# Proteomic analysis reveals the differential histone programs between male germline cells and vegetative cells in *Lilium davidii*

**Hao Yang[1,2], Ning Yang[1] and Tai Wang[1],***

[1]*Key Laboratory of Plant Molecular Physiology, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China, and*
[2]*University of Chinese Academy of Sciences, Beijing 100049, China*

**SUMMARY**

**In flowering plants, male germline fate is determined after asymmetric division of the haploid microspore. Daughter cells have distinct fates: the generative cell (GC) undergoes further mitosis to generate sperm cells (SCs), and the vegetative cell (VC) terminally differentiates. However, our understanding of the mechanisms underlying germline development remains limited. Histone variants and modifications define chromatin states, and contribute to establishing and maintaining cell identities by affecting gene expression. Here, we constructed a lily protein database, then extracted and detailed histone entries into a comprehensive lily histone database. We isolated large amounts of nuclei from VCs, GCs and SCs from lily, and profiled histone variants of all five histone families in all three cell types using proteomics approaches. We revealed 92 identities representing 32 histone variants: six for H1, 11 for H2A, eight for H2B, five for H3 and two for H4. Nine variants, including five H1, two H2B, one H3 and one H4 variant, specifically accumulated in GCs and SCs. We also detected H3 modification patterns in the three cell types. GCs and SCs had almost identical histone profiles and similar H3 modification patterns, which were significantly different from those of VCs. Our study also revealed the presence of multiple isoforms, and differential expression patterns between isoforms of a variant. The results suggest that differential histone programs between the germline and companion VCs may be established following the asymmetric division, and are important for identity establishment and differentiation of the male germline as well as the VC.**

Keywords: histone program, histone variant, male germline, sperm cells, generative cells, vegetative cells, *Lilium davidii*.

## INTRODUCTION

In contrast to animals, in which products of male meiosis develop directly into sperm, flowering plants use a specific post-meiotic mechanism to generate sperm cells (SCs) from haploid microspores produced by meiosis of pollen mother cells (Wilson and Yang, 2004). During the post-meiotic process, the microspore undergoes asymmetric mitosis to produce a large vegetative cell (VC) and a diminutive generative cell (GC) enclosed in the VC. The asymmetric division confers differential identities and fates on daughter cells. The VC terminally differentiates, has highly dispersed chromatin and most of the cytoplasm from the microspore, and acts as a companion cell in subsequent development. The GC has highly condensed chromatin and a small quantity of cytoplasm, and undergoes further mitosis to generate two SCs, which are transported

via VC-derived pollen tubes into the embryo sac. There, one SC fuses to the egg and the other to the central cell during double fertilization (McCormick, 2004; Berger and Twell, 2011).

The development of the male germline, first in the form of the GC and later as the SC (Feng *et al.*, 2013), requires finely tuned cell-cycle control, cell-identity establishment, fate determination and differentiation, as well as genome compaction and genome stability maintenance to support fertilization and post-fertilization development (Berger and Twell, 2011). Genetic studies have identified a body of genes involved in GC division and sperm specification, such as *FBL 17* (*F-box-like 17*), *CDKA;1* (*A-type cyclin-dependent kinase*), *DUO1* (*DUO pollen 1*), *DUO3* (*DUO pollen 3*) and *RBR* (*Retinoblastoma-related*), with *GEM1*

(*GEMINI POLLEN 1*) and *TIO* (*TWO-IN-ONE*) being required for microspore asymmetric division (Park *et al.*, 1998; Oh *et al.*, 2005; Iwakawa *et al.*, 2006; Kim *et al.*, 2008; Brownfield *et al.*, 2009a,b; Chen *et al.*, 2009); however, the mechanisms underlying such regulation are still largely unknown.

The nucleosome is the structural unit of chromatin in eukaryotes, and represents an approximately 147 bp fragment of DNA that is wrapped in 1.7 turns around a protein octamer of highly conserved core histones: H2A, H2B, H3 and H4. The binding of linker histone H1 to DNA entry/exit points of nucleosomes and linker DNA between two nucleosomes facilitates further compaction of chromatin into a higher-order structure. The accessibility and compaction of chromatin in individual cell types, defined by diverse epigenetic ways such as incorporation of histone variants and post-transcriptional modifications (PTMs) of histones, contribute to establishing and maintaining cell identities by affecting gene expression (Henikoff and Smith, 2015). Studies of animals have identified functionally diverse variants of all four histone families, except the H4 family, in almost all eukaryotes and lineage-specific variants (Talbert and Henikoff, 2010). Animal male germ cells express multiple histone variants such as TH2A, H2A.X, TH2B, TH3, H3.3, CenH3, H1T and H1T2. During mammalian spermatogenesis, these variants are incorporated into nucleosomes to restructure chromatin or to mark specific chromatin domains, and are essential for specifying primordial germ cells, meiosis, sex chromosome condensation, X-chromosome inactivation, and further maturation of spermatids (Sasaki and Matsui, 2008). The maturation of spermatids involves replacement of histones by protamine, which results in a tightly packed sperm nucleus that is important for genome integrity and stability, and hence normal sperm function. Large-scale incorporation of new histone variants such as H2AL1/2 and H2A.Z occurred before the exchange, and TH2B plays a key role in mediating the histone-to-protamine packing of the sperm genome, with roles in subsequent fertilization (Montellier *et al.*, 2013). Recent omics analysis demonstrated that 4–10% of the spermatozoa genome remains as nucleosome in humans, and these retained histones are enriched in developmental genes, miRNAs and imprinted genes (Hammoud *et al.*, 2009; Carrell, 2012). In agreement, key developmental genes are bivalently marked with H3K4me3 and H3K27me3, similar to the well-defined epigenetic features of stem cells (Hammoud *et al.*, 2009). The poised state may provide a critical means to define germ cell identity and to transfer epigenetic information to the offspring (Lesch and Page, 2014).

In flowering plants, the two products of microspore asymmetric division have a distinct nucleus appearance: dispersed chromatin for the VC and condensed chromatin for the GC. The daughter cells of the GC also have condensed chromatin. In addition, a number of transposable elements are expressed in VCs but not SCs (Slotkin *et al.*, 2009). These findings suggest distinct accessibility and compaction of chromatin in these cell types, resulting from different histone patterns of individual cell types. However, because of the difficulty in isolating GCs and SCs from pollen (Lu *et al.*, 2015), the knowledge of histone dynamics related to male germline development and histone patterns of each cell type is limited. Several studies of lily pollen identified the GC-preferential histone variants gH2A, gH2B, gcH3, gH3, leH3, soH3-1 and soH3-2 (Ueda and Tanaka, 1994, 1995; Ueda *et al.*, 2000; Sano and Tanaka, 2005; Okada *et al.*, 2006b), and different H3 variants have also been observed between VCs and SCs in Arabidopsis (Ingouff *et al.*, 2007, 2010).

Advances in proteomic technologies have created the opportunity to obtain broad knowledge of histone profiles in various cell types of animals and plants (Boyne *et al.*, 2006; Siuti *et al.*, 2006; Wu *et al.*, 2009; Molden *et al.*, 2015), but use of these technologies in dissecting histone patterns of the developing male germline in plants is hindered by the difficulty in isolating large amounts of GCs, SCs and VC nuclei (VN) (Lu *et al.*, 2015).

Lily pollen is bicellular at anthesis. *In vitro*-cultured lily pollen tubes grow synchronously to allow SC generation via GC mitosis in the growing tube (Zhao *et al.*, 2013). By using an optimized cell purification procedure, we successfully isolated a large amount of GCs and SCs at high purity from lily pollen (Zhao *et al.*, 2013). In the current study, we created a lily histone database using a lily protein database constructed using RNA-seq to analyze histone programs of male germline cells as well as their companion cell VCs. We successfully isolated nuclei of VCs, GCs and SCs, and compared the histone composition and profiles of the three cell types by combining 2D gel electrophoresis, mass spectrometry, and enzymatic digestion with the endoproteinase Glu-C and trypsin. We revealed 92 identities, representing 32 histone sequences, including six variants for H1, 11 for H2A, eight for H2B, five for H3 and two for H4. GCs and SCs showed almost identical histone composition and profiles, with significant differences from VCs. We observed nine male germline-specific variants, and also male germline-specific isoforms of several histone variants. These results provide insight into the histone landscape of the male germline and the companion cell. They also provide clues for understanding mechanisms underlying the establishment and specification of male germline identity and reprogramming chromatin activity during male germline development.

## RESULTS

### Custom-made database of lily histones

Lily is a model system for exploring the molecular physiological mechanisms of pollen tube growth and male

germline development (Kost, 2008; Singh *et al.*, 2008), but its application is limited due to its huge genome and the lack of genomic resources. To identify histone variants in the male germline, we first performed RNA-seq using RNAs pooled from microspores, germinated pollen grains, GCs and SCs, as well as roots. This design is highly suitable for identifying male germline-related histone transcripts. The analysis revealed more than 55 million clean reads, covering approximately 5 billion base pairs, which were assembled into 50 960 unigenes with a mean length of 568 bp. Finally, we created a lily protein database with 31 345 entries (Table S1).

We then constructed a lily histone database by extracting sequences annotated with histone-related keywords from our lily protein database. This non-redundant lily histone database has 84 histone entries: 16 for the H1 family, 28 for the H2A family, 12 for the H2B family, 23 for the H3 family, and five for the H4 family (Table S2). This distribution is compatible with previous studies of histones showing that the H2A family has 20 unique sequences in human

(Arnaudo *et al.*, 2011) and 13 unique sequences in Arabidopsis, with 11 unique sequences of the H2B family in Arabidopsis (Talbert *et al.*, 2012). Together, these results indicate that our database correctly reflects the histone sequences in lily, and will be helpful for further identification of histone variants expressed in male germline cells.

