Abstract: Epigenetic modifications contribute to the determination of cell fate and differentiation. The molecular mechanisms underlying histone variants and post-translational modifications (PTMs) have been studied in the contexts of development, differentiation, and disease. Antibody-based assays have classically been used to target PTMs, but these approaches fail to reveal combinatorial patterns of modifications. In addition, some histone variants are so similar to canonical histones that antibodies have difficulty distinguishing between these isoforms. Mass spectrometry (MS) has progressively developed as a powerful technology for the study of histone variants and their PTMs. Indeed, MS analyses highlighted exquisitely complex combinations of PTMs, suggesting “crosstalk” between them, and also revealed that PTM patterns are often variant-specific. Even though the sensitivity and acquisition speed of MS instruments have considerably increased alongside the development of computational tools for the study of multiple PTMs, it remains challenging to correctly describe the landscape of histone PTMs, and in particular to confidently assign modifications to specific amino acids. Here, we provide an inventory of MS-based strategies and of the pitfalls inherent to histone PTM and variant characterization, while stressing the complex interplay between PTMs and histone sequence variations. We will particularly illustrate the roles played by MS-based analyses in identifying and quantifying histone variants and modifications.

Keywords: histone variants; post-translational modifications; mass spectrometry; bottom-up analysis; computational tools; crosstalk

1. Introduction

The nucleosome core particle packaging cellular DNA is composed of two copies of H2A, H2B, H3, and H4 histones surrounded by 142 base pairs of DNA [1]. At the entry and exit sites on the surface of the nucleosome, DNA is bound by a linker histone H1 to stabilize the chromatin subunit structure [2–4]. Each histone comprises a globular domain, interacting with DNA and amino terminal tails protruding from nucleosome particles.

A large diversity of post-translational modifications (PTMs) has been widely described on histones. They include classic phosphorylation, acetylation, and methylation, but also acyl groups such as propionyl, butyryl, 2-hydroxyisobutyryl groups, and sugar derivatives (O-GlcNAcylation) [5]. They accumulate in the exposed surfaces of histones when incorporated into nucleosomes, namely their N- and C-terminal regions. However, several modifications have been identified within their globular domain (histone fold) and regulate their assembly into nucleosomes [6,7]. Histone PTMs are essential elements of chromatin signaling pathways. They accumulate at specific regions of the genome and participate in the definition of their functionality as active promoters, enhancers or in the maintenance in a robustly repressed status. As an example, the methylation of lysine 4 of H3 (H3K4) is a strong
mark of active transcription. It is deposited by the methyltransferase Set1 from the COMPASS complex. These PTMs can recruit a variety of specific binders, which can either reinforce transcription activation or promote its repression (reviewed in Reference [8]). Another example is the methylation of H3K27 by Polycomb group proteins which establish a strongly repressed chromatin structure. Altogether, histone modifications are key features of chromatin dynamics and transcription regulation mechanisms with crucial roles during embryonic development and cell differentiation [9]. They have also been linked to a subset of diseases, including cancer, neurological, and autoimmune disorders [10].

In mammals, multiple copies of genes encode the canonical histone proteins [11]. These genes share extensive sequence similarities and are expressed during the S phase of the cell cycle. Histone variants emerged in mammals for each family of histones, except for H4. These variants include somatic histone variants (H2A.X, H2A.Z, Macro-H2A, H2A.B, and H3.3) and testis-specific variants (TS H2A.1, H2A.L, TS H2B.1, H2B.L, subH2B, and H3T) [11]. These non-allelic variants differ little in their amino acid sequences and are highly conserved between human and mouse (Figure 1). In contrast to canonical histones, histone variants are usually present as single-copy genes that are expressed throughout the cell cycle [12,13]. They ensure specific functions and confer specific structural properties to the nucleosome. In the particular context of spermatogenesis, for example, histone variants are key players involved in the differentiation of adult germ stem cells into mature spermatozoa. In that specific case, the sequential replacement of canonical histones by histone variants facilitates chromatin compaction. Indeed, the few variations in their amino acid composition confer specific properties to these variants and are either associated to repressive or to active chromatin states by destabilizing the structure of the nucleosome. For instance, H2A.Z was largely associated to active gene expression, while macro-H2A variants were essentially found in repressive chromatin regions [14]. Moreover, some of these variants are expressed in the latest stages of spermatogenesis, suggesting their important role in the packaging of male germ cell DNA [15]. Post-translational modifications on histones can also significantly impact chromatin conformation, and thus regulate the accessibility of DNA to transcription factors. Many of these modifications can occur together to modulate gene expression. The testis-specific histone variant H2B (TS H2B.1) is the major variant form of somatic H2B. Surprisingly, Th2b KO has no effect on male infertility due to a compensatory mechanism mediated by the addition of methyl groups at H4R35, H4R55, H4R67, and H2BR72 [16]. These findings suggest that histone modifications as well as histone variants are “docking sites” and a network for the recruitment of protein factors to facilitate the chromatin remodeling, and thus the regulation of gene expression. However, there is still a long way to go to obtain a complete understanding of the complex interplay between histone variants, PTMs, and protein effectors.
Figure 1. Sequence similarity between histone variants in human and mouse. Core histones H2A (yellow), H2B (red), H3 (blue), and histone linker H1 (purple) are illustrated for mouse (A) and human (B). Sequence data were obtained from and treated as published in Reference [17].

