HHS Public Access

Author manuscript

Cancer Res. Author manuscript; available in PMC 2021 December 01.

Published in final edited form as:

Cancer Res. 2021 June 01; 81(11): 3008–3021. doi:10.1158/0008-5472.CAN-19-4031.

Epigenetic and post-transcriptional modulation of SOS1 can promote breast cancer metastasis through obesity-activated c-Met signaling in African American women

Fei Xing[1,#], **Dan Zhao**[1], **Shih-Ying Wu**[1], **Abhishek Tyagi**[1], **Kerui Wu**[1], **Sambad Sharma**[1], **Yin Liu**[1], **Ravindra Deshpande**[1], **Yuezhu Wang**[1], **Jacob Cleary**[1], **Lance D. Miller**[1], **Amar G. Chittiboyina**[2], **Chinni Yalamanchili**[2], **Yin-Yuan Mo**[3], **Kounosuke Watabe**[1,*,#]

[1]Department of Cancer Biology, Wake Forest University School of Medicine, Winston-Salem, NC 27157, USA

[2]National Center for Natural Products Research, School of Pharmacy, The University of Mississippi, University, MS 38677, USA

[3]Cancer Institute, University of Mississippi Medical Center, Jackson, MS 39216, USA

Abstract

Ethnicity is considered to be one of the major risk factors in certain subtypes of breast cancer. However, the mechanism of this racial disparity remains poorly understood. Here we demonstrate that SOS1, a key regulator of Ras pathway, is highly expressed in African American (AA) breast cancer patients compared to Caucasian American (CA) patients. Because of the higher obesity rate in AA women, increased levels of SOS1 facilitated signal transduction of the c-Met pathway which was highly activated in AA breast cancer patients via HGF secreted from adipocytes. Elevated expression of SOS1 also enhanced cancer stemness through upregulation of PTTG1 and promoted M2 polarization of macrophages by CCL2 in metastatic sites. SOS1 was epigenetically regulated by a super-enhancer identified by H3K27ac in AA patients. Knockout of the super-enhancer by CRISPR in AA cell lines significantly reduced SOS1 expression. Furthermore, SOS1 was post-transcriptionally regulated by miR-483 whose expression is reduced in AA patients through histone tri-methylation (H3K27me3) on its promoter. The natural compound taxifolin suppressed signaling transduction of SOS1 by blocking the interaction between SOS1 and Grb2, suggesting a potential utility of this compound as a therapeutic agent for AA breast cancer patients.

[#]Co-corresponding authors. [*]To whom correspondence should be addressed: Kounosuke Watabe Ph.D. Department of Cancer Biology, Wake Forest University School of Medicine, Winston Salem, NC 27157; Phone: 336-716-0231; Fax: 336-716-0255; kwatabe@wakehealth.edu.

AUTHOR CONTRIBUTIONS

FX and KW designed the study and wrote the manuscript. FX conducted experiments and acquired, analyzed and interpreted the data. AC and CY performed computer based docking. YL assisted with bioinformatics data analysis and data interpretation. DZ performed Chip-PCR. SS, SYW, KWu, FX, YL, DZ, RPD, YZW, LDM, JC, AT, YYM and KW reviewed and edited manuscript and interpreted the data. KW supervised the study.

COMPETING FINANCIAL INTERESTS

No authors have any particular competing interests.

Introduction

The age-adjusted breast cancer incidence in CA women in the US is 128.7 (per 100,000 women per year) compared to 125.5 in AA women (1). However, the mortality rate is 20.8 in CA women and 29.5 in AA women (1). Thus, the cancer death rate in AA females is 14% higher than in CA females despite a 7% lower incidence rate. In addition, AA women have a higher incidence of the most aggressive subtype, known as triple-negative breast cancer, than other racial and ethnic groups (2). AA women also have the highest rates of late-stage disease (1). Therefore, four major disparities are evident between AA and CA patients: (i) early age of onset, (ii) more advanced stage of disease, (iii) more aggressive histologic changes, and (iv) shorter survival times. However, a comprehensive understanding of the basis for these disparities remains elusive.

Non-biological factors, such as socioeconomic characteristics, access to health care, and delivery of treatment modalities, are important in this disparity (3,4). However, many biological and genetic factors have also been identified as factors contributing to the disparity (5,6). These factors include the levels of growth factors and hormones in plasma, expression of their receptors, genetic susceptibility loci, tumor suppressors, epigenetic changes, and chromosomal rearrangement (7). Moreover, microRNAs are powerful post-transcriptional regulators and they have been extensively studied in the context of racial disparities (8). Zhao et. al performed microarray analysis by comparing the levels of circulating miRNAs in plasma samples from 20 breast cancer patients with early stage breast cancer together with matched healthy controls. Interestingly, only two microRNAs are overlapped between AA and CA groups when compared to the normal samples, indicating a distinctive tumor associated microRNA profile between different racial groups (9). A genome-wide miRNA profiling of TNBC tumor tissues from AA and CA also revealed 26 miRNAs that are differentially upregulated in AA patients, and 23 of them were known to be involved in pathways crucial to tumor progression (10). Differences in the epigenetic levels such as DNA methylation of cancer-related genes have also been observed in AA patients (11–13). Mehrotra et al. showed that tumors from AA breast cancer patients have a significantly higher rate of hypermethylation of five tumor suppressor genes (HIN-1, Twist, Cyclin D2, RAR-beta, and RASSF1A) than CA tumors (14). Wang et. al demonstrated that AA women had increased hypermethylation of *CDH13* whose inactivation is associated with early onset breast cancer (13). Histone modifications including acetylation and methylation are critical epigenetic events that affect gene expression (15). However, the underlying mechanisms of racial specific histone modification are virtually unknown.

Cancer is no longer being treated as a solitary disease whose progression also requires the support from the tumor microenvironment (TME) (16). Growing evidence have shown that different racial backgrounds have an impact on the cellular components as well as secretory profile of TME which eventually affect the survival of patients (17). Several groups reported relatively higher levels of cytokines such as IL-6 and VEGF in AA breast cancer patients as compared to CA patients (18,19). Interestingly, both cytokines are known to be excessively secreted by the adipocytes which are more abundant in AA due to their higher obesity rate (20,21). Cancer-associated adipocytes (CAAs) are frequently found adjacent to cancer cells in breast cancer patients, and they affect the growth or invasion of cancer cells through

various secretary factors (22). The crosstalk between adipocytes and cancer cells also affects the gene expression of both cell types, which can further enhance tumor progression (23). Adipocytes are known to secrete high level of HGF which have been shown to regulate multiple cellular processes including proliferation, invasion and angiogenesis (24). The HGF/c-Met axis is able to trigger various downstream signaling pathways in tumor cells, such as PI3K/AKT, JAK/STAT and Ras/MAPK (25). HGF-activated c-Met receptor can interact with the SH2/SH3 domain of Grb2, which recruits guanine nucleotide exchange factors (GEFs) such as SOS1(26). SOS1 can further convert Ras-GTP to its active form by recruiting it from the cell matrix to the membrane (26). Negative feedback of SOS1 and Ras pathway has been reported by many studies. Several groups showed that SOS1 can be phosphorylated upon MAPK activation which disrupts its association with the adaptor proteins Grb2 (27,28) . Thus, Grb2–Sos1 interaction plays a critical role in the activation of the Ras pathway which makes it as an ideal drug target for cancer therapy.

Evidently, multiple racial specific mechanisms mediated by genetic and epigenetic factors together with the tumor microenvironment contribute to the progression of breast cancer in AA patients. However, it is still not clear how those factors affect each other and how are they related to the racial disparities of breast cancer progression. Therefore, we hypothesize that signaling pathways activated by racial associated genes together with their tumor microenvironment contribute to the tumor progression. In this study, through unbiased GSEA-based screening and analyses of multiple clinical cohorts, we found that SOS1 which is essential for the activation of Ras signaling, was highly expressed in AA breast cancer patients. Furthermore, we have identified two potential mechanisms that regulate SOS1 expression in AA patients by activation of a unique super enhancer that is associated with SOS1 gene, and by miR-483 mediated post-transcriptional regulation. In addition, we have demonstrated that the self-renewal of cancer stem cells and the accumulation of M2 macrophages in the lung metastatic sites are both regulated by the upregulation of SOS1. These novel discoveries shed new light into our understanding in breast cancer disparity which leads to an identification of potential therapeutic targets.

Material and Methods

Cell culture and reagents

Human breast carcinoma cell lines, MDA-MB-468 (MDA468), HCC1806, HCC1569, HCC1500, MCF7, T47D, ZR75–1, and MDA-MB231 (MDA231), were purchased from American Type Culture Collection. MCF7, MDA468 and MDA231 were cultured in DMEM medium supplemented with 10% FBS, streptomycin (100 mg/ml) and penicillin (100 units/ml). HCC1806, HCC1569, HCC1500, T47D and ZR75–1 were cultured in RPMI medium supplemented with 10%FBS, streptomycin (100mg/ml) and penicillin (100units/ml). All cells were grown at 37 °C in a 5% CO_2 atmosphere. All cells are tested negative of Mycoplasma by PCR (ATCC Universal Mycoplasma Detection Kit), and they were used between passages 15 and 45. All cell lines were obtained between 2010 and 2019, and they were authenticated by qRT-PCR analysis for the expression of 20 signature genes.

Western Blot

Western blot analysis was performed as described previously (29), using antibodies against SOS1 (1/1000; Cell Signaling Technology), p-AKT (1/1000; Cell Signaling Technology), t-AKT (1/5000; Cell Signaling Technology), t-Erk (1/1000; Cell Signaling Technology),p-Erk(1/1000; Cell Signaling Technology) and PTTG1(1/5,000; Cell Signaling Technology).

Quantitative real-time PCR

Total RNA was isolated from the cells and reverse transcribed as described previously (30). The cDNA was then amplified with a pair of forward and reverse primers for the following genes: SOS1 (5'- GAGTGAATCTGCATGTCGGTT -3' and 5'-CTCTCATGTTTGGCTCCTACAC-3'), PTTG1 (5'- ACCCGTGTGGTTGCTAAGG -3' and 5'- ACGTGGTGTTGAAACTTGAGA-3'), CXCL1(5'-CCCAAGAACATCCAAAGTGTG-3' and 5'-CACATTAGGCACAATCCAGGT-3'), CCL2 (5'- CATCTCCTACACCCCACGAAG -3' and 5'- GGGTTGGCACAGAAACGTC -3'), CD74(5'- GATGACCAGCGCGACCTTATC -3' and 5'- GTGACTGTCAGTTTGTCCAGC -3'), CD68 (5'- GGAAATGCCACGGTTCATCCA-3' and 5'-TGGGGTTCAGTACAGAGATGC-3'), Arg1 (5'- GTGGAAACTTGCATGGACAAC -3' and 5'- AATCCTGGCACATCGGGAATC -3') and CD206 (5'-TCCGGGTGCTGTTCTCCTA -3' and 5'- CCAGTCTGTTTTTGATGGCACT -3'and. PCR reactions were performed using CFX Connect (Bio-Rad) and iTaq Universal SYBR Green Supermix (Bio-Rad). The thermal cycling conditions composed of an initial denaturation step at 95°C for 5 min followed by 40 cycles of PCR using the following profile: 94 oC, 30 s; 60 oC, 30 s; and 72 oC, 30 s.

