Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y

Sonja M. Wiedemann,¹ Silke N. Mildner,¹ Clemens Bönisch,¹ Lars Israel,² Andreas Maiser,³ Sarah Matheisl,¹ Tobias Straub,¹ Rainer Merkl,⁴ Heinrich Leonhardt,^{3,6} Elisabeth Kremmer,⁵ Lothar Schermelleh,³ and Sandra B. Hake^{1,6}

ucleosomal incorporation of specialized histone variants is an important mechanism to generate different functional chromatin states. Here, we describe the identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y. Their messenger RNAs are found in certain human cell lines, in addition to several normal and malignant human tissues. In keeping with their primate specificity, H3.X and H3.Y are detected in different brain regions. Transgenic H3.X and H3.Y proteins are stably incorporated

into chromatin in a similar fashion to the known H3 variants. Importantly, we demonstrate biochemically and by mass spectrometry that endogenous H3.Y protein exists in vivo, and that stress stimuli, such as starvation and cellular density, increase the abundance of H3.Y-expressing cells. Global transcriptome analysis revealed that knockdown of H3.Y affects cell growth and leads to changes in the expression of many genes involved in cell cycle control. Thus, H3.Y is a novel histone variant involved in the regulation of cellular responses to outside stimuli.

Introduction

Chromatin, the storage and regulatory form of genetic information in eukaryotes, consists of nucleosomes that are composed of DNA and octamers of the core histones H2A, H2B, H3, and H4 (van Holde, 1988). To allow changes in chromatin structure, which are necessary to promote different biological functions, several interconnected mechanisms have evolved (for review see Bönisch et al., 2008). Among others, these include the sliding or eviction of nucleosomes by ATP-dependent chromatin remodeling machines (for review see Clapier and Cairns, 2009), posttranslational modifications (PTMs) of histone proteins (Strahl and Allis, 2000), and the exchange of canonical histones with specialized histone variants (for reviews see Pusarla and Bhargava, 2005; Bernstein and Hake, 2006). Histone variants

Correspondence to Sandra B. Hake: Sandra.hake@med.uni-muenchen.de

Abbreviation used in this paper: 3D-SIM, 3D structured illumination microscopy; DG, dentate gyrus; DIC, differential interference contrast; GFAP, glial fibrillary acidic protein; GO, gene ontology; IF, immunofluorescence; IP, immunoprecipitation; LC-MS/MS, liquid chromatography tandem mass spectrometry; NeuN, neuronal nuclei; PTM, posttranslational modification; qPCR, quantitative PCR; RP-HPLC, reversed-phase HPLC; SO, starvation and overgrowth.

differ in sequence and expression timing from their canonical counterparts and are enriched in chromatin of specific functional states, ranging from DNA repair and centromere determination to the regulation of gene expression. In mammals, variants of the H3, H2A, and H2B families of histones have been identified whose incorporation results in nucleosomes with novel functional and structural properties (Suto et al., 2000; Abbott et al., 2001; Angelov et al., 2003; Bao et al., 2004; Gautier et al., 2004).

To date, five different H3 variants have been found in mammals: H3.1, H3.2, H3.3, H3.1t (tH3), and CENP-A. The centromeric H3 variant CENP-A causes changes to the nucleosomal structure (Black et al., 2004) and is crucial for proper chromosome segregation (for review see Allshire and Karpen, 2008). tH3 is a testis-specific histone variant with a putative

© 2010 Wiedemann et al. This article is distributed under the terms of an Attribution–Noncommercial–Share Alike–No Mirror Sites license for the first six months after the publication date (see http://www.rupress.org/terms). After six months it is available under a Creative Commons License (Attribution–Noncommercial–Share Alike 3.0 Unported license, as described at http://creativecommons.org/licenses/by-nc-sa/3.0/).

¹Adolf-Butenandt-Institute, Department of Molecular Biology, and ²Protein Analysis Unit, Ludwig Maximilians University of Munich, 80336 Munich, Germany

³LMU Biocenter, Department of Biology, Ludwig Maximilians University of Munich, 82152 Planegg-Martinsried, Germany

⁴Department of Biochemistry II, University Regensburg, 93053 Regensburg, Germany

⁵Institute of Molecular Immunology, Helmholtz Center Munich, German Research Center for Environmental Health, 81377 Munich, Germany

⁶Center for Integrated Protein Science Munich 5, 81377 Munich, Germany

function in chromatin reorganization during spermatogenesis (Witt et al., 1996). H3.1 and H3.2 sequences are distinguishable by just one amino acid. Although expression of both is replication dependent (Ahmad and Henikoff, 2002a), they differ in their cell type expression levels as well as their enrichment of PTMs (Hake et al., 2006). Furthermore, H3.1 has been implicated in DNA damage response pathways (Polo et al., 2006) and is deposited into chromatin by the chaperone complex CAF-1 (Tagami et al., 2004), whereas H3.3 is expressed and incorporated into chromatin in a replication-independent manner by HIRA (Tagami et al., 2004). The latter variant is highly decorated with modifications associated with gene transcription (McKittrick et al., 2004; Hake et al., 2006) and is thought to be involved in activating gene expression (Ahmad and Henikoff, 2002a) and epigenetic reprogramming (for review see Santenard and Torres-Padilla, 2009).

Here, we describe the identification of two novel primatespecific histone H3 variants (H3.X and H3.Y) that are transcribed at low levels in some human cell lines and in normal and malignant tissues. Their amino acid sequences are highly similar to each other, but differ in several functionally important residues from other H3 variants, e.g., S10 and S28. Transgenic H3.X and H3.Y proteins localize to the nucleus and are stably incorporated into chromatin. We demonstrate that endogenous H3.Y protein is highly expressed in some osteosarcoma cells. Interestingly, the number of cells expressing H3.Y is increased by particular stress stimuli. Knockdown of H3.Y leads to diminished growth and changes in the expression of many genes controlling cell cycle progression. The observed presence of primate-specific histone H3 variants in specific human brain areas opens up fascinating questions about their roles in human cell determination and differentiation.

Results

Identification of novel H3 variant genes

Using the nucleotide sequence of human histone H3.1f (HIST1H3I; available from GenBank/EMBL/DDBJ under accession no. NM_003533), we searched the public database (National Center for Biotechnology Information) and identified two highly similar genes initially annotated as pseudogenes (Fig. S1 A). These two intron-free genes, which we named H3.X(GenBank/EMBL/DDBJ accession no. LOC340096) and H3.Y (GenBank/EMBL/DDBJ accession no. LOC391769), are located on human chromosome 5 (5p15.1). Further database searches revealed the existence of similar genes in primate genomes (Pan troglodytes H3.X, GenBank/EMBL/DDBJ accession no. LOC471464; and H3.Y, GenBank/EMBL/DDBJ accession no. LOC471473; Macaca mulatta H3.X, GenBank/ EMBL/DDBJ accession no. LOC718189; and H3.Y, GenBank/ EMBL/DDBJ accession no. LOC718280; Fig. S1 B). Searches for these genes in other mammalian genomes yielded no positive hits, which suggests that they evolved in evolutionarily younger terms and might constitute primate-specific histones. Both human genes contain a sequence matching the translation initiation start site consensus (underline) for vertebrates (GCC-GCCACCAUGGCG; Kozak, 1991; Nakagawa et al., 2008), and

depending on the search program used ("polyadq" or PolyA_SVM program), H3.X and H3.Y 3' genomic sequences of human and primate origins were predicted to include a conserved poly-A site (Tabaska and Zhang, 1999; Cheng et al., 2006). Alignment of human and primate H3.X and H3.Y coding sequences with respective human H3 variant sequences revealed a higher similarity to H3.3 than to H3.2 and H3.1 (Fig. S1 C).

H3.X and H3.Y genes are predicted to encode proteins of 146 and 135 amino acids, respectively (Fig. 1 A). Both putative variant proteins are highly similar to each other, with differences of only four amino acids in their overlapping region (89.7% identity). H3.X has an unusual long C-terminal tail with no sequence homology to other proteins (Fig. 1 A). H3.X and H3.Y display interesting changes in amino acids that are known to be posttranslationally modified in H3.1, H3.2, and H3.3: serine 10 and 28, which are phosphorylated during mitosis (Hendzel et al., 1997) and immediate early gene induction (Clayton and Mahadevan, 2003), are altered to alanine and arginine, respectively. Regions surrounding lysine 14, usually acetylated and found in actively transcribed genes (Yan and Boyd, 2006), and lysine 79, methylated in transcriptionally active regions (Im et al., 2003) and upon DNA damage (Huyen et al., 2004), are completely altered (Fig. 1 A). In several instances, single amino acid changes between novel and known H3 variants have occurred, thereby generating or eliminating potential PTM sites. The region surrounding aa 87–90 is important for the chaperone-dependent chromatin incorporation of histone variants (Ahmad and Henikoff, 2002b). CAF-1 deposits H3.1, whereas HIRA catalyzes the incorporation of H3.3 (Tagami et al., 2004). The residues of H3.X and H3.Y in this region are identical to H3.3 (Fig. 1 A), leading to the hypothesis that these putative novel variants might be incorporated by HIRA in a replication-independent manner.

To assess whether both genes are expressed in human cells, we analyzed their RNA expression levels in a variety of cell lines by quantitative PCR (qPCR; Fig. 1 B). Because the nucleotide sequences of H3.X and H3.Y are almost identical (Fig. S1 A), we generated one primer set detecting both H3.X and H3.Y mRNAs and one additional primer set that hybridizes with H3.X-specific 3' sequences. We demonstrate that both H3.X and H3.Y genes are transcribed at low levels in the human osteosarcoma cell line U2OS, with H3.Y in higher amounts than H3.X, but \sim 16-fold lower than the histone variant CENP-A (Fig. S1 D). Minimal expression of H3.Y was detected in human HEK293 cells, whereas no expression was seen in HeLa as well as in mouse cells (negative control; Fig. 1 B). These data suggest that H3.X and H3.Y genes are transcribed in some human cell lines, albeit at low total levels.

