Monoallelic expression and epigenetic inheritance sustained by a *Trypanosoma brucei* variant surface glycoprotein exclusion complex

Joana Faria¹, Lucy Glover¹,², Sebastian Hutchinson¹,³, Cordula Boehm¹, Mark C. Field¹ & David Horn¹

The largest gene families in eukaryotes are subject to allelic exclusion, but mechanisms underpinning single allele selection and inheritance remain unclear. Here, we describe a protein complex sustaining variant surface glycoprotein (VSG) allelic exclusion and antigenic variation in *Trypanosoma brucei* parasites. The VSG-exclusion-1 (VEX1) protein binds both telomeric VSG-associated chromatin and VEX2, an ortholog of nonsense-mediated-decay helicase, UPF1. VEX1 and VEX2 assemble in an RNA polymerase-I transcription-dependent manner and sustain the active, subtelomeric VSG-associated transcription compartment. VSG transcripts and VSG coats become highly heterogeneous when VEX proteins are depleted. Further, the DNA replication-associated chromatin assembly factor, CAF-1, binds to and specifically maintains VEX1 compartmentalisation following DNA replication. Thus, the VEX-complex controls VSG-exclusion, while CAF-1 sustains VEX-complex inheritance in association with the active-VSG. Notably, the VEX2-orthologue and CAF-1 in mammals are also implicated in exclusion and inheritance functions. In trypanosomes, these factors sustain a highly effective and paradigmatic immune evasion strategy.
Antigenic variation and host immune evasion by parasites causing malaria\(^1\), giardiasis\(^2\) and African trypanosomiasis\(^3\) depend upon monoallelic and switchable variant surface antigen expression. Indeed, allelic exclusion governs the expression of many of the largest known gene-families, typically encoding cell-surface proteins in protozoa and mammals, and has major impacts on health and disease. In mammals, odour detection depends upon olfactory receptor exclusion\(^4\) and B- and T-cell specificity depends upon immunoglobulin and surface receptor exclusion, respectively\(^5\). Selection of a single allele from a large gene family is thought to be inherently stochastic rather than deterministic and generates tremendous cellular diversity, upon which adaptive mechanisms operate. Notably, the diversity-generating exclusion systems of protozoan antigenic variation, and the host immune response, are mutually adaptive and in direct competition. Despite intense study, the molecular mechanisms facilitating the selection and maintenance of a single active allele, coordinated with heritable exclusion of all others, remain poorly understood.

The African trypanosome, Trypanosoma brucei, is transmitted among mammalian hosts by tsetse-flies and, due to effective immune evasion, causes chronic and lethal infections, specifically sleeping-sickness in humans and nagana in cattle. In these parasites RNA polymerase-I (pol-I) transcribes a single, dominant, telomeric variant surface glycoprotein (VSG) gene\(^3\) as part of a polycistronic transcription unit. Pol-I transcription initiates at several sites but this is attenuated at all but one site\(^6\). T. brucei can switch to activate a new telomeric VSG expression site (ESs) but, significantly, switching occurs at low-frequency and all but one of the ESs typically remain ‘silent’, despite possessing all cis-elements required for expression\(^7\).

The active VSG-ES is associated with an extranucleolar pol-I transcription factory, known as the expression site body (ESB)\(^8,9\). This VSG-ES is depleted of nucleosomes\(^10,11\) and nucleosome-hypersensitive chromatin persists at this locus even if transcription is blocked\(^12\). An HMG chromatin protein is enriched at both the ESB and the nucleolus\(^12,13\) and a focus of SUMOylation associates with the ESB\(^14\). In addition, histones\(^15\) and histone variants\(^16\), telomere-binding proteins, a histone H3-K76 methyltransferase, a histone deacetylase, a chromatin remodeler, histone chaperones, nuclear lamina components, cohesin and the insoluble phosphate pathway all impact VSG-ES silencing to varying degrees (reviewed in ref.\(^17\)). These factors may impact chromatin-dependent silencing or transcriptional permissivity but, notably, a direct role in gene selection has not been demonstrated. More recently, the first factor specifically enriched in association with the active VSG-ES, VSG exclusion 1 (VEX1), was identified\(^17\).

Here, we identify and characterise a VEX1–VEX2 complex. The complex assembles a sub-nuclear domain in a transcription-dependent manner and maintains VSG allelic exclusion by negatively controlling transcription of other telomeric VSGs. Inheritance of VSG exclusion requires maintenance of the VEX-complex during S-phase, which depends upon the conserved chromatin assembly factor, CAF-1.

**Results**

A subtelomere- and VSG-associated VEX-complex. We previously described \(T. brucei\) VEX1 (Tb927.11.16920), the only known protein specifically enriched in association with the ESB\(^17\). VEX1, therefore, is at the heart of the nuclear subdomain mediating antigenic variation in trypanosomes, but VEX1 lacks orthologs in other cell types, it remained unclear why VSG-exclusion was only partially perturbed following VEX1-knockdown and it remained unclear whether other factors were involved in VSG-exclusion.

To identify VEX1 chromatin interactions, we affinity-purified VEX1\(^{9,99}\)-associated chromatin and deep-sequenced the enriched DNA. Reads were accurately aligned to individual VSG-ESs by uniqueness filtering, using MapQ > 118. An examination of protein coding sequences within the hemizygous subtelomeric VSG-ESs\(^9,19\) revealed VSG-2, the active VSG, as the most enriched gene (Fig. 1a, Supplementary Data 1, sheet 1). We also observed particularly strong enrichment of the region downstream of this gene (Fig. 1a). This is consistent with a focus of VEX1 that is adjacent to, rather than coincident with, the extranuclear focus of pol-I at the active VSG-ES\(^17\). Notably, DNA immediately downstream of several silent ES-associated VSGs was also enriched (Fig. 1a, Supplementary Fig. 1). This is also consistent with our model proposing that conserved VSG-associated sequences, including the telomeric repeats, participate in VEX1- and homology-dependent VSG-silencing\(^17\). Thus, chromatin interactions connect VEX1 to the conserved sequences associated with telomeric VSGs.