### Isolation and characterization of VC nuclei, GC nuclei and SC nuclei

To prepare histones from VCs, GCs and SCs, we isolated nuclei from these cells, which displayed a distinct chromatin appearance (Figure 1a–h). We established a method to isolate VN directly from the just-germinated pollen grain, which contains the VN and the GC (Figure 1a), and GC nuclei (GN) and SC nuclei (SN) from GCs and SCs, which were isolated from just-germinated pollen grains and from pollen tubes cultured for 10 h, respectively (Figure 1a,c–h). Isolated VN had no contamination of GCs, were almost spherical (20–30 μm in diameter) and showed light aceto-carmine staining (Figure 1d).
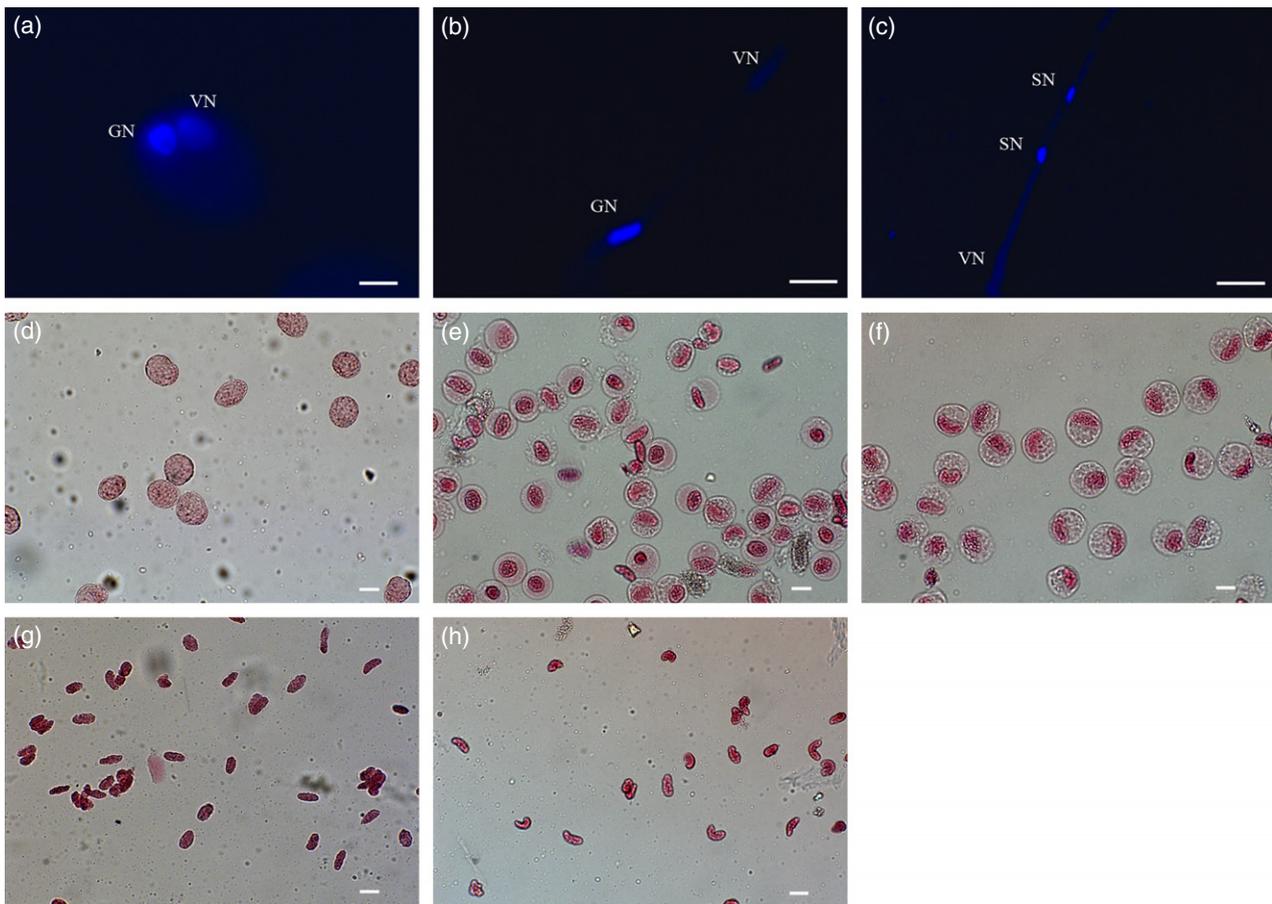


**Figure 1.** Cytological characterization of VC nuclei (VN), GC nuclei (GN) and sperm-cell nuclei (SN) isolated from lily (*Lilium davidii*) pollen.
(a–c) DAPI-stained nuclei in a just-germinated pollen grain (a), and in pollen tubes *in vitro* cultured for 2 h (b) and 10 h (c).
(d–h) Isolated VN (d), GCs (e), SCs (f), GN (g) and SN (h) stained with aceto-carmine.
Scale bars = 20 μm (a), 100 μm (b, c) and 50 μm (d–h).

Isolated GCs and SCs were spherical, with similar size (35–40 μm in diameter) but different cytological features (Figure 1e,f). Most GN appeared to be positioned in the middle of the cell, and SN were close to the cell membrane (Figure 1e,f). In addition, the cytoplasm appearance differed (Zhao *et al.*, 2013). A proteomic study showed that GCs had higher levels of phosphoenolpyruvatecarboxykinase and vacuolar invertase than SCs (Zhao *et al.*, 2013). The levels of the two proteins were higher in isolated GCs than SCs (Figure S1), consistent with the previous report (Zhao *et al.*, 2013). The isolated GCs and SCs were further used to isolate the respective nuclei. The isolated GN and SN were intact and had an ellipsoid shape (long axis of approximately 30 μm) (Figure 1g,h). These results indicate that isolated VN, GN and SN did not contaminate each other. Overall, VN appeared loose but GN and SN appeared dense (Figure 1), suggesting that VN have less condensed chromatin than GN and SN.

### Histone patterns in VCs, GCs and SCs

To compare the histone patterns of VCs, GCs and SCs, we isolated histones from nuclei, and resolved them on two-dimensional TAU/SDS gels with triplicate independent biological repeats for each sample. This method conferred high reproducibility and resolution (Figure 2a and Figure S2). We detected 32 spots in the VC gel and 49 each in the GC and SC gels (Figure 2a). The histone expression profile of GCs (used as the reference gel) showed low correlation with that of VCs ($r = 0.7$) but significantly positive correlation with that of SCs ($r = 0.94$).

Statistical analysis revealed 36 spots with significantly different abundance among the three samples: one (spot 38) was detected only in VCs, 18 were detected exclusively in GCs and SCs, and 17 showed differential abundance in VCs, GCs and SCs. The remaining 14 spots did not show any change in level between all three samples (Figure 2 and Table S3). These results suggest that SCs and GCs have almost identical histone expression patterns, which are significantly different from those of VCs.

### Identification of histone variants and isoforms

To reveal histone variants expressed in developing male germline cells, we analyzed all 49 spots excised from the reference gel (GC gel) and spot 38, which was detected only in the VC gel, by mass spectrometry, and obtained MS/MS spectra for all 50 spots. Using the stringent standard of unused score not <4, we successfully identified proteins from 49 spots but not from spot 38. Among the 49 spots, 20 contained a single protein each, and the remaining 29 had more than one protein each (Table S4). In total, we revealed 92 identities representing 32 unique histone sequences (Table 1 and Table S4). These histone variants were named (Table 1), according to the guidelines for histone nomenclature described previously (Talbert *et al.*, 2012), based on combined phylogenetic, multiple sequence alignment and conserved domain/motif analyses (Figure 3).

We discuss the H2A family as an example. Phylogenetic analysis revealed 11 variants of this family grouped into five branches: one for canonical H2A, five for H2A.W, two for H2A.X, one for H2A.Z, and two for gH2A (Figure 3a and Table 1). In agreement with this result, the members in the H2A.X and H2A.Z branches had an SQEF motif and docking domain, respectively (Figure 3b,c), which are characteristic motifs for these sub-families (Bonisch and Hake, 2012). H2A.W is a plant-specific variant of H2A (Kawashima *et al.*, 2015). The assigned H2A.W variants had putative minor groove-binding motifs (KSPKK) in their C-terminal tails (Figure 3d), as seen in known H2A.W variants from other plants, which may wrap more DNA than other H2A

**Figure 2.** Representative TAU/SDS gel images.
(a) Representative gel image of GC histones used as the reference gel image. Histones were separated by TAU/SDS–PAGE with blue silver staining. Protein spots are numbered using four colors: black, no change in levels in all three samples; green, differential change in levels in two or all three samples; red, exclusively detected in GCs and SCs; pink, detected only in VCs.
(b) Close-up of representative protein spots. Representative spots are indicated by red arrowheads. White arrowheads indicate that the corresponding spot is absent.
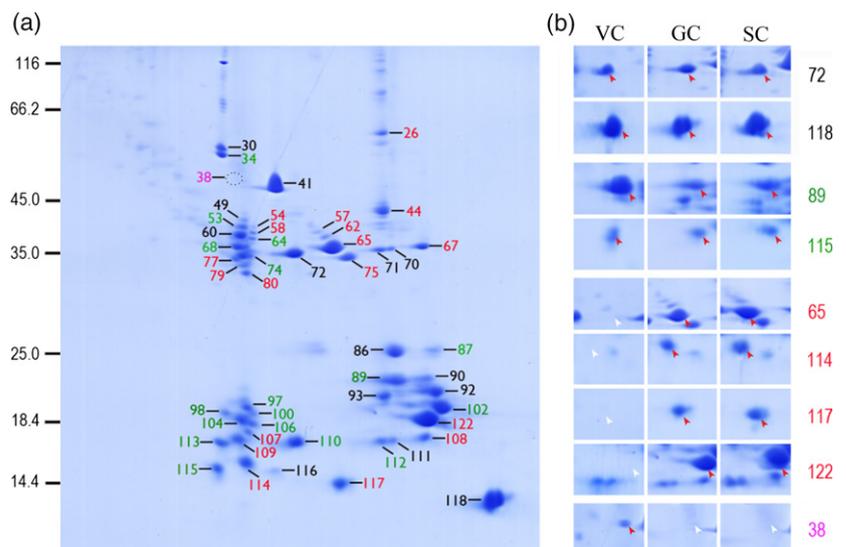
**Table 1** Identified histones and their expression patterns in VCs, GCs and SCs of *Lilium davidii*

