required for prostate basal cell development. Via analyzing different developmental stages, we have provided insights into the evolution of a unique Zeb1 expressing prostate stem cell subpopulation from embryonic development to postnatal development and to adulthood. The identification of the specific Zeb1[+] prostate stem cell, its transcriptome profile and its expression pattern during prostate tumorigenesis could have important implications for our understanding of tissue development and regeneration and for identification of potential cellular origin for cancer.

Methods

Animals. Pregnant SD rats (E17) and athymic nu/nu male mice (7 weeks old) were purchased from Shanghai Slac laboratory animal company. *Zeb1/tdTomato* reporter mice were generated by the CRISPR-Cas9 method at Model Animal Research Center of Nanjing University. The sgRNA sequence we used is GAGGTTGGGAGCTGCACAGCAGG. Exon 8 of Zeb1 followed by a P2A and tdTomato coding sequence, which was flanked by approximately 1.5 kb upstream and 1.5 kb downstream sequence of Zeb1 Exon 8, was cloned to a donor vector. Plasmid containing sgRNA and Cas9 expressing elements and the donor vector were injected to fertilized C57/B6 mouse eggs. F0 generation of mice were genotyped by sequencing and PCR. Positive F0 mice were used to backcross with C57/B6 mice to produce F1 generation of knockin mice. PCR Primers for *Zeb1/tdTomato* mouse genotyping were provided in Supplementary Table 1. *Zeb1-CreERT2* mice were generated by a homologous recombination method at the Cyagen Biosciences Inc. The TAG stop codon of the last exon of Zeb1 was replaced with the "2A-CreERT2" cassette. Homology arms of the targeting vector were generated by PCR using BAC clone RP23-51G9 or RP23-207F18 from the C57BL/6 J library as template. The targeting vector contained the Neo cassette flanked by SDA (self-deletion anchor) sites and DTA for negative selection. C57BL/6 ES cells was used for gene targeting. F0 generation of mice were validated by sequencing and PCR. PCR Primers for genotyping of *Zeb1-CreERT2* mice were provided in Supplementary Table 1. *Zeb1* knockout with deletion of exon 1 mice was purchased from RIKEN BioResource Research Center. The *Rosa-CAG-LSL-tdTomato* mice were purchase from the Jackson laboratory. *Hi-Myc* mice were introduced from the National Cancer Institute (NCI:01XF5). The pathology of Hi-Myc mice was evaluated by a pathologist at Ren Ji hospital, affiliated with School of Medicine, Shanghai Jiao Tong University. *Apc^min* mice were introduced from the Nanjing Biomedical Research Institute of Nanjing University (T001457). All animal experiments were conducted according to the ethical regulations of Ren Ji Hospital. Animal experiment protocol were approved by the Ren Ji Hospital Laboratory Animal Use and Care Committee.

Prostate single cell preparation. Mouse prostates were harvested, minced, then digested with pre-warmed 1X collagenase/hyaluronidase solution (STEMCELL Technologies, #07912) for 3 hr at 37 °C, washed with PBS buffer once and placed in 0.25% Trypsin/EDTA for 6 min at 37 °C. FBS supplemented with 4% FBS was added to quench trypsin reaction followed by centrifuging at 350 *g* for 5 min. Then, the cell pellet was resuspended with pre-warmed Dispase/DNase I solution (STEMCELL Technologies, #07913, #07900) to thoroughly dissociate cells. Single cell suspension was obtained from passing through a 40 μm cell strainer.

Flow cytometry. Mouse prostate single cell suspension was blocked with the Fc blocker (CD16/32 antibody from eBioscience) for 40 min at room temperature. Staining antibodies were diluted in 4% FBS buffer containing Y-27632 ROCK inhibitor (STEMCELL Technologies, #72302) and applied to the prostate single cell suspension for 40 min at room temperature. Flow cytometry analysis or sorting was conducted using BD Accuri C6 or FACSAria III flow cytometer. Antibodies used in the study are listed in Supplementary Table 2. Propidium Iodide (PI, Invitrogen) was added to sample before sorting to gate viable cells.

