s-SHIP Promoter Expression Identifies Mouse Mammary Cancer Stem Cells

Lu Tian, Marie-José Truong, Chann Lagadec, Eric Adriaenssens, Emmanuel Bouchaert, Hélène Bauderlique-Le Roy, Martin Figeac, Xuefen Le Bourhis, Roland Bourette

Version of Record: https://www.sciencedirect.com/science/article/pii/S2213671119301808
Manuscript_251c68716e8c45fb9e1e682eaed1130d

s-SHIP promoter expression identifies mouse mammary cancer stem cells

Lu Tian[1], Marie-José Truong[1,6], Chann Lagadec[2,6], Eric Adriaenssens[2], Emmanuel Bouchaert[3], Hélène Bauderlique-Le Roy[4], Martin Figeac[5], Xuefen Le Bourhis[2], Roland P. Bourette[1*]

[1] Univ. Lille, CNRS, Institut Pasteur de Lille, UMR 8161 - M3T - Mechanisms of Tumorigenesis and Targeted Therapies, F-59000 Lille, France

[2] Univ. Lille, Inserm, U908 - CPAC - Cell Plasticity and Cancer, F-59000 Lille, France

[3] Oncovet Clinical Research, SIRIC ONCOLille , F-59120 Loos, France

[4] BICeL platform, Institut Pasteur de Lille, F-59000 Lille, France

[5] Functional Genomics Platform, Univ. Lille, F-59000 Lille, France

[6] These authors contributed equally to this work

*Correspondence to:

Roland Bourette, Ph.D.

UMR8161, Institut de Biologie de Lille,

1 rue du Professeur Calmette, CS 54447, 59021 Lille Cedex, France

phone: (33) 3 20 87 12 96

E–mail: roland.bourette@ibl.cnrs.fr

Running Title: isolation of s-SHIP^{+} mammary cancer stem cells

Summary

During normal mammary gland development, s-SHIP promoter expression marks a distinct type of mammary stem cells, at two different stages, puberty and early mid-pregnancy. To determine if s-SHIP is a marker of mammary cancer stem cells (CSC), we generated bi-transgenic mice by crossing the C3(1)-SV40 T-antigen transgenic mouse model of breast cancer, and a transgenic mouse (11.5kb-GFP) expressing green fluorescent protein from the s-SHIP promoter. Here we show that in mammary tumors originating in these bi-transgenic mice, s-SHIP promoter expression enriches a rare cell population with CSC activity as demonstrated by sphere-forming assays *in vitro* and limiting dilution transplantation *in vivo*. These s-SHIP-positive CSC are characterized by lower expression of Delta like non-canonical Notch ligand 1 (DLK1), a negative regulator of Notch pathway. Inactivation of *Dlk1* in s-SHIP-negative tumor cells increases their tumorigenic potential, suggesting a role for DLK1 in mammary cancer stemness.

Introduction

Breast cancer is the leading cause of cancer-related death among female across the world. It is a highly complex disease marked by genetic and clinical heterogeneity (Torre et al., 2015). Intertumoral heterogeneity of breast cancers led to a classification into several subsets with varied patient outcomes and implications for treatment. This heterogeneity is associated with an epithelial hierarchy in tumors resembling that observed in normal mammary gland. On top of this hierarchy that forms the tumor bulk is a cell subpopulation described as cancer stem cells (CSC) or tumor-initiating cells (TIC) (Al-Hajj et al., 2003; Ginestier et al., 2007). CSC are defined as cells that retain extensive self-renewal potential through a series of generation and have the ability to recreate the heterogeneity of the original tumor through asymmetric division (Nassar and Blanpain, 2016). The nature of breast CSC is complex with the

demonstration of their heterogeneity both within and between breast tumors and the discovery of their plasticity (Sreekumar et al., 2015). Strong evidence has suggested that they are responsible for the recurrence, metastasis, and drug resistance of high-grade tumors (Peitzsch et al., 2017). A better understanding of breast CSC might, therefore, provide prognostic information for relapse and metastasis and lead to the development of more efficient therapies.

Genetically engineered mouse models (GEMM) have been extremely important in helping to elucidate the mechanisms of pathogenesis of breast cancer (Menezes et al., 2014). Different CSC subpopulations have been identified in these transgenic mouse models of mammary carcinogenesis using combinations of cell surface markers, such as $Lin^{-}CD29^{hi}CD24^{hi}$ (Zhang et al., 2008), $Thy1^{+}CD24^{+}$ (Cho et al., 2008), $CD29^{hi}CD24^{med}$ (Shafee et al., 2008), CD61 (Vaillant et al., 2008). However, most of the sorting schemes require multiparameter fluorescence-activated cell sorting (FACS), yield sometimes rather heterogeneous cell populations, and are unclear in their biological meaning (Medema, 2013). Due to the general lack of unique cell-surface markers, the development of new markers to prospectively identify putative CSC is of the utmost importance (Brooks et al., 2015).

Stem cell-specific expression of s-SHIP (stem-SH2-containing 5'-Inositol Phosphatase) was initially identified in embryonic and hematopoietic stem cells (Tu et al., 2001). A transgenic mouse (Tg 11.5kb–GFP or s-SHIP-GFP) was generated using the 11.5kb s-SHIP promoter that specifically expresses enhanced green fluorescent protein (GFP) in several stem cell populations during embryonic development, including skin epidermis, hair follicles, mammary gland, and prostate (Rohrschneider et al., 2005). In the mammary gland, s-SHIP/GFP labels puberty cap cells and pregnancy basal alveolar bud cells, with a stem cell potential (Bai and Rohrschneider, 2010). Notably, s-SHIP expression in cap cells is associated with the expression of Par3-like polarity protein, which is essential for mammary stem cell maintenance (Huo and Macara, 2014).

The C3(1)/SV40 T-antigen transgenic mouse model (Tg C3(1)-Tag mice) develops mammary hyperplasia by 3 months of age with subsequent development of mammary adenocarcinoma by 5-6 months of age (Maroulakou et al., 1994). C3(1)-Tag mouse is a murine model of human basal-like breast cancer, a complex group of breast cancers with both basal and luminal features (Dontu and Ince, 2015; Gusterson and Eaves, 2018). Using a cross-species genomics approach, murine expression profiles were compared to human breast cancers. The vast majority (≈90%) of the C3(1)-Tag mammary tumors displayed characteristics of human basal-like breast cancer, and 5-10% showed the claudin-low subtype feature (Herschkowitz et al., 2007; Pfefferle et al., 2013, Usary et al., 2016), demonstrating the model's strong resemblance to human disease.

