

Normal and cancerous mammary stem cells evade interferon-induced constraint through the miR-199a–LCOR axis

Toni Celià-Terrassa¹, Daniel D. Liu¹, Abrar Choudhury¹, Xiang Hang¹, Yong Wei¹, Jose Zamalloa^{1,2}, Raymundo Alfaro-Aco¹, Rumela Chakrabarti¹, Yi-Zhou Jiang^{3,4}, Bong Ihn Koh¹, Heath A. Smith¹, Christina DeCoste¹, Jun-Jing Li^{3,4}, Zhi-Ming Shao^{3,4} and Yibin Kang^{1,5}

Tumour-initiating cells, or cancer stem cells (CSCs), possess stem-cell-like properties observed in normal adult tissue stem cells. Normal and cancerous stem cells may therefore share regulatory mechanisms for maintaining self-renewing capacity and resisting differentiation elicited by cell-intrinsic or microenvironmental cues. Here, we show that miR-199a promotes stem cell properties in mammary stem cells and breast CSCs by directly repressing nuclear receptor corepressor LCOR, which primes interferon (IFN) responses. Elevated miR-199a expression in stem-cell-enriched populations protects normal and malignant stem-like cells from differentiation and senescence induced by IFNs that are produced by epithelial and immune cells in the mammary gland. Importantly, the miR-199a–LCOR–IFN axis is activated in poorly differentiated ER[−] breast tumours, functionally promotes tumour initiation and metastasis, and is associated with poor clinical outcome. Our study therefore reveals a common mechanism shared by normal and malignant stem cells to protect them from suppressive immune cytokine signalling.

The mammary gland epithelium is a hierarchically organized tissue with multipotent mammary stem cells (MaSCs) capable of generating luminal and basal epithelial cells¹. It has been hypothesized that regulators of normal stem cell activity may be exploited by tumour-initiating cells (TICs) or cancer stem cells (CSCs)². Indeed, recent studies have revealed several cell fate regulators as such molecular links between MaSCs and breast TICs that drive their renewal activity in both normal and cancerous mammary gland tissues^{3–5}. Poorly differentiated tumours, typically basal-like/clinically low or triple-negative breast tumours^{6,7}, have high TIC activity and are enriched in CD24[−]/CD44⁺ breast CSCs⁸, which resemble some features of normal MaSCs^{5,9,10}.

Besides their cell-intrinsic self-renewal ability, normal stem cells need to adopt additional mechanisms to fend off microenvironmental pressure that may deplete the stem cell pool. Although immune cells have been reported to be critical players in mammary gland development¹¹, it is unknown how MaSCs control their interaction with the immune cells to sustain their stem cell activity. Interestingly, stem cells have been shown to downregulate immunogenic factors, such as major histocompatibility complexes (MHCs), to protect themselves from immune surveillance and ensure

tissue regeneration¹². This mechanism is also used by tumour cells to evade the immune system^{13,14}. However, the molecular mechanism underlying immune evasion by normal and cancerous stem cells in adult tissues remains poorly understood.

MicroRNAs (miRNAs) are critical regulators of development and cancer^{15,16}. MiRNAs have been shown to be expressed in a cell-lineage-specific fashion in the mammary gland¹⁷, and are functionally involved in mammary gland development^{5,18–21}. Likewise, miRNAs display distinct expression patterns in different subtypes of breast cancer and are known to promote or suppress tumorigenesis^{22–24} and regulate breast CSCs²⁵. Despite this progress, relatively little is known about how miRNAs regulate the interaction of stem cells with the immune microenvironment.

In this study, we identified a critical role of miR-199a in promoting MaSC and TIC properties by direct repression of LCOR, a nuclear receptor corepressor that sensitizes cells to interferon-induced differentiation and senescence. The miR-199a–LCOR axis represents a conserved molecular pathway in normal and cancer stem cells that mediates their evasion from autocrine and immune microenvironment suppressive signals.

¹Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544, USA. ²Lewis-Sigler Institute, Princeton University, Princeton, New Jersey 08544, USA. ³Department of Breast Surgery, Fudan University Shanghai Cancer Center, Shanghai 200032, China. ⁴Department of Oncology, Shanghai Medical College, Fudan University, Shanghai 200032, China.

⁵Correspondence should be addressed to Y.K. (e-mail: ykang@princeton.edu)

RESULTS

Systematic screening reveals the MaSC-promoting activity of miR-199a

To identify candidate miRNA regulators of MaSCs, we performed miRNA profiling of the Lin⁻CD24⁺CD29^{high} MaSC-enriched basal population (hereafter denoted as P4) and the Lin⁻CD24⁺CD29^{low} luminal population (P5) of primary mammary epithelial cells (MECs) isolated by fluorescence-activated cell sorting (FACS) from the mouse mammary gland (Supplementary Fig. 1a). Twenty-three miRNAs were significantly upregulated in P4 versus P5 by more than twofold (Fig. 1a). On the basis of the degree of differential expression of miRNAs and their predicted messenger RNA targets of interest, we selected 7 of the 23 upregulated miRNAs (miR-204, miR-211, miR-1a, miR-133a, miR-133b, miR-199a and miR-23a) for further functional analysis. We used lentiviral vectors to transduce primary MECs to test these candidates in mammosphere assays *in vitro* and cleared fat pad (CFP) reconstitution assays *in vivo* (Fig. 1b and Supplementary Fig. 1a). Interestingly, only miR-199a overexpression (OE) led to a significant increase in both assays (Fig. 1b). We confirmed by quantitative PCR (qPCR) higher expression of both mature forms (3p and 5p) of miR-199a in P4 versus P5 cells (Fig. 1c). *In situ* hybridization (ISH) confirmed elevated expression of miR-199a in basal cells compared with luminal cells in the mammary gland (Fig. 1d).

We next examined the ability of miR-199a to induce MaSC activity in total MECs (P3), P4 or P5 cells (Fig. 1e,f and Supplementary Fig. 1b,c). Ectopic miR-199a-OE enhanced stem cell activity of P4 cells in CFP repopulation assays (Fig. 1e) while the opposite was observed with miR-199a knockdown (KD) (Supplementary Fig. 1d). Importantly, ectopic miR-199a expression in P5 cells also increased the repopulation frequency, indicating that miR-199a can induce a stem cell-like state in luminal cells (Fig. 1f). Characterization of the repopulated mammary gland tissues showed an increase in the basal/MaSC marker Keratin14 (Fig. 1g and Supplementary Fig. 1e,f). Serial passage and transplantation assays further confirmed that an increased and sustained capacity of sphere formation, basal phenotype, and regenerating ability was induced by miR-199a in three successive generations of passage (Fig. 1h,i and Supplementary Fig. 1g). Finally, to investigate the relevance of miR-199a in MaSC populations, we used the Lgr5 knock-in EGFP reporter mouse²⁶ to isolate Lgr5⁺ and Lgr5⁻ cells. A significant increase of miR-199a expression was observed in Lgr5⁺ versus Lgr5⁻ P4 cells (Fig. 1j), consistent with reported higher MaSC activity of the Lgr5⁺ P4 population²⁶. Overall, these results suggest that miR-199a functionally promotes MaSC activity.

miR-199a induces stem-cell-like gene signatures and is upregulated in CSC populations

To explore the downstream signalling of miR-199a, we used the immortalized human MEC line HMLE to stably overexpress miR-199a and perform expression profiling. Gene set enrichment analysis (GSEA) revealed that the miR-199a overexpression resulted in enrichment of gene sets related to MaSC²⁷, CSC²⁸, undifferentiated tumour cell populations and the claudin-low (CL) breast cancer signature⁷ (Fig. 2a). Conversely, negative enrichments of luminal differentiated signatures, CSC downregulated genes, undifferentiated

and CL downregulated gene sets were observed after miR-199a overexpression (Fig. 2a). To further investigate the connection of miR-199a with MaSCs and breast cancer, we next applied a Euclidean centroid-based CL predictor and stratified the TCGA data set including the CL subtype, which has been linked to MaSCs on the basis of transcriptomic studies⁷. Notably, we observed a large overlap between CL tumours and MaSCs at the miRNA level, with miR-199a being one of the most significantly upregulated common miRNAs among MaSCs and CL tumours (Supplementary Fig. 2a and Supplementary Table 1).

MaSCs resemble cells that undergo epithelial–mesenchymal transition (EMT) on the basis of gene expression signatures and activated pathways²⁹. Additionally, gene expression profiles of MaSC populations overlap with claudin-low breast cancers, which harbour EMT-like properties and are enriched for TICs^{27,30}. Interestingly, although miR-199a is downstream of EMT signalling, being upregulated by TGF- β and TWIST1 (Supplementary Fig. 2b) as previously reported³¹, it did not induce EMT in HMLE cells by itself (Fig. 2b,c and Supplementary Fig. 2c). However, miR-199a strongly potentiated sphere formation (Fig. 2d) and increased the expression of several stem cell transcription factors and markers (Fig. 2b,e and Supplementary Fig. 2d), indicating miR-199a as a prominent downstream effector of EMT signalling to induce stemness while having no direct role in influencing epithelial or mesenchymal phenotypes. This is consistent with recent findings suggesting that TWIST1 controls tumour maintenance and stemness independently of its function in regulating EMT³², and the existence of mesenchymal-like¹⁰ and epithelial-like CSCs^{33,34}.

To explore whether miR-199a might also be important in TIC/CSC populations, we isolated different TIC populations in various breast cancer models. We used Twist1-induced TICs in HMLE-Neu cells¹⁰, CD24⁺/Thy1⁺ cells in the MMTV-Wnt-1 mouse breast tumour³⁵, and the CD24⁻/CD44⁺ population⁸ in HCI-002 patient-derived xenografts (PDX)³⁶. In all of these TIC/CSC populations, miR-199a displayed elevated expression compared with the bulk population (Fig. 2f–h). These data are consistent with a previous study showing upregulation of miR-199a in human breast CSCs⁵. Overall, these observations suggest miR-199a to be an important regulator of MaSCs and breast CSCs.

LCOR is a direct miR-199a target that suppresses MaSC activity and is downregulated in MaSCs and CSCs

To identify potential targets of miR-199a responsible for the observed phenotype, we focused on genes downregulated in MaSCs (P4) that are: predicted to have direct miR-199a-binding sites (Fig. 3a and Supplementary Table 2), and downregulated after miR-199a-OE in HMLE and NMuMG cells (Fig. 3b). On the basis of these criteria, we selected *Tox3*, *Rbm47* and *Lcor* as candidate functional targets of miR-199a (Fig. 3a). In functional assays for MaSC activity, only *Lcor*-KD increased both sphere formation *in vitro* and mammary gland reconstitution *in vivo* (Fig. 3c). In addition, we validated that *Lcor* is highly expressed in the luminal compartment (Fig. 3d,e and Supplementary Fig. 3a), and especially in mature luminal cells (P5-CD61⁻) compared with luminal progenitors (P5-CD61⁺) (Supplementary Fig. 3b). We next confirmed that transient or stable miR-199a-OE consistently represses *LCOR* in 10 different normal and

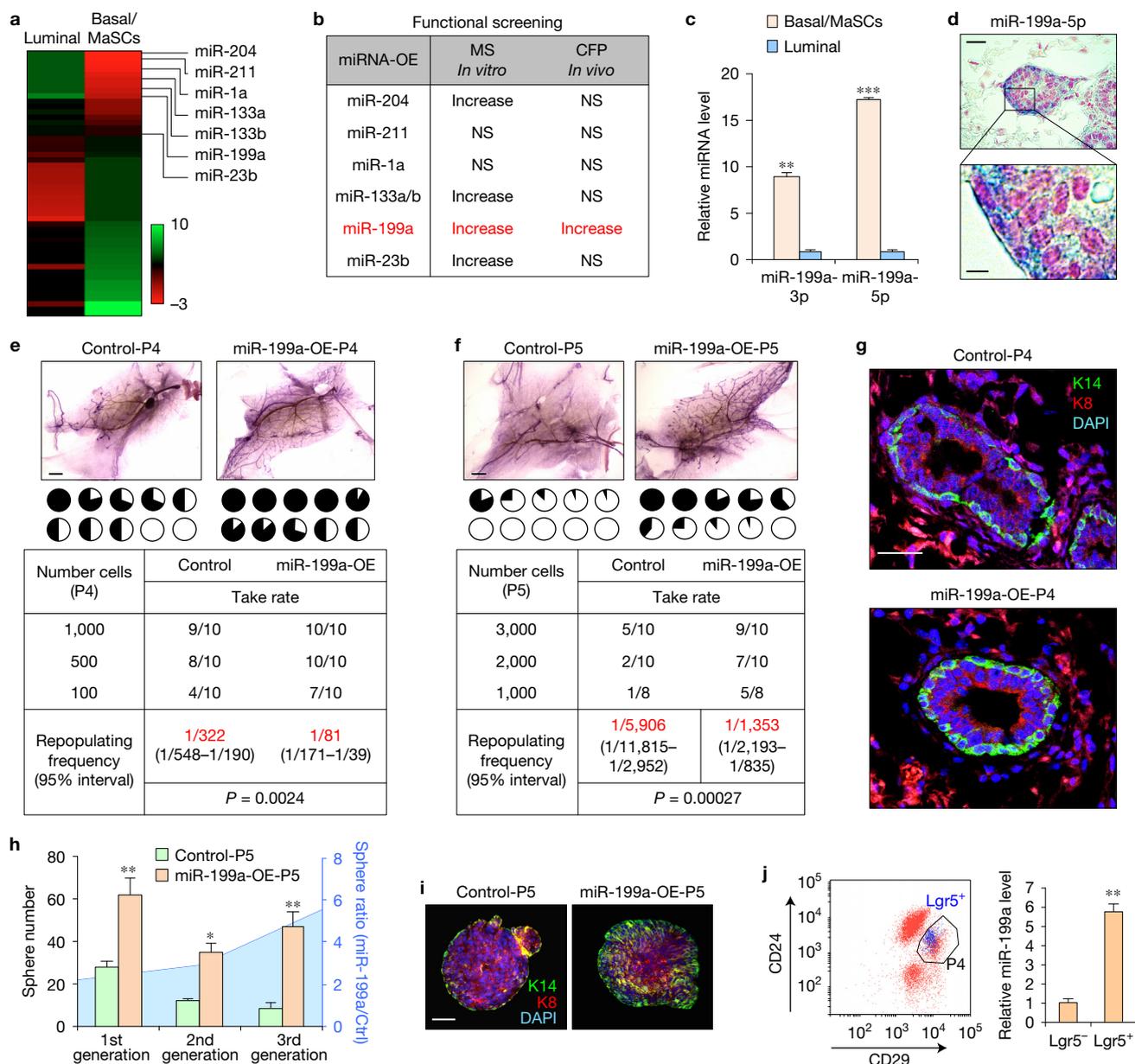


Figure 1 miR-199a is enriched in MaSCs and is functionally critical for MaSC activity. **(a)** Heat map representing miRNAs with >2-fold differential expression between P4 and P5 cells. **(b)** Table of selected miRNAs used for *in vitro* mammosphere (MS) and *in vivo* cleared fat pad (CFP) reconstitution analyses. NS, not significant. **(c)** qRT-PCR analysis of the expression levels of the 3' and 5' arms (3p and 5p) of miR-199a in P4 compared with P5. $n=4$ biologically independent samples; data represent mean \pm s.e.m. **(d)** ISH analysis of miR-199a-5p in the terminal end buds. miR-199a is stained blue and nuclei are stained in red. **(e,f)** P4 **(e)** and P5 **(f)** cells transduced with the indicated constructs are used for limiting dilution CFP reconstitution assay. Representative images show outgrowth. Each pie chart represents a mammary gland with the blackened area denoting the percentage of mammary gland outgrowth. Tables below represent serial dilution injections with the corresponding take rate. n indicates the number of mammary fat pad injections as indicated in the table. Shown in red are the repopulation frequencies for

each condition and P value by Pearson's Chi-squared test, obtained with the ELDA software. **(g)** Krt14 (K14—green) and Krt8 (K8—red) staining with reconstituted mammary outgrowths from control and miR-199a-OE P4 cells. **(h)** Number of P5 mammospheres formed after 3 generations of *in vitro* passage, and the ratio of sphere number between miR-199-OE group versus control. Five thousand cells in the indicated conditions were seeded ($n=3$ biologically independent samples; data represent mean \pm s.e.m.). **(i)** Confocal K14+K8 staining images of mammospheres from control and miR-199-OE P5 cells. **(j)** Left, flow cytometry isolation of P4-Lgr5⁺ and P4-Lgr5⁻ cells from the *Lgr5*-EGFP knock-in reporter mouse glands. Lgr5⁺ cells are represented as blue dots. Right, qRT-PCR analysis of miR-199a expression in Lgr5⁺ and Lgr5⁻ P4 cells ($n=3$ biologically independent samples; data represent mean \pm s.e.m.). Scale bars, 20 μ m (**d** upper panel), 5 μ m (**d** lower panel), 2 mm (**e,f**), 25 μ m (**g**) and 40 μ m (**i**). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$ by two-tailed Student's t -test in bar graphs.

malignant mammary cell lines derived from humans or mice (Fig. 3f,g and Supplementary Fig. 3c,d). Furthermore, to assess the direct repression of *Lcor* by miR-199a, we cloned the mouse *Lcor* 3'UTR

into a luciferase reporter plasmid. The *Lcor* 3'UTR is 8.3 kb long and contains 5 different evolutionarily conserved predicted binding sites for miR-199a: 2 sites for the miR-199a-3p and 3 sites for the

