THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 282, NO. 10, pp. 7641–7655, March 9, 2007

Organismal Differences in Post-translational Modifications in Histones H3 and H4*[S]

Received for publication, August 17, 2006, and in revised form, December 19, 2006 Published, JBC Papers in Press, December 28, 2006, DOI 10.1074/jbc.M607900200

Benjamin A. Garcia[‡1,2], Sandra B. Hake[§1,3], Robert L. Diaz[§], Monika Kauer[§4], Stephanie A. Morris[¶], Judith Recht[§5], Jeffrey Shabanowitz[‡], Nilamadhab Mishra[‖], Brian D. Strahl[¶6], C. David Allis[§7], and Donald F. Hunt[‡8]**

*From the [‡]Department of Chemistry, University of Virginia, Charlottesville, Virginia 22901, the [§]Laboratory of Chromatin Biology, The Rockefeller University, New York, New York 10021, the [¶]Department of Biochemistry and Biophysics, University of North Carolina School of Medicine, Chapel Hill, North Carolina 27599, the [‖]Section on Rheumatology and Clinical Immunology, Department of Internal Medicine, Wake Forest University School of Medicine, Winston-Salem, North Carolina 27157, and the **Department of Pathology, Health Sciences Center, University of Virginia, Charlottesville, Virginia 22908*

Post-translational modifications (PTMs) of histones play an important role in many cellular processes, notably gene regulation. Using a combination of mass spectrometric and immunobiochemical approaches, we show that the PTM profile of histone H3 differs significantly among the various model organisms examined. Unicellular eukaryotes, such as *Saccharomyces cerevisiae* (yeast) and *Tetrahymena thermophila* (Tet), for example, contain more activation than silencing marks as compared with mammalian cells (mouse and human), which are generally enriched in PTMs more often associated with gene silencing. Close examination reveals that many of the better-known modified lysines (Lys) can be either methylated or acetylated and that the overall modification patterns become more complex from unicellular eukaryotes to mammals. Additionally, novel species-specific H3 PTMs from wild-type asynchronously grown cells are also detected by mass spectrometry. Our results suggest that some PTMs are more conserved than previously thought, including H3K9me1 and H4K20me2 in yeast and H3K27me1, -me2, and -me3 in Tet. On histone H4, methylation at Lys-20 showed a similar pattern as H3 methylation at Lys-9, with mammals containing more methylation than the unicellular organisms. Additionally, modification profiles of H4 acetylation were very similar among the organisms examined.

Cellular identity is defined by the characteristic patterns of gene expression and silencing. Inheritance of these transcription patterns through DNA replication and chromatin assembly that accompanies each cell division is crucial for cell survival, but the one or more mechanisms by which this "memory" is achieved are not well understood (reviewed in Ref. 1). A rapidly emerging literature suggests that histone proteins, which function to package genomic DNA into repeating nucleosomal units that are then further folded into higher order chromatin fibers, may be major carriers of epigenetic information (2). Each nucleosome typically contains ~146 bp of DNA wrapped around two copies each of histones H3, H4, H2A, and H2B. Although providing a relative constant packaging theme, subtle changes in nucleosome histone:DNA and histone:histone contacts are likely to provide variation in fiber folding that, in turn, translates into biological readout.

In general, the packaging of DNA into chromatin is recognized to be a major mechanism by which the access of genomic DNA is restricted. This physical barrier to the underlying DNA is precisely regulated (and counteracted), at least in part, by the post-translational modifications (PTMs)[9] of histones. A wide number of studies has revealed that PTMs of histones, especially those located in the N-terminal tails, play a pivotal role in the regulation of chromatin structure necessary for DNA accessibility during gene expression. Remarkable diversity in the histone/nucleosome structure is generated by a variety of PTMs, such as lysine and arginine methylation, lysine acetylation, serine and threonine phosphorylation, and lysine ubiquitination (reviewed in Refs. 3–5). Some PTMs, including acetylation and phosphorylation, are reversible and dynamic and are often associated with inducible gene expression. Other PTMs, such as lysine methylation, are often found to be more stable, participating in long term maintenance of the expression status of regions in the genome that vary in certain developmental contexts (6). Nevertheless, recent work has shown that many lysine methylation "marks" on histones can also be reversed by enzymatic means (7–9).

* This work was supported in part by National Institutes of Health Grants GM 40922 (to C. D. A.), GM 37537 (to D. F. H.), and GM 68088 (to B. D. S.), by the Ford Foundation (to B. A. G. and S. A. M.), and by The Rockefeller University ("Women & Science Fellowship to S. B. H.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S6 and supplemental data pages 1–34.

[1] Both authors contributed equally to the work.

[2] Present address: Institute for Genomic Biology, University of Illinois, Urbana-Champaign, Illinois 61801.

[3] Present address: Adolf-Butenandt-Institut, Molekularbiologie, Ludwig-Maximilians-Universität, 80336 München, Germany.

[4] Present address: Research Institute of Molecular Pathology, The Vienna Biocenter, A-1030 Vienna, Austria.

[5] Present address: Dept. of Oncological Sciences, Mount Sinai School of Medicine, New York, NY 10029.

[6] A Pew Scholar in the Biomedical Sciences.

[7] To whom correspondence may be addressed: Laboratory of Chromatin Biology, The Rockefeller University, Box 78, 1230 York Ave., New York, NY 10021. Tel.: 212-327-7839; Fax: 212-327-7849; E-mail: alliscd@rockefeller.edu.

[8] To whom correspondence may be addressed. Tel.: 434-924-3610; Fax: 434-982-2781; E-mail: dfh@virginia.edu.

[9] The abbreviations used are: PTM, post-translational modification; ac, acetylation; me, methylation; H3, histone H3; H4, histone H4; Tet, *T. thermophila*; MS, mass spectrometry; MS/MS, tandem mass spectrometry; MEF, mouse embryonic fibroblast; RP-HPLC, reversed-phase high-performance liquid chromatography.

Histone Modification Differences among Organisms

Histone PTMs occur on multiple but specific sites, suggesting that histone PTMs can act as signaling platforms for proteins that "read" these marks (10). The "histone-code hypothesis" has been put forward to explain how different combinations of histone PTMs can result in distinct chromatin-regulated functions (11–13). The deciphering of the "code" is complicated further by the observation that certain residues can be either acetylated or methylated, as has been shown for histone H3 lysine 9 in mammalian cells (14). Acetylation of these residues is associated with transcriptional activation and methylation often, but not always (see below), with repression of genes. In contrast, methylation of histone H3 on Lys-4, Lys-36, and Lys-79 correlates with transcriptional activation of genes (15–17). Besides PTMs, the exchange of core histones with histone variants, which differ slightly in their amino acid sequence and which are enriched in different marks (18, 19), adds additional complexity to the structure and function of chromatin.

In this study, we sought to examine the extent to which the histone code is universal, focusing on H3 and H4 from a limited, but select, number of popular organisms as an entry point for our studies. Using both mass spectrometry (MS) and immunobiochemical approaches, we show that the PTM profile of histone H3 greatly differs among species and that the PTM pattern is more complex in mammals as compared with "simple" eukaryotes, such as *Saccharomyces cerevisiae* (yeast) and *Tetrahymena thermophila* (Tet). Unicellular eukaryotes exhibit more marks associated with transcriptional activation or transcriptional competency, whereas mammals contain more modifications linked to gene repression. Additionally, we identified several novel PTMs that are potentially species-specific, such as H3K4ac (observed in Tet, mouse, and human) and H3K79ac (human), H3K23me1 (yeast, mouse, and human), H3K14me2 and H3K64ac (yeast), and H3K56me3 (human). Surprisingly, we observed low amounts of H3K9me1 and H3K27me3 in yeast, and H3K27me1, -me2, and -me3 in Tet; such modifications have not been detected before in these particular organisms.

Additionally, we characterized PTMs of histone H4 from the same set of organisms. We found that H4 is far less modified than H3 and that the H4 acetylation patterns remained consistent across the species examined. In addition, H4K20 methylation was found in higher levels on human and mouse than the unicellular species. Some low abundance novel H4 PTMs were also identified, such as H4K20me2 and H4K20ac in yeast. Collectively, our data reveal a diverse pattern of modification usage on histone H3 that is fundamentally different among unicellular organisms and mammals. We suggest that these differences may be due to the co-development of additional histone-modifying enzymes and histone H3 variants, allowing for additional regulation of more complex genomes.

EXPERIMENTAL PROCEDURES

Cell Cultures—HeLa, HEK293, NIH3T3, and MEF cell lines were grown in Iscove's Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum and penicillin/streptomycin at 37 °C and 5% CO_2. Insect S2 cells were grown in Iscove's insect medium (Invitrogen) supplemented with 10% fetal calf serum at 27 °C. *T. thermophila* and *S. cerevisiae* (strain BY4741) were grown as described previously (20, 21). Splenocytes from MRL-MPJ mice were also prepared as previously described (22).

Histone Extraction—Nuclei and histones from mammalian and *Drosophila* cells were isolated as described earlier (19). *T. thermophila* and yeast histones were isolated as described previously (20, 21).

RP-HPLC—Separation of human and mouse core histones by RP-HPLC was performed as described previously (19). RP-HPLC fractions were resuspended in water, a fraction was analyzed by SDS-PAGE and stained with Coomassie Brilliant Blue as control, and the liquid samples were subjected to MS analysis. Similar gradients were used for *T. thermophila* and yeast histones (20, 21).

Histone Sample Preparation—Histones H3 and H4 were treated with propionylation reagent, digested with trypsin, and re-propionylated as previously described (23).

MS—Propionylated histone digest mixtures were loaded onto capillary precolumns and analyzed essentially as previously reported (24). Briefly, all samples were analyzed by nanoflow HPLC-microelectrospray ionization MS/MS (24) on a Finnigan linear quadrupole ion trap-Fourier transform ion cyclotron resonance mass spectrometer (Thermo Electron, San Jose, CA) operated in data-dependent mode. The HPLC gradient used on an Agilent 1100 series HPLC solvent delivery system (Palo Alto, CA) consisted of 0–60% *B* in 60 min, and then 60–100% *B* in 10 min (*A* = 0.1% acetic acid, *B* = 70% acetonitrile in 0.1% acetic acid).

Immunoblotting—Immunoblots with total acid extracted histones were done as described previously (25). For peptide competition, the H3K4 acetyl antibody was incubated for 1 h at 4 °C with 1 μg/ml of the following peptides: H3 unmodified, amino acids 1–20; H3K4ac, amino acids 1–20 with Lys-4 acetylated; H3K4me2; amino acids 1–20 with Lys-4 dimethylated; H3K9ac, amino acids 1–20 with Lys-9 acetylated; H3K36ac, amino acids 27–46 with Lys-36 acetylated; H4 unmodified, amino acids 1–16; H4K5ac, amino acids 2–20 with Lys-5 acetylated; and H4 tetra-acetyl, amino acids 1–18 with Lys-5, Lys-8, Lys-12, and Lys-16 acetylated. Peptides were synthesized and verified by MS at the Proteomics Resource Center of The Rockefeller University, Upstate Biotechnology, and the Protein Chemistry Core Facility at the University of North Carolina, School of Medicine.