Immunohistochemistry

Human primary breast cancer specimens were obtained from surgical pathology archives of the Wake Forest Comprehensive Cancer Center, and Cooperative Human Tissue Network (CHTN). All tissue sections were obtained by surgical resection. Samples were sectioned with 5μm thickness from the formaldehyde-fixed and paraffin-embedded tissue specimens. The sections were deparaffinized and antigens were retrieved by heating the slides in 10 mM sodium citrate (pH 6.0) at 85 °C for 30min. The slides were treated with 3% H_2O_2 and then incubated overnight at 4 °C with anti-SOS1 antibody (1/500; Cell Signaling Technology). The sections were then incubated with secondary antibodies and visualized using the Envision-plus kit (Dako Corp.).

GSEA

The Gene MatriX file (.gmx) which includes 110 breast cancer specific signatures was generated in our previous study (29). Gene Cluster Text file (.gct) was generated from the TCGA and GSE78958. Categorical class file (.cls) was generated based on the race of each patient. The number of permutations was set to 1000.

Metastasis-free survival analysis

We compiled a microarray dataset of 710 patients from GEO (accession numbers: GSE12276, GSE2034, GSE2603, GSE5327, and GSE14020). These datasets were all

normalized using MAS5.0, and each microarray was centered to the median of all probes. For each patient, metastasis free survival was defined as the time interval between the surgery and the diagnosis of metastasis.

Immunocytochemistry

Cells were fixed with 70% ethanol, washed with PBS and blocked with 2% BSA for 1hr. After blocking, cells were washed again with PBS and incubated with antibodies for, anti-p-c-Met (1/500: R&D), and anti-CCL2 (1/500; Abcam ab9779) overnight at 4 °C. Cells were then incubated with anti-rabbit IgG Alexa Fluor (R) 555 molecular probe (Cell Signaling Technology) for 1hr at room temperature. Fluorescence images were taken by a fluorescent microscope. For frozen section staining, slides were fixed with methanol, washed with PBS and blocked by 2% BSA for 1hr. After blocking, the slides were washed again with PBS and incubated with anti-ARG1 (1/500; Cell Signaling Technology) and anti-CD68 (1/200; Biolegend), for 1hr at room temperature. Fluorescence images were taken by a fluorescent microscope (Olympus IX71).

Establishment of enhancer knockout cell lines

HCC1806 and HCC1569 were transfected with Cas9 protein V2 (ThermoFisher) and paired sgRNAs (sgRNAF1: GCACGACAUGACAGUCCCAA, sgRNA R1: GAGUAUUGCAUCUCAAGUGG) that flank the target sequence based on Synthego's protocol. Single cell derived colonies were selected and the knockout efficiency was measured by PCR with following primers (F1: 5'- CTTCCCTGGCTAGCTCCAAC -3', 5'-GGGACTTTGCAGGTAGGCTT-3') as shown in Supplementary Fig. 2b.

FACS (Fluorescence-activated cell sorting)

For cell surface marker analysis, cells were suspended in FACS buffer (PBS with 0.1% BSA and 0.1% Tirton X100) followed by incubation with FITC conjugated anti-CD24 (eBioscience, Inc.), APC conjugated anti-CD44 (BioLegend), PE conjugated anti-ESA (eBioscience, Inc.) for 15mins at room temperature. Cells were then washed with PBS and re-suspended in PBS for FACS analysis using the C6 Flow cytometer (Accuri LTD).

Sphere forming assay

Cells were plated (1000 cells/ml) in ultra-low attachment plates (Corning, Acton, MA, USA) with DMEM/F12 supplemented with 2% B27 (GIBCO), 20ng/ml EGF (Sigma), and 4μg/ml Insulin (Sigma). Mammospheres with diameters over 100μm were counted and data was represented as the means ± SEM.

Animal experiments

All animal experiments were performed in accordance with a protocol approved by the Wake Forest Institutional Animal Care and Use Committee. For experimental metastasis assay, wild type C57/B6 and BALB/C mice (7–8 weeks) were injected with 2×10^5 luciferase-labeled E0771 or 4T1 cells in PBS into tail vein in a total volume of 100 μl. The lung metastasis progression was monitored and the luminescence was quantified at the indicated time points.

ChIP-seq analysis

For ChIP-seq analysis, reads were first mapped to hg19 human genome by Bowtie version 1.1.0: followed by generating Wig file by using DANPOS 2.2.3. Quantile normalization is performed using DANPOS 2.2.3 to make sure all ChIP-Seq datasets from different sample types were normalized to have the same quantiles. Bigwig is generated using the tool WigToBigWig downloaded from ENCODE project website We then submitted the bigWig file to the UCSC Genome Browser (https://genome.ucsc.edu) to visualize ChIP-Seq signal at each base pair.

ChIP-PCR analysis

ChIP was performed using the Qiagen EpiTect ChIP oneday kit. Briefly, patients' tissues were collected under IRB protocol IRB00048646 at Wake Forest University School of Medicine. Tissues were cryo homogenized in liquid nitrogen and fixed using 1% formaldehyde. Collected cell pellets in cell harvesting buffer was sonicated using Covaris DNA shearer to fragment the chromatin to within a range of 100–400bp. For each immunoprecipitation, 3μg anti-H3K27Ac (ab4729), anti-H3K27me3 (ab6002) or control IgG (ab171870) was used. Realtime PCR was performed using isolated input or ChIP DNA. % Input was calculated according to manufacturer's protocol as 2(-DeltaCt[normalized ChIP]) while deltaCT equals Ct[IP]-(Ct[input]-log2(input dilution factor)). Primer sequences are as follows: for SOS1 enhancer (F: 5'-CCTCCTGTTTGCCTCCTCTG-3' and R:5'-CAAGGCTCACTTGTGTTGGC-3'), for miR-483 promoter (5'-TCAGCCGCAACAACCAGCAA -3' and R:5'-GCCAAGGCCATGCTGAATGC -3')

Single cell RNA sequencing (scRNA-seq)

Lung metastatic nodules derived from C57/B6 mice inoculated with E0771/shCTRL or E0771/shSOS1 cells were dissociated with tumor dissociation kit (Miltenyi) followed by isolation of CD45+ immune cells using CD45 microbeads (Miltenyi). All samples reached 80% of viability measured by trypan blue stain before library preparation. scRNA-seq were performed by the Cancer Genomics Shared Resource of the WFBMC-CCC. Single cell suspension were loaded into a 10× Chromium single cell capture chip followed by library preparation using the Chromium Single Cell 3' Library and Gel Bead kit according to the manufacturer's protocol (10× Genomics). Sequencing libraries were then loaded onto an Illumina NextSeq500 for paired-end sequencing with High Output 150 cycle kit (Illumina). Sample de-multiplexing and alignment are processed by the Cell Ranger Single Cell Software Suite v.2.0.1.

Computational docking

To search for potential inhibitors that can block SOS1 activation, we employed structure-based computational studies at the binding site of the SOS1 in the nSH3 domain of Grb2. Based on the reported solution NMR studies of Grb2 with ten residues of SOS1 peptide (PDB: 4GBQ) structure, molecular docking studies were conducted using the GLIDE module in Schrodinger Suite. Analysis of the ligand-interaction diagram of 10-residues of SOS1 peptide revealed the ionic interactions between Glu16 of Grb2-nSH3 and the Arg8 of

the SOS1 peptide as a critical ionic interaction and hence such hydrogen-bonding interaction with Glu16 used as a requisite for all our docking studies.

Results

SOS1 is highly expressed in African American breast cancer patients

We previously established a novel screening approach based on the gene set enrichment analysis (GSEA) which identifies signaling pathways that enriched in certain phenotype by using expression cohorts from patients and pathway-specific signature gene sets (31). To identify signaling pathways that are specific to AA breast cancer patients, we performed the GSEA for breast cancer patients (740 CA Vs 180 AA) by using The Cancer Genome Atlas (TCGA) and 110 different pathways previously established for breast cancer (31). Using the cutoff criteria (NES>1.2, P<0.05), we identified four pathways (c-Met, Polr2a, E2F5, and Gli1) that were most significantly enriched in AA patients with breast cancer (Figure 1A, B). Next, because of the higher mortality rate in AA patients compared to the CA patients, we examined the correlation of metastasis-free survival with these pathways and found that only the c-Met was positively correlated to lung metastasis but not to bone metastasis in breast cancer patients (Supplementary Figure 1 A, B). Because triple negative cancer is particularly prevalent in AA patients, we examined whether the enrichment of c-Met pathway is specific to this subtype. We found significant enrichment of the c-Met pathway in patients with both basal and luminal subtypes, suggesting that the enrichment of c-Met signaling is not associated with specific cancer subtypes (Supplementary Figure 1C, D). In addition to the TCGA database, we examined enrichment of c-Met pathway in the AA patients by using another cohort (GSE78958), and we found again that the enrichment was independent of the subtypes (Supplementary Figure 1E–G). Because the c-Met signature used in this analysis include over 150 genes, to further identify the genes that are related to racial disparity, we examined 17 genes that are directly associated with the c-Met signaling, by analyzing GSE78958. We found that SOS1 was the only gene significantly upregulated in AA patients (n=96) compared to CA patients (n=304) (Figure 1C). The higher level of SOS1 in AA breast cancer patients compared to other ethnicity groups was further validated using GSE20194 (Figure 1D). Importantly, SOS1 expression is upregulated in AA patients in all subtypes of breast cancer in two independent patient cohorts (GSE20194, GSE78958) (Supplementary Figure 1 H, I). We then examined the expression of SOS1 in tumor tissues from AA (n=15) and CA (n=15) patients by qRT-PCR and immunohistochemistry. We found that the expression of SOS1 was indeed significantly higher in AA patients and it was not correlated with the tumor subtypes (Figure 1 E,F and Supplementary Figure 1J) . Comparison of SOS1 expression in various breast cancer cell lines (44 lines for CA and 8 lines for AA) using GSE41313 also revealed higher expression of SOS1 in AA cell lines (Figure 1G). Next, we then directly examined SOS1 expression in these cell lines by qRT-PCR and western blot and found that SOS1 was indeed highly expressed in AA cell lines (Figure 1H). Notably, SOS1 expression was significantly correlated with worse lung metastasis-free survival of breast cancer patients but not bone metastasis-free survival (Figure 1I, Supplementary Figure 1 K). These data strongly suggest critical roles of SOS1 in breast cancer progression in AA women.

SOS1 is epigenetically controlled by a super enhancer

Histone modifications are one of the essential epigenetic changes that modulate gene expression by affecting the chromatin structures that enhance or block the binding of transcription factors or other elements that control the activity of transcription(15). To examine whether upregulation of SOS1 is due to differential histone modifications between AA and CA, we performed the chromatin immunoprecipitation sequencing (ChIP-seq) analysis for the *SOS1* gene and neighboring regions, using our own CHIP-seq results (HCC1569) for H3K27Ac and H3K27me3 together with existing ChIP-seq data from MDA468, MDA231and MCF7. We did not observed any significant differences between two AA cell lines (HCC1569 and MDA468) and two CA cell lines (MDA231 and MCF7) regarding the H3K27me3 signaling (Supplementary Figure 2A). Strikingly, however, we did observed a strong enrichment of H3K27Ac signaling in the region encompassing ch2: 38662500–38672400 bases which is located 500 kb downstream of the *SOS1* gene in AA cell lines, but not in CA cell lines, suggesting that these regions are highly activated and may affect the expression of neighboring genes, including SOS1 (Figure 2A). Interestingly, we found that this region includes an enhancer that spans approximately 1.3 kb (chr2: 38665459– 38666743) defined by the Enhancer Atlas database (Figure 2B). We also examined breast tumor samples from each race by ChIP-PCR analysis using anti- H3K27Ac antibody and found that this enhancer region is indeed significantly enriched in AA patients compared to CA patients (Figure 2C). We then made deletion of the super enhancer region using the paired gRNA mediated CRISPR technique designed by CRISPETa and validated the successful knockout of the sequence by PCR (Supplementary Figure 2B). As shown in Figure 2D, deletion of the enhancer indeed significantly reduced the SOS1 expression in three AA cell lines. Thus, our results identified a unique race-specific H3K27Ac- mediated gene regulation through the enhancer in AA patients, which promotes the SOS1 expression.

