代表性科技论文及教研教改论文

内容 **页码**

Article

Human HDAC6 senses valine abundancy to regulate DNA damage

https://doi.org/10.1038/s41586-024-08248-5

Received: 18 January 2022

Accepted: 17 October 2024

Published online: 20 November 2024

Check for updates

Jiali Jin[1,11], Tong Meng[1,2,11], Yuanyuan Yu[3,11], Shuheng Wu[4,11], Chen-Chen Jiao[1], Sihui Song[1], Ya-Xu Li[1], Yu Zhang[1], Yuan-Yuan Zhao[1], Xinran Li[1], Zixin Wang[5], Yu-Fan Liu[1], Runzhi Huang[1], Jieling Qin[1], Yihua Chen[6,7], Hao Cao[8], Xiao Tan[1], Xin Ge[1], Cong Jiang[1], Jianhuang Xue[5], Jian Yuan[9], Dianqing Wu[10], Wei Wu[4], Ci-Zhong Jiang[3] & Ping Wang[1]✉

As an essential branched amino acid, valine is pivotal for protein synthesis, neurological behaviour, haematopoiesis and leukaemia progression[1–3]. However, the mechanism by which cellular valine abundancy is sensed for subsequent cellular functions remains undefined. Here we identify that human histone deacetylase 6 (HDAC6) serves as a valine sensor by directly binding valine through a primate-specific SE14 repeat domain. The nucleus and cytoplasm shuttling of human, but not mouse, HDAC6 is tightly controlled by the intracellular levels of valine. Valine deprivation leads to HDAC6 retention in the nucleus and induces DNA damage. Mechanistically, nuclear-localized HDAC6 binds and deacetylates ten-eleven translocation 2 (TET2) to initiate active DNA demethylation, which promotes DNA damage through thymine DNA glycosylase-driven excision. Dietary valine restriction inhibits tumour growth in xenograft and patient-derived xenograft models, and enhances the therapeutic efficacy of PARP inhibitors. Collectively, our study identifies human HDAC6 as a valine sensor that mediates active DNA demethylation and DNA damage in response to valine deprivation, and highlights the potential of dietary valine restriction for cancer treatment.

The abundance of intracellular amino acids is precisely sensed by a complex apparatus to regulate various cellular signalling pathways and functions[4,5]. Insufficient intracellular amino acid levels may block the release of tRNA molecules from general control non-derepressible 2 (GCN2) and consequently activate the GCN2–ATF4 pathway, which is an indirect mechanism for sensing amino acid levels[6]. Additionally, the intracellular levels of some amino acids are sensed by the mechanistic target of rapamycin complex 1 (mTORC1) by directly binding to specific sensors[7–10]. Species-specific sensors for amino acids have been identified[11–13]. However, it remains unclear whether amino acids can also be directly sensed in a tRNA-independent and mTORC1-independent manner. Notably, epigenetics is regulated by the metabolism of amino acids[14,15], but the underlying mechanisms remain elusive. Here we present evidence that human HDAC6 is a valine sensor for regulating DNA damage by dictating TET2-mediated DNA demethylation.

Valine binds to the SE14 domain of HDAC6

To identify the potential sensor for valine, we conjugated biotin to the amino group of valine and affinity-purified potential valine-binding proteins. Identification of valyl-tRNA synthetase validated our pull-down assay (Fig. 1a, Extended Data Fig. 1a and Supplementary Table 1). HDAC6, a microtubule-associated deacetylase[16,17], was identified as a top hit (Fig. 1a, Extended Data Fig. 1a and Supplementary Table 1). We confirmed that valine directly binds to purified recombinant HDAC6 using equilibrium binding assays (Fig. 1b and Extended Data Fig. 1b) and thermal shift assays (Extended Data Fig. 1c,d). Radiolabelled valine also bound to purified HDAC6, which was outcompeted by an excess of non-radioactive valine (Fig. 1b,c and Extended Data Fig. 1b). The dissociation constant (K_d) for the valine–HDAC6 interaction was around 1.95 μM, as determined by a competition binding assay (Fig. 1c). The specific binding between HDAC6 and valine was also confirmed using a biotin pull-down assay (Fig. 1d). Collectively, we conclude that HDAC6 selectively binds valine.

We subsequently examined the structural features of valine binding to HDAC6 by analysing the interaction of several analogues of valine to HDAC6. Modification of the carboxyl group and side chain of valine, but not its N-terminal group, blocked its binding to HDAC6 (Fig. 1e and Extended Data Fig. 1g), To map the valine-binding domain of HDAC6, the Ser-Glu-containing tetradecapeptide (SE14) repeat domain was identified, and its deletion abolished the binding of HDAC6 to valine (Fig. 1f and Extended Data Fig. 1h,i). Isothermal titration calorimetry assays showed that this domain bound to valine with a K_d of around 26.5 μM

[1]Tongji University Cancer Center, Shanghai Tenth People's Hospital, School of Medicine, Tongji University, Shanghai, China. [2]Department of Orthopedics, Shanghai General Hospital, School of Medicine, Shanghai Jiaotong University, Shanghai, China. [3]Key Laboratory of Spine and Spinal Cord Injury Repair and Regeneration of Ministry of Education, Orthopaedic Department of Tongji Hospital, School of Life Sciences and Technology, Tongji University, Shanghai, China. [4]Key Laboratory of Multi-Cell Systems, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, Shanghai, China. [5]Key Laboratory of Spine and Spinal Cord Injury Repair and Regeneration of Ministry of Education, Tongji Hospital affiliated to Tongji University, Frontier Science Center for Stem Cell Research, School of Life Sciences and Technology, Tongji University, Shanghai, China. [6]Shanghai Key Laboratory of Regulatory Biology, East China Normal University, Shanghai, China. [7]School of Pharmaceutical Sciences and Yunnan Key Laboratory of Pharmacology for Natural Products and Yunnan College of Modern Biomedical Industry, Kunming Medical University, Kunming, China. [8]School of Life Science and Bio-Pharmaceutics, Shenyang Pharmaceutical University, Shenyang, China. [9]Department of Biochemistry and Molecular Biology, Tongji University School of Medicine, Shanghai, China. [10]Department of Pharmacology, Yale School of Medicine, New Haven, CT, USA. [11]These authors contributed equally: Jiali Jin, Tong Meng, Yuanyuan Yu, Shuheng Wu. ✉e-mail: wangp@tongji.edu.cn

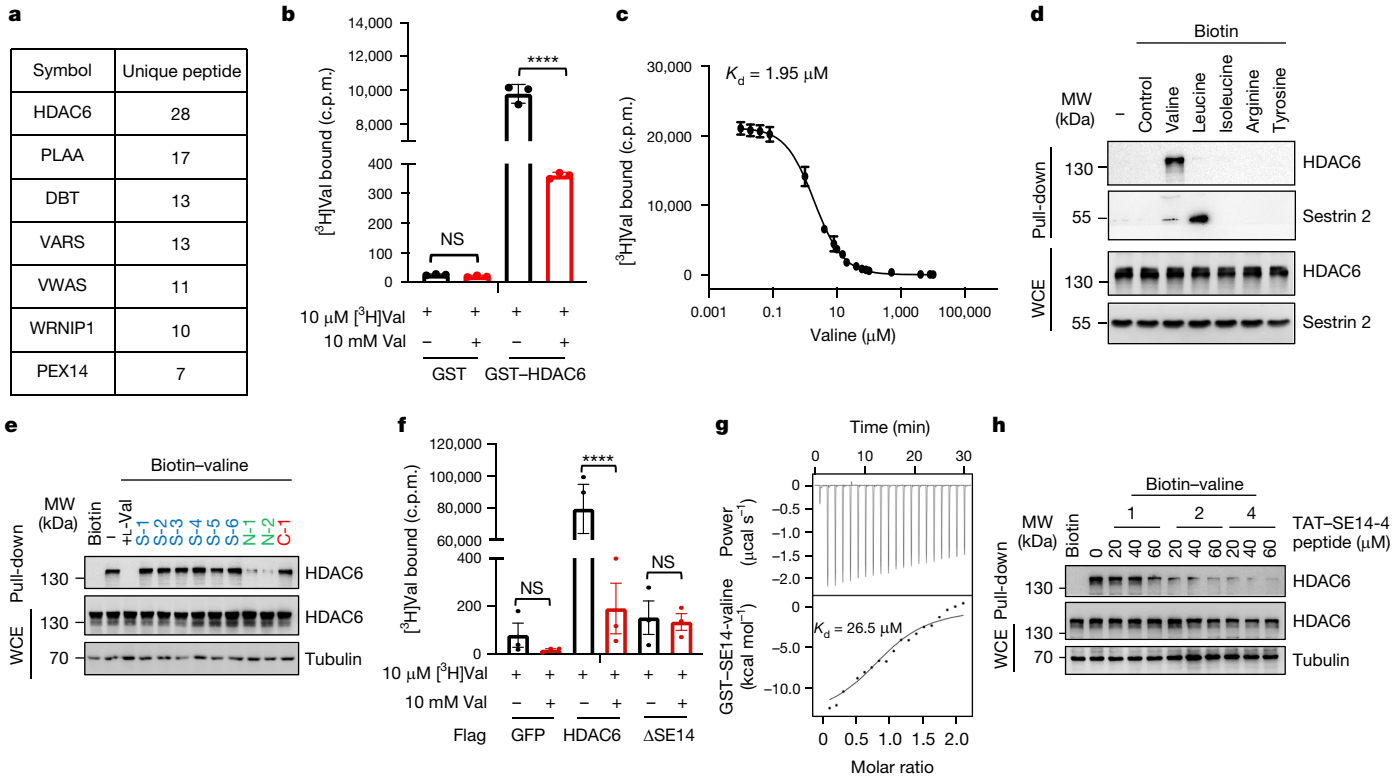


Fig. 1 | Valine directly binds to the SE14 repeat domain of human HDAC6. **a**, Candidate valine-binding proteins with more than seven unique peptides identified by MS of biotin-labelled valine (mass spectral peptide counts shown). **b**, Binding of [^{3}H]valine to GST–HDAC6 but not to the control GST produced from *Escherichia coli*. **c**, Flag–HDAC6 prepared from HCT116 cell extracts were used in binding assays with [^{3}H]valine. HDAC6 binds [^{3}H]valine with a K_d of about 1.95 μM. **d**, Immunoblot analysis of HDAC6 proteins pulled down from HCT116 cells using biotinylated amino acids and avidin-beads. WCE, whole cell extract. **e**, Immunoblot analysis of HDAC6 proteins incubated with valine (10 mM) or its analogue (10 mM) for 6 h with valine deprivation for 6 h (structures are in Extended Data Fig. 1g) pulled down from HCT116 cells through biotinylated valine. **f**, Binding of [^{3}H]valine to full-length, but not to the SE14 repeat domain truncated, HDAC6 immunoprecipitated from HCT116 cell extracts, in which flag-tagged proteins were overexpressed. **g**, Binding of valine to the GST–SE14 repeat domain of HDAC6 produced in *E. coli* was measured by isothermal titration calorimetry. **h**, Effect of these SE14 peptides on the interaction between biotin–valine and HDAC6. Biotin pull-down was prepared from HCT116 cells incubated with the control or the TAT–SE14 peptides of HDAC6 with valine deprivation for 6 h. For **b**, **c** and **f**, data are the mean ± s.d. ($n = 3$ independent experiments). Statistical analysis was performed using one-way analysis of variance (ANOVA); ****$P < 0.0001$. NS, not significant. For gel source data, see Supplementary Fig. 1.

(Fig. 1g). The SE14 repeat domain possesses seven SE14 repeats with the consensus sequence XLXQTXSEXAXGGA (where X represents any amino acid residue), which were categorized into 3 types of peptides with conserved sequences (residues 919–932, 933–946 and 960–974) (Extended Data Fig. 1j). We therefore synthesized these three peptides and investigated their interaction with valine using a competition assay based on biotin pull-down. The SE14-4 peptide (residues 960–974) had the highest binding affinity to valine (Fig. 1h), a result that was further confirmed in an electrochemical assay (Extended Data Fig. 1k).

In a homology modelling analysis of HDAC6 and valine, valine bound to a certain extent to a hydrophobic pocket of the SE14 repeat domain, directly interacting with Ala946, Thr989 and Ala991 (Extended Data Fig. 1l,m). The HDAC6-3M mutant (A946D, T989A and A991D) construct had markedly weakened interactions with valine (Extended Data Fig. 1n). As human HDAC6 possesses seven consecutive SE14 repeats and each may bind to valine (Extended Data Fig. 1o), we mutated all the conserved sites in SE14 repeats (marked in red in Extended Data Fig. 1j), and this HDAC6 mutant (SE-Mut HDAC6) was unable to bind valine (Extended Data Fig. 1p). Together, these data validate the essential role of the SE14 repeat domain in valine binding.

Valine dictates the localization of HDAC6

HDAC6, a class IIb histone deacetylase, is mainly localized in the cytoplasm to regulate tubulin acetylation and mTOR signalling[18,19]. However, valine deprivation led to negligible effects on tubulin acetylation and mTORC1 regulation (Extended Data Fig. 2a,b). As HDAC6 is also found in the nucleus[20,21] and the SE14 repeat domain stably retains human HDAC6 in the cytoplasm[22], we examined whether valine binding affects the subcellular localization of HDAC6. HDAC6 was predominantly cytosolic in HCT116 cells cultured in normal medium, whereas deprivation of valine, but not leucine or isoleucine, induced the nuclear localization of the majority of HDAC6 (Fig. 2a and Extended Data Fig. 2c). Re-supplementation of valine restored the cytoplasmic localization of HDAC6 in a dosage-dependent and time-dependent manner (Extended Data Fig. 2d and Fig. 2b–d). The half-maximal concentration of valine to retain HDAC6 in the cytosol was about 15 μM, whereas valine deprivation reduced cellular valine to 1 μM (Extended Data Fig. 2e–g). This result indicated that valine deprivation-induced HDAC6 nuclear translocation may be due to the loss of valine binding to HDAC6. As extracellular valine is transported into the cytosol through SLC7A5 (ref. 3), we blocked SLC7A5 through genetic knockout or by using a specific inhibitor (JPH203). Both approaches promoted the nuclear translocation of HDAC6, which was rescued by *N*-methyl-L-valine (in which a methyl group is attached to the N terminus of valine), but not by L-valine-OMe (in which a methyl group is attached to the C terminus of valine) (Fig. 2e and Extended Data Fig. 2h,i). This result confirms that nuclear localization of HDAC6 is modulated by the intracellular level of valine.

We next asked whether direct binding of valine is required for the cytosolic sequestration of HDAC6. Alteration of the carboxyl group and

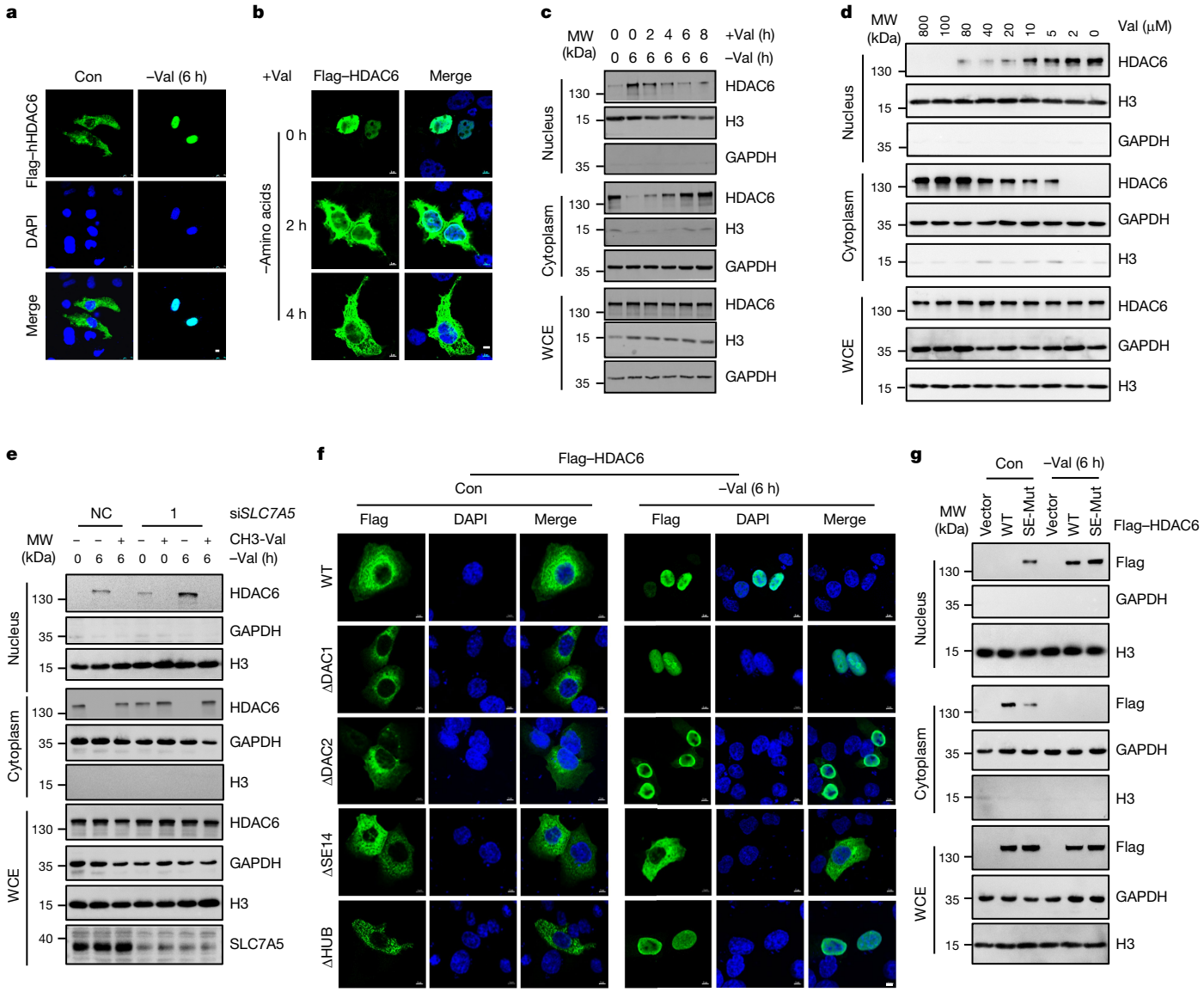


Fig. 2 | Intracellular valine abundancy dictates the subcellular localization of HDAC6. a, Immunofluorescence analysis of subcellular localization of human HDAC6 (hHDAC6) in HCT116 cells expressing Flag–HDAC6 after valine deprivation for 6 h. Con, control. **b**, Immunofluorescence analysis of cellular localization of HDAC6 in HCT116 cells expressing Flag–HDAC6 after amino acid deprivation for 6 h and supplemented with valine for the indicated times. **c**, Cell fractionation analysis of cellular localization of HDAC6 in HCT116 cells pretreated with or without valine for 6 h and supplemented with valine for different times. **d**, Cell fractionation analysis of cellular localization of HDAC6 in HCT116 cells in valine-deprived medium for 6 h and supplemented with valine at the corresponding concentration gradient for 6 h. **e**, Effect of the branched amino acid transporter SLC7A5 on the subcellular localization of HDAC6. The cell fractionation assay was performed using HCT116 cells with SLC7A5 wild-type (WT) or knockdown (si*SLC7A5*) after valine deprivation for 6 h, and rescued with *N*-methyl-L-valine (CH3-Val, 0.8 mM) for 6 h. NC, negative control. **f**, Immunofluorescence images indicating the functional importance of the SE14 repeat domain in valine sensing. **g**, The subcellular localization of HDAC6 WT and valine-binding-deficient mutant (SE-Mut) with or without valine deprivation for 6 h. Scale bar, 10 μm (**a**,**b**,**f**). For gel source data, see Supplementary Fig. 1.

side chain of valine, but not its amino group, resulted in the disruption of its binding to HDAC6 and the capacity to restore the cytoplasmic localization of HDAC6 from the nucleus induced by valine deprivation (Extended Data Fig. 2k,l). Moreover, SE-Mut HDAC6 was largely localized in the nucleus regardless of valine availability (Fig. 2g). Together, these data indicate that binding of valine to HDAC6 through the SE14 repeat domain is required for its cytosolic localization of HDAC6.

We next examined the mechanism underlying valine deprivation-induced HDAC6 nuclear translocation. Valine deprivation did not affect the subcellular localization of HDAC4 or SIRT7, two nuclear–cytoplasmic shuttling proteins, and had a minimal effect on the acetylation of HDAC6 (Extended Data Fig. 3a,b). As HDAC6 possesses one nuclear import sequence (NLS) and two nuclear export sequences (NESs)[22], we examined their effects on valine deprivation-induced nuclear retention. Treatment with the importin inhibitor importazole blocked valine deprivation-induced HDAC6 nuclear translocation (Extended Data Fig. 3c). Deletion of the NLS, but not the NESs, suppressed HDAC6 nuclear translocation induced by either importazole or valine deprivation (Extended Data Fig. 3d). This result indicated that a classical, importin-mediated pathway is required for HDAC6 nuclear translocation. Consistently, binding of importin-α1 and importin-α7 to HDAC6 was readily detected after valine deprivation and was blocked by valine re-supplementation (Extended Data Fig. 3e–g). Deletion of either the NLS or the SE14 repeat domain disrupted the interaction between HDAC6 and importin-α1 after valine deprivation (Extended Data Fig. 3h).

Primate-specific binding of valine to HDAC6

As the SE14 repeat domain is only found in HDAC6 proteins of primates (Extended Data Fig. 3i), we examined whether valine binds to HDAC6 in a species-specific manner. Mouse HDAC6, which does not have the SE14 repeat domain, failed to bind valine, and the subcellular location of mouse HDAC6 was not affected by valine deprivation (Extended Data Fig. 3j,k). A chimera of human SE14 with mouse HDAC6 restored its binding to valine (Extended Data Fig. 3l). We generated chimeric mice in which the human SE14 repeat was knocked into mouse HDAC6, and the cytosolic localization of this chimeric protein was controlled by valine in mouse embryonic fibroblasts (MEFs) (Extended Data Fig. 3m–r). Furthermore, in primate *Rhinopithecus bieti* fibroblasts, the SE14 repeat domain of HDAC6 bound to valine, and valine deprivation increased the nuclear localization of HDAC6 (Extended Data Fig. 3s,t). Therefore, regulation of HDAC6 subcellular localization by valine binding requires the primate-specific SE14 repeat domain.

Nuclear HDAC6 promotes the activity of TET2

To investigate the biological function of valine deprivation-induced HDAC6 nuclear translocation, we examined the nuclear proteins that specifically bind HDAC6 in a valine-dependent manner. Using stable isotope labelling of amino acids in cell culture (SILAC) coupled with mass spectrometry (MS), we identified 356 proteins that showed enhanced binding to HDAC6 after valine deprivation. TET2, one of the DNA dioxygenases that converts 5-methylcytosine (5mC) into 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) to reverse DNA methylation[23], was one of the top ranked hits (Extended Data Fig. 4a, Fig. 3a and Supplementary Table 2). The intracellular interaction between HDAC6 and TET2 after valine deprivation was further confirmed by co-immunoprecipitation and data-independent acquisition-based MS (Extended Data Fig. 4b,c). However, addition of valine to cell extracts did not block the interaction between HDAC6 and TET2 (Extended Data Fig. 4d). This result indicated that the enhanced binding of HDAC6 to TET2 after valine deprivation is not a result of affecting its direct interaction but contingent on the nuclear translocation of HDAC6. Furthermore, HDAC6 bound to the C-terminal cysteine-rich dioxygenase domain of TET2 through its tandem deacetylase domains (DAC1 and DAC2) (Extended Data Fig. 4e,f).

We next investigated whether valine deprivation-induced nuclear localization of HDAC6 affects TET2 activity. As a dioxygenase, TET2 converts 5mC into 5hmC to regulate gene expression, chromatin stability and tumorigenesis[24–26]. Valine deprivation significantly increased the cellular level of 5hmC in a HDAC6-dependent manner (Fig. 3b and Extended Data Fig. 4g–o). In addition, valine deprivation induced a genome-wide and loci-specific increase in 5hmC levels and a corresponding decrease in 5mC levels in wild-type cells but not in cells in which HDAC6 or TET2 was knocked down, as assessed using APOBEC-coupled epigenetic sequencing[27] and whole-genome bisulfite sequencing (Fig. 3c,d, Extended Data Fig. 4o,p and Supplementary Table 3). Together, these data suggest that valine deprivation may promote TET2 activity to induce DNA demethylation in a HDAC6-depedent manner.

As TET2 may bind to DNA to promote DNA demethylation, we used chromatin immunoprecipitation with sequencing (ChIP–seq) to examine the DNA-binding sites of TET2 induced by valine deprivation. Principal component analysis of TET2-binding signals was separated into four groups: wild-type with or without valine deprivation, and HDAC6 knockdown with or without valine deprivation (Extended Data Fig. 4q). We identified 24,774 TET2-binding regions, and a subset of regions (about 3,932, cluster 1) had valine deprivation-increased TET2 binding under HDAC6 regulation (Fig. 3e, Extended Data Fig. 4r and Supplementary Table 4). Furthermore, 5hmC, 5fC and 5caC levels in the valine deprivation-specific enhanced TET2-binding regions were increased after valine deprivation in a TET2-dependent manner (Extended Data Fig. 4s). These data suggest that valine deprivation induces chromatin binding of TET2, which is associated with DNA demethylation and 5hmC, 5fC and 5caC generation.

As HDAC6 is a deacetylase, we explored whether its enzymatic activity modulates TET2 activity. TET2 acetylation was detected in cells cultured with normal medium and was reduced by valine deprivation in a HDAC6-dependent and SE14 domain-dependent manner (Fig. 3f,g). We purified TET2 from HCT116 cells and incubated it with bacterially purified recombinant HDAC6. HDAC6 promoted TET2 deacetylation and significantly increased its activity to generate 5hmC in vitro (Extended Data Fig. 4t,u). The deacetylase activity of HDAC6 was essential for its effect on TET2 deacetylation and activity (Fig. 3h,i). Additionally, deletion of the SE14 repeat domain abolished the valine deprivation-enhanced 5hmC level without affecting HDAC6 deacetylase activity (Extended Data Fig. 4v). Furthermore, depletion of SIRT1, a deacetylase of TET2, had a minor effect on valine deprivation-induced TET2 deacetylation and DNA demethylation in HCT116 cells (Extended Data Fig. 4w–y). Collectively, these data indicate that valine deprivation promotes TET2 activity that depends on both the nuclear translocation of HDAC6 and its deacetylase activity.

TET2 is acetylated at three conserved lysine residues (K1472, K1473 and K1478), and lysine to arginine (KR) mutations increase TET2 catalytic activity[28]. Therefore, we transduced cells with the C-terminal catalytic domain (CD) of TET2 (TET2-CD), which contains a cysteine-rich dioxygenase domain and these acetylation sites. The activity of wild-type TET2-CD was enhanced by valine deprivation in a HDAC6-dependent manner, whereas the 3KR mutant generated increased levels of 5hmC independent of valine deprivation or HDAC6 deletion (Fig. 3j,k). Additionally, re-expression of patient-derived catalytic inactive mutants (H1881Q and R1896S) did not promote active DNA demethylation after valine deprivation. This result provides further support that valine deprivation increases DNA hydroxymethylation in a TET2 activity-dependent manner (Fig. 3l,m). Together, these data suggest that nuclear HDAC6 promotes TET2 activity by deacetylating TET2 after valine deprivation, probably at the three conserved sites (Extended Data Fig. 4z).

VR promotes DNA damage through HDAC6–TET2–TDG

Accumulating evidence has shown that TET2 activity has distinct roles in DNA damage[29–32]. Thus, we examined whether valine restriction (VR) promotes DNA damage through the HDAC6–TET2 axis. Valine deprivation for 24 h significantly increased the levels of H2AX phosphorylation (γH2AX) and poly(ADP)-ribosylation (PAR), two well-characterized markers of DNA damage (Extended Data Fig. 5a–h). Deletion of TET2 or HDAC6 blocked valine deprivation-induced DNA damage (Fig. 4a–c and Extended Data Fig. 5i).

We next examined how TET2-mediated DNA demethylation regulates DNA damage after valine deprivation. TET2-mediated DNA damage mainly depends on the formation of single-strand DNA breaks (SSBs) driven by thymine DNA glycosylase (TDG), an enzyme responsible for the recognition and excision of the oxidized methylcytosines 5fC and 5caC[26,29] (Extended Data Fig. 5j). Thus, we explored the role of TDG-mediated SSBs in valine deprivation-induced DNA damage. First, both valine deprivation and TDG knockdown upregulated 5fC and 5caC levels, as assessed by UPLC–MS/MS and dot blot assays (Extended Data Fig. 5k,l). Additionally, valine deprivation induced the formation of SSBs, as assessed using ddC S1-END-seq (Extended Data Fig. 5m). Furthermore, TDG deficiency almost completely blocked valine deprivation-induced DNA damage (Fig. 4d–f). All these data confirm the crucial role of the TET2–TDG axis in valine deprivation-induced DNA damage.

Cancer cells exhibit both active and passive demethylation mechanisms. Active demethylation (replication-independent)

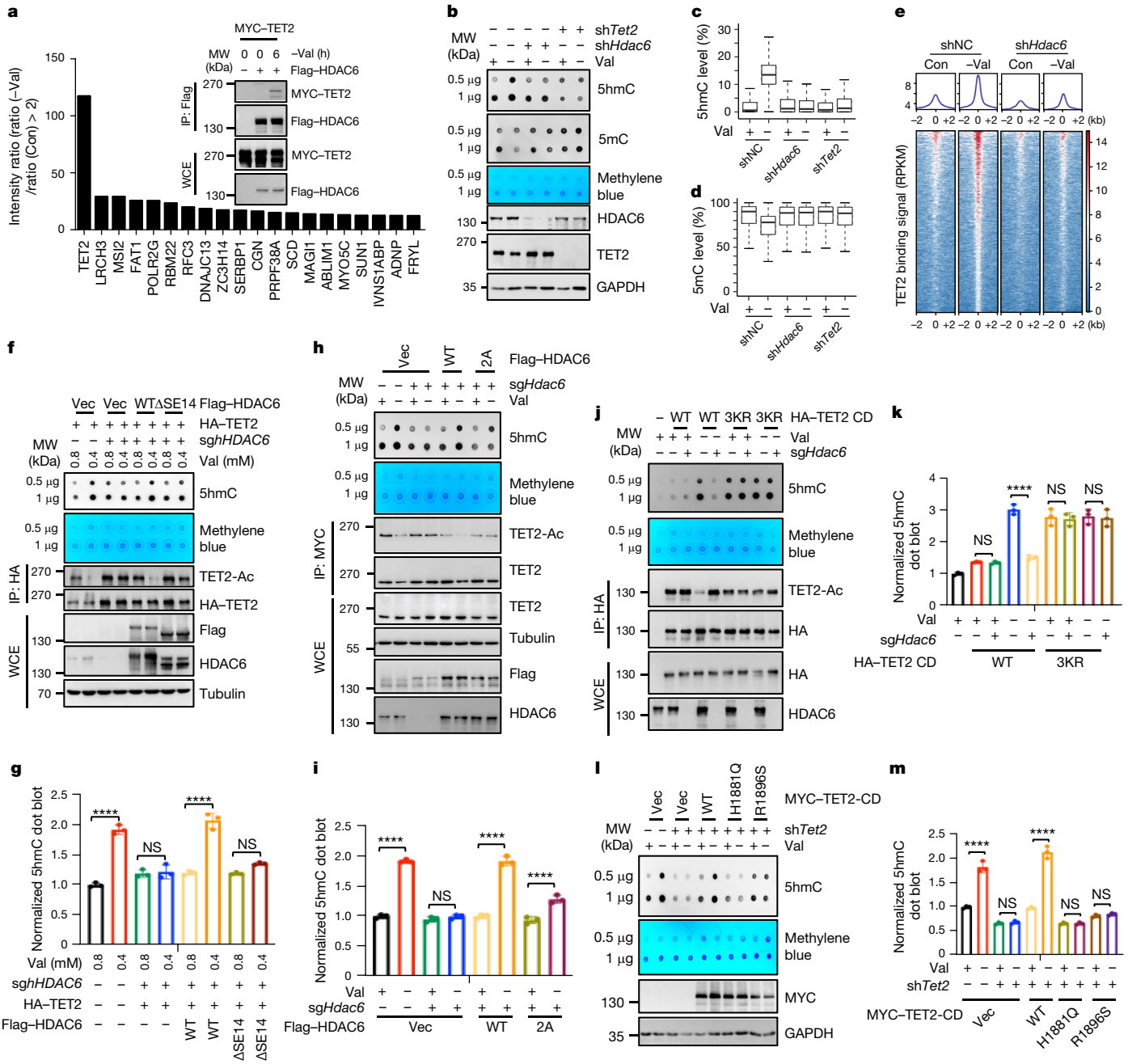


Fig. 3 | Nuclear HDAC6 promotes TET2-dependent active DNA demethylation. **a**, Ranking list of HDAC6-binding proteins identified by proteomics, and the interaction between HDAC6 and TET2 through co-immunoprecipitation (IP). **b**, Effect of valine deprivation (24 h) on the levels of global 5hmC and 5mC after knockdown of HDAC6 (sh*Hdac6*) or TET2 (sh*Tet2*). **c**,**d**, Boxplots showing 5hmC (**c**) and 5mC (**d**) levels at all hyperhydroxymethylation regions (13,058) identified from WT (shNC), HDAC6 or TET2 knockdown HCT116 cells after valine deprivation (24 h). Boxplots represent the median, 25% quantile (Q1) and 75% quantile (Q3). The lower whisker extends to the smallest value within Q1–1.5× the interquartile range (IQR), the upper whisker extends to the largest value within Q3 + 1.5× the IQR. **e**, Average plots (top) and heatmaps (bottom) showing enrichment of TET2-binding signals in WT or HDAC6 knockdown (sg*Hdac6*) HCT116 cells after valine deprivation for 24 h at regions of valine deprivation-specific enhanced TET2 binding. RPKM, reads per kilobase per million. **f**,**g**, The level of 5hmC and acetylation of TET2 (TET2-Ac) treated with 0.8 mM or 0.4 mM valine for 24 h (**f**), and quantification of dot blots (**g**). **h**,**i**, The acetylation of immunoprecipitated TET2 and 5hmC levels in HDAC6 knockout HCT116 cells expressing WT or HDAC6 catalytically dead (2A) with or without valine deprivation for 24 h (**h**), and quantification of dot blots (**i**). **j**,**k**, Immunoprecipitation of the acetylation of TET2 and 5hmC in HDAC6 knockout HCT116 cells expressing WT TET2-CD or TET2-CD-3KR (K1472, 1473 and K1498R) after 24 h of valine deprivation (**j**), and quantification of dot blots (**k**). **l**,**m**, The level of 5hmC in TET2 knockdown HCT116 cells expressing WT or TET2 catalytic dead (H1881Q and R1896S) (**l**), and quantification of dot blots (**m**). For **k**, **g**, **l** and **m**, data are presented as the mean ± s.d. ($n = 3$ independent experiments). Statistical analysis was performed using one-way ANOVA. ****$P < 0.0001$. NS, not significant. Vec, vector. For gel source data, see Supplementary Fig. 1.

is orchestrated by TET methylcytosine dioxygenases and TDG, whereas passive demethylation is mediated by replication-dependent dilution[33]. To exclude the influence of cell replication-mediated passive DNA demethylation on DNA damage, we blocked the cell cycle with palbociclib (also known as PD0332991), an inhibitor of cyclin-dependent kinase 4 and cyclin-dependent kinase 6 (CDK4/6)[34].

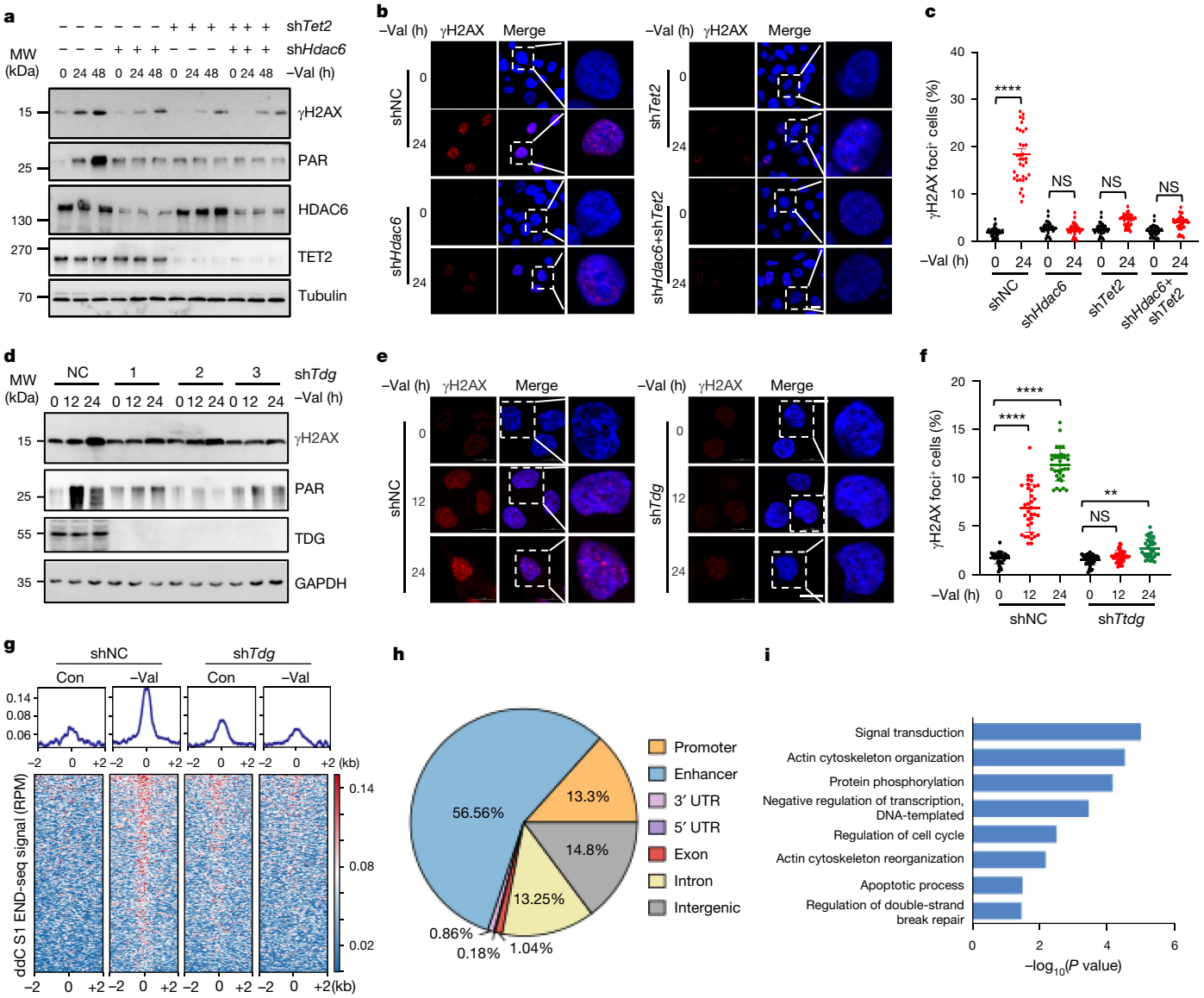


Fig. 4 | Valine deprivation promotes the DNA damage response through the HDAC6–TET2 axis. a, Immunoblot of the DNA damage marker γH2AX and PAR in HDAC6 and TET2 knockdown HCT116 cells after valine deprivation at different time points. **b**,**c**, Representative images from three independent experiments for each group (**b**) and quantification (**c**) of immunofluorescence staining for γH2AX (red) in HDAC6 and TET2 knockdown HCT116 cells after valine deprivation for 24 h. $n = 3$ in each group. **d**, Immunoblot of DNA damage marker γH2AX and PAR in TDG knockdown (sh*Tdg*) HCT116 cells after valine deprivation at different time points. **e**,**f**, Representative images (**e**) and quantification (**f**) of immunofluorescence staining for γH2AX (red) in HDAC6 and TET2 knockdown HCT116 cells after valine deprivation for 12 h and 24 h. $n = 3$ in each group. **g**, Average plots (top) and heatmaps (bottom) illustrate the enrichment of ddC S1-END-seq signals in both WT and TDG knockdown HCT116 cells. These enrichments were observed after 24 h of valine deprivation, followed by 24 h of treatment with palbociclib, specifically at valine deprivation-specific enhanced TET2-binding regions. 0, peak centre. RPM, reads per million. **h**,**i**, Genomic distribution (**h**) and GO analysis (**i**) of valine deprivation-specific enhanced TET2-binding regions, at which 5hmC, 5fC and 5caC levels and DNA damage signals were also increased. UTR, untranslated region. For **c** and **f**, data are the mean ± s.d. ($n = 35$ microscopy views examined across 3 independent experiments). Statistical analysis was performed using Mann–Whitney *U*-test. $^{**}P < 0.01$, $^{****}P < 0.0001$. For GO analysis, one-sided Fisher's exact test *P* values are applied to evaluate gene enrichment in annotation terms. Scale bar, 10 μm (**b**,**e**). For gel source data, see Supplementary Fig. 1.

We then evaluated the role of TDG-mediated active DNA demethylation in valine deprivation-induced DNA damage. Palbociclib arrested 90% of cells in the G1 phase and increased the levels of 5hmC, 5fC and 5caC (Extended Data Fig. 5n,o). Notably, valine deprivation still induced DNA damage in a TDG-dependent manner when treated with palbociclib, as assessed by immunofluorescence staining and ddC S1 END-seq analysis (Extended Data Fig. 5p,q and Fig. 4g). Collectively, these data provide further support that TET2-mediated active DNA demethylation has important roles in valine deprivation-induced DNA damage.

We also analysed the co-localization of SSBs with 5hmC and 5fC sites in the valine deprivation-specific enhanced TET2-binding regions. These DNA damage sites were predominantly located in enhancers (as defined by H3K4me1 peaks), a result that aligned with a previous report in which 5hmC and DNA damage were primarily localized at enhancer regions[29] (Fig. 4h). These findings provide further support for our model that valine deprivation enhances TET2 activity, which in turn promotes active DNA demethylation that leads to SSBs and double-strand DNA breaks. A gene ontology (GO) analysis showed that genes annotated in these SSB-enriched regions were related to

double-strand break repair, cell cycle regulation and apoptotic processes (Fig. 4i and Supplementary Table 5). Consistently, valine deprivation also downregulated the expression of genes related to functions such as base-excision repair, mismatch repair and DNA replication, and induced cell cycle arrest in the G1 phase in a TET2-dependent and HDAC6-dependent manner (Extended Data Fig. 5r–x and Supplementary Tables 6 and 7). Taken together, these data strongly implicate the crucial role of TET2–TDG-mediated active DNA demethylation in valine deprivation-induced DNA damage.

Investigation of the roles of binding between valine and HDAC6 in the induction of DNA damage after valine deprivation showed that the SE14 repeat domain of HDAC6 is required for valine deprivation-induced DNA damage (Extended Data Fig. 6a–g). These results were further confirmed by examining the levels of γH2AX, PAR, tail moment, *pole2*, *eme1* and *lig1* in SE14 repeat domain knock-in MEFs (Extended Data Fig. 6h–r). Consistently, SE-MUT HDAC6, which does not bind to valine and is predominantly located in nucleus, induced DNA demethylation and DNA damage under normal conditions (Fig. 2g and Extended Data Fig. 6t,u).

To further confirm the specific role of HDAC6 nuclear localization in mediating valine deprivation-induced DNA damage, we constructed a nuclear-localized HDAC6 (nHDAC6) by adding six classical NLSs (6×NLS-HDAC6, which contains 2 NLSs of SV40 T antigen, 2 NLSs of MYC and 2 NLSs of HDAC6) to the HDAC6 N terminus and confirmed its nuclear location (Extended Data Fig. 7a–c). nHDAC6 enhanced DNA hydroxymethylation and induced DNA damage in a TET2-dependent manner (Extended Data Fig. 7d–g). ChIP–seq further confirmed the nHDAC6-enhanced genomic distribution of TET2 in HDAC6 knockout HCT116 cells (Extended Data Fig. 7h). These data support the important role of HDAC6 nuclear localization in valine deprivation-induced DNA hydroxymethylation and associated DNA damage.

Dietary VR inhibits cancer progression

Persistent DNA damage can inhibit tumour progression by promoting programmed cell death[35], DNA replication[36] and antitumour immunity[37]. Given that valine deprivation promotes DNA damage, we investigated whether VR could be a therapeutic option for cancer. Valine deprivation reduced the number and size of oncospheres in a HDAC6-dependent manner (Extended Data Fig. 7i,j). SE14 knock-in MEFs proliferated much slower than wild-type cells under valine-deficient conditions (Extended Data Fig. 7k–m). This result provides support for the idea that the SE14 domain has a role in orchestrating the valine-sensitive nuclear translocation of HDAC6 and transmitting valine availability to cell proliferation. Moreover, TDG knockdown reduced the suppressive effect of valine deprivation on cell proliferation (Extended Data Fig. 7n–p), a result consistent with our data that TET–TDG mediates DNA demethylation in valine deprivation-induced DNA damage.

To evaluate the effect of dietary VR on tumour growth in vivo, HCT116 xenograft models were used. A previous study[38] showed that complete valine depletion had obvious inhibitory effects on tumour growth, but led to serious side effects such as fatty liver and body weight loss. To identify an appropriate VR strategy, we switched the diet from normal chow (0.82% valine, w/w) gradually to VR diets (0.70%, 0.41% and 0.16% valine, w/w). This strategy selectively influenced the concentration of valine with no discernible impact on other amino acids in plasma and tumour samples (Extended Data Fig. 8a–f). The effect of different VR concentrations in the diet on body weight and tumour growth were also examined. A dietary VR of 0.16% valine (w/w) induced tumour regression in mice but resulted in up to 50% loss of body weight. By contrast, a dietary VR of 0.70% valine (w/w) resulted in mild body weight loss but minor tumour regression. Notably, a VR diet of 0.41% valine (w/w) significantly inhibited tumour growth and had mild side effects on body weight (Extended Data Fig. 8g–j). This result is consistent with a previous study[2] showing that partial VR was associated with no overt body weight loss. Compared with normal chow, similar or higher amounts of food intake were observed in mice on VR diets (Extended Data Fig. 8k), which implied that inhibitory effects were not due to caloric restriction. To further confirm the therapeutic potential of VR in human cancer, a preclinical patient-derived xenograft (PDX) model of colorectal cancer was used (Fig. 5a). The 0.41% (w/w) VR diet significantly inhibited tumour growth in both prevention and treatment groups (Fig. 5b–e and Extended Data Fig. 8l,m).

The role of the HDAC6–TET2 axis in mediating dietary VR-suppressed tumour growth was explored. The concentration of valine was about 40 μM in tumour tissue from mice fed a normal diet (0.82% valine) compared with about 3 μM in tumour tissue from mice fed the VR diet (0.41% valine) (Extended Data Fig. 8d), lower than the half-maximal concentration of valine (15 μM) to retain HDAC6 in the cytoplasm (Fig. 2d and Extended Data Fig. 2e). Notably, a progressive reduction in valine levels correlated with an incremental increase in the nuclear translocation of HDAC6, as well as the extent of DNA damage within samples (Extended Data Fig. 8n–p). These results further confirm that dietary VR (0.41% valine, 0.16% valine, w/w) can promote the nuclear translocation of HDAC6 in tumours.

In rescue assays performed in HCT116 cell-derived xenograft models, both TET2 and the SE14 repeat domain of HDAC6 played essential parts in inhibiting tumour growth by dietary VR (Fig. 5f,g and Extended Data Fig. 9a–f). Re-expressing human, but not mouse HDAC6, in HDAC6-deficient cells rescued the suppressive effect of dietary VR on tumour growth and DNA damage in tumour xenografts derived from HDAC6-deficient cells (Fig. 5h and Extended Data Fig. 9g–i). These findings indicate that VR inhibits tumour progression in a manner that depends on the SE14 repeat domain of HDAC6.

VR also significantly enhanced 5hmC levels in tumour tissues (Fig. 5i,j and Extended Data Fig. 9j,k), a result consistent with the role of VR in DNA damage (Extended Data Fig. 10a–d). Additionally, VR resulted in increased foci of γH2AX and PAR and higher levels of 5hmC in xenografts originating from wild-type cells compared with those derived from cells lacking the SE14 repeat domain in HDAC6 or with suppressed TET2 (Extended Data Fig. 10e–h). We observed that limiting valine had negligible effects on 5hmC levels, reduced γH2AX and PAR foci, and had minimal effects on tumour size in xenografts derived from cells expressing the 3KR TET2 mutant (Extended Data Fig. 10i–o). These findings further emphasize the crucial role of HDAC6-mediated deacetylation of TET2 in valine deprivation-induced DNA damage.

Increased DNA damage is a hallmark of tumorigenesis, and inducing DNA damage is an anticancer therapy. This effect can be achieved by inhibiting DNA repair using clinically used PAPR inhibitors[39]. We therefore used VR in combination with the PARP inhibitor talazoparib. VR significantly increased the effect of talazoparib both in vitro (Extended Data Fig. 7q–u) and in vivo (Fig. 5k–m and Extended Data Fig. 9l).