Enzyme-linked Immunosorbent Assay—The same peptides used in peptide competition assays were also used for the enzyme-linked immunosorbent assay. Briefly, peptides were diluted in phosphate-buffered saline in concentrations from 0 to 100 ng/ml, and 200 μl was transferred to a 96-well plate (Covance) and incubated overnight at 37 °C. The plate was washed twice with phosphate-buffered saline (containing 0.05% Tween 20), blocked with 200 μl of phosphate-buffered saline/0.05% Tween 20 containing 1% bovine serum albumin, and incubated at 37 °C for 1 h. After two additional wash steps, 100 μl of diluted primary antibody was added to the plate, and the mixture was incubated for 2 h at 37 °C. After washing, 100 μl of secondary horseradish peroxidase-conjugated antibody was added, and the mixture was incubated for 2 h at 37 °C. After two wash steps 100 μl of *O*-phenylenediamine dihydrochloride

substrate (Sigma) was added, and the plate was incubated at room temperature in the dark for 30 min. The detection reaction was stopped with 100 μl of 3 M HCl, and the absorbance was measured at 492 nm in a plate reader (Bioscan).

Antibodies—To develop an antibody against histone H3K4ac, a synthetic histone H3 peptide modified by acetylation at Lys-4 was chemically synthesized, conjugated to KLH, and used to immunize rabbits by Upstate Biotechnology. The positive anti-sera were further purified by immunoaffinity purification. The H3K4ac antibody was diluted 1:1,000 in immunoblots and enzyme-linked immunosorbent assay and 1:100 in immunofluorescence microscopy. The following antibodies from Upstate Biotechnology were used as 1:1,000 dilutions in immunoblots: H3K9me1, -me2, and -me3; H3K27me1, -me2, -me3, and -ac; H3K36me1 and -me2; and H3K79me2. One other antiserum from UBI was used in this study in the following dilution in immunoblot: H3K56ac (1:5,000), and the following antibodies from Abcam were used: H3K4me1, -me2, and -me3 (1:1,000) and H3K79me3 (1:1,000). The anti-H3K36me3 antibody, obtained from UBI, was used at a 1:10,000 dilution.

Immunofluorescence Microscopy—Immunofluorescence microscopy analyses were done as described previously (25).

RESULTS

Comparison of Histone H3 PTMs among Unicellular Eukaryotes and Mammals—To examine the PTM profiles on histone H3 from different species, we extracted and RP-HPLC-purified histones from several model organisms, including yeast, Tet, mouse, and human cells. All of the RP-HPLC H3-containing fractions (regardless of which H3 variants were present, see supplemental Fig. S1) were combined and then subjected to tandem mass spectrometry (MS/MS), which has been successfully used previously to analyze histone H3 (19). This procedure results in the production of predicted peptide sets that could be easily detected by MS (19).

Species Comparison of "ON" Marks—Shown in Fig. 1 are mass spectra that facilitate a comparative analysis of PTMs that occur at lysine 4 of H3, a well known "ON" epigenetic methylation mark (26). Summed mass spectra that record the ion abundances for $[M + 2H]^{2+}$ ions corresponding to residues 3–8 of H3 from human, mouse, Tet, and yeast are presented in Fig. 1, *A–D*, respectively. Because the protein samples have been derivatized with propionic anhydride (addition of 56 Da to unmodified and monomethylated lysine residues, only) prior to enzymatic digestion with trypsin, the signals for the $[M + 2H]^{2+}$ ions of peptides containing H3K4me2, H3K4me3, H3K4, and H3K4me1 appeared at *m*/*z* 394, 401, 408, and 415, respectively.

As indicated in Fig. 1 (*A–D*), the degree of methylation of Lys-4 among the various organisms examined greatly differs with several consistent trends noted. As shown in Fig. 1 (*A* and *B*), human and mouse H3 display similar profiles: unmodified Lys-4 is the most abundant species, whereas K4me1 is very low. Occasionally, contaminant peaks were observed in the mass spectra that, upon fragmentation, produce MS/MS spectra, which do not match any histone peptide such as the singly charged ion around *m*/*z* 396.8 (Fig. 1*B*, *asterisk*). Surprisingly, a small amount of K4ac was seen in human H3 (Fig. 1*A* and supplemental Figs. S2 and S3), as differentiated by accurate mass (−0.25 ppm) and retention time. Accurate mass measurement with the linear quadrupole ion trap-Fourier transform instrument can distinguish between trimethylated and acetylated peptides ($\Delta m = 0.0364$ Da). To our knowledge, H3K4ac has not been reported in histones from any source. We also detected this novel modification as existing in a lower abundance in Tet and mouse, but this mark was not observed with our current detection limits in yeast.

In striking contrast, however, we found that unicellular eukaryotes (Tet and yeast) contained higher degrees of H3K4me1, -me2, and -me3. Unlike in human and mouse H3; Tet (Fig. 1*C*) and yeast (Fig. 1*D*) H3 display clearly visible amounts of K4me2 and -me3. These results agree well with immunoblot analyses showing a large amount of H3K4me2 detected on histone H3 from ciliate and yeast as compared with mouse and human histone (27), as well as robust enzymatic Lys-4 methylation activity found in crude preparations of yeast and ciliate nuclei (28). The degree of methylation on H3K4 is easily differentiated by tandem MS. For example, an MS/MS spectrum of the $[M + 2H]^{2+}$ peptide at 394.7344 *m*/*z* from the Tet sample (Fig. 1*C*) is shown in Fig. 1*E*, and the b_1, b_2, y_4, and y_5 fragment ions indicate that H3K4 is dimethylated on this peptide (3–8 residues). Fig. 1*F* shows an MS/MS spectrum of the doubly charged peptide ion at 401.7425 *m*/*z* generated from the yeast H3 sample (Fig. 1*D*) and was identified as the 3–8 residue fragment containing K4me3 as revealed by the y_4 and y_5 fragment ions. The experimental mass of this peptide is consistent with trimethylation (−1.1 ppm error) and not acetylation (+44.1 ppm error). These data show that mammals have far less H3K4 methylation than unicellular eukaryotes, suggesting that the usage of this particular ON mark was reduced during evolution.

Another modification associated with transcriptional activation is the methylation of H3K79 by the Dot1 histone methyltransferase (29). A comparison of PTMs on a peptide spanning residues 73–83 of histone H3 that contains the Lys-79 residue from the above mentioned four different species can be seen in Fig. 2 The PTM profiles at this site are remarkably different than what we observed on peptides containing H3K4. For example, human and mouse H3 (Fig. 2, *A* and *B*) displayed similar profiles (H3K79me1 is the most abundant species) and H3K79me2 was also observed. Interestingly, a small amount of H3K79ac could be seen as differentiated by accurate mass (+3.2 ppm) on the human sample (see also supplemental data page 6). Nevertheless, the degree of methylation of H3K79 was restricted to only a small amount of monomethylation in Tet (Fig. 2*C*), whereas yeast contained a high degree of H3K79 methylation with the trimethylated form being the most abundant (Fig. 2*D*). A tandem mass spectrum of the $[M + 2H]^{2+}$ peptide at 727.8863 *m*/*z* from the Tet sample is shown in Fig. 2*E*, and fragment ions (b_6, b_7, y_4, and y_5) indicate that H3K79 is monomethylated on this peptide (73–83 residues). Fig. 2*F* shows an MS/MS spectrum of the doubly charged ion at 717.3864 *m*/*z* generated from the yeast H3 sample (Fig. 2*D*). This peptide was identified as the 73–83 residue fragment containing H3K79me3 as determined from observing the same indicator fragment ions (b_6, b_7, y_4, and y_5) as seen on Fig. 2*E*.

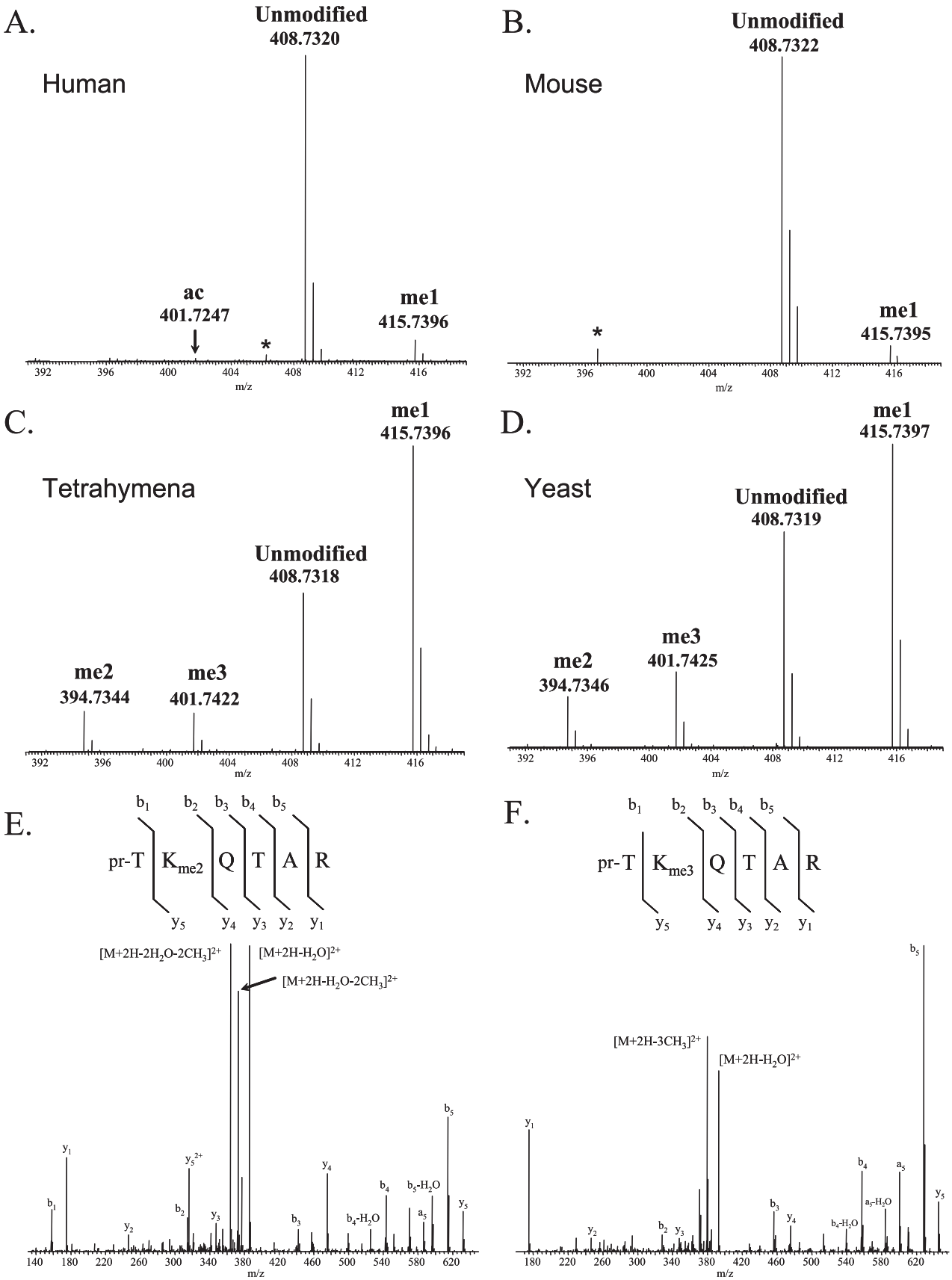

FIGURE 1. **Comparison of full mass spectra depicting H3K4 modifications on the $[M + 2H]^{2+}$ peptide spanning residues 3–8 TKQTAR from human HEK cells (*A*), mouse MEF cells (*B*), Tet (*C*), and yeast (*D*).** *E*, MS/MS spectrum of the $[M + 2H]^{2+}$ 3–8 peptide containing H3K4me2 from the Tet sample. A propionyl group (*pr*) is present on the N terminus. *F*, MS/MS spectrum of the $[M + 2H]^{2+}$ 3–8 peptide containing H3K4me3 from the yeast sample. A propionyl group (*pr*) is present on the N terminus. b and y type ions are labeled in *E* and *F*. Ions arising from non-histone peptides or contamination are labeled with an *asterisk*.