SOS1 is downregulated by racial specific miR-483

Growing evidence showed differential expression of micoRNAs in both solid tumor and biofluid from different ethnic groups (8). To examine whether SOS1 is regulated by any racial specific microRNAs, we compared the expression profile of microRNAs between AA and CA patients by using the TCGA database. We particularly focused on microRNAs that are downregulated in AA breast cancer patients and have the potential binding sites on the 3'UTR of SOS1. We found that miR-483 was significantly downregulated in AA patients and has two binding sites on the 3'UTR of SOS1 gene predicted by TargetScan (Figure 3A, B). In another cohort, miR-483 expression was also lower in AA patients compared to CA breast cancer patients, suggesting that miR-483 may suppress SOS1 gene in AA specific manner (Figure 3C). To examine this hypothesis, we overexpressed miR-483 in two AA cell lines by lenti-virus infection followed by western blot analysis for SOS1 protein. As shown in Figure 3D, ectopic expression of miR-483 significantly suppressed the protein levels of SOS1 in HCC1806 and MDA468 cells. Furthermore, low miR-483 expression was correlated to poorer relapse-free survival in breast cancer patients, indicating a metastasis suppressor function of this microRNA (Figure 3E). We also analyzed miR-483 expression in various breast cancer cell lines derived from different ethnic groups and found that miR-483 was expressed at significantly lower level in cell lines derived from AA patients compared to cell lines derived from CA patients (Figure 3F). We also confirmed this finding by

performing qRT-PCR for various cell lines with different racial background (Figure 3G). A significant reverse correlation between the expression of miR-483 and SOS1 was also observed in a cohort contains 283 breast cancer patients (Figure 3H). To examine whether histone modifications are involved in the down regulation of miR-483 in AA, we analyzed the CHIP-seq data derived from AA and CA cell lines as described in Fig 2. We found a strong enrichment of H3K27me3 signaling located upstream of the miR-483 gene promoter which spans approximately 2kb (Figure 3I red boxed). Importantly, 8 out of 10 transcriptional start sites (red colored) that were predicted by miRStart website are located in this region, indicating a potential transcriptional suppression by histone tri-methylation in the promoter of miR-483 (Figure 3J). To validate this hypothesis in clinical samples, we performed H3K27me3 CHIP-PCR in breast tumor tissues derived from AA and CA patients (n=4 each) and found that this region is indeed highly tri-methylated in histone among AA samples compared to CA samples (Figure 3K). These results strongly suggest that miR-483 is downregulated in a racial specific manner by promoter histone tri-methylation and that this downregulation results in upregulation of SOS1.

Downregulation of SOS1 suppresses lung metastasis in obese mice *in vivo*

The upregulation of the *SOS1* gene alone likely does not have any consequences in AA if no signaling is transduced from the upstream tyrosine kinase receptors which require the binding with their corresponding ligands. As shown in Fig. 1c, we found that c-Met pathway was highly activated in AA patients and there was no difference of HGF expression, the ligand that activates c-Met in tumor cells, between AA and CA. Hence, we speculate that circulating or tumor microenvironment-derived HGF play a critical role in the activation of c-Met in tumor cells. In the US, 46.4% of AA women are obese compare to 26.6% in CA women (32). To examine whether c-Met activation is related to the obesity, we examined HGF expression in a breast cancer cohort (GSE70947) that contains both normal and cancerous tissues separated by laser captured microdissection (LCM) from patient together with their information of body mass index (BMI). As shown in Figure 4A, HGF was highly expressed in normal tissue but not in tumors in both obese and overweight women. To examine which cell types in the normal tissues secret the highest amount of HGF, we compared the HGF expression in four different types of stromal cells that frequently found in the tumor microenvironment. Indeed, we found adipocytes express the highest amount of HGF compared to other stromal cells and cancer cells (Figure 4B, Supplementary Figure 3A). Moreover, conditioned medium derived from adipocytes significantly induced the phosphorylation of c-Met in MDA468 and HCC1806 cells (Figure 4C, Supplementary Figure 3B). Treatment with recombinant HGF also enhanced the phosphorylation of AKT and Erk, and such activation was abolished by knockdown of SOS1 (Figure 4D). We therefore examined the effect of HGF on lung metastasis in an immune competent syngeneic mouse model. We found that administration of HGF significantly augmented lung metastasis of 4T1 cells, but knockdown of the *SOS1* gene abrogated this effect of HGF (Figure 4E), suggesting that the HGF-cMet signaling promotes lung metastasis of breast cancer through activation of SOS1. To further examine the effect of obesity on metastasis, we generated obese C57/B6 mice using a high fat diet and found that the HGF level in the serum of OB mice was significantly higher compared to lean controls (Supplementary Figure 3C). We also found that obesity indeed significantly promoted lung metastasis of E0771 cells and that

this effect was nullified by knockdown of the *SOS1* gene (Figure 4F). Collectively, these results suggest that obesity increases HGF expression from adipocytes, which in turn activates the c-Met-SOS1 signaling. This activation then promotes the progression of breast cancer metastasis.

SOS1 induces PTTG1 expression and promotes stemness of cancer stem-like cells (CSCs)

Next, we attempted to identify key regulatory factors that are influenced by racial disparity using Master Regulator Analysis, an unbiased bioinformatics approach that identifies key genes that are enriched in certain phenotypes from patients' genetic profiles (33) (Figure 5A). We compared the enrichment of various master regulators between AA and CA patients in both cohorts. Nine genes were identified as common master regulators that are associated with race in breast cancer patients in GSE78958 and TCGA. We found that PTTG1 which belongs to the c-Met signature gene analyzed in Figure 1 was the most significantly and differentially expressed regulator in both cohorts (Figure 5B). Moreover, high level of PTTG1 was positively correlated to a worse lung metastasis-free survival of breast cancer patients (Figure 5C). Notably, PTTG1 expression was highly correlated with SOS1 expression (R^2=0.6), suggesting its potential role as a downstream effector of SOS1 mediated signaling (Figure 5D). We then examined the expression of PTTG1 in four AA cell lines with or without knockdown of the *SOS1* gene. The successful knockdown of SOS1 was accomplished by infecting cells with a mixture of shRNAs (#1, #2 and #4) (Supplementary Figure 4A). As shown in Figure 5E and F, PTTG1 expression was significantly suppressed by knockdown of SOS1, suggesting that SOS1 plays a critical role in regulating PTTG1 expression. It has been reported that PTTG1 plays a critical role in the EMT and self-renewal of CSCs (34). To examine whether PTTG1 is essential to the stemness of AA cell lines, we generated PTTG1 knockdown cell lines by using lenti-virus mediated shRNA (Figure 5G). Indeed, we found that knockdown of PTTG1 significantly decreased the CSCs population by both FACS (CD24-/CD44+/ESA+) analyses and mamosphere formation assay (Figure 5G and H and Supplementary Figure 4B). Furthermore, mRNA levels of PTTG1 are positively correlated with CD44 and ESA in breast cancer patients (n=710) (Supplementary Figure 4C).Our data indicates that PTTG1 is a downstream effector of SOS1 which further promotes the stemness of CSCs.

SOS1 enhances M2 polarization of macrophages and alter the immune landscape of the lung

It has been well recognized that metastatic growth of tumor cells requires the adaption and interaction with the tumor microenvironment. To investigate whether SOS1 play a role in remolding the tumor microenvironment in the lung, we examined the immune landscape of metastatic lesions in the lung after implanting E0771 cells with or without knockdown of SOS1 through the tail vein injection into immune competent mice. CD45+ immune cells were sorted by MACS from metastatic tumors followed by single cell sequencing analysis. We found that the knockdown of the SOS1 significantly reduced the number of M2 macrophages (Figure 6 A,B), suggesting that SOS1 may regulate the expression of secreted factors that affect the polarization of macrophages. Notably, conditioned medium from MDA468 cells but not from SOS1 knockdown cells significantly induced CD163 expression, a well-established marker for M2 macrophage, in PMA-treated THP-1 cells

(Figure 6C). To identify the potential factors secreted from SOS1^{high} tumor cells, we performed cytokine/growth factor array analysis using the conditioned medium of breast tumor cells with or without expression of SOS1, and found that knockdown of SOS1 suppressed a group of inflammatory cytokine expression including CXCL1, CCL2, IGFBP2 and CXCL7 (Figure 6D, E). Among these factors, CCL2 but not CXCL1 was found to be highly secreted by cancer cells compared to macrophage, suggesting that CCL2 may have an impact on the surrounding tumor microenvironment (Supplementary Figure 5A). Therefore, we decided to focus on further investigating the role of CCL2 in M2 macrophage polarization. As shown in Figure 6F, knockdown of SOS1 significantly decreased the levels of CCL2 in two AA cell lines. Furthermore, recombinant CCL2 significantly induced the expression of M2-associated genes while decreased M1-associated genes in PMA-treated THP-1 cells (Figure 6G). These results suggest that tumor cells with high SOS1 expression secrete CCL2 that promotes M2 polarization of macrophages, which in turn enhances tumor progression in the lung.

Taxifolin suppresses lung metastasis by inhibiting SOS1 signaling

Our results suggest that upregulated SOS1 in AA women is a potential risk factor for aggressive breast cancer. Our results have also shown that when c-Met is activated by HGF, which is often overexpressed in obese patients, SOS1 signaling promotes tumor progression by affecting both tumor cells and the tumor microenvironment. Therefore, SOS1 is potentially a therapeutic target for AA breast cancer patients. There is a growing body of evidence indicating that natural compounds are very effective chemo-preventive agents due to their lower toxicity and high bioavailability compare to traditional chemotherapies. To search for natural compounds that can block the SOS1 mediated signaling, we used GLIDE docking software to screen natural compounds that may block the interaction of SOS1 and Grb2 using a library containing 840 natural compounds. Top three natural compounds were selected followed by functional assay to validate the screening results. We identified taxifolin as the most promising compound that inhibits the downstream signaling of SOS1 comparing to avicularin and vasicine (Supplementary Figure 6 A) The overlay of docking poses of taxifolin (yellow) and SOS peptide (Grey) was shown in Figure 7 A. Engagement of SOS1 and Grb2 activates Ras and blocking this interaction by taxifolin is expected to suppress the Ras activation. We then tested the inhibitory activity of taxifolin on Akt activation, which is a key downstream event of SOS1-Ras signaling. Taxifolin indeed suppressed phosphorylation of Akt in an HGF-dependent manner (Figure 7B). Furthermore, taxifolin treatment significantly decreased PTTG1 expression in two AA cell lines both in mRNA and protein levels examined by qRT-PCR and Western blot (Figure 7C,D). Immunocytochemical analyses also revealed that taxifolin strongly suppressed CCL2 levels in both AA cell lines (Figure 7E). We also tested the efficacy of taxifolin *in vivo* by administering the compound to immune competent obese mice that had been intravenously injected with E0771 cells. We found that taxifolin effectively suppressed lung metastasis in this model (Figure 7F). Moreover, we found that taxifolin treatment significantly decreased M2 macrophages in the lung metastatic sites by immunocytochemistry using two different sets of antibodies that detect M2 macrophages (CD68$^+$/CD163$^+$ and CD68$^+$/Arg1$^+$) (Figure 7G). Therefore, taxifolin is considered to be a promising drug for treating metastatic breast cancer in AA women with obesity.

