Title: Transcription elongation is dictated by single residues in the histone core domain

Authors: Zachary H. Harvey,[1,*] Kathryn M. Stevens,[2,3] Tobias Warnecke,[2,3,4,**] Frédéric Berger[1,***]

Affiliations:

[1]Gregor Mendel Institute, Austrian Academy of Sciences, Vienna, Austria

[2]MRC Laboratory of Medical Sciences, London, United Kingdom

[3]Institute of Clinical Sciences, Imperial College London, London, United Kingdom

[4]Trinity College, Oxford, United Kingdom

*Correspondence: zachary.harvey@gmi.oeaw.ac.at

**Correspondence: tobias.warnecke@lms.mrc.ac.uk

***Correspondence: frederic.berger@gmi.oeaw.ac.at

Abstract: The chromatin fiber is thought to suppress transcription in eukaryotes by acting as a structural barrier. However, once begun, transcription can readily proceed on chromatin, suggesting this model is insufficient. Here, we establish that the ultra-conserved core domain of the ancestral histone H2A.Z dictates transcription elongation via direct interaction of its loop 2 region with the RNAPII subunit Spt6, rather than biophysical changes to chromatin. Interrogating H2A.Z sequences representing more than a billion years of eukaryotic evolution in a single synthetic host, we show that Spt6 can distinguish even single-residue substitutions within their loop 2, driving either super-repressed or -activated transcriptional states. Our results place the histone core domain at the origin of eukaryotic gene expression, establishing its transformative power to shape transcription.

Main Text

Eukaryotic RNA polymerase II (RNAPII) transcribes mRNA from a chromatin fiber organized into repeating units of ~147 bp of DNA wrapped around a multi-subunit protein complex, the nucleosome (*1*). An enormous body of work has established the importance of nucleosomes in regulating transcription (*2-7*) in development (*8, 9*), disease (*10-12*), and in response to stress (*13*). The prevailing model is that nucleosomes suppress transcription as structural obstacles to RNAPII, and that this conflict must be alleviated to achieve context-specific gene expression (*14-18*). However, recent reports have established that RNAPII has sophisticated mechanisms for navigating nucleosomes without disassembly during transcription elongation (*19, 20*), and once transcription is initiated, eukaryotic RNAPII can readily traverse a chromatinized template (*21-23*), suggesting that the chromatin fiber is not inherently and uniformly suppressive of transcription.

One explanation for the role of the chromatin fiber in shaping transcription is through changes in its composition via incorporation of alternative histone forms, referred to as variants (*4, 5, 24-28*). Consistent with this view, the suite of histone variants found within the transcriptional unit varies according to transcriptional state (*29-31*), with distinct profiles for either repressed (*4, 5*) or activated (*2-4, 11, 26, 32-54*) chromatin domains. Further, histone variants often have divergent biophysical properties *in vitro* (*41, 42*). However, most are highly structurally similar (*55, 56*), and how these physical properties impact transcription *in vivo* remains unresolved.

To address whether and how changes in the chromatin fiber impact transcription, we developed a synthetic complementation paradigm using the histone H2A variant H2A.Z as a model. H2A.Z is an essential feature of eukaryotic chromatin, and has been tightly linked to both transcription activation (*2, 3, 26, 32, 38, 47, 49*) and repression (*11, 33, 48, 50-53*) in different species. We reasoned that we could leverage divergent observations of H2A.Z

function across species as a natural toolkit to generate functional variation in the chromatin fiber. We interrogated H2A.Z sequence variation representing more than a billion years of eukaryotic evolution in a single experimentally tractable synthetic system, the fission yeast *Schizosaccharomyces pombe*, replacing the endogenous *S. pombe* H2A.Z (Pht1) with H2A.Z sequences from diverse eukaryotes. Remarkably, we identify that the RNAPII subunit Spt6 directly reads the primary sequence of the loop 2 (L2) region of the H2A.Z core domain, with even single amino acid changes to its core motif being sufficient to dictate transcription elongation. Together, our findings indicate that an ancient interaction between H2A.Z and Spt6 informs RNAPII directly of the nature of the chromatin template, providing insight into its fundamental role in shaping eukaryotic gene expression, and establishing the nucleosome core as a powerful source of evolutionary innovation.

H2A.Z sequence encodes its transcriptional function

To build our toolkit of H2A.Z sequences, we first constructed a phylogenetic tree of all H2A sequences from diverse eukaryotes in the EukProt database (Fig. S1, S2A, S3; Table S1; Data File S1; Methods). As previously suggested from analyses of smaller datasets (*40, 57*), we found that H2A.Z forms a monophyletic clade that spans nearly the entirety of known eukaryotic diversity (Fig. S2B, S3A). Almost all sequences in this clade, but very few outside of it, harbor the signature DEELD motif located in a C-terminal region of H2A.Zs referred to as the docking domain (DD) (*40*). A small cluster of sequences, mostly from the unicellular flagellate *Carpediemonas membranifera* (phylum Metamonada), branch just outside the H2A.Z clade, contain a near-perfect, but nevertheless variable, DEELD motif and might represent rare, divergent forms of H2A.Z (Fig. S2B). Other H2A variants either have a more restricted phylogenetic distribution (e.g. macroH2A and H2A.W) or cannot be traced to a single ancestral sequence, appearing paraphyletic (e.g. H2A.X) (Fig. S3). Importantly, we found

H2A.Z orthologs in nearly every eukaryotic genome, even in deep-branching clades like Metamonada that were previously thought to lack H2A.Z (Fig. S2B) (*58*). Thus, the phylogenetic distribution points to H2A.Z being present in the last eukaryotic common ancestor (LECA), and broadly retained throughout the evolution of eukaryotes. These findings support the view that H2A.Z is the only H2A variant whose ancestry can be traced back to the LECA.

We selected nine H2A.Z sequences from across the phylogenetic tree (Fig. S2B and S4A, Data File S2). The sequences were selected to maximize the number of eukaryotic supergroups represented, and their total sequence diversity. Although they span more than a billion years of evolution, these representative H2A.Z proteins were ~65–90% identical in their core domain (excluding highly divergent N- and C-terminal tail regions; Fig. S4B), highlighting their high degree of conservation.

To test whether transcription is impacted by sequence diversity in these nine representative H2A.Z proteins, we synthesized *S. pombe* codon-optimized versions of the nine H2A.Z ortholog coding sequences, and generated yeast lines in which the endogenous gene encoding H2A.Z (*pht1*) was replaced by one of these orthologs (Fig. 1A). We then assessed whether these replacement lines phenocopied the wild-type (WT) gene, resembled a H2A.Z deletion mutant (*pht1*$^{-}$), or differed from both.

We assessed global transcriptional changes in the replacement lines by using strand-specific mRNA-seq. Consistent with prior reports (*2*), *pht1*$^{-}$ cells exhibited modest transcriptional changes, affecting ~10% of detected transcripts (Fig. S4C). By contrast, many H2A.Z replacement lines showed changes in ≥20% of detected transcripts when compared with WT (Fig. S4C). To investigate this further, we examined how similar the transcriptomes of the nine H2A.Z replacement lines were to either WT or *pht1*$^{-}$ using principal component analysis. Replacement lines differed markedly both from WT and *pht1*$^{-}$ (Fig. 1B), suggesting that

H2A.Zs from different species are not simply gradations of total loss of function alleles, but give rise to sequence-specific gain-of-function states.

We next asked whether transcriptional variation in H2A.Z from different species had consequence at the level of phenotype. To do this, we assessed the growth of H2A.Z replacement lines using a plate-based assay in the presence of stress conditions targeting metabolism (rapamycin (*59*)), protein folding (low-dose guanidinium hydrochloride (*60*)), cell wall biogenesis (the anti-fungal drug clotrimazole (*61*)), fermentation (ethanol (*62*)), or either directly (doxorubicin (*63*), formamide(*64*)) or indirectly (caffeine (*65*)) transcription and chromatin (Fig. 1C). Consistent with functional variation at the level of transcriptomes, the growth of all nine replacement lines differed substantially from WT and *pht1*$^{-}$ cells across all tested conditions (Fig. 1C), with most lines growing more slowly than *pht1*$^{-}$ cells. Notably, greater sequence divergence of replacement orthologs from *S. pombe* Pht1 correlated with weaker growth in several stress conditions, with the strongest correlation being for conditions known to be linked directly to transcription (doxorubicin, formamide) or stresses known to require a transcriptional response (ethanol (*62*)) (Fig. 1D, S4D). By contrast, we saw no strong correlation between the level of H2A.Z expression and growth (Fig. 1D, S4D), indicating that these effects were not attributable to differential expression of H2A.Z but likely were related to differences in the primary sequence. Together, these data link primary sequence variation in H2A.Z from different species to variation in its function and confirm a tight linkage between H2A.Z and transcription.