Discussion

Valine is an essential amino acid involved in protein synthesis, glucose homeostasis and nutrient-sensitive signalling pathways[14,40,41]. In this study, we demonstrated that valine specifically binds to the primate-specific SE14 repeat domain of HDAC6 to retain HDAC6 in the cytoplasm, implying an evolutionary advance of HDAC6 in primates. Valine deprivation promotes the binding of HDAC6 to importin-α1, leading to the nuclear translocation of HDAC6. Nuclear HDAC6 in turn increases active DNA demethylation and DNA damage by deacetylating and activating TET2 (Extended Data Fig. 10p). We also noted that valine deprivation may affect the expression of genes related to DNA damage repair and promote the binding of HDAC6 to other DNA repair and replication-related proteins, which might also participate in the DNA damage. Together, our study unveils a previously unrecognized

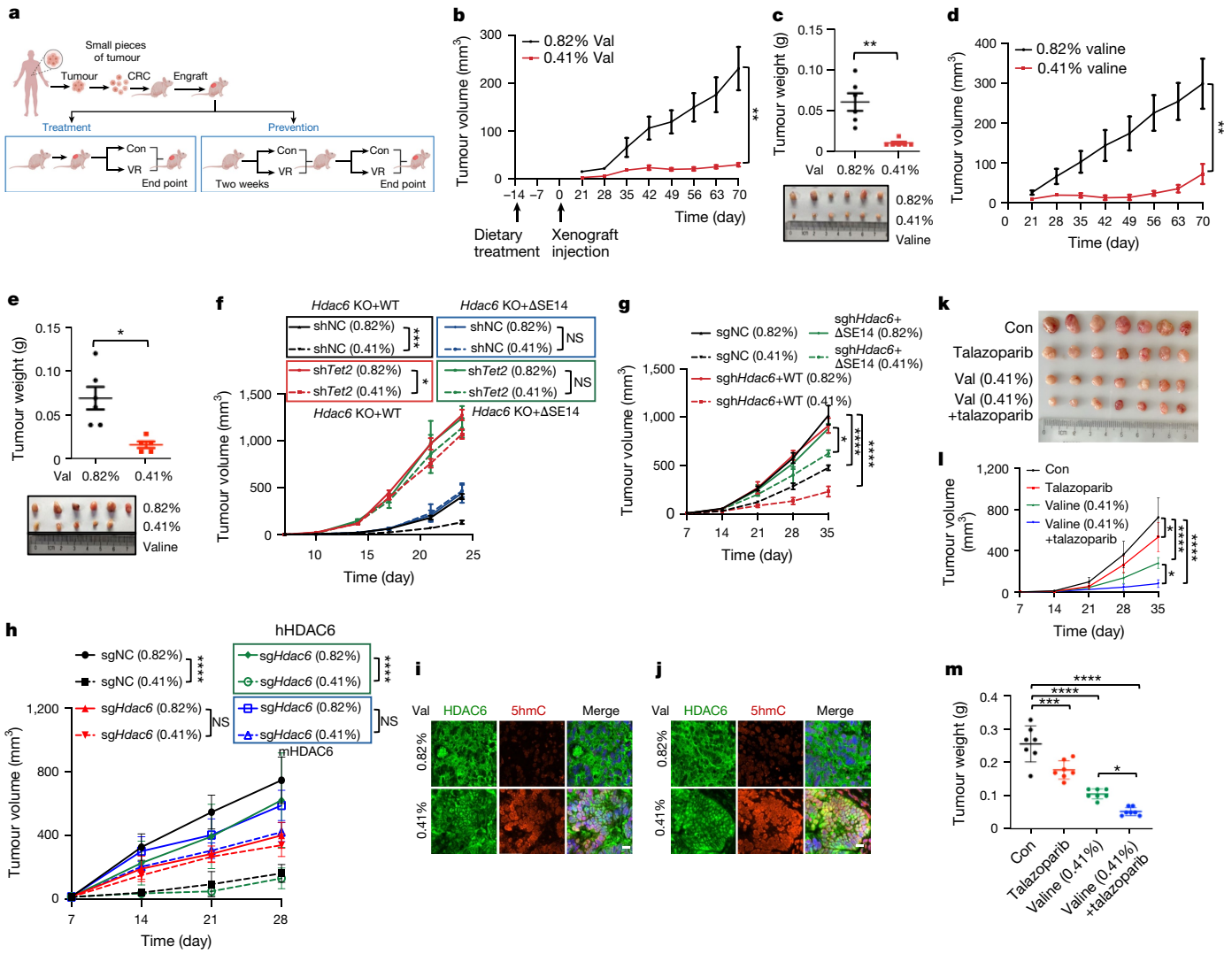


Fig. 5 | Dietary VR inhibits colon cancer progression through the HDAC6–TET2 axis. a, Schematic of the experimental design using colorectal PDXs. Con, valine normal diet (0.82%); VR, valine restriction (0.41%). **b**, Tumour volume of colorectal PDX in the prevention group. **c**, Tumour images and weight at the end point for data in **b**. **d**, Tumour volume of colorectal PDX in the treatment group. **e**, Tumour images and weight at the end point for data in **d**. **f**, Tumour volume of TET2 WT or TET2 knockdown in HDAC6 WT or HDAC6 ΔSE14 stable expressed HCT116 cell-derived tumours in nude mice fed on 0.82% or 0.41% valine diet. **g**, Tumour volume of nude mice inoculated with HDAC6 WT or knockout HCT116 cells infected with HDAC6 WT or SE14 repeat domain truncation virus fed on 0.82% or 0.41% valine diet. **h**, Tumour volume of nude mice inoculated with HDAC6 WT or knockout HCT116 cells infected with human HDAC6 (hHDAC6) or mouse HDAC6 (mHDAC6) virus fed on 0.82% or 0.41% valine diet. **i**,**j**, Immunostaining of tumour sections for HDAC6 (green) and 5hmC (red) and nuclei (blue) from the prevention group (**i**) and the treatment group (**j**). Scale bar, 10 μm. **k**,**l**, Tumour images (**k**) and volume (**l**) of HCT116 cell-derived tumours in nude mice fed on 0.82% or 0.41% valine diet, or treated with talazoparib (0.33 mg kg^{-1}) and 0.82% or 0.41% valine diet. **m**, Tumour weight at the end point for data in **k**. For **b**–**h**, **l** and **m**, data are presented as the mean ± s.d. ($n = 6$ mice (**b**–**e**,**g**,**h**), 8 mice (**f**) or 7 mice (**l**,**m**)). Statistical analysis was performed using unpaired two-tailed *t*-test (**b**–**e**) or one-way ANOVA (**f**–**h**,**l**,**m**). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$; NS, not significant. Schematic in **a** was created using BioRender (https://BioRender.com). For gel source data, see Supplementary Fig. 1.

function of valine in protecting cells from DNA damage and maintaining cell homeostasis, suggesting that HDAC6–TET2 axis may function as a nutrient checkpoint in controlling cell fate by integrating nutrient stress, DNA demethylation and cell death.

The primate-specific regulation of HDAC6 by valine has intriguing evolutionary implication, especially considering that difference of rodents and primates in DNA damage response and repair has been previously reported[42]. Furthermore, branched amino acid levels play essential roles in motor coordination[3], neuronal metabolic state and survival[43], and HDAC6 mitigates neurodegeneration in an autophagy-dependent manner[16]. As such, an evolutionary difference in amino acid sensing may play important roles in highly sophisticated processes such as nervous function, distinguishing primates from other mammals. Our current study also suggests that primates may have evolutionarily developed a mechanism to survey valine deficiency.

Dietary restriction or deprivation of amino acids is emerging as a promising adjuvant therapy for patients with cancer[2,44,45]. The benefits of VR diet in T cell acute lymphoblastic leukaemia (T-ALL) have been reported[1], but its therapeutic effect in solid tumour remains unclear. HDAC6 expression is upregulated in various cancer types, including colon cancer, bladder cancer and lung cancer, and participates in tumorigenesis and development[46]. We speculate that increasing the HDAC6 nuclear localization through VR might be a potential anti-cancer strategy. Dietary VR may also induce DNA damage and inhibit the growth and progression of colon cancer via regulating HDAC6 nuclear entry and subsequent TET2-driven DNA demethylation to alter cancer epigenetics.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-024-08248-5.

1. Thandapani, P. et al. Valine tRNA levels and availability regulate complex I assembly in leukaemia. *Nature* **601**, 428–433 (2022).
2. Taya, Y. et al. Depleting dietary valine permits nonmyeloablative mouse hematopoietic stem cell transplantation. *Science* **354**, 1152–1155 (2016).
3. Tarlungeanu, D. C. et al. Impaired amino acid transport at the blood brain barrier is a cause of autism spectrum disorder. *Cell* **167**, 1481–1494.e18 (2016).
4. Chantranupong, L., Wolfson, R. L. & Sabatini, D. M. Nutrient-sensing mechanisms across evolution. *Cell* **161**, 67–83 (2015).
5. Hu, X. & Guo, F. Amino acid sensing in metabolic homeostasis and health. *Endocr. Rev.* **42**, 56–76 (2021).
6. He, X. D. et al. Sensing and transmitting intracellular amino acid signals through reversible lysine aminoacylations. *Cell Metab.* **27**, 151–166.e6 (2018).
7. Wang, S. et al. Lysosomal amino acid transporter SLC38A9 signals arginine sufficiency to mTORC1. *Science* **347**, 188–194 (2015).
8. Wolfson, R. L. et al. Sestrin2 is a leucine sensor for the mTORC1 pathway. *Science* **351**, 43–48 (2016).
9. Wolfson, R. L. & Sabatini, D. M. The dawn of the age of amino acid sensors for the mTORC1 pathway. *Cell Metab.* **26**, 301–309 (2017).
10. Chen, J. et al. SAR1B senses leucine levels to regulate mTORC1 signalling. *Nature* **596**, 281–284 (2021).
11. Liu, G. Y., Jouandin, P., Bahng, R. E., Perrimon, N. & Sabatini, D. M. An evolutionary mechanism to assimilate new nutrient sensors into the mTORC1 pathway. *Nat Commun.* **15**, 2517 (2024).
12. Gu, X. et al. SAMTOR is an *S*-adenosylmethionine sensor for the mTORC1 pathway. *Science* **358**, 813–818 (2017).
13. Jiang, C. et al. PRMT1 orchestrates with SAMTOR to govern mTORC1 methionine sensing via Arg-methylation of NPRL2. *Cell Metab.* **35**, 2183–2199.e7 (2023).
14. Sivanand, S. & Vander Heiden, M. G. Emerging roles for branched-chain amino acid metabolism in cancer. *Cancer Cell* **37**, 147–156 (2020).
15. Wang, Z. et al. Methionine is a metabolic dependency of tumor-initiating. *cells. Nat. Med.* **25**, 825–837 (2019).
16. Pandey, U. B. et al. HDAC6 rescues neurodegeneration and provides an essential link between autophagy and the UPS. *Nature* **447**, 859–863 (2007).
17. Kawaguchi, Y. et al. The deacetylase HDAC6 regulates aggresome formation and cell viability in response to misfolded protein stress. *Cell* **115**, 727–738 (2003).
18. Hubbert, C. et al. HDAC6 is a microtubule-associated deacetylase. *Nature* **417**, 455–458 (2002).
19. Zhang, T. et al. HDAC6 regulates primordial follicle activation through mTOR signaling pathway. *Cell Death Dis.* **12**, 559 (2021).
20. Liu, Y., Peng, L., Seto, E., Huang, S. & Qiu, Y. Modulation of histone deacetylase 6 (HDAC6) nuclear import and tubulin deacetylase activity through acetylation. *J. Biol. Chem.* **287**, 29168–29174 (2012).
21. Wang, Z. et al. Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. *Cell* **138**, 1019–1031 (2009).
22. Bertos, N. R. et al. Role of the tetradecapeptide repeat domain of human histone deacetylase 6 in cytoplasmic retention. *J. Biol. Chem.* **279**, 48246–48254 (2004).
23. Jones, P. A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**, 484–492 (2012).
24. Tahiliani, M. et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930–935 (2009).
25. Zhang, H. et al. TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. *Cell Res.* **20**, 1390–1393 (2010).
26. He, Y. F. et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**, 1303–1307 (2011).
27. Schutsky, E. K. et al. Nondestructive, base-resolution sequencing of 5-hydroxymethylcytosine using a DNA deaminase. *Nat. Biotechnol.* https://doi.org/10.1038/nbt.4204 (2018).
28. Sun, J. et al. SIRT1 activation disrupts maintenance of myelodysplastic syndrome stem and progenitor cells by restoring TET2 function. *Cell Stem Cell* **23**, 355–369.e9 (2018).
29. Wang, D. et al. Active DNA demethylation promotes cell fate specification and the DNA damage response. *Science* **378**, 983–989 (2022).
30. Shukla, V. et al. TET deficiency perturbs mature B cell homeostasis and promotes oncogenesis associated with accumulation of G-quadruplex and R-loop structures. *Nat. Immunol.* **23**, 99–108 (2022).
31. An, J. et al. Acute loss of TET function results in aggressive myeloid cancer in mice. *Nat. Commun.* **6**, 10071 (2015).
32. Kafer, G. R. et al. 5-Hydroxymethylcytosine marks sites of DNA damage and promotes genome stability. *Cell Rep.* **14**, 1283–1292 (2016).
33. Oswald, J. et al. Active demethylation of the paternal genome in the mouse zygote. *Curr. Biol.* **10**, 475–478 (2000).
34. Fry, D. W. et al. Specific inhibition of cyclin-dependent kinase 4/6 by PD 0332991 and associated antitumor activity in human tumor xenografts. *Mol. Cancer Ther.* **3**, 1427–1438 (2004).
35. Groelly, F. J., Fawkes, M., Dagg, R. A., Blackford, A. N. & Tarsounas, M. Targeting DNA damage response pathways in cancer. *Nat. Rev. Cancer* **23**, 78–94 (2023).
36. Thomas, A. et al. Therapeutic targeting of ATR yields durable regressions in small cell lung cancers with high replication stress. *Cancer Cell* **39**, 566–579.e7 (2021).
37. Harding, S. M. et al. Mitotic progression following DNA damage enables pattern recognition within micronuclei. *Nature* **548**, 466–470 (2017).
38. Nishihira, T. et al. Anti-cancer therapy with valine-depleted amino acid imbalance solution. *Tohoku J. Exp. Med.* **156**, 259–270 (1988).
39. Farmer, H. et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* **434**, 917–921 (2005).
40. She, P. et al. Disruption of BCATm in mice leads to increased energy expenditure associated with the activation of a futile protein turnover cycle. *Cell Metab.* **6**, 181–194 (2007).
41. Ananieva, E. A. & Wilkinson, A. C. Branched-chain amino acid metabolism in cancer. *Curr. Opin. Clin. Nutr. Metab. Care* **21**, 64–70 (2018).
42. Jegga, A. G., Inga, A., Menendez, D., Aronow, B. J. & Resnick, M. A. Functional evolution of the p53 regulatory network through its target response elements. *Proc. Natl Acad. Sci. USA* **105**, 944–949 (2008).
43. Knaus, L. S. et al. Large neutral amino acid levels tune perinatal neuronal excitability and survival. *Cell* **186**, 1950–1967.e25 (2023).
44. Kanarek, N., Petrova, B. & Sabatini, D. M. Dietary modifications for enhanced cancer therapy. *Nature* **579**, 507–517 (2020).
45. Gao, X. et al. Dietary methionine influences therapy in mouse cancer models and alters human metabolism. *Nature* **572**, 397–401 (2019).
46. Li, T. et al. Histone deacetylase 6 in cancer. *J. Hematol. Oncol.* **11**, 111 (2018).

Methods

Ethics statement

This study was approved by the institutional review board of Shanghai Tenth People's Hospital (approval number: 2020-KN171-01). Written informed consent for tissue and clinical data collection was signed by patients or their legal guardians. All mouse experiments were administered based on the guidelines of the Institutional Animal Institutional Animal Care and Use Committees, and all the protocols were approved by Tongji University with approval number TJAA07121103.

Materials

Reagents were obtained from the following sources: HRP-labelled anti-mouse and anti-rabbit secondary antibodies from Beijing Biodragon Immunotechnologies; antibodies against HDAC6 (for western blotting (WB) and immunofluorescence, rabbit isotype), anti-tubulin, Ac-anti-tubulin, GAPDH, TDG, SIRT1, sestrin 2, MYC, HA and Flag epitope from Proteintech; antibodies against 5hmC and 5mC from Active Motif; antibodies against HDAC6 (for immunofluorescence, mouse isotype) from Abcam; antibodies against H3, SLC7A5, acetylated-lysine, TET2 (for WB and immunoprecipitation), γH2AX from Cell Signaling Technology; antibody against TET2 (for ChIP) from Diagenode; antibody against PAR from R&D Systems; antibody against GST prepared by our own laboratory; HA beads and MYC beads from Abmart; Flag M2 affinity gel, EGF, amino acids and amino acid derivatives from Sigma Aldrich; L-valine-OMe and *N*-methyl-L-valine from Bidepharm; palbociclib (PD0332991, S4482) from Selleck; 2′,3′-dideoxycytidine (ddC, D5782) from Sigma-Aldrich; valine-free and amino acid-free DMEM custom made by Shang Zhi Company; DMEM, RPMI, B27, Lipofectamine 2000, Sypro orange dye and DMEM for SILAC from Thermo Fisher; FBS and dialysed (USA) FBS from Thermo Fisher; BCA Protein Quantification kit and GSTSep glutathione agarose resin from Yeasen Biotechnology; [^{3}H]-labelled amino acids from Amersham Life Science; SLC7A5 inhibitor JPH203 from Selleck; importin inhibitor importazole from MedChemExpress; and nuclear export inhibitor leptomycin B from Beyotime Biotechnology.

Animals, diets and tissue collection

Animals were caged in groups of five in a laminar airflow cabinet under specific pathogen-free conditions, fed with sterilized food and water, and kept on a 12-h light–dark cycle. Blinding and randomization were used in the mouse experiments. The mouse HDAC6 with human SE14 knock-in mice on the C57BL6/J background were generated by Shanghai Model Organisms Center. Th primers for genotype identification are provided in Supplementary Table 8.

The special diets with defined valine levels were purchased from Trophic Animal Feed High-Tech. The control diet contained 0.82% methionine (w/w, TP01A0701), and the VR diet contained 0.70%, 0.41% or 0.16% valine (w/w, TP01A0701-15, TP01A0701-50, TP01A0701-80, respectively). Two mouse models were used and are described in the sections 'PDX models of colorectal cancer' and 'Xenograft for colorectal cancer'. For all animal studies, mice were randomized to the control or VR diet, and investigators were blinded to allocation during experiments or outcome assessment. For the VR time-course study in healthy mice, male BALB/c nude mice (Shanghai SLAC Laboratory Animal) were subjected to either the control or the VR diet. At the end point, all mice were euthanized for tissue collection, and plasma samples were collected for amino acid targeted metabolomics.

PDX models of colorectal cancer

In brief, CRC 2019-42252H tumours (ethics code 2020-KN171-01) was resected, washed and minced, and then passaged through nude mice one time. For the dietary studies, CRC 2019-42252H PDX tumours were minced in PBS at 150 mg ml^{-1}, and 200 µl of tumour suspension was subcutaneously injected into the flanks of nude mice obtained from the Shanghai SLAC Laboratory Animal. Mice were subjected to the control or VR diet (0.41% valine) either 2 weeks before the tumour injection or from when the tumour was palpable (when tumours reached approximately 100 mm^3) until the end point (a tumour volume of about 1,500 mm^3). Tumour size was monitored once per week until the end point.

Xenograft for colorectal cancer

HDAC6 WT, HDAC6 knockdown HCT116 cells were infected with lentivirus expressing vector, Flag-human HDAC6 and SE14 repeat domain truncation Flag–HDAC6 ΔSE14. TET2 WT or TET2 knockdown in HDAC6 WT or HDAC6 ΔSE14 were stably expressed in HCT116 cells. HDAC6 WT, HDAC6 knockdown HCT116 cells were infected with lentivirus expressing vector, Flag–human HDAC6 and Flag–mouse HDAC6. Stable expressing or knockdown HCT116 cell lines were trypsinized into single-cell suspensions and resuspended in PBS. Approximately 5×10^6 HCT116 cells in 100 µl were subcutaneously injected into the flanks of nude mice obtained from the Shanghai SLAC Laboratory Animal. Mice were subjected to either the control or VR diet before tumour injection or when the tumour was palpable (when tumours reached approximately 100 mm^3) until the end point (a tumour volume of about 1,500 mm^3). Tumour volume was calculated using the formula: width^2 × length × 0.5.

Combination of valine deprivation and PARP inhibitors to treat tumours

Approximately 5×10^6 HCT116 cells in 100 µl were subcutaneously injected into the flanks of nude mice obtained from the Shanghai SLAC Laboratory Animal. One week after injection, tumours were measured using callipers. Tumour volumes were calculated using the formula: width^2 × length × 0.5. When tumours reached approximately 100 mm^3, mice were randomly assigned to treatment with vehicle, VR diet (0.41% valine), talazoparib (a PARP inhibitor, 0.33 mg kg^{-1}) or a combination of VR diet (0.41% valine) and talazoparib (0.33 mg kg^{-1}). Talazoparib was diluted in diluent (10% DMAC and 6% Solutol in PBS) and administered by oral gavage every other day. For vehicle, an equivalent volume of DMSO was diluted as above. Mouse weight and tumour volume were recorded weekly. After nearly 4 weeks of treatment, mice were euthanized and the tumour weight was measured.

Cell culture and derivation of MEFs

Human embryonic kidney 293T cells (American Type Culture Collection), HCT116 cells (American Type Culture Collection) and MEF cells were cultured in DMEM supplemented with 10% heat-inactivated FBS at 37 °C in 5% CO_2. *Rhinopithecus bieti* fibroblast cells were cultured in DMEM supplemented with 10% heat-inactivated FBS and 0.1 mM NEAA. MEFs were isolated from embryonic day 12.5 embryos. Gonads and internal organs were removed before MEF isolation. The MEFs were grown in DMEM supplemented with 10% (v/v) FBS, 2 mM L-glutamax, 0.1 mM NEAA, 100 U ml^{-1} penicillin and 100 mg ml^{-1} streptomycin. Isolated MEFs in passage 1 were used for further experiments. All of these cell lines have been tested for mycoplasma contamination at regular intervals.

Immunoblotting, IP and pull-down assays

For immunoblotting, whole cell lysates were prepared using SDS loading buffer (250 mM Tris-Cl, pH 6.8, 40% glycerol, 25% SDS, 4% β-mercaptoethanol and 4% bromophenol blue). Cell lysates were separated by SDS–PAGE and proteins were visualized by enhanced chemiluminescence according to the manufacturer's instructions (Chengdu Gezhi Biotechnology). For protein–protein interactions, cells were lysed using RIPA lysis buffer (50 mM Tris-Cl, pH 7.4, 0.5% NP-40, 150 mM NaCl, 1 mM EDTA, 10% glycerophosphate and a cocktail of proteinase inhibitors). Primary antibodies were incubated with protein agarose A/G beads for 30 min at room temperature, followed by incubating with

cell lysates for 3 h with rotation at 4 °C. The beads were washed four times with lysis buffer and analysed by immunoblotting. For biotin–valine or other amino acid pull-down assays, streptavidin beads were preincubated with free biotin or biotin-labelled amino acids in binding buffer (20 mM HEPES, pH 7.9, 150 mM KCl, 1 mM DTT, 1 mM PMSF, 10% glycerol, 0.1% NP-40 and other proteinase inhibitors) for 1 h at room temperature, and then incubated with cell lysates overnight with rotation at 4 °C. The beads were washed 4 times with washing buffer (20 mM HEPES, pH 7.9, 150 mM KCl, 1 mM DTT, 1 mM PMSF, 0.1% NP-40 and other proteinase inhibitors) and analysed by immunoblotting.

Purification of proteins expressed in bacteria

For the purification of proteins expressed in bacteria, recombinant HDAC6 or the SE14 repeat domain of HDAC6 was expressed in *E. coli* (strain BL21 DE3 star) from GST–HDAC6 or GST–SE14 plasmids in pET28a or pGEX-4T-2 vectors, respectively. The bacterial cultures were grown at 37 °C to an optical density of 0.4, at which point the temperature was reduced to 18 °C. After 30 min at 18 °C, the cultures were induced overnight at 18 °C with 0.5 mM IPTG. The cells were subsequently resuspended in lysis buffer (0.5% Triton X-100 in PBS). The cells underwent mechanical homogenization and the lysates were cleared by centrifugation and then loaded onto GST-Sep glutathione agarose resin. After incubation, the resin was washed 3 times with wash buffer (0.1% Triton X-100 in PBS). The proteins were eluted with elution buffer (50 mM Tris-Cl pH 7.4, 0.3% GSH). The GST tag was removed by using thrombin (100 U ml^{-1}). The collected protein was concentrated and immediately used in binding assays or frozen at −80 °C.

Purification of proteins expressed in human cells

Around 5×10^7 HCT116 cells were plated in a 15 cm plate before the experiment. Each plate would produce the protein for one sample. At 12 h after plating, cells were transfected with pCDNA3.1-Flag–HDAC6 expression plasmids (20 μg). After 24 h of transfection, the cells were lysed as previously described[47]. Anti-Flag immunoprecipitates were prepared with the exception that before incubation with lysates, the beads were blocked by rotating in 1 μg μl^{-1} BSA for 20 min at 4 °C and subsequently washed twice in lysis buffer. Next, 30 μl of the 50/50 slurry of beads in lysis buffer was added to each of the clarified cell lysates and incubated as previously described[47].

Equilibrium binding assay

For binding assays performed with bacterially produced proteins, 30 μg HDAC6, or purification of protein expressed in human cells as mentioned as above, was diluted into 500 μl lysis buffer (50 mM Tris pH 7.4, 200 mM NaCl, 5 mM $MgCl_2$ and 0.1% CHAPS) and incubated with 30 μl compact GST resin or Flag M2 affinity gel. The GST resin or Flag M2 affinity gel with bound proteins was washed once with lysis buffer and 3 times with lysis buffer supplemented with 300 mM NaCl. After washing, the liquid was aspirated and the protein bound to the resin or gel was incubated for 1 h on ice with the appropriate amount of [^{3}H]-labelled amino acids and, where indicated, cold amino acids. The tubes were flicked every 5 min. The samples were subsequently washed 3 times after binding with wash buffer (lysis buffer with 300 mM additional NaCl). The resin or gel was aspirated dry and resuspended in 100 μl of wash buffer. The samples were then thoroughly mixed with a cut tip, and 10 μl of each sample was loaded into scintillation fluid in triplicate and quantified with a TriCarb scintillation counter.

Thermal shift assay

The thermal shift (protein melting) assays were performed according to the Bio-Rad CFX Manager instruction manual. In brief, for HDAC6, a combination of 5× Sypro orange dye and HDAC6 at a concentration of 10 μM were combined with or without valine or isoleucine, leucine and arginine (at the indicated concentrations) in thermal shift buffer (100 mM HEPES pH 7.4, 3.4 mM EDTA, 150 mM NaCl and 0.005% Tween-20) to a final volume of up to 10 μl in a well of a genebrick 8-tube strip. Each condition was tested in triplicate. The plate was subjected to a protocol in which the temperature increased from 25 °C to 95 °C at 0.1 °C s^{-1}. Fluorescence was recorded and plotted over time, and melting temperatures were calculated as described in the CFX Manager instruction manual. In brief, the negative first derivative of the curve shown (change in fluorescence/change in temperature) was plotted against the temperature. The peak (that is, lowest point on this curve) reflects the melting temperature. Each reported melting temperature represents the mean ± s.d. for three replicates from one experiment.

Generation of CRISPR–Cas9 genetically modified cells

For HDAC6, TET2 knockout HCT116 cell lines, sgRNAs targeting each gene were cloned into the pX330 vector. Guide RNA sequences were designed using the CRISPR design tool developed by the Zhang laboratory. The oligonucleotide sequences are provided in Supplementary Table 8.

To prepare lentivirus for the knockout experiments, On day one, 2×10^6 HEK293T cells were seeded into 6 wells of a 6-well plate. At 12 h after seeding, each well was transfected with 200 ng of PMD_2G, 800 ng of the pX458 guide construct and 400 ng of PSPAX2 using Lipofectamine 2000. Medium containing the virus was collected 48 h after transfection. The HCT116 cells were cultured in the collected viral supernatant in the presence of polybrene (8 μg ml^{-1}). The following day, cells were trypsinized, pooled in a 10 cm dish and selected with puromycin to eliminate untransfected cells. At 48 h after selection, the medium was aspirated and replenished with fresh medium lacking puromycin. The following day, cells were single-cell sorted with a flow cytometer into the wells of a 96-well plate containing 200 μl of DMEM supplemented with 20% FBS. Cells were grown for 2 weeks and the resultant colonies were trypsinized and expanded. Clones were validated for loss of the relevant protein by immunoblotting.

RNA extraction and quantitative RT–PCR

Total RNA was extracted using TRIzol reagent (Vazyme R401-01). Total RNAs (0.5–1 μg) were subjected to reverse transcription with HiScript II Q Select RT SuperMix (Vazyme, R232-01). To determine relative mRNA level, qPCR was performed using universal SYBR qPCR master mix (Vazyme, Q711-02) and gene expression was normalized to that of actin. Primers used for qPCR are listed in Supplementary Table 8.

Immunofluorescence assays

Cells were seeded on fibronectin-coated glass coverslips in 24-well tissue culture plates. After transfection or infection, the cells were rinsed once with PBS and fixed in 4% paraformaldehyde for 15 min at room temperature. The fixed cells were permeabilized using 0.1% Triton X-100 and rinsed twice with PBS. The coverslips were blocked with blocking buffer (0.3% BSA in PBS) for 1 h and incubated overnight at 4 °C with a primary antibody in blocking buffer. Next, the coverslips were rinsed twice with blocking buffer and incubated with secondary antibodies for 1 h at room temperature in the dark. The glass coverslips were mounted using anti-fade medium containing DAPI and imaged using a Zeiss LSM 510 Meta confocal system.

The tissue specimens were fixed overnight in 4% neutral-buffered formalin and then were dehydrated in increasing concentrations of isopropyl alcohol, followed by clearing of alcohol by xylene. The specimens were subsequently embedded in paraffin wax in cassettes for facilitation of tissue sectioning. For immunohistochemistry, tissue sections were deparaffinized and incubated in citrate buffer at 95 °C for 40 min for antigen retrieval and then incubated overnight at 4 °C with the following primary antibodies: anti-HDAC6 (1:100 dilution), anti-γH2AX (1:100 dilution), anti-PAR (1:100 dilution) and anti-5hmC (1:1,000 dilution). After three washing steps, staining was visualized with secondary antibodies and the images were captured with a digital camera under a confocal microscope. For quantification of the

percentage of γH2AX foci$^+$ cells in each cell line, at least ten microscopy fields were counted in each independent experiment and at least three independent. For quantification of the percentage of γH2AX foci$^+$ cells in tumours, we included at least three noncontinuous sections from each tumour formed in nude mice that had been injected stable overexpressed or knocked out HCT116 cells. At least 3 microscopy fields with more than 300 cells for each tumour were calculated for each section. For statistics on the proportion of γH2AX foci$^+$ cells, we calculated it as (γH2AX foci$^+$ cells/nucleic) × 100.

Cell fractionation assay

About 2×10^6 cells were rinsed with ice-cold PBS and then added buffer A (10 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT, 0.05% NP40 and protease inhibitor cocktail). After 15 min on ice, the cells were centrifuged at 3,000 r.p.m. and 4 °C for 10 min. The resultant supernatant was then used as a cytosolic fraction. The pellet was then homogenized with buffer B (20 mM HEPES pH 7.9, 0.4 mM NaCl, 1 mM EDTA, 1 mM DTT and 1 mM PMSF). After vigorous vortexing for 10 min, the homogenates were centrifuged for 20 min at 13,000 r.p.m. and 4 °C. The resultant supernatants were then used as nuclear fractions. Insoluble chromatin was centrifuged for 5 min at 3,000 r.p.m. and 4 °C, washed once in buffer B and centrifuged again under the same conditions. The nuclear and cytosolic fractions were quantified using a BCA Protein Quantification kit (Yeasen Biotechnology), and 20 μg of each fraction was analysed by SDS–PAGE. H3 and α-tubulin were used as markers for nuclear and cytoplasm, respectively.

Deacetylase activity of HDAC6 using direct detection assay

HDAC6 activity was measured using an Epigenase HDAC Activity/Inhibition Direct Assay kit (Colorimetric) (EpiGentek, P-4034). For detection of HDAC6 and deletion mutant activities, HCT116 cells were transfected with pCDNA3.1-Flag–HDAC6 and the deletion mutants. After transfection for 24 h, the cells were lysed and immunoprecipitated using Flag M2 affinity gel as previously described[47]. The HDAC6 and deletion mutants were eluted with Flag peptide (100 μg ml^{-1}). The immunoprecipitated proteins (1 μg) were assayed for HDAC activity using equal volumes in HDAC assay buffer in duplicate. The plate with HDAC6 and the deletion mutants was incubated at 37 °C for 90 min. The plate was washed and incubated with capture and detection antibodies and developed following the protocol provided by the manufacturer. The activity of HDAC6 (OD min^{-1} mg^{-1} protein) was calculated according to manufacturer's specifications based on duplicate measurements of the optical densities at 450 nm.

Cell cycle arrest and analysis of cell cycle by flow cytometry

About 5×10^5 WT, HDAC6 knockdown, TET2 knockdown and TDG knockdown HCT116 cells were seeded in 12-well plates and treated with palbociclib, a CDK4/6 inhibitor[34]. At 12 h after plating, cells were subjected to valine restriction for 12 h and digested cells with trypsin. The cells were fixed with ice-cold 100% methanol and then incubated at 20 °C for 20 min. Cells were centrifuged at 1,300–1,500 r.p.m. for 5 min. The supernatant was discarded and 1 ml ribonuclease (RNase, 100 μg ml^{-1}) buffer was added to each tube and the cells were resuspended gently by pipetting up and down. The cells were incubated on ice for 30 min, and then centrifuged at 1,300–1,500 r.p.m. for 5 min. The supernatant was discarded and 1 ml PI solution (50 μg ml^{-1}) was added to each tube and the cells resuspended gently by pipetting up and down. The cells were incubated again on ice for 5–10 min, and then resuspended gently by pipetting up and down. The cell suspension was transferred to 5 ml polystyrene round-bottom FACS tubes and then the cell cycle was analysed by flow cytometry.

SILAC labelling of HCT116 cells

For SILAC labelling, cells were grown in SILAC DMEM medium supplemented 10% dialysed (USA) FBS and either with 100 mg ml^{-1} [$^{12}C_6{}^{14}N_2$]lysine, [$^{12}C_6{}^{14}N_4$]arginine (SILAC light) or [$^{13}C_6{}^{15}N_2$]lysine, [$^{13}C_6{}^{15}N_4$]arginine (Cambridge Isotope Laboratories) (SILAC heavy). Cells were collected after passaging 10 times. Flag–HDAC6 overexpressed cell lysate was prepared and peptides were enriched by Flag M2 affinity gel IP. Peptide fractions were analysed by online nanoflow LC–MS/MS.

Proteome sample preparation

About 5×10^7 HCT116 cells were plated in a 15 cm plate before the experiment. Cells were washed 3 times with cold PBS buffer and then lysed in SDT lysis buffer (4% w/v SDS, 100 mM Tris-HCl, 100 mM DTT, pH 7.6). The cell lysates were incubated for 10 min at 95 °C and centrifuged at 15,000*g* for 15 min. The supernatants were collected, and protein concentration was determined by tryptophan fluorescence emission assay as previously described[48]. A of 150 μg protein from each sample was digested by filter-aided sample preparation protocol (FASP) using 10 kDa centrifugal filter tubes as described previously with slight modifications[49]. Peptide mixtures were collected by centrifugation at 15,000*g* for 15 min and quantified by Nanodrop 2000 (Thermo Fisher Scientific). A total of 20 μg peptides from each sample was desalted for LC–MS/MS analysis.

LC–MS/MS analysis

Data-independent acquisition (DIA) MS analysis was performed on an Orbitrap Q-Exactive HF (Thermo Fisher Scientific) platform connected to an online nanoflow EASY nLC1000 HPLC system (Thermo Fisher Scientific). A total of 1 μg peptides was loaded onto a self-packed column (75 μm × 150 mm, 1.9 μm ReproSil-Pur C18 beads, 120 Å, Dr. Maisch) and separated with a 90 min gradient at a flow rate of 300 nl min^{-1} with following gradients: 0–1 min, 1–3% buffer B; 1–71 min, 3–23% buffer B; 71–81 min, 23–32% B; 81–84 min, 32–100% B; 84–90 min, 100% B.

DIA was performed using 120 K resolution MS scan and then followed by 40 MS/MS scans with 30 K resolution. The MS AGC target value was set at 3×10^6 with 100 ms of max injection time by orbitrap mass analyzer (300–1,650 *m/z*), and MS/MS AGC target value was set at 5×10^5 with 45 ms of max injection time. For proteome DIA MS runs, fragment analysis was subdivided into 40 DIA isolation windows. A detailed isolation windows list is provided in Supplementary Table 9.

Database searching and data analysis

Raw data were processed using DIA-NN software (v.1.8)[50] against the UniProt human proteome database, with precursor false discovery rate at 1%. Carbamidomethyl was selected as fixed modifications and N-terminal methionine excision enabled. The peptide length range was 7–30 and one missing cleavage site was allowed. The log level was set as 5. The quantification strategy was Any LC (high accuracy).

Genomic DNA extraction from cultured cells and dot blot assay

Cultured cells were collected and resuspended in digestion buffer (100 mM Tris pH 8.0, 100 mM NaCl, 25 mM EDTA and 0.5% SDS) at 50 °C for 12–18 h. Equal volumes of phenol–chloroform were added to the lysate and vortexed. The upper aqueous phase was separated by centrifugation and transferred to new tubes. A 1/10 volume of sodium acetate and 2 volumes of ethanol were added to let the genomic DNA precipitate. DNA was washed using 70% ethanol and dissolved in TE buffer (10 mM Tris (pH 8.0) and 1 mM EDTA). RNase A was then added to the DNA solution and incubated at 37 °C overnight, after which the phenol–chloroform extraction was repeated. For dot blot assay, 1 μg DNA was first sonicated and levelled to 1,000 ng per 60 μl. Levelled DNA was then loaded into a 96-well plate followed by 2 additional serial dilutions (final volume was 30 μl per well). Next, 20 μl of buffer (1 M NaOH and 25 mM EDTA) was added and samples were incubated at 95 °C for 10 min. Then 50 μl of ammonium acetate (2 M, pH 7.0) was added and cooled on ice for another 10 min. The corresponding amount of DNA dotted on the nitrocellulose membrane. The membrane was then

crosslinked by UV light for 20 min and subjected to the western blot protocol for the remaining steps.

Immunofluorescence staining of 5hmC

For immunofluorescence staining of 5hmC, HCT116 cells were washed with iced PBS before fixation in 4% fresh paraformaldehyde in PBS for 15 min. The cells were treated with 2 M HCl at room temperature for 30 min, followed by neutralizing with 100 mM Tris (pH 8.0) at room temperature for 30 min. Subsequently, the cells were blocked with 5% BSA in PBS for 1 h in a 37 °C incubator. The incubation with 5hmC primary antibody was carried out at 4 °C overnight, and secondary antibody incubation was performed at room temperature for 1 h. DNA was stained with DAPI for 10 min. Images were acquired using a Zeiss LSM 510 Meta confocal system.

5mC, 5hmC and 5fC quantification by UPLC–MS/MS analysis

For quantitative determination of the content of 5mC and 5hmC nucleosides, genomic DNA was first digested using nuclease P1 (NEB) at 37 °C for at least 2 h and then was dephosphorylated with calf intestinal alkaline phosphatase (CIAP, Takara) at 37 °C for an additional 1 h. The samples were centrifuged and the supernatants were collected before being subjected to multiple reaction monitoring (MRM) mode-based UPLC–MS/MS analysis. The MS analyses were performed using a UPLC system (Acquity UPLC I-Class, Waters) coupled to a triple quadrupole mass spectrometer (Triple Quad 6500 + LC–MS/MS, SCIEX). An Acquity UPLC BEH C18 column (130 Å, 1.7 µm, 2.1 mm × 50 mm, Waters) was used for the liquid chromatography separation. The mobile phases (A, 0.1% formic acid in pure water; B, 100% methanol) were used for compound separation. The linear gradient elutes from 1% B (0–1 min), 1–5% B (1–2.5 min), 5–95% B (2.5–3 min), 95–1% B (4–4.1 min), then held at 1% B until 6 min for re-equilibrium. The flow rate was set to 0.2 ml min^{-1}. Optimized MRM transition parameters for each of the nucleosides were obtained using pure compound standards. The quantifier transitions of each nucleoside are 5mC: 242.1/126.1 (CE 17, DP 20); 5hmC: 258.1.1/142.1.1 (CE 15, DP 40); dC: 228.1/112.1 (CE 20, DP 20); dG: 268.1/152.1 (CE 20, DP 60). All compounds were measured on positive ESI mode. Then the retention time for each compound was individually determined by measuring the corresponding MRM transitions on the BEH C18 column. For 5mC: 1.67 min; 5hmC: 1.08 min; dC: 1.08 min; dG: 2.89 min. The amount of each nucleoside was calculated according to the peak areas of quantifier MRM transitions by interpolation from the standard curves.

In vitro deacetylation of TET2 and TET2 enzymatic activity

To prepare the TET2 enzyme, MYC-tagged TET2 was ectopically expressed in HEK293T cells and immunoprecipitated in RIPA lysis buffer (50 mM Tris-Cl, pH 7.4, 0.5% NP-40, 150 mM NaCl, 1 mM EDTA, 10% glycerophosphate and a cocktail of proteinase inhibitors) using anti-MYC affinity beads. Immunoprecipitated and purified MYC-tagged TET2 proteins (immobilized on beads) were incubated with or without recombinant GST–HDAC6 (2, 4, 8, 10 µg) with reaction buffer (60 µM NAD^+, 25 µM Tris-HCl (pH 8.0), 137 mM NaCl, 2.7 mM KCl and 1 mM $MgCl_2$) at 30 °C for 30 min. In vitro assay of TET2 catalytic activity was performed following the procedures previously described using two different substrates: methylated DNA oligonucleotides and total genomic DNA[51]. In brief, to prepare the substrates, two complementary single-stranded oligonucleotides were synthesized at Beijing Tsingke Biotech (GTATGCCTCATGC^{m}CGGACTTAACTGCAGTG-3′, 3′-CATACGGAGTACG^{m}C.

CTGAATTGACGTCAC-5′, where the superscripted 'm' indicates the methylation of following cytosine). Next, 2 nmol of single-stranded oligonucleotides were mixed in annealing buffer (10 mM Tris, pH 8.0, 50 mM NaCl and 1 mM EDTA), boiled for 5 min, transferred to preheated water at 90 °C, gradually cooled to room temperature, and used as substrate. Genomic DNA from HCT116 cells was isolated, sonicated and then used as substrate. MYC-tagged TET2 beads were washed and incubated with methylated DNA oligonucleotides (20 pmol) or total genomic DNA (150 ng) in buffer containing 50 mM HEPES (pH 8.0), 100 mM NaCl, 100 µM Fe $(NH4)_2(SO4)_2$, 2 mM ascorbate, 1 mM DTT and 1 mM ATP at 37 °C for 1 h. DNA was purified, and the levels of 5hmC or 5mC were quantified by dot blot and ELISA using 5hmC or 5mC antibody. MYC-tagged TET2 beads were washed, boiled and subjected to SDS–PAGE analysis.

Comet assay

HCT116, HDAC6-WT or SE14 knock-in MEF cells were seeded and treated with VR at different concentrations and times, and then cells were collected and resuspended in cold PBS at a concentration of 3×10^5 cells per ml. The comet assay then was performed according to the manufacturer's instructions (4250-050-K; Trevigen). For the comet assay, we repeated the experiment three times, with two technical replicates per trial. Furthermore, we calculated the average of the two technical replicates for each biological replicate and derived the statistics from the three biological replicate. DNA damage was measured in terms of tail moments using cometscore software (casplab_1.2.3b2).

Oncosphere formation assay

Oncosphere assays were performed as previously described[47]. Oncospheres were enriched from HCT116 cells. Single-cell suspension of HCT116 cells (200 cells per well, repeat 10 holes in each group) were plated on 96-well ultralow attachment plates (Corning, 3474) and cultured in DMEM/F12 (Gibco) supplemented with 5 mg ml^{-1} insulin (Sigma), 20 ng ml^{-1} EGF (Sigma), 1:50 B27 (Gibco), 10 ng ml^{-1} bFGF and 0.4% BSA for 10 days alone. Floating spheres that grew over 2 weeks were counted. Tumour spheres were visualized under a phase-contrast microscope, photographed, quantified and represented graphically. Tumour spheres larger than 100 µm were counted, and the 'relative per cent of sphere' represents the relative increase in the number of spheres per dish.

Electrochemical detection of the binding of peptides and valine

3-mercaptopropionic acid (MPA), *N*-hydroxy-succinimide (NHS), *N*-(3-dimethyl aminopropyl)-*N*′-ethylcarbodiimide hydrochloride (EDC), morpholino ethanesulfonic acid (MES), $K_3Fe(CN)_6$ and $K_4Fe(CN)_6$ were purchased from Sigma-Aldrich and used without further purification. Chloroauric acid hydrate ($HAuCl_4 \cdot xH_2O$) was purchased from Adamas. PBS (0.01 M PBS, pH 7.4) was prepared from NaCl, KCl, Na_2HPO_4 and KH_2PO_4. A conventional three-electrode system was used with a prepared Au electrode (CH Instruments; diameter: 2 mm), an Ag/AgCl and a platinum plate as the working, reference and counter electrodes, respectively. The 20 mM Fe $(CN)_6^{3-/4-}$ in 0.01 M pH 7.4 PBS was used as the electrolyte unless otherwise indicated. The cyclic voltammetry (CV) test was operated from 0 to 0.5 V at a scan rate of 10 mV s^{-1}. For the electrochemical impedance spectroscopy (EIS) measurement, a sine-wave of 10 mV amplitude was applied to the as-prepared electrode in the frequency range of 0.1 Hz to 100 kHz. The ZSimpWin EIS DATA analysis software (Perkin-Elmer, v.2.00) was used to analyse the obtained EIS data and to fit an equivalent circuit.

For the preparation of the biosensor, a gold disc electrode (2 mm in diameter) was polished according to the literature[52,53], and then rinsed with distilled water before further modification. Afterwards, the electropolished electrodes were dried and immediately immersed in an aqueous solution containing 0.1 M Na_2SO_4 and 30 mM $HAuCl_4 \cdot xH_2O$ for the electrodeposition of gold dendrite (AuD) by chronoamperometry (CA) at a potential of −0.55 V for 600 s (ref. 54). The 25 mM cysteamine hydrochloride was then self-assembled on the AuD through the formation of a gold–sulfur bond[55]. After the activation of the C-terminated group of cysteamine in the MES buffer having EDC and NHS, the 1 mg ml^{-1} NHS-biotin, 0.1 mg ml^{-1} streptavidin and biotin–peptide were successively immobilized on the electrode to form the

AuD/cysteamine–biotin–streptavidin/peptide biosensor. Finally, the biosensor was soaked in 5% BSA blocking buffer for 15 min and ready for the L-valine detection. Four types of peptides were designed in the experiment and marked as biotin–TLAQT ISEAA IGGA (1st), biotin–MLGQT TSEEA VGGA (2nd), biotin–ILDQT TSEDA VGGA (3rd) or biotin–RVTIM PKDIQ LAR (control).

RNA sequencing and data analysis

RNAs were extracted from HCT116 cells using RNA isolater Total RNA extraction reagent (Vazyme, R401-01). RNA sequencing libraries were prepared from 2 μg of RNA per sample as input material for the RNA sample preparations. Sequencing libraries were generated by Annoroad Gene Technology. The sequencing was conducted on an Illumina NovaSeq 6000.

For RNA sequencing data, the quality of raw reads was assessed using FASTQC (v.0.11.8) and the adapters were removed with TrimGalore (v.0.6.4). The resulting clean reads were mapped to the hg38 reference genome using HISAT2 (v.2.1.0). Gene expression levels were quantified to fragments per kilobase of exon model per million mapped fragments (FPKM) by StringTie (v.2.2.1) and counts by prepDE.py based on hg38 annotations. To perform differential expression analysis, the R package DEseq2 (v.1.26.0) was utilized with raw counts from prepDE.py. An adjusted *P* value cut-off of 0.01 and an absolute fold change cut-off of 2 were applied to identify significantly changed genes.

TET2 ChIP–seq assay

HCT116 cells were crosslinked with 1% formaldehyde for 10 min at room temperature and quenched with 125 mM glycine. The fragmented chromatin fragments were pre-cleared and then immunoprecipitated with protein A+G magnetic beads coupled with anti-TET2 antibodies. After reverse crosslinking, ChIPed and input DNA fragments were end-repaired and A-tailed using the NEBNext End Repair/dA-Tailing module (E7442, NEB) followed by adapter ligation with the NEBNext Ultra Ligation module (E7445, NEB). Sequencing libraries were generated by Genefund Biotech. The DNA libraries were amplified for 15 cycles and sequenced using Illumina Hi-Seq with 150 bp paired-end protocol.

ChIP–seq analysis

The ChIP–seq data were checked with FASTQC (v.0.11.8) and trimmed using TrimGalore (v.0.6.4). The resulting clean reads were aligned to hg38 using Bowtie2 (v.2.3.5.1) with the following parameters: --no-unal --no-mixed --no-discordant, and were further filtered with mapping quality of at least 20. PCR duplicates were removed using Sambamba (v.0.7.0). The signal tracks were generated by bamCoverage in deepTools (v.3.3.0) with 25 bp bin size and normalized as RPKM.

For downstream analysis of TET2 ChIP–seq data, we used -p 0.001 for MACS2 to call peaks. As TET2 showed different binding patterns under different treatments, the binding sites were not directly comparable. To enable better comparison of TET2 binding sites among treatments, we first merged peaks from all treatments and the RPKM values were calculated for these regions. Then, we used the regions to perform principal component analysis (PCA) on all samples and retained the regions with eigenvalues of the rotation > 0.006, which contributed mostly for separation of the different treatments. These regions were further clustered based on the TET2-binding signal across the samples using *K*-means clustering. Gene annotations of regions were generated from the R package ChIPseeker (v.1.22.1).

For downstream analysis of H3K4me1 ChIP–seq, we used --broa--broad-cutoff 0.01 -q 0.01 for MACS2 to call peaks. The identified H3K4me1 peaks were defined as enhancer regions.

Functional enrichment analysis

GO analysis was performed by the Database for Annotation, Visualization and Integrated Discovery (DAVID) Bioinformatics Resource, and *P* values were plotted to show the significance.

Metabolomics

The mice were fed a VR diet or not for a corresponding time. Plasma samples were thawed at 4 °C on ice and swirled well for 30 s before sampling. Then, 15 μl of each sample was obtained and placed in an EP tube, 35 μl water added and extracted with 200 μl of extraction solvent (1:1 V methanol:V acetonitrile, internal standard mixture containing isotopes, pre-cooling at −40 °C), vortexed for 30 s, ultrasound treated for 15 min (while being cooled in ice water), and next chilled for 1 h at −40 °C to precipitate the proteins. The samples were then centrifuged at 12,000 r.p.m. for 15 min at 4 °C. The supernatant (100 μl) was rotated to evaporate until dry. Samples were redissolved with 100 μl 50% methanol water, 100 μl derivative added and 50 μl 1 M $NaHCO_3$ added and the mixture swirled well. Samples were incubated in a 40 °C water bath for 1 h, and then removed and cooled to room temperature. Next, 50 μl 2 M HCl was added, and the sample was rotated and evaporated to dry. Finally, 200 μl methanol was added to redissolve the sample, and the supernatant was obtained for the UHPLC-QTOF-MS analysis.

About 20 mg of tumour sample was taken and placed in a EP tube, and 2 small steel balls added, and then 1,000 μl extraction solvent was added containing an internal target (V methanol: V acetonitrile: V water = 2:2:1, containing internal standard, pre-cooled at −40 °C). Samples were homogenized in the ball mill for 4 min at 35 Hz, then ultrasound treated for 5 min (incubated in ice water). After homogenization for 3 times and incubation for 1 h at −40 °C to precipitate proteins, samples were centrifuged at 12,000 r.p.m. for 15 min at 4 °C. Supernatant (100 μl) was transferred into EP tubes. Extracts were rotated to evaporate until dry and redissolved with 100 μl 50% methanol. The extracts had 100 μl derivative and 50 μl 1 M $NaHCO_3$ added, and then swirled well. The extracts were incubated for 1 h in a 40 °C water bath, and then removed and cooled to room temperature. Next, 50 μl 2 M HCl was added, and the samples were rotated and evaporated to dry. Finally, 200 μl methanol was added to redissolve the sample, and the supernatant was obtained for UHPLC-QTOF-MS analysis. Data analysis was performed by Shanghai Biotree.

Whole-genome bisulfite sequencing library preparation

Preparation of the whole-genome bisulfite sequencing (WGBS) library was performed as previously described[56]. genomic DNA of each sample was purified using a Wizard Genomic DNA Purification kit (Promega). Next, 1 μg of genomic DNA was fragmented into 300 bp using a Covaris LE220R. The methylated adapters were ligated to the fragmented DNA. Bisulfite treatment was performed using an EZ DNA Methylation-Gold kit (Zymo Research) according to the instruction manual. The resulting single-stranded DNA was amplificated using KAPA HiFi HotStart Uracil+ ReadyMix (2×) (Roche). The concentration of purified libraries was measured with a Qubit fluorometer dsDNA HS Assay (Thermo Fisher Scientific). Size distribution of the sequencing libraries was analysed using Agilent 2100 Bioanalyzer. WGBS was performed using an Illumina NovaSeq 6000 platform for 150 bp paired-end sequencing.

APOBEC-coupled epigenetic sequencing library preparation

For the APOBEC-coupled epigenetic sequencing (ACE-seq) library construction, 200 ng of genomic DNA was fragmented into 300 bp using a Covaris M220. Then, the T4 phage β-glucosyltransferase (NEB) and uridine diphosphoglucose were used to convert 5hmC into 5ghmC at 37 °C for 2 h. After that, 0.1 M NaOH (5×) was added to the modified fragments for denaturation at 50 °C for 10 min. Then, DEA enzyme mix (Vazyme) was used to deaminate dC or 5mC to U or T, respectively, at 37 °C for 3 h. After deamination, DNA fragments were purified and subjected to library construction with a Scale Methyl-DNA Lib Prep kit according to the instruction manual (Illumina). In brief, the single-stranded heat-denatured DNA fragments were ligated to the dA-tailing module and T7 truncated adapters. Next, the ligated DNA was extended to the second strand synthesis and the methylated adapters were ligated.