Our observation that H3.X and H3.Y mRNAs are expressed in U2OS cells inspired us to examine RNAs from different human tissues by qPCR. Of particular interest were human malignant tissues, as U2OS cells are derived from a moderately differentiated sarcoma of the tibia. We also chose to investigate several regions of the human brain and testis due to the primate-specific appearance of *H3.X* and *H3.Y*. Interestingly, we found that some bone, breast, lung, and ovary tumors express low but significant levels of H3.X/Y mRNA (Fig. 1 C). Analyzing RNA from human brain areas and testis, H3.X/Y expression was

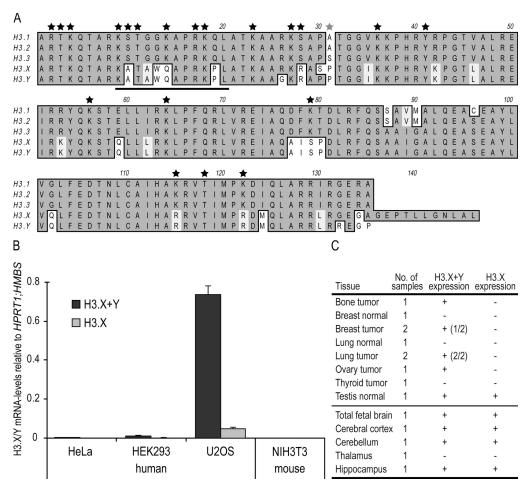


Figure 1. Sequence and mRNA expression analysis of novel H3 variants H3.X and H3.Y. (A) Amino acid sequence alignment of human histone variants H3.1, H3.2, and H3.3 with novel variants H3.X and H3.Y. Alignments were made with ClustalW Alignment (MacVector 10.0.2). Identical amino acids are highlighted in dark gray, similar amino acids are highlighted in light gray, and changes are set apart on a white background. The black bar indicates the peptide sequence used for antibody generation. Black stars mark known PTM sites in H3.1, H3.2, and H3.3. The gray star indicates an H3.3-specific modification site. Amino acid numbers are indicated on top. (B) qPCR analysis with cDNA from different human cell lines shows expression of H3.Y and to a lesser extent H3.X mRNA in U2OS cells. Primer pair H3.X+Y (dark gray) specifically recognizes H3.X and H3.Y nucleotide sequences, whereas another primer pair is H3.X specific (H3.X, light gray). NIH3T3 mouse cDNA was used as a negative control. Data were normalized to HPRT1 and HMBS expression levels. Controls generated without reverse transcription were used to assess amplification threshold. Error bars represent SEM of two independent biological experiments. (C) Expression of H3.X and H3.Y mRNA in normal and malignant human tissues. Commercially available total RNA from human tissues was analyzed for H3.X+Y and H3.X expression in qPCR experiments and compared with results obtained with controls generated without reverse transcription. The number of samples that were positive for H3.X+Y and H3.X expression (+) are indicated in brackets.

observed in all samples except for the thalamus. In contrast to all other tissues tested, H3.X mRNA could also be detected in the brain and testis samples (Fig. 1 C).

In conclusion, mRNAs of the novel primate-specific H3 variants H3.X and H3.Y are not only present in human cell lines, but are also detected in primary cells from different human organs.

Tagged H3.X and H3.Y proteins are similar to canonical histone H3 in their nuclear localization, nucleosome structure, and exchange mobility kinetics

The detection of H3.X and H3.Y mRNA in human cells prompted us to investigate their protein products. We amplified their coding sequences from U2OS cDNA and cloned them under the constitutively active cytomegalovirus promoter in frame with an N-terminal HA or GFP tag. These constructs

were transfected into HeLa cells and, in the case of HA-tagged variants, at least two independent stable cell populations were selected. In both HA-H3.X- and HA-H3.Y-expressing HeLa cells, nuclear HA signals in interphase and staining of condensed chromatin in mitotic cells were observed by immunofluorescence (IF) microscopy (Fig. 2 A). Signals were also present on metaphase chromosome spreads (Fig. 2 B), which suggests that these novel tagged H3 variants are stably incorporated into chromatin and constitute core nucleosomal components. Next, we asked if the amino acid differences between H3.X, H3.Y, and other H3 variants have any impact on nucleosome structure. Thus, we calculated in silico models of H3.X and H3.Y protein structures, using the published structure of Xenopus laevis H3.2 as a template (Davey et al., 2002). With these, we computationally assembled a nucleosome containing H3.X or H3.Y. Our data show only few differences to the published structure of H3.2, with an exception being the unusual C terminus of H3.X

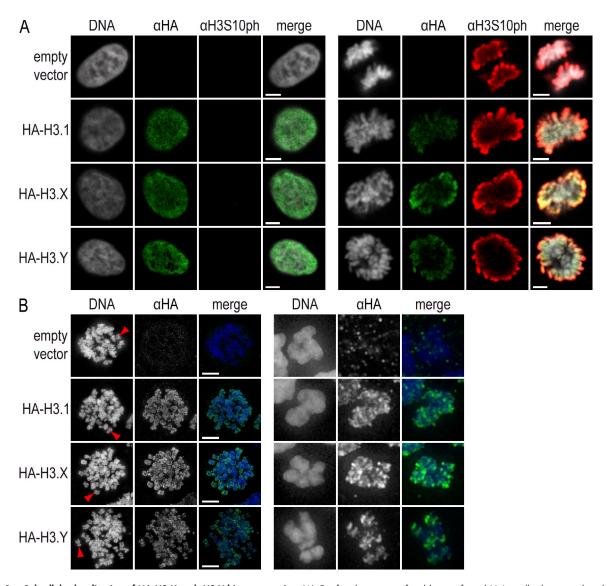


Figure 2. **Subcellular localization of HA-H3.X and -H3.Y histone proteins.** (A) Confocal imaging of stably transfected HeLa cells shows nuclear localization of HA-H3.1, -H3.X, and -H3.Y proteins. Cells were costained with TO-PRO3 (DNA, gray, left), α-HA (green, middle left), and α-H3S10ph (red, middle right). Overlay is shown on the right (merge). The left side shows interphase and the right side shows mitotic cells, as indicated by α-H3S10ph staining. (B) Metaphase spreads of mitotically arrested HeLa cells transfected with empty vector and HA-H3.1, -H3.X, and -H3.Y. Deconvolved wide-field images of chromosomes costained with DAPI (DNA, blue, left) and α-HA (green, middle). Sections containing one chromosome (red arrowhead) stained with α-HA are depicted on the far right. All HA-H3 variants are incorporated into chromosomes. Bars: (A) 5 μm; (B) 10 μm.

(Fig. 3 A). Surprisingly, this extended C-terminal tail of H3.X is proposed to fold into an additional α -helix that fits into the free space of the nucleosomal core structure (Fig. 3 B, left).

Because our in silico models proposed no dramatic changes in nucleosomal structure for H3.Y, but some differences for H3.X, we asked whether H3.X and/or H3.Y are incorporated into nucleosomes in a cellular context and whether they exhibit a similar stability compared with nucleosomes containing H3.1, H3.2, or H3.3. Immunoprecipitation (IP) of mononucleosomes generated from HeLa cells expressing HA-tagged H3 variants using an α -HA antibody confirmed the incorporation of HA-H3.X and -H3.Y into chromatin (Figs. 3 C and S2 A). Additionally, silver staining and immunoblots revealed that HA-H3.X and -H3.Y are present in nucleosomes containing all four core histones. Interestingly, HA-H3.X and -H3.Y containing mononucleosomes also contained endogenous H3 (Fig. 3 C, right), which suggests

that, unlike H3.1 and H3.3 (Tagami et al., 2004), these tagged novel variants form heterotypic nucleosomes with regard to their variant composition. To assess the incorporation stability of these novel variants, we performed FRAP experiments with GFP-tagged H3 variants transiently transfected into HeLa Kyoto cells, which were chosen because of their slow cell motility. These experiments demonstrate similar dissociation kinetics for the novel GFP-tagged H3 variants compared with GFP-H3.1 and -H3.3 (Fig. 3, D and E; and Fig. S2, B–D). GFP fluorescence recovery rates of all H3 variants were very slow (>8 h; Fig. S2, B–D) in comparison to GFP alone (a few seconds; Fig. 3, D and E), which is indicative of stable chromatin-bound immobile populations.

In summary, tagged H3.X and H3.Y protein containing nucleosomes are similar in their predicted structure and indistinguishable from nucleosomes containing H3.1 or H3.3 in their FRAP mobility and chromatin-association behavior.

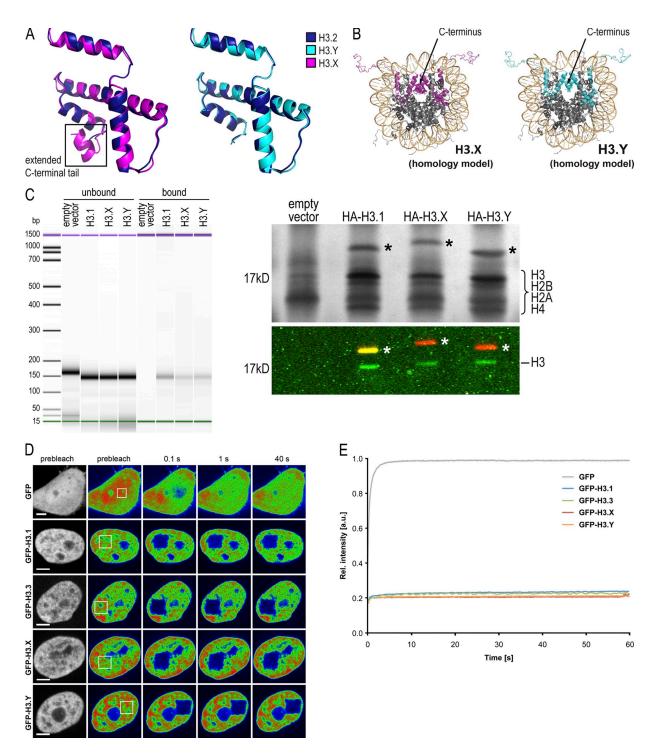


Figure 3. **Structure and stability of H3.X- and H3.Y-containing nucleosomes.** (A) In silico homology model of H3.X (purple, left) and H3.Y (light blue, right) protein structures in overlay with the crystal structure of H3.2 (dark blue). (B) Crystal structure of nucleosome with H3.2 exchanged by in silico homology models of H3.X (purple, left) and H3.Y (light blue, right), respectively. (C) IP of mononucleosomes generated from HeLa cells transfected with empty vector, HA.H3.1, H3.X, and H3.Y shows incorporation of novel H3 variants into nucleosomes. Bioanalyzer evaluation of purified DNA after IP of MNase-treated chromatin (unbound and bound material) shows digestion of chromatin to mononucleosomes and their successful precipitation (left; see also Fig. S2 A for DNA size and quality). Silver stain of 15% SDS-PAGE with α-HA IPs of mononucleosomes revealed successful binding of HA-tagged H3 variants (asterisks) and pull-down of core histones (top, right). Immunoblot of immunoprecipitates with α-HA (red) and α-H3 C-terminal (green) antibodies visualized by the Odyssey infrared imaging system (bottom, right). Notice that endogenous H3 is coimmunoprecipitated with all H3 variants analyzed. (D) FRAP experiments to evaluate nucleosomal stability of novel H3 variants using spinning disk confocal microscopy. HeLa Kyoto cells were transiently transfected with GFP, GFP-H3.1, -H3.3, -H3.X, and -H3.Y constructs. A small nuclear area was photobleached (box) and the recovery of the fluorescent signal was monitored over 1 min and up to 8 h (see Fig. S2, B–D, for long-term FRAP). Depicted is a short-term FRAP series (selected time points are shown) of GFP-tagged H3 variants compared with GFP alone. Bar, 5 μm. (E) Quantification of short-term FRAP series (selected time points are shown) of GFP-tagged H3 variants compared with GFP alone. Bar, 5 μm. (E) Quantification of short-term FRAP experiment. Mean curves of 10–20 individual cells are shown. Standard deviations were very small (in the range of ± 0.02) and were omitted for cla

