expert reviews
in molecular medicine

Epigenetic mechanisms, nuclear architecture and the control of gene expression in trypanosomes

Sam Alsford[1], Kelly duBois[2], David Horn[1] and Mark C. Field[2,*]

The control of gene expression, and more significantly gene cohorts, requires tight transcriptional coordination and is an essential feature of probably all cells. In higher eukaryotes, the mechanisms used involve controlled modifications to both local and global DNA environments, principally through changes in chromatin structure as well as *cis*-element-driven mechanisms. Although the mechanisms regulating chromatin in terms of transcriptional permissiveness and the relation to developmental programmes and responses to the environment are becoming better understood for animal and fungal cells, it is only just beginning to become clear how these processes operate in other taxa, including the trypanosomatids. Recent advances are now illuminating how African trypanosomes regulate higher-order chromatin structure, and, further, how these mechanisms impact on the expression of major surface antigens that are of fundamental importance to life-cycle progression. It is now apparent that several mechanisms are rather more similar between animal and fungal cells and trypanosomes than it originally appeared, but some aspects do involve gene products unique to trypanosomes. Therefore, both evolutionarily common and novel mechanisms cohabit in trypanosomes, offering both important biological insights and possible therapeutic opportunity.

Trypanosomatids constitute a taxonomic grouping of organisms that are highly divergent from animals and fungi (Fig. 1). As a group, trypanosomes have a huge impact on the biota, and representatives include important human disease agents that directly threaten over half-a-billion individuals (Ref. 1). Furthermore, trypanosomatids are increasingly recognised as disease agents of plants, insects and wild animal populations (Refs 2, 3, 4), resulting in an unquantified, but probably enormous, impact on agriculture and the flora and fauna of a given region. In Africa, the African trypanosomes (*Trypanosoma brucei* spp.) have had a major effect

[1]London School of Hygiene and Tropical Medicine, London, UK

[2]Department of Pathology, University of Cambridge, Cambridge, UK

*Corresponding author: Mark C. Field, Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge, UK. E-mail: mcf34@cam.ac.uk

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

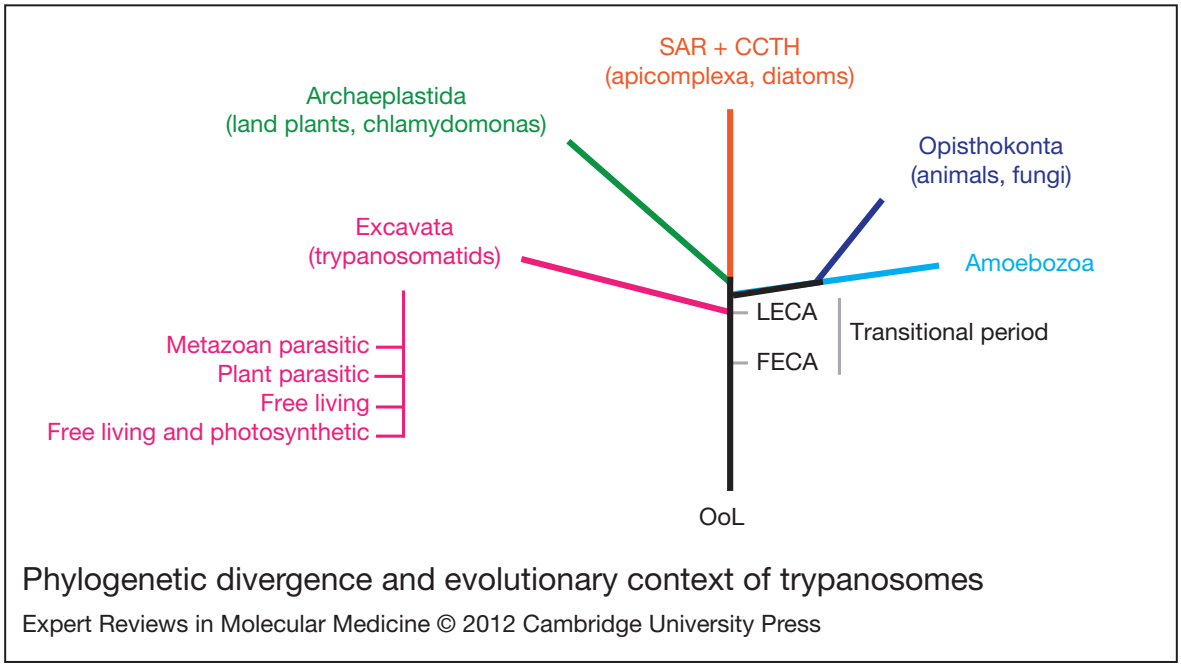

Figure 1. Phylogenetic divergence and evolutionary context of trypanosomes. Schematic representation of the eukaryotic tree of life. Five supergroups are represented reflecting current thinking on large-scale taxonomic groupings. Trypanosomatids are members of the Excavata (pink) and highly divergent from their most significant hosts, the animals (Opisthokonta, blue), and also land plants (Archaeplastida, green). Positioning of the root reflects recent arguments and is highly speculative. Note the transitional period between the eukaryogenesis event represented by the first eukaryotic common ancestor (FECA, equivalent to the eukaryogenesis event) and radiation of modern lineages represented by the last eukaryotic common ancestor (LECA), which probably corresponds to a period of intense biological innovation. SAR + CCTH, stramenopile, alveolate and Rhizaria plus cryptomonads, centrohelids, telonemids and haptophytes; OoL, origin of life.

on the savannah-like ecology and human history (see Refs 5, 6 for discussion on this issue). Even more distantly related organisms such as *Euglena*, which are highly abundant members of aquatic and marine ecosystems, exert great influence over the entire biosphere (Ref. 7). Trypanosomatids are incredibly evolutionarily divergent from animals, fungi and plants, and this is reflected in the multiple unusual and near-unique features of their cell biology, metabolism and gene expression. The eukaryotic root might even lie close to the origin of the trypanosomatid lineage (Fig. 1), suggesting that this divergent biology also reflects some deeper evolutionary origins (Ref. 8). Particularly relevant is the dominance of GPI-anchored proteins and glycoconjugates at the cell surface, and polycistronic transcription and *trans*-splicing of protein-coding genes. The presence of high-abundance surface antigens on several of the better characterised parasitic trypanosomatids, and in many cases, extreme developmental variation resulting in complete remodelling of the surface antigen repertoire during differentiation, suggests that this aspect of trypanosome cell biology particularly relies on strict mechanisms for altering gene expression profiles, which is supported by transcriptome analysis (Refs 9, 10). Because most studies have focused on pathogenic trypanosomes, and in the absence of well-characterised free-living members of the kinetoplastida lineage, it remains an open question as to how much of the unusual biology of trypanosomes arises from their evolutionary position or selective pressure arising from a parasitic lifestyle.

T. brucei is the causative agent of African sleeping sickness in humans and a wasting disease, n'agana, in cattle. The parasite is transmitted by the bite of the Tsetse fly. In both host and vector, a complex differentiation

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

pattern is executed that appears balanced to prevent excess damage to the host and to continuously generate highly infective forms for transmission (Refs 11, 12, 13). In this regard *T. brucei* is poorly adapted to modern humans and European cattle (*Bos taurus*), because the disease is unusually aggressive and invariably fatal unless treated; this is in contrast to many of the endemic mammalian species, including *Bos indicus*, where infections can persist and are comparatively well tolerated. African trypanosomes are unusual pathogens in that they are exclusively extracellular, and therefore exposed to the immune system in mammalian hosts – a behaviour that is seemingly at odds with the ability of the parasite to survive for prolonged periods within the host. Immune evasion is primarily achieved by antigenic variation (Ref. 12). Briefly, the surface of the mammalian infective form of the parasite is dominated by the GPI-anchored variant surface glycoprotein (VSG), which contributes an impressive ~90% of cell surface protein. Switching the active *VSG* gene allows selection of an immunologically distinct VSG coat, which can evade the immune response. The presence of a repertoire of many hundreds of VSG sequences, recombination between them accessing a potentially inexhaustible supply of novel *VSG* genes, plus monoallelic expression achieved by restricting transcription to a single subtelomeric specialised expression site (ES) allows continued immune evasion for many months or even years.

The surface of most stages of *T. brucei* is dominated by superabundant GPI-anchored antigens, and in addition to VSG in the mammalian infective form, procyclin is present in the insect midgut-dwelling procyclic form, and finally brucei stage alanine-rich protein (BARP) in epimastigote stages that reside within the fly proventricular region (Ref. 14). Significantly, transcriptional mechanisms in trypanosomes are rather unusual, with the majority of protein-coding genes arranged in long polycistronic transcriptional units (PTUs) initiating from strand-switch regions, at least in the case of divergently oriented PTUs (Fig. 2). These PTUs are transcribed principally by RNA polymerase II (RNA Pol II) and the resulting nascent RNA is resolved by *trans*-splicing that places an invariant spliced leader at the 5′ end (Ref. 15). 5′ cap methylation and polyadenylation of the 3′ end completes mRNA maturation (Fig. 2). This mechanism at first glance appears to function in the absence of clear promoter-mediated control, and transcription of all open reading frames within a given PTU appears equivalent, with the further consequence being that mRNA abundance is principally mediated by post-transcriptional mechanisms (Ref. 16). Although sequences within the 3′ UTR of mRNAs are clearly involved in regulation of mRNA stability, and hence copy number, it is not established whether this accounts for all mRNA copy number control, and it is probable that mRNA transport, which might be affected by the state of the chromatin and other factors, also has a role (Refs 17, 18). Codon usage bias is also likely to have a large impact on translation control and protein copy number (Ref. 19).

The principal nuclear architectural features are conserved between trypanosomes and higher eukaryotes, and include the double bilayer nuclear envelope, nuclear pores, a single nucleolus and electron-dense chromatin regions close to the nuclear periphery whose arrangement is under developmental control (Refs 20, 21, 22, 23, 24). There is good evidence for positioning of telomeres at the periphery within the nucleus (Ref. 25), and suggestions for variance in the kinetics of the segregation of the different classes of chromosomes (i.e. megabase versus minichromosomes) during mitosis (Refs 26, 27, 28). The location of the telomeres at the nuclear periphery is highly significant because telomeres are also the loci containing the *VSG* ESs. Current models suggest that this location serves in some manner to suppress full expression from these sites, a common aspect of epigenetic control (Refs 29, 30, 31).

