nature communications

Article

https://doi.org/10.1038/s41467-024-55481-7

Defining ortholog-specific UHRF1 inhibition by STELLA for cancer therapy

Received: 6 March 2024

Accepted: 11 December 2024

Published online: 08 January 2025

Check for updates

Wenjing Bai[1,2,3,9], Jinxin Xu[2,3,9], Wenbin Gu[1,2,3,9], Danyang Wang[3,9], Ying Cui[4,9], Weidong Rong[1,2,3], Xiaoan Du[2,3,5], Xiaoxia Li[3,5], Cuicui Xia[1,2,3], Qingqing Gan[2,3], Guantao He[3,5], Huahui Guo[1,2,3,5], Jinfeng Deng[3], Yuqiong Wu[3], Ray-Whay Chiu Yen[4], Srinivasan Yegnasubramanian[4], Scott B. Rothbart[6], Cheng Luo[7,8], Linping Wu[3], Jinsong Liu[2,3] ✉, Stephen B. Baylin[4,6] ✉ & Xiangqian Kong[1,2,3,4] ✉

UHRF1 maintains DNA methylation by recruiting DNA methyltransferases to chromatin. In mouse, these dynamics are potently antagonized by a natural UHRF1 inhibitory protein STELLA, while the comparable effects of its human ortholog are insufficiently characterized, especially in cancer cells. Herein, we demonstrate that human STELLA (hSTELLA) is inadequate, while mouse STELLA (mSTELLA) is fully proficient in inhibiting the abnormal DNA methylation and oncogenic functions of UHRF1 in human cancer cells. Structural studies reveal a region of low sequence homology between these STELLA orthologs that allows mSTELLA but not hSTELLA to bind tightly and cooperatively to the essential histone-binding, linked tandem Tudor domain and plant homeodomain (TTD-PHD) of UHRF1, thus mediating ortholog-specific UHRF1 inhibition. For translating these findings to cancer therapy, we use a lipid nanoparticle (LNP)-mediated mRNA delivery approach in which the short mSTELLA, but not hSTELLA regions are required to reverse cancer-specific DNA hypermethylation and impair colorectal cancer tumorigenicity.

Cancer-specific promoter DNA hypermethylation associated with silencing of tumor suppressor genes (TSGs) is an important driver and hallmark of human malignancy[1–4]. These cancer-specific changes are maintained by DNA methyltransferase 1 (DNMT1), which is recruited to chromatin and DNA replication forks by the multi-domain protein ubiquitin-like with plant homeodomain (PHD) and RING finger domains 1 (UHRF1)[5–8]. Reversing abnormal DNA hypermethylation, via targeting the UHRF1-DNMT1 maintenance axis, represents a potential therapy paradigm for treating both hematologic and solid tumors[9–12]. For this purpose, we have recently defined the histone tail recognition PHD and hemi-methylated DNA binding SET and RING-associated (SRA) domains of UHRF1 as key determinants for maintaining abnormal DNA hypermethylation in colorectal cancer (CRC)[13]. Disrupting either domain phenocopies UHRF1 depletion in impairing tumor growth and metastasis associated with DNA demethylation and upregulation of TSGs[13]. These findings provide the rationale for targeting

[1]Guangdong Provincial Key Laboratory of Stem Cell and Regenerative Medicine, CAS Key Laboratory of Regenerative Biology, China-New Zealand Joint Laboratory on Biomedicine and Health, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China. [2]State Key Laboratory of Respiratory Disease, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China. [3]Institute of Drug Discovery, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China. [4]The Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, The Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA. [5]University of Chinese Academy of Sciences, Beijing 100049, China. [6]Department of Epigenetics, Van Andel Institute, Grand Rapids, MI 49503, USA. [7]State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai 201203, China. [8]School of Pharmacy, Guizhou Medical University, Guiyang 550004, China. [9]These authors contributed equally: Wenjing Bai, Jinxin Xu, Wenbin Gu, Danyang Wang, Ying Cui. ✉e-mail: liu_jinsong@gibh.ac.cn; sbaylin@jhmi.edu; kong_xiangqian@gibh.ac.cn

the chromatin reader functions of UHRF1 for cancer drug development[14,15].

STELLA (also known as DPPA3 or PGC7), is a maternal factor required for early embryogenesis and pluripotency maintenance[16–19], which helps regulate gains and losses of DNA methylation during developmental stages of embryogenesis in mice and is requisite to guide the proper direction of cell lineages[18,20–27]. Accumulating evidence defines mouse STELLA (mSTELLA) as a natural inhibitor of UHRF1 through disrupting UHRF1-chromatin associations. The dynamics for removing DNA methylation occur through the binding of UHRF1 to STELLA, displacement of the conjugates from chromatin in the nucleus, and their translocation and retention into the cytoplasm of oocytes and embryonic stem cells (ESCs)[26,28,29]. These antagonistic effects of mSTELLA on DNA methylation rely on its direct interactions with the PHD domain of UHRF1, which abolishes the histone H3 tail binding and chromatin association capability of this protein[29,30]. The potent inhibition towards maintenance or de novo methylation by mSTELLA raises the hypothesis that its human ortholog may play comparable, indispensable roles in DNA methylation regulation, especially in establishing or maintaining the abnormal DNA methylation patterns in cancer cells. Recent studies indicate ectopically expressed human STELLA (hSTELLA) induces genome-wide DNA demethylation in progenitor-like hepatocellular carcinoma (HCC) cells, or slow-cycling of CRC cells, thereby prompting their oncogenic properties and resistance to chemotherapy[31,32]. However, the methylation changes in these above studies are actually subtle-to-moderate at best. This raises the question of whether hSTELLA is as capable of mediating potent UHRF1 inhibition as mSTELLA dose, thereby providing a major mechanism for regulating oncogenesis via DNA demethylation. Thus, the differential capabilities of endogenous hSTELLA *versus* mSTELLA to block cancer-specific DNA methylation abnormalities remains largely unknown. This is important since its expression in human cancers is heterogenous and generally low in HCC and CRC models[31,32]. In this regard, a recent report highlights that h*STELLA* knockout human ESCs (hESCs) have substantially elevated DNA methylation levels compared with wild-type (WT) controls after converting from a hypermethylated prime state to hypomethylated naïve, eight-cell (8 C) embryo-like totipotent cells. It is noteworthy that hSTELLA depletion blocked such cell fate conversion, and the DNA hypermethylation may reflect the intrinsic DNA methylation differences between the two states. The direct effects from UHRF1 inhibition are thus unknown[33]. In summary, all above studies suggests that the role of mSTELLA and hSTELLA needs continued investigation, especially for roles of these proteins in cancer.

In this work, we uncover distinct differences of hSTELLA *versus* mSTELLA in inhibiting the maintenance or de novo DNA methylation functions of UHRF1 in human cancer cells. Genetic abrogation or overexpression of hSTELLA has limited effects on global DNA methylation and oncogenic functions of UHRF1 in cancer cells, in contrast to the profound effects of mSTELLA. We demonstrate the role of a low level of sequence identities between the two proteins and how their structural differences contribute to their differential binding modes to UHRF1. We then develop a drug inhibitor strategy based on these findings.

Results

hSTELLA depletion has minor impact on global DNA methylation in testicular germ cell tumors (TGCT) cells

Interrogations of the h*STELLA* expression profile in The Cancer Genome Atlas (TCGA) and The Human Protein Atlas (HPA) studies[34–36], reveal high mRNA levels exclusively in testicular germ cell tumors (TGCT) and cell lines compared with other cancer types (Figs. S1a and S1b). These data are consistent with the gonad-restricted expression of h*STELLA* in normal tissues (Fig. S1c)[19,37]. Among the histologic TGCT subtypes transformed or reprogrammed from distinct stages of germline development, embryonal carcinoma, closely resembling the epigenomic features of hESC[38,39], has a remarkably higher h*STELLA* expression than other subtypes in TCGA studies (Fig. 1a)[38]. This may coincide with STELLA being a well-defined pluripotency marker of ESC or induced pluripotent stem cells (iPSCs)[17,18,33]. Given that mSTELLA strongly inhibits UHRF1 to reduce DNA methylation[28,29], we speculated higher h*STELLA* expression may track with lower DNA methylation levels in TGCT tumors. However, we find this not to be the case as the overall DNA methylation levels of embryonal carcinoma, teratoma and yolk sac tumors are comparable, and seminomas retain the most prominent DNA hypomethylation as seen in primordial germ cells (PGCs) (Fig. 1b)[39]. These above results agree with the 5-methylcytosine (5mC) immunostaining results in primary tumor tissues[40,41]. Furthermore, we fail to observe a significant inverse association between h*STELLA* expression and global DNA methylation levels for tumors in each histologic TGCT subtype (Fig. 1c).

Previous studies have demonstrated that mSTELLA is primarily co-localized with UHRF1 in the cytoplasm of mESC and early stage oocytes, and its loss results in an almost exclusive nuclear localization of UHRF1 with genome-wide DNA hypermethylation[28,29]. To query the capability of hSTELLA to inhibit UHRF1 in TGCT cells, we stably depleted this protein in BeWo choriocarcinoma cells, a human, placental subtype of TGCT which has the highest h*STELLA* levels across all cancer cell lines in the HPA database (Fig. S1b)[42]. While hSTELLA and UHRF1 are far more enriched in the nucleus of these cells (Figs. 2a, b and S1d), especially in the chromatin binding portion of BeWo cells (Fig. S1e), our depletion of hSTELLA does not alter this subcellular localization and chromatin association of UHRF1 or of the major DNMTs (Figs. 2a, b and S1d, e). Furthermore, h*STELLA* knockout (Fig. S1f–h) negligibly, though statistically significantly, alters genome-wide DNA methylation at CpG island or non-CpG island promoters, gene body and intergenic regions in either BeWo cells or NCCIT embryonal carcinoma cells (Figs. 2c and S1i). However, abrogation of hSTELLA elicits dramatic transcriptome changes. Gene Ontology (GO) analysis indicates the differentially expressed genes (DEGs) induced upon hSTELLA depletion are significantly enriched in GO terms related to tissue or organ development and morphogenesis, key events during embryogenesis (Figs. 2d, e and S1j)[43]. Also, gene set enrichment analysis (GSEA)-based pathway analysis identifies that a variety of cell-type specific gene expression profiles are altered, again those related to early organ development (*e.g.* sensory and urogenital organs) and tissue morphology (*e.g.* muscular, vascular and glandular tissues) (Fig. S1k). In parallel with the above results, most of these transcriptome alternations are not associated with corresponding promoter DNA methylation changes, with only 6 of 341 down-regulated genes displaying concurrent gain of promoter DNA methylation ($\log_2$ fold change < -1 and $\Delta\beta > 0.2$) in hSTELLA-depleted BeWo cells (Figs. 2f and S1l). However, none of these 6 genes are recognized as potential TSGs based in the TSGene database (Supplementary Data 1)[44]. Moreover, we fail to observe significantly increased promoter DNA methylation for the 474 down-regulated genes in NCCIT cells upon h*STELLA* knockout (Figs. 2f and S1l). In conclusion, the above results indicate a modest role of hSTELLA in establishing and maintaining DNA methylation, and its associated TSG expression in TGCT cells, presumably due to the ineffective UHRF1 inhibition. Instead, hSTELLA may function as a pivotal transcriptional regulator sustaining the developmental potentials or pluripotent states[33], in a largely DNA methylation-independent manner.

Overexpression of mSTELLA, rather than hSTELLA, reverses cancer-specific DNA hypermethylation of CRC cells and inhibits CRC tumor growth

Next, the functional differences between human and mouse STELLA are explored for blocking UHRF1 and the consequential effect on maintaining DNA methylation in CRC cells, in which the maintenance of abnormal promoter DNA hypermethylation and its associated TSG silencing by the histone-binding domains of UHRF1 profoundly fosters

https://doi.org/10.1038/s41467-024-55481-7

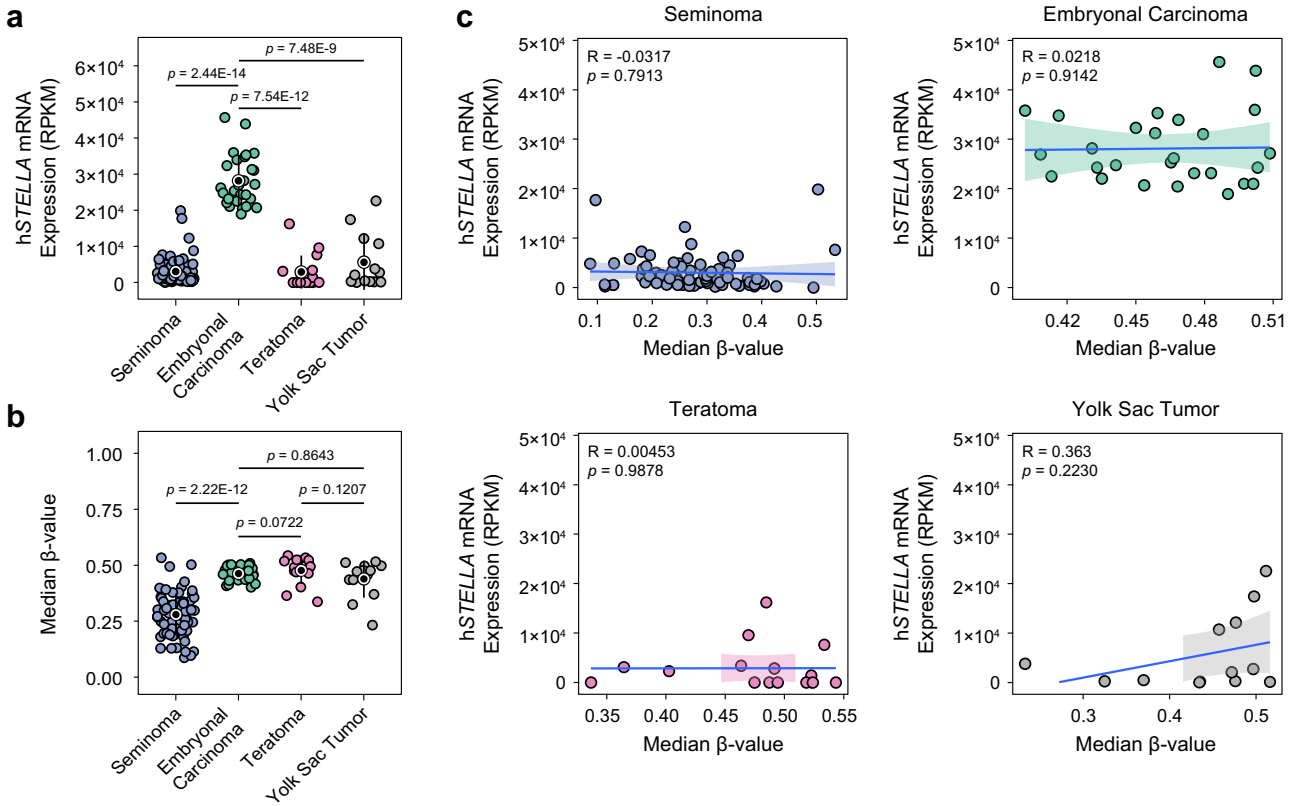


Fig. 1 | No significant correlation between h*STELLA* expression and global DNA methylation levels in each histologic TGCT subtype. a, **b** h*STELLA* mRNA expression levels (RPKM) (**a**) and global DNA methylation levels (median β-value) (**b**) across different TGCT subtypes in TCGA studies. The whiskers represent one standard deviation from the means (black-dots). *P*-values were calculated by unpaired two-sided Wilcoxon rank-sum test. **c** Scatterplots showing the correlation between h*STELLA* expression and global DNA methylation levels in different TGCT subtypes in TCGA studies. Correlation coefficients (R) and two-sided *p*-values were calculated with Pearson's Correlations. The hatched area indicates the 95% confidence interval. For **a**–**c**, the TGCT subtypes include Seminoma ($n = 72$ patients), Embryonal Carcinoma ($n = 27$ patients), Teratoma ($n = 16$ patients), and Yolk Sac Tumor ($n = 13$ patients).

tumorigenicity[4,13,45]. Concordant with the low or absent h*STELLA* expression levels in CRC tumors in TCGA transcriptome data (Fig. S1a), hSTELLA protein is below-detectable level in most of the tumor tissues including CRCs from The Clinical Proteomic Tumor Analysis Consortium (CPTAC) studies[46,47]. This is in contrast to the ubiquitous expression of UHRF1 (Fig. S2a, b). It is noteworthy that a small number of hSTELLA-positive cells might be present but heterogeneously distributed in tumor tissues, as has been seen for immunohistochemistry findings in CRC and HCC tissues[31,32]. However, such heterogeneous expression signals of hSTELLA might be concealed during the TCGA or CPTAC assays using bulk-tumor tissues, especially for the far majority of cancers with a low percentage of hSTELLA-positive cells.

