



# Regulation of plant architecture by a new histone acetyltransferase targeting gene bodies

Xueyong Yang<sup>1,7</sup>, Jianbin Yan<sup>1,2,7</sup>, Zhen Zhang<sup>3,4,7</sup>, Tao Lin<sup>1,2</sup>, Tongxu Xin<sup>1</sup>, Bowen Wang<sup>1,2</sup>, Shenhao Wang<sup>3</sup>, Jicheng Zhao<sup>5</sup>, Zhonghua Zhang<sup>1</sup>, William J. Lucas<sup>6</sup>, Guohong Li<sup>5</sup> and Sanwen Huang<sup>1,2</sup>

**Axillary meristem development determines both plant architecture and crop yield; this critical process is regulated by the PROLIFERATING CELL FACTORS (TCP) family of transcription factors. Although TCP proteins bind primarily to promoter regions, some also target gene bodies for expression activation. However, the underlying regulatory mechanism remains unknown. Here we show that TEN, a TCP from cucumber (*Cucumis sativus* L.), controls the identity and mobility of tendrils. Through its C terminus, TEN binds at intragenic enhancers of target genes; its N-terminal domain functions as a non-canonical histone acetyltransferase (HAT) to preferentially act on lysine 56 and 122 of the histone H3 globular domain. This HAT activity is responsible for chromatin loosening and host-gene activation. The N termini of all tested CYCLOIDEA and TEOSINTE BRANCHED 1-like TCP proteins contain an intrinsically disordered region; despite their sequence divergence, they have conserved HAT activity. This study identifies a non-canonical class of HATs and provides a mechanism by which modification at the H3 globular domain is integrated with the transcription process.**

**T**EOSINTE BRANCHED 1 (TB1), CYCLOIDEA (CYC) and PROLIFERATING CELL FACTORS (TCP) transcription factors constitute a plant-specific gene family involved in a broad range of developmental processes<sup>1</sup>. Among them, the CYC/TB1 clade of the TCP proteins have central roles in controlling development of axillary buds that give rise to either flowers or lateral shoots<sup>1,2</sup>. In maize (*Zea mays* L.), the major domestication gene TB1 suppresses branch outgrowth, a crucial architectural modification that transformed teosinte into a viable crop<sup>3</sup>. Subsequent studies on its homologues in rice<sup>4</sup> and *Arabidopsis thaliana*<sup>5</sup> identified their similar essential roles in repressing axillary bud growth.

Recently, a genome-wide binding profile uncovered a genetic pathway putatively regulated by TB1 (ref. <sup>6</sup>). The study reported that TB1 binds mainly to promoters, with only a few peaks located within gene-body regions. Nevertheless, other studies have also shown that TB1, and its homologue BRANCHED 1 in *Arabidopsis*, can bind to the gene bodies of the target genes *Tassels Replace Upper Ears1* (*Tru1*)<sup>7</sup> and *HOMEBOX PROTEIN53* (ref. <sup>8</sup>), respectively, to activate their expression. However, the mechanism underlying how the intragenic binding of CYC/TB1-like transcription factors regulates gene expression is still unclear. Understanding the conserved regulatory mechanism associated with the function of these CYC/TB1-like proteins would provide insights into their core role in signal integration of axillary bud repression. Such knowledge could broadly benefit crop breeding programmes for tailored plant architecture.

In eukaryotes, enhancers are *cis*-acting DNA sequences which, when bound by specific transcription factors, increase the

transcription in a manner that is independent of their orientation and distance relative to the transcription start site<sup>9</sup>. In *Drosophila melanogaster*, the vast majority (88%) of all enhancers were shown to be located in the vicinity of their targets, of which 30% are upstream, 22% are downstream and 36% are intragenic<sup>10</sup>. These intragenic enhancers appear to mainly (79%) regulate their host genes, with only 21% activating neighbouring genes<sup>10</sup>. The first eukaryotic intragenic enhancer to be discovered was in the immunoglobulin heavy chain gene<sup>11</sup>, and it was shown that this enhancer activity was correlated with an increase in histone acetylation and general sensitivity to digestion by DNase I<sup>12</sup>. Several models have been proposed for the action of enhancers in the regulation of transcription; here we consider the proposed ‘facilitated tracking’ mechanism. In this model, an enhancer-bound complex containing DNA-binding transcription factors and coactivators scans along the chromatin until it encounters the promoter, where a looped chromatin structure is formed. The key points of this tracking mechanism are alteration of a repressive chromatin structure and facilitation of enhancer–promoter communication<sup>9</sup>. Overall, the mechanism of transcriptional regulation by enhancers remains poorly understood<sup>13</sup>.

The dynamics of chromatin structure are strictly regulated by multiple mechanisms, including post-translational modification of histones<sup>14,15</sup>. Tail-based histone acetylation functions as docking sites for the recruitment of transcriptional regulators, whereas recent data suggest that acetylation of lysines 56 and 122 in the globular domain of histone H3 (H3K56 and H3K122) can directly alter histone–DNA interactions, thereby modulating chromatin architecture<sup>16–18</sup> and stimulating transcription<sup>19–21</sup>. In yeast, histone

<sup>1</sup>Key Laboratory of Biology and Genetic Improvement of Horticultural Crops of the Ministry of Agriculture, Sino-Dutch Joint Laboratory of Horticultural Genomics, Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing, China. <sup>2</sup>Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Genome Analysis Laboratory of the Ministry of Agriculture, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China. <sup>3</sup>College of Horticulture, Northwest A&F University, Yangling, China. <sup>4</sup>State Key Laboratory of Crop Stress Adaptation and Improvement, School of Life Sciences, Henan University, Kaifeng, China. <sup>5</sup>University of Chinese Academy of Sciences, National Laboratory of Biomacromolecules, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China. <sup>6</sup>Department of Plant Biology, College of Biological Sciences, University of California, Davis, CA, USA. <sup>7</sup>These authors contributed equally: Xueyong Yang, Jianbin Yan and Zhen Zhang.

chaperone-dependent acetylation of H3K56 by the HAT Rtt109 is required for chromatin assembly during DNA replication<sup>22–25</sup> and for chromatin disassembly during transcriptional activation<sup>19,21</sup>. H3K56 acetylation enhances the unwrapping of DNA close to the DNA entry-exit site of the nucleosome<sup>26</sup> and regulates chromatin at a higher-order level<sup>27</sup>. H3K56 acetylation also appears to be involved in transcription elongation<sup>28,29</sup>. Similarly, in humans and *D. melanogaster*, the HATs CBP and p300 mediate the acetylation of H3K56 in an Asf1-dependent manner, which is required for chromatin assembly during DNA synthesis<sup>30</sup>. H3K122 acetylation (H3K122ac) directly affects histone-DNA binding and stimulates transcription<sup>20</sup>. Recently, it was reported that a subset of active enhancers is marked by histone H3 globular domain acetylation (H3K64ac and H3K122ac)<sup>31</sup>. However, the mechanism underlying *cis* and *trans* determinants of how the histone globular domain acetylation is integrated into specific genes during transcriptional regulation remains unknown<sup>32</sup>.

Intrinsically disordered regions (IDRs) are polypeptide segments that lack sufficient hydrophobic amino acids to mediate cooperative folding, and thus lack an ordered three-dimensional structure<sup>33,34</sup>. IDRs are abundant in eukaryotic proteins, are especially prevalent in transcription factors, and were recently considered to have important roles in gene activation through the formation of biomolecular condensates (phase separation)<sup>13,35</sup>. However, the possible role and mechanism of action of IDRs in transcriptional regulation remain largely unexplored.

Recently, we identified the cucumber (*C. sativus* L.) tendril identity gene *TEN*, which belongs to the CYC/TB1 clade of the TCP gene family<sup>36</sup>. Tendrils are modified branches in which axillary meristems are inhibited from developing and climbing behaviour is acquired. To understand how *TEN* regulates target-gene expression, we investigated its genome-wide binding profiles. We show that *TEN* acts both as an intragenic enhancer-binding transcription factor and as a non-canonical HAT, acting on H3K56 and K122 for host-gene activation. Furthermore, we demonstrate that the N termini of tested CYC/TB1-like proteins contain IDRs, and that despite their sequence divergence, they have conserved HAT activity.

## Results

**Regulation of tendril identity requires N and C termini of TEN.** The cucumber *TEN* mutant *ten-1* forms modified branches instead of tendrils and therefore loses the ability to climb<sup>36</sup>. The *ten-1* mutant harbours a single point mutation (N338Y) in the C terminus of *TEN*, indicating an important function associated with this region<sup>36</sup> (Fig. 1a and Supplementary Fig. 1a,b). To knock out the *TEN* gene and further investigate functions associated with *TEN*, we used clustered regularly interspaced short palindromic repeats (CRISPR)-CRISPR-associated protein 9 (Cas9) to target amino acids 1 to 121 in the N terminus of *TEN* (the N121 region) (Fig. 1a). These *TEN*-edited plants were phenotyped, and a null mutant (*ten-2*) (Fig. 1b) displayed a complete transformation of its tendrils into lateral branches (Fig. 1c,d), equivalent to the *ten-1* mutant phenotype (Supplementary Fig. 1b). This result confirmed the function of *TEN* in control of tendril identity.

In addition to the *ten-2* null mutant, we also identified two other *TEN*-edited plants: *ten-3* has a homozygous in-frame deletion of two amino acids (Gln54 and Asp55) and *ten-4* has a homozygous in-frame deletion of 12 amino acids (Fig. 1b). We observed that in both *ten-3* and *ten-4* plants (Fig. 1e–g), some tendrils retained most of the normal tendril morphology (Fig. 1g); however, others exhibited slight morphological changes, producing axillary meristems on their tendrils (Fig. 1e–g), demonstrating the important role of *TEN* in axillary meristem inhibition during tendril development.

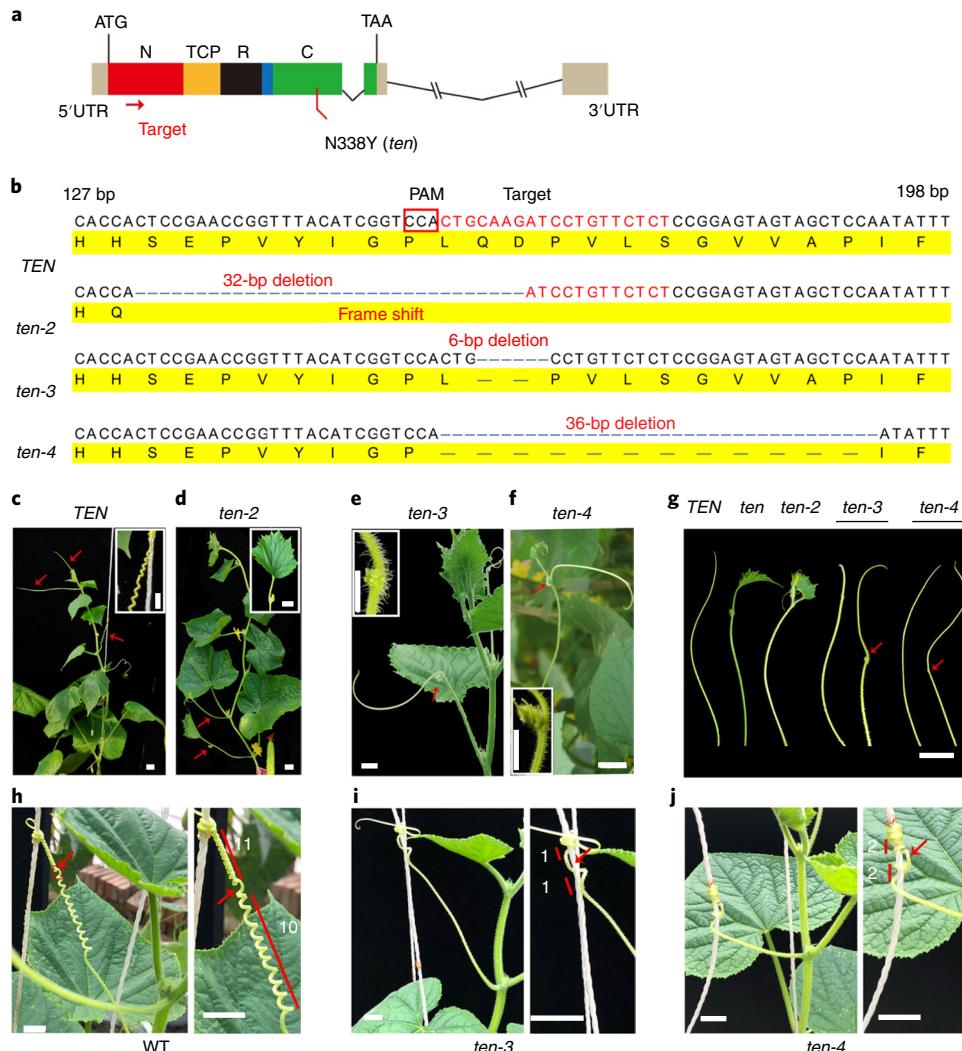
Although tendril morphology was largely unaffected in *ten-3* and *ten-4* plants, their climbing capacity was markedly altered (Fig. 1h–j and Supplementary Video 1). Wild-type (WT) tendrils

form approximately ten helical turns on each side of a perversion point<sup>37</sup> (Fig. 1h and Supplementary Video 1). Although these mutant tendrils could still attach to a support, the free coiling—formed by two oppositely handed helices—was impaired, resulting in a reduced number of helical turns on both sides of the perversion point (Fig. 1i,j and Supplementary Video 1). These results demonstrate that *TEN* controls tendril identity and climbing ability, and that both the N and C terminus are critical for its function.

**TEN C terminus is responsible for binding at intragenic regions of target genes.** To understand how the N and C termini of *TEN* affect transcription factor functions, we first assessed the global binding profiles of *TEN* by chromatin immunoprecipitation with sequencing (ChIP-seq). To this end, we raised polyclonal antibodies and confirmed antibody specificity (Extended Data Fig. 1a–c and Supplementary Table 1). We also performed immunoblot assays and confirmed that the antibody recognized recombinant full-length *TEN* but not its TCP domain (Extended Data Fig. 1d). ChIP-seq assays were performed using tendrils at the coiling stage. Two replicate experiments were performed. They shared a large number of peaks covering more than 59% of peaks in the smaller replicate (Fig. 2a and Supplementary Table 2). Follow-up analysis of genome distribution using the overlapping peaks revealed that these *TEN*-binding sites were highly enriched in intragenic regions (1,257 peaks), which accounted for approximately 74% of all peaks (65.4% in coding regions and 8.3% in intronic regions) (Fig. 2b,c). In addition to intragenic regions, a minor portion of the binding sites was distributed in intergenic regions (15.3%), with only 6.1% in promoter regions located within 5 kb upstream of the transcription start site.

A total of 637 genes, associated with 1,707 peaks were identified, most of which (474 genes; 74.4%) are intragenic targets, with only 7.2% (46 genes) being putatively regulating in promoter regions (Supplementary Table 3). Therefore, in this study, the 474 genes associated with 1,257 intragenic binding sites were designated as the *TEN*-target gene set (Extended Data Fig. 1e and Supplementary Table 4). Gene ontology analysis indicates that some genes involved in axillary bud formation—such as the *SPL* transcription factor genes<sup>38</sup> and homeobox-related transcription factor genes<sup>39</sup>—were also present in the target gene set, consistent with the role of *TEN* in tendril morphology regulation (Supplementary Table 4). Genes involved in ethylene biosynthesis and signal transduction were also significantly enriched ( $P < 0.05$ ) (Extended Data Fig. 1f). Exogenous spraying with ethephon, a plant growth regulator that is converted to ethylene in the plant, induced spontaneous tendril coiling (Extended Data Fig. 1g), consistent with a central role for ethylene in tendril coiling<sup>40</sup>.

Next, we focused on two exemplary target loci, *ACO1* (*Csa6G160180*), which encodes a 1-amincyclopropane-1-carboxylate-oxidase enzyme for ethylene synthesis and *ERF1* (*Csa7G049230*), which encodes an ethylene response factor (Fig. 2d). As revealed by chromatin precipitation with quantitative PCR (ChIP-qPCR), *TEN* was recruited to the exons of both *ACO1* and *ERF1* (Fig. 2e). To provide further genetic evidence that these intragenic targets are genes that are directly regulated by *TEN*, we selected the *ACO1* gene for in-depth investigation. *ACO1* is preferentially expressed in tendril tissue (Extended Data Fig. 1h) and its expression pattern during tendril growth showed that it was upregulated more than 4,000-fold, from the young stage to stretch stage. By contrast, the activation of *ACO1* was significantly repressed in the *ten-1* mutant (Fig. 2f), suggesting an important role for *ACO1* in tendril coiling. We also found that although the expression levels of *TEN* were only slightly upregulated from the young stage to the stretch stage, the level of *TEN* binding on the *ACO1* locus was upregulated by approximately fourfold and the expression of *ACO1* was specifically upregulated at the stretch stage (Extended Data Fig. 1i).

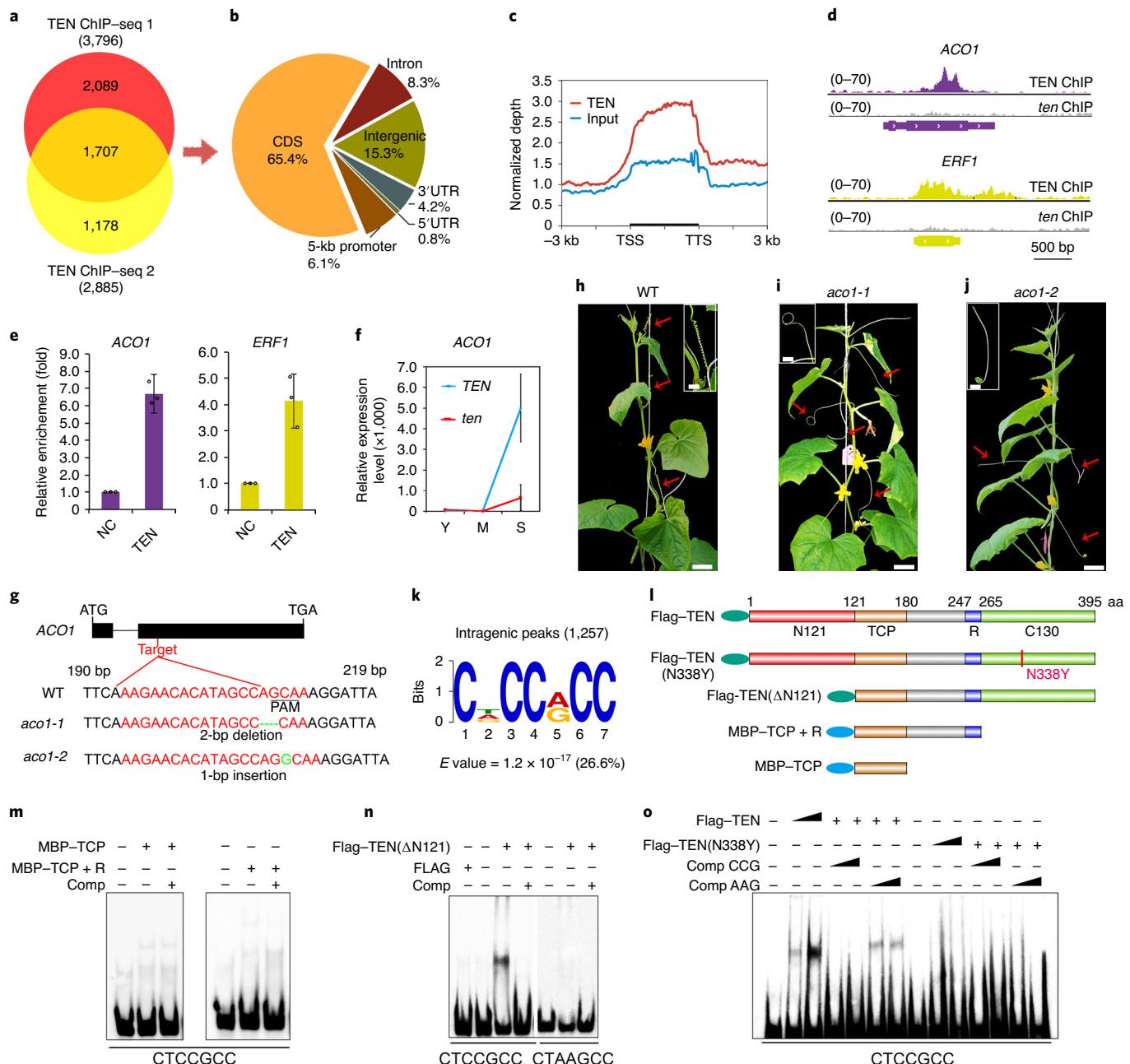


**Fig. 1 | Analysis of CRISPR-Cas9 mutants reveals the in vivo role of TEN.** **a**, Schematic illustrating the single guide RNAs (sgRNAs; red arrow) targeting the N121 region. Coloured boxes represent exons, and black lines represent introns. N338Y designates the *ten-1* mutant that forms modified tendrils. C, C-terminal domain; N, N-terminal domain; R, R domain; UTR, untranslated region. **b**, Identification of the *ten-2* allele as a null mutation carrying a 32-bp deletion, and alleles homozygous for genes encoding proteins with deletions of two amino acids (*ten-3*) and 12 amino acids (*ten-4*). Red text indicates sgRNA targets and the red box indicates the protospacer-adjacent motif (PAM) sequence. **c,d**, WT (*TEN*) plant bearing typical tendrils (red arrows), has the ability to climb (**c**), whereas the *ten-2* mutant bears modified tendrils (red arrows) and is unable to climb (**d**). **e,f**, Examples of *ten-3* (**e**) and *ten-4* (**f**) tendrils on which axillary buds have developed. **g**, Tendril phenotypes of various alleles of the *TEN* gene. **h-j**, Compared with WT plants (**h**), free coiling, formed by two oppositely handed helices, is impaired in *ten-3* (**i**) and *ten-4* (**j**) plants. The number of turns on each side of the perversion point (white numbers) indicates the degree of coiling. Arrows indicate the perversion of coiled tendrils. *ten-3* and *ten-4* plants display a substantial reduction in helical turns on both sides of the perversion points. Scale bars, 2 cm. Images in **c-j** are representative of at least five plants.

