MaTAR25 LncRNA Regulates the *Tensin1* Gene to Impact Breast Cancer Progression

Kung-Chi Chang[1,2], Sarah D. Diermeier[1,8,*], Allen T. Yu[1,3,*], Lily D. Brine[1], Suzanne Russo[1], Sonam Bhatia[1], Habeeb Alsudani[1], Karen Kostroff[4], Tawfiqul Bhuiya[5], Edi Brogi[6], C. Frank Bennett[7], Frank Rigo[7], and David L. Spector[1,2,3,#]

[1]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724
[2]Molecular and Cellular Biology Program, Stony Brook University, Stony Brook, NY 11794
[3]Genetics Program, Stony Brook University, Stony Brook, NY 11794
[4]Department of Surgical Oncology, Northwell Health, Lake Success, NY 11042
[5]Department of Pathology, Northwell Health, Lake Success, NY 11042
[6]Memorial Sloan Kettering Cancer Center, New York, NY 10065
[7]Ionis Pharmaceuticals, Carlsbad, CA 92010

[8]Current address: Department of Biochemistry, University of Otago, Dunedin 9016, New Zealand

[*]These authors contributed equally to this work.

[#]Correspondence: spector@cshl.edu

David L.Spector is a consultant to, and receives research support from, Ionis Pharmaceuticals

SUMMARY

Misregulation of long non-coding RNA genes has been linked to a wide variety of cancer types. Here we report on Mammary Tumor Associated RNA 25 (*MaTAR25*), a nuclear enriched and chromatin associated lncRNA that plays a role in mammary tumor cell proliferation, migration, and invasion, both *in vitro* and *in vivo*. *MaTAR25* functions by interacting with purine rich element binding protein B (PURB), and associating with a major downstream target gene *Tensin 1* (*Tns1*) to regulate its expression in *trans*. Knockout of *MaTAR25* results in down-regulation of *Tns1* leading to a reorganization of the actin cytoskeleton, and a reduction of focal adhesions and microvilli. The human ortholog of *MaTAR25*, *LINC01271*, is upregulated with human breast cancer stage and metastasis.

SIGNIFICANCE

LncRNAs have great potential to reveal new regulatory mechanisms of function as well as having exciting therapeutic capacity given their ease of being targeted by nucleic acid drugs. Our study of *MaTAR25,* and its human ortholog *LINC01271,* reveal an unexpected function of this lncRNA in breast cancer progression by regulating *Tns1* gene expression, whose protein product is a critical component of focal adhesions linking signaling between the extracellular matrix and the actin cytoskeleton. We identified *LINC01271 as* the human ortholog of *MaTAR25*, and importantly, increased expression of *LINC01271* is associated with poor patient prognosis and cancer metastasis. Our findings demonstrate that *LINC01271* represents an exciting therapeutic target to alter breast cancer progression.

INTRODUCTION

Breast cancer is the most common cancer among women in the United States and world-wide, with an estimated 268,600 new cases of invasive disease in women in the United States in 2019 (1). Although breast cancer mortality has been decreasing over the past two decades, it is still the second leading cause of cancer deaths in American women accounting for 15% of all cancer deaths. Breast tumors can be classified into multiple subtypes based on histological evaluation and the most frequent type of breast tumors are ductal carcinomas, which affect the milk ducts of the breast. Ductal carcinomas can further be separated into two groups: non-invasive ductal carcinoma *in situ* (DCIS), and invasive ductal carcinoma (IDC) which accounts for 75% of all breast cancers (2, 3). Breast cancer is recognized as a heterogeneous disease and molecular classification of invasive breast carcinomas can stratify tumors into informative subtypes and provide key prognostic signatures. In addition to traditional pathological characterization and immunohistochemistry (IHC) to examine protein levels of markers such as estrogen receptor (ER), progesterone receptor (PR) and epidermal growth factor receptor-2 (HER2), additional studies evaluating genomic rearrangements and molecular expression profiles of breast cancers have provided further genetic insights to better understand the disease (4-6). These approaches have identified six major molecular subtypes of breast cancer (luminal A, luminal B, HER2-enriched, triple negative/basal-like, normal breast-like, and claudin-low) (7), each displaying different phenotypic and molecular features and which have distinct clinical outcomes.

In recent years, large scale genome-wide studies indicated that thousands of RNAs can be transcribed from the human and mouse genomes that lack protein-coding

capacity (8-10). In particular, long non-coding RNAs (lncRNAs) with a length ≥ 200 nucleotides have been suggested to play key roles in a diverse range of biological processes (11-13). Most lncRNAs are capped, spliced, and poly-adenylated (8). In addition, many lncRNAs are expressed in a tissue-specific and/or cell type specific manner, and are involved in various gene regulatory pathways (14, 15). Furthermore, misregulation of lncRNA expression has been linked to various diseases including neuromuscular diseases, developmental disorders, neurodegenerative diseases, and cancers (16-20). Several lncRNAs have been implicated as regulatory molecules in breast cancer progression and metastasis through different mechanisms (21, 22). For example, the *HOX* antisense intergenic RNA (*HOTAIR*) is overexpressed in primary breast tumors and can alter the localization pattern of Polycomb repressive complex 2 (PRC2) and histone methylation to regulate gene expression in breast carcinoma cells impacting breast cancer progression and metastasis (23). Recent findings suggest that the lncRNA breast cancer anti-estrogen resistance 4 (*BCAR4*) (24) can control GLI family zinc finger 2 (GLI2) gene expression to promote cancer cell migration by interacting with Smad nuclear interacting protein 1 (SNIP1) and serine/threonine-protein phosphatase 1 regulatory subunit 10 (PNUTS). Targeting *BCAR4* by locked nucleic acids (LNA) in mouse models significantly affects cancer cell invasion and reduces lung metastases (25). Genetic knockout or ASO-mediated knockdown of *Metastasis Associated Lung Adenocarcinoma Transcript 1 (Malat1)* was shown to result in differentiation of primary mammary tumors and a significant reduction in metastasis (26). In addition to transcriptional regulation, lncRNAs can have other regulatory roles. For example, the lncRNA *PVT1* has been shown to stabilize the Myc oncoprotein in

breast cancer cells (27), and the lncRNA NKILA can interact with and stabilize the NF-κB/IκB complex and inhibit breast cancer metastasis (28). However, for the majority of lncRNAs, the exact function and molecular mechanism of action in breast cancers still awaits detailed characterization. Previously, we performed an RNA sequencing (RNA-seq) screen to identify differentially expressed lncRNAs between mammary tumor cells and normal mammary epithelial cells. From this screen, we identified 30 previously uncharacterized lncRNAs as Mammary Tumor Associated RNAs (*MaTARs*) 1-30 (29).

Here, we examined the role of *MaTAR25* in mammary tumor progression and metastasis. We find that genetic knockout of *MaTAR25* in highly aggressive 4T1 triple negative (ER-, PR-, HER2-) mammary carcinoma cells results in a reduction in cell proliferation, migration, and invasion. Knockout cells transplanted into the mammary fat pad of BALB/c mice results in a significant decrease in tumor growth as compared to 4T1 control cells. Further, tail vein injection of luciferase labeled *MaTAR25* knockout cells showed reduced homing to the lungs and a significant decrease in metastatic nodules. In a complementary study, antisense oligonucleotide (ASO) mediated knockdown (KD) of *MaTAR25* in the MMTV-Neu-NDL mouse model resulted in a significant decrease in tumor growth and a reduction in lung metastases. Analysis of the molecular function of *MaTAR25* indicates that it regulates the *Tns1* gene at the transcriptional level. Loss of *MaTAR25* results in a reduction of *Tns1* at the RNA and protein levels and a subsequent reorganization of the actin cytoskeleton and a reduction in focal adhesions and microvilli. Together, our data reveal *MaTAR25*, and its identified human ortholog *LINC01271*, as an exciting therapeutic candidate whose expression can be altered to impede breast cancer progression and metastasis.

RESULTS

Characterization of *MaTAR25*, a nuclear enriched lncRNA

We previously performed an RNA-seq screen to identify lncRNAs over-expressed in mammary tumors vs normal mammary epithelial cells as a means to identify potential candidates involved in mammary cancer progression, and to explore their potential as therapeutic targets or key biomarkers in human breast cancer (29). Among those lncRNA genes identified, the *MaTAR25* gene on mouse chromosome 2 was originally annotated as 1200007C13Rik and it encodes a single transcript containing two exons (Fig. 1A). *MaTAR25* is overexpressed in mammary tumors in the MMTV-Neu-NDL (HER2 subtype) model compared to normal mammary epithelial cells and it is also upregulated in luminal and triple negative sub-types of mammary cancer (29). Analysis of ENCODE and FANTOM5 RNA-seq data has shown there is little to no expression of *MaTAR25* in normal mouse tissues compared to MMTV-Neu-NDL tumor cells (29, 30) (Supplementary Fig. S1A). The full length *MaTAR25* transcript was determined to be 1,978 nucleotides by 5' and 3′ rapid amplification of cDNA ends (RACE) and Sanger sequencing (Fig. 1B), which was further confirmed by Northern blot analysis (Fig. 1C).

According to three independent computational coding potential prediction programs, the *MaTAR25* RNA transcript has very low protein coding potential and is suggested to be a non-coding RNA (Supplementary Fig. S1B-S1D). However, there is one predicted open reading frame (ORF) with the potential to generate a 123 amino acid peptide (~13 kDa). In order to assess whether a peptide is encoded by the

MaTAR25 transcript we performed *in vitro* transcription and translation. Compared to a luciferase DNA control (expected size 61 kDa) and a *Xenopus laevis* Histone H2B (HISTH2B) expressing plasmid control (expected size 14 kDa), there was no detectable peptide generated from a plasmid that contained the *MaTAR25* sequence (Fig. 1D). Together, these computational and experimental results confirm that *MaTAR25* does not make a peptide, and thus is a bona fide lncRNA.

In order to determine the localization and abundance of *MaTAR25* we performed single molecule RNA fluorescence in situ hybridization (smRNA-FISH) to detect *MaTAR25* RNA transcripts within MMTV-PyMT (luminal B) and MMTV-Neu-NDL (Her2/neu+) primary mammary tumor cells. The majority of *MaTAR25* transcripts were detected in cell nuclei (Fig. 1E) and each nucleus contained ~10-15 transcript foci. Thus, *MaTAR25* is a nuclear-enriched lncRNA with a potential role in the regulation of gene expression in mammary cancer cells.

MaTAR25 knockout decreases 4T1 cell viability/migration/invasion

To assess the functional role of *MaTAR25*, we proceeded to genetically knockout (KO) *MaTAR25* in highly aggressive 4T1 triple negative (ER-, PR-, HER2-) mammary carcinoma cells using CRISPR/Cas9. We designed gRNA pairs targeting various regions upstream and downstream of the transcription start site (TSS) of *MaTAR25* to create a genomic deletion (Fig. 2A and Supplementary Fig. S2A). *MaTAR25* knockout clones were single cell sorted and selected by Sanger sequencing, qRT-PCR, as well as smRNA-FISH (Fig. 2A and Supplementary Fig. S2B).

After selecting several *MaTAR25* KO clones, we evaluated them for alterations

in cell viability, migration, and invasion as compared to 4T1 control cells. *MaTAR25* KO cells exhibited a significant decrease of 50% in cell viability as compared to 4T1 control cells (Fig. 2B). To further investigate this phenotype, we performed BrdU labeling and FACS analysis, which demonstrated a two-fold increase in G2 cells suggesting that the decreased proliferation phenotype is most likely the result of a lengthened G2 phase (Supplementary Fig. S2C). As cell migration and invasion are critical processes associated with metastasis we were interested in determining whether *MaTAR25* loss might play a role in these events. We used a live cell tracking assay to assess cell migration and we found a 40% reduction in cell motility upon loss of *MaTAR25* (Fig. 2C and Supplementary Fig. S2D). A wound healing assay also corroborated the observed difference in cell migration between 4T1 control and *MaTAR25* KO cells (Supplementary Fig. S2E). Finally, we used a Boyden chamber invasion assay and found that loss of *MaTAR25* resulted in a 45% reduction in invasion ability as compared to 4T1 control cells (Fig. 2D).

In order to exclude the possibility that the phenotypes observed in *MaTAR25* KO cells were caused by disrupting chromatin structure rather than specific loss of the *MaTAR25* transcript, we generated single cell ectopic overexpression clones of *MaTAR25* in 4T1 *MaTAR25* KO cells. Ectopic expression of *MaTAR25* rescued both the cell proliferation and invasion phenotypes (Fig. 2E-2F), indicating that *MaTAR25* RNA plays an important role in these processes *in situ,* and likely exhibits its effect *in trans*. Hence, *MaTAR25* appears to be an important lncRNA impacting mammary tumor cell growth and critical aspects of metastasis. To further explore *MaTAR25's* downstream targets, we performed RNA-seq to identify differentially expressed genes

by comparing *MaTAR25* KO cells with 4T1 control cells (Supplementary Table S1). Pathway analysis of the differentially expressed genes by Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Set Enrichment Analysis (GSEA) revealed alteration in cell cycle and DNA related processes, both related to the phenotypes we observed in *MaTAR25* KO cells (Supplementary Fig. S2F).

MaTAR25* knockout decreases tumor progression/metastasis *in vivo

In order to further evaluate the functional impact and the therapeutic potential of *MaTAR25* in mammary tumor progression, we performed multiple *in vivo* studies. Injection of *MaTAR25* 4T1 KO cells into the mammary fat pad of BALB/c mice resulted in a significant 56% decrease in tumor growth at day 28, compared to the 4T1 control injected group (Fig. 3A-3B). In addition, we performed tail vein injection using *MaTAR25* KO cells expressing a luciferase reporter to track cancer cell homing and metastasis to the lungs in BALB/c mice. The *in vivo* bioluminescence signal in the lungs of mice injected with *MaTAR25* KO cells was reduced (Supplementary Fig. S3A) compared to those injected with 4T1 control cells. At day 21, the number of metastatic nodules in lung samples collected from the *MaTAR25* KO group was also significantly decreased by 62% compared to the 4T1 control group (Fig. 3C).

As a complementary approach to CRISPR/Cas9 KO we designed a series of *MaTAR25* specific antisense oligonucleotides (ASOs), (16mers) comprised of phosphorothioate-modified short S-cEt (S-2 ´ -O-Et-2 ´ ,4 ´ -bridged nucleic acid) gapmer chemistry (31-33). We individually screened multiple ASOs targeting *MaTAR25* to identify the most effective ASOs in terms of knockdown (KD) efficiency by qRT-PCR

after 48 hours and 72 hours of ASO treatment in 4T1 cells. The two most effective *MaTAR25* ASOs achieved a knockdown ranging from 70-90% (Supplementary Fig. S1E). When comparing *MaTAR25* ASO treated cells to mock or scrambled ASO (scASO) treated 4T1 control cells after 72 hours, we found a significant decrease in cell viability using cell counting assays (-45% for ASO1 and -38% for ASO2) (Supplementary Fig. S1F), consistent with our KO studies indicating that *MaTAR25* has a role in mammary cancer cell proliferation.