Endogenous H3.X/Y proteins are expressed in U2OS cells and localize to the nucleus

To characterize endogenous H3.X/Y proteins, we generated a monoclonal antibody against these variants (α -H3.X/Y; see Fig. 1 A, black line, for peptide sequence). α -H3.X/Y specificity was tested in diverse experiments using HeLa cells stably expressing HA-tagged H3 variants. Immunoblot analysis of acid-extracted histones from these cell lines revealed that α -H3.X/Y specifically recognizes HA-H3.X and -H3.Y, but not HA-H3.1, -H3.2, or -H3.3 (Fig. 4 A). In IF microscopy analyses with this antibody, HA-H3.X and -H3.Y showed an exclusive nuclear staining (Fig. S2 E), which could be confirmed to be specific for H3.X/Y by peptide competition assays (Fig. S2 F).

To analyze the expression of endogenous H3.X and H3.Y proteins, we isolated total histones from different cell lines. As a negative control, we used mouse NIH3T3 cells and, as a positive control, HeLa cells expressing HA-H3.X. Immunoblotting of these histones with α -H3.X/Y showed a faint band in the lane containing U2OS histones corresponding to the size of H3, which would be expected for H3.Y (Fig. 4 B). These data are consistent with our previous finding that U2OS cells express mostly H3.Y mRNA (Fig. 1 B). Next, we addressed the subcellular localization of endogenous H3.X and/or H3.Y in U2OS cells by confocal IF microscopy. α-H3.X/Y shows a faint and dotted staining associated with the nuclear rim in all cells, which might be due to cross-reactivity with nuclear pores. Interestingly, we also noticed that few U2OS cells showed a strong nuclear α-H3.X/Y staining, colocalizing with DNA (Fig. 4 C, arrowhead), which was not observed in cell lines of mouse or rat origin (Fig. 4 D). Some few mitotic U2OS cells, as well as metaphase chromosomes, also stained positive with α-H3.X/Y (Fig. 4, E and F, respectively), which suggested that the observed α-H3.X/Y signal marks a stable chromatin component. To shed light on the subnuclear localization of endogenous H3.X/Y in U2OS cells, we performed super-resolution (i.e., below the diffraction limit of conventional optical microscopy) imaging with 3D structured illumination microscopy (3D-SIM; Gustafsson et al., 2008; Schermelleh et al., 2008). As depicted in Fig. 4 G, H3.X/Y is predominantly located outside of DAPI-dense regions, arguing for an association of endogenous H3.X/Y with less condensed, more euchromatic regions.

Nutritional- and growth-associated stress stimuli increase the number of H3.X/Y-expressing cells

Because only \sim 0.1% of U2OS cells showed a general nuclear staining with α -H3.X/Y, we wondered what cellular features distinguished these from other cells. We hypothesized that specific stress stimuli induce the nuclear α -H3.X/Y staining. To test this assumption, we induced DNA damage with different treatments (UV, etoposide, and hydroxyurea) but were unable to detect a significant increase in the percentage of cells positive for α -H3.X/Y staining (unpublished data). Next, we addressed the question of whether cell growth and nutrition could have an impact on H3.X/Y expression. Interestingly, induction of starvation and overgrowth (SO) in U2OS cells for 8 d, but not overgrowth alone (for details, see Materials and methods), led

to an increase in the percentage of cells with nuclear α -H3.X/Y staining (unpublished data). To evaluate if starvation alone, or in combination with high cell density, caused this increase in α-H3.X/Y-positive cell numbers, we either subjected U2OS cells to starvation or cultivated them under SO conditions. In three independent biological experiments, we quantified (1) the percentage of α-H3.X/Y nuclear-stained cells by IF microscopy, (2) H3.X/Y mRNA levels by qPCR, and (3) H3.X/Y protein levels by immunoblot analyses. Interestingly, we observed an approximately sixfold increase in the number of α -H3.X/Y nuclear-stained cells under SO conditions, whereas starvation alone only showed a minor effect (threefold increase; Fig. 5 A). In addition, we could reproducibly show, by qPCR (Fig. 5 B) and immunoblots (Fig. 5 C), that both H3.X/Y RNA and protein are up-regulated upon SO treatment compared with normally growing cells. We observed a stronger increase of H3.X/Ythan H3.X-specific amplicons (Fig. 5 B), which suggests that H3.Y mRNA expression is specifically affected by this particular stress condition. This was further validated by a "primerwalk" experiment with several different primer pairs that span the entire H3.X or H3.Y coding sequence (Fig. S3, A-C). Additionally, the induced protein band observed in immunoblots corresponds to a protein of the size of recombinant H3.Y but not H3.X (Fig. 5 C). To assess the question of which factors in the growth medium cause the observed effect, we also analyzed α-H3.X/Y nuclear stained cells after serum starvation and amino acid depletion. However, we did not observe any significant difference caused by both treatments (unpublished data).

These data show a direct correlation between α -H3.X/Y antibody staining and the increased mRNA and protein levels due to SO treatment. Furthermore, they suggest that H3.Y expression, in particular, is positively affected by these growth conditions.

Endogenous H3.Y protein is expressed after SO treatment

Our experiments indicate that H3.Y expression is mostly, if not exclusively, induced in U2OS cells under SO conditions. To confirm this observation, we purified histones from HEK293, NIH3T3, and SO-treated and normally grown U2OS cells by reversed-phase HPLC (RP-HPLC; Fig. S4 A). Fractions surrounding and including peaks corresponding to H3.2 and H3.3 (peak I) and H3.1 (peak II; Fig. 6 A) were analyzed by immunoblots with α-H3.X/Y. Specific signals were observed in fractions eluting at the beginning of peak I (region A bands) and between peak I and II (region B bands; Fig. 6 B). Interestingly, region A bands correspond to proteins of the molecular weight of histone H3, the predicted size of H3.Y. As expected for H3.Y, these bands are faintly visible in normally grown U2OS cells and strongly enhanced in those cultured under SO conditions. In contrast, no such bands are present in NIH3T3 and HEK293 cells. Region B bands are indicative of proteins of \sim 35 kD and are clearly visible without any change in intensity in all human cell lines, but not in mouse NIH3T3 cells. Fractions corresponding to region A and B bands from SO treated cells from two independent experiments and normally grown U2OS cells were analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS). For region B fractions, no conclusive information

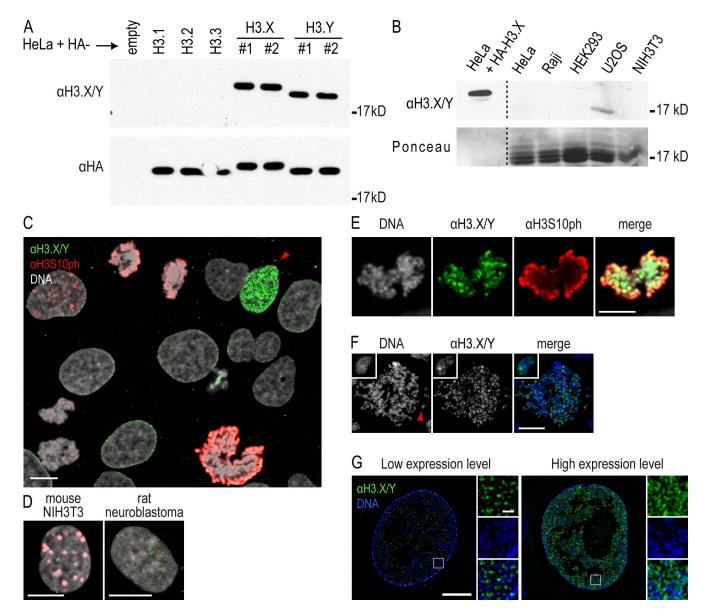


Figure 4. **Detection of endogenous H3.X and/or H3.Y proteins.** (A) A monoclonal antibody against H3.X/Y (α-H3.X/Y) was generated by immunizing rats with an N-terminal peptide specific for H3.X and H3.Y (aa 9–20, see also black line in Fig. 1 A). Immunoblots of acid-extracted histones from HeLa cells stably transfected with empty vector, or vectors containing HA-H3.1, -H3.2, -H3.3, -H3.X, and -H3.Y. Histones from two independently selected HeLa cell populations expressing HA-H3.X and -H3.Y were used (#1 and #2). (A, top) Staining of the membrane with α-H3.X/Y antibody shows only signals in lanes loaded with histones from HeLa cells expressing HA-H3.X and -H3.Y, but not in lanes containing general HA-H3 variants, which demonstrates the specificity of the antibody toward the novel variants in immunoblotting. (A, bottom) Equal loading was controlled by α-HA staining. Note that all bands run slower, as expected, because of the HA tag. HA-H3.X runs even slower than all other HA-H3 variants because of its extended C-terminal tail. (B) Immunoblot analysis of acid-extracted histones from different cell lines with α-H3.X/Y antibody. (B, top) Histones from HeLa cells expressing HA-H3.X served as positive control, and histones from mouse NIH3T3 cells served as a negative control. A faint signal in the lane containing U2OS histones can be seen. (B, bottom) The identical membrane was stained with Ponceau S solution before antibody incubation to control for protein loading. Dotted lines indicate that intervening lanes have been spliced out. (C) Confocal IF analysis of U2OS cells costained with α-H3.X/Y (green), α-H3S10ph (mitosis-specific, red), and TO-PRO3 (DNA, gray) as negative controls. (E) Confocal IF analysis of mitotic U2OS cells costained with α-H3.X/Y (green), α-H3S10ph (mitosis-specific, red), and TO-PRO3 (DNA, gray). (F) Costaining of metaphase chromosomes derived from mitotically arrested U2OS cells with α-H3.X/Y (green), α-H3S10ph (mitosis-specific, red), and TO-PRO3 (DNA, gray). (F) Costaining of metaphase chro

about the protein identity could be obtained. Only one single peptide corresponding to H3.X/Y (aa 53–63) could be observed in B-type fractions, rendering the analysis of these particular fractions inconclusive. To eliminate the possibility that region B bands constitute a dimeric aggregate caused by formation of disulfide bonds between cysteines, we treated these fractions

with increasing amounts of DTT (Fig. S4 B). In contrast to H3.1, H3.2, and H3.3, we did not observe a change in molecular weight for region B bands, which suggests that these particular proteins are not dimers of H3.X and/or H3.Y.

