

Thermodynamic stability of histone H3 is a necessary but not sufficient driving force for its evolutionary conservation

Srinivas Ramachandran^{1,2}, Lisa Vogel¹, Brian D. Strahl^{*1}, Nikolay V. Dokholyan^{*1,2}

¹Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599 USA, ²Program in Molecular and Cellular Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599 USA

Supporting Information

Figure S1. Inter-histone interfaces of H3 and H4. The crystal structure of the yeast nucleosome (PDB 1id3) is rendered in cartoon representation (a, c). The H3-H3' interface is shown in **b** and the H4-H2A interface is shown in **d**. H3 and H3' form a homo-dimer. The H4-H2A interface is formed by a short beta sheet, whose hydrogen bonds are denoted with dashed lines in **d**. The residues considered in this study are shown in stick representation and are labelled. The structures were rendered using PyMOL (<http://www.pymol.org>).

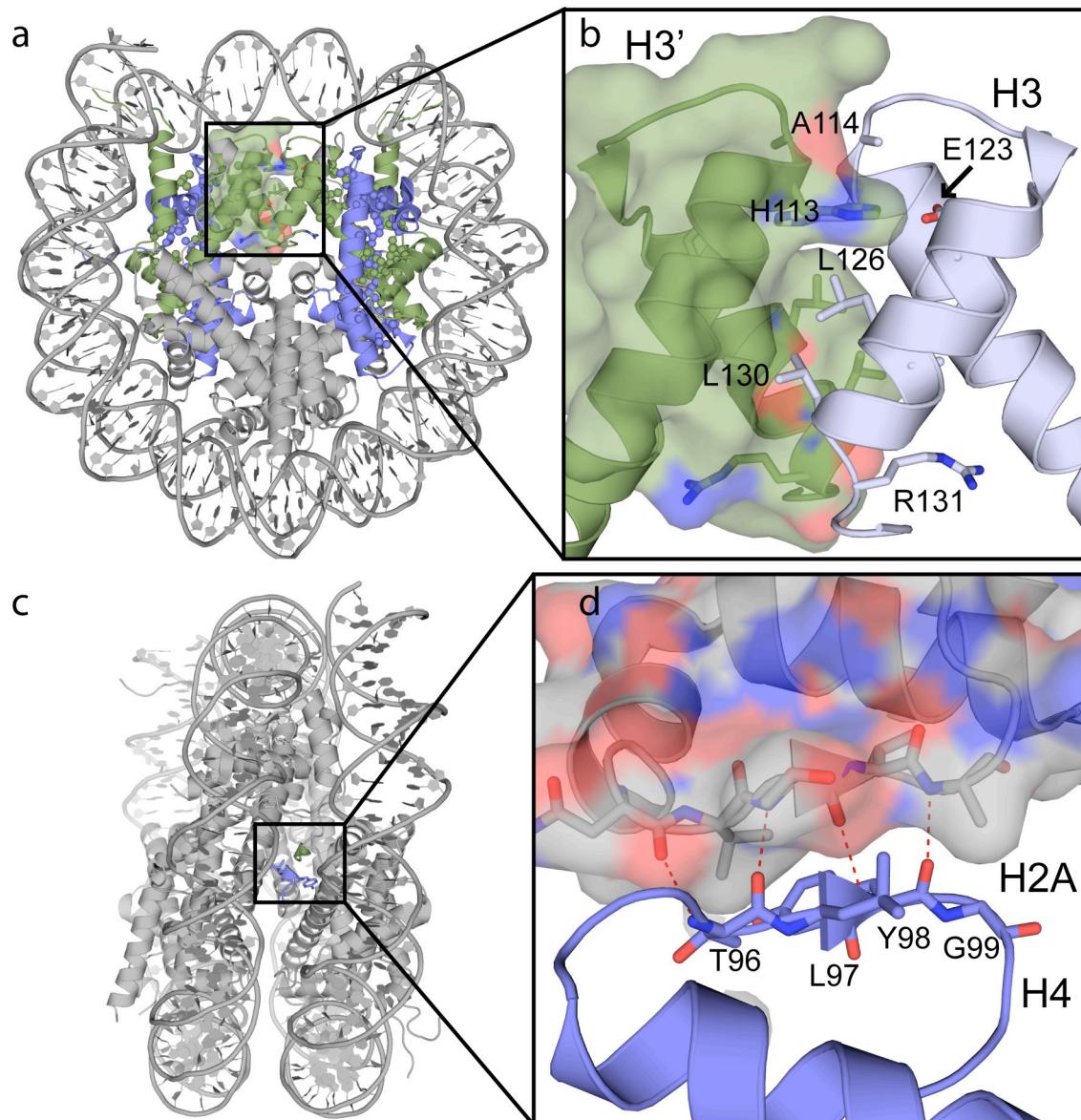


Figure S2. Structure of the yeast mononucleosome and location of H3 and H4 buried and interface residues. The crystal structure of the yeast nucleosome (PDB 1id3) is rendered in cartoon representation and consists of 147 base pairs of DNA wrapped around two copies of each of the four core histone proteins: H2A, H2B, H3, and H4. The buried residues of H3 and H4 are represented as scaled spheres. H3 is colored blue and H4 is colored green. The structure was rendered using PyMOL (<http://www.pymol.org>).

