Quantitative Profiling of Histone H3 Methylation in Human Hepatocellular Carcinoma

Jiajia Chen1,2, Guoquan Yan1, Fang Wang1,2, Xuefei Yin1,2, Hong Jin1,2 and Peng-yuan Yang1,2*

1Department of Chemistry, Fudan University, 220 Handan Road, Shanghai 200433, China
2Institute of Biomedical Science, Fudan University, 138 Yixueyuan Road, Shanghai, 200032, China

Abstract

Alterations of epigenetics including DNA methylation and histone modifications play crucial roles in both initiation and progression of certain types of cancers. Abnormal histone modifications identified within these tumors represent potential biomarkers and therapeutic pathways. In this study, two different types of human hepatocellular carcinoma cell lines HepG2 and MHCC97-H with distinct metastasis capability are used to investigate the association of liver metastasis with cellular histone modification levels. In the HepG2 and MHCC97-H cells, 17 types of histone H3 modification corresponding to methylation are quantitatively profiled with LC coupled to high resolution mass spectrometry. Compared to HepG2 cells without metastasis capability, the combinatorial histone peptides of K9me1K14ac and K9me0K14ac in malignant MHCC97-H cells are found to be significantly down regulated, while the level of H3K9me3 and H3K27me2 are greatly increased. Overall these data provide novel insights into the histone methylation regulatory mechanism in liver metastasis and reveal the epigenetic pathway for HCC disease development.

Keywords: Hepatocellular carcinoma (HCC); Histone methylation; Metastasis capability; LC/MS/MS

Introduction

Histone proteins, as well as DNA, are important components of chromatin representing the physiological form of the genome [1]. Nucleosome is the repeating unit of chromatin in which genomic DNA is wrapped around a core octamer consisting of dimers of four histone proteins (H2A, H2B, H3 and H4 and/or their variant isoforms). Another histone protein H1 serves as a linker to further connect individual nucleosomes into larger chromatin fiber [2]. A large variety of post-translational modifications (PTM) are discovered on histones in the last decade. Histone modification, along with DNA methylation, noncoding RNA and chromatin remodeling contribute to the regulation mechanism of epigenetics [3-6]. They often occur on the exposed N-terminal tail, including methylation, acetylation, phosphorylation, sumoylation, ubiquitylation and some other novel PTMs. Each of these modifications has specific functions in gene regulation and other DNA-templated process and play a crucial role in stabilizing chromosomal structure and regulating gene transcription, replication, DNA repair [7,8].

Biochemically, histone basicity can be modulated by protein methylation [9], influencing any ionic interactions associated with histones residues [10]. It can lead to either gene activation or repression depending upon which residues are modified and which type of modifications is occurred. Tri-methylation of lysine 4 on histone H3 (H3K4me3) is enriched at transcriptionally active gene promoters [11], whereas di- and tri-methylation of H3K9 and tri-methylation of H3K27 are present at gene promoters that are transcriptionally repressed [12-14]. It is also reported that methylation of hisone H3 lysine 36 (H3K36me) can be activating or repressive, depending upon proximity to a gene promoter region [11]. Histone methylation has close pathological links with various types of cancer, such as breast cancer, lung cancer, chronic lymphocytic leukemia and other diseases [15]. In recent years, substantial progress has been made in understanding the molecular pathogenesis of HCC (hepatocellular carcinoma) corresponding to histone modifications [16-18]. Epigenetic silencing of tumor suppressor genes in HCC is generally mediated by DNA hypermethylation of CpG island promoters and histone modifications such as histone deacetylation, methylation of histone H3 lysine 9 (H3K9) and tri-methylation of H3K27 [19].

However, it has not yet been reported about the systematic profiling studies on H3 modifications in HCC cancer cells and its relevancy to cancer metastasis. In our study, two different types of human hepatocellular carcinoma cell lines HepG2 and MHCC97-H with distinct metastasis capability are used to investigate the association of liver metastasis with cellular histone modification levels. In the HepG2 and MHCC97-H cells, 17 types of histone H3 modification corresponding to methylation are quantitatively profiled. As compared to HepG2 cells without metastasis capability, K9me1K14ac and K9me0K14ac levels in malignant MHCC97-H cells are found to be significantly down regulated meanwhile its H3K9me3 and H3K27me2 levels are greatly increased.