Furthermore, to assess the therapeutic potential of reducing the level of *MaTAR25 in vivo*, we evaluated the impact of subcutaneous injection of two independent *MaTAR25* ASOs for their *in vivo* ability to knockdown *MaTAR25* and to impact mammary tumor progression in the MMTV-Neu-NDL mouse model. ASO mediated knockdown of *MaTAR25* resulted in a 59% decrease in tumor growth compared to the scASO control group (Fig. 3D and Supplementary Fig. S3B). By comparing the hematoxylin and eosin (H&E) stained tumor sections collected from the *MaTAR25* ASO injected group with the scASO control group, we observed a strong level of necrosis in the *MaTAR25* ASO treated mammary tumor samples (Fig. 3E) but not in other non-tumor tissues (Supplementary Fig. S3C). Importantly, mammary tumors from the scASO control group lacked any significant necrotic phenotype (Fig. 3E). We also collected lung samples from each group to examine for the presence of micro-metastases, and the H&E stained lung sections showed that KD of *MaTAR25* resulted in a 40% incidence rate of micro-metastatic nodules in lungs from ASO1 or ASO2 treated animals as compared to a 76.9% incidence rate for the scASO control group (Supplementary Fig. S3D). Together, our *in vitro* and *in vivo* data indicate that

MaTAR25 plays a critical role in promoting mammary tumor progression and metastasis.

MaTAR25 is a positive upstream regulator of _Tns1_, a mediator of cell-matrix adhesion and migration

Next, we were interested in revealing aspects of the molecular mechanism of action of *MaTAR25* in regulating mammary tumor progression. Since we previously identified *MaTAR25* to be highly enriched in cell nuclei by smRNA-FISH we went on to perform cell fractionation to isolate cytoplasmic and nucleoplasmic lysates as well as chromatin pellets of 4T1 cells to determine the subcellular enrichment of *MaTAR25* by qRT-PCR analysis. Notably, compared to the enrichment of β-*actin* and *Malat1*, we found a significant enrichment of *MaTAR25* in the nucleoplasmic and chromatin fractions (Fig. 4A), indicating that the molecular mechanism of action of *MaTAR25* may be related to transcriptional regulation. To test this hypothesis, we performed Chromatin Isolation by RNA Purification (ChIRP) (34) to pull down RNA/DNA complexes by using specific biotin-labeled antisense oligonucleotides targeting *MaTAR25* (Fig. 4B) as well as biotin-labeled antisense oligonucleotides targeting housekeeping gene PPIB transcripts as the corresponding control. ChIRP-seq identified *MaTAR25* genomic targeting sites, and revealed that these targets are highly enriched in simple repeats regions and LTRs (log ratio enrichments to input are 2-3 fold) (Supplementary Fig. S4A). According to motif analysis these regions are potential binding sites of the transcription factors ZNF354C, TEAD, GATA1, and REL (data not shown). Combining the *MaTAR25* KO RNA-seq data and ChIRP-seq results, we found a total of 446

overlapping genes (Fig. 4C and Supplementary Table S1), which could be downstream targets regulated by *MaTAR25*. Among these overlapping genes, the top gene ranked by ChIRP-seq data, just under *MaTAR25* itself, is *Tensin1* (*Tns1*) (Fig. 4C and Supplementary Fig. S4B). The *Tns1* gene encodes for a protein that localizes to focal adhesions and positively regulates cell migration and invasion (35, 36). By qRT-PCR and immunoblot analysis, we found that the RNA and protein levels of *Tns1* are significantly lower in *MaTAR25* KO cells than in 4T1 control cells (Fig. 4D). Interestingly, ectopic expression of *MaTAR25* in *MaTAR25* KO cells results in a correspondingly increased level of *Tns1* (Fig. 4E). Hence, we conclude that *Tns1* is a direct downstream target of *MaTAR25* and further confirming that it imparts its function *in trans*. In order to confirm the ChIRP-seq result and to further investigate how *MaTAR25* might regulate the level of *Tns1*, we next performed double-label DNA-FISH to detect the *MaTAR25* (Chr2) and *Tns1* (Chr1) gene loci in cells, and we found no physical interaction between these genomic loci (Supplementary Fig. S4D, upper panels). However, combined *MaTAR25* smRNA-FISH and *Tns1* DNA-FISH in the same cells showed that *MaTAR25* RNA was overlapping with at least one *Tns1* allele in 50% of the cells (Supplementary Fig. S4D, lower panels). This suggests that *MaTAR25* RNA transcripts can bind to the gene body of *Tns1* to regulate its expression. We therefore performed CRISPR/Cas9 knockout using gRNAs targeting *Tns1* in 4T1 cells and selected *Tns1* KO clones for *in vitro* functional assays. We found that the *Tns1* KO cells phenocopied the *MaTAR25* KO cells and exhibited a significant 40% decrease in cell viability (Fig. 4F) and a 30% decrease in cell migration vs control cells (Supplementary Fig. S4E). In addition, ectopic expression of *Tns1* in 4T1 *MaTAR25* KO cells can rescue the cell viability phenotype

(Fig. 4G). Interestingly, high expression of *TNS1* is strongly correlated with poor survival of grade 3 breast cancer patients (37) (Supplementary Fig. S4C). Together, these data indicate that *Tns1* is a critical downstream target of *MaTAR25*.

Since *Tns1* is a key component of focal adhesion complexes and is responsible for cell-cell and cell-matrix interactions as well as cell migration by interacting with actin filaments (38), we examined the organization of actin filaments, as well as the additional focal adhesion complex components paxillin and vinculin (39), in 4T1 control and *MaTAR25* KO cells by immunofluorescence (IF) confocal microscopy. Indeed, the F-actin microfilaments are disrupted (Fig. S4G1 and Supplementary Fig. S4G2) and the distribution of paxillin and vinculin proteins are altered dramatically (Supplementary Fig. S4H) in 4T1 *MaTAR25* KO cells as compared to 4T1 control cells. Interestingly, both ectopic expression of *MaTAR25* or *Tns1* in 4T1 *MaTAR25* KO cells can rescue the actin filament phenotype (Supplementary Fig. S4G3 and S4G4), supporting our finding that *Tns1* is a critical downstream target of *MaTAR25* regulating mammary tumor progression. To further evaluate the phenotype of *MaTAR25* KO cells we used transmission electron microscopy (TEM) (Supplementary Fig. S4F). TEM clearly revealed a dramatic 81% reduction of microvilli over the cell surface of *MaTAR25* KO cells compared to 4T1 control cells, indicating loss of *MaTAR25* expression impacts the actin bundling process (Supplementary Fig. S4H) as well as microvilli formation/maintenance in 4T1 cells.

***MaTAR25* interacts with PURB to carry out its function**

316 It has been suggested that lncRNAs can interact with transcriptional regulators/co-
317 factors to form ribonucleoprotein (RNP) complexes to regulate the expression of
318 downstream genes in the cell nucleus (40). To identify *MaTAR25* interacting proteins we
319 used two different paired sets of biotin-labeled antisense oligonucleotides targeting
320 *MaTAR25* for native RNA antisense oligonucleotide pull-down (RAP) in 4T1 cells
321 followed by qRT-PCR which revealed a 50-60% pull-down efficiency (Supplementary
322 Fig. S5A). Samples were eluted from beads for mass spectrometry isobaric tags for
323 relative and absolute quantitative (MS-iTRAQ) analysis to identify proteins that bind to
324 *MaTAR25*, and PPIB as the corresponding control. We ranked the candidate interactors
325 based on detectable peptides above background in both pair sets of oligonucleotide
326 pull-downs, and selected candidates with at least 2-fold enrichment compared to
327 corresponding PPIB oligo pull-down (Fig. 5A). Among the protein candidates, two
328 transcription co-regulators always appeared on the top list between multiple runs. These
329 are purine rich element binding protein A **(**PURA) and purine rich element binding
330 protein B (PURB), which can form homodimers or heterodimers in the nucleus (41).
331 Additionally, one other protein, Y-box protein 1 (YBX1) also on the candidate list, but
332 which did not pass the enrichment criteria, was shown in a previous study to interact
333 with PURA to form a PURB/PURA/YBX1 heterotrimer (42). To verify our MS result, we
334 first performed immunoblot analysis with PURA and PURB antibodies and we could
335 detect extremely higher signals of PURA and PURB in samples of *MaTAR25*
336 oligonucleotide pull-down than the PPIB oligonucleotide pull-down (Fig. 5B). When RAP
337 was carried out with the same sets of oligonucleotide pairs using NDL cells, the MS
338 result showed a greater enrichment of PURB than PURA (data not shown). Based on

these data, we hypothesized that PURB is the lead protein directly binding to *MaTAR25*. Immunoblot analysis using pull-down samples from 4T1 and 4T1 *MaTAR25* KO cell lysates (Fig. 5B) plus pull-down samples from NDL primary cell lysates (Supplementary Fig. S5B) confirmed the specific interaction between *MaTAR25* and PURB. RNA immunoprecipitation (RIP) using PURB antibodies compared to IgG control also revealed the specificity of the *MaTAR25*-PURB interaction (Fig. 5C). To further confirm the role of PURB in regulating *Tns1* we manipulated the level of PURB in 4T1 cells either through ectopic overexpression (Fig. 5D) or siRNA mediated knockdown and demonstrated upregulation or down-regulation of *Tns1*, respectively (Supplementary Fig. S5C). Next, to determine if the interaction between *MaTAR25* and PURB is related to the expression of *Tns1* we examined the level of *Tns1* expression in each group. The results confirmed that the expression level of *Tns1* is related to the changes in PURB expression level in 4T1 cells, indicating that the *MaTAR25*/PURB RNP complex is essential for the regulation of *Tns1*.

Tns1 isoform 3 was identified as the major isoform expressed in MMTV-Neu-NDL and 4T1 cells (data not shown). Based on our ChIRP-seq result, we were able to go on to identify the promoter region of *Tns1* isoform 3, which contains a very high purine:pyrimidine ratio (3:1) including many potential PURB binding sequence motifs (GGTGG) (43), as the main targeting region in the *Tns1* gene (Supplementary Fig. S5D). Moreover, Hypergeometric Optimization of Motif EnRichment (HOMER) motif analysis based on ChIRP-seq data indicated the top enriched motif sequence of *MaTAR25* interacting genes is GGTGGTGGAGAT further supporting the *MaTAR25* PURB binding motif sequence (Supplementary Fig. S5E). Therefore, we performed Chromatin

362 immunoprecipitation (ChIP) using a PURB antibody and multiple qPCR primer pairs and
363 showed that PURB has a high occupancy capacity over this region of *Tns1*. Importantly,
364 the occupancy was impaired in *MaTAR25* KO cells and was able to be restored upon
365 ectopic expression of *MaTAR25* in 4T1 *MaTAR25* KO cells (Fig. 5E). Together, these
366 results provide compelling evidence indicating the interaction of PURB protein with
367 *MaTAR25* is required for PURB binding to regulatory motifs in the *Tns1* gene in 4T1
368 cells.
369
370 **Human lncRNA *LINC01271* is the human ortholog of *MaTAR25***
371 In order to translate our exciting findings in regard to mouse *MaTAR25* to the
372 human system for potential future clinical applications, we went on to characterize the
373 human ortholog of *MaTAR25* and confirm its function in human breast cancer cells.
374 Based on syntenic conservation between the human and mouse genomes, we
375 previously found three lincRNAs as potential human counterparts of *MaTAR25*:
376 *LINC01270*, *LINC01271*, and *LINC01272* (29) (Fig. 6A). Among these three lncRNAs,
377 only *LINC01271* is transcribed in the same direction as *MaTAR25*. Analysis of The
378 Cancer Genome Atlas (TCGA) data (29) suggests two of these potential orthologs,
379 *LINC01270* and *LINC01271*, are expressed at increased levels in multiple sub-types of
380 breast cancer (Fig. 6A). Therefore, we focused on these two lincRNAs and performed
381 independent ectopic expression of *LINC01271* and *LINC01270* in 4T1 *MaTAR25* KO
382 cells to determine if one of these human lincRNAs could rescue the mouse *MaTAR25*
383 KO phenotype. Cell viability assays indicated that ectopic expression of *LINC01271,* but
384 not *LINC01270,* can rescue the proliferation phenotype of *MaTAR25* KO cells (Fig. 6B).

Invasion assays also showed that ectopic expression of *LINC01271* in 4T1 *MaTAR25* KO cells can rescue the cell invasion phenotype (Fig. 6C). In addition, the expression of *Tns1* can also be restored to a similar level as in 4T1 control cells upon overexpression of *LINC01271* in *MaTAR25* KO cells (Fig. 6D). Immunoprecipitation (RIP) using the PURB antibody indicated a specific interaction between PURB and *LINC01271* (Supplementary Fig. S6A) in human triple negative breast cancer MDA-MB-231 LM2 cells (44). Next, we performed smRNA-FISH to examine the localization of *LINC01271* within MDA-MB-231 LM2 cells and we found that similar to *MaTAR25* it is a nuclear enriched RNA (Supplementary Fig.S6C). Together, these results validate *LINC01271* as the human ortholog of *MaTAR25*.

LINC01271 may play a role in human breast cancer progression and have diagnostic and/or therapeutic potential

We performed ASO mediated KD of *LINC01271* in MDA-MB-231 LM2 cells and selected the three most effective *LINC01271* ASOs to assess a KD phenotype. After 96 hours of ASO treatment to mediate KD of *LINC01271* in MDA-MB-231 LM2 cells, all three independent ASOs decreased cell viability by approximately 32% (Fig. 6E). The KD result supports a role for *LINC01271* in human breast cancer progression.

According to the lncRNA database TANRIC (45), higher expression of *LINC01271* is correlated with poor breast cancer patient survival (Supplementary Fig. S6D). qRT-PCR analysis of the expression level of *LINC01271* in breast tumor organoids vs organoids derived from normal adjacent breast tissue showed a higher expression level of *LINC01271* in tumor-derived organoids (Supplementary Fig. S6E).

Next, we performed smRNA-FISH to localize *LINC01271* in patient breast tumor sections. We found that *LINC01271* expression level was increased with increased breast tumor stage (Fig. 7A and Supplementary Fig. S7A), and we identified the presence of clonal and regional differential expression patterns in most of the breast tumor patient samples (Supplementary Fig. S7B). Most interestingly, lung metastases exhibited higher expression of *LINC01271* than primary tumors from the same patients (Fig. 7B and Supplementary Fig. S7C). Thus, together these findings from patient-derived samples support our hypothesis that *LINC01271* is a potential therapeutic target to impact breast cancer progression and metastasis.

