nature structural & molecular biology

Long noncoding RNA LINP1 regulates repair of DNA double-strand breaks in triple-negative breast cancer

Youyou Zhang[1,9], Qun He[1,9], Zhongyi Hu[1], Yi Feng[1,2], Lingling Fan[1], Zhaoqing Tang[1], Jiao Yuan[1], Weiwei Shan[1], Chunsheng Li[1,3], Xiaowen Hu[1,3], Janos L Tanyi[3], Yi Fan[4], Qihong Huang[5], Kathleen Montone[6], Chi V Dang[2,7,8] & Lin Zhang[1,3,8]

Long noncoding RNAs (lncRNAs) play critical roles during tumorigenesis by functioning as scaffolds that regulate protein-protein, protein-DNA or protein-RNA interactions. Using a clinically guided genetic screening approach, we identified lncRNA in nonhomologous end joining (NHEJ) pathway 1 (LINP1), which is overexpressed in human triple-negative breast cancer. We found that LINP1 enhances repair of DNA double-strand breaks by serving as a scaffold linking Ku80 and DNA-PKcs, thereby coordinating the NHEJ pathway. Importantly, blocking LINP1, which is regulated by p53 and epidermal growth factor receptor (EGFR) signaling, increases the sensitivity of the tumor-cell response to radiotherapy in breast cancer.

Triple-negative breast cancer (TNBC), an aggressive subtype associated with poor clinical outcomes, represents approximately 10–20% of breast cancer cases[1–3]. A large proportion of TNBC (50–75%) exhibits the molecular subtype known as basal-like breast cancer, characterized by high expression of genes that are normally expressed in the basal epithelial layer. The absence of estrogen-receptor and progesterone-receptor expression and Her2 amplification limits the therapeutic options for this disease to surgery with adjuvant chemotherapy and radiotherapy[1–3]. Currently, there are no effective targeted therapies, although *EGFR* amplification, *TP53* mutation, *BRCA1* and *BRCA2* loss and PI3-kinase-pathway activation have been exploited for TNBC treatment. Given the lack of recurrent, targetable genomic alterations, functional characterization of the TNBC genome is crucial to identify driver genomic events[1–3]. The human genome contains ~20,000 protein-coding genes (PCGs), representing less than 2% of the total genome, whereas nearly 70% of the human genome is transcribed into RNA, thus yielding thousands of noncoding RNAs[4]. However, because genomic studies of TNBC have mainly focused on PCGs, the functions of noncoding genes remain largely unknown.

lncRNAs are defined as RNA transcripts >200 nt that lack apparent protein-coding potential[5–13]. More than 15,900 lncRNA genes have recently been identified in the human genome, on the basis of GENCODE annotations[4]. Notably, their expression is strikingly cell-type- or tissue-restricted and, in many cases, is even primate specific. Investigations of lncRNAs have demonstrated that they can serve as scaffolds or guides regulating protein-protein or protein-DNA interactions, as decoys that bind proteins or microRNAs (miRNAs), or as enhancers of gene expression when transcribed within enhancer regions or their neighboring loci. Because of the highly dysregulated expression of lncRNAs in cancer[14,15], lncRNAs have been surmised to contribute to tumorigenesis. In fact, certain lncRNAs have been shown to function as oncogenes or tumor suppressors[5–13]. For example, HOTAIR induces breast cancer metastasis[16] by operating as a tether that links EZH2 (PRC2) and LSD1, thereby coordinating their epigenetic regulatory functions[17]. LINK-A promotes metabolic reprogramming toward glycolysis as well as tumorigenesis, and its expression is increased in TNBC[18].

DNA repair, a collection of processes by which damaged DNA is identified and corrected in cells, is essential to genomic integrity and is involved in tumorigenesis. Although multiple proteins that mediate DNA repair have been identified, it is presently unknown whether RNA molecules are also components of the DNA-repair machinery. NHEJ is one of the major pathways for repairing damaged DNA in cancer cells[19–26]. In response to DNA double-strand breaks (DSBs), the Ku80–Ku70 heterodimer associates with the broken ends, forming a clamp-like complex that recruits the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) to sites of damage. Additional protein factors, including Artemis, DNA ligase IV, XRCC4, and XLF assemble with the Ku80–Ku70–DNA-PK complex and promote processing and ligation of the broken ends[19–26]. To identify lncRNAs that are functionally involved in tumorigenesis of TNBC, we analyzed the expression profile of lncRNAs in The Cancer Genome Atlas (TCGA) breast cancer data sets and performed a clinically guided genetic screening in TNBC cell lines.

[1]Center for Research on Reproduction & Women's Health, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [2]Abramson Family Cancer Research Institute, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [3]Department of Obstetrics and Gynecology, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [4]Department of Radiation Oncology, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [5]Wistar Institute, Philadelphia, Pennsylvania, USA. [6]Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [7]Department of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [8]Abramson Cancer Center, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [9]These authors contributed equally to this work. Correspondence should be addressed to C.V.D. (dangvchi@exchange.upenn.edu) or L.Z. (linzhang@mail.med.upenn.edu).

Received 16 February; accepted 24 March; published online 25 April 2016; doi:10.1038/nsmb.3211

RESULTS

Identification of the TNBC-associated lncRNA LINP1

To identify lncRNAs associated with TNBC, we analyzed differences in lncRNA expression among the distinct pathological and molecular subtypes of breast cancers in the TCGA data set[3] (**Fig. 1a**). We found 330 (fold change ≥2) and 45 (fold change ≥5) lncRNAs whose expression was significantly higher in TNBC compared with non-TNBC tumors. When the nonbasal tumors (luminal A, luminal B and Her2 enriched) were treated as a whole, a total of 402 (fold change ≥2) and 69 (fold change ≥5) lncRNAs were expressed at a significantly higher level in basal tumors. When we compared four molecular subtypes individually, we identified 164 (fold change ≥2) and 75 (fold change ≥5) lncRNAs that were specifically enriched in basal subtypes. By cross-comparing the three gene lists, we found an overlap of 154 (fold change ≥2) and 35 (fold change ≥5) lncRNAs whose expression was enriched in TNBC tumors (**Fig. 1b** and **Supplementary Table 1**), including the most recently identified TNBC-enriched lncRNA, LINK-A[18]. To identify the lncRNAs that are functionally involved in TNBC, we performed short interfering RNA (siRNA) screening in MDA-MB-231 cells. Of the 35 highly enriched lncRNA candidates, expression of 20 was detected in MDA-MB-231 cells. We designed 40 siRNAs targeting 20 lncRNA candidates (**Supplementary Table 2**) and individually transfected them into MDA-MB-231 cells, which we treated with doxorubicin (a first-line chemotherapy drug for TNBC[1–3]) 48 h after siRNA transfection (**Supplementary Fig. 1a**). We measured apoptosis via caspase3 activity 24 h after doxorubicin treatment and identified the lncRNA ENSG00000223784 (LINP1) as a strong candidate.

To corroborate the above findings, we analyzed RNA-seq data of breast cancer cell lines (n = 46) from the Cancer Cell Line Encyclopedia (CCLE) data set[27]. We found that LINP1 was expressed at a significantly higher level in basal lines than in nonbasal lines, in agreement with TCGA data (**Fig. 1c** and **Supplementary Fig. 1b**). We chose two TNBC lines (MDA-MB-231 and MDA-MB-468), a triple-negative immortalized breast line (MCF10A), and one estrogen-receptor-positive line (MCF7) as models for functional assays. Northern analysis confirmed the RNA-seq results, revealing that LINP1 expression was highly expressed in TNBC lines as well as MCF10A cells but was undetectable in MCF7 cells (**Supplementary Fig. 1b**,**c**). As anticipated, transduction of LINP1 siRNAs in the three

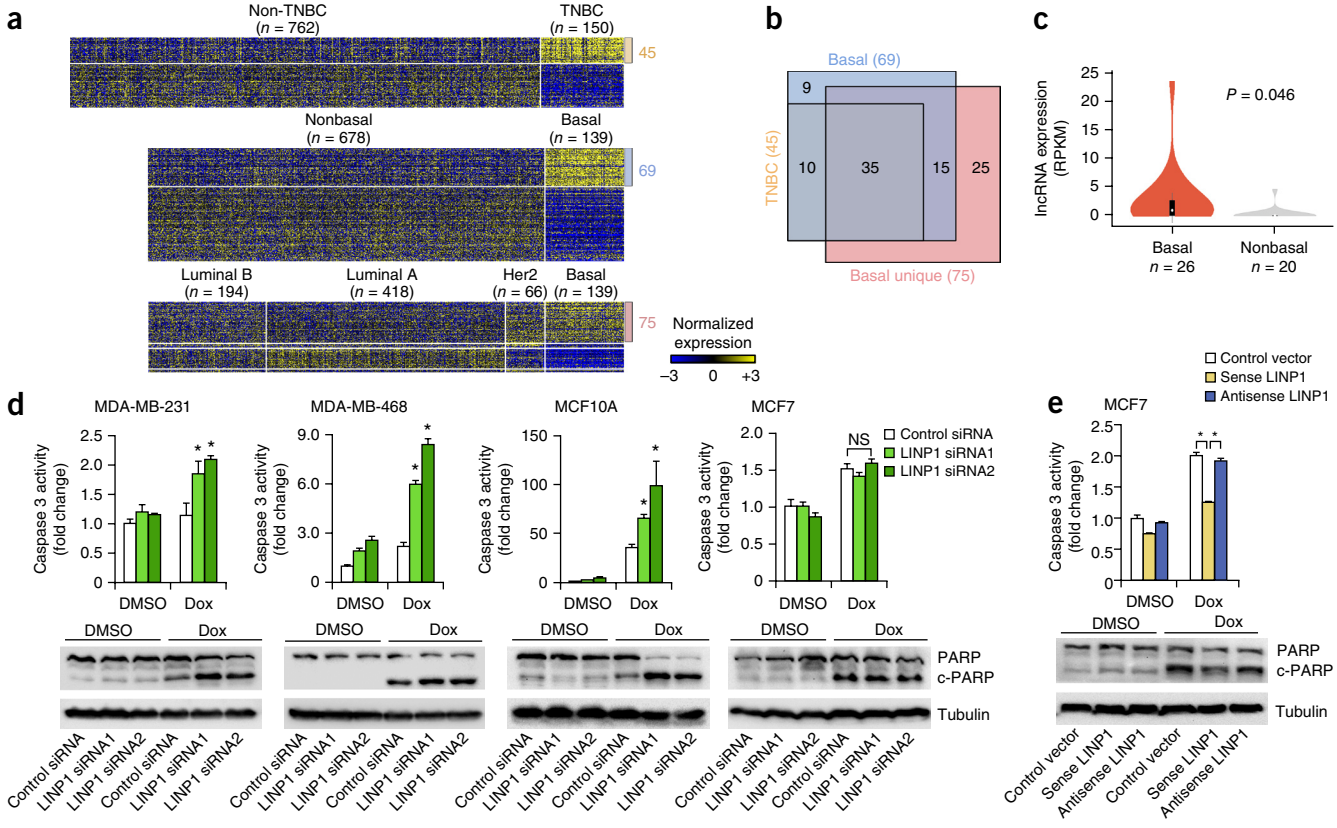

Figure 1 Identification of the TNBC-associated lncRNA LINP1. (**a**) Heat map of lncRNAs with significantly different expression between groups (fold change ≥5; P <0.05 by two-tailed Student's t test), stratified by histological or PAM50 molecular subtypes n, number of tumor samples used for statistical analysis. Detailed information for the TCGA samples is provided in **Supplementary Data Set 1**. Expression differences among the distinct pathological and molecular subtypes of breast cancers were analyzed by BRB-ArrayTools. Top, TNBC versus non-TNBC samples. Middle, basal versus nonbasal samples. Bottom, comparison across all four subtypes. Yellow, high expression; blue, low expression. (**b**) Venn diagram of the three groups of genes identified, through the comparisons in **a**, as having significantly enriched expression in TNBC tumors. (**c**) Violin plot of LINP1 expression in basal (n = 26) versus nonbasal (n = 20) breast cancer cell lines from CCLE. P = 0.046 by two-tailed Student's t test. RPKM, reads per kilobase per million mapped reads. (**d**) Top, caspase3 activity assay in cells expressing control or LINP1 siRNAs, treated with doxorubicin (Dox) or DMSO. Error bars, s.d. *P < 0.05 by two-tailed Student's t test; n = 3 independent cell cultures; NS, not significant. Bottom, expression of poly(ADP-ribose) polymerase (PARP) and cleaved PARP (c-PARP), assessed by western blotting in cells expressing control or LINP1 siRNAs. Tubulin denotes β-tubulin loading control. (**e**) Top, caspase3 activity assay in MCF7 cells expressing control vector, sense LINP1, or antisense LINP1. Error bars, s.d. *P < 0.05 by two-tailed Student's t test; n = 3 independent cell cultures. Bottom, expression of PARP and cleaved PARP, detected by western blotting in MCF7 cells expressing control vector, sense LINP1, or antisense LINP1. Uncropped images of gels are shown in **Supplementary Data Set 2**.