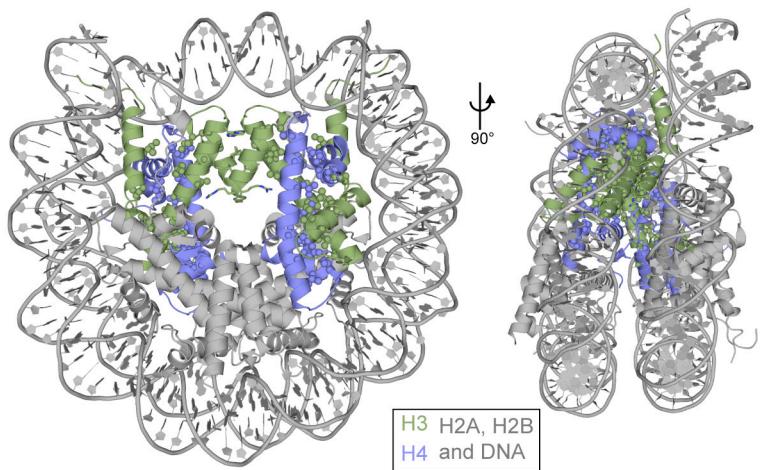


Figure S3. Distinct conservation profiles of three sets of residues in H3. Residues that feature much higher evolutionary entropy compared to other positions are shown as red circles (**A**). Residues featuring low evolutionary entropy, indicating conservation much higher than required by stability are indicated as blue triangles (**A,B**). Residues featuring evolutionary entropy that has modest correlation with Medusa-derived entropy are shown as black open squares (**A,B**). The red line indicates the linear regression ($r=0.64$) between Medusa-derived entropy and evolutionary entropy for positions shown in black open squares (**B**).

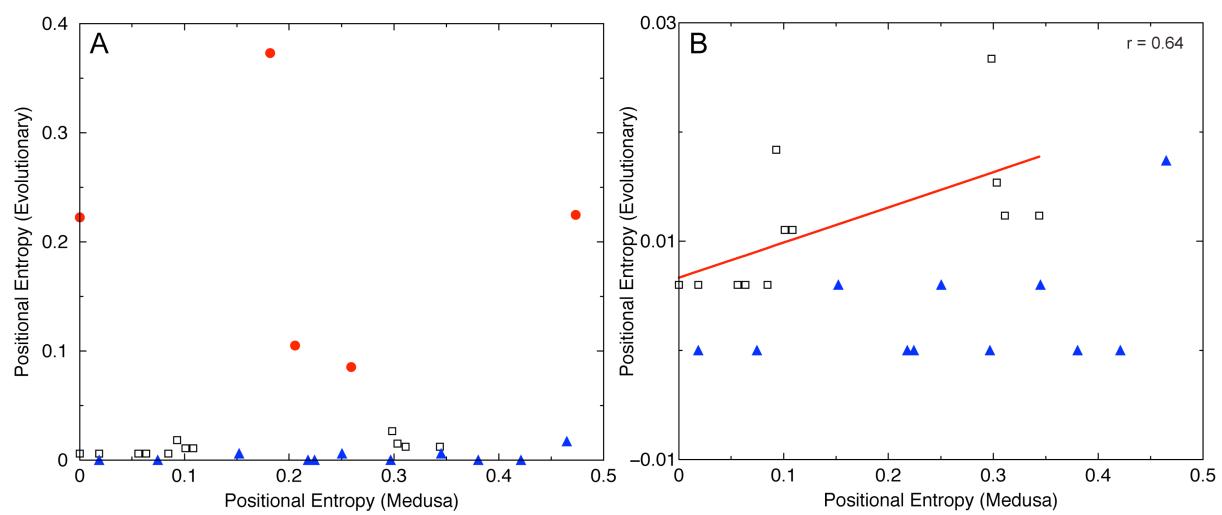


Figure S4. Tree-based conservation profile of position 95 in H3. The conservation of position 95 is determined at each node of the phylogenetic tree constructed from the multiple sequence alignment of H3. We observe that the nodes representing species mostly from kingdom Fungi, have a different preferred amino acid (Serine) compared to other nodes (Alanine), indicating tree-based inheritance.

