A systematic analysis of *Trypanosoma brucei* chromatin factors identifies novel protein interaction networks associated with sites of transcription initiation and termination

Desislava P. Staneva*[1,2], Roberta Carloni*[1,2], Tatsiana Auchynnikava[1] Pin Tong[4], Juri Rappsilber[1,3], A. Arockia Jeyaprakash[1], Keith R. Matthews[2 †] and Robin C. Allshire[1 †]

[1] Wellcome Centre for Cell Biology and Institute of Cell Biology, School of Biological Sciences, University of Edinburgh. Edinburgh EH9 3BF, Scotland, UK

[2] Institute of Immunology and Infection Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, Scotland, UK.

[3] Institute of Biotechnology, Technische Universität, Gustav-Meyer-Allee 25, 13355 Berlin, Germany.

[4] Present address: MRC Human Genetics Unit, Institute of Genetics & Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, Scotland, UK.

* These authors contributed equally to this work.

[†] Co-corresponding authors:
Robin Allshire: robin.allshire@ed.ac.uk
Keith Matthews: keith.matthews@ed.ac.uk

Abstract

Nucleosomes composed of histones are the fundamental units around which DNA is wrapped to form chromatin. Transcriptionally active euchromatin or repressive heterochromatin is regulated in part by the addition or removal of histone post-translational modifications (PTMs) by 'writer' and 'eraser' enzymes, respectively. Nucleosomal PTMs are recognised by a variety of 'reader' proteins which alter gene expression accordingly. The histone tails of the evolutionarily divergent eukaryotic parasite *Trypanosoma brucei* have atypical sequences and PTMs distinct from those often considered universally conserved. Here we identify 65 predicted readers, writers and erasers of histone acetylation and methylation encoded in the *T. brucei* genome and, by epitope tagging, systemically localize 60 of them in the parasite's bloodstream form. ChIP-seq demonstrated that fifteen candidate proteins associate with regions of RNAPII transcription initiation. Eight other proteins exhibit a distinct distribution with specific peaks at a subset of RNAPII transcription termination regions marked by RNAPIII-transcribed tRNA and snRNA genes. Proteomic analyses identified distinct protein interaction networks comprising known chromatin regulators and novel trypanosome-specific components. Notably, several SET- and Bromo-domain protein networks suggest parallels to RNAPII promoter-associated complexes in conventional eukaryotes. Further, we identify likely components of TbSWR1 and TbNuA4 complexes whose enrichment coincides with the SWR1-C exchange substrate H2A.Z at RNAPII transcription start regions. The systematic approach employed provides detail of the composition and organization of the chromatin regulatory machinery in *Trypanosoma brucei* and establishes a route to explore divergence from eukaryotic norms in an evolutionarily ancient but experimentally accessible eukaryote.

Introduction

Nucleosomes are composed of eight highly conserved core histone subunits (two each of H2A, H2B, H3 and H4) around which approximately 147 bp of DNA is wrapped. Nucleosomes are organised into chromatin fibres which provide the dynamic organisational platform underpinning eukaryotic gene expression regulation. Formation of transcriptionally active and silent chromatin states depends on the presence of DNA methylation (Suzuki and Bird, 2008), histone variants (Henikoff and Smith, 2015) and histone post-translational modifications (PTMs) (Bannister and Kouzarides, 2011). Repressive heterochromatin generally concentrates at the nuclear periphery, while active euchromatin localizes to the nuclear interior and can also associate with nuclear pores (Lemaître and Bickmore, 2015; Taddei et al., 2010). Molecular understanding of the composition and function of distinct chromatin types in nuclear architecture and gene expression regulation is most advanced in well-studied eukaryotic models (plants, yeasts, animals) (Allshire and Madhani, 2018). However, these represent only two eukaryotic supergroups while distinct early-branching lineages have highly divergent histones and chromatin-associated regulators. One particularly tractable model for early branching eukaryotes is *Trypanosoma brucei*, the causative agent of human sleeping sickness and livestock nagana in Africa, which has evolved separately from the main eukaryotic lineage for at least 500 million years. Reflecting their evolutionary divergence, detailed analyses of these parasites have revealed numerous examples of biomolecular novelty, including RNA editing of mitochondrial transcripts (Shapiro and Englund, 1995), polycistronic transcription of nuclear genes (Borst, 1986; Tschudi and Ullu, 1988) and segregation of chromosomes via an unconventional kinetochore apparatus comprising components distinct from other eukaryotic groups (Akiyoshi and Gull, 2014).

During its life cycle, *T. brucei* alternates between a mammalian host and the tsetse fly vector, a transition accompanied by extensive changes in gene expression leading to surface proteome alterations as well as metabolic reprogramming of the parasite (Matthews, 2005; Smith et al., 2017). In the mammalian host, bloodstream form (BF) parasites are covered by a dense surface coat made of variant surface glycoprotein (VSG). Only a single VSG gene is expressed from an archive consisting of ~2000 VSG genes and gene fragments (Horn, 2014). Periodically, *T. brucei* switches to express a new VSG protein to which no host antibodies have been produced, contributing to cyclical parasitaemia. Parasites taken up by the tsetse during blood meals differentiate in the fly midgut to the procyclic form (PF) which replaces all VSGs with procyclin surface proteins (Roditi and Liniger, 2002).

The *T. brucei* genome encodes four core histones (H2A, H2B, H3, H4) and a variant for each core histone type (H2A.Z, H2B.V, H3.V, H4.V), but it lacks a centromere-specific CENP-A/cenH3 variant (Akiyoshi and Gull, 2014). All trypanosome histones differ significantly in their amino acid sequence from their counterparts in conventional eukaryotes (Lowell and Cross, 2004; Lowell et al., 2005; Mandava et al., 2007; Thatcher and Gorovsky, 1994). For example, lysine 9 of histone H3, methylation of which specifies heterochromatin formation in many eukaryotes, is not conserved. Nonetheless, many histone PTMs have been detected in *T. brucei* and its relative *T. cruzi* (de Jesus et al., 2016; de Lima et al., 2020; Janzen et al., 2006a; Kraus et al., 2020; Mandava et al., 2007), although only a handful of these have been characterized in some detail.

Unusually for a eukaryote, most trypanosome genes are transcribed in polycistronic units which are resolved by trans-splicing of a spliced leader (SL) RNA sequence to the 5′ end of the mRNA and polyadenylation at the 3′ end (Gunzl, 2010). RNAPII transcription usually initiates from broad (~10 kb) GT-rich divergent Transcription Start Regions (TSRs; comparable to promoters of other eukaryotes) enriched in nucleosomes containing the H2A.Z and H2B.V histone variants as well as the H3K4me (methylation) and H4K10ac (acetylation) PTMs (Siegel et al., 2009; Wedel et al., 2017; Wright et al., 2010). Conversely, RNAPII transcription typically terminates at regions of convergent transcription known as Transcription Termination Regions (TTRs) that are marked by the presence of the DNA modification base J and the H3.V and H4.V histone variants (Schulz et al., 2016; Siegel et al., 2009). Less frequently, *T. brucei* transcription units are arranged head-to-tail requiring termination ahead of downstream TSRs. TTRs between such transcription units are often coincident with RNAPI- or RNAPIII-transcribed genes which may act as boundaries that block the passage of advancing RNAPII into downstream transcription units (Marchetti et al., 1998; Maree and Patterton, 2014; Siegel et al., 2009).

The consensus view is that trypanosome gene expression is regulated predominantly post-transcriptionally via control of RNA stability and translation (Clayton, 2019). Nonetheless, numerous putative chromatin regulators can be identified as coding sequences in the *T. brucei* genome (Berriman et al., 2005), but their functional contexts are largely unexplored. Here we undertake cellular localization, genome-wide chromatin association and proteomics analyses of bioinformatically identified putative readers, writers and erasers of histone acetyl and methyl marks encoded in the trypanosome genome. The results presented provide an entry point for understanding similarities and differences between the transcriptional regulatory machinery of this divergent eukaryote and the eukaryotic norm.

Results and Discussion

Identification of putative chromatin regulators

We identified 65 putative regulators or interpreters of histone lysine acetylation and methylation by interrogating the trypanosome genome database (TriTrypDB) and using additional homology-based searches with known chromatin reader, writer and eraser domains (Table 1; Supplemental Fig. S1A; Supplemental Table S1). These approaches allowed us to detect 16 potential readers with the following domains: Bromo (Haynes et al., 1992), PHD (Aasland et al., 1995), Tudor (Ponting, 1997), Chromo (Paro, 1990; Singh et al., 1991), PWWP (Stec et al., 2000) and Znf-CW (Perry and Zhao, 2003). With respect to writers of histone modifications, we found 6 potential histone acetyltransferases (HATs) belonging to the MYST (MOZ/SAS-related) and GNAT (ELP3-related) families (Lee and Workman, 2007), 29 SET domain (Dillon et al., 2005) and 3 DOT domain (Feng et al., 2002) putative histone methyltransferases (HMTs). We also analyzed several predicted acetyltransferases with non-histone substrates: one lysophospholipid acyltransferase (LPLAT) (Shindou and Shimizu, 2009) and two GNAT family (RimI-related) N-acetyltransferases (NATs) (Vetting et al., 2008) plus the non-catalytic EAF6 component of the NuA4 HAT complex (Hishikawa et al., 2008; Roth et al., 2001). Our searches for potential erasers of histone PTMs identified 7 histone deacetylases (class I, class II and Sir2-related HDACs) (Grozinger and Schreiber, 2002) and 4 JmjC domain demethylases (Klose et al., 2006). The function of some of these putative chromatin regulators has been explored previously (Figueiredo et al., 2009; Maree and Patterton, 2014; Supplemental Table S2).

To examine these putative writers, readers and erasers of histone PTMs, each was YFP-tagged and localized in bloodstream form parasites. AGO1 (Shi et al., 2009) and the putative DNA methyltransferase DMT (Militello et al., 2008) were also included in our analysis since they are associated with chromatin-based silencing in other eukaryotes (Allshire and Madhani, 2018; Kloc and Martienssen, 2008; Suzuki and Bird, 2008).

Additionally, we examined several control proteins with known distinctive nuclear distributions; these were the trypanosome kinetochore protein KKT2 (Akiyoshi and Gull, 2014), the nucleolar protein NOC1 (Alsford and Horn, 2012), the nuclear pore basket protein NUP110/MLP1 (DeGrasse et al., 2009) as well as the telomere repeat-binding factor TRF (Li et al., 2005) and the TATA-binding related protein TBP/TRF4 (Ruan et al., 2004). AGO1 served as a control for cytoplasmic localization (Shi et al., 2004).

Cellular localization of the putative chromatin regulators

Candidate proteins were endogenously tagged with YFP in bloodstream form parasites of the *T. brucei* Lister 427 strain. Proteins were tagged N-terminally to avoid interference with 3′ UTR sequences involved in mRNA stability control (Clayton, 2019). Although the presence or position of the YFP tag might affect individual protein expression and localization patterns, this approach provided a consistent pipeline for the systematic analysis of the proteins on our candidate list. Of the 76 selected proteins (including controls), 74 were successfully tagged while cells expressing YFP-PHD3 and YFP-ELP3c were not obtained. The tagging constructs for SET12, SET30 and DOT1 were correctly integrated but YFP-tagged proteins were not detectable by western analysis suggesting tag failure, low protein abundance or no expression in bloodstream form parasites.

Immunolocalization with anti-GFP antibodies was used to identify proteins residing in the nucleus which might associate with chromatin. The control proteins exhibited localization patterns expected for telomeres (TRF – nuclear foci), TATA-binding protein (TBP – nuclear foci), kinetochores (KKT2 – nuclear foci), nucleolus (NOC1 – single nuclear compartment), nuclear pores (NUP110 – nuclear rim) and cytoplasm (AGO1 – gap in the staining coincident with the nucleus) (Supplemental Fig. S2). Some cytoplasmic signal was also detected for nuclear control proteins, but this may correspond to background staining as evidenced by the signal observed in untagged cells (Fig. 1; Supplemental Fig. S2). Selected proteins were also imaged with and without antibody staining and for different exposure times (Supplemental Fig. S3). From this we concluded that fluorescence signal enhancement through antibody staining was required for optimal imaging of most proteins and that different exposure times were needed to adjust to the expression level of each protein. Images with multiple cells per field in different cell cycle stages are also provided (Supplemental Fig. S4).

Of the YFP-tagged proteins, 20 were exclusively nuclear, 30 exhibited only a cytoplasmic localization and 21 were found in both compartments (Table 1; Supplemental Table S1). Twenty-three candidate and control proteins with exclusive or some nuclear localization were subsequently found to associate with either sites of RNAPII transcription initiation or with a subset of RNAPII transcription termination regions coincident with RNAPIII-transcribed genes (see below; Fig. 1), while the remaining proteins that were not detected on chromatin displayed all three localization patterns (Supplemental Fig. S2). Consistent with previous observations, HAT1-to-3 decorated nuclear substructures (Kawahara et al., 2008), as did the predicted EAF6 subunit of the NuA4 HAT complex. In contrast, LPLAT1

gave both nuclear and cytoplasmic signal whereas both NAT2 and NAT3 localized to the cytoplasm (Supplemental Fig. S2). YFP-tagged ELP3a and ELP3b GNAT acetyltransferases exhibited nuclear and some cytoplasmic signal. GFP-ELP3b was previously reported to be concentrated in the nucleolus (Alsford and Horn, 2011), this difference may be a consequence of using an ectopic overexpression system in that study. Moreover, we did not detect enrichment of ELP3b over rDNA transcription units by ChIP-seq. Of the 29 identified putative SET-domain methyltransferases, only eight exhibited some nuclear localization, while most were concentrated in the cytoplasm. Examination of cells expressing YFP-tagged predicted reader proteins revealed that BDF1-to-3, BDF5-to-7, PHD1, PHD2, PHD4, CRD1, TFIIS2-2 and ZCW1 were exclusively nuclear, whereas BDF4 and PHD5 displayed both nuclear and cytoplasmic localization and the sole Tudor domain protein TDR1 was cytoplasmic. In agreement with previous analyses (Wang et al., 2010), HDAC3 was predominantly nuclear whereas HDAC1 was nuclear/cytoplasmic, and both HDAC2 and HDAC4 resided in the cytoplasm. The Sir2-related proteins Sir2rp1 and Sir2rp2 localized mostly to the cytoplasm whereas Sir2rp3 was detected in the nucleus as well as in the cytoplasm. Of the four identified putative demethylases, JMJ2 was nuclear, JMJ1 and CLD1 were cytoplasmic and LCM1 was found in both compartments. Both AGO1 and DMT were cytoplasmic (Supplemental Fig. S2) suggesting that they are unlikely to be involved in directing chromatin or DNA modifications.

Of the 71 proteins we successfully tagged, expressed and localized in bloodstream form cells, 65 have also been also tagged by the TrypTag project in procyclic form cells (Dean et al., 2017; Table S1). For most proteins, our results are in agreement with the TrypTag data and any discrepancies may be indicative of differences in protein localization between the different developmental forms of *T. brucei*.

