nature communications

Article

https://doi.org/10.1038/s41467-025-59886-w

p53-inducible lncRNA *LOC644656* causes genotoxic stress-induced stem cell maldifferentiation and cancer chemoresistance

Received: 15 April 2024

Accepted: 5 May 2025

Published online: 23 May 2025

Check for updates

Ai Tamura[1,10], Kazuyuki Yamagata[1,10], Takashi Kono[1,2], Masanori Fujimoto[1], Takahiro Fuchigami[1], Motoi Nishimura[3], Masataka Yokoyama[1], Akitoshi Nakayama[1], Naoko Hashimoto[1,2], Ikki Sakuma[1], Nobuyuki Mitsukawa[4], Yusuke Kawashima[5], Osamu Ohara[5], Shinichiro Motohashi[6], Eiryo Kawakami[7], Takashi Miki[2,8], Atsushi Onodera[2,9] & Tomoaki Tanaka[1,2]

Genotoxic stress-induced stem cell maldifferentiation (GSMD) integrates DNA damage responses with loss of stemness and lineage-specific differentiation to prevent damaged stem cell propagation. However, molecular mechanisms governing GSMD remain unclear. Here, we identify the p53-induced long non-coding RNA *LOC644656* as a key regulator of GSMD in human embryonic stem cells. *LOC644656* accumulates in the nucleus upon DNA damage, disrupting pluripotency by interacting directly with POU5F1 and KDM1A/LSD1-NuRD complexes, repressing stemness genes, and activating TGF-β signaling. Additionally, *LOC644656* mitigates DNA damage by binding DNA-PKcs and modulating the DNA damage response. In cancer, elevated *LOC644656* correlates with poor patient survival and enhanced chemoresistance. Our findings demonstrate that *LOC644656* mediates stemness suppression and resistance to genotoxic stress by coordinating DNA damage signaling and differentiation pathways. Thus, *LOC644656* represents a potential therapeutic target for overcoming chemoresistance and advancing stem cell biology.

Stem cells possess a unique ability to self-renew and differentiate into various cell types, a process tightly regulated to maintain tissue homeostasis and repair. DNA damage within stem cell populations can disrupt this balance, often triggering differentiation as a cellular response mechanism. For instance, in hematopoietic stem cells (HSCs), DNA damage activates differentiation checkpoints, thereby limiting their self-renewal capacity[1,2]. Similarly, DNA damage in mammalian neural stem cells promotes differentiation into astrocytes[3]. This phenomenon also holds significance in cancer biology, as DNA damage-induced differentiation of leukemic cells may undermine

[1]Department of Molecular Diagnosis, Graduate School of Medicine, Chiba University, Chiba 260-8670, Japan. [2]Research Institute of Disaster Medicine, Chiba University, Chiba 260-8670, Japan. [3] Division of Laboratory Medicine, Clinical Genetics and Proteomics, Chiba University Hospital, 260-8677 Chiba, Japan. [4]Department of Plastic Surgery, Graduate School of Medicine, Chiba University, Chiba 260-8670, Japan. [5]Department of Applied Genomics, Kazusa DNA Research Institute, Kisarazu, Chiba 292-0818, Japan. [6]Department of Medical Immunology, Graduate School of Medicine, Chiba University, Chiba, Japan. [7]Department of Artificial Intelligence Medicine, Graduate School of Medicine, Chiba University, Chiba 260-8670, Japan. [8]Department of Medical Physiology, Chiba University, Graduate School of Medicine, Chiba 260-8670, Japan. [9]Institute for Advanced Academic Research, Chiba University, Chiba 263-8522, Japan. [10]These authors contributed equally: Ai Tamura, Kazuyuki Yamagata. e-mail: tomoaki@restaff.chiba-u.jp

therapeutic interventions[4]. Beyond cancer therapy, environmental stressors such as aging and metabolic stress can drive stem cells toward differentiation, frequently favoring fibroblast lineages under certain conditions. Although genotoxic stress-induced apoptosis in human embryonic stem cells (hESCs) is well-characterized, molecular mechanisms governing genotoxic stress-induced differentiation remain largely unexplored.

The tumor suppressor p53, recognized as the "guardian of the genome", plays critical roles in preventing malignant transformation and regulating stem cell function and differentiation[5–9]. Upon genotoxic stress, p53 is stabilized and activated, orchestrating cellular responses including cell cycle arrest, DNA repair, and apoptosis, mainly through transcriptional regulation of numerous protein-coding genes[10–12]. Noncoding RNAs, particularly long noncoding RNAs (lncRNAs), have also emerged as essential components of the p53-dependent DNA damage response[13–17]. Several p53-induced lncRNAs regulate cell cycle progression and genomic stability[18,19]; however, their roles in influencing cell plasticity and pluripotency under genotoxic stress remain unclear.

hESCs share properties with adult tissue stem cells and cancer stem cells (CSCs)[20–22]. Unlike somatic cells, mouse ESCs partially respond to p53-dependent G1/S checkpoint arrest, apoptosis, and senescence[23]. Upon activation, p53 drives ESC differentiation by suppressing pluripotency factors such as Nanog[24]. Conversely, pluripotency factors like POU5F1 counteract p53 activity by mechanisms including induction of the histone deacetylase SIRT1, which deacetylates and inactivates p53[25–28]. High-throughput transcriptomic analyses have identified numerous lncRNAs expressed in ESCs, some regulating pluripotency and tumorigenesis[29]. For example, lncRNA *linc00617/TUNAR* forms RNA-protein complexes modulating pluripotency markers including NANOG, SOX2, and FGF4, enhancing properties such as epithelial-mesenchymal transition (EMT) and tumor initiation capacity[30–33]. Another example, lncRNA *lncPRESS1*, downregulated by p53, safeguards pluripotency by disrupting SIRT6-mediated deacetylation in hESCs[34]. Although p53 is critical for regulating lncRNA expression, the impact of p53-induced lncRNAs in hESC differentiation remains largely unexplored.

In this study, we investigate how genotoxic stress impacts pluripotency and differentiation in hESCs, focusing on stem cell maldifferentiation (GSMD) and activation of transforming growth factor-β (TGF-β) signaling. Using integrative bulk and single-cell RNA sequencing, p53 chromatin immunoprecipitation sequencing (ChIP-seq), transposase-accessible chromatin sequencing (ATAC-seq), and interactome analyses, we identify the p53-induced lncRNA *LOC644656* and delineate its functional roles in GSMD. *LOC644656* accumulates in the nucleus under genotoxic stress, interacting directly with pluripotency-associated proteins and DNA repair machinery. This interaction mediates suppression of stemness and promotes resistance to genotoxic stress, suggesting a therapeutic potential in cancer and stem cell biology.

Results

Genotoxic stress induces p53-dependent maldifferentiation of hESCs

Considering the propensity of certain stem cell lineages to exhibit DNA damage-induced differentiation, we examined transcriptional responses in hESCs under p53-dependent genotoxic stress[35,36]. Accordingly, we subjected p53 wild-type (p53WT) and p53 knockout (p53KO) hESCs (Supplementary Fig. 1a) to genotoxic stress using three antineoplastic drugs, 5-fluorouracil (5-FU), adriamycin (ADR), and daunorubicin (DNR). Bulk RNA sequencing (RNA-seq) and fluorescence staining analyses (Fig. 1a) revealed that 24-hour 5-FU treatment significantly reduced pluripotency markers and increased apoptosis in a p53-dependent manner. These effects were accompanied by distinct morphological changes (Fig. 1b and Supplementary Fig. 1b), consistent with findings in mouse ESCs[24]. RNA-seq analysis further demonstrated that genotoxic stress caused aberrant upregulation of triploblastic genes, particularly those associated with the endoderm and mesoderm, alongside the p53-dependent loss of pluripotency (Fig. 1c, d, and Supplementary Fig. 1c–e). Accordingly, we designated this phenomenon as GSMD. These results suggest that p53-induced genes networks drive both the loss of stemness and maldifferentiation in hESCs under genotoxic stress.

Gene Ontology (GO) analysis of 327 upregulated genes co-regulated by p53 (Fig. 1e, f) and associated with cell development (GO_0048468, Fig. 1f) identified TGF-β regulation of the extracellular matrix as the top-ranked pathway (Fig. 1g) and an increase in GSMD-associated processes (Fig. 1h). This finding highlights the activation of TGF-β signaling along with the upregulation of developmental genes during genotoxic stress. Consistent with the suppression of pluripotency upon genotoxic stress, GO analysis of 320 downregulated genes indicated a reduction in embryonic gene expression and a concurrent increase in GSMD-associated lineages (Supplementary Fig. 1f–j).

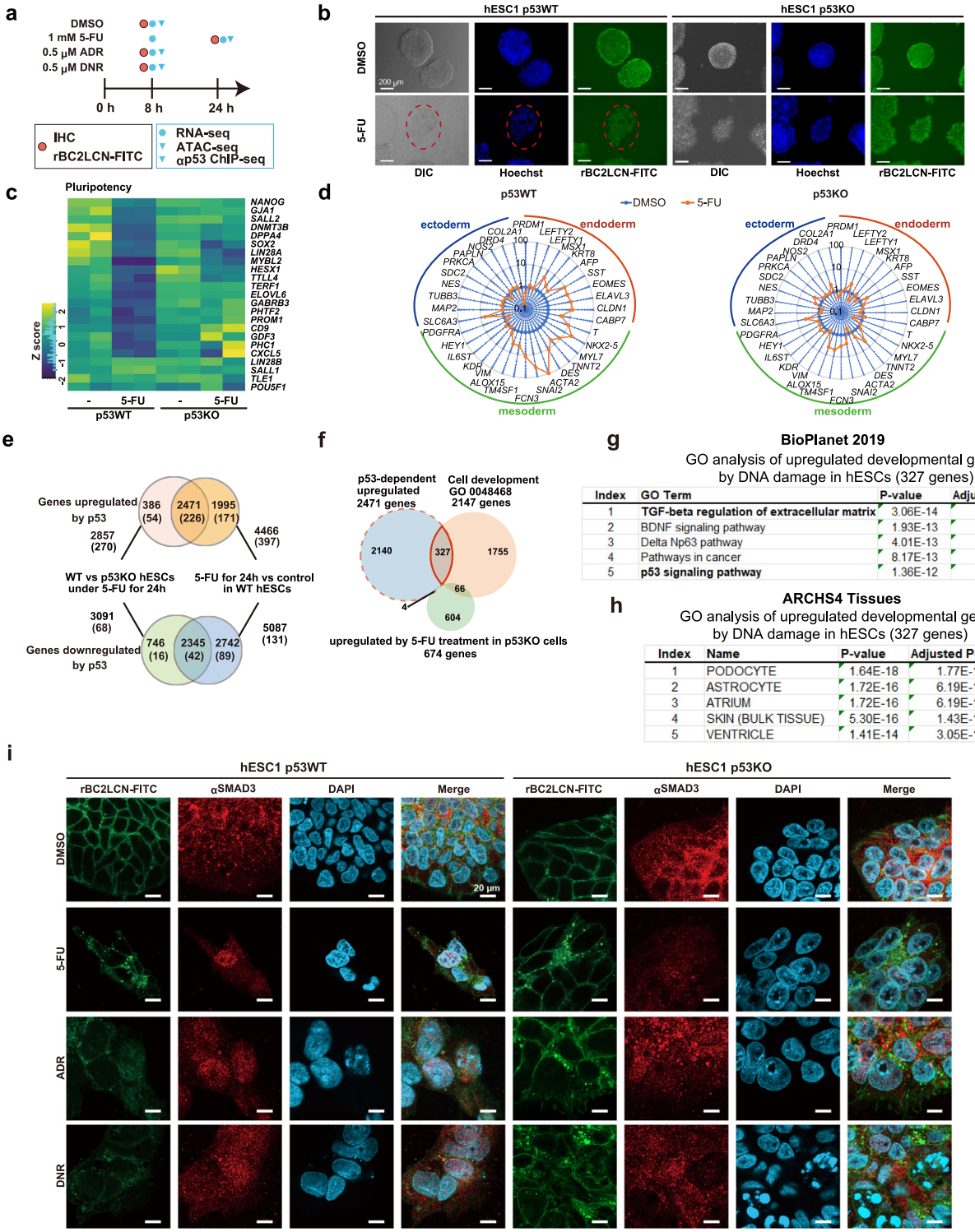
To investigate the relationship between p53 and TGF-β signaling during GSMD, we assessed SMAD3 nuclear accumulation, a hallmark of TGF-β pathway activation. Under genotoxic stress, nuclear SMAD3 accumulation occurred exclusively in p53WT hESCs (Fig. 1i and Supplementary Fig. 1k). Furthermore, phosphorylated SMAD2/3 was observed to be accumulated in the nucleus of p53WT hESCs, but not in p53KO cells (Supplementary Fig. 2a, b), thus confirming the p53-dependent activation of SMAD/TGF-β signaling during GSMD. To exclude cell line-specific effects, these findings were validated using hESC3 cells, which showed similar results (Supplementary Fig. 3a–c). Collectively, these data demonstrate that GSMD in hESCs is mediated by p53-dependent mechanisms, involving coordinated loss of pluripotency and activation of TGF-β signaling in response to genotoxic stress.

Identification of p53-regulated lncRNAs associated with genotoxic stress response in hESCs

Although p53-dependent lncRNA regulation has been studied in cancer cells[13] and the p53-downregulated *LncPRESS1* was shown to maintain pluripotency in hESCs[34], the role of p53-induced lncRNAs in stem cell response to DNA damage remains unclear. To address this, we hypothesized that p53-induced lncRNAs could act as key regulators of stemness and maldifferentiation. Using a multi-omics approach combining RNA-seq, p53 chromatin immunoprecipitation sequencing (ChIP-seq), and transposase-accessible chromatin sequencing (ATAC-seq), we aimed to identify lncRNAs associated with GSMD.

RNA-seq analysis of p53WT and p53KO hESCs treated with DMSO or 5-FU (for 24 h), ADR or DNR (for 8 h) to induce DNA damage (Fig. 1a) revealed that DNA damage substantially altered gene expression profiles in p53WT cells (Supplementary Fig. 3d). Principal component analysis indicated drug-dependent variation along the PC2 axis, while PC1 showed minimal effects on gene expression in p53KO cells (Supplementary Fig. 3d). Consistent with the observation that 5-FU treatment elicited higher specificity for p53-dependent cytotoxicity in differentiated cells compared to other drugs[37–39], our RNA-seq analyses showed that 5-FU treatment affected the expression of fewer genes than ADR or DNR in p53KO cells (Supplementary Fig. 3e–g). Accordingly, 5-FU-treated samples were selected to identify p53-regulated lncRNAs induced during the DNA damage response (DDR). Analysis of differentially expressed genes between 5-FU-treated p53WT and p53KO hESCs, compared with untreated controls, identified 4,816 p53-dependent genes (Fig. 1e). An integrated analysis of RNA-seq and ChIP-seq data, applied to 268 differentially expressed lncRNA genes filtered based on the following criteria: a > 5-fold change in RNA-seq data (DDR vs. DMSO control) and a > 5-fold change in p53-ChIP-seq data (DDR vs. DMSO control), identified 12 candidate lncRNAs (Fig. 2a and

https://doi.org/10.1038/s41467-025-59886-w

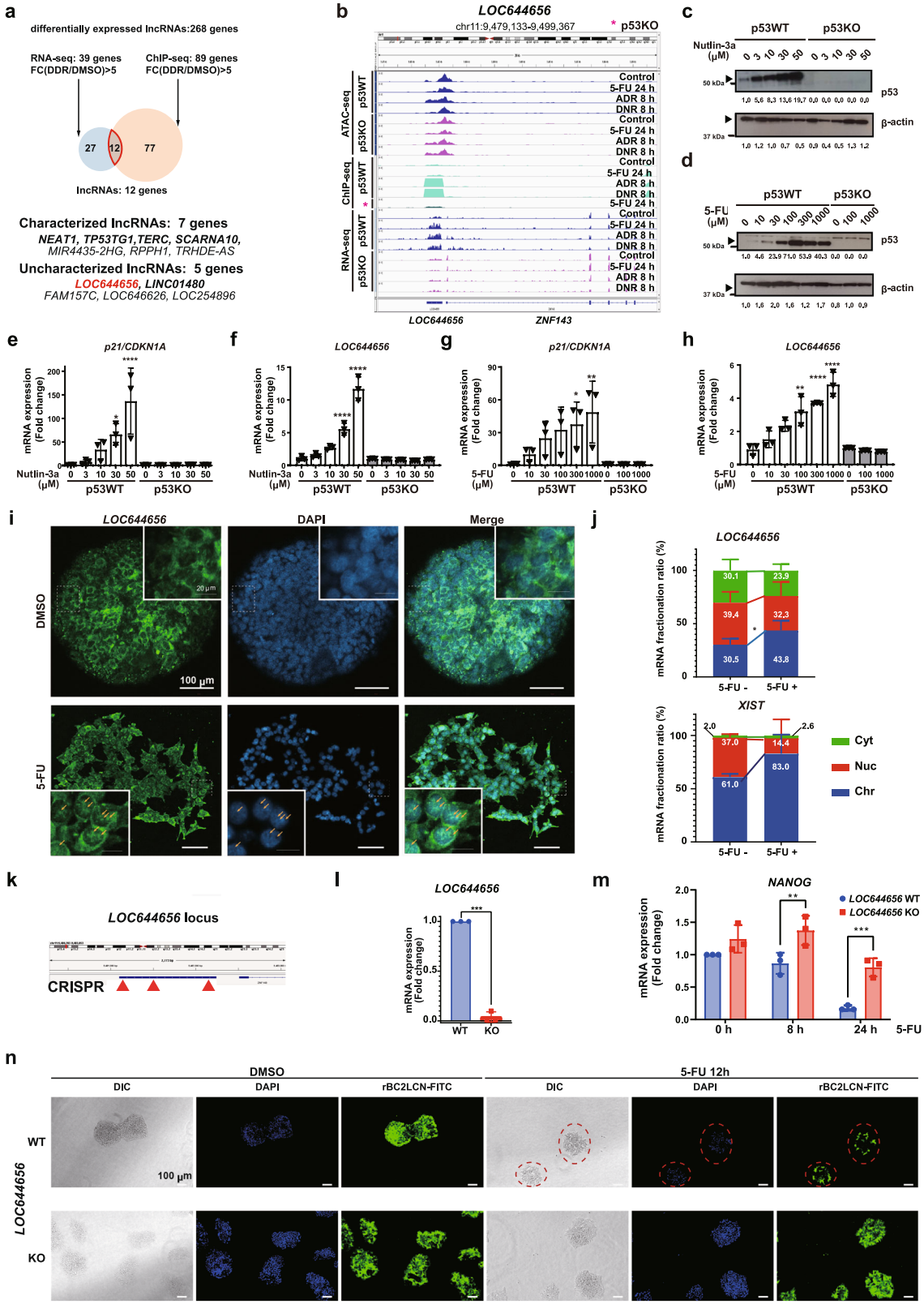


Index	GO Term	P-value	Adjusted P-value
1	**TGF-beta regulation of extracellular matrix**	3.06E-14	2.53E-11
2	BDNF signaling pathway	1.93E-13	8.00E-11
3	Delta Np63 pathway	4.01E-13	1.11E-10
4	Pathways in cancer	8.17E-13	1.69E-10
5	**p53 signaling pathway**	1.36E-12	2.25E-10

Index	Name	P-value	Adjusted P-value
1	PODOCYTE	1.64E-18	1.77E-16
2	ASTROCYTE	1.72E-16	6.19E-15
3	ATRIUM	1.72E-16	6.19E-15
4	SKIN (BULK TISSUE)	5.30E-16	1.43E-14
5	VENTRICLE	1.41E-14	3.05E-13

Fig. 1 | Genotoxic stress-induced stem cell maldifferentiation (GSMD) occurs in hESCs via TGF-β signaling. a Schematic of the workflow. **b** Stemness in hESCs was evaluated using FITC-conjugated rBC2LCN in the presence or absence of 5-FU for 24 h. Images are representative of $n = 3$ independent experiments with similar results. **c** Heatmap displaying the expression of representative mRNAs involved in pluripotency ($n = 3$ biologically independent samples). **d** Radar charts depicting the expression of typical triploblastic genes in hESCs after 5-FU treatment. Data are presented as mean ± SEM from $n = 3$ biologically independent samples. (No statistical test was performed.). **e** Genes differentially expressed in a p53-dependent manner from RNA-seq ($n = 3$ biologically independent experiments). **f** p53-dependent upregulated genes (327) overlapping with cell development-related genes (GO:0048468). The 327 and 2140 genes identified in VENNY 2.1 were subjected to GO analysis (panels **g**, **h**). **g** GO analysis of the 327 upregulated developmental genes using BioPlanet 2019. Statistical significance was determined by two-sided Fisher's exact test with Benjamini–Hochberg (BH) correction; exact p-values are shown. **h** GO analysis of the 327 upregulated developmental genes using ARCHS4 Tissues. Statistical significance was determined by two-sided Fisher's exact test with BH correction; exact p-values are shown. **i** TGF-β signaling was evaluated by immunostaining for SMAD3 and FITC–rBC2LCN under 5-FU, ADR, or DNR. Images are representative of $n = 4$ independent experiments with similar results. Scale bars, 50 μm.

https://doi.org/10.1038/s41467-025-59886-w



Supplementary Fig. 4a–d), including established p53 targets *NEAT1* and *TP53TG1*, alongside five uncharacterized candidates (*LOC644656*, *LINC01480*, *FAM157C*, *LOC644626*, and *LOC254896*). *LOC644656* and *LINC01480* emerged as top candidates due to their robust induction across all genotoxic treatments (Fig. 2b and Supplementary Fig. 4c–e).