Interestingly, in region A fractions, we were able to identify two peptides specific for H3.X/Y (aa 53–63 and 73–83), in

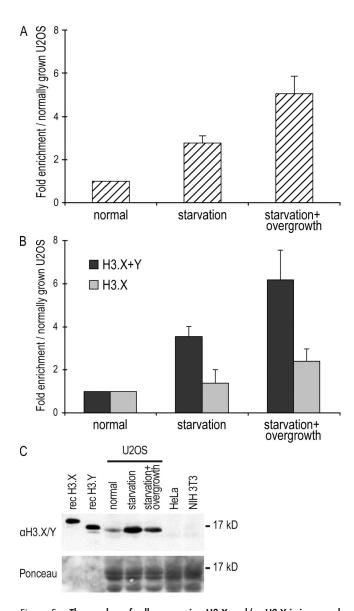


Figure 5. The number of cells expressing H3.X and/or H3.Y is increased by nutritional and proliferative stress. (A) Quantification of the percentage of U2OS cells positive for α -H3.X/Y nuclear staining by IF analyses under different growth conditions. U2OS cells were stained with α -H3.X/Y after 8 d of growth under different conditions: normal, starvation, or SO. The percentage of cells positive for α-H3.X/Y nuclear IF staining was determined (see Materials and methods for details) and plotted. A clear increase in the number of cells containing α -H3.X/Y-positive nuclei can be seen after growth under SO conditions. Error bars represent SEM of three independent biological experiments. (B) qPCR analysis of H3.X and H3.X+Y mRNA expression levels from U2OS cells grown under different conditions. From the same plates described in Fig. $\bar{5}$ A, cells were harvested, RNA was isolated, and cDNA was generated. A clear increase in H3.X+Y (dark gray) but not H3.X mRNA (light gray) under SO conditions can be detected, which is similar to the data obtained with IF analyses (Fig. 5 A). Error bars represent SEM of three independent biological experiments. (C) Immunoblot analysis of H3.X and/or H3.Y proteins isolated from U2OS cells grown under different conditions. From the same plates described in Fig. 5 A, cells were harvested, and histones were acid extracted and immunoblotted with α -H3.X/Y antibody (top). A clear increase in a 17-kD signal can be seen in U2OS cells grown under different stress conditions. Recombinant H3.X and H3.Y proteins serve as positive controls and histones from human HeLa and mouse NIH3T3 cells serve as negative controls. Staining of the same membrane with Ponceau S solution before antibody incubation was performed to ensure similar loading (bottom). One representative blot from three independent biological experiments is shown.

addition to the expected general H3 peptides (Figs. 6 C and S4, C and D). Furthermore, one peptide exclusively present in H3.Y but not H3.X (aa 18–28) was repeatedly detected (Figs. 6 C and S4 E), leading to a combined sequence coverage of 33% for H3.Y. Interestingly, in conjunction with unmodified H3.Y peptide, we were also able to identify acetylation of lysines 18, 23, and 27 (Fig. S4, F and G). These PTMs are also present in the known H3 variants, indicating that H3.Y was a likely part of a nucleosome and present in chromatin fibers.

In summary, these data show that posttranslationally modified H3.Y protein is expressed in vivo, and that under stress-inducing conditions, more H3.Y-expressing cells are present.

Knockdown of H3.Y affects expression of cell cycle-related genes and cell growth

Because we could clearly demonstrate the presence of endogenous H3.Y protein in human cells, we wondered about the biological consequences of its expression. We hypothesized that the natural substitution of amino acids S10, K14, S28, and K79 in H3.Y and its preferential exclusion from DAPI-dense regions (Fig. 4 G) might result in gene expression changes at the sites of chromatin incorporation. Therefore, we used global expression arrays to examine a potential impact on the regulation of gene expression by H3.Y. We used RNAi to significantly reduce H3.Y and H3.X+H3.Y mRNAs and proteins. siRNAs directed against H3.Y, H3.X+Y, and luciferase, as control, were designed and their specificity verified by transfection of HeLa cells expressing HA-tagged H3.X and H3.Y, followed by IF analysis (Fig. 7 A). Next, we sought to quantitatively compare global RNA expression profiles after RNAi in SO-treated U2OS cells using Affymetrix Human Gene 1.0 ST Arrays (performed twice with independent transfections on different days). We verified the proper knockdown of H3.Y (and H3.X) mRNAs by qPCR (Fig. 7 B) and observed a modest but significant deregulation of genes when compared with control luciferase RNAi in global microarray analyses (Fig. 7, C and D; and Tables S1 and S2). Cells transfected with H3.Y-specific siRNAs showed 293 genes up-regulated and 974 genes down-regulated. Cells treated with H3.X+Y-specific siRNAs had 1,106 genes up-regulated and 1,249 genes down-regulated (local false discovery rate cutoff of 0.2). To eliminate off-target and sole H3.X-specific effects and to focus only on H3.Y-specific responses, Venn diagrams of overlapping deregulated genes between H3.Y- and H3.X+Yspecific siRNA changes were used. They revealed up-regulation of 73 and down-regulation of 229 shared genes (Fig. 7 C). Gene ontology (GO) enrichment analysis of the deregulated genes classified shared down-regulated genes to mainly belong to cell cyclecontrolling pathways, but also chromatin organization and metabolic pathways, whereas only few GO groups of shared up-regulated genes could be determined (Tables S1 and S2). Because the majority of genes affected by H3.Y RNAi play a role in cell cycle control (mitosis), as shown by a simplified GO enrichment analysis (Fig. 7 D), and we previously noticed a reduction in cell number after H3.Y and H3.X+Y knockdown, we quantitatively determined cell growth of U2OS cells with reduced H3.Y mRNA and protein levels. Therefore, we transfected U2OS cells with different variantspecific siRNAs and quantitatively monitored cell growth for 96 h using the xCELLigence system (Roche). Interestingly, and in

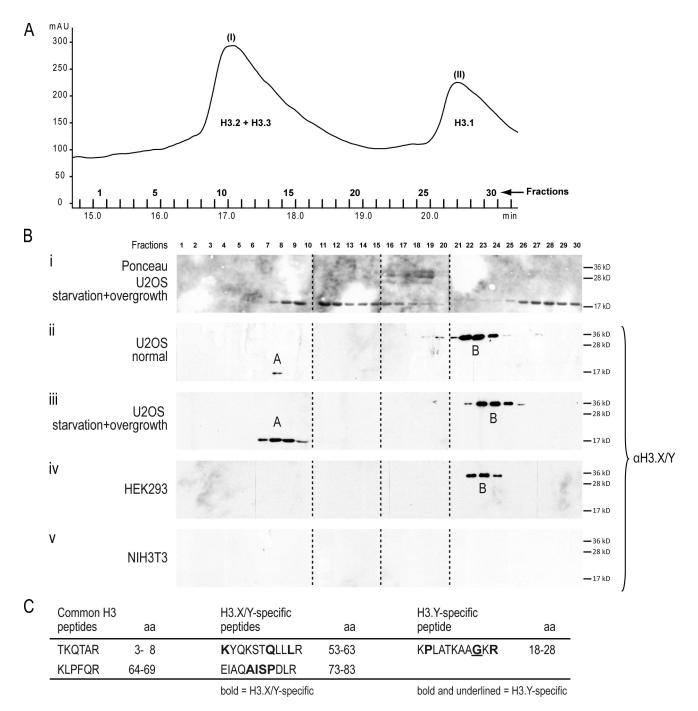


Figure 6. **Purification and identification of endogenous H3.Y variant protein.** (A) RP-HPLC section showing histone H3 peaks (see Fig. S4 A for complete RP-HPLC profile). Acid-extracted histones from starved and overgrown U2OS cells (Fig. 5) were separated by RP-HPLC, and histone H3 peaks (peak I, H3.2+H3.3; peak II, H3.1) spanning fractions 1-31 are shown. (B) Immunoblotting analyses of RP-HPLC fractions 1-30 spanning histone H3 peaks I to II from different cell lines under distinct growth conditions with α -H3.X/Y antibody. (i) Ponceau S staining of membrane containing histone fractions from starved and overgrown U2OS cells to detect H3.1, H3.2, and H3.3 proteins. Immunoblots incubated with α -H3.X/Y from RP-HPLC fractions from U2OS cells (iii), starved and overgrown U2OS cells (iii), HEK293 cells (iv), and mouse NIH3T3 cells (y). Dotted lines indicate that intervening lanes have been spliced out. The two anti-H3.X/Y-positive fractions are indicated with α -nd B. Proteins of both fractions were independently subjected to MS/MS analyses. (C) List of H3-, H3.X/Y, and H3.Y-specific peptides identified by LC-MS/MS from combined band A-corresponding fractions from U2OS cells (normal and SO treated; see iii). Amino acids highlighted in bold are specific for H3.X and H3.Y; a bold and underlined amino acid is found only in H3.Y.

accordance with our global transcriptome analyses, knockdown of H3.Y and H3.X+Y, but not of control RNAi (luciferase), resulted in a significant reduction of cell growth (Fig. 7 E).