A key challenge for scientists studying histones and their variants lies in deciphering the complexity of histone variants, identifying PTM combinations, and understanding their role in pathophysiological contexts. Historically, this challenge has been addressed using antibody-based methods, such as western blot, immunofluorescence analysis, and chromatin immunoprecipitation.
(ChIP). However, the sequence homology between histone variants and the large array of histone PTMs that exist are incompatible with the generation of highly specific antibodies. In addition, antibody-based assays require knowledge of the type and position of the modification of interest. As a result, histone antibodies are not always specific for their targets and often generate cross-reactivity [18]. Additionally, two neighboring PTMs within the same histone may cause epitope occlusion, and thus impede recognition of the target PTM. Finally, these approaches lack high-throughput capabilities, and each PTM must be analyzed separately, which makes the analysis of multiple PTMs occurring within the same histone expensive as well as time- and effort-consuming [19,20]. As an alternative, mass spectrometry (MS)-based proteomics is a valuable approach for the identification and quantification of co-occurring PTMs within a single protein sequence. Historically, three main MS-based approaches, namely bottom-up, middle-down, and top-down methods, were developed and adapted for the investigation of histone PTMs (Figure 2). All of these approaches have been extensively applied to histones and successfully mapped over 500 different PTMs [5]. The PTMs identified include a broad range of classical histone modifications such as lysine or arginine methylation (Kme/Rme), lysine acetylation (Kac), O-GlcNAcylaton (S/T-GlcNAc), ubiquitination (Kub), and serine and threonine phosphorylation (Sph/Tph). The landscape of lysine modifications was significantly broadened by the recent discovery of acyl chains, such as propionylation (Kpr) and butyrylation (Kbu) [21], crotonylation (Kcr) [22], succinylation (Ksucc) [23], malonylation (Kma) [24], 2-hydroxyisobutyrylation (Khib) [25], and glutarylation (Kglu) [26]. Altogether, the large diversity of PTMs, along with their possible combinations on histones, render their MS analysis incredibly complex. Here, we discuss advances made in sample preparation and instrumental methods. Computational methods used to detect and quantify histone variants and their PTMs are also reviewed.

![Figure 2. Representation of bottom-up, middle-down, and top-down mass spectrometry experiments.](image.png)

2. Extracting Histones from Biological Samples

Many protocols have been published to extract histones from mammalian cells and tissues. A landmark protocol was published by Shechter et al. [27] describing two classically used options. The procedure consists in isolating nuclei using a hypotonic lysis buffer. As histones are highly basic proteins, with an isoelectric point close to 10, they can be extracted using either acidic (classically sulfuric acid) or high-salt buffers. However, histone extraction protocols were not
specifically designed for MS analysis, and caution needs to be taken for a successful comparative analysis of endogenous PTMs. First, some PTMs are sensitive to the extraction method, such as the phosphorylations on Ser and Thr residues, which are labile at very low pH [27]. Moreover, it is necessary to avoid/remove MS-incompatible reagents such as detergents and high salt concentrations. The type of lysis mechanism should also be evaluated to maximize histone yield. To investigate the histone PTM patterns between different samples, the extraction should be done as fast as possible to reduce artefactual changes mediated by residual epigenetic effectors. Finally, histone proteins can be further fractionated on an SDS-PAGE gel to remove possibly abundant high-molecular weight proteins that would impede the study of low stoichiometry PTMs; alternatively, extracted histone samples can be loaded on a C8 or C18 column to produce pure fractions of each histone subtype [28,29].

3. Bottom-Up Mass Spectrometry Analysis of Histones

Several MS methods have been developed for use with histones; these methods analyze different sizes of protein fragments. This review will focus on the most widely adopted bottom-up approach; middle-down and top-down approaches are reviewed elsewhere [30,31]. The bottom-up analysis offers reliable comparisons between samples, but for the identification and quantification of PTMs, subtle pitfalls from histone preparation for proteomic analysis to MS data interpretation have to be considered.

The bottom-up approach involves proteolysis of the histones into short peptides (8–30 amino acids), the sequences of which can then be determined by MS/MS fragmentation. Different proteolytic enzymes can be used, but trypsin remains the protease of choice for its efficiency and specificity of cleavage. However, because histones are very rich in lysine and arginine residues, some regions of histones get chopped into peptides which are too short for HPLC retention and successful identification by MS/MS. Garcia et al. [32] developed a protocol to chemically derivatize lysines, thus blocking trypsin cleavage. This chemical treatment adds a propionyl group to unmodified lysine residues and restricts the action of trypsin to arginine residues, mimicking an Arg-C-like proteolysis, but with improved reliability. As a result, peptides suitable for MS-based analysis are generated. However, this derivatization procedure is not compatible with the study of naturally occurring propionylation events [33], and also partially alkylates the hydroxyl groups on serine/threonine residues. In addition, the frequently observed addition of a propionyl group to methylated lysines is strictly isobaric (i.e., of exactly the same mass due to identical compositions in atoms) with butyrylation. This issue can be overcome through the use of an isotope-labeled propionic donor group, such as D5-propionic anhydride [34] or D5-butyryl anhydride [35]. Garcia et al. [33] assessed the ionization efficiency of 12 chemical derivatization groups, and propionic anhydride appeared to be one of the best agents.

In 2011, Zhao and colleagues [22] identified crotonylation as a new lysine PTM and described its role during spermatogenesis. The existence of crotonylation, like the other PTMs reported over the past decade, was validated by analyzing a synthetic peptide of the same sequence and bearing the hypothetical modification, and then verifying strict co-elution during LC separation of the synthetic and endogenous species, and their highly similar fragmentation patterns. This new PTM was searched for using four different sample preparation protocols to increase histone sequence coverage and extend the repertoire of modified sites. Histone extracts were thus (1) digested in-solution using trypsin without chemical propionylation, (2) chemically propionylated after in-solution tryptic digestion, (3) chemically propionylated before in-solution tryptic digestion, and (4) in-gel digested after SDS-PAGE gel separation. Among the four methods tested, in vitro propionylation before tryptic digestion of histones gave the highest sequence coverage for histones H1, H2A, and H2B. In-gel digestion after SDS-PAGE separation provided the best coverage for histones H3 and H4, 75–80% coverage for H2A, H2B and about 60% for H1 [22]. Thus, these protocols, with and without in vitro chemical modification of non-modified Lys residues, can provide complementary PTM identifications.
4. Data Interpretation

