In situ detection of histone variants and modifications in mouse brain using imaging mass spectrometry

Shibojyoti Lahiri[1#], Na Sun[2#],Victor Solis-Mezarino[3], Andreas Fedisch[1], Jovica Ninkovic[4], Annette Feuchtinger[2], Magdalena Götz[4], Axel Walch[2], Axel Imhof[1*]

[1]Protein Analysis Unit (ZfP), Biomedical Center (BMC), Ludwig Maximilians University Munich, Großhaderner Str. 9, 82152 Planegg-Martinsried, Germany

[2]Research Unit Analytical Pathology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany

[3]Gene Center, Ludwig Maximilians University, Feodor-Lynen-Strasse 25 81377 Munich, Germany

[4]Institute of Stem Cell Research, GSF-National Research Center for Environment and Health, Ingolstaedter Landstrasse 1, Neuherberg, Germany and Munich Cluster for Systems Neurology (SyNergy) and Institute of Physiological Genomics, University of Munich, Schillerstrasse 46, 80336 München, Germany and Munich Cluster for Systems Neurology (SyNergy)

*Corresponding author - Prof. Dr. Axel Imhof, Munich Cluster for Systems Neurology (SyNergy) and BioMedical Center (BMC), Ludwig-Maximilians-University of Munich, Großhadernerstr. 9, 82152 Planegg-Martinsried, Germany

[#]These authors contributed equally to this work

e-mail: imhof@lmu.de

Tel.: +49 89 2180 75420, Fax.: +49 89 2180 75440

Total no. of words–6655

ABSTRACT

Histone posttranslational modifications and histone variants control the epigenetic regulation of gene expression and affect a wide variety of biological processes. A complex pattern of such modifications and variants defines the identity of cells within complex organ systems and can therefore be used to characterize cells at a molecular level. However, their detection and identification *in situ* has been limited so far due to lack of specificity, selectivity and availability of anti-histone antibodies. Here, we describe a novel MALDI imaging mass spectrometry (MALDI-IMS) based workflow, which enables us to detect and characterize histones by their intact mass and their correlation with cytological properties of the tissue using novel statistical and image analysis tools. The workflow allows us to characterize the *in situ* distribution of the major histone variants and their modification in the mouse brain. This new analysis tool is particularly useful for the investigation of expression patterns of the linker histone H1 variants for which suitable antibodies are so far not available.

Accepted Article

STATEMENT OF SIGNIFICANCE OF TH&E STUDY

The method introduced in this paper is appropriate for high-throughput, label-free, *in situ* detection of histones, thereby overcoming the limitations posed by antibodies and conventional mass spectrometric approaches. The workflow can, in principle, be applied to any kind of tissue, since it depends on inherent tissue properties and results can be used for prediction of the abundance of histone variants in diseased conditions.

Accepted Article

INTRODUCTION

In the nucleus of a eukaryotic cell, the majority of the DNA mediated processes occur on a tightly packaged array of many histone-DNA complexes called nucleosomes [1, 2]. The nucleosome constitutes of 147 DNA base pairs associated with two copies of each of the core histones H2A, H2B, H3 and H4. The histone H1 binds to the DNA that connects two nucleosomes and is therefore called linker histone [3, 4].The binding of H1 modulates higher order chromatin folding and is therefore considered an important contributor of overall chromatin structure [5]. Extensive research over the last couple of decades has shown that PTMs on the histones affect the fundamental organization of nucleosomes, thereby regulating the aforesaid cellular processes. Modifications at the flexible N-terminal tails of histones act as docking sites for reader proteins that specifically recognize the PTMs and/or their combinations. This in turn recruits other proteins, which act as chromatin modifiers/remodelers and regulates many biological functions [6, 7]. In addition, several histone variants exist that are expressed in a tissue specific manner or at different times of the cell cycle [8].

Epigenetic dysregulation contributes to a wide range of human diseases from neurodegeneration to cancer [9-12].In many cases, altered levels of histone variants and

aberrant combinations of histone modifications have been suggested to play a causal role in the disease. In fact, small molecule inhibitors of histone modifying enzymes or molecules that interfere with the binding of regulatory proteins to chromatin fibers are currently in clinical trials [13]. Understanding the functions of histones and the mechanisms through which they are propagated is, therefore, instrumental in comprehending general cellular physiology and the molecular basis of human diseases. Although a considerable body of knowledge exists regarding the biological functions of histones, the rate of progress is limited by the absence of suitable antibodies for this class of proteins. Anti-histone antibodies are limited in their specificity, selectivity and multiplicity for histone PTMs and variants [14, 15]. Subsequently, researchers have resorted to mass spectrometry as an alternative tool for identification of this diverse family of proteins [16, 17]. However, bulk mass spectrometry approaches lack spatial information thereby limiting its usefulness for the observation of histones in a tissue or organ. We have therefore used MALDI-IMS for *in situ* detection of histones. Based on the fact that few core histones were detectable using MALDI-IMS [18-21], we have developed a method that can be used as a reliable tool for the detailed characterization of core and linker histones in mammalian tissues.

MALDI-IMS combines mass spectrometry and imaging techniques to detect analytes (proteins, peptides, metabolites, etc.) directly from a tissue surface. A tissue section is usually covered with tiny matrix droplets that embed the analytes in their crystals, thereby maintaining intact spatial information. This technique is particularly useful for studying pathological conditions, since diagnostic information remains intact at the tissue level and multiple features can be extracted simultaneously [22]. In the current study, we have used mouse brain as a model organ to demonstrate simultaneous *in situ* detection of multiple histone variants and their modifications. Our findings are in good correlation with the current

literature regarding the abundance of histone variants and their modifications. Moreover, it allows us to precisely measure the distribution of the variants and modifications of the linker histone H1 in adult mouse brain at an unprecedented resolution and multiplicity.

Although in this study we focused on brain, the methods can also be applied to many different tissues. Furthermore, this study could potentially serve as the basis of high-throughput investigation into neuroepigenetics, which is being increasingly implemented in general neurological disorders, including neurodegenerative diseases.

MATERIALS AND METHODS

Materials

Sinapinic acid (SA), acetonitrile (ACN), poly-L-lysine, and Nonidet P-40(octylphenoxypolyethoxylethanol) were purchased from Sigma Aldrich (St. Louis,MO, USA).HPLC-grade water was obtained from Merck KGaA (Darmstadt, Germany). Trifluoroaceticacid (TFA) was purchased from Applied Biosystems (FosterCity, CA, USA) and Carl Roth GmbH (Germany).

