Stem Cell Reports
Article

ISSCR

OPEN ACCESS

HN1L Promotes Triple-Negative Breast Cancer Stem Cells through *LEPR-STAT3* Pathway

Yi Liu,[1,11] Dong Soon Choi,[1,11] Jianting Sheng,[2] Joe E. Ensor,[1] Diana Hwang Liang,[3] Cristian Rodriguez-Aguayo,[4] Amanda Polley,[5] Steve Benz,[5] Olivier Elemento,[6] Akanksha Verma,[6] Yang Cong,[3] Helen Wong,[1] Wei Qian,[1] Zheng Li,[7] Sergio Granados-Principal,[1] Gabriel Lopez-Berestein,[4,9] Melissa D. Landis,[1] Roberto R. Rosato,[1] Bhuvanesh Dave,[1] Stephen Wong,[2] Dario Marchetti,[8] Anil K. Sood,[4,9,10] and Jenny C. Chang[1,*]

[1]Houston Methodist Cancer Center, Houston Methodist Hospital, 6445 Main Street, Floor 24, Houston, TX 77030, USA
[2]Department of Systems Medicine and Bioengineering, Houston Methodist Research Institute, Houston, TX 77030, USA
[3]Department of Surgery, Houston Methodist Hospital, Houston, TX 77030, USA
[4]Center for RNA Interference and Non-Coding RNA, University of Texas, M.D. Anderson Cancer Center, Houston, TX 77030, USA
[5]NantOmics, LLC, Santa Cruz, CA 95060, USA
[6]Department of Physiology and Biophysics, Institute for Computational Biomedicine, Weill Cornell Medical College, New York, NY 10021, USA
[7]Department of Nanomedicine
[8]Biomarker Research Program
Houston Methodist Research Institute, Houston, TX 77030, USA
[9]Department of Experimental Therapeutics
[10]Department of Gynecologic Oncology
University of Texas, M.D. Anderson Cancer Center, Houston, TX 77030, USA
[11]Co-first author
*Correspondence: jcchang@houstonmethodist.org
https://doi.org/10.1016/j.stemcr.2017.11.010

SUMMARY

Here, we show that *HEMATOLOGICAL AND NEUROLOGICAL EXPRESSED 1-LIKE* (*HN1L*) is a targetable breast cancer stem cell (BCSC) gene that is altered in 25% of whole breast cancer and significantly correlated with shorter overall or relapse-free survival in triple-negative breast cancer (TNBC) patients. *HN1L* silencing reduced the population of BCSCs, inhibited tumor initiation, resensitized chemoresistant tumors to docetaxel, and hindered cancer progression in multiple TNBC cell line-derived xenografts. Additionally, gene signatures associated with *HN1L* correlated with shorter disease-free survival of TNBC patients. We defined *HN1L* as a BCSC transcription regulator for genes involved in the *LEPR-STAT3* signaling axis as HN1L binds to a putative consensus upstream sequence of *STAT3*, *LEPTIN RECEPTOR*, and *MIR-150*. Our data reveal that BCSCs in TNBC depend on the transcription regulator *HN1L* for the sustained activation of the *LEPR-STAT3* pathway, which makes it a potentially important target for both prognosis and BCSC therapy.

INTRODUCTION

Triple-negative breast cancer (TNBC), which accounts for 15%–20% of breast cancer cases, is a challenging disease to treat because of tumor heterogeneity, intrinsic resistance to conventional chemotherapy, and lack of targeted therapies (Arnedos et al., 2012; Metzger-Filho et al., 2012). Although optimization of currently available chemotherapeutic agents for TNBC patients has improved patient outcome, novel targeted therapies are urgently needed to improve patient survival (Liedtke and Rody, 2015). Since their discovery through genetic "lineage retracing" in squamous skin tumors, glioblastomas (GBM), and intestinal adenomas, cancer stem cells (CSCs) within the bulk tumor have been associated with multiple phenotypes including chemoresistance, clonogenic potential, and hierarchical tumor growth (Driessens et al., 2012; Chen et al., 2012; Schepers et al., 2012). In this study, we investigated a breast cancer stem cell (BCSC) gene signature, focusing on genes responsible for cancer stem cell self-renewal, and identified *HN1L* as a gene critical for BCSC maintenance.

We have previously published a BCSC gene signature by comparing CSCs and non-CSCs isolated from patient biopsies (Creighton et al., 2009; Dave et al., 2014). This approach narrowed down the number of genes that may influence CSC growth from a genome-wide level to only 493 genes. We then identified 13 genes to be CSC-regulating genes, by screening the effects of short hairpin RNA (shRNA) gene silencing on the mammosphere-forming ability of TNBC cell lines. We recently reported the CSC-regulating roles of *RPL39* and *MLF2*, two candidate genes among the 13 CSC genes (Dave et al., 2014).

HN1L is one of the top 13 CSCs genes whose functions have not been well characterized (Mitchell et al., 2015). Previously, *HN1L*, also known as *L11*, was reported with another 18 uncharacterized genes to be overexpressed in non-small-cell lung cancer (Petroziello et al., 2004). In the report, *L11* overexpression was identified in various cancer types, but the highest expression was found in breast and uterine cancer. While *HN1L* also contributes to embryonic development (Zhou et al., 2004), not much is known about its cellular function since it has no homology

with existing protein domain sequences (Mitchell et al., 2015). Here, we describe the essential role of *HN1L* in BCSC maintenance and explore its mechanism of action in this context. A comprehensive evaluation of *HN1L* function in patient-derived xenograft (PDX) tumors reveals that multiple pathways are regulated by *HN1L*, in which *STAT3* signaling is the key mediator. Taken as a whole, our data reveal the crucial upstream role that *HN1L* plays in promoting multiple cellular survival pathways via influence on the BCSC transcriptome.

RESULTS

High HN1L Expression Correlates with Poor Prognosis in TNBC

HN1L is one of the top candidate genes among the selected 13 genes (Table S1), whose gene silencing significantly reduced the mammosphere-forming ability of MDA-MB-231 cells (Figure S1). Thus, we decided to investigate the expression patterns of *HN1L* in TCGA breast cancer patient database with RNA sequencing (RNA-seq) information. Among the 960 breast cancer patients, approximately 25% of patients have alterations in the *HN1L* gene, consisting of mostly mRNA upregulation or gene copy number amplifications (Figure 1A). We also found that breast tumors express significantly higher levels of *HN1L* mRNA compared with normal breast tissue (Figure 1B). Interestingly, we found, using the TCGA dataset, that the overexpression of *HN1L* was correlated with shorter patient survival ($p = 0.0368$) only in TNBC (Figure 1C), while no survival difference was seen in non-TNBC (Figure S2A). We confirmed the similar survival correlation between *HN1L* expression and the shorter survival trend in TNBC patients using the Curtis patient dataset (Figures S2B and S2C). Additionally, using the previously reported TNBC cohorts (Chen et al., 2014), we confirmed that TNBC patients with high levels of *HN1L* had shorter relapse-free survival than patients with low HN1L levels ($p < 0.01$) (Figure 1D). These results demonstrate that *HN1L* is upregulated in breast cancer and that its enhanced expression correlates with poor clinical outcomes in TNBC patients. Moreover, we recently identified two missense mutations in *HN1L* (P20L and A106V; Table S2) from previously published RNA-seq data (Dave et al., 2014). Based on these findings, we recognized *HN1L* as a potential therapeutic target and chose to study its role in TNBC in greater detail.

Silencing HN1L Reduces the CSC Population in TNBC Cell Lines

We then confirmed that protein levels of *HN1L* were higher in CSC-enriched populations, compared with non-CSCs. $CD44^+/CD24^{-/low}$ population in sorted SUM159 cells only, but not in MDA-MB-231, as more than 90% cells in the MDA-MB-231 cell line are $CD44^+/CD24^{-/low}$ (Figure 2A). Also, TNBC cell lines with stable *HN1L* gene silencing showed reduced $CD44^+/CD24^{-/low}$ population and impaired mammosphere-forming ability (MSFE), implying that *HN1L* may play a significant role in maintaining CSCs (Figures 2B–2D). Knockdown of *HN1L* also impaired cell migration as analyzed by a transwell assay (Figure 2E). To better represent the 3D tumor microenvironment, we first allowed MDA-MB-231 cells to form spheroids and then embedded them in Matrigel, permitting only invasive cells to invade. Stable *HN1L* knockdown cells remained as cell aggregates without invasion into the surrounding matrix, while control cells were more invasive and formed spindle-like protrusions (Figure 2F). No invasion was observed in either control or *HN1L* gene knockdown SUM159 cells (data not shown). In addition, we further confirmed the gene silencing effect (of *HN1L*) on the $CD44^+/CD24^{-/low}$ population and MSFE of three TNBC cell lines using small interfering RNAs (siRNAs) targeting two different regions of *HN1L* (Figures S3A–S3C). These results suggest that HN1L may have an important role in multiple CSC functions, including cell migration and invasion.

Increase in the Cancer Stem Cell Population with HN1L Overexpression *In Vitro*

Because depletion of HN1L reduced CSC function, we hypothesized that *HN1L* overexpression in SUM159 and MDA-MB-468 cells would also increase CSCs. Indeed, overexpression of FLAG-tagged *HN1L* significantly increased the $CD44^+/CD24^{-/low}$ population from 66% to 87% in SUM159 cells ($p < 0.0001$) and from 2.3% to 7.6% in MDA-MB-468 cells ($p < 0.0001$), when compared with the respective control cells transfected with empty vector. The MDA-MB-231 cell line was not used because more than 90% of cells are $CD44^+/CD24^{-/low}$ (Figure 3A). The increased BCSC population, via *HN1L* overexpression, was positively correlated with increased MSFE (Figure 3B) and decreased sensitivity to docetaxel (Figure 3C). MDA-MB-468 cells were equally sensitive to docetaxel regardless of the overexpression of *HN1L*. Additionally, we observed enhanced cell migration and invasion with *HN1L* overexpression, which was subsequently decreased to control levels by the co-transfection of siRNAs against *HN1L* in three different cell lines (Figure 3D). Immunoblotting confirmed the overexpression of FLAG-HN1L and also its siRNA-mediated knockdown (Figure 3E) over the 72 hr of experimental duration. Cells transfected with the overexpression plasmid and the scrambled siRNAs served as a control for overexpressed *HN1L* at 72 hr and were

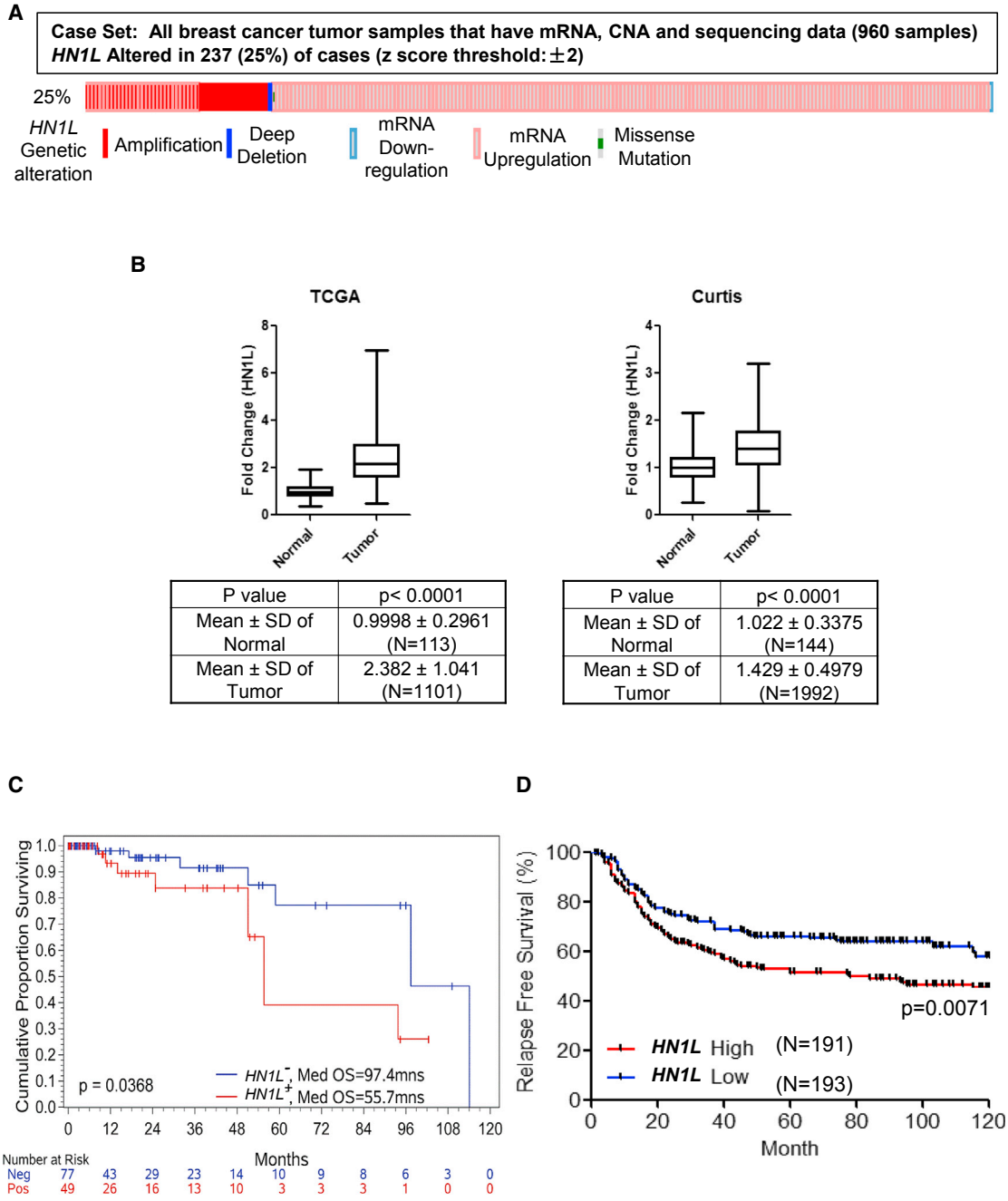


Figure 1. Enhanced Expression of *HN1L* Correlates with Poor Clinical Outcome in TNBC Patients

(A) Data acquired from cBioPortal for Cancer Genomics showed *HN1L* alteration in 960 breast cancer patients.

(B) The expression of *HN1L* mRNA is higher in breast tumors than in normal samples ($p < 0.0001$) in both TCGA and Curtis breast cancer gene expression databases.

(C and D) Higher *HN1L* expression portends significantly poorer overall survival in TCGA TNBC patients (C) and relapse-free survival of previously published TNBC patients (D) (Rody et al., 2011; Chen et al., 2014).

Error bars represent the SD.

compared with cells co-transfected with the overexpression plasmid and the silencing siRNAs for *HN1L* (Figure 3E). These results indicate that increased HN1L expression induced the $CD44^{+}/CD24^{-/low}$ population, increased MSFE and invasive potential, and generated chemoresistance in the cancer cells.

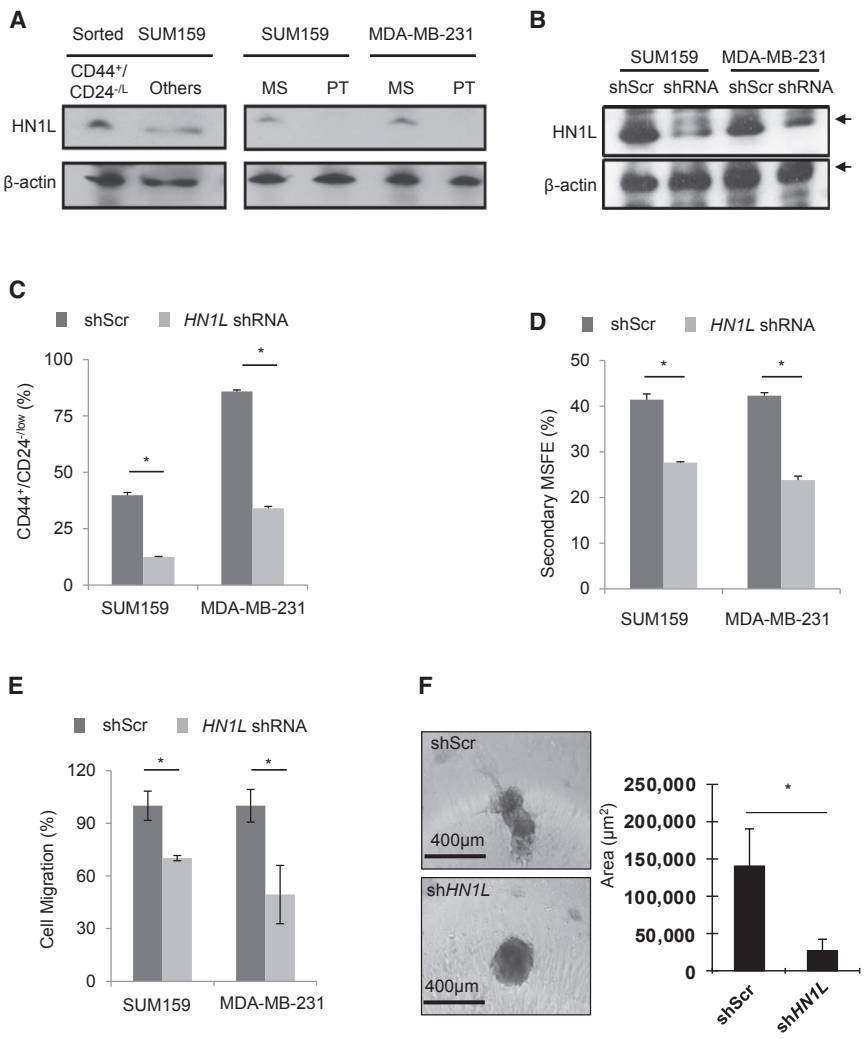


Figure 2. *HN1L* Is Overexpressed and Required for BCSCs

(A) BCSCs were enriched by either flow sorting ($CD44^+/CD24^{-/low}$) or forming mammospheres. Increased HN1L levels were detected in the BCSC-enriched population by immunoblotting. Others, all other cell populations except the $CD44^+/CD24^{-/low}$ BCSC population; MS, mammosphere; PT, prantel cells. β-Actin served as a loading control. n = 3 independent experiments.