| Family | IN | SN | Histone | RNA-seq ID | Detected PTMs | Spot pattern[a] | | | Change[b] (VC/ GC/SC) |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | VCs | GCs | SCs | |
| H1 | 1 | 41 | H1.1 | UniGene23759 | Ac($K_{219}$,$K_{229}$,$K_{251}$,$K_{260}$), Pho($S_{277}$) | + | + | + | N |
| | 2 | 65.2 | H1.1 | | | − | + | + | N |
| | 3 | 57.1 | H1.2 | CL2282.Contig1 | | − | + | + | N |
| | 4 | 57.3 | H1.3 | CL2845.Contig1 | | − | + | + | N |
| | 5 | 62.3 | H1.3 | | | − | + | + | N |
| | 6 | 65.1 | H1.3 | | | − | + | + | N |
| | 7 | 75 | H1.3 | | | − | + | + | N |
| | 8 | 57.4 | H1.4 | UniGene23758 | | − | + | + | N |
| | 9 | 62.2 | H1.4 | | | − | + | + | N |
| | 10 | 57.2 | H1.5 | CL2845.Contig2 | | − | + | + | N |
| | 11 | 62.1 | H1.5 | | | − | + | + | N |
| | 12 | 62.5 | H1.6 | CL6452.Contig1 | | − | + | + | N |
| H2A | 13 | 115 | H2A | CL1479.Contig2 | Ac($K_5$), Me2($K_{126}$) | + | + | + | 1/0.76/0.76 |
| | 14 | 49.1 | H2A.W.1 | CL1133.Contig4 | | + | + | + | N |
| | 15 | 60.1 | H2A.W.1 | | Me1($K_{27}$) | + | + | + | N |
| | 16 | 53.2 | H2A.W.1 | | Me1($K_{27}$), Ac($K_{27}$) | + | + | + | ND |
| | 17 | 68.1 | H2A.W.1 | | Ac($K_{27}$), Me1($K_{102}$) | + | + | + | ND |
| | 18 | 104 | H2A.W.1 | | Me1($K_{102}$) | + | + | + | 1/0.28/0.31 |
| | 19 | 77.2 | H2A.W.1 | | | − | + | + | N |
| | 20 | 60.5 | H2A.W.2 | UniGene23483 | | + | + | + | N |
| | 21 | 53.4 | H2A.W.2 | | | + | + | + | ND |
| | 22 | 64 | H2A.W.2 | | Ac($K_4$) | + | + | + | N |
| | 23 | 74.2 | H2A.W.2 | | Me1($K_{101}$) | + | + | + | ND |
| | 24 | 106 | H2A.W.2 | | Ac($K_4$) | + | + | + | N |
| | 25 | 54.1 | H2A.W.2 | | | − | + | + | N |
| | 26 | 58.2 | H2A.W.2 | | Ac($K_4$) | − | + | + | N |
| | 27 | 98 | H2A.W.3 | CL1133.Contig2 | Me2($K_{131}$), Me3($K_{135}$) | + | + | + | 1/1.1/1.7 |
| | 28 | 60.3 | H2A.W.4 | CL440.Contig1 | | + | + | + | N |
| | 29 | 53.1 | H2A.W.4 | | Ac($K_7$,$K_{26}$), Me1($R_{12}$) | + | + | + | ND |
| | 30 | 97.1 | H2A.W.5 | UniGene22454 | | + | + | + | ND |
| | 31 | 100.1 | H2A.W.5 | | Me1($K_{75}$), Pho($S_{108}$) | + | + | + | ND |
| | 32 | 113.1 | H2A.X.1 | UniGene25543 | Ac($K_{17}$,$K_{39}$), Me1($K_{99}$,$K_{128}$) | + | + | + | ND |
| | 33 | 34.4 | H2A.X.2 | UniGene25544 | | + | + | + | ND |
| | 34 | 110.1 | H2A.X.2 | | | + | + | + | ND |
| | 35 | 109.2 | H2A.X.2 | | | − | + | + | N |
| | 36 | 116 | H2A.Z | UniGene17362 | Ac($K_6$) | + | + | + | N |
| | 37 | 30.4 | gH2A.1 | UniGene18973 | | + | + | + | N |
| | 38 | 60.4 | gH2A.1 | | | + | + | + | N |
| | 39 | 68.2 | gH2A.1 | | Ac($K_{116}$) | + | + | + | ND |
| | 40 | 113.2 | gH2A.1 | | | + | + | + | ND |
| | 41 | 77.1 | gH2A.1 | | | − | + | + | N |
| | 42 | 79.2 | gH2A.1 | | | − | + | + | N |
| | 43 | 107 | gH2A.1 | | | − | + | + | −/1/0.51 |
| | 44 | 109.1 | gH2A.1 | | Me1($K_{98}$) | − | + | + | N |
| | 45 | 68.3 | gH2A.2 | CL549.Contig1 | Me1($K_{90}$) | + | + | + | ND |
| | 46 | 74.1 | gH2A.2 | | | + | + | + | ND |
| | 47 | 77.3 | gH2A.2 | | | − | + | + | N |
| | 48 | 79.1 | gH2A.2 | | | − | + | + | N |
| | 49 | 80 | gH2A.2 | | | − | + | + | N |
| | 50 | 114 | gH2A.2 | | Me1($K_{90}$), Ac($K_{108}$) | − | + | + | N |
| H2B | 51 | 93.2 | H2B.1 | UniGene26027 | Me1($K_{23}$,$K_{33}$,$K_{133}$), Pho($T_{16}$) | + | + | + | N |
| | 52 | 89.2 | H2B.2 | UniGene26029 | Pho($T_{16}$), Me1($K_{23}$,$K_{133}$), Me2($K_{12}$,$K_{23}$), Ac($K_{57}$,$K_{141}$) | + | + | + | ND |
| | 53 | 89.1 | H2B.3 | UniGene1146 | Me1($K_{19}$,$K_{124}$), Me2($K_{19}$), Pho($T_3$), Me3($K_{19}$), Ac ($K_{19}$,$K_{30}$,$K_{48}$,$K_{132}$) | + | + | + | ND |
| | 54 | 92.2 | H2B.4 | CL2535.Contig1 | | + | + | + | N |
| | 55 | 102.1 | H2B.4 | | Me1($K_{10}$), Me2($K_{15}$), Ac($K_{110}$) | + | + | + | ND |

(continued)

**Table 1.** (continued)

| Family | IN | SN | Histone | RNA-seq ID | Detected PTMs | Spot pattern[a] | | | Change[b] (VC/GC/SC) |
|--------|-----|------|-----------|----------------|-------------------------------------------------------|------|-----|-----|------|
| | | | | | | VCs | GCs | SCs | |
| | 56 | 90.2 | H2B.5 | CL779.Contig1 | Me1($K_{25}$), Me2($K_{17}$) | + | + | + | N |
| | 57 | 92.1 | H2B.5 | | Me1($K_{17}$,$K_{24}$,$K_{35}$), Me2($K_{12}$,$K_{17}$,$K_{24}$) | + | + | + | N |
| | 58 | 102.2 | H2B.5 | | Me2($K_{17}$) | + | + | + | ND |
| | 59 | 90.1 | H2B.6 | CL779.Contig2 | Me2($K_{17}$) | + | + | + | N |
| | 60 | 100.2 | H2B.6 | | | + | + | + | ND |
| | 61 | 62.4 | mgH2B.in | UniGene17749 | | − | + | + | N |
| | 62 | 67.1 | mgH2B | UniGene7369 | | − | + | + | N |
| | 63 | 108.2 | mgH2B | | | − | + | + | N |
| | 64 | 122 | mgH2B | | Me1($R_{57}$,$K_{65}$,$K_{103}$,$K_{136}$), Ac($K_{31}$) | − | + | + | N |
| H3 | 65 | 70 | H3.1.1 | CL542.Contig4 | | + | + | + | N |
| | 66 | 111 | H3.1.1 | | | + | + | + | N |
| | 67 | 112.2 | H3.1.1 | | | + | + | + | ND |
| | 68 | 71 | H3.1.2 | UniGene17523 | | + | + | + | N |
| | 69 | 72 | H3.3 | CL542.Contig2 | | + | + | + | N |
| | 70 | 74.3 | H3.3 | | Ac($K_9$,$K_{14}$,$K_{18}$,$K_{23}$) | + | + | + | ND |
| | 71 | 110.2 | H3.3 | | Ac($K_{18}$,$K_{23}$,$K_{27}$) | + | + | + | ND |
| | 72 | 79.3 | H3.3 | | Ac($K_{18}$,$K_{23}$) | − | + | + | N |
| | 73 | 67.2 | H3.3 like | CL542.Contig14 | Ac($K_{27}$,$K_{37}$) | − | + | + | N |
| | 74 | 108.1 | H3.3 like | | | − | + | + | N |
| | 75 | 30.3 | cenH3 | UniGene19707 | | + | + | + | N |
| | 76 | 49.2 | cenH3 | | | + | + | + | N |
| | 77 | 60.2 | cenH3 | | | + | + | + | N |
| | 78 | 93.1 | cenH3 | | Me1($R_{32}$,$R_{97}$), Me2($R_{32}$) | + | + | + | N |
| | 79 | 34.3 | cenH3 | | | + | + | + | ND |
| | 80 | 53.3 | cenH3 | | | + | + | + | ND |
| | 81 | 68.4 | cenH3 | | | + | + | + | ND |
| | 82 | 97.2 | cenH3 | | | + | + | + | ND |
| | 83 | 112.1 | cenH3 | | | + | + | + | ND |
| | 84 | 26 | cenH3 | | Me1($R_{32}$,$R_{97}$,$K_{145}$) | − | + | + | N |
| | 85 | 44.1 | cenH3 | | Ac($K_{142}$) | − | + | + | N |
| | 86 | 54.2 | cenH3 | | | − | + | + | N |
| | 87 | 58.1 | cenH3 | | | − | + | + | N |
| | 88 | 62.6 | cenH3 | | | − | + | + | N |
| | 89 | 108.3 | cenH3 | | | − | + | + | N |
| H4 | 90 | 118 | H4 | CL122.Contig3 | Ac($K_{12}$), Me1($R_{55}$,$R_{77}$) | + | + | + | N |
| | 91 | 44.2 | H4 | | | − | + | + | N |
| | 92 | 117 | mgH4 | CL2634.Contig2 | Me1($R_{54}$), Ac($K_{78}$), Pho($T_{79}$) | − | + | + | −/1/1.42 |

IN, identity number; SN, spot number.
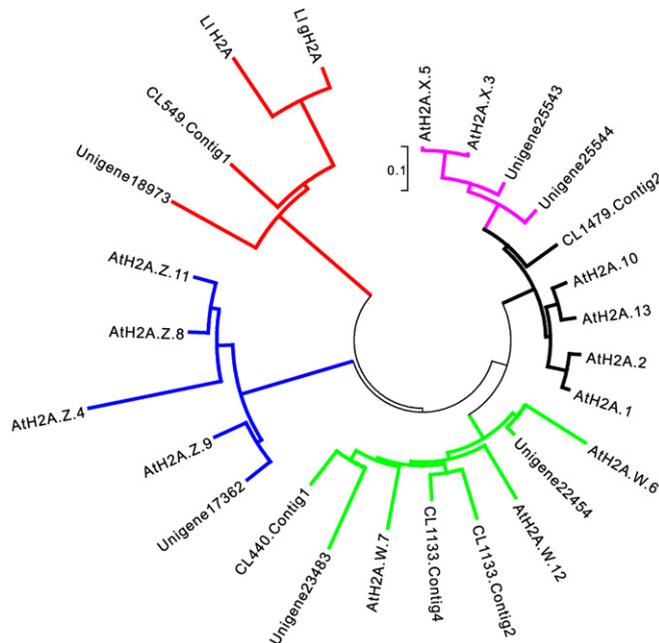[a]Protein spots detected in protein gel images of corresponding cells. −, spot absent; +, spot present.
[b]Change in expression levels of a protein identified in a spot: N, no change; ND, cannot determine whether this identity changed in levels in this spot because this spot was identified as having two or more identities.
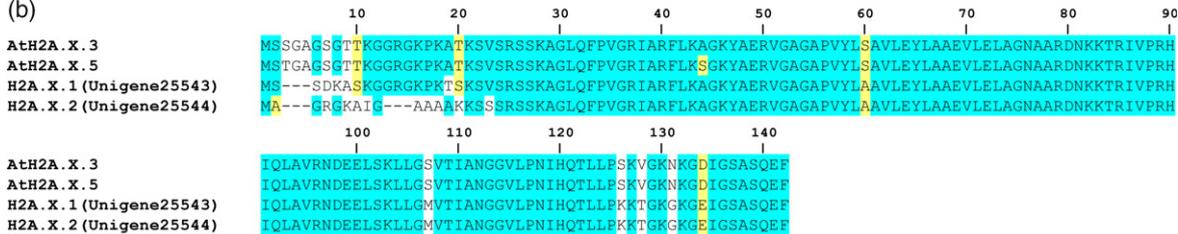
variants and function in defining heterochromatin (Lindsey *et al.*, 1991; Yelagandula *et al.*, 2014). The identified canonical H2A (CL1479.Contig2) was in a branch with the Arabidopsis canonical H2A (AtH2A.1, AtH2A.2, AtH2A.10 and AtH2A.13) (Talbert *et al.*, 2012), and had 82% sequence identity with AtH2A.2 (At4 g27230). gH2A.1 (UniGene18973) and gH2A.2 (CL549.Contig1), together with a previously reported lily H2A and the lily male germ cell-specific H2A (gH2A, NCBInr database accession number GI75311054), formed an independent clade that did not contain Arabidopsis H2A variants and showed 71% and 75% sequence identity with the gH2A, respectively.

