

# HMGA1: A Master Regulator of Tumor Progression in Triple-Negative Breast Cancer Cells

Sandeep N. Shah<sup>1</sup>, Leslie Cope<sup>2,3</sup>, Weijie Poh<sup>4</sup>, Amy Belton<sup>1</sup>, Sujayita Roy<sup>1</sup>, C. Conover Talbot, Jr.<sup>5</sup>, Saraswati Sukumar<sup>2,6</sup>, David L. Huso<sup>7</sup>, Linda M. S. Resar<sup>1,2,4,8,9\*</sup>

**1** Hematology Division, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America, **2** Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America, **3** Department of Biostatistics, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America, **4** Pathobiology Graduate Program, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America, **5** Institute for Basic Biomedical Sciences, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America, **6** Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America, **7** Department of Comparative Molecular & Pathobiology, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America, **8** Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America, **9** Institute for Cellular Engineering, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America

## Abstract

Emerging evidence suggests that tumor cells metastasize by co-opting stem cell transcriptional networks, although the molecular underpinnings of this process are poorly understood. Here, we show for the first time that the *high mobility group A1* (*HMGA1*) gene drives metastatic progression in triple negative breast cancer cells (MDA-MB-231, Hs578T) by reprogramming cancer cells to a stem-like state. Silencing *HMGA1* expression in invasive, aggressive breast cancer cells dramatically halts cell growth and results in striking morphologic changes from mesenchymal-like, spindle-shaped cells to cuboidal, epithelial-like cells. Mesenchymal genes (*Vimentin*, *Snail*) are repressed, while *E-cadherin* is induced in the knock-down cells. Silencing *HMGA1* also blocks oncogenic properties, including proliferation, migration, invasion, and orthotopic tumorigenesis. Metastatic progression following mammary implantation is almost completely abrogated in the *HMGA1* knock-down cells. Moreover, silencing *HMGA1* inhibits the stem cell property of three-dimensional mammosphere formation, including primary, secondary, and tertiary spheres. In addition, knock-down of *HMGA1* depletes cancer initiator/cancer stem cells and prevents tumorigenesis at limiting dilutions. We also discovered an *HMGA1* signature in triple negative breast cancer cells that is highly enriched in embryonic stem cells. Together, these findings indicate that *HMGA1* is a master regulator of tumor progression in breast cancer by reprogramming cancer cells through stem cell transcriptional networks. Future studies are needed to determine how to target *HMGA1* in therapy.

**Citation:** Shah SN, Cope L, Poh W, Belton A, Roy S, et al. (2013) HMGA1: A Master Regulator of Tumor Progression in Triple-Negative Breast Cancer Cells. PLoS ONE 8(5): e63419. doi:10.1371/journal.pone.0063419

**Editor:** Sumitra Deb, Virginia Commonwealth University, United States of America

**Received:** February 11, 2013; **Accepted:** April 4, 2013; **Published:** May 2, 2013

**Copyright:** © 2013 Shah et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** Funding provided by the following: LMSR: NIH R21 CA2149550-0131, NIH RO1 CA092339, Safeway Breast Cancer Fund, Maryland Technology Development Corporation (TEDCO) State of Maryland Stem Cell Initiative Exploratory Award 2009-MSCRFE-0105-00. SNS: Maryland Technology Development Corporation (TEDCO) State of Maryland Stem Cell Research Fund 2011-MSCRFE-0135-00. AB: TEDCO State of Maryland Stem Cell Initiative Exploratory Award 2011-MSCRFE-0102-00. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: lresar@jhmi.edu

## Introduction

Despite advances in our ability to detect and treat breast cancer, it remains a leading cause of death in women with cancer, and the incidence is rising [1]. Approximately 15–20% of all cases are classified as triple negative breast cancer, a subtype that is frequently associated with rapid progression and poor outcomes [1–2]. Triple negative breast cancer refers to the lack of detectable markers for the estrogen receptor (ER), progesterone receptor (PR), and *Her2/neu* amplification. These tumors do not respond to our most effective and least toxic therapies, including hormonal therapy (tamoxifen) or herceptin. Thus, further studies are needed to elucidate the molecular pathways that lead to tumor progression in triple negative breast cancer and could be targeted with novel therapies.

Recent studies identified the *high mobility group A1* (*HMGA1*) oncogene as a key factor enriched in embryonic stem cells, adult stem cells, and refractory or high-grade/poorly differentiated

tumors [3–32]. The *HMGA1* gene encodes the HMGA1a and HMGA1b chromatin remodeling proteins, which result from alternatively spliced messenger RNA [4,14,22,27]. These low molecular weight (thus high mobility group) protein isoforms bind to the minor groove of chromatin at AT-rich regions. HMGA1 proteins modulate gene expression by altering chromatin structure and orchestrating the assembly of transcription factor complexes to enhanceosomes within enhancer or promoter regions throughout the genome. These proteins are highly expressed during embryogenesis with low or absent levels in adult tissues. *HMGA1* is overexpressed in all aggressive cancers studied to date, and high levels portend a poor prognosis in diverse tumors [3–31]. In fact, HMGA1 proteins are the most abundant nonhistone chromatin binding proteins found in cancer cells. A recent landmark paper demonstrated that *HMGA1* is essential for the cellular reprogramming of somatic cells to induced pluripotent stem cells by the four Yamanaka factors (Oct4, Sox2, Klf4, cMyc) [32]. HMGA1 induces expression of key stem cell transcriptional networks in

normal embryonic stem cells and during cellular reprogramming. Together, these findings suggest that HMGA1 could function in tumor progression by reprogramming differentiated cells into poorly differentiated, stem-like cancer cells.

Here, we discovered that HMGA1 is a central factor in reprogramming poorly differentiated, triple-negative breast cancer cells. Our findings further implicate HMGA1 as a master regulator in tumor progression and suggest that targeting HMGA1 pathways could be effective in poorly differentiated, metastatic tumors.