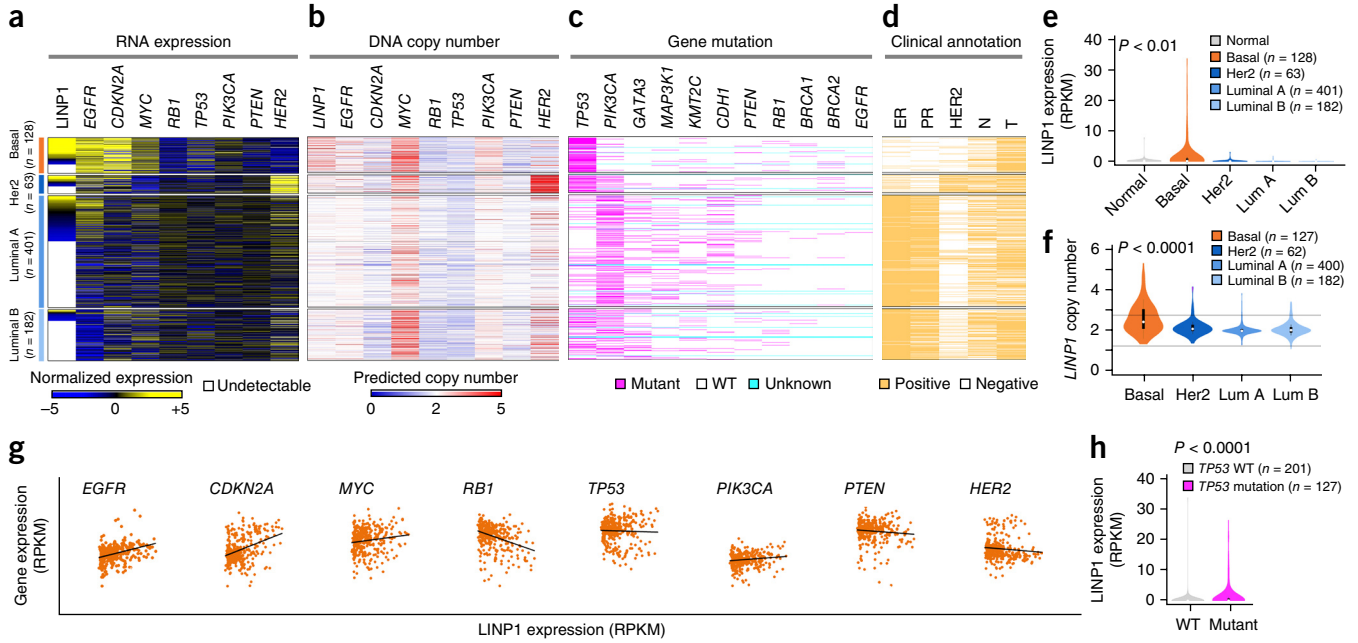

Figure 2 Expression and genomic alteration of LINP1 in breast cancer, (**a**–**d**) Heat maps of the RNA expression (**a**), somatic-gene copy-number alterations (**b**), and mutation status (**c**) of LINP1, and key molecular alterations of breast cancer and the clinical annotation (**d**) in breast cancer samples. The genomic information was retrieved from TCGA and was visualized by Multi-Experiment Viewer. The samples are grouped according to their molecular subtypes in rows and are ranked by LINP1 RNA expression level in each subtype. The genomic alteration of other genes and clinical profile are mapped in columns. (**e**) Violin plot showing LINP1 expression in different breast cancer subtypes. RPKM, reads per kilobase per million mapped reads. $P < 0.01$ by two-tailed Student's *t* test. (**f**) Violin plot showing *LINP1* gene copy-number alterations in different breast cancer subtypes. $P < 0.0001$ by two-tailed Student's *t* test. (**g**) Correlation between LINP1 expression and expression of *EGFR*, *CDKN2A*, *MYC*, *RB1*, *TP53*, *PIK3CA*, *PTEN* and *HER2* (official symbol *ERBB2*) in the breast cancer samples (336 samples in which LINP1 expression was detectable were used in correlation analysis). (**h**) Violin plot showing LINP1 expression in breast cancer with WT or mutant *TP53*. $P < 0.0001$ by two-tailed Student's *t* test. Throughout figure, *n* indicates number of tumor samples used for statistical analysis. Detailed information for the TCGA samples is provided in **Supplementary Data Set 1**.

triple-negative lines significantly enhanced doxorubicin-induced apoptosis but had no effect on MCF7 cells, according to both the caspase3 assay and western blot analysis (**Fig. 1d**). Importantly, overexpression of LINP1 RNA in MCF7 cells by lentiviral infection, compared with either empty vector and antisense controls, protected the cells from doxorubicin-induced apoptosis (**Fig. 1e**). Together, the combination of lncRNA profiling and functional screening identified LINP1 as a lncRNA that may be involved in TNBC.

Expression and genomic alteration of LINP1 in breast cancer

To define the molecular and pathological relevance of LINP1 in TNBC, we extracted the expression and copy-number alteration data for LINP1 and key known breast cancer–associated genomic alterations, along with the clinical annotations, from TCGA (**Fig. 2a**–**d**). We observed significantly higher LINP1 expression in basal breast cancer (**Fig. 2e**), a result consistent with our earlier findings (**Fig. 1a**). Notably, the somatic copy number of *LINP1* was significantly amplified in basal breast cancer (**Fig. 2f**), and there was a significant positive correlation between the gene copy number and RNA expression of LINP1 in the breast cancer samples ($R = 0.26$). These observations suggest that gains in the somatic copy number of the *LINP1* gene is a mechanism by which the RNA expression level of LINP1 is increased in basal breast tumors. We further analyzed the correlation between the LINP1 expression and expression of key molecular markers for breast cancer. We found that whereas LINP1 expression was positively correlated with expression of *EGFR* and *CDKN2A* mRNAs, it was negatively correlated with *RB1* expression (**Fig. 2g**). Intriguingly, we observed that the expression of LINP1 was significantly higher in cells expressing *TP53* mutants rather than wild-type (WT) *TP53* (**Fig. 2h**). Finally, we found that the LINP1 RNA was present in normal breast tissues (**Supplementary Fig. 2a**) and was distributed in both nuclear and cytoplasmic fractions of cells (**Supplementary Fig. 2b**).

LINP1 associates with proteins of the NHEJ pathway

To explore the molecular mechanisms underlying the biological activity of LINP1, we used an RNA pulldown assay followed by MS to identify LINP1-associated proteins (**Fig. 3a** and **Supplementary Table 3**). Interestingly, two proteins involved in the NHEJ pathway[19–26], Ku80 and DNA-PKcs, were present only in LINP1-associated samples (**Supplementary Fig. 3a**). To verify this result, we analyzed the lncRNA-pulldown protein samples by western blotting with antibodies specific to Ku80 or DNA-PKcs. We observed strong signals for Ku80 and DNA-PKcs in proteins pulled down with LINP1 RNA but not in proteins associated with either antisense LINP1 or beads alone (**Fig. 3b**), thus confirming that Ku80 and DNA-PKcs are indeed specifically enriched in the LINP1-associated protein complex. To confirm that the association between LINP1 and Ku80 DNA-PKcs was not an *in vitro* artifact, we tested the interaction between endogenous LINP1 and these two proteins by capture hybridization analysis of RNA targets (CHART)[28] (**Fig. 3c**,**d**). LINP1 was enriched from the cross-linked chromatin extracts by C oligonucleotides (oligos) 1.2 and 1.4 but not by C oligo 1.2S and 1.4S (**Fig. 3e**). Furthermore, western analysis detected Ku80 and DNA-PKcs only in the complexes that were enriched by C oligos 1.2 and 1.4 (**Fig. 3f**). Together, these results strongly suggest that endogenous LINP1 interacts with the Ku80–Ku70 heterodimer and DNA-PKcs. To confirm the interaction of LINP1 and Ku80 DNA-PKcs, we performed RNA-immunoprecipitation assays (RNA-IP) in which RNA–protein complexes were immunoprecipitated