The DNA was amplified for 10 cycles using Illumina 8-bp dual index primers. Finally, the ACE-seq libraries were analysed by Agilent 2100 Bioanalyzer and sequenced on an Illumina NovaSeq 6000 platform by 150 bp paired-end sequencing.

MAB-seq library preparation

We performed this assay as previously described[57,58]. In brief, 500 ng of DNA extracted from HCT116 cells with or without the treatment of valine deprivation was first spiked-in with unmethylated lambda DNA (1:500), and the DNA was fragmented to 600–700 bp. The broken DNA was treated by M.SssI (New England Biolabs, M0226M) in a 50 µl reaction. In the treatment, DNA was first incubated with 1.0 unit μl^{-1} M.SssI methylase (New England Biolabs, M0226M) for 4 h in 50 µl reaction (1 µl of 4 U μl^{-1} M.SssI, 5 µl of 1.6 mM SAM, 5 µl of 10× methyltransferase reaction buffer, and ddH_2O was added up to 50 µl). DNA was purified using DNA Clean & Concentrator kit (Zymo, D4013). M.SssI-treated DNA was then subjected to bisulfite conversion using an EZ DNA Methylation-Direct kit (Zymo, D5021). Bisulfite-treated DNA was purified and subjected to library construction using an EpiArt DNA Methylation Library kit (Vazyme, NE103) and VAHTS Multiplex Oligos Set 4 (Vazyme, N321) according to the instruction manual (Illumina). In brief, the single-stranded heat-denatured DNA fragments were ligated to the 3′ adapter ligation and 5′ adapter ligation. Next, the ligated DNA was extended to the second strand synthesis and the methylated adapters were ligated. The DNA was amplified for 10 cycles using Illumina 8-bp dual index primers. Finally, the MAB-seq libraries were analysed by Agilent 2100 Bioanalyzer and sequenced on an Illumina NovaSeq 6000 platform by 150 bp paired-end sequencing.

WGBS, ACE-seq and MAB-seq analysis

The raw data of WGBS, ACE-seq and MAB-seq were checked with FASTQC (v.0.11.8) and the adapters were removed with TrimGalore (v.0.6.4). The resulting clean reads were mapped to the hg38 using Bismark (v.0.22.1) in paired-end mode. Duplicates were removed using deduplicate_bismark.

For downstream analysis of WGBS, cytosine not converted by bisulfite at each CpG sites was extracted by MethPipe software package (v.3.4.3). Cytosine sites covered by at least three reads and shared between two replications were retained. The mean of $N_C/(N_C + N_T)$ percentage at each CG site in two replications was calculated as WGBS level at each CG site.

For downstream analysis of ACE-seq, 5hmC information at each CpG site was extracted by MethPipe software package (v.3.4.3), and the level of 5hmC at each CpG site equals to $N_C/(N_C + N_T)$. Cytosine sites covered by at least three reads were retained. Hyperhydroxymethylation regions were also identified by MethPipe software package (v.3.4.3) using the hmr command.

For downstream analysis of MAB-seq, 5fC information at each CpG site was extracted by MethPipe software package (v.3.4.3), and the level of 5fC at each CpG site equals to $N_T/(N_C + N_T)$. As the presence of some false positives in MAB-seq, cytosine sites with at least 5 reads coverage and a 5fC level > 20% were retained.

Deriving 5mC and 5hmC levels from WGBS data by combing ACE-seq data

Given that WGBS cannot distinguish between 5mC and 5hmC, we used the MLML tool[59] (v.5.0.1) to estimate 5mC and 5hmC levels by subtracting 5hmC from WGBS data at each CpG site. CpG sites with any negative level or conflict were excluded from the subsequent analysis.

Calculating 5mC, 5hmC, 5fC levels in genomic regions

To compare 5mC, 5hmC and 5fC levels in each given region, the mean of 5mC, 5hmC or 5fC level at each CpG site within the region was calculated.

END-seq and S1-END-seq library preparation

About 4×10^6 HCT116 cells were plated in a 6 cm plate before the experiment. Cells were subjected to valine deprivation for 24 h and treated with ddC (20 µM) for 20 h before collection and washed three times with cold PBS buffer. Cells were collected and gently pipetted 5–10 times and transferred to a conical 15 ml tube. Cells were neutralized with 10 ml of cold PBS/EDTA 5 mM/10% FBS buffer, then pelleted, resuspended in 1 ml of neutralization buffer and transferred to an Eppendorf tube. They were then processed for END-seq or S1-END-seq. For S1-END-seq, cells were collected and embedded in 1% agarose plugs, lysed and digested with proteinase K (1 h at 50 °C, followed by 7 h at 37 °C), washed with TE buffer and then treated with RNAse A for 1 h at 37 °C. Plugs were then washed in EB and equilibrated in S1 nuclease buffer (40 mM sodium acetate pH 5.2, 300 mM NaCl and 2 mM $ZnSO_4$) for 30 min. Next, 1.8 U of S1 nuclease was added to 100 µl of S1 nuclease buffer per plug and incubated on ice for 15 min to allow for the enzyme to diffuse into the plug. The reaction mix was then placed at 37 °C for 20 min before addition of EDTA (10 mM final concentration) to terminate the reaction. Finally, plugs were processed using the standard END-seq protocol[60].

END-seq and S1-END-seq analysis

For END-seq and S1-END-seq analysis, considering that only read 1 contains pertinent information about the broken site, the raw data of read 1 was checked with FASTQC (v.0.11.8) and trimmed using TrimGalore (v.0.6.4). The resulting clean reads aligned to hg38 using Bowtie2 (v.2.3.5.1) in single-end mode, and were further filtered with mapping quality of at least 20. Normalized RPM signals were generated by bed-tools (v.2.28.0) bamToBed and genomecov command.

Statistics and reproducibility

All the software and algorithms used are listed in Supplementary Table 10. Statistical analyses were conducted using GraphPad Prism 8.0 (GraphPad Software) and R (v.3.6.0; https://www.R-project.org/). For all experiments, data are presented as the mean ± s.d. unless otherwise specified. All of the experiments (except those described otherwise in the legend) were performed independently at least three times with a similar outcome. The statistical tests, *n* values and the *P* values are all indicated in the figures and/or legends. Depending on the type of experiment, the *P* values were calculated using two-tailed Student's *t*-test, one-way ANOVA, two-way ANOVA or Mann–Whitney *U*-test. $P < 0.05$ was considered to be significant ($^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$). The exact values of *P* values are shown in Supplementary Table 11. Mice were randomly assigned to each treatment group when the tumour reached a palpable minimum size. Investigators were blinded during image analysis of tumour burden.

Reporting summary

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

Data availability

All sequencing data generated within this study have been uploaded to the NCBI Sequence Read Archive (SRA) and are available under the following accession codes: RNA sequencing: GSE274535; WGBS: GSE274536; ACE-seq: GSE274539; MAB-seq: GSE275023; TET2 ChIP–seq: GSE274881; NLS-HDAC6 TET2 ChIP–seq: GSE274883; END-seq: GSE275061; ddC S1-END-seq: GSE275062; and H3K4me1 ChIP–seq: GSE275024. Additional data that support the findings of this study are available from the corresponding author upon reasonable request. Supplementary tables are provided with this paper. Source data are provided with this paper.

47. Wang, X. et al. AMPK promotes SPOP-mediated NANOG degradation to regulate prostate cancer cell stemness. *Dev. Cell* **48**, 345–360.e7 (2019).
48. Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N. & Mann, M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat. Methods* **11**, 319–324 (2014).

49. Wisniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–362 (2009).
50. Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat. Methods* **17**, 41–44 (2020).
51. Nakagawa, T. et al. $CRL4^{VprBP}$ E3 ligase promotes monoubiquitylation and chromatin binding of TET dioxygenases. *Mol. Cell* **57**, 247–260 (2015).
52. Yang, J. et al. Highly sensitive and selective determination of bisphenol-A using peptide-modified gold electrode. *Biosens. Bioelectron.* **61**, 38–44 (2014).
53. Su, W. Q., Cho, M., Nam, J. D., Choe, W. S. & Lee, Y. Highly sensitive electrochemical lead ion sensor harnessing peptide probe molecules on porous gold electrodes. *Biosens. Bioelectron.* **48**, 263–269 (2013).
54. Qin, J., Kim, S., Cho, M. & Lee, Y. Hierarchical and ultra-sensitive amyloid beta oligomer sensor for practical applications. *Chem. Eng. J.* **401**, 126055 (2020).
55. Qin, J., Jo, D. G., Cho, M. & Lee, Y. Monitoring of early diagnosis of Alzheimer's disease using the cellular prion protein and poly(pyrrole-2-carboxylic acid) modified electrode. *Biosens. Bioelectron.* **113**, 82–87 (2018).
56. Fang, S. et al. Tet inactivation disrupts YY1 binding and long-range chromatin interactions during embryonic heart development. *Nat. Commun.* **10**, 4297 (2019).
57. Wu, H., Wu, X., Shen, L. & Zhang, Y. Single-base resolution analysis of active DNA demethylation using methylase-assisted bisulfite sequencing. *Nat. Biotechnol.* **32**, 1231–1240 (2014).
58. Wu, H., Wu, X. & Zhang, Y. Base-resolution profiling of active DNA demethylation using MAB-seq and caMAB-seq. *Nat. Protoc.* **11**, 1081–1100 (2016).
59. Qu, J., Zhou, M., Song, Q., Hong, E. E. & Smith, A. D. MLML: consistent simultaneous estimates of DNA methylation and hydroxymethylation. *Bioinformatics* **29**, 2645–2646 (2013).
60. Canela, A. et al. DNA breaks and end resection measured genome-wide by end sequencing. *Mol. Cell* **63**, 898–911 (2016).

Acknowledgements We thank X. Jiang, G. Xu, G. Pei, A. Nussenzweig and W. Yang for pre-submission review and insightful comments; X. Zhou for help in equilibrium binding assays; J. Gao and H. Zhou for providing SILAC labelling technology and help with MS analyses; Y. Xu for providing TET2 plasmids; B. Ge for providing importin-α and importin-β plasmids; Z. Shan for help with PDX models of colorectal cancer; X. Zhang for providing *Rhinopithecus bieti* fibroblast cells; J. Yue for helping in multiple alignment of the HDAC6 protein sequence across species; and staff members of the Large-scale Protein Preparation System at the National Facility for Protein Science in Shanghai (NFPS), Shanghai Advanced Research Institute, Chinese Academy of Sciences, China, for providing technical support and assistance in data collection and analyses. This work was supported by the National Key Research and Development Program of China (2022YFC3401500, 2020YFA0803201, 2023YFA1800403, 2022YFA1103203 and 2022YFA1103900), the National Natural Science Foundation of China (31920103007, 82341028, 82073155, 32270858, 82173168, 82003150, 32270639 and 32370575), the Shanghai Rising-Star Program (23QA1407500 and 21QA1407500), the Shanghai Municipal Science and Technology Major Project (Genome Tagging Project), Fundamental Research Funds for the Central Universities (22120220616), and the China National Postdoctoral Program for Innovative Talents (BX20240258).

Author contributions J.J. and P.W. designed and conceived the studies. J.J., T.M., Y.Y. and S.W. performed most of the experiments with help from C.-Z.J., W.W. and P.W. Y.Y., R.H., W.W. and T.M. performed the bioinformatics analyses. J.J., Y.-X.L., C.-C.J. and T.M. generated the PDX model. J.J., Y.-Y.Z. and T.M. carried out immunofluorescence analyses. J.J., C.-C.J. and Y.Z. contributed to animal assays. J.J. and T.M. performed MS analyses. J.J. and S.W carried out END-seq and S1-END-seq library preparation. J.J., X.L., Y.-F.L. and Z.W performed WGBS, ACE-seq and MAB-seq library preparation with help from J.X. J.J. and Z.W carried out 5mC and 5hmC quantification by UPLC–MS/MS analysis. J.Q. and X.T. performed the electrochemical detection. H.C. contributed to the analysis of the structure of the complex of HDAC6 and valine. Y.C. synthesized the biotin–valine probe. J.J., S.S. and C.-C.J. performed the isotope labelling experiment. J.J., T.M., Y.Y., X.G., C.J., J.Y., J.X., W.W., D.W., C.-Z.J. and P.W. analysed the data and wrote the manuscript.

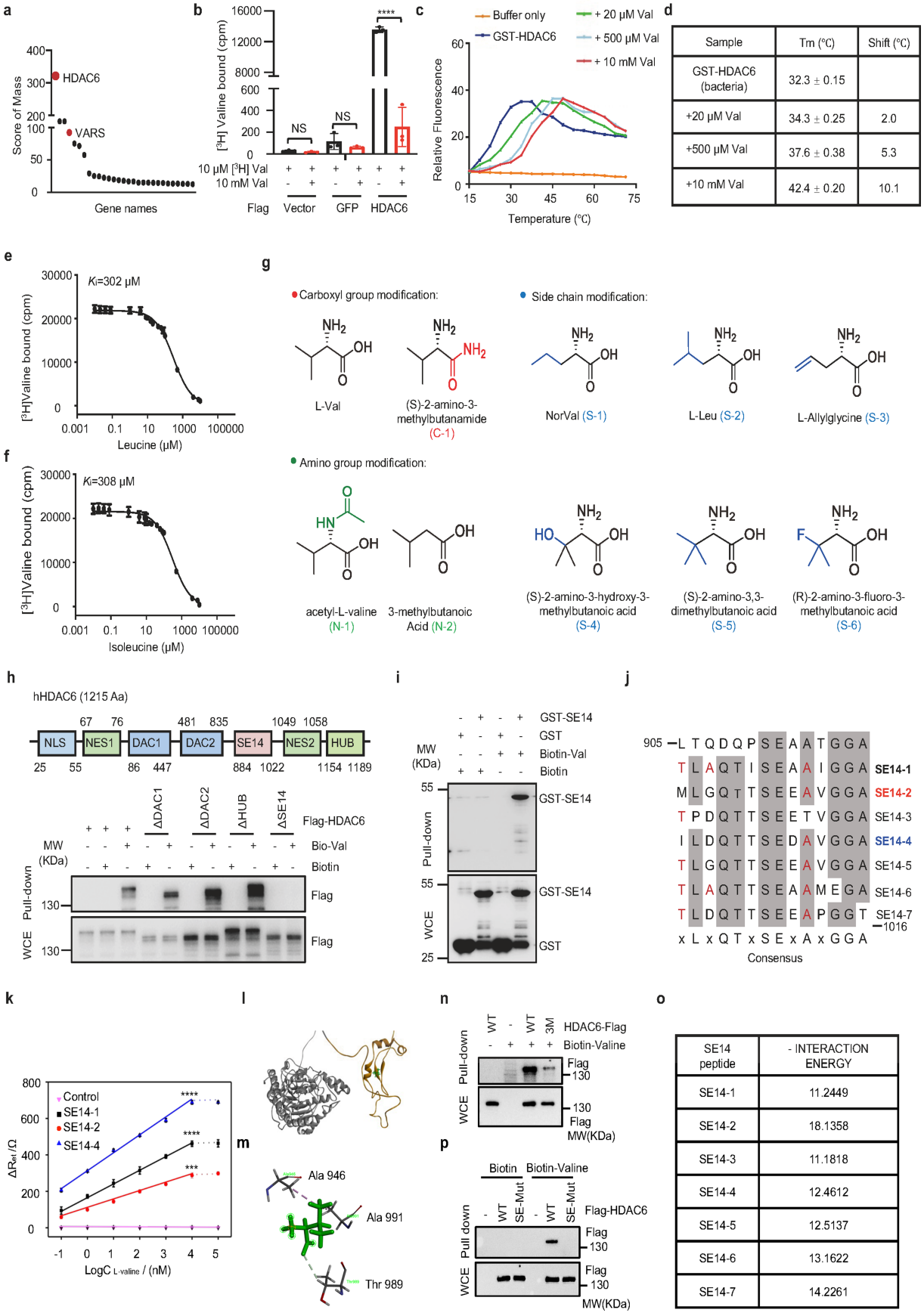
Competing interests The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-024-08248-5.

Correspondence and requests for materials should be addressed to Ping Wang.

Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

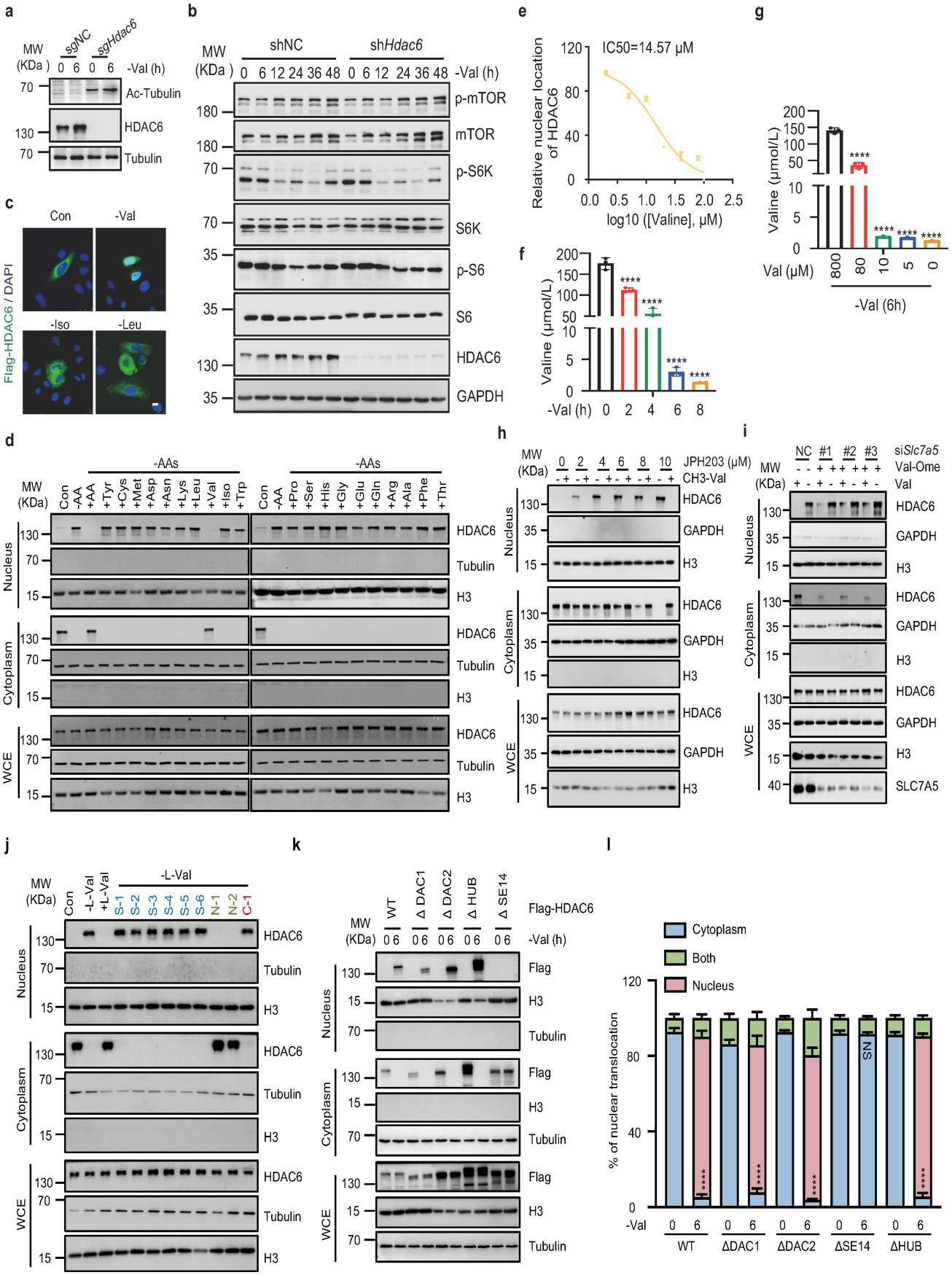


Sample	Tm (℃)	Shift (℃)
GST-HDAC6 (bacteria)	32.3 ± 0.15	
+20 μM Val	34.3 ± 0.25	2.0
+500 μM Val	37.6 ± 0.38	5.3
+10 mM Val	42.4 ± 0.20	10.1

SE14 peptide	- INTERACTION ENERGY
SE14-1	11.2449
SE14-2	18.1358
SE14-3	11.1818
SE14-4	12.4612
SE14-5	12.5137
SE14-6	13.1622
SE14-7	14.2261

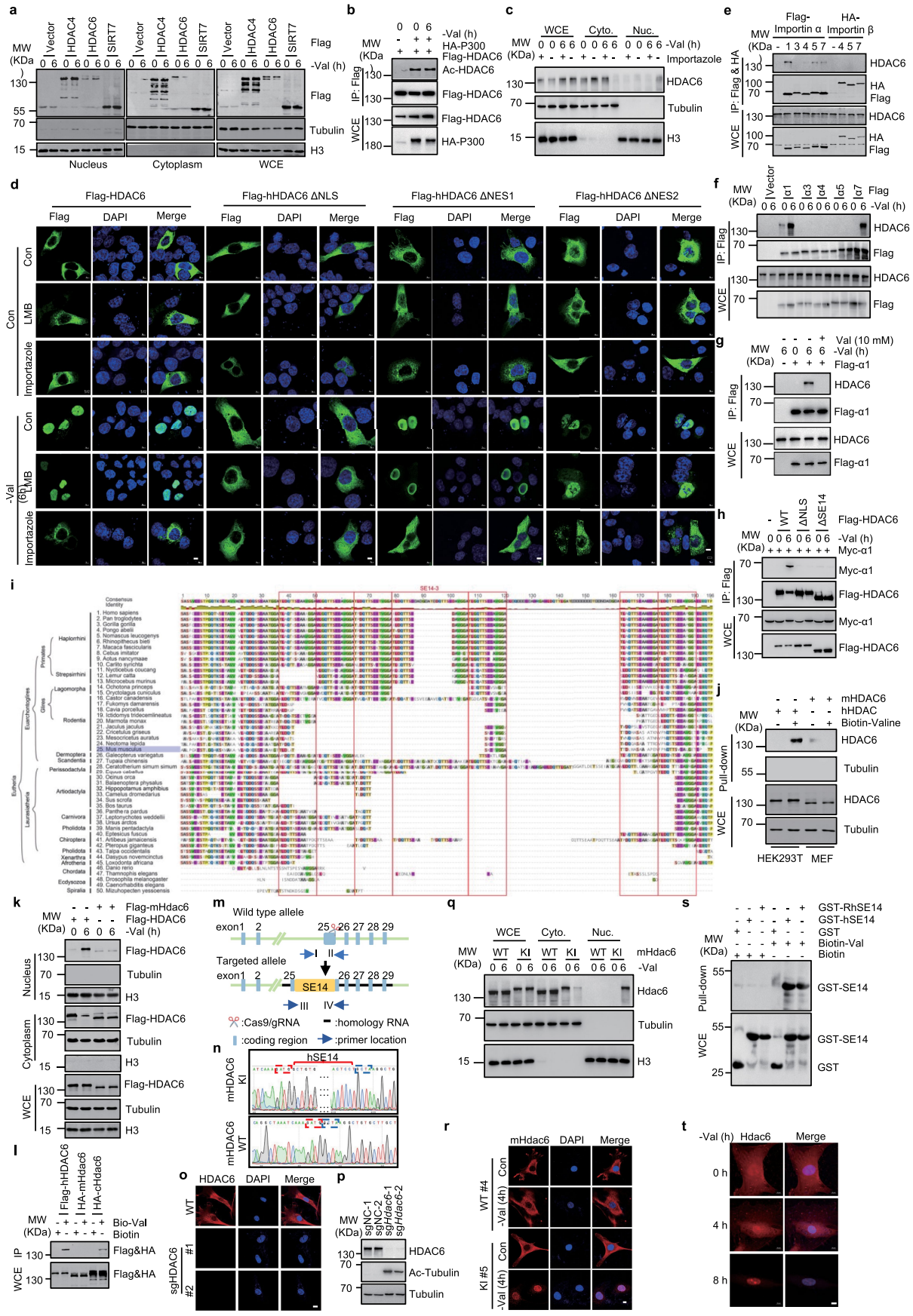
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Valine binds to HDAC6 via the SE14 domain. a, Score of candidate valine binding proteins identified by the mass spectrometry. **b**, Flag immunoprecipitates prepared from HCT116 cell extracts were used in binding assays with [^{3}H] valine. **c**, **d**, Effects of valine on the melting temperature of bacterially produced HDAC6 in a thermal shift assay. GST-HDAC6 produced from *E. coli* was incubated with Sypro Orange dye, with or without valine. **e**, **f**, Binding of [^{3}H] valine to Flag-HDAC6 prepared from HCT116 cells extracts was determined in the presence of unlabeled leucine and isoleucine, respectively. **g**, Structure of valine and its analogue with modifications in the amino terminus, carboxyl terminus or side chain. **h**, Schematic representations of various HDAC6 deletion mutants and immunoblot analysis of HDAC6 and various truncated proteins pulled down by biotin-labelled valine from HCT116 cells expressing Flag-HDAC6 or its various truncations. **i**, The interaction between biotin-valine and GST-SE14 repeat domain of HDAC6 prepared from *E. coli*. **j**, Schematic representation of the sequence for the SE14 repeat domain of HDAC6. **k**, Linear relationship with the sensitivities of between ΔRet value and logarithmic L-valine concentration of Biotin-TLAQT ISEAA IGGA (1st), Biotin-MLGQT TSEEA VGGA (2nd), Biotin-ILDQTTSEDAVGGA (4th), Biotin-RVTIM PKDIQ LAR (control). **l**, Homology modelling for HDAC6. Orange represents the SE14 repeat domain, grey represents other structural parts, and green molecule represent valine. **m**, Close up view of valine bonded to the surrounding core residues Ala946, Thr989, and Ala991. Valine is shown in green stick, and the carbon, nitrogen, oxygen, and hydrogen in the three amino acid residues are shown in gray, blue, red, and white line mode, respectively, with the interaction bonds shown in dashed lines. **n**, Immunoblot analysis of HDAC6 and point mutant (3 M: A946D, T989A and A991D) pulled down by biotin-labelled valine from HCT116 cells expressing Flag-HDAC6 or its point mutant. **o**, Calculated interaction energies of the different peptides in SE14 with valine, respectively. **p**, Immunoblot analysis of HDAC6 and point mutant (SE-Mut: mutate all the sites marked in red in **j**, mutate each threonine to alanine, alanine itself was mutated to aspartic acid) pulled down by biotin-labelled valine from HCT116 cells expressing Flag-HDAC6 or its point mutant. For **b**–**f** and **k**, data are representative of three independent experiments and presented as mean ± s.d. (n = 3 independent experiments). Statistical analysis for **b** was performed using one-way ANOVA, for **k** was performed using two-way ANOVA; $^{****}P < 0.001$, $^{****}P < 0.0001$, NS, not significant. For gel source data, see Supplementary Fig. 1.



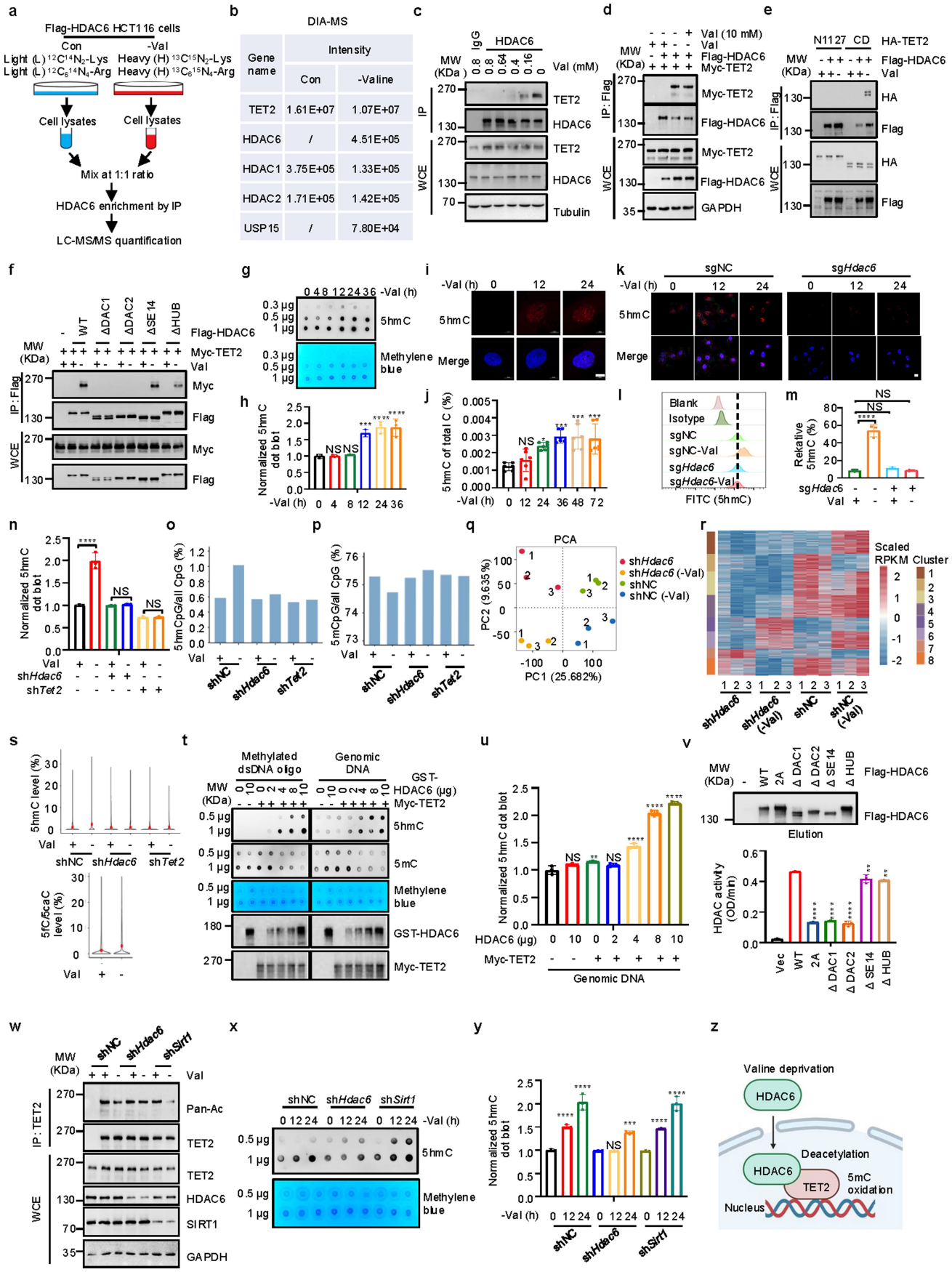
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Intracellular valine abundancy dictates subcellular distribution of HDAC6. a, The negligible effect on the acetylation of tubulin upon valine deprivation for 6 hrs. **b**, Knockdown of HDAC6 in HCT116 cells has little effect on the mTOR signalling with or without valine deprivation. **c**, Effects of BCAA (valine, leucine and isoleucine) on the subcellular localization of HDAC6. HCT116 cells were deprived of valine, leucine, or isoleucine for 6 h followed by immunofluorescence analysis of cellular localization of HDAC6. Scale bar, 10 μm. **d**, Cell fractionation analysis was performed using HCT116 cells deprived of amino acids for 6 h, and then add an indicated amino acid (200 mM) for another 6 h. **e**, The half-maximal concentration of valine at the cellular level to restrain HDAC6 in cytosol of Fig. 2d. **f**, **g**, Absolute quantification of valine in HCT116 cells upon valine deprivation for different times (**f**) or valine restriction at different concentration (**g**). **h**, The localization of HDAC6 was examined via cell fractionation assay after treatment with concentration gradient of SLC7A5 inhibitor JPH203 (0, 2, 4, 6, 8, 10 μM) for 24 h and then rescued with N-Methyl-L-valine (CH3-Val, 0.8 mM) for 6 h. **i**, Effects of BCAA transporter SLC7A5 on the subcellular localization of HDAC6. Cell fractionation assay was performed with SLC7A5 wildtype or knockdown HCT116 cells upon valine deprivation for 6 h and then rescued with L-valine-OMe (Val-Ome, 0.8 mM) for 6 h. **j**, Effects of valine and its analogues on cellular localization of HDAC6 under the condition of valine deprivation. Cell fractionation analysis was performed in HCT116 cells deprived of valine for 6 hrs before valine or one of its analogues (10 mM) were added for 6 hrs. **k**, Effects of valine deprivation on the location of Flag-HDAC6 and its various truncations. Cell fractionation analysis was performed with HCT116 cells transfected with Flag-HDAC6 and its various truncations. **l**, Quantification of HDAC6 nuclear and cytoplasmic localization for Fig. 2f. For **e**–**g**,**l**, data are presented as mean ± s.d. (n = 3 independent experiments). Statistical analysis was performed using one-way ANOVA (**f**, **g**, **l**); ****$P < 0.0001$, NS, not significant. For gel source data, see Supplementary Fig. 1.



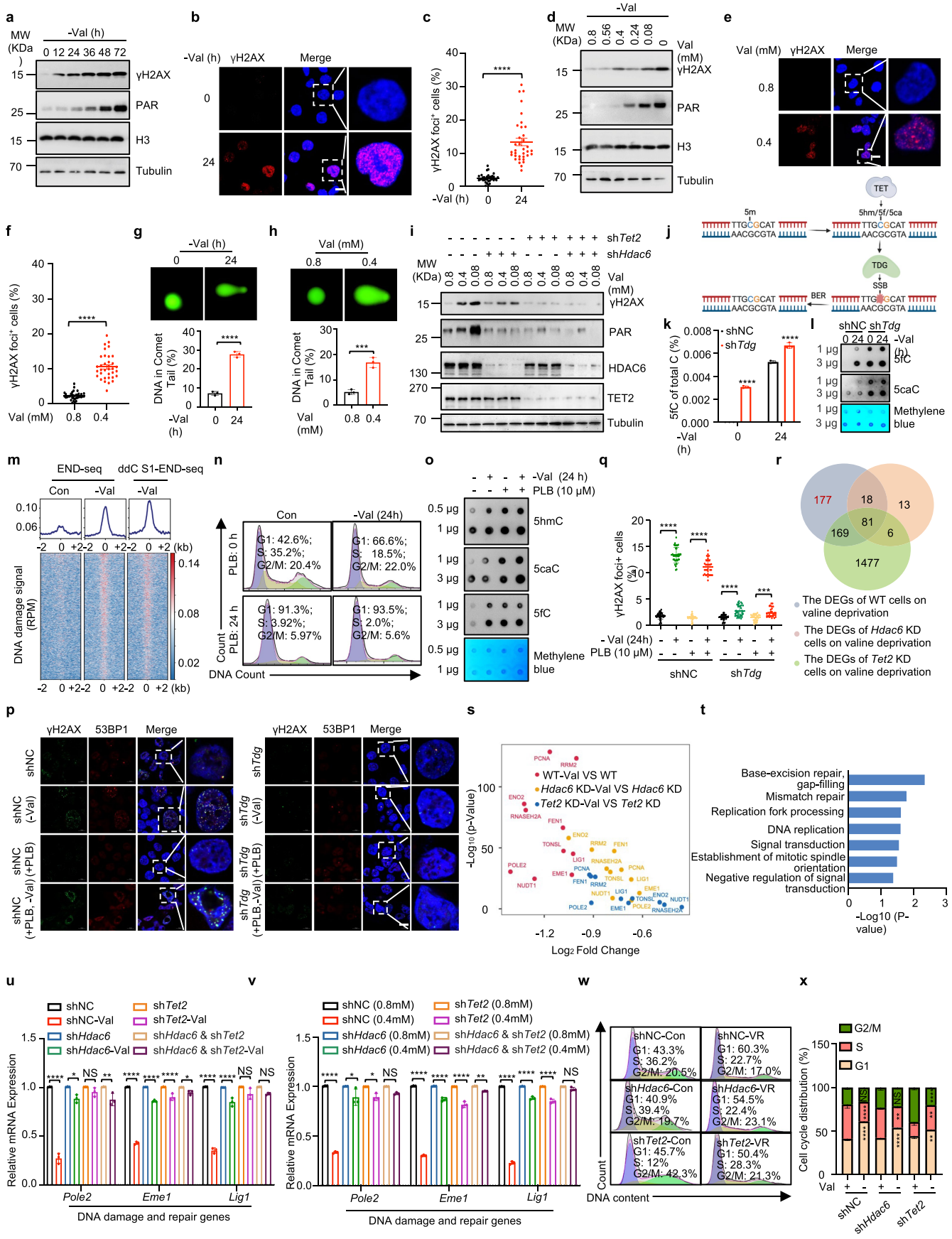
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Importin α1 mediates nuclear translocation of HDAC6 via SE14 domain. a, Effects of valine deprivation on the subcellular localization of histone deacetylases. Cell fractionation analysis was performed with HCT116 cells transfected with Flag-HDAC4, Flag-HDAC6 or Flag-SIRT7. **b**, Effects of valine deprivation on the acetylation of HDAC6. HEK293T cells transfected with Flag-HDAC6 and HA-P300 were subjected to immunoprecipitation with Flag antibody after valine withdraw for 6 h, followed by immunoblotting analysis. **c**, Effects of importin inhibitor on the subcellular localization of HDAC6. Cell fractionation analysis was performed with HCT116 cells pretreated with Importazole (10 μM) for 2 h and then subjected to valine withdraw or not for 6 h. **d**, Immunofluorescence analysis after HCT116 cells expressing HDAC6 or its mutants were pretreated with the nuclear exportin inhibitor Leptomycin B (10 μM) or nuclear importin inhibitor Importazole (10 μM) for 2 h and then subjected to valine deprivation or not for 6 h. Scale bar, 10 μm. **e**, The interaction between HDAC6 and importins. Immunoblot results of cell lysates and anti-Flag or HA immunoprecipitates from HEK293T cells transfected with the indicated Flag–importin α or HA-importin β for 24 h and then subjected to valine deprivation for 6 hrs. **f**, Effects of valine deprivation on the interaction between hHDAC6 and importin α proteins. Immunoprecipitation was performed using HEK293T cells transfected with Flag-importin α for 24 h and subjected to valine deprivation for 6 h. **g**, Effects of valine re-introduction on the interaction between HDAC6 and importin α1. Immunoprecipitation was performed using HEK293T cells transfected with Flag-importin α1 for 24 h with or without re-supplement of valine for 6 h after valine deprivation for 6 h. Valine (10 mM) was add to the cell exacts before co-immunoprecipitation. **h**, Immunoprecipitation was performed using HEK293T cells transfected with Myc-importin α1 and Flag-HDAC6 truncation mutants (ΔNLS and ΔSE14) for 24 h and subjected to valine deprivation for 6 h. **i**, SE14 repeat domain only exists in primates. Multiple alignment of HDAC6 protein sequences across species was analysed by Clustal omega. Seven SE14 repeats are boxed. SE14-3 repeat is also indicated. **j**, Human, but not murine, HDAC6 binds to biotinylated valine. Pull-down assay was performed using HEK293T cells and MEF cells, respectively. **k**, Effects of valine deprivation on the subcellular localization of hHDAC6 and mHdac6. Cell fractionation analysis was performed using HCT116 cells transfected with Flag-hHDAC6 and Flag-mHdac6. **l**, Binding of biotinylated valine to human HDAC6, murine Hdac6 (mHdac6) or the chimera of mHdac6 and the SE14 repeat domain (cHdac6). Pull-down assay was performed using HCT116 cells transfected with Flag-hHDAC6, HA-mHdac6 and its chimaera HA-cHdac6. **m**, Schematic diagram of SE14 repeat domain knocking-in into the m*Hdac6* gene. **n**, The SE14 repeat domain inserted into the 25th exon of m*Hdac6* was verified by DNA sequencing. **o, p**, Validation of HDAC6 knockout efficiency via immunofluorescence (**o**) and immunoblotting (**p**). Scale bars, 10 μm. **q, r**, The function of SE14 repeat domain in sensing valine deprivation. The location of mHdac6 in wildtype (WT) and SE14 repeat domain knock-in (KI) MEF cells exposed to valine deprivation for 6 hrs or not were detected by cell fractionation assay (**q**) and immunofluorescence (**r**). Scale bar, 10 μm. **s**, The interaction between valine and the SE14 repeat domain of human and *Rhinopithecus bieti*. These protein of SE14 repeat domain were prepared from *E. coli*. **t**, Analysis of the subcellular location of Hdac6 in *Rhinopithecus bieti*-fibroblast cells via Immunofluorescence. Scale bar, 10 μm. For gel source data, see Supplementary Fig. 1.



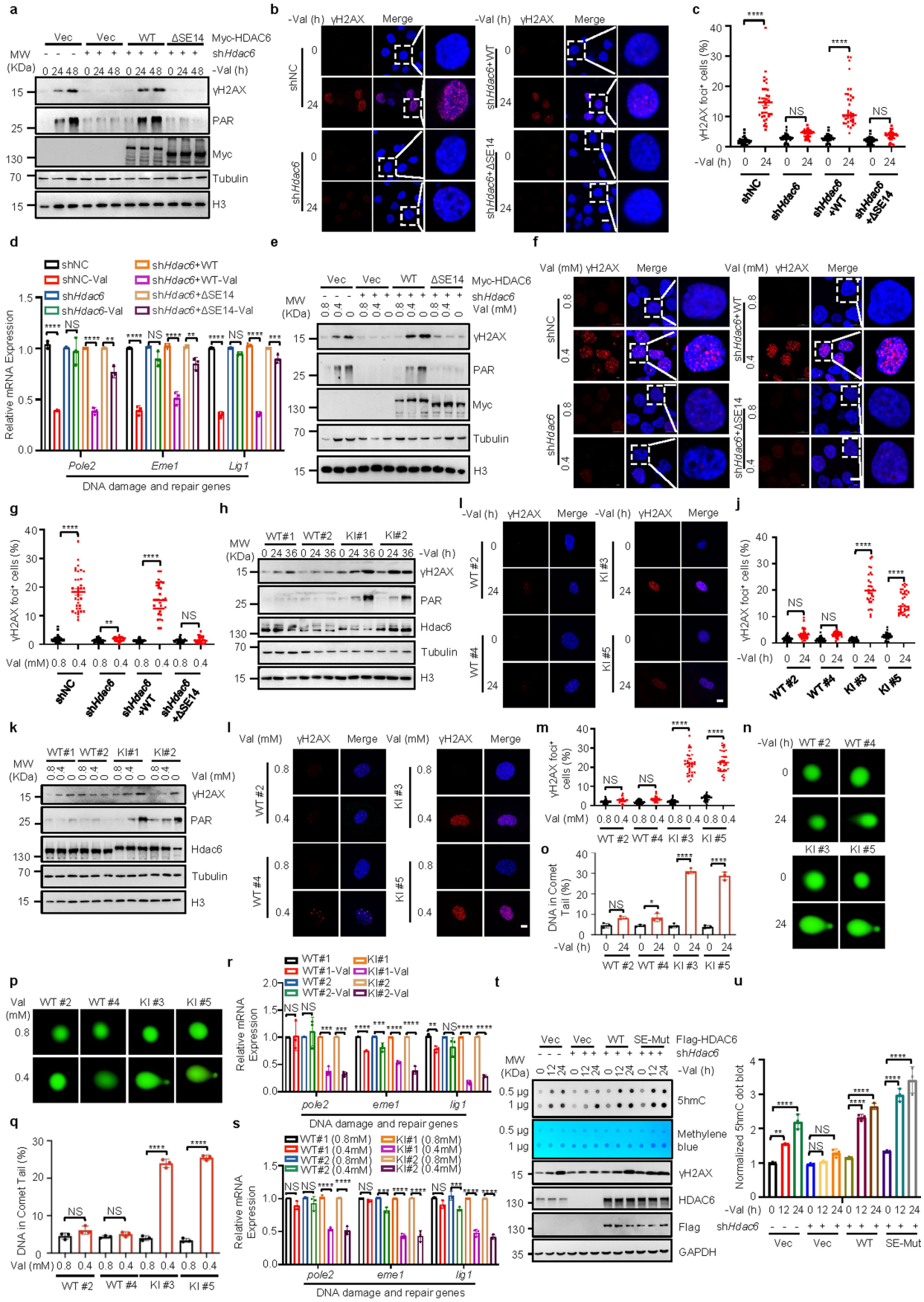
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Valine deprivation increases the activity of TET2 dependent on the nuclear translocation of HDAC6. a, Schematic representation of experimental workflow to identify potential HDAC6 binding proteins under valine deprivation for 6 h by the SILAC quantitative proteomics. **b**, Identification of differential TET2-binding proteins in valine deprivation for 12 h via data-independent acquisition-based mass spectrometry. **c**, Effects of valine restriction on the endogenous interaction between HDAC6 and TET2 detected by co-immunoprecipitation. Valine was deprived for 6 h and then HCT116 cells subjected to add valine in a concentration gradient. **d**, Addition of valine (10 mM) to the cell extract expressing Flag-HDAC6 and Myc-TET2 negatively affects the interaction between HDAC6 and TET2 upon valine deprivation for 12 h via co-immunoprecipitation. **e**, Ectopically expressed HDAC6 binds to the TET2 CD domain via co-immunoprecipitation assay. HDAC6 binds to the CD domain of TET2 upon valine deprivation for 12 h. **f**, TET2 interacts with the exogenous overexpressed DAC domain of HDAC6 upon valine deprivation for 24 h or not. **g**, **h**, Effects of valine deprivation on 5hmC levels in HCT116 cells upon valine deprivation for different time. Quantification of dot blots in **g**. **i**, Effect of valine deprivation on 5hmC levels in *Rhinopithecus bieti*-fibroblast cells upon valine deprivation for different times via Immunofluorescence. **j**, UPLC–MS/MS results showing the effects of valine deprivation in HCT116 cells. **k**, Effects of valine deprivation on 5hmC levels. HCT116 HDAC6 WT or KO cells were exposed to valine deprivation at different time points and the cells were fixed and immunostained with anti-5hmC (red) antibodies. Scale bar, 10 μm. **l**, **m**, Effects of valine deprivation on 5hmC levels in WT or HDAC6 knockout HCT116 cells examined by flow cytometry (**l**). **n**. Quantification of dot blots in Fig. 3b. **o**, **p**, Bar plot showing the fraction of 5hmCG against covered CG sites (**o**) and the fraction of 5mCG against covered CG sites (**p**) in HDAC6 knockdown, TET2 knockdown and WT HCT116 cell lines upon valine deprivation for 24 h or not. **q**, Principal component analysis (PCA) of TET2 ChIP-Seq data. "-Val", valine deprivation for 24 hrs. **r**, Clustering of TET2 binding peaks mainly contributing to the separation of the four groups in PCA. Three replicates for each sample. **s**, Violin plots showing the 5hmC level (**q**) and 5fC/5caC level (**r**) in WT, HDAC6 knockdown, TET2 knockdown HCT116 cell lines upon valine deprivation at the valine deprivation-specific enhanced TET2 binding regions. The red point indicates the mean value of the given data. **t**, **u**, In vitro TET2 catalytic activity assay. Synthesized methylated dsDNA or genomic DNA from HCT116 cells were incubated with purified proteins of TET2 and HDAC6, TET2-catalyzed oxidation was measured by 5mC decrease and 5hmC accumulation as determined by dot-blot assay (**t**). Quantification of dot blots in **u**. **v**, Analysis the deacetylase of HDAC6 and various truncated proteins eluted from HCT116 cells expressing Flag-HDAC6 or its various truncations via immunoprecipitation. **w**, Effects of valine deprivation on the acetylation of TET2 in WT, HDAC6 knockdown or SIRT1 knockdown HCT116 cells upon valine deprivation for 6 hrs. **x**, **y**, Analysis of 5hmC level of HDAC6 or SIRT1 knockdown HCT116 cells upon valine deprivation at different time (**x**). Quantification of dot blots in **y**. **z**. Schematic of HDAC6 promoting TET2 activity in response to valine deprivation. For **h**, **j**, **m**, **n**, **u**, **v** and **y**, data are presented as mean ± s.d. (n = 3 independent experiments for **h**, **m**, **n**, **u**, **v**, **y**; n = 6 independent experiments for **j**). Statistical analysis was performed using one-way ANOVA (**h**, **j**, **n**, **v**, **y**) and two-way ANOVA (**m**, **u**); $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$, NS, not significant. Schematic in **z** was created using BioRender (https://BioRender.com). For gel source data, see Supplementary Fig. 1.