The L2 region of H2A.Z determines its function

Histone proteins are composed of a core histone-fold domain, characterized by a nearly sequence-invariant bundle of three alpha helices ($\alpha 1$, $\alpha 2$, and $\alpha 3$), bounded by mostly unstructured N- and C-terminal tails and separated by two unstructured loop regions, L1 and

L2 (Fig. 2A) (*1*). To pinpoint which sequence domains of H2A.Z are responsible for changes in transcriptional output across the set of orthologs tested above, we first examined sequence conservation across the full H2A.Z dataset (Fig. S3), calculating sequence identity along the length of H2A.Z using a 3-residue sliding window average (Fig. 2A). As expected, we observed high conservation of the core alpha-helical bundle and relatively lower conservation of the tails and the L1 (Fig. 2A). Surprisingly, however, we found that parts of the L2 region, specifically around the core motif KDLKVK, were as highly conserved as the core α-helices (Fig. 2A). This was especially striking when compared with the H2A variant H2A.X, which was notably less conserved in this region.

The unexpectedly high sequence conservation of the L2 region suggested it might play a role in determining the global changes in transcription we observed in the replacement *S. pombe* lines. To test this, we engineered a series of chimeric histone genes whereby the N, C, L1, L2, or docking domain (DD) (*40*) regions of *S. pombe pht1* were substituted into histone H2A (*hta1*) (Fig. 2B, S5A). We expressed each of these chimeric genes by knocking them into the *pht1* locus and assayed their ability to complement the *pht1*$^{-}$ null allele by measuring the growth of the chimeric lines under selected stress conditions based on prior experiments with the representative full-length H2A.Z replacement lines (Fig. 1C). The chimeric histone genes fell into two groups across all phenotypes tested: those that partially complemented the *pht1*$^{-}$ mutant (those encoding H2A/Z^{DD}, H2A/Z^{N}, and H2A/Z^{L1}) and those that fully complemented the *pht1*$^{-}$ mutant in all tested conditions (those encoding H2A/Z^{L2} and H2A/Z^{C}), making them indistinguishable from WT (Fig. 2C). These data indicate that both the L2 region and the C-terminal tail play key roles in determining H2A.Z function.

To investigate whether sequence variation in the L2 region of H2A.Z alone is sufficient to explain the transcriptional differences in our full-length H2A.Z replacement lines, we looked for a correlation between transcriptional divergence and sequence divergence from WT *S.*

pombe in any of the structurally defined primary sequence regions of H2A.Z. Although the short L2 affords limited resolution, we observed a strong correlation between L2 sequence divergence and transcriptional divergence (ρ = 0.63, $p < 10^{-3}$, Fig. 2D). By contrast, the sequence divergence of the C-terminal tail, which fully complemented the *pht1*$^{-}$ mutant in growth assays, did not correlate with transcriptional divergence (ρ = 0.056, p = 0.8, Fig. 2D). There was also a correlation between the sequence divergence of the N-terminal tail and transcriptional divergence (ρ = 0.5, $p < 10^{-2}$; Fig. S5B). This was likely due to co-variation with the L2 sequence, as there was a tight correlation between sequence variation in these two regions (ρ = 0.77, $p < 10^{-5}$, Fig. S5B). Together, these data show that the L2 region of H2A.Z is an important contributor to its function.

Transcription elongation factor Spt6 binds H2A.Z L2

The L2 region of H2A.Z is largely solvent accessible (*55*), particularly during transcription when the DNA is unwrapped from the nucleosome (*19, 20*). We therefore hypothesized that L2 might impact transcription by physically interacting with proteins involved in transcription, rather than by influencing histone–DNA contacts, as is the case for the L2 region of histone H3 (*1, 66*). To identify proteins in the *S. pombe* nuclear proteome that might interact with the L2 region of Pht1, we used AlphaFold2 Multimer (*67*) to predict physical interactions. We first performed a virtual interaction screen of WT Pht1 L2 against the *S. pombe* annotated proteome (*68, 69*) (Fig. 3A), identifying 247 proteins with an interface predicted template modelling (IPTM) score threshold of >0.5 (Table S2). To further refine potential interactors, we reasoned that L2 interactors responsible for driving functional variation in transcriptional output must vary in their interaction in accordance with their L2 primary sequence. We thus selected 17 unique single-residue variants of the L2 core motif from our database of H2A.Z sequences (Fig. S3, Data File S1), performing an AlphaFold2 virtual

interaction screen with correspondingly mutated Pht1 L2 sequences against a subset of 500 nuclear proteins based on their IPTM score with WT Pht1 L2 (Fig. 3B, S6A, Table S3). To assess differences in predicted binding across Pht1 mutants, we then calculated the variance in IPTM score for each candidate interactor protein (Fig. 3C). In most cases, the variance was very low, suggesting that the candidate interacting proteins do not respond to variation in L2 (Fig. 3C, S6A, Table S3). Among the candidate interactors with the highest variance, the transcription elongation factor Spt6 (Fig. 3D–E, Table S3) was the only component of RNAPII (*70, 71*). It interacts with nucleosomes (*72-74*), had highly robust folding models across multiple simulations (Fig. S6B), and closely matched experimentally derived structures (Fig. S6C).

We first experimentally validated the AlphaFold2-predicted interaction between Pht1 and Spt6, using glutathione-S-transferase (GST) as negative control. H2A was previously reported to bind to Spt6 and was used as positive control (*75*). We expressed and purified recombinant *S. pombe* GST-tagged Pht1 and Hta1, and His-tagged Spt6 in *Escherichia coli* (Fig. 3F) and assessed binding *in vitro* to nickel resin (binds the His tag). Pht1 bound Spt6 (Fig. 3G), confirming the Pht1–Spt6 interaction. Further, we noted a qualitative difference in the amount of Hta1 and Pht1 bound to Spt6, suggesting that the two histone proteins might interact differently with Spt6. To verify this finding, we repeated Spt6-Pht1 pulldown experiments under more stringent binding conditions, where the amount of bound Pht1 was considerably less than that of Hta1 (Fig. S6D), reinforcing the suggestion Hta1 and Pht1 may interact with Spt6 differently. Together, these data establish a physical interaction between Spt6 and Pht1.

To test whether the L2 region was sufficient to mediate the Pht1–Spt6 interaction, we synthesized a biotinylated peptide corresponding to this region of *S. pombe* Pht1 and assayed its binding to His-tagged Spt6 (Fig. 3H). As a negative control, we used a peptide corresponding to the docking domain (DD) of *S. pombe* Pht1, which is adjacent to the L2 region

and important for the interaction of Pht1 with H3 and H4 (*55*) and Swr1-C (*76*). Whereas the DD peptide did not bind Spt6, the peptides corresponding to the L2 region of Hta1 or Pht1 bound Spt6 (Fig. 3I), consistent with the results of our experiments with the full-length proteins. These data establish that the L2 region alone is sufficient to explain the interaction of Pht1 with Spt6.

To test whether sequence variation in the L2 region of H2A.Z alters its binding to Spt6, we synthesized L2 peptides from five of the original set of representative H2A.Z sequences and assayed their binding to *S. pombe* Spt6, as above. Consistent with the sequence of L2 determining its interaction with Spt6, we observed variable amount of each peptide bound to Spt6 across species (Fig. 3I), with *A. godoyi*, *T. globosa*, and *P. marinus* H2A.Z L2 peptides having the greatest qualitative difference relative to *S. pombe* Pht1 L2. Although there were clear differences between peptides derived from different species, we did not observe a strong association between Spt6-binding and sequence similarity (Fig. 3I), suggesting that specific sequence features of the peptides, rather than overall sequence divergence, must govern their association with Spt6. Indeed, when we compared the sequence of the three species' L2 peptides that had discernable differences relative to Pht1 to those that did not, we noted that all of them had substitutions within the L2 core motif (Fig. S6F), especially at position L83.

To investigate the role of sequence variation in the core motif around L83 in determining the binding of Pht1 L2 to Spt6, we further examined AlphaFold2 models. Pht1 was predicted to bind to the Spt6 death-like domain (DLD) (Fig. 3D). This domain, whose function has remained uncharacterized, is the most highly conserved feature of the Spt6 surface (*77*), suggesting an important functional role. To assess this further, we gathered Spt6 sequences from the same EukProt database we used previously to identify H2A.Zs, calculating sequence conservation of the DLD across known eukaryotic diversity (Fig. 3J). In the DLD large numbers of hydrophobic residues were not only highly conserved in absolute terms (Fig.