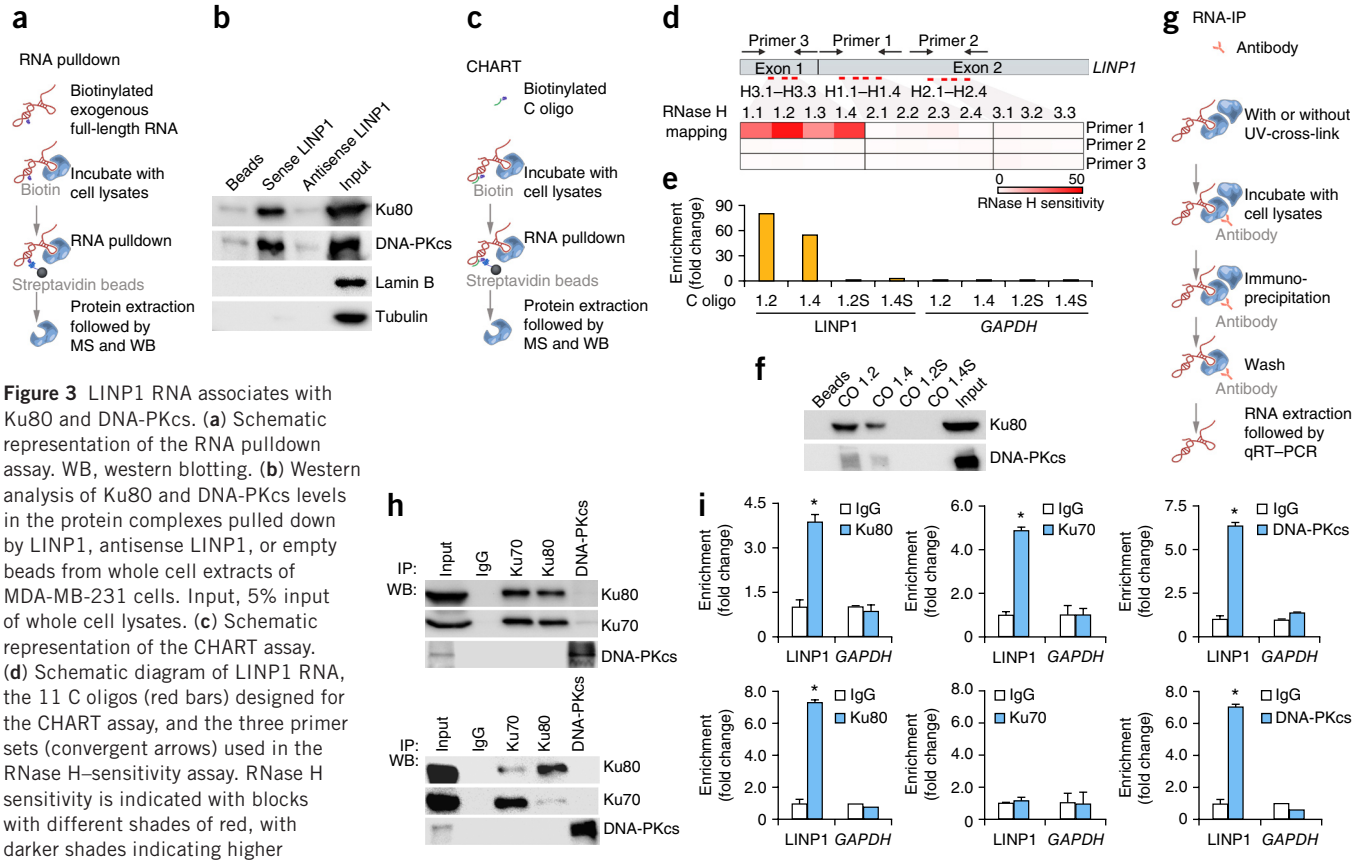

Figure 3 LINP1 RNA associates with Ku80 and DNA-PKcs. (**a**) Schematic representation of the RNA pulldown assay. WB, western blotting. (**b**) Western analysis of Ku80 and DNA-PKcs levels in the protein complexes pulled down by LINP1, antisense LINP1, or empty beads from whole cell extracts of MDA-MB-231 cells. Input, 5% input of whole cell lysates. (**c**) Schematic representation of the CHART assay. (**d**) Schematic diagram of LINP1 RNA, the 11 C oligos (red bars) designed for the CHART assay, and the three primer sets (convergent arrows) used in the RNase H–sensitivity assay. RNase H sensitivity is indicated with blocks with different shades of red, with darker shades indicating higher sensitivity. (**e**) Enrichment of LINP1 with C oligos 1.2 and 1.4, as measured by qRT–PCR. *GAPDH* is a negative control. (**f**) Presence of Ku80 and DNA-PKcs in complexes enriched in C oligos 1.2 and 1.4, detected by western analysis. C oligos 1.2S and 1.4S are nonspecific controls. (**g**) Schematic representation of an RNA-IP assay. (**h**) Results from native RNA-IP (top) and cross-linking and immunoprecipitation (CLIP) (bottom) with Ku80-, Ku70-, and DNA-PKcs-specific antibodies. Western blots showing the immunoprecipitation efficiency of Ku80, Ku70 and DNA-PKcs. (**i**) Results from native RNA-IP (top) and CLIP (bottom) with Ku80-, Ku70-, and DNA-PKcs-specific antibodies. qRT–PCR assays showing the level of LINP1 and *GAPDH* (nonspecific control) in the coprecipitates. Ku80 (left), Ku70 (middle) and DNA-PKcs (right). Error bars, s.d. $*P < 0.05$ by two-tailed Student's *t* test; $n = 3$ technical replicates. Uncropped images of gels are shown in **Supplementary Data Set 2**.

with Ku80-, Ku70-, or DNA-PKcs-specific antibodies (**Fig. 3g**,**h**). Compared with the IgG-bound complexes, the complexes bound by Ku80-, Ku70-, and DNA-PKcs-specific antibodies had significantly higher levels of LINP1 RNA. We observed no enrichment of the negative control *GAPDH* in the complexes immunoprecipitated by antibodies specific to Ku80, Ku70, or DNA-PKcs (**Fig. 3i**). In addition, RNA-IP experiments with UV-cross-linked RNA indicated that LINP1 was associated with the Ku80–Ku70 heterodimer and that LINP1 appeared to directly bind Ku80 but not to Ku70 (**Fig. 3i**). Finally, we mapped the regions of LINP1 that interact with Ku80 and DNA-PKcs by using RNA pulldown assays and identified a 300-nt region in the 5′ region of the LINP1 transcript (nts 1–300) that was essential for interaction with Ku80, and a 317-nt region within the 3′ region (nts 600–917) that was required for interaction with DNA-PKcs (**Supplementary Fig. 3b**). In aggregate, our findings indicate that Ku80 and DNA-PKcs are LINP1-associated proteins and that LINP1 uses distinct regions to interact with these two proteins.

LINP1 serves as a modular scaffold in the NHEJ pathway

Given that LINP1 RNA binds to Ku80 and DNA-PKcs, two proteins with established roles in the NHEJ pathway[19–26], we hypothesized that LINP1 might play a role in DSB repair. We therefore examined the effect of LINP1 knockdown on the repair of ionizing radiation (IR)-induced DNA damage, by using comet assays. Whereas the level of DNA damage gradually returned to the baseline in the control cells 24 h after IR treatment, it remained high in the LINP1-knockdown cells, thus suggesting that DNA repair was delayed in cells with LINP1 inhibition (**Fig. 4a**,**b**). We further confirmed this result on the basis of differences in phosphorylated histone H2AX (γ-H2AX) levels at various time points after IR treatment; the LINP1-knockdown cells, compared with control cells, had higher levels of γ-H2AX for prolonged time periods (**Supplementary Fig. 4a**). Whereas the level of γ-H2AX in the control cells at 24 h after IR treatment was comparable to that at 0 h, γ-H2AX levels remained high in the LINP1-knockdown cells. Furthermore, we counted the number of γ-H2AX-positive foci formed in the control and LINP1-knockdown cells in response to the IR treatment. Consistently with our previous result, the number of γ-H2AX-positive foci quickly diminished in control cells but was sustained in LINP1-knockdown cells (**Fig. 4c**,**d**). Together, these observations suggest that DSB-repair activity is impaired by LINP1 knockdown. To test whether NHEJ is the pathway affected by LINP1, we used an NHEJ reporter assay[29]. In MDA-MB-231 cells, in which LINP1 is highly expressed, we observed a decrease in NHEJ activity after LINP1 knockdown (**Fig. 4e**); conversely, in MCF7 cells, in which LINP1 expression is undetectable, NHEJ activity increased dramatically when the cells were transduced with LINP1 but not with control or antisense LINP1 (**Fig. 4e**). These results support the idea that LINP1 enhances DSB repair via the NHEJ pathway.

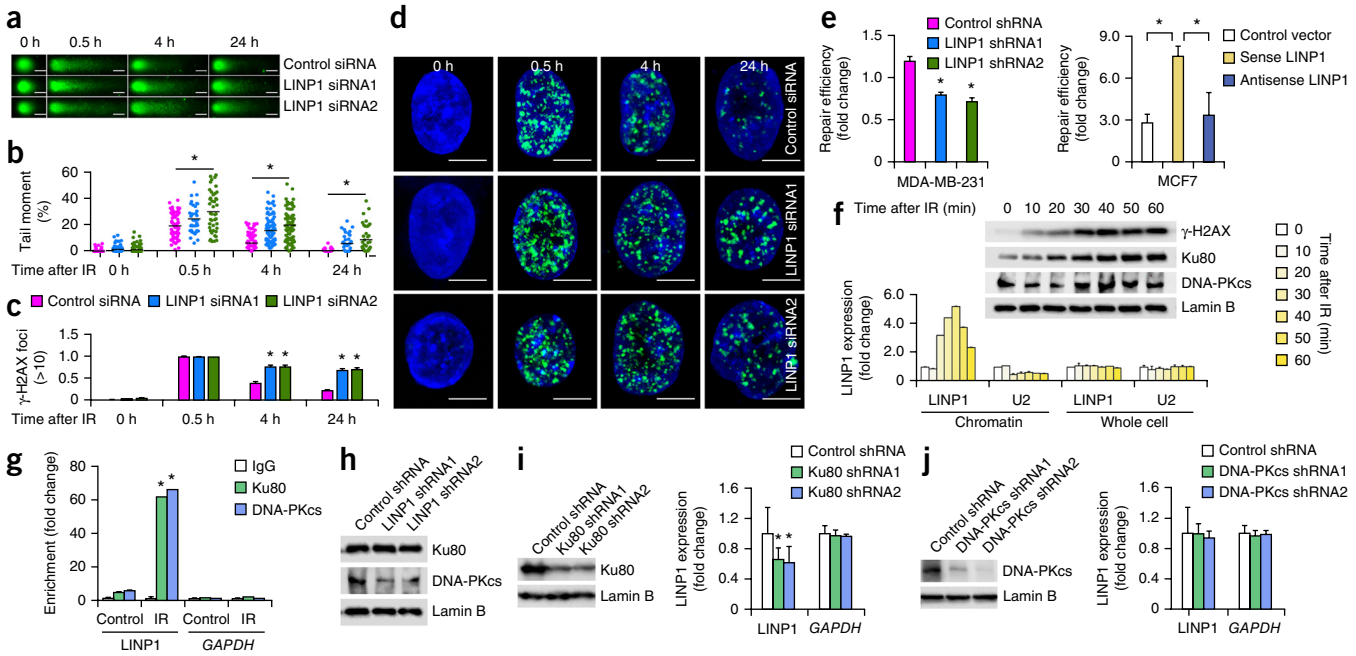

Figure 4 LINP1 serves as a modular scaffold in the NHEJ pathway. (**a**) IR-induced DNA damage in control and LINP1-knockdown MDA-MB-231 cells, as measured by the comet assay. Scale bars, 10 μm. (**b**) Levels of IR-induced DNA damage, quantified by the tail moment in the comet assay. Error bars, s.d. $*P < 0.05$ by two-tailed Student's *t* test; $n = 3$ independent cell cultures. (**c**) Quantification of the number of γ-H2AX-positive foci in the control and LINP1-knockdown cells. Error bars, s.d. $*P < 0.05$ by two-tailed Student's *t* test; $n = 3$ independent cell cultures. (**d**) Representative pictures of γ-H2AX-positive foci in the control and LINP1-knockdown cells. Scale bars, 10 μm. 100 cells were analyzed for each time point. (**e**) NHEJ-mediated DNA-repair activity, measured by NHEJ reporter, in MDA-MB-231 cells expressing control or LINP1-specific shRNAs (left), and in MCF7 cells expressing control vector, sense LINP1 and antisense LINP1 (right). Error bars, s.d. $*P < 0.05$ by two-tailed Student's *t* test; $n = 4$ independent cell cultures. (**f**) Levels of LINP1, Ku80, DNA-PKcs, and γ-H2AX in the chromatin-associated complex at different time points after IR treatment. (**g**) Levels of LINP1 associated with Ku80 and DNA-PKcs in control and IR- treated cells. Error bars, s.d. $*P < 0.05$ by two-tailed Student's *t* test; $n = 3$ technical replicates. (**h**) Levels of chromatin-associated Ku80 and DNA-PKcs in control and LINP1-knockdown cells after IR-treatment. (**i**,**j**) Levels of Ku80, DNA-PKcs, and LINP1 in control and Ku80- or DNA-PKcs-knockdown cells. Left, protein levels of Ku80 (**i**) or DNA-PKcs (**j**) in the control and Ku80-knockdown (**i**) or control and DNA-PKcs-knockdown (**j**) cells. Right, levels of LINP1 and *GAPDH* (as control) in the control and Ku80-knockdown cells (**i**) or in the control and DNA-PKcs-knockdown cells (**j**). Error bars, s.d. $*P < 0.05$ by two-tailed Student's *t* test; $n = 3$ technical replicates. 10 Gy of irradiation was used for all experiments. Uncropped images of gels are shown in **Supplementary Data Set 2**.