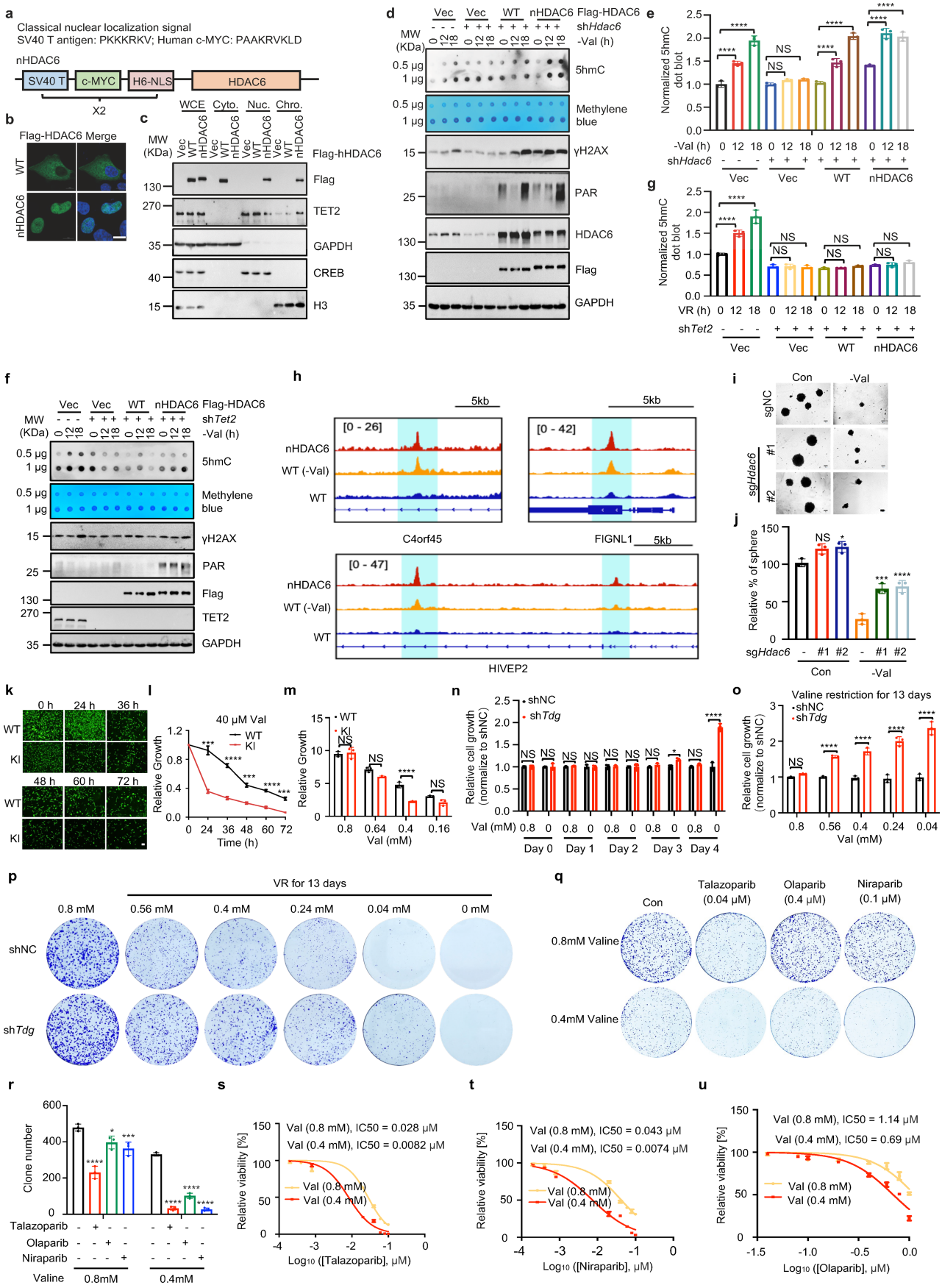
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Valine deprivation promotes the DNA damage response. a, Immunoblot of DNA damage marker γH2AX and PARylation (PAR) in HCT116 cells upon valine deprivation with different time points. **b**, **c**, Representative images (**b**) and quantification (**c**) with immunofluorescence staining for γH2AX (red) in HCT116 cells upon valine deprivation for 24 h. Scale bar, 10 μm. **d**, Immunoblot of DNA damage marker γH2AX and PARylation (PAR) in HCT116 cells cultured with different concentrations of valine for 48 h. **e**, **f**, Representative images (**e**) and quantification (**f**) with immunofluorescence staining for γH2AX (red) in HCT116 cells upon valine deprivation for 48 h at different concentrations. Scale bar, 10 μm. **g**, Representative comet assay (top) and quantification (bottom) showing the tail moment of HCT116 cells upon valine deprivation for 24 h. **h**, Representative comet assay (top) and quantification (bottom) showing the tail moment of HCT116 cell line upon valine deprivation for 48 h at different concentrations. **i**, Immunoblot of DNA damage marker γH2AX and PARylation (PAR) in HDAC6 and TET2 knockdown HCT116 cell line upon valine deprivation at different concentration with 48 h. **j**, Schematic of programmed breaks in active DNA demethylation via TET2-TDG axis. **k**, The levels of 5fC upon valine deprivation in WT, TDG knockdown HCT116 cell lines via UPLC–MS/MS. **l**, The 5caC and 5fC levels in genome upon valine deprivation in WT, TDG knockdown HCT116 cell lines via Dot-blot assay. **m**, Average plots (top) and heatmaps (bottom) showing the enrichment of END-seq and ddC-S1-END-seq signals in HCT116 cell line upon valine deprivation at the valine deprivation-specific enhanced TET2 binding regions. "0", peak centre. **n**, Flow cytometry analysis of cell cycle upon valine deprivation for 24 h with the pretreatment of Palboclib (PLB, PD0332991) for 24 h in HCT116 cell lines. **o**, The 5hmC, 5caC and 5fC levels in genome upon valine deprivation for 24 h with the pretreatment of Palboclib (PLB, PD0332991) for 24 h in HCT116 cell lines via dot-blot assay. **p**, **q**, Representative images (**p**) and quantification (**q**) with immunofluorescence staining for γH2AX (red) and 53BP1 (green) in WT or *Tdg* knockdown HCT116 cell line upon pretreatment with Palboclib (PLB, PD.332991) for 24 h and then valine deprivation for 24 h. Scale bar, 10 μm. **r**, The overlap of differentially expressed genes (DEGs) in WT, HDAC6 knockdown and TET2 knockdown HCT116 cells upon valine deprivation. **s**, Among the 177 valine-modulated and HDAC6-TET2 axis dependent genes, the change expression of DNA damage and repair-related genes in three types of HCT116 cell lines upon valine deprivation. **t**, GO analysis of the 177 valine-modulated and HDAC6-TET2 axis dependent genes in r. **u**, **v**, RT–qPCR validation of a subset of DNA-damage and repair genes upon valine deprivation with 24 h (**u**) or at different valine concentration with 24 h (**v**). **w**, **x**, Flow cytometry analysis of cell cycle (**w**) and quantification (**x**). Schematic in **j** was created using BioRender (https://BioRender.com). For **c**, **f**, and **q**, n = 35 microscopic views examined across 3 independent experiments. For **g**, **h**, **k**, **u**, **v** and **x**, n = 3 independent experiments. Data are presented as mean ± s.d.. Statistical analysis was performed using Mann–Whitney *U*-test (**c**, **f**, **g**, **h**, **q**) and one-way ANOVA (**k**, **u**, **v**, **x**); $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$, NS, not significant. For GO analysis, one-sided Fisher's exact test P-values are applied to evaluate gene enrichment in annotation terms. For gel source data, see Supplementary Fig. 1.



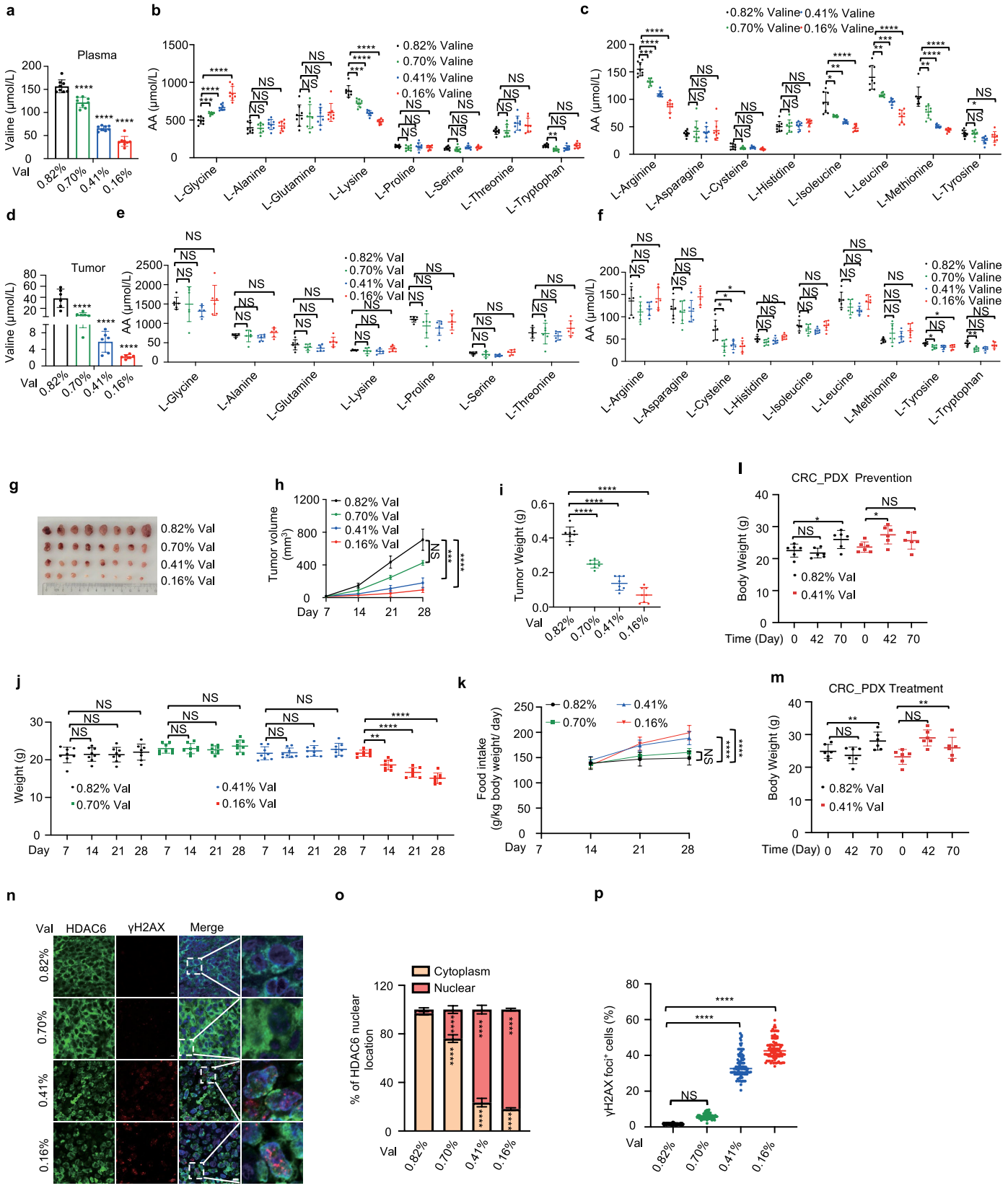
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Valine deprivation promotes the DNA damage response via primates-specific SE14 repeat domain. a, Immunoblot of DNA damage marker γH2AX and PARylation (PAR) in HDAC6 WT, HDAC6 knockdown (KD), HDAC6 WT and HDAC6 SE14 repeat domain deleted mutation expressed in HDAC6 KD HCT116 cell lines upon valine deprivation with different time points. **b**, **c**, Representative images (**b**) and quantification (**c**) with immunofluorescence staining for γH2AX (red) in HDAC6 WT, HDAC6 knockdown (KD), HDAC6 WT and HDAC6 SE14 repeat domain deleted mutation expressed in HDAC6 KD HCT116 cell lines upon valine deprivation for 24 h. Scale bar, 10 μm. **d**, RT–qPCR validation of a subset of DNA-damage and repair genes in HDAC6 WT, HDAC6 knockdown (KD), HDAC6 WT and HDAC6 SE14 repeat domain deleted mutation expressed in HDAC6 KD HCT116 cell lines upon valine deprivation for 48 h. **e**, Immunoblot of DNA damage marker γH2AX and PARylation (PAR) in HDAC6 WT, HDAC6 knockdown (KD), HDAC6 WT and HDAC6 SE14 repeat domain deleted mutation expressed in HDAC6 KD HCT116 cell lines cultured with different concentrations of valine for 48 h. **f**, **g**, Representative images (**f**) and quantification (**g**) with immunofluorescence staining for γH2AX (red) in HDAC6 WT, HDAC6 knockdown (KD), HDAC6 WT and HDAC6 SE14 repeat domain deleted mutation expressed in HDAC6 knockdown HCT116 cell lines cultured with different concentrations of valine for 48 h. Scale bar, 10 μm. **h**, Immunoblot of DNA damage marker γH2AX and PARylation (PAR) in HDAC6 WT and SE14 knock-in MEF cell line upon valine deprivation with different time points. **i**, **j**, Representative images (**i**) and quantification (**j**) with immunofluorescence staining for γH2AX (red) in HDAC6 WT and SE14 knock-in MEF cell line upon valine deprivation for 24 h. Scale bar, 10 μm. **k**, Immunoblot of DNA damage marker γH2AX and PARylation (PAR) in HDAC6 WT and SE14 knock-in MEF cell lines cultured with different concentrations of valine for 48 h. **l**, **m**, Representative images (**l**) and quantification (**m**) with immunofluorescence staining for γH2AX (red) in HDAC6 WT and SE14 knock-in MEF cell line cultured with different concentrations of valine for 48 h. Scale bar, 10 μm. **n**, **o**, Representative comet assay (**n**) and quantification (**o**) showing the tail moment of HDAC6 WT and SE14 knock-in MEF cell line upon valine deprivation for 24 h. **p**, **q**, Representative comet assay (**p**) and quantification (**q**) showing the tail moment of HDAC6 WT and SE14 knock-in MEF cell line cultured with different concentrations of valine for 48 h. **r**, RT–qPCR validation of a subset of DNA-damage and repair genes in HDAC6 WT and SE14 knock-in MEF cell lines upon valine deprivation for 24 h. **s**, RT–qPCR validation of a subset of DNA-damage and repair genes in HDAC6 WT and SE14 knock-in MEF cell lines cultured with different concentrations of valine for 48 h. **t**, **u**, Immunoblot of DNA damage marker γH2AX and the level of 5hmC in HDAC6 WT, HDAC6 knockdown (KD), HDAC6 WT and HDAC6 valine binding sites mutation (SE14-Mut) expressed in HDAC6 KD HCT116 cell lines upon valine deprivation with different time points (**t**). Quantification of dot blots in **u**. For **c**, **g**, **j**, **m**, **n** = 35 microscopic views examined across 3 independent experiments. For **d**, **o**, **q**, **r**, **s** and **u**, n = 3 independent experiments. Data are presented as mean ± s.d.. Statistical analysis was performed using Mann–Whitney *U*-test (**c**, **g**, **j**, **m**, **o**, **q**), one-way ANOVA (**d**, **r**, **s**, **u**); $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$, NS, not significant. For gel source data, see Supplementary Fig. 1.



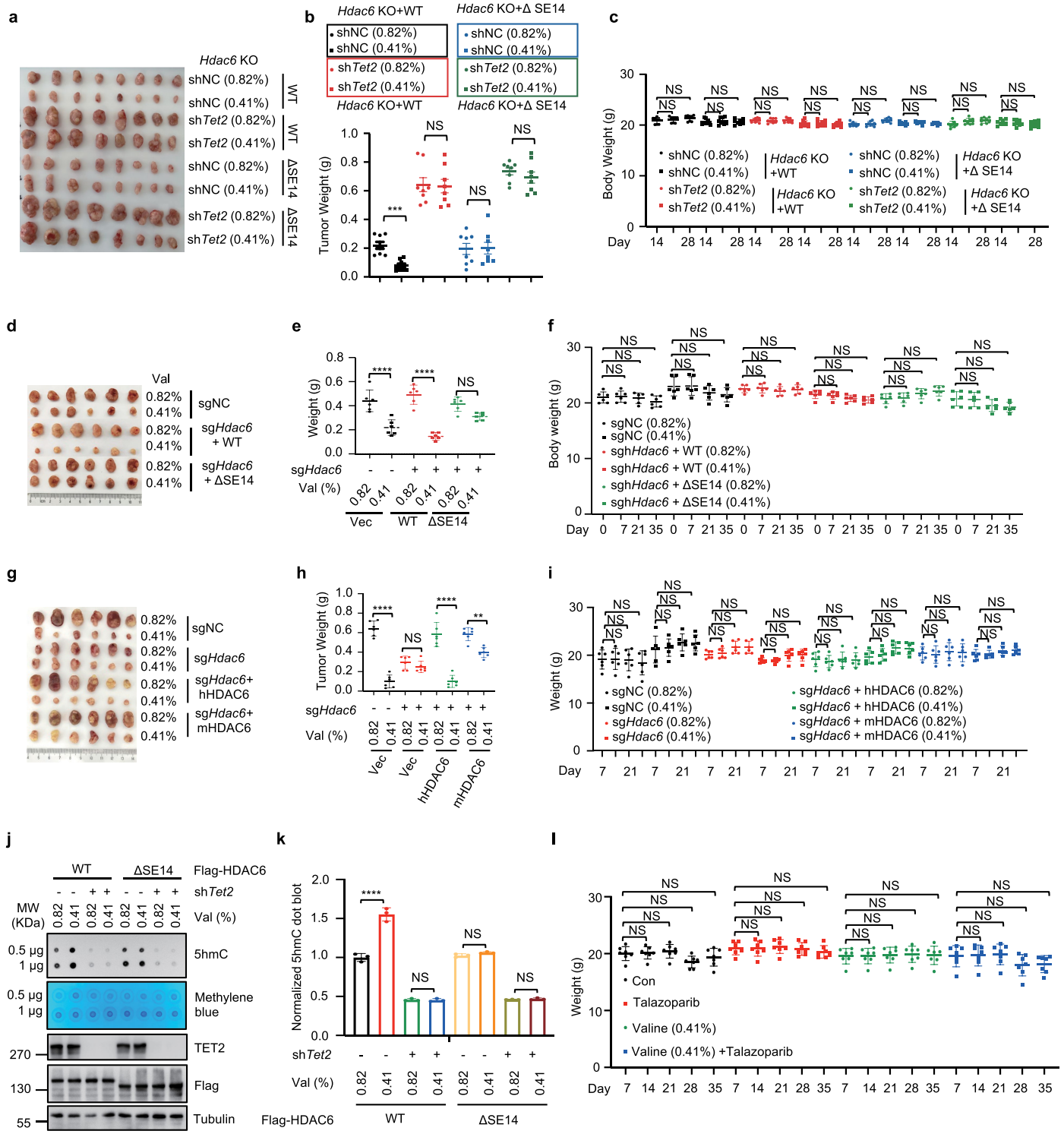
Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | The nuclear localization form of HDAC6 promotes hydroxymethylation and DNA damage. a–**c**, Schematic representations of nuclear localization form of HDAC6 (nHDAC6) (**a**) and verification the location via immunofluorescence (**b**) and cellular fraction (**c**). Scale bar, 10 μm. **d**, **e**, The level of 5hmC and immunoblot of DNA damage marker γH2AX and PARylation (PAR) in HDAC6 WT, HDAC6 knockdown (KD), HDAC6 WT and nHDAC6 expressed in HDAC6 KD HCT116 cell lines upon valine deprivation with different time points via dot blot (**d**). Quantification of dot blots in **e**. **f**, **g**, The level of 5hmC and immunoblot of DNA damage marker γH2AX and PARylation (PAR) in TET2 WT, TET2 knockdown (KD), HDAC6 WT and nHDAC6 expressed in TET2 KD HCT116 cell lines upon valine deprivation with different time points via dot blot (**f**). Quantification of dot blots in **g**. **h**, Genome browser view showing the several instances of TET2 binding peaks among WT, WT (-Val, valine deprivation for 24 hrs) and nHDAC6 HCT116 samples. **i**, Representative oncosphere images of HDAC6 WT or KO HCT116 cells subjected to valine deprivation. Scale bar, 10 μm. **j**, Frequency of tumour spheres formed from HCT116 cells in **i**. Sphere counts are normalized to the control count. **k**, Representative images of EdU incorporation in mHdac6 and mHdac6-hSE14 MEF cells with valine deprivation at indicated time points. Scale bar, 100 μm. **l**, Quantification of EdU incorporation in **k**. **m**, Relative cell growth of mHdac6 and mHdac6-hSE14 MEF cells with valine deprivation at indicated concentration for 5 days. **n**, Relative cell growth of WT and TDG knockdown HCT116 cells with or without treatment of valine deprivation for 4 days via MTT assay. **o**, **p**, Quantification of crystal violet staining (**o**) and representative images of WT and TDG knockdown HCT116 cells (**p**) at indicated concentration for 13 days. **q**, **r**, Representative images of HCT116 cells treated with the indicated PARP inhibitors at different concentration upon valine restriction (0.4 mM) and normal condition (0.8 mM) via crystal violet staining assay (**q**). Quantification of crystal violet staining in **r**. **s**, **t**, **u**, Valine deprivation increases sensitivity to the PARP inhibitor talazoparib (**s**), Olaparib (**t**) and Niraparib (**u**). Cell viability was assessed by the crystal violet staining assay. Data are presented as mean ± s.d.. For **e**, **g**, **j**, **l**, **m**, **n**, **o**, **r**, **s**, **t** and **u**, n = 3 independent experiments. Statistical analysis was performed using one-way ANOVA (**e**, **g**, **j**, **m**, **n**, **o**, **r**), Unpaired two-tailed *t*-test (**l**); $^{*}P < 0.05$, $^{***}P < 0.001$, $^{****}P < 0.0001$, NS, not significant. For gel source data, see Supplementary Fig. 1.

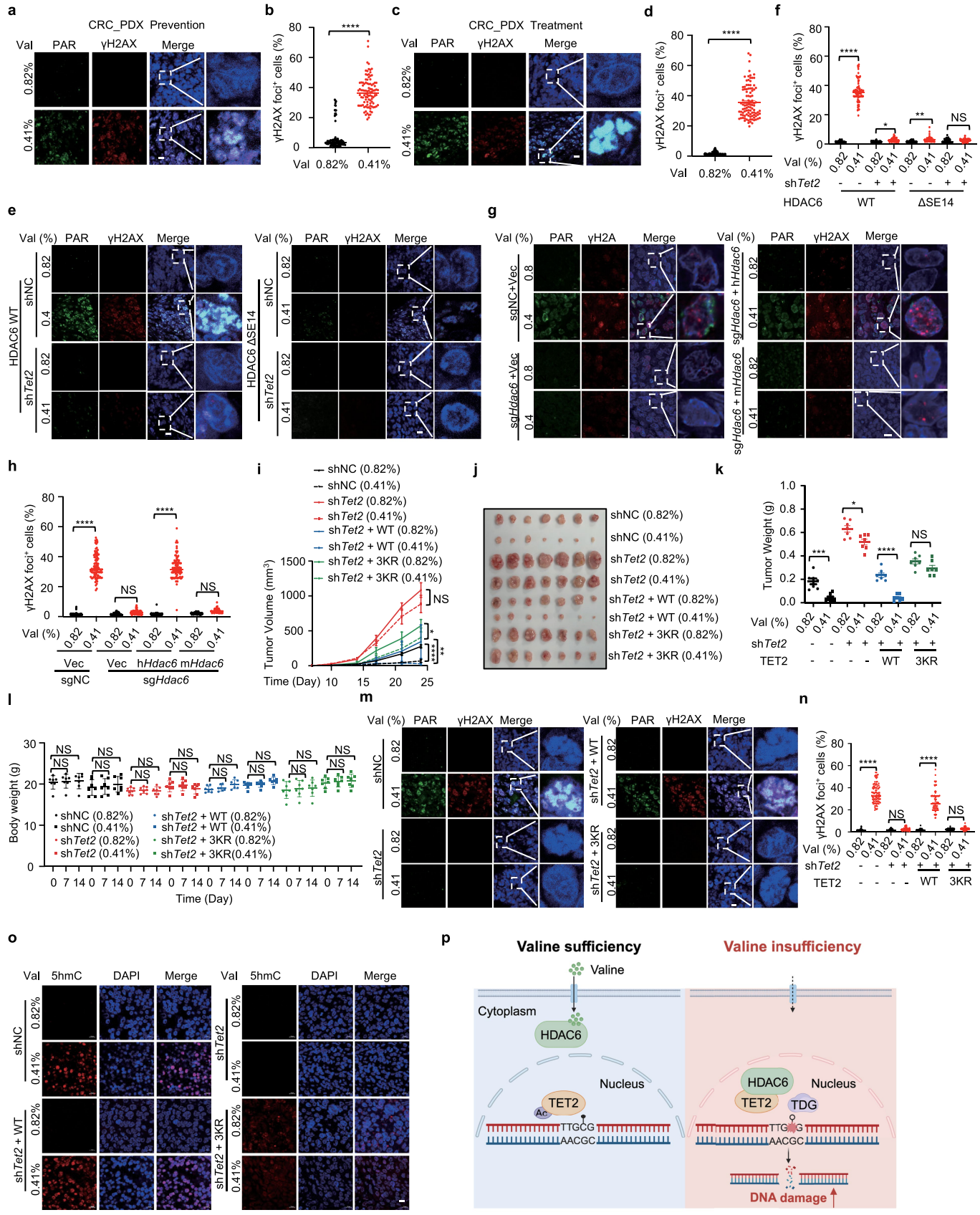


Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Dietary restriction of valine alters valine concentration and HDAC6 nuclear location. a–**c**, Absolute quantification of amino acids in the plasma of the mice fed on diets containing different concentrations of valine was done using targeted metabonomics. **d**–**f**, Absolute quantification of amino acids in the tumour tissue of the mice fed on diets containing different concentrations of valine was done using targeted metabonomics. **g**, **h**, Tumour images (**g**) and volume (**h**) of HCT116 cell-derived tumours in nude mice fed on a special diet containing four concentrations of valine. **i**, Tumour weight at the end point from **g**. **j**, Body weight of nude mice in **g**. **k**, Food intake in **g**. **l**, **m**, Body weight of nude mice grafted with CRC PDX on valine restriction diets in the prevention group (**l**) and treatment group (**m**). **n**, **o**, Representative immunostaining images (**n**) and quantification (**o**) of subcellular localization of HDAC6 in the HCT116 cell-derived tumour sections. Scale bar, 10 μm. **p**, Quantification for γH2AX immunofluorescence staining for γH2AX (red) of tumour sections in **g**. For **a**–**f**, **h**–**m**, **o** and **p**, data are presented as mean ± s.d.. For **a**, **d**, **e**, **f**, **l**, **m**, n = 6 mice. For **b**, **c**, **h**, **i**, **j**, n = 8 mice. For **o**, n = 3 independent experiments. For **p**, n = 100 microscopic views examined for 8 mice. Statistical analysis was performed using one-way ANOVA (**a**, **d**, **h**–**m**, **o**), two-way ANOVA (**b**, **c**, **e**, **f**) and Mann–Whitney *U*-test (**p**); $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$, NS, not significant. For gel source data, see Supplementary Fig. 1.



Extended Data Fig. 9 | The primates-specific SE14 repeat domain is essential for inhibition of cancer progression via valine deprivation. **a**,**b**, Tumour images (**a**) and weight (**b**) at the end point in 5 f. **c**. Body weight of nude mice in 5 f. **d**,**e**, Tumour images (**d**) and weight (**e**) at the end point in 5 g. **f**, Body weight of nude mice in 5 g. **g**,**h**, Tumour images (**g**) and weight (**h**) at the end point in 5 h. **i**, Body weight of nude mice in 5 h. **j**,**k**, The level of 5hmC in tumour sections of 5 g were evaluated by dot blotting (**j**). Quantification of dot blots in **j** (**k**). **l**, Body weight of nude mice in 5k. For **b**, **c**, **e**, **f**, **h**, **i**, **k** and **l**, data are presented as mean ± s.d. (n = 8 mice for **b**, **c**; n = 6 mice for **e**, **f**, **h**, **i**, **l**; n = 3 independent experiments for **k**). Statistical analysis was performed using unpaired two-tailed *t*-test (**b**) and one-way ANOVA (**c**, **e**, **f**, **h**, **i**, **l**, **k**); $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$, NS, not significant. For gel source data, see Supplementary Fig. 1.



Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | Dietary valine restriction inhibits colon cancer progression and DNA damage dependent on the activity of TET2. **a**,**b**, Images (**a**) and quantification for γH2AX (**b**) with immunofluorescence staining for γH2AX (red) and PAR (green) of tumour sections in the prevention group in Fig. 5b. Scale bar, 10 μm. **c**,**d**, Representative images (**c**) and quantification for γH2AX (**d**) with immune-fluorescence staining for γH2AX (red) and PAR (green) of tumour sections in the treatment group in Fig. 5d. Scale bar, 10 μm. **e**,**f**, Representative images (**e**) and quantification for γH2AX (**f**) with immunofluorescence staining for γH2AX (red) and PAR (green) of tumour sections in 5 f. Scale bar, 10 μm. **g**,**h**, Representative images (**g**) and quantification for γH2AX (**h**) with immunofluorescence staining for γH2AX (red) and PAR (green) of tumour sections in 5 h. Scale bar, 10 μm. **i**, Tumour volume of nude mice inoculated with TET2 WT, TET2 knockdown HCT116 cells infected with TET2 WT or mutant virus fed on 0.82% or 0.41% valine diet. **j**,**k**, Tumour images (**j**) and weight (**k**) at the end point from **i**. **l**, Body weight of nude mice in **i**. **m**, **n**, Representative images (**m**) and quantification for γH2AX (**n**) with immunofluorescence staining for γH2AX (red) and PAR (green) of tumour sections in **j**. Scale bar, 10 μm. **o**, Immunostaining of tumour sections for 5hmC (red) and nuclei (blue) from **j**. Scale bar, 10 μm. **p**, Illustration of the proposed mechanism of HDAC6 sensing valine deprivation via the domain of SE14 dictating the epigenetic modification of 5hmC. For **b**, **f**, **h**, **n**, n = 100 microscopic views examined for 6 mice. For d, n = 101 microscopic views examined for 6 mice. For **i**, **k** and **l**, n = 7 mice. Data are presented as mean ± s.d.. Statistical analysis was performed using Mann–Whitney *U*-test (**b**, **d**, **f**, **h**, **n**), one-way ANOVA (**i**, **l**, **k**) and two-way ANOVA (**k**); $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, NS, not significant. Schematic in **p** was created using BioRender (https://BioRender.com). For gel source data, see Supplementary Fig. 1.

nature portfolio

Corresponding author(s): Dr. Ping Wang

Last updated by author(s): Oct 3, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	BD FACSuite were used to collect flow cytometry data. Model CHI760E electrochemical workstation was used for the collection of Electrochemical Impedance Spectroscopy (EIS). BioRad CFX was used to collect data of thermal shift. Illumina NovaSeq 6000 was used to collect data of RNA-Seq, WGBS, ACE-Seq and MAB-Seq. Illumina Hi-Seq was used to collect data of ChIP-Seq. DNBSEQ-T7 was used to collect data of END-seq and S1-END-seq. Cometscore software was used to measure tail moments of DNA damage. Zeiss was used to collect immunofluorescence staining image. Orbitrap QExactive HF (Thermo Fisher Scientific, USA) platform and nanoflow EASY nLC1000 HPLC system were used to collect LC-MS data. SCIEX Triple Quad™ 6500+ LC-MS/MS were used to collect 5mC, 5hmC, and 5fC data.
Data analysis	Flowjo v10 was used for analysis of flow cytometry data. GraphPad Prism v8.0 was used for statistical analysis of all data. ZSimpWin EIS DATA analysis software (Perkin-Elmer, Version 2.00) was used to analyze the obtained EIS data. Cometscore software was used to analyze the tail moments of DNA damage . ImageJ was used for analysis of histology and immunohistochemistry. DIA-NN software was used to analyze LC-MS data. FASTQC (v0.11.8), TrimGalore (v0.6.4) and R (v3.6.0) were used for all sequencing data. HISAT2 (v2.1.0), StringTie (v2.2.1),DEseq2 (v 1.26.0) and DAVID Bioinformatics Resource were used for analysis of RNA-Seq data. Bowtie2 (v2.3.5.1), Sambamba (v0.7.0), deepTools (v3.3.0), MACS2 (v2.1.2) and DAVID Bioinformatics Resource were used to analyze ChIP-Seq data.

Bismark (v0.22.1), MethPipe (v3.4.3) and MLML (v.5.0.1) were used to analyze WGBS, ACE-Seq and MAB-Seq data.
Bowtie2 (v2.3.5.1), bedtools(v2.28.0) were used to analyze END-Seq and S1-END-Seq data.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All omics data, GSE274535, GSE274536, GSE274539, GSE275023, GSE274881, GSE274883, GSE275061, GSE275062, has been released. All other data are available in the article and its Supplementary Information as well as from the corresponding author on reasonable request. Gel source images are shown in Supplementary Fig. 1. Source data are provided in this paper.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	*Use the terms sex (biological attribute) and gender (shaped by social and cultural circumstances) carefully in order to avoid confusing both terms. Indicate if findings apply to only one sex or gender; describe whether sex and gender were considered in study design; whether sex and/or gender was determined based on self-reporting or assigned and methods used.* *Provide in the source data disaggregated sex and gender data, where this information has been collected, and if consent has been obtained for sharing of individual-level data; provide overall numbers in this Reporting Summary. Please state if this information has not been collected.* *Report sex- and gender-based analyses where performed, justify reasons for lack of sex- and gender-based analysis.*
Reporting on race, ethnicity, or other socially relevant groupings	*Please specify the socially constructed or socially relevant categorization variable(s) used in your manuscript and explain why they were used. Please note that such variables should not be used as proxies for other socially constructed/relevant variables (for example, race or ethnicity should not be used as a proxy for socioeconomic status).* *Provide clear definitions of the relevant terms used, how they were provided (by the participants/respondents, the researchers, or third parties), and the method(s) used to classify people into the different categories (e.g. self-report, census or administrative data, social media data, etc.)* *Please provide details about how you controlled for confounding variables in your analyses.*
Population characteristics	Colorectal cancer specimens from one patient were used for PDX model. CRC_PDX: female, 62 years old, KRAS mutation, NRAS and BRAF wildtype.
Recruitment	None selection had been made. Recruitment and collection of human samples adhered to approved IRB protocols. In brief, PDX models of colorectal cancer were developed under the institutional review board of Shanghai Tenth People's Hospital (approval number: 2020-KN171-01). CRC tumours were resected, washed and minced, and then passaged through BALB/c nude mice.
Ethics oversight	This study was conducted according to the Declaration of Helsinki and approved by the the institutional review board of Shanghai Tenth People's Hospital (approval number: 2020-KN171-01). and TUSM institutional review board, with written informed consent obtained from all patients.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical calculation were performed to predetermine simple size. In cases where statistics were derived, sample size was n=3 or more independent biological replicates. Group sizes for in vivo experiments were selected empirically based upon prior knowledge of the intragroup variation of tumor growth, in which group sizes of 6-10 mice or treatment are commonly used in similar studies in literatures. These sample sizes are sufficient to allow for the determination of statistical significance between groups and minimized the number of animals or replicates needed for each experiment. Sample size was determined on the previous studies in the field using similar experiment paradigms

[PMID: 32814895,31371811,31634899, 31634900, 32814895, 35922516], and to give sufficient data values to conduct standard statistical tests.

Data exclusions: No data has been excluded from this study.

Replication: Replicates were used in all experiments as indicated in text, figure legends and methods. All experiments presented in this study were performed at least twice.

Randomization: For animal experiments, animals were randomized for testing different conditions. For experiments not involving mice, cells were randomized into experimental groups.

Blinding: For all animal experiments, all mice were assigned a specific number before data collection, and all data collection and analysis were blindly. For all Mass Spectrometry were performed with the investigators blinded to the condition.
For other vitro experiments, blinding was not done because the investigators needs to know the treatment groups in order to perform the experiments.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used: For IB: The following primary antibodies were used: HDAC6 (1:1000, 12834-1-AP, Proteintech), α-Tubulin (1:1000, 11224-1-AP, Proteintech), GAPDH (1:2000, 10494-1-AP, Proteintech), acetylated Tubulin(Lys40) (1:1000, 66200-1-Ig, Proteintech), GAPDH (1:1000, 10494-1-AP, Proteintech), Sestrin2 (1:1000, 10795-1-AP, Proteintech), Myc (1:1000, 16286-1-AP,Proteintech), HA (1:3000, 51064-2-AP, Proteintech), FLAG (1:3000, 20543-1-AP, Proteintech), H3 (1:5000, 9715, Cell Signaling Technology), SLC7A5 (1:1000, 32683,Cell Signaling Technology), Acetylated-Lysine (1:1000, 9441,Cell Signaling Technology), TET2 (1:1000, 18950, Cell Signaling Technology), γH2AX (1:1000, 9718, Cell Signaling Technology), PAR (1:1000, 4335-MC-100, R&D Systems), 5hmC (1:5000, 39769, Active Motif), 5fC (1:1000, 61224, Active Motif), 5caC (1:1000, 61226, Active Motif), 5mc (1:5000, A-1014, Epigentek), SIRT1 (1:1000, 13161-1-AP, proteintech), TDG (1:1000, 13370-1-AP, Proteintech), Antibody to GST prepared by our own laboratory.
For IP: TET2 (1:500, 41458, diagenode), H3K4me1 (1:1000, 39298, Active Motif)), HA beads (314385, abmart), FLAG M2 affinity gel (A2220, sigma) , Protein A/G beads (sc-2003, Santa Cruze)
For IF or IHC-IF: HDAC6 (1:500, 12834-1-AP, Proteintech,Rabbit isotype), HDAC6 (1:500, 67250-1-Ig, Proteintech, Mouse isotype), 5hmC (1:500, 39769, Active Motif), γH2AX (1:100, 9718, Cell Signaling Technology), PAR (1:100, 4335-MC-100, R&D Systems), 53BP1 (1:500, NB100-305, NOVUS), Alexa Fluor 488-conjugated AffniPure Goat Anti-Mouse lgG + lgM (H+L) (1:500, 115-545-044 , Jackson), Alexa Fluor@ 594-conjugated AffiniPure Goat Anti-Rabbit lgG (H+L) (1:500, 111-585-003 , Jackson).
For flow cytometry: 5hmC (1:200, 39769, Active Motif), Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, FITC (1:200, F-2765, Invitrogen).

Validation: Validation of primary antibodies:
anti-HDAC6 (12834-1-AP, Proteintech) has been validated in previous publication (PMID: 28190767).
anti-HDAC6 (67250-1-Ig, Proteintech) has been validated in previous publication (PMID: 36788143).
anti-acetylated Tubulin(Lys40) (66200-1-Ig, Proteintech) has been validated in previous publication (PMID: 28966044).
anti-Sestrin2 (10795-1-AP, Proteintech) has been validated in previous publication (PMID: 34290409).
anti-cleaved caspase3 (9664, Cell Signaling Technology) has been validated in previous publication (PMID: 29950720).
anti-Acetylated-Lysine (9441,Cell Signaling Technology) has been validated in previous publication (PMID: 24207025).
anti-TET2 (18950, Cell Signaling Technology) has been validated in previous publication (PMID: 32948596).
anti-γH2AX (9718, Cell Signaling Technology) has been validated in previous publication (PMID: 32302590).
anti-PAR (4335-MC-100, R&D Systems) has been validated in previous publication (PMID: 30356214).
anti-53BP1(NB100-305, NOVUS) has been validated in previous publication (PMID: 38203799).
anti-5hmC (39769, Active Motif) has been validated in previous publication (PMID: 32948596).
anti-5mC (A-1014, Epigentek) has been validated in previous publication (PMID: 32948596).
anti-5fC (61224, Active Motif) has been validated in previous publication (PMID: 32948596).
5caC (1:1000, 61226, Active Motif) has been validated in previous publication (PMID: 32948596).
anti-TET2 (41458, diagenode) has been validated in previous publication (GSM2274677).
Validation detail for the other commercial antibodies is available on the manufacture's website that is attached as following:
anti-SLC7A5 (32683,Cell Signaling Technology), https://www.cellsignal.cn/products/primary-antibodies/lat1-slc7a5-e9o4d-rabbit-

mab/32683
α-Tubulin (11224-1-AP, Proteintech), https://www.ptgcn.com/products/TUBA1B-Antibody-11224-1-AP.htm
GAPDH (10494-1-AP, Proteintech), https://www.ptgcn.com/products/GAPDH-Antibody-10494-1-AP.htm
Myc (16286-1-AP,Proteintech), HA (51064-2-AP, Proteintech), https://www.ptgcn.com/products/MYC-tag-Antibody-16286-1-AP.htm
FLAG (20543-1-AP, Proteintech), https://www.ptgcn.com/products/Flag-Tag-Antibody-20543-1-AP.htm
H3 (9715, Cell Signaling Technology), https://www.cellsignal.cn/products/primary-antibodies/histone-h3-antibody/9715
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, FITC (F-2765, Invitrogen), Alexa Fluor 488-conjugated AffniPure Goat Anti-Mouse lgG + lgM (H+L) (115-545-044 , Jackson),https://www.jacksonimmuno.com/catalog/products/115-545-044
Alexa Fluor@ 594-conjugated AffiniPure Goat Anti-Rabbit lgG (H+L) (111-585-003 , Jackson) , https://www.jacksonimmuno.com/catalog/products/111-585-003.
Antibody to GST prepared by our own laboratory have been validated by our laboratory.
These antibodies are further validated and routinely used in our lab.
Antibody validation including western blots and immunostaing is based on the data by comparing control and knockdown or overexpression samples on the same blot in the main Figures and Extended Data Figures.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	HEK293T, HCT116 cell lines were kindly provided by Cell Bank/Stem Cell Bank, Chinese Academy of Sciences. HDAC6 wild type and knock in MEFs were isolated from E12.5 embryos by our own laboratory. Rhinopithecus bieti fibroblast cells were kindly provided by Dr. Xiaoqing Zhang.
Authentication	All cell lines from Bank/Stem Cell Bank, Chinese Academy of Sciences (CAS) were authenticated by STR DNA profiling analysis through Bank/Stem Cell Bank, CAS.
Mycoplasma contamination	All cell lines tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	For all animal studies, mice were housed at ambient temperature (20-23°C) and humidity (40-60% humidity), kept on normal chow and fed ad libitum. The mHDAC6-hSE14 knock in mice on the C57BL6/J background were generated by Shanghai Model Organisms Center, Inc. For tumour xenograft model, male 5-week-old athymic BALB/c nude mice were purchased from Slaccas.
Wild animals	No wild animals involved in this study.
Reporting on sex	Only male mice were used in this study, because male mice have better physical indicators than female mice. Male mice were less affected by reproduction and hormone levels than female mice.
Field-collected samples	This study did not involve field-collected samples.
Ethics oversight	All protocols used in this study were approved by the Institutional Animal Care and Use Committee at Tongji University.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Plants

Seed stocks	*Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.*
Novel plant genotypes	*Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.*
Authentication	*Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.*

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

All sequencing data generated within this study have been uploaded to the NCBI GEO and are available under the following accession codes:
RNA-Seq: GSE274535; WGBS: GSE274536; ACE-Seq: GSE274539; MAB-Seq: GSE275023; TET2 ChIP-Seq: GSE274881; NLS-HDAC6 TET2 ChIP-Seq: GSE274883; H3K4me1 ChIP-Seq: GSE275024; END-Seq and ddC S1-END-Seq: GSE275061, GSE275062.

Files in database submission

All raw fastq files are available for all experiments. The available processed data files include count files for RNA-Seq data, peak files for ChIP-Seq data, and bw tracks for WGBS, ACE-Seq, MAB-Seq, END-Seq, and ddC S1-END-Seq data.

RNA-Seq data consist of the following files:
T2_V_1_Clean_Data1.fq.gz; T2_V_1_Clean_Data2.fq.gz;
T2_V_2_Clean_Data1.fq.gz; T2_V_2_Clean_Data2.fq.gz;
T2_V_3_Clean_Data1.fq.gz; T2_V_3_Clean_Data2.fq.gz;
T2_1_Clean_Data1.fq.gz; T2_1_Clean_Data2.fq.gz;
T2_2_Clean_Data1.fq.gz; T2_2_Clean_Data2.fq.gz;
T2_3_Clean_Data1.fq.gz; T2_3_Clean_Data2.fq.gz;
Tet2KD_gene_count.txt;
H6_V_1_Clean_Data1.fq.gz; H6_V_1_Clean_Data2.fq.gz;
H6_V_2_Clean_Data1.fq.gz; H6_V_2_Clean_Data2.fq.gz;
H6_V_3_Clean_Data1.fq.gz; H6_V_3_Clean_Data2.fq.gz;
H6_1_Clean_Data1.fq.gz; H6_1_Clean_Data2.fq.gz;
H6_2_Clean_Data1.fq.gz; H6_2_Clean_Data2.fq.gz;
H6_3_Clean_Data1.fq.gz; H6_3_Clean_Data2.fq.gz;
Hdac6KD_gene_count.txt;
NC_V_1_Clean_Data1.fq.gz; NC_V_1_Clean_Data2.fq.gz;
NC_V_2_Clean_Data1.fq.gz; NC_V_2_Clean_Data2.fq.gz;
NC_V_3_Clean_Data1.fq.gz; NC_V_3_Clean_Data2.fq.gz;
NC_1_Clean_Data1.fq.gz; NC_1_Clean_Data2.fq.gz;
NC_2_Clean_Data1.fq.gz; NC_2_Clean_Data2.fq.gz;
NC_3_Clean_Data1.fq.gz; NC_3_Clean_Data2.fq.gz;
NC_gene_count.txt

TET2 ChIP-Seq data consist of the following files:
shHDAC6-Val_1_combined_R1.fq.gz; shHDAC6-Val_1_combined_R2.fq.gz; shHDAC6-Val_1_peaks.narrowPeak;
shHDAC6-Val_2_combined_R1.fq.gz; shHDAC6-Val_2_combined_R2.fq.gz; shHDAC6-Val_2_peaks.narrowPeak;
shHDAC6-Val_3_combined_R1.fq.gz; shHDAC6-Val_3_combined_R2.fq.gz; shHDAC6-Val_3_peaks.narrowPeak;
shHDAC6_1_combined_R1.fq.gz; shHDAC6_1_combined_R2.fq.gz; shHDAC6_1_peaks.narrowPeak;
shHDAC6_2_combined_R1.fq.gz; shHDAC6_2_combined_R2.fq.gz; shHDAC6_2_peaks.narrowPeak;
shHDAC6_3_combined_R1.fq.gz; shHDAC6_3_combined_R2.fq.gz; shHDAC6_3_peaks.narrowPeak;
shNC-Val_1_combined_R1.fq.gz; shNC-Val_1_combined_R2.fq.gz; shNC-Val_1_peaks.narrowPeak;
shNC-Val_2_combined_R1.fq.gz; shNC-Val_2_combined_R2.fq.gz; shNC-Val_2_peaks.narrowPeak;
shNC-Val_3_combined_R1.fq.gz; shNC-Val_3_combined_R2.fq.gz; shNC-Val_3_peaks.narrowPeak;
shNC_1_combined_R1.fq.gz; shNC_1_combined_R2.fq.gz; shNC_1_peaks.narrowPeak;
shNC_2_combined_R1.fq.gz; shNC_2_combined_R2.fq.gz; shNC_1_peaks.narrowPeak;
shNC_3_combined_R1.fq.gz; shNC_3_combined_R2.fq.gz; shNC_1_peaks.narrowPeak

NLS-HDAC6 TET2 ChIP-Seq data consist of the following files:
shHDAC6-NLS-1_R1.fastq.gz; shHDAC6-NLS-1_R2.fastq.gz; shHDAC6-NLS-1_peaks.narrowPeak;
shHDAC6-NLS-2_R1.fastq.gz; shHDAC6-NLS-2_R2.fastq.gz; shHDAC6-NLS-2_peaks.narrowPeak

H3K4me1 ChIP-Seq data consist of the following files:
H3K4me1_R1.fastq.gz; H3K4me1_R2.fastq.gz; H3K4me1_peaks.broadPeak

WGBS data consist of the following files:
NC-Con_rep1_R1.fastq.gz; NC-Con_rep1_R2.fastq.gz; NC-Con_rep1.mincov3.bw;
NC-Con_rep2_R1.fastq.gz; NC-Con_rep2_R2.fastq.gz; NC-Con_rep2.mincov3.bw;
NC-VR_rep1_R1.fastq.gz; NC-VR_rep1_R2.fastq.gz; NC-VR_rep1.mincov3.bw;
NC-VR_rep2_R1.fastq.gz; NC-VR_rep2_R2.fastq.gz; NC-VR_rep2.mincov3.bw;
Hdac6KD-Con_rep1_R1.fastq.gz; Hdac6KD-Con_rep1_R2.fastq.gz; Hdac6KD-Con_rep1.mincov3.bw;
Hdac6KD-Con_rep2_R1.fastq.gz; Hdac6KD-Con_rep2_R2.fastq.gz; Hdac6KD-Con_rep2.mincov3.bw;
Hdac6KD-VR_rep1_R1.fastq.gz; Hdac6KD-VR_rep1_R2.fastq.gz; Hdac6KD-VR_rep1.mincov3.bw;
Hdac6KD-VR_rep2_R1.fastq.gz; Hdac6KD-VR_rep2_R2.fastq.gz; Hdac6KD-VR_rep2.mincov3.bw;
Tet2KD-Con_rep1_R1.fastq.gz; Tet2KD-Con_rep1_R2.fastq.gz; Tet2KD-Con_rep1.mincov3.bw;
Tet2KD-Con_rep2_R1.fastq.gz; Tet2KD-Con_rep2_R2.fastq.gz; Tet2KD-Con_rep2.mincov3.bw;
Tet2KD-VR_rep1_R1.fastq.gz; Tet2KD-VR_rep1_R2.fastq.gz; Tet2KD-VR_rep1.mincov3.bw;
Tet2KD-VR_rep2_R1.fastq.gz; Tet2KD-VR_rep2_R2.fastq.gz; Tet2KD-VR_rep2.mincov3.bw

ACE-Seq data consist of the following files:
NC-Con_R1.fastq.gz; NC-Con_R2.fastq.gz; NC-Con.mincov3.bw;
NC-VR_R1.fastq.gz; NC-VR_R2.fastq.gz; NC-VR.mincov3.bw;
Hdac6KD-Con_R1.fastq.gz; Hdac6KD-Con_R2.fastq.gz; Hdac6KD-Con.mincov3.bw;
Hdac6KD-VR_R1.fastq.gz; Hdac6KD-VR_R2.fastq.gz; Hdac6KD-VR.mincov3.bw;
Tet2KD-Con_R1.fastq.gz; Tet2KD-Con_R2.fastq.gz; Tet2KD-Con.mincov3.bw;
Tet2KD-VR_R1.fastq.gz; Tet2KD-VR_R2.fastq.gz; Tet2KD-VR.mincov3.bw

MAB-Seq data consist of the following files:
NC-Con_R1.fastq.gz; NC-Con_R2.fastq.gz; NC-Con.mincov5.clean.bw;
NC-VR_R1.fastq.gz; NC-VR_R2.fastq.gz; NC-VR.mincov5.clean.bw

END-Seq and ddC S1-END-Seq data consist of the following files:
NC-Con_END-seq_R1.fq.gz; NC-Con_END-seq_R2.fq.gz; NC-Con_END-seq.bw;
NC-VR_END-seq_R1.fq.gz; NC-VR_END-seq_R2.fq.gz; NC-VR_END-seq.bw;
NC-VR_ddC-S1-END-seq_R1.fq.gz; NC-VR_ddC-S1-END-seq_R2.fq.gz; NC-VR_ddC-S1-END-seq.bw;
shNC_ddC-S1-END-seq_R1.fq.gz; shNC_ddC-S1-END-seq_R2.fq.gz; shNC_ddC-S1-END-seq.bw;
shNC-VR_ddC-S1-END-seq_R1.fq.gz; shNC-VR_ddC-S1-END-seq_R2.fq.gz; shNC-VR_ddC-S1-END-seq.bw;
shTDG_ddC-S1-END-seq_R1.fq.gz; shTDG_ddC-S1-END-seq_R2.fq.gz; shTDG_ddC-S1-END-seq.bw;
shTDG-VR_ddC-S1-END-seq_R1.fq.gz; shTDG-VR_ddC-S1-END-seq_R2.fq.gz; shTDG-VR_ddC-S1-END-seq.bw

Genome browser session
(e.g. UCSC)

RNA-Seq:
https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE274535&format=file&file=GSE274535%5FHdac6KD%5Fgene%5Fcount%2Etxt%2Egz
https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE274535&format=file&file=GSE274535%5FNC%5Fgene%5Fcount%2Etxt%2Egz
https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE274535&format=file&file=GSE274535%5FTet2KD%5Fgene%5Fcount%2Etxt%2Egz

TET2 ChIP-Seq:
https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE274881&format=file

NLS-HDAC6 TET2 ChIP-Seq:
https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE274883&format=file

H3K4me1 ChIP-Seq:
https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE275024&format=file

WGBS:
https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE274536&format=file

ACE-Seq:
https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE274539&format=file

MAB-Seq:
https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE275023&format=file

END-Seq and ddC S1-END-Seq:
https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE275062&format=file
https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE275061&format=file

Methodology

Replicates

For RNA-Seq amd TET2 ChIP-Seq data, 3 replicates were used per condition. For WGBS and NLS-HDAC6 TET2 ChIP-Seq data, 2 replicates were used per condition.

Sequencing depth

The Sequencing reads information can be checked at following GEO accession:
RNA-Seq: GSE274535; WGBS: GSE274536; ACE-Seq: GSE274539; MAB-Seq: GSE275023; TET2 ChIP-Seq: GSE274881; NLS-HDAC6 TET2 ChIP-Seq: GSE274883; H3K4me1 ChIP-Seq: GSE275024; END-Seq: GSE275061; ddC S1-END-Seq: GSE275062.

Antibodies

For ChIP-seq in HCT116:
anti-TET2 (1:500, 41458, diagenode)

Peak calling parameters

Peak calling for TET2 ChIP-seq was performed using MACS2 with narrow peak mode and parameter: -p 0.001. Peak calling for H3K4me1 ChIP-seq was performed using MACS2 with broad peak mode and parameter: --broad --broad-cutoff 0.01 -q 0.01.

Data quality

Raw reads were trimed with TrimGalore (v0.6.4) and data quality was assessed by FastQC (v0.11.8).