3J), but also relative to other hydrophobic residues in Spt6 ($p < 0.005$, Wilcoxon test; Fig. S6F). These residues were organized into a hydrophobic groove, many of which contacted the Pht1 L2 core motif in AlphaFold2 models (Fig. 3E, S6E). Further, these contacts were predicted to vary according to the H2A.Z L2 sequence, both when examining models for L2s from different species (Fig. S6E), as well as Pht1 mutants to the L83 position we identified earlier (Fig. 3E), suggesting a rationale for why sequence variation at this Pht1 position might affect binding.

To test the contribution of Pht1 L2 sequence variation, particularly at position L83, to the Pht1–Spt6 interaction in the context of the DLD hydrophobic groove, we synthesized three mutated Pht1 L2 peptides with either more hydrophobic (L83F) or more hydrophilic (L83M, L83N) substitutions and tested their binding to Spt6. The more hydrophilic L83N substitution, which corresponds to Hta1 residue N75, does not recapitulate the binding of full-length Hta1 (Fig. 3D), reinforcing our earlier conclusion that numerous other primary sequences differences in Hta1 (Fig. S5A) must contribute to its binding. Whereas L83F mutant L2 peptide bound Spt6 similarly to WT Pht1 peptide, Spt6 bound both more hydrophilic substitutions (L83M/N) to a much lesser extent than WT Pht1 L2 (Fig. 3I). These relative binding abilities corresponded to their hydrophobicities, and are thus consistent with AlphaFold2 predictions that the L2 region of Pht1 interacts with the hydrophobic groove of the DLD of Spt6 (Fig. 3D). Together, these findings suggest that the hydrophobic L2 region of H2A.Z determines its binding to the DLD of Spt6.

H2A.Z L2 encodes transcription elongation states

Since Spt6 is a transcription elongation factor (*70, 71*), we hypothesized that sequence variation in the L2 region of H2A.Z might be sufficient to affect transcription elongation, thereby explaining species-specific differences in reported H2A.Z function. To test this, we first developed a growth-based reporter assay using the uracil biosynthesis gene *ura4*. In this

assay, the amount of transcription of *ura4* is directly related to the amount of growth in media lacking uracil (Fig. 4A). Assaying a growth-based reporter has the advantage that it can be used to test multiple conditions and it is not easily confounded by global transcriptional changes. We first confirmed that our reporter gene accumulates Pht1 by ChIP-qPCR (Fig. S7A). We then introduced point mutations corresponding to those found in the L2 core motif of our nine representative H2A.Z orthologs (Fig. S4A, S6F) into the endogenous *pht1* (H2A.Z) gene in *S. pombe* (Fig. 4B). Whereas *pht1*$^-$ cells gave rise to a modest increase in reporter expression (Fig. 4C), supporting a repressive function of H2A.Z in *S. pombe* as in some other model systems (*11, 48, 50, 53*), half of L2 mutants had a significant increase (by ~1.3–1.5 fold for L83F, L83N and L83Q) or decrease (by ~2 fold for L83M) in reporter expression. These results were not confounded by general fitness defects in the mutant strains (Fig. S7B), and all the mutants were epistatic to Swc2, an essential subunit of the Swr1-C chromatin remodeling complex, which incorporates H2A.Z into nucleosomes (*78, 79*) (Fig. S7C). Together these results establish that reporter expression differences are driven by Pht1 presence at the reporter gene and suggest that L2 sequence variation is sufficient to impact transcription.

We next tested whether differences in reporter expression across Pht1 mutant lines were specifically attributable to transcription elongation through two orthogonal approaches. First, focusing on the four Pht1 mutant lines with the strongest differences in reporter expression relative to WT (L83M, L83N, K86Q, and L83F), we made use of the drug mycophenolic acid (MPA) to specifically interfere with RNAPII elongation by lowering cellular GTP levels (*80*). In this experiment, if a mutant has increased or decreased RNAPII elongation relative to WT, then it should be resistant or sensitive to MPA, respectively. Treatment of WT cells with MPA caused a ~40% reduction in growth relative to the control (Fig. 4D). By contrast, *pht1*$^-$ and the Pht1 mutants L83N, K86Q, and L83F were more resistant to MPA than the WT (~25–35% reduction in growth; Fig. 4D), consistent with increased transcription elongation in these lines.

By contrast, the single mutant that had decreased reporter expression (L83M) was also the most sensitive to MPA treatment (~60% reduction, $p_{adj} < 10^{6}$, Fig. 4D), confirming its reduction in transcription elongation activity. Together, these data establish that Pht1 L2 mutants are sufficient to impact elongating RNAPII.

Finally, as an orthogonal measure to our reporter-based assays, we probed the impact of Pht1 L2 mutants on transcription elongation by measuring nascent transcription using precision run-on sequencing (qPRO-seq) (*81*, *82*). Sequencing nascent transcripts allowed us to position RNAPII in the transcriptional unit and to distinguish between promoter-proximal paused RNAPII and gene body-located RNAPII participating in active elongation across all expressed genes, albeit without accurately quantifying specific parameters of RNAPII activity. Consistent with results of our reporter assays, differential expression analysis of gene body-located qPRO-seq counts in the WT and *pht1*$^{-}$ lines showed a modest increase in active elongation by RNAPII at ~100 genes in *pht1*$^{-}$ cells (Fig. 4E). Similar analyses of the Pht1 point mutant lines identified three classes of mutant. In one class, comprising half of the tested mutants (L83M, L83G, A80S, and V85A), we observed essentially no change in RNAPII elongation, when compared with WT at the level of differentially transcribed genes (Fig. 4E). Notably, this class included the single super-repressive L2 mutant (L83M), suggesting it is likely linked to Spt6's function as part of transcription termination (*70*). The second class comprised a single mutant (K84H), which had a modest increase in RNAPII engaging in elongation, very similar to that of *pht1*$^{-}$ (Fig. 4E). The third class comprised three mutants (L83N, K86Q, and L83F) in which we observed a robust increase in RNAPII engaging in elongation at ~50% more genes than in *pht1*$^{-}$ (Fig. 4E), consistent with their impact on reporter gene transcription. Together, these data establish that sequence variation in the core L2 region of H2A.Z is sufficient to transform its transcriptional impact on the genome.

For much of the past 50 years, investigations into variation in the core histone fold have focused on its interaction with DNA and the stability of the nucleosome (*26, 41, 42, 83-87*). Our findings, by contrast, highlight the importance of the histone core domain for genomic regulation through protein–protein interactions (*66, 88, 89*). Given the extreme depth of conservation of not only H2A.Z and Spt6, but also the entire chromatin and transcription apparatus across eukaryotic evolution (*90*), it seems likely that the nucleosome core of transcriptional regulation originated with eukaryogenesis and diversified with histone variants. That we uncover sequence variation in just seven residues of the L2 core motif of H2A.Z can give rise to enormous phenotypic complexity at the level of transcriptional output and beyond, it is likely that many protein–protein interactions with the core domain of other histone variants remain to be discovered. These will not only broaden our understanding of the importance of the evolution of the nucleosome, but also, given the prevalence of transcription elongation defects in cancers (*91, 92*), provide a new targets to address human disorders.

References and Notes

1. K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent, T. J. Richmond, Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-260 (1997).

2. M. D. Meneghini, M. Wu, H. D. Madhani, Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. *Cell* **112**, 725-736 (2003).

3. C. D. Allis *et al.*, hv1 is an evolutionarily conserved H2A variant that is preferentially associated with active genes. *Journal of Biological Chemistry* **261**, 1941-1948 (1986).

4. R. Yelagandula *et al.*, The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in Arabidopsis. *Cell* **158**, 98-109 (2014).

5. C. Costanzi, J. R. Pehrson, Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. *Nature* **393**, 599-601 (1998).

6. B. D. Giaimo, F. Ferrante, A. Herchenrother, S. B. Hake, T. Borggrefe, The histone variant H2A.Z in gene regulation. *Epigenetics Chromatin* **12**, 37 (2019).

7. V. Turinetto, C. Giachino, Multiple facets of histone variant H2AX: a DNA double-strand-break marker with several biological functions. *Nucleic Acids Res* **43**, 2489-2498 (2015).

8. M. Perino, G. J. Veenstra, Chromatin Control of Developmental Dynamics and Plasticity. *Dev Cell* **38**, 610-620 (2016).

9. R. P. Ghosh, B. J. Meyer, Spatial Organization of Chromatin: Emergence of Chromatin Structure During Development. *Annu Rev Cell Dev Biol* **37**, 199-232 (2021).