Regarding our above findings, we firstly confirm the lack of endogenous STELLA protein expression in a variety of human CRC cell lines, including HCT116 and RKO cells that have high UHRF1 protein levels and the abnormal gene promoter CpG island methylator phenotypes or CIMPs (Fig. S2c, d)[13,48]. Next, we assess the subcellular localization of UHRF1 after inserting largely comparable amounts of exogenous mSTELLA and hSTELLA in HCT116 and RKO cells (Fig. S2e). As observed previously[28,30], a nucleo-cytoplasmic translocation of UHRF1 occurs with ectopically expressed mSTELLA, with UHRF1 being diminished in nuclei and becoming mainly co-localized in the cytoplasm of both cell lines (Figs. 3a, S2f–l). These dynamics are also accompanied by a substantial loss of chromatin binding of UHRF1 (Fig. 3b). In contrast, the exogenous expression of hSTELLA does not induce any of the above events (Figs. 3a, b and S2f–l). Accompanying these results, there is a profound difference between mSTELLA and hSTELLA wherein extensive DNA methylation loss across the genome occurs only with mSTELLA over-expression (Fig. 3c–e and S3a,b). These differences include removal of cancer-specific, abnormal promoter DNA methylation that is maintained by UHRF1 histone binding domains in our previous studies[13] (Figs. 3d, e and S3b). With this promoter de-methylation, there is distinct re-activation of a large number of potential TSGs silenced in association with aberrant promoter DNA hypermethylation in both CRC cell lines (Figs. 3f, g and S3c). These include the negative regulators of CRC proliferation and metastasis (*p16*, *SFRP1*, *UCHL1*, *MATK*, *TFPI2*, *FBLN2*, *ICAM4* and *CDH4*)[13,49–51], the activators of epithelial cell differentiation (*GATA5* and *TFF2*)[52], the oxidative stress antagonists (*GPX3* and *CYGB*)[53,54], and the signal transducer *IRF7* that potentiates anti-tumor immunity[55]. In contrast, for the oncogenes annotated by at least three sources in the OncoKB knowledgebase (Supplementary Data 2)[56], mSTELLA overexpression is ineffective in further prompting their promoter DNA demethylation and upregulation (Fig. S3d–g). In accordance with the limited functions of hSTELLA in regulating the TGCT transcriptome via changes in DNA methylation (Fig. S1l), none of the up-regulated genes induced by hSTELLA over-expression in CRC cells exhibit significant promoter DNA hypomethylation ($\log_2$ fold change > 1 and $\Delta\beta < -0.2$ in Figs. 3g and S3h). Importantly, significantly higher levels of hSTELLA protein are observed in BeWo cells compared to those in NCCIT and hSTELLA-overexpressing CRC cells (Fig. S3i). Among these, BeWo cells exhibit moderately greater alterations in promoter DNA methylation and associated gene expression upon changing hSTELLA expression (Figs. S1l and S3h). Considering the weak yet existing effects of hSTELLA in disrupting UHRF1-chromatin associations (Fig. 3b), these data collectively suggest that, in contrast to mSTELLA, hSTELLA is a far

https://doi.org/10.1038/s41467-024-55481-7

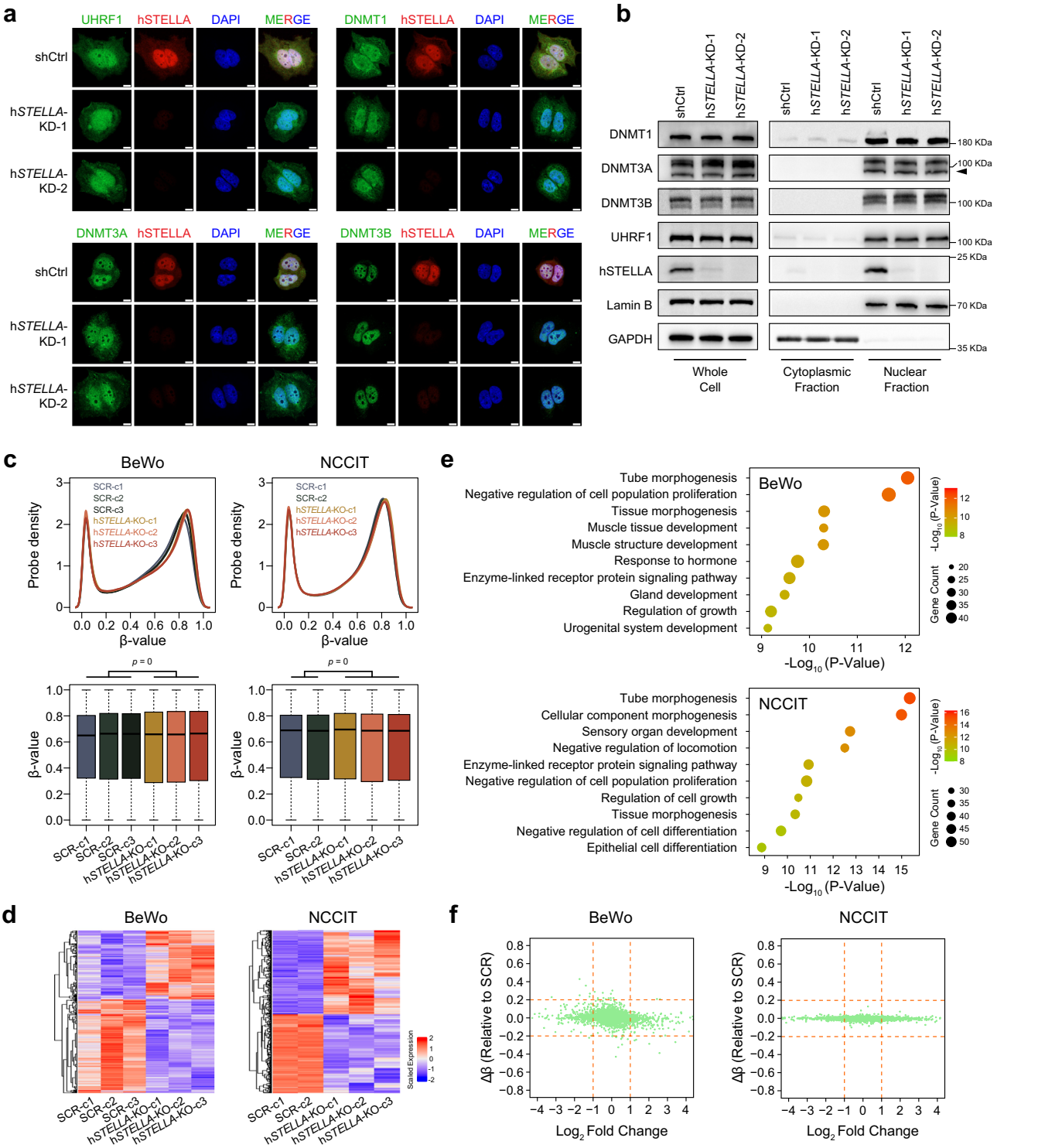
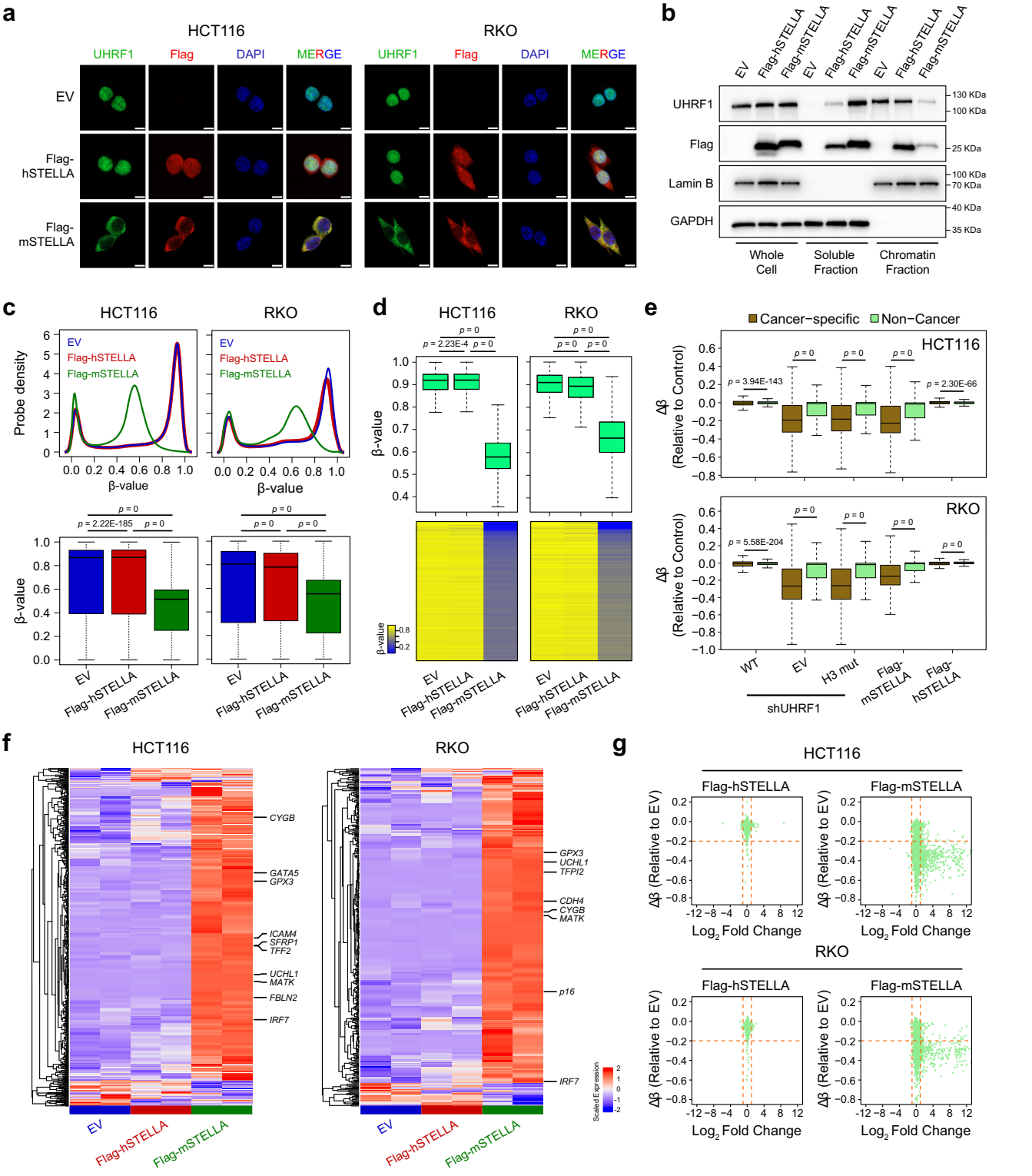


Fig. 2 | hSTELLA is ineffective at modulating UHRF1 subcellular localization and DNA methylation in TGCT cells. a, **b** Immunofluorescence analysis (**a**) and biochemical fractionation analysis (**b**) of the subcellular localizations of indicated proteins in BeWo cells transduced with either non-silencing shRNA (shCtrl) or two shRNA constructs targeting h*STELLA* (h*STELLA*-KD-1 and h*STELLA*-KD-2). Scale bars, 10 μm (**a**). Arrow indicates the specific band of DNMT3A (**b**). **c** Global DNA methylation profiles of scramble-control (SCR) and h*STELLA* knockout (KO) cell clones derived from BeWo and NCCIT cells. β-value (x axis) between 0 and 1 indicates DNA methylation level, and probe density (y axis) describes the distribution of β-values for all probes (top). Boxplots (bottom) are specified as follows: midlines, median; box-limits, 25th-75th percentiles; whiskers, minimum/maximum values, outliers included. *P*-values are calculated by paired two-sided Wilcoxon rank-sum test for the averaged DNA methylation levels of the SCR group and h*STELLA*-KO group, with $n = 864{,}258$ CpG sites for each group in both cell lines. **d** Heatmaps showing the DEGs ($|\log_2$ fold-change$| > 1$ and two-sided Wald test, Benjamini-Hochberg adjusted p-value < 0.05) between h*STELLA*-KO and SCR-control cell clones for both BeWo and NCCIT cells. **e** Metascape-based Gene Ontology (GO) enrichment analysis of DEGs in h*STELLA*-KO BeWo and NCCIT cells, with top 10 significantly enriched terms depicted. *P*-values are calculated by two-sided cumulative hypergeometric distribution. **f** Relationship between gene expression changes (Log_2 Fold Change) and promoter DNA methylation changes (mean difference of β-values, Δβ) in h*STELLA*-KO BeWo and NCCIT cells. The orange vertical lines represent $\log_2$ fold-change equal to -1 and 1, while the orange horizontal lines indicate promoter Δβ values equal to 0.2 and -0.2. Source data are provided as a Source Data file.

https://doi.org/10.1038/s41467-024-55481-7



less effective UHRF1 inhibitor with subtle-to-minimal effects in transcriptional regulation through antagonizing DNA methylation, and such effects depend on its cellular expression levels. In this regard, for past studies, the differential expression levels of hSTELLA and the distinct cell models used, which also have varying UHRF1 levels[13], may all contribute to the discrepancies between our present findings and previous studies showing hSTELLA over-expression prompts subtle-to-moderate global DNA demethylation and UHRF1 delocalization[31,32]. Interestingly, a far less number of DEGs are induced by ectopically expressed hSTELLA in CRC cells than with its depletion in TGCT cells (Figs. S3c and S1j), consistent with the DNA methylation-independent transcription regulatory functions of hSTELLA being more specifically required in germ cell tumors. Taken together, our findings indicate evolutionally diverged and context-dependent actions of STELLA in mouse *versus* humans for orchestrating transcription.

In view of the above differences for maintaining DNA methylation in cancer cells, we reason that mSTELLA and hSTELLA may have separate effects on the phenotypes of the cells. Indeed, ectopic expression of mSTELLA, in contrast to hSTELLA, severely impairs CRC proliferation, and anchorage independent colony growth in soft agar

https://doi.org/10.1038/s41467-024-55481-7

Fig. 3 | Functional differences between hSTELLA and mSTELLA in reversing cancer-specific DNA hypermethylation and TSG silencing. **a** Immunofluorescence analysis of the subcellular localizations of indicated proteins in CRC cells stably expressing empty vector (EV) and Flag-tagged STELLA orthologs. Scale bars, 10 μm. **b** Biochemical fractionation analysis of the subcellular localizations of indicated proteins in 293 T cells transiently transfected with Flag-tagged STELLA proteins. **c** Global DNA methylation profiles of CRC cells stably expressing the STELLA orthologs. β-value between 0 and 1 indicates the DNA methylation level. The probe density describes the distribution of β-values for all probes. *P*-values are calculated by paired two-sided Wilcoxon rank-sum test, with $n = 864,258$ CpG sites for each sample in both cell lines. **d** Promoter demethylation analysis in CRC cells stably expressing the STELLA orthologs. Cancer-specific promoter probes (β-value ≥ 0.75) in EV-transduced cells are used to generate the boxplots and heatmaps. *P*-values are calculated by paired two-sided Wilcoxon rank-sum test, with $n = 89,598$ and 79,462 CpG sites for each sample in HCT116 and RKO cells, respectively. **e** Changes to cancer-specific and non-cancer promoter DNA methylation patterns in CRC cells with endogenous UHRF1 replaced by EV, WT or histone-binding domain mutant (H3 mut) UHRF1 in our previous study[13], as well as those stably expressing the STELLA orthologs. *P*-values are calculated by unpaired two-sided Wilcoxon rank-sum test, with $n = 174,308$ and 118,187 CpG sites for cancer-specific and non-cancer loci in HCT116 cells, respectively, and with $n = 181,297$ and 111,198 CpG sites in RKO cells. **f** Heatmaps showing the DEGs induced by STELLA over-expression in CRC cells, with representative, DNA hypermethylation silenced TSGs labeled. **g** Relationship between gene expression changes (x-axis) and promoter methylation changes (y-axis) in CRC cells stably expressing the STELLA orthologs. The orange vertical lines represent the $\log_2$ fold-change equal to -1 and 1, while the orange horizontal lines indicate Δβ values equal to -0.2. Boxplots in **c**–**e**: midlines, median; box-limits, 25th-75th percentiles; whiskers, minimum/maximum values, outliers excluded. Source data are provided as a Source Data file.

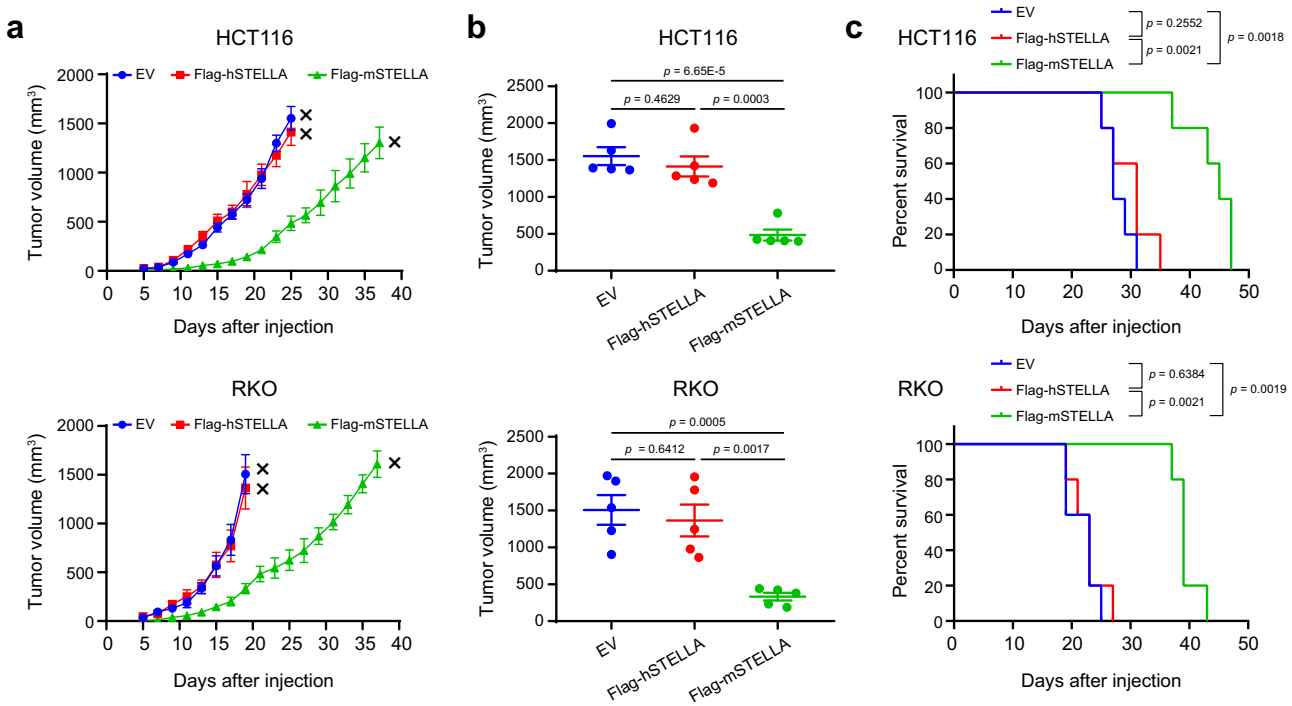


Fig. 4 | Functional differences between hSTELLA and mSTELLA in reversing the oncogenic functions of UHRF1. **a** Tumor growth curves of CRC cells stably expressing the STELLA orthologs in immunodeficient mice. The mark "×" indicates that the observation was stopped for the specific group when at least one mouse was sacrificed due to the tumor volume being approximately 2000 mm^3. **b** Tumor volumes of HCT116 (at day 25) and RKO (at day 19) xenografts from **a**, when at least one mouse's tumor volume in the EV-transduced group was approximately 2000 mm^3, are compared among each group. **c** Survival curves of CRC xenografts from **a**, with $n = 5$ mice per group. Data are represented as mean ± SEM ($n = 5$ mice) in **a** and **b**. *P*-values are calculated by unpaired two-sided t-test in **b**, and log-rank (Mantel-Cox) test in **c**. Source data are provided as a Source Data file.

(Fig. S4a,b). Further, transplantation of CRC cells expressing exogenous mSTELLA into immune-deficient mice, as compared with hSTELLA or empty-vector (EV) expression, markedly reduces tumor burden and decreases the proliferation marker Ki67 in tumor tissues, leading to a prolonged overall survival of tumor bearing mice (Figs. 4a–c and S4c–f). A late subsequent increase of CRC tumors may be due to cells that escaped UHRF1 inhibition and resulted in tumor rebound (Fig. 4a, c) and these dynamics are similar to our previous findings and that from other such studies of key epigenetic proteins essential for tumor growth[13,57,58]. Notably, distinct from past hSTELLA over-expressing CRC xenograft studies, which were performed over a short observation period with limited tumor growth[31], the anti-tumor effects of STELLA orthologs are comprehensively evaluated for a much longer time allowing us to perform survival analysis. Over our observation period, we observe little if any tumor suppressive activity of hSTELLA (Fig. 4a–c). Consistently, in addition to CRC cells, ectopically expressed mSTELLA reverses DNA methylation abnormalities and significantly reduces the proliferation of both lung and breast cancer cells, whereas hSTELLA exhibits minimal such effects in these contexts (Fig. S4g–j). Again, these data provide further credence to the ortholog-specific functions of STELLA proteins in reversing cancer-specific DNA methylation and tumorigenicity, as a consequence of distinct UHRF1 inhibition.

