

ARTICLE

Comparative study of artificial chromosome centromeres in human and murine cells

Daniela Moralli¹, Andrew Jefferson^{1,2}, Emanuela Valeria Volpi¹ and Zoia Larin Monaco^{*1}

Human artificial chromosomes (HAC) are a valuable tool in the analysis of complex chromatin structures such as the human centromere because of their small size and relative simplicity compared with normal human chromosomes. This report includes a comprehensive study of the centromere and chromatin composition of HAC, expressing human genes, generated in human cells and transferred to murine cells. The analysis involved chromatin immuno-precipitation and immuno-FISH on metaphase chromosomes and chromatin fibres. In both the cell types, the HAC consisted of alphoid and non-alphoid DNA and were mainly euchromatic in composition, although a pericentromeric heterochromatic region was present on all the HAC. Fibre-FISH and chromatin immuno-precipitation data indicated that the position of the centromere differed between HAC in human cells and in murine cells. Our work highlights the importance and utilisation of HAC for understanding the epigenetic aspects of chromosome biology.

European Journal of Human Genetics (2013) **21**, 948–956; doi:10.1038/ejhg.2012.296; published online 13 February 2013

Keywords: human artificial chromosomes; centromere; chromatin immuno-precipitation.

INTRODUCTION

In 1997, it was first demonstrated that introducing a centromere-specific repeat DNA, the alpha (α , alphoid) satellite, in human cells led to the formation of small independent molecules. These were entirely composed of the exogenous DNA and were able to assemble a functional centromere and propagate correctly at each cell division.¹ These structures behaved as normal chromosomes and are known as *de novo* human artificial chromosomes (HAC).² Studies on HAC formation and structure have generated valuable insights in the characterisation of chromosome epigenetic composition and centromere and kinetochore structure.^{2–7}

Human chromosomes contain a distinct centromeric domain, which encompasses a block of highly repetitive alpha satellite DNA, and it is characterised by a specific chromatin signature and associated centromeric proteins, including the centromere protein A (CENP A).⁴ CENP A is an epigenetic marker of centromeric chromatin, which promotes centromere identity, and is essential for the chromatin organisation at the core centromere domain. The protein is a histone H3 variant, and the CENP A nucleosomes assemble in an alternate arrangement with histone H3 dimethylated at lysine 4 (H3diK4) nucleosomes to generate a region sharing some of the characteristics of euchromatin. Flanking the core centromeric domain at the pericentromeric region is a distinct area of heterochromatin where predominantly histone H3 dimethylated at lysine 9 binds.⁴

In human cultured cells, the HAC centromeric domain and chromatin organisation mirror the domain structure and chromatin formation of human centromeres.^{4–7} *De novo* HAC have a similar distinct continuous centromeric domain, which spreads over heterogeneous blocks of alphoid and vector DNA. In this region, the HAC centromeric chromatin is assembled at the core, with CENP A alternating with H3diK4 generating a CEN region. The centromeric

domain is also flanked by a pericentromeric DNA forming a heterochromatic region where predominantly histone H3 trimethylated at lysine 9 (H3triK9) binds.

The purpose of this work was to compare HAC structure and composition in human and murine cells. We investigated the centromeric domain and chromatin organisation of *de novo* HAC, which were first generated in human cells then transferred to murine cells by microcell-mediated chromosome transfer (MMCT). Using immuno-FISH and chromatin immuno-precipitation (ChIP) with centromere-specific antibodies, we detected the presence of a distinct centromeric domain on the all HAC spanning alphoid and non-alphoid DNA regions. However, in murine cells, the centromeric region spanned a different region on the HAC. The HAC chromatin showed the same level of organisation as in human chromosomes, with regions of euchromatin at the core centromere and flanking heterochromatin and was independent of HAC size.

MATERIALS AND METHODS

Cell culture

All cell lines were cultured in D-MEM medium supplemented with penicillin/streptomycin, 10% FCS and maintained under selection as previously published.^{8–12} In this study, the human HSV-1 HAC HF15.1 was transferred into murine LA9 and STO cells by MMCT, and the murine SM1-1 HAC was transferred into HT1080 cells as previously detailed.⁹ Selection was applied 24 h after the MMCT as follows: HF15.1 \times LA9/STO G418 650 mg/ml, ouabain 1 mM; SM1-1 \times HT1080 125 μ g/ml G418, 250 ng/ml puromycin, 0.5 \times HAT. After 15–20 days, clones were isolated and expanded.

Cytological analysis

The FISH, fibre-FISH, and immuno-FISH on metaphase spreads were conducted as previously described.^{8–9} The following antibodies were

¹Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK

*Correspondence: Dr Z Larin Monaco, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK. Tel: +44 0 1865287502; Fax: +44 0 1865287501; E-mail: Zoia.Monaco@tufts.edu or zoia@well.ox.ac.uk

²Present address: Department of Cardiovascular Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, USA.

Received 29 August 2012; revised 8 November 2012; accepted 11 December 2012; published online 13 February 2013

used: anti-human CENP-A (Abcam, Cambridge, UK); anti-mouse CENP-A;¹³ anti-Aurora B (BD Biosciences, Oxford, UK); anti-SMC3 (Abcam); anti-histone H3 dimethyl lysine 4 (Upstate; Merck Millipore, Watford, UK); anti-histone H3 trimethyl lysine 9 (Abcam); and anti-histone H3 phospho serine 10 (Upstate).