10. A. M. Valencia, C. Kadoch, Chromatin regulatory mechanisms and therapeutic opportunities in cancer. *Nat Cell Biol* **21**, 152-161 (2019).

11. D. G. Berta *et al.*, Deficient H2A.Z deposition is associated with genesis of uterine leiomyoma. *Nature*, (2021).

12. R. S. Greenberg, H. K. Long, T. Swigut, J. Wysocka, Single Amino Acid Change Underlies Distinct Roles of H2A.Z Subtypes in Human Syndrome. *Cell* **178**, 1421-1436 e1424 (2019).

13. A. Weiner *et al.*, High-resolution chromatin dynamics during a yeast stress response. *Mol Cell* **58**, 371-386 (2015).

14. D. Tremethick, K. Zucker, A. Worcel, The transcription complex of the 5 S RNA gene, but not transcription factor IIIA alone, prevents nucleosomal repression of transcription. *J Biol Chem* **265**, 5014-5023 (1990).

15. G. Almouzni, M. Mechali, A. P. Wolffe, Transcription complex disruption caused by a transition in chromatin structure. *Mol Cell Biol* **11**, 655-665 (1991).

16. J. N. Hirschhorn, S. A. Brown, C. D. Clark, F. Winston, Evidence that SNF2/SWI2 and SNF5 activate transcription in yeast by altering chromatin structure. *Genes Dev* **6**, 2288-2298 (1992).

17. O. I. Kulaeva, F. K. Hsieh, H. W. Chang, D. S. Luse, V. M. Studitsky, Mechanism of transcription through a nucleosome by RNA polymerase II. *Biochim Biophys Acta* **1829**, 76-83 (2013).

18. B. R. Cairns, The logic of chromatin architecture and remodelling at promoters. *Nature* **461**, 193-198 (2009).

19. H. Ehara, T. Kujirai, M. Shirouzu, H. Kurumizaka, S. I. Sekine, Structural basis of nucleosome disassembly and reassembly by RNAPII elongation complex with FACT. *Science* **377**, eabp9466 (2022).

20. T. Kujirai *et al.*, Structural basis of the nucleosome transition during RNA polymerase II passage. *Science* **362**, 595-598 (2018).

21. B. Wasylyk, G. Thevenin, P. Oudet, P. Chambon, Transcription of in vitro assembled chromatin by Escherichia coli RNA polymerase. *J Mol Biol* **128**, 411-440 (1979).

22. J. A. Knezetic, D. S. Luse, The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. *Cell* **45**, 95-104 (1986).

23. Y. Lorch, J. W. LaPointe, R. D. Kornberg, Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. *Cell* **49**, 203-210 (1987).

24. P. B. Talbert, M. P. Meers, S. Henikoff, Old cogs, new tricks: the evolution of gene expression in a chromatin context. *Nat Rev Genet* **20**, 283-297 (2019).

25. S. Henikoff, M. M. Smith, Histone variants and epigenetics. *Cold Spring Harb Perspect Biol* **7**, a019364 (2015).

26. C. M. Weber, S. Ramachandran, S. Henikoff, Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. *Mol Cell* **53**, 819-830 (2014).

27. P. B. Talbert, S. Henikoff, Histone variants--ancient wrap artists of the epigenome. *Nat Rev Mol Cell Biol* **11**, 264-275 (2010).

28. S. B. Hake, C. D. Allis, Histone H3 variants and their potential role in indexing mammalian genomes: the "H3 barcode hypothesis". *Proc Natl Acad Sci U S A* **103**, 6428-6435 (2006).

29. D. Jiang, F. Berger, Variation is important: Warranting chromatin function and dynamics by histone variants. *Curr Opin Plant Biol* **75**, 102408 (2023).

30. D. Jiang *et al.*, The evolution and functional divergence of the histone H2B family in plants. *PLoS Genet* **16**, e1008964 (2020).

31. A. Herchenrother, T. M. Wunderlich, J. Lan, S. B. Hake, Spotlight on histone H2A variants: From B to X to Z. *Semin Cell Dev Biol* **135**, 3-12 (2023).

32. H. Zhang, D. N. Roberts, B. R. Cairns, Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. *Cell* **123**, 219-231 (2005).

33. B. Guillemette *et al.*, Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning. *PLoS Biol* **3**, e384 (2005).

34. R. M. Raisner *et al.*, Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. *Cell* **123**, 233-248 (2005).

35. A. Barski *et al.*, High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823-837 (2007).

36. D. Zilberman, D. Coleman-Derr, T. Ballinger, S. Henikoff, Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. *Nature* **456**, 125-129 (2008).

37. T. N. Mavrich *et al.*, Nucleosome organization in the Drosophila genome. *Nature* **453**, 358-362 (2008).

38. C. D. Allis, C. V. Glover, J. K. Bowen, M. A. Gorovsky, Histone variants specific to the transcriptionally active, amitotically dividing macronucleus of the unicellular eucaryote, Tetrahymena thermophila. *Cell* **20**, 609-617 (1980).

39. M. H. West, W. M. Bonner, Histone 2A, a heteromorphous family of eight protein species. *Biochemistry* **19**, 3238-3245 (1980).

40. H. S. Malik, S. Henikoff, Phylogenomics of the nucleosome. *Nat Struct Biol* **10**, 882-891 (2003).

41. A. Osakabe *et al.*, Histone H2A variants confer specific properties to nucleosomes and impact on chromatin accessibility. *Nucleic Acids Res* **46**, 7675-7685 (2018).

42. N. Horikoshi, T. Kujirai, K. Sato, H. Kimura, H. Kurumizaka, Structure-based design of an H2A.Z.1 mutant stabilizing a nucleosome in vitro and in vivo. *Biochem Biophys Res Commun* **515**, 719-724 (2019).

43. Z. Wen, L. Zhang, H. Ruan, G. Li, Histone variant H2A.Z regulates nucleosome unwrapping and CTCF binding in mouse ES cells. *Nucleic Acids Res* **48**, 5939-5952 (2020).

44. M. Adam, F. Robert, M. Larochelle, L. Gaudreau, H2A.Z is required for global chromatin integrity and for recruitment of RNA polymerase II under specific conditions. *Mol Cell Biol* **21**, 6270-6279 (2001).

45. M. S. Santisteban, T. Kalashnikova, M. M. Smith, Histone H2A.Z Regulates Transcription and Is Partially Redundant with Nucleosome Remodeling Complexes. *Cell* **103**, 411-422 (2000).

46. M. S. Santisteban, M. Hang, M. M. Smith, Histone variant H2A.Z and RNA polymerase II transcription elongation. *Mol Cell Biol* **31**, 1848-1860 (2011).

47. L. A. Stargell *et al.*, Temporal and spatial association of histone H2A variant hv1 with transcriptionally competent chromatin during nuclear development in Tetrahymena thermophila. *Genes Dev* **7**, 2641-2651 (1993).

48. C. Mylonas, C. Lee, A. L. Auld, Cisse, II, L. A. Boyer, A dual role for H2A.Z.1 in modulating the dynamics of RNA polymerase II initiation and elongation. *Nat Struct Mol Biol* **28**, 435-442 (2021).

49. C. D. Allis, Y. S. Ziegler, M. A. Gorovsky, J. B. Olmsted, A conserved histone variant enriched in nucleoli of mammalian cells. *Cell* **31**, 131-136 (1982).

50. D. Coleman-Derr, D. Zilberman, Deposition of histone variant H2A.Z within gene bodies regulates responsive genes. *PLoS Genet* **8**, e1002988 (2012).

51. B. Carter *et al.*, The Chromatin Remodelers PKL and PIE1 Act in an Epigenetic Pathway That Determines H3K27me3 Homeostasis in Arabidopsis. *Plant Cell* **30**, 1337-1352 (2018).

52. B. Jamge *et al.*, Histone variants shape chromatin states in Arabidopsis. *Elife* **12**, (2023).

53. A. Thakar, P. Gupta, W. T. McAllister, J. Zlatanova, Histone variant H2A.Z inhibits transcription in reconstituted nucleosomes. *Biochemistry* **49**, 4018-4026 (2010).

54. K. Ahmad, S. Henikoff, The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol Cell* **9**, 1191-1200 (2002).

55. R. K. Suto, M. J. Clarkson, D. J. Tremethick, K. Luger, Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nat Struct Biol* **7**, 1121-1124 (2000).

56. H. Tachiwana *et al.*, Structures of human nucleosomes containing major histone H3 variants. *Acta Crystallogr D Biol Crystallogr* **67**, 578-583 (2011).

57. P. B. Talbert *et al.*, A unified phylogeny-based nomenclature for histone variants. *Epigenetics Chromatin* **5**, 7 (2012).

