7526–7537 Nucleic Acids Research, 2010, Vol. 38, No. 21 Published online 31 July 2010
doi:10.1093/nar/gkq664

Dynamics of a novel centromeric histone variant CenH3 reveals the evolutionary ancestral timing of centromere biogenesis

Manu Dubin[1], Jörg Fuchs[2], Ralph Gräf[3], Ingo Schubert[2] and Wolfgang Nellen[1,*]

[1]Department of Genetics, University Kassel, Heinrich-Plett-Strasse 40, 34132 Kassel, [2]Leibniz-Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstrasse 3, 06466 Gatersleben and [3]Department of Cell Biology, Institute for Biochemistry and Biology, University of Potsdam, Potsdam-Golm, Germany

Received March 27, 2010; Revised June 20, 2010; Accepted July 14, 2010

ABSTRACT

The centromeric histone H3 variant (CenH3) serves to target the kinetochore to the centromeres and thus ensures correct chromosome segregation during mitosis and meiosis. The *Dictyostelium* H3-like variant H3v1 was identified as the CenH3 ortholog. *Dictyostelium* CenH3 has an extended N-terminal domain with no similarity to any other known proteins and a histone fold domain at its C-terminus. Within the histone fold, α-helix 2 (α2) and an extended loop 1 (L1) have been shown to be required for targeting CenH3 to centromeres. Compared to other known and putative CenH3 histones, *Dictyostelium* CenH3 has a shorter L1, suggesting that the extension is not an obligatory feature. Through ChIP analysis and fluorescence microscopy of live and fixed cells, we provide here the first survey of centromere structure in amoebozoa. The six telocentric centromeres were found to mostly consist of all the DIRS-1 elements and to associate with H3K9me3. During interphase, the centromeres remain attached to the centrosome forming a single CenH3-containing cluster. Loading of *Dictyostelium* CenH3 onto centromeres occurs at the G2/prophase transition, in contrast to the anaphase/telophase loading of CenH3 observed in metazoans. This suggests that loading during G2/prophase is the ancestral eukaryotic mechanism and that anaphase/telophase loading of CenH3 has evolved more recently after the amoebozoa diverged from the animal linage.

INTRODUCTION

Eukaryotic chromosomes contain specialized regions called centromeres where a multiprotein complex, the kinetochore, is formed at the G2/M transition (1). The kinetochore constitutes the attachment site for the spindle microtubules, which connect kinetochores and the centrosomes (microtubule-organizing centres) constituting the two spindle poles. Centrosomes are, in turn, anchored via astral microtubules to the cell cortex. They provide the skewback for physical forces generated by microtubule motors and changes in microtubule length that are needed to distribute sister chromatids to the daughter cells during mitosis and meiosis. The proteins comprising the kinetochore complex are highly conserved and a single evolutionary origin for centromeres early in eukaryotic evolution has been proposed (2). Centromere size is extremely variable, ranging from the 125-bp point centromeres in *Saccharomyces cerevisiae* to the holocentric centromeres of *Caenorhabditis elegans*, which span the entire length of the chromosome (3). Many eukaryotes have regional centromeres consisting of arrays of satellite repeats and/or transposons typically spanning 100 kb to 100 Mb (3). In contrast to the highly conserved kinetochore proteins, the DNA sequence composing the centromeres is highly variable (4,5) and in some species is not even conserved among the different chromosomes (6). It has been hypothesized that the rapid evolution of centromere sequences is a major driver of speciation as sequence divergence leads to incompatibility in meiosis (7). With the exception of point centromeres such as those of *S. cerevisiae,* DNA sequence alone appears insufficient to specify centromere function. Instead, centromere formation seems to require the presence of pericentromeric heterochromatin flanking a core centromere region. In the fission yeast *S. pombe*, pericentromeric heterochromatin formation is RNAi dependent, where small RNAs serve to recruit the repressive RNA-induced transcriptional silencing (RITS) complex which catalyses the heterochromatin-associated histone H3K9me2 modification (8). The RNAi pathway also appears to be necessary for pericentromeric heterochromatin maintenance in mammals and *Drosophila (*3). In the core centromere,

*To whom correspondence should be addressed. Tel: +49 561 804 4805; Fax: +49 561 804 4800; Email: nellen@uni-kassel.de

nucleosomal histone H3 is replaced by a centromere-specific variant know as CSE in fungi, Cid in *Drosophila*, CENP-A in metazoans and Htr12 or CenH3 in plants. This centromeric H3 variant is an essential protein and interacts directly with components of the kinetochore complex. Unlike the highly conserved kinetochore proteins, CenH3 and the centromere DNA sequences rapidly co-evolve and are poorly conserved. Compared with conventional histone H3, the histone fold domain of CenH3 has a divergent and longer L1 and a divergent α2 (9,10). Many CenH3 variants, such as Cid from *Drosophila* have a long N-terminal domain; however, this appears to be dispensable for proper targeting of CenH3 (11,12). In contrast to the surrounding pericentromeric heterochromatin, the core centromere has properties that resemble those of euchromatin such as H3K4 methylation and low levels of H3K9 methylation (1).

In general, the mechanisms controlling centromere specification and formation remain poorly understood and different mechanisms appear to be used in different species. Much of our understanding comes from progress made in dissecting the mechanism in the fission yeast *S. pombe* (13). However, differences such as the absence of orthologs of the RITS complex and the presence of DNA methylation in higher eukaryotes complicate the direct transfer of this knowledge to other organisms.

The social amoeba *Dictyostelium discodeum* is a useful system for studying centromere specification and formation. Its centromeres consist of several retroelement arrays (14). Unlike fungi, which have undergone rapid evolution accompanied by large-scale genome compaction and gene loss (15,16), the amoebozoans appear to have retained more of the ancestral genomic diversity than other members of the crown group of organisms (fungi, plants and metazoans) (17). Many components of the RNAi, chromatin remodelling and DNA damage-repair pathways are conserved between *Dictyostelium* and higher eukaryotes. Examples include small gene families encoding Dicer's, RdRP's, Argonauts, HP1, Aurora kinases, inner centromere protein (INCENP) and components of the centrosome as well as a DNA methyltransferase, many of which have been characterized (18). The predominant localization of H3K9me2, H3K9me3 and the two *Dictyostelium* HP1 homologs, HcpA and HcpB, to a major focus harbouring the centromeres has previously been reported (19–22) by (M Dubin, PhD thesis, University of Kassel, 2010). The predominant localization of H3K9me2, H3K9me3 and the two *Dictyostelium* HP1 homologs, HcpA and HcpB, to a major focus harbouring the centromeres has previously been reported (19,20); (M Dubin, PhD thesis, University of Kassel, 2010).

The euchromatin-associated histone H3K4me modification has a rather homogeneous distribution throughout the nucleus (23). Here, we describe the identification and characterization of a CenH3 ortholog from *Dictyostelium*, which due to its position in the phylogenetic tree provides information on ancestral aspects of centromere biogenesis.

MATERIALS AND METHODS

Dictyostelium cells and culture

The *D. discodeum* strain Ax2-214 (axeA2, axeB2, axeC2) (24) was cultured in petri dishes or shaking culture at 20°C in HL5 medium (Formedium; Hunstaton, UK) supplemented with 100 μg/ml of Ampicillin, 100 μg/ml of Amphotericin-B and the appropriate selective agent (10 μg/ml of Geneticin and/or 10 μg/ml of Blasticidin). Alternatively, *Dictyostelium* cells were grown on bacterial lawns of *Klebsiella aerogenes* on SM agar plates.

Vectors and transformation

Pfu DNA polymerase was used to amplify histones *h3v1* (DDB_G0291185), *h3v2* (DDB_G0277979), *h2AX* (DDB_G0279667) and *h2Bv3* (DDB_G0286509) from genomic DNA using the following primers: MJD87 (5′-AGTCGACAATGGCTAACAAACCCAAACCCT C-3′) and MJD88 (5′-ACTCGAGTTAAAAAAGAAAA TGTCTAGCCCTTTTCC-3′) (*h3v1*), MJD85 (5′-AGTC GACAATGACAAGTGTTAATAATAATATGACAA G-3′) and MJD86 (5′-ACTCGAGTTAATAACGTGGC AAATAAAATGGTTTGATG-3′) (*h3v2*), MJD140 (5′-A GTCGACAATGTCAGAAACCAAACCAGCCTC-3′) and MJD141 (5′-TCTCGAGTTAATAGATTTGAGAT GAACCTTCAGCTG-3′) (*h2AX*), MJD121 (5′-AACAT ATGTCGACAATGGTATTCGTTAAAGGTCAAAAG AAAG-3′) and MJD122 (5′-ACTCGAGTTAGTTTTTG CTTTCAGTTGGATTGTAC-3′) (*h2Bv3*).