We next analyzed the dynamic levels of chromatin-associated LINP1, Ku80, DNA-PKcs, and γ-H2AX in response to IR treatment. We isolated chromatin-associated complexes at 10-min intervals from 0 to 60 min after irradiation. We measured the levels of Ku80, DNA-PKcs, and γ-H2AX in chromatin-associated complexes by western blot analysis and the level of LINP1 by qRT–PCR. In response to IR treatment, the levels of Ku80, DNA-PKcs, γ-H2AX, and LINP1 all increased in the chromatin-associated complex (**Fig. 4f**). Because the level of LINP1 remained unchanged in the whole cell lysates (**Fig. 4f**), this observation suggests that LINP1 was recruited to the chromatin after IR treatment. Furthermore, we found that IR treatment induced the association between LINP1 and Ku80 or DNA-PKcs, as measured by RNA-IP analysis (**Fig. 4g**).

To further define the mechanism responsible, we analyzed the effects of LINP1 knockdown on the levels of chromatin-associated Ku80 and DNA-PKcs after IR treatment. Whereas LINP1 knockdown significantly decreased the level of chromatin-associated DNA-PKcs, it had no effect on Ku80 (**Fig. 4h**). We then knocked down Ku80 or DNA-PKcs to test whether either factor affects the chromatin recruitment of LINP1. Interestingly, we found that Ku80 knockdown resulted in a significant decrease in chromatin-associated LINP1 (**Fig. 4i**), but DNA-PKcs knockdown resulted in no significant changes in LINP1 on chromatin (**Fig. 4j**). Finally, we examined the effect of LINP1 on the interaction of Ku80 and DNA-PKcs by IP followed by western blotting. In LINP1-knockdown cells, there was less association between Ku80 and DNA-PKcs after IR treatment (**Supplementary Fig. 4b**). Collectively, our results suggest that LINP1 may serve as an RNA scaffold that enhances the molecular interaction between Ku80 and DNA-PKcs in the NHEJ pathway.

LINP1 expression is activated by the EGF signaling pathway

The observed correlation between LINP1 and EGFR expression (**Fig. 2g**) is intriguing because overexpressed or amplified *EGFR* has been reported in TNBC[3]. To further confirm this finding, we analyzed LINP1 and *EGFR* RNA expression in the CCLE data set. Expression of LINP1 positively correlated with EGFR expression in cancer cell lines (**Fig. 5a**), and these results were consistent with those from primary specimens (**Fig. 2g**), thus indicating that the EGFR–LINP1 correlation is cell autonomous and has no contribution from tumor stromal RNA. We treated three breast lines with EGF and measured the levels of LINP1 RNA in response to EGF treatment. Whereas EGF significantly induced expression of LINP1 in triple-negative lines MDA-MB-468 and MCF10A, it had no effect in MCF7 (**Fig. 5b**). The downstream effector signaling pathways of EGFR have been well characterized, and multiple small molecules have been developed to inhibit these pathways (**Fig. 5c**). We searched for correlations between the expression of LINP1 and the expression of transcription factors downstream of the EGF pathway in the CCLE data set and found

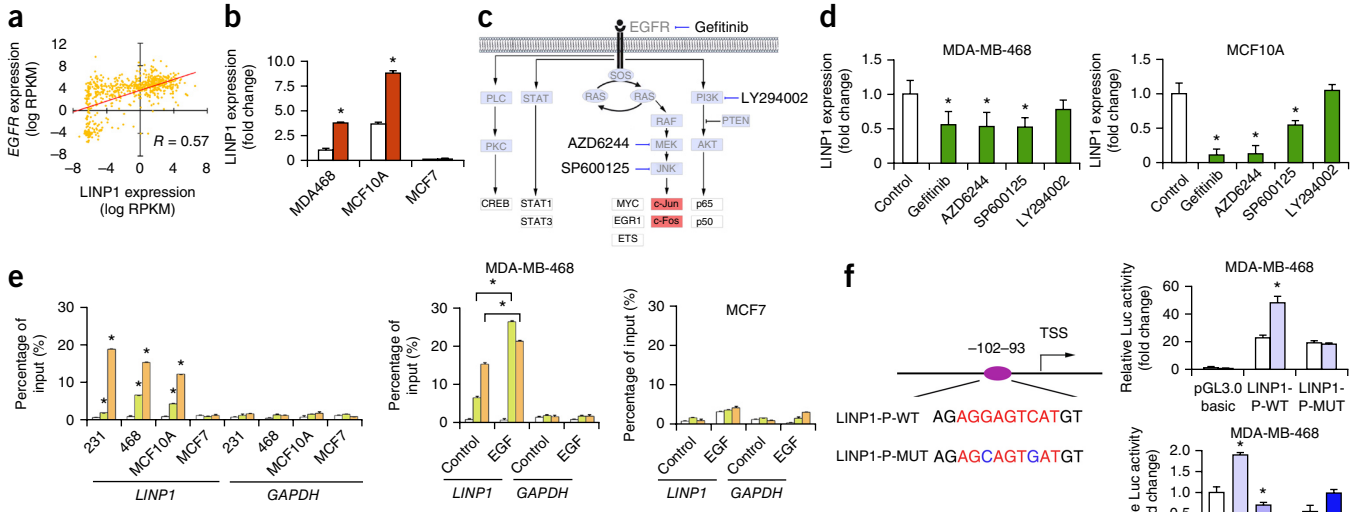

Figure 5 LINP1 is activated by the EGF signaling pathway. (**a**) Correlation between LINP1 and EGFR expression in CCLE. Expression levels of 935 cancer cell lines from CCLE were determined by the RPKM (reads per kilobase per million mapped reads) from RNA-seq. *R* value was calculated by two-sided Pearson's test. (**b**) Expression of LINP1 in MDA-MB-468, MCF10A, and MCF7 cells treated with EGF or control. Error bars, s.d. $*P < 0.05$ by two-tailed Student's *t* test; $n = 3$ technical replicates. (**c**) Schematic diagram of EGFR signaling. Small-molecule inhibitors of EGFR signaling and their specific targets are indicated. (**d**) Expression of LINP1 in MDA-MB-468 and MCF10A cells treated with different small molecules that inhibit different parts of the EGFR pathway. RNA expression was analyzed by qRT–PCR. Error bars, s.d. $*P < 0.05$ by two-tailed Student's *t* test; $n = 3$ technical replicates. (**e**) Left, quantification of the amount of *LINP1* promoter bound to c-Jun or c-Fos in MDA-MB-468, MCF10A, and MCF7 cells. The promoters were pulled down by antibodies to c-Jun or anti-c-Fos and measured by qPCR analysis. *GAPDH* is a negative control. Middle and right, quantification of the amount of *LINP1* promoter bound to c-Jun or c-Fos in MDA-MB-468 (middle) and MCF7 (right) cells, which were treated with EGF. Error bars, s.d. $*P < 0.05$ by two-tailed Student's *t* test; $n = 3$ technical replicates. (**f**) Left, illustration of an AP1-binding site at −102 (red indicates the consensus motif) and the sequences of the AP1 mutations (blue indicates the mutant nucleotides). Middle, luciferase reporter assay of the promoter activities of the *LINP1* core promoter construct and its AP1-mutant counterpart in MDA-MB-468 cells treated with EGF. Right, luciferase (Luc) reporter assay assessing the promoter activities of the *LINP1* core promoter construct in MDA-MB-468 cells with different small molecules that inhibit different parts of the EGFR pathway. Error bars, s.d. $*P < 0.05$ by two-tailed Student's *t* test; $n = 3$ technical replicates.

significant and positive correlations between LINP1 and c-Jun and c-Fos ($P < 0.01$; $R > 0.25$) but not with other transcription factors (**Fig. 5c**), thus suggesting that EGF may regulate LINP1 expression via the RAS–MEK–JNK pathway. We treated MDA-MB-468 and MCF10A cells with inhibitors of the RAS–MEK–JNK pathway as well as the PI3K–AKT pathway as a control. As anticipated, the

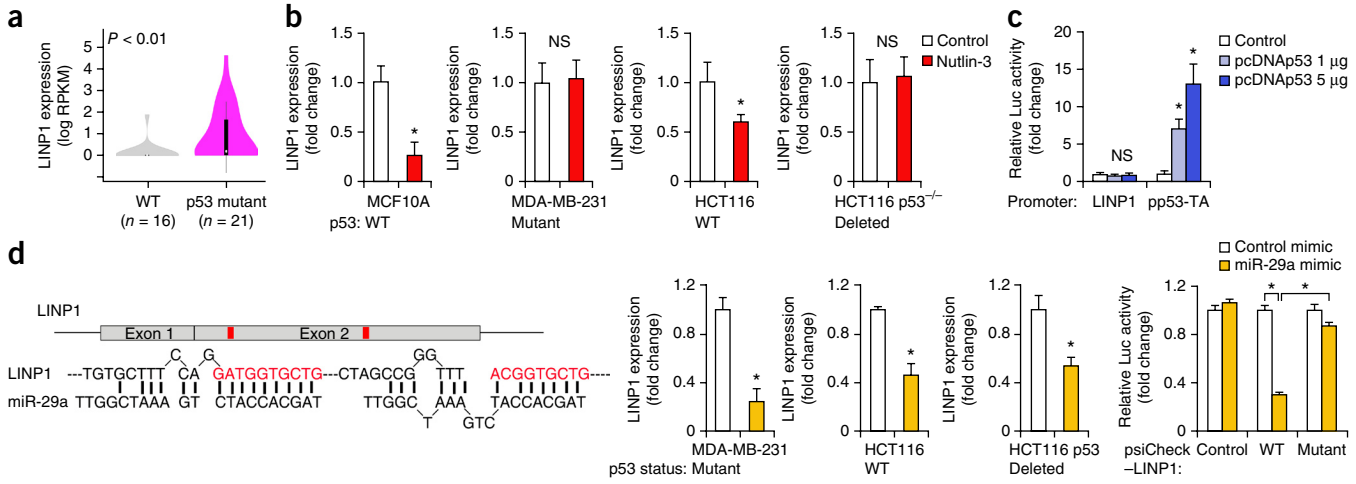

Figure 6 LINP1 is repressed by the p53 signaling pathway. (**a**) LINP1 expression in the CCLE breast cancer cell lines in which the *TP53* mutation status is known. $P < 0.01$ by two-tailed Student's *t* test; *n*, number of breast cancer cell lines used for analysis. (**b**) LINP1 expression in cells of different *TP53* status, in response to nutlin-3a treatment. MCF10A and HCT116, *TP53* WT; MDA-MB-231, *TP53* mutant; HCT116 with *TP53* deletion, *TP53* null. Error bars, s.d. $*P < 0.05$ by two-tailed Student's *t* test; $n = 3$ technical replicates. NS, not significant. (**c**) Luciferase assay measuring the transcription activity of *LINP1* and construct containing a known p53 binding site in cells expressing control vector or WT *TP53*. Error bars, s.d. $*P < 0.05$ by two-tailed Student's *t* test; $n = 3$ independent cell cultures. (**d**) Left, sequence alignment showing the complementarity between LINP1 exon 2 and miR-29. Red, seed sequence of miR-29. Middle, LINP1 expression in cells with different *TP53* status after treatment with vehicle or miR-29. Right, luciferase assay measuring the activity of WT or mutant LINP1-luciferase fusion reporter constructs in response to treatment with miR-29 mimic. Error bars, s.d. $*P < 0.05$ by two-tailed Student's *t* test; $n = 3$ independent cell cultures.