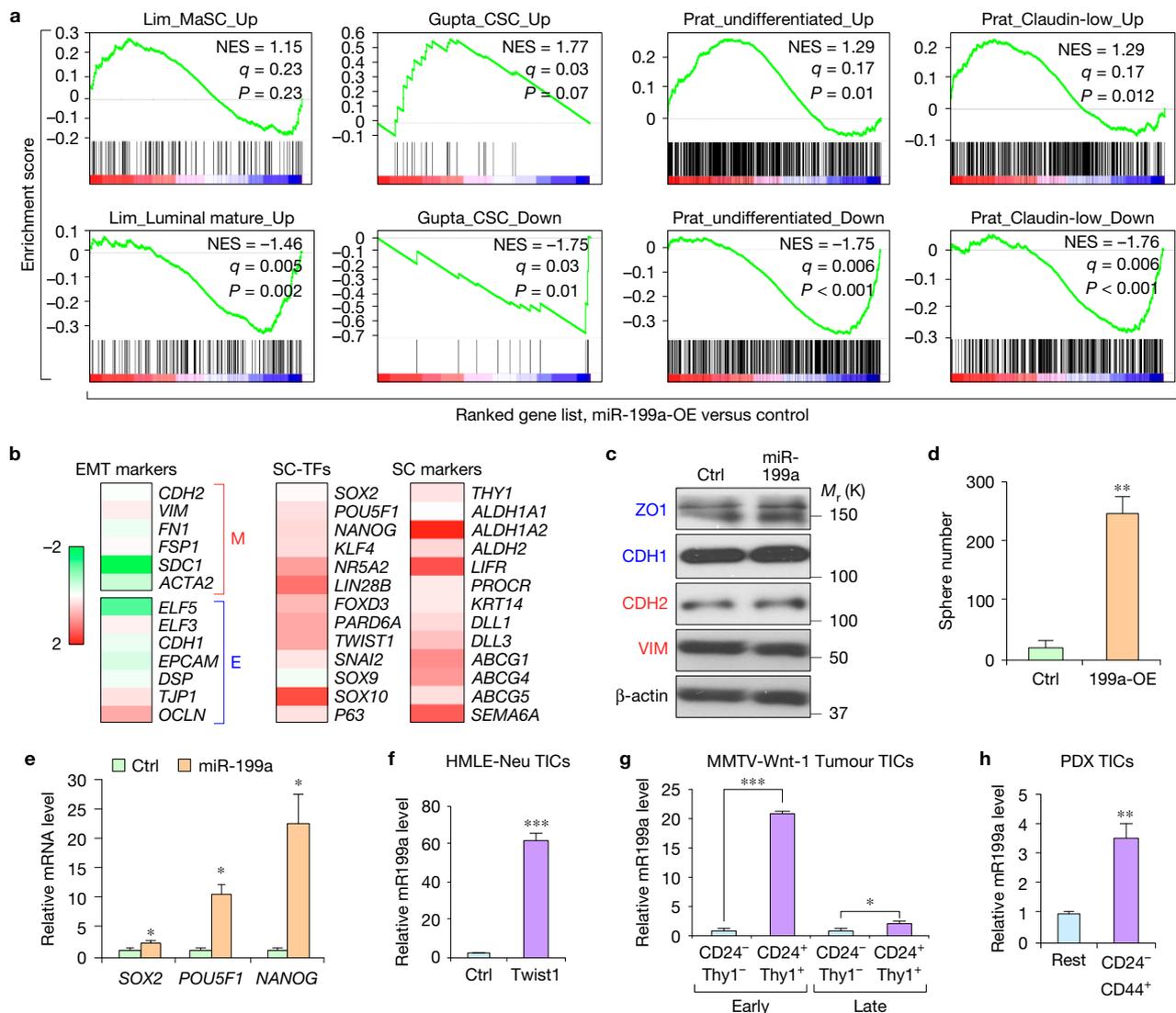


Figure 2 miR-199a induces stem-cell-like gene signatures and is enriched in cancer stem cells. **(a)** GSEA demonstrating the enrichment of gene sets related to MaSC²⁷, CSC²⁸, undifferentiated tumour cells⁷, and claudin-low tumours⁷ in the ranked gene list of miR-199a-OE versus control HMLE cells. NES, normalized enrichment score. **(b)** Heat map of HMLE-miR-199a-OE microarray data representing fold change expression of EMT markers, stem cell transcription factors (SC-TFs), and stem cell (SC) markers. Fold change is represented as \log_2 ratio. **(c)** Western blot analysis of epithelial (blue) and mesenchymal (red) markers. **(d)** *In vitro* quantification of mammospheres formed by 2,000

control or miR-199a-OE HMLE cells seeded. **(e)** qRT-PCR of mRNA extracted from 5-day HMLE control or miR-199a-OE mammospheres. **(f-h)** qRT-PCR of miR-199a levels in HMLE-Neu-Twist1-ER-OE tumour-initiating cells (TICs) **(f)**, CD24⁺/Thy1⁺ TICs isolated from early- and late-stage spontaneous MMTV-Wnt-1 tumours **(g)**, and CD24⁻/CD44⁺ TICs isolated from HCl-002 human breast cancer PDX **(h)** as compared with the non-TIC counterparts ($n = 3$ biologically independent samples; data represent mean \pm s.e.m.) in **d-h**. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$ by two-tailed Student's *t*-test in **d-h**. Unprocessed original scans of blots are shown in Supplementary Fig. 9.

miR-199a-5p (Fig. 3h). Due to the length of the 3'UTR, we cloned it as two separate fragments: UP and DOWN (Fig. 3h,i). We confirmed direct targeting of the three binding sites in DOWN 3'UTR that can be blocked by mutations; however, the two sites in UP 3'UTR did not show repression by miR-199a. Overall, these data implicate Lcor as a major candidate downstream effector of miR-199a.

We next examined the functional role of Lcor in the mammary gland. Ectopic overexpression of Lcor in P4 cells (Supplementary Fig. 4a) decreased MaSC activity as measured in sphere-forming (Fig. 4a) and reconstitution assays (Fig. 4b). Conversely, Lcor-KD in P5 cells increased MaSC activity and phenocopied miR-199a

overexpression (Fig. 4c and Supplementary Fig. 4b). Moreover, Lcor rescue in miR-199a-OE P4 cells completely nullified the MaSC-promoting effect of miR-199a, and suppressed the induction of stem cell transcription factors by miR-199a in HMLE cells (Fig. 4a,d and Supplementary Fig. 4). Likewise, Lcor-KD in HMLE cells increased sphere formation (Supplementary Fig. 4d,e), and Lcor-OE reduced it (Supplementary Fig. 4f). GSEA showed enrichment of the MaSC signature and the CL predictor gene set in HMLE-Lcor-KD cells versus control (Fig. 4e), consistent with the miR-199a-OE transcriptional changes. Moreover, Lcor is strongly downregulated in the Lgr5⁺ MaSC population (Fig. 4f). Consistent with its inverse

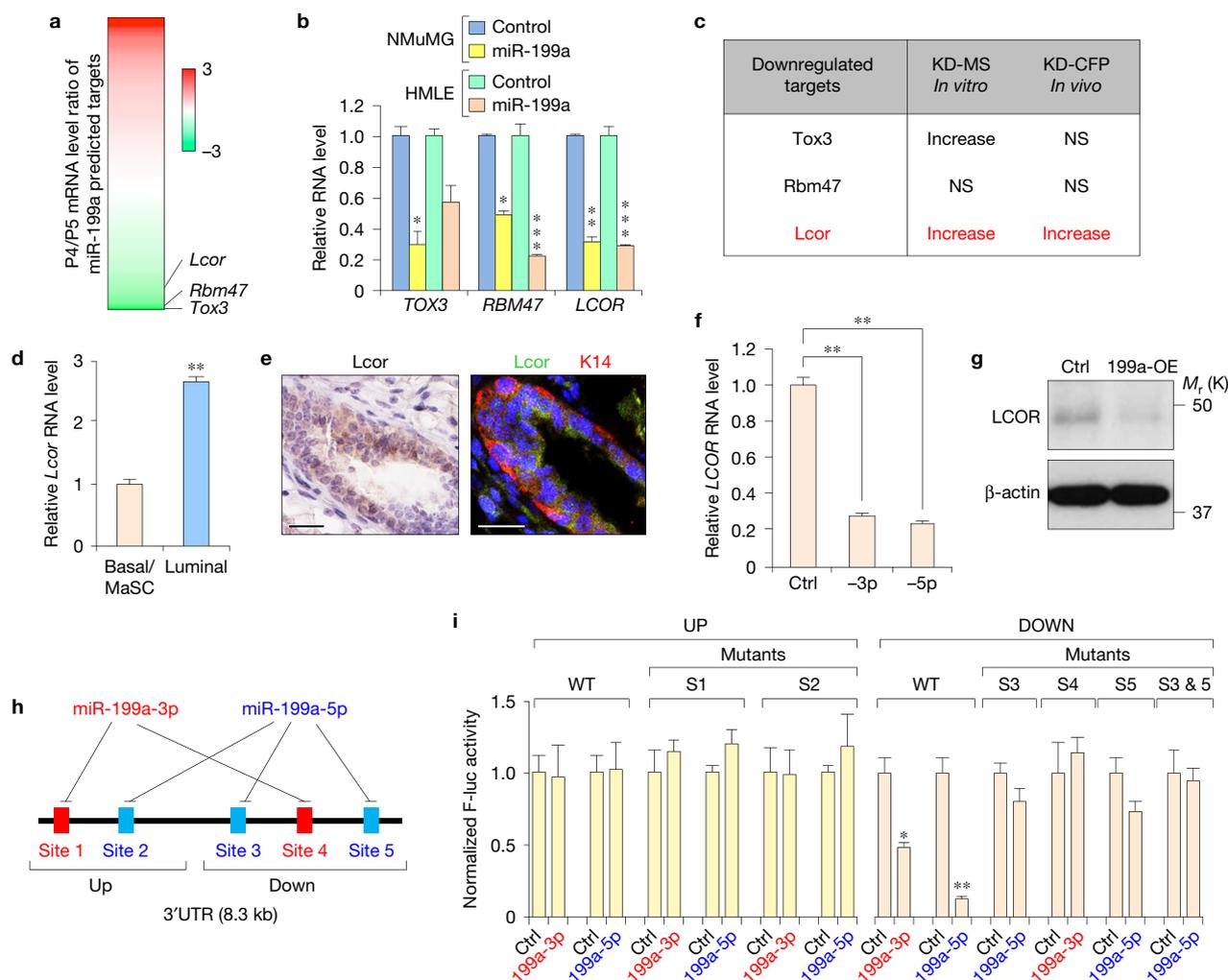


Figure 3 Identification of *LCOR* as a direct target gene of miR-199a. (a) Heat map showing the differential expression of predicted miR-199a target genes (TargetScan v7.0) in P4 versus P5 cells. Fold change is represented as \log_2 ratio. (b) qRT-PCR analysis of top candidate genes in mouse NMuMG and human HMLE cells after transfection with the indicated miRNA mimics. (c) Summary of the knockdown effect of the selected candidate miR-199a targets in functional assays using mouse MECs. MS, mammosphere formation assay; CFP, cleared fat pad mammary reconstitution assay. NS, not significant. (d) qRT-PCR analysis of *Lcor* expression in different lineages of mouse MECs. (e) IHC (left panel) analysis of *Lcor* and immunofluorescence (right panel) of *Lcor* and K14 localization in the mammary ducts. (f) qRT-PCR analysis of *LCOR* after 48 h

transfection of miR-199a-3p and miR-199a-5p in HMLE cells. (g) Western blot analysis of *LCOR* in control and miR-199a-OE HMLE cells. (h) Schematic diagram of the miR-199a-binding sites on the *LCOR* 3'UTR. (i) Normalized activity of the luciferase reporter containing the WT mouse *Lcor* 3'UTR or various miR-199a seed sequence mutants, after co-transfection with miR-199a-3p or miR-199a-5p, or control miRNAs in HeLa cells. The reporters were divided into two groups, UP and DOWN, containing upper and lower halves of the *Lcor* 3'UTR, respectively. Scale bars, 20 μm (e). $n = 3$ biologically independent samples; data represent mean \pm s.e.m. in b,d,f and i. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$ by two-tailed Student's *t*-test in bar graphs. Unprocessed original scans of blots are shown in Supplementary Fig. 9.

expression with miR-199a in normal mammary gland, we also observed that *Lcor* is downregulated in various TIC populations (Fig. 4g,h).

The miR-199a–*LCOR* axis promotes TIC activities in ER⁻ breast cancer

We evaluated the clinical relevance of the miR-199a–*LCOR* axis using the Buffa data set³⁷. Interestingly, miR-199a displayed a poor prognosis value and *LCOR* a good prognosis in the oestrogen-receptor-negative (ER⁻) patients but not in ER⁺ patients (Fig. 5a,b). We confirmed miR-199a and *LCOR* as independent prognostic markers in a triple-negative breast cancer (TNBC)-specific data

set³⁸ (Supplementary Fig. 5a,b). Additionally, *LCOR* mRNA is downregulated in TNBC compared with non-TNBC tumours (Supplementary Fig. 5c), and the analysis of 209 human breast tumour tissues showed high miR-199a and low *LCOR* levels in TNBC by ISH and immunohistochemistry (IHC), respectively (Fig. 5c), with inverse correlation to each other (Supplementary Fig. 5d). In the NKI295 data set, *LCOR* also shows good prognosis for distant metastasis-free survival (Supplementary Fig. 5e). Overall, these analyses indicate that the miR-199a–*LCOR* axis is clinically relevant in ER⁻ breast cancer.

To investigate the functional relevance of the miR-199a–*LCOR* axis in ER⁻ breast cancer, we next used PDX³⁶ in tumoursphere forming assays *in vitro* and mammary fat pad (MFP) tumorigenesis assay *in*

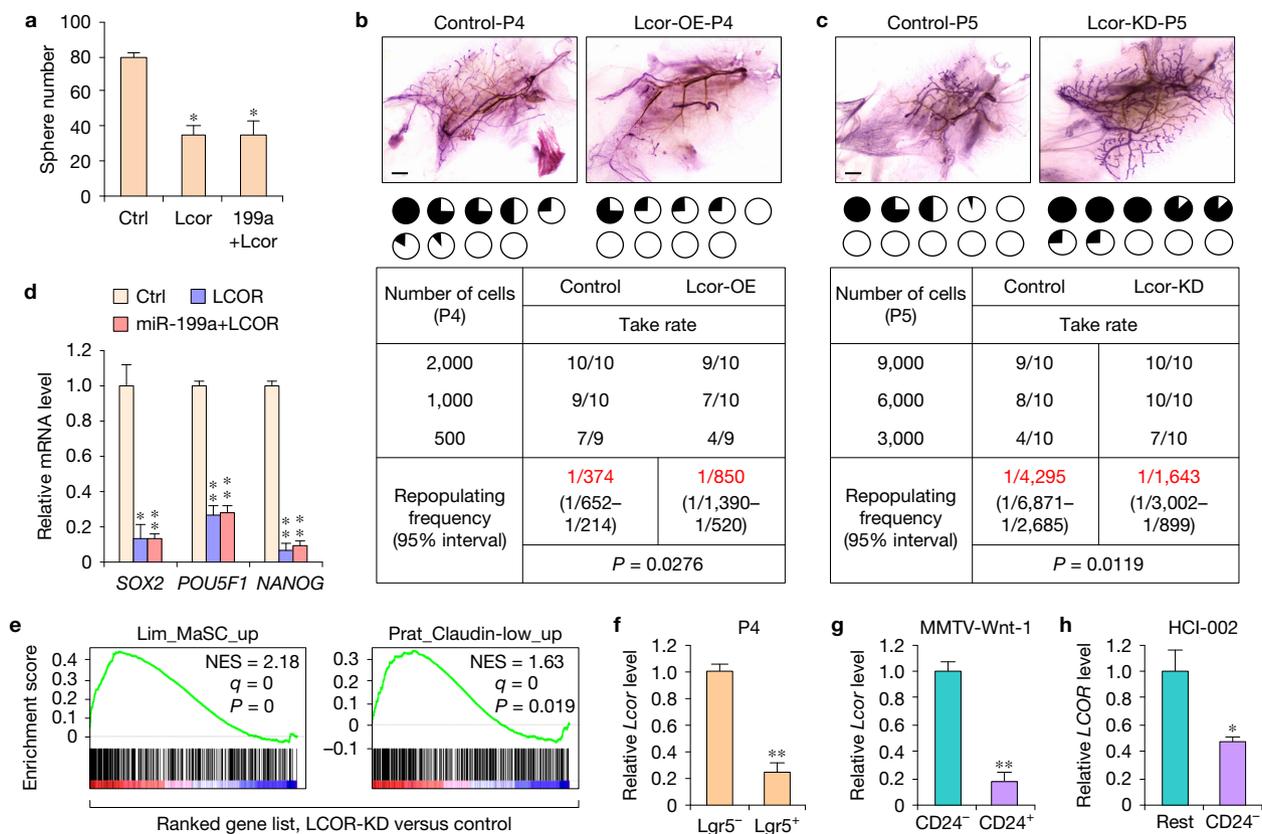


Figure 4 LCOR suppresses MaSC function and is downregulated in stem cell populations. **(a)** Quantification of mammosphere formation of 20,000 P4 cells in the indicated conditions ($n=3$ biologically independent samples; data represent mean \pm s.e.m.). **(b,c)** Limited dilution cleared fat pad reconstitution assay of P4 **(b)** and P5 **(c)** cells after transduction with the indicated constructs. Representative images show the outgrowth. Each pie chart represents a mammary gland with the blackened area showing the percentage of mammary gland outgrowth. Tables below represent serial dilution injections with the corresponding take rate. n indicates the number of mammary fat pad injections as indicated in the table. Shown in red are the repopulation frequencies for each condition and P value by Pearson's

Chi-squared test, obtained with the ELDA software. **(d)** qRT-PCR of mRNA extracted from mammospheres formed by HMLE cells after transduction with the indicated constructs ($n=3$ biologically independent samples; data represent mean \pm s.e.m.). **(e)** GSEA demonstrating the enrichment of gene sets related to MaSC²⁷ and claudin-low tumours⁷ in the ranked gene list of LCOR-KD versus control HMLE cells. **(f–h)** qRT-PCR analysis of *Lcor* expression in Lgr5⁺ MaSC-enriched P4 cells **(f)**, CD24⁺/Thy1⁺ MMTV-Wnt-1 TICs **(g)** and CD24⁻/CD44⁺ TICs isolated from HCI-002 PDX **(h)** as compared with their non-stem cell counterparts ($n=3$ biologically independent samples; data represent mean \pm s.e.m.). Scale bars, 2 mm **(b,c)**. * $P<0.05$, ** $P<0.01$ by two-tailed Student's t -test in bar graphs.

in vivo using NSG mice (Supplementary Fig. 5f). Intriguingly, ectopic miR-199a expression significantly increased tumoursphere formation of the TNBC PDXs (HCI-001; HCI-002; HCI-009; and HCI-010) but not in the ER⁺PR⁺ HCI-003 and HCI-005 (Fig. 5d). Likewise, LCOR-OE reduced tumoursphere formation in the TNBC PDXs, and nullified the tumoursphere-promoting effect of miR-199a (Fig. 5d). Consistent with the tumoursphere assays, miR-199a expression strongly increased tumour initiation of ER⁻ PDXs (Fig. 5e–g and Supplementary Fig. 5g,h), but not in the ER⁺ HCI-003 (Supplementary Fig. 5g). Conversely, LCOR-OE reduced tumour initiation in HCI-001, HCI-002 and HCI-010 (Fig. 5e–g). In addition, miR-199a-OE stimulated the development of systemic metastasis of the weakly metastatic 4T07 mouse tumour cells. Conversely, LCOR overexpression strongly suppressed metastasis of MDA-MB-231 breast cancer cells (Fig. 5h,i). Collectively, these data demonstrate that the miR-199a–LCOR axis, besides its function in regulating normal MaSCs, is also a critical modulator of CSCs/TICs in ER⁻ breast cancer, which are known to be enriched in TIC/CSCs.

miR-199a–LCOR modulates the IFN- α response of normal and cancerous stem cells

LCOR is a corepressor of agonist-bound nuclear receptors (NRs), but can also directly bind to DNA through its HTH domain³⁹. We generated a double-point mutation of the NR box (LSKLL to LSKAA) to abolish NR binding⁴⁰, and a HTH domain deletion mutant defective in direct DNA binding (Fig. 6a). Co-immunoprecipitation revealed near complete loss of interaction of LSKAA with the ER (Supplementary Fig. 6a), and immunofluorescence analysis confirmed that the Δ HTH mutant maintains nuclear localization (Supplementary Fig. 6b). Strikingly, disruption of the NR interaction did not diminish the ability of LCOR to suppress sphere formation of HMLE cells while Δ HTH lost such function (Fig. 6b). These results suggest that LCOR acts through DNA binding to suppress stem cell properties.

