Modifications at K31 on the lateral surface of histone H4 contribute to genome structure and expression in apicomplexan parasites

Fabien Sindikubwabo[1], Shuai Ding[2], Tahir Hussain[1], Philippe Ortet[3], Mohamed Barakat[3], Sebastian Baumgarten[2], Dominique Cannella[1], Andrés Palencia[1], Alexandre Boudgour[1], Lucid Belmudes[4], Yohann Coute[4], Isabelle Tardieux[5], Cyrille Y. Botté[6], Artur Scherf [2] and Mohamed-Ali Hakimi[1] *

[1] Institute for Advanced Biosciences (IAB), Team Host-pathogen interactions & immunity to infection, INSERM U1209, CNRS UMR5309, Université Grenoble Alpes, Grenoble, F-38700.

[2] Unité de Biologie des Interactions Hôte-Parasite, Institut Pasteur, CNRS, ERL 9195, INSERM, Unit U1201, Paris, F-75724.

[3] Aix-Marseille Univ, CEA, CNRS, UMR 7265, BIAM-LEMIRE, F-13108 St-Paul-lez-Durance, France

[4] Univ. Grenoble Alpes, CEA, INSERM, BIG-BGE, F-38000 Grenoble, France

[5] Institute for Advanced Biosciences (IAB), Team Membrane and Cell Dynamics of Host Parasite Interactions, INSERM U1209, CNRS UMR5309, Université Grenoble Alpes, Grenoble, F-38700.

[6] Institute for Advanced Biosciences (IAB), Team ApicoLipid, INSERM U1209, CNRS UMR5309, Université Grenoble Alpes, Grenoble, F-38700.

*Correspondence to: Mohamed-ali Hakimi, Phone +33 4 76 63 74 69, e.mail: mohamed-ali.hakimi@inserm.fr

ORCiD : 0000-0002-2547-8233

Running title: Chromatin regulation by H4K31 modifications

Key words: *Toxoplasma gondii*, *Plasmodium falciparum*, histone core modifications, acetylation, methylation, chromatin, gene expression

Abstract

An unusual genome architecture characterizes the two related human parasitic pathogens *Plasmodium falciparum* and *Toxoplasma gondii*. A major fraction of the bulk parasite genome is packaged as transcriptionally permissive euchromatin with few loci embedded in silenced heterochromatin. Primary chromatin shapers include histone modifications at the nucleosome lateral surface close to the DNA but their mode of action remains unclear. We now identify versatile modifications at Lys31 within the globular domain of histone H4 that crucially determine genome organization and expression in *Apicomplexa* parasites. H4K31 acetylation at the promoter correlates with, and perhaps directly regulates, gene expression in both parasites. In contrast, monomethylated H4K31 is enriched in the core body of *T. gondii* active genes but inversely correlates with transcription while being unexpectedly enriched at transcriptionally inactive pericentromeric heterochromatin in *P. falciparum*, a region devoid of the characteristic H3K9me3 histone mark and its downstream effector HP1.

Introduction

The phylum of *Apicomplexa* clusters thousands of single-celled eukaryotes identified as parasites of metazoans including humans in who they cause or put at risk for major public health problems. Preeminent human pathogens include *Plasmodium spp.* which are responsible for dreadful malaria as well as *Toxoplasma gondii* and *Cryptosporidium spp.* which are leading causes of food-borne and water-borne diseases. A shared characteristic of apicomplexan life cycles is the multiplicity of developmental stages that progress from one to the other along with precise genetic reprogramming to ensure survival and transmission of parasite populations. The emerging concept of a remarkably dynamic nature of gene expression in *Apicomplexa* has risen from the observation that large numbers of mRNAs are exclusively expressed in a given developmental stage (Bozdech et al., 2003; Radke et al., 2005).

Unlike metazoans, *Apicomplexa* genomes have a unique chromatin architecture typified by an unusually high proportion of euchromatin and only a few heterochromatic islands scattered through the chromosome bodies or embedded at telomeres and centromeres. Although alterations in chromatin structure are acknowledged as important for the transcriptional control of commitment to stage differentiation in several *Apicomplexa* as well as for antigenic variation-mediated immune evasion in *P. falciparum*, yet the understanding of chromatin remodeling remains incomplete (Bougdour et al., 2010; Scherf et al., 2008).

In Eukaryotes, the timely opening and closing of chromatin required for gene expression, chromosomal organization, DNA repair or replication is governed by histone turnover and their post-translational modifications (PTMs), such as lysine methylation (me) and acetylation (ac) among many others. PTMs on histone tails were indeed shown, alone or in combination, to alter the accessibility of effector proteins to nucleosomal DNA and thereby impact on chromatin structure, according to the "histone code" hypothesis (Strahl and Allis, 2000; Turner, 2000). In addition, PTMs also act as signals to recruit ATP-dependent remodeling enzymes either to move, eject or reposition nucleosomes. Accounting for the PTMs versatile activity onto chromatin are enzymes carrying antagonist activities: it is the opposite, yet well concerted

activities of histone acetyltransferases (HATs) and histone deacetylases (HDACs) that control acetylation levels in cells.

Apicomplexa have evolved unique ways to modify histones that rival the strategies adopted by the cells they infect and provide zoites with remarkable capacities to multiply or undergo stage differentiation (Bougdour et al., 2010). Some histone-modifying enzymes have acquired gain-of-function mutations that confer broader or enhanced activity on substrates. It is the case of parasite Set8-related proteins endowed with H4K20 mono-, di-, and trimethylase activities that contrast with the mono-methylase-restricted metazoan Set8, and that derive from a single-amino-acid change in the substrate-specific channel (Sautel et al., 2007). Another example is the *Apicomplexa* HDAC3 family which is typified by an AT motif insertion at the entrance of the active-site tunnel of the conserved catalytic domain causing additional substrate/inhibitor recognition and binding properties (Bougdour et al., 2009).

In Eukaryotes, while PTMs have been primarily detected in the histone tails sticking out from the nucleosome, an ever-growing list of PTMs is now identified in the lateral surface of the histone octamer, that directly contacts DNA, whose members were characterized as critical regulators of the chromatin structure and function (Lawrence et al., 2016; Tropberger and Schneider, 2013). Those "core" histone PTMs promote different outcomes on nucleosome dynamics depending on their location. Modifications near the DNA entry-exit region (e.g., H3K36ac) of the nucleosome were shown to favor local unwrapping of DNA from histone octamer thereby providing a better exposure of nucleosomal DNA to chromatin-remodeling and DNA-binding proteins (Neumann et al., 2009). On the other hand, lateral-surface PTMs mapping close to the dyad axis (e.g., H3K122ac) were shown to decrease the affinity of the octamer to DNA and significantly affect nucleosome stability/mobility (Tropberger et al., 2013). Similarly to what has been described for histone tails, different lateral-surface modifications on the same residue can be associated with opposite transcriptional programs. This is the case for the H3K64 residue near the dyad axis that facilitates nucleosome eviction and thereby gene expression when acetylated (Di Cerbo et al., 2014), whereas trimethylation of the same residue acts as a repressive heterochromatic mark (Daujat et al., 2009).

In both *T. gondii* and *P. falciparum* unbiased mass spectrometry has led to uncover the repertoire of the most prevalent histone PTMs including singular marks. However, only few of them were mapped at the outer surface of the octamer (Saraf et al., 2016; Trelle et al., 2009). In this study, we investigated in depth how histone H4 PTMs could influence chromosome organization and gene regulation in apicomplexan parasites. We reported versatile modifications at lysine 31 of histone H4 (H4K31), which lies at the protein-DNA interface close to the dyad axis of the nucleosome. Genome-wide mapping revealed that H4K31 could either be acetylated or methylated and the mark enrichment occurred in a mutually exclusive manner. In *T. gondii* H4K31 residue tended to be acetylated at the promoter of a nearby active gene and to be mono-methylated in the core body of the gene. H4K31me1 occupancy was inversely correlated with gene expression suggesting that the mark acts as a repressive mark impeding RNA polymerase progression. In *P. falciparum*, H4K31ac was also seen at the promoter whereas H4K31me1 occupancy was highly enriched at pericentromeric heterochromatin possibly compensating the absence of H3K9me3 and HP1 in this atypical chromosome structure in order to maintain a constitutive heterochromatin environment.

Results

H4K31 maps at the Dyad Axis of the Nucleosome

While studying the protein content of nucleosomes from *T. gondii* infected cells, we and others mapped an acetylation site on histone H4 lysine 31 (H4K31ac) that was largely underestimated thus far (Fig. 1a) (Jeffers and Sullivan, 2012; Xue et al., 2013). This modification was also identified in both *P. falciparum* proteome (Saraf et al., 2016) and acetylome (Cobbold et al., 2016) throughout the intraerythrocytic developmental cycle of the parasite. H4K31 residue is located at the N-terminus of the H4 α1 helix and its side chain is extended in the major groove of the DNA (Fig. 1b and c). The closed state of chromatin is contributed by interaction of K31 and R35 residues to DNA by a water mediated hydrogen bond. Addition of an acetyl group to the -NH2 group of lysine side chain (K31) abolishes its interaction with DNA mediated by a water molecule (Fig. 1c, panel 2). Acetylation may therefore destabilize the protein-DNA interface close to the dyad axis of the nucleosome where the residue lies and thus presumably open the chromatin. Although the residue H4K31 is well

conserved across species (Fig. 1b), mass spectrometry initially indicates its acetylation to be restricted to metazoans, as an unexpected mark in inflammatory and auto-immune contexts (Garcia et al., 2005; Soldi et al., 2014) since it was found neither in yeast nor in the ciliated protozoan *Tetrahymena* (Garcia et al., 2006). Recent studies including our present data contradict this view as they show this PTM to also arise in the phylum of *Apicomplexa* (Cobbold et al., 2016; Jeffers and Sullivan, 2012; Saraf et al., 2016).

H4K31ac marks euchromatin in mammalian cells and apicomplexan parasites

Although H4K31ac was unequivocally identified by mass spectrometry in both *T. gondii* (Jeffers and Sullivan, 2012; Xue et al., 2013) and *P. falciparum* (Cobbold et al., 2016; Saraf et al., 2016) (Fig. 1a), dynamics and nuclear distribution of the mark during infections remain understudied. To further probe *in situ* the kinetics of this histone mark in apicomplexans, we raised an antibody against a synthetic peptide acetylated at the H4K31 position whose specificity was controlled by dot-blot assays. First, no cross-reactivity with the unmodified peptides (Fig. 1 – figure supplement 1A) or with previously described acetyl and methyl marks in histone tails and globular domains (Fig. 1 – figure supplement 1) was detected. Secondly, H3K14ac-directed antibodies did not cross-react with H4K31 peptides (Fig. 1 –figure supplement 1A) while properly detecting H3K14ac-containing peptides (Fig. 1 – figure supplement 2). Using human primary fibroblasts infected by *T. gondii*, we found H4K31ac to be exclusively and uniformly distributed within the nuclei of both parasite and human cells (Fig. 1d). We also found that exposing cells to histone deacetylase inhibitor (HDACi), e.g., FR235222 significantly increased H4K31ac signal intensity otherwise moderate in parasite nucleosomes (Fig. 1e), another evidence for the specific detection of acetylation. Similarly, response to HDACi treatment was observed by immunofluorescence analysis (Fig. 2). Interestingly, PTMs at histone H3 tail such as H3K14ac and H3K27ac were not altered upon FR235222 treatment under our conditions, which contrasts with the increased signal of H4K31ac and further confirms the specificity of the H4K31ac-directed antibodies (Fig. 2b).

To gain insight on the behavior of H4K31ac during the *P. falciparum* intraerythrocytic developmental cycle (IDC), immunofluorescence assays were

conducted over 48 hours of culture to probe the ring, trophozoite and schizont stages (Fig. 3a). H4K31ac was typified by a nuclear signal throughout the IDC that increased upon HDACi treatment (Fig. 3b and 3c). Overall, H4K31ac showed a nuclear punctate pattern, reminiscent of active loci clusters in specialized 'transcription factories' (Fig. 3b) (Mancio-Silva and Scherf, 2013).

Strikingly, H4K31ac has remained understudied in other eukaryotes thus far. To gain better resolution of any nuclear or chromatin structures with which H4K31ac might be associated, we co-stained murine embryonic fibroblasts (MEFs) for DNA and various chromatin marks. H4K31ac was observed scattered through the nucleoplasm of MEFs but excluded from nucleoli and segregated away from heterochromatic foci similarly to the transcription-associated PTMs H3K4ac, H3K9ac and H3K27ac (Fig. 3d). This pattern is typically euchromatic and opposed to the one revealed by the repressive marks H3K9me3 and H4K20me3, rather associated with regions of highly condensed pericentromeric heterochromatin (Fig. 3e). Taken together, these experiments show that H4K31ac displays an euchromatic pattern in both metazoans and apicomplexans.

GCN5b and HDAC3 enzymes fine-tune the H4K31ac levels in *Toxoplasma gondii*

To our knowledge, the enzymes that acetylate and deacetylate H4K31 remain unknown. FR235222 treatment induced a 5.3-fold increase of H4K31ac signal in *T. gondii* nuclei when compared to other acetylated residues of the histone tails (Fig. 2b), an observation that prompted us to seek for potential deacetylases targeting H4K31ac. To this end, we selected 7 HDACi that cover the entire selectivity range for class I and II deacetylases, albeit with varying specificity profiles against apicomplexan parasites, as determined previously (Bougdour et al., 2009) and included a specific inhibitor of protein translation in *Apicomplexa* halofuginone as a control (Jain et al., 2015). We found that cyclopeptides HDACi strongly enhanced H4K31ac levels in parasite nuclei, whereas other inhibitors had no effect (Fig. 4a and b). These results are consistent with those showing that distinct point mutations at a single locus in apicomplexan conserved region of TgHDAC3 abolishes the enzyme sensitivity to the cyclic tetrapeptide compounds (Bougdour et al., 2009). *T. gondii* is particularly suited for a single gene perturbation strategy, since its genome does not contain extensive HAT and HDAC paralogs unlike mammalian genomes. To identify which of the five class-I/II HDAC homologues in *T. gondii* may account for the deacetylation of this

residue, we systematically performed CRISPR-mediated gene disruption. *TgHDAC3* but not other *TgHDACs* gene inactivation caused hyperacetylation of H4K31 in parasite nuclei (Fig. 5a, b and c), thereby mimicking the effect of the cyclic tetrapeptide HDACi on the enzyme (Fig. 4) (Bougdour et al., 2009).

Reciprocally, we next sought for the one or more responsible HATs targeting this residue. We used a candidate-based approach by systematically depleting the parasite for key members of the three main HAT classes. Apicomplexans possess homologues of the Type A GCN5 and MYST family nuclear HATs as well as the Type B cytoplasmic HAT1 (Vanagas et al., 2012) while missing the mammalian restricted PCAF (p300/CBP-associating factor) family. Intriguingly, *T. gondii* is unique among fellow apicomplexan parasites and other invertebrates in possessing two GCN5 HATs, designated TgGCN5a and b, that exhibit different histone acetylation activities (Vanagas et al., 2012). We first thought of *T. gondii* HAT1 as a promising candidate enzyme as the human HAT4 was shown to acetylate core PTMs *in vivo* and even H4K31 although under *in vitro* conditions (Yang et al., 2011). However, cas9-mediated gene disruption of HAT1 had no effect on H4K31 acetylation (Fig. 6a and b) whereas *TgGCN5b* was the only HAT encoding gene which disruption resulted in a drastic drop of H4K31ac signals in the parasite nuclei (Fig. 6a and b). TgGCN5b is the prototypical GCN5 HAT in *T. gondii* because it can target H3K9, K14 and K18 (Bhatti et al., 2006). Furthermore, we noticed that the amino acid sequence surrounding H4K31 was not homologous to preferred GCN5 consensus sites of acetylation found at H3K14 (Kuo et al., 1996) or H3K36 (Morris et al., 2006), yet the depletion of *TgGCN5b* led to a reduction of both H3K14ac (Fig. 6c and d) and H4K31ac (Fig. 6a and b) signals *in vivo*, suggesting that the repertoire of the lysine residues being acetylated by the GCN5 family is more diverse in *T. gondii*. Altogether these data clearly uncover *TgGCN5b* as H4K31 acetyltransferase whose activity is counteracted by *TgHDAC3*.

H4K31me1 associates *in vivo* with distinct chromatin patterns

It is well appreciated that the targeting of lysine residues by acetylation and methylation or other PTMs cannot occur simultaneously. H4K31 has been identified by proteomic mapping as succinylated in *Drosophila*, yeast, and mammalian cells (Xie et al., 2012) as well as in *T. gondii* (Nardelli et al., 2013). Formylation is another

modification targeting H4K31 in both human (Wiśniewski et al., 2008) and *Apicomplexa* (Nardelli et al., 2013; Saraf et al., 2016). While mass spectrometry-based proteomics strategies have allowed the characterization of H4K31 as a site of monomethylation in plant, budding yeast and metazoan cells, they failed to detect the PTM in *Tetrahymena* (Garcia et al., 2006; Moraes et al., 2015) and *Apicomplexa* (Nardelli et al., 2013; Saraf et al., 2016). Contrariwise, H4K31me2 has been detected in *T. gondii* proteome (Nardelli et al., 2013). Besides from proteomic approaches, the possibility of "dual" modifications occurring on H4K31 has not been yet explored neither in apicomplexans nor in any other species. Therefore, we raised antibodies against synthetic peptides containing mono- and di-methylated H4K31 and controlled for their specificity. The H4K31me2-directed antibodies, although reacting avidly with the peptide antigen, did not detect histone H4 in human or parasite protein extracts (data not shown). On the other hand, antibodies raised against H4K31me1-containing peptide recognized nicely histone H4 in purified *T. gondii* core histone extract (Fig. 1e). The antibody is highly specific for synthetic H4K31me1 peptide (Fig. 1 – figure supplement 3a) over peptides with previously described acetyl and methyl marks in histone tails and globular domains in dot-blot assays (Fig. 1 – figure supplement 3b, c and d). As a control, H4K20me1-directed antibodies (Sautel et al., 2007) did not cross-react with H4K31 peptides (Fig. 1 –figure supplement 3a) while properly detecting H4K20me1-containing peptides in dot-blot assay (Fig. 1 – figure supplement 4). Taken together, these data show that our home-made antibodies are specific of the H4K31me1 epitope.

In situ, the H4K31me1 modification appeared uniformly distributed within the nucleus of dividing parasites but surprisingly no signal was detected in the nucleus of the infected human cell (Fig. 7a), despite that this PTM had been previously detected by mass spectrometry in human samples (Garcia et al., 2006). However, while H4K31me1 was not (or barely) detected in interphase nuclei of either quiescent infected HFFs (Fig. 7a) or uninfected MEFs (Fig. 7b), it decorated mitotic chromosomes providing even more pronounced signals in the chromosome arms than the usual mitotic marker H3S10 phosphorylation (Fig. 7b). These observations argue for a possible role of H4K31 methylation during cell division in mammalian cells.

In *P. falciparum*, H4K31me1 displayed a peculiar condensed punctate pattern (Fig. 7c), similar to the H3K9me3 mark (Lopez-Rubio et al., 2009) at the nuclear periphery, which is reminiscent of heterochromatin/subtelomeric regions clustering (Freitas-Junior et al., 2000). *P. falciparum* centromeres also clustered prior to and throughout mitosis and cytokinesis leading to single nuclear location from early trophozoites to mature schizonts (Hoeijmakers et al., 2012). Therefore, H4K31me1-containing foci could be associated with subtelomeric or/and centromeric regions. The number of foci observed however varied signifying its dynamic changes across parasite developmental stages. The mark was observed in all asexual forms and remained unaffected by treatment of FR235222 (data not shown).

Because H4K31 is also a site of methylation, the transition between H4K31ac and H4K31me1 may represent a novel "chromatin switch" contributing to chromatin structure and function in eukaryotic cells. Yet, a different readout is expected from one species to another. In metazoan, H4K31me1 was temporally regulated during the cell cycle and interplay, if any, with H4K31ac should be restricted to mitotic chromosomes. In *P. falciparum*, H4K31me1 formed discrete immuno-fluorescent *foci* around the nucleus, a pattern quite distinct from that of H4K31ac typified by a diffused signal distributed throughout the parasite nuclei. Since they have distinct nuclear locations and different stoichiometry, H4K31ac being a low abundant species, the transition between H4K31me1 and H4K31ac may be not an issue in *P. falciparum* as it can be in *T. gondii* where both modifications are concomitantly distributed throughout euchromatin.

Nucleosomes with H4K31ac and H4K31me1 are enriched at the promoter and the core body of active genes, respectively

To further explore whether H4K31ac and H4K31me1 are indeed alternative antagonistic PTMs on the same H4 molecules in *T. gondii*, we examined their genome-wide distributions using chromatin immunoprecipitation coupled with next-generation sequencing (ChIP-seq). We first investigated the relative performance of our home-made antibodies in terms of specificity, sensitivity and the number and distribution of peaks. We observed low variability and a high degree of similarity in read coverage between technical replicates, regardless of antibodies used for immunoprecipitation (Fig. 8 – figure supplement 1). We next compared the locations

of the peaks from each antibody type. Visual display of the chromosomal distribution indicated that H4K31ac and H4K31me1 exhibited distinct patterns of enrichment across the chromosomes and were mutually exclusive genome-wide (e.g., Chr. Ib, Fig. 8a). Zooming into detailed gene level revealed that H4K31ac was enriched in distinct peaks at intergenic regions (IGRs) (Fig. 8b), of which 75% percent mapped outside the gene body (Fig. 8c), in line with the euchromatic *in situ* localization (Fig. 1 and Fig. 3). The calculated average profile of H4K31ac showed a pattern strikingly similar to H3K14ac and H3K4me3, characterized by high signals at 5'UTR/promoter that drop sharply after the translational initiation site (Fig. 8d). Conversely, H4K31me1 showed a distinct pattern of enrichment, best discernable at large genes, spanning from the ATG to the entire gene body (Fig. 8b) while being absent at IGRs (Fig. 8c). H4K31me1 computed average density profile (Fig. 8d) fully matched with gene prediction making this mark useful to explicitly detect unannotated genes (Fig. 8 – figure supplement 2). Remarkably, these data allowed identifying H4K31me1 as a novel PTM whose spreading was restricted to gene body in *T. gondii* (Fig. 8b and d). However, this mark was not the only one to decorate chromatin in this fashion as a similar pattern was also found for the genome-wide distribution of H3K4me1 in type I (RH) *T. gondii* strain (Kami Kim, unpublished data available at ToxoDB). We have re-examined the extent and genome-wide scattering of H3K4me1 using a type II (Pru) genetic background. We observed a high degree of similarity in read coverage between H3K4me1 and H4K31me1 (Fig. 8 – supplement figure 3a) and a similar pattern of enrichment over the gene body (Fig. 8 – supplement figure 3b). However, those PTMs are unevenly distributed across the genome, H3K4me1 being occasionally weakly enriched relative to H4K31me1 at few expressed genes (see red dotted square, Fig. 8 – supplement figure 3c and d).

Interplay between H4K31ac and H4K31me1 predicts distinctive patterns of gene expression in *T. gondii*

A closer view of H4K31ac and H4K31me1 chromosomal binding revealed that, at some loci (e.g., *SRS* gene family, Fig. 9a and *GRA1*, Fig. 9b), the latter was absent while the former was enriched at 5'UTR/promoter and unexpectedly spread over a much larger area, overlapping the gene body (see *GRA1* and *MAG1* examples, Fig. 9c). The restricted H4K31ac enrichment at the vicinity of *GRA1 or SRS* genes contrasts with the H4K31ac and H4K31me1 location at their neighboring genes

(*TGME49_233490* or *TGME49_270230*) and this discrepancy may be explained by the higher level of *SRS* or *GRA1* gene expression (Fig. 9a and b). We therefore interrogated whether enrichment patterns of modified H4K31 could specify levels of gene expression in *T. gondii*. We first conducted a global transcriptome analysis by RNA-Seq of tachyzoites during growth phase in murine bone marrow-derived macrophages (BMDMs). Cluster analysis revealed varying levels of gene expression with cluster Q1 displaying the highest level, clusters Q2 and Q3 defining intermediate levels and cluster Q4 displaying the lowest (Fig. 9d and Figure 9-source data 1). High mRNA level (Q1 that includes *GRA1, MAG1* and *SRS* genes) was associated with high level of H4K31ac upstream of the ATG together with an enrichment along the gene body which coincided with the expected lack of H4K31me1 (Fig. 9e and f). In highly expressed relatively long or intron-containing genes (e.g., *MAG1*), H4K31ac spread but did not extend over the entire gene body as observed for *GRA1* (Fig. 9c), indicating a limited spreading of H4K31 acetylation around the translational initiation site. Strikingly, moderate mRNA levels (Q2 and Q3 that include *TGME49_233490* or *TGME49_270230*) related to a relatively high level of H4K31me1 in the gene body and a restricted mapping of H4K31ac at the promoter, thereby arguing for an inverse correlation between the yield of expression and the level of H4K31 methylation (Fig. 9e and f).

Finally, transcriptionally repressed genes clustered in Q4 showed no significant enrichment of either H4K31ac or H4K31me1 (Fig. 9e and f) but were enriched in the repressive mark H3K9me3 (Fig. 9 – supplement figure 1e). In addition, we discovered that genes clustered in Q4 were transcriptionally heterogeneous and the repressive signature H3K9me3 was primarily present at the vicinity of genes typified by low (if not undetectable) RNA levels in tachyzoites and referenced as stage-specific within the *T. gondii* life cycle. As such, *ENO1* and *BAG1* genes well known to be distinctively expressed in the at the bradyzoite stage during the chronic phase of infection (Pittman et al., 2014) were highly enriched in the repressive mark H3K9me3 but they were also unexpectedly enriched in the gene activation hallmark H3K14ac (Fig. 10a and Fig. 10 – supplement figure 1a). Coccidian-specific surface gene SRS families whose expression was restricted to bradyzoite stage (Fig. 10 – supplement figure 1b) also displayed marks of both active and silent chromatin (Fig. 10 – supplement figure 1c and d). This dual histone PTM was also detected on

nucleosomes surrounding genes that have been recognized as exclusively expressed during sexual stages in the definitive feline host (Hehl et al., 2015). The co-enrichment of H3K9me3/H3K14ac was somehow restricted to the transcriptional units that become active in the sexual stages (shown in red, Fig. 10b). However H3K14ac was seen to spread over some tachyzoite-expressed neighboring genes (shown in blue, Fig. 10b) and likely contributes to their activation. Genes carrying this chromatin signature fall into two transcriptomic clusters within the enteroepithelial stages (Hehl et al., 2015): a cluster corresponding to the merozoite genes, i.e. parasites harvested from cat at day 3 (Fig. 10 – supplement figure 2a and b) that includes *GRA11b* (Ramakrishnan et al., 2017) and a cluster of genes typical of the sexual stages among which *AP2X-10* (Hong et al., 2017) and SUB6 are representative examples (Fig. 10 – supplement figure 2c and d). Merozoites, which only infect the feline enterocytes were shown to specifically express a large repertoire of 52 SRS proteins (Hehl et al., 2015). Over the 111 members of the SRS superfamily of proteins annotated in ToxoDB, we identified 66 SRS genes displaying H3K9me3/H3K14ac enrichment, including the aforementioned 52 specific to merozoite (Fig. 10 – supplement figure 3) along with 8 bradyzoite SRS genes (Fig. 10 – supplement figure 1b) whereas the 14 SRS genes exclusively expressed in tachyzoite (e.g. SAG1 cluster) lack the dual PTM (Fig. 9a).

The co-enrichment of H3K14ac and H3K9me3 is somewhat paradoxical. Although this pattern remains unclear, it seems to bookmark genes repressed temporally, which await parasite stage differentiation for stage-specific expression. Thus, H3K14ac and H3K9me3 could form the so-called bivalent chromatin domain capable of silencing developmental genes while keeping them poised for rapid activation upon cell differentiation (Voigt et al., 2013), in a similar manner to the H3K14ac along with the repressive mark H3K27me3 shown to be enriched at a subset of inactive promoters in mouse embryonic stem cells (Karmodiya et al., 2012). Bivalent domains have gathered wide attention, because they might contribute to the precise unfolding of gene expression programs during cell differentiation. Apparently, those bivalent marks are restricted to inactive stage-specific promoters and contrast with both pericentromeric (Fig. 10c and Fig. 10 – figure supplement 4) and telomeric (Fig. 10d) heterochromatin, both decorated by H3K9me3 but missing H3K14ac enrichment. It was previously reported that H3K9me3 typifies centrometric heterochromatin in *T.*

gondii (Brooks et al., 2011) but this study conflicts with our data in reporting enrichment of the mark to "poised" stage-specific genes. This discrepancy could be explained by the genetic background of the parasite strain since Brooks and colleagues infected human cells with a type I (RH) strain that lost the ability to develop into mature cysts while we used infections with a type II strain (Pru) which is more relevant as it does readily develop tissue cysts and latent infections in animals.