Epigenetics

Epigenetics comprises a suite of mechanisms that regulate gene expression but are not directly encoded within the genome, and therefore excluding promotors and other *cis*-based sequence elements (Refs 32, 33). Epigenetic mechanisms appear to exploit the overall organisation of the nucleus and DNA itself, suggesting that they coevolved with the establishment of the nucleus, or soon after (Table 1). Importantly, epigenetic states are heritable and can take many forms (Refs 34, 35, 36), and epigenetic control essentially equates with the organisation of

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

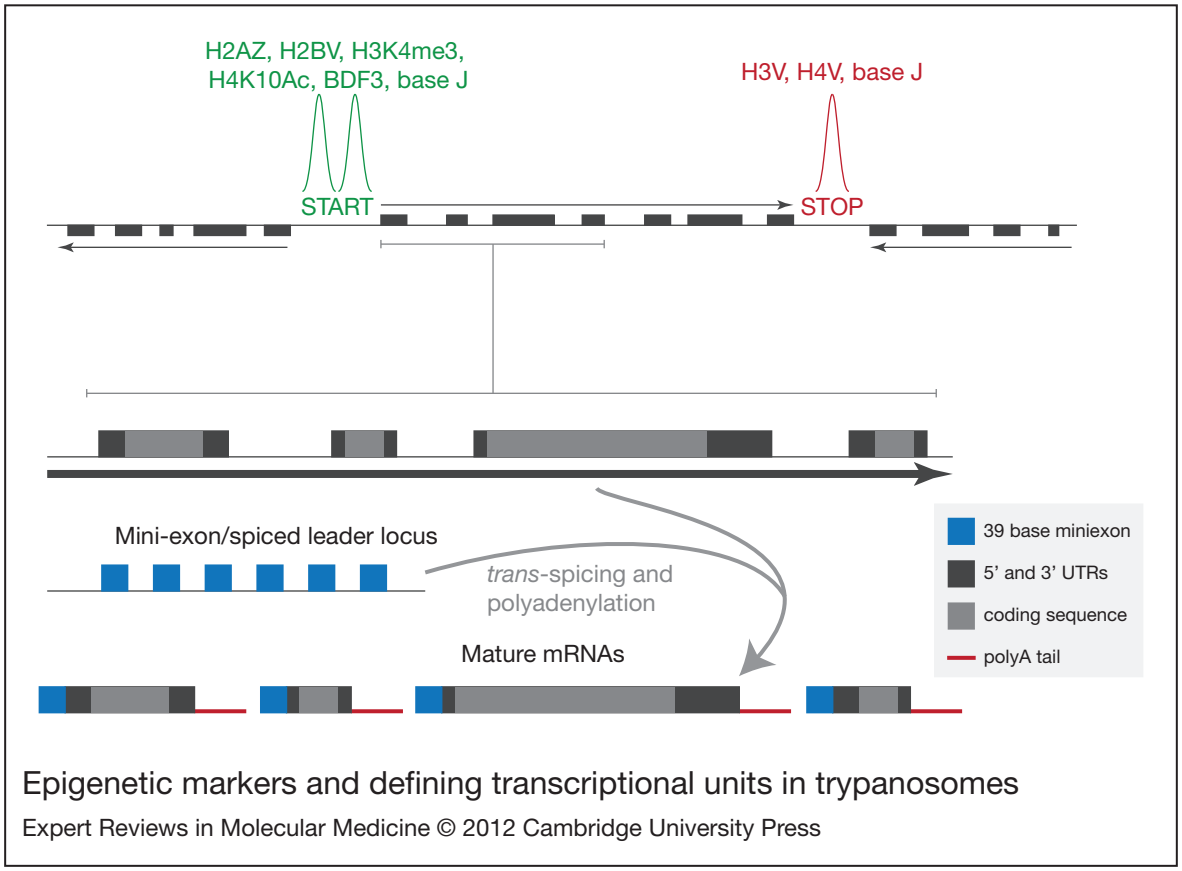

Figure 2. Epigenetic markers and defining transcriptional units in trypanosomes. The top schema shows a region of a megabase chromosome, with two strand-switch regions, one marked 'START' where the two flanking polycistronic transcriptional units (PTUs) are transcribed in a divergent manner and one marked 'STOP', where the PTUs converge. Evidence suggests that transcription initiates at the divergent sites, and terminates at convergent regions; the divergent regions probably resemble RNA-Pol-II-like promoter elements. Recent mapping using ChIP coupled to high-throughput sequencing has found these switch regions to be associated with both variant histones and specific modifications of H3 and H4. The lower schema, connected with a tie bar, shows the resolution of a portion of a polycistronic transcription unit by a combination of *trans*-splicing, adding a mini-exon (blue) and polyadenylation (red). Arrows indicate nascent mRNAs with the arrowhead to show the direction of transcription.

chromatin structure in space and time. The major aspects are threefold: (i) direct modification of DNA itself, for example, by methylation or glucosylation of specific bases; (ii) specification of the density and permissiveness of specific DNA sequences to transcription, frequently based around structural organisation of histones and their assembly into nucleosomes, itself based on post-translational modifications of histones; and (iii) three-dimensional positioning of chromatin with, at least in higher eukaryotes, less transcriptionally active chromatin regions (heterochromatin) tending to be positioned at the nuclear periphery and dependent on intermediate filament and nuclear envelope proteins. All these processes are also probably intimately associated with higher-order organisation of the nuclear contents; in the noncondensed interphase state, metazoan chromosomes are compacted and restricted to 'territories', whereas even the arrangement of chromosomal territories is ordered (Ref. 37).

Epigenetic control is used in a great number of contexts, but highly relevant here is developmental regulation of the expression of gene cohorts, reflecting the extremely stable and high dynamic range of relative expression levels of heterochromatin versus euchromatin (active,

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

Table 1. Chromatin-associated proteins that impact transcription at *T. brucei* telomeres

Category	Protein	Role	Similar functions in other eukaryotes?	Ref.
Chromatin components	Histones	Repressor (histone H3 tested)	Yes (yeast and mammals)	[a]
	NLP	Repressor (nucleoplasmin)	Not demonstrated	108
Chromatin modifiers	DOT1B	Repressor (methyltransferase)	Yes (yeast, Dot1)	110
	SIR2rp1	Repressor (deacetylase)	Yes (yeast, Sir2)	100
	HAT1	Repressor (SIR2rp1 dependent)	Yes (yeast, Sas2, Esa1, mst2+)	101
	DAC1	HAT1-SIR2rp1 antagonism	Yes (yeast, Rpd3 and *Drosophila*, HDAC1)	102
	DAC3	Repressor (deacetylase)	Yes (yeast, Hda1)	102
Chromatin regulators	RAP1	Repressor (telomere binding)	Yes (yeast)	109
	ISWI	Repressor (nucleosome remodeller)	Yes (yeast, Isw2)	105
	FACT	Repressor (histone chaperone)	Centromeric heterochromatin (yeast)	107
	NUP-1	Repressor (lamina component)	Nuclear lamina tethering results in repression	90
	CAF-1b	Repressor (histone chaperone)	Yes (yeast and mammals)	[a]
	ASF1A	Repressor (histone chaperone)	Yes (yeast)	[a]
	Cohesin	Regulation of inheritance	Telomeric heterochromatin (yeast)	27

[a]Our unpublished results. NLP, nucleoplasmin-like protein; DOT1B, disruptor of telomeric silencing 1B; SIR2rp1, silent information regulator related protein 1; HAT1, histone acetyltransferase 1; DAC1–DAC3, histone deacetylase 1–3; RAP1, repressor-activator protein 1; ISWI, imitation switch; FACT, facilitates chromatin transcription; NUP-1, nuclear periphery protein-1; CAF-1b, chromatin assembly factor 1-b; ASF1A, antisilencing factor 1A.

open chromatin) expression levels. Potentially, such mechanisms ensure against ectopic expression of inappropriate gene products, which is critical in tissue differentiation in metazoa and multicellular plants, and the successful completion of ontological development where abnormal cues at early stages can become amplified into gross morphological defects. In addition, the process can retain environmental information; for example, in *Saccharomyces cerevisiae*, as a result of DNA-zip-code-dependent positioning of relevant genes, daughter cells retain a memory of the carbon source their parental cell used, facilitating preferential reactivation should the same carbon source be re-encountered (Refs 38, 39).

In trypanosomes, tight regulation of developmentally expressed surface proteins is a clear requirement. As the parasite progresses

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

through the life cycle, it expresses VSG in the mammalian host, which is replaced by procyclin and then BARP in the insect vector before returning to VSG expression. Although definitive proof is lacking, developmental progression is considered essential for adaptation to specific environments; there is evidence for quality control mechanisms ensuring against ectopic expression, suggesting that trypanosomes are under considerable pressure to fully switch coat composition following differentiation (Ref. 40) or to avoid inappropriate surface antigen expression (Ref. 41). It is not known whether trypanosomes possess epigenetic memory that persists following progression through the life cycle, but the concept that antigenic variation and selective ES utility has a selective advantage has been discussed many times, although its relevance remains unproven (Refs 42, 43, 44).

Given probable origins at the earliest stages of eukaryotic evolution and conservation of several features of nuclear biology between trypanosomes and organisms with well-established epigenetic mechanisms, it is important to understand how and to what extent trypanosomes exploit epigenetics. We focus almost exclusively on *T. brucei*, for the rather pragmatic reason that the vast majority of detailed molecular work has been performed using this organism, and to a first approximation, studies in related trypanosomatids suggest that broadly similar mechanisms probably operate. We will first consider some of the classical work that predates completion of the trypanosome genome sequence, and then more recent studies that have taken candidate-based approaches to analysis of chromatin modifiers. Finally, we present recent data on divergent systems for organisation of heterochromatin and consider recent genome-wide approaches.

