Journal of Visualized Experiments

# Isolation of Histone from Sorghum Leaf Tissue for Top Down Mass Spectrometry Profiling of Potential Epigenetic Markers

--Manuscript Draft--

| | |
|---|---|
| **Article Type:** | Invited Methods Article - JoVE Produced Video |
| **Manuscript Number:** | JoVE61707R1 |
| **Full Title:** | Isolation of Histone from Sorghum Leaf Tissue for Top Down Mass Spectrometry Profiling of Potential Epigenetic Markers |
| **Corresponding Author:** | Ljiljana Pasa-Tolic<br>Pacific Northwest National Laboratory<br>Richland, Washington UNITED STATES |
| **Corresponding Author's Institution:** | Pacific Northwest National Laboratory |
| **Corresponding Author E-Mail:** | ljiljana.pasatolic@pnnl.gov |
| **Order of Authors:** | Mowei Zhou |
| | Shadan H. Abdali |
| | David Dilworth |
| | Lifeng Liu |
| | Benjamin Cole |
| | Neha Malhan |
| | Amir Ahkami |
| | Tanya E Winkler |
| | Joy Hollingsworth |
| | Julie Sievert |
| | Jeff Dahlberg |
| | Robert Hutmacher |
| | Mary Madera |
| | Judith A. Owiti |
| | Kim Hixson |
| | Peggy G. Lemaux |
| | Christer Jansson |
| | Ljiljana Pasa-Tolic |
| **Additional Information:** | |
| **Question** | **Response** |
| Please indicate whether this article will be Standard Access or Open Access. | Standard Access (US$2,400) |
| Please indicate the **city, state/province, and country** where this article will be **filmed**. Please do not use abbreviations. | Richland, WA, USA |

**TITLE:**

# Isolation of Histone from Sorghum Leaf Tissue for Top Down Mass Spectrometry Profiling of Potential Epigenetic Markers

**AUTHORS AND AFFILIATIONS:**

Mowei Zhou[1], Shadan H. Abdali[1], David Dilworth[2], Lifeng Liu[2], Benjamin Cole[2], Neha Malhan[1], Amir Ahkami[1], Tanya E. Winkler[1], Joy Hollingsworth[3], Julie Sievert[3], Jeff Dahlberg[3], Robert Hutmacher[4,5], Mary Madera[6], Judith A. Owiti[6], Kim Hixson[1], Peggy G. Lemaux[6], Christer Jansson[1], Ljiljana Paša-Tolić[1]

[1]Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, USA
[2]DOE-Joint Genome Institute, Lawrence Berkeley Laboratory, Berkeley, CA, USA
[3]Kearney Agricultural Research and Extension Center, University of California Agriculture and Natural Resources, Parlier, CA, USA
[4]West Side Research and Extension Center, University of California, Five Points, CA, USA
[5]Department of Plant Sciences, University of California, Davis, USA
[6]Department of Plant and Microbial Biology, University of California, Berkeley CA, USA

Corresponding author:
Ljiljana Paša-Tolić (Ljiljana.PasaTolic@pnnl.gov)

Email Addresses of Co-Authors:

| | |
|---|---|
| Mowei Zhou | (mowei.zhou@pnnl.gov) |
| Shadan Abdali | (shadan.abdali@pnnl.gov) |
| David Dilworth | (djdilworth@lbl.gov) |
| Lifeng Liu | (lifeng.liu@lbl.gov) |
| Benjamin Cole | (bjcole@lbl.gov) |
| Neha Malhan | (neharmalhan@gmail.com) |
| Amir Ahkami | (amir.ahkami@pnnl.gov) |
| Tanya Winkler | (tanya.winkler@pnnl.gov) |
| Joy Hollingsworth | (joyhollingsworth@ucanr.edu) |
| Julie Sievert | (jasievert@ucanr.edu) |
| Jeff Dahlberg | (jadahlberg@ucanr.edu) |
| Robert Hutmacher | (rbhutmacher@ucdavis.edu) |
| Mary Madera | (mary.madera2@gmail.com) |
| Judith Owiti | (adhiambojudi@yahoo.com) |
| Kim Hixson | (Kim.Hixson@pnnl.gov) |
| Peggy Lemaux | (lemauxpg@berkeley.edu) |
| Christer Jansson | (georg.jansson@pnnl.gov) |


**KEYWORDS:**

drought, epigenetic, histone clipping, post-translational modifications, proteomics, sorghum, top down mass spectrometry

**SUMMARY:**
The protocol has been developed to effectively extract intact histones from sorghum leaf materials for profiling of histone post-translational modifications that can serve as potential epigenetic markers to aid engineering drought resistant crops.

**ABSTRACT:**
Histones belong to a family of highly conserved proteins in eukaryotes. They pack DNA into nucleosomes as functional units of chromatin. Post-translational modifications (PTMs) of histones, which are highly dynamic and can be added or removed by enzymes, play critical roles in regulating gene expression. In plants, epigenetic factors, including histone PTMs, are related to their adaptive responses to the environment. Understanding the molecular mechanisms of epigenetic control can bring unprecedented opportunities for innovative bioengineering solutions. Herein, we describe a protocol to isolate the nuclei and purify histones from sorghum leaf tissue. The extracted histones can be analyzed in their intact forms by top-down mass spectrometry (MS) coupled with online reversed-phase (RP) liquid chromatography (LC). Combinations and stoichiometry of multiple PTMs on the same histone proteoform can be readily identified. In addition, histone tail clipping can be detected using the top-down LC-MS workflow, thus, yielding the global PTM profile of core histones (H4, H2A, H2B, H3). We have applied this protocol previously to profile histone PTMs from sorghum leaf tissue collected from a large-scale field study, aimed at identifying epigenetic markers of drought resistance. The protocol could potentially be adapted and optimized for chromatin immunoprecipitation-sequencing (ChIP-seq), or for studying histone PTMs in similar plants.

**INTRODUCTION:**
The increasing severity and frequency of drought is expected to affect productivity of cereal crops[1,2]. Sorghum is a cereal food and energy crop known for its exceptional ability to withstand water-limiting conditions[3,4]. We are pursuing mechanistic understanding of the interplay between drought stress, plant development, and epigenetics of sorghum [*Sorghum bicolor* (L.) Moench] plants. Our previous work has demonstrated strong connections between plant and rhizosphere microbiome in drought acclimation and responses at the molecular level[5–7]. This research will pave the way for utilizing epigenetic engineering in adapting crops to future climate scenarios. As a part of the efforts in understanding epigenetics, we aim to study protein markers that impact gene expression within the plant organism.

Histones belong to a highly conserved family of proteins in eukaryotes that pack DNA into nucleosomes as fundamental units of chromatin. Post-translational modifications (PTMs) of histones are dynamically regulated to control chromatin structure and influence gene expression. Like other epigenetic factors, including DNA methylation, histone PTMs play important roles in many biological processes[8,9]. Antibody-based assays such as western blots have widely been used to identify and quantify histone PTMs. In addition, the interaction of histone PTMs and DNA can be effectively probed by Chromatin immunoprecipitation – sequencing (ChIP-seq)[10]. In ChIP-seq, chromatin with specific targeted histone PTM is enriched by antibodies against that specific PTM. Then, the DNA fragments can be released from the enriched chromatin and sequenced. Regions

of genes that interact with the targeted histone PTM are revealed. However, all these experiments heavily rely on high quality antibodies. For some histone variants/homologs or combinations of PTMs, development of robust antibodies can be extremely challenging (especially for multiple PTMs). In addition, antibodies can only be developed if the targeted histone PTM is known.[11] Therefore, alternative methods for untargeted, global profiling of histone PTMs are necessary.

Mass spectrometry (MS) is a complementary method to characterize histone PTMs, including unknown PTMs for which antibodies are not available[11,12]. The well-established "bottom-up" MS workflow uses proteases to digest proteins into small peptides prior to liquid chromatography (LC) separation and MS detection. Because histones have large numbers of basic residues (lysine and arginine), the trypsin digestion (protease specific to lysine and arginine) in the standard bottom-up workflow cuts the proteins into very short peptides. The short peptides are technically difficult to analyze by standard LC-MS, and do not preserve the information about the connectivity and stoichiometry of multiple PTMs. The use of other enzymes or chemical labeling to block lysines generates longer peptides that are more suitable for characterization of histone PTMs[13,14].

Alternatively, the digestion step can be completely omitted. In this "top-down" approach, intact protein ions are introduced into the MS by electrospray ionization (ESI) after online LC separation, yielding ions of the intact histone proteoforms. In addition, ions (i.e., proteoforms) of interest can be isolated and fragmented in the mass spectrometer to yield the sequence ions for identification and PTM localization. Hence, top-down MS has the advantage to preserve the proteoform-level information and capture the connectivity of multiple PTMs and terminal truncations on the same proteoform[15,16]. Top-down experiments can also provide quantitative information and offer insights of biomarkers at the intact protein level[17]. Herein, we describe a protocol to extract histone from sorghum leaf and analyze the intact histones by top-down LC-MS.

The example data shown in **Figure 1** and **Figure 2** are from sorghum leaf collected at week 2 after planting. Although variation of yield is expected, this protocol is generally agnostic to specific sample conditions. The same protocol has been successfully used for sorghum plant leaf tissue collected from 2, 3, 5, 8, 9, and 10 weeks after planting.

**PROTOCOL:**

## 1. Preparing sorghum leaf material

NOTE: The sorghum plants were grown in soil in the field in Parlier, CA.

1.1. Collect sorghum leaves from plants into 50 mL centrifuge tubes and immediately freeze the tube in liquid nitrogen. Collect leaf tissue by tearing off the third and fourth fully emerged leaf from the primary tiller.

NOTE: More details of field condition, sample growth, and collection can be found in the published report[18].

1.2. Grind leaves with liquid nitrogen and immediately transfer to a centrifuge tube.

1.3. Store the ground leaf at -80 °C until use. Take about 4 g of cryo-ground leaf powder for histone analysis of each sample.

## 2. Preparing buffers and materials (3–4 h)

NOTE: The high concentration stock solutions can be made ahead of time and stored until use. But all working buffers must be made fresh on the day of the extraction (by dilution from stock and mixing with other contents) and to be placed on ice during the process. The whole experiment should be performed at 4 °C unless recommended otherwise.

2.1. Prepare 2.5 M sucrose by dissolving 42.8 g of sucrose (342.30 g/mol) in 15 mL of sterile water on heat plate in a glass container with continuous stirring. Bring up the volume to 50 mL once the sucrose has dissolved completely. Store the sucrose in 4 °C until use.

2.2. Prepare 1 M Tris pH 8 by dissolving 1.576 g of Tris HCl in 10 mL of $H_2O$ in a 15 mL centrifuge tube. Adjust pH with NaOH to 8 and check with pH paper. Store it at 4 °C until use.

2.3. Prepare 1 M Dithiothreitol (DTT) by weighing 231 mg of DTT (154.25 g/mol) and dissolving it in 1.5 mL sterile water. DTT must be made fresh or use stored frozen aliquots.

