Published in final edited form as:
Trends Parasitol. 2012 May ; 28(5): 202–213. doi:10.1016/j.pt.2012.02.009.

Chromatin modifications, epigenetics, and how protozoan parasites regulate their lives

Matthew M. Croken, **Sheila C. Nardelli**, and **Kami Kim***
Departments of Medicine and Microbiology and Immunology Albert Einstein College of Medicine 1300 Morris Park Ave. Bronx, NY 10128 01 718 430 2611

Abstract

Chromatin structure plays a vital role in epigenetic regulation of protozoan parasite gene expression. Epigenetic gene regulation impacts parasite virulence, differentiation and cell cycle control. Recent work in many laboratories has elucidated the functions of histone modifying proteins that regulate parasite gene expression by chemical modification of constituent nucleosomes. A major focus of investigation has been characterizing post-translational modifications (PTM) of histones and identifying the enzymes that are responsible. Despite conserved features and specificity common to all eukaryotes, parasite enzymes involved in chromatin modification have unique functions that regulate unique aspects of parasite biology.

Epigenetics and gene expression in protozoan parasites

Protozoan parasites constitute a major source of human mortality and morbidity. These parasites have complex life cycles in multiple hosts and must initiate complex developmental programs in response to environmental cues including stress, transition to different hosts, or host defenses. Differentiation events require drastic, and sometimes rapid, alteration of the organism's gene expression profile. In addition, many protozoan parasites use antigenic variation to elude the host immune response. The mechanisms by which parasites regulate transcription and respond to changes in their environment are only partially characterized. Over the past decade, epigenetic regulation of gene expression has emerged as a critical aspect of parasite biology. Although many of the fundamental principles of epigenetic gene regulation are similar to mammalian cells and model systems, protozoan parasites also display unique and diverse mechanisms of epigenetic gene regulation.

The term epigenetics (see Glossary) refers to heritable changes in an organism that alter gene expression levels without altering the DNA sequence. These changes can be mediated by a variety of mechanisms (Box 1) that can be broadly categorized as RNA-based epigenetic gene regulation and epigenetic regulation occurring via post-translational modification of chromatin. The roles of small silencing RNAs in the biology of protozoan parasites has been reviewed recently [1] and appears to be a major mechanism of gene regulation in some but not all protozoa. An intriguing but unresolved enigma is why some

*correspondence: kami.kim@einstein.yu.edu.

parasites have retained RNA-based gene silencing whereas other closely related species do not use small RNAs as a mechanism of gene regulation [2].

An under-studied but potentially significant mechanism of epigenetic regulation is structural based epigenetic inheritance, whereby epigenetic inheritance is via a self-perpetuating spatial or protein structure. Several studies have reported unique nuclear localization of specialized regions of chromatin: *Toxoplasma gondii* centromeres [3], *Plasmodium falciparum* active and silenced *var* genes [4–6] and *Trypanosoma brucei* VSG expression sites [7]. The assembly and propagation of nuclear architecture is likely to be a fruitful area of investigation in protozoan parasites that may also represent an additional mechanism of epigenetic gene regulation.

Epigenetic gene regulation most frequently refers to gene regulation achieved by changes in chromatin structure. Modification of chromatin results in changes in the accessibility of chromatin to transcriptional complexes and polymerases. Either DNA or the proteins associated with DNA, chiefly histones, may be post-translationally modified yielding euchromatin (loose, accessible) or heterochromatin (compact, inaccessible) (Figure 1). Advances in understanding the regulators of epigenetic gene regulation in the metazoa have led to development of new compounds that interfere with pathogenic states. Given the importance of epigenetics for parasite biology, parasite epigenetics machinery has potential as a new avenue of investigation for drug development of new treatments of infections caused by protozoan parasites.

This review will focus upon post-translational modifications (PTM) of chromatin, comparing *P. falciparum, T. gondii* and *T. brucei,* protozoan species in which chromatin modification and epigenetics have been most extensively studied. For more specific details about epigenetics in each organism, several excellent reviews have also been published [8–11]

The histone code: post-translational modification of histones

Histones are small basic proteins that assemble in an octamer consisting of 2 copies of core histone proteins H2A, H2B, H3 and H4 (Table 1). The basic unit of the nucleosome consists of approximately 147 bp of DNA wound twice around the histone octamer core as well the linker DNA connecting nucleosomes. Some parasite species, including trypanosomatids, have a specialized linker histone, H1 [8], but it is not yet clear whether other parasites have H1 proteins. Histone proteins are extensively post-translationally modified, particularly within their N termini (Figure 2). These modifications are important for the recruitment of other chromatin remodeling complexes as well as the transcriptional machinery, and analysis of the post-translational modifications (PTM) of histones in the protozoa show some unexpected differences that are likely to reflect important biological adaptations [12, 13] [14]. While the canonical histones H2A, H2B, H3 and H4 are very well conserved in many parasites, including the Apicomplexa, the histones of trypanosomatids are quite divergent (Figure S1), probably reflecting their unique mechanisms of gene regulation [8].

The `histone code hypothesis' maintains that the protein components of chromatin and their covalent chemical modifications relay critical information to the rest of the cell about the underlying DNA sequences [15], rather than simply being packaging proteins for compacting DNA. Altered histones can delineate the boundaries between promoter regions upstream of genes, gene coding regions, intergenic space and DNA strand breaks. These marks can also be used to distinguish between active and silent genes. The compacted parasite genomes typically lack recognizable orthologues for the large repertoire of transcription factors that their free-living counterparts possess. Recent work indicates that

the `histone code' is a critical regulatory mechanism of gene expression in parasitic microorganisms.

There are many potential chemical groups that may be conjugated histone proteins. Most modify the exposed N-terminal tails of histones, and the combination of these PTM encodes the `histone code'. Modifications that affect gene expression include acetylation, methylation, phosphorylation, glycosylation, ADP-ribosylation, ubiquitination, and sumoylation. These PTM can affect stability or function of the modified proteins and frequently work in concert to regulate gene expression. With the improvement of mass spectrometry methods and greater interest in PTM, comprehensive mass spectrometric surveys of histone PTM have illuminated the complexity and combinatorial potential of the histone code [16]. The most comprehensive studies in parasites have been performed in *P. falciparum* and *T. brucei* and have identified novel features mediated by effectors of epigenetic regulation [12–14]. A similar study underway in *T. gondii* (S.C. Nardelli, F. Che *et al.*, unpublished) shows both conserved and unique features of *T. gondii* histone PTMs.

In many species the presence of histone PTMs has been inferred by activity of recombinant proteins on peptide or recombinant protein substrates, which may not always reflect *in vivo* biological activity. Other studies have relied upon recognition of parasite histones by commercial antibodies recognizing various modifications. While the conclusions of such studies may be valid, a comprehensive survey of 200 antibodies has shown surprising lack of specificity and batch to batch viability for some of these reagents [17], and our own experience in *T. gondii* has shown that not all PTM inferred from histone antibody studies can be verified using mass spectrometry analysis of histones (S.C. Nardelli, F. Che *et al.*, unpublished). Thus results of experiments to infer the presence of histone PTMs should be interpreted carefully and experiments conducted with appropriate controls.

In addition to the canonical histone proteins, a number of variant histones exist in Eukaryota that may be substituted for the canonical core histone (Table 1, Table S1, Figure S1). Many of these variants appear to be conserved in at least some of the parasite species discussed here. These include H3.3, CenH3, H2A.Z, and H2A.X as well as a parasite specific H2B variant (H2Bv) [18–20]. Trypanosomatids also encode unique histone variants. In general the histone variants are less conserved than canonical histones (Table S1, Figure S1). The presence of these variants, as well as covalent modification of variant histone proteins, influences nucleosome stability, chromatin structure and the expression of underlying genes. Studies in *T. gondii, P. falciparum* and *T. brucei* suggest that the distribution of variant histones correlates with the chromatin state, although the specifics are often unique to each parasite (Figure 3). The PTM of variant histones is also linked with transcriptional activity or chromatin, so histone variants, in conjunction with chemical modifications, constitute an `extended histone code'.

The chromatin remodeling machinery in parasites

Bioinformatic analyses have been performed to examine parasite genomes for components of classic transcriptional and epigenetic machinery [21]. The function of core conserved components have been inferred, but not all proteins characterized in mammalian or yeast systems are present, and proteomic characterization of macromolecular complexes involved in transcription or chromatin remodeling are likely to identify unique parasite-specific molecules. Chromatin remodelers are protein complexes that alter the structure and composition of nucleosomes leading to changes in nucleosome position and access of the transcriptional machinery. The remodeling activity is typically ATP dependent and usually includes a core ATPase with domains that interact with chromatin. The enzymes that modify chromatin are frequently called `chromatin writers'. Specialized adaptor proteins encode

modules known as `chromatin readers', such as chromodomains and bromodomains that recognize and bind these modifications to recruit transcription factors and the RNA polymerase complex (Figure 2). Many of these chromatin reader modules are conserved in parasitic protozoa and appear to have similar functions. Additional histone PTM have recently been described in mammalian cells, and the significance of these modifications and the enzymes responsible are areas of active and evolving investigation [22]. Whether these new modifications are present in parasitic protozoa is unknown.