EGFR inhibitor (Gefitinib), the MEK inhibitor (AZD6244), and the JNK inhibitor (SP600125) all significantly reduced the expression of LINP1 in triple-negative lines, whereas the PI3K inhibitor (LY294002) had little effect on the expression of LINP1 (**Fig. 5d**). By analyzing the binding of c-Jun or c-Fos from the chromatin immunoprecipitation–sequencing (ChIP–seq) data from ENCODE, we observed a strong enrichment of c-Jun or c-Fos in MCF10A cells but not in MCF-7 cells (**Supplementary Fig. 5**). We then validated the above ChIP–seq data by ChIP–qPCR and found strong binding of both c-Jun and c-Fos to the promoter region of *LINP1* in MDA-MB-231, MDA-MB-468, and MCF10A cells but not in MCF7 cells (**Fig. 5e**). Importantly, we found that EGF treatment further increased binding between c-Jun or c-Fos and the *LINP1* promoter in MDA-MB-468 cells but had no effect in MCF7 cells (**Fig. 5e**). We found a consensus AP1-binding site at −102 bp from the transcription start site of *LINP1* (**Fig. 5f**). EGF treatment significantly increased the promoter luciferase activity of WT but not that of AP1-mutated *LINP1* (**Fig. 5f**). Consistently with the effects of EGF treatment, the EGFR inhibitor (gefitinib), the MEK inhibitor (AZD6244) and the JNK inhibitor (SP600125) all significantly reduced the activities of the reporter, whereas the PI3K inhibitor (LY294002) had little effect on activity (**Fig. 5f**). Together, these findings demonstrate that EGF signaling, specifically its activation of the RAS–MEK–JNK pathway, is involved in regulating LINP1 expression in TNBC.

LINP1 expression is repressed by the p53 signaling pathway

TNBC shows a high frequency of *TP53* mutations[3]. The differential expression of LINP1 in breast cancer specimens with different *TP53* mutation statuses (**Fig. 2h**) suggests that the p53 pathway might play a role in regulating the expression of LINP1. To delineate the possible link to p53, we looked for a correlation between LINP1 expression and *TP53* mutation status in breast cancer cell lines from CCLE. Consistently with our findings in primary tumors, the breast cancer cells with mutant *TP53* had a significantly higher expression of LINP1 (**Fig. 6a**). Next, we found that nutlin-3a treatment significantly decreased the expression of LINP1 in MCF10A and HCT116 cells, which express WT p53, but not in MDA-MB-231 cells with mutated *TP53* or HCT116 cells with homozygous *TP53* deletion (**Fig. 6b**), thus suggesting that LINP1 expression is negatively regulated by the WT p53 pathway. We used the *LINP1* promoter luciferase reporter

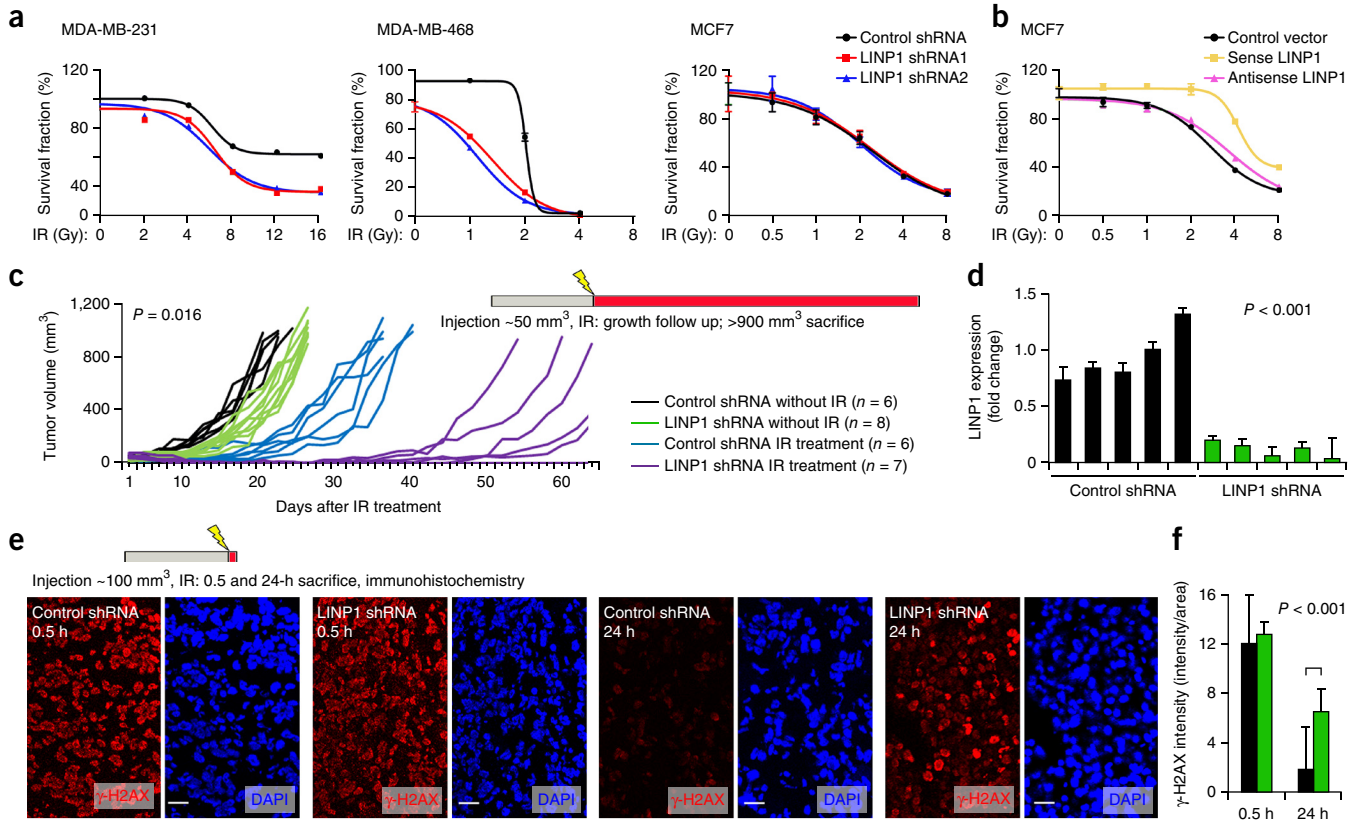

Figure 7 Alteration of LINP1 modulates IR sensitivity. (**a**) Survival of control or LINP1-knockdown MDA-MB-231, MDA-MB-468, and MCF-7 cells in response to IR treatments. Cells expressing control or LINP1 specific shRNAs were treated with different doses of IR, and the survival after treatment was measured with survival assays. Error bars, s.d.; $n = 3$ independent cell cultures. (**b**) Survival of control or LINP1-expressing MCF-7 cells in response to IR treatments. Cells expressing control vector, sense LINP1, or antisense LINP1 were treated with different doses of IR, and the survival after treatment was measured with colony formation assays. Error bars, s.d.; $n = 3$ independent cell cultures. (**c**) Growth of control and LINP1-expressing MDA-MB-231 tumors with or without IR treatment. IR treatment (single dose 8 Gy) was administered when the tumors reached 50 mm^3. Tumor growth was monitored for over 2 months. Animals with tumors larger than 900 mm^3 were euthanized. Number of animals in each group (n) is indicated. (**d**) LINP1 expression in control and LINP1-knockdown tumors, detected by qRT–PCR. Error bars, s.d.; $n = 3$ technical replicates of the qRT-PCR measurements; each histogram bar represents one tumor sample. $^*P < 0.001$ by two-tailed Student's t test between the control shRNA and the LINP1 shRNA groups; $n = 5$ independent tumors. (**e**) γ-H2AX staining in the control and LINP1-expressing MDA-MB-231 tumors at different points after IR treatment. IR treatment was administered when the tumors reached 100 mm^3, and tissues were harvested 0.5 h and 24 h after the IR treatment (single dose 8 Gy). Scale bars, 40 μm. Blue, DAPI nuclear staining; red, γ-H2AX staining. (**f**) Quantification of γ-H2AX staining in control and LINP1-expressing MDA-MB-231 tumors at 0.5 and 24 h after IR treatment. Error bars, s.d. $^*P < 0.001$ by two-tailed Student's t test; $n = 6$ independent tumors.

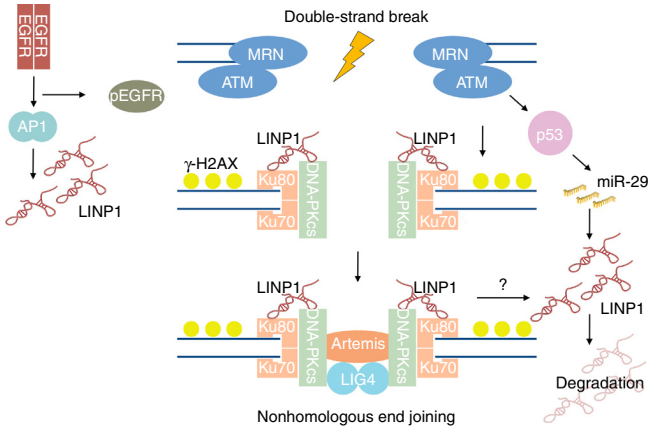

Figure 8 Role of LINP1 in the NHEJ DNA-repair pathway. In response to DSBs, Ku80 and Ku70 associate with the broken ends, thus forming a clamp-like complex that recruits DNA-PKcs to the damage site. Other processing proteins, including Artemis, DNA ligase IV, XRCC4 and XLF then assemble with the Ku80–Ku70–DNA-PK complex, thereby enabling DNA repair. The lncRNA LINP1 enhances NHEJ activity by providing a scaffold for Ku80 and DNA-PKcs. The Ku80 subunit directly interacts with LINP1 and subsequently recruits the lncRNA to broken DNA ends; LINP1 then stabilizes the Ku80–DNA-PKcs complex, thereby increasing NHEJ-mediated DNA-repair activity.

in cells transduced with WT *TP53* cDNA and used a luciferase construct containing a known p53-binding site as the positive control. Whereas the activity of a construct containing a known p53 binding site dramatically increased after *TP53* expression, the luciferase activity of the *LINP1* construct remained unchanged (**Fig. 6c**), thus suggesting that p53 may not directly regulate the transcription of *LINP1*. Consistently with this observation, ChIP–seq data from HCT116 cells also revealed a clear p53 binding signal in the *CDKN1A* (p21) promoter in *TP53* WT cells but not in the *LINP1* promoters in any of the cell lines studied. Together, these results suggest that p53 may regulate LINP1 expression via an indirect pathway.