Defining the Swap1 region as the key sequence determinant for differential UHRF1 binding and inhibition between mSTELLA and hSTELLA

A previous study indicates that nuclear export activity may counteract mSTELLA-mediated UHRF1 inhibition, and an mSTELLA mutant deficient in nuclear exporting of UHRF1 induces more pronounced DNA hypomethylation than the WT protein[29]. Alternatively, the interaction of the C-terminal portion of mSTELLA with a UHRF1 PHD domain which

abrogates UHRF1-histone H3 tail recognition and chromatin binding, could be the key prerequisite for efficient UHRF1 inhibition as previously suggested[29,30]. In this regard, we find that although hSTELLA is predominantly localized in the nucleus (Figs. 2a and 3a)[32], the deletion of a predicted, conserved nuclear export signal (NES) in hSTELLA (residues 34-47, designated as hSTELLA-ΔNES) further enhances its nuclear enrichment rather than its DNA demethylating capability (Fig. S5a–c)[16,59,60]. Moreover, the hSTELLA mutant devoid of the putative nuclear localization signals (NLSs) (residues 93-121, designated as hSTELLA-ΔNLS)[61,62], is far more enriched in the cytoplasm, but is unable to de-localize UHRF1 and induce significant DNA demethylation (Fig. S5a–c). These findings suggest the differential cellular localization of STELLA orthologs may not be the primary cause of their differential inhibition of UHRF1 and maintenance DNA methylation. Rather, we postulated that the discrepant UHRF1 binding potency of STELLA orthologs in the nucleus may explain the above differences. Indeed, similar to the full-length proteins[33], there is low conservation at the amino acid level for the C-terminal part of these proteins (Figs. 5a and S5d). To test this hypothesis and further pinpoint the precise C-terminal sequence determining differential UHRF1 association with the orthologs, we designed two chimeric mutants for C-terminal mSTELLA (residues 80-150) which replace the corresponding sequences from hSTELLA (Fig. 5a, b). The two regions selected for chimeric mutation, Swap1 and Swap2, are based on their comparable sequence length and preservation of potential secondary structures (Fig. S5e)[63], while differing most in sequence conservation among STELLA orthologs (Figs. 5a, b and S5d). Replacing of mSTELLA residues 80-119 with the less conserved, matching hSTELLA residues 75-121 renders the chimeric mSTELLA-Swap1 mutant incompetent for mediating chromatin displacement and cytoplasmic translocation of UHRF1. In contrast, for the more distal and conserved regions contained in the mSTELLA-Swap2 mutants, substitution of residues 120-150 by the corresponding hSTELLA sequence produces virtually no changes as compared to WT mSTELLA (Figs. 5c, d and S5f–j).

Next, inspired by our past findings that lowering DNMT1 expression below a certain threshold is required for maximally reversing promoter DNA hypermethylation and TSGs silencing, and also stimulating viral-like immune signaling[8], we established an isogenic HCT116 cell line with roughly 70% reduced UHRF1 levels (HCT116-UHRF1$^{+/-}$, Fig. S5g). This modification aims to increase sensitivity to UHRF1 inhibition mediated by the chimeric mutants, akin to the enhanced sensitivity to DNMT inhibitors driven by DNMT1-deficiency[64]. In these cells, the ectopic expression of mSTELLA, as well as mSTELLA-Swap2 mutant re-activates the epigenetically silenced TSGs and also induces key immune signaling pathways that have been noted by others[65–67] to potentiate the infiltration of tumor-associated immune cells (Figs. 5e and S5k–m). However, interchanging the Swap1 region with hSTELLA largely abolishes all of these effects (Figs. 5e and S5k). Consistent with our assumptions, the functional discrepancy of human and mouse STELLA, and the above results with our two swapping mutants, closely mirrors their UHRF1 binding capabilities. While mSTELLA manifests robust UHRF1 interaction, the mSTELLA-Swap1 mutant, similar to results seen with hSTELLA, associates poorly with UHRF1 (Fig. 5f). In contrast, the chimeric mutation in the more conserved Swap2 region does not affect UHRF1 binding potency from WT mSTELLA (Fig. 5f). Collectively, our data suggests the differential UHRF1 binding capabilities of mouse and human STELLA appear to underpin their distinct inhibition of maintenance DNA methylation.

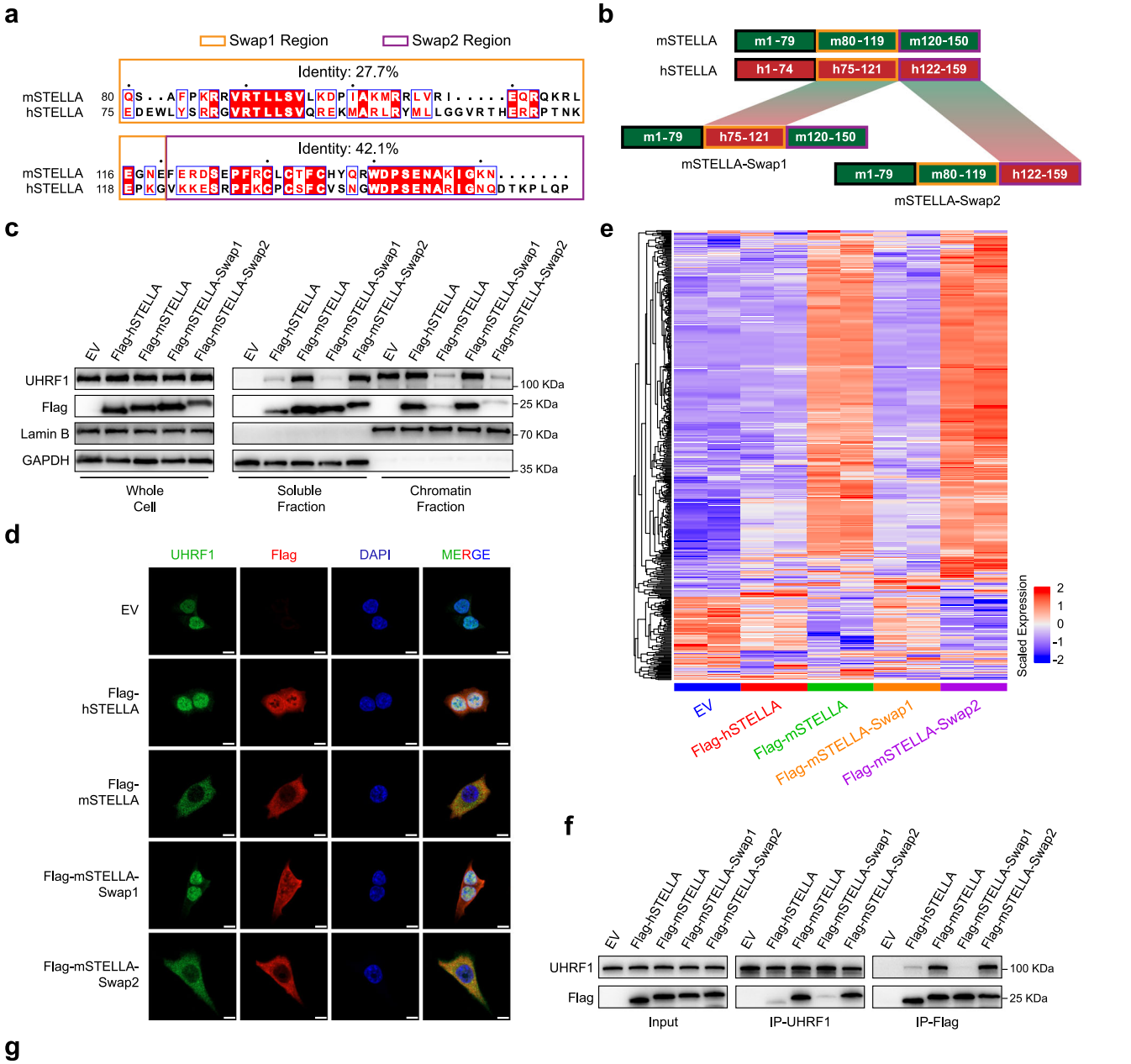
Next, we compared the direct binding affinity between Swap1 regions and UHRF1 histone binding domains. The findings reveal distinct differences for how UHRF1 interacts between the human and mouse orthologs. It has been noted that the UHRF1 tandem Tudor domain (TTD) and PHD fingers are connected by an interdomain linker segment, allowing these regions to function as a single unit to cooperatively recognize H3 tails with trimethylated lysine 9 modification (H3K9me3)[68–70](Fig. S6a). Accordingly, we evaluated the binding affinities of Swap1 peptides with both PHD and linked TTD-PHD domains. Isothermal titration calorimetry (ITC) experiments demonstrate that the peptide of the mSTELLA Swap1 region (hereafter referred to as mSTE-1, with K_d value of 91.16 nM) is nine-fold more potent than the corresponding hSTELLA peptide (hereafter referred to as hSTE-1, with K_d of 818 nM) for associating with the UHRF1 PHD domain at a 1:1 stoichiometry (Figs. 5g and S6b). Intriguingly, the tethering of TTD domain with PHD domain boosts the binding with mSTE-1 more than 15-fold (K_d value of 91.16 nM to 5.84 nM) while such tethering has trivial improvement for hSTE-1 (K_d value of 818 nM to 786 nM), magnifying their UHRF1 binding affinity differences, being more than two orders of magnitude increased (Figs. 5g and S6b,c). These quantitative data indicate the linked TTD-PHD domain, rather than PHD domain alone[30], cooperatively binds to mSTELLA. Such cooperation is similarly observed for H3K9me3 recognition (Fig. 5g), as noted by others[68–70] but is virtually absent for hSTELLA binding (Figs. 5g and S6b,c). Therefore, for competitive binding to TTD-PHD domains, mSTE-1 exhibits greater than 30-fold more potent binding affinity than an H3K9me3 tail while hSTE-1 is 4-fold weaker. Taken together, our findings demonstrate a significant difference between UHRF1 binding to mSTE-1 and hSTE-1, and more importantly, demonstrate for this, is required an ortholog-specific requirement of cooperation between TTD and PHD domains of UHRF1 which has not heretofore been considered or investigated.

Structural studies reveal different binding interactions between mSTELLA and hSTELLA to the TTD-PHD domain of UHRF1

In crystallographic studies, we find stark differences between how peptides from the mSTE-1 and hSTE-1 regions of STELLA orthologs bind to the TTD-PHD domains of human UHRF1, which appear to explain the differential activities of the two orthologues for affecting the maintenance of DNA methylation abnormalities in human cancer cells. We corroborate the previously reported differences in structures for C-terminus mSTELLA versus hSTELLA in binding to the PHD domain[71,72], but more importantly, we provide key insight that an interaction between mSTELLA to the linked TTD-PHD domains of UHRF1 is actually critical (Fig. S7a–d and Table S1).

First, for the binding of mSTELLA to above UHRF1 domains, a truncated peptide within mSTE-1 (residue 85-119, designated mSTE-2 hereafter) maintains comparable UHRF1 binding affinity as mSTE-1 and we utilize this for structural determinations (Figs. 6a and S7e). Overall, this mSTE-2 peptide adopts an L-shaped folding consisting of a short α-helix (αS) connecting to a long helix (αL) through a turn to bind with UHRF1. In this interaction, αS and its preceding ^{88}VRT^{90} motif manifest a similar configuration to how the H3K9me3 tail occupies an acidic pocket in the PHD domain (Figs. 6a and S7f). R89 forms a hydrogen bond (H-bond) which contacts the negatively charged D334 and D337 residues in UHRF1 while V88 and αS in mSTELLA mainly mediate hydrophobic interactions (Figs. 6a and S8a). The importance of these interactions is supported by a co-immunoprecipitation (Co-IP) assay with full length proteins, showing that V88A and R89A/T90A mutations markedly disrupt mSTELLA-UHRF1 binding (Fig. 6b), similar to the reported effects of L91A/V94A mutants[71]. Importantly, residue P98 of mSTE-2 induces a kink between the two α-helices, and partially penetrates into an aromatic cage (F152, Y188 and Y191) of the UHRF1 TTD domain which is known to specifically recognize H3K9me3 (Figs. 6a and S7f)[73,74]. In addition to these similar interactions for H3K9me3 tail recognition (Fig. S7f), αL of mSTE-2, engages an interdomain cavity for additional contacts with both PHD and TTD domains (Figs. 6a and S8b). For these interactions, the positively charged residues K101, R104, R107 and R111 in mSTELLA form polar interactions with D307, C316, D337 and E335 residues of the UHRF1 PHD domain, while L105 and I108 in mSTELLA intimately associate with surrounding UHRF1 residues C302, C316, V315 and M338. Furthermore, the

https://doi.org/10.1038/s41467-024-55481-7



g

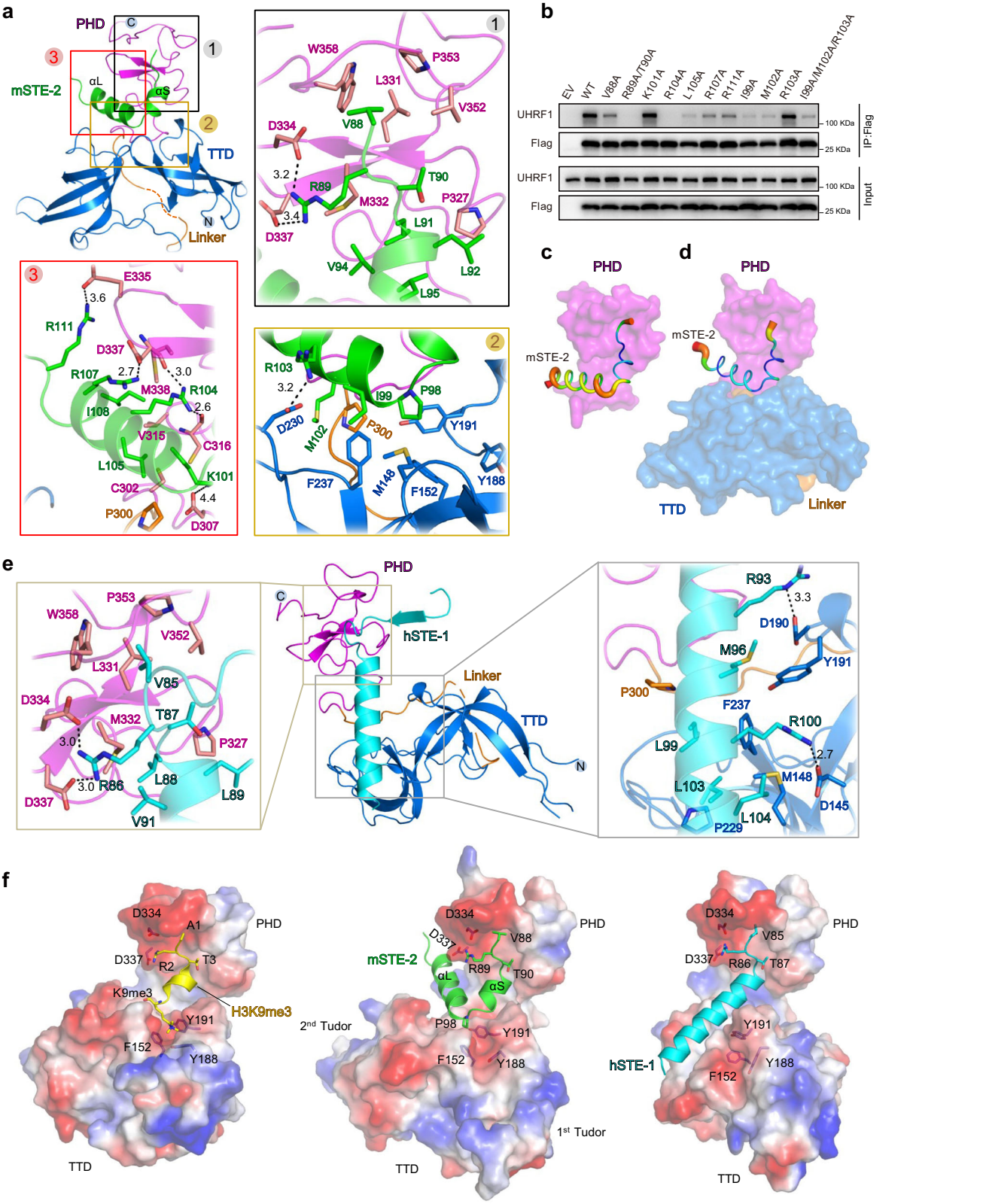
Protein	Peptide	K_d (nM)	N
PHD	mSTE-1	91.16 ± 6.11	1.28 ± 0.06
	hSTE-1	818.00 ± 42.58	0.97 ± 0.28
	H3K9me3	1423.33 ± 289.89	1.16 ± 0.30

Protein	Peptide	K_d (nM)	N
TTD-PHD	mSTE-1	5.84 ± 0.77	1.10 ± 0.11
	hSTE-1	786.00 ± 187.36	1.24 ± 0.10
	H3K9me3	192.33 ± 13.58	0.88 ± 0.06

Fig. 5 | The Swap1 region of STELLA proteins is a key sequence determinant for ortholog-specific UHRF1 binding and inhibition. a Pairwise sequence alignment for the C-terminal parts of mSTELLA (80-150) and hSTELLA (75-159). The identical residues are shaded in red and boxed in blue. The conserved substitutions are in red letters and boxed. The two regions used for swapping experiments, and the sequence identity for each region are indicated. **b** Diagram of the chimeric mutation design for the C-terminal part of mSTELLA. The sequences in Swap1 (residues 80-119) and Swap2 regions (residues 120-150) of mSTELLA are replaced by the hSTELLA residues 75-121 and residues 122-159, respectively. **c** Biochemical fractionation analysis of the subcellular localizations of indicated proteins in 293 T cells transiently transfected with Flag-tagged STELLA proteins. **d** Immunofluorescence analysis of the subcellular localizations of indicated proteins in HCT116-UHRF1$^{+/-}$ cells stably expressing Flag-tagged STELLA proteins or mutants. Scale bars, 10 μm. **e** Heatmaps showing the DEGs (relative to EV) induced by ectopic expression of the STELLA orthologs and mSTELLA swapping mutants in HCT116-UHRF1$^{+/-}$ cells. **f** Reciprocal co-immunoprecipitation between UHRF1 and the ectopically expressed Flag-tagged STELLA proteins in 293 T cells. **g** Summary of the binding affinities (K_d) and stoichiometry (N) between peptides (mSTE-1, hSTE-1, H3K9me3) and UHRF1 PHD or TTD-PHD domains as determined by ITC experiments. Data are represented as mean ± SD ($n = 3$ independent experiments). Source data are provided as a Source Data file.

hydrophobic contacts between mSTELLA αL (I99, M102) and the UHRF1 TTD/Linker regions (M148, F152, F237 and P300), as well as H-bonding between mSTELLA residue R103 and D230 of UHRF1 are observed. Compared with the PHD/mSTE-2 complex alone, these TTD/Linker interactions shift the mSTELLA αL further toward the UHRF1 PHD domain to form a more stable TTD-PHD/mSTE-2 protein complex (Fig. S8b). This is also reflected by the relatively lower B-factor for mSTE-2, a key parameter that quantifies the uncertainty in the position

https://doi.org/10.1038/s41467-024-55481-7



of atoms and structure motion in protein crystallography, in associating with the UHRF1 TTD-PHD domains (Fig. 6c, d). Indeed, disrupting the contacts of αL with either the UHRF1 TTD or PHD domains results in severely defective UHRF1 binding (Fig. 6b), consistent with cooperative and functional interactions of the above essential inter-domains, for mSTE-2 binding to the linked TTD-PHD domains of UHRF1 (Fig. 5g).