## Results

### Silencing *HMGA1* halts cell proliferation and reprograms invasive, mesenchymal-like cells

To define the role of *HMGA1* in oncogenic properties and tumor progression, we silenced *HMGA1* expression using lentiviral-mediated delivery of short hairpin RNA (shRNA) [32] in cell lines derived from aggressive, triple negative breast cancers (MDA-MB-231, Hs578T; Fig. 1A). Control cells were transduced with a control lentiviral vector [10,32]. We discovered that cell proliferation was rapidly halted in both cell lines (Fig. 1B) within the first 4 days. Surprisingly, there was a dramatic change in cell morphology whereby the spindle-shaped, fibroblast-like cells became cuboidal and epithelial-like in appearance (Fig. 1C). Because these morphologic changes are consistent with a mesenchymal-epithelial transition, we investigated the expression of genes involved in a mesenchymal-epithelial transition [10,33]. In MDA-MB-231 cells, we found that silencing *HMGA1* led to a significant repression in the mesenchymal genes, *Snail* and *Vimentin*, while there was an increase in the gene expressing the epithelial marker, E-Cadherin (Fig. 1D). Similarly, in Hs578T cells, *E-Cadherin* was induced when *HMGA1* was silenced. We also assessed tumor progression properties, including invasion and migration. In both cell lines, there was a marked reduction in migration and invasion in cells with silencing of *HMGA1* (Fig. 1E). Together, these findings indicate that silencing *HMGA1* results in a profound decrease in proliferation, migration, and invasion, as well as morphologic and gene expression changes consistent with a mesenchymal-epithelial transition.

### Silencing *HMGA1* interferes with orthotopic tumorigenicity and metastatic progression

Next, we assessed the role of HMGA1 on tumorigenesis using *in vivo* models of triple negative breast cancer. First, we assessed tumor growth following mammary fat pad implantation. We found that silencing *HMGA1* in the aggressive MDA-MB-231 cells leads to a dramatic decrease in tumor growth following mammary fat pad implantation (Fig. 2Ai). Specifically, cells ( $10^5$ ) transduced with control virus reached a volume of  $0.53 \text{ cm}^3 \pm 0.34$  at 8 weeks following mammary fat pad implantation. In contrast, the tumors from cells transduced with *HMGA1* shRNA (shHMGA1) were significantly smaller at 8 weeks following implantation ( $0.037 \text{ cm}^3 \pm 0.058$ ;  $p = 0.016$ ). Because there was a dramatic effect on primary tumorigenesis, we also sought to determine if silencing *HMGA1* interferes with metastatic progression. We therefore evaluated the lungs histopathologically for tumor foci after necropsy. Strikingly, we discovered almost no metastatic lesions to the lungs in the mice implanted with the shHMGA1 cells as compared to the mice implanted with control cells in which there were extensive, coalescing sheets of metastatic tumor cells throughout the lungs following mammary implantation with  $10^5$  cells (Fig. 2Aii). We also assessed metastatic progression following mammary fat pad implantation with a greater number of cells ( $10^7$ ) from a repeat transduction experiment, and sacrificed the

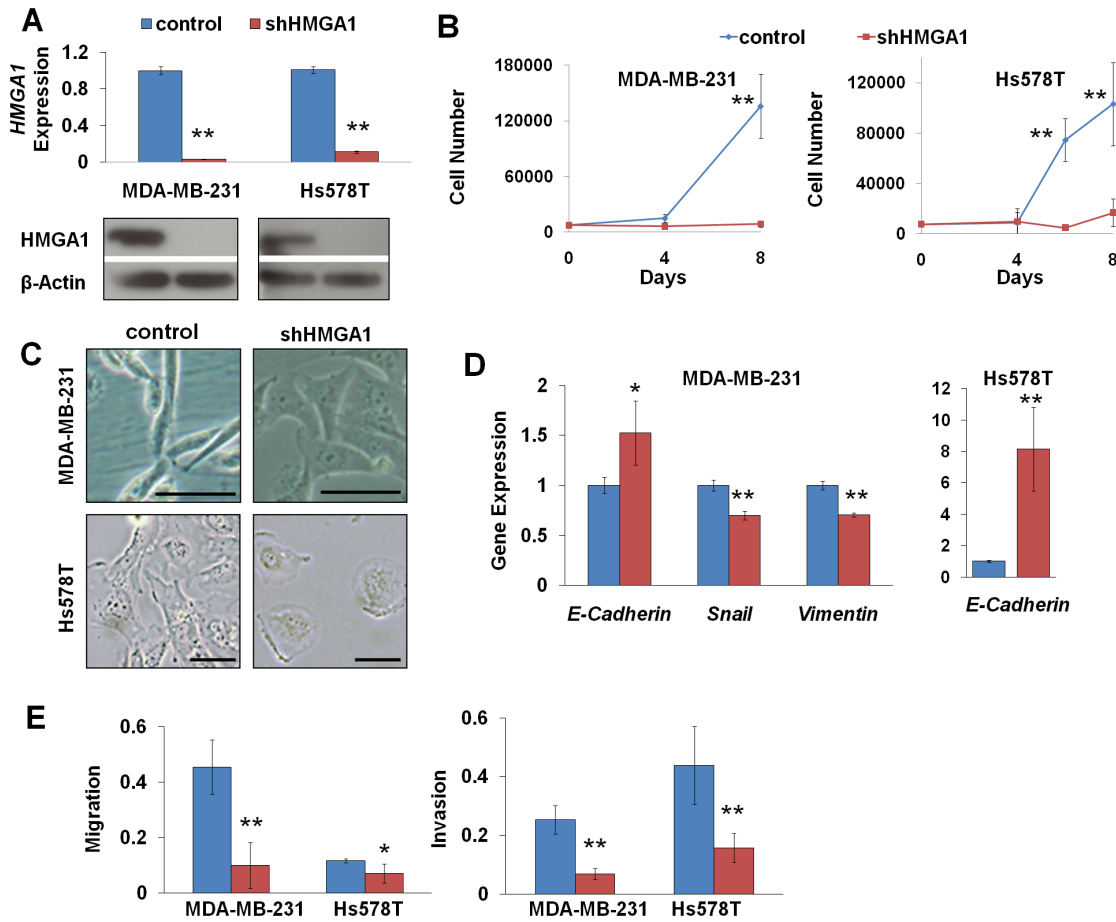
mice after 5 weeks. With the higher number of cells, tumors formed from all injections (3/3 in controls and 3/3 in shHMGA1 cells (Fig. 2Bi). Although tumors were slightly smaller from the shHMGA1 cells, the difference was not significant ( $0.64 \pm 0.27$  in controls versus  $0.17 \pm 0.072$  in shHMGA1 cells,  $p = 0.08$ ). Despite the similar tumor volumes, we observed a dramatic decrease ( $>100$ -fold) in metastatic foci in the shHMGA1 cells as compared to controls ( $0.67 \pm 1.15$  versus  $>100$  in all controls;  $p = 0.00004$ ; Fig. 2Bi & 2Bii). We also assessed lung foci following tail vein injection of control or shHMGA1 cells ( $10^6$ ) after 3 weeks. Only one lung focus was observed after injection of the shHMGA1 cells, while there were numerous foci in the control cells ( $0.25 \pm 0.5$  versus  $99.3 \pm 15.0$ ;  $p = 0.007$ ; Fig. S1).