The dark ages before the genome sequence

The structure of the nucleus of *T. brucei* is similar to that of other eukaryotes (Ref. 45) and is bound by a membrane punctuated by 200–300 nuclear pores, contains the cell's chromatin and has a single nucleolus for rRNA transcription and ribosome assembly. In contrast, the *T. brucei* nucleus lacks a metazoan-type nuclear lamina, contains more than 120 chromosomes and extrachromosomal elements, classes of which exhibit different modes of segregation (Ref. 26), hosts extranucleolar RNA Pol I transcription (Ref. 29), and in common with many unicellular taxa, the nuclear membrane and nucleolus maintain integrity throughout mitosis (Ref. 28).

A suspicion that trypanosomatid DNA was organised in a manner distinct from the higher eukaryotes has been a common refrain, even reaching back to the (with hindsight) bizarre suggestion that trypanosome DNA is not associated with histones (Ref. 46). This refrain of radically divergent mechanisms persisted, being supported by several reports of an absence of histone H1 (Refs 47, 48) and suggestions that the absence of conventional histone complements correlates with greater micrococcal nuclease sensitivity than mammalian chromatin (Ref. 49). The parallel discoveries of *trans*-splicing and polycistronic transcription also lent credence to conventional wisdom that trypanosome biology was highly divergent from metazoan and fungal systems, with trypanosomes perhaps using completely distinct machinery for gene expression.

This view began to erode when it became clear that H1 was indeed present, initially by biochemical isolation, and a decade later by identification of histone gene orthologues (Refs 50, 51). Unexpected levels of sequence divergence probably explain the earlier detection issues, whereas differential micrococcal nuclease sensitivity between BSF and PCF DNA suggests life-cycle remodelling of chromatin organisation on a wide scale (Ref. 52). Furthermore, *trans*-splicing is performed using essentially the same splicing factors and biochemistry as conventional *cis*-splicing, indicating that although trypanosomes are indeed unusual, they are essentially exploiting molecular mechanisms with strong commonality to higher eukaryotes.

In the early 1990s, a modified base, β-D-glucopyranosyloxymethyluracil (or J) was reported in trypanosome DNA, and was suggested to associate specifically with telomeric DNA bearing inactive *VSG* genes (Refs 53, 54, see Ref. 55 for a discussion on this issue). Evidence that J was essentially absent from the insect form was consistent with a role in *VSG* repression, and further implied that the mechanism of *VSG* silencing was distinct between the two life stages. Unfortunately, this

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

concept proved too simplistic because J is present in telomeric and nontelomeric DNA across and beyond the kinetoplastids, with no clear pattern emerging that correlates antigenic variation with J presence and/or location. Recent knockouts of J-binding proteins 1 and 2 failed to relieve *VSG* repression and the viability of these knockouts demonstrates that base J is nonessential, in vitro at least. This, of course, begs the question of why the biosynthetic pathway for J is retained, although a potential association with strand-switch regions suggests a role in general transcription (see below Ref. 56).

Biochemical probes revealed over 30 years ago that not all *VSG* genes are in the same physical environment, a concept that has yielded important and mechanistic insights. Specific interest in telomeres clearly arises from the demonstration that active *VSG* genes are associated with telomeric regions, and provides an important focus for study. Several *VSG* genes have enhanced DNAse sensitivity, which extends beyond the telomere repeats and VSG open reading frames to encompass additional protein-coding sequences – a subtelomeric structure now known as an ES (Refs 57, 58, 59). DNAse hypersensitivity implies some opening up of chromatin structure, but not necessarily increased transcriptional activity. Trypanosome telomeres appear to cluster at the nuclear periphery, based on the number of puncta detectable with fluorescent probes being substantially less than the number of telomeres (Ref. 60); however, the clusters are dispersed and not concentrated within one or two regions at interphase. There is also potential differential organisation of megabase and minichromosome telomeres, which further adds to the evidence that these two classes of chromosome are distinct in their positioning and behaviour at mitosis (Ref. 26). Furthermore, the arrangement of telomeres varies between life stages, with evidence for increased frequency of peripheral telomere clusters in procyclic forms compared with bloodstream stages, where there is a greater presence within the interior of the nucleus. This relocation was suggested to correlate with repression of *VSG* ESs in insect forms, and to perhaps reflect a more accessible state for even inactive *VSG* ESs in the bloodstream stage (Ref. 31). Critically, it is known that transcription from multiple ESs does take place, with only the 'active' ES mRNAs progressing to maturity (Ref. 61). Taken together, these data suggest a modified, more open chromatin environment surrounding the *VSG* subtelomeric ES, and which is not restricted to the active ES.

The concept of subnuclear positioning and a relationship to chromatin structure in mediating control of *VSG* expression is further supported by demonstrations that a metacyclic *VSG* promoter is more active if placed at a chromosomal internal location compared with the normal subtelomeric location (Ref. 62). Furthermore, investigations of the location of the active *VSG* gene using fluorescence in situ hybridisation demonstrated that the active ES was non-nucleolar, which is highly unusual for a Pol-I-transcribed gene. In addition, the procyclin locus also appears to be perinucleolar, but is located closer to the nucleolus than the active *VSG* gene (Ref. 63). Significantly, exclusion of the putative active ES from the nucleolus was independent of the promoter because replacement of the *VSG* promoter with an rRNA promoter did not lead to incorporation into the nucleolus, suggesting that nonpromoter sequence elements, perhaps DNA zip codes, are required for targeting. Moreover, many *VSG* ESs, presumed to be inactive, were located at more peripheral sites in the nucleus (Ref. 29).

The extranucleolar spot was subsequently demonstrated to contain an RNA Pol I component shortly afterwards and definitive evidence for the specific inclusion of the active ES within this extranucleolar structure, the expression site body (ESB), was obtained (Ref. 30). These studies have led to a simple model for monoallelic *VSG* expression by the inclusion of a single ES within the privileged environment of the ESB. Before the mechanism for targeting of the active ES to the ESB or recruitment of ESB factors to the active ES can be unravelled, detailed characterisation of ESB protein composition and the precise DNA sequences incorporated within the structure need to be determined. The specific telomeric association of a rather unusual variant H3 histone (which is distinct from mammalian H3 variants such as CenH3 that also associate with centromeres) unfortunately also fails to provide a clue here because knockout does not noticeably affect ES silencing (Ref. 64); because this is the only H3V in the *T. brucei* genome, we

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

can confidently exclude H3 variants as having prominent roles in control of ES expression, although the role of this histone appears to be rather extensive (see below). Overall, studies prior to completion of genome sequencing indicated that positioning of specific DNA elements was vital for control of *VSG* gene expression. However, a number of red herrings and perhaps overenthusiastic focus on antigenic variation resulted in a significant level of confusion, and many fundamental questions remained unaddressed.

Genome organisation and segregation

T. brucei has a haploid genome of approximately 32 Mbp distributed across more than 120 chromosomes (Ref. 65). These can be divided into three classes based on size: 11 pairs of megabase chromosomes (MBCs) of at least a megabase in size, 1–5 intermediate-sized chromosomes (ICs) and more than 100 minichromosomes (MCs) of 50–150 kbp (Refs 66, 67). The MBCs contain all the known 'housekeeping' genes, whereas many of the MCs contain *VSG* genes, and so contribute to antigenic variation by providing a potential library of distinct variants. The MCs are highly repetitive in nature, being constructed around a palindromic arrangement of a characteristic 177 bp direct repeat (Ref. 67). The intermediate chromosomes also carry *VSG* genes, although in contrast to the MCs, these are present in ESs (Ref. 68).

In addition to the chromosomes described above, the *T. brucei* nucleus also contains an unknown number of circular extrachromosomal DNAs termed as NlaIII repeat (NR)-elements of up to 400 kbp in size; although present in many strains, these appear to be absent or undetectable in the 927 genome strain (Ref. 69). The NR elements contribute approximately 6% of the total DNA content of the nucleus (Ref. 69), but have no known function. In common with the MCs, they are highly repetitive, being made up of short repeats of variable size delineated by *Nla*III restriction sites.

The chromosomes of *T. brucei* apparently do not condense during mitosis to a detectable degree as in metazoa. However, trypanosome chromosomes do appear to have a degree of subnuclear organisation that exhibits some variation through parasite development. As described above, the telomeres are differentially positioned during both the cell and life cycles, and the MCs, which also possess telomeres and therefore make up a significant proportion of the telomere signal, clearly contribute to this positioning and are predominantly peripheral at interphase (Refs 26, 31). In contrast, the circular NR elements are dispersed throughout the interphase nucleus and show no colocalisation with MCs (Ref. 69). In spite of their different structures, the linear MCs and circular NR elements are segregated by a similar mechanism, which is distinct from that of the larger chromosomes. The MBCs appear to be segregated by a spindle–kinetochore interaction through microtubules that are peripheral to the main central spindle (Ref. 26). The MCs and NR elements remain tightly associated with the central spindle and rapidly segregate to the poles of the daughter nuclei, after which the MCs maintain an asymmetrical distribution and the NR elements become dispersed throughout the interphase nucleus (Refs 26, 69).

The lack of a correlation between the small number of kinetochores observable by electron microscopy in the *T. brucei* nucleus (Ref. 28) and the much larger number of chromosomes, the manner of MC segregation and the apparent absence of many centrosomal or kinetochore protein-coding genes from *T. brucei*, led to the proposal of a lateral stacking model (Ref. 70). This envisages lateral association between the MCs and the central spindle through kinesin motor proteins, which would not require kinetochores. Recently Kif13-1, a member of the kinesin-13 family, has been identified. It is essential for regulation of spindle assembly in *T. brucei* (Refs 71, 72). However, motor proteins that could fulfil the role postulated in the lateral stacking model are yet to be identified. The association between the MCs and the spindle might also be dependent upon the 177 bp repeat that makes up a substantial proportion of their sequence and, although unrelated to this repeat, the highly repetitive nature of the NR elements might allow them to use a similar segregation mechanism. Although the details of MC and NR element segregation are still unclear, their segregation occurs with a high degree of fidelity (Refs 69, 73).