To further investigate the function of *ACO1*, we used CRISPR-Cas9 to target the *ACO1* N terminus. We identified two null mutants, *aco1-1* and *aco1-2* (Fig. 2g). Although tendril morphology was largely unaffected in *aco1-1* and *aco1-2* plants, their climbing capacity was substantially altered (Fig. 2h–j and Supplementary Video 2). WT tendrils could attach to the climbing supports and showed normal free-coiling activity (Fig. 2h and Supplementary Video 2); however, these mutant tendrils displayed irregular coiling and could not attach to their supports (Fig. 2i,j and Supplementary Video 2). These results demonstrated that *ACO1* is an authentic direct target of TEN, providing genetic evidence that TEN directly regulates its intragenic targets.

To further analyse the summits of intragenic peaks, we identified a statistically overrepresented motif, CDCCRCC (Fig. 2k). We also showed that both ChIP-seq repeats had good reproducibility and

identified the same motifs. Meanwhile, these motifs were more abundant in the intergenic peaks than other peaks (Supplementary Fig. 2). We expressed and purified a series of truncated TEN proteins, using Sf9 insect cells, to investigate TEN-binding activity on this motif (Fig. 2l and Extended Data Fig. 2). Electrophoretic mobility shift assays (EMSA) and surface plasmon resonance established that the TCP and R domains bind the previously described TCP-binding motif GGTCCC with high affinity but had no significant affinity for a 50-bp probe containing the CTCCGCC motif (Fig. 2m and Extended Data Fig. 3a–d). However, purified TEN( $\Delta$ N121) (containing the TCP and R domains, and amino acids 265 to 395 in the C terminus (the C130 domain)) and full-length TEN protein bound to the DNA fragment containing the CTCCGCC motif, supporting an essential role of the C130 region in the sequence-specific DNA-binding activity of TEN (Fig. 2n and Extended Data Fig. 3e–g).



**Fig. 2 |** TEN is a transcription factor with intragenic binding capacity. **a**, Overlap of TEN-binding sites in two TEN ChIP-seq replicates. **b**, Distribution of overlapping TEN-binding peaks in the cucumber genome. **c**, TEN-binding peaks are highly enriched in the intragenic regions of coiling tendrils. TSS, transcription start sites; TTS, transcription termination sites. **d**, Two examples of TEN-binding profiles in the gene bodies of *ACO1* and *ERF1*. **e**, ChIP-qPCR analysis of TEN recruitment to the indicated intragenic region (mean  $\pm$  s.d.,  $n=3$ ). NC, negative control. **f**, Relative *ACO1* mRNA expression levels during tendril growth detected by RT-qPCR (mean  $\pm$  s.d.,  $n=3$ ). Y, young; M, medium; S, stretch. **g**, Identification of the *aco1-1* and *aco1-2* alleles as two independent null mutations. Red font highlights sgRNA targets, and underline indicates the PAM sequence. **h-j**, Compared with WT (**h**), the tendrils in *aco1-1* (**i**) and *aco1-2* (**j**) form irregular coiling and could not attach to their supports. Arrows indicate the coiled tendrils and insets illustrate the differences in coiling between WT and mutants. Scale bars, 5 cm. **k**, The enriched motif CDCCRCC. Enrichment was analysed by MEME-ChIP, using a Fisher's exact test ( $n=1,257$ ). **l**, Schematics of WT TEN, TEN(N338Y) and three truncated proteins. aa, amino acids. **m**, EMSA showing that TCP and TCP-R do not bind to a DNA probe containing the CTCCGCC motif. Comp, competitor. **n**, Flag-TEN(ΔN121) specifically binds to DNA containing the CTCCGCC motif. **o**, Flag-TEN specifically binds to DNA containing the CTCCGCC motif, but Flag-TEN(N338Y) does not. Comp CCG, competitor (unlabelled CTCCGCC probe); comp AAG, mutant competitor (unlabelled CTAAGCC probe); closed triangle, increasing amount of protein (1 or 4  $\mu$ g) or competitor (100- or 1,000-fold probe concentration). Images in **h-j** and **m-o** are representative of experiments that were repeated at least two times with similar results.

EMSA performed with full-length TEN established that it binds both the GGTCCC motif and the CTCCGCC motif; however, the dissociation constant ( $K_d$ ) was lower for CTCCGCC motif, reflecting stronger binding to this newly identified motif (Extended Data Fig. 3f,g).

In addition, we also showed that the TEN protein could bind specifically to the CTCCGCC motif without competition from a GGTCCC probe (Extended Data Fig. 3h). Furthermore, although the N338Y mutation in the C terminus of TEN had no effect on TEN

binding to the GGTCCC motif (Extended Data Fig. 3i), it abolished binding to the CTCCGCC motif (Fig. 2o), reflecting the phenotypic changes induced by the N338Y mutation in *ten-1* (ref. <sup>36</sup>). Lastly, we established that a purified fusion protein containing the TEN C terminus and maltose-binding protein (MBP) could not bind to CTCCGCC probes, indicating that despite the essential role of the TEN C terminus in binding the CTCCGCC motif, the C terminus alone is not sufficient for this binding capacity (Extended Data Fig. 3j,k). These findings provided strong support for the notion that TEN is a transcription factor with intragenic binding capacity.

**The TEN-binding site is a novel intragenic enhancer for host-gene activation.** To investigate the role of TEN in regulating expression of intragenic target genes, the transcriptomes of candidate genes were analysed in WT versus *ten-1* mutant plants. Among these 474 genes, 132 were downregulated significantly in mutant tendrils (more than 1.5-fold change;  $P < 0.05$ ); interestingly, no gene was upregulated significantly in the mutant (less than 0.67-fold change;  $P < 0.05$ ; Supplementary Table 4). This result suggests that the intragenic binding sites (CDCCRCC) of TEN have an enhancer activity for the host genes.

To explore whether binding of TEN to the potential intragenic enhancer CDCCRCC can regulate neighbouring genes, we investigated the expression of genes flanking the intragenic targets of TEN (three genes upstream and downstream) in WT and *ten-1* mutant plants. These experiments showed that TEN activation occurred predominantly on its specific target genes (Fig. 3a). To validate these data, quantitative PCR with reverse transcription (RT-qPCR) was performed on two exemplary targets—*ACO1* and *ERF1*—and their flanking genes in tendrils of WT, *ten-1*, *ten-2* and *ten-3* plants. Our results showed that dysfunction of TEN in these mutants leads to a marked reduction in both *ACO1* and *ERF1* expression (Fig. 3b and Extended Data Fig. 4a), whereas there were no significant effects on the expression of the flanking genes located upstream or downstream of the *ACO1* and *ERF1* loci (Fig. 3b and Extended Data Fig. 4a).

To demonstrate *in vivo* that the intragenic regulatory elements bound by TEN are a novel enhancer sequence and to test whether the CDCCRCC motifs are required for the observed enhancer activity, we assayed reporter transgenic lines<sup>41</sup>. To this end, we selected *ACO1* full-length genomic DNA sequence as a candidate enhancer, and each construct contained an expression cassette with TEN or a hygromycin (*Hyg*) gene under the control of the CaMV 35S promoter and another expression cassette containing the candidate enhancer (or mutant enhancer in which we disrupted the respective motifs with point mutations), minimal promoter and luciferase (*LUC*) reporter gene (Fig. 3c and Extended Data Fig. 4b). All four constructs were integrated, independently, into the cucumber genome. Importantly, construct B (TEN+enhancer) exhibited 20-fold higher LUC activity than construct A, the negative control (*Hyg*+enhancer), and construct D (TEN+mutated enhancer) had strongly reduced LUC activity, compared with construct B (TEN+enhancer) (Fig. 3d).

To further confirm that the transformed TEN protein regulates the intragenic enhancer activity through binding directly and specifically to the intragenic enhancer, we investigated the binding capacity with ChIP-qPCR, using leaves from plants expressing one of the stably integrated constructs. We demonstrated that TEN could bind the intragenic enhancer *in vivo* and that mutation in the CTCCNCCN motif largely impaired this TEN-binding capacity (Fig. 3e). These findings demonstrate that TEN binds on a previously unidentified type of intragenic enhancer and further validate the functional importance of the CDCCRCC motifs in the enhancer.

**TEN is a non-canonical HAT.** Having demonstrated that TEN binds to intragenic enhancers of host genes via its C terminus, we next

investigated the functional role of its N terminus. PSI-BLAST analysis revealed that the N121 region shares a modest similarity (26% identity) with the transferase domain of an *Arabidopsis* HXXXD acyltransferase (At1G03940; Extended Data Fig. 5a). Considering that the intragenic enhancer-binding transcription factors might participate in shaping chromatin structure<sup>42</sup>, we speculated that TEN might have HAT activity.

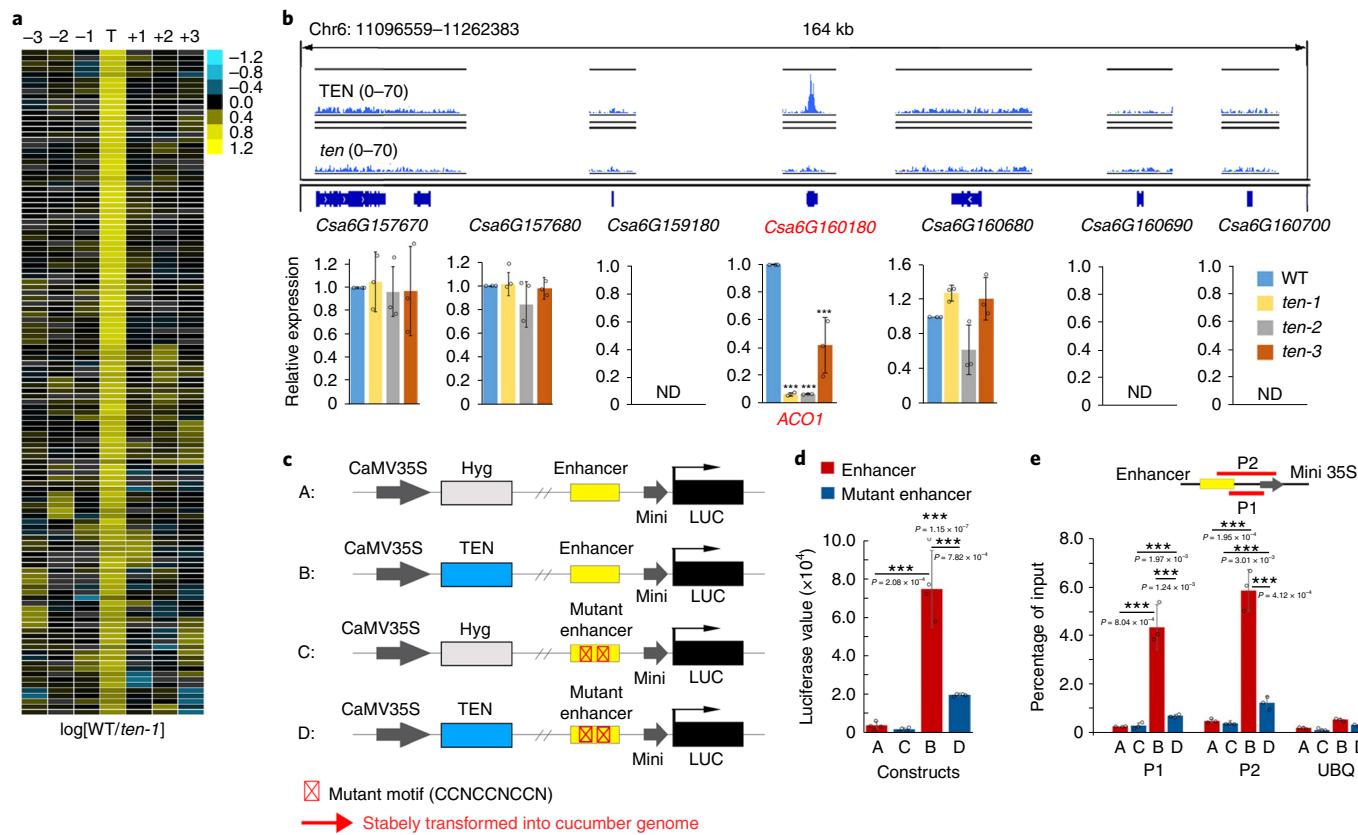
To assess this notion, we tested the activity of recombinant TEN protein and determined that it acetylated all four core histones when histone H3 or free core histones were used as substrates (Fig. 4a). Histone tetramers or octamers enhanced the acetylation ability of TEN, which occurred predominantly in histone H3 (Fig. 4a). TEN also efficiently acetylated mononucleosomes assembled on the 208-bp 5S rDNA, with a preference for nucleosomal histone H3 (Fig. 4b).

To identify the intrinsic HAT domain in TEN, we performed HAT activity assays using recombinant TEN, TEN(N338Y), TEN N121 and TEN(ΔN121) (Extended Data Fig. 5b). WT TEN, TEN(N338Y) and TEN(N121) acetylated histone H3 in H3–H4 tetramers, whereas TEN(ΔN121) had no detectable HAT activity (Fig. 4c). We next assessed the acetylation site specificity of TEN by quantitative mass spectrometry<sup>43</sup> to measure the acetylation levels of individual lysine residues on histones, using the recombinant TEN N121 protein expressed in and purified from *Escherichia coli*. Our *in vitro* assays indicated that N121 acetylated K23 and K36 in the tail domain of histone H3, and K56, K79 and K122 in its globular domain, with a preference for K56, K79 and K122 (Fig. 4d and Extended Data Fig. 5c–e).

To confirm the results of these liquid chromatography with tandem mass spectrometry (LC–MS/MS) studies, we performed a combination of *in vitro* HAT assays and immunoblotting with antibodies specific for different acetylated sites. These assays confirmed that full-length TEN and N121 both acetylated the identified lysines of histone H3, with a preference for H3K56 and K122 (Fig. 4e,f and Extended Data Fig. 5f,g). In addition, the *in vivo* acetylation patterns of H3 lysine residues by TEN were assessed by transient expression of exogenous TEN in *Nicotiana tabacum* leaves; we observed an increase in acetylation of H3K56 and H3K122 (Fig. 4g and Extended Data Fig. 5h). Furthermore, TEN overexpression led to a significant increase in nuclear H3K56ac and H3K122ac levels, as revealed by immunolabeling of *N. tabacum*, indicating that TEN acetylates chromatin-bound H3K56 and K122 *in vivo* (Extended Data Fig. 5i,j). Collectively, these *in vitro* and *in vivo* assays established that TEN N121 contains a HAT domain, which preferentially acetylates H3K56 and K122 within nucleosomes.

N121 shares no sequence homology with any other known HAT. To further study its enzymatic mechanism, we therefore focused on specific amino acid residues, including Gln14, Asp55, Ser91 and Cys104, that generally form part of the catalytic centre. Point mutations of these candidate residues (Q14N, D55N, S91A and C104A) (Fig. 4h and Supplementary Fig. 3) were tested using *in vitro* HAT assays to assess their effects on H3K56 acetylation. All mutations showed a varying degree of reduction in HAT activity. In particular, the C104A mutant protein had almost no ability to acetylate H3K56 (Fig. 4h,i). These findings demonstrate that C104 is essential, and that D55, Q14 and S91 also contribute to HAT activity. The critical role of Asp55 for the HAT activity of TEN N121 suggested a biochemical basis underlying the phenotypic changes induced by the Asp55 deletion in the *ten-3* and *ten-4* mutants.

To obtain biochemical evidence that N121 is a bona fide HAT, we next assayed steady-state kinetics of H3K56 and K122 acetylation<sup>43</sup>. Here, acetylation of both H3K56 and K122 showed saturation kinetics (Fig. 4j,k). The lower Michaelis constant ( $K_m$ ) for H3K56ac indicated a higher affinity of N121 for H3K56, whereas the catalytic efficiency of H3K56ac was about 60% of that for H3K122ac (Fig. 4l). We compared the acetylation kinetics of N121 and Rtt109, a HAT



**Fig. 3 |** TEN binds to intragenic enhancers of its target genes. **a**, Expression ratio of 132 genes directly upregulated by TEN; comparison of WT and *ten-1*, showing that TEN regulates predominantly host genes. T, intragenic target genes; -1, -2 and -3, three genes upstream of the target gene; +1, +2 and +3, three genes downstream of the target gene. **b**, Expression of the putative enhancer target gene *ACO1* and flanking genes, assayed by RT-qPCR (mean  $\pm$  s.d.,  $n=3$ ). *UBQ* was used as internal control. ND, not detected. **c**, Schematic showing the constructs used for in vivo enhancer validation, with reporter transgenic lines. *LUC*, luciferase gene. **d**, *LUC* activity of positive transgenic cucumber leaves (mean  $\pm$  s.d.,  $n=4$ ). Four independent transgenic plants per construct were used. **e**, ChIP-qPCR analysis of TEN recruitment to the indicated regions (P1 and P2) of the intragenic enhancer (mean  $\pm$  s.d.,  $n=3$ ). *UBQ* was used as internal control. Data were compared using a one-sided Student's *t*-test. \*\*\* $P < 0.01$ .

required for H3K56 acetylation in yeast. On the basis of a previous study, we established that N121 and Rtt109 exhibited similar levels of activity on the H3–H4 tetramer<sup>44</sup> (Fig. 4*j*).

We also directly compared the HAT activity of N121 with the canonical HAT P300 by quantitative mass spectrometry. These results showed that N121 acts preferentially on lysine 56 and 122 of histone H3 (Fig. 4*m* and Extended Data Fig. 6*a*). The HAT activity of TEN N121 on H3K56 and H3K122 was approximately seven times lower than that of the canonical HAT P300 (Fig. 4*m* and Extended Data Fig. 6*a*).

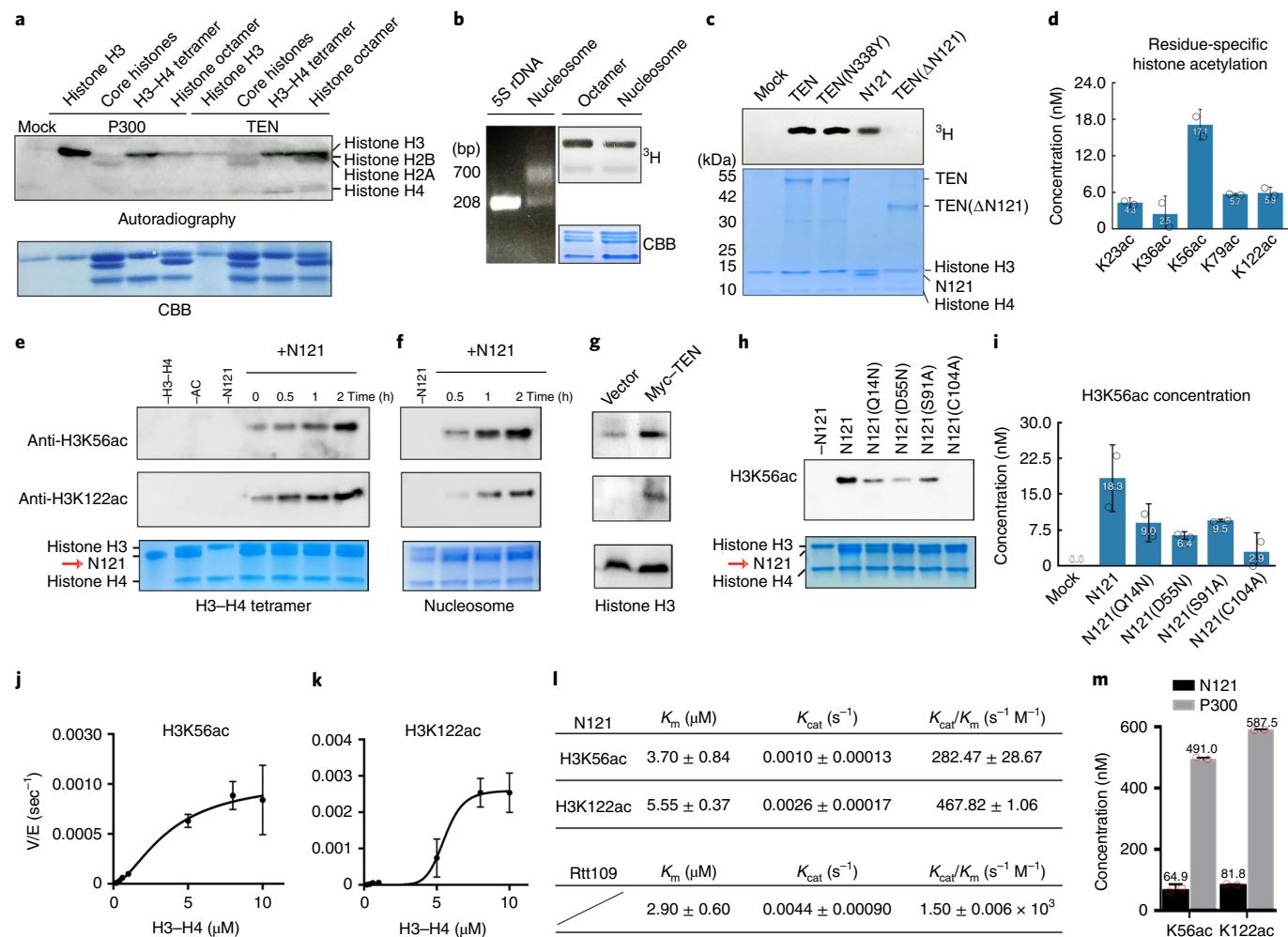
Together, these findings demonstrate that the N121 domain of TEN is a histone acetyltransferase which functions preferentially at the globular domain of histone H3 within the nucleosome.