Taken together, our data highlight unique chromatin signatures associated with the transcription rate in *T. gondii*. Genes clustered in Q1 are primarily defined by H4K31me1 low enrichment and enhanced acetylation at both promoter and 5' proximal gene body, while those from Q2 and Q3 are markedly typified by the presence of H4K31 methylation in the gene body and an acetylation mark restricted to the promoter. In this context, H4K31ac would be predicted to disrupt histone-DNA interaction thereby affecting nucleosome stability while promoting RNA polymerase progression across transcribed units. Conversely, H4K31me1 could act as a transcription-linked repressive mark that may hypothetically slow the progress of the RNA polymerase on active genes, likely by modulating the transcription-dependent histone turnover, but this still needs to be established. Nevertheless, the mark does not elicit its predicted repressive effect on constitutively repressed genes. Otherwise, we identified a subset of genes typified by their exclusive expression in chronic and sexual stages that are displaying typical bivalent chromatin domains characterized by H3K9me3 and H3K14ac enrichments in acute-phase tachyzoite.

H4K31me1 enrichment, a blueprint for unannotated genes and uncharacterized long non-coding RNAs

As mentioned previously, H4K31me was mainly detected throughout the body of active genes with translation start and stop codons as boundaries and its enrichment was inversely correlated to the yield of mRNA. These features should allow this mark to explicitly predict unannotated genes even when the low level of expression impedes detection by RNA profiling (Fig. 8 – figure supplement 2).

Although H4K31me1 rarely covered IGRs (Fig. 8c), the mark was found enriched occasionally in chromosomal regions devoid of any predicted protein-coding genes (Fig. 10d and Fig. 10 – figure supplement 5). This enrichment correlated with

extensive transcription of large RNA transcripts ranging from 20 to 70 kB reminiscent of long noncoding RNAs (lncRNAs) in other eukaryotic cells (Azzalin and Lingner, 2015). Those *T. gondii* lncRNAs are stand-alone transcription units with a proper chromatin signature, i.e., H4K31ac and H3K4me3 at the promoter and H4K31me1 along the transcribed length (Fig. 10d and Fig. 10 – figure supplement 5). Considering their distribution at both telomere-adjacent regions (Fig. 10d) and chromosome arms (Fig. 10 – figure supplement 5), those lncRNAs may work in *cis* near the site of their production (e.g., functions in telomere homeostasis) or act in *trans* to alter chromatin shape and gene expression at distant loci, as reported in other model organisms (Azzalin and Lingner, 2015).

Distribution of H4K31 modifications across the *P. falciparum* genome reveals H4K31me1 as a novel pericentromeric PTM

The *P. falciparum* genome is primarily maintained in a decondensed euchromatic state with perinuclear heterochromatin islands. Those heterochromatin-based gene silencing regions are used for the regulation of monoallelic expression of clonally variant genes (e.g. *var* and *rifin*) and are enriched in H3K9me3 which binds HP1 (Voss et al., 2014). We observed an apparent non-overlapping staining for acetylated and methylated H4K31 and more specifically a discrete focal distribution of H4K31me1 at the nuclear periphery (Fig. 3b and Fig. 7c). To get a comprehensive view of the genomic distribution of those PTMs across *P. falciparum* genome, we also performed ChIP-seq analyses during the IDC. As for *T. gondii*, we observed low variability and high similarity in read coverage between technical replicates for all the antibodies used (Fig. 11 – figure supplement 1).

We next compared peak location for each antibody type. H4K31ac displayed a rather even distribution throughout the genome similarly to the euchromatic mark H3K4me3 (Fig. 11a and enhanced view at Fig. 11 – figure supplement 2). As for *T. gondii*, H4K31ac matched with the gene annotation, i.e., high at promoter and low at gene body of active genes (e.g., *GAPDH*, Fig. 11b). Consistent with this, the H4K31ac and the repressive mark H3K9me3 were found inversely correlated (Fig. 11a). However, while H4K31ac displayed a relatively narrow enrichment restricted to transcribed promoters, H3K4me3 was instead enriched in a large fraction of the genome (Fig. 11c) as already described (Salcedo-Amaya et al., 2009).

Interestingly, the methylation of H3K9 and the properties to bind HP1 which have emerged as hallmarks of pericentromeric heterochromatin in model systems, including *T. gondii* (Brooks et al., 2011; Gissot et al., 2012) have not been detected in *P. falciparum*, leading to the view that the parasite may lack pericentric heterochromatin (Flueck et al., 2009; Lopez-Rubio et al., 2009; Salcedo-Amaya et al., 2009). While our ChIP-seq analysis confirmed the absence of pericentric enrichment of both H3K9me3 and HP1, it clearly highlighted a remarkable enrichment of H4K31me1 at pericentromeric regions that flank the cenH3-enriched centromeres (Fig. 11d, Fig. 11 – figure supplement 3 and Fig.12a). It is therefore possible that H4K31me1 constraints PfCENH3 to the centromeres in *P. falciparum* thus replacing the H3K9me3/HP1 functions described in most of the species. In addition to the pericentromeric localization H4K31me1 was also enriched to few subtelomeric regions and more specifically at telomere-associated repetitive element (TARE, Fig. 12b) repeat blocks shown to encode lncRNAs (Fig. 12c) (Sierra-Miranda et al., 2012). The presence of the mark at pericentric chromatin combined with its absence at transcriptionally permissive loci (e.g. *GAPDH*, Fig. 11b) suggest H4K31me1 as a novel hallmark of heterochromatin in *P. falciparum*, but not similar to H3K9me3/HP1 in subtelomeric regions and at *var* genes.

Discussion

In this study, we provide in depth understanding of the interaction between the core histone H4 and the template DNA by functionally characterizing novel modifications of H4K31, a residue exposed on the outer surface of the nucleosome in close proximity to the DNA entry-exit point. Proteome-wide mapping of acetylome/methylome as well as nucleosome protein content analyses allowed identifying H4K31 as a site for both acetylation and methylation across a wide range of species including those from the apicomplexan parasitic phylum. The K31 residue lies at the N-terminus of the histone H4 α1 helix and its positively charged side chain forms water-mediated interactions with the DNA phosphate backbone (Fig. 1c). Its acetylation was predicted to trigger substantial conformational changes in the nucleosome by shifting the side chain of lysine from unacetylated to acetylated state and causing the loss of the water-mediated interactions K31 establishes with DNA and

the residue R35 (Fig. 1c). However, this prediction was not validated since X-ray crystallography did not indicate large structural changes into nucleosomes when glutamine was substituted to lysine to mimic the acetylated state (H4Q31, Fig. 1c) (Iwasaki et al., 2011). Alternatively, H4K31ac may increase DNA unwrapping at the entry-exit point of the nucleosome thus giving access to the ATP-dependent chromatin remodelers that act on nucleosome disassembly and turnover as proposed by Chatterjee et al., 2015. The latter assumption would fit with the "regulated nucleosome mobility" model (Cosgrove et al., 2004), which predicts that outer surface PTMs (e.g., H3K36ac, Williams et al., 2008) regulate the equilibrium between mobile and relatively stationary nucleosomes by altering histone-DNA molecular interplay.

In both *T. gondii* and *P. falciparum*, genome wide studies pinpointed a local enrichment of H4K31ac at active gene promoters, in line with the cooperative contribution of acetylation and other PTMs to shape a transcriptionally permissive chromatin state. While H4K31ac relieves nucleosomal repression thus facilitating the access of the transcriptional machinery to the DNA template, H4K31 monomethylation likely locks the nucleosome in a repressed conformation which maintains chromatin in a closed or semi-closed state also called poised-state, while it also prevents GCN5-related HAT to catalyze acetylation of the residue. Interestingly, in *T. gondii*, apart from its predicted effects on the nucleosome mobility and chromatin state, we found that H4K31ac also prevents methylation at the body of highly expressed genes, thereby ensuring maximal efficacy of the RNA polymerase progression and activity. Indeed it is only in the transcribed coding sequence of a subset of genes typically associated with limited activity of the RNA polymerase II that we found enrichment in H4K31me1. In a model where the nucleosome disassembles in front of transcribing RNA polymerase II to allow its physical progression across transcribed units, it is plausible that H4K31me1 by stabilizing the wrapping of DNA around the histone octamer slows down the RNA Pol II processing along the fiber hence reducing the level of transcription.

Aside from specific patterns of PTMs, histone chaperones significantly contribute to control how the RNA polymerase II engages the nucleosome in and around a promoter and during the elongation step. For instance, the FACT (Facilitates Chromatin Transactions) histone chaperone was shown to assist first in the removal

of nucleosomes ahead of the transcribing RNA Pol II and next in their reassembly after polymerase passage. While we brought evidence that H4K31 modifications may contribute to gene regulation at least in *T. gondii*, studies in other Eukaryotes have underlined H4K31 as instrumental in the recruitment/mobilization of histone chaperone at transcribed genes. In budding yeast, H4K31 along with two proximal residues on the side of the nucleosome (i.e., H4R36 and H3L61) promotes the recruitment of the yFACT subunit Spt16 across transcribed genes as assessed by the typical change in Spt16 distribution which occupancy shifts toward the 3' ends of transcribed genes in the H4K31E yeast mutant (Nguyen et al., 2013).

The versatility of H4K31 goes even beyond these modifications since ubiquitylation of H4K31 has been reported in human cells as an additional regulatory PTM for transcription elongation (Kim et al., 2013). Indeed, it was shown that the histone H1.2 subtype while localized at target genes interacts with the elongating RNA Polymerase II, typified by phosphorylation of Ser2 on its carboxy terminal domain (CTD). Indeed, it was shown that upon interaction with the Ser2-phosphorylated carboxy terminal domain CTD of the active RNA Pol II, the histone H1.2 subtype becomes able to recruit the Cul4A E3 ubiquitin ligase and PAF1 elongation complexes. In turn, those stimulate H4K31 ubiquitylation that influences positively the accumulation of the H3K4me3/H3K79me2 signature, thereby leading to more productive elongation phase of transcription. Importantly, blocking H4K31 ubiquitylation by K31R mutation markedly reduces H3K4 and H3K79 methylation and consequently impairs gene transcription (Kim et al., 2013).

In order to test the functional significance of H4K31 modifications *in vivo*, we tried but remained unsuccessful at substituting H4K31 in *T. gondii* genome with alanine or glutamine to mimic acetyl lysine or with arginine to mimic nonacetylated lysine (data not shown). Engineered budding yeast with those substitutions did not significantly affect cell viability but led to an unexpected increase of telomeric and ribosomal DNA silencing (Hyland et al., 2005) that both argue for the mutations driving a non-permissive chromatin state. This does not fit with our working model in which, H4K31Q should promote open chromatin. It is however plausible that the substitutions did not faithfully mimic the effects of the modifications in these

instances. In sharp contrast with the aforementioned substitutions, glutamic acid (E) that mimics succinylated lysine was shown to severely compromise the growth in budding yeast (Xie et al., 2012) maybe as a consequence of the alteration in the distribution of Spt16 across yeast genes (Nguyen et al., 2013). The succinylation on H4K31 has also been detected by mass spectrometry in *T. gondii* (Li et al., 2014; Nardelli et al., 2013). The modification could drastically impact intranucleosomal structure and induce "abnormal" histone-DNA interactions (Fig. 1c).

In this context, H4K31 methylation would counteract the activating effect of H4K31 acetylation and succinylation, by preventing the nucleosome from adopting an open conformation permissive to gene expression. The analysis in *P. falciparum* revealed remarkable features of H4K31me1 by stressing a much more pronounced repressive character as the modification was exclusively restricted to non-permissive silenced chromosomal zones. Originally, *P. falciparum* heterochromatin in which clusters of genes are maintained in a silent state was singularly defined by increased nucleosomal occupancy, histone deacetylation, H3K9me3 and the binding of PfHP1 (Scherf et al., 2008). While most of the genome can be characterized as euchromatin, those silenced regions were organized towards the periphery of the nucleus and contain among others the *var*, rif and stevor families that cluster *together,* proximal to each telomere. The repression of the *var* genes for instance involves the trimethylation of H3K9 and its spreading to the next-door nucleosome by the action of HP1 (Scherf et al., 2008). H4K31me1 enrichment was detected, yet unevenly and at low rates in the vicinity of few *var* (5 over 64) and *rifin* genes (Fig. 12c). However, the modification does not spread while its enrichment fades quickly and remains likely limited to the site of heterochromatin initiation where both H3K9me3 and HP1 levels culminate (Fig. 12c). The lack of spread of H4K31me1 along a series of nucleosomes may be explained by the absence of a competent protein reader that specifically recognizes the PTM and recruits the H4K31me1-catalyzing methyltransferase. So far, no H4K31me1-reading protein was identified although the PTM is not buried and hence accessible for regulatory factor binding. In fact, the bromodomain of BRD4 is able to recognize the acetylated isoform of H4K31 (Filippakopoulos et al., 2012).

While H4K31me1 occupancy is overall limited across *P. falciparum* genome, the modification is by far the most promiscuous PTMs found at pericentromeric zones of all chromosomes (Fig. 11d and Fig. 12 – figures supplement 1 and 2). As such, both H4K31me1 (Fig. 7c) and centromeres (Hoeijmakers et al., 2012) were found to be clustered towards nuclear periphery. *P. falciparum* centromeres were originally described as displaying a unique epigenetic status typified by the noteworthy absence of the canonical pericentromeric PTM H3K9me3 (Hoeijmakers et al., 2012) present in all species including *T. gondii* (Fig. 10c and Fig. 10 – figure supplement 4) (Brooks et al., 2011). Clearly this study has emphasized an unusual role of H4K31me1 in pericentromeric heterochromatin in *P. falciparum* and have provided new insights on the mechanism of transcriptional regulation in *T. gondii*.

In metazoan, H4K31me1 was shown to decorate the mitotic chromosome arms (Fig. 7b). The PTM is in this regard a novel mitotic marker that targets newly synthesized H4 that may have be involved in the regulation of chromosomal condensation and segregation during mitosis. H4K31 is structurally very close to H3K56 (Fig. 1c), the acetylation of which reported to increase the binding affinity of H3 toward histone chaperones, thereby promoting nucleosome assembly during S phase of the cell cycle (Li et al., 2008). Collectively, our results argue for a similar role for H4K31me1 in chromatin assembly during DNA replication in metazoan. However the picture appears more complex since H4K31 methylation, unlike H3K56ac, is predicted to prevent histone exchange, thereby slowing histone turnover rate behind the replication forks which overall contributes to stabilize newly incorporated nucleosomes into chromatin.

In conclusion, we demonstrate that H4K31 acetylation and methylation are associated to very distinct nuclear functions in *T. gondii* and *P. falciparum*. Moreover, we demonstrate the evolvement of distinct epigenetic strategies in apicomplexan parasites to organize chromosome regions that are essential for cell division and gene expression.

Materials and Methods

Parasites and host cells culture

HFF primary cells (Bougdour et al., 2009) were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Invitrogen) supplemented with 10% heat inactivated Fetal Bovine Serum (FBS) (Invitrogen), 10mM (4-(2-hydroxyethyl)-1-piperazine ethanesulphonic acid) (HEPES) buffer pH 7.2, 2 mM L-glutamine and 50 µg ml of penicillin and streptomycin (Invitrogen). Cells were incubated at 37°C in 5% CO2. Type I (RH wild type and RH Δ*ku80*) and type II strains (Pru Δ*ku80*) of *T. gondii* were maintained *in vitro* by serial passage on monolayers of HFFs. *P. falciparum* 3D7 strain was grown in RPMI 1640 media supplemented with 0.5% Albumax II, 0.1mM Hypoxanthine and Gentamicin 10 mcg/ml. The culture was maintained at 2% hematocrit and 5% parasitemia. The parasites were grown at 37°C and at 1% O2, 5% CO2 and 94% N2 gas mixture concentration. The cultures were free of mycoplasma, as determined by qualitative PCR.

Immunofluorescence microscopy

T. gondii infecting HFF cells grown on coverslips were fixed in 3% formaldehyde for 20 min at room temperature, permeabilized with 0.1% (v/v) Triton X-100 for 15 min and blocked in Phosphate buffered saline (PBS) containing 3% (w/v) BSA. The cells were then incubated for 1 hour with primary antibodies followed by the addition of secondary antibodies conjugated to Alexa Fluor 488 or 594 (Molecular Probes). Nuclei were stained for 10 min at room temperature with Hoechst 33258. Coverslips were mounted on a glass slide with Mowiol mounting medium, and images were acquired with a fluorescence ZEISS ApoTome.2 microscope and images were processed by ZEN software (Zeiss). *P. falciparum* asexual blood life stages were washed with phosphate-buffered saline (PBS) and fixed in solution with 4% paraformaldehyde and 0.0075% glutaraldehyde in PBS for 30 min. After one wash with PBS, cells were permeabilized with 0.1% Triton X-100 in PBS for 10 min. Cells were washed twice with PBS, blocked with 3% bovine serum albumin (BSA) in PBS for 1 hour. The cells were then incubated for 1 hour with primary antibodies followed by the addition of secondary antibodies conjugated to Alexa Fluor 488 or 594 (Molecular Probes). Nuclei were stained for 30 min at room temperature with Hoechst 33258. The parasites were finally washed 2-3 times before loading on to glass slides

mixed with fluoro-gel (Electron Microscopy Sciences). Images were acquired with a fluorescence ZEISS ApoTome.2 microscope and images were processed by ZEN software (Zeiss).

HDACi treatments

The final concentration of histone deacetylase inhibitors dissolved in DMSO was, as described (Bougdour et al., 2013), FR-235222 (90nM), apicidin (100nM), HC-toxin (100nM), trichostatin A (100nM), scriptaid (100nM), APHA (100mM) and sodium butyrate (5mM). They were added to infected HFF cells for 18 hours. Halofuginone (10 nM) was shown to inhibit prolyl-tRNA synthetase (Jain et al., 2015) and was used as a control.

Plasmid constructs

To construct the vector pLIC-ENO1-HAFlag, the coding sequence of ENO1(TGME49_268860) was amplified using primers LIC-268860_Fwd (TACTTCCAATCCAATTTAGCgaacatgcaggcaatggcttggctcttc) and LIC-268860_Rev (TCCTCCACTTCCAATTTTAGCttttgggtgtcgaaagctctctcccgcg) using Pruku80 genomic DNA as template. The resulting PCR product was cloned into the pLIC-HF-dhfr vector using the LIC cloning method as reported previously (Bougdour et al., 2013).

Cas9-mediated gene disruption in *Toxoplasma gondii*

The plasmid pTOXO_Cas9-CRISPR was described in (Sangaré et al., 2016). For gene disruption using CRISPR/Cas9 system, the genes of interests (GOI) were: GCN5A (*TGGT1_254555*), GCN5B (*TGGT1_243440*), MYST-A (*TGGT1_318330*), MYST-B (*TGGT1_207080*), HAT1 (*TGGT1_293380*), HDAC1 (*TGGT1_281420*), HDAC2 (*TGGT1_249620*), HDAC3 (*TGGT1_227290*), HDAC4 (*TGGT1_257790*) and HDAC5 (*TGGT1_202230*). Twenty mers-oligonucleotides corresponding to specific GOI were cloned using Golden Gate strategy. Briefly, primers TgGOI-CRISP_FWD and TgGOI-CRISP_REV containing the sgRNA targeting TgGOI genomic sequence were phosphorylated, annealed and ligated into the linearized pTOXO_Cas9-CRISPR plasmid with BsaI, leading to pTOXO_Cas9-CRISPR::sgTgGOI. *T. gondii* tachyzoites were then transfected with the plasmid and grown on HFF cells for 18-36 hours. Cloning oligonucleotides used in this study:

TgHDAC1-CRISP-FWD : 5'- AAGTTGCGTCGCCGTTCTCTCACGCG -3'

TgHDAC1-CRISP-REV : 5'- AAAACGCGTGAGAGAACGGCGACGCA -3'

TgHDAC2-CRISP-FWD : 5'- AAGTTGCGCCCGTCGCCTCCCCCGCG -3'

TgHDAC2-CRISP-REV : 5'- AAAACGCGGGGGAGGCGACGGGCGCA -3'

TgHDAC3-CRISP-FWD : 5'- AAGTTGATATCGGAAGTTACTACTAG -3'

TgHDAC3-CRISP-REV : 5'- AAAACTAGTAGTAACTTCCGATATCA -3'

TgHDAC4-CRISP-FWD : 5'- AAGTTGCTGTTGCTGAAGCCCAGGCG -3'

TgHDAC4-CRISP-REV : 5'- AAAACGCCTGGGCTTCAGCAACAGCA -3'

TgHDAC5-CRISP-FWD : 5'- AAGTTGGCGAGACCGGGGCAGCCGCG -3'

TgHDAC5-CRISP-REV : 5'- AAAACGCGGCTGCCCCGGTCTCGCCA -3'

TgGCN5A-CRISP-FWD : 5'- AAGTTGCGTGACGAACGACAGGCAAG -3'

TgGCN5A-CRISP-REV : 5'- AAAACTTGCCTGTCGTTCGTCACGCA -3'

TgGCN5B-CRISP-FWD : 5'- AAGTTGGGTTTCCTGTGTCGAGACCG -3'

TgGCN5B-CRISP-REV : 5'- AAAACGGTCTCGACACAGGAAACCCA -3'

TgMYSTA-CRISP-FWD : 5'- AAGTTGGCTGCTCCGCGACTCAGCGG -3'

TgMYSTA-CRISP-REV : 5'- AAAACCGCTGAGTCGCGGAGCAGCCA -3'

TgMYSTB-CRISP-FWD : 5'- AAGTTGCGCGAAGAAGGGAGAGAGCG -3'

TgMYSTB-CRISP-REV : 5'- AAAACGCTCTCTCCCTTCTTCGCGCA -3'

TgHAT1-CRISP-FWD : 5'- AAGTTGCCGACGGGTCACGGAGACTG -3'

TgHAT1-CRISP-REV : 5'- AAAACAGTCTCCGTGACCCGTCGGCA -3'

***Toxoplasma gondii* transfection**

T. gondii RH, RH Δ*ku80* and Pru Δ*ku80* were electroporated with vectors in cytomix buffer (120mM KCl, 0.15mM $CaCl_2$, 10mM K_2HPO_4/ KH_2PO_4 pH7.6, 25mM HEPES pH7.6, 2mM EGTA, 5mM $MgCl_2$) using a BTX ECM 630 machine (Harvard Apparatus). Electroporation was performed in a 2mm cuvette at 1.100V, 25Ω and 25µF. Stable transgenic parasites were selected with 1µM pyrimethamine, single-cloned in 96 well plates by limiting dilution and verified by immunofluorescence assay.

Antibodies

Primary antibodies : rabbit home-made anti-TgHDAC3 described in (Bougdour et al., 2009; RRID:AB_2713903), mouse anti-HA (Roche, RRID:AB_2314622), rabbit anti-H4K8ac (Millipore, RRID:AB_310524), rabbit anti-H4K12ac (Millipore,

RRID:AB_11215637), rabbit anti-H3K4ac (Diagenode, RRID:AB_2713904), rabbit anti-H3K9ac (Diagenode RRID:AB_2713905), rabbit anti-H3K14ac (Diagenode, RRID:AB_2713906), rabbit anti-H3K18ac (Diagenode, RRID:AB_2713907), rabbit anti-H3K14ac (Millipore, RRID:AB_1977241) and mouse anti-H3K27ac (Diagenode, RRID:AB_2713908), H4K20me3 (Diagenode, RRID:AB_2713909), H3K9me3 (Millipore, RRID:AB_916348), H3K4me1 (Diagenode, RRID:AB_2637078) and H3K4me3 (Diagenode, RRID:AB_2616052). Western blot secondary antibodies were conjugated to alkaline phosphatase (Promega), while immunofluorescence secondary antibodies were coupled with Alexa Fluor 488 or Alexa Fluor 594 (Thermo Fisher Scientific). We also raised homemade H4K31acetylation and H4K31monomethylation-specific antibodies in rabbit against linear peptides corresponding to amino acid residues 23-35 of histone H4 and carrying modified residue K31: C-DNIQGITKme1PAIR; C-DNIQGITKacPAIR and C-RDNIQGITKacPAIR. They were produced by Eurogentec and used for immunofluorescence, immunoblotting and chromatin immunoprecipitation.

Histones purification, Immunoblotting and mass spectrometry-based proteomic analysis

For histone purification, HFF cells were grown to confluence and infected with PruΔ*ku80* parasites. Intracellular tachyzoites were treated with histone deacetylase HDAC3 inhibitor, 90nM FR235222 for 18 hours. As appropriate control, we treated tachyzoites with 0.1% DMSO. Histones were extracted and purified using histone purification kit (Active motif) according to manufacturer's protocol. For western blotting, histone proteins were run on a NuPAGE 4-12% Bis-Tris polyacrylamide gels in MES-SDS running buffer (Invitrogen) and transferred to a polyvinylidene fluoride PVDF membrane (Immobilon-P; Millipore) using NuPAGE transfer buffer (Invitrogen). The blots were probed using primary antibodies: pan acetyl H4, H4K31ac and H4K31me1, followed by phosphatase-conjugated goat secondary antibodies (*Promega*). The expected band of histones were detected using NBT-BCIP (*Amresco*). Nucleosomes from *T. gondii*-infected cells were purified and proteins separated by SDS-PAGE. The band corresponding to H4 was excised and its protein content digested using trypsin. Resulting peptides were submitted to mass spectrometry-based proteomic analysis (U3000 RSLCnano coupled to Q-Exactive HF,

Thermo Scientific). Peptides and proteins were identified using Mascot software (Matrix Science).

Immunodot blot assay

The MODifiedTM Histone Peptide Array (Activemotif) contains 59 different post-translational modifications for histone acetylation, methylation, phosphorylation and citrullination on the N-terminal tails of histones H2A, H2B, H3 and H4. Each 19mer peptide may contain up to four modifications each. Five control spots are included on each array : biotin peptide, c-Myc tag, no histone peptide and two background spots containing a mixture of modifications that are present on the array. Arrays were blocked with TTBS (10 mM Tris [pH 7.4], 150 mM NaCl, 0.05% Tween 20) plus 5% nonfat dry milk. Antibodies were diluted in TTBS. Primary antibodies were detected using HRP-conjugated anti-IgG antibodies (R&D systems). The blots were developed with the SuperSignal West Pico Chemiluminescent Substrate kit (Thermo Fisher Scientic). The Array Analyze software program designed by Activemotif was used to analyze the spot intensity of the interactions and to generate graphical analysis of the histone peptide-antibody interactions. The Spot Statistics tab gives a comparison of the intensities of each spot in the left and right array. The data points ideally should be on a straight line connecting 0,0 and 1,1 which indicates perfect duplicates. Outliers (or poor replicates) will not fall along the linear line. The distribution errors displays the errors of the intensities between the right and left spot, normalized to the maximum intensity. It is ideal to have most of the peptides contain an error range of 0-5% (which means very little variation between the left and right sides). The Reactivity tab displays the background subtracted intensity values for all single modified peptides, as well as for control spots. Selecting a modification from the pull down menu scales the intensity of the selected modification to 1.0 (Y-axis). The impact of neighboring modifications on the single modified peptides was further investigated using the Modification Analysis tab and selecting for the modification of interest. The results were graphed as a Specificity factor, which is the ratio of the average intensity of all spots containing the mark divided by the intensity of all spots not containing the mark. While all modifications have been accounted for in determining the specificity factor, only the 10 modifications with the highest specificity factor values are graphed. The larger the difference in specificity factor

values for the mark of interest versus other marks, the more specific the interaction. Non-specific signals will decrease the specificity factor values.

Chromatin Immunoprecipitation and Next Generation Sequencing in *Toxoplasma gondii*

HFF cells were grown to confluence and infected with type II (PruΔ*ku80*) strain. Harvested intracellular parasites were crosslinked with formaldehyde (final concentration 1%) for 8 min at room temperature and the crosslinking was stopped by addition of glycine (final concentration 0.125M) for 5 min at room temperature. Crosslinked chromatin was lysed in ice-cold lysis buffer (50mM HEPES KOH pH7.5, 140mM NaCl, 1mM EDTA, 10% glycerol, 0.5%NP-40, 0.125% triton X-100, protease inhibitor cocktail) and sheared in shearing buffer (1mM EDTA pH8.0, 0.5mM EGTA pH8.0, 10mM Tris pH8.0, protease inhibitor cocktail) by sonication using a Diagenode Biorupter. Samples were sonicated, for 16 cycles (30 seconds ON and 30 seconds OFF), to 200-500 base-pair average size. Immunoprecipitation was carried out using sheared chromatin, 5% BSA, protease inhibitor cocktail, 10% triton X-100, 10% deoxycholate, DiaMag Protein A-coated magnetic beads (Diagenode) and antibodies (H4K31ac, H4K31me1, pan acetyl H4, H4K20me3, H3K9me3, H3K4me3, H3K4me1, H3K14ac). A rabbit IgG antiserum was used as a control mock. After overnight incubation at 4°C on rotating wheel, chromatin-antibody complexes were washed and eluted from beads by using iDeal ChIP-seq kit for Histones (Diagenode) according to the manufacturer's protocol. Samples were decrosslinked by heating for 4 hours at 65°C. DNA was purified by using IPure kit (Diagenode) and quantified by using Qubit Assays (*Thermo Fisher Scientific*) according to the manufacturer's protocol. For ChIP-seq, purified DNA was used to prepare libraries and then sequenced by Arraystar (USA, http://www.arraystar.com/).