Polymerase chain reaction (PCR) products were cloned into pJET (Fermentas; Burlington, Canada) and confirmed by sequencing. The genes were excised using the SalI and XhoI sites introduced during the PCR amplification (underlined in the primer sequences). Sequences encoding *h3v1* and *h3v2* were ligated into the vector pDneo2a-GFP using the same two sites, while H2B was ligated into the extrachromosomal vector pDbsrXP-RFP. *Dictyostelium* A × 2 cells were transformed using electroporation (25) and selected in HL5 medium supplemented with 10 μg/ml of Geneticin. After ~10 days, transformants were subcloned and a clonal line was used for further experiments. Cells expressing GFP-DdCenH3 were super-transformed with the pDbsrXP-RFP-H2B construct and selected in HL5 medium supplemented with 10 μg/ml of Geneticin and 10 μg/ml of Blasticidin.

Identification of homologs, alignments

Homologs of histone genes were identified in the *Dictyostelium* genome (17) using Basic Local Alignment Search Tool (BLAST) search (26), a reciprocal BLAST search against the National Center for Biotechnology Information (NCBI) databank (non-redundant protein sequences), performed to confirm identity. Orthologs of H3v1 were identified using a position-specific iterative BLAST (PSI-BLAST) search (27) against the NCBI databank (non-redundant protein sequences). Sequence alignments were performed using Clustal W (28). Phylogenetic trees were inferred using Mr Bayes (v3.1.2) (29,30). A mixed amino acid model was run for 1 000 000 iterations using default settings. Similar results were

obtained using Clustal X (31). Phylogenetic trees were presented using Dendroscope (v1.4) (32).

Immunofluorescence

Cells were grown overnight on coverslips to ~80% confluence and fixed in 4% (w/v) paraformaldehyde (PFA) dissolved in 20-mM phosphate buffer (pH 6.7) for 10 min at 22°C. The PFA solution was removed and cells were permeabilized for 5 min in 0.2% (v/v) Triton X-100 in 20-mM phosphate buffer (pH 6.7). Cells were blocked with 2% (w/v) bovine serum albumin (BSA) in phosphate-buffered saline (PBS), pH 7.6, for 30 min at room temperature. The coverslips were incubated overnight at 4°C with the primary antibody diluted in 1% (w/v) BSA in PBS pH 7.6, washed 3 × 5 min in PBS and then incubated with the secondary antibody diluted in 1% (w/v) BSA in PBS pH 7.6 for 1 h at room temperature. The coverslips were washed 3 × 5 min in PBS, and mounted on a slide with a drop of mounting media (250 ng/ml of 4′,6-Diamidino-2-phenylindole (DAPI) in 90% (v/v) glycerol, 20-mM Tris–HCL, 1 μg/ml DABCO, pH 8.3) and examined by fluorescence microscopy.

Fluorescence *in situ* hybridization

The Fluorescence *in situ* hybridization (FISH) protocol was adapted from Moerman and Klein (33). Cells were allowed to settle on slides for 10 min, fixed in 3:1 methanol:acetic acid at −20°C for 60 min and air dried. Slides were placed at 60°C for 30 min on a hot plate and then treated with RNaseA (200 μg/ml in 2 × SSC for 40 min at 37°C). Slides were washed 2 × 5 min in 2 × SSC, 1 × 5 min in PBS then post-fixed with 4% PFA in PBS for 10 min at room temperature. Slides were washed 2 × 5 min in PBS and then dehydrated using an ethanol series of 70, 90 and 99.7% for 3 min each and then air dried. Approximately 30 ng of fluorescent-labelled probe was resuspended in 12 μl of hybridization solution (50% formamide, 10% dextran sulphate and 2 × SSC), denatured for 5 min at 95°C and cooled on ice. The hybridization mix was placed on the slide, overlaid with a coverslip and sealed with rubber cement. The slide was placed on a 80°C heat block to denature the genomic DNA for 2 min and then incubated at 37°C overnight. Slides were washed twice in 2 × SSC for 5 min at 42°C and three times for 5 min in 2 × SSC, 50% formamide at 42°C, dehydrated in an ethanol series and air dried. Mounting media (see above) was placed on the slides and overlaid with a coverslip.

The DIRS-1 the Y41 clone (34) was used as a probe by direct labelling using FITC-dUTP according to Lysak *et al.* (35).

Combined immunofluorescence–FISH

Cells were immunostained as described above and then fixed with 4% (w/v) PFA, 4% (w/v) sucrose in 20-mM phosphate buffer (pH 6.7) for 10 min at room temperature. Slides were washed 1 × 5min with PBS, 2 × 5 min in 2 × SSC and processed according to the FISH protocol starting from the RNaseA treatment step.

Microscopy

Images were acquired on a Leica DMIRB inverted microscope equiped with a DC350 camera and IM50 Acquisition software (Leica Microsystems; Wetzlar, Germany). Alternatively, live cells were observed using an Axiovert 200 M CellObserver HS system (Carl Zeiss MicroImaging GmbH, Göttingen, Germany) equipped with a Sutter DG-4 light source (Sutter instruments, Novato, CA, USA), an ASI piezo stage (Applied Scientific Instruments, Eugene, OR, USA), a Zeiss LCI PlanNeo 63 × /1.3 NA water immersion lens, a Zeiss Axiocam MRm Rev. 3 CCD camera and Axiovision 4.6.3 software. Non-saturated Z-stacks (seven slices, 700 nm between slices were acquired every 20 or 60 s. Images were quantified and prepared for presentation using ImageJ v1.42n-v1.43b (http://rsbweb.nih.gov/ij/). A maximum-intensity projection of the stack was derived (without deconvolution). Background subtraction was achieved by (linearly) adjusting the levels of each channel so that areas outside the nucleus had intensity values of zero. For each channel, images in the time series were normalized (so each frame has the same average pixel intensity) to correct for photobleaching.

To measure the centromere (GFP) signal, a circular ROI slightly larger than the centromere was manually centred over the centromere and the total signal measured. The size/dimensions of the ROI were identical for every centromere and every time point. For the nuclear (RFP) signal, a circular ROI slightly larger than the nucleus was manually centred over the nucleus and the total signal measured. The total intensity values of the three (before anaphase) or six (after anaphase) centromeres/nuclei over three consecutive time points (a minimum of nine data points) were averaged and plotted together with the standard error (SE, $P = 0.05$).

To avoid artefacts by synchronization, observations during the cell cycle were done by selecting appropriate cells and following their fate in an unsynchronized culture. We found ~5% of the cells in mitosis.

Chromatin immunoprecipitation

All steps were performed on ice unless otherwise stated. The 1×10^8 cells were fixed in 1% (w/v) PFA in phosphate buffer at 22°C. After 10 min, glycine was added to a final concentration of 200 mM and the cells were incubated for a further 5 min. Cells were washed with ice-cold phosphate buffer and with 1% Triton X-100 in phosphate buffer. Cells were resuspended in 500 μl of nuclear lysis buffer [50 mM of HEPES, 10 mM of ethylene-diamine-tetra-acetic acid (EDTA), 1% Triton X-100, 1% SDS 0.1 mM of PMSF, complete protease inhibitors (Roche) pH 8.0] in an Eppendorf tube. Samples were sonicated with a UP 200S sonicator (Dr Hielscher GmbH, Stansdorf, Germany) in an ice-water bath (5 × 10 s, 25% output, 45% duty cycle) with a 50-s pause between bursts. 900 μl of ChIP dilution buffer [16.7 mM HEPES, 1% Triton X-100, 1.2 mM of EDTA, 167 mM of sodium chloride (NaCl), 0.1 mM of PMSF, complete protease inhibitors (Roche) pH 8.0] was added per 100 μl of sonicated lysate. Samples were centrifuged for 2 min, 16 000*g* at