58. M. C. Dalmasso, W. J. Sullivan, Jr., S. O. Angel, Canonical and variant histones of protozoan parasites. *Front Biosci (Landmark Ed)* **16**, 2086-2105 (2011).

59. J. Heitman, N. R. Movva, M. N. Hall, Targets for cell cycle arrest by the immunosuppressant rapamycin in yeast. *Science* **253**, 905-909 (1991).

60. Y. O. Chernoff, S. L. Lindquist, B. Ono, S. G. Inge-Vechtomov, S. W. Liebman, Role of the chaperone protein Hsp104 in propagation of the yeast prion-like factor [psi+]. *Science* **268**, 880-884 (1995).

61. P. D. Crowley, H. C. Gallagher, Clotrimazole as a pharmaceutical: past, present and future. *J Appl Microbiol* **117**, 611-617 (2014).

62. W. Samakkarn, K. Ratanakhanokchai, N. Soontorngun, Reprogramming of the Ethanol Stress Response in Saccharomyces cerevisiae by the Transcription Factor Znf1 and Its Effect on the Biosynthesis of Glycerol and Ethanol. *Appl Environ Microbiol* **87**, e0058821 (2021).

63. F. Yang, C. J. Kemp, S. Henikoff, Doxorubicin enhances nucleosome turnover around promoters. *Curr Biol* **23**, 782-787 (2013).

64. R. Hoyos-Manchado *et al.*, RNA metabolism is the primary target of formamide in vivo. *Scientific reports* **7**, 15895 (2017).

65. S. Torres-Garcia *et al.*, Epigenetic gene silencing by heterochromatin primes fungal resistance. *Nature* **585**, 453-458 (2020).

66. K. M. Stevens *et al.*, Histone variants in archaea and the evolution of combinatorial chromatin complexity. *Proceedings of the National Academy of Sciences* **117**, 33384-33395 (2020).

67. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).

68. M. A. Harris *et al.*, Fission stories: using PomBase to understand Schizosaccharomyces pombe biology. *Genetics* **220**, (2022).

69. C. The Gene Ontology, The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res* **47**, D330-D338 (2019).

70. A. Narain *et al.*, Targeted protein degradation reveals a direct role of SPT6 in RNAPII elongation and termination. *Mol Cell* **81**, 3110-3127 e3114 (2021).

71. K. Zumer *et al.*, Two distinct mechanisms of RNA polymerase II elongation stimulation in vivo. *Mol Cell* **81**, 3096-3109 e3098 (2021).

72. A. Bortvin, F. Winston, Evidence that Spt6p controls chromatin structure by a direct interaction with histones. *Science* **272**, 1473-1476 (1996).

73. S. M. McDonald, D. Close, H. Xin, T. Formosa, C. P. Hill, Structure and biological importance of the Spn1-Spt6 interaction, and its regulatory role in nucleosome binding. *Mol Cell* **40**, 725-735 (2010).

74. M. W. Adkins, J. K. Tyler, Transcriptional activators are dispensable for transcription in the absence of Spt6-mediated chromatin reassembly of promoter regions. *Molecular cell* **21**, 405-416 (2006).

75. C. Jeronimo, S. Watanabe, C. D. Kaplan, C. L. Peterson, F. Robert, The Histone Chaperones FACT and Spt6 Restrict H2A.Z from Intragenic Locations. *Mol Cell* **58**, 1113-1123 (2015).

76. J. Hong *et al.*, The catalytic subunit of the SWR1 remodeler is a histone chaperone for the H2A.Z-H2B dimer. *Mol Cell* **53**, 498-505 (2014).

77. D. Close *et al.*, Crystal structures of the S. cerevisiae Spt6 core and C-terminal tandem SH2 domain. *J Mol Biol* **408**, 697-713 (2011).

78. W. H. Wu *et al.*, Swc2 is a widely conserved H2AZ-binding module essential for ATP-dependent histone exchange. *Nat Struct Mol Biol* **12**, 1064-1071 (2005).

79. E. Luk *et al.*, Stepwise histone replacement by SWR1 requires dual activation with histone H2A.Z and canonical nucleosome. *Cell* **143**, 725-736 (2010).

80. R. J. Shaw, J. L. Wilson, K. T. Smith, D. Reines, Regulation of an IMP dehydrogenase gene and its overexpression in drug-sensitive transcription elongation mutants of yeast. *J Biol Chem* **276**, 32905-32916 (2001).

81. G. T. Booth, I. X. Wang, V. G. Cheung, J. T. Lis, Divergence of a conserved elongation factor and transcription regulation in budding and fission yeast. *Genome Res* **26**, 799-811 (2016).

82. J. Judd *et al.*, A rapid, sensitive, scalable method for Precision Run-On sequencing (PRO-seq). *bioRxiv*, (2020).

83. C. Bonisch, S. B. Hake, Histone H2A variants in nucleosomes and chromatin: more or less stable? *Nucleic Acids Res* **40**, 10719-10741 (2012).

84. S. Chakravarthy, Y. Bao, V. A. Roberts, D. Tremethick, K. Luger, Structural characterization of histone H2A variants. *Cold Spring Harb Symp Quant Biol* **69**, 227-234 (2004).

85. Y. Arimura *et al.*, Structural basis of a nucleosome containing histone H2A.B/H2A.Bbd that transiently associates with reorganized chromatin. *Scientific reports* **3**, 3510 (2013).

86. N. Horikoshi, Y. Arimura, H. Taguchi, H. Kurumizaka, Crystal structures of heterotypic nucleosomes containing histones H2A.Z and H2A. *Open Biol* **6**, (2016).

87. H. Kohestani, J. Wereszczynski, Effects of H2A.B incorporation on nucleosome structures and dynamics. *Biophys J* **120**, 1498-1509 (2021).

88. A. K. Michael *et al.*, Cooperation between bHLH transcription factors and histones for DNA access. *Nature* **619**, 385-393 (2023).

89. S. Sanulli *et al.*, HP1 reshapes nucleosome core to promote phase separation of heterochromatin. *Nature* **575**, 390-394 (2019).

90. X. Grau-Bove *et al.*, A phylogenetic and proteomic reconstruction of eukaryotic chromatin evolution. *Nat Ecol Evol* **6**, 1007-1023 (2022).

91. V. Modur *et al.*, Defective transcription elongation in a subset of cancers confers immunotherapy resistance. *Nat Commun* **9**, 4410 (2018).

92. T. E. Miller *et al.*, Transcription elongation factors represent in vivo cancer dependencies in glioblastoma. *Nature* **547**, 355-359 (2017).

Acknowledgements: We thank the entire Berger group and Antoine Hocher, for their insightful, considerate, and helpful discussions, as well as Carol Featherstone from Life Science Editors and Matt Watson for editorial assistance during the preparation of the manuscript. We especially thank Zdravko Lorković, Elin Axelsson-Ekker, and Chung Hyun Cho for their invaluable assistance preparing antibodies, biochemical assays, and bioinformatic analyses. We also thank John T. Lis for advice with qPRO-seq. We thank the Molecular Biology Service and Media Kitchen for a constant supply of plates, basic reagents, and cloning and sequencing services, as well as the Peptide Synthesis Service. Additionally, we thank the Vienna BioCenter Core Facilities, in particular the Next Generation Sequencing facility for their advice and swift handling of all our requests and the Protein Technologies facility for their expertise, helpful discussions, and recombinant protein production.

Funding:

Gregor Mendel Institute, Austrian Academy of Sciences Core Funding (FB)

Österreichischer Wissenschaftsfonds FWF TAI304 (FB)

Österreichischer Wissenschaftsfonds FWF ESP213 (ZHH)

EMBO Postdoctoral Fellowship ALTF169-2020 (ZHH)

United Kingdom Research and Innovation's Medical Research Council MC_A658_5TY40 (TW)

Author contributions: Conceptualization: ZHH, FB, and TW; Methodology: ZHH, KS, and TW; Validation: ZHH and KS; Formal Analysis: ZHH, KS, and TW; Investigation: ZHH, KS, and TW; Data Curation: ZHH, KS, and TW; Writing – original draft: ZHH, TW, and FB; Visualization: ZHH, KS, and TW; Supervision: FB, TW; Project Administration: ZHH; Funding Acquisition: ZHH, FB, and TW.

Competing interests: Z.H.H. and F.B. are inventors on a patent submitted by the Gregor Mendel Institute at the Austrian Patent Office (A50392/2024) pertaining to the modification of transcription and fitness by engineering H2A.Z L2.