The interplay between modified histones, chromatin modifying machinery, and general or sequence specific transcription factors is an essential element of transcriptional regulation. To date few transcription factors have been identified in the parasitic protozoa [21] and fewer have been experimentally validated. Gene expression in the kinetoplastids is regulated mainly at the post-transcriptional level. In the Apicomplexa, the AP2 family of plant-like transcription factors have been proposed to be the major sequence-specific regulators of gene expression [23], and a recent study using protein-binding arrays indicates the AP2 DNA-binding domains of many PfAP2 proteins do bind specific DNA sequences [24]. Analysis of yeast two-hybrid screens in *P. falciparum* reveals many proteins with the AP2 DNA-binding domain interact directly or indirectly with chromatin remodelers such as GCN5, a histone acetyl transferase (HAT) [25]. In *T. gondii,* TgHDAC3 immunoprecipitates with TgCRC-350 (TGME49_072710) which is predicted to have an AP2 domain [26], and ongoing studies have confirmed that several TgAP2 associate with other chromatin remodeling proteins, including GCN5B (W.J. Sullivan and K. Kim, unpublished). Some of the AP2 family members may have functions separate from transcription but important for epigenetic gene regulation. *P. falciparum* AP2 PFF0200c (or PfSIP2) appears to be important for heterochromatin formation and genome integrity of subtelomeric regions [27] whereas PfAP2 PF11_0091 binds the conserved intron of *var* genes and is implicated in targeting of silenced *var* genes to heterochromatin-rich silenced *var* clusters in the nuclear periphery [28].

The functions of many histone PTM appear to be evolutionarily conserved. In general, acetylation of histones is associated with gene activation, whereas histone methylation can be associated with either gene repression or gene activation depending upon the residue and histone modified. Although many of the basic concepts of the histone code are preserved in parasites, there is surprising diversity in the molecular players and histone marks that regulate epigenetic gene regulation in the protozoa, even among species that are relatively close phylogenetically. The histones and histone PTM of trypanosomatids are particularly divergent (Figure 2) [8]. Among the most conserved modifications is Histone 3 trimethylation of lysine 4 (H3K4me3). This is an established marker of transcriptionally active promoters, and *T. gondii*, *P. falciparum*, and *T. brucei* all possess this modification at putative transcription start sites [19, 29, 30]. *T. brucei* histones are also distinguished by acetylations that are not phylogenetically conserved (Figure 2) [8].

Opposed to `activation marks' are modifications associated with silent genes and densely packed heterochromatin. Notable among these are H3K9 methylation. As a general rule, protozoan parasite histones are more highly enriched in the activation marks associated with euchromatin with lower abundance of histone modifications associated with heterochromatin [3, 30, 31]. In the Apicomplexa, heterochromatin marks have localized to specialized regions of the chromatin such as telomeres or centromeres; however, they have not been universally associated with silenced genes. For example, the heterochromatin mark H3K9me3 and heterochromatin have been identified in both *P. falciparum* and *T. gondii,* but these have different biological functions and chromosomal locations (see sections below discussing each parasite). By contrast, *T. brucei* expresses a single *VSG* and silences of most

of its telomeric *VSG* genes during regulation of antigenic variation but does not have H3K9me3 [8].

The state of chromatin and histone modification has been mapped to specific genetic loci using a technique called chromatin immunoprecipitation (ChIP), whereby proteins and DNA are immunoprecipitated as a complex, the DNA isolated, and the relative enrichment of DNA is compared to input DNA using quantitative-PCR (q-PCR), hybridization to a microarray (ChIP-chip) or high throughput sequencing (ChIP-seq). In parallel the transcriptome is often monitored using similar platforms to correlate gene expression with different chromatin states. Mass spectrometry and yeast two hybrid studies have also been used to gain insight into the macromolecular complexes that participate in transcription regulation, gene activation or silencing and chromatin remodeling. A major challenge is integration of the various datasets to obtain an integrated view of the importance and significance of particular chromatin states in the biology of parasites.

Antigenic variation in parasites: epigenetic gene regulation

Antigenic variation is a common strategy used by pathogens that are exposed to host immune factors, particularly host antibodies. The pathogen varies the molecules that the host encounters and switches its coat as the host mounts an antibody response. The molecular mechanisms of switching of antigens vary in eukaryotic pathogens, but usually rely upon epigenetic mechanisms that help ensure that all genes are not simultaneously expressed. This allows for a somewhat stochastic but regulated process of switching that facilitates responses to environmental cues and to changes within the host. Among the parasites that use antigenic variation are *Giardia lamblia, P. falciparum* and *T. brucei.* Each organism differs in the specifics of how switching of antigens is achieved although epigenetic mechanisms appear to be critical in each of these cases. This is an area of intensive inquiry since interventions that dysregulate antigenic switching in parasites could be a potential strategy to induce or improve immune responses of infected hosts [32]. Epigenetic modifiers are an area of active investigation for treatment of many human diseases, including cancer, and therefore `piggybacking' current anti-parasitic drug development efforts onto other active areas of drug development may be feasible as we achieve more insights into the importance of parasite epigenetic gene regulation.

Histone modifying enzymes

Candidate histone modifying enzymes and chromatin remodelers can be identified in bioinformatic analyses of the protozoan genomes [21]. Although the function of most of the predicted proteins has not been established, for many, the predicted function and specificity has been inferred bioinformatically [9, 10, 33]. The activity of the candidate modifying enzymes is predicted to induce changes in the chromatin structure and, therefore, could potentially affect the transcriptional profile of the organism. Opening of the chromatin to make regions accessible for transcription factors and the RNA polymerase machinery is necessary, although perhaps not sufficient, for gene transcription. The specialized functions of chromatin remodelers have been implicated in critical biological processes like antigenic variation, life cycle development, and cell cycle control.

Another evolving area of investigation is the modification by `histone modifiers' of non-histone substrates [34, 35], including RNA binding proteins, tRNA, ribosomal proteins and transcription factors. These activities could impact gene expression by affecting macromolecular complex formation, heterochromatin assembly, RNA stability, or efficiency of translation. For example, TgPRMT1, originally described as methylating H4R3 [26], is associated with and methylates TgAgo, the *T. gondii* Argonaute orthologue, to regulate recruitment of a Tudor nuclease [36, 37]. Because of the expanding list of significant non-

histone substrates, studies performed with inhibitors of `histone modifiers' should be interpreted with caution.

Plasmodium falciparum

Both ChIP-chip and ChIP-seq have been used to characterize the epigenome of *P. falciparum*. The *P. falciparum* genome is AT-rich, and interpretation of the results of these studies may be complicated by technical obstacles presented by the extremely AT-rich intergenic regions (approximately 90%). Such regions may be amplified, cross-hybridize on genomic arrays or be difficult to map unambiguously onto the genome when high throughput sequencing approaches are used. Two groups have reported that nucleosome occupancy is low in intergenic regions [38, 39], whereas another group, using techniques less subject to amplification artifacts, has found that intergenic regions have H2AZ, H3K4me3, and H3K9ac [40]. Differences in techniques probably explain the divergent results, but the differences of many results in protozoa from what has been observed in other eukaryotes underscores the importance of careful systematic studies in each biological system.

H3K4me3 and H3K9ac are enriched in a large proportion of the *P. falciparum* genome including in open reading frames for much of the erythrocytic cycle [31, 40], unlike what is observed in other eukaryotes or even the Apicomplexan *T. gondii* [30]. In ring stages, the H3K4me3 and H3K9ac marks are associated with intergenic regions of both active and inactive genes and do not correlate with transcriptional activity [31, 40] as inferred by comparison with steady state mRNA levels. As parasites progress to schizont stages, both constitutively active and stage-specific genes have enrichment of these marks at the 5' region [31, 40].

H3K9me3, a marker of heterochromatin, is associated with central *var* intrachromosomal clusters and with subtelomeric regions that are enriched in variant surface antigen families such as *var, stevor* and *rifins* [41], but this mark is not associated with developmentally silenced genes such as *csp* or *trap/ssp2* [31]. The H3K9me3 in chromosome end clusters and internally is bound by PfHP1 (heterochromatin protein 1) [42, 43], a conserved chromodomain protein that recognizes H3K9me3 and is associated with telomeres and centromeres in other organisms. Neither HP1 nor H3K9me3 are associated with centromeres in *Plasmodium* [41–43].

