LETTER

doi:10.1038/nature19759

Acetylation–regulated interaction between p53 and SET reveals a widespread regulatory mode

Donglai Wang[1]*, Ning Kon[1]*, Gorka Lasso[2], Le Jiang[1], Wenchuan Leng[3], Wei-Guo Zhu[4], Jun Qin[3,5], Barry Honig[2] & Wei Gu[1]

Although lysine acetylation is now recognized as a general protein modification for both histones and non-histone proteins[1–3], the mechanisms of acetylation-mediated actions are not completely understood. Acetylation of the C-terminal domain (CTD) of p53 (also known as TP53) was an early example of non-histone protein acetylation[4] and its precise role remains unclear. Lysine acetylation often creates binding sites for bromodomain-containing 'reader' proteins[5,6]. Here we use a proteomic screen to identify the oncoprotein SET as a major cellular factor whose binding with p53 is dependent on CTD acetylation status. SET profoundly inhibits p53 transcriptional activity in unstressed cells, but SET-mediated repression is abolished by stress-induced acetylation of p53 CTD. Moreover, loss of the interaction with SET activates p53, resulting in tumour regression in mouse xenograft models. Notably, the acidic domain of SET acts as a 'reader' for the unacetylated CTD of p53 and this mechanism of acetylation-dependent regulation is widespread in nature. For example, acetylation of p53 also modulates its interactions with similar acidic domains found in other p53 regulators including VPRBP (also known as DCAF1), DAXX and PELP1 (refs. 7–9), and computational analysis of the proteome has identified numerous proteins with the potential to serve as acidic domain readers and lysine-rich ligands. Unlike bromodomain readers, which preferentially bind the acetylated forms of their cognate ligands, the acidic domain readers specifically recognize the unacetylated forms of their ligands. Finally, the acetylation-dependent regulation of p53 was further validated *in vivo* by using a knock-in mouse model expressing an acetylation-mimicking form of p53. These results reveal that acidic-domain-containing factors act as a class of acetylation-dependent regulators by targeting p53 and, potentially, other proteins.

Although the physiological consequences of acetylation at positions K120 and K164 within the DNA-binding domain have been established in studies of p53 acetylation-defective mutant mice[10,11], the *in vivo* functions of CTD acetylation remain unclear. By examining mutant mice expressing C-terminal truncated forms of p53, two recent studies have shown that loss of the CTD results in p53 activation[12,13], suggesting that the CTD may act as a docking site for negative regulators of p53. Nevertheless, the identity of the negative regulators and the consequences of CTD acetylation remain unknown. To identify proteins that bind to p53 in a manner dependent on the CTD acetylation status of p53, we synthesized both unacetylated (Un-Ac) and fully-acetylated (Ac) biotin-conjugated CTD peptides and used the immobilized peptides as affinity columns to purify cellular factors (Fig. 1a). We failed to identify any proteins enriched in the acetylated p53 CTD column (Fig. 1b). Instead, coomassie blue staining of the bound fraction revealed a major band of approximately 38 kDa from the unacetylated p53 column that was completely absent in the acetylated column. Mass spectrometry analysis of this band revealed 28 unique peptides identical to SET (Fig. 1c and Extended Data Fig. 1a), an oncoprotein that is activated by translocation-associated gene fusions in patients with acute myeloid leukaemia[14]. Although a previous study reported an interaction between p53 and SET[15], the impact of CTD acetylation on the functional consequences of this interaction are unclear.

Acetylation-dependent disruption of the p53–SET interaction was confirmed *in vitro* with purified SET protein (Fig. 1d). Moreover, expression of CREB-binding protein (CBP), the enzyme responsible for CTD acetylation, completely abrogated the formation of SET complexes with wild-type p53 ($p53^{WT}$), but not with a CTD acetylation-deficient p53 ($p53^{KR}$) mutant, confirming that CTD acetylation is crucial for the p53–SET interaction in cells (Fig. 1e). Notably, other modifications on the CTD lysine residues, including methylation, ubiquitination, sumoylation and neddylation, had no effect on this binding, underscoring the specificity of the acetylation-dependent control of p53–SET interactions (Extended Data Fig. 1b–e).

Next, we tested whether SET acts as a transcriptional cofactor by forming a p53–SET complex on the p53 target promoter. Although SET alone showed no obvious DNA-binding activity (Fig. 1f), in the presence of both p53 and SET, a slower-migrating SET/p53–DNA complex was formed and super-shifted by antibodies against p53 or SET. Further binding-domain mapping indicated that the CTD of p53 interacts directly with the acidic domain of SET (Extended Data Fig. 1f–h). To determine the impact of SET on the transcriptional activity of p53, we measured transactivation of a p53-responsive reporter gene. Indeed, p53-mediated transactivation was abrogated upon co-expression of wild-type SET, but not a SET mutant lacking the acidic domain required for p53 binding (Fig. 1g). Conversely, wild-type SET-mediated repression was abrogated when a p53 mutant lacking the CTD was expressed (Fig. 1g). Notably, the interaction of endogenous p53 and SET was easily detected in unstressed cells; however, upon DNA damage, despite increased p53 levels, the p53–SET interaction was largely diminished, probably owing to the induction of CTD acetylation (Fig. 1h). Moreover, chromatin immunoprecipitation (ChIP) assays revealed that the recruitment of SET to the promoter of p53 targets was largely inhibited (Fig. 1i and Extended Data Fig. 1i–k). Together, these data indicate that SET acts as a transcriptional co-repressor of p53. However, acetylation of the CTD upon DNA damage leads to abrogation of this repression through disruption of the p53–SET interaction (Fig. 1j).

We further investigated whether inactivation of SET influences the activities of p53 in human cancer cells. RNA-interference-mediated depletion of SET markedly elevated the expression of p53 targets, such as cyclin dependent kinase inhibitor 1A (CDKN1A, also known as p21) and p53 upregulated modulator of apoptosis (PUMA, also

[1]Institute for Cancer Genetics, Department of Pathology and Cell Biology, Herbert Irving Comprehensive Cancer Center, College of Physicians & Surgeons, Columbia University, 1130 Nicholas Ave, New York, New York 10032, USA. [2]Department of Biochemistry and Molecular Biophysics and Systems Biology, Center for Computational Biology and Bioinformatics, Howard Hughes Medical Institute, Columbia University, 1130 Nicholas Ave, New York, New York 10032, USA. [3]State Key Laboratory of Proteomics, National Center for Protein Sciences (The PHOENIX Center, Beijing), Beijing, 102206, China. [4]Department of Biochemistry and Molecular Biology, Shenzhen University School of Medicine, Shenzhen 518060, China. [5]Alkek Center for Molecular Discovery, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, Texas 77030, USA.
*These authors contributed equally to this work.

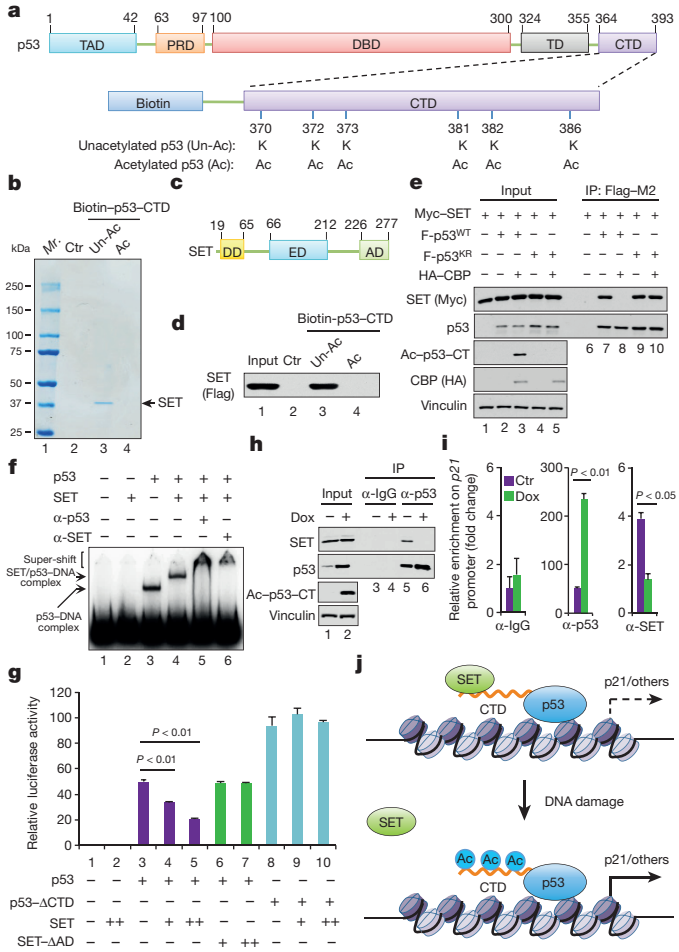


Figure 1 | Identification of SET as a specific co-repressor of C-terminal unacetylated p53. a, Schematic diagram of the synthesized biotin-conjugated p53 CTD. **b**, Coomassie blue staining of the protein complex bound with the p53 CTD. **c**, Schematic diagram of SET. DD: dimerization domain; ED: earmuff domain; AD: acidic domain. **d**, *In vitro* binding assay of p53 CTD and purified SET. **e**, Western blot analysis of the interaction between p53 and SET in the nuclear fraction of H1299 cells. **f**, Electrophoretic mobility shift assay showing the SET/p53–DNA complex formation *in vitro*. **g**, Luciferase assays of SET-mediated regulation of p53 transactivity in H1299 cells. **h**, Western blot analysis of the endogenous interaction between p53 and SET upon doxorubicin (Dox) treatment of HCT116 cells. **i**, ChIP analysis of p53 or SET recruitment onto the *p21* promoter upon Dox treatment of HCT116 cells. **j**, A model of dynamic promoter-recruitment of SET regulated by p53 CTD acetylation status. Error bars indicate mean ± s.d., $n = 3$ for technical replicates. Data are shown as representative of three experiments. Uncropped blots can be found in Supplementary Fig. 1.

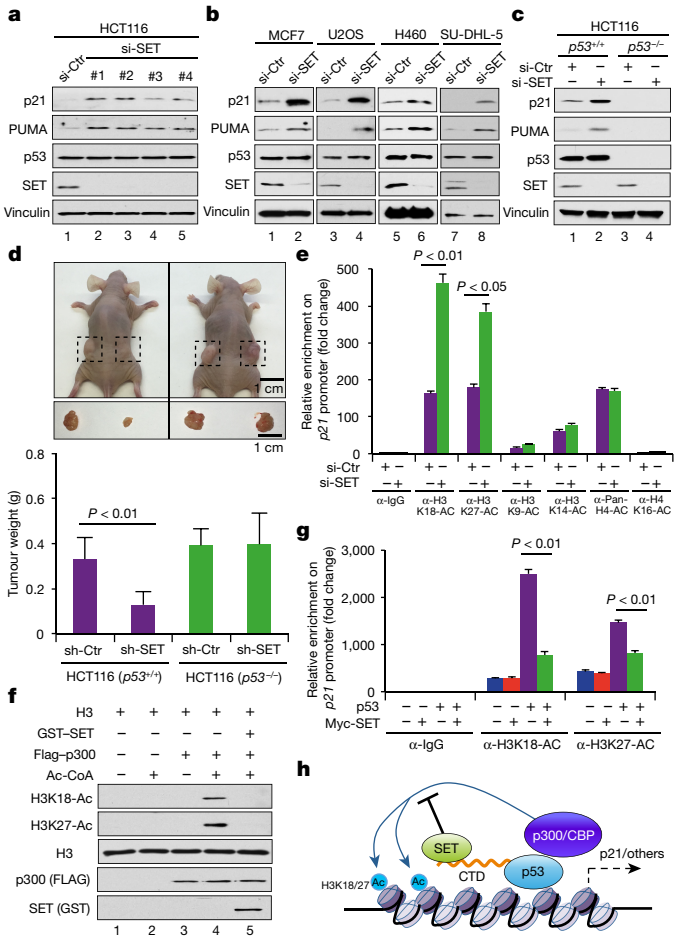


Figure 2 | SET negatively regulates p53 transactivity by inhibiting p300/CBP-mediated H3K18 and H3K27 acetylation on the p53 target promoter. a–c, Western blot analysis of the effect of SET knockdown on p53 activity in cells. si-Ctr: control siRNA. **d**, Xenograft analysis of SET-mediated effect on growth of control and p53-deficient HCT116 tumours. Top, representative images of mice (NU/NU; left flank: control knockdown cells; right flank: SET knockdown cells). Insert: images of dissected HCT116 tumours from the mice shown above. Bottom, analysis of tumour weight growing from $p53^{+/+}$ and $p53^{-/-}$ HCT116 cells after SET depletion in xenografted mice. sh-Ctr: control shRNA; sh-SET: human SET-specific shRNA. Scale bars, 1 cm. **e**, ChIP analysis of the SET knockdown-mediated effect on histone modifications at the *p21* promoter in HCT116 cells. **f**, *In vitro* acetylation assay of the effect of SET on p300-mediated H3K18 and H3K27 acetylation. **g**, ChIP analysis of the SET-mediated effect on p53-dependent H3K18 and H3K27 acetylation on the *p21* promoter in H1299 cells. **h**, A model of SET-mediated regulation on p53 transactivity. Error bars indicate mean ± s.d., $n = 3$ for technical replicates in **e** and **g**; $n = 5$ ($p53^{+/+}$ group) or $n = 3$ ($p53^{-/-}$ group) for biological replicates in **d**. Data are shown as representative of three experiments. Uncropped blots can be found in Supplementary Fig. 1.

known as Bcl-2-binding component 3), without affecting the steady-state levels of endogenous p53 in HCT116 colorectal carcinoma cells (Fig. 2a). Similar effects were obtained in other human cancer cell lines that express wild-type p53, including MCF7 (breast carcinoma), U2OS (osteosarcoma), H460 (lung carcinoma) and SU-DHL-5 (B-cell lymphoma) (Fig. 2b). Moreover, this induction of p21 and PUMA expression was completely abrogated in isogenic HCT116 $p53^{-/-}$ cells (Fig. 2c), indicating that the SET-mediated effects are p53-dependent. Further analysis of U2OS and p53-null U2OS cells that had SET knocked down identified a number of p53 targets that were upregulated upon inactivation of SET in a p53-dependent manner; SET knockdown induced p53-dependent cell growth repression in those cells (Extended Data Figs 2a–c, 3a, b). To examine the effect of SET on p53-mediated tumour suppression, we tested whether SET depletion affected cell growth in xenograft tumour models in immunodeficient mice (NU/NU). SET knockdown strongly suppressed tumour growth of HCT116 cells, but not isogenic HCT116 $p53^{-/-}$ cells (Fig. 2d). Moreover, the p53-dependent effects were further validated in HCT116 p53 knockout cells generated by the CRISPR/Cas9-mediated genome editing technique (Extended Data Fig. 3c–e). These data indicated that the p53–SET interaction is crucial for the tumour growth suppression induced by p53.