Data and materials availability: PRO-seq pipeline and scripts used for data analysis are available as a container on Github, and all other scripts used in analysis are available upon request to the lead contact. NGS data (PRO-seq, mRNA-seq) are deposited in GEO under accession number GSE265749 and GSE265775, respectively. For material request (e.g. strains), please contact the lead (zachary.harvey@gmi.oeaw.ac.at).

Supplementary Materials

Materials and Methods

Figs. S1 to S7

Supplementary Data Tables S1 to S4

Supplementary Data File S1 to S3

References ()

FIGURES AND FIGURE LEGENDS

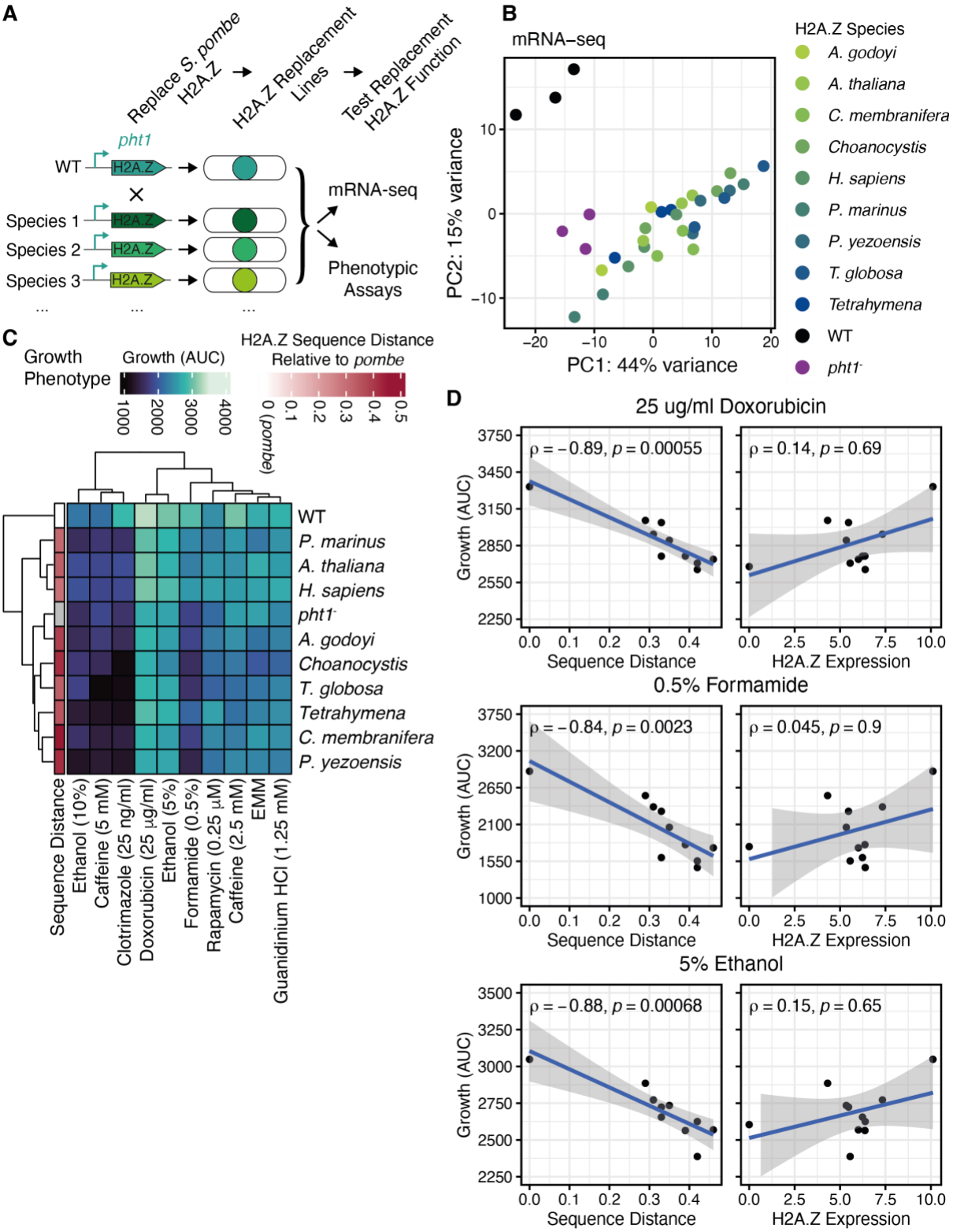
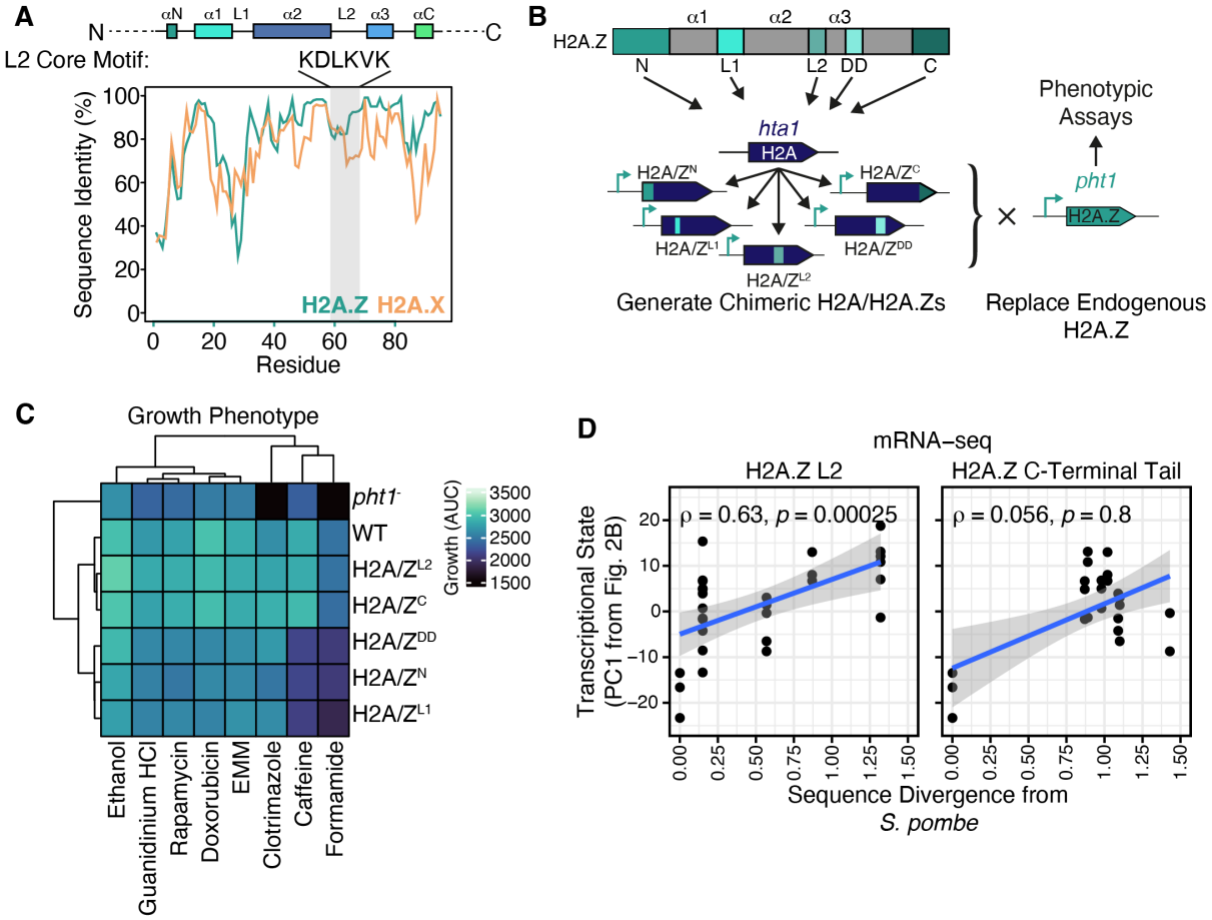

A
Replace S. pombe H2A.Z
H2A.Z Replacement Lines
Test Replacement H2A.Z Function
pht1
WT
H2A.Z
Species 1
Species 2
Species 3
mRNA-seq
Phenotypic Assays
B
mRNA–seq
PC2: 15% variance
PC1: 44% variance
H2A.Z Species
A. godoyi
A. thaliana
C. membranifera
Choanocystis
H. sapiens
P. marinus
P. yezoensis
T. globosa
Tetrahymena
WT
pht1-
C
Growth Phenotype
Growth (AUC)
H2A.Z Sequence Distance Relative to pombe
Sequence Distance
Ethanol (10%)
Caffeine (5 mM)
Clotrimazole (25 ng/ml)
Doxorubicin (25 μg/ml)
Ethanol (5%)
Formamide (0.5%)
Rapamycin (0.25 μM)
Caffeine (2.5 mM)
EMM
Guanidinium HCl (1.25 mM)
D
25 ug/ml Doxorubicin
ρ = −0.89, p = 0.00055
ρ = 0.14, p = 0.69
0.5% Formamide
ρ = −0.84, p = 0.0023
ρ = 0.045, p = 0.9
5% Ethanol
ρ = −0.88, p = 0.00068
ρ = 0.15, p = 0.65
Growth (AUC)
Sequence Distance
H2A.Z Expression