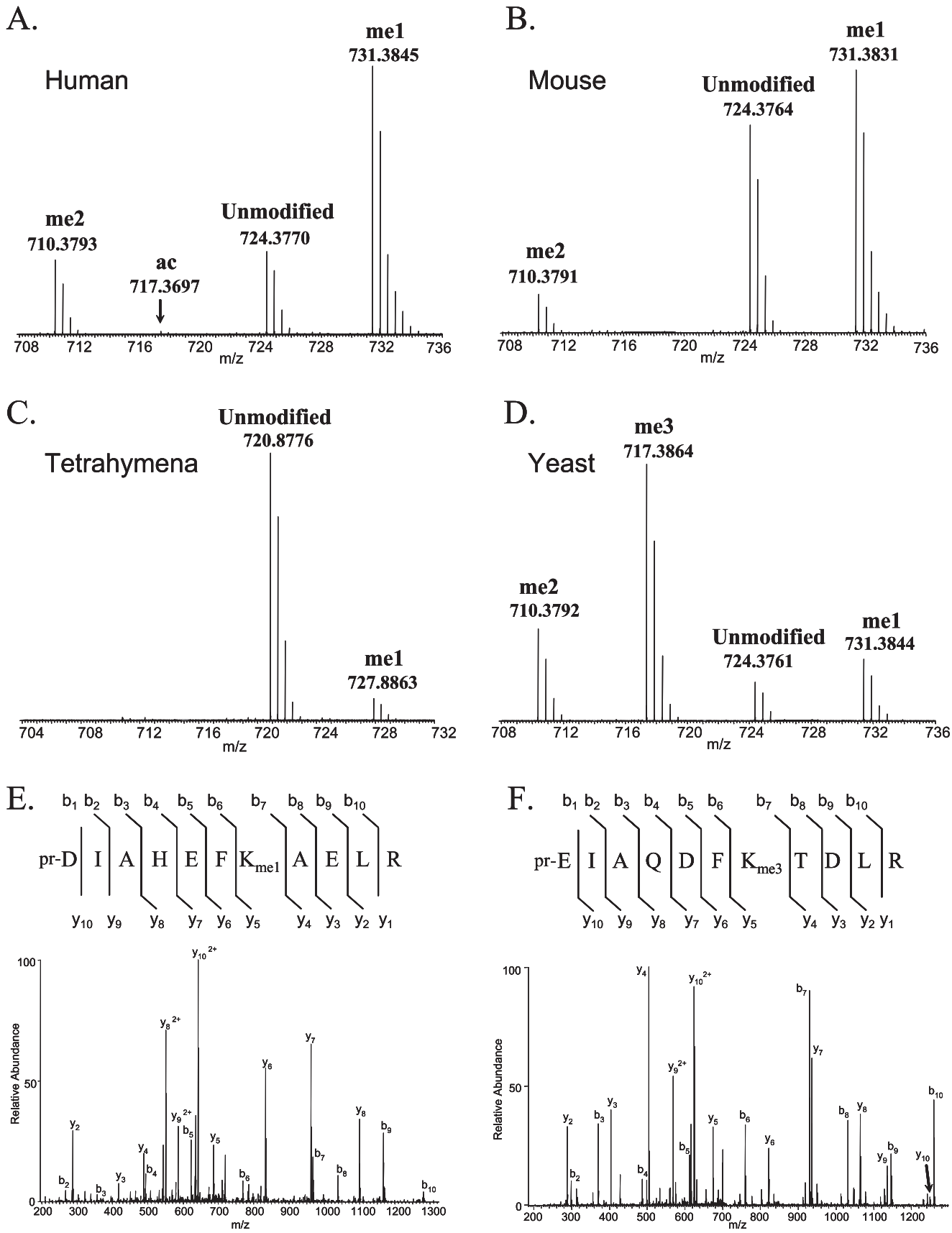

FIGURE 2. **Comparison of full mass spectra depicting H3K79 modifications on the $[M + 2H]^{2+}$ peptide spanning residues 73–83 (EIAQDFKTDLR) from human HEK cells (*A*), mouse MEF cells (*B*), Tet (DIAHEFKAELR) (*C*), and yeast (*D*).** *E*, MS/MS spectrum of the $[M + 2H]^{2+}$ 73–83 peptide containing H3K79me1 from the Tet sample. A propionyl group (*pr*) is present on the N terminus. *F*, MS/MS spectrum of the $[M + 2H]^{2+}$ 73–83 peptide containing H3K79me3 from yeast. A propionyl group (*pr*) is present on the N terminus. b and y type ions are labeled in *E* and *F*.

The experimental mass of this peptide (717.3864 *m*/*z*) determined that a trimethylation (+1.1 ppm error) and not acetylation (+26.5 ppm error) mark was present.

These data suggest that, similar to the H3K4 modification profile, mouse and human are almost identical in their H3K79 modifications. On the other hand, in contrast to H3K4 methylation, yeast and Tet are different from each other in their PTM profile, possibly due to the lack of Dot1 and sequence differences in Tet. In conclusion, mammals are similar to each other, but differ greatly from unicellular eukaryotes with regard to their H3K4 and Lys-79 modification profiles.

Species Comparison of OFF Marks—A well established modification that is associated with transcriptional repression and constitutive heterochromatin formation is the methylation of H3K9 (30). We wondered if organisms differ in their H3K9 modification profile from each other, as we have seen for ON marks (Figs. 1 and 2). A species comparison of modifications at Lys-9 on histone H3 is shown in Fig. 3, a residue known to exhibit both acetylation (ON) or methylation (OFF). Fig. 3 shows the summed full mass spectra around the 520–550 *m*/*z* range, which display the $[M + 2H]^{2+}$ peptides spanning residues 9–17 from human (*A*), mouse (*B*), Tet (*C*), and yeast (*D*) samples. In contrast to the modification profiles obtained on H3K4 from the same species, an increasing shift in the degree of methylation at H3K9 was seen in mammals as compared with Tet and yeast. Both human and mouse contain H3K9, -me1, -me2, -me3, and -ac, with H3K9me2 being the most abundant modification.

Peptides containing Lys-9 from Tet and yeast H3 (Fig. 3, *C* and *D*) show that this residue is not as heavily modified, because the unmodified peptide is the most abundant species in both organisms. Acetylated peptides account for the most abundant modification in both unicellular organisms, with H3K9ac in Tet being present at relatively higher levels (Fig. 3*C*). Unexpectedly, based on the existing literature, a low amount of H3K9me3 is detected in Tet, and H3K9me1 (see also supplemental data page 3) in yeast to far less of an extent and degree of methylation at Lys-9 that characterizes mammals. Localization of the modified site was accomplished by MS/MS, and Fig. 3*E* depicts an MS/MS spectrum of the $[M + 2H]^{2+}$ 9–17 residue peptide at 535.3129 *m*/*z* from the Tet sample. Fragment ion b_1 indicates the presence of H3K9me3. A tandem mass spectrum is displayed in Fig. 3*F* of the $[M + 2H]^{2+}$ 9–17 residue peptide at 521.3056 *m*/*z* generated from the human sample, and the b_1 ion determined that it contained H3K9me2.

These data suggest that, as observed for ON modifications, the H3K9 modification profiles of mammals are almost identical to each other but differ greatly from the ones detected in unicellular eukaryotes. In contrast to the ON marks, which are present in very low abundance in human and mouse, OFF marks are easily detectable in more complex organisms, whereas the opposite is true for unicellular eukaryotes.

Lastly, we examined the modification profile of H3 peptides spanning residues 27–40, which contain both Lys-27 and Lys-36, two well known modification sites that when methylated have been linked to gene repression and activation, respectively (16, 31). Shown in Fig. 4 are mass spectra that facilitate a comparative analysis of modifications that occur on Lys-27 and Lys-36 of H3. Summed mass spectra, which record the ion abundances for $[M + 2H]^{2+}$ ions corresponding to residues 27–40 of H3 from human, mouse, Tet, and yeast, are presented in *panels A–D*, respectively. In samples from human and mouse cells, we found several different isomeric forms of H3. Of the H3 isomeric forms detected in mammals the one acetylated at Lys-27 was by far the least abundant. Isomeric forms without modifications on Lys-27 and Lys-36 and those that contain K27me1, K27me2, K36me1, and K36me2 and modifications on both Lys-27 and Lys-36 were of intermediate abundance. H3 with K27me3 was the most abundant isomeric form in the mixture.

Fig. 4 (*C* and *D*) shows the mass spectral data acquired on H3 isomeric forms isolated from Tet and yeast. We found these samples were enriched in H3 not modified at either Lys-27 or Lys-36. Both samples were also enriched for the isomeric form containing K27ac. Isomeric forms that contained K27me1, K27me2, and K27me3 were abundant in Tet but not readily detected in yeast. The finding of methylated H3 Lys-27 in Tet is novel and is currently under investigation.[10] Isomeric forms that contain K36me1, -me2, and -me3 were abundant in yeast, but only K36me1 could be detected in Tet. Evidence for H3K36ac was found in both Tet and yeast samples, albeit at low levels, and the function of this particular new mark is currently under investigation (see companion report (52)).

Although several peptides containing the same nominal masses spanning residues 27–40 were observed, their slightly different elution profiles allow for MS/MS spectra to be recorded from these unique species thus unambiguously identifying the modified site, allowing for PTM assignment to either H3K27 or H3K36. For example, Fig. 4*E* shows an MS/MS spectrum of the doubly charged ion at *m*/*z* 830.4792 from the yeast sample, and the fragment ions b_9, b_{10}, y_4, and y_5 identify Lys-36 as a trimethylation site. The b_1 ion, however, was used to identify the trimethylation site on Lys-27 from the MS/MS spectrum of the human sample (Fig. 4*F*). We have also included several representative MS/MS spectra, which demonstrate the ability of our MS platform to distinguish peptides with modifications at either H3K27 or H3K36, as supplemental data pages 7–24.