Library Preparation, Sequencing and Data analysis (Arraystar)

ChIP-Sequencing library preparation was performed according to Illumina's protocol Preparing Samples for ChIP Sequencing of DNA. **Library Preparation**: 10 ng DNA of each sample was converted to phosphorylated blunt-ended with T4 DNA polymerase, Klenow polymerase and T4 polymerase (NEB); An 'A' base was added to the 3' end of the blunt phosphorylated DNA fragments using the polymerase activity of Klenow (exo minus) polymerase (NEB); Illumina's genomic adapters were

ligated to the A tailed DNA fragments; PCR amplification was performed to enrich ligated fragments using Phusion High Fidelity PCR Master Mix with HF Buffer (Finnzymes Oy). The enriched product of ~200-700 bp was cut out from gel and purified. **Sequencing:** The library was denatured with 0.1M NaOH to generate single-stranded DNA molecules, and loaded onto channels of the flow cell at 8pM concentration, amplified in situ using TruSeq Rapid SR cluster kit (#GD-402-4001, Illumina). Sequencing was carried out by running 100 cycles on Illumina HiSeq 4000 according to the manufacturer's instructions. **Data Analysis:** After the sequencing platform generated the sequencing images, the stages of image analysis and base calling were performed using Off-Line Basecaller software (OLB V1.8). After passing Solexa CHASTITY quality filter, the clean reads were aligned to *T. gondii* reference genome (Tgo) using BOWTIE (V2.1.0). Aligned reads were used for peak calling of the ChIP regions using MACS V1.4.0. Statistically significant ChIP-enriched regions (peaks) were identified by comparison of two samples, using a p-value threshold of 10^{-5}. Then the peaks in each sample were annotated by the overlapped gene using the newest *T. gondii* database. The EXCEL/BED format file containing the ChIP-enriched regions was generated for each sample. **Data visualization:** The mapped 100 bp reads represent enriched DNA fragments by ChIP experiment. Any region of interest in the raw ChIP-seq signal profile can be directly visualized in genome browser. We use 10-bp resolution intervals (10-bp bins) to partition the stacked reads region, and count the number of reads in each bin. All the 10 bp resolution ChIP-seq profiles of each sample are saved as UCSC wig format files, which can be visualized in *T. gondii* Genome Browser (http://protists.ensembl.org/Toxoplasma_gondii/Info/ Index). All these raw and processed files can be found at Series GSE98806.

Chromatin Immunoprecipitation and Next Generation Sequencing in *P. falciparum*

Chromatin from synchronous rings stage parasites of 3D7 clone G7 was prepared and $3*10^8$ cells per ChIP used for the previously described protocol (Lopez-Rubio et al., 2013). Briefly, chromatin was crosslinked in 1% formaldehyde for 10 min (Sigma-Aldrich, #SZBD1830V), sheared to an average length of 300 bp using the BioRuptor Pico and individual histone modifications were pulled down using 0.5 µg of antibody for H3K4me3 (Diagenode, cat # K2921004), H3K9me3 (Millipore, cat # 257833),

and home-made rabbit polyclonal anti-PfHP1. 5 µl rabbit polyclonal anti-H4K31me1 and 15 µl anti-H4K31ac were used for each experiment. To generate Illumina-compatible sequencing libraries, the immunoprecipitated DNA and input was processed using the MicroPlex Library Preparation Kit (Diagenode C05010014) according to manufacturer's instructions. The optimized library amplification step was used KAPA Biosystems HIFI polymerase (KAPA Biosystems KK2101). Pooled, multiplexed libraries were sequenced on an Illumina NextSeq® 500/550 system as a 150 nucleotide single-end run. The raw data were demultiplexed using bcl2fastq2 (Illumina) and converted to fastq format files for downstream analysis. Two biological replicates were analyzed for each antibody.

***Plasmodium falciparum* ChIP-seq Data Analysis**

Sequencing reads were mapped to the *P. falciparum* 3D7 genome assembly (PlasmoDB v3.0) with Burrows-Wheeler Alignment tool (BWA) using default settings, and then sequences were quality filtered at Q20 Phred quality score. ChIP-seq peak calling was performed using the MACS2 algorithm. For genome-wide representation of each histone mark's distribution, the coverage was calculated as average reads per million over bins of 1000 nucleotides using bamCoverage from the package deepTools. Correlation of the different biological replicates were calculated by performing Pearson's and Spearman's correlation analysis of pairwise comparison of BAM alignment files, and ChIP-seq peak enrichment scores (log2) using MACS2 and deepTools. Circular and linear coverage plots were generated using Circos and Integrated Genomics Viewer, respectively. All these raw and processed files can be found at NCBI Bioproject ID PRJNA386433.

Acknowledgments

This work was supported by the Agence Nationale Recherche, France [grant ANR Blanc 2012 TOXOHDAC – grant number ANR-12-BSV3-0009-01]; the Laboratoire d'Excellence ParaFrap, France [grant number ANR-11-LABX-0024]; Atip-Avenir and Finovi programs [Apicolipid projects to CYB]; the European Research Council [ERC Consolidator Grant N°614880 Hosting TOXO to M.A.H] and [ERC AdG N° 670301 PlasmoSilencing to A.S.). Proteomic experiments were partly supported by the ProFi grant (ANR-10-INBS-08-01).

References

Azzalin, C.M., and Lingner, J. (2015). Telomere functions grounding on TERRA firma. Trends Cell Biol. *25*, 29–36.

Bhatti, M.M., Livingston, M., Mullapudi, N., and Sullivan, W.J. (2006). Pair of Unusual GCN5 Histone Acetyltransferases and ADA2 Homologues in the Protozoan Parasite Toxoplasma gondii. Eukaryot. Cell *5*, 62–76.

Bougdour, A., Maubon, D., Baldacci, P., Ortet, P., Bastien, O., Bouillon, A., Barale, J.-C., Pelloux, H., Ménard, R., and Hakimi, M.-A. (2009). Drug inhibition of HDAC3 and epigenetic control of differentiation in Apicomplexa parasites. J. Exp. Med. *206*, 953–966.

Bougdour, A., Braun, L., Cannella, D., and Hakimi, M.-A. (2010). Chromatin modifications: implications in the regulation of gene expression in Toxoplasma gondii. Cell. Microbiol. *12*, 413–423.

Bougdour, A., Durandau, E., Brenier-Pinchart, M.-P., Ortet, P., Barakat, M., Kieffer, S., Curt-Varesano, A., Curt-Bertini, R.-L., Bastien, O., Coute, Y., et al. (2013). Host cell subversion by Toxoplasma GRA16, an exported dense granule protein that targets the host cell nucleus and alters gene expression. Cell Host Microbe *13*, 489–500.

Bozdech, Z., Llinás, M., Pulliam, B.L., Wong, E.D., Zhu, J., and DeRisi, J.L. (2003). The Transcriptome of the Intraerythrocytic Developmental Cycle of Plasmodium falciparum. PLoS Biol. *1*, e5.

Brooks, C.F., Francia, M.E., Gissot, M., Croken, M.M., Kim, K., and Striepen, B. (2011). Toxoplasma gondii sequesters centromeres to a specific nuclear region throughout the cell cycle. Proc. Natl. Acad. Sci. *108*, 3767–3772.

Chatterjee, N., North, J.A., Dechassa, M.L., Manohar, M., Prasad, R., Luger, K., Ottesen, J.J., Poirier, M.G., and Bartholomew, B. (2015). Histone Acetylation near the Nucleosome Dyad Axis Enhances Nucleosome Disassembly by RSC and SWI/SNF. Mol. Cell. Biol. *35*, 4083–4092.

Cobbold, S.A., Santos, J.M., Ochoa, A., Perlman, D.H., and Llinás, M. (2016). Proteome-wide analysis reveals widespread lysine acetylation of major protein complexes in the malaria parasite. Sci. Rep. *6*, 19722.

Cosgrove, M.S., Boeke, J.D., and Wolberger, C. (2004). Regulated nucleosome mobility and the histone code. Nat. Struct. Mol. Biol. *11*, 1037–1043.

Daujat, S., Weiss, T., Mohn, F., Lange, U.C., Ziegler-Birling, C., Zeissler, U., Lappe, M., Schübeler, D., Torres-Padilla, M.-E., and Schneider, R. (2009). H3K64 trimethylation marks heterochromatin and is dynamically remodeled during developmental reprogramming. Nat. Struct. Mol. Biol. *16*, 777–781.

Di Cerbo, V., Mohn, F., Ryan, D.P., Montellier, E., Kacem, S., Tropberger, P., Kallis, E., Holzner, M., Hoerner, L., Feldmann, A., et al. (2014). Acetylation of histone H3 at lysine 64 regulates nucleosome dynamics and facilitates transcription. Elife *3*, e01632.

969 Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J.-P., Barsyte-
970 Lovejoy, D., Felletar, I., Volkmer, R., Müller, S., Pawson, T., et al. (2012). Histone
971 Recognition and Large-Scale Structural Analysis of the Human Bromodomain
972 Family. Cell *149*, 214–231.

973 Flueck, C., Bartfai, R., Volz, J., Niederwieser, I., Salcedo-Amaya, A.M., Alako,
974 B.T.F., Ehlgen, F., Ralph, S.A., Cowman, A.F., Bozdech, Z., et al. (2009).
975 Plasmodium falciparum Heterochromatin Protein 1 Marks Genomic Loci Linked to
976 Phenotypic Variation of Exported Virulence Factors. PLoS Pathog. *5*, e1000569.

977 Freitas-Junior, L.H., Bottius, E., Pirrit, L.A., Deitsch, K.W., Scheidig, C., Guinet, F.,
978 Nehrbass, U., Wellems, T.E., and Scherf, A. (2000). Frequent ectopic recombination
979 of virulence factor genes in telomeric chromosome clusters of P. falciparum. Nature
980 *407*, 1018–1022.

981 Garcia, B.A., Busby, S.A., Shabanowitz, J., Hunt, D.F., and Mishra, N. (2005).
982 Resetting the Epigenetic Histone Code in the MRL- *lpr* / *lpr* Mouse Model of Lupus
983 by Histone Deacetylase Inhibition. J. Proteome Res. *4*, 2032–2042.

984 Garcia, B.A., Hake, S.B., Diaz, R.L., Kauer, M., Morris, S.A., Recht, J., Shabanowitz,
985 J., Mishra, N., Strahl, B.D., Allis, C.D., et al. (2006). Organismal Differences in Post-
986 translational Modifications in Histones H3 and H4. J. Biol. Chem. *282*, 7641–7655.

987 Gissot, M., Walker, R., Delhaye, S., Huot, L., Hot, D., and Tomavo, S. (2012).
988 Toxoplasma gondii Chromodomain Protein 1 Binds to Heterochromatin and
989 Colocalises with Centromeres and Telomeres at the Nuclear Periphery. PLoS ONE *7*,
990 e32671.

991 Hehl, A.B., Basso, W.U., Lippuner, C., Ramakrishnan, C., Okoniewski, M., Walker,
992 R.A., Grigg, M.E., Smith, N.C., and Deplazes, P. (2015). Asexual expansion of
993 Toxoplasma gondii merozoites is distinct from tachyzoites and entails expression of
994 non-overlapping gene families to attach, invade, and replicate within feline
995 enterocytes. BMC Genomics *16*, 66.

996 Hoeijmakers, W.A.M., Flueck, C., Françoijs, K.-J., Smits, A.H., Wetzel, J., Volz,
997 J.C., Cowman, A.F., Voss, T., Stunnenberg, H.G., and Bártfai, R. (2012). *Plasmodium*
998 *falciparum* centromeres display a unique epigenetic makeup and cluster prior to and
999 during schizogony: Centromeres of *P . falciparum*. Cell. Microbiol. *14*, 1391–1401.

1000 Hong, D.-P., Radke, J.B., and White, M.W. (2017). Opposing Transcriptional
1001 Mechanisms Regulate *Toxoplasma* Development. mSphere *2*, e00347-16.

1002 Hyland, E.M., Cosgrove, M.S., Molina, H., Wang, D., Pandey, A., Cottee, R.J., and
1003 Boeke, J.D. (2005). Insights into the Role of Histone H3 and Histone H4 Core
1004 Modifiable Residues in Saccharomyces cerevisiae. Mol. Cell. Biol. *25*, 10060–10070.

1005 Iwasaki, W., Tachiwana, H., Kawaguchi, K., Shibata, T., Kagawa, W., and
1006 Kurumizaka, H. (2011). Comprehensive Structural Analysis of Mutant Nucleosomes
1007 Containing Lysine to Glutamine (KQ) Substitutions in the H3 and H4 Histone-Fold
1008 Domains. Biochemistry (Mosc.) *50*, 7822–7832.

Jain, V., Yogavel, M., Oshima, Y., Kikuchi, H., Touquet, B., Hakimi, M.-A., and Sharma, A. (2015). Structure of Prolyl-tRNA Synthetase-Halofuginone Complex Provides Basis for Development of Drugs against Malaria and Toxoplasmosis. Struct. Lond. Engl. 1993 *23*, 819–829.

Jeffers, V., and Sullivan, W.J. (2012). Lysine Acetylation Is Widespread on Proteins of Diverse Function and Localization in the Protozoan Parasite Toxoplasma gondii. Eukaryot. Cell *11*, 735–742.

Karmodiya, K., Krebs, A.R., Oulad-Abdelghani, M., Kimura, H., and Tora, L. (2012). H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. BMC Genomics *13*, 424.

Kim, K., Lee, B., Kim, J., Choi, J., Kim, J.-M., Xiong, Y., Roeder, R.G., and An, W. (2013). Linker Histone H1.2 Cooperates with Cul4A and PAF1 to Drive H4K31 Ubiquitylation-Mediated Transactivation. Cell Rep. *5*, 1690–1703.

Kuo, M.H., Brownell, J.E., Sobel, R.E., Ranalli, T.A., Cook, R.G., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996). Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. Nature *383*, 269–272.

Lawrence, M., Daujat, S., and Schneider, R. (2016). Lateral Thinking: How Histone Modifications Regulate Gene Expression. Trends Genet. *32*, 42–56.

Li, Q., Zhou, H., Wurtele, H., Davies, B., Horazdovsky, B., Verreault, A., and Zhang, Z. (2008). Acetylation of Histone H3 Lysine 56 Regulates Replication-Coupled Nucleosome Assembly. Cell *134*, 244–255.

Li, X., Hu, X., Wan, Y., Xie, G., Li, X., Chen, D., Cheng, Z., Yi, X., Liang, S., and Tan, F. (2014). Systematic Identification of the Lysine Succinylation in the Protozoan Parasite *Toxoplasma gondii*. J. Proteome Res. *13*, 6087–6095.

Lopez-Rubio, J.-J., Mancio-Silva, L., and Scherf, A. (2009). Genome-wide Analysis of Heterochromatin Associates Clonally Variant Gene Regulation with Perinuclear Repressive Centers in Malaria Parasites. Cell Host Microbe *5*, 179–190.

Lopez-Rubio, J.-J., Siegel, T.N., and Scherf, A. (2013). Genome-wide chromatin immunoprecipitation-sequencing in Plasmodium. Methods Mol. Biol. Clifton NJ *923*, 321–333.

Mancio-Silva, L., and Scherf, A. (2013). In situ fluorescence visualization of transcription sites and genomic Loci in blood stages of Plasmodium falciparum. Methods Mol. Biol. Clifton NJ *923*, 335–351.

Miao, J., Lawrence, M., Jeffers, V., Zhao, F., Parker, D., Ge, Y., Sullivan, W.J., and Cui, L. (2013). Extensive lysine acetylation occurs in evolutionarily conserved metabolic pathways and parasite-specific functions during *Plasmodium falciparum* intraerythrocytic development: The malaria parasite acetylome. Mol. Microbiol. *89*, 660–675.

Moraes, I., Yuan, Z.-F., Liu, S., Souza, G.M., Garcia, B.A., and Casas-Mollano, J.A. (2015). Analysis of histones H3 and H4 reveals novel and conserved post-translational modifications in sugarcane. PloS One *10*, e0134586.

Morris, S.A., Rao, B., Garcia, B.A., Hake, S.B., Diaz, R.L., Shabanowitz, J., Hunt, D.F., Allis, C.D., Lieb, J.D., and Strahl, B.D. (2006). Identification of Histone H3 Lysine 36 Acetylation as a Highly Conserved Histone Modification. J. Biol. Chem. *282*, 7632–7640.

Nardelli, S.C., Che, F.-Y., Silmon de Monerri, N.C., Xiao, H., Nieves, E., Madrid-Aliste, C., Angel, S.O., Sullivan, W.J., Angeletti, R.H., Kim, K., et al. (2013). The Histone Code of Toxoplasma gondii Comprises Conserved and Unique Posttranslational Modifications. mBio *4*, e00922-13-e00922-13.

Neumann, H., Hancock, S.M., Buning, R., Routh, A., Chapman, L., Somers, J., Owen-Hughes, T., van Noort, J., Rhodes, D., and Chin, J.W. (2009). A Method for Genetically Installing Site-Specific Acetylation in Recombinant Histones Defines the Effects of H3 K56 Acetylation. Mol. Cell *36*, 153–163.

Nguyen, H.-T.T., Wharton, W., Harper, J.A., Dornhoffer, J.R., and Duina, A.A. (2013). A Nucleosomal Region Important for Ensuring Proper Interactions Between the Transcription Elongation Factor Spt16 and Transcribed Genes in Saccharomyces cerevisiae. G3amp58 GenesGenomesGenetics *3*, 929–940.

Pittman, K.J., Aliota, M.T., and Knoll, L.J. (2014). Dual transcriptional profiling of mice and Toxoplasma gondii during acute and chronic infection. BMC Genomics *15*, 806.

Radke, J.R., Behnke, M.S., Mackey, A.J., Radke, J.B., Roos, D.S., and White, M.W. (2005). The transcriptome of Toxoplasma gondii. BMC Biol. *3*, 26.

Ramakrishnan, C., Walker, R.A., Eichenberger, R.M., Hehl, A.B., and Smith, N.C. (2017). The merozoite-specific protein, TgGRA11B, identified as a component of the Toxoplasma gondii parasitophorous vacuole in a tachyzoite expression model. Int. J. Parasitol.

Salcedo-Amaya, A.M., van Driel, M.A., Alako, B.T., Trelle, M.B., van den Elzen, A.M., Cohen, A.M., Janssen-Megens, E.M., van de Vegte-Bolmer, M., Selzer, R.R., Iniguez, A.L., et al. (2009). Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of Plasmodium falciparum. Proc. Natl. Acad. Sci. *106*, 9655–9660.

Sangaré, L.O., Alayi, T.D., Westermann, B., Hovasse, A., Sindikubwabo, F., Callebaut, I., Werkmeister, E., Lafont, F., Slomianny, C., Hakimi, M.-A., et al. (2016). Unconventional endosome-like compartment and retromer complex in Toxoplasma gondii govern parasite integrity and host infection. Nat. Commun. *7*, 11191.

Saraf, A., Cervantes, S., Bunnik, E.M., Ponts, N., Sardiu, M.E., Chung, D.-W.D., Prudhomme, J., Varberg, J.M., Wen, Z., Washburn, M.P., et al. (2016). Dynamic and

Combinatorial Landscape of Histone Modifications during the Intraerythrocytic Developmental Cycle of the Malaria Parasite. J. Proteome Res. *15*, 2787–2801.

Sautel, C.F., Cannella, D., Bastien, O., Kieffer, S., Aldebert, D., Garin, J., Tardieux, I., Belrhali, H., and Hakimi, M.-A. (2007). SET8-mediated methylations of histone H4 lysine 20 mark silent heterochromatic domains in apicomplexan genomes. Mol. Cell. Biol. *27*, 5711–5724.

Scherf, A., Lopez-Rubio, J.J., and Riviere, L. (2008). Antigenic Variation in *Plasmodium falciparum*. Annu. Rev. Microbiol. *62*, 445–470.

Sierra-Miranda, M., Delgadillo, D.M., Mancio-Silva, L., Vargas, M., Villegas-Sepulveda, N., Martínez-Calvillo, S., Scherf, A., and Hernandez-Rivas, R. (2012). Two long non-coding RNAs generated from subtelomeric regions accumulate in a novel perinuclear compartment in Plasmodium falciparum. Mol. Biochem. Parasitol. *185*, 36–47.

Soldi, M., Cuomo, A., and Bonaldi, T. (2014). Improved bottom-up strategy to efficiently separate hypermodified histone peptides through ultra-HPLC separation on a bench top Orbitrap instrument. Proteomics *14*, 2212–2225.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41–45.

Trelle, M.B., Salcedo-Amaya, A.M., Cohen, A.M., Stunnenberg, H.G., and Jensen, O.N. (2009). Global histone analysis by mass spectrometry reveals a high content of acetylated lysine residues in the malaria parasite Plasmodium falciparum. J. Proteome Res. *8*, 3439–3450.

Tropberger, P., and Schneider, R. (2013). Scratching the (lateral) surface of chromatin regulation by histone modifications. Nat. Struct. Mol. Biol. *20*, 657–661.

Tropberger, P., Pott, S., Keller, C., Kamieniarz-Gdula, K., Caron, M., Richter, F., Li, G., Mittler, G., Liu, E.T., Bühler, M., et al. (2013). Regulation of Transcription through Acetylation of H3K122 on the Lateral Surface of the Histone Octamer. Cell *152*, 859–872.

Turner, B.M. (2000). Histone acetylation and an epigenetic code. BioEssays News Rev. Mol. Cell. Dev. Biol. *22*, 836–845.

Vanagas, L., Jeffers, V., Bogado, S.S., Dalmasso, M.C., Sullivan, W.J., and Angel, S.O. (2012). *Toxoplasma* histone acetylation remodelers as novel drug targets. Expert Rev. Anti Infect. Ther. *10*, 1189–1201.

Voigt, P., Tee, W.-W., and Reinberg, D. (2013). A double take on bivalent promoters. Genes Dev. *27*, 1318–1338.

Voss, T.S., Bozdech, Z., and Bártfai, R. (2014). Epigenetic memory takes center stage in the survival strategy of malaria parasites. Curr. Opin. Microbiol. *20*, 88–95.

Williams, S.K., Truong, D., and Tyler, J.K. (2008). Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation. Proc. Natl. Acad. Sci. *105*, 9000–9005.

Wiśniewski, J.R., Zougman, A., and Mann, M. (2008). N ε -Formylation of lysine is a widespread post-translational modification of nuclear proteins occurring at residues involved in regulation of chromatin function. Nucleic Acids Res. *36*, 570–577.

Xie, Z., Dai, J., Dai, L., Tan, M., Cheng, Z., Wu, Y., Boeke, J.D., and Zhao, Y. (2012). Lysine Succinylation and Lysine Malonylation in Histones. Mol. Cell. Proteomics *11*, 100–107.

Xue, B., Jeffers, V., Sullivan, W.J., and Uversky, V.N. (2013). Protein intrinsic disorder in the acetylome of intracellular and extracellular Toxoplasma gondii. Mol. Biosyst. *9*, 645.

Yang, X., Yu, W., Shi, L., Sun, L., Liang, J., Yi, X., Li, Q., Zhang, Y., Yang, F., Han, X., et al. (2011). HAT4, a Golgi Apparatus-Anchored B-Type Histone Acetyltransferase, Acetylates Free Histone H4 and Facilitates Chromatin Assembly. Mol. Cell *44*, 39–50.

Figure legends

Figure 1. The residue K31 on the lateral surface of histone H4 is a novel PTM. (**a**) The high resolution of MS/MS spectrum of H4K31ac peptide generated from histone H4. H4K31ac was identified using Mascot search engine in the DNIQGITK(ac)PAIR peptide. (**b**) Sequence alignment of the first 42 residues of histone H4 from the indicated organisms. Yellow boxes highlight the conserved residue H4K31. (**c**) Structural analysis of H4K31 modifications. Nucleosome core particle with key H3 and H4 lysine residues that are known to be modified shown in ball-and-sphere representation. The histone proteins of the nucleosome (PDB code: 3AFA) are colour coded as follows: H2A cyan, H2B grey, H3 orange and H4 blue. The H4K31 residue, highlighted in red, is placed at the dyad axis and mediates key interactions to the DNA (in green). The bottom panel is rotated 90 degrees around the molecular dyad axis. On the right, close-up of the interactions established by H4K31 with a water molecule (red sphere) and residue R35; and impact of the modifications: methylation, acetylation and succinylation (mimicked by mutant K31E). The mutant H4K31Q (PDB code: 3AZI) partially mimics lysine acetylation. (**d**) Immunofluorescence analysis of H4K31ac (in red) in both human foreskin fibroblast cells and parasites nuclei. DNA (top) was stained

with Hoechst. Scale bar, 10 μm. (**e**) Immunoblots of native purified nucleosomes from *T. gondii* parasites treated with FR235222 or DMSO for 18 hours. Data are representative of two independent experiments.

Figure 2 Immunofluorescence analysis of histone modifications in human cell infected with *T. gondii*. (**a**) Confluent monolayers of HFF cells were infected with tachyzoites in the presence of FR235222 and DMSO as a control. IFAs were carried out with antibodies against specific acetylated histones H3 and H4 lysine residues as indicated. All modifications, showed specific and distinct localization in both parasite and host cell nuclei (in red). Scale bar, 20 μm. (b) Quantification of the intensity of the aforementioned PTMs staining in each parasite nucleus following FR235222 stimulation. Each symbol marks the PTM density of a single parasite nucleus. The results are represented as mean ± standard deviations from two independent experiments ; the number of nuclei quantified was at least n=50. Asterisks indicate statistical for each individual PTM significance between FR235222-treated sample and the corresponding control (DMSO) as determined by an unpaired two-tailed Student t test, **P<0.05 and ***P<0.0001 ; n.s., not significant. Fold induction of each PTM by FR235222 in parasite nuclei was indicated.

Figure 3 Immunofluorescence analysis of histone modifications in *P. falciparum* and MEF cells. (**a**) The blood stages of *P. falciparum* characterized by initial Ring, followed by mature trophozoite and segmented schizont stage. The three developmental stages represent the predominant asexual phase of the malaria parasite. (**b**) Immunofluorescence analysis of H4K31ac (in red) in asexual stages following 12 hours of treatment with DMSO (vehicle) or FR235222 HDACi. Parasite nuclear DNA was stained with Hoechst (blue). Scale bar, 10 μm. (**c**) Quantification of the intensity of H4K31ac staining in each *P. falciparum* nucleus following FR235222 stimulation of asexual stages. The results are represented as mean ± standard deviations from four independent experiments; the number of nuclei quantified was at least n=25. Asterisks indicate statistical for H4K31ac significance between FR235222-treated sample and the corresponding control (DMSO) in ring, trophozoite and schizont as determined by an unpaired two-tailed Student t test, ***P<0.0001 ; n.s., not significant. (**d**) and (**e**) Immunofluorescence analysis of H4K31ac in mouse

embryonic fibroblasts. DNA was stained with DAPI (blue); the bright foci mark pericentromeric heterochromatin. The signal for H4K31ac along with H3K27ac, H3K4ac or H3K9ac are enriched in euchromatic regions as shown in the merge. The mark is excluded from the DAPI dense foci that are associated with H3K9me3 and H4K20me3. Scale bar, 10 μm. Data are representative of three independent experiments.

Figure 4 Chemical inactivation of TgHDAC3 promotes H4K31ac accumulation in *T. gondii* nuclei. (**a**) Immunofluorescence analysis of H4K31ac (in red) in HFF cells infected with parasites expressing a HAFlag (HF)-tagged copy of the bradyzoite gene *ENO1* treated for 18 hours with vehicle (DMSO) or individual HDAC inhibitors, including short-chain fatty acids (sodium butyrate), cyclic tetrapeptides and hydroxamic acids classes. Halofuginone (HF), a non-HDACi anticoccidial compound was used as a relevant control. ENO1 expression was detected by IFA in parasite nuclei (anti-HA, in green). Host-cell and parasite nuclei were stained with Hoechst. Scale bar, 20 μm. (**b**) Quantification of the intensity of H4K31ac staining in each parasite nucleus following HDAC inhibitors and vehicle (DMSO) stimulations. Each symbol marks the PTM density of a single parasite nucleus. The results are represented as mean ± s.d. from three independent experiments ; the number of nuclei quantified was at least n = 50. Asterisks indicate statistical for H4K31ac significance between DMSO and each HDACi as determined by a two-way ANOVA with Bonferroni's multiple comparison test, ***P<0.0001 ; n.s., not significant.

Figure 5 Genetic inactivation of TgHDAC3 promotes H4K31ac accumulation in *T. gondii* nuclei. (**a**) Representative micrographs showing intracellular tachyzoites in which the *TgHDAC3* gene was disrupted by transient transfection of CRISPR/Cas9. The efficiency of TgHDAC3 disruption in Cas9-expressing parasites was monitored by the anti-TgHDAC3 staining (in pink) and cas9-GFP expression (in green). The levels of H4K31ac (in red) were monitored in *TgHDAC3*-disrupted parasites (GFP positive) and compared to untransfected parasites (GFP negative). Scale bar, 10 μm. (**c**) Levels of H4K31ac (in red) were monitored in *TgHDAC* knockout parasites. Scale bar, 10 μm. Data are representative of four independent experiments. (**b**) Quantification of the intensity of the H4K31ac nuclear staining of *T. gondii*

expressing gRNA directed against TgHDAC as indicated (GFP positive cells) or left untransfected. Each symbol marks the H4K31ac density of a single parasite nucleus. The results are represented as mean ± s.d. from four independent experiments ; the number of nuclei quantified was at least n=60. Asterisks indicate statistical for H4K31ac significance between cas9-GFP positive and untransfected parasites as determined by an unpaired two-tailed Student *t* test, ***P<0.0001 ; n.s., not significant.