Mass spectrometry fragmentation (MS/MS) data is interpreted using dedicated search engines which compare experimental MS/MS spectra to theoretical peptide fragmentation patterns and identify the best match. A number of search engines are now publicly available on the web (summarized here https://omictools.com/database-search-category) alongside the commercial solutions. These search engines rely on a “protein sequence database” which is explored to identify proteins and their PTMs from the experimentally acquired data. These databases should be comprehensive, non-redundant, and well annotated to allow optimal MS/MS data interpretation. Two major repositories exist, hosted by the National Center for Biotechnology Information (NCBI: http://www.ncbi.nlm.nih.gov/) and the Universal Protein Resource (UniProt: http://www.uniprot.org/). However, the current versions of these public databases contain many redundant entries (UniProtKB or NCBI). The NCBIINr or UniRef100 databases contain only non-redundant proteins, but their histone nomenclature does not follow the current community standards [36]. Recently, a proteomics-oriented and manually curated resource of mouse and human sequences, named MS_histoneDB, was released [11]. Each histone entry in this database is annotated in accordance with the current nomenclature and unified with the “HistoneDB2.0” database [36,37]. This resource is provided in a format that can be directly read by MS/MS data interpretation programs. Our team used MS_histoneDB to interpret proteomics data for histones extracted from mouse testis. Several histone variants that had so far only been inferred by homology or detected at the RNA level, were detected by MS analysis, thus confirming their existence at the protein level [11].

Another critical parameter to specify before launching the interpretation of MS/MS data for histone peptides is the list of expected modifications. Close to 20 different structures of PTMs have thus far been described on histones [38]. However, they cannot all be considered simultaneously by most search engines, which only accommodate a limited number of variable PTMs (e.g., a maximum of nine for Mascot). The performance of eight programs was compared with respect to their capacity to identify PTMs on histone peptides, and pFind and Mascot appeared to yield good performances [39]. As an alternative to matching experimental MS/MS spectra to theoretical spectra deduced from known protein sequences, programs designed to perform de novo sequencing of MS/MS data have been successfully used for sequences containing few modifiable residues. For peptide sequences such as the N-terminal peptide of H4 GKGGKGLGKGGAKR, however, the several variable PTMs on the four lysines make the search space far too broad to obtain satisfactory results. Finally, a few groups have developed software tools to identify mass shifts associated with as-yet undescribed PTMs. PTMap is one such tool based on a sequence alignment algorithm; it selects sequence-information-rich MS peaks, infers precise localization of PTM sites, and assigns a score based on spectrum quality [40]. PTMap was successfully used to discover propionylation and butyrylation on histones [41].

5. The Puzzle of Histone PTMs

Over the past ten years, >500 individual histone PTMs have been mapped by MS analysis [5]. The complexity of investigating PTMs within histone proteolytic peptides is due to: (i) the frequent existence of strictly isobaric peptides, (ii) the uncertain positioning of PTMs on neighboring modifiable residues, (iii) the difficulty of distinguishing between several isobaric PTM combinations, (iv) the potential masking of isobaric histone amino acid variations by PTM combinations (Figure 3). Finally, the above aspects not only complicate identification of the correct sequence and of the PTM sites, they also make it very difficult to quantify the relative abundance of positional isomers that often co-elute, at least partially.
First, at the protein sequence level, some histone variants generate strictly isobaric peptides that can only differ by their amino acid order. For example, the mammalian H2A.L.1 isoforms share very high similarity. Those encoded by the H2al1a and H2al1j genes differ by a swap of an amino acid triplet: peptides GEFPLSVDR from H2A.L.1 (H2al1j) and GELPSLVDR from H2A.L.1 (H2al1a) differ by the swap of FPL residues to LPF. They co-elute using a classical elution gradient on a C18 column, and can only be qualitatively and quantitatively distinguished by targeted proteomics (Selected Reaction Monitoring or Parallel Reaction Monitoring) using discriminating fragment ions, such as y6 and y7 ions.
y7 ions [17]. Second, a given amino acid sequence containing several modifiable residues (such as Lys and Arg) and bearing a set of PTMs can exist as multiple positional isomers (i.e., same modification occurring on different residues of a given peptide sequence), which complicates the identification and relative quantification of these diverse forms, as detailed below.

5.2. Uncertain Positioning of PTMs on Neighboring Modifiable Residues

LC-MS/MS analysis of the N-terminal fragment of histone H4 can identify acetyl modification across four lysine sites (K5, K12, K8 and K16) [42]. In this case, it is difficult to confidently assign di- or tri-acetylation to specific lysine residues, because positional isomers frequently co-elute, and as a result co-fragment. Identification results produced by the Mascot search engine on this sequence (Figure 3, case 1) demonstrate that several possible sequences only differing by their PTM positions were suggested with similar identification scores. Feller et al. [43] developed an optimized protocol to elucidate the problem of PTM positioning within a given peptide. They used chemical acetylation of all modified and mono-methylated lysines with deuterated acetic anhydride, which adds a 3-Da mass shift and distinguishes endogenous from chemical acetylation. The acetic derivatization gives the originally unmodified peptides the same chemical properties as the endogenously acetylated peptides to make comparison possible.

5.3. Ambiguity between Several Isobaric PTM Combinations

H3 residues K27 and K36 have been described as harboring numerous combinations of acetylation/methylation states (ranging from me1 to me3 on either H3K27 or H3K36) [44]. In fact for a precursor of the same mass, peptide K27SAPSTGGVK36K37PHR may in theory be modified as follows: K27me1-K36me2, K27me1-K36me1-K37me1 and K27me1-K37me2. The absence of efficient HCD fragmentation between residues K36 and K37 makes PTM positioning very difficult, if not impossible (Figure 3, case 2). Moreover, K27me1-K36me2 and K27me2-K36me1 co-elute, making quantification of these two species difficult when using a classical exploratory approach. In such cases, the positional isomers must be quantified by considering specific fragments of each peptide form, which consists of implementing a targeted proteomic analysis. Indeed, these two positional isomers can be distinguished by monitoring fragment ions y8 to y10 (see details in Figure 3 case 3).