2.4. (Optional) Prepare the additional inhibitors by mixing three different salts. Prepare 18.38 mg of sodium orthovanadate (183.91 g/mol) in 1 mL of sterile water, then prepare separately sodium butyrate by adding 11.008 mg of sodium butyrate (110.09 g/mol) in 1 mL of sterile water. Prepare the final salt by adding 4.199 mg of sodium fluoride (41.99 g/mol) in 1 mL of water. Mix the three salt solutions together in equal volume as stock solution for “additional inhibitors” (33 mM of each of the three chemicals).

NOTE: Sodium vanadate polymerizes at concentrations higher than 0.1 mM under neutral pH. It is advised to activate sodium vanadate to depolymerize it for maximum efficacy following published protocols[19]. Alternatively, activated sodium vanadate is commercially available. Herein, sodium vanadate was not activated intentionally, so the efficacy does not get reduced. Activated sodium vanadate has not been tested for this protocol yet.

2.5. Prepare 1 M of $MgCl_2$ by dissolving 0.952 g of anhydrous magnesium chloride (95.2 g/mol) in 10 mL of $H_2O$ in a 15 mL centrifuge tube. Store 1 M $MgCl_2$ at 4 °C until use.

2.6. Prepare 10% (v/v) Triton X-100 by mixing 53.5 g of Triton X-100 with 35 mL of sterile water, bring up to 50 mL with water and store it at room temperature.

2.7. Prepare 5% Guanidine buffer pH7 (referred as “Gdn buffer”) that will be used to condition the resin at least overnight – prepare 0.1 M potassium hydrogen phosphate dibasic ($K_2HPO_4$) by weighing 870 mg of $K_2HPO_4$ and dissolving in 50 mL of sterile water and store at 4 °C.

2.8. Weigh 0.7 g of guanidine hydrochloride and dissolve in 0.1 M $K_2HPO_4$ to a final volume of 14 mL. Adjust pH to 7 by checking with pH paper.

2.9. Soak the dry weak cation exchange (WCX) resin in 5% Guanidine buffer pH 7 overnight. Remove the supernatant and refill with fresh 5% Gdn buffer and soak it again overnight to let the resin fully equilibrate (until the supernatant has the same pH as the original buffer).

2.10. Before starting the experiment in the next section, mix the reagents to make EB1, EB2A, and EB2B buffer based on **Table 1**. Add all inhibitors and DTT fresh just before use.

[Place **Table 1** here]

2.11. Make the Nuclei Lysis Buffer (NLB) based on **Table 2**. Prepare NLB in advance and store at 4 °C until use. Add PI tablets fresh just before use at 1x (0.5 tablet per 5 mL). See **Table 2** for specific volumes.

[Place **Table 2** here]

## 3. Nuclei isolation procedure

NOTE: It is recommended to perform steps 3.1–3.3 of the first day (2–3 h), save the nuclei in NLB buffer at -80 °C and resume the following day (or later) for protein purification (4 h). The nuclei isolation steps in this protocol were adapted from a sorghum ChIP-seq protocol being used at the Joint Genome Institute. Additional washes and sucrose gradient separation may be required to ensure nuclei purity for ChIP-seq applications.

### 3.1. Filtration of debris (~0.5 h)

3.1.1. Weigh ground leaf powder ~4 g, ensuring it remains frozen by placing on dry ice or liquid nitrogen until ready to use.

3.1.2. Add protease inhibitor tablets to EB1 to a final concentration of 0.2x (0.5 tablet for 25 mL per sample). Use a miniature plastic pestle or a pipette tip to pre-crush tablets in a microcentrifuge tube prior to adding to buffers to aid in dissolution of the tablet in the buffer. To prevent material loss, add the PI tablet and sonicate the buffer to dissolve the tablet.

3.1.3. Add 20 mL of EB1 to the frozen ground leaf powder, gently vortex and mix them until the powder is completely suspended. Keep mixing gently for ~10 min.

3.1.4. Filter through mesh 100, rinsing the filtered material twice with 2 mL of EB1 each time.

NOTE: Both the filtrate and the filtered debris should be green. If tracking using a microscope, one should be able to see intact nuclei and intact chloroplasts in the filtrate at this point. Majority of large debris should be absent/depleted. Mix dyes such as methylene blue with sample. Nuclei are easily observable as ~3–5 µm diameter dark blue/aquamarine spheres when visualized using a 20x, 40x, and/or 100x objective. Relative to nuclei, chloroplasts are similar in size, but greenish in color and often more oval in shape. Vacuoles are also similar to nuclei in size and shape, but they will not readily take up the Methylene blue dye.

3.1.5. Centrifuge the combined filtrate at 3,000 *x g* for 10 min at 4 °C in a swinging bucket rotor to pellet debris and large subcellular organelles, including nuclei and chloroplasts.

NOTE: It is recommended to prepare EB2A during this spin (see step 3.2.1).

3.1.6. Decant the supernatant, being careful to not disturb the pellet.

NOTE: As no detergent has yet been added, the pellet should remain intense green and the supernatant should be, at most, pale green/yellow.

3.2 Lysis of non-target organelles (~0.5 h)

3.2.1 Prepare EB2A by adding protease inhibitors to a final concentration of 0.4x (0.5 tablet per 12.5 mL EB2A).

3.2.2 Resuspend the pellet from step 3.1.6 in 5 mL of EB2A and incubate on ice for 10 min with gentle mixing.

NOTE: The detergent concentration needs to be optimized to preferentially lyse intact cells and chloroplasts but not nuclei. The amount required can vary among organisms. It is recommended to check for lysis of chloroplasts and retention of intact nuclei under microscope.

3.2.3 Centrifuge at 2,100 *x g* for 15 min at 4 °C in a swinging bucket rotor to pellet debris and nuclei.

NOTE: At this stage, the supernatant should be intensely green, and the pellet should be much less green than observed in the previous stages due to the lysis of chloroplasts and chlorophyll release into the cytosol.

3.2.3 Decant the supernatant, being careful to not disturb the pellet.

**3.3 Isolation of nuclei from remaining cytoplasmic contaminants (~0.5 h)**

3.3.1 Prepare EB2B by adding protease inhibitors to a final concentration of 1x (0.5 tablet per 5 mL EB2B).

3.3.2 Resuspend crude nuclear pellet from step 3.2.3 in 2 mL of EB2B.

NOTE: EB2B does not contain Triton X-100, so no additional lysis should occur at this point.

3.3.3 Centrifuge at 2,100 *x g* for 15 min at 4 °C in a swinging bucket rotor to pellet debris and nuclei.

NOTE: Small organelles and cytoplasmic components should not pellet, so they should remain in the supernatant.

3.3.4 Decant the supernatant, being careful to not disturb the pellet.

3.3.5 Resuspend the pellet using 250 µL of NLB (add 0.5 protease inhibitor tablet fresh for 5 mL).

NOTE: The goal is to resuspend the nuclei in a minimum amount of NLB without significant material loss. Because NLB is very viscous and the pellets contain a large amount of insoluble debris, it is very difficult to pipette and tends to cling to the inside of pipette tips. For this reason, it is recommended to reuse the same pipette tip whenever possible. If concerned with residual material in a pipette tip, simply hang the pipette from a shelf or rack for ~1 min to allow gravity to collect material at the opening of the tip. Do not aggressively pipette to resuspend the pellets. Instead, use the pipette tip as a stir rod until the pelleted material can be aspirated into the pipette tip. i.e., it is perfectly fine for large pellet clumps to stay at this stage so long as it can be drawn into a pipette tip.

3.3.6 Vortex 15 s at max to homogenize and partially resuspend the material. Sonicate for 5 min at 4 °C, then store at -80 °C.

NOTE: For subsequent steps, keep in mind that the total amount of NLB added is 250 µL, but the total apparent volume of the sample can be up to twice as much due to insoluble debris. The sample is frozen and thawed to assist in the lysis of nuclei.

3.4 Nuclei lysis and histone extraction (~4 h)

3.4.1 Add 750 µL of 5% Gdn buffer to the thawed sample. Sonicate for 15 min at 4 °C.

3.4.2 Transfer sample into a single 2 mL tube and spin 10,000 *x g* for 10 min at 4 °C.

NOTE: The supernatant will likely look green. The following chromatography steps should remove most of the pigments from the protein.

3.4.3 While waiting on step 3.4.1 and 3.4.2, prepare the column for ion exchange chromatography clean up. Rinse the chromatography column with 2 mL of acetonitrile and 4 mL of water to minimize contamination on surface.

3.4.4 Load 200~300 µL of WCX resin (pre-conditioned with 5% Gdn buffer) onto the chromatography column. Let the resin settle. Wash four times with 1 mL of 5% Gdn buffer. Keep the tube and column on ice for the rest of the purification steps.

3.4.5 Put the chromatography column on a 2 mL collection tube. Load the supernatant from step 3.4.2 slowly onto the resin bed without disrupting the resin (try to slowly drop from the side of the tubes). Let the solution flow through by gravity. As the solution is flowing through, load the eluent back to the top of the column 6–8 times to allow maximum binding to the resin. Then, discard the eluent.

3.4.6 Load 2 mL of 5% Gdn buffer to wash non-histone proteins off the column. Discard the eluent.

3.4.7 Elute histones with 1 mL 20% Gdn buffer. Collect the eluent, which contains histone proteins.

3.4.8 Use 3 kDa molecular weight cut off (MWCO) spin filter (0.5 mL) to desalt the eluent from step 3.4.6. Before use, load 500 µL wash solvent (0.2% formic acid in 3% ACN) and spin it down twice to clean the filter.

NOTE: It is recommended to start washing the MWCO filter while performing the resin chromatography steps to save time. The following spin filter steps take ~3–4 h.

3.4.9 First load 500 µL of histone sample, spin at 14,000 *x g* for ~25 min to reduce volume down to ~100 µL. Then load another 400 µL of sample and spin at 14,000 *x g* again for ~25 min. Load the final 100 µL of sample, rinse the sample tube with 300 µL wash solvent and load the solvent into the filter. Spin at 14,000 *x g* again for ~25 min.

3.4.10 Load 400 µL wash solvent, spin at 14, 000 *x g* for ~25 min to reduce volume to ~100 µL or less. Each cycle reduces the salt concentration by one-fifth. Repeat for another three cycles to bring guanidine concentration to ~0.01%. Reverse the filter into a clean collection tube and spin at 1,000 *x g* for 2 min. Save the purified histone sample at -20 °C or -80 °C for analysis.

NOTE: It is recommended to spin longer (30–40 min) at the last step to minimize sample volume to obtain higher concentration. The volume should be able to go down to 50–70 µL.

