

Original Article

Identification of dual histone modification-binding protein interaction by combining mass spectrometry and isothermal titration calorimetric analysis



Pu Chen^a, Zhenchang Guo^a, Cong Chen^a, Shanshan Tian^a, Xue Bai^a, Guijin Zhai^a, Zhenyi Ma^a, Huiyuan Wu^b, Kai Zhang^{a,*}

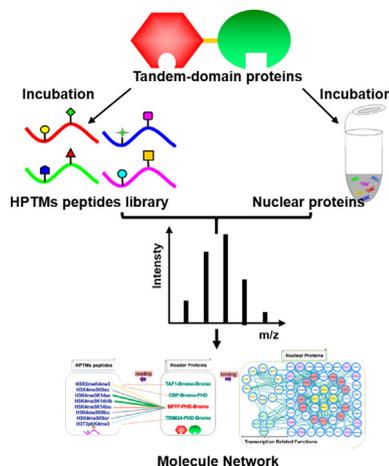
^a2011 Collaborative Innovation Center of Tianjin for Medical Epigenetics, Tianjin Key Laboratory of Medical Epigenetics, Key Laboratory of Immune Microenvironment and Disease (Ministry of Education), Department of Biochemistry and Molecular Biology, Tianjin Medical University, Tianjin 300070, China

^bSchool of Pharmaceutical Science and Technology, Tianjin University, Tianjin 300072, China

HIGHLIGHTS

- The interaction between combinatorial histone modifications and tandem-domain reader proteins was identified. Four tandem-domain proteins (BPTF-PB, CBP-BP, TRIM24-PB, TAF1-BB) could read the peptides with dual-modifications.
- The binding affinities were detected by isothermal titration calorimetry. The interaction between BPTF-PB and peptides with PTMs is the strongest.
- The binding proteins to the tandem-domains were quantified. 78 enriched proteins were further characterized.
- The molecule network of “histone modification-reader-binding proteins” was analyzed.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 6 September 2019
 Revised 29 October 2019
 Accepted 10 November 2019
 Available online 13 November 2019

Keywords:

Posttranslational modification
 Histone
 Protein-protein interaction
 Reader proteins
 Binding proteins

ABSTRACT

Histone posttranslational modifications (HPTMs) play important roles in eukaryotic transcriptional regulation. Recently, it has been suggested that combinatorial modification codes that comprise two or more HPTMs can recruit readers of HPTMs, performing complex regulation of gene expression. However, the characterization of the multiplex interactions remains challenging, especially for the molecular network of histone PTMs, readers and binding complexes. Here, we developed an integrated method that combines a peptide library, affinity enrichment, mass spectrometry (MS) and bioinformatics analysis for the identification of the interaction between HPTMs and their binding proteins. Five tandem-domain-reader proteins (BPTF, CBP, TAF1, TRIM24 and TRIM33) were designed and prepared as the enriched probes, and a group of histone peptides with multiple PTMs were synthesized as the target peptide library. First, the domain probes were used to pull down the PTM peptides from the library, and then the resulting product was characterized by MS. The binding interactions between PTM peptides and domains were further validated and measured by isothermal titration calorimetry analysis (ITC). Meanwhile, the binding proteins were enriched by domain probes and identified by HPLC-MS/MS. The

Peer review under responsibility of Cairo University.

* Corresponding author at: Department of Biochemistry and Molecular Biology, Tianjin Medical University, Tianjin 300070, China.

E-mail address: kzhang@tmu.edu.cn (K. Zhang).

<https://doi.org/10.1016/j.jare.2019.11.003>

2090-1232/© 2020 Production and hosting by Elsevier B.V. on behalf of Cairo University.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

interaction network of histone PTMs-readers-binding complexes was finally analyzed via informatics tools. Our results showed that the integrated approach combining MS analysis with ITC assay enables us to understand the interaction between the combinatorial HPTMs and reading domains. The identified network of “HPTMs-reader proteins-binding complexes” provided potential clues to reveal HPTM functions and their regulatory mechanisms.

© 2020 Production and hosting by Elsevier B.V. on behalf of Cairo University. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Recently, a number of histone posttranslational modifications (HPTMs) were reported (e.g., acetylation (ac), methylation (me), phosphorylation (ph), butyrylation (bu) [1], crotonylation (cr) [2], succinylation (su) [3], 2-hydroxybutyrylation (hib) [4] or lactylation [5]). Meanwhile, it has been suggested that two or more histone modifications may form combinatorial HPTMs, which act as a recognition platform to recruit reader proteins and further regulate gene transcription [6–9]. For example, the methylation of lysine H3 and its adjacent phosphorylation (H3S10ph) can modulate the binding of hetero-chromatin protein 1 with histone H3 and thus alter chromosome alignment and segregation [10]. Therefore, the readers that display distinct binding abilities to different HPTMs play critical roles in translating the complex PTM codes to certain meaningful biological functions [8,9,11], including transcription, cell cycle progression, cell growth and differentiation, and apoptosis [12]. The majority of reader proteins can recognize combinatorial HPTMs by their multiple domains. For example, a *trans*-histone PTM platform that was formed by tri-methylation of lysine 4 on histone H3 (H3K4me3) and acetylation of lysine 16 on histone H4 (H4K16ac) to coordinately interact with the PHD finger and Bromodomain of BPTF [13]. Recently, a dual histone methyl-lysine binding module of SHORT LIFE was reported to recognize both tri-methylation on H3K4 and H3K27 via its BAH and PHD domains, respectively [14]. Moreover, tandem-domain-reader proteins not only lead to multivalent binding of the combinatorial histone modifications but also interact with nuclear proteins to form large multiprotein complexes, which are involved in many chromatin-dependent functions [6,12,15–17]. A further complication is the fact that the “HPTMs-reader proteins-binding complexes” coordinately interact with each other to reveal epigenetic codes [8,18]. Previous reports have indicated that the network of “HPTMs-reader proteins-binding complexes” is related to numerous diseases [19].

Currently, research on the interactions between HPTMs and proteins has developed rapidly, and most studies have been performed on the peptide level, such as peptide microarrays. However, the interactions are usually transient and characterized by modest tens-to-hundreds micromolar affinity [19,20]. Peptide microarrays are largely limited to peptides containing individual PTMs, and the integrity and spatial orientation of the peptides also have some nonspecific influence on the interaction [21]. Another popular technology, chromatin immunoprecipitation (ChIP)-based methods [22], has enabled the mapping and understanding of histone modifications at the genomic level. However, ChIP has also been limited by the weakness of the antibodies, which exhibit cross-reactivity and epitope closure [23,24]. Recently high throughput strategy based on the semi-synthesis of DNA-barcoded nucleosome libraries had been developed for screening the recruitment and modulation of HPTM binders, as well as known combinatorial HPTM crosstalks [25,26]. The method is very attractive, however the construction of analysis tool still has a high requirement in histone protein synthesis and the reassemble nucleosome. To date, more and more methods, such as surface plasmon resonance imaging technique [27] and NMR spectroscopy [28] have been

implemented to profile readers of combinatorial HPTMs or to reveal their interactions. Notably, MS, as a more detailed, high-throughput, and unbiased method [29,30], plays important roles in the study of HPTMs and has promising potential to elucidate the communication of “HPTMs-reader proteins-binding complexes”.

In this work, we focused on several tandem-domain-reader proteins and analyzed their readout of HPTMs peptides and binding nucleoproteins to profile the complicated interactions between combinatorial HPTMs and tandem domains. The interactions were identified by MS, and then the binding abilities were quantified by isothermal titration calorimetry (ITC). In addition, tandem-domain proteins were used as probes to pull down the nucleoproteins. The interaction network of the histone PTMs-readers-binding complexes was analyzed by bioinformatics analysis. A schematic view of the work is provided in the supplemental information (Fig. S1). The results showed that four transcription-associated proteins, such as MBB1A, HELLS, PRKDC and TRRAP, could modify the histones alone or as a component of the complex. Finally, we constructed the network of “HPTMs-reader proteins-binding complexes” and revealed the molecular mechanism of the effects of tandem-domain protein-mediated histone crosstalk on epigenetic regulation.

Material and methods

Protein expression and purification

The chromatin-associating domains from human BPTF (PHD-Bromo 8326–8832), human CBP (Bromo-PHD 4038–4745), human TAF1 (Bromo-Bromo 4033–4908), human TRIM24 (PHD-Bromo 2658–3341), and human TRIM33 (PHD-Bromo 2789–3511) were N-terminally fused to GST. All proteins were expressed in *E. coli* Rosetta and induced overnight by 1 mM isopropyl β-D-thiogalactoside at 16 °C in LB medium supplemented with 50 mM ZnCl₂ [13]. After cell lysis and centrifugation, the supernatant was applied to Glutathione Sepharose 4B agarose (GE Healthcare, Pittsburgh, USA). The resultant plasmid sequences (BPTF-PB, CBP-BP, TRIM24-PB, TRIM33-PB, and TAF1-BB) are provided in the supplemental information (Figs. S8–S12).