Chromatin fibres for immuno-FISH were prepared as described.¹⁴ Briefly, cells were resuspended in KCl 75 mM and spun onto slides at 800 rpm for 4 min, using a Shandon Cytospin machine. Chromatin was released from nuclei by treatment in a urea lysis buffer (25 mM urea, 25 mM Tris-HCl pH 7.5, 0.5 M NaCl, 1% Triton X-100), and fixed in 2% formaldehyde in PBS. After detection of the primary antibody, the chromatin was fixed for a second time in 2% formaldehyde in PBS, followed by a 20-s treatment in 0.2 N NaOH. Finally, probe and chromatin DNA were denatured simultaneously at 85 °C for 6 min. Post-hybridisation washes and detection were as described previously.⁹ All cytological preparations were analysed using an Olympus BX60 microscope for epifluorescence equipped with a Sensys CCD camera (Photometrics, Tucson, AZ, USA). Images were collected using either MacProbe 4.3 (Perceptive Systems, Houston, TX, USA) or Genus Cytovision software (Leica Microsystems, Milton Keynes, UK). Some images were pseudo-coloured using Adobe Photoshop 12.0 (San Jose, CA, USA).

ChIP and real-time quantitative PCR analysis

Native chromatin was prepared from 10 million cells by lysis in TBS buffer (10 mM Tris-HCl pH 7.5, 150 mM NaCl, 3 mM CaCl₂, 2 mM MgCl₂), 0.5% Tween 40, followed by homogenisation with a Dounce homogeniser. Nuclei were separated from cellular debris by centrifugation on a two-phase gradient (25% sucrose in TBS, over 50% sucrose in TBS). Di-trinucleosomes fragments were obtained by brief digestion with micrococcal nuclease (GE-Healthcare, Amersham, UK) 1 unit/ μ g of chromatin and checked on a 1.5% agarose gel.

In all, 100 μ g of chromatin were then immuno-precipitated overnight using 40 μ g of anti-human CENP-A (ChIP Grade, Abcam) for the human clones, or 100 μ g of anti-mouse CENP-A for the mouse clones, in incubation buffer (50 mM NaCl, 20 mM Tris HCl pH 7.5, 5 mM EDTA, 1 \times Protease Inhibitor Complete, Roche, Burgess Hill, UK). An equal amount of chromatin was treated in the same way, but without any antibody (mock sample). After overnight incubation at 4 °C, the chromatin fraction bound by the antibody was purified using 200 μ l of either 50%v/v protein A-Sepharose (GE Healthcare) or 50%v/v protein G-Agarose (Upstate), salmon sperm DNA blocked. Following a 3-h incubation, the agarose/sepharose beads were pelleted and washed in buffers at increasing NaCl content (50 mM Tris HCl pH7.5, 10 mM EDTA, 50 mM NaCl; 50 mM Tris HCl pH7.5, 10 mM EDTA, 100 mM NaCl; 50 mM Tris HCl pH7.5, 10 mM EDTA, 150 mM NaCl). The bound chromatin was then released by washing twice in incubation buffer, 1% SDS. The DNA was purified by phenol-chloroform extraction followed by precipitation, using 50 μ g of glycogen as carrier. The relative enrichment of the IP_{CENP-A} DNA versus IP_{mock} DNA was tested by quantitative real-time PCR, using the Sybr GreenER kit (Invitrogen, Paisley, UK) on a ICycler machine (Bio-Rad, Hemel Hempstead, UK). The primer sets are listed in Supplementary Table S1. The Ct values obtained were compared using the 2^{- $\Delta\Delta$ Ct} method. For each cell line, at least four batches of N-chromatin were analysed, in three replicate independent ChIP experiments. The real-time PCR quantifications were conducted in triplicate.

RESULTS

In this study, we characterised the structure, centromere formation, chromatin composition and stability of three human HAC and compared them with their murine derivatives, obtained following MMCT to LA-9 or STO cells (Figure 1a). In previous work, the AG6-1 HAC was obtained in HT1080 cells, following lipofection delivery of an input vector containing 220 kb of 17 alpha DNA and the whole *HPRT* genomic locus (200 kb).⁸ The AG6-1 HAC was transferred by MMCT to STO murine cells, to generate Sag1.2 (Moralli *et al*⁹). The LJ2-1 HAC was obtained in HT1080 from a vector containing 220 kb of 17 α DNA on a pBeloBAC11 backbone.¹⁰ Following transfer to STO cells by MMCT, we obtained the LJ2-1 murine derivative SM1-1

(Alazami *et al*¹¹). The HF15.1 HAC was generated in HT1080 cells using the HSV-1 amplicon delivery system¹⁵ and contained chromosome 21 α DNA (60 kb) the *GFP* and the *HPRT* minigene.¹² In this report, the HF15.1 HAC was also transferred to LA-9 and STO cells by MMCT to generate HAC containing murine cells. Thirty-nine LA-9 and 10 STO clones were isolated, and HAC were detected in all clones although HAC DNA had also integrated into the murine host chromosomes. The LHF2 HAC (LA-9) was chosen for further studies as it had the highest percentage of independent HAC (20%). All clones used in this study contained on average one HAC per cell (Figure 1a). For clarity purposes, the names of all HAC have been simplified as reported in Table 1, where each HAC in human cells has been identified by a number dash H, and the corresponding murine HAC are defined by the same number dash M.

HAC fibre-FISH analysis

The HAC DNA organisation was characterised by fibre-FISH on extended DNA fibres (Figure 1b), which were hybridised to the following probes: 17 α and *HPRT* genomic DNA for HAC 1-H and 1-M1; 17 α and pBeloBAC11 DNA for HAC 2-H and 2-M; 21 α DNA and a vector carrying the *HPRT* minigene (pHGHPR12) for HAC 3-H. The presence of integrated DNA in HAC 3-M prevented the analysis of the HAC structure by fibre-FISH, as at this level the HAC and integration signals are undistinguishable.