## 4. Mass spectrometry of purified histones

### 4.1 Liquid chromatography mass spectrometry (LC-MS) data acquisition

4.1.1 Estimate protein concentration by Bicinchoninic Acid (BCA) assay following the manufacturer's protocol.

NOTE: BCA can only provide an estimate of total protein concentration, but not the quality of histone purification. If MS instrumentation is not readily available for checking the quality of histone purification, western blot can be used. Reversed-phase LC coupled with 210 nm ultraviolet absorbance detection as described in our previous report can be also used[20]. The chromatogram can be compared with a known standard for checking sample quality. However, different organisms can have different elution profiles. Therefore, using histone standards from similar organisms is highly recommended.

4.1.2 Connect a C18 reversed phase (RP) analytical column (e.g., 3 µm 300 Å, column inner diameter 75 µm, outer diameter 360 µm, length 70 cm) and a C18 trap column (e.g., 3.6 µm, column inner diameter 150 µm, outer diameter 360 µm, length 5cm) to a dual-pump nanoflow liquid chromatography system (e.g., Waters NanoAcquity). The binary solvents are A: 0.1% formic acid in water, and B: 0.1% formic acid in acetonitrile.

NOTE: The dual pump LC includes a wash pump and a gradient pump. Both pumps go through two stages in each analysis—a trapping stage followed by the analytical stage. In the trapping stage, the wash pump flows into the trap column and the gradient pump flows into the analytical column. In the analytical stage, the trap column is coupled with the analytical column, and the gradient pump flows into both columns. The wash pump then goes to the waste.

4.1.3 Trapping stage: Set up the LC method to first load 1–2 µg of histone sample onto the trap column. Desalt the sample by the wash pump at 3 µL/min 5% solvent B for 10 min. Set the analytical pump at 0.3 µL/min 5% solvent B for equilibration.

4.1.4 Analytical stage: Set the gradient pump (0.3 µL/min) to start at from 5% B and ramp to 30% at 15 min. Then, increase to 41% B at 100 min before a high organic wash up to 95% B at the end.

NOTE: The gradient can be optimized depending on the different retention profiles on individual columns. Typically, full-length histones elute around 30%–40% B on the specified LC conditions. Longer gradients can be used to increase the numbers of MS2 spectra to capture more histone proteoforms.

4.1.5 Set up data-dependent acquisition method on a high-resolution MS (e.g., Thermo Orbitrap Fusion Lumos or similar) with electron transfer dissociation (ETD) capability. Use the intact protein mode and perform all the necessary calibrations as suggested by the manufacturer. Critical parameters are described below. This will be specific to the instrument used.

4.1.5.1. MS1: scan range 600–2,000 m/z, resolution 120k (at *m/z* 200), 4 microscans, AGC target 1E6, max injection 50 ms.

4.1.5.2. MS2: resolution 120k; 1 microscan; AGC target 1E6; data dependent MS/MS: alternating ETD (25 ms reaction time, max injection time 500 ms) and higher-energy collisional dissociation (HCD, 28% normalized collision energy with ±5% stepped energy, max injection time 100 ms); isolation window of 0.6 Da; priority on highest charge states.

4.1.5.3. Dynamic exclusion: 120 s time window, ±0.7 Da mass window. Exclude charge states lower than 5 and undetermined charge states.

4.1.6. Run a few injections of peptide or histone standards on new columns to equilibrate and check the system, before running the actual samples. For running large number of samples, add short blanks or washes in between samples to minimize carry over. Let the columns equilibrate for 15–20 min at the starting condition (5% solvent B) before the next sample.

NOTE: Longer LC gradients and higher max injection time for MS2 can improve the spectral quality for identifying more histone proteoforms.

## 4.2 LC-MS data processing and proteoform identification

4.2.1 Obtain the (sorghum) protein sequence in FASTA format from JGI (https://genome.jgi.doe.gov) or UniProt (https://www.uniprot.org/).

4.2.2 Use MSConvert[21] (http://proteowizard.sourceforge.net/tools.shtml) to convert the instrument raw data files (*.raw) into mzML format.

4.2.3 Download TopPIC suite[22] (http://proteomics.informatics.iupui.edu/software/toppic/) for data processing. The program can be run in either command line or through the graphical interface.

4.2.4 Use TopFD in the TopPIC suite to deconvolute the spectra from the mzML file from step 4.2.2. The default parameters can be used. But the “precursor window” (-w) needs to be reduced to 1 m/z because a narrow isolation window is used.

4.2.5 Use TopPIC in the TopPIC suite to identify proteoforms. Most of the default parameters can be used. Set the spectrum and proteoform cutoff type to FDR (false discovery rate) and set the cutoff value to 0.01 (1% FDR) or as desired. Set the “proteoform error tolerance” to 5 (Dalton). Load the FASTA file from step 4.2.1 and the “*_ms2.msalign” file from step 4.2.4. Then start the search.

NOTE: The “proteoform error tolerance” setting will combine proteoforms with similar masses (± 5 Da) as one. This helps reduce redundancy in the proteoform counts. However, it should be used with caution because large tolerance will merge proteoforms with small or no mass differences. This parameter is only available in TopPIC version 1.3 or later.

4.2.6 The identified proteoforms can be examined in the "*_proteoform.csv" file or visualized using the Topview module under the "*_html" folder of the output.

4.2.7 The proteoforms list generated from the steps above using TopPIC annotates the histone PTMs as mass shifts. In order to localize individual PTMs, a modification file list must be included. Detailed description can be found in the TopPIC manual. Alternatively, proceed to the next step to perform a complementary data analysis using the Informed-Proteomics package[23] (https://github.com/PNNL-Comp-Mass-Spec/Informed-Proteomics).

4.2.8 Follow the instructions and use the PbfGen module to convert the instrument raw data to a PBF file. Then deconvolute the MS1 data using ProMex module to output a ms1ft file (feature list, each feature represents a unique combination of mass and retention time).

4.2.9 Create a focused FASTA for Informed-Proteomics using the identified protein list from TopPIC in step 4.2.6.

NOTE: Searching the entire genome using Informed-Proteomics with large number of variable PTMs can be extremely slow and may cause crashes. Therefore, it is recommended to reduce the size of FASTA by only including the target proteins.

4.2.10 Create a targeted modification list to search for histone PTMs following the format in the example file. The common PTMs to include are: Lysine acetylation, lysine mono-methylation, lysine di-methylation, lysine tri-methylation, serine/threonine/tyrosine phosphorylation, protein N-terminal acetylation, methionine/cysteine oxidation. For sorghum, protein N-terminal mono-methylation, di-methylation, and trimethylation should be added.

NOTE: Informed-Proteomics only looks for PTMs specified in the list. If unspecified PTMs are present, the proteoform may not be identified, or may be misidentified to other proteoforms. However, the PTM list should be kept as short as possible to minimize the search time.

4.2.11 Execute the MSPathFinder module to identify proteoforms using the files from step 4.2.8, the focused FASTA from step 4.2.9, and the modification list from step 4.2.10. The default parameters can be used.

4.2.12 The results can be visualized in LcMsSpectator by loading all the result files.

NOTE: Other bioinformatics tools are available for processing and visualizing top-down data, each with its own strengths[24–28]. Sorghum and many other organisms have limited known information regarding histone PTMs in the database. Use TopPIC first to identify mass shifts from PTMs. This analysis can readily discover both known and unknown PTMs. Then, the detected PTMs can be searched in a targeted fashion either by specifying a PTM list in TopPIC, or with other complementary tools.

**REPRESENTATIVE RESULTS:**

Following the protocol, the histones can be extracted and identified using the LC-MS analysis. The raw data and processed results are available at MassIVE (https://massive.ucsd.edu/) via accession: MSV000085770. Based on the TopPIC results from the representative sample (available also from MassIVE), we identified 303 histone proteoforms (106 H2A, 72 H2B, 103 H3, and 22 H4 proteoforms). Co-purified ribosomal proteoforms have also been detected, typically eluting early in the LC. They usually consist of ~20% of the identified proteoforms, but do not overlap with the histone proteoforms eluting in the later stage of the LC gradient. The results can be easily visualized with the latest TopPIC or Informed-Proteomics packages. For demonstration, we will focus on the data visualization using the Informed-Proteomics package, which can be used to directly load raw MS files and manually examine proteoform identifications. Please note that both software packages use different algorithms and parameters. The reported numbers of proteoforms will not be identical. We recommend reporting the proteoform counts from TopPIC because it is more conservative, and it does consider unknown PTMs. Informed-Proteomics package has integrated data processing and visualization for easy manual validation. For organisms with well-annotated PTMs, we recommend ProSightPC[24] for best site localization. Combining the results using multiple tools can increase the number of and the confidence of proteoform identifications.

After processing the data with Informed-Proteomics, the LC-MS feature map can be visualized in LcMsSpectator, which displays the deconvoluted protein masses against the LC retention time. By clicking on the identified proteoforms in the software, the associated feature will be highlighted with a small green rectangle in the feature map. Major histone proteins should be seen in specific regions of the map, which indicates the success of the experiment. **Figure 1a** shows a representative LC-MS feature map of intact histones. Full-length histone proteoforms are highlighted in the dashed boxes. Most proteoforms detected can be confidently identified using $MS^2$ data.

**Figure 1b** shows the zoom in of the region with H2A and H2B proteoforms. Most of them have N-terminal modifications of 42 Da. This nominal mass corresponds to either trimethylation (42.05 Da) or acetylation (42.01 Da), which are commonly seen for histones. Their accurate masses differ only by 0.04 Da and are difficult to differentiate at the intact protein level (~2 ppm). In high resolution $MS^2$ spectra, the PTMs can be easily differentiated and confirmed because of the lower mass of the fragments[29]. In addition, H2A and H2B histones have multiple homologs with very similar sequences as noted by the different UniProt accession numbers in **Figure 1b**. Again, high resolution LC-MS analysis can readily identify and differentiate them. Two types of H2As were identified for sorghum histones. The 16 kDa H2A histones in **Figure 1b** have extended terminal tails in the non-conserved regions of histones. Another group of H2A histones without the extended tails (14 kDa) can be seen in **Figure 1c**.

For H4 histones, N-terminal acetylation was identified as the major PTM. Additional lysine acetylations and methionine oxidations can be also observed simply by examining the mass differences of the features in **Figure 1d**. We also observed an unknown modification of 112.9 Da in addition to the N-terminal acetylation (the feature above “3Ac” in **Figure 1d**). This is likely some unknown adducts from the reagent used in the preparation. We have previously detected sulfate

ion adducts on H4, which may be attributed to residual salts combined with high basicity of histone proteins. For H3, two protein sequences were identified H3.3 and H3.2 (**Figure 1e**). Although these two protein sequences differ at only 4 residues (32, 42, 88, and 91), they can still be easily distinguished in LC-MS based on the separation in both dimensions, mass, and retention time. H3 proteins are heavily modified by varying degrees of methylation and acetylation. The high degree of modification can be easily visualized by the dense, parallel lines in the feature map, which are 14 Da apart. However, three methylation groups (14*3 Da) have the equal nominal mass to one acetylation (42 Da). Because these PTMs cannot be easily resolved at intact protein level, they are referred to as "methyl equivalents" (i.e., multiples of 14 Da; one acetylation equals three methyl equivalents). In **Figure 1e**, H3 proteoforms are labeled in the form of methyl equivalents based on their intact mass. Due to limited resolution of the RPLC separation, many different H3 proteoforms are likely co-eluting and fragmented in the same spectrum. The method presented here will only identify the most abundant combinations of methylation and acetylation as illustrated in **Figure 2**. For more comprehensive characterization of H3, more targeted analysis is still required[30,31].