Figure 6 TgGCN5b acetylates H4K31 in *Toxoplasma gondii* (**a**) Levels of H4K31ac (in red) were monitored in both host cell and parasite nuclei following CRISPR/Cas9-mediated disruption of individual *T. gondii* HAT enzymes. Transfected vacuole in which H4K31 acetylation was impaired are indicated by a white arrow. Scale bar, 10 μm. (**b**) Quantification of the intensity of the H4K31ac nuclear staining of *T. gondii* expressing the gRNAs directed against TgHAT as indicated (GFP positive cells) or left untransfected. Each symbol marks the H4K31ac density of a single parasite nucleus. The results are represented as mean ± s.d. from three independent experiments ; the number of nuclei quantified was at least n=40. Asterisks indicate statistical for H4K31ac significance between cas9-GFP positive and untransfected parasites as determined by an unpaired two-tailed Student *t* test, ***P<0.0001 ; n.s., not significant. (**c**) Levels of H3K14 acetylation (in red) were monitored in *TgGCN5b* knockout parasites. Scale bar, 10 μm. (**d**) Quantification of the intensity of the H3K14ac nuclear staining of *T. gondii* expressing the gRNAs directed against *gcn5b* (GFP positive cells) or left untransfected. Data are representative of three independent experiments. ***P<0.0001 (unpaired two-tailed Student *t* test).

Figure 7 Distribution of H4K31me1 in host cell and parasites nuclei (**a**) Immunofluorescence analysis of H4K31me1 (in green) in intracellular parasite nuclei. DNA was stained with Hoechst (blue). Scale bar, 10 μm. (**b**) The localization of H4K31me1 (in red), H3K27ac (in green) and H3S10P (in green) were tested by immunofluorescence in prophase, metaphase, anaphase and telophase MEFs. DNA was stained with Hoechst (blue). Scale bar, 10 μm. Quantification of H4K31me1 and H3S10P staining for each mouse embryonic cell analysed in its respective phase of

the cell cycle. In total, 45 cells were scored for each phase. Each dot represents a single cell nucleus. The results are represented as mean ± s.d. from three independent experiments. Asterisks indicate statistical for H4K31me1 or H3S10P significance between interphase (background control) and each phase of the cell cycle as determined by a two-way ANOVA with Bonferroni's multiple comparison test, ***P<0.0001 ; n.s., not significant. (**c**) Immunofluorescence analysis of H4K31me1 (in red) or H3K9me3 in asexual stages of Pf-3D7. Scale bar, 5 μm. Data are representative of three independent experiments.

Figure 8 Genome-wide analysis of H4K31ac and H4K31me1 chromatin occupancy in *Toxoplasma gondii*. (**a**) A genome browser (IGB) snapshot showing normalized reads for different histone marks across *T. gondii* chromosome 1b reveals peak-like distribution of H4K31ac and H4K31me1 ChIP-seq enrichments. The y-axis depicts read density. Genes are depicted above the profiles in black. (**b**) A zoomed-in view of Chr. Ib region (yellow box in **a**) showing the distribution of the aforementioned PTMs. (**c**) Distribution of PTMs occupied regions relative to the *T. gondii* reference genome annotation. (**d**) Correlation of H4K31 modifications enrichment with other marks. The average signal profiles of each histone modification was plotted over a -2-kb to +10-kb region with respect to *T. gondii* genes ATG. The y-axis shows the average tag count of the enrichment. The vertical dashed line indicates the position of the ATG.

Figure 9 The enrichment of H4K31ac and H4K31me1 at transcribed genes correlates with gene expression levels in *Toxoplasma gondii*. (**a**) and (**b**) IGB screenshots of *T. gondii* Chr. VIII genomic regions showing reads for various histone marks as well as RNA-seq data (in black). (**c**) A zoomed-in view of *T. gondii GRA1-MAG1* locus. The y-axis depicts read density. (**d**) Boxplot showing the normalized expression distribution of *T. gondii* genes in tachyzoite stage subdivided into four ranges of expression (cluster Q1 to Q4). Genomewide H4K31ac (**e**) and H4K31me1 (**f**) occupancy profiles at peri-ATG regions are plotted for the gene groups ranked by their mRNA levels. The y-axis shows the average tag count of the enrichment. The vertical dashed line indicates the position of the ATG.

Figure 10 H4K31me1 marks long non-coding RNA but not pericentromeric heterochromatin in *Toxoplasma gondii*. (**a**) Left panel : IGB screenshot of *T. gondii* *ENO1* and *ENO2* locus on Chr. VIII showing reads for various histone marks as well as RNA-seq data (in black) and predicted genes (in magenta). The y-axis depicts read density. Right panel : bar graphs showing expression (FPKM values) of *ENO1* and *ENO2* genes during acute (tachyzoite) or chronic (bradyzoite) infection in mice (data source : ToxoDB; Pittman et al., 2014). (**b**) Left panel : IGB screenshot of two sexual stages genes (in red) surrounding an house-keeping gene (in blue). Right panel : bar graphs showing expression (FPKM values) of the genes in cultured tachyzoite (T) and parasites harvested from cat at day 3 (D3 ; merozoite), day 5 (D5 ; merozoite and sexual stages) and day 7 (D7 ; sexual stages and oocysts) (data source : ToxoDB; Hehl et al., 2015). (**c**) A zoomed-in view of *T. gondii* Chr. X peri-centromeric region and neighboring genes including the gene *TGME49_223262* which expression is restricted to sexual stages as seen in the bar graph. (**d**) IGB screenshot of *T. gondii* Chr. III genomic region showing reads for various histone marks as well as RNA-seq data. A predicted lncRNA of 72-kb is indicated in magenta.

Figure 11 Genome-wide analysis of H4K31ac and H4K31me1 chromatin occupancy in *Plasmodium falciparum*. (**a**) Chromosomal projection of H4K31ac, H3K9me3 and H3K4me3 occupancies in *P. falciparum*. The full set of chromosomes is represented as the circular plot. (**b**) Zoomed-in view of PTMs and HP1 enrichment along the *gapdh* locus. (**c**) Genomewide H4K31ac and H3K4me3 occupancy profiles at peri-TSS (Transcription Start Sites) and –TTS (Transcription Termination Sites) regions were plotted. (**d**) Chromosomal projection of H4K31me1, H3K9me3 and HP1 occupancies in *P. falciparum*. The full set of chromosomes is represented as the circular plot where CenH3 locations (black arrow) and var genes (in green) are indicated. (**e**) Genomewide H4K31me1 occupancy profiles at peri-TSS (Transcription Start Sites) and –TTS (Transcription Termination Sites) regions were plotted. (**f**) IGB view of a section of chromosome 10 showing enrichment of H4K31me1, H3K9me3 and HP1 at *var* gene.

Figure 12 H4K31me1 singularly marks peri-centromeric heterochromatin in *Plasmodium falciparum.* (**a**) (top) Chromosome-wide coverage plot of histone modifications and PfHP1 on *P. falciparum* Chr. 7. CenH3 was mapped according to Hoeijmakers et al., 2012 and *var* genes were indicated. (bottom) Zoomed-in views of *var* genes-containing internal locus (left panel) and centromeric (right panel) from *P. falciparum* Chr. 7. (**b**) Genomic organization and nuclear position of var genes and telomere-associated repeat elements (TAREs) in *P. falciparum*. (**c**) IGB screenshots of *P. falciparum* sub-telomeric regions of chromosome 10. Rifin and var genes as well as TAREs are highlighted.

Figure 1 – figure supplement 1 Specific binding of to home-made H4K31ac-directed antibody to H4K31 acetylated peptide *in vitro*. (**a**) Peptides with acetylated (ac) and unmodified (um) H4K31 were spotted at 10 or 1000 pmol and detected with home-made H4K31ac-directed antibody or with the control anti-H3K14ac. (**b**) A 59 PTMs-containing MODified Histone Peptide Array (from *Activemotif*) was incubated with H4K31ac-directed antibody (1/2000) and (**c - d**) the signals were analyzed according to the manufacturer (see material and method).

Figure 1 – figure supplement 2 Specific binding of to H3K14ac-directed antibody to H3K14ac-containing peptides *in vitro*. (**a**) A 59 PTMs-containing MODified Histone Peptide Array (from *Activemotif*) was incubated with H3K14ac-directed antibody (1/2000, from *Chemicon*) and (**b - d**) the signals were analyzed according to the manufacturer (see material and method).

Figure 1 – figure supplement 3 Specific binding of to home-made H4K31me1-directed antibody to H4K31 methylated peptide *in vitro*. (**a**) Peptides with monomethylated (me1) and unmodified (um) H4K31 were spotted at 10 or 1000 pmol and detected with home-made H4K31me1-directed antibodies or with the control anti-H4K20me1 (described in Sautel et al., 2007). (**b**) A 59 PTMs-containing MODified Histone Peptide Array (from *Activemotif*) was incubated with H4K31me1-directed antibody (1/1000) as well as the internal control c-Myc antibody (1/2000) and (**c - d**) the signals were analyzed according to the manufacturer (see material and method).

Figure 1 – figure supplement 4 Specific binding of to H4K20me1-directed antibody to H4K20me1-containing peptides *in vitro*. (**a**) A 59 PTMs-containing MODified Histone Peptide Array (from *Activemotif*) was incubated with H4K20me1-directed antibody (1/3000; described in Sautel et al., 2007) and (**b - d**) the signals were analyzed according to the manufacturer (see material and method).

Figure 8 – figure supplement 1 ChIP-seq enrichments between biological replicates are highly correlated. Scatterplot comparing the enrichment difference of H4K31ac (**a**) or H4K31me1 (**b**) measured in the two independent replicate experiments. The x- and y-axis show the average tag count of the enrichment. IGB screenshot of *T. gondii* Chr. X (**c**) and Ia (**d**) genomic regions showing reads for H4K31ac (replicates R1 and R2) and H4K31me1 (replicates R1 and R2).

Figure 8 – figure supplement 2 H4K31me1 explicitly predicts unannotated genes. IGB screenshot of *T. gondii* Chr. VIIb (**a**) and VIII (**b**) genomic regions showing reads for various histone marks as well as RNA-seq data (in black). Predicted genes are indicated in magenta along with their putative translated sequence.

Figure 8 – figure supplement 3 H4K31me1 and H3K4me1 are both enriched at gene body in *T. gondii*. (**a**) Scatterplot comparing the enrichment difference of H4K31me1and H3K4me1. The x- and y-axis show the average tag count of the enrichment. (**b**) Genomewide correlation of H4K31me1and H3K4me1 enrichment at peri-ATG regions. The average signal profiles of each histone modification was plotted over a -2-kb to +10-kb region with respect to *T. gondii* genes ATG. The y-axis shows the average tag count of the enrichment. The vertical dashed line indicates the position of the ATG. (**c**) and (**d**) IGB screenshot of *T. gondii* Chr. XI (**c**) and VIIb (**d**) genomic regions showing reads for various histone marks as well as RNA-seq data (in black). Predicted genes are indicated in magenta along with their putative translated sequence.

Figure 9 – figure supplement 1 PTMs distribution and gene expression in *Toxoplasma gondii*. Genomewide PTM occupancy profiles at peri-ATG regions are plotted for the gene groups ranked by their mRNA levels (**a**). H3K14ac (**b**),

H3K4me3 (**c**), H3K4me1 (**d**), and H3K9me3 (**e**) are shown. The y-axis shows the average tag count of the enrichment. The vertical dashed line indicates the position of the ATG.

Figure 10 – figure supplement 1 (**a**) IGB screenshot of *T. gondii BAG1* locus on Chr. VIIb showing reads for various histone marks as well as RNA-seq data (in black) and predicted genes (in magenta). The y-axis depicts read density. Right panel : bar graphs showing expression (FPKM values) of genes during acute and chronic infection in mice. (**b**) Heatmap representation of chronic-stage bradyzoite SRS genes expression in different life cycle stages. Gene expression values were mean log2 transformed and median centered for clustering. Transcriptomic data from cultured tachyzoite (T) and sexual stages isolated from mice at day 3 (D3 ; merozoite), 5 (D5 ; merozoite and sexual stages) and 7 (D7 ; sexual stages and oocysts) post-infection (Hehl et al., 2015) as well as those from acute and chronic murine toxoplasmosis (Pittman et al., 2014) are available on ToxoDB. (**c**) and (**d**) IGB screenshots of selected bradyzoite SRS genes showing reads for various histone marks as well as RNA-seq data (in black) and predicted genes (in magenta). The y-axis depicts read density.

Figure 10 – figure supplement 2 (**a**) Heatmap representation of *T. gondii* genes specifically expressed in sexual stages as described above. (**b**), (**c**) and (**d**) IGB screenshots of selected *T. gondii* genes whose expression is restricted to sexual stages. Reads are shown for various histone marks as well as RNA-seq data (in black) and predicted genes (in magenta). The y-axis depicts read density.

Figure 10 – figure supplement 3 (**a**) Heatmap representation of *T. gondii* SRS genes specifically expressed in sexual stages as described above. (**b**), (**c**) (**d**) and (**e**) IGB screenshots of SRS genes whose expression is restricted to sexual stages. Reads are shown for various histone marks as well as RNA-seq data (in black) and predicted genes (in magenta). The y-axis depicts read density.

Figure 10 – figure supplement 4 *Toxoplasma gondii* peri-centromeric regions (**a**) IGB screenshots of *T. gondii* peri-centromeric region of chromosomes Ia, II, III, VI, VIIa, VIII, IX and X.

Figure 10 – figure supplement 5 H4K31me1 marks long non-coding RNAs. IGB screenshot of *T. gondii* Chr. VIIa (**a**) and XI (**b**) genomic regions showing reads for various histone marks as well as RNA-seq data. Predicted lncRNAs of (**a**) 70-kb and (**b**) 22-kb are indicated in magenta.

Figure 11 – figure supplement 1 Correlation matrix between *Plasmodium falciparum* ChIP-seq experiments. Heatmap displaying (**a**) Pearson and (**b**) Spearman rank correlations between all pairwise comparisons for all *P. falciparum* ChIPs. Spearman correlations were calculated using the normalized read depth across the entire set of binding sites identified for all ChIP-seq experiments.

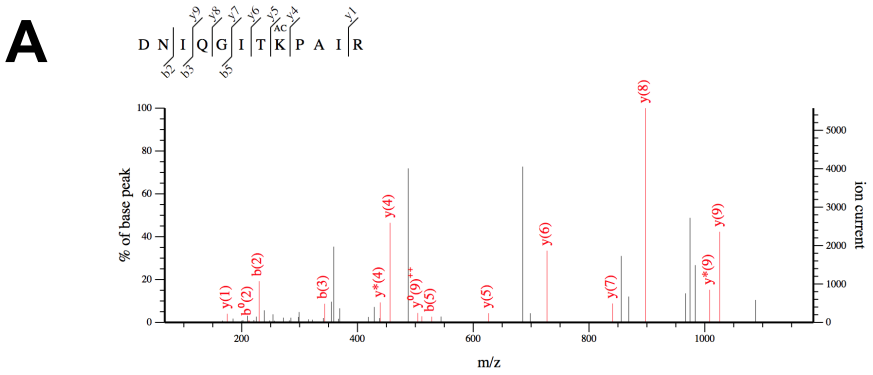
Figure 11 – figure supplement 2 Chromosomal projection of H4K31ac, H3K9me3 and H3K4me3 occupancies in *P. falciparum*. The full set of chromosomes is represented as the circular plot.

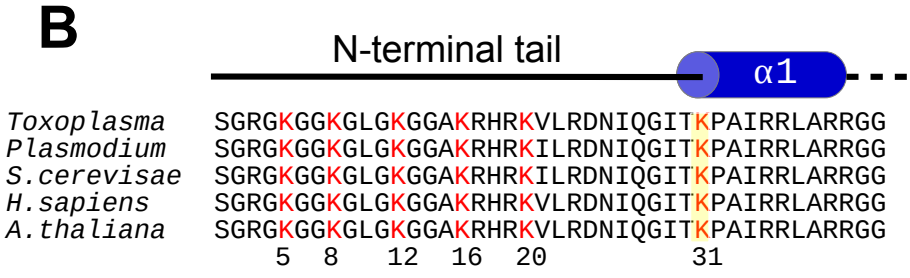
Figure 11 – figure supplement 3 Chromosomal projection of H4K31me1, H3K9me3 and HP1 occupancies in *P. falciparum.* The full set of chromosomes is represented as the circular plot, where centromeric regions are marked by a black arrow.

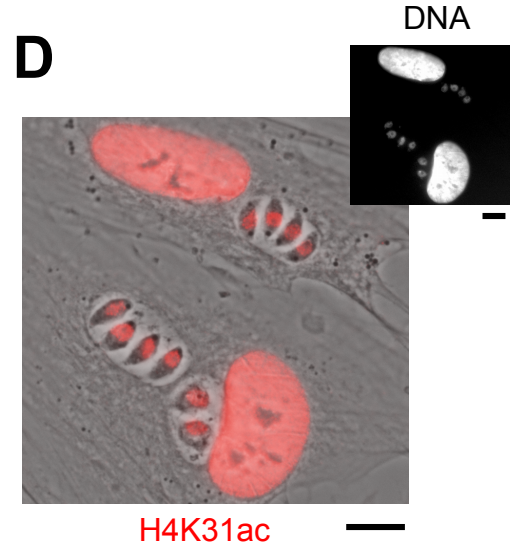
Figure 12 – figure supplement 1 H4K31me1 marks pericentromeric chromatin in *Plasmodium falciparum*. Zoomed-in views of centromeric and peri-centromeric chromatin from *P. falciparum* chromosomes.

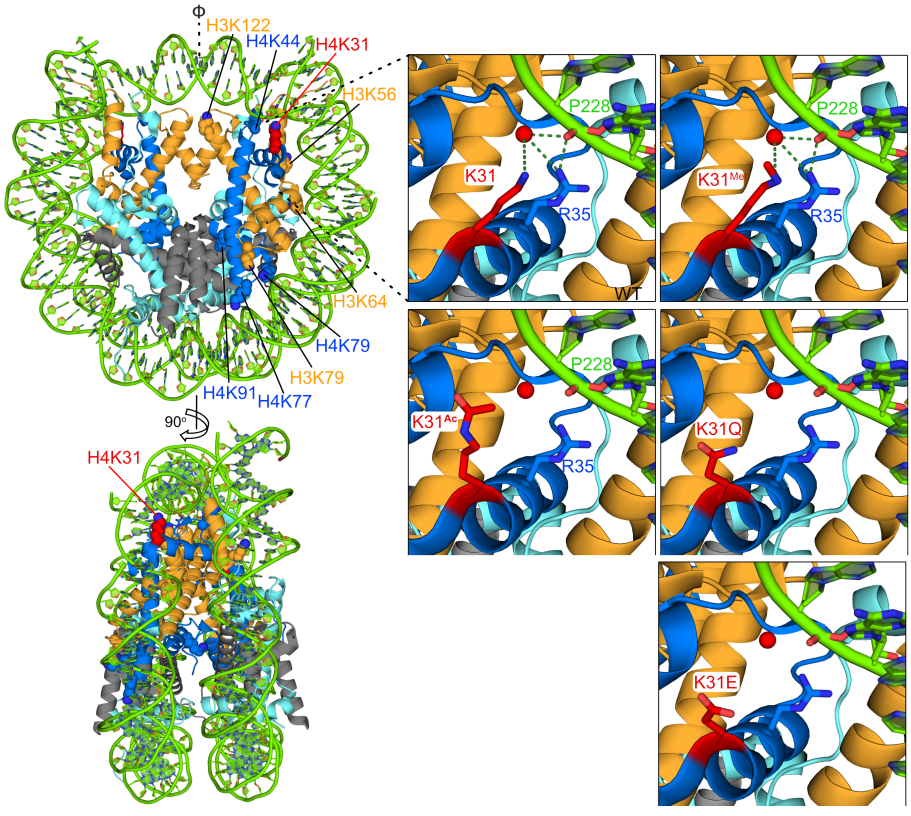
Figure 12 – figure supplement 2 H4K31me1 marks pericentromeric chromatin in *Plasmodium falciparum*. Zoomed-in views of centromeric and peri-centromeric chromatin from *P. falciparum* chromosomes.

Figure 9-source data 1 Table corresponding to Figure 9D with normalized expression distribution of *T. gondii* genes in tachyzoite stage subdivided into four ranges of expression (cluster Q1 to Q4).


A
D N I Q G I T K P A I R
% of base peak
ion current
m/z


B
N-terminal tail
α1
Toxoplasma SGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGG
Plasmodium SGRGKGGKGLGKGGAKRHRKILRDNIQGITKPAIRRLARRGG
S.cerevisae SGRGKGGKGLGKGGAKRHRKILRDNIQGITKPAIRRLARRGG
H.sapiens SGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGG
A.thaliana SGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGG
5 8 12 16 20 31


D
DNA
H4K31ac


C
Φ
H3K122
H4K44
H4K31
H3K56
H3K64
H4K79
H3K79
H4K91
H4K77
90°
H4K31
P228
K31
R35
WT
K31Me
K31Ac
K31Q
K31E

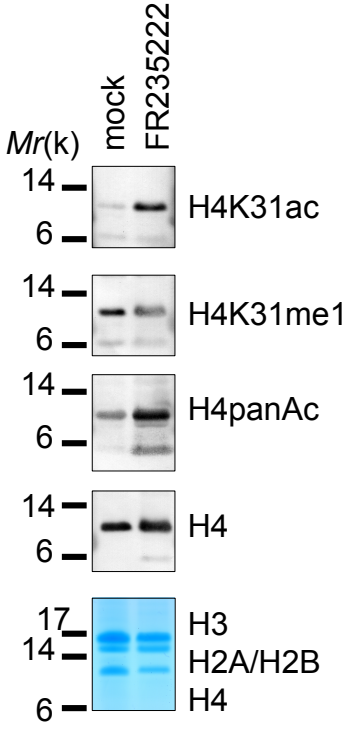

E
Mr(k)
mock
FR235222
14
6
H4K31ac
H4K31me1
H4panAc
H4
17
H3
H2A/H2B

Figure 1

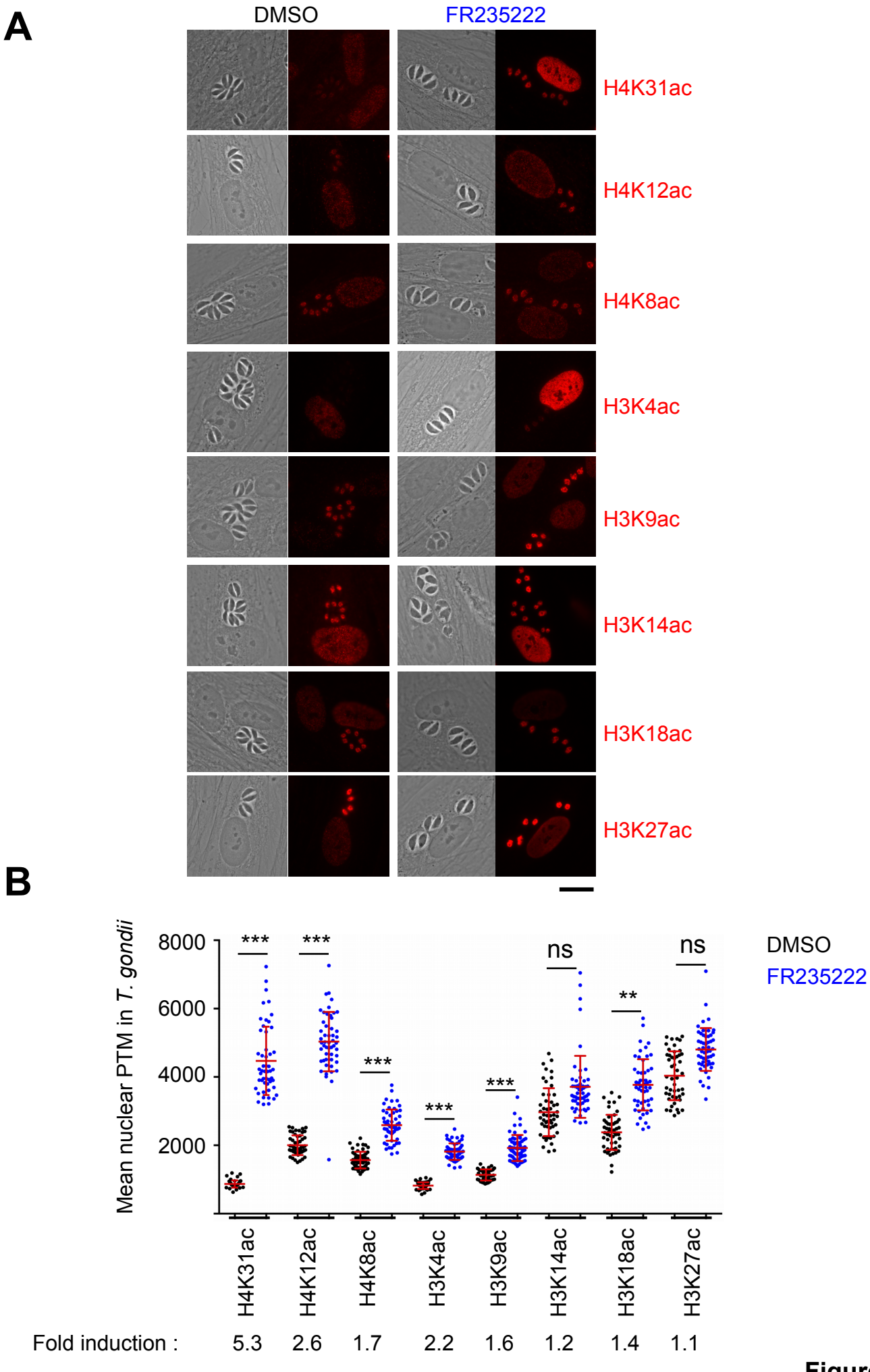
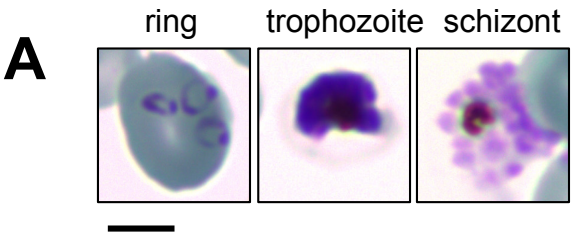
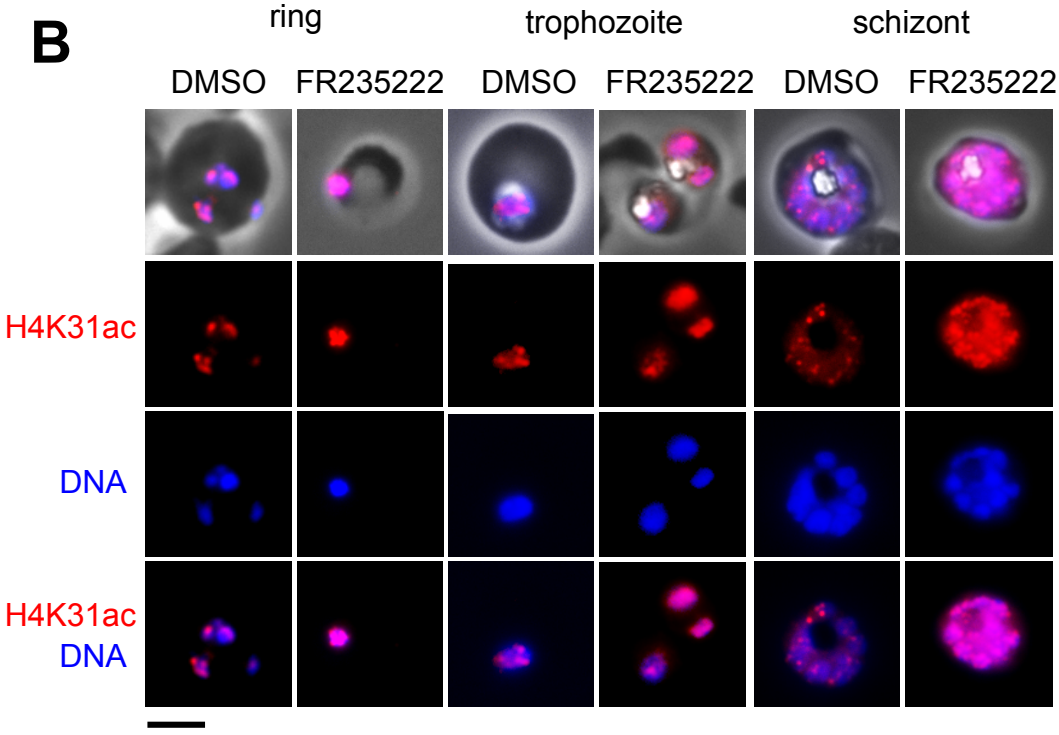
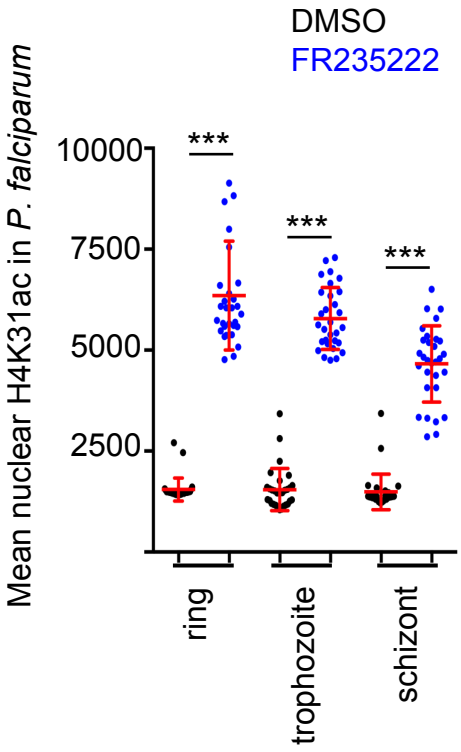
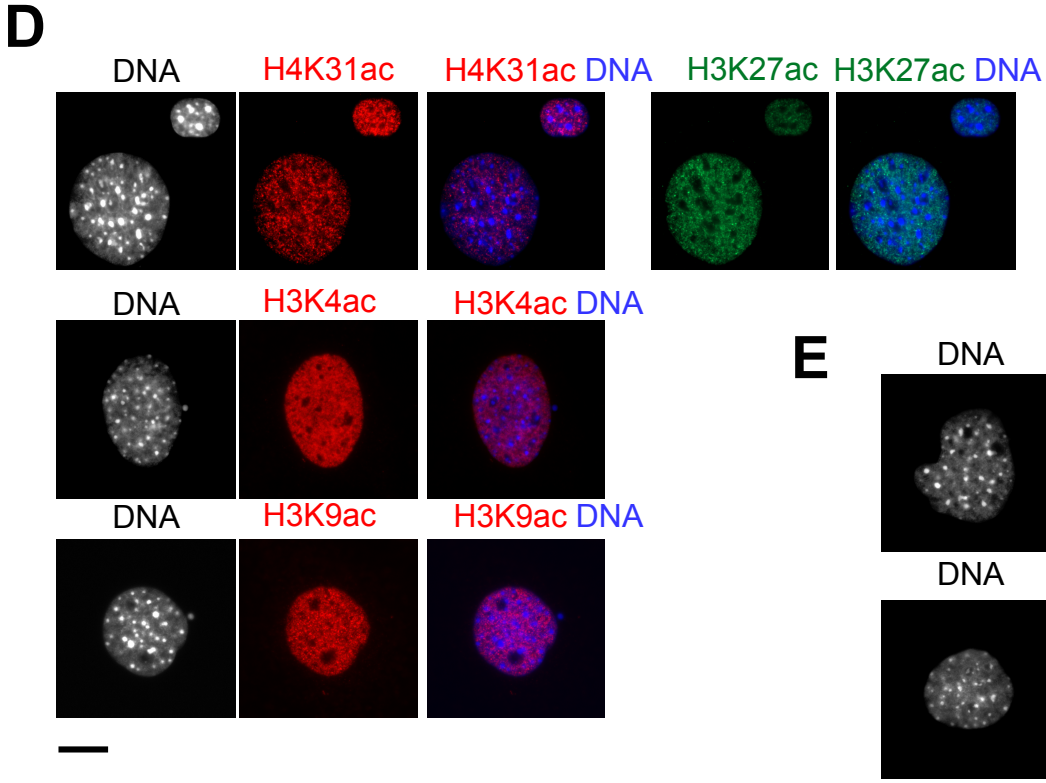