Many putative chromatin regulators accumulate at RNAPII TSRs

We performed ChIP-seq for all expressed candidate and control proteins except TDR1 and NOC1 (69 in total) to determine which proteins associate with chromatin and assess their distribution across the *T. brucei* genome. Our rationale was that ChIP-seq might register chromatin association even if cellular localization analysis reported a protein to be predominantly cytoplasmic. As expected, the kinetochore control protein KKT2 was specifically enriched over centromeric regions (Fig. 2A). Consistent with their cytoplasmic localizations, AGO1 and DMT registered no ChIP-seq signal (Supplemental Fig. S5). Moreover, under our standard fixation and ChIP-seq conditions, no specific enrichment over any genomic region was detected for 45 of the YFP-tagged proteins, including several that exhibited clear nuclear localization (Supplemental Fig. S5; Supplemental Table S3).

In total, 24 of the YFP-tagged proteins assayed (including the KKT2 control) gave specific enrichment patterns across the *T. brucei* genome. Fifteen of our candidate proteins displayed enrichment that was adjacent to, or overlapping with, the previously reported H2A.Z peaks indicating that these proteins are enriched at known RNAPII TSRs (Fig. 2A; Supplemental Fig. S6A). Our ChIP-seq data confirmed that the largest RNAPII subunit (RPB1) does indeed exhibit peaks at the same locations. The TSR-associated proteins include BDF1-to-6, CRD1, EAF6, HAT1, HAT2, HDAC1, HDAC3, SET26, SET27, and ZCW1. These fifteen proteins were always enriched together at RNAPII TSRs. Some of these proteins also presented peaks that coincide with those of RNAPIII/TTR-enriched factors in a few locations (see below; Supplemental Fig. S6C) but otherwise they did not display significant association with any other genomic regions. The fact that six Bromodomain proteins, two histone acetyltransferases and a NuA4 component are included in this set is consistent with these acting together to promote RNAPII-mediated transcription, a hallmark of which is histone acetylation (Roth et al., 2001). Indeed, BDF1, BDF3 and BDF4 have previously been shown to be enriched at sites of *T. brucei* RNAPII transcription initiation where nucleosomes exhibit HAT1-mediated acetylation of H2A.Z and H2B.V and HAT2-mediated acetylation of histone H4, which are important for normal RNAPII transcription from these regions (Kraus et al., 2020; Schulz et al., 2015; Siegel et al., 2009). The remaining 10 proteins we identified as being associated with TSRs have not been previously shown to act at kinetoplastid RNAPII promoters.

Putative chromatin regulators exhibit two distinct TSR association patterns

In yeast, where many RNAPII genes are transcriptionally regulated, a group of chromatin regulators are enriched specifically at promoters where they assist in transcription initiation while others travel with RNAPII into gene bodies aiding transcription elongation, splicing and termination (Carrozza et al., 2005; Cheung et al., 2008; Jonkers and Lis, 2015; Joshi and Struhl, 2005; Kaplan et al., 2003; Keogh et al., 2005; Li et al., 2007; Mason and Struhl, 2003; Venkatesh and Workman, 2013). We therefore compared the enrichment profiles of the TSR-associated proteins relative to that of the Chromo domain protein CRD1 which displayed the sharpest and highest peak signal. We identified 177 CRD1 peaks across the *T. brucei* Lister 427 genome which overlap with 136 of the 148 annotated RNAPII TSRs (Fig. 2B). For each TSR-associated factor, normalized reads were assigned to 10 kb windows upstream and downstream of all CRD1 peak summits. The general peak profile for each protein was then displayed as a metagene plot and the read distribution around individual CRD1 peaks represented as a heatmap (Fig. 2C,D; Supplemental Fig. S7). This analysis indicated that CRD1, SET27, BDF4 and, to a lesser extent, BDF1 and BDF3 exhibit sharp

peaks at all RNAPII TSRs, similar to RNAPII/RPB1 itself. We refer to these as Class I TSR-associated factors (Fig. 2C). The remaining 10 proteins were more broadly enriched over the same regions with a slight trough evident in the signal for at least five (BDF2, BDF5, HAT1, HAT2, SET26), perhaps indicative of two adjacent peaks as observed for H2A.Z (Fig. 2D). We refer to these as Class II TSR-associated factors. Class II factor profiles are similar to the previously reported H2A.Z enrichment pattern (Siegel et al., 2009), with the signal gradually declining over 5-10 kb regions on either side of the CRD1 peak summit. We observed similar peak widths of Class II factors across all TSRs regardless of polycistron length, thus there is no apparent relationship between TSR size and the length of the downstream polycistronic transcription unit.

Of the 148 annotated *T. brucei* RNAPII promoters, 99 are bidirectional, initiating production of stable transcripts in both directions, and 49 are unidirectional, driving transcription in just one direction. To investigate the relationship between transcription directionality and enrichment of our candidate proteins, we sorted the heatmaps of Class II factors by their distribution around CRD1 peaks. We then compared the sorted heatmaps with previously published RNA-seq data from which the direction of RNAPII transcription was derived (Naguleswaran et al., 2018). This analysis demonstrated that Class II TSR-associated factors exhibit specific enrichment in the same direction as RNAPII transcription initiated from all uni- and bi-directional promoters (Fig. 3A,B).

Proteins associated with TSRs participate in discrete interaction networks

The analyses presented above suggest that, as in yeast, proteins which exhibit either narrow (Class I) or broad (Class II) association patterns across RNAPII promoter regions might play different roles such as defining sites of RNAPII transcription initiation or facilitating RNAPII processivity through their association with chromatin and interactions with RNAPII auxiliary factors. Determining how these various activities are integrated through association networks should provide insight into how the distinct sets of proteins might influence RNAPII transcription. Therefore, we affinity selected each tagged protein that registered a specific ChIP-seq signal at TSRs and identified their interacting partners by mass spectrometry.

Below we detail the interaction networks for the TSR-associated factors (Fig. 4A,B; Supplemental Table S4), the homologies identified in their key interacting partners through HHpred searches (Soding et al., 2005; Supplemental Table S5) and discuss possible functional implications arising from these proteomics data. Interacting partners identified by mass spectrometry analysis are also included for ten proteins (PHD1, PHD5, HAT3, AGO1,

NUP110, SET13, SET15, SET20, SET23 and SET25) for which no specific ChIP-seq signal was obtained (Supplemental Fig. S8).

Class I: CRD1, SET27, BDF4, BDF1 and BDF3

CRD1 and SET27

Affinity selection of YFP-CRD1 and YFP-SET27 (Fig. 4A,B; Supplemental Table S4) revealed that they both associate strongly with each other, with four uncharacterized proteins (Tb927.1.4250, Tb927.3.2350 Tb927.11.11840, Tb927.11.13820) and with JBP2. JBP2 is a TET-related hydroxylase that catalyses thymidine oxidation on route to the synthesis of the DNA modification base J which is found at transcription termination regions and telomeres in trypanosomes (Cliffe et al., 2010; Reynolds et al., 2016; Schulz et al., 2016). In addition, SET27 associates with JBP1 - another TET-related thymidine hydroxylase involved in base J synthesis (Borst and Sabatini, 2008). The Chromo domain of CRD1 exhibits marginal similarity when aligned with other Chromo domain proteins (Supplemental Fig. S1B), however its reciprocal association with SET27 suggests that they might function together as a reader-writer pair at TSRs. In yeast and human cells, the Set1/SETD1 methyltransferase installs H3K4 methylation at promoters (Shilatifard, 2012) and *T. brucei* SET27 might play a similar role at TSRs. Recently, methylated histone lysine residues were found to be prevalent at trypanosome TSRs (Kraus et al., 2020). SET27 likely catalyses the methylation of at least one of these lysines which may then be bound by CRD1, ensuring SET27 recruitment and persistence of the methylation event(s) that it installs on histones within resident TSR nucleosomes. The association of the RPB1 and RPB3 RNAPII subunits with CRD1 underscores its likely involvement in linking such chromatin modifications with transcription.

BDF1 and BDF4

Consistent with their colocalization in Class I ChIP-seq peaks, YFP-BDF1 and YFP-BDF4 showed strong reciprocal association with each other (Fig. 4A,B; Supplemental Table S4). BDF4 also exhibited weak association with BDF3 and the Class II factor BDF5. Bromo domains are known to bind acetylated histones (Zaware and Zhou, 2019), and are thus presumably attracted to TSRs due to the presence of highly acetylated histones, particularly H2A.Z, H2B.V and H4, in resident nucleosomes (Kraus et al., 2020). The coincidence of the H2A.Z variant with histone modifications associated with active transcription in this evolutionarily distinct eukaryote suggests that they act together to recruit various chromatin remodelling and modification activities to ensure efficient transcription.

BDF3 (Class I), BDF5 (Class II) and HAT2 (Class II)

BDF3, BDF5 and HAT2 reciprocally associate with each other and a set of six uncharacterized proteins (Tb927.3.4140, Tb927.4.2340, Tb927.6.1070, Tb927.7.2770, Tb927.9.13320, Tb927.11.5230) suggesting that these nine proteins may act together in a complex (Fig. 4A,B; Supplemental Table S4). BDF3 interacts with both Class I and Class II proteins, suggesting that it may straddle the interface between both classes of factors at RNAPII TSRs. HAT2 mediates acetylation of histone H4 and promotes normal transcription initiation by RNAPII (Kraus et al., 2020). The Bromo domains of BDF3 and BDF5 may guide HAT2 to pre-existing acetylation at TSRs to maintain the required acetylated state for efficient transcription. We also note that Tb927.3.4140 displays similarity to Poly ADP Ribose Polymerase (PARP; Supplemental Table S5), suggesting that ribosylation might contribute to TSR definition by promoting chromatin decompaction as seen upon *Drosophila* heat shock puff induction (Sawatsubashi et al., 2004; Tulin et al., 2003; Tulin and Spradling, 2003) and at some mammalian enhancers-promoter regions (Benabdallah et al., 2019). Tb927.11.5230 contains an EMSY ENT domain whose structure has been determined (Mi et al., 2018; Supplemental Table S5); such domains are present in several chromatin regulators. Tb927.4.2340 exhibits similarity to the C-terminal region of the vertebrate TFIID TAF1 subunit (Supplemental Table S5). Metazoan TAF1 bears two Bromo domains in its C-terminal region whereas in yeast the double Bromo domain component of TFIID is contributed by the separate Bdf1 (or Bdf2) proteins (Matangkasombut et al., 2000; Timmers, 2020). *T. brucei* BDF5 contains two Bromo domains and may thus be equivalent to the yeast TFIID Bdf1 subunit.

<u>Class II: BDF2, BDF5, BDF6, EAF6, HAT1, HAT2, HDAC1, HDAC3, SET26, ZCW1</u>

BDF2 and HDAC3

We found that both TSR-enriched histone variants H2A.Z and H2B.V strongly associate with BDF2 and HDAC3, which interact with each other as well as with four uncharacterized proteins (Tb927.3.2460, Tb927.6.4330, Tb927.9.4000, Tb927.9.8520; Fig. 4A,B; Supplemental Table S4). We note that Tb927.9.8520 exhibits similarity to DNA Polymerase Epsilon (DNAPolE; Supplemental Table S5) and both DNA Polymerase Theta (DNAPolQ) and DNA Primase were also enriched along with the PARN3 poly(A) specific ribonuclease and Casein Kinase I (CKI). Moreover, Tb927.3.2460 shows similarity to a nuclear pore protein while Tb927.6.4330, Tb927.9.4000 and DNAPolQ were previously shown to associate with the telomere proteins TRF and TelAP1, and DNA primase is known to interact with TelAP1 (Reis et al., 2018). Here we find that the telomere-associated proteins TRF,

TIF2, TelAP1 and RAP1 are also enriched in both BDF2 and HDAC3 affinity selections. The significance of this association is unknown, however BDF2 and HDAC3 have previously been shown to be required for telomeric VSG expression site silencing (Schulz et al., 2015; Wang et al., 2010) and RNAi knock-down of Tb927.6.4330 causes defects in telomere-exclusive VSG gene expression (Glover et al., 2016). It was also unexpected to observe HDAC3 enrichment at RNAPII TSRs given that actively transcribed regions tend to be hyperacetylated (Kouzarides, 2007). HDAC3 may be required to reverse acetylation associated with newly deposited histones during S phase (Stewart-Morgan et al., 2020) or to remove acetylation added during expected H2A-H2B:H2A.Z-H2B.V dimer-dimer exchange events at TSR regions (Bonisch and Hake, 2012; Millar et al., 2006).

SET26 and *ZCW1*

SET26 and ZCW1 associated with each other as well as with most histones and histone variants (Fig. 4A,B; Supplemental Table S4). In addition, both SET26 and ZCW1 showed strong interaction with the SPT16 and POB3 subunits of the FACT (Facilitates Chromatin Transcription) complex which is involved in trypanosome VSG silencing through increased histone occupancy at VSG expression sites (Denninger and Rudenko, 2014) and is known to aid transcription elongation in other eukaryotes (Belotserkovskaya et al., 2003). SET26 could perform an analogous role to the yeast Set2 H3K36 histone methyltransferase which travels with RNAPII and, together with FACT, ensures that chromatin integrity is restored behind advancing RNAPII. These activities are known to prevent promiscuous transcription initiation events from cryptic promoters within open reading frames (Carrozza et al., 2005; Cheung et al., 2008; Joshi and Struhl, 2005; Kaplan et al., 2003; Keogh et al., 2005; Li et al., 2007; Mason and Struhl, 2003; Venkatesh and Workman, 2013). ZCW1 also exhibits strong association with apparent orthologs of several SWR1/SRCAP/EP400 remodelling complex subunits (Scacchetti and Becker, 2020; Willhoft and Wigley, 2020), including Swr1/SRCAP (Tb927.11.10730), Swc6/ZNHIT1 (Tb927.11.6290), Swc2/YL1 (Tb927.11.5830), two RuvB-related helicases (Tb927.4.2000; Tb927.4.1270), actin and actin-related proteins (Tb927.4.980, Tb927.10.2000, Tb927.3.3020), and the possible equivalents of the Swc4/DMAP1 (Tb927.7.4040) and Yaf9/GAS41 YEATS domain protein (Tb927.10.11690, designated YEA1) subunits (Supplemental Table S5; Supplemental Table S6). Most of these putative TbSWR1-C subunits were also detected as being associated with BDF2 (Fig. 4A,B; Supplemental Table S4; Supplemental Table S5). Thus, since the yeast Bdf1 and human BRD8 Bromodomain proteins also associate with SWR1-C/SRCAP-C/EP400, it is likely that TbBDF2 performs a similar function in engaging acetylated histones. The SWR/SRCAP remodelling complexes are well known for being required to direct the replacement of H2A with H2A.Z in nucleosomes residing close to transcription start sites (Mizuguchi et al., 2004;

Ruhl et al., 2006). The prevalence of both H2A.Z and H2B.V with affinity selected ZCW1 suggests that it may also play a role in directing the *T. brucei* SWR/SRCAP complex to TSRs to ensure incorporation of H2A.Z-H2B.V in place of H2A-H2B in resident nucleosomes.

