Proteomic Analysis of Histones H2A/H2B and Variant Hv1 in *Tetrahymena thermophila* Reveals an Ancient Network of Chaperones

Kanwal Ashraf,[†,1] Syed Nabeel-Shah,[†,‡,2] Jyoti Garg,[1] Alejandro Saettone,[2] Joanna Derynck,[2] Anne-Claude Gingras,[3,4] Jean-Philippe Lambert,[5,6] Ronald E. Pearlman,[*,1] and Jeffrey Fillingham[*,2]

[1]Department of Biology, York University, Toronto, ON, Canada

[2]Department of Chemistry and Biology, Ryerson University, Toronto, ON, Canada

[3]Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada

[4]Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, ON, Canada

[5]Department of Molecular Medicine and Cancer Research Centre, Université Laval, Québec, QC, Canada

[6]CHU de Québec Research Center, CHUL, Québec, QC, Canada

[†]These authors contributed equally to this work.

[‡]Present address: Donnelly Centre, University of Toronto, Toronto, ON, Canada; Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada

***Corresponding authors:** E-mails: jeffrey.fillingham@ryerson.ca; ronp@yorku.ca.

Associate editor: Amanda Larracuente

Abstract

Epigenetic information, which can be passed on independently of the DNA sequence, is stored in part in the form of histone posttranslational modifications and specific histone variants. Although complexes necessary for deposition have been identified for canonical and variant histones, information regarding the chromatin assembly pathways outside of the Opisthokonts remains limited. *Tetrahymena thermophila*, a ciliated protozoan, is particularly suitable to study and unravel the chromatin regulatory layers due to its unique physical separation of chromatin states in the form of two distinct nuclei present within the same cell. Using a functional proteomics pipeline, we carried out affinity purification followed by mass spectrometry of endogenously tagged *T. thermophila* histones H2A, H2B and variant Hv1.We identified a set of interacting proteins shared among the three analyzed histones that includes the FACT-complex, as well as H2A- or Hv1-specific chaperones. We find that putative subunits of *T. thermophila* versions of SWR- and INO80-complexes, as well as transcription-related histone chaperone Spt6^{Tt} specifically copurify with Hv1. We also identified importin β6 and the *T. thermophila* ortholog of nucleoplasmin 1 (cNpl1^{Tt}) as H2A–H2B interacting partners. Our results further implicate Poly [ADP-ribose] polymerases in histone metabolism. Molecular evolutionary analysis, reciprocal affinity purification coupled to mass spectrometry experiments, and indirect immunofluorescence studies using endogenously tagged Spt16^{Tt} (FACT-complex subunit), cNpl1^{Tt}, and PARP6^{Tt} underscore the validity of our approach and offer mechanistic insights. Our results reveal a highly conserved regulatory network for H2A (Hv1)–H2B concerning their nuclear import and assembly into chromatin.

Key words: **histone, FACT, histone chaperone, histone H2A, histone H2B.**

Article

Introduction

The eukaryotic genome is packaged in the form of a nucleoprotein complex called chromatin. The primary repeating unit of chromatin is the nucleosome which is formed when ~146 bp of DNA is wrapped around a core histone octamer consisting of two histone H2A–H2B heterodimers and one H3/H4 tetramer (Luger et al. 1997). Chromatin structure influences all DNA-mediated cellular processes, including gene transcription, replication, recombination, and repair (reviewed by Venkatesh and Workman [2015]). Histones carry posttranslational modifications which have important roles in gene expression regulation. For example, an enrichment of histone H3 trimethylated at lysine (K) 9 or K27 (H3K9me3 or H3K27me3) has been associated with heterochromatic chromatin regions whereas H3K4me3 and H3K36me3 posttranslational modifications have been linked with transcriptional activity (Allshire and Madhani 2018).

Variant histones have been described across species that differ in their primary amino acid sequences (Talbert et al. 2012). Canonical histones are only expressed during S-phase and are deposited onto chromatin in a DNA replication-dependent (RD) manner, whereas variants are expressed throughout the cell cycle and are deposited onto chromatin in a replication-independent (RI) manner (Mendiratta et al. 2018). Interestingly, RI variants, including H3.3 and H2A.Z for

histones H3 and H2A, respectively, have been described to have nonrandom distribution along the chromatin. For example, H3.3 is enriched in the euchromatic regions associated with transcriptionally active genes (Goldberg et al. 2010; Ray-Gallet et al. 2011). Genome-wide studies have indicated the enrichment of H2A.Z/H3.3 double variants within regions of high chromatin accessibility such as active promoters, enhancers, and insulator regions (Jin et al. 2009), highlighting their role in gene expression regulation.

Chromatin assembly is a fundamental process that may affect a broad range of gene regulatory processes such as DNA repair, DNA replication, and progression through the cell cycle (Mendiratta et al. 2018). Protein factors known as histone chaperones are thought to have key roles in regulation of chromatin assembly (Grover et al. 2018). For example, chromatin assembly factor-1 and histone-regulator-A have been shown to mediate RD and RI chromatin assembly processes to deposit either H3–H4 or H3.3–H4 (Hoek and Stillman 2003; Tagami et al. 2004; Jullien et al. 2012), respectively, whereas the SWR-complex specifically targets H2A.Z–H2B onto chromatin (Gerhold et al. 2015). Histone chaperones have specific preferences for binding to either H3–H4 or H2A–H2B (Keck and Pemberton 2012; Grover et al. 2018). For example, nucleosome assembly protein 1 (Nap1) and nucleoplasmin 1 (Npm1) are both H2A/H2B-specific chaperones (Straube et al. 2010; Hammond et al. 2017) whereas antisilencing factor 1 (Asf1) is an H3/H4-specific chaperone (English et al. 2006; Mendiratta et al. 2018). It is currently unclear how chaperones target a certain histone variant to a specific genomic region and how mechanistically this task is achieved. Several previous studies have utilized functional proteomics approaches to identify and examine the role of histone-binding proteins (Tagami et al. 2004; Latreille et al. 2014; Hammond et al. 2017). However, considering the complexity of chromatin assembly and gene expression regulatory layers, it is conceivable that many yet to be identified chaperones might have roles in these processes. Comparative proteomics is a powerful tool that has been widely employed to study the evolution and functional conservation of proteins (Boekhorst et al. 2008; Lotan et al. 2014). However, the extent to which the role of previously identified histone chaperones is conserved across the eukaryotic species remains unexplored (Grover et al. 2018).

The unicellular ciliate protozoan *Tetrahymena thermophila* provides an excellent experimental system to study chromatin dynamics and identify new factors involved in these processes. The *T. thermophila* genome is amenable to tractable alterations enabling the endogenous tagging of genes of interest. Ciliates are considered evolutionarily divergent organisms (Orias et al. 2011; Gao et al. 2016) and are therefore well-suited to examine the functional conservation of known histone chaperones. The *T. thermophila* single cell features a physical separation of two structurally and functionally distinct chromatin states in the form of a germ-line diploid micronucleus (MIC) and a polyploid somatic macronucleus (MAC). Functionally, the MAC regulates gene expression whereas the MIC ensures stable genetic inheritance (Martindale et al. 1982). The two nuclei originate from the same zygotic nucleus during sexual development (conjugation) of the cell, and subsequently embark on unique developmental pathways leading toward distinct chromatin organization within each nucleus (Martindale et al. 1982). The alterations in the chromatin states, including DNA rearrangements and removal of internally eliminated sequences during *T. thermophila* development (Yao et al. 1984, 1990, 2003; Mochizuki and Gorovsky 2004) share similarities with epigenetic changes that occur to mammalian chromatin during development.

The *T. thermophila* genome encodes two major histone H2A genes (*HTA1* and *HTA2*) which at the protein level are nearly identical with only three amino acid differences in the central core region (Liu et al. 1996). Furthermore, neither *HTA1* nor *HTA2* alone is essential for *T. thermophila* vegetative growth suggesting that the function of the encoded proteins is redundant (Liu et al. 1996). However, the C-termini of the two proteins differ significantly from each other as H2A.1 (encoded by *HTA1*) has an additional five residues (Liu et al. 1996). These additional five residues include an SQ motif which is conserved across species (as in mammalian H2A.X) and provides a target site for phosphorylation by a specific protein kinase family (Song et al. 2007). The SQ motif phosphorylation has been shown to function in double-strand break repair during mitosis, meiosis, and amitosis in *T. thermophila* (Song et al. 2007). Thus, *T. thermophila* H2A.1 can be considered an H2A.x ortholog although it differs from mammals where the H2A.X histone variant is a quantitatively minor component (Rogakou et al. 1998). *Tetrahymena thermophila* H2B.1 and H2B.2, encoded by *HTB1* and *HTB2* respectively, are nonallelic variants of H2B and only differ at three positions. Similar to H2A, *T. thermophila* cells lacking either *HTB1* or *HTB2* alone are viable and do not exhibit any growth defects indicating the functional redundancy of H2Bs (Wang et al. 2009).

The *T. thermophila* H2A variant Hv1 (H2A.Z and Htz1 in humans and yeast, respectively), has been found to be essential for growth (Liu et al. 1996). Hv1 localizes to the transcriptionally active MAC during vegetative growth and is found in the MIC only during early conjugation events (Stargell et al. 1993), prior to the stage when MIC becomes transcriptionally active (Martindale et al. 1985). Thus, the localization patterns of Hv1 suggest a role in transcription regulation. The mechanistic details of how Hv1 is targeted to the MAC (and MIC during early conjugation) remain elusive.

In this study, we employed a functional proteomics workflow to examine the histone-interactome for the first time in *T. thermophila*. Affinity purifications of *T. thermophila* H2A.1 (*HTA1*), H2B.1 (*HTB1*), and Hv1 followed by mass spectrometry analysis (AP-MS) revealed both new histone-interacting factors as well as a set of chaperones that have been previously identified only in Opisthokonts, indicating the evolutionarily conserved histone metabolism regulatory networks. Specifically, we identified *T. thermophila* FACT-, SWR-, and INO80-complexes suggesting an ancient origin for these proteins. We carried out detailed molecular evolutionary analyses of several histone-interacting proteins which further reinforced the idea that dedicated chaperones arose very early

during eukaryotic evolution to regulate histone metabolism. We validated several of the identified interactions by reciprocal affinity purification coupled to mass spectrometry (AP-MS) analyses and indirect immunofluorescence (IF) studies.

Results

Identification of *T. thermophila* H2A/H2B-Interacting Proteome

We generated stable *T. thermophila* lines expressing H2A.1 (TTHERM_00790790) (H2A hereafter) and H2B.1 (TTHERM_00633360) (H2B hereafter) with a C-terminal FZZ epitope tag from their native MAC chromosomal loci. The FZZ epitope tag contains 2 protein A moieties and a 3xFLAG separated by a TEV cleavage site, permitting affinity purification of the fusion protein and analysis of the copurifying proteins by Western blotting and/or mass spectrometry. To accomplish this, we engineered constructs that included ~1 kb of DNA sequence upstream and downstream of the predicted stop codons of *HTA1* and *HTB1*. The engineered FZZ constructs (supplementary fig. 1, Supplementary Material online) were used to transform growing *T. thermophila* cells using biolistic transformation. Homologous recombination mediates the gene replacement of the wild type (WT) *HTA1* and *HTB1* loci by FZZ constructs (Cassidy-Hanley et al. 1997). The polyploid MAC divides amitotically and does not afford an equal segregation of alleles (reviewed by Karrer [2012]). Homozygocity in the polyploid MAC of the transformed cells can be achieved through "phenotypic assortment" (reviewed by Karrer [2012]). Western blotting analysis using anti-FLAG antibody demonstrated successful expression of the epitope-tagged proteins in whole-cell extracts (WCEs) from H2A- and H2B-FZZ-expressing strains, compared with the WCEs prepared from untagged control cells (fig. 1A, left panels). To test the possibility that the presence of the FZZ tag might interfere in the localization of the tagged histones, we carried out indirect IF analysis on H2A- and H2B-FZZ in growing *T. thermophila* cells. Previously, H2A and H2B have been shown to localize to both the MAC and MIC (Song et al. 2007; Wang et al. 2009). Our IF analysis indicated that H2A- and H2B-FZZ also localize to both the MAC and MIC (fig. 1A, right panels) supporting that the FZZ tag does not interfere with their function.

We performed affinity purification in biological replicates on H2A- and H2B-FZZ expressing strains. The recovery of the baits was confirmed by Western blotting using the affinity-purified material from either the untagged WT cells or H2A- and H2B-FZZ cells (fig. 1B). To define H2A/H2B protein–protein interaction (PPIs) networks, a gel-free liquid chromatography coupled to tandem mass spectrometry (LC–MS/MS) analysis was carried out using the affinity purified material. The mass spectrometry data were evaluated with SAINTexpress, which uses semiquantitative spectral counts for assigning a confidence value to individual PPIs (Teo et al. 2014). Application of SAINTexpress to the AP-MS data for two biological replicates of H2A- and H2B-FZZ affinity purifications from growing *T. thermophila* cells scored against numerous control AP-MS experiments revealed several interaction partners that pass the cutoff confidence value (Bayesian FDR $\leq$ 1%) (supplementary file 2, Supplementary Material online).

This analysis revealed that H2A- and H2B-FZZ copurify with 14 and 17 significant interacting partners, respectively (fig. 1C). Three interaction partners, TTHERM_00283330, TTHERM_00049080, and TTHERM_00726470, copurified with both H2A and H2B. TTHERM_00283330 and TTHERM_00049080 proteins are the orthologs of yeast Spt16 (suppressor of Ty 16; SUPT16H in humans) and Pob3 (Pol1-binding protein; SSRP1 in humans) subunits of the FACT-complex, respectively. The FACT-complex is a well-characterized transcriptional regulator that functions as an H2A–H2B dimer chaperone (Belotserkovskaya et al. 2003). The third H2A/H2B shared interacting partner TTHERM_00726470 is *T. thermophila* Poly [ADP-ribose] polymerase 2 ($PARP2^{Tt}$). PARPs are functionally diverse proteins with critical roles in chromatin architecture, mRNA processing and histone ADP-ribosylation (Hassa and Hottiger 2008).