Peptide pull-down

All histone peptides bearing combinational modifications were purchased from SciLight Biotechnology (China) and derived by chemical synthesis. The combinatorial HPTMs peptides were designed based on previous reports [1–4,31–34]. The peptide mixture was prepared by mixing equivalent amounts of different peptides. GST-tagged proteins were first attached to Glutathione Sepharose 4B beads (GE Healthcare, Pittsburgh, USA). After washing the beads five times (wash buffer, 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 0.05% NP40), the peptide mixture was incubated with the beads. During the peptide pull-down, the peptide mixture (100 mg/ml, 5 μl) was incubated with the GST-tagged proteins (10 mg/ml, 20 μl). After incubation, all the beads were washed by

wash buffer I two times (wash buffer I, 50 mM Tris-HCl (pH 8.0), 350 mM NaCl, 0.5% NP40), and by wash buffer II three times (wash buffer II, 50 mM Tris-HCl (pH 8.0), 350 mM NaCl, 0.05% NP40) to remove false positive binding peptides. Finally, 30% acetic acid was added, and the enriched peptides were eluted and detected by MALDI-TOF MS.

Matrix-assisted laser desorption ionization time-of-flight analysis (MALDI-TOF)

Matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) analysis was performed using Autoflex III TOF/TOF mass spectrometer (Bruker Daltonics, Leipzig, Germany) (mass tolerance: ≤ 4 ppm). The measurements were conducted in reflex positive-ion mode with delayed ion extraction. Prior to analysis, the instrument was externally calibrated with a mixture of peptide standards. 2,5-Dihydroxybenzoic acid (DHB) was used as the matrix for the analysis of peptides. Sample aliquots of 1.0 μ l were placed onto the MALDI plate. Then, 1.0 μ l of the DHB matrix was added and dried at room temperature. MS data were analyzed using Flexanalysis software (3.3.65.0) for spectral processing and peak detection.

Isothermal titration calorimetry (ITC)

For ITC measurement, synthetic histone peptides (SciLight Biotechnology, Beijing, China) and proteins were extensively dialyzed against the ITC buffer: PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4 , 2 mM KH_2PO_4) (pH 6.5). During the ITC assay procedure, four GST tagged proteins (75 μ M, 350 μ l) were titrated with each peptide (1.5 mM, 80 μ l), separately. The titration experiment was monitored using a MicroCal iTC200 system (GE Healthcare, Pittsburgh, USA) at 25 . Each ITC titration comprised 18 successive injections. Each peptide was titrated into different proteins and tested by ITC. The resultant ITC curves were processed using Origin (v.8.0) software (OriginLab) in accordance with the "One Set of Sites" fitting model.

Protein pull-down experiment

All GST-tagged proteins (BPTF-PB, CBP-BP, TRIM24-PB, and TAF1-BB) were first incubated with Glutathione Sepharose 4B beads. After washing the beads five times (wash buffer, 50 mM

Table 1
The list of HPTM peptides.

Peptide	Name	Sequence and Modification
H3(1–17)	Control	ARTKQTARKSTGGKAPR
H3(1–17)	H3K4me3	ARTK(me3)QTARKSTGGKAPR
H3(1–17)	H3K9ac	ARTKQTARK(ac)STGGKAPR
H3(1–17)	H3K9cr	ARTKQTARK(cr)STGGKAPR
H3(1–17)	H3K9bu	ARTKQTARK(bu)STGGKAPR
H3(1–17)	H3K9hib	ARTKQTARK(hib)STGGKAPR
H3(1–17)	H3R2meK4me3	AR(me)TK(me3)QTARKSTGGKAPR
H3(1–17)	H3T3phK4me3	ART(ph)K(me3)QTARKSTGGKAPR
H3(1–17)	H3K4me3K9ac	ARTK(me3)QTARK(ac)STGGKAPR
H3(1–17)	H3K4me3K9bu	ARTK(me3)QTARK(bu)STGGKAPR
H3(1–17)	H3K4me3K9cr	ARTK(me3)QTARK(cr)STGGKAPR
H3(1–17)	H3K4me3K9hib	ARTK(me3)QTARK(hib)STGGKAPR

^aAbbreviation of acylation modifications: methylation (me), tri-methylation (me3), acetylation (ac), phosphorylation (ph), butyrylation (bu), crotonylation (cr), succinylation (su), and 2-hydroxybutyrylation (hib).

Tris-HCl (pH 8.0), 150 mM NaCl, 0.05% NP40), 1 mg HEK293T nuclear extract was added to the beads which were previously bound to the tandem domain proteins and incubated overnight at 4. The nuclear proteins enriched by the tandem-domain-protein probes (BPTF-PB, CBP-BP, TRIM24-PB, and TAF1-BB) served as sample groups. The beads incubated only with 1 mg nuclear extracts served as the negative control group. After incubation, all the beads were washed with wash buffer I two times (wash buffer I, 50 mM Tris-HCl (pH 8.0), 350 mM NaCl, 0.5% NP40) and with wash buffer II three times (wash buffer II, 50 mM Tris-HCl (pH 8.0), 350 mM NaCl, 0.05% NP40) to remove false positive binding peptides. Finally, 5 \times loading buffer was added to the beads, and the mixture was boiled at 95 for 5 min. Then, the enriched proteins were separated by a 10–12% gradient PAGE gel. The gel was dealt with silver staining and subjected to LC-MS/MS analysis.

LC-MS/MS analysis

All proteins were first subject to in-gel trypsin digestion. Then, each sample of peptides was reconstituted in 7 μ l HPLC buffer A (0.1% (v/v) formic acid in water), and 5 μ l was injected into a Nano-LC system (EASY-nLC 1000, Thermo Fisher Scientific, Waltham, USA). We used C18 columns (50- μ m inner diameter \times 15 cm, 2 μ m C18) to separate each sample via an 85-minute HPLC-gradient (linear gradient from 2 to 35% HPLC buffer B and 0.1% formic acid in acetonitrile for 75 min and then to 90% buffer B in 10 min). The HPLC elution was electro-sprayed to an Orbit rap Q-Exactive mass spectrometer (Thermo Fisher Scientific, Waltham, USA). The source was operated at 1.8 kV. We carried out mass spectrometric analysis in a data-dependent mode with an automatic switch between a full MS scan and an MS/MS scan in the orbit rap. The automatic gain control (AGC) target was 3e6, and the scan range was from 400 to 1350 with a resolution of 70,000 in the full MS survey scan. We selected the 10 most intense peaks with a charge state of 2 and above for fragmentation by higher-energy collision dissociation (HCD) with a normalized collision energy of 27%. The MS2 spectra were acquired with 17,500 resolutions. Finally, we searched the MS/MS data results against the UniProt database using MaxQuant software (v1.5.2.8) with a less than 1% overall false discovery rate (FDR) for peptides. The mass tolerance of LC-MS/MS is 0.05 Da. The peptide sequences were searched using trypsin specificity with a maximum of two missed cleavages. We performed three replicate experiments to evaluate experimental reproducibility. The mass spectrometry proteomics data were deposited to the ProteomeXchange Consortium via the PRIDE [35] partner repository with dataset identifier PXD014909.

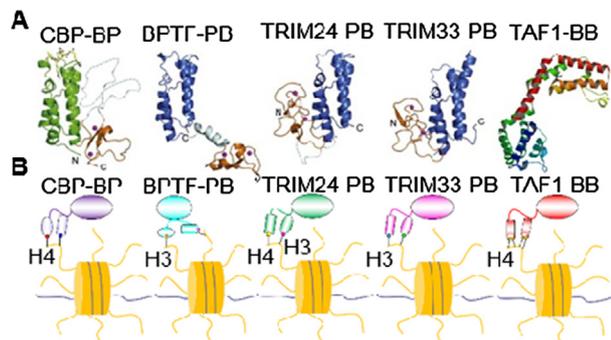


Fig. 1. The protein structures and readout of combinatorial HPTMs peptides. (A). The BPTF [13,36], TAF1 [37], CBP [38], TRIM24 [39] and TRIM33 [40,54] proteins have conservative tandem domains (BPTF PDB code 2F6J, CBP PDB code 4N4F, TAF1 PDB code 6FIC, TRIM24 PDB code 3O33, and TRIM33 PDB code 3U5M); (B). These proteins can also read combinatorial HPTMs peptides, i.e., BPTF-PH-D-Bromo (BPTF-PB) reads H3K4me3/H4K16ac [15], CBP-Bromo-PHD (CBP-PB) reads H4K12ac/K16ac [48], TRIM24-PHD-Bromo (TRIM24-PB) reads H3K4me3K23ac [39], TRIM33-PHD-Bromo (TRIM33-PB) reads H3K9me3K18ac [37,40] and TAF1-Bromo-Bromo (TAF1-BB) reads H4K5ac/K8ac [37].