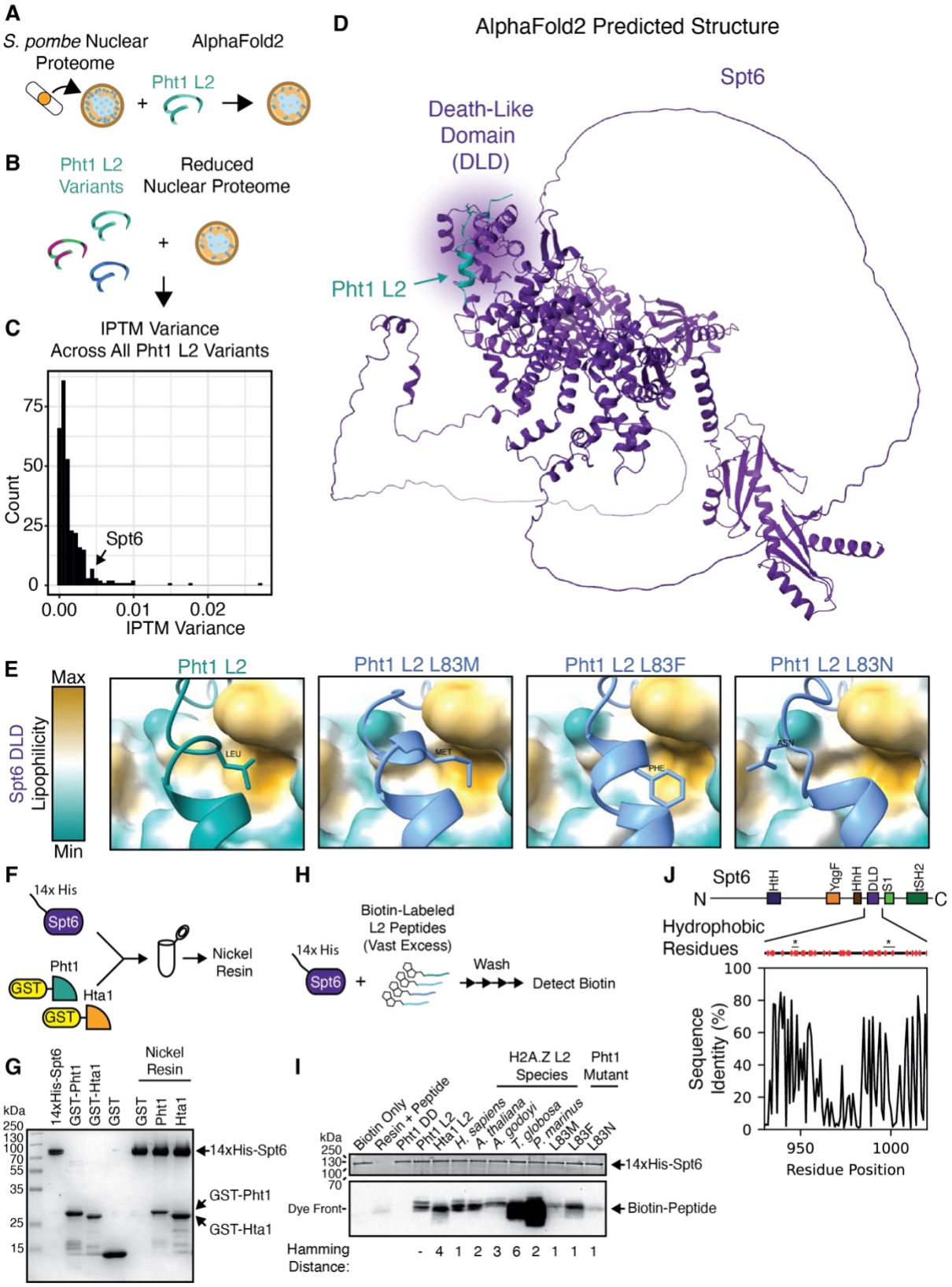
Figure 1. Phylogenetically diverse H2A.Z sequences drive distinct gain-of-function transcriptional states. A, Experimental scheme describing the replacement of endogenous *S. pombe pht1* by representative H2A.Zs taken from the H2A phylogeny (Fig. S2B, S3A). **B**, Principal component analysis (PCA) for the top 500 variable genes from mRNA-seq data generated for H2A.Z replacement lines. **C**, Growth phenotypes for H2A.Z replacement lines in various stress conditions, annotated with the sequence distance of each H2A.Z from that of *S. pombe* Pht1 (which has a sequence distance of zero). Plotted values are the mean of four biological replicates. Growth is quantified as the area under the curve (AUC) from respective growth curves. H2A.Zs are annotated with their sequence distance relative to *S. pombe* Pht1, with zero being Pht1 and larger numbers indicating greater divergence. **D**, Example Spearman's correlations between the growth from (C) with either sequence distance (as in C) of each H2A.Z relative to WT *S. pombe*, or the expression of the respective H2A.Z (from mRNA-seq).


A
N
αN
α1
L1
α2
L2
α3
αC
C
L2 Core Motif:
KDLKVK
Sequence Identity (%)
100
80
60
40
20
0
H2A.Z
H2A.X
0
20
40
60
80
Residue
B
H2A.Z
α1
α2
α3
N
L1
L2
DD
C
hta1
H2A
H2A/Z^N
H2A/Z^C
H2A/Z^L1
H2A/Z^DD
H2A/Z^L2
Generate Chimeric H2A/H2A.Zs
Phenotypic Assays
pht1
H2A.Z
Replace Endogenous H2A.Z
C
Growth Phenotype
pht1-
WT
H2A/Z^L2
H2A/Z^C
H2A/Z^DD
H2A/Z^N
H2A/Z^L1
Growth (AUC)
3500
3000
2500
2000
1500
Ethanol
Guanidinium HCl
Rapamycin
Doxorubicin
EMM
Clotrimazole
Caffeine
Formamide
D
mRNA–seq
H2A.Z L2
H2A.Z C-Terminal Tail
ρ = 0.63, p = 0.00025
ρ = 0.056, p = 0.8
Transcriptional State (PC1 from Fig. 2B)
20
10
0
−10
−20
0.00
0.25
0.50
0.75
1.00
1.25
1.50
Sequence Divergence from S. pombe