Here we aim to assess if s-SHIP promoter expression could enable the isolation of a subpopulation of tumors cells with CSC characteristics *i.e*. higher ability to self-renew and to form spheres *in vitro*, and higher ability to form secondary tumors phenotypically similar to the primary tumor, following their transplantation into immunodeficient mice. By crossing Tg 11.5kb-GFP mice with Tg C3(1)-Tag mice, we showed the presence of a subpopulation of GFP^+ cells in mammary tumors. These GFP^+ mammary cancer cells are $CD24^+/CD49f^+/CD29^+$, and enriched for sphere-forming activity *in vitro*. Importantly, they can regenerate heterogeneous tumors that displayed properties similar to the primary tumor upon subsequent transplantation. Transcriptomic analysis showed that these s-SHIP-positive CSC were characterized by lower expression of Delta like non-canonical Notch ligand 1 (DLK1), a negative regulator of Notch pathway. Inactivation of *Dlk1* in s-SHIP-negative tumor cells increased their sphere-forming capacity and their tumorigenic potential, suggesting a role for DLK1 in mammary cancer cell stemness. Altogether, these results demonstrate that s-SHIP promoter expression offers a valuable marker for the isolation and characterization of mammary cancer stem cells.

Results

s-SHIP-GFP/C3(1)-Tag bi-transgenic mice develop mammary tumors containing a rare s-SHIP/GFP$^+$ cell subpopulation

We generated a bi-transgenic mouse model by crossing homozygous Tg 11.5kb-GFP mice with hemizygous Tg C3(1)-Tag mice. Progressive mammary gland lesions were observed in female mice that carried the T Ag-containing transgene, from ductal hyperplasia to adenocarcinoma (Figure 1A) (Figure S1A). All female mice developed multiple mammary tumors by 4-5 months of age. GFP$^+$ cells were detected on frozen sections (Figure 1B) and by flow cytometry after enzymatic digestion of tumors (Figure 1C) (1.03% ± 0.64% of total cells, n=10). The vast majority of GFP$^+$ cells were negative for lineage markers (Lin$^+$ GFP$^+$ cells = 0.08% ± 0.08% of total cells, n=5). These results indicated that s-SHIP promoter drives GFP expression specifically in a subpopulation of mammary tumor cells. Moreover, GFP expression correlated with the endogenous s-SHIP mRNA expression since sorted tumor GFP$^+$ cells expressed higher levels of s-SHIP mRNA as compared to sorted tumor GFP$^-$ cells (Figure S1C). Analysis of luminal (cytokeratin 8, K8) and basal/myoepithelial (cytokeratin 14, K14) markers showed that few tumors expressed K8 while all tumors displayed expression of K14 (Figure 1B). Importantly, GFP$^+$ tumor cells expressed K14 (Figure 1B). Similarly to s-SHIP/GFP expression at puberty during normal mammary gland development (Bai and Rohrschneider, 2010), some K14$^+$ mammary basal cells of 7-week-old biTg mice also expressed GFP (Figure S1B). We next examined the expression of cell surface markers historically associated with stem/progenitor cells in the mammary gland. Previous studies using flow cytometry to isolate mouse mammary stem cells have shown the majority of these cells have a CD49f^{hi} CD29^{hi} CD24$^+$ EpCAM$^+$ Sca-1^{neg} cell surface marker phenotype (Shackleton et al., 2006; Shehata et al., 2012; Sleeman et al., 2005; Stingl et al., 2006). Independent tumors were dissociated to single cell suspensions and stained for CD24, CD29,

CD49f, and EpCAM cell surface markers. Tumors displayed distinct FACS profiles showing heterogeneous expression for different markers but with enrichment for $CD24^{+}CD29^{+}$ and $CD24^{+}EpCAM^{+}$ cell subsets (Figure 1D) (Figure S1D). GFP^{+} cell population was homogeneous and the majority of cells were located in $Lin^{-}CD24^{+}$ cell subset, and expressed CD29, CD49f, and EpCAM cell surface markers (Figure 1D). Moreover, $Lin^{-}GFP^{+}$ cells showed a higher percentage of double positive $[CD24\ CD49f]^{+}$ as compared to total Lin^{-} tumor cells (Figure S1D).

s-SHIP/GFP^{+} tumor cells are enriched for sphere-forming and tumorigenic cells

To further characterize GFP^{+} cells, we used CD49f expression to separate GFP^{+} epithelial cells ($CD49f^{high}$) from the few GFP^{+} vascular smooth muscle cells ($CD49f^{-/low}$) as previously reported (Bai and Rohrschneider, 2010) (Figure S2A). To compare the frequency of tumor-initiating cells in GFP^{+} versus GFP^{-} cell subsets, cells were isolated by FACS and either cultured under serum-free suspension conditions to form mammospheres, or subcutaneously injected into SCID mice (Figure S2B). Since the majority of GFP^{+} cells were located in the Lin^{-} cell fraction, we compared $Lin^{-}CD49f^{+}GFP^{+}$ cells to $Lin^{-}CD49f^{+}GFP^{-}$ cells (Figure S2B). GFP^{+} cell populations showed a significant higher sphere-forming potential (2.55% ± 1.92%) as compared to the control GFP^{-} cell populations (0.12% ± 0.22%) (Figure 2A, left panel, n=7). Spheres derived from GFP^{+} cells were heterogeneous in size but most often larger compared to those derived from GFP^{-} cells (Figure 2A, right panels). To examine the *in vitro* self-renewal potential of GFP^{+} cells, primary spheres were dissociated into single cell suspensions and cells were plated in the same condition as for primary spheres. Secondary spheres derived from GFP^{+} subgroups were more numerous and larger as compared to secondary spheres derived from GFP^{-} subgroups (Figure 2B, n=3). Spheres initially derived from GFP^{+} cell subsets can be maintained through at least 4 passages (not shown). It is to note that few GFP^{+} cells were always observed in the spheres at all passages (Figures 2, S3A).

We next performed serial transplantation studies to evaluate the tumorigenic potential of GFP^+ cells versus GFP^- cells and to determine whether GFP^+ cells were able to self-renew *in vivo* and to recapitulate the cell heterogeneity of the original tumors. When injected subcutaneously in the region next to the nipple of the third thoracic gland of immunodeficient SCID mice, GFP^+ cells displayed significantly higher tumorigenic potential as compared to GFP^- cells (Figure 3A). The calculated TIC frequency in the $Lin^-CD49f^+GFP^-$ cell population was of 1 in 638 tumor cells whereas it was of 1 in 118 in the $Lin^-CD49f^+GFP^+$ cell population. No difference was observed in tumor latency (Figure S3B). Similar TIC frequency was observed $CD49f^+GFP^+$ without lineage depletion (Figure S2C). For $CD49f^+GFP^-$ cells, TIC frequency was lower compared to $Lin^-CD49f^+GFP^-$ cells. This could be explained by a dilution effect of numerous Lin^+ cells present in the tumor, but also may indicate an inhibitory effect of some Lin^+ cells on tumorigenicity. C3(1)-Tag mammary tumor cells and tumor-derived cell lines cannot be transplanted into immunocompetent FVB mice since they are rejected likely due to the expression of the SV40 T antigen (Aprelikova et al., 2016). However, we have successfully transplanted primary tumor cells into FVB mice but with a significantly lower engraftment rate as well as longer tumor formation period (four to seven months) compared to SCID mice (one to three months, Figure S3B). Nevertheless, $Lin^-CD49f^+GFP^+$ cells were again capable of generating more tumors than $Lin^-CD49f^+GFP^-$ cells (Figure S2D).