Software

FastQC v0.11.8, MultiQC v1.13.dev0, TrimGalore v0.6.4, Bowtie2 v2.3.5.1, SAMtools v1.9, deepTools v3.3.0, sambamba v0.7.0, bedTools v2.28.0, MACS2 v2.1.2, HISAT2 v2.1.0, StringTie v2.2.1, DAVID 2021, R v3.6.0, IGV v2.8.9, DEseq2 v1.26.0, ggplot2 v3.3.5, Pheatmap v1.0.12, ChIPseeker v1.22.1, Bismark v0.22.1, MethPipe v3.4.3

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Cells were seeded in 24-well tissue culture plates and treated as shown in the figures. Cells were re-suspended in FACS buffer (PBS with 1% BSA, 2 mM EDTA) and incubated with Fc blocker for 10 min on ice. After washing with FACS buffer, cells were fixed and permeabilized using the Cell Fixation/permeabilization kit (BD Biosciences). DNA was denatured by adding 2N HCl for 20 min and then neutralized with 10 mM Tris-HCl (pH 8.0) for 20 min. Anti-5hmC antibody (1:200, 39769, Active Motif) and FITC conjugated goat anti-rabbit secondary antibody (1:200, F-2765, Invitrogen) were used for 5hmC staining. Data were analyzed using the FlowJo v10 software.
Instrument	BD FACSVerse
Software	Using BD FACSuite to collect data and FlowJo v10 software to analyze data.
Cell population abundance	At least 5,000 cells were analyzed for each sample.
Gating strategy	Initial cell population gating (FSC-A Vs SSC-A) was drawn to exclude debris, then analyzed for 5hmC using the FITC filter.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

A pair of dopamine neurons mediate chronic stress signals to induce learning deficit in *Drosophila melanogaster*

Jia Jia[a,1], Lei He[a,1], Junfei Yang[a,1], Yichun Shuai[b], Jingjing Yang[a], Yalan Wu[a], Xin Liu[a], Tianli Chen[a], Guaxiu Wang[a], Xingyu Wang[a], Xiaoxu Song[a], Zhaowen Ding[a], Yan Zhu[c,d], Li Zhang[e], Peng Chen[e,2], and Hongtao Qin[a,2]

[a]State Key Laboratory of Chemo/Biosensing and Chemometrics, College of Biology, Hunan University, Changsha 410082, P.R. China; [b]HHMI, Janelia Research Campus, Ashburn, VA 20147; [c]State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, P.R. China; [d]College of Life Sciences, University of the Chinese Academy of Sciences, Beijing 100049, P.R. China; and [e]School of Pharmaceutical Science & Yunnan Key Laboratory of Pharmacology for Natural Products, Kunming Medical University, Kunming, Yunnan 650500, P.R. China

Edited by Rachel I. Wilson, Harvard Medical School, Boston, MA, and approved August 27, 2021 (received for review November 14, 2020)

Chronic stress could induce severe cognitive impairments. Despite extensive investigations in mammalian models, the underlying mechanisms remain obscure. Here, we show that chronic stress could induce dramatic learning and memory deficits in *Drosophila melanogaster*. The chronic stress–induced learning deficit (CSLD) is long lasting and associated with other depression-like behaviors. We demonstrated that excessive dopaminergic activity provokes susceptibility to CSLD. Remarkably, a pair of PPL1-γ1pedc dopaminergic neurons that project to the mushroom body (MB) γ1pedc compartment play a key role in regulating susceptibility to CSLD so that stress-induced PPL1-γ1pedc hyperactivity facilitates the development of CSLD. Consistently, the mushroom body output neurons (MBON) of the γ1pedc compartment, MBON-γ1pedc>α/β neurons, are important for modulating susceptibility to CSLD. Imaging studies showed that dopaminergic activity is necessary to provoke the development of chronic stress–induced maladaptations in the MB network. Together, our data support that PPL1-γ1pedc mediates chronic stress signals to drive allostatic maladaptations in the MB network that lead to CSLD.

Drosophila melanogaster | chronic stress | learning and memory | dopamine neuron | depression

Stress has significant and complex effects on cognitive function. In general, these effects follow an inverted U–shaped dose–response relationship in intensity and duration. So that moderate acute stress could promote learning and memory, while chronic stress very often induces detrimental effects (1). Since chronic stress–induced learning and memory impairments are closely associated with many neural disorders, such as depression, schizophrenia, and Alzheimer's disease, understanding the underlying neurobiology is of importance for developing effective drugs and treatments (2, 3). To this end, animal models, especially mammalian models, have been extensively investigated. Current findings suggest that the effects of chronic stress on learning and memory could be influenced by many internal and external factors that involve multiple brain regions, genes, and complex mechanisms that have not yet been fully elucidated (3–6).

Stress could have consequential effects on aversive olfactory memory in *Drosophila melanogaster*. For example, moderate fast promotes long term memory (LTM) formation (7, 8), while sleep deprivation promotes forgetting and impairs memory capacity (9–11). Recent reports have shown that chronic stress can induce depression-like symptoms in *Drosophila*, as manifested by characteristic behaviors that indicate anhedonia, lack of motivation, prone to despair, and sleep disorder (12–14). However, investigation of the effect of chronic stress on *Drosophila* learning and memory is still lacking.

In the present study, we report that a 4-d chronic stress treatment (CST) effectively induces strong learning and memory deficits in *Drosophila*. We focused on the learning deficit phenotype and found that the *Drosophila* dopaminergic (DAergic) system plays an important role in modulating susceptibility to chronic stress–induced learning deficit (CSLD), suggesting that DAergic modulation is an evolutionary conserved chronic stress–coping mechanism. We pinpointed the key CSLD regulating dopamine neurons (DANs) to a pair of PPL1-γ1pedc neurons that project to the mushroom body (MB) γ1pedc compartment and further showed that MBON-γ1pedc>α/β, the output neurons of γ1pedc compartment, modulates susceptibility to CSLD as well. Imaging studies identified chronic stress–induced abnormal neural activities in learning-related neurons, which require DAergic activity during CST. Overall, our studies delineate a model that chronic stress signals can be mediated by a pair of DANs, PPL1-γ1pedc, to drive maladaptations in the MB network that lead to CSLD.

Significance

Chronic stressful life events could induce learning and memory impairments and increase the risk of developing psychiatric disorders such as depression. Understanding the underlying mechanism is critical for developing effective drugs and treatments. Here, we show that chronic stress induces learning and memory deficits in *Drosophila melanogaster*. Furthermore, the dopaminergic system is important for regulating susceptibility to chronic stress– induced learning deficit (CSLD). Significantly, a single pair of dopamine neurons, PPL1-γ1pedc neurons, are indispensable for CSLD. We show that PPL1-γ1pedc mediates chronic stress signals to induce abnormal neural activities in mushroom bodies that lead to a learning deficit. Together, these suggest that *Drosophila melanogaster* can be a powerful model organism for studying the etiology of chronic stress–induced memory impairments.

Author contributions: J.J., L.H., Junfe Yang, P.C., and H.Q. designed research; J.J., L.H., Junfei Yang, Y.S., Jingjing Yang, Y.W., X.L., T.C., G.W., X.W., X.S., Z.D., Y.Z., L.Z., and H.Q. performed research; Y.S. and H.Q. contributed new reagents/analytic tools; J.J., L.H., Junfei Yang, Jingjing Yang, Y.W., X.L., T.C., X.S., Z.D., L.Z., and H.Q. analyzed data; J.J., Y.S., P.C., and H.Q. wrote the paper; and H.Q. secured funding.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]J.J., L.H., and Junfei Yang contributed equally to this work.

[2]To whom correspondence may be addressed. Email: qinhongtao@hnu.edu.cn or ynkmcp@gmail.com.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2023674118/-/DCSupplemental.

Published October 15, 2021.

Results

Chronic Stress Induces Olfactory Learning and Memory Deficits in *Drosophila*. To investigate the impact of chronic stress on learning and memory in adult *Drosophila*, we first established a 4-d chronic stress procedure (Fig. 1*A*). During the 4 d, groups of about 100 flies were cultured in small food vials and subjected to a 10-min mechanical shock each day. In each minute of the 10-min treatment, flies were vortexed (500 rpm) for either 5, 10, or 15 s (Video S1). The vortex was both uncontrollable and unpredictable, as it started randomly within each minute. After the 4-d CST (1 d after the last vortex), flies were tested for olfactory learning (3-min memory) or middle-term memory (3-h memory) with a classical conditioning paradigm as previously described (15, 16). The 4-d CST induced remarkable deficits in both learning (Fig. 1*B*) and middle-term memory (Fig. 1*C*), while 1-d stress treatment had no effect (*SI Appendix*, Fig. S1). The learning deficit appeared as vortex-time dependent, as 10 s of vortex shows a trend of stronger learning deficit than that of 5 s of vortex. As 15 s of vortex did not further diminish olfactory learning, we chose 10-s vortex as the standard mechanical shock condition for our CSTs. With this standard condition, we did not detect major body damage, brain cell death nor intestinal barrier damage after chronic stress (ACS) treatment (*SI Appendix*, Figs. S2–S4), indicating that the learning and memory deficits induced by chronic stress are unlikely the consequence of traumatic brain injury (13, 17–19). Importantly, the task relevant sensorimotor responses (odor acuity and shock reactivity) necessary to perform this task also appeared normal (*SI Appendix*, Fig. S5), suggesting that the standard CST specifically impaired associative learning and memory.

Chronic Stress Induces a Long-Lasting Depression-Like State in *Drosophila*. Recent reports had shown that chronic stress could induce depression-like states in *Drosophila* as evidenced by lack of motivation and anhedonia (12, 13). To determine if our CST induces a depression-like state, we tested three other behavioral assays that had been previously used to assess depression-like symptoms in *Drosophila*. We first monitored the effect of chronic stress on courtship behavior (20). Courtship latency was quantified by measuring the time lag for the male to display the first wing extension. Compared with no treatment controls, chronically stressed flies showed significantly longer courtship latency (Fig. 1*D*), suggesting that these flies were less motivated to court a virgin female. We next investigated anhedonia-like behavior with the stop-for-sweet assay (13). Flies were allowed to stop and feed on a stripe of sweet-tasting glycerol while performing negative geotaxis. Chronically stressed flies made significantly fewer stops compared with controls, suggesting that they were lacking interest in enjoyment (Fig. 1*E*). Third, we tested forced swimming test (FST) (21), a classic behavioral assay for evaluation of antidepression activity. As shown in Fig. 1 *G* and *H*, after the 4-d CST, the latency to the first immobility period was significantly decreased while the duration of immobility was enhanced, indicating a tendency of despair-like behavior. Together with the finding that chronic stress induces learning and memory deficits, these data are consistent with previous reports that chronic stress could induce depression-like behaviors in *Drosophila* (13). To examine if our chronic stress procedure induces the formation of a long-lasting internal state, we tested olfactory learning at either 1 or 2 d ACS treatment. As shown in Fig. 1*I*, although the olfactory learning performance ACS is significantly lower than no treatment control, it is not significantly different from those tested in 1 d (1 d ACS) or 2 d (2 d ACS) later. We also tested FST at either 1 or 2 d ACS. Consistently, compared with no treatment control, the latency to immobility was significantly reduced when tested at each of the three time points (CST, 1 d ACS, or 2 d ACS) (Fig. 1*G*), while the duration of immobility was significantly increased (Fig. 1*H*). Therefore, the effect of chronic stress on these depression-like behaviors persisted for at least 2 d after treatment, suggesting that chronic stress induces a long-lasting depression-like state.

The DAergic System Modulates Susceptibility to CSLD. Since DAergic signaling is important for olfactory learning and responsive to vibration stimulus (22, 23), we reasoned that DAergic signaling might be important for the regulation of CSLD. To test this idea, we first took a pharmacological approach and blocked the rate-limiting enzyme for dopamine (DA) biosynthesis. We fed the flies with 3-iodotyrosine (3-IY), a tyrosine hydroxylase (TH)-specific inhibitor, in the last 2 d of the 4-d CST. Strikingly, down-regulating DA level during the process of CST significantly alleviated CSLD (Fig. 2*B*). Since DA is indispensable during training to form aversive olfactory memory, this counterintuitive observation suggested that chronic stress–induced hyperdopaminergia could wreck the flies' capacity to form associative memory. We then fed the flies with L-DOPA (L-3,4-dihydroxyphenylalanine) plus carbidopa to increase brain DA level (24). Consistent with the above hypothesis, feeding L-DOPA/carbidopa in the last 2 d of CST further decreased the learning performance of chronically stressed flies (Fig. 2*C*).

Second, we examined whether the activity of Dop1R1, a D1-like DA receptor whose function is essential for olfactory learning, is important for CSLD. $Dop1R1^{dumb2}$ is a hypomorphic allele in which the expression of Dop1R1 in MBs is largely abolished (25). We used $Dop1R1^{dumb2/+}$ to down-regulate the expression level of Dop1R1 (25, 26). Untreated $Dop1R1^{dumb2/+}$ showed normal olfactory learning. However, when tested ACS treatment, $Dop1R1^{dumb2/+}$ animals exhibited significantly higher learning performance than wild-type controls (Fig. 2*D*), suggesting that reducing Dop1R1 level alleviates CSLD. These data also indicate that Dop1R1 hyperactivity might dampen olfactory learning. To verify this idea, we overexpressed Dop1R1 with Gal4 drivers driving *UAS-Dop1R1* in otherwise wild-type flies. Remarkably, overexpressing Dop1R1 with OK107, an MB-specific Gal4, driving *UAS-Dop1R1* resulted in a significant learning deficit (Fig. 2*E*). Adding an extra MB Gal4, R75F05, to drive even higher Dop1R1 expression further exacerbated the learning deficit (Fig. 2*E*), suggesting dosage sensitivity. Moreover, introducing an MB-Gal80 transgene into the above line effectively suppressed the learning deficit (Fig. 2*E*), indicating that Dop1R1 overexpression in MBs was necessary for inducing learning deficit. Importantly, sensorimotor responses to the odors or footshock used in the conditioning assay were not affected by Dop1R1 overexpression (*SI Appendix*, Fig. S6 *A*–*C*). Taken together, these data suggested that chronic stress–induced supernormal Dop1R1 activity in MBs precipitates susceptibility to CSLD.

The above experimental data support the idea that stress-induced hyperdopaminergic signaling during CST is important for the development of CSLD. However, both drug feeding and gene expression manipulation lacked fine temporal resolution. To directly test this hypothesis, we targeted genetically encoded tools to DANs for acute neural activity manipulation. To block neurotransmission from the majority of DANs, we used TH-Gal4 to drive $UAS\text{-}Shi^{TS1}$ that encodes a temperature-sensitive and dominant-negative mutant form of dynamin. Synaptic release from DANs could be conditionally inhibited by shifting to the restrictive temperature (27). As shown in Fig. 2 *F* and *G*, interrupting synaptic release from TH-labeled DANs only during the daily mechanical shock treatments was sufficient to prevent CSLD, while CSLD stayed significant if flies were shocked at permissive temperature (18 °C). These data suggest that stress-induced hyperactivity of DANs is

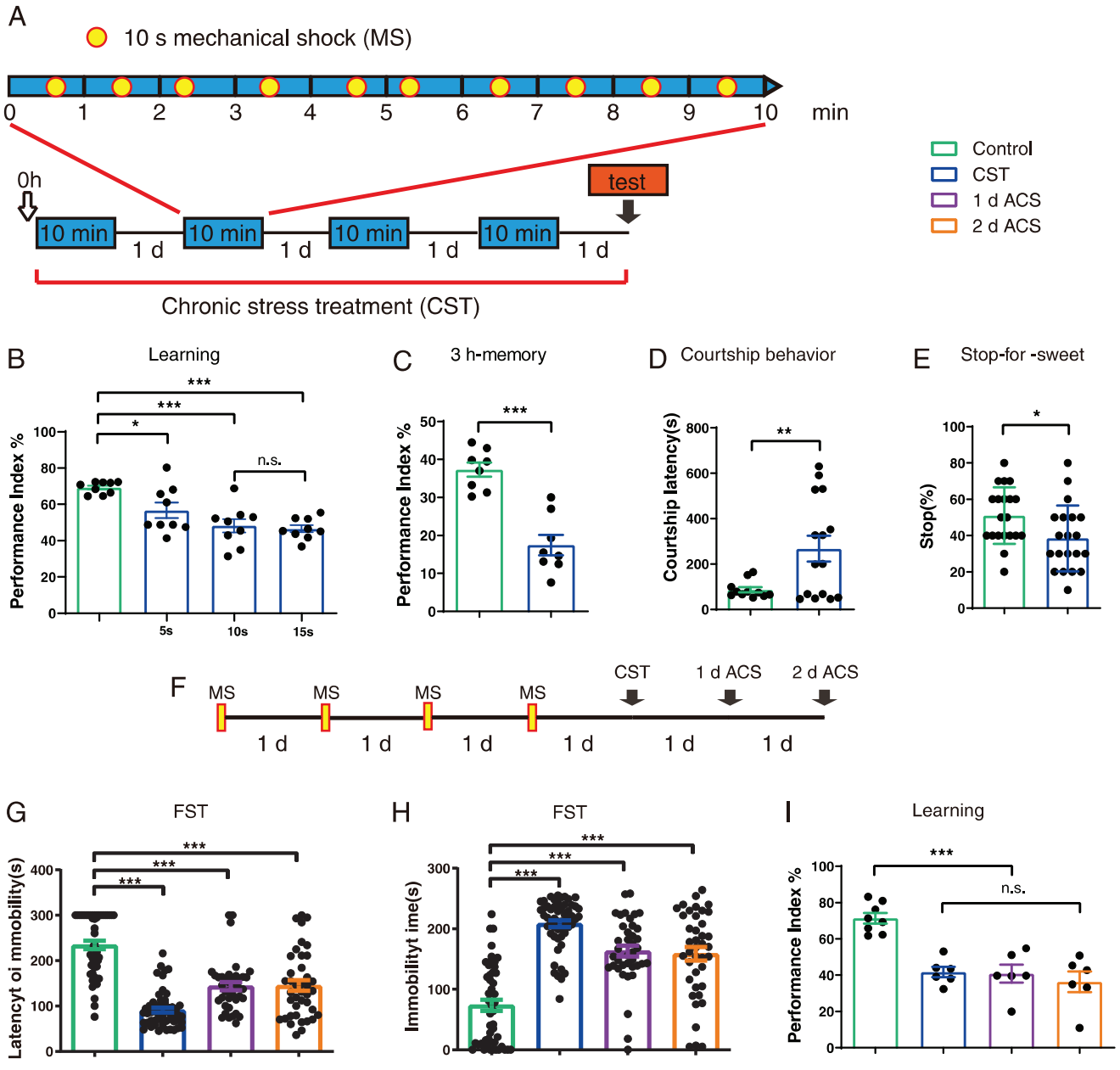


Fig. 1. Chronic stress induces olfactory learning and memory deficits as well as other depression-like behaviors in *Drosophila*. (*A*) Schematic of the standard CST procedure used in this study. During the 4-d CST, flies were cultured in small food vials and received 10 min of uncontrollable and unpredictable mechanical shock treatment each day. Behavioral assays were tested after treatment. (*B*) The effects of 4-d chronic stress with different vortex time on olfactory learning. Significant learning deficit was induced in each condition (one-way ANOVA, $F_{(3,\ 32)} = 11.75$, $P < 0.001$, $n = 9$; Tukey's test, for 5 s, $P < 0.05$; for 10 s, $P < 0.001$; for 15 s, $P < 0.001$). The learning deficit induced by 10-s vortex was not significantly different from that of 15-s vortex (Tukey's test, $P > 0.05$, $n = 9$). The 10-s vortex was chosen as the standard mechanical shock condition for the CSTs in the rest of this study unless otherwise noted. (*C*) Chronic stress induced significant 3-h memory deficit (*t* test, $P < 0.001$, $n = 8$). (*D*) The courtship latency was longer in chronically stressed males than controls (Welch corrected *t* test, $P < 0.01$, $n \geq 11$). (*E*) Chronically stressed flies tended to ignore the sweet substance in the stop-for-sweet assay (*t* test, $P < 0.05$, $n = 20$). (*F*) Schematic of the experimental design for *G–I*. In the FST, flies showed a long-lasting tendency of despair-like behavior, as indicated by shorter latency to immobility (*G*) and longer immobility time (*H*) ACS treatment, 1 d ACS, or 2 d ACS (for *G*, Kruskal–Wallis test, $H_{(3)} = 83.69$, $P < 0.001$; Dunn's test, $P < 0.001$; for *H*, Kruskal–Wallis test, $H_{(3)} = 76.86$, $P < 0.001$, $n \geq 39$ for all groups). (*I*) CSLD is long lasting, as the learning deficit could be detected not only after the 4-d CST but also at 1 d ACS or 2 d ACS (one-way ANOVA, $F_{(3,\ 22)} = 17.06$, $P < 0.001$, $n \geq 6$; Tukey's test, $P < 0.001$ for all groups). Moreover, the learning performances of these three time points are not significantly different from each other (Tukey's test, $P > 0.05$, $n = 6$ for all groups). Canton-S flies were used for courtship behavioral assays, while *w1118(isoCJ1)* flies were used for all the other experiments. Data are represented as mean ± SEM. The stars indicate significant differences (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$); n.s., not significant. MS stands for mechanical shock; CST, chronic stress treatment; ACS, after chronic stress; FST, forced swimming test.

indispensable for chronic stress to erode olfactory learning. To examine if chronic activation of DANs could precipitate the development of CSLD, we drove the expression of the heat-activated ion channel TrpA1 (*UAS-TrpA1*) with TH-Gal4. In the course of CST, the mechanical shock was replaced with daily 1 h high-temperature exposure (30 °C) to activate the DANs. After 4 d of treatment, the learning performance of *TH-Gal4/UAS-TrpA1* flies was significantly decreased, whereas shock reactivity and odor acuity appeared normal (*SI Appendix*, Fig. S6 *D–F*), and +/*UAS-TrpA1* controls were not affected

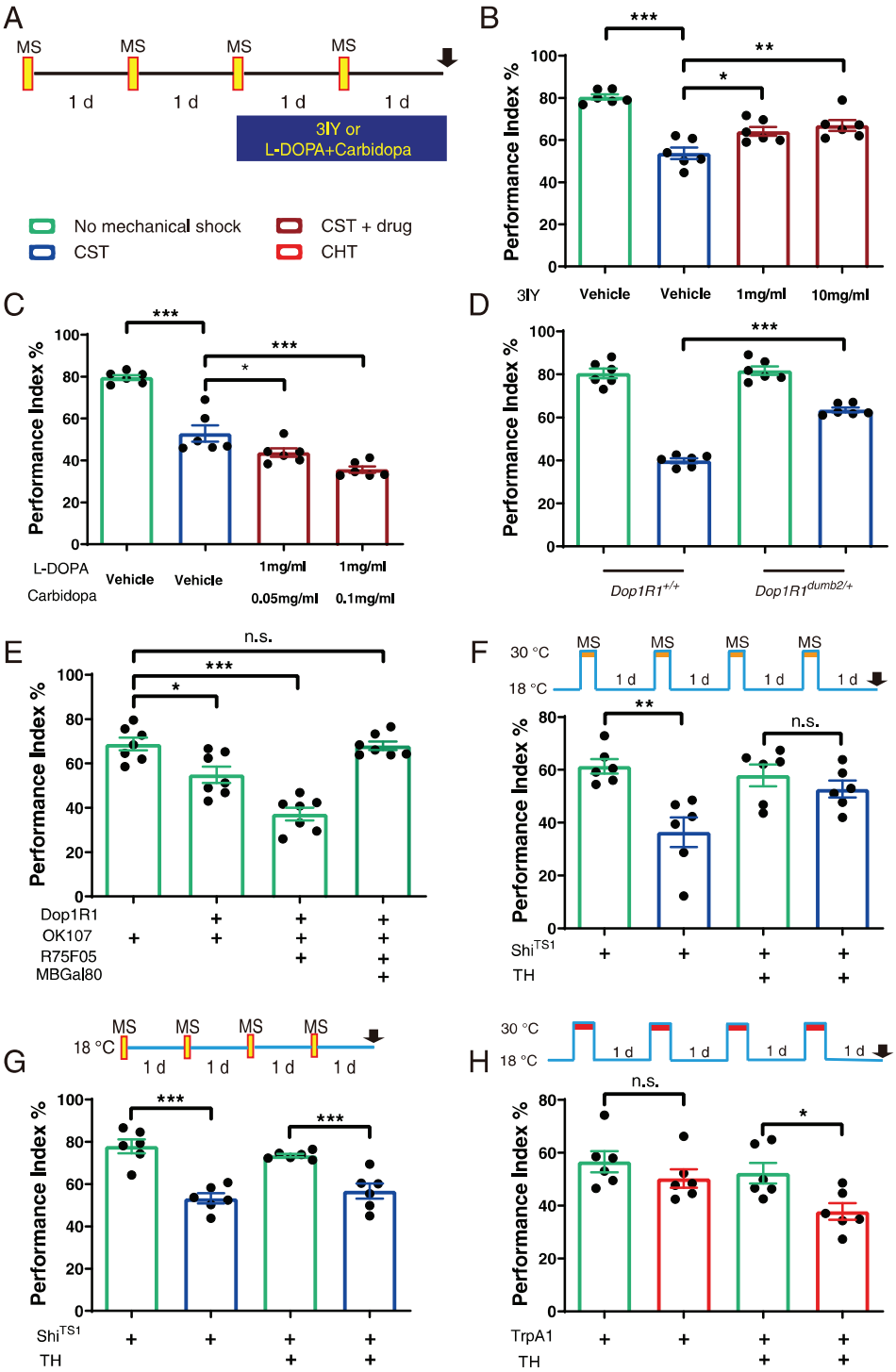


Fig. 2. Susceptibility to CSLD is modulated by dopaminergic system. (*A*) Schematic of the experimental design for *B* and *C*. (*B*) Feeding of DA synthesis inhibitor 3-IY in the last 2 d of the chronic stress procedure alleviated CSLD (one-way ANOVA, $F_{(3,\ 20)} = 23.34$, $P < 0.001$, $n = 6$; Dunnett's test, for 1 mg/mL, $P < 0.05$; for 10 mg/mL, $P < 0.01$). (*C*) Feeding of DA precursor L-DOPA plus carbidopa aggravated CSLD (one-way ANOVA, $F_{(3,\ 20)} = 63.78$, $P < 0.001$, $n = 6$; Dunnett's test, for 0.05 mg/mL carbidopa, $P < 0.05$; for 0.1 mg/mL carbidopa, $P < 0.001$). (*D*) *Dop1R1*$^{dumb2/+}$ flies showed a CSLD resistant phenotype (*t* test, $P < 0.001$, $n = 6$). (*E*) Overexpressing Dop1R1 with MB-specific Gal4 drivers in otherwise wild-type flies resulted in significant learning deficit (one-way ANOVA, $F_{(3,\ 24)} = 26.13$, $P < 0.001$, $n = 7$; Dunnett's test, for *OK107/UAS-Dop1R1*, $P < 0.05$; for *OK107, R75F05/UAS-Dop1R1*, $P < 0.001$), which can be rescued by adding an MB-Gal80 transgene (Dunnett's test, $P > 0.05$, $n = 7$). (*F*) Blocking synaptic release from TH-Gal4–labeled DANs during mechanical shock was sufficient to prevent CSLD (*t* test, $P > 0.05$, $n = 6$). (*G*) *TH/UAS-Shi*^{TS1} showed significant CSLD phenotype in permissive temperature (18 °C) (*t* test, $P < 0.001$, $n = 6$). (*H*) Replacing mechanical shock with thermogenetic activation of TH-labeled DANs by 1 h per day chronic heat treatment was sufficient to induce learning deficit (*t* test, $P < 0.05$, $n = 6$). Schematics that illustrate the experimental design are shown on the top of each panel. Data are represented as mean ± SEM. The stars indicate significant differences (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$); n.s., not significant. MS stands for mechanical shock; CST, chronic stress treatment; CHT, chronic heat treatment; CSLD, chronic stress–induced learning deficit.

(Fig. 2*H*). Thus, excessive DANs activity during chronic stress is sufficient to induce learning deficit. Taken together, our experimental data on three levels, including DA, DA receptor, and DANs activity, are all in agreement with the notion that stress-induced excess DAergic activity promotes susceptibility to CSLD.

PPL1-γ1pedc Neurons Are the Key DANs That Precipitate Susceptibility to CSLD. The adult fly brain contains ∼280 DANs that project to diverse brain regions (28). To pinpoint the DAN subtypes that are involve in CSLD, we focused on two DAN clusters, PPL1 and PAM. Both of them project to MBs and are important for learning and memory. We used the Shi^{TS1} thermogenetic approach to acutely block DAN subsets during the mechanical shock treatments. R58E02-Gal4 marks the majority of PAM neurons (29, 30). Blocking PAM neurons with *R58E02-Gal4/UAS-*Shi^{TS1} showed no significant impact on CSLD (Fig. 3*B*), suggesting that the activity of PAM neurons is dispensable for CSLD. MB065B split Gal4 drives expression in a subset of PPL1 cluster neurons that project to the vertical lobes of MBs, including PPL1-γ2α′1, PPL1-α′2α2, PPL1-α3, and PPL1-α′3 (31, 32). Blocking these neurons with *MB065B/UAS-*Shi^{TS1} failed to alleviate CSLD as well (Fig. 3*C*). However, blocking neurotransmission with a PPL1-γ1pedc–specific split Gal4 MB320C driving *UAS-*Shi^{TS1} prevented CSLD (Fig. 3*D*). Furthermore, mechanical shock treatments performed at permissive temperature (18 °C) did not have a preventive effect on CSLD (Fig. 3*E*). These data suggest that PPL1-γ1pedc activity is indispensable for the development of CSLD. Besides strong expression in PPL1-γ1pedc neurons, MB320C split Gal4 also weakly marks PPL1-α′2α2 neurons (31, 33). To corroborate the indispensability of PPL1-γ1pedc for the development of CSLD, we drove *UAS-*Shi^{TS1} with VT50733, another PPL1-γ1pedc–specific driver that does not express in PPL1-α′2α2 neurons (*SI Appendix*, Fig. S7), and found that blocking PPL1-γ1pedc neurotransmission during mechanical shock with *VT50733-Gal4/UAS-*Shi^{TS1} also prevented CSLD (Fig. 3*F*). Importantly, chronic stress effectively induced learning deficit if these flies were shocked at permissive temperature (18 °C) (Fig. 3*G*). These data thus suggested that stress-induced PPL1-γ1pedc hyperactivity facilitates the development of CSLD. We verified this hypothesis by showing that chronic thermogenetic activation of PPL1-γ1pedc with *VT50733-Gal4/UAS-TrpA1* was sufficient to induce significant olfactory learning deficit in the absence of mechanical shock (Fig. 3*H*). Note that here we used a long protocol in which flies received daily 1-h thermogenetic activation in the first 2 d of the 4-d treatment followed by a 2-d constant thermogenetic activation. To confirm this finding, we used two additional PPL1-γ1pedc split Gal4 drivers, MB320C or MB438B (31), to chronically activate the neurons with the long thermogenetic activation protocol and showed that each was sufficient to induce a significant learning deficit in the absence of mechanical shock (Fig. 3*I*). Importantly, chronic thermogenetic activation of these neurons did not perturb sensorimotor responses to the odors or footshock (*SI Appendix*, Fig. S8 *G–O*). Based on these data, we concluded that PPL1-γ1pedc neurons are the key DANs whose activity mediates stress signals to precipitate susceptibility to CSLD.

MBON-γ1pedc>α/β Neurons Modulate Susceptibility to CSLD. PPL1-γ1pedc neurons project to the γ1pedc compartments of MBs, while MBON-γ1pedc>α/β neurons are the output neurons of γ1pedc compartments that furnish feedforward inhibition to multiple MBONs. Notably, PPL1-γ1pedc neural activity could depress MBON-γ1pedc>α/β activity (33, 34). We thus investigated whether MBON-γ1pedc>α/β neurons are important for regulating susceptibility to CSLD. Since stress-induced PPL1-γ1pedc activity precipitates susceptibility to CSLD, we hypothesized that stress-induced inhibition of MBON-γ1pedc>α/β activity should promote the development of CSLD. To test this hypothesis, we acutely activated MBON-γ1pedc>α/β with *R83A12-Gal4/UAS-TrpA1* (30, 34). Indeed, activating MBON-γ1pedc>α/β neurons only during mechanical shock treatments prevented CSLD (Fig. 4*B*), while CSLD appeared normal under permissive condition (18 °C) (Fig. 4*C*). On the other hand, chronic inhibition of MBON-γ1pedc>α/β synaptic release with *R83A12-Gal4/UAS-*Shi^{TS1} in the absence of mechanical shock was sufficient to induce a significant learning deficit (Fig. 4*D*), while shock reactivity and odor acuity were not diminished (*SI Appendix*, Fig. S9). These data suggested that stress-induced MBON-γ1pedc>α/β inhibition facilitates susceptibility to CSLD. Together with the above finding that PPL1-γ1pedc mediates stress signals to precipitate susceptibility to CSLD, our data indicated that the PPL1-γ1pedc–MBON-γ1pedc>α/β axis might be an important pathway for modulating susceptibility to CSLD.

Chronic Stress Induces Abnormal Neural Activity in the MB Network. Since normal MB function is required for olfactory learning, we speculated that chronic stress might induce abnormal neuronal activities in the MB network. To investigate this, we first took advantage of the CaLexA (calcium-dependent nuclear import of LexA) system that had been previously used to monitor accumulative neuronal activity in *Drosophila* (*SI Appendix*, Fig. S10 *A* and *B*) (35). CaLexA is composed of a genetically encoded GFP reporter whose expression is controlled by the intracellular calcium-dependent nuclear import of a chimeric transcription factor, the nuclear factor of activated T cells (NFAT) (35). We drove CaLexA and UAS-myr-tdTomato (to normalize the GFP signal) expression with MB subtype-specific Gal4s and focused on fluorescence in the lobe region (30, 36, 37). With VT30604, we found significantly reduced GFP/RFP signal in α′/β′ lobes of the chronically stressed flies compared with no treatment control (Fig. 5 *A* and *F*), indicating reduced accumulative neuronal activity in this brain region. However, when CaLexA was expressed via c739, the GFP/RFP signal in α/β lobes was not significantly different between chronic stress and control (*SI Appendix*, Fig. S10 *C* and *D*), suggesting that the neuronal activity of α/β does not change ACS. We used R14H06-Gal4 to drive CaLexA expression in γ-lobes and noted that the basal expression level between male and female flies was considerably different (30) (Fig. 5 *B* and *C*). We thus examined the γ-lobe GFP/RFP signal in male and female flies separately. Interestingly, significant enhancement was noted in the chronically stressed males compared with control (Fig. 5 *B* and *G*), yet no significant difference was detected in females (Fig. 5 *C* and *H*), suggesting sexual dimorphism of chronic stress response in this brain region.

In our attempts to investigate the neural activity of MB DANs and MBON-γ1pedc>α/β neurons, we noted that driving CaLexA with TH or MB112C, an MBON-γ1pedc>α/β–specific split Gal4, resulted in unreliably weak green fluorescent expression (31). To examine the impact of chronic stress on the activity of these neurons, we turned to ANF-GFP, a rat atrial natriuretic factor (ANF) and GFP fusion transgene that had previously been used to monitor neuropeptidergic vesicle trafficking and chronic neuronal activity (38–40). As neuronal excitation induces the release of the ANF-GFP containing vesicles and decreases GFP level in the terminal, the reduction of green fluorescence signal denotes increasing neuronal activity. To verify that the ANF-GFP signal reports the activity of neurons, we used TH to drive both ANF-GFP and TrpA1 in DANs and focused on γ1pedc, the MB compartment that PPL1-γ1pedc projects to. In line with previous reports, thermogenetic activation of DANs led to a significant reduction of ANF-GFP signal

NEUROSCIENCE

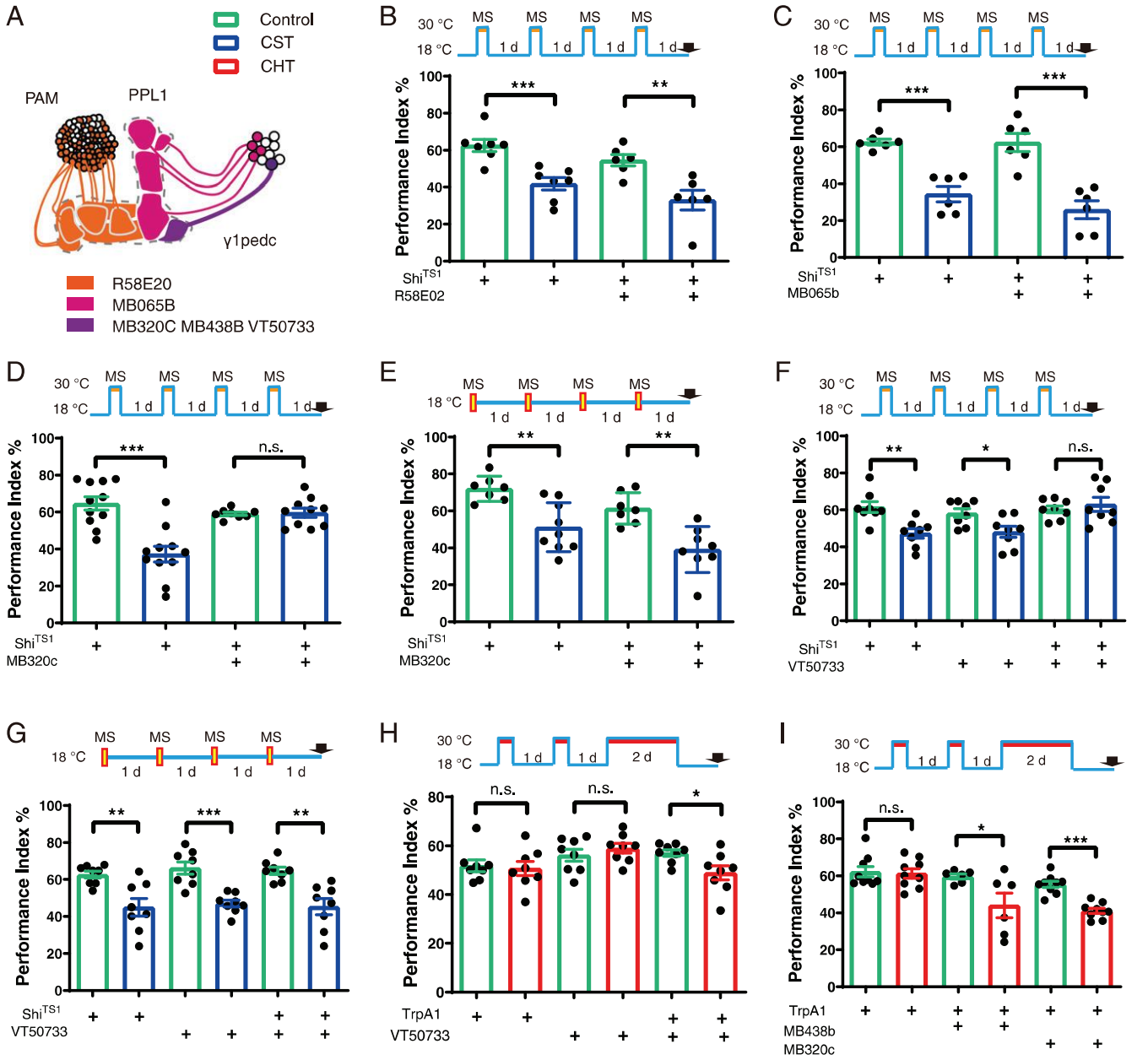


Fig. 3. The activity of PPL1-γ1pedc DANs provokes susceptibility to CSLD. (*A*) Schematic of expression patterns of Gal4 and Split-GAL4 drivers that each labels a subset of DANs that project to MB lobes. (*B*) Blocking synaptic release from PAM DANs with *R58E02-Gal4/UAS-Shi^{TS1}* could not prevent CSLD (*t* test, $P < 0.01$, $n = 6$). (*C*) Blocking synaptic release from PPL1-DANs that spare the γ1pedc subtype with *MB065B/UAS-Shi^{TS1}* could not prevent CSLD (*t* test, $P < 0.001$, $n = 6$). (*D*) In contrast, interrupting synaptic release from PPL1-γ1pedc DANs with *MB320C/UAS-Shi^{TS1}* only during mechanical shock treatments was sufficient to prevent CSLD (*t* test, $P > 0.05$, $n \geq 8$). (*E*) *MB320C/UAS-Shi^{TS1}* showed significant CSLD phenotype in permissive temperature (18 °C) (*t* test, $P < 0.01$, $n \geq 7$). (*F*) Interrupting synaptic release from PPL1-γ1pedc DANs with *VT50733/UAS-Shi^{TS1}* was also sufficient to prevent CSLD (*t* test, $P > 0.05$, $n \geq 8$). (*G*) *VT50733/UAS-Shi^{TS1}* showed significant CSLD phenotype in permissive temperature (18 °C) (*t* test, $P < 0.01$, $n = 8$). (*H*) Chronic activation of PPL1-γ1pedc with *VT50733/UAS-TrpA1* was sufficient to induce significant learning deficit in the absence of mechanical shock (*t* test, $P < 0.05$, $n = 8$). (*I*) Chronic activation of PPL1-γ1pedc with *MB438B/UAS-TrpA1* or *MB320C/UAS-TrpA1* was sufficient to induce significant learning deficit as well (*t* test, for *MB438B*, $P < 0.05$, $n = 6$; for *MB320C*, $P < 0.001$, $n = 9$). Data are represented as mean ± SEM. The stars indicate significant differences (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$); n.s., not significant. MS stands for mechanical shock; CST, chronic stress treatment; CHT, chronic heat treatment; CSLD, chronic stress–induced learning deficit.

(38) (*SI Appendix*, Fig. S11 *A* and *C*). We then drove the expression of UAS-ANF-GFP and UAS-myr-tdTomato (to normalize the GFP signal) with TH and focused on the MB lobe region. TH-labeled DANs project to multiple MB compartments, including γ1pedc, γ2α′1, α2α′2, α′3, and α3 (28, 31). We examined each compartment separately. There was no significant difference in GFP/RFP signal between the chronic stress and control in each of the five compartments (*SI Appendix*, Fig. S11 *B* and *D*). These data suggest that the accumulative activity of DANs is not affected ACS. To probe MBON-γ1pedc>α/β neurons, we drove UAS-ANF-GFP and UAS-myr-tdTomato with MB112C (31). The axon of MBON-γ1pedc>α/β projects to α- and β-lobes. We found significant GFP/RFP signal reduction in the MBON-γ1pedc>α/β neuronal processes that project to the horizontal lobe region in the chronic stress group compared with no treatment control (Fig. 5 *D* and *I*), which may reflect

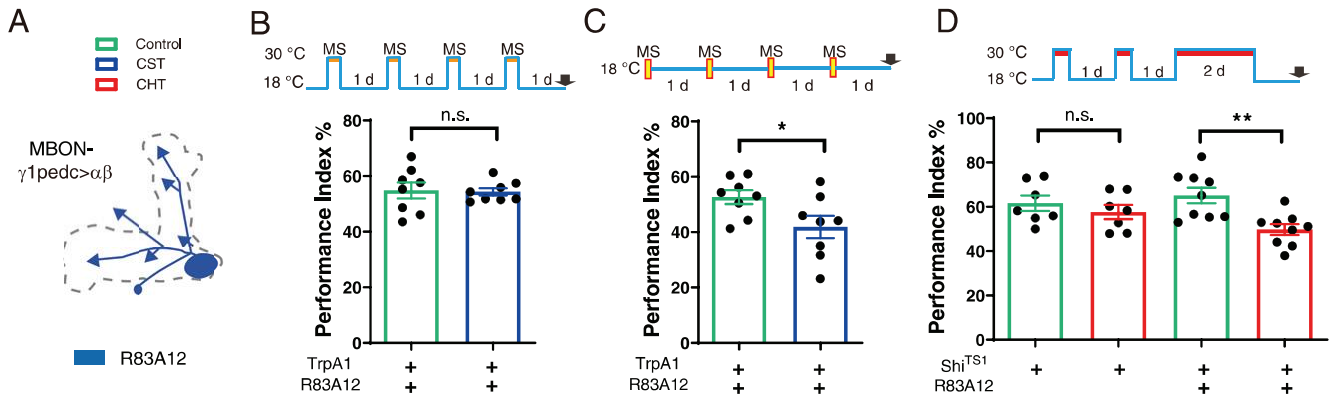


Fig. 4. MBON-γ1pedc>α/β neurons modulate susceptibility to CSLD. (*A*) Schematic of expression patterns of R83A12-Gal4 drivers that labels MBON-γ1pedc>α/β. (*B*) Thermogenetic activation of MBON-γ1pedc>α/β neurons only during the mechanical shock treatments prevented CSLD (*t* test, $P > 0.05$, $n = 8$). (*C*) The CSLD phenotype of *R83A12/UAS-TrpA1* under permissive condition (18 °C) was undiminished (*t* test, $P < 0.05$, $n = 8$). (*D*) Conversely, chronic interruption of MBON-γ1pedc>α/β synaptic transmission was sufficient to induce significant learning deficit in the absence of mechanical shock (*t* test, $P < 0.01$, $n = 9$). Data are represented as mean ± SEM. The stars indicate significant differences (*$P < 0.05$, **$P < 0.01$); n.s., not significant. MS stands for mechanical shock; CST, chronic stress treatment; CHT, chronic heat treatment; CSLD, chronic stress–induced learning deficit.

an increased accumulative synaptic release from the MBON-γ1pedc>α/β to the MBONs in that region (34, 41).

Taken together, we identified abnormal neural activities in the MB network after flies had undergone CST. The abnormalities were remarkable especially in α′/β′ KCs and MBON-γ1pedc>α/β neurons, while aberration was also detected in the γ-KCs of male flies. Notably, these neurons all play important roles in olfactory learning, as blocking α′/β′ KCs synaptic release, activating γ KCs, or activating MBON-γ1pedc>α/β each could impair 3-min memory performance (34, 42–44). Our data thus demonstrate that chronic stress induces maladaptations in the MB network that are detrimental to olfactory learning.

DAergic Activity Is Indispensable for the Development of Chronic Stress–Induced Abnormal Neural Activity. Given that DAergic activity promotes susceptibility to CSLD and chronic stress induces abnormal neural activity in learning-related neurons, we predicted that DAergic activity should be important for the development of chronic stress–induced abnormal neural activity. To test this hypothesis, we blocked the biosynthesis of DA by feeding flies with 3-IY in the last 2 d of the chronic stress procedure and probed α′/β′ lobes with VT30604 driving CaLexA. As shown in Fig. 5 *E* and *J*, down-regulating DA level during the process of CST prevented the chronic stress–induced reduction of GFP/RFP signal in α′/β′ lobes, suggesting that DAergic activity is indispensable for chronic stress to induce abnormal neural activity in α′/β′ KCs. These data, together with all the above findings, are consistent with a model that stress-induced excess DAergic activity promotes susceptibility to CSLD by driving maladaptations in the MB network that lead to learning deficit.

Discussion

Depression-like symptoms in *Drosophila* can be induced by a variety of approaches, such as genetic manipulation, drug feeding, and stress treatments (12, 13, 45–49). Notably, a recent report has articulated that chronic stress could induce a depression-like state in *Drosophila* (13). In the report, a 3-d chronic vibration stress protocol induces depression-like behaviors and decreased serotonin activity. Moreover, these symptoms can be relieved by feeding antidepressant 5-hydroxy-L-tryptophan. Consistently, a later study reports that depression-like behaviors can also be induced by a 10-d chronic unpredictable mild stress paradigm and reverted by feeding fluoxetine (12). These findings suggest that chronic stress–induced *Drosophila* depression-like models could have great validity (50). In the present study, with a 4-d chronic stress protocol, we showed that CST could induce substantial learning and memory impairment, a core depression-like symptom not addressed in previous *Drosophila* chronic stress studies. Besides learning and memory impairment, other depression-like symptoms, such as lack of motivation, anhedonia, and prone to despair were also evidenced. Consistent with the idea that a depression-like state was induced, the learning deficit and prone to despair phenotypes appeared to be long lasting. Furthermore, we showed that stress-induced supranormal DAergic activity is a key etiologic factor for the development of CSLD, which nicely mirrors the important roles of the mammalian DAergic system in the etiology and maintaining of depression symptoms (51–54). Our findings thus provide additional evidence to corroborate the idea that *Drosophila* chronic stress paradigm could be a valid depression-like animal model.

Drosophila DAergic systems respond to various stress stimuli, including mechanical shock and electric shock (9, 22, 55, 56). To demonstrate DAergic activity is indispensable for the development of CSLD, we provided three lines of evidence: pharmacological manipulation of DA synthesis, genetic intervention of Dop1R1 function, and thermogenetic manipulation of DAN activity. All the manipulations were bidirectional, and the data consistently suggested that excess DAergic activity precipitates susceptibility to CSLD. Importantly, blocking DAN synaptic release only during mechanical shock prevents CSLD. Conversely, chronic DAN activation without mechanical shock is sufficient to induce a learning deficit. Since DAergic activity could be induced by vibration (22), these findings indicate that DANs might play an anxiogenic-like role and mediate chronic stress signals to provoke the development of learning deficit presumably by inducing allostatic maladaptation in learning-related neural circuits.