Confidently assigning PTMs to specific histone residues is made even more challenging as multiple PTM combinations can correspond to the very same mass increment. For instance, as already mentioned above, the mass of butyrylation equals the sum of the masses of methylation and propionylation. Other possible correspondences are indicated in Figure 4. Notably, both PTMs occurring in vivo and chemical modifications that may be added during protein/peptide sample processing can affect the peptide sequences submitted to LC-MS/MS analysis. Formylation and acetylation can exist on the side chain of Lys residues, but can also modify the N-termini of peptides when formic acid and acetic acid are used during in-gel digestion and LC separation, for example. Confidently assigning a PTM to the Lys residue present at the N-terminal end of a peptide (such as K27SAPSTGGVK36) can thus be uncertain.
5.4. Distinguishing between Isobaric Amino Acid Variations and PTM Combinations

Canonical histones and their variants are highly similar and can differ by a small number of amino acid variations (Figure 1). In some cases, the mass shift induced by the amino acid change between histone variants equals the difference of mass between two PTMs. Thus, the combination of amino acid variations and PTMs can generate isobaric peptides, making the confident assignment of

Figure 4. Examples of isobaric PTM combinations. Differentiating structures in the Lys side chain are highlighted in yellow with an indication of their corresponding masses.
PTMs even more complicated. For example, the H3 family includes two canonical histones, H3.1/H3.2, and three variants, H3.3/H3.5/TS H3.4. These proteins are highly similar, as shown in Figures 1 and 5. TS H3.4 differs from canonical H3.1/2 by two amino acid residues, namely A24V and R42H. H3.1K23 has been described as acetylated in mouse studies [45]. The combination of acetylation and of a valine residue produces a peptide with the same mass as a butyrylation and an alanine residue. Hence, incomplete fragmentation makes it difficult, if not impossible, to distinguish between the TS H3.4 variant bearing an acetylation within the K23VAR sequence and canonical H3 bearing a butyrylation in the K23AAR region (Figure 3, case 4). Interestingly, modified amino acids can generate immonium ions, which, after neutral losses, yield unique diagnostic ions. These ions can be used to detect the presence of a modification within peptides. Closer examination of the MS/MS spectrum in case 4 from Figure 3 revealed the presence of an ion characteristic of acetylated lysine at m/z 126.091, whereas no signal was detectable at m/z 154.123, which would have indicated the presence of a butyrylated Lys residue in the fragmented peptide. This observation validated the presence of an acetylated Lys in the analyzed peptide, and thus identified the sequence as K18QLATK23acVAR from TS H3.4.

**Figure 5.** Sequence alignment for mouse histone H3 and its variants.

Scrutinizing these diagnostic ions is also very useful for the delicate case of peptides which have a Lys residue in their N-terminal position. Indeed, they can discriminate between a peptide with an N-terminal lysine bearing an acetylation on its side chain, and a peptide combining a methylation on the lysine side chain and a formylation at its N-terminus as a result of using formic acid during sample preparation and analysis. Several diagnostic ions have been reported for most of the commonly investigated PTMs (Table 1). The production of immonium and diagnostic ions can be improved by using the fragmentation mode Higher energy Collisional Dissociation (HCD) rather than Collision-Induced Dissociation (CID) [46]. In addition, whereas CID spectra often lack these fragments due to the instability of lower-mass species in the ion trap, HCD spectra include them with a high mass accuracy. The large numbers of CID spectra acquired in recent years on instruments such as the hybrid LTQ-Orbitrap instruments actually provide ambiguous identifications due to the strictly isobaric nature of several PTM combinations. More frequent use of HCD fragmentation now should eliminate ambiguities by focusing on diagnostic ions. The SPIID program (Spectral Immonium Ion Detection) was recently developed to facilitate the mapping of known diagnostic ions through the binning of high-resolution MS/MS spectra, and may facilitate this task [47].
Table 1. Diagnostic ions useful for PTM assignment. Immonium ions (IM); side-chain fragmentations (SC); neutral losses (NL).

| Amino Acid Residues | PTMs            | Diagnostic Ion Type       | m/z   | MS Fragmentation | References   |
|---------------------|-----------------|--------------------------|-------|-----------------|--------------|
| Lysine              | Acetylation     | IM-NH<sub>3</sub>         | 126.091 | CID, HCD        | [48–51]      |
|                     | Methylation     | IM-NH<sub>3</sub>         | 98.0964 | CID, HCD        | [52,53]      |
|                     | Dimethylation   | IM                        | 112.1  | CID, HCD        | [52–54]      |
|                     | Trimethylation  | NL                        | 59.0735 | CID             | [51–54]      |
|                     | Proponylation   | IM                        | 140.106 | CID             | [45]         |
|                     | Crotonyl        | --                        | --     | --              | --           |
|                     | Butyryl         | IM                        | 154.123 | HCD             | [45]         |
|                     | Malonyl         | --                        | --     | --              | --           |
|                     | 2-hydroxyisobutryl | --                | --     | --              | --           |
|                     | Succinyl        | --                        | --     | --              | --           |
|                     | Glaraylt        | IM                        | 112.0756 | HCD           | [47,51,55]  |
|                     | Formyl          | --                        | --     | --              | --           |
|                     | Methylation     | NL                        | 73.064  | ETD             | [52–54,56–58]|
|                     | Dimethylation   | SC                        | 46.0651 | CID             | [59,60]      |
| Arginene            | Dimethylation-symmetric/asymmetric | SC | 71.0604 | CID | [59,60] |
|                     | Dimethyl-symmetric | NL                     | 31.0417 | CID | [56,57,61] |
|                     | Dimethylation-symmetric | NL | 46.0651 | CID | 59,60,64     |
|                     | Dimethylation-asymmetric | NL | 45.0573 | CID | [56,57,61] |
| Tyrosine            | phosphoryl      | NL                        | 216.0426 | CID             | [47,62]      |