[Place **Figure 1** here]

A representative example of proteoform identification is shown in **Figure 2** using MSPathfinder and visualized in LcMsSpectator. The fragmentation spectrum in **Figure 2a** was generated using ETD, which yields c and z type ions along the protein backbone. HCD of the same precursor can be used to validate the identification, but HCD generally provides limited sequence coverage[20]. The precursor ions in the previous and next MS1 spectra are shown in **Figure 2b,c**, with their matched isotope peaks highlighted in purple. The sequence coverage map in **Figure 2d** can help localize any possible PTMs. A high-confidence identification should have most of the fragments matched, precursor ion matched, and good sequence coverage to help localize PTMs. In this example, an H3.2 proteoform was identified with two PTMs—di-methylation on K9 and methylation on K27. Following the same method, other proteoforms with different PTMs and terminal truncations can be manually validated.

[Place **Figure 2** here]

Quantitative comparison of the detected histone proteoforms can reveal potential epigenetic markers. We have applied this protocol previously to 48 sorghum samples collected from the field ("additional inhibitors" were not used in this study)[29]. Two different genotypes of sorghum were compared in response to pre-flowering or post-flowering droughts. By comparing the relative abundance of the proteoforms, we discovered some interesting changes of truncated histone proteoforms that are specific to sample conditions as shown in **Figure 3**. C-terminal truncation of H4 was observed only in weeks 3 and 9 for some of the samples (**Figure 3a,b**). For H3.2, N-terminal truncated proteoforms were generally more abundant in week 10 (**Figure 3c,d**). In contrast, C-terminal truncated H3.2 tend to be seen in earlier time points (**Figure 3c**). More importantly, both the genotypes did not respond in the exact same way. The H4 C-terminal truncated proteoforms were significantly more abundant in BTx642 than in RTx430 (**Figure 3b**).

Such data reveals potential epigenetic markers of plant development and stress tolerance that can be further tested with other techniques.

[Place **Figure 3** here]

**FIGURE AND TABLE LEGENDS:**

**Table 1: Composition for extraction buffers (EBs).**

**Table 2: Composition for the nuclei lysis buffer (NLB).**

**Figure 1: LC-MS feature map on intact histones extracted from sorghum leaves.** The figure shows LC retention time (in minutes) vs. the molecular mass for all detected proteoforms. The log abundance is shown by the color scale next to the top map (log 10 abundance). (**a**) The major histone peaks are labeled by the dashed boxes. Most of the features outside the boxes are truncated histones and ribosomal proteins. Zoom-in views for each group of histones: (**b**) H2B and 16 kDa H2A, (**c**) H3, (**d**) 14 kDa H2A, and (**e**) H3. The UniProt accession numbers are noted alongside each feature, followed by detected PTMs. "Ac", "me", "+O" indicate acetylation, methylation, and oxidation, respectively. In (**b**), two truncated H2A C5YZA9 proteoforms are labeled, which had one or two C-terminal alanine clipped (shown as -A*, and -AA*).

**Figure 2: Representative example of an identified histone H3.2 proteoform**. H3.2 preteoform with its (**a**) ETD spectrum, (**b**) precursor ion in the previous MS1 spectrum, (**c**) precursor ion in the next MS1 spectrum, and (**d**) sequence coverage map. The c ions from the N-terminus are labeled in cyan, and the z ions from the C-terminus are in pink. Two PTMs were identified and highlighted in yellow in (**d**) with their mass shifts annotated.

**Figure 3: Quantitative comparison of histone proteoforms.** (**a**) Heatmap of histone H4 proteoforms across different samples. For each proteoform, the abundance extracted from top-down MS data was normalized to the sum of all identified H4 proteoforms in each analysis, yielding the "relative abundance". The values were then scaled to the maximum of each row to better show the changes in low abundance proteoforms. The scaled relative abundance is denoted in the color key at the bottom of the heatmap. Growth conditions are noted on the horizontal axis (Pre: pre-flowering drought, Post: post-flowering drought). Three replicates are grouped together and are separated by black vertical stripes from other conditions. For samples labeled with asterisks, only technical replicates were acquired. Proteoforms are represented on the vertical axis, in the format "starting residue – ending residue: mass; putative modification". (**b**) Relative abundance plot of the truncated H4 proteoforms 2–99 (proteoforms highlighted in bold in (**a**) are summed) at different conditions. The key to the symbols is shown in the legend in the top-right corner. Filled dots in the middle of the error bars are the average values. (**c**) Heatmap of H3.2 proteoforms and (**d**) abundance plot for all identified N-terminal truncated H3.2 are shown in the same format as those for H4. Proteoforms smaller than 8 kDa in (**C**) were omitted for simplicity. The N-terminal and C-terminal truncated H3.2 proteoforms showed

different responses across the growth conditions. Reprinted with permission from ELSEVIER from ref.[29].

**DISCUSSION:**

The presented protocol describes how to extract histones from sorghum leaf (or more generally plant leaf) samples. The average histone yield is expected to be 2–20 µg per 4–5 g sorghum leaf material. The materials are sufficiently pure for the downstream histone analysis by LC-MS (mostly histones with ~20% ribosomal protein contamination). Lower yield may be obtained due to sample variations, or potential mishandling/failures throughout the protocol. Maintaining the integrity of the nuclei before the nuclei lysis step is critical; therefore, aggressive vortexing and pipetting should be avoided before adding NLB. In addition, loss of nuclei may occur when removing the supernatants from the pellets. Care must be taken to not disrupt the pellets when pipetting. The Triton X-100 concentration of 1% was optimized to selectively lyse the non-targeted organelles but not the nuclei (step 3.2). Optimal detergent concentration for other tissue or organisms may be different and need to be experimentally determined. Color change of the supernatant during the filtration process could indicate potential issues such as inefficient release of chloroplast or insufficient grinding of leaf. If possible, use a microscope to check for lysis of chloroplasts and retention of intact nuclei after each step to further optimize the protocol (especially if modifying the protocol for other tissues or plants). This protocol has only been tested with sorghum leaf tissue. It does not work for sorghum root tissue likely due to interference from soil. Application to other plant leaf tissues has not been tested and application to different plants may need additional optimization. For adapting the nuclei isolation protocol for ChIP-seq applications, an additional sucrose gradient density separation after step 3.3.4 (before using NLB) is advised to reduce cytoplasmic contamination. Because of the extensive clean-up steps, small amounts of residual non-nuclei materials are not expected to cause significant interference for histone analysis in LC-MS and can be left with the pellet.

Several initial trials failed when using commercial tablets of phosphatase inhibitors (e.g., PhosSTOP). The supernatant in step 3.1.6 appeared to be intense green when the tablets were used in the extraction buffer. The final extract showed low number of identified histones. We suspect the proprietary ingredients in the tablets may have caused nuclei lysis before step 3.4, reducing the overall histone yield. Another possible reason for failure is the incompatibility of the ingredients in the histone purification step with the ion exchange resin (step 3.4). We have used this protocol to consistently extract high purity histones for subsequent LC-MS over 150 samples. On average, we were able to obtain higher yield without using the “additional inhibitors” (unpublished data). Therefore, it is advised to cautiously test new inhibitors when modifying or adapting this protocol for other purposes. If phosphorylation is not of interest, the phosphatase inhibitors can be omitted in the extraction buffers.

The steps in 3.4 can take 3–4 h or more. It is recommended to break the protocol in 2 days—freeze the nuclei pellet from step 3.3 and perform the purification on day 2 (or later). The freeze-thaw cycle may partially help the nuclei lysis. The MWCO filter steps (3.4.7) can be very time consuming but can be easily scaled up by preparing multiple samples in parallel. Do not add the protease inhibitor tablets in step 3.4. Many commercial tablets contain polymers (e.g.,

polyethene glycol) as fillers, which will interfere with LC-MS analysis. At this step, the most other proteins should have been removed or denatured, so enzyme inhibitors are not critical. However, it is still necessary to keep the samples at 4 °C or frozen to minimize degradation.

Following this protocol, histones can be successfully extracted from sorghum leaves. Histone PTMs can be characterized with LC-MS. The method can be potentially applied to large scale studies for comparing histone PTMs between different biological samples (e.g., different genotypes, plants grown under different conditions, etc.) as shown by the example data in **Figure 3**. However, data processing still requires extensive manual analysis for confidently assigning proteoforms, especially for unexpected (or novel) PTMs. New developments in bioinformatics tools are anticipated to automate the workflow and significantly increase the throughput for large-scale studies. Another limitation is that the top-down MS method, currently, cannot easily differentiate many proteoforms of hyper-modified H3 (e.g., multiple sites of mono/di/tri-metlylation and acetylation). The single dimension reversed-phase LC cannot fully separate the different H3 proteoforms. Therefore, the MS2 spectra of H3 will typically contain fragments from multiple proteoforms and cannot be easily and confidently deconvoluted. Combining top-down with bottom-up or middle-down methods[30,32,33] can be especially beneficial for characterization of histone H3. Alternatively, multi-dimensional separation can be considered to improve the depth of top-down MS[34–36].

Histone PTM profiling by LC-MS enables discovery of novel epigenetic markers for designing chromatin modifiers and improve the resilience of plants to severe environmental conditions. A pilot study using sorghum from two cultivars and grown under drought conditions in the field indicated that selective histone terminal clipping in leaf may be related to drought acclimation and plant development[29]. The identified histone markers may serve as targets by complementary techniques such as ChIP-seq. Comprehensive understanding of epigenetic factors gained from these complementary techniques would be indispensable for engineering innovative solutions to crops in response to environmental changes.

**ACKNOWLEDGMENTS:**
We thank Ronald Moore and Thomas Fillmore for helping with mass spectrometry experiments, and Matthew Monroe for data deposition. This research was funded by grants from US Department of Energy (DOE) Biological and Environmental Research through the Epigenetic Control of Drought Response in Sorghum (EPICON) project under award number DE-SC0014081, from the US Department of Agriculture (USDA; CRIS 2030-21430-008-00D), and through the Joint BioEnergy Institute (JBEI), a facility sponsored by DOE (Contract DE-AC02-05CH11231) between Lawrence Berkeley National Laboratory and DOE. The research was performed using Environmental Molecular Sciences Laboratory (EMSL) (grid.436923.9), a DOE Office of Science User Facility sponsored by the Office of Biological and Environmental Research.