BDF6, EAF6 and HAT1

BDF6, EAF6 and HAT1 exhibit robust association with each other, with a second YEATS domain protein - the possible ortholog of acetylated-H2A.Z binding Yaf9/Gas41 (Tb927.7.5310, designated YEA2) and a set of five other uncharacterized proteins (Tb927.1.650, Tb927.6.1240, Tb927.8.5320, Tb927.10.14190, Tb927.11.3430; Fig. 4A,B; Supplemental Table S4). Tb927.1.650 is an MRG domain protein with similarity to yeast Eaf3 while Tb927.10.14190 appears to be orthologous to yeast Epl1, both of which along with Yaf9 are components of the yeast NuA4 complex (Supplemental Table S5; Supplemental Table S7). The yeast Yaf9 YEATS protein contributes to both the NuA4 HAT and SWR1 complexes and can bind acetylated or crotonylated histone tails (Arrowsmith and Schapira, 2019; Timmers, 2020). In *T. brucei*, it appears that distinct YEATS domain proteins contribute to putative SWR1 (YEA1) and NuA4 (YEA2) complexes. HAT1 associates with both BDF6 and EAF6 whereas HAT3, EAF6 and PHD1 reciprocally associate with each other and share Tb927.11.7880, a YNG2-related ING domain protein, as a common interactor (Supplemental Fig. S8; Supplemental Table S5; Supplemental Table S7). Thus, HAT3-EAF6-PHD1-YNG2 perhaps represents a *T. brucei* subcomplex analogous to yeast piccolo-NuA4 while HAT1-EAF6-BDF6-EAF3-EPL1-YEA2 may form the larger NuA4 complex (Doyon and Cote, 2004; Wang et al., 2018). In *T. brucei*, the Esa1/TIP60 catalytic MYST acetyltransferase function may be shared between HAT1 and HAT3. Although we have clearly identified NuA4-like complexes, no association was detected with a *T. brucei* Eaf1/EP400-related Helicase-SANT domain protein that provides the platform for the assembly of distinct modules of the yeast and metazoan NuA4 complexes (Levi et al., 1987; Scacchetti and Becker, 2020).

Together, BDF6, EAF6, and HAT1 appear part of a putative NuA4-related complex that functions at *T. brucei* TSRs. HAT1 has been shown to be required for H2A.Z and H2B.V acetylation and efficient RNAPII engagement and transcription (Kraus et al., 2020). BDF6, YEA2 or both may bind acetylated histones at TSRs to promote stable association of interacting chromatin modification and remodelling activities that enable the required histone

dynamics to take place in these highly specialized regions and thereby facilitate efficient transcription.

HDAC1

We also readily detect HDAC1 enriched broadly over RNAPII TSRs, but notably it does not associate with any of the other factors we found in these regions (Fig. 4A; Supplemental Table S4). However, five uncharacterized proteins (Tb927.3.890, Tb927.4.3730, Tb927.6.3170, Tb927.7.1650, Tb927.9.2070) reproducibly associated with HDAC1. HHpred searches detected some similarity of Tb927.3.890, Tb927.4.3730 and Tb927.6.3170 to chromatin-associated proteins (Supplemental Table S5) and Tb927.9.2070 was also enriched with CRD1. HDAC1 was previously shown to be an essential nuclear protein whose knock down increases silencing of telomere-adjacent reporters in bloodstream form parasites (Wang et al., 2010). However, a general role for HDAC1 at RNAPII TSRs was not anticipated. HDAC1 has been reported to be mainly cytoplasmic in procyclic cells (Wang et al., 2010) and it is therefore expected to be absent from RNAPII TSRs in insect form parasites.

Eight proteins are specifically enriched over RNAPII TTRs coincident with RNAPIII-transcribed genes

Our analysis of ChIP-seq association patterns also revealed a distinct set of eight proteins which displayed sharp peaks at a subset of RNAPII transcription termination regions that coincide with RNAPIII-transcribed genes (Fig. 5; Supplemental Fig. S6B). These RNAPIII/TTR-associated factors included six of our candidate readers and writers (BDF7, ELP3b, PHD2, PHD4, TFIIS2-2 and DOT1A) as well as two of the selected control proteins (TRF and TBP). A total of 154 RNAPII TTRs have been annotated in the Lister 427 genome (Müller et al., 2018). We observed enrichment of some or all of these eight proteins at the 20 TTRs which overlap with tRNA and/or snRNA genes and also at the single TTR6 (sTTR6) which lacks such genes (Fig. 5A,B). Consistent with *T. brucei* snRNA and tRNA genes being transcribed by RNAPIII (Nakaar et al., 1997; Tschudi and Ullu, 2002), we detected enrichment of RPC1 (the largest RNAPIII subunit) at these locations (Fig. 5A). However, although RPC1 associates with sTTR6, it is not enriched at any of the other 133 TTRs lacking annotated RNAPIII-transcribed genes. Our analysis suggests that an unannotated RNAPIII transcript is probably produced within sTTR6 but that the majority of RNAPII TTRs are not associated with RNAPIII transcription or with any of the eight highlighted proteins.

These eight RNAPIII/TTR-associated factors were also clearly enriched at the U2 (Tb927.2.5680) and U6 (Tb927.4.1213) snRNA genes and at 53 of the 69 tRNA genes annotated in the Lister 427 genome assembly (Fig. 5A,B; Supplemental Fig. S6B; Supplemental Table S8). We observed enrichment of only some of these proteins at 11 of the other tRNA genes and no association with the remaining 5. Moreover, not all tRNA genes bound by these proteins coincide with an RNAPII TTR (Supplemental Fig. S6B; Supplemental Table S8).

RNAPIII/RPC1 was also enriched over the 5S rRNA cluster together with TBP but not the other seven members of this protein group (Fig. 5C). However, all eight proteins gave prominent peaks over the spliced leader gene locus (Fig. 5D). The fifteen RNAPII promoter-associated factors discussed above also exhibit sharp peaks that appear to coincide with those of the RNAPIII/TTR-associated factors at some locations (Supplemental Fig. S6C); the significance of this colocation is unknown.

It was unexpected that the terminal $(TTAGGG)_n$ telomere repeat-binding protein TRF was one of the eight proteins enriched at internal chromosomal regions overlapping RNAPIII-associated TTRs. We hypothesized that enrichment of TRF at this subset of TTRs could result from the presence of underlying sequence motifs with similarity to canonical $(TTAGGG)_n$ telomere repeats, however sequence scrutiny revealed no significant matches. *T. brucei* contains approximately 115 linear chromosomes and consequently has an abundance of telomeres and telomere binding proteins that cluster at the nuclear periphery (Akiyoshi and Gull, 2013; DuBois et al., 2012; Reis et al., 2018; Yang et al., 2009). The tethering of RNAPIII-transcribed genes to the nuclear periphery, as observed in yeast (Chen and Gartenberg, 2014; Iwasaki et al., 2010), would bring them in close proximity to telomeres offering a potential explanation for the association of TRF with these nucleosome depleted regions.

In yeasts, both the cohesin and condensin complexes, which shape chromosome architecture, are enriched or loaded at highly transcribed regions such as tRNA genes (D'Ambrosio et al., 2008; Gard et al., 2009; Haeusler et al., 2008; Iwasaki et al., 2010). Since the *T. brucei* Scc1 cohesin subunit is also enriched over tRNA genes (Muller et al., 2018) it is possible that the plethora of factors associated with RNAPIII-transcribed genes mediate the formation of nucleosome depleted boundary structures that facilitate transcription termination of polycistronic units by obstructing the passage of RNAPII. Since none of the eight proteins identified decorate RNAPII TSRs, they likely contribute to RNAPII transcription termination and/or facilitate RNAPIII transcription of tRNA and snRNA genes.

Interaction networks of RNAPIII/TTR-associated proteins

To gain more insight into the functional context of the eight proteins found to be enriched at the subset of TTRs coinciding with RNAPIII-transcribed genes, we applied the same approach used to identify interacting partners of the promoter-associated factors. Below we detail the interaction networks of the RNAPIII/TTR-enriched proteins and discuss their potential functional implications (Fig. 5E,F; Supplemental Table S4; Supplemental Table S5).

Affinity selection of the telomere binding protein TRF resulted in enrichment of previously identified TRF- and telomere-associated proteins (Reis et al., 2018). The interaction of HDAC3 with TRF and BDF2 and its enrichment over TSRs suggests that HDAC3 functions at both RNAPII promoters and at subtelomeric regions (Fig. 5E,F). Our proteomic analyses detected no significant interactions of ELP3b, PHD2 and PHD4 with other proteins (Fig.5E; Supplemental Table S4). Indeed, none of the eight RNAPIII/TTR-associated factors displayed reciprocal interactions with each other. This lack of crosstalk suggests that each protein performs distinct functions at these locations.

BDF7 and NAP proteins

BDF7 is a Bromodomain protein containing an AAA^+ ATPase domain, equivalent to Yta7, Abo1 and ATAD2 of budding and fission yeast, and vertebrates, respectively. These proteins have been implicated in altering nucleosome density to facilitate transcription, and Abo1 has recently been shown to mediate H3-H4 deposition onto DNA *in vitro* (Cho et al., 2019; Lombardi et al., 2011; Murawska and Ladurner, 2020). Affinity selected YFP-BDF7 showed strong association with three Nucleosome Assembly Proteins (Tb927.1.2210, designated NAP1; Tb927.3.4880, designated NAP2; and Tb927.10.15180, designated NAP3; Fig. 5E,F; Supplemental Table S4), supporting a potential role for BDF7 as a histone chaperone involved in nucleosome formation. H3.V and H4.V tend to be enriched at the end of trypanosome polycistronic units where RNAPII transcription is terminated (Siegel et al., 2009), and we find that BDF7 is associated with a subset of these TTRs (Fig. 5B). Thus, it is possible that BDF7 acts with NAP1-to-3 to mediate H3.V-H4.V deposition at *T. brucei* RNAPIII-associated transcription termination regions. Distinct nucleosome depleted regions are formed over tRNA genes and, as discussed above, the termination of some RNAPII transcription occurs in regions coincident with RNAPIII-transcribed genes (Marchetti et al., 1998; Maree and Patterton, 2014; Siegel et al., 2009). As suggested previously, it is possible that a subset of genes transcribed by RNAPIII act as boundaries that block the passage of advancing RNAPII into convergent or downstream transcription units (Siegel et al., 2009).

BDF7 may act with associated NAP proteins to promote nucleosome depletion and termination over such regions.

TFIIS2-2 and the PAF1 complex

TFIIS2-2 was clearly enriched in the vicinity of RNAPIII-transcribed genes and associated TTRs and upon affinity selection exhibited strong interaction with PAF1 Complex components (LEO1/Tb927.9.12900, CTR9/Tb927.3.3220, CDC73/Tb927.11.10230) and with several RNAPII subunits (Fig. 5E,F; Supplemental Table S4). Yeast Paf1C acts with TFIIS to enable transcription elongation through chromatin templates (Schier and Taatjes, 2020; Van Oss et al., 2017). The accumulation of PAF1 complex subunits at these TTR regions presumably reflects the role that these proteins are known to play in transcription termination and 3′ end processing of RNAPII transcripts.

DOT1A and the RNase H2 complex

The DOT1A and DOT1B histone methyltransferases direct trypanosome H3K76 di- and tri-methylation, respectively (Janzen et al., 2006b). DOT1A is involved in cell cycle progression while DOT1B is necessary for maintaining the silent state of inactive VSGs and for rapid transcriptional VSG switching (Figueiredo et al., 2008; Janzen et al., 2006b). Our proteomic analysis revealed that DOT1A associates with all three subunits of the RNase H2 complex (RH2A, RH2B and RH2C; Fig. 5E,F; Supplemental Table S4). The RH1 and RH2 complexes are necessary for resolving R-loops formed during transcription (Cerritelli and Crouch, 2009). While both RH1 and RH2 complexes are involved in antigenic variation, only RH2 has a role in trypanosome RNAPII transcription (Briggs et al., 2019). A recent study suggests that DOT1B is also required to clear R-loops by suppressing RNA-DNA hybrid formation and resulting DNA damage (Eisenhuth et al., 2020). DOT1A, which we find to be enriched at tRNA and snRNA genes (see above), may act with RH2 to prevent the accumulation of RNA-DNA hybrids at these RNAPIII-transcribed regions.

TBP, BRF1 and RNAPIII-transcribed genes

TBP (TATA-box related protein) has largely been studied with respect to its role in RNAPII transcription from spliced leader (SL) RNA gene promoters (Das et al., 2005). However, TBP was also previously shown to associate strongly with the TFIIIB component BRF1 (Schimanski et al., 2005) and thus, like BRF1, TBP may also promote RNAPIII transcription (Vélez-Ramírez et al., 2015). Indeed, directed ChIP assays indicate that TBP associates with specific RNAPIII-transcribed genes (Ruan et al., 2004; Vélez-Ramírez et al., 2015). Consistent with a dual role, we confirmed that YFP-TBP interacts with both SNAP complex

components involved in SL transcription and with BRF1 (Fig. 5E,F; Supplemental Table S4). Our ChIP-seq analysis shows that TBP is concentrated in sharp peaks that coincide with both RNAPII promoters for SL RNA genes as well as RNAPIII-transcribed tRNA and snRNA genes (Fig. 5A,D). Additionally, TBP was significantly enriched at arrays of RNAPIII-transcribed 5S rRNA genes (Fig. 5C). Thus, apart from SL RNA gene promoters, TBP appears to mark all known sites of RNAPIII-directed transcription.

Conclusions

Using protein domain homology searches, we identified a collection of 65 putative chromatin regulators in *Trypanosoma brucei* that were predicted to act as writers, readers or erasers of histone post-translational modifications. Many of these proteins exhibited a discernible nuclear localization and displayed distinct patterns of association across the genome, frequently coinciding with regions responsible for RNAPII transcription initiation and termination or RNAPIII transcription (Fig. 6). Robust proteomic analyses allowed the interaction networks of these proteins to be identified thereby providing further insight into their possible functions at specific genomic locations by revealing putative complexes likely involved in the distinct phases of transcription: initiation, elongation and termination. Counterparts of yeast SWR1 H2A-to-H2A.Z exchange complex and NuA4 HAT complex components were identified, some of which were enriched where H2A.Z is prevalent at TSRs. A recent report indicates that both TbSWR1 and TbNuA4 complexes are indeed involved in the deposition of *T. brucei* H2A.Z (Vellmer et al. 2021).

We also demonstrate that two predicted SET domain methyltransferases associate with putative histone modification reader proteins with which they occupy RNAPII TSRs. Moreover, our analyses reveal that six of the seven Bromodomain proteins are involved in four interaction networks enriched at TSRs while the BDF7 network alone marks a subset of TTRs that are coincident with sites of RNAPIII transcription. The association of SET and Bromodomain proteins with conserved RNAPII subunits, histone acetyltrasnferases, chromatin chaperones and remodelling factors suggests that the networks identified play pivotal roles in defining sites of transcription initiation and termination.

The data presented here provide a comprehensive depiction of the operational context of chromatin writers, readers and erasers at important genomic regulatory elements in this experimentally tractable but divergent eukaryote. Critically, our analyses identify many novel proteins unrelated to, or divergent from, known chromatin regulators of conventional eukaryotes. This highlights the utility of our approach to reveal novelty in the composition of

T. brucei chromatin regulatory complexes which differ from the paradigms established using conventional eukaryotic models.