As SET had no apparent effect on protein stability, DNA binding or acetylation levels of p53 (Extended Data Fig. 4a–c), we examined whether SET suppressed p53-mediated transactivation by affecting chromatin modifications at p53 target promoters. ChIP analysis revealed that SET depletion significantly increased the acetylation levels of H3K18 and H3K27 at the promoters of *p21* and *PUMA* in

HCT116 cells without affecting H3K9, H3K14, H4K16 or inducing pan-H4 acetylation (Fig. 2e and Extended Data Fig. 4d). p300/CBP, which target H3K18 and H3K27 acetylation *in vivo*[16,17], act as a key co-activators of p53-mediated transcriptional activation[18–20]. We tested whether SET suppressed p300/CBP-mediated acetylation of H3K18 and H3K27, as SET had no obvious effect on the recruitment of p300/CBP (Extended Data Fig. 4e). Indeed, *in vitro* acetylation assays revealed that SET effectively suppressed p300-dependent acetylation of H3K18 and H3K27 (Fig. 2f) and these findings were further verified for p53 target promoters by ChIP analysis (Fig. 2g and Extended Data Fig. 4f). Together, these data indicate that SET represses p53-mediated transactivation by inhibiting p300/CBP-dependent acetylation of H3K18 and H3K27 on p53 target promoters (Fig. 2h).

Numerous studies have indicated that lysine acetylation often creates docking sites for 'reader' proteins that possess a bromodomain, a structural motif that forms a recognition surface for acetylated lysine[5,6]. Our analysis of the p53–SET interaction suggests that the acidic domain of SET serves as a 'converse reader' that binds the lysine-rich CTD of p53 in a manner that can be specifically abrogated upon acetylation of these lysine residues. To further evaluate this model, we tested whether p53 interacts with other proteins in a similar manner. Several transcription cofactors known to interact directly with p53, including VPRBP, DAXX and PELP1 (refs. 7–9), also contain acidic domains similar to that of the SET protein (Fig. 3a and Extended Data Fig. 5a). Their acidic domains also readily bound unacetylated, but not acetylated, p53 CTD (Fig. 3b–d). Similar results were also obtained when the full-length proteins of VPRBP, DAXX and PELP1 were tested (Extended Data Fig. 5b). More importantly, the interactions of VPRBP, DAXX and PELP1 with wild-type p53, but not the acetylation-deficient $p53^{KR}$ mutant, were inhibited by CBP-induced acetylation in human cells (Extended Data Fig. 5c–e).

Previous studies showed that SET also regulates the activities of several other cellular factors, including histone H3, KU70 and FOXO1, through direct interactions with these proteins[21–23]. Notably, the binding region of all three proteins contains a lysine-rich domain (KRD) similar to the CTD of p53 (Fig. 3e). These lysine residues have also been reported to be acetylated *in vivo*[24–26]. To test whether SET-mediated interactions with these factors are also regulated by acetylation, we performed *in vitro* binding assays of the acidic domain of SET with unacetylated or acetylated KRDs of H3, KU70 and FOXO1. The acidic domain of SET interacted with unacetylated, but not acetylated, KRDs of H3, KU70 and FOXO1 (Fig. 3f–h). Similar results were also obtained when the full-length SET protein was used in the binding assays (Extended Data Fig. 5f–h), suggesting that the interaction of SET with H3, KU70 and FOXO1 were abrogated by acetylation in a manner analogous to that of p53 binding to SET. Since VPRBP, DAXX and PELP1 have also been implicated in transcription regulation, we investigated whether these factors could interact with H3 in a similar manner. VPRBP, DAXX and PELP1 specifically bound unacetylated H3 whereas, as expected, bromodomain proteins such BRD4 and BRD7 recognized only acetylated H3 (Extended Data Fig. 5i, j).

Our data indicate that this mechanism of acetylation-dependent regulation is widespread in nature. As the positive charge within the KRD can attract the negative charge of the acidic domain, these lysine clusters form a docking site for acidic-domain-containing regulators. However, upon acetylation, the positive charge of the lysine sidechains is neutralized, abolishing the docking site for the acidic-domain-containing regulators. Conversely, deacetylation of these lysine residues reverses this effect and promotes the recruitment of acidic-domain-containing regulators (Fig. 3i). Thus, unlike bromodomain readers, which preferentially bind the acetylated forms of their cognate ligands, the acidic domain readers specifically recognize the unacetylated forms of their ligands.

To corroborate this notion, we compared the SET-binding properties of the acetylation-deficient mutant $p53^{KR}$ with an acetylation-mimicking mutant, $p53^{KQ}$ (Extended Data Fig. 6a)., The $p53^{KR}$ mutant,

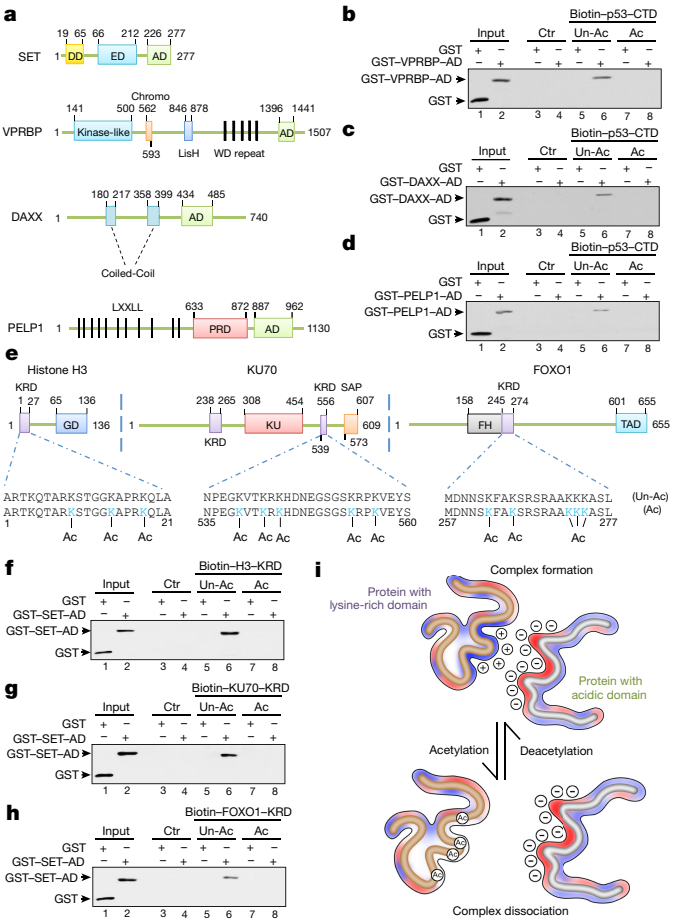


Figure 3 | Acidic-domain-containing proteins represent a new class of 'reader' for their unacetylated ligands. a, Schematic diagrams of the acidic-domain (AD)-containing proteins SET, VPRBP, DAXX and PELP1. **b**–**d**, *In vitro* binding assay of p53 CTD and acidic domains of VPRBP (**b**), DAXX (**c**) and PELP1 (**d**). Empty streptavidin beads were used as negative binding control (Ctr). **e**, Schematic diagrams of the KRD-containing proteins histone H3, KU70 and FOXO1. **f**–**h**, *In vitro* binding assay between the purified SET acidic domain and KRDs of H3 (**f**), KU70 (**g**) and FOXO1 (**h**). **i**, A model of acetylation-dependent regulation of the interactions between KRD-containing proteins and their acidic-domain-containing 'readers'. Uncropped blots can be found in Supplementary Fig. 1.

like unacetylated p53, strongly bound SET (Extended Data Fig. 6b); conversely, the $p53^{KQ}$ mutant, like acetylated p53, did not interact with SET. Similar results were also obtained upon analysis of the acetylation-modulated interactions of p53 with VPRBP, DAXX and PELP1 (Extended Data Fig. 6c–e).

To further determine the physiological importance of these interactions *in vivo*, we generated $p53^{KQ/KQ}$-mutant mice (Extended Data Fig. 7a–d). Although heterozygous $p53^{+/KQ}$ mice displayed normal postnatal development, $p53^{KQ/KQ}$ homozygous mice showed neonatal lethality (Extended Data Fig. 7e). All newborn $p53^{KQ/KQ}$ pups were slightly smaller than their $p53^{+/+}$ littermates (Fig. 4a), lacked milk in their stomachs and died within one day of birth, apparently owing to dehydration from lack of maternal nourishment. In addition, live $p53^{KQ/KQ}$ mice also displayed uncoordinated movements, consistent with neurological impairments. Indeed, the brains of $p53^{KQ/KQ}$ mice appeared smaller than those of $p53^{+/+}$ mice (Fig. 4b).

Immunohistochemistry analysis of $p53^{KQ/KQ}$ brain sections revealed a marked induction of cleaved caspase 3 staining without an obvious increase in p53 protein levels (Fig. 4c and Extended Data Fig. 7f), suggesting that the neurological defects of $p53^{KQ/KQ}$ mice may reflect

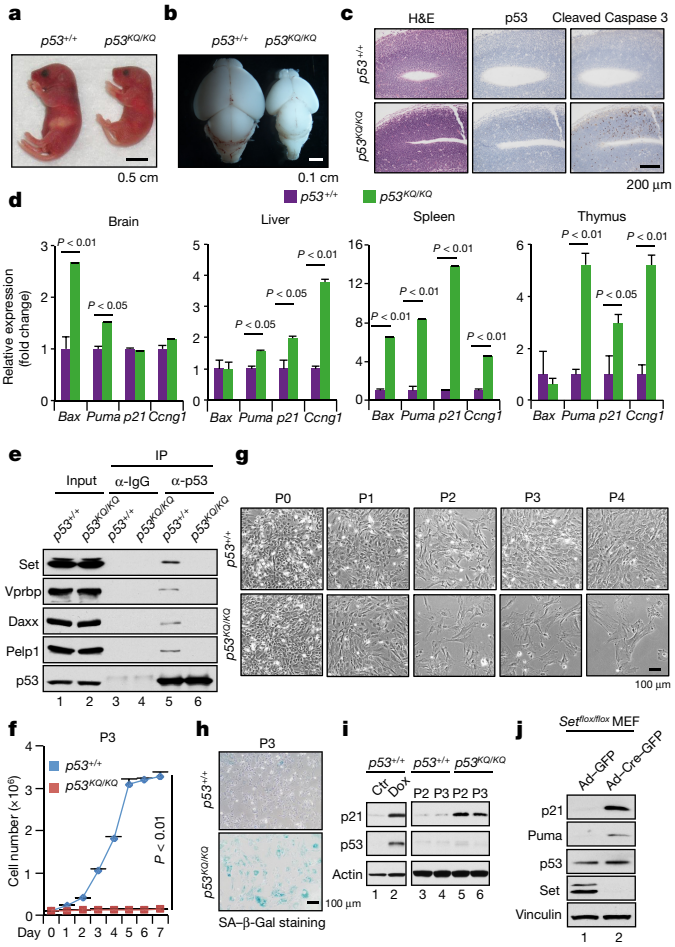


Figure 4 | The physiological significance of acetylation-dependent dissociation of p53 from its acidic-domain-containing 'readers'. **a**, Newborn $p53^{+/+}$ and $p53^{KQ/KQ}$ mice. Scale bar, 0.5 cm. **b**, The brains of newborn $p53^{+/+}$ and $p53^{KQ/KQ}$ mice. Scale bar, 0.1 cm. **c**, Immunohistochemistry analysis of brain sections from $p53^{+/+}$ and $p53^{KQ/KQ}$ embryos. Scale bar, 200 μm. **d**, RT–qPCR analysis of gene expression of p53 targets in $p53^{+/+}$ and $p53^{KQ/KQ}$ tissues. **e**, Western blot analysis of the interaction between p53 and acidic-domain-containing proteins in $p53^{+/+}$ or $p53^{KQ/KQ}$ MEFs treated with the proteasome inhibitor epoxomicin. **f**, Cell growth analysis of $p53^{+/+}$ or $p53^{KQ/KQ}$ MEFs at passage 3 (P3). **g**, Morphological representative images of $p53^{+/+}$ and $p53^{KQ/KQ}$ MEFs from P0 to P4. Scale bar, 100 μm. **h**, SA–β-gal staining of $p53^{+/+}$ and $p53^{KQ/KQ}$ MEFs (P3). Scale bar, 100 μm. **i**, Western blot analysis of p21 and p53 expression in $p53^{+/+}$ and $p53^{KQ/KQ}$ MEFs. **j**, Western blot analysis of p53 targets in *Set* conditional knockout MEFs. Error bars indicate mean ± s.d., $n = 3$ for technical replicates in **d**; $n = 3$ for biological replicates in **f**. Data are shown as representative of three experiments. Uncropped blots can be found in Supplementary Fig. 1.

increased apoptosis due to deregulation of the $p53^{KQ}$ protein. In accordance with this notion, the major apoptotic transcriptional targets of p53, namely *Bax* and *Puma*, were significantly upregulated in $p53^{KQ/KQ}$ brain tissue (Fig. 4d). Indeed, various tissues of $p53^{KQ/KQ}$ mice displayed distinct patterns of induction of different p53 target genes, suggesting tissue-specific activation of target genes by $p53^{KQ}$ *in vivo* (Fig. 4d).

The p53–SET interaction was readily detected in $p53^{+/+}$, but not $p53^{KQ/KQ}$, mouse embryonic fibroblasts (MEFs) (Fig. 4e). Similar results were also obtained for the other acidic-domain-containing cofactors (VPRBP, DAXX and PELP1), suggesting that the $p53^{KQ}$ mutant recapitulates the activity of acetylated p53 *in vivo*. Moreover, $p53^{KQ/KQ}$ MEFs displayed a severe proliferation defect (Fig. 4f) and exhibited clear signs of senescence, including a flat and enlarged morphology with large multinucleated nuclei and marked senescence-associated β-galactosidase (SA-β-Gal) staining (Fig. 4g, h and Extended Data Fig. 7g, h). In addition, western blot analysis revealed an increase in the steady-state levels of p21 protein in $p53^{KQ/KQ}$ MEFs (Fig. 4i). To directly address the role of SET *in vivo*, we generated *Set*-mutant mice (Extended Data Fig. 8a, b). Although the characterization of these mice was not complete (Extended Data Fig. 8c–e), we prepared $Set^{flox/flox}$ MEFs for functional analysis. As shown in Fig. 4j, upon Cre-mediated *Set* deletion, the expression of p53 target genes, such as p21 and Puma, was markedly induced, indicating that SET is a critical regulator of p53 *in vivo*. Together, these data validate the key role of CTD acetylation in p53 activation *in vivo*.

Previous studies showed that a $p53^{KR}$ knock-in mutant targeting the same CTD lysine residues does not significantly affect mouse development or p53 activity in mouse tissues or embryonic fibroblasts[27,28]. Thus, loss of modifiable CTD lysines may neutralize the overall effect on p53 function by abrogating both the negative and positive effects of regulation through different types of CTD modification. Surprisingly, $p53^{KQ}$ knock-in mice died shortly after birth with substantial p53 activation. Like $p53^{KR}$, $p53^{KQ}$ also eliminates other types of modification on these lysine residues; however, $p53^{KQ}$ mimics the acetylated form while $p53^{KR}$ resembles unacetylated p53. Thus, the difference between the phenotypes of $p53^{KQ}$ and $p53^{KR}$ mutant mice underscores the role of CTD acetylation *in vivo*.