Nuclear structure and transcription in trypanosomes

It is well recognised that the nucleus of many taxa possesses internal organisation, with the most

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

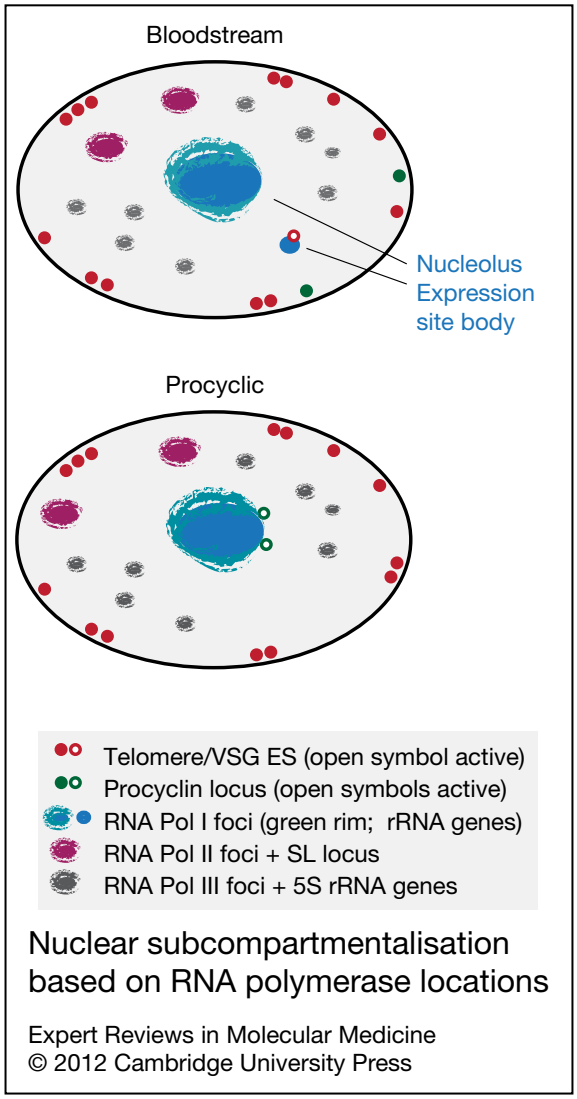

Figure 3. Nuclear subcompartmentalisation based on RNA polymerase locations. (*See next column for legend.*)

Figure 3. Nuclear subcompartmentalisation based on RNA polymerase locations. (*See previous column for figure.*) Foci for the three RNA polymerases are shown as teal, magenta and grey (RNA Pol I, II and III, respectively. The nuclear envelope is in black, positions of telomeres in red and procyclin loci are in green. Actively transcribed procyclin or telomeric expression sites (ESs) are shown as open symbols whereas transcriptionally repressed elements are closed symbols. Note that RNA Pol I is restricted to the nucleolus and ES body and the rRNA loci are shown as a fuzzy rim at the periphery of the nucleolus–RNA-Pol-I territory, whereas the major sites for RNA Pol II are coincident with the spliced leader loci. In the case of the latter, the localisation is very extensive because of the extensive involvement in transcription of protein-coding genes, and only the most concentrated regions are shown. RNA Pol III is found in numerous small foci throughout the nucleoplasm, but is excluded from the nucleolus. Diagram is based on data from Refs 22, 75, 76, 122. The peripheral localisation of *VSG* ES telomeres and procyclin loci in bloodstream cells is more speculative.

prominent features being the nucleolus and differentiation of DNA into euchromatin and heterochromatin based around density in electron micrographs. Furthermore, localisation of specific proteins and DNA sequence elements has led to the suggestion that the metazoan nucleus is organised into regions rich in transcription, splicing and other RNA processing activities. RNA Pol III and tRNA genes cluster and nuclear regions containing highly expressed genes and RNA Pol II have been identified. However, there is uncertainty as to whether these constitute structurally distinct regions of the nucleus that recruit specific loci, that is 'factories', or self-assembling transcription zones that form around transcriptionally active loci (Ref. 74).

Similar to other eukaryotes, the trypanosome nucleolus is constructed around the RNA-Pol-I-transcribed *rDNA* arrays (Ref. 65), although this appears to be more complex because of the presence of RNA-Pol-I-transcribed protein-coding genes (Fig. 3). There is also evidence that RNA Pol II and III exhibit a degree of compartmentalisation in trypanosomes. The large subunit of RNA Pol III localises to several foci in the interphase nucleus (Ref. 75). The large subunit of RNA Pol II is distributed throughout the *T. brucei* nucleus, but also forms several macrofoci (Ref. 76), consistent with the constitutive nature of trypanosome transcription where it generally initiates from divergent strand-switch regions (Ref. 77). *Trans*-splicing adds a common 39 nucleotide 5′-cap containing splice leader (SL) sequence to every mRNA in *T. brucei*, and the SL genes are organised as an array on chromosome 9 with the only recognisable trypanosome RNA Pol II promoters (Fig. 2) (Ref. 78). The macrofoci containing RNA Pol II colocalise with the SL gene array; interestingly, the transcription elongation factors, TFIIS1 and TFIIS2-1, are not enriched at these

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

loci, suggesting that they are primarily involved in the generation of polycistronic RNAs (Ref. 76). RNA Pol II has a similar distribution in the nucleus of the related kinetoplastid *Trypanosoma cruzi*, and although dispersed throughout the nucleus as in *T. brucei*, forms a strong focus that colocalises with the SL gene array in a region that appears to exclude other loci (Refs 79, 80). More recently, the tSNAP complex, which is involved in SL transcript processing (Refs 81, 82) has been localised to a distinct region of the *T. brucei* nucleus (Refs 75, 83). Taken together, these data indicate that an SL transcription and processing factory might exist in the trypanosome nucleus.

Nuclear envelope and nucleoskeletal interactions

In metazoa, the nucleus is structurally supported by a family of intermediate filament proteins called lamins (Ref. 84). These proteins, together with many associated factors, are responsible for the construction of a filamentous lamina. The importance of lamins and the lamina in both the structural maintenance of the nucleus and specific roles in the positioning of the chromosomes is well established. Most significant here are observations that the positioning of telomeres is heavily dependent on the correct expression, targeting and assembly of lamins. Inborn defects of lamin targeting and assembly are associated with multiple pathologies in humans, the best known of which is progeria, which results in premature ageing as a result of misregulation of gene expression, telomere shortening and nuclear structural abnormalities. Lamins are restricted to metazoa, with the result that the molecular players controlling nuclear structure and chromosomal positioning in trypanosomes have remained unknown.

A second important element of the nuclear periphery is the nuclear pore complex (NPC), large multisubunit organelles that span the double-membrane nuclear envelope and which act as the gateway for all nucleocytoplasmic transport (Ref. 85). NPCs are responsible for export of mRNAs and tRNAs, and for import of DNA replication proteins, transcription factors and the proteins that make and modulate chromatin. Recent work indicates a physical association between NPCs and chromatin, providing a further important player in chromosomal organisation at the nuclear periphery.

The vast majority of nucleoporins have been identified in trypanosomes by proteomics (Ref. 86). Sequence searches had suggested little similarity between trypanosomes and higher eukaryotes, but surprisingly the NUP composition of *T. brucei* is similar to animals and fungi, reflecting greater conservation of structure rather than sequence, and also suggesting both an ancient common origin for many NUPs and hinting at potential conservation of functions encompassing both mechanisms of nucleocytoplasmic transport and potential interactions with other nuclear processes. An example of the latter is *T. brucei* NUP92; this protein is a TPR-repeat-containing protein, and uniquely among *T. brucei* NUPs, has a cell-cycle-dependent location, being exclusively at the NPC during interphase and at a polarised location that is likely to be the spindle-anchoring site during mitosis (Fig. 4). This behaviour is highly reminiscent of Mad-1 and Mad-2 and Tpr1–Tpr2 from *S. cerevisiae* and these early experiments suggest that a subset of trypanosome NUPs are mobile, and thus participate in the control of mitotic events (Refs 87, 88, 89). Trypanosome nuclear lamina components have begun to emerge. A giant 450 kDa coiled-coil protein, NUP-1, is a component of fibres at the inner nuclear envelope, and is required to maintain NPC and telomere positioning. Suppression of NUP-1 has highly specific effects on telomeric *VSG* genes, with clear misregulation of normally silenced *VSGs* (Ref. 90). Preliminary evidence suggests that at least two additional players interact with NUP-1; one of these, NUP-2, is also a coiled-coil protein that appears to be part of the NUP-1-associated fibrous network. Together, these data indicate that repression by the specific positioning of genes within regions of peripheral heterochromatin is probably a strategy shared with higher eukaryotes. It will be interesting to connect the functions of the NUP-1–NUP-2 networks with factors controlling *VSG* silencing. Importantly, although the lamins and NUP-1–NUP-2 bear no obvious common ancestry, these studies do indicate that similar, filament-based nucleoskeletons are a basis for much of the higher-order control of nuclear function in animals, fungi and trypanosomes.