4°C and the supernatant was transferred to a new tube. The supernatant was pre-cleared with protein-A sepharose for 60 min at 4°C, then centrifuged and transferred to a new tube. The supernatant was divided into 1 ml fractions (corresponding to 2×10^7 cells) and 20 µl of anti-GFP affinity resin (36) (GFP-Trap provided by ChromoTek, Martinsried) was added to one fraction and 20 µl of rabbit immunoglobulin G (IgG)-sepharose added to another fraction to serve as the negative control. Samples were rotated for 60 min at 4°C. Beads were washed twice with low-salt wash buffer (150 mM of NaCl, 0.1% SDS, 1% Triton X-100, 2 mM of EDTA, 20 mM of HEPES, pH 8.0), once with high-salt wash buffer (500 mM of NaCl, 0.1% SDS, 1% Triton X-100, 2 mM of EDTA, 20 mM of HEPES, pH 8.0), twice with LiCl wash buffer (250 mM of LiCl, 1% NP40, 1% sodium deoxycholate, 1 mM EDTA, 10 mM HEPES, pH 8.0) and twice with TE buffer (10 mM of Tris–HCl, 1 mM of EDTA, pH 8.0). After the final wash, 250 µl of elution buffer [1% (w/v) SDS, 100 mM of sodium bicarbonate ($NaHCO_3$)] was added to the beads and the resultant slurry was incubated at 65°C for 15 min under gentle agitation. After a brief centrifugation to pellet the beads, the supernatant was transferred to a new tube. Extraction of the beads was repeated with 250 µl of elution buffer. The two elutions were combined, 20 µl of 5M NaCl was added and incubated overnight at 65°C to reverse the cross-linking. Then 20 µl of proteinaseK buffer (250 mM of EDTA, 1 M Tris–HCl pH 6.5) and 2 µl of proteinaseK (10 mg/ml) were added and the solution was incubated at 45°C for 3 h. After phenol extraction, the DNA was ethanol precipitated and resuspended in 50 µl of double-distilled water (DDW) containing 10 ng/ml of RnaseA. Oligonucleotide primers against the loci of interest were designed using the Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3www.cg). Minimum and maximum primer GC content was set to 40 and 60%, respectively, and amplified fragment length to 150–200 bp. Otherwise, standard settings were used. Primers are listed in the Supplementary Data. Quantification was performed using a Real-Time PCR machine from Eppendorf (Hamburg, Germany) or Bio-rad (Hercules, CA, USA) using Eva-Green nucleic acid dye (Jena Bioscience; Jena, Germany) according to the manufacturer's instructions. For each locus, the PCR assay was performed in quadruplicate and each ChIP was performed at least twice. An aliquot of sheared chromatin was prepared to be analysed by gel electrophoresis the average size of DNA fragments. This was in the range of 750–250 bp with the majority between 300 bp and 400 bp (data not shown).

RESULTS

Identification of a putative *Dictyostelium* CenH3 ortholog

Attempts to identify a *Dictyostelium* CenH3 ortholog by BLAST search with CenH3 from various species as the query sequence were unsuccessful, since in each case one of the core histone H3.3 variants H3a, H3b and H3c was found to be the *Dictyostelium* protein with the highest homology. In addition, two genes encoding histone H3-like domains are present in the *Dictyostelium* genome (*h3v1*: DDB_G0291185 and *h3v2*: DDB_G0277979; Supplementary Figure S1a). H3v1 is significantly longer than a typical histone H3 (619 versus 136 amino acids) and contains a histone H3-like domain at the C-terminus while H3v2 is considerably smaller (96 versus 136 amino acids) and contains an intron. H3c, H3v1 and H3v2, were cloned and expressed with an N-terminal GFP tag. GFP-H3c was distributed throughout the nucleus as would be expected for a core histone and GFP-H3v2 did not localize to the nucleus at all (data not shown). In interphase cells, GFP-H3v1 localized to a single focus (Supplementary Figure S1b) adjacent to the centrosome (Figure 1B) suggesting that it represents a *Dictyostelium* ortholog of CenH3.

The expression pattern of endogenous H3v1 throughout the developmental cycle has been determined by deep sequencing and is represented on DictyExpress (http://www.ailab.si/dictyexpress/run/index.php?gene=DDB_G0291185&db=rnaseq). These data show a low expression level with a peak at 12 h in development. This can be explained by a rather synchronous single S-phase of many cells at this stage.

H3v1 contains a large N-terminal domain with no homology to other known proteins. Long, unique N-terminal domains are found in many, but not all, CenH3 proteins and have been proposed to recruit components of the kinetochore complex to the centromeres (5). A position-specific iterative BLAST (PSI-BLAST) search of *Dictyostelium* H3v1 against Genebank (November 2008) revealed that the most closely related protein is *Drosophila teissieri* Cid (CenH3, AAK20217). The similarity is restricted to the histone fold domain and is quite low for a histone ($4e^{-14}$). The histone domain of H3v1 was aligned with histone H3 and CenH3 proteins of other species (Figure 1A). Like the other known and putative CenH3 proteins, H3v1 has a poorly conserved α2, with weak similarity to that of the putative CenH3 from *Entamoeba histolytica* and *Drosophila teissieri* Cid. Unlike all other CenH3 proteins examined, H3v1 does not have an extended L1 (Figure 1A), which has previously been proposed to be an essential feature for centromere targeting (11,37). A phylogenetic tree was calculated using Bayesian methods (29). While conventional H3 histones, including *Dictyostelium* H3a, H3b and H3c clustered close together, H3v1 was on a separate, more dispersed branch with the CenH3 orthologs from other species (Supplementary Figure S2). Therefore, we will refer to H3v1 as DdCenH3.

Localization of GFP-DdCenH3 throughout the cell cycle

DdCenH3 was expressed with an N-terminal GFP, 3xFLAG, or 3xHA tag, and its localization was examined by immunofluorescence (IF) microscopy. Western blotting confirmed that the expressed protein was of the expected size (Supplementary Figure S3a). In all cases, DdCenH3 localized to a single focus at the edge of the nucleus next to the centrosome during interphase (Figure 1B and Supplementary S3b) suggesting that the