Methods

Animals

Adult wild type male mice between 8-10 weeks of age were used for all experiments. The animals were transcardially perfused with 50 ml ice cold PBS following which the brains were removed and frozen in liquid nitrogen. The organs were then stored at -80°C until sectioning. Animal handling was in accordance with German and European guidelines for use

of animals for research purposes. Experiments were further approved by the institutional animal care committee.

MALDI-IMS

Sample preparation

The fresh frozen mice brains were cryo-sectioned to 12 μm thick sections using CM1950 cryostat (Leica Microsystems; Wetzlar, Germany). Sections were thaw-mounted onto conductive indium-tin-oxide (ITO) coated glass slides (Bruker Daltonik GmbH; Bremen, Germany) pretreated with poly-lysine (1:1 in water with 0.1% NP-40). The slides were then stored at -80°C until further analysis. Tissue sections were washed prior to matrix spraying to remove lipids and salts, which act as interferents in MALDI measurements. Washing and drying was performed as previously reported [23] with some modifications, briefly; sections were fixed by a sequential wash in 70% and 100% ethanol (for 30 s each) followed by a 2 min wash in Carnoy's fluid (6:3:1 Ethanol:Chloroform:Acetic acid, v/v/v) to remove lipids. Tissues were put back in 100% ethanol for 30 s to remove the excess chloroform. Further washing in dH_2O was performed to remove salts and the excess water was removed by washing them in 100% ethanol for 30 s. Subsequently, the slides were dried in vacuum at RT for at least 2 hrs before spraying the matrix.

Matrix (10 mg/ml SA in 60% ACN and 0.2% TFA (v/v)) was sprayed on the brain sections using an ImagePrep (Bruker Daltonik GmbH; Bremen,Germany) automated sprayer following the manufacturer's protocol with minor modifications.

Mass spectrometry measurements

Imaging experiments were performed in an Ultraflex III MALDI-TOF/TOF mass spectrometer (Bruker Daltonik GmbH; Bremen,Germany) equipped with a SmartBeam laser, in a positive linear mode with 50 µm spatial resolution and within a mass range of 2500-25000 Da. Each measurement was pre-calibrated externally using a commercial protein calibrant mixture (Bruker Daltonik GmbH; Bremen, Germany) spotted on the same target as the tissues at multiple positions.

After measurements, the matrix was removed by washing the slides in 70% ethanol and H&E counterstaining was performed. High-resolution images of the stained sections were obtained from Mirax Scan system (Carl Zeiss MicroImaging, Munich, Germany) and co-registered with MALDI-IMS data for histological correlation.

Generation of mass lists

All measured datasets were visualized and partly analyzed in FlexImaging 4.0 (Bruker Daltonik GmbH; Bremen, Germany). The mMass5.5.0software package [24] was used for peak picking on the exported spectra (TIC normalized) from FlexImaging software. Baseline correction was performed using a precision of 20 and relative offset value of 50 followed by a moving average smoothing algorithm with 2 cycles and m/z window of 5.0. Peak-picking was done using an s/n ratio of 2.0 and a relative intensity threshold of 0.5% with a picking height of 90. Spectra from the ROIs were recalibrated with respect to the average spectrum in mMass because of inherent mass shifts in MALDI-TOF measurements. The m/z lists with their corresponding intensities were exported as *.txt files and used for further analysis.

Unsupervised hierarchical clustering analysis (HCA)

Hierarchical clustering was performed in FlexImaging 4.0 on all the IMS datasets with a mass tolerance of 250 ppm, normal spatial smoothing and no entries in the mass whitelist and/or mass blacklist. Segmentation maps (clusters) were analyzed based on the histological features of the same section obtained from H&E stained images. Clusters centered on cerebellum and olfactory bulb (COC) and cerebral cortex (Ctx) were selected for further analysis.

Principal component analysis (PCA)

Spectra from COC and Ctx were exported to ClinProTools 3.0(Bruker Daltonik GmbH; Bremen, Germany) and PCA was performed. Unit variance scaling method was used for all analyses. Further analysis with the scores and loadings plot were also performed in ClinProTools 3.0.

Calculation of nuclear density

Images obtained by H&E staining were analyzed using the image analysis software Definiens TissueStudio3 (Definiens AG; Munich, Germany). All stained slides were scanned at 20x objective magnification using a Mirax Desk digital slide scanner (Carl Zeiss MicroImaging, Munich, Germany), and the resulting images were imported into the image analysis software. In a first step, ROIs were manually annotated in order to select differentially stained areas from all over the tissue, which also encompasses different cell types. Within each ROI, a specific rule set was then defined to detect and quantify the nuclei, based on morphology, size, pattern, shape, neighborhood and special color features. The quantified parameter 'nuclear density' was calculated according to the following formula:

Nuclear Density (%) = [Nuclear Sum Area (μm^2)/Total ROI Area (μm^2)]*100 (Equation 1)

Data analysis

To compare the experimental mass list with a theoretical mass list and to generate data for the correlation analysis we used two python based scripts ProtFinder and OrgPlotter respectively.

The tools, in their present form, are available upon request. The package, apart from the script, contains a theoretical mouse proteome mass list, an example input experimental file (in list format) and a PTM file containing me, ac and phos with their possible combinations and corresponding mass shifts and detailed instructions.

LC-MS/MS

Sample preparation

In order to validate the masses annotated in MALDI-IMS, we performed LC-MS/MS experiments on the proteins extracted in the matrix layer on the tissue sections ('matrix proteome'). We followed the same protocol for sample preparation as described above on serial mouse brain sections. Following the matrix spray, we extracted the matrix as described before[25]; in brief, the matrix layer on the tissue were extracted in two fractions containing increasing concentration of ACN(7.5% and 60% v/v) and 0.2% (v/v) TFA. Matrix extractions from 6 serial mouse brain sections were pooled together and 50% of the extracts were used for MS analysis.

In a bottom-up proteomic approach, the fractions were dried, resuspended in Laemmli buffer and SDS-PAGE was performed in a 4-20% gradient denaturing polyacrylamide gel followed by Coomassie staining. Regions corresponding to control histone bands (Figure S4) were excised from both 7.5% and 60% fractions. In-gel tryptic digestion was performed after destaining the gel pieces and propionylation of the lysine residues [17]. The peptides were

chromatographically purified using C18 stage tips (3M Empore™, Germany), solvent

evaporated and resuspended in 15 µl of 0.1% TFA.