Similar to the analysis of HAC in previous work,⁸ HAC 1-H and its murine derivative HAC 1-M1 contained 17 α DNA interspersed with *HPRT* sequences. The FISH signals were of variable length, confirming that the input DNA had undergone some amplification and rearrangement (Figure 1b). Following the transfer from 1-H to 1-M1, the size of the HAC, as estimated by measuring the length of at least 20 fibres in pixels using ImageJ software (NIH, Bethesda, MD, USA) (Table 2), did not change significantly ($P=0.4759$, Student's *t*-test for independent samples). The HAC 2-H fibres showed an alternate arrangement of 17 α DNA interspersed with vectors sequences. The composition of the murine derivative HAC 2-M (Figure 1b) differed from the HAC 2-H, and contained a higher ratio of vectors sequences, as the number of FISH vector signals from the murine derivative HAC 2-M fibre was approximately five times higher than the vector signals in human HAC 2-H (as measured using ImageJ software). The overall length of the HAC 2-M fibre was similar to 2-H ($P=0.194$, Student's *t*-test for independent samples; Table 2). The 3-H HAC fibres were also composed of alternating 21 α and pHGHPR12 DNA, yet the fibres were approximately 8.6 times shorter than in the HAC 1-H (Table 2) and 7 times shorter than HAC 2-H (Table 2).

In summary, the fibre-FISH experiments showed that all HAC were composed of alternating stretches of the input DNA and that there was no change in HAC length following transfer from a human background to murine cells.

Centromere localisation

To identify where the HAC centromere was localised, immuno-FISH experiments were carried out on fibres from the HAC 1-H, 1-M1, 2-H, 2-M and 3-H, by staining with an antibody against CENP A, a histone H3 variant, which marks centromeric chromatin. The region of the HAC fibre that bound CENP-A was identified by FISH using different components of the HAC DNA as probes (ie, 17 α and *HPRT* genomic DNA for HAC 1-H and 1-M1; 17 α and pBeloBAC11 vector for HAC 2-H and 2-M; 21 α and *HPRT* minigene DNA for HAC 3-H). The results are shown in Figure 1c. In 1-H cells, the HAC centromere (green signal) was formed both on aliphoid and non-aliphoid DNA