Global transcriptomic profiling clustered LCOR-OE cells together with LCOR-LSKAA, and LCOR- Δ HTH with control cells (Fig. 6c). Using GSEA, we observed that the most enriched gene sets in LCOR-OE cells were related to the IFN- α and IFN- γ responses

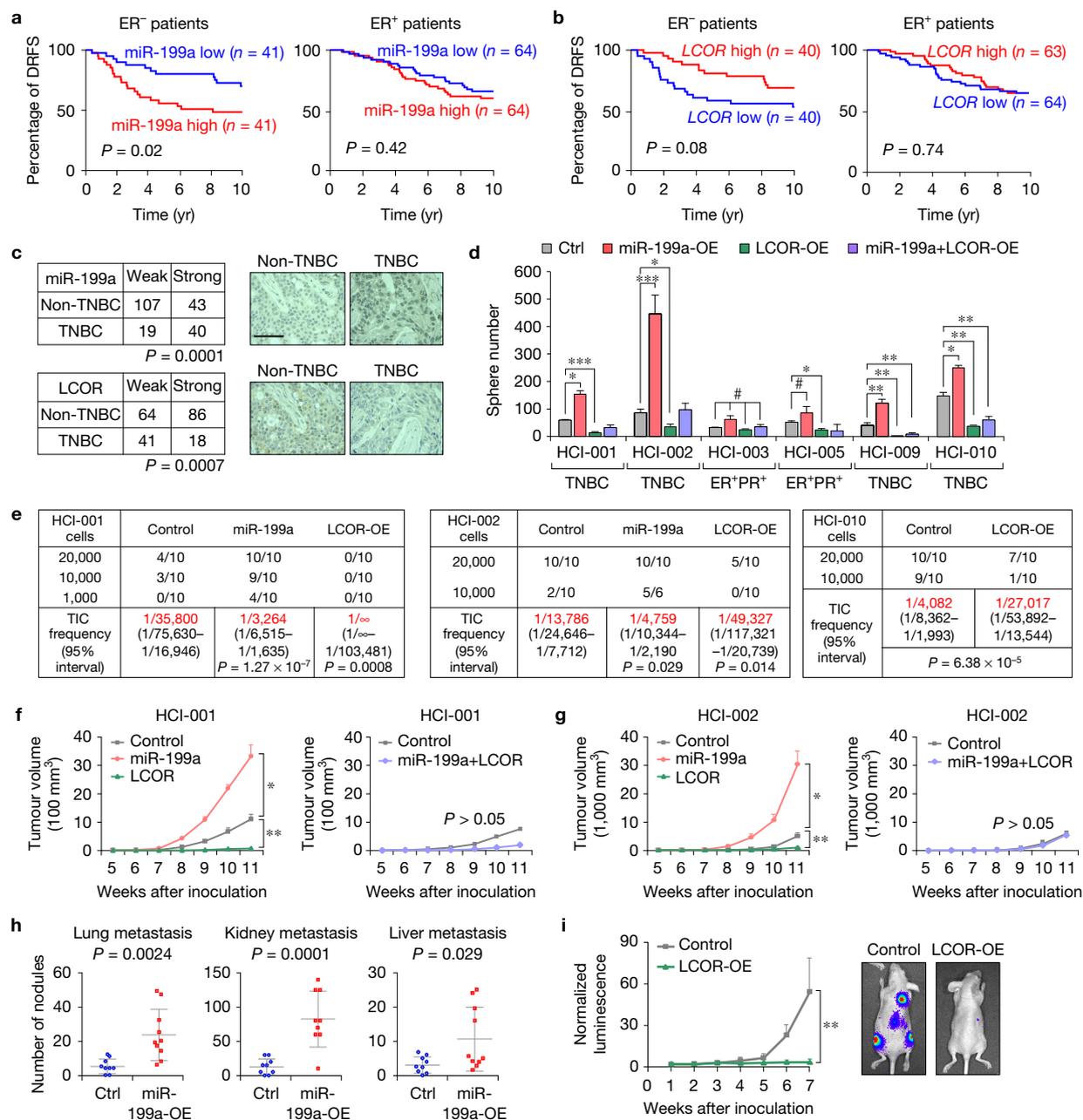


Figure 5 miR-199a and LCOR functionally influence the initiation of ER⁻ breast tumours *in vivo*. (**a,b**) Kaplan–Meier distant relapse-free survival (DRFS) curve of breast cancer patients with higher or lower than median RNA expression levels of miR-199a (**a**) and *LCOR* (**b**) in their tumours³⁷. (**c**) miR-199a and *LCOR* protein expression levels in TNBC ($n = 59$ patient samples) and non-TNBC tumours ($n = 150$ patient samples). Each sample was scored as weak (low expression) or strong (high expression) according to staining intensities of miR-199a by ISH and *LCOR* by IHC. (**d**) Quantification of tumourspheres formed by 10,000 cells from multiple human breast cancer PDXs in different tumour subtypes with the indicated conditions ($n = 3$ biologically independent samples; data represent mean \pm s.e.m.). (**e**) Tumour take rate of HCI-001, HCI-002 and HCI-010 following MFP injection of the indicated cells. n indicates the number of MFP injections

as indicated in the table. Tumour-initiating cell (TIC) frequency calculated by the ELDA software is indicated in red. (**f,g**) Tumour growth of HCI-001 (**f**) and HCI-002 (**g**) following MFP injection of 20,000 cells in the indicated conditions ($n = 10$ mouse mammary glands). (**h**) Metastatic nodule counts in the indicated organs 10 days after intracardiac injection of 100,000 4T07 cells in Balb/c mice ($n = 10$ mice). Each dot represents a value and the lines represent the mean and s.d. (**i**) Bioluminescence imaging quantification of the metastatic growth of the control and *LCOR*-KD MDA-MB-231 cells after intracardiac injection of 100,000 cells in Ncr-nu/nu mice ($n = 10$ mice). Scale bar, 100 μ m (**c**). P value by log-rank test in **a** and **b**, Fisher's exact test in **c**. P value by Pearson's Chi-squared test in **e**. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$ by two-tailed Student's t -test in **d** and **f–i**. # $P > 0.05$ in **d**.

(Supplementary Table 3). Importantly, the Δ HTH mutant, but not LSKAA, completely lost this enrichment of the IFN- α signature (Fig. 6d and Supplementary Table 3). Consistently, *LCOR*-KD HMLE

and MDA-MB-231 cells showed a strong negative enrichment of the IFN- α response gene set (Fig. 6e). Such negative enrichment is also observed in HMLE-miR-199a-OE cells, suggesting that miR-199a,

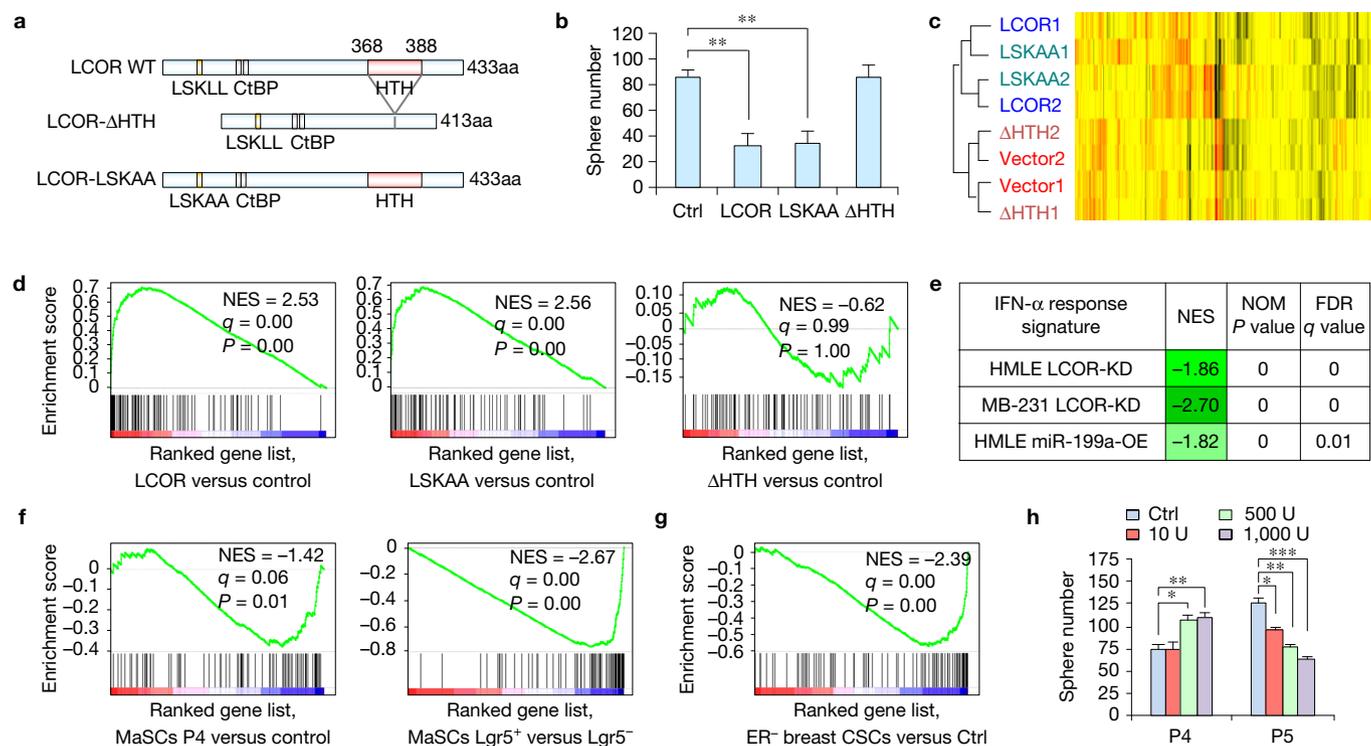


Figure 6 LCOR primes the IFN- α response. **(a)** Schematic representation of LCOR mutants. **(b)** Quantification of mammospheres formed by 5,000 HMLE cells with ectopic expression of the indicated LCOR constructs ($n=3$ biologically independent samples; data represent mean \pm s.e.m.). **(c)** Unsupervised hierarchical clustering of the HMLE expressing various LCOR constructs based on transcriptomic profiles. **(d)** GSEA of the IFN- α response gene set (M5911) in the ranked gene list of LCOR, LSKAA and Δ HTH versus control HMLE cells. **(e)** GSEA of the IFN- α response gene set in the ranked gene list of the LCOR-KD or miR-199a-OE

versus control HMLE and MDA-MB-231 cells. **(f)** GSEA of the IFN- α response gene set in the indicated gene list from the current study. **(g)** GSEA of the IFN- α response gene set in the gene list of CD24⁻/CD44⁺ CSCs versus non-CSCs in ER⁻ breast cancer⁹. **(h)** Quantification of mammospheres formed by P4 (20,000) and P5 (10,000) cells treated with different doses of IFN- α , from 10 to 1,000 U ml⁻¹ in 8 days ($n=3$ biologically independent samples; data represent mean \pm s.e.m.). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$ by two-tailed Student's t -test in **b** and **h**.

through LCOR repression, reprogrammes the transcriptome to suppress the IFN- α response. We further performed GSEA of the same gene sets in stem-cell-enriched populations versus non-stem cell populations. Strikingly, the IFN- α signature showed negative enrichment in P4-MaSC, and this negative enrichment bias was even more accentuated in the Lgr5⁺ MaSCs (Fig. 6f). Similar negative enrichment of the IFN- α signature was also observed in ER⁻ CD24⁻/CD44⁺ breast CSCs⁹ (Fig. 6g). These findings further suggest a muted interferon response in MaSCs and breast CSCs.

To directly evaluate the functional importance of interferon signalling in stem cell regulation, we performed mammosphere assays using P4 and P5 (luminal cells have high sphere formation capacity due to progenitor proliferation⁴¹) treated with IFN- α and IFN- γ . The P5 cells, which have higher LCOR levels, responded to the IFN- α treatment with reduced sphere formation ability (Fig. 6h and Supplementary Fig. 6c). In contrast, P4 cells showed an increase in sphere formation (Fig. 6h and Supplementary Fig. 6c), suggesting that MaSCs and luminal cells respond differently to IFN- α .

Interestingly, stem-cell-related genes (in red) were upregulated in P4 spheres and downregulated in P5 spheres following IFN- α treatment (Fig. 7a). Moreover, luminal differentiation genes (in blue) were upregulated in P5, and not in P4, following IFN- α treatment

(Fig. 7a). GSEA revealed enrichment of luminal and differentiated gene sets in P5 cells after IFN- α treatment, while MaSC and stemness gene sets were enriched in P4 treated cells (Fig. 7b). Consistent with the genomic profiling result, IFN- α induced the basal/MaSC marker K14 in P4 spheres, and the luminal marker K8 in Lcor-OE P4 spheres (Supplementary Fig. 7a). To further investigate the IFN- α effects *in vivo*, we performed CFP injections of P4 and P5, followed by IFN- α administration for 3 weeks. P4 cell reconstitution was slightly increased by the subcutaneous IFN- α treatment while P5 cell reconstitution was severely suppressed (Fig. 7c,d), again highlighting the differential response of P4 and P5 cells to IFN- α *in vivo*.

We next determined how miR-199a and ectopic LCOR expression affects these responses to IFN- α . Ectopic expression of miR-199a in P5 prevented mammosphere reduction following IFN- α treatment (Fig. 7e), while Lcor expression in P4 and HMLE cells sensitized them to the suppressive effect of IFN- α (Fig. 7f and Supplementary Fig. 7b). Similarly, IFN- α treatment in the HCI-010, a PDX highly enriched in TICs (Fig. 5e), did not reduce sphere formation; however, LCOR-OE sensitized HCI-010 to IFN- α -induced reduction of tumourspheres (Supplementary Fig. 7c,d). In contrast, IFN- α treatment in miR-199a-OE or LCOR-KD HCI-010 cells increased