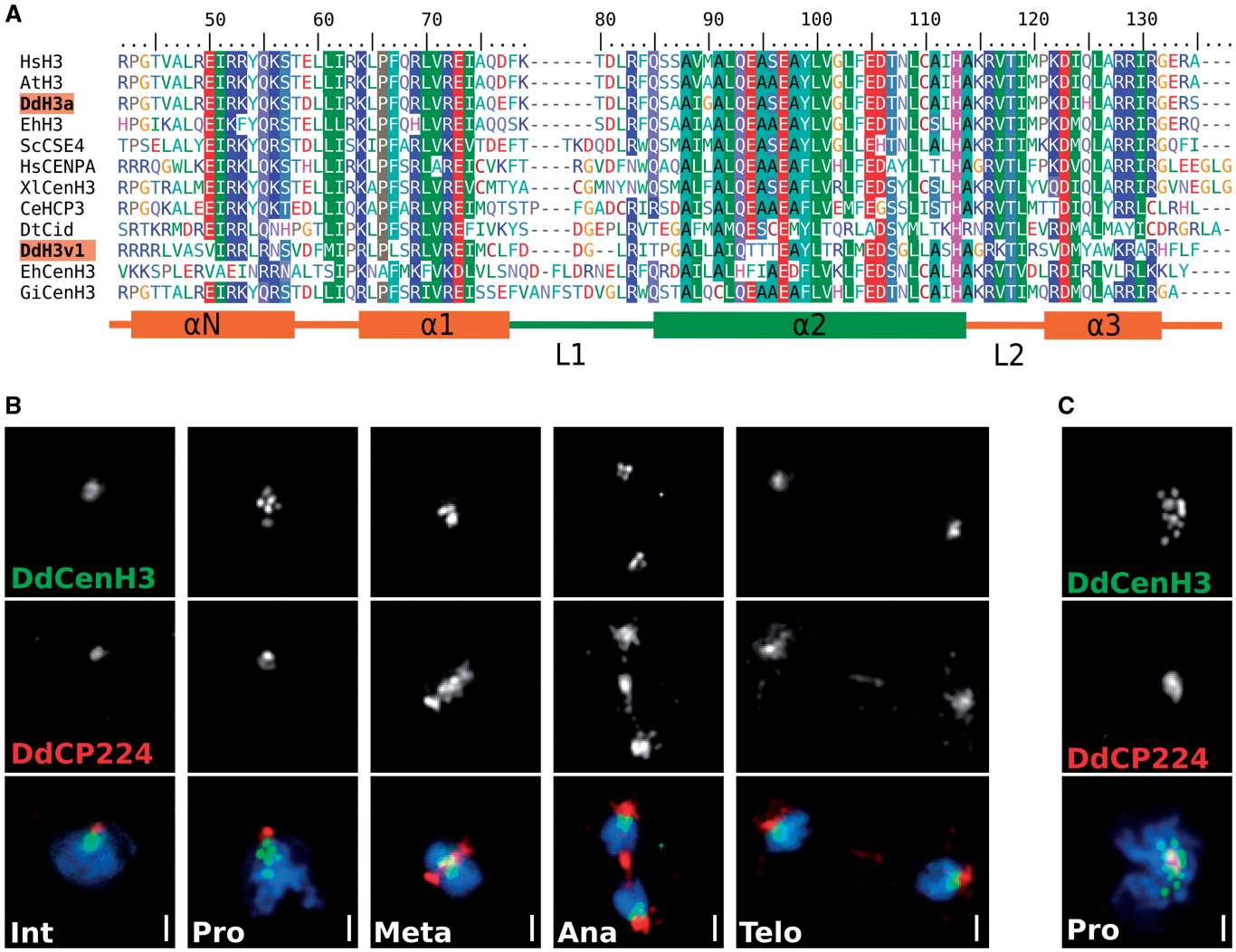

Figure 1. The *Dictyostelium* centromeric histone H3 variant. (**A**) *Dictyostelium* H3a and DdCenH3 (H3v1) are highlighted in orange on the left. Below the alignment is a schematic presentation of the domain organization of histone H3, with the N-terminal α-helix (αN), α-helixes 1–3 (α1–3) and loops 1 and 2 (L1, L2) indicated. L1 and α2, which have been shown to be responsible for targeting of CenH3 to centromeres are shaded green. Hs, *Homo sapiens;* At, *Arabidopsis thaliana*; Dd, *Dictyostelium discodeum*; Eh, *Entamoeba histolytica*; Sc, *Saccharomyces cerevisiae*; Xl, *Xenopus laevis*; Ce, *Caenorhabditis elegans*; Dt, *Drosophila teissieri*; Gi, *Giardia intestinalis*. (**B**) The organization of the centromeres was examined in fixed cells expressing GFP-DdCenH3 and immunostained with the centrosome marker DdCP224 at the indicated stages of the cell cycle. (**C**) Prophase of a diploid cell immunostained with DdCP224 and expressing GFP-DdCenH3 displays up to 12 GFP-DdCenH3 labelled foci presumably corresponding to the 12 centromeres. Scale bar = 1 μm.

interphase centromeres remain clustered and attached to the centrosome as in *S. cerevisiae* (38) and *S. pombe* (39). This is in contrast to the situation in many metazoans and plants as well as in *Giardia intestinalis* and *Plasmodium falciparum,* where single centromeres are usually detectable during interphase (40,41).

The localization of GFP-DdCenH3 during the cell cycle was examined in cells that had been fixed at different stages and immunostained with antibodies against H3K9me3, tubulin or DdCP224 (42). As cells enter prophase, the chromosomes begin to condense and up to six GFP-DdCenH3 foci were visible at the edge of the nucleus close to the centrosome (Figure 1B and Supplementary Figure S4). These six foci presumably correspond to the centromeres of the six chromosomes of the haploid *Dictyostelium* genome. When diploid cells were examined, up to 12 foci were visible (Figure 1C). During metaphase chromosome condensation, the six GFP-DdCenH3 foci form a ring around the mitotic spindle until they can no longer be clearly distinguished. In anaphase and telophase cells, the sister chromatids separated and a focus of GFP-DdCenH3 was found at the leading edge of the chromosomes immediately behind the centrosomes. Double labelling with an antibody against the pericentromeric heterochromatin marker histone H3K9me3 shows that GFP-DdCenH3 has a very similar, but not identical, distribution pattern (Figure 2A). In both interphase and telophase cells, a cross section through the nucleus showed that the centre of intensity of GFP-DdCenH3 is slightly closer to the

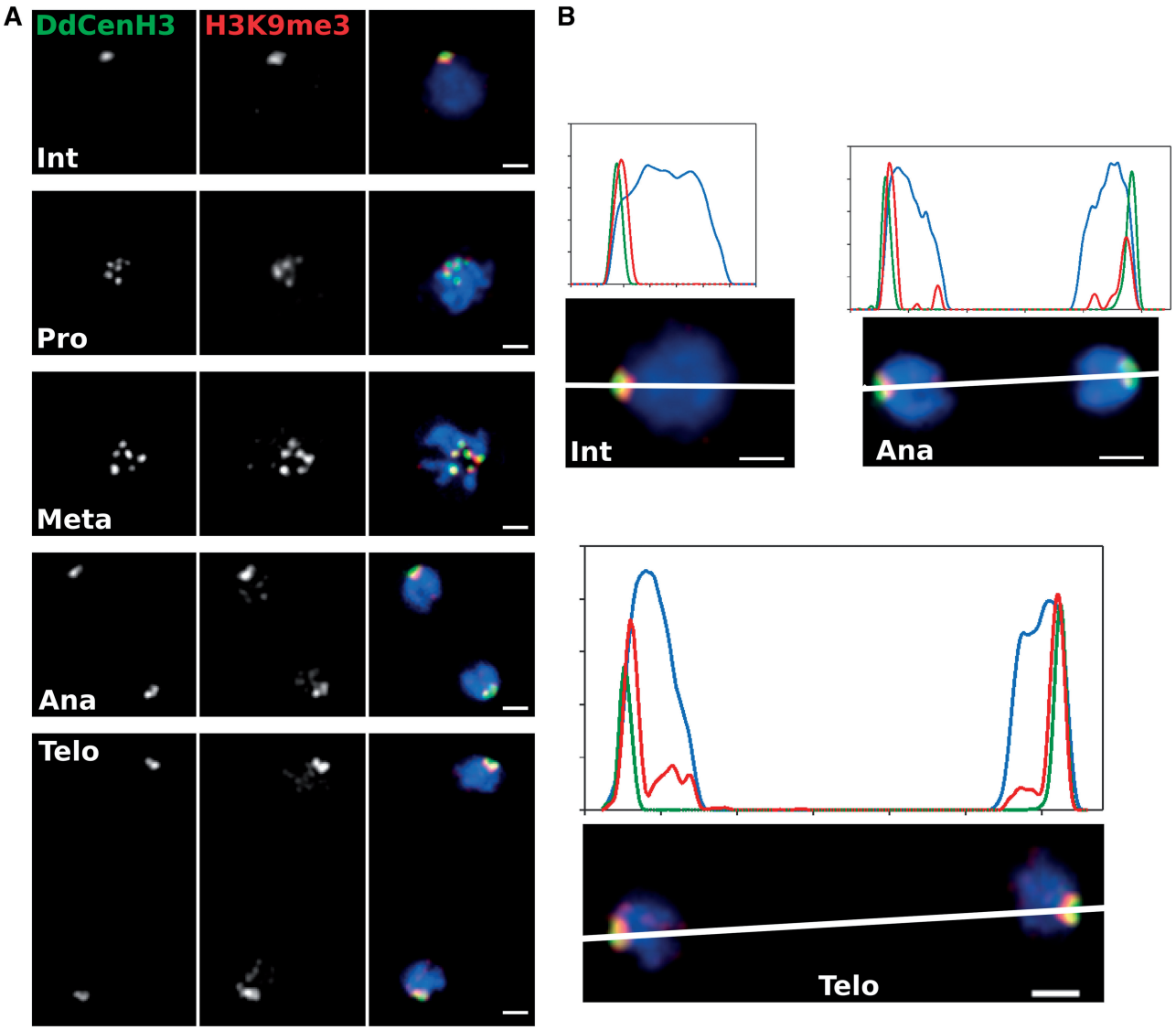

Figure 2. Centromere behaviour during mitosis. (**A**) Cells expressing GFP-DdCenH3 and labelled with the heterochromatin marker H3K9me3. DNA is stained with DAPI. (**B**) Cross sections through cells at different stages of the mitotic cycle show that GFP-DdCenH3 (green) is slightly closer to the leading edge of the separating chromatids than the heterochromatin marker H3K9me3 (red). Scale bar = 1 μm.

centrosome than that of H3K9me3 (Figure 2B). This suggests that the core centromere domain containing DdCenH3 is flanked by or interspersed with pericentromeric heterochromatin enriched in H3K9me3 as has been described for other species (1).