Next, we observed two regions in LINP1 exon 2 (**Fig. 6d**) that are fully complementary to the seed sequence of miR-29, which is positively and directly regulated by p53 (ref. 30). To determine whether p53 regulates LINP1 expression via miR-29, we treated the cells with nutlin-3a and confirmed an increase in miR-29 expression in *TP53* WT but not mutant or null cells (**Supplementary Fig. 6**). Then we transduced a miR-29 mimic into MCF10A and HCT116 cells, as well as into HCT116 cells bearing a homozygous *TP53* deletion, and measured LINP1 levels. In all three lines, expression of the miR-29 mimic decreased expression of LINP1 (**Fig. 6d**). Finally, we generated reporter constructs in which WT LINP1 or LINP1 with an miR-29 seed-sequence mutation was inserted into the 3′ untranslated region of a luciferase reporter gene. After we cotransfected the reporter construct with miR-29 mimics into cells, we found that the expression of the WT construct was significantly reduced in cells cotransfected with the miR-29 mimic. Importantly, unlike the WT LINP1 control construct, the mutant LINP1 construct retained a high level of expression despite the expression of miR-29 mimic (**Fig. 6d**). Together, these results confirm a role for the p53 pathway in repressing LINP1 expression and indicate that miR-29 is a mediator of p53-regulated LINP1 expression.

Alteration of LINP1 modulates radiation sensitivity

Radiation treatment is currently one of the standard therapies for patients with TNBC[1–3], and previous studies have demonstrated that NHEJ is a key determinant of IR resistance in cancer cells[19–26]. We hypothesized that LINP1 may regulate the IR response by increasing NHEJ activity. First, we assessed the effect of LINP1 knockdown on the IR sensitivities of three cancer cell lines: MDA-MB-231 and MDA-MB-468, which express high levels of LINP1, and MCF7, which has undetectable levels of LINP1 expression. We introduced LINP1-specific short hairpin RNAs (shRNAs) into these cells, which we then treated with different doses of IR. We then assessed cell survival one week after the IR treatments. LINP1 shRNA expression significantly decreased survival after IR treatment in MDA-MB-231 and MDA-MB-468 but not in MCF7 cells (**Fig. 7a**). In contrast, expression of LINP1 RNA in MCF7 cells rendered the cells more resistant to IR (**Fig. 7b**). Next, we established control or LINP1-shRNA-expressing MDA-MB-231-derived xenografts in nude mice. We administered a single dose of 8 Gy IR to each tumor when the tumors reached 50 mm^3 and monitored the mice until the tumors reached 900 mm^3 in size. In the untreated groups, we observed a slight delay of growth in the LINP1-knockdown tumors compared with the controls (**Fig. 7c**). In the IR-treated groups, however, the regrowth of LINP1-knockdown tumors was significantly attenuated. At the time point at which all IR-treated control tumors reached 900 mm^3, the LINP1-knockdown tumors had just started to regrow. In fact, two of the seven LINP1-knockdown tumors became undetectable after IR treatment and never reemerged during the 64-day observation period after IR (**Fig. 7c**). qRT–PCR analysis demonstrated that the tumors expressing LINP1 shRNAs had lower endogenous LINP1 expression than controls (**Fig. 7d**). We next implanted MDA-MB-231-derived tumors as described above and administered a single dose of 8 Gy IR when the tumors reached 100 mm^3. We harvested the tumor tissues either 0.5 h or 24 h after the IR treatment and used γ-H2AX staining to assess the level of DSBs over time. At 0.5 h after IR treatment, a substantial amount of γ-H2AX was present in both control and LINP1-knockdown tumors. At 24 h,the level of γ-H2AX in the control tumors was significantly decreased, but the level in LINP1-knockdown cells remained high (**Fig. 7e**,**f**). These observations indicated that in LINP1-knockdown tumors, compared with control tumors, the ability to repair DSBs was reduced. Collectively, our results suggest that suppression of LINP1 expression impairs DNA-repair activity *in vivo*, thereby sensitizing tumors to IR treatment.

DISCUSSION

Triple-negative breast cancer is a clinically challenging disease involving multistep changes in the genome[1–3]. To date, changes in PCGs in TNBC genomes have been the major focus[3]. By sifting through genomic alterations and distinguishing 'driver' from 'passenger' alterations, a number of key PCG hubs have been uncovered, including gain of *EGFR* signaling or loss of *TP53*. However, despite these pivotal findings, resistance of TNBC to standard therapy, particularly radiation and chemotherapy, has remained poorly understood at a mechanistic level. Because PCGs constitute only 2% of the human genome, it is likely that noncoding RNAs play as-yet-undefined roles in the TNBC 'phenome' of therapeutic resistance. In this regard, we devised a clinically guided genetic screening approach to identify functional lncRNAs in TNBC. Using the lncRNA expression profile as an initial clinical filter, we generated a relatively short list of lncRNA candidates for more extensive testing in siRNA-based functional genetic screening. On the basis of screening for apoptosis induced by the chemotherapy drug doxorubicin, we identified LINP1 as a potential lncRNA candidate that may be involved in cell death and the DNA-damage response in TNBC. Importantly, LINP1 enhances NHEJ activity by providing a scaffold for Ku80 and DNA-PKcs. Once a DSB occurs, the Ku80–Ku70 heterodimer recruits LINP1 to the damaged DNA;

LINP1 then stabilizes the Ku80–DNA-PKcs complex, thereby increasing NHEJ-mediated DNA-repair activity (**Fig. 8**). Because cells lacking LINP1 expression (for example, MCF7) are still competent for NHEJ-mediated repair, LINP1 does not appear to be a prerequisite for the NHEJ process. However, expression of LINP1 in non-LINP1-expressing cells enhances NHEJ repair activity. Interestingly, Ting *et al.* have shown that the Ku80–Ku70 complex also interacts with hTR (TERC), a lncRNA component of telomerase, in human cells[31].

We also uncovered functional links between the noncoding (LINP1 lncRNA) and protein-coding (*EGFR* and *TP53*) genomic hubs. *EGFR* has been reported to be highly amplified in TNBC and to serve as a potential target for treatment[1–3]. Notably, the EGFR pathway is known to enhance NHEJ-mediated DNA repair, and high EGFR activity is associated with radiation resistance[32,33]. Here, we uncovered an additional mechanism for EGFR-induced radiation resistance, whereby EGFR activation results in the upregulation *LINP1* transcription via the activation of the RAS–MEK–ERK pathway and AP1 transcription factors. Thus, in cells with EGFR activation, increased levels of LINP1 stabilize the interaction between Ku80 and DNA-PKcs and enhance NHEJ-mediated DNA-repair activity. Our study also revealed that p53 activation downregulates LINP1 expression via induction of miR-29, which targets LINP1 RNA. Because LINP1's enhancement of NHEJ activity takes place immediately after DNA damage, whereas the mir29-mediated LINP1 downregulation occurs at a much later time point, we speculate that the p53- and mir29-mediated LINP1 regulation may serve as a negative feedback mechanism restricting the level of NHEJ-mediated DNA-repair activity in cells long after damage. High frequencies of *EGFR* amplification and *TP53* mutations in TNBC may increase LINP1 expression at the transcriptional and post-transcriptional levels, respectively. Moreover, copy-number amplification of the *LINP1* gene itself may further enhance the response to increased *EGFR* activity and loss of *TP53* repression in TNBC.

Because of limited therapeutic targets, TNBC is typically treated with surgery and a combination of radiation and chemotherapy, which induce various types of DNA damage[1–3]. The NHEJ pathway, which repairs DSBs in DNA, is one of the major pathways in tumor cells that respond to radiation treatment and chemotherapeutic agents[19–26]. Inhibition of the NHEJ pathway has been proposed to synergize with DNA-damaging therapies for TNBC[1–3]. In addition, the NHEJ pathway may also be a key source of genomic rearrangement and instability[19–26], which are fundamental features of TNBC[1–3]. We believe that a better understanding of the role of lncRNA in the NHEJ pathway will not only provide a deeper understanding TNBC development but also will help to refine the classification and treatment of this disease.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank the TCGA and CCLE project teams. We thank Dr. B. Vogelstein (Johns Hopkins University) for providing HCT116 WT and HCT116 TP53 cell lines. This work was supported, in whole or in part, by the Basser Center for BRCA (L.Z.), the US National Institutes of Health (R01CA142776 to L.Z., R01CA190415 to L.Z., P50CA083638 to L.Z., P50CA174523 to L.Z., R01CA148759 to Q.H., and R01NS094533 to Y.F.), the Breast Cancer Alliance (L.Z. and C.V.D.), the Ovarian Cancer Research Fund (X.H.), the Foundation for Women's Cancer (X.H.), and the Marsha Rivkin Center for Ovarian Cancer Research (L.Z.). S.W. was supported by the China Scholarship Council.

AUTHOR CONTRIBUTIONS

Y.Z., C.V.D. and L.Z. conceptualized and designed the experiments. Y.Z., Q. He, L.F., W.S., and C.L. performed the molecular and cellular biology experiments. Z.H., Y. Fan, Z.T. and J.Y. performed bioinformatics analysis. Y.Z., Z.H., J.Y., X.H., J.L.T., Y.F., Q. Huang, and K.M. analyzed and interpreted data. Y.Z., Y. Feng, C.V.D., and L.Z. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Foulkes, W.D., Smith, I.E. & Reis-Filho, J.S. Triple-negative breast cancer. *N. Engl. J. Med.* **363**, 1938–1948 (2010).
2. Carey, L., Winer, E., Viale, G., Cameron, D. & Gianni, L. Triple-negative breast cancer: disease entity or title of convenience? *Nat. Rev. Clin. Oncol.* **7**, 683–692 (2010).
3. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
4. Derrien, T. *et al.* The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* **22**, 1775–1789 (2012).
5. Mercer, T.R., Dinger, M.E. & Mattick, J.S. Long non-coding RNAs: insights into functions. *Nat. Rev. Genet.* **10**, 155–159 (2009).
6. Rinn, J.L. & Chang, H.Y. Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* **81**, 145–166 (2012).
7. Guttman, M. & Rinn, J.L. Modular regulatory principles of large non-coding RNAs. *Nature* **482**, 339–346 (2012).
8. Batista, P.J. & Chang, H.Y. Long noncoding RNAs: cellular address codes in development and disease. *Cell* **152**, 1298–1307 (2013).
9. Ørom, U.A. & Shiekhattar, R. Long noncoding RNAs usher in a new era in the biology of enhancers. *Cell* **154**, 1190–1193 (2013).
10. Ulitsky, I. & Bartel, D.P. lincRNAs: genomics, evolution, and mechanisms. *Cell* **154**, 26–46 (2013).
11. Tay, Y., Rinn, J. & Pandolfi, P.P. The multilayered complexity of ceRNA crosstalk and competition. *Nature* **505**, 344–352 (2014).
12. Sahu, A., Singhal, U. & Chinnaiyan, A.M. Long noncoding RNAs in cancer: from function to translation. *Trends Cancer* **1**, 93–109 (2015).
13. Ling, H. *et al.* Junk DNA and the long non-coding RNA twist in cancer genetics. *Oncogene* **34**, 5003–5011 (2015).
14. Iyer, M.K. *et al.* The landscape of long noncoding RNAs in the human transcriptome. *Nat. Genet.* **47**, 199–208 (2015).
15. Yan, X. *et al.* Comprehensive genomic characterization of long non-coding RNAs across human cancers. *Cancer Cell* **28**, 529–540 (2015).
16. Gupta, R.A. *et al.* Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**, 1071–1076 (2010).
17. Tsai, M.C. *et al.* Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **329**, 689–693 (2010).
18. Lin, A. *et al.* The LINK-A lncRNA activates normoxic HIF1α signalling in triple-negative breast cancer. *Nat. Cell Biol.* **18**, 213–224 (2016).
19. Helleday, T., Petermann, E., Lundin, C., Hodgson, B. & Sharma, R.A. DNA repair pathways as targets for cancer therapy. *Nat. Rev. Cancer* **8**, 193–204 (2008).
20. Jackson, S.P. & Bartek, J. The DNA-damage response in human biology and disease. *Nature* **461**, 1071–1078 (2009).
21. Lieber, M.R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* **79**, 181–211 (2010).
22. Lord, C.J. & Ashworth, A. The DNA damage response and cancer therapy. *Nature* **481**, 287–294 (2012).
23. Chiruvella, K.K., Liang, Z. & Wilson, T.E. Repair of double-strand breaks by end joining. *Cold Spring Harb. Perspect. Biol.* **5**, a012757 (2013).
24. Deriano, L. & Roth, D.B. Modernizing the nonhomologous end-joining repertoire: alternative and classical NHEJ share the stage. *Annu. Rev. Genet.* **47**, 433–455 (2013).
25. Mehta, A. & Haber, J.E. Sources of DNA double-strand breaks and models of recombinational DNA repair. *Cold Spring Harb. Perspect. Biol.* **6**, a016428 (2014).
26. Goldstein, M. & Kastan, M.B. The DNA damage response: implications for tumor responses to radiation and chemotherapy. *Annu. Rev. Med.* **66**, 129–143 (2015).
27. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).
28. Simon, M.D. *et al.* The genomic binding sites of a noncoding RNA. *Proc. Natl. Acad. Sci. USA* **108**, 20497–20502 (2011).
29. Seluanov, A., Mittelman, D., Pereira-Smith, O.M., Wilson, J.H. & Gorbunova, V. DNA end joining becomes less efficient and more error-prone during cellular senescence. *Proc. Natl. Acad. Sci. USA* **101**, 7624–7629 (2004).
30. Ugalde, A.P. *et al.* Aging and chronic DNA damage response activate a regulatory pathway involving miR-29 and p53. *EMBO J.* **30**, 2219–2232 (2011).
31. Ting, N.S., Yu, Y., Pohorelic, B., Lees-Miller, S.P. & Beattie, T.L. Human Ku70/80 interacts directly with hTR, the RNA component of human telomerase. *Nucleic Acids Res.* **33**, 2090–2098 (2005).
32. Nyati, M.K., Morgan, M.A., Feng, F.Y. & Lawrence, T.S. Integration of EGFR inhibitors with radiochemotherapy. *Nat. Rev. Cancer* **6**, 876–885 (2006).
33. Chen, D.J. & Nirodi, C.S. The epidermal growth factor receptor: a role in repair of radiation-induced DNA damage. *Clin. Cancer Res.* **13**, 6555–6560 (2007).