**In vivo evidence that TEN facilitates chromatin accessibility.** To further probe the relationship between TEN binding and acetylation of H3K56 and H3K122 in vivo, we mapped genome-wide H3K56ac and H3K122ac levels by ChIP-seq and compared the results to the distribution of TEN-binding sites of tendrils at the coiling stage. The vast majority (more than 99%) of H3K56ac and H3K122ac peaks overlapped (Fig. 5*a*). The intragenic binding sites of TEN were colocalized predominantly with the H3K56ac (1,073 out of 1,257 peaks) and H3K122ac peaks (1,054 out of 1,257 peaks; Fig. 5*a*). In addition, we observed that there were 1,043 TEN-binding peaks that colocalized with both H3K56ac and H3K122ac peaks (Extended Data Fig. 6*b* and Supplementary

Table 5). H3K56ac and H3K122ac had a higher probability of being on exons and introns than on promoters and intergenic regions bound by TEN (Fig. 5*b*). These results indicate that the binding of TEN to gene bodies is related to acetylation of the histone H3 globular domain.

To assess the in vivo effect of the TEN C terminus on TEN binding and HAT activities, we next examined the 1,043 intragenic TEN-binding sites that were simultaneously enriched for H3K56ac and H3K122ac in WT versus *ten-1* plants (Fig. 5*c–f*), as well as 1,043 randomly selected H3K56ac and H3K122ac peaks lacking TEN binding, as an internal control (Fig. 5*g–j*). Most of the binding peaks were absent in the *ten-1* mutant (Fig. 5*c*), consistent with our in vitro binding assays (Fig. 2*a*). This reduction of TEN binding in the *ten-1* mutant was associated with a decrease in H3K56ac levels (Fig. 5*d*). As H3K56ac and H3K122ac are implicated in the loosening and eviction of nucleosomes<sup>20,26–28</sup>, we reasoned that TEN may promote chromatin accessibility and, thereby, stimulate transcription of its target genes.

To test this notion, chromatin accessibility was tracked in the tendrils genome using formaldehyde-assisted isolation of regulatory elements with sequencing (FAIRE-seq). We observed a strong correlation between TEN binding, histone acetylation and chromatin accessibility (Fig. 5*c–e*); these functions were markedly impeded in *ten-1* compared to WT plants (Fig. 5*e*). Importantly, RNA-seq analysis further showed that most genes that corresponded to the TEN intragenic peaks were downregulated in *ten-1* (Fig. 5*f* and



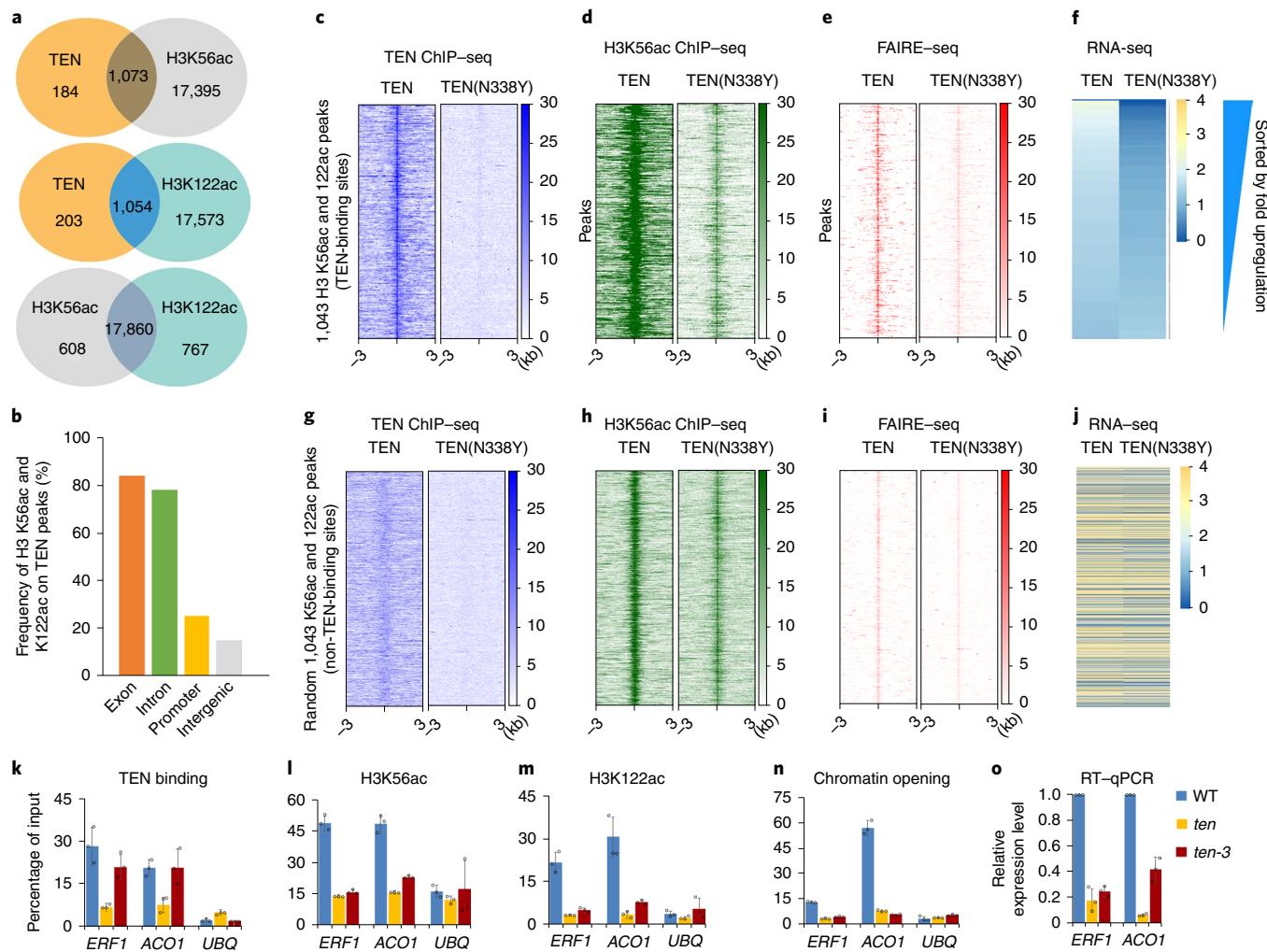
**Fig. 4 |** TEN is a HAT. **a**, HAT assays with radiolabelled acetyl CoA and recombinant P300, Flag-TEN or mock (control). CBB, Coomassie brilliant blue. **b**, Mononucleosomes (~700 bp) assembled on 5S ribosomal DNA (rDNA) (208 bp), visualized in ethidium bromide-stained gels (left), subjected to HAT assays with recombinant Flag-TEN with  $^3\text{H}$ -labelled acetyl CoA and visualized by autoradiography or CBB staining (right). **c**, HAT assays with radiolabelled acetyl CoA and Flag-TEN(N338Y), Flag N121 and Flag( $\Delta$ N121) or mock (control). **d**, Quantification by mass spectrometry of acetylation levels of individual lysines on histone H3 after the HAT assay using the recombinant N121 protein expressed in and purified from *E. coli* (mean  $\pm$  s.d.,  $n=2$ ). **e,f**, Acetylation of H3K56 and K122 within the H3-H4 tetramer (**e**) and nucleosome (**f**) by N121 purified from *E. coli*, determined by immunoblotting analysis. AC, acetyl CoA. The red arrow highlights the bands corresponding to the recombinant N121 protein. **g**, Overexpression of Myc-TEN in tobacco leaves increased the acetylation levels of H3K56 and H3K122. **h,i**, Mutant N121 proteins purified from *E. coli* reduced acetylation on H3K56 (mean  $\pm$  s.d.,  $n=2$ ). The red arrow highlights the bands corresponding to the recombinant N121 protein. **j-l**, Steady-state kinetics of HAT activity of N121 purified from *E. coli* on H3K56 (**j**) and H3K122 (**k**) (mean  $\pm$  s.d.,  $n=3$ ). Derived kinetic parameters for  $K_m$  and  $K_{cat}$  are shown and compared with those for Rtt109 (ref. 42) (**l**) (mean  $\pm$  s.d.,  $n=3$ ). V, reaction rate at a substrate concentration;  $K_{cat}$ , turnover number expressed as number of substrate molecules turned into product per enzyme site per minute; E, enzyme concentration. **m**, A direct comparison between N121 HAT activity and that of a canonical HAT, P300, by quantitative mass spectrometry. HAT assays with 600  $\mu\text{M}$  acetyl CoA, 0.2  $\mu\text{M}$  N121 or P300, and 2.0  $\mu\text{M}$  H3-H4 tetramer at 30 °C for 3 h. The y-axis indicates the calculated concentration of acetylation at specific lysine residues (mean  $\pm$  s.d.,  $n=2$ ). Images in **a-c** and **e-h** are representative of experiments that were repeated at least twice with similar results. Data were compared using a one-sided Student's t-test.

Supplementary Table 6). By contrast, the internal controls showed no significant difference between WT and *ten-1* (Fig. 5g–j). The genomic view of two exemplary targets, *ACO1* and *ERF1*, also supported these findings (Extended Data Fig. 6c, d).

To survey the *in vivo* effect of the N121 domain on TEN binding, H3K56ac and H3K122ac, chromatin accessibility and target-gene expression, we performed ChIP-qPCR, FAIRE-qPCR and RT-qPCR at the TEN-target genes *ACO1* and *ERF1* in WT and *ten-3* (which has an in-frame deletion of Asp55 in the N121 domain), using *ten-1* as a negative control (Fig. 5k–o). As expected, both genes showed substantial decreases in TEN binding in their TEN-binding regions (Fig. 5k), H3K56ac and H3K122ac levels

(Fig. 5l,m) and FAIRE signals (Fig. 5n), as well as substantially reduced expression—about 7-fold for *ERF1* and 17-fold for *ACO1*—in the *ten-1* mutant (Fig. 5o).

Furthermore, we established that deletion of Asp55 in *ten-3* did not significantly affect the TEN binding levels (Fig. 5k). However, H3K56ac and H3K122ac levels, and chromatin accessibility were markedly reduced (Fig. 5l–n). As predicted, these two genes were also downregulated by about fivefold for *ERF1* and threefold for *ACO1* compared to the WT (Fig. 5o), consistent with our biochemical results showing that Asp55 was a major contributor to the HAT activity of TEN N121 (Fig. 4h,j). These findings provide strong support for the hypothesis that, in tendrils, TEN binds specifically to its



**Fig. 5 |** TEN promotes chromatin accessibility. **a**, Metagene analysis showing genome-wide colocalization of TEN intragenic peaks with H3K56ac and H3K122ac in tendrils. **b**, H3K56ac and H3K122ac are enriched in intragenic regions of TEN peaks, but not in promoter or intergenic regions. **c–f**, Read density heat maps showing the intensity of TEN peaks (**c**), H3K56ac signals (**d**), chromatin-opening signals (**e**) and RNA-seq signals (**f**) in WT and *ten-1* at 1,043 overlapped peaks spanning  $\pm 3$  kb from the centre of the TEN peaks. Analysed peaks were organized from top to bottom on the basis of the fold downregulation in *ten-1*. **g–i**, Read density heat maps showing the intensity of TEN peaks (**g**), H3K56ac signals (**h**), FAIRE-seq signals (**i**) and RNA-seq signals (**j**) in WT and *ten-1* at 1,043 random non-TEN-binding peaks spanning  $\pm 3$  kb from the centre of the H3K56ac peaks. **k–o**, Validation of TEN binding, histone acetylation, chromatin opening and gene expression. Association of TEN protein with *ACO1* and *ERF1* loci and control regions (UBQ, *Csa3G778350*). ChIP was performed on WT, *ten-1* and *ten-3* plants with a TEN polyclonal antibody (**k**). H3K56ac (**l**) and H3K122ac (**m**) levels at the *ACO1*, *ERF1* and *UBQ* loci in WT, *ten-1* and *ten-3*. Chromatin opening detected at *ACO1*, *ERF1* and *UBQ* loci in WT, *ten-1* and *ten-3* plants by FAIRE-qPCR (**n**). Relative mRNA expression levels of *ACO1* and *ERF1* genes among WT, *ten-1* and *ten-3* plants detected by RT-qPCR. UBQ was used as internal control (mean  $\pm$  s.d.,  $n = 3$ ) (**o**).

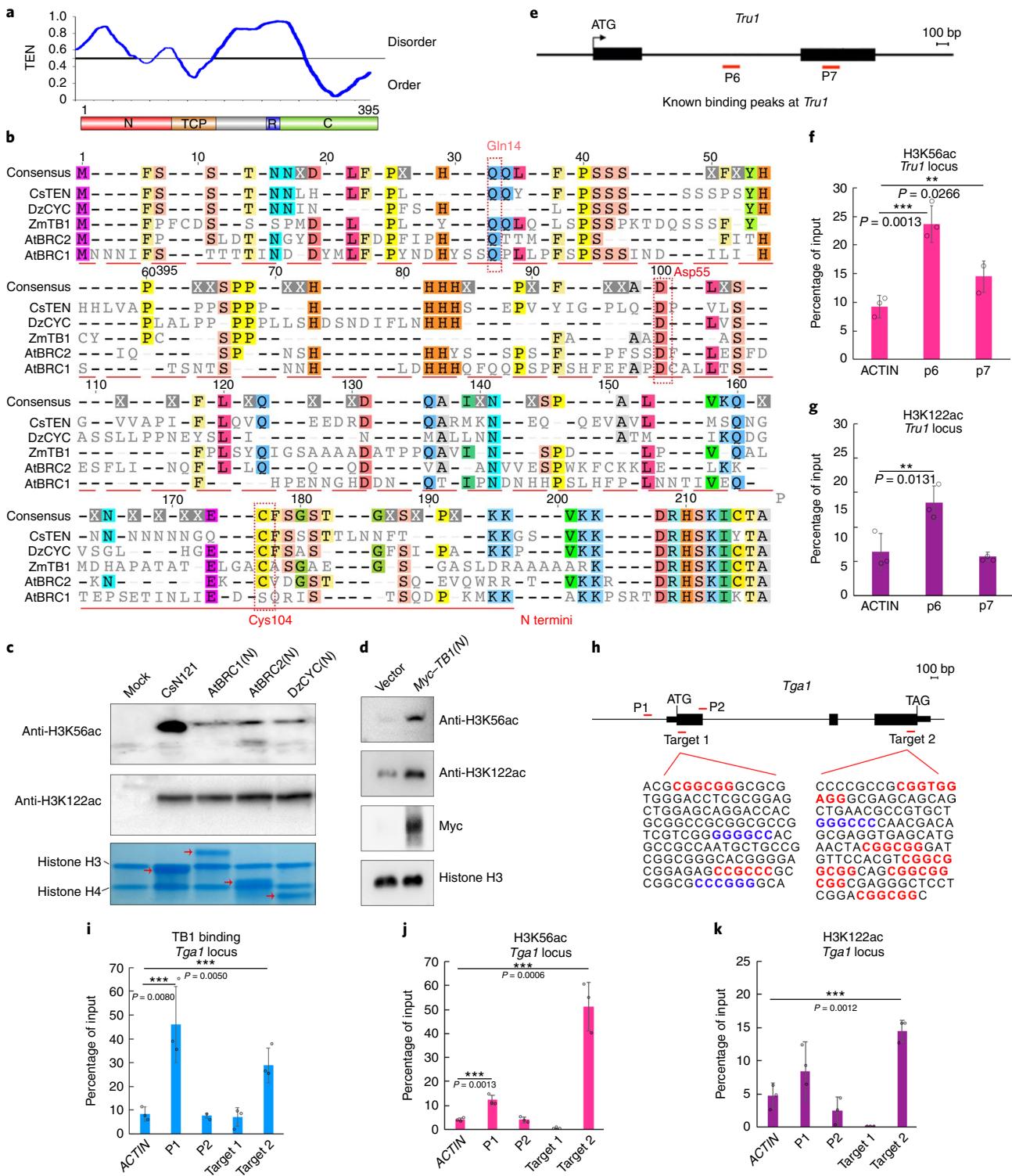
target intragenic regions, where it acetylates the globular domain of histone H3 to facilitate chromatin accessibility, thereby activating target-gene expression.

Equally importantly, combined with the phenotypic changes in tendrils induced by the mutations in the C and N termini, our findings provide in planta evidence that TEN binding, and the associated acetylation of the globular domain of H3, are critical for normal tendril architecture and behaviour.

**Conserved HAT function of CYC/TB1 transcription factors conferred by IDRs.** To explore whether the N-terminal regions of CYC/TB1-like proteins have conserved HAT activity, we performed N-terminal alignments of several CYC/TB1-like transcription factors from angiosperm species (Extended Data Fig. 7a). Cucumber TEN, *Durio* CYC (DzCYC), *Arabidopsis* BRC1 (AtBRC1) and BRC2 (AtBRC2), and maize TB1 (ZmTB1)

were selected for further analysis. All these N-terminal regions contain a large proportion of IDRs (Fig. 6a and Extended Data Fig. 7b) and share little sequence homology (Fig. 6b). However, these regions do contain some amino acids, including Gln14, Asp55 and Cys104 that are responsible for the catalytic activity of TEN N121 (Figs. 4h and 6b). As IDRs generally evolve rapidly at the primary sequence level, it is currently difficult to predict whether their functions have been preserved during the evolution of CYC/TB1-like proteins<sup>45</sup>.

N-terminal domains of DzCYC, AtBRC1 and AtBRC2 were expressed and purified to test their effects on Histone H3 K56 and K122 acetylation (Supplementary Fig. 4). Our results showed that each protein acetylated both K56 and K122 in the histone H3 globular domain (Fig. 6c and Extended Data Fig. 8). Transient expression of ZmTB1 in *N. tabacum* leaves similarly resulted in increased H3K56 and H3K122 acetylation (Fig. 6d). These findings support



**Fig. 6 |** The N termini of CYC/TB1-like proteins have conserved HAT activities. **a**, IDR of TEN. Plot of intrinsic disorder (PONDR VL3-BA) for TEN protein. PONDR VL3-BA score (y-axis) and amino acid position (x-axis) are shown, indicating the intrinsically disordered and ordered regions. **b**, Alignment of the N terminus of CYC/TB1-like proteins among four angiosperm species; cucumber TEN (CsTEN), durian CYC DzCYC, maize TB1 (ZmTB1), *Arabidopsis* BRC1 (AtBRC1) and *Arabidopsis* BRC2 (AtBRC2). Red dashed boxes indicate the Gln14, Asp55 and Cys104 amino acid residues. Coloured boxes beneath letters highlight relatively conserved amino acid residues. **c**, Acetylation of H3K56 and K122 within the H3-H4 tetramer by N-terminal proteins of CsTEN, AtBRC1, AtBRC2 and DzCYC, determined by immunoblotting analysis. Red arrows indicate the bands corresponding to the recombinant N-terminal proteins. **d**, Overexpression of Myc fusion with N terminus of TB1 (Myc-TB1(N)) in tobacco leaves increased acetylation levels of H3K56 and H3K122. **e**, The *Tru1* genic region. Black boxes indicate exons, and lines between boxes represent introns. **f,g**, H3K56ac (**f**) and H3K122ac (**g**) levels at P6 and P7 sites of *Tru1* locus bound by TB1 (mean  $\pm$  s.d.,  $n=3$ ). **h**, The *Tga1* genic region. Locations of the amplicons (targets 1 and 2) used for ChIP-qPCR are marked below. Potential TB1-binding motifs are highlighted in blue (GGGCC) and red (CCNNNN). **i-k**, TB1 binding (**i**) and H3K56ac (**j**) and H3K122ac (**k**) levels at target 1 and 2 sites of *Tga1*. ChIP was performed with antibodies to TB1, H3K56ac and H3K122ac (mean  $\pm$  s.d.,  $n=3$ ). Experiments in **c,d** were repeated at least twice with similar results. Data in **f,g,i-k** were compared using a one-sided Student's t-test. \*\*\* $P < 0.01$ , \*\* $P < 0.05$ .

the hypothesis that such homologous IDRs retain similar functions, despite extensive sequence divergence<sup>45</sup>.