Dictyostelium centromeres contain DIRS-1

DIRS-1, the most abundant retrotransposon in the *Dictyostelium* genome (14), has previously been proposed to indicate the centromeres positions as it is present in six clusters, one on each chromosome (43). DIRS-1 sequences are methylated (21) and are associated with H3K9me2 (20). FISH with DIRS-1 sequences yielded six foci per interphase nucleus (17), while DdCenH3, H3K9me3, the HP1 homologs HcpA and HcpB (19) and Cenp68 (44), all yield a single signal cluster. To further investigate this discrepancy, FISH was performed using Cy3 or FITC-labelled PCR fragments or a plasmid containing a full-length DIRS-1 clone (Y41) (34). All interphase cells displayed a single DIRS-1 focus at the nuclear periphery similar to that seen with GFP-DdCenH3. As cells entered prophase, up to six foci became visible. These condensed to a single focus during metaphase and appeared at the leading edge of the separating chromatids during anaphase and telophase (Figure 3A).

DdCenH3 colocalizes with DIRS-1 retrotransposons

Immuno-FISH experiments revealed colocalization of DIRS-1 with both HA-tagged DdCenH3 (Figure 3B) and FLAG-tagged DdCenH3 (Supplementary Figure S5). Both marks remain associated with DIRS- during the entire cell cycle, suggesting a constitutive association of DdCenH3 with DIRS-1 sequences. ChIP was performed on cells expressing GFP-DdCenH3 with the

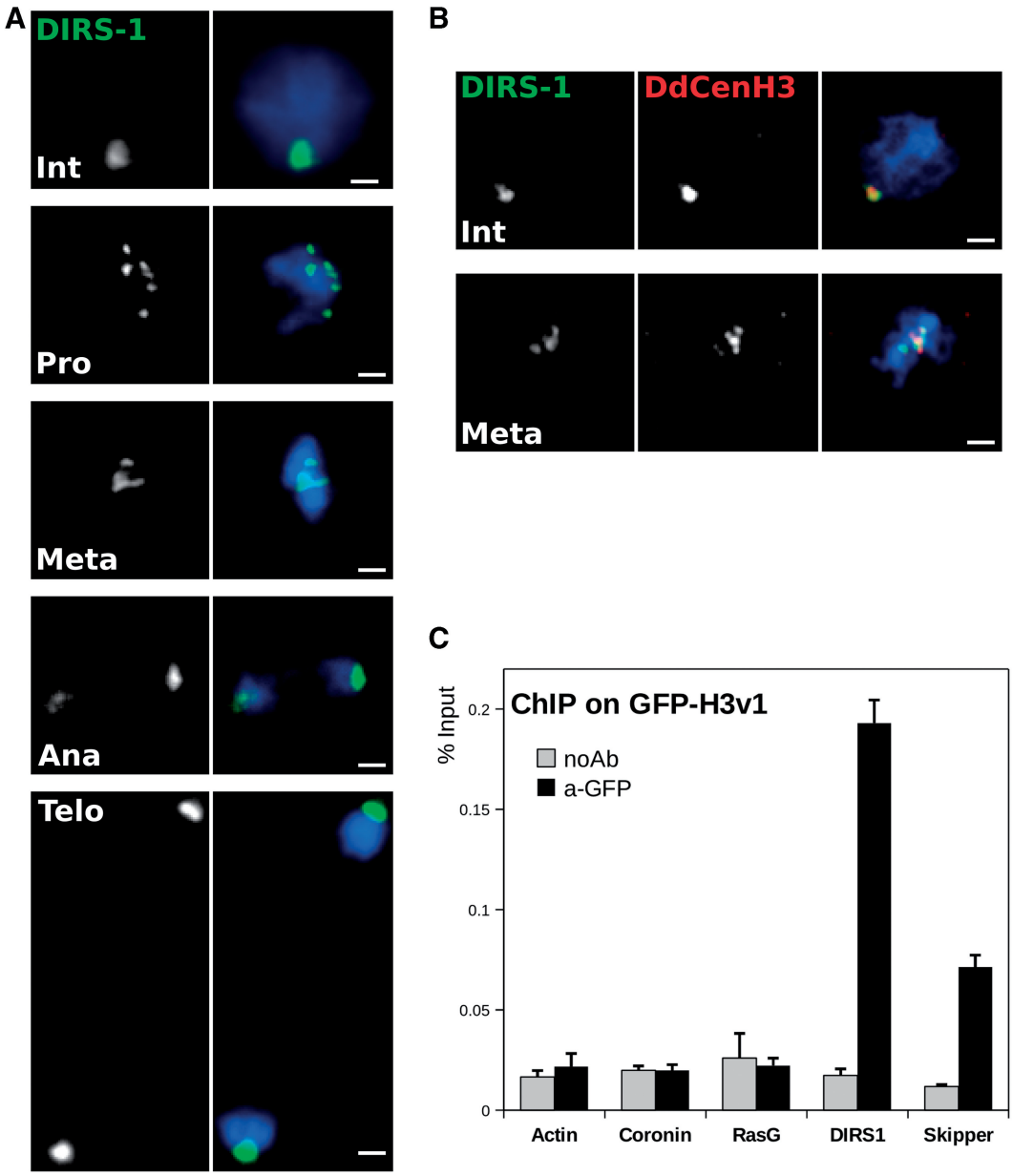

Figure 3. *Dictyostelium* centromeres contain DIRS-1. (**A**) FISH on cells at different stages of the cell cycle using a directly FITC-labelled probe for DIRS-1. (**B**) Combined Immuno-FISH showing colocalization HA-tagged DdCenH3 and DIRS1 retrotransposon. Scale bar = 1 μm. (**C**) ChIP of GFP-DdCenH3. Real-time qPCR was performed with primers against the actively transcribed *actin*, *coronin* and *rasG* genes or the DIRS-1 and *skipper* retrotransposons. noAb = no antibody control.

GFP affinity resin. DIRS-1, and to a lesser extent the *skipper* retrotransposon, were enriched in DdCenH3-containing chromatin while the active genes *coronin*, *actin* and *rasG* were not enriched at all (Figure 3C). These data provide direct evidence that the 200–300 kb cluster of DIRS-1 transposons at one end of each chromosome represents the centromeres of *Dictyostelium*.

Loading of DdCenH3 onto centromeres occurs before sister chromatid separation

Unlike the core histones, which are incorporated into nucleosomes during S-phase and unlike the replication-independent deposition of the histone variants H3.3 and H2AZ, which are deposited in a transcription-dependent manner, the mechanism of CenH3 incorporation into centromeric nucleosomes remains poorly understood. Experiments with *Drosophila* and mammalian cell lines suggested that CenH3 loading occurs in anaphase and telophase/early G1, respectively (45,46). More recently, it has been shown that the Holiday Junction Repair Protein, HJURP is the chaperone responsible for *de novo* CenH3 deposition in mammals (47,48). The yeast protein SpScm3 shares ancestry with HJURP and is involved in CenH3 loading in this species

(49). Homologs of HJURP and SpScm3 appear to be absent from plants and *Dictyostelium* (data not shown). In *Arabidopsis* loading of CenH3 occurs in late G2 phase (12). Also in other higher plants, in red algae and in fission yeast CenH3 loading occurs prior to separation of sister chromatids [for review see (50)]. Since amoebozoa diverged from the opisthokont lineage shortly after the plant–animal split, determination of the CenH3 loading time in the ancestral eukaryote *Dictyostelium* may elucidate when loading of CenH3 occurred in ancestral eukaryotes and thus provide information as to the evolution of the CenH3 loading mechanism. Cells expressing GFP-DdCenH3 and the cell-cycle marker RFP-PCNA (51) were fixed and the intensity of GFP at the centromere was measured at various stages of the cell cycle (Figure 4A). After mitosis, *Dictyostelium* cells enter S-phase almost immediately and in late S-phase RFP-PCNA stains the late-replicating heterochromatin (51). The intensity of GFP-DdCenH3 was low in late S-phase and not significantly higher in G2. The intensity doubled in prophase, remained constant in metaphase and became halved in telophase nuclei (where both daughter nuclei were measured separately). Thus *de novo* CenH3 loading occurs at the G2-prophase transition in *Dictyostelium*.