ATAC-seq analysis revealed constitutively open chromatin at the *LOC644656* locus, independent of DNA damage status (Fig. 2b). Transcription start site analysis using the Integrative Genomics Viewer confirmed p53-dependent transcriptional activation of *LOC644656* and *LINC01480* in p53WT hESCs (Fig. 2b and Supplementary Fig. 4e). Taken together, these findings establish that p53-mediated transcriptional programs activate specific lncRNAs during genotoxic stress, suggesting their potential roles as regulators of GSMD and other cellular responses in hESCs.

Fig. 2 | p53-induced *LOC644656* is required for genotoxic stress-mediated suppression of pluripotency in hESCs. a Venn diagram of 268 differentially expressed (DE) lncRNAs identified from RNA-seq and p53 ChIP-seq analyses (DDR vs DMSO > 5-fold). **b** Integrative Genomics Viewer tracks for p53 ChIP-seq, ATAC-seq, and RNA-seq at the *LOC644656* locus. The pink asterisk indicates the α-p53 ChIP-seq track in p53KO hESCs. **c**, **d** Immunoblot analyses of p53 after treatment with Nutlin-3a (**c**) or 5-FU (**d**). Molecular weight markers (kDa) are shown on the left. Blots are representative of $n = 3$ independent experiments with similar results; uncropped blots are provided in the Source Data file. The samples derive from the same experiment but different gels for p53 and β-actin in parallel. **e**–**h** Real-time RT-PCR analysis of *p21/CDKN1A* (**e**, **g**) and *LOC644656* (**f**, **h**) relative to *ACTB* under Nutlin-3a or 5-FU treatment. Data are presented as mean ± SEM from $n = 3$ biologically independent samples. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$ (two-sided one-way ANOVA with Dunnett's post hoc test). **i**, RNA-FISH of *LOC644656* in hESCs ± 5-FU. Yellow arrows indicate nuclear speckles. Images are representative of $n = 3$ independent experiments with similar results. Scale bars: 100 µm (main images), 20 µm (insets). **j** Subcellular fractionation of *LOC644656* in cytoplasmic (Cyt), nuclear (Nuc), and chromatin (Chr) fractions. XIST is used as a nuclear control. *Data are mean ± SEM ($n = 3$), $p < 0.05$ by two-sided Student's t-test vs untreated. **k** Schematic of CRISPR/Cas9-based *LOC644656* knockout (KO). **l** *LOC644656* expression in WT vs KO hESCs. **Data are mean ± SEM ($n = 4$), $^{**}p < 0.0001$ by two-sided Student's t-test. **m** *NANOG* expression ± 5-FU in WT vs KO hESCs. **Data are mean ± SEM ($n = 3$); two-way ANOVA (two-sided) with Tukey's post hoc test, $^{**}p < 0.01$, $^{*}p < 0.001$ vs untreated WT. **n** FITC-rBC2LCN staining of pluripotency ± 5-FU for 12 h. Images are representative of $n = 3$ independent experiments with similar results. Scale bar, 100 µm. Exact p-values and 95% confidence intervals: Fig. 2e: $p = 0.017$, 95% CI [16.10, 114.3] (0 µM vs 30 µM p53WT); $p < 0.0001$, 95% CI [86.74, 185.0] (0 µM vs 50 µM p53WT); Fig. 2f: $p < 0.0001$, 95% CI [−6.118, −2.894] (0 µM vs 30 µM p53WT); $p < 0.0001$, 95% CI [−12.29, −9.061] (0 µM vs 50 µM p53WT); Fig. 2g: $p = 0.0391$, 95% CI [−71.64, −4.455] (0 µM vs 300 µM p53WT); $p = 0.0055$, 95% CI [−82.77, −12.59] (0 µM vs 1000 µM p53WT); Fig. 2h: $p < 0.0001$, 95% CI [−3.369, −1.235] (0 µM vs 100 µM p53WT); $p < 0.0001$, 95% CI [−3.852, −1.719] (0 µM vs 300 µM p53WT); $p < 0.0001$, 95% CI [−4.994, −2.860] (0 µM vs 1000 µM p53WT); Fig. 2j: $p = 0.0341$, 95% CI [2.451, 24.09]; Fig. 2l: $p = 0.001$, 95% CI [−1.091, −0.8315]; Fig. 2m: $p = 0.0052$, 95% CI [−0.8581, −0.1572] (*LOC644656* WT vs KO, 8 h); $p = 0.001$, 95% CI [−0.9778, −0.2769] (*LOC644656* WT vs KO, 24 h).

p53-induced lncRNA *LOC644656* as a key regulator of GSMD: chromatin localization and loss of pluripotency

To validate p53's role in inducing *LOC644656*, *LINC01480*, and other lncRNAs, we treated p53WT and p53KO cells with varying concentrations of 5-FU and Nutlin-3a, an MDM2 inhibitor that activates p53-dependent DNA damage signaling[40]. Immunoblot analysis demonstrated dose-dependent induction of p53 exclusively in p53WT cells for both treatments (Fig. 2c, d). This was accompanied by the dose-dependent upregulation of *LOC644656*, *LINC01480*, and the canonical p53 target gene *p21/CDKN1A* specifically in p53WT cells (Fig. 2e–h and Supplementary Fig. 5a–e). Statistical analysis confirmed significant induction of these genes in p53WT cells (Fig. 2e: $p = 0.017$ for 30 µM vs 0 µM; $p < 0.0001$ for 50 µM vs 0 µM). Similarly, known p53-induced lncRNAs, including *NEAT1*, *TP53TG1*, and *PURPL*[41,42] were upregulated in a p53-dependent manner (Supplementary Fig. 5b, c). These results confirmed the p53-dependent expression of both known and newly identified p53-target lncRNAs.

Given the nuclear localization of several lncRNAs implicated in GSMD, we investigated the subcellular distribution of *LOC644656* and *LINC01480*. RNA fluorescence in situ hybridization (RNA-FISH) in hESCs revealed that *LOC644656* displayed dual cytoplasmic and nuclear localization under basal conditions, but 5-FU treatment induced pronounced nuclear translocation and chromatin accumulation (Fig. 2i). In contrast, *LINC01480* exhibited predominantly nuclear localization under basal conditions, with 5-FU treatment leading to its redistribution to the cytoplasm (Supplementary Fig. 5g). Subcellular fractionation and RT-qPCR analyses, using *DANCR* and *XIST* as cytoplasmic and chromatin controls respectively, validated these observations, confirming 5-FU-induced chromatin enrichment of *LOC644656* (Fig. 2j, $p = 0.0341$) and increased cytoplasmic localization of *LINC01480* (Supplementary Fig. 5i). To assess the functional role of LOC644656 in GSMD, we generated *LOC644656* knockout (KO) hESCs using CRISPR/Cas9 (Fig. 2k), which showed significant reduction in *LOC644656* expression (Fig. 2l, $p = 0.001$). Notably, *LOC644656* deletion significantly attenuated the 5-FU-induced reduction in *NANOG* expression, while the expression of canonical p53 target genes, including *CDKN1A* and *PIG3*, remained unaffected (Fig. 2m and Supplementary Fig. 5j). The effect on *NANOG* expression was statistically significant at both 8 h ($p = 0.0052$) and 24 h ($p = 0.001$) after 5-FU treatment. Furthermore, *LOC644656*-KO hESCs displayed marked resistance to 5-FU-mediated suppression of pluripotency markers as assessed by FITC-rBC2LCN staining (Fig. 2n).

Collectively, these findings establish *LOC644656* as a p53-induced lncRNA that accumulates on chromatin under genotoxic stress and plays a critical role in regulating the loss of pluripotency during GSMD in hESCs.

LOC644656 interacts with the POU5F1 complex and attenuates pluripotency in hESCs

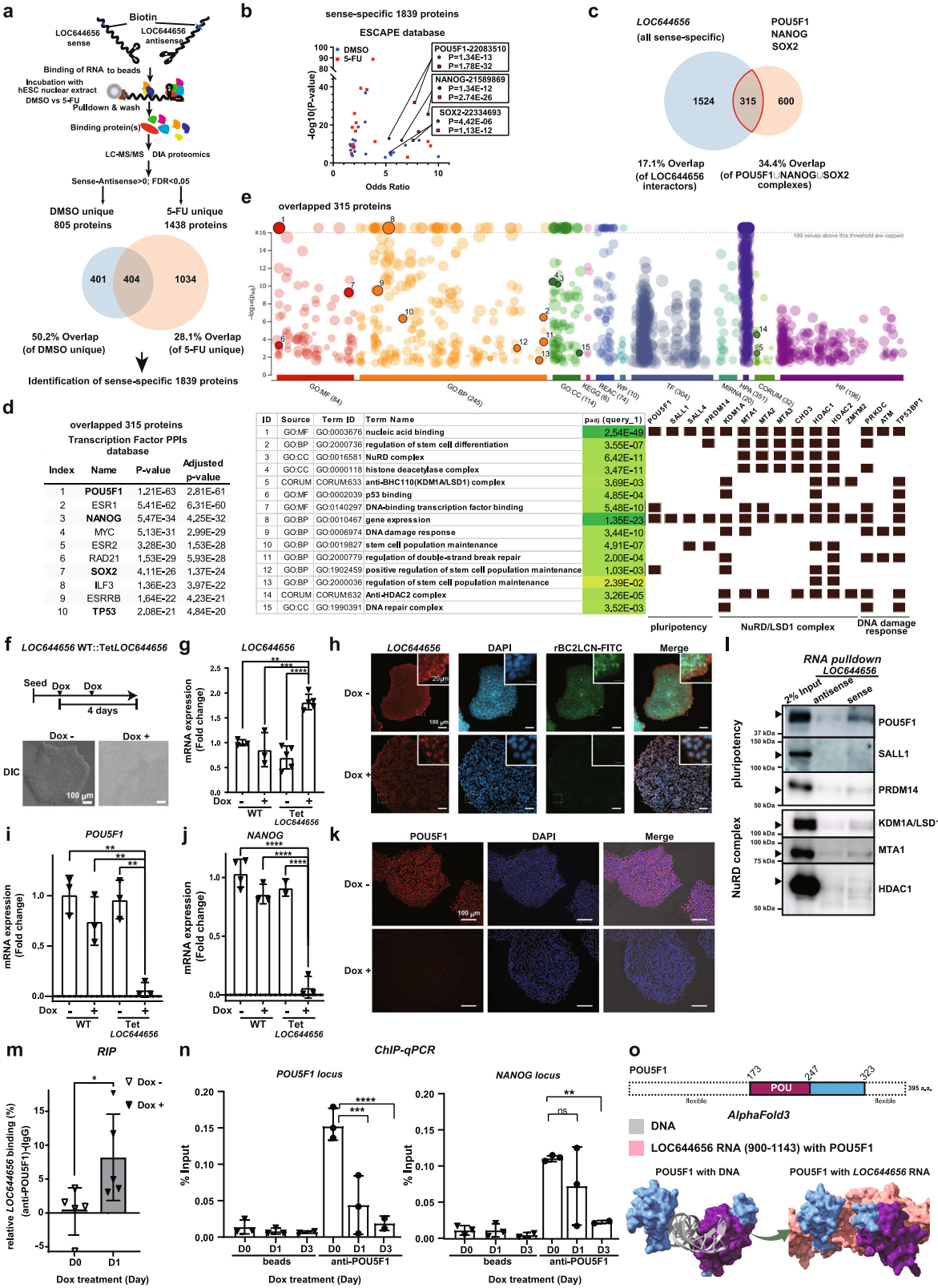
To elucidate the role of *LOC644656* in regulating pluripotency, we employed proteomic approaches to identify its interaction partners in hESCs under both normal and DNA damage conditions. RNA pulldown assays utilizing sense and antisense *LOC644656* RNA, followed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis, identified 805 and 1,438 *LOC644656*-interacting proteins in undamaged and damaged cells, respectively (Fig. 3a). Analysis of sense *LOC644656*-specific protein complexes using transcription factor protein-protein interaction networks and ESCAPE database[43,44] revealed a significant enrichment of key pluripotency factors, including *POU5F1, NANOG*, and *SOX2* (Fig. 3b). Comparative analysis demonstrated substantial overlap between *LOC644656* and POU5F1/NANOG/SOX2 interactomes, identifying 315 shared proteins (Fig. 3c, d). These results suggest that *LOC644656* predominantly associates with pluripotency-regulating factors, particularly the POU5F1 transcription factor complexes (Fig. 3d).

Additionally, Gene Ontology analysis of *LOC644656* interactors revealed enrichment of chromatin-modifying complexes, including the KDM1A/LSD1-NuRD repressor complex (Fig. 3e; ID3-5, 14) and DNA damage response proteins such as the PRKDC/DNA-PKcs complex (Fig. 3e; ID9, 11, 15). Collectively, *LOC644656* was found to associate with at least three major functional complexes: the POU5F1 transcriptional complex, the KDM1A/LSD1-NuRD repressor complex, and the PRKDC/DNA-PKcs complex. This indicates that *LOC644656* may function as a molecular hub, integrating pluripotency regulation with DNA damage signaling during GSMD.

Given that certain lncRNAs function as nuclear decoys[45], we hypothesized that *LOC644656* might suppress pluripotency by sequestering *POU5F1* and its associated factors in hESCs. Using a Tet-inducible system for *LOC644656* expression, we observed a nuclear accumulation of *LOC644656* transcripts that led to a reduction in hESC stemness, independent of p53 (Fig. 3f–h and Supplemental Fig. 6a, b; Fig. 3g *LOC644656*: $p = 0.001$, 95% CI [−1.282, −0.3534]). *LOC644656* induction suppressed the expression of *POU5F1* and *NANOG* (Fig. 3i, j *POU5F1*: $p = 0.0011$, 95% CI [0.4657, 1.430]; Fig. 3j *NANOG*: $p < 0.0001$, 95% CI [0.7270, 1.213]), which was corroborated by immunofluorescence showing reduced POU5F1 protein levels (Fig. 3k).

RNA pulldown assays confirmed the direct interaction of sense *LOC644656* RNA with the POU5F1- and KDM1A/LSD1-NuRD complexes in vitro (Fig. 3l). RNA immunoprecipitation (RIP) assays further validated the specific binding of LOC644656 to POU5F1 in cells (Fig. 3m, $p = 0.0368$, 95% CI [0.7986, 15.18]). Functionally, *LOC644656* induction inhibited POU5F1 recruitment to regulatory regions of *POU5F1* and

https://doi.org/10.1038/s41467-025-59886-w



NANOG locus (Fig. 3n, POU5F1 locus: $p = 0.0002$, 95% CI [0.06560, 0.1564]; *NANOG* locus: $p = 0.0057$, 95% CI [0.02939, 0.1483]).

To further delineate the functional domain of *LOC644656*, a series of doxycycline (Dox)-inducible *LOC644656* deletion mutants was generated and assessed for their impact on stemness maintenance (Supplementary Fig. 6c–e). Deletions in the 5′ region of *LOC644656* had minimal impact on its ability to reduce stemness, whereas deletions in the 3′ region significantly abrogated its suppressive activity (Supplementary Fig. 6d, e, $p < 0.0001$). This indicates that the region spanning 900-1143 bp in *LOC644656* is essential for its pluripotency reducing function.

To document the mechanism by which *LOC644656* inhibits POU5F1 activity, we performed structural modeling of POU5F1/DNA/*LOC644656* complexes using AlphaFold3[46]. Docking simulations, which

Fig. 3 | *LOC644656* interacts with the POU5F1 complex to attenuate pluripotency in hESCs. a Schematic of the RNA pulldown assay. **b** ESCAPE database analysis of the 1,839 *LOC644656*-interacting proteins (±5-FU). Key pluripotency transcription factors are highlighted. **c** Overlap between LOC644656 interactors and the POU5F1/NANOG/SOX2 complex. **d** Transcription factor protein–protein interaction network of the 315 overlapping proteins. **e** Functional categories and GO terms enriched among the identified interactors. **f** Experimental workflow for hESC::Tet*LOC644656* induction. **g** *LOC644656* expression ± doxycycline (Dox). Data are mean ± SEM ($n = 3$), $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$ by two-sided one-way ANOVA. **h** RNA-FISH of *LOC644656* (green) and FITC–rBC2LCN (red) ± Dox for four days. Images are representative of $n = 3$ independent experiments with similar results. Scale bars: 100 µm (main images), 20 µm (insets). **i**, **j** RT–PCR analysis of *POU5F1*(i) and *NANOG* (**j**) ± Dox. Data means SEM ($n = 3$), $^{**}p < 0.01$ by two-sided one-way ANOVA. **k** Immunofluorescence for POU5F1 ± Dox for three days. Representative of $n = 3$ independent experiments. Scale bar, 50 µm. **l** RNA pulldown assay using biotinylated sense or antisense *LOC644656*. Representative of $n = 3$ independent experiments. Uncropped blots are provided in Source Data. The samples derive from the same experiment but different gels for POU5F1, NANOG, another for LSD1 and another for HDAC1 were processed in parallel. **m** RIP assay with an anti-POU5F1 antibody in hESC::Tet*LOC644656* ± Dox. Data means SEM ($n = 3$), $^{*}p < 0.05$ by two-sided paired t-test. **n** ChIP assay of POU5F1 binding at target loci ± Dox. Data means SEM ($n = 3$), $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$ by two-sided one-way ANOVA. **o** AlphaFold3 structural modeling of POU5F1 bound to DNA (left) or *LOC644656* RNA (right). Exact p-values and 95% confidence intervals: Panel **g**: $p = 0.001$, 95% CI [−1.282, −0.3534] (WT Dox - vs WT::Tet*LOC644656* Dox +); $p = 0.0003$, 95% CI[−1.423, −0.4947] (WT Dox + vs WT::Tet*LOC644656* Dox +); $p < 0.0001$, 95% CI [−1.517, −0.7135] (WT::Tet*LOC644656* Dox - vs WT::Tet*LOC644656* Dox +). Panel **i**: $p = 0.0011$, 95% CI [0.4657, 1.430] (WT Dox - vs WT::Tet*LOC644656* Dox +); $p = 0.0082$, 95% CI[0.2014, 1.166] (WT Dox + vs WT::Tet*LOC644656* Dox +); $p = 0.0015$, 95% CI [0.4160, 1.381] (WT::Tet*LOC644656* Dox - vs WT::Tet*LOC644656* Dox +). Panel **j**: $p < 0.0001$, 95% CI [0.7270, 1.213] (WT Dox - vs WT::Tet*LOC644656* Dox +); $p < 0.0001$, 95% CI[0.5330, 1.052] (WT Dox + vs WT::Tet*LOC644656* Dox +); $p < 0.0001$, 95% CI [0.5568, 1.138] (WT::Tet*LOC644656* Dox - vs WT::Tet*LOC644656* Dox +). Panel **m**: $p = 0.0368$, 95% CI [0.7986, 15.18] (Day 0 vs Day 1, Dox +). Panel **n**: POU5F1 locus: $p = 0.0002$, 95% CI [0.06560, 0.1564] (Day 0 vs Day 1, Dox +); $p < 0.0001$, 95% CI[0.08524, 0.1868] (Day 0 vs Day 3, Dox +); NANOG locus: $p = 0.0057$, 95% CI [0.02939, 0.1483] (Day 0 vs Day 3, Dox +).

closely aligns with previous crystal structure analyses from Protein Data Bank (8G87; Supplementary Fig. 7a, b), displayed that *LOC644656* binding induces conformational changes in POU5F1 (Supplementary Fig. 7c–e), compromising its DNA recognition capability (Fig. 3o and Supplementary Fig. 7d). These results suggest that nuclear-localized *LOC644656* functions as a molecular decoy by directly binding to POU5F1 and disrupting its transcriptional activity in hESCs.