(B) Western blot showing HN1L knockdown by shRNA in TNBC cell lines. n = 3 independent experiments.

(C) *HN1L* shRNA transduction decreased $CD44^+/CD24^{-/low}$ population upon flow-cytometry analysis. n = 3 independent experiments with 3 technical replicates; *p < 0.05.

(D) *HN1L* shRNA transduction reduced MSFE. n = 3 independent experiments with 6 technical replicates; *p < 0.05.

(E and F) *HN1L* shRNA transduction reduced cell migration (E) and 3D cancer invasion (F). n = 3 independent experiments with 3 technical replicates; *p < 0.05.

Error bars represent the SD.

HN1L Depletion Reduces the Cancer Stem Cell Population and Lung Metastasis *In Vivo*

We examined the treatment efficacy of *HN1L* siRNA in MDA-MB-231 and SUM159 TNBC cell line xenograft tumor models. Neutral 1,2-dioleoyl-sn-glycero-3-phosphatidylcholine (DOPC) nanoliposomes were used for efficient *in vivo* siRNA delivery (Tanaka et al., 2010; Tekedereli et al., 2012). We initially treated MDA-MB-231 tumor-bearing mice with *HN1L* liposomal siRNA for a short-term study of 21 days. We confirmed the *in vivo* gene silencing effects using immunohistochemistry (Figure S3D). Significant reductions in tumor volume, $CD44^+/CD24^{-/low}$ population, and secondary MSFE were observed with HN1L knockdown (Figures 4A–4C), and a statistically non-significant decrease in $ALDF^+$ cells was also apparent (Figure S3E). Similar results were observed in the SUM159 xenograft model (Figures S3F–S3H). We confirmed with a limiting dilution assay (LDA) that HN1L is critical for maintenance of BCSCs in TNBC. The gene silencing by *HN1L* siRNA-DOPC decreased the BCSC frequencies in MDA-MB-231 cells about 2- to 3-fold compared with the scrambled siRNA-DOPC as measured by these LDAs(Figure 4D).

We further evaluated the effects of silencing *HN1L* on BCSCs using a patient-derived TNBC xenograft model, BCM2665, with known drug resistance to docetaxel up to 20 mg/kg (Zhang et al., 2013; Dave et al., 2014). Unlike the cell line tumor models, *HN1L* siRNA-DOPC as a single agent did not affect tumor volume in a 21-day study (Figure 4E), although we confirmed the gene silencing effect by western blot analysis of tumors from each treatment group (Figure S4A). This result is likely due to the fact that the proportion of BCSCs in primary PDX tumors is low compared with cell lines such as MDA-MB-231. Previously, CSC in BCM2665 was defined by the ALDF and MSFE due to lack of $CD44^+/CD24^{-/low}$ subpopulation (Dave et al., 2014). Despite there being no effect on tumor growth, we observed a reduction of the $ALDF^+$ population

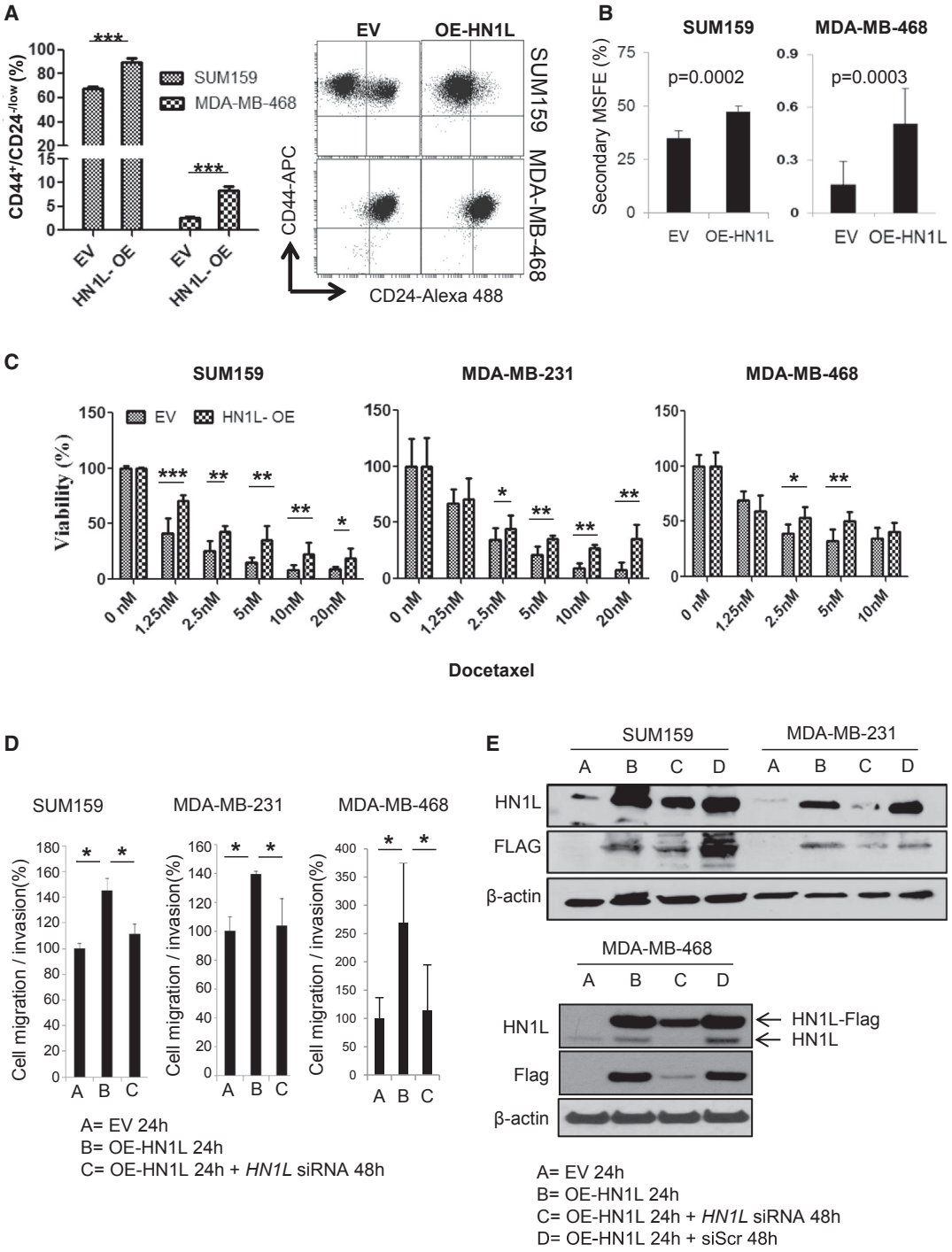


Figure 3. *HN1L* Overexpression Increases the BCSC Population *In Vitro*

TNBC cells were transfected with a plasmid to overexpress Myc-DDK tagged HN1L.

(A) Flow analysis revealed that the $CD44^{+}/CD24^{-/low}$ population increased with HN1L overexpression (OE) after 48 hr. n = 3 independent experiments with 3 technical replicates. ***p < 0.001. The data present average results of three biological repeats with SD.

(B) Secondary MSFE increased with *HN1L* overexpression. n = 3 independent experiments with 12 technical replicates.

(C) Cells with or without *HN1L* overexpression were treated with varying doses of docetaxel for 48 hr. MTT assays showed decreased sensitivity of *HN1L*-overexpressed cells to docetaxel treatment. *p < 0.05, **p < 0.01, ***p < 0.001. n = 3 independent experiments with 8 technical replicates.

(legend continued on next page)

and a significant decrease in MSFE at the end of *HN1L* siRNA treatment (Figures 4F and 4G), suggesting an effect on BCSC self-renewal function. Based on these results, we examined whether silencing *HN1L* would increase the sensitivity of drug-resistant BCM2665 to docetaxel (33.3 mg/kg) by targeting the BCSC subpopulation. As expected, the combination of *HN1L* siRNA-DOPC and docetaxel significantly inhibited tumor growth compared with docetaxel treatment alone (Figure 4H). Using an LDA, we confirmed that the synergistic effect of the combination treatment with a reduction in BCSC subpopulation. Tumors transplanted after the combination treatment showed a significant decline in BCSCs as indicated by decreased rates of tumor initiation and BCSC frequency by these assays (Figure 4I). Additionally, combination treatment with a lower docetaxel dose schedule (20 mg/kg) sensitized MDA-MB-231 cells to chemotherapy. While docetaxel alone did not significantly affect tumor growth, the combination treatment caused a statistically significant tumor growth delay (Figure S4B).

BCSCs have been found to drive metastasis (Liu et al., 2010; Liu et al., 2014). We investigated whether suppression of *HN1L* was able to impair lung metastasis in TNBC xenografts using a luciferase-tagged MDA-MB-231 cell line that develops spontaneous lung metastases from the orthotopically transplanted primary tumor (Choi et al., 2014). *HN1L* siRNA-DOPC was administered as either a single agent or in combination with docetaxel. The results indicated that delivery of *HN1L* siRNA-DOPC alone (or in combination with docetaxel) reduced tumor burden in lungs compared with scrambled siRNA (or scrambled siRNA with docetaxel) and significant reduction was seen only for the combination treatment (Figure S4C). These results suggest that *HN1L* inhibition not only decreased the BCSC population and increased chemosensitivity, but also reduced metastasis from primary tumors to lungs *in vivo*. Altogether, these data indicate that *HN1L* is a promising target for patients with TNBC for whom chemotherapy is the only option.

Inhibition of HN1L Is Associated with Better TNBC Patient Prognosis

Since HN1L depletion reduces cancer burden in our models, we sought to determine genes that may be influenced by the absence of HN1L. The corresponding gene signature could then be correlated with patient survival to determine whether the altered gene set can predict prognosis. To explore gene expression profiles when *HN1L* expression is silenced, we performed microarray analysis on BCM2665 tumor samples from scrambled siRNA- and *HN1L* siRNA-treated xenografts (n = 10). This analysis identified 74 upregulated and 62 downregulated genes under *HN1L* knockdown conditions (fold change > 1.5, $p < 0.05$) (Figure 5A). This altered gene set was defined as the *HN1L* knockdown gene signature. On applying this gene signature to the 383-patient cohort, we found that patients with an activated *HN1L* knockdown gene signature demonstrated a significantly longer relapse-free survival (log-rank test, $p = 0.00583$) (Figure 5B). These findings reaffirm the effectiveness of HN1L inhibition on TNBC progression and indicate that *HN1L*-mediated signaling pathways correlate with TNBC patient survival. With further validation, this signature may be a useful prognostic marker in the future.

HN1L Regulates an LEPR-STAT3 Signaling Pathway

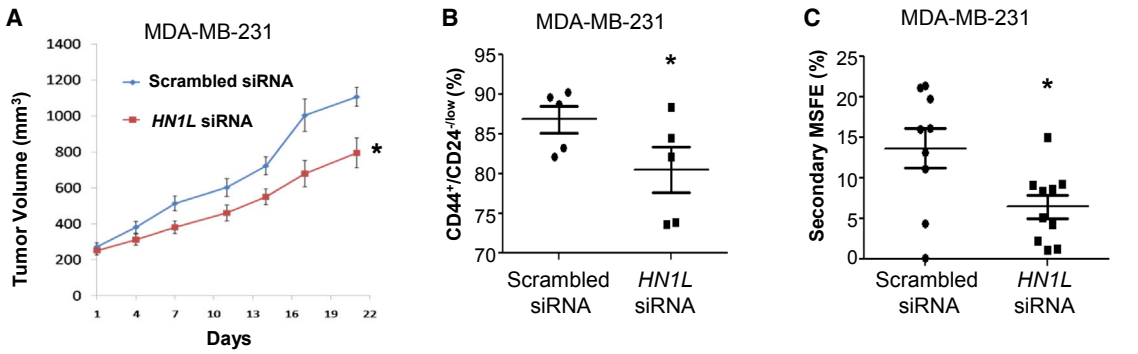
We performed a pathway enrichment analysis using Ingenuity Pathway Analysis (IPA) on the identified gene signature to identify the top signaling pathways. Using this method, the most significantly altered gene set identified was leptin signaling (Figure 5C). Leptin receptor (LEPR) is well known to be upregulated and essential in maintaining BCSCs in TNBC (Feldman et al., 2012; Zheng et al., 2013). *STAT3* signaling is a key downstream mediator of *LEPR* pathways (Ohba et al., 2010; Park and Scherer, 2011; Chang et al., 2015). We then examined *JAK-STAT* pathway activity with gene set enrichment analysis (GSEA), whereby a significant reduction of *JAK-STAT* signaling was observed in *HN1L* knockdown tumors (Figure 5D). These results led us to investigate the correlation between *HN1L*, *LEPR*, and *STAT3*. A mutual exclusivity analysis using microarray and reverse-phase protein assay (RPPA) data on 1,104 breast cancer patients from TCGA revealed that *HN1L* was highly statistically likely to act through the same signaling pathway with *LEPR* and *STAT3* (Figure 5E).

HN1L Is a BCSC Transcription Regulator of the LEPR-STAT3 Signaling Pathway

To confirm the relevance of this pathway *in vitro*, *HN1L* silencing experiments using siRNA were performed in SUM159, MDA-MB-231, and MDA-MB-468 cells. LEPR

(D) *HN1L* overexpression increased cell motility, which was inhibited by the simultaneous *HN1L* gene silencing (*$p < 0.05$). n = 3 independent experiments with 3 technical replicates.

(E) Both endogenous HN1L and overexpressed Myc-DDK(FLAG)-tagged HN1L levels were detected after overexpression and then were reduced with *HN1L* siRNA treatment by 48 hr. β-Actin served as a loading control. Cell migration/invasion in BME-coated Boyden chambers was examined. EV, empty mock plasmid; OE-*HN1L*, *HN1L* overexpression plasmid; siScr, scrambled siRNA; *HN1L* siRNA, siRNA against *HN1L*. Error bars represent the SD.



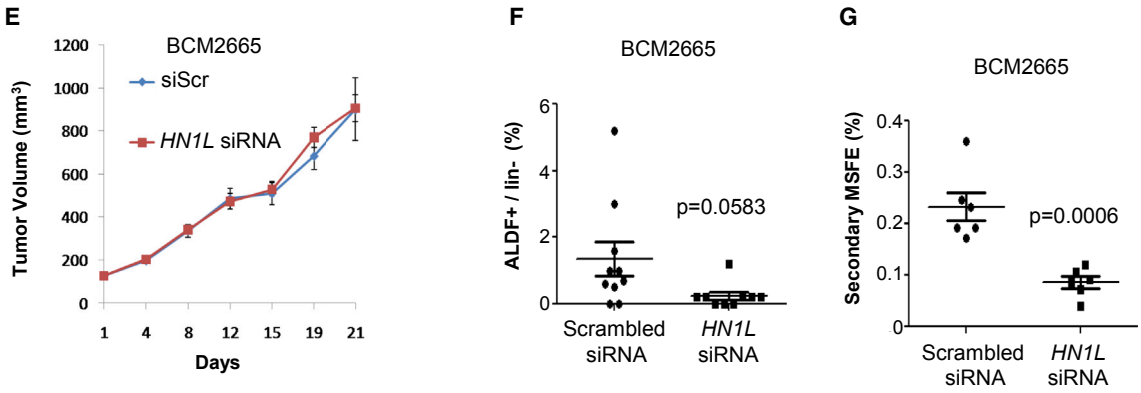
D

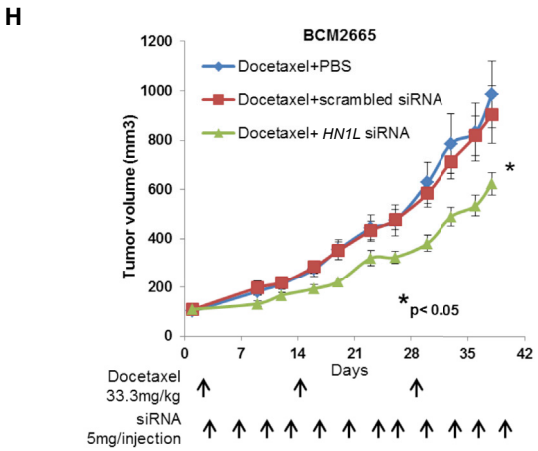
Limiting Dilution Assay in MDA-MB-231

Treatment / Cell number	Scrambled siRNA	*HN1L* siRNA
$2x10^4$	14/16	11/16
$1x10^4$	3/8	1/8
$5x10^3$	5/8	0/8

Confidence Interval of Stem Cell Frequency

	Lower	Estimate	Upper
siScr	6.101E-05	9.718E-05	1.548E-04
si*HN1L*	2.200E-05	3.863E-05	6.782E-05
p-value= 0.0123			





I

Limiting Dilution Assay in BCM2665

Treatment / Cell number	Docetaxel + siScr	Docetaxel + *HN1L* siRNA
$6x10^5$	12/12	11/12
$3x10^5$	9/12	6/12
$1x10^5$	4/12	0/12

Confidence Interval of Stem Cell Frequency

	Lower	Estimate	Upper
Docetaxel + siScr	3.214E-06	5.103E-06	8.101E-06
Docetaxel + *HN1L* siRNA	1.516E-06	2.462E-06	3.998E-06
p-value= 0.0373			

Figure 4. *HN1L* siRNA Treatment in MDA-MB-231 Xenografts and BCM2665 Patient-Derived Xenografts Decreases BCSC Function

MDA-MB-231 cells were injected, and BCM2665 tumor pieces were orthotopically transplanted into SCID-Beige mice. Mice (n = 10 per group) were injected with the respective DOPC liposomal siRNA (scrambled siRNA or HN1L siRNA) by intraperitoneal injection at 5 μg/injection twice a week (A and E). Tumor volumes were measured twice per week for 3 weeks, and tumors were collected after euthanizing animals for fluorescence-activated cell sorting analysis of $CD44^+/CD24^{-/low}$ (B) or ALDF + CSCs (F). The remaining tumor cells were used to assess mammosphere-forming ability (C and G). Another set of animals with BCM2665 tumors were randomized into three groups (n = 15):

(legend continued on next page)

and phospho-STAT3, but not leptin, were reduced at the protein level upon *HN1L* silencing (Figure 6A). Consistently, transfection with a *HN1L* overexpression plasmid for 48 hr increased the expression of phospho-STAT3, STAT3, LEPR, and leptin (Figures 6B and S5A). Also, *LEPR* depletion inhibited STAT3 activation while not affecting HN1L levels (Figure 6C), suggesting that HN1L may work upstream of LEPR. Similar results were seen *in vivo*, in which PDX tumors depleted of HN1L had decreased LEPR and STAT3 signaling (Figure S4A). Together, these data support the existence of an HN1L-LEPR-STAT3 pathway.