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

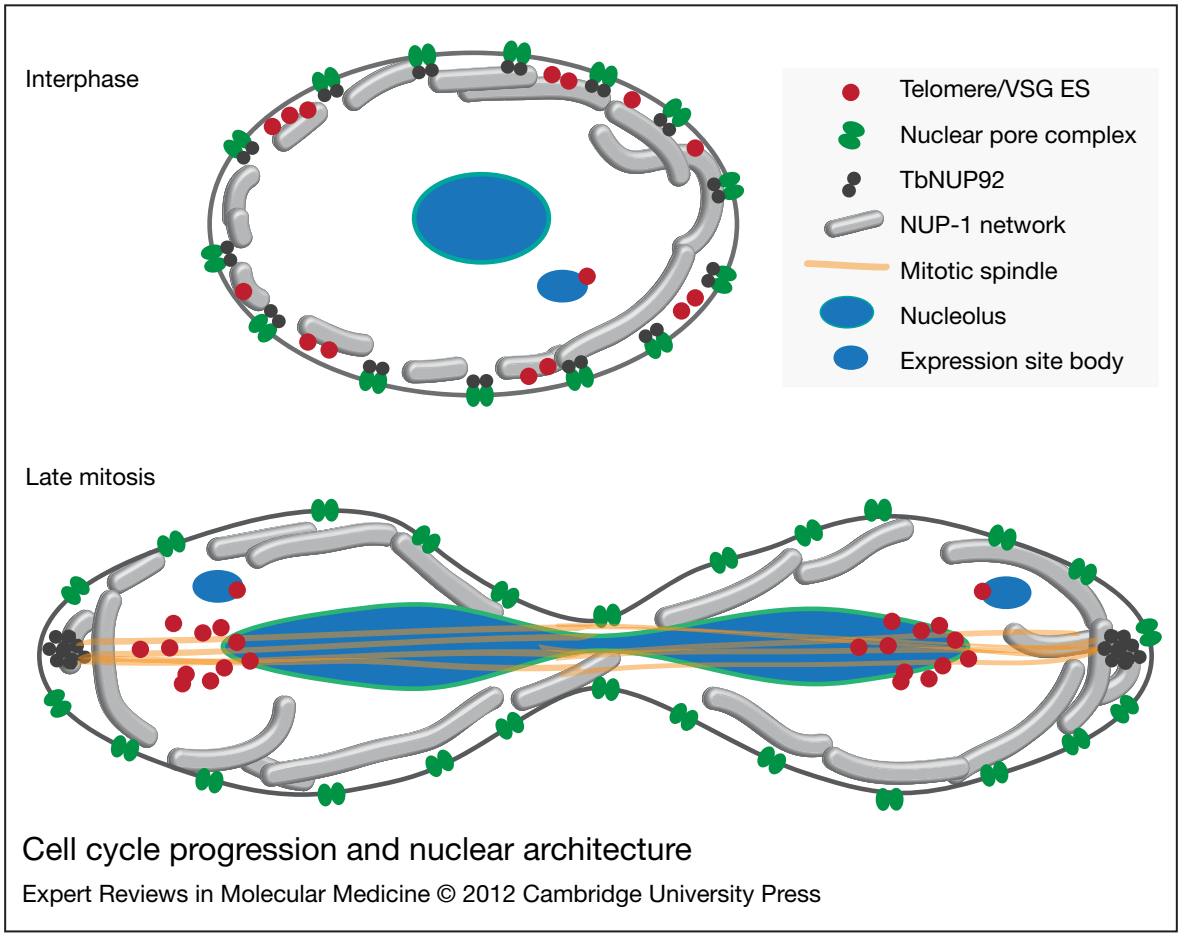

Figure 4. Cell cycle progression and nuclear architecture. Top panel shows an interphase bloodstream stage nucleus in schematic, with the focus on the nuclear pore complex (NPC) (green) the putative lamina protein NUP-1 (grey tubes) and positioning of telomeres (red) within small clusters at the nuclear periphery. A single nucleolus and an expression site body (ESB) are shown, both containing RNA Pol I subunits (blue), and in the case of the ESB, an active *VSG* and associated telomeric region. Note the presence of *T. brucei* NUP92, a probable nuclear basket component at the NPC (dark grey). Lower panel shows a late mitotic nucleus, where the telomeres have nearly completed migration from the midbody to the opposite poles of the daughter nuclei. The mitotic spindle is shown in orange. Note that NUP92 relocates from the NPC to polarised regions that might correlate with the trypanosome spindle pole body, whereas the remaining NPC subunits remain at the nuclear periphery. The ESB replicates late during mitosis, and only following sister chromatid separation, which is delayed with reference to the remaining chromosomes and the inactive ESs (Ref. 27).

Subtelomeres and the control of *VSG* ES silencing

Although the precise mechanisms of allelic exclusion itself are only partially uncovered, our understanding of the mechanism(s) underlying *VSG* silencing has advanced considerably. It appears that two distinct mechanisms operate to repress gene expression at telomeres and *VSG* ESs (Ref. 91). Telomeric silencing was first reported in the budding yeast, *S. cerevisiae* (Ref. 92), but subsequent studies showed that genes located within a few kilobase pairs of a telomere are prone to repression in many eukaryotes. The mechanism has been particularly well characterised in *S. cerevisiae* (see Refs 93, 94). Briefly, telomeric silencing depends upon binding of the telomere-binding protein, repressor-activator protein 1 (RAP1) and consequent recruitment of a silencing complex containing a histone deacetylase, silent information regulator 2 (Sir2). Thus, telomeric heterochromatin is defined, at one level, by hypoacetylated chromatin.

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

Histones both package chromosomes and regulate access to factors involved in a wide range of chromatin transactions, including transcription, DNA replication and repair. Reversible modifications, such as N-terminal acetylation and methylation, extending beyond the nucleosome core, are particularly important for these regulatory functions. Although histone modifications do occur in trypanosomes, they are divergent from higher eukaryote canonical histone modifications (Ref. 32). Chip-seq analysis provided the first direct evidence that specific histone modifications have a role in controlling trypanosome gene expression (Ref. 77), and despite their divergence from higher eukaryotes, trypanosomatid histones are subject to reversible modification, in particular for methylation (Refs 95, 96). Furthermore, several histone acetyltransferases, deacetylases, methyltransferases and demethylases are encoded in the trypanosome genome (Ref. 97). Because histone variants also have roles in transcription regulation and functional specialisation of chromatin domains, similar histone variants were also of interest in trypanosomatids. As noted above, the cenH3 variant is absent, but novel histone variants are present. Trypanosome H2AZ and H2BV are enriched at transcription start sites (TSSs), whereas H3V and H4V are enriched at transcription termination sites (TTSs) (Ref. 77) (Fig. 2). This finding represented a breakthrough in understanding transcriptional control in the apparently promoterless, trypanosome PTUs. The same study also revealed roles for histone acetylation and methylation at initiation sites. Thus, knowledge of trypanosome chromatin biology has advanced considerably in recent years (Ref. 32), but many fundamental questions remain, not least of which is how centromeres are defined in the absence of cenH3.

The first direct indication that telomeric silencing operates in *T. brucei* came from genetic manipulation of the bloodstream form. Using selectable marker genes as reporters inserted adjacent to native telomeres, dramatic repression was observed compared with identical reporters at the active locus (Ref. 98). Notably, this positional effect was developmentally regulated, more consistent with repression by facultative rather than constitutive heterochromatin. It is also important to note a key distinction in the constructs used to interrogate apparent telomere-position effects in *T. brucei* and in other organisms. Promoters recruiting RNA Pol I were used in *T. brucei*, but RNA Pol II promoters were used elsewhere. This reflects the unusual properties of transcription units in trypanosomatids and that *VSGs* are transcribed by RNA Pol I. Because RNA Pol I transcription operates at a much higher rate relative to RNA Pol II transcription, these findings indicated that the positional effect identified at *T. brucei* telomeres was able to repress even these 'strong' promoters.

On the basis of the data at hand, two important questions arise. First, is the telomere itself responsible for silencing, or are additional sequences involved, and second, what factors contribute to this repression? Because it seems that there are two distinct mechanisms operating at these loci, and therefore possibly two distinct sets of factors recruited by distinct sequences, we will consider sequence dependence first. Almost all native telomeres in *T. brucei* appear to have an adjacent *VSG* gene or *VSG* ES. By creating de novo telomeres without *VSGs*, it was found that the telomere itself can exert negative control on local gene expression (Ref. 99), but significantly, repression was severely diminished at only a short distance from the telomere, immediately suggesting that *VSGs* or other features of *VSG* ESs contribute to the more robust and developmentally regulated silencing.

We can also consider the factors known to contribute to this 'basal' (*VSG* and *VSG* ES independent) telomeric silencing and compare them with factors specifically contributing to *VSG* ES silencing. At least 14 proteins are linked to the control of transcription at telomeres in *T. brucei* (Table 1). Telomeric silencing in budding yeast requires Sir2, and consistent with an evolutionarily conserved mechanism, the *T. brucei* Sir2 orthologue SIR2rp1 also contributes to basal telomeric silencing (Ref. 100). Analysis of several other histone acetyltransferases and deacetylases revealed two additional proteins that modulate SIR2rp1-dependent basal telomeric silencing; the histone acetyltransferase HAT1, thought to constrain SIR2rp1 to telomeric domains through histone H4K10 acetylation (Ref. 101), and the deacetylase DAC1, which antagonises SIR2rp1 action at telomeres (Ref. 102). Consistent with the paradigm that *VSG* ES silencing is

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

mechanistically distinct, none of these three factors has a detectable impact on transcription from the *VSG* ES promoter (Refs 100, 101, 102).

Basal telomeric silencing is a heterochromatin-based phenomenon and heterochromatin is clearly a likely participant in *VSG* silencing. Indeed, nucleosomes are specifically depleted from the single active *VSG* ES (Refs 103, 104). So, what factors repress transcription from the *VSG* ES promoter, and how do these differ from basal telomeric silencing? The first protein linked to *VSG* ES silencing was a chromatin-remodelling factor known as Imitation SWI (ISWI) (Ref. 105). ISWI knockdown, however, leads to quite modest derepression, and *VSG* genes remain repressed to a large degree (Ref. 106). Other proteins showing similar knockdown phenotypes include DAC3, a histone deacetylase (Ref. 102), the chromatin chaperone FACT (Ref. 107), a nucleoplasmin-like protein, NLP1 (Ref. 108). Knockdown of a telomere-binding protein, RAP1, also generates a similar phenotype (Ref. 109), but in this case, multiple telomeric *VSGs* achieve sufficient expression levels to be detected at the cell surface, which has not been reported for the other proteins; furthermore, the phenotype also includes the production of additional ESBs, which is more consistent with the idea that the ESB associates around an active ES rather than being a stable factory. Our more recent work indicates a role for the nuclear lamina components NUP-1 and NUP2 in regulating telomeric silencing, suggesting that telomere tethering is important for the control of chromatin structure and modification (Ref. 90).