Trypanosome gene expression is generally considered to be regulated post-transcriptionally with a plethora of factors dedicated to sculpting mature mRNAs from nascent polycistronic transcripts (Clayton, 2019). The complexity of chromatin regulatory factors that we have found to be enriched at TSRs and some TTRs may simply represent the core set of proteins required to mediate efficient and constitutive eukaryotic transcription initiation, elongation and termination in a chromatin context. An alternative possibility is that these proteins, with potentially antagonistic functions, operate in a more complex regulatory landscape where transcriptional control contributes alongside post-transcriptional mechanisms in ensuring optimal trypanosome gene expression.

Figure Legends

Fig. 1. Cellular localization of chromatin-associated *T. brucei* candidate proteins.
The indicated YFP-tagged proteins expressed in bloodstream Lister 427 cells from their endogenous genomic loci were detected with an anti-GFP primary antibody and an Alexa Fluor 568 labelled secondary antibody (red). Nuclear and kinetoplast (mitochondrial) DNA were stained with DAPI (green). Staining of untagged 427 parasites serves as a negative control. Representative images are shown for those candidate proteins that gave a specific ChIP-seq signal. The images are ordered according to ChIP-seq patterns shown in Fig. 2A and Fig. 5A. Images for all other tagged proteins are included in Supplemental Fig. S2.
Scale bar = 5 μm.

Fig. 2. ChIP-seq reveals two classes of proteins at *T. brucei* transcription start regions.
A. A region of Chromosome 7 (coordinates as indicated, kb) is shown with ChIP-seq reads mapped for the indicated proteins. A single replicate is shown for each protein. Tracks are scaled separately as reads per million (values shown at the end of each track). The kinetochore protein KKT2 is included as a positive control and is enriched at centromeric regions. ChIP-seq performed in Lister 427 cells expressing no tagged protein (untagged) provides a negative control. ChIP-seq profiles for the different proteins are ordered

according to their patterns. Previous H2A.Z data (Wedel et al. 2017) and our RPB1 ChIP-seq allowed Transcription Start Regions (TSRs) to be identified. No input data was available for normalization of H2A.Z reads resulting in a higher read scale. The position of bidirectional/divergent and unidirectional/single TSRs is indicated with arrows showing the direction of transcription. The position of convergent and single Transcription Termination Regions (TTRs) is shown with arrows indicating the direction from which transcription is halted. The position of tRNA genes (blue bars) and the Chromosome 7 centromere (CEN, grey oval) are marked. Positions of the various genomic elements (TSRs, TTRs, tRNAs, CEN) were obtained from annotations of the Lister 427 genome (Muller et al, 2018).

B. Most TSRs annotated in the Lister 427 genome overlap with YFP-CRD1 ChIP-seq peaks.

C. Enrichment profiles of Class I TSR factors. The metagene plots (top) show normalized read density around all CRD1 peak summits, with individual replicates for each protein shown separately. Note the different scale for CRD1. The heatmaps (bottom) are an average of all replicates for each protein and show protein density around individual CRD1 peaks. Scale bars represent reads that were normalized to input and library size.

D. As in (C) for Class II TSR factors.

Fig. 3. Enrichment of Class II proteins at TSRs follows the direction of RNAPII transcription.

A. SET26 is used as a representative protein of Class II TSR-associated factors. Comparison of SET26 ChIP-seq data with strand specific RNA-seq data (Naguleswaran et al., 2018) shows that SET26 reads are enriched in the same direction as RNAPII transcript reads. Heatmaps show from top to bottom: minus strand reads from unidirectional TSRs (top), plus and minus strand reads from bidirectional TSRs or from two adjacent unidirectional TSRs (middle), and plus strand reads from unidirectional TSRs (bottom).

B. Examples from the different heatmap regions described in (A). A single replicate is shown for each protein. Tracks are scaled separately as reads per million (values shown at the end of each track).

Fig. 4. Class I and Class II TSR-associated factors define distinct interaction networks.

A. YFP-tagged proteins found to be enriched at TSRs were analysed by LC-MS/MS to identify their protein interactions. The data for each plot is based on three biological replicates. Cut-offs used for significance: $\log_2$ (tagged/untagged) > 2 or < -2 and $p < 0.01$ (Student's *t*-test). Enrichment scores for proteins identified in each affinity selection are presented in Supplemental Table S4. Plots in the same box show reciprocal interactions.

Proteins of interest are indicated by red font. Uncharacterized proteins common to several affinity selections are depicted in yellow, green, purple and cyan. Black squares represent members of the SWR1/SRCAP complex found in the BDF2 and ZCW1 affinity selections.
B. Key proteins identified as being associated with the indicated YFP-tagged bait proteins (thick oval outlines). Rectangles contain proteins common to several affinity purifications. Lines denote interactions between proteins. The interactions of BDF3 and BDF5 with BDF4 (dashed lines); BDF5 with ZCW1 (dashed lines); and BDF7 (grey) with ZCW1 and SET26 were not confirmed by reciprocal affinity selections.

Fig. 5. Proteins enriched over RNAPII TTRs coinciding with RNAPIII-transcribed genes define distinct interaction networks.

The ChIP-seq tracks show a single replicate for each protein and are scaled separately as reads per million (values shown at the end of each track).

A. Examples of protein enrichment over an snRNA (top panel) and tRNAs (bottom panel).

B. Overlap between TTRs, tRNAs, snRNAs and RNAPIII/TTR-associated factors. The numbers under the horizontal bar graphs refer to the number of tRNA or snRNA genes overlapping with each TTR. Presence and absence of overlap with RNAPIII/TTR-associated factors is indicated by green and empty squares, respectively. cTTRs – convergent TTRs, sTTRs – single TTRs.

C. Enrichment of the RNAPIII/TTR-associated factors and RPC1 at the 5S rRNA gene cluster.

D. Enrichment of the RNAPIII/TTR-associated factors and RPC1 at the spliced leader gene cluster.

E. YFP-tagged proteins found to be enriched at a subset of TTRs were analyses by LC-MS/MS to identify their protein interactions. The data for each plot is based on three biological replicates. Cut-offs used for significance: $\log_2$ (tagged/untagged) > 2 or < -2 and $p < 0.01$ (Student's *t*-test). Enrichment scores for proteins identified in each affinity selection are presented in Supplemental Table S4. Significantly enriched proteins are indicated by black or coloured dots. Proteins of interest are indicated by red font.

F. Key proteins identified as being associated with the indicated YFP-tagged bait proteins (thick oval outlines). Lines denote interactions between proteins.

Fig. 6. Model depicting distribution of chromatin regulators across a trypanosome polycistronic transcription unit.

Diagram shows the five Class I (sharp; green) and ten Class II (broad; purple) TSR-associated factors at a unidirectional/single RNAPII promoter. The arrow indicates the direction of transcription. The grey rectangle represents a single polycistron. Class II proteins are enriched in the direction of transcription. The eight proteins found at RNAPII TTRs associated with RNAPIII-transcribed genes are shown in blue. tRNA and snRNA genes are represented by the brown rectangle. Proteins within each box were found to interact in the proteomic experiments.

Materials and Methods

Cell culture

Lister 427 bloodstream form *T. brucei* was used throughout this study. Parasites were grown in HMI-9 medium (Hirumi and Hirumi, 1989) at 37^{o}C and 5% CO_2. Cell lines with YFP-tagged proteins were grown in the presence of 5 µg/ml blasticidin. The density of cell cultures was maintained below 3 x 10^6 cells/ml.

Protein tagging

Candidate proteins were tagged endogenously on their N termini with YFP using the pPOTv4 plasmid (Dean et al., 2015). Tagging constructs were produced by fusion PCR of three fragments: a ~500 bp fragment homologous to the end of the 5′ UTR of each gene, a region of the pPOTv4 plasmid containing a blasticidin resistance cassette and a YFP tag, and a ~500 bp fragment homologous to the beginning of the coding sequence of each gene. Table S9 lists the primers used for tagging. Fusion constructs were transfected into bloodstream form parasites by electroporation as previously described (Burkard et al., 2007). The cell lines obtained after blasticidin selection were tested for correct integration of the tagging constructs by PCR and for expression of the tagged proteins via western blotting analysis.

Fluorescent immunolocalization

Cells were fixed with 4% paraformaldehyde for 10 min on ice. Fixation was stopped with 0.1 M glycine. Cells were added to polylysine-coated slides and permeabilised with 0.1% Triton X-100. The slides were blocked with 2% BSA. Rabbit anti-GFP primary antibody (A-11122; Thermo Fisher Scientific) was used at 1:500 dilution and secondary Alexa fluor 568 goat anti-rabbit antibody (A-11036; Thermo Fisher Scientific) was used at 1:1000 dilution. Images were taken with a Zeiss Axio Imager microscope.

Chromatin immunoprecipitation and sequencing (ChIP-seq)

4 x 10^8 parasites were fixed with 0.8% formaldehyde for 20 min at room temperature. Cells were lysed and sonicated in the presence of 0.2% SDS for 30 cycles (30 s on, 30 s off) using the high setting on a Bioruptor sonicator (Diagenode). Cell debris were pelleted by centrifugation and SDS in the lysate supernatants was diluted to 0.07%. Input samples were taken before incubating the rest of the cell lysates overnight with 10 µg rabbit anti-GFP antibody (A-11122; Thermo Fisher Scientific) and Protein G Dynabeads. The beads were washed, and the DNA eluted from them was treated with RNase and Proteinase K. DNA was then purified using a QIAquick PCR Purification Kit (Qiagen) and libraries were

prepared using NEXTflex barcoded adapters (Bioo Scientific). The libraries were sequenced on Illumina HiSeq 4000 (Edinburgh Genomics), Illumina NextSeq (Western General Hospital, Edinburgh) or Illumina MiniSeq (Allshire lab). In all cases, 75 bp paired-end sequencing was performed. Our subsequent analyses were based on three replicates for BDF6, CRD1 and DMT; one replicate for ELP3a, NAT2 and SET10; and two replicates for the remaining YFP-tagged proteins and for the untagged control.

ChIP-seq data analysis

Sequencing reads were de-duplicated with pyFastqDuplicateRemover (Webb et al., 2018; https://git.ecdf.ed.ac.uk/sgrannem/pycrac) and subsequently aligned to the Tb427v9.2 genome (Muller et al., 2018) with Bowtie 2 (Langmead and Salzberg, 2012). The default mode of Bowtie 2 was used which searches for multiple alignments and reports the best one or, if several alignments are deemed equally good, reports one of those randomly. The ChIP samples were normalized to their respective inputs (ratio of ChIP to input reads) and to library size (reads per million). TSR and TTR regions were defined based on annotations of the Lister 427 genome (Muller et al., 2018). CRD1 peak summits were called using the narrow peak mode of MACS2 (Feng et al., 2012) followed by manual filtering of false positives which included peaks not present in all CRD1 replicates, peaks present in the untagged control and/or peaks with fold enrichment < 6.5. 10 kb regions upstream and downstream of CRD1 peak summits were divided into 50 bp windows. The metagene plots display individual ChIP-seq replicates separately and were generated by summing normalized reads in each 50 bp window and representing them as density centered around CRD1. The average metagene plots were generated analogously, except that the reads around all CRD1 peaks were averaged before plotting. Heatmaps represent normalized reads around individual CRD1 peaks and were generated as an average of all replicates for each protein.

Affinity purification and LC-MS/MS proteomic analysis

4 x 10^8 cells were lysed per IP in the presence of 0.2% NP-40 and 150 mM KCl. Lysates were sonicated briefly (3 cycles, 12 s on, 12 s off) at a high setting in a Bioruptor (Diagenode) sonicator. The soluble and insoluble fractions were separated by centrifugation, and the soluble fraction was incubated for 1 h at 4^{o}C with beads crosslinked to mouse anti-GFP antibody (11814460001; Roche). Resulting immunoprecipitates were washed three times with lysis buffer and protein was eluted with RapiGest surfactant (Waters) at 55°C for 15 min. Next, filter-aided sample preparation (FASP) (Wiśniewski et al., 2009) was used to digest the protein samples for mass spectrometric analysis. Briefly, proteins were reduced with DTT and then denatured with 8 M Urea in Vivakon spin (filter) column 30K cartridges.

Samples were alkylated with 0.05 M IAA and digested with 0.5 µg MS Grade Pierce Trypsin Protease (Thermo Fisher Scientific) overnight, desalted using stage tips (Rappsilber et al., 2007) and resuspended in 0.1%TFA for LC-MS/MS. Peptides were separated using RSLC Ultimate3000 system (Thermo Fisher Scientific) fitted with an EasySpray column (50 cm; Thermo Fisher Scientific) utilising 2-40-95% non-linear gradients with solvent A (0.1% formic acid) and solvent B (80% acetonitrile in 0.1% formic acid). The EasySpray column was directly coupled to an Orbitrap Fusion Lumos (Thermo Fisher Scientific) operated in DDA mode. "TopSpeed" mode was used with 3 s cycles with standard settings to maximize identification rates: MS1 scan range - 350-1500 mz, RF lens 30%, AGC target $4.0e^5$ with intensity threshold $5.0e^3$, filling time 50 ms and resolution 60000, monoisotopic precursor selection and filter for charge states 2-5. HCD (27%) was selected as fragmentation mode. MS2 scans were performed using Ion Trap mass analyzer operated in rapid mode with AGC set to $2.0e^4$ and filling time to 50 ms. The resulting shot-gun data were processed using Maxquant 1.6.1.0 (*T. brucei* proteome from 14 May 2019) and visualized using Perseus 1.6.1.3 (Tyanova et al., 2016).

Data Access

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE150253.

The proteomics data generated in this study have been submitted to the Proteomics Identifications Database (PRIDE; https://www.ebi.ac.uk/pride/) under accession number PXD026743.

Acknowledgements

We thank members of both the Allshire and Matthews labs for useful input and discussion. In particular, we thank Manu Shukla and Sito Torres-Garcia for advice on ChIP-seq; Alison Pidoux and Eleanor Silvester for advice on immunolocalization and microscopy; and Julie Young for supplying HMI-9 media. We would also like to thank Stefan Bresson who provided valuable input on the bioinformatic analysis, Samuel Dean who sent us the pPOTv4 plasmid used for YFP tagging and Esteban Serra for discussion on acetyltransferase nomenclature. We also thank Edinburgh Genomics (NERC, R8/H10/56; MRC, MR/K001744/1; BBSRC, BB/J004243/1) and Genetics Core, Edinburgh Clinical Research Facility for their valued sequencing services. This research was made possible by core funding to the Wellcome Centre for Cell Biology (203149); a Wellcome 4-year PhD in Cell Biology studentship (102336) to DPS; a Wellcome Senior Research Fellowship (103139) and Wellcome

Instrument grant (108504) to JR; a Wellcome Senior Research Fellowship (202811) to AAJ; a Wellcome Investigator award (103740) to KRM; a Wellcome Trust Principal Research Fellowship (095021 and 200885) to RCA; and an MRC Research Grant (MR/T04702X/1) to RCA and KRM.

Disclosure Declaration

The authors have no conflicts of interest or competing interests to declare.

References

Aasland R, Gibson TJ, Stewart AF. 1995. The PHD finger: implications for chromatin-mediated transcriptional regulation. *Trends Biochem Sci* **20**: 56-59.