Two major families of histone acetyltransferases (HAT) have been characterized in *Plasmodium* parasites: MYST and GCN5. These are lysine methyltransferases (or KAT). PfMYST appears to be essential, and overexpression is deleterious for parasites [44]. Recombinant PfMYST is capable of acetylating multiple lysines on H4, marks commonly associated with gene activation. Overexpression disrupts the cell cycle, resulting in accelerated schizogony and fewer merozoites per infected red blood cell [44].

PfGCN5 plays a key role in the acetylation of H3K9 and K14 [45]. As is the case in other species, these are essential marks of gene activation. While PfGCN5 is thought to be essential, parasites can survive treatment with curcumin [45], selected as an inhibitor of GCN5, but now appreciated to have other targets. After treatment with curcumin, *Plasmodium* parasites showed greatly reduced H3K9 acetylation genome-wide and several reference genes were downregulated.

Opposing the action of HAT enzymes are histone deactylases (HDAC). Most of these enzymes have not been characterized but treatment of *P. falciparum* parasites with apicidin, a HDAC inhibitor, led to hyperacetylation of histones and major dysregulation of the transcriptome in asexual blood stages [46]. Sir2 is an NAD+-dependent HDAC (or sirtuin)

that was first identified in yeast as important for silencing telomeric genes. Since NAD+ is a key cofactor and metabolite, Sir2 may act as a sensor for environmental or nutrient changes. PfSir2A localizes to heterochromatin in the nuclear periphery as well as to the nucleolus and its activity has been linked to spreading of heterochromatin regions in telomeric regions [5, 6].

In *Plasmodium*, there are two Sir2 orthologs, PfSir2A and PfSir2B [47] that work in concert to regulate *var* silencing. *Plasmodium* parasites that lack PfSir2A and B are unable to effectively silence undesired *var* genes, and transcript levels of *var* genes are generally elevated [47, 48]. While regulation of *var* gene expression is seriously compromised, parasites are still able to switch *var* genes, suggesting that other regulatory mechanisms exist and work in concert with Sir2 [47]. Not all *var* genes are subtelomeric, and while centrally located *var* localize to heterochromatin, PfSir2a deletion does not affect expression of centrally located *vars* as markedly [5, 6]. Thus there may be epigenetic factors that regulate subtelomeric *var* because of their chromosomal location and other factors that regulate silencing and activation of centrally located *var*. However, there must be some intersection of the regulation of these different sets of *var* genes since *P. falciparum* is able to mutually exclude *var* gene expression. This additional level of regulation may be accomplished by targeting of *var* genes to specialized nuclear compartments. Silenced *var* loci are organized in several perinuclear clusters, and a protein complex that includes actin and PfAP2 PF11_0091 recognizes a conserved element within the *var* intron that is important for recruitment of silenced *var* to these regions [28]. Actin is critical for the spatial repositioning from silent clusters to active. Active *var* are also localized to the nuclear periphery [28].

Protein arginine methyltransferases (PRMT) are responsible for histone methylation in many eukaryotes. PRMTs methylate histones but have a number of other nonhistone substrates including RNA binding proteins (see [34] for an excellent review of PRMT in parasites). *P. falciparum* has at least three conserved PRMTs annotated in PlasmoDB [9]. Of these, only PfPRMT1 has been biochemically characterized. PfPRMT1 is capable of targeted methylation of arginine 3 on the N terminus of histone 4, a mark associated with gene activation [49]. Genetic studies will be required to determine the extent of the role of PfPRMT1 in parasite survival, however, treatment of cultured parasites with PRMT enzyme inhibitors also inhibits *Plasmodium* growth [49].

Methylation of lysine residues on histones is linked to both transcriptional upregulation and silencing, depending on the targeted residue. In general, there are two major families of histone lysine methyltransferases (KMT): those which contain the SET domain and those related to Dot1. In *P. falciparum* (and *T. gondii*), no Dot1 orthologue is evident, but there appear to be ten SET genes (PfSET1–10) [9]. The activity of KMTs is opposed by lysine demethylases. The Apicomplexa encode orthologues of both major groups of these enzymes, lysine specific demethylases (LSD1) and the Jumonji C histone demethylases (JHDM). The specific functions of these *Plasmodium* enzymes have not been described, but presumably these enzymes oppose the KMT activity upon the various substrates of lysine methylases, and thus demethylases should be important for the transition of a *var* gene to either the active or silenced state.

Biochemically defining the activity of all of these enzymes in *Plasmodium* will be a significant undertaking. Volz *et al*. have localized many nuclear proteins providing complementary insight into the potential functions [50]. Cui *et al*. were able to demonstrate recombinant PfSET8 methylates H4K20 [51]. Volz *et al*. have identified PfSET10 as a H3K4me4 transferase that localizes to a specialized region of the nucleus with the active *var* gene, keeping the locus in a poised state during the *P. falciparum* cell cycle [52]. The biochemistry of most of the remaining KMT enzymes remains unconfirmed; however, the

site specificities of several have been predicted based on similarity to SET enzymes in other organisms [51].

In addition to the four canonical histones, *P. falciparum* possesses CenH3, H3.3, H2A.Z, and H2Bv [19]. PfH2A.Z stably localizes to the entire intergenic space regardless of the transcriptional status of the genes [40]. This contrasts with other organisms, including *T. gondii* and *T. brucei* (Figure 3) [18, 53], where H2A.Z appears dynamically during gene activation and typically occupies only the transcription start site [54]. PfH2A.Z has a distinct role in *var* gene activation. In this case, H2AZ is exchanged for canonical H2A during activation, opposing the *var* silencing activity of PfSir2A [55]

Toxoplasma gondii

T. gondii is an apicomplexan parasite, like *Plasmodium*. Bioinformatic analysis of the *T. gondii* genome predicts epigenetic machinery that is similar to *Plasmodium* [10, 56], but characterization of the *T. gondii* epigenome has shown some surprising differences, which may reflect differences in biology or host range. *In vivo, Toxoplasma* responds to stress from the host by transitioning from the tachyzoite to a latent bradyzoite form. The bradyzoites form tissue cysts that can persist for years within the host, and these latent cysts can reactivate into tachyzoites resulting in a severe acute infection. The role of epigenetic regulation in the developmental transition of *T. gondii* from tachyzoites to bradyzoite forms is an area of intense investigation. *T. gondii* is unusual in that there appears to be continuous interconversion between tachyzoite and bradyzoite forms within latently infected hosts. Although *T. gondii* encodes a large family of stage-specific surface antigens, it does not undergo classic antigenic variation as seen in *Plasmodium* or *T. brucei*.

The function and distribution of histone modifications and variant histones of *T. gondii* more closely resembles that of classic model organisms than *Plasmodium* does, but the *T. gondii* epigenome also has unique features. ChIP-chip studies of *T. gondii* revealed that the classic gene activation marks H4ac, H3K9ac, and H3K4me3 co-localize and mark the promoters of actively transcribed genes [30], but *T. gondii* has a relative paucity of heterochromatin [3]. The euchromation marks correlate with active genes, as determined by transcriptome analysis of steady state mRNA levels [30]. Stage specific bradyzoite and sporozoite promoters are not marked with these euchromatin histone PTMs [30] nor are they marked by other known heterochromatin marks such as H3K9me3 [3]. In contrast to *P. falciparum, T. gondii* H3K9me2 and H3K9me3 are not associated with subtelomeric heterochromatin regions, but instead flank the centromeric histone CenH3 within the centromeric heterochromatin region on each chromosome [3]. These marks are not associated with centromeres in *Plasmodium,* and *T. gondii* does not appear to have classic heterochromatin in its subtelomeric regions [3]. The variant histones H2A.Z and H2Bv are associated with promoters of active genes (Figure 3), rather than representing a stable intergenic mark as in *Plasmodium* [18]. As observed in other eukaryotes, TgH2A.X is associated with DNA damage [18] and its expression is increased under stress conditions that favor bradyzoite formation.

There are two MYST family members present in *T. gondii*, TgMYST-A and TgMYST-B, both of which have histone H4 acetylation activity [57, 58]. TgMYST-B also appears to regulate the rate of parasite replication as well as aid in DNA repair via upregulation of TgATM kinase [57, 58]. By analogy to yeast, MYST HATs have been hypothesized to be important for nutritional sensing and developmental transitions in response to nutrient starvation [59].