The acidic-domain-containing proteins in this study consist of a specific group of proteins that harbour long clusters of acidic amino acids. Searching the Uniprot database with our motif-finding algorithm[29], we identified 49 polypeptides with highly acidic domains similar to SET, many of which are involved in transcriptional regulation and chromatin remodelling (Extended Data Table 1). In addition, by using the Species-Specific Prediction of lysine (K) Acetylation program (SSPKA)[30], we also identified 49 proteins containing a cluster of lysine residues that can potentially bind these acidic domains in an acetylation-modulated manner (Extended Data Table 2). On the basis of our data, we propose that acetylation-mediated regulation, whereby acetylation of p53 abrogates its association with the acidic-domain-containing cofactors, can be expanded to a general mode of post-translational control for protein interactions that involve other acidic-domain-containing factors and their ligands, which can be modified by acetylation.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 8 January; accepted 19 August 2016.
Published online 14 September 2016.

1. Zhao, S. *et al.* Regulation of cellular metabolism by protein lysine acetylation. *Science* **327,** 1000–1004 (2010).
2. Choudhary, C. *et al.* Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325,** 834–840 (2009).
3. Kim, S. C. *et al.* Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol. Cell* **23,** 607–618 (2006).
4. Gu, W. & Roeder, R. G. Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. *Cell* **90,** 595–606 (1997).
5. Dhalluin, C. *et al.* Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399,** 491–496 (1999).
6. Marmorstein, R. & Zhou, M. M. Writers and readers of histone acetylation: structure, mechanism, and inhibition. Cold Spring Harb. *Perspect. Biol.* **6,** a018762 (2014).
7. Kim, K. *et al.* Vpr-binding protein antagonizes p53-mediated transcription via direct interaction with H3 tail. *Mol. Cell. Biol.* **32,** 783–796 (2012).
8. Zhao, L. Y. *et al.* Negative regulation of p53 functions by Daxx and the involvement of MDM2. *J. Biol. Chem.* **279,** 50566–50579 (2004).
9. Nair, B. C. *et al.* Proline, glutamic acid and leucine-rich protein-1 is essential for optimal p53-mediated DNA damage response. *Cell Death Differ.* **21,** 1409–1418 (2014).
10. Li, T. *et al.* Tumor suppression in the absence of p53-mediated cell-cycle arrest, apoptosis, and senescence. *Cell* **149,** 1269–1283 (2012).
11. Jiang, L. *et al.* Ferroptosis as a p53-mediated activity during tumour suppression. *Nature* **520,** 57–62 (2015).
12. Simeonova, I. *et al.* Mutant mice lacking the p53 C-terminal domain model telomere syndromes. *Cell Reports* **3,** 2046–2058 (2013).

13. Hamard, P. J. *et al.* The C terminus of p53 regulates gene expression by multiple mechanisms in a target- and tissue-specific manner *in vivo*. *Genes Dev.* **27,** 1868–1885 (2013).
14. von Lindern, M. *et al.* Can, a putative oncogene associated with myeloid leukemogenesis, may be activated by fusion of its 3′ half to different genes: characterization of the set gene. *Mol. Cell. Biol.* **12,** 3346–3355 (1992).
15. Kim, J. Y. *et al.* Inhibition of p53 acetylation by INHAT subunit SET/TAF-Iβ represses p53 activity. *Nucleic Acids Res.* **40,** 75–87 (2012).
16. Tang, Z. *et al.* SET1 and p300 act synergistically, through coupled histone modifications, in transcriptional activation by p53. *Cell* **154,** 297–310 (2013).
17. Jin, Q. *et al.* Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *EMBO J.* **30,** 249–262 (2011).
18. Kruse, J. P. & Gu, W. Modes of p53 regulation. *Cell* **137,** 609–622 (2009).
19. Vousden, K. H. & Prives, C. Blinded by the light: The growing complexity of p53. *Cell* **137,** 413–431 (2009).
20. Berger, S. L. Keeping p53 in check: a high-stakes balancing act. *Cell* **142,** 17–19 (2010).
21. Matsumoto, K., Nagata, K., Okuwaki, M. & Tsujimoto, M. Histone- and chromatin-binding activity of template activating factor-I. *FEBS Lett.* **463,** 285–288 (1999).
22. Kim, K. B. *et al.* Inhibition of Ku70 acetylation by INHAT subunit SET/TAF-Iβ regulates Ku70-mediated DNA damage response.*Cell. Mol. Life Sci.* **71,** 2731–2745 (2014).
23. Chae, Y. C. *et al.* Inhibition of FoxO1 acetylation by INHAT subunit SET/TAF-Iβ induces p21 transcription. *FEBS Lett.* **588,** 2867–2873 (2014).
24. Kouzarides, T. Chromatin modifications and their function. *Cell* **128,** 693–705 (2007).
25. Cohen, H. Y. *et al.* Acetylation of the C terminus of Ku70 by CBP and PCAF controls Bax-mediated apoptosis. *Mol. Cell* **13,** 627–638 (2004).
26. Daitoku, H. *et al.* Silent information regulator 2 potentiates Foxo1-mediated transcription through its deacetylase activity. *Proc. Natl Acad. Sci. USA* **101,** 10042–10047 (2004).
27. Krummel, K. A., Lee, C. J., Toledo, F. & Wahl, G. M. The C-terminal lysines fine-tune P53 stress responses in a mouse model but are not required for stability control or transactivation. *Proc. Natl Acad. Sci. USA* **102,** 10188–10193 (2005).
28. Feng, L., Lin, T., Uranishi, H., Gu, W. & Xu, Y. Functional analysis of the roles of posttranslational modifications at the p53 C terminus in regulating p53 stability and activity. *Mol. Cell. Biol.* **25,** 5389–5395 (2005).
29. UniProt Consortium. UniProt: a hub for protein information. *Nucleic Acids Res.* **43,** D204–D212 (2015).
30. Li, Y. *et al.* Accurate *in silico* identification of species-specific acetylation sites by integrating protein sequence-derived and functional features. *Sci. Rep.* **4,** 5765 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank F. Giancotti, X. Yang, R. K. Vadlamudi and W. An for providing reagents for this work. We also thank R. Baer for discussion and suggestions. This work was supported by the National Cancer Institute of the National Institutes of Health under Award 5R01CA193890, 5RO1CA190477, 5RO1CA085533 and 2P01CA080058 to W.G. and GM030518 and CA121852 to B.H. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author Contributions The experiments were conceived and designed by D.W., N.K., G.L. and W.G. The experiments were performed mainly by D.W. and N.K. Bioinformatic analysis was performed by G.L. Mass spectrometry analysis was performed by W.L. The xenograft assay was performed by D.W. and L.J. Data were analysed and interpreted by D.W., N.K., G.L., W.-G.Z., J.Q., B.H. and W.G. The manuscript was written by D.W., N.K., G.L. and W.G.

Author Information RNA-seq data is available through NCBI Gene Expression Omnibus (GEO) database with the accession number GSE83635. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to W.G. (wg8@cumc.columbia.edu).

METHODS

General data reports. No statistical methods were used to pre-evaluate the sample size in this study. The experiments (including animal experiments) were not randomized. The investigators were not blinded to experiments. No samples/data were excluded except any obviously unhealthy xenografted mice.

Cell culture, plasmid generation, transfection and reagent treatment. H1299, U2OS, MCF7, H460 and HCT116 cell lines were cultured in DMEM supplemented with 10% (vol/vol) FBS. The SU-DHL-5 cell line was cultured in IMDM supplemented with 10% (vol/vol) FBS. MEFs were cultured in DMEM supplemented with 10% (vol/vol) heat-inactivated FBS. All the cell lines were obtained from ATCC and have been proven to be negative for mycoplasma contamination. No cell lines used in this work were listed in the ICLAC database. The cell lines were freshly thawed from the purchased seed cells and were cultured for no more than 2 months. The morphology of cell lines was checked every week and compared with the ATCC cell line image to avoid cross-contamination or misuse of cell lines. SET stable knockdown cells were generated by lentivirus-based infection of shRNA. *SET* cDNA was purchased from Addgene (Plasmid number 24998) and the full-length cDNA or the various fragments were sub-cloned into pWG-F-HA, pCMV-Myc or PGEX-2TL vectors. Each p53 plasmid was generated by sub-cloning human *p53* cDNA (including full-length or various fragments) into pWG-F-HA, pcDNA3.1 or PGEX-2TL vectors. The point-mutation constructs (including p53-KR and -KQ) were generated by using a site-directed mutagenesis Kit (Stratagene, 200521). Introduction of the expressing construct and siRNA transfection were performed by Lipofectamine 2000 (Invitrogen, 11668-019) according to the manufacturer's protocol. To transfer oligos into SU-DHL-5 cells, we used electroporation following the manufacturer's protocol (Lonza PBC3-00675). The DNA damage inducer doxorubicin was used at 1 μM for 24 h. The proteasome inhibitor epoxomicin was used at 100 nM for 6 h. Cells were treated with TSA (1 μM) and nicotinamide (5 mM) for 6 h to inhibit HDAC activity in the assays in which p53 acetylation needed to be maintained. Ad–GFP and Ad–Cre–GFP viruses were purchased from Vector Biolabs (Catalogue numbers 1761 and 1710).

Mouse model. To generate the knock-in mice, W4/129S6 mouse embryonic stem (ES) cells (Taconic) were electroporated with a targeting vector containing homologous regions flanking the mouse p53 exon 11, in which all 7 lysines were mutated to glutamines ($p53^{KQ}$ allele). A neomycin-resistance gene cassette flanked by two LoxP sites (LNL) was inserted into intron 10 to allow selection of targeted ES cell clones with G418. ES cell clones were screened by Southern blotting with EcoRI-digested genomic DNA, using a probe generated from PCR amplification in the region outside the homologous region in the targeting vector. The correctly targeted ES cell clones containing the K-to-Q mutations were injected into C57BL/6 blastocysts, which were then implanted into pseudopregnant females to generate chimaeras. Germ-line transmission was accomplished by breeding chimaeras with C57BL/6 mice. Subsequently, mice containing the targeted allele were bred with Rosa26-Cre mice to remove the LNL cassette and to generate mice with only the K-to-Q mutations. To confirm the mutations inserted in $p53^{+/KQ}$ mice, we sequenced p53 cDNA derived from mRNA isolated from $p53^{+/KQ}$ spleen. All seven K-to-Q mutations were confirmed and no additional mutations were found. The offspring were genotyped by PCR using the following primer set, forward: 5′-GGGAGGATAAACTGATTCTCAGA-3′, reverse: 5′-GATGGCTTCTACTATGGGTAGGGAT-3′.

To generate a *Set* conditional knockout mouse, exon 2 of the *Set* gene was floxed and deletion of exon 2 resulted in a frameshift and the truncation of the C-terminal domain. The targeting vector of Set contained 10 kb genomic DNA spanning exon 2; a neomycin-resistance gene cassette and loxP sites were inserted flanking exon 2. To increase targeting frequency, a diphtheria toxin A cassette was inserted at the 3′ end of the targeting vector to reduce random integration of the modified *Set* genomic DNA. A new BglII restriction site was also inserted to facilitate Southern blot screening. Of the 200 mouse ES cell clones screened, eight were identified to have integrated the floxed exon 2 by Southern blot using a 5′ probe, which detects a 14-kb band for the wild-type allele and an 11-kb band for the floxed exon 2 allele (Set^{flox}). Two of the clones were then injected into blastocysts to generate *Set* chimaera mice and they were bred to produce germ-line transmission of the floxed exon 2 allele. $Set^{flox/+}$ mice were intercrossed to generate Set homozygous conditional knockout mice ($Set^{flox/flox}$).

Maintenance and experimental procedures of mice were approved by the Institutional Animal Care and Use Committee (IACUC) of Columbia University.

***In vitro* binding assay.** For the *in vitro* peptide binding assay: equal amounts of each synthesized biotin-conjugated peptide (made as column or as batch) were incubated with highly concentrated HeLa nuclear extract (NE) or purified proteins for 1 h or overnight at 4 °C. After washing with BC100 buffer (20 mM Tris-HCl pH 7.9, 100 mM NaCl, 10% glycerol, 0.2 mM EDTA, 0.1% triton X-100) three times, the binding components were eluted in high-salt buffer (20 mM Tris-HCl pH 7.9, 1,000 mM NaCl, 1% DOC, 10% glycerol, 0.2 mM EDTA, 0.1% triton X-100) or by boiling with 1 × Laemmli buffer for further analysis. For the *in vitro* GST-fusion protein binding assay: *Escherichia coli* containing GST or GST-fusion protein expressing constructs were grown in a shaking incubator at 37 °C until the OD_{600} was about 0.6. Next 0.1 mM IPTG was added and the *E. coli* were incubated at 25 °C for 4 h or overnight, to induce GST or GST-fusion protein expression. After purification by GST·Bind Resin (Novagen, 70541), equal amounts of immobilized GST or GST-fusion proteins were incubated with other purified proteins for 1 h at 4 °C, followed by washing with BC100 buffer three times. The binding components were eluted by boiling with 1 × Laemmli buffer and were analysed by western blot.

Co-immunoprecipitation assay (Co-IP). Whole cellular extracts (WCE) were prepared in BC100 buffer with sonication. Nuclear extract (NE) was prepared by sequentially lysing cells with HB buffer (20 mM Tris-HCl pH 7.9, 10 mM KCl, 1.5 mM $MgCl_2$, 1 mM PMSF, 1 × protease inhibitor (Sigma)) for the cytosolic fraction and BC400 buffer (20 mM Tris-HCl pH 7.9, 400 mM NaCl, 10% Glycerol, 0.2 mM EDTA, 0.5% triton X-100, 1 mM PMSF, 1 × protease inhibitor) for nuclear fraction. The salt concentration of NE was adjusted to 100 mM. 2 μg of the indicated antibody (or 20 μl Flag M2 Affinity Gel (Sigma, A2220)) was added into WCE or NE and incubated overnight at 4 °C, followed by addition of 20 μl protein A/G agarose (Santa Cruz, sc-2003; only for IP with unconjugated antibodies mentioned above) for 2 h. After washing with BC100 buffer three times, the binding components were eluted using Flag peptide (Sigma, F3290), 0.1% trifluoroacetic acid (TFA, Sigma, 302031) or by boiling with 1 × Laemmli buffer, and were analysed by western blot.