LC-MS/MS measurements

For mass spectrometry measurements, 5 out of 15µl were injected in an Ultimate 3000 HPLC

system (Dionex) and separated at 295 nl/min in a 15 cm analytical C18 column (75µm ID

home-packed with ReproSil-Pur C18-AQ 2.4 µm from Dr. Maisch) with a 50 min gradient

from 4 to 48% ACN in 0.1% formic acid. The effluent from the HPLC was directly electro-

sprayed into a LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific). The MS

instrument was operated in data dependent mode to automatically switch between full scan

MS and MS/MS acquisition. Survey full scan MS spectra (from m/z 250–1800) were

acquired in the Orbitrap with resolution R = 60000 at m/z 400 (after accumulation to a 'target

value' of 500,000 in the linear ion trap). The seven most intense peptide ions with charge

states between 2 and 4 were sequentially isolated to a target value of 10,000 and fragmented

in the linear ion trap by collision induced dissociation (CID). All fragment ion spectra were

recorded in the LTQ part of the instrument. For all measurements with the Orbitrap detector,

1 lock-mass ion from ambient air (445.12002) was used for internal calibration as described

before [26]. Typical MS conditions were: spray voltage, 1.5 kV; no sheath and auxiliary gas

flow; heated capillary temperature, 200ºC; normalized CID energy 35%; activation q = 0.25;

activation time = 30 ms.

Data analysis

Protein identification was performed by MaxQuant 1.5.3.12 software package. Parent ion and

fragment mass tolerances were 8 ppm and 0.7 Da respectively and allowance for 3 missed

cleavages was made. Mouse canonical protein database from Uniprot (release May, 2015)

was used for the searches. Regular MaxQuant conditions were the following: Peptide FDR, 0.01; Protein FDR, 0.05; Min. peptide Length, 7; Variable modifications, Oxidation (M), Acetyl (N-Term), Acetyl (K), Dimethyl (KR), Phospho (STY), Propyl (K), Propyl/Methyl (K), Trimethyl (K); Peptides for protein quantitation, razor and unique; Min. peptides, 2; Min. ratio count, 2. Proteins were validated on the basis of at least 1 unique peptide detected in the 'matrix proteome' in all the 3 replicates or in at least 2 of the 3 replicates. Histones identified on the basis of peptides from the control only were excluded from further analysis.

Relative quantification of LC-MS/MS data enabled us to estimate the relative distribution pattern of H1 variants in mouse brain and compare it with that obtained from MALDI-IMS. Intensities of peptides belonging to all the different H1 isoforms, with the exception of shared peptides (across different variants), were summed and normalized to the total number of tissue sections (Supplemental file: H1_Analysis_LC_MS_MS_pmic.201500345.xlsx). In case of MALDI-IMS, corresponding intensities from the overall average spectrum (originating from one whole brain section) were averaged over 3 replicates and calculated for a single section.

RESULTS

To use MALDI-IMS for *in situ* analysis of intact histone variants and histone modifications we optimized sample preparation, ionization and detection parameters of the MALDI-TOF instrument to detect masses larger than 10 kDa. Optimization involved a series of washing and drying steps (as described in *Materials and methods*) of the tissue sections followed by a pulsed-ion-extraction (PIE) with a delay time of 200 ns. This led to a significant enhancement of signals, especially in the high mass range (Figures 1a-e and Figure S1). As a result, we

observed significantly higher number of peaks (Figure 1f) and 2-3 fold increase in intensities (Figure 1g). In addition, a substantially higher resolution was achieved owing to a narrower energy distribution of the desorbed ions (Figure 1h).

We applied this improved, high mass MALDI-IMS experimental protocol to sagittal sections of mouse brain tissue as it contains a large variability of cell types and histone modification patterns [27, 28] making it very difficult to draw conclusion on the regulation of histone modifications in bulk analysis of histones. To investigate if different regions of the brain differ significantly by their spectral composition, we performed an unsupervised clustering of spectra across the whole section. Interestingly, the clustering allowed us to distinguish different brain regions solely based on spectral features (Figures 2a-c) resulting in a detailed segmentation map of the brain. Visualization of the segmentation maps reveals that a single type of cluster (blue) is centered on the cerebellum and the olfactory bulb regions (COC, Figures 2d-f) whereas another (green) is located primarily at the cerebral cortex (Ctx, Figures 2g-i). Histological correlation with the H&E stained image shows that the COC regions have a much higher concentration of nuclei, when compared to the Ctx region (Figures 2e,f and 2h,i). To better understand the spectral features distinguishing the nucleus rich COC region from the Ctx region, we randomly picked 500 spectra from each of these clusters as input for automated feature extraction by PCA (Figure 2j). A scores plot (Figure 2j) shows the distribution of the spectra from the two regions in the 3D subspace determined by the extracted spectral features. Prominent variance in the two datasets is reflected in the 3D scores plot (Figure 2j), as spectra from the two groups (COC or Ctx) are distinctly separated from each other, while they are clustered within their own group.

When calculating the principal components, the spectral features corresponding to defined m/z values, acquire loading values, which represent their contribution towards the variance

between two datasets. A thorough analysis of the loadings plot (Figure 3a) therefore allowed us to identify the masses responsible for the distinction between COC and Ctx. Within the plot, each point represents an m/z value and the total number of points corresponds to the total number of peaks considered for the calculation. Higher loading values for a data point in the loadings plot (Figure 3a) therefore reflects a larger contribution to the distinct protein profiles. Points with loading values higher and lower than +1 and -1 were therefore considered to be the key molecules creating the difference between COC and Ctx. By doing this we noticed that the m/z values of peaks with a high abundance within COC were restricted to three distinct regions of the entire mass spectrum (Figures 3b-d). These regions coincide with the mass ranges of the histone variants and their PTMs, which we termed 'histone regions' (H4, Figure 3b; H2A_B/H3, Figure 3c and H1, Figure 3d). Therefore, we concluded that native histones and/or their post-translationally modified forms are responsible for the observed proteomic distinction of COC and Ctx.