Unlike *Plasmodium*, which encodes a single GCN5 orthologue, there are two lysine histone acetyltransferases in *T. gondii* orthologous to GCN5 [60]. While TgGCN5-B appears to be

essential (W.J. Sullivan, personal communication), TgGCN5-A may be deleted without an obvious phenotype in tissue culture [60, 61]. However, TgGCN5-A is upregulated during alkaline stress and Δ*gcn5*-a parasites do not properly modulate their transcriptional profile when exposed to elevated pH, a classic experimental condition used to induce bradyzoite formation. Thus, TgGCN5-A may be important in *T. gondii* stress response and life cycle progression [61]. The molecular basis for bradyzoite differentiation remains an open question in the field and the role of TgGCN5-A, other stress responsive chromatin regulators, and their co-effectors action continues to be a promising avenue of inquiry.

The activity and functions of most *Toxoplasma* histone deacetylases (HDAC) have not been fully explored. One study suggests that HDAC3 is an essential regulator of gene expression and parasite differentiation, and its gene cannot be disrupted [26]. When TgHDAC3 activity is inhibited with an enzyme-specific drug (FR235222), H4 becomes hyper-acetylated and the parasites appear to differentiate into latent form bradyzoites [62]. This effect was not seen when the investigators tested a strain that had been transfected with a HDAC3 allele that was resistant to the compound, providing strong evidence that the effect was specific for HDAC3. What is especially interesting about this result is that it suggests that the activity of HDAC3 actively prevents bradyzoite differentiation. FR235222 also affected growth of *P. berghei* and *P. falciparum* suggesting that *Plasmodium* HDAC3 may be effectively targeted by the compound.

There are five PRMT in *T. gondii* of which two have been characterized, PRMT1 and CARM1 (or PRMT4). Recombinant PRMT1 could methylate H4R3 and CARM1 could methylate H3R17 in *in vitro* assays [26], consistent with the activity of orthologues in yeast and *Plasmodium*. CARM1 appears to be essential and treatment of tachyzoites with AMI-1, a PRMT inhibitor with activity against CARM1 but not PRMT1, induced bradyzoite formation [26]. ChIP-chip studies with H3R17 antibodies, showed the association of H3R17, a marker of gene activation in other systems, with a subset of active promoters that are active in tachyzoites and bradyzoites[30]. These studies support the hypothesis that CARM1 activity on histones may be important for tachyzoite-bradyzoite transition, although a direct role for CARM1 has not been proven and it may have additional substrates that are important for this transition.

PRMT1 appears to have important functions beyond histone modification. This PRMT is not essential [37] (K. el Bissati *et al.*, unpublished), but appears to have a role in regulation of cell division. PRMT1 co-immunoprecipitates with TgAgo, the Argonaute orthologue that is implicated in RNAi in *T. gondii* [36, 37]. Intriguingly, TgPRMT1 appears to methylate the N-terminal RGG-repeat domain of TgAgo, and this methylation correlates with the recruitment of Tudor staphylococcal nuclease, an endonuclease with more robust activity than TgAgo [37]. While PRMT1 may methylate histones, its nonhistone substrates may be more significant for survival of the parasite.

T. gondii encodes 20 proteins with the lysine methyltransferase (KMT) SET domain. While some of these SET proteins are orthologous to known histone modifying enzymes [10] or have PHD, chromo or bromo reader domains that suggest a nuclear function, most of these KMTs have not been functionally characterized. *T. gondii* protein methylases may have other roles beyond histone methylation and regulation of gene expression. Some commercial histone methyl antibodies cross react with other structures in *T. gondii*, including the apical region [63, 64] and at least one SET (i.e. AKMT or apical lysine methyl transferase) is implicated in regulation of apical cytoskeletal structures [65]. *T. gondii* tubulin is also methylated, and the enzymes responsible for this modification have not yet been identified [64].

TgSET8, orthologous to PfSET8, also mono-, di-, or trimethylates H4K20 [63]. Both TgSET8 and the H4K20 monomethyl mark are cell cycle regulated [63]. TgSET8 protein expression peaks during the G1 phase and is degraded during mitosis, while H4K20me1 peaks during mitosis [63]. H4K20me1/3 are established markers of heterochromatin in metazoans and are hypothesized to serve the same function in Apicomplexa [63].

Trypanosoma brucei

Regulation of transcription in *T. brucei* (and all kinetoplastids) is highly divergent from other eukaryotes and remains poorly understood [66]. *T. brucei* and the other kinetoplastids organize genes into polycistronic transcription units (PTU). The highly expressed mRNAs for blood stage variant surface antigen VSG and the procyclic surface protein procyclin are transcribed by RNA Polymerase I (PolI), while most other mRNAs are transcribed by RNA Polymerase II (PolII). The polycistronic transcripts are trans-spliced to create a mature mRNA with a 5' spliced leader sequence. Conventional PolII promoters have not been identified, and sites of Pol II initiation and termination have been mapped to strand switch sites in both *T. brucei* and *Leishmania* [67, 68]. Despite the absence of transcriptional regulation, steady state mRNAs vary during different life cycle stages [67, 69], indicating post-transcriptional mechanisms are the major mode of gene regulation.

The histones of the kinetoplastids are highly divergent from most histone sequences (Figure S1), and thus nearly all of the reagents that have proven useful across species are not useful for the study of the epigenome of kinetoplastids. Several RNA-seq and ChIP-seq studies have illuminated aspects of gene expression and the epigenome of kinetoplastids. *T. brucei* possesses H2A.Z as well as three parasite specific histone variants: H3V, H4V, and H2Bv [70, 71]. H2AZ, H2BV, and H4K10Ac are all enriched upstream of PTUs, as is a bromodomain-containing protein, TbBDF3 [53]. H3V and H4V are both enriched at likely transcription termination sites [53] although most of H3v is associated with telomeres, not with transcription termination. RNAi knockdown of TbBDF3 leads to severe growth defects [53]. *T. brucei* encodes five PRMT but there is no evidence for histone arginine methylation, and these PRMT appear to have roles in regulation of RNA metabolism [34].

T. brucei lacks any apparent GCN5 homologue, but does possess three MYST family members, TbHAT1, TbHAT2, and TbHAT3 [72]. TbHAT1 and 2 are essential, and RNAi knockdown disrupts the cell cycle and reduces parasite growth [72]. Interestingly, knockdown of TbHAT1 results in increased expression of a reporter gene near telomeres. Telomeres are often associated with densely packed heterochromatin, which silences local genes. TbHAT3, while not essential for parasite survival, is responsible for acetylation of H4K4, Since most of chromatin associated H4K4 is acetylated, these data indicate that, contrary to expectation, H4K4 acetylation is not critical as an important transcription regulatory mark [73].

Like *P. falciparum*, *T. brucei* relies on antigenic variation in order to evade host immunity. This is achieved through periodically switching the variant surface glycoprotein (VSG) that the parasite expresses. These VSG expression sites are subtelomeric, and only one of the *VSG* genes is expressed at a time, while the other *VSG* genes are silenced. TbHDAC1 and TbHDAC3 are important for regulating telomeric VSG silencing [74]. TbHDAC1 works against VSG silencing in the blood stage form of the parasite. TbHDAC3 silences VSG transcription and prevents more than one VSG from being expressed at the same time. RNAi depletion of either enzyme results in a growth defect [74]. There are 3 sirtuin orthologues with Sir2rp1 being nuclear localized, and Sir2rp2 and Sir2rp3 are localized to the mitochondria [75]. Sir2rp1 is implicated in control of DNA repair and PolI mediated telomeric silencing but does not appear to play a role in antigenic variation of *VSG* loci [75].

By contrast, tbRAP1 does play a critical role in both telomeric integrity and silencing VSG expression sites [76].

DOT1 enzymes are methyltransferases targeting H3K79 in most organisms. Two DOT1 homologues exist in *T. brucei*, TbDOT1A and TbDOT1B [77]. DOT1A is responsible for dimethylation of H3K76 (equivalent to H3K79 in other eukaryotes) and appears to be essential in the bloodstream form of the parasites [77, 78]. Dimethylated TbH3K76 appears only during mitosis [77]. When TbDOT1A is depleted, severe growth and cell cycle defects arise with a proportion of the normally diploid cells appearing to be haploid [77, 78]. These results suggest TbDOT1A is responsible for preventing cytokinesis before DNA replication is complete.

TbDOT1B trimethylates TbH3K76 and does not require the histone to be previously methylated (i.e. it is independent of TbDOT1A). The enzyme does not appear to be cell cycle regulated and can be disrupted without an obvious phenotype during the bloodstream stage [77]. TbDOT1B is essential for successful life cycle progression from bloodstream trypomastigotes to the procyclic form in the tsetse fly midgut [77], and it is also required for maintenance of strict VSG silencing and transcription switching of VSG [78].