Structural diversity for matching regions of hSTE-1 compared to mSTE-2 strongly emphasizes the importance of our above findings. In contrast to the L-shaped conformation of mSTE-2 interactions with UHRF1, the comparative region of hSTE-1 (residues 75-121) forms only a single long straight α-helix for associating with the UHRF1 TTD-PHD domains (Figs. 6e and S7b). The interactions with the PHD domain are as found by others (Fig. S7d)[75]. The ^{85}VRT^{87} motif and N-terminal

Fig. 6 | Structural basis on the differences between mSTELLA and hSTELLA interactions with UHRF1 TTD-PHD domains. a Detailed interactions of UHRF1 TTD-PHD domains with mSTE-2. TTD domain, PHD domain and linker are colored in blue, magenta and orange, respectively. mSTE-2 is shown as green cartoon. The polar interactions including H-bonds are indicated by dashed lines, and the distances (in angstroms) between the interacting atoms are labeled. **b** Co-immunoprecipitation between UHRF1 and the ectopically expressed Flag-tagged mSTELLA mutants in 293 T cells. **c-d** B-factor putty representation of mSTE-2 from crystal structure of UHRF1 PHD/mSTE-2 complex (**c**) and UHRF1 TTD-PHD/mSTE-2 complex (**d**). The narrow blue tubes and broad red tubes represent regions with low B-values and high B-values, respectively. PHD and TTD domains are shown as magenta surface and blue surface, respectively. **e** Detailed interactions of UHRF1 TTD-PHD domains with hSTE-1. hSTE-1 is shown as cyan cartoon. TTD domain, PHD domain and linker are colored as in (**a**). The polar interactions including H-bonds are indicated by dashed lines, and the distances (in angstroms) between the interacting atoms are labeled. **f** Overall structures of UHRF1 TTD-PHD domains bound to H3K9me3 (left, PDB: 4GY5), mSTE-2 (middle) and hSTE-1 (right). The UHRF1 TTD-PHD domains are represented in surface contours which are colored according to the electrostatic potential (blue, positive; white, neutral; red, negative). H3K9me3, mSTE-2 and hSTE-1 peptides are shown in cartoon modes, and colored by yellow, green and cyan, respectively. The side chains are shown as sticks for residues D334 and D337 on UHRF1 PHD domains, and for aromatic cage residues F152, Y188 and Y191 on TTD domains. For peptides, the ART (in H3K9me3 complex) or VRT motifs (in mSTE-2 and hSTE-1 complexes), as well as residues K9me3 (in H3K9me3 complex) or P98 (in mSTE-2 complex) interacting with the aromatic cage on TTD domains are labeled. Source data are provided as a Source Data file.

portion (residues 88-91) of the α-helix, conserved between mSTELLA and hSTELLA orthologs (Fig. 5a), similarly occupy the H3 binding pocket in the UHRF1 PHD domain as the VRT-αS segment of mSTE-2 (Figs. 6a, e and S7f). However, one amino acid difference, the substitution of P98 in mSTE-2 by a lysine residue K95 in hSTE-1, tracks with the difference from the mSTE-2 interactions with UHRF1 TTD-PHD domains. This difference corresponds to the kinked conformation in mSTE-2 due to the unique dihedral angle constraints from the proline backbone[76]. In comparison, the extension of the α-helix of hSTE-1 along the TTD domain (Fig. 6e), mediated by the lysine residue K95 substitution with a strong α-helical propensity[77], fails to achieve the contacts with the UHRF1 PHD induced by αL in mSTE-2 (Fig. 6a). This later hSTE-1 variation is consistent with its remarkably weaker PHD domain binding affinity than for mSTE-1/2 (Figs. 5g and S7e). The components of these differences involve hSTE-1 mediating both polar (R93, R100) and hydrophobic (M96, L99, L103, L104) interactions with residues on the UHRF1 TTD domain surface (D190, D145, M148, P229, F237, P300 and Y191) (Figs. 6e and S8c). These contacts induce a relative movement of the C-terminal α-helix by the PHD/hSTE-1 complex structure that not only abrogates the H-bond between K95 and S301 in the PHD/hSTE-1 complex, but also shifts the PHD domain loop residues (D356 and E355) outward away from H-bond interacting with G84 and V85 of hSTELLA (Fig. S8d). Given the importance of these PHD loop interactions in recognizing the N-terminus of the histone H3 tail and binding to PAF15 by UHRF1[70,78,79], the loss of these key H-bonds may offset the UHRF1 TTD domain interactions, leading to a negligible contribution of this region in hSTELLA-UHRF1 interactions (Fig. 5g).

Lastly, a distinct interdomain configuration of UHRF1 accompanies the above conformational differences of mSTE-2 and hSTE-1. Structural superposition of UHRF1 TTD-PHD domains complexed with mSTE-2 versus hSTE-1 with an H3K9me3 complex shows that the individual TTD or PHD domains share highly similar folds (Fig. S8e), while their relative orientation differs remarkably (Fig. S8f). For mSTE-2, the clockwise rotation (~15°) of the UHRF1 PHD domain enlarges the cavity between it and a 2nd Tudor module for fitting αL that mediates coordinated contacts with TTD-PHD domains (Figs. 6f and S8b). These dynamics lead to a much higher UHRF1 binding affinity for mSTE-2. In contrast, a large-scale anticlockwise rotation (~62°) of the PHD domain is observed when comparing the TTD-PHD/hSTE-1 complex with the H3K9me3 complex (Fig. S8f). Such spatial arrangement helps to accommodate the long, rigid α-helix of hSTE-1, but paradoxically disrupts its H-bond interactions to the PHD loop as aforementioned and limits its contacts with the aromatic cage in the UHRF1 TTD domain (Fig. 6f). Collectively, the immediately above configuration changes result in a defective UHRF1 binding of hSTE-1 comparing with the H3K9me3 tail.

In summary, our data suggests that although both hSTE-1 and mSTE-2 orthologs commonly occupy the H3K9me3 binding site but the species-specific, cooperative interactions with the linked TTD-PHD domains underpins their distinct capabilities in competitive inhibition of UHRF1-H3 interactions. Importantly, these findings underscore the potential that these inter-domain dynamics could be leveraged to develop potent UHRF1 inhibitors through conformationally mimicking the multifaceted, cooperative contacts of mSTELLA with UHRF1. This potential is explored in detail in the section below.

Lipid nanoparticle (LNP) delivery of mRNA encoding mSTELLA effectively reprograms the DNA methylome of CRC cells

We lastly explore the feasibility of utilizing our findings for mSTELLA interactions with UHRF1 to craft a strategy for designing a potential UHRF1 inhibitor for cancer therapy. As a proof-of-principle study, we have designed a peptidic nanoparticle (PNP) drug delivery approach utilizing a nanoparticle cargo coated with SV40 nuclear localization signal (NLS) peptides. This nanoparticle delivery of the short mSTE-1 peptide alone (residues 80-119), corresponding to the mSTELLA Swap1 region essential for both UHRF1 interactions and inhibition (Fig. 5c–g), into UHRF1-deficient HCT116-UHRF1$^{+/-}$cells (Fig. S5g) fails to re-activate the promoter DNA hypermethylation silenced TSGs (Fig. S9a,b). This suggests that other flanking regions of mSTELLA might be required for maintaining a proper conformation of mSTE-1 for efficient UHRF1 interactions in cells. We thus turned to an LNP mediated mRNA delivery strategy utilizing mRNA encoding both WT mSTELLA and the UHRF1 binding deficient mSTELLA-Swap1 mutant in which the L-shaped, cooperative TTD-PHD binding region of mSTELLA (residues 80-119) is replaced by the long α-helical hSTELLA sequence (residues 75-121) (Figs. 5b, 6a, e). For this, we utilize LNPs encapsulated through a microfluidic system (Fig. 7a, Tables S2 and S3)[80,81], wherein only LNP-mSTELLA robustly stimulates TSG expression in UHRF1-deficient cells (Fig. S9c). Furthermore, inhibition of UHRF1-H3 binding in parental HCT116 cells by LNP-mSTELLA elicits a fast and durable reversion of the CRC specific, abnormal promoter DNA hypermethylation, reflected by an extensive DNA demethylation at Day 2 followed by a slowly progressive decline of DNA methylation levels from Day 6 to Day 10 (Figs. 7b and S9d). The DNA methylation dynamics largely correlate with transcriptomic changes showing noticeable gene re-activation at Day 2 post-treatment of LNP-mSTELLA and then reaching a plateau between Day 6 and Day 10 (Fig. 7c). The efficiency of this LNP strategy for UHRF1 inhibition is further highlighted by the more potent DNA demethylation exerted towards hypermethylated probes and stronger correlating transcriptional activation of gene expression, when comparing with the effects induced by lentivirus mediated mSTELLA overexpression (Fig. S9e,f). In contrast, these above results are in striking contrast to those produced by the LNP-mSTELLA-Swap1 mutation which replaces mSTE-1 peptide residues with the comparative hSTE-1 peptide residues (Fig. 5b). This switch produces a sharp decrease in mSTELLA engagement with the UHRF1 TTD-PHD domain (Fig. 5c, f) and this mutant is virtually incapable of derepressing the hypermethylated promoter mediated epigenetic silencing (Fig. 7b, c).

The above comparisons are emphasized when analyzing the findings for effects on specific, potentially important TSG genes which are DNA hypermethylated in their CpG island promoters with

https://doi.org/10.1038/s41467-024-55481-7

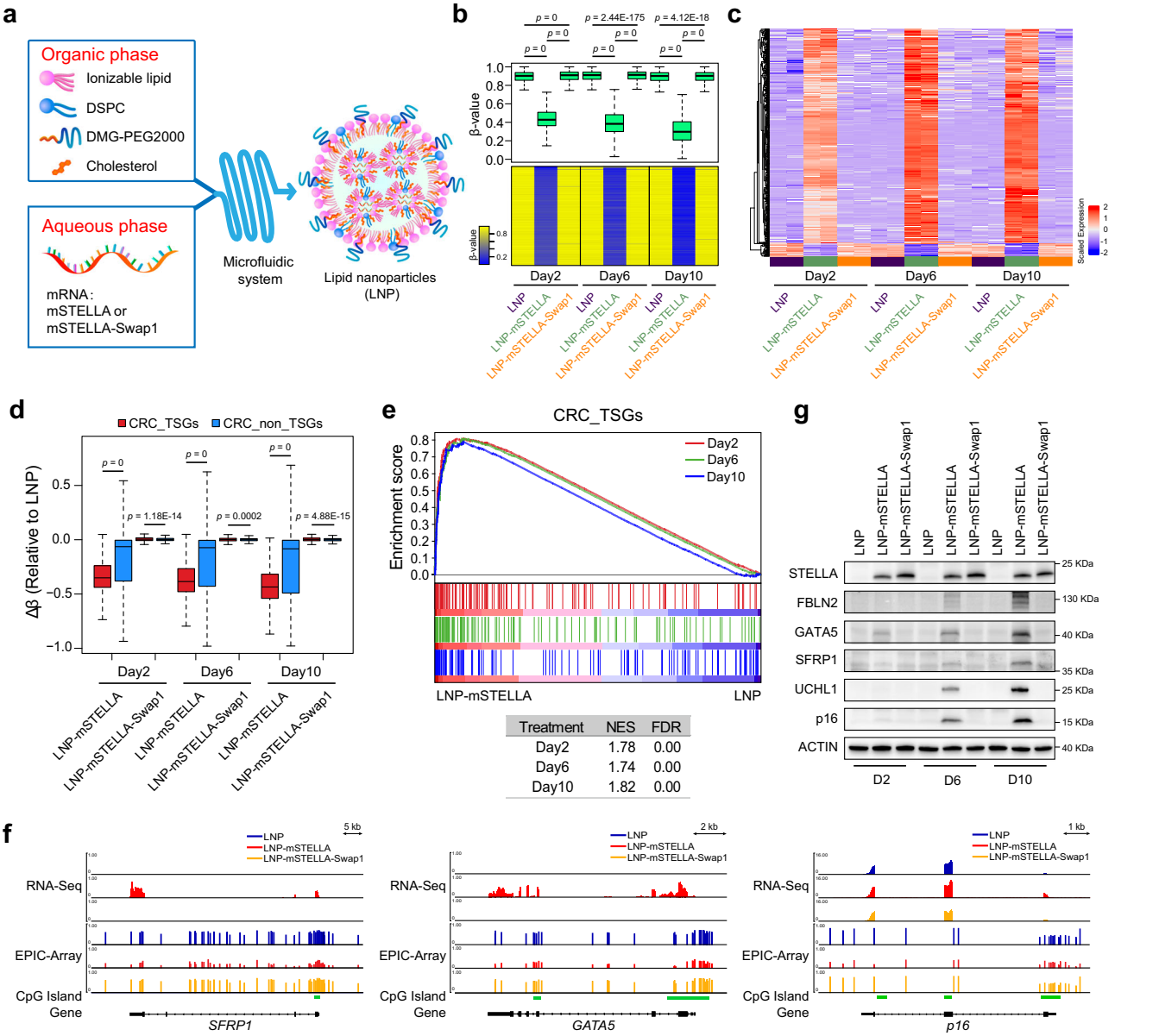
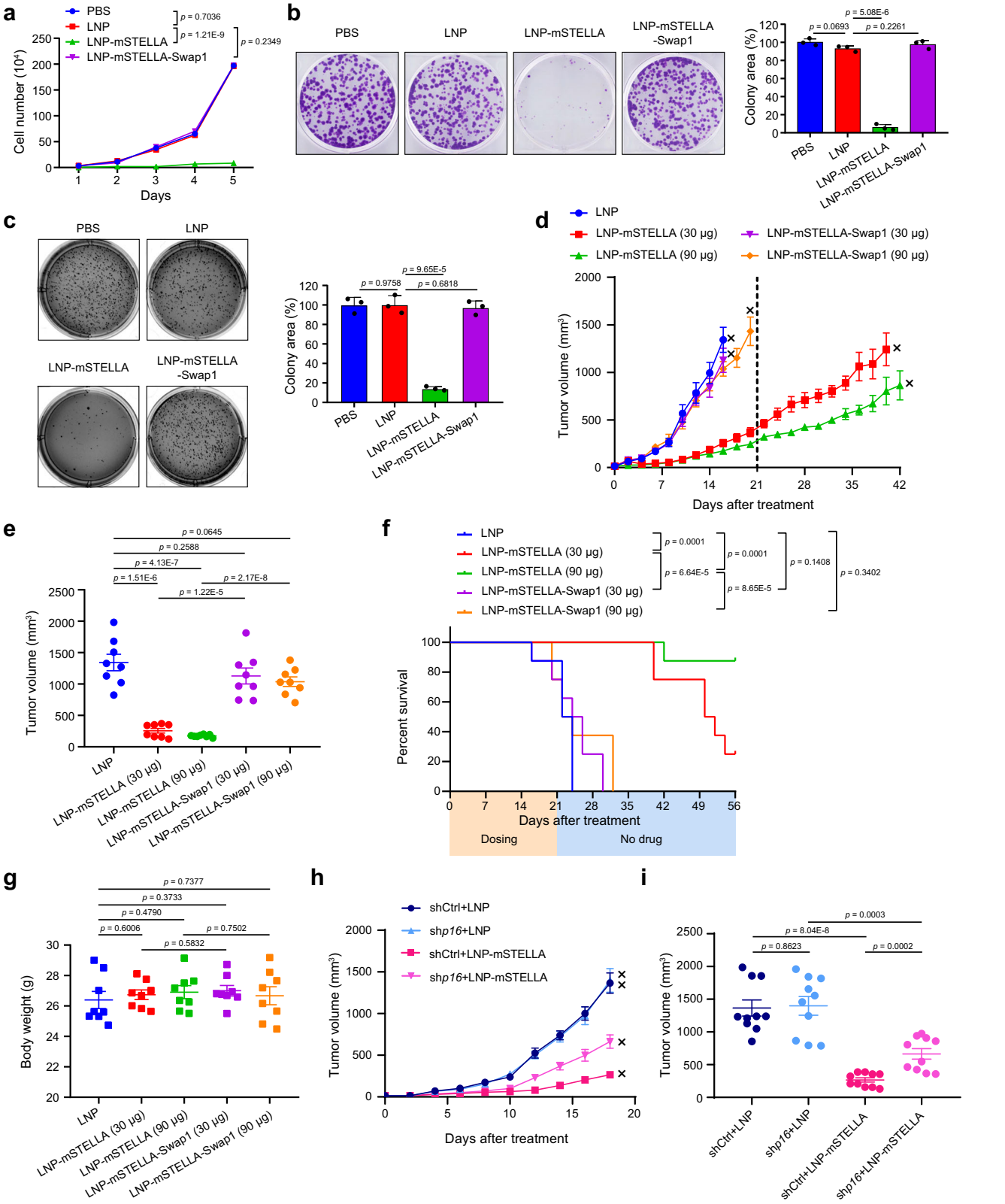


Fig. 7 | LNP delivery of mRNA encoding mSTELLA reverses DNA methylation abnormalities and reactivates epigenetically silenced TSGs in CRC cells. **a** Schematic illustration of the preparation of LNP for delivery of mRNAs encoding mSTELLA and mSTELLA-Swap1. **b** Promoter DNA demethylation analysis at Days 2, 6 and 10 following LNP, LNP-mSTELLA and LNP-mSTELLA-Swap1 treatments in HCT116 cells. Cancer-specific promoter probes (β-value ≥ 0.75) in LNP-treated HCT116 cells (Day 2) are used to generate the boxplot and heatmap. Boxplot is specified as follows: midlines, median; box-limits, 25th-75th percentiles; whiskers, minimum/maximum values, outliers excluded. *P*-values are calculated by paired two-sided Wilcoxon rank-sum test, with $n = 89{,}598$ CpG sites for each sample. **c** Heatmaps showing the DEGs (relative to LNP) induced by LNP-mSTELLA and LNP-mSTELLA-Swap1 treatments in HCT116 cells at Days 2, 6 and 10. **d** Promoter DNA demethylation analysis for CRC_TSGs and CRC_non_TSGs at Days 2, 6 and 10 following LNP-mSTELLA and LNP-mSTELLA-Swap1 treatments in HCT116 cells. Boxplot is specified as follows: midlines, median; box-limits, 25th-75th percentiles; whiskers, minimum/maximum values, outliers excluded. *P*-values are calculated by unpaired two-sided Wilcoxon rank-sum test, with $n = 5{,}486$ and 287,139 CpG sites for CRC_TSGs and CRC_non_TSGs, respectively. **e** GSEA analysis of the transcriptomic changes induced by LNP-mSTELLA treatments at Days 2, 6 and 10 showing enrichment with the CRC_TSGs gene set. NES, normalized enrichment score; FDR, false discovery rate. **f** Transcriptome (RNA-seq) and DNA methylome (EPIC-Array) profiles at *SFRP1* (left), *GATA5* (middle) and *p16* (right) gene loci in LNP, LNP-mSTELLA and LNP-mSTELLA-Swap1 treated HCT116 cells. The CpG island regions for each gene are highlighted in green. **g** Western blot analysis of the indicated TSGs and STELLA proteins at Days 2, 6 and 10 following LNP, LNP-mSTELLA and LNP-mSTELLA-Swap1 treatments in HCT116 cells. Source data are provided as a Source Data file.

concurrent transcriptional silencing in CRC samples compared with adjacent normal tissues in TCGA studies (Fig. S9g,h)[82]. These genes, which also manifest profound promoter hypermethylation in the CRC cells used, are defined in a similar manner as in our previous study (Fig. S9g-h)[13], and designated in studies below as CRC_TSGs (Supplementary Data 3). Importantly in this regard, a large percentage of the above genes are normally controlled during development in ESC by the polycomb group (PcG) proteins and basally have what is termed "bivalent" histone modifications comprising simultaneously, both the transcriptionally repressive histone modification mark H3K27me3 and the active mark H3K4me3 in their non-methylated promoters in ESCs. This chromatin scenario maintains expression of the genes at a low poised expression level, but allows an inducible transcription state to allow repression of stem cell characteristics and promote of



differentiation dynamics in key cell lineages during development[83]. We and others have shown that these bivalent genes are predisposed to promoter CpG island hypermethylation in human cancers, which abolishes their inducibility thus contributing to CRC tumorigenicity[84–86].