In order to further validate the identification of native and modified histones and account for possible false positives within the 'histone regions', we applied a correlation analysis to the recorded data sets (Figure 4). To do this, we grouped spectra from several ROIs and calculated their nuclear and total ROI area based on the detected nuclei in the H&E counterstained tissue images (Figure 4). The ROIs were selected such that a) they encompass areas with different cell types (Table S1) and b) they cover regions from all over the tissue (to alleviate discrepancies arising from unevenness of the surface). Correlation analysis was then performed between the mean nuclear densities (from 3 replicates, equation 1) and the total intensities of the 30 m/z values detected within the 'histone regions' in all replicates (Figure 4 and Table S2). Resulting correlation plots were linear or non-linear, based on the detectability of the particular mass in all the ROIs (Figure 4). 12 of the 30 m/z values did not show a

correlation or even an anti-correlation and were therefore not considered histone derived. However, 11 of the peaks detected within the 'histone regions' showed a positive linear correlation ($R^2 > 0.8$) with nuclear density, suggesting that they indeed reflect modified histones or histone variants (Table S2). Interestingly, we also observe 7 peaks that show a positive but non-linear correlation with increasing nuclear density or a medium or weak linear positive correlation with intermediate R^2 values (Figure4, Figure S2 and Table S2). These masses are probably derived from histones that are expressed in a highly cell type specific manner or modifications of histones that only occur in selected regions of the brain.

Analyzing the 'matrix proteome' by bottom-up LC-MS/MS enabled us to confirm the identity of 10 of 11 peaks that have a strong positive linear correlation with nuclear density to be histones   (Table   S3,   Supplemental   file: Histone_Identification_LC_MS_MS_pmic.201500345). Also, 2 out of 7 peaks having a weak or non-linear positive correlation could be identified as histones. Among the 4 peaks showing a moderate or weak correlation, only one (m/z – 22011, $R^2 = 0.57$) could be identified in all replicates as histone and another (m/z – 21178, $R^2 = 0.26$) could be identified in one of the replicates. The remaining two masses ($R^2 < 0.1$) could not be validated as histones. In case of the 3 protein masses showing a non-linear positive correlation with nuclear density, one (m/z – 15329) could be unambiguously identified as being derived from histone H3, whereas the other two were detected in only one of the three replicates.

Our analysis provides a full measurement of the major histone modifications and histone variants in mouse brain tissue with a spatial resolution of 50 μm. In agreement with the current literature, the prominent form of histone H4 for example is the one that carries two methyl groups and one acetyl group (Figure 5a), which are very likely at H4K20 (me2) and at the N-terminus (ac). The second most abundant isoform carries an additional acetylation

mark, presumably at H4K16 [29]. Both H4 isoforms show a strong correlation with the nuclear density (Figure S2) and an almost identical distribution (Figure 5a) suggesting that they are not involved in regulating cell type identity in the brain.

As there is little information on the distribution of H1 isoforms in the brain[30], we tried to use our method to characterize the distribution of all H1 isoforms and their respective modified variants in the brain. We detect masses for all ubiquitously expressed H1 isoforms except H1.1 with a strong or moderate positive correlation to nuclear density (Figure 5b and Figure S2). Similar to what has been observed when H1 is analyzed in tissue culture cells or from tissue homogenates [31]; we find most H1 isoforms with the exception of the differentiation specific form H1.0 in a monoacetylated form, which in all likelihood resides at the N-terminus. In contrast to H4, multiple variants of the linker histone H1 exist in mammals, which have partially redundant but also non overlapping functions [3, 5, 32].

Due to the limitation of suitable antibodies the distribution of H1 variants in complex tissues has so far not been thoroughly analyzed [5]. The use of MALDI-IMS now allows us to determine the relative contribution of individual isoforms to the total H1 pool in individual brain regions (Figure 5c). A comparison of the integrated peak areas corresponding to the N-terminally acetylated H1 isoforms reveals H1.4 as the most abundant isoform overall (Figure 5d, Left panel) and in 5 out of the 6 ROIs analyzed. In the olfactory bulb area more H1.3 than H1.4 is detected. Compared to different H1 variants, the H1.5 isoform is only weakly expressed in all ROIs. Relative quantification of H1 family of proteins analyzed in LC-MS/MS shows an almost identical distribution of the H1 isoforms (Figure 5d, Right panel) as compared to MALDI-IMS in mouse brain. A probable hemoglobin variant was used as the positive control for the spatial distribution studies in the perfused brains (Figure S3).

As this differential distribution of H1 variants in the brain suggest that they might be involved in defining cellular identities and regulating the epigenetic regulation in the brain, it will be interesting to investigate changes in the histone composition upon brain injury or pathological changes as the ones observed in Alzheimer's or Parkinson's disease. As our newly developed analysis pipeline is independent of specific antibodies and allows the simultaneous measurement of multiple histone features, specific pattern recognition algorithms may help to detect molecular changes in the brain upon such neurological challenges and get a better understanding of epigenetic pathways in the brain.

DISCUSSION

The *in situ* detection of multiple histone variants and their PTMs in a single tissue section and with a high spatial resolution requires extensive IMS of intact proteins and comprehensive data analysis. Imaging mass spectrometry for intact proteins has been limited so far by lack of spectral sensitivity and resolution on one hand and that of in-depth data analysis on the other. A recent study shows IMS of intact proteins could be performed with very high accuracy and mass resolution [33], but it is limited by the mass range in which it can be applied and the instrumental set-up. The current study uses a combination of tissue pre-treatment steps and optimum measurement parameters that provides significantly higher resolution and sensitivity in the high mass ranges. Maximum effect is observed in case of molecules with molecular weights starting from 20000 Da. This makes IMS particularly well suited for the analysis of histones including the linker histone H1,as all the H1 variants have their molecular weights between 20000 Da. and 25000 Da [5].

Even without an a priori focus on the masses derived from intact histones, tissue areas with similar spectral features and hence proteome composition could be identified by HCA of the datasets [34]. An automated feature extraction from two clusters (COC and Ctx) pointed to a number of distinguishing m/z values within specific regions of the whole mass spectrum. These spectral regions were similar to the known histone masses, suggesting histones as the probable discriminating protein component within different brain regions. A correlation analysis between the discriminating m/z values and cytological features such as the nuclear density is capable of further validating the identified values as being of nuclear origin and therefore in all likelihood histones. This is substantiated by the complementary LC-MS/MS experiments, which identify 10 out of 11 strongly correlated peaks to be histones. The above mentioned correlation plots can also be used as standard curves for estimating levels of

histones outside the ROIs. This is particularly useful in diseased conditions, where verified histone masses can be correlated with other microscopic features such as a staining for apoptotic cells or plaque formation. Although this correlation analysis provides an excellent filter to eliminate false positive assignments, it also has certain drawbacks regarding false negatives. In fact, we do observe peaks in the spectrum that only show a weak correlation with overall nuclear density due to their highly tissue specific expression pattern. As we also do not identify these potential histones in the LC MS/MS analysis, we were not sufficiently confident to assign a particular histone modification or variant to these peaks. Therefore, we cannot exclude that some of the masses that are excluded from further investigation were indeed derived from cell type specific histone variants. In particular, non-nuclear H1 has been shown to have functional implications in neurons [35], which would result in a weaker correlation of the peak intensity with the nuclear density. Another technical limitation in the current set up is the lack of resolution in the H3 mass range (15000-16000 Da.). Flanking of this region by sharp hemoglobin peaks (m/z - 14980, 15616 and 15848) makes it difficult to resolve the H3 modifications. However, the method should be competent to detect them in systems, where this technical limitation is not present.