Post-translational modification of DNA: methylation and glycosylation

In higher eukaryotes, epigenetic gene regulation involves both DNA methylation and post-translational modification of histones. Methylation of DNA is a common mechanism for epigenetic gene silencing in plants and mammals, but is not universally present in unicellular organisms. Methylation of cytosines typically occurs near promoters and represses gene expression. Several studies have been unable to detect methylated cytosine in Apicomplexa including *T. gondii*, *Plasmodium* and *Cryptosporidium* [79, 80], although the genomes do encode candidate methyltransferases that could potentially methylate DNA or RNA. It is possible that DNA methylation is restricted to a small number of loci that were not detected by the methodology used, but current data suggest that DNA methylation is not a major mechanism to regulate gene expression in the Apicomplexa.

By contrast, *T. brucei* does have methylated DNA detectable by mass spectrometry, as well as a DNA methyltransferase [81], although the significance of methylated DNA has not been explored. *Entamoeba histolytica* DNA is methylated and is hypothesized to control transposon-like repetitive DNA [82, 83]. Differential DNA methylation also has been linked to differences in virulence in *Entamoeba* strains [83, 84].

Glycosylation of thymine, called `J' base or β-d-glucosyl-hydroxymethyluracil occurs in *T. brucei*, *Trypanosoma cruzi*, as well as *Leishmania* [85]. In *Trypanosoma*, J occurs only during the bloodstream form of the parasites and accounts for about one percent of total genomic thymine [86]. J was initially hypothesized to play a role in telomeric repeat silencing and variation of sub-telomeric VSG antigens in *T. brucei*. More recent work has identified JBP1 and JBP2, a SWI2/SNF2-like protein, [87, 88], the enzymes essential for J base biosynthesis. Although J is enriched in telomeric regions, J does not appear to be critical for VSG regulation, and J is not essential for *T. brucei* viability [89]. More recent studies show that DNA glycosylation marks the regions between *Trypanosoma cruzi* polycistronic transcription units and regulates PolII transcription [90, 91]. Knockout of these enzymes leads to increased RNA Pol II activity and drastically decreased *T. cruzi* virulence Intriguingly, ongoing studies in *Leishmania* indicate that Base J has a more crucial role in parasite survival than in *T. brucei* (P. Borst, personal communication).

Concluding remarks

As DNA sequencing and proteomics technologies have improved, we have been able to develop a more detailed understanding of the role of epigenetics in protozoan parasite biology. While many of the basic themes are similar to those gleaned from studies of model organisms, protozoan parasites have evolved a surprisingly diverse repertoire of mechanisms to modify chromatin structure to regulate gene expression. These differences undoubtedly reflect the different evolutionary pressures and biological niches occupied by each organism. The challenge for the future will be to integrate the large amounts of data acquired from high throughput datasets to develop an integrated and mechanistic view of the biological processes regulated by epigenetics in each of these organisms.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank William Sullivan for critical review of the manuscript. MMC was supported by the Training Program in Cellular and Molecular Biology and Genetics, funded by NIH T32 GM007491 awarded to the Albert Einstein College of Medicine. Some of this work will be published in a thesis submitted in partial fulfillment of the requirements for a Doctor of Philosophy conferred by the Sue Golding Graduate Division of the Albert Einstein College of Medicine (M.M.C.). This work was partially supported by NIH grants R01 AI087625, RC4 AI092801 (to K.K.) and the Einstein-Montefiore Center for AIDS Research, funded by P30AI051519.

Glossary

ChIP	Proteins associated with DNA are immunoprecipitated as a protein-DNA complex, and the abundance of associated DNA is compared to input DNA. X-ChIP, which involves crosslinking of DNA-protein complexes with formaldehyde, or N-ChIP, which involves native immunoprecipitation conditions (frequently after micrococcal nuclease digestion), are common approaches.
ChIP-chip or ChIP-seq	Techniques by which proteins associated with DNA are immunoprecipitated and the associated DNA identified by hybridization to genome-wide arrays (ChIP-chip) or high through-put sequencing (ChIP-seq).
Chromatin	The complex of DNA and associated proteins, including histones, that makes up nuclear DNA. Histones are highly conserved proteins, and many reagents developed for study in yeast or mammalian cells can be adapted for use in parasites.
Epigenetics	Heritable changes in an organism that alter gene expression levels without altering the DNA sequence. Commonly taken to mean changes in chromatin structure involving DNA methylation or histone modifications, although this is not strictly true (see Box 1).
Epigenome	Epigenetic modifications on a genome wide scale.
High-throughput sequencing (massively	Numerous platforms by which millions of short sequencing reads (35–100 bp) are performed and assembled to survey DNA or RNA

parallel sequencing)	abundance in a sample. Among the most commonly used platforms are Illumina/Solexa, SoLiD/ABI, and 454/Roche (longer but fewer reads).
Histone	Small basic proteins that make the protein core that DNA is wound around. The classic or canonical histones are H2A, H2B, H3 and H4. Histone variants include H2A.Z, H2A.X (DNA repair), H3.3 (active genes) and CenH3 (or CenPA, centromeric), as well as parasite-specific histones. Trypanosomatid histones are very divergent from those of other eukaryotes (See Supplementary materials Figure S1 for alignments).
Histone code hypothesis	The idea that the combination of post-translational modifications (PTM) on histones, primarily on the histone N termini, confers epigenetic information that determines transcriptional activation or repression. Among the important modifications are: (i) methylation, acetylation, ubiquitination of lysine; (ii) methylation of arginine; and (iii) phosphorylation of serine. The combination and sequence of addition of PTM affects the interaction of chromatin with macromolecular complexes. Many new modifications have been reported in the past two years, increasing the potential complexity of the histone code.
Nucleosome	The basic unit of chromatin structure that consists of 147 bp of DNA wound twice around an octamer of histones. The strict definition includes the DNA linking the histone-DNA unit to the next nucleosome. Linker DNA may also have a specialized H1 linker histone that aids in DNA compaction. The histone core is composed of two H2A–H2B heterodimers and one H3–H4 heterotetramer.
Post-translational modifications (PTM)	Enzymatic modification of amino acid residues that occur after translation. These often occur in combination or in specific sequences that affect histone function, histone lifespan or interaction with other proteins. Although PTM that affect epigenetic processes are classically studied on histones, recent studies indicate that PTM of other proteins including transcription factors and chromatin remodelers have significant impact.

References

1. Atayde VD, et al. The emerging world of small silencing RNAs in protozoan parasites. Trends in parasitology. 2011; 27:321–327. [PubMed: 21497553]
2. Lye LF, et al. Retention and loss of RNA interference pathways in trypanosomatid protozoans. PLoS Pathog. 2010; 6:e1001161. [PubMed: 21060810]
3. Brooks CF, et al. *Toxoplasma gondii* sequesters centromeres to a specific nuclear region throughout the cell cycle. Proc Natl Acad Sci U S A. 2011; 108:3767–3772. [PubMed: 21321216]
4. Ralph SA, et al. Antigenic variation in *Plasmodium falciparum* is associated with movement of var loci between subnuclear locations. Proc Natl Acad Sci U S A. 2005; 102:5414–5419. [PubMed: 15797990]
5. Freitas-Junior LH, et al. Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. Cell. 2005; 121:25–36. [PubMed: 15820676]
6. Duraisingh MT, et al. Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in *Plasmodium falciparum*. Cell. 2005; 121:13–24. [PubMed: 15820675]