Time-lapse imaging was performed on cells co-expressing GFP-DdCenH3 and RFP-tagged histone H2AX or H2B. Cells were tracked as they went through mitosis and the fluorescence intensity of GFP-DdCenH3 and of RFP-histone were measured every 20 or 60 s (Figure 4B; Supplementary S6 and Supplementary movie S1). The intensity of the core histone remained constant until cells entered mitosis and dropped to approximately half the initial value in anaphase when sister chromatids are separated. In contrast, the intensity of the GFP-DdCenH3 signal approximately doubled towards the G2-prophase transition and was halved in anaphase nuclei, indicating that loading of DdCenH3 occurs at the G2-prophase transition. This behaviour clearly differs from that of the core histones (H2AX, H2B, H3 and H4) which are incorporated in nucleosomes during S-phase. To exclude S-phase loading of DdCenH3, GFP-DdCenH3 was co-expressed with RFP-PCNA and cells were monitored by time-lapse from the exit of mitosis until the end of S-phase (Figure 5). No increase in GFP-DdCenH3 intensity was observed during S-phase. Due to the light sensitivity of *Dictyostelium* cells, we were unable to track an individual amoeba through the entire cell cycle. Thus, we cannot exclude some loading of DdCenH3 earlier in G2. However, the strong increase in intensity observed at the G2/prophase transition strongly suggests that the majority of DdCenH3 is incorporated into centromeric nucleosomes at the onset of mitosis.

DISCUSSION

The *Dictyostelium* CenH3 ortholog, H3v1 (DdCenH3) is longer than a typical histone H3 (619 versus 136 amino acids) and contains a histone H3-like domain at the C-terminus. Tagged DdCenH3 co-localizes with the heterochromatin marker H3K9me3 during the entire cell cycle suggesting that it is a bonafide CenH3 ortholog and marks the core centromeres. The histone fold of DdCenH3 shares some similarity to the CenH3 from *Drosophila* family members, possibly due to convergent evolution. Previously it has been shown that most amino acid substitutions within L1 abolished centromere targeting of CenH3 in *Drosophila* cells. Likewise, shortening L1 by even one residue impaired targeting (11). Crystallographic studies suggest that L1 forms a DNA binding domain that binds the minor grove of the DNA warped around the nucleosome (52) and it has been proposed that extended L1 of CenH3 gives specificity for the centromere sequence of that particular species (11). In contrast to all other CenH3 orthologs examined, L1 of the DdCenH3 histone fold is not longer than that of a conventional histone H3. ChIP experiments suggest DdCenH3 has some specificity for DIRS-1, suggesting that at least in *Dictyostelium*, an extended L1 is dispensable for correct CenH3 targeting.

The DIRS-1 retrotransposons are believed to represent the centromeric sequences in *Dictyostelium* (43). FISH experiments revealed that all of DIRS-1 localizes to a single focus at the nuclear periphery in interphase. The separate centromere foci observed in previous experiments (17) may be due to a disruption of tightly associated centromeres during preparation combined with higher resolution microscopy. DIRS-1 colocalized throughout the cell cycle with DdCenH3. DIRS-1 was enriched in DdCenH3-containing chromatin, suggesting that it is a component of the core centromere.

Like plant and yeast CenH3 (50), DdCenH3 is loaded onto the centromeres from G2 to prophase. This is in contrast to the situation in metazoans, where CenH3 is loaded after the separation of sister chromatids in telophase/G1 (45,46) by the HJURP chaperone complex (47,48) or in yeast were CenH3 is loaded in S-phase and G2 by Scm3, a HJURP homolog (52). HJURP/Scm3 homologs have so far only been identified in the opisthokont lineage and telophase/G1 or S-phase/G2 loading of CenH3 has only been described in this lineage to date. This suggests that incorporation of CenH3 into centromeres during late G2 is the ancestral mechanism and it was only after the divergence of amoebozoans from the opisthokonts that an alternative mechanism involving the HJURP/Scm3 complex evolved in the latter lineage. The mechanism of CenH3 loading in non-metazoans remains elusive.

Taken together, for the first time, we have provided a survey of centromere structure in amoebozoans and obtained clear evidence that acquisition of the centromere-specific H3 to centromeric DNA at the G2/M transition is the ancestral eukaryotic mechanism of centromere formation.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

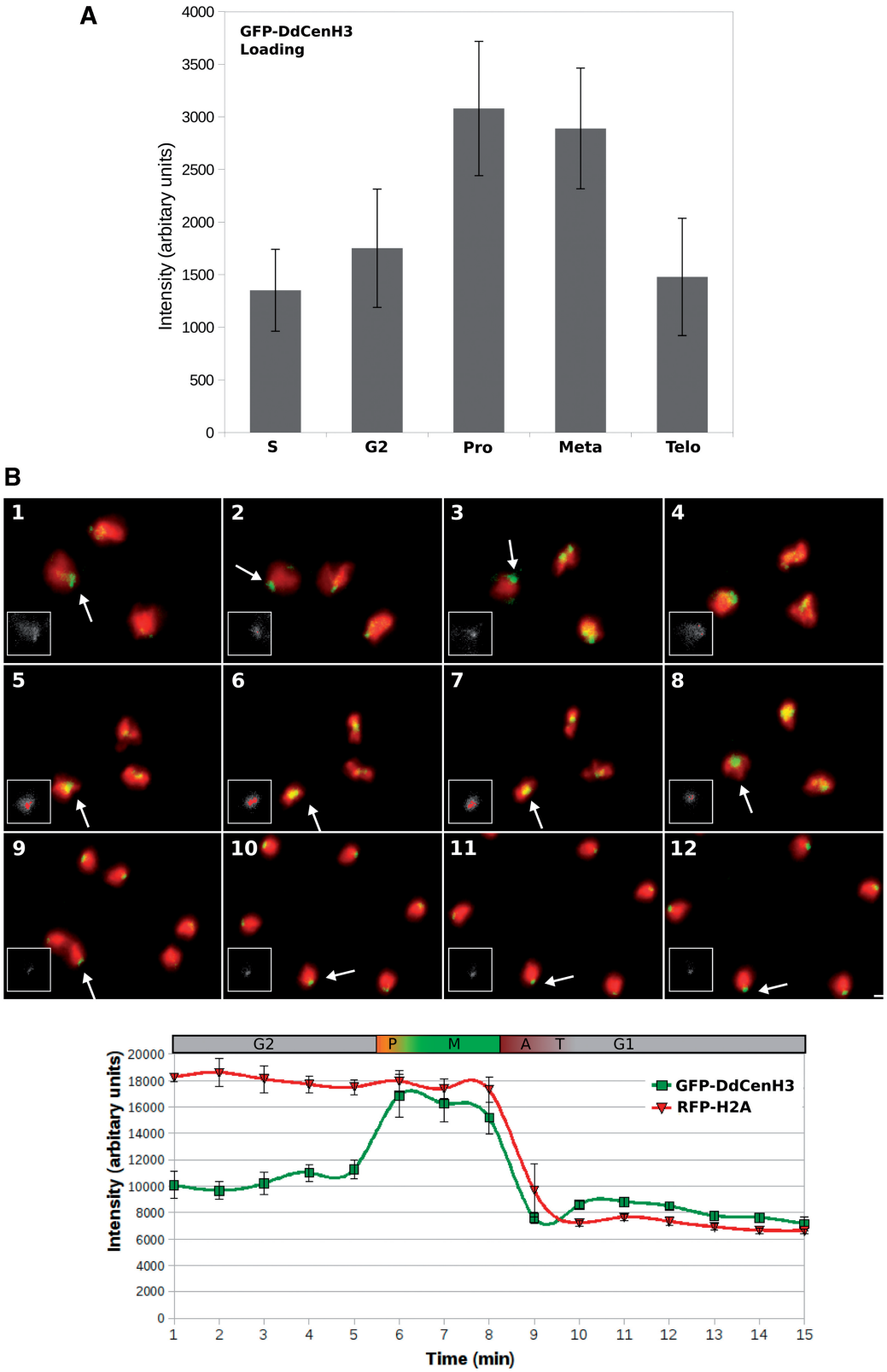

Figure 4. Loading of GFP-DdCenH3 onto centromeres. (**A**) Cells co-expressing GFP-DdCenH3 and the cell cycle marker RFP-PCNA were fixed and the normalized intensity of the GFP-DdCenH3 signal at the centromeres at different stages of the cell cycle was plotted. (**B**) Time-lapse images of a cell with three nuclei co-expressing GFP-DdCenH3 and RFP-H2AX. Images were acquired every 20 s and the overlay of the GFP and RFP signals from every third acquisition (every 60 s) is displayed. The GFP signal for one centromere at each time point (indicated with an arrow) is shown as an insert in grey scale. Signal intensities above a certain arbitrary threshold are shaded red. The average signal intensities from the three nuclei averaged over three time points (1 min) are plotted together with the standard error. The bar above the charts indicates the stage of the cell cycle Scale bar = 1 μm. Note that the values in the graph are calculated from a minimum of nine individual data points and do not strictly correspond to the single-intensity data displayed in the inserts.