DISCUSSION

We identified *MaTAR25* as a lncRNA that is upregulated in Her2+, luminal, and TNBC as compared to normal mammary epithelial cells. Genetic KO or ASO KD of *MaTAR25* results in a reduction in cell proliferation, migration, and invasion. Introduction of *MaTAR25* 4T1 KO cells into mice by mammary fat pad injection or tail vein injection results in smaller tumor growth and a significant reduction in lung metastases. The human ortholog of *MaTAR25* was identified as *LINC01271* and it is expressed in primary breast tumors and at even higher levels in lung metastases. Increased expression of the human ortholog of *MaTAR25* is associated with poor patient prognosis (44).

MaTAR25 regulates expression of the *Tns1* gene

In order to elucidate the molecular mechanism by which *MaTAR25* imparts its function, we used ChIRP-seq and RNA-seq to identify and validate the *Tns1* gene as a direct downstream target of *MaTAR25*. *MaTAR25* positively regulates the expression of *Tns1* in mammary tumor cells through binding to its DNA sequence in *trans*. Tns1 has been shown to localize to focal adhesions and to assist in mediating signaling between the extracellular matrix and the actin cytoskeleton to impact cell movement and proliferation (35, 36). Several reports have shown that loss of *Tns1* can cause a decrease in cell motility in many cell types (46-48), and *Tns1* has also been shown to be involved in epithelial-mesenchymal transition (EMT) of cancer cells (49). The general relationship between the expression of *Tns1* and different stages/subtypes of breast cancer has been unclear. However, Kaplan-Meier survival analysis indicates that the expression of *Tns1* is increased in grade 3 breast tumors (37) supporting our findings of a positive role of *Tns1* in breast tumor progression mediated by *MaTAR25*. Although *Tns1* is the top target of *MaTAR25* with the highest statistical significance we cannot rule out the possibility that *MaTAR25* also regulates additional genes given the other candidates identified in our ChIRP-seq analysis. These candidates will be the focus of future studies.

***MaTAR25* partners with purine rich element binding protein B**

By performing RNA antisense oligonucleotide pull-downs (RAP) in 4T1 cells combined with iTRAQ mass spectrometry, to determine the absolute quantitation of *MaTAR25* associated proteins, we identified a specific interaction between *MaTAR25* and purine rich element binding protein B (PURB) which appears to be crucial for the

453 downstream regulation of *MaTAR25*. PURB has been reported to be a transcriptional
454 co-activator and binds to the single strand of the repeated purine-rich element PUR
455 which is present in promoter regions and as such has been implicated in transcriptional
456 control. PURB plays different roles in many physiological and pathological processes
457 (50-52). For example, PURB has been shown to be over-expressed in several different
458 cancer types (53). In addition, a previous study showed PURB to act as a transcriptional
459 co-factor that can be recruited by *linc-HOXA1* to mediate its transcriptional regulation in
460 embryonic stem cells (43) supporting a critical role of PURB with different lincRNAs in
461 different biological contexts. We identified the *MaTAR25* target region of the *Tns1* gene
462 to have high ratio of purine and pyrimidine bases (3:1), and our ChIP-qPCR result
463 indicated a high occupancy capacity of PURB over this targeting region, further
464 supporting a functional role of PURB in partnering with *MaTAR25* to regulate the *Tns1*
465 gene. The results of ectopic over-expression or siRNA mediated knockdown of PURB,
466 resulting in altered expression (upregulation or down-regulation, respectively) of *Tns1* in
467 4T1 mammary tumor cells indicates a transcriptional regulatory role of PURB in this
468 context. As the interaction of PURB with *MaTAR25* is essential for PURB binding to
469 *Tns1* DNA this suggests that *MaTAR25* acts as a chaperone and/or scaffold for the
470 *MaTAR25*/PURB/*Tns1* DNA complex, which is critical for transcriptional regulation of
471 *Tns1* thereby impacting cancer progression.

472 PURB can form a homodimer, a heterodimer with PURA, or a heterotrimer with
473 PURA and Y-box protein 1 (YBX1). Interestingly, these two additional proteins were
474 also identified in our MS-iTRAQ analysis of *MaTAR25* interactors, and have been
475 studied in many cancer types (54-56). Future investigation of these potential

interactions may provide further insights into the molecular mechanism of the *MaTAR25* PURB complex in cancer cells.

LINC01271* is the human ortholog of *MaTAR25

Based upon synteny and further validation we identified *LINC01271* as the human ortholog of *MaTAR25.* Interestingly, *LINC01271* has been identified as one of 65 new genetic loci that are related to overall breast cancer risk (57). *LINC01271* expression is increased in breast cancer and correlates with poor clinical outcome based on the analysis of patient clinical data (44). In addition, our results from examining sections of breast tumors and corresponding lung metastasis from the same patients showed a positive correlation between the high expression of *LINC01271* and breast cancer stage. Our finding of an even higher level of expression in lung metastases from the same patients was especially interesting. Metastasis is the major cause of cancer related deaths, particularly in breast cancer patients (58, 59), and developing an efficient treatment strategy to target and reduce breast cancer metastasis still remains the key challenge for this disease. Our ability to use ASOs to knockdown *MaTAR25* in the Her2/neu mouse model resulting in necrotic tumors and a significant reduction in metastasis has therapeutic implications. ASO targeting as a therapeutic approach has been applied to many diseases (60-62) including cancer (26, 63, 64) in recent years. For example, a cEt ASO targeting the transcription factor STAT3 has shown robust single agent activity in highly treatment-refractory lymphoma and non-small cell lung cancer studies (64). The STAT3 ASO (AZD9150) has advanced into multiple Phase I and II clinical trials (NCT01563302, NCT02983578, NCT02549651). In

addition, an antisense drug targeting all forms of the androgen receptor for the treatment of advanced metastatic prostate cancer has entered a clinical trial (NCT02144051). In this study, we developed three ASOs targeting *LINC01271* for functional assays *in vitro*. Ultimately, *LINC01271* represents an ideal candidate to be exploited as a potential prognostic/therapeutic target and *LINC01271* specific ASOs as therapeutics to impact breast cancer progression and metastasis.

METHODS

EXPERIMENTAL MODELS

Cell Culture

Murine 4T1 cells, murine NF639 (MMTV-cNEU) cells, human MDA-MB-231 cells, and human MDA-MB-231 LM2 cells were cultured in DMEM with 10% FBS, and 1% penicillin/strptomycin.

Organoid Culture

Surgically removed tumor samples from breast cancer patients along with adjacent normal tissue were collected from Northwell Health in accordance with Institutional Review Board protocol IRB-03-012 (TAP16-08). Informed consent ensured that the de-identified materials collected, the models created, and data generated from them can be shared without exceptions with researchers in the scientific community. Tumor and normal organoids were developed using a previously published protocol (Sachs et al., 2018). The tissues were manually cut into smaller pieces and treated with Collagenase IV at 37°C. The samples were manually broken down by pipetting into smaller

fragments and seeded in a dome of matrigel. Organoids were grown in culture media which contained 10% R-Spondin 1 conditioned media, 5 nM Neuregulin 1, 5 ng/ml FGF7, 20 ng/ml FGF10, 5 ng/ml EGF, 100 ng/ml Noggin, 500 nM A83-01, 5 μM Y-27632, 1.2 μM SB 202190, 1x B27 supplement, 1.25 mM N-Acetyl-cysteine, 5 mM Nicotinamide, 1x Glutamax, 10 mM Hepes, 100U/ml Penicillin/streptomycin, 50 μg/ml Primocin in 1x Advanced DMEM-F12 (Sachs et al., 2018). Cultures were passaged every 2-4 weeks using TrypLETM to break down the organoids into smaller clusters of cells and replating them in Matrigel domes.

Mice

All animal procedures and studies were approved by the Cold Spring Harbor Laboratory Animal Use Committee in accordance to IACUC procedures. Briefly, male MMTV-Neu-NDL mice (FVB/N background) were kindly provided by Dr. William Muller (McGill University, Canada). Male MMTV-Neu-NDL mice were crossed with wild type FVB/N female mice purchased from the Jackson Laboratory for breeding. PCR genotyping was applied to select female mice with heterogeneous genotypes for MMTV-Neu-NDL for later *in vivo* ASO injection experiments. 4-6 week old BALB/c female mice were purchased from the Jackson Laboratory for *in vivo* 4T1 cell mammary fat pad (MFP) injection and tail-vein injection experiments.

METHODS

Details of materials and reagents are listed in Supplementary Table S2

Details of oligonucleotide sequences are listed in Supplementary Table S3

545

546 **5'/3' Rapid Amplification of cDNA Ends (RACE)**

547 5' and 3' RACE of *MaTAR25* transcripts was performed on TRIzol-extracted RNA using
548 the Ambion FirstChoice RLM-RACE kit according to the manufacturer's instructions.
549 Briefly, fragments were amplified by nested PCR using AmpliTaq Polymerase. PCR
550 products were separated on 2% agarose, bands excised, gel purified, sub-cloned into
551 pGEM-T Easy (Promega) and 4 or more clones per fragment were sequenced using
552 standard Sanger sequencing. Primer sequences are provided in Table S1.

553

554 **Northern Blot Analysis**

555 *MaTAR25*-specific radiolabeled DNA probes were generated using dCTP P32 in a
556 random primed labeling reaction. Total RNA was extracted by the TRIzol method.
557 Analysis of RNA expression was performed by following NorthernMax® Kit manual.
558 Briefly, 20 µg and 30 µg total RNA samples were electrophoresed on a 1% agarose gel
559 and was transferred to a positively charged nylon membrane (NC). The RNA was then
560 fixed to the NC membrane using UV crosslinking. The cross-linked membrane was
561 prehybridized with ultrahyb-oligo hybridization buffer and hybridized with the *MaTAR25*-
562 specific radiolabeled DNA probe. After washing with SSC wash buffers several times,
563 the wrapped membrane was exposed to a PhosphorImager screen in a cassette or X
564 ray film for detecting signals.

565

566 **Cell Lysate Preparation for Immunoblot Analysis**

Cells were trypsinized, and harvested cell pellets were lysed in RIPA buffer (25 mM Tris-HCl pH 7.6, 150 mM NaCl, 1% NP-40 substitute, 1% sodium deoxycholate and 0.1% SDS) supplemented with 1X Roche protease inhibitor cocktail. The cell lysate was incubated on ice for 15 minutes, then sonicated for 5 minutes before centrifugation at 13000xg. The supernatant was collected and quantified using the BCA protein assay.

In vitro Transcription/Translation

T7 promoter containing DNA or plasmids were used in a TNT® Quick Coupled Transcription/Translation System following the manufacture's protocol. Briefly, 1 µg Plasmid DNA Template was mixed with 40 µl TNT® T7 Quick Master Mix, 1 µl Methionine (1 mM), 1 µl Transcend™ Biotin-Lysyl-tRNA, and nuclease free water for the final volume of 50 µl per reaction. The reaction tube was incubated at 30°C for 90 minutes, and 1 µl reaction product was added into diluted 2x Laemmli sample buffer for immunoblot analysis. Samples were loaded on 4–20% Mini-PROTEAN® TGX™ Precast Protein Gels, and the signals were detected by Streptavidin-HRP.

RNA Isolation and Quantitative Real-Time PCR (qRT-PCR) Assays

Total RNA was extracted using TRIzol following the manufacture's protocol. 1 µg total RNA was treated with DNaseI and reverse transcribed into cDNA using TaqMan Reverse Transcription Reagent kit, followed by qPCR with SYBR green PCR master mix on an ABI QuantStudio 6 Flex Real-Time PCR System. qRT-PCR conditions were as follows: 30 minutes at 50°C for reverse transcription, 15 minutes at 95°C for the initial activation step followed by 40 cycles of 15 seconds at 94°C, 30 seconds at 60°C.

590 Mouse peptidylprolyl isomerase B (cyclophilin B; PPIB), and human GAPDH and
591 RPL13A were used as endogenous controls to normalize each sample. A list of primers
592 used is provided in Table S1.
593
594 **Cell Fractionation, Cytoplasmic/Nucleoplasmic/Chromatin-related RNA Isolation**
595 Cell fractionation was done using a standardized protocol previously described (66).
596 Briefly, cultured cells were harvested and lysed in NP-40 substitute lysis buffer (10 mM
597 Tris pH 7.5, 150 mM NaCl, and 0.15% NP-40 substitute). The cell lysate was overlaid
598 on top of sucrose buffer (10 mM Tris pH7.4, 150 mM NaCl, and 24% sucrose) and
599 centrifuged at 3500xg for 10 minutes to separate the cytoplasmic fraction and nuclei
600 pellet. The nuclei pellet was rinsed with PBS-EDTA once, and resuspended in glycerol
601 buffer (20 mM Tris pH7.4, 75 mM NaCl, 0.5 mM EDTA, and 50% Glycerol) mixed with
602 Urea buffer (1 M Urea, 0.3 M NaCl, 7.5 mM $MgCl_2$, 0.2 mM EDTA, and 1% NP-40
603 substitute) on ice for 2 minutes. The lysate was centrifuged at 13000xg for 2 minutes to
604 separate the nucleoplasmic fraction and chromatin pellet. The chromatin later was
605 resuspended in TRIzol reagent and fully solubilized by passing through the 21-gauge
606 needle and syringe. The cytoplasmic fraction and nucleoplasmic fraction were also used
607 for RNA extraction using TRIzol reagent. RNA extracted from different fractions were
608 applied for cDNA synthesis and qRT-PCR. Primer sequences are provided in Table S1.
609
610 **CRISPR/Cas9 Genetic Knockout**
611 To generate a genetic knockout of *MaTAR25*, two sgRNAs targeting the promoter
612 region were combined, creating a deletion including the TSS. Both sgRNAs were

designed using http://crispr.mit.edu/. The sgRNA targeting the gene body of *MaTAR25* was cloned into a pSpCas9(BB)-2A-GFP vector and the sgRNA targeting the upstream promoter region was cloned into a pSpCas9(BB)- 2A-mCherry vector. 4T1 cells were transfected with both plasmids using 4D-Nucleofector X Unit, using program code "CN-114". To select for cells expressing both gRNAs, GFP and mCherry double positive cells were sorted 40 hours post transfection, as single cell deposition into 96-well plates using a FACS Aria (SORP) Cell Sorter (BD). Each single cell clone was propagated and analyzed by genomic PCR and qRT-PCR to select for homozygous knockout clones. Cells transfected with a sgRNA targeting *Renilla* luciferase were used as a negative control. Sequences for sgRNAs and primers are provided in Table S1.

Cell Counting Viability Assay

Cultured cells were harvested and the same number of cells were seeded into each well of a 12 well tissue culture plate at day 0. Trypan Blue-treated cell suspensions were collected and applied to a hemocytometer for manual counting at different time points.