To determine whether HN1L influences the transcription of *STAT3* and *LEPR*, we performed qRT-PCR in MDA-MB-231 cells after silencing or overexpressing *HN1L*. We limited our investigation to *STAT3* and *LEPR* because these two genes were commonly affected by the *HN1L* gene silencing and overexpression. We found that silencing *HN1L* reduced the mRNA levels by half for both *STAT3* and *LEPR* (Figures 6D and S5B) while the overexpression caused a 2.5-fold increase in the transcription of the two genes compared with the respective controls (Figure S5C). We further confirmed the transcriptional activity of *HN1L* using a *STAT3* luciferase reporter system. *STAT3*-dependent transcription was strongly stimulated by *HN1L* overexpression in a dose-dependent manner (Figure 6E). We confirmed the regulatory roles of HN1L on the STAT3 signaling pathway through a rescue experiment by co-transfecting SUM159 cells with *HN1L* siRNAs and a plasmid carrying the constitutively active *STAT3-GFP* gene. Cells co-transfected with scrambled siRNA and *GFP* plasmid or siRNA and *STAT3-GFP* plasmid served as controls. *HN1L* gene silencing reduced both the $CD44^{+}/CD24^{-/low}$ CSCs and the MSFE in SUM159 cells while the co-transfection with *STAT3-GFP* rescued the cells from the gene silencing, as indicated by the recovered CSC population and the increased MSFE (Figures S5D and S5E). We further examined the activity of *HN1L* on *STAT3* by assessing the stem cell-regulating transcription factors, *KLF4*, *Sox2*, *Sox9*, and *Twist1*, regulated by *STAT3* in an *HN1L* knockdown stable cell line (Zhang et al., 2010; Yu et al., 2011; Guo et al., 2012; Li et al., 2012; Yang et al., 2012; Sarkar and Hochedlinger, 2013), whose levels were significantly reduced when *HN1L* levels were depleted by shRNA (Figure 6F). Furthermore, knocking down *LEPR* was sufficient to reduce TNBC cell migration and MSFE (Figures 6G and 6H). Altogether, the data support our hypothesis that HN1L is an essential regulator of BCSCs as an upstream regulator of the LEPR-STAT3 signaling pathway.

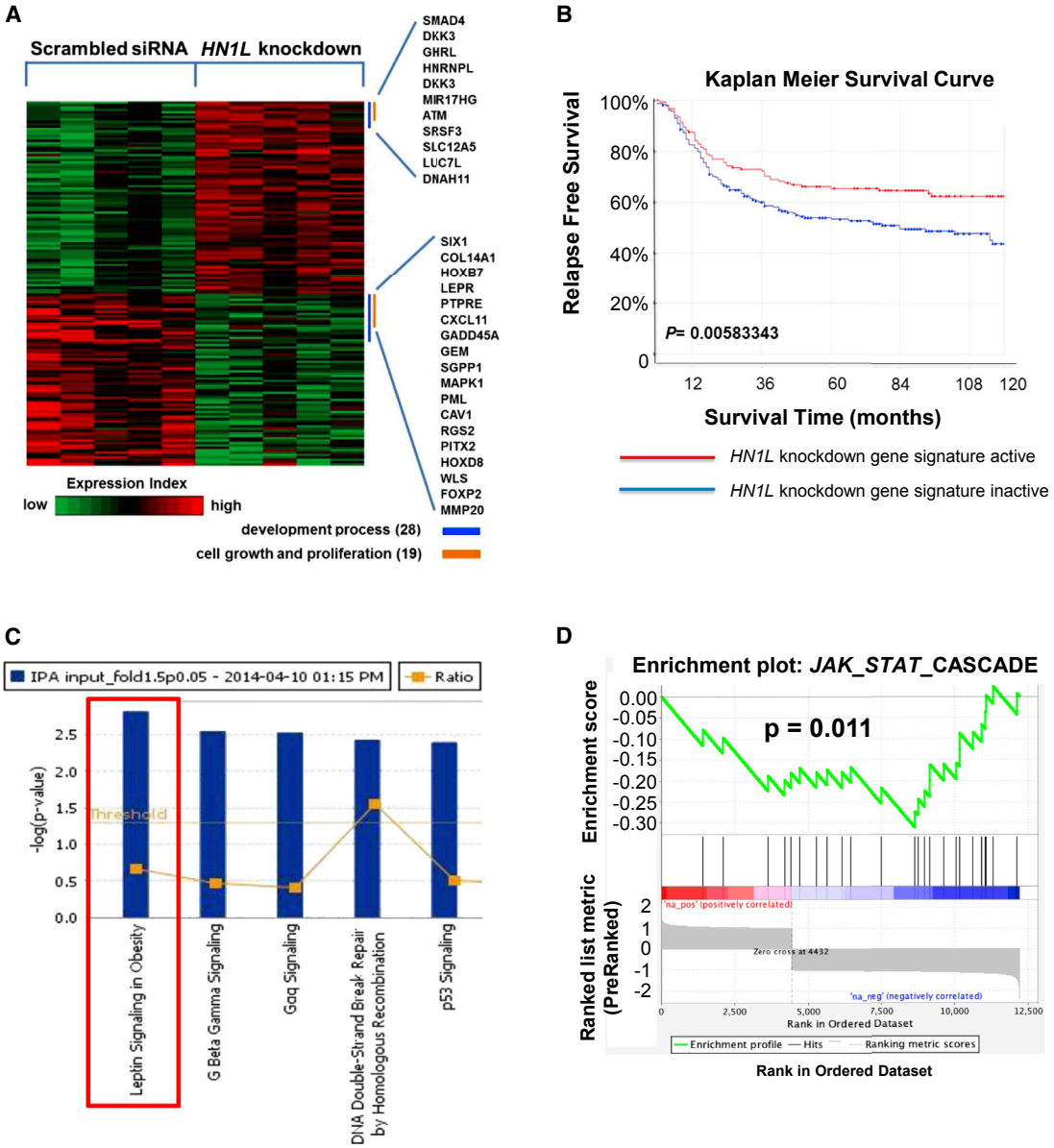
STAT3 Is the Key Mediator in HN1L Function

Since *HN1L* is a gene whose gene product has no predictable conserved domains (Mitchell et al., 2015), the mechanism of action of *HN1L* is enigmatic. The subcellular localization database, COMPARTMENTS, indicated nuclear localization of HN1L with high confidence (Binder et al., 2014). Also, nuclear localization of GFP-tagged HN1L (called L11 in their paper) has been observed (Petroziello et al., 2004). Based on this, we hypothesized that HN1L may play a regulatory role in gene expression by acting as a transcription factor or co-factor. Hence, chromatin immunoprecipitation coupled with ultra-high-throughput DNA sequencing (ChIP-seq) was performed to identify DNA sequences to which it may bind (Figure S6A). Due to the unavailability of a ChIP-grade HN1L antibody, we overexpressed FLAG-tagged HN1L in SUM159 cells and performed ChIP using anti-FLAG antibodies. HN1L showed 2,249 binding peaks from 10,000 bp upstream to 5,000 bp downstream of the transcription start site, among which 35 genes overlapped with a previously published BCSC gene signature (Creighton et al., 2009), 10 genes overlapped with the *HN1L* knockdown gene signature, and 8 genes are well established CSC transcription factors (Kim and Orkin, 2011; Ell and Kang, 2013) (Figure 7A). When applying overlapped genes in the STRING10 pathway analysis database (Szklarczyk et al., 2015), a protein interaction network centered with STAT3 was obtained (Figure 7B). *STAT3* and *FGFR2* peaks were validated by qPCR (Figure 7C). Model-based analysis of ChIP-seq (MACS) (Zhang et al., 2008; Liu, 2014) was then used to confirm the peaks found by Hypergeometric Optimization of Motif EnRichment (HOMER) (Heinz et al., 2010). Besides the binding peaks found in HOMER, another peak was called by MACS within *LEPR* (Figure S6B) and was also validated by qPCR (Figure S6C). These findings indicate that HN1L may act as a transcription factor through binding to enhancer regions of *STAT3* and other *STAT3* regulators.

Notably, a pathway analysis on the overlapped genes showed that RNA polymerase II-dependent transcription was a significant biological process involved ($p = 0.0000575$) (Figure S6D). RNA polymerase II synthesizes not only mRNA but also other RNA molecules including microRNA (miRNA) (Lee et al., 2004). This finding directed us to further investigate the role of HN1L in regulating miRNA. After comparing miRNA expression profiles from PDX BCM2665 xenograft tumors with or without *HN1L* knockdown, we found that miR-150, an

docetaxel + vehicle (PBS), docetaxel + scrambled siRNA, docetaxel + *HN1L* siRNA (H). siRNAs were delivered by intraperitoneal injection at 5 μg/injection twice a week for 6 weeks. All animals received docetaxel (33.3 mg/kg at days 1, 15, and 30). At the end of the study, tumors were collected from euthanized mice and subjected to limiting dilution assays (D and I).

Error bars in (B), (C), (F), and (G) represent the SD. Error bars in (A), (E), and (H) represent the SEM. *$p < 0.05$.



E

Mutual exclusivity Breast Invasive Carcinoma (TCGA, Provisional), All Tumors (1104 samples) /3 Genes

Gene A	Gene B	p-Value	Log Odds Ratio	Association
HN1L	***STAT3***	1.51E-05	1.781759	Tendency towards co-occurrence (**Significant**)
HN1L	***LEPR***	0.008033	1.418311	Tendency towards co-occurrence (**Significant**)
STAT3	***LEPR***	0.259881	0.771265	Tendency towards co-occurrence

Figure 5. Genes Regulated by HN1L Are Enriched for the *JAK-STAT3* Pathway

(A) Five tumor samples from scrambled siRNA-treated mice and *HN1L* siRNA-treated mice were used for microarray analysis. Genes significantly changed ($p < 0.05$, fold change > 1.5) were identified as the *HN1L* knockdown gene signature and are shown in the heatmap. (B) The *HN1L* knockdown gene signature correlated with significantly improved relapse-free survival in a cohort of 383 TNBC patients.

(legend continued on next page)

important miRNA known to be overexpressed in BCSCsm, was significantly downregulated (Figures S7A–S7C) (Huang et al., 2013; Isobe et al., 2014). To understand the role of miR-150 in the *HN1L-LEPR-STAT3* network, we analyzed all experiment-validated-targets of miR-150 (Hsu et al., 2014), together with *LEPR* and *STAT3*, in STRING10. Surprisingly, the miR-150 target genes were clustered with *STAT3* through protein-protein interactions (Figure S7D). Overall, our data indicate that HN1L is able to regulate multiple pathways including *LEPR* signaling and miR-150, which eventually converge into STAT3 to execute its function in maintaining BCSCs.

DISCUSSION

Here, we demonstrate that *HN1L* expression is increased in breast cancer and that its amplification is associated with poor patient prognosis in TNBC. Knockdown of *HN1L* in TNBC cell lines and PDX models significantly reduced BCSC population, lung metastasis, and tumor volume when combined with chemotherapy. The gene signature obtained from patient tumors with *HN1L* knockdown correlated with better survival in TNBC patients. In particular, patients with copy number gains of *HN1L* have a higher rate of relapse and metastasis and lower survival rate. Mechanistically, HN1L works through multiple pathways regulating downstream *STAT3* signaling, which is an established BCSC regulatory pathway (Marotta et al., 2011; Dave et al., 2012; Carpenter and Lo, 2014; Yu et al., 2014).

Adipose tissue is a major component of breast and is an active endocrine organ secreting cytokines such as leptin to influence the mammary gland microenvironment (Park and Scherer, 2011). The important role of LEPR in breast cancer has been addressed in multiple studies (Ishikawa et al., 2004; Rene Gonzalez et al., 2009), not only in supporting tumor cell proliferation (Yin et al., 2004) and angiogenesis (Rene Gonzalez et al., 2009), but also in maintaining CSC self-renewal properties (Park and Scherer, 2011; Feldman et al., 2012; Zheng et al., 2013). Our data showed that the LEPR pathway is the top canonical pathway associated with *HN1L* knockdown, and depleting LEPR levels sufficiently reduced the BCSC population. Although *LEPR* is a promising drug target, drugs that block LEPR are currently unavailable. As HN1L is an important regulator of *LEPR*, targeting *HN1L* might provide another angle to impair overactive *LEPR* pathways in cancer.

HN1L has multiple cellular functions, including binding to DNA sequences of many BCSC-related genes. Although the details of how HN1L works as a transcription regulator are still unclear, the important role of HN1L in BCSC function necessitates further study on its specific mechanisms. It is noteworthy that among the HN1L-associated genes detected by ChIP-seq, *FGFR2* has been reported to be indispensable for BCSC self-renewal (Kim et al., 2013), suggesting another possible mechanism of HN1L through regulation of *FGFR2* expression.

A common downstream pathway activated by LEPR, miR-150, and FGFR2 is STAT3 (Fletcher et al., 2013), which has been confirmed here. Although the importance of STAT3 signaling in BCSCs has been established, and the development of *STAT3* inhibitors is being pursued, no STAT3 antagonist is available to-date. Therefore, our findings are crucial since we have identified HN1L as an upstream regulator of *STAT3* and *LEPR* signaling. Therapies targeting HN1L have the potential to interfere with multiple downstream effectors and exert more potent effects on this signaling cascade. Also, upstream inhibition may avoid the activation of compensatory signaling from other downstream effectors causing drug resistance (Toyoshima et al., 2012). Moreover, *STAT3* activation by HN1L is not through a single pathway, as both transcriptional and translational levels of *STAT3* are affected by multiple mechanisms. HN1L inhibition may therefore have a more deleterious impact on *STAT3* action than targeting single pathways. These findings support *HN1L* as an important therapeutic target in TNBC.

In conclusion, our study describes the functions of the gene *HN1L* in BCSC self-renewal. Inhibition of *HN1L* reduces tumor cell proliferation, CSC self-renewal and migration, tumor growth, and metastasis and is associated with improved patient survival. We demonstrate that *HN1L* depletion deactivates *STAT3*-regulated gene networks and reduces BCSC self-renewal. Importantly, *HN1L* overexpression has also been observed in lung, colon, ovarian, pancreatic, prostate, and uterine cancer (Petroziello et al., 2004). Further studies are required to prove the clinical significance of *HN1L* as a therapeutic target.

(C) Top canonical pathways identified by IPA based on microarray analysis of siRNA-treated BCM2665 xenograft tumors. The red box indicates the most significantly altered signaling pathway.

(D) Broad Institute GSEA analysis of *JAK2-STAT3* pathway in BCM2665 xenografts treated with siRNA against *HN1L*, showing that the pathway was significantly downregulated.

(E) Mutual exclusivity analysis of *HN1L*, *LEPR*, and *STAT3* in the TCGA breast cancer database demonstrated a significant tendency toward co-occurrence for *HN1L* with *LEPR* and *STAT3*. mRNA expression (microarray) and protein/phosphoprotein level (RPPA) data profiles were selected in all 1,104 tumor samples from cBioPortal for Cancer Genomics to generate the table.

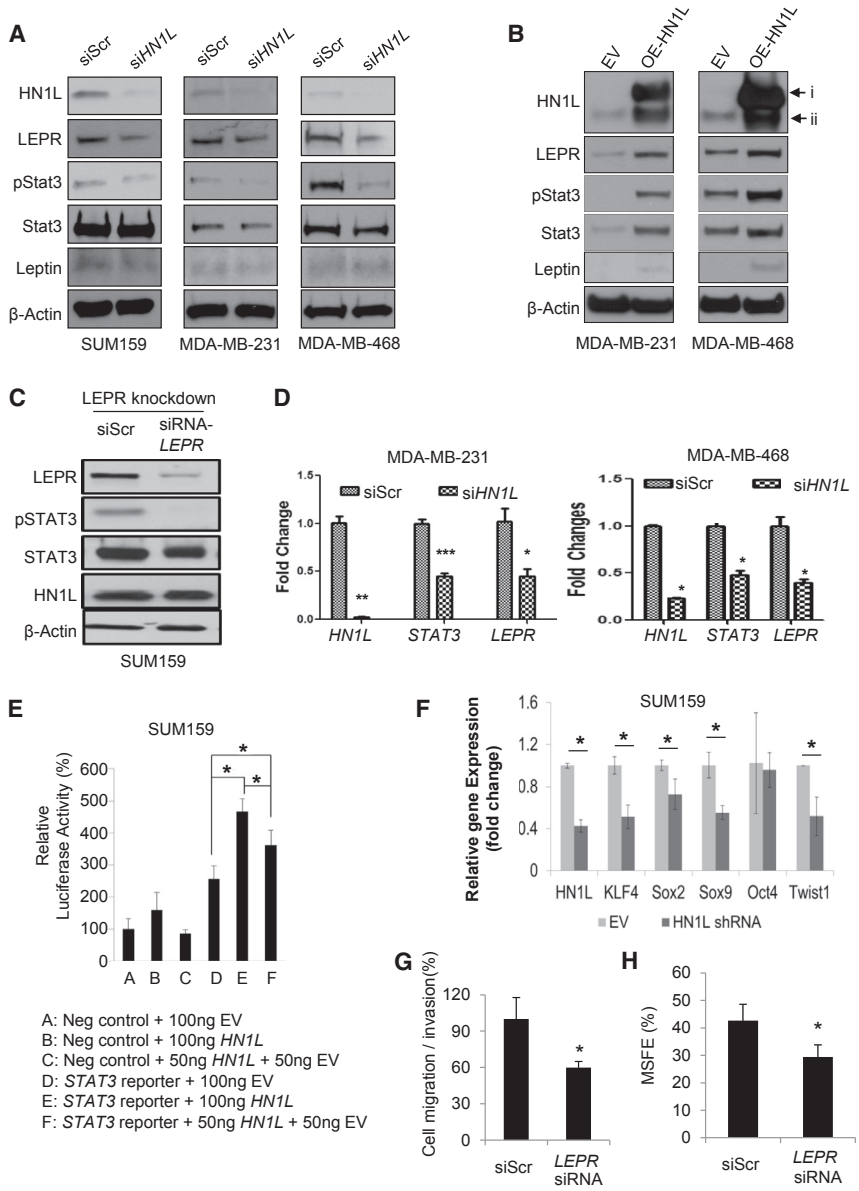


Figure 6. HN1L Activates the STAT3 Pathway by Regulating LEPR Expression

(A–C) Western blot data assessing HN1L, LEPR, and STAT3 signaling in TNBC cell lines under different treatment conditions; siScr, scrambled siRNA; siHN1L, silencing RNA against HN1L. β-Actin served as a loading control. i, HN1L-Myc-DDK; ii, endogenous HN1L.