6. Positional Isomer Quantification

Co-eluting positional isomers cannot be quantified by classical bottom-up data-dependent acquisition (DDA) because their respective signals merge at the MS1 level. To determine the relative abundance of co-eluting isobaric peptides, targeted approaches like SRM (Selected Reaction Monitoring) or PRM (Parallel Reaction Monitoring) must be implemented. Recently, Wei Li and colleagues [63] developed an SRM approach to quantify combinatorial PTMs born by histone H3 K<sup>27</sup> and K<sup>36</sup>. They successfully applied their method to investigate the distribution of modified H3K<sup>27</sup>-K<sup>36</sup> in mouse organs. Other studies also successfully used SRM and PRM to assess the level of histone H3 acetylation in human brain tissue from patients with advanced Alzheimer’s disease or to quantify low-abundance histone modifications [64,65]. Jaffe et al. [66] conducted a large-scale analysis of histone modifications and implemented PRM to quantify histone H3 modifications from 115 cell lines contained in the Cancer Cell Line Encyclopedia.

A few software tools have been developed to address the challenging process of quantifying differentially modified isobaric peptide forms. Garcia et al. [67] developed the EpiProfile program to distinguish positional isomers, using discriminating fragment ions in their MS/MS spectra and extracting the chromatographic peak area under the curve based on information available on peptide retention times. The software successfully differentiated between H3 K9/K14 mono-acetylated peptides; H3K27me1 versus H3K36me1 peptides; and H4 K5/K8/K12/K16 mono-acetylated or di-acetylated peptides.

Data-Independent Acquisition (DIA) has been effectively used to investigate histone PTMs [68–70], offering the combined benefits of SRM/PRM and DDA analysis modes. The DIA approach fragments all precursor ions within a given retention time and m/z window. As a result, DIA is unlikely to produce missing values, and thus provides higher quantitative reproducibility. Moreover, DIA can differentiate between isobaric and co-eluting peptides as data analysis and quantitation can be performed from both MS1 and MS2 scans. For instance, it can resolve and quantify isobaric H3 peptides such as the sequence K<sub>18</sub>QLATK<sub>23</sub>AAR acetylated on either K18 or K23 [68]. DIA was successfully used to profile 62 PTMs in histones H3 and H4 and to assess dynamic changes in methylation and acetylation in response to exposure to HDAC inhibitors [70]. In addition, Sidoli et al. [71] used DIA to differentially quantify isobaric phosphopeptides. It should be noted, however, that the application of this method is
still limited due to the difficulty represented by the combination of several species in the same MS/MS spectrum, the sensitivity required to detect unique fragment ions of the same type for the two isobaric isoforms, and limitations of the computational tool when calculating the relative ratio between more than two isoforms. The increasing interest in DIA should lead to the development of software for data processing which can identify and quantify isobaric peptide modifications.

7. Conclusions

Over the past decades, MS-based proteomics workflows have been established and refined for the investigation of a broad range of histone PTMs. The significant improvements to computational tools have provided new insight into developmental biology, through the identification of histone PTM crosstalk, but further developments are still needed. The discovery of an array of new PTMs (e.g., acylations) and the need to consider subtle amino acid variations among histone sequences, especially in specialized tissues such as the testis, still make the deciphering of histone PTM patterns an exquisitely complex game. Being able to comprehensively and reliably identify and quantify these combinations will pave the way for clinical epigenetics, through profiling of histone modifications for diagnostic and prognostic purposes.

Author Contributions: D.P., J.G., and S.E.K. defined the topic of the review. S.E.K. designed figures and prepared the manuscript. D.P. and J.G. corrected, edited, and finalized the manuscript. D.P., M.C., and S.E.K. interpreted mass spectrometry data. All authors read and approved the final manuscript.

Funding: S.E.K. was awarded a Ph.D. fellowship from CEA (Irtelis program) and MC benefitted from a Ph.D. fellowship from Univ. Grenoble Alpes (UGA, AGIR program). This work was supported by grants from the “Investissement d’Avenir Infrastructures Nationales en Biologie et Santé” program (ProFI project, ANR-10-INBS-08) and the French National Research Agency (GRAL project, ANR-10-LABX-49-01).

Acknowledgments: We are grateful to colleagues in EDyP for their informatics support and for technical help on the MS instruments. We thank Maighread Gallagher-Gambarelli for editing services.

Conflicts of Interest: None of the authors have any conflict of interest to declare.

References
1. Luger, K.; Mäder, A.W.; Richmond, R.K.; Sargent, D.F.; Richmond, T.J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* 1997, 389, 251–260. [CrossRef] [PubMed]
2. Kowalski, A.; Pałyga, J. Linker histone subtypes and their allelic variants. *Cell Biol. Int.* 2012, 36, 981–996. [CrossRef] [PubMed]
3. Hergeth, S.P.; Schneider, R. The H1 linker histones: Multifunctional proteins beyond the nucleosomal core particle. *EMBO Rep.* 2015, 16, 1439–1453. [CrossRef] [PubMed]
4. Bednar, J.; Garcia-Saez, I.; Boopathi, R.; Cutter, A.R.; Papai, G.; Reymer, A.; Syed, S.H.; Lone, I.N.; Tonchev, O.; Crucifix, C.; et al. Structure and Dynamics of a 197 bp Nucleosome in Complex with Linker Histone H1. *Mol. Cell* 2017, 66, 384–397. [CrossRef] [PubMed]
5. Zhao, Y.; Garcia, B.A. Comprehensive catalog of currently documented histone modifications. *Cold Spring Harb. Perspect. Biol.* 2015, 7. [CrossRef] [PubMed]
6. Lawrence, M.; Daujat, S.; Schneider, R. Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet.* 2016, 32, 42–56. [CrossRef] [PubMed]
7. Tropberger, P.; Schneider, R. Scratching the (lateral) surface of chromatin regulation by histone modifications. *Nat. Struct. Mol. Biol.* 2013, 20, 657–661. [CrossRef] [PubMed]
8. Berger, S.L. The complex language of chromatin regulation during transcription. *Nature* 2007, 447, 407–412. [CrossRef] [PubMed]
9. Vaquerizas, J.M.; Torres-Padilla, M.-E. Developmental biology: Panoramic views of the early epigenome. *Nature* 2016, 537, 494–496. [CrossRef] [PubMed]
10. Quenet, D. Histone Variants and Disease. *Int. Rev. Cell Mol. Biol.* 2018, 335, 1–39. [CrossRef] [PubMed]
11. El Kennani, S.; Adrait, A.; Shaytan, A.K.; Khochbin, S.; Bruley, C.; Panchenko, A.R.; Landsman, D.; Pfieger, D.; Govin, J. MS_HistoneDB, a manually curated resource for proteomic analysis of human and mouse histones. *Epigenet. Chromatin* 2017, 10, 2. [CrossRef] [PubMed]
12. Banaszynski, L.A.; Allis, C.D.; Lewis, P.W. Histone Variants in Metazoan Development. Dev. Cell 2010, 19, 662–674. [CrossRef] [PubMed]