We also identified TTHERM_00429890 as an interaction partner for H2A-FZZ. TTHERM_00429890 shares sequence similarity with a known human H2A/H2B chaperone proto-oncogene NPM1 (Okuwaki et al. 2001), suggesting a conserved histone-binding function for this protein. The remaining H2A-FZZ copurifying proteins include a *Tetrahymena*-specific protein TTHERM_00242240 which does not have an identifiable ortholog in any other organism, a DNA-binding AT-Hook domain protein, a VWA domain-containing protein, a MutS family protein which shares sequence similarity with yeast MSH6, a POZ domain protein, histones H2B, H3, H4, and two PARPs including $PARP6^{Tt}$ and $PARP3^{Tt}$ (fig. 1C).

The *T. thermophila* genome encodes at least 13 importin (imp) α- and 11 impβ-like proteins (Malone et al. 2008). Our SAINTexpress analysis indicated that H2B copurifies with TTHERM_00962200 which encodes an Impβ6 protein. We previously have shown that Impβ6 interacts with Asf1 and likely functions in the H3/H4 transport pathway (Garg et al. 2013). Among the interacting partners detected for H2B-FZZ, were three hypothetical proteins: TTHERM_00532520, TTHERM_00657290, and THERM_00648920. TTHERM_00532520 is a ciliate-specific protein with an ortholog in *Paramecium tetrarelia* without any recognizable domains, whereas TTHERM_00657290 appears to carry an SMC-N terminal domain suggesting a role in chromatin structural maintenance. THERM_00648920 is a predicted ~32 kDa *Tetrahymena*-specific protein which has several stretches of acidic residues similar to other histone chaperones, for example, NPMs. We named THERM_00648920 as "histone-interacting acidic protein 1" (Hiap1). Additionally, other notable H2B-FZZ copurifying proteins include a Basic Leucine Zipper Domain-transcription factor (bZIP1), an Alba2-domain DNA-binding protein, DExD/H box RNA helicase Drh29, Mak21, an apoptosis-antagonizing transcription factor AATF, an ARM-repeat protein, and an MutS family protein

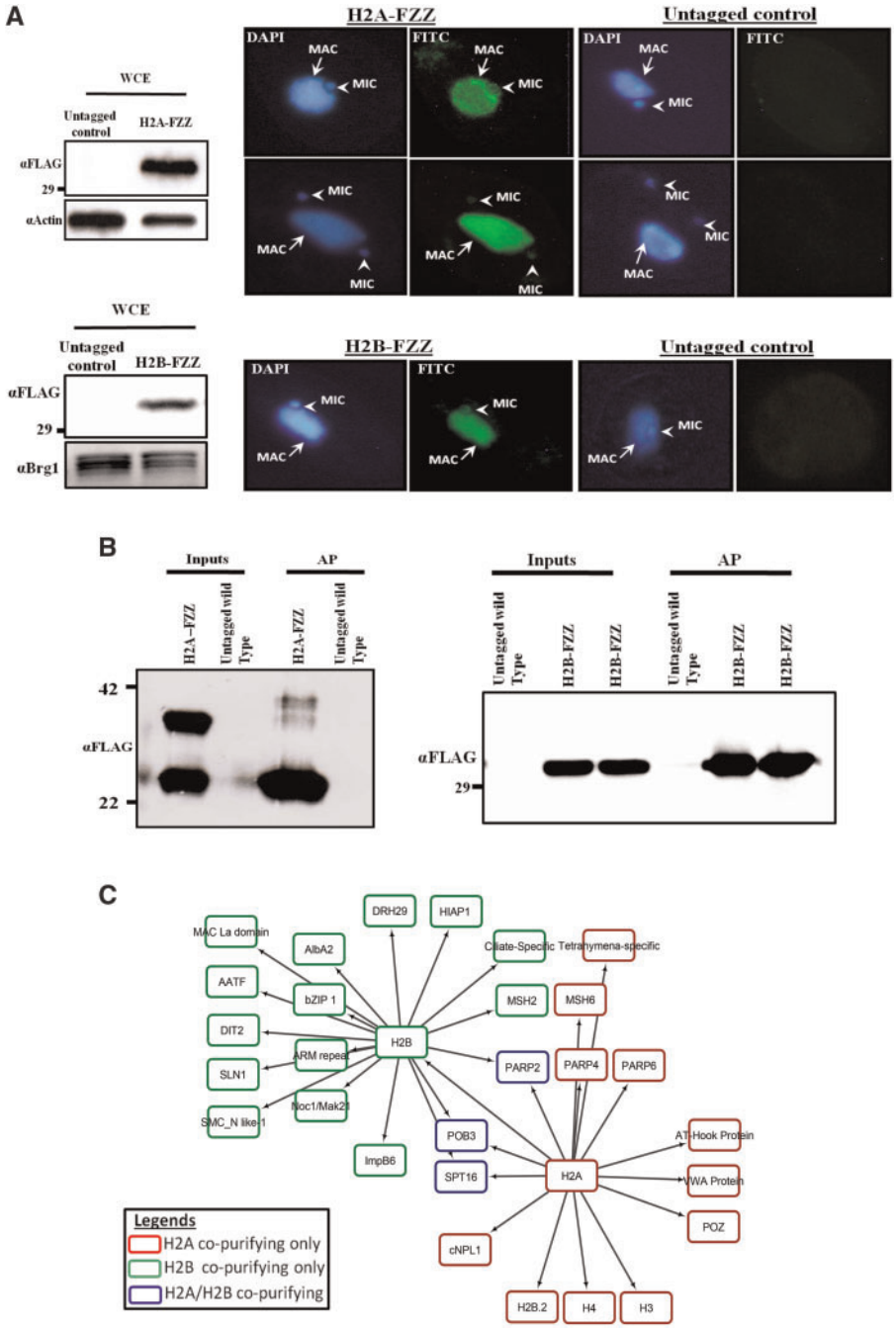


Fig. 1. H2A-FZZ and H2B-FZZ expression and affinity purification. (*A*) Left panels: Expression analysis of H2A-FZZ (H2A ~14.77 kDa + FZZ ~18 kDa) and H2B-FZZ in comparison to the untagged controls by Western blotting using WCEs. Blots were probed with anti-FLAG antibody for FZZ detection whereas anti-Actin and anti-Brg1 (146 kDa) were used as loading controls. Right panels: H2A- and H2B-FZZ localize to both MAC and MIC. Note: For H2A-FZZ IF images, the lower panel demonstrates dividing cells. DAPI was used to stain the nuclei and the position of the MAC and MIC is indicated with arrows and arrow-heads, respectively. (*B*) Western blotting analysis indicating the recovery of the affinity purified (AP) H2A-FZZ (left) and H2B-FZZ (right). The top panels were probed with anti-FLAG antibody to examine the recovery of the baits. No signal was detected in the WT. Anti-Actin and anti-Brg1 were used as loading controls. Two bands in the H2A-FZZ input likely represent dimers. (*C*) Network representation of H2A- and H2B-FZZ copurifying proteins. Node border legend is provided. The MS data were searched against the *Tetrahymena* Genome Database (www.ciliate.org; last accessed September 24, 2018) (TGD). Full-length protein sequences were retrieved from TGD and searched against yeast or human proteins to annotate them (see Materials and Methods for details).

MSH2. Yeast MSH2 and MSH6 proteins are known to interact with each other and have critical roles in DNA-mismatch repair (Studamire et al. 1998) (fig. 1C; supplementary file 2, Supplementary Material online, for details on all H2B interaction partners).

We utilized publicly available microarray expression data to compare gene expression profiles of H2A and H2B with those of the genes encoding their copurifying proteins. As expected, H2A, H2B, H3, and H4 cluster together due to their similar expression profiles in S-phase (supplementary fig. 2*A*

and *B*, Supplementary Material online). Spt16^{Tt} and Pob3^{Tt} also clustered with H2A and H2B, consistent with a role of the FACT-complex in histone metabolism. In addition, PARP6^{Tt}, PARP3^{Tt}, Poz, Msh2, Hiap1, and Impβ6 also exhibited similarities in their expression profiles with those of the H2A and H2B (supplementary fig. 2*A* and *B*, Supplementary Material online), suggesting functional linkage of these proteins with histones. The observation that putative *T. thermophila* FACT-complex subunits, PARPs, and NPM1-like proteins copurify with H2A–H2B suggests an evolutionarily conserved role of these proteins in histone metabolism. Considering their central role in a number of chromatin-related processes and relevance to human diseases we further characterized these proteins.

The FACT-Complex Is Conserved across Eukaryotes

The FACT-complex is a critical transcription regulator and an H2A/H2B chaperone (Mason and Struhl 2003; Hsieh et al. 2013). The copurification of the putative FACT-complex subunits with H2A- and H2B-FZZ in an evolutionarily divergent eukaryote highlights the conserved nature of its role in chromatin-related processes. However, evidence regarding the origin of the FACT-complex is currently lacking. To gauge the evolutionary history of the FACT-complex, we carried out extensive database searches and identified the putative orthologs of the FACT subunits, that is, Spt16 and Pob3, throughout the eukaryotic supergroups including the basal eukaryote *Giardia lamblia* (supplementary file 1, Supplementary Material online). This suggests that Spt16 and Pob3 were already present in the last eukaryotic common ancestor. Given that FACT subunits were likely present in the last eukaryotic common ancestor, we wanted to further examine their evolutionary patterns, and reconstructed the phylogenetic trees of Spt16 and Pob3. Clustering in the resulting phylogenetic trees (fig. 2*A*) appears highly similar to that of the eukaryotic classification system (Adl et al. 2012). Both proteins, that is, Spt16 and Pob3, follow nearly identical phylogenetic paths with a few minor exceptions. For example, the amoebozoan lineages corresponding to Pob3 form a monophyletic group below the metazoans whereas Spt16 amoebozoans take a basal position below the opisthokonts (fig. 2*A*). Such differences likely represent the isolated cases where lineage-specific functional constraints might have been operating on both proteins independently of each other. Nevertheless, similarities in the phylogenetic histories strongly suggest that both proteins together experienced strong purifying selection to retain their structural and functional features. To examine the selective constraints operating on the FACT-complex, we used nucleotide coding sequences of Spt16 and Pob3 from the representative lineages and carried out codon-based *Z*-test of selection by comparing synonymous and nonsynonymous variations. We found extensive synonymous variations that were considerably higher than nonsynonymous variations ($P < 0.001$) in all comparisons for both Spt16 and Pob3 indicating the presence of purifying selection (supplementary file 3, Supplementary Material online). Extensive silent variations that we observed at the nucleotide level also resulted in a subsequent overall decrease in codon usage bias (supplementary file 3, Supplementary Material online), consistent with the idea of strong functional constraints operating at the protein level. Previous high-throughput studies have reported the phosphorylation of human and mouse Spt16 and Pob3 at highly conserved serine residues (supplementary file 3, Supplementary Material online). Interestingly, we found that the serine residues in Spt16 are preferentially encoded by the codon UCU across all the taxa. For Pob3 serine residues, "AGC" is the preferred codon within Opisthokonts whereas UCU and AGU are preferentially used in plants and protist lineages (supplementary file 3, Supplementary Material online). These results indicate the strong purifying selection operating not only at the protein level to maintain the structural features but also by the usage of preferred codons for functionally important positions.

Spt16 contains a signature Spt16_domain (SMART accession: SM001286), an N-terminal lobe (SM001285), a peptidase (pfam: PF00557), and an Rtt106 domain (SM001287) which is also found in Pob3 (fig. 2*B*, left). The "peptidase" and Rtt106 domains are known to function as histone-binding modules (Stuwe et al. 2008; Zunder et al. 2012). We examined the structural features of the Spt16^{Tt} and Pob3^{Tt}. We aligned Spt16^{Tt} and Pob3^{Tt} against budding yeast and human homologs and observed that the domain organization in both proteins is highly conserved (fig. 2*B*, left). In fact, Spt16^{Tt} and Pob3^{Tt}, respectively, exhibit more than 30% and 20% sequence identities to their homologs both in the budding yeast and humans. Of note, Pob3 in tetrapods has gained a high-mobility group (HMG) domain whereas unicellular eukaryotes, for example, budding yeast, FACT-complex interact with an HMG protein Nhp6 to provide the same activity (Formosa et al. 2001). Ciliates and humans diverged ~1,781 Ma (Kumar et al. 2017), and such a degree of sequence and structural conservation points toward possible functional similarities that might exist among the distant homologs. To further investigate this possibility, we used the strategy described above to engineer *T. thermophila* cells stably expressing C-terminally epitope tagged Spt16^{Tt}-FZZ from its native chromosomal locus (fig. 2*B*, right). As shown in figure 2*B* (lower panel), Spt16^{Tt}-FZZ localizes to both the MAC and MIC in growing *T. thermophila* cells. Affinity purification on growing Spt16^{Tt}-FZZ strains and SAINTexpress analysis of the LC–MS/MS data confirmed the copurification of Pob3^{Tt} with Spt16^{Tt}-FZZ (supplementary file 2, Supplementary Material online). We also detected two subunits of RNA polymerase I and III (RNAP) Rpac1 and Rpa2, consistent with a role in transcription regulation. Additionally, a *T. thermophila*-specific TTHERM_01046850 protein also copurified with Spt16^{Tt}. TTHERM_01046850 encodes a predicted ~53 kDa protein and does not have any identifiable domains. We named this protein as "FACT-interacting mysterious protein 1" (Fimp1) (supplementary file 2, Supplementary Material online, for all Spt16^{Tt} interactions). Consistent with their copurification, Spt16^{Tt} and Pob3^{Tt} share nearly identical gene expression profiles. Similarly, Fimp1 also clusters along with the FACT-complex (supplementary fig. 3, Supplementary Material online). Further analysis will be required to understand the mechanistic details of Fimp1