Akiyoshi B, Gull K. 2013. Evolutionary cell biology of chromosome segregation: insights from trypanosomes. *Open Biol* **3**: 130023.

Akiyoshi B, Gull K. 2014. Discovery of unconventional kinetochores in kinetoplastids. *Cell* **156**: 1247-1258.

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

Alsford S, Horn D. 2011. Elongator protein 3b negatively regulates ribosomal DNA transcription in African trypanosomes. *Mol Cell Biol* **31**: 1822-1832.

Alsford S, Horn D. 2012. Cell-cycle-regulated control of VSG expression site silencing by histones and histone chaperones ASF1A and CAF-1b in Trypanosoma brucei. *Nucleic Acids Res* **40**: 10150-10160.

Alsford S, Kawahara T, Isamah C, Horn D. 2007. A sirtuin in the African trypanosome is involved in both DNA repair and telomeric gene silencing but is not required for antigenic variation. *Mol Microbiol* **63**: 724-736.

Alsford S, Turner DJ, Obado SO, Sanchez-Flores A, Glover L, Berriman M, Hertz-Fowler C, Horn D. 2011. High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome Res* **21**: 915-924.

Arrowsmith CH, Schapira M. 2019. Targeting non-bromodomain chromatin readers. *Nat Struct Mol Biol* **26**: 863-869.

Bannister AJ, Kouzarides T. 2011. Regulation of chromatin by histone modifications. *Cell Res* **21**: 381-395.

Baron DM, Ralston KS, Kabututu ZP, Hill KL. 2007. Functional genomics in Trypanosoma brucei identifies evolutionarily conserved components of motile flagella. *J Cell Sci* **120**: 478-491.

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

Benabdallah NS, Williamson I, Illingworth RS, Kane L, Boyle S, Sengupta D, Grimes GR, Therizols P, Bickmore WA. 2019. Decreased Enhancer-Promoter Proximity Accompanying Enhancer Activation. *Mol Cell* **76**: 473-484 e477.

Berriman M Ghedin E Hertz-Fowler C Blandin G Renauld H Bartholomeu DC Lennard NJ Caler E Hamlin NE Haas B et al. 2005. The genome of the African trypanosome Trypanosoma brucei. *Science* **309**: 416-422.

Bonisch C, Hake SB. 2012. Histone H2A variants in nucleosomes and chromatin: more or less stable? *Nucleic Acids Res* **40**: 10719-10741.

Borst P. 1986. Discontinuous transcription and antigenic variation in trypanosomes. *Annu Rev Biochem* **55**: 701-732.

Borst P, Sabatini R. 2008. Base J: discovery, biosynthesis, and possible functions. *Annu Rev Microbiol* **62**: 235-251.

Branche C, Kohl L, Toutirais G, Buisson J, Cosson J, Bastin P. 2006. Conserved and specific functions of axoneme components in trypanosome motility. *J Cell Sci* **119**: 3443-3455.

Briggs E, Crouch K, Lemgruber L, Hamilton G, Lapsley C, McCulloch R. 2019. Trypanosoma brucei ribonuclease H2A is an essential R-loop processing enzyme whose loss causes DNA damage during transcription initiation and antigenic variation. *Nucleic Acids Res* **47**: 9180-9197.

Burkard G, Fragoso CM, Roditi I. 2007. Highly efficient stable transformation of bloodstream forms of Trypanosoma brucei. *Molecular and Biochemical Parasitology* **153**: 220-223.

Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, Shia WJ, Anderson S, Yates J, Washburn MP et al. 2005. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**: 581-592.

Cerritelli SM, Crouch RJ. 2009. Ribonuclease H: the enzymes in eukaryotes. *FEBS J* **276**: 1494-1505.

Chen M, Gartenberg MR. 2014. Coordination of tRNA transcription with export at nuclear pore complexes in budding yeast. *Genes Dev* **28**: 959-970.

Cheung V, Chua G, Batada NN, Landry CR, Michnick SW, Hughes TR, Winston F. 2008. Chromatin- and transcription-related factors repress transcription from within coding regions throughout the Saccharomyces cerevisiae genome. *PLoS Biol* **6**: e277.

Cho C, Jang J, Kang Y, Watanabe H, Uchihashi T, Kim SJ, Kato K, Lee JY, Song JJ. 2019. Structural basis of nucleosome assembly by the Abo1 AAA+ ATPase histone chaperone. *Nat Commun* **10**: 5764.

Clayton C. 2019. Regulation of gene expression in trypanosomatids: living with polycistronic transcription. *Open Biol* **9**: 190072.

Cliffe LJ, Siegel TN, Marshall M, Cross GA, Sabatini R. 2010. Two thymidine hydroxylases differentially regulate the formation of glucosylated DNA at regions flanking polymerase II polycistronic transcription units throughout the genome of Trypanosoma brucei. *Nucleic Acids Res* **38**: 3923-3935.

D'Ambrosio C, Schmidt CK, Katou Y, Kelly G, Itoh T, Shirahige K, Uhlmann F. 2008. Identification of cis-acting sites for condensin loading onto budding yeast chromosomes. *Genes Dev* **22**: 2215-2227.

D'Archivio S, Wickstead B. 2017. Trypanosome outer kinetochore proteins suggest conservation of chromosome segregation machinery across eukaryotes. *J Cell Biol* **216**: 379-391.

Das A, Zhang Q, Palenchar JB, Chatterjee B, Cross GA, Bellofatto V. 2005. Trypanosomal TBP functions with the multisubunit transcription factor tSNAP to direct spliced-leader RNA gene expression. *Mol Cell Biol* **25**: 7314-7322.

de Jesus TC, Nunes VS, Lopes Mde C, Martil DE, Iwai LK, Moretti NS, Machado FC, de Lima-Stein ML, Thiemann OH, Elias MC et al. 2016. Chromatin Proteomics Reveals Variable Histone Modifications during the Life Cycle of Trypanosoma cruzi. *J Proteome Res* **15**: 2039-2051.

de Lima LP, Poubel SB, Yuan ZF, Roson JN, Vitorino FNL, Holetz FB, Garcia BA, da Cunha JPC. 2020. Improvements on the quantitative analysis of Trypanosoma cruzi histone post translational modifications: Study of changes in epigenetic marks through the parasite's metacyclogenesis and life cycle. *J Proteomics* **225**: 103847.

Dean S, Moreira-Leite F, Varga V, Gull K. 2016. Cilium transition zone proteome reveals compartmentalization and differential dynamics of ciliopathy complexes. *Proc Natl Acad Sci U S A* **113**: E5135-5143.

Dean S, Sunter J, Wheeler RJ, Hodkinson I, Gluenz E, Gull K. 2015. A toolkit enabling efficient, scalable and reproducible gene tagging in trypanosomatids. *Open Biol* **5**: 140197.

Dean S, Sunter JD, Wheeler RJ. 2017. TrypTag.org: A Trypanosome Genome-wide Protein Localisation Resource. *Trends Parasitol* **33**: 80-82.

DeGrasse JA, DuBois KN, Devos D, Siegel TN, Sali A, Field MC, Rout MP, Chait BT. 2009. Evidence for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic ancestor. *Mol Cell Proteomics* **8**: 2119-2130.

Denninger V, Rudenko G. 2014. FACT plays a major role in histone dynamics affecting VSG expression site control in Trypanosoma brucei. *Mol Microbiol* **94**: 945-962.

Dillon SC, Zhang X, Trievel RC, Cheng X. 2005. The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol* **6**: 227.

Doyon Y, Cote J. 2004. The highly conserved and multifunctional NuA4 HAT complex. *Curr Opin Genet Dev* **14**: 147-154.

DuBois KN, Alsford S, Holden JM, Buisson J, Swiderski M, Bart JM, Ratushny AV, Wan Y, Bastin P, Barry JD et al. 2012. NUP-1 Is a large coiled-coil nucleoskeletal protein in trypanosomes with lamin-like functions. *PLoS Biol* **10**: e1001287.

Eisenhuth N, Vellmer T, Butter F, Janzen CJ. 2020. A DOT1B/Ribonuclease H2 protein complex is involved in R-loop processing, genomic integrity and antigenic variation in Trypanosoma brucei. bioRxiv doi: 10.1101/2020.03.02.969337.

ElBashir R, Vanselow JT, Kraus A, Janzen CJ, Siegel TN, Schlosser A. 2015. Fragment ion patchwork quantification for measuring site-specific acetylation degrees. *Anal Chem* **87**: 9939-9945.

Feng J, Liu T, Qin B, Zhang Y, Liu XS. 2012. Identifying ChIP-seq enrichment using MACS. *Nature Protocols* **7**: 1728-1740.

Feng Q, Wang H, Ng HH, Erdjument-Bromage H, Tempst P, Struhl K, Zhang Y. 2002. Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. *Curr Biol* **12**: 1052-1058.

Figueiredo LM, Cross GA, Janzen CJ. 2009. Epigenetic regulation in African trypanosomes: a new kid on the block. *Nat Rev Microbiol* **7**: 504-513.

Figueiredo LM, Janzen CJ, Cross GA. 2008. A histone methyltransferase modulates antigenic variation in African trypanosomes. *PLoS Biol* **6**: e161.

Fritz M, Vanselow J, Sauer N, Lamer S, Goos C, Siegel TN, Subota I, Schlosser A, Carrington M, Kramer S. 2015. Novel insights into RNP granules by employing the trypanosome's microtubule skeleton as a molecular sieve. *Nucleic Acids Res* **43**: 8013-8032.

Gard S, Light W, Xiong B, Bose T, McNairn AJ, Harris B, Fleharty B, Seidel C, Brickner JH, Gerton JL. 2009. Cohesinopathy mutations disrupt the subnuclear organization of chromatin. *J Cell Biol* **187**: 455-462.

Gheiratmand L, Brasseur A, Zhou Q, He CY. 2013. Biochemical characterization of the bi-lobe reveals a continuous structural network linking the bi-lobe to other single-copied organelles in Trypanosoma brucei. *J Biol Chem* **288**: 3489-3499.

Glover L, Hutchinson S, Alsford S, Horn D. 2016. VEX1 controls the allelic exclusion required for antigenic variation in trypanosomes. *Proc Natl Acad Sci U S A* **113**: 7225-7230.

Goos C, Dejung M, Janzen CJ, Butter F, Kramer S. 2017. The nuclear proteome of Trypanosoma brucei. *PLoS One* **12**: e0181884.

Grozinger CM, Schreiber SL. 2002. Deacetylase enzymes: biological functions and the use of small-molecule inhibitors. *Chem Biol* **9**: 3-16.

Gunzl A. 2010. The pre-mRNA splicing machinery of trypanosomes: complex or simplified? *Eukaryot Cell* **9**: 1159-1170.

Guther ML, Urbaniak MD, Tavendale A, Prescott A, Ferguson MA. 2014. High-confidence glycosome proteome for procyclic form Trypanosoma brucei by epitope-tag organelle enrichment and SILAC proteomics. *J Proteome Res* **13**: 2796-2806.

Haeusler RA, Pratt-Hyatt M, Good PD, Gipson TA, Engelke DR. 2008. Clustering of yeast tRNA genes is mediated by specific association of condensin with tRNA gene transcription complexes. *Genes Dev* **22**: 2204-2214.

Halliday C, Billington K, Wang Z, Madden R, Dean S, Sunter JD, Wheeler RJ. 2019. Cellular landmarks of Trypanosoma brucei and Leishmania mexicana. *Mol Biochem Parasitol* **230**: 24-36.

Haynes SR, Dollard C, Winston F, Beck S, Trowsdale J, Dawid IB. 1992. The bromodomain: a conserved sequence found in human, Drosophila and yeast proteins. *Nucleic Acids Res* **20**: 2603.

Henikoff S, Smith MM. 2015. Histone variants and epigenetics. *Cold Spring Harb Perspect Biol* **7**: a019364.

Hirumi H, Hirumi K. 1989. Continuous cultivation of Trypanosoma brucei blood stream forms in a medium containing a low concentration of serum protein without feeder cell layers. *The Journal of parasitology*: 985-989.

Hishikawa D, Shindou H, Kobayashi S, Nakanishi H, Taguchi R, Shimizu T. 2008. Discovery of a lysophospholipid acyltransferase family essential for membrane asymmetry and diversity. *Proc Natl Acad Sci U S A* **105**: 2830-2835.

Horn D. 2014. Antigenic variation in African trypanosomes. *Mol Biochem Parasitol* **195**: 123-129.

Ingram AK, Horn D. 2002. Histone deacetylases in Trypanosoma brucei: two are essential and another is required for normal cell cycle progression. *Mol Microbiol* **45**: 89-97.

Iwasaki O, Tanaka A, Tanizawa H, Grewal SI, Noma K. 2010. Centromeric localization of dispersed Pol III genes in fission yeast. *Mol Biol Cell* **21**: 254-265.

Janzen CJ, Fernandez JP, Deng H, Diaz R, Hake SB, Cross GA. 2006a. Unusual histone modifications in Trypanosoma brucei. *FEBS Lett* **580**: 2306-2310.

Janzen CJ, Hake SB, Lowell JE, Cross GA. 2006b. Selective di- or trimethylation of histone H3 lysine 76 by two DOT1 homologs is important for cell cycle regulation in Trypanosoma brucei. *Mol Cell* **23**: 497-507.

Jehi SE, Li X, Sandhu R, Ye F, Benmerzouga I, Zhang M, Zhao Y, Li B. 2014. Suppression of subtelomeric VSG switching by Trypanosoma brucei TRF requires its TTAGGG repeat-binding activity. *Nucleic Acids Res* **42**: 12899-12911.

Jonkers I, Lis JT. 2015. Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16**: 167-177.

Joshi AA, Struhl K. 2005. Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Mol Cell* **20**: 971-978.

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

Kawahara T, Siegel TN, Ingram AK, Alsford S, Cross GA, Horn D. 2008. Two essential MYST-family proteins display distinct roles in histone H4K10 acetylation and telomeric silencing in trypanosomes. *Mol Microbiol* **69**: 1054-1068.

Keogh MC, Kurdistani SK, Morris SA, Ahn SH, Podolny V, Collins SR, Schuldiner M, Chin K, Punna T, Thompson NJ et al. 2005. Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* **123**: 593-605.

Kloc A, Martienssen R. 2008. RNAi, heterochromatin and the cell cycle. *Trends Genet* **24**: 511-517.

Klose RJ, Kallin EM, Zhang Y. 2006. JmjC-domain-containing proteins and histone demethylation. *Nat Rev Genet* **7**: 715-727.

Kouzarides T. 2007. Chromatin Modifications and Their Function. *Cell* **128**: 693-705.

Kovarova J, Horakova E, Changmai P, Vancova M, Lukes J. 2014. Mitochondrial and nucleolar localization of cysteine desulfurase Nfs and the scaffold protein Isu in Trypanosoma brucei. *Eukaryot Cell* **13**: 353-362.

Kramer S, Bannerman-Chukualim B, Ellis L, Boulden EA, Kelly S, Field MC, Carrington M. 2013. Differential localization of the two T. brucei poly(A) binding proteins to the nucleus and RNP granules suggests binding to distinct mRNA pools. *PLoS One* **8**: e54004.