Single cell RNA-seq and data analysis. Lineage[−] Sca-1[+] CD49f[hi] PI[−] prostate cells from postnatal day 14 mice were sorted and counted manually before processing to the single cell RNA-seq library preparation. Libraries were constructed following the instruction of the Chromium single cell 3' solution (10x Genomics). The libraries were sequenced on Illumina Hiseq X Ten platform. 10x Genomics workflow and Cell Ranger Single Cell Software was used to process raw sequencing

data and to align reads to the mm10 mouse reference genome. Raw gene expression matrix produced by 10x Genomics workflow were first processed by the Seurat package. Cells with less than 5000 unique molecular identifiers (UMIs) or less than 1400 genes detected or more than 5% UMI mapped to mitochondria genes were removed (Supplementary Fig. 5e, f). This led to 9833 high-quality single-cell RNA-seq data with an average gene detected to be around 2703 (Supplementary Fig. 5f). 1606 variable genes were selected based on their expression and dispersion (expression cutoff = 0.0125 and dispersion cutoff = 0.5). The first 12 principal components were used for clustering analysis (resolution = 0.5) and t-SNE projection. This identified a total of 10 clusters (Supplementary Fig. 6a). Violin plots of epithelial, stromal, endothelial and immune cell associated genes for those 10 different cell clusters were presented in Supplementary Fig. 6b–m. Clusters which expressed clear markers of non-epithelial cells were removed in the following analyses. Cluster 5 was labeled as endothelial cells based on the expression of Eng, S1pr1 and Emcn (Supplementary Fig. 6i–k). Cluster 10 was labeled as immune cells based on the expression of Cd74 and Cd72 (Supplementary Fig. 6l, m). After the removal of above-mentioned non-epithelial cells, we re-ran the Seurat work flow to generate new cell clusters and t-SNE projections (the first 12 principal components were used). A total of 9 clusters were identified (Fig. 6a). The expression levels of prostate basal cell markers, Krt5 and Krt14 were shown in violin plots (Supplementary Fig. 6n, o). We used the FindAllMarkers script in the Seurat package to identify genes that are enriched in a specific cluster (the specific cluster vs. the rest of clusters) with default settings. PCA matrix with the first 12 principal components and clusters from the second Seurat running were taken as input. The cluster representing putative stem cells (Zeb1 expressing cluster 7) was chosen as the root node. Totally, three branches were identified. GSEA analysis was performed by using fgsea package in R[55] with ranked genes fold change values created by FindMarkers module (default parameter except min.pct = 0.2) in Seurat package[56]. Gene Sets used here were collected from previous publications[57–61]. The curated gene sets (C2) of the Molecular Signature Database (MSigDB) version 6.2 provided by the Broad Institute (http://www.broad.mit.edu/gsea/)[62]. Similar clusters and the structure of a common origin could also be inferred by using Monocle or Diffusionmap with standard parameters (Supplementary Fig. 7). Codes used are deposited to the website [https://github.com/HelenHeZhu/StemCell].

Human prostate clinical specimens. Freshly dissected human prostate specimens (non-BPH, BPH and prostate cancer specimens) were obtained from the department of Urology at Ren Ji Hospital with informed consent from patients. All Human sample experiments were conducted according to the ethical regulations of Ren Ji Hospital. Human sample collection and handling protocols were approved by the Ren Ji Ethics committee.

UGM stromal cell preparation. The UGM isolation procedures has been described previously[63]. Briefly, E17 embryos from pregnant SD rats were sacrificed and urogenital sinuses were collected. The UGM was separated from the urogenital sinus epithelium and then digested with 1X collagenase/hyaluronidase solution on a shaker for 90 min at 37 °C. The UGM was washed with PBS once and then placed in 0.25% Trypsin/EDTA for 6 min at 37 °C. UGM single cell suspension was seeded to cell culture dishes and cultured in DMEM (Thermo Fisher Scientific, C11965500BT) supplemented with 10% fetal bovine serum (FBS), 2 mM glutamine, 100U/ml penicillin and 100 mg/ml streptomycin in vitro for at least 1 week.