7. Navarro M, Gull K. A pol I transcriptional body associated with VSG mono-allelic expression in *Trypanosoma brucei*. Nature. 2001; 414:759–763. [PubMed: 11742402]
8. Figueiredo LM, et al. Epigenetic regulation in African trypanosomes: a new kid on the block. Nat Rev Microbiol. 2009; 7:504–513. [PubMed: 19528957]
9. Cui L, Miao J. Chromatin-mediated epigenetic regulation in the malaria parasite *Plasmodium falciparum*. Eukaryotic cell. 2010; 9:1138–1149. [PubMed: 20453074]
10. Bougdour A, et al. Chromatin modifications: implications in the regulation of gene expression in *Toxoplasma gondii*. Cell Microbiol. 2010; 12:413–423. [PubMed: 20109158]
11. Dixon SE, et al. A decade of epigenetic research in *Toxoplasma gondii*. Molecular and biochemical parasitology. 2010; 173:1–9. [PubMed: 20470832]
12. Mandava V, et al. Histone modifications in *Trypanosoma brucei*. Molecular and biochemical parasitology. 2007; 156:41–50. [PubMed: 17714803]
13. Janzen CJ, et al. Unusual histone modifications in *Trypanosoma brucei*. FEBS letters. 2006; 580:2306–2310. [PubMed: 16580668]
14. Trelle MB, et al. Global histone analysis by mass spectrometry reveals a high content of acetylated lysine residues in the malaria parasite *Plasmodium falciparum*. Journal of proteome research. 2009; 8:3439–3450. [PubMed: 19351122]
15. Jenuwein T, Allis CD. Translating the histone code. Science. 2001; 293:1074–1080. [PubMed: 11498575]
16. Sidoli S, et al. Proteomics in chromatin biology and epigenetics: Elucidation of post-translational modifications of histone proteins by mass spectrometry. Journal of proteomics. 2012
17. Egelhofer TA, et al. An assessment of histone-modification antibody quality. Nature structural & molecular biology. 2011; 18:91–93.
18. Dalmasso MC, et al. Toxoplasma H2A variants reveal novel insights into nucleosome composition and functions for this histone family. Journal of molecular biology. 2009; 392:33–47. [PubMed: 19607843]
19. Miao J, et al. The malaria parasite *Plasmodium falciparum* histones: organization, expression, and acetylation. Gene. 2006; 369:53–65. [PubMed: 16410041]
20. Mandava V, et al. Trypanosome H2Bv replaces H2B in nucleosomes enriched for H3 K4 and K76 trimethylation. Biochemical and biophysical research communications. 2008; 368:846–851. [PubMed: 18261990]
21. Iyer LM, et al. Comparative genomics of transcription factors and chromatin proteins in parasitic protists and other eukaryotes. International journal for parasitology. 2008; 38:1–31. [PubMed: 17949725]
22. Tan M, et al. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell. 2011; 146:1016–1028. [PubMed: 21925322]
23. Balaji S, et al. Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. Nucleic Acids Res. 2005; 33:3994–4006. [PubMed: 16040597]
24. Campbell TL, et al. Identification and genome-wide prediction of DNA binding specificities for the ApiAP2 family of regulators from the malaria parasite. PLoS Pathog. 2010; 6:e1001165. [PubMed: 21060817]
25. LaCount DJ, et al. A protein interaction network of the malaria parasite *Plasmodium falciparum*. Nature. 2005; 438:103–107. [PubMed: 16267556]
26. Saksouk N, et al. Histone-modifying complexes regulate gene expression pertinent to the differentiation of the protozoan parasite *Toxoplasma gondii*. Molecular and cellular biology. 2005; 25:10301–10314. [PubMed: 16287846]
27. Flueck C, et al. A major role for the *Plasmodium falciparum* ApiAP2 protein PfSIP2 in chromosome end biology. PLoS Pathog. 2010; 6:e1000784. [PubMed: 20195509]
28. Zhang Q, et al. A critical role of perinuclear filamentous actin in spatial repositioning and mutually exclusive expression of virulence genes in malaria parasites. Cell host & microbe. 2011; 10:451–463. [PubMed: 22100161]

29. Wright JR, et al. Histone H3 trimethylated at lysine 4 is enriched at probable transcription start sites in *Trypanosoma brucei*. Molecular and biochemical parasitology. 2010; 172:141–144. [PubMed: 20347883]
30. Gissot M, et al. Epigenomic modifications predict active promoters and gene structure in *Toxoplasma gondii*. PLoS Pathog. 2007; 3:e77. [PubMed: 17559302]
31. Salcedo-Amaya AM, et al. Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of *Plasmodium falciparum*. Proc Natl Acad Sci U S A. 2009; 106:9655–9660. [PubMed: 19497874]
32. Rivero FD, et al. Disruption of antigenic variation is crucial for effective parasite vaccine. Nature medicine. 2010; 16:551–557. 551p following 557.
33. Sullivan WJ Jr. et al. Histones and histone modifications in protozoan parasites. Cell Microbiol. 2006; 8:1850–1861. [PubMed: 17026479]
34. Fisk JC, Read LK. Protein arginine methylation in parasitic protozoa. Eukaryotic cell. 2011; 10:1013–1022. [PubMed: 21685318]
35. Choudhary C, et al. Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science. 2009; 325:834–840. [PubMed: 19608861]
36. Braun L, et al. A complex small RNA repertoire is generated by a plant/fungal-like machinery and effected by a metazoan-like Argonaute in the single-cell human parasite *Toxoplasma gondii*. PLoS Pathog. 2010; 6:e1000920. [PubMed: 20523899]
37. Musiyenko A, et al. PRMT1 methylates the single Argonaute of *Toxoplasma gondii* and is important for the recruitment of Tudor nuclease for target RNA cleavage by antisense guide RNA. Cell Microbiol. 2012
38. Westenberger SJ, et al. Genome-wide nucleosome mapping of *Plasmodium falciparum* reveals histone-rich coding and histone-poor intergenic regions and chromatin remodeling of core and subtelomeric genes. BMC genomics. 2009; 10:610. [PubMed: 20015349]
39. Ponts N, et al. Nucleosome landscape and control of transcription in the human malaria parasite. Genome Res. 2010; 20:228–238. [PubMed: 20054063]
40. Bartfai R, et al. H2A.Z demarcates intergenic regions of the *Plasmodium falciparum* epigenome that are dynamically marked by H3K9ac and H3K4me3. PLoS Pathog. 2010; 6:e1001223. [PubMed: 21187892]
41. Lopez-Rubio JJ, et al. Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. Cell host & microbe. 2009; 5:179–190. [PubMed: 19218088]
42. Perez-Toledo K, et al. *Plasmodium falciparum* heterochromatin protein 1 binds to trimethylated histone 3 lysine 9 and is linked to mutually exclusive expression of var genes. Nucleic Acids Res. 2009; 37:2596–2606. [PubMed: 19270070]
43. Flueck C, et al. *Plasmodium falciparum* heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. PLoS Pathog. 2009; 5:e1000569. [PubMed: 19730695]
44. Miao J, et al. The MYST family histone acetyltransferase regulates gene expression and cell cycle in malaria parasite *Plasmodium falciparum*. Mol Microbiol. 2010; 78:883–902. [PubMed: 20807207]
45. Cui L, et al. PfGCN5-mediated histone H3 acetylation plays a key role in gene expression in *Plasmodium falciparum*. Eukaryotic cell. 2007; 6:1219–1227. [PubMed: 17449656]
46. Chaal BK, et al. Histone deacetylases play a major role in the transcriptional regulation of the *Plasmodium falciparum* life cycle. PLoS Pathog. 2010; 6:e1000737. [PubMed: 20107518]
47. Tonkin CJ, et al. Sir2 paralogues cooperate to regulate virulence genes and antigenic variation in *Plasmodium falciparum*. PLoS biology. 2009; 7:e84. [PubMed: 19402747]
48. Merrick CJ, et al. The effect of *Plasmodium falciparum* Sir2a histone deacetylase on clonal and longitudinal variation in expression of the var family of virulence genes. International journal for parasitology. 2010; 40:35–43. [PubMed: 19666023]
49. Fan Q, et al. Characterization of PRMT1 from *Plasmodium falciparum*. The Biochemical journal. 2009; 421:107–118. [PubMed: 19344311]