Purification of Ub-, Sumo- or Nedd-p53 conjugates from cells. For preparation of Ub-p53: H1299 cells were co-transfected with p53, MDM2 and 6 × HA-Ub (human) expressing plasmids for 48 h. The cells were lysed with Flag lysis buffer (50 mM Tris-HCl pH 7.9, 137 mM NaCl, 10 mM NaF, 1 mM Na_3VO_4, 10% glycerol, 0.5 mM EDTA, 1% triton X-100, 0.2% sarkosyl (sodium lauroyl sarcosinate), 0.5 mM DTT, 1 mM PMSF, 1 × protease inhibitor) and total Ub-conjugated proteins were purified by anti-HA-agarose (Sigma, A2095) and eluted by 1 × HA peptide (Sigma I2149). For the preparation of Sumo-p53 or Nedd-p53: H1299 cells were co-transfected with p53, MDM2 (only for Nedd-p53 preparation) and 6 × His-HA-Sumo1 (human) or 6 × His-HA-Nedd8 (human) expressing plasmids for 48 h. The cells were lysed with guanidine lysis buffer (6 M guanidin-HCl, 0.1 M Na_2HPO_4, 6.8 mM NaH_2PO_4, 10 mM Tris-HCl pH 8.0, 0.2% triton-X100, freshly supplemented with 10 mM β-mercaptoethanol and 5 mM imidazole) with mild sonication. After overnight pull-down by Ni^+-NTA agarose (Qiagen 30230), the binding fractions were sequentially washed with guanidine lysis buffer, urea buffer I (8 M urea, 0.1 M Na_2HPO_4, 6.8 mM NaH_2PO_4, 10 mM Tris-HCl pH 8.0, 0.2% triton-X100, freshly supplemented with 10 mM β-mercaptoethanol and 5 mM imidazole) and urea buffer II (8 M urea, 18 mM Na_2HPO_4, 80 mM NaH_2PO_4, 10 mM Tris-HCl pH 6.3, 0.2% triton-X100, freshly supplemented with 10 mM β-mercaptoethanol and 5 mM imidazole). Precipitates were eluted in elution buffer (0.5 M imidazole, 0.125 M DTT). All purified proteins were dialysed against BC100 buffer before use in the subsequent pull-down assay. After the pull-down assay, the interaction between SET and each p53-conjugate was detected by western blot with anti-p53 (DO-1) antibody.

Mass spectrometry assay. The protein complex was separated by SDS–PAGE and stained with GelCode Blue reagent (Pierce, 24592). The visible band was cut and digested with trypsin and then subjected to liquid chromatography (LC)-MS/MS analysis.

Luciferase assay. A firefly reporter (p21-Luci reporter) and a Renilla control reporter were co-transfected with indicated constructs in H1299 cells for 48 h and the relative luciferase activity was measured by dual-luciferase assay protocol (Promega, E1910).

Electrophoretic mobility shift assay. Highly purified p53 or SET was incubated with a ^{32}P-labelled probe (160 bp) containing the p53-binding element of the *p21* promoter in 1× binding buffer (10 mM HEPES, pH 7.6, 40 mM NaCl, 50 μM EDTA, 6.25% glycerol, 1 mM $MgCl_2$, 1 mM spermidine, 1 mM DTT, 50 ng μl^{-1} BSA, 5 ng μl^{-1} sheared single strand salmon DNA) for 20 min at room temperature (RT). For the super-shift assay, α-p53 or α-SET antibody was pre-incubated with purified p53 and SET in the reaction system without probe for 30 min at RT and then the probe was added for a further 20 min. The complex was analysed by 4% Tris-Borate-EDTA buffer–polyacrylamide gel electrophoresis (TBE–PAGE) and visualized by autoradiography. The probe was obtained by PCR, labelled by T4 kinase (NEB, M0201S) and purified by Bio-Spin column (Bio-Rad, 732-6223).

Chromatin immunoprecipitation (ChIP) assay. Cells were fixed with 1% formaldehyde for 10 min at room temperature and lysed with ChIP lysis buffer (50 mM Tris-HCl pH 8.0, 5 mM EDTA, 1% SDS, 1× protease inhibitor) for 10 min at 4 °C.

After sonication, the lysates were centrifuged, and the supernatants were collected and pre-cleaned by salmon sperm DNA saturated protein A agarose (Millipore, 16-157) in dilution buffer (20 mM Tris-HCl pH 8.0, 2 mM EDTA, 150 mM NaCl, 1% triton X-100, 1× protease inhibitor) for 1 h at 4 °C. The pre-cleaned lysates were aliquoted equally and incubated with indicated antibodies overnight at 4 °C. Saturated protein A agarose was added into each sample and incubated for 2 h at 4 °C. The agarose was washed with TSE I (20 mM Tris-HCl pH 8.0, 2 mM EDTA, 150 mM NaCl, 0.1% SDS, 1% triton X-100), TSE II (20 mM Tris-HCl pH 8.0, 2 mM EDTA, 500 mM NaCl, 0.1% SDS, 1% triton X-100), buffer III (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.25 M LiCl, 1% DOC, 1% NP40), and buffer TE (10 mM Tris-HCl pH 8.0, 1 mM EDTA), sequentially. The binding components were eluted in 1% SDS and 0.1 M $NaHCO_3$ and reverse cross-linkage was performed at 65 °C for at least 6 h. DNA was extracted using the PCR purification Kit (Qiagen, 28106). Real-time PCR was performed to detect relative enrichment of each protein or modification on indicated genes.

Cell growth assay. Approximately 10^5 MEFs or U2OS cells, as indicated in each figure, were seeded into 6-well plates with three replicates. Their cell growth was monitored on consecutive days, as indicated, by using the Countess automated cell counter (Invitrogen) or by staining with 0.1% crystal violet. For quantitative analysis of the crystal violet staining, the crystal violet was extracted from cells using 10% acetic acid and the relative cell number was measured by detecting the absorbance at 590 nm.

Xenograft model. 10^6 HCT116-derived cells, as indicated in each figure, were mixed with Matrigel (Corning, 354248) in a 1:1 ratio in a total volume of 200 μl. The cell–matrix complex was subcutaneously injected into nude mice (NU/NU; 8 weeks old; female; strain 088; Charles River). After 3 weeks, the mice were killed and weight of the tumours was measured. The experimental procedures were approved by the Institutional Animal Care and Use Committee (IACUC) of Columbia University. None of the experiments were exceeded the limit for tumour burden (10% of total bodyweight or 2 cm in diameter).

RT–qPCR. Total RNA was extracted by TRIzol (Invitrogen, 15596-026) and precipitated in ethanol. 1 μg of total RNA was reverse transcribed into cDNA using the SuperScript III First-Strand Synthesis SuperMix (Invitrogen, 11752-50). The relative expression of each target was measured by qPCR and the data were normalized by the relative expression of *GAPDH* or *ActB*.

Immunohistochemistry (IHC). FFPE sections of mouse brain tissue samples were stained with indicated antibodies and visualized by DAB exposure.

Protein purification. The Flag-tagged p53 or SET construct was transfected into H1299 cells for 48 h and the cells were lysed in Flag lysis buffer. After centrifugation, the Flag M2 Affinity Gel was added to supernatant and incubated for 1h at 4 °C. After washing with Flag lysis buffer six times, the purified proteins were eluted with Flag peptide. For purification of acetylated p53, the construct CBP was co-transfected with the p53 vector for 48 h. TSA and nicotinamide were added into the medium for the last 6 h and the cells were harvested in Flag lysis buffer supplemented with TSA and nicotinamide. The C-terminal unacetylated p53 was removed by p53-PAb421 antibody and then the acetylated p53 was purified as described above.

***In vitro* acetylation assay.** 0.5 μg recombinant H3 was incubated with 20 ng purified p300 in 1× HAT buffer (50 mM Tris-HCl, pH 7.9; 1 mM DTT; 10 mM sodium butyrate, 10% glycerol) containing 0.1 mM Ac-CoA for 30 min at 30 °C. After the reaction, the products were assayed by western blot with indicated antibodies. To measure the effect of SET on p300-mediated H3 acetylation, H3 and purified SET (1 μg) were pre-incubated in 1× HAT buffer for 20 min at room temperature before addition of the other components (p300 and Ac-CoA) for the subsequent *in vitro* acetylation assay.

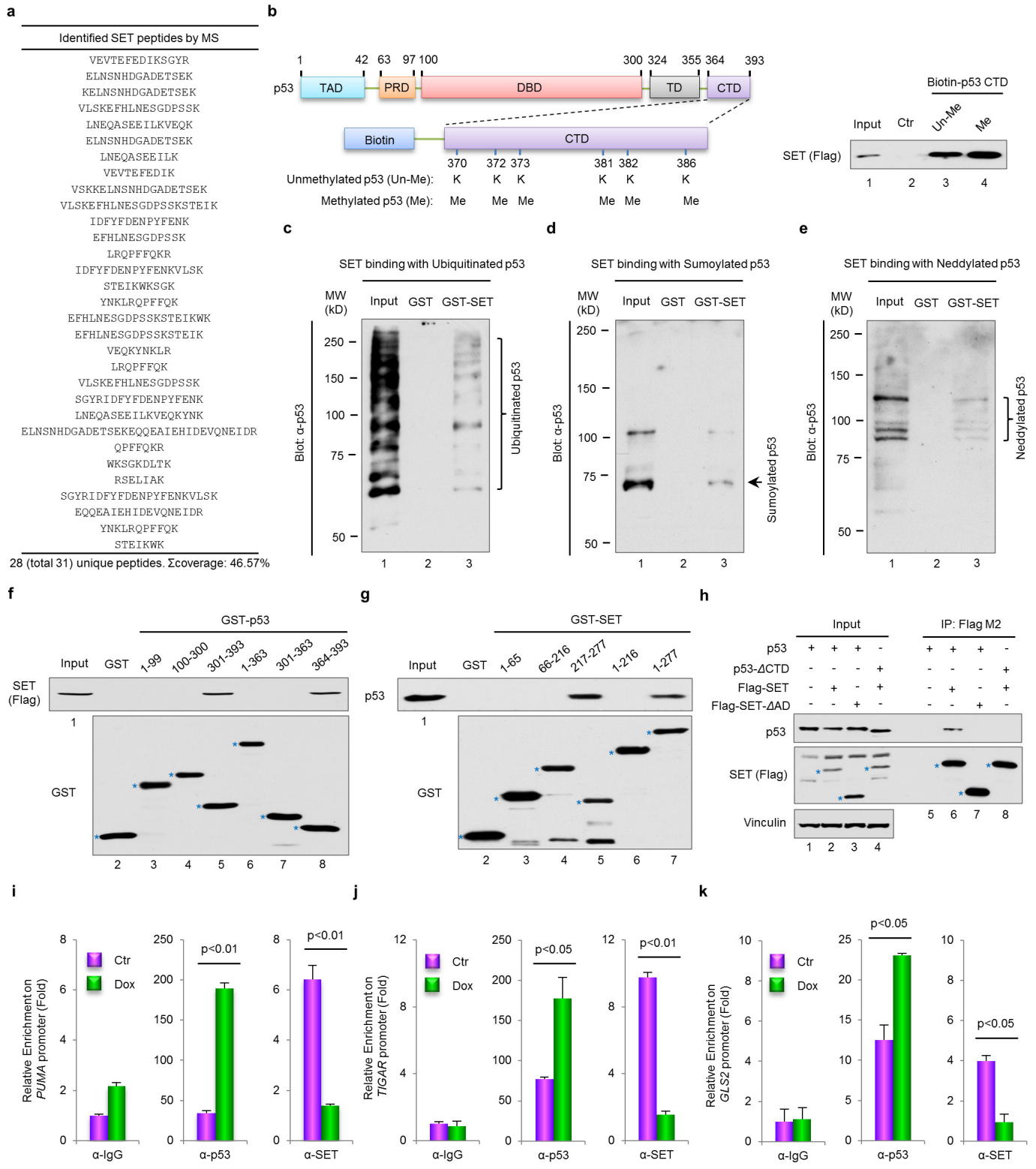
Generation of the p53 knockout (p53-KO) cell line using the CRISPR/Cas9 technique. Cells were transfected with constructs expressing Cas9-D10A (Nickase) and control sgRNAs or sgRNAs targeting p53 exon3 (Santa Cruz: sc-437281 for control; sc-416469-NIC for targeting of p53). After 48 h of transfection, cells were suspended, diluted and re-seeded to ensure single clone formation. More than 30 clones were picked up and the expression of p53 in each single clone was evaluated by western blot with both α-p53 (DO-1) and α-p53 (FL-393) antibodies. Further verification of positive clones was done by sequencing the genomic DNA to make sure that the functional genomic editing occured (insertion or deletion-mediated frame-shift of the *p53* open reading frame (ORF)). Two (U2OS) or three (HCT116) clones were finally selected for subsequent experiments. The p53 knockout-mediated effect was verified to be reproducible in these independent clones. The targeting sequences of p53 loci for the sgRNAs were: 1) TTGCCGTCCCAAGCAATGGA; 2) CCCCGGACGATATTGAACAA.

RNA-seq. U2OS (CRISPR Ctr or CRISPR p53-KO) cells were transfected with control siRNA or SET-specific siRNA (three oligos) for 4 days. Each sample group had at least two biological replicates. Total RNA was prepared using TRIzol (Invitrogen, 15596-026). The RNA quality was evaluated by Bioanalyzer (Agilent) and confirmed that the RIN > 8. Before performing RNA-seq analysis, a small aliquot of each sample was analysed by RT–qPCR to confirm SET knockdown efficiency. RNA-seq analysis was performed at the Columbia Genome Center. Specifically, from total RNA samples, mRNAs were enriched by poly-A pull-down and then processed for library preparation by using the Illumina TruSeq RNA prep kit (Illumina RS-122-2001). Libraries were then sequenced using the Illumina HiSeq2000. Samples were multiplexed in each lane and yielded targeted number of single-end 100-bp reads for each sample. RTA (Illumina) was used for base calling and bcl2fastq (version 1.8.4) was used for converting BCL to fastq format, coupled with adaptor trimming. Reads were mapped to a reference genome (Human: NCBI/build37.2) using TopHat (version 2.0.4). Relative abundance of genes and splice isoforms were determined using Cufflinks (version 2.0.2) using the default settings. Differentially expressed genes were tested under various conditions using DEseq, an R package based on a negative binomial distribution that models the number reads from RNA-seq experiments and tests for differential expression. To further analyse the differentially expressed genes in a more reliable interval, the following filter strategies were applied: 1) the average of FPKM (Fragments per kilobase of transcript per million mapped reads) in either sample group exceeded 0.1; 2) the fold change between the CRISPR Ctr/si-Ctr group and the CRISPR Ctr/si-SET group exceeded 2; 3) the *P* value between the CRISPR Ctr/si-Ctr group and the CRISPR Ctr/si-SET group < 0.01.