13. Weber, C.M.; Henikoff, S. Histone variants: Dynamic punctuation in transcription. Genes Dev. 2014, 28, 672–682. [CrossRef] [PubMed]

14. Bonisch, C.; Hake, S.B. Histone H2A variants in nucleosomes and chromatin: More or less stable? Nucleic Acids Res. 2012, 40, 10719–10741. [CrossRef] [PubMed]

15. Hoghoughi, N.; Barral, S.; Vargas, A.; Rousseaux, S.; Khochbin, S. Histone variants: Essential actors in male genome programming. J. Biochem. 2018, 163, 97–103. [CrossRef] [PubMed]

16. Montellier, E.; Boussouar, F.; Rousseaux, S.; Zhang, K.; Buchou, T.; Fenaille, F.; Shiota, H.; Debernardi, A.; Hery, P.; Curtet, S.; et al. Chromatin-to-nucleoprotamine transition is controlled by the histone H2B variant TH2B. Genes Dev. 2013, 27, 1680–1692. [CrossRef] [PubMed]

17. El Kennani, S.; Adrait, A.; Permiakova, O.; Hesse, A.-M.; Ialy-Radio, C.; Ferro, M.; Brun, V.; Cocquet, J.; Govin, J.; Pfieger, D. Systematic quantitative analysis of H2A and H2B variants by targeted proteomics. Epigenet. Chromatin 2018, 11, 2. [CrossRef] [PubMed]

18. Baker, M. Reproducibility crisis: Blame it on the antibodies. Nature 2015, 521, 274–276. [CrossRef] [PubMed]

19. Peach, S.E.; Rudomin, E.L.; Udeshi, N.D.; Carr, S.A.; Jaffe, J.D. Quantitative assessment of chromatin immunoprecipitation grade antibodies directed against histone modifications reveals patterns of co-occurring marks on histone protein molecules. Mol. Cell. Proteom. 2012, 11, 128–137. [CrossRef] [PubMed]

20. Egelhofer, T.A.; Minoda, A.; Klugman, S.; Lee, K.; Kolasinska-Zwierz, P.; Alekseyenko, A.A.; Cheung, M.-S.; Day, D.S.; Gadel, S.; Gorchakov, A.A.; et al. An assessment of histone-modification antibody quality. Nat. Struct. Mol. Biol. 2011, 18, 91–93. [CrossRef] [PubMed]

21. Chen, Y.; Sprung, R.; Tang, Y.; Ball, H.; Sangras, B.; Kim, S.C.; Falck, J.R.; Peng, J.; Gu, W.; Zhao, Y. Lysine Propionylation and Butyrylation Are Novel Post-translational Modifications in Histones. Mol. Cell. Proteom. 2007, 6, 812–819. [CrossRef] [PubMed]

22. Tan, M.; Luo, H.; Lee, S.; Jin, F.; Yang, J.S.; Montellier, E.; Buchou, T.; Cheng, Z.; Rousseaux, S.; Rajagopal, N.; et al. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell 2011, 146, 1016–1028. [CrossRef] [PubMed]

23. Zhang, Z.; Tan, M.; Xie, Z.; Dai, L.; Chen, Y.; Zhao, Y. Identification of lysine succinylation as a new post-translational modification. Nat. Chem. Biol. 2011, 7, 58–63. [CrossRef] [PubMed]

24. Xie, Z.; Dai, J.; Dai, L.; Tan, M.; Cheng, Z.; Wu, Y.; Boeke, J.D.; Zhao, Y. Lysine Succinylation and Lysine Malonylation in Histones. Mol. Cell. Proteom. 2012, 11, 100–107. [CrossRef] [PubMed]

25. Dai, L.; Peng, C.; Montellier, E.; Lu, Z.; Chen, Y.; Ishii, H.; Debernardi, A.; Buchou, T.; Rousseaux, S.; Jin, F.; et al. Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark. Nat. Chem. Biol. 2014, 10, 365–370. [CrossRef] [PubMed]

26. Tan, M.; Peng, C.; Anderson, K.A.; Chhoy, P.; Xie, Z.; Dai, L.; Park, J.; Chen, Y.; Huang, H.; Zhang, Y.; et al. Lysine glutarylation is a protein posttranslational modification regulated by SIRT5. Cell Metab. 2014, 19, 605–617. [CrossRef] [PubMed]

27. Shechter, D.; Dormann, H.L.; Allis, C.D.; Hake, S.B. Extraction, purification and analysis of histones. Nat. Protoc. 2007, 2, 1445–1457. [CrossRef] [PubMed]

28. Sidoli, S.; Bhanu, N.V.; Karch, K.R.; Wang, X.; Garcia, B.A. Complete Workflow for Analysis of Histone Post-translational Modifications Using Bottom-up Mass Spectrometry: From Histone Extraction to Data Analysis. J. Vis. Exp. 2016, 54112. [CrossRef] [PubMed]

29. Gupta, N.; Pentakota, S.; Mishra, L.N.; Jones, R.; Rao, M.R.S. Chapter Seven—Identification of Posttranslational Modifications of Endogenous Chromatin Proteins From Testicular Cells by Mass Spectrometry. In Proteomics in Biology, Part B; Shukla, A.K., Ed.; Methods in Enzymology; Academic Press: New York, NY, USA, 2017; Volume 586, pp. 115–142.