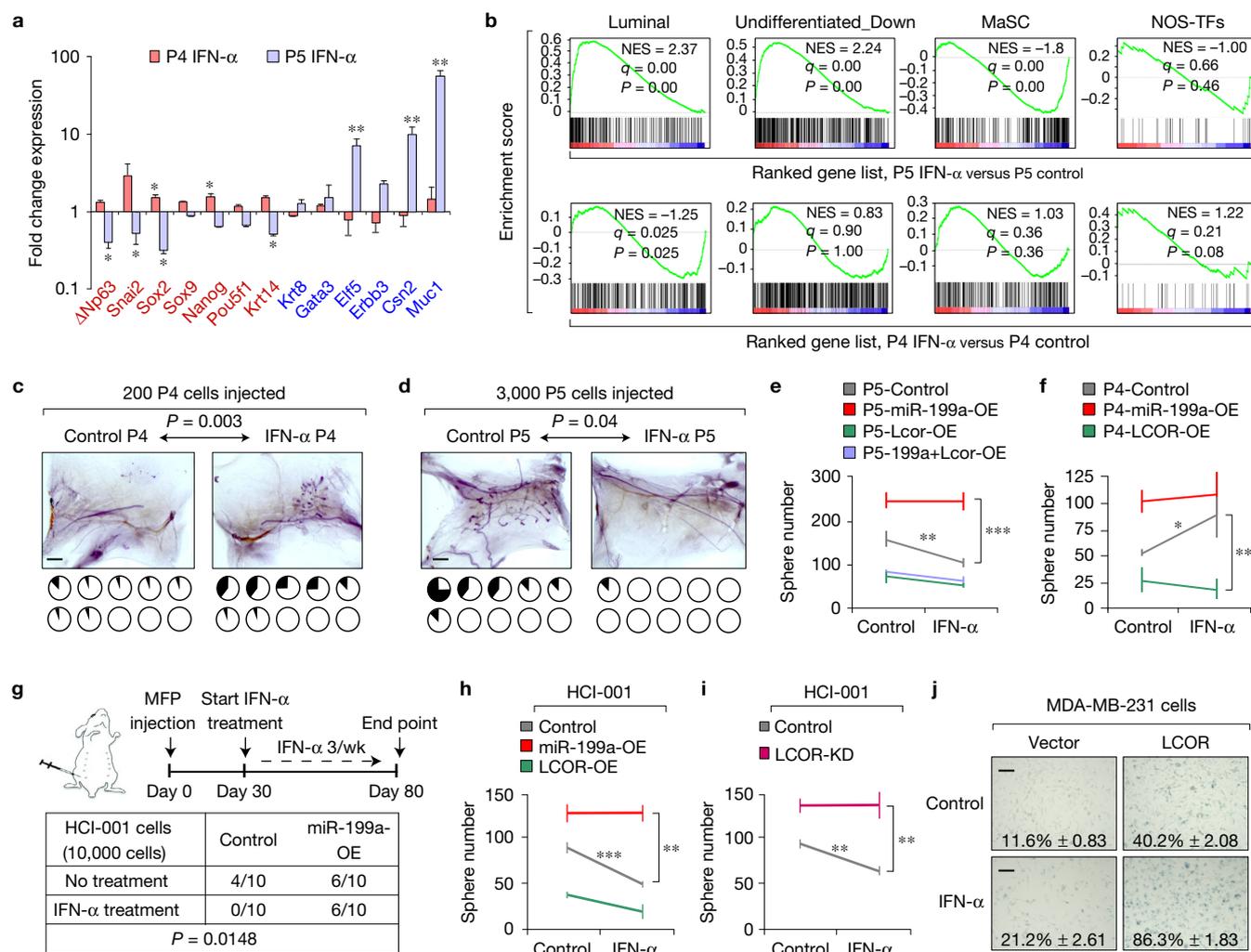


Figure 7 Stem cells and differentiated cells respond differently to the IFN- α signalling. **(a)** qRT-PCR result showing the fold change of stem-cell-related genes (red) and luminal differentiation-related genes (blue) in mammospheres formed by P4 and P5 with or without IFN- α 1,000 U ml $^{-1}$ treatment for 6 days ($n=3$ biologically independent samples; data represent mean \pm s.e.m. of fold change in IFN- α versus Ctrl). **(b)** GSEA of the MaSC and luminal upregulated gene sets generated in this study, as well as Nanog-Oct4-Sox2 transcriptional factor (NOS-TF) target gene set⁶ and undifferentiated downregulated genes⁷ in the ranked gene list of IFN- α -treated P5 and P4 cells versus control. **(c,d)** Cleared fat pad reconstitution assay of 200 P4 **(c)** and 3,000 P5 **(d)** cells after thrice a week treatment of 100,000 U IFN- α for 3 weeks. Representative images show the outgrowth. Each pie chart represents a mammary gland with the blackened area showing the percentage of mammary gland outgrowth ($n=10$ mammary glands injected); P value by two-tailed Student's t -test). **(e,f)** Quantification of mammospheres formed by 10,000 P5 **(e)** and P4 **(f)**

cells with or without IFN- α treatment after transduction of the indicated constructs ($n=3$ biologically independent samples; data represent mean \pm s.e.m.). Plots represent the enhanced difference of treated conditions versus the control conditions (without treatment). **(g)** Tumour take rate following MFP injection of 10,000 control and miR-199a-OE HCl-001 cells, with or without treatment with 100,000 U IFN- α as indicated in the schematics. P value calculated by one-way ANOVA of the tumour incidence. **(h,i)** Quantification of PDX cell tumourspheres formed by 10,000 HCl-001 cells, with the indicated conditions ($n=3$ biologically independent samples; data represent mean \pm s.e.m.). **(j)** Senescence-associated β -galactosidase (SA- β gal) assay of MDA-MB-231 cells comparing control and LCOR-OE cells, and with or without IFN- α treatment for 72 h ($n=3$ biologically independent samples; data represent mean \pm s.e.m.; source of data in Supplementary Table 4). Scale bars, 2 mm **(c,d)** and 100 μ m **(j)**. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$ by two-tailed Student's t -test in **a,e,f,h** and **i**.

sphere formation (Supplementary Fig. 7c,d), suggesting that the TIC-enriched population resists the interferon blockage on sphere formation. Another PDX line (HCl-001), which has relatively lower TIC activity, is sensitive to IFN- α treatment *in vivo* and *in vitro* (Fig. 7g-i), but was rendered insensitive to IFN- α after miR-199a-OE (Fig. 7m,n) or LCOR-KD (Fig. 7i).

Overexpression of WT but not the Δ HTH mutant of LCOR partially induced senescence in MDA-MB-231 and HMLE cells, and

strongly sensitizes them to interferon-mediated senescence (Fig. 7j and Supplementary Fig. 7e). Consistent with these findings, GSEA revealed a significant enrichment of the senescence gene set in LCOR-OE but not in Δ HTH-OE HMLE cells, and a negative enrichment in miR-199a-OE HMLE cells (Supplementary Fig. 7f). Overall, these findings indicate that the miR-199a-LCOR axis modulates the sensitivity of normal and cancerous stem cells to the differentiation and senescence effects of IFN- α .

MaSCs and breast CSCs are protected from suppressive effects of immune and autocrine IFN- α

To investigate the expression levels and sources of IFN- α in the mammary gland during different physiological and malignant states, we performed flow cytometry after co-staining intracellular IFN- α with various lineage markers of immune cells. Importantly, virgin mice already had a presence of IFN- α -expressing cells in the mammary gland and a substantial increase was observed during pregnancy, lactation and involution (Fig. 8a and Supplementary Fig. 8a). Interestingly, most of the IFN- α -expressing cells were macrophages across all of the different stages of the mammary gland, while T cells and dendritic cells were a relatively insignificant source of IFN- α (Fig. 8b). Importantly, high levels of infiltration (Fig. 8c) and activation (Fig. 8d) of macrophages in the mammary gland were observed during lactation and involution. In addition, the macrophages from the virgin mammary glands already expressed more interferons than peritoneal macrophages (Supplementary Fig. 8b). Besides immune cells, a small source of IFN- α -expressing cells corresponded to the Lin⁻CD24⁺ epithelial cells (Fig. 8b). This IFN- α -positive population increases significantly during pregnancy, consistent with the major expansion of the epithelial tissue at this stage. In breast tumours, different tumour types also showed elevated IFN- α levels and increased infiltration of IFN- α -positive macrophages (Fig. 8a,c,e and Supplementary Fig. 8a).

We next isolated F4/80⁺ macrophages from the mammary gland at different states to generate conditioned media (CM) and validated IFN- α secretion by enzyme-linked immunosorbent assay (ELISA) (Fig. 8e). Interestingly, CM of mammary gland macrophages significantly increased P4 and decreased P5 sphere formation, but such an effect was lost after miR-199a-OE in P5 cells (Fig. 8f,g). The effect of CM from mammary gland macrophages on P5 spheres was abolished after treatment with neutralizing antibodies against IFN- α / β (Fig. 8g), proving that these effects were mediated by IFNs. Moreover, neutralizing antibodies against IFN- α / β consistently increased P5 and P4-Lcor-OE sphere formation even without the presence of immune cells (Fig. 8h), indicating that P5 or P4-Lcor cells are also constrained by autocrine IFN- α / β signalling. Similarly, mammary gland macrophage CM also reduced the HCI-001 tumoursphere formation, and this was avoided by miR-199a-OE (Fig. 8i). These results demonstrate that, by virtue of elevated miR-199a expression and reduced LCOR levels, both normal and cancerous stem cells are protected from immune or autocrine/paracrine interferon-mediated suppressive effects.

To further explore the clinical significance of the miR-199a-LCOR-IFN-related regulatory pathway in breast cancer, we generated an IFN-Stem Cell-Down signature (ISDS) from the Interferon- α -response gene set (M5911) to represent interferon- α -responsive genes that are regulated by the miR-199a-LCOR axis and also contribute to negative enrichment of the M5911 in normal and cancerous stem cells. The combined 27 genes constitute the ISDS (Fig. 8j) and have a good prognosis value for relapse-free survival, overall survival and distant metastasis-free survival in ER⁻ breast cancer patients from the KM plotter data set⁴² (Fig. 8k-m). The individual analysis of the ISDS genes in ER⁻ breast cancer showed a good prognosis value for all of them, with 21 out of 27 being significant (Supplementary Fig. 8c), highlighting the clinical relevance of the miR-199a-LCOR-IFN axis in ER⁻ breast cancer.

DISCUSSION

In this study, we identified a miRNA-199a-mediated pathway shared by both MaSCs and breast CSCs to maintain their self-renewal competence and avoid differentiation or senescence induced by suppressive immune cytokines such as IFN- α (Fig. 8n). MiR-199a has been reported to have either tumour-suppressive functions^{43,44} or tumour-promoting activities⁴⁵ across different cancer types, including breast cancer^{46,47}. Here, we provide evidence for an important functional role of miR-199a in promoting MaSC activity by directly repressing LCOR, a nuclear receptor corepressor. LCOR has been proposed as a tumour suppressor in prostate cancer⁴⁸; however, there is no previous report of LCOR function in mammary gland development or breast cancer. We identified LCOR as a direct functional target of miR-199a in regulating MaSC and breast CSC activities. Our experimental and clinical data show that the miR-199a-LCOR axis mainly influences tumorigenesis of ER⁻ breast cancer, suggesting that LCOR is not acting through ER binding. This is also consistent with our results showing the maintenance of LCOR function in the LSKAA mutant. Therefore, the LCOR action on stem cells and breast cancer is independent of the ER, which is consistent with its function in MaSCs (which are ER⁻) and ER⁻ breast cancer.

Exploring the downstream effects of miR-199a-LCOR, we showed that LCOR negatively regulates stem cells by sensitizing them to interferon responses. Interferon signalling is known to be critical in anticancer immune surveillance of primary tumours and metastases^{49,50}. Interferons can also induce tumour cell-intrinsic inhibitory effects, including differentiation, growth arrest, and cell death^{50,51}. However, less is known about the effects of interferon on stem cells. Some studies have found different effects of IFN- α on haematopoietic stem cells, as it can activate dormant haematopoietic stem cells (HSCs) but inhibit active HSCs⁵², or drive exhaustion of quiescent HSCs⁵³. Therefore, IFNs can play opposite roles in stem cell fate in the haematopoietic system, depending on target cell status and on acute or chronic signalling⁵². Whether this is also the case in other adult stem cell systems and in cancer stem cells was unknown. Here we show that IFN can have different effects on MaSCs or differentiated cells, depending on the status of the miR-199a-LCOR axis. Remarkably, we found that IFN- α response is attenuated in normal and malignant stem cell populations on the basis of GSEA, indicating that low IFN sensitivity is a critical and general mechanism to maintain the stem cell phenotype.

Taken together, our study reveals a miR-199a-LCOR-IFN-dependent mechanism that is commonly used by MaSCs and CSCs to escape from differentiation and senescence induced by IFN signalling, which is particularly relevant during mammary gland lactation and involution⁵⁴, and in immune cell-rich claudin-low and TNBC tumours^{7,55}. Moreover, normal stem cells may use this mechanism to acquire immune privilege properties, as they do by downregulating MHCs to ensure tissue homeostasis. Since MHCs are regulated by IFN¹², stem cells may downregulate MHCs by suppressing IFN signalling. Accordingly, CSCs are less responsive to IFN and can escape immune surveillance, which is a critical ability during tumour and metastasis initiation events^{56,57}. However, CSCs may have a defective antiviral interferon-mediated response, which may explain why oncolytic viruses specifically target CSC populations^{58,59}. In fact, the interferon response is frequently defective in multiple cancer types

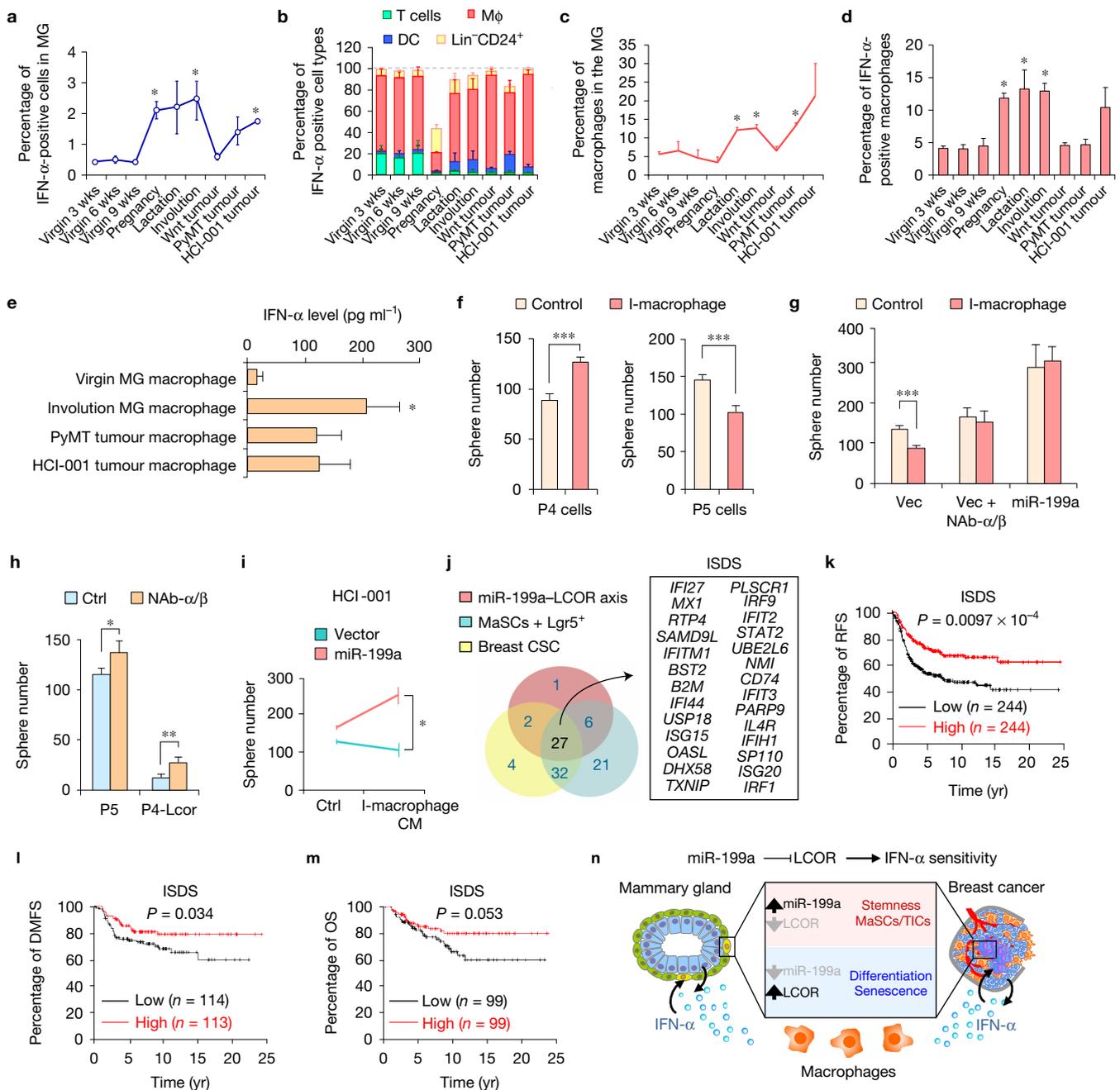


Figure 8 Immune and autocrine IFN-related effects on mammary gland and tumour cells. **(a)** Percentage of total IFN- α -expressing cells from digested mammary glands (MG) and tumours at the indicated stages, analysed by flow cytometry after intracellular IFN- α staining of single-cell suspensions. **(b)** Relative percentage of IFN- α -positive cell types of the mammary gland, analysed by flow cytometry after co-staining intracellular IFN- α with CD3e (T cells), CD11c (dendritic cells), F4/80 (macrophages) and Lin⁻CD24⁺ (epithelial cells). **(c)** Flow cytometry analysis showing the percentage of F4/80 (macrophages)-positive cells in the mammary gland at different stages. **(d)** Flow cytometry analysis of the percentage of IFN- α -positive cells within the total macrophage population. In **a–d** $n=4$ biologically independent samples; data represent mean \pm s.e.m. **(e)** Quantification of IFN- α levels in the CM of the indicated cells, detected by ELISA ($n=3$ biologically independent samples; data represent mean \pm s.e.m.). **(f,g)** Quantification of mammospheres formed by P4 (20,000 cells), P5 (10,000 cells) **(f)** and P5-miR-199a-OE (10,000 cells) cells treated 1:3 with CM from involution macrophages and neutralizing antibodies (NAb) against IFN- α/β (2.5 μ g ml⁻¹) **(g)** ($n=3$ biologically independent samples; data represent mean \pm s.e.m.). **(h)** Quantification of mammospheres formed by 10,000 P5 and P4-Lcor cells with or without treatment with NAb against IFN- α/β (2.5 μ g ml⁻¹) ($n=5$ biologically independent samples; data represent mean \pm s.e.m.). **(i)** Quantification of PDX cell tumourspheres formed by 10,000 HCI-001 treated 1:3 with CM from involution macrophages and the conditions indicated ($n=3$ biologically independent samples; data represent mean \pm s.e.m.). **(j)** Schematic diagram showing the compilation of the 27-gene ISDS. See Methods for details. **(k–m)** Kaplan–Meier relapse-free survival (RFS) **(k)**, distant metastasis-free survival (DMFS) **(l)**, and overall survival (OS) **(m)** analysis of the ISDS gene signature in ER⁻ breast cancer using the KM plotter⁴². **(n)** Schematic model for the conserved function of the miR-199a-LCOR axis in allowing the evasion of normal mammary gland and breast cancer cells from macrophage-derived and autocrine IFN- α . * $P < 0.05$ by Student's t -test in **a**, **c** and **d** with respect to the virgin 9-week condition. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$ by Student's t -test in **e–i**. P value by log-rank tests in **k–m**.

against IFN- α/β (2.5 μ g ml⁻¹) **(g)** ($n=3$ biologically independent samples; data represent mean \pm s.e.m.). **(h)** Quantification of mammospheres formed by 10,000 P5 and P4-Lcor cells with or without treatment with NAb against IFN- α/β (2.5 μ g ml⁻¹) ($n=5$ biologically independent samples; data represent mean \pm s.e.m.). **(i)** Quantification of PDX cell tumourspheres formed by 10,000 HCI-001 treated 1:3 with CM from involution macrophages and the conditions indicated ($n=3$ biologically independent samples; data represent mean \pm s.e.m.). **(j)** Schematic diagram showing the compilation of the 27-gene ISDS. See Methods for details. **(k–m)** Kaplan–Meier relapse-free survival (RFS) **(k)**, distant metastasis-free survival (DMFS) **(l)**, and overall survival (OS) **(m)** analysis of the ISDS gene signature in ER⁻ breast cancer using the KM plotter⁴². **(n)** Schematic model for the conserved function of the miR-199a-LCOR axis in allowing the evasion of normal mammary gland and breast cancer cells from macrophage-derived and autocrine IFN- α . * $P < 0.05$ by Student's t -test in **a**, **c** and **d** with respect to the virgin 9-week condition. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.005$ by Student's t -test in **e–i**. P value by log-rank tests in **k–m**.

by genetic or epigenetic alteration of related genes^{60,61}, suggesting that defective IFN responses are advantageous for tumours. As IFNs have been widely used as adjuvant therapy in multiple cancer types, such treatments may become more effective if the IFN-resistant CSCs can be rendered sensitive by targeting the miR-199a–LCOR axis. □

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the [online version of this paper](#).