A
DMSO
FR235222
H4K31ac
H4K12ac
H4K8ac
H3K4ac
H3K9ac
H3K14ac
H3K18ac
H3K27ac
B
Mean nuclear PTM in T. gondii
8000
6000
4000
2000
***
***
***
***
***
ns
**
ns
DMSO
FR235222
H4K31ac
H4K12ac
H4K8ac
H3K4ac
H3K9ac
H3K14ac
H3K18ac
H3K27ac
Fold induction :
5.3
2.6
1.7
2.2
1.6
1.2
1.4
1.1

Figure 2


A
ring
trophozoite
schizont


B
ring
trophozoite
schizont
DMSO
FR235222
DMSO
FR235222
DMSO
FR235222
H4K31ac
DNA
H4K31ac
DNA


C
DMSO
FR235222
Mean nuclear H4K31ac in P. falciparum
10000
7500
5000
2500
***
***
***
ring
trophozoite
schizont


D
DNA
H4K31ac
H4K31ac DNA
H3K27ac
H3K27ac DNA
DNA
H3K4ac
H3K4ac DNA
DNA
H3K9ac
H3K9ac DNA

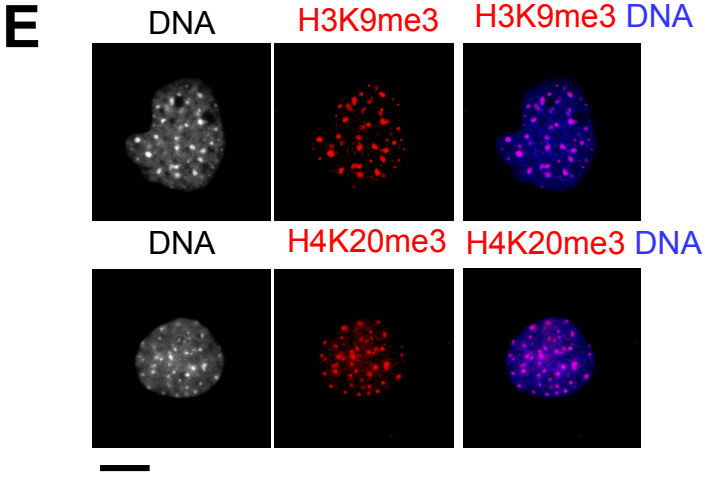

E
DNA
H3K9me3
H3K9me3 DNA
DNA
H4K20me3
H4K20me3 DNA

Figure 3

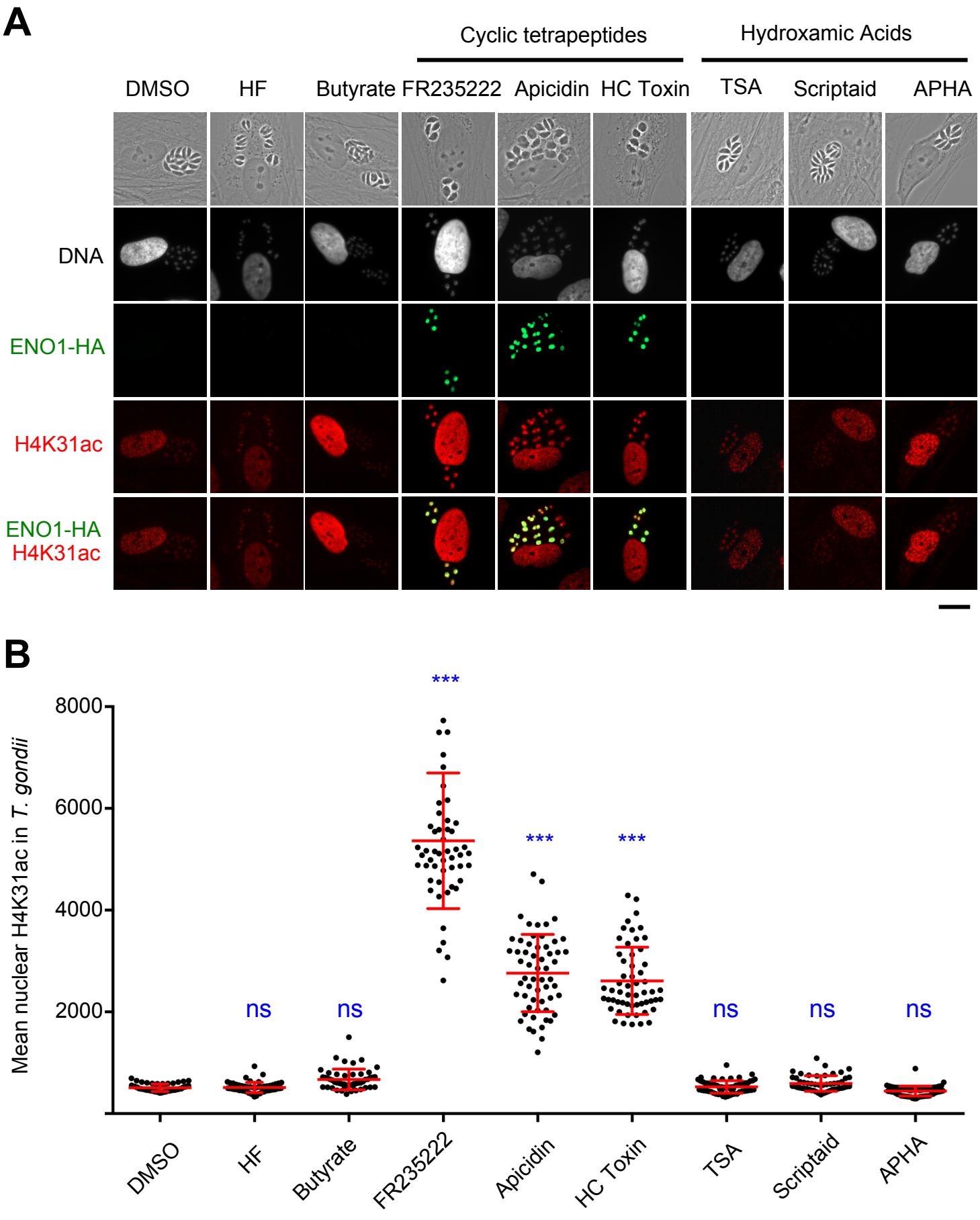

A
Cyclic tetrapeptides
Hydroxamic Acids
DMSO
HF
Butyrate
FR235222
Apicidin
HC Toxin
TSA
Scriptaid
APHA
DNA
ENO1-HA
H4K31ac
ENO1-HA
H4K31ac
B
Mean nuclear H4K31ac in T. gondii
8000
6000
4000
2000
***
***
***
ns
ns
ns
ns
ns
DMSO
HF
Butyrate
FR235222
Apicidin
HC Toxin
TSA
Scriptaid
APHA

Figure 4

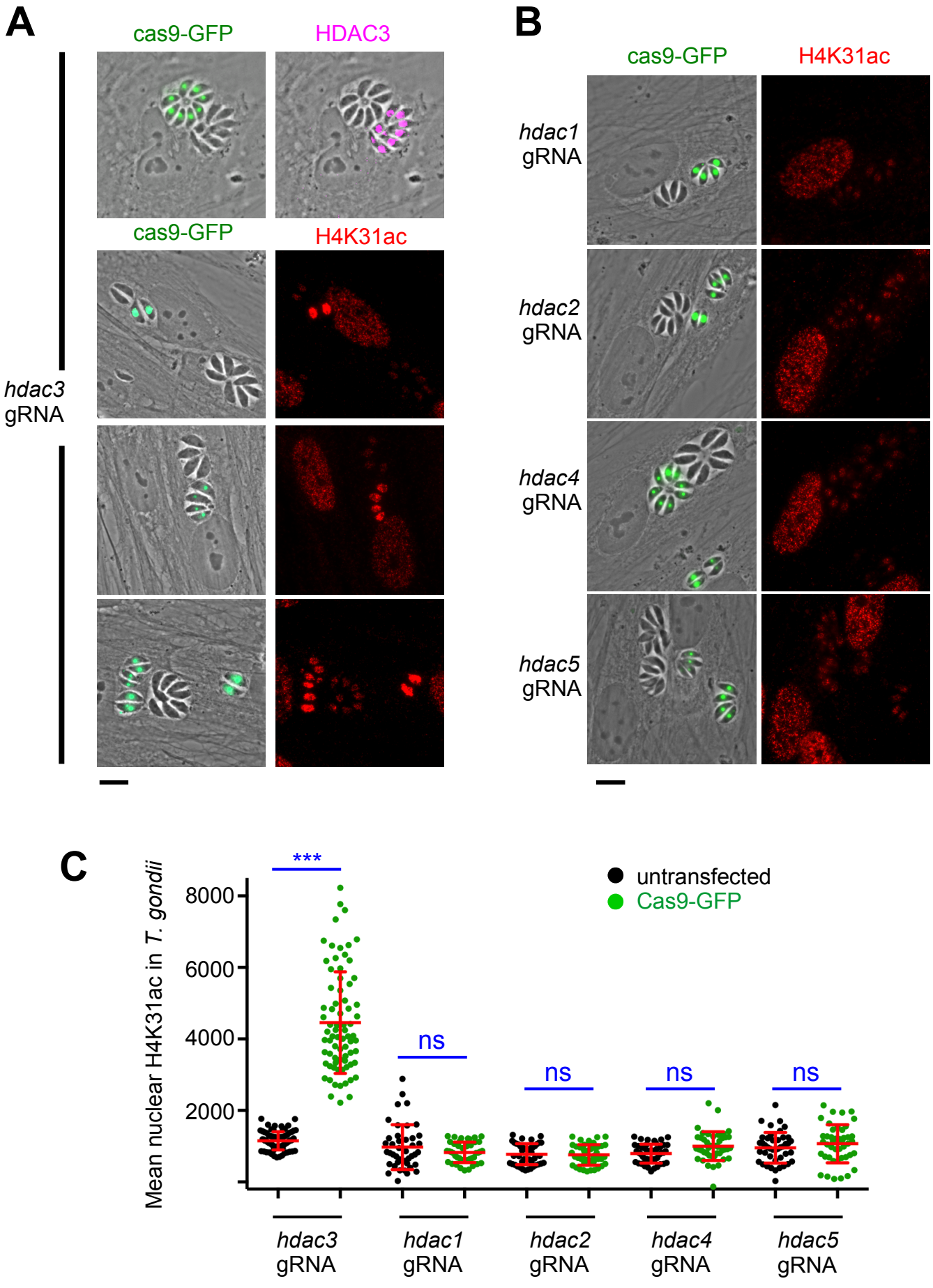

A
cas9-GFP
HDAC3
cas9-GFP
H4K31ac
hdac3
gRNA
B
cas9-GFP
H4K31ac
hdac1
gRNA
hdac2
gRNA
hdac4
gRNA
hdac5
gRNA
C
Mean nuclear H4K31ac in T. gondii
8000
6000
4000
2000
***
ns
ns
ns
ns
untransfected
Cas9-GFP
hdac3
gRNA
hdac1
gRNA
hdac2
gRNA
hdac4
gRNA
hdac5
gRNA

Figure 5

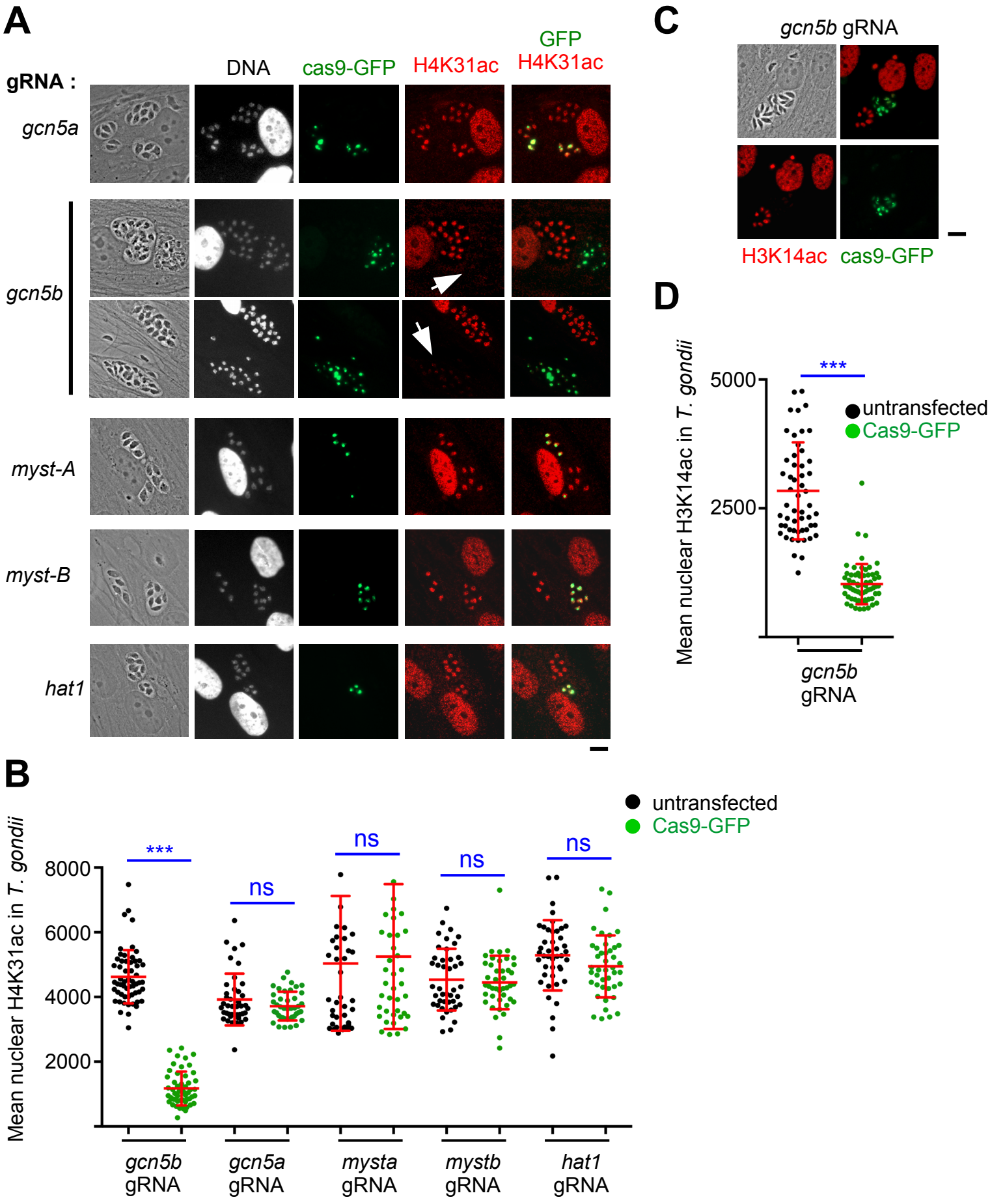

A
gRNA :
DNA
cas9-GFP
H4K31ac
GFP H4K31ac
gcn5a
gcn5b
myst-A
myst-B
hat1
C
gcn5b gRNA
H3K14ac
cas9-GFP
D
Mean nuclear H3K14ac in T. gondii
5000
2500
***
untransfected
Cas9-GFP
gcn5b gRNA
B
Mean nuclear H4K31ac in T. gondii
8000
6000
4000
2000
***
ns
ns
ns
ns
untransfected
Cas9-GFP
gcn5b gRNA
gcn5a gRNA
mysta gRNA
mystb gRNA
hat1 gRNA

Figure 6

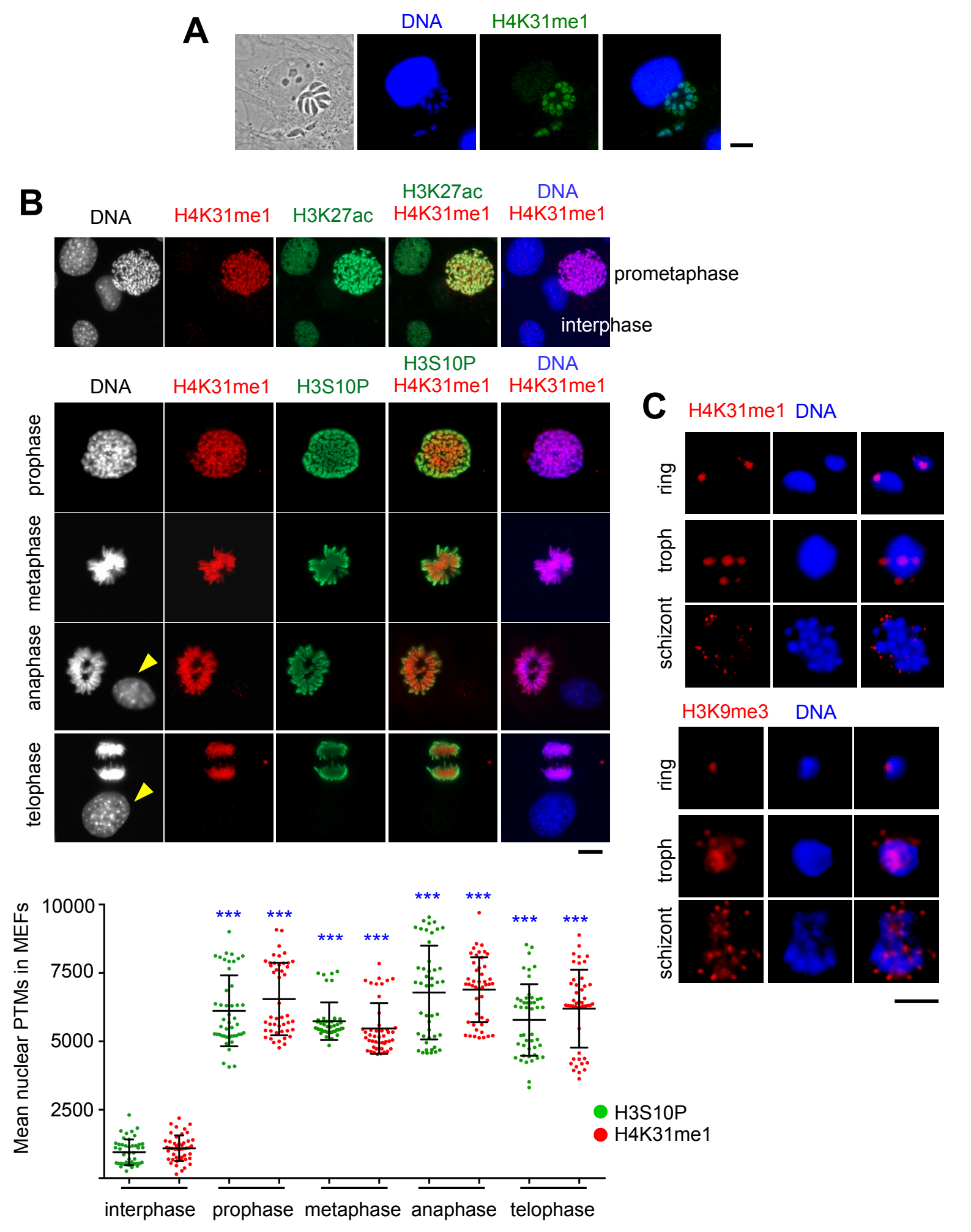

A
DNA
H4K31me1
B
DNA
H4K31me1
H3K27ac
H3K27ac
H4K31me1
DNA
H4K31me1
prometaphase
interphase
DNA
H4K31me1
H3S10P
H3S10P
H4K31me1
DNA
H4K31me1
prophase
metaphase
anaphase
telophase
C
H4K31me1
DNA
ring
troph
schizont
H3K9me3
DNA
ring
troph
schizont
Mean nuclear PTMs in MEFs
10000
7500
5000
2500
***
H3S10P
H4K31me1
interphase
prophase
metaphase
anaphase
telophase

Figure 7

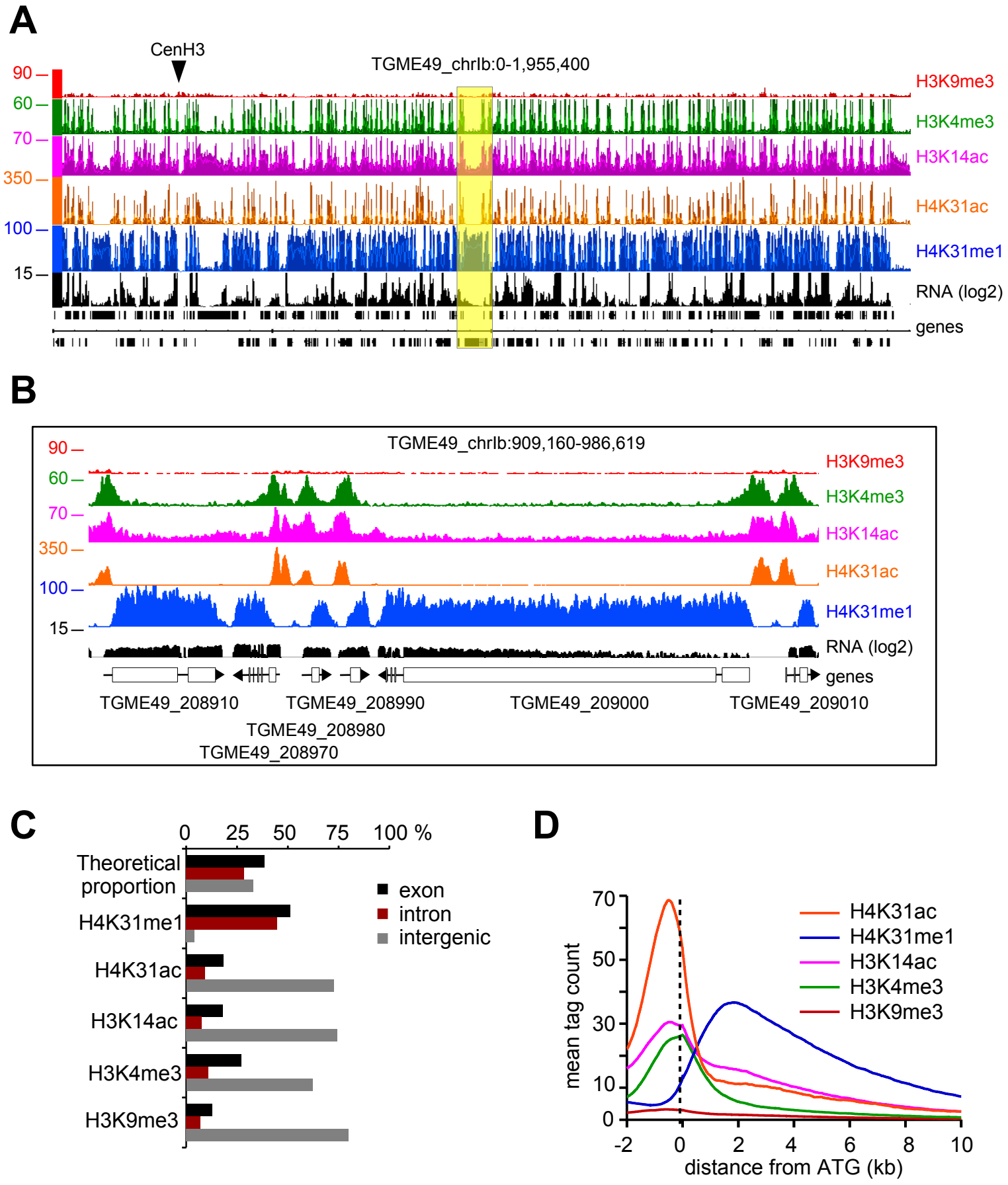

A
CenH3
TGME49_chrIb:0-1,955,400
90
60
70
350
100
15
H3K9me3
H3K4me3
H3K14ac
H4K31ac
H4K31me1
RNA (log2)
genes
B
TGME49_chrIb:909,160-986,619
90
60
70
350
100
15
H3K9me3
H3K4me3
H3K14ac
H4K31ac
H4K31me1
RNA (log2)
genes
TGME49_208910
TGME49_208990
TGME49_209000
TGME49_209010
TGME49_208980
TGME49_208970
C
0
25
50
75
100 %
Theoretical proportion
H4K31me1
H4K31ac
H3K14ac
H3K4me3
H3K9me3
exon
intron
intergenic
D
mean tag count
70
50
30
10
0
-2
0
2
4
6
8
10
distance from ATG (kb)
H4K31ac
H4K31me1
H3K14ac
H3K4me3
H3K9me3