Cell Cycle Analysis

Cell cycle analysis was performed using BD bromodeoxyuridine (BrdU) FITC assay kit following the manufacturer's protocol. Briefly, cultured cells were incubated with BrdU containing medium for 30 minutes, and FITC conjugated anti-BrdU antibody was applied for labeling actively synthesizing DNA. 7-aminoactinomycin D (7-AAD) was used for labeling total DNA. The labeled cell samples were analyzed on the BD LSR II flow cytometer, and BrdU FITC-A vs DNA 7-AAD dot plot with gates was used to

636 encompass the G0/G1, S, and G2/M populations. The collected cytometry data were
637 analyzed with FACSDiva™ and FlowJo software.
638
639 **Migration Assay**
640 Live cell tracking was performed to examine cell migration. Cultured cells were
641 harvested and the same number of cells were seeded into 6 wells of a tissue culture
642 plate at day 0. Images were collected every 5 minutes (5 viewpoints were selected from
643 each well) using Zeiss AxioObserver microscope for 8 hours, and the images of
644 individual sample were converted to videos using ImageJ. Videos were analyzed by
645 CellTracker image processing software. The mean relative migration distance (μm) of
646 three independent replicates of 4T1 control groups and *MaTAR25* KO groups were
647 calculated.
648
649 **Scratch Wound Healing Assay**
650 Cultured cells were harvested and seeded into each well of a 24 well tissue culture
651 plate. Cells were incubated until they reached at least 90% confluence. The wound line
652 was created by “scratch” with a p200 micropipette tip, then cells were washed with PBS
653 twice to remove the debris and then each well was imaged. After 12 hours incubation,
654 each well was images again and the migration areas in each well was measured by
655 ImageJ for comparison.
656
657 **Invasion Assay**

Invasion assays with 4T1 *MaTAR25* knockout cells and 4T1 control cells were carried out using the Cultrex® 24 well BME Cell Invasion Assay (Trevigen) following the manufacturer's protocol. Briefly, cells were harvested and seeded at a density of 1 x 10^5 cells/well into the invasion chamber. As a negative control, serum-free medium was used that did not stimulate cell invasion through the BME. The plate was incubated at 37°C for 24 hours and the assay was performed. The tumor cells that invaded through the BME layer and attached to the bottom of the invasion chamber were collected using cell dissociation solution and stained with Calcein AM solution. The fluorescence was measured with a SpectraMax i3 Multi-Mode Detection Platform (Molecular Devices) using the 480/520 nm filter set.

Cloning

Specific gene overexpression plasmids were constructed using NEBuilder HiFi DNA Assembly following the manufacture's protocol. PCR products were cloned into pCMV6-entry plasmids digested with Sgf I and Fse I. Assembled plasmids were introduced into NEB stable competent E. *coli* using heat shock transformation and kanamycin selection. 4 or more colonies per plate were picked and sequenced using standard Sanger sequencing.

Antisense Oligonucleotide (ASO) and siRNA-Mediated Knockdown (KD)

Specific 16mer antisense oligonucleotides (ASOs) comprised of phosphorothioate-modified short S-cEt (S-2´-O-Et-2´,4´-bridged nucleic acid) gapmer chemistry targeting *MaTAR25* and *LINC01271* were designed and provided by Ionis

Pharmaceuticals, Inc. Briefly, cultured cells were harvested and seeded into culture dishes. Transfection-free uptake of ASOs was accomplished by adding 4 μM of either *MaTAR25/LINC01271*-specific ASOs or scrambled ASO (scASO) to the culture medium immediately after seeding the cells. Cells were incubated for indicated time points and RNA was isolated using TRIzol reagent for qRT-PCR to check the knockdown efficiency. For siRNA mediated knockdown, siRNAs (27mers) targeting mouse PURB were purchased from ORIGENE, and siRNA transfection was done using Lipofectamine 2000 following the manufacture's protocol. RNA was extracted at different time points for qRT-PCR to check the knockdown efficiency. ASO sequences and primer sequences are provided in Table S1.

Chromatin Isolation by RNA purification (ChIRP)-Seq

For Chromatin Isolation by RNA purification (ChIRP), we followed a previously described protocol (34). Briefly, 20 million cells were harvested and fixed in 1% glutaraldehyde solution for each reaction. ChIRP was performed using biotinylated oligo probes designed against mouse *MaTAR25* using the ChIRP probe designer (Biosearch Technologies). Independent even and odd probe pools were used to ensure lncRNA-specific retrieval (refer to Table S1 for odd and even sequences targeting *MaTAR25*, and probes targeting mouse PPIB transcripts which were used as negative controls). ChIRP-Seq libraries were constructed using the Illumina TruSeq ChIP Library Preparation Kit. Sequencing libraries were barcoded using TruSeq adapters and sequenced on Illumina NextSeq instruments.

704 **ChIRP-Seq Data analysis**

705 Data quality was assessed using FastQC
706 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and paired-end reads were
707 mapped to GRCm38 using Bowtie2 (67) with parameters --end-to-end --sensitive --fr,
708 resulting in a 90% or higher overall alignment rate. ChIRP seq analysis was performed
709 using HOMER (65). Differential ChIRP peaks were called using the
710 getDifferentialPeaksReplicates.pl script, with negative control (PPIB) pull-down samples
711 as background and parameters -style histone -f 50. Peaks identified with at least a 50-
712 fold enrichment were processed further using the annotatePeaks.pl script and the
713 GENCODE vM16 annotation. Both known and *de novo* motif analysis was carried out
714 with the findMotifsGenome.pl script on the repeat-masked GRCm38 genome, +/- 500 bp
715 around the identified ChIRP peaks.

716

717 **RNA-seq Library Construction**

718 RNA was extracted using TRIzol following the manufacture's protocol. RNA quality was
719 assayed by running an RNA 6000 Nano chip on a 2100 Bioanalyzer. 1 µg total RNA
720 was used for constructing each RNA-seq library using the Illumina TruSeq sample prep
721 kit v2 following the manufacture's protocol. Briefly, RNA was polyA selected and
722 enzymatically fragmented. cDNA was synthesized using Super Script II master mix,
723 followed by end repair, A-tailing and PCR amplification. Each library was high-
724 throughput single-end sequenced on Illumina NextSeq instruments.

725 .

726

727 **RNA-Seq Data analysis**

728 Data was analyzed as previously described (29). Briefly, the quality of FASTQ files was
729 assessed using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/).
730 Reads were mapped to GRCm38 using STAR (68), and the reads per gene record were
731 counted using HTSeq-count (69) and the GENCODE vM5 annotation. Differential gene
732 expression was performed with DESeq2 (70), and an adjusted p-value of < 0.05 was set
733 as threshold for statistical significance. KEGG pathway and GO term enrichment and
734 was carried out using the R/Bioconductor packages GAGE 71) and Pathview (72).

735

736 **RNA Antisense Pulldown and Mass Spectrometry**

737 Cells were lysed in a 10 cm culture dish in 1 ml IP lysis buffer (IPLB, 25 mM Tris-HCl pH
738 7.4, 150 mM NaCl, 1% NP-40, 1 mM EDTA, 5% glycerol, supplemented with 100 U/ml
739 SUPERase-IN and 1X Roche protease inhibitor cocktail) for 10 minutes, and lysate was
740 centrifuged at 13,000xg for 10 minutes. Cell lysate was adjusted to 0.3 mg/ml (Pierce
741 BCA Protein Assay). A total of 100 pmol of biotinylated oligo was added to 500 µl of
742 lysate and incubated at room temperature for 1 hour with rotation. 100 µl streptavidin
743 Dynabeads were washed in IPLB, added to the lysate, and incubated for 30 minutes at
744 room temperature with rotation. Beads were washed three times with 1 ml lysis buffer.
745 For determining temperature for optimal elution, beads were then resuspended in 240 µl
746 of 100 mM TEAB and aliquoted into eight PCR tubes. Temperature was set on a veriflex
747 PCR block and incubated for 10 minutes. Beads were captured and TRIzol was added
748 to the eluate and beads. Once optimal temperature is established, the beads were
749 resuspended in 90 µl of 100 mM TEAB, and incubated at 50° C for 10 minutes. TRIzol

was added to 30 µl of the eluate, another 30 µl was kept for immunoblots, and the last 30 µl aliquot was sent directly to the Cold Spring Harbor Laboratory Mass Spectrometry Shared Resource for analysis.

RNA Immunoprecipitation (RIP)

RIP was performed following RIP the Abcam protocol with minor modifications. Briefly, cultured cells were harvested and 40 million cells were washed once with cold PBS, then the cells were resuspended in 8 ml PBS, 8 ml nuclear isolation buffer (1.28 M sucrose, 40 mM Tris-HCl pH 7.5, 20 mM $MgCl_2$, and 4% Triton X-100 supplemented with 100 U/ml SUPERase-IN and 1X Roche protease inhibitor cocktail), and 24 ml nuclease free water on ice for 20 min with frequent mixing. The cleared lysates were pelleted by centrifugation at 2,500xg for 15 min. Pellets resuspended in 4 ml RIP buffer (150 mM KCl, 25 mM Tris pH 7.4, 5 mM EDTA, 0.5 mM DTT, and 0.5% NP40 substitute supplemented with 100 U/ml SUPERase-IN and 1X Roche protease inhibitor cocktail) and sonicated for 5 minutes using BioRuptor Pico water bath sonicator (30 s ON/OFF) at 4°C. The lysates were cleaned by centrifugation at 13,000 rpm for 10 minutes. The supertanant was collected and separated, then incubated with 4 µg PURB antibody or rabbit isotype IgG control for 2 hours to overnight at 4°C with gentle rotation. 80 µl of protein A beads for rabbit antibody then added into the reactions and incubate for 1 hour at 4°C with gentle rotation. After washing three times with RIP buffer and once with

PBS, beads were collected for immunoblot analysis and RNA extraction for qRT-PCR. Primers for RIP qRT-PCR can be found in Table S1.

Chromatin Immunoprecipitation (ChIP) coupled with quantitative PCR (ChIP-qPCR)

For Chromatin Immunoprecipitation (ChIP) we followed protocols previously described (73). Briefly, 30 million 4T1 cells were harvested and crosslinked in 1% formaldehyde at room temperature for 20 minutes, then the reaction was quenched using 0.125 M glycine. Cells were incubated with cell lysis buffer (10 mM Tris pH8.0, 10 mM NaCl, 0.2% NP-40 substitute) and then resuspended and sonicated in 1.5 ml of nuclei lysis buffer (50 mM Tris pH8.0, 10 mM EDTA, 1% SDS) for 15 min using BioRuptor Pico water bath sonicator (30 s ON/OFF) at 4°C. For one IP, 1.5 ml of sonicated chromatin from 30 million cells were diluted with 21 ml IP-Dilution buffer (20 mM Tris pH 8.0, 2 mM EDTA, 150 mM NaCl, 1% Triton X-100, 0.01% SDS,) and incubated with 4 µg of PURB antibody or rabbit isotype IgG control, and 80 µl of protein A beads for rabbit antibody at 4°C overnight. After washing once with IP-wash 1 buffer (20 mM Tris pH8.0, 2 mM EDTA, 50 mM NaCl, 1% Triton X-100, 0.1% SDS), twice with High-salt buffer (20 mM Tris pH 8.0, 2 mM EDTA, 500 mM NaCl, 1% Triton X-100, 0.01% SDS), once with IP-wash 2 buffer (10 mM Tris pH 8.0, 1 mM EDTA 0.25 M LiCl, 1% NP-40 substitute, 1% sodium deoxycholate), twice with TE buffer (10 mM Tris-Cl, 1 mM EDTA, pH 8.0), beads bound chromatin were eluted in 800 µl nuclei lysis buffer by heating at 65 °C for 15 minutes. 48 µl of 5 M NaCl was added to the 800 µl eluted chromatin, followed by incubation at 65°C overnight for reverse cross-linking. After reverse cross-linking, DNA

795 was treated with RNaseA and proteinase K, followed by purification using QIAGEN PCR
796 purification kit. qPCR was performed on ABI QuantStudio 6 Flex Real-Time PCR
797 System. ChIP-qPCR primers can be found in Table S1.
798
799 **Single Molecule RNA Fluorescence In Situ Hybridization (FISH)**
800 For single-molecule RNA FISH, custom Type-6 primary probes targeting *MaTAR25*,
801 *LINC01271* and other lncRNAs were designed and synthesized by Affymetrix. For RNA-
802 FISH on cultured cell samples, Affymetrix View ISH Cell Assay Kit reagents were used.
803 Cultured cells were harvested and seeded onto acid-cleaned #1.5 glass coverslips for
804 24 hours incubation to 70% confluence, cell samples then were fixed in freshly-prepared
805 4% paraformaldehyde (PFA). Cells were then permeabilized and protease digested
806 before hybridization. For RNA-FISH on formalin-fixed paraffin-embedded (FFPE) tissue
807 sections of breast tumors and metastases, Affymetrix ViewRNA ISH Tissue 1-Plex
808 Assay kit reagents were applied. Sections on slides were deparaffinized, protease
809 digested, and fixed with 10% NBF before hybridization. QuantiGene ViewRNA probe
810 hybridizations were performed at 40°C for 3 hours. The hybridization and signal
811 amplification steps were performed according to the manufacturer's instructions, and
812 nuclei were counter-stained with DAPI. Coverslips and tissue sections were mounted in
813 ProLong Gold Antifade mounting medium before detection. Imaging was performed on
814 Zeiss LSM 710/780 Confocal Microscope systems.
815
816 **DNA FISH**

Different mouse BAC clones (RPCI-23) were used as template included (*MaTAR25*), and (*Tensin1*). 1 µg BAC DNA was used as template for random priming reaction to generate amine-modified DNA, and amine-modified DNA was labelled with a reactive fluorescent dye as fluorescent probes according to the protocol provided with ARES™ Alexa Fluor™ DNA Labeling Kit. For DNA FISH, we followed protocols previously described (Hogan et al., 2015). Briefly, cultured cells were seeded onto 22mm^2 glass coverslips (Corning), and coverslips were fixed with freshly prepared 4% PFA for 20 minutes at room temperature, and permeablized in 0.5% Triton X-100/1X PBS for 5 minutes on ice. Probes were prepared for hybridization by mixing 2µl probe with 5µl each of sheared salmon sperm DNA, mouse Cot1 DNA, and yeast tRNA, dehydrating the probe mixture in the speed-vac, and then resuspending the probe in 10µl deionized formamide (Ambion). Just prior to hybridization, the probes were denatured at 95°C for 10 minutes, transferred to ice for 5 minutes, and then mixed with 10µl 2X Hybridization Buffer (4X SSC, 20% dextran sulfate) and pipetted onto slides so that coverslips could be placed cell-side-down on the probe mixture for the hybridization reaction. After several washes, nuclei were counter-stained with DAPI. Coverslips were mounted in ProLong Gold Antifade mounting medium before detection. Coverslips were imaged using Zeiss LSM 710/780 Confocal Microscope systems.