Our analysis revealed six variants for H1, 11 for H2A, eight for H2B, five for H3, and two for H4 (Table 1). These data are in agreement with previous observations that, among the four core histones, the H2A family has the highest number of variants in a species (Millar, 2013). Our studies showed that among the six H1 variants, only one was common to all three cell types, and the remaining five were detected only in GCs and SCs (Table 1). Our data also revealed the male germline-specific variants mgH2B.in, mgH2B, H3.3-like and mgH4, and identified sequence features of these variants (Figure 4 and Table 1).
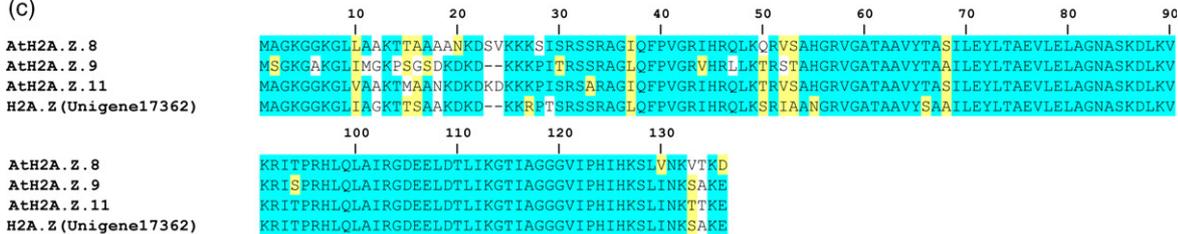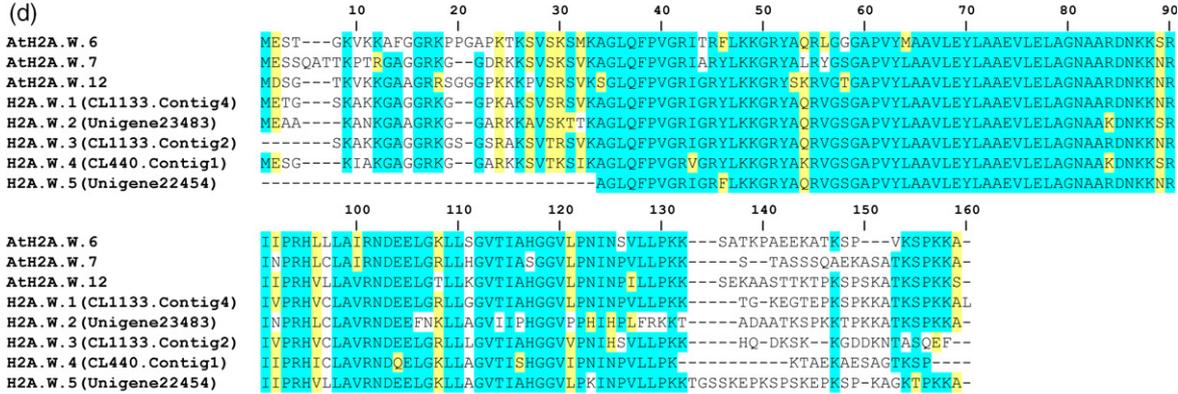
(a)

LlH2A
LlgH2A
CL549.Contig1
Unigene18973
AtH2A.X.5
AtH2A.X.3
Unigene25543
Unigene25544
CL1479.Contig2
AtH2A.10
AtH2A.13
AtH2A.2
AtH2A.1
AtH2A.W.6
Unigene22454
AtH2A.W.12
CL1133.Contig2
CL1133.Contig4
AtH2A.W.7
Unigene23483
CL440.Contig1
Unigene17362
AtH2A.Z.9
AtH2A.Z.4
AtH2A.Z.8
AtH2A.Z.11

0.1

| Color | Group |
|---|---|
| Black | H2A |
| Green | H2A.W |
| Blue | H2A.Z |
| Magenta | H2A.X |
| Red | gH2A |

(b)

```
                10        20        30        40        50        60        70        80        90
AtH2A.X.3         MSSGAGSGTTKGGRGKPKATKSVSRSSKAGLQFPVGRIARFLKAGKYAERVGAGAPVYLSAVLEYLAAEVLELAGNAARDNKKTRIVPRH
AtH2A.X.5         MSTGAGSGTTKGGRGKPKATKSVSRSSKAGLQFPVGRIARFLKSGKYAERVGAGAPVYLSAVLEYLAAEVLELAGNAARDNKKTRIVPRH
H2A.X.1(Unigene25543) MS---SDKASKGGRGKPKTSKSVSRSSKAGLQFPVGRIARFLKAGKYAERVGAGAPVYLAAVLEYLAAEVLELAGNAARDNKKTRIVPRH
H2A.X.2(Unigene25544) MA---GRGKAIG---AAAAKKSSSRSSKAGLQFPVGRIARFLKAGKYAERVGAGAPVYLAAVLEYLAAEVLELAGNAARDNKKTRIVPRH

                100       110       120       130       140
AtH2A.X.3         IQLAVRNDEELSKLLGSVTIANGGVLPNIHQTLLPSKVGKNKGDIGSASQEF
AtH2A.X.5         IQLAVRNDEELSKLLGSVTIANGGVLPNIHQTLLPSKVGKNKGDIGSASQEF
H2A.X.1(Unigene25543) IQLAVRNDEELSKLLGMVTIANGGVLPNIHQTLLPKKTGKGKGEIGSASQEF
H2A.X.2(Unigene25544) IQLAVRNDEELSKLLGMVTIANGGVLPNIHQTLLPKKTGKGKGEIGSASQEF
```

(c)

```
                10        20        30        40        50        60        70        80        90
AtH2A.Z.8         MAGKGGKGLLAAKTTAAAANKDSVKKKSISRSSRAGIQFPVGRIHRQLKQRVSAHGRVGATAAVYTASILEYLTAEVLELAGNASKDLKV
AtH2A.Z.9         MSGKGAKGLIMGKPSGSDKDKD--KKKPITRSSRAGLQFPVGRVHRLLKTRSTAHGRVGATAAVYTAAILEYLTAEVLELAGNASKDLKV
AtH2A.Z.11        MAGKGGKGLVAAKTMAANKDKDKDKKKPISRSRAGIQFPVGRIHRQLKTRVSAHGRVGATAAVYTASILEYLTAEVLELAGNASKDLKV
H2A.Z(Unigene17362) MAGKGGKGLIAGKTTSAAKDKD--KKRPTSRSSRAGLQFPVGRIHRQLKSRIAANGRVGATAAVYSAAILEYLTAEVLELAGNASKDLKV

                100       110       120       130
AtH2A.Z.8         KRITPRHLQLAIRGDEELDTLIKGTIAGGGVIPHIHKSLVNKVTKD
AtH2A.Z.9         KRISPRHLQLAIRGDEELDTLIKGTIAGGGVIPHIHKSLINKSAKE
AtH2A.Z.11        KRITPRHLQLAIRGDEELDTLIKGTIAGGGVIPHIHKSLINKTTKE
H2A.Z(Unigene17362) KRITPRHLQLAIRGDEELDTLIKGTIAGGGVIPHIHKSLINKSAKE
```

(d)

```
                10        20        30        40        50        60        70        80        90
AtH2A.W.6         MEST---GKVKKAFGGRKPPGAPKTKSVSKSMKAGLQFPVGRITRFLKKGRYAQRLGGGAPVYMAAVLEYLAAEVLELAGNAARDNKKSR
AtH2A.W.7         MESSQATTKPTRGAGGRKG--GDRKKSVSKSVKAGLQFPVGRIARYLKKGRYALRYGSGAPVYLAAVLEYLAAEVLELAGNAARDNKKNR
AtH2A.W.12        MDSG---TKVKKGAAGRKRSGGGPKKKPVSRSVKSGLQFPVGRIGRYLKKGRYSKRVGTGAPVYLAAVLEYLAAEVLELAGNAARDNKKNR
H2A.W.1(CL1133.Contig4) METG---SKAKKGAGGRKG--GPKAKSVSRSVKAGLQFPVGRIGRYLKKGRYAQRVGSGAPVYLAAVLEYLAAEVLELAGNAARDNKKNR
H2A.W.2(Unigene23483)   MEAA----KANKGAAGRKG--GARKKAVSKTTKAGLQFPVGRIGRYLKKGRYAQRVGSGAPVYLAAVLEYLAAEVLELAGNAAKDNKKSR
H2A.W.3(CL1133.Contig2) -------SKAKKGAGGRKGS-GSRAKSVTRSVKAGLQFPVGRIGRYLKKGRYAQRVGSGAPVYLAAVLEYLAAEVLELAGNAAKDNKKSR
H2A.W.4(CL440.Contig1)  MESG---KIAKGAGGRKG--GARKKSVTKSIKAGLQFPVGRVGRYLKKGRYAKRVGSGAPVYLAAVLEYLAAEVLELAGNAAKDNKKSR
H2A.W.5(Unigene22454)   ----------------------------AGLQFPVGRIGRFLKKGRYAQRVGSGAPVYLAAVLEYLAAEVLELAGNAARDNKKNR

                100       110       120       130       140       150       160
AtH2A.W.6         IIPRHLLLAIRNDEELGKLLSGVTIAHGGVLPNINSVLLPKK---SATKPAEEKATKSP---VKSPKKA
AtH2A.W.7         INPRHLCLAIRNDEELGRLLHGVTIASGGVLPNINPVLLPKK-----S--TASSSQAEKASATKSPKKA
AtH2A.W.12        IIPRHVLLAVRNDEELGTLLKGVTIAHGGVLPNINPILLPKK-----SEKAASTTKTPKSPSKATKSPKKS
H2A.W.1(CL1133.Contig4) IVPRHVCLAVRNDEELGRLLGGVTIAHGGVLPNINPVLLPKK-----TG-KEGTEPKSPKKATKSPKKAL
H2A.W.2(Unigene23483)   INPRHLCLAVRNDEEFNKLLAGVIIPHGGVPPHIHPLFRKKT-----ADAATKSPKKTPKKATKSPKKA
H2A.W.3(CL1133.Contig2) IVPRHVCLAVRNDEELGRLLLGVTIAHGGVVPNIHSVLLPKK-----HQ-DKSK--KGDDKNTASQEF-
H2A.W.4(CL440.Contig1)  IIPRHICLAVRNDQELGKLLAGVTISHGGVIPNINPVLLPK-----------KTAEKAESAGTKSP----
H2A.W.5(Unigene22454)   IIPRHVLLAVRNDEELGKLLAGVTIAHGGVLPKINPVLLPKKTGSSKEPKSPSKEPKSP-KAGKTPKKA-
```

**Figure 3.** Phylogenetic analysis and multiple sequence alignment of identified H2A variants.

(a) Phylogenetic relationship of identified H2A sequences (RNA-seq ID) and known H2A sequences from *Lilium longiflorum* (Ll) and *Arabidopsis thaliana* (At). The five distinct clades in the tree correspond to five H2A sub-families: H2A, H2A.Z, H2A.X, H2A.W and gH2A. The phylogenetic tree was constructed using MEGA5.

(b) Sequence alignment of identified H2A.X and known H2A.X from *A. thaliana* with the conserved signature motif SQEF at the C-terminal tails.

(c) Sequence alignment of identified H2A.Z and known H2A.Z from *A. thaliana* with the conserved docking domain at the C-terminal tails.