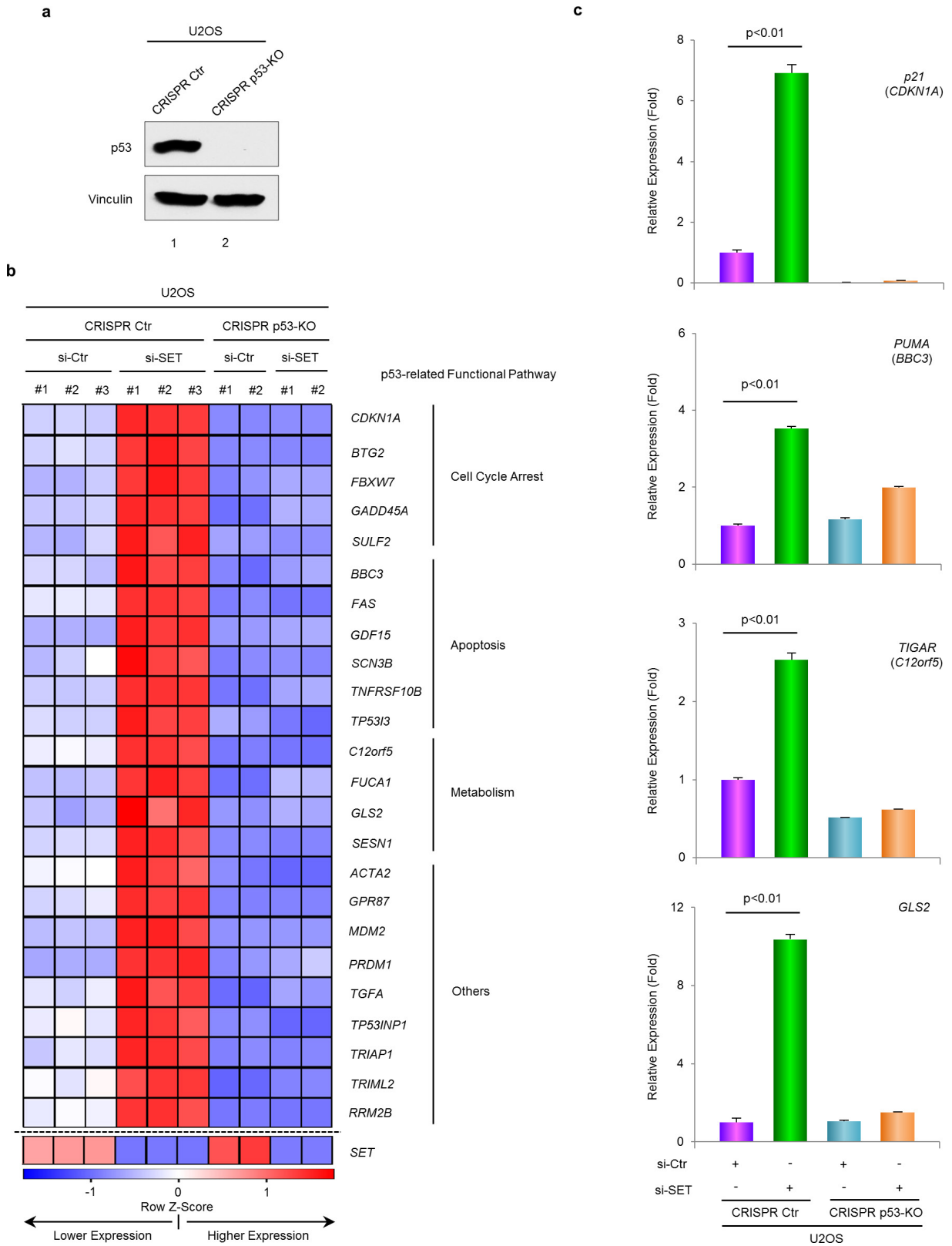
To retrieve potential p53 target genes which were repressed by SET in a p53-dependent manner, we searched the filtered RNA-seq results using the following strategies: 1) the expression level in the CRISPR Ctr/si-SET group was at least 2-fold higher than that in the CRISPR Ctr/si-Ctr group; 2) the expression level in the CRISPR Ctr/si-SET group was at least 2-fold higher than that in the CRISPR p53-KO/si-SET group. The filtered genes which were also verified as p53 target genes from the literature were collected and presented as a heatmap.

Bioinformatic analysis. For the discovery of acidic domains in the human proteome: our motif-finding algorithm initially searched for sequence motifs with a minimum acidic composition of 76% using a sliding window of 36 residues, as dictated by experimental results. Motifs found to be partially overlapping were merged into single motifs. Flanking non-acidic residues were subsequently cropped-out from the final motif. Motif discovery was carried out using the UniProt database, which contains 20,187 canonical human proteins, that have been manually annotated and reviewed. For prediction of proteins that bound acidic domain-containing proteins and were regulated by acetylation: we identified proteins that can potentially bind long acidic domains in a similar way to p53: using a K-rich region whose binding properties can be regulated by acetylation. We used the training set assembled in SSPKA, which combines lysine acetylation annotations from multiple resources obtained either experimentally or in the scientific literature. This dataset individually lists all annotated acetylation sites for a given protein. We generated acetylation motifs with multiple acetylation sites by clustering those sites found to within a maximum distance of 11 residues in sequence. Following this, we searched for acetylation motifs with five or more lysines where at least three of them are annotated as acetylation sites.

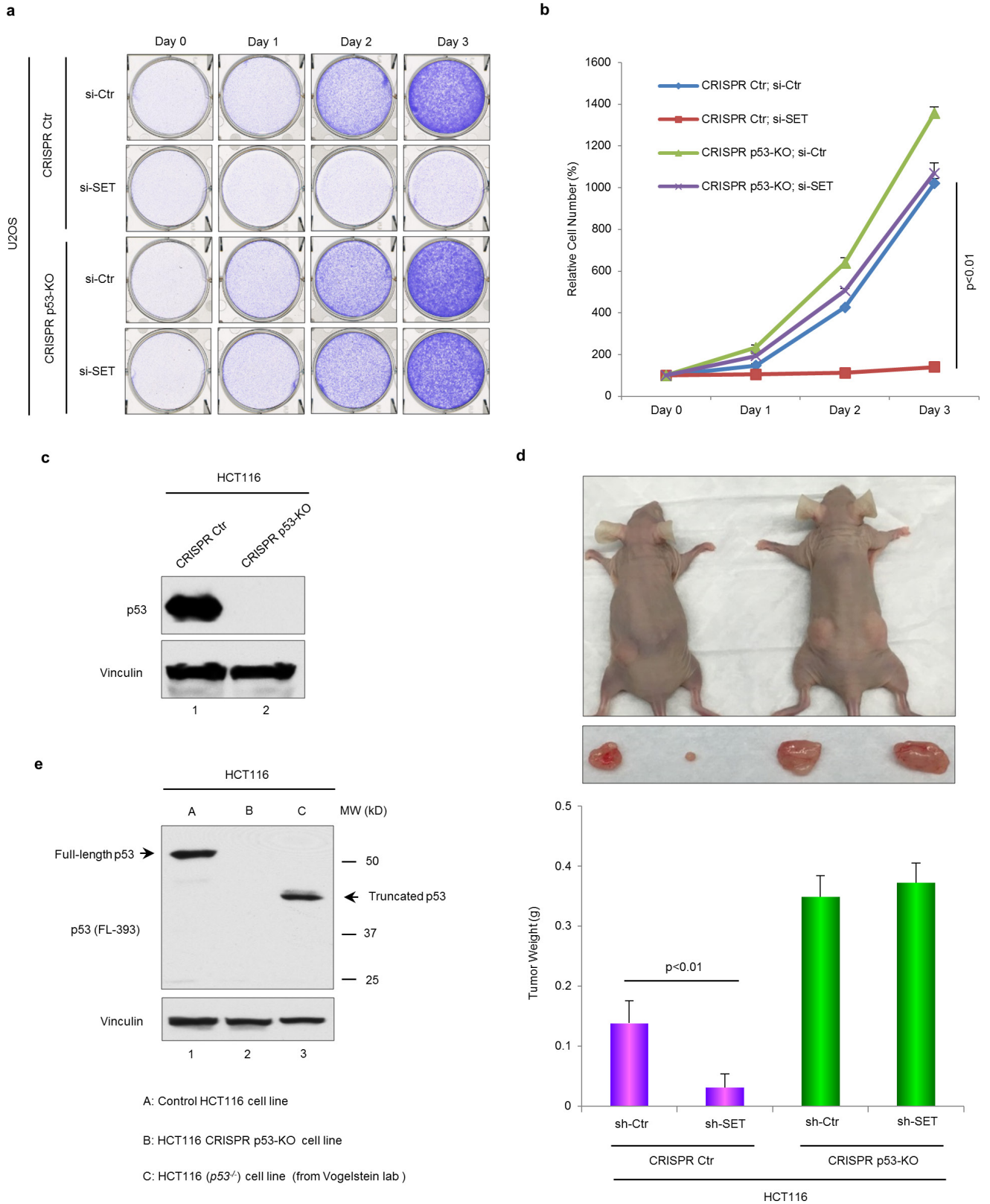
Statistical analysis. Results are shown as means ± s.d. Statistical significance was determined by using a two-tailed, unpaired Student *t*-test in all figures except those described below. In Fig. 1g, significance was determined by one-way ANOVA with a Bonferroni post hoc test. In Fig. 2d and g and Extended Data Figs 2c, 3b, d, 4f and 7h, statistical significance was measured by two-way ANOVA with a Bonferroni post hoc test. All statistical analysis was performed using GraphPad Prism software. $P < 0.05$ was denoted as statistically significant.



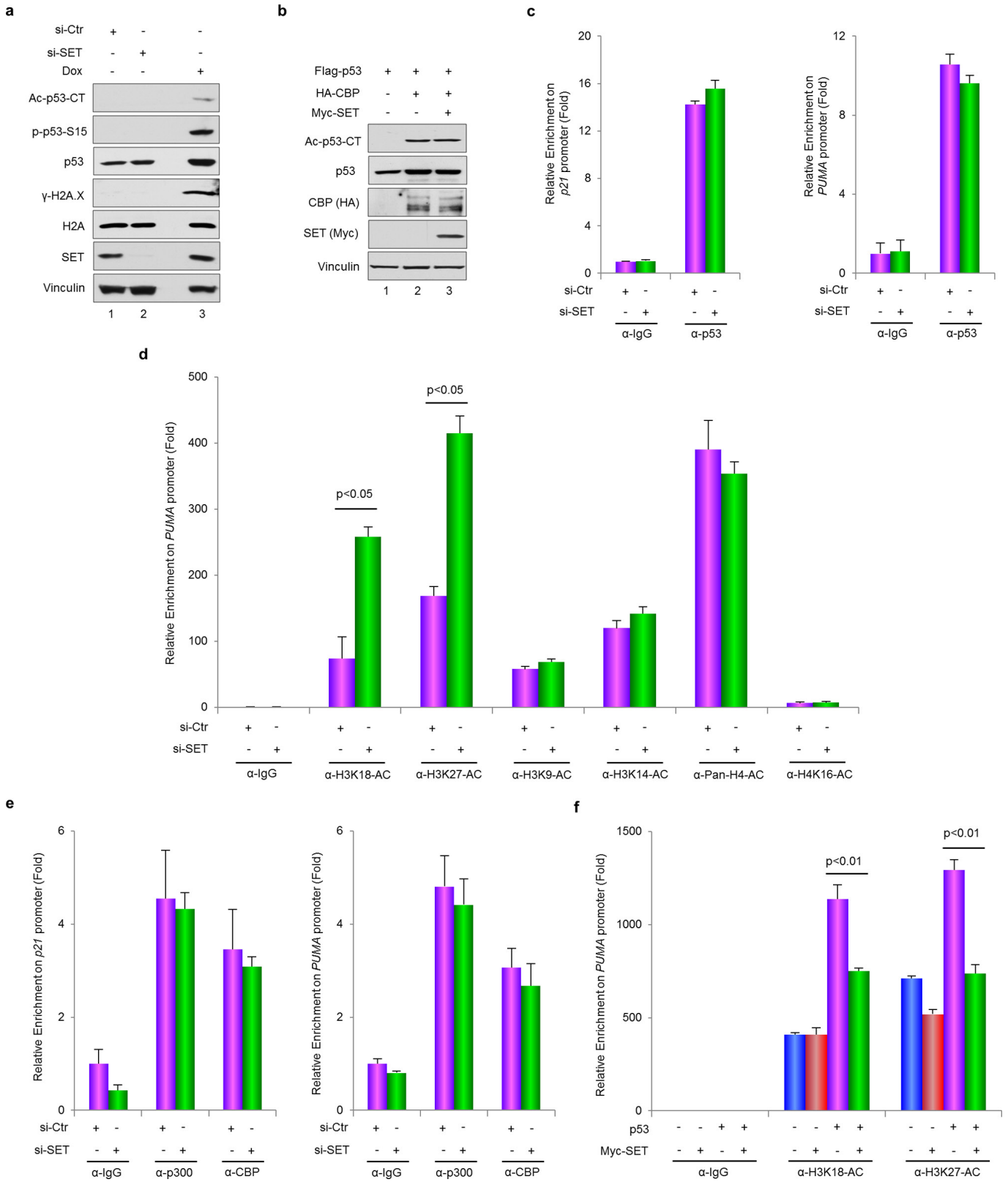
Extended Data Figure 1 | Further analysis of p53–SET interaction. **a**, A list of SET peptides identified by mass spectrometry. **b**, *In vitro* binding assay of methylated p53 CTD and purified SET. **c**–**e**, *In vitro* binding assay between SET and the purified ubiquitinated, sumoylated or neddylated forms of p53. **f**, **g**, Western blot analysis of p53 and SET domains for their interaction. *In vitro* binding assay was performed by incubating immobilized GST, GST–p53 or GST–SET with each purified SET or p53 protein, as indicated. **h**, Western blot analysis of the interaction between p53 and SET in cells. H1299 cells were co-transfected with indicated constructs and the nuclear extract was analysed by co-IP assay. **i**–**k**, ChIP analysis of p53 or SET recruitment onto the *PUMA* (**i**), *TIGAR* (**j**) or *GLS2* (**k**) promoter. HCT116 cells were treated with or without 1 μM doxorubicin for 24 h and then the cellular extracts were analysed by ChIP assay with indicated antibodies. Asterisks indicate the specific bands of indicated proteins. Error bars indicate mean ± s.d., $n = 3$ for technical replicates. Data are shown as representative of three experiments. Uncropped blots can be found in Supplementary Fig. 1.



Extended Data Figure 2 | RNA-seq analysis to identify genes regulated by p53–SET interplay. a, Western blot analysis of the expression of p53 in U2OS-derived CRISPR control cells or CRISPR p53-KO cells. **b**, Heat map of genes regulated by the p53–SET interplay. U2OS (CRISPR Ctr or CRISPR p53-KO) cells were transfected with control siRNA or SET-specific siRNA for 4 days and the total RNA was prepared for RNA-seq analysis with two or three biological replicates, as indicated. Known p53 target genes which were also repressed by SET in a p53-dependent manner were selected and presented as a heat map. The relative *SET* expression is shown in the last row of the heat map. **c**, qPCR validation of the genes regulated by the p53–SET interplay. Error bars indicate mean ± s.d., $n = 3$ for technical replicates. Data are shown as representative of three experiments. Uncropped blots can be found in Supplementary Fig. 1.