Discussion

Racial disparity is one of the major risk factors that affect the tumor progression for breast cancer patients, and AA patients are known to develop more aggressive metastatic diseases compared to CA patients (1). In this study, we found that SOS1, a guanine nucleotide exchange protein which is essential for the activation of Ras protein, was significantly upregulated in AA patients. Importantly, the expression of SOS1 is controlled through the super enhancer which is epigenetically activated in an AA-dependent manner. To the best of our knowledge, this is the first study to demonstrate that differential gene expression in AA and CA can be regulated by racial specific histone modifications. Increased levels of SOS1 was previously reported in AA patients with prostate cancer, and this gene was also found to be one of the 18 genes that are upregulated in AA breast cancer patients (35,36). Missense mutations in SOS1 have been reported in individuals with Noonan syndrome but not in breast cancer patients (37,38). Racial specific SNPs have been shown to have a substantial impact on gene expression in various cancers (39–41). Bensen and colleagues performed a large scale genome wide SNP study by analyzing 8,350 women including 3,663 breast cancer patients and 4,687 controls in the cohort of African American Breast Cancer Epidemiology and Risk (AMBER) Consortium and found only four racial specific SNPs that affect the expression of miR3065 and BAIAP2, but not SOS1(42). Interestingly, Moncini et al. identified a heterozygous c.755T>C variant in *SOS1* that was able to enhance the activity of Ras-Erk pathway in Noonan syndrome patients (43). However, because this SNP affects only the activity of SOS1 protein but does not alter the mRNA expression, it would be necessary to analyze other potential SNPs that are located in the promoter regions or gene body of *SOS1* in AA and CA patients.

Epigenetic modification, particularly DNA methylation has been shown to play a critical role in gene regulation in the context of racial disparities (44,45). Fraser et al. examined the DNA methylation status around the promoter of over 14, 000 genes in 180 cell lines isolated from AA and CA populations and found racial specific DNA methylation in over one third of those genes(46). Furthermore, Benevolenskaya et al. identified hypermethylation in the promoter regions of *FZD9, MME , BCAP31, HDAC9, PAX6 , SCGB3A1 , PDGFRA , IGFBP3* , and *PTGS2* genes from 75 breast tumor samples including 21 CA, 31 AA, and 23 Hispanic patients (47). Regardless of the drastic difference of methylation patterns among AA and CA, the expression levels of methyltransferases were unchanged, indicating that other mechanisms may be involved in this event (Supplementary Figure 7A). Here we demonstrated that *SOS1* was upregulated by a super enhancer identified by H3K27Ac signaling that is specifically enriched in AA cell lines and tumor tissues from AA patients. Furthermore, we have identified miR483 as a microRNA targeting SOS1. Importantly, we found that miR483 expression was regulated by H3K27me3 histone methylation in the promoter region. Therefore, both key components, miR-483 and the super enhancer, that control the expression of SOS1 is regulated by racial specific histone modifications. However, the overall histone modification levels and the amount of histone deacetyltransferases (HDACs) and histone methyltransferases (HMT) are at similar levels between those two races (Supplementary Figure 7B,C).

In addition to the histone modifications that lead to an enhanced transcriptional activity of SOS1 in AA tissues, we have identified miR-483 as a post-transcriptional regulator of SOS1 by targeting its 3'UTR. Both tumor promoting and suppressive functions of miR-483 have been reported in different cancers. MiR-483 is also known to be embedded in the IGF2 gene and it binds directly to the 5'UTR of *IGF2* mRNA, which promotes its transcription by enhancing the association of the RNA helicase DHX9 with the *IGF2* transcript (48). However, we found no difference between AA and CA patients (Supplementary Figure 7D). We also found that the expression of miR-483 was significantly decreased in AA cell lines and tumor samples, which is validated by multiple cohorts and our clinical samples. Furthermore, higher level of miR-483 was correlated with a better relapse-free survival in breast cancer patients. Differential expression of several microRNAs between AA and CA have been shown by others by genetic profiling of tumor samples. However, a mechanism involved in racial specific microRNA regulation is vastly unknown, and our study is the first to show that racial specific histone modification plays a critical role in microRNA modulation in different ethnicity groups.

The nonbiological factors such as socioeconomic factors, obesity, and plasma hormone levels are known to correlate with the aggressiveness and survival of breast cancer regardless of race (4). Obesity is more prevalent in African American women compared to other races, and it has been associated with a more advanced stage of breast cancer (49,50). Our GSEA-based screening showed that the c-Met pathway was highly activated in AA patients. We also showed that HGF, the ligand of c-Met, is highly expressed in obese patients. Furthermore, Bousquenaud et al. recently showed that obesity promotes breast cancer metastasis by enhancing neutrophil infiltration and EMT in syngeneic breast cancer models with high fat diet (51). Since AA women have significantly higher incidence of obesity than CA women, the higher amount of HGF in the circulation of AA patients is likely to activate the c-Met pathway in tumor cells, which would be further augmented by overexpression of SOS1. Therefore, obesity is considered as one of racial factors that contribute to the progression of breast cancer. In addition to the various secreted growth factors and ligands that directly activate corresponding receptors and signaling pathways, obesity may also play a role in epigenetic regulation of gene expression. Wahl and colleagues found that BMI was associated with the DNA methylation status in 10,261 healthy donors and that adiposity is the cause of DNA methylation (52). We have demonstrated that overexpression of SOS1 amplifies the signaling that was transduced from various tyrosine kinase receptors such as c-Met which is further activated by HGF secreted from adipocytes.

The roles of M2 macrophage have been well characterized in both primary and metastatic tumors in various types of cancers (53). Our single-cell RNA sequencing and cytokine array analysis revealed that SOS1 promotes the lung metastasis through secreted CCL2 which enhances the M1–M2 transition of macrophages. It has been reported that obesity increases the amount of adipose tissue macrophages that are mainly pro-inflammatory (54). Moreover, Tashiro et al demonstrated that saturated fatty acid increased the number of lung macrophages in mice fed with high-fat diet (55). Therefore, it is highly plausible that increased amount of macrophages in the lung may serve as a pre-metastatic niche for the outgrowth of tumor cells, which will further induce the M1–M2 switch of macrophages. In addition to macrophages, neutrophils have also been reported to be recruited in the lung of

mice fed with high-fat diet in a CXCL1 dependent manner (56). Manicone and colleagues showed that high fat diet increased CXCL1 and CXCL2 but not CCL2 levels in the plasma of obese mice (57). Interestingly, our cytokine array data also showed a strong decrease of CXCL1 expression after knock down of SOS1 in tumor cells. Since tumor cells are not abundant source of CXCL1compaerd to immune cells (Supplementary Figure 5A), it is still possible that CXCL1 directly promotes tumor metastasis in an autocrine manner, which has been previously reported.

Ras protein itself is known to be “undruggable” and most of the efforts have been focused on targeting Ras membrane interaction and subcellular localization (58). In addition, inhibitors of downstream pathways of Ras such as RAF–MEK–ERK are still under clinical evaluation (59). Increasing evidence showed that targeting the upstream transducers of Ras signaling may be a more effective approach. Hillig et al. performed a high-throughput screening using the Bayer drug library which consists of over 3 million compounds and identified quinazoline as a potential Ras inhibitor by blocking the interaction between Ras and SOS1 (60). Yu et al. developed a dimeric peptide that forms a covalent bond with cytosol Grb2 protein which leads to the inhibition of cell mobility and viability (61) . Cussac et al. reported another approach to target tumorigenic Ras-dependent processes by using a proline-rich Sos-derived peptide dimer which inhibits the recognition between Sos and the two SH3 domains of Grb2 (62) . We have identified taxifolin which blocks the interaction between SOS1 and Grb2 by using computer-aided molecular docking approach by searching a most comprehensive natural compound library that has ever been reported. Tumor suppressive function of taxifoline has previously been reported in UV-induced skin cancer by targeting EGFR mediated PI3K pathway, which is consistent with our findings that taxifolin attenuates HGF-induced phosphorylation of AKT (63). Notably, Tumurkhuu et al. showed that mutation of SOS1 affected the expression levels of multiple genes that are involved in both RAS/MAPK and PI3K/AKT pathways (64). In addition, SOS1 expression is indispensable to the activation of the HRAS/PI3K/AKT pathway and maintaining nuclear translocation of p-AKT in liver cancer (65). Therefore, it is possible that taxifolin attenuates tumor growth through multiple mechanisms. Nevertheless, since SOS1 and HGF mediated c-Met signaling is highly activated in AA patients, taxifolin may serve as an effective chemo-preventive drug for AA patients with obesity (Figure 7H).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

This work was supported by NIH grant RO1CA173499, R01CA185650, R01CA205067 (to KW) and R37CA230451 (to FX). This study used various Shared Resources including Tumor Tissue and Pathology, Cancer Genomics, Flow Cytometry, and Cell Engineering that are supported by the Comprehensive Cancer Center of Wake Forest University NCI, National Institutes of Health Grant (P30CA012197).

References

1. DeSantis CE, Ma J, Goding Sauer A, Newman LA, Jemal A. Breast cancer statistics, 2017, racial disparity in mortality by state. CA Cancer J Clin 2017;67(6):439–48. [PubMed: 28972651]

2. Siddharth S, Sharma D. Racial Disparity and Triple-Negative Breast Cancer in African-American Women: A Multifaceted Affair between Obesity, Biology, and Socioeconomic Determinants. Cancers (Basel) 2018;10(12).

3. Gerend MA, Pai M. Social determinants of Black-White disparities in breast cancer mortality: a review. Cancer Epidemiol Biomarkers Prev 2008;17(11):2913–23. [PubMed: 18990731]

4. Danforth DN, Jr. Disparities in breast cancer outcomes between Caucasian and African American women: a model for describing the relationship of biological and nonbiological factors. Breast Cancer Res 2013;15(3):208. [PubMed: 23826992]

5. Gupta V, Haque I, Chakraborty J, Graff S, Banerjee S, Banerjee SK. Racial disparity in breast cancer: can it be mattered for prognosis and therapy. J Cell Commun Signal 2018;12(1):119–32. [PubMed: 29188479]

6. Yedjou CG, Tchounwou PB, Payton M, Miele L, Fonseca DD, Lowe L, et al. Assessing the Racial and Ethnic Disparities in Breast Cancer Mortality in the United States. Int J Environ Res Public Health 2017;14(5).