Recent studies have shown that ZmTB1 binds to the gene bodies of the target gene *Tru1* (ref. <sup>7</sup>). The acetylation levels of TB1-binding sites on *Tru1* in maize tiller buds were assessed by ChIP-qPCR, using the previously reported primers<sup>7</sup> (Fig. 6e). Levels of both H3K56ac and H3K122ac were markedly enriched at the intronic P6 site (Fig. 6f,g), consistent with previous results showing that TB1 binds to the P6 site with high affinity<sup>7</sup>. *Teosinte glume architecture1* (*Tga1*) is another maize domestication gene that is downstream of TB1 and upregulated by TB1 (ref. <sup>46</sup>). By using a TB1 antibody, we established that in addition to binding to the known p1 site in the promoter, TB1 could bind to a DNA element located in the third exon of the *Tga1* locus (Fig. 6h,i). Furthermore, we showed that both H3K56ac and H3K122ac modifications were significantly enriched at the TB1 intragenic binding site (Fig. 6j,k). H3K56ac was also identified to be significantly enriched at the p1 promoter region bound by TB1. However, although TB1 showed stronger binding at the promoter, the level of H3K56ac was lower than that in the intragenic binding site (Fig. 6i,j). We inferred that this might be attributable to acetylation at the histone H3 globular domain, such as at H3K56ac or H3K122ac in coding regions, thereby evading the surveillance of the classic histone deacetylation pathway<sup>14</sup>. Recent studies have indicated that AtBRC1 also binds to the introns of targets *HB53* and *HB21* (ref. <sup>8</sup>), and activates their expression. Our study implicates a conserved mechanism for gene activation by CYC/TB1-like transcription factors, which function both as transcription factors binding at intragenic enhancer sites and as HATs acting on the histone globular domain (Supplementary Fig. 5). This discovery provides important insights into the regulatory mechanism involved in axillary bud development and thus in determining plant architecture.

## Discussion

Every genome results from a distinct evolutionary trajectory, and offers a unique window through which to investigate a range of biological processes. In this study, the cucumber genome provides a lens through which to observe the molecular functions of the CYC/TB1-like protein family of transcription factors to gain molecular insight into how intragenic binding by a transcription factor can regulate gene expression and to explore the function of IDRs within transcription factors.

A recent genome-wide ChIP-seq assay reported that TB1 mainly binds to promoters, and only a few peaks were located within gene-body regions<sup>6</sup>. In the current study, TEN predominantly bound to the gene body. We have previously shown that the important N338 residue in the C terminus of TEN, which is essential for its capacity to bind to the identified CDCCRCC motif (Fig. 2o), is located within a highly conserved seven-amino-acid motif (CNNFYFP) in the Cucurbitaceae, but is absent from all other known plant TCP proteins<sup>36</sup>. Given that TB1 and TEN have diverged over a substantial period of evolution, the difference in their binding specificities was not unexpected. In addition, approximately 1-mm-long arrested buds were used for TB1 ChIP-seq, whereas tendrils at the coiling stage were used for TEN ChIP-seq assays. We also observed that the anti-TEN antibody was raised against a peptide partially overlapping the TCP domain. Although the antibody did not recognize the TCP domain protein in vitro (Extended Data Fig. 1d), when TEN protein is bound to the TCP consensus binding sequence through its TCP domain in vivo, the steric hindrance or associated changes in TCP domain folding may subsequently make TEN protein unrecognizable to the antibody. This is a possibility that may explain why the TCP consensus binding sequence, GGNCCC, was not identified from the ChIP-seq peaks.

We further compared the regulatory datasets of TB1 and TEN by analysing the functional categories of the TEN targets with identified TB1 regulatory datasets<sup>6</sup>. Gene ontology analyses indicated that

the targets of TB1 and TEN could be classified mainly in categories including transcription factors, sugar metabolism, signal to stress, biosynthetic metabolism, phytohormone pathway or development. TB1 has been shown to directly modulate sugar-signalling genes in the bud, including *TREHALOSE-6-PHOSPHATE SYNTHETASE* (*trps*), which regulates the metabolism of trehalose 6-phosphate, an important signalling molecule that is involved in allowing bud growth. Interestingly, in this regard, we found the *trps* genes in the TEN-target gene set (Supplementary Table 7). In addition, our gene ontology analyses also indicated that both TEN and TB1 appear to directly regulate phytohormones, although they target mainly different genes, and that these target genes were regulated by TB1 through binding of their promoter in maize, whereas in cucumber, their regulation by TEN occurs through binding of their intragenic enhancer. TB1 was reported to directly modulate the biosynthesis and signalling of abscisic acid (ABA), jasmonic acid (JA) and gibberellin (GA)<sup>6</sup>. TEN was also able to regulate ABA biosynthesis, JA signalling and GA signalling; however, TEN specifically regulates ethylene biosynthesis and signalling as well as auxin transport and signalling, indicating the divergence in the target networks between maize TB1 and cucumber TEN (Extended Data Fig. 9 and Supplementary Table 7).

In this study, we reveal three features of TEN: (1) its primary function may be as a transcriptional activator, (2) it binds intragenic enhancers of target genes, and (3) it is a non-canonical HAT that acts on the histone globular domain. TEN thus links these three activities, and therefore, provides a mechanism by which the modification at the H3 globular domain is integrated with the transcriptional process<sup>32</sup>. Tail-based histone acetylation sites function as sites for the recruitment of transcriptional regulators, which usually occurs around transcription start sites. Our study supports the notion that acetylation of the histone globular domain, in concert with the maintenance of accessible chromatin, are important determinants of transcriptional regulation via intragenic enhancers.

In view of these findings, we investigated the genomic distribution of two human enhancer-binding transcription factors, heat shock factor (HSF-1)<sup>47</sup> and oestrogen receptor alpha (ER $\alpha$ )<sup>20,48</sup>, both of which can interact and recruit CBP/p300. We showed that H3K56ac or H3K122ac modification was more likely to occur on exons than in the other regions bound by the HSF-1 or ER $\alpha$  (Extended Data Fig. 10a,b). Using K56ac or K122ac could also be advantageous, these sites can evade the surveillance of the classic histone deacetylation pathway in coding regions<sup>14</sup>. Thus, our findings add critical information that advances knowledge of how the binding of a transcription factor within the intragenic region of a gene can influence expression. The acetylation and loosening of chromatin at specific intragenic targets by TEN could act to facilitate productive RNA elongation by RNA polymerase (Supplementary Fig. 5). TEN-dependent epigenetic regulation on these intragenic targets could also potentially have directed codon choice and affected protein evolution<sup>49</sup> during the evolutionary process of tendril formation.

HATs constitute a diverse collection of enzymes characterized by their sequence homology and structural features, including the GNAT, MYST and p300/CBP families<sup>50</sup>. These classical HATs are recruited mainly to target promoters through physical interactions with sequence-specific transcription factors<sup>51</sup>, whereas in mammalian systems, some HATs also possess DNA-binding activity<sup>52,53</sup>. In our study, we identified TEN, a non-canonical plant DNA-binding HAT that acetylates the histone H3 globular domain at intragenic enhancers through its N terminus, which harbours a significant portion of IDRs (Fig. 6a and Extended Data Fig. 7b). We also showed that the kinetics of the acetylation activity of the N121 domain on H3K56 and H3K122 are similar to that of Rtt109, approximately one-eighth that of P300. Vps75 greatly enhances the HAT activity of Rtt109 (ref. <sup>44</sup>), and therefore, it is possible that HAT activity of

TEN in vivo may also increase through its interactions with other proteins or via post-translational modifications.

Owing to their highly divergent primary sequences, the function and evolution of these IDRs has remained largely unexplored. Our findings that the intrinsically disordered N termini of all tested CYC/TB1-like proteins have conserved HAT activity provide a mechanism of action of transcription factors with IDRs<sup>35</sup>. In addition, IDRs have recently been shown to have an important role in the compartmentalization of the transcription apparatus via liquid–liquid phase separation, which is related to the nature of super-enhancers<sup>13</sup>. In this regard, our study also suggests the possibility of testing whether the typical intragenic enhancer functions via a similar mechanism.

## Methods

**Experimental materials.** The cucumber (*Cucumis sativus* L.) inbred line 404 (WT) and its BC<sub>3</sub>S<sub>2</sub> mutant tendril near-isogenic line 404-38 (*ten*) were used for genome-wide ChIP-seq, FAIRE-seq and RNA-seq analyses. After seed germination, 100 plants of each line were grown in a greenhouse in pots containing mixed peat moss and vermiculite (v/v = 1:1) and were transplanted to soil at the three-leaf stage. Pest control was performed according to standard management practices.

The cucumber inbred line CU2 was used in cucumber transformation. Seeds were soaked in distilled water at 50 °C for 30 min. Seed coats were removed, and the naked seed was then surface-sterilized by sequential immersion in 70% ethanol for 15 s and 0.6% sodium hypochlorite solution for 15 min, followed by 8 rinses in sterile distilled water. Sterilized seeds were spread on 1× Murashige and Skoog medium (Phytotech, M519) supplemented with 2 mg l<sup>-1</sup> 6-benzylaminopurine (Sigma, B3408) and 1 mg l<sup>-1</sup> ABA (Phytotech, A102) for 2 d at 28 °C. Cotyledons were excised from germinated seedlings and infected with *Agrobacterium*. Subsequently, after shoot regeneration, elongation and rooting processes, the root plants were transplanted to the greenhouse.

For transient expression analysis, tobacco plants (*Nicotiana benthamiana*) were grown in pots containing mixed peat moss and vermiculite (v/v = 1:1) in a growth chamber with a light regime of 16 h:8 h light:dark at 22 °C.

**Plasmid construction and plant transformation.** To generate CRISPR–Cas9 engineered mutations in the *TEN* gene, a binary CRISPR–Cas9 vector, pBSE402, plus a 35S-GFP expression cassette was modified from pBSE401a<sup>54</sup>. For assembly of *TEN* sgRNA into pBSE402, equal volumes of 100 μM forward and reverse primers were mixed, incubated at 95 °C for 5 min, and slowly cooled to room temperature, resulting in a double stranded DNA fragment with sticky BsaI ends. This short DNA fragment was then assembled into pBSE402 by restriction fragment ligation using BsaI and T4 Ligase (New England Biolabs). Primers are shown in Supplementary Table 8. *Agrobacterium tumefaciens* strain EH105 carrying a pBSE402-TEN construct was used to transform the cucumber inbred line CU2 using cotyledonary nodes as explants, as previously described<sup>15</sup>. Shoot regeneration, elongation and rooting processes strictly followed normative procedures.

Genomic DNA was extracted from the positive transgenic plants using the DNeasy Plant Mini Kit (Qiagen, 69104). PCR was performed using gene-specific primers (Supplementary Table 8). PCR products were cloned into pEASY-Blunt Zero (Transgen Biotech, CB501) and the various alleles of the *TEN* gene were identified by sequencing.

**ChIP-seq and ChIP-qPCR.** ChIP assays were performed as previously described<sup>56</sup>, with some modifications. In brief, normal tendrils (WT) at coiling stage (the stage at which the tendril attaches to a support, but before free tendril coiling occurs), mutant tendrils of *ten*-3 (at the corresponding growth stage where the tendril would normally be attaching to a support) and *ten*-1 mutants (at the corresponding growth stage where the modified tendrils show slight curling of petioles) (Fig. 1g) were used in ChIP assays for two biological replicates. Collected tendrils (30 g of each material divided into 10 equal samples) were fixed in cross-linking buffer (10 mM sodium phosphate, pH 7.0, 50 mM NaCl, 0.1 M sucrose and 1% formaldehyde) under vacuum for 10 min. Fixation was stopped by incubation in 0.25 M glycine for an additional 10 min. Tendril material was then ground in liquid nitrogen, and 3 g aliquots of powdered tissue were resuspended in 30 ml extraction buffer (0.4 M sucrose, 10 mM Tris-HCl pH 8.0, 5 mM β-mercaptoethanol, 1 mM PMSF and protease inhibitor cocktail). Chromatin isolation was performed as previously described<sup>56</sup>, and chromatin was sonicated (Ningbo Scientz Biotechnology, JY96-IIIN) for 12 cycles, each with a 15 s pulse at 70% of maximal power, followed by a 45 s cooling period on ice, to achieve an average DNA size of 200 bp for immunoprecipitation. The following antibodies were used for ChIP assays: TEN antibody, anti-H3K56ac (Active Motif, 39282) and anti-H3K12ac (Abcam, Ab33309) (Supplementary Table 9). ChIP products were combined and eluted into 50 μl Tris-EDTA buffer for ChIP-seq (5 ng DNA) or ChIP-qPCR (1 μl aliquot).

ChIP-qPCR was performed as previously described<sup>56</sup>. Primers are listed in Supplementary Table 8; *UBQ* (*Csa3G778350*) was used as a negative control. The qPCR signals derived from the ChIP samples were normalized to the signals derived from the input DNA control sample. Amount of DNA (percentage of input) was calculated by the 2<sup>-ΔCt</sup> method.

**FAIRE-seq.** FAIRE assays were performed as previously described<sup>57</sup>. Normal tendrils (WT) at the coiling stage and mutant tendrils (*ten*-3 and *ten*-1 mutants) at the corresponding growth stage were used. Two grams of tissue was fixed with formaldehyde and regulatory elements were isolated. FAIRE DNA was dissolved into 30 μl Tris-EDTA buffer for FAIRE-seq (5 ng DNA) or FAIRE-qPCR (1 μl aliquot). Two biological replicates were performed for each FAIRE assay.

qPCR was performed on crosslinked and non-crosslinked (input) FAIRE samples. The ubiquitin gene, *UBQ*, was used as a negative control. Lists of all primers used are given in Supplementary Table 8. The value for DNA accessibility over that of input was obtained by the 2<sup>-ΔCt</sup> method.

**RT-qPCR analyses.** Total RNA was isolated using an RNA extraction Kit (Qiagen, 74903). First-strand cDNA was synthesized from 1 μg total RNA using the M-MLV Reverse Transcriptase (Promega, M1705) (primers are listed in Supplementary Table 8). Primer specificity was checked by sequencing and BLAST analysis. qPCRs were performed on an ABI 7900 using SYBR Premix (Roche, 4913914001), according to the manufacturer's instructions. Three technical replicates and three independent biological experiments were performed in all cases. Relative gene expression was assessed using the comparative 2<sup>-ΔΔCt</sup> method. *UBQ* was used as an internal reference gene.

**High-throughput sequencing.** ChIP-seq or FAIRE-seq libraries were prepared using the Illumina ChIP-seq DNA Sample Prep kit according to the manufacturer's instructions with the following modifications: mRNA adaptor indexes from the TruSeq RNA kit were used, and enrichment PCR was also performed with reagents from the Illumina TruSeq mRNA kit. The enriched libraries were purified with AMPure magnetic beads (Agencourt), the concentrations were checked with Qubit (Invitrogen), and the distribution and size of fragments were confirmed with a Bioanalyzer (Agilent). Four samples were pooled in equimolar quantity and sequenced on a HiSeq 2500 (single read, 50 bp) to yield up to 30 million reads per sample. The obtained reads were demultiplexed with Illumina CASAVA 1.8 software.

RNA-seq libraries were developed from five biological replicates from tendrils (WT) and mutant tendrils (*ten*). The 100-bp paired-end reads (2.4 Gb, 10×) for each sample were generated from the RNA-seq libraries with an Illumina HiSeq 2500 sequencer.

**Mapping of sequencing reads and data analysis.** All sequencing reads were mapped to the cucumber genome using bowtie software (<http://bowtie-bio.sourceforge.net>)<sup>58</sup> with default parameters, except that multiple loci-matching reads that might introduce error signals by repeat counting were discarded.

For computational processing of ChIP-seq: peaks for TEN and TEN(N338Y) were identified by the model-based analysis software MACS (<https://github.com/macs3-project/MACS/>)<sup>59</sup> using input DNA as a control. MACS default parameters were used, except for detecting more reliable TEN association signals with *-mfold*=0, 30 and fold enrichment ≥ 2. Peaks for H3K56ac and H3K12ac were identified with RSEG software (<https://github.com/smithlabcode/rseg>)<sup>60</sup> with default parameters using input DNA as a control. For computational processing of FAIRE-seq data: data peaks were identified by F-Seq software (<http://fureyelab.unc.edu/software/fseq/>)<sup>61</sup> using input DNA as a control. Heat map graphs of peaks were plotted using deeptools software (<https://github.com/deeptools/deepTools>) with normalization set to 1x.

For computational processing, RNA-seq data were mapped to the cucumber genome using tophat2 software (<http://ccb.jhu.edu/software/tophat/index.shtml>)<sup>62</sup>, with default parameters. According to the cucumber genome annotation, all mapped reads were then assembled into known transcripts by Cufflink software. Next, the expression of transcripts was calculated in fragments per kilobase of exon model per million mapped fragments.

The putative DNA-binding motifs in the TEN-binding peaks were searched using MEME-ChIP software<sup>63</sup> (online version 5.0.5) using the default background model of MEME-ChIP. The background model is normalized for biased distribution of letters in the input sequences.

**Purification of TEN protein from Sf9 insect cells and *E. coli*.** The cDNAs encoding TEN, TEN(N338Y), TEN N121 and TEN(ΔN121) were cloned into pFast-FH vector (inserting a Flag tag into pFast-HTB; Life Technologies, 10712-024). Primers for constructing these vectors are shown in Supplementary Table 8. Bacmid preparation and insect cell transfections were conducted using the Bac-to-Bac Baculovirus Expression System, according to the manufacturer's instructions. The isolated P3 recombinant baculoviruses were added to the cultured Sf9 cells at a volume ratio of 1:100. Cells were collected after a further 48–60 h of cultivation at 27 °C and 110 rpm in Nalgene conical flasks.

Insect cells that expressed recombinant proteins were resuspended and sonicated in lysis buffer (50 mM HEPES-KOH, pH 7.5, 500 mM NaCl, 5% glycerol). Cell lysates were centrifuged at 16,000g for 60 min at 4°C. The supernatants from cell lysates were loaded onto a gravity column (Bio-Rad, 732-1010) filled with 2 ml anti-Flag M2 Affinity Gel (Sigma, A2220), and bound proteins were eluted with a Flag elution buffer (50 mM HEPES-KOH, pH 7.5, 500 mM NaCl, 5% glycerol, 0.4 mg ml<sup>-1</sup> 1× Flag peptide). The eluate was concentrated to 2 ml and then loaded onto a Superdex 200 column (GE Healthcare, 10034543) for size-exclusion chromatography (S200 buffer: 50 mM Hepes-KOH, pH 7.5, 500 mM NaCl, 5% glycerol, 1 mM DTT). The fractions containing target protein were collected and constituted the purified protein.

The cDNA encoding a series of N-terminal sequences, including N121, the point mutated N121 and N termini of the TB1 transcription factor family, were cloned into the pET22b vector to create a fusion protein with an N-terminal 6× His-tag. The resultant plasmid was transformed into *E. coli* BL21(DE3), identified by PCR, digested with NdeI and XhoI, and sequenced. Expression of the recombinant N121 protein was induced with 0.2 mM isopropyl thiogalactoside for 5 h at 26°C. The protein was then affinity purified in lysis buffer (25 mM Tris-HCl, pH 7.5, 500 mM NaCl) with an Ni<sup>2+</sup>-chelating Sepharose Fast Flow column (Amersham Biosciences), following the manufacturer's instructions.

**Electrophoretic mobility shift assays.** EMSA was performed using recombinant MBP-TCP, MBP-TCP-R, Flag-TEN(ΔN121), Flag-TEN or Flag-TEN(N338Y) protein purified from insect cells. DNA probes containing the CTCCGCC motif, mutant CTAAGCC motif or a GTGGTCCCCAC motif used for the EMSA, were synthesized and amplified by PCR using the biotin-labelled primers listed in Supplementary Table 8. Binding reactions were performed in 20 μl binding buffer, composed of 10 mM Tris-HCl, pH 7.5, 200 mM NaCl, 10 mM KCl, 1 mM MgCl<sub>2</sub>, 10 μM ZnCl<sub>2</sub>, 0.5 mg ml<sup>-1</sup> BSA, 0.02 mg ml<sup>-1</sup> poly(deoxyinosinic-deoxycytidyl) sodium salt (Thermo Scientific, 20148E), 1 mM DTT and 10% glycerol. Binding reactions were carried out using 0, 100 or 200 ng of recombinant protein and 5 nM of each biotin-labelled probe at 4°C for 1 h. Samples were separated on 6% polyacrylamide gels (19:1 acryl:bisacrylamide) in Tris-borate-EDTA, at 4°C. After the sample transfer, the PVDF membrane was exposed under ultraviolet light to cross-link the samples, and then blocked with blocking reagent (provided in the kit) for 15 min. Biotin signal was visualized using a Chemiluminescent Nucleic Acid Detection Module kit (Thermo Scientific, 89880). Competition experiments were performed using 100- or 1,000-fold levels of unlabelled fragments.

**Expression and purification of recombinant histones.** Recombinant histones were expressed in BL21 (DE3) pLysS. Single colonies were grown in 1 l of lysogeny broth medium at 37°C until they reached an optical density at 600 nm of 0.6 and induced with 0.5 mM IPTG for 2 h. Cells were collected by centrifugation at 5,000g and resuspended in lysis buffer (50 mM Tris, 100 mM NaCl, 1 mM EDTA, 1 mM 2-mercaptoethanol, pH 7.5). Cells were lysed by sonication and then centrifuged at 30,000g.