### Silencing *HMGA1* blocks mammosphere formation and depletes tumor-initiator cells

Because silencing *HMGA1* has profound effects on oncogenic properties *in vitro*, primary tumorigenesis and metastatic progression *in vivo*, and expression of genes involved in epithelial-mesenchymal transition, we sought to determine its role in cancer stem cell characteristics. To this end, we explored the epithelial stem cell property of mammosphere formation [34] in the control and shHMGA1-treated cells (Fig. 3A–B). We found that growth of primary, secondary, and tertiary mammospheres was significantly impaired in the MDA-MB-231 cells with silencing of *HMGA1*. Similarly, we observed that there was a significant decrease in primary mammosphere formation in the Hs578T cells treated with shHMGA1. (Secondary or tertiary mammospheres do not form in control Hs578T cells, precluding analysis of these phenotypes). Next, we performed orthotopic implantations and assessed tumorigenicity with limiting dilutions. As presented above, tumors formed in both control and shHMGA1 cells when  $10^7$  or  $10^5$  cells were implanted. In contrast, no tumors formed in the MDA-MB-231 cells with silencing of *HMGA1* when  $10^4$  cells were injected (0/3), while tumors formed in all control injections (3/3; Fig. 3C). These results indicate that silencing *HMGA1* in MDA-MB-231 cells depletes the tumor-initiator or cancer stem-like cells and further underscores the role of *HMGA1* as a key regulator of stem cell properties in aggressive, triple-negative breast cancer cells.

### HMGA1 induces a stem cell signature in triple negative breast cancer cells

To globally define the transcriptional networks regulated by HMGA1, we performed gene expression profile analysis in MDA-MB-231 cells with or without *HMGA1* knock-down. To this end, we used siRNA [21,25] and observed a rapid and significant reduction in *HMGA1* expression (Fig. S2A). *HMGA1* mRNA falls dramatically by 48 hours, with persistent decreases at 72 hours (Fig. S2A). There was also a marked decrease in HMGA1 protein at 48 and 72 hours (Fig. S2B). We therefore performed global gene expression profile analysis at 48 hours using an Affymetrix exon array (GeneChip Human Exon 1.0 ST Array) with RNA from three independent replicates of each experimental condition. To define an HMGA1 signature in breast cancer, we identified the 100 transcripts that were most differentially expressed. These 100 transcripts correspond to 63 unique genes. Because *HMGA1* is enriched in embryonic stem cells and our functional studies showed that it is required for cancer stem cell properties, we compared the HMGA1 signature of 63 genes to gene expression profiles from diverse pluripotent stem cells and differentiated cells, including embryonic stem cells (ESCs), induced pluripotent stem cells, embryoid bodies, and fibroblasts [35]. As shown, un-