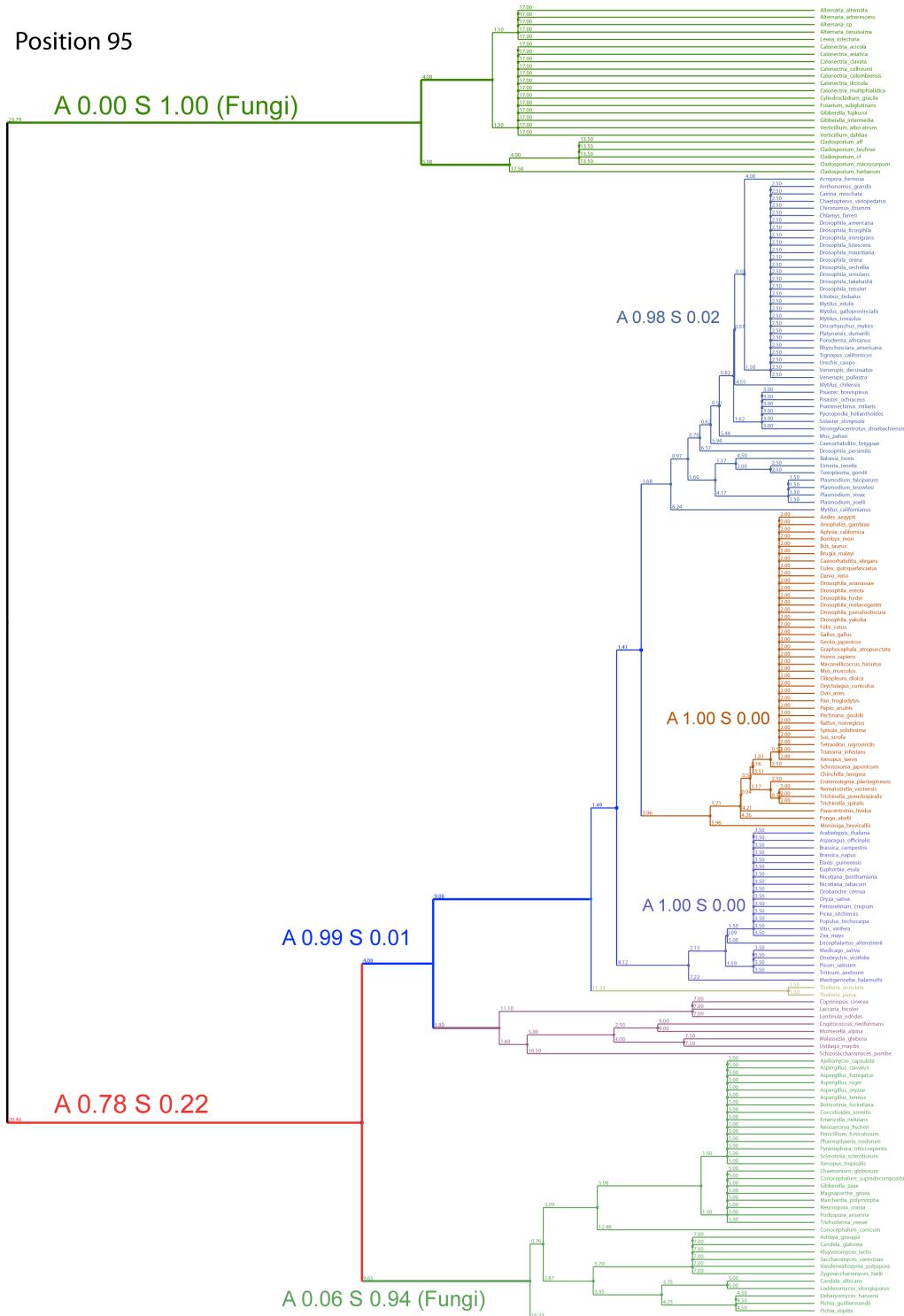


Figure S5. Tree-based conservation profile of position 96 in H3. The conservation of position 96 is determined at each node of the phylogenetic tree constructed from the multiple sequence alignment of H3. We observe that the nodes representing species mostly from kingdom Fungi, have a different preferred amino acid (Serine) compared to the node containing plant kingdom (Alanine), the node containing *Homo sapiens* and many species of genus *Drosophila* (Serine), indicating tree-based inheritance.