Besides the danger of eliminating histone masses due to a false negative assignment there is also the potential of a false positive assignment. As our analysis is based on matching the m/z value of the intact molecule with a theoretical value derived from a histone database, there is the potential that a peak is wrongly assigned as a histone peak due to its similar mass. We tried to eliminate this by a correlation with the nuclear density and measurements with ppm values of maximally 250. Moreover, the current pipeline lacks positional information of the modification within the intact histone. To do this, either a top-down MS/MS approach of the intact molecule as it is released from the section or an *in situ* digestion would be required,

which is difficult using the current MS hardware. However, bottom-up LC-MS/MS analysis of the 'matrix proteome' could identify 12 out of the 18 peaks assigned as histones to be derived from histones. Results show that chance of false positive assignment is less than 10% for masses in 'histone regions' having a strong correlation with nuclear density ($R^2 > 0.8$).

Keeping in mind, the inability of the method to distinctly identify cell type specific histones, our MALDI-IMS method is currently the only reliable method that allows a simultaneous analysis of histones, histone PTMs and histone variants from a single tissue section, including the linker histone H1. The relative abundance of H1 variants as estimated by MALDI-IMS was strongly corroborated by quantitative LC-MS/MS studies (Figure 5d). This further consolidates the effectiveness, accuracy and applicability of the method. Furthermore, our improved experimental set-up and data analysis tools (ProtFinder and OrgPlotter) can not only be used to analyze histone proteins *in situ* but can be applied to the whole proteome in general. These predictive tools are particularly useful for the MALDI-IMS community when combined with other cytological tools such as the detection of specific organelles like nuclei or mitochondria or the use of antibodies to delineate specific features of the studied tissue. With ProtFinder, it is possible to predict one or more proteins corresponding to an experimental mass or mass list. The second script, OrgPlotter, uses an experimental mass list and correlates the ROI intensities with the nuclear density or any other recognizable microscopy markers identifying specific cell types or other organelles. Therefore, the workflow reported here along with the data analysis tools and validation using an orthogonal LC-MS/MS method, provides a comprehensive platform to simultaneously analyze multiple histone variants in a label-free manner from biological tissue sections.

REFERENCES

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

[2] Richmond, T. J., Davey, C. A., The structure of DNA in the nucleosome core. *Nature* 2003, *423*, 145-150.

[3] Izzo, A., Kamieniarz, K., Schneider, R., The histone H1 family: specific members, specific functions? *Biological chemistry* 2008, *389*, 333-343.

[4] Song, F., Chen, P., Sun, D., Wang, M.*, et al.*, Cryo-EM Study of the Chromatin Fiber Reveals a Double Helix Twisted by Tetranucleosomal Units. *Science* 2014, *344*, 376-380.

[5] Harshman, S. W., Young, N. L., Parthun, M. R., Freitas, M. A., H1 histones: current perspectives and challenges. *Nucleic acids research* 2013, *41*, 9593-9609.

[6] Musselman, C. A., Lalonde, M.-E., Cote, J., Kutateladze, T. G., Perceiving the epigenetic landscape through histone readers. *Nat Struct Mol Biol* 2012, *19*, 1218-1227.

[7] Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D., Patel, D. J., How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* 2007, *14*, 1025-1040.

[8] Kamakaka, R. T., Biggins, S., Histone variants: deviants? *Genes & Development* 2005, *19*, 295-316.

[9] Feng, Y., Jankovic, J., Wu, Y. C., Epigenetic mechanisms in Parkinson's disease. *Journal of the neurological sciences* 2015, *349*, 3-9.

[10] Fraga, M. F., Ballestar, E., Villar-Garea, A., Boix-Chornet, M.*, et al.*, Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nature genetics* 2005, *37*, 391-400.

[11] Waldmann, T., Schneider, R., Targeting histone modifications--epigenetics in cancer. *Current opinion in cell biology* 2013, *25*, 184-189.

[12] Wang, J., Yu, J. T., Tan, M. S., Jiang, T., Tan, L., Epigenetic mechanisms in Alzheimer's disease: implications for pathogenesis and therapy. *Ageing research reviews* 2013, *12*, 1024-1041.

[13] Wagner, J. M., Hackanson, B., Lübbert, M., Jung, M., Histone deacetylase (HDAC) inhibitors in recent clinical trials for cancer therapy. *Clinical Epigenetics* 2010, *1*, 117-136.

[14] Nishikori, S., Hattori, T., Fuchs, S. M., Yasui, N.*, et al.*, Broad ranges of affinity and specificity of anti-histone antibodies revealed by a quantitative peptide immunoprecipitation assay. *Journal of molecular biology* 2012, *424*, 391-399.

[15] Rothbart, S. B., Lin, S., Britton, L.-M., Krajewski, K.*, et al.*, Poly-acetylated chromatin signatures are preferred epitopes for site-specific histone H4 acetyl antibodies. *Scientific Reports* 2012, *2*, 489.

[16] Britton, L.-M. P., Gonzales-Cope, M., Zee, B. M., Garcia, B. A., Breaking the histone code with quantitative mass spectrometry. *Expert Review of Proteomics* 2011, *8*, 631-643.

[17] Villar-Garea, A., Israel, L., Imhof, A., *Current Protocols in Protein Science*, John Wiley & Sons, Inc. 2001.

[18] Hanrieder, J., Wicher, G., Bergquist, J., Andersson, M., Fex-Svenningsen, Å., MALDI mass spectrometry based molecular phenotyping of CNS glial cells for prediction in mammalian brain tissue. *Anal Bioanal Chem* 2011, *401*, 135-147.