Flow cytometry analysis of secondary tumors generated by $Lin^-CD49f^+GFP^+$ cells injected into SCID mice showed that the expression patterns of cell surface markers were similar to those of the original tumor (Figure 3B). Interestingly, GFP^+ cells were also present in these secondary tumors (Figure 3B). The percentage of GFP^+ cells in secondary tumors was higher but not significantly different (Figure 3C). Transplantation experiments showed that both $Lin^-CD49f^+GFP^+$ and $Lin^-CD49f^+GFP^-$ cells isolated from these secondary tumors had higher

tumor-forming capacities than their corresponding cell populations from primary tumors (Figure 3A,D). This suggests that both cell populations changed toward a more aggressive phenotype. Importantly, $Lin^-CD49f^+GFP^+$ cells still exhibited a higher tumorigenic potential in these secondary injections (1 in 20) as compared to $Lin^-CD49f^+GFP^-$ cells (1 in 157) (Figure 3D).

Microarray analysis revealed DLK1 as a negative regulator of stemness

To identify molecular differences between $Lin^-CD49f^+GFP^+$ and $Lin^-CD49f^+GFP^-$ cells, we performed transcriptional profiling using Agilent microarray technology. As shown in Figure 4A, $Lin^-CD49f^+GFP^+$ and $Lin^-CD49f^+GFP^-$ cell subpopulations shared similar gene expression pattern (GEO accession number GSE108373). Nevertheless, microarray analysis revealed several genes whose expression differs significantly between the two subpopulations (Figure 4B) (Table S1). Granzyme E (GZME), was significantly upregulated in $Lin^-CD49f^+GFP^+$ cells (fold change of 29.8) (Figure 4B). Granzymes are a family of serine proteases expressed by immune cells (Arias et al., 2017). As a first approach, we delivered siRNA to knock-down *Gzme* expression in $Lin^-CD49f^+GFP^+$ cells and investigated the outcome of *Gzme* silencing on the sphere-forming potential of these cells. As shown in Figure S3C, the down-regulation of *Gzme* didn't affect the sphere-forming potential of GFP^+ cells. Then, we were interested in the Delta like non-canonical Notch ligand 1 (*Dlk1*) gene that was significantly downregulated in $Lin^-CD49f^+GFP^+$ cells (fold change of 4.5) (Figure 4B). *Dlk1* gene encodes for a single-spanning transmembrane protein in the epidermal growth factor (EGF)-like family, which includes NOTCH receptors and their ligands (Tanimizu et al., 2003). DLK1 is a non-canonical Notch ligand and has been shown to act as an inhibitor of Notch signaling *in vitro* (Nueda et al., 2017). Microarray analysis revealed the expression of *Notch-1*, *-3*, and *-4* genes in Lin^-CD49f^+ cell populations. Chemical inhibition of NOTCH processing with the γ-secretase inhibitor Compound E decreased sphere-forming potential of Lin^-

$CD49f^+GFP^+$ cells (Figure S3D). Thus, we hypothesized that repression of the Notch negative regulator DLK1 in $Lin^-CD49f^+GFP^+$ cells might favor their CSC phenotype. The difference in *Dlk1* expression between $Lin^-CD49f^+GFP^-$ and $Lin^-CD49f^+GFP^+$ subpopulations was first confirmed by RT-qPCR using total RNA isolated from these two subpopulations. Similar to microarray result, we observed a fold-change of 11 ± 8 (Figure 4C, n=4 different tumors). We thus delivered siRNA to knock-down *Dlk1* expression in $Lin^-CD49f^+GFP^-$ cells and performed sphere-forming assays. $Lin^-CD49f^+GFP^-$ cell subpopulation was transfected with si*Dlk1* or siScramble (siScr) and were then cultured in serum-free sphere medium under non-adherent conditions for 7 to 10 days. A transient down-regulation of *Dlk1* expression was obtained with 77 % of inhibition after 2 days of culture, and only 10 % of inhibition in the generated spheres at day-8 (Figure 4D). We observed a significant increase of the sphere-forming potential of si*Dlk1*-treated cells, (0.82 % ± 0.21 %) as compared to siScr-treated cells (0.42 % ± 0.12 %) (Figure 4E). On the contrary, a down-regulation of 80 % of *Dlk1* expression in $Lin^-CD49f^+GFP^+$ cells (Figure 4D) did not interfere with their sphere-forming potential (Figure 4E), suggesting that the level of *Dlk1* expression in these cells may be too low to have any inhibitory effect on their sphere-forming capacity. Since si*Dlk1* enhanced *in vitro* sphere-forming potential of $Lin^-CD49f^+GFP^-$ cells, we next performed transplantation studies to evaluate the tumorigenic potential of si*Dlk1*- and siScr-transfected $Lin^-CD49f^+GFP^-$ cells. The calculated TIC frequency in the si*Dlk1* transfected $Lin^-CD49f^+GFP^-$ cell population was of 1 in 340 tumor cells whereas it was of 1 in 859 in the si*Scr* transfected $Lin^-CD49f^+GFP^-$ cells (Figure 4F). No difference was observed in the tumor latency (Figure S3E). Thus, down-regulation of *Dlk1* in $Lin^-CD49f^+GFP^-$ cells significantly increased both their sphere-forming potential *in vitro* and their tumorigenic potential *in vivo* as compared to control cells. Altogether, these results suggest that DLK1 could act as a negative regulator of stemness of mammary cancer cells.