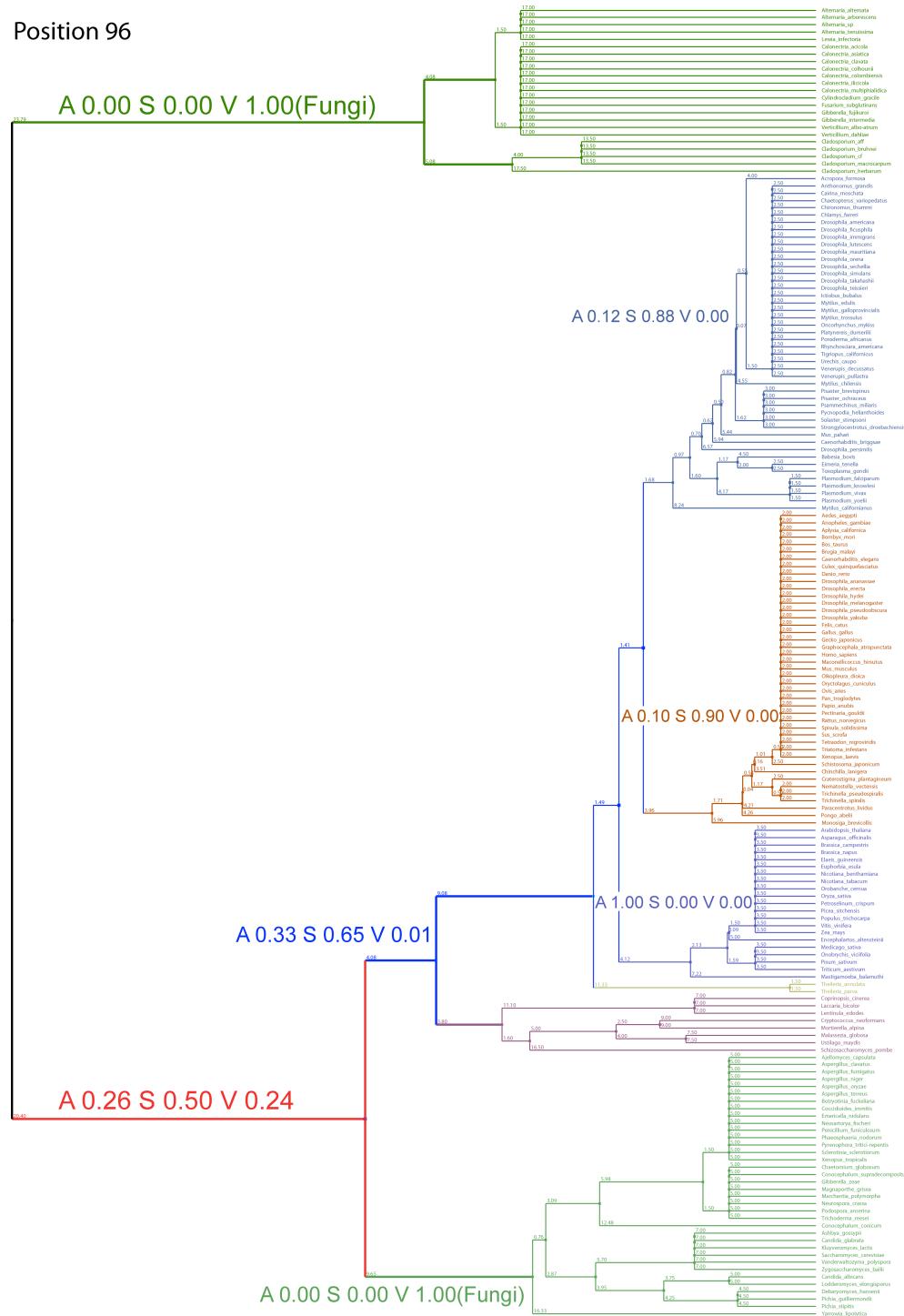


Figure S6. Tree-based conservation profile of position 130 in H3. The conservation of position 130 is determined at each node of the phylogenetic tree constructed from the multiple sequence alignment of H3. We observe that the nodes representing species mostly from kingdom Fungi, have a different preferred amino acid (Leucine) compared to other nodes (Isoleucine), indicating tree-based inheritance.