7. Huo D, Hu H, Rhie SK, Gamazon ER, Cherniack AD, Liu J, et al. Comparison of Breast Cancer Molecular Features and Survival by African and European Ancestry in The Cancer Genome Atlas. JAMA Oncol 2017;3(12):1654–62. [PubMed: 28472234]

8. Gong Z, Wang J, Wang D, Buas MF, Ren X, Freudenheim JL, et al. Differences in microRNA expression in breast cancer between women of African and European ancestry. Carcinogenesis 2019;40(1):61–69. [PubMed: 30321299]

9. Zhao H, Shen J, Medico L, Wang D, Ambrosone CB, Liu S. A pilot study of circulating miRNAs as potential biomarkers of early stage breast cancer. PLoS One 2010;5(10):e13735. [PubMed: 21060830]

10. Sugita B, Gill M, Mahajan A, Duttargi A, Kirolikar S, Almeida R, et al. Differentially expressed miRNAs in triple negative breast cancer between African-American and non-Hispanic white women. Oncotarget 2016;7(48):79274–91. [PubMed: 27813494]

11. Song MA, Brasky TM, Marian C, Weng DY, Taslim C, Dumitrescu RG, et al. Racial differences in genome-wide methylation profiling and gene expression in breast tissues from healthy women. Epigenetics 2015;10(12):1177–87. [PubMed: 26680018]

12. Ahmad A, Azim S, Zubair H, Khan MA, Singh S, Carter JE, et al. Epigenetic basis of cancer health disparities: Looking beyond genetic differences. Biochim Biophys Acta Rev Cancer 2017;1868(1):16–28. [PubMed: 28108348]

13. Wang S, Dorsey TH, Terunuma A, Kittles RA, Ambs S, Kwabi-Addo B. Relationship between tumor DNA methylation status and patient characteristics in African-American and European-American women with breast cancer. PLoS One 2012;7(5):e37928. [PubMed: 22701537]

14. Mehrotra J, Ganpat MM, Kanaan Y, Fackler MJ, McVeigh M, Lahti-Domenici J, et al. Estrogen receptor/progesterone receptor-negative breast cancers of young African-American women have a higher frequency of methylation of multiple genes than those of Caucasian women. Clin Cancer Res 2004;10(6):2052–7. [PubMed: 15041725]

15. Audia JE, Campbell RM. Histone Modifications and Cancer. Cold Spring Harb Perspect Biol 2016;8(4):a019521. [PubMed: 27037415]

16. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;144(5):646–74. [PubMed: 21376230]

17. Deshmukh SK, Srivastava SK, Tyagi N, Ahmad A, Singh AP, Ghadhban AAL, et al. Emerging evidence for the role of differential tumor microenvironment in breast cancer racial disparity: a closer look at the surroundings. Carcinogenesis 2017;38(8):757–65. [PubMed: 28430867]

18. Deshmukh SK, Srivastava SK, Bhardwaj A, Singh AP, Tyagi N, Marimuthu S, et al. Resistin and interleukin-6 exhibit racially-disparate expression in breast cancer patients, display molecular association and promote growth and aggressiveness of tumor cells through STAT3 activation. Oncotarget 2015;6(13):11231–41. [PubMed: 25868978]

19. Lindner R, Sullivan C, Offor O, Lezon-Geyda K, Halligan K, Fischbach N, et al. Molecular phenotypes in triple negative breast cancer from African American patients suggest targets for therapy. PLoS One 2013;8(11):e71915. [PubMed: 24260093]

20. Incio J, Ligibel JA, McManus DT, Suboj P, Jung K, Kawaguchi K, et al. Obesity promotes resistance to anti-VEGF therapy in breast cancer by up-regulating IL-6 and potentially FGF-2. Sci Transl Med 2018;10(432).
21. Cao Y Angiogenesis modulates adipogenesis and obesity. J Clin Invest 2007;117(9):2362–8. [PubMed: 17786229]
22. Wu Q, Li B, Li Z, Li J, Sun S. Cancer-associated adipocytes: key players in breast cancer progression. J Hematol Oncol 2019;12(1):95. [PubMed: 31500658]
23. Cai Z, Liang Y, Xing C, Wang H, Hu P, Li J, et al. Cancerassociated adipocytes exhibit distinct phenotypes and facilitate tumor progression in pancreatic cancer. Oncol Rep 2019;42(6):2537–49. [PubMed: 31638193]
24. Ziegler KM, Considine RV, True E, Swartz-Basile DA, Pitt HA, Zyromski NJ. Adipocytes enhance murine pancreatic cancer growth via a hepatocyte growth factor (HGF)-mediated mechanism. Int J Surg 2016;28:179–84. [PubMed: 26957017]
25. Organ SL, Tsao MS. An overview of the c-MET signaling pathway. Ther Adv Med Oncol 2011;3(1 Suppl):S7–S19. [PubMed: 22128289]
26. Downward J Targeting RAS signalling pathways in cancer therapy. Nat Rev Cancer 2003;3(1):11–22. [PubMed: 12509763]
27. Ueki K, Matsuda S, Tobe K, Gotoh Y, Tamemoto H, Yachi M, et al. Feedback regulation of mitogen-activated protein kinase kinase kinase activity of c-Raf-1 by insulin and phorbol ester stimulation. J Biol Chem 1994;269(22):15756–61. [PubMed: 8195229]
28. Corbalan-Garcia S, Yang SS, Degenhardt KR, Bar-Sagi D. Identification of the mitogen-activated protein kinase phosphorylation sites on human Sos1 that regulate interaction with Grb2. Mol Cell Biol 1996;16(10):5674–82. [PubMed: 8816480]
29. Bandyopadhyay S, Pai SK, Hirota S, Hosobe S, Takano Y, Saito K, et al. Role of the putative tumor metastasis suppressor gene Drg-1 in breast cancer progression. Oncogene 2004;23(33):5675–81. [PubMed: 15184886]
30. Bandyopadhyay S, Wang Y, Zhan R, Pai SK, Watabe M, Iiizumi M, et al. The tumor metastasis suppressor gene Drg-1 down-regulates the expression of activating transcription factor 3 in prostate cancer. Cancer research 2006;66(24):11983–90. [PubMed: 17178897]
31. Xing F, Liu Y, Sharma S, Wu K, Chan MD, Lo HW, et al. Activation of the c-Met Pathway Mobilizes an Inflammatory Network in the Brain Microenvironment to Promote Brain Metastasis of Breast Cancer. Cancer Res 2016;76(17):4970–80. [PubMed: 27364556]
32. Consedine NS, Magai C, Conway F, Neugut AI. Obesity and awareness of obesity as risk factors for breast cancer in six ethnic groups. Obes Res 2004;12(10):1680–9. [PubMed: 15536232]
33. Lefebvre C, Rajbhandari P, Alvarez MJ, Bandaru P, Lim WK, Sato M, et al. A human B-cell interactome identifies MYB and FOXM1 as master regulators of proliferation in germinal centers. Mol Syst Biol 2010;6:377. [PubMed: 20531406]
34. Yoon CH, Kim MJ, Lee H, Kim RK, Lim EJ, Yoo KC, et al. PTTG1 oncogene promotes tumor malignancy via epithelial to mesenchymal transition and expansion of cancer stem cell population. J Biol Chem 2012;287(23):19516–27. [PubMed: 22511756]
35. Timofeeva OA, Zhang X, Ressom HW, Varghese RS, Kallakury BV, Wang K, et al. Enhanced expression of SOS1 is detected in prostate cancer epithelial cells from African-American men. Int J Oncol 2009;35(4):751–60. [PubMed: 19724911]
36. Field LA, Love B, Deyarmin B, Hooke JA, Shriver CD, Ellsworth RE. Identification of differentially expressed genes in breast tumors from African American compared with Caucasian women. Cancer 2012;118(5):1334–44. [PubMed: 21800289]
37. Lepri F, De Luca A, Stella L, Rossi C, Baldassarre G, Pantaleoni F, et al. SOS1 mutations in Noonan syndrome: molecular spectrum, structural insights on pathogenic effects, and genotype-phenotype correlations. Hum Mutat 2011;32(7):760–72. [PubMed: 21387466]
38. Swanson KD, Winter JM, Reis M, Bentires-Alj M, Greulich H, Grewal R, et al. SOS1 mutations are rare in human malignancies: implications for Noonan Syndrome patients. Genes Chromosomes Cancer 2008;47(3):253–9. [PubMed: 18064648]

39. Troutman SM, Sissung TM, Cropp CD, Venzon DJ, Spencer SD, Adesunloye BA, et al. Racial disparities in the association between variants on 8q24 and prostate cancer: a systematic review and meta-analysis. Oncologist 2012;17(3):312–20. [PubMed: 22382457]
40. Datta S, Sherva RM, De La Cruz M, Long MT, Roy P, Backman V, et al. Single Nucleotide Polymorphism Facilitated Down-Regulation of the Cohesin Stromal Antigen-1: Implications for Colorectal Cancer Racial Disparities. Neoplasia 2018;20(3):289–94. [PubMed: 29471289]
41. Bhardwaj A, Srivastava SK, Khan MA, Prajapati VK, Singh S, Carter JE, et al. Racial disparities in prostate cancer: a molecular perspective. Front Biosci (Landmark Ed) 2017;22:772–82. [PubMed: 27814645]
42. Bensen JT, Graff M, Young KL, Sethupathy P, Parker J, Pecot CV, et al. A survey of microRNA single nucleotide polymorphisms identifies novel breast cancer susceptibility loci in a case-control, population-based study of African-American women. Breast Cancer Res 2018;20(1):45. [PubMed: 29871690]
43. Moncini S, Bonati MT, Morella I, Ferrari L, Brambilla R, Riva P. Differential allelic expression of SOS1 and hyperexpression of the activating SOS1 c.755C variant in a Noonan syndrome family. Eur J Hum Genet 2015;23(11):1531–7. [PubMed: 25712082]
44. Wang X, Ji P, Zhang Y, LaComb JF, Tian X, Li E, et al. Aberrant DNA Methylation: Implications in Racial Health Disparity. PLoS One 2016;11(4):e0153125. [PubMed: 27111221]
45. Raut JR, Guan Z, Schrotz-King P, Brenner H. Whole-blood DNA Methylation Markers for Risk Stratification in Colorectal Cancer Screening: A Systematic Review. Cancers (Basel) 2019;11(7).
46. Fraser HB, Lam LL, Neumann SM, Kobor MS. Population-specificity of human DNA methylation. Genome Biol 2012;13(2):R8. [PubMed: 22322129]
47. Benevolenskaya EV, Islam AB, Ahsan H, Kibriya MG, Jasmine F, Wolff B, et al. DNA methylation and hormone receptor status in breast cancer. Clin Epigenetics 2016;8:17. [PubMed: 26884818]
48. Liu M, Roth A, Yu M, Morris R, Bersani F, Rivera MN, et al. The IGF2 intronic miR-483 selectively enhances transcription from IGF2 fetal promoters and enhances tumorigenesis. Genes Dev 2013;27(23):2543–8. [PubMed: 24298054]
49. Lincoln KD, Abdou CM, Lloyd D. Race and socioeconomic differences in obesity and depression among Black and non-Hispanic White Americans. J Health Care Poor Underserved 2014;25(1):257–75. [PubMed: 24509025]
50. Picon-Ruiz M, Morata-Tarifa C, Valle-Goffin JJ, Friedman ER, Slingerland JM. Obesity and adverse breast cancer risk and outcome: Mechanistic insights and strategies for intervention. CA Cancer J Clin 2017;67(5):378–97. [PubMed: 28763097]
51. Bousquenaud M, Fico F, Solinas G, Ruegg C, Santamaria-Martinez A. Obesity promotes the expansion of metastasis-initiating cells in breast cancer. Breast Cancer Res 2018;20(1):104. [PubMed: 30180888]
52. Wahl S, Drong A, Lehne B, Loh M, Scott WR, Kunze S, et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. Nature 2017;541(7635):81–86. [PubMed: 28002404]
53. Aras S, Zaidi MR. TAMeless traitors: macrophages in cancer progression and metastasis. Br J Cancer 2017;117(11):1583–91. [PubMed: 29065107]
54. Lauterbach MA, Wunderlich FT. Macrophage function in obesity-induced inflammation and insulin resistance. Pflugers Arch 2017;469(3–4):385–96. [PubMed: 28233125]
55. Tashiro H, Takahashi K, Sadamatsu H, Kato G, Kurata K, Kimura S, et al. Saturated Fatty Acid Increases Lung Macrophages and Augments House Dust Mite-Induced Airway Inflammation in Mice Fed with High-Fat Diet. Inflammation 2017;40(3):1072–86. [PubMed: 28365872]
56. Sawant KV, Xu R, Cox R, Hawkins H, Sbrana E, Kolli D, et al. Chemokine CXCL1-Mediated Neutrophil Trafficking in the Lung: Role of CXCR2 Activation. J Innate Immun 2015;7(6):647–58. [PubMed: 26138727]
57. Manicone AM, Gong K, Johnston LK, Giannandrea M. Diet-induced obesity alters myeloid cell populations in naive and injured lung. Respir Res 2016;17:24. [PubMed: 26956558]
58. Gysin S, Salt M, Young A, McCormick F. Therapeutic strategies for targeting ras proteins. Genes Cancer 2011;2(3):359–72. [PubMed: 21779505]