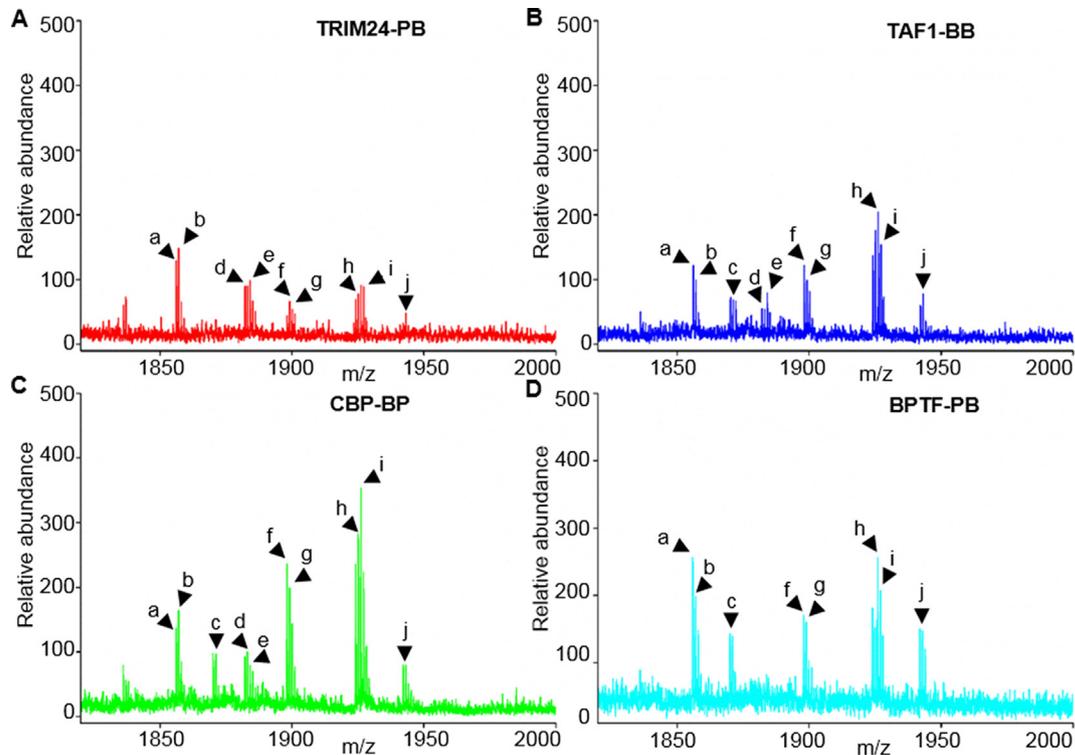


Fig. 2. Identification of combinatorial HPTM peptides by MALDI-TOF MS screening. (A). The peptides interacted with TRIM24-PB; (B). The peptides interacted with TAF1-BB; (C). The peptides interacted with CBP-BP; (D). The peptides interacted with BPTF-PB; (a. H3K4me3; b. H3K9ac; c. H3R2meK4me3; d. H3K9bu; e. H3K9cr; f. H3K9hib; g. H3K4me3K9ac; h. H3K4me3K9bu; i. H3K4me3K9cr; j. H3K4me3K9hib).

Label-free analysis

The nuclear proteins enriched by the GST-tagged tandem domain probes (BPTF-PB, CBP-BP, TRIM24-PB, and TAF1-BB) served as the sample group, while nuclear proteins enriched only by the Glutathione Sepharose 4B beads (GE Healthcare, Pittsburgh, USA) served as the control group. After SDS-PAGE, the proteins of sample group and control group were digested in gel and identified by LC-MS/MS. Every identified protein has an LFQ intensity after searching with MaxQuant software (v1.5.2.8). The ratio was obtained when the experimental LFQ intensity divided by the control LFQ intensity of each identified protein. A ratio greater than 2.0-fold (ratio > 2.0) was defined as indicative of “enriched” proteins. Then, bioinformatics analysis of enriched proteins was further carried out with the DAVID and KEGG database.

Bioinformatics analysis

Categorical annotation was performed using Gene Ontology (GO). Database for Annotation Visualization and Integrated Discovery (DAVID) was used to analyze the biological process (BP), molecular function (MF) and cellular component (CC) of the proteins. The distribution of the different proteins in the metabolic pathways was demonstrated by Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways.

Results and discussion

The structures of tandem-domain proteins and readout of combinatorial HPTMs by these proteins

The proteins BPTF [13,36], TAF1 [37], CBP [38], TRIM24 [39] and TRIM33 [40] all have tandem PHD fingers and Bromodomain, except TAF1, which contains tandem Bromodomain. The plant homeodomain (PHD) finger recognizes methylated lysine

[13,36,41–43], while Bromodomain recognizes acetylated lysine, which is involved in the suppression or activation of transcription [19,44,45]. Located in tandem with other reader domains, these domains often perform dual recognition of HPTMs, suggesting a potential cross-talk among readers [46]. PHD fingers and Bromodomain are conservative in different proteins or species [47,48]. However, the interactions or orientation in the Bromodomain-PHD or PHD-Bromodomain tandem modules are different, as shown in Fig. 1A. This finding implies that the unique function of a single domain or the whole function of tandem domains may be distinguished. It has been reported that these tandem domain proteins can read combinatorial HPTMs peptides, namely, BPTF-PHD-Bromo (BPTF-PB) reads H3K4me3/H4K16ac [15], CBP-Bromo-PHD (CBP-BP) reads H4K12acK16ac [48], TRIM24-PHD-Bromo (TRIM24-PB) reads H3K4me3K23ac [39], TRIM33-PHD-Bromo (TRIM33-PB) reads H3K9me3K18ac [40] and TAF1-Bromo-Bromo (TAF1-BB) reads H4K5ac/K8ac [37]. Here, we present schematics of the binding models of the five proteins (Fig. 1B).

To explain the relationship of the “HPTMs-reader proteins-binding complexes”, we carried out the experiments in two parts (Fig. S1). First, to profile the complicated interactions between combinatorial HPTMs and reader proteins, we incubated the peptide library with tandem-domain proteins using a new integrated method that screened and quantified the interactions by MALDI-TOF MS and ITC. Then, we used the tandem domain proteins as probes to enrich the interactive nucleoproteins and identified these proteins by HPLC-MS/MS analysis.

The MALDI-TOF MS screening of the peptides recognized by tandem-domain proteins

To reveal the transient and slight interactions between combinatorial HPTMs peptides and tandem-domain proteins, we used an integrated method that combines MS and ITC analysis.