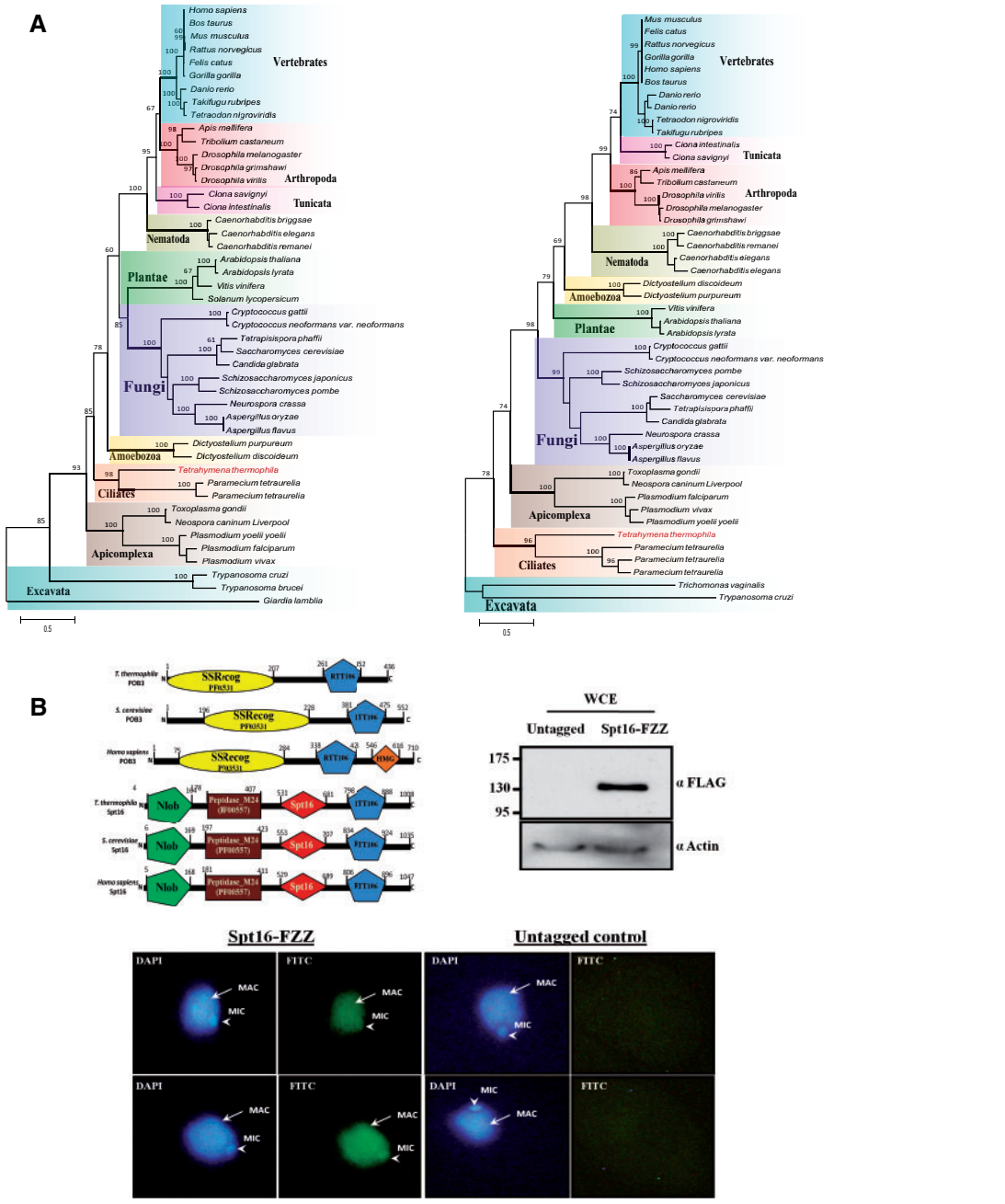
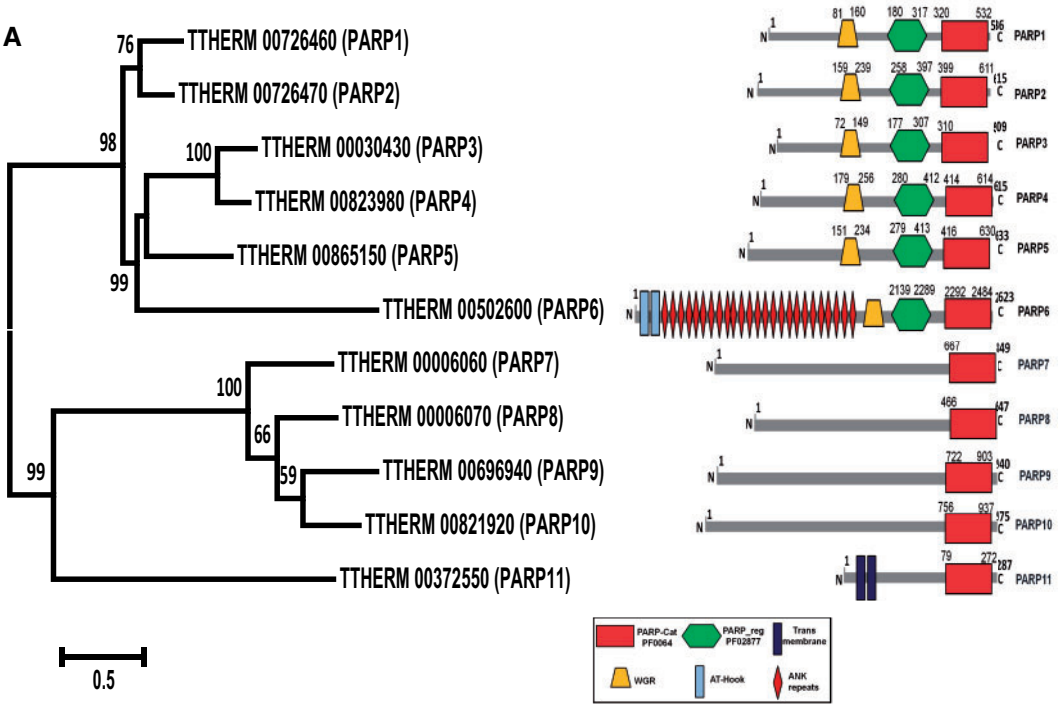


Fig. 2. Phylogenetic analysis of the FACT-complex Spt16 and Pob3 subunits. (*A*) Protein phylogenies representing the evolutionary patterns for Spt16 (left) and Pob3 (right) FACT-complex subunits under LG+G model of evolution. Numbers on the left side of each branch represent the confidence values based on 1,000 bootstrap replicas (only reported when at least $\geq$ 50%). Different taxonomic groups are highlighted in different colors. *Tetrahymena thermophila* is indicated in red. The scale bar shows the number of substitutions per site. (*B*) Left: Comparative domain analysis of *T. thermophila* Spt16^{Tt} and Pob3^{Tt} against human and budding yeast homologs. Right: Expression analysis of Spt16^{Tt}-FZZ (Spt16^{Tt} ~116 kDa + FZZ ~18 kDa) in comparison to the untagged controls by Western blotting using WCEs. Blot was probed with anti-FLAG antibody for FZZ detection, and anti-Actin was used as a loading control. Bottom panel: Indirect IF analysis of Spt16^{Tt}-FZZ. Spt16^{Tt}-FZZ localizes to both the MAC and the MIC, whereas no signal was detected in the untagged cells. DAPI was used to stain the nuclei. Arrows represent MAC, whereas arrow heads denote MIC.

interaction with the FACT-complex. We conclude that Spt16^{Tt} and Pob3^{Tt} constitute the *T. thermophila* FACT-complex with possible roles in histone H2A/H2B chaperoning and transcription regulation.

PARP Proteins in *T. thermophila*

The observation that certain PARPs copurified with histones (fig. 1C) prompted us to examine the full repertoire of PARP proteins in *T. thermophila*. Our query against the *T. thermophila* genome database using human PARP1 identified at least 11 proteins with a PARP-catalytic (PF0064) domain (fig. 3A, right). Multiple sequence alignment indicated that catalytic residues (HYE) within PARP-catalytic domains are highly conserved with the exception of PARPs7–9 where the third residue aspartic acid (E) has been mutated (fig. 3B). These observations suggest that at least some of these PARPs might



B

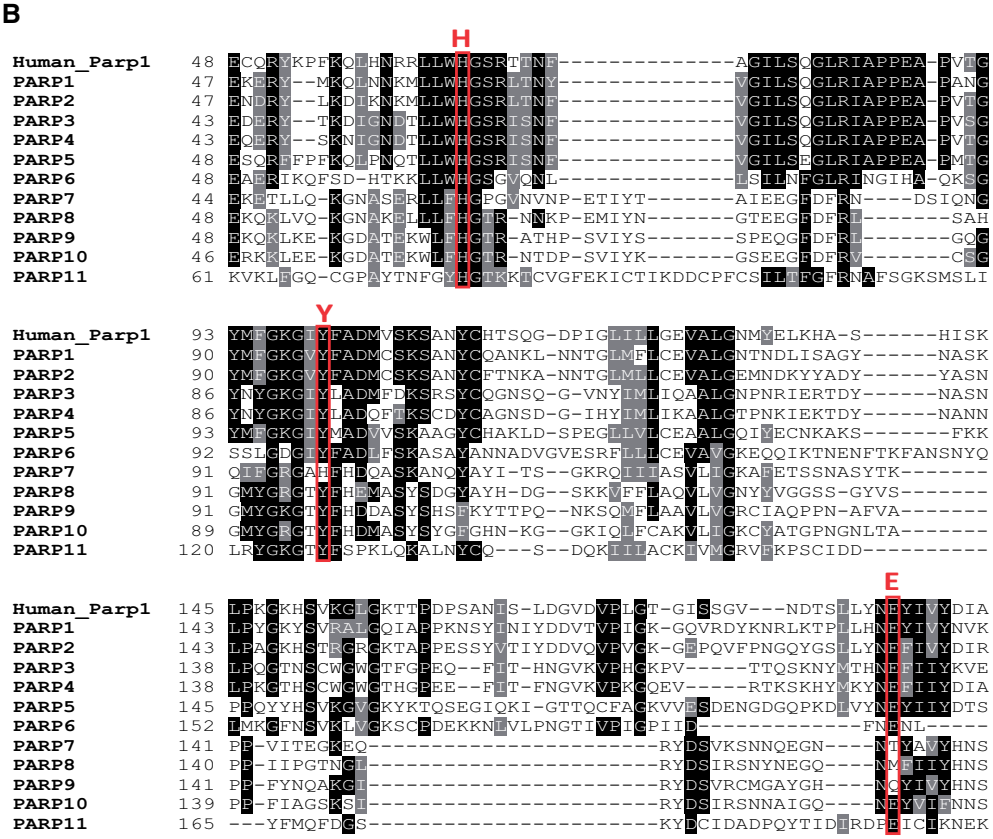


Fig. 3. Domain analysis of *Tetrahymena thermophila* PARP proteins. (*A*) Left: Protein phylogenetic analysis of putative PARPs using the identified PARP-catalytic domain sequences under LG+G model of evolution. *Tetrahymena thermophila* genome database accession numbers along with protein names are indicated. Tree topology represents ML estimations and confidence values are based on 1,000 bootstrap replicas (only reported when at least ≥50%). The scale bar indicates the number of substitutions per site. Right: Domain analysis of the *T. thermophila* PARPs. The analysis was carried out using the SMART database (see Materials and Methods) and numbers represent the amino acid positions for each identified domain. Domain legend is provided in the box. (*B*) Multiple sequence alignments of PARP-catalytic domains of *T. thermophila* PARPs. The human PARP1 catalytic domain is used as a reference to examine the conservation. The catalytic residues are highlighted as red boxes.