Serial isolation and renal capsule transplantation. We used a previously described procedure to isolate the primary prostate epithelial cells and to perform renal capsule transplantation[63]. For the multi-cellular renal capsule assay, 1,000 Lineage[−] Sca-1[+] CD49f[hi] tdTomato[+] or Lineage[−] Sca-1[+] CD49f[hi] tdTomato[−] prostate basal cells were sorted via FACS and then mixed with purified rat UGM cells (250,000 cells per graft) in rat tail Collage Type I (4.42 mg/ml, Corning, 354236). Next, the mixture was plated as a drop of 10 μl into the center of a well of a 6-well plate. The plate was placed in a 37 °C cell culture incubator for at least 60 min to make sure that the Collage type I solidified. Then pre-warmed DMEM (containing 10% FBS) medium was gently pipetted into each well. Each collagen graft was embedded underneath the renal capsule of a nude mouse the next day. For the first renal capsule transplantation, single Lineage[−] Sca-1[+] CD49f[hi] tdTomato[+] or Lineage[−] Sca-1[+] CD49f[hi] tdTomato[−] prostate basal cell was sorted into each well of a 96-well plate via FACS and then examined under a light microscope. The viable sorted cell was mixed with purified rat UGM cells (250,000 cells per graft) in rat tail Collage Type I (4.42 mg/ml, Corning, 354236). For the secondary or tertiary renal capsule transplantation, grafts from last transplantation were digested with 1X collagenase/hyaluronidase solution for 90 min and 0.25% Trypsin/EDTA for 6 min at 37 °C. 50 Lineage[−] Epcam[+] Zeb1[+] or Lineage[−] Epcam[+] Zeb1[−] cells were sorted via FACS and then mixed with UGM stromal cells (250,000 cells per graft). All serial transplantation grafts were harvested at 8 weeks after implantation. UGS from Zeb1[−/−] and Zeb1[+/+] embryos were implanted under the renal capsule of male nude mice. The transplants were collected at 20 days post implantation and sectioned for further analysis.

Cell lineage tracing. To trace the fate of Zeb1[+] basal cells, *Zeb1-CreERT2/tdTomato* mice were administrated with 1.25 mg tamoxifen (40 mg/mL solution,

dissolved into corn oil, Sigma) via intraperitoneal injection at postnatal day 3. Two days and 12 days after the induction, prostates were collected and sectioned for immunofluorescence staining of different prostate epithelial cell markers. To examine the functional role of Zeb1+ basal cells during prostate regeneration in adults, adult Zeb1CreERT2/tdTomato mice were given tamoxinfen to induce expression of tdTomato in Zeb1+ cells. Four days after the induction, some of the mice were sacrificed and the prostates were collected and sectioned to examine the induction rate of the Cre recombinase. The rest of mice were castrated at 1 week after the injection of tamoxifen and reconstituted with testosterone pellets at 3 weeks post castration. Two weeks after the regeneration, prostates were collected and sectioned for immunofluorescent staining of different prostate epithelial cell markers to determine the differentiating fate of Zeb1+ basal cells.

Organoid culture. We utilized a previously reported protocol to culture organoids in vitro and perform organoid frozen sections for immunostaining analysis[15]. FACS purified LSCZeb1+ and LSCZeb1− cells were centrifuged at 350 g for 5 min at 4 °C and resuspended in complete organoid culture media (10 ng/ml epidermal growth factor, 10 μM Y-27632, 1x glutamax, 5% Matrigel, 5% charcoal-stripped FBS, 100 nM DHT, heat-inactivated at 55 °C for 1 hr before use) with the density of 5000 to 10,000 prostate epithelial cells per 100 μL media. The medium was refreshed every 3 days. After approximately 2 weeks, organoids were washed with cold PBS buffer repeatedly to remove the matrigel, collected via centrifugation at 350 g for 5 min, then dissociated using 0.25% Trypsin/EDTA for 7 min at 37 °C. The single cells were centrifuged and resuspended in medium following the aforehand described procedure for a serial passaging. For section preparation, organoids were collected, fixed with 4% PFA for 30 min, washed with cold PBS buffer twice, then centrifugated and resuspended with 50 μL collagen type I and incubated at 37 °C for 1 hour to solidify the organoid/collagen slurry. The slurry was transferred into 4%PFA solution for fixation, 30% sucrose for dehydration and finally O.C.T for embedding and section. Five-micrometer frozen sections were used for following immunostaining.