Immunofluorescence (IF)

For IF, we followed protocols previously described with minor modifications depended on applied antibody (74). Briefly, cultured cells were harvested and seeded onto acid-cleaned #1.5 glass coverslips for 24 hours incubation to 70% confluence, cell samples

then were fixed in 4% formaldehyde for 20 minutes. Samples were permeabilized in 0.2% Triton X-100 plus 1% Bovine Serum Albumin (BSA) in PBS for 5 minutes on ice. After incubated in 1% Bovine Serum Albumin (BSA) in PBS for 30 minutes blocking, samples were incubated in the appropriate concentration (1:50-1:200 followed by mamufacturer's recommendations) of primary antibody for 1-2 hours at room temperature, and incubated in diluted secondary antibody solution (Alexa 594 conjugated) with phalloidin (Alexa 488 conjugated) for 1 hour in a humidified chamber at room temperature. After several washes, nuclei were counter-stained with DAPI. Coverslips were mounted in ProLong Gold Antifade mounting medium before detection. Coverslips were imaged on the Zeiss LSM 710/780 Confocal Microscope systems.

Transmission Electron Microscopy (TEM)

Cultured cells were harvested and seeded onto 10 cm culture dishes for 24 hours, and fixed with 2.5% glutaraldehyde (EM grade) in 0.1 M phosphate buffer pH 7.4 at room temperature for 1 hour. The fixed cells were collected by using a cell scraper, washed several times in 0.1M phosphate buffer and post fixed in 1% Osmium tetroxide. Samples were dehydrated in ethanol (30%, 50%, 70%, 80%, 95%, and 100%), embedding in EMbed 812 Resin and polymerized. 70-90 nm sections were cut on a Reichert-Jung ultramicrotome using a diamond knife (DiATOME). Sections were collected on copper grids, stained with UranyLess and lead citrate and imaged with a Hitachi H-7000 TEM.

In Vivo Mouse Model ASO Injection

Three month old MMTV-Neu-NDL mice were divided into three cohorts (7-12 mice each), and each mouse in the cohort received either scASO or *MaTAR25* specific ASO1 or ASO2 via subcutaneous injection 50mg/kg/day twice per week. The injections were carried out for a period of 7 weeks, upon which at least one tumor from most of the control mice reached 2 cm in size. During the course of treatment, tumors were measured twice per week. At the end of the treatment period the animals were euthanized and the primary tumors, lungs, livers, spleens, and intestines were collected. Collected lungs were fixed in 4% paraformaldehyde and incubated in 20-30% sucrose solution overnight, then frozen in OCT solution. The lung OCT blocks were cross-sectioned 2 mm apart, and the lung sections were embedded horizontally to obtain serial sections of the entire lung. Other tissues were cut into two pieces. One parts of each issue was snap frozen using liquid nitrogen for further RNA extraction. The remaining tissues were fixed in 4% paraformaldehyde and paraffin embedded, then formalin-fixed, paraffin embeeded (FFPE) tissue blocks were sectioned. All sections were stained with Hematoxlin and Eosin (H&E) following standard protocol, and slides were scanned and analyzed using an Aperio ImageScope pathology slide viewing system. All samples were processed and stained at the Cold Spring Harbor Laboratory Histology Shared Resource.

In Vivo 4T1 Cells Injection

5-6 week old female BALB/c mice and 4T1 cells (control and *MaTAR25* KO cells) expressing luciferase were used for 4T1 mammary fat pad and tail-vein injection

886 experiments. For mammary fat pad injection, $1x10^5$ 4T1 control or *MaTAR25* KO cells
887 were injected orthotopically into the mammary fat pad of female BALB/c mice. Mice
888 were monitored and primary tumors were measured every week. Mice were sacrificed
889 and tumors were collected at day 28 to compare the tumor growth rate between 4T1
890 control groups and *MaTAR25* KO groups.
891 For tail-vein injection, female BALB/c mice were injected intravenously with $1x10^5$ 4T1
892 control or *MaTAR25* KO cells in the tail vein. Mice were monitored every week and
893 sacrificed at day 21. The mouse lungs were collected and imaged, and lung metastatic
894 nodules were counted to compare the metastatic ability between 4T1 control groups
895 and *MaTAR25* KO groups.
896
897 **QUANTIFICATION AND STATISTICAL ANALYSIS**
898 Statistics tests were performed and analysed using Microsoft Excel and GraphPad
899 Prism 7.0. p value was calculated by paired Student's t-test, two-tailed. Significance
900 was defined as $p < 0.05$.
901
902 **DATA AND SOFTWARE AVAILABILITY**
903 The accession number for the RNA-seq and ChIRP-seq data reported in this study is
904 GEO: GSE142169
905
906
907 **Disclosure of Potential Conflicts of Interest**

D.L.Spector is a consultant to, and receives research support from, Ionis Pharmaceuticals.

Authors' Contributions:

Conception and design: K.Chang and D.L. Spector

Analysis of the RNA-seq, ChIRP-seq, and TCGA human BC patient data: S.D. Diermeier

Cell cycle analysis, and technical support of CRISPR mediated KO and RNA antisense pulldown: A.T. Yu

Tissue RNA-FISH: L.D. Brine

Technical support (*in vivo* mouse model ASO injections and 4T1 mammary fat pad injections): S. Russo

Human organoid culture and organoid RNA extraction: S. Bhatia.

TEM sample preparation and imaging: H. Alsudani

Material support (i.e., human BC patient samples for organoid culture and FFPE slides): K. Kostroff, T. Bhuiya, E. Brogi

Material support (designing and providing ASOs for *in vitro* and *in vivo* experiments): C. F. Bennett, F. Rigo

Writing of the manuscript: K.Chang and D.L. Spector

Study supervision: D.L. Spector

Acknowledgments

We thank members of the Spector lab for critical discussions and advice throughout the course of this study. We also acknowledge Dr. J. Erby Wilkinson (University of Michigan Medical School) for help with pathology and histology analysis. We thank Dr. William Muller (McGill University, Montreal) for providing MMTV-Neu-NDL mice, Dr. Mikala Egeblad (CSHL) for providing MMTV-PyMT mice, Dr. Fred R. Miller (Wayne State University) for providing 4T1 cells, Dr. Nicholas Tonks for providing SKBR3 and BT-474 cells, Dr. Joan Massagué (Sloan Kettering Institute) for providing MDA-MB-231 LM2 cells, Dr. Bruce Stillman (CSHL) for providing the pET3a-Xenopus laevis histone H2B plasmid, and Dr. Scott Lyons (CSHL) for providing Ef1α-Luc2-mStrawberry plasmid. We would also like to thank the CSHL Cancer Center Shared Resources (Microscopy, Flow Cytometry, Animal, Histology, and Next-Gen Sequencing) for services and technical expertise (NCI 2P3OCA45508). This research was supported by NCI 5P01CA013106-Project 3 (D.L.S.), Susan G. Komen postdoctoral fellowship (S.D.D.) and NCI 1K99CA215362 (S.D.D.), CSHL/Northwell Health (D.L.S.), and Manhasset Woman's Coalition Against Breast Cancer (S.B.).

REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin 2019;69:7-34.
2. Li CI, Uribe DJ, Daling JR. Clinical characteristics of different histologic types of breast cancer. Br J Cancer 2005;93:1046-52.
3. Malhotra GK, Zhao X, Band H, Band V. Histological, molecular and functional subtypes of breast cancers. Cancer Biol Ther 2010;10:955-60.

4. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012;486:346-52.
5. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature 2000;406:747-52.
6. Wallden B, Storhoff J, Nielsen T, Dowidar N, Schaper C, Ferree S, et al. Development and verification of the PAM50-based Prosigna breast cancer gene signature assay. BMC Med Genomics 2015;8:54.
7. Eroles P, Bosch A, Perez-Fidalgo JA, Lluch A. Molecular biology in breast cancer: intrinsic subtypes and signaling pathways. Cancer Treat Rev 2012;38(6):698-707.
8. Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. Genome Res 2012;22:1775-89.
9. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, et al. Landscape of transcription in human cells. Nature 2012;489:101-8.
10. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 2009;458:223-7.
11. Wilusz JE, Sunwoo H, Spector DL. Long noncoding RNAs: functional surprises from the RNA world. Genes Dev 2009;23:1494-504.
12. Bergmann JH, Spector DL. Long non-coding RNAs: modulators of nuclear structure and function. Curr Opin Cell Biol 2014;26:10-8.

13. Ulitsky I, Bartel DP. lincRNAs: genomics, evolution, and mechanisms. Cell 2013;154:26-46.
14. Mercer TR, Dinger ME, Sunkin SM, Mehler MF, Mattick JS. Specific expression of long noncoding RNAs in the mouse brain. Proc Natl Acad Sci U S A 2008;105:716-21.
15. Hu W, Yuan B, Flygare J, Lodish HF. Long noncoding RNA-mediated anti-apoptotic activity in murine erythroid terminal differentiation. Genes Dev 2011;25:2573-8.
16. Amodio N, Raimondi L, Juli G, Stamato MA, Caracciolo D, Tagliaferri P, et al. MALAT1: a druggable long non-coding RNA for targeted anti-cancer approaches. J Hematol Oncol 2018;11:63.
17. Arun G, Diermeier SD, Spector DL. Therapeutic Targeting of Long Non-Coding RNAs in Cancer. Trends Mol Med 2018;24:257-77.
18. Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. Nat Rev Genet 2014;15:7-21.
19. Morlando M, Fatica A. Alteration of Epigenetic Regulation by Long Noncoding RNAs in Cancer. Int J Mol Sci 2018;19.
20. Sun M, Nie FQ, Wang ZX, De W. Involvement of lncRNA dysregulation in gastric cancer. Histol Histopathol 2016;31:33-9.
21. Tian T, Wang M, Lin S, Guo Y, Dai Z, Liu K, et al. The Impact of lncRNA Dysregulation on Clinicopathology and Survival of Breast Cancer: A Systematic Review and Meta-analysis. Mol Ther Nucleic Acids 2018;12:359-69.
22. Zhang H, Chen Z, Wang X, Huang Z, He Z, Chen Y. Long non-coding RNA: a new player in cancer. J Hematol Oncol 2013;6:37.

998 23. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, et al. Long non-coding
999 RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature
1000 2010;464:1071-6.
1001 24. Godinho M, Meijer D, Setyono-Han B, Dorssers LC, van Agthoven T.
1002 Characterization of BCAR4, a novel oncogene causing endocrine resistance in
1003 human breast cancer cells. J Cell Physiol 2011;226:1741-9.
1004 25. Xing Z, Park PK, Lin C, Yang L. LncRNA BCAR4 wires up signaling transduction in
1005 breast cancer. RNA Biol 2015;12:681-9.
1006 26. Arun G, Diermeier S, Akerman M, Chang KC, Wilkinson JE, Hearn S, et al.
1007 Differentiation of mammary tumors and reduction in metastasis upon Malat1 lncRNA
1008 loss. Genes Dev 2016;30:34-51.
1009 27. Tseng YY, Moriarity BS, Gong W, Akiyama R, Tiwari A, Kawakami H, et al. PVT1
1010 dependence in cancer with MYC copy-number increase. Nature 2014;512:82-6.
1011 28. Liu B, Sun L, Liu Q, Gong C, Yao Y, Lv X, et al. A cytoplasmic NF-kappaB
1012 interacting long noncoding RNA blocks IkappaB phosphorylation and suppresses
1013 breast cancer metastasis. Cancer Cell 2015;27:370-81.
1014 29. Diermeier SD, Chang KC, Freier SM, Song J, El Demerdash O, Krasnitz A, et al.
1015 Mammary Tumor-Associated RNAs Impact Tumor Cell Proliferation, Invasion, and
1016 Migration. Cell Rep 2016;17:261-74.
1017 30. Petryszak R, Keays M, Tang YA, Fonseca NA, Barrera E, Burdett T, et al.
1018 Expression Atlas update--an integrated database of gene and protein expression in
1019 humans, animals and plants. Nucleic Acids Res 2016;44:D746-52.

1020 31. Murray S, Ittig D, Koller E, Berdeja A, Chappell A, Prakash TP, et al. TricycloDNA-
1021 modified oligo-2'-deoxyribonucleotides reduce scavenger receptor B1 mRNA in
1022 hepatic and extra-hepatic tissues--a comparative study of oligonucleotide length,
1023 design and chemistry. Nucleic Acids Res 2012;40:6135-43.
1024 32. Seth PP, Siwkowski A, Allerson CR, Vasquez G, Lee S, Prakash TP, et al. Short
1025 antisense oligonucleotides with novel 2'-4' conformationaly restricted nucleoside
1026 analogues show improved potency without increased toxicity in animals. J Med
1027 Chem 2009;52:10-3.
1028 33. Teplova M, Minasov G, Tereshko V, Inamati GB, Cook PD, Manoharan M, et al.
1029 Crystal structure and improved antisense properties of 2'-O-(2-methoxyethyl)-RNA.
1030 Nat Struct Biol 1999;6:535-9.
1031 34. Chu C, Qu K, Zhong FL, Artandi SE, Chang HY. Genomic maps of long noncoding
1032 RNA occupancy reveal principles of RNA-chromatin interactions. Mol Cell
1033 2011;44:667-78.
1034 35. Chen H, Ishii A, Wong WK, Chen LB, Lo SH. Molecular characterization of human
1035 tensin. Biochem J 2000;351 Pt 2:403-11.
1036
1037 36. Hall EH, Daugherty AE, Choi CK, Horwitz AF, Brautigan DL. Tensin1 requires
1038 protein phosphatase-1alpha in addition to RhoGAP DLC-1 to control cell polarization,
1039 migration, and invasion. J Biol Chem 2009;284:34713-22.
1040 37. Gyorffy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q, et al. An online
1041 survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer

1042 prognosis using microarray data of 1,809 patients. Breast Cancer Res Treat
1043 2010;123:725-31.
1044 38. Mouneimne G, Brugge JS. Tensins: a new switch in cell migration. Dev Cell
1045 2007;13:317-9.
1046 39. Parsons JT, Horwitz AR, Schwartz MA. Cell adhesion: integrating cytoskeletal
1047 dynamics and cellular tension. Nat Rev Mol Cell Biol 2010;11:633-43.
1048 40. Balas MM, Johnson AM. Exploring the mechanisms behind long noncoding RNAs
1049 and cancer. Noncoding RNA Res 2018;3:108-17.
1050 41. Johnson EM, Daniel DC, Gordon J. The pur protein family: genetic and structural
1051 features in development and disease. J Cell Physiol 2013;228:930-7.
1052 42. Lasham A, Lindridge E, Rudert F, Onrust R, Watson J. Regulation of the human fas
1053 promoter by YB-1, Puralpha and AP-1 transcription factors. Gene 2000;252:1-13.
1054 43. Maamar H, Cabili MN, Rinn J, Raj A. linc-HOXA1 is a noncoding RNA that represses
1055 Hoxa1 transcription in cis. Genes Dev 2013;27:1260-71.
1056 44. Minn AJ, Gupta GP, Siegel PM, Bos PD, Shu W, Giri DD, et al. Genes that mediate
1057 breast cancer metastasis to lung. Nature 2005;436:518-24.
1058 45. Li J, Han L, Roebuck P, Diao L, Liu L, Yuan Y, et al. TANRIC: An Interactive Open
1059 Platform to Explore the Function of lncRNAs in Cancer. Cancer Res 2015;75:3728-
1060 37.
1061 46. Bernau K, Torr EE, Evans MD, Aoki JK, Ngam CR, Sandbo N. Tensin 1 Is Essential
1062 for Myofibroblast Differentiation and Extracellular Matrix Formation. Am J Respir Cell
1063 Mol Biol 2017;56:465-76.