50. Volz J, et al. Potential epigenetic regulatory proteins localise to distinct nuclear sub-compartments in *Plasmodium falciparum*. International journal for parasitology. 2010; 40:109–121. [PubMed: 19765590]
51. Cui L, et al. Histone lysine methyltransferases and demethylases in *Plasmodium falciparum*. International journal for parasitology. 2008; 38:1083–1097. [PubMed: 18299133]
52. Volz JC, et al. PfSET10, a *Plasmodium falciparum* Methyltransferase, Maintains the Active var Gene in a Poised State during Parasite Division. Cell host & microbe. 2012; 11:7–18. [PubMed: 22264509]
53. Siegel TN, et al. Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. Genes & development. 2009; 23:1063–1076. [PubMed: 19369410]
54. Yuan G, Zhu B. Histone variants and epigenetic inheritance. Biochimica et biophysica acta. 2011
55. Petter M, et al. Expression of P. falciparum var genes involves exchange of the histone variant H2A.Z at the promoter. PLoS Pathog. 2011; 7:e1001292. [PubMed: 21379342]
56. Sullivan WJ Jr. Hakimi MA. Histone mediated gene activation in *Toxoplasma gondii*. Molecular and biochemical parasitology. 2006; 148:109–116. [PubMed: 16644030]
57. Smith AT, et al. MYST family histone acetyltransferases in the protozoan parasite *Toxoplasma gondii*. Eukaryotic cell. 2005; 4:2057–2065. [PubMed: 16339723]
58. Vonlaufen N, et al. MYST family lysine acetyltransferase facilitates ataxia telangiectasia mutated (ATM) kinase-mediated DNA damage response in *Toxoplasma gondii*. J Biol Chem. 2010; 285:11154–11161. [PubMed: 20159970]
59. Hutson SL, et al. T. gondii RP promoters & knockdown reveal molecular pathways associated with proliferation and cell-cycle arrest. PLoS One. 2010; 5:e14057. [PubMed: 21124925]
60. Bhatti MM, et al. Pair of unusual GCN5 histone acetyltransferases and ADA2 homologues in the protozoan parasite *Toxoplasma gondii*. Eukaryotic cell. 2006; 5:62–76. [PubMed: 16400169]
61. Naguleswaran A, et al. *Toxoplasma gondii* lysine acetyltransferase GCN5-A functions in the cellular response to alkaline stress and expression of cyst genes. PLoS Pathog. 2010; 6:e1001232. [PubMed: 21179246]
62. Bougdour A, et al. Drug inhibition of HDAC3 and epigenetic control of differentiation in Apicomplexa parasites. The Journal of experimental medicine. 2009; 206:953–966. [PubMed: 19349466]
63. Sautel CF, et al. SET8-mediated methylations of histone H4 lysine 20 mark silent heterochromatic domains in apicomplexan genomes. Molecular and cellular biology. 2007; 27:5711–5724. [PubMed: 17562855]
64. Xiao H, et al. Post-translational modifications to *Toxoplasma gondii* alpha- and beta-tubulins include novel C-terminal methylation. Journal of proteome research. 2010; 9:359–372. [PubMed: 19886702]
65. Heaslip AT, et al. The motility of a human parasite, *Toxoplasma gondii*, is regulated by a novel lysine methyltransferase. PLoS Pathog. 2011; 7:e1002201. [PubMed: 21909263]
66. Rudenko G. Epigenetics and transcriptional control in African trypanosomes. Essays in biochemistry. 2010; 48:201–219. [PubMed: 20822495]
67. Siegel TN, et al. Genome-wide analysis of mRNA abundance in two life-cycle stages of *Trypanosoma brucei* and identification of splicing and polyadenylation sites. Nucleic Acids Res. 2010; 38:4946–4957. [PubMed: 20385579]
68. Thomas S, et al. Histone acetylations mark origins of polycistronic transcription in *Leishmania major*. BMC genomics. 2009; 10:152. [PubMed: 19356248]
69. Jensen BC, et al. Widespread variation in transcript abundance within and across developmental stages of *Trypanosoma brucei*. BMC genomics. 2009; 10:482. [PubMed: 19840382]
70. Lowell JE, Cross GA. A variant histone H3 is enriched at telomeres in *Trypanosoma brucei*. Journal of cell science. 2004; 117:5937–5947. [PubMed: 15522895]
71. Lowell JE, et al. Histone H2AZ dimerizes with a novel variant H2B and is enriched at repetitive DNA in *Trypanosoma brucei*. Journal of cell science. 2005; 118:5721–5730. [PubMed: 16303849]

72. Kawahara T, et al. Two essential MYST-family proteins display distinct roles in histone H4K10 acetylation and telomeric silencing in trypanosomes. Mol Microbiol. 2008; 69:1054–1068. [PubMed: 18631159]
73. Siegel TN, et al. Acetylation of histone H4K4 is cell cycle regulated and mediated by HAT3 in *Trypanosoma brucei*. Mol Microbiol. 2008; 67:762–771. [PubMed: 18179414]
74. Wang QP, et al. Histone deacetylases play distinct roles in telomeric VSG expression site silencing in African trypanosomes. Mol Microbiol. 2010; 77:1237–1245. [PubMed: 20624217]
75. Alsford S, et al. A sirtuin in the African trypanosome is involved in both DNA repair and telomeric gene silencing but is not required for antigenic variation. Mol Microbiol. 2007; 63:724–736. [PubMed: 17214740]
76. Yang X, et al. RAP1 is essential for silencing telomeric variant surface glycoprotein genes in *Trypanosoma brucei*. Cell. 2009; 137:99–109. [PubMed: 19345190]
77. Janzen CJ, et al. Selective di- or trimethylation of histone H3 lysine 76 by two DOT1 homologs is important for cell cycle regulation in *Trypanosoma brucei*. Molecular cell. 2006; 23:497–507. [PubMed: 16916638]
78. Figueiredo LM, et al. A histone methyltransferase modulates antigenic variation in African trypanosomes. PLoS biology. 2008; 6:e161. [PubMed: 18597556]
79. Gissot M, et al. *Toxoplasma gondii* and *Cryptosporidium parvum* lack detectable DNA cytosine methylation. Eukaryotic cell. 2008; 7:537–540. [PubMed: 18178772]
80. Choi SW, et al. LC/ESI-MS demonstrates the absence of 5-methyl-2'-deoxycytosine in *Plasmodium falciparum* genomic DNA. Molecular and biochemical parasitology. 2006; 150:350–352. [PubMed: 16934885]
81. Militello KT, et al. African trypanosomes contain 5-methylcytosine in nuclear DNA. Eukaryotic cell. 2008; 7:2012–2016. [PubMed: 18791035]
82. Harony H, et al. DNA methylation and targeting of LINE retrotransposons in *Entamoeba histolytica* and *Entamoeba invadens*. Molecular and biochemical parasitology. 2006; 147:55–63. [PubMed: 16530279]
83. Fisher O, et al. Characterization of cytosine methylated regions and 5-cytosine DNA methyltransferase (Ehmeth) in the protozoan parasite *Entamoeba histolytica*. Nucleic Acids Res. 2004; 32:287–297. [PubMed: 14715927]
84. Ali IK, et al. Growth of the protozoan parasite *Entamoeba histolytica* in 5-azacytidine has limited effects on parasite gene expression. BMC genomics. 2007; 8:7. [PubMed: 17207281]
85. Borst P, Sabatini R. Base J: discovery, biosynthesis, and possible functions. Annual review of microbiology. 2008; 62:235–251.
86. van Leeuwen F, et al. The modified DNA base beta-D-glucosyl-hydroxymethyluracil is not found in the tsetse fly stages of *Trypanosoma brucei*. Molecular and biochemical parasitology. 1998; 94:127–130. [PubMed: 9719516]
87. Cliffe LJ, et al. Two thymidine hydroxylases differentially regulate the formation of glucosylated DNA at regions flanking polymerase II polycistronic transcription units throughout the genome of *Trypanosoma brucei*. Nucleic Acids Res. 2010; 38:3923–3935. [PubMed: 20215442]
88. Kieft R, et al. JBP2, a SWI2/SNF2-like protein, regulates de novo telomeric DNA glycosylation in bloodstream form *Trypanosoma brucei*. Molecular and biochemical parasitology. 2007; 156:24–31. [PubMed: 17706299]
89. Cliffe LJ, et al. JBP1 and JBP2 are two distinct thymidine hydroxylases involved in J biosynthesis in genomic DNA of African trypanosomes. Nucleic Acids Res. 2009; 37:1452–1462. [PubMed: 19136460]
90. Ekanayake D, Sabatini R. Epigenetic regulation of polymerase II transcription initiation in *Trypanosoma cruzi*: modulation of nucleosome abundance, histone modification, and polymerase occupancy by O-linked thymine DNA glucosylation. Eukaryotic cell. 2011; 10:1465–1472. [PubMed: 21926332]
91. Ekanayake DK, et al. Epigenetic regulation of transcription and virulence in *Trypanosoma cruzi* by O-linked thymine glucosylation of DNA. Molecular and cellular biology. 2011; 31:1690–1700. [PubMed: 21321080]

92. Epp C, et al. Chromatin associated sense and antisense noncoding RNAs are transcribed from the var gene family of virulence genes of the malaria parasite *Plasmodium falciparum*. RNA. 2009; 15:116–127. [PubMed: 19037012]

93. Li F, et al. Nuclear non-coding RNAs are transcribed from the centromeres of *Plasmodium falciparum* and are associated with centromeric chromatin. J Biol Chem. 2008; 283:5692–5698. [PubMed: 18165241]

94. Kolev NG, et al. RNA interference in protozoan parasites: achievements and challenges. Eukaryotic cell. 2011; 10:1156–1163. [PubMed: 21764910]

Box 1. Mechanisms of epigenetic regulation

Chromatin

DNA modification: *E. histolytica* (DNA methylation) [83], *Trypanosoma brucei* (DNA methylation) [81], *Kinetoplastids* (DNA glycosylation) [85]

Histone post-translational modifications: All eukaryotes

RNA-based heterochromatin maintenance: *Plasmodium?* [92, 93], *Toxoplasma*? [36]

RNA

RNA interference: *T. brucei, E. histolytica, G. lamblia, T. gondii* [1, 94]

Protein

Structural/Protein based *T. brucei*? [7], *P. falciparum*? [4–6], *T. gondii?*[3]