(D) Real-time PCR results demonstrating silencing effects of *HN1L* on *STAT3* and *LEPR*. n = 3 independent experiments with 3 technical replicates.

(E) *STAT3* transcriptional activity was determined by Cignal luciferase reporter assay using *STAT3* reporter in SUM159 cells transfected with the indicated amounts of the *HN1L* overexpression plasmid. EV, empty vector. n = 3 independent experiments with 3 technical replicates.

(F) Real-time PCR on transcription factors that are downstream of *STAT3* and are essential for stem cell function was performed in *HN1L* knockdown SUM159 cells. n = 3 independent experiments with 3 technical replicates.

(G and H) Silencing *LEPR* decreased cell migration/invasion and MSFE in SUM159 cells. n = 3 independent experiments with 3 technical replicates in migration/invasion assay. n = 3 independent experiments with 12 technical replicates in MSFE assay.

Error bars represent the SD. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

EXPERIMENTAL PROCEDURES

Study Design

The objectives of this study were to elucidate unidentified functions and clinical significance of *HN1L*, as a BCSC gene, in *in vitro* and *in vivo* models. First, public breast cancer patient datasets were analyzed to compare the expression of *HN1L* gene and patient outcomes. To investigate the functions of *HN1L* in BCSCs, we utilized two cell line models and one PDX TNBC orthotopic tumor model, *in vivo* limiting dilution, flow-cytometric analysis, gene silencing or overexpression, gene expression microarrays, ChIP-seq analysis, western blot assays, and qRT-PCR. Each experiment was repeated at least two or three times using biological or technical repeats and confirmed by two or three different co-authors. Animals bearing tumors (150–200 mm^3) were randomly selected into groups, treated, and measured as indicated in Supplemental Experimental Procedures. The endpoints of animal experiments were determined prior to each experiment to mimic clinical treatment schedules. Animals with ulcerated tumors or >10% body weight loss were excluded from the randomization. We performed blinded bioinformatics or statistical analyses.

Materials and Cell Culture

TNBC cell lines (SUM159, MDA-MB-231, and MDA-MB-468) were purchased from ATCC (Manassas, VA, USA). Cells were cultured in DMEM (Invitrogen, Grand Island, NY) with 10% fetal bovine serum (Thermos Scientific Hyclone, Rockford, IL) in a humidified 5% CO_2 incubator at 37°C. Mammospheres were grown in MammoCult

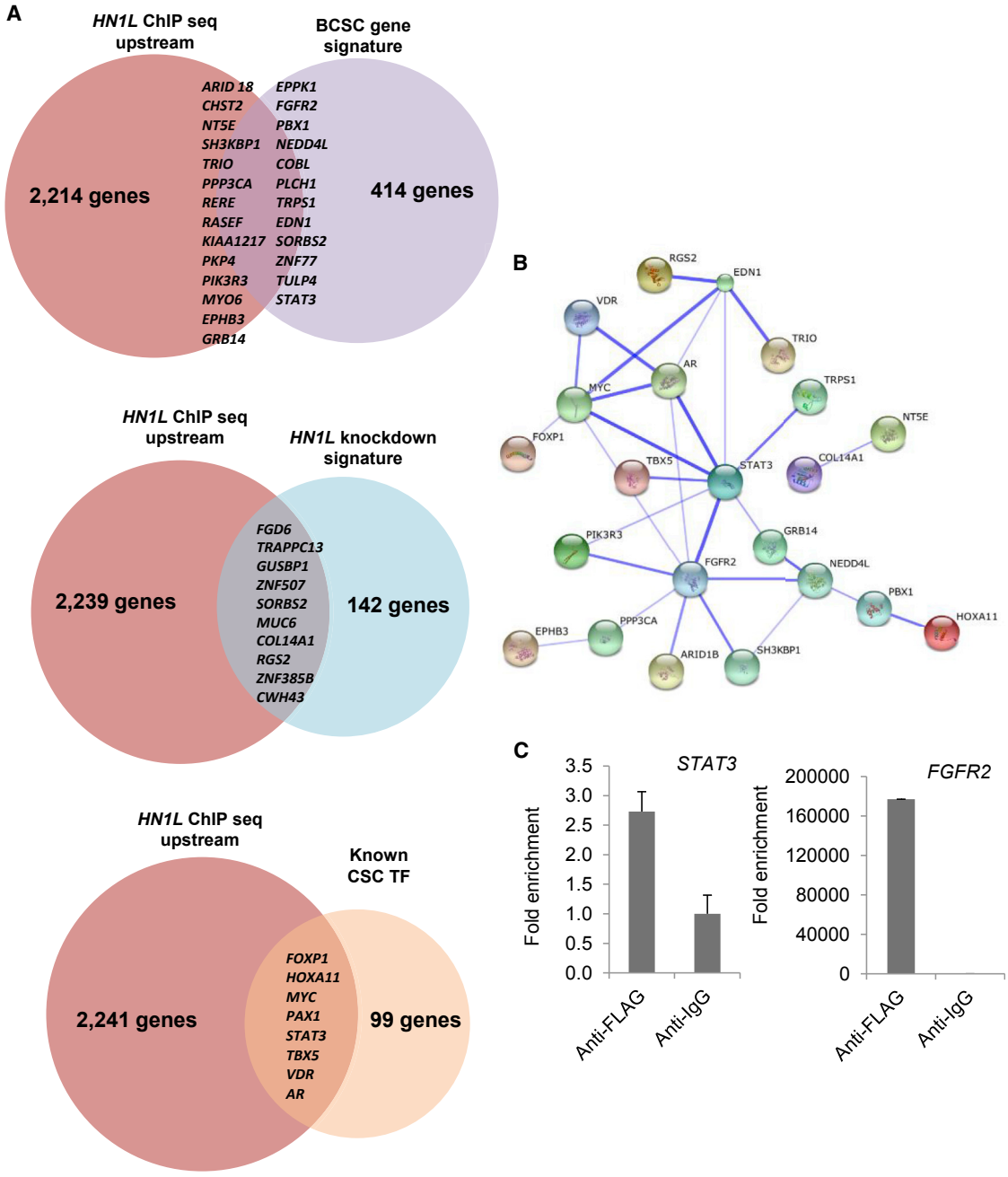


Figure 7. *STAT3* Is a Hub in HN1L Function

(A) Venn diagrams showing the overlaps between *HN1L* targets in SUM159 cells and either the BCSC gene signature, the *HN1L* knockdown gene signature, or known CSC transcription factors. Upstream is defined as *HN1L* enrichment 10 kb upstream from the open reading frames of the identified genes.

(B) Protein-protein interactions of the overlapped genes from (A) and their association with STAT3 is presented by STRING10. Thicker lines indicate the stronger confidence of associations.

(C) ChIP-seq analysis suggested that HN1L binds to upstream sequences of *STAT3* and *FGFR2*. Validation was performed by qPCR, and the y axis represents the relative abundance of DNA fragments in FLAG-ChIP (HN1L) samples over the control IgG-ChIP. n = 3 independent experiments with 3 technical replicates.

Error bars represent the SD.

Human Medium with 0.5% MethoCult, both of which were purchased from STEMCELL Technologies (Vancouver, Canada). All CD marker antibodies and mouse immunoglobulin G (IgG) isotype antibodies (APC mouse anti-human CD44, #559942; CD24 PE-CY7, #561646; mouse anti-human CD2, #555327; mouse anti-human CD31, #555446; mouse anti-human CD18, #555924; mouse anti-human CD16, #555407; mouse anti-human CD19, #555413; mouse anti-human CD45, #555483; CD140b, #558821; CD3, #555333; BD CompBead negative control, #552843) used in flow-cytometry analysis were purchased from BD Biosciences (San Jose, California). SYTOX Blue Nucleic Acid Stain for dead cell staining was purchased from Invitrogen (catalog #S34857). Antibodies used for immunoblotting, including anti-STAT3 (catalog #9132) and anti-pSTAT3-Tyr705 (catalog #9131), were purchased from Cell Signaling Technology (Danvers, MA). Anti-HN1L (catalog #HPA041908) and anti-FLAG (catalog #F3165-1MG) were obtained from Sigma (St. Louis, MO). Anti-leptin receptor antibody (catalog #AB177469) and anti-leptin antibody (catalog #AB3583) were purchased from Abcam (Cambridge, MA).

siRNA and shRNA Knockdown

The shRNA lentiviral plasmid against *HN1L* was transduced as previously described (Dave et al., 2014). All siRNAs were purchased from Ambion (Austin, TX). Sequences of shRNA and siRNAs are available in Supplemental Experimental Procedures.

Mammosphere Formation Assays

MSFE was evaluated by mammosphere assay as previously described (Choi et al., 2014; Dave et al., 2014). Mammospheres were grown in MammoCult medium with 0.5% methylcellulose. Secondary mammospheres formed in SUM159 cells were counted on day 10 and those in MDA-MB-231 cells were counted on day 14. Detailed methods for these assays are available in Supplemental Experimental Procedures.

Fluorescence-Activated Cell Sorting Analysis

$CD44^{+}/CD24^{-/low}$ population and ALDH activity in cell lines and xenograft tumors were measured by flow cytometry as described previously (Choi et al., 2014) (see Supplemental Experimental Procedures)

Cell Migration and Invasion Assays

Cell migration, cell invasion, and 3D invasion assays were carried out using Cultrex 96-well cell-migration assay (Trevigen, #3465-096-K), Cultrex 96-well BME cell-invasion assay kit (Trevigen, #3455-096-K), and Cultrex 3D spheroid cell-invasion assay (Trevigen, #3500-096-K), respectively, following the manufacturer's instructions. Further details can be found in Supplemental Experimental Procedures.

Overexpression

Cancer cells were transfected with a *HN1L* plasmid tagged with *Myc-DDK* (with *DDK* encoding a *FLAG* sequence) (Origene #RC209294) using TurboFectin transfection reagent (Origene). An empty pCMV6 vector was used as a control. One to two micrograms of plasmid was used for each transfection in 6-well plates. For a 100-mm dish, we used 20 μg of plasmid and selected cells using 3 mg/mL G418 (Invitrogen) for 3 days. For western blot assay, MDA-MB-231 and MDA-MB-468 cells were serum-starved for 24 hr before harvest.

In Vivo Studies

The Houston Methodist Hospital Research Institute Animal Care and Use Review Office approved this study. All tumor models were developed orthotopically in the mammary fat pad of SCID-Beige mice and handled as described previously (Tanaka et al., 2010; Choi et al., 2014; Dave et al., 2014). In brief, after tumors were grown orthotopically to 150–300 mm^3 in volume, mice were randomized into groups of (i) scrambled siRNA-DOPC, (ii) *HN1L* siRNA-DOPC, (iii) docetaxel + PBS, (iv) docetaxel + scrambled siRNA-DOPC, or (v) docetaxel + *HN1L* siRNA-DOPC. Groups (i), (ii), (iv), and (v) were treated with 5 μg/mouse DOPC nanoliposomal siRNA injection intraperitoneally twice a week for 3 weeks. Groups (iii) to (v) were given 20 mg/kg docetaxel injection intraperitoneally every 2 weeks with the same respective 5 μg/mouse DOPC liposomal siRNAs. DOPC nanoliposomal siRNA was prepared as previously described (Landen et al., 2005; Tanaka et al., 2010). In brief, Scrambled or *HN1L* siRNA was incorporated into DOPC at a ratio of 1:10 (w/w) in the presence of excess tertiary butanol. Liposomal siRNA powder was reconstituted with PBS immediately before *in vivo* administration. More information is available in Supplemental Experimental Procedures.

Western Blot Assay

Western blot assay was performed as previously described (Choi et al., 2014; Dave et al., 2014). In brief, all primary antibodies were diluted to 1:1,000 in 5% BSA in a washing buffer (PBS with 1% Tween 20 [PBST]), and the secondary antibodies conjugated with horseradish peroxidase to 1:3,000 in 5% skim milk in PBST. Each protein was detected using Western Blotting Luminol Reagent (Santa Cruz Biotechnology). More information is available in Supplemental Experimental Procedures.

Immunohistochemistry

IHC staining for HN1L target engagement *in vivo* was performed using a protocol similar to one previously published (Dave et al., 2014). In each treatment group, 10 FFPE tumor samples were stained, only 3 of which were randomly chosen to show images due to space limitations.

Gene Expression Microarray Analysis

Microarrays were performed on 10 snap-frozen tumor samples from scrambled siRNA-treated and *HN1L* siRNA-treated BCM2665 xenografts (5 samples from each treatment group), as described in Supplemental Experimental Procedures.

HN1L Expression and Patient Survival Analysis

We downloaded level-3 RNA-seq V2 data from the TCGA and Curtis dataset for *HN1L* from Oncomine (Rhodes et al., 2004). We normalized the gene expression levels of *HN1L* to the normal breast. Overall survival was estimated for each *HN1L* expression cohort (high versus low) within each database (TCGA and Curtis)

by means of the Kaplan-Meier method. Survival was truncated at 10 years. Cohort survival differences were assessed by means of the log-rank test. A low expresser were defined as *HN1L* < 1.85 and *HN1L* < 2.0 for the TCGA and the Curtis data, respectively. Additionally, relapse-free survival analysis either with *HN1L* expression or with the *HN1L* knockdown gene signature was performed according to a previously published method and in the same 383-patient TNBC cohort (Chen et al., 2014).

ChIP and ChIP-Seq

FLAG-tagged HN1L was overexpressed in SUM159 cells. ChIP was carried out using the EZ-ChIP kit (EMD Millipore, #17-371) following the manufacturer's protocol. Further details are available in Supplemental Experimental Procedures.

Real-Time PCR

Real-time PCR analysis of *HN1L*, *STAT3*, *LEPR*, *LEP*, and *STAT3* downstream regulators was performed using TaqMan Gene Expression Assays. Validation of ChIP peaks was done in triplicate with the SYBR Green method (Life Technologies, #4309155). Primer sequences are available in Supplemental Experimental Procedures. Validation of miR-150 in BCM2665 tumors was performed with TaqMan MicroRNA assays following the manufacturer's instructions (primer #000473).

Statistics

All *in vitro* data are presented as means ± SD and statistical significance was analyzed by two-tailed Student's t test. All experiments had at least 3 replicates and were repeated twice. *In vivo* data are presented as mean ± SEM. Different treatment groups were compared by one-way ANOVA. Patient survival was derived by the Kaplan-Meier method and the log-rank test was used to assess the statistical significance of survival between groups. Cox regression was also used to assess the impact of gene expression values on survival probabilities when analyzing the *HN1L* knockdown gene signature. A p value of less than 0.05 was considered statistically significant.

ACCESSION NUMBERS

The GEO SuperSeries accession number for this microarray and microRNA array data is GEO: GSE106200. The GEO accession number for this ChIP-seq data is GEO: GSE105446.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and three tables and can be found with this article online at https://doi.org/10.1016/j.stemcr.2017.11.010.

AUTHOR CONTRIBUTIONS

Conceptualization, Y.L., D.S.C., and J.C.C.; Methodology, Y.L., D.S.C., M.D.L., and J.C.C.; Investigation, Y.L., D.S.C., J.E.E., J.S., B.D., S.G-P., C.R.-A., H.W., W.Q., and Z.L.; Formal Analysis, Y.L., D.S.C., J.E.E., J.S., S.B., A.P., O.E., A.V., Y.C., and S.W.; Resources, D.S.C., B.D., M.D.L., A.K.S., S.W., G.L.-B., and J.C.C.; Writing – Original Draft, Y.L. and D.S.C.; Writing – Review & Editing, D.S.C., D.H.L., J.E.E., J.S., R.R.R., B.D., D.M., and J.C.C.; Supervision, J.C.C.; Funding Acquisition, J.C.C.

ACKNOWLEDGMENTS

We thank the Flow Cytometry Core of Houston Methodist Research Institute for assisting with flow-cytometry analysis. We also thank the Epigenomics Core of Weill Cornell Medical College for assisting with sequencing of ChIP samples. We thank Dr. Gary Deyter for critical reading and editing of the manuscript. This work was supported by NIH/NCI grants R01 CA138197 and U54 CA149196, the Breast Cancer Research Foundation, the Chan Soon-Shiong Institute for Advanced Health, and Golfers against Cancer.