Collectively, these findings establish *LOC644656* as a regulatory RNA that modulates stem cell identity by acting as a specific decoy for *POU5F1*, thereby interfering with its chromatin-associated functions.

LOC644656 induces SMAD/ TGF-β pathway activation to drive fibroblast-like differentiation in hESCs

To understand how *LOC644656* expression coordinately affects stemness and differentiation at the single-cell level, we performed single-cell RNA sequencing (scRNA-seq) on Tet*LOC644656* hESCs under three conditions: control (PBS; Day 0), three-day doxycycline treatment (Day 3), and six-day doxycycline treatment (Day 6). Analysis of 12,707 cells identified 13 distinct clusters (Fig. 4a, b). *LOC644656* induction significantly altered the proportional distribution of cells among clusters. Doxycycline-enriched clusters (0, 2, 4, 7, 8, and 10) exhibited reduced expression of pluripotency markers such as *POU5F1*, *NANOG*, *SOX2*, and *DNMT3B* (Fig. 4c, d, and Supplementary Fig. 8, Fig. 4d POU5F1: $p < 0.0001$).

To elucidate *LOC644656*-regulated pathways, we performed unbiased gene clustering analysis of the scRNA-seq data, revealing 16 distinct gene modules. Hierarchical clustering analysis demonstrated distinct patterns between control cells (Day 0; clusters 1, 3, 5, 6, 9, 11) and doxycycline-treated cells (Days 3 and 6; clusters 0, 2, 4, 7, 8, 10), with treated cells predominantly associated with modules 3, 8, 11, and 16 (Fig. 4e). Gene Ontology analysis of these modules highlighted enrichment of TGF-β signaling components (Fig. 4f). Differential expression analysis identified TGF-β and its target genes, such as *TGFB1*, *VIM*, and *TAGLN*, among the top differentially expressed genes (DEGs) in doxycycline-induced clusters 2 and 8 (Fig. 4g). UMAP and violin plots further confirmed the upregulation of *TGFB1*, *VIM*, and *TAGLN* following *LOC644656* induction (Supplementary Fig. 9a–g, all $p < 0.0001$). Notably, cluster 2 DEGs were associated with SMAD3 and SMAD4 emerged as key upstream regulators (Supplementary Fig. 9h) and myoblast and fibroblast signatures (Supplementary Fig. 9i). Consistently, *LOC644656* induction increased SMAD2/3 phosphorylation and nuclear localization in both p53WT and p53KO hESCs (Fig. 4h–j and Supplementary Fig. 9j, k, Fig. 4j: $p < 0.0001$, 95% CI [−46.02, −18.81] (Day 0 vs Day 3, $n > C$); $p < 0.0001$, 95% CI[−112.9, −85.67] (Day 0 vs Day 6, $n > C$)).

Mechanistically, given that KDM1A/LSD1 forms a repressor complex with SMAD2/3 to negatively regulate EMT-related genes[47] and *LOC644656* interact with KDM1A/LSD1-NuRD complex (Fig. 3e, l), we hypothesized that *LOC644656* functions as a molecular decoy to disrupt this repressor complex in hESCs. Supporting this hypothesis, SMAD2 immunoprecipitation showed that *LOC644656* induction abrogated the endogenous interaction of SMAD2 and KDM1A/LSD1 (Fig. 4k). As a result, *LOC644656* expression enhanced TGF-β target gene activation in a time-dependent manner (Fig. 4l). These findings suggest that *LOC644656* acts as a molecular decoy to derepress the SMAD-KDM1A/LSD1 complex, thereby facilitating SMAD/TGF-β pathway activation.

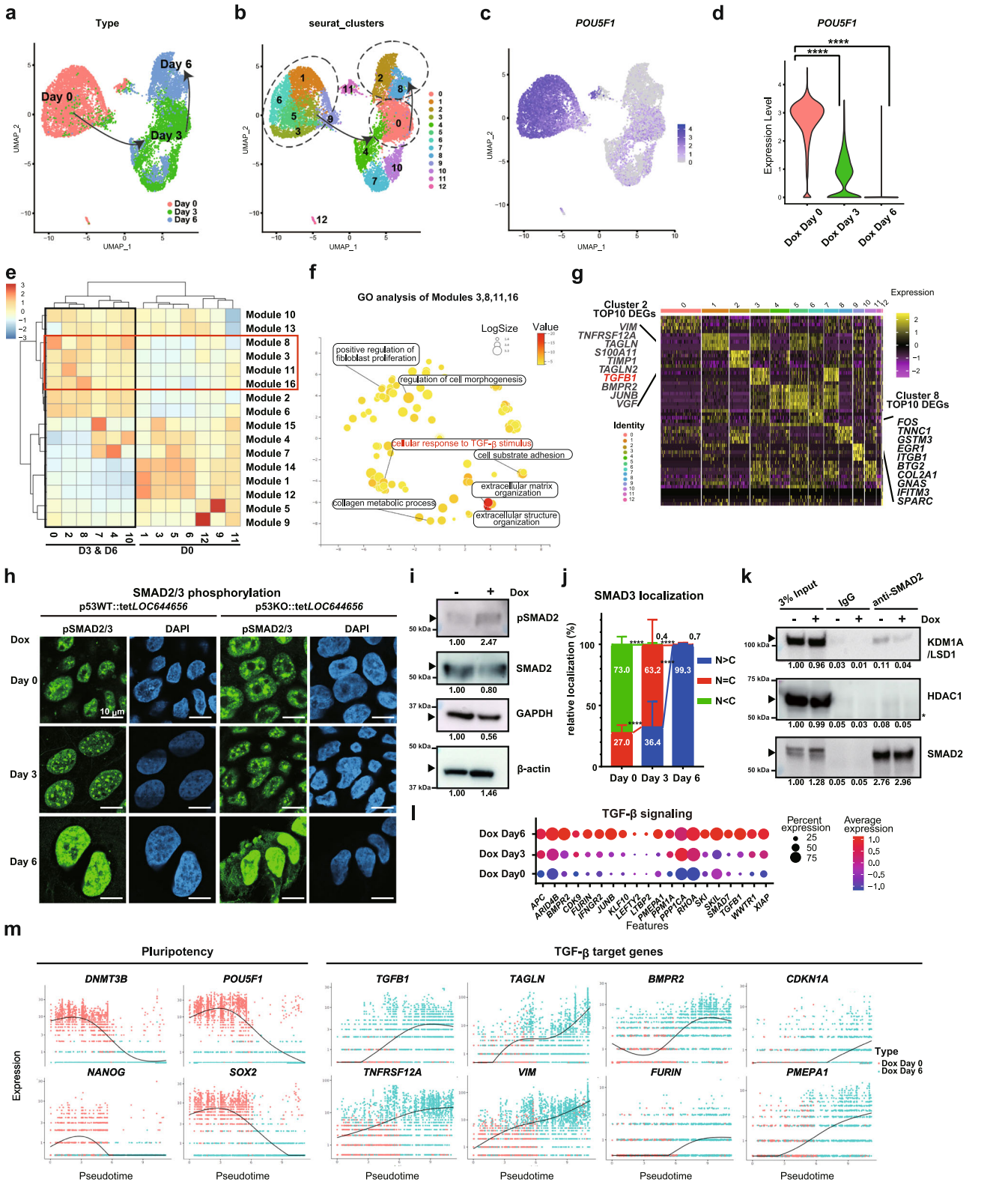
Given the established role of TGF-β pathways in cell cycle regulation[48], we investigated whether *LOC644656* induction affects cell cycle-related genes associated with stemness. Integrated analysis of scRNA-seq data from control and day 6 doxycycline-treated cells revealed an increased proportion of G1-phase cells following *LOC644656* induction (Supplementary Fig. 9l, m). To examine the temporal relationship between stemness suppression and TGF-β pathway activation during *LOC644656*-induced differentiation, we performed pseudotime analysis. As differentiation predominantly occurs in G1 phase[49], G1-phase cells from cluster 3 at Day 0 were designated as the root state, and their developmental trajectory was tracked following *LOC644656* expression (Supplementary Fig. 9n, o). This analysis revealed a sequential process in which *LOC644656* first suppresses stemness markers, followed by activation of TGF-β signaling pathways (Fig. 4m). These results demonstrate that *LOC644656* coordinately drives cell cycle arrest and fibroblast-like differentiation through SMAD/TGF-β signaling activation, recapitulating key features of GSMD.

LOC644656 attenuates DDR signaling through DNA-PKcs interaction to prevent genotoxic stress-induced apoptosis

Given its interaction with DNA-PKcs (PRKDC; Fig. 3e), we explored whether *LOC644656* modulates the DNA damage response (DDR). Single cell RNA-seq analysis of Tet*LOC644656* hESCs under three conditions [control (DMSO), doxycycline(-)5-FU, or doxycycline(+) 5-FU] revealed 11 distinct clusters across a total of 9,891 cells (Fig. 5a, b, and Supplementary Fig. 10a). Cells treated with 5-FU alone were enriched in clusters 6 and 8, whereas those treated with both 5-FU and doxycycline were predominantly found in clusters 0 and 2, with G1 phase accumulation (Fig. 5a, b, and Supplementary Fig. 10b).

Unbiased gene module analysis identified 15 gene modules, with modules 9–11 being relatively enriched in doxycycline (+)5-FU-treated

https://doi.org/10.1038/s41467-025-59886-w



cells (Fig. 5c and Supplementary Fig. 10c). REVIGO analysis[50] of the genes within these modules revealed significant enrichment in pathways associated with signaling by p53 class mediator, DNA damage response, DNA replication, and mitosis-related genes (Fig. 5d, e).

Notably, RNA pulldown assays and docking simulations utilizing AlphaFold3, which closely aligns with previous crystal structure analyses from Protein Data Bank (Supplementary Fig. 11a, b), demonstrated that *LOC644656* interacts with DNA-PKcs, inducing conformational changes in its own structure (Supplementary Fig. 11c). This interaction specifically occludes both the kinase domain of DNA-PKcs and the phosphorylation sites critical for its activity (Fig. 5f and Supplementary Fig. 11c). Supporting this simulation model, sense *LOC644656* RNA bound to DNA-PKcs proteins in vitro (Fig. 5g), and in vitro kinase assays further confirmed that sense *LOC644656* significantly inhibited the autophosphorylation of DNA-PKcs (Fig. 5h, $p = 0.0034$, 95% CI [−1.054, −0.2608]).

Fig. 4 | *LOC644656* induction suppresses stemness and activates TGF-β signaling in hESCs at the single-cell level. a–d hESC::Tet*LOC644656* cells were cultured ± Dox for 0, 3, or 6 days and analyzed by scRNA-seq ($n = 12{,}707$ cells, one library). **a** UMAP plot showing cell distribution colored by time point (Day 0, Day 3, Day 6). **b** UMAP plot with cluster assignments (numbered 0-12). **c** UMAP visualization of *POU5F1* expression across clusters. **d** Violin plot comparing *POU5F1* expression levels between time points. Statistical analysis by two-sided Wilcoxon rank sum test. **e** Heatmap showing module-cluster relationships with modules 3, 8, 11, and 16 notably upregulated at Days 3 and 6. **f** Gene Ontology analysis (REVIGO) of genes in modules 3, 8, 11, and 16. **g** Heatmap of differentially expressed genes by cluster. **h** Representative immunostaining images of phospho-SMAD2/3 ± Dox at Day 3 or Day 6. Images are from $n = 3$ independent experiments with similar results. Scale bar, 50 µm. **i** Immunoblot analysis of SMAD2 phosphorylation ± Dox for 6 days. Representative of $n = 3$ independent experiments. Samples derive from the same experiment and were processed on the same gel. Uncropped blots are provided in the Source Data file. **j** Quantification of SMAD3 localization ($n >$ C: nuclear>cytoplasmic; $n =$ C: nuclear=cytoplasmic; $n <$ C: nuclear<cytoplasmic). Data are mean ± SEM ($n = 3$). Statistical analysis by two-sided one-way ANOVA with Tukey's post hoc test. **k** Co-immunoprecipitation of SMAD2/LSD1/HDAC1 complex in nuclei ± LOC644656 induction. Representative of $n = 3$ independent experiments. Samples derive from the same experiment but SMAD2, LSD1, and HDAC1 were analyzed on separate gels in parallel. Uncropped blots are provided in the Source Data file. **l** Dot plot showing TGF-β target gene expression across clusters and treatment conditions. Dot size indicates percentage of expressing cells; color intensity shows average expression level. Statistical analysis by two-sided Wilcoxon rank sum test. **m** Pseudotime trajectory analysis (Monocle 3) showing progressive loss of pluripotency and gain of TGF-β signaling. Exact p-values and 95% confidence intervals: Panel **d** POU5F1: $p < 0.0001$ (Day 3 vs Day 0 and Day 6 vs Day 0). Panel **j** SMAD3 localization: $n >$ C pattern: Day 0 vs Day 3: $p < 0.0001$, 95% CI [−46.02, −18.81]; Day 0 vs Day 6: $p < 0.0001$, 95% CI [−112.9, −85.67]; Day 3 vs Day 6: $p < 0.0001$, 95% CI [−79.03, −54.69]. $n =$ C pattern: Day 0 vs Day 3: $p < 0.0001$, 95% CI [−52.83, −25.62]; Day 0 vs Day 6: $p < 0.0001$, 95% CI [13.45, 40.66]; Day 3 vs Day 6: $p < 0.0001$, 95% CI [54.11, 78.45]. $n <$ C pattern: Day 0 vs Day 3: $p < 0.0001$, 95% CI [58.04, 85.25]; Day 0 vs Day 6: $p < 0.0001$, 95% CI [58.62, 85.83]. Panel **l**: Complete statistical analysis with exact p-values is provided in the Source Data file.

In cells expressing *LOC644656*, 5-FUinduced phosphorylation of DNA-PKcs and its downstream kinases was suppressed, accompanied by reduced activation of p53 following genotoxic stress (Fig. 5i, pDNA-PKcs/DNA-PKcs: $p = 0.0082$, 95% CI [−1.091, −0.3419]; pCHK1/β-actin: $p = 0.0145$, 95% CI [−1.238, −0.2731]; p53/β-actin: $p = 0.0044$, 95% CI [−1.200, −0.6516]). Additionally, *LOC644656* induction suppressed DNA double-strand break (DSB)-induced γH2AX activation after Neocarzinostatin treatment, independently of p53 (Supplementary Fig. 12). In contrast, *LOC644656* knockout cells displayed enhanced phosphorylation of DNA-PKcs and CHK1 in response to genotoxic stress (Fig. 5j, pDNA-PKcs/DNA-PKcs: $p = 0.0363$, 95% CI [−1.557, −0.1296]; pCHK1/CHK1: $p = 0.0271$, 95% CI [0.1147, 1.360]).

LOC644656 induction specifically downregulated key DNA damage-sensing genes, including *PRKDC*, *CHEK1*, and *CHEK2*, while *ATM* and *ATR* expression remained unchanged, even after genotoxic stress conditions (Supplementary Fig. 10d–n, all $p < 0.0001$). Differential expression analysis revealed the upregulation of p53 target genes, such as *FDXR* and *GLS2* in clusters 6 and 8 under 5-FU treatment without doxycycline [Dox (−)/5-FU (+)], whereas these genes were downregulated in cluster 10 under doxycycline induction [Dox (+)/5-FU (+)] (Fig. 5k and Supplemental Fig. 10o–s, all $p < 0.0001$). Collectively, these results suggest that *LOC644656* attenuates genotoxic stress-induced p53 signaling by modulating the DDR pathway through inhibiting DNA damage-sensing proteins.

To further investigate *LOC644656*'s role in apoptosis regulation, we examined its effects on genotoxic stress-induced apoptotic pathways. *LOC644656* expression prevented the 5-FU-induced upregulation of the pro-apoptotic factor *PUMA/BBC3* (Fig. 5l, $p = 0.0031$, 95% CI [−322.2, −65.49]). Flow cytometric analysis of Annexin V staining revealed that, while ADR treatment increased the proportion of Annexin V and DAPI double-positive apoptotic cells, *LOC644656* induction significantly attenuated this ADR-induced apoptotic response (Fig. 5m, n, Fig. 5n: $p < 0.0001$, 95% CI [54.57, 92.58]). These results demonstrate that p53-induced *LOC644656* acts as a negative feedback regulator by directly binding to and inhibiting DNA-PKcs. Through this mechanism, *LOC644656* protects hESCs from genotoxic stress-induced apoptosis by suppressing the DDR pathway and mitigating pro-apoptotic signaling.

Shared TGF-β activation and cancer-related gene networks in *LOC644656*-induced differentiation and chemotherapy-mediated GSMD

To elucidate the shared molecular basis underlying *LOC644656*-induced stem cell differentiation and GSMD triggered by chemotherapeutic agents, we performed comparative scRNA-seq analysis across control (DMSO), *LOC644656*-expressing, and 5-FU-treated conditions.

UMAP analysis (Supplementary Fig. 13a) and subsequent identification of treatment-specific differentially expressed genes (DEGs) revealed 287 commonly upregulated and 509 commonly downregulated genes between the 5-FU and *LOC644656* conditions (Supplementary Fig. 13b, c). Gene Ontology analysis demonstrated that the commonly upregulated genes were significantly enriched in TGF-β-associated pathways (Supplementary Fig. 13d–j), whereas the commonly downregulated genes were associated with cell cycle regulation, DNA replication, and the loss of pluripotency (Supplementary Fig. 13k–q). Notably, these DEGs showed significant overlap with gene signatures linked to neoplasm metastasis and carcinogenesis (Supplementary Fig. 13i, p), suggesting a mechanistic connection between *LOC644656*-induced stem cell response, DDR in hESCs and tumor progression.

To further validate these observations, we confirmed that both 5-FU treatment and *LOC644656* induction led to G1 phase arrest (Supplementary Fig. 14a–d). These results supported that p53-mediated *LOC644656* expression activates TGF-β signaling and induces cell cycle arrest. To better characterize the gene sets involved in these processes, we employed a scoring method incorporating pathways such as p53 signaling, epithelial-mesenchymal transition (EMT), and related mechanisms. As shown in Supplementary Fig. 14e–l, genes associated with EMT, focal adhesion, TGF-β signaling, and neoplasm metastasis were significantly upregulated under both *LOC644656* expression and 5-FU treatment compared to control conditions (all $p < 0.0001$).