Figure 8

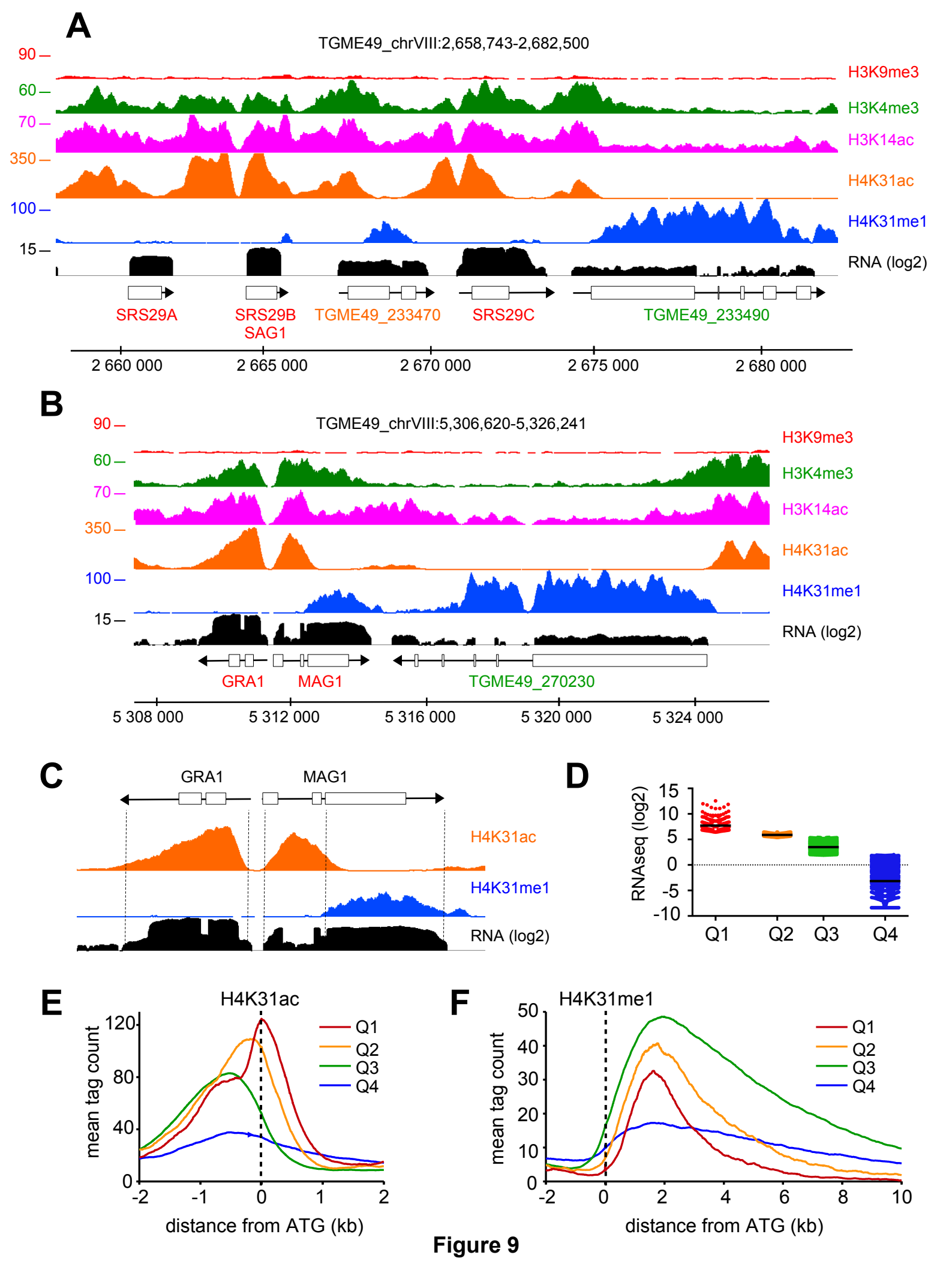

A
TGME49_chrVIII:2,658,743-2,682,500
90
60
70
350
100
15
H3K9me3
H3K4me3
H3K14ac
H4K31ac
H4K31me1
RNA (log2)
SRS29A
SRS29B
SAG1
TGME49_233470
SRS29C
TGME49_233490
2 660 000
2 665 000
2 670 000
2 675 000
2 680 000
B
TGME49_chrVIII:5,306,620-5,326,241
90
60
70
350
100
15
H3K9me3
H3K4me3
H3K14ac
H4K31ac
H4K31me1
RNA (log2)
GRA1
MAG1
TGME49_270230
5 308 000
5 312 000
5 316 000
5 320 000
5 324 000
C
GRA1
MAG1
H4K31ac
H4K31me1
RNA (log2)
D
RNAseq (log2)
15
10
5
0
-5
-10
Q1
Q2
Q3
Q4
E
H4K31ac
mean tag count
120
80
40
0
-2
-1
0
1
2
distance from ATG (kb)
Q1
Q2
Q3
Q4
F
H4K31me1
mean tag count
50
40
30
20
10
0
-2
0
2
4
6
8
10
distance from ATG (kb)
Q1
Q2
Q3
Q4

Figure 9

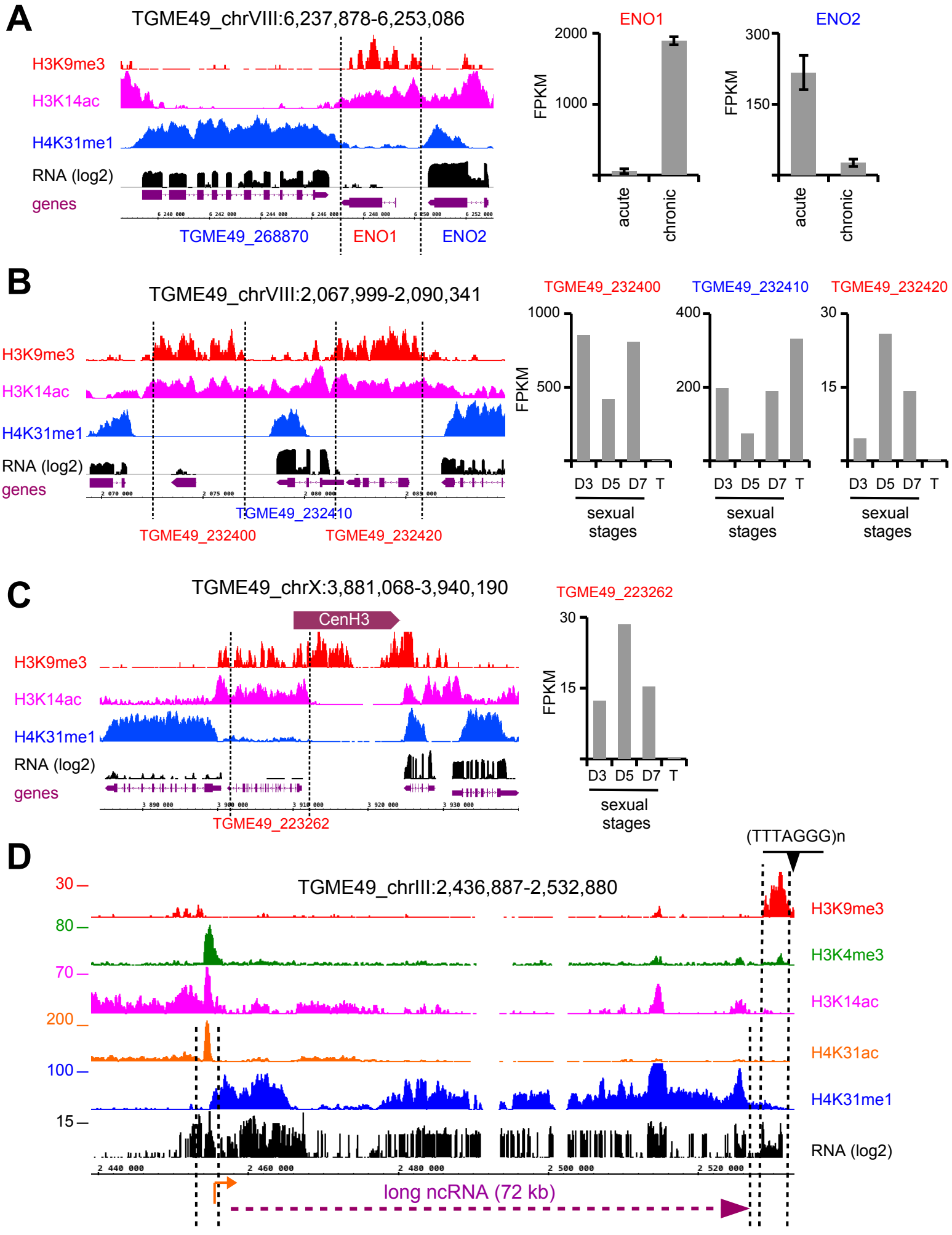

A
TGME49_chrVIII:6,237,878-6,253,086
H3K9me3
H3K14ac
H4K31me1
RNA (log2)
genes
TGME49_268870
ENO1
ENO2
ENO1
ENO2
FPKM
2000
1000
300
150
acute
chronic
B
TGME49_chrVIII:2,067,999-2,090,341
TGME49_232400
TGME49_232410
TGME49_232420
1000
500
400
200
30
15
D3 D5 D7 T
sexual stages
C
TGME49_chrX:3,881,068-3,940,190
CenH3
TGME49_223262
D
TGME49_chrIII:2,436,887-2,532,880
(TTTAGGG)n
H3K4me3
H4K31ac
long ncRNA (72 kb)

Figure 10

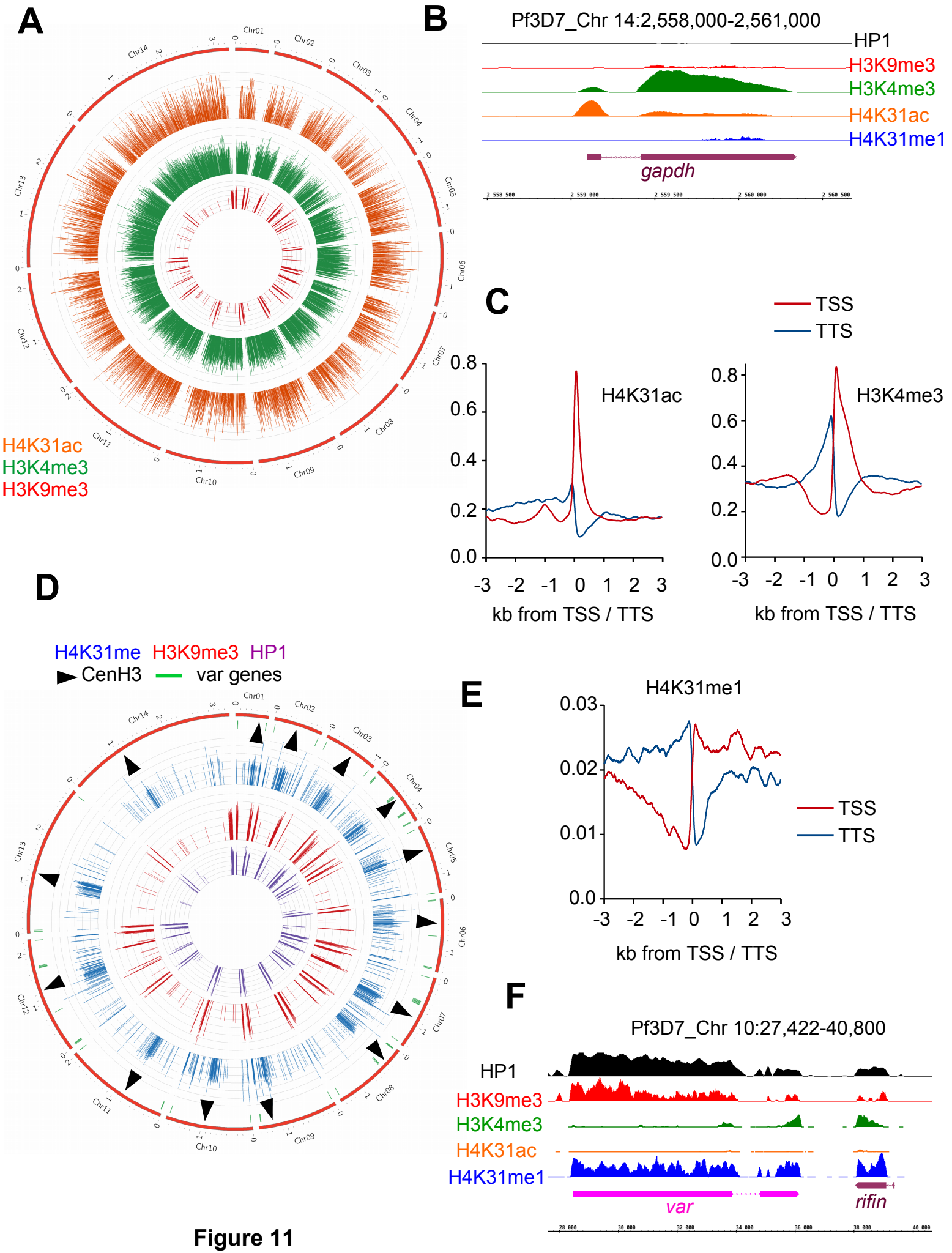

A
Chr01
Chr02
Chr03
Chr04
Chr05
Chr06
Chr07
Chr08
Chr09
Chr10
Chr11
Chr12
Chr13
Chr14
H4K31ac
H3K4me3
H3K9me3
B
Pf3D7_Chr 14:2,558,000-2,561,000
HP1
H3K9me3
H3K4me3
H4K31ac
H4K31me1
gapdh
2 558 500
2 559 000
2 559 500
2 560 000
2 560 500
C
TSS
TTS
H4K31ac
H3K4me3
0.8
0.6
0.4
0.2
0.0
-3 -2 -1 0 1 2 3
kb from TSS / TTS
D
H4K31me H3K9me3 HP1
CenH3
var genes
E
H4K31me1
0.03
0.02
0.01
0.0
TSS
TTS
F
Pf3D7_Chr 10:27,422-40,800
HP1
H3K9me3
H3K4me3
H4K31ac
H4K31me1
var
rifin
28 000
30 000
32 000
34 000
36 000
38 000
40 000

Figure 11

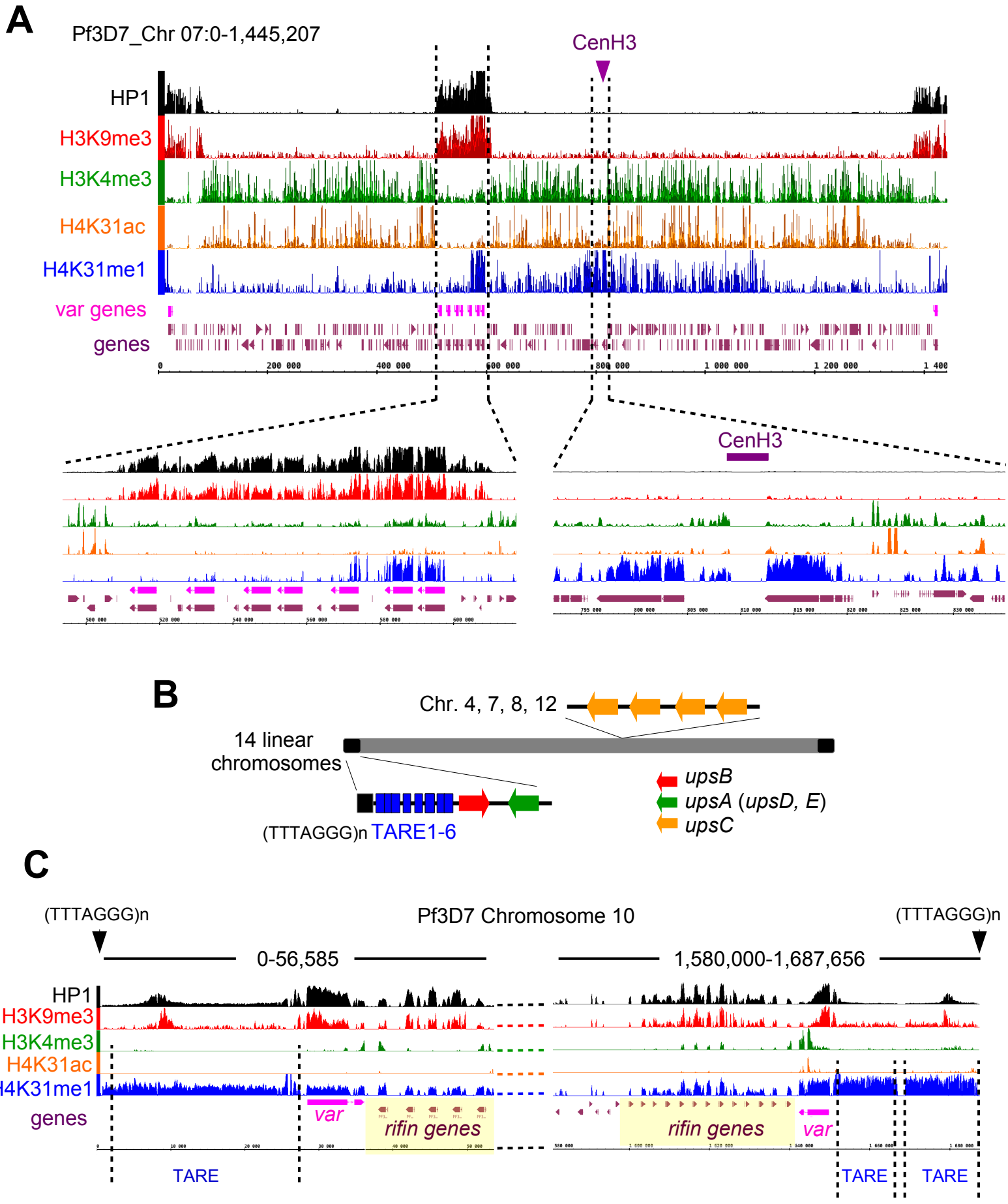

A
Pf3D7_Chr 07:0-1,445,207
CenH3
HP1
H3K9me3
H3K4me3
H4K31ac
H4K31me1
var genes
genes
0
200 000
400 000
600 000
800 000
1 000 000
1 200 000
1 400
CenH3
B
Chr. 4, 7, 8, 12
14 linear chromosomes
(TTTAGGG)n TARE1-6
upsB
upsA (upsD, E)
upsC
C
(TTTAGGG)n
Pf3D7 Chromosome 10
(TTTAGGG)n
0-56,585
1,580,000-1,687,656
HP1
H3K9me3
H3K4me3
H4K31ac
H4K31me1
genes
var
rifin genes
rifin genes
var
TARE
TARE
TARE

Figure 12

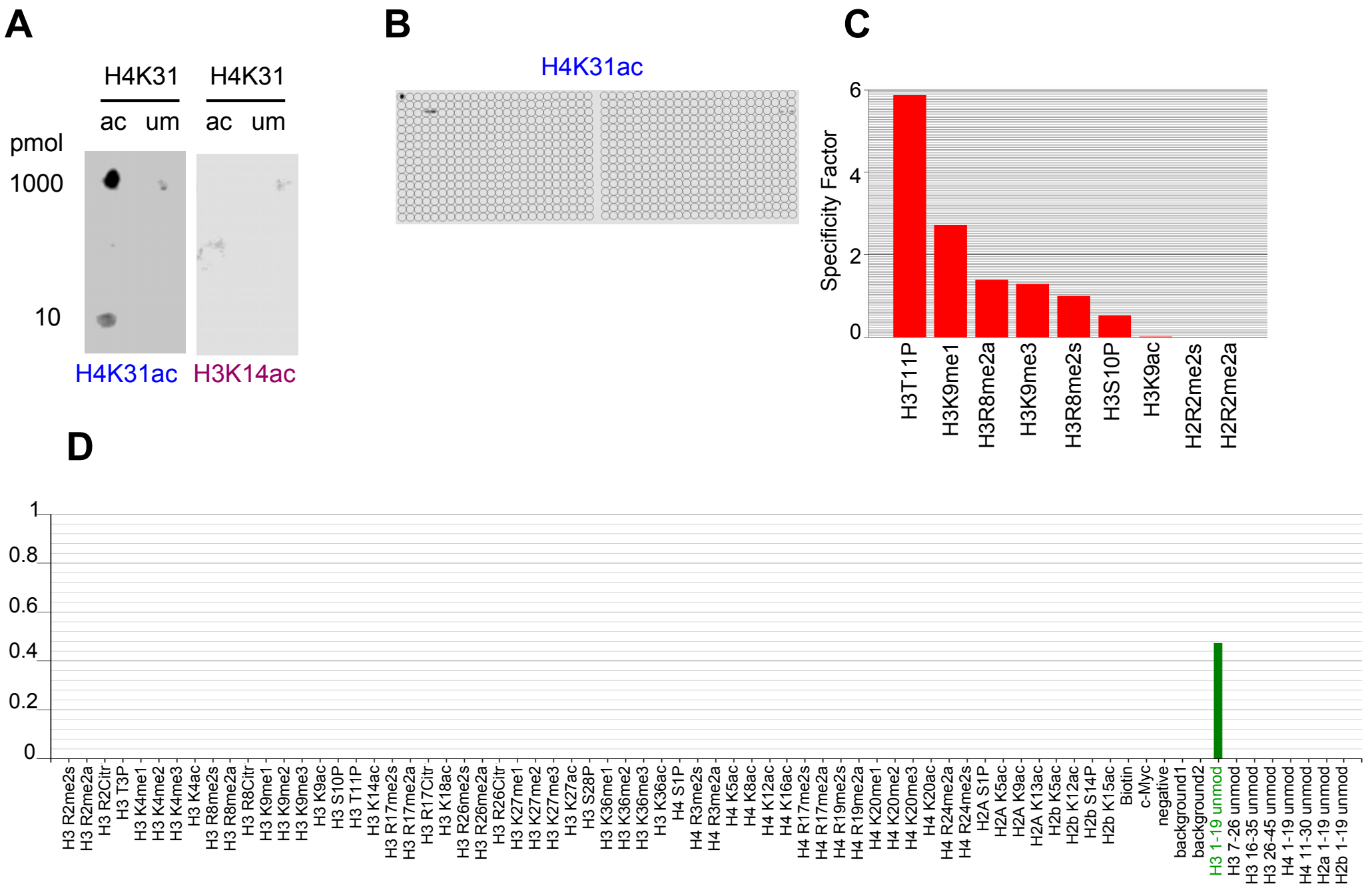


Figure 1 – figure supplement 1

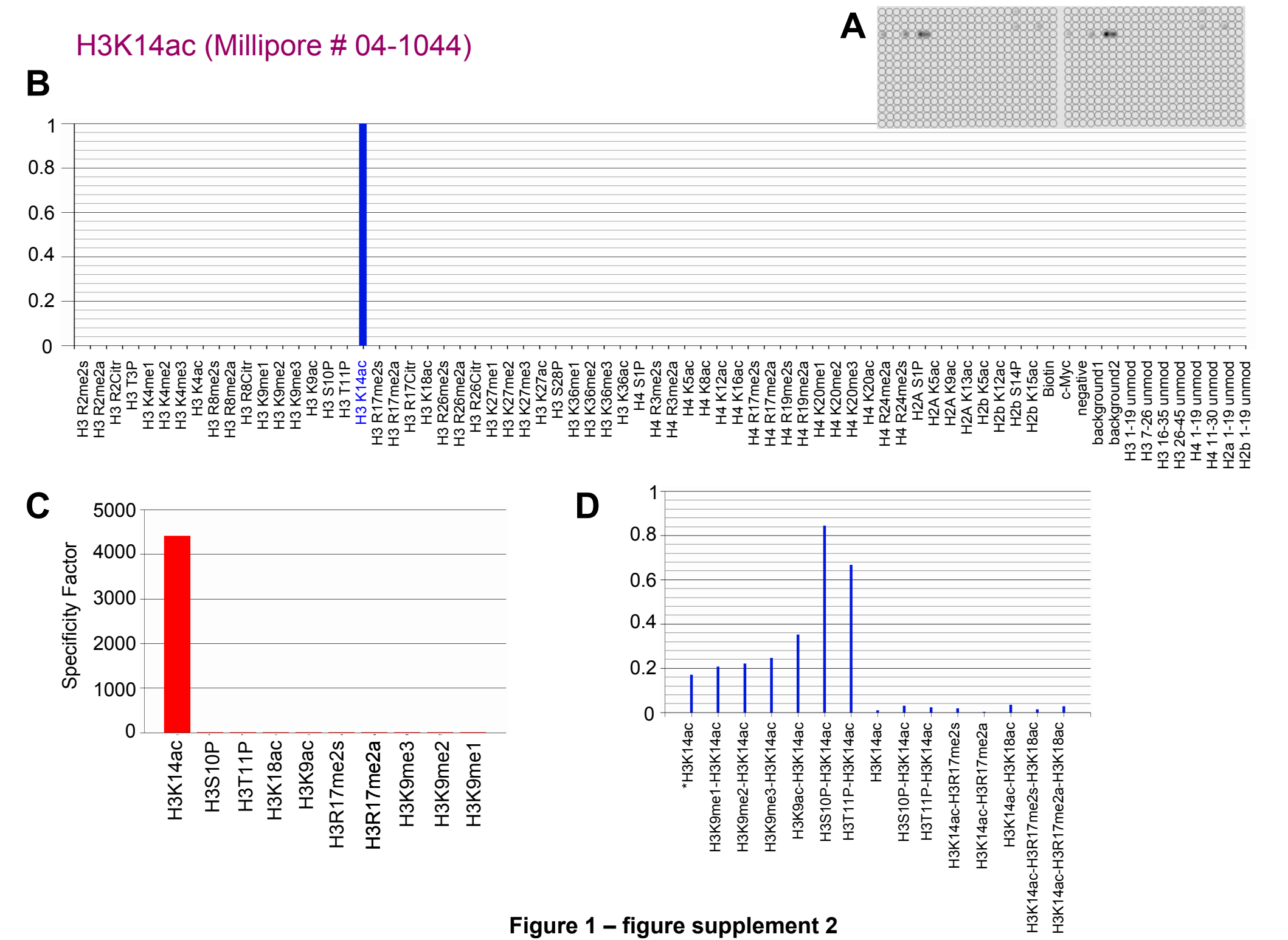


Figure 1 – figure supplement 2

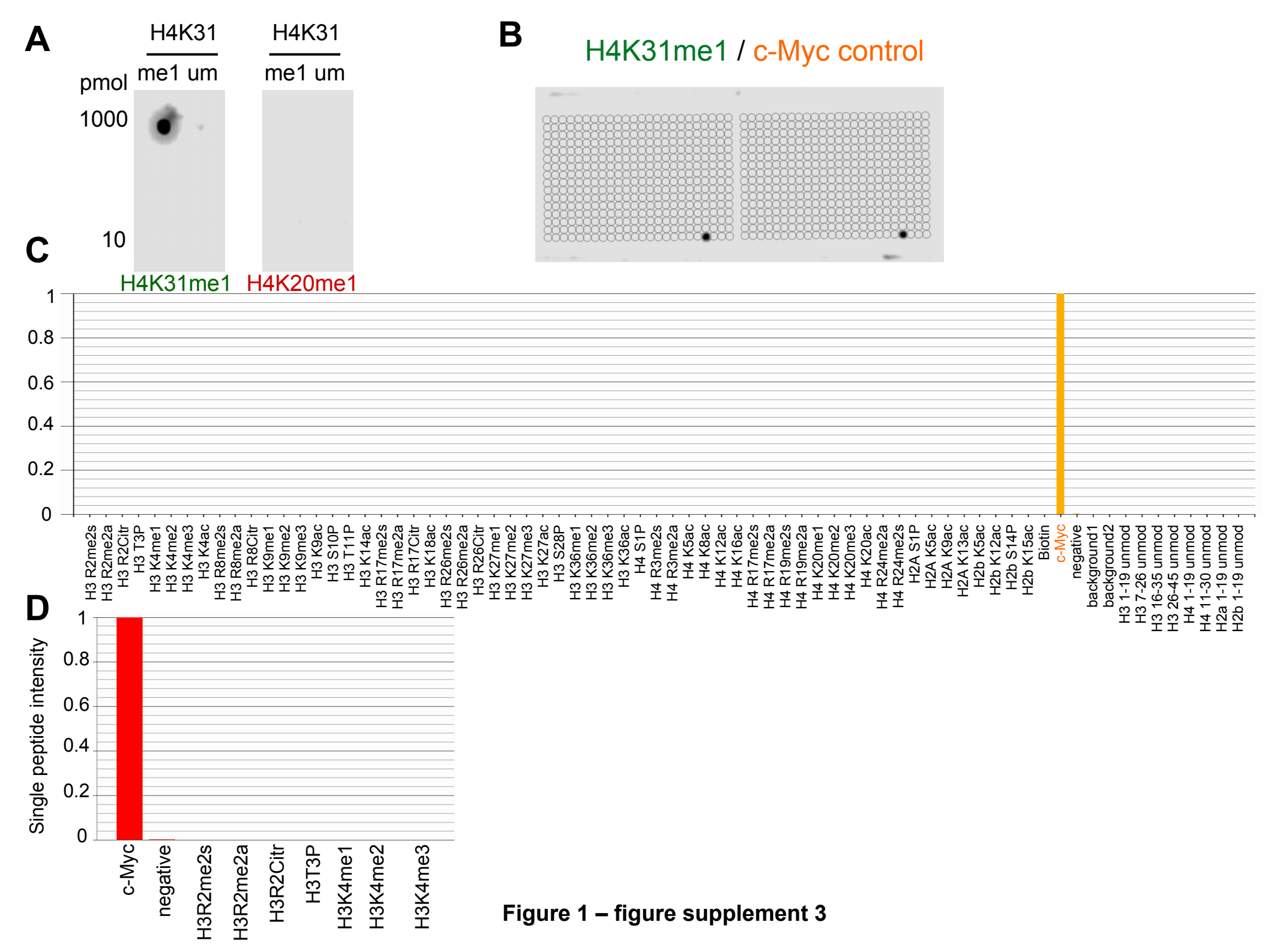
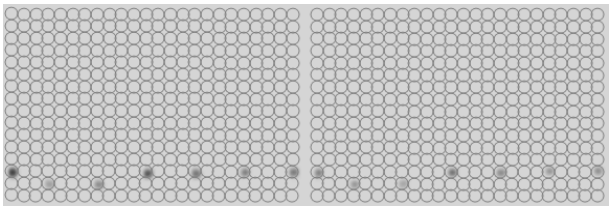
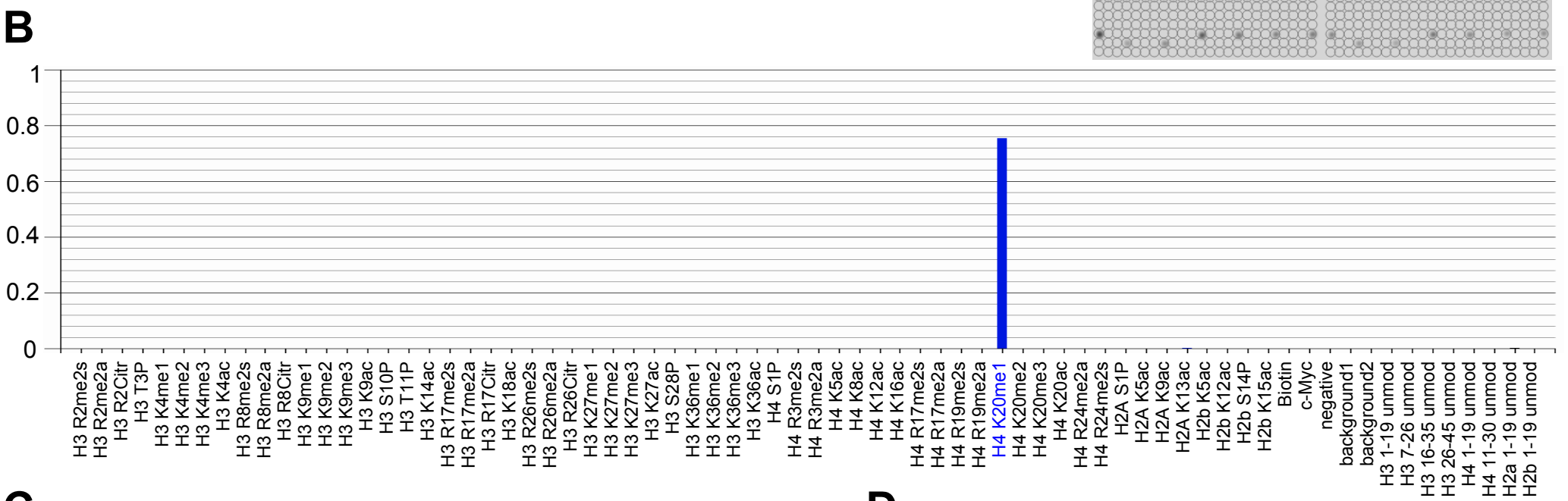
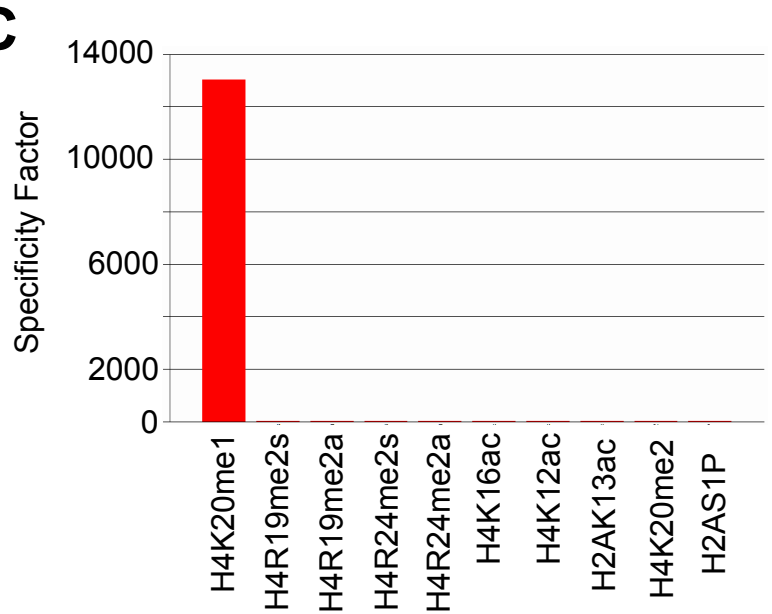