Kraus AJ, Vanselow JT, Lamer S, Brink BG, Schlosser A, Siegel TN. 2020. Distinct roles for H4 and H2A.Z acetylation in RNA transcription in African trypanosomes. *Nat Commun* **11**: 1498.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357-359.

Lee KK, Workman JL. 2007. Histone acetyltransferase complexes: one size doesn't fit all. *Nat Rev Mol Cell Biol* **8**: 284-295.

Lemaître C, Bickmore WA. 2015. Chromatin at the nuclear periphery and the regulation of genome functions. *Histochem Cell Biol* **144**: 111-122.

Levi J, Malachi T, Djaldetti M, Bogin E. 1987. Biochemical changes associated with the osmotic fragility of young and mature erythrocytes caused by parathyroid hormone in relation to the uremic syndrome. *Clin Biochem* **20**: 121-125.

Li B, Espinal A, Cross GA. 2005. Trypanosome telomeres are protected by a homologue of mammalian TRF2. *Mol Cell Biol* **25**: 5011-5021.

Li B, Gogol M, Carey M, Pattenden SG, Seidel C, Workman JL. 2007. Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription. *Genes Dev* **21**: 1422-1430.

Lombardi LM, Ellahi A, Rine J. 2011. Direct regulation of nucleosome density by the conserved AAA-ATPase Yta7. *Proc Natl Acad Sci U S A* **108**: E1302-1311.

Lowell JE, Cross GA. 2004. A variant histone H3 is enriched at telomeres in Trypanosoma brucei. *J Cell Sci* **117**: 5937-5947.

Lowell JE, Kaiser F, Janzen CJ, Cross GA. 2005. Histone H2AZ dimerizes with a novel variant H2B and is enriched at repetitive DNA in Trypanosoma brucei. *J Cell Sci* **118**: 5721-5730.

Mandava V, Fernandez JP, Deng H, Janzen CJ, Hake SB, Cross GA. 2007. Histone modifications in Trypanosoma brucei. *Mol Biochem Parasitol* **156**: 41-50.

Marchetti MA, Tschudi C, Silva E, Ullu E. 1998. Physical and transcriptional analysis of the Trypanosoma brucei genome reveals a typical eukaryotic arrangement with close

interspersionof RNA polymerase II- and III-transcribed genes. *Nucleic Acids Res* **26**: 3591-3598.

Maree JP, Patterton HG. 2014. The epigenome of Trypanosoma brucei: a regulatory interface to an unconventional transcriptional machine. *Biochim Biophys Acta* **1839**: 743-750.

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

Matangkasombut O, Buratowski RM, Swilling NW, Buratowski S. 2000. Bromodomain factor 1 corresponds to a missing piece of yeast TFIID. *Genes Dev* **14**: 951-962.

Matthews KR. 2005. The developmental cell biology of Trypanosoma brucei. *J Cell Sci* **118**: 283-290.

Mi J, Yang X, Zhang J, Zhang X, Xu C, Liao S, Tu X. 2018. Crystal structure of an ENT domain from Trypanosoma brucei. *Biochem Biophys Res Commun* **505**: 755-760.

Militello KT, Wang P, Jayakar SK, Pietrasik RL, Dupont CD, Dodd K, King AM, Valenti PR. 2008. African trypanosomes contain 5-methylcytosine in nuclear DNA. *Eukaryot Cell* **7**: 2012-2016.

Millar CB, Xu F, Zhang K, Grunstein M. 2006. Acetylation of H2AZ Lys 14 is associated with genome-wide gene activity in yeast. *Genes Dev* **20**: 711-722.

Mizuguchi G, Shen X, Landry J, Wu WH, Sen S, Wu C. 2004. ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* **303**: 343-348.

Mony BM, MacGregor P, Ivens A, Rojas F, Cowton A, Young J, Horn D, Matthews K. 2014. Genome-wide dissection of the quorum sensing signalling pathway in Trypanosoma brucei. *Nature* **505**: 681-685.

Muller LSM, Cosentino RO, Forstner KU, Guizetti J, Wedel C, Kaplan N, Janzen CJ, Arampatzi P, Vogel J, Steinbiss S et al. 2018. Genome organization and DNA accessibility control antigenic variation in trypanosomes. *Nature* **563**: 121-125.

Murawska M, Ladurner AG. 2020. Bromodomain AAA+ ATPases get into shape. *Nucleus* **11**: 32-34.

Naguleswaran A, Doiron N, Roditi I. 2018. RNA-Seq analysis validates the use of culture-derived Trypanosoma brucei and provides new markers for mammalian and insect life-cycle stages. *BMC Genomics* **19**: 227.

Nakaar V, Gunzl A, Ullu E, Tschudi C. 1997. Structure of the Trypanosoma brucei U6 snRNA gene promoter. *Mol Biochem Parasitol* **88**: 13-23.

Nanavaty V, Sandhu R, Jehi SE, Pandya UM, Li B. 2017. Trypanosoma brucei RAP1 maintains telomere and subtelomere integrity by suppressing TERRA and telomeric RNA:DNA hybrids. *Nucleic Acids Res* **45**: 5785-5796.

Ohshima S, Ohashi-Suzuki M, Miura Y, Yabu Y, Okada N, Ohta N, Suzuki T. 2010. TbUNC119 and its binding protein complex are essential for propagation, motility, and morphogenesis of Trypanosoma brucei procyclic form cells. *PLoS One* **5**: e15577.

Pandey M, Huang Y, Lim TK, Lin Q, He CY. 2020. Flagellar targeting of an arginine kinase requires a conserved lipidated protein intraflagellar transport (LIFT) pathway in Trypanosoma brucei. *J Biol Chem* **295**: 11326-11336.

Paro R. 1990. Imprinting a determined state into the chromatin of Drosophila. *Trends Genet* **6**: 416-421.

Perry J, Zhao Y. 2003. The CW domain, a structural module shared amongst vertebrates, vertebrate-infecting parasites and higher plants. *Trends Biochem Sci* **28**: 576-580.

Ponting CP. 1997. Tudor domains in proteins that interact with RNA. *Trends Biochem Sci* **22**: 51-52.

Ralston KS, Lerner AG, Diener DR, Hill KL. 2006. Flagellar motility contributes to cytokinesis in Trypanosoma brucei and is modulated by an evolutionarily conserved dynein regulatory system. *Eukaryot Cell* **5**: 696-711.

Rappsilber J, Mann M, Ishihama Y. 2007. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols* **2**: 1896-1906.

Reis H, Schwebs M, Dietz S, Janzen CJ, Butter F. 2018. TelAP1 links telomere complexes with developmental expression site silencing in African trypanosomes. *Nucleic Acids Res* **46**: 2820-2833.

Reynolds D, Hofmeister BT, Cliffe L, Alabady M, Siegel TN, Schmitz RJ, Sabatini R. 2016. Histone H3 Variant Regulates RNA Polymerase II Transcription Termination and Dual Strand Transcription of siRNA Loci in Trypanosoma brucei. *PLoS Genet* **12**: e1005758.

Roditi I, Liniger M. 2002. Dressed for success: the surface coats of insect-borne protozoan parasites. *Trends Microbiol* **10**: 128-134.

Roth SY, Denu JM, Allis CD. 2001. Histone acetyltransferases. *Annu Rev Biochem* **70**: 81-120.

Ruan JP, Arhin GK, Ullu E, Tschudi C. 2004. Functional characterization of a Trypanosoma brucei TATA-binding protein-related factor points to a universal regulator of transcription in trypanosomes. *Mol Cell Biol* **24**: 9610-9618.

Ruhl DD, Jin J, Cai Y, Swanson S, Florens L, Washburn MP, Conaway RC, Conaway JW, Chrivia JC. 2006. Purification of a human SRCAP complex that remodels chromatin by incorporating the histone variant H2A.Z into nucleosomes. *Biochemistry* **45**: 5671-5677.

Sawatsubashi S, Maki A, Ito S, Shirode Y, Suzuki E, Zhao Y, Yamagata K, Kouzmenko A, Takeyama K, Kato S. 2004. Ecdysone receptor-dependent gene regulation mediates histone poly(ADP-ribosyl)ation. *Biochem Biophys Res Commun* **320**: 268-272.

Scacchetti A, Becker PB. 2020. Variation on a theme: Evolutionary strategies for H2A.Z exchange by SWR1-type remodelers. *Curr Opin Cell Biol* **70**: 1-9.

Schier AC, Taatjes DJ. 2020. Structure and mechanism of the RNA polymerase II transcription machinery. *Genes Dev* **34**: 465-488.

Schimanski B, Nguyen TN, Gunzl A. 2005. Characterization of a multisubunit transcription factor complex essential for spliced-leader RNA gene transcription in Trypanosoma brucei. *Mol Cell Biol* **25**: 7303-7313.

Schulz D, Mugnier MR, Paulsen EM, Kim HS, Chung CW, Tough DF, Rioja I, Prinjha RK, Papavasiliou FN, Debler EW. 2015. Bromodomain Proteins Contribute to Maintenance of Bloodstream Form Stage Identity in the African Trypanosome. *PLoS Biol* **13**: e1002316.

Schulz D, Zaringhalam M, Papavasiliou FN, Kim HS. 2016. Base J and H3.V Regulate Transcriptional Termination in Trypanosoma brucei. *PLoS Genet* **12**: e1005762.

Shapiro TA, Englund PT. 1995. The structure and replication of kinetoplast DNA. *Annu Rev Microbiol* **49**: 117-143.

Shi H, Chamond N, Djikeng A, Tschudi C, Ullu E. 2009. RNA interference in Trypanosoma brucei: role of the n-terminal RGG domain and the polyribosome association of argonaute. *J Biol Chem* **284**: 36511-36520.

Shi H, Djikeng A, Tschudi C, Ullu E. 2004. Argonaute protein in the early divergent eukaryote Trypanosoma brucei: control of small interfering RNA accumulation and retroposon transcript abundance. *Mol Cell Biol* **24**: 420-427.

Shilatifard A. 2012. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu Rev Biochem* **81**: 65-95.

Shindou H, Shimizu T. 2009. Acyl-CoA:lysophospholipid acyltransferases. *J Biol Chem* **284**: 1-5.

Siegel TN, Hekstra DR, Kemp LE, Figueiredo LM, Lowell JE, Fenyo D, Wang X, Dewell S, Cross GA. 2009. Four histone variants mark the boundaries of polycistronic transcription units in Trypanosoma brucei. *Genes Dev* **23**: 1063-1076.

Singh PB, Miller JR, Pearce J, Kothary R, Burton RD, Paro R, James TC, Gaunt SJ. 1991. A sequence motif found in a Drosophila heterochromatin protein is conserved in animals and plants. *Nucleic Acids Res* **19**: 789-794.

Smid O, Horakova E, Vilimova V, Hrdy I, Cammack R, Horvath A, Lukes J, Tachezy J. 2006. Knock-downs of iron-sulfur cluster assembly proteins IscS and IscU down-regulate the active mitochondrion of procyclic Trypanosoma brucei. *J Biol Chem* **281**: 28679-28686.

Smith TK, Bringaud F, Nolan DP, Figueiredo LM. 2017. Metabolic reprogramming during the Trypanosoma brucei life cycle. *F1000Res* **6**.

Soding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* **33**: W244-248.

Stec I, Nagl SB, van Ommen GJ, den Dunnen JT. 2000. The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? *FEBS Lett* **473**: 1-5.

Stewart-Morgan KR, Petryk N, Groth A. 2020. Chromatin replication and epigenetic cell memory. *Nat Cell Biol* **22**: 361-371.

Suzuki MM, Bird A. 2008. DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet* **9**: 465-476.

Taddei A, Schober H, Gasser SM. 2010. The budding yeast nucleus. *Cold Spring Harb Perspect Biol* **2**: a000612.

Thatcher TH, Gorovsky MA. 1994. Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. *Nucleic Acids Res* **22**: 174-179.

Timmers HTM. 2020. SAGA and TFIID: Friends of TBP drifting apart. *Biochim Biophys Acta Gene Regul Mech* doi:10.1016/j.bbagrm.2020.194604: 194604.

Tschudi C, Ullu E. 1988. Polygene transcripts are precursors to calmodulin mRNAs in trypanosomes. *EMBO J* **7**: 455-463.

Tschudi C, Ullu E. 2002. Unconventional rules of small nuclear RNA transcription and cap modification in trypanosomatids. *Gene Expr* **10**: 3-16.

Tulin A, Chinenov Y, Spradling A. 2003. Regulation of chromatin structure and gene activity by poly(ADP-ribose) polymerases. *Curr Top Dev Biol* **56**: 55-83.

Tulin A, Spradling A. 2003. Chromatin loosening by poly(ADP)-ribose polymerase (PARP) at Drosophila puff loci. *Science* **299**: 560-562.

Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, Cox J. 2016. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**: 731-740.

Van Oss SB, Cucinotta CE, Arndt KM. 2017. Emerging Insights into the Roles of the Paf1 Complex in Gene Regulation. *Trends Biochem Sci* **42**: 788-798.

Vélez-Ramírez DE, Florencio-Martínez LE, Romero-Meza G, Rojas-Sánchez S, Moreno-Campos R, Arroyo R, Ortega-López J, Manning-Cela R, Martínez-Calvillo S. 2015. BRF1, a subunit of RNA polymerase III transcription factor TFIIIB, is essential for cell growth of Trypanosoma brucei. *Parasitology* **142**: 1563-1573.

Vellmer T, Hartleb L, Sola AF, Kramer S, Meyer-Natus E, Butter F, Janzen CJ. 2021. A novel SNF2 ATPase complex in Trypanosoma brucei with a role in H2A.Z-mediated chromatin remodelling. bioRxiv doi: 10.1101/2021.04.06.438560.

Venkatesh S, Workman JL. 2013. Set2 mediated H3 lysine 36 methylation: regulation of transcription elongation and implications in organismal development. *Wiley Interdiscip Rev Dev Biol* **2**: 685-700.

Vetting MW, Bareich DC, Yu M, Blanchard JS. 2008. Crystal structure of RimI from Salmonella typhimurium LT2, the GNAT responsible for N(alpha)-acetylation of ribosomal protein S18. *Protein Sci* **17**: 1781-1790.

Wang QP, Kawahara T, Horn D. 2010. Histone deacetylases play distinct roles in telomeric VSG expression site silencing in African trypanosomes. *Mol Microbiol* **77**: 1237-1245.

Wang R, Gao J, Zhang J, Zhang X, Xu C, Liao S, Tu X. 2019. Solution structure of TbTFIIS2-2 PWWP domain from Trypanosoma brucei and its binding to H4K17me3 and H3K32me3. *Biochem J* **476**: 421-431.

Wang X, Ahmad S, Zhang Z, Cote J, Cai G. 2018. Architecture of the Saccharomyces cerevisiae NuA4/TIP60 complex. *Nat Commun* **9**: 1147.

Webb S, Kudla G, Granneman, S. 2018. The pyCRAC Manual. Available online at: https://git.ecdf.ed.ac.uk/sgrannem/pycrac [Accessed on 02 December 2019].

Wedel C, Forstner KU, Derr R, Siegel TN. 2017. GT-rich promoters can drive RNA pol II transcription and deposition of H2A.Z in African trypanosomes. *EMBO J* **36**: 2581-2594.