These results suggest that loss of H3.Y especially affects the expression of genes involved in cell cycle control, leading to diminished cell growth. Neuronal cell subpopulations in human hippocampus express novel H3 variant proteins

Encouraged by our findings that endogenous H3.Y protein exists in vivo and plays an essential role in cell growth and gene regulation, we wondered if H3.Y and/or H3.X proteins

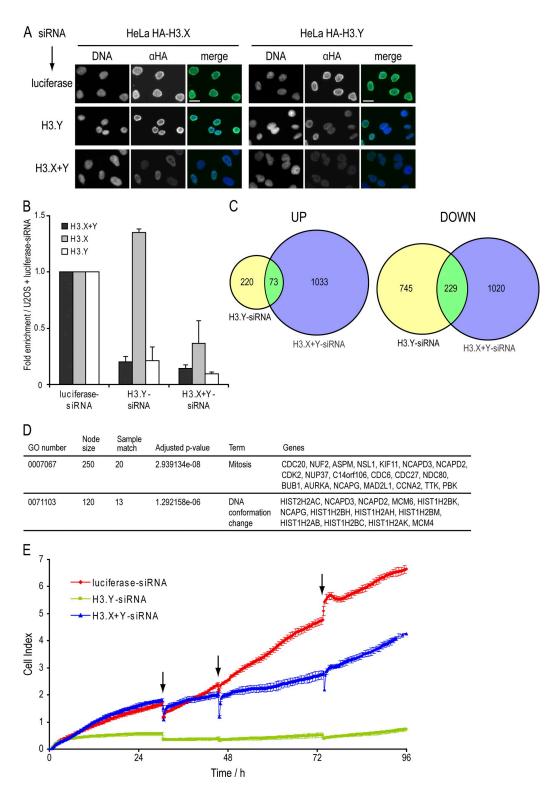


Figure 7. Influence of H3.Y expression on global gene regulation and cell growth. (A) Specificity determination of siRNAs against novel variants. IF microscopy using α -HA (green) and DAPI (blue) staining of HeLa cells expressing HA-H3.X and -H3.Y 4 d after RNAi treatment with indicated siRNAs. Bar, 20 μ m. (B) qPCR analysis to verify efficient H3.X and H3.Y RNAi knockdown before global transcriptome analysis. Primer pair H3.X+Y (dark gray) specifically recognizes H3.X and H3.Y nucleotide sequences, whereas two other primer pairs are H3.X- (light gray) or H3.Y-specific (white). Data were normalized to HPRT1 and HMBS expression levels and depict fold enrichment of expression in comparison to luciferase control RNAi. Controls generated without reverse transcriptase were used to assess amplification threshold. Error bars represent SEM of two independent biological experiments. (C) Venn diagrams of genes deregulated after H3.X+Y (blue) and H3.Y (yellow) RNAi in SO-treated U2OS cells, as identified by microarray analyses of two independent experiments when compared with luciferase control knockdown. Digits indicate numbers of genes significantly up- (left) or down-regulated (right) in comparison to luciferase control knockdown. (D) Simplified GO analysis of overlapping genes after H3.X+Y and H3.Y knockdown. Detailed GO lists are shown in Tables S1 and S2. Node size = total number of genes analyzed in this node (GO term/group). (E) Growth curve of U2OS cells after RNAi (red, luciferase control siRNA; blue, H3.X+Y siRNA; green, H3.Y siRNA). Arrows mark changes of growth medium.

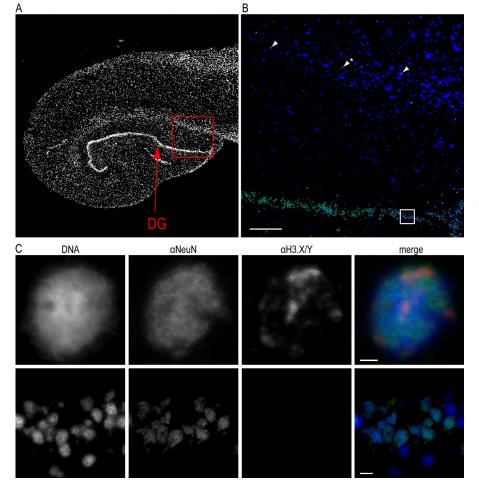


Figure 8. H3.X/Y protein expression in human brain. (A) Overview IF picture of commercially available human hippocampus section stained with DAPI (DNA, gray). (B) Human hippocampus sections were costained with α -H3.X/Y (red), α -NeuN (neuronal marker, green), and DAPI (DNA, blue). The boxed section from A is shown. Arrows indicate neuronal cells with positive α -H3.X/Y staining in the region above DG. One out of three representative stainings is shown. (C) The boxed section and the α -H3.X/Y-positive cell marked with an asterisk in B are shown in higher resolution. Costainings with astrocyte marker antibody (α -GFAP) are shown in Fig. S5. Bars: (A and B) 200 μm; (C, top) 2 μm; (C, bottom) 10 μm.

are expressed in human tissue. Based on our qPCR results (Fig. 1 C), and the observation that many primate-specific genes are expressed in the brain and reproductive organs (Tay et al., 2009), we chose to analyze sections of human hippocampus in IF (Fig. 8 A).

Interestingly, some few cells in the region above the dentate gyrus (DG) of the hippocampus stained positive with α -H3.X/Y (Fig. 8 B). We determined their cellular origin by costaining with antibodies against neurons (α neuronal nuclei [α -NeuN] and astrocytes (α glial fibrillary acidic protein [α -GFAP]; Fig. S5). H3.X/Y proteins were identified in a subpopulation of neurons outside of the DG (Fig. 8, B and C). This finding is in accordance with our observation that only few U2OS cells express large amounts of H3.Y. Surprisingly, α -H3.X/Y staining of neurons displayed an enrichment in certain chromatin areas, different from that observed in U2OS cells (Fig. 4, C and G).

In conclusion, H3.X and/or H3.Y proteins are expressed in a subpopulation of neurons in the human hippocampus, highlighting the possibility that these novel histone variants might have cell type–specific functions.

Collectively, our data demonstrate that the novel histone variant H3.Y is expressed in human cells and specialized tissues, is incorporated into chromatin, is posttranslationally modified, and impacts the regulation of many genes implicated in cell cycle progression.

Discussion

More than 20 yr ago, the mammalian histone H3 variants CENP-A (Palmer et al., 1987) and tH3 (Trostle-Weige et al., 1984) were found. Since then several variants of the mammalian H2A family, such as macroH2A (Pehrson and Fried, 1992) and H2A.Bbd (Chadwick and Willard, 2001), have been described, but it was long thought that all members of the histone H3 family had been identified. To better understand the role of H3 variants in chromatin-related processes, we set out to investigate if there might be even more, yet unknown, H3 variants present in mammals.

H3.X and H3.Y are primate-specific histone genes

Using the DNA sequence of human histone *H3.1* in genomic searches, we have now identified two novel intron-free histone H3 variant genes on human chromosome 5, which we named *H3.X* and *H3.Y*. Sequences for *H3.X* and *H3.Y* genes were, in addition to humans, only found in chimpanzee and macaque primate genomes; sequence information of other primate genomes is unfortunately not yet available. Additional Basic Local Alignment Search Tool (BLAST) searches in mouse, rat, and other genomes did not yield any positive hits, which suggests that *H3.X* and *H3.Y* genes are present in primates, but not in other mammals or even lower eukaryotes. Both genes might be duplications

from one ancestral gene, as their sequences are highly conserved even in their 5' and 3' untranslated regions, with only some differences in their putative promoter sequences. Phylogenetic analyses of H3.X and H3.Y coding and 3' genomic sequences revealed a higher sequence homology to H3.3 genes than to other H3 variants, leading to the speculation that H3.X and H3.Y might be evolutionary derivatives of H3.3, although they do not contain any introns. To our knowledge, these are the first histone genes that are primate specific outside of testes and likely confer specialized chromatin functions unique to these higher vertebrates.

H3.X and H3.Y genes are both transcribed, but only H3.Y protein can be detected in vivo

So far we have obtained evidence that H3.X and especially H3.Y mRNAs are present at low but significant amounts in some human bone, breast, lung, and ovary tumor tissues, as well as in testis and certain areas of the brain. Therefore, their expression is not only found in transformed cell lines (U2OS) but also in primary human tissues. Surprisingly, we were able to verify only the existence of H3.Y but not H3.X protein in human cell lines thus far. One possibility is that alternative transcription of H3.X may occur. Two previous annotations of the H3.X locus in the NCBI database predicted several introns and different splice sites (Fig. S3 D); however, our primer walk experiment and negative qPCR data using primers specific for these splice forms (unpublished data) provided no evidence that alternative H3.X transcripts exist. Based on our results, H3.X is transcribed only at low levels in U2OS cells, and because no protein could be detected thus far, it is feasible that this particular variant is a pseudogene without any functional consequences. Alternatively and more interestingly, it is also possible that its mRNA, although it is present only at low levels in some cells, may have unknown regulatory functions and may even be influencing H3.Y gene expression. Future studies will hopefully shed more light onto a putative function of H3.X mRNA.

The number of H3.Y-expressing U2OS cells is increased by stress conditions

We showed that H3.Y mRNA and protein is present in some U2OS cells (\sim 0.1%) and that the number of cells expressing high levels of this novel histone variant can be increased by a stress response that involves nutritional starvation in combination with high cellular density. It is interesting to note that when calculating the amount of endogenous H3.Y mRNA per cell, we observe an ~40-fold higher expression compared with a cell expressing HA-H3.Y mRNA. This suggests that once a cell receives a specific signal to start transcribing endogenous H3.Y or reaches a certain threshold of an unknown factor, this variant is expressed in extremely high amounts. It will be crucial to identify the physiological trigger of H3.Y expression in future studies to learn more about H3.Y's cellular function.

Loss of H3.Y expression impacts transcriptional regulation and cell growth To get a first glimpse on H3.Y's potential impact on gene regulation, we knocked down H3.Y in U2OS cells using RNAi.