59. Zhu Z, Golay HG, Barbie DA. Targeting pathways downstream of KRAS in lung adenocarcinoma. Pharmacogenomics 2014;15(11):1507–18. [PubMed: 25303301]

60. Hillig RC, Sautier B, Schroeder J, Moosmayer D, Hilpmann A, Stegmann CM, et al. Discovery of potent SOS1 inhibitors that block RAS activation via disruption of the RAS-SOS1 interaction. Proc Natl Acad Sci U S A 2019;116(7):2551–60. [PubMed: 30683722]

61. Yu Y, Nie Y, Feng Q, Qu J, Wang R, Bian L, et al. Targeted Covalent Inhibition of Grb2-Sos1 Interaction through Proximity-Induced Conjugation in Breast Cancer Cells. Mol Pharm 2017;14(5):1548–57. [PubMed: 28060514]

62. Cussac D, Vidal M, Leprince C, Liu WQ, Cornille F, Tiraboschi G, et al. A Sos-derived peptidimer blocks the Ras signaling pathway by binding both Grb2 SH3 domains and displays antiproliferative activity. FASEB J 1999;13(1):31–8. [PubMed: 9872927]

63. Oi N, Chen H, Ok Kim M, Lubet RA, Bode AM, Dong Z. Taxifolin suppresses UV-induced skin carcinogenesis by targeting EGFR and PI3K. Cancer Prev Res (Phila) 2012;5(9):1103–14. [PubMed: 22805054]

64. Tumurkhuu M, Saitoh M, Takita J, Mizuno Y, Mizuguchi M. A novel SOS1 mutation in Costello/CFC syndrome affects signaling in both RAS and PI3K pathways. J Recept Signal Transduct Res 2013;33(2):124–8. [PubMed: 23528009]

65. Liu K, Jiang T, Ouyang Y, Shi Y, Zang Y, Li N, et al. Nuclear EGFR impairs ASPP2-p53 complex-induced apoptosis by inducing SOS1 expression in hepatocellular carcinoma. Oncotarget 2015;6(18):16507–16. [PubMed: 25980493]

Significance

Findings elucidate the signaling network of SOS1-mediated metastasis in AA patients, from the epigenetic upregulation of SOS1 to the identification of taxifolin as a potential therapeutic strategy against SOS1-driven tumor progression.

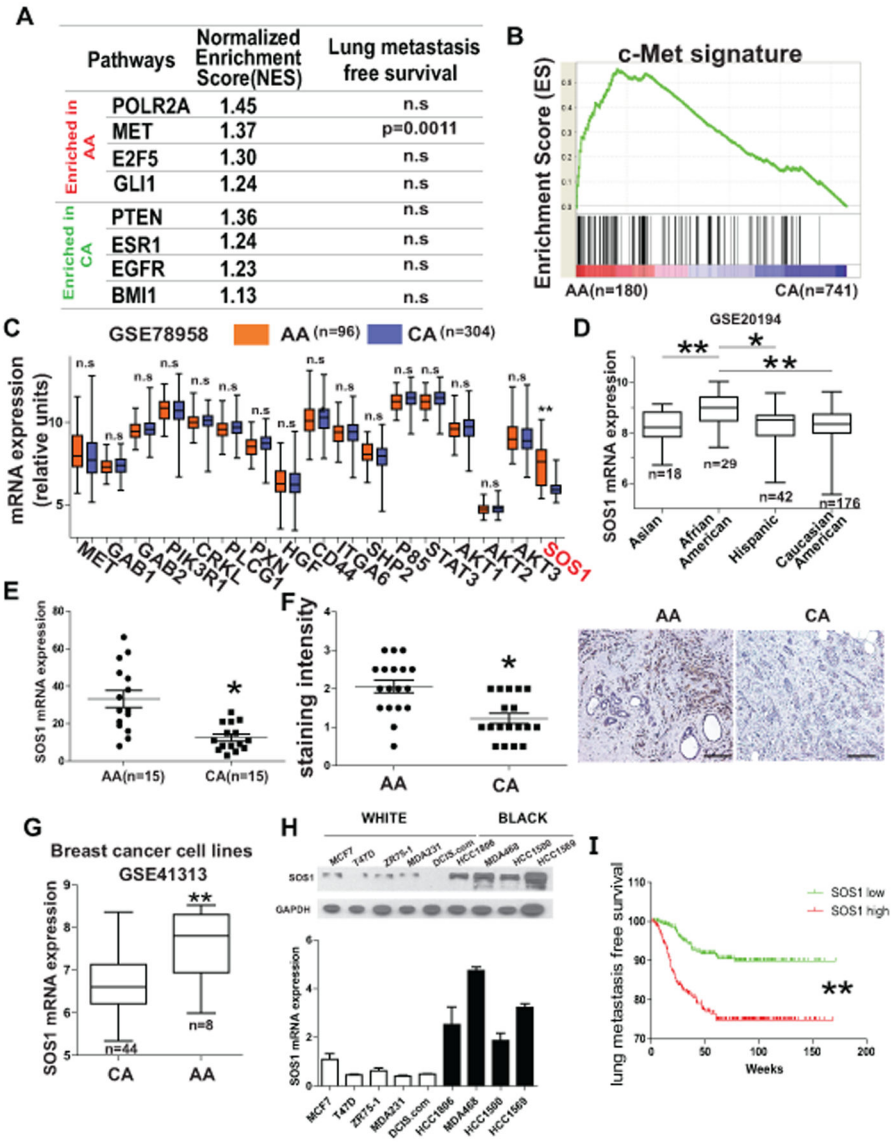


Figure 1. c-Met-SOS1 pathway is activated in AA breast cancer.

(A) GSEA was performed using the TCGA database for 110 breast cancer specific signaling pathways. List of enriched gene signatures in AA and CA patients are shown. **(B)** The result of GSEA for the c-Met signature in AA and CA breast cancer patients. **(C)** Expression of 17 genes directly associated with the c-Met pathway in AA and CA patients were examined using GSE78958 database. **(D)** SOS1 expression was examined in breast cancer patients with different races using GSE20194. **(E)** Expression of the *SOS1* gene was examined for AA and CA breast cancer patients by Taqman PCR. **(F)** SOS1 protein levels in AA and CA breast cancer patients were examined by IHC. Right panels show representative images from two patients each from AA and CA. scale bar : 100 μm**(G)** SOS1 expression was examined in various breast cancer cell lines derived from AA or CA patients. **(H)** mRNA and protein expression of SOS1 were examined for various breast cancer cell lines derived from AA and CA patients. **(I)** Lung metastasis-free survival was examined for 710 breast cancer patients

with high or low SOS1 expression using the previously published cohort data (31). * P<0.05, ** P<0.001.