A role for RAP1 in basal telomeric silencing has not been defined, but on the basis of mechanisms in *S. cerevisiae*, is expected to recruit SIR2rp1, as well as additional factors that spread further along *VSG* ESs. This rather vague model underlines the fact that mechanisms controlling 'spreading' of a repressed domain, probably occurring in *cis* along the DNA strand, are not particularly well understood in any system. Thus, only modest derepression, which in the case of RAP1 is ~100 times less than full activity (Ref. 109), is a common phenotype emerging from knockdowns of candidate regulators for *VSG* ES repression. It is unclear as to whether this reflects redundancy or the presence of an additional major pathway that remains to be uncovered.

Several factors do have a more direct impact on the monoallelic *VSG* expression mechanism. First, DOT1B, a nonessential histone lysine methyltranferase, has a role in in situ switching, where transcription is transferred from one *VSG* ES to another without requiring a DNA rearrangement (Ref. 110). Null *dot1b* cells show very slow in situ switching, suggesting that histone H3K76 methylation is important in inactivating transcription at the original ES, activating transcription at the new ES, or both. Second, chromosomal cohesion appears to have a role in inheritance of the activation status of the *VSG* ES (Ref. 27). Perturbed sister chromatid cohesion disrupts what is normally very stable inheritance, leading to an increased frequency of in situ switching. Third, an intact nuclear lamina also appears to be important for this stable inheritance, which is disrupted in cells depleted for NUP-1 (Ref. 90).

Overall, much of our understanding here comes from candidate-based approaches, and despite being highly informative, are probably limited in scope, providing only restricted snapshots of the overall processes of telomere repression, mechanisms controlling ES activity and broader issues in terms of epigenetic control. Furthermore, because candidate approaches by necessity target proteins with domain similarity to known participants in epigenetic processes in other systems, the identification of putative trypanosome-specific pathways that exploit truly novel factors is only just beginning.

Genome-wide approaches

Recently, more global approaches have become available for the analysis of post-transcriptional gene expression control. For example, ChIP-Chip has been used to describe the activity of the RNA-binding protein DRBD3 in stabilising *T. brucei* mRNAs (Ref. 111) and ChIP-qPCR has been used to show that *T. brucei* ISWI regulates RNA-Pol-I-transcribed loci and is also present at RNA Pol II transcription boundaries (Ref. 106). Furthermore, arginine methyltransferases modify histone tails in higher eukaryotes, among many other probable functions, and thus affect gene regulation. At least five of these enzymes are also present in trypanosomes, and one, *T. brucei* PRMT6, was found to interact directly with histones by immunoprecipitation, and in vitro assays indicate that PRMT6 can methylate histones (Refs 112, 113). Furthermore,

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

PRMT7 can methylate both histones and RNA-binding proteins, suggesting multiple impacts on the gene expression machinery (Ref. 114). Arginine modification of trypanosome histones has been reported in *T. cruzi* (Ref. 115), but the precise contributions of the *T. brucei* PRMT family and potential roles in chromatin maintenance and modification remain to be elucidated.

Most significantly, high-throughput sequencing is significantly advancing our understanding of epigenetics and allows global analysis. Although these studies frequently remain at the candidate level, the results are more wide ranging. For example, J synthesis and maintenance requires two thymidine hydroxylases, JBP1 and JBP2 (Ref. 116). Anti-J immunoprecipitation and high-throughput sequencing showed that JBP1 deposits J at chromosome internal sites and there is some association with strand-switch regions, whereas JBP2 acts at telomeric regions (Ref. 55). This study suggests that JBP1 induces J synthesis, but maintenance is dependent upon JBP2, which binds DNA through a SWI2–SNF2 domain and hence epigenetically remodels chromatin.

Several recent studies have used Illumina/Solexa high-throughput cDNA sequencing (RNA-seq) to identify splice acceptor sites and polyadenylation signals (Refs 117, 118, 119, 120, reviewed in Ref. 121). Together, these studies redefined the major splice acceptor site for most genes and identified a surprisingly large number of alternative splice sites. Moreover, a considerable proportion map within the coding region, with clear potential to alter the sequence of the expressed protein. Mechanisms that underpin splice-site selection, and the nature or even presence of signalling pathways that can influence the frequency of splicing at each site are completely unknown. Again, the detailed analysis of these phenomena is only just beginning.

Conclusions and future directions

The relationship between chromatin structure, gene expression and antigenic variation in trypanosomes is becoming clearer, with many proteins now implicated in the silencing and repression of *VSG* ESs. The genome sequence has done much to refocus the field, with new knowledge and the ability to construct hypotheses based on sound information rather than vague or very incomplete data, as was apparent in earlier studies. There is now excellent evidence placing histone modification as the mediator of telomeric silencing and control of ESs, as well as in the definition of PTUs, which suggests a greater similarity than was previously appreciated between trypanosome TSS regions and conventional eukaryotic promoters, especially in their bidirectional activity. It is also clear that the candidate-based approach still has power because the list is far from exhausted, and further study of orthologues is likely to be highly informative. Falling costs in sequencing mean that genome-wide, rather than target-gene-based, analysis of impacts on transcription is now possible and desirable.

However, several important questions remain: (i) How do the factors operating at telomeres, chromosome internal TSSs and TTSs and at the ESB and other nuclear internal subcompartments interact with each other? Specifically, what are the protein–DNA interaction networks that subtend these functions? (ii) How are these networks incorporated into interactions with the nucleoskeleton, the nuclear envelope and lamina, which in other systems provide much of the assembly framework for high-level organisation of the nucleus, and is also emerging as a participant in transcriptional control and epigenetic memory in higher eukaryotes? (iii) What novel (or taxon-restricted) players are present, and how do they interact and function in the context of the more widely conserved cohorts of gene products? (iv) How are these networks controlled in a spatiotemporal manner, and how does this integrate with life-cycle progression and responses to alterations in the environment? (v) Is the inheritance of *VSG* transcription status by daughter cells based on the simple retention of all bar one ES at the nuclear periphery through mitosis, or is there a more complex mechanism, perhaps related to the inheritance of nutritional memory in *S. cerevisiae*?

Addressing these questions will require unbiased sampling of long-range complex interactions with, typically, the application of proteomic and cell biological approaches to identification of new and established players, assessment of post-translational modifications and identification of the DNA elements with

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

which these factors associate and which they modulate in a functional manner. Major advances in our understanding of mammalian and fungal gene regulatory networks and how they impact at the cellular level can provide a framework, but the novel aspects of trypanosome biology should be our priority; technical advances in nucleic acid sequencing, proteomics and immunoisolation methods and RNAi-based screens will offer many opportunities. Indeed, here is where the novel biology and the understanding of genetic mechanisms that kinetoplastids use to express their surface antigens reside with the potential to provide insight to disease and even therapeutic opportunities.

Acknowledgements and funding

Work in the authors' laboratories is mainly supported by the Wellcome Trust, which is gratefully acknowledged. We are grateful to the anonymous reviewers for constructive comments on the early drafts of this article.

References

1 Chatelain, E. and Ioset, J.R. (2011) Drug discovery and development for neglected diseases: the DNDi model. Drug Design, Development and Therapy 5, 175-181
2 Catarino, L.M. et al. (2001) Classification of trypanosomatids from fruits and seeds using morphological, biochemical and molecular markers revealed several genera among fruit isolates. FEMS Microbiology Letters 201, 65-72
3 Lloyd-Smith, J.O. et al. (2009) Epidemic dynamics at the human–animal interface. Science 326, 1362-1367
4 Sturm, N.R. et al. (2007) Rational sub-division of plant trypanosomes (*Phytomonas* spp.) based on minicircle conserved region analysis. Infection, Genetics and Evolution 7, 570-576
5 Hide, G. (1999) History of sleeping sickness in East Africa. Clinical Microbiology Reviews 12, 112-125
6 Rogers, D. and Randolph, S.E. (1988) Tsetse flies in Africa: bane or boon? Conservation Biology 2, 57-65
7 Duttagupta, S., Gupta, S. and Gupta, A. (2004) Euglenoid blooms in the floodplain wetlands of Barak Valley, Assam, North eastern India. Journal of Environmental Biology 25, 369-373
8 Cavalier-Smith, T. (2010) Kingdoms Protozoa and Chromista and the eozoan root of the eukaryotic tree. Biology Letters 6, 342-345
9 Jensen, B.C. et al. (2009) Widespread variation in transcript abundance within and across developmental stages of *Trypanosoma brucei*. BMC Genomics 10, 482
10 Koumandou, V.L. et al. (2008) The trypanosome transcriptome is remodelled during differentiation but displays limited responsiveness within life stages. BMC Genomics 9, 298
11 Field, M.C. and Carrington, M. (2009) The trypanosome flagellar pocket. Nature Reviews Microbiology 7, 775-786
12 Gadelha, C. et al. (2011) Specializations in a successful parasite: what makes the bloodstream-form African trypanosome so deadly? Molecular and Biochemical Parasitology 179, 51-58
13 MacGregor, P. et al. (2011) Transmission stages dominate trypanosome within-host dynamics during chronic infections. Cell Host and Microbe 9, 310-318
14 Urwyler, S. et al. (2007) A family of stage-specific alanine-rich proteins on the surface of epimastigote forms of *Trypanosoma brucei*. Molecular Microbiology 63, 218-228
15 Lee, M.G. and Van der Ploeg, L.H. (1997) Transcription of protein-coding genes in trypanosomes by RNA polymerase I. Annual Review of Microbiology 51, 463-489
16 Haile, S., Estevez, A.M. and Clayton, C. (2003) A role for the exosome in the in vivo degradation of unstable mRNAs. RNA 9, 1491-1501
17 Kramer, S. and Carrington, M. (2011) *Trans*-acting proteins regulating mRNA maturation, stability and translation in trypanosomatids. Trends in Parasitology 27, 23-30
18 Serpeloni, M. et al. (2011) An essential nuclear protein in trypanosomes is a component of mRNA transcription/export pathway. PLoS One 6, e20730
19 Horn, D. (2008) Codon usage suggests that translational selection has a major impact on protein expression in trypanosomatids. BMC Genomics 9, 2
20 Belli, S.I. (2000) Chromatin remodelling during the life cycle of trypanosomatids. International Journal of Parasitology 30, 679-687
21 Elias, M.C. et al. (2001) Transcription rate modulation through the *Trypanosoma cruzi* life cycle occurs in parallel with changes in nuclear organisation. Molecular and Biochemical Parasitology 112, 79-90
22 Navarro, M., Penate, X. and Landeira, D. (2007) Nuclear architecture underlying gene expression in *Trypanosoma brucei*. Trends in Microbiology 15, 263-270