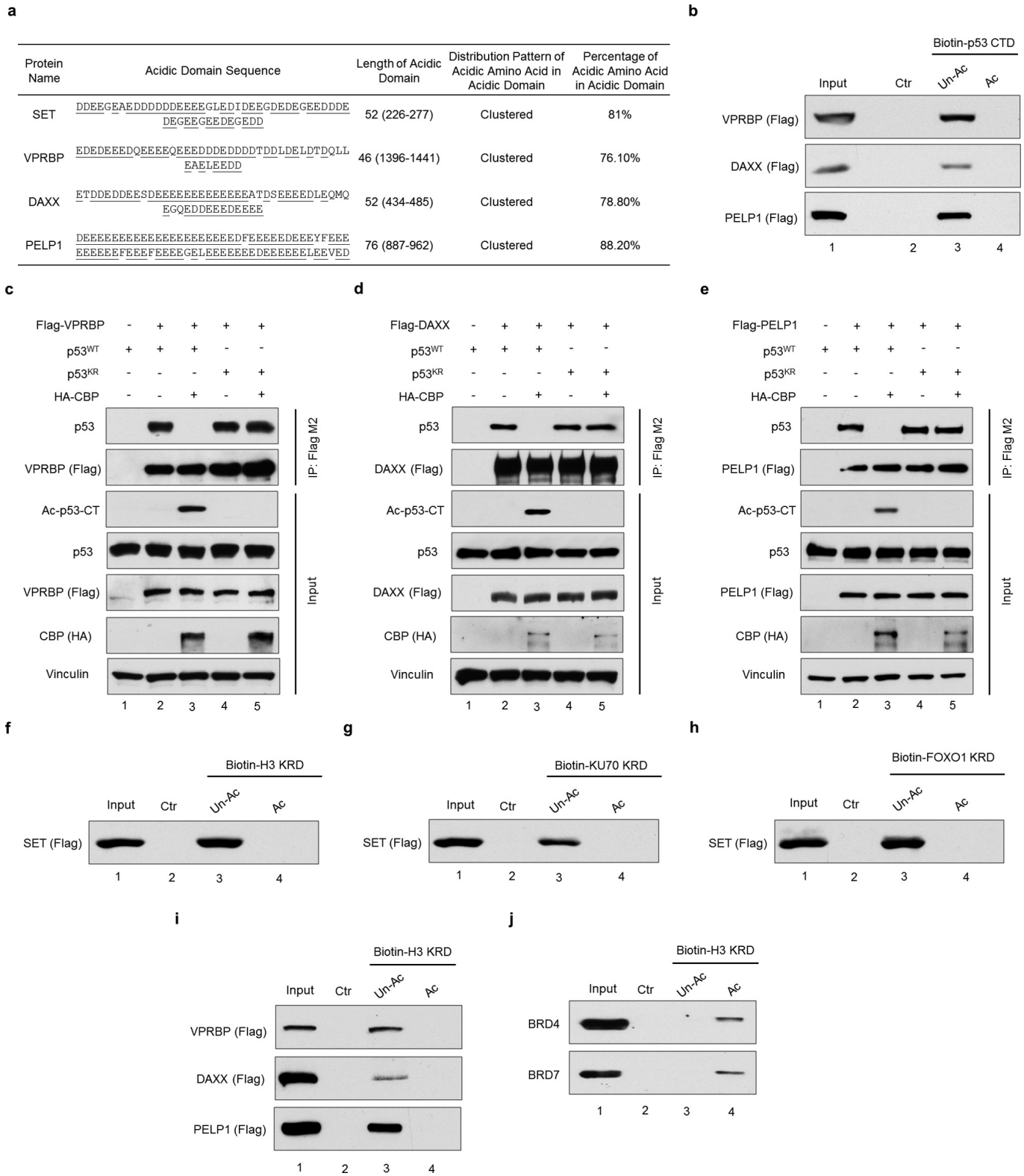
Extended Data Figure 3 | SET-mediated effects on cell proliferation and tumour growth. a, **b**, Representative image (**a**) or quantitative analysis (**b**) of the SET knockdown-mediated effect on cell growth of U2OS-derived CRISPR control cells or CRISPR p53-KO cells. **c**, Western blot analysis of the expression of p53 in HCT116-derived CRISPR control cells or CRISPR p53-KO cells. **d**, Xenograft analysis of the SET-mediated effect on tumour growth by HCT116-derived CRISPR control cells or CRISPR p53-KO cells. **e**, Western blot analysis of p53 expression in control or derived HCT116 cell lines, as indicated. Error bars indicate mean ± s.d., $n = 3$ in **b** or $n = 5$ in **d** for biological replicates. Uncropped blots can be found in Supplementary Fig. 1.



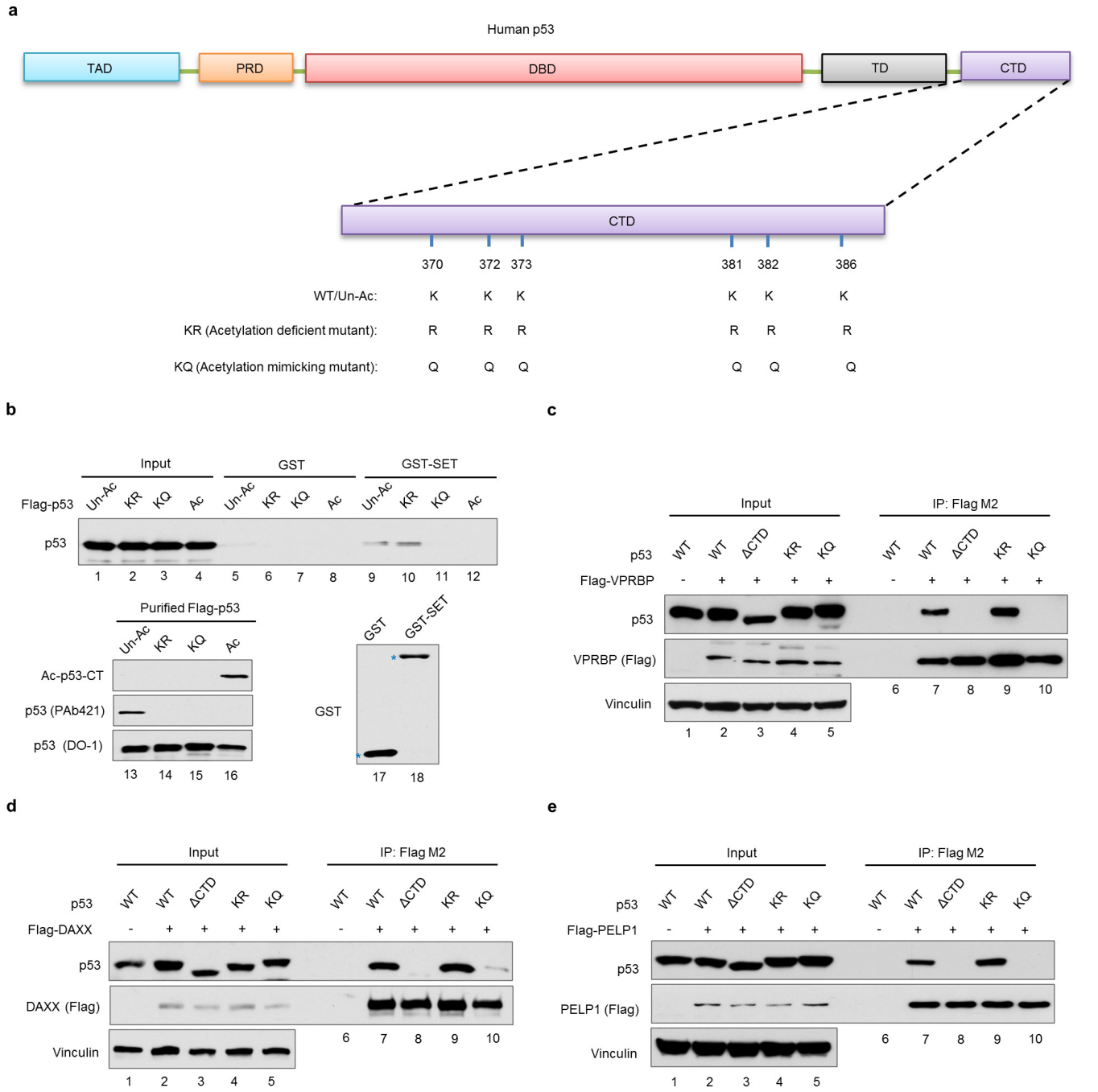
Extended Data Figure 4 | SET regulates histone modifications on p53 target promoter. a, Western blot analysis of the SET knockdown-mediated effect on the p53 C-terminal acetylation in HCT116 cells. Doxorubicin (Dox)-treated cells were also analysed in parallel as a positive control. **b**, Western blot analysis of the SET-mediated effect on the CBP-induced p53 C-terminal acetylation in H1299 cells. **c**, **e**, ChIP analysis of promoter-recruitment of p53 (**c**) or p300/CBP (**e**) upon SET depletion in HCT116 cells. **d**, ChIP analysis of the SET-knockdown-mediated effect on histone modifications in the *PUMA* promoter in HCT116 cells. **f**, ChIP analysis of the SET-mediated effect on p53-dependent H3K18 and H3K27 acetylation in the *PUMA* promoter. Error bars indicate mean ± s.d., $n = 3$ for technical replicates. Data are shown as representative of three experiments. Uncropped blots can be found in Supplementary Fig. 1.

a

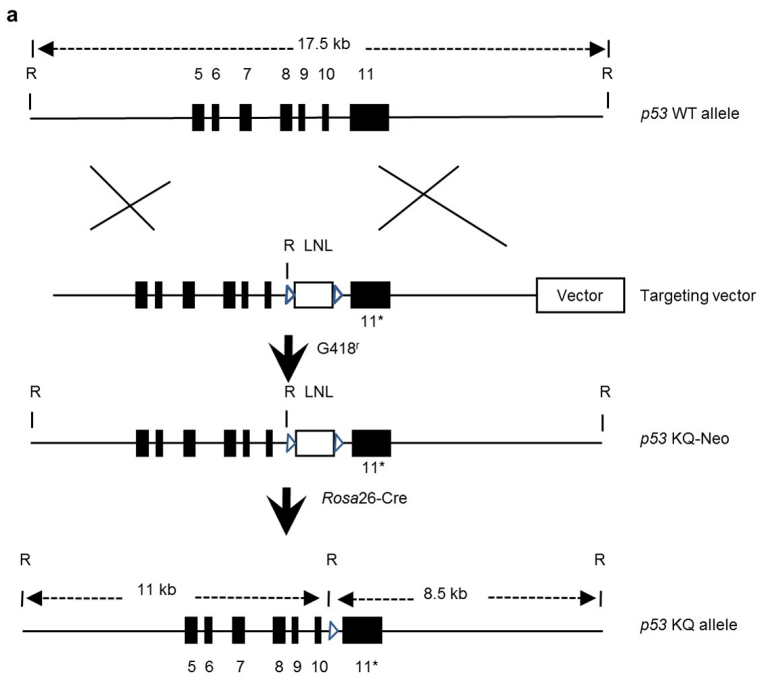
Protein Name	Acidic Domain Sequence	Length of Acidic Domain	Distribution Pattern of Acidic Amino Acid in Acidic Domain	Percentage of Acidic Amino Acid in Acidic Domain
SET	DDEEGEAEDDDDDDEEEEGLEDIDEEGDEDEGEEDDDE DEGEEGEEDEGEDD	52 (226-277)	Clustered	81%
VPRBP	EDEDEEEDQEEEEQEEEDDDEDDDDTDDLDELDTDQLL EAELEEDD	46 (1396-1441)	Clustered	76.10%
DAXX	ETDDEDDEESDEEEEEEEEEEEEEATDSEEEEDLEQMQ EGQEDDEEEDEEEE	52 (434-485)	Clustered	78.80%
PELP1	DEEEEEEEEEEEEEEEEEEEEEDFEEEEDEEEYFEEE EEEEEEFEEEFEEEEGELEEEEEEEDEEEEELEEVED	76 (887-962)	Clustered	88.20%

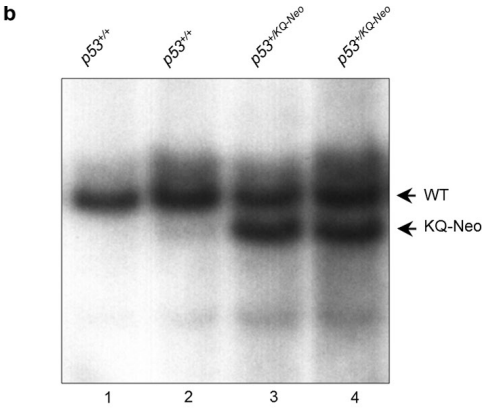


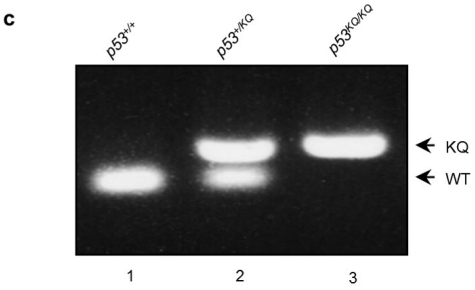
Extended Data Figure 5 | Acetylation regulates the interaction between acidic-domain-containing proteins and their acetylatable ligands. **a**, A summary table of characteristic features of the acidic-domain-containing proteins SET, VPRBP, DAXX and PELP1. The acidic amino acids are underlined. **b**, *In vitro* binding assay of p53 CTD and purified full-length VPRBP, DAXX or PELP1. **c**–**e**, Western blot analysis of the interaction between p53 and VPRBP (**c**), DAXX (**d**) or PELP1 (**e**) in the nuclear fraction of H1299 cells. **f**–**h**, *In vitro* binding assay between purified SET and KRD of H3 (**f**), KU70 (**g**) or FOXO1 (**h**). **i**, *In vitro* binding assay of the H3 KRD and purified VPRBP, DAXX or PELP1. **j**, *In vitro* binding assay of the H3 KRD and BRD4 or BRD7 (nuclear extract). Uncropped blots can be found in Supplementary Fig. 1.

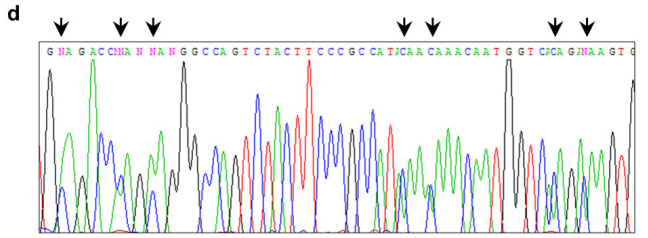


Extended Data Figure 6 | p53^{KQ} mutant mimics acetylated p53. **a**, Schematic diagram of human unacetylated p53 and the acetylation-deficient and acetylation-mimicking mutants of p53. **b**, *In vitro* binding assay of SET and different types of p53, as indicated. **c**–**e**, Western blot analysis of the interaction between acidic-domain-containing proteins (**c**, VPRBP; **d**, DAXX; **e**, PELP1) and different types of p53 in cells. H1299 cells were co-transfected with indicated constructs, and the nuclear extract was analysed by Co-IP assay. Asterisks indicate the purified proteins. Uncropped blots can be found in Supplementary Fig. 1.