Willhoft O, Wigley DB. 2020. INO80 and SWR1 complexes: the non-identical twins of chromatin remodelling. *Curr Opin Struct Biol* **61**: 50-58.

Wiśniewski JR, Zougman A, Nagaraj N, Mann M. 2009. Universal sample preparation method for proteome analysis. *Nature methods* **6**: 359-362.

Wright JR, Siegel TN, Cross GA. 2010. Histone H3 trimethylated at lysine 4 is enriched at probable transcription start sites in Trypanosoma brucei. *Mol Biochem Parasitol* **172**: 141-144.

Yang X, Figueiredo LM, Espinal A, Okubo E, Li B. 2009. RAP1 is essential for silencing telomeric variant surface glycoprotein genes in Trypanosoma brucei. *Cell* **137**: 99-109.

Zaware N, Zhou MM. 2019. Bromodomain biology and drug discovery. *Nat Struct Mol Biol* **26**: 870-879.

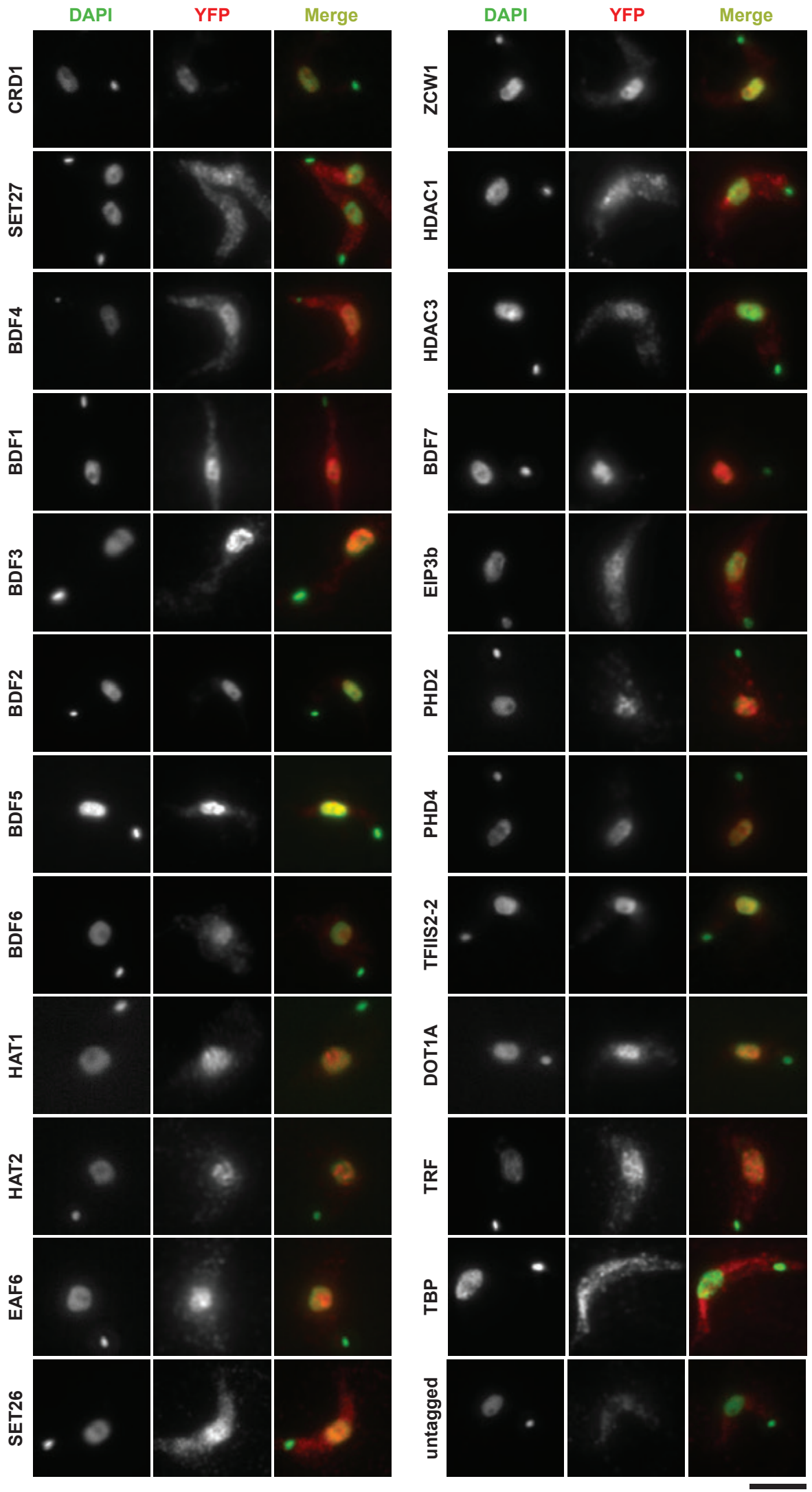
DAPI
YFP
Merge
CRD1
SET27
BDF4
BDF1
BDF3
BDF2
BDF5
BDF6
HAT1
HAT2
EAF6
SET26
ZCW1
HDAC1
HDAC3
BDF7
EIP3b
PHD2
PHD4
TFIIS2-2
DOT1A
TRF
TBP
untagged
5 µm

Fig. 1

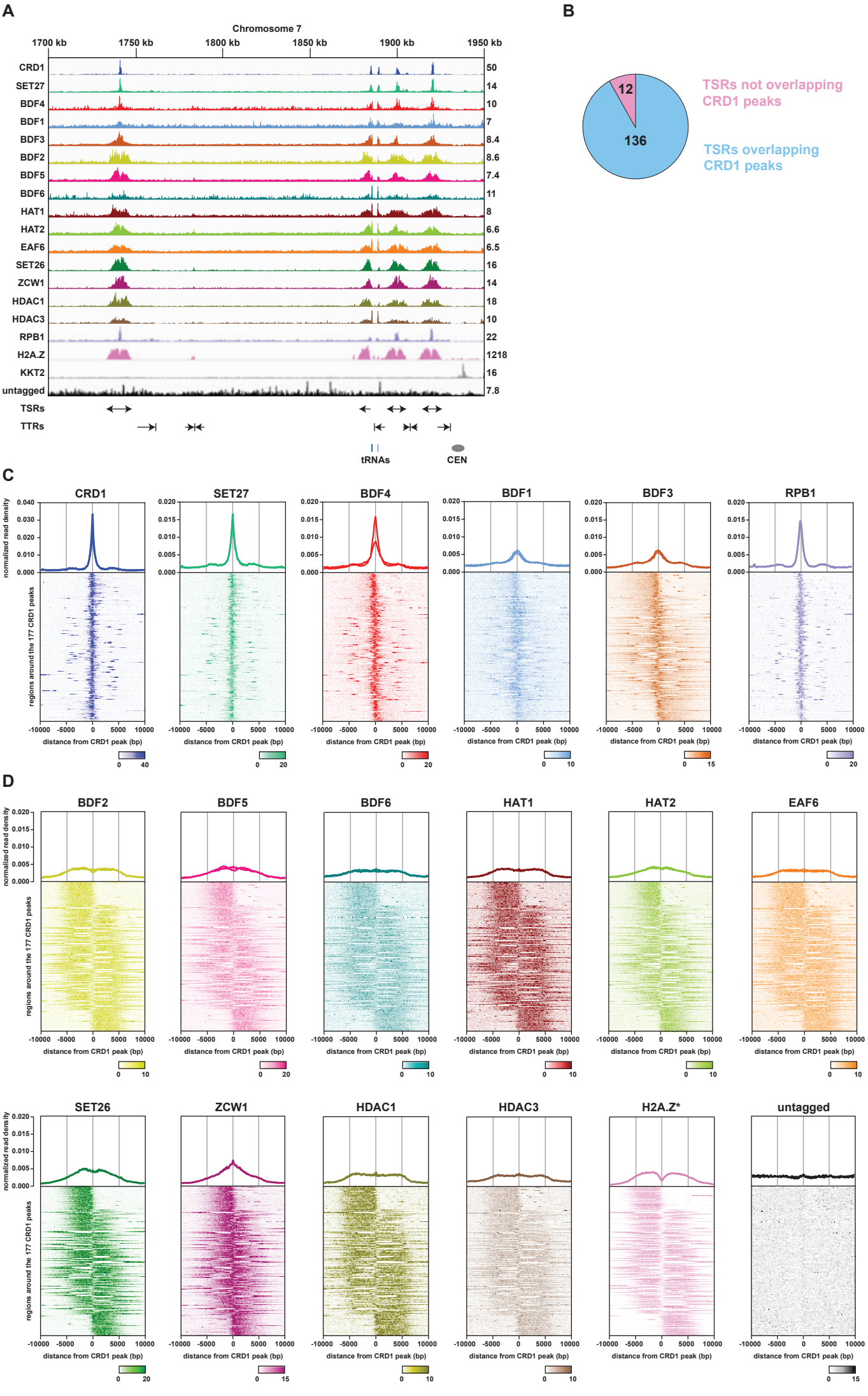
A
Chromosome 7
1700 kb
1750 kb
1800 kb
1850 kb
1900 kb
1950 kb
CRD1
SET27
BDF4
BDF1
BDF3
BDF2
BDF5
BDF6
HAT1
HAT2
EAF6
SET26
ZCW1
HDAC1
HDAC3
RPB1
H2A.Z
KKT2
untagged
50
14
10
7
8.4
8.6
7.4
11
8
6.6
6.5
16
14
18
10
22
1218
16
7.8
TSRs
TTRs
tRNAs
CEN
B
12
136
TSRs not overlapping CRD1 peaks
TSRs overlapping CRD1 peaks
C
CRD1
SET27
BDF4
BDF1
BDF3
RPB1
normalized read density
regions around the 177 CRD1 peaks
distance from CRD1 peak (bp)
D
BDF2
BDF5
BDF6
HAT1
HAT2
EAF6
SET26
ZCW1
HDAC1
HDAC3
H2A.Z*
untagged

Fig. 2

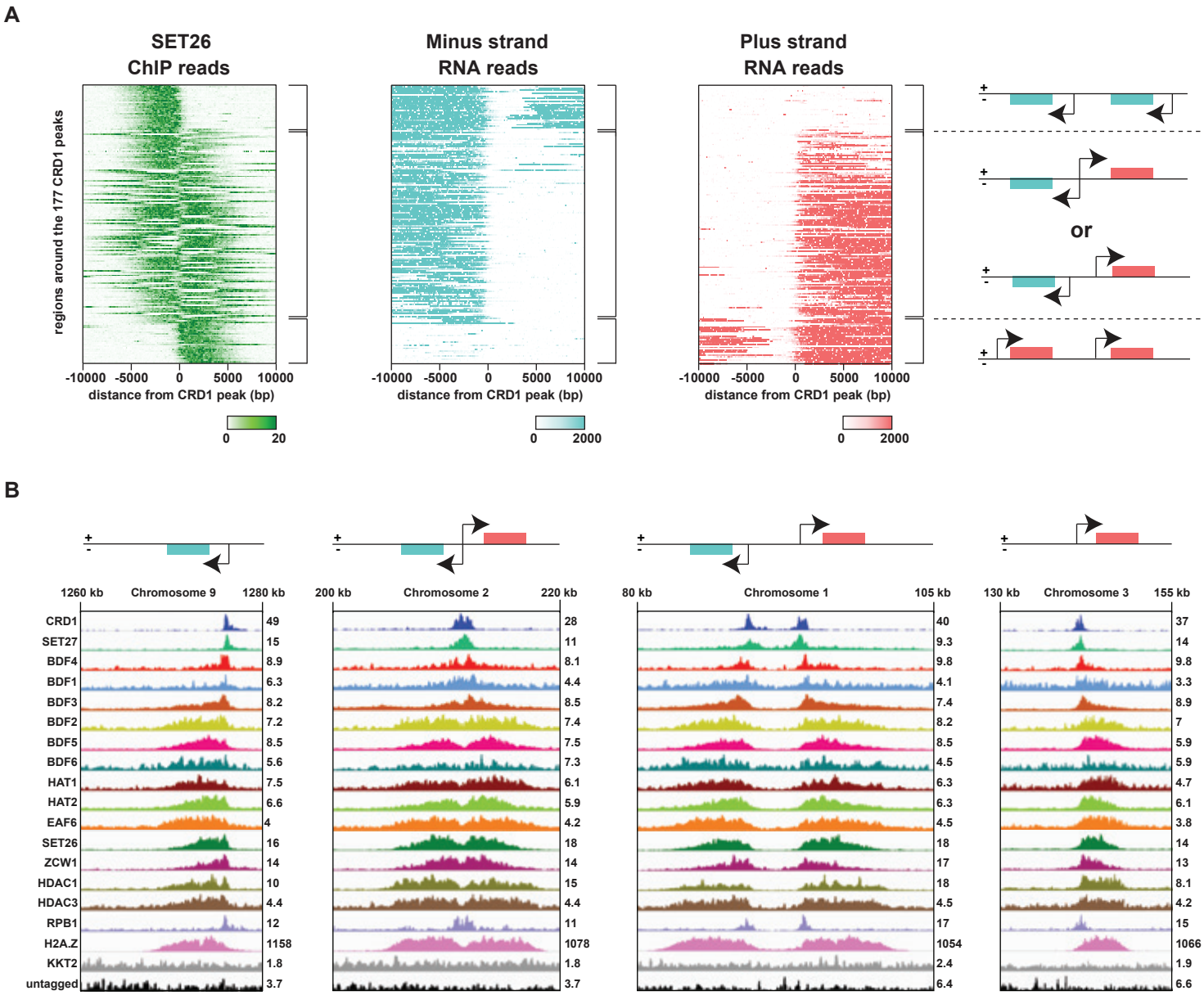
A
SET26
ChIP reads
Minus strand
RNA reads
Plus strand
RNA reads
regions around the 177 CRD1 peaks
-10000 -5000 0 5000 10000
distance from CRD1 peak (bp)
0 20
0 2000
0 2000
or
B
1260 kb Chromosome 9 1280 kb
200 kb Chromosome 2 220 kb
80 kb Chromosome 1 105 kb
130 kb Chromosome 3 155 kb
CRD1 49 28 40 37
SET27 15 11 9.3 14
BDF4 8.9 8.1 9.8 9.8
BDF1 6.3 4.4 4.1 3.3
BDF3 8.2 8.5 7.4 8.9
BDF2 7.2 7.4 8.2 7
BDF5 8.5 7.5 8.5 5.9
BDF6 5.6 7.3 4.5 5.9
HAT1 7.5 6.1 6.3 4.7
HAT2 6.6 5.9 6.3 6.1
EAF6 4 4.2 4.5 3.8
SET26 16 18 18 14
ZCW1 14 14 17 13
HDAC1 10 15 18 8.1
HDAC3 4.4 4.4 4.5 4.2
RPB1 12 11 17 15
H2A.Z 1158 1078 1054 1066
KKT2 1.8 1.8 2.4 1.9
untagged 3.7 3.7 6.4 6.6