Received: November 14, 2016
Revised: November 13, 2017
Accepted: November 14, 2017
Published: December 14, 2017

REFERENCES

Arnedos, M., Bihan, C., Delaloge, S., and Andre, F. (2012). Triple-negative breast cancer: are we making headway at least? Ther. Adv. Med. Oncol. *4*, 195–210.

Binder, J.X., Pletscher-Frankild, S., Tsafou, K., Stolte, C., O'Donoghue, S.I., Schneider, R., and Jensen, L.J. (2014). COMPARTMENTS: unification and visualization of protein subcellular localization evidence. Database (Oxford) *2014*, bau012.

Carpenter, R.L., and Lo, H.W. (2014). STAT3 target genes relevant to human cancers. Cancers (Basel) *6*, 897–925.

Chang, C.C., Wu, M.J., Yang, J.Y., Camarillo, I.G., and Chang, C.J. (2015). Leptin-STAT3-G9a signaling promotes obesity-mediated breast cancer progression. Cancer Res. *75*, 2375–2386.

Chen, J., Li, Y., Yu, T.S., McKay, R.M., Burns, D.K., Kernie, S.G., and Parada, L.F. (2012). A restricted cell population propagates glioblastoma growth after chemotherapy. Nature *488*, 522–526.

Chen, X., Iliopoulos, D., Zhang, Q., Tang, Q., Greenblatt, M.B., Hatziapostolou, M., Lim, E., Tam, W.L., Ni, M., Chen, Y., et al. (2014). XBP1 promotes triple-negative breast cancer by controlling the HIF1alpha pathway. Nature *508*, 103–107.

Choi, D.S., Blanco, E., Kim, Y.S., Rodriguez, A.A., Zhao, H., Huang, T.H., Chen, C.L., Jin, G., Landis, M.D., Burey, L.A., et al. (2014). Chloroquine eliminates cancer stem cells through deregulation of Jak2 and DNMT1. Stem Cells *32*, 2309–2323.

Creighton, C.J., Li, X., Landis, M., Dixon, J.M., Neumeister, V.M., Sjolund, A., Rimm, D.L., Wong, H., Rodriguez, A., Herschkowitz, J.I., et al. (2009). Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. Proc. Natl. Acad. Sci. USA *106*, 13820–13825.

Dave, B., Granados-Principal, S., Zhu, R., Benz, S., Rabizadeh, S., Soon-Shiong, P., Yu, K.D., Shao, Z., Li, X., Gilcrease, M., et al. (2014). Targeting RPL39 and MLF2 reduces tumor initiation and metastasis in breast cancer by inhibiting nitric oxide synthase signaling. Proc. Natl. Acad. Sci. USA *111*, 8838–8843.

Dave, B., Landis, M.D., Tweardy, D.J., Chang, J.C., Dobrolecki, L.E., Wu, M.F., Zhang, X., Westbrook, T.F., Hilsenbeck, S.G., Liu, D., and Lewis, M.T. (2012). Selective small molecule Stat3 inhibitor reduces breast cancer tumor-initiating cells and improves recurrence free survival in a human-xenograft model. PLoS One *7*, e30207.

Driessens, G., Beck, B., Caauwe, A., Simons, B.D., and Blanpain, C. (2012). Defining the mode of tumour growth by clonal analysis. Nature *488*, 527–530.

Ell, B., and Kang, Y. (2013). Transcriptional control of cancer metastasis. Trends Cell Biol. *23*, 603–611.

Feldman, D.E., Chen, C., Punj, V., Tsukamoto, H., and Machida, K. (2012). Pluripotency factor-mediated expression of the leptin receptor (OB-R) links obesity to oncogenesis through tumor-initiating stem cells. Proc. Natl. Acad. Sci. USA *109*, 829–834.

Fletcher, M.N., Castro, M.A., Wang, X., de Santiago, I., O'Reilly, M., Chin, S.F., Rueda, O.M., Caldas, C., Ponder, B.A., Markowetz, F., and Meyer, K.B. (2013). Master regulators of FGFR2 signalling and breast cancer risk. Nat. Commun. *4*, 2464.

Guo, W., Keckesova, Z., Donaher, J.L., Shibue, T., Tischler, V., Reinhardt, F., Itzkovitz, S., Noske, A., Zurrer-Hardi, U., Bell, G., et al. (2012). Slug and Sox9 cooperatively determine the mammary stem cell state. Cell *148*, 1015–1028.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Hsu, S.D., Tseng, Y.T., Shrestha, S., Lin, Y.L., Khaleel, A., Chou, C.H., Chu, C.F., Huang, H.Y., Lin, C.M., Ho, S.Y., et al. (2014). miRTarBase update 2014: an information resource for experimentally validated miRNA-target interactions. Nucleic Acids Res. *42* (Database issue), D78–D85.

Huang, S., Chen, Y., Wu, W., Ouyang, N., Chen, J., Li, H., Liu, X., Su, F., Lin, L., and Yao, Y. (2013). miR-150 promotes human breast cancer growth and malignant behavior by targeting the pro-apoptotic purinergic P2X7 receptor. PLoS One *8*, e80707.

Ishikawa, M., Kitayama, J., and Nagawa, H. (2004). Enhanced expression of leptin and leptin receptor (OB-R) in human breast cancer. Clin. Cancer Res. *10*, 4325–4331.

Isobe, T., Hisamori, S., Hogan, D.J., Zabala, M., Hendrickson, D.G., Dalerba, P., Cai, S., Scheeren, F., Kuo, A.H., Sikandar, S.S., et al. (2014). miR-142 regulates the tumorigenicity of human breast cancer stem cells through the canonical WNT signaling pathway. Elife *3*, e01977. https://doi.org/10.7554/eLife.01977.

Kim, J., and Orkin, S.H. (2011). Embryonic stem cell-specific signatures in cancer: insights into genomic regulatory networks and implications for medicine. Genome Med. *3*, 75.

Kim, S., Dubrovska, A., Salamone, R.J., Walker, J.R., Grandinetti, K.B., Bonamy, G.M., Orth, A.P., Elliott, J., Porta, D.G., Garcia-Echeverria, C., and Reddy, V.A. (2013). FGFR2 promotes breast tumorigenicity through maintenance of breast tumor-initiating cells. PLoS One *8*, e51671.

Landen, C.N., Jr., Chavez-Reyes, A., Bucana, C., Schmandt, R., Deavers, M.T., Lopez-Berestein, G., and Sood, A.K. (2005). Therapeutic EphA2 gene targeting in vivo using neutral liposomal small interfering RNA delivery. Cancer Res. *65*, 6910–6918.

Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H., and Kim, V.N. (2004). MicroRNA genes are transcribed by RNA polymerase II. EMBO J. *23*, 4051–4060.

Li, C.W., Xia, W., Huo, L., Lim, S.O., Wu, Y., Hsu, J.L., Chao, C.H., Yamaguchi, H., Yang, N.K., Ding, Q., et al. (2012). Epithelial-mesenchymal transition induced by TNF-alpha requires NF-kappaB-mediated transcriptional upregulation of Twist1. Cancer Res. *72*, 1290–1300.

Liedtke, C., and Rody, A. (2015). New treatment strategies for patients with triple-negative breast cancer. Curr. Opin. Obstet. Gynecol. *27*, 77–84.

Liu, H., Patel, M.R., Prescher, J.A., Patsialou, A., Qian, D., Lin, J., Wen, S., Chang, Y.F., Bachmann, M.H., Shimono, Y., et al. (2010). Cancer stem cells from human breast tumors are involved in spontaneous metastases in orthotopic mouse models. Proc. Natl. Acad. Sci. USA *107*, 18115–18120.

Liu, S., Cong, Y., Wang, D., Sun, Y., Deng, L., Liu, Y., Martin-Trevino, R., Shang, L., McDermott, S.P., Landis, M.D., et al. (2014). Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. Stem Cell Reports *2*, 78–91.

Liu, T. (2014). Use model-based Analysis of ChIP-Seq (MACS) to analyze short reads generated by sequencing protein-DNA interactions in embryonic stem cells. Methods Mol. Biol. *1150*, 81–95.

Marotta, L.L., Almendro, V., Marusyk, A., Shipitsin, M., Schemme, J., Walker, S.R., Bloushtain-Qimron, N., Kim, J.J., Choudhury, S.A., Maruyama, R., et al. (2011). The JAK2/STAT3 signaling pathway is required for growth of CD44(+)CD24(-) stem cell-like breast cancer cells in human tumors. J. Clin. Invest. *121*, 2723–2735.

Metzger-Filho, O., Tutt, A., de Azambuja, E., Saini, K.S., Viale, G., Loi, S., Bradbury, I., Bliss, J.M., Azim, H.A., Jr., Ellis, P., et al. (2012). Dissecting the heterogeneity of triple-negative breast cancer. J. Clin. Oncol. *30*, 1879–1887.

Mitchell, A., Chang, H.Y., Daugherty, L., Fraser, M., Hunter, S., Lopez, R., McAnulla, C., McMenamin, C., Nuka, G., Pesseat, S., et al. (2015). The InterPro protein families database: the classification resource after 15 years. Nucleic Acids Res. *43* (Database issue), D213–D221.

Ohba, S., Lanigan, T.M., and Roessler, B.J. (2010). Leptin receptor JAK2/STAT3 signaling modulates expression of Frizzled receptors in articular chondrocytes. Osteoarthritis Cartilage *18*, 1620–1629.

Park, J., and Scherer, P.E. (2011). Leptin and cancer: from cancer stem cells to metastasis. Endocr. Relat. Cancer *18*, C25–C29.

Petroziello, J., Yamane, A., Westendorf, L., Thompson, M., McDonagh, C., Cerveny, C., Law, C.L., Wahl, A., and Carter, P. (2004). Suppression subtractive hybridization and expression profiling identifies a unique set of genes overexpressed in non-small-cell lung cancer. Oncogene *23*, 7734–7745.

Rene Gonzalez, R., Watters, A., Xu, Y., Singh, U.P., Mann, D.R., Rueda, B.R., and Penichet, M.L. (2009). Leptin-signaling inhibition results in efficient anti-tumor activity in estrogen receptor positive or negative breast cancer. Breast Cancer Res. *11*, R36.

Rhodes, D.R., Yu, J., Shanker, K., Deshpande, N., Varambally, R., Ghosh, D., Barrette, T., Pandey, A., and Chinnaiyan, A.M. (2004). ONCOMINE: a cancer microarray database and integrated data-mining platform. Neoplasia *6*, 1–6.

Rody, A., Karn, T., Liedtke, C., Pusztai, L., Ruckhaeberle, E., Hanker, L., Gaetje, R., Solbach, C., Ahr, A., Metzler, D., et al. (2011). A clinically relevant gene signature in triple negative and basal-like breast cancer. Breast Cancer Res. *13*, R97.

Sarkar, A., and Hochedlinger, K. (2013). The sox family of transcription factors: versatile regulators of stem and progenitor cell fate. Cell Stem Cell *12*, 15–30.

Schepers, A.G., Snippert, H.J., Stange, D.E., van den Born, M., van Es, J.H., van de Wetering, M., and Clevers, H. (2012). Lineage tracing reveals $Lgr5^{+}$ stem cell activity in mouse intestinal adenomas. Science *337*, 730–735.

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K.P., et al. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. *43* (Database issue), D447–D452.

Tanaka, T., Mangala, L.S., Vivas-Mejia, P.E., Nieves-Alicea, R., Mann, A.P., Mora, E., Han, H.D., Shahzad, M.M., Liu, X., Bhavane, R., et al. (2010). Sustained small interfering RNA delivery by mesoporous silicon particles. Cancer Res. *70*, 3687–3696.

Tekedereli, I., Alpay, S.N., Tavares, C.D., Cobanoglu, Z.E., Kaoud, T.S., Sahin, I., Sood, A.K., Lopez-Berestein, G., Dalby, K.N., and Ozpolat, B. (2012). Targeted silencing of elongation factor 2 kinase suppresses growth and sensitizes tumors to doxorubicin in an orthotopic model of breast cancer. PLoS One *7*, e41171.

Toyoshima, Y., Kakuda, H., Fujita, K.A., Uda, S., and Kuroda, S. (2012). Sensitivity control through attenuation of signal transfer efficiency by negative regulation of cellular signalling. Nat. Commun. *3*, 743.

Yang, C.M., Chiba, T., and Groner, B. (2012). Expression of reprogramming factors in breast cancer cell lines and the regulation by activated Stat3. Horm. Mol. Biol. Clin. Investig. *10*, 241–248.

Yin, N., Wang, D., Zhang, H., Yi, X., Sun, X., Shi, B., Wu, H., Wu, G., Wang, X., and Shang, Y. (2004). Molecular mechanisms involved in the growth stimulation of breast cancer cells by leptin. Cancer Res. *64*, 5870–5875.

Yu, F., Li, J., Chen, H., Fu, J., Ray, S., Huang, S., Zheng, H., and Ai, W. (2011). Kruppel-like factor 4 (KLF4) is required for maintenance of breast cancer stem cells and for cell migration and invasion. Oncogene *30*, 2161–2172.

Yu, H., Lee, H., Herrmann, A., Buettner, R., and Jove, R. (2014). Revisiting STAT3 signalling in cancer: new and unexpected biological functions. Nat. Rev. Cancer *14*, 736–746.

Zhang, P., Andrianakos, R., Yang, Y., Liu, C., and Lu, W. (2010). Kruppel-like factor 4 (Klf4) prevents embryonic stem (ES) cell differentiation by regulating Nanog gene expression. J. Biol. Chem. *285*, 9180–9189.

Zhang, X., Claerhout, S., Prat, A., Dobrolecki, L.E., Petrovic, I., Lai, Q., Landis, M.D., Wiechmann, L., Schiff, R., Giuliano, M., et al. (2013). A renewable tissue resource of phenotypically stable, biologically and ethnically diverse, patient-derived human breast cancer xenograft models. Cancer Res. *73*, 4885–4897.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-seq (MACS). Genome Biol. *9*, R137.

Zheng, Q., Banaszak, L., Fracci, S., Basali, D., Dunlap, S.M., Hursting, S.D., Rich, J.N., Hjlemeland, A.B., Vasanji, A., Berger, N.A., et al. (2013). Leptin receptor maintains cancer stem-like properties in triple negative breast cancer cells. Endocr. Relat. Cancer *20*, 797–808.

Zhou, G., Wang, J., Zhang, Y., Zhong, C., Ni, J., Wang, L., Guo, J., Zhang, K., Yu, L., and Zhao, S. (2004). Cloning, expression and subcellular localization of HN1 and HN1L genes, as well as characterization of their orthologs, defining an evolutionarily conserved gene family. Gene *331*, 115–123.

Stem Cell Reports, Volume *10*

Supplemental Information

HN1L Promotes Triple-Negative Breast Cancer Stem Cells through *LEPR-STAT3* Pathway

Yi Liu, Dong Soon Choi, Jianting Sheng, Joe E. Ensor, Diana Hwang Liang, Cristian Rodriguez-Aguayo, Amanda Polley, Steve Benz, Olivier Elemento, Akanksha Verma, Yang Cong, Helen Wong, Wei Qian, Zheng Li, Sergio Granados-Principal, Gabriel Lopez-Berestein, Melissa D. Landis, Roberto R. Rosato, Bhuvanesh Dave, Stephen Wong, Dario Marchetti, Anil K. Sood, and Jenny C. Chang

Supplementary Materials

Table S1: List of 13 CSC genes

Gene Symbols	Gene name
RPL39	ribosomal protein L39
MLF2	myeloid leukemia factor 2
HN1L	hematological and neurological expressed 1-like
MAGI3	membrane associated guanylate kinase, WW and PDZ domain containing 3
GNAZ	G protein subunit alpha z
HMGXB3	HMG-box containing 3
ZBTB16	zinc finger and BTB domain containing 16
KIF16B	kinesin family member 16B
TRBV19	T cell receptor beta variable 19
MARVELD2	MARVEL domain containing 2
MAP7	microtubule associated protein 7
SHB	Src homology 2 domain containing adaptor protein B
PLCH1	phospholipase C eta 1

Table S2: List of mutations identified using RNA-seq in patients with lung metastasis.

Gene	AA. variant	Chr.	Base Position	refBase	altBase
HN1L	p.P20L	16	1735454	C	T
HN1L	p.A106V	16	1741967	C	T

Table S3: siRNA sequences targeting the 11 BCSC genes.

Name	Sequence (Sense)	Sequence (Antisense)
HN1L	CGCCUGUAUUUGGAAGAUUUAA	UUAAAUCUUCCAAAUACAGGCA
MAGI3	CCCUUCUGAGGUCUACCUGAAA	UUUCAGGUAGACCUCAGAAGGA
SHB	ACCUUCUUUGCUGGCUUUAUUA	UAAUAAAGCCAGCAAAGAAGGG
KIF16B	CGGCUGAGAAGUUUCAGAUAUU	AAUAUCUGAAACUUCUCAGCCU
GNAZ	CGCUAAGUGUCUUGGUAUUUAA	UUAAAUACCAAGACACUUAGCU
PLCH1	CGCUCAGUACCUGAAAGGAAUA	UAUUCCUUUCAGGUACUGAGCA
ZBTB16	ACCCUUCAGUCUCCACUUCAUU	AAUGAAGUGGAGACUGAAGGGC
MAP7	AUCUUACAUAAUGUAUUUAUAA	UUAUAAAUACAUUAUGUAAGAG
MARVELD2	AUGCUACUAUCCGUUAUUUAAU	AUUAAAUAACGGAUAGUAGCAG
TRBV19	AACCCUGAGUUGUGAACAGAAU	AUUCUGUUCACAACUCAGGGUC
HMGXB3	GCCUGUCUAUGUGGUAGAU	AUCUACCACAUAGACAGGC
Scrambled	AUCUCGCUUGGGCGAGAGUAAG	CUUACUCUCGCCCAAGCGAGAG

Supplementary Experimental Procedures

Screening BCSC candidate genes.