**DISCLOSURES:**
None.

**REFERENCES:**

1. Farooq, M., Wahid, A., Kobayashi, N., Fujita, D., Basra, S. M. A. Plant drought stress: Effects, mechanisms and management. *Agronomy for Sustainable Development*. 153–188 (2009).
2. Dai, A. Drought under global warming: a review. *Wiley Interdisciplinary Reviews: Climate Change*. **2** (1), 45–65 (2011).
3. Rooney, W. L., Blumenthal, J., Bean, B., Mullet, J. E. Designing sorghum as a dedicated bioenergy feedstock. *Biofuels, Bioproducts and Biorefining*. **1** (2), 147–157 (2007).
4. Mullet, J. E., Klein, R. R., Klein, P. E. Sorghum bicolor - an important species for comparative grass genomics and a source of beneficial genes for agriculture. *Current Opinion in Plant Biology*. **5** (2), 118–121 (2002).
5. Xu, L. et al. Drought delays development of the sorghum root microbiome and enriches for monoderm bacteria. *Proceedings of the National Academy of Sciences of the United States of America*. **115** (18), E4284–E4293 (2018).
6. Gao, C. et al. Strong succession in arbuscular mycorrhizal fungal communities. *ISME Journal*. **13** (1), 214–226 (2019).
7. Gao, C. et al. Fungal community assembly in drought-stressed sorghum shows stochasticity, selection, and universal ecological dynamics. *Nature Communications*. **11** (1) (2020).
8. Bannister, A. J., Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Research*. **21** (3), 381–395 (2011).
9. Yuan, L., Liu, X., Luo, M., Yang, S., Wu, K. Involvement of histone modifications in plant abiotic stress responses. *Journal of Integrative Plant Biology*. **55** (10), 892–901 (2013).
10. Park, P. J. ChIP-seq: advantages and challenges of a maturing technology. *Nature Reviews. Genetics*. **10** (10), 669–680 (2009).
11. Huang, H., Lin, S., Garcia, B. A., Zhao, Y. Quantitative proteomic analysis of histone modifications. *Chemical Reviews*. **115** (6), 2376–2418 (2015).
12. Moradian, A., Kalli, A., Sweredoski, M. J., Hess, S. The top-down, middle-down, and bottom-up mass spectrometry approaches for characterization of histone variants and their post-translational modifications. *Proteomics*. **14** (4–5), 489–497 (2014).
13. Sidoli, S., Garcia, B. A. Characterization of individual histone posttranslational modifications and their combinatorial patterns by mass spectrometry-based proteomics strategies. *Methods in Molecular Biology (Clifton, N. J.)*. **1528**, 121–148 (2017).
14. Maile, T. M. et al. Mass spectrometric quantification of histone post-translational modifications by a hybrid chemical labeling method. *Molecular & Cellular Proteomics*. **14** (4), 1148–1158 (2015).
15. Dang, X. et al. The first pilot project of the consortium for top-down proteomics: a status report. *Proteomics*. **14** (10), 1130–1140 (2014).
16. Schaffer, L. V. et al. Identification and quantification of proteoforms by mass spectrometry. *Proteomics*. **19** (10), 1800361 (2019).
17. Cupp-Sutton, K. A., Wu, S. High-throughput quantitative top-down proteomics. *Molecular Omics* (2020).
18. Varoquaux, N. et al. Transcriptomic analysis of field-droughted sorghum from seedling to maturity reveals biotic and metabolic responses. *Proceedings of the National Academy of Sciences of the United States of America*. **116** (52), 27124 LP–27132 (2019).
19. Gordon, J. A. Use of vanadate as protein-phosphotyrosine phosphatase inhibitor.

*Methods in Enzymology*. **201** (C), 477–482 (1991).
20. Zhou, M. et al. Profiling changes in histone post-translational modifications by top-down mass spectrometry. *Methods in Molecular Biology (Clifton, N. J.)*. **1507**, 153–168 (2017).
21. Chambers, M. C. et al. A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology*. **30** (10), 918–920 (2012).
22. Kou, Q., Xun, L., Liu, X. TopPIC: a software tool for top-down mass spectrometry-based proteoform identification and characterization. *Bioinformatics (Ocford, England)*. **32** (22), btw398 (2016).
23. Park, J. et al. Informed-Proteomics: open-source software package for top-down proteomics. *Nature Methods*. **14** (9), 909–914 (2017).
24. LeDuc, R. D. et al. The C-Score: a bayesian framework to sharply improve proteoform scoring in high-throughput top down proteomics. *Journal of Proteome Research*. **13** (7), 3231–3240 (2014).
25. Fornelli, L. et al. Advancing top-down analysis of the human proteome using a benchtop quadrupole-orbitrap mass spectrometer. *Journal of Proteome Research*. **16** (2), 609–618 (2017).
26. Sun, R.-X. et al. pTop 1.0: A high-accuracy and high-efficiency search engine for intact protein identification. *Analytical Chemistry*. **88** (6), 3082–3090 (2016).
27. Xiao, K., Yu, F., Tian, Z. Top-down protein identification using isotopic envelope fingerprinting. *Journal of Proteomics*. **152**, 41–47 (2017).
28. Cai, W. et al. MASH Suite Pro: A comprehensive software tool for top-down proteomics. *Molecular & Cellular Proteomics: MCP*. **15** (2), 703–714 (2016).
29. Zhou, M. et al. Top-down mass spectrometry of histone modifications in sorghum reveals potential epigenetic markers for drought acclimation. *Methods (San Diego, Calif.)*. (2019).
30. Garcia, B. A., Pesavento, J. J., Mizzen, C. A., Kelleher, N. L. Pervasive combinatorial modification of histone H3 in human cells. *Nature Methods*. **4** (6), 487–489 (2007).
31. Zheng, Y. et al. Unabridged analysis of human histone H3 by differential top-down mass spectrometry reveals hypermethylated proteoforms from MMSET/NSD2 overexpression. *Molecular & Cellular Proteomics: MCP*. **15** (3), 776–790 (2016).
32. Garcia, B. A. et al. Chemical derivatization of histones for facilitated analysis by mass spectrometry. *Nature Protocols*. **2** (4), 933–938 (2007).
33. Holt, M. V., Wang, T., Young, N. L. One-pot quantitative top- and middle-down analysis of GluC-digested histone H4. *Journal of the American Society for Mass Spectrometry*. **30** (12), 2514–2525 (2019).
34. Tian, Z. et al. Enhanced top-down characterization of histone post-translational modifications. *Genome Biology*. **13** (10), R86 (2012).
35. Wang, Z., Ma, H., Smith, K., Wu, S. Two-dimensional separation using high-pH and low-pH reversed phase liquid chromatography for top-down proteomics. *International Journal of Mass Spectrometry*. **427**, 43–51 (2018).
36. Gargano, A. F. G. et al. Increasing the separation capacity of intact histone proteoforms chromatography coupling online weak cation exchange-HILIC to reversed phase LC UVPD-HRMS. *Journal of Proteome Research*. **17** (11), 3791–3800 (2018).