[19] Hardesty, W. M., Kelley, M. C., Mi, D., Low, R. L., Caprioli, R. M., Protein signatures for survival and recurrence in metastatic melanoma. *Journal of proteomics* 2011, *74*, 1002-1014.

[20] Morita, Y., Ikegami, K., Goto-Inoue, N., Hayasaka, T.*, et al.*, Imaging mass spectrometry of gastric carcinoma in formalin-fixed paraffin-embedded tissue microarray. *Cancer Science* 2010, *101*, 267-273.

[21] Poté, N., Alexandrov, T., Le Faouder, J., Laouirem, S.*, et al.*, Imaging mass spectrometry reveals modified forms of histone H4 as new biomarkers of microvascular invasion in hepatocellular carcinomas. *Hepatology* 2013, *58*, 983-994.

[22] Aichler, M., Walch, A., MALDI Imaging mass spectrometry: current frontiers and perspectives in pathology research and practice. *Lab Invest* 2015, *95*, 422-431.

[23] Deutskens, F., Yang, J., Caprioli, R. M., High Spatial Resolution Imaging Mass Spectrometry and Classical Histology on a Single Tissue Section. *Journal of mass spectrometry : JMS* 2011, *46*, 568-571.

[24] Strohalm, M., Hassman, M., Košata, B., Kodíček, M., mMass data miner: an open source alternative for mass spectrometric data analysis. *Rapid Communications in Mass Spectrometry* 2008, *22*, 905-908.

[25] Maier, S. K., Hahne, H., Gholami, A. M., Balluff, B.*, et al.*, Comprehensive Identification of Proteins from MALDI Imaging. *Molecular & Cellular Proteomics* 2013, *12*, 2901-2910.

[26] Olsen, J. V., de Godoy, L. M. F., Li, G., Macek, B.*, et al.*, Parts per Million Mass Accuracy on an Orbitrap Mass Spectrometer via Lock Mass Injection into a C-trap. *Molecular & Cellular Proteomics* 2005, *4*, 2010-2021.

[27] Taniura, H., Sng, J. C. G., Yoneda, Y., Histone modifications in the brain. *Neurochemistry International* 2007, *51*, 85-91.

[28] Rumbaugh, G., Miller, C., in: Roberson, E. D. (Ed.), *Alzheimer's Disease and Frontotemporal Dementia*, Humana Press 2011, pp. 263-274.

[29] Taipale, M., Rea, S., Richter, K., Vilar, A.*, et al.*, hMOF Histone Acetyltransferase Is Required for Histone H4 Lysine 16 Acetylation in Mammalian Cells. *Molecular and Cellular Biology* 2005, *25*, 6798-6810.

[30] Doenecke, D., Albig, W., Bode, C., Drabent, B.*, et al.*, Histones: genetic diversity and tissue-specific gene expression. *Histochemistry* 1997, *107*, 1-10.

[31] Wisniewski, J. R., Zougman, A., Kruger, S., Mann, M., Mass spectrometric mapping of linker histone H1 variants reveals multiple acetylations, methylations, and phosphorylation as well as differences between cell culture and tissue. *Molecular & cellular proteomics : MCP* 2007, *6*, 72-87.

[32] Rupp, R. A., Becker, P. B., Gene regulation by histone H1: new links to DNA methylation. *Cell* 2005, *123*, 1178-1179.

[33] Spraggins, J., Rizzo, D., Moore, J., Rose, K.*, et al.*, MALDI FTICR IMS of Intact Proteins: Using Mass Accuracy to Link Protein Images with Proteomics Data. *J. Am. Soc. Mass Spectrom.* 2015, *26*, 974-985.

[34] Alexandrov, T., Becker, M., Deininger, S. O., Ernst, G.*, et al.*, Spatial segmentation of imaging mass spectrometry data with edge-preserving image denoising and clustering. *Journal of proteome research* 2010, *9*, 6535-6546.

[35] Bolton, Russelakis, C., Betmouni, Perry, Non-nuclear histone H1 is upregulated in neurones and astrocytes in prion and Alzheimer's diseases but not in acute neurodegeneration. *Neuropathology and Applied Neurobiology* 1999, *25*, 425-432.

ACKNOWLEDGEMENTS

Accepted Article

CONFLICT OF INTEREST

The authors declare no conflict of interests.