MB is the olfactory memory center in the fly central brain, and two clusters of DANs, PPL1 and PAM, send presynaptic projects to MB lobes. Among these DANs, we identified a pair of PPL1-γ1pedc neurons as the key DANs that are indispensable for evoking CSLD. Consistent with the presumptive function of relaying chronic stress signals to drive allostatic adaptation in the MB network, PPL1-γ1pedc neurons are known to respond to external noxious stimuli (such as electric shock, heat, and bitter taste) and mediate the aversive US for writing associative memory in MBs (28, 57–60). Furthermore,

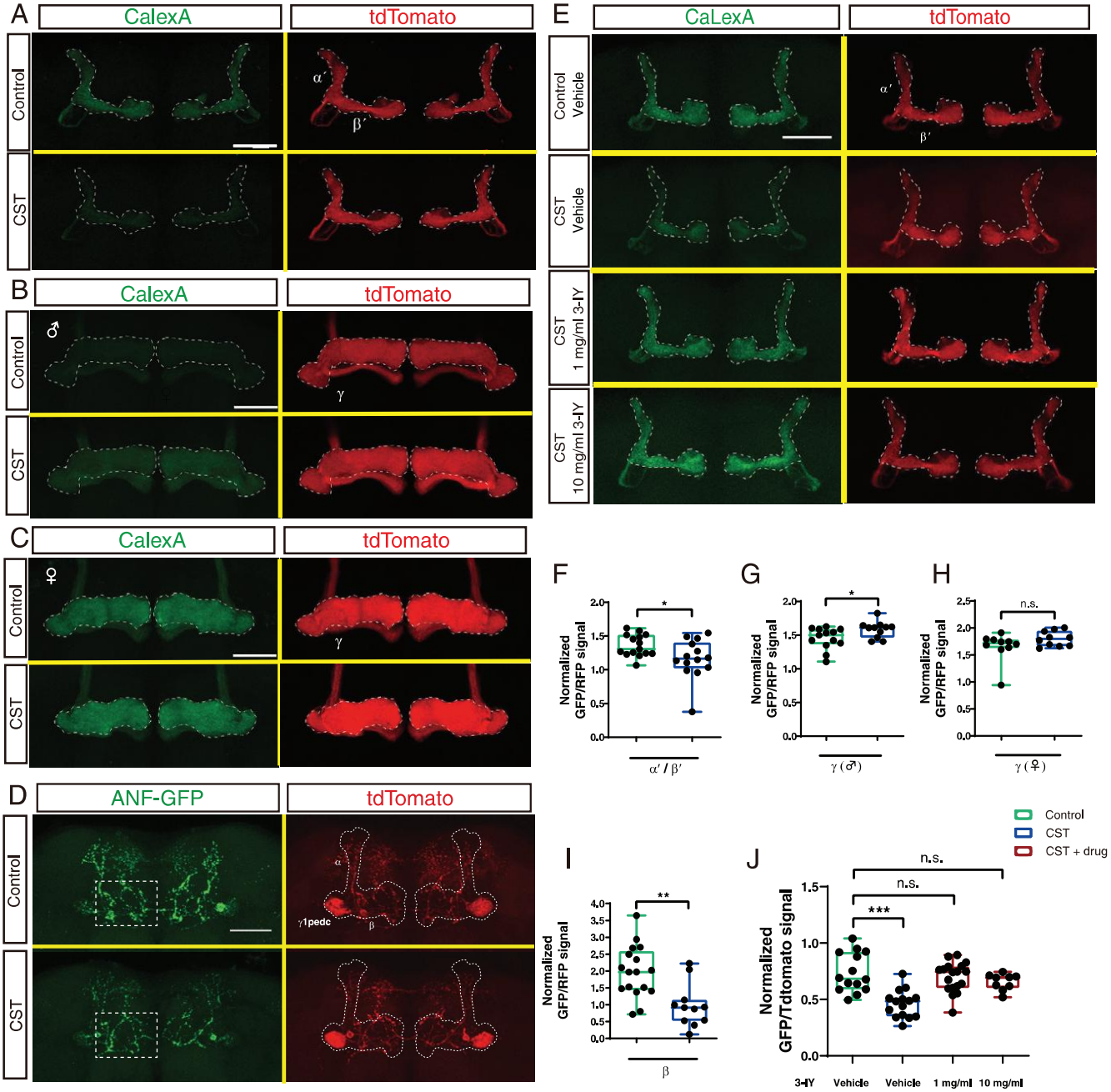


Fig. 5. Chronic stress induces abnormal neural activity in the MB network, which requires DAergic activity during CST. (*A*) Chronic stress induces the reduction of Ca^{2+} activity in α′/β′ lobes as measured with the CaLexA technique. Representative confocal images of CaLexA (*Left*) and tdTomato (*Right*) in the MB region of *VT30604-Gal4/CaLexA; UAS-myr-tdTomato* flies were shown. Dotted lines highlight MB α′/β′ lobes. (*B* and *C*) Chronic stress induces the increase of Ca^{2+} activity in the γ-lobes of male but not female flies as measured with CaLexA. Representative confocal images of CaLexA and tdTomato in the MB region of *R14H06-Gal4/CaLexA; UAS-myr-tdTomato* flies were shown in *B* for males and in *C* for females. Dotted lines highlight MB γ-lobes. (*D*) MBON-γ1pedc>α/β neurons are more active after the 4-d CST as indicated by diminished ANF-GFP signal in the β-lobe region. Representative confocal images of ANF-GFP (*Left*) and tdTomato (*Right*) in the MB area of *MB112C-Gal4/UAS-ANF-GFP; UAS-myr-tdTomato* flies were shown. Dotted lines highlight MB (*Right*) and β-lobe regions (*Left*). (*E*) Feeding DA synthesis inhibitor 3-IY in the last 2 d of the chronic stress procedure prevented chronic stress–induced Ca^{2+} activity reduction in α′/β′ lobes. Representative confocal images of CaLexA (*Left*) and tdTomato (*Right*) in the MB region of *VT30604-Gal4/CaLexA; UAS-myr-tdTomato* flies were shown. Dotted lines highlight MB α′/β′ lobes. (*F–J*) Quantitative analysis of normalized CaLexA fluorescence intensity in the region of α′/β′ lobes (*F*) (Welch corrected *t* test, $P < 0.05$, $n \geq 14$), γ-lobes (male) (*G*) (*t* test, $P < 0.05$, $n \geq 11$), and γ-lobes (female) (*H*) (Mann–Whitney *U* test, $P > 0.05$, $n = 10$). (*I*) Quantitative analysis of normalized ANF-GFP fluorescence intensity in the region of β-lobes (*t* test, $P < 0.01$, $n \geq 11$). (*J*) Quantitative analysis of the effect of 3-IY on the chronic stress–induced CaLexA signal reduction in the region of α′/β′ lobes (one-way ANOVA, $F(3, 52) = 12.52$, $P < 0.001$, $n \geq 9$; Dunnett's test, for CST, $P < 0.001$; for 1 mg/mL, $P > 0.05$; for 10 mg/mL, $P > 0.05$). In all confocal images, maximum projections of *Z*-stack sections with all the frames in the same parameter are shown. Scale bars in all images represent 50 μm unless otherwise noted. Box plots show the median (line inside the box), 25 and 75% quartiles (box), and minimum and maximum score (whiskers). The stars indicate significant differences (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$); n.s., not significant. CST stands for chronic stress treatment.

PPL1-γ1pedc neurons display sustained rhythmic activity even in the absence of external stimulus, which has been implicated in the internal state (satiation) and memory processing after acquisition (8, 61). In the present paper, our imaging studies identified abnormal accumulative neural activities in the MB network of the chronically stressed flies, including reduced neural activity in α′/β′ KCs and enhanced neural activities in γ-KCs (male) and MBON-γ1pedc>α/β neurons (Fig. 5 *A*, *B*, *D*, *F*, *G*, and *I*). It is interesting to note that sleep deprivation also increases the spontaneous activity of γd KCs and decreases that of α′/β′ KCs, indicating that these neurons are vulnerable points on the MB network (62). As previously reported, either blocking α′/β′ KCs synaptic release, activating γ-KCs, or activating MBON-γ1pedc>α/β results in learning deficiency (34, 42–44). Thus, the superimposed effects of these chronic stress–induced maladaptations may well underlie CSLD. Importantly, similar to CSLD, the chronic stress–induced maladaptation in α′/β′ KCs also requires excessive DAergic activity during CST (Fig. 5 *E* and *J*). Thus, our findings support the model that PPL1-γ1pedc activity mediates chronic stress signals to drive allostatic maladaptations in the MB network that lead to a learning deficit.

Although further investigations are needed to fully understand how excess PPL1-γ1pedc activity induces maladaptations in the MB network, anatomical and functional connectivity suggest that PPL1-γ1pedc could directly activate or inhibit several neural cell types in the γ1pedc compartment, such as α/β KCs, γ KCs, and MBON-γ1pedc>α/β (31, 33, 34, 41, 63). These neurons could potentially relay PPL1-γ1pedc activity to other regions of the MB network so that PPL1-γ1pedc could broadly impact the MB network. GABAergic MBON-γ1pedc>α/β neurons, for instance, not only feedback to γ1pedc compartment but also send feedforward inhibition over MBONs of other compartments, which allows MBON-γ1pedc>α/β neurons to exert widespread impact on the MB network (41, 64, 65). As an example, the release of MBON-γ1pedc>α/β feedforward inhibition over MBON-γ5β′2a is implicated in aversive short term memory retrieval, and the MBON-γ5β′2a disinhibition, in turn, might activate PAM-γ5 DANs to form extinction memory (34, 66). Considering that stress-induced MBON-γ1pedc>α/β inhibition also facilitates susceptibility to CSLD (Fig. 4), the PPL1-γ1pedc–MBON-γ1pedc>α/β axis might play an important role in mediating the chronic stress signal to drive allostatic maladaptations in multiple MB compartments, which needs to be validated in future studies. Interestingly, accumulated stress-related memories (memories of stressful events) are believed to contribute to major depressive disorder (67). Since the PPL1-γ1pedc–MBON-γ1pedc>α/β axis is functionally important for memory formation (8, 57, 58, 64), modulations of the PPL1-γ1pedc–MBON-γ1pedc>α/β axis might have prevented CSLD by interrupting the vortex stress–related memory. Although this speculation still waits for a formal demonstration, it raises the possibility that the potential aversive vortex-odor associative memory formed during CST, if any, might interfere with later olfactory conditioning and contribute to CSLD. Testing whether DA signaling is also important for developing other chronic stress–induced depressive-like behavior would help to resolve this issue. We thus examined the forced swimming phenotype after overexpressing Dop1R1 with MB-specific GAL-4 drivers. As shown in *SI Appendix*, Fig. S13, Dop1R1 hyperactivity in MBs did not affect latency to immobility and total immobility time in the FST, suggesting independent mechanisms are involved.

The gross DA level in the fly head is remarkably downregulated following chronic unpredictable mild stress (12). Similarly, our 4-d CST also leads to diminished DA concentration in the fly head (*SI Appendix*, Fig. S12), indicating that DAergic activity might be attenuated ACS. Consistent with this idea, imaging studies indicate that the inhibitory feedback to PPL1-γ1pedc from GABAergic MBON-γ1pedc>α/β neurons might be enhanced ACS treatment (64) (Fig. 5 *D* and *I*). Therefore, it is surprising that no change of accumulative spontaneous activity was detected in the PPL1 neurons that project to MB. Interestingly, in rodents, the effects of chronic stress on the ventral tegmental area (VTA) DAN activity depend on chronic stress protocols, even though all the protocols are capable to induce depression-like symptoms (68). Generally, CSTs using relatively mild stressors, such as cold or mild food shock, tend to induce long-lasting attenuated spontaneous VTA DAN activity, while those using more severe stressors, such as social defeat, are prone to induce enhanced spontaneous VTA DAN activity (51, 69, 70). These indicate that complex mechanisms underlie depression-associated maladaptations in DANs. We, therefore, speculate that ANF-GFP might not be sensitive enough to detect the chronic stress–induced change in DANs, or maladaptive response of the DAergic system might involve mechanisms other than DAN spontaneous activity in *Drosophila*.

Drosophila genome encodes 4 DA receptors, Dop1R1, Dop1R2, Dop2R, and DopEcR. Dop1R1 is highly expressed in MB and indispensable for aversive olfactory learning and memory (25, 71). A previous study has suggested an "inverted-U" dosage response of Dop1R1 level on long term memory (72). In the current study, we show that a similar Dop1R1 "inverted-U" dosage relationship also applies to learning, as Dop1R1 overexpression in MB KCs results in olfactory learning impairment (Fig. 2*E*). Importantly, down-regulating Dop1R1 expression with $Dop1R1^{dumb2/+}$ alleviates CSLD (Fig. 2*D*), suggesting that excess Dop1R1 signaling in KCs is a key determinant of susceptibility to CSLD. A similar inverted U–shaped dose–response relationship between DA receptor D1 (D1R) and working memory is well known in the mammalian prefrontal cortex (73). And stress-induced supranormal D1R activity could markedly impair working memory (74, 75). Furthermore, in the mammalian amygdala, DA also mediates anxiety via D1R and D2R (76). These similarities suggest that the anxiogenic effects of D1R activity might have evolutionarily conserved mechanisms. Whether Dop1R2, Dop2R, and DopEcR are involved in CSLD remains an open question.

Collectively, in this study, we establish a *Drosophila* model for studying the impact of chronic stress on learning and memory. Our investigation of the etiology of CSLD has demonstrated that the DAergic system plays essential roles in modulating susceptibility. Especially, we identified a single pair of DANs, PPL1-γ1pedc, that mediate stress signals to induce allostatic overload of the MB network that results in abnormal neural activities and learning deficits. These suggest that with abundant genetic tools and complete connectome, the *Drosophila* model can provide a unique opportunity to study conserved signaling pathway mechanisms underlying chronic stress–induced cognitive impairments in a learning and memory center whose anatomy is well-annotated.

Materials and Methods

Fly Stocks. Flies were cultured in cornmeal fly food at 23 °C. The following fly stocks were used in this study: *w1118(isoCJ1)*, Canton-S, NP2758, Elav, TH, OK107, c739, *UAS-myr::GFP*, *UAS-Shibire^{TS1}*, and *UAS-TrpA1*. The following FlyLight Split-GAL4 lines were shared by Yi Zhong, Tsinghua University, Beijing: MB065B, MB320C, MB438B, MB131B, and MB112C (31). *Dop1R1^{dumb2}*, *UAS-Dop1R1*, and MBGal80 were gifts from Josh Dubnau, Stony Brook University, New York. CaLexA (*LexAop-CD8-GFP-2A-CD8-GFP; UAS-mLexA-VP16-NFAT, lexAop-rCD2-GFP*) was from Zhefeng Gong, Zhejiang University, Hangzhou. *UAS-ANF-GFP*, *UAS-myr-tdTomato,* R58E02, R75F05, R83A12, and R14H06 were ordered from Bloomington *Drosophila* Stock Center (BDSC). VT30604, VT49246, and VT50733 were ordered from Vienna *Drosophila* Resource Center. *w1118(isoCJ1)*, a white isogenic line derived from Canton-S, was used as the wild-type control unless otherwise noted. *Dop1R1^{dumb2}*, *UAS-Dop1R1*, MBGal80, *UAS-Shibire^{TS1}*, *UAS-TrpA1*, Elav, TH, OK107, and C739 strains had been equilibrated to *w1118(isoCJ1)* for at least five generations before being utilized.

CST. Groups of about 100 flies were collected and cultured in small food vials (restricted in a space of 25 mm in diameter and about 55 mm in height). During the CST, flies received a 10-min mechanical shock every day. Before the mechanical shock, flies were transferred into empty 10-mL centrifuge tubes (16 mm in diameter). The tubes were put on the perforated foam cushion installed on a vortex shaker (IKA GENIUS 3). Vortexes of 500 rpm were applied for 10 s in each minute of the 10-min treatment. To avoid habituation, vortexes were started randomly within the minute. After the 10-min vortex, flies were transferred back to the food vials. The same treatments were repeated for 4 consecutive days. Behavioral assays were tested at the end of the 4-d treatment (1 d after the last vortex).

Behavior Assays. All behavioral experiments were carried out in an environmental chamber with 70% humidity. The temperature was set at 25 °C except for Shibire^{TS1} and TrpA1 experiments.

For aversive olfactory learning and memory, flies were trained to associate odors with electric shock in a T-maze apparatus with a Pavlovian conditioning paradigm as previously described (15, 16, 71). Each individual *n* consisted of ~200 flies, with half of the flies trained to one odor and half to the other odor. For learning, flies were trained with a single training session and tested immediately after training. For 3-h memory, flies were transferred into food vials after training and kept in dark for 3 h before testing. For odor acuity, odor avoidance to 3-octanol (OCT, Sigma-Aldrich, product no. 218405) or 4-methylcyclohexanol (MCH; Sigma-Aldrich, product no. 218405) was quantified as previously described (15, 16, 77). For shock reactivity, shock avoidance to 60 V was quantified as previously described (15, 16, 77).

Male courtship assay was adapted from a previous report (20). Briefly, 6- to 8-d-old stressed or control virgin flies were cold anesthetized and individually loaded into two-layer round chambers (diameter: 1cm; height: 2.5 mm per layer, a gift from Yufeng Pan, Southeast University, Nanjing). To let flies adapt to the chamber environment, males were separated from target females by a transparent plastic barrier for 1 h before the courtship test. Courtship tests were videoed with a camcorder (Canon HF R806). The time lag to the first courtship display by the male after pairing with the female was recorded.

FST was adapted from a previous report (21). Briefly, a single fly was gently aspirated into a chamber (35 × 10 mm) filled with 2.5 mm height of 0.08% sodium dodecyl sulfate solution (Sangon Biotech, product no. A600485). Each fly was videotaped for 5 min (Canon HF R806). Latency to the first period of immobility and duration of immobility were recorded for final analyses.

Stop-for-sweet was tested according to previous description (13). In brief, qualitative filter papers (medium speed) were cut into rectangles of 55 × 20 mm. Along the midline of the paper, 5-mm-wide traces of glycerol (Sinopharm, product no. 10010618) were painted. The papers were held upright in a chamber with an angle of 110° to 120°. Stressed or no-stressed control flies, with their wings cut before test, were allowed to climb up the filter papers and stop to eat when they walked by the glycerol. Each fly was tested 10 times, and the number of stops was scored.

Confocal Microscopy. Brain samples were prepared as previously described (71). Adult fly brains were dissected in cold phosphate buffer saline (PBS) and then fixed in 4% paraformaldehyde for 20 min. Fixed brains were transferred into phosphate buffer saline with 0.5% Triton X-100 (PBST) and vacuum twice in a vacuum desiccator, 10 min each. Brains were mounted in VECTASHIELD mounting medium (Vector Laboratories, Inc.). The confocal images were acquired with an Olympus FV1200 or a Nikon TI-E+A1 SI confocal microscope. The image data were processed with ImageJ or Fiji ImageJ and later manipulated as figures in Adobe Photoshop and Illustrator (Adobe Systems Incorporated). To measure fluorescent intensities, we manually selected the region of interest (ROI) and measured the total intensity. To correct for background, background fluorescent intensity from the adjacent area of the ROI was measured and subtracted. To compare CaLexA or ANF-GFP signals between individual flies, the GFP signals were normalized to the tdTomato signals from the same ROIs: $F = (F_{gfp} - F_{background\text{-}gfp})/(F_{tdtomato}\text{-}F_{background\text{-}tdtomato})$.

Pharmacology. 3-IY (I8250, Sigma), L-DOPA (D9628, Sigma), or S-(−)-Carbidopa (C1335, Sigma) were dissolved in 5% sucrose plus 2% yeast solution. For drug feeding, adult flies were transferred to food vials containing a tissue (4 × 4 cm) soaked with 2 mL of the sucrose solution. Control flies were fed with vehicle (5% sucrose plus 2% yeast solution) for the same amount of time.

TUNEL Staining. Dissected fly brains were treated with 20 μg/L proteinase K (Sangon Biotech, Shanghai) in PBS buffer for 25 min. After washing three times with PBST, terminal deoxynucleotidyl transferase–mediated deoxyuridine triphosphate nick end-labeling (TUNEL) staining was done by following the instruction from the manufacturer (In Situ Cell Death Detection Kit, Fluorescein, Roche).

Smurf Assay. Smurf assay was adapted from a previous report (78). Flies were maintained on standard corn meal medium containing 2.5% brilliant blue dye (Shanghai Dyestyffs Research Institute Co., Ltd) for 1 d. Smurf phenotype was determined if the blue color was observed outside of the digestive tract.

Enzyme-Linked Immunosorbent Assay. For each independent experiment, the heads of 200 flies were collected and homogenized in PBS (pH 7.2). The samples were then centrifuged at 10,000 *g* for 20 min at 4 °C, and supernatants were collected. Enzyme-linked immunosorbent assays (ELISAs) were done following the instruction from the manufacturer (Cloud-Clone Corp., product no. CEA851Ge). A multimode plate reader (PerkinElmer EnSpire) was used to detect the optical intensities.

Statistics. Statistical analysis was performed with GraphPad Prism 9 (GraphPad Software). According to the central limit theorem, the performance indices (PIs) of the T-maze assays (the average of two half PIs) should be normally distributed (79, 80). For the rest of the data in this study, the normality of data were determined with the D'Agostino and Pearson test. For data sets that are normally distributed, Student's *t* test or Welch corrected *t* test was used for comparisons between two groups, and one-way ANOVA followed by Tukey's test or Dunnett's test was used for comparisons of multiple groups. For data sets that are not normally distributed, Mann–Whitney *U* test was used for comparisons between two groups, and Kruskal–Wallis test followed by Dunn's test was used for comparisons of multiple groups. The threshold for statistical significance was set at $P < 0.05$. In all bar graphs, data are presented as means ± SEM. Box plots show the median (line inside the box), 25 and 75% quartiles (box), minimum and maximum score (whiskers).

Data Availability. All study data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. We thank Wanhe Li, Yan Li, Yongjun Tan, Zhefeng Gong, and members of H.Q.'s laboratory for stimulating discussions. We are grateful for the flies from Yi Zhong, Yi Rao, Josh Dubnau, Wanhe Li, Lisha Shao, Cheng Huang, Yufeng Pan, and Yulong Li as well as the BDSC, Shanghai Institute of Biochemistry and Cell Biology Core Facility of Drosophila Resource and Technology, and Tsinghua Fly Center. This work was financially supported by National Natural Science Foundation of China (31371069), National Natural Science Foundation of China (81860641), Natural Science Foundation of Hunan Province of China (2018JJ2035), The Key Project of Natural Science Foundation of Yunnan Province, China (202001AS070035), Hunan University Interdisciplinary Research Project (2015-039), and Yunnan Key Laboratory of Pharmacology for Natural Products Open Research Fund (2019G009).

1. B. S. McEwen *et al.*, Mechanisms of stress in the brain. *Nat. Neurosci.* **18**, 1353–1363 (2015).
2. M. F. Marin *et al.*, Chronic stress, cognitive functioning and mental health. *Neurobiol. Learn. Mem.* **96**, 583–595 (2011).
3. G. Fink, *Stress: Concepts, Cognition, Emotion, and Behavior, Handbook of Stress* (Academic Press, Boston, 2016).
4. C. D. Conrad, A critical review of chronic stress effects on spatial learning and memory. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **34**, 742–755 (2010).
5. B. Roozendaal, B. S. McEwen, S. Chattarji, Stress, memory and the amygdala. *Nat. Rev. Neurosci.* **10**, 423–433 (2009).
6. B. S. McEwen, H. Akil, Revisiting the stress concept: Implications for affective disorders. *J. Neurosci.* **40**, 12–21 (2020).
7. Y. Hirano *et al.*, Fasting launches CRTC to facilitate long-term memory formation in Drosophila. *Science* **339**, 443–446 (2013).
8. P. Y. Plaçais, T. Preat, To favor survival under food shortage, the brain disables costly memory. *Science* **339**, 440–442 (2013).
9. J. A. Berry, I. Cervantes-Sandoval, M. Chakraborty, R. L. Davis, Sleep facilitates memory by blocking dopamine neuron-mediated forgetting. *Cell* **161**, 1656–1667 (2015).
10. X. Li, F. Yu, A. Guo, Sleep deprivation specifically impairs short-term olfactory memory in Drosophila. *Sleep* **32**, 1417–1424 (2009).
11. L. Seugnet, Y. Suzuki, L. Vine, L. Gottschalk, P. J. Shaw, D1 receptor activation in the mushroom bodies rescues sleep-loss-induced learning impairments in Drosophila. *Curr. Biol.* **18**, 1110–1117 (2008).
12. S. M. Araujo *et al.*, Chronic unpredictable mild stress-induced depressive-like behavior and dysregulation of brain levels of biogenic amines in Drosophila melanogaster. *Behav. Brain Res.* **351**, 104–113 (2018).
13. A. S. Ries, T. Hermanns, B. Poeck, R. Strauss, Serotonin modulates a depression-like state in Drosophila responsive to lithium treatment. *Nat. Commun.* **8**, 15738 (2017).
14. S. Lall *et al.*, Adult crowding induces sexual dimorphism in chronic stress-response in *Drosophila melanogaster. bioRxiv* [Preprint] (2019). https://doi.org/10.1101/702357 (Accessed 14 August 2019).

15. T. Tully, W. G. Quinn, Classical conditioning and retention in normal and mutant Drosophila melanogaster. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **157**, 263–277 (1985).
16. T. Tully, T. Preat, S. C. Boynton, M. Del Vecchio, Genetic dissection of consolidated memory in Drosophila. *Cell* **79**, 35–47 (1994).
17. A. Barekat *et al.*, Using Drosophila as an integrated model to study mild repetitive traumatic brain injury. *Sci. Rep.* **6**, 25252 (2016).
18. R. J. Katzenberger *et al.*, Death following traumatic brain injury in Drosophila is associated with intestinal barrier dysfunction. *eLife* **4**, e04790 (2015).
19. R. J. Katzenberger *et al.*, A Drosophila model of closed head traumatic brain injury. *Proc. Natl. Acad. Sci. U.S.A.* **110**, E4152–E4159 (2013).
20. D. Chen *et al.*, Genetic and neuronal mechanisms governing the sex-specific interaction between sleep and sexual behaviors in Drosophila. *Nat. Commun.* **8**, 154 (2017).
21. W. S. Neckameyer, A. R. Nieto-Romero, Response to stress in Drosophila is mediated by gender, age and stress paradigm. *Stress* **18**, 254–266 (2015).
22. W. S. Neckameyer, J. S. Weinstein, Stress affects dopaminergic signaling pathways in Drosophila melanogaster. *Stress* **8**, 117–131 (2005).
23. S. Waddell, Dopamine reveals neural circuit mechanisms of fly memory. *Trends Neurosci.* **33**, 457–464 (2010).
24. K. Cichewicz *et al.*, A new brain dopamine-deficient Drosophila and its pharmacological and genetic rescue. *Genes Brain Behav.* **16**, 394–403 (2017).
25. Y. C. Kim, H. G. Lee, K. A. Han, D1 dopamine receptor dDA1 is required in the mushroom body neurons for aversive and appetitive learning in Drosophila. *J. Neurosci.* **27**, 7640–7647 (2007).
26. T. Lebestky *et al.*, Two different forms of arousal in Drosophila are oppositely regulated by the dopamine D1 receptor ortholog DopR via distinct neural circuits. *Neuron* **64**, 522–536 (2009).
27. T. Kitamoto, Conditional modification of behavior in Drosophila by targeted expression of a temperature-sensitive shibire allele in defined neurons. *J. Neurobiol.* **47**, 81–92 (2001).
28. Z. Mao, R. L. Davis, Eight different types of dopaminergic neurons innervate the Drosophila mushroom body neuropil: Anatomical and physiological heterogeneity. *Front. Neural Circuits* **3**, 5 (2009).
29. C. Liu *et al.*, A subset of dopamine neurons signals reward for odour memory in Drosophila. *Nature* **488**, 512–516 (2012).
30. A. Jenett *et al.*, A GAL4-driver line resource for Drosophila neurobiology. *Cell Rep.* **2**, 991–1001 (2012).
31. Y. Aso *et al.*, The neuronal architecture of the mushroom body provides a logic for associative learning. *eLife* **3**, e04577 (2014).
32. P. Masek, K. Worden, Y. Aso, G. M. Rubin, A. C. Keene, A dopamine-modulated neural circuit regulating aversive taste memory in Drosophila. *Curr. Biol.* **25**, 1535–1541 (2015).
33. T. Hige, Y. Aso, M. N. Modi, G. M. Rubin, G. C. Turner, Heterosynaptic plasticity underlies aversive olfactory learning in Drosophila. *Neuron* **88**, 985–998 (2015).
34. E. Perisse *et al.*, Aversive learning and appetitive motivation toggle feed-forward inhibition in the Drosophila mushroom body. *Neuron* **90**, 1086–1099 (2016).
35. K. Masuyama, Y. Zhang, Y. Rao, J. W. Wang, Mapping neural circuits with activity-dependent nuclear import of a transcription factor. *J. Neurogenet.* **26**, 89–102 (2012).
36. P. T. Lee *et al.*, Serotonin-mushroom body circuit modulating the formation of anesthesia-resistant memory in Drosophila. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 13794–13799 (2011).
37. M. Y. Yang, J. D. Armstrong, I. Vilinsky, N. J. Strausfeld, K. Kaiser, Subdivision of the Drosophila mushroom bodies by enhancer-trap expression patterns. *Neuron* **15**, 45–54 (1995).
38. Q. Liu, S. Liu, L. Kodama, M. R. Driscoll, M. N. Wu, Two dopaminergic neurons signal to the dorsal fan-shaped body to promote wakefulness in Drosophila. *Curr. Biol.* **22**, 2114–2123 (2012).
39. S. Rao, C. Lang, E. S. Levitan, D. L. Deitcher, Visualization of neuropeptide expression, transport, and exocytosis in Drosophila melanogaster. *J. Neurobiol.* **49**, 159–172 (2001).
40. D. Shakiryanova, A. Tully, R. S. Hewes, D. L. Deitcher, E. S. Levitan, Activity-dependent liberation of synaptic neuropeptide vesicles. *Nat. Neurosci.* **8**, 173–178 (2005).
41. S. Y. Takemura *et al.*, A connectome of a learning and memory center in the adult *Drosophila* brain. *eLife* **6**, e26975 (2017).
42. S. Zhang, G. Roman, Presynaptic inhibition of gamma lobe neurons is required for olfactory learning in Drosophila. *Curr. Biol.* **23**, 2519–2527 (2013).
43. Y. Wang, A. Mamiya, A. S. Chiang, Y. Zhong, Imaging of an early memory trace in the Drosophila mushroom body. *J. Neurosci.* **28**, 4368–4376 (2008).
44. Y. Ueoka, M. Hiroi, T. Abe, T. Tabata, Suppression of a single pair of mushroom body output neurons in *Drosophila* triggers aversive associations. *FEBS Open Bio* **7**, 562–576 (2017).
45. M. D. Jiang, Y. Zheng, J. L. Wang, Y. F. Wang, Drug induces depression-like phenotypes and alters gene expression profiles in Drosophila. *Brain Res. Bull.* **132**, 222–231 (2017).
46. G. E. Brown, A. L. Mitchell, A. M. Peercy, C. L. Robertson, Learned helplessness in Drosophila melanogaster? *Psychol. Rep.* **78**, 962 (1996).
47. H. O. Lawal *et al.*, Drosophila modifier screens to identify novel neuropsychiatric drugs including aminergic agents for the possible treatment of Parkinson's disease and depression. *Mol. Psychiatry* **19**, 235–242 (2014).
48. A. P. Mendes-Silva *et al.*, Brain-enriched MicroRNA-184 is downregulated in older adults with major depressive disorder: A translational study. *J. Psychiatr. Res.* **111**, 110–120 (2019).
49. Z. Yang, F. Bertolucci, R. Wolf, M. Heisenberg, Flies cope with uncontrollable stress by learned helplessness. *Curr. Biol.* **23**, 799–803 (2013).
50. P. Willner, The validity of animal models of depression. *Psychopharmacology (Berl.)* **83**, 1–16 (1984).
51. D. Chaudhury, H. Liu, M. H. Han, Neuronal correlates of depression. *Cell. Mol. Life Sci.* **72**, 4825–4848 (2015).
52. P. Belujon, A. A. Grace, Dopamine system dysregulation in major depressive disorders. *Int. J. Neuropsychopharmacol.* **20**, 1036–1046 (2017).
53. T. F. Marton, V. S. Sohal, Of mice, men, and microbial opsins: How optogenetics can help hone mouse models of mental illness. *Biol. Psychiatry* **79**, 47–52 (2016).
54. S. Lammel, K. M. Tye, M. R. Warden, Progress in understanding mood disorders: Optogenetic dissection of neural circuits. *Genes Brain Behav.* **13**, 38–51 (2014).
55. T. Riemensperger, T. Völler, P. Stock, E. Buchner, A. Fiala, Punishment prediction by dopaminergic neurons in Drosophila. *Curr. Biol.* **15**, 1953–1960 (2005).
56. C. Huang *et al.*, Long-term optical brain imaging in live adult fruit flies. *Nat. Commun.* **9**, 872 (2018).
57. Y. Aso *et al.*, Specific dopaminergic neurons for the formation of labile aversive memory. *Curr. Biol.* **20**, 1445–1451 (2010).
58. Y. Aso, G. M. Rubin, Dopaminergic neurons write and update memories with cell-type-specific rules. *eLife* **5**, e16135 (2016).
59. D. S. Galili *et al.*, Converging circuits mediate temperature and shock aversive olfactory conditioning in Drosophila. *Curr. Biol.* **24**, 1712–1722 (2014).
60. G. Das *et al.*, Drosophila learn opposing components of a compound food stimulus. *Curr. Biol.* **24**, 1723–1730 (2014).
61. P. Y. Musso, P. Tchenio, T. Preat, Delayed dopamine signaling of energy level builds appetitive long-term memory in Drosophila. *Cell Rep.* **10**, 1023–1031 (2015).
62. D. Sitaraman *et al.*, Propagation of homeostatic sleep signals by segregated synaptic microcircuits of the drosophila mushroom body. *Curr. Biol.* **25**, 2915–2927 (2015).
63. C. H. Tsao, C. C. Chen, C. H. Lin, H. Y. Yang, S. Lin, *Drosophila* mushroom bodies integrate hunger and satiety signals to control innate food-seeking behavior. *eLife* **7**, e35264 (2018).
64. A. Pavlowsky, J. Schor, P. Y. Placais, T. Preat, A GABAergic feedback shapes dopaminergic input on the drosophila mushroom body to promote appetitive long-term memory. *Curr. Biol.* **28**, 1783–1793.e4 (2018).
65. Y. Aso *et al.*, Mushroom body output neurons encode valence and guide memory-based action selection in Drosophila. *eLife* **3**, e04580 (2014).
66. J. Felsenberg *et al.*, Integration of parallel opposing memories underlies memory extinction. *Cell* **175**, 709–722.e15 (2018).
67. G. Richter-Levin, L. Xu, How could stress lead to major depressive disorder? *IBRO Rep.* **4**, 38–43 (2018).
68. O. Valenti, K. M. Gill, A. A. Grace, Different stressors produce excitation or inhibition of mesolimbic dopamine neuron activity: Response alteration by stress pre-exposure. *Eur. J. Neurosci.* **35**, 1312–1321 (2012).
69. A. A. Grace, Dysregulation of the dopamine system in the pathophysiology of schizophrenia and depression. *Nat. Rev. Neurosci.* **17**, 524–532 (2016).
70. E. N. Holly, K. A. Miczek, Ventral tegmental area dopamine revisited: Effects of acute and repeated stress. *Psychopharmacology (Berl.)* **233**, 163–186 (2016).
71. H. Qin *et al.*, Gamma neurons mediate dopaminergic input during aversive olfactory memory formation in Drosophila. *Curr. Biol.* **22**, 608–614 (2012).
72. W. Li *et al.*, MicroRNA-276a functions in ellipsoid body and mushroom body neurons for naive and conditioned olfactory avoidance in Drosophila. *J. Neurosci.* **33**, 5821–5833 (2013).
73. S. Vijayraghavan, M. Wang, S. G. Birnbaum, G. V. Williams, A. F. Arnsten, Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. *Nat. Neurosci.* **10**, 376–384 (2007).
74. J. Zahrt, J. R. Taylor, R. G. Mathew, A. F. Arnsten, Supranormal stimulation of D1 dopamine receptors in the rodent prefrontal cortex impairs spatial working memory performance. *J. Neurosci.* **17**, 8528–8535 (1997).
75. A. F. Arnsten, Stress weakens prefrontal networks: Molecular insults to higher cognition. *Nat. Neurosci.* **18**, 1376–1385 (2015).
76. M. P. de la Mora, A. Gallegos-Cari, Y. Arizmendi-García, D. Marcellino, K. Fuxe, Role of dopamine receptor mechanisms in the amygdaloid modulation of fear and anxiety: Structural and functional analysis. *Prog. Neurobiol.* **90**, 198–216 (2010).
77. W. Song *et al.*, The interruptive effect of electric shock on odor response requires mushroom bodies in Drosophila melanogaster. *Genes Brain Behav.* **18**, e12488 (2019).
78. M. Rera, R. I. Clark, D. W. Walker, Intestinal barrier dysfunction links metabolic and inflammatory markers of aging to death in Drosophila. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 21528–21533 (2012).
79. R. R. Sokal, F. J. Rohlf, *Biometry: The Principles and Practice of Statistics in Biological Research* (Freeman, New York, ed. 3, 1995).
80. T. Tully, D. Gold, Differential effects of dunce mutations on associative learning and memory in Drosophila. *J. Neurogenet.* **9**, 55–71 (1993).

PNAS BRIEF REPORT | EVOLUTION OPEN ACCESS

Pleiotropy of positive selection in ancient ACE2 suggests an alternative hypothesis for bat-specific adaptations to host coronaviruses

Yuan-Ting Guo[a,b,1], Ji-Bin Jiang[b,c,1], Guan-Rong Qiao[d,1], Rong-Hua Luo[e], Xin Zhou[b,c], Rong Hua[b,c], Chang-Bo Zheng[d,2], and Zhen Liu[b,f,g,2]

Edited by John Speakman, Chinese Academy of Sciences, Shenzhen, China; received December 15, 2023; accepted April 30, 2024

Angiotensin-convertingenzyme 2 (ACE2) has dual functions, regulating cardiovascular physiology and serving as the receptor for coronaviruses. Bats, the only true flying mammals and natural viral reservoirs, have evolved positive alterations in traits related to both functions of ACE2. This suggests significant evolutionary changes in ACE2 during bat evolution. To test this hypothesis, we examine the selection pressure in ACE2 along the ancestral branch of all bats (AncBat-ACE2), where powered flight and bat-coronavirus coevolution occurred, and detect a positive selection signature. To assess the functional effects of positive selection, we resurrect AncBat-ACE2 and its mutant (AncBat-ACE2-mut) created by replacing the positively selected sites. Compared to AncBat-ACE2-mut, AncBat-ACE2 exhibits stronger enzymatic activity, enhances mice's performance in exercise fatigue, and shows lower affinity to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Our findings indicate the functional pleiotropy of positive selection in the ancient ACE2 of bats, providing an alternative hypothesis for the evolutionary origin of bats' defense against coronaviruses.

bat | adaptive evolution | pleiotropy

The coronavirus disease 2019 (COVID-19) pandemic has highlighted the role of angiotensin-converting enzyme 2 (ACE2) as the receptor for SARS-CoV-2, the causative coronavirus (1). However, it is of note that ACE2's original function is to regulate cardiovascular physiological processes. Deficiency in ACE2 is associated with severe pathological conditions such as dilated cardiomyopathy and cardiac dysfunction (2). Interestingly, there appears to be a close correlation between the two functions of ACE2. For example, both hamsters infected with SARS-CoV-2 and patients with severe COVID-19 commonly exhibit cardiac injuries characterized by functional and structural abnormalities, including arrhythmias, myocardial fibrosis, and heart failure (3).

Bats, which are the only group of mammals capable of true powered flight and serve as the unique natural viral reservoirs, are believed to have evolved positive alterations in traits related to the two functions of ACE2. On the one hand, although bats have been suggested as natural hosts for several coronaviruses that cause infectious diseases, they show minimal or no symptomatic signs of diseases (4, 5). Importantly, ACE2 has been found to undergo episodic selection in residues that interact with coronaviruses among bats (6). On the other hand, the origin and evolution of powered flight have critically impacted the morphology and physiological function of bat hearts, resulting in larger hearts, higher heart rates, and higher metabolic rates (7, 8). These findings suggest that ACE2 in the last common ancestor of bats (AncBat-ACE2) may have experienced remarkable genetic changes driven by selection, potentially influencing its two functions. Our focus on the ancestral branch of all bats is justified because of the long evolutionary history of bats coexisting with viral pathogens (9, 10) and the origination of flight in the last common ancestor of bats (11).

Results and Discussion

To test our hypothesis, we used a likelihood method based on a modified branch-site model (12) to estimate the selective pressure on AncBat-ACE2. We compiled a dataset that included high-quality protein-coding sequences of ACE2 from 102 representative mammalian species (Dataset S1). The foreground branch was designated as the ancestral branch of all bats, while the other branches served as background branches. The ω value was estimated to be 58.6, and a total of 10 sites were identified as being under positive selection in ACE2 along the ancestral branch of all bats (Dataset S2). The alternative model showed a significantly higher likelihood than the null model ($P = 0.0012$; two-tailed χ^2 test), suggesting that ACE2 underwent positive selection along the ancestral branch of all bats. Notably, rhinolophid bats have been more frequently identified as hosts for SARS-related viruses compared to

Author affiliations: [a]College of Life Sciences, Northwest University, Xi'an 710127, China; [b]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650201, China; [c]University of Chinese Academy of Sciences, Beijing 100049, China; [d]School of Pharmaceutical Science and Yunnan Key Laboratory of Pharmacology for Natural Products, Kunming Medical University, Kunming 650500, China; [e]Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650201, China; [f]Yunnan Key Laboratory of Biodiversity Information, Kunming 650201, China; and [g]Key Laboratory of Genetic Evolution and Animal Models, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming 650201, China

Author contributions: C.-B.Z. and Z.L. designed research; Y.-T.G., J.-B.J., G.-R.Q., R.-H.L., X.Z., and R.H. performed research; Y.-T.G., J.-B.J., G.-R.Q., R.-H.L., X.Z., and R.H. analyzed data; and Z.L. wrote the paper.

The authors declare no competing interest.

[1]Y.-T.G., J.-B.J., and G.-R.Q. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: zhengchangbo@kmmu.edu.cn or zhenliu@mail.kiz.ac.cn.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2321619121/-/DCSupplemental.

Published June 4, 2024.

other bat clades (13). We thus examined the evolutionary pressure on ACE2 along the branch of rhinolophid bats and did not find a significant signal of positive selection. This suggests that rhinolophid bats may have evolved alternative mechanisms for hosting SARS-related viruses. Together, these findings indicate that ACE2 has undergone positive Darwinian selection along the ancestral branch of all bats, potentially resulting in functional changes in AncBat-ACE2.

Among the 10 positively selected sites (PSSs) on the ancestral branch of all bats, one site (H34T) is involved in contacting SARS-CoVs, five sites (D367E, I407V, A412V, I421M, and T445N) are located in the catalytic domain, two sites (A99I and A193G) are in the zinc-containing subdomain, and the remaining two sites (Y521F and S607H) are in the C-terminus-containing subdomain (14). These locations suggest that the positive selection may be related to both enzymatic activity and coronavirus binding of AncBat-ACE2.

To assess whether the PSSs influence the enzymatic activity of AncBat-ACE2, we used the maximum likelihood method to infer the protein-coding sequence of AncBat-ACE2. The inference was considered reliable as the mean posterior probabilities for the entire sequence of AncBat-ACE2 exceeded 95%. Subsequently, we generated a corresponding mutant of AncBat-ACE2 by artificially replacing the PSSs in AncBat-ACE2 with those found in outgroups (AncBat-ACE2-mut; Dataset S3). The two protein-coding sequences of AncBat-ACE2 and AncBat-ACE2-mut were synthesized and used in the enzymatic activity assay. We conducted fluorescence intensity measurements generated from catalyzed substrates using equal amounts of proteins from AncBat-ACE2 and AncBat-ACE2-mut (Dataset S4). Upon comparing the results, we observed that the slope of the standard curve from the relative fluorescence units (Fig. 1*A*), representing ACE2 activity against substrate per unit time, was significantly larger for AncBat-ACE2 than for AncBat-ACE2-mut

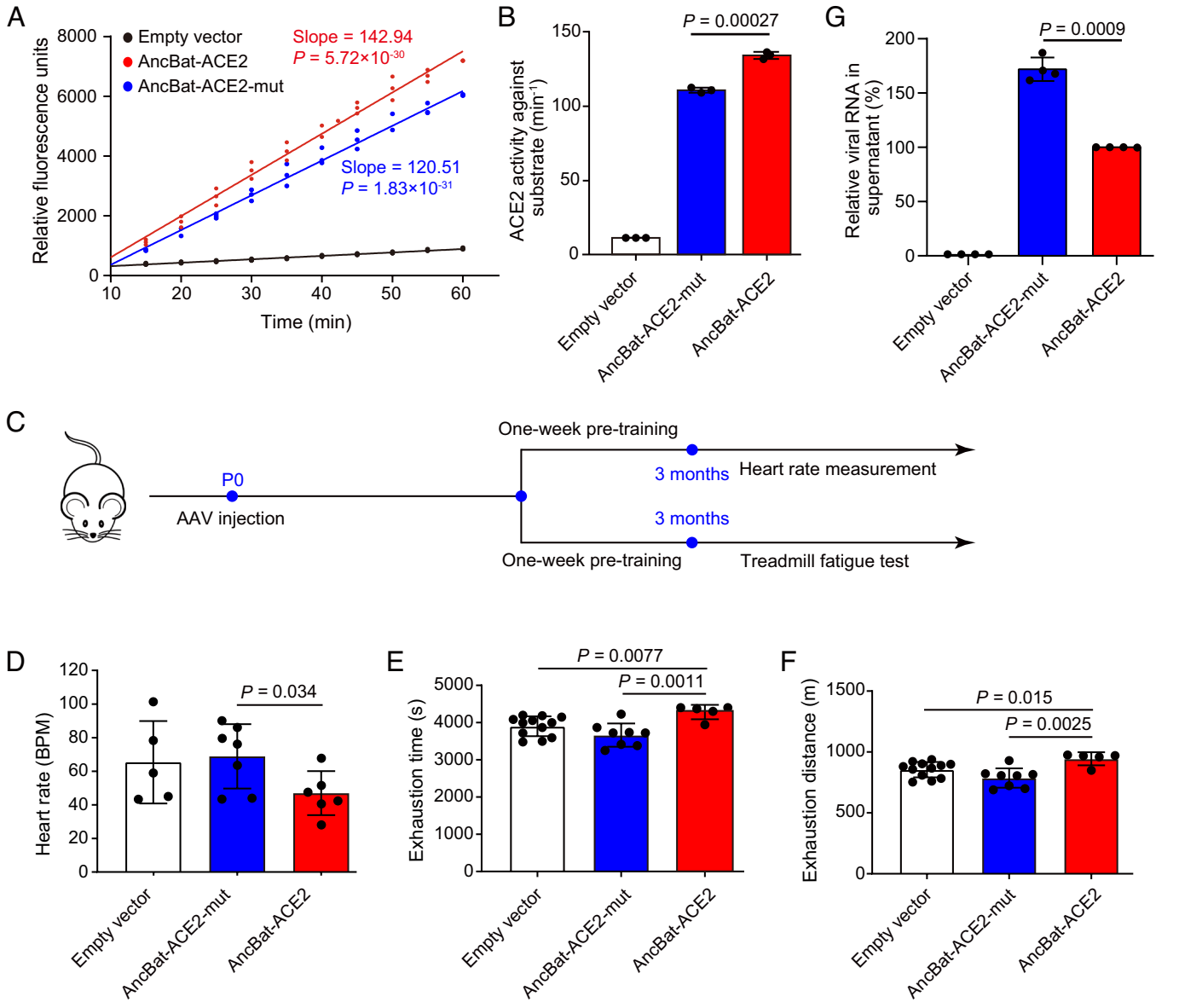


Fig. 1. Functional pleiotropy of positive selection in AncBat-ACE2. (*A*) Detection of enzymatic activities of AncBat-ACE2 and AncBat-ACE2-mut. The curves of the relative fluorescence intensity of the catalytic product over time for AncBat-ACE2 and AncBat-ACE2-mut are presented. The slope of the curve represents the enzymatic activity of ACE2. (*B*) The enzymatic activity of AncBat-ACE2 is significantly higher than that mediated by AncBat-ACE2-mut. (*C*) An illustration of behavioral experiments of mice carrying AncBat-ACE2 and AncBat-ACE2-mut. (*D*) The heart rates for mice carrying AncBat-ACE2 and AncBat-ACE2-mut are compared before and after injecting dobutamine hydrochloride. BPM represents beats per minute. (*E*) The exhaustion time of mice carrying AncBat-ACE2 and AncBat-ACE2-mut is compared. (*F*) The exhaustion distance of mice carrying AncBat-ACE2 is significantly greater than that of mice carrying AncBat-ACE2-mut. (*G*) Comparison of the relative amount of SARS-CoV-2 RNA in the supernatant mediated by AncBat-ACE2 and by AncBat-ACE2-mut. The dots indicate the number of biological replicates. The *P* values are from the two-tailed Student's *t* test.

($P = 0.00027$, two-tailed Student's t test; Fig. 1*B*). This result suggests that positive selection has led to the enhanced enzymatic activity of AncBat-ACE2, as the PSSs were the only sequence differences between AncBat-ACE2 and AncBat-ACE2-mut.

The increased enzymatic activity of AncBat-ACE2 may reflect adaptations to flight in bats, such as enhanced cardiac function and energy metabolism. To test this hypothesis, we generated mouse models using in vivo gene transfer mediated by adeno-associated virus 9 (AAV9) (Fig. 1*C*). Neonatal mice were treated with intravenous injections of AAV vectors carrying AncBat-ACE2 and AncBat-ACE2-mut at a dose of 1×10^{11} virus copies per pup, which was sufficient to transduce cardiac cells, as confirmed by PCR using specific primers after 3 mo of AAV transduction. We then measured heart rates in two groups of mice, one carrying AncBat-ACE2 and the other carrying AncBat-ACE2-mut. While there was no significant difference in resting heart rates, the change in heart rate before and after injecting dobutamine hydrochloride, a commonly used inotropic agent for increasing cardiac output, was significantly lower in mice carrying AncBat-ACE2 compared to mice carrying AncBat-ACE2-mut ($P = 0.034$; two-tailed Student's t test; Fig. 1*D*). This result suggests that mice carrying AncBat-ACE2 may have a stronger capacity for regulating heart rate and cardiac function. To confirm this suggestion, we conducted a treadmill test on the two groups of mice and measured the time and distance they were able to run until exhaustion. On average, mice carrying AncBat-ACE2 ran for 4,117.4 ± 327 (mean ± SD) seconds and covered a distance of 944.7 ± 54 m before exhaustion, which were significantly larger than those from mice carrying AncBat-ACE2-mut ($P = 0.0011$, $P = 0.0025$; two-tailed Student's t tests; Fig. 1 *E* and *F*). Overall, our findings indicate that positive selection leads to stronger enzymatic activity of AncBat-ACE2 and, more importantly, improves cardiac function and energy metabolism in vivo.

Given the suggestion of functional pleiotropy reflected by the locations of the PSSs, we investigated whether positive selection influence the change in infection efficiency of SARS-CoV-2 mediated by AncBat-ACE2. We first conducted protein docking between AncBat-ACE2 and the receptor binding domain (RBD) of SARS-CoV-2, as well as between AncBat-ACE2-mut and the RBD of SARS-CoV-2, using the ZDOCK server (15). Both the binding energy and the dissociation constant suggest a higher binding affinity of AncBat-ACE2-mut with the RBD of SARS-CoV-2 compared to AncBat-ACE2 (Dataset S5). This result was highly consistent with that between the RBD of SARS-CoV-2 and the AncBat-ACE2-mut-2, which was created by artificially replacing the five PSSs not in the catalytic domain of ACE2.

To confirm this result, we, respectively, transfected AncBat-ACE2 and AncBat-ACE2-mut in the A549 cell line and successfully expressed them at comparable levels. The relative quantity of SARS-CoV-2 RNA in the supernatant, mediated by AncBat-ACE2, was significantly lower than that mediated by AncBat-ACE2-mut ($P = 0.0009$; two-tailed Student's t test; Fig. 1*G*). This suggests that the PSSs in AncBat-ACE2 are closely associated with the change in infection efficiency of SARS-CoV-2, further supporting the pleiotropy of positive selection of ACE2 on the ancestral branch of all bats.

In summary, our study identified adaptive selection occurring in ACE2 of the last common ancestor of bats. Through a combination of gene functional experiments and animal behavioral experiments, we demonstrated the functional pleiotropy of positive selection in the ancient ACE2 of all bats. It is challenging to definitively determine which PSSs are more crucial for enhanced enzymatic activity or weaker binding to coronaviruses of AncBat-ACE2. Assuming that the PSSs contribute individually and specifically to these two functional aspects of AncBat-ACE2, the consistently higher binding affinity between AncBat- ACE2-mut and AncBat-ACE2-mut-2 suggests that the five PSSs in the catalytic domain of ACE2 may enhance the enzymatic activity of AncBat-ACE2 and improve the performance of mice in treadmill fatigue tests. This suggestion will require further experiments to be tested in the future. Our findings suggest a potential link between host defenses against coronaviruses and improved cardiac function following the origin of flight in bats, providing an alternative hypothesis for the evolutionary origin of bats' defense against coronaviruses.

Materials and Methods

Based on the ACE2 protein-coding sequences of 102 representative mammalian species, we examined the evolutionary pressures on the ancestral branch of all bats and reconstructed the ACE2 protein-coding sequence of the last common ancestor of all bats (AncBat-ACE2) using the branch-site likelihood method. We synthesized and cloned the protein-coding sequences of AncBat-ACE2 and its mutant (AncBat-ACE2-mut) into expression vectors to perform their enzymatic activity assay and viral infection experiments. After delivering AncBat-ACE2 and AncBat-ACE2-mut to mouse hearts using AAV, we measured the heart rate of the mice and conducted treadmill fatigue tests. Details are available in *SI Appendix*.

Data, Materials, and Software Availability. All study data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. This research was funded by grants from the National Key R&D Program of China (2021YFC2301300), the National Natural Science Foundation of China (32192422, 32330014, 82200550, 81960662), Yunnan Fundamental Research Projects (202201AS070058, 202102AA310055, 202101-AT070154, 202301AT070270), the CAS "Light of West China Program" (xbzg-zdsys-202113), and the Key Research Program of the Chinese Academy of Sciences (KJZD-SW-L11).