(d) Sequence alignment of identified H2A.W and known H2A.W from *A. thaliana* with the conserved signature motif KSPKK at the C-terminal tails.

Alignments were performed using ClustalX and BioEdit. Identical and similar amino acids are shaded blue and yellow, respectively. The signature motif or domain is underlined in red.



**Figure 4.** Sequence characteristics of male germline-specific (mg) H2B, H3 and H4 variants.

(a) Sequence alignments of identified mgH2Bs and universal H2B variants, indicating amino acid variation between male germline-specific variants and universal variants and the five insertions in mgH2B.in (underlined in red).

(b) Sequence alignments of H3.1 and H3.3 from Arabidopsis and this study, and male germline-specific H3 variants H3-like (this study), AtH3.10, and OsH3.709 from *Oryza sativa*. The alignments show signatures (red dots) that discriminate H3.1 (A31–F41–S87–A90) from H3.3 (T31–Y41–H87–L90), and the highly variable region (red box) that is common to H3-like, AtH3.10 and OsH3.709.

(c) Sequence alignment of mgH4, H4 and AtH4 to illustrate preferential arginine residue substitutions at the N- and C-termini of mgH4 compared with canonical H4.

Alignments were performed using ClustalX and BioEdit. Identical and similar amino acids are shaded blue and yellow, respectively.

To validate these identities and expression patterns obtained using proteomic approaches, we prepared antibodies against H1.3, mgH2B and mgH4, and examined their expression profiles by Western blot analysis. The three proteins were detected in GCs and SCs but not VCs (Figure 5), in agreement with the 2D TAU/SDS and MS analyses, indicating the reliability of the 2D TAU/SDS-based proteomic results.

## Presence of multiple isoforms

Two-dimensional gel electrophoresis is a powerful approach to identify isoforms of a protein. In plants, this approach has revealed isoforms in proteomes of numerous tissues or cells such as pollen grains, GCs and SCs (Holmes-Davis *et al.*, 2005; Chu *et al.*, 2006; Dai *et al.*, 2007; Mechin *et al.*, 2007; Xu *et al.*, 2008; Zhao *et al.*, 2013). Generally, diverse PTMs are important contributors to histone isoforms (Phanstiel *et al.*, 2008; Arnaudo *et al.*, 2011). Among our identified 32 variants, 20 had isoforms: nine had two isoforms and 11 had three or more. Among these variants, cenH3 had the most isoforms (15), and the four variants of H2A had the second highest number (6–8 isoforms). Furthermore, among the 20 variants with isoforms, nine had isoforms with a distinct distribution in the VC and the male germline (Table 1).

To evaluate PTMs related to the occurrence of isoforms, we re-analyzed MS/MS data and revealed diverse PTMs in 37 of the 92 identities (40%): one PTM in 13 identities and combined PTMs in 24 identities (Table 1 and Table S5).

The PTMs we examined included acetylation, phosphorylation, and mono-, di or tri-methylation; most involved acetylation and methylation (Table 1 and Table S5). The observations suggested that these PTMs were involved in the generation of histone isoforms, consistent with another study that found that combined PTMs led to the generation of multiple H4 isoforms in human embryonic stem cells (Phanstiel *et al.*, 2008).

## Western blot examination of H3 PTMs

Methylations at K4, K9, K27 and K36 of H3 are important epigenetic markers, involved in silent chromatin (H3K27me1/me2/me3), heterochromatin organization (H3K9me1/me2), and active chromatin (H3K4me3, H3K36me3) (Fransz *et al.*, 2006 and Kouzarides, 2007). To determine the changed expression patterns of these epigenetic markers, we detected them in VCs, GCs and SCs using histones extracted from nuclei and antibodies against modifications at H3K4, H4K9, H3K27 and H3K36. Among the examined markers, the levels of most were not significantly changed in the three cell types (Figure 6 and Figure S3). Levels were decreased for H3K36me1 and H3K36me2, sharply decreased for H3K9me2, H3K27me2, H3K27me3 and H3K9ac, and undetectable for H3K9me3 in GCs and SCs compared with VCs. Levels of H3K9ac and H3K27me3 were higher in GCs than



**Figure 5.** Western blot analysis of histone expression.
Histones were prepared from nuclei of VCs, GCs and SCs, separated by SDS–PAGE, transferred to PVDF membranes, and immunodetected using primary rabbit antibodies against H1.3, mgH2B and mgH4. The antibody against H3 was used as a loading control. Three independent biological repeats were performed for each experiment.



**Figure 6.** Detection of the global H3 methylation and acethylation modification levels in VCs, GCs and SCs.
Histones were prepared from corresponding nuclei, separated by SDS–PAGE, transferred to PVDF membranes, and immunodetected using primary antibodies against the modifications indicated on the left. The antibody against H4 was used as a loading control. The image is representative of three independent biological repeats.

SCs (Figure 6 and Figure S3). Therefore, GCs and SCs had similar levels of H3 epigenetic markers, which significantly differed from those of VCs.

## DISCUSSION

We extended our recent study of isolated lily GCs and SCs (Zhao *et al.*, 2013) by isolating nuclei from GCs, SCs and VCs from lily pollen and constructing a lily protein database, with the aim of elucidating histone-based reprogramming of chromatin during male germline development. Male germline cell fate is determined after asymmetric division of the haploid microspore (Berger and Twell, 2011). The asymmetric division is essential for fate establishment and correct development of both cell types: VCs show highly dispersed chromatin, decondensed centromeric heterochromatin, and activation of transposable elements, whereas GCs have highly condensed chromatin and expression of specific genes involved in GC mitosis and further SC specification, such as *DUO1*, which also regulates the specific accumulation of H3.10 (also called HTR10) in GCs and subsequent SCs (Brownfield *et al.*, 2009a; Borg and Berger, 2015). Accordingly, we found that, after microspore asymmetric division, the two daughter cells, the GC and VC, significantly differed in histone composition and patterns (Figure 2 and Table 1), and the GCs specifically accumulated a set of histone variants and isoforms of some histone variants (Table 1). Among six examined H1 variants, five specifically accumulated in GCs and SCs, and only one was also present in VCs (Table 1). This finding was indirectly supported by previous immunocytochemistry observations that VCs showed a significantly lower level of H1 than did GCs in *Lilium longiflorum* and *Tulipa gesneriana* (Tanaka *et al.*, 1998). H1 histones bind linker DNA and the core nucleosome, and thus are involved in chromatin condensation and in determining the accessibility of regulatory proteins, including chromatin remodeling factors and DNA damage response-related factors, to nucleosomal components (Catez *et al.*, 2006). In mammals, H1 variants are required for cell fate determination and differentiation (Happel and Doenecke, 2009; Zhang *et al.*, 2012). In Arabidopsis megaspore mother cells, cell fate transition from differentiation to meiosis inception followed transient eviction of H1 variants and *de novo* incorporation, consistent with the respective decondensed and condensed chromatin states (She *et al.*, 2013). In tobacco (*Nicotiana tabacum*), H1A and H1B deficiency led to defects in male meiosis and subsequent asymmetric division of microspores (Prymakowska-Bosak *et al.*, 1999). In yeast (*Saccharomyces cerevisiae*), the linker histone Hho1 was selectively enriched in promoters of early meiotic genes and was found to be essential for meiosis and full compaction of the spore genome (Bryant *et al.*, 2012). In agreement, studies in mammals have suggested that individual H1 variants appear to localize non-

randomly in chromatin and function in negative or positive control of specific loci (Happel and Doenecke, 2009). Furthermore, H1 mediates the dependence of heterochromatic DNA methylation on the nucleosome remodeler DDM1 (deficient in DNA methylation 1), which allows stable silencing of transposable elements in cooperation with RNA-directed DNA methylation in Arabidopsis (Zemach *et al.*, 2013) and promotes formation of epigenetic silencing markers in mammalian cells (Yang *et al.*, 2013). These lines of evidence suggest that diverse germline-specific H1 variants may be involved in the high chromatin compaction of GCs and SCs and silencing of transposable elements in these cells. In addition, we found that germline-specific variants had unique features that were different from the universal variants, which may suggest functional differences (Figure 4). Together, these data suggest that the differential histone programs established in the two daughter cells after asymmetric division are important for identity establishment and differentiation of the male germline as well as the VC.

Studies of Arabidopsis have revealed differential patterns of H3 variants between VCs and SCs: the former expresses H3.3 and H3.1 and the latter expresses H3.3, cenH3 and the SC-specific H3.10 (Ingouff *et al.*, 2007, 2010). Similarly, our study revealed the presence of H3.3, cenH3 and germline-specific H3.3-like in GCs and SCs, and H3.3 in VCs. By contrast, we observed H3.1 in VCs, GCs and SCs, and cenH3 in VCs (Table 1). CenH3 and H3K9me2 are repressive markers and hallmarks of constitutive heterochromatin (Jasencakova *et al.*, 2003), and are absent in Arabidopsis VCs (Ingouff *et al.*, 2007; Schoft *et al.*, 2009). Their removal from VCs corresponded to centromeric heterochromatin decondensation (Ingouff *et al.*, 2007; Schoft *et al.*, 2009). We found cenH3 as well as H3K9me2 in lily VCs (Figure 6 and Table 1), which is consistent with previous observations in VCs of barley (*Hordeum vulgare*) and rye (*Secale cereale*) (Houben *et al.*, 2011; Pandey *et al.*, 2013). However, the isoforms of cenH3 detected in VCs and germline cells in lily appeared distinct (Table 1). The difference in levels for the two marker types in VCs from different plants suggests that mechanisms of chromatin remodeling related to male germline development are not conserved universally across angiosperms, reflecting the differential mechanisms for establishing the male germline and its companion cell VCs and/or plasticity of the histone patterns underlying specific functions. The identified H3.3-like variant (Table 1), together with Arabidopsis H3.10 and H3.709 of rice (*Oryza sativa*), which are preferentially transcribed in SCs (Brownfield *et al.*, 2009a; Anderson *et al.*, 2013), contain a motif signature that is similar to H3.3 (T31–Y41–H87–L90) (Figure 4b) but with several unique sequence characters (the presence of a variable region and an amino acid residue substitution) compared with H3.3 (Figure 4b). Considering the specificity of expression in

male germline cells, these sequence features may facilitate acquisition of a specific function in male germline differentiation and sperm specification.

In mammals, the spermatid genome is extremely compacted by the arginine-rich highly basic protamine, which replaces histone-based nucleosomes (Kimmins and Sassone-Corsi, 2005; DeRouchey *et al.*, 2013). Transcription is silenced in the resulting mature spermatids, which contain plentiful protamine with low levels of histone variants, which are positioned on important developmental genes ready for early embryo development (Rathke *et al.*, 2014). However, flowering plants do not encode protamine, and relatively little is known about how the plant SC genome is compacted (Okada *et al.*, 2005; Russell *et al.*, 2012; Anderson *et al.*, 2013). Our study revealed that SCs had almost identical histone composition and profiles to GCs, but with altered expression of several histone variants (Table 1). This finding suggests that, after asymmetric division, the mechanisms underlying histone expression and incorporation in GCs persist, and that the histone pattern of the GC is transmitted to the SC via GC mitosis, possibly according to a semi-conserved nucleosome replication model (Saze, 2008).