Materials and Methods

Cell culture

HepG2 and MHCC97-H cell lines were obtained from Research Institute of hepatocellular carcinoma, Zhongshan Hospital. HepG2 was cultured in Eagle’s Minimum Essential Medium (DMEM) containing...
10% fetal bovine serum (FBS, Gibro) and 100 µg/mL penicillin/streptomycin (Gibro). HepG2 was treated with 0.05% trypsin-EDTA (Gibro) and harvested by centrifuging at 1200 rpm, 4°C. MHCC97-H was cultured with the same condition as HepG2, except that MHCC97-H was treated with 0.25% trypsin-EDTA. Cell pellets were store at -80°C for subsequent analysis.

**Histone extraction**

5×10^5-1×10^6 cells were collected for consequent analysis. Nuclei were isolated, and histone proteins were fractionated as described by Shechter et al. [20]. To avoid the cross-contamination of cytoplasmic proteins, nuclei were initially isolated. Briefly, cells were washed in 5 volumes of PBS (Gibro), and incubated in 10 mM Tris-Cl pH 8.0, 1 mM KCl, 1.5 mM MgCl_2, and 1 mM DTT with rotation. Nuclei were collected and re-suspended in 2 times volume of 0.4M HCl and incubated at 4°C with rotation. Collect the supernatant fraction containing acid soluble proteins and discard the pellet. Neutralize the acid extracted proteins by adding 2/5 the total volume of neutralization buffer (1M Sodium phosphate, dibasic, pH 12.5, 2.5 mM DTT, 10 mM PMSF). Store the protein at -80°C for long-term storage.

**HPLC purification and MALDI validation**

The acid extracted proteins were separated on a C8 reversed (150 mm × 4.6 mm, Agilent) column, using an Agilent 1200 system (Waldbronn, Germany) as described previously. The flow rate was 0.8 mL/min and a gradient was performed at 5% B for 5 min, 5-35% B for 30 min, 35% B maintain for 8 min, 35-55% B for 42 min, 55-90% B for 2 min, 90% B maintain for 8 min, 90% B for 2 min, and 5% B for 8 min (A=0.1% TFA in dd H_2O and B=0.1% TFA in ACN). Fractions containing histone H2A, H2B, H3 and H4 were collected separately, and the proteins were vacuum-dried completely in a SpeedVac. The collected histones were validated by MALDI-TOF mass spectrometric in a Bruker ultrafleXtreme instrument (Bruker, Bremen, Germany). A 1.0 µl aliquot of the protein solution was mixed with 1.0 µl of α-cyano-4-hydroxyquinamic acid matrix (5 mg/mL in 1:1 acetonitrile/water and 0.1% formic acid; Sigma).

**NHS-Propionate Histone Derivatization and trypsin digestion**

Reconstitute in de-ionized water the RP-HPLC-purified histone sample that had been completely dried (keep the concentration around 0.2-0.5 µg/µl). Derivatization steps are carried out according to Liao et al. [21]. Aliquot about 20 µl of histone sample and dilute with the same volume of 100mM ammonium bicarbonate buffer (pH 8). The derivatized reagent (200 mM NHS-propionate in ACN) should be freshly prepared for every reaction. Equal volume of NHS-propionate was added to histone sample and vortex to mix completely. Incubate the reaction at 50°C for 30 min and SpeedVac to dryness. The propionylated histones were reconstituted in 25 mM ammonium bicarbonate buffer (pH 8). Trypsin was added with protein: trypsin ratio of 20:1 (wt/wt). After digestion for 12 h, the samples were SpeedVac to dryness. To block the newly generated free amines, the same derivatization procedure was conducted on digested peptide.

**Results and Discussion**

**Acid extraction of nuclei**

Histone proteins are mostly located in the nucleus; therefore the nuclei extraction can enrich histone proteins to some extent. To reduce the contamination from cytoplasmic proteins, cells were first incubated in hypertonic buffer to break the cell membrane. The cytoplasm was removed and nuclei were collected. The histone proteins are basic due to its lysine and arginine amino acid in the primary sequence. Since most proteins in the nucleus are mostly acid or neutral, acid extraction is optimal for histone extraction with high efficiency and specificity. The acid extracted proteins were analyzed by BCA Protein Assay Kit (Pierce® BCA Protein Assay Kit, Thermo scientific), and the amount of extracted proteins were 103 µg for HepG2 and 59 µg for MHCC97-H respectively.