30. Sidoli, S.; Garcia, B.A. Middle-down proteomics: A still unexploited resource for chromatin biology. Expert Rev. Proteom. 2017, 14, 617–626. [CrossRef] [PubMed]

31. Toby, T.K.; Fornelli, L.; Kelleher, N.L. Progress in Top-Down Proteomics and the Analysis of Proteoforms. Annu. Rev. Anal. Chem. 2016, 9, 499–519. [CrossRef] [PubMed]

32. Garcia, B.A.; Mollah, S.; Ueberheide, B.M.; Busby, S.A.; Muratore, T.L.; Shabanowitz, J.; Hunt, D.F. Chemical derivatization of histones for facilitated analysis by mass spectrometry. Nat. Protoc. 2007, 2, 933–938. [CrossRef] [PubMed]
33. Witalison, E.; Thompson, P.; Hofseth, L. Drawbacks in the use of unconventional hydrophobic anhydrides for histone derivatization in bottom-up proteomics PTM analysis. Curr. Drug Targets 2015, 16, 700–710. [CrossRef] [PubMed]

34. Garcia, B.A.; Shabanowitz, J.; Hunt, D.F. Characterization of histones and their post-translational modifications by mass spectrometry. Curr. Opin. Chem. Biol. 2007, 11, 66–73. [CrossRef] [PubMed]

35. Goudarzi, A.; Zhang, D.; Huang, H.; Barral, S.; Kwon, O.K.; Qi, S.; Tang, Z.; Buchou, T.; Vitte, A.-L.; He, T.; et al. Dynamic Competing Histone H4 K5K8 Acetylation and Butyrylation Are Hallmarks of Highly Active Gene Promoters. Mol. Cell 2016, 62, 169–180. [CrossRef] [PubMed]

36. Talbert, P.B.; Ahmad, K.; Almouzni, G.; Ausió, J.; Berger, F.; Bhalla, P.L.; Bonner, W.M.; Cande, W.Z.; Chadwick, B.P.; Chan, S.W.L.; et al. A unified phylogeny-based nomenclature for histone variants. Epigenet. Chromatin 2012, 5, 7. [CrossRef] [PubMed]

37. Draizen, E.J.; Shaytan, A.K.; Mariño-Ramírez, L.; Talbert, P.B.; Landsman, D.; Panchenko, A.R. HistoneDB 2.0: A histone database with variants—An integrated resource to explore histones and their variants. Database 2016, 2016. [CrossRef] [PubMed]

38. Huang, H.; Lin, S.; Garcia, B.A.; Zhao, Y. Quantitative proteomic analysis of histone modifications. Chem. Rev. 2015, 115, 2376–2418. [CrossRef] [PubMed]

39. Yuan, Y.; Chen, W.; Cobb, M.H.; Zhao, Y. PTMap—A sequence alignment software for unrestricted, accurate, and full-spectrum identification of post-translational modification sites. Proc. Natl. Acad. Sci. USA 2009, 106, 761–766. [CrossRef] [PubMed]

40. Zhang, K.; Chen, Y.; Zhang, Z.; Zhao, Y. Identification and Verification of Lysine Propionylation and Butyrylation in Yeast Core Histones Using PTMap Software. J. Proteome Res. 2009, 8, 900–906. [CrossRef] [PubMed]

41. Pesavento, J.J.; Bullock, C.R.; LeDuc, R.D.; Mizzen, C.A.; Kelleher, N.L. Combinatorial modification of human histone H4 quantitated by two-dimensional liquid chromatography coupled with top down mass spectrometry. J. Biol. Chem. 2008, 283, 14927–14937. [CrossRef] [PubMed]

42. Feller, C.; Forné, I.; Imhof, A.; Becker, P.B. Global and Specific Responses of the Histone Acetylome to Systematic Perturbation. Mol. Cell 2015, 57, 559–571. [CrossRef] [PubMed]

43. Nien, L.; Shuai, L.; Zhu, M.; Liu, P.; Xie, Z.-F.; Jiang, S.; Jiang, H.-W.; Li, J.; Zhao, Y.; Li, J.-Y.; et al. The Landscape of Histone Modifications in a High-Fat Diet-Induced Obese (DIO) Mouse Model. Mol. Cell. Proteom. 2017, 16, 1324–1334. [CrossRef] [PubMed]

44. Kwak, H.-G.; Suzuki, T.; Dohmae, N. Global mapping of post-translational modifications on histone H3 variants in mouse testes. Biochem. Biophys. Rep. 2017, 11, 1–8. [CrossRef] [PubMed]

45. Olsen, J.V.; Macek, B.; Lange, O.; Makarof, A.; Horning, S.; Mann, M. Higher-energy C-trap dissociation for peptide modification analysis. Nat. Methods 2007, 4, 709–712. [CrossRef] [PubMed]

46. Kelstrup, C.D.; Frese, C.; Heck, A.J.R.; Olsen, J.V.; Nielsen, M.L. Analytical Utility of Mass Spectral Binning in Proteomic Experiments by SPerctonal Immonium Ion Detection (SPIID). Mol. Cell. Proteom. 2014, 13, 1914–1924. [CrossRef] [PubMed]

47. Griffiths, J.R.; Unwin, R.D.; Evans, C.A.; Latch, S.H.; Corfe, B.M.; Whetton, A.D. The Application of a Hypothesis-driven Strategy to the Sensitive Detection and Location of Acetylated Lysine Residues. J. Am. Soc. Mass Spectrom. 2007, 18, 1423–1428. [CrossRef] [PubMed]

48. Evans, C.A.; Ow, S.Y.; Smith, D.L.; Corfe, B.M.; Wright, P.C. Application of the CIRAD Mass Spectrometry Approach for Lysine Acetylation Site Discovery. In Protein Acetylation: Methods and Protocols; Hake, S.B., Janzen, C.J., Eds.; Humana Press: Totowa, NJ, USA, 2013; pp. 13–23. ISBN 978-1-62703-305-3.