1064 47. Chen H, Duncan IC, Bozorgchami H, Lo SH. Tensin1 and a previously
1065 undocumented family member, tensin2, positively regulate cell migration. Proc Natl
1066 Acad Sci U S A 2002;99:733-8.
1067 48. Saintigny G, Bernard FX, Juchaux F, Pedretti N, Mahe C. Reduced expression of
1068 the adhesion protein tensin1 in cultured human dermal fibroblasts affects collagen
1069 gel contraction. Exp Dermatol 2008;17:788-9.
1070 49. Okayama A, Miyagi Y, Oshita F, Ito H, Nakayama H, Nishi M, et al. Identification of
1071 Tyrosine-Phosphorylated Proteins Upregulated during Epithelial-Mesenchymal
1072 Transition Induced with TGF-beta. J Proteome Res 2015;14:4127-36.
1073 50. Becker C, Wirtz S, Ma X, Blessing M, Galle PR, Neurath MF. Regulation of IL-12
1074 p40 promoter activity in primary human monocytes: roles of NF-kappaB,
1075 CCAAT/enhancer-binding protein beta, and PU.1 and identification of a novel
1076 repressor element (GA-12) that responds to IL-4 and prostaglandin E(2). J Immunol
1077 2001;167:2608-18.
1078 51. Johnson EM. The Pur protein family: clues to function from recent studies on cancer
1079 and AIDS. Anticancer Res 2003;23:2093-100.
1080 52. Kelm RJ, Jr., Lamba GS, Levis JE, Holmes CE. Characterization of purine-rich
1081 element binding protein B as a novel biomarker in acute myelogenous leukemia
1082 prognostication. J Cell Biochem 2018;119:2073-83.
1083 53. Shi J, Cheng C, Ma J, Liew CC, Geng X. Gene expression signature for detection of
1084 gastric cancer in peripheral blood. Oncol Lett 2018;15:9802-10.

54. Murugesan SN, Yadav BS, Maurya PK, Chaudhary A, Singh S, Mani A. Expression and network analysis of YBX1 interactors for identification of new drug targets in lung adenocarcinoma. J Genomics 2018;6:103-12.
55. Shibata T, Tokunaga E, Hattori S, Watari K, Murakami Y, Yamashita N, et al. Y-box binding protein YBX1 and its correlated genes as biomarkers for poor outcomes in patients with breast cancer. Oncotarget 2018;9:37216-28.
56. Wang LG, Johnson EM, Kinoshita Y, Babb JS, Buckley MT, Liebes LF, et al. Androgen receptor overexpression in prostate cancer linked to Pur alpha loss from a novel repressor complex. Cancer Res 2008;68:2678-88.
57. Michailidou K, Lindstrom S, Dennis J, Beesley J, Hui S, Kar S, et al. Association analysis identifies 65 new breast cancer risk loci. Nature 2017;551:92-4.
58. Gupta GP, Massague J. Cancer metastasis: building a framework. Cell 2006;127:679-95.
59. Jin X, Mu P. Targeting Breast Cancer Metastasis. Breast Cancer (Auckl) 2015;9:23-34.
60. Hua Y, Sahashi K, Rigo F, Hung G, Horev G, Bennett CF, et al. Peripheral SMN restoration is essential for long-term rescue of a severe spinal muscular atrophy mouse model. Nature 2011;478:123-6.
61. Pandey SK, Wheeler TM, Justice SL, Kim A, Younis HS, Gattis D, et al. Identification and characterization of modified antisense oligonucleotides targeting DMPK in mice and nonhuman primates for the treatment of myotonic dystrophy type 1. J Pharmacol Exp Ther 2015;355:329-40.

62. Wheeler TM, Leger AJ, Pandey SK, MacLeod AR, Nakamori M, Cheng SH, et al. Targeting nuclear RNA for in vivo correction of myotonic dystrophy. Nature 2012;488:111-5.

63. Bennett CF, Baker BF, Pham N, Swayze E, Geary RS. Pharmacology of Antisense Drugs. Annu Rev Pharmacol Toxicol 2017;57:81-105.

64. Hong D, Kurzrock R, Kim Y, Woessner R, Younes A, Nemunaitis J, et al. AZD9150, a next-generation antisense oligonucleotide inhibitor of STAT3 with early evidence of clinical activity in lymphoma and lung cancer. Sci Transl Med 2015;7:314ra185.

65. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 2010;38:576-89.

66. Conrad T, Ørom UA. Cellular Fractionation and Isolation of Chromatin-Associated RNA. Methods Mol Biol 2017;1468:1-9.

67. Langmead, B., and Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. Nature Methods 2012;9:357–359.

68. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15-21.

69. Anders, S., Pyl, P.T., and Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 2015;31:166–169.

70. Love, M.I., Huber, W., and Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.

71. Luo, W., Friedman, M.S., Shedden, K., Hankenson, K.D., and Woolf, P.J. GAGE: generally applicable gene set enrichment for pathway analysis. BMC Bioinformatics 2009;10:161.

72. Luo, W., and Brouwer, C. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. Bioinformatics 2013;29:1830–1831.

73. Huang YH, Klingbeil O, He XY, Wu XS, Arun G, Lu B, et al. POU2F3 is a master regulator of a tuft cell-like variant of small cell lung cancer. Genes Dev 2018;32:915-28.

74. Spector DL, Smith HC. Redistribution of U-snRNPs during mitosis. Exp Cell Res 1986;163:87-94.

Figure 1. Characterization of *Mammary Tumor Associated RNA 25* (*MaTAR25*)

(A) Representation of the *MaTAR25* gene locus. *MaTAR25* is an intergenic lncRNA gene located on mouse chromosome 2, and the *MaTAR25* RNA transcript contains 2 exons and a poly (A) tail.

(B) 5' and 3' rapid amplification of cDNA ends (RACE) was performed to identify the full length *MaTAR25* transcript.

(C) The full length *MaTAR25* transcript was confirmed by Northern blot analysis to be ~2000 nt. 20 µg or 30 µg total RNA extracted from MMTV-PyMT primary cells was electrophoresed on a 1% agarose gel and probed.

(D) *In vitro* transcription and translation reactions were performed to confirm that *MaTAR25* does not produce a peptide. The reaction products were run on a 4-20% gradient SDS-PAGE gel, and the signals were detected by HRP-conjugated Streptavidin. Luciferase control DNA and *Xenopus laevis* Histone H2B (HISTH2B) expressing plasmids were used as positive controls and empty vector as a negative control.

(E) Representative smRNA-FISH images showing localization of *MaTAR25* RNA transcripts within nuclei of MMTV-PyMT and MMTV-Neu-NDL primary cells. Scale bars are 10 µm.

Figure 2. *MaTAR25* knockout affects 4T1 cell viability, migration, and invasion *in vitro*; all of which can be rescued by ectopic expression of *MaTAR25* in knockout cells

(A) CRISPR/Cas9 was used to generate *MaTAR25* KO clones in 4T1 cells. Pairs of sgRNAs were introduced targeting upstream and downstream of the transcription start site of *MaTAR25*, resulting in 390-620 bp genomic delections, and a *Renilla* Luciferase sgRNA was used as a negative control. Knockout clones were selected by genomic PCR and Sanger sequencing for homozygous genomic deletion. qRT-PCR and representative images of smRNA-FISH are shown to confirm *MaTAR25* KO. Scale bars are 5 µm.

(B) 4T1 cells were seeded at the same cell density in 12-well tissue culture plates at day 0 and cell counting was performed at different time points. The mean cell numbers of three independent replicates of 4T1 control groups and *MaTAR25* KO groups is shown ± SD (n=3). *$p < 0.05$ (student's t-test).

(C) Live cell tracking was performed over time to examine cell migration. Images were collected every 5 minutes for a total of 8 hours and analyzed by CellTracker image processing software. The mean relative migration distance (μm) of two independent replicates of 4T1 control groups and *MaTAR25* KO groups is shown ± SD (n=3). *$p < 0.05$

(D) 24-well Boyden chamber invasion assay (24 hours). The mean relative cell invasion of three independent replicates of 4T1 control groups and *MaTAR25* KO groups is shown ± SD (n=3). *$p < 0.05$ (student's t-test).

(E) Ectopic expression of *MaTAR25* or GFP was used as positive and negative controls to assess rescue in a cell viability assay, or (F) cell invasion assay. The mean cell numbers and mean relative cell invasion of three independent replicates of 4T1 Control1, *MaTAR25* KO2, *MaTAR25* KO2 with GFP expression, and *MaTAR25* KO2 with *MaTAR25* ectopic expression is shown ± SD (n=3) *$p < 0.05$ (student's t-test).

Figure 3. *MaTAR25* knockout impairs tumor growth and metastasis *in vivo*

(A) 4T1 control or *MaTAR25* KO cells were injected orthotopically into the mammary fat pad of female BALB/c mice. Primary tumors were measured every week over a period of four weeks and the mean tumor volume of 8 mice per group is shown ± SE. *$p < 0.05$ (student's t-test).

(B) Mice were sacrificed and tumors were collected at day 28 to compare the tumor growth rate between the control group and *MaTAR25* KO groups. Tumors derived from *MaTAR25* KO cells showed a 56% reduction in tumor growth. The mean tumor wet weight is shown ± SE. *$p < 0.05$ (student's t-test).

(C) Female BALB/c mice were injected into the tail vein with 4T1 Control1 or *MaTAR25* KO cells. Mice were monitored every week and sacrificed at day 21. Mouse lungs were collected and imaged (left panel), and lung metastatic nodules were counted to

compare the metastatic ability between the control group and *MaTAR25* KO group (right panel). Mice injected with *MaTAR25* KO cells exhibited a 62% reduction in lung metastatic nodules.

(D) Schematic showing the approach for ASO mediated knockdown of *MaTAR25* in MMTV-Neu-NDL mice. Two independent *MaTAR25* ASOs or a control scASO were used for subcutaneous injection. Primary tumors were measured twice per week and the mean tumor volume of 7 mice per group is shown ± SE. *$p < 0.05$ (student's t-test).

(E) Representative hematoxylin and eosin (H&E) stained tumor images showing the different histological phenotypes between tumor samples from scASO injected mice and *MaTAR25* ASO injected groups.

Figure 4. *MaTAR25* is a positive upstream regulator of *Tns1*

(A) Cell fractionation was performed to isolate cytoplasmic, nucleoplasmic, and chromatin associated RNA. qRT-PCR was used to determine the subcellular localization ratio of *MaTAR25* transcripts. *β-actin* and *Malat1* were used as marker RNAs for quality control of cell fractionation.

(B) Schematic diagram showing the targeting of biotin labeled oligonucleotides binding *MaTAR25* transcripts for chromatin isolation by RNA purification (ChIRP)-seq. Odd and even oligo pools (7 oligos binding different regions within the *MaTAR25* transcript in each pool) were used for ChIRP-seq, and qRT-PCR was performed to assess RNA purification enrichment.

(C) Venn diagram showing differentially expressed genes in *MaTAR25* KO cells identified from RNA-Seq overlapped with *MaTAR25* ChIRP-seq data. A total 446 overlapping genes were identified, and the top candidate genes are listed.

(D) Validation of *Tns1* as a *MaTAR25* targeted gene by qRT-PCR and immunoblotting in 4T1 control and *MaTAR25* KO cells.

(E) The RNA expression level of *Tns1* is rescued upon ectopic expression of *MaTAR25* in *MaTAR25* KO cells as determined by qRT-PCR.

(F) CRISPR/Cas9 targeting was used in 4T1 cells to generate *Tns1* knockout clones. The upper panel shows expression levels of *Tns1* in 4T1 control, *Tns1* KO Clone1, and *Tns1* KO Clone2 by immunoblotting. The lower panel shows the cell counting viability assay results of three independent replicates of 4T1 Control1, *Tns1* KO1, and *Tns1* KO2. Results are mean ± SD (n=3) *$p < 0.05$ (student's t-test).

(G) Ectopic expression of *Tns1* in *MaTAR25* KO cells rescues the cell viability defect. The top panel shows expression levels of *Tns1* in 4T1 control1, *MaTAR25* KO2, *MaTAR25* KO2 with *Tns1* ectopic expression Clone1. The bottom panel shows the cell counting viability assay results of three independent replicates of 4T1 Control1, *MaTAR25* KO2, *MaTAR25* KO2 with *MaTAR25* ectopic expression, and *MaTAR25* KO with *Tns1* ectopic expression Clone1-3. Results are mean ± SD (n=3) *$p < 0.05$ (student's t-test).

Figure 5. *MaTAR25* interacts with PURB to carry out its function

(A) Scatterplot depicts the fold enrichment of protein candidates from isobaric tags for the relative and absolute quantitation (iTRAQ) analysis comparing two independent oligo pair sets targeting *MaTAR25* RNA transcripts vs PPIB RNA transcripts.

(B) Upper: immunoblot analysis of PURA and PURB following pull-down of *MaTAR25* or PPIB from 4T1 cells. Lower: immunoblot analysis of PURB following the pull-down of *MaTAR25* or PPIB from 4T1 cells or 4T1 *MaTAR25* KO cells.

(C) *MaTAR25*, PPIB, and *Gapdh* transcripts were assessed by qRT-PCR in endogenous PURB, or IgG (negative control) immunoprecipitates from 4T1 cells. Fold enrichment over IgG signal is shown ± SD (n=3). Immunoblot analysis of PURB was performed as a control.

(D) qRT-PCR analysis and immunoblotting of *Tns1* expression in 4T1 cells following ectopic over-expression of PURB. The relative expression levels are shown ± SD (n=3).

(E) ChIP-qPCR analysis of PURB occupancy over the identified *MaTAR25* targeting region and non-targeting region of the *Tns1* DNA locus by ChIRP-seq analysis.

ChIP-qPCR was performed in 4T1 control cells, 4T1 *MaTAR25* KO cells, and upon ectopic expression of *MaTAR25* in *MaTAR25* KO cells. Primers for a *MaTAR25* non-targeting region and the *Gapdh* TSS were used as negative controls. Bar graphs represent the mean ± SD (n=2).

Figure 6. *LINC01271* is the human ortholog of *MaTAR25*

(A) All potential human orthologs of *MaTAR25* (*hMaTAR25*) were identified based on conservation of genomic location (synteny). RNA-seq data from The Cancer Genome Atlas (TCGA) was analyzed to evaluate the expression status of all potential *hMaTAR25* candidates by comparison of 1128 TCGA breast tumor datasets to 113 normal breast tissue controls. Fold change and statistical significance was calculated for the entire data matrix using DESeq2 (29).