**LC-MS/MS Analysis**

The dried peptide mixtures were dissolved in 0.1% formic acid. The peptides were loaded on a C18 reversed-phase column (150 mm × 4.6 mm, 5 µm, Agilent). The derivatized peptide were HPLC separated by a Shimadzu LC20A instrument with chromatographic gradient as follows: 3% B for 3 min, 3-35% B for 39 min, 45-95% B for 3 min, 95% B for 3 min, 95-3% B for 2 min, and maintain 3% B for 10 min (phase A=0.1% TFA in H_2O ; phase B=0.15% TFA in ACN). The flow rate is 0.3 ml/min in 3 min and 0.4 ml/min from 3-60 min. The HPLC system was coupled to Q-Exact mass spectrometry instrument (ThermoFisher Scientific, San Jose, CA) acquiring a full mass scan (m/z 200-2000) in the Orbitrap analyzer with a resolution of 70,000 after accumulating 10^6 ions followed by high energy collision dissociation (HCD) fragmentation of ten most intense ions isolated by quadrupole, which is also analyzed in Orbitrap, with the resolution of 17,500. The normalized collision energy was set at 27%. All data were collected in a data-dependent mode.

**Data analysis**

The HCD mass spectra were searched by using the Proteome Discoverer 1.3 software (ThermoFisher Scientific, San Jose, CA) against human histone protein database derived from sequences obtained from the National Center for Biotechnology Information (NCBI) database. Enzyme specificity was set to trypsin, fully enzymatic, allowing for up to 3 missed cleavage sites (this is due to endogenous and chemical modification of lysine residues which hinders enzymatic digestion). Propionylation (+56.026 Da) on the N-terminus of the peptides was set as a static modification, while oxidation of methionine (+15.995 Da) was set as a variable modification for all searches. For histone PTM searches, propionylation (+56.026 Da), acetylation (+42.010 Da), mono- (+70.042 Da, as the sum of propionylation (+56.026 Da) and methylation (+41.016 Da)), di- (+28.031 Da) and trimethylation (+42.046 Da) of lysine residues were selected as variable modifications. Parent mass tolerance was set to 10 ppm and fragment ion tolerance was set to 0.05 Da. All MS/MS spectra from modified peptides were also manually inspected for accurate mass and correct fragment assignment. With Xcalibur Qual Browser (ThermoFisher Scientific, San Jose, CA), relative quantification of histone modifications was determined by measuring the area under the extracted ion chromatogram peak (XIC) corresponding to a specific modified peptide normalized with three different normalization methods (normalizing to sum of the peak areas corresponding to all observed modified forms of such peptide) housekeeping peptide) all observed modified forms of peptide respectively.

**HPLC separation, collection of histone H3 and MALDI-TOF validation**

To further separate H1, H2A, H2B, H3 and H4, the acid extracted histones were separated by C8 reversed-phase column chromatography in binary pump HPLC system. The fractions were collected using a
UV detector at a wavelength of 214 nm and collected in a time-based collection mode, with time slice of 1 min. Figure 1 shows the HPLC result of histones from MHCC97H- (up) and HepG2 (down). The collection started from 20 min and end at 55 min, to ensure all the histone proteins were collected in the elution. According to the HPLC chromatographic result and designated fractionation tubes, histone proteins can be collected separately. Elution fractions with histones are lyophilized and preserved for follow-up analysis. In the gradient elution of two samples, histone H1 was eluted first, then H2B, H2A, H4 were sequentially washed off, and eventually histone H3 was collected between 50–56 min. The retention time of histones from both samples is consistent with each other.

The collected histone proteins from MHCC97-H and HepG2 were identified by MALDI-TOF-TOF-MS from their molecular weight (Mw). According to the amino acid sequence of histone proteins and previous literature of their retention time in HPLC, we can locate H1, H2A, H2B, H3 and H4 on the HPLC chromatography [21]. The molecular weight of H3, H4H2A, H2B was about 15 KD, 11 KD, 14 KD and 13.8 KD respectively. Figure 2 showed the MALDI-TOF spectrum of each type of histone from HepG2 cells.

**LC-MS result of H3 from MHCC97-H&HepG2**

As shown in table 1, the histone endogenous peptides LVR and EIR were chosen as house-keeping peptides. Three propionylated H3 tryptic peptides (T,KQTAR, K,STGGKAPR, K,SAPATGGVKKPRH) with various modification types are identified by mass spectrometry. Table 1 summarized 17 types of histone H3 modification corresponding to various modification types are identified by mass spectrometry. Table 1 summarized 17 types of histone H3 modification corresponding to methylation and acetylation in both MHCC97-H and HepG2 cells. All peptides MS/MS spectra were manually inspected.