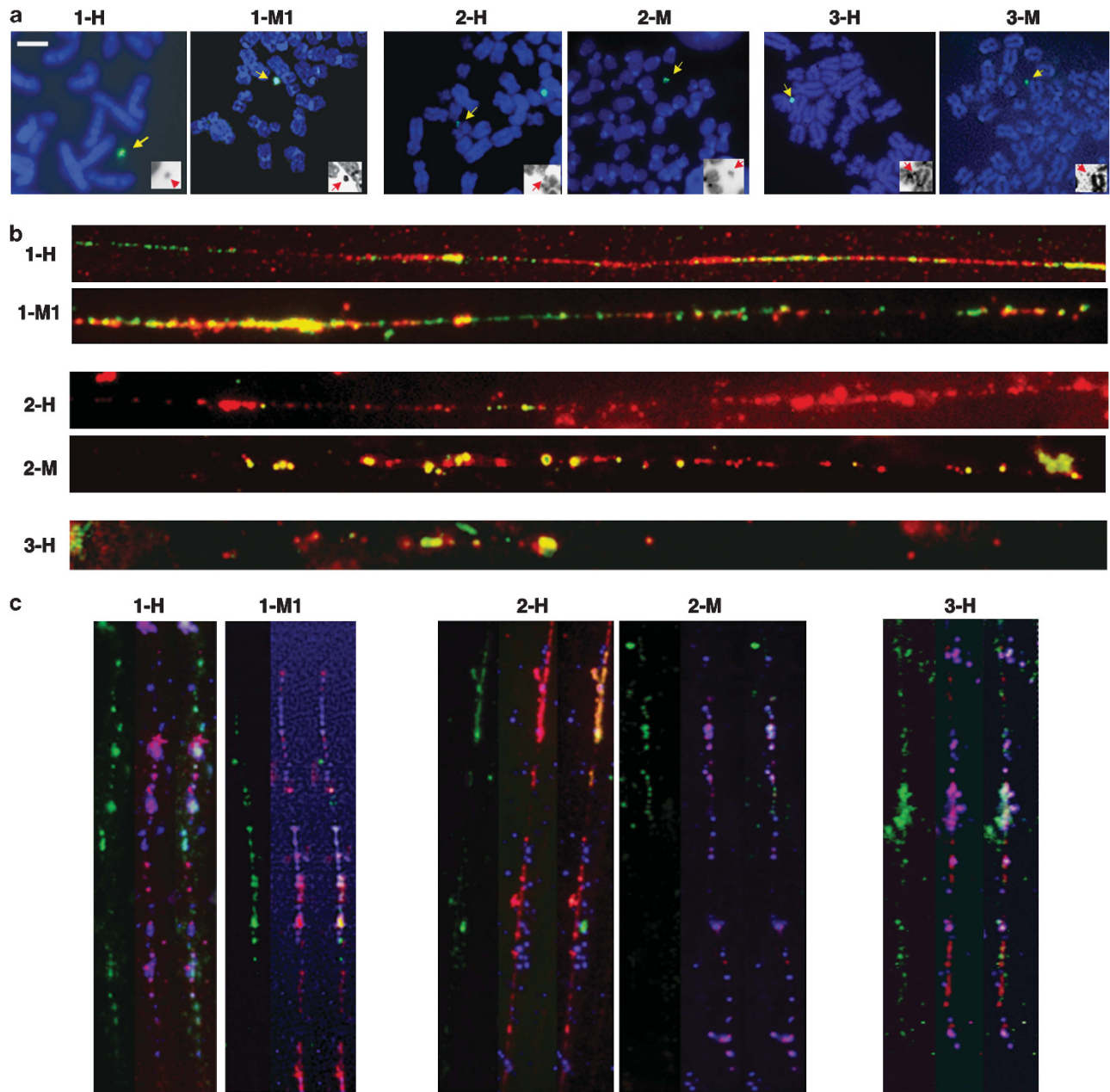


Figure 1 Fluorescence *in situ* hybridisation (FISH) on chromosome metaphase spreads and HAC DNA fibres. The human HAC are grouped with their murine derivatives. (a) FISH on metaphase spreads with a HAC-specific probe (HAC 3-H and 3-M, green signal) or 17 α DNA (HAC 1-H, 1-M1, 2-H, 2-M, green signal). Chromosomes are counterstained by DAPI (4,6-diamidino-2-phenylindole; blue). Yellow arrows denote the HAC. The inset shows DNA staining only, with the HAC identified by the red arrow. (b) Fibre-FISH on DNA fibres, hybridised with alphoid DNA (red signal, 17 α for human HAC 1-H, 2-H, and murine HAC 1-M1, 2-M; 21 α for human HAC 3-H) and HPRT genomic DNA (green signal, HAC 1-H and 1-M1) or vector DNA (green signal, HAC 2-H, 2-M and 3-H). (c) Immuno-fibre-FISH analysis of chromatin fibres. On all the HAC, CENP A (green signal) localises to areas containing both alphoid DNA (red signal, 17 α , HAC 1-H, 2-H, 1-M1, 2-M; 21 α HAC 3-H) and non-alphoid (blue signal, HPRT genomic in human HAC 1-H and murine derivative HAC 1-M1; vector DNA in human HAC 2-H, 3-H and murine 2-M). For each clone, the left panel shows CENP A signal only, the middle panel the FISH signal only (alpha and non alpha DNA probes) and the right panel the merged picture.

and no major difference was observed between the HAC 1-H and the murine derivative HAC 1-M1 using this technique. Similarly, the signal from the human HAC 2-H included both alphoid and pBeloBAC11 vector DNA, yet, a shorter tract of non-alphoid/vector DNA signal in the region bound by CENP A was detected, suggesting that the centromere was localised mostly on alphoid DNA. However, in the murine derivative HAC 2-M, the pattern of immuno-FISH signals indicated that the region of non-alphoid DNA forming the

centromere was greater than in HAC 2-H, yet, the majority of the CENP A-binding region was on alphoid DNA. Although the HAC 3-H fibres were shorter than the HAC 1-H and 2-H fibres, the CENP-A signal covered most of the fibre and spanned both alphoid and non-alphoid DNA.

To confirm the immuno-fibre-FISH observations, chromatin immuno-precipitation experiments were undertaken to identify the centromeric region with higher precision. Native unfixed chromatin

Table 1 HAC features

HAC clone name	Composition	Cell line	Simplified name
AG6-1	17 α DNA, HPRT genomic locus, G418 resistance gene	HT1080 (human)	1-H
Sag1.2	17 α DNA, HPRT genomic locus, G418 resistance gene	STO (murine)	1-M1
Sag1.1	17 α DNA, HPRT genomic locus, G418 resistance gene	STO (murine)	1-M2
Sag2.2	17 α DNA, HPRT genomic locus, G418 resistance gene	STO (murine)	1-M3
Sag2.3	17 α DNA, HPRT genomic locus, G418 resistance gene	STO (murine)	1-M4
LJ2-1	17 α DNA, G418 resistance gene	HT1080 (human)	2-H
SM1-1	17 α DNA, G418 resistance gene	STO (murine)	2-M
HF15.1	21 α , HPRTminigene, GFP and G418 resistance genes	HT1080 (human)	3-H
LHF2	21 α , HPRTminigene, GFP and G418 resistance genes	LA9 (murine)	3-M

Table 2 Length measurement of HAC chromatin fibres (in pixel)

Fibre	Average length (pixel)	SE
HAC 1-H (h)	600	± 31.9
HAC 1-M1 (m)	680	± 54.8
HAC 2-H (h)	490	± 24.4
HAC 2-M (m)	540	± 23.6
HAC 3-H (h)	70	± 8.24

Abbreviations: h, human parental HAC; m, murine derivative HAC; SE, standard error.

was prepared from all cell lines and digested with micrococcal nuclease to obtain fragments corresponding on average to di/tri-nucleosomes. Centromere-associated DNA was then co-immunoprecipitated using an anti-CENP A antibody. As a control, equal amounts of chromatin were subjected to the same treatments but without any antibody (mock). The ChIP DNA was then analysed by quantitative real-time PCR, using primers designed on alpha satellite DNA, the vector sequence or on the HPRT genomic locus. The relative enrichments in IP_{CENPA} DNA versus IP_{mock} DNA for all the primer sets are detailed in Figure 2, and in Supplementary Table S2. As the alpha satellite primers amplify the endogenous chromosomes centromere too, the $-$ fold increase was generally much higher than that observed for non-centromeric sequences. For this reason, the positive (alpha satellite) and negative (18S) controls are not shown on the graphs in Figure 2 but are presented in Supplementary Table S2.