CRISPR/Cas9 mediated Zeb1 knockout prostate organoids. The lentiCRISPRv2-mCherry plasmid containing cas9 and Cherry fragments was purchased from Addgene (#99154, kindly deposited by Dr.Agata Smogorzewska). Then the Cherry fragment was exchanged for GFP fragment via enzymatic digestion and connection followed by sgRNA cloning. The single guide RNA sequences are ACTGCTTATATGTGAGCTAT in the upstream and GGAA-CAACCTGAAGTTGACT in the downstream of the sixth exon of the mouse Zeb1 gene. The Zeb1-cas9 lentivirus was transfected into primary prostate cells using a spinoculation method at 3,000 rpm and room temperature for 2 hours. Cells were then transferred to a 15 mL tube and spinned down at 400 g for 5 min. The pellet was resuspended with the complete organoid culture media and seeded at a density of 10,000 cells per well of a 96-well plate. After 3 weeks of culture, the organoids were harvested and dissociated with pre-warmed 0.25% Trypsin-EDTA. Single cell suspension was stained with anti-lineage (CD45.2, CD31, Ter119), anti-Sca-1 and anti-CD49f antibody. Lin− Sca1+ CD49fhi GFP+ cells (sgRNA successfully transfected prostate epithelial cells) were FACS sorted for organoid culture. Organoids were collected when reaching 200μm in diameter, and then fixed, dehydrated and embedded with O.C.T for frozen section preparation.

RNA extraction and quantitative-PCR analysis. Total RNA was extracted from FACS purified LSCZeb1+ or LSCZeb1− cells using a RNeasy micro kit (Qiagen). cDNA was synthesized using the PrimeScript RT Reagent Kit (Takara, RR037A). qPCR was conducted using SYBR Premix Ex Taq (Takara, RR420A). Relative transcript abundance was quantified by the comparative CT method using Actin as an internal reference gene. Primers utilized in RT-PCR were shown in Supplementary Table 1.

Immunofluorescence staining. Human or mouse prostates were fixed in 4% paraformaldehyde for 20 min and dehydrated overnight in 30% sucrose solution. Tissues were embedded in Optional Cutting Temperature (O.C.T.) compound and placed in a −80 °C refrigerator for 10 min or longer depending on the tissue size. Frozen sections were cut at a thickness of 6 um. Sections were washed with PBS buffer and placed into 0.01 M sodium citrate (PH 6.0) buffer for heat-induced antigen retrieval. Slides were then subjected to a blocking step in 10% donkey serum solution for 1 h at room temperature. Primary antibodies, diluted in 1% donkey serum, were added to sections overnight at 4 °C. After thorough wash with PBS buffer, secondary antibodies were applied and incubated for 1 hr at room temperature. Sections were rinsed at least three times with PBS before being mounted by Vector Shield mounting medium containing DAPI. Antibodies (including the dilutions) used in the study is listed in Supplementary Table 2.

Immunohistochemistry. Paraffin-embedded prostate tissue sections were deparaffinized and rehydrated following conventional methods. Harris modified hematoxylin solution (Sigma) was applied to sections for 5 min followed by water washing. Then 1% Eosin solution (Sigma) was applied to sections for 3 min. After thorough wash, sections were mounted using neutral balsam.

Castration and androgen replacement. Zeb1 reporter mice were surgically castrated. Three weeks later, mice were given dihydrotestosterone (MCE, HY-A0120) dissolved in sterile corn oil via intraperitoneal injection twice a day (50 ug/d) for prostate regeneration. Prostate collected at indicated regeneration stages (regressed, regenerated, and recovered stages post testosterone administration) for analysis.

Statistical analysis. We used the ImageJ 1.46r software to quantify positive stained cells. All statistical analysis was evaluated using a two-tailed Student's T-test and a p-value < 0.05 was defined to be significant.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The single cell RNA-seq raw data used for this study are available at the GEO web with the accession number GSE111429. Codes used are deposited to the website [https://github.com/HelenHeZhu/StemCell]. All other data is available in the main text or the supplementary materials. The source data underlying Figs. 1b, h–j, 2c, e–i, 3p, 4i, 5b, d, h, 6, 7 and Supplementary Figs. 1c, 2, 3f, 8, 9d–f, 11 are provided as a Source Data file.