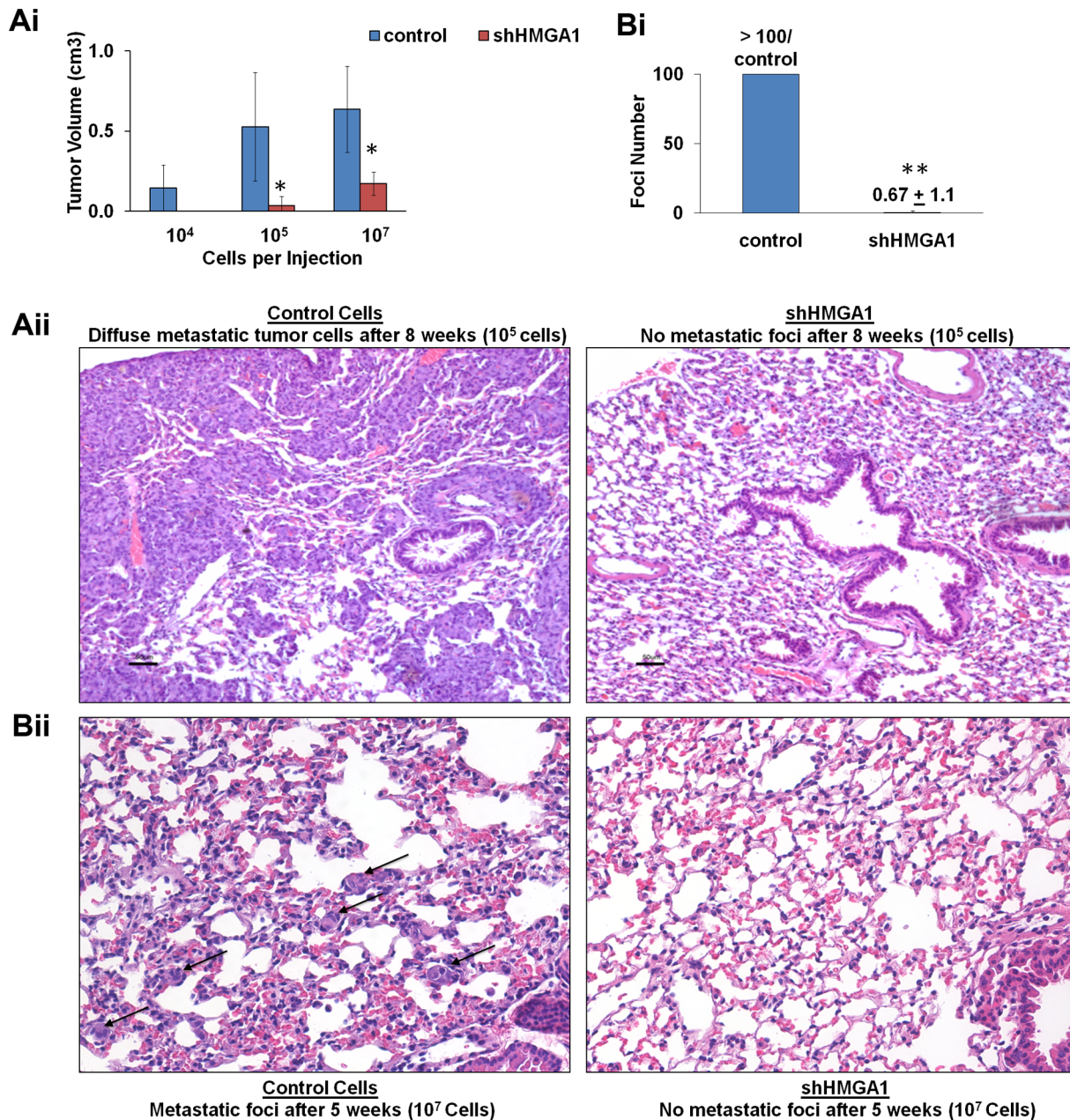


**Figure 1. Silencing *HMGA1* expression halts cell growth and induces dramatic changes in cell morphology and gene expression. A)** Lentiviral-mediated delivery of shRNA to *HMGA1* (denoted shHMGA1) results in a marked decrease in *HMGA1* mRNA and protein in triple negative breast cancer cell lines (MDA-MB-231, Hs578T). **B)** Proliferation is disrupted in cancer cell lines following silencing of *HMGA1*. **C)** Mesenchymal, fibroblast-like cancer cells undergo dramatic morphologic changes within 4 days after treatment with shHMGA1. Striking changes were observed in MDA-MB-231 (top panels) and Hs578T cells (bottom panels). Bar: 50  $\mu$ m. **D)** Alterations in EMT genes with silencing of *HMGA1*. **E)** Migration and invasion is decreased with silencing of *HMGA1*. \* $P < 0.05$ ; \*\* $P < 0.01$ . doi:10.1371/journal.pone.0063419.g001

supervised cluster analysis of these genes separates the samples by cell type with a clear distinction between pluripotent stem cells and differentiated cells. Moreover, the *HMGA1* signature is highly enriched in pluripotent/embryonic stem cells ( $p < 0.001$ ; Fig. 4A). We validated a subset of the *HMGA1* signature genes using quantitative RT-PCR, and found differential expression similar to the microarray gene expression results in all cases (Fig. S2C). These findings suggest that *HMGA1* drives tumor progression by inducing stem cell transcriptional networks.

To elucidate cellular pathways regulated by *HMGA1* in breast cancer, we analyzed the *HMGA1* signature with Ingenuity Pathway Analysis (IPA, Ingenuity Systems, <http://www.ingenuity.com>). From the top list of differentially regulated genes, 2 pathways had significant network scores (69 and 46, respectively; Fig. 4B and Fig. S3). The highest scoring network was embryonic development, tissue development, and cellular development. The top molecular and cellular functions were cell death and survival and cellular movement, while the top physiologic system development and functions included: 1. nervous system development and function, 2. organ morphology, and 3. embryonic development. In this network, the most down-regulated molecule was ARL2BP or ADP-ribosylation factor (ARF)-like 2 binding protein. This protein is a member of a functionally distinct group

of RAS-related GTPases, called the ARF family. ARL2BP protein binds to ARL2.GTP with high affinity and plays a role in the nuclear translocation, retention and transcriptional activity of STAT3 [36]. Notably, we showed that *HMGA1* induces *STAT3* expression in lymphoid tumorigenesis, and *STAT3* inhibitors are cytotoxic to the *HMGA1*-driven tumor cells [24]. TMCO1 or transmembrane and coiled-coil domains 1 protein was the most up-regulated protein in this network. Although its function is not known, it is associated with breast cancer cells [8882811 GEO Profiles-NCBI]. TGF $\beta$ 1 is a major node and this protein is up-regulated in diverse cancers and thought to promote invasion, migration, EMT and tumor progression [37]. EGFR and MAPK are other important nodes that are activated in cancer and mediate proliferative signals [9]. Another central node was HIF-1 alpha, a key factor involved in angiogenesis during tumor progression and vascular development during embryogenesis [38]. In addition, Myc was identified as a major node and prior studies found that not only does cMYC induce *HMGA1* expression [12], but *HMGA1* also directly up-regulates cMYC expression [32]. Myc also has a well-defined role in breast [39–40] and other diverse cancers [41] as well as in embryonic stem cells [41–42]. Thus, our pathway analysis further confirms the important role for



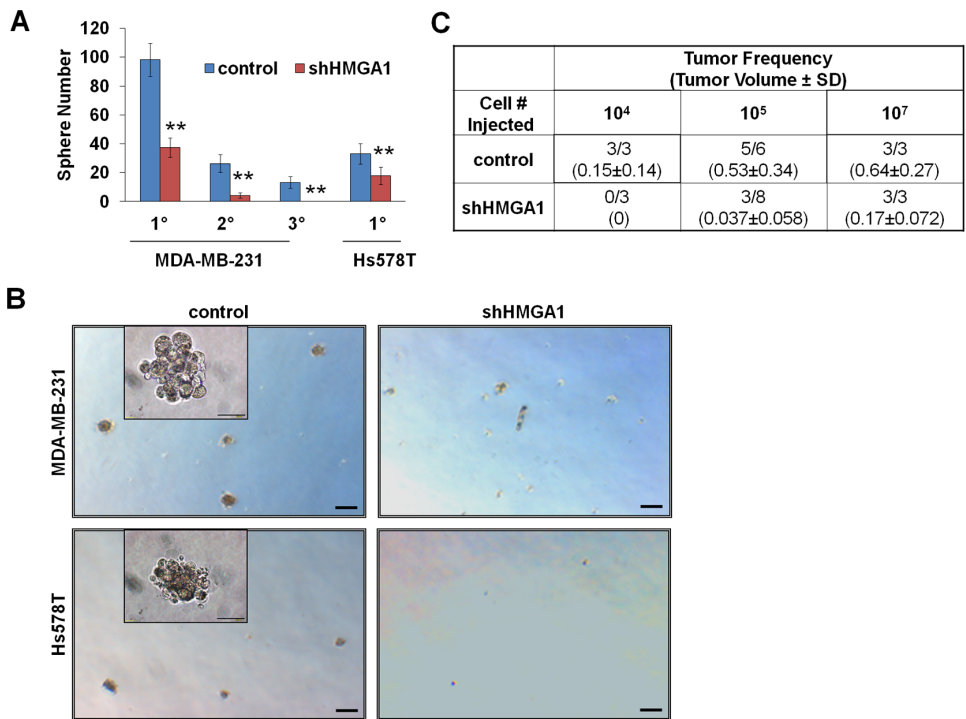
**Figure 2. Silencing *HMGA1* interferes with orthotopic tumorigenicity and metastatic progression.** **A)** Silencing *HMGA1* impairs orthotopic tumorigenicity. Tumor volumes  $\pm$  standard deviations are shown. No tumors formed from shHMGA1 cells when  $10^4$  cells were implanted. (For injections with  $10^4$  cells,  $n=3$  for control or shHMGA1 cells; for injections with  $10^5$  cells,  $n=5$  for control and  $n=8$  for shHMGA1 cells; and for injections with  $10^7$  cells,  $n=3$  for control and shHMGA1 cells. **B)** Metastatic progression is almost completely abrogated in cells that do not express *HMGA1*. This graph shows the number metastatic foci to the lung 5 weeks following implantation of MDA-MB-231 cells ( $10^7$ ) into mammary fat pads following treatment with control shRNA or shHMGA1. **C)** The top photographs show the lungs 8 weeks following implantation into mammary fat pads. There are coalescing sheets of metastatic tumor cells in the lungs of mice injected with control cells (left) as compared to mice injected with shHMGA1 cells (right). Due to the widespread tumor cells, individual foci could not be counted. Bar: 50  $\mu$ m. **D)** The bottom panels show multiple, discrete foci in the lungs 5 weeks following implantation of control cells into mammary fat pads (left) as compared to mice injected with shHMGA1 cells (right). \* $P<0.05$ ; \*\* $P<0.0001$ . doi:10.1371/journal.pone.0063419.g002