In agreement with the immuno-fibre-FISH results, the ChIP data indicated that in HAC 1-H, the region that formed the centromere spanned a large portion of the pBeloBAC11 vector and parts of the HPRT genomic region as well as the alphoid DNA. In HAC 1-M1, the centromere was formed on pBeloBAC11 vector DNA together with some HPRT DNA, but the non-alphoid region showing enrichment was smaller than in the HAC 1-H (Figure 2). To further analyse possible differences between the 1-H human HAC and its murine counterpart, we characterised three further clones derived from MMCT transfer of the HAC from 1-H to STO cells (HAC 1-M2, 1-M3, 1-M4)⁹ by chromatin immuno-precipitation. The HAC 1-M4 had the CENP A distribution pattern most similar to HAC 1-H, with the centromere formed on a large portion of the vector DNA and parts of the HPRT gene (Figure 3). In HAC 1-M2 most of CENP A was found to associate with the alpha satellite DNA and a small part of one HPRT intron. In the HAC 1-M3, no significant association between CENP A and non-alphoid DNA was identified (Figure 2). In the CENP A precipitated chromatin from HAC 3-H, the highest enrichment occurred on the EGFP gene within the HSV-1 vector

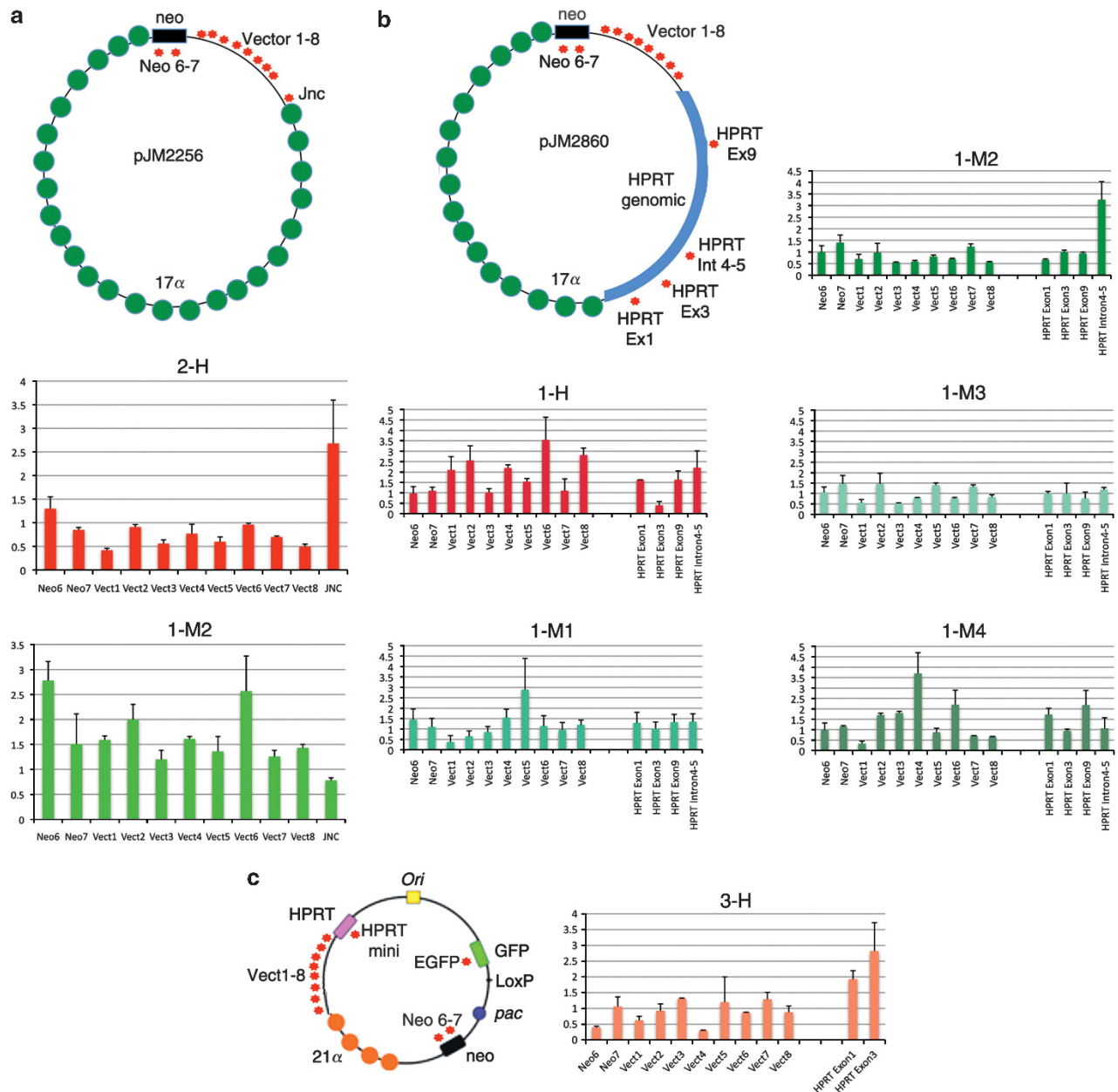
DNA (Figure 2). In agreement with the immuno-fibre-FISH results, the ChIP data indicated that in HAC 2-H only a minimal section of the vector (the junction area between the alpha satellite DNA and the insert) was contained within the centromere, confirming that the centromeric region spanned mainly alphoid DNA. The murine derivative HAC 2-M had the highest amount of non alphoid DNA at the centromere among all HAC, as most of the vector primers analysed were co-precipitated by CENP-A, indicating a shift in the location of the centromere region.