Note: Supplementary Information is available in the online version of the paper

ACKNOWLEDGEMENTS

We thank A. Welm for providing the PDX and A. Prat for technical advice for CL subtype classification. This work was supported by a Susan G. Komen Fellowship to T.C.-T. (PDF15332075), and grants from the Brewster Foundation, the Breast Cancer Research Foundation, Department of Defense (BC123187), and the National Institutes of Health (R01CA141062) to Y.K. This research was also supported by the Genomic Editing and Flow Cytometry Shared Resources of the Cancer Institute of New Jersey (P30CA072720).

AUTHOR CONTRIBUTIONS

T.C.-T. and Y.K. designed experiments. T.C.-T., D.D.L., A.C., X.H., Y.W., R.A.-A., R.C., B.I.K. and H.A.S. performed the experiments. J.Z., Y.-Z.J., C.D., J.-J.L. and Z.-M.S. provided crucial samples and technical advice. T.C.-T. and Y.K. wrote the manuscript. All authors discussed the results and commented on the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://dx.doi.org/10.1038/nbc3533>

Reprints and permissions information is available online at www.nature.com/reprints
 Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

- Visvader, J. E. & Stingl, J. Mammary stem cells and the differentiation hierarchy: current status and perspectives. *Genes Dev.* **28**, 1143–1158 (2014).
- Nguyen, L. V., Vanner, R., Dirks, P. & Eaves, C. J. Cancer stem cells: an evolving concept. *Nat. Rev. Cancer* **12**, 133–143 (2012).
- Chakrabarti, R. *et al.* Δ Np63 promotes stem cell activity in mammary gland development and basal-like breast cancer by enhancing Fzd7 expression and Wnt signalling. *Nat. Cell Biol.* **16**, 1004–1015 (2014).
- Guo, W. *et al.* Slug and Sox9 cooperatively determine the mammary stem cell state. *Cell* **148**, 1015–1028 (2012).
- Shimono, Y. *et al.* Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells. *Cell* **138**, 592–603 (2009).
- Ben-Porath, I. *et al.* An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat. Genet.* **40**, 499–507 (2008).
- Prat, A. *et al.* Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res.* **12**, R68 (2010).
- Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. & Clarke, M. F. Prospective identification of tumorigenic breast cancer cells. *Proc. Natl Acad. Sci. USA* **100**, 3983–3988 (2003).
- Liu, S. *et al.* Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. *Stem Cell Rep.* **2**, 78–91 (2014).
- Mani, S. A. *et al.* The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* **133**, 704–715 (2008).
- Gyorki, D. E., Asselin-Labat, M. L., van Rooijen, N., Lindeman, G. J. & Visvader, J. E. Resident macrophages influence stem cell activity in the mammary gland. *Breast Cancer Res.* **11**, R62 (2009).
- Drukker, M. & Benvenisty, N. The immunogenicity of human embryonic stem-derived cells. *Trends Biotechnol.* **22**, 136–141 (2004).
- Cordon-Cardo, C. *et al.* Expression of HLA-A,B,C antigens on primary and metastatic tumor cell populations of human carcinomas. *Cancer Res.* **51**, 6372–6380 (1991).
- Bruttel, V. S. & Wischhusen, J. Cancer stem cell immunology: key to understanding tumorigenesis and tumor immune escape? *Front. Immunol.* **5**, 360 (2014).
- Gangaraju, V. K. & Lin, H. MicroRNAs: key regulators of stem cells. *Nat. Rev. Mol. Cell Biol.* **10**, 116–125 (2009).
- Lujambio, A. & Lowe, S. W. The microcosmos of cancer. *Nature* **482**, 347–355 (2012).
- Pal, B. *et al.* Integration of microRNA signatures of distinct mammary epithelial cell types with their gene expression and epigenetic portraits. *Breast Cancer Res.* **17**, 85 (2015).
- Greene, S. B., Gunaratne, P. H., Hammond, S. M. & Rosen, J. M. A putative role for microRNA-205 in mammary epithelial cell progenitors. *J. Cell Sci.* **123**, 606–618 (2010).
- Ibarra, I., Erlich, Y., Muthuswamy, S. K., Sachidanandam, R. & Hannon, G. J. A role for microRNAs in maintenance of mouse mammary epithelial progenitor cells. *Genes Dev.* **21**, 3238–3243 (2007).
- Llobet-Navas, D. *et al.* The miR-424(322)/503 cluster orchestrates remodeling of the epithelium in the involuting mammary gland. *Genes Dev.* **28**, 765–782 (2014).
- Ucar, A. *et al.* miR-212 and miR-132 are required for epithelial stromal interactions necessary for mouse mammary gland development. *Nat. Genet.* **42**, 1101–1108 (2010).
- Bockmeyer, C. L. *et al.* MicroRNA profiles of healthy basal and luminal mammary epithelial cells are distinct and reflected in different breast cancer subtypes. *Breast Cancer Res. Treat.* **130**, 735–745 (2011).
- Dvinge, H. *et al.* The shaping and functional consequences of the microRNA landscape in breast cancer. *Nature* **497**, 378–382 (2013).
- Zhu, M. *et al.* Integrated miRNA and mRNA expression profiling of mouse mammary tumor models identifies miRNA signatures associated with mammary tumor lineage. *Genome Biol.* **12**, R77 (2011).
- Liu, S., Clouthier, S. G. & Wicha, M. S. Role of microRNAs in the regulation of breast cancer stem cells. *J. Mammary Gland Biol. Neoplasia* **17**, 15–21 (2012).
- Plaks, V. *et al.* Lgr5-expressing cells are sufficient and necessary for postnatal mammary gland organogenesis. *Cell Rep.* **3**, 70–78 (2013).
- Lim, E. *et al.* Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways. *Breast Cancer Res.* **12**, R21 (2010).
- Gupta, P. B. *et al.* Identification of selective inhibitors of cancer stem cells by high-throughput screening. *Cell* **138**, 645–659 (2009).
- Visvader, J. E. Keeping abreast of the mammary epithelial hierarchy and breast tumorigenesis. *Genes Dev.* **23**, 2563–2577 (2009).
- Taube, J. H. *et al.* Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proc. Natl Acad. Sci. USA* **107**, 15449–15454 (2010).
- Lee, Y. B. *et al.* Twist-1 regulates the miR-199a/214 cluster during development. *Nucleic Acids Res.* **37**, 123–128 (2009).
- Beck, B. *et al.* Different levels of Twist1 regulate skin tumor initiation, stemness, and progression. *Cell Stem Cell* **16**, 67–79 (2015).
- Celia-Terrassa, T. *et al.* Epithelial-mesenchymal transition can suppress major attributes of human epithelial tumor-initiating cells. *J. Clin. Invest.* **122**, 1849–1868 (2012).
- Ocana, O. H. *et al.* Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prx1. *Cancer Cell* **22**, 709–724 (2012).
- Cho, R. W. *et al.* Isolation and molecular characterization of cancer stem cells in MMTV-Wnt-1 murine breast tumors. *Stem Cells* **26**, 364–371 (2008).
- DeRose, Y. S. *et al.* Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. *Nat. Med.* **17**, 1514–1520 (2011).
- Buffa, F. M. *et al.* microRNA-associated progression pathways and potential therapeutic targets identified by integrated mRNA and microRNA expression profiling in breast cancer. *Cancer Res.* **71**, 5635–5645 (2011).
- Jiang, Y. Z. *et al.* Transcriptome analysis of triple-negative breast cancer reveals an integrated mRNA-lncRNA signature with predictive and prognostic value. *Cancer Res.* **76**, 2105–2114 (2016).
- Calderon, M. R. *et al.* Ligand-dependent corepressor (LCoR) recruitment by Kruppel-like factor 6 (KLF6) regulates expression of the cyclin-dependent kinase inhibitor CDKN1A gene. *J. Biol. Chem.* **287**, 8662–8674 (2012).
- Fernandes, I. *et al.* Ligand-dependent nuclear receptor corepressor LCoR functions by histone deacetylase-dependent and -independent mechanisms. *Mol. Cell* **11**, 139–150 (2003).
- Lim, E. *et al.* Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat. Med.* **15**, 907–913 (2009).
- Gyorffy, B. *et al.* An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res. Treat.* **123**, 725–731 (2010).
- Liu, R. *et al.* miR-199a-3p targets stemness-related and mitogenic signaling pathways to suppress the expansion and tumorigenic capabilities of prostate cancer stem cells. *Oncotarget* **7**, 56628–56642 (2016).
- Yin, G. *et al.* TWISTing stemness, inflammation and proliferation of epithelial ovarian cancer cells through MIR199A2/214. *Oncogene* **29**, 3545–3553 (2010).
- Alemdehy, M. F. *et al.* ICL-induced miR139-3p and miR199a-3p have opposite roles in hematopoietic cell expansion and leukemic transformation. *Blood* **125**, 3937–3948 (2015).
- Cuiffo, B. G. *et al.* MSC-regulated microRNAs converge on the transcription factor FOXP2 and promote breast cancer metastasis. *Cell Stem Cell* **15**, 762–774 (2014).
- Chen, J. *et al.* miR-199a-5p confers tumor-suppressive role in triple-negative breast cancer. *BMC Cancer* **16**, 887 (2016).
- Asim, M. *et al.* Ligand-dependent corepressor acts as a novel androgen receptor corepressor, inhibits prostate cancer growth, and is functionally inactivated by the Src protein kinase. *J. Biol. Chem.* **286**, 37108–37117 (2011).
- Bidwell, B. N. *et al.* Silencing of Irf7 pathways in breast cancer cells promotes bone metastasis through immune escape. *Nat. Med.* **18**, 1224–1231 (2012).

50. Zitvogel, L., Galluzzi, L., Kepp, O., Smyth, M. J. & Kroemer, G. Type I interferons in anticancer immunity. *Nat. Rev. Immunol.* **15**, 405–414 (2015).
51. McNab, F., Mayer-Barber, K., Sher, A., Wack, A. & O'Garra, A. Type I interferons in infectious disease. *Nat. Rev. Immunol.* **15**, 87–103 (2015).
52. Essers, M. A. *et al.* IFN α activates dormant haematopoietic stem cells *in vivo*. *Nature* **458**, 904–908 (2009).
53. Sato, T. *et al.* Interferon regulatory factor-2 protects quiescent hematopoietic stem cells from type I interferon-dependent exhaustion. *Nat. Med.* **15**, 696–700 (2009).
54. Schedin, P. Pregnancy-associated breast cancer and metastasis. *Nat. Rev. Cancer* **6**, 281–291 (2006).
55. Lehmann, B. D. *et al.* Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J. Clin. Invest.* **121**, 2750–2767 (2011).
56. Celia-Terrassa, T. & Kang, Y. Distinctive properties of metastasis-initiating cells. *Genes Dev.* **30**, 892–908 (2016).
57. Malladi, S. *et al.* Metastatic latency and immune evasion through autocrine inhibition of WNT. *Cell* **165**, 45–60 (2016).
58. Schatton, T., Frank, N. Y. & Frank, M. H. Identification and targeting of cancer stem cells. *Bioessays* **31**, 1038–1049 (2009).
59. Eriksson, M. *et al.* Oncolytic adenoviruses kill breast cancer initiating CD44⁺CD24^{-low} cells. *Mol. Ther.* **15**, 2088–2093 (2007).
60. James, C. D. *et al.* Chromosome 9 deletion mapping reveals interferon alpha and interferon beta-1 gene deletions in human glial tumors. *Cancer Res.* **51**, 1684–1688 (1991).
61. Lu, R., Au, W. C., Yeow, W. S., Hageman, N. & Pitha, P. M. Regulation of the promoter activity of interferon regulatory factor-7 gene. Activation by interferon and silencing by hypermethylation. *J. Biol. Chem.* **275**, 31805–31812 (2000).

METHODS

Animal studies. Animal procedures were approved by the Institutional Animal Care and Use Committee (IACUC) of Princeton University. For mammary stem cell isolation, 8- to 9-week-old female FVB virgin mice were used. The Lgr5-EGFP-IRES-creERT2 (C57/B6 background) model was used for MaSCs Lgr5⁺ isolation (Lin⁻CD24⁺CD29^{high}GFP⁺) as previously described²⁶. For cleared fat pad injections, 3- to 4-week-old female FVB mice were anaesthetized and minimal incisions were made to expose the mammary gland. Randomization among litters was performed before the injection time, and animals were similar age and female sex (applies to the rest of the experiments using mice). No statistical method was used to predetermine the sample size (applies to the rest of the experiments using mice). Following the standard protocol, the mammary gland clearing was done above the central lymph node in inguinal gland no. 4, and different quantities of cells (as indicated in the figures or figure legends) were injected into the cleared fat pad as previously described⁶². The investigators were not blinded to allocation during experiments and outcome assessment. For the IFN- α treatment experiment, 100,000 U of recombinant mouse IFN- α 2 was subcutaneously administered 48 h after surgery, three times a week for three weeks. A MMTV-Wnt1 mouse model of spontaneous breast cancer was used to isolate mouse tumour-initiating cells. MMTV-PyMT and MMTV-Wnt1 mouse models of spontaneous breast cancer were used to measure the IFN- α levels in various immune cell populations in tumours and to isolate macrophages. For orthotopic mammary tumour experiments of human patient-derived xenografts (PDX; kindly provided by A. Welm, Huntsman Cancer Institute, University of Utah, Salt Lake City, Utah, USA) including HCI-001, HCI-002, HCI-003, HCI-005, HCI-009 and HCI-010, immunocompromised NOD Scid Gamma (NSG) mice were used. Note that NSG mice lack mature T cells, B cells and NK cells, but they have dendritic cells and macrophages, although with reduced activity^{63,64}. We followed the standard protocol for PDX transplantation, maintenance and digestion of the tumours³⁶. The lentiviral transduction and orthotopic injections of PDX single-cell suspension were optimized for primary tumour initiation experiments. Six to ten mice or glands were used for each experimental group and the primary tumours were monitored weekly by palpation. For the IFN- α treatment experiment, subcutaneous injection of 100,000 U of recombinant human IFN- α 2A was administered three times a week for 50 days, starting at day 30 after tumour inoculation (prior to the formation of palpable tumours) and ended at the endpoint of the experiment (Fig. 6m). Tumour monitoring and measurement were performed by trained technicians in a blinded fashion. Tumours were measured by callipers for calculation of tumour volumes ($\pi \times \text{length} \times \text{width}^2 / 6$). For systemic metastasis experiments using 4T07 and the MDA-MB-231 cell line, intracardiac injection of 100,000 cells in the left ventricle was performed in anaesthetized female athymic Ncr-nu/nu mice. Development of metastases was monitored by blinded investigators and measuring the photon flux of metastatic lesions based on bioluminescence imaging as previously described⁶⁵, and nodule counts were obtained after dissection of the different organs (investigators were not blinded to outcome assessment).

Cell lines, culture conditions and treatments. All cell lines used in the study, including mammary epithelial cell lines (human HMLE, MCF10A and mouse NMuMG cell lines), breast cancer cell lines (human MDA-MB-231, HMLE-Neu, T47D, MCF7, BT474, and mouse 4T07 cell lines) and other cell lines (HEK293T, H29 and HeLa), were cultured using the standard conditions according to the American Type Culture Collection (ATCC) instructions. HMLE and HMLE-Neu cells were obtained from R. Weinberg at MIT, USA. iMMEC cells were obtained from V. Karantzis at CINJ, USA. Primary isolated mammary epithelial cells (MECs) were cultured with MEGM (Lonza) and immortalized murine MECs (iMMECs) were cultured as described previously⁶⁶. No cell lines used in this study were found in the database of commonly misidentified cell lines that is maintained by ICLAC and NCBI Biosample. The cell lines were not authenticated. Mycoplasma contamination was routinely checked for (monthly) in the laboratory by PCR analysis; all cell lines used in the study were confirmed to be mycoplasma negative. TGF- β 1 (R&D systems) and Z-4-hydroxytamoxifen (4-OHT) (Sigma-Aldrich) cell culture treatments *in vitro* were performed for 12 days, at 100 pM and 20 nM, respectively.

Limiting dilution assays. For mammary gland reconstitution assays, we prepared single-cell suspensions of MECs and sorted for P4/P5 cells by flow cytometry. Lentivirally transduced cells were injected in serial dilution numbers into cleared mammary fat pads. Cells were injected in 50% Matrigel and outgrowth was analysed after 6 weeks. ELDA (extreme limiting dilution analysis) software was used to calculate the frequency of MaSCs with 95% of confidence. The same quantitative method was used in limiting dilution mammary fat pad injections of PDX cells to calculate TIC frequency. The PDX transduced single-cell suspensions were also injected in 50% Matrigel into the mammary fat pad of NSG mice. Tumours were considered established when they became palpable for 2 consecutive weeks.

Viral production and infection of cell lines and primary cells. Lentiviral plasmids were transfected into HEK293T cells together with the envelope plasmid (VSVG) and gag-pol plasmid (pCMV-dR8.91) following the standard lentiviral packaging protocol to generate lentiviruses. For retrovirus, the pWZL-ER-Blast retroviral vectors were transfected into the H29 packaging cell line and viruses were collected 48 and 72 h after transfection. Primary cells were spin infected in conical tubes for 2 h at 1,000g at 4 °C with concentrated viruses in media containing 8 $\mu\text{g ml}^{-1}$ Polybrene. After spin infection, cells were counted and used for *in vitro* or *in vivo* experiments. For established cell lines, cells were transduced in culture and selected with the corresponding antibiotic resistance.

Immune cell isolation. F4/80⁺ macrophages were obtained by FACS from 3 day involuting mammary glands of 9–10-week FVB mice. Primary cells were cultured in mammosphere media (MSM) for 48 h to generate the CM, which was then used to treat mammospheres.