HMGA1 in regulating embryonic stem cell networks during tumor progression in breast cancer.

## Discussion

Here, we report for the first time that silencing *HMGA1* induces a rapid and dramatic reprogramming of highly proliferative,

invasive, mesenchymal-like breast cancer cells to more differentiated, slowly growing, epithelial-like cells. We also found that knock-down of *HMGA1* has profound effects on oncogenic properties associated with both tumor initiation (orthotopic tumorigenesis) and tumor progression (migration, invasion, and metastatic progression). In fact, the *in vivo* effects on metastatic progression were even more pronounced than the effects on



**Figure 3. Silencing *HMGA1* blocks mammosphere formation and depletes tumor-initiator cells.** **A)** Silencing *HMGA1* blocks mammosphere formation in MDA-MB-231 cells (1°, 2°, 3°) and Hs578T cells (1°). **B)** Photographs of mammospheres following treatment of breast cancer cells with control or shHMGA1. Silencing *HMGA1* significantly inhibits mammosphere formation in MDA-MB-231 and Hs578T cells. Bars: 200 μm (large panels) and 50 μm (insets). **C)** Tumor numbers at limiting dilutions show that silencing *HMGA1* depletes the tumor initiator/cancer stem cells in MDA-MB-231 cells. Note that no tumors formed following injection of 10<sup>4</sup> cells treated with shHMGA1, while tumors formed in all cases when control cells were injected. Both tumor frequency and tumor volumes (± standard deviations) are shown. \*P<0.05; \*\*P<0.01. doi:10.1371/journal.pone.0063419.g003

primary tumorigenesis, thus highlighting the role of *HMGA1* in tumor progression. The changes induced by silencing *HMGA1* are among the most striking alterations reported to date with knock-down of *HMGA1* or most other oncogenes for that matter, both in degree and rate of onset. The profound effects could be related to our efficient, viral-mediated delivery of shRNA to repress *HMGA1*. In addition, triple negative breast cancer cells may be highly dependent upon *HMGA1* and related pathways for their oncogenic properties. Indeed, a study from the Broad Institute at MIT identified *HMGA1* as a key transcription factor enriched in triple negative breast cancer [3]. Moreover, expression of *HMGA1* and 8 additional genes predicted poor outcomes in breast cancer, as well as brain and bladder cancer. Prior studies using antisense or dominant-negative approaches in triple-negative breast cancer cells (MDA-MB-231 or Hs578T) also showed that anchorage-independent cell growth or colony formation are inhibited by *HMGA1* repression [15–16]. We also have preliminary evidence demonstrating that *HMGA1* expression correlates with more advanced nuclear grade in primary tumors (Asch & Resar, unpublished data).

Emerging evidence further indicates that *HMGA1* is important in maintaining a de-differentiated, pluripotent stem-like state [32]. A recent landmark paper demonstrated that *HMGA1* is required for cellular reprogramming of somatic cells to induced pluripotent stem cells (iPSCs) by the Yamanaka factors [31]. Blocking *HMGA1* expression or function prevents the derivation of iPSCs. In normal embryonic stem cells in culture and during the reprogramming process to iPSCs, *HMGA1* activates expression of stem cell transcriptional networks. Recent studies also found that tumor progression and an epithelial-mesenchymal transition (EMT)

involves transcriptional networks important in stem cells [3,9–10,32–33]. The first evidence linking *HMGA1* to EMT came from an important study in 2001 in MCF-7 breast cancer cells, which demonstrated that forced expression of *HMGA1* results in metastatic progression and histologic changes consistent with EMT in the epithelial MCF-7 breast cancer cell line [16]. This group also found that *HMGA1* induces changes in classes of genes involved in tumor progression. More recently, studies in colon cancer showed that *HMGA1* is required for tumor progression and stem cell properties [10]. Here, we discovered that *HMGA1* is required for mammosphere formation, including secondary and tertiary mammospheres in MDA-MB-231 cells. We also found that silencing *HMGA1* depletes tumor initiator/cancer stem cells, indicating that targeting *HMGA1* in breast cancer therapy could have an important impact on the cancer stem cell population, which is believed to be the basis for refractory disease in diverse tumors. These functional studies are corroborated by the *HMGA1* signature and pathway analysis demonstrating that *HMGA1* orchestrates transcriptional networks important in stem cells and metastatic progression.