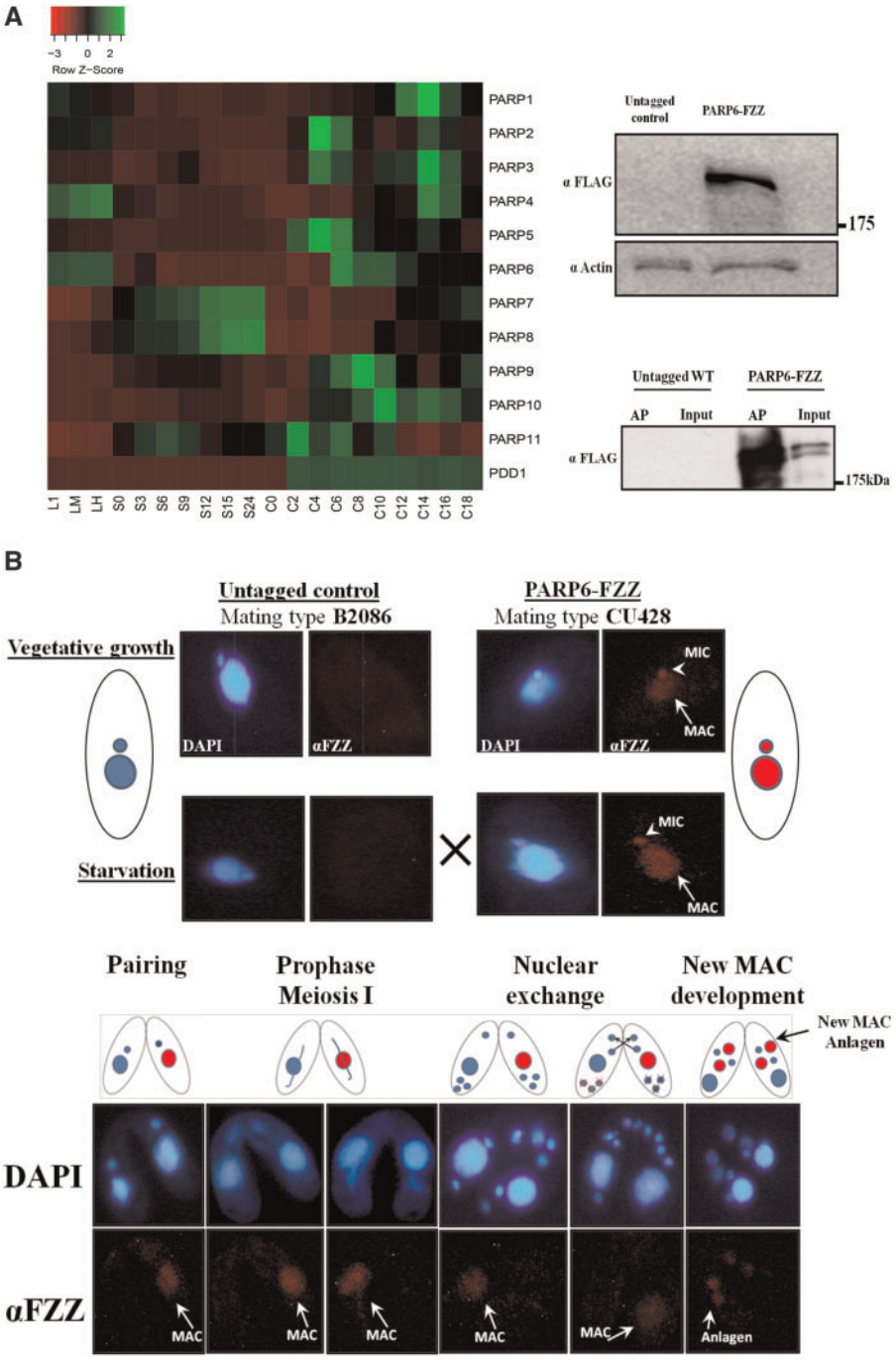


Fig. 4. Expression analysis of *Tetrahymena thermophila* PARP proteins and PARP6^{Tt} localization profile during development. (*A*) Left: Heat map representation of microarray expression values for PARP1-11^{Tt}. *Z*-scores were calculated across the rows for each PARP to examine its differential expression across growth, starvation, and various developmental stages. L1–LH represent growth phase, S0–24 represent starvation for 24 h, and C stands for conjugation where 0–18 denote hours postmixing the different mating types. PDD1 is used as a conjugation-specific marker. Right: Top, Expression analysis of PARP6^{Tt}-FZZ (PARP6^{Tt} ~300 kDa + FZZ ~18 kDa) in comparison to the untagged controls by Western blotting using WCEs. Blot was probed with anti-FLAG antibody for FZZ detection, and anti-Actin was used as a loading control. Bottom panel: Western blotting analysis indicating the recovery of the affinity purified PARP6^{Tt}-FZZ in comparison to a control purification. The blot was probed with anti-FLAG. (*B*) PARP6^{Tt}-FZZ localizes to both MAC and MIC during vegetative growth and starvation. PARP6^{Tt}-FZZ cells were mated with untagged WT cells of different mating type. Nuclear events are depicted above the images taken for conjugating cells during various developmental stages. DAPI was used to stain the nuclei. PARP6^{Tt}-FZZ localizes to only MAC during early conjugation events. At the onset of new MAC development (anlagen), PARP6^{Tt}-FZZ loses signal in the parental MAC and is found within developing MACs. Note: The signal observed in both mating pairs (PARP6^{Tt}-FZZ and controls) at the anlagen stage indicates mixing of cellular contents between the pairing cells. CU428 and B2086 refer to the stock strain numbers of the different mating types, as adopted from the *Tetrahymena* stock center Cornell University (http://tetrahymena.vet.cornell.edu/; last accessed September 24, 2018).

be catalytically active. Based on the domain architecture and phylogenetic analysis (fig. 3A, left), we assigned these putative PARPs into subgroups and established a systematic nomenclature. Notably, PARP1 to PARP5 appear closely related to each other consistent with their similar domain architecture. Expression analysis using publicly available RNA-seq and microarray data showed that the *T. thermophila* PARPs have distinct expression profiles (supplementary fig. 4*A* and *B*, Supplementary Material online). Most of the PARPs are weakly expressed during vegetative growth with the exception of PARP4 and PARP6 (fig. 4*A*). PARP7 and PARP8 are highly expressed during starvation whereas PARP1, 2, and 4 have relatively higher expression levels during late developmental stages (14–16 h postmixing) (fig. 4*A*; supplementary fig. 4*A* and *B*, Supplementary Material online). This suggests that PARP expression levels are tightly coordinated during growth and various developmental stages. In vertebrates PARP proteins, including human PARP1, also contain PADR (PF08063) and zinc finger (zf)–PARP domains (PF00645). The zf–PARP domain binds to DNA, whereas the function of the PADR1 domain remains unknown (Citarelli et al. 2010). Interestingly, none of the *T. thermophila* putative PARPs carries any PADR1 and zf–PARP domains. Instead, we identified six additional proteins carrying PADR1 and zf–PARP domains (supplementary fig. 4C, Supplementary Material online). Thus, *T. thermophila* PARPs might require additional protein factors for their proper functioning.

Among the identified PARP proteins, PARP6^{Tt} (THERM_00502600), which copurified with H2A, piqued our interest due to its unique domain architecture. PARP6^{Tt} contains 25 tandem ankyrin repeats (ANK) as well as two DNA binding AT-hook domains in addition to the PARP-catalytic and PARP-regulatory (PF02877) domains (fig. 3A). This domain organization is unique to Amoebozoa (*Dictyostelium*), Opisthokonta (fungi), and Chromalveolates (ciliates) and has been categorized as the PARP1 subfamily (Citarelli et al. 2010). Interestingly, human PARP5a, b (known as Tankyrase 1 and 2, respectively), also contain tandem ANK repeats as well as a PARP-catalytic domain but lack PARP-regulatory and AT-hook domains. Tankyrase 1 and 2 function in maintenance of telomeres (Chiang et al. 2008). To gain functional insights, we generated a strain of *T. thermophila* stably expressing PARP6^{Tt}-FZZ from its native MAC locus (fig. 4A, right). We performed AP-MS analysis on growing cells to investigate the PARP6^{Tt}-interacting proteins. The recovery of the bait was examined using Western blotting analysis (fig. 4A, right). The SAINTexpress analysis revealed nine high-confidence PARP6^{Tt}-FZZ copurifying proteins including histone H2A. Additionally, H3 and ribosomal proteins were identified as PARP6^{Tt}-FZZ copurifying partners (see supplementary file 2, Supplementary Material online, for details). The copurification of H2A with PARP6^{Tt}-FZZ reciprocally verifies the interaction between the two proteins. PARP6^{Tt} and histones H2A and H3 cluster together based on their gene expression profiles further indicating a role of PARP6^{Tt} in histone metabolism (supplementary fig. 4*D*, Supplementary Material online).

PARP6^{Tt} is expressed throughout the *T. thermophila* life cycle with relatively low expression levels during early conjugation (1–2 h postmixing the cells) as examined by using previously published expression data (fig. 4*A*, left) (Miao et al. 2009; Xiong et al. 2012). The expression levels increase between 6 and 8 h postmixing, a time of new MAC development. We performed IF staining in growing and conjugating *T. thermophila* cells to examine the PARP6^{Tt}-FZZ localization during development. PARP6^{Tt}-FZZ localized to both the MAC and MIC in growing and starved *T. thermophila* (fig. 4*B*). Interestingly, we observed that PARP6^{Tt}-FZZ loses signal in the MIC and localizes exclusively to the MAC during conjugation when the cells have formed pairs (fig. 4*B*). More specifically it localizes to the parental MAC during early nuclear developmental stages including meiosis before switching to the anlagen which corresponds to midway through development (fig. 4*B*). The localization of PARP6^{Tt}-FZZ in the parental MAC is lost at the onset of MAC development, a stage where the two anterior nuclei (the anlagen) have become visibly larger than the posterior nuclei (fig. 4*B*). This pattern of localization is strikingly similar to that of Ibd1 (Interactive Bromo-Domain protein 1) protein which we recently reported to function as a recruitment hub for various transcription regulators and chromatin remodeling complexes (Saettone et al. 2018). The PARP6^{Tt} subcellular localization appears to correlate with transcriptional activity during nuclear development. Further studies will be needed to explore the role of PARP6^{Tt} in transcription regulation and histone metabolism.

Nucleoplasmin Has an Ancient Origin

NPM-family proteins are histone H2A/H2B chaperones with critical roles in various cellular processes (Box et al. 2016). NPM-family proteins have been linked to a number of human diseases, including acute myeloid leukemia, and are the subject of anticancer drug development (Box et al. 2016). Previous work has shown that among vertebrates the NPM-family has greatly diversified giving rise to three members (NPM1–3) whereas invertebrates, such as *Drosophila*, contain only a single Npm-like protein (NLP) (Eirín-López et al. 2006). To date, no orthologs have been detected in *Arabidopsis thaliana*, *Saccharomyces cerevisiae*, or *Caenorhabditis elegans*. Little is known however about the evolution and origin of NPM proteins, and as such they have not been studied in unicellular model organisms. Deciphering the evolutionary history often provides meaningful insights into protein function. To trace their evolutionary origin, we carried out database searches and identified putative NPM homologs throughout the basal unicellular eukaryotes, including chromalveolates and excavates (supplementary file 1, Supplementary Material online). We reconstructed a protein phylogeny using the identified homologs and found that these proteins have a monophyletic origin and share a common ancestry (fig. 5A). Importantly, the identification of NPM homologs in the earliest branching eukaryotes, such as kinetoplastids, confirms an ancient origin of this protein family.

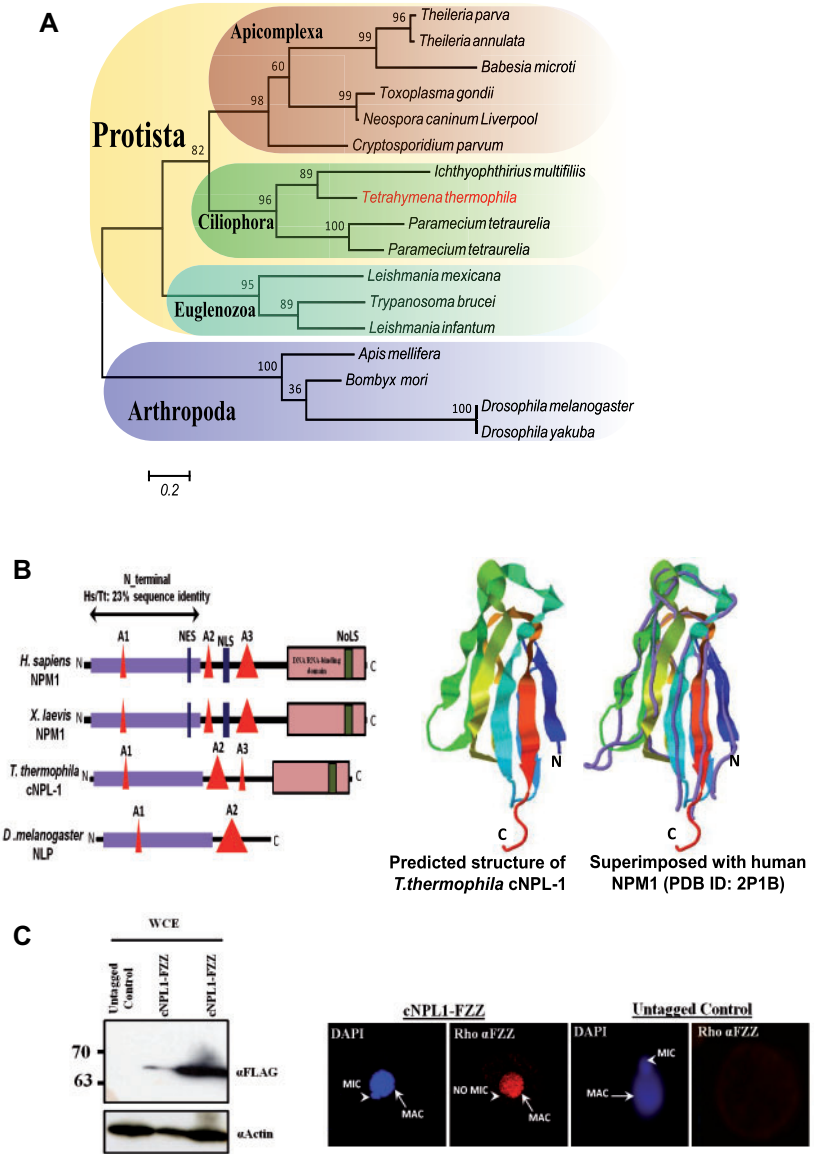


Fig. 5. Phylogenetic relationship among NPM-family proteins. (*A*) Protein phylogeny of NPM-family members in Protista under LG+G model of evolution. Different taxonomic groups are highlighted in colors. Arthropoda NPMs are used to represent the metazoan sequences. Tree topology represents the ML estimations based on 1,000 bootstrap replicas (confidence value only reported when at least ≥50%). The scale bar indicates the number of substitutions per site. (*B*) Left: Domain organization of cNpl1 in comparison to human and *Xenopus laevis* NPM1 proteins and *Drosophila melanogaster* NLP. "A" represents acidic stretches shown in red triangles, and NES and NLS stand for nuclear export and import signals. Nucleolar localization signal is denoted as NoLS. NPM core N-terminal domain (PF03066) is shown in light blue, and the C-terminal region is shown in red accent color. Note: cNpl1^{Tt} NoLS was predicted using the "NOD" web server (http://www.compbio.dundee.ac.uk/www-nod/index.jsp; last accessed September 24, 2018). Right: Cartoon diagram shows the predicted structure of the cNpl1 core domain in rainbow color. The predicted cNpl1^{Tt} structure shown in rainbow color was superimposed with the human NPM1 crystal structure (PDB ID: 2P1B) depicted in violet backbone format. N- and C-termini are indicated. (*C*) Left: Expression analysis of cNpl1^{Tt}-FZZ (cNPL1 ~40 kDa + FZZ ~18 kDa) in comparison to the untagged controls by Western blotting using WCEs. Blot was probed with anti-FLAG antibody for FZZ detection, whereas anti-Actin was used as a loading control. Right: Indirect IF analysis of cNpl1^{Tt}-FZZ. cNpl1^{Tt} primarily localizes to MAC. No signal was detected in the untagged control cells. DAPI was used to stain the nuclei. Arrows represent MAC, whereas arrow heads denote MIC.