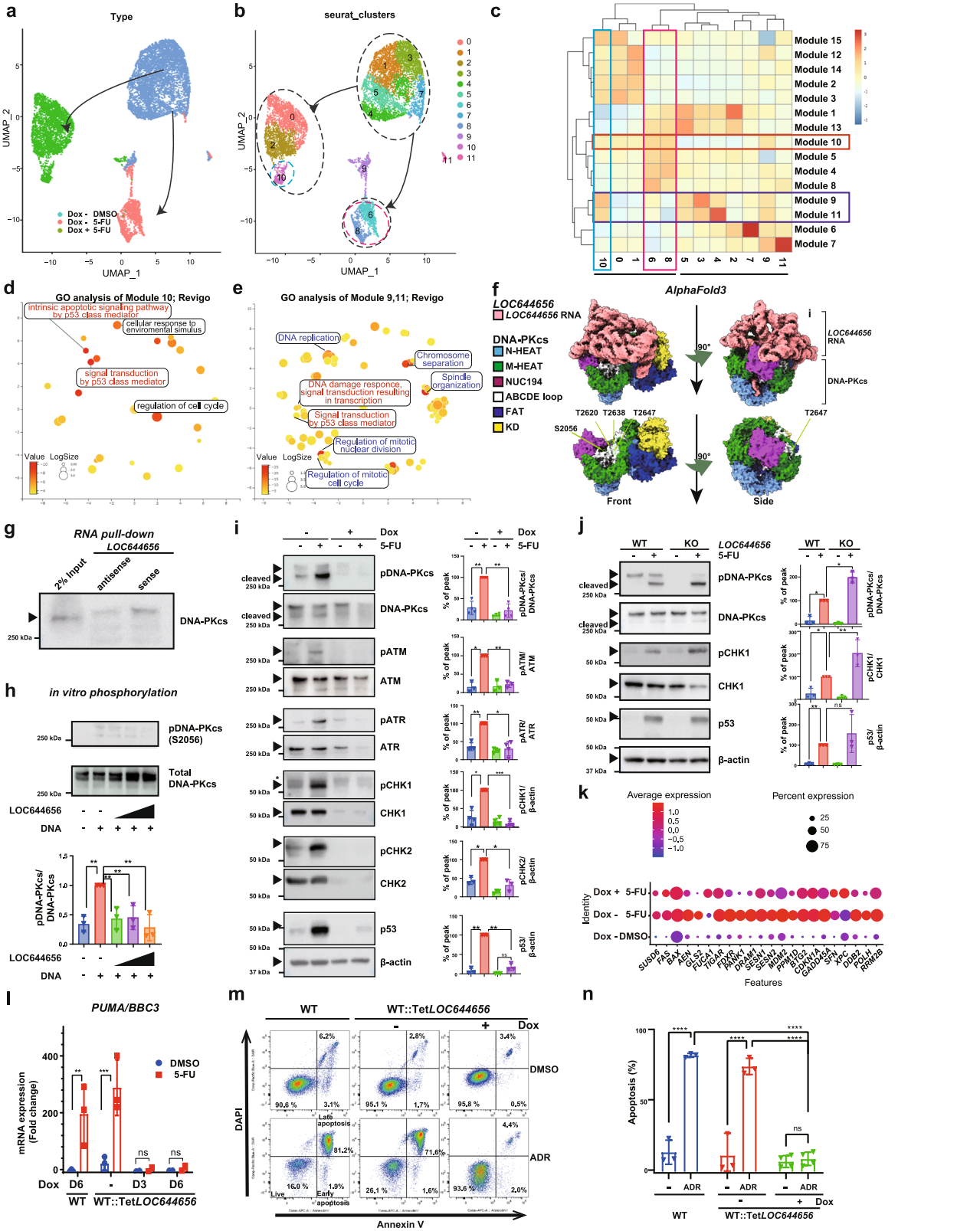
Taken together, these findings suggest that *LOC644656* expression and genotoxic stress converge on tumorigenesis-related gene networks through TGF-β activation, highlighting a shared mechanism driving cancer progression and metastasis.

LOC644656 induces GSMD-like features in cancer and correlates with poor prognosis

To investigate the clinical relevance of *LOC644656*, we analyzed its expression in tumor samples using TNM-plotter algorithm[51] and The Cancer Genome Atlas (TCGA). *LOC644656* expression was markedly upregulated in multiple tumor types compared to corresponding normal tissues (Supplementary Fig. 15a), and was associated with poor prognosis in several cancers, including liver hepatocellular carcinoma (LIHC; HR = 1.45), kidney renal clear cell carcinoma (KIRC; HR = 2.44), and breast invasive carcinoma (BRCA; HR = 1.29; Fig. 6a–c; Supplementary Fig. 15b, Fig. 6b LIHC: logrank $p = 0.037$; Fig. 6c KIRC: logrank $p = 2.6e\text{-}09$).

TCGA analysis revealed a complex relationship between *LOC644656* expression and patient outcomes. While high *LOC644656* expression exhibited a trend toward improved overall survival in 12

https://doi.org/10.1038/s41467-025-59886-w



tumor types, high *LOC644656* expression was associated with a tendency toward poor prognosis in 10 distinct cancer types (Fig. 6a). In addition to LIHC and KIRC, elevated *LOC644656* expression was significantly associated with poor prognosis in colon adenocarcinoma (COAD; HR = 1.6; Supplemental Fig. 15c, logrank $p = 0.048$), esophageal carcinoma (ESCA; HR = 1.85; Supplementary Fig. 15d), and prostate adenocarcinoma (PRAD; HR > 10; Supplementary Fig. 15e, logrank $p = 0.014$). Notably, analysis of molecular subtypes in COAD revealed a significant correlation between high *LOC644656* expression and poor prognosis, specifically within the CMS3, metabolic subtype (HR = 2.62; Supplementary Fig. 15l, logrank $p = 0.00094$).

To elucidate the mechanism basis of *LOC644656*'s role in cancer prognosis, we analyzed *LOC644656*-correlated genes across 22 tumor types, identifying 7,650 genes positively correlated with *LOC644656*

Fig. 5 | *LOC644656* induction prevents genotoxic stress-induced apoptosis in hESCs. a, b hESC::Tet*LOC644656* cells (± Dox for 5 days) were treated with DMSO or 5-FU for 24 h, then subjected to scRNA-seq ($n = 9{,}891$ cells, one library). **a** UMAP plot colored by treatment condition. **b** UMAP with cluster assignments (numbered 0-10). **c** Hierarchical clustering of gene modules versus cell clusters. Red box indicates modules enriched in clusters 6 & 8; blue box highlights module 10. **d, e** Gene Ontology analysis (REVIGO) of module 10 (d) and modules 9 & 11 (**e**). **f** AlphaFold3 structural model showing DNA-PKcs interaction with *LOC644656*. **g** RNA pulldown assay demonstrating DNA-PKcs binding to sense *LOC644656*. Representative of $n = 3$ independent experiments. Uncropped blots are provided in the Source Data file. **h** In vitro kinase assay measuring DNA-PKcs autophosphorylation ± *LOC644656*. Data are mean ± SEM ($n = 3$), $^{*}p < 0.05$, $^{**}p < 0.01$ by two-sided repeated measures ANOVA. Each experiment was separately analyzed with phospho-DNA-PKcs and DNA-PKcs antibodies. Uncropped blots are provided in the Source Data file. **i** Immunoblot analysis of DNA damage response (DDR) proteins in hESC::Tet*LOC644656* cells ± Dox for 2 days, then 5-FU for 24 h. Data are mean ± SEM ($n = 3$), $^{*}p < 0.05$, $^{**}p < 0.01$ by two-sided one-way ANOVA. The samples derive from the same experiment but different gels for each antibody were processed in parallel. Uncropped blots are provided in the Source Data file. **j** Immunoblot analysis in *LOC644656* WT vs KO hESCs after 5-FU treatment for 6 **h** Data are mean ± SEM ($n = 3$), $^{*}p < 0.05$, $^{**}p < 0.01$ by two-sided two-way ANOVA. The samples derive from the same experiment but different gels for each antibody were processed in parallel. Uncropped blots are provided in the Source Data file. **k** Dot plot showing expression of major p53 target genes from scRNA-seq analysis. Dot size represents percentage of expressing cells; color intensity shows average expression level. **l** Real-time RT-PCR analysis of PUMA/BBC3 expression ± Dox, ± 5-FU. Data are mean ± SEM ($n = 3$), $^{**}p < 0.01$, $^{***}p < 0.001$ by two-sided two-way ANOVA. Complete source data are provided in the Source Data file. **m, n** Flow cytometric analysis of apoptosis using Annexin V staining in hESC::Tet*LOC644656* cells ± Dox for 24 h, followed by ADR treatment for 24 h. Data are mean ± SEM ($n = 3$), $^{****}p < 0.0001$ by two-sided one-way ANOVA with Tukey's post hoc test. Exact p-values and 95% confidence intervals: Panel **h** (kinase assay): lane 1 vs lane 2: $p = 0.0034$, 95% CI [−1.054, −0.2608]; lane 2 vs lane 3: $p = 0.0083$, 95% CI [−0.9610, −0.1683]; lane 2 vs lane 4: $p = 0.0099$, 95% CI [−0.9436, −0.1509]; lane 2 vs lane 5: $p = 0.0019$, 95% CI [−1.115, −0.3218]. Panel **i** (key protein ratios): pDNA-PKcs/DNA-PKcs: lane 1 vs lane 2: $p = 0.0082$, 95% CI [−1.091, −0.3419]; lane 2 vs lane 4: $p = 0.0067$, 95% CI [0.3992, 1.163]. pATM/ATM: lane 1 vs lane 2: $p = 0.0216$, 95% CI [−1.404, −0.2994]; lane 2 vs lane 4: $p = 0.0086$, 95% CI [0.4656, 1.092]. pATR/ATR: lane 1 vs lane 2: $p = 0.0043$, 95% CI [−0.9063, −0.3710]; lane 2 vs lane 4: $p = 0.0191$, 95% CI [0.2063, 1.184]. pCHK1/β-actin: lane 1 vs lane 2: $p = 0.0145$, 95% CI [−1.238, −0.2731]; lane 2 vs lane 4: $p = 0.0005$, 95% CI [0.7428, 1.107]. pCHK2/β-actin: lane 1 vs lane 2: $p = 0.0145$, 95% CI [−0.9435, −0.1945]; lane 2 vs lane 4: $p = 0.0384$, 95% CI [0.08776, 1.285]. p53/β-actin: lane 1 vs lane 2: $p = 0.0044$, 95% CI [−1.200, −0.6516]; lane 2 vs lane 4: $p = 0.0096$, 95% CI [0.4764, 1.182]. Panel **j** (WT vs KO): pDNA-PKcs/DNA-PKcs: lane 1 vs lane 2: $p = 0.0363$, 95% CI [−1.557, −0.1296]; lane 2 vs lane 4: $p = 0.0452$, 95% CI [−1.915, −0.05092]. pCHK1/CHK1: lane 1 vs lane 2: $p = 0.0271$, 95% CI [0.1147, 1.360]; lane 2 vs lane 4: $p = 0.0065$, 95% CI [−1.661, −0.4154]. p53/β-actin: lane 1 vs lane 2: $p = 0.0011$, 95% CI [−1.036, −0.7781]. Panel **l** (RT-PCR): DMSO vs 5-FU, WT: $p = 0.0031$, 95% CI [−322.2, −65.49]; DMSO vs 5-FU, WT::Tet*LOC644656*: $p = 0.0001$, 95% CI [−405.8, −149.2]. Panel **n** (Annexin V): DMSO vs ADR, WT: $p < 0.0001$, 95% CI [−89.22, −48.58]; DMSO vs ADR, WT::Tet*LOC644656*: $p < 0.0001$, 95% CI [−83.65, −43.01]; WT ADR vs WT::Tet*LOC644656* Dox + ADR: $p < 0.0001$, 95% CI [54.57, 92.58]; Dox − ADR vs Dox + ADR, WT::Tet*LOC644656*: $p < 0.0001$, 95% CI [46.70, 84.72].

expression (Fig. 6d). Of these, 147 genes overlapped with the 647 GSMD-related genes that were up- or downregulated in a p53-dependent manner. GO analysis of these 147 GSMD genes revealed mesenchymal stem cell-like characteristics (Fig. 6e), consistent with the upregulation of mesodermal genes in GSMD (Fig. 1d). Furthermore, these genes exhibited properties of CSCs (Fig. 6e) and were linked to various tumor types (Fig. 6f). Correlation matrix analysis revealed that specific clusters were associated with poor prognosis in distinct cancers (Fig. 6g and Supplementary Fig. 16a). Notably, high *LOC644656* expression was upregulated in advanced cancer stages (Supplementary Fig. 16b–m).

Pluripotency-associated genes, including *POU5F1* and *NANOG*, are highly expressed in hCSCs compared to normal tissues[52]. To validate the impact of *LOC644656* on cancer stemness, we analyzed its correlation with *POU5F1*. While *POU5F1* negatively correlated with *LOC644656* in normal liver and kidney tissues, positive correlations were observed in LIHC and KIRC tumors (Fig. 6h and Supplementary Fig. 17c; Fig. 6h: LIHC Tumor $p = 3.82e\text{-}08$). Experimental validation confirmed that *LOC644656* expression was regulated in a p53-dependent manner in various cancer cells (Supplementary Fig. 17a, b). Overexpression of *LOC644656* increased *POU5F1* and *NANOG* expression in SK-HEP1 and 786-O cells (Fig. 6i and Supplementary Fig. 17e, Fig. 6i: *LOC644656* $p = 0.0064$, 95% CI [0.7682, 2.512]; *POU5F1* $p = 0.0087$, 95% CI [0.2493, 1.334]; *NANOG* $p = 0.0354$, 95% CI [0.06535, 1.666]), while its knockdown decreased their expression (Fig. 6j, *LOC644656* $p = 0.0005$, 95% CI [−0.3779, −0.1488]; *POU5F1* $p = 0.0015$, 95% CI [−0.6490, −0.2510]) in SK-HEP1 cells.

Given that *LOC644656* activates TGF-β signaling in hESCs (Fig. 4), we hypothesized a similar role in cancer. Consistent with this, *LOC644656*-overexpressing LIHC and KIRC cells exhibited maldifferentiation (Supplementary Fig. 17d). Additionally, *TGFB1* and downstream target genes, such as *TAGLN* and *TWIST2*, were activated upon *LOC644656* overexpression (Fig. 6k: *TGFB1* $p = 0.0031$, 95% CI [0.7513, 3.024]; *TAGLN* $p < 0.0001$, 95% CI [2.058, 2.689]) and suppressed following antisense oligo (ASO)-mediated knockdown of *LOC644656* (Fig. 6l: *TGFB1* $p = 0.0262$, 95% CI [−0.6054, −0.05461]). Moreover, *TGFBR1* inhibitors significantly suppressed the induction of *POU5F1, NANOG*, and TGF-β signaling and its downstream targets induced by *LOC644656* overexpression (Fig. 6m, n, Fig. 6m *POU5F1*: $p < 0.0001$, 95% CI [−1.202, −0.7496]; Fig. 6n *TGFB1*: $p < 0.0001$, 95% CI [−2.877, −1.853]). *LOC644656* overexpression also enhanced cell migration in vitro (Supplementary Fig. 17g, $p < 0.0001$, 95% CI [−76.19, −37.57]). Collectively, these results indicate that *LOC644656* primarily drives a GSMD-like phenotype characterized by stemness-related features and TGF-β signaling, contributing to poor prognosis in cancer types such as LIHC and KIRC.

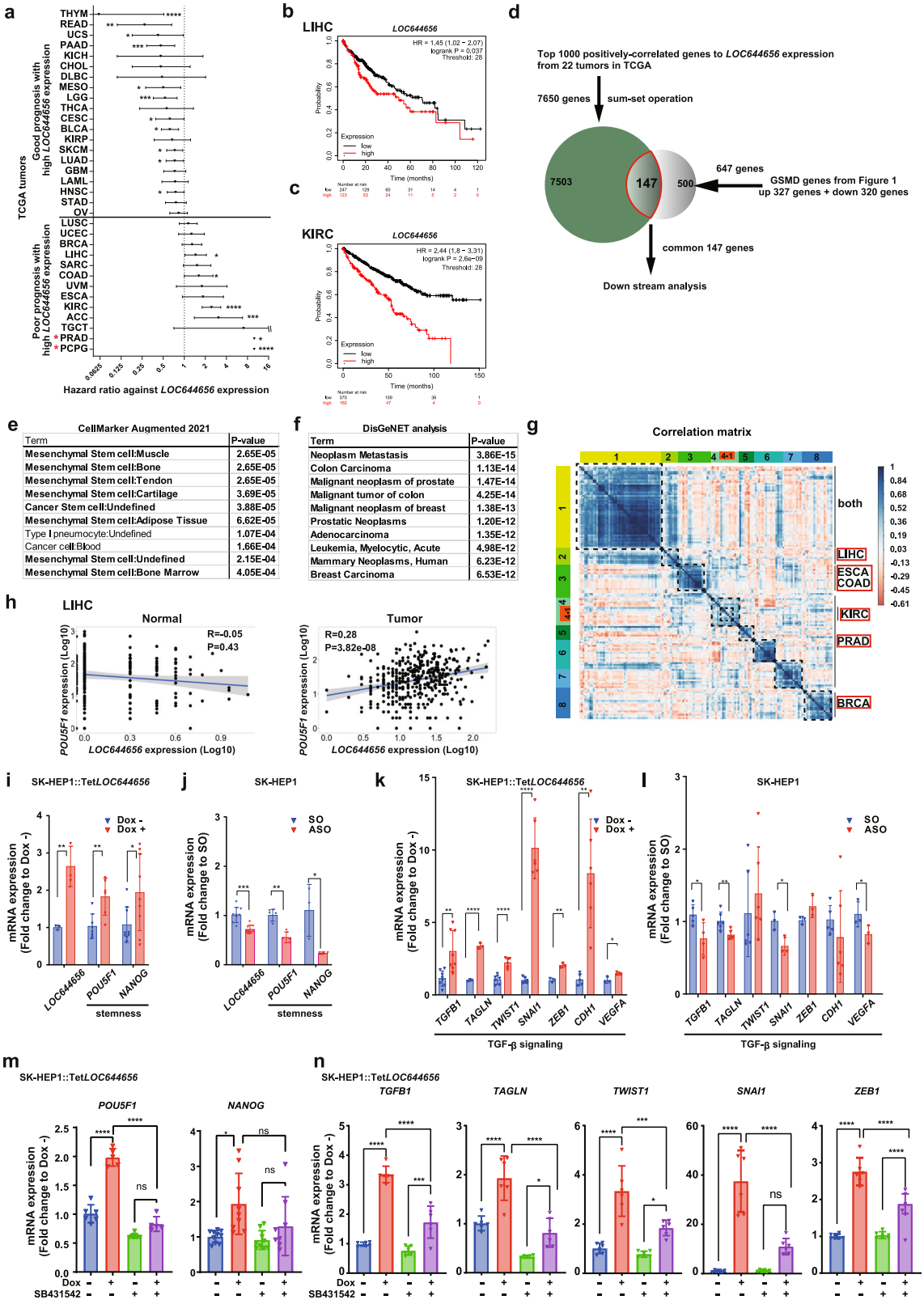
LOC644656 as a regulator of DDR signaling and chemoresistance in cancer models

To validate the role of *LOC644656* in suppressing DDR signaling and promoting chemoresistance, we generated multiple Tet*LOC644656*-inducible cancer cell lines, including HepG2 and SK-HEP1 cells (LIHC models), MCF-7 cells (BRCA model), and 786-O cells (KIRC model). LOC644656 induction significantly decreased cell death caused by 5-FU, ADR, and DNR in HepG2 cells, SK-HEP1 cells, MCF-7 breast cancer cells, and 786-O renal carcinoma cells (Fig. 7a–d and Supplementary Fig. 18a–g, Fig. 7a: $p < 0.0001$, 95% CI [−31.34, −11.03] (0 μM vs 200 μM)). Conversely, ASO-mediated knockdown of *LOC644656* enhanced the chemosensitivity of these cell lines to 5-FU (Fig. 7e–h, Fig. 7e: $p = 0.0002$, 95% CI [15.13, 55.41] (0 μM vs 100 μM)).