MDA-MB-231 cells were screened for mammosphere forming ability against 11 genes from the 13 genes (Table S1) using each specific siRNA (Table S3). Cells (200,000 cells/well) in an ultralow-attachment 24-well plate were transfected in 6 replicates using siPORT (Invitrogen) according to the manufacturer's instruction. The final concentration of each siRNA was 30nM per well. Methyl cellulose (1%) was used to for the mammosphere formation assay. For the secondary mamm0sphere formation, 5000 cells were used for each well in 6 replicates per gene. We repeated the screening twice with similar gene silencing effects on the mammosphere formation of MDA-MB-231

siRNA and shRNA knockdown.

HN1L shRNA sequences are: 5'-CGCCTGTATTTGGAAGATTTAA-3' and 3'-TTAAATCTTCCAAATACAGGCA-5'. siRNA was tranfected *in vitro* using Lipofectamine 2000 (Invitrogen) or TransIt-TKO (Mirus Bio, Madison, WI). The optimal *HN1L* siRNA sequences delivered were obtained from Ambion; *HN1L* siRNA1 :5'-CCAAGGAUCAUGUUUUCUU-3' and 3'- AAGAAAACAUGAUCCUUGG-5', HN1L siRNA2: 5'-CCUCAGAACAUACCCAAGA-3' and 3'-UCUUGGGUAUGUUCUGAGG-5'. Scrambled siRNA sequences are 5'- CGUGAACACGCAACUAAGG-3' and 3'-CCUUAGUUGCGUGUUCACG-5'. *LEPR* siRNA were also purchased from Ambion, and sequences are 5'-GAGUGAUCAUGUUAGCAAA-3' and 3'-UUUGCUAACAUGAUCACUC-5'. For *in vivo* delivery, we used *HN1L* siRNA1. For testing gene silencing effects on MSFE, cells were treated with siRNAs (25 to 50nM) for 48 hours in advance. When plated into the assay

plates, cells were treated again with 25nM of siRNAs. For the secondary mammosphere assay, no siRNA transfection was used.

Fluorescence-activated cell sorting (FACS) analysis.

The gating was performed as previously described (Li, Lewis et al. 2008, Creighton, Li et al. 2009, Choi, Blanco et al. 2014, Dave, Granados-Principal et al. 2014). Briefly, side scatter and forward scatter were used to eliminate debris and doublets, and Sytox-Blue staining was used to differentiate live and dead cells. The remaining tumor cells were further analyzed using antibodies, CD44-APC and CD24-FITC or -PE-Cy7 or by measuring aldehyde dehydrogenase activity. For In Vivo tumors, tumor cells negative for H2kD were analyzed for CSCs. Data analysis was performed with FACS Diva (BD Biosciences, San Jose, C.A., USA). $CD44^{+}/CD24^{-/low}$ (BCSC) and other cells (non-BCSC) were sorted as previously described (Creighton, Li et al. 2009). Both flow analysis and sorting were performed at the Houston Methodist Research Institute Flow Cytometry Core, using BD FACS Fortessa for analysis and BD FACS Aria II for cell sorting. All in vitro experiments were repeated three times with 3 replicates. The average values of a single experiment were shown in the figure.

Mammosphere Assay

Cells treated with either siRNA or plasmid for different durations were trypsinized, collected and counted for mammosphere assay. Mammospheres were growing in MammoCult medium with 0.5% methylcellulose. Cells were seeded at 1,000 to 5,000 cells per well in 500µl Mammocult medium in 24-well ultra-low attachment plate (Corning, Lowell, MA). Every three days, 100 µl of fresh Mammocult medium was added into the well. After 5-14 days, depending on the cells,

mammospheres formed and were counted with GelCount (Oxford Optronix, Oxford, UK) and its bundled software. For the secondary mammosphere assay, cells were collected with 0.05% trypsin for 5 minutes followed by neutralization with 10% FBS. The cells were then re-suspended in MammoCult medium and seeded at 1,000 cells per well. Secondary mammosphere formed in SUM159 cells were counted on day 10 and in MDA-MB-231 cells were counted on day 14. Mammosphere assays were repeated with 6-12 replicates for each treatment group.

Western blot

Cells treated with siRNA or plasmids with different durations were lysed in a lysis buffer (1.5% Triton X-100 and 10% glycerol in DPBS) containing a proteinase and phosphatase inhibitor cocktail (Thermo Scientific Pierce Protein Biology, Rockford, IL). 30-500 μg of protein extracts were loaded for western blot.

Cell migration and invasion assays.

50,000 cells were seeded in each well after overnight starvation in serum-free medium, and migration was measured 6 hours after adding DMEM+10%FBS as a chemoattractant. For the invasion assay, the Boyden Chamber was coated with 0.1x BME solution and invasion was measured after 24 hours. In the 3D invasion assay, 5,000 cells were plated in each well and cells invaded for 4 days in invasion matrix. Data was quantified using the Image J Program.

***In vivo* Experiments.**

MDA-MB-231 cell line tumor model: $3x10^6$ MDA-MB-231 cells injected into the mammary fat pad of SCID-Beige mice were grown to 150-300 mm^3. Then, mice were randomized into 5 groups: (*i*)scrambled siRNA-DOPC, (*ii*)*HN1L* siRNA-DOPC, (*iii*)docetaxel+PBS, (*iv*)docetaxel +scrambled siRNA-DOPC, (*v*)docetaxel+*HN1L* siRNA-DOPC. Groups *(i), (ii), (iv), and (v)* (n=10) were treated with 5 µg/mouse DOPC nanoliposomal siRNA intraperitoneal (IP) injection twice a week for 3 weeks. Mice were sacrificed on day 21. Tumors were harvested and analyzed for BCSCs using FACS, MSFE and limiting dilution assays as previously described (Schott, Landis et al. 2013). Limiting dilution assays were analyzed by Extreme Limiting Dilution Analysis (ELDA) (Hu and Smyth 2009). Groups (*iii*) - (*v*) (n=13) were given 20 mg/kg docetaxel IP injection every 2 weeks for a total of 6 weeks. Respective 5 µg/mouse DOPC-liposomal siRNA was also given twice a week for a total of 6 weeks. Treatment ended on day 42, but mice were continuously monitored without any treatment until day 58.

In order to investigate *HN1L* siRNA treatment effects on lung metastasis, $3x10^6$ MDA-MB-231 cells transfected with luciferase were injected into the mammary fat pad of each SCID-Beige mouse, and the mice were randomized into 4 groups (n=10) when the primary tumor reached 150-300mm^3: (*i*)scrambled siRNA-DOPC, (*ii*)*HN1L* siRNA-DOPC, (*iii*)docetaxel+scrambled siRNA-DOPC, (*iv*)docetaxel+*HN1L* siRNA-DOPC. Treatment schedule was the same as above. At the end of the study, mice were sacrificed and lungs were harvested and imaged as previously described (Choi, Blanco et al. 2014).

SUM159 cell line tumor model: $1x10^6$ SUM159 cells injected into the mammary fat pad of SCID-Beige mice were grown to ~200 mm^3. Then, mice were randomized into 2 groups: (*i*)scrambled siRNA-DOPC, (*ii*)*HN1L* siRNA-DOPC. Groups *(i) and (ii),* (n=9), were treated with 5 µg/mouse DOPC nanoliposomal siRNA intraperitoneal (IP) injection twice a week for 3 weeks. Mice were sacrificed on day 21. Tumors were harvested and analyzed for BCSCs using FACS, and MSFE analysis in vitro (Schott, Landis et al. 2013).

BCM2665 tumor model: In docetaxel-resistant BCM2665 PDX xenografts, tumors were transplanted into the mammary fat pad of SCID-Beige mice. Mice were randomized into 5 groups when tumor volume reached 150-200mm^3: (*i*)scrambled siRNA-DOPC, (*ii*)*HN1L* siRNA-DOPC, (*iii*)docetaxel+PBS, (*iv*)docetaxel+scrambled siRNA-DOPC, (*v*)docetaxel+*HN1L* siRNA-DOPC. Groups (*i*) and (*ii*) (n=10) were treated with 5 µg/mouse DOPC nanoliposomal siRNA IP injection twice a week for 3 weeks. Mice were sacrificed on day 21. Tumors were harvested and analyzed for BCSC using FACS and MSFE. Groups (*iii*) - (*v*) (n=15) were given 33.3 mg/kg docetaxel IP injection on day 1, 8 and 22. DOPC-liposomal siRNA (5 µg/mouse) was given twice a week for 5 weeks. At the end of the study, mice were sacrificed and tumors were harvested for limiting dilution assays as previously described (Chen, Iliopoulos et al. 2014). Tumor pieces containing $6x10^5$, $3x10^5$ or $1x10^5$ cells were transplanted with basal membrane extract into mammary glands on both sides of nude mice. Tumor incidence was reported at 4 weeks after transplantation on the counts of established tumors (equal or larger than 50mm^3). Limiting dilution assays were analyzed by Extreme Limiting Dilution Analysis (ELDA) (Hu and Smyth 2009). DOPC nanoliposomal siRNA was prepared as previously described (Landen, Chavez-Reyes et al. 2005, Tanaka, Mangala et al. 2010).

ChIP and ChIP-seq.

ChIP DNA was prepared into libraries and sequenced by the Epigenomics Core of Weill Cornell Medical College using SR50 lane. Antibodies used are anti-FLAG (Sigma, #3165) and anti-mouse IgG (EMD Millipore, #12-371). ChIP-Seq analysis began with mapping the sequenced reads to the genome. We utilized the Burrows-Wheeler Aligner (BWA) MEM algorithm to align the sequence reads against the human genome GRCh37/hg19 Assembly (Li 2013). We next used the Hypergoemetric Optimization of Motif Enrichment (HOMER) suite of tools to find and annotate peaks, and identify enriched motifs. First, we utilized HOMER's findPeaks tool to perform peak calling. Peak calling identifies the regions in the genome where a significant number of sequencing reads are found. These peaks were visualized in bigWig track file format in the UCSC genome browser. This UCSC-accepted file was created by first running HOMER's makeUCSCfile tool followed by UCSC's bedGraphToBigWig script. Next, HOMER's annotatePeaks.pl program was used to associate peaks with nearby genes. From here we compared the list of nearby genes with the HN1L and BCSC gene signatures, as well as CSC TF's. The final stage of the ChIP-Seq analysis involved using HOMER's findMotifsGenome.pl program to find enriched motifs, and ultimately, consensus sequences in the ChIP-Seq peaks. The results include a ranked list of de novo and known motifs. The "best" motifs are those with p-values significantly smaller than 1e-50. The p-value in this application is a measure of the ratio of target peaks containing the motif to background peaks containing the motif (Heinz, Benner et al. 2010). Using the chosen motif file from HOMER, we used the R program "seqLogo" to create a visually informative motif logo (Bindewald, Schneider et al. 2006). To confirm the peaks found by HOMER's findPeaks tool, and

to check for additional peaks, we used another peak-finding algorithm, the Model-based Analysis for ChIP-Seq (MACS) program, to find peaks in our sample. Following this, we used the online tool PAVIS to annotate these peaks (Zhang, Liu et al. 2008, Huang, Loganantharaj et al. 2013). The GEO accession number for this ChIP-seq data is GSE105446

Gene expression microarray analysis.

Microarrays were performed, using Affymetrix genechip U133plus 2.0. Normalization and evaluation of the data was performed as previously described (Dave, Granados-Principal et al. 2014). Differentially expressed genes were identified as the *HN1L* knockdown gene signature with cutoff *p* value less than 0.05 and fold-change greater than 1.5. Further functional and pathway analyses were done by Ingenuity Pathway Analysis (IPA) tools. Gene Set Enrichment Analysis (GSEA) was used to determine the alternation of the *JAK-STAT* pathway in *HN1L* siRNA-treated tumors compared with scrambled siRNA-treated tumors.

MicroRNA array analysis.

miRNA expression values obtained from affymetrix – miRNA array 4.0 was used to perform differential expression analysis across the 10 tumor samples. Bioinformatic analysis was performed to study changes in miRNA profiles between scrambled siRNA-treated mice and *HN1L* siRNA-treated mice tumor samples. Expression values were processed and normalized using the *affy* library in R statistical software. Only human probes were isolated, from the multispecies miRNA profiling by the miRNA array 4.0, for the purpose of the analysis. Given the 6631 human probes, supervised expression based clustering was performed to remove outlier samples present in the sample groups. Differential expression analysis was then performed using *limma* package

in R. The most significantly differentially expressed miRNAs were selected based on a cut off log fold change > 1 and FDR < 0.2. The GEO SuperSeries accession number for this microRNA array data is GSE106200.

Real-time PCR analysis.

All primers used are listed below and were designed using Primer3 and synthesized by Sigma. *STAT3* forward 5'-CCAAGATAGCGCCACTGC-3'; reverse 5'-ACATGTATCCTGTTAATTGACTTGC-3'. *FGFR2* forward 5'-TGCACTATTCACCCAACTTTCT-3'; reverse 5'-AGGAATGTGTTTGTGGCCAC-3'. All TaqMan Gene Expression Assays were purchased from Invitrogen; *STAT3* (Assay ID: Hs01051722_s1), *LEPR* (Assay ID: Hs00174492_m1), *LEP* (Assay ID: Hs00174877_m1), and *HN1L* (Assay ID: Hs00375909_m1). Eukaryotic 18S rRNA Endogenous Control (Invitrogen) was used as an internal control. Gene expression was analyzed using a standard curve for each gene as described previously (Choi, Blanco et al. 2014, Dave, Granados-Principal et al. 2014), PCR were repeated two to three times with three technical repeats. Representative results are shown in the Fig. with standard deviation.

Rescue experiment by constitutively active *STAT3*.

EF.STAT3DN.Ubc.GFP (Addgene plasmid # 24984) and EF.STAT3C.Ubc.GFP (Addgene plasmid # 24983) were gifts from Linzhao Cheng. SUM159 cells ($2x10^6$ cells) were co-transfected with these plasmids and siRNAs (50nM) using Lipofectamine 2000 (Invitrogen) according to the manufacturer's manual. After 48 hours later, cancer cells were analyzed for $CD44^+/CD24^{low/-}$ breast cancer stem cells by FACS or for MSFE as described earlier. For FACS analysis, the same

gating strategy was used as described in the Fluorescence-activated cell sorting (FACS) analysis section with an additional EGFP positive gate. The FACS data were the average values of three biological repeats. The mammosphere experiments were repeated twice with 6 replicates. The average values of one experiment were presented in the figure.

MicroRNA array analysis.

MicroRNA arrays were performed on 10 snap-frozen tumor samples from scrambled siRNA-treated and *HN1L*-siRNA treated BCM2665 xenografts (5 samples from each treatment group), as described in the Supplemental Experimental Procedures.

Supplementary Figure Legends

Fig. S1. *HN1L* is one of the top candidate genes critical for breast cancer mammosphere forming ability. MDA-MB-231 cells were transfected with each specific siRNAs (30nM) prior to the primary mammosphere assay. Scrambled siRNA was used as the control. After the primary mammosphere formation, the cancer cells (5000 cells/well) were re-seeded for the secondary mammosphere. Each MS sample included 6 replicants, and experiments were repeated 3 times independently.

Fig. S2. Non-TNBC patients show no survival correlation with the expression levels of HN1L. (A and C) The expression *HN1L* is not correlated with the overall survival of patients with non-TNBC. TNBC patients with the higher expression level of *HN1L* tend to show the shorter overall survival than those with lower *HN1L* expression (B). (A) TCGA; (B) and (C): Curtis.

Fig. S3. *HN1L* silencing has anti-CSC effects.

(A) Western blot showing *HN1L* knockdown by two different *HN1L* siRNAs in TNBC cell lines, SUM159, MDA-MB-231, and MDA-MB-468. (B) *HN1L* siRNA silencing decreased $CD44^{+}/CD24^{-/low}$ population by flow cytometry analysis in SUM159, MDA-MB-231, and MDA-MB-468 cell lines. (C) *HN1L* siRNA silencing reduced primary and secondary MSFE in SUM159 and MDA-MB-231 cell lines. Each experimental group had 6 replicates, and all experiments were repeated three times. (D) Immunohistochemistry staining of tumor samples using HN1L antibody to show target engagement. (E) A statistically non-significant decrease in ALDF+ cells was apparent in MDA-MB-231 cells with *HN1L* knockdown. (F-H) SUM159 cells were injected into SCID Beige mice and allowed tumors to grow to ~200mm3 before were randomized into different groups. These mice were injected with the respective DOPC liposomal

siRNA by i.p. injection at 5μg/injection twice a week for 3 weeks. Two treatment groups (n=10): scrambled siRNA, *HN1L* siRNA. (F) Tumor volume was measured. Mice were sacrificed 3 weeks later, and tumors were collected and processed for flow analysis of CD44+/CD24-/low cells (G), and ALDF+ cells (H). (*p<0.05, **p<0.01, ***p<0.001). All experiments were repeated three times with three technical repeats. For the purpose of publication, representative data of a repeat is presented. For the multiple comparison, Tukey's multiple comparison tests for one-way ANOVA was performed with Graphpad Prism 5.0 (Graphpad Software Inc., La Jolla, CA, USA).