Next, we affinity-purified VEX1\(^{9,99}\)-associated proteins using cryo-milling and high-affinity nanobodies. The procedure was carried out initially in insect-stage \(T. brucei\) (Supplementary Fig. 2a, b), for which protocols were established\(^20\), and then in bloodstream form \(T. brucei\) (Fig. 1b). Quantitative proteomic analysis (Fig. 1b, Supplementary Fig. 2a, b, Supplementary Data 1, sheet 2) revealed tag-dependent enrichment of green fluorescent protein (GFP) and the same set of five proteins in four independent experiments; VEX1 was enriched, as expected, but also Tb927.11.13380, an ortholog of the nonsense-mediated mRNA-decay ATP-dependent superfamily 1-type helicases, UPF1/SMG2/NAM7/Rem1\(^{21}\) (Fig. 1c, Supplementary Fig. 2c–e), for which protocols were established\(^20\), and then in bloodstream form \(T. brucei\) (Fig. 1b). The putative \(T. brucei\) canonical UPF1 has been studied but it remains unclear whether classical nonsense-mediated decay operates\(^22\). VEX2 contains a UPF1-related helicase domain towards the C-terminus, including the canonical motifs involved in ATP binding and hydrolysis and nucleic acid binding (Supplementary Fig. 2d). Structural prediction suggested the presence of an α-solenoid architecture at the N-terminus (Supplementary Fig. 2e), in place of the typical UPF2-interacting domain\(^23\); α-solenoids are frequently involved in protein–protein or protein–RNA interactions. The other three proteins that displayed tag-dependent enrichment were all three components of chromatin assembly factor 1 (CAF-1; CAF-1a, Tb927.8.3980; CAF-1b, Tb927.8.3980; CAF-1c, Tb927.11.4970; Fig. 1b, Supplementary Fig. 2a), the evolutionary conserved heterotrimeric replication-associated histone chaperone\(^24\). Notably, no pol-I components were enriched.

Super-resolution fluorescence microscopy revealed that, like VEX1, VEX2 localised to a single major sub-nuclear focus (detected in ~80% of nuclei) that was replicated and segregated at all but one site\(^6\). T. brucei can switch to activate a new telomeric VSG expression site (ESs) but, significantly, switching occurs at low-frequency and all but one of the ESs typically remain ‘silent’, despite possessing all cis-elements required for expression\(^7\).

VEX-complex compartmentalisation is transcription-dependent. VEX1 foci are dispersed in cells treated with transcription inhibitors, indicating that active transcription is required for assembly\(^17,25\). However, it was unclear from these studies whether VEX1 was degraded or redistributed throughout the nucleoplasm. Protein blotting indicated that neither VEX1 (Fig. 2a) nor VEX2 (Fig. 2b) were degraded following pol-I specific transcription
inhibition with BMH-21. Indeed, super-resolution confocal microscopy revealed that both VEX1 (Fig. 2a, Supplementary Fig. 3a) and VEX2 (Fig. 2b, Supplementary Fig. 3b) displayed significant levels of redistribution across the nucleus following inhibition of pol-I transcription, or inhibition of general transcription with actinomycin D (Supplementary Fig. 3c–f). Thus, recruitment of VEX1 and VEX2 to the VSG transcription compartment requires active pol-I transcription. VEX1 and VEX2 proteins remain within the nucleus when transcription is prevented, suggesting availability for retargeting once pol-I transcription reinitiates.

**VEX2 depletion results in multi-VSG expression.** VEX1 was initially identified from a high-throughput RNA interference (RNAi) screen for subtelomeric loss-of-silencing. Another genome-wide RNAi screen indicated minimal fitness-cost following VEX1-knockdown, but a significant and major fitness-cost following VEX2-knockdown, specifically in bloodstream-form cells and not insect-stage cells. This likely explains why only VEX1, and not VEX2, was identified in the loss-of-silencing screen, and indicates a bloodstream-form fitness-cost following VEX2-knockdown, specifically in bloodstream-form cells and not insect-stage cells. To explore the function of VEX2, we assembled three independent inducible RNAi expression sites (Fig. 3, Supplementary Fig. 4b). Knockdown was associated with a significant growth defect (Fig. 3a), but little perturbation of cell cycle distribution (Supplementary Fig. 4c). In these cells, normally silent VSG-6 was strongly derepressed after only 24 h, as assessed by protein immunoblotting (Fig. 3b), immunofluorescence microscopy (Fig. 3c) and immunostaining followed by flow cytometry (Fig. 3d). Notably, abundance of the initially active VSG was also clearly reduced following VEX1-knockdown, but a significant and major fitness-cost following VEX2-knockdown, specifically in bloodstream-form cells and not insect-stage cells. This likely explains why only VEX1, and not VEX2, was identified in the loss-of-silencing screen, and indicates a bloodstream-form fitness-cost specific function for VEX2; VSG is not normally expressed in the insect-stage.

**Fig. 1** A subtelomere- and VSG-associated VEX-complex. **a** Affinity purification of VEX1::myc-associated chromatin followed by sequencing (ChIP-seq). Enrichment traces over active (VSG-2) and silent (VSG-6, VSG-13) VSG-ESs with 1 kb non-overlapping bins; red circles, promoters; red boxes, VSGs; grey boxes, expression-site associated genes (ESAGs). Box-plot depicting enrichment of the indicated CDSs. Centre lines show the medians; box limits indicate the 25th and 75th percentiles; whiskers extend from the 10th to the 90th percentile, outliers are shown. Silent VSGs, n = 18; ESAGs, n = 129. Metagene plot and the associated heat-map for active and silent telomeric VSGs. The red bar on the metagene plot indicates the location of a highly conserved sequence in the VSG 3'-untranslated region. **b** VEX1::GFP immunoprecipitation and proteomics. GFP-tag dependent enrichment from bloodstream form T. brucei. emPAI, exponentially modified protein abundance index. The inset shows immunoprecipitates and tag-dependent VEX1 (red) and VEX2 (green) bands highlighted in a silver-stained gel. **c** Schematic representation of VEX1 (Tb927.11.16920), VEX2 (Tb927.11.13380) and CAF-1 subunits (a, Tb927.8.3980; b, Tb927.10.7050; c, Tb927.11.4970) indicating conserved domains and motifs. **d** Immunofluorescence analysis during the cell cycle and protein-blot analysis of mycVEX2 expression. N nucleus; K kinetoplast (mitochondrial genome). **e**, **f** Immunofluorescence-based colocalisation studies of VEX2::GFP and VEX1::myc (e) or a nucleolar (N) and ESB marker (pol-l, largest subunit), a nucleolar marker (NOG1) or Histone H3 (f). In e, G1 or G2 cells are shown. **d–f** DNA was counter-stained with DAPI; the images correspond to maximal 3D projections of 0.1 μm stacks; scale bars 2 μm; the data are representative of independent biological replicates and independent experiments. Source data are provided as a Source Data file.
after 48 h of VEX2-knockdown (Fig. 3b) and VSG-6 appeared to accumulate in intracellular compartments as well as at the surface (Fig. 3c, right-hand side). To survey additional VSGs, we used quantitative proteomics of proteins released from the cell-surface. The active VSG on wild-type cells, VSG-2, displayed a relative abundance of 99.9%. Following VEX2-knockdown, VSG-2 remained the most abundant VSG, but VSG-6 and several additional ES-associated VSGs were derepressed (Fig. 3e, Supplementary Data 1, sheet 3). Because antigenic variation occurs at low frequency, T. brucei clones are typically homogeneous and express a single VSG (Fig. 3c–e). Following VEX2-knockdown, however, >90% of cells simultaneously expressed VSG-2, and VSG-6 at the cell surface as determined by microscopy (Fig. 3c) and flow cytometry (Fig. 3d). Taken together, these results indicate major disruption of allelic exclusion and the simultaneous expression of multiple VSGs on the surface of individual cells following VEX2-knockdown.