Figure 1



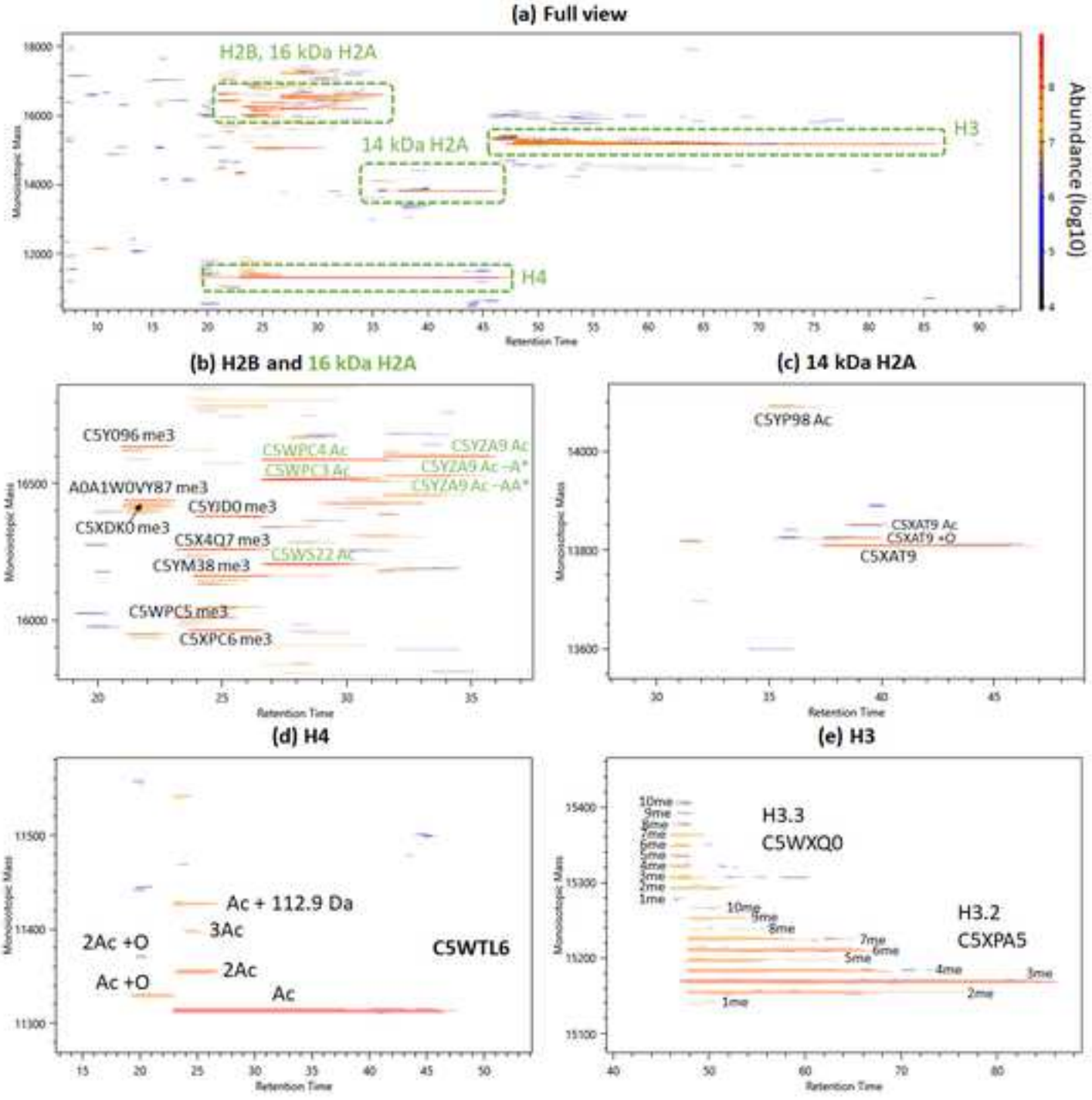


Figure 2



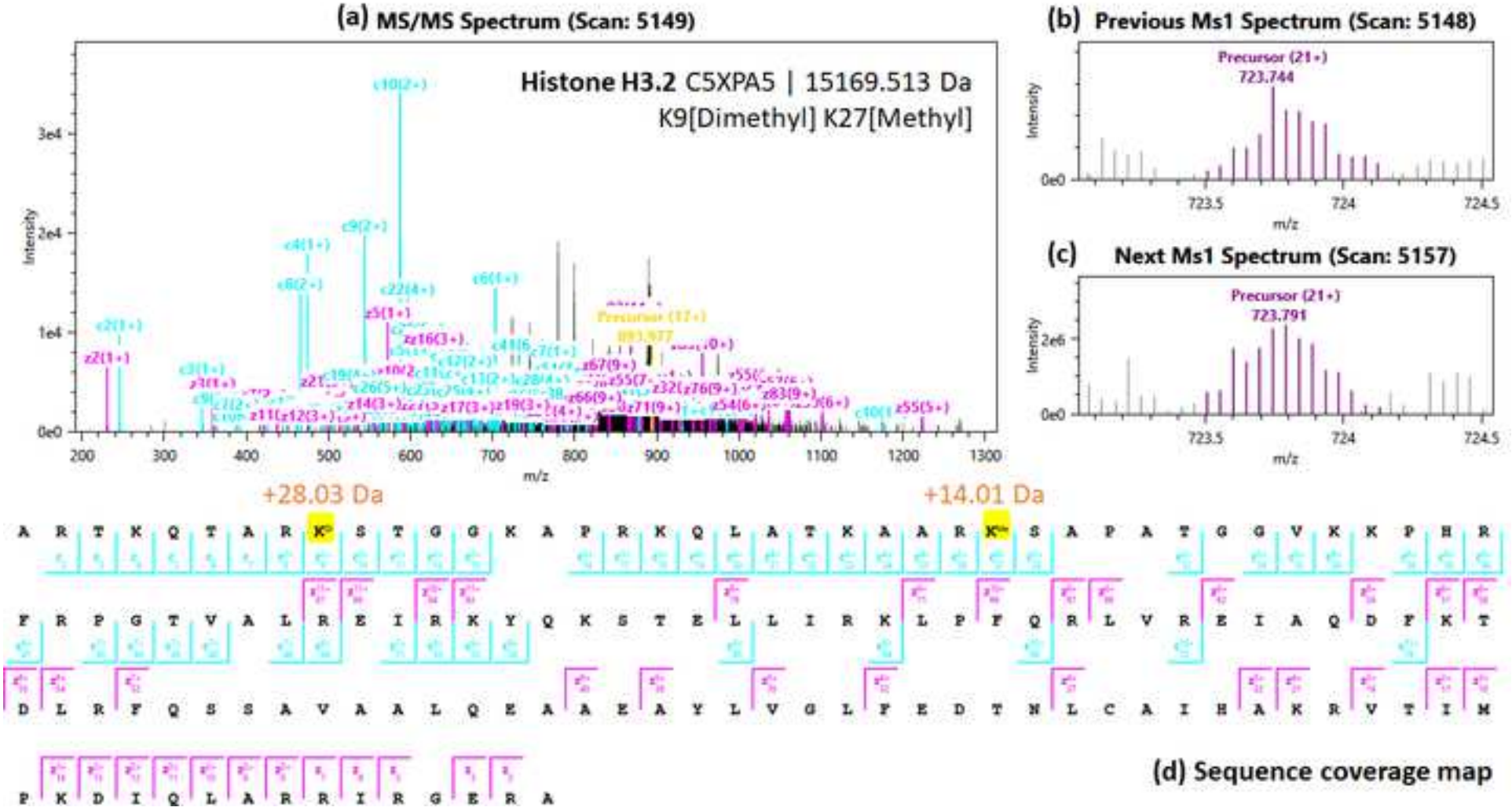


Figure 3



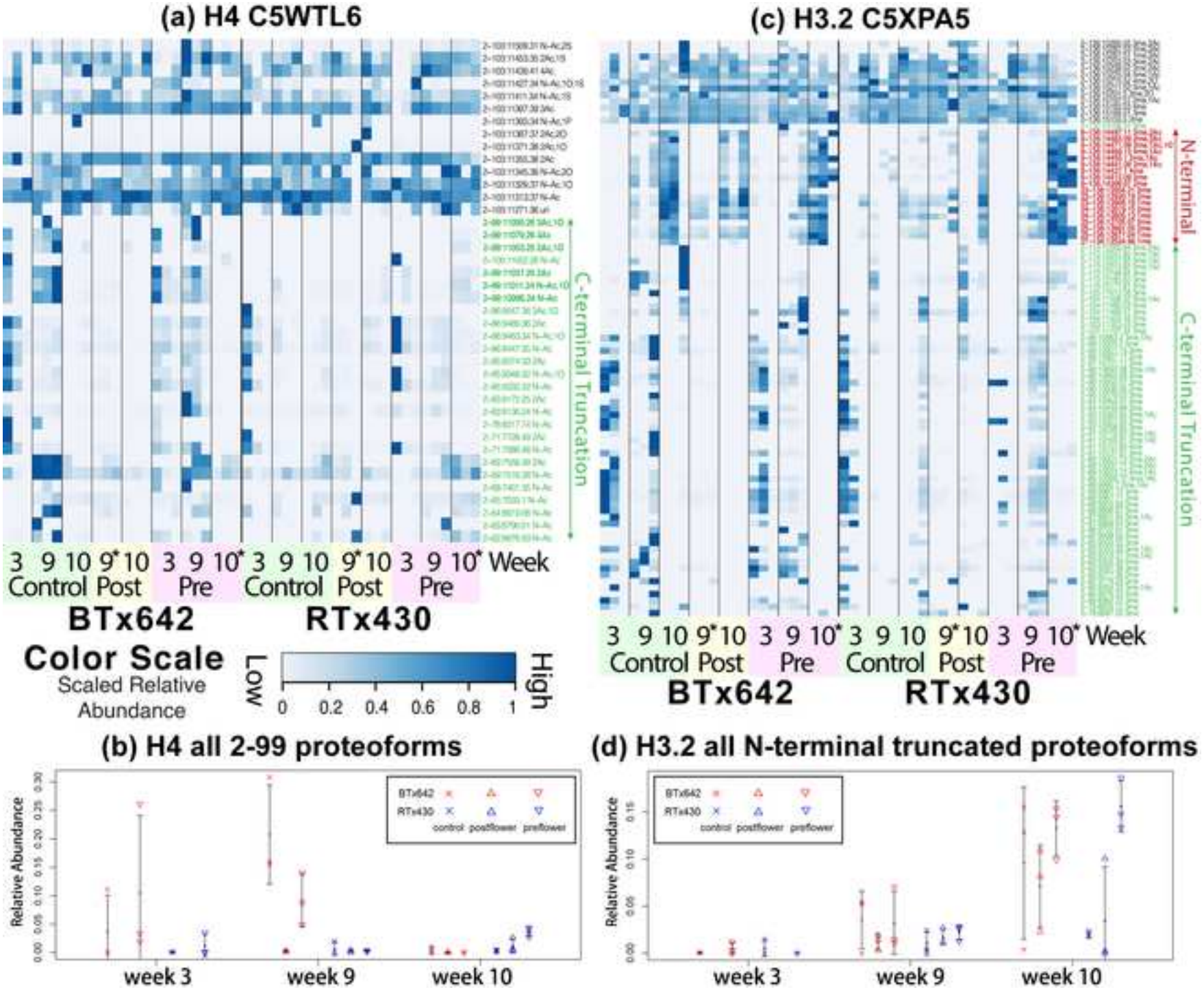


**Table 1. Composition for extraction buffers (EBs).**

| **Reagents** | Stock concentration | **EB1** | **EB2A** | **EB2B** |
|---|---|---|---|---|
| | | Volume (mL) | Volume (mL) | Volume (mL) |
| Sucrose | 2.5M | 4.4 | 1.25 | 0.5 |
| Tris HCl pH8 | 1M | 0.25 | 0.125 | 0.05 |
| DTT | 1M | 0.125 | 0.0625 | 0.025 |
| $H_2O$ | | 20.225 | 9.6875 | 4.375 |
| protease inhibitor (PI) tablet | | 0.5 pill | 0.5 pill | 0.5 pill |
| Additional inhibitors (Optional) | 33mM | 0.25 | 0.125 | 0.05 |
| $MgCl_2$ | 1M | | 0.125 | 0.05 |
| Triton X100 | 10% | | 1.25 | |
| **Overall Volume** | | 25 mL | 12.5 mL | 5 mL |

**Table 2. Composition for the nuclei lysis buffer (NLB).**

| **NLB** | Stock concentration | Volume (mL) |
|---|---|---|
| NaCl | 5M | 0.4 |
| Tris HCl pH8 | 1M | 0.05 |
| Triton X100 | 10% | 0.5 |
| EDTA | 0.5M | 0.2 |
| $H_2O$ | | 3.85 |
| PI tablets | | 0.5 pill |
| Additional inhibitors (optional) | 33mM | 0.05 |
| **Overall Volume** | | 5 mL |

| Name of Material/Equipment | Company | Catalog Number | Comments/Description |
|---|---|---|---|
| Acetonitrile | Fisher Chemical | A955-4L | |
| Dithiothreitol (DTT) | Sigma | 43815-5G | |
| EDTA, 500mM Solution, pH 8.0 | EMD Millipore Corp | 324504-500mL | |
| Formic Acid | Thermo Scientific | 28905 | |
| Guanidine Hydrochloride | Sigma | G3272-100G | |
| MgCl2 | Sigma | M8266-100G | |
| Potassium phosphate, dibasic | Sigma | P3786-100G | |
| Protease Inhibitor Cocktail, cOmplete tablets | Roche | 5892791001 | |
| Sodium butyrate | Sigma | 303410-5G | Used for histone deacetylase inhibitor |
| Sodium Chloride (NaCl) | Sigma | S1888 | |
| Sodium Fluoride | Sigma | S7020-100G | Used for phosphatase inhibitor |
| Sodium Orthovanadate | Sigma | 450243-10G | Used for phosphatase inhibitor |
| Sucrose | Sigma | S7903-5KG | |
| Tris-HCl | Fisher Scientific | BP153-500 g | |
| Triton X-100 | Sigma | T9284-100ML | |
| Weak cation exchange resin, mesh 100-200 analytical (BioRex70) | Bio-Rad | 142-5842 | |
| **Disposables** | | | |
| Chromatography column (Bio-Spin) | BIO-RAD | 732-6008 | |
| Mesh 100 filter cloth | Millipore Sigma | NY1H09000 | This is part of the Sigma kit (catalog # CELLYTPN1 |
| Micropipette tips (P20, P200, P1000) | Sigma | | |
| Tube, 50mL/15mL, Centrifuge, Conical | Genesee Scientific | 28-103 | |
| Tube, Microcentrifuge, 1.5/2 mL | Sigma | | |
| **Equipment** | | | |
| Analytical Balance | Fisher Scientific | 01-912-401 | |

| | | |
|---|---|---|
| Beakers (50mL – 2L) | | |
| Microcentrifuge with cooling | Fisher Scientific | 13-690-006 |
| Micropipettes | | |
| Swinging-bucket centrifuge with cooling | Fisher Scientific | |
| Vortex | Fisher Scientific | 50-728-002 |
| Water bath Sonicator | Fisher Scientific | 15-336-120 |

l) for plant nuclei extraction. Similar filters with the same mesh size can be used.

www.emsl.pnl.gov

Environmental Molecular Sciences Laboratory
902 Battelle Boulevard • P.O. Box 999 • Richland, WA 99352
Tel: (509) 371-6585 ● Fax: (509) 371-6555 ● ljiljana.pasatolic@pnnl.gov

Aug 10, 2020

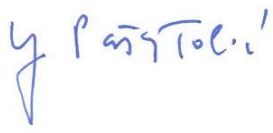
Dear Vineeta,

We are submitting our revised manuscript titled "Isolation of Histones from Sorghum Leaf Tissue for Top-Down Mass Spectrometry Profiling of Potential Epigenetic Markers". We appreciate all the critical comments from reviewers to help improve our manuscript, especially some of the technical comments from reviewer 3. We also reorganized some of the text to improve clarity.