Senescence assays. Cell lines were cultured and senescence was evaluated after 2–3 days of IFN- α treatment (1,000 U ml⁻¹ human IFN- α 2A) using the β -Galactosidase staining kit (Cell Signaling Technology) following the manufacturer's instructions. Images were taken with a 10 \times objective in bright field.

Molecular cloning and plasmids. Multiple miRNA (mouse miR199a-2, miR-204, miR-211, miR-1a, miR-133a, miR-133b, miR-23b) and gene (mouse *Lcor* and human *LCOR-HA*, *LCOR-LSKAA-HA* and *LCOR- Δ HTH-HA*) expression constructs were generated using the pLEX-MCS-Puro lentiviral vector. cDNA was introduced into pLEX using the SpeI and AgeI cutting sites. For *LCOR-HA*, *LCOR* cDNA was first introduced into pRVPTO-HA using EcoRI and NotI. The resultant *LCOR-HA* was then subcloned into the pLEX-MCS using SpeI and AgeI. The *LCOR* mutants (*LCOR-LSKAA-HA* and *LCOR-HTH-HA*) were generated from the pLEX-*LCOR-HA* construct by two separate PCR amplifications with the mutation region joining the two PCR fragments. The 5' fragment was amplified using a forward primer paired with specifically designed mutation-containing reverse primers, which end on the intended mutation region or right before it. The 3' fragment was amplified using the mutation-containing forward primers paired with a reverse primer. To enable blunt end ligation and cloning, the mutant-harboring primers were phosphorylated using T4 kinase before the PCR reactions. The amplified 5' and 3' fragments were digested with SpeI and AgeI respectively and ligated into the pLEX-MCS. Retroviral tamoxifen-inducible overexpression plasmids pWZL-ER-Blast and pWZL-Twist1-ER-Blast were obtained from Addgene. For miR-199a knockdown, miRZIP technology from System Biosciences was used to generate pGreenPuro (scrambled hairpin control), miR-ZIP-199a-5p and miR-ZIP-199a-3p. For gene knockdown studies, shRNA lentiviral vectors were purchased from Sigma-Aldrich for targeting mouse genes *Lcor* (TRCN0000085107), *Tox3* (TRCN0000413123) and *Rbm47* (TRCN0000123514), and human *LCOR* (TRCN0000016306 and TRCN0000436034). The pMIR-REPORT vector (Ambion) was used to generate luciferase reporters for miRNA targeting activity. Wild-type and mutant mouse *Lcor* 3'UTRs were cloned into pMIR-REPORT using SpeI and HindIII sites. The 3'UTR-*Lcor* sequence was divided into UP and DOWN segments, as the full-length 3'UTR exceeded 8 kb and could not be cloned in its entirety. Single and double mutants were generated for the five different predicted seed sequences. Mutated 3'UTR seed sequences were generated using the QuikChange multi site-directed mutagenesis kit (Stratagene). All primer sequences used for cloning are listed in Supplementary Table 5 (restriction sites are in green and mutated nucleotides are in red and bold).

Flow cytometry analysis and cell sorting. For mammary gland lineage cell isolation, we followed the standard protocol used previously⁶². Briefly, mammary glands from 8–9-week-old female mice were digested to form single-cell suspensions of primary MECs. These cells were sorted using Lin⁻CD24⁺CD29^{high} markers to obtain P4 (MaSCs/basal) cells, Lin⁻CD24⁺CD29^{low} markers to obtain P5 (luminal) cells, and total Lin⁻ for P3 (total MECs). We used CD61⁺ to isolate luminal progenitors and CD61⁻ for mature luminal cells within the P5 population. For Lgr5⁺ MaSC isolation, the Lgr5-EGFP-IRES-creERT2 mice were used and Lgr5⁺ MaSCs were isolated by FACS using the Lin⁻CD24⁺CD29^{high}GFP⁺ marker combination. F4/80⁺ primary macrophages were isolated by FACS from single-cell suspensions of digested mammary glands from FVB mice. Mouse TICs were obtained from MMTV-Wnt1 tumours after sorting for the CD45⁻CD24⁺Thy1⁺ population. Digestion of PDX tumours and preparation of single-cell suspensions were performed using the standard protocol³⁶ and human primary TICs were isolated by sorting the Lin⁻CD24⁻CD44⁺ population.

Intracellular IFN- α flow cytometry analysis. Mammary glands and tumours were digested and brought to single-cell suspension as described above and with the presence of GolgiPlug as indicated by the manufacturer (BD Biosciences). After

fixation with 2% paraformaldehyde 30 min on ice, cells were permeabilized and blocked with 0.2% saponin, 5 mM EDTA, 2 mM NaN_3 , 5% NGS and $4 \mu\text{g ml}^{-1}$ anti-Fc γ receptor (Clone 2.4 G2; BD Biosciences). IFN- α was stained using FITC-conjugated anti-mouse IFN- α (Clone RMMA-1; PBL)⁶⁷. T cells were stained with PE anti-mouse CD3 ϵ (Clone 145-2C11; BioLegend), macrophages with APC/Cy7 anti-mouse F4/80 (Clone BM8; BioLegend), and dendritic cells with APC anti-mouse CD11c (Clone HL3; BD Pharmingen). FITC-conjugated rat IgG1 was used as a negative isotype control (Clone RTK2071; BioLegend). Note, pregnancy, lactation and involution cell samples display small levels of autofluorescence for reasons unknown. We thoroughly excluded this population from our analysis and applied the same gating for all the samples.

Histological analysis, immunohistochemistry (IHC), immunofluorescence (IF) and *in situ* hybridization (ISH). Histology, IHC and IF analysis of mouse mammary and tumour tissue samples was done as previously described⁶⁸ using the antibody and dilution ratios listed in Supplementary Table 6.

For IHC analysis of clinical specimens, paraffin slides of $4 \mu\text{m}$ thickness or tissue microarrays were baked overnight at 60°C . Tissue slides were washed with PBS after deparaffinization and hydration and then boiled in citrate buffer at 100°C for 40 min. After treatment with 3% H_2O_2 for 30 min to block endogenous peroxidase, slides were incubated at 4°C overnight with rabbit anti-human LCOR antibody (Sigma). Following washes with PBS, slides were then incubated with HRP-conjugated goat anti-rabbit secondary antibody (Genetech) for 30 min at room temperature. Sections were stained by DAB and then counterstained with Gill haematoxylin.

IF analysis of cell culture was performed in HMLE cells using anti-HA to determine the localization of ectopically expressed LCOR-HA. Sterile coverslips placed at the bottom of 24-well plates were seeded, washed with PBS, and fixed for 1 h with methanol at -20°C . After fixation, samples were washed with acetone, then 5 times with PBS, blocked for 30 min with blocking buffer (5% normal goat serum, 0.5% Triton X-100 in PBS) and incubated with anti-HA for 2 h at room temperature. This was followed by PBS washes, and 1 h incubation with secondary antibody conjugated with Alexa Fluor 488. Images were taken using a Nikon A1 confocal microscope and Zeiss fluorescence microscope. For mammosphere staining, spheres were collected with a 5 min spin at $450g$, fixed with PFA 4% for 15 min at 4°C and washed with PBS, followed by centrifuge colony precipitation. Blocking was performed using PBS-Tween 0.3% and M.O.M. kit blocking reagent (Vector Laboratories), followed by 1 h co-staining with anti-Keratin-14 and anti-Keratin-8, and then by 1 h incubation with the respective species-specific secondary antibodies.

For mammary gland ISH experiments, we used the miRCURY LNA microRNA ISH Optimization Kit from Exiqon. LNA probes were double DIG labelled to specifically detect miR-199-5p and U6 snRNA as a positive control (Exiqon). The manufacturer's protocol was strictly followed to perform the ISH in mouse and human samples. For clinical breast cancer ISH, paraffin slides ($4\text{-}\mu\text{m}$ thick) of paraformaldehyde-fixed tissues were baked overnight at 60°C , deparaffinized and hydrated, and then washed with phosphate-buffered saline (PBS). To block endogenous peroxidase activity, slides were treated with 3% H_2O_2 for 10 min at room temperature. After pepsin digestion for 30 min at 37°C , slides were incubated with pre-hybridization buffer (all reagents from sensitivity enhanced *in situ* hybridization kits, Boster) for 3 h at 37°C . Slides were hybridized with double digoxigenin-labelled probes (30 nM for miR-199a, 30 nM for the internal control U6, Exiqon) in hybridization buffer at 65°C overnight, and then washed sequentially with $2\times$ SSC, $1\times$ SSC and $0.2\times$ SSC buffers. Slides were incubated sequentially with the following reagents: blocking buffer, biotinylated digoxin, streptavidin-biotin complex, and peroxidase. Sections were stained by 3,3'-diaminobenzidine (DAB) and then counterstained with Gill haematoxylin.

Mammosphere and tumoursphere assays. For mammosphere and tumoursphere assays, single cells were plated in ultralow-attachment plates (Corning) with the standard mammosphere media⁶⁹. The number of cells plated is indicated for each specific experiment in the figures. The mammospheres were counted 5–12 days later depending on the experiment, which is indicated in figure legends. Tumourspheres were counted after 5–8 days. For multiple-generation sphere formation experiments, colonies were collected by 2 min centrifugation at $200g$ and dissociated with trypsin for 5 min at 37°C . Single dissociated cells were then centrifuged and counted to allow seeding of equal numbers of cells for the next round of sphere formation assays.

Interferon treatment ($1,000 \text{ U ml}^{-1}$) was started 24 h after cell seeding. Recombinant mouse IFN- $\alpha 2$ and IFN- γ were purchased from Novoprotein. Neutralizing antibodies against mouse IFN- α (clone RMMA-1, PBL) and IFN- β (clone RMMB-1, PBL) were used at $2.5 \mu\text{g ml}^{-1}$. Conditioned medium (CM) was generated by culturing primary immune cells in mammosphere medium (MSM) for 48 h, using 1 ml of MSM for every 600,000 primary cells seeded.

Tissue microarrays and other tumour samples. A total of 200 stage I to III primary breast cancer samples from females with invasive ductal carcinoma were randomly collected at the Fudan University Shanghai Cancer Center (FUSCC) between March 2003 and January 2008. Tissue microarrays were constructed using paraffin-embedded blocks of these samples, consisting of duplicate cores from different areas of the same tumour to compare staining patterns. We also included nine cases of primary metaplastic carcinoma diagnosed in the same time period. Paraffin-embedded sections were used for *in situ* hybridization and immunohistochemical staining. We used the staining index to interpret the staining of miR-199a and LCOR. Briefly, the staining score was determined by three independent researchers blinded to the tumour information. Each sample was scored as weak (1) or strong (2) according to staining intensities and the average of the resultant scores was computed. Our study was approved by the independent ethics committee/institutional review board of FUSCC (Shanghai Cancer Center Ethics Committee). All patients gave their written informed consent before inclusion.

Murine IFN- α ELISA. Specific cell populations (500,000 cells) were plated for 48 h with subsequent collection of conditioned media for cytokine quantification. IFN- α levels were quantified from the CM for all cell cultures using the mouse IFN- α Platinum ELISA kit (Fisher/eBioscience) following the manufacturer's instructions.

Luciferase reporter assays. Wild-type and mutant pMIR-LCOR-3'UTR reporters were transfected into HeLa cells together with the *Renilla*-luciferase control plasmid (Ambion). Reporter plasmid (200 ng) was co-transfected with the *Renilla*-luciferase control plasmid and 10 pM of miRNA mimics (Applied Biosystems; Life Technologies). Lipofectamine 2000 was used as the transfection reagent. Cells were lysed 24 h after transfection and analysed for luciferase activity using the Glomax 96 Luminometer (Promega).

qRT-PCR analyses. Total mRNA and miRNAs were isolated using the mirVana miRNA isolation Kit (Ambion). mRNA reverse transcription was done using Superscript III kit (Invitrogen) and real-time quantitative PCR performed using the Power SYBR green PCR master mix (Applied Biosystems). miRNAs were reverse transcribed using the TaqMan Reverse Transcription Kit (Applied Biosystems) and followed by real-time qPCR using TaqMan miRNA assays (Applied Biosystems). All analyses were performed using an ABI 7900HT PCR machine. mRNA expression was normalized by the expression of *GAPDH*, and miRNA expression by RNU6B in each sample. qRT-PCR primers used are listed in Supplementary Table 5.

Immunoprecipitation and western blot analysis. For immunoprecipitation (IP) experiments, cells were lysed in IP lysis buffer (20 mM Tris pH 7.4, 0.15 M NaCl, 1 mM EDTA, 1 mM EGTA, 1% Tx-100) with complete protease inhibitor cocktail (Roche). Cell lysates were incubated on ice for 20 min, centrifuged, and incubated with $5 \mu\text{g}$ of anti-HA (Abcam, Ab9110) at 4°C overnight. Protein G magnetic beads (Life Technologies) were pre-cleared and $20 \mu\text{l}$ were added to the sample for 2 h at 4°C . Beads were washed 5 times and boiled with SDS Laemmli buffer to elute bound protein for western blotting (WB). For WB analysis of cultured cells, proteins were extracted using RIPA buffer and SDS Laemmli buffer. Gel preparation and immunoblotting were performed according to standard protocols. Antibodies and dilutions for WB are listed in Supplementary Table 6.

Microarray analysis. The P4 and P5 subpopulations of mammary epithelial cells (MECs) were isolated from the mammary glands (4–5 mammary glands from each group) of 8–9-week-old virgin FVB or Lgr5-EGFP-IRES-creERT2 mice. MECs were isolated using FACS as previously described⁶². Total RNA was prepared from MECs and various human cell lines including HMLE-miR-199a, -LCOR variants and -LCOR-KDs or MDA-MB-231 LCOR-KD cells, using the mirVana kit as described above. P4 and P5 with IFN- α treatment ($1,000 \text{ U ml}^{-1}$) were maintained in mammosphere conditions during 8 days of mammosphere culture. The expression of miRNA in MEC subpopulations was determined using Agilent mouse miRNA_V19 array (G4872A). The RNA samples were labelled and hybridized using a miRNA complete labelling and hybridization kit (Agilent, 5190-0456). The expression of mRNA in MECs was determined with Agilent Mouse GEV1 8x60K Microarray (G4852A). The mRNA expression in human cell lines was determined with Agilent human GEV2 8x60K microarray (G4851B). The mRNA microarray analyses were performed using a two-colour system. Briefly, the RNA samples and universal mouse reference RNA (Agilent 740100) were labelled with CTP-cy5 and CTP-cy3, respectively, using the Agilent Quick Amp Labeling Kit. Labelled testing and reference RNA samples were mixed in equal proportions, and hybridized to the arrays as described above. After hybridization, the miRNA and mRNA arrays were scanned with an Agilent G2565BA scanner and raw data were extracted using Agilent Feature Extraction software (v10.7). Data were analysed using the GeneSpring GX software (Agilent). In brief, for one-colour miRNA array data, the intensity values of multiple probes for the same miRNA are first summed up and

then \log_2 transformed. The values were then normalized to the 90th percentile of each sample and finally baseline transformed with the median across all samples. For two-colour mRNA array data, for each probe, the $\log_2(\text{Cy5}/\text{Cy3})$ ratio is computed and used as the expression value.

Gene set enrichment analysis (GSEA). Normalized microarray \log_2 ratio expression data were rank-ordered by differential expression between cell populations using a fold change metric. Multiple probes for the same gene were collapsed into one value by the highest probe reading when there were fewer than 3 probe matches, and median when there were 3 or more probe matches. Interrogated signatures include HALLMARK gene sets from the MSigDB database v5.1 release; MaSCs (GSE19446) with 489_UP and 428_DOWN genes qualifying for >1.5-FC (fold change) and FDR < 0.05; and Senescence data set (M9143) with 77_UP genes. Other interrogated data sets include the CSC data set (GSE17215) with 25_UP and 14_DOWN genes qualifying for >3-FC; Undifferentiated tumour cell data set (GSE18229) with 558_UP and 490_DOWN and claudin-low data set (GSE18229) with 437_UP and 370_DOWN genes qualifying for FDR < 0.05. The NOS_TFs gene set with 37 genes was derived from a published gene list⁶, and the MaSCs 230_UP genes and Luminal 230_UP genes were derived from the current study (GSE85808). In addition, the published breast cancer stem cell expression data set⁹ was extracted from Gene Expression Omnibus (GEO) GSE52262 and analysed using GeneSpring GX software. Using this data set, we generated an ER⁻ BCSC data set by a median compilation of the 4 different ER⁻ breast cancer cell lines and xenografts (HCC1954, MC1, SUM149 and SUM159) isolated by CD24⁻/CD44⁺ (ref. 9). Gene signatures were tested using default enrichment GSEA statistics and compared with enrichment results from 1,000 random permutations to obtain the *P* value, *q* value and normalized enrichment score (NES).

The IFN-Stem Cell-Down signature (ISDS) was generated to represent genes that are regulated by the miR-199a–LCOR axis and also contribute to negative enrichment of the Interferon- α response gene set (M5911) in normal and cancerous stem cells. We first identify three subsets of genes from the M5911 Interferon- α response gene set that represent: common genes downregulated by miR-199a-OE and LCOR-KD, and upregulated by LCOR-OE in HMLE cells; genes downregulated in MaSCs or in Lgr5⁺-MaSCs; and genes downregulated in ER⁻ breast CSCs (HCC1954, MC1, SUM149, and SUM159). The 27-gene ISDS represents the common overlap among these three sets of genes.

Clinical data set analysis. The Cancer Genome Atlas (TCGA) data portal was used to access the TCGA breast cancer expression data. The RNA-Seq by Expectation Maximization (RSEM) expression data were median centred and all samples were standardized to zero-mean and unit variance before subtype classification, as recommended by previous authors to remove platform biases⁷. The TCGA breast cancer population analysed contains 794 ER⁺ (77%) and 233 ER⁻ (23%) samples. Prior to any subtype classification, ER populations were balanced using 233 ER⁻ samples and 233 randomly sampled ER⁺ samples to calculate a median gene expression, and then normalized by subtracting this median gene expression. TCGA miRNA expression data were normalized with the R voom package from the limma library⁷⁰. Comparisons between subtype miRNA expression levels were performed using a Wilcoxon unpaired two-sample test. Intrinsic subtype classification of breast cancer samples was performed using the previously described PAM50 centroid-based classifier. The claudin-low (CL) classifier was constructed according to ref. 7. Briefly, for each sample we calculated the Euclidean distance to the nice-cell CL predictor 'CL' and 'others' centroids. The samples were classified on the basis of their proximity to the nearest centroid.