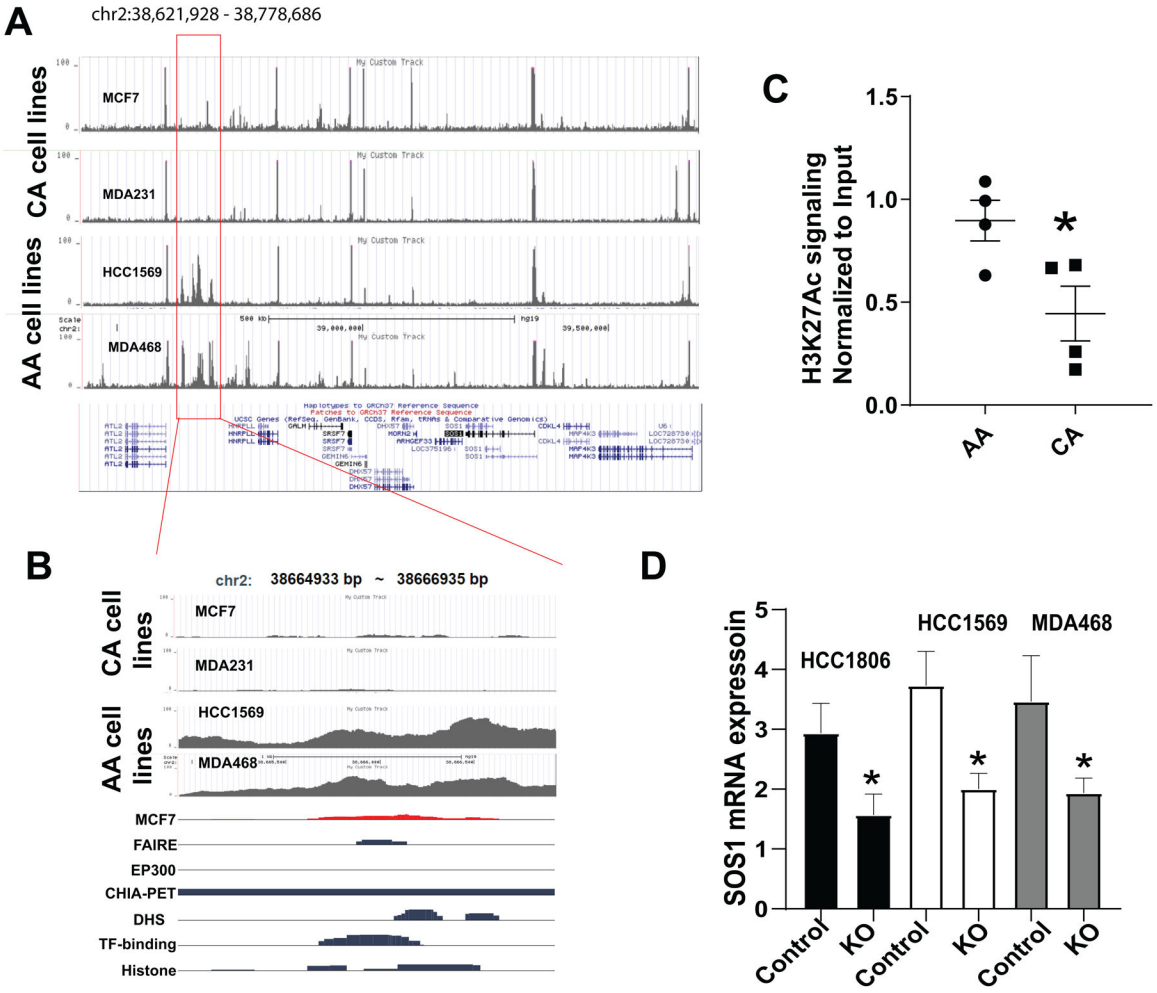
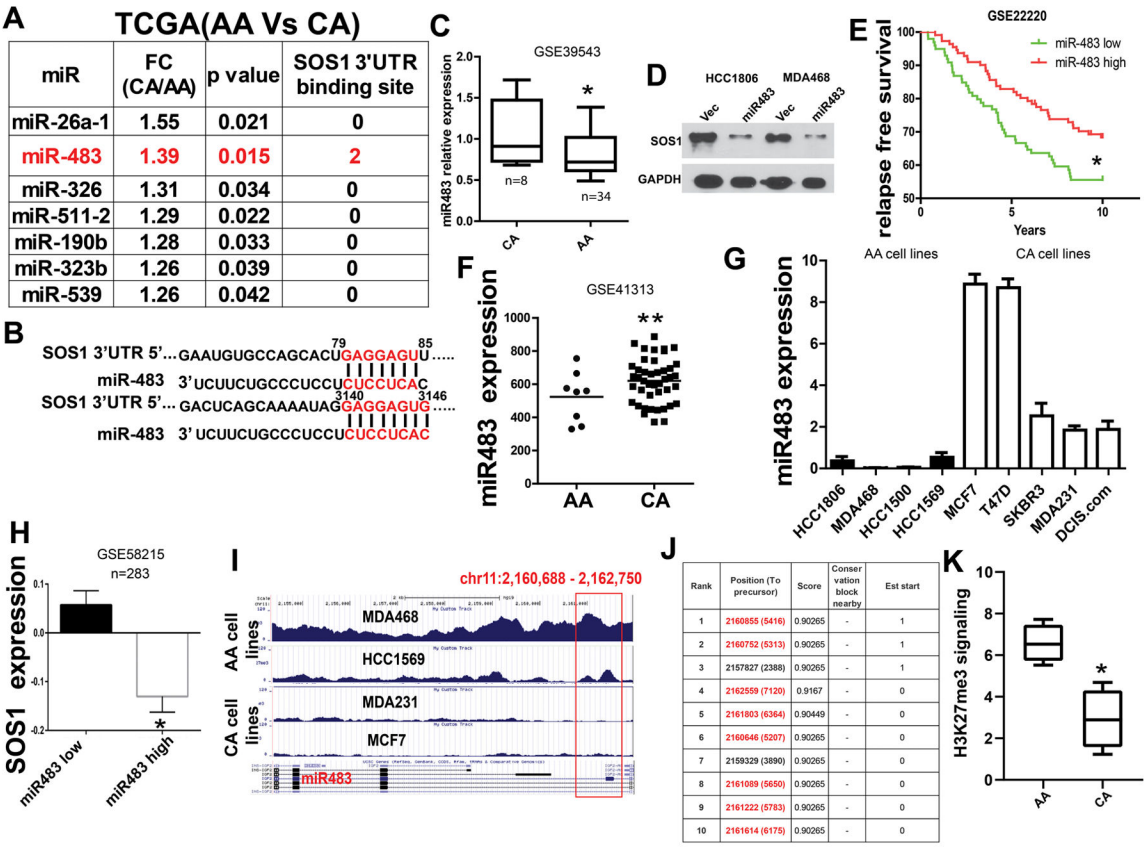


Figure 2. SOS1 is regulated by a super enhancer.

(A) ChIP-seq data of H3K27Ac signaling for two AA and two CA cell lines. The region most significantly changed between AA and CA is indicated by a red box (upper panel). **(B)**The red boxed region in Fig. 2a is expanded, and the super enhancer region of SOS1 is identified based on EnhanerAtlas. **(C)** H3K27Ac ChIP-PCR analysis for the super enhancer region was performed for clinical samples from AA and CA breast cancer patients. **(D)** Super enhancer was deleted using the paired-end sgRNA-mediated CRISPR technology in three AA cell lines, and the expression of SOS1 was examined by qRT-PCR. * $p<0.05$, ** $p<0.001$.



miR	FC (CA/AA)	p value	SOS1 3'UTR binding site
miR-26a-1	1.55	0.021	0
miR-483	1.39	0.015	2
miR-326	1.31	0.034	0
miR-511-2	1.29	0.022	0
miR-190b	1.28	0.033	0
miR-323b	1.26	0.039	0
miR-539	1.26	0.042	0

Figure 3. Down regulation of miR-483 promotes the SOS1 expression in AA patients.

(A) List of microRNAs that are differentially expressed in AA and CA breast cancer patients based on TCGA database. The number of potential binding sites of each microRNA on 3'UTR of the SOS gene was examined by TargetScan. **B)** Two predicted miR-483 binding sites on 3'UTR of the *SOS1* gene by TargetScan is shown. **(C)** Expression of miR-483 was examined in AA and CA breast cancer patients using GSE39543. **(D)** Ectopic expression of miR-483 suppressed SOS1 in two AA breast cancer cell lines. The expression was examined by western blot. **(E)** Relapse-free survival of breast cancer patients with high or low miR-483 expression was examined using GSE22220 (n=210). **(F)** miR-483 expression in breast tumor cell lines derived from AA or CA patients was examined using GSE41313. **(G)** miR-483 expression in the indicated breast cancer cell lines was measured by Taq-man PCR. **(H)** Reverse correlation between SOS1 and miR-483 expression was examined in breast cancer patients using GSE58215. **(I)** H3K27me3 signaling in the promoter (red boxed) of miR-483 of the indicated AA and CA cell lines. **(J)** Potential transcriptional start sites of miR-483 predicted by miRStart. **(K)** H3K27me3 ChIP-PCR analysis for the promoter region of miR-483 was done using clinical samples from AA (n=5) and CA (n=5) breast cancer patients. * $p<0.05$, ** $p<0.001$.

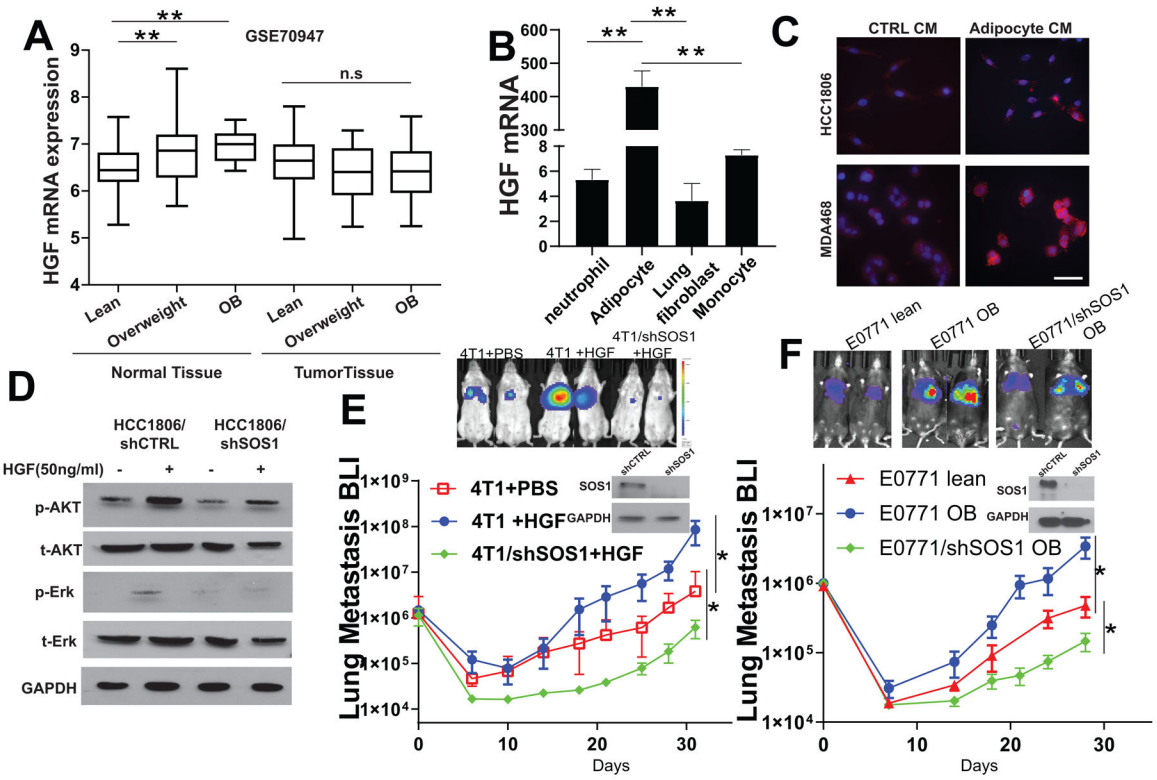


Figure 4. Knockdown of SOS1 suppressed lung metastasis in obese mice.

(A) Expressions of HGF in breast tumor tissues and matched adjacent normal tissues were analyzed using GSE70947. **(B)** Expression of HGF was examined by qRT-PCR in various tissues. **(C)** MDA468 and HCC1806 cells were treated with or without conditioned medium from adipocytes for 30mins followed by immunocytochemical analysis for phosphor-c-Met. Scale bar: 20μm. **(D)** Western blot analyses for phosphor-AKT and phosphor-Erk in HGF-treated HCC1806 cells with or without knockdown of SOS1. **(E)** 4T1 or 4T1/shSOS1 cells ($2X10^5$) were implanted by tail vein injection into BALB/C mice, and they were treated with or without recombinant HGF (2 mg/kg twice/weekly). Lung metastasis signal was measured by IVIS. Knockdown of SOS1 was verified by western blot (inserted fig). **(F)** E0771 cells ($2X10^5$) were implanted through tail vein into normal or obese C57/B6 mice followed by measuring lung metastasis signal at different time points. Knockdown of SOS1 was verified by western blot (inserted fig). * $p<0.05$, ** $p<0.001$.