**Normalization with three different methods: on sum of same primary sequence, on sum of all peptides and on house-keeping peptides**

For the relative quantification of modified peptides from both MHCC97-H and HepG2 cells, XIC peak areas corresponding to each modified peptide were measured. The XIC peak areas corresponding to a specific modified peptide were further normalized with three different normalization methods: 1. Normalization based on sum of all modified forms of the same primary peptide sequence. 2. Normalization based on average peak areas of house-keeping peptides (LVR & EIR). 3. Normalization based on the peak areas of all observed modified forms of peptide.

As shown in figure 3A-3C, we found that the peptide quantitation ratios fluctuate very little with different normalization methods. Also...
the global trend of H3 methylation change on different sites (H3K4, H3K9, H3K27) keep consistent. Interestingly, the data shown in figure 3A-3C indicates that the levels of K9me1K14ac0 and K9me2K14ac1 of histone H3 are significantly down-regulated in malignant MHCC97-H cells compared to HepG2 cells with no metastasis capability. In contrast, the peptide levels of H3K9me3 and H3K27me2 are greatly up-regulated in our malignant cancer cell model MHCC97-H cells.

Figure 4 showed the base peak intensity of H3 for malignant MHCC97- H and HepG2 cells respectively. All the LC/MS/MS data are carried out from HPLC-Q-Exactive platform with triplicate. From figure 4, the BPC intensity of histone H3 from MHCC97-H and HepG2 is 8.18e+7 and 9.6e+7 respectively. In addition, the derivatization of three replicates is very specific and efficient as we did not find any side products of peptide derivatization. Meanwhile we observed that the pattern of base peak and the peak intensity from both samples were about at the similar level to facilitate relative quantitation.

In figure 5, the R.T. of K9(me)STGGK(ac)APR (left) is located at 20.31min and 20.75min in MHCC97-H & HepG2 cells respectively. The peak areas in MHCC97-H & HepG2 are manually calculated. When normalized to house-keeping peptide, log_{MHCC97-H/HepG2} results in -1.9512 (shown in figure 3); The R.T. of K9(me)STGGK(ac)APR (right) is located at 18.53min and 18.49min in MHCC97-H & HepG2 cells respectively. The peak areas in MHCC97-H & HepG2 are manually calculated, resulting log_{MHCC97-H/HepG2} value to -2.1496 (also shown in figure 3, normalize to house-keeping peptide). The value of log_{MHCC97-H/HepG2} shows that K9me1K14ac and K9me0K14ac levels in malignant MHCC97-H cells are significantly down-regulated compared to HepG2.

In figure 6, the R.T. of K17(me)STGGK(ac)APR (left) is located at 15.18min and 15.15min in MHCC97-H & HepG2 cells respectively. The peak areas in MHCC97-H & HepG2 are manually calculated. When normalized to house-keeping peptide, log_{MHCC97-H/HepG2} value to 1.93 (also shown in figure 3); The R.T. of K17(me)APATGGVK(me)KPHR(right) is located at 19.31min and 19.28min in MHCC97-H & HepG2 cells respectively. The peak areas in MHCC97-H & HepG2 are manually calculated, resulting log_{MHCC97-H/HepG2} value to 1.93 (also shown in figure 3, normalization to house-keeping peptide). We can conclude from log_{MHCC97-H/HepG2} data that H3K9me3 and H3K27me2 levels are greatly increased.

When the sample is of high complexity, the XIC chromatograms will include other ions, furthermore ionization properties of a particular modified peptide varies among different samples. Therefore, specific extraction and purification steps prior to histone derivatization are conducted in our experiment. Nuclear extraction can remove proteins from cytoplasm, and with acid extraction, most other nuclear proteins and nucleic acids will precipitate while histones are preferentially soluble in acids. Furthermore, HPLC can not only separate different type of histone but also help us to purify histone H3. So we can confirm that the system is rather simple before derivatization. Additionally, the novel derivatization method can achieve high specificity and efficiency in a reproducible manner, ensuring the simplicity and consistency among samples. As for data analysis, parent mass tolerance is set to 10 ppm and fragment ion tolerance is 0.05 Da, thus reducing the influence caused by other ions with similar m/z to some extent.