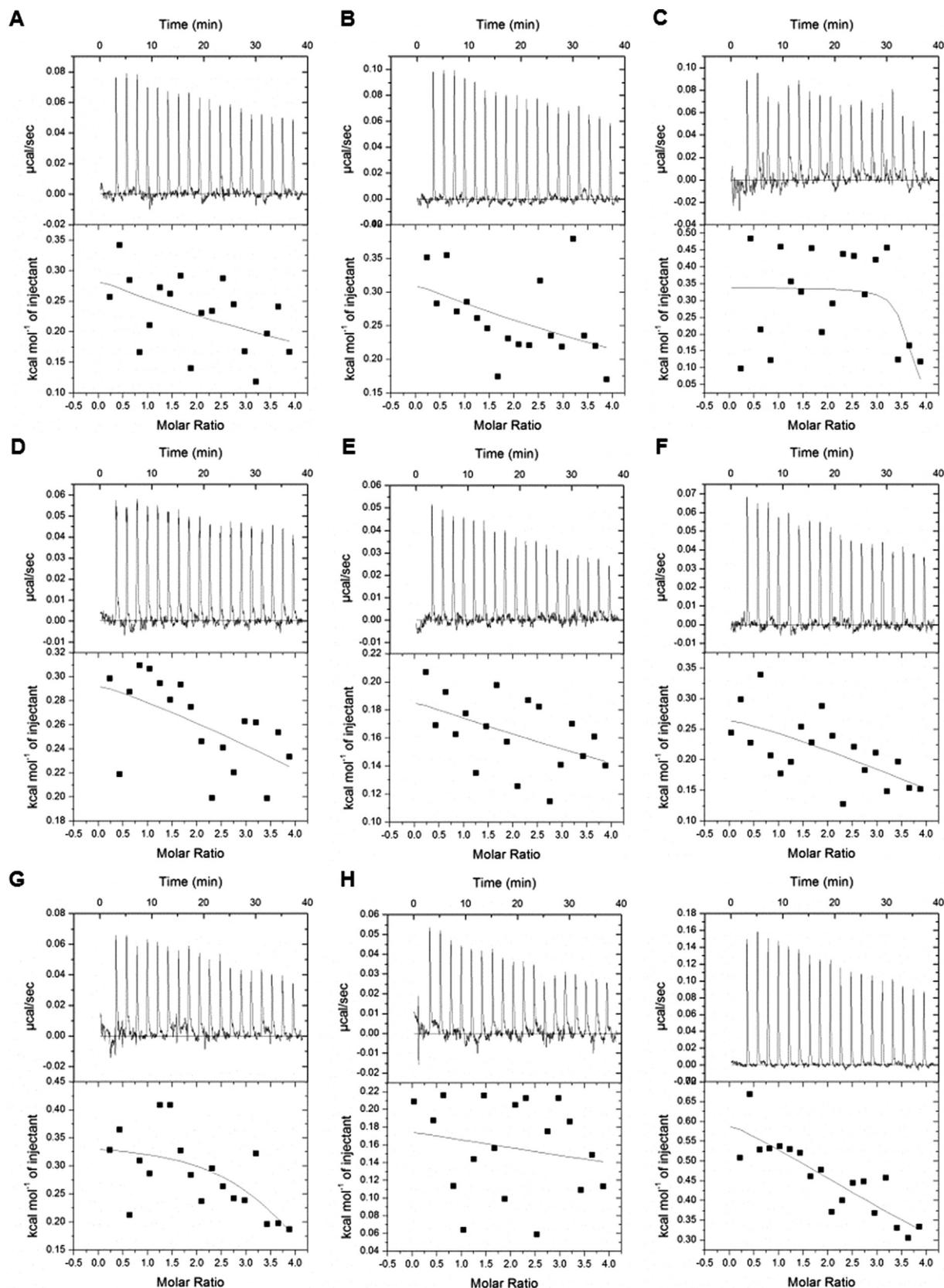


Fig. 3. Quantitative analysis of the binding ability was determined by isothermal titration calorimetry (ITC). Each tandem-domain protein was titrated with HPTM peptides. The titration in the same group was conducted under the same experimental parameters. (A). TRIM24-PB titrated with H3K4me3K9ac; (B) TRIM24-PB titrated with H3K4me3K9bu; (C). TRIM24-PB titrated with H3K4me3K9cr; (D) TRIM24-PB titrated with H3K4me3K9hib; (E). TAF1-BB titrated with H3K4me3K9ac; (F). TAF1-BB titrated with H3K4me3K9bu; (G). TAF1-BB titrated with H3K4me3K9cr; (H) TAF1-BB titrated with H3K4me3K9hib; (I). CBP-PB titrated with H3R2meK4me3; (J). CBP-PB titrated with H3K4me3K9ac; (K) CBP-PB titrated with H3K4me3K9bu; (L). CBP-PB titrated with H3K4me3K9cr; (R) CBP-PB titrated with H3K4me3K9hib; (M). BPTF-PB titrated with H3R2meK4me3; (N). BPTF-PB titrated with H3K4me3K9ac; (O) BPTF-PB titrated with H3K4me3K9bu; (P). BPTF-PB titrated with H3K4me3K9cr; (Q) BPTF-PB titrated with H3K4me3K9hib.

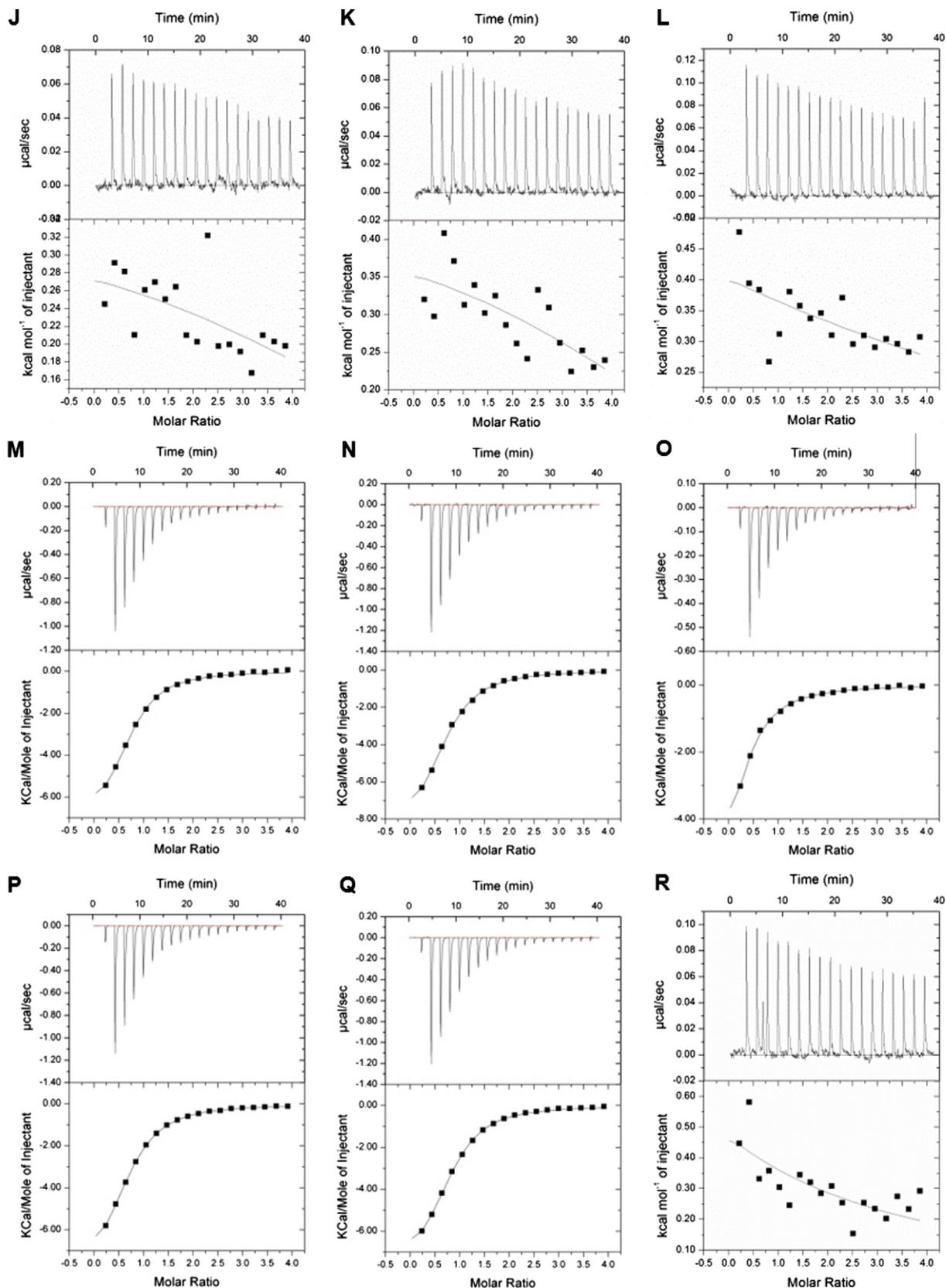


Fig. 3 (continued)

The peptide library and highly sensitive MS were used to measure the interaction between the HPTMs peptides and the tandem-domain proteins. Then, ITC was used to quantitatively verify the interactions.

We incubated the proteins with a peptide library (Table 1) that contains combinatorial modifications on different amino acid sites. The identification of tandem-domain proteins and combinatorial HPTMs peptides by MS is provided in the supplemental