Similar to effects of pharmacologically reducing DNA methylation in cancer cells, treatment of CRC cells by LNP-mSTELLA induces preferential promoter DNA demethylation for these above TSGs in CRC and the key bivalent developmental genes[85], associating with significant transcriptional up-regulation (Figs. 7d, e and S9i,j). Examples include the WNT signaling antagonist *SFRP1*[49], intestinal epithelium differentiating transcription factor *GATA5*[87], and apoptosis or angiogenesis regulatory genes (*UCHL1* and *FBLN2*) whose promoter hypermethylation tracks with CRC progression (Figs. 7f, g and S9k)[13], and their reactivation is also

Fig. 8 | LNP-mSTELLA inhibits the tumorigenicity of CRC cells both in vitro and in vivo. a Cell growth curves of HCT116 cells with the indicated treatments. Cell numbers are counted daily from Day 1 to Day 5. Data are represented as mean ± SD ($n = 3$ independent experiments). **b, c** Colony formation (**b**) and soft agar assays (**c**) of HCT116 with the indicated treatments. The percentages of colony area (relative to PBS-treatment) are shown in bar plots. Data are represented as mean ± SD ($n = 3$ independent experiments). **d, e** Tumor growth curves of HCT116 mice xenografts peritumorally injected with the indicated LNP formulations every other day (**d**), and tumor volumes at Day 16 from (**d**), when at least one mouse's tumor volume in the LNP-treated group was approximately 2000 mm^3 (marked as "×"), are compared (**e**). The vertical dashed line indicates the time at which the dosing ceased (**d**). Data are represented as mean ± SEM ($n = 8$ mice). **f** Survival curves of HCT116 xenografts from (**d**) where dosing occurred during the first 21 days, with $n = 8$ mice per group. **g** Body weights of HCT116 xenografts at Day 16 from (**d**) are compared among each group. Data are represented as mean ± SEM ($n = 8$ mice). **h, i** Mice xenografts are established by subcutaneous injection of HCT116 cells expressing the indicated shRNA, and peritumorally injected with the indicated LNP formulations (30 μg/dose) every other day. Tumor growth curves are recorded (**h**) and tumor volumes at Day 18 from (**h**), when at least one mouse's tumor volume in the shCtrl+LNP group was approximately 2000 mm^3 (marked as "×"), are compared (**i**). Data are represented as mean ± SEM ($n = 10$ mice). Unpaired two-sided t-test in **a** (Day 5), **b, c, e, g** and **i**, and log-rank (Mantel-Cox) test in **f**. Source data are provided as a Source Data file.

observed in CRC cells transduced with mSTELLA (Fig. 3f). Additionally, the cell cycle inhibitor p16 and blocker of stem cell differentiation during development is one of the most important TSG's epigenetic silencing events in human cancer[88–90]. In mice, silencing of *p16* alone sufficiently drives tumorigenesis[91,92]. In HCT116 cells, this gene is frameshift mutated for the constitutively expressed and unmethylated allele that produces no functional protein, while the other WT allele is silenced by DNA hypermethylation[93]. Reversion of cancer-specific DNA methylation by LNP-mSTELLA, but not LNP-mSTELLA-Swap1, reactivates the expression of p16 protein from the hypermethylated WT allele (Fig. 7f, g). Additionally, we and others have recently shown inhibiting maintenance DNA methylation in cancer cells provokes an anti-tumor immune response through potentiating type I/III interferon signaling. These effects are triggered by upregulating the demethylated endogenous retrovirus which in turn stimulates the viral defense pathway, a phenomenon termed viral mimicry[55,66,67,94]. Indeed, we observed a strong induction of the viral defense and interferon signaling by LNP-mSTELLA (Fig. S9l,m), congruent with previous findings from UHRF1 histone binding domain mutations[13].

LNP-mSTELLA impairs CRC tumorigenicity by demethylating key TSGs

The above results motivated us to further evaluate the therapeutic potential of pharmacological UHRF1 inhibition for CRC tumorigenicity. Indeed, there are considerably attenuated cell proliferation rates and anchorage-independent growth of HCT116 cells for LNP delivered mSTELLA, but not by the LNP-mSTELLA-Swap1 mutant (Fig. 8a–c). These phenotypic changes may not involve a significant alternation of ten-eleven translocation (TET) enzymatic activities (Fig. S10a), as LNP-mSTELLA treatment imparts comparable antiproliferative effects for both TETs-deficient and WT CRC cells (Fig. S10b–d). Considering the extensive hepatic accumulation of LNPs when systemically administrated[95,96], in-vivo mouse delivery of the above nanoparticles is achieved by peritumor injection of LNP-mSTELLA every other day. Such treatment regimen substantially reduces tumor burden in an HCT116 xenograft model, accompanying by the reversion of cancer-specific promoter DNA hypermethylation and TSG silencing in tumors (Figs. 8d–f and S10e–i). It is noteworthy that the tumor growth is still suppressed after treatment ended at Day 21, probably through an epigenetic memory effect observed with clinically relevant low doses of DNMT inhibitors[97,98], leading to significantly extended overall survival time (Fig. 8d, f). While the 30 μg/dose group of LNP-mSTELLA exhibits a 26 ~ 28-day survival benefit, a > 33-day such benefit is achieved by the 90 μg dosing group with 87.5% of mice surviving 5 weeks after treatment ceases. On the contrary, vehicle or the LNP-mSTELLA-Swap1 control which replaces the mSTELLA residues 80-119 with the corresponding hSTELLA residues 75-121 shows no survival benefits.

Administration of therapeutic proteins, especially those produced from different species, may induce severe immunological reactions leading to reduced tolerability and therapeutic efficacy, characterized by the release of high levels of pro-inflammatory cytokines[99,100]. Nevertheless, LNP-mediated expression of mSTELLA or mSTELLA-Swap1 does not lead to the acute upregulation of a variety of pro-inflammatory genes (Fig. S10j). This is further supported by enzyme-linked immunosorbent assay (ELISA) conducted on representative cytokines secreted by HCT116 and human monocyte THP-1 cells. Specifically, both LNP-mSTELLA and LNP-mSTELLA-Swap1 do not increase the levels of IFN-γ and IL-2, similar to the response observed with lipopolysaccharide (LPS) treatments. Moreover, for TNF-α and IL-6 that are effectively stimulated by LPS in THP-1 cells, the above LNP formulations remain incapable of prompting their secretion (Fig. S10k). Consistently, there are no apparent adverse effects on either body weight or major organs of mice during the 8-week in vivo treatment (Figs. 8g and S10l). In summary, our findings suggest that pharmacological blocking of the TTD-PHD domain of UHRF1 via a mSTELLA-mimicry strategy is a generally tolerated and promising one for targeting and reversing abnormal DNA methylation in cancer.

Finally, we interrogate the possible contribution of the DNA demethylated TSGs in mSTELLA-mediated inhibition of the oncogenic functions of UHRF1 as studied above, by focusing on the five representative TSGs that are readily de-repressed by LNP-mSTELLA in HCT116 cells (Figs. 7f, g and S9k). Consistently, LNP-mSTELLA treatment remarkably reduces cell growth and colony formation of CRC cells, which is partially but significantly rescued upon blunting the induction of each individual TSG (*SFRP1*, *UCHL1*, *FBLN2* and *p16*) by shRNA, with least effect observed for *GATA5* (Fig. S10m–o). This type of an approach has been used to study the role of UHRF1 silenced TSGs in CRC growth and metastasis[13]. More importantly, when we prevent the re-expression of *p16* that displays the most effective rescuing effects in vitro, the CRC tumor growth of LNP-mSTELLA treated group is significantly promoted (Figs. 8h, i and S10p). These data corroborate our previous findings that the maintenance of promoter DNA hypermethylation and silencing of a variety of key TSGs by UHRF1 collectively prompts CRC tumorigenicity[13], and importantly, further underline the values of cooperatively targeting the UHRF1 TTD-PHD domains, as achieved by LNP-mSTELLA, for cancer management.

Discussion

We have introduced earlier that developing means to inhibit UHRF1 is an increasingly recognized need as an epigenetic based cancer therapy approach. To provide for discussion of our basic results in this paper and the above translation goal, a brief recapitulation of the study findings is useful. We have uncovered key differences between the hSTELLA and mSTELLA proteins for blocking the ability of UHRF1 to inhibit the maintenance or de novo establishment of cancer-specific DNA methylation in cancer cells. Accompanying this difference is the much weaker capability of hSTELLA to block oncogenic properties of CRC cells as mSTELLA does. We pinpoint these above dynamics to a short C-terminus Swap1 region, in which a single proline/lysine switch might be one of the most important determinants for the ortholog-specific UHRF1 binding modes (Figs. 5a and S5d). This difference contributes to the ability of mSTELLA, but not hSTELLA to form the L-shaped binding conformation which brings together essential

contacts with the UHRF1 PHD and TTD domains (Fig. 6) to inhibit DNA methylation and prevent UHRF1 from acting as a cancer oncogene.

Examining the homologies for STELLA throughout multiple species provides important clues which further accentuate understanding of the structural variations we have defined. Interestingly, although STELLA orthologs are specifically present in boreoeutherian mammals[29], only *Rodentia* (*Mus musculus*, *Rattus norvegicus*) has a proline residue in the short C-terminus peptide we have focused upon. The presence of this residue apparently endows the ability of mSTELLA to form a kink and thereby the L-shaped binding motif which mediates the multifaceted, cooperative interactions with the UHRF1 PHD-TTD domain (Fig. 6). A matching lysine or arginine residue, both having high α-helical propensity[77], exists among other boreoeutherian species, including *Homo sapiens*, *Pan troglodytes*, *Vulpes vulpes* and *Bos taurus* (Fig. S5d). This homology search above suggests the role of STELLA in safeguarding the DNA hypomethylation at transcriptionally inactive regions, an unique and conserved feature of oocytes[101,102], which might be more efficiently achieved in *Rodentia*, but might act through more complex mechanisms in non-rodent species. This may require a sufficiently high amount of STELLA proteins being enriched at the transcriptionally inert loci for antagonizing DNA methylation, as seen from the far less effective UHRF1 inhibitory activities of hSTELLA depending on its cellular expression levels (Figs. S1l, S3h and S3i). Nevertheless, a very small subset of up-regulated genes (UBE2S, PLCD1, ENOSF1, TCEA3, MSI1 and JUN) exhibits promoter hypomethylation following hSTELLA loss in BeWo cells (Fig. 2f), while no genes show such corresponding changes in NCCIT and CRC cells with either hSTELLA depletion or over-expression. This implicates that hSTELLA may suppress TET enzymes[103] to regulate DNA methylation at limited loci through a context-dependent manner in human cancer cells. Furthermore, histone H3K36 demethylases KDM2A/2B may antagonize de novo DNA methylation at the CpG-rich, transcriptionally inert regions in human oocytes[104], implicating proper histone modifications may also help to guard the human oocyte methylome. Notably, our data strongly suggests that the DNA methylation-independent transcription regulatory functions of hSTELLA, as we see in TGCT cells (Figs. 1–2), warrants further studies in both early human embryonic development and cancer models with progenitor-like or slow-cycling properties[31–33]. Such functions may involve the modulation of the higher-order chromatin structures, which has been explored for mSTELLA in mouse female germline stem cells[23], and the delicate coordination with key transcriptional factors or histone modifications[23,33].

The UHRF1 TTD-PHD domains function as a single unit to cooperatively recognize H3K9me3 tail, essential for both the maintenance and oncogenic functions of UHRF1[13,68,79]. Having now defined such histone binding module as a potentially druggable site for UHRF1 inhibitors development, our present findings provide several important insights and evidence for such a goal. First, although the complex structures of UHRF1 TTD-PHD domains or singular TTD/PHD domains with various substrates including H3K9me3, PAF15, DNA ligase 1, and chemical compounds, have been solved[68,78,79,105,106], it is still pharmacologically challenging to competitively inhibit the interactions between UHRF1 and highly redundant histone H3 in cells. To date, no small molecule or peptide inhibitors have been reported to target the histone-binding domains of UHRF1 for DNA demethylation. Here, our structural and functional data clearly indicate that the occupation of the acidic H3 tail binding pocket in the UHRF1 PHD domain alone is insufficient to disrupt the chromatin binding and maintenance functions of UHRF1. Instead, the conformational dynamics between TTD and PHD domains of UHRF1 may set up a target for therapeutic interventions. Such dynamics may be overlooked in structures of STELLA orthologs that interact with the PHD domain alone (Fig. S7c,d)[71,72], but can be more thoroughly studied based on structures of STELLA complexed with the linked TTD-PHD domains of UHRF1, as we reported here. The STELLA orthologs induce differential relative orientations between the TTD and PHD domains, allowing a tight, productive complex with UHRF1 only formed by mSTELLA. This is exemplified by the coordinated contacts mediated by the αL of mSTELLA with both TTD and PHD domains, in comparison with the singular PHD domain, leading to more efficient inhibition of UHRF1-H3K9me3 interactions. In contrast, the extensive TTD-PHD inter-domain rotation attenuates hSTELLA-PHD domain contacts, resulting in negligible cooperation between the TTD and PHD domains when associating with hSTELLA. In this regard, the differential capabilities in cooperative binding to the linked, dynamic TTD-PHD domains underpins the ortholog-specific UHRF1 inhibition by STELLA proteins. This is further validated by our LNP-mediated mRNA delivery strategy wherein replacing the mSTELLA motif that cooperatively interact with UHRF1 (residues 80-119) by the corresponding hSTELLA sequences (residues 75-121) cannot mediate such cooperation, rendering it unable to reverse abnormal DNA methylation and oncogenic functions of UHRF1. Therefore, our findings provide deeper insights into STELLA-UHRF1 recognition than those from structural studies focusing solely on the PHD domain[71,72]. More importantly, our findings emphasize the importance to target the inter-domain dynamics and cooperativity of the histone reader modules to develop potent UHRF1 inhibitors, which could be achieved through rationally designing at least two approaches going forward as detailed below.

First, the nanoparticle delivery approach we now introduce may be developed as an approach to take forward. In this regard, nanoparticle mRNA expression approaches are either in clinical usage and key examples include two current LNP-mRNA vaccines against coronavirus disease 2019 (COVID-19). Meanwhile, a variety of cancer vaccines or immunomodulatory therapeutics are intensely being pursued in ongoing drug developmental efforts. These are also based on designing LNP-mRNA formulations some of which are in clinical trials and demonstrating encouraging clinical benefits[107,108]. Related specifically to our present study, LNP-mSTELLA effectively reverses cancer-specific DNA hypermethylation, reactivates epigenetically silenced TSG, and elicits durable inhibition of CRC tumor growth without obvious toxicity. Furthermore, an interferon-driving anti-tumor immune response is provoked by LNP-mSTELLA, suggesting its value in improving tumor microenvironment dynamics and sensitizing immune checkpoint inhibitors[65]. Based on these proof-of-principle findings, sequence optimization in both regulatory elements (*e.g.* untranslated regions in 5′ or/and 3′ end) and coding regions of mSTELLA mRNA may further improve the nanoparticle delivery expression efficiency and therapeutic windows, while also circumventing the immunogenicity arising from the exogenous mRNA sequence per se or its translated protein from other species[99,100,109,110]. It is noteworthy that, although our current treatment regimen demonstrates low activity in inducing an acute, high lever release of pro-inflammatory cytokines with reduced tolerability, this issue may need further investigation during long-term exposure. Finally, in vivo tumor-specific delivery carriers for such optimized STELLA mRNA sequences, through modulating the lipid structures or surface modification of LNP[111,112], may lead to more effective and safer mRNA-based hypomethylating agents for preclinical evaluation, either alone or in combination with immune checkpoint therapy.

Second, another translational possibility from our present study could move forward exploiting the potential to rationally develop a peptidomimetic or small molecule mimicry by utilizing a structure-based drug discovery (SBDD) approach[113,114]. These inhibitors could be designed, based on our TTD-PHD/mSTE-2 complex, to mimic the essential, L-shaped pharmacophores of mSTE-2 by cooperatively interacting with both the H3 tail binding pocket and the interdomain cavity of UHRF1 TTD-PHD domains. While the reciprocal synergy between the UHRF1 histone- and DNA-binding domains orchestrates UHRF1-

chromatin association[115,116], it will be interesting to also study the allosteric effects of these mSTELLA-mimicking, histone-binding domain inhibitors in interfering with the essential function of the hemi-methylated DNA binding SRA domain. Nevertheless, a DNA hypomethylating agent may demethylate and upregulate an oncogene in some patients with myelodysplastic syndrome[117], suggesting that monitoring oncogene activity might still be warranted during hypomethylating therapy, such as the mSTELLA-mimicry strategy proposed here.

In summary, we delineate the distinct differences of STELLA orthologs in antagonizing DNA methylation abnormalities in cancer cells through UHRF1 inhibition, and how these species-specificities drive the development of cancer management strategies.

Methods

Ethics statement

This study complies with all relevant ethical regulations of Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences. All animal experiment protocols were approved and conducted in accordance with guidelines of the Institutional Animal Care and Use Committee (IACUC, No. 2020100) of the Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences.