590

591

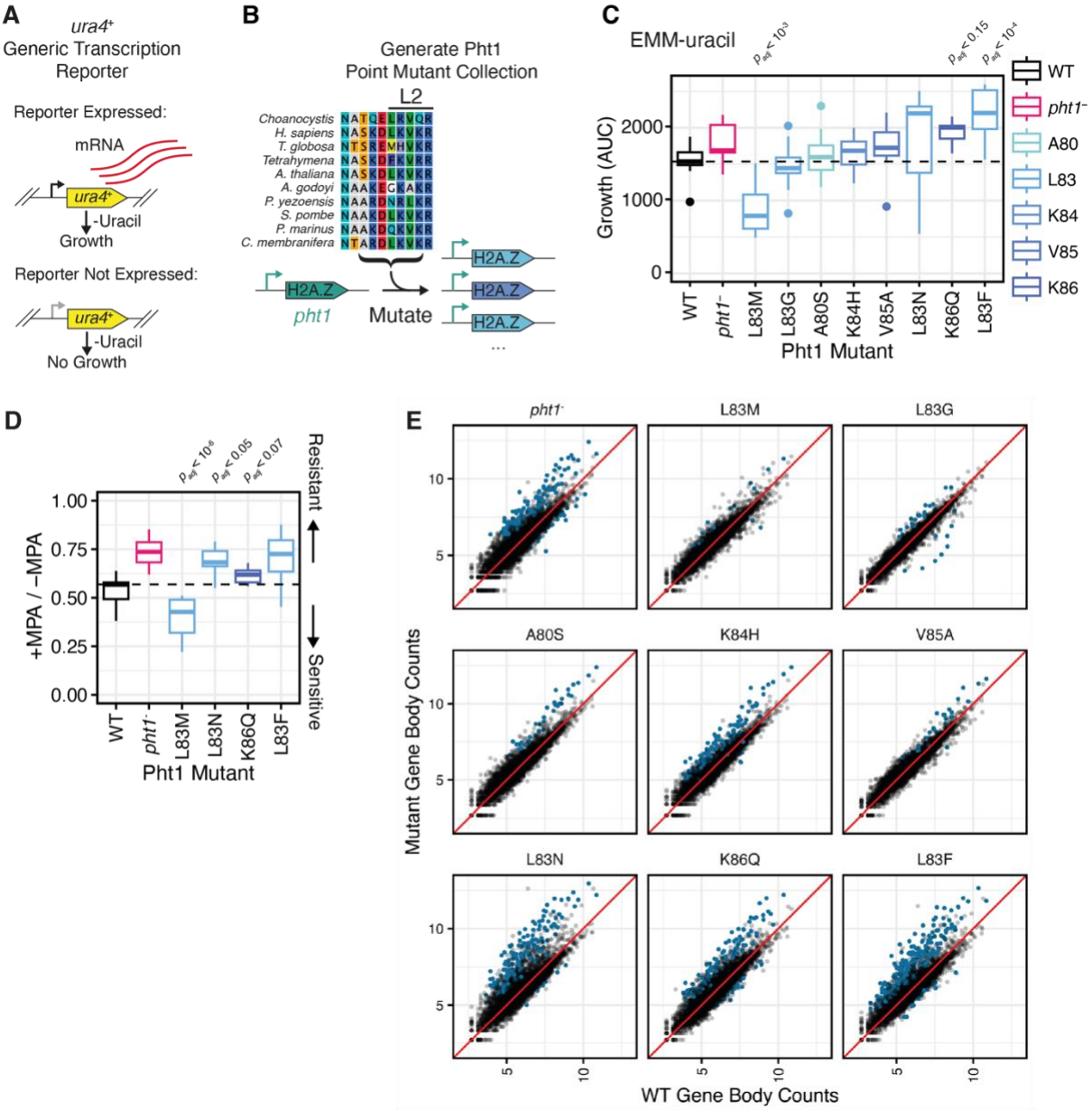
Figure 2. The L2 region of H2A.Z is a key determinant of its function. A, Mean pairwise sequence identity of H2A.Z and H2A.X residues across eukaryotes, averaged using a 3-amino acid sliding window. The L2 region is highlighted in grey on the histone secondary structure, with the core domain we identify indicated. Dashed lines indicate the N- and C-terminal tails that were not included in this analysis due to variable length and resulting gappy alignment. **B**, Experimental scheme for generation of Hta1/Pht1 chimeras. **C**, Mean growth (AUC) for four biological replicates in each indicated condition, which are as in Fig. 1B. **D**, Spearman's correlation analysis between the sequence divergence from *S. pombe* Pht1 of either the L2 or C-terminal tail region and principal component 1 from mRNA-seq analysis from Fig. 1B. In this analysis, the *S. pombe* Pht1 sequence corresponds to a value of zero, with a larger number corresponding to a greater divergence. See also Fig. S5B.


A
S. pombe Nuclear Proteome
AlphaFold2
Pht1 L2
B
Pht1 L2 Variants
Reduced Nuclear Proteome
C
IPTM Variance Across All Pht1 L2 Variants
Count
Spt6
IPTM Variance
D
AlphaFold2 Predicted Structure
Spt6
Death-Like Domain (DLD)
Pht1 L2
E
Max
Min
Spt6 DLD Lipophilicity
Pht1 L2
Pht1 L2 L83M
Pht1 L2 L83F
Pht1 L2 L83N
F
14x His
Spt6
Pht1
GST
Hta1
Nickel Resin
H
Biotin-Labeled L2 Peptides (Vast Excess)
Wash
Detect Biotin
J
Spt6
N
C
Hydrophobic Residues
Sequence Identity (%)
Residue Position
G
Nickel Resin
14xHis-Spt6
GST-Pht1
GST-Hta1
GST
Pht1
Hta1
kDa
I
H2A.Z L2 Species
Pht1 Mutant
Biotin Only
Resin + Peptide
Pht1 DD
Pht1 L2
Hta1 L2
H. sapiens
A. thaliana
A. godoyi
T. globosa
P. marinus
L83M
L83F
L83N
Dye Front
Biotin-Peptide
Hamming Distance:
- 4 1 2 3 6 2 1 1 1

604

605

Figure 3. The L2 region of H2A.Z binds to transcription elongation factor Spt6. A, Scheme explaining the AlphaFold2 virtual interaction screen against *S. pombe* Pht1. **B,** Scheme explaining the AlphaFold2 approach used to calculate the IPTM variance for 17 Pht1 L2 variants, which is plotted in **C**. The reduced nuclear proteome was defined by the top 500 WT *S. pombe* H2A.Z interactors from (A). **D,** AlphaFold2 predicted structure of the Pht1 L2 interaction with Spt6, with the the death-like-domain (DLD) highlighted with shading. **E,** Example AlphaFold2 structures of *S. pombe* Pht1, Hta1, and Pht1 mutants. Spt6 DLD is colored with a lipophilicity surface, and each H2A position 83 (e.g. L83 for Pht1) is labelled to highlight its interaction with the Spt6 hydrophobic groove. **F,** Experimental scheme for *in vitro* interaction assays of full-length GST-Pht1, GST-Hta1, with 14xHis-Spt6(282-1159). **G,** Coomassie-stained SDS-PAGE gel of Ni-NTA-pulldowns of recombinant *S. pombe* GST-Pht1 or GST-Hta1 with *S. pombe* His-Spt6(282-1159). Detection done with colloidal Coomassie staining. **H,** Experimental scheme showing the pulldown experiment modified for biotinylated peptides. **I,** Binding assay of Pht1-derived peptides to *S. pombe* 14xHis-Spt6(282-1159) immobilized on Ni-NTA resin. Detection was done either by colloidal Coomassie staining (for 14xHis-Spt6 loading control) or by a streptavidin-HRP conjugate (for biotinylated peptides). Representative of five experiments, each with similar results. **J,** Conservation analysis of Spt6 DLD. Conservation is calculated as the mean pairwise amino acid identity across Spt6 orthologs (see Methods). The location of hydrophobic residues are highlighted in red, with those predicted to interact with Pht1 L2 further indicated by an asterisk (*). The following residues are considered hydrophobic for the purpose of this analysis: A, V, I, L, F, M, W, Y.


A
ura4+
Generic Transcription Reporter
Reporter Expressed:
mRNA
ura4+
-Uracil
Growth
Reporter Not Expressed:
ura4+
-Uracil
No Growth
B
Generate Pht1 Point Mutant Collection
L2
Choanocystis
H. sapiens
T. globosa
Tetrahymena
A. thaliana
A. godoyi
P. yezoensis
S. pombe
P. marinus
C. membranifera
H2A.Z
pht1
Mutate
H2A.Z
H2A.Z
H2A.Z
...
C
EMM-uracil
Growth (AUC)
Pht1 Mutant
WT
pht1-
L83M
L83G
A80S
K84H
V85A
L83N
K86Q
L83F
A80
L83
K84
V85
K86
D
+MPA / −MPA
Resistant
Sensitive
Pht1 Mutant
E
pht1-
L83M
L83G
A80S
K84H
V85A
L83N
K86Q
L83F
Mutant Gene Body Counts
WT Gene Body Counts

627

628

Figure 4. H2A.Z L2 sequence variation explains species-specific effects on transcription. **A,** Experimental scheme explaining the *ura4*$^{+}$ reporter gene. **B,** Experimental scheme describing the generation of H2A.Z point mutants. **C,** Growth of H2A.Z point mutant strains bearing the transcriptional reporter from (D) in EMM-Uracil. The mean of WT is represented with a dashed line. Each boxplot represents 12 biological replicates, adjusted *p*-values (p_{adj}) are calculated using a two-way ANOVA followed by a post-hoc Tukey test. **D,** Growth of indicated *S. pombe* Pht1 mutants in the presence of 12.5 µM mycophenolic acid (MPA) expressed as the ratio of their paired growth in the absence of the drug. Data are 9 biological replicates and the experiment was repeated three times with similar results. Boxplots represent the median bounded by quartiles. Adjusted *p*-values (p_{adj}) are calculated using a two-way ANOVA followed by a post-hoc Tukey test. **E,** Differential count analysis for gene body qPRO-seq signal calculated using DESeq2. Red line, diagonal (Y=X) line drawn to aid in interpretation. Statistically significant (p_{adj}<0.1) changes are indicated in blue.