To elucidate the molecular mechanisms underlying this chemoresistance, we monitored DDR signaling in 5-FU-treated HepG2::Tet*LOC644656* cells. *LOC644656* induction reduced genotoxic stress-induced phosphorylation of DNA-PKcs and subsequent activation of γH2AX, consistent with meta-analyses of LIHC patient data (Fig. 7i and Supplementary Fig. 18j, Fig. 7i pDNA-PKcs/DNA-PKcs: $p = 0.0033$, 95% CI [−1.124, −0.6438] (lane 1 vs lane 2)). In contrast, *LOC644656* knockdown increased DDR signaling and p53 activation (Fig. 7j, pDNA-PKcs/β-actin: $p = 0.0166$, 95% CI [−1.256, −0.3476] (lane 1 vs lane 4), further corroborating the role of *LOC644656* in promoting chemoresistance in vitro.

Finally, we assessed the effect of *LOC644656* on chemoresistance in 3D culture and xenograft models. In 3D culture, 5-FU-induced cell death was significantly suppressed in MCF-7::Tet*LOC644656* cells following *LOC644656* induction (Fig. 7k, l, Fig. 7l: $p < 0.0001$, 95% CI [33.64, 51.16] (5-FU Dox - vs 5-FU Dox +)). In an ex vivo xenograft model, *LOC644656* expression in transplanted tumor cells did not affect tumor

https://doi.org/10.1038/s41467-025-59886-w



size but significantly increased resistance to chemotherapy-mediated genotoxic stress compared to control tumors (Fig. 7m, n, Fig. 7n: $p < 0.0001$, 95% CI [1024, 2292] (Dox - /5-FU - vs Dox -/5-FU +); $p = 0.0017$, 95% CI [−1601, −332.8] (Dox - /5-FU + vs Dox +/5-FU +)).

Collectively, these findings establish *LOC644656* as a key mediator of chemoresistance in cancer cells, functioning through the suppression of DNA-PKcs signaling in response to genotoxic stress.

Discussion

Our findings highlight the pivotal role of the p53-induced lncRNA *LOC644656* in regulating pluripotency and differentiation in hESCs. By directly inhibiting the POU5F1 complex and the KDM1A/LSD1-NuRD repressor complex, *LOC644656* activates TGF-β/SMAD signaling, driving fibroblast-like differentiation, a process we term GSMD. In cancer, however, *LOC644656* downregulates genes involved in the genotoxic

Fig. 6 | *LOC644656* is highly expressed in tumors and primarily correlates with poor prognosis. a Kaplan–Meier survival analysis comparing high versus low *LOC644656* expression across multiple TCGA tumor types (n = 80–500 patients per tumor type). Hazard ratios (HRs) and log-rank p-values are shown. Red asterisks indicate reference HR values in smaller cohorts. **b, c** Kaplan–Meier survival plots for liver hepatocellular carcinoma (LIHC, b) and kidney renal clear cell carcinoma (KIRC, c). HRs and p-values determined by two-sided log-rank tests using Kaplan–Meier Plotter. **d** Identification of 7650 common genes correlated with *LOC644656* expression across 22 tumor types. The Venn diagram highlights 147 genes that overlap with the 647 GSMD-related genes identified in Fig. 1f. **e, f** Gene Ontology analysis of the 147 overlapping genes using CellMarker Augmented 2021 (**e**) and DisGeNET (**f**). Statistical significance determined by two-sided Fisher's exact test with Benjamini-Hochberg correction. **g** Correlation matrix showing tumor/normal fold change relationships among the 147 genes. Red boxes indicate gene clusters associated with specific tumor types. **h** Pearson correlation analysis of *LOC644656* and *POU5F1* expression in LIHC tumor samples from TCGA database. **i, j** Real-time RT-PCR analysis of *LOC644656*, *POU5F1*, and *NANOG* expression in SK-HEP1::Tet*LOC644656* cells ± Dox (i), or SK-HEP1 cells transfected with sense oligonucleotides (SO) or antisense oligonucleotides (ASO) (**j**). Data are mean ± SEM (n = 3). Statistical analysis by two-sided Student's t-test. **k, l** Expression analysis of *TGFB1* and its downstream targets in SK-HEP1::Tet*LOC644656* cells ± Dox (**k**) or SK-HEP1 cells ± *LOC644656*-ASO (**l**). Data are mean ± SEM (n = 3). Statistical analysis by two-sided Student's t-test. **m, n** Effect of TGFBR1 inhibitor (SB431542) treatment on SK-HEP1::Tet*LOC644656* cells ± Dox for 72 h. Expression of pluripotency markers (**m**) and TGF-β pathway genes (**n**) were analyzed by RT-PCR. Data are mean ± SEM (n = 3). Statistical analysis by two-sided one-way ANOVA with Tukey's post hoc test. Exact p-values and 95% confidence intervals**:** Panel **a** (survival analysis): BLCA: p = 0.00084; HNSC: p = 0.043; KIRC: p = 2.6e-09; LIHC: p = 0.037; LUAD: p = 0.035; PAAD: p = 3e-04; PCPG: p = 3.1e-05; READ: p = 0.0024; STAD: p = 0.043; THYM: p = 0.00052. Panel **h**: LIHC Tumor: p = 3.82e-08. Panel **i**: *LOC644656*: p = 0.0064, 95% CI [0.7682, 2.512] (Dox − vs Dox +). *POU5F1*: p = 0.0087, 95% CI [0.2493, 1.334] (Dox − vs Dox +). *NANOG*: p = 0.0354, 95% CI [0.06535, 1.666] (Dox − vs Dox +). Panel **j**: *LOC644656*: p = 0.0005, 95% CI [−0.3779, −0.1488] (SO vs ASO). *POU5F1*: p = 0.0015, 95% CI [−0.6490, −0.2510] (SO vs ASO). *NANOG*: p = 0.0359, 95% CI [−1.520, −0.08622] (SO vs ASO). Panel **k**: *TGFB1*: p = 0.0031, 95% CI [0.7513, 3.024] (Dox − vs Dox +). *TAGLN*: p < 0.0001, 95% CI [2.058, 2.689] (Dox − vs Dox +). *TWIST1*: p < 0.0001, 95% CI [0.7678, 1.605] (Dox − vs Dox +). *SNAI1*: p < 0.0001, 95% CI [7.195, 11.01] (Dox − vs Dox +). *ZEB1*: p = 0.0012, 95% CI [0.6746, 1.372] (Dox − vs Dox +). *CDH1*: p = 0.002, 95% CI [3.474, 11.17] (Dox − vs Dox +). *VEGFA*: p = 0.0286, 95% CI [0.07875, 0.8412] (Dox − vs Dox +). Panel **l**: *TGFB1*: p = 0.0262, 95% CI [−0.6054, −0.05461] (SO vs ASO). *TAGLN*: p = 0.0071, 95% CI [−0.3183, −0.06505] (SO vs ASO). *SNAI1*: p = 0.0261, 95% CI [−0.6258, −0.06748] (SO vs ASO). *VEGFA*: p = 0.0449, 95% CI [−0.7579, −0.01206] (SO vs ASO). Panel **m**: *POU5F1*: p < 0.0001, 95% CI [−1.202, −0.7496] (Dox − vs Dox +). p < 0.0001, 95% CI [0.9179, 1.390] (Dox + vs Dox +/SB431542 +). *NANOG*: p < 0.0001, 95% CI [−2.433, −0.6762] (Dox − vs Dox +). Panel **n**: *TGFB1*: p < 0.0001, 95% CI [−2.877, −1.853] (Dox − vs Dox +). p < 0.0001, 95% CI [1.108, 2.132] (Dox + vs Dox +/SB431542 +). p = 0.0002, 95% CI [−1.475, −0.4514] (SB431542 + vs Dox +/SB431542 +). *TAGLN*: p < 0.0001, 95% CI [−1.373, −0.4674] (Dox − vs Dox +). p < 0.0001, 95% CI [0.6591, 1.564] (Dox + vs Dox +/SB431542 +). p = 0.0302, 95% CI [−0.9443, −0.03908] (SB431542 + vs Dox +/SB431542 +). *TWIST1*: p < 0.0001, 95% CI [−3.209, −1.434] (Dox − vs Dox +). p = 0.0007, 95% CI [0.6062, 2.381] (Dox + vs Dox +/SB431542 +). p = 0.0152, 95% CI [−1.952, −0.1778] (SB431542 + vs Dox +/SB431542 +). *SNAI1*: p < 0.0001, 95% CI [−46.84, −25.98] (Dox − vs Dox +). p < 0.0001, 95% CI [16.00, 36.86] (Dox + vs Dox +/SB431542 +). *ZEB1*: p < 0.0001, 95% CI [−2.140, −1.370] (Dox − vs Dox +). p < 0.0001, 95% CI [0.4885, 1.258] (Dox + vs Dox +/SB431542 +). p < 0.0001, 95% CI [−1.235, −0.4652] (SB431542 + vs Dox +/SB431542 +).

stress-sensing machinery, contributing to chemoresistance in KIRC, LIHC, and BRCA models. Notably, elevated *LOC644656* expression correlates with poor overall survival in these cancers and promotes chemoresistance in vitro.

These findings align with the dual role of TGF-β signaling in cellular differentiation and tumor progression, where it can act as both a tumor suppressor and promoter depending on the cellular context[48]. This duality is mediated by context-dependent transcriptional networks and epigenetic modifiers[53]. Furthermore, our observations that *LOC644656* suppresses DDR signaling and enhances chemoresistance are consistent with studies demonstrating that genotoxic stress responses, such as those induced by 5-FU, depend on mismatch repair and TP53 status[35]. However, the molecular mechanisms underlying the suppression of pluripotency by *LOC644656* in hESCs and its promotion of pluripotency factors such as POU5F1 and NANOG in cancer cells are complex and highly context-dependent.

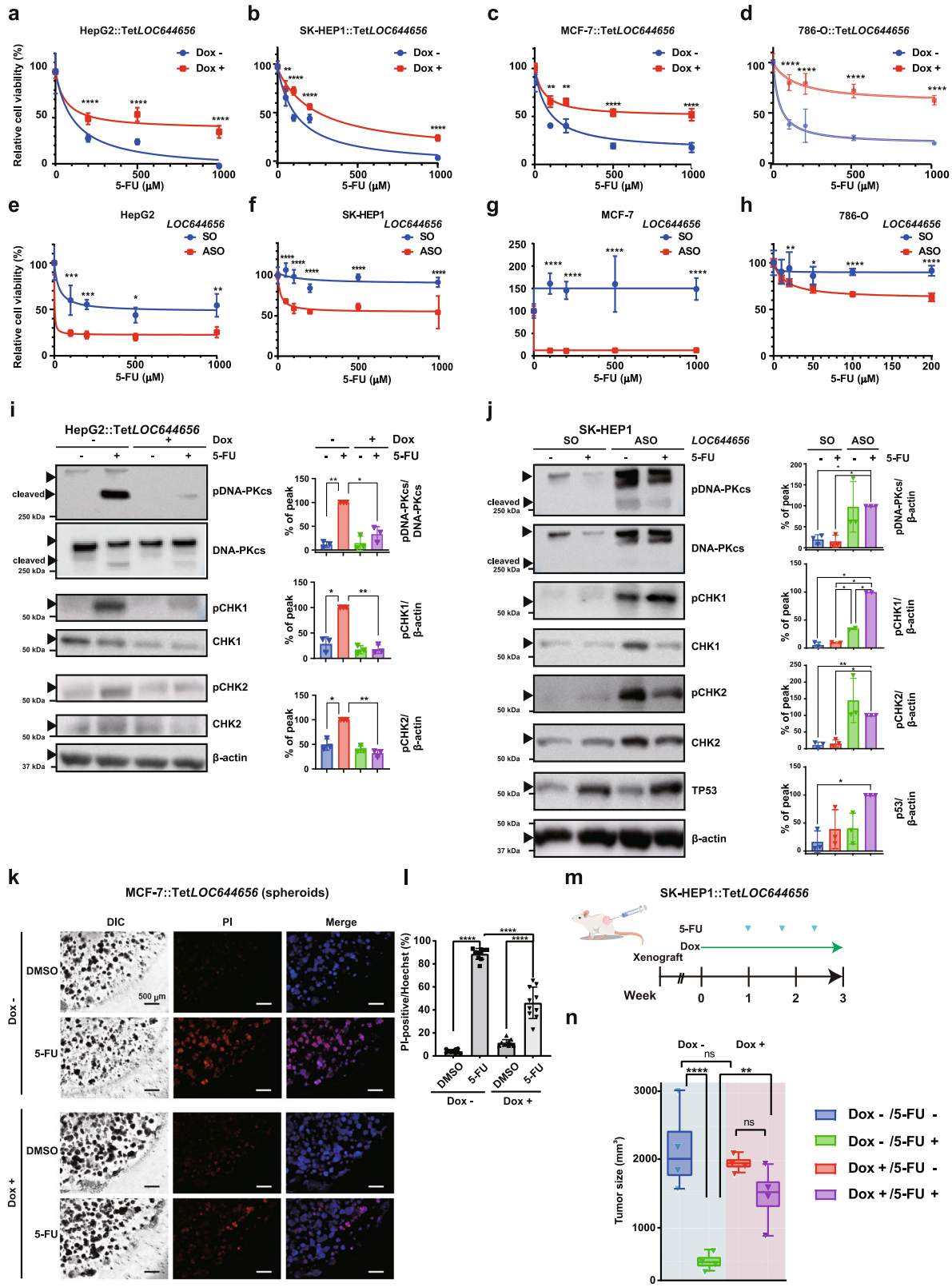
In hESCs, *LOC644656* directly interacts with pluripotency factor complexes, including POU5F1, inhibiting its self-activation and promoting differentiation through the modulation of TGF-β signaling. Conversely, in tumor cells, particularly SK-Hep cells, where the baseline expression of POU5F1 is approximately 100-fold lower than in hESCs, *LOC644656* activates TGF-β signaling. This activation induces the expression of pluripotency factors, including POU5F1 and NANOG, promoting stemness and drug resistance. POU5F1, a well-established marker of stemness, is implicated in chemotherapy resistance in various cancers[54], further supporting our findings. These results suggest that the cellular context, particularly the relative abundance of target proteins, dictates the functional outcome of *LOC644656* activation. While transcriptional modulation of TGF-β signaling appears to be the primary mechanism, potential post-translational effects cannot be completely excluded.

The role of *LOC644656* in promoting differentiation and preventing apoptosis in hESCs aligns with p53-mediated regulation of these processes, as previously reported[55]. However, the divergent roles of *LOC644656* also highlight the intricate interplay between TGF-β signaling and cell-specific transcriptional networks. In hESCs, TGF-β signaling shifts the transcriptional program toward differentiation by upregulating mesodermal genes and suppressing pluripotency factors. In cancer cells, however, TGF-β signaling often cooperates with oncogenic pathways, such as those involving MYC and β-catenin, to enhance stemness and EMT. This observation is consistent with TGF-β's role in modulating EMT and enhancing metastatic potential[56,57]. Additionally, TGF-β-induced EMT is closely linked to VEGFC/VEGFR3 axis activation, further promoting cancer progression[58]. These context-dependent effects may involve the differential recruitment of transcriptional cofactors or chromatin modifiers in response to *LOC644656*, resulting in cell-type-specific gene expression outcomes. Further investigation into the epigenetic landscape and cofactor interactions in hESCs versus cancer cells could provide deeper insights into these mechanisms.

LOC644656 also functions as a safeguard against genotoxic stress in hESCs by driving differentiation and preventing apoptosis. This protective mechanism likely involves reducing pluripotency, inducing maldifferentiation, and attenuating DDR signaling, thereby preserving the integrity of undamaged hESCs. Notably, the nuclear localization of *LOC644656* strongly activates the SMAD/TGF-β pathway, which is critical for somatic stem cell maintenance and differentiation. This activation mimics the fibroblast-like differentiation observed in GSMD, where SMAD2/3 dissociation from repressor complexes promotes mesodermal differentiation[53]. Additionally, TGF-β1-mediated fibroblast-to-cancer-associated fibroblast (CAF) transition facilitates metastasis, as reported in gastrointestinal stromal tumors[57]. Thus, *LOC644656*-induced TGF-β expression may exacerbate EMT and negatively impact overall survival.

Paradoxically, while *LOC644656* suppresses pluripotency in hESCs, it increases the expression of *POU5F1* and *NANOG* in cancers. This divergence underscores the dual role of TGF-β signaling, acting as both a tumor suppressor and a tumor promoter depending on cellular context[59]. In cancer, transcription factors such as TWIST and SNAI, along with SNAI2-mediated regulation of NANOG and MYC, may amplify TGF-β-induced EMT and chemoresistance[60]. Furthermore, the plasticity of cancer stem cells may allow TGF-β signaling to promote a

https://doi.org/10.1038/s41467-025-59886-w



stem-like state, even in the presence of a generally differentiation-promoting environment.

The dual role of p53 in tumorigenesis adds complexity to *LOC644656*'s function. While p53 typically suppresses tumors, wild-type p53 may promote chemoresistance by inducing *LOC644656*, as evidenced by its elevated expression in tumor tissues compared to normal tissues. This complexity aligns with the dynamic and tissue-specific nature of p53 signaling[61–63]. *LOC644656* appears to suppress DDR signaling, including DNA-PKcs activity, and mitigate p53-mediated apoptosis, thereby contributing to cancer cell survival under genotoxic stress.