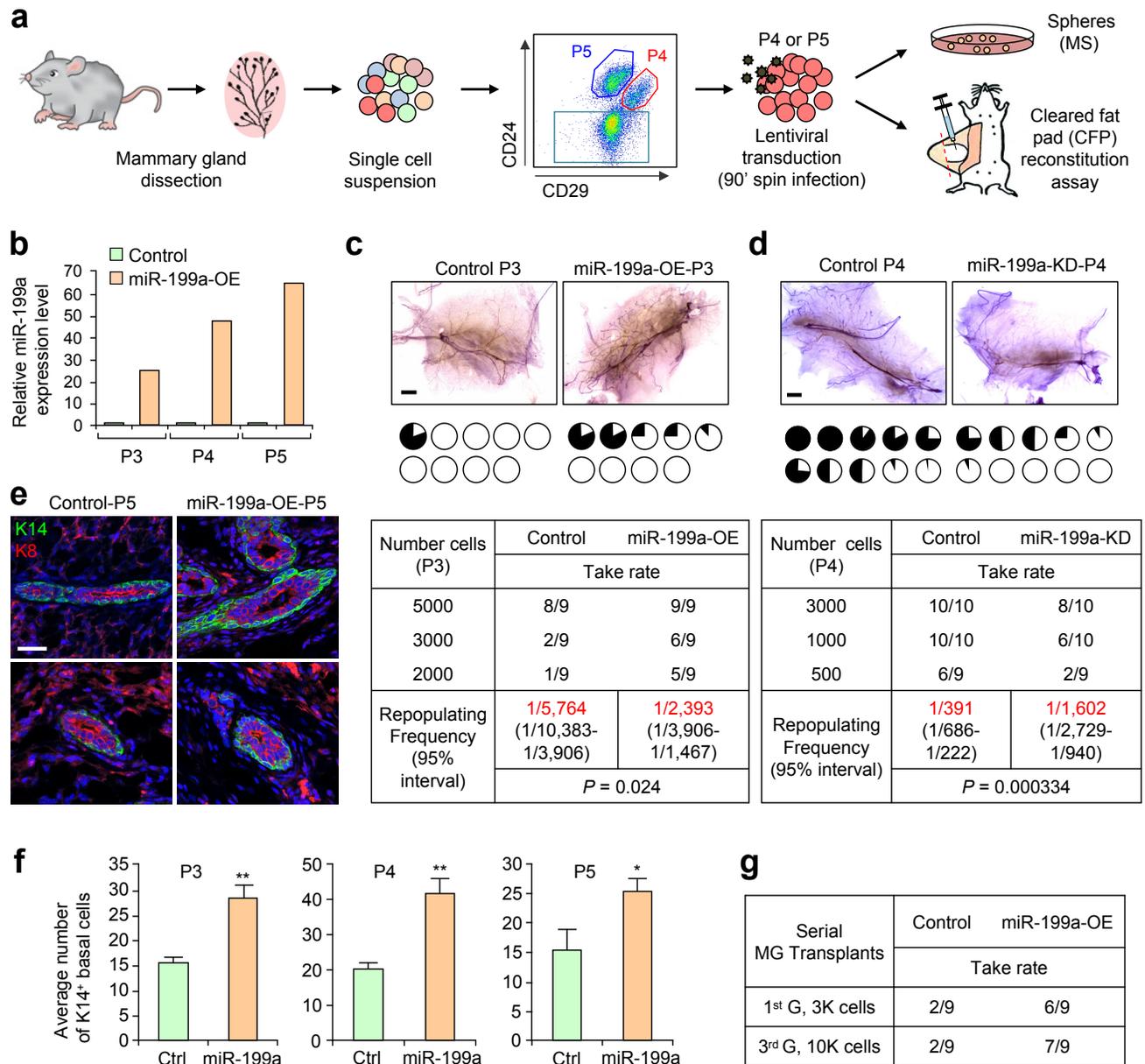
To evaluate the prognosis value of matched miRNA-target samples in breast cancer, the Buffa data set³⁷ was used, with 210 patients (82 ER⁻) and 10-year follow-up. The patients were stratified by median expression of miR-199a and *LCOR*. Other mRNA data sets were also used: the Jiang *et al.* mRNA data set with a total of 168 triple-negative breast cancer (TNBC) and 5 year follow-up, samples were collected at the Fudan University Shanghai Cancer Center (FDUSCC) as previously described³⁸. We used 65 frozen sample from the Jiang data set to extract miRNAs with Trizol and perform qRT-PCR analysis of miR-199a using TaqMan miRNA assays (Applied

Biosystems). The NKI295 data set gene expression was also used to determine distant metastasis-free survival prognosis of a total of 295 patients. Patients were stratified by median expression of *LCOR* (*C10orf12*). To evaluate the prognosis of gene signatures in ER⁻ breast cancer, the KM plotter with a total of 807 ER⁻ patients was used⁴². Patients were stratified by the median score.

Statistics and reproducibility. Results are represented as indicated in the figure legends, generally as mean \pm s.e. (standard error). For experiments with two groups, a small sample size (less than 30), and normally distributed data, the significance was evaluated using a two-tailed unpaired Student's *t*-test under the assumption of unequal variance. Asterisks denote *P*-value significance: **P* < 0.05; ***P* < 0.01; ****P* < 0.005. For multiple independent groups, one-way ANOVA was evaluated. Non-parametric data sets were evaluated using the Mann-Whitney-Wilcoxon *U* test. Stem cell and TIC frequency was calculated with the ELDA software by Pearson's χ^2 test. For free-survival analysis, Kaplan-Meier plots and significance with *P* log-rank test were used. For correlation analysis of clinical samples, χ^2 was used to assess significance. For the clinical multivariate analysis, Cox proportional hazards modelling was used to assess independent prognosis value. All statistics were calculated with the commercial software: GraphPad Prism6 and Microsoft Excel, except for the stem cell and TIC frequency using the online ELDA software. All of the experiments with images (bioluminescence imaging, FACS, IF, IHC and senescence) were repeated >3 times and representative images are shown. If applicable, data corresponding to representative images have been included in Supplementary Table 4.

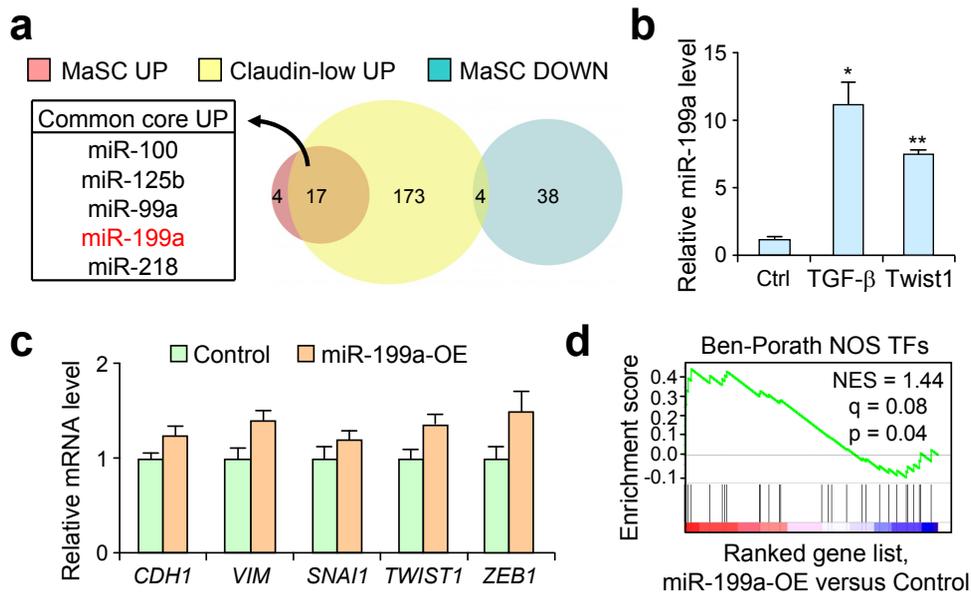
Data availability. All microarray data generated in this study have been deposited as a superseries at the NCBI Gene Expression Omnibus with the accession code GSE85808. Previously published microarray data that were reanalysed or used for GSEA are available under the origin accession codes: GSE76250 (Jiang data set), GSE22220 (Buffa data set), GSE19446, GSE17215, GSE18229 and GSE52262 at the Gene Expression Omnibus. Other gene sets used for GSEA analysis are found in the MSigDB database v5.1 release under the code: M5911, M9143 and the Hallmark gene set collection. Other analysed prognosis sources: KM plotter breast cancer (<http://kmplot.com/analysis>). Previously published RNA-seq data reanalysed are available in the TCGA Genomic Data Commons (<https://gdc-portal.nci.nih.gov/projects/TCGA-BRCA>): TCGA-BRCA (Breast Invasive Carcinoma) containing mRNA and miRNA data. Source data for Supplementary Figs 1b, 7j and 7e have been provided as Supplementary Table 4. All other data supporting the findings of this study are available from the corresponding author on request.

62. Shackleton, M. *et al.* Generation of a functional mammary gland from a single stem cell. *Nature* **439**, 84–88 (2006).
63. Behan, J. W. *et al.* Activation of adipose tissue macrophages in obese mice does not require lymphocytes. *Obesity* **21**, 1380–1388 (2013).
64. Shultz, L. D. *et al.* Multiple defects in innate and adaptive immunologic function in NOD/LtSz-scid mice. *J. Immunol.* **154**, 180–191 (1995).
65. Chakrabarti, R. *et al.* E1f5 inhibits the epithelial-mesenchymal transition in mammary gland development and breast cancer metastasis by transcriptionally repressing Snail2. *Nat. Cell Biol.* **14**, 1212–1222 (2012).
66. Karantza-Wadsworth, V. & White, E. A mouse mammary epithelial cell model to identify molecular mechanisms regulating breast cancer progression. *Methods Enzymol.* **446**, 61–76 (2008).
67. Bonnefoy, F. *et al.* Plasmacytoid dendritic cells play a major role in apoptotic leukocyte-induced immune modulation. *J. Immunol.* **186**, 5696–5705 (2011).
68. Choi, Y. S., Chakrabarti, R., Escamilla-Hernandez, R. & Sinha, S. E1f5 conditional knockout mice reveal its role as a master regulator in mammary alveolar development: failure of Stat5 activation and functional differentiation in the absence of E1f5. *Dev. Biol.* **329**, 227–241 (2009).
69. Dontu, G. *et al.* *In vitro* propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev.* **17**, 1253–1270 (2003).
70. Ritchie, M. E. *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).



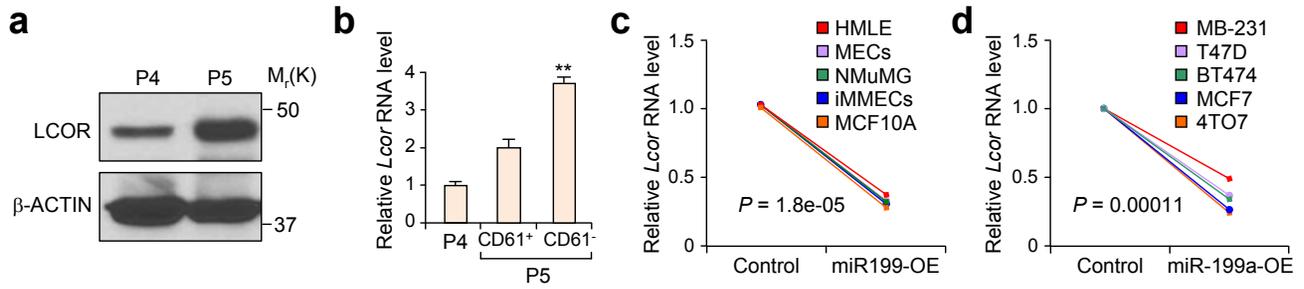
Supplementary Figure 1 miR-199a promotes MaSC activity and basal-like features. **(a)** Mammary gland dissociation of different MEC (P3: Lin⁻ epithelial cells) populations and the lentiviral strategy to conduct gain- and loss-of function experiments. P4 denotes the MaSC-enriched Lin⁻CD24⁺CD29^{high} basal population and P5 the Lin⁻CD24⁺CD29^{low} luminal population. **(b)** qRT-PCR test of the miR-199a ectopic expression after lentiviral transduction of pLEX-miR-199a expression construct in different mammary epithelial cells (MECs). P3: total MECs; P4: basal/MaSC-enrich MECs; and P5: luminal MECs. The test was performed at least in 3 independent experiments for each cell population, and the figure shows one representative test. Source of data in Supplementary Table 4. **(c)** P3 cells were isolated and transduced with the control or miR-199a lentiviruses and subjected to limited dilution cleared fat pad reconstitution assay. **(d)** P4 cells were isolated and transduced with the indicated miR-ZIP constructs, and subjected to limited dilution cleared fat pad reconstitution assay. In **c**

and **d**, Representative images show outgrowth. Each pie chart represents a mammary gland with the blackened area denoting the percentage of mammary gland outgrowth. Tables below represent serial dilution injections with the corresponding take rate. n= number of mammary fat pad injections as indicated in the table. Shown in red are the repopulation frequencies for each condition and P value by Pearson's Chi-squared test, obtained with the ELDA software. **(e)** Keratin-14 (K14-green) and Keratin-8 (K8-red) staining with reconstituted mammary outgrowths from control and miR-199a-OE P5 cells. **(f)** Quantification of the K14⁺ basal cells of mammary outgrowth from control and miR-199a-OE P3, P4 and P5 cells. n=10 ducts and terminal end buds (TEB) sections scored from 3 mammary outgrowths. **(g)** Table representing the mammary gland take-rate after cleared fat pad injection of unsorted MECs overexpressing miR-199a or control vector after 3 generations. Injected cell numbers are indicated. Scale bars: 2 mm in **c** and **d**; 25 μ m in **e**. * $P < 0.05$, ** $P < 0.01$ by two-tailed Student's *t*-test in **f**.



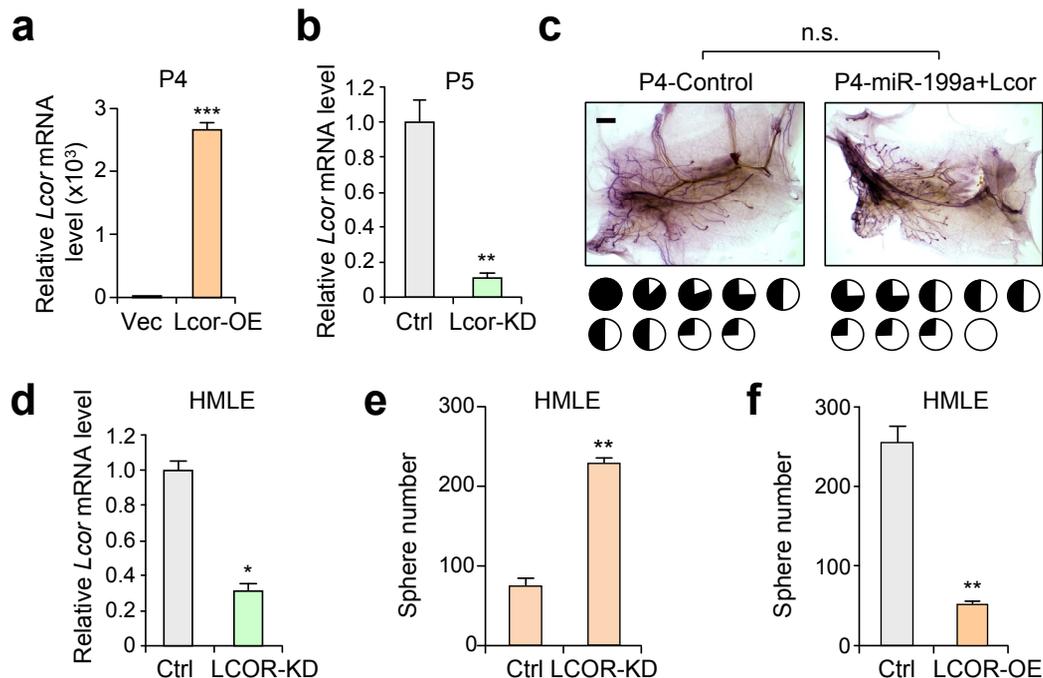
Supplementary Figure 2 miR-199a enhances stem cell-like properties without inducing epithelial-mesenchymal transition (EMT). **(a)** Venn diagram representation of the overlap of up-regulated miRNAs in P4-MaSCs with the up-regulated miRNAs in Claudin-low (CL) tumors based on the TCGA dataset (2015 release). The common core represents the top-5 common miRNAs in the CL subtype. **(b)** qRT-PCR quantification of miR-199a levels in control, TGF- β -treated and

Twist1-ER-OE HMLE cells. **(c)** qRT-PCR analysis of EMT markers and transcription factors in miR-199a-OE and control HMLE cells. **(d)** GSEA of the reported transcriptional factors Nanog-Oct4-Sox2 (NOS-TFs) targets gene set¹ in the ranked gene list of miR-199a-OE HMLE cells vs. control cells. n=3 biologically independent samples; data represents mean \pm SEM in **b** and **c**. * $P < 0.05$, ** $P < 0.01$ by two-tailed Student's t -test in **b** and **c**.