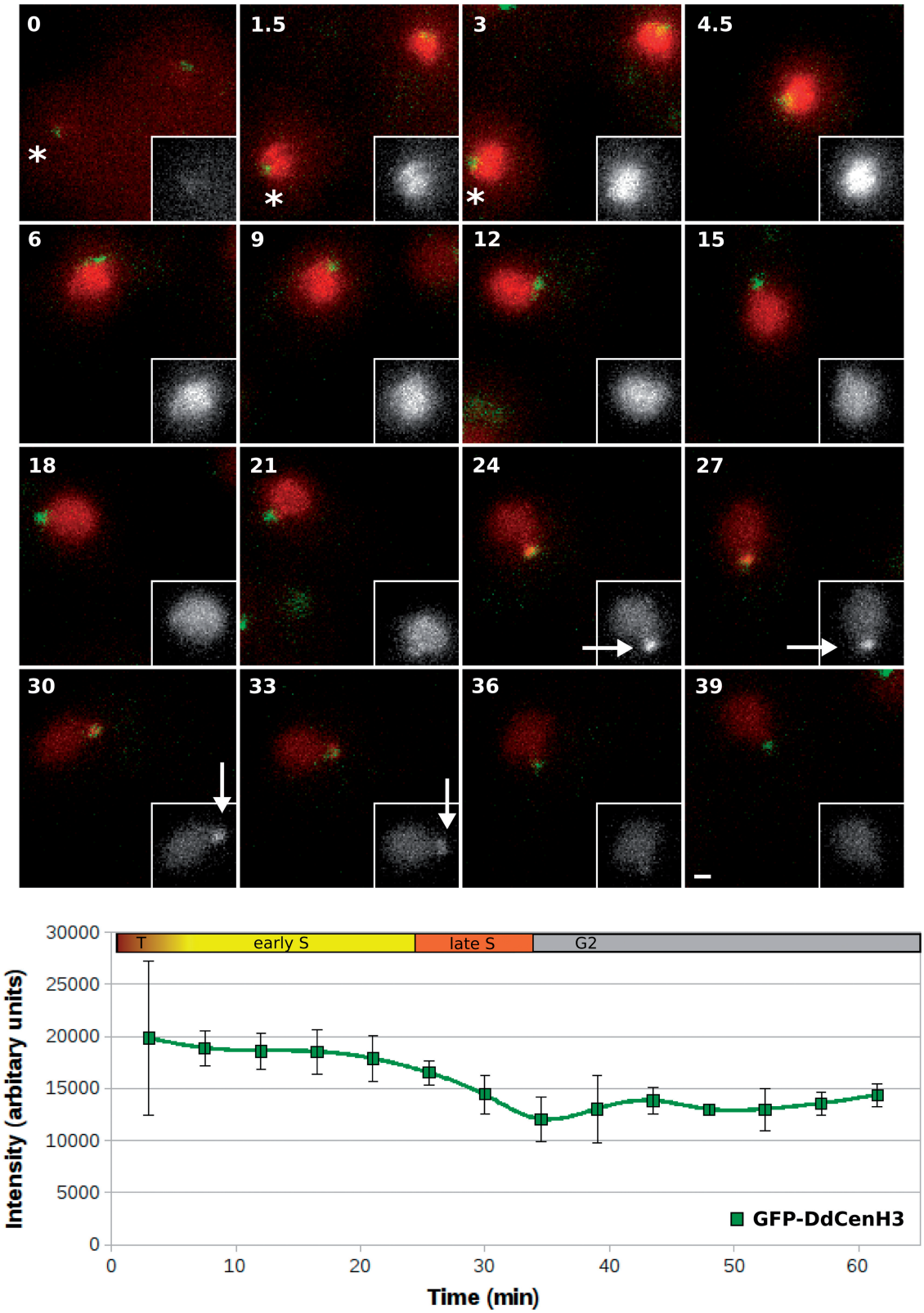

Figure 5. S-phase Behaviour of GFP-DdCenH3. Time-lapse images of cells co-expressing GFP-DdCenH3 and RFP-PCNA starting from telophase. Images were acquired every 90 s and the overlay of the GFP an RFP signals is displayed. The time elapsed since the start of the acquisition is indicated in the upper left hand corner of each frame. The RFP signal for one centromere at each time point (indicated with an asterisk) is shown as an insert in greyscale. The strongly stained spot of RFP-PCNA characteristic of late S-phase is indicated with an arrow. The intensity of the GFP-DdCenH3 signal is averaged over three time points (4.5 min) and plotted below. The bar above the chart indicates the stage of the cell cycle. Scale bar = 1 μm.

ACKNOWLEDGEMENTS

We thank Johnathan Chubb for generously providing the PCNA constructs and Gernot Glöckner for communicating data prior to publication. Markus Maniak is acknowledged for supplying various reagents, for advice and helpful discussions. Reinhard Lührmann, Wolfgang Fischel and Christian Stegmann are acknowledged for giving access to the qPCR machine and for support. We thank Ulrich Rothbauer for generously providing the Alpaca GFP antibody. This work would not have been possible without the bioinformatics resources provided by DictyBase and material provided by the Dicty stock centre.

FUNDING

Funding for open access charge: Deutsche Forschungsgemeinschaft grant (Ne 285/8, SPP 1129 to W.N.).

Conflict of interest statement. None declared.