Clinically, *LOC644656* is significantly associated with poor prognosis in KIRC, LIHC, and PRAD. Its expression could serve as a biomarker for predicting therapeutic responses and overall survival in

Fig. 7 | *LOC644656* expression prevents genotoxic stress-induced death of cancer cells. a–d Cell viability (CCK-8) assays for HepG2::Tet*LOC644656* (**a**), SK-HEP1::Tet*LOC644656* (**b**), MCF-7::Tet*LOC644656* (**c**), and 786-O::Tet*LOC644656* (**d**) cells ± Dox for 1 day, followed by 5-FU treatment for 24–72 h. Data are mean ± SEM ($n = 3$). Statistical analysis by two-sided two-way ANOVA with Tukey's post hoc test. **e–h** Dose-response curves for HepG2 (**e**), SK-HEP1 (**f**), MCF-7 (**g**), and 786-O (**h**) cells transfected with *LOC644656* sense oligonucleotides (SO) or antisense oligonucleotides (ASO) for 48 h, then treated with 5-FU for 24 h. Data are mean ± SEM ($n = 3$). Statistical analysis by two-sided two-way ANOVA with Tukey's post hoc test. **i**, Immunoblot analysis of DNA damage response proteins in HepG2::Tet*LOC644656* cells ± Dox, then 5-FU. Data are mean ± SEM ($n = 3$). Statistical analysis by two-sided one-way ANOVA with Tukey's post hoc test. The samples derive from the same experiment but different gels for each antibody were processed in parallel. Uncropped blots are provided in the Source Data file. **j** Immunoblot analysis of genotoxic stress-sensing proteins in SK-HEP1 cells transfected with SO or ASO for 48 h, then 5-FU for 12 h. Data are mean ± SEM ($n = 3$). Statistical analysis by two-sided one-way ANOVA with Tukey's post hoc test. The samples derive from the same experiment but different gels for each antibody were processed in parallel. Uncropped blots are provided in the Source Data file. **k, l** 3D spheroid assay of MCF-7::Tet*LOC644656* cells ± Dox for 24 h, then 5-FU for 24 h. Spheroids were stained with propidium iodide (PI)/Hoechst. Data are mean ± SEM ($n = 3$). Statistical analysis by two-sided one-way ANOVA with Tukey's post hoc test. **m** Schematic of the mouse xenograft experimental design. Cells were randomly allocated into four groups, and transplantation was independently performed by two investigators to minimize allocation bias. **n** Tumor size measurements in each treatment group. Data are mean ± SEM ($n = 3$). Statistical analysis by two-sided one-way ANOVA with Tukey's post hoc test. Exact p-values and 95% confidence intervals: Panels **a–d** (viability assays): Panel a (HepG2): 0 vs 200 μM: $p < 0.0001$, 95% CI [−31.34, −11.03]; 0 vs 500 μM: $p < 0.0001$, 95% CI [−39.94, −19.62]; 0 vs 1000 μM: $p < 0.0001$, 95% CI [−47.23, −26.91]. Panel **b** (SK-HEP1): 0 vs 50 μM: $p = 0.0039$, 95% CI [−15.63, −2.230]; 0 vs 100 μM: $p < 0.0001$, 95% CI [−34.60, −21.20]; 0 vs 200 μM: $p < 0.0001$, 95% CI [−19.27, −5.406]; 0 vs 1000 μM: $p < 0.0001$, 95% CI [−27.68, −13.35]. Panel **c** (MCF-7): 0 vs 100 μM: $p = 0.0014$, 95% CI [−28.93, −5.778]; 0 vs 200 μM: $p = 0.0012$, 95% CI [−29.14, −5.985]; 0 vs 500 μM: $p < 0.0001$, 95% CI [−42.49, −19.34]; 0 vs 1000 μM: $p < 0.0001$, 95% CI [−42.34, −19.19]. Panel d (786-O): 0 vs 100 μM: $p < 0.0001$, 95% CI [−53.90, −29.02]; 0 vs 200 μM: $p < 0.0001$, 95% CI [−54.24, −29.36]; 0 vs 500 μM: $p < 0.0001$, 95% CI [−59.06, −34.18]; 0 vs 1000 μM: $p < 0.0001$, 95% CI [−55.18, −30.30]. Panels **e–h** (ASO effects): Panel e (HepG2): 0 vs 100 μM: $p = 0.0002$, 95% CI [15.13, 55.41]; 0 vs 200 μM: $p = 0.0005$, 95% CI [12.64, 52.92]; 0 vs 500 μM: $p = 0.0149$, 95% CI [3.588, 43.87]; 0 vs 1000 μM: $p = 0.0023$, 95% CI [8.777, 49.06]. Panel **f** (SK-HEP1): 0 vs 50 μM: $p < 0.0001$, 95% CI [22.42, 54.76]; 0 vs 100 μM: $p < 0.0001$, 95% CI [23.37, 55.72]; 0 vs 200 μM: $p < 0.0001$, 95% CI [13.00, 45.35]; 0 vs 500 μM: $p < 0.0001$, 95% CI [20.24, 52.38]; 0 vs 1000 μM: $p < 0.0001$, 95% CI [21.00, 53.34]. Panel **g** (MCF-7): 0 vs 100 μM: $p < 0.0001$, 95% CI [102.6, 195.7]; 0 vs 200 μM: $p < 0.0001$, 95% CI [88.06, 181.1]; 0 vs 500 μM: $p < 0.0001$, 95% CI [101.2, 194.3]; 0 vs 1000 μM: $p < 0.0001$, 95% CI [89.61, 182.7]. Panel **h** (786-O): 0 vs 20 μM: $p = 0.0038$, 95% CI [3.730, 27.16]; 0 vs 50 μM: $p = 0.0132$, 95% CI [2.018, 25.45]; 0 vs 100 μM: $p < 0.0001$, 95% CI [12.09, 35.52]; 0 vs 200 μM: $p < 0.0001$, 95% CI [16.17, 41.21]. Panel **i** (protein ratios): pDNA-PKcs/DNA-PKcs: lane 1 vs lane 2: $p = 0.0033$, 95% CI [−1.124, −0.6438]; lane 2 vs lane 4: $p = 0.0446$, 95% CI [0.03890, 1.297]. pCHK1/β-actin: lane 1 vs lane 2: $p = 0.032$, 95% CI [−1.288, −0.1482]; lane 2 vs lane 4: $p = 0.0092$, 95% CI [0.4796, 1.159]. pCHK2/β-actin: lane 1 vs lane 2: $p = 0.0419$, 95% CI [−0.9614, −0.04393]; lane 2 vs lane 4: $p = 0.0087$, 95% CI [0.4070, 0.9570]. Panel **j** (ASO effects on signaling): pDNA-PKcs/β-actin: lane 1 vs lane 4: $p = 0.0166$, 95% CI [−1.256, −0.3476]; lane 2 vs lane 4: $p = 0.0264$, 95% CI [−1.466, −0.2393]. pCHK1/β-actin: lane 1 vs lane 4: $p = 0.0407$, 95% CI [−1.771, −0.1221]; lane 2 vs lane 3: $p = 0.0423$, 95% CI [−0.4667, −0.02649]; lane 2 vs lane 4: $p = 0.0211$, 95% CI [−0.9051, −0.8952]; lane 3 vs lane 4: $p = 0.0211$, 95% CI [−0.8687, −0.4384]. pCHK2/β-actin: lane 1 vs lane 4: $p = 0.0082$, 95% CI [−1.233, −0.5412]; lane 2 vs lane 4: $p = 0.0139$, 95% CI [−1.282, −0.4078]. p53/β-actin: lane 1 vs lane 4: $p = 0.044$, 95% CI [−1.620, −0.05384]. Panel **l** (3D spheroid): DMSO Dox - vs 5-FU Dox -: $p < 0.0001$, 95% CI [−92.97, −75.85]; 5-FU Dox - vs 5-FU Dox +: $p < 0.0001$, 95% CI [33.64, 51.16]; DMSO Dox + vs 5-FU Dox +: $p < 0.0001$, 95% CI [−43.36, −25.84]. Panel **n** (xenograft): Dox - /5-FU - vs Dox -/5-FU +: $p < 0.0001$, 95% CI [1024, 2292]; Dox - /5-FU + vs Dox +/5-FU +: $p = 0.0017$, 95% CI [−1601, −332.8].

these cancers. Furthermore, its role in attenuating p53's protective effects against genotoxic stress highlights its potential as a therapeutic target. Antisense oligonucleotides (ASOs) targeting *LOC644656* may offer a strategy to overcome chemoresistance, particularly in patients with KIRC or LIHC.

Methods

Ethics statement

The study was conducted in compliance with the ARRIVE 2.0 guidelines and regulations for the use of animals in research. Animal experiments were approved by the ethics committee at Chiba University (registration number: A5-161). The use of human embryonic stem (hES) cells was approved by the Life Science and Medical Ethics Review Committee of the Graduate School of Medicine, Chiba University (approval ID: Chiba University Inohana Research No. 710). hESCs were obtained from Riken BioResource Center (Japan) with appropriate approvals and consent.

Cell lines and cell culture

hESCs, namely hESC1 (KhES1) and hESC3 (KhES3), were obtained from Riken BioResource Center (Japan) and maintained on hESC-qualified Matrigel (BD Biosciences, USA) in mTeSR1 medium (STEMCELL Technologies, USA) or on iMatrix-511 (Nippi)-coated plates in StemFit AK02N medium (Ajinomoto, Tokyo, Japan) in 5 % CO_2 at 37 °C following the manufacturer's instructions. On the days of passaging, the Rho kinase inhibitor Y-27632 (10 μM, Cayman Chemical/Funakoshi, Tokyo, Japan) was added to the medium. MCF-7, HepG2, and 786-O cells were obtained from the American Type Culture Collection (ATCC) (Funakoshi, Tokyo, Japan). MCF-7 and HepG2 cells were maintained in Dulbecco's modified Eagle medium (DMEM; Sigma-Aldrich, Tokyo, Japan) supplemented with 10% fetal bovine serum (FBS, Gibco, Tokyo, Japan) and 1% penicillin/streptomycin (Sigma-Aldrich); 786-O cells were maintained in Roswell Park Memorial Institute (RPMI-1640) medium (Sigma-Aldrich) supplemented with 10% FBS and 1% penicillin/streptomycin. The cell lines were routinely tested for and confirmed to be free of *Mycoplasma* contamination using a MycoAlert assay (performed as recommended by the manufacturer; Lonza, Tokyo, Japan).

Mice

Six-week-old, immunodeficient female mice (NOD.Cg-Prkdcscid/Il2rgtm1Wjl/SzJ; also known as NSG) were obtained from Oriental Yeast Co., Ltd (Japan). Animals were maintained in specific pathogen-free conditions in a temperature-controlled room (22 ± 2 °C) with a 12-h/12-h light/dark cycle and relative humidity of 40-70%. The mice were provided food and water ad libitum. All animal experiments were approved by the Institutional Animal Care and Use Committee at Chiba University (Protocol #: A5-161) and conducted in accordance with the ARRIVE 2.0 guidelines and institutional regulations for animal care and use.

Tumor growth was monitored three times per week by measuring tumor dimensions with calipers. Tumor volume was calculated using the formula: 4π(length × width × height)/3. Animal welfare was assessed during each measurement using a standardized scoring system. In accordance with our approved protocol, mice were euthanized when any of the following predefined humane endpoints were reached. Tumor-related criteria for euthanasia included tumor volume exceeding 2000 mm^3, presence of ulceration, necrosis, or infection at the tumor site, or tumors that interfered with vital physiological functions such as mobility, feeding, or elimination. Clinical criteria necessitating euthanasia comprised body weight loss ≥20% from baseline, persistent signs of distress (including hunching posture and piloerection), severe lethargy or unresponsiveness to external stimuli, respiratory distress manifested by labored breathing or cyanosis, significant deterioration in grooming behavior, and self-isolation or

abnormal behavior patterns. Animals were monitored daily for these signs by trained personnel. In three cases where tumor size exceeded 2000 mm^3 between scheduled measurements, mice were immediately euthanized. All euthanasia was performed by cervical dislocation under isoflurane anesthesia by trained personnel, in accordance with the American Veterinary Medical Association (AVMA) Guidelines for the Euthanasia of Animals (2020 Edition). Death was confirmed by cessation of vital signs before tissue collection.

CRISPR/Cas9 system

To generate the hESC1 p53KO and LOC644656KO cell lines, the backbone vectors pX459 pSpCas9 (BB)−2A-Puro and pX458 pSpCas9 (BB)−2A-Puro were obtained from Addgene (MA, USA). Target guide RNA sequences were designed based on exon 2 of *TP53* and three sites in *LOC644656* (Supplemental Data 1). For insertion of the *TP53* target sequence into pX459 and the *LOC644656* target sequence into pX458, the oligonucleotides listed in Supplementary Data 1 were obtained from Eurofins Genomics (Tokyo, Japan). The oligonucleotides were annealed with T4 DNA Ligase Reaction Buffer and phosphorylated with T4 Polynucleotide Kinase (both from New England Biolabs, Tokyo, Japan). The pX458 and pX459 vectors were digested using *Bbs*1, purified using gel electrophoresis, and ligated to the oligonucleotides using a Quick Ligation Kit (New England Biolabs). Stbl3 cells were transformed with the constructed plasmids and cultured; the amplified plasmid was then collected using NucleoBond Xtra Midi (Takara, Shiga, Japan). A plasmid targeting the *TP53* locus was transfected into hESC1 cells using the Neon™ transfection system (Thermo Fisher Scientific, Tokyo, Japan) according to the manufacturer's protocol. Subsequently, cells were selected using puromycin (Nacalai, Kyoto, Japan), and p53KO clones were selected. Independent cell lines were established from at least two independent clones.

Construction of lncRNA expression plasmids

The backbone vector Xlone-GFP was purchased from Addgene (plasmid #96930, MA, USA). The GFP sequence was replaced with the *LOC644656* gene (63 to 1143 bp) via an IN-FUSION reaction using the In-Fusion Snap Assembly Master Mix (Takara). The blasticidin resistance gene from Xlone-GFP was replaced with the GFP-tagged puromycin resistance gene via a P2A signal peptide. The first-strand cDNA prepared from hESCs was synthesized via the same protocol used for conventional PCR to amplify *LOC644656*. The full-length *LOC644656* PCR product was inserted into the pGEM-T Easy Vector (Promega), which was subsequently digested using *Not*I for subcloning into the pcDNA3.1(+) vector in the sense and antisense orientations to construct the RNA-FISH probes.

Generation of hESC1::Tet*LOC644656* cell lines

The Xlone-Tet*LOC644656* plasmid was transfected into hESC1 cells with the PiggyBac transposon vector (System Bioscience/Funakoshi) using Lipofectamine Stem reagent (Thermo Fisher Scientific) following the method described by Qiu et al.[64]. Cells were then selected using puromycin (Nacalai), and hESC::Tet*LOC644656* clones were selected. Cell lines were established from at least two independent clones.

Real-time RT-PCR analysis and primer information

The expression levels of gene-specific mRNAs were measured relative to *ACTB*, *18S*, or *L32* as internal controls, as described previously[65]. Briefly, total RNA was extracted from hESCs using an Rneasy Kit (Qiagen, Tokyo, Japan) or from 786-O, HepG2, SK-HEP1, and MCF-7 cells with FastGene (Nippon Genetics, Tokyo, Japan). RNA quality was measured using a NanoDrop spectrophotometer (Thermo Fisher Scientific). Subsequently, 1 µg of RNA was reverse transcribed using a ReverTra Ace qPCR RT kit (Toyobo, Osaka, Japan) and specific primers (Supplementary Data 2). cDNA products (10 ng) were subjected to real-time RT-PCR using the Step One Plus Real-Time PCR System (Applied Biosystems, Tokyo, Japan). Relative expression was measured with ΔΔCT method using *ACTB*, *18S*, and *RPL32* genes mentioned in specific results.

Endogenous coimmunoprecipitation

hESCs were plated in a 10-cm dish. At 70-80% confluence, the cells were treated with 1 µg/mL Dox for two days. The cells were washed with PBS, collected into a 1.5 mL tube, incubated with 1 mL of swelling buffer (25 mM HEPES [pH 7.9], 1.5 mM $MgCl_2$, 10 mM KCl, 0.1% NP-40, and protease inhibitor cocktail) on ice for 20 min with pipetting, and collected using centrifugation (700 × g, 4 °C, 5 min). The nuclei pellets were resuspended in coimmunoprecipitation buffer (20 mM HEPES [pH 7.9], 100 mM KCl, 1 mM EDTA, 0.1% NP-40, and protease inhibitor cocktail) and sonicated using a QSONICA Q700 (Waken Btech, Shiga, Japan) in cold water with 15 s on/45 s off cycles and total sonication energy of 1000 J. The sonicated chromatin was clarified using two rounds of centrifugation at 14,000 × g for 20 min each, and the supernatant was subjected to immunoprecipitation with anti-FLAG (for coimmunoprecipitation, Sigma-Aldrich) or anti-SMAD2 antibodies (for endogenous coimmunoprecipitation, Cell Signaling Technology, Tokyo, Japan)/Protein G Dynabeads (Veritus, Tokyo, Japan) over night at 4 °C on a rotating wheel. The magnetic beads were washed with coimmunoprecipitation buffer four times using a magnetic stand (Promega, Tokyo, Japan), and SDS sample buffer (CosmoBio, Tokyo, Japan) was added. The sample buffer/beads were boiled for 5 min to elute binding proteins, and the proteins were subjected to western blot analysis.

Immunoblot analysis and antibodies

Western blot analysis was performed as previously described[66]. Briefly, cell pellets were lysed with RIPA buffer (50 mM Tris-HCl, 300 mM KCl, 1% NP-40, 0.5% sodium deoxycholate, and 1× protease inhibitor cocktail; Nacalai) on ice for 30 min. After centrifugation (14,000 × g, 20 min, 4 °C), the protein concentration in the lysates was measured, and SDS sample buffer (CosmoBio) was added. The following primary antibodies were used: anti-p53, clone DO1 (Santa Cruz Biotechnology/Nacalai Tesque), anti-OCT3/4 (POU5F1) (Santa Cruz Biotechnology/Nacalai Tesque), anti-NANOG (Abcam, Tokyo, Japan), and anti-actin (Sigma-Aldrich). The details of antibodies were listed in Supplementary Data 3.

RNA-seq and data analysis

Total RNA was extracted from hESCs using an Rneasy Plus Mini Kit or an Rneasy Plus Micro Kit (Qiagen). RNA concentration and integrity were measured using an Agilent 2100 bioanalyzer. Library construction and sequencing were performed as previously described[67]. The TopHat program (version 1.3.2; https://ccb.jhu.edu/software/tophat/index.shtml) was used with default parameters to align the reads to the hg19 human reference genome. Subsequently, gene expression values were calculated as FPKM read values using Cufflinks (version 2.0.2). DEGs were identified using R (TCC library)/Bioconductor software (version 4.0.4; FDR < 0.05). A heatmap and dendrogram were also generated using R software (gplots).

ChIP-seq

ChIP-seq analysis of treated hESCs was performed as previously described, with slight modifications[68]. Specifically, hESCs were plated in a 10-cm dish. At 70-80% confluence, the cells were treated with 1 mM 5-FU for 24 h, 0.5 µM ADR for 8 h, or 0.5 µM DNR for 8 h. Cells were fixed using 1% methanol-free formaldehyde (Thermo Fisher Scientific) for 10 min at 37 °C, after which the reaction was quenched with 0.125 M glycine for 5 min at room temperature (approximately 25 °C). The cells were washed thrice with ice-cold PBS. The cells were then incubated with swelling buffer (25 mM HEPES [pH 7.9], 1.5 mM $MgCl_2$, 10 mM KCl, 0.1% NP-40, and protease inhibitor cocktail) on ice for 10 min,

harvested by scraping, and collected via centrifugation (5000 × *g*, 4 °C, 5 min). Cell pellets were resuspended in sonication buffer (50 mM HEPES [pH 7.9], 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate, 0.1% SDS, and protease inhibitor cocktail) and sonicated using a QSONICA Q700 (Waken Btech) in cold water with 15 s on/45 s off cycles and total sonication energy of 6000 J. The sonicated chromatin was clarified using two rounds of centrifugation at 16,000 × *g* for 5 min each, and the supernatant was then precleared with 20 μL of protein G magnetic beads (Dynabeads, Veritas, Tokyo, Japan) for 2 h at 4 °C with rotation. The supernatants were collected using a magnetic stand (Promega) and subjected to immunoprecipitation with anti-p53 antibodies (mixture of anti-p53 monoclonal antibodies from clones DO-1 and 1801) for 24 h at 4 °C on a rotating wheel. The magnetic beads were washed with 1 mL of each of the following wash buffers for 5 min at 4 °C: sonication buffer (one wash), wash buffer A (50 mM HEPES [pH 7.9], 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Na-deoxycholate and 0.1% SDS; one wash), wash buffer B (20 mM Tris-HCl [pH 8.0], 250 mM LiCl, 0.5% NP-40 and 0.5% Na-deoxycholate; one wash) and TE buffer (two washes). To reverse the crosslinking and release the captured DNA, beads were added to freshly prepared elution buffer (50 mM Tris-HCl [pH 7.5], 1 mM EDTA, and 1% SDS), incubated at 65 °C for 5 min, and mixed at 1200 rpm for 15 min at room temperature. Isolated DNA fragments were incubated with proteinase K (Wako, Osaka, Japan) at 45 °C for 2 h and recovered using a PCR cleanup kit (Favorgen/Chiyoda Science, Tokyo, Japan). The DNA content was then quantified using fluorimetry with a Qubit fluorometer (Thermo Fisher Scientific), and 1 ng of enriched fragmented DNA was used to prepare libraries for next-generation sequencing. ChIP-seq libraries were constructed using a KAPA ChIP-Seq Library Preparation Kit (Nippon Genetics) according to the manufacturer's instructions. Deep sequencing was performed on the Illumina NextSeq 500 platform using a TruSeq Rapid SBS kit (Illumina, Tokyo, Japan) to generate 60-base single-end reads according to the manufacturer's protocol. The sequenced reads obtained in the ChIP-seq analysis were mapped to the UCSC human genome (hg19) using Bowtie 1.2.2, and duplicate reads were removed with Picard tools. Peak calling and motif analysis were performed using HOMER software (http://homer.salk.edu/homer/index.html), which was also used to identify differential peaks. Enhancers for the nearest genes were annotated with the Genomic Regions Enrichment of Annotations Tool (GREAT) (http://bejerano.stanford.edu/great/public/html/index.php). ChIP-qPCR analysis was conducted using the specific primers for POU5F1 and NANOG promoter regions as detailed in Supplementary Data 4.