Drosophila NLP (dNLP) also binds H2A/H2B dimers and assembles histone octamers (Namboodiri et al. 2003), suggesting functional conservation among distantly related family members. To gain functional insights, we compared the structural features of the putative *T. thermophila* Npm1 with those of human NPMs and dNLP. We observed that *T. thermophila* Npm1 domain organization is highly conserved and nearly identical to that of human NPM1 (fig. 5*B*, left). In fact, the *T. thermophila* Npm1 predicted N-terminal core domain can be structurally superimposed to that of the human NPM1 (fig. 5*B*, right). We named the putative *T. thermophila* homolog as conserved nucleoplasmin-like 1 (cNpl1). We engineered *T. thermophila* cell lines stably expressing cNPL1-FZZ from its native chromosomal locus. The expression of the tagged protein was examined by Western blotting (fig. 5*C*, left) and AP-MS experiments which

successfully recovered the bait (not shown) without any other significant interaction partners (see Discussion). IF analysis showed that cNpl1^{Tt} primarily localizes to the transcriptionally active MAC (fig. 5C, right), consistent with known roles of human NPM1 in transcription- and chromatin-related processes. We conclude that NPMs are a structurally/functionally conserved family of proteins which arose very early during the eukaryotic diversification.

Identification of *T. thermophila* Hv1-Interacting Proteome

We next focused on delineating the PPIs of transcription associated histone H2A variant Hv1 (H2A.Z in humans) in *T. thermophila*. We utilized our above described strategy to generate *T. thermophila* strains stably expressing Hv1-FZZ from their native MAC locus. The expression of the tagged protein was monitored by Western blotting analysis using WCEs prepared from Hv1-FZZ expressing cells in comparison to the untagged control cell lysates (fig. 6A, left). Hv1 has previously been reported to exclusively localize to the MAC during growth (Stargell et al. 1993). Our IF analysis of the Hv1-FZZ expressing cells showed an exclusive MAC signal indicating that the FZZ tag does not interfere with the protein localization (fig. 6B).

We subjected the Hv1-FZZ expressing cells to our AP-MS pipeline. Recovery of the bait was monitored by Western blotting (fig. 6A, right). SAINTexpress analysis of the LC–MS/MS data revealed that Hv1 copurifies with 106 significant interacting partners (BFDR $\leq$ 1%). We annotated these hits either by homology searches against the *S. cerevisiae* and human genomes or by using *T. thermophila* genome database annotations (supplementary file 2, Supplementary Material online, for annotations and conservation of interaction data; fig. 6C). ATP-dependent chromatin-remodeling complexes including SWR- and INO80-complexes are known to antagonistically modulate H2A.Z (Htz1 in yeast) dynamics. The SWR-C is specialized to deposit H2A.Z onto chromatin (Krogan et al. 2003; Kobor et al. 2004) whereas INO80-C mediates the reverse of this reaction (Papamichos-Chronakis et al. 2011), mainly at nonpromoter sites (reviewed by Gerhold and Gasser [2014]). Both the SWR-C and INO80-C have several shared as well as distinct subunits (reviewed by Gerhold and Gasser [2014]). Interestingly, SAINTexpress analysis of the Hv1-FZZ AP-MS data revealed the copurification of a set of proteins that based on similarity to *S. cerevisiae* orthologs comprise the putative subunits of *T. thermophila* INO80-C and SWR1-C. The identified INO80-C putative subunits include Arp8, Actin1 (also shared with SWR-C), Yuh1, and Ino80 (fig. 6C). In addition, we also identified the RuvB1 (also shared with SWR-C) and Ies2 subunits of the INO80-C, albeit at a slightly relaxed SAINTexpress value (BFDR $\leq$ 3%). We have recently purified *T. thermophila* SWR-C via Swc4-FZZ and identified at least 12 subunits (Saettone et al. 2018). In addition to Actin1 and RuvB1, SAINTexpress identified Swr1, Swc2, and Arp5 subunits of SWR-C as high confidence interacting proteins (fig. 6C). These data indicate that Hv1 deposition and eviction from the chromatin are tightly regulated by a highly conserved network of chromatin-remodeling complexes.

Other high-confidence Hv1 copurifying proteins with chromatin-related functions (inferred by sequence similarity to proteins in yeast and humans) could be broadly divided into four groups: 1) putative transcription and chromatin assembly regulators including, Spt16 and Pob3 (FACT-complex), Spt6, Cys2-His2 zf transcription factor ZAP1, TAF6, HMG protein Ixr1, transcription factors bZIP1 and bZIP2; 2) chromatin remodeling SWI/SNF complex subunits Swi3 and Snf12; 3) PARP proteins including PARP1, PARP2, and PARP5; and 4) proteins with various DNA- and RNA-related functions such as putative Alba2 DNA-binding protein, RNA-helicases, and topoisomerases (fig. 6C). Furthermore, we also identified a POZ-domain protein, Hiap1 and 8 additional *Tetrahymena*-specific hypothetical proteins without any recognizable domains. We named these proteins as "hypothetical histone copurifying proteins (HHCP1–8)" (fig. 6C) (see supplementary file 2, Supplementary Material online, for conserved and novel interactions).

We clustered the Hv1-FZZ copurifying proteins based on their gene expression profiles (supplementary fig. 5, Supplementary Material online). Our analysis suggests that proteins with key roles in histone metabolism, such as histone chaperones, share highly similar expression profiles and cluster together with Hv1 whereas factors with diverse functions (as inferred by similarities with yeast or human proteins) such as RNA-helicases, topoisomerases, and kinases are less likely to have expression patterns comparable with those of the histones (supplementary fig. 5, Supplementary Material online). Notably, consistent with their known role(s) in histone metabolism, INO80-C, SWR1-C, FACT-complex, Spt6, and SWI/SNF-complex subunits cluster together with Hv1 due to their very similar gene expression profiles, further reinforcing the idea that these proteins are functionally conserved in *T. thermophila*. We conclude that variant Hv1 in *T. thermophila* forms several functional links that might influence the transcriptional landscape of the cell, and furthermore, Hv1 distribution along the chromatin is regulated via a highly conserved network of chaperones.

Discussion

Although the deposition complexes for histones H2A/H2B and H2A variant H2A.Z have been identified (Zhang et al. 2017), information regarding the histone chaperoning network(s) outside of Opisthokonta remains limited. Considering the complexity of the histone deposition pathways, new factors are likely to be found to have key roles in these processes. *Tetrahymena thermophila* is an evolutionarily divergent unicellular eukaryote and is particularly suitable to study histone dynamics (Orias et al. 2011; Gao et al. 2016). In fact, initial clues regarding the transcription-related role(s) of H2A variants emerged from *T. thermophila* following the observations that Hv1 resides within the transcriptionally active nuclei (Martindale et al. 1985; Stargell et al. 1993). As per our ongoing efforts to understand the histone deposition pathways, here we report the first comprehensive

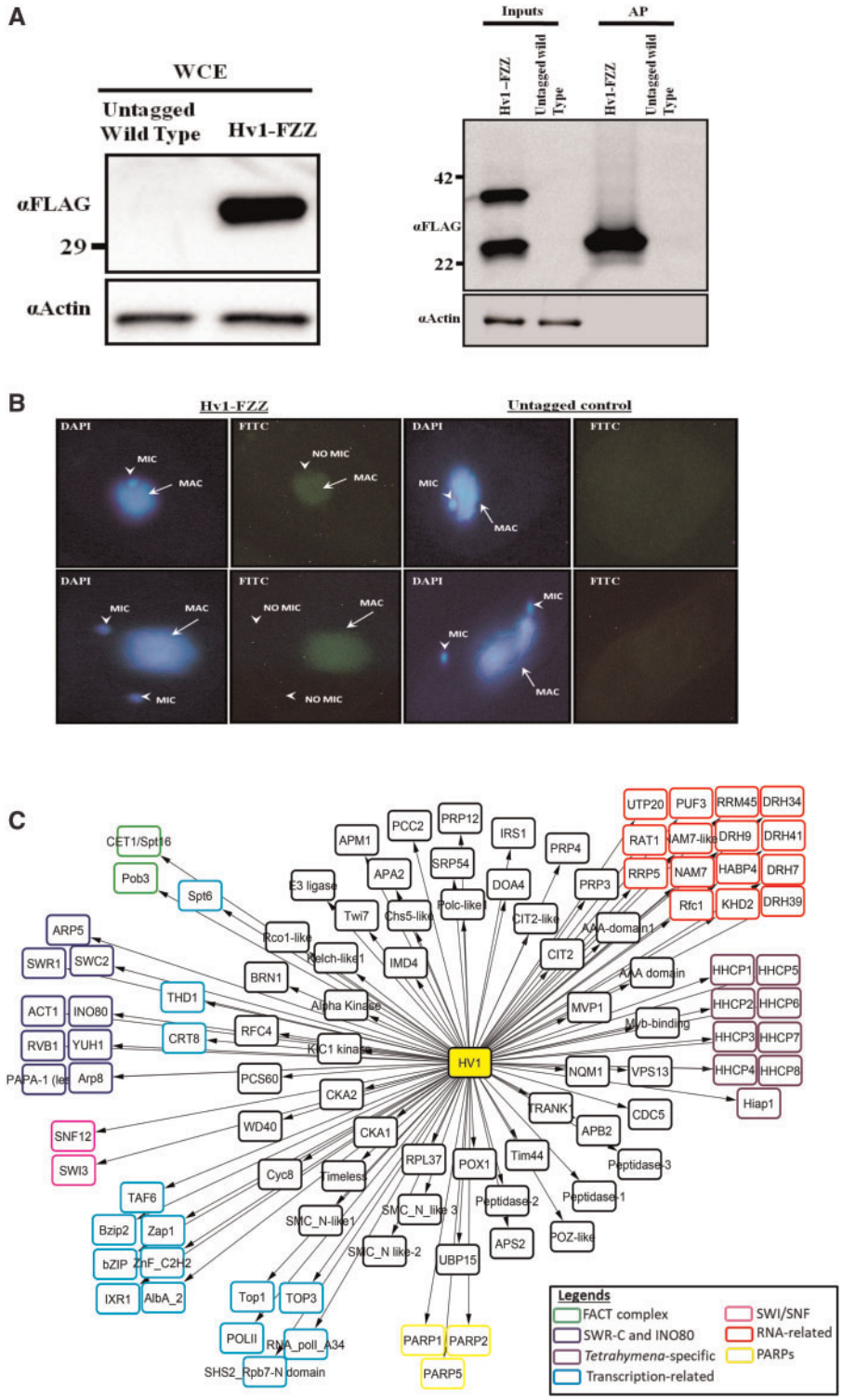


Fig. 6. Hv1-FZZ expression and affinity purification. (*A*) Left: Expression analysis of Hv1-FZZ (H2A ~15 kDa + FZZ ~18 kDa) in comparison to the untagged controls by Western blotting using WCEs. Blot was probed with anti-FLAG antibody for FZZ detection, whereas anti-Actin was used as a loading control. Right: Western blotting analysis indicating the recovery of the affinity purified (AP) Hv1-FZZ. The blot was probed with the indicated antibodies. No signal was detected in the WT lanes. Note: Two bands in the Hv1-FZZ input lane could represent dimerized histones. (*B*) Indirect IF analysis of Hv1-FZZ. Hv1-FZZ exclusively localized to MAC only during growth. The lower panel indicates dividing cells. No signal was detected in the untagged control cells. DAPI was used to stain the nuclei. Arrows represent MAC, whereas arrow heads denote MIC. (*C*) Network view of Hv1-FZZ PPIs. Bait node is shown in yellow. Prey node borders are colored according to their putative functions or protein complexes. Network legend is provided in the box.

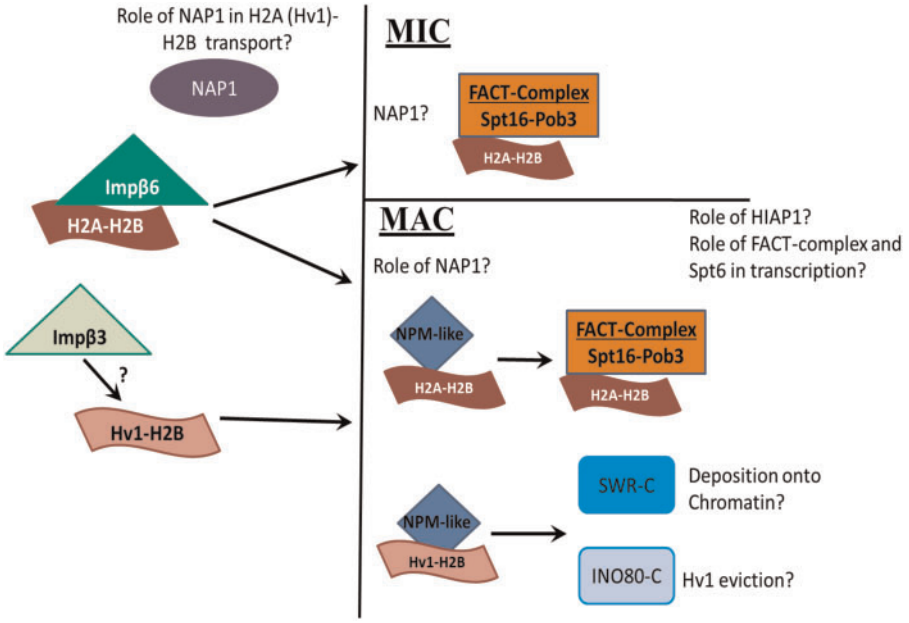


Fig. 7. Model for H2A (Hv1)–H2B nuclear transport in *Tetrahymena thermophila*.

PPI network for H2A, its variant Hv1 and H2B in *T. thermophila*.