References

1. Shen, M. M. & Abate-Shen, C. Molecular genetics of prostate cancer: new prospects for old challenges. *Genes Dev.* **24**, 1967–2000 (2010).
2. Isaacs, J. T. *Control of cell proliferation and cell death in the normal and neoplastic prostate: A stem cell model.* Vol. II, 85–94. (National Institute of Health, Report No. 87–2881, 1987).
3. Xin, L., Lawson, D. A. & Witte, O. N. The Sca-1 cell surface marker enriches for a prostate-regenerating cell subpopulation that can initiate prostate tumorigenesis. *Proc. Natl Acad. Sci. USA* **102**, 6942–6947 (2005).
4. Leong, K. G., Wang, B. E., Johnson, L. & Gao, W. Q. Generation of a prostate from a single adult stem cell. *Nature* **456**, 804–808 (2008).
5. Lawson, D. A., Xin, L., Lukacs, R. U., Cheng, D. & Witte, O. N. Isolation and functional characterization of murine prostate stem cells. *Proc. Natl Acad. Sci. USA* **104**, 181–186 (2007).
6. Goldstein, A. S. et al. Trop2 identifies a subpopulation of murine and human prostate basal cells with stem cell characteristics. *Proc. Natl Acad. Sci. USA* **105**, 20882–20887 (2008).
7. Wang, J. et al. Symmetrical and asymmetrical division analysis provides evidence for a hierarchy of prostate epithelial cell lineages. *Nat. Commun.* **5**, 4758 (2014).
8. Karthaus, W. R. et al. Identification of multipotent luminal progenitor cells in human prostate organoid cultures. *Cell* **159**, 163–175 (2014).
9. Kwon, O. J., Zhang, L. & Xin, L. Stem Cell Antigen-1 Identifies a Distinct Androgen-Independent Murine Prostatic Luminal Cell Lineage with Bipotent Potential. *Stem cells (Dayt., Ohio)* **34**, 191–202 (2016).
10. Zhang, D. et al. Histone 2B-GFP Label-Retaining Prostate Luminal Cells Possess Progenitor Cell Properties and Are Intrinsically Resistant to Castration. *Stem cell Rep.* **10**, 228–242 (2018).
11. Wang, X. et al. A luminal epithelial stem cell that is a cell of origin for prostate cancer. *Nature* **461**, 495–500 (2009).
12. Choi, N., Zhang, B., Zhang, L., Ittmann, M. & Xin, L. Adult murine prostate basal and luminal cells are self-sustained lineages that can both serve as targets for prostate cancer initiation. *Cancer Cell* **21**, 253–265 (2012).
13. Ousset, M. et al. Multipotent and unipotent progenitors contribute to prostate postnatal development. *Nat. Cell Biol.* **14**, 1131–1138 (2012).
14. Wang, Z. A. et al. Lineage analysis of basal epithelial cells reveals their unexpected plasticity and supports a cell-of-origin model for prostate cancer heterogeneity. *Nat. Cell Biol.* **15**, 274–283 (2013).
15. Chua, C. W. et al. Single luminal epithelial progenitors can generate prostate organoids in culture. *Nat. Cell Biol.* **16**, 951–961 (2014). 951-954.
16. Yoo, Y. A. et al. Bmi1 marks distinct castration-resistant luminal progenitor cells competent for prostate regeneration and tumour initiation. *Nat. Commun.* **7**, 12943 (2016).
17. Vasioukhin, V., Degenstein, L., Wise, B. & Fuchs, E. The magical touch: genome targeting in epidermal stem cells induced by tamoxifen application to mouse skin. *Proc. Natl Acad. Sci. USA* **96**, 8551–8556 (1999).
18. Kretzschmar, K. & Watt, F. M. Markers of epidermal stem cell subpopulations in adult mammalian skin. *Cold Spring Harb. Perspect. Med.* **4**, a013631 (2014).

19. Pignon, J. C. et al. p63-expressing cells are the stem cells of developing prostate, bladder, and colorectal epithelia. *Proc. Natl Acad. Sci. USA* **110**, 8105–8110 (2013).

20. Shin, K. et al. Hedgehog/Wnt feedback supports regenerative proliferation of epithelial stem cells in bladder. *Nature* **472**, 110–114 (2011).