Our global microarray data imply that H3.Y influences the transcription of several genes, either directly or indirectly. Interestingly, loss of H3.Y led to more genes being down-rather than up-regulated, which suggests that H3.Y might be involved in transcriptional activation of some genes. This is in accordance with our finding that H3.Y mainly localizes to euchromatic regions (Fig. 4 G). The majority of genes affected by H3.Y reduction are components of cell cycle- and chromatin structure-regulating pathways, leading to a significant impairment of cell growth. It is difficult to unequivocally assign a gene or group of genes as direct targets of H3.Y depletion, which would require the identification of genomic regions containing H3.Y on a global scale by chromatin IP followed by chip hybridization or sequencing. Unfortunately, IP of endogenous H3.Y protein with α -H3.X/Y antibody turned out to be insufficient with standard protocols, and therefore more time and testing will be required to work out a functional ChIP protocol. It was surprising to us that the whole U2OS cell population was equally impaired in its cell growth after H3.Y knockdown although only few cells appeared to strongly express H3.Y protein, as shown by IF quantification (Figs. 4 C and 5 A). It is therefore possible that all U2OS cells express a minimal amount of H3.Y RNA and protein, which is beyond our detection limit, but nevertheless crucial for the cell to traverse through the cell cycle. To our knowledge, only one other histone variant showed a similar effect on the expression of cell cycle-related genes after RNAi depletion: knockdown of the linker histone variant H1.2 altered the expression of \sim 2% of genes genome wide (most of them being repressed), including a relevant proportion of cell cycle-related genes (Sancho et al., 2008). It is therefore not completely unlikely that H3.Y plays an important role in the regulation of genes involved in cell cycle control. It is tempting to speculate that the stress-related appearance of more U2OS cells expressing H3.Y is caused by continued cell proliferation or the prevention of cell death. One interesting observation was that knockdown of H3.Y led to stronger cell growth defects than combined H3.X+H3.Y RNAi. Because we repeatedly saw an increase of H3.X mRNA after H3.Y depletion, one speculative idea would be that H3.X and H3.Y are connected by some sort of regulatory feedback mechanism. It will be of great interest in future studies to shed light on the regulatory expression pathways of these novel variants, particularly in the context of cellular stress and their existence in neuronal subpopulations of the human hippocampus.

Materials and methods

Cell culture, cloning of H3.X and H3.Y expression constructs,

Human HeLa, HeLa Kyoto, HEK293, U2OS, mouse NIH3T3, and rat neuroblastoma cells were grown in DME medium (PAA) supplemented with 10% FCS (Sigma-Aldrich) and 1% penicillin/streptomycin (PAA) at 37°C and 5% CO₂. Human Raji cells were cultured in RPMI-1640 medium (Invitrogen) supplemented with 10% FCS (Sigma-Aldrich) and 1% penicillin/ streptomycin (PAA) at 37°C and 5% CO_2 . H3.1, H3.2, H3.3, H3.Y, and H3.X cDNA were cloned into pEGFP-C1 (Takara Bio Inc.) and pIRES-neo vector (Takara Bio Inc.) containing an HA tag (a gift from H. Dormann, Rockefeller University, New York, NY) to generate N-terminally tagged H3 variants after transfection into cell lines. For simplicity, we will refer to the EGFP-tagged constructs as GFP-tagged throughout the text. All constructs were sequenced (MWG-Biotech AG) to verify cloning and amplification accuracy.

Transfections were performed with FuGene HD (Roche) according to the manufacturer's instructions. In brief, cells were incubated with the transfection complex for 1 d, after which selection medium containing 600 μ g/ml G-418 sulfate (PAA) was added. Stable cell lines were maintained in medium containing 400 μ g/ml G-418 sulfate.

Expression of recombinant H3.X and H3.Y proteins in Escherichia coli

BL21-CodonPlus (DE3)-RIL bacteria (Agilent Technologies) were transformed with pET-21a(+) plasmids (EMD) containing H3.X and H3.Y cDNA. Expression of recombinant proteins was induced by incubation with 0.1% IPTG (Carl Roth) overnight and controlled through SDS-PAGE analysis of boiled bacteria (Coomassie and immunoblot).

RNA expression analysis

Total RNA was isolated using the RNeasy mini kit (QIAGEN) according to the manufacturer's instructions. Genomic DNA contaminations were removed by performing an on-column DNase I digest (RNase-free DNase; QIAGEN). cDNA was synthesized using the ProtoScript First Strand cDNA Synthesis kit (New England Biolabs, Inc.), priming with random nonamers. aPCR analysis was performed on a LightCycler 480 (Roche) using the LightCycler 480 SYBR green I mastermix (Roche). Data analysis was performed with the advanced relative quantification tool of the LightCycler 480 software; results were normalized to HPRT1 and HMBS levels. H3.X and H3.X primer sequences (Sigma-Aldrich) are listed in Fig. S3 C. Total RNA from different human tissues was commercially acquired from Applied Biosystems (normal lung, breast and tumor breast, and lung and ovary) and BioChain (tumor lung, breast, thyroid and bone, normal testis, cerebellum, cerebral cortex, hippocampus, thalamus, and total fetal brain).

Histone extraction and mononucleosome IP

Histones were acid extracted as described previously (Shechter et al., 2007). In brief, nuclei were isolated by hypotonic lysis and extracted using 0.4 M sulfuric acid. Soluble histones were precipitated with trichloroacetic acid and resuspended in water.

Mononucleosomes for IP experiments were generated as described previously (Wysocka et al., 2001), with the following changes: For cell lysis, 0.1% NP-40 was used instead of Triton X-100. All centrifugation steps before MNase treatment were performed at 3,200 g. Mononucleosomes were generated by digestion of chromatin with 0.25 U MNase (Sigma-Aldrich) for 15 min in buffer A (10 mM Hepes, pH 7.9, 10 mM KCl, 1.5 mM MgCl₂, 0.34 M sucrose, 10% glycerol [vol/vol], 1 mM DTT, and protease inhibitor cocktail [Roche] plus 1 mM CaCl₂) and stopped by the addition of EGTA (final concentration of 2 mM). Centrifugation was performed at 20,000 g for 20 min. Supernatants of four MNase digests were combined, and salt concentration was adjusted to 150 mM KCl. Magnetic beads (Invitrogen) were washed four times with buffer C (20 mM Hepes/KOH, pH 7.9, 20% [vol/vol] glycerol, 0.2 mM EDTA, 0.2% [vol/vol] Triton X-100, 300 mM KCl, and protease inhibitor cocktail [Roche]). One part was evaluated for DNA size and quality using DNA 1000 reagents (Agilent Technologies) with the 2100 Bioanalyzer (Agilent Technologies), another part was analyzed by immunoblotting using the Odyssey infrared imaging system (LI-COR Biosciences) and evaluated with Odyssey Software Version 2.1 (LI-COR Biosciences), and the last part was analyzed by silver staining.

Antibodies

To generate an antibody against H3.X and H3.Y (α -H3.X/Y), a peptide spanning amino acids KATAWQAPRKLP of histone H3.X and H3.Y was synthesized (Peptide Specialty Laboratories GmbH) and coupled to BSA and ovalbumin, respectively. Rats were immunized subcutaneously and intraperitoneally with a mixture of 50 µg peptide-ovalbumin, 5 nmol CPG oligonucleotide (TIB MOLBIOL GmbH), 500 µl PBS, and 500 µl of incomplete Freund's adjuvant. A boost without adjuvant was given 6 wk after the primary injection. Fusion was performed using standard procedures (Köhler and Milstein, 1975). Supernatants were tested by differential ELISA with the histone peptide coupled to BSA and an irrelevant peptide coupled to the same carrier. Monoclonal antibodies that reacted specifically with the peptide were further analyzed in immunoblot and IF studies. α -H3.X/Y clone 8H6-2111 of rat IgG2a subclass was deployed in this study.

The following primary antibodies were used in this study: α -HA (clone 12CA5; Roche), α -H3S10ph (Millipore), α -H3 (C terminus; Abcam), α -NeuN (Millipore), and α -GFAP (Dako).

The following secondary antibodies were used: α -rat HRP and α -rabbit HRP (GE Healthcare); α -mouse–IRDye700DX and α -rabbit–IRDye800DX (Rockland Immunochemicals, Inc.); α -rat–Alexa Fluor 488, α -mouse–Alexa Fluor 488, and streptavidin-Alexa 555 (Invitrogen); and

 α -rabbit–Rhodamine red X and α -rabbit–Cy5 (Dianova, Inc.); and α -ratbiotin (Vector Laboratories).

IF microscopy

Adherent mammalian cells were grown on coverslips, washed, fixed in 1% or 3.7% formaldehyde–PBS solution, permeabilized with 0.1% or 0.5% Triton X-100–PBS solution, and blocked with 1% BSA in 0.1% or 0.5% Triton X-100–PBS solution. After stepwise incubation with primary and then secondary fluorescent antibody, cells were stained with DAPI (Invitrogen) and mounted with a ProLong Antifade kit (Invitrogen). Samples for 3D-SIM were prepared on precision cover glass No. 1.5 (thickness 0.170 \pm 0.005 mm; Carl Roth) and embedded in Vectashield mounting medium (Vector Laboratories). Chromosome spreads were generated as described previously (Hake et al., 2005). For confocal microscopy, samples were stained with To-Pro3 (Invitrogen) and mounted with Vectashield mounting medium (Vector Laboratories).

FRAP

For FRAP experiments, human HeLa Kyoto cells were transiently transfected with GFP-tagged histone H3 constructs. After 24 h, cells were seeded in Lab-Tek chamber slides (Thermo Fisher Scientific) and incubated for several hours or overnight before performing FRAP experiments. FRAP was performed on a spinning disc microscope (UltraVIEW VoX; PerkinElmer) with an integrated FRAP PhotoKinesis accessory (PerkinElmer) assembled to an Axio Observer D1 stand (Carl Zeiss, Inc.). The microscope was equipped with a heated environmental chamber set to 37°C and CO₂ perfusion set to 5%. For photobleaching experiments, several square bleach regions with a size of 5×5 µm were positioned on selected cell nuclei within the field of view. Photobleaching was performed using two iterations with the acoustooptical tunable filter of the 488-nm and the 514-nm laser line set to 100% transmission. To determine long-term recovery kinetics, 3D image stacks of 8-µm height and a z distance between image planes of 1 µm were recorded with an initial speed of 1 frame/min for the first 10 frames, followed by intervals of 10 min.

Quantitative evaluations were performed with ImageJ (http://rsb.info.nih.gov/ij/). Lateral and rotational movements of the cell nucleus were corrected by image registration using the StackReg plug-in of ImageJ. Mean intensities over time were extracted from the total nuclear area (7). The background ROI outside of the cell was defined manually from the initial field of view. The mean gray values over time were measured, background subtracted, and normalized to the respective means of the last prebleach values. The resulting postbleach B values were then divided by the respective T value to correct for the superimposed gain or loss of total fluorescence during postbleach acquisition, potentially caused by newly synthesized GFP-histones, bleaching-by-acquisition, and flux of residual fluorescence from above and below the recorded optical plane.

Peptide competition

α-H3.X/Y antibody was incubated for 4 h at 4°C with the following peptides (1 μg/ml) prior to addition to the fixed cells: biotin-coupled H3.3 unmodified, aa 22–41 (Proteomics Resource Center of the Rockefeller University); and H3.X/Y unmodified, aa 9–20 (Peptide Specialty Laboratories GmbH).