FIGURE LEGENDS
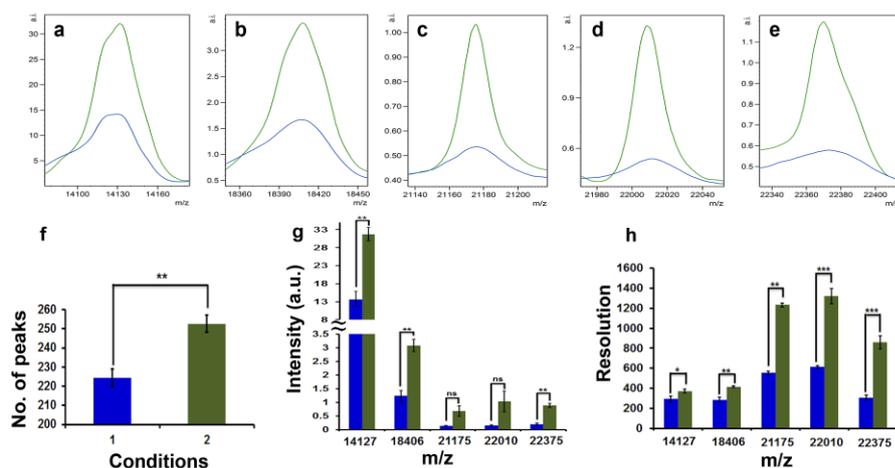
Figure 1: Comparison between two experimental conditions (**Old** – Extensive washing as described in Materials and Methods, air drying and 100 ns PIE; **New** – Extensive washing, vacuum drying and 200 ns PIE). *(a-e)* Peaks from different regions of mass spectrum show improved intensity and resolution upon changing of conditions. *(f)* No. of peaks significantly increases in the New conditions (p = 0.0037) (**Old** – 1, **New** – 2). *(g)* Intensity of the peaks increases substantially in altered conditions. *(h)* Mass resolution shows significant improvement for the 'New' conditions. **Symbols** – ns for p > 0.05, * for p ≤ 0.05, ** for p ≤ 0.01, *** for p ≤ 0.001 (Two-tailed students' t-test with unequal variance. Plots: mean ± SD).
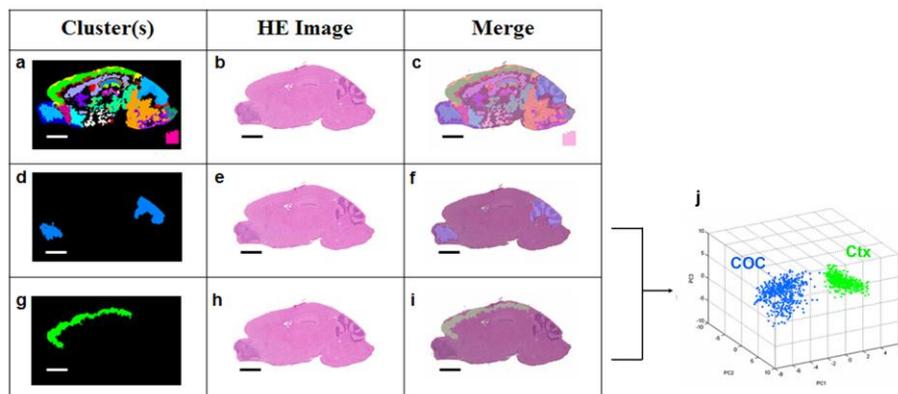
Figure 2: *(a)* Segmentation map showing clustering of brain regions according to their proteomic profiles. *(b)* H&E image of the same section and *(c)* merged image. *(d-f)* A single type of cluster centered on cerebellum and olfactory bulb (COC) and *(g-i)* cerebral cortex (Ctx). *(j)* PCA of COC and Ctx and the resulting 3D scorings plot. Scale bars = 2 mm.
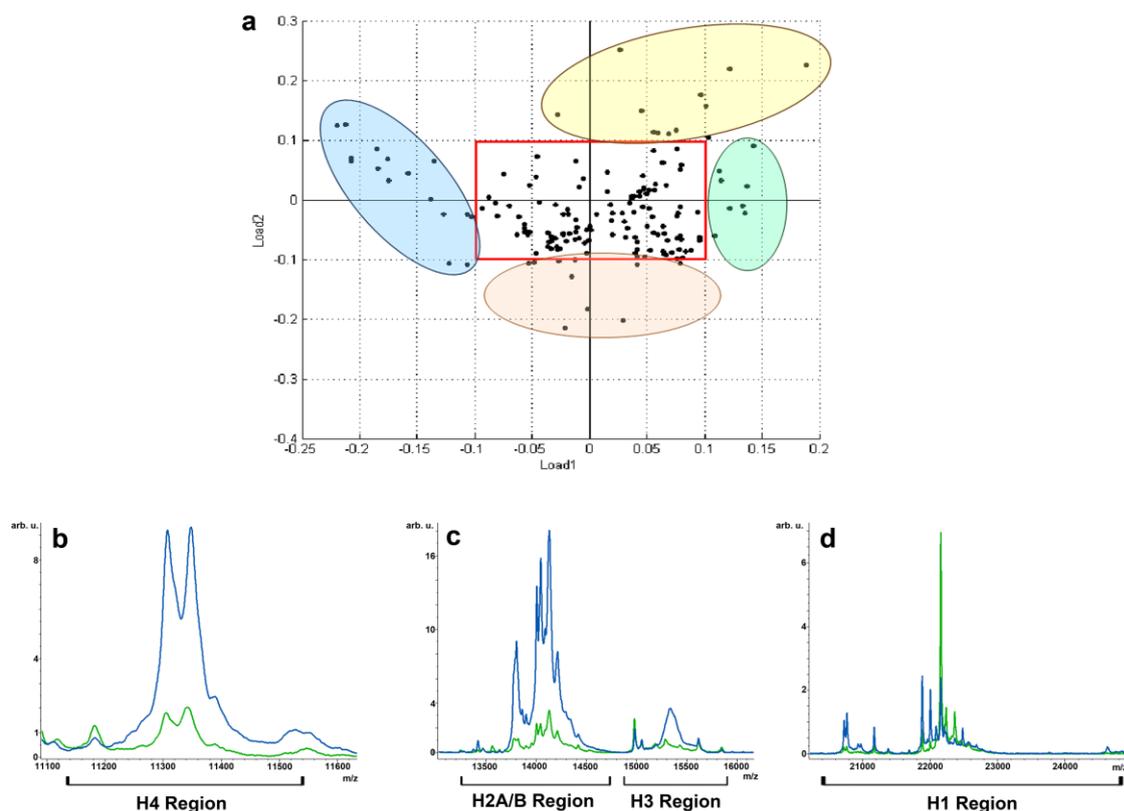
Figure 3: *(a)* Loadings plot of the PCA. Red rectangle defines the region, masses outside of which are considered to contribute maximally to the explained variance. The colored ovals