These germline-specific histone variants may be important for sperm genome compaction and genome stability for a number of reasons. First, the male germline specifically expressed multiple H1 variants, and such variants have been shown to function in stable silencing of transposable elements and formation of silencing markers (Yang *et al.*, 2013; Zemach *et al.*, 2013). Second, the male germline-specific variants show considerable divergence compared with the corresponding universal variants (Figure 4), which may allow them to acquire specific chromatin functions (Borg and Berger, 2015). Third, several male germline-specific variants appeared to have evolved additional sequence signatures, such as mgH2B.in. This variant has five insertions compared with other H2B variants, and the insertions are predicted to form strand and helix structures (Figure 4a). In addition, SCs still have gene expression to some extent, in contrast to spermatids, in which transcription is silenced (Engel *et al.*, 2003; Okada *et al.*, 2006a; Borges *et al.*, 2008; Borg *et al.*, 2011; Rathke *et al.*, 2014). Comparative proteomic analysis revealed that SCs and GCs had similar protein expression profiles with several differentially expressed proteins (Zhao *et al.*, 2013). Consistently, among the identified 92 identities, only four showed differential expression between SCs and GCs (IN18, H2A.W.1; IN27, H2A.W.3; IN43, gH2A.1; IN92, mgH4) (Table 1). SCs and GCs also had similar PTM patterns at H3K4, 9, 27 and 36; two markers, the permissive marker H3K9ac and the silencing marker H3K27me3, showed altered levels in GCs in comparison with SCs (Figure 6). In contrast, the two cell types significantly differed from VCs in terms of histone patterns and histone PTM patterns.

Thus, genome compaction of the male germline may occur mainly at the GC stage, and SC development from its precursor GC may involve fine-tuning chromatin remodeling, rather than significant changes.

In summary, our study revealed differences in composition and expression patterns of histones between VCs, GCs and SCs and the male germline-specific histone variants. SCs had almost identical histone composition and expression profiles to the precursor GCs but these were significantly different from those of the VCs. The differential histone programs possibly established in the two daughter cells after asymmetric division may be important for establishing identity and differentiation of the male germline as well as the companion VCs. The differences in histone landscape for the three cell types provide a basis for further understanding mechanisms of identity establishment and differentiation of the male germline.

## EXPERIMENTAL PROCEDURES

### Isolation of VC, GC and SC nuclei

Mature pollen grains were collected from lily (*Lilium davidii* var. *unicolor*) at anthesis and stored at −80°C. The stored mature pollen grains were pre-hydrated, washed with 15% sucrose solution at room temperature for 5 min to remove lipid materials enclosing the pollen grains as described previously (Zhao *et al.*, 2013), and then used to isolate VN, GN and SN. All subsequent procedures were performed at 4°C for VN isolation and at room temperature for GC and SC isolation, unless otherwise specified.

To isolate VN, pre-hydrated pollen grains were incubated in 15% sucrose solution at 27°C for 50 min to generate short pollen tubes (Zhao *et al.*, 2013). These just-germinated pollen grains were pooled using a 300-mesh hydrated screen, then osmotically shocked in pre-cooled isolation buffer (IB) (10 mM MES/KOH, pH 6.0, 5 mM EDTA, 10 mM NaCl, 0.15 mM spermine, 0.5 mM spermidine, 0.5 mM phenylmethanesulfonyl fluoride, 1 mM dithiothreitol, 7.5% sucrose) for 5 min to release VN. After removal of cell debris using a 400-mesh screen, VN in the filtrate were collected by centrifugation at 1000 *g* for 5 min, and purified using a Percoll gradient (7.5% Percoll in IB) at 1500 *g* for 20 min. After the gradient centrifugation, VN was partitioned onto the upper interface of the gradient and carefully transferred to a clean centrifuge tube, snap-frozen in liquid nitrogen and stored at −80°C.

GCs and SCs were prepared as described previously (Zhao *et al.*, 2013). In brief, just-germinated pollen grains prepared as above were osmotically shocked in IB to release GCs, then filtered through a 400-mesh screen to remove cell debris. GCs in the filtrate were collected by centrifugation, washed with IB, then purified using a 18%/24% Percoll gradient in IB by use of centrifugation at 500 *g* for 5 min. GCs partitioned onto the interface of 18 and 24% Percoll were collected and washed. For isolation of SCs, pre-hydrated pollen grains were germinated and pollen tubes were cultured in germination medium (1.6 mM H₃BO₃, 1 mM KCl, 0.5 mM CaCl₂, 15% sucrose) for 10 h to allow mitosis of GCs to generate SCs. Thereafter, pollen tubes were collected using an 80-mesh screen, osmotically shocked in 15% sucrose medium, then filtered through a 400-mesh screen. SCs in the filtrate were collected by centrifugation at 600 *g* for 5 min, washed with IB, then purified using a 8%/13% Percoll gradient in

IB by use of centrifugation at 500 *g* for 5 min. SCs at the interface of 8 and 13% Percoll were collected by use of a glass pipet.

Isolated GCs and SCs were treated with pre-cooled 0.1% Triton X-100 in IB for 10 min. The treatment led to cell membrane disruption and release of nuclei into IB. After centrifugation at 1000 *g* for 5 min, pelleted nuclei were washed twice with IB. Purified GN and SN were snap-frozen in liquid nitrogen and stored at −80°C.

Using these methods, we obtained approximately $3 \times 10^4$ VN, $4 \times 10^4$ GN and $5 \times 10^4$ SN each from 4 g pollen grains, containing approximately 25, 40 and 45 μg histones, respectively.

Morphological characteristics and the purity of nuclei were examined under a microscope (Axio Scope.A1, Zeiss, http://www.zeiss.com/microscopy/) by staining with aceto-carmine (1% carmine in 45% glacial acetic acid).

### Extraction of histones

Histones were extracted by the acid extraction method (Shechter *et al.*, 2007). In brief, isolated nuclei were re-suspended in 0.2 M $H_2SO_4$ and disrupted by vortexing until the suspension became clear. Thereafter, the suspension was incubated on a rotator at 4°C for 2 h, and centrifuged at 16 000 *g* for 20 min. Proteins in the supernatant were transferred to a new centrifuge tube, and precipitated using a 2-D Clean-Up Kit (GE Healthcare Life Sciences, www.gelifesciences.com) according to the manufacturer's instructions. Protein pellets were dissolved in acetic acid/urea sample buffer (6 M urea, 0.02% Pyronin Y (Sigma-Aldrich, http://www.sigmaaldrich.com/), 5% glacial acetic acid, 12.5 mg/ml protamine sulfate). Protein concentration was determined by the Bradford method (Bradford, 1976) using a DU640 UV-visible spectrophotometer (Beckman, http://www.beckmancoulter.cn/), with BSA as the standard. Prepared histone extracts were stored in aliquots (70 μg each) at −80°C. For each sample, triplicate independent protein preparations were used.

### TAU/SDS–PAGE separation of histones

Histones were solved on 2D gels as described previously (Shechter *et al.*, 2007). In brief, histones (70 μg) were first separated using a Triton/acetic acid/urea (TAU) polyacrylamide gel (15% polyacrylamide, 6 M urea, 5% acetic acid, 0.4% Triton X-100) on a Rubby SE600 electrophoresis unit (GE Healthcare Life Sciences) under 20 mA constant current. After electrophoresis, the sample lane was cut out, then equilibrated three times for approximately 5 min each in 0.125 M Tris/HCl, pH 8.8. The equilibrated strips were carefully transferred to 15% SDS/polyacrylamide gels and sealed with 4.5% SDS/polyacrylamide gel. The 2D gel was run under a 25 mA constant current on the Rubby SE600 unit. Low-molecular-mass protein markers (Thermo Fisher/MBI Fermentas, https://www.thermofisher.com/) were co-electrophoresed to indicate the protein relative molecular mass (MM). Protein in gels was visualized by blue silver staining as described previously (Candiano *et al.*, 2004). Gel images at 400 dpi resolution were obtained by scanning stained gels using an image scanner (UMAX, Power-Look 1120, http://www.umax.com/scaner). Triplicate biological repeats of protein preparations underwent independent separation by TAU/SDS–PAGE.

### Protein spot quantification

All gel images were analyzed using ImageMaster 2D Platinum version 5.0 (GE Healthcare Life Sciences), with the gel image from the GC sample used as a reference. After background subtraction and automatic detection of spots, nine images (three samples each with triplicate biological repeats) were individually matched to the reference image using three spot pairs that were well matched among all gels. Spots from gels that matched with those in the reference gel were grouped and used to analyze the expression profile. The mean relative volume was calculated for each spot in each sample; spots with a mean relative volume change ≥1.2-fold (Student's *t* test, $P \leq 0.05$) between samples were considered as showing a significant change in expression. Scatterplots showing correlation of histone patterns were drawn on the basis of linear dependence between the normalized spot values for one gel in comparison with the reference gel.

### Identification of proteins by ESI-QUAD-TOF MS/MS

Protein spots were excised from TAU/SDS gels and digested using trypsin (Roche Diagnostics, http://www.roche.com/) or Glu-C (Sigma-Aldrich, http://www.sigmaaldrich.com/) as described previously (Shevchenko *et al.*, 2006; Xu *et al.*, 2010). The resulting peptides were first desalted using a 100 μm × 20 mm trap column, then eluted on an analytical 75 μm × 150 mm column using a Eksigent NanoLC Ultra 2D Plus platform (AB Sciex, http://sciex.com/). Both columns were filled with MAGIC C18AQ with particle diameter 5 μm and pore size 200 Å (Bruker Michrom Bioresources, https://www.bruker.com/). Trapping and desalting were performed at 2 μl/min for 7 min using 100% buffer A (0.1% formic acid). The nanoLC was run at a flow rate of 300 nl/min for 60 min, with a gradient from 8–20% buffer B (100% acetonitrile, 0.1% formic acid) over 30 min, to 32% buffer B over the following 7 min. After the peptide elution window, the gradient was increased to 80% buffer B over 1 min, and maintained for 8 min. Thereafter, initial chromatography conditions were restored over 1 min, and maintained for 13 min.

The nanoLC system was coupled to a TripleTOF 5600+ mass spectrometer (AB Sciex) interfaced to a nanoIII source (AB Sciex). The source parameters were set as: ionspray voltage floating, 2500 V; curtain gas, 25; ion source gas, 5; interface heater temperature, 150°C; declustering potential, 100 V. All data were acquired in information-dependent acquisition mode using Analyst TF version .6 (AB Sciex). For information-dependent acquisition parameters, MS spectra were acquired across the mass range of 350–1500 *m/z* in high-resolution mode (>30 000) with a 250 msec accumulation time per spectrum. A maximum of 40 precursors per cycle was chosen for fragmentation from each MS spectrum, with a 50 msec accumulation time for each precursor. Dynamic exclusion was set as half the peak width (approximately 18 sec). Tandem mass spectra were recorded in high sensitivity mode (resolution >15 000) with rolling collision energy on.