Both reviewer 2 and 3 requested additional analysis to extract biologically significant information from the results. We are glad that they found the technique interesting and potentially useful to address important biological questions. However, the focus of this manuscript is the protocol itself, and not the specific biological question. In addition, the reviewer 2 was questioning the reproducibility of our protocol. To address these concerns, we have now added Figure 3 from our published manuscript describing quantitative comparison between control and drought samples (https://doi.org/10.1016/j.ymeth.2019.10.007). We believe this inclusion better demonstrates that the protocol can generate histones for LC-MS analysis reproducibly in a large-scale study setting. Figure 3 also addresses the reviewer 3 comments regarding quantitative analysis of the data, and we thank both of the reviewers for this improvement of the manuscript.

Detailed responses are attached below in this letter. We believe we have addressed all of the reviewers' concerns and that the manuscript is now ready for publication. We are looking forward to hearing from you soon.

Sincerely,



Ljiljana Paša-Tolić, Ph.D.
Laboratory Fellow and Deputy for Technology
W.R. Wiley Environmental Molecular Sciences Laboratory
Pacific Northwest National Laboratory

## Response to comments

Original comments are shown in gray Calibri. Responses are in black Arial font.

---

Editorial comments:

NOTE: Please read this entire email before making edits to your manuscript. Please include a line-by-line response to each of the editorial and reviewer comments in the form of a letter along with the resubmission.

• Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammatical errors.

• Protocol Detail: Please note that your protocol will be used to generate the script for the video, and must contain everything that you would like shown in the video. Please ensure that all specific details (e.g. button clicks for software actions, numerical values for settings, etc) have been added to your protocol steps. There should be enough detail in each step to supplement the actions seen in the video so that viewers can easily replicate the protocol.

1) 1.1: Mention plant growth conditions and environment. Mention age at plucking.

The protocol is generally agnostic to growth conditions. Details are now added in step 1.

• Protocol Numbering: Add a one-line space between each protocol step.

Added.

• Protocol Highlight: After you have made all of the recommended changes to your protocol (listed above), please re-evaluate the length of your protocol section. Please highlight ~2.5 pages or less of text (which includes headings and spaces) in yellow, to identify which steps should be visualized to tell the most cohesive story of your protocol steps.

1) The highlighting must include all relevant details that are required to perform the step. For example, if step 2.5 is highlighted for filming and the details of how to perform the step are given in steps 2.5.1 and 2.5.2, then the sub-steps where the details are provided must be included in the highlighting.

2) The highlighted steps should form a cohesive narrative, that is, there must be a logical flow from one highlighted step to the next.

3) Please highlight complete sentences (not parts of sentences). Include sub-headings and spaces when calculating the final highlighted length.

4) Notes cannot be filmed and should be excluded from highlighting.

Updated as requested.

• Discussion: JoVE articles are focused on the methods and the protocol, thus the discussion should be similarly focused. Please ensure that the discussion covers the following in detail and in paragraph form (3-6 paragraphs): 1) modifications and troubleshooting, 2) limitations of the technique, 3) significance with respect to existing methods, 4) future applications and 5) critical steps within the protocol.

We believe the discussion already addressed all of these points. We have modified the text to expand on a few points the reviewers asked about.

• Figures: Please remove the embedded figures from the manuscript. Figure legends, however, should remain within the manuscript text, directly below the Representative Results text.

Updated as requested.

• References:

1) Please spell out journal names.

2) Please make sure that your references comply with JoVE instructions for authors. Citation formatting should appear as follows: (For less than 6 authors, list all authors. For more than 6 authors, list only the first author then et al.): [Lastname, F.I., LastName, F.I., LastName, F.I. Article Title. Source. Volume (Issue), FirstPage – LastPage, (YEAR).]

Updated as requested.

• Commercial Language: JoVE is unable to publish manuscripts containing commercial sounding language, including trademark or registered trademark symbols (TM/R) and the mention of company brand names before an instrument or reagent. Examples of commercial sounding language in your manuscript are BioRex70, Amicon, Phenomenex , Phenomenex Aeris,

1) Please use MS Word's find function (Ctrl+F), to locate and replace all commercial sounding language in your manuscript with generic names that are not company-specific. All commercial products should be sufficiently referenced in the table of materials/reagents. You may use the generic term followed by "(see table of materials)" to draw the readers' attention to specific commercial names.

All commercial sounding language has been removed from the main text.

• Table of Materials: Sort the list alphabetically.

Updated as requested.

• If your figures and tables are original and not published previously or you have already obtained figure permissions, please ignore this comment. If you are re-using figures from a previous publication, you must obtain explicit permission to re-use the figure from the previous publisher (this can be in the form

of a letter from an editor or a link to the editorial policies that allows you to re-publish the figure). Please upload the text of the re-print permission (may be copied and pasted from an email/website) as a Word document to the Editorial Manager site in the "Supplemental files (as requested by JoVE)" section. Please also cite the figure appropriately in the figure legend, i.e. "This figure has been modified from [citation]."

All figures in the submitted manuscript were new, generated from previously unpublished data. In response to the reviewer comments, we added a published figure as Figure 3 in the revised manuscript and uploaded the re-print permission.

---

Reviewers' comments:

Reviewer #1:

The authors describe a protocol to isolate the nuclei and purify histones from sorghum leaf tissue and analyze the intact histones by top-down LC-MS. The manuscript is well written.

Reviewer #2:

Manuscript Summary:

This paper reports the extraction of histone from sorghum leaves and its application in the post-translational modification of histone. The manuscript does report some data that may be considered important and interesting, but it needs substantial improvement before it can be published. First of all, English is often incorrect. Many sentences have logical errors that make the manuscript difficult to understand. Professional language editors must be consulted before submitting revisions. The following suggestions are put forward:

We have done our best to revise the text and we hope it now addressed this concern.

Major Concerns:

1st major point

The experimental materials must be clear, such as sorghum varieties, planting environment and sampling time. Due to different periods and different states of sampling, histone extraction and analysis will be affected to some extent. Therefore, it is necessary to make clear all kinds of information about sorghum materials.

We agree that sample condition can affect the histone extraction efficiency. We added the information in the note to step 1.1 to describe the condition of the sample(s) we used to generate the data shown:

“NOTE: The sorghum plants were grown in soil in the field in Parlier, CA. Leaf tissue was collected by tearing off the third and fourth fully emerged leaf from the primary tiller. More details of field condition, sample growth, and collection can be found in published report.[18] The example data shown in Figure 1 and Figure 2 were from sorghum leaf collected at week 2 after planting. Although variation of yield is expected, we believe this protocol is generally agnostic to specific sample conditions. We have successfully used the same protocol for sorghum plant leaf tissue collected 2, 3, 5, 8, 9, and 10 weeks after planting.”

2nd major point

From the perspective of operational steps, it is not seen that this experiment has been repeated for three times to ensure the accuracy and scientific nature of the experiment.

We have used this protocol to extract histones from ~50 samples in previously published work (reference 29) and from over 100 field samples in a follow-up large-scale study (manuscript in preparation). Current manuscript is focused on the detailed protocol for histone extraction. We revised the last part of the abstract to emphasize that this protocol has been used for a larger study before:

“We have applied this protocol previously to profile histone PTMs from sorghum leaf tissue collected from a large-scale field study, aimed at identifying epigenetic markers of drought resistance. The protocol could potentially be adapted and optimized for chromatin immunoprecipitation – sequencing (ChIP-seq), or for studying histone PTMs in similar plants.”

3rd major point

It is not rigorous to determine the integrity of nucleus only by observing the color change of chloroplast during filtration. We should use some qualitative and quantitative data to ensure the accuracy of the experiment.

This protocol originated from another protocol initially developed for the ChIP-seq workflow. Initial optimization was performed with more rigorous testing. The color change is mentioned in the notes to help troubleshoot any potential issues. With the conditions already sorted out, we don’t feel it is necessary to check the integrity of the nuclei in every step. We agree it would be valuable to check the integrity of the nuclei at various points in the protocol, especially for the first time user, but this would be inefficient for large number of samples (as often encounter is such studies). We mentioned in the discussion that checking integrity of nucleus is advised, especially when transferring the protocol to other systems. We reworded the sentence to avoid confusion (highlighted in red):

“Optimal detergent concentration for other tissue or organisms may be different and need to be experimentally determined. Color change of the supernatant during the filtration process could indicate potential issues such as inefficient release of chloroplast or insufficient grinding of leaf. If possible, use a microscope (e.g. with methylene blue stain) to check for lysis of chloroplasts and retention of intact nuclei after each step to further optimize the protocol (especially if modifying the protocol for other tissues or plants).”

Additional details are added in NOTES for 3.1.4 and 3.2.2:

“NOTE: Both the filtrate and the filtered debris should be green. If tracking using a microscope, you should be able to see intact nuclei and intact chloroplasts in the filtrate at this point. The majority of large debris should be absent/depleted. Mix dyes such as methylene blue with sample. Nuclei are easily observable as ~3-5 µm diameter dark blue/aquamarine spheres when visualized using a using a 20X, 40X and/or 100X objective. Relative to nuclei, chloroplasts are similar in size, but greenish in color and often more oval in shape. Vacuoles are also similar to nuclei in size and shape, but they will not readily take up the Methylene blue dye.”

“NOTE: The detergent concentration needs to be optimized to preferentially lyse intact cells and chloroplasts but not nuclei. The amount required can vary among organisms. It is recommended to check for lysis of chloroplasts and retention of intact nuclei under microscope.”

4th major point

According to the author's operation steps, histone can be extracted successfully, but the output of histone is low. It is suggested to change reagents or carry out experiments at low temperature to reduce the degradation and unnecessary loss of histone.