describe the outliers.*(b-d)* Spectral comparison shows masses from three distinct regions of the mass spectrum to be enriched in COC as compared to Ctx. Comparison of observed masses with theoretical histone database delineates the same three regions of the mass spectrum as 'histone regions' (H4 region, *b*; H2A_B/H3 region, *c* and H1 region, *d*).
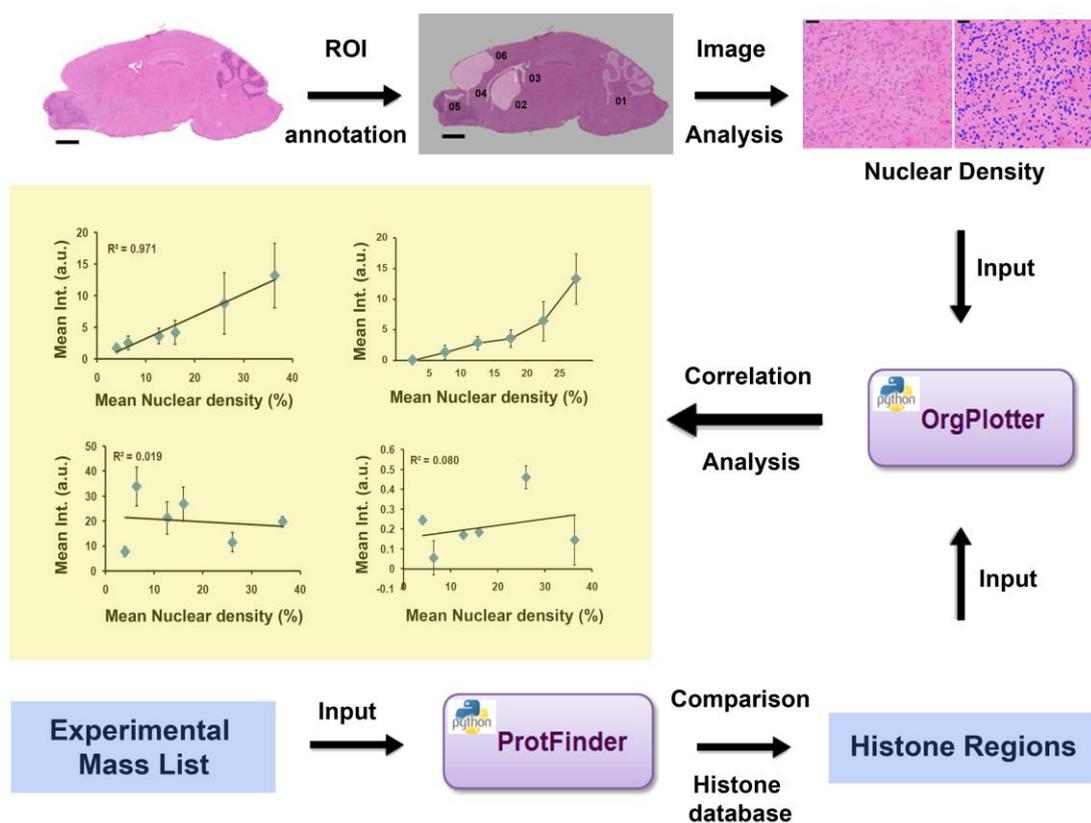
Figure 4: Workflow for nucleus detection from tissues and correlation analysis of experimental masses with nuclear density. Scale bars (whole tissue) = 2 mm, Scale bars (nuclear density panel) = 50 μm.
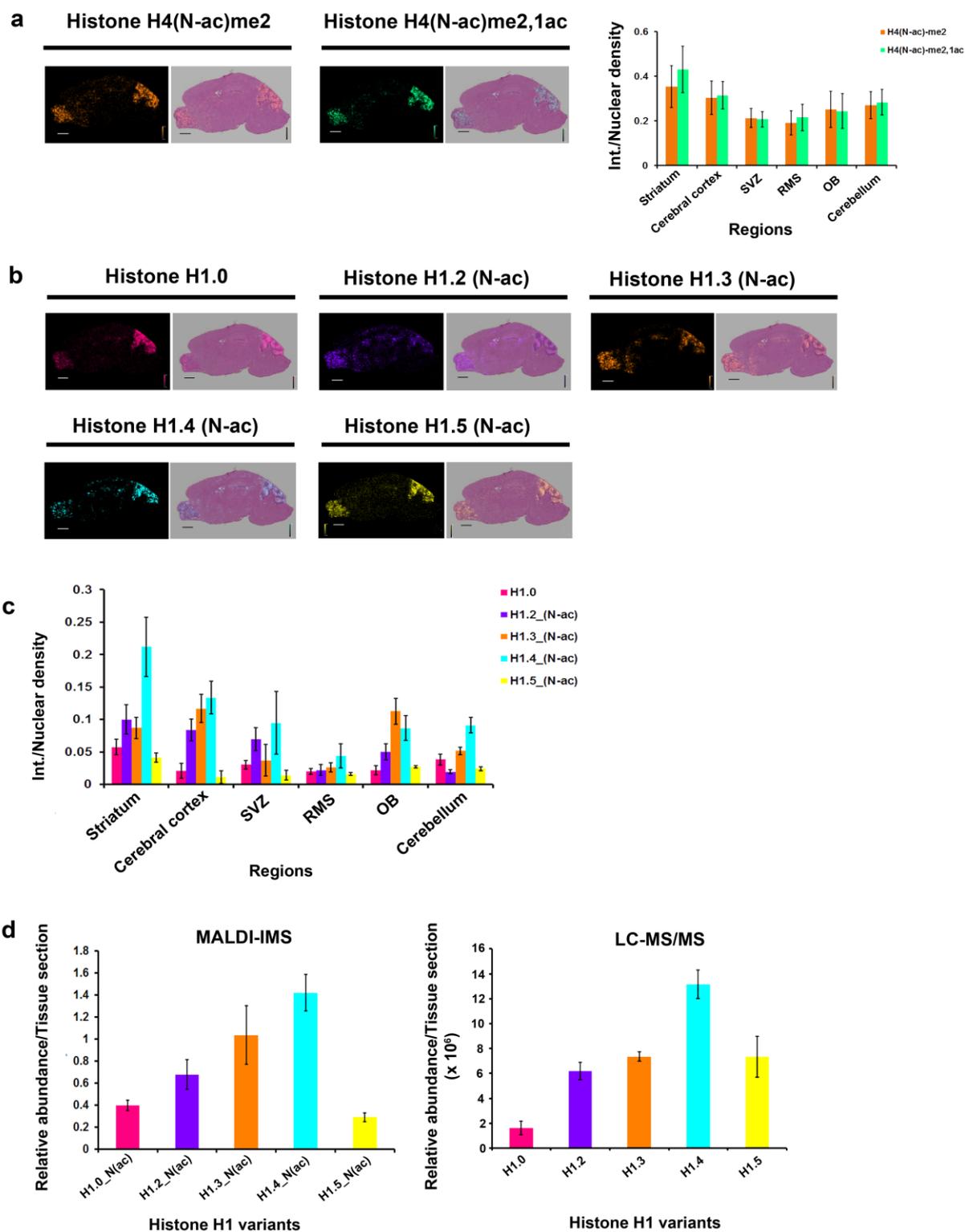
Figure 5: *(a)* Spatial distribution patterns of modified forms of histone H4 and relative

distribution of these forms in the selected brain regions (n = 3). *(b)* Spatial distribution

patterns of unmodified (or N-terminally acetylated) forms of different histone H1 variants and *(c)* their relative distribution in the selected brain regions (n = 3) *(d)* Relative distribution of H1 variants in mouse brain as measured in MALDI-IMS (left panel) and LC-MS/MS (Right panel)(Relative abundance for MALDI-IMS – Mean intensity of the protein peaks per tissue section, n= 3; and for LC-MS/MS – Sum of intensities of peptides belonging to the corresponding H1 variant per section, n = 10). All Plots: mean ± SEM. Scale bars = 2 mm.