Ancient Histone Chaperones

An interesting outcome of our work is that *T. thermophila* histones H2A (Hv1)/H2B are connected to a network of highly conserved chaperones and karyopherins. We have previously reported that Impβ6 physically interacts with Asf1 both of which localize to both MAC and MIC with a significantly stronger signal in the MIC indicating that Impβ6 functions in the H3/H4 transport pathway (Garg et al. 2013). The copurification of Impβ6 with H2B highlights the idea that it might be a more generalized karyopherin in *T. thermophila* for core histone transport pathways. It will be important to test this hypothesis by depleting Impβ6 and determining whether core histone can enter the MAC or MIC. Apart from cNpl1 which was found exclusively in the MAC, most of the H2A and H2B interacting partners that we characterized in this work localized to both the MAC and MIC. We expect RD histone-binding proteins to be found in the MAC and MIC as core histones are found within both nuclei (Song et al. 2007; Wang et al. 2009). The variant Hv1 is known to have nuclear-specific functions (Martindale et al. 1985; Stargell et al. 1993). We found that another karyopherin Impβ3 (TTHERM_00550700) copurified with Hv1 (though it fell below our stringent confidence threshold), and like Hv1, it localizes to MAC only (supplementary fig. 6, Supplementary Material online) consistent with a functional link between the two proteins. We suggest that transport of *T. thermophila* H2A (Hv1)/H2B to the nuclei and their subsequent assembly onto chromatin is mediated by an interplay among conserved karyopherins, histone chaperones, and chromatin-remodeling complexes (fig. 7), consistent with what has been proposed in humans and yeast. It will be important to determine the complete PPI networks for Impβ6 and Impβ3 whether by AP-MS or orthogonal methods such as Bio-ID. Future work should focus on understanding the nuclear-specific, replication-independent, chromatin assembly pathways and the role of chaperones, such as cNpl1, in these processes.

Numerous chaperones such as NASP, NPMs, and yeast Asf1 possess long acidic stretches, consistent with their potential to bind basic histones (reviewed by De Koning et al. [2007]). Hiap1^{Tt} also possesses several acidic stretches with an overall net negative charge (not shown), suggesting a possibility to function as a histone-binding protein. We suggest that Hiap1 functions as an H2A/H2B chaperone in *T. thermophila*. It is also worthwhile to note here that the *T. thermophila* ortholog of Nap1 also copurified with H2B and Hv1 (though it fell below our high-confidence threshold). Nap1 is a histone chaperone with a known function in H2A/H2B transport (Mosammaparast et al. 2002). Further work beyond the scope of this report will be required to examine the role of *T. thermophila* Nap1 and Hiap1 proteins in H2A/H2B metabolism. It will be important to express Hiap1 as a recombinant protein and examine whether it binds histones.

We have previously reported that histone chaperones including Asf1 and NASP are highly conserved throughout evolution (Nabeel-Shah et al. 2014), likely representing innovations to specifically regulate eukaryotic H3/H4 dynamics. Our present study has highlighted several aspects regarding the conserved nature of chromatin-remodeling and H2A/H2B assembly complexes. The FACT-complex is of particular interest due to its important roles in chromatin- and transcription-related processes. FACT is a histone chaperone and facilitates transcription elongation by colocalizing with RNAPII (Mason and Struhl 2003). Our evolutionary analysis indicated that FACT was already present in the last common ancestor of all eukaryotes indicating its functional importance. The similarities between the FACT evolutionary profile and the species phylogeny highlight the role of histone chaperones in eukaryotic evolution. This hypothesis is consistent

with previous work indicating that chromatin architectural HMG protein (González-Romero et al. 2015), histones (Eirín-López et al. 2012) and their chaperones including Asf1, NASP (Nabeel-Shah et al. 2014), NPMs (Eirín-López et al. 2006; Frehlick et al. 2007) might have played critical roles during eukaryotic evolution. Previous work has shown that *T. thermophila* Spt16^{Tt} associates with transcriptionally active MAC chromatin in vitro (Fujiu and Numata 2004). Consistent with a role in transcription, we also found that Spt16^{Tt} stably interacts with RNA polymerase subunits and localizes to the MAC. Spt16^{Tt} localization to MIC likely represents transcription-independent function(s) of the FACT-complex. Consistent with this hypothesis, FACT also functions in an array of processes including DNA replication and repair (Charles Richard et al. 2016; Yang et al. 2016; Kurat et al. 2017).

The *T. thermophila* genome encodes TTHERM_00216040 which shares sequence similarity to yeast HMG domain-containing protein Nhp6. We did not recover any HMG protein to copurify with Spt16^{Tt} (when enforcing an FDR cut-off of 1%). Our comparative analysis indicated that Pob3 within vertebrates, arthropods, tunicates, and plants carries an HMG domain whereas lineages representing fungi, amoebazoa, ciliates, apicomplexa, and excavatas do not possess this domain. We suggest that HMG was not present in the ancestral FACT-complex and was later acquired to meet the demands of complex regulatory layers of chromatin.

Human NPM1 is known to function in an array of processes, including histone chaperoning, chromatin remodeling, transcription regulation, genome stability, apoptosis, and embryogenesis (Okuwaki et al. 2001; Grisendi et al. 2005; Swaminathan et al. 2005; Box et al. 2016). Owing to its loss in widely studied eukaryotic microbial model organisms (e.g., *S. cerevisiae*), previous studies have been restricted to cultured cells. Furthermore, earlier attempts to decipher the evolutionary history of the NPMs have been limited to metazoans (Eirín-López et al. 2006). Our finding that cNpl1^{Tt} copurifies with H2A in *T. thermophila* combined with the observations that NPMs are highly conserved throughout the basal eukaryotes paves the way to study their function in easily tractable eukaryotic model organisms. Human NPM1 is thought to have key roles in cell cycle regulation (Zhao et al. 2015; Pfister and D'Mello 2016). *Tetrahymena thermophila* cells lacking Cyc2 and Cyc17 are arrested at early crescent (∼2–3.5 h postmixing) and diakinesis-like metaphase I (∼5 h post-meiotic induction) stages of meiosis, respectively (Xu et al. 2016; Yan et al. 2016). Interestingly, cNpl1^{Tt} expression levels are significantly upregulated at these meiotic stages in Cyc2 and Cyc17 knockouts as examined using publicly available RNA-seq data (supplementary fig. 7, Supplementary Material online). This suggests a role for cNpl1^{Tt} in cell cycle regulation. Our AP-MS experiments using cNpl1-FZZ successfully recovered the bait; however, further work is required to reveal the full scope of its interactions and unravel potential role(s) during development. To this end, carrying out BioID, an orthogonal approach to AP-MS that identifies proteins proximal to the bait in the cell (Kim et al. 2016) during growth and development will be informative and is in progress.

Role of PARPs in Histone Metabolism

Our study also implicates PARPs in histone metabolism. PARPs are functionally diverse proteins with critical roles in a number of processes including DNA break repair (Langelier et al. 2012), cell cycle regulation (Masutani et al. 1995), mRNA binding (Melikishvili et al. 2017), transcription regulation (Ko and Ren 2012; Chen et al. 2014), and maintenance of chromatin architecture (for review Bai 2015). The observation that the *T. thermophila* genome encodes 11 putative PARPs and their expression is temporally regulated suggests that these proteins might be important for distinct cellular processes during various stages of the *Tetrahymena* life cycle. Previous studies have reported that *T. thermophila* histones are highly ADP-ribosylated (Levy-Wilson 1983). It was recently reported that in humans newly synthesized histones H3/H4 carry poly (ADP-ribosylated) marks (Alvarez et al. 2011). In this study, it was proposed that poly (ADP-ribosylation) might help to keep histones H3 and H4 folded in the absence of the other histones (Alvarez et al. 2011). The copurification of certain PARPs with histones in *T. thermophila* is consistent with these earlier findings. Another hypothesis is that certain *T. thermophila* PARPs might function as well as a histone chaperone similar to what has been shown for human PARP1 (Muthurajan et al. 2014). PARP6^{Tt} is of particular interest due to its domain architecture and expression patterns. The PARP6^{Tt} contains 25 tandem ANK repeats similar to its distantly related human Tankyrases 1 and 2 which function in telomere maintenance (Chiang et al. 2008). The PARP6^{Tt} localization pattern during early conjugation correlates with the transcriptional state of the nuclei, suggesting a role in transcription regulation. As the human tankyrases are actively being pursued as drug targets, it will be informative to further examine the PARP6^{Tt} functions through phenotypic analysis of a PARP6^{Tt} knockout.

Conserved Regulatory Network for Variant Hv1

The *T. thermophila* H2A variant Hv1 localization profile has been reported to be correlated with the transcriptional state of the nuclei (Stargell et al. 1993). Consistently, recent genome-wide studies reported a strong enrichment of Hv1 near the transcription start sites (Wang et al. 2017). The SWR- and INO80-complexes are known to function antagonistically to regulate the Htz1 (or H2A.Z in humans) chromatin occupancy (Gerhold and Gasser 2014). We suggest that similar to humans and yeast, *T. thermophila* Hv1 chromatin occupancy is guided by evolutionarily conserved SWR- and INO80-complexes. Based on expression profiles, the subunits of SWR- and INO80-complexes cluster with Hv1 supporting their functional link. Our recent report suggests that a bromo-domain protein Ibd1 in *T. thermophila* might be responsible for recruiting SWR-complex to highly expressed genes (Saettone et al. 2018). *Tetrahymena thermophila* encodes at least 14 bromo-domain proteins, and it will be interesting to examine the potential role of bromo-domain proteins in INO-80 recruitment/function.

In addition to the FACT-complex, Spt6^{Tt} was also recovered as a significant interacting protein in Hv1 AP-MS data. *Saccharomyces cerevisiae* Spt6 has a well-documented role as

a histone chaperone during transcription (Bortvin and Winston 1996; Hartzog et al. 1998). Spt6 physically interacts with RNAPII and functions to reassemble nucleosomes in the wake of RNAPII passage (Kaplan et al. 2003). Recent evidence indicates that the FACT-complex and Spt6 inhibit the widespread chromatin incorporation of H2A.Z by preventing the pervasive recruitment of SWR-C to gene bodies (Jeronimo et al. 2015). The copurification of $Spt6^{Tt}$ with Hv1 suggests that $Spt6^{Tt}$ might have similar functions to regulate the transcription and safeguard the Hv1 occupancy across chromatin. The observation that $Spt6^{Tt}$ and FACT-complex have very similar expression profiles further reinforces the possibility that these proteins are functionally linked. $Spt6^{Tt}$ knockout analysis followed by monitoring SWR-C and Hv1 chromatin occupancy will be instrumental to test this hypothesis.

Conclusions

Our study has provided the first comprehensive view of *T. thermophila* histones H2A, variant Hv1 and H2B protein-interaction networks. Providing new insights into ciliates' histone metabolism, our study also highlighted the conserved nature of chromatin regulatory networks involving H2A (Hv1)–H2B-specific chaperones, thus underscoring the broad utility of these results. Further work is warranted to understand the mechanistic details of conserved chaperones and chromatin-remodeling complexes that we have identified here.

Materials and Methods

Cell Strains

Tetrahymena thermophila strains CU428 [Mpr/Mpr (VII, mp-s)] and B2086 [Mpr^{+}/Mpr^{+} (II, mp-s)] of inbreeding line B were obtained from the *Tetrahymena* Stock Center, Cornell University, Ithaca, NY (http://tetrahymena.vet.cornell.edu/). Cells cultured in 1 × SPP were maintained axenically at 30 °C as previously described (Fillingham et al. 2001).

Bioinformatics and Molecular Evolutionary Analyses

Amino acid sequences for yeast Spt16, Pob3, and human NPM1 were acquired from the UniprotKB and were used as a query to search the NCBI nonredundant database using PSI-BLAST with default parameters. Protein sequences retrieved were analyzed at the Pfam (http://pfam.sanger.ac.uk/; last accessed September 24, 2018) (Finn et al. 2016) and SMART (http://smart.embl-heidelberg.de/; last accessed September 24, 2018) (Letunic and Bork 2018) databases to examine the domain architecture (supplementary file S1, Supplementary Material online, for accession numbers). To reconstruct a protein phylogeny, we used amino acid sequences of the identified conserved domains (as identified by SMART analysis) present within Spt16 (FACT-Spt16_Nlob, Peptidase_M24 (PF00557), Spt16 signature and Rtt106 domains), and Pob3 (SSrecog [PF03531] and Rtt106 domain) orthologs. For the NPM-family phylogeny, complete protein sequences were used. For phylogenetic trees, we also included all the paralogous genes that were identified within a given species. Multiple sequence alignments were built using MUSCLE with default parameters. All protein phylogenetic analyses were carried out using the maximum likelihood (ML) method under LG+G model using MEGA 7 (Kumar et al. 2016). The reliability of the resulting phylogenetic trees was assessed using the bootstrap method (1,000 replicas for each tree). cNpl1 structural prediction and superimposition were carried out using I-TASSER server (Yang et al. 2015). Molecular evolutionary analyses were carried out using MEGA 7 (Kumar et al. 2016). To identify putative PARPs, we used the human PARP1 catalytic domain amino acid sequence as a query against the *T. thermophila* genome. (Please refer to supplementary methods, Supplementary Material online, for further details on molecular evolutionary analyses for Spt16, Pob3, and PARPs.)

Macronuclear Gene Replacement

Epitope tagging vectors for H2A, H2B, Hv1, $Spt16^{Tt}$, $Parp6^{Tt}$, cNpl1, and Impβ3 were constructed by amplifying two separate ~1-kb fragments up- and downstream of the predicted stop codons using WT *T. thermophila* genomic DNA as template. Upstream and downstream PCR products were digested with *Kpn*I and *Xho*I or *Not*I and *Sac*I, respectively. The digested products were cloned into the appropriate sites within the tagging vector (pBKS-FZZ) provided by Dr Kathleen Collins (University of California, Berkeley, CA). The resulting plasmid was again digested with *Kpn*I and *Sac*I prior to transformation. One micrometer gold particles (60 mg/ml; Bio-Rad) were coated with 5 μg of the digested plasmid DNA which was subsequently introduced into the *T. thermophila* MAC using biolistic transformation with a PDS-1000/He Biolistic particle delivery system (Bio-Rad). The transformants were selected using paromomycin (60 μg/ml). To achieve MAC homozygousity, cells were grown in increasing concentrations of paromomycin to a final concentration of 1 mg/ml.