In vivo mouse studies

All animal experiment protocols were approved and conducted in accordance with guidelines of the Institutional Animal Care and Use Committee (IACUC, No. 2020100) of the Guangzhou Institutes of Biomedicine and Health. Mice were housed under standard conditions in the specific pathogen-free (SPF)-grade animal facility with 12 h/12 h light/dark cycles, a temperature of 21-26 °C and humidity of 40-70%. We used 5- to 6-week-old male NOD-Prkdc^{em26Cd52}Il2rg^{em26Cd22}/Gpt (NCG) mice (GemPharmatech) in these studies. To align the gender of the patients from whom the cell lines used for the xenograft models were isolated and to avoid potential impacts of the physiological cycle of female mice on the experimental results, we chose male mice for our studies. 1×10^6 viable HCT116 and RKO cells expressing the indicated constructs (for experiments in Figs. 4a–c, S4d–f and 8h, i) or untreated HCT116 cells (for experiments in Figs. 8d–g and S10f–i) in 0.1 mL of PBS with Matrigel (1:1) were subcutaneously injected into one flank of each mouse. The number of mice in each of the above experiments has been indicated in the figure legends. For the in vivo LNP experiments, once tumors were palpable, mice were randomized into each group with the indicated dosing. Mice were peritumorally injected with 0.1 mL LNP-mRNA formulations or LNP vehicle control every other day for the specified duration, and the tumor growth of LNP-mSTELLA groups was monitored for an additional five-week observation period after dosing ended as depicted in Fig. 8f. Tumor volume was measured every two days using digital calipers and calculated with the formula: $0.5 \times (L \times W \times H)$. Mice were euthanized when tumor volume was approximately 2000 mm^3 or as indicated, or at the end of observation period in the LNP-mRNA experiments, whichever came first. During the course of the experiment, mice were monitored every other day for changes in tumor volume, weight loss and other signs of distress. This monitoring was conducted daily when the tumor volume was between 1500 mm^3 and 2000 mm^3 to avoid the following distressing symptoms that necessitate euthanizing the animals: > 20% body weight loss, poor grooming and mobility, hypothermia, hunching, and significantly reduced food/water consumption, and to ensure that the tumor size in all experiments did not exceed the maximum tumor size (2000 mm^3) permitted by the approved protocols, as shown in the Source Data file.

Cell culture

Cell lines BeWo (ATCC CCL-98), NCCIT (ATCC CRL-2073), HCT116 (ATCC CCL-247), RKO (ATCC CRL-2577), MDA-MB-231 (ATCC HTB-26), MDA-MB-453 (ATCC HTB-131), A549 (ATCC CCL-185) and THP-1 (ATCC TIB-202) were obtained from ATCC. NCI-H1299 cells (SCSP-589) were obtained from the Cell Bank of the Chinese Academy of Sciences. NCCIT, NCI-H1299, and THP-1 cells were cultured in RPMI-1640 medium, BeWo cells in Ham's F-12K medium, HCT116 cells in McCoy's 5 A medium, and RKO cells in Minimum Essential Medium (MEM). MDA-MB-231, MDA-MB-453 and A549 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM). All mediums were supplemented with FBS at 10% v/v. Cells were maintained at 37 °C with 5% CO_2 and were periodically tested mycoplasma, ensuring they remained negative. Cell line identity was confirmed using short tandem repeat (STR) profiling.

Lentivirus-mediated shRNA knockdown, CRISPR knockout and stable over-expression

The lentiviral shRNA constructs were obtained from Sigma-Aldrich. For CRISPR/Cas9-mediated genome editing, sgRNAs were synthesized and cloned into the lentiCRIPSR v2 vector (Addgene #52961) based on the recommended protocol from Addgene (Cambridge, MA, USA). The complete list of shRNA and gRNA sequences were shown in Table S4. The cDNA of full-length mSTELLA (UniProt: Q8QZY3) and hSTELLA (UniProt: Q6W0C5) was obtained from VectorBuilder. The swapping mutants (mSTELLA-Swap1 and mSTELLA-Swap2) were generated using the ClonExpress II One Step Cloning Kit (Vazyme). These cDNAs were then subcloned into pLenti-III-EF1alpha vector with a Puromycin selection marker (Applied Biological Materials #LV043). The lentiviral plasmids, and packing plasmids including psPAX2 and pMD2.G, were co-transfected into HEK293T cells by using Lipofectamine 3000 (Thermo Fisher Scientific). The lentiviruses were harvested three times at 24 h, 48 h and 72 h post transfection, and filtered through a 0.45 μm filter. The viruses were concentrated by PEG 8000 and stored at −80°C.

TGCT cell lines (BeWo and NCCIT), CRC cell lines (HCT116, RKO and HCT116-UHRF1$^{+/-}$), lung cancer cell lines (A549 and NCI-H1299) and breast cancer cell lines (MDA-MB-231 and MDA-MB-453) were transduced with lentiviruses expressing the indicated shRNAs/sgRNAs and cDNAs. After three days, the cells were selected with 1–2 μg/mL Puromycin (InvivoGen) for 48 h to establish stable cell pools. The selected stable cell pools were harvested at day 9 post transduction and used for the following experiments. For CRISPR knockout in TGCT cells, single-cell clones were selected through serial dilution to ensure clonal purity. These clones were then verified by western blot and sequencing.

Subcellular fractionation isolation

The subcellular fractionation isolation assay was performed to separate cellular components into various fractions[118,119]. Cell pellets were homogenized on ice in Buffer A [10 mM HEPES (pH 8.0), 10 mM KCl, 1.5 mM $MgCl_2$, 0.34 M sucrose, 10% glycerol, 0.2% IGEPAL (v/v), 1 × protease inhibitor cocktail (TargetMol), 1 × phosphatase Inhibitor cocktail (Sigma-Aldrich), N-ethyl-maleimide (Sigma-Aldrich)], and nuclei were collected by centrifugation (4000 × g for 10 min at 4 °C), washed with Buffer B (Buffer A without IGEPAL). The collected supernatant was used as the cytoplasmic fraction. For the preparation of total nuclear protein extracts, the nuclear pellets were lysed with 4% SDS (w/v) and passed through homogenizer columns (Omega). Instead, the nuclear pellets were lysed in Buffer C (3 mM EDTA, 2 mM EGTA, inhibitors) for 30 min at 4 °C, and the nuclear soluble fractions were separated from the tight chromatin pellets by centrifugation (14000 × *g* for 10 min at 4°C). The tight chromatin pellets were lysed with 4% SDS as aforementioned. GAPDH and Lamin B1 immunoblotting served as cytoplasmic and nuclear controls, respectively.

Immunofluorescence staining

BeWo and NCCIT cells with hSTELLA-depletion, and CRC cells (HCT116, RKO, HCT116-UHRF1$^{+/-}$) stably expressing exogenous Flag-tagged proteins were cultured on gelatin-coated coverslips to enhance cell adhesion. Cells were washed twice with PBS and fixed with 4% (v/v) paraformaldehyde in PBS for 30 min at room temperature. Coverslips were rinsed three times with PBS and then

permeabilized with 0.5% (v/v) Triton X-100-containing PBS for 10 min at room temperature. Following permeabilization, coverslips were rinsed three times with PBS and cells were blocked with 1% bovine serum albumin for 1 h and then incubated with specific primary antibodies overnight at 4 °C. The following antibodies were used for immunofluorescence staining: mouse anti-STELLA (1:100, Santa Cruz, sc-376862), rabbit anti-STELLA (1:100, Invitrogen, PA5-34601), rabbit anti-UHRF1 (1:100, Thermo Fisher Scientific, PA5-29884), mouse anti-UHRF1 (1:100, Santa Cruz, sc-373750), rabbit anti-DNMT1 (1:1000, Abcam, ab92314), rabbit anti-DNMT3A (1:100, Sigma-Aldrich, HPA026588), rabbit anti-DNMT3B [1:500, A rabbit polyclonal antibody (QCB/BioSource International) raised against a fusion protein containing residues 376–390 (ENKTRRRTADDSATS)], rabbit anti-Flag (1:400, Sigma-Aldrich, F7425), and mouse anti-Flag (1:400, Sigma-Aldrich, F1804). After rinsing three times with PBS, the coverslips were incubated with Alexa Fluor 488- or 568-conjugated secondary antibodies (1:500; Thermo Fisher Scientific) along with 1 μg/mL DAPI for nucleus staining. The incubation was carried out at room temperature for 1 h, protected from light. The slides were mounted with fluorescence mounting medium (DAKO) and kept from light at 4 °C. The images were visualized and captured by an LSM 800 (Carl Zeiss) laser scanning confocal microscope (40× magnification). Nuclear localization levels were analyzed using ImageJ software, measuring integrated signal intensities in the mid-section of nucleus with maximal area for calculating the nuclear distribution ratio ($Intensity_{nucleus}/Intensity_{whole\_cell}$).

Co-immunoprecipitation

The Flag-tagged mSTELLA point mutants (V88A, R89A/T90A, K101A, R104A, L105A, R107A, R111A, I99A, M102A, R103A, I99A/M102A/R103A) and swapping mutants (mSTELLA-Swap1 and mSTELLA-Swap2) were generated by using the ClonExpress II One Step Cloning Kit (Vazyme). These constructs were transiently transfected into HEK293T cells in exponential growth phase. Cells were harvested at 48 h post-transfection. Collected cell pellets were washed with prechilled PBS, and then resuspended in prechilled Cell lysis buffer [20 mM Tris-HCl (pH 8.0), 150 mM NaCl, 2.5 mM EDTA (pH 8.0), 1% IGEPAL (v/v)] with phosphatase and protease inhibitors added. The suspension was incubated on ice for 30 min, followed by sonication using a Bioruptor sonicator (Diagenode SA, Belgium) to achieve efficient cell lysis. Cell lysates were cleared by centrifugation at 12000 g for 20 min at 4°C, and the supernatants were diluted into IP dilution buffer [20 mM Tris-HCl (pH 8.0), 150 mM NaCl, 2.5 mM EDTA (pH 8.0)] for immunoprecipitations. 5% input was taken for each sample. The diluted extracts were incubated with anti-UHRF1 antibody (1:200, Thermo Fisher Scientific, PA5-29884) or anti-Flag antibody (1:400, Sigma-Aldrich, F1804) at 4°C overnight. The antibody-protein complexes were captured by Protein A/G magnetic beads (DynaBeads) at 4°C for 3 h. The beads were washed three times with the Washing buffer [20 mM Tris-HCl (pH 8.0), 150 mM NaCl, 2.5 mM EDTA (pH 8.0), 0.5% IGEPAL(v/v)] for 5 min at 4°C. Complexes were eluted and denatured with SDS Laemmli buffer followed by SDS-PAGE separation and immunoblot with the indicated antibodies.

Western blot assay

Cells were washed in ice-cold PBS, and then lysed by 4% SDS. Cell lysates were homogenized with homogenizer columns (Omega) and total proteins were quantified by Pierce BCA Protein Assay (Thermo Fisher Scientific). Samples were separated by 10% SDS-PAGE, transferred onto a PVDF membrane (Sigma-Aldrich), and blocked in 5% non-fat dry milk in TBST (Tris-buffered saline with Tween 20) solution. The membrane was incubated overnight at 4°C with the following antibodies: anti-STELLA (1:300, Sigma-Aldrich, MAB4388), anti-UHRF1 (1:5000, Sigma-Aldrich, MABE308), anti-DNMT3A (1:1000, Novusbio, NB120-13888), anti-DNMT3B [1:1000, A rabbit polyclonal antibody (QCB/BioSource International) raised against a fusion protein containing residues 376–390 (ENKTRRRTADDSATS)], anti-DNMT1 (1:5000, Sigma-Aldrich, D4692), anti-Flag (1:1000, Sigma-Aldrich, F1804), anti-p16 (1:1000, Abcam, ab108349), anti-Lamin B1 (1:1000, Santa Cruz, sc-377000), anti-GAPDH (1:5000, Bioworld, AP0063), anti-SFRP1 (1:1000, Abcam, ab267466), anti-GATA5 (1:500, Proteintech, 55433-1-AP), anti-UCHL1 (1:1000, CST, 13179), anti-FBLN2 (1:500, Novusbio, NBP1-33479), anti-TET2 (1:1000, Sigma, MABE462), anti-TET3 (1:1000, GeneTex, GTX121453) and anti-β-actin (1:5000, Sigma-Aldrich, A5441). The membrane was washed in TBST solution for five times, each for 5 min. The membrane was then incubated with a horseradish peroxidase (HRP)-conjugated secondary antibody, incubated with ECL substrate (Pierce) and developed.

Dot blot assay

Genomic DNA was extracted utilizing the Wizard Genomic DNA Purification Kit (Promega) and rehydrated in TE buffer. The genomic DNA was denatured by treatment with 0.4 M NaOH and 10 mM EDTA at 95 °C for 10 min, followed by neutralization with an equal volume of cold 2 M NH_4OAc (pH 7.0). The samples were applied to a pre-wetted nitrocellulose membrane in two-fold serial dilutions by using a Bio-Dot Apparatus Assembly (Bio-Rad). The membrane was briefly rinsed in 2 × SSC solution, air-dried, baked at 65 °C for 5 min, fixed using UV light at 1.2 J/m^2 for 5 min, blocked in 5% non-fat milk for 1 h at room temperature, and incubated overnight at 4°C with anti-5mC antibody (CST, 28692, 1:2000) or anti-5hmC antibody (Active Motif, 39769, 1:2000). After washing in TBST solution for three times (each for 5 min), the membrane was incubated with HRP-conjugated anti-rabbit IgG secondary antibody, incubated with ECL substrate and developed.

Cytokines measurement by enzyme-linked immunosorbent assay (ELISA)

HCT116 and THP-1 cells were seeded in six-well plates. After 24 h, the cells were treated with the indicated LNP formulations (0.1 μg/mL) or vehicle control for 48 h. The lipopolysaccharide (LPS) solution (1 μg/mL) was included as a positive control. The concentrations of IFN-γ, TNF-α, IL-2, and IL-6 in the supernatants were quantified using human-specific ELISA kits: IFN-γ (Abclonal, RK00015), TNF-α (Abclonal, RK00030), IL-2 (Abclonal, RK04123), and IL-6 (Abclonal, RK00004), following the protocols provided by the manufacturer. Optical density at 450 nm was measured using a microplate reader (Bio Tek), with a reference wavelength of 630 nm set to correct for baseline absorbance.

Cell proliferation, colony formation and soft agar colony formation assay

Single-cell suspensions of CRC cells (HCT116 and RKO), lung cancer cells (A549 and NCI-H1299) and breast cancer cells (MDA-MB-231 and MDA-MB-453) expressing the indicated constructs were seeded into six-well plates at a density of 0.8×10^4 cells per well (NCI-H1299), 1×10^4 cells per well (HCT116 and RKO), 3×10^4 cells per well (A549 and MDA-MB-231), and 5×10^4 cells per well (MDA-MB-453). The cell numbers were counted daily using a hemocytometer from Day 1 to Day 5. For the LNP-mRNA treatment assay, HCT116 cells were pretreated with LNP-mRNA or vehicle control for 4 days. Subsequently, cells were seeded into six-well plates with complete medium containing LNP-mRNA or vehicle control for further analysis. The cell numbers were calculated similarly.

For colony formation assays, suspensions of 1000 single cells were seeded into six-well plates, shaken, and cultured at 37 °C with 5% CO_2. After 10 days, the medium was removed, and cells were stained with 0.1% crystal violet (dissolved in autoclaved water with 20% MeOH) for visualization and quantification. In the soft agar colony formation assays, 5000 single cells were seeded in 0.4% soft agar medium and cultured for 14-21 days. Colonies were then fixed with 4% formaldehyde and stained with 1 μg/mL ethidium bromide. The area

covered by the colonies was calculated using the ColonyArea plugin in ImageJ software[120].

Immunohistochemistry and H&E staining

Histological sections were isolated from murine xenografts or the major organs, immunohistochemistry and H&E staining were performed as follows. The tissues were fixed in 10% formalin overnight at room temperature and then placed in 70% ethanol, embedded in paraffin according to standard protocols. For immunohistochemistry staining, slides were incubated with primary antibody (anti-Ki67, Servicebio GB111141, 1:1000) at 4 °C overnight. DAB chromogen was used to develop signals. For H&E staining, the slides were stained with hematoxylin and eosin using standard protocols. Subsequently, images of the stained slides were acquired using a LEICA DMi8 microscope. The area of Ki67-positive cells was quantitated using the ImageJ software.

Protein expression and purification

The cDNA encoding the PHD domain (298-367), TTD-PHD (134-366), and TTD-PHD (deletion 167-175) of UHRF1 were each cloned into pGEX 6P-1. The protein was expressed in *E.coli* strain Rosetta2 at 16 °C overnight after induction with 0.3 mM IPTG. Cells were collected and lysed in lysis buffer (20 mM Tris pH 8.0, 300 mM NaCl, 1 mM TCEP, 50 μM zinc acetate). The protein was purified using a glutathione affinity column with standard protocols. The GST tag was removed by PreScission Protease (PSP). The protein was further purified by Source Q, and then followed by gel filtration using a Superdex 75 column (GE Healthcare) equilibrated against a buffer containing 20 mM Tris (pH 7.5), 100 mM NaCl and 0.5 mM TCEP.

Crystallization, data collection and structure determination

For crystallization, PHD domain or TTD-PHD (deletion 167-175) was mixed with peptide (mSTE-2 or hSTE-1) with a mole ratio of 1:1.5. The complex was then concentrated to 20 mg/mL. Crystals were grown at 293 K with sitting drop vapor diffusion method. The crystals of PHD/mSTE-2 were grown in 0.1 M Potassium thiocyanate, 30% PEG MME2000. The crystals of TTD-PHD/mSTE-2 were grown in 0.6 M NaCl, 0.1 M MES (pH 6.5), 20% PEG 4000. The crystals of PHD/hSTE-1 were grown in 0.2 M Ammonium acetate, 0.1 M HEPES (pH 7.5), 25% PEG 3350. The crystals of TTD-PHD/hSTE-1 were grown in 4% v/v Tacsimate (pH 7.0), 12% PEG 3350.

All crystals were cryo-protected in mother liquid with addition of 30% glycerol, prior to collecting data at 100 K. Diffraction data for TTD-PHD/mSTE-2 was collected at beam line 02U1 of the Shanghai Synchrotron Radiation Facility. Diffraction data for PHD/mSTE-2, PHD/hSTE-1 and TTD-PHD/hSTE-1 were collected at beam line 19U1 of the Shanghai Synchrotron Radiation Facility[121]. Data for PHD/mSTE-2 and TTD-PHD/mSTE-2 were auto-processed by Aquarium pipeline[122], and date for PHD/hSTE-1 and TTD-PHD/hSTE-1 were processed with mosflm[123]. All processed data were scaled with Aimless[124]. The structure of TTD-PHD/mSTE-2 was solved by molecular replacement using structure of TTD-PHD/H3K9me3 (PDB code 3ASK) as search mode. The structures of PHD/ mSTE-2, PHD/ hSTE-1 and TTD-PHD/ hSTE-1, were solved by molecular replacement with structure of TTD-PHD/mSTE-2 as search mode. The refinement of the structures was conducted using Refmac5[125]. Data collection and refinement statistics are shown in Table S1.

Isothermal titration calorimetry (ITC)

ITC experiments were performed using a MicroCal PEAQ-ITC instrument. The proteins and peptides were dissolved in the buffer containing 20 mM Tris (pH 7.5), 100 mM NaCl, 0.5 mM TCEP. The measurements were conducted at 293 K. Protein in syringe at concentration of 300–400 μM was titrated into peptide solution in cell at concentration of 20–45 μM in 20 × 2 μL. The data were analyzed with the MicroCal PEAQ-ITC analysis software using a one-site model.

Real-Time qPCR

Total RNA was extracted by using the RNAiso Plus reagent (Takara), and 1 μg of total RNA was then utilized for cDNA synthesis with the PrimeScript™ RT Master Mix (Takara). For qPCR assays, cDNA aliquots were utilized for detecting the differential expression of *UCHL1*, *SFRP1*, *GATA5* and *FBLN2* using the SYBR green PCR kit (Takara). *β-actin* was selected as the reference gene for normalization in qPCR assays. The ΔΔCt method was used to calculate the fold change over control cells. All qPCR assays were performed in triplicate. The primer sequences used in qPCR assays are provided in Table S5.