Discussion

Stem-cell specific expression of s-SHIP was initially identified in hematopoietic and embryonic stem cells (Tu et al. 2001, Rohrschneider et al., 2005). Using s-SHIP-GFP promoter reporter, its stem cell-specific expression has been demonstrated in prostate tissues (Bauderlique-Le Roy et al., 2015; Brocqueville et al., 2016; Al Shareef et al., 2018). In postnatal mammary gland development, GFP^+ cap cells in puberty and basal alveolar bud cells in pregnancy each exhibit stem cell properties (Bai and Rohrschneider, 2010). In the same study, the authors described the presence of GFP^+ tumor cells in heterogeneous luminal/basal-like mammary tumors originating in bi-transgenic s-SHIP-GFP/MMTV-Wnt1 mice. Subsequent studies have demonstrated the value of s-SHIP promoter to investigate normal mammary stem cells (Kogata et al., 2013; Huo and Macara, 2014; Roarty et al., 2015). Yet s-SHIP/GFP usefulness in isolating mammary cancer stem cells remained to be determined. Here we used a mouse model of basal-like breast cancer to show that s-SHIP is a marker of mammary CSC. In our s-SHIP-GFP/C3(1)-Tag bi-transgenic mice, s-SHIP/GFP^+ cells expressed commonly used mammary CSC cell surface markers, exhibited a higher sphere-forming potential *in vitro* and a higher ability to reform secondary tumors when transplanted into recipient mice. Interestingly, transplanted s-SHIP/GFP^+ cells generated heterogeneous tumors that displayed properties similar to the primary tumors, including the presence of a subset of s-SHIP/GFP^+ cells with CSC properties. Bai and Rohrschneider (2010) did not observe s-SHIP/GFP^+ cells in tumors of bi-transgenic s-SHIP-GFP/MMTV-ErbB2/neu/HER2 mice exhibiting luminal features. Considering breast CSC heterogeneity, it remains to determine if s-SHIP expression is restricted or not to CSC in basal-like breast cancer. To address this question, we aim to characterize the exact nature of s-SHIP/GFP^+ tumor cells that have been observed in heterogeneous mammary tumors developed by bi-

transgenic s-SHIP-GFP/MMTV-Wnt1 mice (Bai and Rohsrchneider, 2010), together with an in-depth analysis of TCGA data for breast cancer.

Transcriptional profiling showed few genes differentially expressed between $Lin^-CD49f^+GFP^+$ and $Lin^-CD49f^+GFP^-$ cells. Among them, the significant decrease of *Dlk1* expression in GFP^+ cells has addressed our attention. DLK1 acts as a non-canonical ligand of the Notch signaling pathway and is expressed at a high frequency in various human tumors including breast carcinoma (Yanai et al., 2010). Several studies have demonstrated that DLK1 can act as an inhibitor of Notch signaling *in vitro*, including in breast cancer cells (Nueda et al., 2017). Our results strongly suggest that DLK1 acts as a negative regulator of stemness in breast cancer, which agrees with several studies pointing to a role of Notch pathway in breast cancer, especially in subtypes with stem-like features and EMT (Harrison et al., 2010). Based on our results, we can hypothesize that DLK1 interacts with Notch receptor expressed by tumor cells and can act as a negative regulator of Notch activation and signaling in these tumors cells. It is likewise possible that DLK1 action implicates cells of the tumor microenvironment. Interestingly, a recent study showed that Notch ligand Dll1 mediates crosstalk between mammary stem cells and tissue-resident macrophages (Chakrabarti *et al.* 2018*).* This implies that DLK1 may also be involved in the modulation of the oncogenic potential of mammary tumor cells through the inhibition of Notch signaling in surrounding macrophages. To address this issue, it would be important to determine the relationship between breast CSC, macrophages, DLK1, and Notch signaling.

Taken together, our results strongly support the relationship between mammary CSC and s-SHIP expression. This is in accordance with high s-SHIP expression in murine claudin-low tumors that are enriched for CSC signature (Herschkowitz et al., 2012). s-SHIP protein is a shorter isoform of SHIP1 protein that lacks the SH2 domain in its N-terminal region (Tu et al., 2001). Yet the function of s-SHIP protein remains unknown. SHIP2, the ubiquitous

homolog of SHIP1 is crucial for maintaining the ER-negative breast cancer stem cells through activation of Akt and JNK (Fu et al., 2014). Since our data indicate s-SHIP as a potential clinical target for breast cancer therapy, expression and role of s-SHIP in human breast cancers will be the subject of our future investigation.

Experimental procedures

All animal procedures were performed under a protocol approved by the Animal Protocol Review Committees of the Institut Pasteur de Lille (France) in accordance with European regulation (No. 01989.02).

Statistics

Data are expressed as means ± standard error of mean (SEM) of at least 3 independent experiments. The statistical analysis was done by using Two-way ANOVA or Student's t-test with GraphPad Prism5 and p-value < 0.05 was considered significant. For *in vivo* tumor-initiating potential, the results were analyzed by Extreme Limiting Dilution Analysis (ELDA) (Hu and Smyth, 2009).

Accession number

The accession number for the microarray dataset reported in this paper is GSE108373

Supplemental information

Supplemental information includes Supplemental Experimental Procedures, three figures and one Table.

Author contributions

L.T., M.J.T., E.B., H.B.L., M.F., R.P.B did the experiments L.T, C.L, E.A., X.L.B., R.P.B designed the study. L.T and R.P.B wrote the main manuscript text, and all authors reviewed the manuscript.

Acknowledgments

We thank S. Sebda and F. Leprêtre (Functional Genomic platform, Lille) for microarrays hybridization and for GEO repository Assistance, Q. Pascal (OCR, Loos) for anatomopathological analysis, S. Salomé-Desnoulez (BICel), M-C. Bouchez (CNRS) and T. Chassat (IPL) for their help.

Funding

This study was supported by grants from the SIRIC ONCOLille (INCa-DGOS-Inserm 6041), the Ligue Contre le Cancer (comités du Nord et du Pas-de-Calais), and the Cancéropole Nord-Ouest. L.T was supported by fellowships from the Région Hauts-de-France and the Fondation ARC pour la Recherche sur le Cancer.

References

Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J., and Clarke, M.F. (2003). Prospective identification of tumorigenic breast cancer cells. Proc. Natl. Acad. Sci. USA *100*, 3983–3988.

Al Shareef, Z., Kardooni, H., Murillo-Garzón, V., Domenici, G., Stylianakis, E., Steel, J.H., Rabano, M., Gorroño-Etxebarria, I., Zabalza, I., Vivanco M.D., et al. (2018). Protective effect of stromal dickkopf-3 in prostate cancer: opposing roles for TGFBI and ECM-1. Oncogene *37*, 5305-5324.

Aprelikova, O., Tomlinson, C.C., Hoenerhoff, M., Hixon, J.A., Durum, S.K., Qiu, T., He, S., Burkett, S., Liu, Z.-Y., Swanson, S.M., et al. (2016). Development and preclinical application of an immunocompetent transplant model of basal breast cancer with lung, liver and brain metastases. PLoS One *11*, e0155262.

Arias, M., Martinez-Lostao, L., Santiago, L., Ferrandez, A., Granville, D.J., and Pardo, J. (2017). The untold story of granzymes in oncoimmunology: novel opportunities with old acquaintances. Trends Cancer *3*, 407-422.