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

23 Rout, M.P. and Field, M.C. (2001) Isolation and characterization of subnuclear compartments from *Trypanosoma brucei*. Identification of a major repetitive nuclear lamina component. Journal of Biological Chemistry 276, 38261-38271
24 Vickerman, K. and Preston, T.M. (1970) Spindle microtubules in the dividing nuclei of trypanosomes. Journal of Cell Science 6, 365-383
25 Dossin Fde, M. et al. (2008) Automated nuclear analysis of *Leishmania major* telomeric clusters reveals changes in their organization during the parasite's life cycle. PLoS One 3, e2313
26 Ersfeld, K. and Gull, K. (1997) Partitioning of large and minichromosomes in *Trypanosoma brucei*. Science 276, 611-614
27 Landeira, D. et al. (2009) Cohesin regulates *VSG* monoallelic expression in trypanosomes. Journal of Cell Biology 186, 243-254
28 Ogbadoyi, E. et al. (2000) Architecture of the *Trypanosoma brucei* nucleus during interphase and mitosis. Chromosoma 108, 501-513
29 Chaves, I. et al. (1998) Subnuclear localization of the active variant surface glycoprotein gene expression site in *Trypanosoma brucei*. Proceedings of the National Academy of Sciences of the United States of America 95, 12328-12333
30 Navarro, M. and Gull, K. (2001) A pol I transcriptional body associated with *VSG* mono-allelic expression in *Trypanosoma brucei*. Nature 414, 759-763
31 Perez-Morga, D. et al. (2001) Organization of telomeres during the cell and life cycles of *Trypanosoma brucei*. Journal of Eukaryotic Microbiology 48, 221-226
32 Figueiredo, L.M., Cross, G.A. and Janzen, C.J. (2009) Epigenetic regulation in African trypanosomes: a new kid on the block. Nature Reviews. Microbiology 7, 504-513
33 Rando, O.J. and Chang, H.Y. (2009) Genome-wide views of chromatin structure. Annual Review of Biochemistry 78, 245-271
34 Danchin, E. et al. (2011) Beyond DNA: integrating inclusive inheritance into an extended theory of evolution. Nature Reviews. Genetics 12, 475-486
35 Nowacki, M., Shetty, K. and Landweber, L.F. (2011) RNA-mediated epigenetic programming of genome rearrangements. Annual Review of Genomics and Human Genetics 12, 367-389
36 Skinner, M.K. (2011) Environmental epigenetic transgenerational inheritance and somatic epigenetic mitotic stability. Epigenetics 6, 838-842
37 Geyer, P.K., Vitalini, M.W. and Wallrath, L.L. (2011) Nuclear organization: taking a position on gene expression. Current Opinion in Cell Biology 23, 354-359
38 Brickner, D.G. et al. (2007) H2A.Z-mediated localization of genes at the nuclear periphery confers epigenetic memory of previous transcriptional state. PLoS Biology 5, e81
39 Light, W.H. et al. (2010) Interaction of a DNA zip code with the nuclear pore complex promotes H2A.Z incorporation and *INO1* transcriptional memory. Molecular Cell 40, 112-125
40 Engstler, M. and Boshart, M. (2004) Cold shock and regulation of surface protein trafficking convey sensitization to inducers of stage differentiation in *Trypanosoma brucei*. Genes and Development 18, 2798-2811
41 Leung, K.F. et al. (2011) Ubiquitylation and developmental regulation of invariant surface protein expression in trypanosomes. Eukaryotic Cell 10, 916-931
42 Gerrits, H. et al. (2002) The physiological significance of transferrin receptor variations in *Trypanosoma brucei*. Molecular and Biochemical Parasitology 119, 237-247
43 Salmon, D. et al. (2005) *Trypanosoma brucei*: growth differences in different mammalian sera are not due to the species-specificity of transferrin. Experimental Parasitology 109, 188-194
44 Steverding, D. (2003) The significance of transferrin receptor variation in *Trypanosoma brucei*. Trends in Parasitology 19, 125-127
45 Ersfeld, K. (2011) Nuclear architecture, genome and chromatin organisation in *Trypanosoma brucei*. Research in Microbiology 162, 626-636
46 Beck, J.S. and Walker, P.J. (1964) Antigenicity of *Trypanosome Nuclei*: evidence that DNA is not coupled to histone in these Protozoa. Nature 204, 194-195
47 Astolfi Filho, S., Martins de Sá, C. and Gander, E.S. (1980) On the chromatin structure of *Trypanosoma cruzi*. Molecular and Biochemical Parasitology 1, 45-53
48 Rubio, J., Rosado, Y. and Castaneda, M. (1980) Subunit structure of *Trypanosoma cruzi* chromatin. Canadian Journal of Biochemistry 58, 1247-1251
49 Hecker, H. et al. (1989) Instability of the nuclear chromatin of procyclic *Trypanosoma brucei brucei*. Molecular and Biochemical Parasitology 37, 225-234
50 Burri, M. et al. (1993) Biochemical and functional characterization of histone H1-like proteins in procyclic *Trypanosoma brucei* brucei. Parasitology Research 79, 649-659

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

51 Horn, D. (2001) Nuclear gene transcription and chromatin in *Trypanosoma brucei*. International Journal of Parasitology 31, 1157-1165
52 Schlimme, W. et al. (1993) *Trypanosoma brucei brucei*: differences in the nuclear chromatin of bloodstream forms and procyclic culture forms. Parasitology 107, 237-247
53 Gommers-Ampt, J.H. et al. (1993) ®-*D*-glucosyl-hydroxymethyluracil: a novel modified base present in the DNA of the parasitic protozoan *T. brucei*. Cell 75, 1129-1136
54 van Leeuwen, F. et al. (1997) Localization of the modified base J in telomeric *VSG* gene expression sites of *Trypanosoma brucei*. Genes and Development 11, 3232-3241
55 Borst, P. and Sabatini, R. (2008) Base J: discovery, biosynthesis, and possible functions. Annual Review of Microbiology 62, 235-251
56 Cliffe, L.J. et al. (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 Research 38, 3923-3935
57 Kooter, J.M. et al. (1988) Boundaries of telomere conversion in *Trypanosoma brucei*. Gene 69, 1-11
58 Pays, E., Lheureux, M. and Steinert, M. (1981) The expression-linked copy of a surface antigen gene in *Trypanosoma* is probably the one transcribed. Nature 292, 265-267
59 Pays, E. et al. (1989) The genes and transcripts of an antigen gene expression site from *T. brucei*. Cell 57, 835-845
60 Chung, H.M. et al. (1990) Architectural organization in the interphase nucleus of the protozoan *Trypanosoma brucei*: location of telomeres and mini-chromosomes. EMBO Journal 9, 2611-2619
61 Vanhamme, L. et al. (2000) Differential RNA elongation controls the variant surface glycoprotein gene expression sites of *Trypanosoma brucei*. Molecular Microbiology 36, 328-340
62 Graham, S.V., Terry, S. and Barry, J.D. (1999) A structural and transcription pattern for variant surface glycoprotein gene expression sites used in metacyclic stage *Trypanosoma brucei*. Molecular and Biochemical Parasitology 103, 141-154
63 Landeira, D. and Navarro, M. (2007) Nuclear repositioning of the *VSG* promoter during developmental silencing in *Trypanosoma brucei*. Journal of Cell Biology 176, 133-139
64 Lowell, J.E. and Cross, G.A. (2004) A variant histone H3 is enriched at telomeres in *Trypanosoma brucei*. Journal of Cell Science 117(Pt 24), 5937-5947
65 Berriman, M. et al. (2005) The genome of the African trypanosome *Trypanosoma brucei*. Science 309, 416-422
66 Melville, S.E. et al. (2000) The molecular karyotype of the megabase chromosomes of *Trypanosoma brucei* stock 427. Molecular and Biochemical Parasitology 111, 261-273
67 Wickstead, B., Ersfeld, K. and Gull, K. (2004) The small chromosomes of *Trypanosoma brucei* involved in antigenic variation are constructed around repetitive palindromes. Genome Research 14, 1014-1024
68 Ersfeld, K., Melville, S.E. and Gull, K. (1999) Nuclear and genome organization of *Trypanosoma brucei*. Parasitology Today 15, 58-63
69 Alsford, N.S. et al. (2003) The identification of circular extrachromosomal DNA in the nuclear genome of *Trypanosoma brucei*. Molecular Microbiology 47, 277-289
70 Gull, K., Alsford, S. and Ersfeld, K. (1998) Segregation of minichromosomes in trypanosomes: implications for mitotic mechanisms. Trends in Microbiology 6, 319-323
71 Chan, K.Y., Matthews, K.R. and Ersfeld, K. (2010) Functional characterisation and drug target validation of a mitotic kinesin-13 in *Trypanosoma brucei*. PLoS Pathogens 6, e1001050
72 Wickstead, B. et al. (2010) The expanded Kinesin-13 repertoire of trypanosomes contains only one mitotic Kinesin indicating multiple extra-nuclear roles. PLoS One 5, e15020
73 Alsford, S. et al. (2001) Diversity and dynamics of the minichromosomal karyotype in *Trypanosoma brucei*. Molecular and Biochemical Parasitology 113, 79-88
74 Sutherland, H. and Bickmore, W.A. (2009) Transcription factories: gene expression in unions? Nature Reviews. Genetics 10, 457-466
75 Alsford, S. and Horn, D. (2011) Elongator protein 3b negatively regulates ribosomal DNA transcription in African trypanosomes. Molecular and Cellular Biology 31, 1822-1832
76 Uzureau, P. et al. (2008) Identification and characterization of two trypanosome TFIIS proteins exhibiting particular domain architectures and differential nuclear localizations. Molecular Microbiology 69, 1121-1136
77 Siegel, T.N. et al. (2009) Four histone variants mark the boundaries of polycistronic transcription units