ATAC-seq

ATAC-seq analysis of 5-FU-, ADR-, and DNR-treated hESCs was performed as previously described with minor modifications[69]. The cells were trypsinized using TrypLE select (Thermo Fisher Scientific), neutralized by AK02N, and centrifugated at 700 × g for 5 min. A total of 50,000 cells were pelleted and washed with 50 μL of 1× PBS prior to incubation with 50 μL of 2× lysis buffer. Nuclei were resuspended in 40 μL of Tagment DNA Enzyme/Buffer supplemented with 2.5 μL of Tn5 transposase (Illumina, 15027865) to tag and fragment accessible chromatin. The mixture was incubated at 37 °C for 30 min; the DNA fragments were purified with a gel/PCR DNA purification system (Viogene) and amplified by 9-12 cycles of PCR based on the amplification curve. The libraries were purified with a gel/PCR DNA purification system (VioGENE/Nacalai) and subjected to size selection with a BluePippin system (Nippon Genetics) following the manufacturer's protocol. The libraries were subjected to 60 sequencing cycles (single-end reads) on the Illumina HiSeq 1500 platform.

Subcellular fractionation analysis

The procedure for separating the cytoplasmic, nucleoplasmic, and chromatin fractions was based on previously described protocol[70,71]. Briefly, hESCs treated with or without 1 mM 5-FU were harvested and washed with PBS. The cells were then resuspended in nuclear isolation buffer (0.32 M sucrose, 10 mM Tris-HCl [pH 7.5], 5 mM $MgCl_2$, and 1% Triton X-100) and incubated on ice for 20 min. Nuclei were collected by centrifugation (10 min, 500 × *g*, 4 °C), and the supernatant was recovered as the cytoplasmic fraction. The nuclei were washed once with nuclear isolation buffer, lysed for 10 min in extraction buffer (3 mM EDTA and 0.2 mM EGTA), and vortexed for 2 min. The insoluble chromatin pellet and nucleoplasmic fraction were separated using centrifugation (5 min, 1700 × g, 4 °C). RNA was isolated from each fraction using ISOGEN (Nippon Genetics) according to the manufacturer's instructions. Total RNA (1 μg) was reverse transcribed using a ReverTra Ace qPCR RT kit (TOYOBO). The cDNA products were subjected to RT-PCR using the Step One Plus Real-Time PCR System (Applied Biosystems).

RNA-FISH

Full-length sense and antisense digoxigenin (DIG)-labeled RNAs were transcribed in vitro using DIG RNA Labeling Mix (Sigma-Aldrich) and T7 RNA polymerase (Takara), treated with DNase I (Qiagen) and purified using a RNeasy Mini Kit (Qiagen). hESCs were grown on coverslips in 4-well chambers (Lab-Tek, Thermo Fisher Scientific) coated with Matrigel and iMatrix and then treated with 1000 μM 5-FU for 24 h. Alternatively, hESCs were grown on coverslips in 4-well chambers (Lumox x-well, SARSTEDT, Tokyo, Japan) coated with iMatrix. The cells were then rinsed once with PBS, fixed with 4% paraformaldehyde in PBS (Wako) for 10 min at room temperature, and washed thrice with PBS. The fixed cells were permeabilized with 0.5% Triton X-100 in PBS for 5 min at room temperature and washed twice with PBS. Prehybridization was performed with prehybridization solution (1× Denhardt's solution [Nacalai], 50% formamide, 10 mM EDTA, 300 ng/mL salmon sperm DNA, and 0.01% Tween 20 in 2× saline–sodium citrate [SSC]) in a humidified chamber at 54 °C for 2 h. Subsequently, hybridization was performed at 55 °C overnight in hybridization solution (5% dextran sulfate in prehybridization solution) containing each probe at a final concentration of 2 μg/mL. After hybridization, the coverslips were transferred to fresh plates and washed with washing buffer (0.01% Tween 20 and 50% formamide in 2× SSC; one wash), buffer A (0.01% Tween 20 in 2× SSC; one wash), and buffer B (0.01% Tween 20 in 0.1× SSC; two washes) at 55 °C for 30 min each. The coverslips were blocked with 2% bovine serum albumin containing 0.1% Tris-buffered saline with Tween 20 (TBST) for 1 h at room temperature in a humidified chamber. The coverslips were washed thrice with TBST and incubated with a 1/100 dilution of an anti-DIG antibody (Sigma-Aldrich) at 4 °C overnight in a humidified chamber. The coverslips were washed thrice with 0.1% TBST for 5 min each and incubated with a secondary antibody (goat anti-mouse Alexa Fluor 488, Invitrogen, Tokyo, Japan) in the dark at room temperature for 1 h in a humidified chamber. Finally, the coverslips were rinsed thrice with 0.1% TBST for 5 min and mounted with DAPI (Vector Laboratories, CA, USA). Images were acquired using a confocal laser scanning microscope (LSM 5 Pascal or LSM 980, Carl Zeiss, Jena, Germany).

RNA pulldown and MS analysis

Biotin-labeled full-length *LOC644646* and its antisense strand were transcribed in vitro with Biotin RNA Labeling Mix (Sigma-Aldrich) and T7 RNA polymerase (Takara), treated with DNase I (Roche, Tokyo, Japan) and purified using an Rneasy Mini Kit (Qiagen). hESCs were cultured in 10 cm dish, treated with or without 1000 μM 5-FU for 24 h, and pelleted by centrifugation (500 × *g*, 4 °C, 5 min). Then, cell pellet was incubated with swelling buffer (25 mM HEPES [pH 7.9], 1.5 mM $MgCl_2$, 10 mM KCl, 0.1% NP-40, and protease inhibitor cocktail) on ice for 10 min, pipetted well, and collected via centrifugation (700 × g, 4 °C, 5 min). Each nuclei was incubated with RNA binding buffer (25 mM TRIS [pH 7.9], 150 mM KCl, 0.5% NP-40, 5 mM EDTA, and

protease inhibitor cocktail), homogenized on ice with an ultrasonic homogenizer QSONICA Q700 (Waken Btech) in cold water with 15 s on/45 s off cycles and total sonication energy of 1000 J, and centrifuged at 14,000 × *g* for 10 min at 4 °C; supernatants were then collected. Biotinylated RNAs were immobilized on the surface of Dynabeads MyOne Streptavidin C1 Beads (Veritas) according to the manufacturer's protocol. The RNA-linked beads were washed once with RNA binding buffer and incubated with nuclei lysate diluted in RNA binding buffer supplemented with 40 U of RNase inhibitor (Nacalai) and 30 mM EDTA for 3 h at 4 °C with rotating. The precipitated RNA-binding protein complexes were washed five times with RNA binding buffer, once with water, and incubated with 50 µL reaction buffer (50 mM Tris-HCl pH 8.0, 0.02% lauryl maltose neopentyl glycol and 10 mM $CaCl_2$) (all washes for 5 min each at 4 °C).

Beads binding proteins were digested with 500 ng of trypsin (Trypsin Platinum, Promega Madison, WI, USA) in 50 mM Tris-HCl pH 8.0, 0.02% lauryl maltose neopentyl glycol and 10 mM $CaCl_2$, then incubated at 37 °C for 14 hours. The supernatant was transferred to a new tube and treated with 10 mM TCEP and 40 mM 2-chloroacetamide at 80 °C for 15 min. The sample was acidified with 5% TFA, then lauryl maltose neopentyl glycol and salts were removed using SDB-StageTip as previously reported[72], followed by drying with a centrifugal evaporator. The dried peptides were redissolved in 0.02% decyl maltose neopentyl glycol with 0.1% TFA.

The 500 ng of digested peptides were directly injected onto a nanoLC column (C18, 75 µm ID, 30 cm length, 1.7 µm beads; CoAnn Technologies, Richland, WA) at 60 °C and then separated with a 100-min gradient at a flow rate of 150 nL/min using an UltiMate 3000 RSLCnano LC system. The peptides eluted from the column were analyzed by DIA mode using an Orbitrap Exploris 480 with an InSpIon system[73]. MS1 spectra were collected in the range of m/z 495–745 at a 60,000 resolution to set an AGC targets of 3×10^6 and a maximum injection time of "Auto". MS2 spectra were collected at m/z 200-1800 at a 45,000 resolution to set an AGC targets of 3×10^6, a maximum injection time of "Auto", and normalized collision energy of 26%. The isolation width for MS2 was set to 4 Th. The MS files were analyzed by DIA-NN v1.8.1[74]. DIA-NN parameters were as previously reported (PMID: 37036810). Proteins with significant DIA-NN scores and a high signal/noise ratio (sense-antisense > 0, protein intensity >2× 10^5, identified peptides>2) were considered specific LOC644656-interacting proteins.

RIP assay

The RIP assay was performed per the RIP-Assay kit (MBL, Tokyo, Japan) instructions with slight modifications. Pellets of hESC::Tet*LOC644656* cells cultured in the absence or presence of Dox for 1 d (Dox does not affect POU5F1 protein levels at this time point) were lysed with RIPA buffer (50 mM Tris-HCl, 300 mM KCL, 1% NP-40, 0.5% sodium deoxycholate, 1.5 mM dithiothreitol [DTT], 1×Protease Inhibitor Cocktail; Nacalai) and Rnase inhibitor; Nacalai) on ice for 30 min. After centrifugation (14,000 × *g*, for 20 min at 4 °C), the lysates were diluted with lysis buffer from the kit and subjected to immunoprecipitation with normal mouse IgG or an anti-POU5F1 antibody conjugated to Dynabeads Protein A (Veritas) for 3 h. The beads were then washed, and each RNA was purified according to the manufacturer's instructions. The RNA was treated with DNase I for 3 h and purified using FastGene Premium (Nippon Genetics, Tokyo, Japan). RNA was reverse transcribed using a ReverTra Ace qPCR RT kit (Toyobo). The cDNA products were subjected to RT–PCR using the Step One Plus Real-Time PCR System (Applied Biosystems). Relative binding of *LOC644656* RNA was calculated as the % input, and the value for each anti-POU5F1 sample was subtracted from the value for each normal IgG sample.

scRNA-seq library preparation and sequencing

scRNA-seq analysis of 5-FU- and Dox-treated hESCs was performed as previously described with slight modifications[69,75]. Single-cell hESC suspensions were prepared via digestion using TrypLE Select; live cells were then collected via sorting as the DAPI-negative population. scRNA-seq libraries were prepared according to 10× Genomics specifications (10× Chromium Next GEM Single-Cell 3′ Reagent Kits and Library Construction Kit, Tokyo, Japan). Briefly, cell suspensions were loaded into the 10× Genomics Chromium Controller to generate gel beads in emulsion (GEMs). The samples were then incubated at 53 °C for 45 min in a thermal cycler (Veriti, Thermo Fisher Scientific) to generate poly(A) cDNAs barcoded at the 5′-end by the addition of a template switching oligo (TSO) linked to a cell barcode and a unique molecular identifier (UMI). The GEMs were dissolved, and the single-stranded cDNA was cleaned with Dynabeads MyOne Silane Beads (Thermo Fisher Scientific). The cDNA was amplified (98 °C for 3 min; 11 cycles at 98 °C for 15 s and 63 °C for 20 s; 72 °C for 1 min), and cDNA quality was assessed using an Agilent TapeStation system. The cDNA was subjected to enzymatic fragmentation, end repair, poly(A) tailing, and double-sided size selection with SPRIselect beads (Beckman Coulter, Tokyo, Japan) and was then ligated to the adaptors provided in the kit. A unique sample index for each library was introduced through 14 cycles of PCR amplification using the indexes provided in the kit (98 °C for 45 s; 14 cycles at 98 °C for 20 s, 54 °C for 30 s, and 72 °C for 20 s; 72 °C for 1 min; and holding at 4 °C). The indexed libraries were subjected to a second round of double-sided size selection and were then quantified and quality-controlled with the Agilent TapeStation system. The libraries were submitted to GENEWIZ, and the reads were clustered on a paired-end flow cell using the NovaSeq platform and sequenced first based on R1 reads (10× barcodes and UMIs) and then via eight cycles based on the I7 index (sample index) and 89 cycles based on R2 reads (transcripts). The 10× Genomics Cell Ranger Single Cell Software was used for sample demultiplexing, alignment to the human reference genome (hg19), filtering, UMI counting, single-cell 3′-end gene counting, and quality control according to the manufacturer's parameters.

scRNA-seq data analysis

FASTQ files were aligned using the 10X Genomics CellRanger 6.0.0. pipeline to the human GRCh38 reference transcriptome to generate gene-expression count matrices with the "exclude the introns" option. The R package Seurat (v4.03) was used to cluster cells in a merged matrix. A docker image from GitHub (https://github.com/rnakato/docker_singlecell) was used to set up the scRNA-seq data analysis. hESC1s with <10% mitochondrial gene transcripts were first filtered as low-quality cells. The individual gene counts for each cell were divided by the total gene count for each cell, multiplied by a scale factor of 10,000, and subjected to natural log transformation. The FindVariableFeatures function was used with default parameters to select DEGs. The ScaleData function was used to scale and center the counts in the dataset. Principal component analysis was performed on the DEGs, and 50 principal components were used for cell clustering (resolution = 0.5) and UMAP dimensionality reduction. Cluster markers were identified using the FindAllMarkers function, and cell types were annotated manually based on the cluster markers. The cell cycle phase was identified with the cell cycle scoring and regression method[76].

Apoptosis assay

Harvested cells were stained with an Annexin V-633 Apoptosis Detection Kit (Nacalai) according to the manufacturer's protocol. Briefly, 1×10^5 cells were incubated with 5 µL of annexin V and 5 µL of DAPI at room temperature for 15 min in the dark. Subsequently, 1000 cells for each group were analyzed immediately with a flow cytometer (FACSCantoII; BD Biosciences, Tokyo, Japan). The data were analyzed with FlowJo (BD Biosciences). Gating strategy for flow cytometric analysis was demonstrated in Supplementary Fig. 19.

Gene ontology analysis
Gene Ontology (GO) analysis was performed using Enrichr (https://maayanlab.cloud/Enrichr/)[77] as a hub for further GO analyses. The details of each GO analysis are described in the corresponding Figures. In some cases, Metascape (https://metascape.org/)[78] and g:profiler (https://biit.cs.ut.ee/gprofiler/)[79] were employed to visualize the GO results.

Docking simulations
Docking simulations between POU5F1 proteins (8G87-based 1-360 a.a) and LOC644656 RNA (900-1143 bp) was performed using AlphaFold3[46]. Likewise, docking simulations between DNA-PKcs protein (7OTP-based 1-4128 a.a.) and LOC644656 RNA (200-1143 bp) was performed using AlphaFold3. Then, the simulation result were drawn by ChimeraX software[80].

Pan-cancer analysis of *LOC644656* expression in the TCGA database
Pan-cancer analysis of *LOC644656* gene expression across all tissues in all available normal and tumor RNA-seq data was performed using the TNM plot algorithm (https://tnmplot.com/analysis)[51]. Correlation analysis between the expression of *LOC644656* and that of GSMD genes was also performed using the TNM-plotter algorithm. Overall survival in the high- and low-*LOC644656*-expression groups was analyzed using Kaplan-Meier Plotter (https://kmplot.com/analysis)[81] with auto-selection to determine the best cutoff. In some tumors, namely ACC, MESO, and LGG, the data from TCGA were downloaded from MEXPRESS, where the gene name of *LOC644656* was input as "ENSG00000268403" (https://mexpress.be/)[82]. The hazard ratio was then calculated in the custom data analysis of the KM-plotter algorithm. The top 1000 correlated genes from all available 22 tumors (LUSC, UCEC, BRCA, LIHC, COAD, ESCA, KIRC, ACC, TGCT, PRAD, READ, UCS, PAAD, KICH, THCA, BLCA, KIRP SKCM, LUAD, LAML, STAD, and OV) were operated "sum-set" calculation and identified 7650 genes. Subsequently, Venn diagrams were created using Venny 2.1.0 - BioinfoGP (https://bioinfogp.cnb.csic.es/tools/venny/) between *LOC644656* correlated genes and GSMD genes, and 147 *LOC644656*-correlated GSMD genes were extracted. Next, GO analysis was performed. Alternatively, 147 genes with Log2(Tumor/Normal) values of each gene and tumor from the TNM plot were added. The correlation matrix was created using corrplot (https://cran.r-project.org/web/packages/corrplot/vignettes/corrplot-intro.html) of R. Finally, the genes from the correlation matrix were analyzed using GO analysis.

Sense oligos (SO) and antisense oligos transfection
The sense oligos and antisense oligos were designed by Integrated DNA Technologies (IDT, Tokyo, Japan) to target 933-953 bp of *LOC644656*. The transfection of SO/ASO to HepG2, SK-HEP1, MCF-7, and 786-O cells was performed using RNAiMAX (Thermo Fisher Scientific) according to the manufacturer's protocol. For cell viability assay, cells were spread on a 96-well plate after 24 h transfection and further incubated with 24 h. Cells were treated with 5-FU at indicated concentrations for 24 h, and cell viability was determined.

Cell viability assay
Relative cell viability was quantified using a 2-(2-methoxy-4-nitrophenyl)−3-(4-nitrophenyl)−5-(2,4-disulfophenyl)−2*H*-tetrazolium, monosodium salt (WST-8) assay using a Cell Counting Kit-8 (Dojindo, Tokyo, Japan) according to the manufacturer's protocol[83]. Briefly, 1×10^5 MCF-7::Tet*LOC644656* and HepG2::Tet*LOC644656* cells were separately plated in a 96-well plate and incubated with PBS or Dox for 24 h. The cells were treated with 1 mM 5-FU in the presence or absence of Dox for 48 h (MCF-7) or 72 h (HepG2). The cells were incubated with 10 µL of WST-8 solution in a 96-well plate for 2 h at 37 °C. The absorbance values of the samples at 450 nm were obtained using a SYNERGY 2 microplate reader (Bio-Tek) or a Varioskan LUX plate reader (Thermo Fisher Scientific).

Spheroid culture
The spheroid culture was performed according to a previously described protocol[66]. Briefly, 10,000 MCF7::Tet*LOC644656* cells were embedded in 40 µL of Matrigel and spotted in one well of a 24-well plate. The cells were cultured in MammoCult Human Medium (STEMCELL Technologies, 05620/Veritas) for five days until spheroids had formed. Spheroids were treated with Dox (1 µg/mL) for one day, then 5-FU (0.5 mM) was added to the culture for one day. Spheroids were left unfixed and incubated with propidium iodide (PI; 10 µg/mL) and Hoechst (5 µg/mL) in PBS for 30 min at 37 °C. Images of the spheroids were acquired with a Keyence BZ-X800 microscope after three washes with PBS. The areas of spheroids stained with PI or Hoechst were determined from digitized images with a BZ-X800 Analyzer (Keyence, Osaka, Japan). The apoptotic effect of 5-FU was evaluated by calculating the ratio of the area of PI-positive spheroids to that of Hoechst-positive spheroids. All ratios calculated for different conditions were normalized to the ratio for MCF7::Tet*LOC644656* cells without Dox treatment.