Figure 1 – figure supplement 3

H4K20me1 (Sautel *et al.*, 2007)







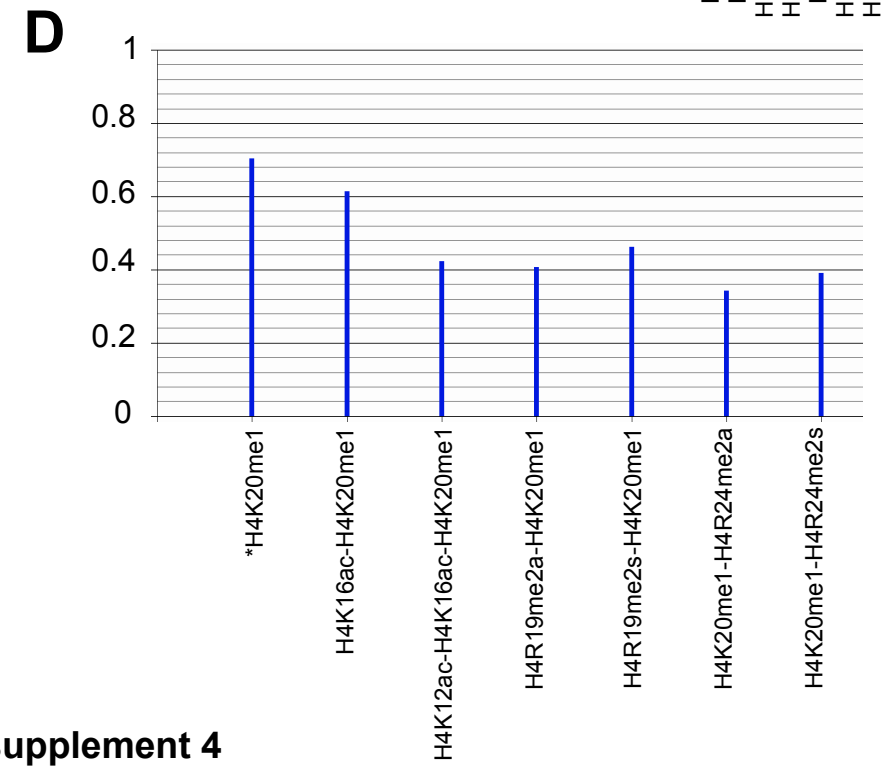


Figure 1 – figure supplement 4

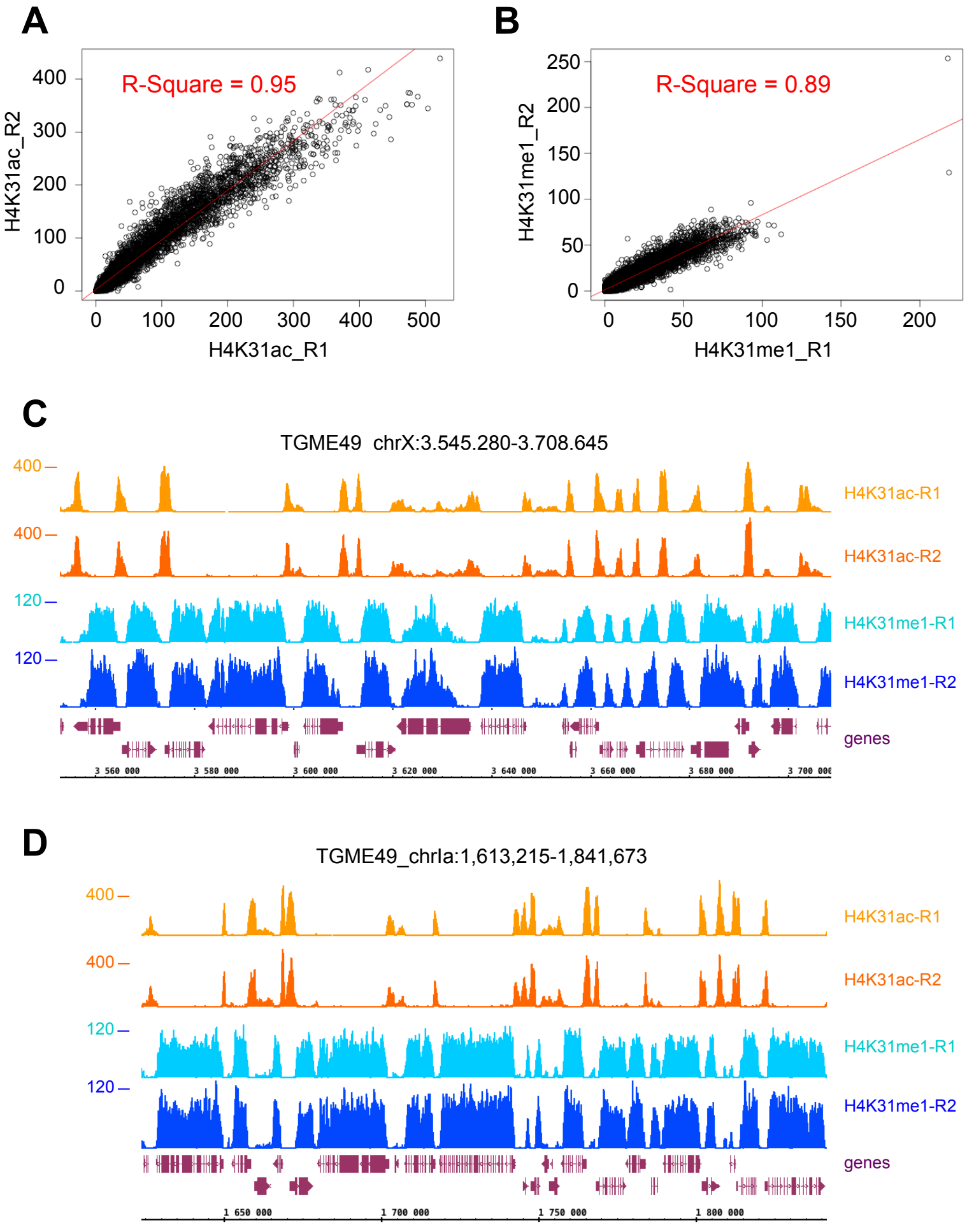


Figure 8 – figure supplement 1

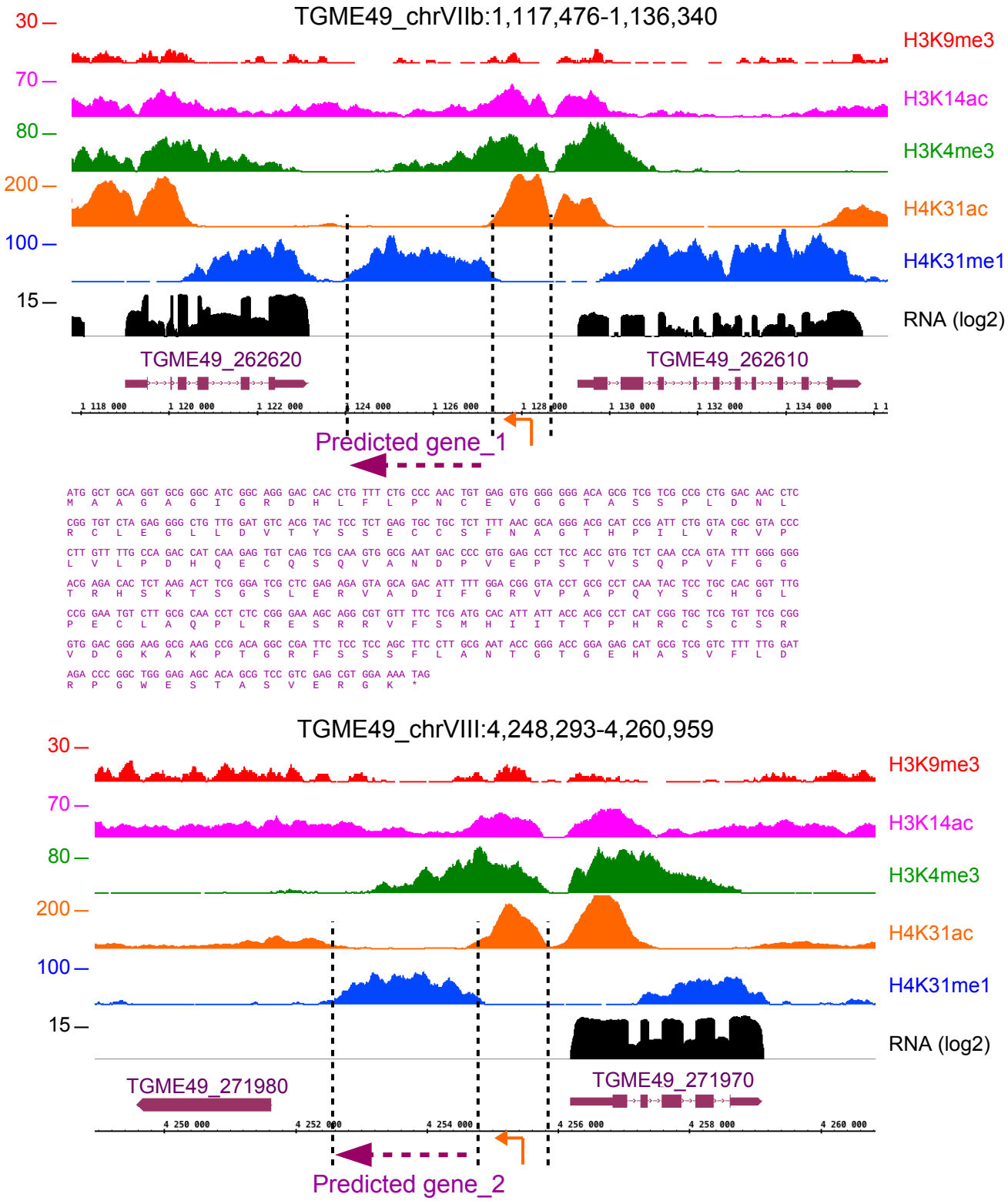


Figure 8 – figure supplement 2

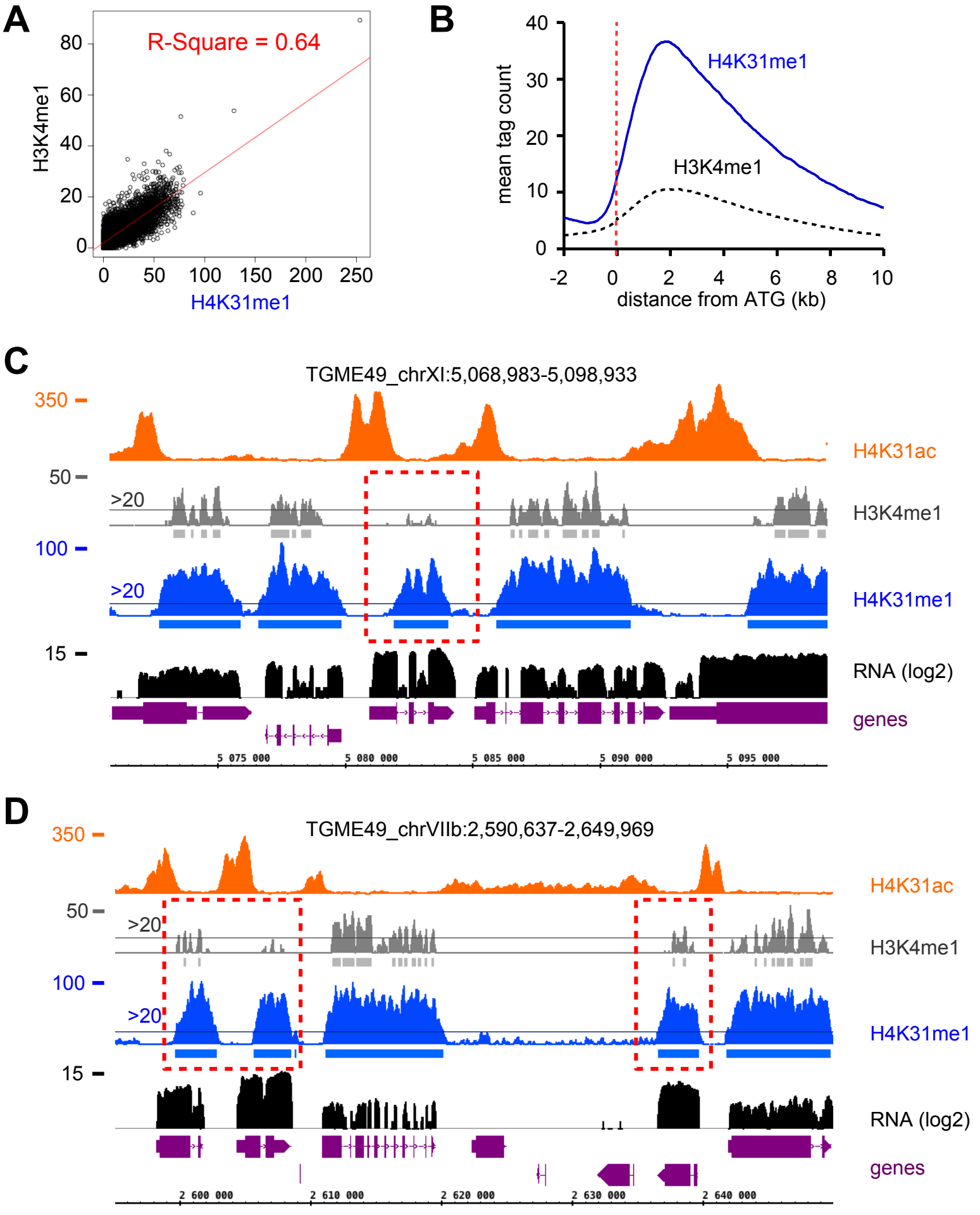


Figure 8 – figure supplement 3

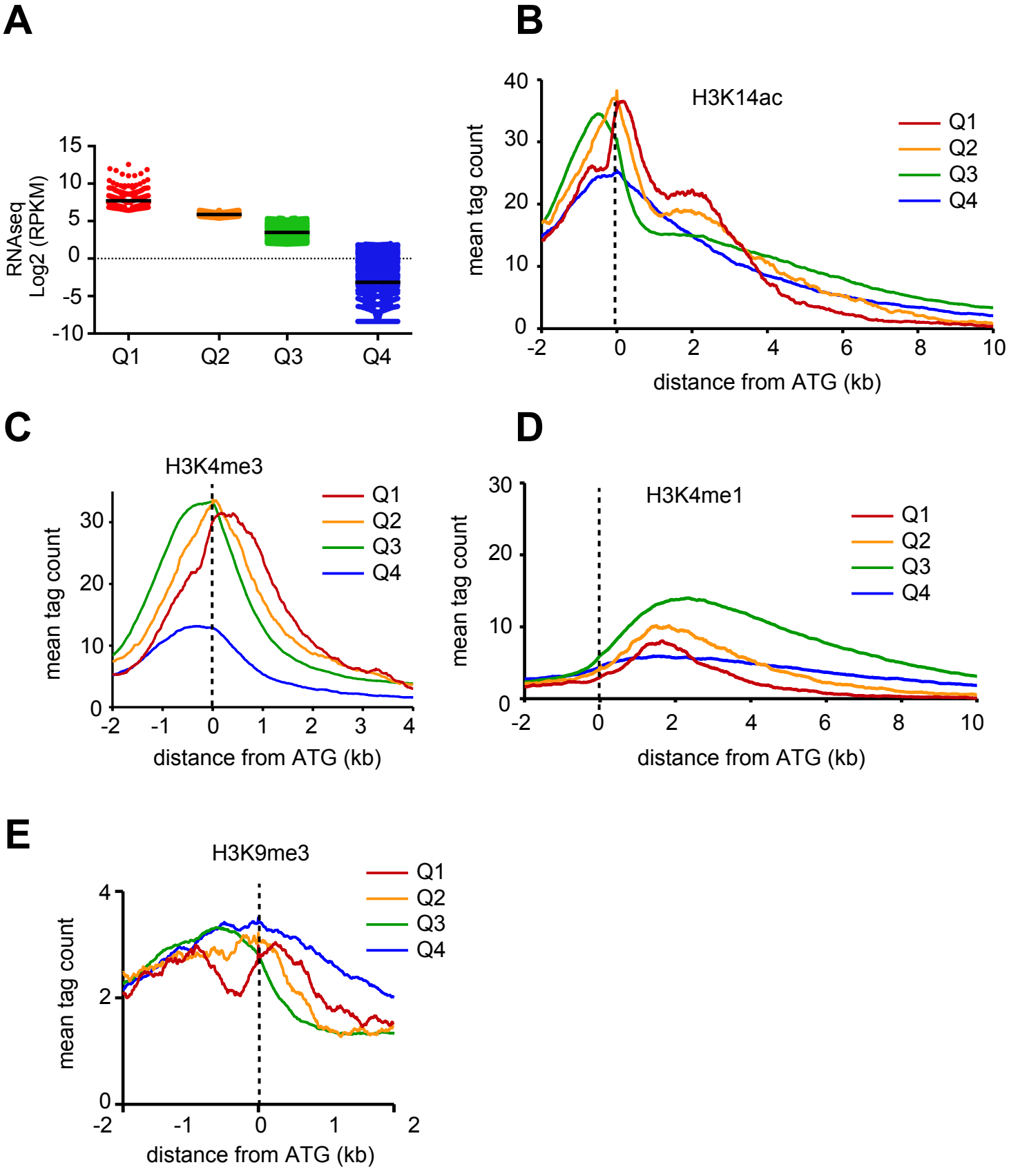


Figure 9 – figure supplement 1

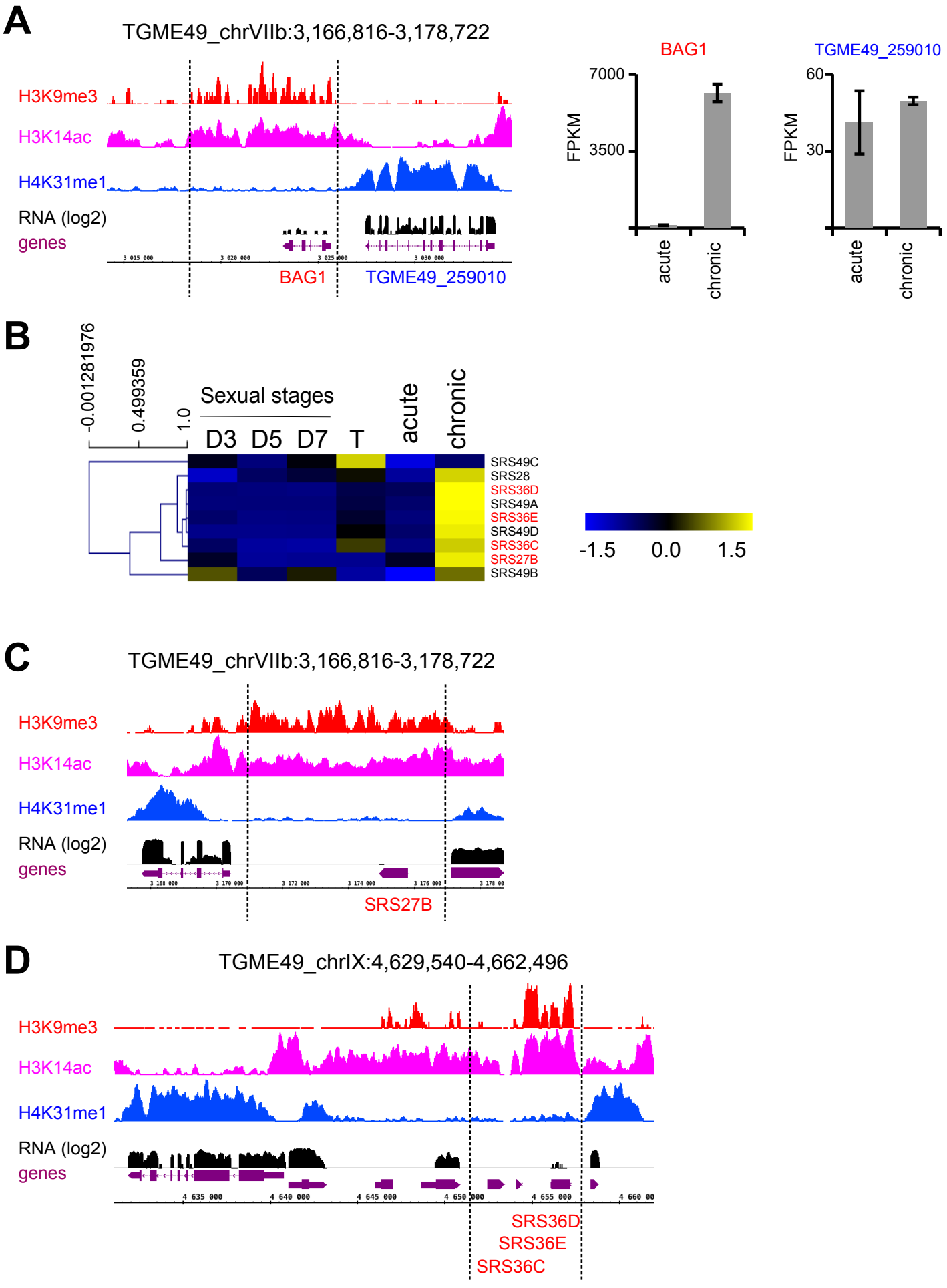


Figure 10 – figure supplement 1

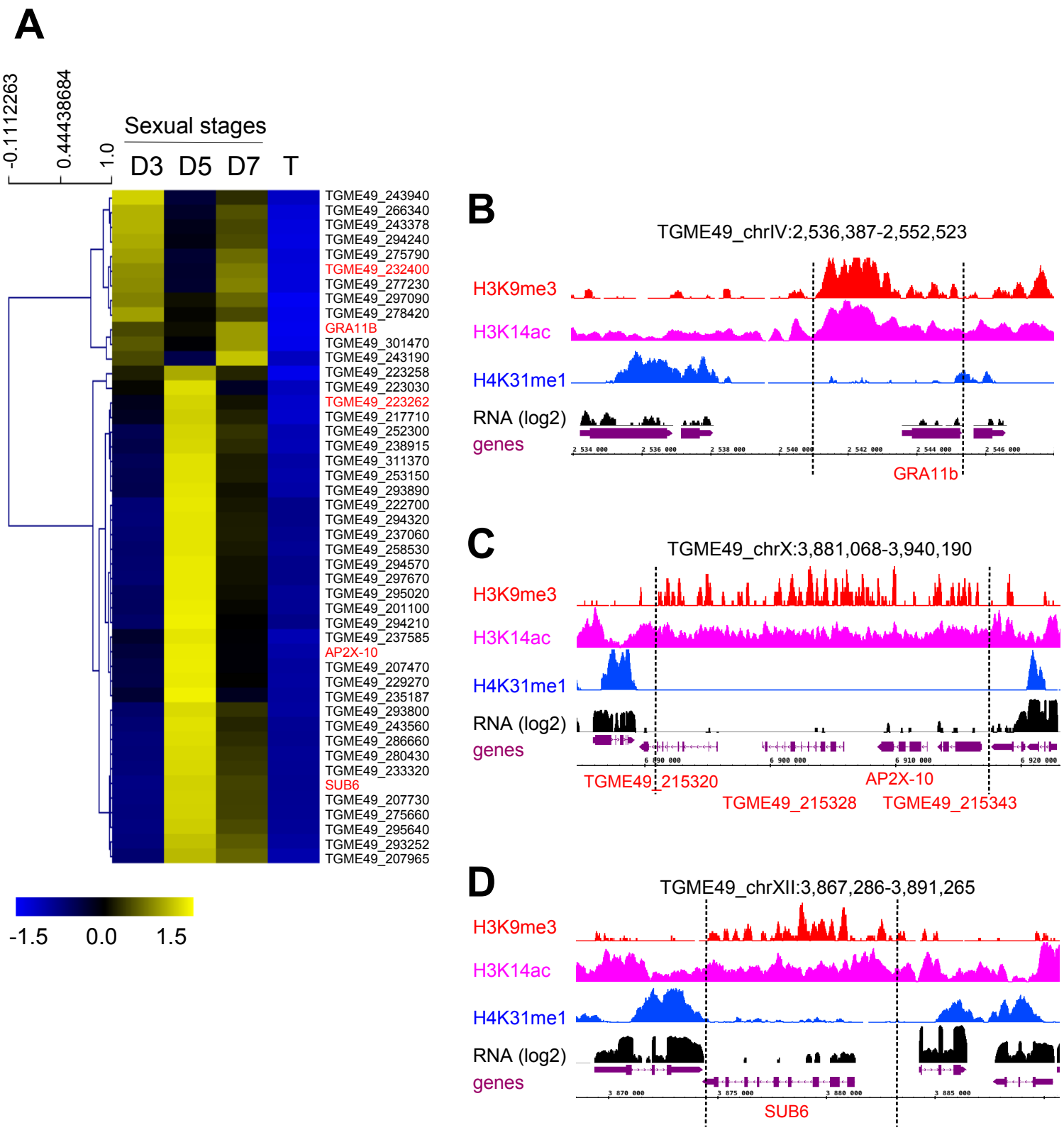


Figure 10 – figure supplement 2