In summary, the quantitative PCR analysis of IP_{CENPA} DNA confirmed that the centromere region in the human and murine HAC was formed on alphoid DNA, yet spanned a large area of non-alphoid DNA that was variable in size in different clones, suggesting a shift in the location of the region.

Centromeric proteins and chromatin analysis

The binding of specific centromere/kinetochore proteins, and the presence of eu/heterochromatin markers on the large HAC 1-H, 1-M1, 2-H and 2-M was previously characterised, and found to be similar to that of the endogenous chromosomes.⁹ To compare the smaller HAC 3-H with the larger HAC, metaphase chromosome spreads were prepared from 1-H, 1-M1, 2-H, 2-M, 3-H and 3-M and analysed with specific centromere antibodies, including CENP-A, Aurora B (AIM-1), and SMC3. Identification of the HAC was confirmed by simultaneous hybridisation with a 17 α or 21 α DNA probes (Figure 3). All HAC bound similar levels of CENP A and the chromosomal passenger protein Aurora B (involved in mitotic checkpoints and chromosome segregation, Figure 3). The level of the protein SMC3 present on the HAC was also investigated as it is involved with chromatid cohesion and separation. Using an antibody against SMC3, all the HAC in human and murine cells bound the protein in various stages of mitosis (Figure 3). No obvious difference was detected in signal intensity between proteins bound to HAC compared with the endogenous chromosomes.

The chromatin composition on the HAC 3-H and 3-M was also characterised in detail and compared with that of the larger HAC (Figure 3). Various reports^{16–17} showed the importance of epigenetic modifications of histone H3 for chromosome stability and centromere formation and maintenance. The histone H3 dimethylated in lysine 4 is a modification typically associated with open chromatin, accessible to transcription. Histone H3diK4 is found at the centromere, where it is interspersed with CENP-A.⁴ In this study, the HAC were labelled by the anti-H3diK4 antibody along the length of the chromosome, suggesting that they were mainly euchromatic in composition (Figure 3). The histone H3 trimethylated



in lysine 9 is found principally on heterochromatin and mostly prevalent in the pericentromeric area. Grimes *et al*¹⁷ showed that some small HAC were depleted in H3triK9. However, the 3-H and 3-M HAC bound the anti-H3triK9 antibody, suggesting that they contained a pericentromeric domain, although the signal intensity appeared lower compared with the intensity of H3diK4 (Figure 3).

Lastly, we investigated the presence of histone H3 phosphorylated in serine 10 (Figure 3). This modification is produced by Aurora B kinase activity and is typically not only associated with chromosome condensation before mitosis¹⁸ but also with DNA transcription. The modification was found on all HAC, at levels similar to the other chromosomes, and no differences were observed between HAC in human and mouse cells.

In summary, the immuno-FISH staining showed that no difference was detected between the HAC and the endogenous chromosomes in human or murine cells, indicating that the chromatin organisation was similar.

Mitotic stability

The stability of the HAC 3-H and 3-M was monitored for at least 150 days in the absence of selection, as previously done for HAC 1-H, 1-M1, 2-H and 2-M.⁹ The daily loss rate is shown in Table 3. The human parental HAC were relatively stable (daily loss rate of 0.07% for HAC 1-H,⁹ 0.04% for HAC 2-H⁹ and 0.125% for HAC 3-H), and the smaller size of the human HAC 3-H did not significantly affect its stability in this study. The derivative