Position 130

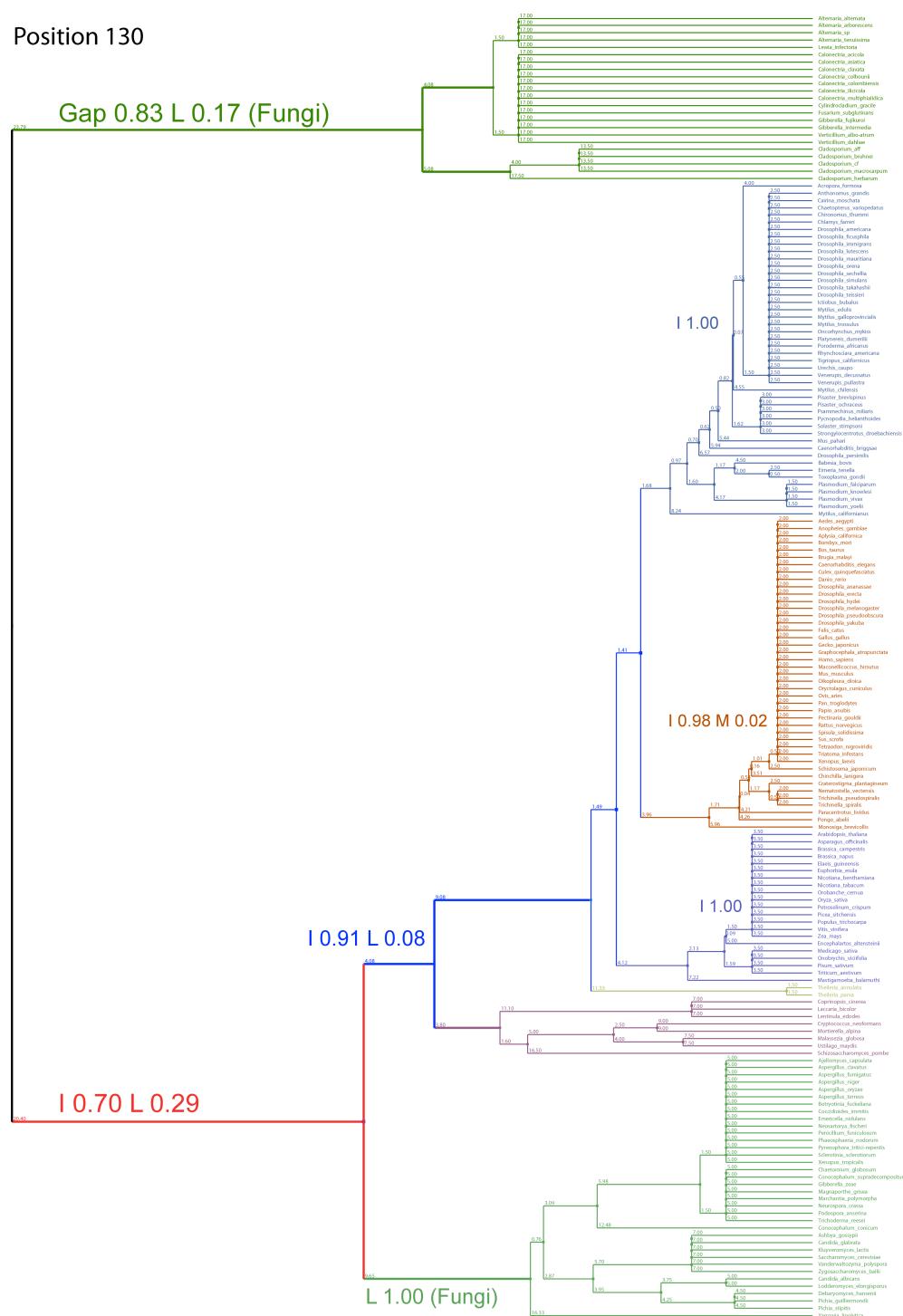


Figure S7. Organisms whose H3 sequences were used to construct the phylogenetic tree. The color coding of organisms is based on the nodes represented in the tree presented in Figures S3, S4 and S5.