In sum, Tet primarily utilizes H3K27 methylation, whereas yeast mostly uses H3K36 methylation. Additionally, unicellular eukaryotes also do not display the complex array of peptides consisting of different combination of concurrent modifications at H3K27 and H3K36 as do mammals.

Table 1 provides a summary of all covalent modifications detected on H3 from yeast, Tet, mouse, and human cells. Our analyses show a remarkably large number of differences in the usage of covalent modifications among species (+ = easily detected in both experiments, ND = not detected, $+^{L}$ = detected in low abundance, and $+^{V}$ = detected in only one experiment).

Immunobiochemical Analysis of Abundant H3 PTMs—We first analyzed the MS data for the presence of very abundant

[10] Taverna, S. D., Ueberheide, B. M., Lin, Y., Tackett, A. J., Diaz, R. L., Shabanowitz, J., Chait, B. T., Hunt, D. F., and Allis, C. D. (2007) *Proc. Natl. Acad. Sci. U.S.A.*, in press.

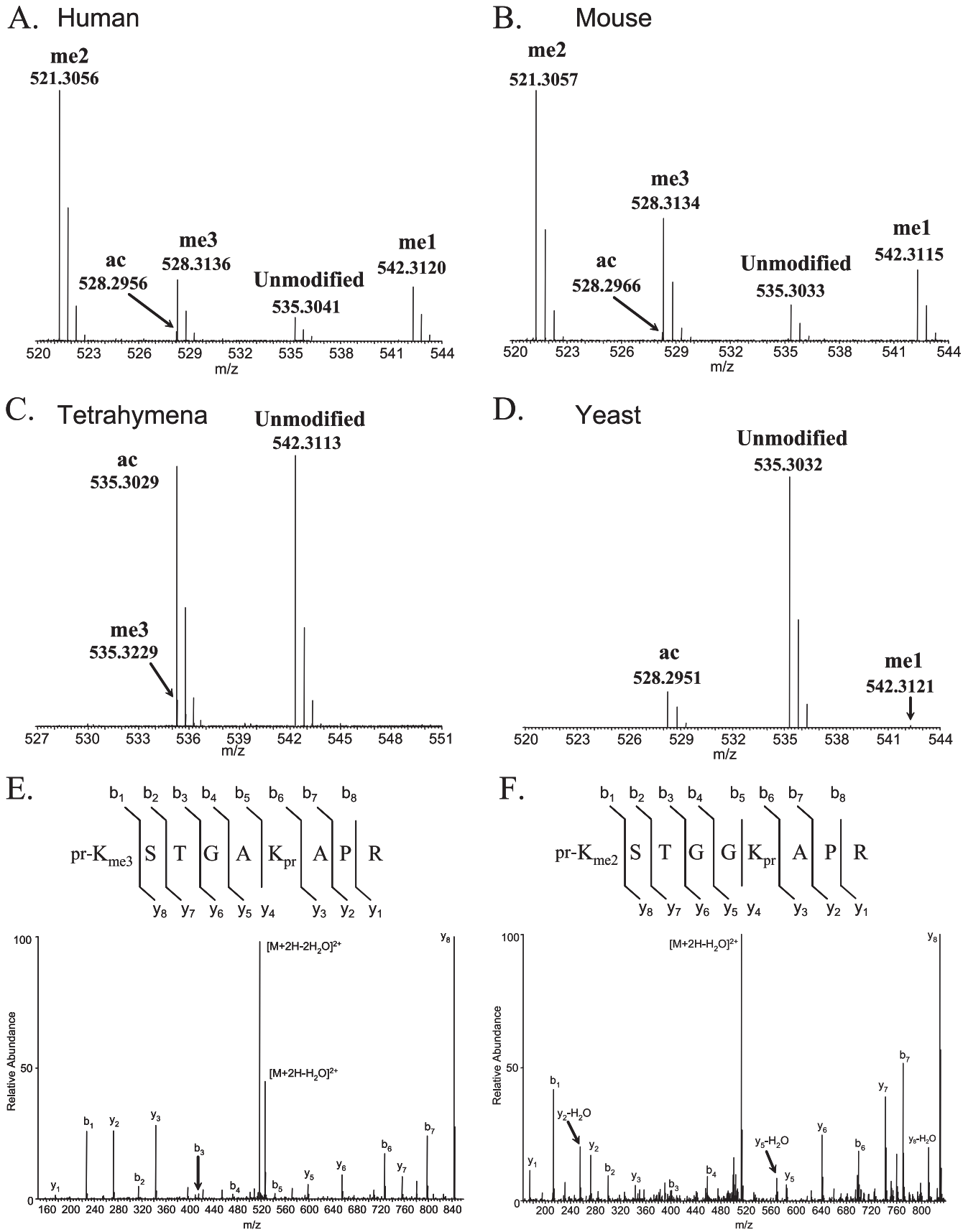

FIGURE 3. **Comparison of full mass spectra depicting H3K9 modifications on the $[M + 2H]^{2+}$ peptide spanning residues 9–17 KSTGGKAPR from human HEK cells (*A*), mouse MEF cells (*B*), Tet (KSTGAKAPR) (*C*), and yeast (*D*).** *E*, MS/MS spectrum of the $[M + 2H]^{2+}$ 9–17 peptide containing H3K9me3 from the Tet sample. Propionyl groups (*pr*) are present on the N terminus and on Lys-14. *F*, MS/MS spectrum of the $[M + 2H]^{2+}$ 9–17 peptide containing H3K9me2 from the human sample. A propionyl group (*pr*) is present on the N terminus and Lys-14. b and y type ions are labeled in *E* and *F*.

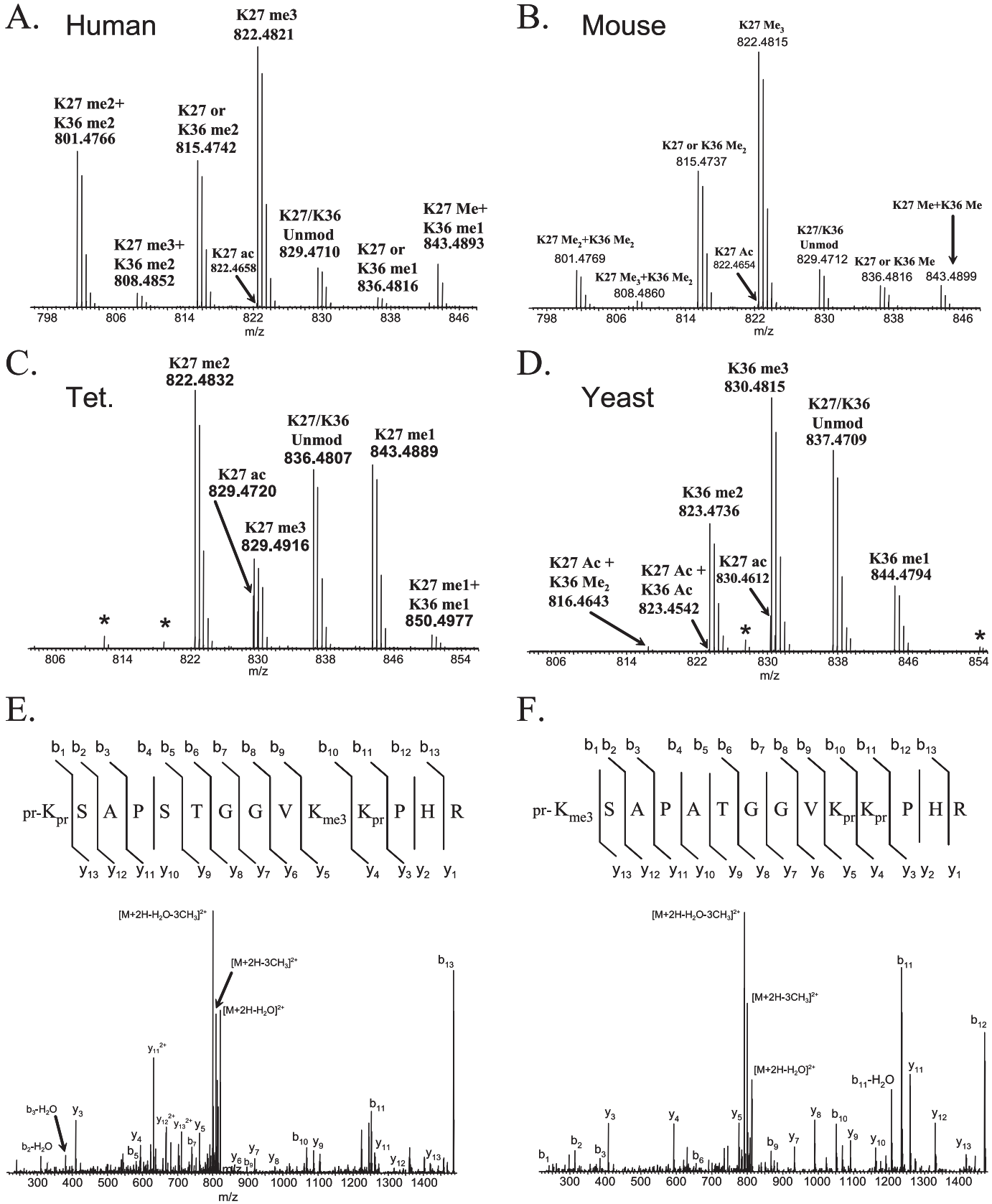

FIGURE 4. **Comparison of full mass spectra depicting H3K27 and Lys-36 modifications on the $[M + 2H]^{2+}$ peptide spanning residues 27–40 (KSAPATGGVKKPHR) from human HEK cells (*A*), mouse MEF cells (*B*), Tet (*C*), and yeast (KSAPSTGGKKPHR) (*D*).** *E*, MS/MS spectrum of the $[M + 2H]^{2+}$ 27–40 peptide containing H3K36me3 from the yeast sample. Propionyl groups (*pr*) are present on the N terminus, Lys-27 and Lys-37. *F*, MS/MS spectrum of the $[M + 2H]^{2+}$ 27–40 peptide containing H3K27me3 from the human sample. Propionyl groups (*pr*) are present on the N terminus, Lys-36 and Lys-37. b and y type ions are labeled in *E* and *F*. Ions arising from non-histone peptides or contamination are labeled with an *asterisk*.

TABLE 1

Post-translational modifications of H3

Data obtained from comparative analysis experiments conducted with tandem mass spectrometer (linear quadrupole ion trap-Fourier transform) on one H3 sample (Tet) and on two independent H3 samples isolated from yeast (*S. cerevisiae)*, mouse (NIH3T3 and MRL-MPJ spleen cells) and human (HEK293) cells. All modifications were localized by nano-liquid chromatography-MS/MS experiments.