**VEX2 coordinates allele-specific VSG-ES transcription.** To further explore gene expression defects following VEX2-knockdown, we analysed the transcriptome, comparing three independent knockdown strains to three wild-type sub-clones (Supplementary Fig. 5, Supplementary Data 1, sheets 4–5). Since RNAi may not degrade the entire mRNA, we show efficient disruption of VEX2 transcript in Fig. 4a. Among approximately 7600 genes encoded in the genome, 166 displayed >5-fold and significantly (p < 10^{-2}) increased expression after 24 h knockdown (Fig. 4b–e), including every known pol-I transcribed protein-coding locus. There are 18 known silent, telomeric, pol-I promoter-associated VSGs in the genome; 13 in polycistronic expression sites and 5 in monocistronic expression sites, with the latter only normally activated in metacyclic cells in the insect salivary gland\(^\text{18}\). Total transcript abundance from this full set of 18 VSGs (Fig. 4b, c), and 130 silent expression-site associated genes (ESAGs) (Fig. 4d), was increased by 212 and 32-fold,
Differential control had been identified in VSGs with abundant transcripts, and no factor responsible for this was previously identified; VSG mRNA is >140-fold more abundant than the mean ESAG mRNA levels. Our transcriptome analysis now indicates post- or co-transcriptional suppression of these ESAGs mediated by VEX2. We speculate, based on sequence similarity of VEX2 to UPF1, that negative control by VEX2 may be related to nonsense-mediated mRNA-decay.

We next asked what impact VEX2-knockdown and VSG-ES derepression had on RNA pol-I localisation. Immunofluorescence microscopy revealed a remarkable loss of the nucleolar pol-I signal after only 24 h of knockdown and a substantial loss of detectable extranucleolar pol-I signals; the remaining cells displayed no detectable signal. DNA was counter-stained with DAPI; the images correspond to maximal 3D projections of 0.1 μm stacks; scale bars 2 μm. Error bar, SD; *p < 0.05; **p < 0.01 (two-tailed paired Student’s t test); data are averages from three independent biological replicates and representative of independent experiments. Source data are provided as a Source Data file.

**VEX-complex knockdown yields a severe exclusion defect.**

VEX1 knockdown yielded a moderate VSG derepression defect relative to the VEX2 knockdown phenotype (Figs. 3 and 4). We, therefore, knocked down both VEX1 and VEX2 simultaneously and carried out a similar analysis to that described above. VEX1–VEX2 double knockdown was associated with a very severe growth defect that was cytocidal after 72 h (Fig. 5a). In these cells, silent VSG-6 was again strongly derepressed, as assessed by protein immunoblotting; the active VSG signal was derepressed and redistributed to extranucleolar sites. Immuno–fluorescence and protein-blot analysis of pol-I expression and localisation following VEX2 RNAi. The bar graph indicates the proportion of nuclei (n ≥ 100) with nucleolar and/or extranucleolar pol-I signals; the remaining cells displayed no detectable signal. DNA was counter-stained with DAPI; the images correspond to maximal 3D projections of 0.1 μm stacks; scale bars 2 μm. Error bar, SD; *p < 0.05; **p < 0.01 (two-tailed paired Student’s t test); data are averages from three independent biological replicates and representative of independent experiments. Source data are provided as a Source Data file.

**Fig. 4 VEX2 coordinates allele-specific VSG-ES transcription.** a–e RNA-seq analysis following VEX2-knockdown (24 h). Total number of reads at VEX2 locus in the parental population or following VEX2 RNAi (target fragment indicated). a VEX2 RNAi. b–e Values are averages from three independent RNAi strains relative to wild-type controls and numbers in brackets indicate increase in transcript abundance for each gene cohort. Red square, VSG-2: black circle, VEX2. Lower panels show transcript abundance at the silent VSG-6, active VSG-2 and procyclin loci, respectively. Red circles, promoters; red boxes, VSGs; grey boxes, other genes. RPM, reads per kilobase per million. f Immunofluorescence and protein-blot analysis of pol-I expression and localisation following VEX2 RNAi. The bar graph indicates the proportion of nuclei (n ≥ 100) with nucleolar and/or extranucleolar pol-I signals; the remaining cells displayed no detectable signal. DNA was counter-stained with DAPI; the images correspond to maximal 3D projections of 0.1 μm stacks; scale bars 2 μm. Error bar, SD; *p < 0.05; **p < 0.01 (two-tailed paired Student’s t test); data are averages from three independent biological replicates and representative of independent experiments. Source data are provided as a Source Data file.

respectively. Even the ESAGs at the active transcribed locus (n = 18) were significantly (p < 10−50) upregulated (9-fold, Fig. 4d), as were the non-telomeric, pol-I transcribed procyclin (associated) genes (Fig. 4e); procyclin genes produce abundant surface proteins that are normally expressed only in insect mid-gut stage cells. Highlighting the magnitude of the VEX2-knockdown phenotype, derepressed VSGs contributed >24% of total VSG transcripts, compared to only 0.1% in control cells (Supplementary Data 1, sheets 4–5); the super-abundant active VSG mRNA is >200-fold more abundant than the average mRNA encoding ribosomal proteins. Notably, several previously silent VSGs in ESs with dual promoters (VSG-6, 8, 15 and 17; see ref. 7) were particularly derepressed and produced the four most abundant cellular mRNAs, after VSG-2; these VSGs were also readily detected at the cell-surface following VEX2-knockdown (Fig. 3e). The same VSG-ESs were activated at high rates during VSG switching27. Thus, an additional VSG-ES promoter likely facilitates the derepression or activation of a VSG-ES. We conclude that VEX2 coordinates VSG allele exclusion and has a profound impact on differential VSG expression.