ONLINE METHODS

RNA-seq data processing. The poly(A)$^+$ RNA-seq (Illumine) data, in BAM format, for human breast tumor specimens were generated and processed by the University of North Carolina (UNC) as part of the TCGA project. The poly(A)$^+$ RNA-seq (Illumina HiSeq) data, in BAM format, for 935 human cancer cell lines across 21 cancer types was generated and processed by the Cancer Cell Line Encyclopedia (CCLE) project, a collaboration between the Broad Institute, the Novartis Institutes for Biomedical Research, and the Genomics Institute of the Novartis Research Foundation. RNA-seq files were downloaded from the Cancer Genomics Hub (http://cghub.ucsc.edu/). We imported the aligned reads of each BAM file to the Partek Genomic Suite (http://www.partek.com/) to obtain the expression levels for genes by summarizing the reads per kilobase per million mapped reads (RPKM) values. GENCODE annotations (version 18; http://www.gencodegenes.org/releases/18.html/) were used to define lncRNAs and PCGs. The log-transformed RPKM values of genes were further analyzed with Partek Genomic Suite and BRB-ArrayTools (http://linus.nci.nih.gov/BRB-ArrayTools.html).

SNP array data processing and copy-number analysis. The TCGA SNP array (Affymetrix Genome-Wide Human SNP Array 6.0) data in CEL format of patients' paired breast tumor and germline-derived DNA specimens was downloaded from the TCGA Data Portal (https://tcga-data.nci.nih.gov/tcga/). The CEL files were imported into the Partek Genomic Suite for subsequent segmentation and calculation of the predicted copy number for each given gene. A predicted copy number larger than 2.3 or smaller than 1.7 was considered to be a copy-number gain or loss for each gene, respectively. Amplified segments and GISTIC scores were visualized with IGV (http://www.broadinstitute.org/igv/).

Cell culture. Cancer cell lines were purchased from the ATCC without further authentication. HCT116 WT and HCT116 *TP53* cell lines were from B. Vogelstein (Johns Hopkins University) without further authentication. MDA-MB-231, MDA-MB-468, MCF7, HCT116 WT and HCT116 *TP53* were cultured in RPMI1640 medium (Invitrogen) supplemented with 10% FBS (FBS, Invitrogen). MCF10A was cultured in DMEM/F12 medium (Invitrogen) containing 5% horse serum, 20 ng/ml EGF, 0.5 mg/ml hydrocortisone, 100 ng/ml cholera toxin, and 10 μg/ml insulin. Cells were routinely tested for mycoplasma contamination with a Mycoplasma Plus PCR Primer Set (Agilent) and were found to be negative.

siRNA screening in the MDA-MB-231 cell line. A total of 20 lncRNAs were included in our initial screening in MDA-MB-231 cells (**Supplementary Table 2**). To reduce the off-target effect of the siRNAs, we designed two independent siRNA sequences targeting each lncRNA gene candidate. qRT–PCR was used to monitor the siRNAs' knockdown efficiency. We found that 11 of 20 (55%) lncRNAs were efficiently knocked down by siRNAs. LINP1 and ENSG00000227036 were initially identified from the screening. Nontargeting siRNA controls (which do not target any human or mouse genes) were used as negative controls in the screening. LINP1 was the only positive lncRNA identified from the initial screening. Cells were seeded in six-well plates in antibiotic-free medium overnight, and transfections were performed with Lipofectamine RNAiMAX transfection reagent (Invitrogen). 24 h after transfection with the indicated siRNAs, cells were trypsinized and plated in 96-well plates in triplicate. 24 h later, 1 μM doxorubicin was added. After cells were allowed to grow for another 24 h, caspase3 activity was assessed with a Caspase-Glo3 Assay Kit (Promega). MTT assays were conducted in parallel for normalization with the Cell Proliferation Kit (I) (Roche) according to the manufacturer's instructions. Caspase3 activity was measured with a Fluoroskan Ascent FL (Thermo), and MTT was quantified with an ELx800 Absorbance Microplate Reader (BioTek) at 570 nm.

Plasmid construction. For pulldown assays, full length (1–917 bp), 5′ (1–300 bp), 5′ (1–600 bp), 5′ (300–600 bp), 5′ (600–917 bp) and full-length antisense LINP1 were cloned between the 5′ BamHI and 3′ XhoI sites of the pBluescript II SK (+) vector. For LINP1 overexpression, full-length sense and antisense LINP1 were cloned into the CD513B vector (System Biosciences). For the TP53-activation luciferase reporter, the *LINP1* promoter (−4000 to +300 bp) was cloned into pGL3-basic (Promega); −200 to +300 bp was cloned into pGL3-basic for the studies of EGF activation, drug inhibition and AP1 transactivation. To generate the AP1-binding-site-mutation reporter vector, two point mutations were introduced into the putative AP1-binding site of the LINP1-promoter-WT (−200 to +300 bp) vector by QuikChange Lightning Site-Directed mutagenesis kit (Agilent Technologies). The following primers were used for the LINP1-P-MUT vector: LINP1-P-MUT F, GAGCCAATGGGTAACATCACTGCTCTGT TCTTAGCCTCT; LINP1-P-MUT R, AGAGGCTAAGA ACAGAGCAGTGAT GTTACCCATTGGCTC. For miRNA luciferase reporter assays of LINP1, double-stranded DNA sequence from LINP1 containing two WT (psiCheck-LINP1 WT: TCGAGTGTGCTT TCCAGGATGGTGCTGAGATCTTAGCCGGGTTTTACGG TGCTGGC) or two mutant target sequences (psiCheck-LINP1 mutant: TCGAGTG TGCTTTCCAGGATATGAAGTAGATCTTAGCCG GG TTTTACATGAAGTGC) of miR-29 were synthesized by IDT and cloned into the psiCheck 2 vector. The underlining indicates the miR-29 seed sequences in LINP1.

shRNA lentiviral transduction. The Lentiviral vector (pLKO.1) and packaging vectors were transfected into 293T cells. The medium was changed 8 h after transfection, and the medium containing lentivirus was collected 48 h later. Cancer cells were infected with lentivirus in the presence of 8 mg/ml polybrene. The shRNA oligonucleotide sequences are shown in **Supplementary Data Set 3**.

RNA isolation and qRT–PCR. Total RNA was extracted with TRIzol Reagent (Invitrogen) and reverse transcribed with a High Capacity RNA-to-cDNA Kit (Applied Biosystems).cDNA was quantified with an ABI ViiA 7 System (Applied Biosystems).

Protein isolation and western blotting. Western blotting was performed with the following primary antibodies: anti-Ku80 (cat no. MA5-12933, clone no. 111, Thermo); anti-Ku70 (cat no. MA5-13110, clone no. N3H10, Thermo); anti-DNA-PKcs (cat no. MA5-13404, Thermo); anti-phospho-H2AX(S139) (cat no. 05-636, clone no. JBW301, Millipore); anti-PARP (cat no. 9542, CST); anti-β-tubulin (cat no. 2128, clone no. 9F3, CST); anti-Lamin B (cat no. ab8982, clone no. 119D5-F1, Abcam) were used; this was followed by incubation with secondary antibodies conjugated with horseradish peroxidase (anti-rabbit IgG HRP-linked antibody (cat no. 7074S, CST) and anti-mouse IgG HRP-linked antibody (cat no. NA931V, GE Healthcare Life Sciences)). Immunoreactive proteins were visualized with the LumiGLO chemiluminescent substrate (Cell Signaling). The antibody and validation information is provided in **Supplementary Data Set 4**.

Northern blot. 561-bp LINP1 cDNA fragment (266–826 bp) was cloned into pBluescript II SK(+). DIG-labeled RNA probe was transcribed *in vitro* and purified. 20 μg of the total RNA was fractionated on a 2% agarose gel containing 1× Denaturing Gel Buffer (Invitrogen). After visualization of 28S and 18S rRNAs by SYBR Gold staining to verify the integrity of RNA samples and equal loading, the RNA was blotted onto a nylon membrane (Whatman). After UV cross-linking, membranes were placed into a hybridization bag containing prewarmed DIG Easy Hyb buffer (Roche) and incubated for 30 min at 68 °C; this was followed by incubation with DIG-labeled LINP1 RNA probe (final concentration 50 ng/ml) for 14 h at 68 °C. The membranes were washed in 2× SSC and 0.1% SDS for 10 min twice at room temperature and in 0.1× SSC and 0.1% SDS for 10 min twice at 68 °C, and detection was performed with ready-to-use CDP-Star buffer (Roche).

RNA pulldown assay. The cDNA sequence of LINP1 was cloned into pBluescript II SK (+). Biotin-labeled RNAs were transcribed *in vitro* and purified. 3 μg of biotinylated RNA was mixed with precleared human MDA-MB-231 whole cell lysate (containing 1 mg proteins) in 500 μl RIP buffer and then mixed with 50 μl washed streptavidin agarose beads at RT for 1 h. Beads were washed briefly with RIP buffer five times and boiled in SDS buffer. Then the retrieved proteins were detected by western blotting or by MS identification.