Peptidic nanoparticle (PNP) formulation

1,2-distearoyl-sn-glycero-3-phosphocholine (DSPC), 1,2-dimyristoyl-rac-glycero-3-methoxypolyethylene glycol-2000 (DMG-PEG2000), and L-α-Phosphatidylethanolamine-N-(lissamine rhodamine B sulfonyl) (Rhod PE) were obtained from Avanti (USA). Cholesterol was purchased from Sigma-Aldrich (USA). The palmitic acid (Pal), which forms the nanoparticles with other helper lipids (DSPC, cholesterol, DMG-PEG2000), were conjugated with the nuclear localization signal (NLS) peptide derived from SV40 large T-antigen that facilitates the PNP to enter the nucleus where DNA hypermethylation occurs ((Pal)-GPKKKRKV, designated as Pal-NLS). The mSTE-1 peptide were synthesized and HPLC-purified by ChinaPeptides CO., Ltd (Shanghai, China) with purities > 98%. Moreover, Rhod PE was employed as a fluorescence label allowing for the detection of PNP localization in the cells using fluorescence microscopy.

The Pal-NLS peptide, mSTE-1 peptide, DSPC, cholesterol, DMG-PEG2000 and Rhod PE were mixed in ethanol at a molar ratio of 33.3:16.7:10:37.5:1.5:1, respectively. Ethanol was eliminated by freeze-drying technique. Then, dried lipid film was hydrated with PBS and ultrasonicated for complete dispersion. The diameter size, and polydispersity index (PDI) of peptidic nanoparticles of mSTE-1 (PNP-mSTE-1) were analyzed using the dynamic light scattering (DLS) technique (Zetasizer Nano ZS, Malvern Panalytical, USA) at the room temperature.

mRNA synthesis

mRNA encoding the full-length sequence of mSTELLA or mSTELLA-Swap1 was produced by in vitro transcription (IVT) of DNA template using T7 RNA polymerase, which contains the open reading frame encoding the target proteins flanked by 5′ and 3′ untranslated regions. Co-transcriptional capping added cap analogs simultaneously to the 5′ end of mRNA. Poly(A) tails were added by *E. coli* Poly(A) Polymerase I to the 3′ termini of transcribed RNA. RNA quality was analyzed for RNA length, poly (A) length, purity, dsRNA, endotoxin and capping efficiency (Table S2). All the mRNAs were produced with the modified nucleobase N1-methylpseudouridine triphosphate (N1-MepUTP) instead of Uridine-5′-triphosphate (UTP).

Lipid nanoparticle (LNP) formulation

Lipid nanoparticles (LNP) were prepared by microfluidic mixing[126]. Briefly, lipids were dissolved in ethanol at a molar ratio of 48.5:10:40:1.5 for ionizable lipid, DSPC, cholesterol, and DMG-PEG2000. The lipid mixture was rapidly combined with a buffer of 50 mM sodium citrate (pH 5.0) containing mRNA at a volume ratio 1:3 of aqueous: ethanol using a microfluidic mixer (Inano E, Micro&Nano, Shanghai, China). Formulations were dialyzed against PBS (pH 7.2) in the dialysis cassettes (Thermo Scientific, Rockford, IL, USA) for at least 18 h. Formulations were diluted with PBS (pH 7.2) to reach a required concentration and then passed through a 0.22 μm filter and stored at 4 °C until use. To analyze the particle size and polydispersity index (PDI) (Table S3), the LNP formulations were diluted with PBS (pH 7.2) to reach a mRNA concentration of 5 μg/mL, which were subjected to Dynamic light scattering (DLS) analyzer (Zetasizer Nano ZS, Malvern Panalytical, USA). Data analysis was performed by Zetasizer Software (v. 7.11). The amount of mRNA was quantified by Quant-it™ RiboGreen

RNA Assay Kit (Invitrogen™, Thermo Fisher Scientific, Paisley, UK), for which the unencapsulated mRNA was obtained from LNP solution diluted in TE buffer, and the total mRNA was obtained from LNP solution diluted in TE buffer supplemented with 1% Triton X-100. The encapsulation efficiency (EE) was calculated by the following formula: EE (%) = (1- unencapsulated mRNA /total mRNA) × 100%.

Genome-wide DNA methylation analysis

Genomic DNA was isolated from cultured cells using the Wizard Genomic DNA Purification Kit (Promega) following the manufacturer's instructions. Extracted genomic DNA was bisulfite converted using EZ DNA Methylation Kit (Zymo Research)[127]. Genome wide DNA methylation status was determined using Illumina's Infinium MethylationEPIC BeadChip v1.0 (Figs. 2c, f, 3c–e, 3g, 7b, d, f, S1i, S1l, S3a,b, S3d,e, S3h, S9d,e, S9h,i, S9k) and v2.0 microarray (Figs. S4i and S10g-h). The Bioconductor minfi and SeSAMe packages were used to preprocess and normalize the EPIC v1.0 and EPIC v2.0 raw data, respectively[128,129]. Probes with detection p values > 0.01 were excluded from subsequent analysis and methylation β-values were defined as following: β value = (signal intensity of methylation-detection probe)/(signal intensity of methylation-detection probe + signal intensity of non-methylation-detection probe), ranging from 0 to 1.0 for each CpG site (methylation level from 0-100%, respectively). Probes were initially mapped to the Human Genome Assembly GRCh37/hg19, and annotated using the IlluminaHumanMethylationEPICanno.ilm10b2.hg19 package. The enhancer probes, as well as cancer-specific and non-cancer methylation probes defined in our previous study[13], were used.

RNA-seq based transcriptome analysis

Total RNA was isolated from cell pallets using RNAiso Plus reagent (Takara). After RNA extraction, purity was quantified with NanoDrop followed by integrity evaluation via Agilent 4200 TapeStation analysis according to manufacturer's protocol. For mRNA library preparation, 1 μg of total RNA per sample was utilized as input material. The mRNA isolation was using oligo (dT) magnetic beads and fragmentation was carried out using divalent cations. The first strand of cDNA was synthesized using a reverse transcriptase system and the second strand cDNA synthesis was performed using DNA Polymerase I and RNase H. The resulting double-stranded cDNA underwent sequential steps including end repair, adenylation of 3' ends, and ligation of sequencing adapters. The library fragments were then purified with AMPure XP system (Beckman Coulter, Beverly, USA) to select cDNA fragments of preferentially 200-300 bp in length. This was followed by PCR amplification using High-Fidelity DNA polymerase, and the PCR product was subjected to further purification to yield the final library. The quality of the library was subsequently assessed using Agilent 4200 TapeStation.

Adapter removal and low-quality read filtering were performed using Fastp v0.21.0[130]. Mapping of clean reads to the GRCh38 reference genome was conducted using HISAT2 v2.0.4[131]. StringTie v2.1.4 was employed to quantify gene expression, generating gene counts and normalized expression values (TPM)[132]. Differential expression analysis was performed using the DESeq2 package v4.3[133], applying the criteria of absolute $\log_2$ (fold change relative to control) > 1 and padj <0.05. Visualization and illustration of specific gene loci were conducted using the Integrative Genomics View (IGV) genome browser v2.16x.

The cancer genome atlas (TCGA) analysis

The DNA methylation and RNA-seq data of TCGA testicular germ cell tumors (TGCT) and colon adenocarcinoma (COAD) studies were obtained from Broad Institute TCGA Genome Data Analysis Center [doi:10.7908/C11G0KM9] and UCSC Xena Hub, respectively[36]. Four histologic subtypes (seminoma, embryonal carcinoma, teratoma and yolk sac tumor) from 128 primary TGCT tumors as previously reported were used for analysis[38]. Pearson Wilcox test was used to determine the p-values. The limma package v3.50.3 was employed for TCGA-COAD RNA-seq analysis to identify genes that were down-regulated (fold change (normal / tumor) ≥ 1.5 and p value < 0.05) in CRC tissues (n = 471) compared to adjacent normal tissues (n = 41)[134]. For DNA methylation analysis, the probes in the Infinium Human Methylation 450 K array available in TCGA-COAD project were used. Probes were mapped to Human Genome Assembly GRCh37/hg19 reference genome and were annotated as described above. The mean DNA methylation changes ($\Delta\beta_{promoter}$) for the probes within CpG island promoter regions between CRC tumors and adjacent normal tissues were determined. Genes with DNA hypermethylated CpG island promoters ($\Delta\beta_{promoter}$ (tumor - normal) > 0.15) were retained.

Enrichment analysis and gene set enrichment analysis (GSEA)

Gene Ontology (GO) enrichment analysis of differentially expressed genes (DEGs) was performed by Metascape webserver (https://metascape.org/gp/index.html) with default settings[135]. A pre-ranked gene list, based on their expression changes, was utilized in GSEA. The GSEA analysis was performed on the gene sets in Molecular Signatures Database (MSigDB) using GSEA v.4.2.3 from the Broad Institute (https://www.gsea-msigdb.org/gsea/index.jsp), and the figures were generated by clusterProfiler R package (Bioconductor)[136]. Gene sets with False Discovery Rate (FDR) < 0.25 and absolute Normalized Enrichment Score (NES) value > 1.2 were taken as significantly enriched terms.

Statistics & reproducibility

Statistical analyses were conducted by using GraphPad Prism (v8.0.1) software (http://www.graphpad.com/), Microsoft Excel (v2409) or R statistical software (v4.3.3, http://www.r-project.org). P-values of less than 0.05 were considered statistically significant. Pearson's correlation analysis was used to determine the two-sided p-values and correlation coefficients (R) in Fig. 1c. DESeq2 (v4.3) was used to determine the DEGs ($|\log_2$ fold-change$|$ > 1 and two-sided Wald test, Benjamini-Hochberg adjusted p-value < 0.05), and the empirical Bayes moderated t-statistic (two-sided) in limma package (v3.50.3) was used for the expression analysis of CRC_TSGs based on TCGA_COAD studies in Supplementary Data 3. Two-sided cumulative hypergeometric distribution was used in the Metascape-based Gene Ontology (GO) enrichment analysis. Paired or unpaired two-sided Wilcoxon rank-sum tests were used for other DNA methylome and transcriptional analysis. Log-rank (Mantel-Cox) tests were used for mice survival analysis in xenograft experiments. Unpaired two-sided t-tests were used for comparisons between two groups in other experiments. No statistical methods were used to predetermine sample size. All boxplots are shown without outliers to facilitate data visualization unless otherwise noted. All immunoblotting analysis were repeated at least twice with consistent results. The immunofluorescence staining experiments performed on BeWo, HCT116, RKO and HCT116-UHRF1$^{+/-}$ cells were repeated three times with consistent results, and representative results are shown. The Ki67 staining for the histological sections from HCT116 xenograft tumors was performed on five mice in each group with consistent results, and representative results from one mouse are shown. The H&E staining for the tissue sections from HCT116 mice xenografts was performed on three mice in each treatment group with consistent results, and representative results from one mouse are shown. The ITC analyses were repeated at least twice with consistent results, and representative figures from one experiment are shown. The colony formation and soft agar assays were repeated three times with consistent results, and representative results from one experiment are shown. All other experiments were repeated as indicated in the figure legends.

Reporting summary

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

Data availability

Structure coordinates have been deposited in the Protein Data Bank (https://www.rcsb.org/) with accession code 8XV4, 8XV6, 8XV7, and 8XV8. The DNA methylation and gene expression data have been deposited in the Gene Expression Omnibus (GEO) database with accession code: GSE255083, GSE255105, and GSE273176. The publicly available data on DNA methylation and gene expression of TCGA testicular germ cell tumors (TGCT) studies were obtained from Broad Institute TCGA Genome Data Analysis Center (https://gdac.broadinstitute.org/runs/stddata_2016_01_28/data/TGCT/20160128/) [doi:10.7908/C11G0KM9]. The publicly available data on DNA methylation and gene expression of TCGA colon adenocarcinoma (COAD) studies were obtained from UCSC Xena Hub (https://xenabrowser.net/)[36]. The publicly available data of DNA methylation in CRC cells with endogenous UHRF1 replaced by exogenous WT/domain mutants were downloaded and reanalyzed from the GEO database under the accession GSE118970[13]. The publicly available data of the annotated TSGs were obtained from the TSGene knowledgebase (https://bioinfo.uth.edu/TSGene/)[44]. The publicly available proteomic data of the indicated cancer types from CPTAC studies were obtained from LinkedOmicsKB knowledgebase (https://kb.linkedomics.org/)[47]. The publicly available data of the annotated oncogenes were obtained from the cancer gene list (updated on 07/04/2024) from OncoKB knowledgebase (https://www.oncokb.org/cancer-genes)[56]. The remaining data are available with the Article, Supplementary Information and source data file. Source data are provided with this paper.

References

1. Baylin, S. B. & Jones, P. A. Epigenetic Determinants of Cancer. *Cold Spring Harb. Perspect. Biol.* **8**, https://doi.org/10.1101/cshperspect.a019505 (2016).
2. De Carvalho, D. D. et al. DNA methylation screening identifies driver epigenetic events of cancer cell survival. *Cancer Cell* **21**, 655–667 (2012).
3. Hanahan, D. Hallmarks of cancer: new dimensions. *Cancer Discov.* **12**, 31–46 (2022).
4. Tao, Y. et al. Aging-like spontaneous epigenetic silencing facilitates wnt activation, stemness, and braf(v600e)-induced tumorigenesis. *Cancer Cell* **35**, 315–328 e316 (2019).
5. Citterio, E. et al. Np95 is a histone-binding protein endowed with ubiquitin ligase activity. *Mol. Cell Biol.* **24**, 2526–2535 (2004).
6. Bostick, M. et al. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317**, 1760–1764 (2007).
7. Sharif, J. et al. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450**, 908–912 (2007).
8. Cai, Y. et al. Critical threshold levels of DNA methyltransferase 1 are required to maintain DNA methylation across the genome in human cancer cells. *Genome Res* **27**, 533–544 (2017).
9. Jones, P. A., Issa, J. P. & Baylin, S. Targeting the cancer epigenome for therapy. *Nat. Rev. Genet* **17**, 630–641 (2016).
10. Loo Yau, H., Ettayebi, I. & De Carvalho, D. D. The cancer epigenome: exploiting its vulnerabilities for immunotherapy. *Trends Cell Biol.* **29**, 31–43 (2019).
11. Hogg, S. J., Beavis, P. A., Dawson, M. A. & Johnstone, R. W. Targeting the epigenetic regulation of antitumour immunity. *Nat. Rev. Drug Discov.* **19**, 776–800 (2020).
12. Morel, D., Jeffery, D., Aspeslagh, S., Almouzni, G. & Postel-Vinay, S. Combining epigenetic drugs with other therapies for solid tumours - past lessons and future promise. *Nat. Rev. Clin. Oncol.* **17**, 91–107 (2020).
13. Kong, X. et al. Defining UHRF1 domains that support maintenance of human colon cancer dna methylation and oncogenic properties. *Cancer Cell* **35**, 633–648 e637 (2019).
14. Zhao, S., Allis, C. D. & Wang, G. G. The language of chromatin modification in human cancers. *Nat. Rev. Cancer* **21**, 413–430 (2021).
15. Li, Y., Chen, X. & Lu, C. The interplay between DNA and histone methylation: molecular mechanisms and disease implications. *EMBO Rep.* **22**, e51803 (2021).
16. Payer, B. et al. Stella is a maternal effect gene required for normal early development in mice. *Curr. Biol.* **13**, 2110–2117 (2003).
17. Hayashi, K., de Sousa Lopes, S. M. C., Tang, F., Lao, K. & Surani, M. A. Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. *Cell Stem Cell* **3**, 391–401 (2008).
18. Xu, X. et al. Dppa3 expression is critical for generation of fully reprogrammed iPS cells and maintenance of Dlk1-Dio3 imprinting. *Nat. Commun.* **6**, 6008 (2015).
19. Bowles, J., Teasdale, R. P., James, K. & Koopman, P. Dppa3 is a marker of pluripotency and has a human homologue that is expressed in germ cell tumours. *Cytogenet Genome Res* **101**, 261–265 (2003).
20. Nakamura, T. et al. PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* **486**, 415–419 (2012).
21. Nakamura, T. et al. PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat. Cell Biol.* **9**, 64–71 (2007).
22. Wei, X. et al. PGC7 Regulates Genome-Wide DNA Methylation by Regulating ERK-Mediated Subcellular Localization of DNMT1. *Int. J. Mol. Sci.* **24**, https://doi.org/10.3390/ijms24043093 (2023).
23. Hou, C., Zhao, X., Tian, G. G. & Wu, J. Stella regulates the development of female germline stem cells by modulating chromatin structure and DNA methylation. *Int J. Biol. Sci.* **18**, 3006–3018 (2022).
24. Han, L., Ren, C., Zhang, J., Shu, W. & Wang, Q. Differential roles of Stella in the modulation of DNA methylation during oocyte and zygotic development. *Cell Discov.* **5**, 9 (2019).
25. Han, L. et al. Embryonic defects induced by maternal obesity in mice derive from Stella insufficiency in oocytes. *Nat. Genet* **50**, 432–442 (2018).
26. Funaki, S. et al. Inhibition of maintenance DNA methylation by Stella. *Biochem Biophys. Res Commun.* **453**, 455–460 (2014).
27. Nakashima, H. et al. Effects of dppa3 on DNA methylation dynamics during primordial germ cell development in mice. *Biol. Reprod.* **88**, 125 (2013).
28. Li, Y. et al. Stella safeguards the oocyte methylome by preventing de novo methylation mediated by DNMT1. *Nature* **564**, 136–140 (2018).
29. Mulholland, C. B. et al. Recent evolution of a TET-controlled and DPPA3/STELLA-driven pathway of passive DNA demethylation in mammals. *Nat. Commun.* **11**, 5972 (2020).
30. Du, W. et al. Stella protein facilitates DNA demethylation by disrupting the chromatin association of the RING finger-type E3 ubiquitin ligase UHRF1. *J. Biol. Chem.* **294**, 8907–8917 (2019).
31. Cuesta-Borras, E. et al. DPPA3-HIF1alpha axis controls colorectal cancer chemoresistance by imposing a slow cell-cycle phenotype. *Cell Rep.* **42**, 112927 (2023).
32. Yan, Q. et al. PGC7 promotes tumor oncogenic dedifferentiation through remodeling DNA methylation pattern for key developmental transcription factors. *Cell Death Differ.* **28**, 1955–1970 (2021).
33. Mazid, M. A. et al. Rolling back human pluripotent stem cells to an eight-cell embryo-like stage. *Nature* **605**, 315–324 (2022).
34. Uhlen, M. et al. A pathology atlas of the human cancer transcriptome. *Science* **357**, https://doi.org/10.1126/science.aan2507 (2017).
35. Cancer Genome Atlas Research, N. et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nat. Genet* **45**, 1113–1120, (2013).