Residue/peptide	Modification	Tet	Yeast	Mouse	Human
TK[4]QTAR	Monomethyl	+[a]	+	+	+
	Dimethyl	+	+	ND[b]	ND
	Trimethyl	+	+	ND	ND
	Acetyl[c]	**+[L]**	**ND**	**+[L]**	**+**
K[9]STGGKAPR	Monomethyl	ND	+[d]	+	+
	Dimethyl	ND	ND	+	+
	Trimethyl	+	ND	+	+
	Acetyl	+	+[V]	+[L]	+[L]
KSTGGK[14]APR	**Dimethyl**	**ND**	**+[L,V]**	**ND**	**ND**
	Acetyl	+	+[V]	+	+
K[9]STGGK[14]APR	Monomethyl + acetyl	ND	ND	+[V]	+
	Dimethyl + acetyl	ND	ND	+	+
	Trimethyl + acetyl	ND	ND	+	+
	Acetyl + acetyl	+	+[L]	+	+
K[18]QLATKAAR	Monomethyl	ND	+	+	+
	Acetyl	+	+	+	+
KQLATK[23]AAR	Monomethyl	ND	+	+[V]	+
	Acetyl	+	+	+	+
K[18]QLATK[23]AAR	Acetyl + acetyl	+	+	+	+
K[27]SAPATGGVKKPHR	Monomethyl	+[d]	ND	+	+
	Dimethyl	+[d]	ND	+	+
	Trimethyl	+[d]	+[Ld]	+	+
	Acetyl	+	+	+	+[L]
KSAPATGGVK[36]KPHR	Monomethyl	+	+	+	+
	Dimethyl	ND	+	+	+
	Trimethyl	ND	+	ND	ND
	Acetyl	+[L]	+[L]	ND	+[L]
K[27]SAPATGGVK[36]KPHR	Monomethyl + monomethyl	ND	ND	+	+
	Monomethyl + dimethyl	ND	ND	+	+
	Monomethyl + trimethyl	ND	ND	ND	+
	Monomethyl + acetyl	ND	ND	ND	+[V]
	Dimethyl + monomethyl	ND	ND	+	+
	Dimethyl + dimethyl	ND	ND	+	+
	Trimethyl + monomethyl	ND	ND	+	+
	Trimethyl + dimethyl	ND	ND	+[V]	+
	Acetyl + dimethyl	ND	+[L]	ND	ND
	Acetyl + acetyl	ND	+[L]	ND	ND
YQK[56]STELLIR	Monomethyl	+	ND	ND	+
	Trimethyl	**ND**	**ND**	**ND**	**+[V]**
	Acetyl	+	+	ND	+[L]
K[64]LPFQR	Monomethyl	ND	ND	ND	+
	Acetyl	**ND**	**+[L]**	**ND**	**ND**
EIAQDFK[79]TDLR	Monomethyl	+	+	+	+
	Dimethyl	ND	+	+	+
	Trimethyl	ND	+	ND	ND
	Acetyl	**ND**	**ND**	**ND**	**+[L]**
VTIMPK[122]DIQLAR	Monomethyl	ND	ND	ND	+

[a] +, detected in both experiments; +[L], detected, but low abundance; and +[V], detected in only one experiment.
[b] ND, not detected.
[c] Novel modifications are highlighted in bold letters.
[d] Known modifications that were not detected previously in one organism.

and functionally well characterized methylation modifications among species (a summary is depicted in Fig. 5*A*, abundant monomethyl = *one red circle*, dimethyl = *two red circles*, and trimethyl = *three red circles*). Especially di- and trimethylation of H3K4 and H3K36, as well as dimethylation of H3K79, have been shown to be associated with gene activation (16), whereas methylation of Lys-9 and Lys-27 on H3, especially di- and trimethylation, are found to be involved in transcriptional silencing (32). As depicted in Fig. 5*A*, H3 (H3.3) from yeast was enriched in activation marks (*light gray boxes*). In contrast, histone H3 (mixture of H3.1, H3.2, and H3.3) from both mammals was enriched in silencing modifications (*dark gray boxes*). Interestingly, we observed a pattern of abundant methyl marks on histone H3 (mixture of H3.2 and H3.3) from Tet that shows characteristics from both yeast and mammals. Tet H3 had abundant K4me1, -me2, and -me3 (active marks), but also K9me3 and to our surprise K27me1, -me2, and -me3; the latter have not been described previously (see Table 1). These results suggest that an evolutionary trend might exist from activation toward silencing PTMs, which also correlates with the appearance of different histone H3 variants.

To extend the analysis of abundant PTMs among different species, we additionally used antibodies against the MS/MS-detected abundant methyl marks (see Fig. 5*A*). Fig. 5*B* shows the results of one representative (of three) "zoo-blot." Histones were acid-extracted and subjected to immunoblotting, and recombinant H3 from *Xenopus* was used as a negative control. We also included histones from *Drosophila melanogaster* (fly), as an additional multicellular eukaryote that contains only H3.2 and H3.3, but not H3.1. With this immunobiochemical method, we observed a similar pattern of abundant methyl marks, as we have found by MS/MS. Significantly, H3K4me2

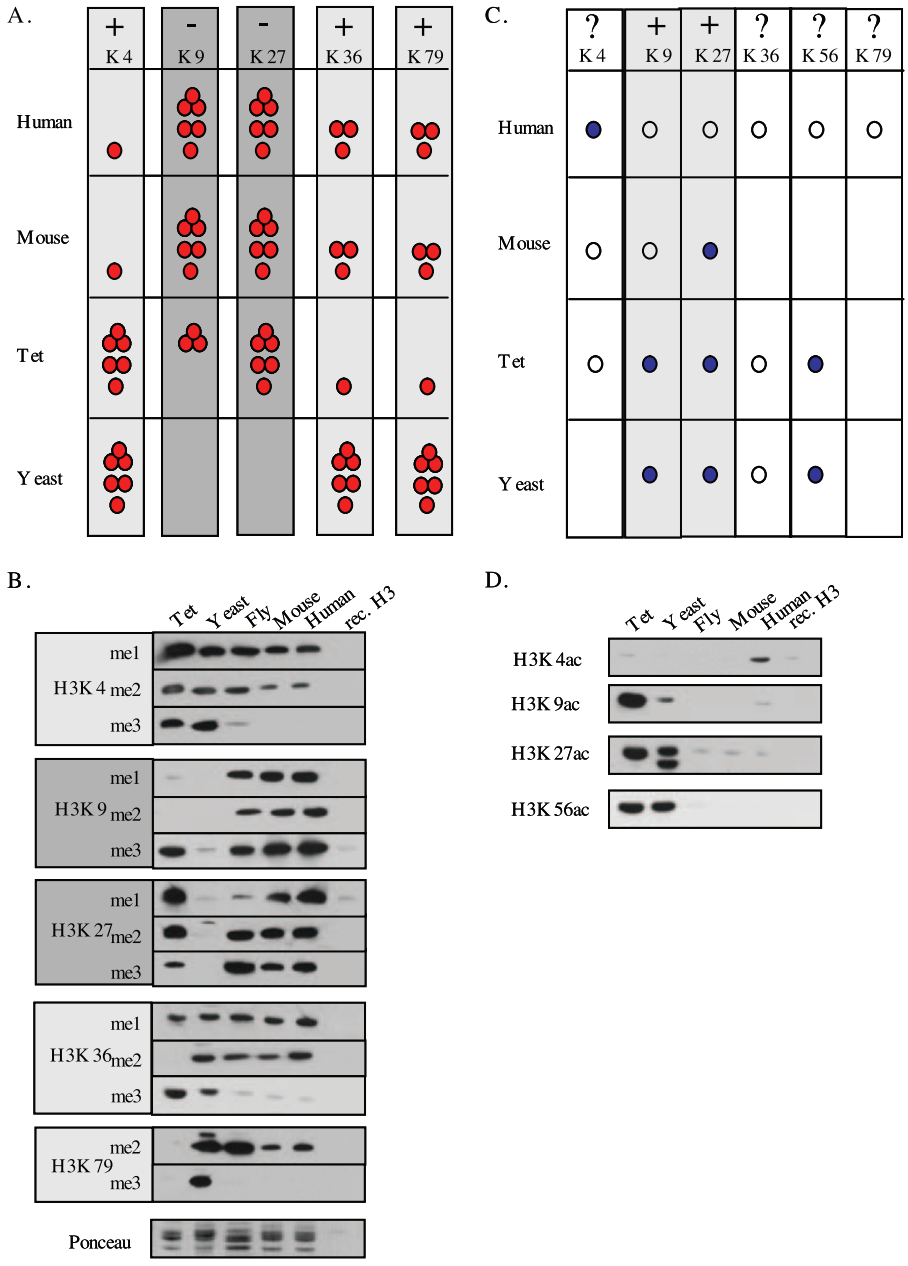

FIGURE 5. *A*, comparison of abundant histone H3 methylation modifications at Lys-4, Lys-9, Lys-27, Lys-36, and Lys-79 from human, mouse, Tet, and yeast. H3 from human and mouse is highly modified at Lys-9, whereas Tet and yeast are more highly modified at Lys-4. Tet is hypomodified at Lys-79, whereas yeast is the only species to contain Lys-79 trimethylation and no Lys-27 modifications. *B*, immunoblots with acid-extracted histones from different species as marked *above* and several antibodies as described on the *left side* are shown. 1 μg of recombinant histone H3 from *Xenopus* (H3.2) was used as negative control (*right lane*). Staining of these blots with Ponceau was performed to ensure equal loading (*blot below*). Modifications that have been associated with transcriptional activation are *boxed* in *light gray*, and modifications that are associated with transcriptional silencing are shown in *dark gray boxes*. *C*, comparison of histone H3 acetylation modifications at Lys-4, Lys-9, Lys-27, Lys-36 (52), Lys-56, and Lys-79 from human, mouse, Tet, and yeast (*filled blue circle*, abundant acetyl marks; *circle with blue dots*, not abundant acetyl marks). *D*, immunoblots stripped from *B* were used for analysis of acetylation profile among species.

and H3K4me3, associated with transcriptional activation, were very abundant in Tet and yeast when compared with mouse and human histone H3. Somewhat surprisingly, H3K4me2 and H3K4me3 were not detected by MS/MS in mouse and human samples but were observed by immunoblotting, albeit at a low level when compared with Tet and yeast. In contrast to our MS analyses, we could not detect the H3K9me1 mark in yeast by immunoblotting suggesting that this mark was either present only at very low abundance or that this modification could not be detected with antibodies because of detection or epitope occlusion issues. In contrast, and in accordance with our MS results, fly, mouse, and human H3 contained high levels of PTMs associated with transcriptional silencing, such as H3K9- and H3K27me1, -me2, and -me3. In conclusion, most of the data we obtained by MS/MS analyses of histone H3 from different organisms could be reproduced by immunobiochemical experiments and strengthen our observation that organisms differ in their histone H3 lysine methylation profiles dramatically. In general, less complex eukaryotes contain more marks associated with transcriptional activation, whereas more complex eukaryotes tend to have more modifications that are involved in gene silencing.

Fig. 5*C* depicts some of the acetylation marks found on histone H3 among different species by MS/MS (abundant acetyl mark = *filled blue circle*, less abundant acetyl mark = *circle with blue dots*). To further evaluate the H3 acetylation data we obtained by MS/MS, we performed zoo-blots with antibodies against some of these acetylation marks as described in Fig. 5*D*. We developed an antibody against the novel H3K4ac mark that unfortunately showed cross-reactivity against H4K5ac (for the characterization of this antibody see supplemental Fig. S3). Nevertheless, our immunobiochemical analysis further supported the observed acetylation pattern by MS/MS. Interestingly, we again noticed that unicellular eukaryotes contained more acetylation marks associated with gene transcription, such as H3K9ac and H3K27ac, whereas the novel H3K4ac was only detected on human H3. It should be noted that some antibodies (anti-H3K27me3, anti-H3K36me1, anti-K36me2, and anti-K36me3) gave reproducibly varied results in these immunobiochemical analyses.