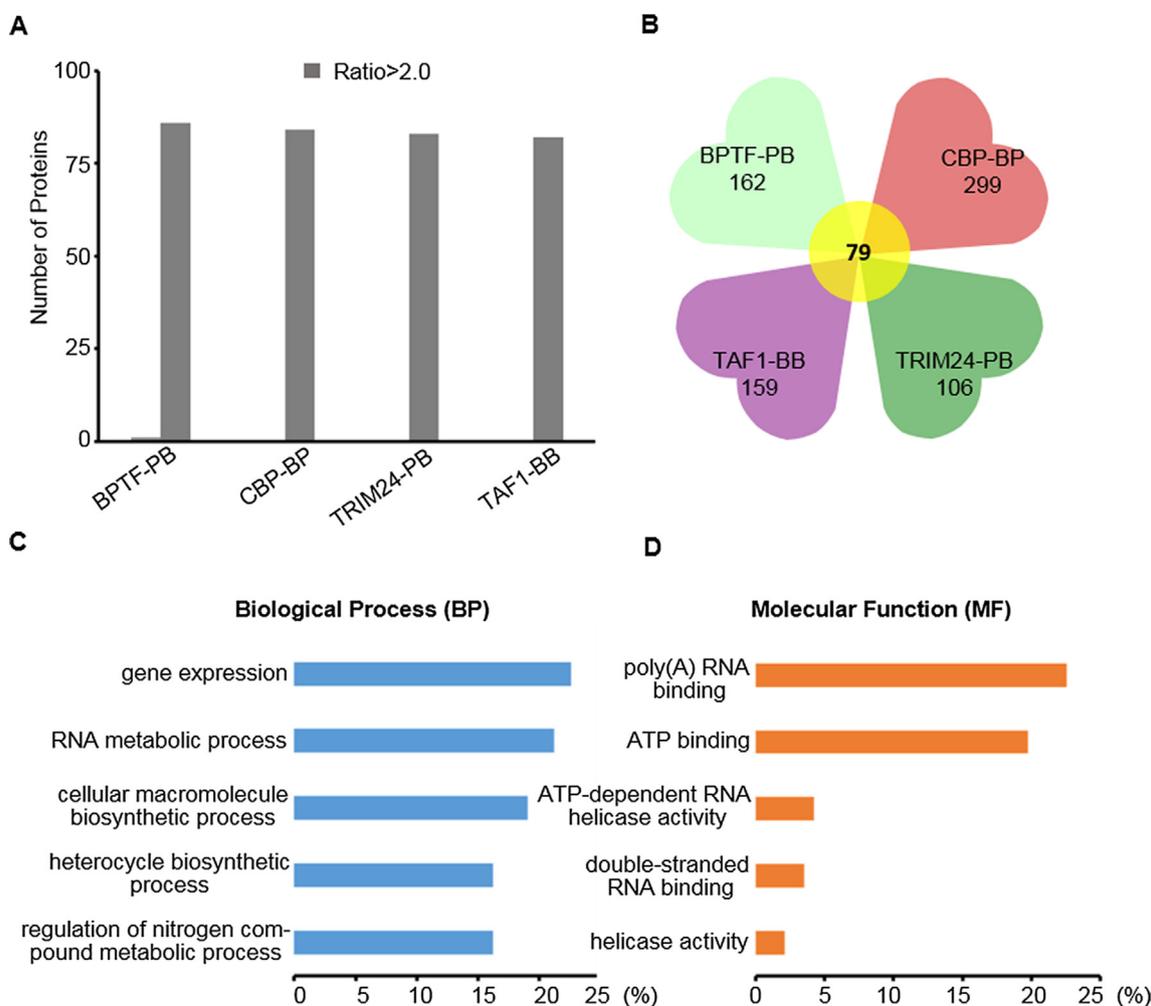


Fig. 4. Quantitative and qualitative analyses of differential proteins using label-free. (A). Enriched proteins screened by label-free quantitative analysis; (B). Quantitative analysis of proteins enriched by tandem-domain proteins; (C). The biological process of the proteins enriched by BPTF-PB, TAF1-BB, CBP-BP and TRIM24-PB were analyzed by DAVID; (D). The molecular function of the proteins enriched by BPTF-PB, TAF1-BB, CBP-BP and TRIM24-PB were analyzed by DAVID.

information (Fig. S2, Figs. S3 and S7). To exclude nonspecific binding peptides, we incubated peptides only with GST beads as the negative control for comparison with peptides incubated with BPTF-PB (Fig. S4). Then, the enriched interactive peptides were identified by MALDI-TOF MS (Fig. 2). The results showed that the peptides H3K4me3, H3K9ac, H3R2meK4me3, H3K9bu, H3K9cr, H3K9hib, H3K4me3K9ac, H3K4me3K9bu, H3K4me3K9cr and H3K4me3K9hib could be enriched by the four tandem-domain proteins, but the intensities of these peptides were different. This finding is consistent with previous reports that the domain-domain interactions or orientation between PHD fingers and Bromodomain affected the binding properties of these peptides [20]. However, we did not detect any peptides readout by TAIM33-PB.

The HPTMs readout patterns of the tandem-domain proteins revealed different readout properties. For example, the distribution of H3T3ph is opposite to that of H3K4me3, which reveals a Tph-mediated binary switch mechanism in active genes [49]. The diverse outcomes of the histone combinatorial readout by tandem-domain-reader proteins indicates a complex regulation based on multiple histone modifications. Using the integrated approach in this research, we could understand the interaction between combinatorial HPTMs and reading proteins by combining the sensitive MS technique with ITC analysis.

The quantitative analysis of the binding ability of combinatorial HPTM peptides with tandem-domain proteins by ITC

To quantify the binding affinities, we carried out ITC assays between the peptides containing combinatorial modifications and tandem-domain proteins based on the MS screening results (Fig. 3). The tandem-domain proteins used in ITC were digested by thrombin proteases to remove the GST tag and the digestion was verified by SDS-PAGE (Fig. S5). We found that only BPTF-PB could efficiently bind different combinatorial HPTM peptides. No apparent heat change was detected in the other three proteins during ITC. K_D and other thermodynamic parameters are reported in the supplemental information (Table S4).

BPTF-PB could interact with several combinatorial HPTM peptides in our peptide library: H3R2meK4me3 (K_D 13 μ M), H3K4me3K9ac (K_D 15.8 μ M), H3K4me3K9bu (K_D 27.6 μ M), H3K4me3K9cr (K_D 17.5 μ M) and H3K4me3K9hib (K_D 14.1 μ M) (Fig. 3 M–Q). The measured interactions between BPTF-PB and H3R2meK4me3, H3K4me3K9ac or H3K4me3K9cr are consistent with the previously reported interaction [13,33]. The other peptides (H3K4me3K9bu and H3K4me3K9hib) were found to interact with BPTF-PB for the first time. As known, the BPTF-PB readout of combinatorial HPTM peptides mainly depended on the PHD finger

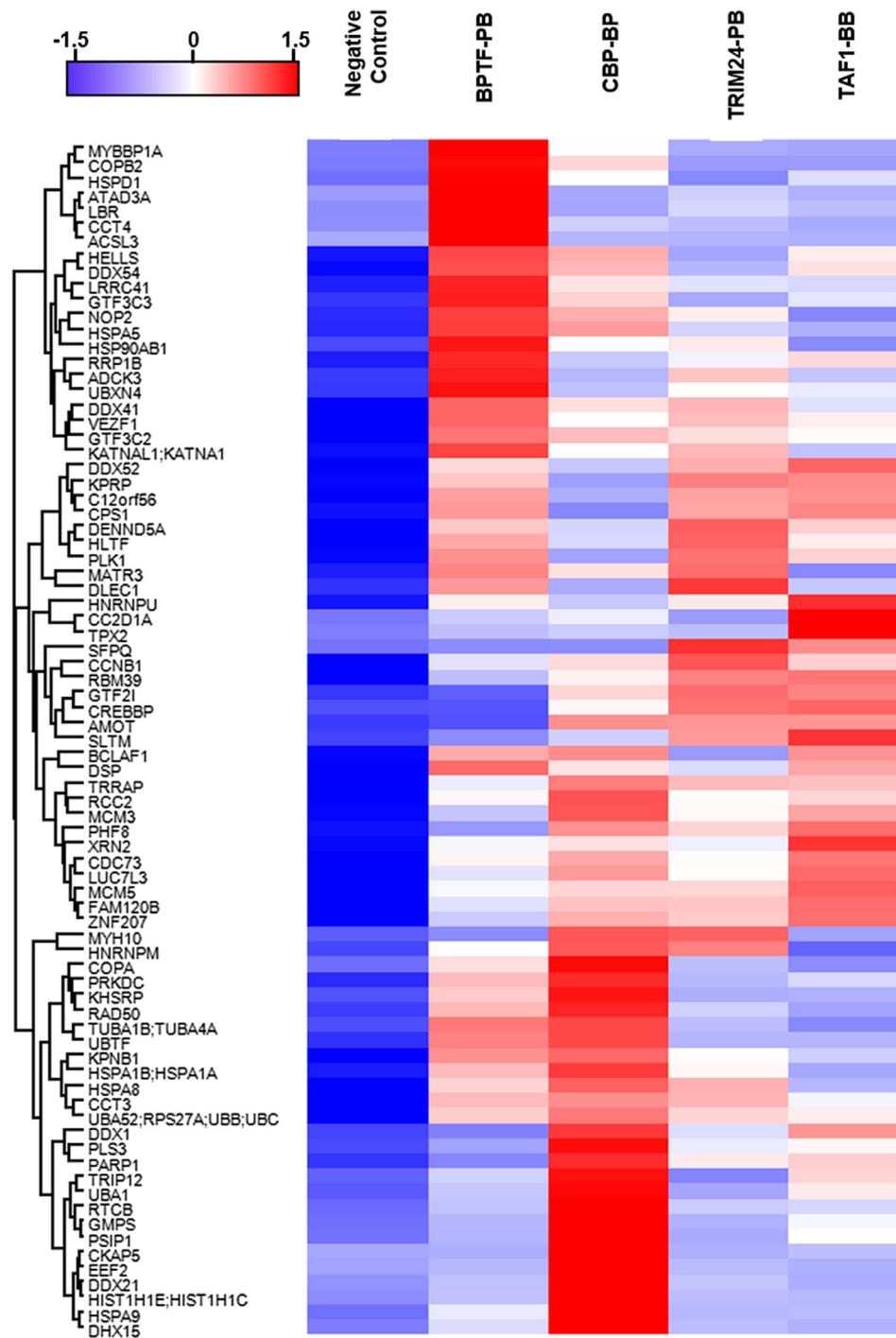


Fig. 5. Cluster analysis of the enriched proteins of BPTF-PB, CBP-BP, TRIM24-PB and TAF1-BB by Perseus software. The proteins enriched by BPTF-PB, CBP-BP, TRIM24-PB and TAF1-BB exhibited some common features. All these proteins can interact with these four tandem domain proteins.

readout of H3K4me3 [36], and the Bromodomain of BPTF-PB could read acetylated lysine, which exhibits a synergistic effect [36]. These acetylation modifications (butyrylation, crotonylation, or 2-hydroxybutyrylation) are similar to lysine acetylation but exhibit different hydrocarbon chain lengths and hydrophobicity or charges [33,34]. Thus, the modifications might have the potential to stretch into the binding pocket of Bromodomain and ultimately lead to different binding abilities of BPTF-PB. Thus, given the strong interaction between PHD and H3K4me3, the peptides with combined modifications (H3K4me3K9bu, H3K4me3K9cr and H3K4me3K9hib) also interacted with BPTF-PB. In contrast, the other three proteins CBP-BP, TRIM24-PB, and TAF1-BB did not

exhibit binding ability when titrated with the peptides in ITC. We hypothesized that these proteins lack of the molecular basis of BPTF-PB for which the interaction mainly relied on the PHD finger readout of H3K4me3 [36]. The extended hydrocarbon chains of butyrylation, crotonylation and hydroxybutyrylation on lysine increased the hydrophobicity and the bulk of the modified lysine residues in histones compared with lysine acetylation [33]. Thus, the interaction between these proteins and peptides was weak and transient, and it was difficult to detect. In general, compared with ITC, MALDI-TOF MS is a more sensitive technology to detect weak and transient interactions, and this integrated approach enabled us to detect the signal of the interactions between the