Fig. S4. *HN1L* silencing effect *in vivo*. (A) Western blot data assessing HN1L, LEPR and STAT3 signaling in BCM2665 xenograft tumors under different treatment conditions; siScr: scrambled siRNA. β-actin serves as a loading control. (B) Tumor volume of mice from 3 treatment groups (n=13): docetaxel+PBS, docetaxel+scrambled siRNA, docetaxel+*HN1L* siRNA. Docetaxel was given every 2 weeks for 3 cycles. Liposomal siRNA was delivered twice a week for 6 weeks. Tumor volumes were still monitored for 2 more weeks after treatment was stopped. (C) Effects of *HN1L* knockdown on metastasis *in vivo*. *Ex vivo* imaging on lungs from each group on day 21 were presented on the upper panel when *HN1L* siRNA was used as single agent. *Ex vivo* imaging on lungs from each group on day 58 were shown on the lower panel when *HN1L* siRNA was combined with Docetaxel. (*p<0.05)

Fig. S5. HN1L regulates expression of STAT3-LEPR signaling pathway (A) Overexpression of HN1L cells upregulated phosphorylation of STAT3 and LEPR protein expression in SUM159. (B) Conversely, *HN1L* gene silencing reduced the mRNA expression of *STAT3*. (C) Overexpression of *HN1L* increases the mRNA levels of *STAT3* and *LEPR* in MDA-MB-231 and

SUM159. (D-E) Co-transfection with *HN1L* siRNA and a plasmid with constitutively active *STAT3-GFP* gene rescues the anti-CSC effects of *HN1L* silencing. For the mammosphere formation assay, each experimental group had 6 replicates, and all experiments were repeated three times. All experiments were repeated three times with three technical repeats. For the purpose of publication, representative data of a repeat is presented. For the multiple comparison, Tukey's multiple comparison tests for one-way ANOVA was performed with Graphpad Prism 5.0 (Graphpad Software Inc., La Jolla, CA, USA).

Fig. S6. ChIP peaks in *LEPR* called by MACS. (A) Identification of *HN1L* motif in HN1L overexpressed SUM159 ChIP-seq data. Anti-FLAG antibody was used for ChIP. Matrices predicted by HOMER Motif Analysis. (B)Visualized peaks shown in both input and anti-FLAG samples. The peak found by MACS and validated by QPCR (C) was indicated by the red arrow. (D) Pathway analysis by STRING 10 revealed top pathways regulated by the overlapped genes from Fig. 7.

Fig. S7. HN1L regulates *LEPR* and *miR150* pathways which converge into *STAT3*. (A) Top most differentially expressed miRNAs between scrambled siRNA-treated mice tumors and *HN1L* siRNA-treated tumors in BCM2665 xenografts (n=4 for each treatment arm). miRNA were selected based on a cut-off log fold change >1 and FDR <0.2. (B) Table showing the ID, log fold change and p-value of the top 4 most upregulated (positive value in logFC) or downregulated (negative value) miRNA. (C) Validation of reduced miR-150 in *HN1L* siRNA-treated tumors by qPCR. (D) Protein-protein interactions of known miR-150 targets with LEPR-STAT3 are presented by STRING 10. Thicker lines indicate the stronger confidence of associations.

References

Bindewald, E., T. D. Schneider and B. A. Shapiro (2006). "CorreLogo: an online server for 3D sequence logos of RNA and DNA alignments." Nucleic Acids Res **34**(Web Server issue): W405-411.

Chen, X., D. Iliopoulos, Q. Zhang, Q. Tang, M. B. Greenblatt, M. Hatziapostolou, E. Lim, W. L. Tam, M. Ni, Y. Chen, J. Mai, H. Shen, D. Z. Hu, S. Adoro, B. Hu, M. Song, C. Tan, M. D. Landis, M. Ferrari, S. J. Shin, M. Brown, J. C. Chang, X. S. Liu and L. H. Glimcher (2014). "XBP1 promotes triple-negative breast cancer by controlling the HIF1alpha pathway." Nature **508**(7494): 103-107.

Choi, D. S., E. Blanco, Y. S. Kim, A. A. Rodriguez, H. Zhao, T. H. Huang, C. L. Chen, G. Jin, M. D. Landis, L. A. Burey, W. Qian, S. M. Granados, B. Dave, H. H. Wong, M. Ferrari, S. T. Wong and J. C. Chang (2014). "Chloroquine eliminates cancer stem cells through deregulation of Jak2 and DNMT1." Stem cells **32**(9): 2309-2323.

Creighton, C. J., X. Li, M. Landis, J. M. Dixon, V. M. Neumeister, A. Sjolund, D. L. Rimm, H. Wong, A. Rodriguez, J. I. Herschkowitz, C. Fan, X. Zhang, X. He, A. Pavlick, M. C. Gutierrez, L. Renshaw, A. A. Larionov, D. Faratian, S. G. Hilsenbeck, C. M. Perou, M. T. Lewis, J. M. Rosen and J. C. Chang (2009). "Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features." Proc Natl Acad Sci U S A **106**(33): 13820-13825.

Dave, B., S. Granados-Principal, R. Zhu, S. Benz, S. Rabizadeh, P. Soon-Shiong, K. D. Yu, Z. Shao, X. Li, M. Gilcrease, Z. Lai, Y. Chen, T. H. Huang, H. Shen, X. Liu, M. Ferrari, M. Zhan, S. T. Wong, M. Kumaraswami, V. Mittal, X. Chen, S. S. Gross and J. C. Chang (2014). "Targeting RPL39 and MLF2 reduces tumor initiation and metastasis in breast cancer by inhibiting nitric oxide synthase signaling." Proc Natl Acad Sci U S A **111**(24): 8838-8843.

Heinz, S., C. Benner, N. Spann, E. Bertolino, Y. C. Lin, P. Laslo, J. X. Cheng, C. Murre, H. Singh and C. K. Glass (2010). "Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities." Mol Cell **38**(4): 576-589.

Hu, Y. and G. K. Smyth (2009). "ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays." J Immunol Methods **347**(1-2): 70-78.

Huang, W., R. Loganantharaj, B. Schroeder, D. Fargo and L. Li (2013). "PAVIS: a tool for Peak Annotation and Visualization." Bioinformatics **29**(23): 3097-3099.

Landen, C. N., Jr., A. Chavez-Reyes, C. Bucana, R. Schmandt, M. T. Deavers, G. Lopez-Berestein and A. K. Sood (2005). "Therapeutic EphA2 gene targeting in vivo using neutral liposomal small interfering RNA delivery." Cancer Res **65**(15): 6910-6918.

Li, H. (2013). "Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM." eprint arXiv:1303.3997.

Li, X., M. T. Lewis, J. Huang, C. Gutierrez, C. K. Osborne, M. F. Wu, S. G. Hilsenbeck, A. Pavlick, X. Zhang, G. C. Chamness, H. Wong, J. Rosen and J. C. Chang (2008). "Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy." J Natl Cancer Inst **100**(9): 672-679.

Schott, A. F., M. D. Landis, G. Dontu, K. A. Griffith, R. M. Layman, I. Krop, L. A. Paskett, H. Wong, L. E. Dobrolecki, M. T. Lewis, A. M. Froehlich, J. Paranilam, D. F. Hayes, M. S. Wicha and J. C. Chang (2013). "Preclinical and clinical studies of gamma secretase inhibitors with docetaxel on human breast tumors." Clin Cancer Res **19**(6): 1512-1524.

Tanaka, T., L. S. Mangala, P. E. Vivas-Mejia, R. Nieves-Alicea, A. P. Mann, E. Mora, H. D. Han, M. M. Shahzad, X. Liu, R. Bhavane, J. Gu, J. R. Fakhoury, C. Chiappini, C. Lu, K. Matsuo, B. Godin, R. L. Stone, A. M. Nick, G. Lopez-Berestein, A. K. Sood and M. Ferrari (2010). "Sustained small interfering RNA delivery by mesoporous silicon particles." Cancer Res **70**(9): 3687-3696.
Zhang, Y., T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li and X. S. Liu (2008). "Model-based analysis of ChIP-Seq (MACS)." Genome Biol **9**(9): R137.

Fig. S1

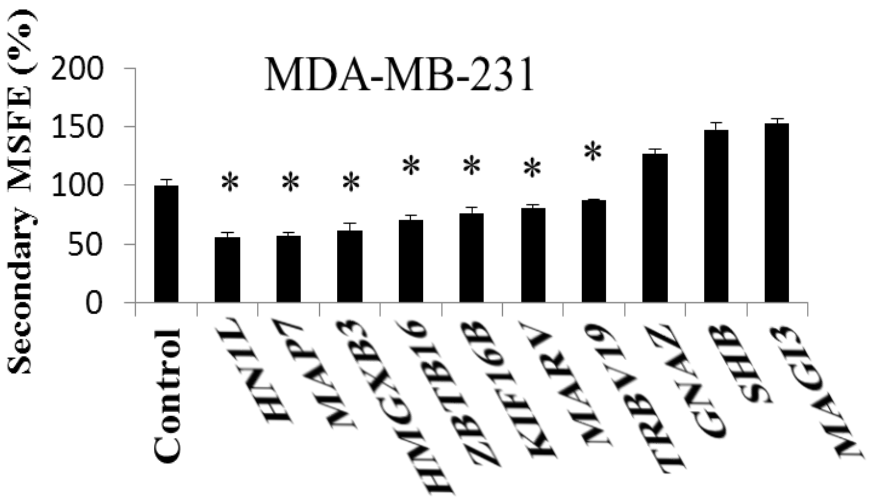
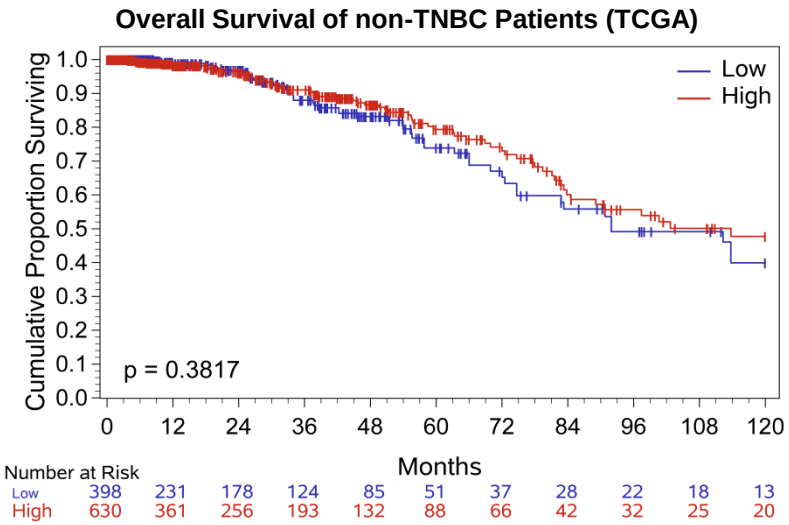
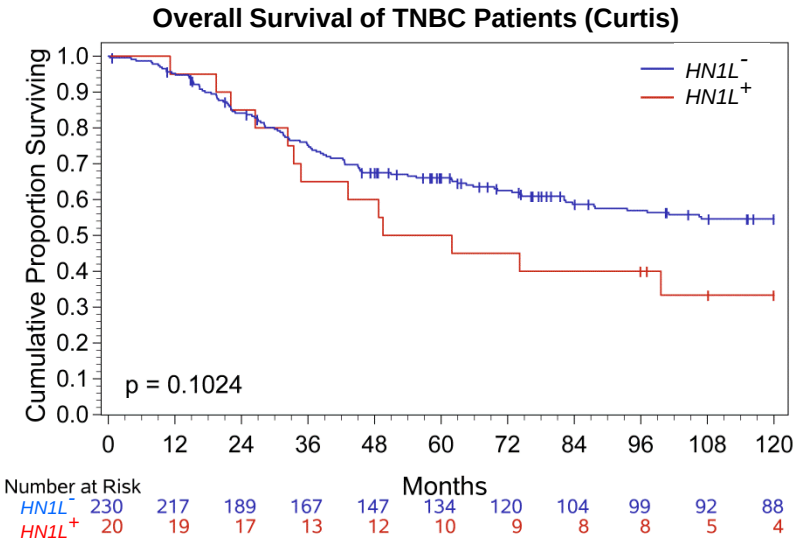

MDA-MB-231
Secondary MSFE (%)
200
150
100
50
0
*
*
*
*
*
*
*
Control
HN1L
MAP7
MAJN
HMGXB3
ZBTB16
KIF16B
MARV
TRBV19
GNAZ
SHB
MAGI3

Fig. S2

A


Overall Survival of non-TNBC Patients (TCGA)
Cumulative Proportion Surviving
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Low
High
p = 0.3817
0
12
24
36
48
60
72
84
96
108
120
Months
Number at Risk
Low 398 231 178 124 85 51 37 28 22 18 13
High 630 361 256 193 132 88 66 42 32 25 20

B


Overall Survival of TNBC Patients (Curtis)
Cumulative Proportion Surviving
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
HN1L-
HN1L+
p = 0.1024
0
12
24
36
48
60
72
84
96
108
120
Months
Number at Risk
HN1L- 230 217 189 167 147 134 120 104 99 92 88
HN1L+ 20 19 17 13 12 10 9 8 8 5 4

C

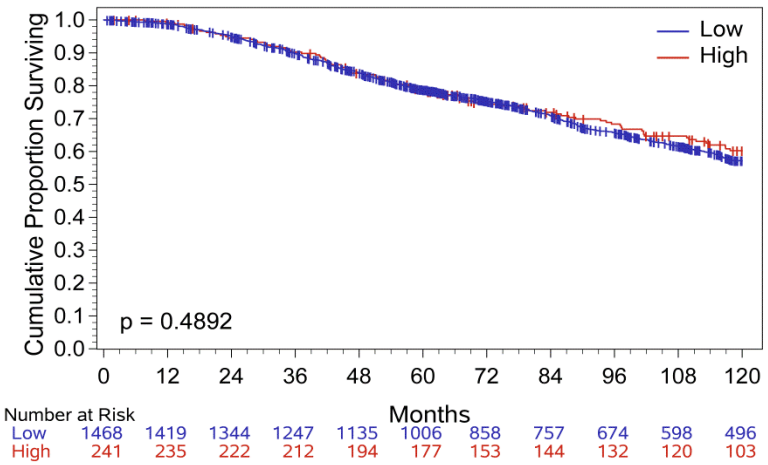

Overall Survival of non-TNBC Patients (Curtis)
Cumulative Proportion Surviving
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Low
High
p = 0.4892
0
12
24
36
48
60
72
84
96
108
120
Months
Number at Risk
Low 1468 1419 1344 1247 1135 1006 858 757 674 598 496
High 241 235 222 212 194 177 153 144 132 120 103

Fig. S3

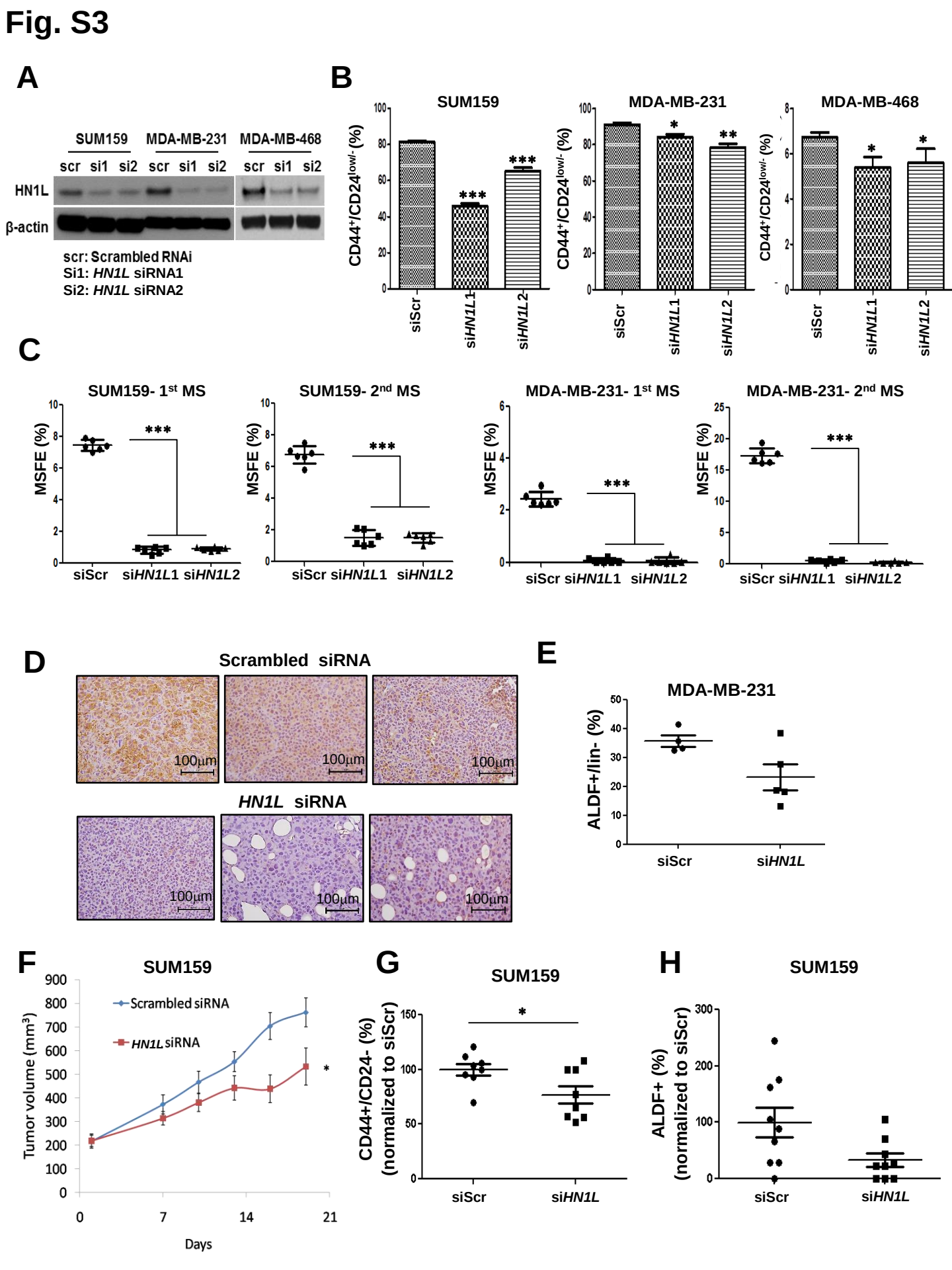

A
SUM159
MDA-MB-231
MDA-MB-468
scr si1 si2 scr si1 si2 scr si1 si2
HN1L
β-actin
scr: Scrambled RNAi
Si1: HN1L siRNA1
Si2: HN1L siRNA2
B
SUM159
MDA-MB-231
MDA-MB-468
CD44+/CD24low/- (%)
siScr
siHN1L1
siHN1L2
C
SUM159- 1st MS
SUM159- 2nd MS
MDA-MB-231- 1st MS
MDA-MB-231- 2nd MS
MSFE (%)
siScr siHN1L1 siHN1L2
D
Scrambled siRNA
HN1L siRNA
100µm
E
MDA-MB-231
ALDF+/lin- (%)
siScr
siHN1L
F
SUM159
Scrambled siRNA
HN1L siRNA
Tumor volume (mm3)
Days
G
SUM159
CD44+/CD24- (%) (normalized to siScr)
siScr
siHN1L
H
SUM159
ALDF+ (%) (normalized to siScr)
siScr
siHN1L