<i>Alternaria alternata</i>	<i>Acropora formosa</i>	<i>Aedes aegypti</i>	<i>Arabidopsis thaliana</i>	<i>Ajellomyces capsulata</i>
<i>Alternaria arborescens</i>	<i>Anthonomus grandis</i>	<i>Anopheles gambiae</i>	<i>Asparagus officinalis</i>	<i>Aspergillus clavatus</i>
<i>Alternaria tenuissima</i>	<i>Cairina moschata</i>	<i>Aplysia californica</i>	<i>Brassica campestris</i>	<i>Aspergillus fumigatus</i>
<i>Lewia infectoria</i>	<i>Chaetopterus variopedatus</i>	<i>Bombyx mori</i>	<i>Brassica napus</i>	<i>Aspergillus niger</i>
<i>Calonectria acicola</i>	<i>Chironomus thummi</i>	<i>Brugia malayi</i>	<i>Elaeis guineensis</i>	<i>Aspergillus oryzae</i>
<i>Calonectria asiatica</i>	<i>Chlamys farreri</i>	<i>Caenorhabditis elegans</i>	<i>Euphorbia esula</i>	<i>Aspergillus terreus</i>
<i>Calonectria clavata</i>	<i>Drosophila americana</i>	<i>Culex quinquefasciatus</i>	<i>Nicotiana benthamiana</i>	<i>Botryotinia fuckeliana</i>
<i>Calonectria colhounii</i>	<i>Drosophila ficusphila</i>	<i>Drosophila ananassae</i>	<i>Nicotiana tabacum</i>	<i>Coccidioides immitis</i>
<i>Calonectria colombiensis</i>	<i>Drosophila immigrans</i>	<i>Drosophila erecta</i>	<i>Orobanche cernua</i>	<i>Emericella nidulans</i>
<i>Calonectria ilicicola</i>	<i>Drosophila lutescens</i>	<i>Drosophila hydei</i>	<i>Petroselinum crispum</i>	<i>Neosartorya fischeri</i>
<i>Calonectria multiphialidica</i>	<i>Drosophila mauritiana</i>	<i>Drosophila melanogaster</i>	<i>Picea sitchensis</i>	<i>Penicillium funiculosum</i>
<i>Cylindrocladium gracile</i>	<i>Drosophila orena</i>	<i>Drosophila pseudoobscura</i>	<i>Populus trichocarpa</i>	<i>Phaeosphaeria nodorum</i>
<i>Fusarium subglutinans</i>	<i>Drosophila sechellia</i>	<i>Drosophila yakuba</i>	<i>Vitis vinifera</i>	<i>Pyrenophora tritici-repentis</i>
<i>Gibberella fujikuroi</i>	<i>Drosophila simulans</i>	<i>Gallus gallus</i>	<i>Encephalartos altensteinii</i>	<i>Sclerotinia sclerotiorum</i>
<i>Gibberella intermedia</i>	<i>Drosophila takahashi</i>	<i>Gecko japonicus</i>	<i>Medicago sativa</i>	<i>Xenopus tropicalis</i>
<i>Verticillium albo-atrum</i>	<i>Drosophila teissieri</i>	<i>Graphocephala atropunctata</i>	<i>Onobrychis viciifolia</i>	<i>Chaetomium globosum</i>
<i>Verticillium dahliae</i>	<i>Ictiobus bubalus</i>	<i>Homo sapiens</i>	<i>Pisum sativum</i>	<i>Conocephalum supradecompositum</i>
<i>Cladosporium aff</i>	<i>Mytilus edulis</i>	<i>Maconellicoccus hirsutus</i>	<i>Mastigamoeba balamuthi</i>	<i>Gibberella zeae</i>
<i>Cladosporium bruhnei</i>	<i>Mytilus galloprovincialis</i>	<i>Mus musculus</i>	<i>Theileria annulata</i>	<i>Magnaporthe grisea</i>
<i>Cladosporium cf</i>	<i>Mytilus trossulus</i>	<i>Oikopleura dioica</i>	<i>Theileria parva</i>	<i>Marchantia polymorpha</i>
<i>Cladosporium macrocarpum</i>	<i>Oncorhynchus mykiss</i>	<i>Oryctolagus cuniculus</i>	<i>Coprinopsis cinerea</i>	<i>Neurospora crassa</i>
<i>Cladosporium herbarum</i>	<i>Platynereis dumerilii</i>	<i>Pan troglodytes</i>	<i>Laccaria bicolor</i>	<i>Podospora anserina</i>
	<i>Poroderma africanus</i>	<i>Papio anubis</i>	<i>Lentinula edodes</i>	<i>Trichoderma reesei</i>
	<i>Rhynchosciara americana</i>	<i>Pectinaria gouldii</i>	<i>Cryptococcus neoformans</i>	<i>Conocephalum conicum</i>
	<i>Tigriopus californicus</i>	<i>Rattus norvegicus</i>	<i>Mortierella alpina</i>	<i>Ashbya gossypii</i>
	<i>Urechis caupo</i>	<i>Spisula solidissima</i>	<i>Malassezia globosa</i>	<i>Candida glabrata</i>
	<i>Venerupis decussatus</i>	<i>Tetraodon nigroviridis</i>	<i>Ustilago maydi</i>	<i>Kluyveromyces lactis</i>
	<i>Venerupis pullastra</i>	<i>Triatoma infestans</i>	<i>Schizosaccharomyces pombe</i>	<i>Saccharomyces cerevisiae</i>
	<i>Mytilis chilensis</i>	<i>Xenopus laevis</i>		<i>Vanderwaltzyna polyspora</i>
	<i>Pisaster brevispinus</i>	<i>Schistosoma japonicum</i>		<i>Zygosaccharomyces bailii</i>
	<i>Pisaster ochraceus</i>	<i>Chinchilla lanigera</i>		<i>Candida albicans</i>
	<i>Psammechinus miliaris</i>	<i>Craterostigma plantagineum</i>		<i>Lodderomyces elongisporus</i>
	<i>Pycnopodia helianthoides</i>	<i>Nematostella vectensis</i>		<i>Debaryomyces hansenii</i>
	<i>Solaster stimpsoni</i>	<i>Trichinella spiralis</i>		<i>Pichia stipitis</i>
	<i>Strongylocentrotus droebachiensis</i>	<i>Paracentrotus lividus</i>		<i>Yarrowia lipolytica</i>
	<i>Caenorhabditis briggsae</i>	<i>Pongo abelii</i>		
	<i>Drosophilida persimilis</i>	<i>Monosiga brevicollis</i>		
	<i>Babesia bovis</i>			
	<i>Eimeria tenella</i>			
	<i>Toxoplasma gondii</i>			
	<i>Plasmodium falciparum</i>			
	<i>Plasmodium knowlesi</i>			
	<i>Plasmodium vivax</i>			
	<i>Plasmodium yoelii</i>			
	<i>Mytilus californianus</i>			