Novel PTMs and PTM Patterns on Histone H3—Several novel PTMs were identified by MS/MS, which are summarized in Fig. 6. Most of these novel PTMs were of low abundance and are therefore depicted as *dotted circles* (*blue* = acetylation, and

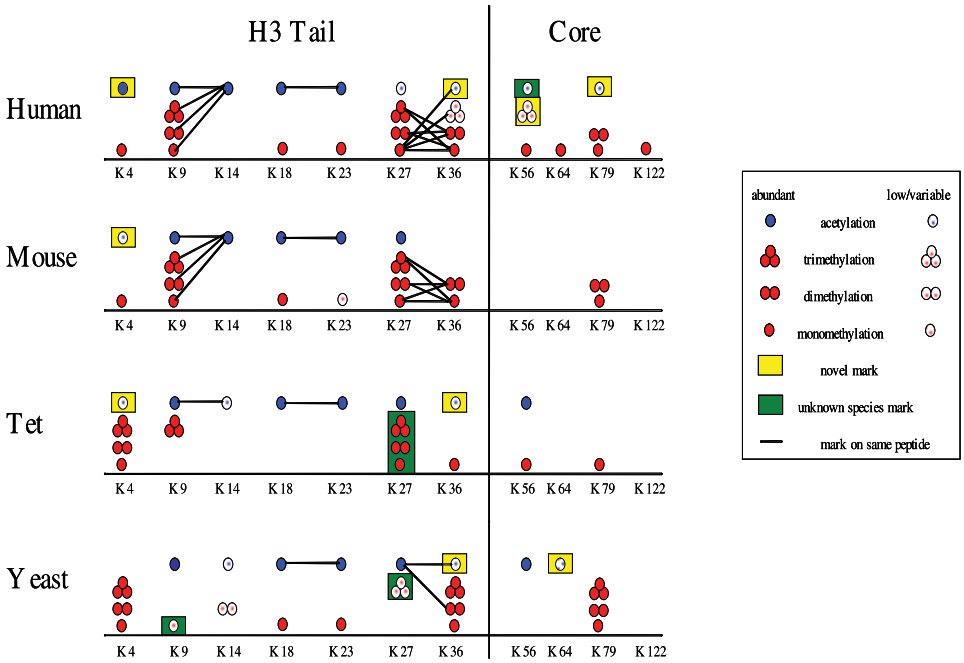

FIGURE 6. **Summary of histone H3 modifications detected in MS/MS experiments from human, mouse, Tet, and yeast.** Methyl marks are depicted as *red circles* (*one circle*, mono-methyl; *two circles*, di-methyl; and *three circles*, tri-methyl), and acetylation marks are depicted as *blue circles*. *Colored circles* represent abundant marks, whereas *dotted circles* show modifications that are present in low abundance or that have been detected in only one of the two runs. Novel modifications are in *yellow boxes*, and modifications, which have not been described previously for a specific organism, but are otherwise known, are surrounded by *green boxes*.

red = methylation) and include the identification of the following acetylation marks: H3K4ac in all organisms except yeast, with the strongest signal in humans and H3K79ac only in humans. In addition to these results, the companion report (52) shows the existence of H3K36 acetylation as well. Supplemental Fig. S2 shows the MS/MS spectra of a peptide containing H3K4ac. These novel acetylation marks are particularly interesting, because they occur on lysine residues, which, when methylated, are associated with transcriptional activation. A differential read-out between methylation and acetylation of the same residue often results in a functional "switch" between transcriptional activation or silencing, as has been described for H3K9 and H3K27 (33). Our observation that H3K4 and H3K79 residues can also be acetylated at very low abundances might suggest that these acetylation marks counteract methylation of these residues leading to altered states of gene expression or participation in other downstream processes that remain poorly appreciated and poorly understood.

In addition to the novel acetylation marks mentioned above, we also identified novel methylation marks on H3: K9me1+K14me2 (supplemental data page 3), and K27me3 (supplemental data page 24) in yeast, K27me1, -me2, and -me3 in Tet (supplemental data pages 21–23), and K56me3 in humans (supplemental data page 5). All these marks, except Lys-27 methylation in Tet, were at a very low occupancy, which might explain why these PTMs have not been identified previously.

Next the methylation and acetylation patterns were compared among the selected species. As depicted in Fig. 6, many differences between PTM patterns exist among species. Interestingly, most lysine residues in humans can be either methylated or acetylated, with exceptions being H3K14, H3K64, and H3K122. Because it is possible that modifications might act together to achieve a particular biological outcome, we sought to identify PTM patterns on H3 peptides that contained more than one modifiable residue. PTMs that co-exist on the same peptide are schematically shown in Fig. 6 by *line connections*. This scheme shows that the complexity of PTM patterns on histone H3 increases from unicellular eukaryotes to mammals. Modest differences between mouse and human H3 PTM patterns may reflect subtle differences in epigenetic signatures in mice that are not always mirrored in humans.

Comparison of PTMs on Histone H4 among Unicellular Eukaryotes and Mammals—Because our studies showed that species differ in their PTMs and PTM pattern for histone H3, we wondered if this was the result of the development of more enzymes that "write" and proteins that "read" these marks and/or the difference between the numbers of functionally distinct histone H3 variants among organisms. To investigate which of these possibilities or a combination of both may be correct, we turned our attention to histone H4. H4 was selected in part, because it is the only histone protein of which no other histone variants have been identified in these organisms. Histone H4 is highly conserved in its amino acid sequence among species (supplemental Fig. S4*A*) and similar to other species as compared by sequence identity analysis (supplemental Fig. S4*B*).

As described above for histone H3, we subjected RP-HPLC-purified histone H4 from the four different organisms to MS/MS analysis following propionylation and trypsin digestion. Fig. 7 displays the summed full mass spectra around the 745–780 m/z range (for human (*A*), mouse (*B*), and yeast (*D*)), which show the $[M + 2H]^{2+}$ peptides spanning residues 4–17 and the 840–872 range for Tet (*C*) (peptide-spanning residues 1–16). The peptide-spanning residues 4–17 (1–16 Tet) contain four acetylation sites (Lys-5, Lys-8, Lys-12, and Lys-16 in human, mouse, and yeast, and Lys-4, Lys-7, Lys-11, and Lys-15 in Tet). As seen in Fig. 7*A*, the most abundant form of the 4–17 peptide for the human H4 is the unmodified form, followed by the one acetyl form, then two, three, and four acetyl forms. Remarkably, a very similar pattern of acetyl PTMs can also be observed on H4 from the other species (Fig. 7, *B–D*). Fig. 7*E* displays a tandem MS of the human sample seen in Fig. 7*A*. The fragment b and y ions show that all lysine residues on this 4–17 peptide are acetylated. MS/MS experiments also determined that the major pattern of acetylation on human, mouse, and yeast H4 peptides essentially followed the trend that mono-acetylation is mainly at Lys-16, di-acetylation mainly at Lys-16 and Lys-12, tri-acetylation at Lys-16, Lys-12, and Lys-8 or

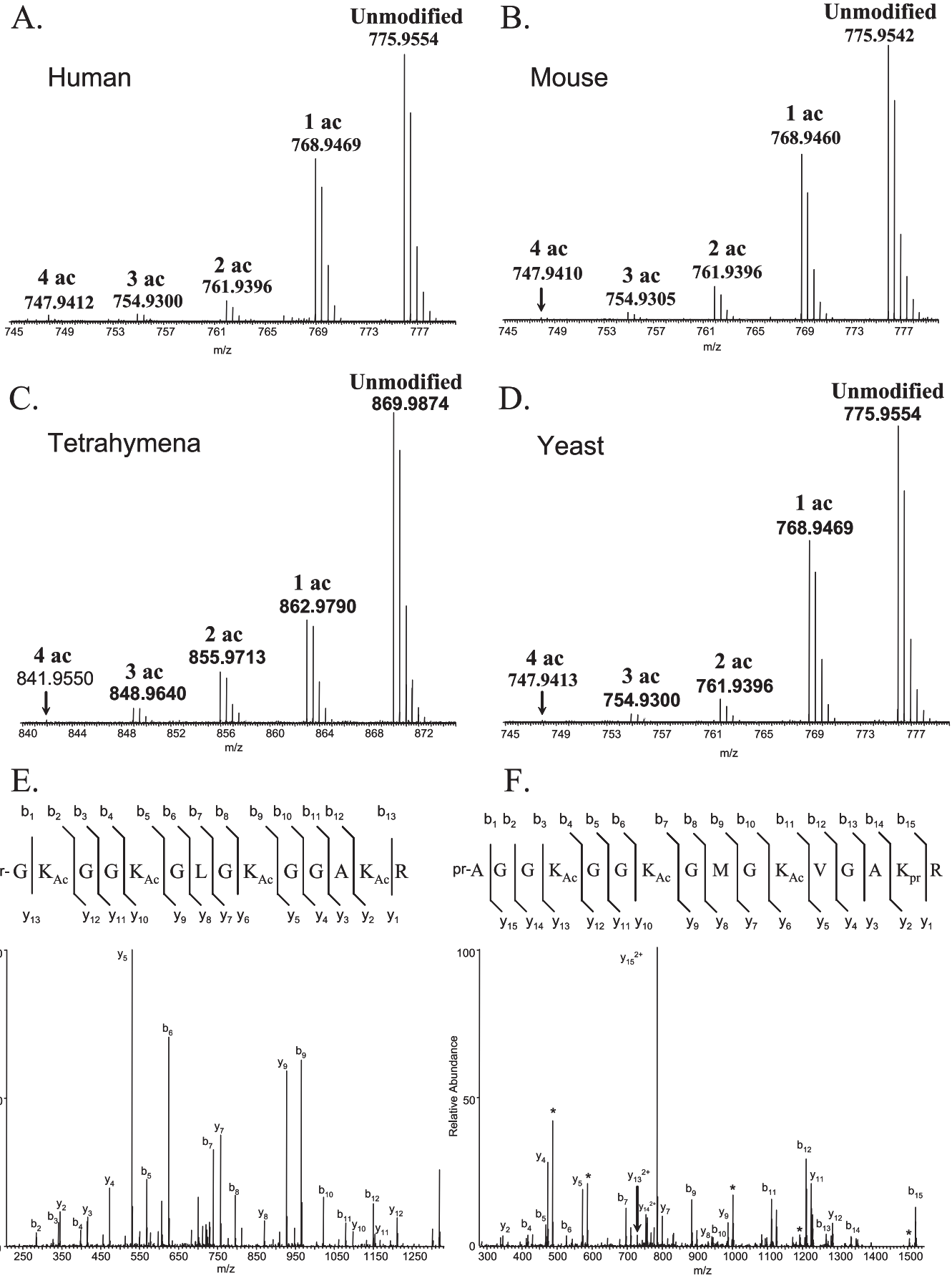
A.
Human
Unmodified
775.9554
1 ac
768.9469
2 ac
761.9396
3 ac
754.9300
4 ac
747.9412
745 749 753 757 761 765 769 773 777
m/z
B.
Mouse
Unmodified
775.9542
1 ac
768.9460
2 ac
761.9396
3 ac
754.9305
4 ac
747.9410
745 749 753 757 761 765 769 773 777
m/z
C.
Tetrahymena
Unmodified
869.9874
1 ac
862.9790
2 ac
855.9713
3 ac
848.9640
4 ac
841.9550
840 844 848 852 856 860 864 868 872
m/z
D.
Yeast
Unmodified
775.9554
1 ac
768.9469
2 ac
761.9396
3 ac
754.9300
4 ac
747.9413
745 749 753 757 761 765 769 773 777
m/z
E.
pr-G K_Ac G G K_Ac G L G K_Ac G G A K_Ac R
Relative Abundance
m/z
F.
pr-A G G K_Ac G G K_Ac G M G K_Ac V G A K_pr R
Relative Abundance
m/z

TABLE 2

Post-translational modifications of H4

Data obtained from comparative analysis experiments conducted with tandem mass spectrometer (linear quadrupole ion trap-Fourier transform) on one H4 sample (Tet, yeast, mouse (NIH3T3 and MRL-MPJ), spleen cells, and human HEK293 cells). All modifications were localized by nano-liquid chromatography-MS/MS experiments.