### Histone database establishment

The lily protein database was constructed using *de novo* RNA-seq data. Total RNAs from microspores, germinated pollen grains, GCs, SCs and roots were extracted using a mirVana miRNA isolation kit (Thermo Fisher/Life Technologies, https://www.thermofisher.com/). Aliquots of 5 μg RNA from each sample were pooled for RNA-seq using a HiSeq 2000 sequencing system (Illumina, www.illumina.com) performed by Huada Genomics (http://www.genomics.cn/). Raw reads were cleaned by removing adaptors, empty reads and low-quality reads, then clean reads were assembled using Trinity (https://github.com/trinityrnaseq/trinityrnaseq/wiki) to produce unigenes. Unigenes were annotated by a BLASTx search (cut-off e-value <0.00001) against protein databases: NCBInr (http://www.ncbi.nlm.nih.gov/protein/), SwissProt (http://www.uniprot.org/blast/), KEGG (http://www.genome.jp/kegg/) and COG (http://www.ncbi.nlm.nih.gov/COG/). The alignment results were also used to predict the direction and coding

sequence of unigenes. A lily protein database was constructed by translating all coding sequences of unigenes into protein sequences. A lily histone database was constructed by extracting histone-related keywords from the lily protein database and removing redundant histone sequences. We also manually checked correctness and completion of each histone coding sequence.

## MS data analysis

For protein identification, ProteinPilot version 4.5 (AB Sciex) was used to analyze all raw MS/MS data files (*.wiff) by searching against the custom-made lily histone database. Parameters in the Paragon algorithm of ProteinPilot used for the search were set as follows: sample type, identification; Cys alkylation, iodoacetamide; digestion, trypsin or Glu-C; instrument: TripleTOF 5600; special factors, none; ID focus, biological modifications; search effort, thorough ID.

For PTM identification, the MS/MS data (*.wiff) were converted to mgf format using MS Data Converter version 1.3 (AB Sciex), and were searched against the custom-made lily histone database using Mascot server V2.4 (Matrix Science, http://www.matrix-science.com/). Parameter settings for the search were: MS tolerance, 10 ppm; MS/MS tolerance, 0.05 Da; instrument type, ESI-QUAD-TOF; fixed modification, carbamidomethyl at cysteine; variable modifications, mono- and di-methylation on lysine and arginine residues, tri-methylation and acetylation on lysine, phosphorylation on serine and threonine, oxidation on methionine. All identified modifications were validated by manually checking the MS/MS spectrum.

## Bioinformatics analysis

Identified histone sequences belonging to the same family were further aligned to compare sequence similarity using MUSCLE (http://www.ebi.ac.uk/Tools/msa/muscle/), and used to construct neighbor-joining trees using MEGA5 (http://www.megasoftware.net/), with known variants in a given histone family from *Arabidopsis thaliana* and *Lilium longiflorum* as references. Newly identified variants were named according to their relationship to references in the phylogenetic tree and guidelines proposed previously (Talbert *et al.*, 2012).

The amino acid composition of histone variants was analyzed using DNAMAN (LynnonBiosoft, http://www.lynnon.com/). The secondary structure of variants was predicted using SERp (http://services.mbi.ucla.edu/SER/) and PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/). The protein sequence alignment was analyzed using ClustalX version 1.83 (http://www.clustal.org/) and processed using BioEdit version 6.05 (http://www.mbio.ncsu.edu/bioedit/).

## Western blot analysis

Histones prepared from VCs, GCs and SCs were separated by 12% SDS–PAGE. Proteins in a gel were electrophoretically transferred onto PVDF membrane (Thermo Fisher/Pierce, https://www.thermofisher.com/) using a transfer buffer containing 25 mM Tris, 192 mM glycine and 20% methanol, then immunodetected as described previously (Dai *et al.*, 2007). For each Western blot, three biological repeats were used. The primary antibodies against H1.3, mgH2B and mgH4 were raised in rabbits by use of an antigenic determinant specific for the respective protein, produced by B&M Biotech (http://bio-med.biomart.cn/), and used at a dilution of 1:10 000 for H1.3 and mgH4, and 1:20 000 for mgH2B. The antibodies used to detect H3 PTMs (according to the manufac-

turer's instructions) were 07-436 for H3K4me1, 07-030 for H3K4me2, 07-473 for H3K4me3, 07-450 for H3K9me1, 07-441 for H3K9me2, 07-442 for H3K9me3, 07-448 for H3K27me1, 07-452 for H3K27me2, 07-449 for H3K27me3, 07-548 for H3K36me1, 07-274 for H3K36me2, 06-942 for H3K9Ac, 06-599 for H3Ac and 04-858 for H4 (all Merck-Millipore, http://www.emdmillipore.com/), plus ab1791 for H3 and ab9050 for H3K36me3 (both Abcam, http://www.abcam.com/).

## ACCESSION NUMBERS

The RNA-seq data have been deposited in the Sequence Read Archive (SRA) under the accession number SRP066393. Histone sequences for our lily histone database may be found in the European Nucleotide Archive under accession numbers LN906598–906629.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.
**Figure S1.** Western blot examination of proteins.
**Figure S2.** Gel images for the three replicates for nuclei from each cell type .
**Figure S3.** Western blot examination of global H3 methylation and acetylation modification levels.
**Table S1.** Summary of the RNA-seq assembly data.
**Table S2.** Entries in the custom-made lily histone database.
**Table S3.** Protein spot quantification analysis.
**Table S4.** MS/MS identification of proteins.
**Table S5.** PTMs of histones.

## REFERENCES

**Anderson, S.N., Johnson, C.S., Jones, D.S., Conrad, L.J., Gou, X.P., Russell, S.D. and Sundaresan, V.** (2013) Transcriptomes of isolated *Oryza sativa* gametes characterized by deep sequencing: evidence for distinct sex-dependent chromatin and epigenetic states before fertilization. *Plant J.* **76**, 729–741.

**Arnaudo, A.M., Molden, R.C. and Garcia, B.A.** (2011) Revealing histone variant induced changes via quantitative proteomics. *Crit. Rev. Biochem. Mol. Biol.* **46**, 284–294.

**Berger, F. and Twell, D.** (2011) Germline specification and function in plants. *Annu. Rev. Plant Biol.* **62**, 461–484.

**Bonisch, C. and Hake, S.B.** (2012) Histone H2A variants in nucleosomes and chromatin: more or less stable? *Nucleic Acids Res.* **40**, 10719–10741.

**Borg, M. and Berger, F.** (2015) Chromatin remodelling during male gametophyte development. *Plant J.* **83**, 177–188.

**Borg, M., Brownfield, L., Khatab, H., Sidorova, A., Lingaya, M. and Twell, D.** (2011) The R2R3 MYB transcription factor DUO1 activates a male germline-specific regulon essential for sperm cell differentiation in Arabidopsis. *Plant Cell*, **23**, 534–549.

**Borges, F., Gomes, G., Gardner, R., Moreno, N., McCormick, S., Feijo, J.A. and Becker, J.D.** (2008) Comparative transcriptomics of Arabidopsis sperm cells. *Plant Physiol.* **148**, 1168–1181.

**Boyne, M.T. 2nd, Pesavento, J.J., Mizzen, C.A. and Kelleher, N.L.** (2006) Precise characterization of human histones in the H2A gene family by top-down mass spectrometry. *J. Proteome Res.* **5**, 248–253.

**Bradford, M.M.** (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. *Anal. Biochem.* **72**, 248–254.

**Brownfield, L., Hafidh, S., Borg, M., Sidorova, A., Mori, T. and Twell, D.** (2009a) A plant germline-specific integrator of sperm specification and cell cycle progression. *PLoS Genet.* **5**, e1000430.

**Brownfield, L., Hafidh, S., Durbarry, A., Khatab, H., Sidorova, A., Doerner, P. and Twell, D.** (2009b) Arabidopsis DUO POLLEN3 is a key regulator of male germline development and embryogenesis. *Plant Cell*, **21**, 1940–1956.

**Bryant, J.M., Govin, J., Zhang, L., Donahue, G., Pugh, B.F. and Berger, S.L.** (2012) The linker histone plays a dual role during gametogenesis in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **32**, 2771–2783.

**Candiano, G., Bruschi, M., Musante, L., Santucci, L., Ghiggeri, G.M., Carnemolla, B., Orecchia, P., Zardi, L. and Righetti, P.G.** (2004) Blue silver: a very sensitive colloidal Coomassie G-250 staining for proteome analysis. *Electrophoresis*, **25**, 1327–1333.

**Carrell, D.T.** (2012) Epigenetics of the male gamete. *Fertil. Steril.* **97**, 267–274.

**Catez, F., Ueda, T. and Bustin, M.** (2006) Determinants of histone H1 mobility and chromatin binding in living cells. *Nat. Struct. Mol. Biol.* **13**, 305–310.

**Chen, Z., Hafidh, S., Poh, S.H., Twell, D. and Berger, F.** (2009) Proliferation and cell fate establishment during Arabidopsis male gametogenesis depends on the Retinoblastoma protein. *Proc. Natl Acad. Sci. USA*, **106**, 7257–7262.

**Chu, F., Nusinow, D.A., Chalkley, R.J., Plath, K., Panning, B. and Burlingame, A.L.** (2006) Mapping post-translational modifications of the histone variant MacroH2A1 using tandem mass spectrometry. *Mol. Cell Proteomics*, **5**, 194–203.

**Dai, S., Chen, T., Chong, K., Xue, Y., Liu, S. and Wang, T.** (2007) Proteomics identification of differentially expressed proteins associated with pollen germination and tube growth reveals characteristics of germinated *Oryza sativa* pollen. *Mol. Cell Proteomics*, **6**, 207–230.

**DeRouchey, J., Hoover, B. and Rau, D.C.** (2013) A comparison of DNA compaction by arginine and lysine peptides: a physical basis for arginine-rich protamines. *Biochemistry*, **52**, 3000–3009.

**Engel, M.L., Chaboud, A., Dumas, C. and McCormick, S.** (2003) Sperm cells of *Zea mays* have a complex complement of mRNAs. *Plant J.* **34**, 697–707.

**Feng, X., Zilberman, D. and Dickinson, H.** (2013) A conversation across generations: soma-germ cell crosstalk in plants. *Dev. Cell*, **24**, 215–225.

**Fransz, P., ten Hoopen, R. and Tessadori, F.** (2006) Composition and formation of heterochromatin in *Arabidopsis thaliana*. *Chromosome Res.* **14**, 71–82.

**Hammoud, S.S., Nix, D.A., Zhang, H., Purwar, J., Carrell, D.T. and Cairns, B.R.** (2009) Distinctive chromatin in human sperm packages genes for embryo development. *Nature*, **460**, 473–478.

**Happel, N. and Doenecke, D.** (2009) Histone H1 and its isoforms: contribution to chromatin structure and function. *Gene*, **431**, 1–12.

**Henikoff, S. and Smith, M.M.** (2015) Histone variants and epigenetics. *Cold Spring Harb. Perspect. Biol.* **7**, a019364.

**Holmes-Davis, R., Tanaka, C.K., Vensel, W.H., Hurkman, W.J. and McCormick, S.** (2005) Proteome mapping of mature pollen of *Arabidopsis thaliana*. *Proteomics*, **5**, 4864–4884.

**Houben, A., Kumke, K., Nagaki, K. and Hause, G.** (2011) CENH3 distribution and differential chromatin modifications during pollen development in rye (*Secale cereale* L.). *Chromosome Res.* **19**, 471–480.

**Ingouff, M., Hamamura, Y., Gourgues, M., Higashiyama, T. and Berger, F.** (2007) Distinct dynamics of HISTONE3 variants between the two fertilization products in plants. *Curr. Biol.* **17**, 1032–1037.