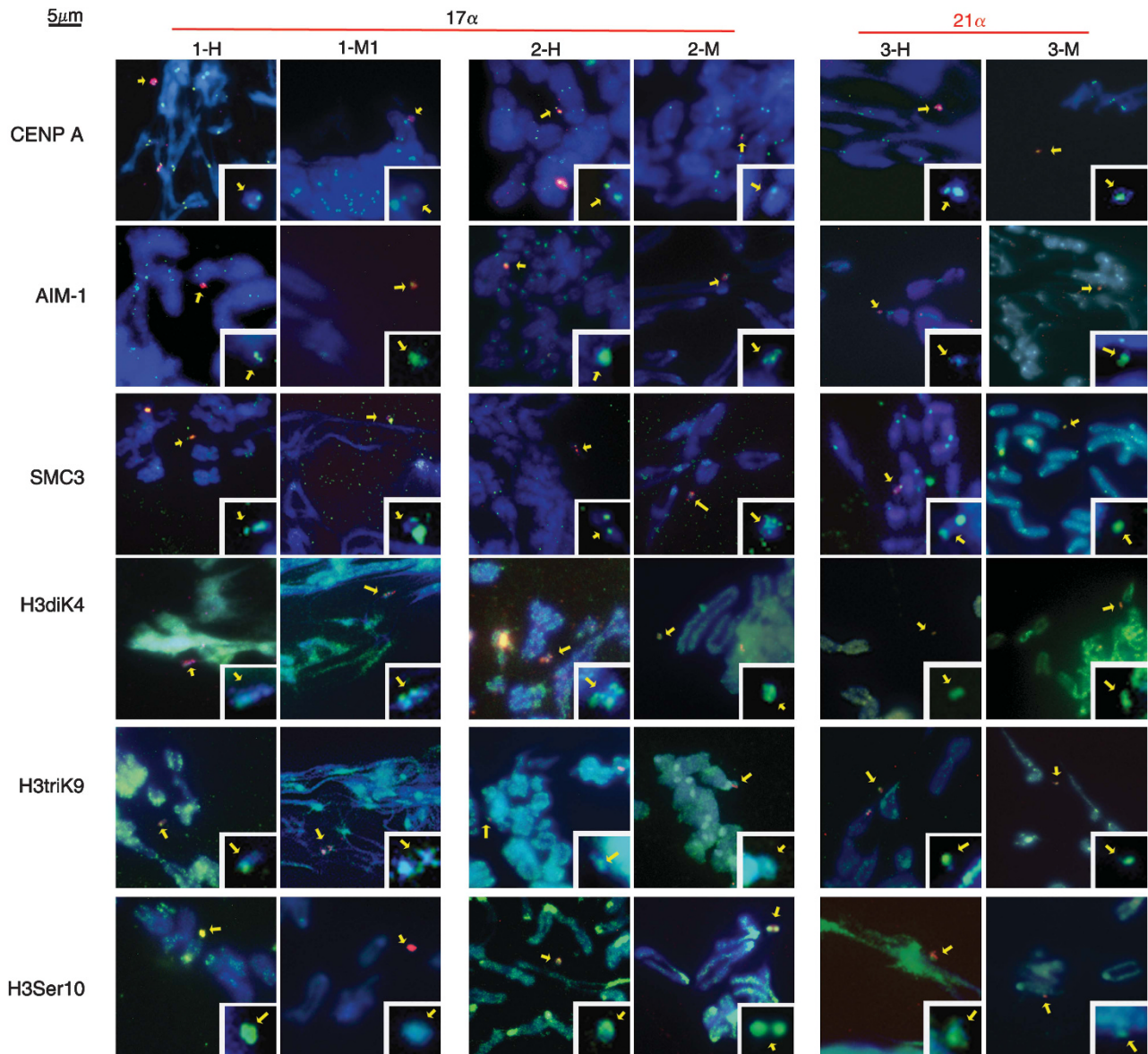


Figure 3 Immunofluorescence (immunofluorescence *in situ* hybridization) on metaphase spreads for each of the HAC clones. The antibody signal is green, and the FISH (either with 17 α or 21 α) signal red. The insets show a three times enlargement of the HAC (yellow arrow) without FISH signal.

HAC in murine cells exhibited an increased level of instability, with the murine derivative HAC 3-M showing an extremely high loss rate (7.4%), resulting in complete loss over a period of 60 days, similarly to the previously observed loss rate of murine HAC 2-H (5.2%)⁹.

Transfer and characterisation of a murine HAC into a human background

To investigate whether the change in structure and/or centromere repositioning were linked to the reduced HAC stability observed in the murine background, the HAC 2-M, showing the highest degree of instability among those fully characterised, was transferred by MMCT back into HT1080 human cells. Following selection, we isolated 33 clones, containing HAC and a variable number of mouse chromosomes. Among the clones, we selected, for further studies, the clone 2-MH, which contained HAC in 81% of the cells, and no mouse chromosomes (Figure 4a).

The stability of the HAC 2-MH was followed over a period of 120 days in the absence of selection. Compared with the murine 2-M, the HAC in 2-MH was very stable, showing a daily loss rate of 0.03%, similar to that observed in the human parental HAC 2-H (Figure 4b).

The centromere positioning of HAC 2-MH was analysed by ChIP, and it was found that CENP A was mostly associated with alpha satellite DNA. The only non-alphaoid, vector sequences co-precipitated by CENP A were the junction area between the alpha satellite DNA and the vector (similar to 2-H) and a portion of the vector immediately adjacent (Figure 4c, Supplementary Table S3).

In summary, these experiment showed that the high level of instability displayed by the HAC 2-M in murine cells was not due to structural rearrangements occurred during the MMCT. Furthermore, the data further indicated that the centromere region may have shifted again to include a larger region of alphaoid DNA following the MMCT transfer of HAC 2-M back into a human background.

DISCUSSION

In this study, we compared the structure, function and chromatin composition of HAC in human and murine cells. All the HAC contained irregular, alternating, blocks of alphoid/non alphoid DNA resulting from amplification and rearrangement of the initial input DNA into human HT1080 cells. The HSV-1 HAC 3-H was up to 7–8 times smaller in size than the HAC 1-H or 2-H. The difference was not fully accounted for by the different size of the input DNA, (400 kb for HAC 1-H, 240 Kb for HAC 2-H and 75 kb for HAC 3-H).

Table 3 HAC loss rate after 150 days in absence of selection, calculated by the formula $N_n = N_0 \times (1-R)^n$, where N_0 is the number of metaphase chromosome spreads showing HAC in the cells cultured under selection, N_n is the number of HAC-containing metaphase chromosome spreads after n days of culture in the absence of selection, and R is the daily rate of loss

HAC	Daily loss rate (%)
1-H (h)	0.07 ^a
1-M1(m)	1.69 ^a
2-H (h)	0.04 ^a
2-M (m)	5.20 ^{a,b}
3-H (h)	0.12
3-M (m)	7.4 ^b

Each table cell shows the human parental HAC (h) and the murine derivative HAC (m).

^aMoralli *et al*⁸.

^bLoss rate after 30 days. The HAC is completely lost after 60 days off selection.

Even though the measurement of HAC's fibre is subject to variations due to different stretching of the chromatin fibre along the glass slide, the results were fairly consistent. This finding suggested that the generation of the HAC 3-H involved a lower degree of amplification of the input DNA. Despite the small size, the HAC 3-H was highly stable. For both HAC 3-H and 3-M, there was no significant difference in centromeric protein-binding capacity or chromatin composition compared with the endogenous host chromosomes in human and mouse cells. A similar result was also previously observed for all the other HAC.⁹ All the HAC bound centromeric and kinetochore proteins in an amount similar to each other and to the endogenous chromosomes, unlike Spence *et al*,¹⁹ who reported that HAC of different sizes bound different amounts of the Aurora B. Analysis of the HAC 1-H by PFGE indicated that it was circular,⁸ yet the HAC cohesin complex characterised by the binding of SMC1, SMC3, Scc1 and Scc3 proteins remained intact throughout the different cell-cycle stages.^{18,20} The two chromatids are linked together for the whole chromosome length, but at metaphase, they are joined only at the centromere region. As no difference was observed between the HAC and the endogenous chromosomes, we concluded that the circular structure of the HAC did not interfere with the cohesin complex formation and anaphase chromatid resolution.