21. Shin, K. et al. Cellular origin of bladder neoplasia and tissue dynamics of its progression to invasive carcinoma. *Nat. Cell Biol.* **16**, 469–478 (2014).

22. Stingl, J. et al. Purification and unique properties of mammary epithelial stem cells. *Nature* **439**, 993–997 (2006).

23. Shackleton, M. et al. Generation of a functional mammary gland from a single stem cell. *Nature* **439**, 84–88 (2006).

24. Moad, M. et al. Multipotent Basal Stem Cells, Maintained in Localized Proximal Niches, Support Directed Long-Ranging Epithelial Flows in Human Prostates. *Cell Rep.* **20**, 1609–1622 (2017).

25. Kwon, O. J. & Xin, L. Prostate epithelial stem and progenitor cells. *Am. J. Clin. Exp. Urol.* **2**, 209–218 (2014).

26. Goldstein, A. S., Huang, J., Guo, C., Garraway, I. P. & Witte, O. N. Identification of a cell of origin for human prostate cancer. *Science* **329**, 568–571 (2010).

27. Xin, L., Lukacs, R. U., Lawson, D. A., Cheng, D. & Witte, O. N. Self-renewal and multilineage differentiation in vitro from murine prostate stem cells. *Stem Cells* **25**, 2760–2769 (2007).

28. Stoyanova, T. et al. Prostate cancer originating in basal cells progresses to adenocarcinoma propagated by luminal-like cells. *Proc. Natl Acad. Sci. USA* **110**, 20111–20116 (2013).

29. English, H. F., Santen, R. J. & Isaacs, J. T. Response of glandular versus basal rat ventral prostatic epithelial cells to androgen withdrawal and replacement. *Prostate* **11**, 229–242 (1987).

30. Aigner, K. et al. The transcription factor ZEB1 (deltaEF1) promotes tumour cell dedifferentiation by repressing master regulators of epithelial polarity. *Oncogene* **26**, 6979–6988 (2007).

31. Shibue, T. & Weinberg, R. A. EMT, CSCs, and drug resistance: the mechanistic link and clinical implications. *Nat. Rev. Clin. Oncol.* **14**, 611–629 (2017).

32. Mayer, C. et al. Developmental diversification of cortical inhibitory interneurons. *Nature* **555**, 457–462 (2018).

33. Zheng, C. et al. Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. *Cell* **169**, 1342–1356 e1316 (2017).

34. Nieto, M. A., Huang, R. Y., Jackson, R. A. & Thiery, J. P. EMT: 2016. *Cell* **166**, 21–45 (2016).

35. Zhang, L., Zhao, W., Valdez, J. M., Creighton, C. J. & Xin, L. Low-density Taqman miRNAs array reveals differentially expressed in prostatic stem cells and luminal cells. *Prostate* **70**, 297–304 (2010).

36. Mani, S. A. et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* **133**, 704–715 (2008).

37. Li, P. et al. Zeb1 promotes androgen independence of prostate cancer via induction of stem cell-like properties. *Exp. Biol. Med.* **239**, 813–822 (2014).

38. Salm, S. N. et al. TGF-{beta} maintains dormancy of prostatic stem cells in the proximal region of ducts. *J. Cell Biol.* **170**, 81–90 (2005).

39. Tsujimura, A. et al. Proximal location of mouse prostate epithelial stem cells: a model of prostatic homeostasis. *J. Cell Biol.* **157**, 1257–1265 (2002).

40. Donjacour, A. A. & Cunha, G. R. Assessment of prostatic protein secretion in tissue recombinants made of urogenital sinus mesenchyme and urothelium from normal or androgen-insensitive mice. *Endocrinology* **132**, 2342–2350 (1993).

41. Xin, L., Ide, H., Kim, Y., Dubey, P. & Witte, O. N. In vivo regeneration of murine prostate from dissociated cell populations of postnatal epithelia and urogenital sinus mesenchyme. *Proc. Natl Acad. Sci. USA* **100**(Suppl 1), 11896–11903 (2003).

42. Richardson, G. D. et al. CD133, a novel marker for human prostatic epithelial stem cells. *J. Cell Sci.* **117**, 3539–3545 (2004).