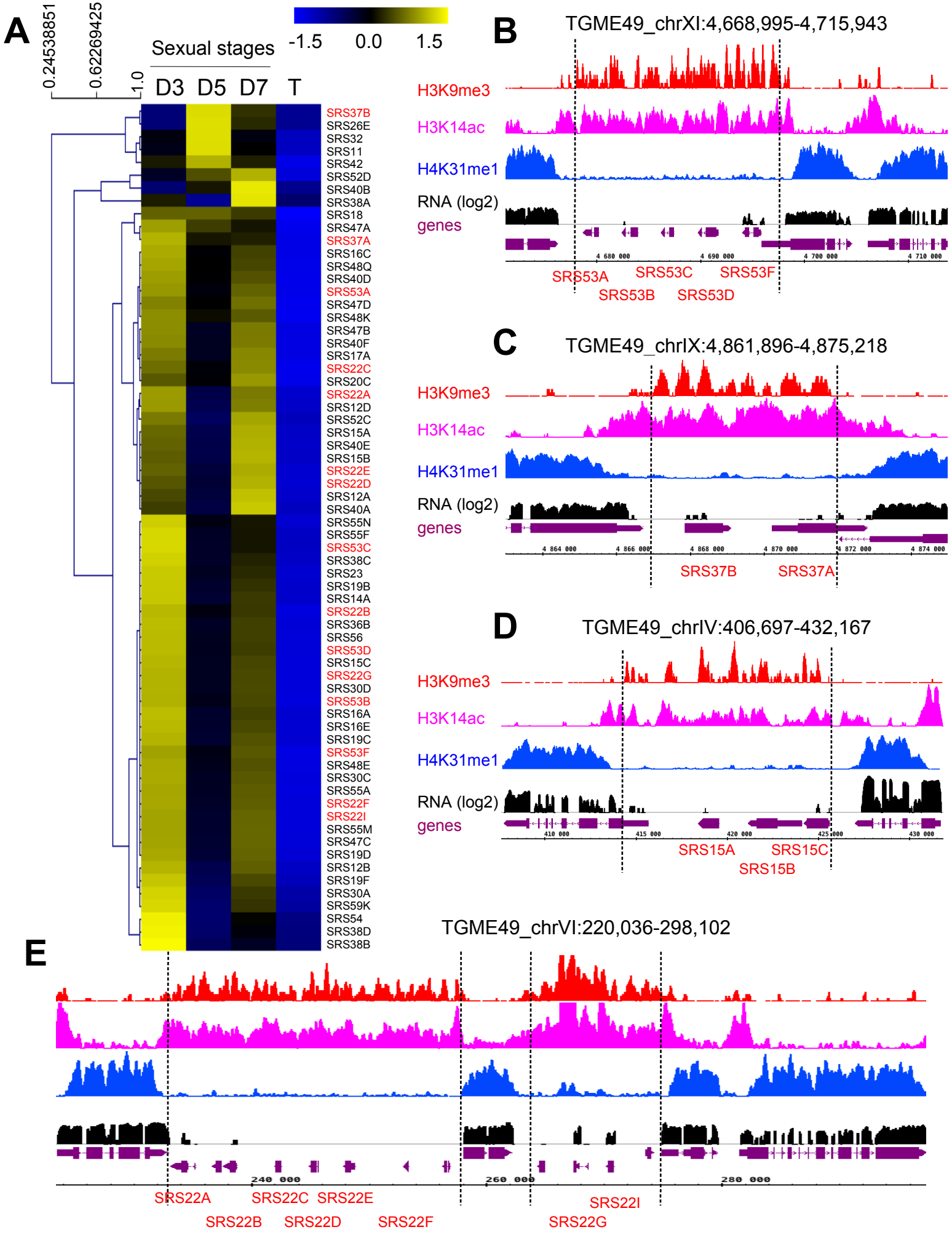


Figure 10 – figure supplement 3

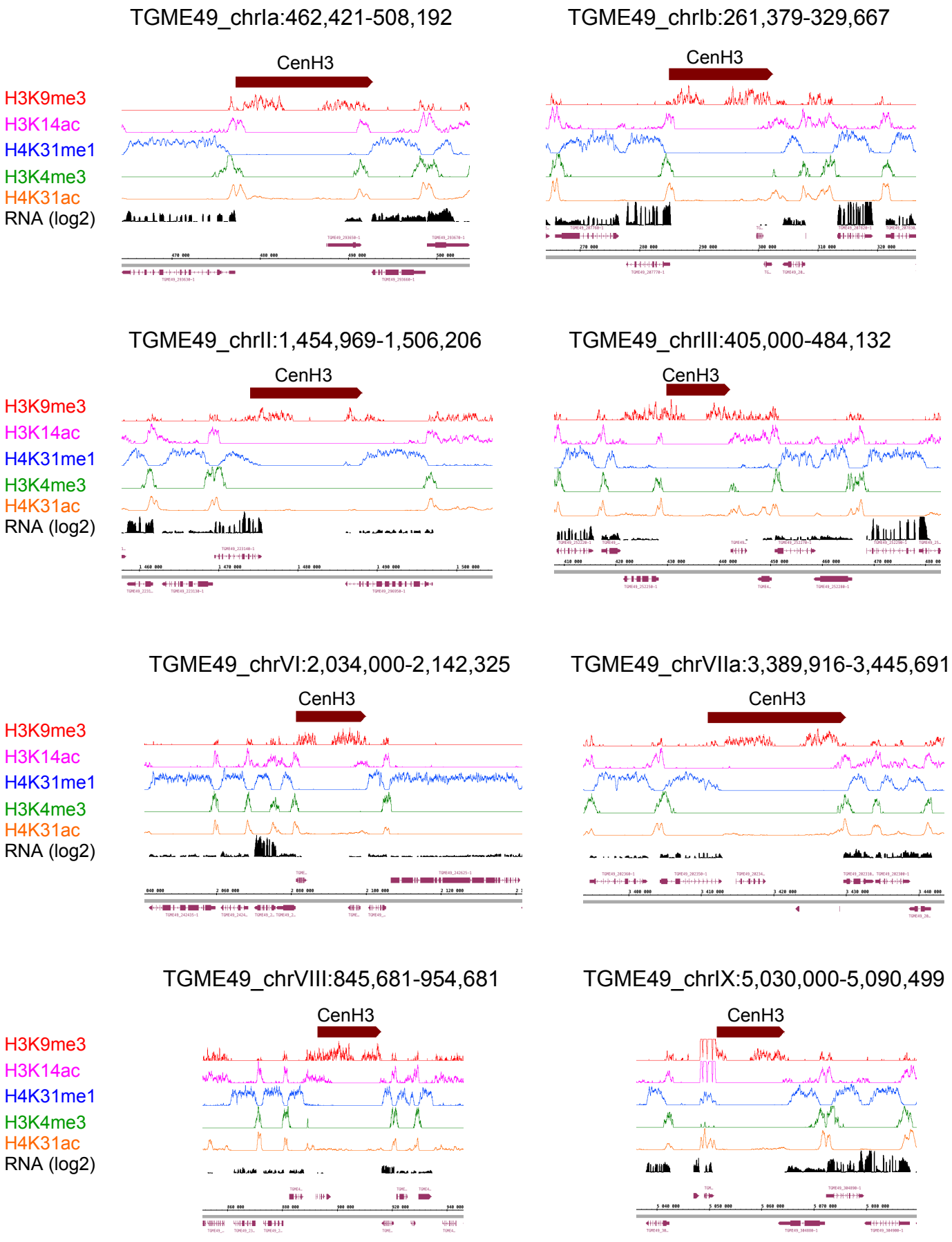


Figure 10 – figure supplement 4

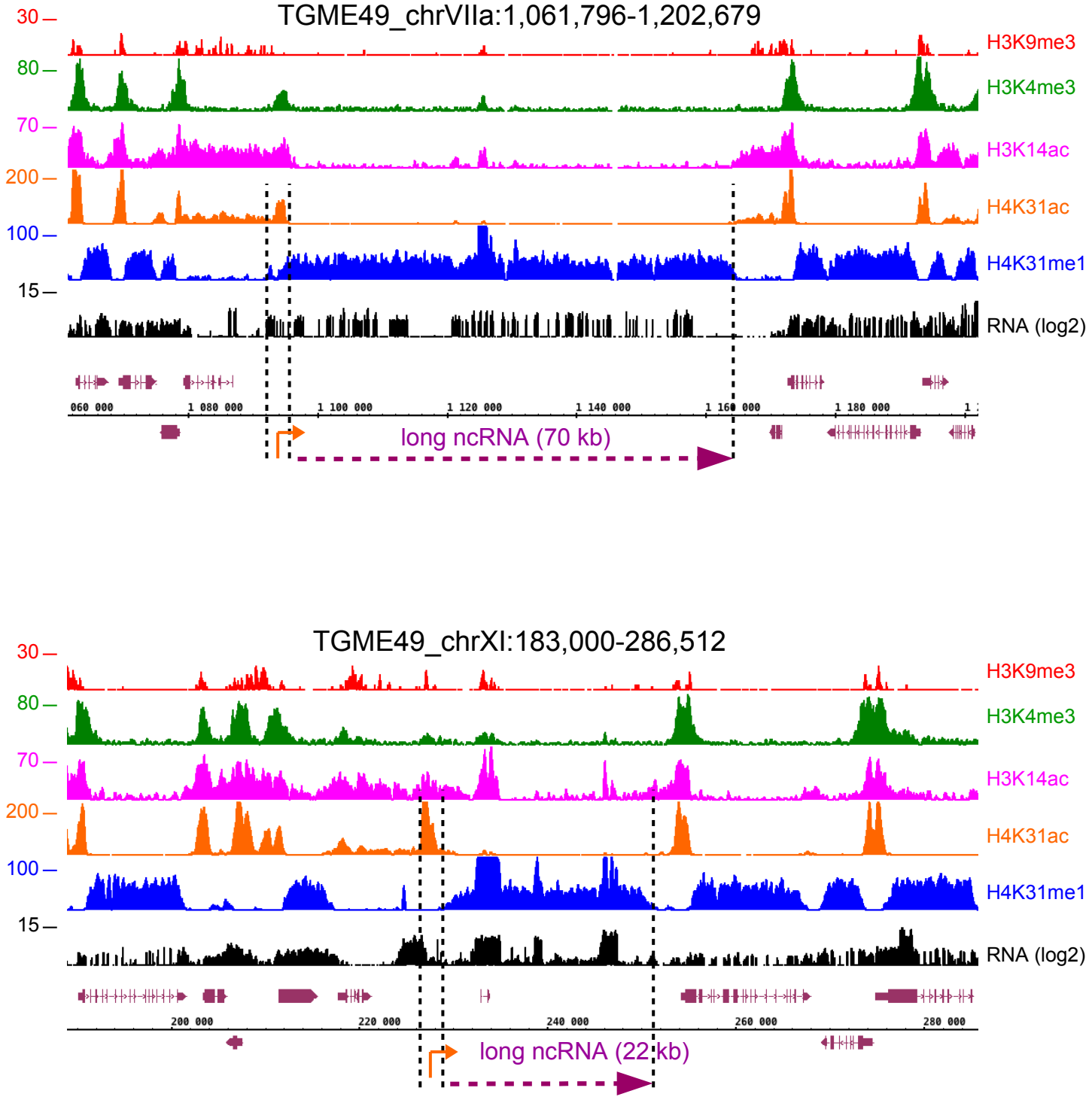


Figure 10 – figure supplement 5

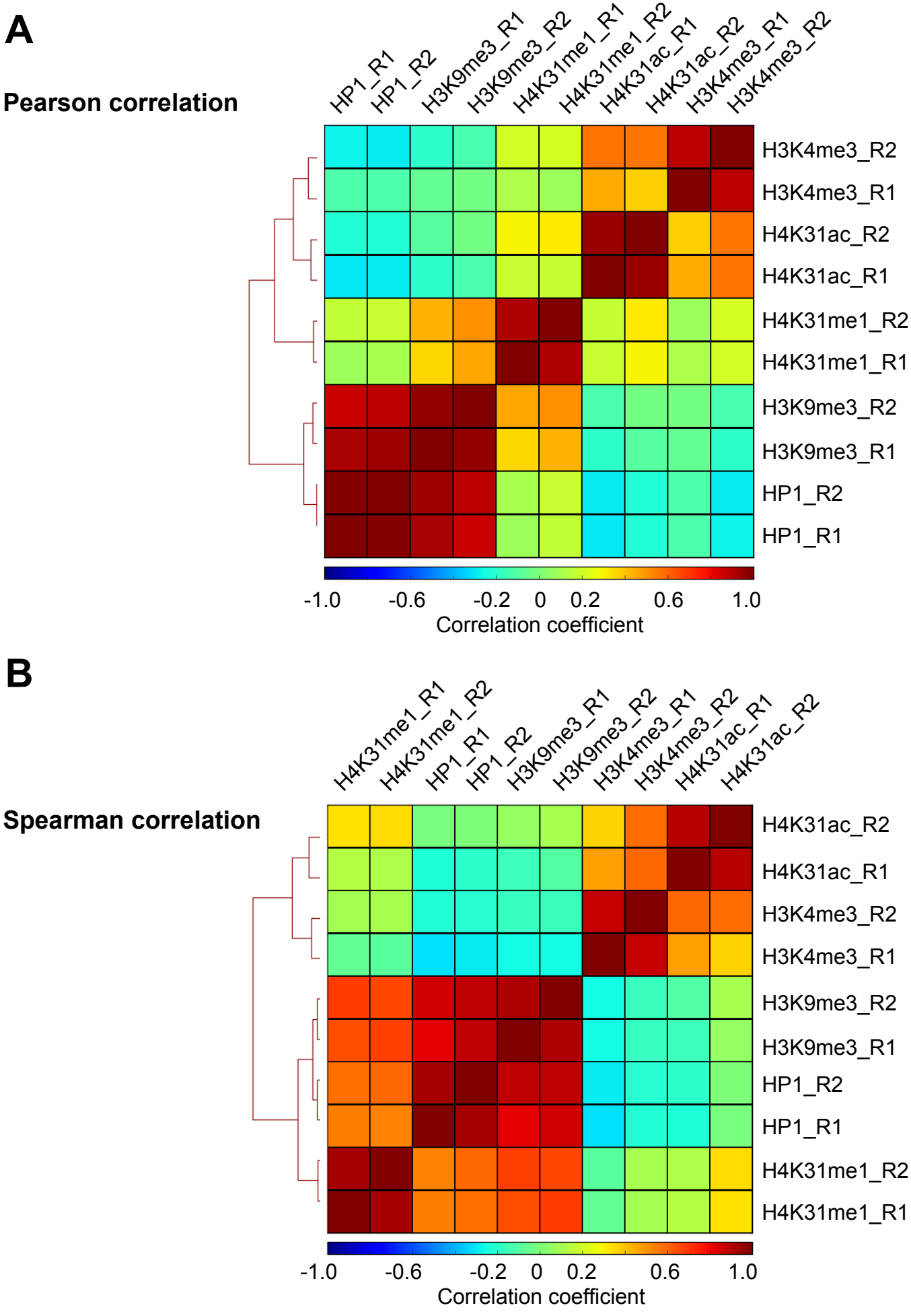


Figure 11 – figure supplement 1

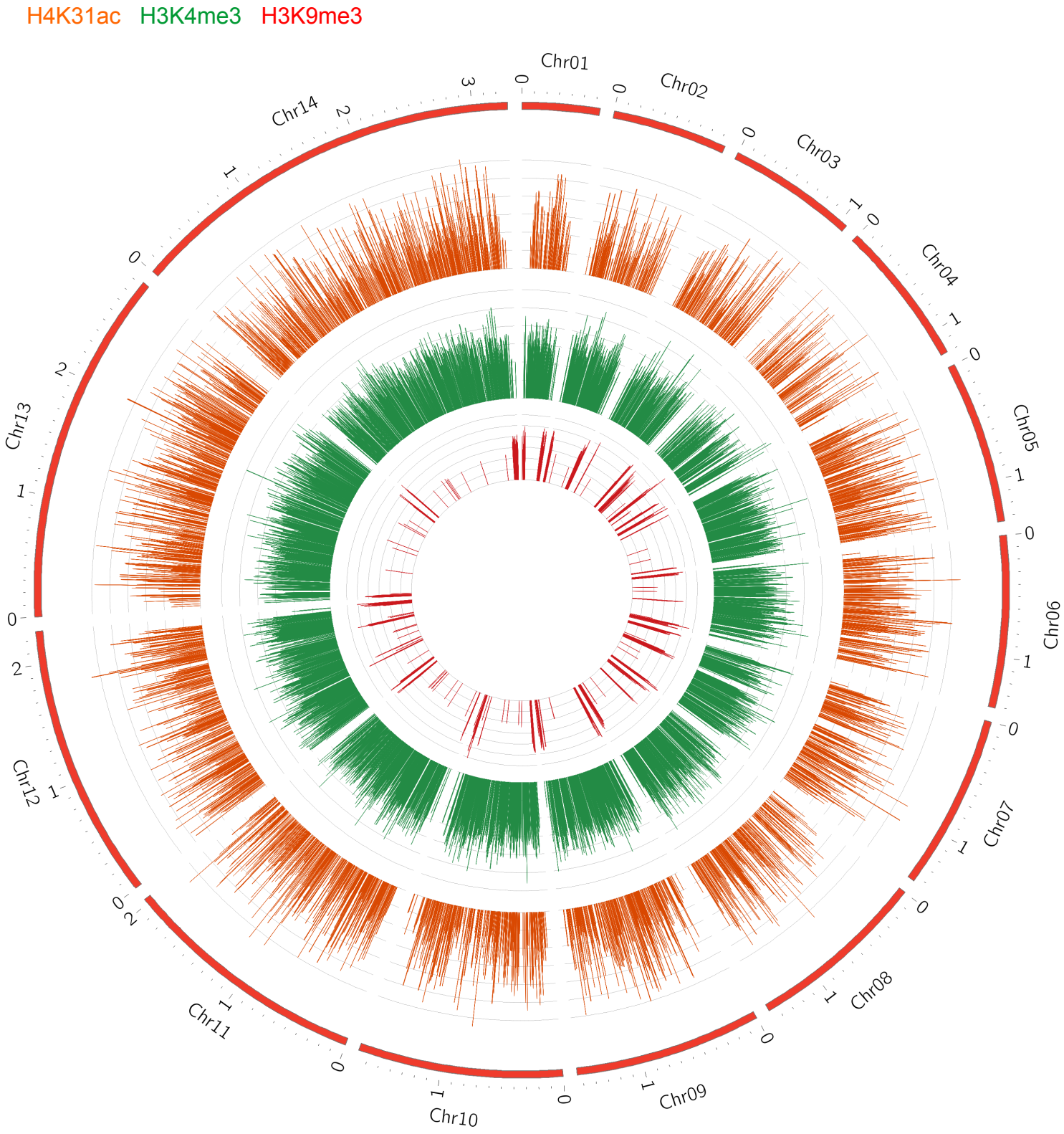


Figure 11 – figure supplement 2

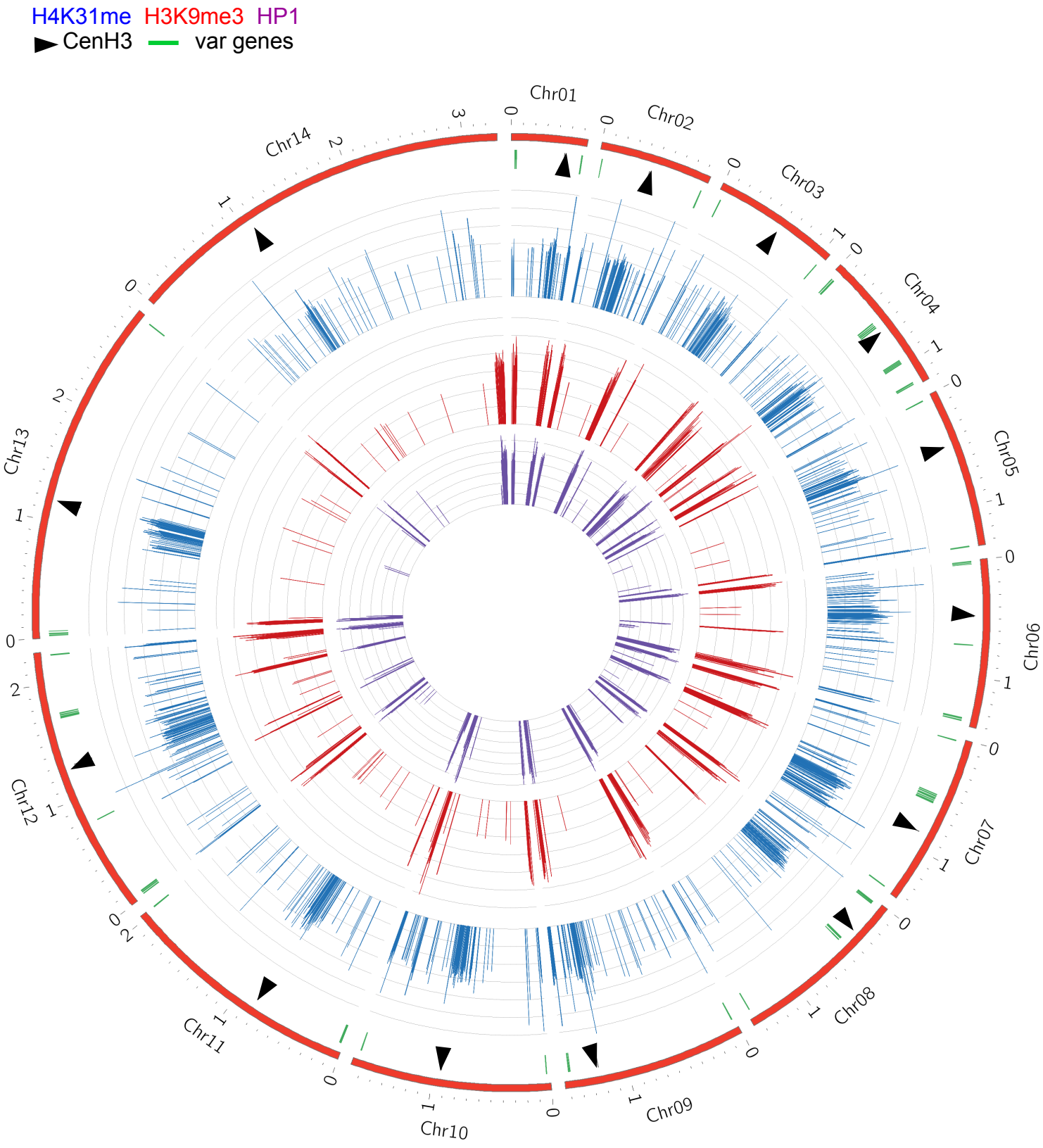


Figure 11 – figure supplement 3

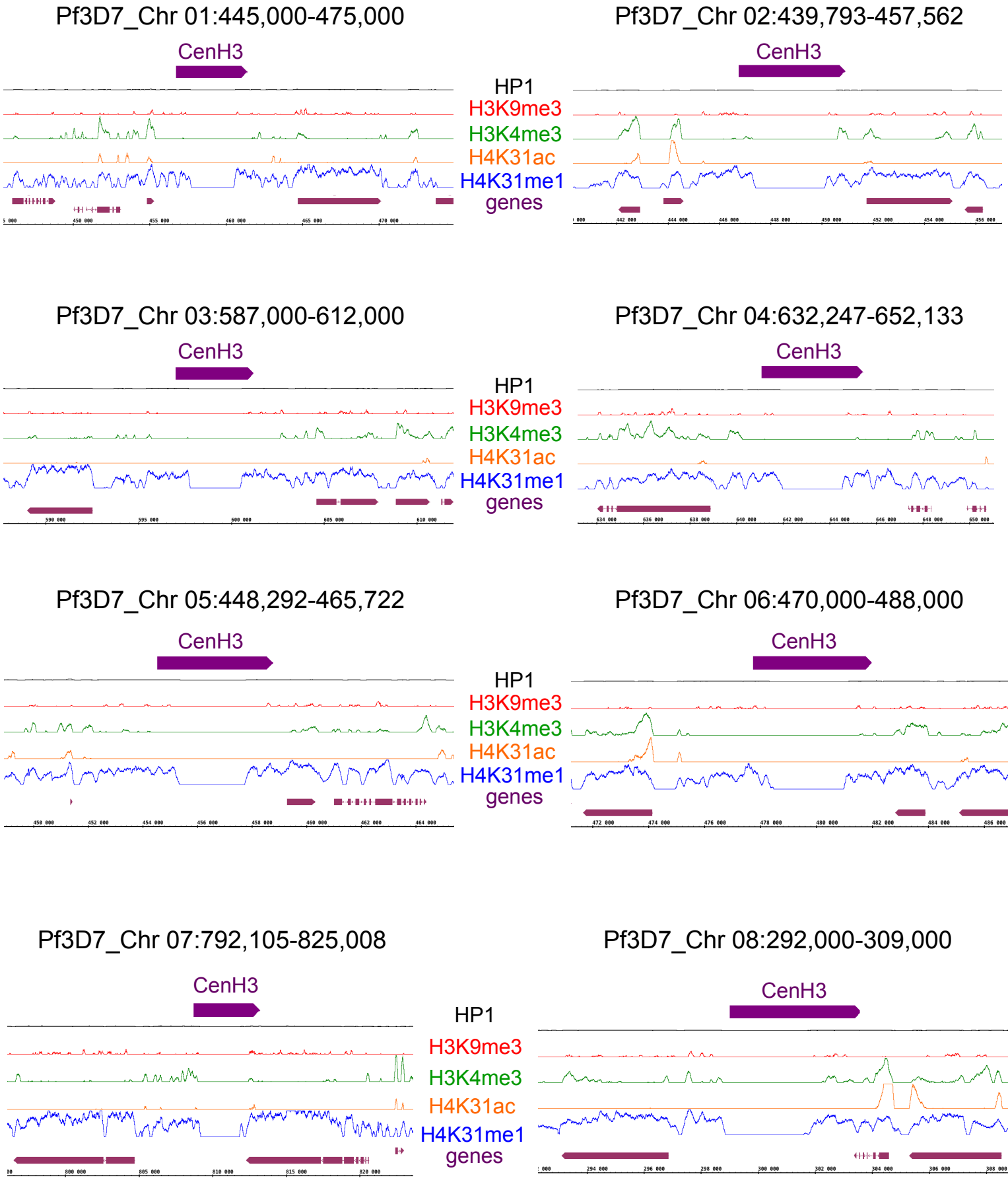


Figure 12 – figure supplement 1

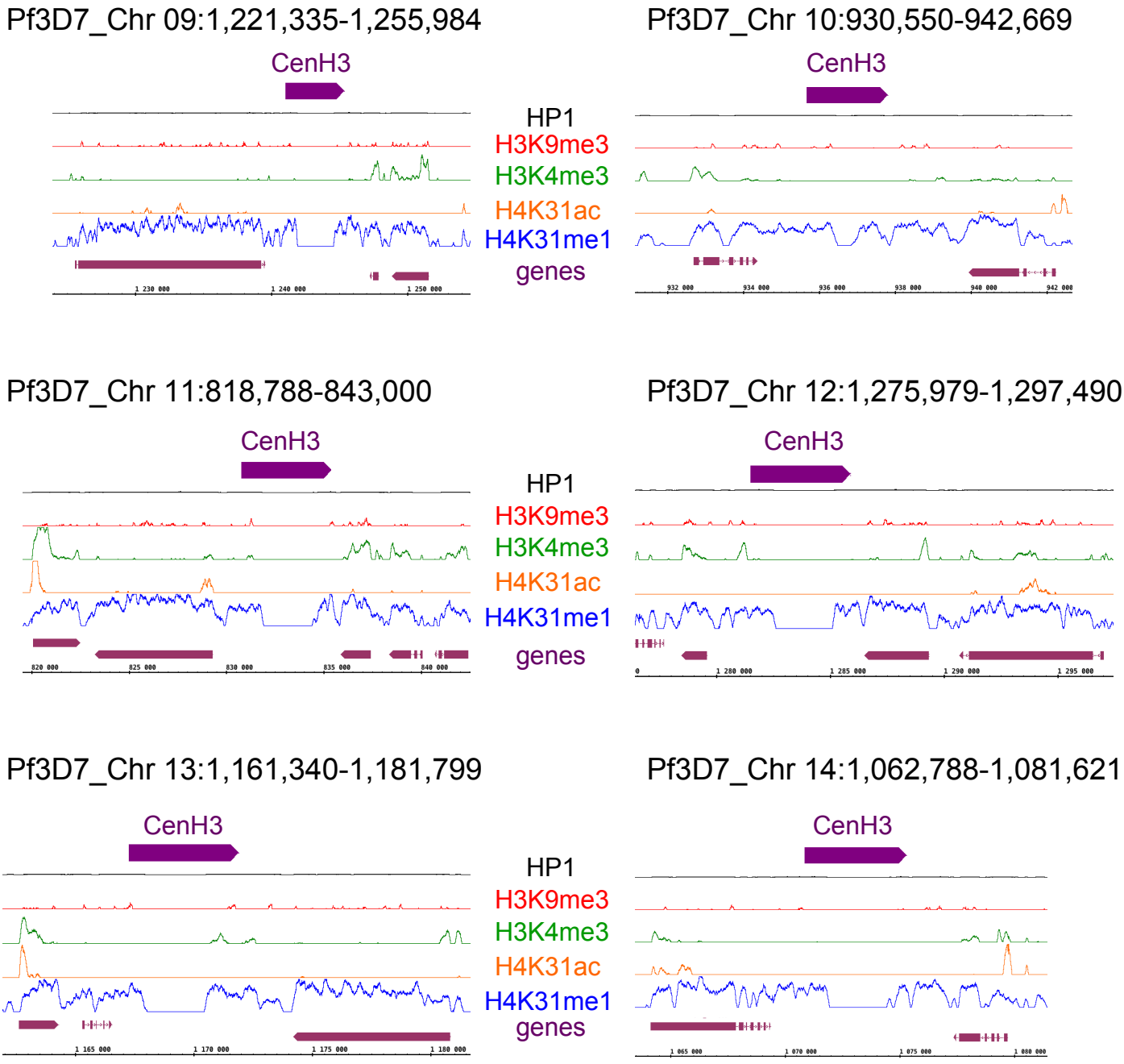


Figure 12 – figure supplement 2