Residue/peptide	Modification	Tet	Yeast	Mouse	Human
$GK^{5}GGK^{8}GLGK^{12}GGAK^{16}R$	1 Acetyl (mostly Lys-16)	+[a]	+	+	+
	2 Acetyl (mostly Lys-16 and Lys-12)	+	+	+	+
	3 Acetyl (mostly Lys-16, Lys-12, and Lys-8)	+	+	+	+
	4 Acetyl (Lys-16, Lys-12, Lys-8, and Lys-5)	+	+	+	+
$K^{20}VLR$	Monomethyl	ND[b]	+[c]	+	+
	Dimethyl	ND	+[c]	+	+
	Trimethyl	ND	ND	+	+
	Acetyl[d]	ND	+[c]	ND	ND
$DNIQGITK^{31}PAIR$	Monomethyl	ND	+	+	+
	Dimethyl	ND	ND	ND	ND
	Trimethyl	ND	ND	ND	ND
	Acetyl	ND	ND	+	+

[a] +, detected.
[b] ND, not detected.
[c] Known modifications that were not detected previously in one organism.
[d] Novel modifications are highlighted in bold letters.

Lys-5, and tetra-acetylation at Lys-16, Lys-12, Lys-8, and Lys-5 (supplemental data pages 29–31), although other lower level combinations exist. This pattern has been previously reported and termed the "zipper acetylation" motif (34). Overall acetylation levels of histone H4 from Tet were similar to other species, but higher abundant different combinations of acetylation were also detected as shown in Fig. 7*F* with a tri-acetylated 1–16 peptide containing acetyl marks at Lys-4, Lys-7, and Lys-11 as revealed by the numerous b and y type fragment ions and shown in supplemental data pages 32–34 (Tet also contains no N-terminal acetylation, which is different from all other organisms). A similar methylation pattern was observed on H4K20 as was seen on H3K9, because human and mouse contained much higher amounts of this repressive mark (especially dimethylation) than yeast and Tet, which were mostly unmodified (supplemental Fig. S5). The results of our tandem MS experiments are summarized in Table 2 and are depicted in Fig. 8. Novel histone H4 modification sites such as H4K20ac in yeast and others are also displayed as supplemental data pages 25–28. The lower number of PTMs and the consistency observed in the acetylation patterns of H4 was not seen in our analysis of H3, suggesting that histone H3 displays more organismal differences ("epigenetic signatures") than its dimeric partner H4.

DISCUSSION

Many studies to date have analyzed PTMs of histones, in particular H3, in different organisms by separate techniques. Although these studies enriched and promoted the understanding of the function of histone marks in biological pathways, to our knowledge, few analyses have sought to provide a systematic comparison of histone H3 and H4 PTMs among different model organisms. Several conclusions can be drawn from our study: 1) Overall, the H3 PTM profile among unicellular eukaryotes and mammals differs significantly, and the PTM patterns become more complex from "simple" to more "complex" eukaryotes. 2) With few exceptions, most lysines in H3 show the general tendency to be either methylated or acetylated although to varying extents depending upon site and organism. The significance of extremely low levels of one or the other of these modifications remains unclear. The argument that low levels of PTM reflect physiologically irrelevant "noise" that is detected by high sensitivity methods, cannot be ruled out. 3) Unicellular eukaryotes, notably yeast and Tet, exhibit more modifications associated with transcriptional activation, whereas mammals have more marks involved in gene silencing. This observation mirrors the studies that show that the majority of the genome in unicellular eukaryotes is transcriptionally competent, whereas >60% of the mammalian genome is permanently silenced, and only ~3% of its DNA encodes structural genes (reviewed in Ref. 35). 4) Surprisingly, we identified by MS what appear to be species-specific PTMs such as: H3K14me2 (yeast), H3K56me3 (human), H3K64me1 (human), H3K64ac (yeast), H3K79me3 (yeast), H4K20ac (yeast), H3K79ac (human), and H3K122me1 (human). 5) We also found novel modifications: H3K4ac (human, and to a lesser extent, mouse, and Tet), H3K56me3 (human), H3K64ac (yeast), H4K20ac (yeast), and H3K79ac (human). Collectively, these data argue against there being a universal histone code and underscore the need to avoid general conclusions obtained from one organism.

One important difference between budding yeast and other eukaryotes is that, although it contains well defined "silenced" chromatin (36), several hallmark features of constitutive heterochromatin found in more complex eukaryotes (*e.g.* the presence of HP1 that has been shown to specifically bind to H3K9me2 and -me3 with its chromodomain (37–39)) have yet to be observed (40). This finding correlates well with the presence of only an "active" H3.3 variant in yeast. On the other hand, Tet has been reported to contain H3K9 methylation and the expression of an HP1-like binding protein (Pdd1p) (41, 42).

FIGURE 7. **Comparison of full mass spectra depicting histone H4 acetylation modifications on the $[M + 2H]^{2+}$ peptide spanning residues 4–17 (GKGGKGLGKGGAKR) from human HEK cells (*A*), mouse MEF cells (*B*), Tet (1–16 residue peptide (*C*), AGGKGGKGMGKVGAKR), and yeast (*D*).** *E*, MS/MS spectrum of the $[M + 2H]^{2+}$ 4–17 peptide of H4 containing tetra-acetylation at Lys-5, Lys-8, Lys-12, and Lys-16 from the human sample. A propionyl group (*pr*) is present on the N terminus. *F*, MS/MS spectrum of the $[M + 2H]^{2+}$ 1–16 peptide of H4 containing acetylation at Lys-4, Lys-7, and Lys-11 from the Tet sample. Propionyl groups (*pr*) are present on the N terminus and on Lys-15. b and y type ions are labeled in *E* and *F*. *Asterisks*, ions that indicate that an isomeric peptide containing three acetyl groups at different residues is also present.

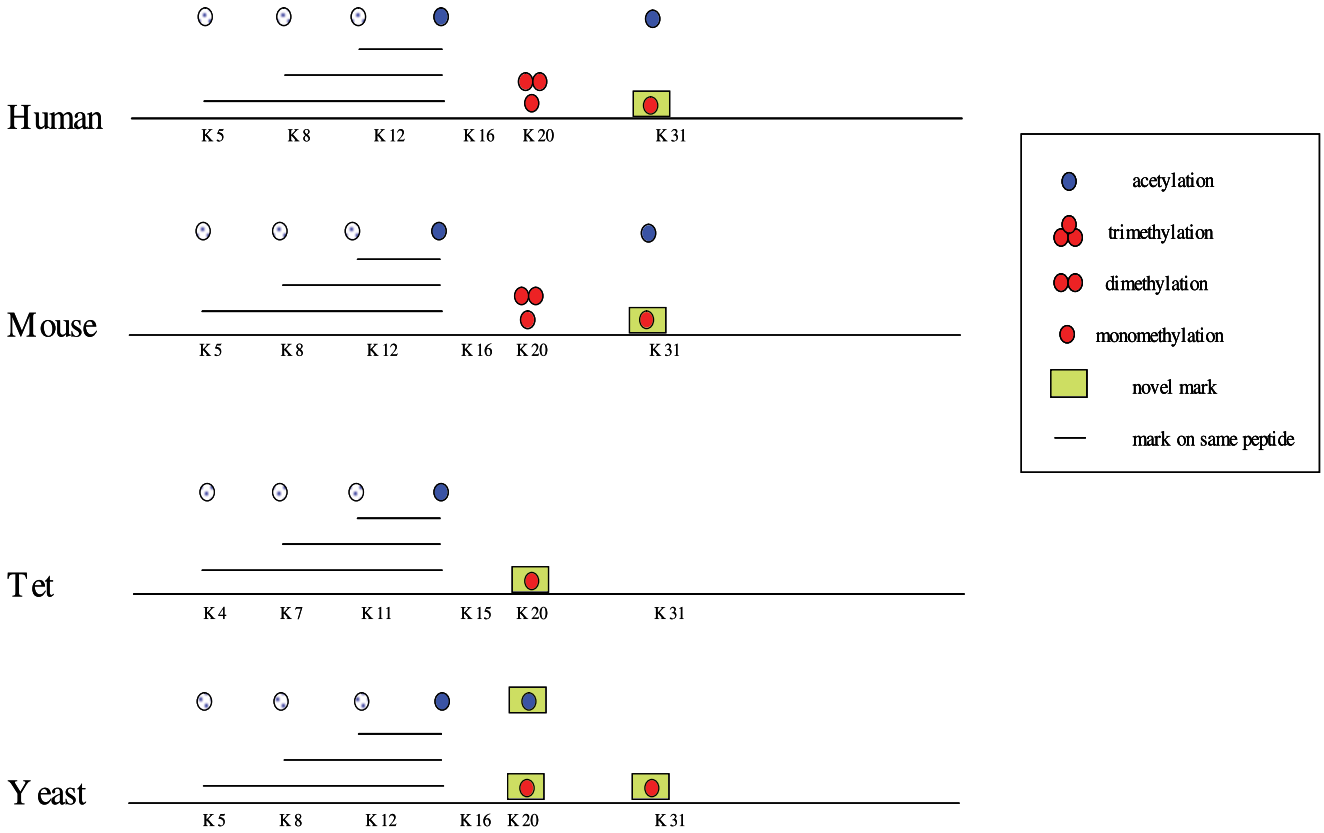

FIGURE 8. **Summary of histone H4 modifications detected by MS/MS from human, mouse, Tet, and yeast samples.** Methyl marks are depicted as *red circles* (*one circle*, me1; *two circles*, me2; and *three circles*, me3) and acetylation marks are depicted as *blue circles*. *Colored circles* represent abundant marks, whereas *dotted circles* show modifications that are present in low abundance or which have been detected in only one of the two runs. Novel modifications are in *yellow boxes*.