RNA immunoprecipitation (RNA-IP). For native RNA-IP, MDA-MB-231 extract was incubated with 10 μg of anti-Ku70 (cat no. MA5-13110, clone no. N3H10, Thermo), anti-Ku80 (cat no. MA5-12933, clone no. 111, Thermo), anti-DNA-PKcs (cat no. MA5-13404, Thermo) antibody or control IgG (cat no. 5415S, CST) and then with Protein A–Sepharose beads. After a total of three washes in RNA-IP buffer, beads were boiled in SDS buffer for western blotting or were resuspended in TRIzol reagent for real-time RT-PCR. UV-cross-linking RNA-IP (CLIP) was performed as previously described[34–36]. Briefly, UV-irradiated

doi:10.1038/nsmb.3211

MDA-MB-231 cells were lysed in RSB-Triton buffer, incubated with anti-Ku70 (cat no. MA5-13110, clone no. N3H10, Thermo), anti-Ku80 (cat no. MA5-12933, clone no. 111, Thermo), anti-DNA-PKcs (cat no. MA5-13404, Thermo) antibody or control IgG (cat no. 5415S, CST) and then precipitated with Protein A–Sepharose beads. Beads were then extracted for western blotting or real-time RT-PCR.

Capture hybridization analysis of RNA targets (CHART). Experiments were performed as previously described[28]. Briefly, cells were washed with PBS, and nuclei were enriched by disrupting cells with a Dounce homogenizer in sucrose buffer, diluted with an equal volume of glycerol buffer, and layered on top of glycerol buffer (4 mL). The cross-linked nuclei were collected by centrifugation and further cross-linked in 3% formaldehyde diluted in PBST for 30 min. Cross-linked nuclei were washed in PBST and resuspended in sonication buffer and then sheared with a Misonix sonicator 3000. CHART nuclear extracts were diluted 1:4 in NRB buffer. RNase H–mapping reactions were performed and analyzed as previously described[28]. RNase H–mapping oligonucleotides and sequences of qPCR primers are listed in **Supplementary Data Set 3**. Capture oligonucleotides were synthesized (Integrated DNA Technologies) to incorporate an internal hexa-ethyleneglycol spacer (iSp18) and a 3′ biotin label with an extended spacer arm (3Bio-TEG); sequences are listed in **Supplementary Data Set 3**.

For each CHART reaction, 100 pmol of capture oligonucleotides was added to the extract from 10^7 cells and hybridized overnight at room temperature with gentle shaking. Hybridized material was captured with 60 µl streptavidin resin (Invitrogen) 8 h at room temperature. Bound material was washed five times with WB250 buffer. Streptavidin resin was boiled in SDS buffer for western blotting or was resuspended in TRIzol reagent for real-time RT-PCR.

Chromatin fractionation. MDA-MB-231 cells were fractionated as previously described[37] with modification. Briefly, cells were resuspended in cytoplasmic extract (CE) buffer and incubated on ice for 5 min. Cell lysates were centrifuged at 300*g* for 2 min, and the supernatant (cytoplasm fraction) was removed. The remaining pellet (enriched with nuclei) was washed with CE buffer once and then lysed in buffer B on ice for 5 min. The nuclei lysate was then centrifuged at 1700*g* for 4 min, and the supernatant (soluble nuclear fraction) was removed. The final pellet is the chromatin fraction.

Comet assays. MDA-MB-231 cells were transfected with anti-LINP1 siRNAs or control siRNA 48 h before irradiation. Cells were treated with 10 Gy of IR, and harvested at 0 h (before radiation), 0.5 h, 4 h, or 24 h after IR. Neutral comet assays with SYBR Gold staining (Invitrogen) were performed. The quantification of tail DNA was performed with CASP software.

Immunofluorescence. siRNA-treated cells were seeded on coverslips, treated with 10 Gy of IR the next day, and then harvested at 0 h (before radiation), 0.5 h, 4 h, or 24 h for immunofluorescence. Cells were fixed in solution containing 3% paraformaldehyde, 2% sucrose for 10 min at room temperature. Cells were subsequently permeabilized with 0.5% Triton solution for 5 min at 4 °C and then incubated with anti-γH2AX antibody (cat no. ab81299, Abcam; cat no. 05-636, clone no. JBW301, Millipore) at a dilution of 1:1,000 in PBST buffer (PBS plus 0.1% Tween-20, and 0.02% NaN_3) overnight at 4 °C. Cells were then washed three times with PBST and then incubated with secondary antibody for 1 h at room temperature. After being washed four times with PBST, coverslips were mounted onto glass slides with Vectashield mounting medium containing DAPI (Vector Laboratories) and visualized with an Axiovert 200M inverted microscope (Zeiss).

Coimmunoprecipitation (Co-IP). MDA-MB-231 cells with stable expression of LINP1 shRNA1, LINP1 shRNA2 or control shRNA were treated with 10 Gy of IR. Cells were recovered in a 37 °C incubator for 0.5 h after IR, then lysed in Co-IP buffer (50 mM Tris HCl, pH 7.5, 150 mM NaCl, 2 mM EDTA, 10% glycerol, 0.5% NP40, and 1× PIC) with disruptive sonication. After preclearing, 10 µg of anti-Ku80 (cat no. MA5-12933, clone no. 111, Thermo), anti-DNA-PKcs (cat no. MA5-13404, Thermo) antibody or control IgG (cat no. 5415S, CST) was added to 5 mg supernatant and incubated overnight at 4 °C with gentle rotation. 50 µg supernatant from each samples was saved as input for the following western blot. Protein A–Sepharose beads were added to each sample and incubated at 4 °C for 1 h. After three washes, proteins were extracted for western blotting.

Nonhomologous end-joining assay and FACS analysis. The experimental strategy for the NHEJ assay was as previously described[29]. LINP1-shRNA-treated MDA-MB-231 or LINP1-overexpressing MCF7 cells were transfected with HindIII-digested plasmid along with 0.1 µg of control pDsRed2-N1. Expression of GFP and DsRed was monitored by fluorescence microscopy (Nikon, Eclipse TE2000-U). 48–72 h after transfection, cells were harvested, resuspended in 0.5 ml of PBS, pH 7.4 (Gibco, Invitrogen), and analyzed by FACS (BD FACS Canto).

TaqMan miRNA assays. Briefly, single-stranded cDNA was synthesized from 5 ng of total RNA in a 15-µl reaction volume with a TaqMan miRNA reverse-transcription kit (Applied Biosystems). Quantitative PCR was performed with TaqMan microRNA assay mix.

miRNA-mimic transfection. The hsa-miR-29a mimic and control mimic were purchased from Sigma. For transient transfections, cells were plated 24 h before transfection at 50% confluence. miRNA-mimic transfections were performed with Lipofectamine RNAiMAX (Invitrogen).

Luciferase assays. For TP53 activation assays of the *LINP1* promoter, 500 ng pGL3.0-basic or pGL3.0-LINP1 vector plus 5 ng of the *Renilla* luciferase plasmid with or without 1 µg or 5 µg pcDNA3.1-TP53 vector was transfected to HEK293 cells with FuGENE 6 (Roche). For measurement of promoter activity in cells treated with EGF or inhibitors, MDA-MB-468 cells were first transfected with LINP1-P-WT or LINP1-P-MUT for 24 h before addition of EGF or inhibitors for another 24 h incubation. For miRNA luciferase reporter assays, HEK293 cells were plated on a 24-well plate 24 h before transfection at 50% confluence. 30 nM miR-29a mimics or control mimics (Sigma) was transfected with Lipofectamine RNAiMAX (Invitrogen). 24 h after transfection, 0.125 µg of psiCheck-LINP1 WT or psiCheck-LINP1 MUT reporter vector was transfected with FuGENE6 transfection reagent (Roche). 48 h after reporter vector transfection, cells were harvested, and reporter assays were performed with a dual luciferase reporter assay system (Promega) with a Fluoroskan Ascent FL fluorometer (Thermo Fisher Scientific).

Chromatin immunoprecipitation (ChIP). ChIP was performed as previously described[38] with the following modifications. 3×10^7 of MDA-MB-231, MDA-MB-468, MCF10A, or MCF7 cells treated with or without 200 ng/ml of EGF were harvested for ChIP experiments. Cells were cross-linked with 1% formaldehyde at room temperature for 10 min and then neutralized with 125 mM glycine for 5 min. Cells were rinsed with ice-cold PBS twice and scraped into 1 ml of ice-cold PBS. Cells were resuspended in 0.3 ml of lysis buffer and sonicated. After centrifugation, supernatants were collected and diluted in IP dilution buffer, and this was followed by immunoclearing with protein A–Sepharose for 2 h at 4 °C. 5 µg anti-c-Jun (cat no. 9165S, clone no. 60A8, CST) or anti-c-Fos (cat no, 2250S, CST) antibody (Cell Signaling Technology) or control IgG (cat no. 2729S, CST) was used for immunoprecipitation. After immunoprecipitation, 45 µl protein A–Sepharose was added and incubated for another 1 h. Precipitates were washed, and DNA was purified after removal of cross-links for real-time PCR. Primers are listed in **Supplementary Data Set 3**.

***In vivo* tumor experiments.** MDA-MB-231 cells were transduced with lentiviruses expressing anti-LINP1 or control shRNA, and selected in puromycin for 7 d. Three million tumor cells were injected subcutaneously into 6-week-old athymic female nu/nu mice (stock no 002019, Jackson Labs). An 8-Gy single dose was precisely delivered to the tumors of anaesthetized mice with a small animal radiation research platform (SARRP) after tumors had grown to approximately 50 mm^3. Tumor growth was monitored every other day with a digital vernier caliper, and tumor volumes were calculated according to the formula: tumor volume (mm^3) = $(1/6) \times \pi \times$ (tumor length) × (tumor width)^2. For xenograft immunofluorescence, an 8-Gy single dose was delivered after tumors had grown to 100 mm^3. 0.5 h or 24 h after irradiation, mice were sacrificed, and tumors were harvested for immunofluorescence of γ-H2AX. The intensity of fluorescence was quantified with ImageJ. Statistical significance of the differences was evaluated with two-tailed Student's *t* tests. For all the xenograft studies, the sample size of each group is indicated in the figures. We performed pilot experiments with a few mice per group and then performed larger studies if needed to reach statistical

significance; we repeated experiments to ensure reproducibility. Owing to the nature of the performed experiments, no randomization and no blinding were used because they were deemed unfeasible. However, the resulting tumors were analyzed in a blinded manner. We treated a *P* value of less than 0.05 as a significant difference. All experiments were performed at least twice. All animal procedures were in accordance with protocols approved by the Institutional Animal Care and Use Committee of the University of Pennsylvania.

Statistical analysis. Statistical analysis was performed with SPSS and SAS software. All results were expressed as mean ± s.d., and $P < 0.05$ indicated significance.

34. Ule, J., Jensen, K., Mele, A. & Darnell, R.B. CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods* **37**, 376–386 (2005).
35. Zhao, J. *et al.* Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Mol. Cell* **40**, 939–953 (2010).
36. Zhang, C. & Darnell, R.B. Mapping *in vivo* protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. *Nat. Biotechnol.* **29**, 607–614 (2011).
37. Méndez, J. & Stillman, B. Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. *Mol. Cell. Biol.* **20**, 8602–8612 (2000).
38. Shang, Y., Hu, X., DiRenzo, J., Lazar, M.A. & Brown, M. Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. *Cell* **103**, 843–852 (2000).

doi:10.1038/nsmb.3211