For histone purification, inclusion bodies from 1 l bacterial culture were washed by resuspension in 100 ml wash buffer (lysis buffer containing 1% Triton X-100) and then centrifuged for 10 min at 4°C at 23,000g. This step was repeated, once, with wash buffer, and twice with lysis buffer. The pellet was solubilized in 30 ml unfolding buffer (7 M guanidium hydrochloride, 20 mM Tris, pH 7.5, 10 mM DTT) for 1 h at room temperature. After centrifugation, the supernatant was analysed by SDS-PAGE.

**Histone tetramer and octamer reconstitution.** For histone H3–H4 tetramer or histone octamer reconstitution, histones H3, H4 or H2A, H2B, H3 and H4 were mixed in equimolar amounts and then dialysed at 4°C against three changes of 2 l of freshly-prepared refolding buffer (2 M NaCl, 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 5 mM β-mercaptoethanol) for 12 h. Next, the mixture was dialysed for another 12 h, in 2 l fresh refolding buffer. After centrifugation, the supernatant was collected and concentrated to a final volume of 500 μl. Following a minimum of three centrifugation steps, the cleared supernatant was loaded onto a Superdex 200 gel filtration column and equilibrated with refolding buffer. The purity and stoichiometry of eluted fractions were checked by SDS-PAGE. Histone H3–H4 tetramer or histone octamer peak fractions were collected and stored at -80°C.

**In vitro nucleosome assembly.** Mononucleosomes were assembled on 208-bp (5S rDNA) DNA fragments. Before adding octamer, the salt concentration of the DNA solution was adjusted to 2 M using 5 M NaCl and Tris-EDTA. DNA and histone octamers were mixed at a 1:1.05 molar ratio in 2 M NaCl buffer. A peristaltic pump was used for continuous dialysis against 450 ml refolding buffer (2 M NaCl, 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, 5 mM β-mercaptoethanol) for 16 h at 4°C under constant stirring with continuous addition of Tris-EDTA buffer into the dialysis buffer to reduce the salt concentration to 0.6 M. Samples were collected after final dialysis in Hepes-EDTA buffer (10 mM Hepes, pH 8.0, 0.1 mM EDTA) for 4 h. The assembled nucleosomes were visualized on 2% agarose gels.

**Histone acetyltransferase assay.** HAT assays were performed in a 30 μl reaction volume using either 0.5 μg of recombinant histone H3 (Millipore, 14–411),

2 μg of chicken core histones (Millipore, 14–411), 1 μg H3–H4 tetramer, 2 μg histone octamer or 2 μg mononucleosome in the presence of 1 μCi of <sup>3</sup>H-acetyl CoA (ARC, 0213A-50 μCi) or 200 μM acetyl CoA. Enzymatic reactions were performed using 100 nM of purified Flag-N121, and the same amount of Flag-TEN, Flag-TEN(N338Y), Flag(ΔN121) or commercial P300 protein. Reactions were incubated at 30°C for 2 h, and then 7.5 μl 5× SDS-PAGE sample buffer was added, followed by boiling for 5 min. After separation on 15% SDS-PAGE, aliquots were analysed by LC-MS/MS, autoradiography or immunoblotting. For autoradiography, proteins were separated and transferred to a PVDF membrane using a semi-dry blotter (TE70, GE Life Sciences). The <sup>3</sup>H signal was detected with a BioMax Transcreen Intensifying Screen LE (Sigma, Z374318) and BIOMAX MS films. For immunoblotting, the antibodies specific for different acetylated lysine residues are listed in Supplementary Table 9.

**LC-MS/MS analysis.** Equal protein amounts were separated by SDS-PAGE. The gel bands of histone H3 protein were excised, reduced with 25 mM DTT and alkylated with 55 mM iodoacetamide, followed by addition of propionic anhydride and in-gel digestion at 37°C overnight with sequencing-grade modified trypsin. Peptides were extracted twice with 0.1% trifluoroacetic acid in 50% acetonitrile aqueous solution for 30 min and then dried in a speedvac. Peptides were redissolved in 25 μl 0.1% trifluoroacetic acid and 6 μl of extracted peptides were analysed with a Q Exactive HF-X mass spectrometer.

In LC-MS/MS analysis, digestion products were separated by a 120 min gradient elution at a flow rate of 0.300 μl min<sup>-1</sup> using a Dionex 3000 nano-HPLC system, which was directly interfaced with a Thermo Q Exactive HF-X mass spectrometer. The analytical column was a fused silica capillary column (75 μm internal diameter, 150 mm length; packed with C-18 resin). Mobile phase A consisted of 0.1% formic acid, and mobile phase B consisted of 80% acetonitrile and 0.08% formic acid. The Q Exactive mass spectrometer was operated in the data-dependent acquisition mode, using Xcalibur4.1 software, and a single full-scan mass spectrum in the Orbitrap (300–1,800 *m/z*, 12,000 resolution) was followed by 40 data-dependent MS/MS scans. The MS/MS spectra from each LC-MS/MS run were searched against the selected database, using the Proteome Discoverer searching algorithm (v.1.4).

The MS/MS spectra from each LC-MS/MS run were searched against the Histone H3.fasta file. The search criteria were as follows: full tryptic specificity was required; two missed cleavages were allowed; carbamidomethylation was set as the fixed modification; the oxidation (M), propionyl (P) and acetyl (K) were set as the variable modification; precursor-ion mass tolerances were set at 20 ppm for all mass spectra acquired in the Orbitrap mass analyser; and the fragment-ion mass tolerance was set at 0.02 Da for all MS2 spectra acquired. The peptide false discovery rate (FDR) was calculated using Percolator provided by Proteome Discoverer. When the *q* value was smaller than 1%, the peptide spectrum match was considered to be correct. FDR was determined on the basis of PSMs when searched against the reverse, decoy database. Peptides assigned only to a given protein group were considered as unique. The FDR was also set to 0.01 for protein identifications. The peak areas of fragment ions were used to calculate the relative intensity of the precursor ion for selected peptides.

**Quantification of residue-specific histone acetylation.** Quantification of residue-specific histone acetylation was performed as previously described<sup>41</sup>. The fraction of specific peptide (*F*<sub>s</sub>) was calculated using equation (1), where *I*<sub>s</sub> is the intensity of an acetylated peptide state, and *I*<sub>p</sub> is the intensity of any state of that peptide:

$$F_s = I_s / \left( \sum I_p \right). \quad (1)$$

The concentration of acetylation at specific lysine residues was quantified and calculated by multiplying the fraction *F*<sub>s</sub> by the initial concentration of histone.

For steady-state kinetic analyses, all models were fitted to the data using Prism (v.7.0). The initial rates (*V*) of acetylation were calculated from the linear stage in acetylation for a 10 min reaction time. To measure steady-state parameters for the H3–H4 tetramer, *K*<sub>cat</sub> and *K*<sub>m</sub> were determined based on the equation:

$$\frac{V}{[E]} = K_{cat} \frac{[S]}{K_m + [S]}, \quad (2)$$

where [S] is the concentration of substrate (H3–H4 tetramer), [E] is the concentration of N121 protein, and *V* is the initial rate of acetylation.

**Protein extraction and detection by immunoblotting.** The 18-residue peptide LNNFTKKGSVKKDRHSKC, spanning the N121 and TCP domains, was selected as antigen for polyclonal antibody production. Two grams of cucumber tendrils or tobacco leaves was collected and ground to fine power in liquid nitrogen. Total protein was extracted using a protein extraction buffer (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 4 M Urea, 10% glycerol, 5 mM DTT, 1 mM PMSF and 1× protease inhibitor cocktail). The samples were centrifuged at 12,000g at 4°C for 30 min. Total protein was quantified using the BCA Protein Assay Kit according to the manufacturer's instructions. Twenty to fifty micrograms of protein sample was resolved on 12.5% SDS-PAGE and then transferred to a PVDF membrane using

a semi-dry blotter. Western blot analyses were performed using the antibodies listed in Supplementary Table 9. The immunoblotting signal was visualized using a SuperSignal West Femto kit (Thermo Scientific).

**Immunolabelling.** Half of each tobacco leaf was infiltrated with *A. tumefaciens* strains GV3101 expressing Myc-TEN, and the other half without infiltration was used as a control. Seventy-two hours after infiltration, the entire tobacco leaf was cut into 0.5 cm × 1 cm pieces and fixed in cold 4% paraformaldehyde in Tris-HCl buffer (10 mM Tris, pH 7.5, 100 mM NaCl, 10 mM EDTA) for 20 min. Leaf pieces were then washed twice using ice-cold Tris-HCl buffer for 10 min each and nuclei were released by finely chopping in LB01 buffer (15 mM Tris-HCl, pH 7.5, 2 mM EDTA, 0.5 mM spermine, 80 mM KCl, 20 mM NaCl, 0.1% Triton X-100), followed by filtration through a cell strainer cup (BD falcon). Nuclei in the flow through were then 1:4 diluted in sorting buffer (100 mM Tris, pH 7.5, 50 mM KCl, 2 mM MgCl<sub>2</sub>, 0.05% Tween 20, 5% sucrose), spotted on microscopy slides and air-dried. After post-fixation with 4% paraformaldehyde in PBS buffer (10 mM sodium phosphate, pH 7.0, 143 mM NaCl), slides were used for immunolabelling. Double labelling was performed using the Myc antibody (1:200), H3K56ac antibody (1:500) and H3K122ac antibody (1:500). Myc-TEN was detected with FITC-conjugated goat anti-rabbit (1:200, ZSGB-BIO) secondary antibodies, and each specific histone acetylation was visualized by TRITC-conjugated goat anti-rabbit (1:200, ZSGB-BIO) secondary antibodies. After staining, slides were mounted in mounting medium with DAPI and then photographed on a Leica TCS SP8 confocal microscope. More than 50 pairs of transfected nuclei versus non-transfected nuclei, in the same field of view, were observed to collect consistent results.

**Immunoprecipitation.** Immunoprecipitation was performed to enrich the TEN protein for detection and LC-MS/MS analysis. Ten milligrams of total tendril protein from the supernatants was incubated with an excess amount of anti-TEN antibody at 4 °C overnight with rotation, followed by addition of 50 µl Protein A Dynabeads (Thermo Scientific, 10002D) for an additional 2 h. Beads were then washed three times with extraction buffer. Half of the immunoprecipitates was analysed by immunoblotting with anti-TEN antibody, and the other half was resolved on 12.5% SDS-PAGE for LC-MS/MS analysis.

**Transient expression in tobacco leaves.** TEN, TEN N121, TEN(ΔN121) and TEN(ΔC130) fused to 5× Myc-tag peptides (EQKLISEEDL), were cloned into a binary vector (pCAMBIA1300) downstream of the 35S promoter using the primers listed in Supplementary Table 8. Constructs were transformed into *A. tumefaciens* strain GV3101. After cultivation overnight, cells were collected by centrifugation and resuspended in 10 mM MES (pH 5.6) buffer containing 10 mM MgCl<sub>2</sub> and 200 µM acetosyringone (Sigma, D134406) at OD<sub>600</sub> of 1.0. After incubation at room temperature for 3 h in the dark, the *Agrobacterium* suspension was infiltrated into leaves of 1-month-old tobacco plants from the adaxial side using a needless syringe. Leaf samples were harvested after 3 d and used for immunoblotting or immunolabelling analyses. These experiments were repeated independently at least twice with similar results.

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

## Data availability

The data supporting the findings in this study are available from the corresponding author upon reasonable request. Source data are provided with this paper.

Received: 8 November 2019; Accepted: 29 May 2020;

Published online: 13 July 2020

## References

- Martín-Trillo, M. & Cubas, P. TCP genes: a family snapshot ten years later. *Trends Plant Sci.* **15**, 31–39 (2010).
- Howarth, D. G. & Donoghue, M. J. Phylogenetic analysis of the “ECE”(CYC/TB1) clade reveals duplications predating the core eudicots. *Proc. Natl Acad. Sci. USA* **103**, 9101–9106 (2006).
- Doebley, J., Stec, A. & Hubbard, L. The evolution of apical dominance in maize. *Nature* **386**, 485–488 (1997).
- Takeda, T. et al. The *OstTB1* gene negatively regulates lateral branching in rice. *Plant J.* **33**, 513–520 (2003).
- Aguilar-Martínez, J. A., Poza-Carrión, C. & Cubas, P. *Arabidopsis BRANCHED1* acts as an integrator of branching signals within axillary buds. *Plant Cell* **19**, 458–472 (2007).
- Dong, Z. et al. The regulatory landscape of a core maize domestication module controlling bud dormancy and growth repression. *Nat. Commun.* **10**, 1–15 (2019).
- Dong, Z. et al. Ideal crop plant architecture is mediated by *tassels Replace Upper ears1*, a BTB/POZ ankyrin repeat gene directly targeted by TEOSINTE BRANCHED1. *Proc. Natl Acad. Sci. USA* **114**, 8656–8664 (2017).
- González-Grandío, E. et al. Abscisic acid signaling is controlled by a BRANCHED1/HD-ZIP I cascade in *Arabidopsis* axillary buds. *Proc. Natl Acad. Sci. USA* **114**, 245–254 (2017).
- Blackwood, E. M. & Kadonaga, J. T. Going the distance: a current view of enhancer action. *Science* **281**, 60–63 (1998).
- Kvon, E. Z. et al. Genome-scale functional characterization of *Drosophila* developmental enhancers in vivo. *Nature* **512**, 91–95 (2014).
- Gillies, S. D., Morrison, S. L., Oi, V. T. & Tonegawa, S. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell* **33**, 717–728 (1983).
- Hebbes, T., Clayton, A., Thorne, A. & Crane-Robinson, C. Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. *EMBO J.* **13**, 1823–1830 (1994).
- Hahn, S. Phase separation, protein disorder, and enhancer function. *Cell* **175**, 1723–1725 (2018).
- Li, B., Carey, M. & Workman, J. L. The role of chromatin during transcription. *Cell* **128**, 707–719 (2007).
- Liu, C., Lu, F., Cui, X. & Cao, X. Histone methylation in higher plants. *Annu. Rev. Plant Biol.* **61**, 395–420 (2010).
- Rajagopal, M., Balasubramanian, S., Ioshikhes, I. & Ramaswamy, A. Structural dynamics of nucleosome mediated by acetylations at H3K56 and H3K115,122. *Eur. Biophys. J.* **46**, 471–484 (2017).
- Suzuki, Y., Horikoshi, N., Kato, D. & Kurumizaka, H. Crystal structure of the nucleosome containing histone H3 with crotonylated lysine 122. *Biochem. Biophys. Res. Comm.* **469**, 483–489 (2016).
- Tessarz, P. & Kouzarides, T. Histone core modifications regulating nucleosome structure and dynamics. *Nat. Rev. Mol. Cell Biol.* **15**, 703 (2014).
- Rufiange, A., Jacques, P.-E., Bhat, W., Robert, F. & Nourani, A. Genome-wide replication-independent histone H3 exchange occurs predominantly at promoters and implicates H3 K56 acetylation and Asf1. *Mol. Cell* **27**, 393–405 (2007).
- Tropberger, P. et al. Regulation of transcription through acetylation of H3K122 on the lateral surface of the histone octamer. *Cell* **152**, 859–872 (2013).
- Williams, S. K., Truong, D. & Tyler, J. K. Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation. *Proc. Natl Acad. Sci. USA* **105**, 9000–9005 (2008).
- Chen, C.-C. et al. Acetylated lysine 56 on histone H3 drives chromatin assembly after repair and signals for the completion of repair. *Cell* **134**, 231–243 (2008).
- Driscoll, R., Hudson, A. & Jackson, S. P. Yeast Rtt109 promotes genome stability by acetylating histone H3 on lysine 56. *Science* **315**, 649–652 (2007).
- Han, J. et al. Rtt109 acetylates histone H3 lysine 56 and functions in DNA replication. *Science* **315**, 653–655 (2007).
- Zhang, L. et al. Multilist substrate recognition in Asf1-dependent acetylation of histone H3 K56 by Rtt109. *Cell* **174**, 818–830 (2018).
- Neumann, H. et al. A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. *Mol. Cell* **36**, 153–163 (2009).
- Watanabe, S. et al. Structural characterization of H3K56Q nucleosomes and nucleosomal arrays. *Biochim. Biophys. Acta* **1799**, 480–486 (2010).
- Schwabish, M. A. & Struhl, K. Asf1 mediates histone eviction and deposition during elongation by RNA polymerase II. *Mol. Cell* **22**, 415–422 (2006).
- Värv, S. et al. Acetylation of H3 K56 is required for RNA polymerase II transcript elongation through heterochromatin in yeast. *Mol. Cell Biol.* **30**, 1467–1477 (2010).
- Das, C., Lucia, M. S., Hansen, K. C. & Tyler, J. K. CBP/p300-mediated acetylation of histone H3 on lysine 56. *Nature* **459**, 113–117 (2009).
- Pradeepa, M. M. et al. Histone H3 globular domain acetylation identifies a new class of enhancers. *Nat. Genet.* **48**, 681–686 (2016).
- Venkatesh, S. & Workman, J. L. Histone exchange, chromatin structure and the regulation of transcription. *Nat. Rev. Mol. Cell Biol.* **16**, 178–189 (2015).
- Tompa, P. Intrinsically disordered proteins: a 10-year recap. *Trends Biochem. Sci.* **37**, 509–516 (2012).
- Wright, P. E. & Dyson, H. J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **16**, 18–29 (2015).
- Roeder, R. G. 50+ years of eukaryotic transcription: an expanding universe of factors and mechanisms. *Nat. Struct. Mol. Biol.* **26**, 1–9 (2019).
- Wang, S. et al. A rare SNP identified a TCP transcription factor essential for tendril development in cucumber. *Mol. Plant* **8**, 1795–1808 (2015).
- Gerbode, S. J., Puzey, J. R., McCormick, A. G. & Mahadevan, L. How the cucumber tendril coils and overwinds. *Science* **337**, 1087–1091 (2012).
- Wang, Y. & Li, J. Branching in rice. *Curr. Opin. Plant Biol.* **14**, 94–99 (2011).
- Vroemen, C. W., Mordhorst, A. P., Albrecht, C., Kwaaitaal, M. A. & de Vries, S. C. The CUP-SHAPED COTYLEDON3 gene is required for boundary and shoot meristem formation in *Arabidopsis*. *Plant Cell* **15**, 1563–1577 (2003).
- Jaffe, M. J. in *Hormonal Regulation of Plant Growth and Development* (ed. Purohit, S. S.) 353–367 (Springer, 1985).

41. Inoue, F. et al. A systematic comparison reveals substantial differences in chromosomal versus episomal encoding of enhancer activity. *Genome Res.* **27**, 38–52 (2017).
42. Weatheritt, R. J. & Babu, M. M. The hidden codes that shape protein evolution. *Science* **342**, 1325–1326 (2013).
43. Kuo, Y.-M., Henry, R. A. & Andrews, A. J. A quantitative multiplexed mass spectrometry assay for studying the kinetic of residue-specific histone acetylation. *Methods* **70**, 127–133 (2014).
44. Berndsen, C. E. et al. Molecular functions of the histone acetyltransferase chaperone complex Rtt109–Vps75. *Nat. Struct. Mol. Biol.* **15**, 948–956 (2008).
45. Zarin, T., Tsai, C. N., Ba, A. N. N. & Moses, A. M. Selection maintains signaling function of a highly diverged intrinsically disordered region. *Proc. Natl Acad. Sci. USA* **114**, 1450–1459 (2017).
46. Studer, A. J., Wang, H. & Doebley, J. F. Selection during maize domestication targeted a gene network controlling plant and inflorescence architecture. *Genetics* **207**, 755–765 (2017).
47. Lo, K. A. et al. Genome-wide profiling of H3K56 acetylation and transcription factor binding sites in human adipocytes. *PLoS ONE* **6**, e19778 (2011).
48. Carroll, J. S. et al. Genome-wide analysis of estrogen receptor binding sites. *Nat. Genet.* **38**, 1289–1297 (2006).
49. Stergachis, A. B. et al. Exonic transcription factor binding directs codon choice and affects protein evolution. *Science* **342**, 1367–1372 (2013).
50. Sabari, B. R., Zhang, D., Allis, C. D. & Zhao, Y. Metabolic regulation of gene expression through histone acylations. *Nat. Rev. Mol. Cell Biol.* **18**, 90–101 (2017).
51. Lee, K. K. & Workman, J. L. Histone acetyltransferase complexes: one size doesn't fit all. *Nat. Rev. Mol. Cell Biol.* **8**, 284–295 (2007).
52. Doi, M., Hirayama, J. & Sassone-Corsi, P. Circadian regulator CLOCK is a histone acetyltransferase. *Cell* **125**, 497–508 (2006).
53. Kawasaki, H. et al. ATP-2 has intrinsic histone acetyltransferase activity which is modulated by phosphorylation. *Nature* **405**, 195–200 (2000).
54. Chen, Y. et al. CRISPR/Cas9-mediated base-editing system efficiently generates gain-of-function mutations in *Arabidopsis*. *Sci. China Life Sci.* **60**, 520–523 (2017).
55. Hu, B. et al. Engineering non-transgenic gynoecious cucumber using an improved transformation protocol and optimized CRISPR/Cas9 system. *Mol. Plant.* **10**, 1575–1578 (2017).
56. ZhuJ.-Y., Sun, Y. & WangZ.-Y. Genome-wide identification of transcription factor-binding sites in plants using chromatin immunoprecipitation followed by microarray (ChIP-chip) or sequencing (ChIP-seq). *Methods Mol. Biol.* **876**, 173–188 (2012).
57. Simon, J. M., Giresi, P. G., Davis, I. J. & Lieb, J. D. Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA. *Nat. Protoc.* **7**, 256–267 (2012).
58. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
59. Zhang, Y. et al. Model-based analysis of ChIP-seq (MACS). *Genome Biol.* **9**, R137 (2008).
60. Song, Q. & Smith, A. D. Identifying dispersed epigenomic domains from ChIP-seq data. *Bioinformatics* **27**, 870–871 (2011).
61. Boyle, A. P., Guinney, J., Crawford, G. E. & Furey, T. S. F-Seq: a feature density estimator for high-throughput sequence tags. *Bioinformatics* **24**, 2537–2538 (2008).
62. Trapnell, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578 (2012).
63. Bailey, T. L., Williams, N., Misleh, C. & Li, W. W. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.* **34**, 369–373 (2006).