We did not anticipate increased expression of ESAGs at the active VSG-ES in response to VEX2-knockdown (Fig. 4d). However, it has long been known that ESAGs upstream of VSGs, despite co-transcription in the same polycistron, yield far less abundant transcripts, and no factor responsible for this differential control had been identified previously; VSG mRNAderepression had on RNA pol-I localisation. Immunofluorescence microscopy revealed a remarkable loss of the nucleolar pol-I signal after only 24 h of knockdown and a substantial loss of detectable extranucleolar pol-I foci; protein blotting indicated that pol-I was still present following VEX2 knockdown (Fig. 4f). These results suggest that derepressed VSG-ESs deplete the nucleolus and the ESB of pol-I, redistributing the polymerase to multiple extranucleolar sites.
once again clearly reduced after 48 h of knockdown (Fig. 5b). Following 48 h of VEX1-VEX2 double knockdown, immunofluorescence microscopy revealed that almost all cells expressed both VSG-2 and VSG-6, while >20% of cells simultaneously expressed VSG-2, VSG-6 and VSG-3 (Fig. 5c); VSG-13 was also expressed by >20% of these cells (Supplementary Fig. 6a). We further monitored VEX1-knockdown cells over an extended period, during which cells remain viable. This revealed >25% of the population expressing both VSG-2 and VSG-6 throughout the time-course, with no evidence for switching to VSG-6 expression (Supplementary Fig. 6b). Thus, VEX-complex knockdown disrupts allelic exclusion and does not simply increase the VSG switching rate. Quantitative proteomic analysis following VEX1–VEX2 knockdown revealed derepression of 15 pol-I promoter-associated VSGs (Fig. 5d, Supplementary Data 1, sheet 3). A comparison of relative VSG expression revealed that derepressed VSGs contributed approx. 40% of total VSG following VEX1–VEX2 knockdown (Fig. 5e, Supplementary Data 1, sheet 7).
3) compared to only approximately 10% of total VSG following VEX2 knockdown. As for VEX2-knockdown (Fig. 4f), double knockdown was associated with loss of both the nucleolar pol-I signal and detectable extranucleolar pol-I foci (Supplementary Fig. 6c).

Transcriptome analysis following VEX1–VEX2 knockdown (Fig. 5f, Supplementary Fig. 7, Supplementary Data 1, sheets 4–5) revealed a similar pattern of derepression as observed above for VEX2 knockdown alone (Fig. 5g–j). Again, all known pol-I transcribed protein-coding loci were derepressed, and previously silent VSGs linked to dual VSG-ES promoters produced the four most abundant cellular mRNAs, after VSG-2 (Fig. 5g). These data demonstrate that the VEX-complex coordinates VSG allelic exclusion. The allelic exclusion system collapses following depletion of the VEX-complex, yielding multi-VSG expression, highly heterogeneous VSG coats and cell-death.

VEX-complex interactions impact abundance and location. We next asked what impact VEX1 and VEX2 have on each other. VEX1-knockdown had no detectable impact on VEX2 expression or focal localisation (Fig. 6a). In contrast, VEX1 over-expression substantially diminished VEX2 abundance; VEX1 was highly overexpressed in most nuclei (~80%) and VEX2 foci were no longer detectable by immunofluorescence (Fig. 6b). These data...
likely explain the previously reported more severe VSG derepression phenotype following VEX1 over-expression compared with knockdown17 and also indicate that VEX1 can limit VEX2 abundance; this latter feature may be important for preventing activation of a second VSG. We also tested the impact of VEX2-knockdown on VEX1 localisation and found that VEX1 was redistributed across multiple sub-nuclear puncta following loss of VEX2 (Fig. 6c). Thus, VEX1 association with the VSG transcription compartment is VEX2-dependent.

VEX-complex compartmentalisation is stage-specific. Insect-stage T. brucei cells neither transcribe VSG nor exhibit allelic exclusion. Since we show above that VEX-complex association with the VSG transcription compartment requires pol-I transcription, we asked whether VEX2 association with this compartment is absent in insect-stage cells. Like VEX117, VEX2 was indeed redistributed through the nuclear compartment during differentiation (Fig. 7a). Protein blotting revealed substantial redistribution of VEX2 as the VSG signal diminished (Fig. 7a) and these differentiated cells displayed multiple VEX2-foci, some of which were coincident with VEX1-foci (Fig. 7b). Redistributed VEX2 was similarly observed in long-term established insect-stage cells. Thus, both VEX1 and VEX2 associate with the VSG transcription compartment in a pol-I transcription-dependent and life cycle stage-specific manner.

Colocalisation of CAF-1 and the VEX-complex during S phase. We next considered the interaction between the VEX-complex and CAF-1; all three components of this conserved chromatin chaperone were co-immunoprecipitated with VEX1 (see Fig. 1b). DNA replication presents both an opportunity to retain or reset epigenetic states. We, therefore, asked if CAF-1 plays a role in the inheritance of the VEX-complex-dependent epigenetic state during DNA replication; the active VSG-ES is replicated early during S-phase while the silent VSG-ESs are replicated later24. Significantly, S-phase specific VSG-ES derepression was previously reported following CAF-1b knockdown15. By analysing the localisation of a tagged CAF-1b subunit, either CAF-1bGFP or CAF-1bmyc, we observed major CAF-1b foci, despite an overall punctate nuclear distribution. Notably, these foci co-localised with the sub-nuclear compartment defined by both VEX1 (Fig. 8a, Supplementary Movie 2) and VEX2 (Fig. 8b) and coincident signals were significantly enriched in S-phase cells (Fig. 8a, b); the interaction between the VEX-complex and CAF-1 was also confirmed by immunoprecipitation followed by protein blotting (Fig. 8c).