REFERENCES

1. Gieni,R.S., Chan,G.K. and Hendzel,M.J. (2008) Epigenetics regulate centromere formation and kinetochore function. *J. Cellular Biochem.*, **104**, 2027–2039.
2. Tyler-Smith,C. and Floridia,G. (2000) Many paths to the top of the mountain, diverse evolutionary solutions to centromere structure. *Cell*, **102**, 5–8.
3. Ekwall,K. (2007) Epigenetic control of centromere behavior. *Annu. Rev. Genet.*, **41**, 63–81.
4. Cooper,J.L. and Henikoff,S. (2004) Adaptive evolution of the histone fold domain in centromeric histones. *Mol. Biol. Evol.*, **21**, 1712–1718.
5. Malik,H.S. and Henikoff,S. (2001) Adaptive evolution of cid, a centromere-specific histone in *Drosophila*. *Genetics*, **157**, 1293–1298.
6. Sanyal,K., Baum,M. and Carbon,J. (2004) Centromeric DNA sequences in the pathogenic yeast Candida albicans are all different and unique. *Proc. Natl Acad. Sci.*, **101**, 11374–11379.
7. Henikoff,S., Ahmad,K. and Malik,H.S. (2001) The centromere paradox: stable inheritance with rapidly evolving DNA. *Science*, **293**, 1098–1102.
8. Zofall,M. and Grewal,S.I.S. (2006) RNAi-mediated heterochromatin assembly in fission yeast. *Cold Spring Harb. Symp. Quant. Biol.*, **71**, 487–496.
9. Sullivan,K.F., Hechenberger,M. and Masri,K. (1994) Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. *J. Cell. Biol.*, **127**, 581–592.
10. Black,B.E. and Bassett,E.A. (2008) The histone variant CENP-A and centromere specification. *Curr. Opin. Cell Biol.*, **20**, 91–100.
11. Vermaak,D., Hayden,H.S. and Henikoff,S. (2002) Centromere targeting element within the histone fold domain of Cid. *Mol. Cell. Biol.*, **22**, 7553–7561.
12. Lermontova,I., Schubert,V., Fuchs,J., Klatte,S., Macas,J. and Schubert,I. (2006) Loading of *Arabidopsis* centromeric histone CENH3 occurs mainly during g2 and requires the presence of the histone fold domain. *Plant Cell*, **18**, 2443–2451.
13. Folco,H.D., Pidoux,A.L., Urano,T. and Allshire,R.C. (2008) Heterochromatin and RNAi are required to establish CENP-A chromatin at centromeres. *Science*, **319**, 94–97.
14. Glöckner,G. and Heidel,A.J. (2009) Centromere sequence and dynamics in *Dictyostelium discoideum*. *Nucleic Acids Res.*, **37**, 1809–1816.
15. Aravind,L., Anantharaman,V. and Venancio,T.M. (2009) Apprehending multicellularity: Regulatory networks, genomics, and evolutiony. *Birth Defects Res. Part C: Embryo Today: Rev.*, **87**, 143–164.
16. Galagan,J.E., Henn,M.R., Ma,L., Cuomo,C.A. and Birren,B. (2005) Genomics of the fungal kingdom: Insights into eukaryotic biology. *Genome Res.*, **15**, 1620–1631.
17. Eichinger,L., Pachebat,J.A., Glöckner,G., Rajandream,M.A., Sucgang,R., Berriman,M., Song,J., Olsen,R., Szafranski,K. and Xu,Q. (2005) The genome of the social amoeba *Dictyostelium discoideum*. *Nature*, **435**, 43–57.
18. Martens,H., Novotny,J., Oberstrass,J., Steck,T.L., Postlethwait,P. and Nellen,W. (2002) RNAi in *Dictyostelium*: the role of RNA-directed RNA polymerases and double-stranded RNase. *Mol. Biol. Cell*, **13**, 445–453.
19. Kaller,M., Euteneuer,U. and Nellen,W. (2006) Differential effects of heterochromatin Protein 1 isoforms on mitotic chromosome distribution and growth in *Dictyostelium discoideum*. *Eukaryotic Cell*, **5**, 530–543.
20. Kaller,M., Földesi,B. and Nellen,W. (2007) Localization and organization of protein factors involved in chromosome inheritance in *Dictyostelium discoideum*. *Biol. Chem.*, **388**, 355–365.
21. Kuhlmann,M., Borisova,B.E., Kaller,M., Larsson,P., Stach,D., Na,J., Eichinger,L., Lyko,F., Ambros,V., Soderbom,F. *et al.* (2005) Silencing of retrotransposons in *Dictyostelium* by DNA methylation and RNAi. *Nucleic Acids Res.*, **33**, 6405–6417.
22. Gräf,R., Daunderer,C. and Schulz,I. (2004) Molecular and functional analysis of the *Dictyostelium* centrosome. *Int. Rev. Cytol.*, **241**, 155–202.
23. Chubb,J.R., Bloomfield,G., Xu,Q., Kaller,M., Ivens,A., Skelton,J., Turner,B.M., Nellen,W., Shaulsky,G., Kay,R.R. *et al.* (2006) Developmental timing in *Dictyostelium* is regulated by the Set1 histone methyltransferase. *Dev. Biol.*, **292**, 519–532.
24. Watts,D.J. and Ashworth,J.M. (1970) Growth of myxameobae of the cellular slime mould *Dictyostelium discoideum* in axenic culture. *Biochem. J.*, **119**, 171–174.
25. Howard,P.K., Ahern,K.G. and Firtel,R.A. (1988) Establishment of a transient expression system for *Dictyostelium discoideum*. *Nucleic Acids Res.*, **16**, 2613–2623.
26. Karlin,S. and Altschul,S.F. (1993) Applications and statistics for multiple high-scoring segments in molecular sequences. *Proc. Natl Acad. Sci. USA*, **90**, 5873–5877.
27. Altschul,S.F., Madden,T.L., Schäffer,A.A., Zhang,J., Zhang,Z., Miller,W. and Lipman,D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
28. Thompson,J.D., Higgins,D.G. and Gibson,T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673–4680.
29. Huelsenbeck,J.P., Ronquist,F., Nielsen,R. and Bollback,J.P. (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. *Science*, **294**, 2310–2314.
30. Huelsenbeck,J.P. and Ronquist,F. (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, **17**, 754–755.
31. Thompson,J.D., Gibson,T.J., Plewniak,F., Jeanmougin,F. and Higgins,D.G. (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.*, **25**, 4876–4882.
32. Huson,D.H., Richter,D.C., Rausch,C., Dezulian,T., Franz,M. and Rupp,R. (2007) Dendroscope: an interactive viewer for large phylogenetic trees. *BMC Bioinformatics*, **8**, 460.
33. Moerman,A.M. and Klein,C. (1998) *Dictyostelium discoideum* Hsp32 is a resident nucleolar heat-shock protein. *Chromosoma*, **107**, 145–154.
34. Cappello,J., Handelsman,K. and Lodish,H.F. (1985) Sequence of Dictyostelium DIRS-1: an apparent retrotransposon with inverted terminal repeats and an internal circle junction sequence. *Cell*, **43**, 105–115.
35. Lysak,M., Fransz,P. and Schubert,I. (2006) Cytogenetic analyses of *Arabidopsis*. *Methods Mol. Biol.*, **323**, 173–186.
36. Rothbauer,U., Zolghadr,K., Muyldermans,S., Schepers,A., Cardoso,M.C. and Leonhardt,H. (2008) A versatile nanotrap for

biochemical and functional studies with fluorescent fusion proteins. *Mol. Cell Proteomics*, **7**, 282–289.
37. Black,B.E., Jansen,L.E., Maddox,P.S., Foltz,D.R., Desai,A.B., Shah,J.V. and Cleveland,D.W. (2007) Centromere identity maintained by nucleosomes assembled with histone H3 containing the CENP-A targeting domain. *Molecular Cell*, **25**, 309–322.
38. Schulz,I., Baumann,O., Samereier,M., Zoglmeier,C. and Gräf,R. (2009) *Dictyostelium* Sun1 is a dynamic membrane protein of both nuclear membranes and required for centrosomal association with clustered centromeres. *Eur. J. Cell Biol.*, **88**, 621–638.
39. Appelgren,H., Kniola,B. and Ekwall,K. (2003) Distinct centromere domain structures with separate functions demonstrated in live fission yeast cells. *J. Cell. Sci.*, **116**, 4035–4042.
40. Li,F., Sonbuchner,L., Kyes,S.A., Epp,C. and Deitsch,K.W. (2008) Nuclear non-coding RNAs are transcribed from the centromeres of *Plasmodium falciparum* and are associated with centromeric chromatin. *J. Biol. Chem.*, **283**, 5692–5698.
41. Dawson,S., Sagolla,M. and Cande,W. (2007) The cenH3 histone variant defines centromeres in *Giardia intestinalis*. *Chromosoma*, **116**, 175–184.
42. Gräf,R., Daunderer,C. and Schliwa,M. (2000) *Dictyostelium* DdCP224 is a microtubule-associated protein and a permanent centrosomal resident involved in centrosome duplication. *J. Cell. Sci.*, **113 (Pt 10)**, 1747–1758.
43. Loomis,W.F., Welker,D., Hughes,J., Maghakian,D. and Kuspa,A. (1995) Integrated maps of the chromosomes in *Dictyostelium discoideum*. *Genetics*, **141**, 147–157.
44. Schulz,I., Erle,A., Gräf,R., Krüger,A., Lohmeier,H., Putzler,S., Samereier,M. and Weidenthaler,S. (2009) Identification and cell cycle-dependent localization of nine novel, genuine centrosomal components in *Dictyostelium discoideum*. *Cell Motil Cytoskeleton*, **66**, 915–928.
45. Schuh,M., Lehner,C.F. and Heidmann,S. (2007) Incorporation of *Drosophila* CID/CENP-A and CENP-C into centromeres during early embryonic anaphase. *Curr. Biol.*, **17**, 237–243.
46. Jansen,L.E., Black,B.E., Foltz,D.R. and Cleveland,D.W. (2007) Propagation of centromeric chromatin requires exit from mitosis. *J. Cell Biol.*, **176**, 795–805.
47. Foltz,D.R., Jansen,L.E., Bailey,A.O., Yates,Y.R. III, Bassett,E.A., Wood,S., Black,B.E. and Clevel,D.W. (2009) Centromere-specific assembly of CENP-A nucleosomes is mediated by HJURP. *Cell*, **137**, 472–484.
48. Dunleavy,E.M., Roche,D., Tagami,H., Lacoste,N., Ray-Gallet,D., Nakamura,Y., Daigo,Y., Nakatani,Y. and Almouzni-Pettinotti,G. (2009) HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. *Cell*, **137**, 485–497.
49. Lermontova,I., Fuchs,J., Schubert,V. and Schubert,I. (2007) Loading time of the centromeric histone H3 variant differs between plants and animals. *Chromosoma*, **116**, 507–510.
50. Muramoto,T. and Chubb,J.R. (2008) Live imaging of the Dictyostelium cell cycle reveals widespread S phase during development, a G2 bias in spore differentiation and a premitotic checkpoint. *Development*, **135**, 1647–1657.
51. Luger,K., Mader,A.W., Richmond,R.K., Sargent,D.F. and Richmond,T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*, **389**, 251–260.
52. Sanchez-Pulido,L., Pidoux,A.L., Ponting,C.P. and Allshire,R.C. (2009) Common ancestry of the CENP-A chaperones Scm3 and HJURP. *Cell*, **137**, 1173–1174.