There is a dire need to understand the molecular underpinnings of metastatic progression because this is the major cause of death in patients with cancer. Although cancer is a highly complex and heterogeneous disease, with significant heterogeneity even within a single tumor, increasing evidence indicates that common, central pathways exist that could serve as “Achilles heels” or rational therapeutic targets in diverse tumors. Our studies presented here underscore the fundamental role for *HMGA1* in tumor progression in preclinical models for aggressive, triple negative breast cancers. This work, together with prior studies in diverse tumors,



**Figure 4. The HMGA1 signature is enriched in pluripotent stem cells, including embryonic and induced pluripotent stem cells. A)** The HMGA signature derived from genes with the greatest expression changes in the control versus *HMGA1* knock-down cells displayed as a heat map. Green depicts down-regulation in expression, while red depicts up-regulation; black denotes little or no change in expression. The HMGA1 signature overlaps with pluripotent stem cell genes that distinguish human embryonic stem cells (hESCs) and induced pluripotent stem cells (iPSCs) from fibroblasts and embryoid bodies (EB). Genes ( $n=63$ ) were selected for the greatest changes in expression in the breast cancer cell lines with *HMGA1* knock-down as compared to the control breast cancer lines (Fig. S1). In a hierarchical clustering of fibroblasts, hESCs, iPSCs, and EBs derived from the hESCs, these genes distinguish samples by type. The majority of the HMGA1 signature genes, represented in blue along the left margin, are significantly differentially expressed between fibroblasts and human pluripotent stem cells (hESC/iPSCs;  $p<0.001$ ). **B)** HMGA1 network derived from the list of differentially expressed genes using Ingenuity Pathway Analysis (IPA) with microarray gene expression data from control and *HMGA1* knock-down in MDA-MB-231 cells. From 63 differentially expressed genes as the focus gene set, the highest-scoring network was Embryonic Development, Tissue Development, and Cellular Development (score = 69). Red nodes indicate up-regulation; green nodes indicate down-regulation. Arrows and lines denote interactions between specific genes within the network. A, activation; E, expression regulation; I, inhibition; L, proteolysis; LO, localization; M, biochemical modification; MB, membership of a group or complex; P, phosphorylation; PD, protein-DNA interaction; PP, protein-protein interaction; PR, protein-RNA interaction; RB, regulation of binding; RE, reaction; T, transcription; TR, translocation.  
doi:10.1371/journal.pone.0063419.g004

provide compelling evidence that HMGA1 is a master regulator in the evolution of primary tumors to metastatic disease. Further studies are now needed to develop approaches to target HMGA1 in cancer.

## Materials and Methods

### Ethics statement

All animal experiments were conducted in accordance with a protocol approved by the Johns Hopkins University Animal Care and Use Committee (protocol #MO11M270). Mice were housed in a sterile environment where they had free access to food and water as outlined in our institutional guidelines.

### Cell culture and proliferation assay

Cells (MDA-MB-231 and Hs578T) were cultured as recommended (ATCC). For proliferation assays, cells were seeded (7,500/well) and counted using an automated cell counter (Nexcelom). Each experiment was done in triplicate and performed at least twice.

### RNA Interference

The short-hairpin RNA interference vector for HMGA1 targets 5'-CAACTCCAGGAAGGAAACCAA-3' and has been described elsewhere [43]. Virus was prepared as previously described. The empty vector was used as a negative control as we described [10]. Polyclonal, transduced cells were selected and maintained in puromycin (1  $\mu\text{g}/\text{ml}$ ). To ensure that the effects of silencing HMGA were not a result of a single clone, independent, polyclonal transductions were done at least twice for each experiment. All functional experiments were performed in duplicate or triplicate and replicated after a repeat transduction experiment and polyclonal selection of shRNA or control cells. Repression of HMGA1 was confirmed in each case at the level of gene expression (qRT-PCR) and Western analysis.

### Migration and invasion assays

Invasion assays were performed as previously described [25] with the following modifications. Briefly, 15,000 cells were resuspended in serum-free media (500  $\mu\text{l}$ ) and placed in the upper chamber of a 24-well BD BioCoat™ Matrigel™ Invasion Chamber coated with Matrigel. Invasion was calculated as the percentage of total cells that invaded into the bottom chamber containing complete media with serum. Migration was performed similarly, except that Matrigel was omitted.

### Orthotopic tumorigenicity and metastatic foci experiments

Cells (suspended in 75  $\mu\text{l}$  phosphate-buffered saline (PBS)) were implanted into murine (NOD-scid IL2R $\gamma$ <sup>null</sup>) mammary fat pads with an equal volume of Matrigel. Tumor volumes (calculated by  $4/3 \pi \times \text{length}/2 \times \text{width}/2 \times \text{depth}/2$ ) were monitored daily until they reached 1–1.5  $\text{cm}^3$ , after which mice were euthanized. The presence of tumor foci within the lung was analyzed histopathologically. For tail vein injection experiments, cells ( $10^6$ ) were resuspended in PBS (150  $\mu\text{l}$ ). Mice were euthanized after 3 weeks, and lungs were examined histopathologically.

### Mammosphere Assay

Mammosphere assays were performed as previously described [34] and spheres ( $>50 \mu\text{m}$ ) were counted.

### Western Analysis

Western blots were performed as previously described [10,12], using commercial antibodies to HMGA1 (Abcam) and  $\beta$ -Actin (Cell Signaling), both at a 1:1000 dilution.

### Gene Expression Analysis with Quantitative, Reverse Transcription PCR

Total RNA was isolated using the Direct-zol RNA MiniPrep kit (Zymo) and analyzed by qRT-PCR as we previously described. The expression level of each gene was normalized to the human *RPLP0* (Applied Biosystems) or  $\beta$ -actin gene. Primers for *E-Cadherin*, *Snail*, and *Vimentin* were previously described [10,33].

### HMGA1 knock-down and gene expression profile analysis

*HMGA1* was knocked-down in MDA-MB-231 cells using siRNA as we previously described [10,21,25]. RNA was isolated and hybridized to the Affymetrix exon array (GeneChip Human Exon 1.0 ST Array) as previously described. Expression data was pre-processed using the robust microarray (RMA) algorithm [44], as implemented in the *oligo* software package [45] available from Bioconductor [46] and annotated to the most recent human genome using the *getNetAffx* function in the *oligo* package. Microarray data were uploaded to Gene Expression Omnibus (GSE45483). Expression profiles from HMGA1 knock-down cells were compared to control cells treated with the vector RNA using Bayes modified t-tests and the *limma* package from Bioconductor [47]. We focused our analysis on the 100 most differentially expressed transcripts.

To compare this signature to genes expressed in a panel of embryonic stem cells, induced pluripotent stem cells, embryoid bodies, and fibroblasts, we downloaded expression profiles for a 43 sample study (GSE25970) [35] from the Gene Expression

Omnibus [48]. Unsupervised cluster analysis was performed on all probes annotated to the 63 genes in the HMGA1 panel using agglomerative clustering with complete linkage. Euclidean distance and t-tests were used to compare probe-specific expression in stem cells and fibroblasts.