43. Jiao, J. et al. Identification of CD166 as a surface marker for enriching prostate stem/progenitor and cancer initiating cells. *PloS ONE* **7**, e42564 (2012).

44. Wang, D. et al. Identification of multipotent mammary stem cells by protein C receptor expression. *Nature* **517**, 81–84 (2015).

45. Takagi, T., Moribe, H., Kondoh, H. & Higashi, Y. DeltaEF1, a zinc finger and homeodomain transcription factor, is required for skeleton patterning in multiple lineages. *Development* **125**, 21 (1998).

46. Zhang, K. et al. WNT/beta-Catenin Directs Self-Renewal Symmetric Cell Division of hTERT(high) Prostate Cancer Stem Cells. *Cancer Res* **77**, 2534–2547 (2017).

47. Scheel, C. et al. Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. *Cell* **145**, 926–940 (2011).

48. Ye, X. et al. Distinct EMT programs control normal mammary stem cells and tumour-initiating cells. *Nature* **525**, 256–260 (2015).

49. Guen, V. J. et al. EMT programs promote basal mammary stem cell and tumor-initiating cell stemness by inducing primary ciliogenesis and Hedgehog signaling. *Proc. Natl Acad. Sci. USA* **114**, E10532–e10539 (2017).

50. Guo, W. et al. Slug and Sox9 cooperatively determine the mammary stem cell state. *Cell* **148**, 1015–1028 (2012).

51. Kurita, T., Medina, R. T., Mills, A. A. & Cunha, G. R. Role of p63 and basal cells in the prostate. *Development* **131**, 4955–4964 (2004).

52. Li, P., Yang, R. & Gao, W. Q. Contributions of epithelial-mesenchymal transition and cancer stem cells to the development of castration resistance of prostate cancer. *Mol. cancer* **13**, 55 (2014).

53. Barros-Silva, J. D. et al. Single-Cell Analysis Identifies LY6D as a Marker Linking Castration-Resistant Prostate Luminal Cells to Prostate Progenitors and Cancer. *Cell Rep.* **25**, 3504–3518.e3506 (2018).

54. Henry, G. H. et al. A Cellular Anatomy of the Normal Adult Human Prostate and Prostatic Urethra. *Cell Rep.* **25**, 3530–3542.e3535 (2018).

55. Sergushichev, A. A. et al. GAM: a web-service for integrated transcriptional and metabolic network analysis. *Nucleic acids Res.* **44**, W194–W200 (2016).

56. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).

57. Huper, G. & Marks, J. R. Isogenic normal basal and luminal mammary epithelial isolated by a novel method show a differential response to ionizing radiation. *Cancer Res* **67**, 2990–3001 (2007).

58. Schaefer, C. F. et al. PID: the Pathway Interaction Database. *Nucleic acids Res.* **37**, D674–D679 (2009).

59. Wong, D. J. et al. Module map of stem cell genes guides creation of epithelial cancer stem cells. *cell stem cell* **2**, 333–344 (2008).

60. Lim, E. et al. Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways. *Breast cancer Res.: BCR* **12**, R21 (2010).

61. Zhang, D. et al. Stem cell and neurogenic gene-expression profiles link prostate basal cells to aggressive prostate cancer. *Nat. Commun.* **7**, 10798 (2016).

62. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).

63. Lukacs, R. U., Goldstein, A. S., Lawson, D. A., Cheng, D. & Witte, O. N. Isolation, cultivation and characterization of adult murine prostate stem cells. *Nat. Protoc.* **5**, 702–713 (2010).

Acknowledgements

Author contributions

H.H.Z. and W.Q.G. conceived the study; X.W. performed the experiments. H.X. and D. W. conducted bioinformatic analyses; Z.J. and J.W. assisted in data interpretation; C.C. and Y.S. helped in cell culture experiments; H.Z. helped microscopic imaging; X.L. performed mouse genotyping and breeding; Y.S. helped organoid culture; Y.H. assisted in animal experiments. L.F., B.D. and W.X. collected human prostate samples and assisted in histological analysis of human sample; H.H.Z., W.Q.G. and X.W. interpreted the data and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-14296-y.

Correspondence and requests for materials should be addressed to W.-Q.G. or H.H.Z.

Peer review information *Nature Communications* thanks Dean Tang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.