## Acknowledgements

We thank Y. Huang and J. Doebley for comments on the manuscript, J. Lai, K. Liu, H. Deng, S. Fan, X. Chen, L. Guo and Y. Zhao for experimental assistance, and H. Wang for providing TB1 antibodies. This work was supported by National Key R&D Program of China (2019YFA0906200), the National Natural Science Foundation of China (31922076 to X.Y.) and the Central Public-Interest Scientific Institution Basal Research Fund (Y2017PT52). Additional support was provided by the Chinese Academy of Agricultural Science (ASTIP-CAAS, CAAS-XTCX2016001 and the Elite Young Scientists Program and The Agricultural Science and Technology Innovation Program), the Leading Talents of Guangdong Province Program (00201515 to S.H.), the Shenzhen Municipality (The Peacock Plan KQTD2016113010482651) and the Dapeng district government.

## Author contributions

S.H. and X.Y. designed the research. X.Y., Zhen Zhang and B.W. made major contributions to biochemical analyses and ChIP assays. J.Y. contributed to protein purification. T.L., X.Y. and Zhonghua Zhang led bioinformatic analyses. X.Y. and T.X. led genetic transformation of plants. S.W. helped to collect tendrill materials. G.L., J.Z. and Zhen Zhang contributed to histone purification and assembly. S.H., X.Y., G.L., W.J.L. and J.Y. analysed the data and wrote the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Extended data** is available for this paper at <https://doi.org/10.1038/s41477-020-0715-2>.

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41477-020-0715-2>.

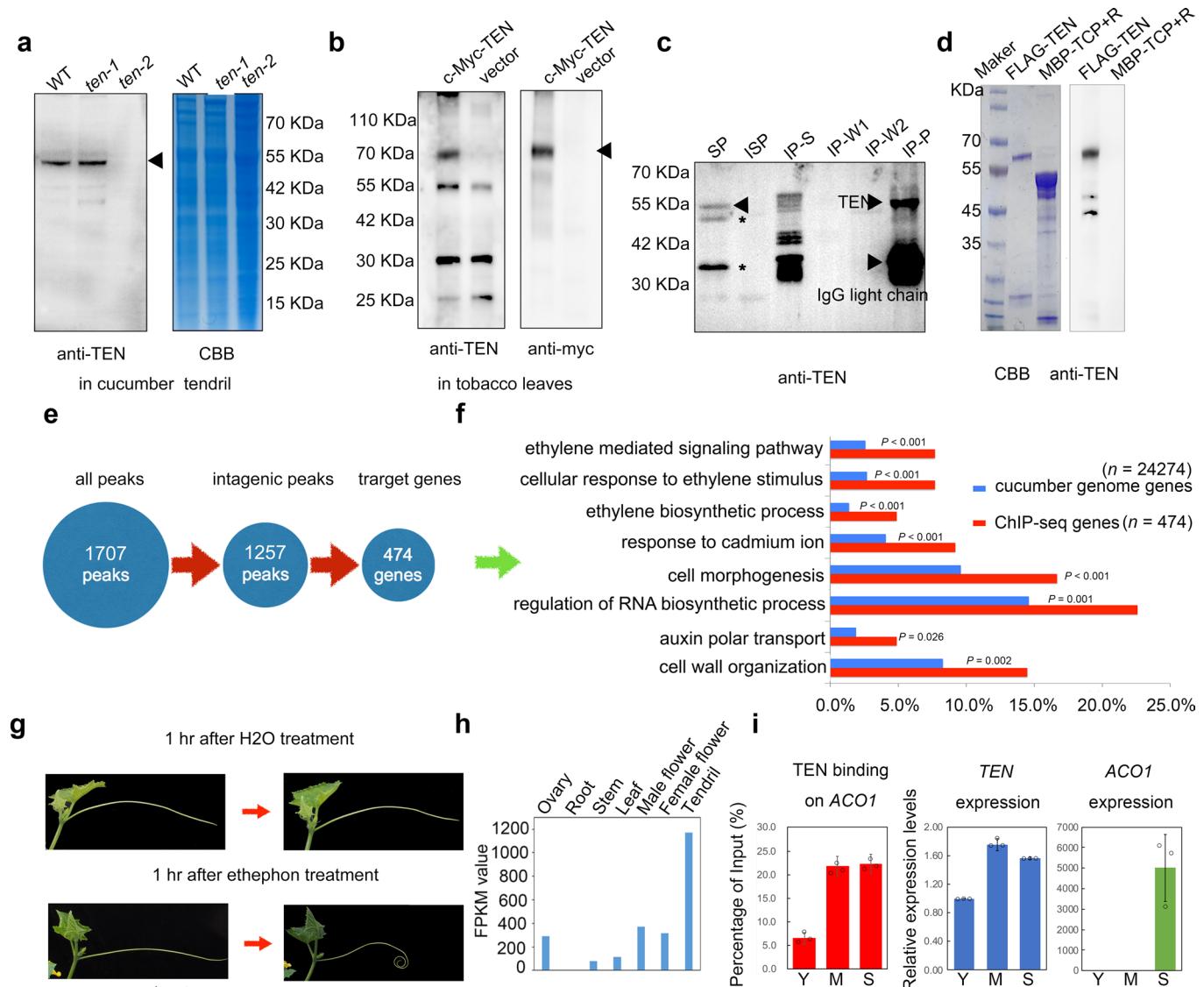
**Correspondence and requests for materials** should be addressed to S.H.

**Peer review information** *Nature Plants* thanks Pilar Cubas, Richard Immink and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

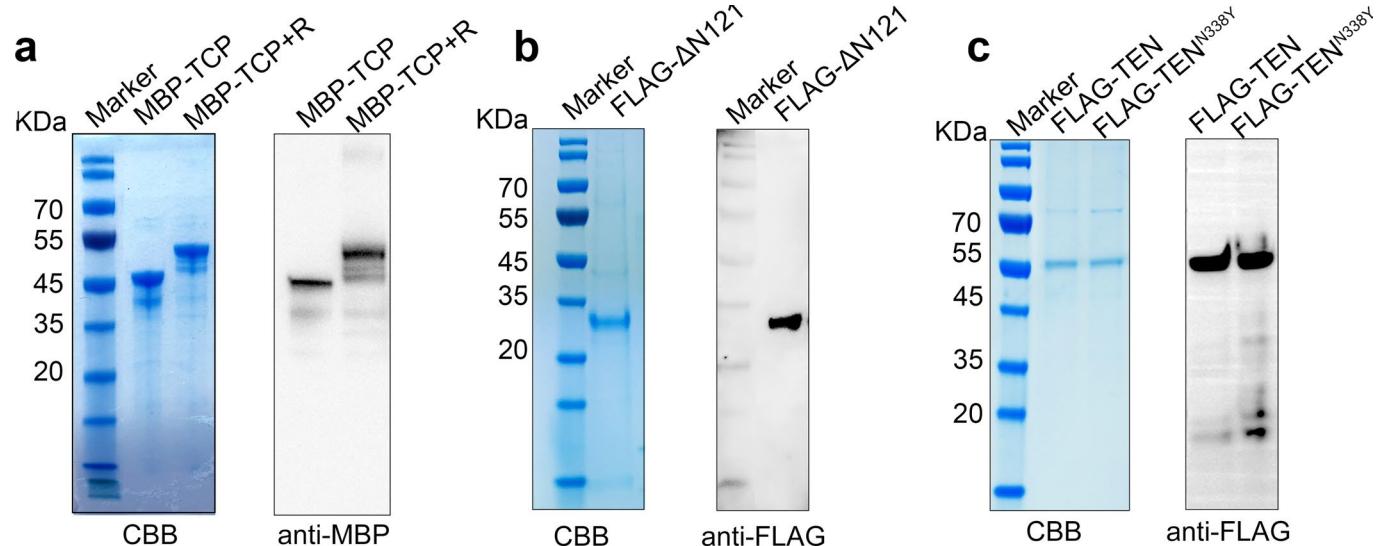
**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

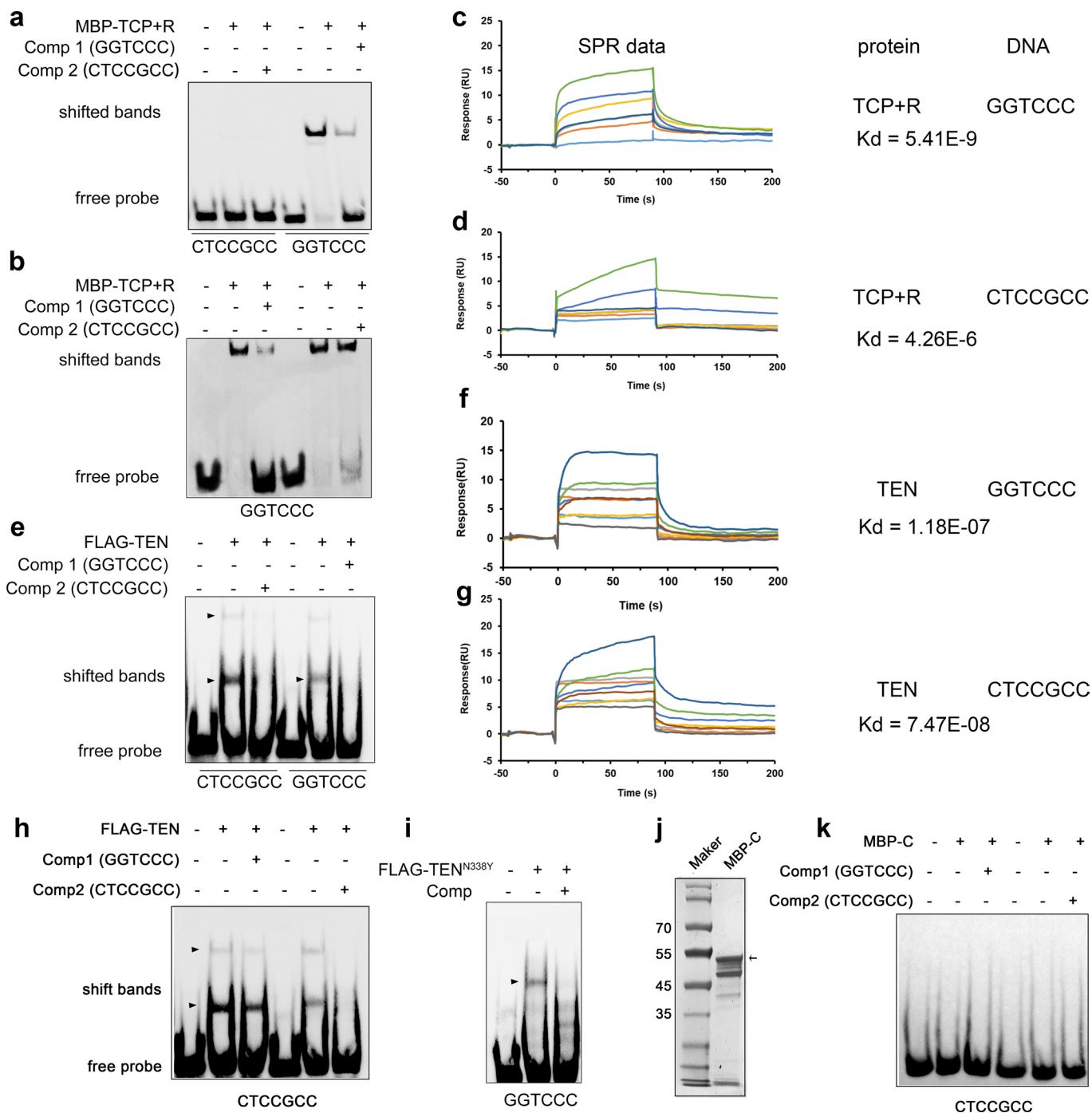
© The Author(s), under exclusive licence to Springer Nature Limited 2020



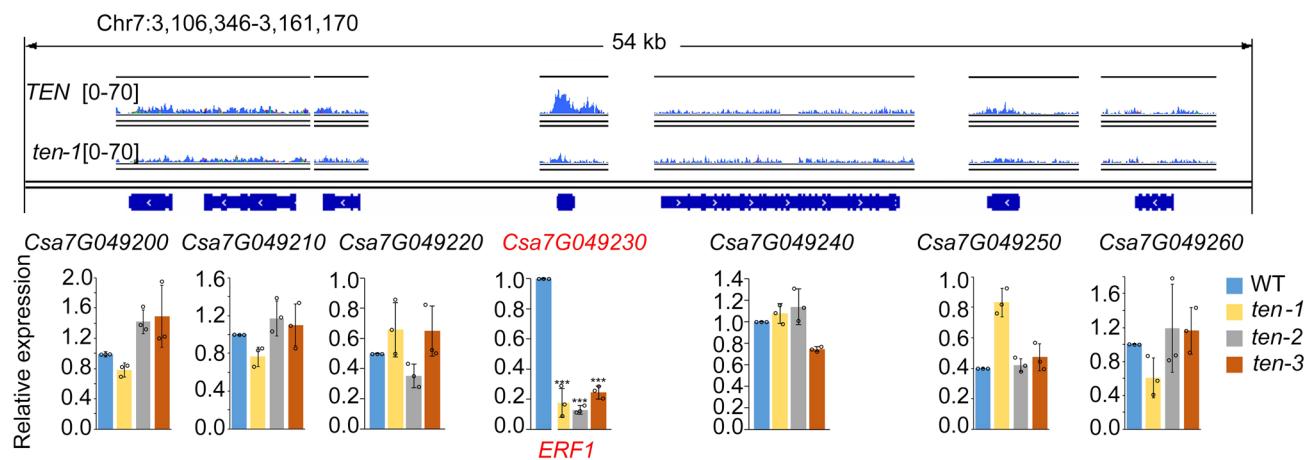
**Extended Data Fig. 1 | Identification of direct target genes of TEN and purification of the TEN truncated proteins.** **a**, Total protein extracts from tendrils or modified tendrils from WT, *ten*<sup>-1</sup> and *ten*<sup>-2</sup> plants were subjected to immunoblotting, using anti-TEN polyclonal antibody. The antibody recognized specifically a 50 KDa band (theoretical molar mass is approx. 45 KDa) in both WT and the *ten*<sup>-1</sup> mutant, but not in the *ten*<sup>-2</sup> null mutant. CBB, Coomassie Brilliant Blue staining. **b**, *pCAMBIA 1300-c-Myc-TEN*, and the corresponding vector, were transiently expressed in tobacco leaves. Total protein was extracted for immunoblotting, using anti-TEN and anti-c-Myc antibodies. Both antibodies recognized specifically the same bands at 70 KDa (5x-c-Myc-tag peptide is 20 KDa). **c**, Total protein extracts from WT tendrils were subjected to immunoprecipitation with anti-TEN antibody. Input and immunoprecipitated protein fractions were analyzed by immunoblotting, using anti-TEN antibodies. SP, soluble protein; ISP, insoluble protein; IP-S, immunoprecipitation-supernatant fraction; W, wash fraction; P, pellet fraction; closed triangles indicate the target bands; asterisk, degradation of TEN protein. **d**, The TEN antibody recognized recombinant full-length TEN, but not its TCP domain. **e**, Schematic of the TEN binding peaks and target genes. **f**, Selected enrichment of GO terms for the 474 intragenic target genes. Significant analysis was done using a Fisher's exact test. **g**, Exogenous treatment of ethephon promotes spontaneous tendril coiling. Spraying with water was carried out as control. **h**, Fragments Per Kilobase of exon per Million fragments mapped (FPKM) values of ACO1 gene in *ten*<sup>-1</sup> different cucumber tissues, showing preferential expression in the tendril. **i**, Relative *TEN* expression levels, *TEN* binding levels to ACO1 and ACO1 expression levels during tendril growth (mean ± s.d., n=3). The images in a-d and g were repeated at least twice, with similar results.



**Extended Data Fig. 2 | Purification of recombinant TEN and the TEN truncated proteins.** (a-c) Purification of MBP-TCP and MBP-TCP+R (a), FLAG-ΔN121 (b), and FLAG-TEN and FLAG-TEN<sup>N338Y</sup> (c) proteins from transfected insect cells. The experiments were repeated at least three times, with similar results.



**Extended Data Fig. 3 | The binding specificity and affinity of the full length TEN to DNA containing the GGTCCC motif and the CTCCGCC motif identified by EMSA and SPR.** **a**, EMSA showing that TCP+R binds to a DNA probe containing the GGTCCC motif but does not bind to a DNA probe containing the CTCCGCC motif. **b**, EMSA showing binding of the TCP domain, to the GGTCCC motif, could not compete with the CTCCGCC motif. **c** and **d**, Surface plasmon resonance (SPR) curves for interactions of TCP+R with the DNA probes containing GGTCCC motif (**c**), or CTCCGCC motif (**d**). **e**, EMSA showing that full length TEN binds to a DNA probe containing the CTCCGCC motif and the GGTCCC motif. **f** and **g**, Surface plasmon resonance (SPR) curves for interactions of full length TEN with the DNA probes containing GGTCCC motif (**f**), or CTCCGCC motif (**g**). **h**, EMSA showing that the C terminus binding to the CTCCGCC motif could not be competed with the GGTCCC motif. **i**, The N338Y mutation, in the C-terminus of TEN, had no effect on TEN binding to the GGTCCC motif. **j**, Purification of MBP-tagged C-terminus of TEN. **k**, MBP-C protein could not bind to CTCCGCC probes. Comp, competitor (unlabeled probe); +/−, presence/absence of protein or competitor. The experiments were repeated twice, with similar results.

**a****b**

## ACO1 genomic sequence

GAGTTCCAATCATCACTTGGAGAACGTTAATGGTGAAGAGAGACTAACTATCTAGCCAAA  
TCAAAGATCTGTGAAACTGGGTTCTTGAGGTAAATTATTTACGTTCTTCTATATAT  
ATATATATAATATATATATATGATCATATAGTTAATTATTTATTAATGATGTTGAGA  
ATATATATGTTGATTAAAGTTGATTAACCTTGTGATTAAATAAGGACTGAACACCGA  
ATAGATATAGAAATTCTAGACAAAGTAGAGAAATTAAACAAAGAACATTACAAGAAATGATGG  
AAGAGAGATTCAAAGAACATAGCCAGAAAGGATTAAACGACGCTGAACAGAAGTCACG  
ATGTTGACTGGAAAGCATTCTTCTCCCGTCAATCTCCGATCTCCCG  
GAATCATCCGACGATTACAATCCGTAATGAAACAAATTCCGGTCAAATTACAAGTACTCCCG  
AACTCCCTGGATTGCTCTCGAAAATCTGACACCTCCAAAATCATCTGAAAAACGCTTC

CACGGATCCAACGGCCCAACTTCGGGACAAAGTCAGCAATTACCCACCTTGCCAAACCCA  
GATCTAAATAAGGCTCCGAGCACACCCGACGCCGGAGGGATAATCTCTATTCAAGAC  
GACAAAGTGAATGGTTGCAACTGTTGAAGGAGGGAGTGGGGACGTGCCGGTCCG  
GCACCTCGATCTGGTGAATATCGGGATCAATTGGAAGTATTACAACGGAAATATAAAAG  
TGTTTGCAACAGAGTAATTGGCGACCCGGAAGGGAAAGGGAGAATGTCGTTGGGTGTTTAC  
AATCCGGAAGTGACCGGTGATATTCTCCCGTCAATCTCCGATCTCCCG  
ATATACCCAAAATTGTATTGAGGATTATGAAAGCTTATGTCGACTGAAATTGACCCAA  
GGAGCCGAGGTTGAAGCCATGAAGGGATAATGTCAAATTGATCCTATTGCAACCGTT

Intragenic enhancer  
candidate of ACO1

**K E G E**

CTCCNCCN motif  
amino acid sequence

GACAAAGTGACTGGTTGCAACTGTTGAAGAACCGAGAATGGGTGGACGTGCCGGTCCGGCA  
CTCGATCTGTGTAATCGGGATCAATTGGAAGTATTACAACGGAAATATAAAAGTGTTC  
CACAGAGTAATTGCGCAGCCGGAAGGGAGAATGTCGTTGGCGTGTGTTACAATCCGG  
AAAGTGACCGGGTGTATTCGGCGCCGAGTTCTGGGGCCAGGAAGGGAGAATACCCAAA  
ATTGTATTGAGGATTATGAAAGCTTATGCTGGA

**L V A E E**

Synonymous mutations

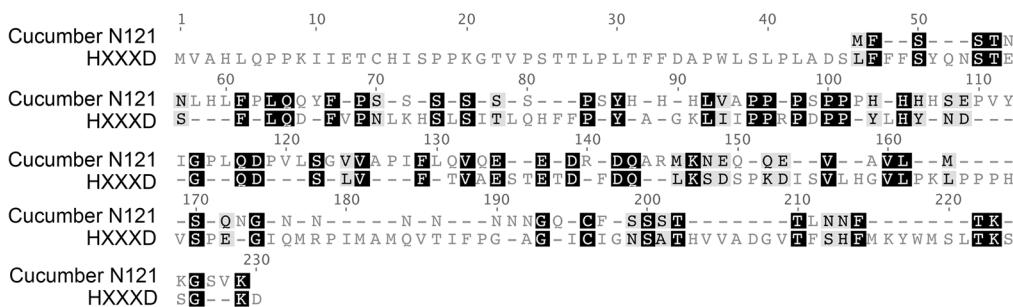
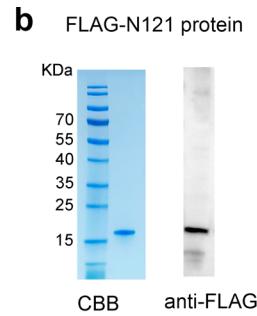
Mutant intragenic enhancer  
candidate of ACO1

**K E G E**

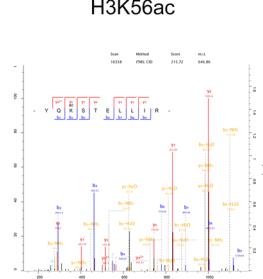
GACAAAGTGAGTGGTTGCAACTGTTGAAGAACCGAGAATGGGTGGACGTGCCGGTCCGGCA  
CTCGATCTGTGTAATCGGGATCAATTGGAAGTATTACAACGGAAATATAAAAGTGTTC  
CACAGAGTAATTGCGCAGCCGGAAGGGAGAATGTCGTTGGCGTGTGTTACAATCCGG  
AAAGTGACCGGGTGTATTCGGCGCCGAGTTCTGGGGCCAGGAAGGGAGAATACCCAAA  
ATTGTATTGAGGATTATGAAAGCTTATGCTGGA

**L V A E E**

**Extended Data Fig. 4 | Functional validation for the importance of intragenic enhancer in ACO1 gene expression.** **a**, Expression of putative enhancer target gene, *ERF1*, and flanking genes, assayed by RT-qPCR (mean  $\pm$  s.d.,  $n=3$ ). *UBQ* was used as internal control. **b**, Schematic showing the design of synonymous mutations in the CTCCNCCN motifs for validation of ACO1 intragenic enhancers.