1. P. Zhou *et al.*, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
2. J. Zhong *et al.*, Angiotensin-converting enzyme 2 suppresses pathological hypertrophy, myocardial fibrosis, and cardiac dysfunction. *Circulation* **122**, 717–728 (2010).
3. A. Viveiros *et al.*, SARS-CoV-2 infection downregulates myocardial ACE2 and potentiates cardiac inflammation in humans and hamsters. *Am. J. Physiol. Heart Circ. Physiol.* **323**, H1262–H1269 (2022).
4. A. T. Irving, M. Ahn, G. Goh, D. E. Anderson, L. F. Wang, Lessons from the host defences of bats, a unique viral reservoir. *Nature* **589**, 363–370 (2021).
5. A. Banerjee, A. C. Doxey, K. Mossman, A. T. Irving, Unraveling the zoonotic origin and transmission of SARS-CoV-2. *Trends Ecol. Evol.* **36**, 180–184 (2021).
6. H. K. Frank, D. Enard, S. D. Boyd, Exceptional diversity and selection pressure on coronavirus host receptors in bats compared to other mammals. *Proc. Biol. Sci.* **289**, 20220193 (2022).
7. M. T. O'Mara *et al.*, Cyclic bouts of extreme bradycardia counteract the high metabolism of frugivorous bats. *Elife* **6**, e26686 (2017).
8. M. Canals, C. Atala, B. Grossi, J. Iriarte-Diaz, Relative size of hearts and lungs of small bats. *Acta Chiropterol.* **7**, 65–72 (2005).
9. D. Jebb *et al.*, Six reference-quality genomes reveal evolution of bat adaptations. *Nature* **583**, 578–584 (2020).
10. E. C. Skirmuntt, M. Escalera-Zamudio, E. C. Teeling, A. Smith, A. Katzourakis, The potential role of endogenous viral elements in the evolution of bats as reservoirs for zoonotic viruses. *Annu. Rev. Virol.* **7**, 103–119 (2020).
11. N. B. Simmons, K. L. Seymour, J. Habersetzer, G. F. Gunnell, Primitive early Eocene bat from Wyoming and the evolution of flight and echolocation. *Nature* **451**, 818–821 (2008).
12. Z. Yang, PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
13. S. Zhou *et al.*, ZOVER: The database of zoonotic and vector-borne viruses. *Nucleic Acids Res.* **50**, D943–D949 (2022).
14. P. Towler *et al.*, ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis. *J. Biol. Chem.* **279**, 17996–18007 (2004).
15. B. G. Pierce *et al.*, Accelerating protein docking in ZDOCK using an advanced 3D convolution library. *PLoS One* **6**, e24657 (2011).

BMC Medical Education

RESEARCH Open Access

Assessment and predictive analysis of health professions students' proficiency in health economics and pharmacoeconomics in Southwest China: a cross-sectional study

Jian Yang[1,2,3†], Quanzhi Wei[2,3,5†], Xin Yang[1,2,3†], Jingyi Jiao[1,2,3†], Zaixian Yang[1,3,4†] and Fan Li[2,3,4*]

Abstract

Objectives This study addresses the growing importance of health economics and pharmacoeconomics in preparing healthcare professionals for economic decision-making, particularly in China's evolving healthcare system. Despite their significance, these subjects are inconsistently included in health education in Southwest China, making it crucial to assess students' proficiency and educational needs.

Methods A cross-sectional study was conducted from July to October 2023, involving medical and pharmaceutical programs across universities in Sichuan, Chongqing, Guizhou, and Yunnan. Data on students' knowledge, attitudes, and proficiency were captured using a structured questionnaire. Advanced statistical methods, including descriptive analysis, ANOVA, regression analysis, PCA, and Spearman correlation heatmaps, were employed for data analysis using IBM SPSS and Python. The sample consisted of 781 students, with the survey facilitated via the Questionnaire Star platform.

Results The survey revealed moderate comprehension of health economics and pharmacoeconomics among 781 health-related students, with an average knowledge score of 2.9. The majority were pharmacy students (45.58%) and aged 21–24 years (65.81%) with a significant female majority (64.92%). Academic year significantly influenced understanding, while gender differences were minimal. Spearman correlation indicated a strong linkage (0.78) between Supply and Demand and Cost Effectiveness, with minimal correlations between knowledge and attitudes towards the educational importance of health economics. Multiple regression analysis highlighted the impact of gender and academic progression on students' comprehension of Supply and Demand concepts, as well as the influence of knowledge levels on their attitudes towards the importance of health economics education. PCA indicated the complex structure of knowledge and attitudes.

Conclusions The study found moderate knowledge levels in health economics among health professions students in Southwest China, with academic year affecting comprehension and minimal gender differences. Pharmacy students had higher proficiency in pharmaceutical-related areas, but gaps in complex topics like 'Economic Evaluation'

†Jian Yang, Quanzhi Wei, Xin Yang, Jingyi Jiao and Zaixian Yang contributed equally to the study.

*Correspondence:
Fan Li
1425256831@qq.com
Full list of author information is available at the end of the article

highlight the need for targeted educational improvements to better prepare students for healthcare economic decision-making.

Keywords Health economics, Pharmacoeconomics, Medical education, Curriculum development, Knowledge assessment

Introduction

The burgeoning advancements in the health sector are inherently intertwined with economic implications, thereby accentuating the imperative role of health economics and pharmacoeconomics in sculpting proficient healthcare professionals. Health economics, broadly encapsulating the allocation, utilization, and behavior towards healthcare resources, and pharmacoeconomics, focusing on identifying the most efficient means to allocate resources in the field of pharmacy, stand pivotal in determining healthcare strategies, policies, and decision-making processes [1].

The importance of health economics and pharmacoeconomics extends beyond theoretical knowledge, intertwining with practical aspects of healthcare delivery, policy-making, and strategic development in the health sector. A nuanced understanding of these disciplines enables healthcare professionals to make informed decisions, ensuring the optimal utilization of resources, facilitating cost-effective healthcare delivery, and ultimately enhancing healthcare outcomes [2].

Internationally, there has been a surge in the integration of health economics and pharmacoeconomics into the curricula of health professions education. In the United States, the incorporation of these subjects is recognized as essential in fostering a comprehensive understanding of healthcare systems among medical and pharmacy students. Studies have indicated that while students acknowledge the importance of economic knowledge in healthcare, their proficiency levels vary significantly [3, 4]. This variation can be attributed to differences in educational frameworks, the depth of the curriculum, and the pedagogical approaches employed.

In European contexts, similar trends have been observed. A study across several European universities revealed that health economics is increasingly recognized as a vital component of healthcare education, though the extent and depth of coverage vary considerably [5]. In the United Kingdom, the emphasis on pharmacoeconomics has grown in response to the National Health Service's demand for cost-effective healthcare solutions, leading to more focused educational initiatives in this area [6].

While health economics and pharmacoeconomics are globally recognized as essential components of healthcare education, their integration into curricula varies across regions. Beyond examples from the USA, UK, and Europe, studies from other Asian and Middle Eastern countries offer valuable insights. In Japan, Suzuki et al. (2023) emphasized socioeconomic disparities among medical students and the growing need for equitable healthcare resource distribution through health economics education [7]. Similarly, Kim et al. (2023) found that socioeconomic differences influence perceptions of healthcare inequalities in South Korea, underscoring the importance of integrating health economics education to bridge these gaps [8].

In the Middle East, Alzarea et al. (2022) reviewed barriers to implementing pharmacoeconomics in high- and low-income countries, highlighting challenges such as the lack of governing bodies and experts, but also the opportunities provided by collaborations between researchers and policymakers [9]. In Jordan, Hammad et al. (2020) found that pharmacy students had the highest awareness of health economics, while medical and nursing students showed a need for further training [10]. Despite growing recognition of the importance of health economics education in Asia and the Middle East, its inclusion in curricula remains inconsistent.

The scenario in Asian countries, including China, presents a unique perspective. The rapid evolution of the healthcare system in China, marked by increasing healthcare demands and economic constraints, necessitates a workforce proficient in health economics and pharmacoeconomics. However, research on the proficiency levels of health professions students in these areas within the Chinese context, especially in the southwest region, remains limited [11, 12].

In the context of China, particularly its southwestern region, health professions students face a diverse and dynamic healthcare environment, shaped by rapid technological advancements and complex healthcare needs. This region's unique demographic and socioeconomic profile, marked by varied healthcare demands across both urban and rural settings, creates a complex tableau. In such a context, the role of health economics and pharmacoeconomics becomes critical for ensuring equitable, efficient, and sustainable healthcare delivery. Although these subjects are globally recognized as essential components of healthcare education, they have not yet been universally included in health professions curricula in Southwest China. While some students may have encountered these topics through elective courses,

seminars, or guest lectures, formal instruction remains inconsistent [13]. This study aims to uncover the current state of students' awareness, knowledge, and attitudes towards these crucial subjects. Identifying gaps, challenges, and opportunities in the existing educational framework is vital for enhancing educational strategies and curricula that meet the unique needs of students and the regional healthcare context.

Methods

Study design and setting

This cross-sectional study, conducted from July 17 to October 27, 2023, aimed to assess the proficiency of health economics and pharmacoeconomics among health professions students in universities across Sichuan, Chongqing, Guizhou, and Yunnan provinces in Southwest China. The study targeted universities offering medical and pharmaceutical programs to gain comprehensive insights into students' understanding and application of these essential disciplines.

To collect data, a carefully constructed questionnaire was developed based on an extensive literature review and expert consultations [14, 15]. The questionnaire was designed to evaluate students' awareness, knowledge, attitudes, and perceived barriers in health economics and pharmacoeconomics. It included Likert scale questions to capture a broad range of relevant data [16, 17]. To ensure inclusivity, the questionnaire was made available in both English and Mandarin.

A pilot study with 50 students from health-related programs (not included in the final sample) was conducted to assess and refine the questionnaire. Feedback from this phase helped improve the survey design. Reliability, measured by Cronbach's alpha, yielded a value of 0.82, indicating satisfactory internal consistency. Expert opinions from professionals in health economics and education ensured content validity, and the questionnaire was revised to improve clarity and alignment with the study's objectives. After these validation steps, the final questionnaire was deployed to collect data from students across the targeted provinces, providing valuable insights for academic development and policy formulation in health education. (The English version of the questionnaire has been uploaded as a supplementary file.)

Participants and data collection

This quantitative cross-sectional study targeted students enrolled in medical and health-related programs across universities in Sichuan, Chongqing, Guizhou, and Yunnan, Southwest China. The survey was designed to assess whether students had formally studied health economics or pharmacoeconomics, as well as whether they had encountered these topics through other means such as elective courses or seminars. Our primary aim was to evaluate and predictively analyze their proficiency in health economics and pharmacoeconomics.

We used the "Questionnaire Star platform," a widely utilized online tool in China, to collect data in both the preliminary and main phases. This platform was selected to facilitate comprehensive data collection and detect knowledge gaps that could inform curriculum improvements. The survey was distributed via university email systems and academic platforms to ensure a representative sample from the student population.

Inclusion and exclusion criteria

Students enrolled in any health-related program were included in the study. We excluded participants who submitted incomplete or invalid questionnaires, defined as responses that did not meet certain quality criteria, such as logical inconsistencies or excessive time taken to complete the survey. Specifically, questionnaires taking more than 10 min to complete (the minimum completion time tested by our team) and those answering two key logical questions correctly were deemed valid and included in the analysis. Otherwise, they were excluded.

For determining the sample size, the formula used was $\boldsymbol{n}=\boldsymbol{z}^2\times\boldsymbol{p}\times(1-\boldsymbol{p})/\boldsymbol{d}^2$, where 'z' is the z-score for a 95% confidence level (1.96), 'p' is the estimated proportion of the population exhibiting the attribute (assumed to be 0.5), and 'd' is the margin of error (set at 0.04). According to this formula, a sample size of at least 601 was required. We surveyed 838 individuals, and after applying the inclusion and exclusion criteria, 57 responses were excluded, resulting in 781 valid responses for analysis.

Data analysis

The questionnaire used a Likert-scale format, where participants rated their responses from 1 (strongly disagree) to 5 (strongly agree). These responses were summed within each domain to calculate an overall score. The total score for each participant was normalized by dividing it by the number of items in that domain, resulting in a standardized score ranging from 1 to 5. To ensure accuracy, incomplete or missing responses were excluded from the analysis. Specifically, if more than 20% of items in a domain were unanswered, the data for that domain were excluded. Responses that were logically inconsistent or contradictory were also removed.

This study analyzed the proficiency of health professions students in health economics and pharmacoeconomics in Southwest China using IBM SPSS 22.0 and Python 3.10. Descriptive statistics summarized participant demographics, while ANOVA, regression analysis, and K-Means Clustering identified key patterns in student proficiency. Principal Component Analysis (PCA)

simplified the data for better interpretation, while correlation and heatmap analyses visually represented relationships between educational variables. All analyses were performed using a two-tailed test with a significance level of $P < 0.05$ to ensure robustness [18].

Ethical considerations

This study adhered to stringent ethical standards with approval from the Institutional Review Board of Kunming Medical University (Protocol Number: 202300279). Prior to data collection, informed consent was obtained from each participant, ensuring awareness of the study's objectives and participant rights. This procedure was vital for maintaining the research's ethical integrity and safeguarding participant confidentiality, contributing to a respectful and trustworthy research environment.

Results

Demographic overview of survey participants

The survey targeting university students with a focus on health economics and pharmacoeconomics included a sample of 781 participants. The demographics revealed a predominant female representation, with 507 female students compared to 274 male students. Most respondents (514) fell within the 21–24 age bracket, indicating a primary undergraduate demographic. Academic program participation was led by Pharmacy with 356 students, highlighting its significance in the surveyed areas. Furthermore, 335 students were in their fourth year of study, suggesting the survey's appeal to undergraduates nearing degree completion. The survey encompassed a wide array of academic disciplines, extending to 20 different programs such as Public Health, Clinical Medicine, and Pharmacy Administration. This diverse academic representation provided a comprehensive view of the perspectives and knowledge levels regarding health economics and pharmacoeconomics among university students. Among the respondents, 58% reported having taken a formal course in health economics or pharmacoeconomics, while the remaining 42% indicated that they had not received formal instruction in these subjects. However, some students mentioned that they had encountered related content through elective courses, seminars, or guest lectures. Pharmacy students were more likely to have encountered these topics in their coursework, with 65% of pharmacy students reporting formal education in health economics. This data, along with the demographic breakdown, offers valuable insight into the educational exposure and diversity of the student population. (Refer to Table 1 for detailed demographic and academic profiles).

Table 1 Demographic and academic profile

Categories	Count	Percentage
Sex		
Male	274	35.08%
Female	507	64.92%
Age		
18–20	249	31.88%
21–24	514	65.81%
25–29	12	1.54%
30 or above	6	0.77%
Program		
Clinical Medicine	79	10.12%
Pharmacy	356	45.58%
Nursing	3	0.38%
Public health	252	32.27%
Other	91	11.65%
Year of Study		
First year	43	5.51%
Second year	142	18.81%
Third year	200	25.61%
Fourth year	335	42.89%
Fifth year or above	4	0.51%
Master	48	6.15%
Doctor	9	1.15%

Table 2 Descriptive statistics of knowledge scores in health economics and pharmacoeconomics domains

Domain	Average Score	Standard Error
Supply and Demand Knowledge	2.95	0.04
Cost Effectiveness Knowledge	2.99	0.05
Insurance Reimbursement Knowledge	2.99	0.05
Economic Evaluation Knowledge	2.89	0.04
Pricing Strategies Knowledge	2.81	0.04
Budget Impact Knowledge	2.82	0.04
Pharmaceutical Expenditure Knowledge	2.95	0.05

Quantitative analysis of student proficiency and influencing factors in health economics and pharmacoeconomics education

Quantitative assessment of student knowledge in economic healthcare disciplines

The quantitative analysis evaluates university students' knowledge in health economics and pharmacoeconomics based on the findings from Table 2 and Table 3. The average knowledge scores are around 2.9, with moderate variability across different domains, indicating basic proficiency in areas such as 'Supply and Demand,' 'Cost

Table 3 Comparative knowledge scores of pharmacy vs. non-pharmacy students (Mean ± SE)

Domain	Pharmacy Students	Non-Pharmacy Students	*P*-value
Supply and Demand Knowledge	2.96 ± 0.067	2.93 ± 0.056	< 0.05
Cost Effectiveness Knowledge	3.15 ± 0.063	2.86 ± 0.056	< 0.05
Insurance Reimbursement Knowledge	3.08 ± 0.066	2.92 ± 0.054	< 0.05
Economic Evaluation Knowledge	3.17 ± 0.065	2.66 ± 0.060	< 0.01
Pricing Strategies Knowledge	3.03 ± 0.066	2.63 ± 0.062	< 0.01
Budget Impact Knowledge	2.99 ± 0.066	2.68 ± 0.060	< 0.05
Pharmaceutical Expenditure Knowledge	3.16 ± 0.066	2.77 ± 0.060	< 0.01

Effectiveness,' and 'Insurance Reimbursement.' Table 2 highlights key areas such as 'Economic Evaluation,' 'Pricing Strategies,' and 'Budget Impact,' all showing room for improvement across student groups. The 'Pharmaceutical Expenditure' domain exhibits broader variability in knowledge, as demonstrated in Table 3, which provides a detailed comparison between pharmacy and non-pharmacy students. In the Pharmaceutical Expenditure domain, pharmacy students scored significantly higher (3.16 ± 0.066) than non-pharmacy students (2.77 ± 0.061). This difference indicates that pharmacy students benefit from a curriculum that emphasizes pharmaceutical economics more than that of their non-pharmacy counterparts. Moreover, pharmacy students also outperformed non-pharmacy students in areas like 'Cost Effectiveness' (3.15 for pharmacy students vs. 2.86 for non-pharmacy students) and 'Economic Evaluation' (3.17 for pharmacy students vs. 2.66 for non-pharmacy students), suggesting a stronger overall understanding of economic principles in healthcare.

Knowledge differences across academic years in health economics and pharmacoeconomics: a comparative analysis

Table 4 shows the average scores for key domains like 'Supply and Demand,' 'Cost Effectiveness,' 'Insurance Reimbursement,' 'Economic Evaluation,' 'Pricing Strategies,' 'Budget Impact,' and 'Pharmaceutical Expenditure.' Students in later academic years, particularly doctoral students, demonstrated significantly higher understanding across most domains. For instance, doctoral students scored higher in 'Cost Effectiveness' (Mean: 3.89 ± 0.35) and 'Economic Evaluation' (Mean: 4.00 ± 0.33), while first-year students scored lower across all domains. This trend reflects a consistent improvement in knowledge with academic progression, especially in complex areas like 'Cost Effectiveness' and 'Economic Evaluation.' Earlier-year students, particularly those in their first and second years, scored lower in advanced topics such as 'Pharmaceutical Expenditure' and 'Pricing Strategies.' This suggests that early targeted interventions may improve comprehension in these areas. ANOVA analysis confirmed significant differences in knowledge across academic years, with F-values showing statistical significance ($P < 0.05$) in all key domains, highlighting the need for structured, progressive education as students advance in their academic careers (refer to Table 5).

Correlation patterns in health economics education perspectives

The Spearman correlation heatmap, shown in Fig. 1a, is essential for understanding the relationships between various knowledge domains and educational attitudes in health economics and pharmacoeconomics. Significant correlations are evident among

Table 4 Knowledge scores across academic years in key domains (Mean ± SE)

Academic Year	Supply and Demand	Cost Effectiveness	Insurance Reimbursement	Economic Evaluation	Pricing Strategies	Budget Impact	Pharmaceutical Expenditure
Doctor	3.11 ± 0.35	3.89 ± 0.35	3.33 ± 0.17	4.00 ± 0.33	3.22 ± 0.32	3.00 ± 0.24	3.67 ± 0.33
Fifth year or above	2.50 ± 0.65	2.75 ± 0.63	3.00 ± 0.71	2.25 ± 0.75	2.50 ± 0.65	2.75 ± 0.63	3.00 ± 0.71
First year	2.07 ± 0.19	2.28 ± 0.21	2.30 ± 0.20	2.12 ± 0.21	2.05 ± 0.19	2.21 ± 0.22	2.42 ± 0.21
Fourth year	3.05 ± 0.06	3.04 ± 0.06	2.96 ± 0.06	2.86 ± 0.07	2.81 ± 0.07	2.81 ± 0.06	2.90 ± 0.07
Master	2.56 ± 0.17	2.88 ± 0.18	2.94 ± 0.17	2.92 ± 0.19	2.67 ± 0.18	2.63 ± 0.19	2.79 ± 0.20
Second year	2.95 ± 0.10	2.96 ± 0.10	3.04 ± 0.10	3.08 ± 0.10	2.89 ± 0.11	2.93 ± 0.10	3.09 ± 0.10
Third year	3.05 ± 0.09	3.08 ± 0.09	3.15 ± 0.09	2.93 ± 0.09	2.93 ± 0.09	2.95 ± 0.09	3.05 ± 0.09

Table 5 ANOVA results of knowledge scores by academic year (F-Values and *P*-Values)

Domain	F-value	*P*-value
Supply and Demand	5.615	0.000010
Cost Effectiveness	3.947	0.000677
Insurance Reimbursement	3.308	0.003194
Economic Evaluation	4.758	0.000090
Pricing Strategies	3.293	0.003315
Budget Impact	2.514	0.020424
Pharmaceutical Expenditure	2.512	0.020542

specific knowledge areas, such as the strong correlation coefficient of 0.78 between 'Supply_Demand' and 'Cost_Effectiveness,' indicating that proficiency in one domain often correlates with competence in another. This interconnectedness is further illustrated by the correlation of 0.73 between 'Insurance_Reimbursement' and 'Economic_Evaluation.'

Conversely, the relationships between knowledge levels in these domains and attitudes toward the importance of health economics education are relatively weaker. For instance, 'Knowledge_Supply_Demand' has a low correlation (0.12) with the perceived importance of health economics education. However, moderate correlations exist among attitudes, particularly a 0.79 correlation between the belief in including pharmacoeconomics in the curriculum and the understanding of its career enhancement potential. This indicates that students recognize the career benefits associated with pharmacoeconomics education. These findings highlight the complex nature of students' educational experiences and perceptions, suggesting opportunities for curriculum development to better integrate knowledge acquisition with the practical significance of these subjects.

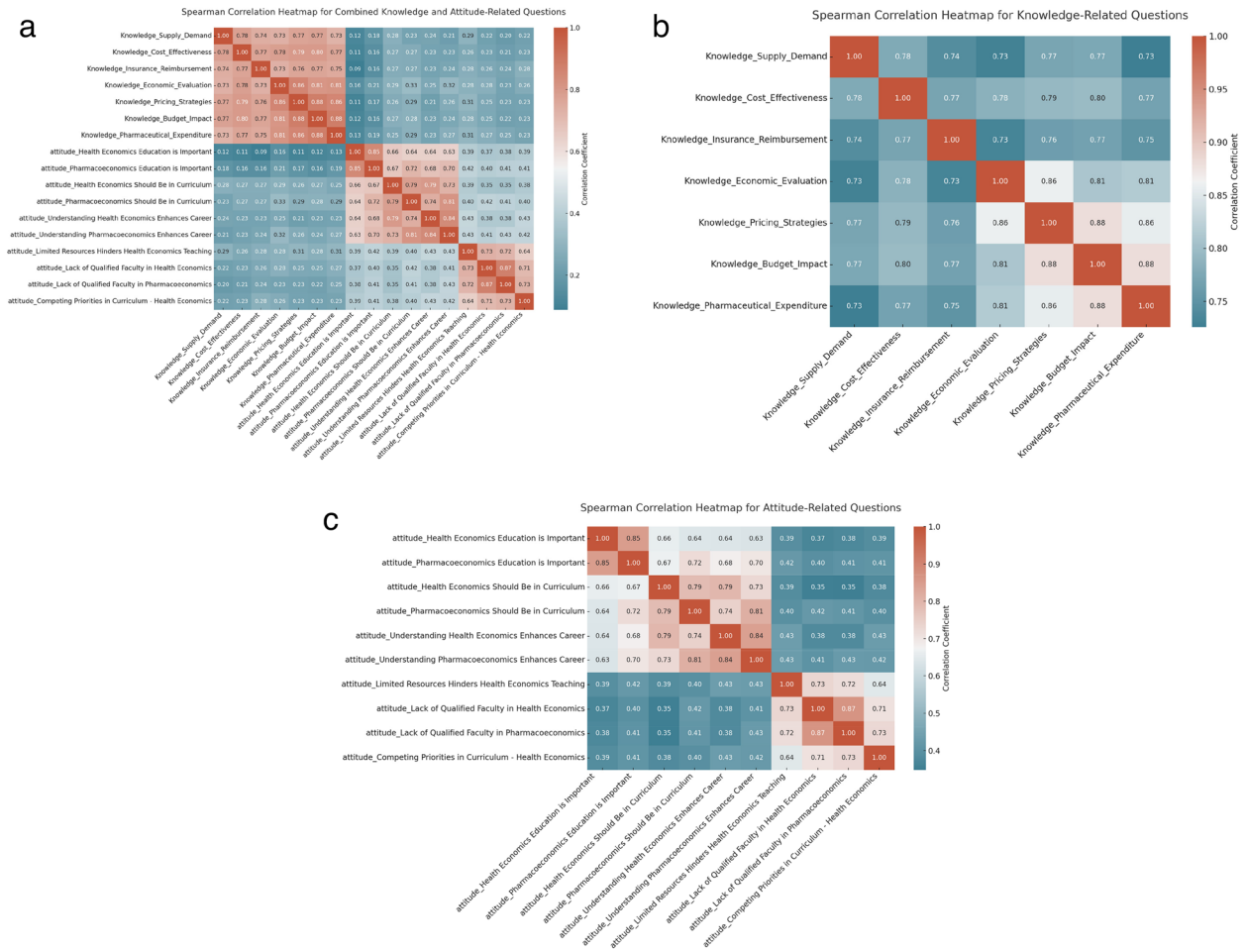


Fig. 1 **a** Spearman Correlation Heatmap of Survey Responses on Health Economics and Pharmacoeconomics Education. **b** Correlation Matrix Heatmap: Interrelationships in Student Knowledge of Health Economics and Pharmacoeconomics Domains. **c** Spearman Correlation Heatmap for Attitude-Related Questions in Health Economics Education

Correlation matrix of knowledge domains for health economics and pharmacoeconomics survey

The Spearman Correlation Heatmap, shown in Fig. 1b, is a crucial tool for analyzing the relationships among different knowledge domains within Health Economics and Pharmacoeconomics. The heatmap reveals significant correlations, ranging from 0.73 to 0.80, underscoring the interconnections between various conceptual areas. A strong understanding in one area, such as supply and demand, often correlates with a similar level of comprehension in related topics like cost-effectiveness and insurance reimbursement. The heatmap highlights particularly strong correlations between knowledge of pricing strategies and economic evaluation (0.86), as well as between budget impact and pharmaceutical expenditure (0.88). This suggests that these areas are either taught together or more easily understood by students in a cohesive manner. The consistency of these strong correlations indicates a systematic and integrated approach to learning Health Economics and Pharmacoeconomics. These findings provide valuable insights for educators and curriculum planners, helping them to enhance educational content and foster a more interconnected understanding of these essential subjects.

Analysis of the spearman correlation heatmap for attitude-related questions in health economics education

The Spearman Correlation Heatmap, shown in Fig. 1c, highlights the strong interconnections in students' attitudes toward Health Economics and Pharmacoeconomics education. With high correlation values, peaking at 0.85, the heatmap reflects a strong consensus on the importance of including these subjects in academic curricula. Correlations of 0.66 and 0.72 indicate solid agreement on the integration of these fields, underscoring their perceived significance for both academic and professional development. The heatmap also reveals correlations between 0.64 and 0.70, linking students' understanding of these subjects with their perception of enhanced career prospects. Additionally, moderate correlations, such as 0.73 related to limited resources and 0.71–0.73 for challenges like inadequate faculty and competing curriculum priorities, highlight perceived institutional barriers. This emphasizes the strong student belief in the value of Health Economics and Pharmacoeconomics education while pointing to the need to address institutional challenges to effectively integrate these subjects into the curriculum.

Dynamics of knowledge and attitudes in health economics education: a multifaceted regression analysis

Regression insights on educational progress and supply–demand comprehension in health economics

The regression analysis explores the effects of gender and academic progression on the understanding of supply and demand concepts in health economics. The intercept value of 3.0258 reflects the baseline level of comprehension, suggesting that the average student scores around 3.0 out of 5 in understanding these concepts. Gender plays a significant role, with male students showing a higher level of understanding, as indicated by a regression coefficient of 0.2246, a result that is statistically significant ($P < 0.05$). Academic progression also has a significant impact on comprehension. First-year students, with a regression coefficient of -1.0657, show a much lower understanding compared to students in higher academic years, suggesting that comprehension improves progressively with academic advancement. For students in their fifth year and beyond, the regression coefficient is close to zero, indicating that their understanding level is near the baseline, as indicated by the intercept. The R-squared value of 0.050 indicates that gender and academic year together explain only 5% of the variance in students' understanding, implying that other factors also contribute to their grasp of supply and demand. As illustrated in Fig. 2a, the analysis highlights the influence of both gender and academic progression on students' economic comprehension. However, additional variables likely play a role in shaping this understanding. The results point to the need for targeted educational interventions, especially for earlier-year students, to strengthen their grasp of key economic principles in health economics.

Insights into knowledge influences on health economics educational attitudes

This regression analysis delves into how students' knowledge across various domains influences their attitudes towards the importance of health economics education. The results, as depicted in Fig. 2b, show a positive relationship between knowledge in Economic Evaluation and educational attitudes, with a coefficient of 0.167. This suggests that students with a stronger grasp of economic evaluation are more likely to value health economics education. Similarly, comprehension of Supply and Demand (coefficient = 0.075) and Budget Impact (coefficient = 0.042) also contributes positively to students' attitudes, though their influence is somewhat smaller compared to economic evaluation. In contrast, knowledge in areas such as Cost Effectiveness and Insurance Reimbursement displays an inverse relationship with educational attitudes, with coefficients ranging from -0.049 to -0.086. This indicates that proficiency in these areas may unexpectedly reduce the perceived importance of health economics education, suggesting a possible disconnect between expertise in these domains and the perceived broader value of the subject. The model's intercept at 3.701 reflects a general baseline appreciation for the importance of health economics education across

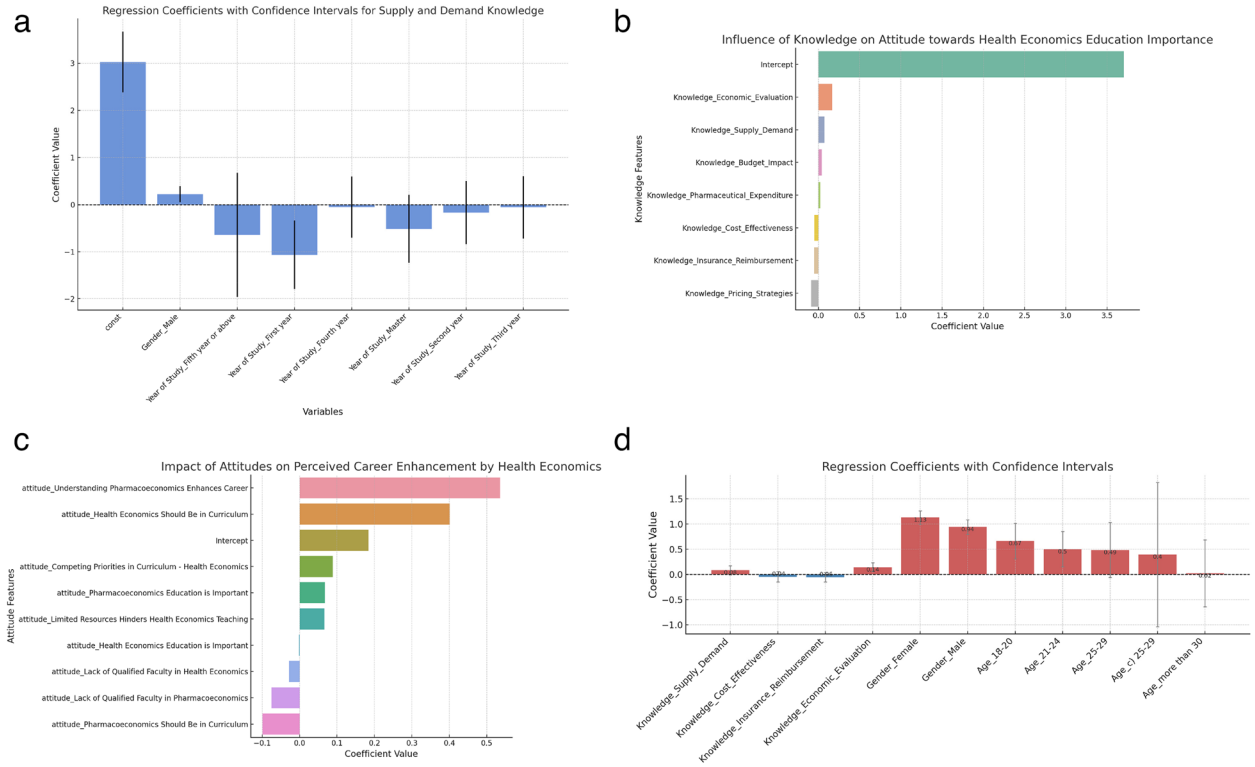


Fig. 2 **a** Regression Coefficients with Confidence Intervals for Supply and Demand Knowledge. **b** Regression Coefficients with Confidence Intervals for Educational Attitudes in Health Economics. **c** Regression Coefficients with Confidence Intervals for Attitudinal Determinants of Career Enhancement in Health Economics Education. **d** Regression Coefficients and Confidence Intervals for Factors Influencing Attitudes Towards Health Economics Education

the student population. These insights, as visualized in Fig. 2b, highlight a complex interplay between knowledge and attitudes, indicating that while understanding in some areas reinforces the value of health economics education, in other areas, it may paradoxically diminish it. This suggests the need for a balanced educational approach that aligns students' knowledge with the perceived importance of these concepts.

Influence of educational attitudes on career prospects in health economics

This regression analysis investigates how students' attitudes towards health economics and pharmacoeconomics education influence their perceived career prospects. As shown in Fig. 2c, there is a strong positive relationship between valuing Pharmacoeconomics Education and the belief that it enhances career prospects, with a regression coefficient of 0.536. This indicates that students who recognize the importance of pharmacoeconomics education are more likely to perceive it as beneficial for their careers. Furthermore, the belief that Health Economics Should Be in the Curriculum is also positively correlated with career outlook, with a coefficient of 0.402, emphasizing the role of curriculum design in shaping students' professional expectations. However, concerns about Competing Priorities in the Curriculum and Lack of Qualified Faculty in health economics show negative effects, with coefficients ranging from -0.075 to -0.099. This suggests that students facing institutional challenges may have lower career expectations from health economics education. The intercept, set at 0.184, represents the baseline perception of career relevance. These results highlight opportunities for academic institutions to enhance the perceived career benefits of health economics education by addressing curriculum gaps and resource limitations.

Impact of demographic and knowledge factors on attitudes toward health economics education

Our multiple regression analysis, depicted in Fig. 2d, critically examines the influence of demographic factors (gender, age) and knowledge levels (supply and demand, cost-effectiveness, insurance reimbursement, economic evaluation) on students' attitudes towards the importance of health economics education. The model, accounting for 5.5% of the variance (Adj. R-squared:

0.044), reveals significant insights. Notably, 'Knowledge_Economic_Evaluation' emerges as a strong positive influencer, suggesting that a deeper understanding of economic evaluation correlates with a heightened appreciation of health economics in education. Gender differences are also pronounced; both male and female students exhibit significant differences in their attitudes, with females showing a higher tendency to value health economics education. Age-wise, younger groups ('18–20' and '21–24') demonstrate a more positive attitude compared to their older counterparts, indicating a generational shift in the perception of this field. The depicted coefficients and their confidence intervals highlight the critical elements shaping students' perspectives, thereby providing a roadmap for tailored educational interventions in health economics. The model's moderate explanatory power and the suggestion of multicollinearity warrant cautious interpretation, implying the potential influence of other unmeasured factors.

Principal component analysis of knowledge levels and attitudes towards health economics and pharmacoeconomics education

Principal Component Analysis (PCA) was utilized to discern the multidimensional structure of respondents' knowledge and attitudes towards health economics and pharmacoeconomics education. This method effectively reduced the dataset's dimensionality, unveiling the underlying patterns in the data. The first principal component accounted for the majority of variance, followed by the second component, together capturing a significant share of the total variance. Figure 3 revealed a broad dispersion of respondents across the principal components, indicative of a diverse array of knowledge and attitudinal profiles. A concentration of data points near the plot's origin suggests shared characteristics among many respondents, while outliers reflect unique educational experiences and perceptions. This variance emphasizes the individual differences in comprehension and valuation of these educational fields. The PCA's insights into the variability of respondents' understanding and beliefs are instrumental for crafting bespoke educational strategies. By identifying key dimensions of differentiation in perceptions, PCA establishes a groundwork for targeted interventions, aiming to bridge knowledge gaps and align attitudes with educational objectives, thereby enhancing the efficacy of health economics and pharmacoeconomics curricula.

Latent structures in educational perceptions: a factor analysis study

In the analysis of "Defining Educational Perspectives: Factor Loadings in Health Economics and Pharmacoeconomics," a detailed factor analysis reveals the complex educational dimensions within these fields. Figure 4 highlights four key factors shaping knowledge and attitudes. Factor 1 is dominated by knowledge-focused variables, with strong negative loadings (-0.68 to -0.75), representing a dimension inversely linked to fundamental understanding. Factor 2 centers on the perceived importance of education, with positive loadings (0.38 to 0.53), highlighting attitudes toward its significance. Factor 3

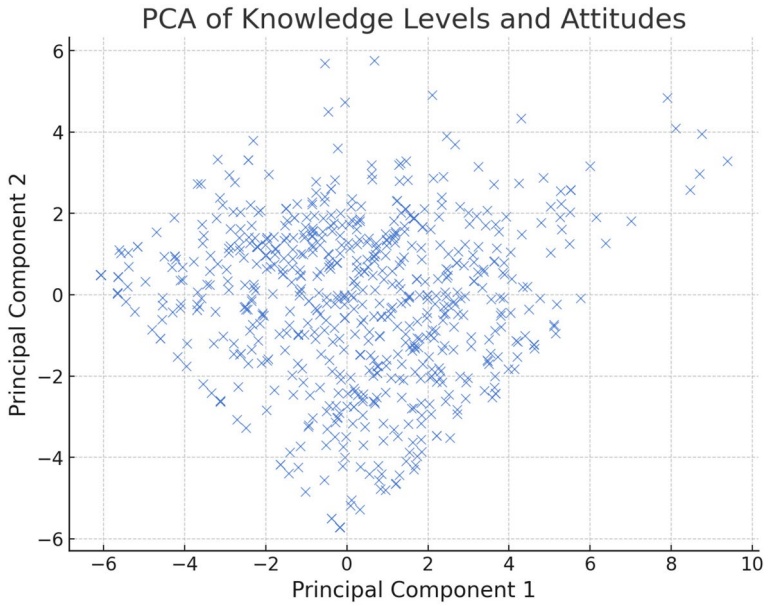


Fig. 3 Principal component analysis of knowledge levels and attitudes

reflects concerns about resource limitations and faculty qualifications, with loadings around 0.39, indicating perceived educational barriers. Factor 4 balances the perceived value of health economics in the curriculum with the challenges of integrating it, with loadings between 0.35 and -0.27. This analysis underscores the need for comprehensive educational strategies: (1) Strengthening Core Knowledge: Factor 1 highlights the need to enhance students' foundational understanding through targeted interventions. (2) Fostering Positive Attitudes: Factor 2 suggests focusing on shaping positive attitudes toward health economics' relevance to career development. (3) Overcoming Institutional Barriers: Factor 3 stresses the importance of addressing resource and faculty constraints. (4) Curriculum Balance: Factor 4 emphasizes the need to carefully integrate health economics into the curriculum while managing competing academic priorities. These strategies are crucial for improving health economics and pharmacoeconomics education.

Discussion

Evolving dynamics in health economics and pharmacoeconomics education

The demographic trends observed among health professions students in Southwest China reflect significant changes in the landscape of health economics and pharmacoeconomics education. The predominant participation of female students signifies a transformative gender balance, challenging traditional norms in these fields. This shift may indicate broader societal changes and a more inclusive educational environment in healthcare-related disciplines. The focus on younger undergraduates suggests that concepts of health economics are being introduced early in their academic careers, potentially shaping their professional outlook and expertise [19].

Moreover, the high proportion of pharmacy students, constituting 45.58% of the participants, could have influenced the results, particularly in areas related to pharmaceutical economics and expenditure. Pharmacy students generally receive more in-depth education in these areas, as reflected by their higher knowledge scores in the 'Pharmaceutical Expenditure' domain. This concentration of pharmacy students may introduce a bias, favoring results that are more representative of pharmaceutical knowledge rather than general health economics proficiency. To account for this, subgroup analyses were performed to compare the knowledge levels of pharmacy and non-pharmacy students, confirming that while pharmacy students excel in certain areas, trends in general economic comprehension hold across disciplines. Nevertheless, this highlights the importance of tailored educational strategies that address both general and specialized knowledge across diverse health disciplines.

A comparison with other studies reveals notable trends. Research from Japan and South Korea shows similar gender disparities and the early integration of health economics into academic programs [20, 21]. In the UK, where pharmacoeconomics is more systematically embedded in medical and pharmacy curricula, proficiency levels are comparable, especially in

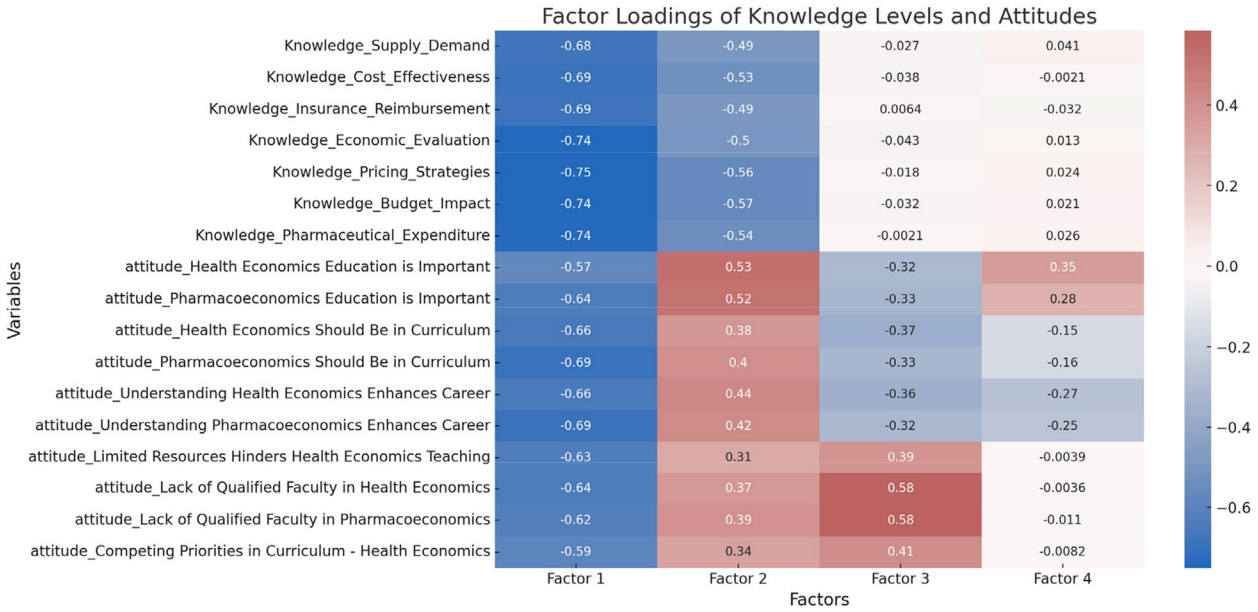


Fig. 4 Defining educational perspectives: factor loadings in health economics and pharmacoeconomics

pharmaceutical fields [22]. However, studies from Western countries like the UK and the US report higher baseline proficiency in health economics, likely due to their more established programs [23]. These global findings highlight the growing inclusion of health economics in education while emphasizing the need for localized strategies to address region-specific needs.

The diversity in academic backgrounds, including disciplines like Public Health and Clinical Medicine, underscores an interdisciplinary approach in understanding economic impacts on healthcare. This indicates a paradigm shift towards a more integrated curriculum, preparing students to tackle complex health challenges with a comprehensive understanding of both clinical and economic aspects. These evolving educational trends are pivotal for cultivating future healthcare professionals who are adept at navigating the increasingly complex intersections of healthcare delivery, economics, and policy [24].

Enhancing educational approaches in health economics and pharmacoeconomics

A quantitative analysis of student proficiency in health economics and pharmacoeconomics highlights the necessity for sophisticated educational methodologies. Observations indicate a moderate understanding in areas such as 'Supply and Demand,' 'Cost Effectiveness,' and 'Insurance Reimbursement.' This implies a foundational grasp among students. Nonetheless, a prevalent inclination toward mere basic proficiency, particularly in intricate fields like 'Economic Evaluation,' 'Pricing Strategies,' and 'Budget Impact,' necessitates specialized educational interventions. These should be designed to augment comprehension and analytical aptitudes, transcending basic knowledge to cultivate expertise in these pivotal areas [2].

Furthermore, the impact of academic advancement on students' understanding, especially in 'Supply and Demand' and 'Cost Effectiveness,' underscores the significance of a progressively complex curriculum. Such a curriculum would initially introduce elementary concepts, subsequently escalating to more complex topics, thus ensuring thorough subject matter comprehension. Moreover, the minimal gender differences observed denote an inclusive educational setting that offers equal opportunities, a vital factor in fostering gender equality in healthcare education. The varied proficiency levels in 'Pharmaceutical Expenditure' further necessitate diverse teaching methods, including case-based learning and practical applications, to address disparate understanding levels and achieve uniform knowledge acquisition across all sectors. These insights indicate the crucial role of ongoing curriculum evaluation and modification in addressing the dynamic educational requirements of students in these fields [25].

The finding that only 58% of students received formal education in health economics or pharmacoeconomics indicates that these subjects are not universally integrated into the health professions curriculum in Southwest China. While some students may have encountered these topics through elective courses or seminars, the lack of formal instruction suggests that there is a need for more structured and consistent educational strategies. Expanding the formal inclusion of health economics and pharmacoeconomics into the curriculum could better prepare students for the economic aspects of healthcare.

Furthermore, a new guiding major, Pharmaceutical Economics and Management, was recently introduced by China's Ministry of Education. This program emphasizes the integration of pharmaceutical sciences, economics, and management, aiming to cultivate professionals with expertise in both drug economics and policy, aligned with national healthcare priorities. By incorporating courses like health economics, medical statistics, and real-world research design, universities can address the growing need for interdisciplinary talent in health economics and pharmacoeconomics [26]. These findings provide essential guidance for educational institutions preparing future healthcare professionals. Implementing these strategies will better prepare students for the multifaceted aspects of healthcare delivery, economics, and policy, equipping them to effectively confront the healthcare industry's challenges.

Integrating knowledge and attitudes in health economics education

The analysis of Spearman correlation heatmaps offers an intricate understanding of the interrelationships within health economics and pharmacoeconomics education, thereby informing potential enhancements to the curriculu m. The observed strong correlations among various knowledge domains indicate a cohesively structured conceptual framework. For instance, proficiency in 'Supply_Demand' often correlates with adeptness in 'Cost_Effectiveness.' This significant interconnectivity underscores the necessity for an integrated teaching approach that ensures a holistic understanding across these interlinked areas. Furthermore, the correlation between specific knowledge sectors and corresponding attitudes towards the significance of health economics education is of particular interest. Despite a relatively low correlation between knowledge levels and the perceived importance of health economics education, the moderate to high correlations among attitudes reflect a collective acknowledgment of the practical relevance and career

benefits of pharmacoeconomics education among students [27].

This correlation pattern reveals the students' academic comprehension and their recognition of the real-world application and professional merits of these disciplines. The weaker links between particular knowledge domains and the perceived importance of health economics education present an opportunity for educators to fortify these connections. By enriching the curriculum with content that explicitly highlights the practical relevance and career implications of health economics and pharmacoeconomics, educators can cultivate a deeper appreciation and enthusiasm for these fields in students. Additionally, the strong correlations within attitude-related aspects suggest a student consensus on the essential integration of these subjects into their academic curriculum. This collective sentiment, combined with the awareness of institutional barriers, advocates for educational strategies that focus not only on knowledge transmission but also actively address perceived challenges, effectively and cohesively embedding these vital subjects in health economics education [28].

Interpretative insights from regression analysis on health economics education

The multifaceted regression analyses conducted on the dynamics of knowledge and attitudes in health economics education yield profound insights that have significant implications for pedagogical strategies. The observed gender differences in understanding supply and demand concepts, with males displaying higher proficiency, underscore the need for gender-sensitive approaches in teaching these economic principles. This finding suggests underlying differences in learning styles or interests that educators should consider. The evident academic learning curve, wherein proficiency increases with each academic year, highlights the progressive nature of learning in health economics. This progression necessitates a curriculum design that incrementally builds complexity, allowing students to develop a deeper understanding as they advance through their educational journey [29].

Moreover, the intricate relationship between various knowledge domains and educational attitudes reveals the complex nature of learning in health economics. Students with higher expertise in Economic Evaluation tend to value the education in this field more, suggesting that deeper knowledge leads to a greater appreciation of the subject's relevance. Contrarily, proficiency in certain areas inversely correlates with the perceived importance of health economics, indicating potential mismatches between curriculum content and students' perceived value [30]. This necessitates curriculum adjustments to align the teaching of these domains more closely with their practical significance.

The positive correlation between attitudes towards pharmacoeconomics education and career prospects underscores the role of curriculum design in shaping professional expectations. Addressing concerns about overlapping academic priorities and insufficient specialized faculty could significantly enhance students' perceptions of the career benefits of health economics. Lastly, demographic factors, such as gender and age, influence attitudes toward health economics education, revealing generational shifts and gender-specific preferences. These insights demand a nuanced approach to curriculum development, one that considers demographic variations and evolves in response to changing student needs and expectations in health economics and pharmacoeconomics education [31].

Educational strategies inferred from PCA on health economics knowledge

Principal Component Analysis (PCA) has provided a multifaceted view of students' knowledge and attitudes in health economics and pharmacoeconomics, revealing varied levels of comprehension and underscoring the need for customizable and stratified educational approaches. The broad spread of data along the PCA's primary axis reflects diverse student understanding, necessitating tailored educational methods that cater to different learning stages, from beginners to more advanced learners. The reduced variance seen in the second principal component indicates subtler knowledge distinctions among students, which may benefit from more focused and targeted teaching strategies.

The diminishing explanatory power of subsequent principal components suggests that a limited set of fundamental concepts currently dominates the educational focus, possibly leading to a plateau in foundational knowledge mastery. This observation is critical for curriculum planners, suggesting the expansion of course content to cover an increased range of complex subjects, thereby deepening both theoretical and practical understanding in health economics. The integration of advanced topics and real-world applications is essential to mirror the evolving dynamics of the healthcare industry.

Enhancing the curriculum in this manner promises to develop students' critical thinking and analytical capabilities, preparing them to effectively contribute to health policy formulation and economic decision-making. This alignment of educational content with the healthcare sector's current and future needs is paramount to cultivating well-prepared professionals equipped to address the intricate challenges of healthcare economics [32].

Factorial insights for adaptive curriculum development in health economics

The factor analysis undertaken to assess perceptions in health economics and pharmacoeconomics education sheds light on underlying educational dimensions, necessitating adaptive curriculum development. The inverse correlation with foundational knowledge identified by Factor 1 may signal a shift in educational needs, advocating for a dynamic curriculum that evolves from imparting basic principles to promoting applied, analytical skills [33]. Meanwhile, Factor 2's positive loadings on the value of education underscore the influence of student engagement on learning outcomes, indicating the need for curricular content that is both relevant and motivating [34]. Factor 3's emphasis on external barriers to education, such as resource limitations and faculty qualifications, points to the pivotal role of institutional support in enhancing the learning environment. Factor 4 brings attention to the delicate balance between the value of health economics in the curriculum and the challenges posed by integrating it with other academic priorities, pointing to the need for a curriculum that aligns educational content with real-world applicability while managing practical constraints [35]. Collectively, these factors argue for a multi-dimensional, responsive educational approach that is attuned to the diverse needs and perceptions of students, integrating foundational knowledge with practical skills, while also addressing resource adequacy and faculty expertise. This approach should foster an educational dialogue that aligns curricular objectives with operational practicality, ensuring the curriculum's relevance and effectiveness in preparing students for the complexities of health economics [36].