Fig. S4

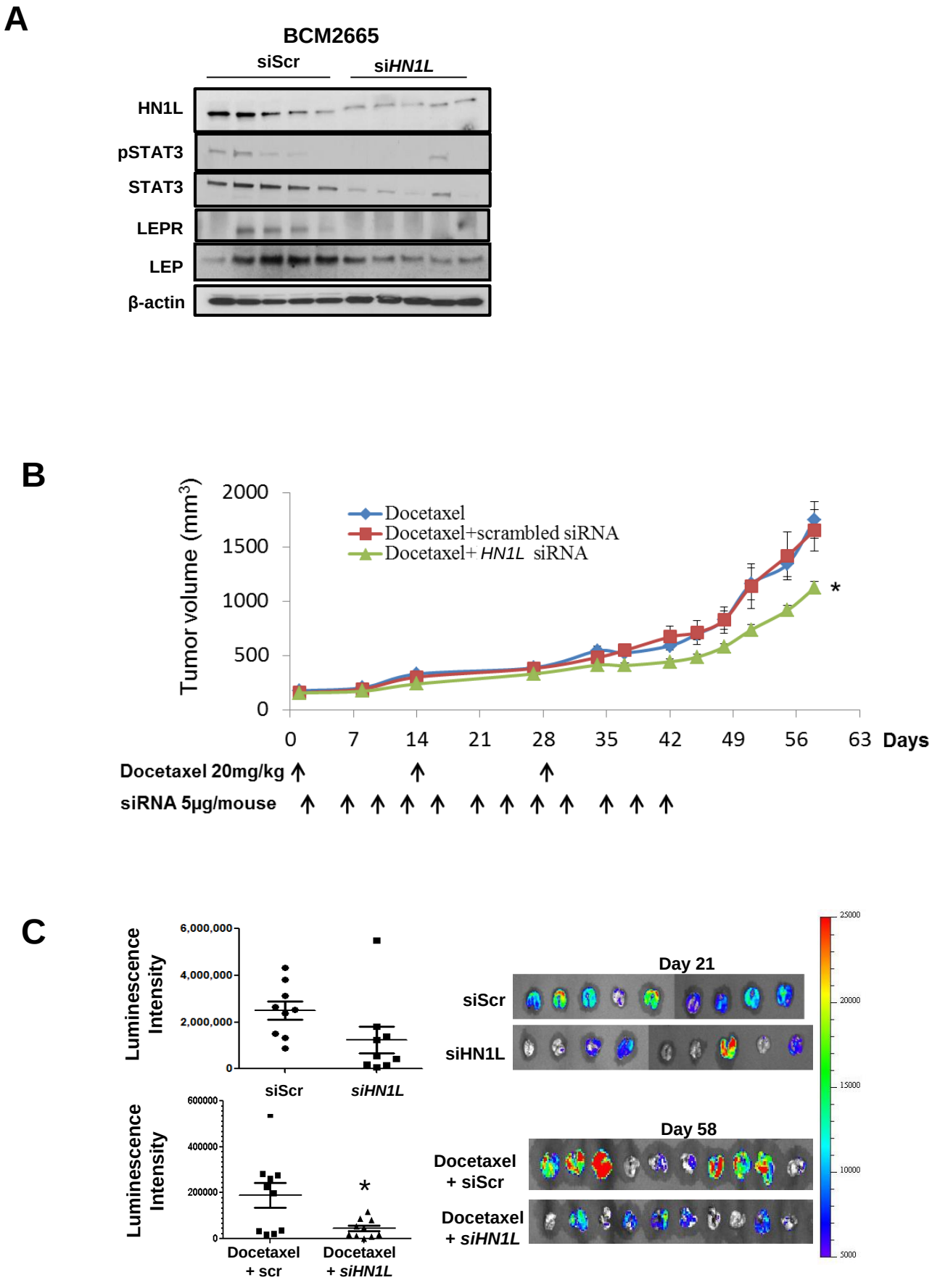

A
BCM2665
siScr
siHN1L
HN1L
pSTAT3
STAT3
LEPR
LEP
β-actin
B
Tumor volume (mm3)
2000
1500
1000
500
0
Docetaxel
Docetaxel+scrambled siRNA
Docetaxel+ HN1L siRNA
*
0
7
14
21
28
35
42
49
56
63
Days
Docetaxel 20mg/kg
siRNA 5μg/mouse
C
Luminescence Intensity
6,000,000
4,000,000
2,000,000
0
siScr
siHN1L
600000
400000
200000
0
*
Docetaxel + scr
Docetaxel + siHN1L
Day 21
siScr
siHN1L
Day 58
Docetaxel + siScr
Docetaxel + siHN1L
25000
20000
15000
10000
5000

Fig. S5

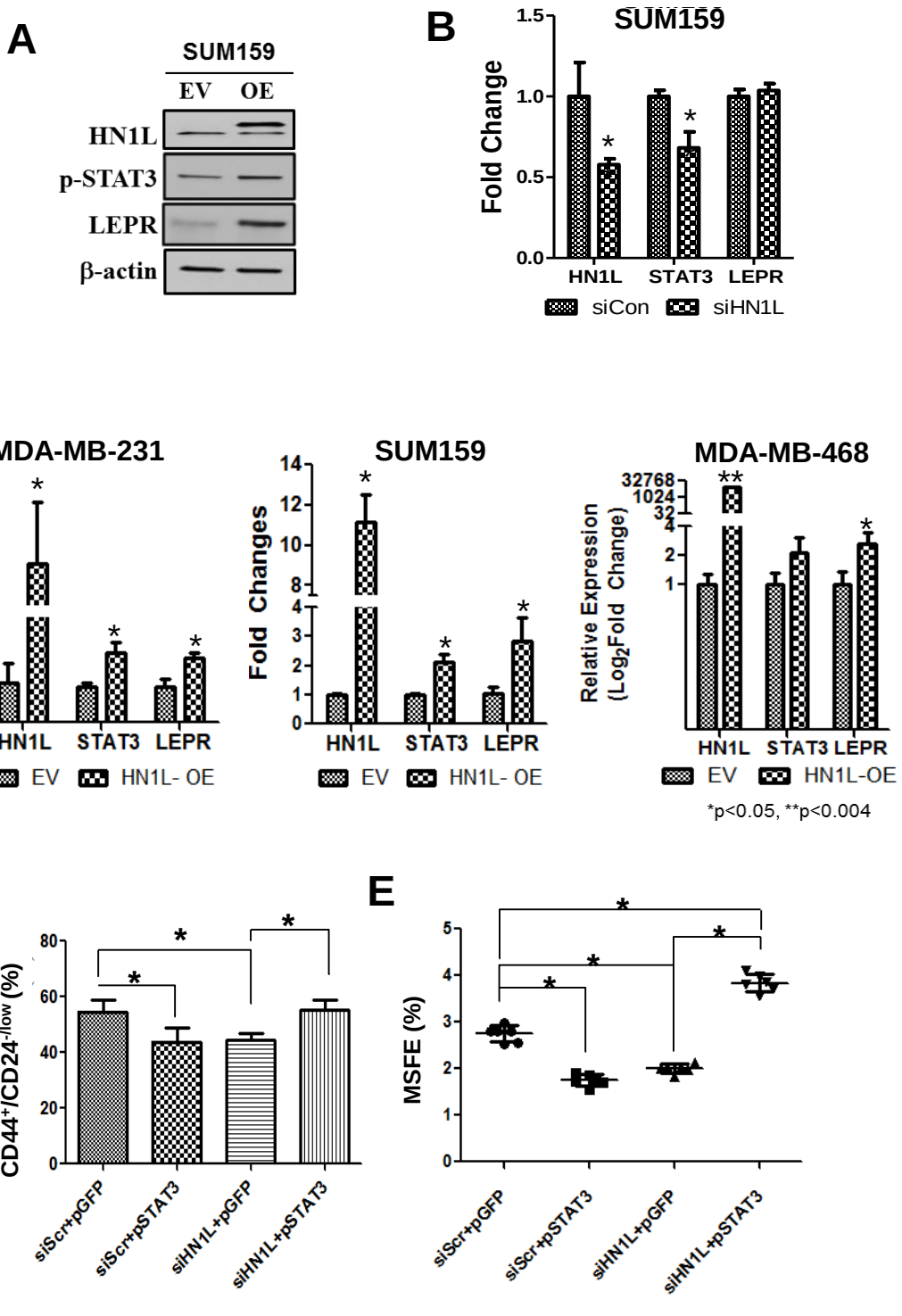
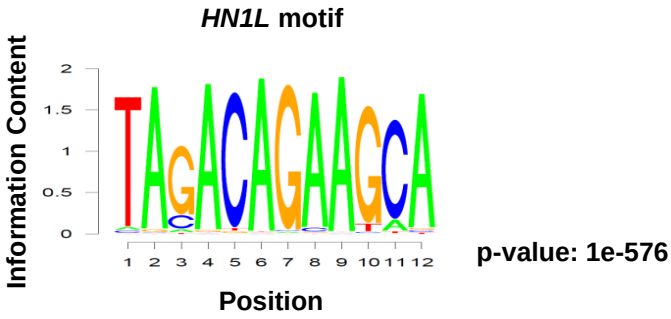
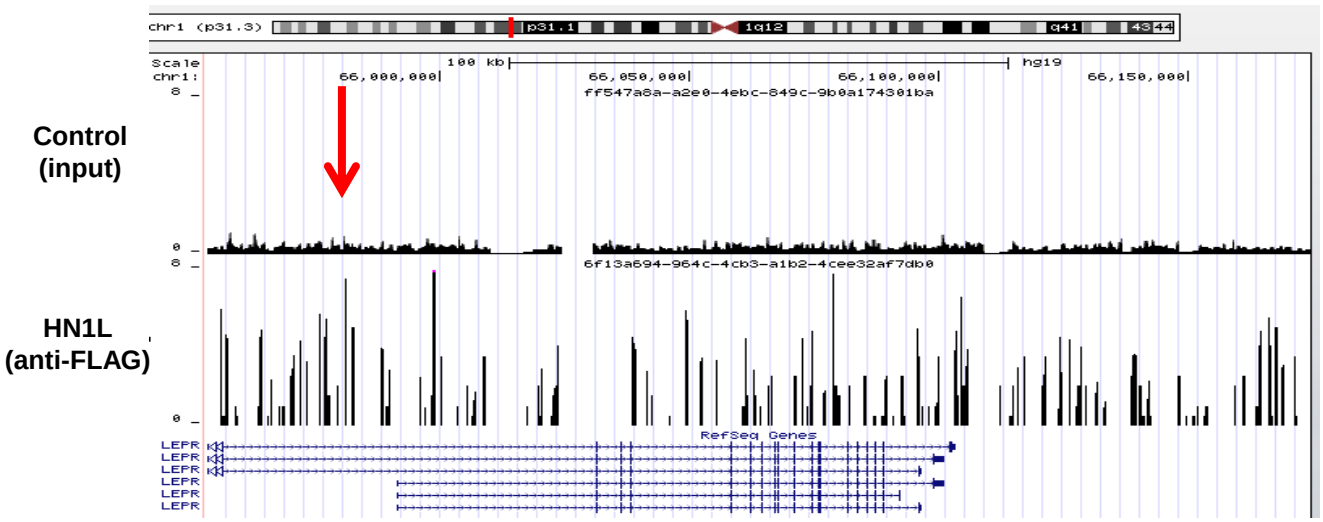

A
SUM159
EV
OE
HN1L
p-STAT3
LEPR
β-actin
B
SUM159
Fold Change
HN1L
STAT3
LEPR
siCon
siHN1L
C
MDA-MB-231
Fold Changes
HN1L
STAT3
LEPR
EV
HN1L- OE
SUM159
Fold Changes
HN1L
STAT3
LEPR
EV
HN1L- OE
MDA-MB-468
Relative Expression (Log2Fold Change)
HN1L
STAT3
LEPR
EV
HN1L-OE
*p<0.05, **p<0.004
D
CD44+/CD24-/low (%)
siScr+pGFP
siScr+pSTAT3
siHN1L+pGFP
siHN1L+pSTAT3
E
MSFE (%)
siScr+pGFP
siScr+pSTAT3
siHN1L+pGFP
siHN1L+pSTAT3

Fig. S6

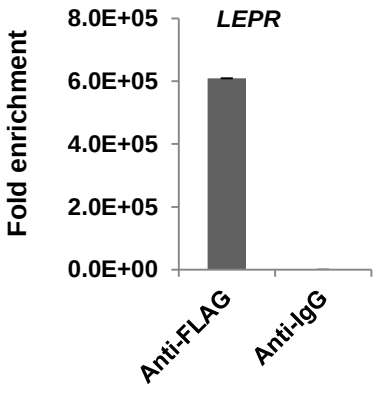
A



B



C

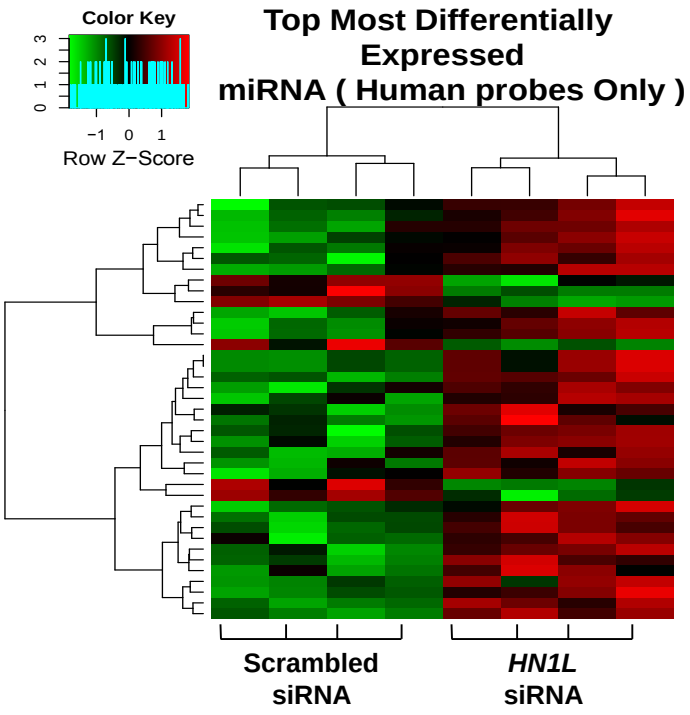


D

Pathway ID (GO)	Pathway Description	False Discovery Rate	Matching Proteins in the Network
0045893	Positive regulation of transcription, DNA-templated	5.75E-05	AR,ARID1B,EDN1,FGFR2,FOXP1, HOXA11,MYC,MYO6,PAX1,PBX1, PPP3CA,RERE, STAT3,TBX5,VDR
0045944	Positive regulation of transcription from RNA polymerase II promoter	5.75E-05	AR,EDN1,FGFR2,FOXP1,MYC,MYO6, PAX1,PBX1,PPP3CA,RERE,STAT3, TBX5,VDR
0009891	positive regulation of biosynthetic process	6.27E-05	AR,ARID1B,EDN1,FGFR2,FOXP1, HOXA11,MYC,MYO6,NT5E,PAX1,PBX1, PPP3CA, RERE, STAT3,TBX5,VDR
0022603	Regulation of anatomical structure morphogenesis	6.27E-05	EDN1,EPHB3,FGD6,FGFR2,FOXP1, HOXA11,MYC,NEDD4L,PPP3CA, SH3KBP1,TBX5,VDR
0001763	Morphogenesis of a branching structure	0.000104	EDN1,FGFR2,HOXA11,MYC,PBX1, RERE, VDR

Fig. S7

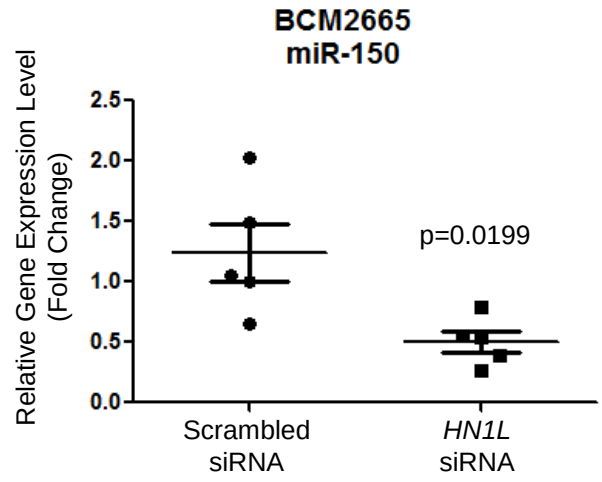
A



B

Transcript.ID. Array Design	logFC	t	p Value
hsa-miR-150-5p	-2.183121	7.984466	1.59E-05
hsa-miR-3921	-1.606919	5.161487	0.000493
hsa-miR-5195-3p	-1.173754	4.828139	0.000792
hsa-miR-3188	-1.150942	4.696959	0.000959
hsa-miR-6746-3p	1.7503	-5.45189	0.00033
hsa-miR-6504-5p	1.76714	-4.73776	0.000904
hsa-miR-2113	2.16653	-4.11002	0.002328
hsa-miR-6868-3p	2.32171	-4.73751	0.000904

C



D