**a****b****c**

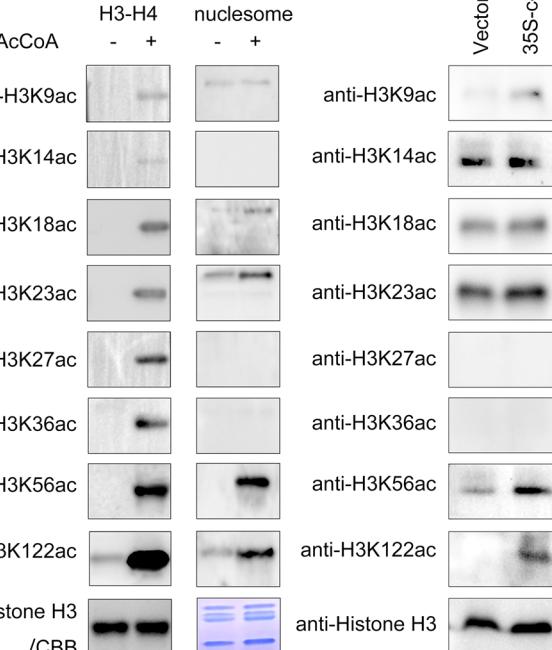
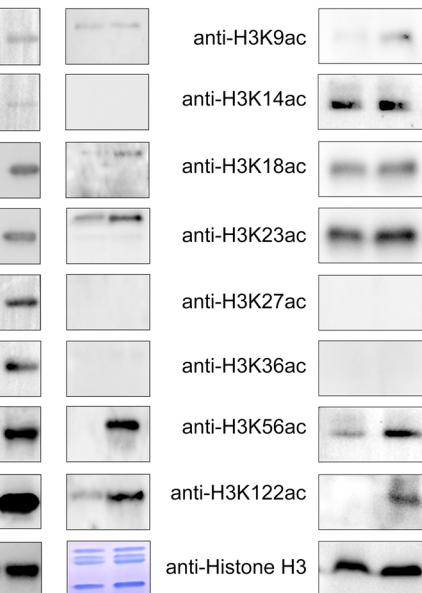
H3K56ac

**f**

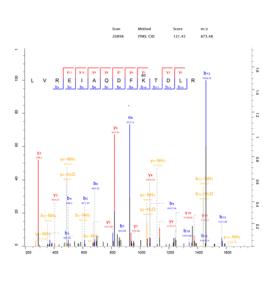
FLAG-N121

**g**

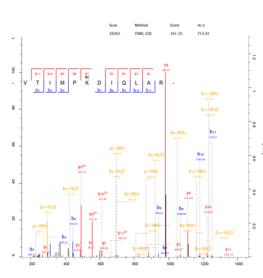
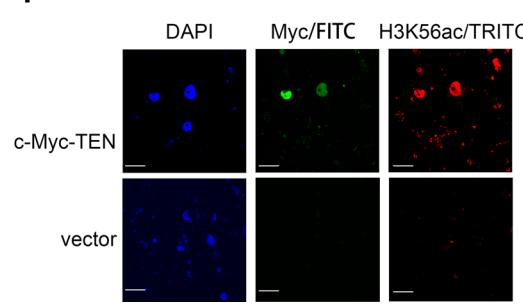
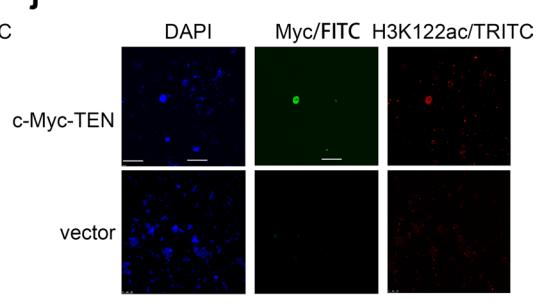
FLAG-TEN

**h**Vector  
35S-c-Myc-TEN**d**

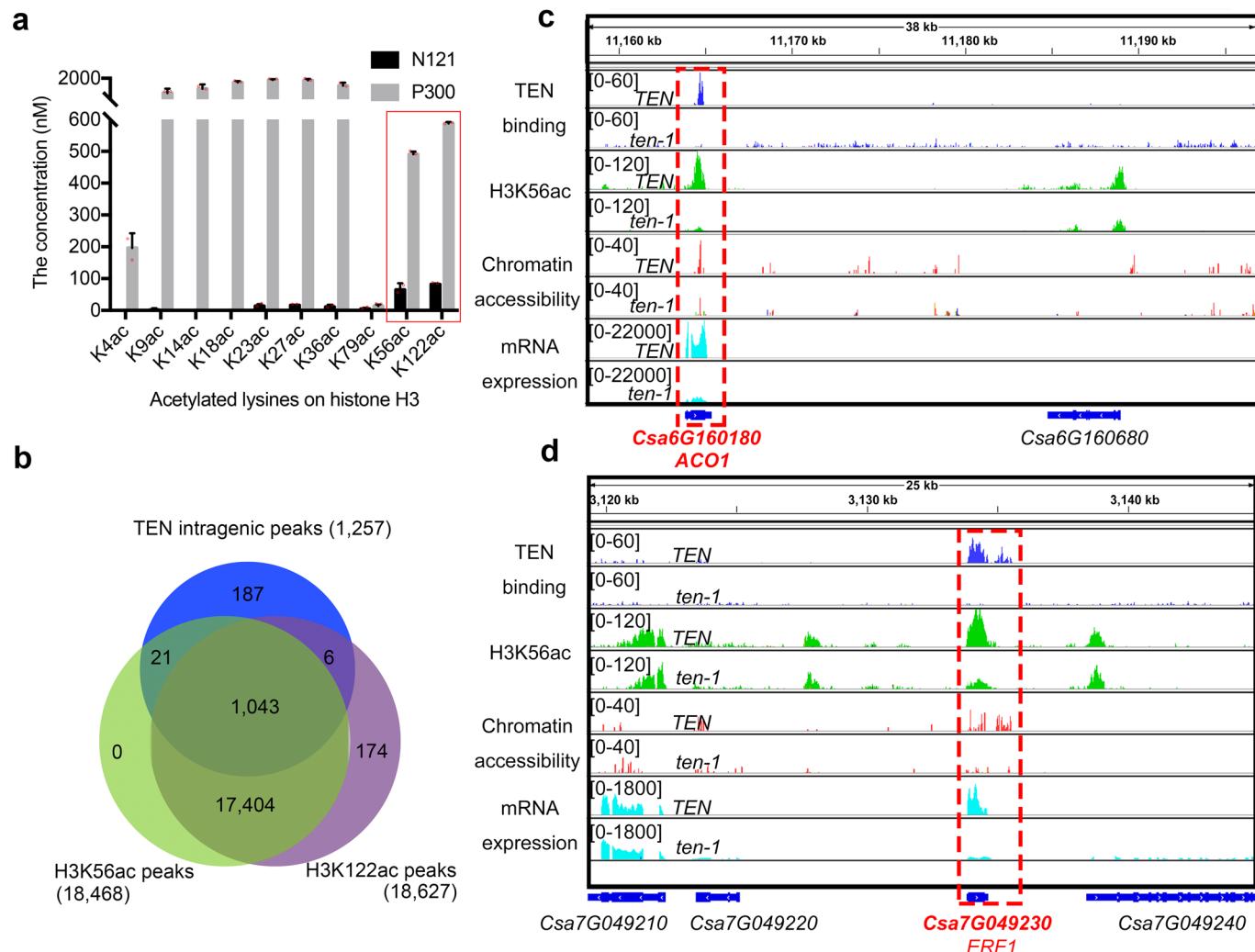
H3K79ac

**e**

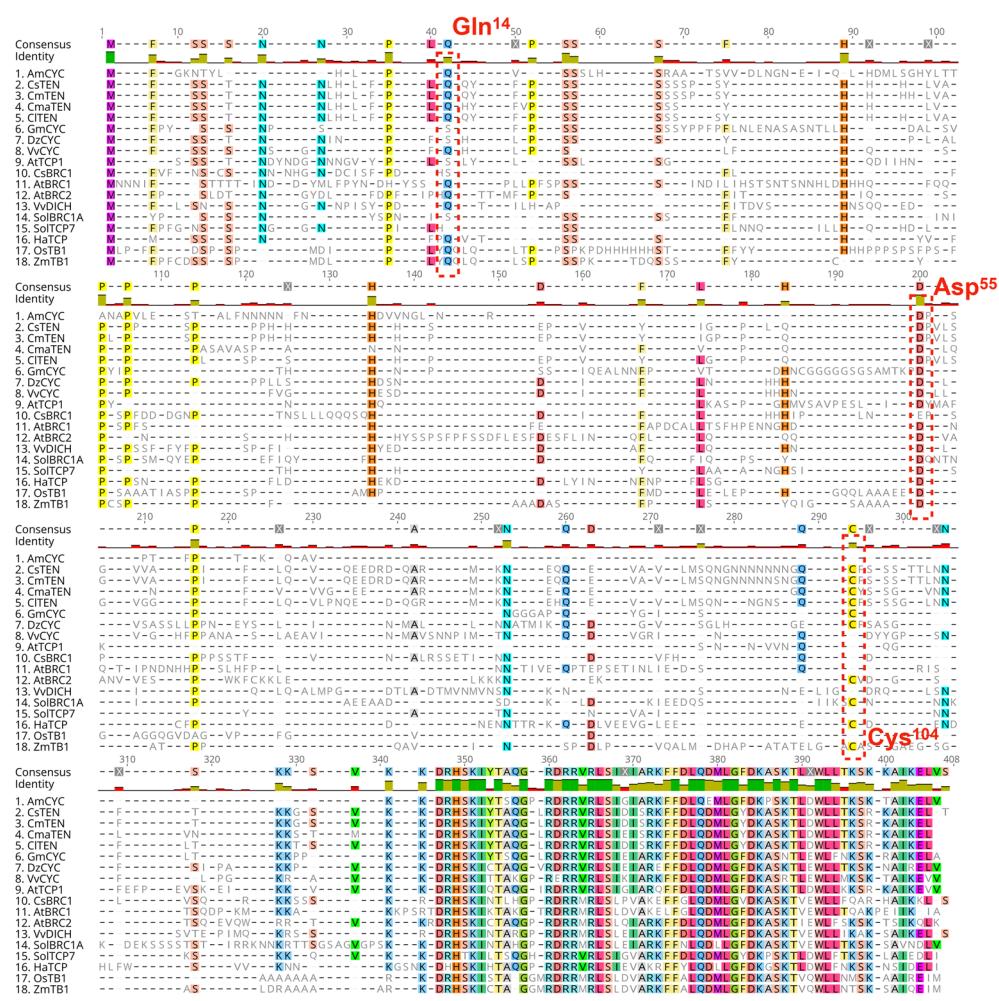
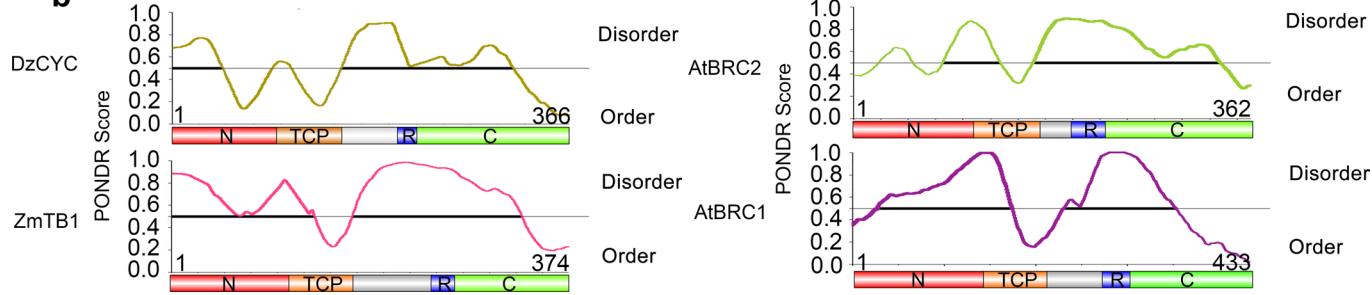
H3K122ac

**i****j**

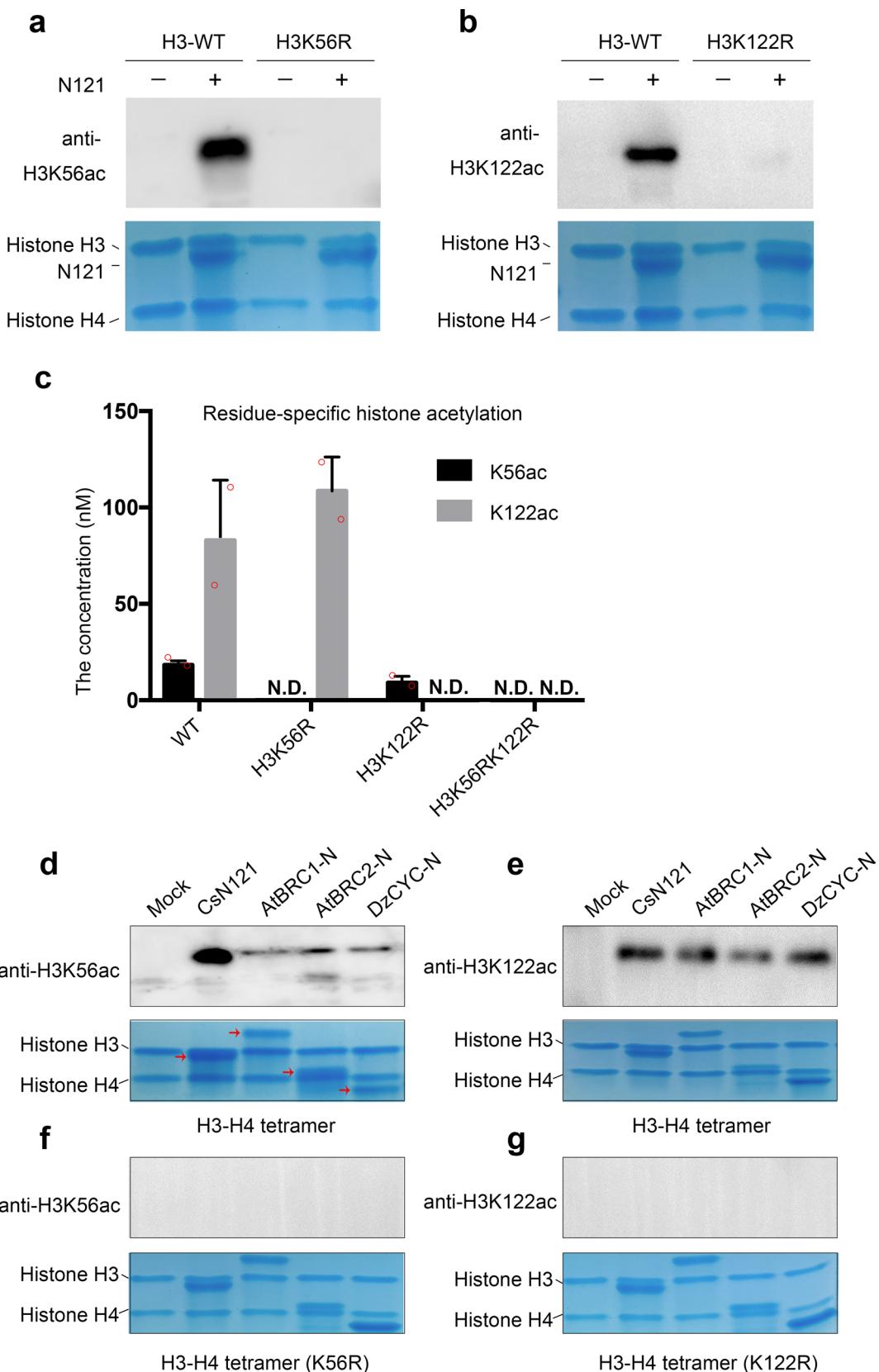
**Extended Data Fig. 5 | HAT activity of TEN demonstrated by LC-MS/MS analysis and immunoblotting assays.** **a**, PSI-BLAST analysis showing sequence similarity between N121 and the transferase domain within an *Arabidopsis* HXXXD acetyltransferase, At1G03940. **b**, Purification of FLAG-N121 protein from insect cells; protein was identified by immunoblotting using anti-FLAG antibody. **c-e**, MS/MS spectrum of the charged ions corresponding to the H3 peptides where K56, K79 or K122 is acetylated. **f-g**, HAT assays of H3-H4 tetramer or nucleosome with FLAG-N121 (f) or FLAG-TEN proteins (g), followed by immunoblotting with antibodies specific for different acetylated sites. **h**, Acetylation of different lysine sites in histone H3, analyzed in tobacco leaves overexpressing c-Myc-TEN. Membrane blotted with H3K122ac was re-probed with anti-H3 to confirm equal loading. **i-j**, Overexpression of c-Myc-TEN promoted acetylation of chromatin-bound H3K56 and K122. Histone acetylation is shown in red (right), and nuclei transfected with c-Myc-TEN were visualized by FITC signal (green; middle). Nuclei were stained with DAPI (blue; left). Scale bar, 2 μm. The experiments in b-j were repeated at least 3 times, with similar results.



**Extended Data Fig. 6 |** TEN promotes histone acetylation and chromatin accessibility at its gene targets. **a**, Quantitative comparison of the histone acetyltransferase activity of N121 with a canonical HAT, CBP/P300, by quantitative mass spectrometry (mean  $\pm$  s.d.,  $n=2$ ). **b**, Metagene analysis showing genome-wide colocalization of TEN intragenic peaks with H3K56ac and H3K122ac in tendrils. **c, d**, Genomic browser tracks showing the TEN binding levels, H3K56ac levels, chromatin accessibility and mRNA expression at the TEN intragenic binding targets of *ACO1* (**c**) and *ERF1* (**d**).