VEX-complex reassembly and inheritance requires CAF-1. By analysing phenotypes associated with CAF-1b knockdown, we confirmed two further specific predictions in support of CAF-1 dependent inheritance of the VEX-complex. First, VEX1 distributed across multiple sub-nuclear puncta following CAF-1b knockdown (12 h, Fig. 9a); while VEX2 notably remained primarily compartmentalised in a single focus in each nucleus under the same conditions (Fig. 9b). Indeed, we observed more intense VEX2 foci following CAF-1b knockdown, and protein blotting confirmed that VEX2 abundance is specifically increased when CAF-1b is depleted (Fig. 9c, Supplementary Fig. 8a, b). Thus, CAF-1 limits VEX2 abundance, an effect that may be enhanced by excess VEX1 (see above). Second, following CAF-1b knockdown (see Fig. 9d), transcriptome analysis revealed specific derepression of VSGs and other pol-I transcribed protein-coding genes (Fig. 9e–i, Supplementary Fig. 8c, d, Supplementary Data 1, sheets 4–5). Following CAF-1b knockdown for 24 h (see Supplementary Fig. 8c), 115 genes displayed >5-fold and significantly (p < 10^-2) increased expression, >80% of which were also significantly increased following VEX2 knockdown (Supplementary Data 1, sheets 4, 5). In this case, we observed a more pronounced increase in total transcript abundance for promoter-proximal VSGs (89-fold, n = 5, Fig. 9g) relative to promoter-distal VSGs (9-fold, n = 5, Fig. 9g). Similarly, we observed a more pronounced increase in total transcript abundance for promoter-proximal ESAGs (48-fold, n = 34, Fig. 9b) relative to promoter-distal ESAGs (4-fold, n = 96, Fig. 9b). Procyclin (associated genes) were also derepressed (6-fold, Fig. 9b), while ESAGs at the active locus displayed significantly (p < 10^-3) reduced expression (44 ± 19%, n = 12; Fig. 9b, Supplementary Data 1, sheets 4–5), consistent with increased VEX2 abundance and enhanced VEX2-mediated negative control at this locus (see above). Thus, CAF-1 knockdown has a major impact on pol-I transcribed protein-coding loci; promoter-proximal genes may be particularly susceptible to derepression due to a genome-wide histone chaperone defect.

A model for VSG allelic exclusion by the VEX-complex. Finally, combining RNA-seq data, we compared the relative contributions of VEX1, VEX2 and CAF1-b to VSG silencing (Fig. 10a) and also calculated relative contributions to an allelic exclusion index; the
**Fig. 8** Colocalisation of CAF-1 and the VEX-complex during S phase. **a, b** Immunofluorescence-based colocalisation studies of VEX1myc/CAF-1bmyc and VEX2GFP/CAF-1bmyc. Representative images from G1 and S phase cells are shown. Nuclei (≥ 70) were scored for CAF-1b puncta coincident with VEX1myc/VEX2GFP foci. DNA was counter-stained with DAPI; scale bars, 2 μm. Error bars, SD; *p < 0.05; **p < 0.01; ***p < 0.001 (one-way ANOVA test). Values are averages from two independent biological replicates and are representative of independent experiments. The images on the right-hand side correspond to maximal 3D projections of 0.1 μm stacks. **c** Co-immunoprecipitation of VEX2GFP/CAF-1bmyc followed by protein-blot analysis. Green circle, GFP; black circles, myc; I, input; E, elution. Source data are provided as a Source Data file.

**Fig. 9** VEX-complex reassembly and inheritance requires CAF-1. VEX1myc (a), but not mycVEX2 (b), is distributed following tetracycline (Tet) induced CAF-1b knockdown (12 h), as assessed by immunofluorescence analysis. DNA was counter-stained with DAPI; scale bars, 2 μm. The images correspond to maximal 3D projections of 0.1 μm stacks. Error bars, SD; ns, not significant; **p < 0.01 (two-tailed paired Student’s t test). Values are averages from three independent biological replicates and are representative of independent experiments (the remaining cells displayed no detectable signal). **c** Protein-blot analysis of CAF-1bmyc, mycVEX2 and VEX1myc expression following CAF-1b knockdown. EF1α was used as a loading-control. **d–i** RNA-seq analysis following CAF-1b knockdown (24 h). Total number of reads across CAF-1b locus in the parental population or following CAF-1b RNAi. **e–i** Values are averages from two independent RNAi strains relative to wild-type controls. Panel **e** shows transcript abundance at the silent VSG-6 ES. Red circles, promoters; red boxes, VSGs; grey boxes, other genes. RPM/RPKM, reads per (kilobase per) million (e–i). Panels **f–i** highlight different gene cohorts and numbers in square brackets indicate increase/decrease in total transcript abundance for each category; red square, VSG-2; black circle, CAF-1b. Promoter-proximal VSGs and ESAGs (6, 7 and 10) are within 5 kbp of a promoter, distal VSGs are 46+/-4 kbp from the promoter. Source data are provided as a Source Data file.
ratio of reads from the active VSG relative to all 18 silent expression-site associated VSGs (Fig. 10b). VEX2 clearly makes the major contribution to VSG silencing (Fig. 10a, b), establishing a 1000-fold expression differential in wild-type cells, which, following VEX2 or VEX-complex depletion, is diminished to a 2–4-fold differential (Fig. 10b). As illustrated by our model (Fig. 10c), we conclude that the VEX-complex associates with the active VSG in a pol-I transcription-dependent manner and thereby coordinates nuclear pol-I compartments and VSG exclusion. VEX1 compartmentalisation is VEX2-dependent, while retention during DNA replication is also CAF-1 dependent.

Discussion

An improved understanding of immune evasion by pathogens has profound importance to health and disease. Despite intense study, mechanisms underlying stochastic activation of one allele and the heritable exclusion of all others remain to be fully defined. This has been the case for antigenic variation in the parasites that cause malaria,2, giardiasis,2 and African trypanosomiasis;3 and also in mammals where olfactory receptor allelic exclusion underpins odour detection4 and where immunoglobulin T cells5.

Several observations suggest potentially shared mechanisms with allelic exclusion in mammals. The VEX2-related UPF1 orthologue in humans (Rent1) influences early nuclear events in mRNA biogenesis, including splicing23. Mammalian Rent1 is enriched at telomeres and negatively controls the association between telomeric repeat-containing RNA (TERRA) transcripts and chromatin at these sites; it has been proposed that these transcripts promote heterochromatin assembly, similar to Xist RNA promotion of X-chromosome inactivation.25 In addition, murine UPF1 is required for X-inactivation and formation of Xist RNA domains, a function involving Xist splicing control23. Taken together, these findings suggest a conserved role for this helicase family in regulating lncRNA-chromatin associations, and possibly in gene expression choices. The enrichment of VEX1 and VEX2 in association with the active VSG-ES is also reminiscent of...
nuclear sub-structures in other eukaryotic cells, such as PML bodies and nuclear speckles. It remains unclear how the activity state of a gene is reproduced on daughter chromatids following DNA replication. The conserved hetero-trimeric histone H3–H4 specific chaperone, CAF-1, participates in chromatin assembly directly behind the passing DNA polymerase during DNA replication, integrating DNA synthesis with conservative reassembly of chromatin in eukaryotes. This involves reincorporating the majority of maternal histones within 400 bp of their pre-replication locus and locus-specific bookmarking by epigenetic regulatory complexes to maintain transcriptional programs. Mammalian CAF-1 sustains somatic cell type-specific and epigenetic memory by interacting with heterochromatin protein 1a, a histone H3 methyltransferase and KRAB-ZFP-associated protein 1; this complex is also associated with sub-nuclear foci. Human CAF-1 can also activate transcription independent of its histone deposition function and the conserved CAF-1 subunit-associated DREAM complex plays a role in specifying and maintaining olfactory receptor gene expression in Drosophila. Thus, in addition to housekeeping roles in chromatin assembly, several lines of evidence suggest an evolutionary conserved role for CAF-1 in sustaining site-specific molecular memories of epigenetic states. Our findings indicate that this can involve CAF-1 dependent and locus-specific compartmentalisation of non-histone, chromatin-associated regulators as epigenetic marks. Consistent with this view, a cohesin-dependent delay in active VSG sister chromatid separation facilitates epigenetic inheritance. Thus, premature segregation of sister chromatids following cohesin depletion may yield free VEX1 and present opportunities for activation of other VSG-ESs.