(B) Attempted rescue of 4T1 *MaTAR25* KO cells upon independent ectopic expression of two transcript isoforms of *LINC01270* (*LINC01270.1*, *LINC01270.2*), *or LINC01271* in cell viability assays. The mean cell numbers of three independent replicates of 4T1 control, *MaTAR25* KO, *MaTAR25* KO with GFP, *LINC01270.1*, *LINC01270.2*, and *LINC01271* is shown ± SD (n=3) **p* < 0.05 (student's t-test). Only ectopic expression of *LINC01271* can rescue the *MaTAR25* KO cell viability phenotype.

(C) 4T1 *MaTAR25* KO cells with ectopic expression of GFP was used as a control to assess rescue in a cell invasion assay. The mean relative cell invasion of two independent replicates of 4T1 control, *MaTAR25* KO, *MaTAR25* KO with GFP, *LINC01271* ectopic expression is shown ± SD (n=2) **p* < 0.05 (student's t-test). Ectopic expression of *LINC01271* can rescue the *MaTAR25* KO cell invasion phenotype.

(D) RNA expression level of *Tns1* was determined in *MaTAR25* KO cells ectopically expressing *LINC01270.1*, *LINC01270.2*, or *LINC01271* by qRT-PCR. The protein level of TNS1 was also examined in *MaTAR25* KO cells with ectopic expression of *LINC01271* by immunoblot analysis.

(E) Three different ASOs targeting *LINC01271* were used to independently knockdown *LINC01271* in MDA-MB-231 LM2 cells. Left panel: the knockdown efficiency is shown ± SD (n=3) by qRT-PCR after 24 hours treatment of ASOs. Cells were seeded at the same density ($5x10^4$/well) in 12-well tissue culture plates at day 0, ASOs were added to the culture medium and cell counting was performed at different time points to measure cell numbers. Right panel: The mean cell numbers of three independent replicates of MDA-MB-231 LM2 mock treated control cells, MDA-MB-231 LM2 cells treated with scrambled ASO, MDA-MB-231 LM2 cells independently treated with 3 different *LINC01271* ASOs is shown ± SD (n=3) *$p < 0.05$ (student's t-test).

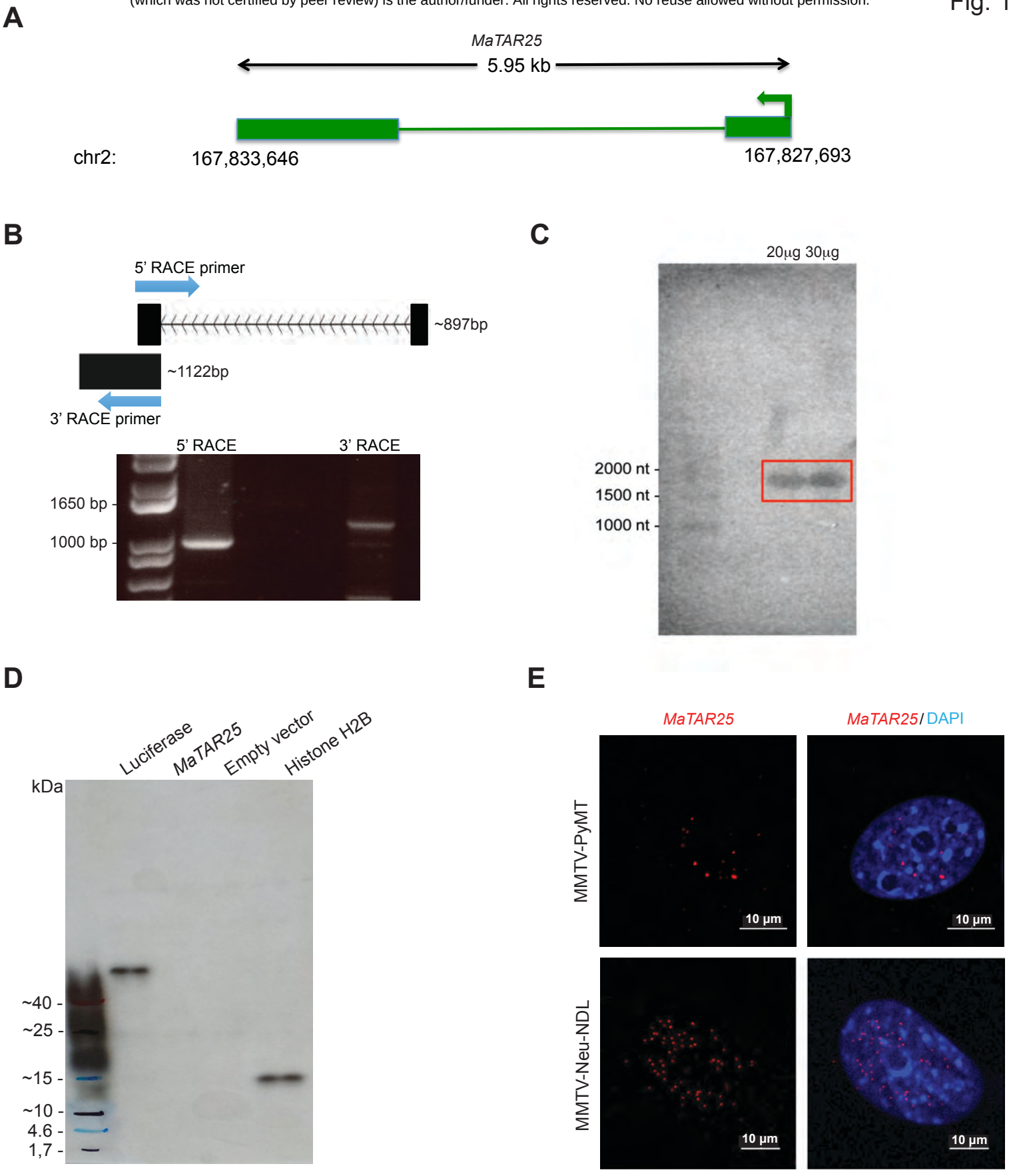
Figure 7. *LINC01271* expression in breast tumors and lung metastases

(A) Representative smRNA-FISH images showing the expression of *LINC01271* in human breast tumor sections from different stages of breast cancer. Scale bars are 20 μm.

(B) Representative smRNA-FISH images showing the expression pattern of *LINC01271* within luminal subtype human primary breast tumors and lung metastases sections from the same patients. Scale bars are 20 μm.

(C) Working model of *MaTAR25* function.

Fig. 1



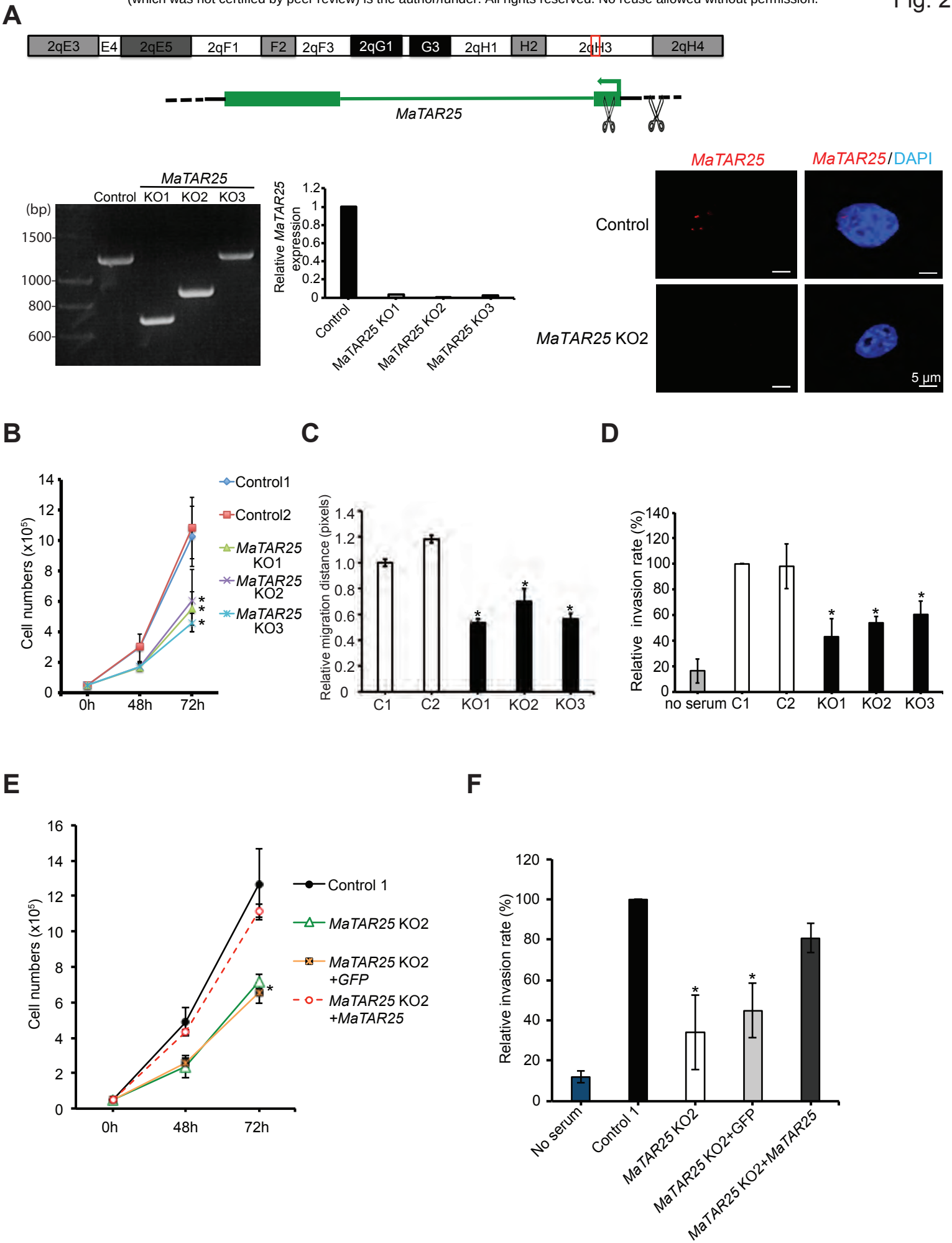


Fig. 3

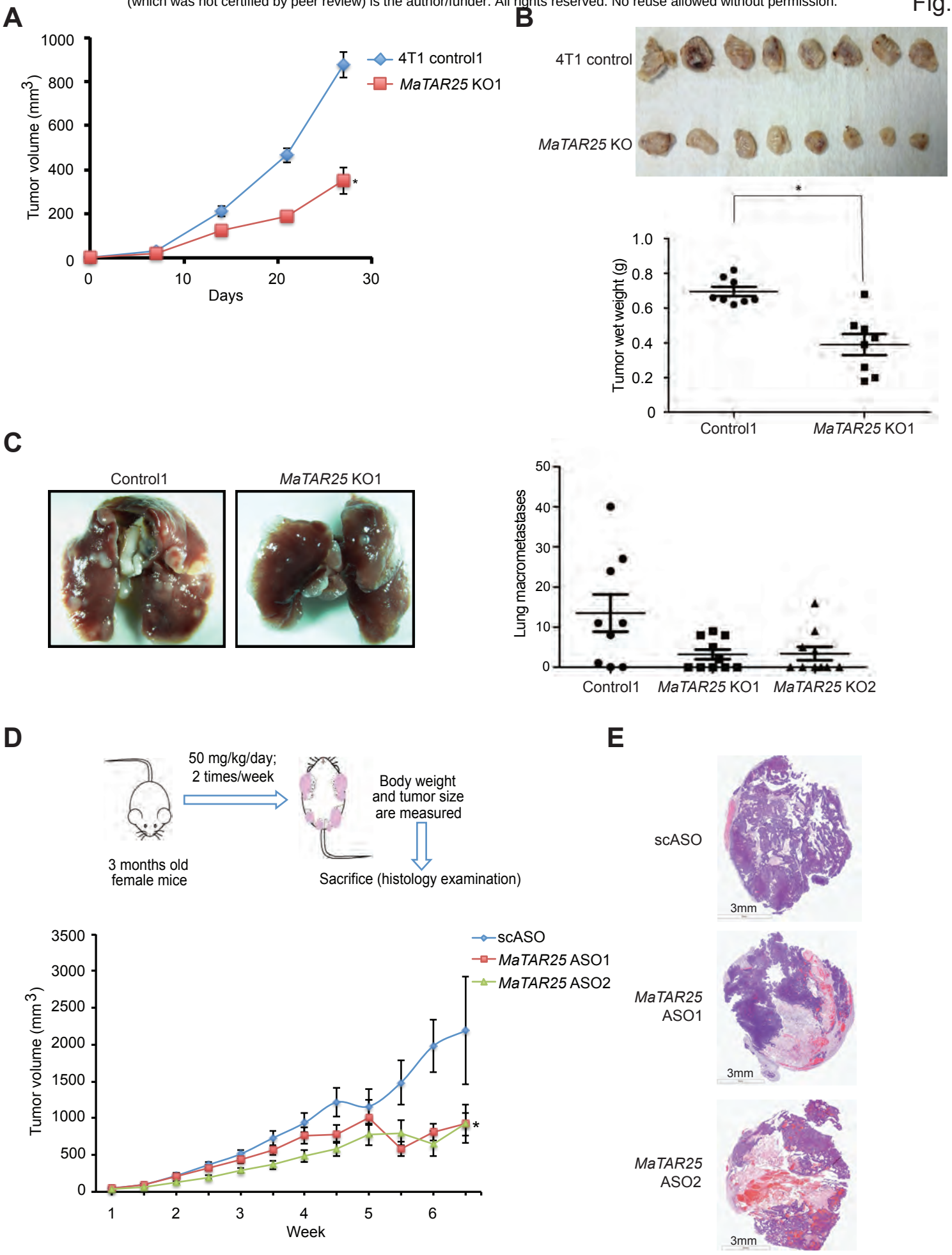
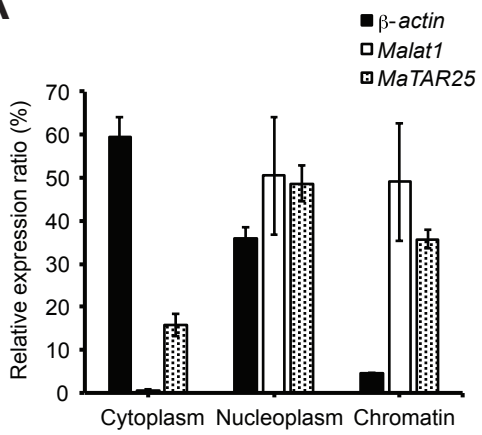
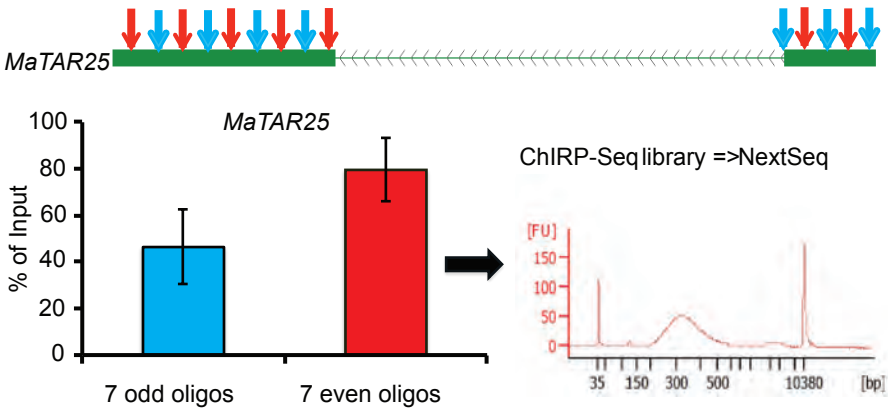