Study limitations and future directions

This study has several limitations. One major limitation is the reliance on self-reported data, which may introduce response bias as participants might overestimate or underestimate their knowledge in certain areas. Additionally, the cross-sectional design restricts the ability to infer causality between educational factors and knowledge levels. Future longitudinal studies would be beneficial to observe the progression of knowledge over time. The disproportionate representation of pharmacy students may have influenced the results, particularly in pharmaceutical-related domains. Furthermore, given the regional focus of this study in Southwest China, caution should be exercised in applying the findings to other geographical or educational contexts. To address these limitations, future research should incorporate more diverse samples and objective measures to provide a more comprehensive understanding of educational needs across different health disciplines.

Conclusion

This study assessed the knowledge and proficiency of health professions students in Southwest China in health economics and pharmacoeconomics. The findings revealed moderate proficiency in key areas such as 'Supply and Demand,' 'Cost Effectiveness,' and 'Insurance Reimbursement,' with academic year influencing comprehension—senior students showing greater understanding. Minimal gender differences were observed, and pharmacy students demonstrated higher proficiency in pharmaceutical-related areas, like 'Pharmaceutical Expenditure.' However, gaps in complex areas like 'Economic Evaluation' and 'Budget Impact' indicate the need for targeted educational interventions. Regression analysis highlighted a stronger appreciation for health economics among students proficient in 'Economic Evaluation,' while areas like 'Cost Effectiveness' showed inverse relationships with attitudes. The introduction of the 'Pharmaceutical Economics and Management' program by the Ministry of Education underscores the growing importance of these fields. These findings suggest the need for localized educational strategies to prepare students for economic decision-making in healthcare.

Abbreviations

ANOVA	Analysis of Variance
PCA	Principal Component Analysis
SPSS	Statistical Package for the Social Sciences
K-Means	K-Means Clustering
IBM	International Business Machines Corporation

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12909-024-06214-3.

Supplementary Material 1.

Acknowledgements

The authors would like to express heartfelt gratitude to the institutions and individuals whose support was crucial in this study. Special thanks to Guizhou Medical University, particularly the Pharmaceutical Science School and School of Public Health, and to Professor Zhihao Yang for his invaluable assistance. We also acknowledge Sichuan University, especially the West China Pharmaceutical College and School of Public Health, Chongqing Medical University, Yunnan Traditional Chinese Medicine University, Dali University, Kunming University of Science and Technology, and Kunming Medical University for their significant contributions. Our sincere appreciation extends to the staff and faculty members at these institutions for their guidance and insights, which were instrumental in enhancing our researc's hdepth and quality.

Authors' contributions

Jian Yang and Quanzhi Wei conceived the study, led the study design, data collection, analysis, and manuscript drafting. Xin Yang contributed to the study design, data analysis, and interpretation, and provided critical manuscript revisions. Jingyi Jiao, Zaixian Yang, and Fan Li were involved in data collection, analysis, and interpretation, offering valuable feedback. Fan Li and

Jian Yang, as senior authors, provided guidance and oversight throughout the research process. All authors read and approved the final manuscript.

Funding
This research was supported by the following: [1] The Research Project on Education and Teaching of Kunming Medical University, 2023 (Project Number: 2023-JY-Y-090); [2] The project "Exploring the Establishment of a Model Modern Biomedical Industry College" (Project Number: JG2023001); [3] First-Class Discipline Team of Kunming Medical University (Research and Practice Team of Pharmaceutical Policy), 2024 (Project Number: 2024XKTDPY21); [4] Undergraduate Teaching Quality and Teaching Reform Project of Kunming Medical University, 2024 (Course: Clinical Pharmacokinetics, Project No. 2024KCSZS-FXM022); [5] Yunnan Provincial Department of Education, Provincial First-Class Undergraduate Course, Offline First-Class Course (Pharmaceutical Administration), Project No. 2019-2-125. These funding bodies had no role in the study's design, data collection, analysis, interpretation, or manuscript writing.

Data Availability
No datasets were generated or analysed during the current study.

Declarations

Ethics Approval and Consent to Participate
This study was conducted with ethical approval from the biomedical ethics committee of Kunming Medical University, Kunming, Yunnan, China[NO.202300279]. Information about the study's objectives and contents, and the right to refuse participation, was provided to all participants before data collection. Written informed consent was obtained from all participants, and for those under 16 years old, consent was obtained from their legal guardians. All methods followed relevant guidelines and regulations.

Consent for publication
Not Applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]School of Pharmaceutical Sciences & Yunnan Provincial Key Laboratory of Pharmacology for Natural Products, Kunming Medical University, Kunming, Yunnan 650500, People's Republic of China. [2]Yunnan Provincial Center for Drug Policy Research, Kunming, Yunnan 650500, People's Republic of China. [3]College of Modern Biomedical Industry, Kunming Medical University, Kunming, Yunnan 650500, People's Republic of China. [4]Technology Transfer Center, Kunming Medical University, Kunming, Yunnan 650500, People's Republic of China. [5]The Third Affiliated Hospital, Kunming Medical University, Kunming, Yunnan 650500, People's Republic of China.

Received: 22 July 2024 Accepted: 17 October 2024
Published online: 31 October 2024

References

1. Tappenden P, Chilcott JB. Avoiding and identifying errors and other threats to the credibility of health economic models. Pharmacoeconomics. 2014;32(10):967–79. https://doi.org/10.1007/s40273-014-0186-2.
2. Tonin FS, Aznar-Lou I, Pontinha VM, Pontarolo R, Fernandez-Llimos F. Principles of pharmacoeconomic analysis: the case of pharmacist-led interventions. Pharm Pract (Granada). 2021 Jan-Mar;19(1):2302. https://doi.org/10.18549/PharmPract.2021.1.2302.
3. Adunlin G, Skiera J, Cupp CS, Ali AA, Afeli SAY. The State of Pharmacoeconomics Education in the Doctor of Pharmacy Curriculum amid the Changing Face of Pharmacy Practice. Healthcare (Basel). 2023;11(22):2923. https://doi.org/10.3390/healthcare11222923.
4. Li M, Schulz R, Wang J, Lu ZK. Pharmacy students' intentions to utilize pharmacoeconomics, pharmacoepidemiology, and health outcomes in future jobs. Curr Pharm Teach Learn. 2019Oct;11(10):995–1001. https://doi.org/10.1016/j.cptl.2019.06.001.
5. Rascati KL, Drummond MF, Annemans L, Davey PG. Education in pharmacoeconomics: an international multidisciplinary view. Pharmacoeconomics. 2004;22(3):139–47. https://doi.org/10.2165/00019053-200422030-00001.
6. Elsisi GH, Ahmed N. Fourth Pharmacoeconomics International Masterclass: Vital Initiative. Glob J Med Therap. 2021;3(1):6–10. https://doi.org/10.46982/gjmt.2021.102.
7. Suzuki Y, Tsunekawa K, Takeda Y, Cleland J, Saiki T. Impact of medical students' socioeconomic backgrounds on medical school application, admission and migration in Japan: a web-based survey. BMJ Open. 2023;13(9). https://doi.org/10.1136/bmjopen-2023-073559.
8. Kim N, Yun HW, Park J, Nari F, Jin Wang H, Jun JK, Choi KS, Suh M. Socioeconomic differences in the perception of inequalities in healthcare utilization and health in South Korea. Prev Med Rep. 2023;36: 102445. https://doi.org/10.1016/j.pmedr.2023.102445.
9. Alzarea AI, Khan YH, Alanazi AS, Butt MH, Almalki ZS, AlAhmari AK, Alsahali S, Mallhi TH. Barriers and Facilitators of Pharmacoeconomic Studies: A Review of Evidence from the Middle Eastern Countries. Int J Environ Res Public Health. 2022;19(13):7862. https://doi.org/10.3390/ijerph19137862.
10. Hammad EA, Mousa R, Hammad AA, Al-Qudah M. Awareness, knowledge, and attitudes of health professions students toward health economics and pharmacoeconomics education in Jordan. Curr Pharm Teach Learn. 2020;12(9):1072–80. https://doi.org/10.1016/j.cptl.2020.04.027.
11. Yue X, Li Y, Wu J, Guo JJ. Current Development and Practice of Pharmacoeconomic Evaluation Guidelines for Universal Health Coverage in China. Pharmacoeconomics. 2020. https://doi.org/10.1016/j.vhri.2020.07.580.
12. Singh N, Shukla R, Acharya S, Shukla S. The past, present, and future of health economics in India. J Family Med Prim Care. 2022Dec;11(12):7513–6. https://doi.org/10.4103/jfmpc.jfmpc_2266_21.
13. Hammad EA, Mousa R, Hammad AA, Al-Qudah M. Awareness, knowledge, and attitudes of health professions students toward health economics and pharmacoeconomics education in Jordan. Curr Pharm Teach Learn. 2020Sep;12(9):1072–80. https://doi.org/10.1016/j.cptl.2020.04.027.
14. Postma MJ, Kwik JJ, Rutten WJ, de Jong-van den Berg LT, Brouwers JR. Agreement between guidelines for pharmaco-economic research and never-before-published health-economics evaluations. Ned Tijdschr Geneeskd. 2002;146(23):1082–7.
15. Gavaza P, Shepherd M, Shcherbakova N, Khoza S. The state of health economics and pharmaceoconomics research in Russia: A systematic review. J Pharm Health Serv Res. 2010;1:113–21. https://doi.org/10.1111/j.1759-8893.2010.00023.x.
16. Eljilany I, El-Dahiyat F, Curley LE, Babar ZU. Evaluating quantity and quality of literature focusing on health economics and pharmacoeconomics in Gulf Cooperation Council countries. Expert Rev Pharmacoecon Outcomes Res. 2018Aug;18(4):403–14. https://doi.org/10.1080/14737167.2018.1479254.
17. D'Egidio V, Patrissi R, De Vivo G, Chiarini M, Grassi MC, La Torre G. Global Health Professions Student Survey among Healthcare students: a cross sectional study. Ann Ig. 2020 Jul-Aug;32(4):419–429. https://doi.org/10.7416/ai.2020.2365.
18. Dixon, Thomas., Mickaël, Hiligsmann., Denny, John., Ola, Ghaleb, Al, Ahdab., Hong, Li., Hong, Li. (2019). Pharmacoeconomic Analyses and Modeling. 261–275. https://doi.org/10.1016/B978-0-12-814276-9.00018-0.
19. Murakami K, Kuriyama S, Hashimoto H. Economic, cognitive, and social paths of education to health-related behaviors: evidence from a population-based study in Japan. Environ Health Prev Med. 2023;28:9. https://doi.org/10.1265/ehpm.22-00178.
20. Dronina Y, Yoon YM, Sakamaki H, Nam EW. Health System Development and Performance in Korea and Japan: A Comparative Study of 2000–2013. J Lifestyle Med. 2016Mar;6(1):16–26. https://doi.org/10.15280/jlm.2016.6.1.16. (Epub 2016 Mar 31).
21. Oppong R, Mistry H, Frew E. Health economics education in undergraduate medical training: introducing the health economics education (HEe) website. BMC Med Educ. 2013Sep;13(13):126. https://doi.org/10.1186/1472-6920-13-126.
22. Gray E, Lorgelly PK. Health economics education in undergraduate medical degrees: an assessment of curricula content and student knowledge. Med Teach. 2010;32(5):392–9. https://doi.org/10.3109/01421590903480113.

23. Farid S, Baines D. Pharmacoeconomics Education in the Middle East and North Africa Region: A Web-Based Research Project. Value Health Reg Issues. 2021Sep;25:182–8. https://doi.org/10.1016/j.vhri.2021.05.006.
24. Rosselli D. The, "Pharmacoeconomics" of Delivering Babies. Value Health Reg Issues. 2022Jul;30:168–9. https://doi.org/10.1016/j.vhri.2022.05.005.
25. Haines T, Isles R, Jones A, Jull G. Economic Consequences in Clinical Education. Focus on Health Professional Education. 2011Mar;12(3):53–63. https://doi.org/10.3316/ielapa.908656457525740.
26. Ministry of Education General Office. (2023, December 25). Notice on the Issuance of the "Guidelines for Professional Programs for Talent Cultivation in Health Services and Health Industry" (Document No. [2023] 26). Ministry of Education of the People's Republic of China. Retrieved from http://www.moe.gov.cn/srcsite/A08/moe_740/s3864/202401/t20240104_1098007.html
27. Amrita., Ram, Babu, Roy. Integrating knowledge, attitude and practice in decision model for maternity care in emerging economy. (2015).71–78. https://doi.org/10.1109/GHTC.2015.7343957
28. Shojaeezadeh D, Heshmati H. Integration of Health Education and Promotion Models for Designing Health Education Course for Promotion of Student's Capabilities in Related to Health Education. Iran J Public Health. 2018Sep;47(9):1432–3.
29. Ridhwan MM, Nijkamp P, Ismail A, M Irsyad L. The effect of health on economic growth: a meta-regression analysis. Empir Econ. 2022;63(6):3211–3251. https://doi.org/10.1007/s00181-022-02226-4.
30. Worsley A, Wang WC, Yeatman H, Byrne S, Wijayaratne P. Does school health and home economics education influence adults' food knowledge? Health Promot Int. 2016Dec;31(4):925–35. https://doi.org/10.1093/heapro/dav078.
31. Francisca, Rosendo, de, Carvalho, e, Silva., Marta, Simões., João, Sousa, Andrade. Health investments and economic growth: a quantile regression approach. International Journal of Development Issues, (2018).;17(2):220–245. https://doi.org/10.1108/IJDI-12-2017-0200
32. Siddique HM, Mohey-ud-din G, Kiani A. Health, education and economic growth nexus: evidence from middle income countries. Global Social Sciences Review: GSSR. 2018;3(4):68–86.
33. Bhowmik D. Determinants of India's health expenditure: an econometric analysis. International Journal on Recent Trends in Business and Tourism (IJRTBT). 2020;4(1):13–23.
34. Quintero RS, Ruiz RL, Legrá GN, Legrá MN. Una mirada a la economía de la salud. Revista Información Científica. 2017;96(5):978–86.
35. Shakila, Khanum., Bilal, Anjum, Ansari., Anila, Hussain., Huma, Qasim. Factor analysis of socio-economic determinants of diseases. International Journal of Advances in Scientific Research, (2016).;5(1):13–17. https://doi.org/10.14419/IJSW.V5I1.6611
36. Gurtner S. An analysis of the influence of framework aspects on the study design of health economic modeling evaluations. Eur J Health Econ. 2013Apr;14(2):221–30. https://doi.org/10.1007/s10198-011-0363-6.

Publisher's Note

BMC Medical Education

RESEARCH Open Access

Career choices and career mobility of chinese pharmacy graduates: a cross-sectional study

Hongyu Gong[1], Yan Li[2], Wenyue Duan[3], Ying Chen[4], Jianyu Zhou[3,5*] and Shuming Cao[3,5,6*]

Abstract

Background This study investigates the career mobility and employment status of Chinese pharmacy graduates, aiming to address workforce challenges and career planning within the pharmacy profession. Given the rapid expansion of the healthcare sector in China, understanding factors influencing job changes and career decisions is crucial for ensuring workforce stability and professional development. The study covers pharmacy graduates enrolled at Kunming Medical University from 1996 to 2012.

Methods A cross-sectional online survey collected data from 231 respondents in April 2022, focusing on demographic characteristics, career choices, job changes, and the factors influencing these decisions. Multivariate logistic regression and Chi-square (χ^2) tests were used to analyze the data, with a particular focus on the pharmacy practice, pharmaceutical sciences, and clinical pharmacy.

Results Of the 231 respondents, 52.4% ($n=121$) had never changed jobs, and 32.5% ($n=75$) had changed jobs once or twice. The majority preferred working in governmental institutions ($n=146$, 63.2%), with fewer graduates choosing domestic ($n=48$, 20.8%) or overseas pharmacy enterprises ($n=20$, 8.7%). Income and personal growth were significant factors influencing job changes ($p<0.05$). Respondents who had changed jobs three or more times placed a greater emphasis on personal growth (Relative Risk Ratio [RRR]: 2.96; 95% Confidence Interval [CI]: [1.6, 5.49]).

Conclusions This study reveals significant variations in job mobility and career choices among Chinese pharmacy graduates. Graduates in governmental institutions showed lower mobility, while those in pharmacy enterprises and non-pharmaceutical sectors experienced higher job changes. Income, personal growth, and work environment were key factors.

Keywords Pharmacy, Career Choice, Career mobility, Job changes; China

*Correspondence:
Jianyu Zhou
zjykm9206@163.com
Shuming Cao
15969477153@163.com
[1]Technology Achievement Transformation Center , Kunming Medical University, Chunrong West Road 1168, Kunming City, China
[2]The Third Primary School in Chenggong District, No. 1246, Ruixiang West Street, Kunming City, China
[3]School of Pharmaceutical Science & Yunnan Key Laboratory of Pharmacology for Natural Products, Kunming Medical University, Chunrong West Road 1168, Kunming City, China
[4]Faculty of Humanities and Management, Kunming Medical University, Chunrong West Road, Kunming City 1168, China
[5]College of Modern Biomedical Industry, Kunming Medical University, Chunrong West Road, Kunming City 1168, China
[6]Kunming Medical University, Chunrong west road 1168, Kunming City, China

Introduction

In the continuously evolving healthcare landscape, pharmacy professionals play a critical role in ensuring the safe and effective use of medications [1]. Many studies have examined the factors influencing career choices - defined here as the jobs, roles, or organizations pharmacy graduates choose - and career mobility across various healthcare professions, including pharmacy [2, 3]. Research has demonstrated that job satisfaction, income, and opportunities for professional development are key drivers of career decisions among healthcare professionals. For example, general practitioners who report higher job satisfaction experience reduced stress and burnout, greater engagement, and longer retention in their roles [4]. Additionally, income has been identified as a critical determinant, with competitive salaries and financial incentives positively correlating with job retention and commitment [5].

In pharmacy, factors such as workplace environment, career advancement prospects, and job security significantly shape career mobility [6, 7]. However, declining interest in pursuing pharmacy as a career has raised concerns, attributed largely to factors like limited autonomy, repetitive tasks, and a lack of fulfillment in many work settings [8].

Job mobility, defined as the frequency and likelihood of professionals changing positions within or between organizations, has become a critical focus in workforce stability and satisfaction studies [9]. Frequent job changes can disrupt continuity of care, lead to inefficiencies, and challenge workforce planning, particularly in healthcare professions like pharmacy, where specialized training and experience are vital [10]. Understanding why pharmacists change jobs and how often they do so is essential for improving job satisfaction and maintaining a stable workforce [11, 12].

In China, where the healthcare system is undergoing rapid reforms, both career choices and job mobility among pharmacy graduates are particularly important. Career choices refer to the specific roles, organizations, or sectors in which graduates work, while job mobility focuses on how frequently and why individuals change these positions. The demand for highly trained pharmacy professionals is expanding, driven by the need to improve healthcare delivery and ensure workforce stability [13, 14]. However, despite this growing demand, literature on pharmacy graduates' career mobility in China is limited, with most studies focusing on clinical healthcare professionals or broader healthcare sectors rather than pharmacy specifically [15].

Previous studies on job mobility in healthcare professions suggest that the frequency and reasons for job changes are influenced by various factors, such as job satisfaction, work-life balance, and opportunities for professional development [12, 16, 17]. Rapidly growing sectors like pharmacy often see professionals seeking roles with greater autonomy, better compensation, or more fulfilling work environments [11, 18]. Factors influencing job changes include income, location, personal growth, work environment, regulations, and commuting time. These interconnected factors affect job satisfaction, career mobility, and retention, highlighting the need for targeted strategies to enhance workforce stability.

This study aims to investigate the factors influencing job changes and career choices among Chinese pharmacy graduates. By analyzing the relationship between demographic characteristics, workplace types, job-change frequency, and the factors affecting career decisions, this research provides insights into the career mobility of pharmacy professionals. The findings will contribute to workforce planning and inform policy decisions related to career development within the pharmacy profession in China.

Materials and methods

Design

The research underwent registration and approval from the Kunming Medical University (Reference No: KMMU-2022MEC201). A literature review was undertaken to develop a questionnaire to identify pharmacy graduates' career choices and career mobility and to examine the factors that influence these choices. Three pharmacy practitioners and experts were then invited to advise on the content and wording of the questionnaire. Finally, ten pharmacy graduates were randomly selected to pilot the survey to confirm whether the content needed to be modified. The final questionnaire was attached in the supplementary material (Supplementary Questionnaire).

The online survey began with a brief paragraph informing participants that their involvement was voluntary and that they were not obligated to complete the questionnaire. This paragraph also outlined the study's purpose and confirmed that the findings would be used solely for research purposes. The survey was disseminated via WeChat to pharmacy graduates of Kunming Medical University, targeting individuals who enrolled between 1996 and 2012. To ensure the accuracy and completeness of the data, the web surveyed form was set to disallow incomplete submissions. The questionnaire consisted of three parts.

The questionnaire consisted of three parts. The first part collected demographic characteristics, including participants' age, gender, graduation year, and current employment status. Additionally, it gathered information on years of work experience, and whether they were employed in a pharmacy-related position. The second section of the questionnaire focused on participants' employment conditions and job characteristics. This

section included questions about their current occupational category, workplace type, salary range, job satisfaction, and frequency of job changes. The third part included seven questions that assessed how important various subjective factors were in influencing the respondents' career choices. This part used a 5-point Likert-type scale (1: not important at all, 2: not important, 3: neutral, 4: important, 5: very important).

Key terms

This study utilized several key terms that are defined as follows:

Major In the context of pharmacy education, a major refers to a specific area of academic focus or specialization within the broader field of pharmacy. It typically encompasses a structured set of courses that students complete to gain in-depth knowledge in a particular aspect of pharmacy practice or pharmaceutical sciences. Common pharmacy majors include **pharmacy practice**, which focuses on clinical work and patient care; **pharmaceutical sciences**, which emphasizes drug development and research; and **clinical pharmacy**, which involves working directly with healthcare teams to optimize medication use for patients.

Clinical pharmacy it is centered on patient care, where pharmacists collaborate with healthcare teams in clinical settings to optimize medication regimens, ensuring treatments are safe, effective, and tailored to individual patient needs.

Research and Teaching Refers to academic or educational positions, primarily focused on conducting research or providing instruction in a university or research institution setting.

Pharmaceutical Care Includes positions that involve providing pharmaceutical care, such as in hospitals, community pharmacies, or other healthcare settings where pharmacists ensure proper medication use and patient care. Graduates from pharmaceutical practice and clinical pharmacy programs are well-equipped to work in pharmaceutical care roles.

Pharmaceutical sales Encompasses roles related to the sale and marketing of pharmaceutical products, typically in commercial settings or within pharmaceutical companies.

Pharmaceutical Production Involves positions related to the manufacturing, formulation, and processing of pharmaceutical drugs in production facilities or factories.

Pharmaceutical Administration Refers to roles involved in the regulatory and administrative aspects of the pharmaceutical industry, including positions in government agencies or regulatory bodies that oversee drug safety and compliance.

Other Pharmaceutical-related work Includes any other roles related to the pharmaceutical industry that do not fall under the previously defined categories.

Other non-pharmaceutical-related work Refers to occupations outside the pharmaceutical field, encompassing diverse sectors unrelated to pharmacy or healthcare.

Domestic Pharmacy Enterprise Refers to privately-owned pharmaceutical companies that operate within the country. These enterprises may be involved in the research, production, marketing, or distribution of pharmaceutical products and services within the domestic market.

Overseas Pharmacy Enterprise Refers to international pharmaceutical companies or organizations based outside the country. These may include multinational pharmaceutical corporations that engage in various activities such as drug development, production, and distribution on a global scale.

Others Refers to any other types of organizations or institutions not classified under the above categories. This can include non-governmental organizations, freelance work, or employment in sectors unrelated to pharmacy or healthcare.

Data collection

Data collection was conducted in April 2022. If respondents had just graduated, they would not have had enough time to determine whether they could adapt to work or change jobs. This could have resulted in outcome bias. Therefore, we chose respondents who completed the questionnaire to have graduated for at least 10 years. To ensure comprehensive coverage, graduation years were divided into "1996–2000, 2001–2005, 2006–2009, and 2010–2012," reflecting the significant increase in pharmacy-related enrollments after 2010. A cross-section of pharmacy-major graduates from 1996 to 2012 completed the questionnaire using the online survey tool Wenjuanxing. Ten graduates from each grade in three different majors (pharmacy practice, pharmaceutical sciences, and clinical pharmacy) were randomly selected based on the student identity document (ID) of the school, which was a unique identifier assigned to each student during their enrollment. If a graduate could not be contacted, a graduate with the following student

ID was chosen. The survey was initially distributed to 240 potential participants, including 170 pharmacy practice graduates from 1996 to 2012, 40 pharmaceutical sciences graduates from 2009 to 2012, and 30 clinical pharmacy graduates from 2010 to 2012. After excluding three graduates who declined to participate and six graduates who submitted incomplete responses (spending less than 40 s or selecting the same option for five consecutive questions), a total of 231 valid questionnaires were collected, resulting in a response rate of 96.25%. Sample goodness of fit was tested by G*Power3.1 using Chi-square (χ^2) tests and obtained a type *I* error value less than 0.01 which indicated a reasonable experimental design (Supplementary Table 1) [19, 20]. The factors influencing the career choices of the 231 respondents demonstrated adequate reliability, with a Cronbach's alpha of 0.72, indicating reasonable internal consistency and correlation between the items in the questionnaire. Additionally, the construct validity of the 7-item scale was supported by a Measure of Sampling Adequacy (MSA) of 0.71, confirming that the overall data was suitable for factor analysis, as MSA values above 0.5 are generally considered acceptable (Supplementary Tables 2 and Table 1).

Statistical analysis

Data were analyzed using the packages epiDisplay, ggplot2, and nnet in R 4.2.1. Categorical data were presented as frequencies and percentages. The descriptive variables in the questionnaire were tested using χ^2 tests. Based on the respondents with different job change and preferred career factors, a multivariate logistic regression was used. In addition, Fisher's exact tests were used to examine the association between categorical variables in cases where the sample size was small. The results were considered statistically significant at a p-value of less than 0.05.

Results

Graduates' demographic characteristics

The respondents' characteristics are shown in Table 2. The final sample included 161 pharmacy practice graduates (69.7%), 40 pharmaceutical sciences graduates (17.3%), and 30 clinical pharmacy graduates (13.0%). The majority of respondents ($n=121$; 52.4%) had never changed jobs, while 32.5% ($n=75$) had changed jobs once or twice, and 15.1% ($n=35$) had changed jobs three times or more. Among the respondents, approximately 50.0% had never changed jobs across all three majors. Specifically, 54.0% of pharmacy practice graduates ($n=87$ out of 161), 47.5% of pharmaceutical sciences graduates ($n=19$ out of 40), and 50.0% of clinical pharmacy graduates ($n=15$ out of 30) reported that they had never changed jobs. These percentages reflected the proportion of respondents within each major who had not experienced job changes. Of the total 231 respondents, 17.3% ($n=40$) were enrolled between 1996 and 2000, 19.0% ($n=44$) between 2001 and 2005, 22.9% ($n=53$) between 2006 and 2009, and 40.7% ($n=94$) between 2010 and 2012. Women accounted for 69.7% ($n=161$) of the total respondents. The majority of respondents were aged from 30 to 35 years ($n=86$; 37.2%).

There were significant differences in the frequency of job changes among respondents based on whether they work in pharmacy-related jobs and their age. Statistical analysis was also conducted on the working years of respondents. A significant difference in the frequency of job changes was observed between respondents who had worked for 15–20 years and those who had worked for 3–8 years. This was measured using a χ^2 test, which compared the distribution of job changes across different work experience groups. The test results indicated a statistically significant difference for both the 15–20 year group ($p=0.022$) and the 3–8 year group ($p=0.011$), suggesting that length of work experience was associated with the likelihood of job changes.

Graduates' career choices

The data on respondents' workplace type, occupational category, and salary range are shown in Tables 3 and 1, and 4. Respondents were more inclined to work in governmental institutions, including public hospitals, community health centers, universities, and government

Table 1 Frequency of job changes by Occupational Category

Occupational Category	Total (%)	No Job Changes (%)	1–2 Job Changes (%)	≥ 3 Job Changes (%)	Test Statistic	*P*-value
Scientific Research and Teaching	29 (12.5%)	19 (15.7%)	7 (9.3%)	3 (8.6%)	χ^2 (12 df) = 43.68	< 0.001***
Pharmaceutical Care	113 (48.9%)	65 (53.7%)	40 (53.3%)	8 (22.9%)		
Pharmaceutical Sales	19 (8.2%)	2 (1.7%)	7 (9.3%)	10 (28.6%)		
Pharmaceutical Production	5 (2.2%)	1 (0.8%)	4 (5.3%)	0 (0.0%)		
Pharmaceutical Administration	14 (6.1%)	8 (6.6%)	5 (6.7%)	1 (2.9%)		
Other Pharmaceutical-Related Work	11 (4.8%)	7 (5.8%)	2 (2.7%)	2 (5.7%)		
Other Non-Pharmaceutical-Related Work	40 (17.3%)	19 (15.7%)	10 (13.3%)	11 (31.4%)		

*$p=0.05$; **$p=0.01$; ***$p=0.001$;

Table 2 Job-change frequency by Pharmacy graduates' demographic characteristics

Demographic Characteristics	Total (%)	No Job Changes (%)	1–2 Job Changes (%)	≥ 3 Job Changes (%)	Test Statistic	*P*-value
Gender						
Male	70 (30.3%)	37 (52.9%)	22 (31.4%)	11 (15.7%)	χ^2 (2 df) = 4.98	0.067
Female	161 (69.7%)	84 (52.2%)	53 (32.9%)	24 (14.9%)	χ^2 (2 df) = 5.83	0.068
Age						
40–45	49 (21.2%)	33 (67.3%)	8 (16.3%)	8 (16.3%)	χ^2 (2 df) = 6.10	0.041*
35–40	47 (20.3%)	21 (44.7%)	19 (40.4%)	7 (14.9%)	χ^2 (2 df) = 4.54	0.126
30–35	86 (37.2%)	44 (51.2%)	27 (31.4%)	15 (17.4%)	χ^2 (2 df) = 4.41	0.077
25–30	49 (21.2%)	23 (46.9%)	21 (42.9%)	5 (10.2%)	χ^2 (2 df) = 7.05	0.026*
Majors						
Pharmacy	161 (69.7%)	87 (54.0%)	47 (29.2%)	27 (16.8%)	χ^2 (2 df) = 5.12	0.073
Pharmaceutical Sciences	40 (17.3%)	19 (47.5%)	15 (37.5%)	6 (15.0%)	Fisher's Exact Test	0.117
Clinical Pharmacy	30 (13.0%)	15 (50.0%)	13 (43.3%)	2 (6.7%)	Fisher's Exact Test	0.036
Pharmacy Programme Enrollment Period (Years)						
1996–2000	40 (17.3%)	25 (62.5%)	6 (15.0%)	9 (22.5%)	Fisher's Exact Test	0.065
2001–2005	44 (19.0%)	24 (54.5%)	15 (34.1%)	5 (11.4%)	χ^2 (2 df) = 6.44	0.056
2006–2009	53 (22.9%)	23 (43.4%)	21 (39.6%)	9 (17.0%)	χ^2 (2 df) = 3.89	0.159
2010–2012	94 (40.7%)	49 (52.1%)	33 (35.1%)	12 (12.8%)	χ^2 (2 df) = 9.63	0.051
Pharmacy Related Work						
Yes	182 (78.8%)	95 (52.2%)	64 (35.2%)	23 (12.6%)	χ^2 (2 df) = 7.30	0.035*
No	49 (21.2%)	26 (53.1%)	11 (22.4%)	12 (24.5%)	χ^2 (2 df) = 2.98	0.25
Years of Work Experience						
21–26	13 (5.6%)	7 (53.8%)	1 (7.7%)	5 (38.5%)	Fisher's Exact Test	0.182
15–20	50 (21.6%)	31 (62.0%)	13 (26.0%)	6 (12.0%)	χ^2 (2 df) = 6.84	0.022*
9–14	64 (27.7%)	28 (43.8%)	23 (35.9%)	13 (20.3%)	χ^2 (2 df) = 3.09	0.229
3–8	91 (39.4%)	48 (52.7%)	35 (38.5%)	8 (8.8%)	χ^2 (2 df) = 9.91	0.011*
< 3	13 (5.6%)	7 (53.8%)	3 (23.1%)	3 (23.1%)	Fisher's Exact Test	0.324
Total	231 (100%)	121 (52.4%)	75 (32.5%)	35 (15.1%)	χ^2 (2 df) = 5.87	0.052

*p = 0.05; **p = 0.01; ***p = 0.001;

Table 3 Frequency of job changes by workplace type

Workplace Type	Total (%)	No Job Changes (%)	1–2 Job Changes (%)	≥ 3 Job Changes (%)	Test Statistic	*P*-value
Governmental Institutions	146 (63.2%)	96 (79.3%)	42 (56.0%)	8 (22.9%)	χ^2 (12 df) = 53.49	< 0.001***
Domestic Pharmacy Enterprise	48 (20.8%)	12 (9.9%)	21 (28.0%)	15 (42.8%)		
Oversea Pharmacy Enterprise	20 (8.7%)	3 (2.5%)	8 (10.7%)	9 (25.7%)		
Others	17 (7.4%)	10 (8.3%)	4 (5.3%)	3 (8.6%)		

*p = 0.05; **p = 0.01; ***p = 0.001;

Table 4 Frequency of job changes by Salary Range *p = 0.05; **p = 0.01; ***p = 0.001; US: the United States

Salary Range (US$)	Total (%)	No Job Changes (%)	1–2 Job Changes (%)	≥ 3 Job Changes (%)	Test Statistic	*P*-value
245–735 (US$)	14 (6.1%)	5 (4.1%)	7 (9.3%)	2 (5.7%)	χ^2 (10 df) = 28.77	0.001***
735–1176 (US$)	98 (42.4%)	54 (44.6%)	33 (44.0%)	11 (31.4%)		
1176–2500 (US$)	94 (40.7%)	58 (47.9%)	25 (33.3%)	11 (31.4%)		
2500–4411 (US$)	10 (4.3%)	2 (1.7%)	4 (5.3%)	4 (11.4%)		
4411–5882 (US$)	6 (2.6%)	1 (0.8%)	3 (4.0%)	2 (5.7%)		
5882–8823 (US$)	9 (3.9%)	1 (0.8%)	3 (4.0%)	5 (14.3%)		

departments (n = 146; 63.2%), followed by domestic pharmacy enterprises (n = 48; 20.8%) and overseas pharmacy enterprises (n = 20; 8.7%). Other respondents chose to work for themselves, work in other sectors, or had a freelance personnel relationship (n = 17; 7.4%). A freelance person generally refered to someone who was not affiliated with any company or had an exclusive contract with a specific company. A significant difference in job-change frequency across different workplaces was identified using a χ^2 test (p < 0.001), indicating that the

type of workplace type was significantly associated with the likelihood of job changes. Among those who had never changed jobs, the majority worked in governmental institutions ($n=96$; 79.3%). However, the frequency of job changes was notably higher among those working in pharmacy enterprises. Specifically, 12.4% ($n=15$) of respondents in pharmacy enterprises had never changed jobs, 38.7% ($n=29$) had changed jobs once or twice, and 68.5% ($n=24$) had changed jobs three times or more, indicating a higher rate of job changes compared to those working in other sectors.

Most of the respondents were engaged in pharmaceutical care ($n=113$; 48.9%), and the second largest group worked in non-pharmaceutical-related jobs, accounting for 17.3% ($n=40$). These respondents included lawyers, 5G technology staff, flight attendants, a security consultant, and the self-employed. It was worth noting that despite the school's offering of a pharmaceutical sciences major, only 2.2% of respondents were engaged in pharmaceutical production. There was also a significant difference between job-change frequency and occupational category was identified using a χ^2 test ($p<0.001$), indicating that the likelihood of job changes varied significantly across different occupational groups. Among those who had never changed jobs or had changed jobs once or twice, 53.7% ($n=65$) and 53.3% ($n=40$) respectively worked in pharmaceutical care. Respondents who worked in non-pharmaceutical-related jobs were more likely to change their career ($n=11$ out of 35; 31.4%), while those working in pharmaceutical sales were the next most likely ($n=10$ out of 35; 28.6%). Finally, a χ^2 test ($p=0.001$) revealed a significant association between salary range and job-change frequency. Respondents earning 735–1176 US$ per month were the largest group across all categories ($n=98$; 42.4%). Higher salary levels (2500–8823 US$) were associated with more frequent job changes, with 14.3% ($n=5$) of those earning 5882–8823 US$ having changed jobs three times or more, compared to 0.8% ($n=1$) who never changed jobs.

Factors affecting career choices

Income was rated as the most important factor, with a mean score of 4.1 (interquartile range [IQR]: 4–5) on a 5-point Likert scale, where 1 indicated "not important at all" and 5 indicated "very important." This was followed by personal growth and work environment (mean score: 3.9, IQR: 3–5). Regional factors, such as a preference for developed cities or hometowns, had the least impact (mean score: 3.5, IQR: 3–4).

The relationship between job-change frequency and consideration factors was analyzed by logistic regression (Table 5). Respondents who had never changed jobs were used as the reference group. Those who had changed jobs once or twice showed greater concern for income (Relative Risk Ratio [RRR]: 1.62; 95% Confidence Interval [CI]: [1.05, 2.49]) but were less concerned about regional location (RRR: 0.74; 95% CI: [0.55, 0.99]). Respondents who had changed jobs three times or more were more concerned about personal growth (RRR: 2.96; 95% CI: [1.6, 5.49]), but not about the work environment (RRR: 0.55; 95% CI: [0.31, 0.98]).

Career choices and job-change frequency in high-income graduates

This study collated the income level of respondents in different occupations, and the results are shown in Supplementary Fig. 1. Since income was an important factor in career choices, research was conducted on the relationship between occupations and job-change frequency among high-income respondents. According to the 2022 China Statistical Yearbook, China's per capita gross domestic product (GDP) was US$12,600 in 2021; therefore, this study classified an income as high if it was more than US$1,176 per month. The results are shown in Table 6. Fisher's exact tests revealed that individuals

Table 5 Factors influencing graduates' choices when changing job

Frequency of Job Changes	**1 ~ 2**		**>=3**	
Outcomes	Coeff./SE	RRR (95%CI)	Coeff./SE	RRR (95%CI)
(Intercept)	-1.69/1.124	-	-4.36/1.563**	-
Income	0.48/0.22*	1.62 (1.05,2.49)	0.09/0.269	1.1 (0.65,1.86)
Region	-0.3/0.149*	0.74 (0.55,0.99)	0.04/0.212	1.04 (0.69,1.58)
Personal Growth	0.18/0.189	1.2 (0.83,1.74)	1.09/0.315***	2.96 (1.6,5.49)
Interest	-0.05/0.204	0.95 (0.64,1.42)	-0.08/0.259	0.93 (0.56,1.54)
Work Environment	0.22/0.24	1.24 (0.78,1.99)	-0.6/0.297*	0.55 (0.31,0.98)
Law and Regulation	-0.09/0.201	0.92 (0.62,1.36)	0.15/0.285	1.16 (0.66,2.02)
Commuting Time	-0.22/0.194	0.8 (0.55,1.17)	0.02/0.27	1.02 (0.6,1.74)

*$p=0.05$; **$p=0.01$; ***$p=0.001$; reference group (no job change): $p=0$; residual deviance: 425.86

Coeff./SE: This represents the coefficient (Coeff.) and standard error (SE). The coefficient shows the strength and direction of the relationship between the independent variable and the dependent variable, while the standard error measures the precision of the coefficient estimate

RRR (95%CI): Relative Risk Ratio (RRR) with a 95% Confidence Interval (CI). The RRR indicates the likelihood of an outcome occurring in one group compared to a reference group. The 95% confidence interval provides a range within which the true RRR is likely to fall, with 95% certainty

Table 6 Job changes among high-income graduates in different occupations

Frequency of Job Changes	0 (%)	1 ~ 2 (%)	>=3 (%)	*p* value
Scientific Research and Teaching	9 (14.5%)	1 (1.6%)	2 (3.2%)	0.494
Pharmaceutical Care	34 (54.8%)	12 (19.4%)	3 (4.8%)	0.002**
Pharmaceutical Sales	1 (1.6%)	7 (11.3%)	8 (12.9%)	0.018*
Pharmaceutical Production	1 (1.6%)	3 (4.8%)	0 (0.0%)	0.090
Pharmaceutical Administration	1 (1.6%)	3 (4.8%)	1 (1.6%)	0.192
Other Pharmaceutical-Related Work	2 (3.2%)	2 (3.2%)	1 (1.6%)	0.324
Other Non-Pharmaceutical-Related Work	14 (22.6%)	7 (11.3%)	7 (11.3%)	0.021*

*p=0.05; **p=0.01; ***p=0.001;

with high incomes who worked in pharmaceutical care or non-pharmaceutical-related work were less likely to switch jobs (p=0.002 and p=0.021, respectively) than those who worked in pharmaceutical sales (p=0.018).

Discussion

There is an increasing number of undergraduate pharmacy students in China, leading to a growing number of graduates within the pharmacy sectors [21, 22]. However, previous research has not fully addressed career status, job selection considerations, and job-change frequency pharmacy majors [23, 24]. The findings from this cross-sectional study provide valuable insights into the career choices and mobility among Chinese pharmacy graduates.

The gender distribution among respondents shows a higher representation of females, consistent with the global trend of increasing female participation in the pharmacy profession [25, 26].

In terms of career choices, the data illustrate a preference for employment in governmental institutions, particularly public hospitals and community health centers. This preference may be attributed to factors such as perceived job security, opportunities for career advancement, and the availability of resources for professional development [26]. Conversely, a smaller proportion of Chinese graduates opted for positions in domestic or overseas pharmacy enterprises, possibly reflecting variations in job characteristics, work environments, and career trajectories across different sectors. This contrasts with findings from other regions, such as in developed countries, where pharmaceutical enterprises and industry roles are more commonly sought after due to higher pay and better opportunities for research and innovation [27, 28].

The distribution of respondents across occupations within the pharmacy profession further highlights the diversity of career paths pursued by pharmacy graduates. The majority of respondents were employed in pharmaceutical care, and they were less likely to change jobs compared to those in other sectors. Interestingly, a considerable number of respondents who were in pharmaceutical sales had a lot of changes in their jobs. Specifically, this group had a higher proportion of individuals who changed jobs multiple times, likely due to the impact of the centralized drug procurement and the emphasis on sales targets and performance-based roles [29].

Several factors were found to influence graduates' career choices and job-changing behaviors. Income emerged as a significant determinant, with higher income levels associated with reduced job-changing frequency. This finding aligns with previous research highlighting the role of financial incentives in shaping career decisions among healthcare professionals [30]. Additionally, personal growth emerged as a key motivator for graduates who had changed jobs multiple times, emphasizing the importance of opportunities for professional development and career advancement in retaining talent within the pharmacy workforce. A study on pharmacists leaving the profession also identified professional stagnation and lack of growth opportunities as critical reasons for job dissatisfaction and turnover, echoing the importance of continuous development opportunities to retain talent [12].

The results also suggest potential implications for policy and practice aimed at enhancing career satisfaction and retention among pharmacy graduates. Strategies to promote supportive work environments, provide avenues for continuous learning and skill development, and offer competitive compensation packages may help reduce the likelihood of job turnover and foster long-term career engagement among pharmacy professionals. In line with the results of our study, another study reported that higher monthly income in Romania increased the likelihood that a person was satisfied with his or her with job [11]. Almaghaslah and Alsayari found that the main factors that motivated medical representatives to join pharmaceutical sales included receiving a high salary, and further career development [31].

One limitation of this study was the small sample size of 231. Although the sample consisted of graduates from different majors, enrollment periods, age groups, and genders, insufficient data on the number of job-change graduates made it difficult to generalize from the findings, as the sample may not be representative of the

general situation. A further limitation was that the cross-sectional study designed limited data collection to a single point in time, so changes over time were not assessed. Finally there were limited published articles nationally and internationally on the career choices and career mobility of pharmacy graduates, which made comparisons and conclusions challenging. The findings of this study might be generalizable only to other countries with similar social and cultural factors.

Conclusion

This study highlights significant variations in career choices and job mobility among Chinese pharmacy graduates, with distinct trends observed across different occupational categories. Graduates working in governmental institutions showed lower job mobility, while those in pharmacy enterprises and non-pharmaceutical sectors experienced higher job-change rates. Factors such as income, and personal growth played pivotal roles in influencing career decisions, with income being the most critical determinant. These findings provide valuable insights into the career trajectories of pharmacy professionals and underscore the need for targeted workforce strategies to address mobility and retention challenges in the pharmacy sector.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12909-024-06448-1.

Supplementary Material 1

Acknowledgements
Not applicable.

Author contributions
Hongyu Gong planned the study, participated in data collection, analysis and drafted the manuscript; Yan Li and Ying Chen made substantial contributions to writing the manuscript; Wenyue Duan participated in data analysis and manuscript review; Shuming Cao and Jianyu Zhou provided over all supervision for the study. All authors reviewed the manuscript.

Funding
This study was supported by the Ministry of Education for Employment Program (20220104684), Yunnan Provincial Department of Education Scientific Research Fund Project (2024J0172), the Innovation Team of Kunming Medical University (CXTD202219) and First-Class Discipline Team of Kunming Medical University (2024XKTDPY21).

Data availability
The datasets used and analysed during the current study available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Kunming Medical University (Reference No: KMMU2022MEC201) [32]. Before interviewing each subject, the participants were clearly briefed about the aims and procedures of the research. Written informed consent was obtained from all participants of the study.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 1 March 2024 / Accepted: 3 December 2024
Published online: 18 December 2024

References

1. Wang CY, Li MZ, Huang YK, Xi XY. Factors influencing clinical pharmacists' integration into the clinical multidisciplinary care team. Front Pharmacol. 2023;14.
2. Liaw SY, Wu LT, Chow YL, Lim S, Tan KK. Career choice and perceptions of nursing among healthcare students in higher educational institutions. Nurse Educ Today. 2017;52:66–72.
3. Childs-Kean LM, Bhatt S, Egelund EF. Factors influencing career paths in Infectious Disease and/or HIV pharmacy: an exploratory study of pharmacy school graduates. Pharm Educ. 2021;21(1):577–81.
4. Le Floch B, Bastiaens H, Le Reste JY, Nabbe P, Le Floch P, Cam M, et al. Job satisfaction criteria to improve general practitioner recruitment: a Delphi consensus. Fam Pract. 2022;41(4):554–63.
5. Liu SM, Li SP, Li YJ, Wang HP, Zhao JJ, Chen G. Job preferences for healthcare administration students in China: a discrete choice experiment. PLoS ONE. 2019;14(1).
6. Wilby KJ, Rainkie DC. Global comparison of professional positioning by pharmacy programs towards prospective students. Res Social Adm Pharm. 2020;16(12):1658–63.
7. Smolina VA, Novokreshchenova IG, Novokreshchenov IV. Job satisfaction among pharmacists. Russian Open Med J. 2021;10(3).
8. Steeb DR, Cain J, Haines ST. Reconsidering pharmacy: we need to turn Career Regret into Career options. Am J Pharm Educ. 2024;88(4).
9. Hakim C. Labour mobility and employment stability. Rhetoric and reality on the sex differential in labour-market behaviour. Eur Sociol Rev. 1996;12(1):1–31.
10. Cohn A, Maréchal MA, Schneider F, Weber RA, FREQUENT JOB CHANGES CAN, SIGNAL POOR WORK ATTITUDE AND REDUCE EMPLOYABILITY. J Eur Econ Assoc. 2021;19(1):475–508.
11. Mullen R, Hassell K, Noyce P. Primary care pharmacist workforce mobility: why do pharmacists want to work in primary care and how do these reasons differ for community and hospital pharmacists? Int J Pharm Pract. 2005;13(4):281–8.
12. Aspden TJ, Silwal PR, Marowa M, Ponton R. Why do pharmacists leave the profession? A mixed-method exploratory study. Pharm Practice-Granada. 2021;19(2).
13. Li J, Xiao C, Hou J, Zhao Y, Gong H, Zhang B, et al. Clinical pharmacy undergraduate education in China: a comparative analysis based on ten universities' training programs. BMC Med Educ. 2023;23(1):83.
14. Lebovitz L, Eddington ND. Trends in the pharmacist workforce and Pharmacy Education. Am J Pharm Educ. 2019;83(1).
15. Tomo A, De Simone S. Using the job demands-resources approach to assess employee well-being in healthcare. Health Serv Manage Res. 2019;32(2):58–68.
16. Le Floch B, Bastiaens H, Le Reste JY, Lingner H, Hoffman R, Czachowski S et al. Which positive factors give general practitioners job satisfaction and make general practice a rewarding career? A European multicentric qualitative research by the European general practice research network. BMC Fam Pract. 2019;20(1).
17. Baum T. Skills, training and development within an insular labour market the changing role of catering managers in the healthcare environment. J Manage Dev. 2007;26(2):132–47.
18. Hassell K. Destination, future intentions and views on practice of British-based pharmacists 5 and 10 years after qualifying. Pharm World Sci. 2006;28(3):116–22.
19. Kang H. Sample size determination for repeated measures design using G Power software. Anesth Pain Med. 2015;10(1):6–15.
20. Brysbaert M, Stevens M. Power Analysis and effect size in mixed effects models: a Tutorial. J Cognition. 2018;1(1):9.
21. DiPiro JT, Allen DD, Lin A, Scott SA, Sorensen TD, Maine LL. Impact of Social Forces and Environmental Factors on Healthcare and Pharmacy Education:

the Report of the 2021–2022 AACP Argus Commission. Am J Pharm Educ. 2023;87(1):ajpe9452.
22. Lynn Crismon M, West-Strum DS, Dowling-McClay K, Drame I, Hastings TJ, Jumbo-Lucioni P, et al. The report of the 2021–2022 AACP Research and Graduate Affairs Committee. Am J Pharm Educ. 2023;87(1):ajpe9454.
23. Ives RC, Klein KC, Mason NA. Career and professional development services for pharmacy students. Currents Pharm Teach Learn. 2020;12(9):1110–5.
24. Alhomoud FK, AlGhalawin L, AlGofari G, AlDjani W, Ameer A, Alhomoud F. Career choices and preferences of Saudi Pharmacy undergraduates: A Cross Sectional Study. Saudi Pharm J. 2019;27(4):467–74.
25. Macias-Moriarity LZ, Sinclair SM, Walker D, Purnell M. Impostor phenomenon and grit as predictors of job satisfaction in female pharmacy Faculty. Pharm Res. 2023.
26. Liu P, Liu SM, Gong TT, Li Q, Chen G, Li SP. Job preferences of undergraduate pharmacy students in China: a discrete choice experiment. Hum Resour Health. 2021;19(1).
27. Arbab AH, Eltahir YAM, Elsadig FS, Yousef BA. Career preference and Factors Influencing Career Choice among undergraduate pharmacy students at University of Khartoum, Sudan. Pharmacy. 2022;10(1).
28. Alolga RN, Kassim SA, Dramou P. Study Destination Choice and Career Preferences of International Students at China Pharmaceutical University. Pharmacy. 2022;10(6).
29. Wang JL, Zhang SY, Wang CH, Li J, Wang R, Zhu L. The vacated space of volume/price of the drugs centralized procurement with quantity in secondary and above public hospitals of China. BMC Health Serv Res. 2024;24(1).
30. Scott A, Connelly LB. Financial incentives and the health workforce. Aust Health Rev. 2011;35(3):273–7.
31. Almaghaslah D, Alsayari A. A cross-sectional study on Saudi pharmacists working as medical representatives: what attracted them and what is keeping them in this sector—misconceptions and reality. Front Public Health. 2023;11.
32. Ehni HJ, Wiesing U. The declaration of Helsinki in Bioethics literature since the last revision in 2013. Bioethics. 2024;38(4):335–43.

Publisher's note