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

in *Trypanosoma brucei*. Genes and Development 23, 1063-1076
78 Gilinger, G. and Bellofatto, V. (2001) Trypanosome spliced leader RNA genes contain the first identified RNA polymerase II gene promoter in these organisms. Nucleic Acids Research 29, 1556-1564
79 Dossin Fde, M. and Schenkman, S. (2005) Actively transcribing RNA polymerase II concentrates on spliced leader genes in the nucleus of *Trypanosoma cruzi*. Eukaryotic Cell 4, 960-970
80 Schenkman, S., Pascoalino Bdos, S. and Nardelli, S.C. (2011) Nuclear structure of *Trypanosoma cruzi*. Advances in Parasitology 75, 251-283
81 Das, A. et al. (2005) Trypanosomal TBP functions with the multisubunit transcription factor tSNAP to direct spliced-leader RNA gene expression. Molecular and Cellular Biology 25, 7314-7322
82 Schimanski, B., Nguyen, T.N. and Gunzl, A. (2005) Characterization of a multisubunit transcription factor complex essential for spliced-leader RNA gene transcription in *Trypanosoma brucei*. Molecular and Cellular Biology 25, 7303-7313
83 Lustig, Y. et al. (2007) Spliced-leader RNA silencing: a novel stress-induced mechanism in *Trypanosoma brucei*. EMBO Reports 8, 408-413
84 Shimi, T. et al. (2010) Nuclear lamins in cell regulation and disease. Cold Spring Harbor Symposia on Quantitative Biology 75, 525-531
85 Capelson, M., Doucet, C. and Hetzer, M.W. (2010) Nuclear pore complexes: guardians of the nuclear genome. Cold Spring Harbor Symposia on Quantitative Biology 75, 585-597
86 DeGrasse, J.A. et al. (2009) Evidence for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic ancestor. Molecular and Cellular Proteomics 8, 2119-2130
87 Iouk, T. et al. (2002) The yeast nuclear pore complex functionally interacts with components of the spindle assembly checkpoint. Journal of Cell Biology 159, 807-819
88 Niepel, M. et al. (2005) The nuclear pore complex-associated protein, Mlp2p, binds to the yeast spindle pole body and promotes its efficient assembly. Journal of Cell Biology 170, 225-235
89 Strambio-de-Castillia, C., Blobel, G. and Rout, M.P. (1999) Proteins connecting the nuclear pore complex with the nuclear interior. Journal of Cell Biology 144, 839-855
90 DuBois, K.N. et al. (2012) NUP-1 Is a large coiled-coil nucleoskeletal protein in trypanosomes with lamin-like functions. PLoS Biology 10, e1001287
91 Horn, D. and McCulloch, R. (2010) Molecular mechanisms underlying the control of antigenic variation in African trypanosomes. Current Opinion in Microbiology 13, 700-705
92 Gottschling, D.E. et al. (1990) Position effect at *S. cerevisiae* telomeres: reversible repression of Pol II transcription. Cell 63, 751-762
93 Grunstein, M. (1997) Molecular model for telomeric heterochromatin in yeast. Current Opinion in Cell Biology 9, 383-387
94 Tham, W.H. and Zakian, V.A. (2002) Transcriptional silencing at *Saccharomyces* telomeres: implications for other organisms. Oncogene 21, 512-521
95 Alsford, S. and Horn, D. (2004) Trypanosomatid histones. Molecular Microbiology 53, 365-372
96 Janzen, C.J. et al. (2006) Unusual histone modifications in *Trypanosoma brucei*. FEBS Letters 580, 2306-2310
97 Ivens, A.C. et al. (2005) The genome of the kinetoplastid parasite, *Leishmania major*. Science 309, 436-442
98 Horn, D. and Cross, G.A. (1995) A developmentally regulated position effect at a telomeric locus in *Trypanosoma brucei*. Cell 83, 555-561
99 Glover, L. and Horn, D. (2006) Repression of polymerase I-mediated gene expression at *Trypanosoma brucei* telomeres. EMBO Reports 7, 93-99
100 Alsford, S. et al. (2007) A sirtuin in the African trypanosome is involved in both DNA repair and telomeric gene silencing but is not required for antigenic variation. Molecular Microbiology 63, 724-736
101 Kawahara, T. et al. (2008) Two essential MYST-family proteins display distinct roles in histone H4K10 acetylation and telomeric silencing in trypanosomes. Molecular Microbiology 69, 1054-1068
102 Wang, Q.P., Kawahara, T. and Horn, D. (2010) Histone deacetylases play distinct roles in telomeric *VSG* expression site silencing in African trypanosomes. Molecular Microbiology 77, 1237-1245
103 Figueiredo, L.M. and Cross, G.A. (2010) Nucleosomes are depleted at the *VSG* expression site transcribed by RNA polymerase I in African trypanosomes. Eukaryotic Cell 9, 148-154
104 Stanne, T.M. and Rudenko, G. (2010) Active *VSG* expression sites in *Trypanosoma brucei* are depleted of nucleosomes. Eukaryotic Cell 9, 136-147
105 Hughes, K. et al. (2007) A novel ISWI is involved in *VSG* expression site downregulation in African trypanosomes. EMBO Journal 26, 2400-2410
106 Stanne, T.M. et al. (2011) TbISWI regulates multiple polymerase I (Pol I)-transcribed loci and is present

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

at Pol II transcription boundaries in *Trypanosoma brucei*. Eukaryotic Cell 10, 964-976
107 Denninger, V. et al. (2010) The FACT subunit TbSpt16 is involved in cell cycle specific control of *VSG* expression sites in *Trypanosoma brucei*. Molecular Microbiology 78, 459-474
108 Narayanan, M.S. et al. (2011) NLP is a novel transcription regulator involved in *VSG* expression site control in *Trypanosoma brucei*. Nucleic Acids Research 39, 2018-2031
109 Yang, X. et al. (2009) RAP1 is essential for silencing telomeric variant surface glycoprotein genes in *Trypanosoma brucei*. Cell 137, 99-109
110 Figueiredo, L.M., Janzen, C.J. and Cross, G.A. (2008) A histone methyltransferase modulates antigenic variation in African trypanosomes. PLoS Biology 6, e161
111 Estevez, A.M. (2008) The RNA-binding protein TbDRBD3 regulates the stability of a specific subset of mRNAs in trypanosomes. Nucleic Acids Research 36, 4573-4586
112 Fisk, J.C. and Read, L.K. (2011) Protein arginine methylation in parasitic protozoa. Eukaryotic Cell 10, 1013-1022
113 Fisk, J.C. et al. (2010) TbPRMT6 is a type I protein arginine methyltransferase that contributes to cytokinesis in *Trypanosoma brucei*. Eukaryotic Cell 9, 866-877
114 Fisk, J.C. et al. (2009) A type III protein arginine methyltransferase from the protozoan parasite *Trypanosoma brucei*. Journal of Biological Chemistry 284, 11590-11600
115 da Cunha, J.P. et al. (2006) Post-translational modifications of *Trypanosoma cruzi* histone H4. Molecular and Biochemical Parasitology 150, 268-277
116 Cliffe, L.J. et al. (2009) JBP1 and JBP2 are two distinct thymidine hydroxylases involved in J biosynthesis in genomic DNA of African trypanosomes. Nucleic Acids Research 37, 1452-1462
117 Kolev, N.G. et al. (2010) The transcriptome of the human pathogen *Trypanosoma brucei* at single-nucleotide resolution. PLoS Pathogens 6, e1001090
118 Nilsson, D. et al. (2010) Spliced leader trapping reveals widespread alternative splicing patterns in the highly dynamic transcriptome of *Trypanosoma brucei*. PLoS Pathogens 6, e1001037
119 Siegel, T.N. et al. (2010) Genome-wide analysis of mRNA abundance in two life-cycle stages of *Trypanosoma brucei* and identification of splicing and polyadenylation sites. Nucleic Acids Research 38, 4946-4957
120 Veitch, N.J. et al. (2010) Digital gene expression analysis of two life cycle stages of the human-infective parasite, *Trypanosoma brucei gambiense* reveals differentially expressed clusters of co-regulated genes. BMC Genomics 11, 124
121 Siegel, T.N. et al. (2011) Gene expression in *Trypanosoma brucei*: lessons from high-throughput RNA sequencing. Trends in Parasitology 27, 434-441
122 Daniels, J.P., Gull, K. and Wickstead, B. (2010) Cell biology of the trypanosome genome. Microbiology and Molecular Biology Reviews 74, 552-569

Further reading, resources and contacts

DuBois, K.N. et al. (2012) NUP-1 Is a large coiled-coil nucleoskeletal protein in trypanosomes with lamin-like functions. PLoS Biology (in press)

Demonstration that the large coiled-coil protein NUP-1 performs a similar role in trypanosomes as lamins in metazoa; suggests that despite huge divergence that overall mechanisms for defining peripheral heterochromatin may be conserved across eukaryotes and be involved in control of VSG expression.

Siegel, T.N. et al. (2009) Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. Genes and Development 23, 1063-1076

A groundbreaking study that demonstrates preferential localisation of specific histones and histone variants with transcriptional start and stop sites, with profound implications for the organisation of trypanosome transcriptional units and the role of chromatin in defining such units.

Yang, X. et al. (2009) RAP1 is essential for silencing telomeric variant surface glycoprotein genes in *Trypanosoma brucei*. Cell 137, 99-109

The conserved telomere-binding protein RAP-1 is required for peripheral localisation of silent VSG ESs, suggesting that part of the mechanism for telomeric silencing is highly conserved across euakryotes.

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012

Features associated with this article

Figures

Table

Citation details for this article

Sam Alsford, Kelly duBois, David Horn and Mark C. Field (2012) Epigenetic mechanisms, nuclear architecture and the control of gene expression in trypanosomes. Expert Rev. Mol. Med. Vol. 14, e13, May 2012, doi:10.1017/erm.2012.7

Accession information: doi:10.1017/erm.2012.7; Vol. 14; e13; May 2012