VEX2 makes the greatest contribution to differential VSG expression, as assessed at mRNA and protein levels, while loss of both VEX1 and VEX2 leads to collapse of the exclusion system. Transcription sustains VEX-complex assembly at the active locus and this sequestration through transcription may also drive the establishment of a single active site. The VEX-complex may also be self-limiting, in the sense that CAF-1 and VEX1 can negatively control VEX2 expression; this may be important to minimise activation at other sites and hence secure selectivity. We previously demonstrated homology-dependent silencing by VEX1 and suspect that communication among VSG alleles requires a trans-acting RNA component. Although the current findings demonstrate robust negative control of silent VSGs by the VEX-complex, a more detailed dissection of the proposed positive and negative controls, and their coordination to achieve allelic exclusion, will require further study.

In summary, the crosstalk among VSG-ESs and inheritance of allelic exclusion in trypanosomes requires transcription, the VEX-complex and CAF-1 dependent partitioning of VEX1. These factors collectively sustain a specific assembly in association with the active VSG, resulting in exclusion of all other VSG alleles. CAF-1 links the VEX-complex to the active site to produce a molecular memory. Our findings reveal the factors underpinning a winner-takes-all paradigm for the establishment, maintenance and inheritance of nuclear bodies and allele-specific epigenetic states. To our knowledge, this is the first characterisation of a protein complex directly responsible for single gene choice in an allelic exclusion system.

**Methods**

**T. brucei growth and manipulation.** Bloodstream-form *T. brucei*, Lister 427 and 2T1 cells, both wild-type with respect to VEX1, VEX2 and CAF-1 subunits, were grown in BHM-11 medium and genetically manipulated using electroporation. Cytoxins were used at 2, 2.5 and 10 µg ml⁻¹ for selection of recombinant clones; and at 1, 1, and 2 µg ml⁻¹ for maintaining those clones, respectively. Cumulative growth curves were generated from cultures seeded at 10⁶ cells ml⁻¹, in BHM-11 medium and diluted back to 10⁶ cells ml⁻¹ as necessary. For differentiation of bloodstream form to procyclic form, 2 × 10⁶ cells were resuspended in DTM medium, 2.5 mg ml⁻¹ of haemin, 300 mM cis-aconitate and incubated at 27 °C. Established procyclic form T. brucei, Lister 427 cells were grown in SDM-79 at 27 °C and genetically manipulated using electroporation as described previously. Blasticidin was used at 10 and 2 µg ml⁻¹ for selection and maintenance, respectively.

**Plasmids.** For RNAi primers were selected from ORF sequences using CLC Viewer v. 7.8 and BLAST analysis to minimise potential off-target effects. A specific RNAi target fragment for VEX2 (Tb927.11.13380, 471 bp) was amplified and cloned into pPRα[3,4]. The VEX1 (Tb927.11.16920, 574 bp) and CAF-1b fragments were excised following digestion by Ascl. For epitope-tagging at the native locus, a 710 bp fragment of VEX2 was amplified and cloned into pNATAG[45] to add an N-terminal 6× c-myc or GFP tag and a fragment of 918 bp was amplified and cloned into pNATAG[45] to add a C-terminal GFP tag. The vectors were linearised with Mool and HpsI, respectively. The VEX1[12myc]c and CAF-1b[12myc]C-terminal tagging vectors were linearised with SphI and Ndel, respectively. The VEX1 and CAF-1b GFP C-terminal tagging vectors were made by replacing the 12× c-myc tag and were also linearised with SphI and Ndel, respectively. The VEX1-C terminal myc-tag overexpression cassette was excised prior to electroporation by digesting with Ascl. Linearised RNAi and overexpression constructs, under the control of tetracycline-inducible promoters, were transfected into 2T1 cells, which allow for targeting to a single genomic locus validated for robust inductive expression.

**ChIp-seq.** ChIP and cell lysis[46] was carried out with the following modifications. Briefly, 2 × 10⁸ *T. brucei* bloodstream form cells expressing a C-terminal 12-myc tagged endogenous copy of VEX1 were cross-linked with 1% formaldehyde for 20 min at RT. DNA was sonicated using a Bioruptor (Diagenode) with sonication beads (Diagenode, C01020031) for 5 cycles of 30 s on/30 s off. C-terminal 12-myc VEX1 was immunoprecipitated with a -Myct antibody coupled to Dynabeads Protein G (2.8 µm). Antibody coupling to the Dynabeads was carried out according to the manufactures recommendations. The beads were then washed with RIPA buffer (50 mM HEPES-KOH, pH 7.5, 550 mM NaCl, 1 mM EDTA, 1% NP-40, 0.7% Na-Deoxycholate) and eluted DNA was purified by phenol chloroform extraction and ethanol precipitation. Reads were mapped to bloodstream VSG-ES[7] and metacyclic VSG-ES[47] from the Lister 427 strain. Bowtie 2-mapping[49] was with the parameters --very-sensitive --no-discordant --phred33, and with a MapQ value of >1. Alignment files were manipulated with SAMtools[50]. Alignments were inspected with the Artemis genome browser[51]. Telsmorphic VSG coding sequences with a mapping quality filter of >1 were aligned using deepdepts computeMatrix scale-regions. The bedgraph for log2 fold change was generated using deepTools[25]. Files were assembled in Illustrator. The heat map was generated using deepTools. Reads were counted using BEDTools[53], counting reads in non-overlapping 1 kb bins. Files were normalised by counting total read counts per library prior to fold change calculation for each bin. Locus maps were generated by exporting vector graphics views or regions of interest from Artemis genome browser and manipulation in Adobe illustrator. The box plot was generated in Graphpad Prism v7.0.