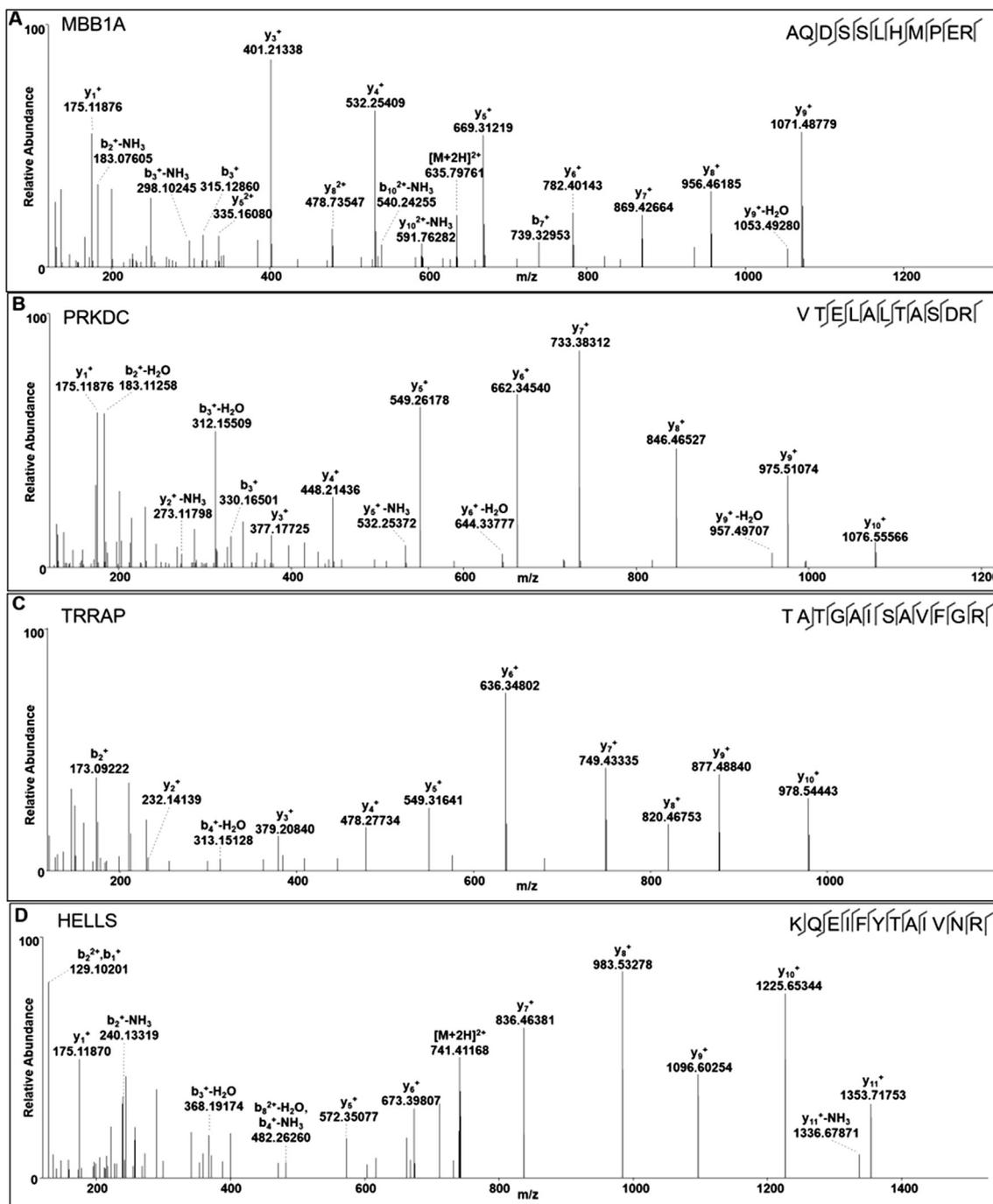


Fig. 6. Identification of the special peptides of the screened binding proteins by MS. Four proteins were screened and identified in the EpiFactors database. (A) Myb-binding protein 1A (MBB1A); (B) DNA-dependent protein kinase catalytic subunit (PRKDC); (C) transformation/transcription domain-associated protein (TRRAP); (D) Lymphoid-specific helicase (HELLS, also known as LSH).

combinatorial HPTMs and reading domains, which could not be easily observed *in vivo*.

The qualitative and quantitative analysis of differential enriched-proteins of these tandem-domain proteins by label-free analysis

To explore the nuclear proteins enriched by the tandem-domain proteins, we identified and analyzed these nucleoproteins by HPLC-MS/MS. Here, we used four different proteins, BPTF-PB, TAF1-BB, CBP-BP and TRIM24-PB, as probes that were incubated with HEK293T cell nuclear extracts.

Nucleoproteins enriched by tandem-domain-protein probes (BPTF-PB, CBP-BP, TRIM24-PB and TAF1-BB) served as the sample group, while nuclear proteins enriched only by the beads served as the control group (Fig. S6). The ratio was obtained when the experimental LFQ intensity divided by the control LFQ intensity of each identified protein. The proteins which ratio > 2.0 were defined as enriched proteins and we found that almost all the differential proteins belonged to enriched proteins (Fig. 4A) [40]. All these differential nucleoproteins could be separately enriched by BPTF-PB, TAF1-BB, CBP-BP or TRIM24-PB (Tab S1 and Tab S2), and these differential proteins were further analyzed by bioinformatics

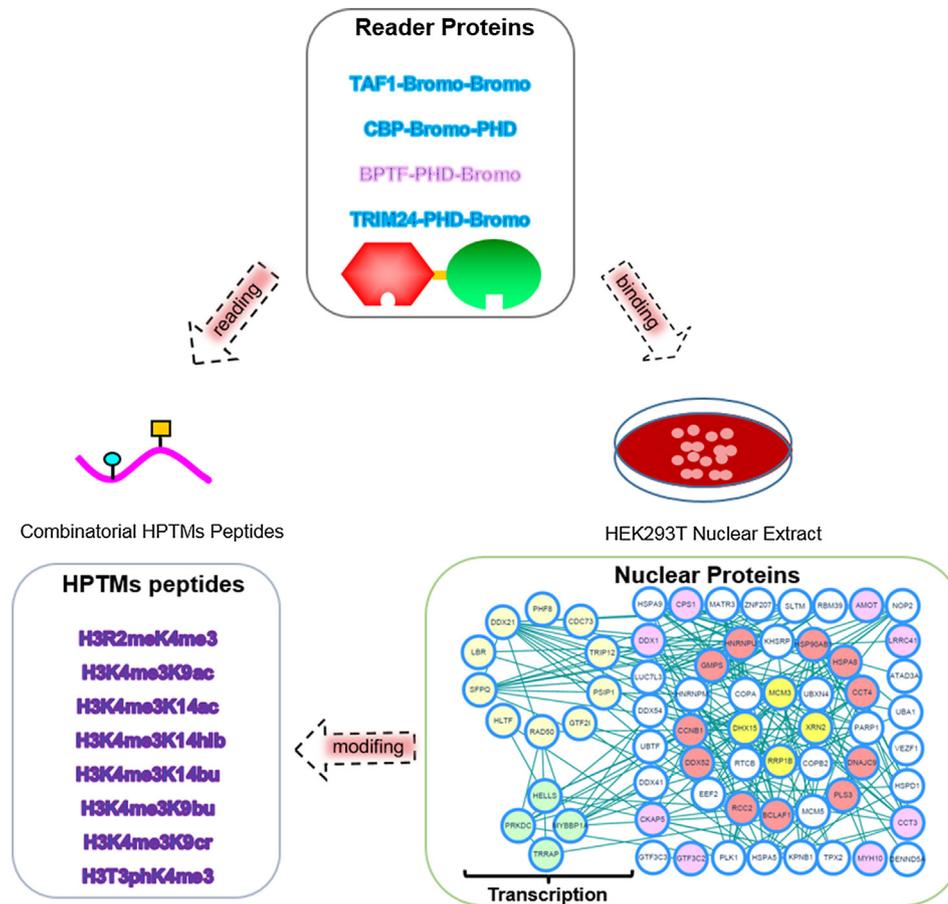


Fig. 7. The molecular network of “HPTMs-reader proteins-binding proteins”. From the network, the tandem-domain-reader proteins could read different combinatorial HPTMs peptides and interact with nucleoproteins. The interactive nucleoproteins were separated into two parts based on function, namely, transcription and other functions. Among the 78 proteins, 14 proteins (HELLS, MBB1A, PRKDC, LBR, SFPQ, PSIP1, GTF2I, HLTf, TRIP2, RAD50, CDC73, DDX21, and PHF8) have functions related to transcription, and in particular, and 4 of these proteins (HELLS, MBB1A, PRKDC, and TRRAP) modified histones directly.