Fig. 3

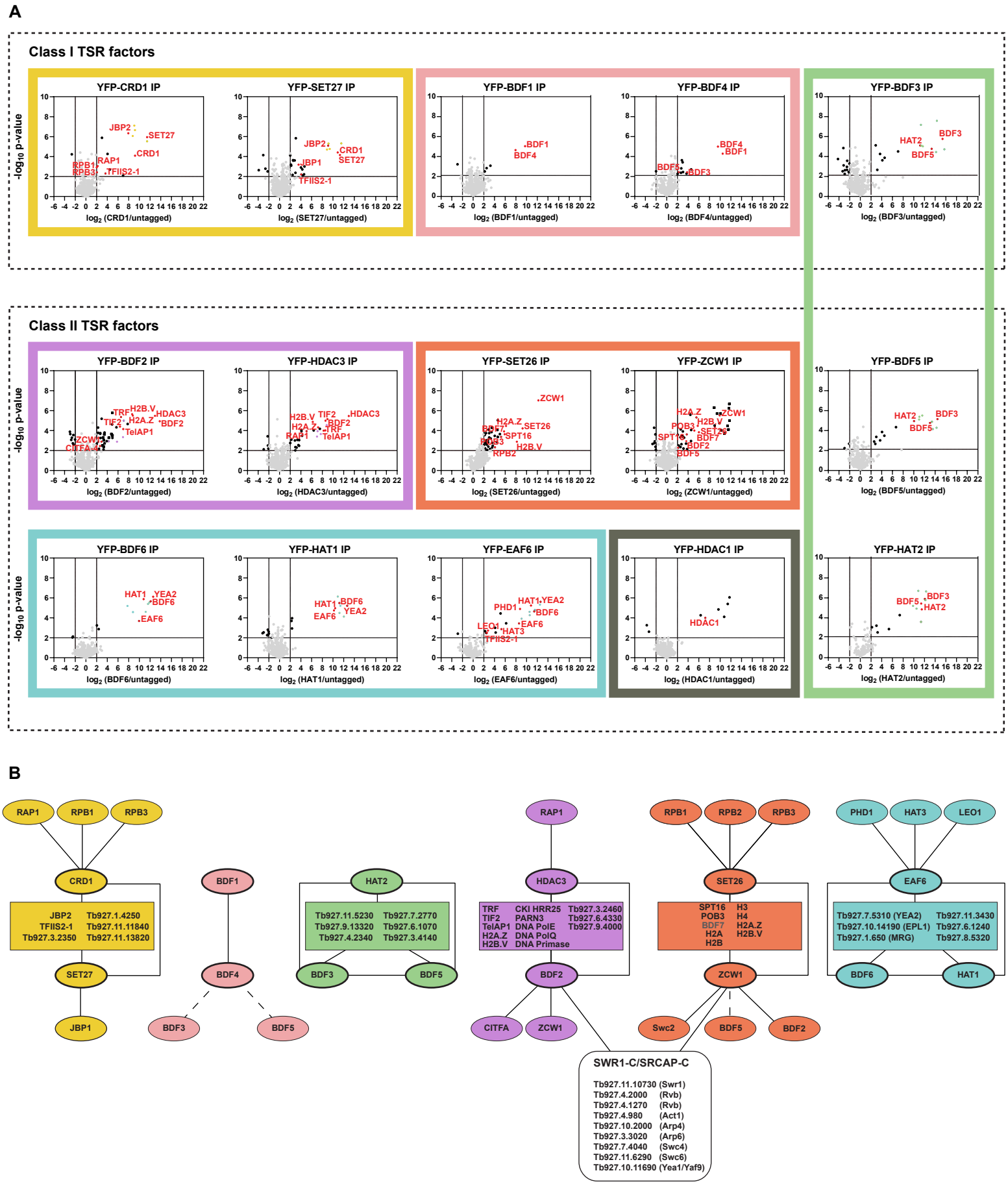
A
Class I TSR factors
YFP-CRD1 IP
YFP-SET27 IP
YFP-BDF1 IP
YFP-BDF4 IP
YFP-BDF3 IP
Class II TSR factors
YFP-BDF2 IP
YFP-HDAC3 IP
YFP-SET26 IP
YFP-ZCW1 IP
YFP-BDF5 IP
YFP-BDF6 IP
YFP-HAT1 IP
YFP-EAF6 IP
YFP-HDAC1 IP
YFP-HAT2 IP
-log10 p-value
log2 (CRD1/untagged)
log2 (SET27/untagged)
log2 (BDF1/untagged)
log2 (BDF4/untagged)
log2 (BDF3/untagged)
log2 (BDF2/untagged)
log2 (HDAC3/untagged)
log2 (SET26/untagged)
log2 (ZCW1/untagged)
log2 (BDF5/untagged)
log2 (BDF6/untagged)
log2 (HAT1/untagged)
log2 (EAF6/untagged)
log2 (HDAC1/untagged)
log2 (HAT2/untagged)
B
RAP1
RPB1
RPB3
CRD1
JBP2 Tb927.1.4250
TFIIS2-1 Tb927.11.11840
Tb927.3.2350 Tb927.11.13820
SET27
JBP1
BDF1
BDF4
BDF3
BDF5
HAT2
Tb927.11.5230 Tb927.7.2770
Tb927.9.13320 Tb927.6.1070
Tb927.4.2340 Tb927.3.4140
BDF3
BDF5
RAP1
HDAC3
TRF CKI HRR25 Tb927.3.2460
TIF2 PARN3 Tb927.6.4330
TelAP1 DNA PolE Tb927.9.4000
H2A.Z DNA PolQ
H2B.V DNA Primase
BDF2
CITFA
ZCW1
RPB1
RPB2
RPB3
SET26
SPT16 H3
POB3 H4
BDF7 H2A.Z
H2A H2B.V
H2B
ZCW1
Swc2
BDF5
BDF2
PHD1
HAT3
LEO1
EAF6
Tb927.7.5310 (YEA2) Tb927.11.3430
Tb927.10.14190 (EPL1) Tb927.6.1240
Tb927.1.650 (MRG) Tb927.8.5320
BDF6
HAT1
SWR1-C/SRCAP-C
Tb927.11.10730 (Swr1)
Tb927.4.2000 (Rvb)
Tb927.4.1270 (Rvb)
Tb927.4.980 (Act1)
Tb927.10.2000 (Arp4)
Tb927.3.3020 (Arp6)
Tb927.7.4040 (Swc4)
Tb927.11.6290 (Swc6)
Tb927.10.11690 (Yea1/Yaf9)

Fig. 4

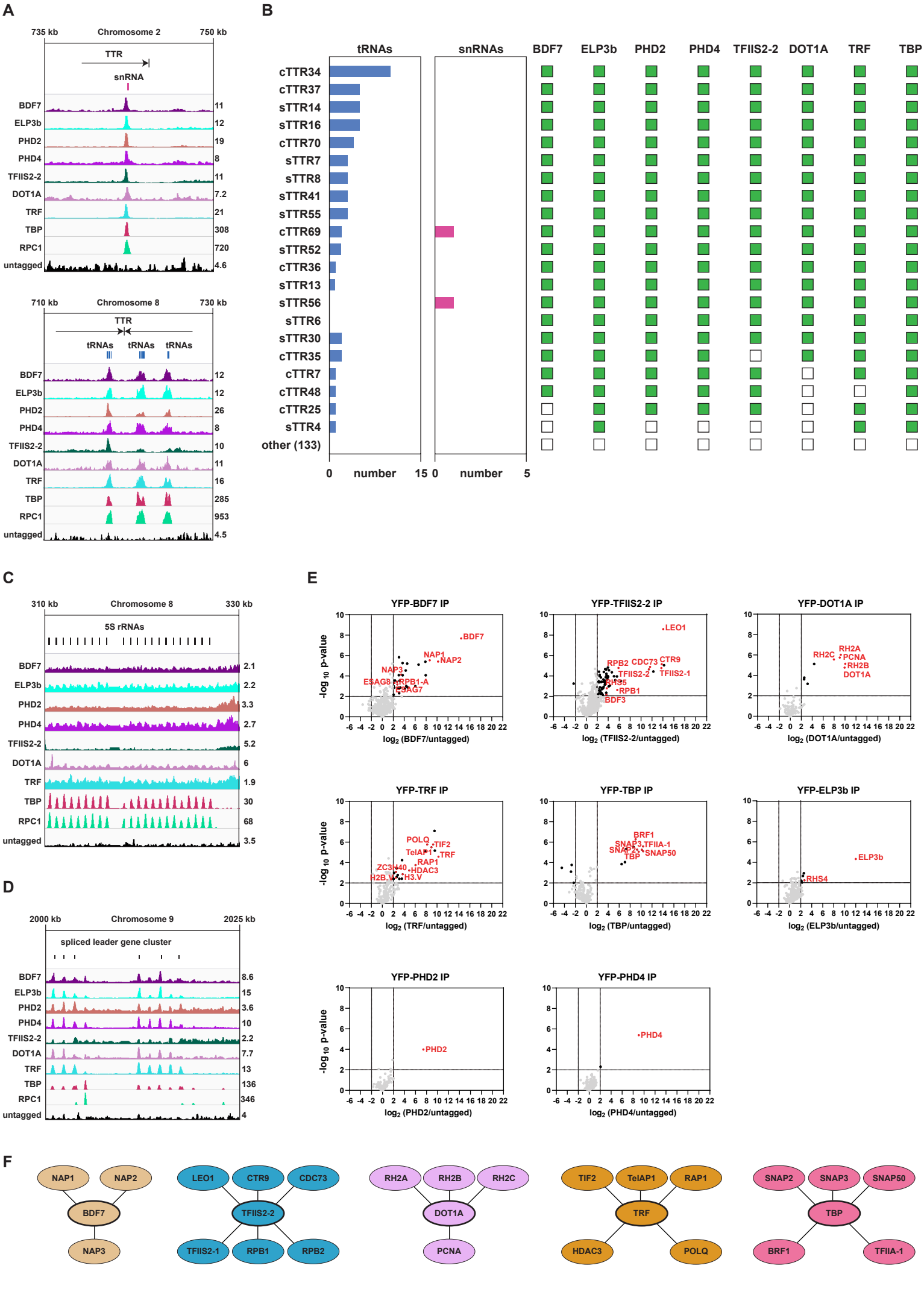
A
735 kb
Chromosome 2
750 kb
TTR
snRNA
BDF7 11
ELP3b 12
PHD2 19
PHD4 8
TFIIS2-2 11
DOT1A 7.2
TRF 21
TBP 308
RPC1 720
untagged 4.6
710 kb
Chromosome 8
730 kb
TTR
tRNAs tRNAs tRNAs
BDF7 12
ELP3b 12
PHD2 26
PHD4 8
TFIIS2-2 10
DOT1A 11
TRF 16
TBP 285
RPC1 953
untagged 4.5
B
tRNAs
snRNAs
BDF7
ELP3b
PHD2
PHD4
TFIIS2-2
DOT1A
TRF
TBP
cTTR34
cTTR37
sTTR14
sTTR16
cTTR70
sTTR7
sTTR8
sTTR41
sTTR55
cTTR69
sTTR52
cTTR36
sTTR13
sTTR56
sTTR6
sTTR30
cTTR35
cTTR7
cTTR48
cTTR25
sTTR4
other (133)
0 number 15
0 number 5
C
310 kb
Chromosome 8
330 kb
5S rRNAs
BDF7 2.1
ELP3b 2.2
PHD2 3.3
PHD4 2.7
TFIIS2-2 5.2
DOT1A 6
TRF 1.9
TBP 30
RPC1 68
untagged 3.5
D
2000 kb
Chromosome 9
2025 kb
spliced leader gene cluster
BDF7 8.6
ELP3b 15
PHD2 3.6
PHD4 10
TFIIS2-2 2.2
DOT1A 7.7
TRF 13
TBP 136
RPC1 346
untagged 4
E
YFP-BDF7 IP
BDF7
NAP1
NAP2
NAP3
ESAG8
RPB1-A
ESAG7
-log10 p-value
log2 (BDF7/untagged)
YFP-TFIIS2-2 IP
LEO1
RPB2
CDC73
CTR9
TFIIS2-2
TFIIS2-1
RPB5
RPB1
BDF3
log2 (TFIIS2-2/untagged)
YFP-DOT1A IP
RH2C
RH2A
PCNA
RH2B
DOT1A
log2 (DOT1A/untagged)
YFP-TRF IP
POLQ
TIF2
TelAP1
TRF
RAP1
ZC3H40
HDAC3
H2B.V
H3.V
log2 (TRF/untagged)
YFP-TBP IP
BRF1
SNAP3
TFIIA-1
SNAP2
SNAP50
TBP
log2 (TBP/untagged)
YFP-ELP3b IP
ELP3b
RHS4
log2 (ELP3b/untagged)
YFP-PHD2 IP
PHD2
log2 (PHD2/untagged)
YFP-PHD4 IP
PHD4
log2 (PHD4/untagged)
F
NAP1
NAP2
BDF7
NAP3
LEO1
CTR9
CDC73
TFIIS2-2
TFIIS2-1
RPB1
RPB2
RH2A
RH2B
RH2C
DOT1A
PCNA
TIF2
TelAP1
RAP1
TRF
HDAC3
POLQ
SNAP2
SNAP3
SNAP50
TBP
BRF1
TFIIA-1

Fig. 5

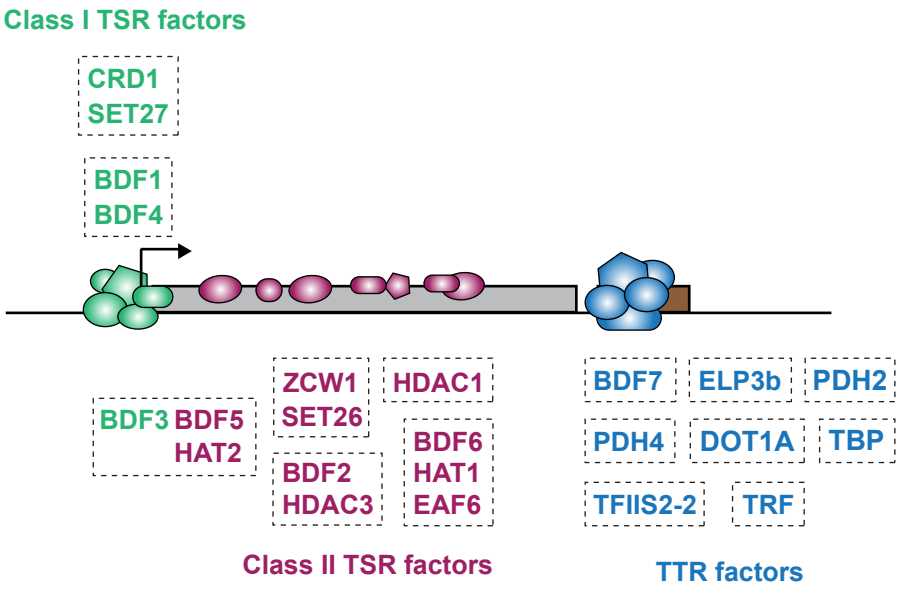
Class I TSR factors
CRD1
SET27
BDF1
BDF4
BDF3
BDF5
HAT2
ZCW1
SET26
HDAC1
BDF2
HDAC3
BDF6
HAT1
EAF6
Class II TSR factors
BDF7
ELP3b
PDH2
PDH4
DOT1A
TBP
TFIIS2-2
TRF
TTR factors

Fig. 6