**Ingouff, M., Rademacher, S., Holec, S., Soljic, L., Xin, N., Readshaw, A., Foo, S.H., Lahouze, B., Sprunck, S. and Berger, F.** (2010) Zygotic resetting of the HISTONE 3 variant repertoire participates in epigenetic reprogramming in Arabidopsis. *Curr. Biol.* **20**, 2137–2143.

**Iwakawa, H., Shinmyo, A. and Sekine, M.** (2006) Arabidopsis *CDKA;1*, a *cdc2* homologue, controls proliferation of generative cells in male gametogenesis. *Plant J.* **45**, 819–831.

**Jasencakova, Z., Soppe, W.J.J., Meister, A., Gernand, D., Turner, B.M. and Schubert, I.** (2003) Histone modifications in Arabidopsis – high methylation of H3 lysine 9 is dispensable for constitutive heterochromatin. *Plant J.* **33**, 471–480.

**Kawashima, T., Lorkovic, Z.J., Nishihama, R., Ishizaki, K., Axelsson, E., Yelagandula, R., Kohchi, T. and Berger, F.** (2015) Diversification of histone H2A variants during plant evolution. *Trends Plant Sci.* **20**, 419–425.

**Kim, H.J., Oh, S.A., Brownfield, L., Hong, S.H., Ryu, H., Hwang, I., Twell, D. and Nam, H.G.** (2008) Control of plant germline proliferation by SCF (FBL17) degradation of cell cycle inhibitors. *Nature*, **455**, 1134–1137.

**Kimmins, S. and Sassone-Corsi, P.** (2005) Chromatin remodelling and epigenetic features of germ cells. *Nature*, **434**, 583–589.

**Kost, B.** (2008) Spatial control of Rho (Rac-Rop) signaling in tip-growing plant cells. *Trends Cell Biol.* **18**, 119–127.

**Kouzarides, T.** (2007) Chromatin modifications and their function. *Cell*, **128**, 693–705.

**Lesch, B.J. and Page, D.C.** (2014) Poised chromatin in the mammalian germ line. *Development*, **141**, 3619–3626.

**Lindsey, G.G., Orgeig, S., Thompson, P., Davies, N. and Maeder, D.L.** (1991) Extended C-terminal tail of wheat histone H2A interacts with DNA of the 'linker' region. *J. Mol. Biol.* **218**, 805–813.

**Lu, Y., Wei, L. and Wang, T.** (2015) Methods to isolate a large amount of generative cells, sperm cells and vegetative nuclei from tomato pollen for 'omics' analysis. *Front. Plant Sci.* **6**, 391.

**McCormick, S.** (2004) Control of male gametophyte development. *Plant Cell*, **16**(Suppl.), S142–S153.

**Mechin, V., Thevenot, C., Le Guilloux, M., Prioul, J.L. and Damerval, C.** (2007) Developmental analysis of maize endosperm proteome suggests a pivotal role for pyruvate orthophosphate dikinase. *Plant Physiol.* **143**, 1203–1219.

**Millar, C.B.** (2013) Organizing the genome with H2A histone variants. *Biochem. J.* **449**, 567–579.

**Molden, R.C., Bhanu, N.V., LeRoy, G., Arnaudo, A.M. and Garcia, B.A.** (2015) Multi-faceted quantitative proteomics analysis of histone H2B isoforms and their modifications. *Epigenetics Chromatin*, **8**, 15.

**Montellier, E., Boussouar, F., Rousseaux, S. *et al.*** (2013) Chromatin-to-nucleoprotamine transition is controlled by the histone H2B variant TH2B. *Genes Dev.* **27**, 1680–1692.

**Oh, S.A., Johnson, A., Smertenko, A., Rahman, D., Park, S.K., Hussey, P.J. and Twell, D.** (2005) A divergent cellular role for the FUSED kinase family in the plant-specific cytokinetic phragmoplast. *Curr. Biol.* **15**, 2107–2111.

**Okada, T., Endo, M., Singh, M.B. and Bhalla, P.L.** (2005) Analysis of the histone H3 gene family in Arabidopsis and identification of the male-gamete-specific variant AtMGH3. *Plant J.* **44**, 557–568.

**Okada, T., Bhalla, P.L. and Singh, M.B.** (2006a) Expressed sequence tag analysis of *Lilium longiflorum* generative cells. *Plant Cell Physiol.* **47**, 698–705.

**Okada, T., Singh, M.B. and Bhalla, P.L.** (2006b) Histone H3 variants in male gametic cells of lily and H3 methylation in mature pollen. *Plant Mol. Biol.* **62**, 503–512.

**Pandey, P., Houben, A., Kumlehn, J., Melzer, M. and Rutten, T.** (2013) Chromatin alterations during pollen development in *Hordeum vulgare*. *Cytogenet. Genome Res.* **141**, 50–57.

**Park, S.K., Howden, R. and Twell, D.** (1998) The *Arabidopsis thaliana* gametophytic mutation *gemini pollen1* disrupts microspore polarity, division asymmetry and pollen cell fate. *Development*, **125**, 3789–3799.

**Phanstiel, D., Brumbaugh, J., Berggren, W.T., Conard, K., Feng, X., Levenstein, M.E., McAlister, G.C., Thomson, J.A. and Coon, J.J.** (2008) Mass spectrometry identifies and quantifies 74 unique histone H4 isoforms in differentiating human embryonic stem cells. *Proc. Natl Acad. Sci. USA*, **105**, 4093–4098.

**Prymakowska-Bosak, M., Przewloka, M.R., Slusarczyk, J., Kuras, M., Lichota, J., Kilianczyk, B. and Jerzmanowski, A.** (1999) Linker histones play a role in male meiosis and the development of pollen grains in tobacco. *Plant Cell*, **11**, 2317–2329.

**Rathke, C., Baarends, W.M., Awe, S. and Renkawitz-Pohl, R.** (2014) Chromatin dynamics during spermiogenesis. *Biochim. Biophys. Acta*, **1839**, 155–168.

**Russell, S.D., Gou, X., Wong, C.E., Wang, X., Yuan, T., Wei, X., Bhalla, P.L. and Singh, M.B.** (2012) Genomic profiling of rice sperm cell transcripts

reveals conserved and distinct elements in the flowering plant male germ lineage. *New Phytol.* **195**, 560–573.

Sano, Y. and Tanaka, I. (2005) A histone H3.3-like gene specifically expressed in the vegetative cell of developing lily pollen. *Plant Cell Physiol.* **46**, 1299–1308.

Sasaki, H. and Matsui, Y. (2008) Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nat. Rev. Genet.* **9**, 129–140.

Saze, H. (2008) Epigenetic memory transmission through mitosis and meiosis in plants. *Semin. Cell Dev. Biol.* **19**, 527–536.

Schoft, V.K., Chumak, N., Mosiolek, M., Slusarz, L., Komnenovic, V., Brownfield, L., Twell, D., Kakutani, T. and Tamaru, H. (2009) Induction of RNA-directed DNA methylation upon decondensation of constitutive heterochromatin. *EMBO Rep.* **10**, 1015–1021.

She, W., Grimanelli, D., Rutowicz, K., Whitehead, M.W., Puzio, M., Kotlinski, M., Jerzmanowski, A. and Baroux, C. (2013) Chromatin reprogramming during the somatic-to-reproductive cell fate transition in plants. *Development*, **140**, 4008–4019.

Shechter, D., Dormann, H.L., Allis, C.D. and Hake, S.B. (2007) Extraction, purification and analysis of histones. *Nat. Protoc.* **2**, 1445–1457.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V. and Mann, M. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856–2860.

Singh, M.B., Bhalla, P.L. and Russell, S.D. (2008) Molecular repertoire of flowering plant male germ cells. *Sex. Plant Reprod.* **21**, 27–36.

Siuti, N., Roth, M.J., Mizzen, C.A., Kelleher, N.L. and Pesavento, J.J. (2006) Gene-specific characterization of human histone H2B by electron capture dissociation. *J. Proteome Res.* **5**, 233–239.

Slotkin, R.K., Vaughn, M., Borges, F., Tanurdzic, M., Becker, J.D., Feijo, J.A. and Martienssen, R.A. (2009) Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell*, **136**, 461–472.

Talbert, P.B. and Henikoff, S. (2010) Histone variants – ancient wrap artists of the epigenome. *Nat. Rev. Mol. Cell Biol.* **11**, 264–275.

Talbert, P.B., Ahmad, K., Almouzni, G. *et al.* (2012) A unified phylogeny-based nomenclature for histone variants. *Epigenetics Chromatin*, **5**, 7.

Tanaka, I., Ono, K. and Fukuda, T. (1998) The developmental fate of angiosperm pollen is associated with a preferential decrease in the level of histone H1 in the vegetative nucleus. *Planta*, **206**, 561–569.

Ueda, K. and Tanaka, I. (1994) The basic proteins of male gametic nuclei isolated from pollen grains of *Lilium longiflorum*. *Planta*, **192**, 446–452.

Ueda, K. and Tanaka, I. (1995) Male gametic nucleus-specific H2b and H3 histones, designated Gh2b and Gh3, in *Lilium longiflorum*. *Planta*, **197**, 289–295.

Ueda, K., Kinoshita, Y., Xu, Z.J., Ide, N., Ono, M., Akahori, Y., Tanaka, I. and Inoue, M. (2000) Unusual core histones specifically expressed in male gametic cells of *Lilium longiflorum*. *Chromosoma*, **108**, 491–500.

Wilson, Z.A. and Yang, C. (2004) Plant gametogenesis: conservation and contrasts in development. *Reproduction*, **128**, 483–492.

Wu, T., Yuan, T., Tsai, S.N., Wang, C., Sun, S.M., Lam, H.M. and Ngai, S.M. (2009) Mass spectrometry analysis of the variants of histone H3 and H4 of soybean and their post-translational modifications. *BMC Plant Biol.* **9**, 98.

Xu, S.B., Li, T., Deng, Z.Y., Chong, K., Xue, Y. and Wang, T. (2008) Dynamic proteomic analysis reveals a switch between central carbon metabolism and alcoholic fermentation in rice filling grains. *Plant Physiol.* **148**, 908–925.

Xu, S.B., Yu, H.T., Yan, L.F. and Wang, T. (2010) Integrated proteomic and cytological study of rice endosperms at the storage phase. *J. Proteome Res.* **9**, 4906–4918.

Yang, S.M., Kim, B.J., Toro, L.N. and Skoultchi, A.I. (2013) H1 linker histone promotes epigenetic silencing by regulating both DNA methylation and histone H3 methylation. *Proc. Natl Acad. Sci. USA*, **110**, 1708–1713.

Yelagandula, R., Stroud, H., Holec, S. *et al.* (2014) The histone variant H2A.W defines heterochromatin and promotes chromatin condensation in Arabidopsis. *Cell*, **158**, 98–109.

Zemach, A., Kim, M.Y., Hsieh, P.H., Coleman-Derr, D., Eshed-Williams, L., Thao, K., Harmer, S.L. and Zilberman, D. (2013) The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell*, **153**, 193–205.

Zhang, Y., Cooke, M., Panjwani, S. *et al.* (2012) Histone H1 depletion impairs embryonic stem cell differentiation. *PLoS Genet.* **8**, e1002691.

Zhao, X., Yang, N. and Wang, T. (2013) Comparative proteomic analysis of generative and sperm cells reveals molecular characteristics associated with sperm development and function specialization. *J. Proteome Res.* **12**, 5058–5071.