Table S1. List of mutations in H3 extracted from the HistoneHits database.

Mutation	Experimental Score ¹	Medusa $\Delta\Delta G$ (kcal/mol)	S.E.M ²
I51A	-2.0	4.4	0.4
L61A	0.0	2.0	0.5
F67A	-0.7	6.3	0.4
L70A	0.0	5.5	0.4
V71A	0.0	3.6	0.4
I74A*	-0.5	4.0	0.4
A75S	0.0	-0.1	0.4
A88S*	-0.5	0.5	0.5
A91S*	-0.5	0.4	0.5
L92A*	-1.0	4.6	0.5
Q93A	-1.3	2.8	0.4
Q93E	0.0	2.0	0.5
S95A	0.0	-1.3	0.4
S95D	-2.0	1.5	0.5
V96A	0.0	1.0	0.5
L100A	0.0	5.5	0.5
L103A	-1.3	3.7	0.4
F104A*	-2.0	4.9	0.5
T107A	0.0	2.7	0.4
T107D	-2.0	9.3	0.4
A110S*	-0.5	-0.4	0.4
A111S*	-0.5	2.0	0.4
H113A	-1.5	8.1	0.5
H113Q	-2.0	3.9	0.5
A114S	-0.5	0.1	0.5
I119A	-1.3	2.8	0.4
D123A	-1.3	-0.4	0.4
D123N	-2.0	0.3	0.4
I124A*	-2.0	4.2	0.4
L126A	-1.0	4.8	0.4
A127S	-0.5	1.4	0.5
L130A	-1.3	10.1	0.4
R131A	-1.0	0.0	0.4
R131K	-0.5	0.3	0.4

¹Average of the 5-point phenotypic value from all the lethality assays of the mutant submitted to the database. A score of -0.67 was considered significantly affecting viability. *The results of temperature sensitivity assay are used for these mutants, as these mutants were lethal only at higher temperatures. ²The standard error of mean (SEM) obtained for the Medusa calculations.

Table S2. List of mutations in H4 extracted from the HistoneHits database.

Mutation	Experimental Score ¹	Medusa $\Delta\Delta G$ (kcal/mol)	S.E.M ²
I29A*	-1.0	3.6	0.4
A33S	0.0	0.3	0.4
I34A	-1.3	5.8	0.5
L37A*	-2.0	4.0	0.5
A38S*	-1.0	0.7	0.4
V43A	0.0	4.3	0.5
I50A	0.0	4.4	0.5
V54A	0.0	2.4	0.5
L58A*	-0.5	4.6	0.4
L62A*	-1.0	5.6	0.5
V65A	0.0	3.0	0.4
I66A*	-1.0	3.8	0.5
S69A	0.0	-2.0	0.5
S69D	-2.0	4.2	0.6
T73A	0.0	0.5	0.5
T73D	-2.0	3.6	0.5
V81A*	-0.5	2.4	0.4
V86A*	-0.5	2.5	0.5
A89S	0.0	3.8	0.4
L90A	-2.0	4.0	0.4
T96A	0.0	-0.5	0.4
T96D	0.0	1.3	0.4
L97A	0.0	2.4	0.4
Y98A	-0.5	7.5	0.5
Y98E	-1.0	10.5	0.5
Y98F	0.0	-2.8	0.5
G99A*	-0.5	4.7	0.5

¹Average of the 5-point phenotypic value from all the lethality assays of the mutant submitted to the database. A score of -0.67 was considered significantly affecting viability. *The results of temperature sensitivity assay are used for these mutants, as these mutants were lethal only at higher temperatures. ²The standard error of mean (SEM) obtained for the Medusa calculations.

Table S3. Evolutionary and Medusa positional entropy values of buried and interface residues of H3.