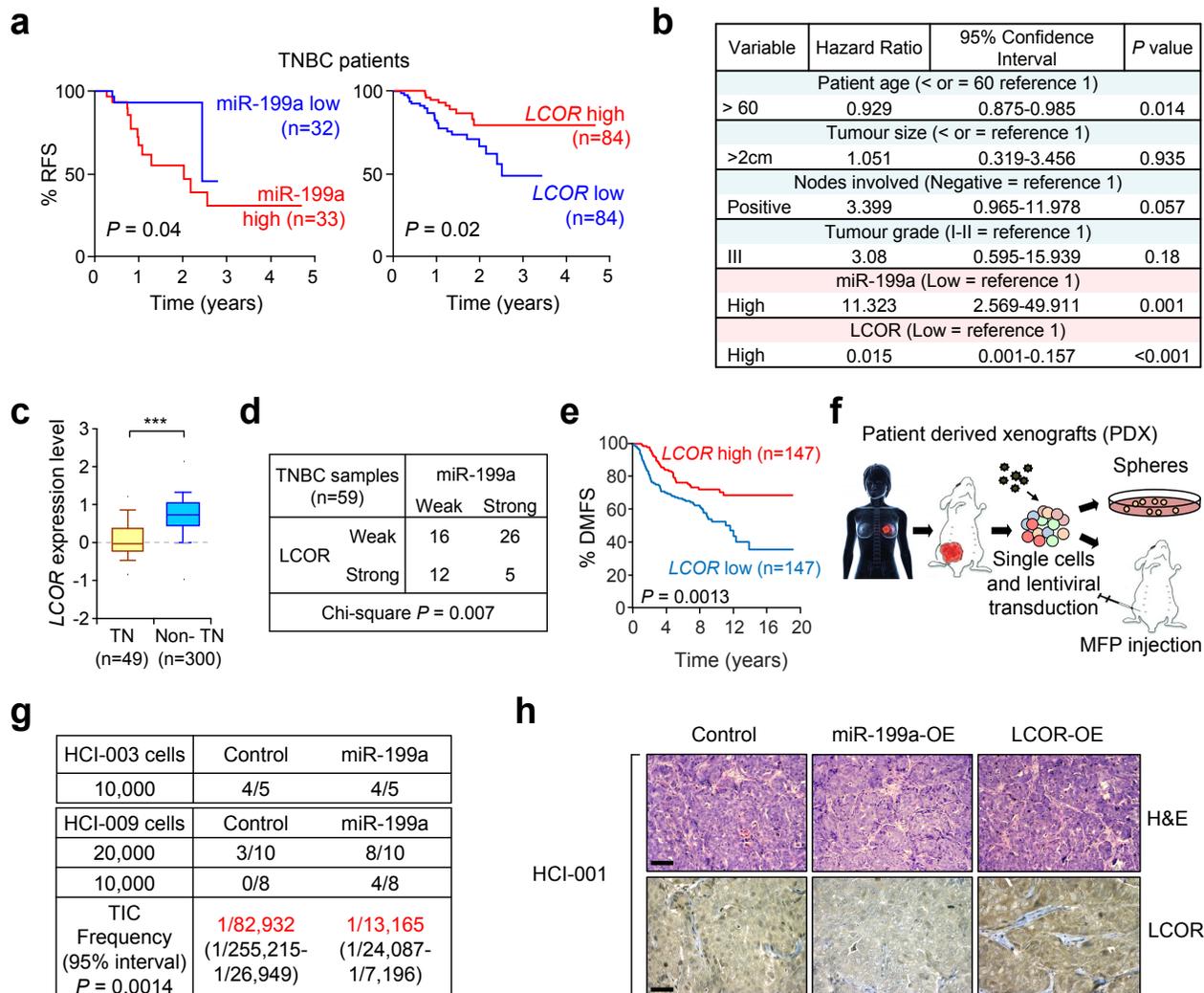
Supplementary Figure 3 Evaluation of LCOR as a miR-199a target. (a) Western blot analysis of LCOR protein level in P4 (MaSCs) and P5 (Luminal) cells. (b) qRT-PCR analysis of *Lcor* in MaSC-enriched P4 cells, luminal progenitor cells (P5-CD61⁺) and luminal mature cells (P5-CD61⁻); n=3 biologically independent samples; data represents

mean ± SEM. (c-d) qRT-PCR analysis of *LCOR* expression upon miR-199a ectopic expression in multiple mammary epithelial cell lines (c) and breast cancer cell lines (d). * $P < 0.05$, ** $P < 0.01$ by two-tailed Student's *t*-test in b and *P* value by two-tailed Student's *t*-test in c, d.



Supplementary Figure 4 LCOR is a potent MaSC suppressor. (a-b) qRT-PCR analysis of *Lcor* mRNA level in P4 transduced with pLEX vector or pLEX-*Lcor* (a) and P5 transduced with *Lcor* shRNA (b). (c) P4 cells were isolated and transduced with the indicated constructs, followed by cleared fat pad injections of 1,000 cells ($n=9$ mouse mammary glands). Representative images and pie charts show the outgrowth and mammary gland filling percentage. Two-tailed Student's *t*-test showed non-significance (n.s.). (d)

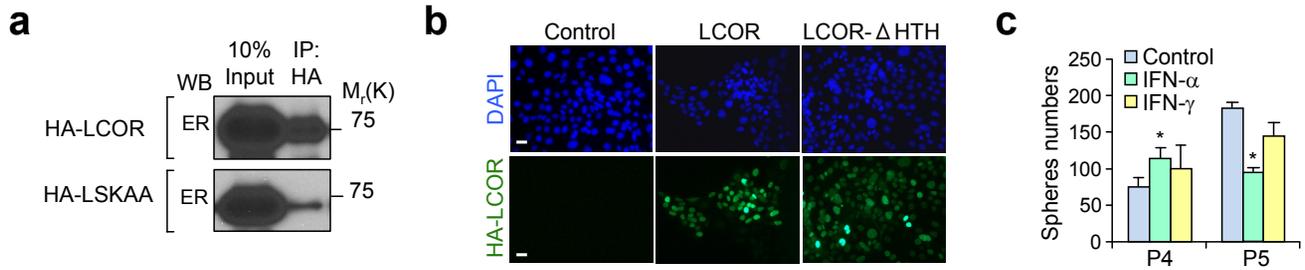
qRT-PCR analysis of *LCOR* expression in control and LCOR-KD HMLE cells. (e) Quantification of mammospheres formed by 5,000 control and LCOR-KD HMLE cells. (f) Quantification of mammosphere formation of 10,000 control or LCOR-OE HMLE cells; counted at 8 days in e and f. Scale bars: 2 mm in c. $n=3$ biologically independent samples; data represents mean \pm SEM and * $P<0.05$, ** $P<0.01$, *** $P<0.005$ by two-tailed Student's *t*-test in a, b, d, e and f.



Supplementary Figure 5 miR-199a and LCOR clinical analysis and *in vivo* functional validation. **(a)** Kaplan-Meier relapse free-survival (RFS) curve of triple-negative breast cancer (TNBC) patients stratified by higher or lower than the median miR-199a ($n = 65$ patients) and *LCOR* expression ($n = 168$ patients) in a previously described TNBC patient cohort². **(b)** Multivariate analysis adjusted for age, tumor size, lymph nodes, tumor grade, miR-199a and *LCOR* expression in the TNBC samples² ($n = 168$ patients). **(c)** OncoPrint analysis of *LCOR* log₂ median centered expression in triple-negative breast cancer (TN; $n = 49$ patients) compared to non-triple-negative breast cancer ($n = 300$ patients) (TCGA dataset). The box represents 75th, 50th and 25th percentile of values, and the whiskers represent maximum and minimum data points. **(d)** Inverse correlation of the expression of miR-199a and *LCOR* in TNBC, as analyzed by ISH (miR-199a) and IHC (*LCOR*). Samples ($n = 59$ tumors) were scored as weak (low expression) or strong

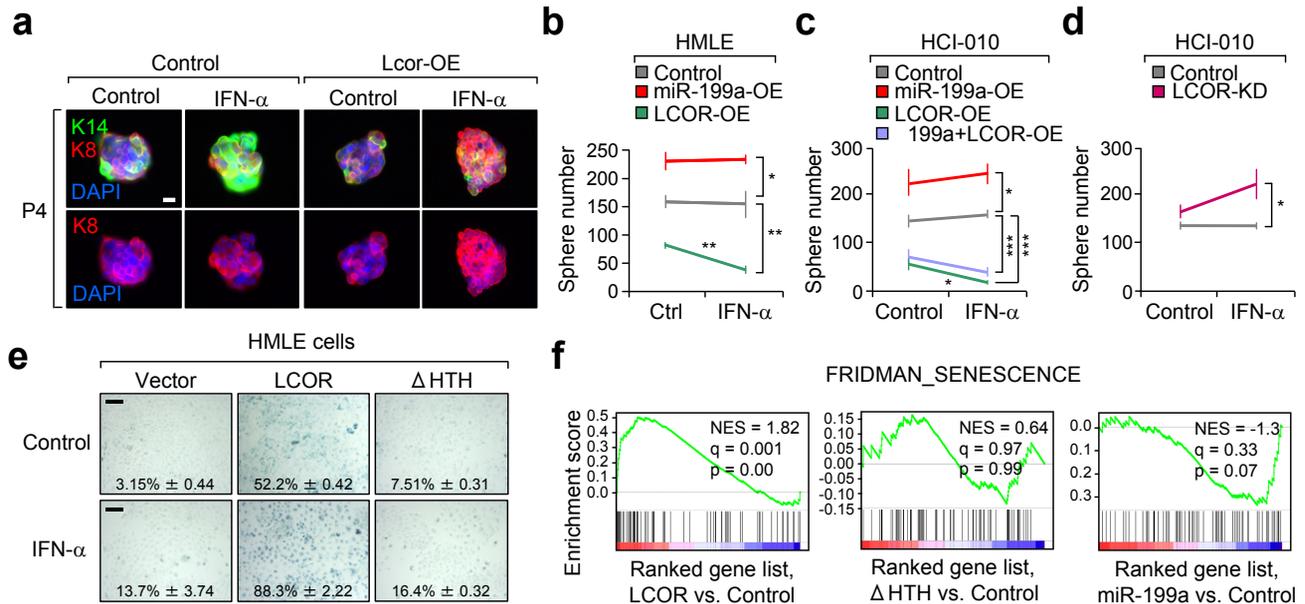
(high expression) according to staining intensities. **(e)** Kaplan-Meier distant metastasis-free survival (DMFS) curve of breast cancer patients stratified by higher or lower than the median *LCOR* expression using the NKI295 dataset³ ($n = 147$ patients). **(f)** Schematic diagram illustrating the procedure of patient-derived xenograft (PDX) maintenance in NSG mice, transduction and functional assays. **(g)** Tumor incidence of HCI-003 (ER⁺PR⁺) and HCI-009 (TNBC) upon mammary fat pad injection of indicated cells. n = number of mammary fat pad injections as indicated in the table. Tumor-initiating cell (TIC) frequency calculated by the ELDA software is indicated in red and P value by Pearson's Chi-squared test. **(h)** Hematoxylin-eosin and *LCOR* IHC analysis of mammary tumors formed by mammary fat pad injection of the indicated PDX cells with or without the overexpression of miR-199a or *LCOR*. Scale bars: 50 μ m in **h**. Log-rank test in **a**, **e**; Cox proportional hazard in **b**, Wilcoxon un-paired test in **c**, and Chi-square test in **d**.

SUPPLEMENTARY INFORMATION



Supplementary Figure 6 LCOR mutant functionality and response to interferons. **(a)** Co-immunoprecipitation of co-transfected estrogen receptor (ER) and wild-type or LSKAA mutant LCOR in HMLE cells. LCOR and LSKAA are HA tagged and were immunoprecipitated with anti-HA antibody and immunoblotted for ER. Lanes correspond to: 10% Input and anti-HA pull-down. **(b)** Immunofluorescence analysis of

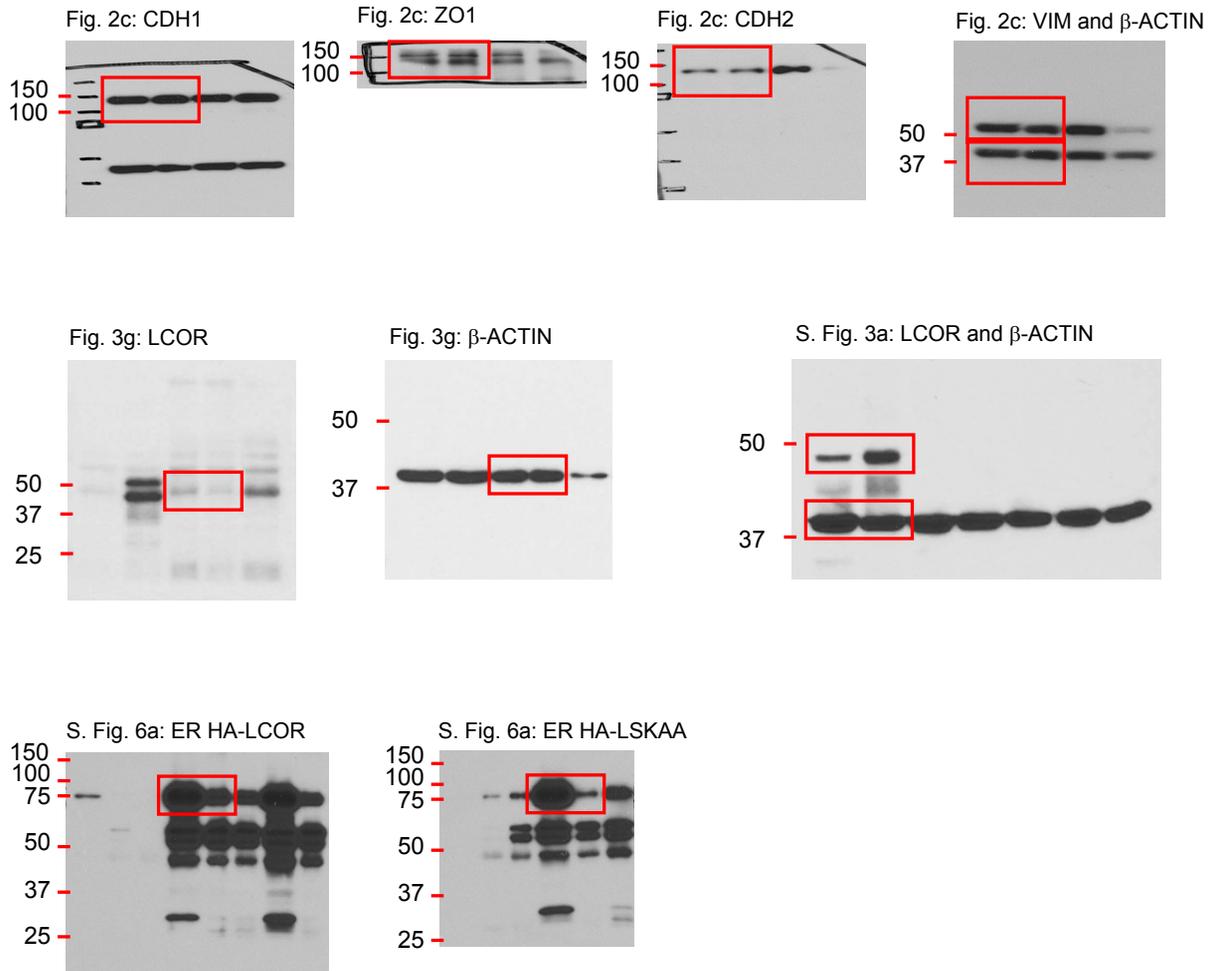
WT and HTH mutant HA-LCOR using anti-HA antibody in HMLE cells. **(c)** Quantification of mammospheres formed by P4 (20,000 cells) and P5 (10,000 cells) mammospheres treated with 1000 U/ml IFN- α or IFN- γ (n=5 biologically independent samples; data represented mean \pm SEM). Scale bars: 20 μ m in **b**. * $P < 0.05$ by two-tailed Student's *t*-test in **c**.



Supplementary Figure 7 LCOR induces an interferon response and senescent-differentiation state. **(a)** K14 (green) and K8 (red) immunofluorescence staining of mammospheres formed by control or Lcor-OE P4 cells with or without IFN- α (1000 U/ml) treatment. **(b)** Quantification of mammospheres formed by HMLE after transduction with the indicated expression constructs, and with or without treatment with 1000 U/ml IFN- α (n=3 biologically independent samples; data represents mean \pm SEM), 10,000 cells were seeded and mammospheres counted at 8 days. **(c-d)** Quantification of PDX cell tumorspheres formed by 10,000 HCl-010 cells, with the indicated

conditions (n=3 biologically independent samples; data represents mean \pm SEM). **(e)** Senescence-associated β -galactosidase (SA- β gal) assay of HMLE cells with or without wt and Δ HTH mutant LCOR expression, and with or without IFN- α treatment for 48 hours (n=3 biologically independent samples; data represents mean \pm SEM; source of data in Supplementary Table 4). **(f)** GSEA of the senescence up-regulated gene set (M9143)⁴ in the ranked gene list of LCOR, LCOR- Δ HTH or miR-199a overexpressing HMLE cells vs. control. Scale bars: 25 μ m in **a**, and 100 μ m in **e**. * P <0.05, ** P <0.01, *** P <0.005 by two-tailed Student's t -test in **b**, **c** and **d**.

SUPPLEMENTARY INFORMATION



Supplementary Figure 9 Western blot scanned films. Boxes highlight lanes used in figures.

SUPPLEMENTARY INFORMATION

Supplementary Tables Legends

Supplementary Table 1 Significantly upregulated miRNAs in Claudin-low tumors and MaSCs.

Supplementary Table 2 Differential expression of the miR-199a predicted targets in MaSCs (P4) vs Luminal cells (P5).

Supplementary Table 3 Top ten correlated HALLMARK-MSigDB pathways by GSEA of the indicated conditions.

Supplementary Table 4 Statistics Source Data. Raw data for the miR-199a ectopic overexpression control tests in different mammary epithelial cells (Supplementary Fig. 1b); values represent relative miR-199a expression levels of each biologically independent experiment. Raw data for the senescence-associated β -galactosidase (SA- β gal) assays of HMLE (Supplementary Fig. 7e) and MDA-MB-231 (Fig. 7j) cells. Values represent the percentage of β -galactosidase positive cells; (n=3 biologically independent samples; data represents mean \pm SEM).

Supplementary Table 5 Information of primers used in cloning, mutagenesis and qRT-PCR.

Supplementary Table 6 Information of antibodies used in IHC, IF and western blot experiments.

References

1. Ben-Porath, I. *et al.* An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet* **40**, 499-507 (2008).
2. Jiang, Y.Z. *et al.* Transcriptome analysis of triple-negative breast cancer reveals an integrated mRNA-lncRNA signature with predictive and prognostic value. *Cancer research* (2016).
3. Chang, H.Y. *et al.* Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 3738-3743 (2005).
4. Fridman, A.L. & Tainsky, M.A. Critical pathways in cellular senescence and immortalization revealed by gene expression profiling. *Oncogene* **27**, 5975-5987 (2008).
5. Gyorffy, B. *et al.* An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1, 809 patients. *Breast Cancer Res Treat* **123**, 725-731 (2010).