Immunohistochemistry

Commercially available frozen sections of human hippocampus (Biochain) were thawed and blocked with 0.5% BSA in 0.5% Triton X-100–PBS solution. After stepwise incubation with primary and then secondary fluorescent antibodies, cells were stained with DAPI (Invitrogen) and mounted with Aqua-Poly/Mount (Polysciences, Inc.). To enhance \(\alpha\text{-H3.X/Y}\) signals, samples were incubated with a biotin-coupled \(\alpha\text{-rat}\) antibody (Vector Laboratories), followed by streptavidin coupled to an Alexa 555 fluorophore (Invitrogen).

Microscopes

Unless stated otherwise, samples were kept at room temperature during image acquisition. Stained cells were analyzed on a microscope (Axiovert 200M) with an EC Plan-Neofluar 40×/1.3 NA oil Ph3 (differential interference contrast [DIC] III) objective (both from Carl Zeiss, Inc.). Images were processed with AxioVision software (Carl Zeiss, Inc.). Confocal imaging was performed with a confocal microscope (LSM 510 META; Carl Zeiss, Inc.) equipped with an argon-ion and two helium-ion lasers using a Plan-Apochromat 63×/1.4 NA oil DIC or a Plan-Neofluar 40×/1.3 NA oil DIC objective lens (both from Carl Zeiss, Inc.). Images were processed with LSM 510 META software (Carl Zeiss, Inc.). Images of metaphase chromosome spreads were acquired using a personalDV wide-field epifluorescence microscope (Applied Precision) equipped with a 60×/1.42 NA Plan-Apochromat oil objective lens (Olympus) and a CoolSNAP HQ²

charge-coupled device (CCD) camera (Photometrics). Typically, stacks with a z distance of 200 nm were recorded and then subjected to a constrained iterative deconvolution (enhanced ratio, 10 cycles, medium noise filtering) using the SoftWoRx 3.7 imaging software package (Applied Precision). For a description about the microscopic setup used for FRAP, see the description for in the FRAP section. Imaging with 3D-SIM was performed with an DeltaVision OMX v3 prototype (Applied Precision) equipped with highpower diode lasers with wavelengths of 405, 488 and 592.2 nm; a 100×/ 1.4 NA Plan-Apochromat oil objective lens (Olympus); and Cascade II:512 EM CCD cameras (Photometrics). Laser light is directed through a movable optical grating to generate a fine-striped illumination pattern on the sample plane. Image stacks with 15 images per plane (five phases, three angles) and a z distance of 125 nm were acquired and subjected to a computational reconstruction to obtain a high-resolution 3D dataset with a twofold enhanced optical resolution in all three special directions compared with conventional light microscopy (Gustafsson et al., 2008; Schermelleh et al., 2008). Immunohistochemistry samples were analyzed using a confocal LSM 710 microscope (Carl Zeiss, Inc.) equipped with three argon, one diode-pumped solid-state, one helium, and one diode laser. Images were acquired using an EC Plan-Neofluar 10x/0.3 NA M27, an LD C-Apochromat 40x/1.1 W Korr M27 (water), or a LCI Plan-Neofluar 63x/1.3-mm Korr Ph3 M27 (water) objective lens (all from Carl Zeiss, Inc.), and Image processing was performed with ZEN 2009 (Carl Zeiss, Inc.) software.

For downstream image processing of all microscope images, Photoshop (Adobe) software packages were used.

In silico modeling

To generate homology models from protein sequences, we used the I-TASSER server (Zhang, 2008). I-TASSER combines threading, assembly, and refinement to generate 3D models. For sequences of both H3.X and H3.Y, the server automatically selected as a template the chain A (histone H3.2), PDB accession no. 1kx5. This structure was obtained using X. laevis histones assembled into nucleosomes with 147 bp of DNA from Homo sapiens (Davey et al., 2002). The model for H3.X had a C score of 0.07, and the model for H3.Y had a C score of 1.50. In both cases, the second-best model had a considerably lower C score. Datasets of the models were downloaded and visualized by means of the Swiss-PdbViewer (Guex and Peitsch, 1997) and PyMOL (DeLano Scientific LLC; http://www.pymol.org).

RP-HPLC and MS/MS

Histones from U2OS, HEK293, and NIH3T3 cells were separated by RP-HPLC on a C4 column (250 \times 4.6-mm Jupiter, 10 μ m, 300 Å; Phenomenex) using a linear gradient of 35-53% solvent B (solvent A, 0.1% trifluoroacetic acid; solvent B, 99.92% acetonitrile and 0.08% trifluoroacetic acid) over 30 min at 1.0 ml/min on a Biotech Ettan microLC (GE Healthcare). The H3-containing fractions were dried under vacuum and stored at −20°C. For mass spectrometry analysis, RP-HPLC fractions containing histones were resolved in 0.1 M ammonium bicarbonate and digested. Proteolytic digestions were performed overnight at 37°C with sequencing grade trypsin (Promega). In this case, histones were chemically modified beforehand by treatment with propionic anhydride (Merck) to convert free amino groups of lysine residues to propionic amides. After digestion, samples were directly loaded onto a nano electrospray ionization (ESI) LC-MS/MS for protein identification. Each sample was first separated on a C18 reversed phase column via a linear acetonitrile (Sigma-Aldrich) gradient (UltiMate 3000 system [Dionex] and an LC Packings column; 75 µm inner diameter × 15 cm, packed with C18 PepMap, 3 µm, 100 Å) before MS and MS/MS spectra were recorded on an Orbitrap mass spectrometer (Thermo Fisher Scientific). The resulting data were analyzed via the Mascot Software (Matrix Science) using a home-made database that contains the diverse H3 sequences. Fragment spectra were also interpreted manually.

RNAi and growth curve

siRNAs were designed using siDESIGN Center (http://www.dharmacon.com/designcenter/designcenterpage.aspx; Thermo Fisher Scientific) and synthesized (MWG-Biotech AG). siRNAs have been prevalidated to confirm their targeting specificity to H3.Y or H3.X+H3.Y and to reduce the chance of off-target effects by BLAST searches (NCBI). The following double-stranded siRNAs, which had differences of at least three nucleotides from other targets, were used. Luciferase, 5'-CUJACGCUGAGUACUUCGA-3'; H3.Y, 5'-CCGCAGAGAGGUCCUUAA-3'; H3.X+Y, 5'-GCGGGAAUCCAGAAAGUAC-3'. Cells were transfect dwice with siRNAs using oldectamine (Invitrogen) according to the manufacturer's instructions. 4 d after transfection, cells were used for various assays. Determination of cell growth after RNAi was performed using xCELLigence system (Roche). Cells were seeded in E plates 1 d after the second siRNA transfection. Cell growth was

monitored over the next $4\,\mathrm{d}$. To eliminate side effects caused by the lack of nutrients or the accumulation of waste products, growth medium was exchanged every day.

Microarray hybridization

Total RNA preparations were further purified with the RNeasy MinElute Cleanup kit (QIAGEN). 100 ng of RNA were used as starting material for all target preparations. RNA amplification, labeling, and hybridization to Human Gene 1.0 ST Arrays (Affymetrix) were performed according to the Human Gene 1.0 ST Array kit protocol (Affymetrix). The raw microarray data were processed in R/Bioconductor (http://www.bioconductor .org/) as follows: gene-based expression values were calculated using the robust multichip average (RMA) method provided by the oligo package. Genes that had a log2 expression value of at least 4 in at least one of the treatment conditions were kept for downstream analyses. Differential expression estimation was based on a moderated t statistic (limma package) with subsequent calculation of the local false discovery rate (lfdr; locfdr package). Genes were classified as responders by an lfdr cutoff of 0.2. GO enrichment analysis was performed using a hypergeometric distribution test and subsequent Bonferroni correction as supplied by the GOHyperGALL script (http://faculty.ucr.edu/~tgirke/Documents/R_BioCond/My_R_Scripts/ GOHyperGAll.txt). We reduced term redundancy by applying the GOHyperGAll_Simplify function with a P-value cutoff of 0.001.

Online supplemental material

Fig. S1 shows H3.X and H3.Y alignments, evolutionary origin, and expression level determination. Fig. S2 shows analysis of H3.X- and H3.Y-containing nucleosomes, and α-H3.X/Y antibody specificity determination. Fig. S3 shows evaluation of human H3.X sequences and inducible endogenous H3.X and H3.Y gene expression. Fig. S4 shows purification and mass spectrometrical identification of endogenous H3.Y protein. Fig. S5 shows H3.X/Y expression in the human hippocampus. Table S1 shows a GO list with shared up-regulated genes after H3.Y knockdown. Table S2 shows a GO list with shared down-regulated genes after H3.Y knockdown. Online supplemental material is available at http://www.jcb.org/cgi/content/full/jcb.201002043/DC1.

We are grateful to Ignasi Forne for technical advice concerning RP-HPLC, and Dietmar Martin and Kerstin Maier of the Gene Center Affymetrix Microarray Platform (Cramer Laboratory) for help with microarray experiments. We thank Axel Imhof for his remarks concerning MS data, Robert Loewe for his valuable advice about statistical analysis of qPCR data, and Holger Dormann for the pIRES-neo-HA plasmid. We are indebted to Katrin Schneider for help with FRAP analysis, Mariacristina Chioda and Sandra Vengadasalam for help with the confocal microscope, and Christiane Simon for immunohistochemistry advice. We thank Dirk Eick, Martin Heidemann, and Corinna Hintermair for their help with the xCELLigence instrument. We appreciate Antonia Jack's, Christian Janzen's, and Emily Bernstein's critical examination of the manuscript. We thank especially the Hake laboratory, members of the Adolf-Butenand+Institute, and Peter Becker, Stefan Jentsch, and Gunnar Schotta for constructive discussions.

This work was supported by the Deutsche Forschungsgemeinschaft (part of the Sonderforschungsbereich Chromatin Transregio 5) to S.B. Hake, H. Leonhardt, L. Schermelleh, and E. Kremmer; and by the Center for Integrated Protein Science Munich to S.B. Hake and H. Leonhardt. S.M. Wiedemann is a fellow of the Munich International Max Planck Research School for Molecular and Cellular Life Sciences (IMPRS-LS).