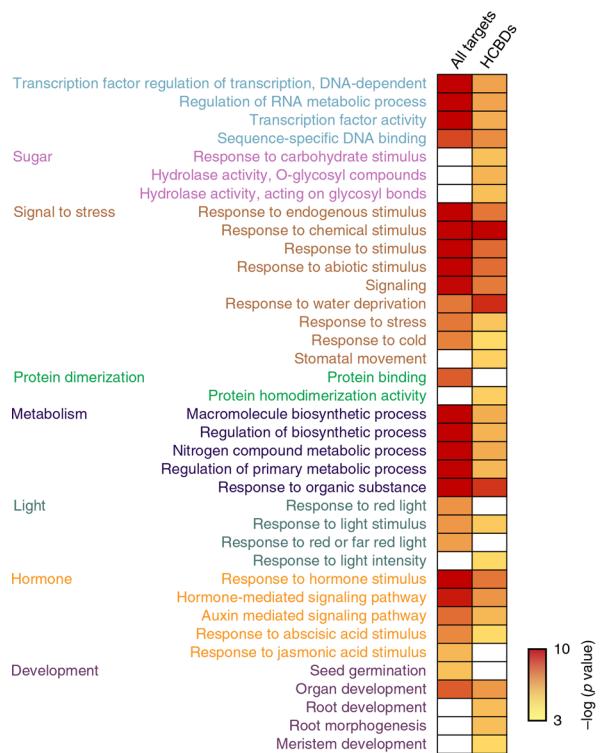
**a****b**

**Extended Data Fig. 7 | Alignment of the IDRs of TB1-like proteins.** **a**, Alignment of the N-terminus of CYC/TB1-like proteins among selected angiosperm species, including *Antirrhinum majus* CYC, cucumber TEN, melon TEN, pumpkin TEN, watermelon TEN, soybean CYC, *Durio* CYC, grape CYC, *Arabidopsis* TCP1, cucumber BRC1, *Arabidopsis* BRC1, *Arabidopsis* BRC2, grape DICH, potato BRC1A, potato TCP7, sunflower TCP, rice TB1 and maize TB1. Red dashed boxes indicate the Gln<sup>14</sup>, Asp<sup>55</sup> and Cys<sup>104</sup> amino acid residues. **b**, Schematic of the different domains of four TB1-like proteins. Graphs plotting intrinsic disorder (PONDR VL3-BA) for selected TB1-like proteins. PONDR VL3-BA score (y-axis) and amino acid position (x-axis) are shown, indicating the intrinsically disordered and ordered regions.

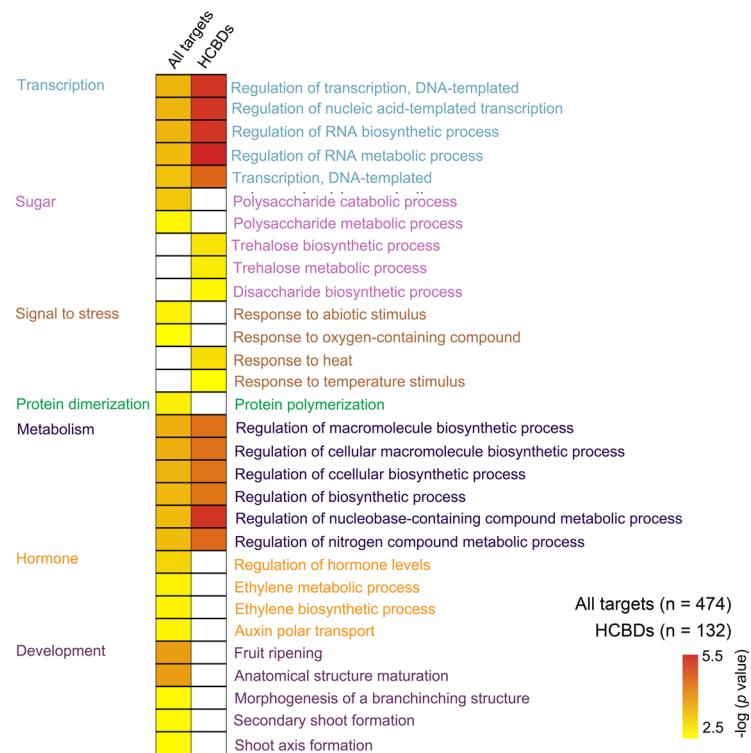


**Extended Data Fig. 8 | HAT assays with mutant H3-H4 tetramer (H3K56R and H3K122R) by N121.** **a, b**, Acetylation of H3K56 and K122 within the wild type H3-H4 tetramer or mutant H3-H4 tetramer (H3K56R or H3K122R) by N121, determined by immunoblotting analysis. **c**, Acetylation of H3K56 and K122 within the wild type H3-H4 tetramer or mutant H3-H4 tetramer (H3K56R or H3K122R) by N121, determined by LC-MS/MS (mean  $\pm$  s.d.,  $n=2$ ). **d-g**, Acetylation of H3K56 and K122 within the wild type H3-H4 tetramer or mutant H3-H4 tetramer (H3K56R or H3K122R) by N terminus proteins of TEN, BRC1, BRC2 and durian CYC, determined by immunoblotting analysis. The representative images in a-b and d-g were repeated twice, with similar results.

## Maize-TB1



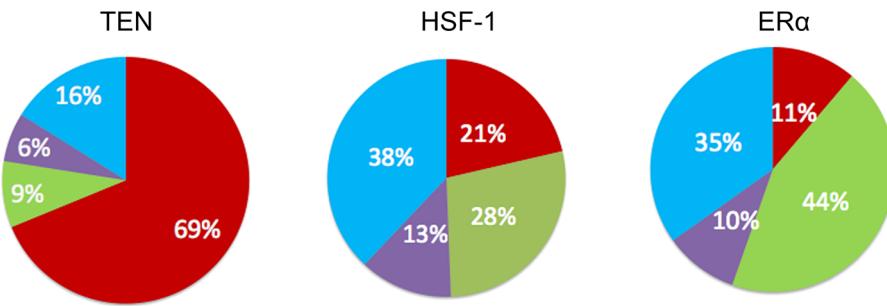
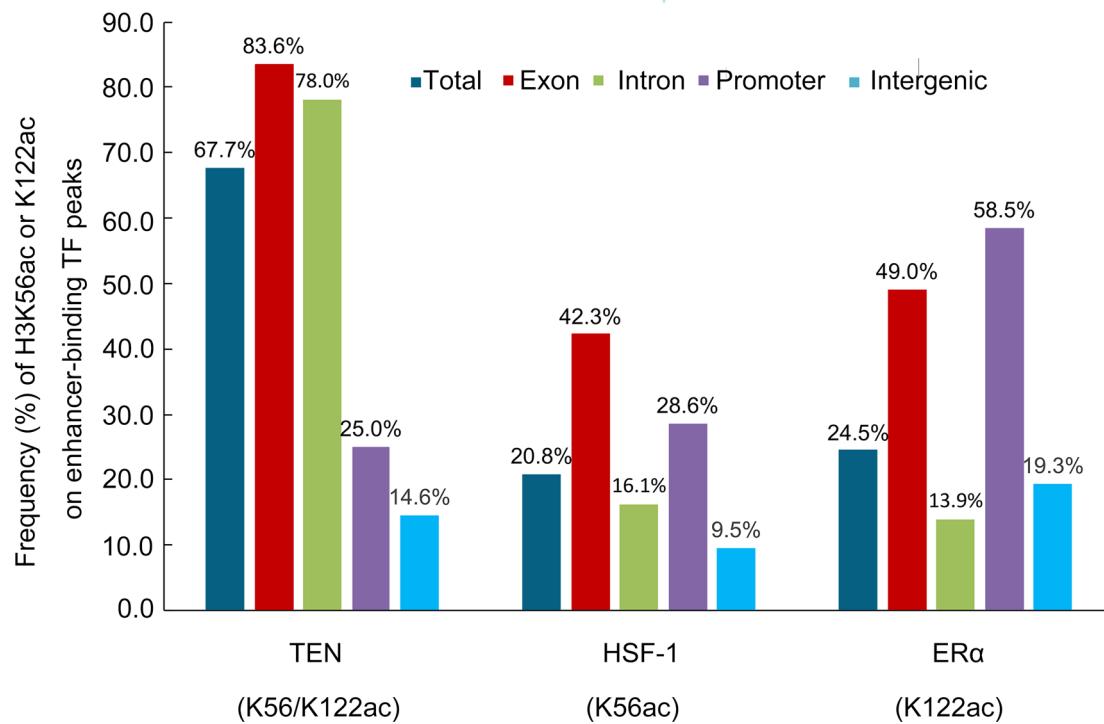
## Cucumber-TEN



**Extended Data Fig. 9 | Comparison of the regulatory datasets between TB1 and TEN.** All targets of TB1 and TEN could be classified within the categories including transcription factors, sugar metabolism, signal to stress, biosynthetic metabolism, phytohormone pathway or development. TEN target datasets did not show GO enrichment in light response, which was present in the GO results for the TB1 targets. HCBD, high confidence bound DEGs. Significance analysis was done using a Fisher's exact test.

**a**

## Peaks distribution of enhancer-binding TFs

**b**

**Extended Data Fig. 10 | H3K56ac and H3K122ac are more likely to occur on exonic regions bound by TEN, HSF-1 and ER $\alpha$ .** **a**, Distribution of binding peaks of enhancer-binding TFs, TEN, HSF-1 and ER $\alpha$ . **b**, Frequency of H3K56ac and/or H3K122ac on TEN, HSF-1 and ER $\alpha$  peaks, showing that histone globular acetylation is a general phenomenon for transcriptional regulation on exonic enhancers.

Corresponding author(s): Sanwen Huang

Last updated by author(s): May 21, 2020

## Reporting Summary

Nature Research 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 Research policies, see [Authors & Referees](#) 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

The reads of High-throughput sequencing were collected using Illumina CASAVA 1.8 software. The MS/MS spectra from each LC-MS/MS run were collected using Proteome Discovery searching algorithm (version 1.4).

Data analysis

Data were analysed using bowtie software (<http://bowtie-bio.sourceforge.net>), MACS (<http://liulab.dfci.harvard.edu/MACS/>), RSEG software (<https://github.com/smithlabcode/rseg>), F-Seq software (<http://fureylab.web.unc.edu/software/fseq/>), deeptools software (<https://github.com/deeptools/deepTools>), tophat2 software (<http://tophat.ccbcb.umd.edu/>), MEME-ChIP software (the online version 5.0.5), cufflinks (version 2.2.1) and Prism (version 7.0).

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/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [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 list of figures that have associated raw data
- A description of any restrictions on data availability

The data supporting the findings in this study are available from the corresponding author upon reasonable request.

# 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](http://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	The sample size used in a study is determined based on the need to have sufficient statistical power, and based on accepted sample sizes in relevant published reports within this field. . The sample size for each experiment was described completely in the Figure and Table legends or in Methods.
Data exclusions	Due to the several steps of the ChIP protocol, sometimes the assays failed and these results were not analyzed.
Replication	All attempts at replication were successful and similar results were obtained.
Randomization	As a general rule, plants (wt or transgenic) were all exposed to the same growth condition and treatment. And the sampling was random.
Blinding	The investigators were blinded to group allocation during data collection and analysis.

## 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		Methods	
n/a	Involved in the study	n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies	<input type="checkbox"/>	<input checked="" type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Eukaryotic cell lines	<input type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology	<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants		
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data		

## Antibodies

Antibodies used	<p>The TEN protein antibody was homemade.</p> <p>Commercial antibodies used:</p> <p>Anti-Histone H3 (acetyl K9) , Abcam, Cat#ab4441, Lot#GR304381-1, Rabbit polyclonal antibody, 1:10000 dilution in vitro, 1:10000 dilution in vivo;</p> <p>Anti-Histone H3 (acetyl K14) ,Abcam, Cat#ab52946, Lot#GR302893-1, Rabbit monoclonal antibody, 1:5000 dilution in vitro, 1:5000 dilution in vivo;</p> <p>Anti-Histone H3 (acetyl K18), Merck-Millipore, Cat#07-354, Lot#2842177, Rabbit polyclonal antibody, 1:10000 dilution in vitro, 1:10000 dilution in vivo;</p> <p>Anti-Histone H3 (acetyl K23), Merck-Millipore, Cat#07-355, Lot#2733792, Rabbit polyclonal antibody, 1:2000 dilution in vitro, 1:2000 dilution in vivo;</p> <p>Anti-Histone H3 (acetyl K27), Merck-Millipore, Cat#07-360, Lot#2806129, Rabbit monoclonal antibody, 1:10000 dilution in vitro, 1:10000 dilution in vivo;</p> <p>Anti-Histone H3 (acetyl K36), Active motif, Cat#39379, Lot#29108001, Rabbit polyclonal antibody, 1:2000 dilution in vitro, 1:2000 dilution in vivo;</p> <p>Anti-Histone H3 (acetyl K56), Active motif, Cat#39282, Lot#25817004, Rabbit polyclonal antibody, 1:3000 dilution in vitro, 1:1000 dilution in vivo;</p> <p>Anti-Histone H3 (acetyl K122), Abcam, Cat#ab33309, Lot#GR314063-2, Rabbit polyclonal antibody, 1:5000 dilution in vitro, 1:500 dilution in vivo;</p> <p>Anti-Histone H3, Abcam, Cat#ab1791, Lot#GR300976-3, Rabbit polyclonal antibody, 1:5000 dilution in vitro, 1:5000 dilution in vivo;</p> <p>Anti-Myc-tag mAb, MBL, Cat#M192-3, Lot#005, Mouse monoclonal antibody, 1:3000 dilution in vivo;</p> <p>Anti-FLAG mAb, Sigma, Cat#F3165, Mouse monoclonal antibody, 1:5000 dilution in vitro;</p> <p>Anti-MBP mAb, NEB, Cat#E8032L, Mouse monoclonal antibody, 1:10000 dilution in vitro.</p>
-----------------	---

## Validation

The TEN protein antibody was homemade, and the species and validation has been showed in the manuscript. The other antibodies are commercially available, have been validated by the respective company and/or in prior publications. These antibodies are commonly used in molecular biology research.

Anti-Histone H3 (acetyl K9) (Abcam, ab4441): <https://www.abcam.com/histone-h3-acetyl-k9-antibody-chip-grade-ab4441.html>. This antibody is published in plant science-related studies (PMID: 31338715). It has dot blot verified specificity and has been validated in ChIP, WB and dot blot.

Anti-Histone H3 (acetyl K14) (Abcam, ab52946): <https://www.abcam.com/histone-h3-acetyl-k14-antibody-ep964y-chip-grade-ab52946.html>. This antibody is published in plant science-related studies (PMID: 31182557, PMID: 30346270). It has dot blot verified specificity and has been validated in ChIP, WB and IHC-Wholmount.

Anti-Histone H3 (acetyl K18) (Merck-Millipore, 07-354): [https://www.merckmillipore.com/NL/en/product/Anti-acetyl-Histone-H3-Lys18-Antibody,MM\\_NF-07-354?ReferrerURL=https%3A%2F%2Fwww.google.com%2F&bd=1](https://www.merckmillipore.com/NL/en/product/Anti-acetyl-Histone-H3-Lys18-Antibody,MM_NF-07-354?ReferrerURL=https%3A%2F%2Fwww.google.com%2F&bd=1). This antibody is published in plant science-related studies (PMID: 19643030, PMID: 23564230, PMID: 22733760). It has been validated in ChIP, ChIP-seq, WB and dot blot.

Anti-Histone H3 (acetyl K23) (Merck-Millipore, 07-355): [https://www.merckmillipore.com/NL/en/product/Anti-acetyl-Histone-H3-Lys23-Antibody,MM\\_NF-07-355?ReferrerURL=https%3A%2F%2Fwww.google.com%2F&bd=1](https://www.merckmillipore.com/NL/en/product/Anti-acetyl-Histone-H3-Lys23-Antibody,MM_NF-07-355?ReferrerURL=https%3A%2F%2Fwww.google.com%2F&bd=1). This antibody is published in plant science-related studies (PMID: 19643030, PMID: 19954517, PMID: 21552333). It does not recognize unacetylated recombinant histone H3 and has been validated in ChIP and WB.

Anti-Histone H3 (acetyl K27) (Merck-Millipore, 07-360): [https://www.merckmillipore.com/NL/en/product/Anti-acetyl-Histone-H3-Lys27-Antibody-Trial-Size,MM\\_NF-07-360-S?ReferrerURL=https%3A%2F%2Fwww.google.com%2F&bd=1](https://www.merckmillipore.com/NL/en/product/Anti-acetyl-Histone-H3-Lys27-Antibody-Trial-Size,MM_NF-07-360-S?ReferrerURL=https%3A%2F%2Fwww.google.com%2F&bd=1). This antibody is published in plant science-related studies (PMID: 21317377, PMID: 17085686). It has dot blot verified specificity and has been validated in ChIP, ChIP-seq, WB and Mplex.

Anti-Histone H3 (acetyl K36) (Active motif, 39379): <https://www.activemotif.com/catalog/details/39379/histone-h3-acetyl-lys36-antibody-pab>. This antibody is published in plant science-related studies (PMID: 30765479, PMID: 29187567). It has dot blot verified specificity and has been validated in ChIP, WB and Immunofluorescence.

Anti-Histone H3 (acetyl K56) (Active motif, 39282): <https://www.activemotif.com/catalog/details/39281/histone-h3-acetyl-lys56-antibody-pab>. This antibody is commonly used in molecular biology research (PMID: 31216030, PMID: 29759984). It has dot blot verified specificity and is validated in ChIP, ChIP-seq and WB.

Anti-Histone H3 (acetyl K122) (Abcam, ab33309): <https://www.abcam.com/histone-h3-acetyl-k122-antibody-ab33309.html>. This antibody is published in molecular biology research (PMID: 23415232, PMID: 27436229). It is validated in ChIP, ChIP-seq and WB. This antibody shows no cross reactivity with acetyl K56 in Western Blot. Slight cross reactivity with acetyl K56 may be observed in Dot Blot.

Anti-Histone H3 (Abcam, ab1791): <https://www.abcam.com/histone-h3-antibody-nuclear-loading-control-and-chip-grade-ab1791.html>. This antibody is commonly used in plant science-related studies (PMID: 30712008, PMID: 30670606, PMID: 31184697) and is validated in ChIP, ChIP-seq and WB.

Anti-Myc-tag mAb (MBL, M192-3): <https://ruo.mbl.co.jp/bio/e/dt1/A/?pcd=M192-3>. This antibody is commonly used in plant science-related studies (PMID: 29093216) and is validated in ChIP, Immunoprecipitation and WB.

Anti-FLAG mAb (Sigma, F3165) : [https://www.sigmal Aldrich.com/catalog/product/sigma/f3165?lang=en&region=NL&gclid=EAAlQobChMlvPj\\_rK7v5wlVGed3Ch2UBgFaEAAYASAAEgJWQPD\\_BwE](https://www.sigmal Aldrich.com/catalog/product/sigma/f3165?lang=en&region=NL&gclid=EAAlQobChMlvPj_rK7v5wlVGed3Ch2UBgFaEAAYASAAEgJWQPD_BwE). It detects a single band of protein on a Western blot from an E. coli crude cell lysate and is validated in ChIP, ChIP-seq and WB.

Anti-MBP mAb (NEB, E8032L): <https://international.neb.com/products/e8032-anti-mbp-monoclonal-antibody#Product%20Information>. This antibody is published in plant science-related studies (PMID: 32296066, PMID: 32240168). It is verified for use in both Western blotting and ELISA.

## Eukaryotic cell lines

Policy information about [cell lines](#)

Cell line source(s)

Sf9 cells were purchased from ThermoFisher Scientific.

Authentication

The cell line that was used was a commercial sf9 cell line (ThermoFisher). No further characterization of the cell lines were performed.

Mycoplasma contamination

Sf9 cells were tested regularly for Mycoplasma contamination.

Commonly misidentified lines  
(See [ICLAC](#) register)

No commonly misidentified cell lines were used.

## 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.

We have deposited the bam files of ChIP-seq data to the Sequence Read Archive: <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA520931>

Files in database submission

CUC-FAIRE, F-INP, CUC-56, CUC-122, TEN-INP, H3-INP, CUC-TEN1, CUC-TEN2

Genome browser session  
(e.g. [UCSC](#))

N.A.

## Methodology

### Replicates

Two replicates for ChIP-seq. the Pearson correlation coefficients ( $r$ ) between the read coverage of peaks from two of the replicates are greater than 0.7.

### Sequencing depth

For all ChIP-seq, 50SE, Hiseq 2500 and at least 30M reads data.

### Antibodies

The following antibodies were used for ChIP assays: TEN antibody, anti-H3K56ac (active motif, Cat. #39282) and anti-H3K122ac (Abcam, Cat. #Ab33309). The antibodies information was provided in the supplementary table. We validated the TEN's antibody, and the validation statements for both antibodies can be found on the manufacturer's website.

### Peak calling parameters

The command line program and parameters used for read mapping of ChIP and control (input) sample:  
bowtie2 -x tair10 -1 ChIP\_R1\_paired.fastq.gz -2 ChIP\_R2\_paired.fastq.gz | samtools view -Sh -q 30 -F 4 - | grep -v 'XS:' | samtools view -Shub | samtools sort - -T ChIP -o ChIP.bam;  
bowtie2 -x tair10 -1 input\_R1\_paired.fastq.gz -2 input\_R2\_paired.fastq.gz | samtools view -Sh -q 30 -F 4 - | grep -v 'XS:' | samtools view -Shub | samtools sort - -T input -o input.bam.  
The command line program and parameters used for peak calling: macs14 -t ChIP.bam -c input.bam -g 119667750 -p 0.01 -nomodel.

### Data quality

Total number of peaks at FDR 5%.  
We calculated the read coverage of peaks and performed the correlation analysis to ensure the reproducibility of the biological replicates.

### Software

bowtie software, <http://bowtie-bio.sourceforge.net>; MACS, <http://liulab.dfci.harvard.edu/MACS/>; RSEG, <http://smithlabresearch.org/software/rseg/>; F-Seq software, <http://fureylab.web.unc.edu/software/fseq/>; tophat2 software, <http://tophat.cbcb.umd.edu/>; Cufflink software, <http://cole-trapnell-lab.github.io/cufflinks/>.