We agree the protocol may be further optimized to improve yield. However, this is out of the scope of our current study. The current protocol on average yields sufficient material for the downstream mass spectrometry analysis as outlined in the manuscript. We revised discussion to note there is still room for improvement, especially regarding the enzyme inhibitors (edited text in red):

“We have used this protocol to consistently extract high purity histones for subsequent LC-MS analyses from over 150 samples. On average we were able to obtain higher yield without the use of “additional inhibitors” (unpublished data). Therefore, it is advised to cautiously test the new inhibitors when modifying or adapting this protocol for other purposes. If phosphorylation is not of interest, the phosphatase inhibitors can be omitted in the extraction buffers.”

5th major point

The key word "drought" given by the author is not involved in the experiment, but the prospect of histone modification of sorghum under drought condition is put forward. It would be more convincing if the author could carry out experiments on Sorghum leaves under both drought and normal conditions at the same time.

We have previously performed and published the comparison study (reference 29). To avoid confusion, we now added Figure 3 and a paragraph in the result section citing the published data for quantitative comparison of 48 samples. This manuscript focuses on the detailed protocol for histone extraction. We encourage the reader to check our published manuscript for more details.

6th major point

It is suggested that in the introduction part, the whole content written by the author should be divided into clear paragraphs according to a certain logical order, so that the readers can read more clearly.

We thank reviewer for this comment and have now divided the introduction into smaller paragraphs.

Minor Concerns:

Line 106, line 108, line 123, etc.

Loss of ℃ in 4 ℃.

Corrected.

Reviewer #3:

Manuscript Summary:

The manuscript of Zhou et al. presents a protocol for purification of histone proteoforms for top-down proteomics. The topic is definitely of high interest as there is a necessity to integrate data from various approaches to extend knowledge in the field of plant histones. Generally, the methods presented seem to be appropriate and feasible. However, the authors should specify / clarify / correct several points.

Major Concerns:

1) The authors should comment on the purity of isolated histones and provide the graph showing the proportion of histones within all identified proteins.

We added a comment in the result section. Based on our previous experience we usually see ribosomal proteins as contamination, but they can be separated by LC and generally do not interfere with histone analysis.

“Following the protocol, the histones can be extracted and identified using the LC-MS analysis. Based on the TopPIC results from the representative sample, we identified 303 histone proteoforms (106 H2A, 72 H2B, 103 H3, and 22 H4 proteoforms). Co-purified ribosomal proteoforms have also been detected, typically eluting early in the LC. They usually represent

~20% of the identified proteoforms, but do not overlap with the histone proteoforms eluting in the later stage of the LC gradient."

2) Figure 1: The authors should describe the proteoforms in the figure properly, or explain in the legend or in the text in more details what's behind. It should be mentioned that acetylations in Figure 1d represent both N-terminal and lysine acetylation. Similarly, methylations presented in Figure 1e represent me1, me2, and me3. No acetylation is indicated either in Figure 1e or 2, however, identification of acetylated forms of H3 is mentioned in the text. This should be clarified.

We thank the reviewer for this comment and have revised the discussion to clarify the labeling as follows:

"However, three methylation groups (14*3 Da) have the equal nominal mass to one acetylation (42 Da). Because these PTMs cannot be easily resolved at intact protein level, they are referred to as "methyl equivalents" (i.e. multiples of 14 Da; one acetylation equals three methyl equivalents). In Figure 1e, H3 proteoforms are labeled in the form of methyl equivalents based on their intact mass."

3) The authors admit that many proteoforms cannot be identified due to co-elution (especially proteoforms of H3 - lines 425-432; 509-516). On the other hand, they could identify altogether 303 proteoforms which is extremely high number. With this respect it would be useful to add the list of all identified proteoforms into supplementary data. This would be really interesting information not only from analytical but also from biological point of view.

Many of the proteoforms are truncated histones. At this point, we don't know the exact origin of these truncated proteoforms. Some of them may be biologically relevant, but some may be result of degradation during the sample processing steps. The number of identifications is directly taken from TopPIC output without extensive manual examination and filtering. We now included these results in the deposited data.

4) The authors should comment on quantification issue. The number of unambiguously identified proteoforms which could have been reliably quantified should be added in the text. In addition, those proteoforms together with co-eluting proteoforms should be marked in the list in supplementary data. The authors should discuss if it's feasible to estimate the quantity of individual sequential variants and calculate their ratios (e.g., H3.1 : H3.3 ratio). This would be useful and interesting information for plant biologists.

We have performed intensity-based label free quantitation in a previous study (reference 29) using the same protocol. We were able to observe significant differences in abundance for several proteoforms. The data is now included as Figure 3 to better demonstrate the quantitative aspects of the study, i.e. without having to look up the published study.

We agree with the reviewer that this is an interesting aspect to estimate the histone homologs/variants. However, our current manuscript is focused on the protocol for histone extraction. The example data is to show what a successful extraction looks like. The raw data

including different biological conditions for the previous study are available for the readers who are interested in researching these aspects.

5) The authors should deposit all MS data into PRIDE.

Example dataset is now uploaded at:
https://massive.ucsd.edu/ProteoSAFe/dataset.jsp?task=dedbebf3dc124315b3de47af05607b3c
The original data for the comparative study (reference 29) are also available online as described in the manuscript.

Minor Concerns:

1) Line 95: Change „4 g of cryo-ground leaf" for „5 g of cryo-ground leaf" in order the amount of starting material is consistent with information in line 150.

Thanks for pointing out the inconsistency. For this data set the amount was 4 g.

2) Line 112: The reviewer recommends to change "Sodium Orthovanadate" for "activated Sodium Orthovanadate". The authors should add the information that Sodium Orthovanadate has to be activated - depolymerized before use to enhance the ability to inhibit phosphatases (see Gordon JA. Use of vanadate as protein-phosphotyrosine phosphatase inhibitor. Methods Enzymol. 1991;201:477-82. DOI:10.1016/0076-6879(91)01043-2). Alternatively, the activated liquid form of Sodium Orthovanadate is also available (5086050004 Sigma-Aldrich).

We thank the reviewer for this comment. We only used sodium orthovanadate without intentionally depolymerize it. We have now added the note in step 2.4.

3) Line 113: Sodium butyrate is an HDAC inhibitor, not phosphatase inhibitor. Make a correction in the text and in Table 1 accordingly.

Corrected.

4) Line 121-125: The text in 2.7 and 2.8 is confusing. The reviewer recommends to merge 2.7 with 2.8 as both paragraphs refer to Gdn buffer preparation. Change "Prepare 5% Guanidine buffer pH7" for "Prepare 5% Guanidine in phosphate buffer, pH7". Change "Adjust pH to 7 by checking with pH paper" for "Check pH with pH paper and adjust to 7 using ….". More precisely, Potassium phosphate monobasic and dibasic should be mixed to adjust pH7 while keeping desired ion concentration.

Corrected.

5) Tables 1 and 2:

Change "Tris pH8" for "Tris-HCl pH8".

All inhibitors of enzymatic activity have to be added freshly just before use - add this information into 2.10. Transfer "Make the Nuclei Lysis Buffer (NLB) based on Table 2." from 2.10 to 2.11.

Add "overall volume" into Table 2 - similarly as it is in Table 1.

Corrected.

6) Line 143: Change "steps 1-3" for "steps 3.1-3.3".

Corrected.

7) Line 150: Change "leaf material" for "ground leaf powder".

Corrected.

8) Line 152: Add volume of EB1. Apparently, EB1 is prepared for all following steps - altogether 24 ml are need. Adjust the amount of PI accordingly.

Lines 175, 188: Add volume of EB2A and EB2B, respectively. Adjust the amount of PI accordingly.

Line 198: Add information that PI has to be added into NLB.

Corrected.

ELSEVIER LICENSE
TERMS AND CONDITIONS

Jul 14, 2020

---

This Agreement between Mowei Zhou ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

| | |
|---|---|
| License Number | 4867940150563 |
| License date | Jul 14, 2020 |
| Licensed Content Publisher | Elsevier |
| Licensed Content Publication | Methods |
| Licensed Content Title | Top-down mass spectrometry of histone modifications in sorghum reveals potential epigenetic markers for drought acclimation |
| Licensed Content Author | Mowei Zhou,Neha Malhan,Amir H. Ahkami,Kristin Engbrecht,Gabriel Myers,Jeffery Dahlberg,Joy Hollingsworth,Julie A. Sievert,Robert Hutmacher,Mary Madera,Peggy G. Lemaux,Kim K. Hixson,Christer Jansson,Ljiljana Paša-Tolić |
| Licensed Content Date | Available online 23 October 2019 |
| Licensed Content Volume | n/a |
| Licensed Content Issue | n/a |

| | |
|---|---|
| Licensed Content Pages | 1 |
| Start Page | 0 |
| End Page | 0 |
| Type of Use | reuse in a journal/magazine |
| Requestor type | academic/educational institute |
| Portion | figures/tables/illustrations |
| Number of figures/tables /illustrations | 1 |
| Format | electronic |
| Are you the author of this Elsevier article? | Yes |
| Will you be translating? | No |
| Title of new article | Isolation of Histone from Sorghum Leaf Tissue for Top Down Mass Spectrometry Profiling of Potential Epigenetic Markers |
| Lead author | Mowei Zhou |
| Title of targeted journal | JoVE |
| Publisher | JoVE |

| | |
|---|---|
| Expected publication date | Dec 2020 |
| Portions | part of Figure 4 |
| Requestor Location | Mowei Zhou<br>3335 Innovation Blvd K8-98<br><br>RICHLAND, WA 99354<br>United States<br>Attn: Mowei Zhou |
| Publisher Tax ID | 98-0397604 |
| Total | 0.00 USD |

Terms and Conditions

## INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

## GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of

any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

## LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. **Posting licensed content on any Website**: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.

**Posting licensed content on Electronic reserve**: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors:** the following clauses are applicable in addition to the above:

**Preprints:**

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted

Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

**Accepted Author Manuscripts:** An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
    - via their non-commercial person homepage or blog
    - by updating a preprint in arXiv or RePEc with the accepted manuscript
    - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
    - directly by providing copies to their students or to research collaborators for their personal use
    - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- After the embargo period
    - via non-commercial hosting platforms such as their institutional repository
    - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

**Published journal article (JPA):** A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

**Subscription Articles:** If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.

Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional

private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

**Gold Open Access Articles:** May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's posting policy for further information.

18. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

## Elsevier Open Access Terms and Conditions

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our open access license policy for more information.

### Terms & Conditions applicable to all Open Access articles published with Elsevier:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

### Additional Terms & Conditions applicable to each Creative Commons user license:

**CC BY:** The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the

user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by/4.0.

**CC BY NC SA:** The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at http://creativecommons.org/licenses/by-nc-sa/4.0.

**CC BY NC ND:** The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by-nc-nd/4.0. Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. **Other Conditions**:

v1.9

**Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.**