Submitted: 8 February 2010 Accepted: 11 August 2010

References

Abbott, D.W., V.S. Ivanova, X. Wang, W.M. Bonner, and J. Ausió. 2001. Characterization of the stability and folding of H2A.Z chromatin particles: implications for transcriptional activation. J. Biol. Chem. 276:41945–41949. doi:10.1074/jbc.M108217200

Ahmad, K., and S. Henikoff. 2002a. Histone H3 variants specify modes of chromatin assembly. *Proc. Natl. Acad. Sci. USA*. 99:16477–16484. doi:10.1073/pnas.172403699

Ahmad, K., and S. Henikoff. 2002b. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol. Cell*. 9:1191–1200. doi:10.1016/S1097-2765(02)00542-7

Allshire, R.C., and G.H. Karpen. 2008. Epigenetic regulation of centromeric chromatin: old dogs, new tricks? Nat. Rev. Genet. 9:923–937. doi:10.1038/nrg2466

- Angelov, D., A. Molla, P.Y. Perche, F. Hans, J. Côté, S. Khochbin, P. Bouvet, and S. Dimitrov. 2003. The histone variant macroH2A interferes with transcription factor binding and SWI/SNF nucleosome remodeling. *Mol. Cell*. 11:1033–1041. doi:10.1016/S1097-2765(03)00100-X
- Bao, Y., K. Konesky, Y.J. Park, S. Rosu, P.N. Dyer, D. Rangasamy, D.J. Tremethick, P.J. Laybourn, and K. Luger. 2004. Nucleosomes containing the histone variant H2A.Bbd organize only 118 base pairs of DNA. EMBO J. 23:3314–3324. doi:10.1038/sj.emboj.7600316
- Bernstein, E., and S.B. Hake. 2006. The nucleosome: a little variation goes a long way. *Biochem. Cell Biol.* 84:505–517. doi:10.1139/O06-085
- Black, B.E., D.R. Foltz, S. Chakravarthy, K. Luger, V.L. Woods Jr., and D.W. Cleveland. 2004. Structural determinants for generating centromeric chromatin. *Nature*. 430:578–582. doi:10.1038/nature02766
- Bönisch, C., S.M. Nieratschker, N.K. Orfanos, and S.B. Hake. 2008. Chromatin proteomics and epigenetic regulatory circuits. *Expert Rev. Proteomics*. 5:105–119. doi:10.1586/14789450.5.1.105
- Chadwick, B.P., and H.F. Willard. 2001. A novel chromatin protein, distantly related to histone H2A, is largely excluded from the inactive X chromosome. J. Cell Biol. 152:375–384. doi:10.1083/jcb.152.2.375
- Cheng, Y., R.M. Miura, and B. Tian. 2006. Prediction of mRNA polyadenylation sites by support vector machine. *Bioinformatics*. 22:2320–2325. doi:10.1093/bioinformatics/btl394
- Clapier, C.R., and B.R. Cairns. 2009. The biology of chromatin remodeling complexes. Annu. Rev. Biochem. 78:273–304. doi:10.1146/annurev .biochem.77.062706.153223
- Clayton, A.L., and L.C. Mahadevan. 2003. MAP kinase-mediated phosphoacetylation of histone H3 and inducible gene regulation. FEBS Lett. 546:51–58. doi:10.1016/S0014-5793(03)00451-4
- Davey, C.A., D.F. Sargent, K. Luger, A.W. Maeder, and T.J. Richmond. 2002. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. *J. Mol. Biol.* 319:1097–1113. doi:10.1016/ S0022-2836(02)00386-8
- Gautier, T., D.W. Abbott, A. Molla, A. Verdel, J. Ausio, and S. Dimitrov. 2004. Histone variant H2ABbd confers lower stability to the nucleosome. EMBO Rep. 5:715–720. doi:10.1038/sj.embor.7400182
- Guex, N., and M.C. Peitsch. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*. 18:2714–2723. doi:10.1002/elps.1150181505
- Gustafsson, M.G., L. Shao, P.M. Carlton, C.J. Wang, I.N. Golubovskaya, W.Z. Cande, D.A. Agard, and J.W. Sedat. 2008. Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. *Biophys. J.* 94:4957–4970. doi:10.1529/biophysj.107.120345
- Hake, S.B., B.A. Garcia, M. Kauer, S.P. Baker, J. Shabanowitz, D.F. Hunt, and C.D. Allis. 2005. Serine 31 phosphorylation of histone variant H3.3 is specific to regions bordering centromeres in metaphase chromosomes. *Proc. Natl. Acad. Sci. USA*. 102:6344–6349. doi:10.1073/pnas.0502413102
- Hake, S.B., B.A. Garcia, E.M. Duncan, M. Kauer, G. Dellaire, J. Shabanowitz, D.P. Bazett-Jones, C.D. Allis, and D.F. Hunt. 2006. Expression patterns and post-translational modifications associated with mammalian histone H3 variants. J. Biol. Chem. 281:559–568. doi:10.1074/jbc .M509266200
- Hendzel, M.J., Y. Wei, M.A. Mancini, A. Van Hooser, T. Ranalli, B.R. Brinkley, D.P. Bazett-Jones, and C.D. Allis. 1997. Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation. *Chromosoma*. 106:348–360. doi:10.1007/s004120050256
- Huyen, Y., O. Zgheib, R.A. Ditullio Jr., V.G. Gorgoulis, P. Zacharatos, T.J. Petty, E.A. Sheston, H.S. Mellert, E.S. Stavridi, and T.D. Halazonetis. 2004. Methylated lysine 79 of histone H3 targets 53BP1 to DNA double-strand breaks. *Nature*. 432:406–411. doi:10.1038/nature03114
- Im, H., C. Park, Q. Feng, K.D. Johnson, C.M. Kiekhaefer, K. Choi, Y. Zhang, and E.H. Bresnick. 2003. Dynamic regulation of histone H3 methylated at lysine 79 within a tissue-specific chromatin domain. J. Biol. Chem. 278:18346–18352. doi:10.1074/jbc.M300890200
- Köhler, G., and C. Milstein. 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*. 256:495–497. doi:10.1038/256495a0
- Kozak, M. 1991. An analysis of vertebrate mRNA sequences: intimations of translational control. J. Cell Biol. 115:887–903. doi:10.1083/jcb.115.4.887
- McKittrick, E., P.R. Gafken, K. Ahmad, and S. Henikoff. 2004. Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc. Natl. Acad. Sci. USA*. 101:1525–1530. doi:10.1073/pnas.0308092100
- Nakagawa, S., Y. Niimura, T. Gojobori, H. Tanaka, and K. Miura. 2008. Diversity of preferred nucleotide sequences around the translation initiation codon in eukaryote genomes. *Nucleic Acids Res.* 36:861–871. doi:10.1093/nar/gkm1102

- Palmer, D.K., K. O'Day, M.H. Wener, B.S. Andrews, and R.L. Margolis. 1987. A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. J. Cell Biol. 104:805–815. doi:10.1083/ icb.104.4.805
- Pehrson, J.R., and V.A. Fried. 1992. MacroH2A, a core histone containing a large nonhistone region. Science. 257:1398–1400. doi:10.1126/science .1529340
- Polo, S.E., D. Roche, and G. Almouzni. 2006. New histone incorporation marks sites of UV repair in human cells. *Cell.* 127:481–493. doi:10.1016/ j.cell.2006.08.049
- Pusarla, R.H., and P. Bhargava. 2005. Histones in functional diversification. Core histone variants. FEBS J. 272:5149–5168. doi:10.1111/j.1742-4658.2005.04930.x
- Sancho, M., E. Diani, M. Beato, and A. Jordan. 2008. Depletion of human histone H1 variants uncovers specific roles in gene expression and cell growth. PLoS Genet. 4:e1000227. doi:10.1371/journal.pgen.1000227
- Santenard, A., and M.E. Torres-Padilla. 2009. Epigenetic reprogramming in mammalian reproduction: contribution from histone variants. *Epi-genetics*. 4:80–84. doi:10.4161/epi.4.2.7838
- Schermelleh, L., P.M. Carlton, S. Haase, L. Shao, L. Winoto, P. Kner, B. Burke, M.C. Cardoso, D.A. Agard, M.G. Gustafsson, et al. 2008. Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science*. 320:1332–1336. doi:10.1126/science.1156947
- Shechter, D., H.L. Dormann, C.D. Allis, and S.B. Hake. 2007. Extraction, purification and analysis of histones. *Nat. Protoc.* 2:1445–1457. doi:10.1038/nprot.2007.202
- Strahl, B.D., and C.D. Allis. 2000. The language of covalent histone modifications. *Nature*. 403:41–45. doi:10.1038/47412
- Suto, R.K., M.J. Clarkson, D.J. Tremethick, and K. Luger. 2000. Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. Nat. Struct. Biol. 7:1121–1124. doi:10.1038/81971
- Tabaska, J.E., and M.Q. Zhang. 1999. Detection of polyadenylation signals in human DNA sequences. *Gene*. 231:77–86. doi:10.1016/S0378-1119 (99)00104-3
- Tagami, H., D. Ray-Gallet, G. Almouzni, and Y. Nakatani. 2004. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell.* 116:51–61. doi:10 .1016/S0092-8674(03)01064-X
- Tay, S.K., J. Blythe, and L. Lipovich. 2009. Global discovery of primate-specific genes in the human genome. *Proc. Natl. Acad. Sci. USA*. 106:12019– 12024. doi:10.1073/pnas.0904569106
- Trostle-Weige, P.K., M.L. Meistrich, W.A. Brock, and K. Nishioka. 1984. Isolation and characterization of TH3, a germ cell-specific variant of histone 3 in rat testis. *J. Biol. Chem.* 259:8769–8776.
- van Holde, K.E. 1988. Chromatin. Springer, New York. 497 pp.
- Witt, O., W. Albig, and D. Doenecke. 1996. Testis-specific expression of a novel human H3 histone gene. Exp. Cell Res. 229:301–306. doi:10.1006/ excr.1996.0375
- Wysocka, J., P.T. Reilly, and W. Herr. 2001. Loss of HCF-1-chromatin association precedes temperature-induced growth arrest of tsBN67 cells. *Mol. Cell. Biol.* 21:3820–3829. doi:10.1128/MCB.21.11.3820-3829.2001
- Yan, C., and D.D. Boyd. 2006. Histone H3 acetylation and H3 K4 methylation define distinct chromatin regions permissive for transgene expression. *Mol. Cell. Biol.* 26:6357–6371. doi:10.1128/MCB.00311-06
- Zhang, Y. 2008. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 9:40. doi:10.1186/1471-2105-9-40