Pathway analysis of differentially expressed genes was performed using Ingenuity Pathway Analysis (IPA, Ingenuity Systems, <http://www.ingenuity.com>) as we described [9]. IPA scores were generated for each network and indicate the likelihood that the focus genes present in the network could occur by chance alone. A score of  $\geq 3$  is considered significant because it represents a 1/1,000 chance that the network contains specific focus genes by random chance alone [9].

## Supporting Information

**Figure S1 Silencing HMGA1 in MDA-MB-231 blocks the formation of foci to the lung following tail vein injections.** **A)** Lung foci were enumerated 3 weeks following tail vein injections of control or shHMGA1 MDA-MB-231 cells ( $n = 3$  for control mice;  $n = 4$  for shHMGA1 mice). **B)** Graph of the mean number of tumor foci  $\pm$  standard deviation shows a striking decrease in foci following injection of shHMGA1 MDA-MB-231 cells as compared to controls ( $p = 0.007$ ).

(TIF)

**Figure S2 Silencing HMGA1 in MDA-MB-231 results in significant repression in HMGA1 mRNA and protein, with alterations in gene expression.** **A)** Independent replicate experiments of MDA-MB-231 cells with or without HMGA1 knock-down result in silencing HMGA1 at the level of

mRNA. **B)** HMGA1 protein is also repressed following treatment with siRNA. **C)** Validation of genes in the HMGA1 signature shows that gene expression assessed by quantitative RT-PCR (qRT-PCR) parallels that of the microarray results. **D)** Table comparing differential expression of the HMGA1 signature identified by microarray and qRT-PCR.

(TIF)

**Figure S3 HMGA1 network derived from differentially expressed genes in MDA-MB-231 with or without HMGA1 knock-down.** From 63 differentially expressed genes as the focus gene set, the second highest-scoring network was Cardiovascular Disease, Cell Death and Survival, and Nervous System Development and Function (score = 46). Colors, arrows, lines and abbreviations are described under Figure 4B. NF- $\kappa$ B, ERK, and MAPK are major nodes, which have been identified in prior studies of global gene expression profiles mediated by HMGA1 [9].

(TIF)

## Acknowledgments

This work is dedicated to the memory of Dr. Wenlan Wang, a friend and colleague who courageously battled breast cancer. She will forever inspire all who had the opportunity to know her personally or through her work.

## Author Contributions

Conceived and designed the experiments: SNS SS DLH LMSR. Performed the experiments: SNS WP AB SR. Analyzed the data: SNS LC WP AB SR CCT DLH LMSR. Contributed reagents/materials/analysis tools: LC CCT. Wrote the paper: SNS LMSR.

## References

- Siegel R, Naishadham D, Jemal A (2013) Cancer statistics. *CA Cancer J Clin* 63: 11–30.
- Lee JS, Fackler MJ, Lee JH, Choi C, Park MH, et al. (2010) Basal-like breast cancer displays distinct patterns of promoter methylation. *Cancer Biol Ther* 9: 1017–1024.
- Ben-Porath I, Thomson MW, Carey VJ, Ge R, Bell GW, et al. (2008) An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat Genet* 40: 499–507.
- Resar LMS (2010) The High Mobility Group A1 Gene: Transforming Inflammatory Signals into Cancer? *Cancer Research* 70: 436–439.
- Chou B-K, Mali P, Huang X, Ye Z, Dowey S, et al. (2011) Efficient human iPSC cell derivation by a non-integrating plasmid from blood cells with unique epigenetic and gene expression signatures. *Cell Res* 21: 518–529.
- Zhou G, Chen J, Lee S, Clark T, Rowley JD, et al. (2001) The pattern of gene expression in human CD34(+) stem/progenitor cells. *Proc Natl Acad Sci U S A* 98: 13966–13971.
- Karp JE, Smith BD, Resar LS, Greer JM, Blackford A, et al. (2011) Phase I and pharmacokinetic study of bolus-infusion flavopiridol followed by cytosine arabinoside and mitoxantrone for acute leukemias. *Blood* 117: 3302–3310.
- Nelson DM, Joseph B, Hillion J, Segal J, Karp JE, et al. (2011) Flavopiridol induces BCL-2 expression and represses oncogenic transcription factors in leukemic blasts from adults with refractory acute myeloid leukemia. *Leuk Lymphoma* 52: 1999–2006.
- Schuldenfrei A, Belton A, Kowalski J, Talbot CC, Di Cello F, et al. (2011) HMGA1 Drives Stem Cell, Inflammatory Pathway, and Cell Cycle Progression Genes During Lymphoid Tumorigenesis. *BMC Genomics* 12: 549.
- Belton A, Gabrovsky A, Iacobuzio-Donahue C, Huso DL, Resar LMS (2012) HMGA1 induces intestinal polyposis in transgenic mice and drives tumor progression and stem cell properties in colon cancer cells. *PLoS One* 7: e30034.
- Shah SN, Resar LMS (2012) High Mobility Group A1 and Cancer: Potential Biomarker and Therapeutic Target. *Histol Histopathol* 27: 567–579.
- Wood IJ, Mukherjee M, Dolde CE, Xu Y, Maher JF, et al. (2000) HMG-I/Y, a new c-Myc target gene and potential oncogene. *Mol Cell Biol* 20: 5490–5502.
- Pedula ML, Treff NR, Resar LM, Reeves R (2001) Sequence and analysis of the murine Hmg1y (Hmg1) gene locus. *Gene* 271: 51–58.
- Reeves R, Beckerbauer L (2001) HMGI/Y proteins: flexible regulators of transcription and chromatin structure. *Biochim Biophys Acta* 1519: 13–29.
- Reeves R, Edberg DD, Li Y (2001) Architectural transcription factor HMGI(Y) promotes tumor progression and mesenchymal transition of human epithelial cells. *Mol Cell Biol* 21: 575–594.
- Dolde CE, Mukherjee M, Cho C, Resar LMS (2002) HMG-I/Y in human breast cancer cell lines. *Breast Cancer Research and Treatment* 71: 181–191.
- Dhar A, Hu J, Reeves R, Resar LM, Colburn NH (2004) Dominant-negative c-Jun (TAM67) target genes: HMGA1 is required for tumor promoter-induced transformation. *Oncogene* 23: 4466–4476.
- Takaha N, Resar LMS, Vindivich D, Coffey DS (2004) High mobility group protein HMGI(Y) enhances tumor cell growth, invasion, and matrix metalloproteinase-2 expression in prostate cancer cells. *The Prostate* 60: 160–167.
- Hommura F, Katabami M, Leaner VD, Donniger H, Sumter TF, et al. (2004) HMG-I/Y is a c-Jun/activator protein-1 target gene and is necessary for c-Jun-induced anchorage-independent growth in Rat1a cells. *Mol Cancer Res* 2: 305–314.
- Xu Y, Sumter TF, Bhattacharya R, Tesfaye A, Fuchs EJ, et al. (2004) The HMG-I oncogene causes highly penetrant, aggressive lymphoid malignancy in transgenic mice and is overexpressed in human leukemia. *Cancer Res* 64: 3371–3375.
- Tesfaye A, Di Cello F, Hillion J, Ronnett BM, Elbahloul O, et al. (2007) The high-mobility group A1 gene up-regulates cyclooxygenase 2 expression in uterine tumorigenesis. *Cancer Res* 67: 3998–4004.
- Fusco A, Fedele M (2007) Roles of HMGA proteins in cancer. *Nat Rev Cancer* 7: 899–910.
- Di Cello F, Hillion J, Kowalski J, Ronnett BM, Aderinto A, et al. (2008) Cyclooxygenase inhibitors block uterine tumorigenesis in HMGA1a transgenic mice and human xenografts. *Molecular Cancer Therapeutics* 7: 2090–2095.
- Hillion J, Dhara S, Sumter TF, Mukherjee M, Di Cello F, et al. (2008) The high-mobility group A1/signal transducer and activator of transcription-3 axis: an achilles heel for hematopoietic malignancies? *Cancer Res* 68: 10121–10127.
- Hillion J, Wood IJ, Mukherjee M, Bhattacharya R, Di Cello F, et al. (2009) Upregulation of MMP-2 by HMGA1 promotes transformation in undifferentiated, large-cell lung cancer. *Mol Cancer Res* 7: 1803–1812.
- Hristov AC, Cope L, Di Cello F, Reyes MD, Singh M, et al. (2010) HMGA1 correlates with advanced tumor grade and decreased survival in pancreatic ductal adenocarcinoma. *Mod Pathol* 23: 98–104.
- Reeves R (2010) Nuclear functions of HMG proteins. *Biochim Biophys Acta* 1799: 3–14.
- Di Cello F, Dhara S, Hristov A, Kowalski J, Elbahloul O, et al. (2013) Inactivation of the *Cdkn2a* locus cooperates with HMGA1 to drive T-cell leukemogenesis. *Leuk Lymphoma* [Epub ahead of print].
- Williams MD, Reeves R, Resar LMS, Hill HH Jr (2013) Metabolomics of colorectal cancer (CRC): Past and current analytical platforms. *Anal Bioanal Chem* [Epub ahead of print].