Bai, L., and Rohrschneider, L.R. (2010). s-SHIP promoter expression marks activated stem cells in developing mouse mammary tissue. Genes Dev. *24*, 1882–1892.

Bauderlique-Le Roy, H., Vennin, C., Brocqueville, G., Spruyt, N., Adriaenssens, E., and Bourette, R.P. (2015). Enrichment of human stem-like prostate cells with s-SHIP promoter activity uncovers a role in stemness for the long noncoding RNA H19. Stem Cells Dev. *24*, 1252–1262.

Brocqueville, G., Chmelar, R.S., Bauderlique-Le Roy, H., Deruy, E., Tian, L., Vessella, R.L., Greenberg, N.M., Rohrschneider, L.R., and Bourette, R.P. (2016). s-SHIP expression identifies a subset of murine basal prostate cells as neonatal stem cells. Oncotarget *7*, 29228–29244.

Brooks, M.D., Burness, M.L., and Wicha, M.S. (2015). Therapeutic implications of cellular heterogeneity and plasticity in breast cancer. Cell Stem Cell *17*, 260-271.

Chakrabarti, R., Celià-Terrassa, T., Kumar, S., Hang, X., Wei, Y., Choudhury, A., Hwang, J., Peng, J., Nixon, B., Grady, J.J., et al (2018). Notch ligand Dll1 mediates cross-talk between mammary stem cells and the macrophageal niche. Science *360*, eaan4153

Cho, R.W., Wang, X., Diehn, M., Shedden, K., Chen, G.Y., Sherlock, G., Gurney, A., Lewicki, J., and Clarke, M.F. (2008). Isolation and molecular characterization of cancer stem cells in MMTV-Wnt-1 murine breast tumors. Stem Cells *26*, 364–371.

Dontu, G., and Ince, T.A. (2015). Of mice and women: a comparative tissue biology perpective of breast stem cells and differentiation. J. Mammary Gland Biol. Neoplasia *20*, 51-62.

Fu, C.-H., Lin, R.-J., Yu, J., Chang, W.-W., Liao, G.-S., Chang, W.-Y., Tseng, L.-M., Tsai, Y.-F., Yu, J.-C., and Yu, A.L. (2014). A novel oncogenic role of inositol phosphatase SHIP2 in ER-negative breast cancer stem cells: involvement of JNK/vimentin activation. Stem Cells *32*, 2048–2060.

Gusterson, B., and Eaves, C.J. (2018). Basal-like breast cancers: from pathology to biology and back again. Stem Cell Rep. *10*, 1676-1686.

Ginestier, C., Hur, M.H., Charafe-Jauffret, E., Monville, F., Dutcher, J., Brown, M., Jacquemier, J., Viens, P., Kleer, C.G., Liu, S., et al. (2007). ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell Stem Cell *1*, 555–567.

Harrison, H., Farnie, G., Brennan, K.R., and Clarke, R.B. (2010). Breast cancer stem cells: something out of notching? Cancer Res. *70*, 8973-8976.

Herschkowitz, J.I., Simin, K., Weigman, V.J., Mikaelian, I., Usary, J., Hu, Z., Rasmussen, K.E., Jones, L.P., Assefnia, S., Chandrasekharan, S., et al. (2007). Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biol. *8*, R76.

Herschkowitz, J.I., Zhao, W., Zhang, M., Usary, J., Murrow, G., Edwards, D., Knezevic, J., Greene, S.B., Darr, D., Troester, M.A., et al. (2012). Comparative oncogenomics identifies breast tumors enriched in functional tumor-initiating cells. Proc. Natl. Acad. Sci. USA *109*, 2778–2783.

Hu, Y., and Smyth, G.K. (2009). ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. J. Immunol. Methods *347*, 70-78.

Huo, Y., and Macara, I.G. (2014). The Par3-like polarity protein Par3L is essential for mammary stem cell maintenance. Nat. Cell Biol. *16*, 526–534.

Kogata, N., Zvelebil, M., and Howard, B.A. (2013). Neuregulin 3 and Erbb signalling networks in embryonic mammary gland development. J. Mammary Gland Biol. Neoplasia *18*, 149–154.

Maroulakou, I.G., Anver, M., Garrett, L., and Green, J.E. (1994). Prostate and mammary adenocarcinoma in transgenic mice carrying a rat C3(1) simian virus 40 large tumor antigen fusion gene. Proc. Natl. Acad. Sci. U. S. A. *91*, 11236–11240.

Medema, J.P. (2013). Cancer stem cells: the challenges ahead. Nat. Cell Biol. *15*, 338–344.

Menezes, M.E., Das, S.K., Emdad, L., Windle, J.J., Wang, X.Y., Sarkar, D., and Fisher, P.B. (2014). Genetically engineered mice as experimental tools to dissect the critical events in breast cancer. Adv. Cancer Res. *121*, 331–382.

Nassar, D., and Blanpain, C. (2016). Cancer stem cells: basic concepts and therapeutic implications. Annu. Rev. Pathol. *11*, 47–76.

Nueda, M.L., Naranjo, A.I., Baladrón, V., and Laborda, J. (2017). Different expression levels of DLK1 inversely modulate the oncogenic potential of human MDA-MB-231 breast cancer cells through inhibition of NOTCH1 signaling. FASEB J. *31*, 3484–3496.

Peitzsch, C., Tyutyunnykova, A., Pantel, K., and Dubrovska, A. (2017). Cancer stem cells: the root of tumor recurrence and metastases. Semin. Cancer Biol. *44*, 10–24.

Pfefferle, A.D., Herschkowitz, J.I., Usary, J., Harrell, J.C., Spike, B.T., Adams, J.R., Torres-Arzayus, M.I., Brown, M., Egan, S.E., Wahl, G.M., et al. (2013). Transcriptomic classification of genetically engineered mouse models of breast cancer identifies human subtype counterparts. Genome Biol. *14*, R125.

Roarty, K., Shore, A.N., Creighton, C.J., and Rosen, J.M. (2015). Ror2 regulates branching, differentiation, and actin-cytoskeletal dynamics within the mammary epithelium. J. Cell Biol. *208*, 351–366.

Rohrschneider, L.R., Custodio, J.M., Anderson, T.A., Miller, C.P., and Gu, H. (2005). The intron 5/6 promoter region of the *ship1* gene regulates expression in stem/progenitor cells of the mouse embryo. Dev. Biol. *283*, 503–521.

Shackleton, M., Vaillant, F., Simpson, K.J., Stingl, J., Smyth, G.K., Asselin-Labat, M.L., Wu, L., Lindeman, G.J., and Visvader, J.E. (2006). Generation of a functional mammary gland from a single stem cell. Nature *439*, 84–88.