Generation of WCEs and Western Blotting

We used 10% trichloroacetic acid to prepare WCEs by incubation on ice for 30 min. The WCEs were resuspended in 100 μl of SDS loading dye. To neutralize the solution, 10 μl of 1 N NaOH was added. WCEs were subjected to electrophoresis through 10% SDS-PAGE. The proteins were transferred to nitrocellulose and probed with indicated antibodies after blocking in 5% skim milk. Antibodies and dilutions used were anti-Flag (1:4,000; Sigma), anti-Actin (1:10,000; Abcam), and anti-Brg1 (1:1,000, as described by Fillingham et al. [2006]).

Experimental Design for Mass Spectrometry Experiments

For each analysis, at least two biological replicates of each bait were processed independently. These were analyzed alongside negative controls in each batch of samples processed. *Tetrahymena* cells expressing no tagged bait (i.e., empty cells) were used as control. To minimize carry-over issues, extensive washes were performed between each sample (see details for each instrumentation type); and the order of sample acquisition on the mass spectrometer was reversed for the second replicate to avoid systematic bias. On the LTQ mass

spectrometer, a freshly made column was used for each sample as described (Saettone et al. 2018).

Affinity Purification and Mass Spectrometry Sample Preparation

Affinity purification was carried out essentially as described (Garg et al. 2013). Briefly, *T. thermophila* were grown in ~500 ml of 1× SPP to a final concentration of 3×10^5 cells/ml, were pelleted and frozen at −80 °C. The pellets were thawed on ice and resuspended in lysis buffer (10 mM Tris–HCl pH 7.5, 1 mM $MgCl_2$, 300 mM NaCl, and 0.2% NP40 plus yeast protease inhibitors [Sigma]). Benzonase (Sigma E8263) was added (500 units) and extracts were rotated for 30 min at 4 °C. WCEs were clarified by centrifugation at 16,000× g for 30 min, and resulting soluble material was incubated with 50 μl of packed M2-agarose (Sigma) at 4 °C for 3–4 h. The M2-agarose was washed once with 10 ml IPP300 (10 mM Tris–HCl pH 8.0, 300 mM NaCl, 0.1% NP40), two times with 5 ml of IP100 buffer (10 mM Tris–HCl pH 8.0, 100 mM NaCl, 0.1% NP40), and two times with 5 ml of IP100 buffer without detergent (10 mM Tris–HCl pH 8.0, 100 mM NaCl). Five hundred microliters of 0.5 M NH4OH was used to elute the proteins by rotating for 20 min at room temperature. Preparation of protein eluates for mass spectrometry acquisition was essentially as previously described (Saettone et al. 2018). (Please refer to supplementary methods, Supplementary Material online, for details.)

MS Data Visualization and Archiving

Interaction networks were generated using Cytoscape (V3.4.0; Cline et al. 2007). Individual nodes were manually arranged in physical complexes. The annotation of the copurifying partners was carried out using BLAST searches as well as SMART domain analysis (http://smart.embl-heidelberg.de/; last accessed September 24, 2018) of the predicted amino sequences as acquired from the *Tetrahymena* genome database (www.ciliate.org; last accessed September 24, 2018). All MS files used in this study were deposited at MassIVE (http://massive.ucsd.edu; last accessed February 15, 2018). Additional details (including Mass IVE accession numbers and FTP download links) can be found in supplementary table S2*F*, Supplementary Material online. For gene expression analysis, microarray data (accession number GSE11300) was acquired (http://tfgd.ihb.ac.cn/; last accessed September 24, 2018) and the expression values were represented in the heatmap format. Hierarchical clustering was performed to assess the similarities in gene expression profiles.

Indirect IF

Cells were grown and fixed during vegetative growth, 24-h starvation, and 2, 4, 6, and 7.5 h postmixing after starvation to perform indirect IF as previously described (Garg et al. 2013). (Please refer to supplementary methods, Supplementary Material online, for details.)

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

We thank Dr Takahiko Akematsu for his assistance with microscopy. We also thank Anita Samardzic for her technical assistance with *Tetrahymena* growth media preparations. Work in the Fillingham and Lambert laboratories was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grants RGPIN-2015-06448 and RGPIN-2017-06124, respectively. J.-P.L. holds a Junior 1 salary award from the Fonds de Recherche du Québec-Santé (FRQ-S) and was also supported through a John R. Evans Leaders Fund from the Canada Foundation for Innovation (37454). Work in the Pearlman laboratory was supported by Canadian Institutes of Health Research (CIHR) (MOP13347) and Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant 539509. Work in the Gingras laboratory was supported by the Canadian Institutes of Health Research (CIHR) Foundation Grant (FDN 143301). The authors declare no conflict of interest.

Author Contributions

K.A. generated H2A-FZZ, Spt16-FZZ, PARP6-FZZ, and Hv1-FZZ cell lines and performed Western blots, affinity purifications, IF microscopy, data analysis, participated in manuscript drafting and in overall study design with J.F. and R.E.P.'s feedback. S.N.-S. performed evolutionary analysis, participated in study design with feedback from J.F., R.E.P. and K.A, prepared all the final figures, wrote the manuscript, and coordinated the edits from all the authors. J.G. generated cNpl1-FZZ cell line, performed IF analysis on cNpl1-FZZ and affinity purification on Hv1-FZZ. A.S. generated H2B-FZZ, performed IFs and affinity purification on H2B-FZZ. J.D. participated in H2B-FZZ generation. J.-P.L. processed and analyzed samples for mass spectrometry, provided feedback on data figures, and edited the manuscript. A.-C.G. participated in manuscript editing and mass spectrometry. R.E.P. cosupervised the project, provided reagents, monitored the overall progress, and participated in manuscript editing. J.F. envisioned and designed the study, cosupervised the project, coordinated the overall progress of the study, and edited the manuscript. All authors have read and approved the final manuscript.

References

Adl SM, Simpson AGB, Lane CE, Lukeš J, Bass D, Bowser SS, Brown MW, Burki F, Dunthorn M, Hampl V. 2012. The revised classification of eukaryotes. *J Eukaryot Microbiol*. 59(5): 429–493.

Allshire RC, Madhani HD. 2018. Ten principles of heterochromatin formation and function. *Nat Rev Mol Cell Biol*. 19(4): 229–244.

Alvarez F, Muñoz F, Schilcher P, Imhof A, Almouzni G, Loyola A. 2011. Sequential establishment of marks on soluble histones H3 and H4. *J Biol Chem*. 286(20): 17714–17721.

Bai P. 2015. Biology of poly(ADP-ribose) polymerases: the factotums of cell maintenance. *Mol Cell*. 58(6): 947–958.

Belotserkovskaya R, Oh S, Bondarenko VA, Orphanides G, Studitsky VM, Reinberg D. 2003. FACT facilitates transcription-dependent nucleosome alteration. *Science* 301(5636): 1090–1093.

Boekhorst J, van Breukelen B, Heck AJ, Snel B. 2008. Comparative phosphoproteomics reveals evolutionary and functional conservation of phosphorylation across eukaryotes. *Genome Biol*. 9(10): R144.

Bortvin A, Winston F. 1996. Evidence that Spt6p controls chromatin structure by a direct interaction with histones. *Science* 272(5267): 1473–1476.

Box JK, Paquet N, Adams MN, Boucher D, Bolderson E, O'Byrne KJ, Richard DJ. 2016. Nucleophosmin: from structure and function to disease development. *BMC Mol Biol*. 17(1): 19.

Cassidy-Hanley D, Bowen J, Lee JH, Cole E, VerPlank LA, Gaertig J, Gorovsky MA, Bruns PJ. 1997. Germline and somatic transformation of mating *Tetrahymena thermophila* by particle bombardment. *Genetics* 146(1): 135–147.

Charles Richard JL, Shukla MS, Menoni H, Ouararhni K, Lone IN, Roulland Y, Papin C, Ben Simon E, Kundu T, Hamiche A, et al. 2016. FACT assists base excision repair by boosting the remodeling activity of RSC. Bianchi M, editor. *PLoS Genet* 12(7): e1006221.

Chen H, Ruiz PD, Novikov L, Casill AD, Park JW, Gamble MJ. 2014. MacroH2A1.1 and PARP-1 cooperate to regulate transcription by promoting CBP-mediated H2B acetylation. *Nat. Struct. Mol. Biol*

Chiang YJ, Hsiao SJ, Yver D, Cushman SW, Tessarollo L, Smith S, Hodes RJ. 2008. Tankyrase 1 and tankyrase 2 are essential but redundant for mouse embryonic development. *PLoS One* 3(7): e2639.

Citarelli M, Teotia S, Lamb RS. 2010. Evolutionary history of the poly(ADP-ribose) polymerase gene family in eukaryotes. *BMC Evol. Biol* 10:308.

Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, Christmas R, Avila-Campilo I, Creech M, Gross B. 2007. Integration of biological networks and gene expression data using Cytoscape. *Nat. Protoc* 2(10): 2366–2382.

De Koning L, Corpet A, Haber JE, Almouzni G. 2007. Histone chaperones: an escort network regulating histone traffic. *Nat Struct Mol Biol*. 14(11): 997–1007.

Eirín-López JM, Frehlick LJ, Ausió J. 2006. Long-term evolution and functional diversification in the members of the nucleophosmin/nucleoplasmin family of nuclear chaperones. *Genetics* 173(4): 1835–1850.

Eirín-López JM, Rebordinos L, Rooney AP, Rozas J. 2012. The birth-and-death evolution of multigene families revisited. In: Genome Dynamics Vol. 7:170–196.

English CM, Adkins MW, Carson JJ, Churchill MEA, Tyler JK. 2006. Structural basis for the histone chaperone activity of Asf1. *Cell* 127(3): 495–508.

Fillingham JS, Bruno D, Pearlman RE. 2001. Cis-acting requirements in flanking DNA for the programmed elimination of mse2.9: a common mechanism for deletion of internal eliminated sequences from the developing macronucleus of *Tetrahymena thermophila*. *Nucleic Acids Res*. 29(2): 488–498.

Fillingham JS, Garg J, Tsao N, Vythilingum N, Nishikawa T, Pearlman RE. 2006. Molecular genetic analysis of an SNF2/brahma-related gene in *Tetrahymena thermophila* suggests roles in growth and nuclear development. *Eukaryot Cell*. 5(8): 1347–1359.

Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, et al. 2016. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res*. 44(D1): D279–D285.

Formosa T, Eriksson P, Wittmeyer J, Ginn J, Yu Y, Stillman DJ. 2001. Spt16-Pob3 and the HMG protein Nhp6 combine to form the nucleosome-binding factor SPN. *EMBO J*. 20(13): 3506–3517.

Frehlick LJ, Eirín-López JM, Ausió J. 2007. New insights into the nucleophosmin/nucleoplasmin family of nuclear chaperones. *Bioessays* 29(1): 49–59.

Fujiu K, Numata O. 2004. Identification and molecular cloning of *Tetrahymena* 138-kDa protein, a transcription elongation factor homologue that interacts with microtubules in vitro. *Biochem Biophys Res Commun*. 315(1): 196–203.

Gao F, Warren A, Zhang Q, Gong J, Miao M, Sun P, Xu D, Huang J, Yi Z, Song W. 2016. The all-data-based evolutionary hypothesis of ciliated protists with a revised classification of the phylum Ciliophora (Eukaryota, Alveolata). *Sci Rep*. 6:24874.

Garg J, Lambert JP, Karsou A, Marquez S, Nabeel-Shah S, Bertucci V, Retnasothie DV, Radovani E, Pawson T, Gingras AC, et al. 2013. Conserved Asf1-importin β physical interaction in growth and sexual development in the ciliate *Tetrahymena thermophila*. *J Proteomics*. 94:311–326.

Gerhold C-B, Hauer MH, Gasser SM. 2015. INO80-C and SWR-C: guardians of the Genome. *J Mol Biol*. 427(3): 637–651.

Gerhold CB, Gasser SM. 2014. INO80 and SWR complexes: relating structure to function in chromatin remodeling. *Trends Cell Biol*. 24(11): 619–631.

Goldberg AD, Banaszynski LA, Noh K-M, Lewis PW, Elsaesser SJ, Stadler S, Dewell S, Law M, Guo X, Li X, et al. 2010. Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* 140(5): 678–691.

González-Romero R, Eirín-López JM, Ausió J. 2015. Evolution of high mobility group nucleosome-binding proteins and its implications for vertebrate chromatin specialization. *Mol Biol Evol*. 32(1): 121–131.

Grisendi S, Bernardi R, Rossi M, Cheng K, Khandker L, Manova K, Pandolfi PP. 2005. Role of nucleophosmin in embryonic development and tumorigenesis. *Nature* 437(7055): 147–153.

Grover P, Asa JS, Campos EI. 2018. H3–H4 Histone Chaperone Pathways. Annu. Rev. Genet. 52:109–130.

Hammond CM, Strømme CB, Huang H, Patel DJ, Groth A. 2017. Histone chaperone networks shaping chromatin function. *Nat Rev Mol Cell Biol*. 18(3): 141–158.

Hartzog GA, Wada T, Handa H, Winston F. 1998. Evidence that Spt4, Spt5, and Spt6 control transcription elongation by RNA polymerase II in *Saccharomyces cerevisiae*. *Genes Dev*. 12(3): 357–369.

Hassa PO, Hottiger MO. 2008. The diverse biological roles of mammalian PARPS, a small but powerful family of poly-ADP-ribose polymerases. *Front Biosci*. 13:3046–3082.

Hoek M, Stillman B. 2003. Chromatin assembly factor 1 is essential and couples chromatin assembly to DNA replication in vivo. *Proc Natl Acad Sci U S A*. 100(21): 12183–12188.