**Affinity-enrichment of VEX-interactors.** 1 × 10¹⁰ insect or bloodstream form *T. brucei* cells with or without a C-terminal GFP-tagged endogenous copy of VEX1 were washed three times in ice cold PBS with protease inhibitors (Roche, EDTA free) and subsequently cryomilled into a fine grindate, in a planetary ball mill (Retsch); protein–protein interactions are effectively preserved under these conditions. Six aliquots of 50 mg of frozen grindate were thawed into ice cold lysis buffer (20 mM HEPES, pH 7.4, 1 mM MgCl₂, 10 µM CaCl₂, protease inhibitor cocktail) containing either 100 mM NaCl or 250 mM citrate and either 0.1% Tween, 0.1% Brij58, 0.1% Triton X-100 or 0.1% Chaps, and pipetted to homogeneity. The samples were then subjected to sonication (three pulses at 60% amp) and placed on ice. Samples were spun for 10 min at 20,000g at 4 °C to pellet the debris. The supernatant was removed and added to a 1.5 ml eppendorf containing dimeric a-GFP nanobodies[44] conjugated to magnetic beads (Dynabeads—Thermo Scientific) and agitated at 4 °C for 2 h. The samples were then placed on a magnetic rack and washed three times with ice-cold lysis buffer. Samples were eluted with a non-reducing sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) loading buffer or equivalent. The resulting proteins were fractionated by 1D SDS-PAGE and visualised by silver staining.

**Quantitative mass spectrometry.** Affinity-enriched samples were run 2 cm into a NuPAGE® Novex Bis–Tris 10% gel with NuPAGE® MOPS SDS running buffer (Life Technologies) and subjected to overnight (16 h) trypsin digestion (Modified Sequencing Grade, Roche). Peptides were then extracted, dried in a SpeedVac (Thermo Scientific), resuspended in 50 µl 1% formic acid, cleaned and transferred to high-performance liquid chromatography (HPLC) vials. 5 µl was typically analysed on a Q Exactive HF Hybrid Quadrupole-Orbitrap Mass Spectrometer.
Fig. 5c, cells were analysed by confocal microscopy, using a Leica TCS SP8 confocal laser scanning microscope and the Leica Application Suite X (LASX) software (1:1000). For the colocalisation studies in Fig. 8a, b and the quantification in Fig. 5c, cells were analysed by confocal microscopy, using a Leica TCS SP8 confocal laser scanning microscope and the Leica Application Suite X (LASX) software (1:1000).

Flow cytometry. Flow cytometry was carried out according to standard procedures. The primary antibodies were as follows: rat α-VSG-2 (1:10,000) and rabbit α-VSG-6 (1:10,000). Secondary antibodies were goat α-rat Alexa Fluor 467 and goat α-rabbit Alexa Fluor 488 (both 1:2000). DNA was stained with propidium iodide at 10 µg ml⁻¹. Samples were analysed on a BD LSRSFortessa (BD Biosciences) and data were visualised using FlowJo software. More than 40,000 events were analysed to determine the percentage of cells in each quadrant.

Transcriptome analysis. RNA-seq analysis was performed using 2T1 cells and uninduced or induced clones of V2E (24 and 48 h), VEX/VEX (24 and 48 h) or CAFl-1b (CAFl). RNAi. Briefly, polyadenylated transcripts were enriched using poly-dT beads and reverse-transcribed before sequencing on a HiSeq platform (Illumina). Reads were mapped to a hybrid assembly consisting of the T. brucei 927 reference genome plus the bloodstream VSG-E5 and metacyclic VSG-E5s from the Lister 427 strain. Bowtie 2-mapping was with the parameters --very-sensitive --no-discard --phred33. Alignment files were manipulated with SAMTools. Per-gene read counts were derived using the Artemis genome browser, MapQ > 1.8. Read counts were normalised using edgeR and differential expression was determined with classic edgeR. RPKM values were derived from normalised read counts in edgeR. Base pair resolution plots were generated using an in-house script.

Statistical analysis. All statistical analyses were performed using GraphPad Prism software (version 7.0), except the transcriptomic analysis (described above).

Microscopy. Immunofluorescence microscopy was carried out according to standard protocols, using 12-well 5 mm (Thermo Scientific) or 18-well μ-slides (Ibidi) for wide field, and confocal microscopy, respectively. For super-lowest speed; pinhol microscopy, the cells were attached to poly-l-lysine treated coverslips (Ibidi) for wide field microscopy, using a Zeiss Axiovert 200 M microscope with an AxioCam MRm camera and the ZEN Pro software (Carl Zeiss, Germany). The images were acquired as z-stacks (0.1–0.2 µm) and in the case of the wide field microscopy, further deconvolved using the fast iterative algorithm in Zen Pro. For all quantifications, images were acquired with the same settings and equally processed. Corrected total nuclear fluorescence = integrated density – (area of selected nucleus × mean nuclear fluorescence). The images were opened, adjusted for brightness and contrast and the areas of selected nuclei were analysed using an in-house script.

Data availability

RNA-seq and ChIP-seq data have been deposited in the European Nucleotide Archive with Accession nos. PRJEB21615 and PRJEB25352, respectively. Proteomics data have been deposited in the PRoteomics IDentiﬁcations (PRIDE) database www.ebi.ac.uk/pride. ProteomeXchange accession no. PXD013304. The source data underlying Figs. 1d, 2a, b, 3a–e, 4f, 5a, 6a–c, 7a, 8a, 9a–c, 10a, 11b, Supplementary Data 3–5, 9a, 12a, 16a, 8a, b, d are provided as a Source Data file.

References

1. Guijtiett, J. & Scherf, A. Silence, activate, poise and switch! Mechanisms of antigenic variation in African trypanosomes. Nat. Rev. Microbiol. 13, 718–726 (2015).
2. Gargantini, P. R., Serradell Mdel, C., Rios, D. N., Tenaglia, A. H. & Lujan, H. D. Antigenic variation in the intestinal parasite Giardia lamblia. Curr. Opin. Microbiol. 32, 52–58 (2016).
3. Horn, D. Antigenic variation in African trypanosomes. Mol. Biochem. Parasitol. 195, 123–129 (2014).
8. Chaves, I. et al. Subnuclear localization of the active variant surface glycoprotein expression site in *Trypanosoma brucei*. *Proc. Natl Acad. Sci. USA* 111, 8941–8946 (2014).