Shafee, N., Smith, C.R., Wei, S., Kim, Y., Mills, G.B., Hortobagyi, G.N., Stanbridge, E.J., and Lee, E.Y. (2008). Cancer stem cells contribute to cisplatin resistance in Brca1/p53–mediated mouse mammary tumors. Cancer Res. *68*, 3243–3250.

Shehata, M., Teschendorff, A., Sharp, G., Novcic, N., Russell, I.A., Avril, S., Prater, M., Eirew, P., Caldas, C., Watson, C.J., et al. (2012). Phenotypic and functional characterisation of the luminal cell hierarchy of the mammary gland. Breast Cancer Res. *14*, R134.

Sleeman, K.E., Kendrick, H., Ashworth, A., Isacke, C.M., and Smalley, M.J. (2005). CD24 staining of mouse mammary gland cells defines luminal epithelial, myoepithelial/basal and non-epithelial cells. Breast Cancer Res. *8*, R7.

Sreekumar, A., Roarty, K., and Rosen, J.M. (2015). The mammary stem cell hierarchy: a looking glass into heterogeneous breast cancer landscapes. Endocr. Relat. Cancer *22*, T161-T176.

Stingl, J., Eirew, P., Ricketson, I., Shackleton, M., Vaillant, F., Choi, D., Li, H.I., and Eaves, C.J. (2006). Purification and unique properties of mammary epithelial stem cells. Nature *439*, 993–997.

Tanimizu, N., Nishikawa, M., Saito, H., Tsujimura, T., and Miyajima, A. (2003). Isolation of hepatoblasts based on the expression of Dlk/Pref-1. J. Cell Sci. *116*, 1775–1786.

Torre, L.A., Bray, F., Siegel, R.L., Ferlay, J., Lortet-Tieulent, J., and Jemal, A. (2015). Global cancer statistics, 2012. CA Cancer J. Clin. *65*, 87–108.

Tu, Z., Ninos, J.M., Ma, Z., Wang, J.W., Lemos, M.P., Desponts, C., Ghansah, T., Howson, J.M., and Kerr, W.G. (2001). Embryonic and hematopoietic stem cells express a novel SH2–containing inositol 5′-phosphatase isoform that partners with the Grb2 adapter protein. Blood *98*, 2028–2038.

Usary, J., Darr, D.B., Pfefferle, A.D., and Perou, C.M. (2016). Overview of genetically engineered mouse model of distinct breast cancer subtypes. Curr. Protoc. Pharmacol. *72*, 14.38.1-14.38.11.

Vaillant, F., Asselin-Labat, M.-L., Shackleton, M., Forrest, N.C., Lindeman, G.J., and Visvader, J.E. (2008). The mammary progenitor marker CD61/β3 integrin identifies cancer stem cells in mouse models of mammary tumorigenesis. Cancer Res. *68*, 7711–7717.

Yanai, H., Nakamura, K., Hijioka, S., Kamei, A., Ikari, T., Ishikawa, Y., Shinozaki, E., Mizunuma, N., Hatake, K., and Miyajima, A. (2010). Dlk-1, a cell surface antigen on foetal

hepatic stem/progenitor cells, is expressed in hepatocellular, colon, pancreas and breast carcinomas at a high frequency. J. Biochem. *148*, 85-92.

Zhang, M., Behbod, F., Atkinson, R.L., Landis, M.D., Kittrell, F., Edwards, D., Medina, D., Tsimelzon, A., Hilsenbeck, S., Green, J.E., et al. (2008). Identification of tumor-initiating cells in a p53 null mouse model of breast cancer. Cancer Res. *68*, 4674-4682.

Figure legends

Figure 1. s-SHIP/GFP expression is detected in mammary tumors of s-SHIP-GFP/C3(1)-Tag bi-transgenic mice. (A) H&E staining of paraffin-embeded sections of mammary tumors illustrating different stages of tumor development: a, ductal hyperplasia; b, ductal carcinoma *in situ*; c, infiltrative solid carcinoma; d, solid carcinoma; (a,b,d, scale bars = 50µm ; c, scale bar = 100µm) (see also Figure S1A). (B) Representative images (n=4) of frozen sections of mammary tumors containing GFP^+ cells (green) and stained with cytokeratin 14 (K14, red, upper panels), cytokeratin 8 (K8, red, lower panels) and with DAPI nuclear stain (blue) (scale bar = 100µm left panels and 10 µm right panels). Images of 7-week-old mammary gland stained with K14 and K8 are shown in Figure S1B. (C) Representative flow cytometry analysis (n=5) of GFP expression in dissociated mammary cancer cells isolated from a 5-month-old bi-transgenic mouse. P1 gate is set as to exclude dead cells (Zombie $violet^+$). See also Figure S1C. (D) Representative flow cytometry analysis (n=3) of total Lin^- cells, Lin^- GFP^- and Lin^-GFP^+ cells from two different tumors: T1 and T2. Cells were analyzed for the expression of CD24, CD29, CD49f, and EpCAM cell surface markers. See also Figure S1D.

Figure 2. s-SHIP/GFP^+ cells have higher sphere-forming potential and *in vitro* self-renewal capacity. (A) Primary mammospheres derived from $Lin^-CD49f^+GFP^-$ and Lin^- $CD49f^+GFP^+$ cells isolated from bi-Tg mammary tumors. Cells were seeded by limited dilution and grown in suspension for 7 to 10 days. Data represent the mean ± SEM of seven

independent experiments, *p* values were determined by Student's t-test ***$p < 0.001$. Representative pictures of primary spheres after 7 to 10 days of culture (right panels). (B) Primary spheres derived from $Lin^-CD49f^+GFP^-$ or $Lin^-CD49f^+GFP^+$ cells were dissociated into single cells. Cells were seeded at 200 cells per well in triplicate and grown in suspension for 7 to 10 days for secondary mammosphere formation. Data represent mean values ± SEM of three different experiments, *p* values were determined by Student's t-test *$p<0.05$. Representative pictures of secondary spheres after 7 to 10 days of culture (right panels). (scale bar = 100µm). See also Figure S2A,B for cell sorting strategy and experimental design and Figure S3A.