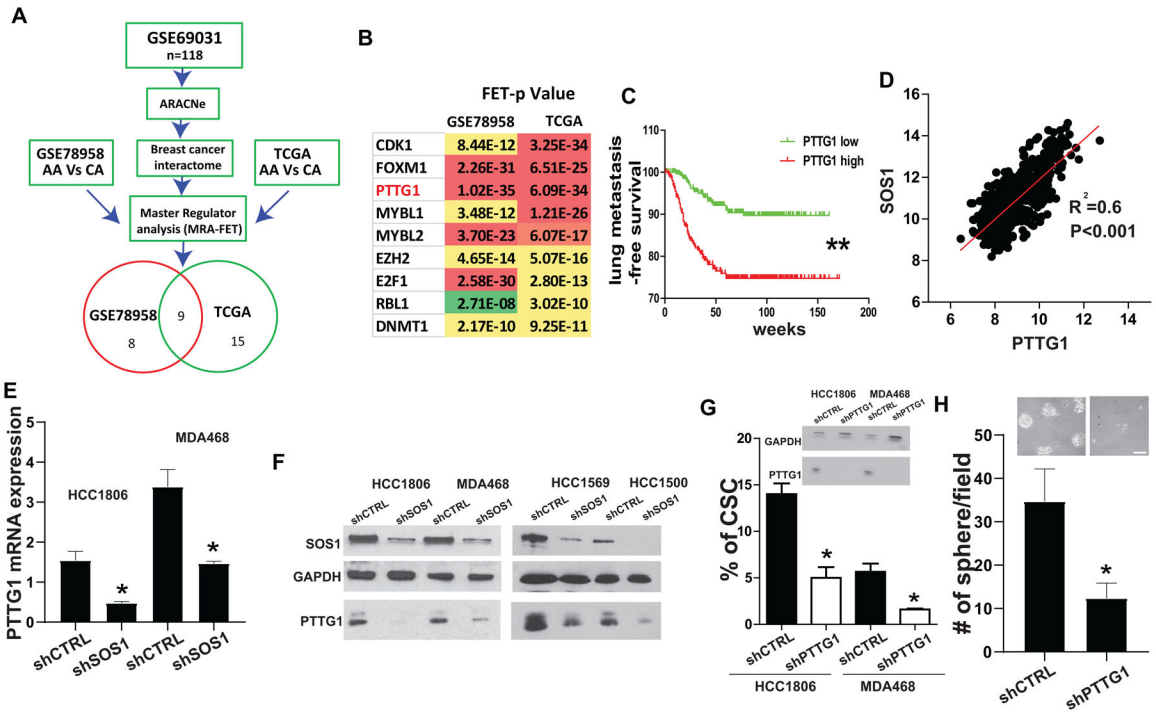
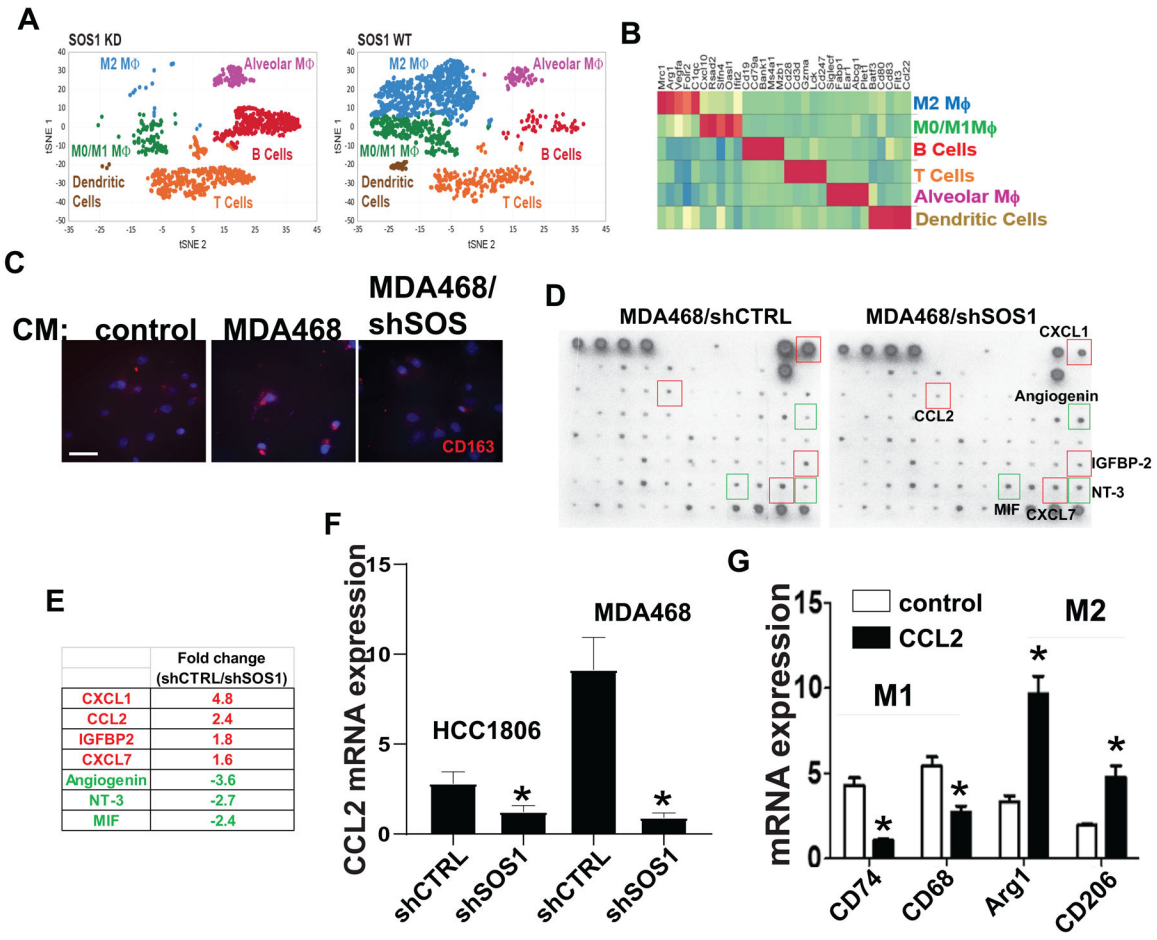


Figure 5. SOS1 mediated PTTG1 upregulation and enhanced stemness of cancer cells.

(A) Flowchart for identifying racial-specific genes using the tool for Master regulator analysis from GeWorkbench. **(B)** List of most significantly enriched master regulators in AA vs CA. **(C)** Lung metastasis-free survival of patients with high or low PTTG1 expression was examined. **(D)** Correlation between SOS1 and PTTG1 expressions in breast cancer patients was examined using the previously published cohort database (n=710). **(E)** PTTG1 expression was examined in cancer cells with or without knock-down of SOS1 by qRT-PCR. **(F)** PTTG1 expression was examined in cancer cells with or without knock-down of SOS1 by Western Blot. **(G)** Cancer stem cell population was measured for HCC1806 or MDA468 cells with or without knockdown of PTTG1 by FACS. Knock down of PTTG1 was examined by Western Blot (inserted figure). **(H)** Mammosphere forming assay was performed using MDA468 cells with or without knockdown of PTTG1. Scale bar: 100μm * $p<0.05$, ** $p<0.001$.



	Fold change (shCTRL/shSOS1)
CXCL1	4.8
CCL2	2.4
IGFBP2	1.8
CXCL7	1.6
Angiogenin	-3.6
NT-3	-2.7
MIF	-2.4

Figure 6. SOS1-induced CCL2 promotes lung metastasis by polarizing macrophages

(A) E0771 cells with or without knockdown of the *SOS1* was injected into obese C57/B6 mice via tail vein. After 30 days, lungs were removed, dissociated, and CD45+ cells were then selected by MACS sorting. The isolated CD45+ cells were then subjected to single cell RNA sequencing analysis. **(B)** Expression profile of immune specific gene sets was analyzed for different types of immune cells. **(C)** CD163 expression in THP-1 cells that were incubated with the conditioned medium derived from cells with or without knockdown of SOS1 was examined by immunocytochemistry. Scale bar: 20μm. **(D)** Human growth factor and cytokine array analysis (Raybiotech) was performed for the conditioned medium of MDA468 with or without knockdown of SOS1. **(E)** List of secretary factors that most up- or down-regulated by SOS1. **(F)** qRT-PCR analysis for CCL2 in breast tumor cells with or without knockdown of SOS1. **(G)** M1- and M2-associated genes were examined in THP-1 cells treated with or without recombinant CCL2, by qRT-PCR * $p<0.05$, ** $p<0.001$.

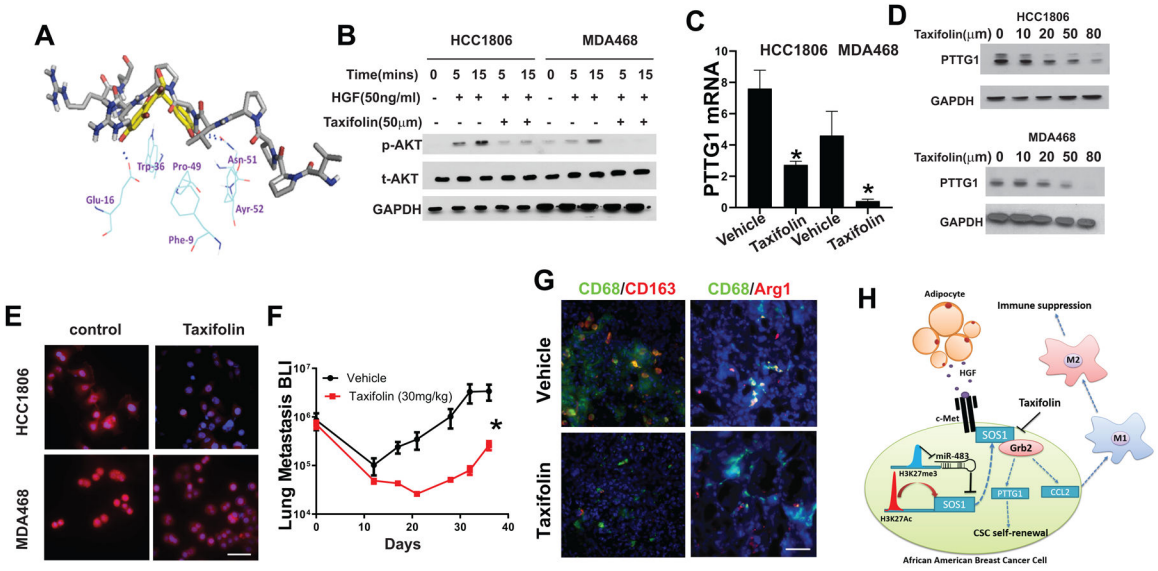


Figure 7. Taxifolin suppresses lung metastasis by disrupting SOS1-Grb2 interaction.

(A) Computational docking analysis showing that taxifolin (yellow) blocks the interaction of SOS1 (gray) and Grb2 (light blue). **(B)** Western blot analysis for p-AKT in AA cancer cells with or without treatment of HGF and taxifolin. **(C)** mRNA level of PTTG1 was examined in tumor cells treated with or without taxifolin by qRT-PCR. **(D)** PTTG1 protein expression was examined by western blot in tumor cells treated with or without taxifolin. **(E)** CCL2 expression was examined by immunocytochemical analysis in tumor cells that were treated with or without taxifolin. Scale bar: 20μm **(F)** E0771 cells (2×10^5) were implanted into obese C57/B6 mice through tail vein followed by taxifolin treatment (30 mg/kg twice a week). Lung metastasis signal was measured at different time points by IVIS. **(G)** M1 and M2 macrophages in the lung metastatic sites from the mice treated with or without taxifolin were examined by immunocytochemistry using anti-Arg1/CD68 and anti-CD163/CD68 antibodies. Scale bar: 20μm **(H)** Proposed mechanism of epigenetic regulation of SOS1 and its role in lung metastasis. * $p<0.05$, ** $p<0.001$.