36. Goldman, M. J. et al. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat. Biotechnol.* **38**, 675–678 (2020).
37. Consortium, G. T. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet* **45**, 580–585 (2013).
38. Shen, H. et al. Integrated molecular characterization of testicular germ cell tumors. *Cell Rep.* **23**, 3392–3406 (2018).
39. Killian, J. K. et al. Imprints and DPPA3 are bypassed during pluripotency- and differentiation-coupled methylation reprogramming in testicular germ cell tumors. *Genome Res* **26**, 1490–1504 (2016).
40. Haffner, M. C. et al. Hypomethylation, endogenous retrovirus expression, and interferon signaling in testicular germ cell tumors. *Proc. Natl Acad. Sci. USA* **115**, E8580–E8582 (2018).
41. Netto, G. J. et al. Global DNA hypomethylation in intratubular germ cell neoplasia and seminoma, but not in nonseminomatous male germ cell tumors. *Mod. Pathol.* **21**, 1337–1344 (2008).
42. Pattillo, R. A. & Gey, G. O. The establishment of a cell line of human hormone-synthesizing trophoblastic cells in vitro. *Cancer Res* **28**, 1231–1236 (1968).
43. Gerri, C., Menchero, S., Mahadevaiah, S. K., Turner, J. M. A. & Niakan, K. K. Human embryogenesis: a comparative perspective. *Annu Rev. Cell Dev. Biol.* **36**, 411–440 (2020).
44. Zhao, M., Kim, P., Mitra, R., Zhao, J. & Zhao, Z. TSGene 2.0: an updated literature-based knowledgebase for tumor suppressor genes. *Nucleic Acids Res* **44**, D1023–1031, (2016).
45. Yang, L. et al. Tissue-location-specific transcription programs drive tumor dependencies in colon cancer. *Nat. Commun.* **15**, 1384 (2024).
46. Ellis, M. J. et al. Connecting genomic alterations to cancer biology with proteomics: the NCI Clinical Proteomic Tumor Analysis Consortium. *Cancer Discov.* **3**, 1108–1112 (2013).
47. Liao, Y. et al. A proteogenomics data-driven knowledge base of human cancer. *Cell Syst.* **14**, 777–787 e775 (2023).
48. Ahmed, D. et al. Epigenetic and genetic features of 24 colon cancer cell lines. *Oncogenesis* **2**, e71 (2013).
49. Suzuki, H. et al. Epigenetic inactivation of SFRP genes allows constitutive WNT signaling in colorectal cancer. *Nat. Genet* **36**, 417–422 (2004).
50. Glockner, S. C. et al. Methylation of TFPI2 in stool DNA: a potential novel biomarker for the detection of colorectal cancer. *Cancer Res* **69**, 4691–4699 (2009).
51. Miotto, E. et al. Frequent aberrant methylation of the CDH4 gene promoter in human colorectal and gastric cancer. *Cancer Res* **64**, 8156–8159 (2004).
52. Akiyama, Y. et al. GATA-4 and GATA-5 transcription factor genes and potential downstream antitumor target genes are epigenetically silenced in colorectal and gastric cancer. *Mol. Cell Biol.* **23**, 8429–8439 (2003).
53. Barrett, C. W. et al. Tumor suppressor function of the plasma glutathione peroxidase gpx3 in colitis-associated carcinoma. *Cancer Res* **73**, 1245–1255 (2013).
54. Shivapurkar, N. et al. Cytoglobin, the newest member of the globin family, functions as a tumor suppressor gene. *Cancer Res* **68**, 7448–7456 (2008).
55. Roulois, D. et al. DNA-demethylating agents target colorectal cancer cells by inducing viral mimicry by endogenous transcripts. *Cell* **162**, 961–973 (2015).
56. Chakravarty, D. et al. OncoKB: A Precision Oncology Knowledge Base. *JCO Precis Oncol.* **2017**, https://doi.org/10.1200/PO.17.00011 (2017).
57. Wan, L. et al. ENL links histone acetylation to oncogenic gene expression in acute myeloid leukaemia. *Nature* **543**, 265–269 (2017).
58. Zuber, J. et al. RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* **478**, 524–528 (2011).
59. Xu, D. et al. LocNES: a computational tool for locating classical NESs in CRM1 cargo proteins. *Bioinformatics* **31**, 1357–1365 (2015).
60. Kosugi, S., Yanagawa, H., Terauchi, R. & Tabata, S. NESmapper: accurate prediction of leucine-rich nuclear export signals using activity-based profiles. *PLoS Comput Biol.* **10**, e1003841 (2014).
61. Kosugi, S., Hasebe, M., Tomita, M. & Yanagawa, H. Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc. Natl Acad. Sci. USA* **106**, 10171–10176 (2009).
62. Nguyen Ba, A. N., Pogoutse, A., Provart, N. & Moses, A. M. NLStradamus: a simple Hidden Markov Model for nuclear localization signal prediction. *BMC Bioinforma.* **10**, 202 (2009).
63. Drozdetskiy, A., Cole, C., Procter, J. & Barton, G. J. JPred4: a protein secondary structure prediction server. *Nucleic Acids Res* **43**, W389–394 (2015).
64. Egger, G. et al. Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival. *Proc. Natl Acad. Sci. USA* **103**, 14080–14085 (2006).
65. Topper, M. J., Vaz, M., Marrone, K. A., Brahmer, J. R. & Baylin, S. B. The emerging role of epigenetic therapeutics in immunooncology. *Nat. Rev. Clin. Oncol.* **17**, 75–90 (2020).
66. Topper, M. J. et al. Epigenetic therapy ties myc depletion to reversing immune evasion and treating lung cancer. *Cell* **171**, 1284–1300 e1221 (2017).
67. Chiappinelli, K. B. et al. Inhibiting DNA methylation causes an interferon response in cancer via dsRNA including endogenous retroviruses. *Cell* **162**, 974–986 (2015).
68. Houliston, R. S. et al. Conformational dynamics of the TTD-PHD histone reader module of the UHRF1 epigenetic regulator reveals multiple histone-binding states, allosteric regulation, and druggability. *J. Biol. Chem.* **292**, 20947–20959 (2017).
69. Rothbart, S. B. et al. Multivalent histone engagement by the linked tandem Tudor and PHD domains of UHRF1 is required for the epigenetic inheritance of DNA methylation. *Genes Dev.* **27**, 1288–1298 (2013).
70. Cheng, J. et al. Structural insight into coordinated recognition of trimethylated histone H3 lysine 9 (H3K9me3) by the plant homeodomain (PHD) and tandem tudor domain (TTD) of UHRF1 (ubiquitin-like, containing PHD and RING finger domains, 1) protein. *J. Biol. Chem.* **288**, 1329–1339 (2013).
71. Hata, K. et al. Structural basis for the unique multifaceted interaction of DPPA3 with the UHRF1 PHD finger. *Nucleic Acids Res* **50**, 12527–12542 (2022).
72. Shiraishi, N. et al. Structure of human DPPA3 bound to the UHRF1 PHD finger reveals its functional and structural differences from mouse DPPA3. *Commun. Biol.* **7**, 746 (2024).
73. Rothbart, S. B. et al. Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nat. Struct. Mol. Biol.* **19**, 1155–1160 (2012).
74. Nady, N. et al. Recognition of multivalent histone states associated with heterochromatin by UHRF1 protein. *J. Biol. Chem.* **286**, 24300–24311 (2011).
75. Shiraishi, N. et al. Structure of human DPPA3 bound to the UHRF1 PHD finger reveals its functional and structural differences from mouse DPPA3. *bioRxiv*, 2024.2001.2011.575164, https://doi.org/10.1101/2024.01.11.575164 (2024).
76. Morris, A. L., MacArthur, M. W., Hutchinson, E. G. & Thornton, J. M. Stereochemical quality of protein structure coordinates. *Proteins* **12**, 345–364 (1992).
77. Pace, C. N. & Scholtz, J. M. A helix propensity scale based on experimental studies of peptides and proteins. *Biophys. J.* **75**, 422–427 (1998).

78. Nishiyama, A. et al. Two distinct modes of DNMT1 recruitment ensure stable maintenance DNA methylation. *Nat. Commun.* **11**, 1222 (2020).
79. Arita, K. et al. Recognition of modification status on a histone H3 tail by linked histone reader modules of the epigenetic regulator UHRF1. *Proc. Natl Acad. Sci. USA* **109**, 12950–12955 (2012).
80. Ma, Q. et al. SARS-CoV-2 bivalent mRNA vaccine with broad protection against variants of concern. *Front Immunol.* **14**, 1195299 (2023).
81. Wang, D. & Wu, L. P. Nanomaterials for delivery of nucleic acid to the central nervous system (CNS). *Mater. Sci. Eng. C. Mater. Biol. Appl* **70**, 1039–1046 (2017).
82. Cancer Genome Atlas, N. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337 (2012).
83. Bernstein, B. E. et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006).
84. Schlesinger, Y. et al. Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. *Nat. Genet* **39**, 232–236 (2007).
85. Easwaran, H. et al. A DNA hypermethylation module for the stem/progenitor cell signature of cancer. *Genome Res* **22**, 837–849 (2012).
86. Xia, L. et al. CHD4 has oncogenic functions in initiating and maintaining epigenetic suppression of multiple tumor suppressor genes. *Cancer Cell* **31**, 653–668 e657 (2017).
87. Hellebrekers, D. M. et al. GATA4 and GATA5 are potential tumor suppressors and biomarkers in colorectal cancer. *Clin. Cancer Res* **15**, 3990–3997 (2009).
88. He, S., Nakada, D. & Morrison, S. J. Mechanisms of stem cell self-renewal. *Annu Rev. Cell Dev. Biol.* **25**, 377–406 (2009).
89. Janzen, V. et al. Stem-cell ageing modified by the cyclin-dependent kinase inhibitor p16INK4a. *Nature* **443**, 421–426 (2006).
90. Ohm, J. E. et al. A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. *Nat. Genet* **39**, 237–242 (2007).
91. Yu, D. H. et al. Targeted p16(Ink4a) epimutation causes tumorigenesis and reduces survival in mice. *J. Clin. Invest* **124**, 3708–3712 (2014).
92. Merlo, A. et al. 5′ CpG island methylation is associated with transcriptional silencing of the tumour suppressor p16/CDKN2/MTS1 in human cancers. *Nat. Med* **1**, 686–692 (1995).
93. Myohanen, S. K., Baylin, S. B. & Herman, J. G. Hypermethylation can selectively silence individual p16ink4A alleles in neoplasia. *Cancer Res* **58**, 591–593 (1998).
94. Sheng, W. et al. LSD1 ablation stimulates anti-tumor immunity and enables checkpoint blockade. *Cell* **174**, 549–563 e519 (2018).
95. Akinc, A. et al. Targeted delivery of RNAi therapeutics with endogenous and exogenous ligand-based mechanisms. *Mol. Ther.* **18**, 1357–1364 (2010).
96. Akinc, A. et al. The Onpattro story and the clinical translation of nanomedicines containing nucleic acid-based drugs. *Nat. Nanotechnol.* **14**, 1084–1087 (2019).
97. Tsai, H. C. et al. Transient low doses of DNA-demethylating agents exert durable antitumor effects on hematological and epithelial tumor cells. *Cancer Cell* **21**, 430–446 (2012).
98. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6–21 (2002).
99. Dingman, R. & Balu-Iyer, S. V. Immunogenicity of Protein Pharmaceuticals. *J. Pharm. Sci.* **108**, 1637–1654 (2019).
100. Sauna, Z. E., Lagasse, D., Pedras-Vasconcelos, J., Golding, B. & Rosenberg, A. S. Evaluating and Mitigating the Immunogenicity of Therapeutic Proteins. *Trends Biotechnol.* **36**, 1068–1084 (2018).
101. Demond, H. & Kelsey, G. The enigma of DNA methylation in the mammalian oocyte. *F1000Res* **9**, https://doi.org/10.12688/f1000research.21513.1 (2020).
102. Zhu, P. et al. Single-cell DNA methylome sequencing of human preimplantation embryos. *Nat. Genet* **50**, 12–19 (2018).
103. Bian, C. & Yu, X. PGC7 suppresses TET3 for protecting DNA methylation. *Nucleic Acids Res* **42**, 2893–2905 (2014).
104. Yan, R. et al. Decoding dynamic epigenetic landscapes in human oocytes using single-cell multi-omics sequencing. *Cell Stem Cell* **28**, 1641–1656 e1647 (2021).
105. Kori, S. et al. Structure of the UHRF1 tandem tudor domain bound to a methylated non-histone protein, LIG1, reveals rules for binding and regulation. *Structure* **27**, 485–496 e487 (2019).
106. Senisterra, G. et al. Discovery of small-molecule antagonists of the h3k9me3 binding to uhrf1 tandem tudor domain. *SLAS Discov.* **23**, 930–940 (2018).
107. Carvalho, T. Personalized anti-cancer vaccine combining mRNA and immunotherapy tested in melanoma trial. *Nat. Med* **29**, 2379–2380 (2023).
108. Zong, Y., Lin, Y., Wei, T. & Cheng, Q. Lipid nanoparticle (LNP) enables mRNA delivery for cancer therapy. *Adv. Mater.* **35**, e2303261 (2023).
109. Zhang, H. et al. Algorithm for optimized mRNA design improves stability and immunogenicity. *Nature* **621**, 396–403 (2023).
110. Leppek, K. et al. Combinatorial optimization of mRNA structure, stability, and translation for RNA-based therapeutics. *Nat. Commun.* **13**, 1536 (2022).
111. Naidu, G. S. et al. A combinatorial library of lipid nanoparticles for cell type-specific mRNA delivery. *Adv. Sci. (Weinh.)* **10**, e2301929 (2023).
112. Hou, X., Zaks, T., Langer, R. & Dong, Y. Lipid nanoparticles for mRNA delivery. *Nat. Rev. Mater.* **6**, 1078–1094 (2021).
113. Ou-Yang, S. S. et al. Computational drug discovery. *Acta Pharm. Sin.* **33**, 1131–1140 (2012).
114. Jiang, H. et al. Peptidomimetic inhibitors of APC-Asef interaction block colorectal cancer migration. *Nat. Chem. Biol.* **13**, 994–1001 (2017).
115. Harrison, J. S. et al. Hemi-methylated DNA regulates DNA methylation inheritance through allosteric activation of H3 ubiquitylation by UHRF1. *Elife* **5**, https://doi.org/10.7554/eLife.17101 (2016).
116. Fang, J. et al. Hemi-methylated DNA opens a closed conformation of UHRF1 to facilitate its histone recognition. *Nat. Commun.* **7**, 11197 (2016).
117. Liu, Y. C. et al. Demethylation and up-regulation of an oncogene after hypomethylating therapy. *N. Engl. J. Med* **386**, 1998–2010 (2022).
118. O'Hagan, H. M. et al. Oxidative damage targets complexes containing DNA methyltransferases, SIRT1, and polycomb members to promoter CpG Islands. *Cancer Cell* **20**, 606–619 (2011).
119. Kuo, A. J. et al. The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome. *Nature* **484**, 115–119 (2012).
120. Guzman, C., Bagga, M., Kaur, A., Westermarck, J. & Abankwa, D. ColonyArea: an ImageJ plugin to automatically quantify colony formation in clonogenic assays. *PLoS One* **9**, e92444 (2014).
121. Zhang, W.-Z. et al. The protein complex crystallography beamline (BL19U1) at the Shanghai Synchrotron Radiation Facility. *Nucl. Sci. Tech.* **30**, 170 (2019).
122. Yu, F. et al. Aquarium: an automatic data-processing and experiment information management system for biological macromolecular crystallography beamlines. *J. Appl. Crystallogr.* **52**, 472–477 (2019).
123. Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. iMOSFLM: a new graphical interface for diffraction-image

processing with MOSFLM. *Acta Crystallogr D. Biol. Crystallogr* **67**, 271–281 (2011).
124. Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr D. Biol. Crystallogr* **67**, 282–292 (2011).
125. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D. Biol. Crystallogr* **67**, 355–367 (2011).
126. Patel, S., Ryals, R. C., Weller, K. K., Pennesi, M. E. & Sahay, G. Lipid nanoparticles for delivery of messenger RNA to the back of the eye. *J. Control Release* **303**, 91–100 (2019).
127. Xie, W. et al. DNA Methylation Patterns Separate Senescence from Transformation Potential and Indicate Cancer Risk. *Cancer Cell* **33**, 309–321 e305 (2018).
128. Aryee, M. J. et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363–1369 (2014).
129. Zhou, W., Triche, T. J. Jr., Laird, P. W. & Shen, H. SeSAMe: reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. *Nucleic Acids Res 46, e123* https://doi.org/10.1093/nar/gky691 (2018).
130. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
131. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
132. Pertea, M. et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
133. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
134. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).
135. Zhou, Y. et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).
136. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).

Acknowledgements

Research was supported by grants from National Natural Science Foundation of China, No. 22107101 to X.K.; Investigator initiated grant from Bristol Myers-Squibb – Celgene to S.B.B.; National Institute of Environmental Health Sciences (R01ES011858) to S.B.B.; the Samuel Waxman Research Foundation to S.B.B.; The Hodson Trust to S.B.B.; National Key R&D Program of China (2022YFA1303100 to J. L., 2023YFF0724200 to L.W.); Guangdong Basic and Applied Basic Research Foundation (2021B1515420002, 2023B1212060050) to X.K.; the Science and Technology Projects in Guangzhou (2024A04J4358) to W.B.; the Pearl River Talents Plan (2021QN02Y734) to X.K., the Open Project of State Key Laboratory of Respiratory Disease (SKLRD-OP-202213) to X.K., the Lingang Laboratory (LG-QS-202205-07) to X.K., the Basic Research Project of Guangzhou Institutes of Biomedicine and Health (No. GIBHBRP23-03, GIBHBRP24-03) to X.K. and L.W., and National Cancer Institute, National Institutes of Health (R01CA283463) to S.B.R. S.B.B. is an inventor of MSP which is licensed to MDxHealth in agreement with Johns Hopkins University (to J.H.U.). S.B.B. and J.H.U. are entitled to royalty sale shares. We acknowledge the help of Dr. Qianqian Fan in formatting the LNP diagram figure, and the help of Haozhao Liang for IHC and H&E staining, and the Johns Hopkins SKCC Experimental and Computational Genomics Core. We thank the staff members from BL19U1 beamline at the National Center for Protein Science Shanghai and BL02U1 beamline at the Shanghai Synchrotron Radiation Facility for assistance during the diffraction data collection.

Author contributions

W.B., J.X., W.G., D.W., Y.C., J.L., S.B.B., and X.K. conceived and designed the research; X.K., W.B., W.G., and Y.C. performed in vitro assays and animal experiments; W.G. performed bioinformatics analyses; J.X., Q.G., and X.D. performed crystallographic and ITC studies; D.W., X.L., G.H., Y.W., and J.D. prepared the LNP and PNP formulations; X.K., W.B., J.X., W.G., D.W., J.L., and S.B.B. wrote the manuscript; C.L., S.Y., S.B.R., and L.W. provided critical advice on this study and revised the manuscript; W.R., H.G., C.X., and R.-W.C.Y. provided important research reagents and technical advice. All authors revised and approved the 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/s41467-024-55481-7.

Correspondence and requests for materials should be addressed to Jinsong Liu, Stephen B. Baylin or Xiangqian Kong.

Peer review information *Nature Communications* thanks Youqiong Ye, Yizhou Dong and the other, anonymous, reviewers for their contribution to the peer review of this work. A peer review file is available.