Figure 3. s-SHIP/GFP$^+$ cells are enriched for tumor-initiating cells and generate tumors that mimic the parental tumors. (A) Flow cytometry was used to separate dissociated tumor cells based on cell surface markers and GFP expression (see Figure S2B for experimental procedure). The collected cell populations were injected subcutaneously into recipient SCID female mice in a limiting dilution manner. Mice were monitored until tumors were observed or up to 7 months if no tumors were detected. See also Figure S2C,D and Figure S3B. (B) Secondary tumors derived from $Lin^-CD49f^+GFP^+$ cell fraction have similar flow cytometric profiles as the primary tumors. Representative flow cytometry analysis of dissociated cells from a primary tumor (parental) and two respective secondary tumors (secondary 1 and 2) were presented (C) the percentage of $Lin^-CD49f^+GFP^+$ cells in secondary tumors and corresponding primary parental tumors Data represent mean values ± SEM of four different experiments, *p* values were determined by Student's t-test, ns: not significant (D) Cells from secondary tumors generated by $Lin^-CD49f^+GFP^+$ fraction were freshly digested and FACS sorted into $Lin^-CD49f^+GFP^-$ or $Lin^-CD49f^+GFP^+$ subpopulations and injected in SCID mice as described in (A). Tumor-initiating cell frequencies were generated by ELDA: extreme

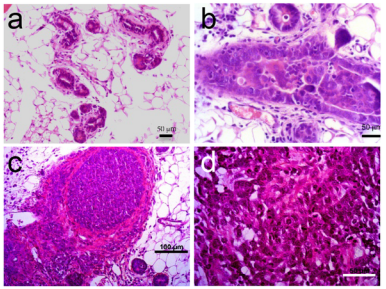
limiting dilution analysis (A,D). See also Figure S2A,B for cell sorting strategy and experimental design.

Figure 4. Gene expression analysis of Lin$^-$CD49f$^+$GFP$^-$ versus Lin$^-$CD49f$^+$GFP$^+$ tumor cells reveals the role of DLK1 in the regulation of stemness. (A and B) Fifteen genes were selected from the master list of genes differentially regulated (at least two-fold with a *p* value < 0.05). A list of 355 deregulated probes (fold change ≥1) is presented in Table S1. For the heatmap, each column represents a probe; each row represents various subpopulations from biological replicates. The red color indicates high-level expression while blue indicates a low level of expression. (C) RT-qPCR confirmed the downregulation of *Dlk1* expression in Lin$^-$CD49f$^+$GFP$^+$ cells as compared to Lin$^-$CD49f$^+$GFP$^-$ cells. Data represent mean values ± SEM from four different tumors (T1 to T4). (D and E) Inhibition of *Dlk1* expression in Lin$^-$CD49f$^+$GFP$^-$ cells increased their sphere-forming potential. si*Dlk1* or control siScramble (siScr) were introduced into Lin$^-$CD49f$^+$GFP$^-$ and Lin$^-$CD49f$^+$GFP$^+$ cells isolated from bi-Tg mammary tumors. (D) The level of *Dlk1* mRNA was determined by RT-qPCR, 2 days or 8 days after siRNA transfection. (E) Lin$^-$CD49f$^+$GFP$^-$ and Lin$^-$CD49f$^+$GFP$^+$ cells were treated by siScr or si*Dlk1*, and plated under sphere conditions. Spheres were counted after 7 to 10 days in culture. (D and E) Data represent mean values ± SEM of six independent experiments, *p* values were determined by Student's t-test ***$p<0.001$ **$p<0.01$ ns: not significant. See Figure S3C,D for additional experiments. (F) siScr or si*Dlk1* transfected Lin$^-$CD49f$^+$GFP$^-$ cells were injected subcutaneously into recipient SCID female mice in a limiting dilution manner. Mice were monitored until tumors were observed or up to 3 months if no tumors were detected. Tumor-initiating cell frequencies were generated by ELDA. See also Figure S3E.

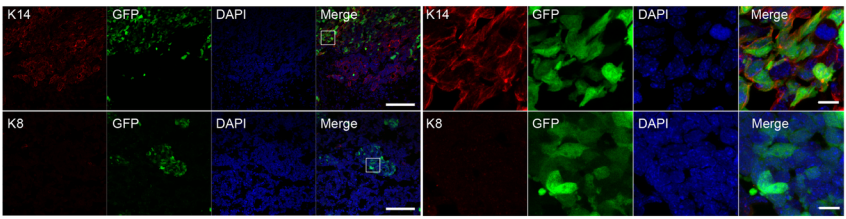
TABLE S1: List of the 355 deregulated probes identified by micro-array analysis (p-value (p) less than 0.01 and fold change (FC) above or equal 1). The p-value was calculated with a paired t-test. RELATED TO FIGURE 4

Figure 1

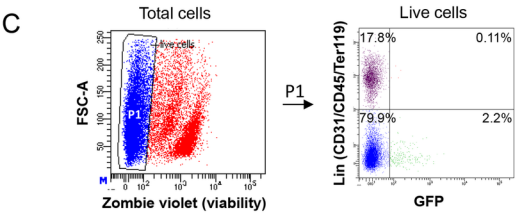
A



B



C



D

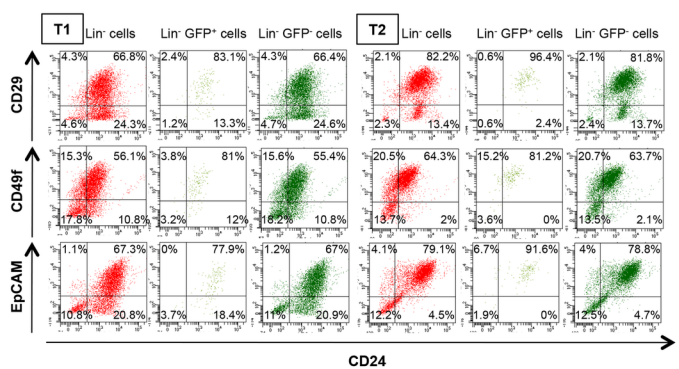
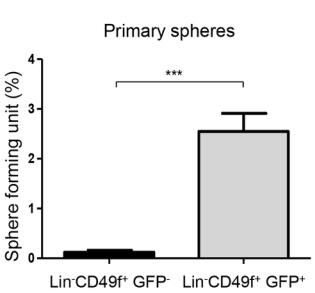
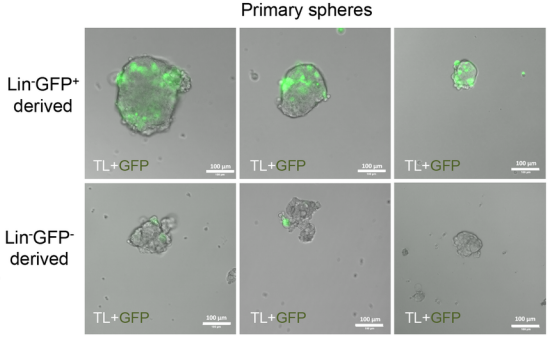