The relative quantitative profiling of histone H3 methylation modifications found on the two HCC metastasis cell line with distinct metastasis capability was obtained through LC-Q-Exactive mass
Applications of Mass Spectrometry in Epigenetics

Figure 5: Two significantly down-regulated peptides in MHCC97-H chromatogram K9 (me1)STGGK(ac1)APR17 (left) and K9 (me0)STGGK(ac0)APR17 (right) from MHCC97-H & HepG2, and MS/MS fragment spectrum K9 (me1)STGGK(ac1)APR17 (left) and K9 (me0)STGGK(ac0)APR17 (right) for further validation.
Figure 6: Two significantly up-regulated peptides in MHCC97-H. Extracted ion chromatogram of K\textsubscript{9} (me\textsubscript{3})STGGK(ac\textsubscript{0})APR\textsubscript{17} (left) and K\textsubscript{27} (me\textsubscript{2})SAPATGGVK(me\textsubscript{0})KPHR\textsubscript{40} (right) from MHCC97-H & HepG2, and MS/MS fragment spectrum K\textsubscript{9} (me\textsubscript{3})STGGK(ac\textsubscript{0})APR\textsubscript{17} (left) and K\textsubscript{27} (me\textsubscript{2})SAPATGGVK(me\textsubscript{0})KPHR\textsubscript{40} (right) for further validation.
spectrometry platform. Peptides with expression level change > 2 or < 0.5 (log_2 (change)) = 1 are defined as differentially modified. The result shows that K9me1K14ac, K9me0K14ac, H3K9me3 and H3K27me2 changes greatly in hepatocellular carcinoma cells with high metastasis. The methylation site mentioned above might eventually shed light on the future biomarker searching for hepatocellular carcinoma diagnosis, prognosis and therapy.

However, in order to understand the histone code consisting of a combinatorial set of PTMs and further reveal their biological relevance, it will be necessary to combine more advanced separation methods to distinguish between individual histone isoforms with more sensitive and accurate analytical approaches [22].

Overall, a newly developed specific derivatization method of histones with NHS-propionate is used for the research of histone H3 in two different HCC cell lines. This relatively new technology increases the sensitivity, accuracy and data acquisition time. It is helpful to optimize their histone methylation sites determination system. Moreover, this label free quantitative profiling method avoided any isobaric labeling of histone modified peptides, and is more suitable for high throughput application. It can also be used for the determination of specific histone modifications and the mechanism of gene regulation. Four quantitatively identified methylation type (K9me1K14ac, K9me0K14ac, H3K9me3 and H3K27me2) on histone H3 shows significant difference in 97H and HepG2 cell lines in our study, which represent a novel insight for the regulation of biological function on histone methylation in liver metastasis. And this can eventually provide new therapist method for HCC.

Acknowledgements

This work is supported by the National Basic Research Program of China s973 2012CB910504.

References

1. Jenuwein T, Allis CD (2001) Translating the histone code. Science 293: 1074-1080.
2. Burgess RJ, Zhang Z (2013) Histone chaperones in nucleosome assembly and human disease. Nat Struct Mol Biol 20: 14-22.
3. Black JC, Van Rechem C, Whetstine JR (2012) Histone lysine methylation dynamics: establishment, regulation, and biological impact. Mol Cell 48: 491-507.
4. Reik W (2007) Stability and flexibility of epigenetic gene regulation in mammalian development. Nature 447: 425-432.
5. Sedivy JM, Banumathy G, Adams PD (2008) Aging by epigenetics—a consequence of chromatin damage? Exp Cell Res 314: 1909-1917.
6. Young NL, DiMaggio PA, Plazas-Mayorca MD, Balbina RC, Floudas CA, et al. (2009) High throughput characterization of combinatorial histone codes. Mol Cell Proteomics 8: 2266-2284.
7. Hake SB, Garcia BA, Duncan EM, Kauer M, Delaire G, et al. (2006) Expression patterns and post-translational modifications associated with mammalian histone H3 variants. J Biol Chem 281: 559-568.
8. Latham JA, Dent SY (2007) Cross-regulation of histone modifications. Nat Struct Mol Biol 14: 1017-1024.
9. Paik WK, Kim S (1971) Protein methylation. Science 174: 114-119.
10. Erce MA, Pang CN, Hart-Smith G, Wilkins MR (2012) The methylpyroteine and the intracellular methylation network. Proteomics 12: 566-586.