Fig. 4

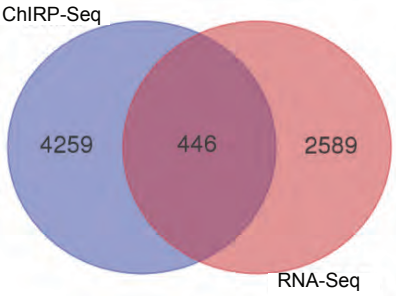
A



B



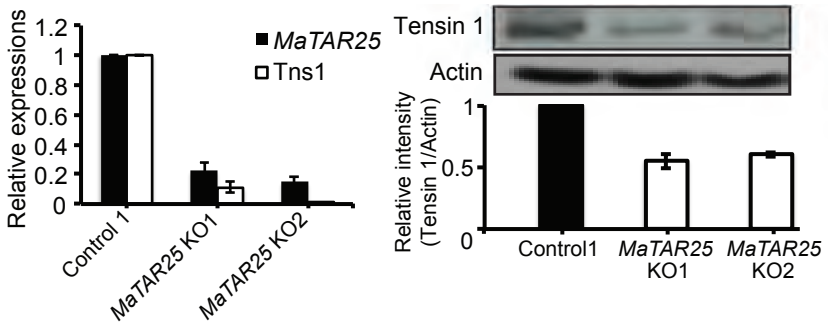
C



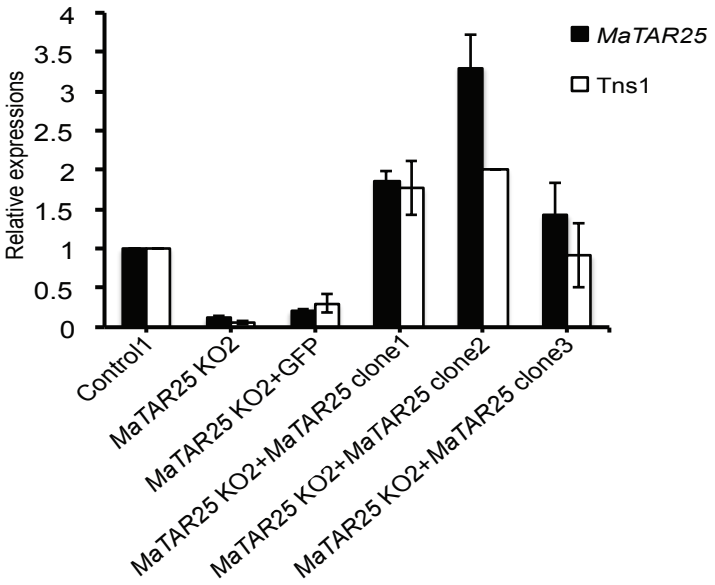
MaTAR25 odd & even oligo pools overlapping targets

Chr	gene	fold change in KO	p-value
2	MaTAR25	0.219515585	2.38E-05
1	Tns1	0.379721873	0.0093756
15	Pvt1	0.624304762	0.0025525
8	Lonp2	0.625035638	0.0007314
15	Zhx2	0.428373444	1.24E-07
11	Ggnbp2	0.653081803	0.0148347
10	Ilvbl	0.755908045	0.0143116
12	Immp2l	0.57074968	0.0027005
7	Cebpg	0.360909244	6.84E-05

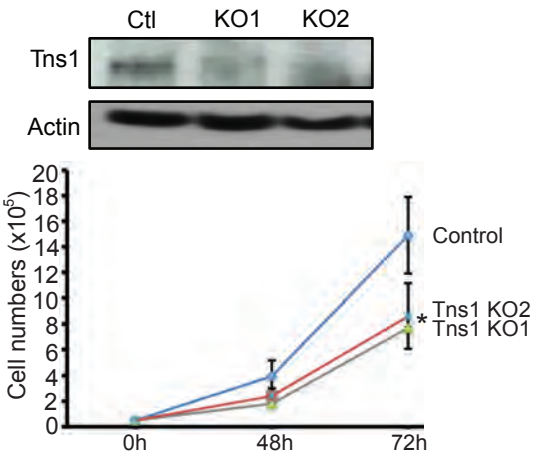
D



E



F



G

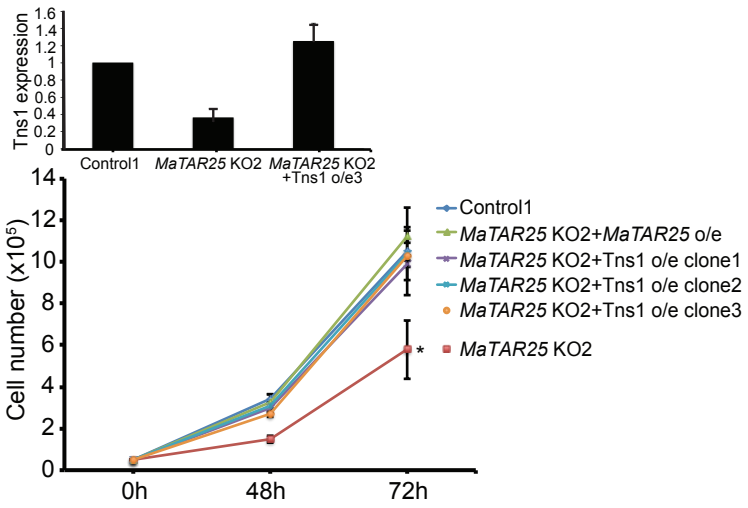
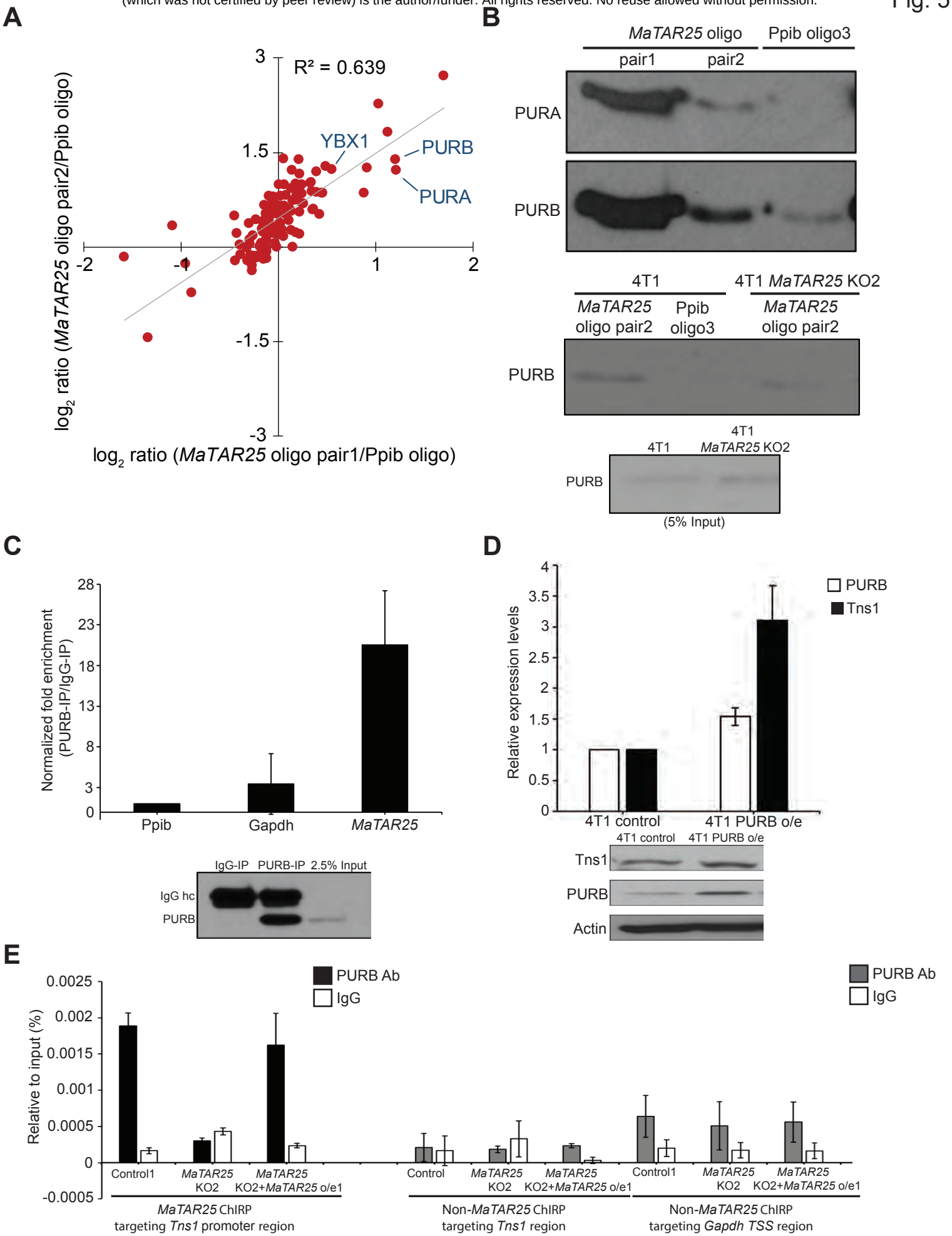
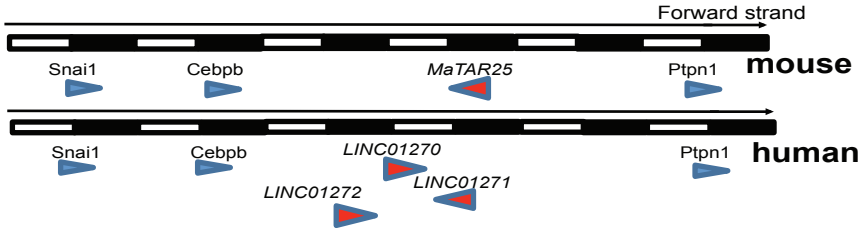


Fig. 5

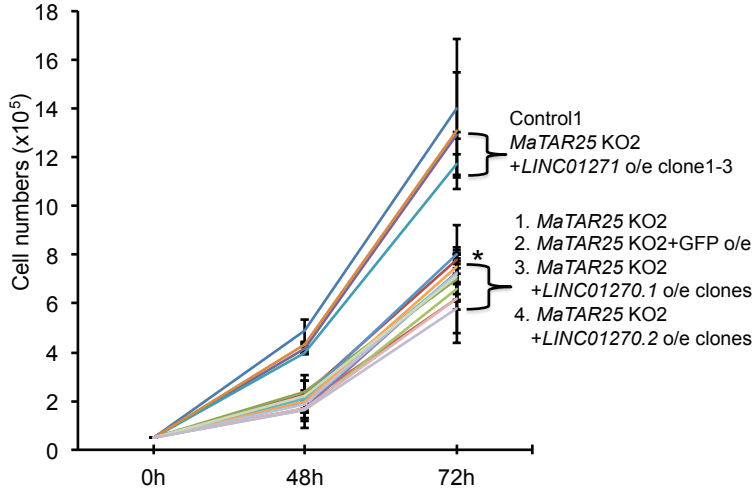


A

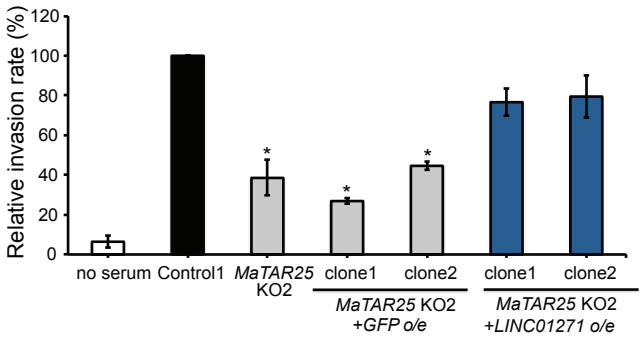


	Potential human counterpart	Fold change	P value	
hMaTAR25.1	ENSG00000203999	1.6	1.88E-06	*LINC01270*
hMaTAR25.2	ENSG00000233077	1.7	8.67E-06	*LINC01271*
hMaTAR25.3	ENSG00000224397	1.04	7.18E-01	*LINC01272*

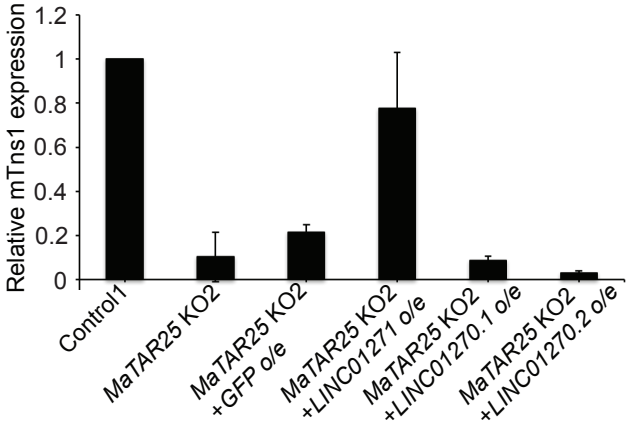
B

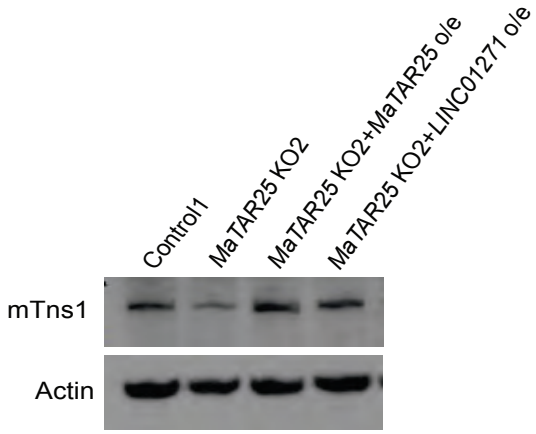


C



D





E