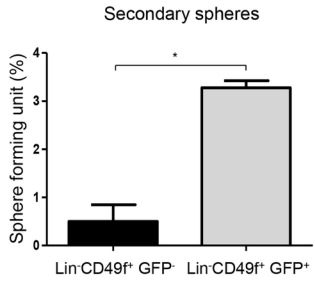
Figure 2

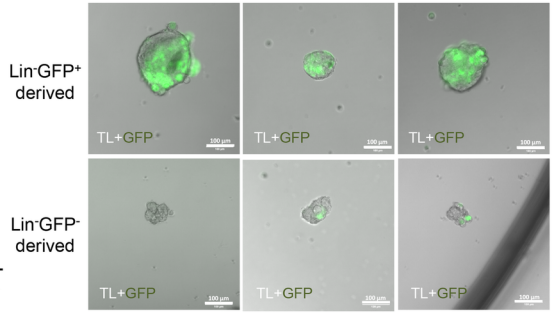
A


Primary spheres
Sphere forming unit (%)
4
3
2
1
0
***
Lin-CD49f+ GFP-
Lin-CD49f+ GFP+


Primary spheres
Lin-GFP+ derived
Lin-GFP- derived
TL+GFP
100 µm

B


Secondary spheres
Sphere forming unit (%)
4
3
2
1
0
*
Lin-CD49f+ GFP-
Lin-CD49f+ GFP+


Secondary spheres
Lin-GFP+ derived
Lin-GFP- derived
TL+GFP
100 µm

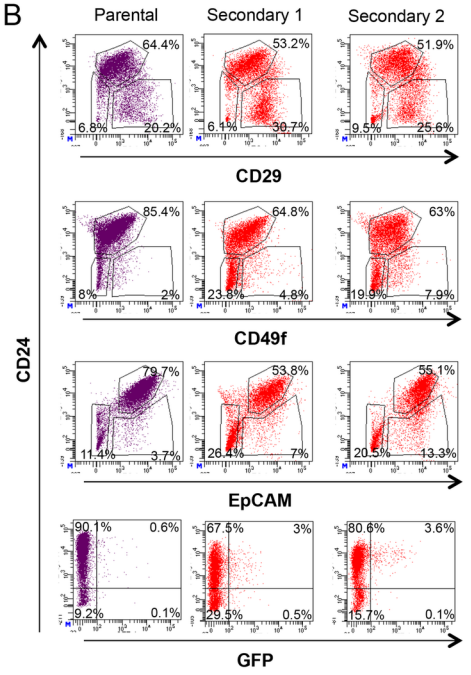
A

Primary injection in SCID mice

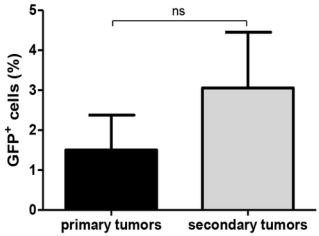
	1000 cells	100 cells	50 cells	Tumor initiating cell frequency (95% CI)
Lin$^-$ CD49f$^+$ GFP$^-$	5/6	3/22	1/17	**1/638** (1/1257 - 1/324)
Lin$^-$ CD49f$^+$ GFP$^+$	6/6	10/22	9/17	**1/118** (1/188 - 1/74)

P=2.65^{e-05}

B



C



D

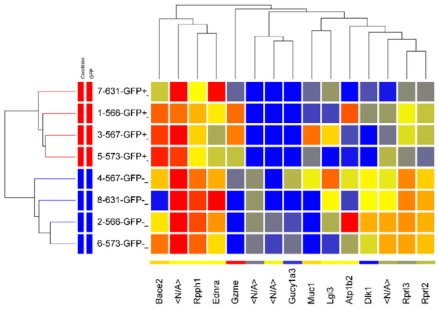
Secondary injection in SCID mice

	1000 cells	100 cells	50 cells	Tumor initiating cell frequency (95% CI)
Lin$^-$ CD49f$^+$ GFP$^-$	4/4	6/11	2/11	**1/157** (1/311 - 1/80)
Lin$^-$ CD49f$^+$ GFP$^+$	4/4	11/11	10/11	**1/20** (1/40 - 1/10)

P=5.32^{e-06}

Figure 4

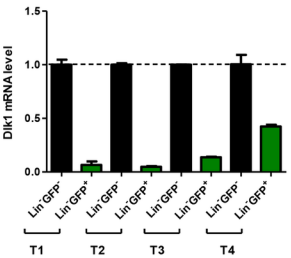
A



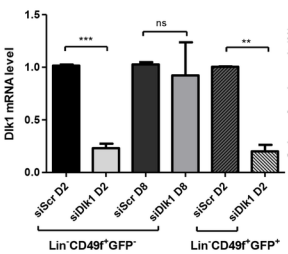
B

GeneSymbol	Description	p	Regulation	FC (abs)
Gzme	granzyme E	**0.00991789**	up	29.83715
Dlk1	delta-like 1 homolog (Drosophila)	**0.00414383**	down	4.5239267
Gucy1a3	guanylate cyclase 1, soluble, alpha 3	**0.00643437**	down	4.0955086
	Mus musculus cDNA clone IMAGE:40090117, [BC128469]	**0.00405105**	down	3.5768538
Rprl3	ribonuclease P RNA-like 3	**0.00658998**	down	3.2771783
	lincRNA:chr13:22110479-22116629 reverse strand	**0.00622081**	down	3.1527982
Rprl2	ribonuclease P RNA-like 2	**0.00696632**	down	2.9873924
Rpl7	ribosomal protein L7A	**0.00653786**	down	2.50248
Rpph1	ribonuclease P RNA component H1	**0.00963502**	down	2.4559698
Bace2	beta-site APP-cleaving enzyme 2	**0.00738185**	up	2.404046
Muc1	mucin 1, transmembrane	**0.00493007**	up	2.170539
Atp1b2	ATPase, Na+/K+ transporting, beta 2 polypeptide	**0.00842456**	down	2.1331136
Ednra	endothelin receptor type A	**0.00942573**	down	2.0767322
	lincRNA:chr6:47687025-47728525 forward strand	**0.00465488**	down	2.0415895
Lgi3	10 days neonate medulla oblongata cDNA, RIKEN full-length enriched library, clone:B830040F02 product:leucine-rich repeat LGI family, member 3, full insert sequence, [AK163664]	**3.07E-06**	down	2.016158

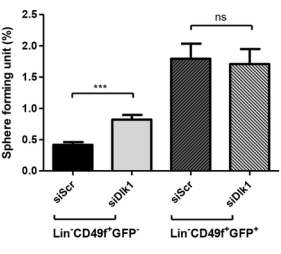
C



D



E



F

	600 cells	300 cells	100 cells	Tumor initiating cell frequency (95% CI)
Lin$^-$CD49f$^+$GFP$^-$ siScr	13/28	7/28	5/20	**1/859** (1/1283 – 1/575)
Lin$^-$CD49f$^+$GFP$^-$ siDlk1	20/28	21/28	5/20	**1/340** (1/467 – 1/248)

P=2.09^{e-04}