Xenograft model
SK-HEP1::Tet*LOC644656* cells were harvested from cell cultures, washed, and 1×10^6 cells were resuspended in 20 µL of PBS and 80 µL of Matrigel without bubbles on ice. Mice were anesthetized using isoflurane, and the cells were subcutaneously injected into the fascia lata with a syringe equipped with a 28 G needle. Tumor growth was monitored by measuring the size of the tumors with a ruler. Once the tumors were at least 6 mm in diameter, regular or Dox-containing chow was provided *ad libitum* for one week. Subsequently, mice were intraperitoneally injected with 50 mg/kg 5-FU thrice weekly for two weeks. The mice were sacrificed, and tumors were collected for analysis.

Statistics and reproducibility
No statistical method was used to predetermine sample size. Sample sizes were based on previous studies in the field and were sufficient to observe statistically significant differences between experimental groups. No data were excluded from the analyses. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

All data are presented as the mean ± SEM of multiple independent experiments with biological replicates. For normally distributed data, statistical analysis was performed using Student's t-test, one-way ANOVA, and two-way ANOVA (as applicable) to evaluate differences between groups. For non-normally distributed data, non-parametric tests were used instead. In cases where statistically significant interactions were identified, Tukey's or Bonferroni's post hoc test was applied for multiple comparisFlue threshold of 0.05 was considered statistically significant. Statistical analyses were conducted using GraphPad Prism 9 software (MA, USA), and figures were generated accordingly. In the TNM plot, p-values were automatically calculated using the Mann-Whitney U-test. For the violin plots, we used the Wilcoxon rank sum test from the AllFindMarker function of Seurat.

All experiments were repeated at least three times independently with similar results unless otherwise specified in the figure legends. For representative images (microscopy, blots, etc.), the number of independent experiments with similar results is indicated in the corresponding figure legends. All tests were two-sided unless otherwise specified.

Reporting summary
Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The raw and processed datasets from bulk and single-cell RNA sequencing of hESCs have been deposited in the DNA Data Bank of Japan (DDBJ) under the following accession numbers: ATAC-seq: DRA018484. RNA-seq: DRA018537. ChIP-seq: DRA018538. scRNA-seq: DRA018542 Source data are provided with this paper.

References

1. Mandal, P. K. & Rossi, D. J. DNA-damage-induced differentiation in hematopoietic stem cells. *Cell* **148**, 847–848 (2012).
2. Santos, M. A. et al. DNA-damage-induced differentiation of leukaemic cells as an anti-cancer barrier. *Nature* **514**, 107–111 (2014).
3. Schneider, L. et al. DNA damage in mammalian neural stem cells leads to astrocytic differentiation mediated by BMP2 signaling through JAK-STAT. *Stem Cell Rep.* **1**, 123–138 (2013).
4. Wang, J. et al. A differentiation checkpoint limits hematopoietic stem cell self-renewal in response to DNA damage. *Cell* **148**, 1001–1014 (2012).
5. Vousden, K. H. & Prives, C. Blinded by the Light: The growing complexity of p53. *Cell* **137**, 413–431 (2009).
6. Aylon, Y. & Oren, M. The paradox of p53: What, how, and why? *Cold spring harb perspect med* **6** (2016).
7. Kastenhuber, E. R. & Lowe, S. W. Putting p53 in context. *Cell* **170**, 1062–1078 (2017).
8. Mello, S. S. & Attardi, L. D. Deciphering p53 signaling in tumor suppression. *Curr. Opin. Cell Biol.* **51**, 65–72 (2018).
9. Oda, K. et al. p53AIP1, a potential mediator of p53-dependent apoptosis, and its regulation by Ser-46-phosphorylated p53. *Cell* **102**, 849–862 (2000).
10. Di Micco, R. et al. Oncogene-induced senescence is a DNA damage response triggered by DNA hyper-replication. *Nature* **444**, 638–642 (2006).
11. Tanaka, T., Ohkubo, S., Tatsuno, I. & Prives, C. hCAS/CSE1L associates with chromatin and regulates expression of select p53 target genes. *Cell* **130**, 638–650 (2007).
12. Majidinia, M. & Yousefi, B. DNA damage response regulation by microRNAs as a therapeutic target in cancer. *DNA Repair (Amst.)* **47**, 1–11 (2016).
13. Huarte, M. et al. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* **142**, 409–419 (2010).
14. Wang, W., Cheng, B., Miao, L., Mei, Y. & Wu, M. Mutant p53-R273H gains new function in sustained activation of EGFR signaling via suppressing miR-27a expression. *Cell Death Dis.* **4**, e574 (2013).
15. Idogawa, M. et al. Identification and analysis of large intergenic noncoding RNAs regulated by p53 family members through a genome-wide analysis of p53-binding sites. *Hum. Mol. Genet* **23**, 2847–2857 (2014).
16. Leveille, N. et al. Genome-wide profiling of p53-regulated enhancer RNAs uncovers a subset of enhancers controlled by a lncRNA. *Nat. Commun.* **6**, 6520 (2015).
17. Schmitt, A. M. et al. An inducible long noncoding RNA amplifies DNA damage signaling. *Nat. Genet* **48**, 1370–1376 (2016).
18. Hu, W. L. et al. GUARDIN is a p53-responsive long non-coding RNA that is essential for genomic stability. *Nat. Cell Biol.* **20**, 492–502 (2018).
19. Yang, Y. et al. TRMP, a p53-inducible long noncoding RNA, regulates G1/S cell cycle progression by modulating IRES-dependent p27 translation. *Cell Death Dis.* **9**, 886 (2018).
20. Lobo, N. A., Shimono, Y., Qian, D. & Clarke, M. F. The biology of cancer stem cells. *Annu Rev. Cell Dev. Biol.* **23**, 675–699 (2007).
21. Kreso, A. & Dick, J. E. Evolution of the cancer stem cell model. *Cell Stem Cell* **14**, 275–291 (2014).
22. O'Connor, M. L. et al. Cancer stem cells: A contentious hypothesis now moving forward. *Cancer Lett.* **344**, 180–187 (2014).
23. Aladjem, M. I. et al. ES cells do not activate p53-dependent stress responses and undergo p53-independent apoptosis in response to DNA damage. *Curr. Biol.* **8**, 145–155 (1998).
24. Lin, T. et al. p53 induces differentiation of mouse embryonic stem cells by suppressing Nanog expression. *Nat. Cell Biol.* **7**, 165–171 (2005).
25. Zhang, Z. N., Chung, S. K., Xu, Z. & Xu, Y. Oct4 maintains the pluripotency of human embryonic stem cells by inactivating p53 through Sirt1-mediated deacetylation. *Stem Cells* **32**, 157–165 (2014).
26. Guttman, M. et al. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat. Biotechnol.* **28**, 503–510 (2010).
27. Cabili, M. N. et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* **25**, 1915–1927 (2011).
28. Hafner, S. J., Talvard, T. G. & Lund, A. H. Long noncoding RNAs in normal and pathological pluripotency. *Semin Cell Dev. Biol.* **65**, 1–10 (2017).
29. Ulitsky, I., Shkumatava, A., Jan, C. H., Sive, H. & Bartel, D. P. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* **147**, 1537–1550 (2011).
30. Lin, N. et al. An evolutionarily conserved long noncoding RNA TUNA controls pluripotency and neural lineage commitment. *Mol. Cell* **53**, 1005–1019 (2014).
31. Loewer, S. et al. Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nat. Genet* **42**, 1113–1117 (2010).
32. Cheng, E. C. & Lin, H. Repressing the repressor: a lincRNA as a MicroRNA sponge in embryonic stem cell self-renewal. *Dev. Cell* **25**, 1–2 (2013).
33. Wang, Y. et al. Endogenous miRNA sponge lincRNA-RoR regulates Oct4, Nanog, and Sox2 in human embryonic stem cell self-renewal. *Dev. Cell* **25**, 69–80 (2013).
34. Jain, A. K. et al. LncPRESS1 Is a p53-Regulated LncRNA that safeguards pluripotency by disrupting SIRT6-mediated de-acetylation of histone H3K56. *Mol. Cell* **64**, 967–981 (2016).
35. Adamsen, B. L., Kravik, K. L. & De Angelis, P. M. DNA damage signaling in response to 5-fluorouracil in three colorectal cancer cell lines with different mismatch repair and TP53 status. *Int J. Oncol.* **39**, 673–682 (2011).
36. Liu, J. C. et al. High mitochondrial priming sensitizes hESCs to DNA-damage-induced apoptosis. *Cell Stem Cell* **13**, 483–491 (2013).
37. Longley, D. B., Harkin, D. P. & Johnston, P. G. 5-fluorouracil: mechanisms of action and clinical strategies. *Nat. Rev. Cancer* **3**, 330–338 (2003).
38. Adamsen, B. L., Kravik, K. L., Clausen, O. P. & De Angelis, P. M. Apoptosis, cell cycle progression and gene expression in TP53-depleted HCT116 colon cancer cells in response to short-term 5-fluorouracil treatment. *Int J. Oncol.* **31**, 1491–1500 (2007).
39. Ju, J., Schmitz, J. C., Song, B., Kudo, K. & Chu, E. Regulation of p53 expression in response to 5-fluorouracil in human cancer RKO cells. *Clin. Cancer Res* **13**, 4245–4251 (2007).
40. Vassilev, L. T. et al. In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **303**, 844–848 (2004).
41. Diaz-Lagares, A. et al. Epigenetic inactivation of the p53-induced long noncoding RNA TP53 target 1 in human cancer. *Proc. Natl Acad. Sci. USA* **113**, E7535–E7544 (2016).
42. Li, X. L. et al. Long noncoding RNA PURPL suppresses basal p53 levels and promotes tumorigenicity in colorectal cancer. *Cell Rep.* **20**, 2408–2423 (2017).
43. Xu, H., Ang, Y. S., Sevilla, A., Lemischka, I. R. & Ma'ayan, A. Construction and validation of a regulatory network for pluripotency and self-renewal of mouse embryonic stem cells. *PLoS Comput Biol.* **10**, e1003777 (2014).

44. Kuleshov, M. V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* **44**, W90–W97 (2016).
45. Huarte, M. The emerging role of lncRNAs in cancer. *Nat. Med* **21**, 1253–1261 (2015).
46. Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
47. Pezone, A. et al. Targeted DNA oxidation by LSD1-SMAD2/3 primes TGF-beta1/ EMT genes for activation or repression. *Nucleic Acids Res* **48**, 8943–8958 (2020).
48. Zhang, Y., Alexander, P. B. & Wang, X. F. TGF-beta family signaling in the control of cell proliferation and survival. *Cold spring harb perspect Biol* **9,** a022145 (2017).
49. Calder, A. et al. Lengthened G1 phase indicates differentiation status in human embryonic stem cells. *Stem Cells Dev.* **22**, 279–295 (2013).
50. Supek, F., Bosnjak, M., Skunca, N. & Smuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* **6**, e21800 (2011).
51. Bartha, A. & Gyorffy, B. TNMplot.com: A web tool for the comparison of gene expression in normal, tumor and metastatic tissues. *Int. J. Mol. Sci.* **22** (2021).
52. Villodre, E. S., Kipper, F. C., Pereira, M. B. & Lenz, G. Roles of OCT4 in tumorigenesis, cancer therapy resistance and prognosis. *Cancer Treat. Rev.* **51**, 1–9 (2016).
53. Beyer, T. A. et al. Switch enhancers interpret TGF-beta and Hippo signaling to control cell fate in human embryonic stem cells. *Cell Rep.* **5**, 1611–1624 (2013).
54. Mohiuddin, I. S., Wei, S. J. & Kang, M. H. Role of OCT4 in cancer stem-like cells and chemotherapy resistance. *Biochim Biophys. Acta Mol. Basis Dis.* **1866**, 165432 (2020).
55. Qin, H. et al. Regulation of apoptosis and differentiation by p53 in human embryonic stem cells. *J. Biol. Chem.* **282**, 5842–5852 (2007).
56. Drabsch, Y. & ten Dijke, P. TGF-beta signalling and its role in cancer progression and metastasis. *Cancer Metastasis Rev.* **31**, 553–568 (2012).
57. Yoon, H. et al. TGF-beta1-mediated transition of resident fibroblasts to cancer-associated fibroblasts promotes cancer metastasis in gastrointestinal stromal tumor. *Oncogenesis* **10**, 13 (2021).
58. Duan, L. et al. VEGFC/VEGFR3 axis mediates TGFbeta1-induced epithelial-to-mesenchymal transition in non-small cell lung cancer cells. *PLoS One* **13**, e0200452 (2018).
59. Xie, F., Ling, L., van Dam, H., Zhou, F. & Zhang, L. TGF-beta signaling in cancer metastasis. *Acta Biochim Biophys. Sin. (Shanghai)* **50**, 121–132 (2018).
60. Findlay, V. J. et al. SNAI2 modulates colorectal cancer 5-fluorouracil sensitivity through miR145 repression. *Mol. Cancer Ther.* **13**, 2713–2726 (2014).
61. Kim, J. et al. Wild-Type p53 Promotes Cancer Metabolic Switch by Inducing PUMA-Dependent Suppression of Oxidative Phosphorylation. *Cancer Cell* **35**, 191–203 e198 (2019).
62. Marei, H. E. et al. p53 signaling in cancer progression and therapy. *Cancer Cell Int* **21**, 703 (2021).
63. Stewart-Ornstein, J. et al. p53 dynamics vary between tissues and are linked with radiation sensitivity. *Nat. Commun.* **12**, 898 (2021).
64. Ma, Y. et al. High-efficiency siRNA-based gene knockdown in human embryonic stem cells. *RNA* **16**, 2564–2569 (2010).
65. Nagano, H. et al. p53-inducible DPYSL4 associates with mitochondrial supercomplexes and regulates energy metabolism in adipocytes and cancer cells. *Proc. Natl Acad. Sci. USA* **115**, 8370–8375 (2018).
66. Ochiiwa, H. et al. TAS4464, a NEDD8-activating enzyme inhibitor, activates both intrinsic and extrinsic apoptotic pathways via c-Myc-mediated regulation in acute myeloid leukemia. *Oncogene* **40**, 1217–1230 (2021).
67. Rizq, O. et al. Dual Inhibition of EZH2 and EZH1 sensitizes PRC2-dependent tumors to proteasome inhibition. *Clin. Cancer Res* **23**, 4817–4830 (2017).
68. Yamagata, K. & Kobayashi, A. The cysteine-rich domain of TET2 binds preferentially to mono- and dimethylated histone H3K36. *J. Biochem* **161**, 327–330 (2017).
69. Fujimoto, M. et al. Liver group 2 innate lymphoid cells regulate blood glucose levels through IL-13 signaling and suppression of gluconeogenesis. *Nat. Commun.* **13**, 5408 (2022).
70. Zhang, A. et al. The human long non-coding RNA-RoR is a p53 repressor in response to DNA damage. *Cell Res* **23**, 340–350 (2013).
71. Gao, S. et al. ROR functions as a ceRNA to regulate Nanog expression by sponging miR-145 and predicts poor prognosis in pancreatic cancer. *Oncotarget* **7**, 1608–1618 (2016).
72. Kawashima, Y., Ishikawa, M., Konno, R., Nakajima, D. & Ohara, O. Development of a simple and stable NanoESI spray system using suction wind from the ms inlet. *J. Proteome Res* **22**, 1564–1569 (2023).
73. Konno, R. et al. Universal pretreatment development for low-input proteomics using lauryl maltose neopentyl glycol. *Mol. Cell Proteom.* **23**, 100745 (2024).
74. Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat. Methods* **17**, 41–44 (2020).
75. Aersilan, A. et al. MicroRNA-874 targets phosphomevalonate kinase and inhibits cancer cell growth via the mevalonate pathway. *Sci. Rep.* **12**, 18443 (2022).
76. Nestorowa, S. et al. A single-cell resolution map of mouse hematopoietic stem and progenitor cell differentiation. *Blood* **128**, e20–e31 (2016).
77. Xie, Z. et al. Gene set knowledge discovery with enrichr. *Curr. Protoc.* **1**, e90 (2021).
78. Zhou, Y. et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).
79. Kolberg, L. et al. g:Profiler-interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Res,* https://doi.org/10.1093/nar/gkad347 (2023).
80. Meng, E. C. et al. UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci.* **32**, e4792 (2023).
81. Nagy, A., Munkacsy, G. & Gyorffy, B. Pancancer survival analysis of cancer hallmark genes. *Sci. Rep.* **11**, 6047 (2021).
82. Koch, A., Jeschke, J., Van Criekinge, W., van Engeland, M. & De Meyer, T. MEXPRESS update 2019. *Nucleic Acids Res* **47**, W561–W565 (2019).
83. Yamagata, K. et al. Arginine methylation of FOXO transcription factors inhibits their phosphorylation by Akt. *Mol. Cell* **32**, 221–231 (2008).

Acknowledgements

We are deeply grateful to Siti Zhahara for her technical assistance throughout the experiments in this study. We also thank M. Matsumoto and K. Nakayama for their technical support with LC–MS/MS analysis, and the laboratory of A. Kaneda for their contributions to next-generation sequencing procedures and data processing. We further acknowledge R. Yuki for his support in ATAC-seq analysis. We would like to express our appreciation to Naoki Takahashi for his technical assistance. This work was supported by grants from the Ministry of Education, Culture, Sports, Science and Technology (Japan) (Grants-in-Aid: for Scientific Research [B] #22300325, #19H03708, #21H02974, #23H02809, #23K24147; [C] #25K11780, #25K10723, #25K10638, #24K10279, 23K09074, #22K07205, #22K08619; Challenging Research [Pioneering] and [Exploratory] #23K17429, #21K19398; JSPS Core-to-Core Program, (JPJSCCA20200006). This work was supported by the

Japan Science and Technology Agency (JST) (grant no. JPMJPF2301). This work was partly supported by the Uehara Memorial Foundation, the Mochida Memorial Foundation for Medical and Pharmaceutical Research, the Naito Foundation, the Mitsui Life Social Welfare Foundation, the Princes Takamatsu Cancer Research Fund, the Takeda Science Foundation, the Senshin Medical Research Foundation, the Kose Cosmetology Research Foundation, the Japan Diabetes Foundation, the Yamaguchi Endocrine Research Foundation, the Cell Science Research Foundation, the Ichiro Kanehara Foundation for the Promotion of Medical Sciences and Medical Care, the Yasuda Memorial Medical Foundation, the Hamaguchi Foundation for the Advancement of Biochemistry, the Ono Medical Research Foundation, the Novartis Foundation (Japan) for the Promotion of Science and the Medical Institute of Bioregulation Kyushu University Cooperative Research Project Program.

Author contributions

A.T. identified LOC644656 from hESCs. K.Y. discovered GSMD and characterized the whole functions of LOC644656 in hESCs and cancer cell lines. A.T., K.Y., T.K., A.O., and T.T. designed experiments and compiled the manuscript. A.T. and K.Y. maintained hESCs, optimized protocols, and performed biochemical and molecular assays, RNA-seq, and omics analysis. K.Y. analyzed the pan-cancer functions of LOC644656 in the TCGA database and performed ChIP-seq and ATAC-seq analyses. K.Y. and M.F. prepared scRNA-seq libraries and analyzed the data. T.F. performed confocal imaging and biochemical analysis. M.N. and I.S. constructed the expression vectors and established stable cell lines. M.Y. performed FACS analysis and sorting. K.Y., M.Y., and A.N. established the xenograft model. N.H. prepared and analyzed spheroid cultures. A.O. analyzed data from ATAC-seq and ChIP-seq. N.M. supervised cohort analysis. S.M. provided research funding and conceptual advice. Y.K. and O.O. supervised genomic analyses and provided analytical tools. E.K. and T.M. supervised data analyses. T.T. supervised the study. All authors discussed the results and contributed to the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-59886-w.

Correspondence and requests for materials should be addressed to Tomoaki Tanaka.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.