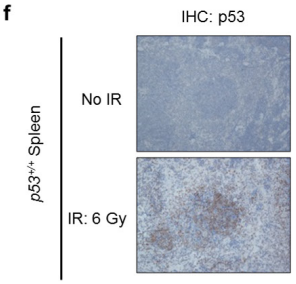


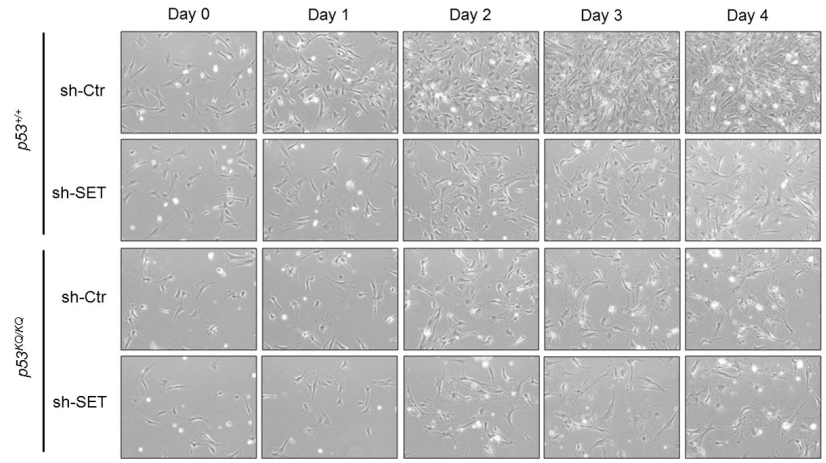
e

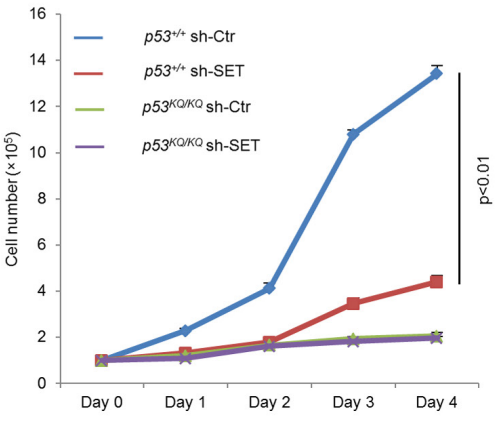
$p53^{+/KQ}$ intercross

	$p53^{+/+}$	$p53^{+/KQ}$	$p53^{KQ/KQ}$
E13.5	30	40	19
P0.5	32	50	18*
P19.5	12	26	0

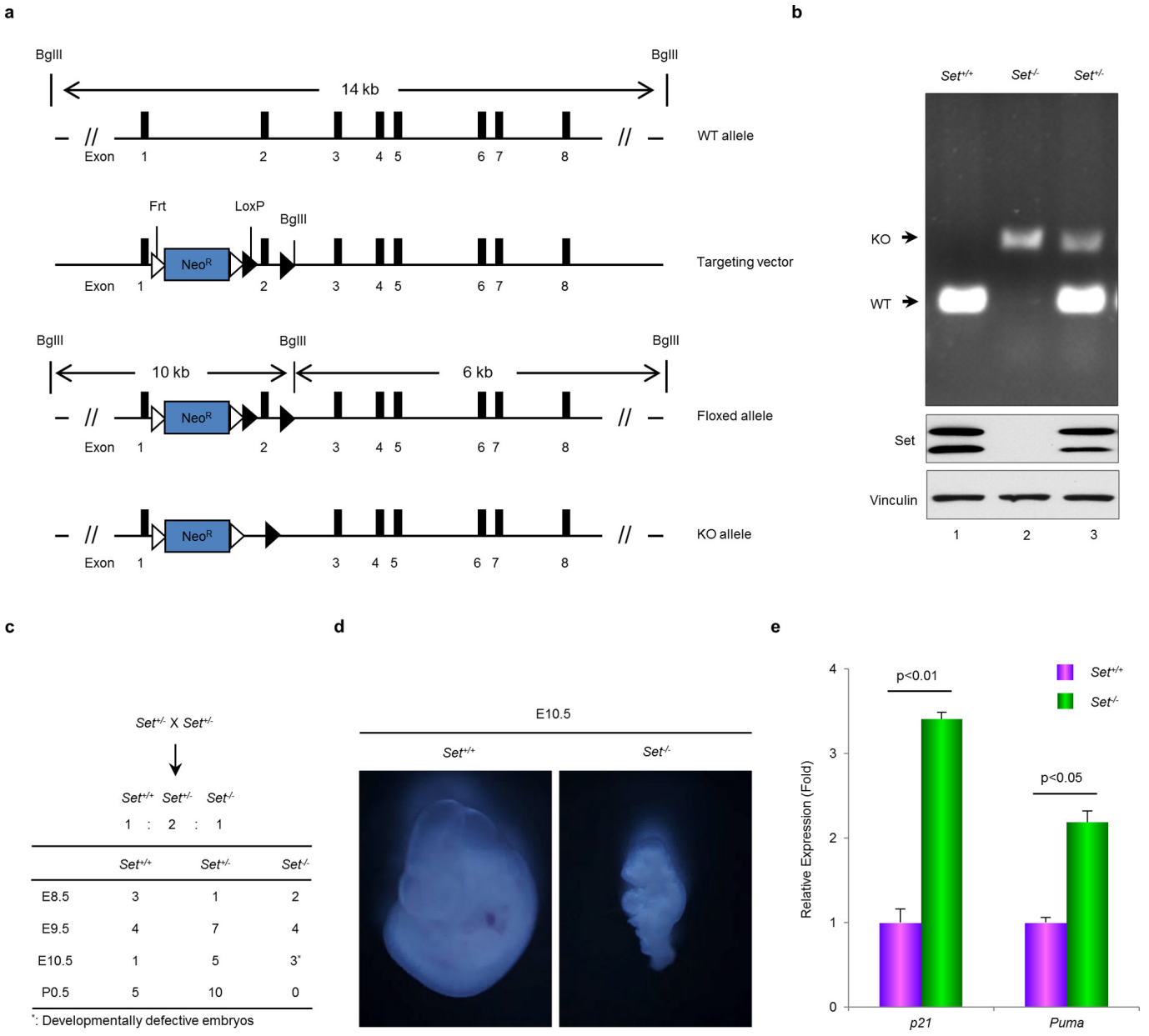
*: Neonatal lethality; E13.5: Embryonic day 13.5; P0.5, P19.5: Postnatal day 0.5, 19.5.







Extended Data Figure 7 | Generation of $p53^{KQ/KQ}$ mice. a, Schematic diagram of the gene targeting strategy to replace the p53 C-terminal 7 lysines with 7 glutamines in mouse *p53*. **b**, Southern blot screening of ES cells to identify $p53^{+/KQ}$ clones. **c**, PCR genotyping analysis of wild-type (110 bp), $p53^{+/KQ}$ heterozygous (110 bp and 150 bp), and $p53^{KQ/KQ}$ homozygous mice (150 bp only). **d**, Sequencing analysis of the transcripts prepared from the $p53^{+/KQ}$ heterozygous mouse spleen. **e**, A summary table of observed numbers of mice from $p53^{+/KQ}$ heterozygous intercrosses. **f**, Positive control for p53 staining in the IHC assay. The spleen tissue sections of $p53^{+/+}$ mice treated with or without 6 Gy γ-radiation was stained with p53 (CM-5) antibody. **g**, **h**, Representative image (**g**) or quantitative analysis (**h**) of SET-knockdown-mediated cell growth of $p53^{+/+}$ or $p53^{KQ/KQ}$ MEFs (P2). Error bars indicate mean ± s.d., $n = 3$ for biological replicates. Uncropped blots can be found in Supplementary Fig. 1.



	$Set^{+/+}$	$Set^{+/-}$	$Set^{-/-}$
E8.5	3	1	2
E9.5	4	7	4
E10.5	1	5	3*
P0.5	5	10	0

*: Developmentally defective embryos

Extended Data Figure 8 | Characterization of *Set* conditional knockout mice. a, Schematic diagram of the strategy to generate *Set* conditional knockout mice. **b**, Validation of *Set* knockout in embryos (E8.5) by genotyping and western blot analysis. **c**, A summary table of observed numbers of embryos or pups from $Set^{+/-}$ intercrosses. **d**, Representative pictures of $Set^{+/+}$ and $Set^{-/-}$ embryos (E10.5). **e**, qPCR analysis of the expression of p53 target genes in $Set^{+/+}$ and $Set^{-/-}$ embryos (E10.5). Error bars indicate mean ± s.d., $n = 3$ for technical replicates. Data are shown as representative of three experiments. Uncropped blots can be found in Supplementary Fig. 1.

Extended Data Table 1 | A list of human proteins containing acidic domains with a minimum percentage of acidic residues of 76% within a 36-residue window

	UniProt ID	Protein Name	Acidic Domain Position	Acidic Domain Sequence	Biological Function (GO)
Proteins Involved in Gene Expression Control through DNA Binding, Transcription Regulation and Chromatin Remodeling	Q8IZL8	Proline-, glutamic acid- and leucine-rich protein 1	886 - 963	DEEEEEEEEEEEEEEEEEEEEDFEEEEDEEEYFEEEEEEEFEEEFEEEEGELEEEEEEEDEEEEEELEEVEDLE	Chromatin binding, Transcription factor binding, poly(A) RNA binding
	Q92688	Acidic leucine-rich nuclear phosphoprotein 32 family member B	156 - 232	DSDAEVDGVDEEEEDEEGEDEEDEDDEDGEEEEFDEEDDEDEVEGDEDDDEVSEEEEEFGLDEEDEDEDEDEEEEE	Protein binding, Histone binding, RNA polymerase binding
	Q9UL68	Myelin transcription factor 1-like protein	107 - 169	EEYSEDNDEPGDEDEEDEEGDREEEEEIEEEDEDDDEDGEDVEDEEEEEEEEEEEEEEEENED	Sequence-specific DNA binding, Transcription factor, Zinc binding
	Q01538	Myelin transcription factor 1	257 - 315	EEEDEEEEEEEEEEEEDEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEAAPDVIFQED	Sequence-specific transcription factor, Zinc binding
	A1YPR0	Zinc finger and BTB domain-containing protein 7C	124 - 178	EIMEPGGDGGEEDDKEDDDDDEDDDDEEDEEEEEEEEEDDDDTEDFADQENLPD	Nucleic acid binding, Metal ion binding
	Q86V15	Zinc finger protein castor	1671 - 1724	ESSQEDEEEELELPEEEAEDDEDEDDDEDDDDDEDDDEDDDEDLRTDSEESLPE	DNA binding, Metal ion binding
	Q01105	Protein SET	236 - 289	DMDDEEGEGEEDDDDDEEEEGLEDIDEEGDEDEGEEDEDDDEGEEGEEDEGEDD	DNA binding, Histone binding, Phosphatase inhibitor
	P0DME0	Protein SETSIP	248 - 301	DDEEGGEDDDDDDDDGDEGEEELEDIDEGDEDEGEEDEDDDEGEEGEEDEGEDD	Chromatin binding
	Q7Z6M4	Transcription termination factor 4, mitochondrial	332 - 380	DDKRASLDEDEDDDDEEDNDEDDNDEDDDDEDDDEAEDNDEDEDDDEEE	DNA binding, RNA binding, Protein binding
	Q6PL18	ATPase family AAA domain-containing protein 2	242 - 288	ESSEEGEDQEHEDDGEDEDDDEDDDDDDDDDDDDDEDDEDEEDGEEE	Histone binding, Chromatin binding, Hydrolase
	Q9BTT0	Acidic leucine-rich nuclear phosphoprotein 32 family member E	158 - 203	EEEDDEDGDEDDEEEEENEAGPPEGYEEEEEEEEEEDEDEDEDE	Histone binding, Phophatase inhibitor activity
	Q7Z6Z7	E3 ubiquitin-protein ligase HUWE1	2425 - 2469	EDEDDSQDEEEEEEEDEEDQEDDEGEEGDEDDDDDGSEMELDED	DNA binding, ligase activity, poly(A) RNA binding
	Q12873-3	Isoform 3 of Chromodomain-helicase-DNA-binding protein 3	6-48	DEEEEEEEMVVSEEEEEEEEGDEEEEEEVEAADEDDEEDDDE	DNA binding, Helicase activity, poly(A) RNA binding
	Q96KQ7	Histone-lysine N-methyltransferase EHMT2	289 - 331	EVEALTEQLSEEEEEEEEEEEEEEEEEEEEEEEEEDEESGNQSD	Histone methyltransferase, p53 binding, C2H2 Zinc finger domain
	Q8IX15	Homeobox and leucine zipper protein Homez	507 - 549	EVVVCLDEEEEEEEEELPEDDEEEEEEEEEDDDDDDDVIIQD	DNA binding, Sequence-specific transcription factor, Transcription co-repressor
	Q8WYB5	Histone acetyltransferase KAT6B	1062 - 1103	ELSKESSEEEEEEDEEEEEEEEEEEEEEDEEEEEEEEEEEE	DNA binding, Histone acyltransferase, Transcription factor binding
	P19338	Nucleolin	233 - 274	EDEDEEEDDEDEDDDDDEDDEDDDDEDDEEEEEEEEEPVKE	DNA binding, RNA binding, Protein binding
	Q5H9L4	Transcription initiation factor TFIID subunit 7-like	326 - 367	DSRSNNDDDEDEDDEDEDEDEDEDEDEDKEEEEDCSEEYLE	Transcription coactivator, Transcription factor binding, Histone acyltransferase binding
	Q13029	PR domain zinc finger protein 2	261 - 301	EVNDLGEEEEEEEEEDEEEEEDDDDELEDEGEEEASMPNE	DNA binding, sequence-specific DNA binding transcription factor, Zinc binding, histone-lysine N-methyltransferase
	P27797	Calreticulin	368 - 407	EEEEDKKRKEEEEAEDKEDDEDKDEDEEDEEDKEEDEEED	androgen receptor binding, carbohydrate binding, complement component C1q binding
	Q9UER7	DAXX_HUMAN Death domain-associated protein 6	433 - 471	ETDDEDDEESDEEEEEEEEEEEEEATDSEEEEDLEQMQE	Androgen receptor binding, Heat shock protein binding, Histone binding
	Q4LE39	AT-rich interactive domain-containing protein 4B	528 - 566	DETNKEEDEDDEEAEEEEEEEEEEEDEDDDDNNEEEEFE	DNA binding, Protein binding, Transcription regulatory region DNA binding
	Q9UPS6-2	SET1B_HUMAN Isoform 2 of Histone-lysine N-methyltransferase SETD1B	1042 - 1079	EEQESTEEEEEAEEEEEEEDDDDDSDDRDESENDDED	Histone-lysine N-methyltransferase, Nuycleotide binding, RNA binding
	P39687	AN32A_HUMAN Acidic leucine-rich nuclear phosphoprotein 32 family member A	164 - 201	EGLDDEEEDEDEEEYDEDAQVVEDEEDEDEEEEGEEED	Gene expression, Intracellular signal transduction, nucleocytoplasmic transport
	P09429	High mobility group protein B1	178 - 214	EKSKKKKEEEEDEDEDEDEEEEEDEDEDEEEDDDDE	DNA binding, Protein binding, Transcription factor binding
	Q9BT43	DNA-directed RNA polymerase III subunit RPC7-like	157 - 192	EEEVTSEEDEEKEEEEEKEEEEEEEYDEEEHEEETD	Gene expression, Innate immune response, RNA polymerase III activity
	P17480	UBF1_HUMAN Nucleolar transcription factor 1	710 - 745	ESSSEDESEDGDENEEDDEDEDDDEDDDEDEDNESE	poly(A) RNA binding, RNA pol I CORE element seq-specific DNA binding, RNA pol I upstream control element seq-specific DNA binding
	Q15911	Zinc finger homeobox protein 3	453 - 488	EKVEPAEEEAEEEEEEEEAEEEEEEEEEEEEEEDE	DNA binding, sequence-specific DNA binding transcription factor, Protein binding
	Q9UK99	F-box only protein 3	417 - 451	DEYEEMEEEEEEEEEEDEDDDSADMDESDEDDEEE	ubiquitin-protein transferase activity
	Q9Y4B6	Protein VPRBP	1395 - 1429	EDEDEEEDQEEEEQEEEDDDEDDDTDDLDELDTD	Histone kinase, Ser/Thr kinase, Protein binding
DNA-related (Replication, Repair)	P07199	Major centromere autoantigen B	403 - 446	EGEEEEEEEEEEEEEEGEGEEEEEEGEEEEEEGGEGEELGEEEE	Centromeric DNA binding, Chromatin binding, DNA binding
			504 - 537	EGGEDSDSDSEEEDDEEEDDEDEDDDDDEEDGDE	
	P20962	Parathymosin	38 - 74	EEEENGAEEEEEETAEDGEEEDEGEEEDEEEEEEDDE	DNA replication, Immune system process
RNA-related (Processing, Translation)	Q96MU7	YTH domain-containing protein 1	198 - 264	ENEEEGVEEDVEEDEEVEEDAEEDEEVDEDGEEEEEEEEEEEEEEEEEEEEYEQDERDQKEEGNDYD	poly(A) RNA binding, RNA binding
	O60841	Eukaryotic translation initiation factor 5B	528 - 566	ENPEEEEEEEEEEEEDEESEEEEEEGESEGSEGDEEDE	GTPase activity, poly(A) RNA binding, GTP binding
	P12270	Nucleoprotein TPR	1948 - 1983	DDEEEDDDENDGEHEDYEEDEEDDDDDEDDTGMGDE	chromatin binding, heat shock protein binding, mRNA binding
	Q6ZU64	Coiled-coil domain-containing protein 108	1768 - 1803	EEEEEELEEEEEEEEETEEEELGKEEIEEKEEERDE	poly(A)RNA binding
	Q9NW13	RNA-binding protein 28	223 - 257	EEEDMEEEENDDDDDDDDEEDGVFDDEDEEEENIE	nucleotide binding, poly(A) RNA binding
	Q9UQ88	Cyclin-dependent kinase 11A	291 - 323	EEEEEEEEEEEEEGSTSEESEEEEEEEEEEEE	ATP binding, cyclin-dependent protein ser/thr kinase
	P21127	Cyclin-dependent kinase 11B	303 - 335	EEEEEEEEEEEEEGSTSEESEEEEEEEEEEE	ATP binding, cyclin-dependent protein ser/thr kinase, poly(A) RNA binding
Other	Q5TCY1	Tau-tubulin kinase 1	732 - 779	EEEEEEEEDEEEEEDEEEEEEEEEEEEEEEEEEEEEAAAAVALGE	ATP binding, protein serine/threonine kinase activity
	P46060	Ran GTPase-activating protein 1	358 - 404	DDEDEEEEEEGEEEEEEAEEEEEEDEEEEEEEEEEEEPQQRGQGE	GTPase activator activity
	Q5JTC6	APC membrane recruitment protein 1	369 - 410	EEMALPDDDDEEEEEEEEVELEEEEEEVKEEEEDDDLEYLWE	beta-catenin binding, phosphatidylinositol-4,5-bisphosphate binding
	O60721	Sodium/potassium/calcium exchanger 1	854 - 894	DGGDSEEEEEEEEEQEEEEEEEEQEEEEEEEEEEEEKGNEE	calcium, potassium:sodium antiporter activity, symporter activity
	P21817	RYR1_HUMAN Ryanodine receptor 1	1872 - 1911	EEEEEEDEEEEGEEEDEEEKEEDEEETAQEKEDEEKEEEE	Calcium ion channel, Calmoduling binding
	O43847	NRDC_HUMAN Nardilysin	141 - 179	DDEEEEEVEEEEDDDEDSGAEIEDDDEEGFDDEDEFDDE	Epidermal growth factor binding, Metalloendopeptidase, Zinc ion binding
Function not clear	Q86TY3	Uncharacterized protein C14orf37	604 - 651	DQLESEEGQEDEDEEDEDEDEEEEDEEEDEEDKDADSLDEGLDGDTE	Membrane
	Q7L0X2	Glutamate-rich protein 6	16 - 63	DQKESEEELEEEEEEEEVEEEEEEVEEEEEEVEEEEEEVVEEELVGEE	NA
	Q8TC90	Coiled-coil domain-containing glutamate-rich protein 1	301 - 344	EEEEEVEDEEEEVEDEEEEEVEEAEYVEEGEEELEEEELEEEEE	NA
	P0C7V8	DDB1- and CUL4-associated factor 8-like protein 2	107 - 146	EEETEREEDEEIQEEGGEEEEEEEEEEEEEEEEEEEE	NA