Hsieh F-K, Kulaeva OI, Patel SS, Dyer PN, Luger K, Reinberg D, Studitsky VM. 2013. Histone chaperone FACT action during transcription through chromatin by RNA polymerase II. *Proc Natl Acad Sci U S A*. 110(19): 7654–7659.

Jeronimo C, Watanabe S, Kaplan CD, Peterson CL, Robert F. 2015. The histone chaperones FACT and Spt6 restrict H2A.Z from intragenic locations. *Mol Cell*. 58(6): 1113–1123.

Jin C, Zang C, Wei G, Cui K, Peng W, Zhao K, Felsenfeld G. 2009. H3.3/H2A.Z double variant-containing nucleosomes mark "nucleosome-free regions" of active promoters and other regulatory regions. *Nat Genet*. 41(8): 941–945.

Jullien J, Astrand C, Szenker E, Garrett N, Almouzni G, Gurdon JB. 2012. HIRA dependent H3.3 deposition is required for transcriptional reprogramming following nuclear transfer to *Xenopus* oocytes. *Epigenetics Chromatin* 5(1): 17.

Kaplan CD, Laprade L, Winston F. 2003. Transcription elongation factors repress transcription initiation from cryptic sites. *Science* 301(5636): 1096–1099.

Karrer KM. 2012. Nuclear dualism. *Methods Cell Biol*. 109:29–52.

Keck KM, Pemberton LF. 2012. Histone chaperones link histone nuclear import and chromatin assembly. *Biochim Biophys Acta*. 1819(3–4): 277–289.

Kim DI, Jensen SC, Noble KA, Kc B, Roux KH, Motamedchaboki K, Roux KJ. 2016. An improved smaller biotin ligase for BioID proximity labeling. *Mol Biol Cell*. 27(8): 1188–1196.

Ko HL, Ren EC. 2012. Functional aspects of PARP1 in DNA repair and transcription. *Biomolecules* 2(4): 524–548.

Kobor MS, Venkatasubrahmanyam S, Meneghini MD, Gin JW, Jennings JL, Link AJ, Madhani HD, Rine J. 2004. A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. *PLoS Biol*. 2(5): E131.

Krogan NJ, Keogh M-C, Datta N, Sawa C, Ryan OW, Ding H, Haw RA, Pootoolal J, Tong A, Canadien V, et al. 2003. A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. *Mol Cell*. 12(6): 1565–1576.

Kumar S, Stecher G, Suleski M, Hedges SB. 2017. TimeTree: a resource for timelines, timetrees, and divergence times. *Mol Biol Evol*. 34(7): 1812–1819.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol*. 33(7): 1870–1874.

Kurat CF, Yeeles JTP, Patel H, Early A, Diffley JFX. 2017. Chromatin controls DNA replication origin selection, lagging-strand synthesis, and replication fork rates. *Mol Cell*. 65(1): 117–130.

Langelier M-F, Planck JL, Roy S, Pascal JM. 2012. Structural basis for DNA damage-dependent poly(ADP-ribosyl)ation by human PARP-1. *Science* 336(6082): 728–732.

Latreille D, Bluy L, Benkirane M, Kiernan RE. 2014. Identification of histone 3 variant 2 interacting factors. *Nucleic Acids Res*. 42(6): 3542–3550.

Letunic I, Bork P. 2018. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res*. 46(D1): D493–D496.

Levy-Wilson B. 1983. Glycosylation, ADP-ribosylation, and methylation of *Tetrahymena* histones. *Biochemistry* 22(2): 484–489.

Liu X, Li B, Gorovsky MA. 1996. Essential and nonessential histone H2A variants in *Tetrahymena thermophila*. *Mol Cell Biol*. 16:4305–4311.

Lotan T, Chalifa-Caspi V, Ziv T, Brekhman V, Gordon MM, Admon A, Lubzens E. 2014. Evolutionary conservation of the mature oocyte proteome. *EuPA Open Proteomics*. 3:27–36.

Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* 389(6648): 251–260.

Malone CD, Falkowska KA, Li AY, Galanti SE, Kanuru RC, LaMont EG, Mazzarella KC, Micev AJ, Osman MM, Piotrowski NK, et al. 2008. Nucleus-specific importin alpha proteins and nucleoporins regulate protein import and nuclear division in the binucleate *Tetrahymena thermophila*. *Eukaryot Cell*. 7(9): 1487–1499.

Martindale DW, Allis CD, Bruns PJ. 1982. Conjugation in *Tetrahymena thermophila*. A temporal analysis of cytological stages. *Exp Cell Res*. 140(1): 227–236.

Martindale DW, Allis CD, Bruns PJ. 1985. RNA and protein synthesis during meiotic prophase in *Tetrahymena thermophila*. *J Protozool*. 32(4): 644–649.

Mason PB, Struhl K. 2003. The FACT complex travels with elongating RNA polymerase II and is important for the fidelity of transcriptional initiation in vivo. *Mol Cell Biol*. 23(22): 8323–8333.

Masutani M, Nozaki T, Wakabayashi K, Sugimura T. 1995. Role of poly(ADP-ribose) polymerase in cell-cycle checkpoint mechanisms following gamma-irradiation. *Biochimie* 77(6): 462–465.

Melikishvili M, Chariker JH, Rouchka EC, Fondufe-Mittendorf YN. 2017. Transcriptome-wide identification of the RNA-binding landscape of the chromatin-associated protein PARP1 reveals functions in RNA biogenesis. *Cell Discov*. 3:17043.

Mendiratta S, Gatto A, Almouzni G. 2018. Histone supply: multitiered regulation ensures chromatin dynamics throughout the cell cycle. *J Cell Biol*. 218(1):39–54.

Miao W, Xiong J, Bowen J, Wang W, Liu Y, Braguinets O, Grigull J, Pearlman RE, Orias E, Gorovsky MA. 2009. Microarray analyses of gene expression during the Tetrahymena thermophila life cycle. *PLoS One* 4(2):e4429.

Mochizuki K, Gorovsky MA. 2004. Small RNAs in genome rearrangement in *Tetrahymena*. *Curr Opin Genet Dev*. 14(2): 181–187.

Mosammaparast N, Ewart CS, Pemberton LF. 2002. A role for nucleosome assembly protein 1 in the nuclear transport of histones H2A and H2B. *EMBO J*. 21(23): 6527–6538.

Muthurajan UM, Hepler MRD, Hieb AR, Clark NJ, Kramer M, Yao T, Luger K. 2014. Automodification switches PARP-1 function from chromatin architectural protein to histone chaperone. *Proc Natl Acad Sci U S A*. 111(35): 12752–12757.

Nabeel-Shah S, Ashraf K, Pearlman RE, Fillingham J. 2014. Molecular evolution of NASP and conserved histone H3/H4 transport pathway. *BMC Evol Biol*. 14:139.

Namboodiri VMH, Dutta S, Akey IV, Head JF, Akey CW. 2003. The crystal structure of *Drosophila* NLP-core provides insight into pentamer formation and histone binding. *Structure* 11(2): 175–186.

Okuwaki M, Matsumoto K, Tsujimoto M, Nagata K. 2001. Function of nucleophosmin/B23, a nucleolar acidic protein, as a histone chaperone. *FEBS Lett*. 506(3): 272–276.

Orias E, Cervantes MD, Hamilton EP. 2011. *Tetrahymena thermophila*, a unicellular eukaryote with separate germline and somatic genomes. *Res Microbiol*. 162(6): 578–586.

Papamichos-Chronakis M, Watanabe S, Rando OJ, Peterson CL. 2011. Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. *Cell* 144(2): 200–213.

Pfister JA, D'Mello SR. 2016. Regulation of neuronal survival by nucleophosmin 1 (NPM1) is dependent on its expression level, subcellular localization, and oligomerization status. *J Biol Chem*. 291(39): 20787–20797.

Ray-Gallet D, Woolfe A, Vassias I, Pellentz C, Lacoste N, Puri A, Schultz DC, Pchelintsev NA, Adams PD, Jansen LET, et al. 2011. Dynamics of histone H3 deposition in vivo reveal a nucleosome gap-filling mechanism for H3.3 to maintain chromatin integrity. *Mol Cell*. 44(6): 928–941.

Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM. 1998. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem*. 273(10): 5858–5868.

Saettone A, Garg J, Lambert J-P, Nabeel-Shah S, Ponce M, Burtch A, Thuppu Mudalige C, Gingras A-C, Pearlman RE, Fillingham J. 2018. The bromodomain-containing protein Ibd1 links multiple chromatin-related protein complexes to highly expressed genes in *Tetrahymena thermophila*. *Epigenetics Chromatin* 11(1): 10.

Song X, Gjoneska E, Ren Q, Taverna SD, Allis CD, Gorovsky MA. 2007. Phosphorylation of the SQ H2A.X motif is required for proper meiosis and mitosis in *Tetrahymena thermophila*. *Mol Cell Biol*. 27(7): 2648–2660.

Stargell LA, Bowen J, Dadd CA, Dedon PC, Davis M, Cook RG, Allis CD, Gorovsky MA. 1993. Temporal and spatial association of histone H2A variant hv1 with transcriptionally competent chromatin during nuclear development in *Tetrahymena thermophila*. *Genes Dev*. 7(12B): 2641–2651.

Straube K, Blackwell JS, Pemberton LF. 2010. Nap1 and Chz1 have separate Htz1 nuclear import and assembly functions. *Traffic* 11(2): 185–197.

Studamire B, Quach T, Alani E. 1998. *Saccharomyces cerevisiae* Msh2p and Msh6p ATPase activities are both required during mismatch repair. *Mol Cell Biol*. 18(12): 7590–7601.

Stuwe T, Hothorn M, Lejeune E, Rybin V, Bortfeld M, Scheffzek K, Ladurner AG. 2008. The FACT Spt16 "peptidase" domain is a histone H3-H4 binding module. *Proc Natl Acad Sci U S A*. 105(26): 8884–8889.

Swaminathan V, Kishore AH, Febitha KK, Kundu TK. 2005. Human histone chaperone nucleophosmin enhances acetylation-dependent chromatin transcription. *Mol Cell Biol*. 25(17): 7534–7545.

Tagami H, Ray-Gallet D, Almouzni G, Nakatani Y. 2004. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* 116(1): 51–61.

Talbert PB, Ahmad K, Almouzni G, Ausió J, Berger F, Bhalla PL, Bonner WM, Cande W, Chadwick BP, Chan SWL, et al. 2012. A unified phylogeny-based nomenclature for histone variants. *Epigenetics Chromatin* 5(1): 7.

Teo G, Liu G, Zhang J, Nesvizhskii AI, Gingras A-C, Choi H. 2014. SAINTexpress: improvements and additional features in Significance Analysis of INTeractome software. *J Proteomics*. 100:37–43.

Venkatesh S, Workman JL. 2015. Histone exchange, chromatin structure and the regulation of transcription. *Nat Rev Mol Cell Biol*. 16(3): 178–189.

Wang Y, Chen X, Sheng Y, Liu Y, Gao S. 2017. N6-adenine DNA methylation is associated with the linker DNA of H2A.Z-containing well-

positioned nucleosomes in Pol II-transcribed genes in *Tetrahymena*. *Nucleic Acids Res.* 45(20): 11594–11606.

Wang Z, Cui B, Gorovsky MA. 2009. Histone H2B ubiquitylation is not required for histone H3 methylation at lysine 4 in *Tetrahymena*. *J Biol Chem*. 284(50): 34870–34879.

Xiong J, Lu X, Zhou Z, Chang Y, Yuan D, Tian M, Zhou Z, Wang L, Fu C, Orias E, et al. 2012. Transcriptome analysis of the model protozoan, *Tetrahymena thermophila*, using Deep RNA sequencing. *PLoS One* 7(2): e30630.

Xu Q, Wang R, Ghanam AR, Yan G, Miao W, Song X. 2016. The key role of CYC2 during meiosis in *Tetrahymena thermophila*. *Protein Cell* 7(4): 236–249.

Yan G-X, Dang H, Tian M, Zhang J, Shodhan A, Ning Y-Z, Xiong J, Miao W. 2016. Cyc17, a meiosis-specific cyclin, is essential for anaphase initiation and chromosome segregation in *Tetrahymena thermophila*. *Cell Cycle* 15(14): 1855–1864.

Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. 2015. The I-TASSER Suite: protein structure and function prediction. *Nat Methods*. 12(1): 7–8.

Yang J, Zhang X, Feng J, Leng H, Li S, Xiao J, Liu S, Xu Z, Xu J, Li D, et al. 2016. The histone chaperone FACT contributes to DNA replication-coupled nucleosome assembly. *Cell Rep*. 14(5): 1128–1141.

Yao M-C, Fuller P, Xi X. 2003. Programmed DNA deletion as an RNA-guided system of genome defense. *Science* 300(5625): 1581–1584.

Yao M-CC, Choi J, Yokoyama S, Austerberry CF, Yao C-HH. 1984. DNA elimination in *Tetrahymena*: a developmental process involving extensive breakage and rejoining of DNA at defined sites. *Cell* 36(2): 433–440.

Yao MC, Yao CH, Monks B. 1990. The controlling sequence for site-specific chromosome breakage in *Tetrahymena*. *Cell* 63(4): 763–772.

Zhang Y, Ku WL, Liu S, Cui K, Jin W, Tang Q, Lu W, Ni B, Zhao K. 2017. Genome-wide identification of histone H2A and histone variant H2A.Z-interacting proteins by bPPI-seq. *Cell Res*. 27(10): 1258–1274.

Zhao X, Ji J, Yu L-R, Veenstra T, Wang XW. 2015. Cell cycle-dependent phosphorylation of nucleophosmin and its potential regulation by peptidyl-prolyl cis/trans isomerase. *J Mol Biochem*. 4:95–103.

Zunder RM, Antczak AJ, Berger JM, Rine J. 2012. Two surfaces on the histone chaperone Rtt106 mediate histone binding, replication, and silencing. *Proc Natl Acad Sci U S A*. 109(3): E144–E153.