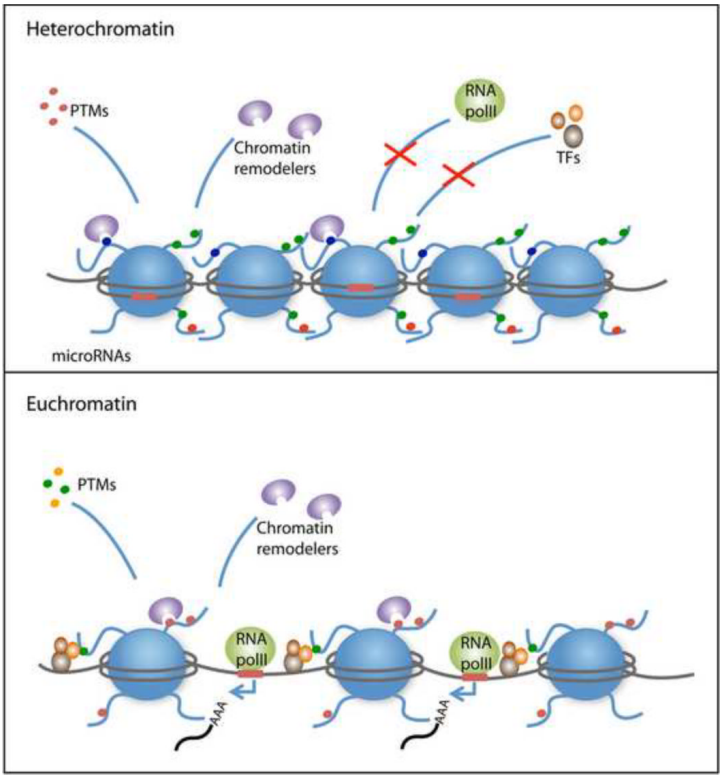

Figure 1.

Heterochromatin and euchromatin

Chromatin consists of heterochromatin and euchromatin. Heterochromatin is compact with characteristic histone post-translational modifications that prevent the accessibility of RNA polymerase and transcription factors. Heterochromatin frequently is maintained or established with the cooperation of microRNAs, and these mechanisms may be operative in maintenance of centromeric heterochromatin in the Apicomplexa [93]. For simplicity in this diagram, chromatin remodelers are depicted as histone modifying enzymes, but would include ATP-dependent remodelers, histone modifying enzymes and other specialized macromolecular complexes, which are important for maintaining a heterochromatin or euchromatin state. Euchromatin, an open chromatin state, is also maintained by macromolecular complexes, enabling access of sequence-specific transcription factors and the RNA polymerase complex. In parasites, the sequence of events and components of these complexes is not known, although orthologues of similar complexes in yeast and mammals are present.

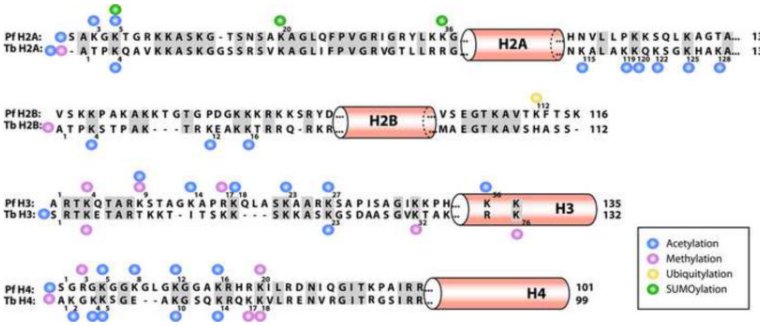

Figure 2.

Comparison of histone posttranslational modifications of canonical histones in *Plasmodium falciparum* and *Trypanosoma brucei*

The histone PTMs have been thoroughly characterized in *P. falciparum* [14] and *T. brucei* [12, 13] using proteomics techniques. The canonical histones H2A, H2B, H3 and H4 are highly conserved throughout evolution, but *T. brucei* histones are less well conserved (See alignments in Figure S1). Most of the PTM seen in *P. falciparum* have been described in other eukaryotes, but the modifications of *T. brucei* are divergent. This figure compares the PTM of the 4 canonical histones of *P. falciparum* and *T. brucei*. Although not fully characterized, variant histones are also marked by PTM. Many PTM on histones are substochiometric, reflecting the dynamic properties of chromatin. As more sensitive mass spectrometry techniques are utilized, this list is likely to expand. The specificity of interactions with chromatin remodelers and transcription factors is likely to be modulated by the combination of modifications on each histone (i.e. the histone code).

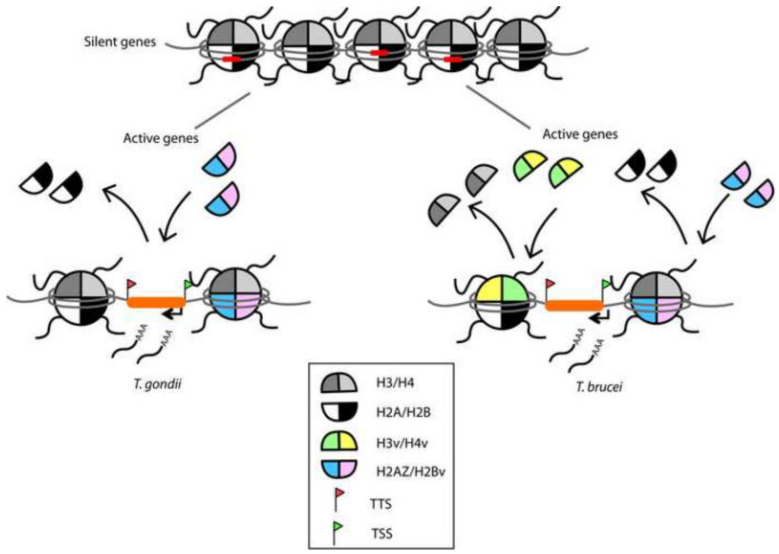

Figure 3.

Variant histones in *T. gondii* and *T. brucei*

Activation of genes in eukaryotes is often associated with displacement of nucleosomes composed of canonical histones with nucleosomes containing variant histones. Nucleosomes composed of variant histones are often more unstable, facilitating exchange. In *T. brucei* the transcription start site (TSS) of active genes are marked with H2AZ/H2Bv dimers, whereas, transcription termination sites (TTS) are marked with H3v/H4v dimers [53]. H3v is also highly enriched in telomeres. Transcription in *T. brucei* is polytranscriptional. In contrast, *T. gondii* active genes are marked with H2AZ/H2Bv at promoters [18], but to date no marks distinguishing TTS have been discovered. For silent genes, inactive chromatin is depicted as compact and inaccessible (similar to heterochromatin in Figure 1), but inactive genes do not have classic heterochromatin marks and the histone marks and protein complexes associated with silent genes are not characterized.

Table 1

Histones in parasites, yeast, and humans

	H. sapiens[a]	**S. cerevisiae**[b]	**T. brucei**[a]	**T. gondii**[c]	**P. falciparum**[c]
H3	AAA52651.1	not found[b]	Tb927.1.2430	TGME49_061240	PFF0510w
H3.3	CAA29288.1	AAS64349.1	not found	TGME49_018260	PFF0865w
CenH3	AAH02703.1	CAA81884.1	not found	TGME49_025410	PF13_0185
H4	AAA52652.1	EDV12280.1	Tb927.5.4170	TGME49_039260	PF11_0061
H4V	parasite specific	parasite specific	Tb927.2.2670	not found	not found
H2A	AAN59974.1	not found[b]	Tb927.7.2820	TGME49_061250	PFF0860c
H2A.Z	AAA35984.1	EDV10597.1	Tb927.7.6360	TGME49_100200	PFC0920w
H2A.X	NP_002096.1	CAA81267.1	not found	TGME49_061580	not found
H2B-A	CAA41051.1	CAA81268.1	Tb927.10.10460	TGME49_105160	PF11_0062
H2B-B	NA[d]	NA[d]	NA[d]	TGME49_051870	NA[d]
H2Bv	parasite specific	parasite specific	Tb11.02.5250	TGME49_009910	PF07_0054
H3V	parasite specific	parasite specific	Tb927.10.15350	not found	not found

[a]Only one copy of each of the canonical histones H2A, H2B, H3 and H4 was selected for analysis in *H. sapiens* and *T. brucei*, which have multiple nearly identical copies of histone genes.

[b]*Saccharomyces cerevisiae* H3 and H2A appear more closely related to H3.3 and H2A.X in metazoa although they function as canonical histones. There are 2 identical copies of each H3, H4, H2A, and H2B gene in *S. cerevisiae.*

[c]The Apicomplexa have only one copy of each histone, except for TgH2B.

[d]*T. gondii* has two copies of H2B. They differ by five residues on the N-termini and are otherwise highly conserved. Expression of H2B-A appears to be cell cycle dependent while H2B-B is constitutively expressed, according to data from www.ToxoDB.org.