(Fig. 4B). Among the enriched proteins, HELLS [50], PRKDC [51], PLK1 [52], and PSIP1 [53] have been reported to interact with BPTF, while TRIP12 [54], BPTF [55], and PARP1 [56] interact with CBP.

To annotate and assess the biological roles of these proteins, we used Database for Annotation Visualization and Integrated Discovery (DAVID) to analyze the biological process, molecular function and cellular components of the proteins (Fig. 4C and 4D). The distribution map of cluster analysis reveals that the molecular functions of these 78 enriched proteins were associated with various binding processes, such as ATP binding or RNA binding (Fig. 4D). GO biological process analysis results showed that these proteins were significantly enriched in many cellular gene expression and transcription processes (Fig. 4C). In particular, 14 of these 78 enriched proteins could bind with DNA, RNA, histone or chromatin, which directly participated in transcription-related epigenetics processes (Fig. 7). The distribution of different proteins in metabolic pathways was analyzed by Kyoto Encyclopedia of Proteins and Genomes (KEGG) pathways. These enriched proteins were mainly involved in five different metabolic pathways, particularly cell cycle and spliceosome (Tab S3). Collectively, these data revealed a high-probability role of these enriched proteins in the regulation of cellular transcription.

The cluster analysis of these enriched proteins

To discover the regulation patterns revealed by the similarity of enriched proteins, we carried out cluster analysis. We used Perseus

software (v1.5.6) to cluster these 78 enriched proteins and to identify the relationship among BPTF-PB, TAF1-BB, CBP-BP and TRIM24-PB. The result reveals that the intensities of these proteins was relatively similar, even they were enriched by different tandem-domain proteins (Fig. 5). In addition, all these nuclear proteins have functions related to multiple cellular processes, such as transcription regulation and chromatin remodeling [12]. In general, the cluster analysis of these enriched proteins is of great significance in establishing the molecular network.

The special binding proteins in the network of “HPTMs-reader proteins-binding complexes”

To further reveal the relationship of the “binding complexes” and “HPTMs”, we tried to screen and identify the proteins associated with both of them using the EpiFactors database. From the EpiFactors database, four nucleoproteins (MBB1A, HELLS, PRKDC and TRRAP) were identified that directly participated in epigenetic transcriptional regulation. These four proteins were perfectly identified by HPLC/MS-MS. The specific peptides of each protein are presented in Fig. 6.

Myb-binding protein 1A (MBB1A) [57] interacts with sequence-specific DNA binding proteins to activate or inhibit transcription. MBB1A is one of the components of the histone phosphorylation-modifying complex and chromatin-remodeling complex that can phosphorylate H2AX142. DNA-dependent protein kinase catalytic subunit (PRKDC) [58] is a serine/threonine protein kinase. As a

molecular sensor of DNA damage, PRKDC participates in double-strand breaks and repairs. PRKDC is a HPTM writer that can phosphorylate histones, especially H2AXS139 and H2AFXS139. Transformation/transcription domain-associated protein (TRRAP) [59] is a histone acetylation writer cofactor and provides specific tags for transcriptional activation. Lymphoid-specific helicase (HELLS, also as LSH) [60–62] participates in the formation of heterochromatin and chromatin remodeling. HELLS could also maintain DNA methylation and modify histones. Thus, all of these four proteins that can interact with reader proteins can also modify the histones. Finally, direct links of the network “HPTMs-reader proteins-binding complexes” have been established by these four proteins.

The molecular network of “HPTMs-reader proteins-binding complexes”

Finally, the two functions of reader proteins, which read combinatorial histone modifications and interact with nucleoproteins, were linked together to explain their relationship. As an important linker, the tandem-domain-reader proteins connect the HPTMs with binding complexes to form the network (Fig. 7). From the network, the tandem-domain-reader proteins could read different combinatorial HPTMs peptides and interact with nucleoproteins. The interactive nucleoproteins were separated into two parts based on function, namely, transcription and other functions.

These reader proteins often exist in protein complexes associated with chromatin [12,47]. In addition to HPTM-dependent histone bindings, reader proteins could also mediate other protein-protein or protein-DNA interactions, ultimately contributing to the overall chromatin interaction [46,63]. Given the presence of multiple proteins in the complex, their multivalent interaction increases the affinity and specificity and leads to different biological outcomes [15,16]. The knowledge of this network is beneficial for interpreting the mechanism and for translating this molecular network language into downstream functions given their significant roles in gene activation.

Conclusions

In the study, we established an integrated method to identify the molecular network of the “HPTMs-reader proteins-binding complexes” for understanding the molecular mechanism of HPTM. Our results showed that the method was beneficial for the comprehensive analysis of combinatorial HPTMs crosstalk. Focusing on the interaction between tandem-domain proteins and peptides containing combinatorial modifications, we found that all the tandem-domain proteins (BPTF-PB, CBP-BP, TRIM24-PB, TRIM33-PB, and TAF1-BB) have the potential to read dual-modification peptides, especially BPTF-PB, which could interact strongly with H3R2meK4me3, H3K4me3K9ac, H3K4me3K9bu, H3K4me3K9cr and H3K4me3K9hib. The results also indicated the effect of the new modifications (such as butyrylation, crotonylation, or 2-hydroxybutyrylation) on such interactions. Then, we analyzed the relationship between the tandem-domain proteins and nuclear proteins. Using GST pull-down and HPLC-MS/MS identification, we identified 78 enriched proteins and screened out differential enrichment of the reading domains. Together, we identified the interaction network of “HPTMs-reader proteins-binding complexes”, which will be helpful to understand the crosstalk between combinatorial HPTMs, to interpret the binding mechanisms of tandem domain histone readers, and to dissect the regulatory mechanism of dual HPTMs and the downstream activities regulated by histone codes.

Compliance with ethics requirements

This article does not contain any studies with human or animal subjects

Declaration of Competing Interest

The authors declare no conflicts of interest.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Nos. 21874100 and 21904097), Tianjin Municipal Science and Technology Commission (Nos. 19JCZDJC35000 and 19JCQNJC08900), and Talent Excellence Program from Tianjin Medical University.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jare.2019.11.003>.

References

- [1] Chen Y, Sprung R, Tang Y, Ball H, Sangras B, Kim SC, et al. Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Mol Cell Proteom* 2007;6:812–9.
- [2] Tan MJ, Luo H, Lee S, Jin FL, Yang JS, Montellier E, et al. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* 2011;146:1015–27.
- [3] Park J, Chen Y, Tishkoff DX, Peng C, Tan M, Dai L, et al. SIRT5-mediated lysine desuccinylation impacts diverse metabolic pathways. *Mol Cell* 2013;50:919–30.
- [4] Dai LZ, Peng C, Montellier E, Lu ZK, Chen Y, Ishii H, et al. Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark. *Nat Chem Biol* 2014;10:365–U73.
- [5] Zhang D, Tang Z, Huang H, Zhou G, Cui C, Weng Y, et al. Metabolic regulation of gene expression by histone lactylation. *Nature* 2019;574:575–80.
- [6] Kouzarides T. Chromatin modifications and their function. *Cell* 2007;128:693–705.
- [7] Strahl BD, Allis CD. The language of covalent histone modifications. *Nature* 2000;403:41–5.
- [8] Jenuwein T, Allis CD. Translating the histone code. *Science* 2001;293:1074–80.
- [9] Liu Y, Li M, Fan M, Song Y, Yu H, Zhi X, et al. Chromodomain Y-like protein-mediated histone crotonylation regulates stress-induced depressive behaviors. *Biol Psych* 2019;85:635–49.
- [10] Flanagan JF, Mi LZ, Chruszcz M, Cymborowski M, Clines KL, Kim YC, et al. Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature* 2005;438:1181–5.
- [11] Berger SL. The complex language of chromatin regulation during transcription. *Nature* 2007;447:407–12.
- [12] Musselman CA, Lalonde ME, Cote J, Kutateladze TG. Perceiving the epigenetic landscape through histone readers. *Nat Struct Mol Biol* 2012;19:1218–27.
- [13] Li HT, Ilin S, Wang WK, Duncan EM, Wysocka J, Allis CD, et al. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* 2006;442:91–5.
- [14] Qian SM, Lv XC, Scheid RN, Lu L, Yang ZL, Chen W, et al. Dual recognition of H3K4me3 and H3K27me3 by a plant histone reader SHL. *Nat Commun* 2018;9:2425.
- [15] Ruthenburg AJ, Li H, Patel DJ, Allis CD. Multivalent engagement of chromatin modifications by linked binding modules. *Nat Rev Mol Cell Biol* 2007;8:983–94.
- [16] Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* 2007;14:1025–40.
- [17] Yap KL, Zhou MM. Keeping it in the family: diverse histone recognition by conserved structural folds. *Crit Rev Biochem Mol Biol* 2010;45:488–505.
- [18] Longbotham JE, Chio CM, Dharmarajan V, Trnka MJ, Torres IO, Goswami D, et al. Histone H3 binding to the PHD1 domain of histone demethylase KDM5A enables active site remodeling. *Nat Commun* 2019;10:94.
- [19] Sanchez R, Zhou MM. The role of human bromodomains in chromatin biology and gene transcription. *Curr Opin Drug Discov Dev* 2009;12:659–65.
- [20] Filippakopoulos P, Picaud S, Mangos M, Keates T, Lambert JP, Barsyte-Lovejoy D, et al. Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* 2012;149:214–31.
- [21] Wilkinson AW, Gozani O. Histone-binding domains: strategies for discovery and characterization. *Biochim Biophys Acta* 2014;1839:669–75.