Residue	Evolutionary Entropy ¹	Medusa Entropy ²
I51	0.00	0.30
L61	0.00	0.42
F67	0.01	0.25
L70	0.01	0.11
V71	0.01	0.31
I74	0.02	0.30
A75	0.00	0.38
A88	0.00	0.22
A91	0.00	0.22
L92	0.00	0.07
Q93	0.01	0.34
S95	0.22	0.47
V96	0.37	0.18
L100	0.01	0.15
L103	0.02	0.09
F104	0.01	0.02
T107	0.01	0.34
A110	0.11	0.21
A111	0.01	0.06
H113	0.01	0.00
A114	0.02	0.46
I119	0.01	0.08
D123	0.03	0.30
I124	0.09	0.26
L126	0.01	0.06
A127	0.01	0.10
L130	0.22	0.00
R131	0.00	0.02

¹Entropy values obtained from HSSP database have been normalized by $\ln(20)$, the maximal possible entropy, so that the range of entropy values is between 0-1. ²Normalized entropy obtained using the residue propensities in Medusa calculations as described in the Methods.

Table S4. Evolutionary and Medusa positional entropy values of buried and interface residues of H4.

Residue	Evolutionary Entropy ¹	Medusa Entropy ²
I29	0.08	0.44
A33	0.16	0.51
I34	0.09	0.04
L37	0.08	0.25
A38	0.02	0.19
V43	0.01	0.04
I50	0.15	0.15
V54	0.26	0.45
L58	0.06	0.16
L62	0.09	0.12
V65	0.12	0.29
I66	0.17	0.31
S69	0.22	0.34
T73	0.09	0.36
V81	0.03	0.24
V86	0.05	0.19
A89	0.09	0.08
L90	0.02	0.03
T96	0.18	0.39
L97	0.07	0.08
Y98	0.01	0.02
G99	0.02	0.00

¹Entropy values obtained from HSSP database have been normalized by $\ln(20)$, the maximal possible entropy, so that the range of entropy values is between 0-1. ²Normalized entropy obtained using the residue propensities in Medusa calculations as described in the Methods.

Table S5. Yeast strains used in the study.

Residue	Evolutionary Entropy ¹	Medusa Entropy ²	Buried /Interface
I51	0.00	0.30	Buried
L61	0.00	0.42	Buried
F67	0.01	0.25	Buried
A75	0.00	0.38	Buried
A88	0.00	0.22	Buried
A91	0.00	0.22	Buried
L92	0.00	0.07	Buried
Q93	0.01	0.34	Buried
L100	0.01	0.15	Buried
A114	0.02	0.46	Interface
R131	0.00	0.02	Interface

¹Entropy values obtained from HSSP database have been normalized by ln(20), the maximal possible entropy, so that the range of entropy values is between 0-1. ²Normalized entropy obtained using the residue propensities in Medusa calculations as described in the Methods.

Table S6. Yeast strains used in the study.

Strain	Genotype	References
WZY42	<i>MATa, ura3-52, lys2-801, ade2-101, trp1Δ63, his3Δ200, leu2Δ1, hht1-hhf1::pWZ405-F2F9-LEU2, hht2-hhf2::pWZ403-F4F10-HIS3, Ycp50-copyII (HHT2-HHF2)</i>	Zhang et al. 1998 ¹
WZY42 H3-H113A	Isogenic to WZY42, plus Ycp50-copyII (HHT2(H113A) -HHF2)	This study
WZY42 H3-L126A	Isogenic to WZY42, plus Ycp50-copyII (HHT2(L126A) -HHF2)	This study
WZY42 H3-L130A	Isogenic to WZY42, plus Ycp50-copyII (HHT2(L130A) -HHF2)	This study
YBL574	<i>MATa, leu2Δ1, his3Δ200, ura3-52, trp1Δ63, lys2-128δ, (hht1-hhf1)ΔLEU2(hht2-hhf2)Δ::HIS3 Ty912Δ35-lacZ::his4, {pDM9-HHT1-HHF1-URA3}</i>	Carrozza et al. 2005 ²
YBL574 H3-H113A	Isogenic to YBL574, plus Ycp50-copyII (HHT2(H113A) -HHF2)	This study
YBL574 H3-L126A	Isogenic to YBL574, plus Ycp50-copyII (HHT2(L126A) -HHF2)	This study
YBL574 H3-L130A	Isogenic to YBL574, plus Ycp50-copyII (HHT2(L130A) -HHF2)	This study

¹Zhang, W; Bone, JR; Edmondson, DG; Turner, BM; Roth, SY. Essential and redundant functions of histone acetylation revealed by mutation of target lysines and loss of the Gcn5p acetyltransferase. *EMBO J.* 1998;17:3155–3167.

²M.J. Carrozza, B. Li, L. Florens, T. Suganuma, S.K. Swanson, K.K. Lee, W.J. Shia, S. Anderson, J. Yates, M.P. Washburn and J.L. Workman, Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription, *Cell* 123 (2005), 581–592.