Proteins are clustered into different categories depending on the biological process in which they are involved. Each protein is described by UniProt accession code (1st column), protein name (2nd column) and a list of GO terms (5th column). The corresponding acidic domains are described by their position in the coding sequence (3rd column) and their sequence (4th column).

Extended Data Table 2 | A list of human proteins containing KRDs with at least five lysines where three or more lysines are annotated as acetylation sites in the SSPKA database

	UniProt ID	Protein Name	Acetylated Lysines	Sequence of Lysine-rich Domain
Transcription Factor	O15525	Transcription factor MafG	53, 60, 71, 76	EEIVQLKQRRRTLKNRGYAASCRVKRVTQKEELEKQ
	P18146	Early growth response protein 1	422, 424, 425	KIHLRQKDKKADKSVV
	P52630	Signal transducer and activator of transcription 2	182, 184, 194, 197	RYKIQAKGKTPSLDPHQTKEQKILQETL
	Q16236	Nuclear factor erythroid 2-related factor 2	533, 536, 538, 541, 543, 548, 554, 555	QDLDHLKDEKEKLLKEKGENDKSLHLLKKQLSTLY
	Q9Y2Y9	Krueppel-like factor 13	166, 168, 180	LESPQRKHKCHYAGCEKVYGKSSHLKA
	P04150	Glucocorticoid receptor	480, 492, 494, 495	PACRYRKCLQAGMNLEARKTKKKIKGIQ
	P43694	Transcription factor GATA-4	312, 319, 321, 323	RPLAMRKEGIQTRKRKPKNLNKSK
	P06733*	Alpha-enolase	60, 71, 80, 89	KTRYMGKGVSKAVEHINKTIAPALVSKKLNVTEQEKIDKLMI
	P23769	Endothelial transcription factor GATA-2	389, 390, 399, 403, 405, 406, 408, 409	NRPLTMKKEGIQTRNRKMSNKSKKSKKGAECFE
Transcriptional Regulation (Except Transcription Factor), Chromatin Remodeling	O60563	Cyclin-T1	380, 386, 390	SQKQNSKSVPSAKVSLKEYRAKH
	P04406*	Glyceraldehyde-3-phosphate dehydrogenase	251, 254, 259, 260	LTCRLEKPAKYDDIKKVVKQAS
	P06748*	Nucleophosmin	141, 150, 154, 155	LLSISGKRSAPGGGSKVPQKKVKLAAD
			250, 257, 267, 273	VEDIKAKMQASIEKGGSLPKVEAKFINYVKNCFRMT
	P09874	Poly [ADP-ribose] polymerase 1	498, 505, 508	VVAPRGKSGAALSKKSKGQVKEE
	P19338	Nucleolin	70, 79, 87	VVVSPTKKVAVATPAKKAAVTPGKKAAATP
			102, 109, 116, 124, 132	KTVTPAKAVTTPGKKGATPGKALVATPGKKGAAIPAKGAKNGK
	P51531	Probable global transcription activator SNF2L2	996, 997, 999, 1003	DGSEKDKKGKGGAKTLMNTI
			1547, 1551, 1553, 1555, 1556	LNKKDDKGRDKGKGKKRPNRGK
	Q00987	E3 ubiquitin-protein ligase Mdm2	466, 467, 469, 470	ACFTCAKKLKKRNKPCP
	Q13547	Histone deacetylase 1	432, 438, 439, 441	EGEGGRKNSSNFKKAKRVKTED
	Q92793	CREB-binding protein	1797, 1806, 1809	SLPSCQKMKRVVQHTKGCKRKTNGG
			1583, 1586, 1587, 1588, 1591, 1592, 1595, 1597	GSQGDSKNAKKKNNKKTNKNKSSISRA
	Q92831	Histone acetyltransferase KAT2B	416, 428, 430, 441, 442	SSSPACKASSGLEANPGEKRKMTDSHVLEEAKKPRVMGD
	P27695*	DNA-(apurinic or apyrimidinic site) lyase	24, 27, 31, 32, 35	RTEPEAKKSKTAAKKNDKEAAGEG
	P62805	Histone H4	6, 9, 13, 17, 21, 32	MSGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRL
	Q92922	SWI/SNF complex subunit SMARCC1	345, 346, 354, 359	SRKKSGKKGQASLYGKRRSQKEEDEQE
	P26358	DNA (cytosine-5)-methyltransferase 1	1111, 1113, 1115, 1117, 1119, 1121	SPGNKGKGKGKGKGKPKSQACEP
	Q13569	G/T mismatch-specific thymine DNA glycosylase	83, 84, 87	KKPVESKKSGKSAKSKE
	Q8TEK3	Histone-lysine N-methyltransferase, H3 lysine-79 specific	397, 398, 401	PSKARKKKLNKKGRKMA
	Q92841	Probable ATP-dependent RNA helicase DDX17	108, 109, 121, 129	GGGLPPKKFGNPGERLRKKKWDLSELPKFEKNFY
	P68431	Histone H3.1	5, 10, 15, 19, 24, 28, 37, 38	MARTKQTARKSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRP
	Q92522	Histone H1x	179, 182, 185	KKGAGAKKDKGGKAKKTAA
	P46100	Transcriptional regulator ATRX	1933, 1935, 1936, 1939	YTKKKKKGKKGKKDSSSSG
	Q6DN03	Putative histone H2B type 2-C	13, 16, 17, 21, 24	FAPAPKKGSKKAVTKAQKKDGKKR
	P05114	Non-histone chromosomal protein HMG-14	3, 5, 14, 18, 27, 31, 38, 42, 48, 53, 55, 59, 61	MPKRKVSSAEGAAKEEPKRRSARLSAKPPAKVEAKPKKAAAKDKSSDKKVQTKGKRGAK GKQAEVAN
DNA Repair and Integrity	P12956	X-ray repair cross-complementing protein 6	539, 542, 544, 553, 556	DYNPEGKVTKRKHDNEGSGSKRPKVEYSEE
	Q9UQE7	Structural maintenance of chromosomes protein 3	105, 106, 113, 114	RRVIGAKKDQYFLDKKMVTKND
	P27695*	DNA-(apurinic or apyrimidinic site) lyase	24, 27, 31, 32, 35	RTEPEAKKSKTAAKKNDKEAAGEG
Other DNA Related Function	O94761	ATP-dependent DNA helicase Q4	376, 380, 382, 385, 386	RSRLLRKQAWKQKWRKKGECFGG
Ribosome Biogenesis	P06748*	Nucleophosmin	141, 150, 154, 155	LLSISGKRSAPGGGSKVPQKKVKLAAD
			250, 257, 267, 273	VEDIKAKMQASIEKGGSLPKVEAKFINYVKNCFRMT
Specific Molecular/Biological Function Uncertain	P81534	Beta-defensin 103	48, 54, 61, 66, 67	VLSCLPKEEQIGKCSTRGRKCCRRKK
	Q3BBV0	Neuroblastoma breakpoint family member 1	1101, 1103, 1105, 1106	VGEIEKKGKGKKRRGRRS
	Q8N7X0	Androglobin	337, 340, 343	KDGKEVKDVKEFKPESSLT
	Q6ZQR2	Uncharacterized protein C9orf171	237, 240, 246	EQKATQKAIKLEKKQKVVLGKL
Others	P04406*	Glyceraldehyde-3-phosphate dehydrogenase	251, 254, 259, 260	LTCRLEKPAKYDDIKKVVKQAS
	P09622	Dihydrolipoyl dehydrogenase, mitochondrial	267, 271, 273, 277	FQRILQKQGFKFKLNTKVTGATK
	P40939	Trifunctional enzyme subunit alpha, mitochondrial	350, 353, 359	HGQVLCKKNKFGAPQKDVKHLA
	Q9NP61	ADP-ribosylation factor GTPase-activating protein 3	223, 228, 229	KPNQAKKGLGAKKGSLGAQ
	Q9Y6F6	Protein MRVI1	398, 402, 405	EKRFAGKAGGKLAKAPGLKD
	P02768	Serum albumin	205, 214, 223, 229, 236	AACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVAR
			543, 548, 560, 565, 569, 581, 584, 588, 597, 598	ICTLSEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEE GKKLVAASQ
	P62328	Thymosin beta-4	4, 12, 15	MSDKPDMAEIEKFDKSKLKKT
	Q13576	Ras GTPase-activating-like protein IQGAP2	1467, 1471, 1474	SIKLDGKGEPKGAKRAKPVK
	Q15283	Ras GTPase-activating protein 2	208, 209, 211	PSRNDQKKTKVKKKTS
	Q99075	Proheparin-binding EGF-like growth factor	96, 97, 99, 104	EHGKRKKKGKGLGKKRDPCLR
	P06733*	Alpha-enolase	60, 71, 80, 89	KTRYMGKGVSKAVEHINKTIAPALVSKKLNVTEQEKIDKLMI
	P15692	Vascular endothelial growth factor A	142, 147, 149, 152	RARQEKKSVRGKGKGQKRKRKKS
	P10636	Microtubule-associated protein tau	571, 574, 576, 584, 591, 597, 598, 607, 615	VPMPDLKNVKSKIGSTENLKHQPGGGKVQIINKKLDLSNVQSKCGSKDNIKHVPGGG

Each protein is described by its UniProt accession code and protein name (1st and 2nd column, respectively). Acetylated motifs are described by the position of their annotated acetylation sites within the coding sequence and their sequence (3rd and 4th column, respectively).