9. Kassem, A., Pays, E. & Vanhamme, L. Transcription is initiated on silent variant surface glycoprotein expression sites despite monoallelic expression in *Trypanosoma brucei*. *Proc. Natl Acad. Sci. USA* 111, 4004–4009 (2014).

10. Figueiredo, L. M. & Cross, G. A. Nucleosomes are depleted at the VSG expression site transcribed by RNA polymerase I in African trypanosomes. *Eukaryot. Cell* 9, 148–154 (2010).

11. Stanne, T. M. & Rudenko, G. Active transcription factor GATA1. *Nucleic Acids Res.* 43, 1109–1118 (2016).

12. Aarest-Branco, F., Pimenta, S. & Figueiredo, L. M. A transcription-independent epigenetic mechanism is associated with antigenic switching in *Trypanosoma brucei*. *Nucleic Acids Res.* 44, 3131–3146 (2016).

13. Narayanay, M. S. & Rudenko, G. TDP1 is an HMG chromatin protein facilitating RNA polymerase I transcription in African trypanosomes. *Nucleic Acids Res.* 41, 2981–2992 (2013).

14. Lopez-Farfan, D., Bart, J. M., Rojas-Barros, D. I. & Navarro, M. SUMOylation by the E3 ligase ThbSI2/P1AS1 positively regulates VSG expression in *Trypanosoma brucei*. *PLoS Pathog.* 10, e1004545 (2014).

15. Alford, S. & Horn, D. Cell-cycle-regulated control of VSG expression site silencing by histones and histone chaperones ASF1A and CAP1-b in *Trypanosoma brucei*. *Nucleic Acids Res.* 40, 10150–10160 (2012).

16. Muller, L. S. M. et al. Genome organization and DNA accessibility control antigenic variation in trypanosomes. *Nature* 563, 121–125 (2018).

17. Glover, L., Hutchinson, S., Alford, S. & Horn, D. VEX1 controls the allelic exclusion required for antigenic variation in trypanosomes. *Proc. Natl Acad. Sci. USA* 113, 7225–7230 (2016).

18. Hutchinson, S., Glover, L. & Horn, D. High-resolution analysis of multi-copy variant surface glycoprotein gene expression sites in African trypanosomes. *BMC Genom.* 17, 806 (2016).

19. Callejas, S., Leech, V., Retter, C. & Melville, S. Hemizygous subtelomeres of an African trypanosome chromosome may account for over 75% of chromosome length. *Genome Res.* 16, 1109–1118 (2006).

20. Obado, S. O., Field, M. C., Chait, B. T. & Rout, M. P. High-efficiency isolation of nuclear envelope protein complexes from trypanosomes. *Methods Mol. Biol.* 1411, 67–80 (2016).

21. He, F. & Jacobson, A. Nonsense-mediated mRNA decay: degradation of decayed transcripts is only part of the story. *Annu. Rev. Genet.* 49, 339–366 (2015).

22. Delhi, P., Queiroz, R., Inchaustegui, D., Carrington, M. & Clayton, C. Is there a classical nonsense-mediated decay pathway in trypanosomes? *PLoS ONE* 6, e25112 (2011).

23. Lykke-Andersen, S. & Jensen, T. H. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptionites. *Nat. Rev. Mol. Cell Biol.* 16, 665–677 (2015).

24. Hammond, C. M., Strome, C. B., Huang, H., Patel, D. J. & Groth, A. Histone chaperone networks shaping chromatin function. *Nat. Rev. Mol. Cell Biol.* 18, 141–158 (2017).

25. Kerry, L. et al. Selective inhibition of RNA polymerase I transcription as a potential approach to treat African trypanosomiasis. *PLoS Negl. Trop. Dis.* 11, e0003432 (2017).

26. Alford, S. et al. High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. *Genome Res.* 21, 915–924 (2011).

27. Atcheson, N. et al. VSG switching in *Trypanosoma brucei*: antigenic variation analysed using RNAi in the absence of immune selection. *Mol. Microbiol.* 57, 1608–1622 (2005).

28. Devlin, R. et al. Mapping replication dynamics in *Trypanosoma brucei* reveals a link with telomere transcription and antigenic variation. *Elife* 5, e12765 (2016).

29. Mendell, J. T., ap Rheys, C. M. & Dietz, H. C. Separable roles for ren1/hUpf1 in altered splicing and decay of nonsense transcripts. *Science* 298, 419–422 (2002).

30. Azzalin, C. M., Reichenbach, P., Khoraiaud, L., Giulotto, E. & Lingner, J. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Nat. Cell Biol.* 18, 795–801 (2007).

31. Claudio, C. et al. Nuclear mRNA degradation pathway(s) are implicated in X chromosome inactivation. *PLoS Genet.* 2, e94 (2006).
Acknowledgements
We thank T. Owen-Hughes, N. Wiechens and V. Singh for assistance with ChIP-seq library preparation; A. Cassidy for Illumina sequencing; R. Clark of the Flow Cytometry and Cell Sorting Facility at the University of Dundee for cell sorting; D. Lamont, K. Beattie and S. Kosto of the FingerPrints Proteomics Facility at the University of Dundee for assistance with quantitative proteomics; the Dundee Imaging Facility; J. Rouse for access to the Leica Confocal SP8 Hyvolution microscope; S. Alsford (London School of Hygiene & Tropical Medicine) for the CAF-1 RNAi and CAF-1 tagging constructs; Marilyn Parsons (University of Washington) for the α-NOG1 antisera and C. Marques for discussions. The work was funded by Wellcome Trust Investigator Awards to D.H. [100320/Z/12/Z] and M.C.F. [204697/Z/16/Z] with additional support from a Wellcome Trust Centre Award [203134/Z/16/Z]. The University of Dundee Flow Cytometry and Imaging Facilities are supported by a Wellcome Trust award [097418/Z/11/Z], and the MRC Next Generation Optical Microscopy award [MR/K015869/1], respectively; while both the Proteomics and Imaging Facilities are supported by a Wellcome Trust Technology Platform award [097945/B/11/Z].

Author contributions
Conceptualisation: J.F., L.G., M.C.F. and D.H. Data curation: J.F. and S.H. Investigation: J.F., L.G. and C.B. Supervision: D.H. Original draft: J.F. and D.H. Review and editing: J.F., L.G., S.H., M.C.F. and D.H.

Additional information
Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-019-10823-8.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Peer review information: Nature Communications thanks Luisa Figueiredo, Sanjeeva Srivastava and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019