Interestingly, this organism contains in addition to H3.3 another H3 variant (H3.2), suggesting that there might be correlations between the occurrence of heterochromatin and specific H3 variants each decorated with PTM signatures. The finding of H3K9me1 and H4K20me2 in budding yeast was unexpected, given the absence of known writers and readers for this mark in this organism. Whether novel enzyme and effector systems exist or whether certain known activities are simply more promiscuous in yeast due to differences in gene regulation or nuclear architecture as compared with other organisms is not known. Similarly, the amino acid sequence adjacent to H3K79 is not conserved in Tet (see supplemental Fig. S1*A*), and additionally, a Dot1 homologue, the enzyme that methylates H3K79 (29) is not found in Tet.[11] It will be of interest to identify the enzyme responsible for H3K79me1 in Tet.

A different picture emerged from our data about lysine modifications and their pattern on histone H4 among organisms. Notably, the H4 PTM profile and pattern are not as complex as those shown for histone H3 among species. Although we also identified novel marks on H4 (H4K20Me2 and H4K20ac in yeast, H4K31me1 in all organisms examined except yeast, and H4K31ac in mouse and humans), the number of PTMs is much lower for H4 than H3.

This may indicate that a combination of two different mechanisms has evolved. We suggest that, not only have more enzymes that set these marks ("writers") and proteins that specifically bind to these modifications ("readers") evolved during evolution, but also the emergence of more functionally distinct histone H3 variants may play a role in the here observed modification differences among organisms (reviewed in Ref. 43). It has been shown that more histone-modifying enzymes and proteins that recognize these marks have evolved over time (reviewed in Ref. 44). For example, the number of genes encoding for SET domain-containing proteins differs greatly among different species. It has been estimated that the human and mouse genomes each encode ≥50 predicted SET domain proteins (45), whereas *Schizosaccharomyces pombe* contains only ~10 putative SET domain histone methyltransferases, and *S. cerevisiae* has no more than 7 (27). On the other hand, besides an increase in the number of genes encoding histone-modifying proteins, there has also been a shift in the number of histone variants during the evolution from unicellular toward multicellular eukaryotes. Over the last several years, many studies have shown that variants from different histone families, most prominently the family of H2A histones, play distinct roles in diverse biological processes (reviewed in Refs. 46 and 47).

Interestingly, even though in contrast to H2A variants, which differ greatly in their primary sequence, H3 variants contain only small amino acid changes between each other, and their function might also be separate. The H3.3 variant, expressed replication independently, has been shown to be associated with transcriptional activation by its specific PTM profile (18, 19, 48) and biochemical experiments, such as chromatin immu-

[11] Y. Liu, personal communication.

noprecipitation and chromatin immunoprecipitation-chip assays (49–51). Another study in *Arabidopsis thaliana* showed that H3.2 contains a modification profile linking it to transcriptional repression (48). Previously, we showed by quantitative MS/MS analyses that, not only H3.2, but also the mammalian-specific H3.1 variant, have modifications associated with gene repression, but that these variants are enriched in different "silencing" PTMs, suggesting that these highly similar H3 variants might have separate biological functions (19).

In summary, our data suggest that, during evolution, in addition to the emergence of more histone-modifying enzymes, the appearance of more histone H3 PTMs and H3 variants led to a more complex and diverse histone code over time (47–49). Taken together, our findings underscore the need to use multiple approaches to identify PTMs with confidence, particularly those that occur at low levels.

Acknowledgments—We thank members of the Allis, Hunt, and Strahl laboratories for insightful discussions. We especially thank Christian Janzen for critical review of the manuscript. We also want to thank Upstate Biotechnologies for their continuous support.

REFERENCES

1. Cheung, P., and Lau, P. (2005) *Mol. Endocrinol.* **19,** 563–573
2. Felsenfeld, G., and Groudine, M. (2003) *Nature* **421,** 448–453
3. Iizuka, M., and Smith, M. M. (2003) *Curr. Opin. Genet. Dev.* **13,** 154–160
4. Berger, S. L. (2002) *Curr. Opin. Genet. Dev.* **12,** 142–148
5. Zhang, L., Eugeni, E. E., Parthun, M. R., and Freitas, M. A. (2003) *Chromosoma* **112,** 77–86
6. Bernstein, E., and Allis, C. D. (2005) *Genes Dev.* **19,** 1635–1655
7. Whetstine, J. R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M., and Shi, Y. (2006) *Cell* **125,** 467–481
8. Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P., and Zhang, Y. (2006) *Nature* **439,** 811–816
9. Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., Cole, P. A., and Casero, R. A. (2004) *Cell* **119,** 941–953
10. Cheung, P., Allis, C. D., and Sassone-Corsi, P. (2000) *Cell* **103,** 263–271
11. Strahl, B. D., and Allis, C. D. (2000) *Nature* **403,** 41–45
12. Turner, B. M. (2000) *BioEssays* **22,** 836–845
13. Jenuwein, T., and Allis, C. D. (2001) *Science* **293,** 1074–1080
14. Nicolas, E., Roumillac, C., and Trouche, D. (2003) *Mol. Cell. Biol.* **23,** 1614–1622
15. Wang, H., Cao, R., Xia, L., Erdjument-Bromage, H., Borchers, C., Tempst, P., and Zhang, Y. (2001) *Mol. Cell* **8,** 1207–1217
16. Bannister, A. J., Schneider, R., Myers, F. A., Thorne, A. W., Crane-Robinson, C., and Kouzarides, T. (2005) *J. Biol. Chem.* **280,** 17732–17736
17. Feng, Q., Wang, H., Ng, H. H., Erdjument-Bromage, H., Tempst, P., Struhl, K., and Zhang, Y. (2002) *Curr. Biol.* **12,** 1052–1058
18. McKittrick, E., Gafken, P. R., Ahmad, K., and Henikoff, S. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 1525–1530
19. Hake, S. B., Garcia, B. A., Duncan, E. M., Kauer, M., Dellaire, G., Shabanowitz, J., Bazett-Jones, D. P., Allis, C. D., and Hunt, D. F. (2006) *J. Biol. Chem.* **281,** 559–568
20. Medzihradszky, K. F., Zhang, X., Chalkley, R. J., Guan, S., McFarland, M. A., Chalmers, M. J., Marshall, A. G., Diaz, R. L., Allis, C. D., and Burlingame, A. L. (2004) *Mol. Cell. Proteomics* **3,** 872–886
21. Recht, J., Tsubota, T., Tanny, J. C., Diaz, R. L., Berger, J. M., Zhang, X., Garcia, B. A., Shabanowitz, J., Burlingame, A. L., Hunt, D. F., Kaufman, P. D., and Allis, C. D. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103,** 6988–6993
22. Garcia, B. A., Busby, S. A., Shabanowitz, J., Hunt, D. F., and Mishra, N. (2005) *J. Proteome Res.* **4,** 2032–2042
23. Syka, J. E., Marto, J. A., Bai, D. L., Horning, S., Senko, M. W., Schwartz, J. C., Ueberheide, B., Garcia, B., Busby, S., Muratore, T., Shabanowitz, J., and Hunt, D. F. (2004) *J. Proteome Res.* **3,** 621–626
24. Martin, S. E., Shabanowitz, J., Hunt, D. F., and Marto, J. A. (2000) *Anal. Chem.* **72,** 4266–4274
25. Hake, S. B., Garcia, B. A., Kauer, M., Baker, S. P., Shabanowitz, J., Hunt, D. F., and Allis, C. D. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102,** 6344–6349
26. Bernstein, B. E., Humphrey, E. L., Erlich, R. L., Schneider, R., Bouman, P., Liu, J. S., Kouzarides, T., and Schreiber, S. L. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 8695–8700
27. Briggs, S. D., Bryk, M., Strahl, B. D., Cheung, W. L., Davie, J. K., Dent, S. Y., Winston, F., and Allis, C. D. (2001) *Genes Dev.* **15,** 3286–3295
28. Strahl, B. D., Ohba, R., Cook, R. G., and Allis, C. D. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 14967–14972
29. van Leeuwen, F., Gafken, P. R., and Gottschling, D. E. (2002) *Cell* **109,** 745–756
30. Lachner, M., and Jenuwein, T. (2002) *Curr. Opin. Cell Biol.* **14,** 286–298
31. Chow, J. C., and Brown, C. J. (2003) *Cell Mol. Life Sci.* **60,** 2586–2603
32. Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2002) *Genes Dev.* **16,** 2893–2905
33. Yang, Y., Li, T., Vu, T. H., Ulaner, G. A., Hu, J. F., and Hoffman, A. R. (2003) *Endocrinology* **144,** 5658–5670
34. Zhang, K., Williams, K. E., Huang, L., Yau, P., Siino, J. S., Bradbury, E. M., Jones, P. R., Minch, M. J., and Burlingame, A. L. (2002) *Mol. Cell Proteomics* **1,** 500–508
35. Jenuwein, T., Laible, G., Dorn, R., and Reuter, G. (1998) *Cell Mol. Life Sci.* **54,** 80–93
36. Rusche, L. N., Kirchmaier, A. L., and Rine, J. (2003) *Annu. Rev. Biochem.* **72,** 481–516
37. Jacobs, S. A., and Khorasanizadeh, S. (2002) *Science* **295,** 2080–2083
38. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001) *Nature* **410,** 116–120
39. Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T. (2001) *Nature* **410,** 120–124
40. Avramova, Z. V. (2002) *Plant Physiol.* **129,** 40–49
41. Taverna, S. D., Coyne, R. S., and Allis, C. D. (2002) *Cell* **110,** 701–711
42. Madireddi, M. T., Coyne, R. S., Smothers, J. F., Mickey, K. M., Yao, M. C., and Allis, C. D. (1996) *Cell* **87,** 75–84
43. Hake, S. B., and Allis, C. D. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103,** 6428–6435
44. Lachner, M., O'Sullivan, R. J., and Jenuwein, T. (2003) *J. Cell Sci.* **116,** 2117–2124
45. Kouzarides, T. (2002) *Curr. Opin. Genet. Dev.* **12,** 198–209
46. Kamakaka, R. T., and Biggins, S. (2005) *Genes Dev.* **19,** 295–310
47. Pusarla, R. H., and Bhargava, P. (2005) *FEBS J.* **272,** 5149–5168
48. Johnson, L., Mollah, S., Garcia, B. A., Muratore, T. L., Shabanowitz, J., Hunt, D. F., and Jacobsen, S. E. (2004) *Nucleic Acids Res.* **32,** 6511–6518
49. Ahmad, K., and Henikoff, S. (2002) *Mol. Cell* **9,** 1191–1200
50. Daury, L., Chailleux, C., Bonvallet, J., and Trouche, D. (2006) *EMBO Rep.* **7,** 66–71
51. Jin, C., and Felsenfeld, G. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103,** 574–579
52. Morris, S. A., Rao, B., Garcia, B. A., Hake, S. B., Diaz, R. L., Shabanowitz, J., Hunt, D. F., Allis, C. D., Lieb, J. D., and Strahl, B. D. (2007) *J. Biol. Chem.* **282,** 7632–7640