- [22] Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, et al. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res* 2012;22:1813–31.
- [23] Egelhofer TA, Minoda A, Klugman S, Lee K, Kolasinska-Zwierz P, Alekseyenko AA, et al. An assessment of histone-modification antibody quality. *Nat Struct Mol Biol* 2011;18:91–3.
- [24] Fuchs SM, Krajewski K, Baker RW, Miller VL, Strahl BD. Influence of combinatorial histone modifications on antibody and effector protein recognition. *Curr Biol* 2011;21:53–8.
- [25] Dann GP, Liszczak GP, Bagert JD, Muller MM, Nguyen UTT, Wojcik F, et al. ISWI chromatin remodellers sense nucleosome modifications to determine substrate preference. *Nature* 2017;548:607–11.
- [26] Liszczak G, Muir TW. Nucleic acid-barcoding technologies: converting DNA sequencing into a broad-spectrum molecular counter. *Angew Chem Int Ed* 2019;58:4144–62.
- [27] Zhao S, Zhang B, Yang M, Zhu J, Li H. Systematic profiling of histone readers in *Arabidopsis thaliana*. *Cell Rep* 2018;22:1090–102.
- [28] Gatchalian J, Wang XD, Ikebe J, Cox KL, Tencer AH, Zhang Y, et al. Accessibility of the histone H3 tail in the nucleosome for binding of paired readers. *Nat Commun* 2017;8:1489.
- [29] Yuan ZF, Arnaudo AM, Garcia BA. Mass spectrometric analysis of histone proteoforms. *Annu Rev Anal Chem (Palo Alto Calif)* 2014;7:113–28.
- [30] Huang H, Lin S, Garcia BA, Zhao Y. Quantitative proteomic analysis of histone modifications. *Chem Rev* 2015;115:2376–418.
- [31] Ali M, Yan K, Lalonde ME, Degerny C, Rothbart SB, Strahl BD, et al. Tandem PHD fingers of MORF/MOZ acetyltransferases display selectivity for acetylated histone H3 and are required for the association with chromatin. *J Mol Biol* 2012;424:328–38.
- [32] Minguez P, Letunic I, Parca L, Bork P. PTMcode: a database of known and predicted functional associations between post-translational modifications in proteins. *Nucleic Acids Res* 2013;41:D306–11.
- [33] Lu Y, Xu Q, Liu Y, Yu Y, Cheng ZY, Zhao Y, et al. Dynamics and functional interplay of histone lysine butyrylation, crotonylation, and acetylation in rice under starvation and submergence. *Genome Biol* 2018;19:144.
- [34] Rousseaux S, Khochbin S. Histone acylation beyond acetylation: terra incognita in chromatin biology. *Cell J* 2015;17:1–6.
- [35] Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res* 2019;47:D442–50.
- [36] Ruthenburg AJ, Li HT, Milne TA, Dewell S, McGinty RK, Yuen M, et al. Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. *Cell* 2011;145:692–706.
- [37] Jacobson RH, Ladurner AG, King DS, Tjian R. Structure and function of a human TAFII250 double bromodomain module. *Science* 2000;288:1422–5.
- [38] Kremer EA, N Gaur, MA Lee, O Engmann, J Bohacek, IM Mansuy. Interplay between TETs and microRNAs in the adult brain for memory formation. *Scientific Reports*. 2018;8:1678.
- [39] Tsai WW, Wang Z, Yiu TT, Akdemir KC, Xia W, Winter S, et al. TRIM24 links a non-canonical histone signature to breast cancer. *Nature* 2010;468:927–32.
- [40] Xi Q, Wang Z, Zaromytidou AI, Zhang XH, Chow-Tsang LF, Liu JX, et al. A poised chromatin platform for TGF-beta access to master regulators. *Cell* 2011;147:1511–24.
- [41] Pascual J, Martinez-Yamout M, Dyson HJ, Wright PE. Structure of the PHD zinc finger from human Williams-Beuren syndrome transcription factor. *J Mol Biol* 2000;304:723–9.
- [42] Shi X, Hong T, Walter KL, Ewalt M, Michishita E, Hung T, et al. ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. *Nature* 2006;442:96–9.
- [43] Wysocka J, Swigut T, Xiao H, Milne TA, Kwon SY, Landry J, et al. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* 2006;442:86–90.
- [44] Dhalluin C, Carlsson JE, Zeng L, He C, Aggarwal AK, Zhou MM. Structure and ligand of a histone acetyltransferase bromodomain. *Nature* 1999;399:491–6.
- [45] Sartor GC, Malvezzi AM, Kumar A, Andrade NS, Wiedner HJ, Vilca SJ, et al. Enhancement of BDNF expression and memory by HDAC inhibition requires BET bromodomain reader proteins. *J Neurosci* 2019;39:612–26.
- [46] Wang Z, Patel DJ. Combinatorial readout of dual histone modifications by paired chromatin-associated modules. *J Biol Chem* 2011;286:18363–8.
- [47] Yun M, Wu J, Workman JL, Li B. Readers of histone modifications. *Cell Res* 2011;21:564–78.
- [48] Plotnikov AN, Yang S, Zhou TJ, Rusinova E, Frasca A, Zhou MM. Structural insights into acetylated-histone H4 recognition by the bromodomain-PHD finger module of human transcriptional coactivator CBP. *Structure* 2014;22:353–60.
- [49] Bai X, Bi WJ, Dong HY, Chen P, Tian SS, Zhai GJ, et al. An integrated approach based on a DNA Self-assembly technique for characterization of crosstalk among combinatorial histone modifications. *Anal Chem* 2018;90:3692–6.
- [50] Kimura M, Morinaka Y, Imai K, Kose S, Horton P, Imamoto N. Extensive cargo identification reveals distinct biological roles of the 12 importin pathways. *Elife* 2017;6.
- [51] Frey WD, Chaudhry A, Slepicka PF, Ouellette AM, Kirberger SE, Pomerantz WCK, et al. BPTF maintains chromatin accessibility and the self-renewal capacity of mammary gland stem cells. *Stem Cell Rep* 2017;9:23–31.
- [52] Galdeano C, Ciulli A. Selectivity on-target of bromodomain chemical probes by structure-guided medicinal chemistry and chemical biology. *Future Med Chem* 2016;8:1655–80.
- [53] Bochynska A, Luscher-Firzlaff J, Luscher B. Modes of interaction of KMT2 histone H3 lysine 4 methyltransferase/COMPASS complexes with chromatin. *Cells* 2018;7:E17.
- [54] Park HY, Lee SB, Yoo HY, Kim SJ, Kim WS, Kim JI, et al. Whole-exome and transcriptome sequencing of refractory diffuse large B-cell lymphoma. *Oncotarget* 2016;7:86433–45.
- [55] Kim RN, Moon HG, Han W, Noh DY. Perspective insight into future potential fusion gene transcript biomarker candidates in breast cancer. *Int J Mol Sci* 2018;19:E502.
- [56] Perniola R. Twenty years of AIRE. *Front Immunol* 2018;9:98.
- [57] Smith RS, Meyers DA, Peters SP, Moore WC, Wenzel SA, Bleecker ER, et al. Sequence analysis of HSPA1A and HSPA1B in a multi-ethnic study population. *DNA Seq* 2007;18:47–53.
- [58] Chen MH, Tan KT, Cheng JH, Fang WL, Yeh YC, Yeh CN, et al. A new candidate for checkpoint blockade immunotherapy?. *J Clin Oncol* 2017;35:3022.
- [59] Tapias A, Zhou ZW, Shi Y, Chong Z, Wang P, Groth M, et al. Trapp-dependent histone acetylation specifically regulates cell-cycle gene transcription to control neural progenitor fate decisions. *Cell Stem Cell* 2014;14:632–43.
- [60] Geiman TM, Durum SK, Muegge K. Characterization of gene expression, genomic structure, and chromosomal localization of Hells (Lsh). *Genomics* 1998;54:477–83.
- [61] Han YX, Ren JK, Lee E, Xu XP, Yu WS, Muegge K. Lsh/HELLS regulates self-renewal/proliferation of neural stem/progenitor cells. *Sci Rep* 2017;7:1136.
- [62] Litwin I, Bakowski T, Maciaszyk-Dziubinska E, Wysocki R. The LSH/HELLS homolog Irc5 contributes to cohesin association with chromatin in yeast. *Nucleic Acids Res* 2017;45:6404–16.
- [63] Su Z, Denu JM. Reading the Combinatorial Histone Language. *ACS Chem Biol* 2016;11:564–74.