49. Karch, K.R.; Zee, B.M.; Garcia, B.A. High resolution is not a strict requirement for characterization and quantification of histone post-translational modifications. J. Proteome Res. 2014, 13, 6152–6159. [CrossRef] [PubMed]

50. Kwak, H.-G.; Dohmae, N. Proteomic characterization of histone variants in the mouse testis by mass spectrometry-based top-down analysis. Biosci. Trends 2016, 10, 357–364. [CrossRef] [PubMed]
52. Zhang, K.; Yau, P.M.; Chandrasekhar, B.; New, R.; Kondrat, R.; Imai, B.S.; Bradbury, M.E. Differentiation between peptides containing acetylated or tri-methylated lysines by mass spectrometry: An application for determining lysine 9 acetylation and methylation of histone H3. *Proteomics* 2004, 4, 1–10. [CrossRef] [PubMed]

53. Couttas, T.A.; Raftery, M.J.; Bernardini, G.; Wilkins, M.R. Immonium Ion Scanning for the Discovery of Post-Translational Modifications and Its Application to Histones. *J. Proteome Res.* 2008, 7, 2632–2641. [CrossRef] [PubMed]

54. Hirota, J.; Satomi, Y.; Yoshikawa, K.; Takao, T. Epsilon-N,N,N-trimethyllysine-specific ions in matrix-assisted laser desorption/ionization-tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* 2003, 17, 371–376. [CrossRef] [PubMed]

55. Dave, K.A.; Hamilton, B.R.; Wallis, T.P.; Furness, S.G.B.; Whitelaw, M.L.; Gorman, J. Identification of N,Nε-dimethyl-lysine in the murine dioxin receptor using MALDI-TOF/TOF- and ESI-LTQ-Orbitrap-FT-MS. *Int. J. Mass Spectrom.* 2007, 268, 168–180. [CrossRef] [PubMed]

56. Snijders, A.P.L.; Hung, M.-L.; Wilson, S.A.; Dickman, M.J. Analysis of arginine and lysine methylation utilizing peptide separations at neutral pH and electron transfer dissociation mass spectrometry. *J. Am. Soc. Mass Spectrom.* 2010, 21, 88–96. [CrossRef] [PubMed]

57. Brame, C.J.; Moran, M.F.; McBroom-Cerajewski, L.D.B. A mass spectrometry based method for distinguishing between symmetrically and asymmetrically dimethylated arginine residues. *Rapid Commun. Mass Spectrom.* 2004, 18, 877–881. [CrossRef] [PubMed]

58. Ong, S.-E.; Mittler, G.; Mann, M. Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. *Nat. Methods* 2004, 1, 119–126. [CrossRef] [PubMed]

59. Gehrig, P.M.; Hunziker, P.E.; Zahariev, S.; Pongor, S. Fragmentation pathways of NG-methylated and unmodified arginine residues in peptides studied by ESI-MS/MS and MALDI-MS. *J. Am. Soc. Mass Spectrom.* 2004, 15, 142–149. [CrossRef] [PubMed]

60. Darwanto, A.; Curtis, M.P.; Schrag, M.; Crofton, A.; Trivedi, R.; Vinters, H.; Kirsch, W. Targeted proteomics for quantification of histone acetylation in Alzheimer’s disease. *Proteomics* 2012, 12, 1261–1268. [CrossRef] [PubMed]

61. Jaffe, J.D.; Wang, Y.; Chan, H.M.; Zhang, J.; Huether, R.; Kryukov, G.V.; Bhang, H.C.; Taylor, J.E.; Hu, M.; Englund, N.P.; et al. Global chromatin profiling reveals widespread full phosphorylation site occupancy during mitosis. *Sci. Signal.* 2010, 3. [CrossRef] [PubMed]

62. Yuan, Z.-F.; Lin, S.; Molden, R.C.; Cao, X.-J.; Bhanu, N.V.; Garcia, B.A. EpiProfile quantifies histone modifications by extracting retention time and intensity in high-resolution mass spectra. *Mol. Cell. Proteom.* 2015, 14, 1696–1707. [CrossRef] [PubMed]

63. Sidoli, S.; Lin, S.; Xiong, L.; Bhanu, N.V.; Karch, K.R.; Johansen, E.; Hunter, C.; Mollah, S.; Garcia, B.A. Sequential Window Acquisition of all Theoretical Mass Spectra (SWATH) Analysis for Characterization and Quantification of Histone Post-translational Modifications. *Mol. Cell. Proteom.* 2015, 14, 2420–2428. [CrossRef] [PubMed]
69. Sidoli, S.; Simithy, J.; Karch, K.R.; Kulej, K.; Garcia, B.A. Low Resolution Data-Independent Acquisition in an LTQ-Orbitrap Allows for Simplified and Fully Untargeted Analysis of Histone Modifications. *Anal. Chem.* 2015, *87*, 11448–11454. [CrossRef] [PubMed]

70. Krautkramer, K.A.; Reiter, L.; Denu, J.M.; Dowell, J.A. Quantification of SAHA-Dependent Changes in Histone Modifications Using Data-Independent Acquisition Mass Spectrometry. *J. Proteome Res.* 2015, *14*, 3252–3262. [CrossRef] [PubMed]

71. Sidoli, S.; Fujiwara, R.; Kulej, K.; Garcia, B.A. Differential quantification of isobaric phosphopeptides using data-independent acquisition mass spectrometry. *Mol. Biosyst.* 2016, *12*, 2385–2388. [CrossRef] [PubMed]

© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).