30. Pomeroy SL, Tamayo P, Gaasenbeek M, Sturla LM, Angelo M, et al. (2002) Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature* 415: 436–442.
31. Flohr AM, Rogalla P, Bonk U, Puettmann B, Buerger H, et al. (2003) High mobility group protein HMGA1 expression in breast cancer reveals a positive correlation with tumour grade. *Histol Histopathol* 18: 999–1004.
32. Shah SN, Kerr C, Cope L, Zambidis E, Liu C, et al. (2012) HMGA1 reprograms somatic cells into pluripotent stem cells by inducing stem cell transcriptional networks. *PLoS One* 7: e48533.
33. Mani SA, Guo W, Liao MJ, Ng E, Ayyanan A, et al. (2008) The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 133: 704–715.
34. Shaw FL, Harrison H, Spence K, Ablett MP, Simões BM, et al. (2012) A detailed mammosphere assay protocol for the quantification of breast stem cell activity. *J Mammary Gland Biol Neoplasia* 17: 111–117.
35. Bock C, Kisinis E, Verstappen G, Gu H, Boulting G, et al. (2011) Reference maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell* 144: 439–452.
36. Muromoto R, Sekine Y, Imoto S, Ikeda O, Okayama T, et al. (2008) BART is essential for nuclear retention of STAT3. *Int Immunol* 20: 395–403.
37. Massague J (2008) TGFbeta in cancer. *Cell* 134: 215–230.
38. Semenza GL (2012) Cancer stromal cell interactions mediated by hypoxia-inducible factors promote angiogenesis, lymphangiogenesis, and metastasis. *Oncogene* in press.
39. Thibodeaus CA, Liu X, Disbrow GL, Zhang Y, Rone JD, et al. (2009) Immortalization and transformation of human mammary epithelial cells by a tumor-derived Myc mutant. *Breast Cancer Res Treat* 116:281–294.
40. Tront JS, Huang Y, Fornace AJ, Hoffman B, Liebermann D (2010) *Cadd45a* functions as a promoter or suppressor of cancer dependent on the oncogenic stress. *Cancer Res* 70:9671–9681.
41. Dang CV (2012) MYC on the path to cancer. *Cell* 2:304–307.
42. Nie Z, Hu G, Wei G, Yamane A, Resch A, et al. (2012) c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell* 151:68–79.
43. Liao S-S, Jazag A, Whang EE (2006) HMGA1a is a determinant of cellular invasiveness & *in vitro* metastasis in pancreatic adenocarcinoma. *Cancer Res* 66: 11613–11622.
44. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, et al. (2003) Summaries of Affymetrix GeneChip Probe Level Data. *Nucleic Acids Res* 31: e15.
45. Carvalho BS, Irizarry RA (2010) A framework for oligonucleotide microarray preprocessing. *Bioinformatics* 26: 2363–2367.
46. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5: R80.
47. Smyth GK (2005) Limma: linear models for microarray data, in R. Gentleman; R. V. Carey; S. Dudoit; R. Irizarry; concordance at the top (CAT) plot & W. Huber, eds., 'Bioinformatics and Computational Biology Solutions using R and Bioconductor', Springer-Verlag: 397–420.
48. Barrett T, Suzek TO, Troup DB, Wilhite SE, Ngau W-C, et al. (2005) NCBI GEO: mining millions of expression profiles—database and tools. *Nucleic Acids Res* 33: D562–D566.