The characterisation of HAC chromatin demonstrated that both euchromatic and heterochromatic-specific histone H3 modifications were present on each of the HAC. We observed that heterochromatin flanked this region on all HAC, including the smallest HAC 3-H, unlike Grimes *et al*,¹⁷ who found that only the larger HAC

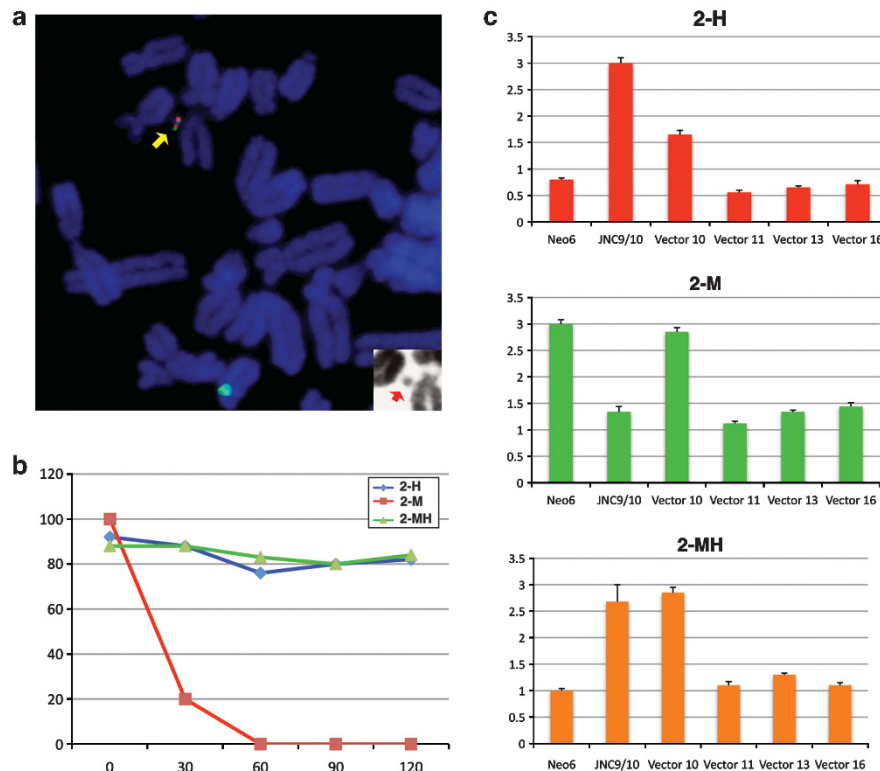


Figure 4 Analysis of 2-MH HAC. (a) FISH on metaphase spreads with 17 α DNA (green signal) and vector probe (red signal). The HAC (yellow arrow) is labelled by both. Chromosomes are counterstained by DAPI (4,6-diamidino-2-phenylindole; blue). The inset shows DNA staining only, with the HAC identified by the red arrow. (b) 2-MH HAC loss rate over a period of 120 days in the absence of selection compared with its murine parental HAC 2-M and human parental HAC 2-H. Clone 2-MH (green line) replicates the pattern of clone 2-H (blue line). (c) Quantitative PCR analysis of ChIP_{Cenp A} DNA, showing the relative fold-enrichment in the ChIP_{Cenp A} versus ChIP_{mock}. Clone 2-MH follows the same pattern as clone 2-H.

contained heterochromatin and the small HAC <3 Mb were euchromatin-rich only.

In human cells, the HAC centromere was localised both on aliphoid and non-aliphoid DNA, albeit at different degrees in the different HAC, but this did not seem to affect HAC stability. In most cases, the centromere localisation of the murine derivatives differed from their human counterparts. Furthermore, when different HAC deriving from the same MMCT experiments were characterised, they appeared to differ from one another in organisation. Based on the ChIP data, we could not find a definite correlation between low stability of HAC in murine cells and amount of non-alpha DNA-forming part of the centromere. However, it is possible that this is due to the small sample size we analysed.

It is possible that the differences observed in centromere positioning between human HAC and murine derivatives were due to rearrangements, which occurred during the MMCT. Indeed, we have showed previously that the HAC 2-M had a different structure from HAC 2-H, as determined by Southern blotting.¹¹ However, in a similar analysis of HAC 1-H and its murine derivatives HAC 1-M1, 1-M2, 1-M3 and 1-M4, no such difference in structure was identified.⁹ An alternative explanation for the different positioning of the centromere in human HAC compared with the murine derivatives could be that the epigenetic inheritance of centromere identity was inhibited due to the disruption of mitosis during MMCT. The MMCT procedure is based on the synchronization of cells with high doses of colchicine, which prevents the cells from progressing further from metaphase. Following centrifugation to form microcells, they are fused to an unsynchronized population of receiving cells. When the HAC were transferred to the murine cells, the centromere epigenetic marking may have been lost, because the chromosomes contained in the microcells did not exit from mitosis. After the fusion with receiving cells, this could have lead to the formation of a centromeric domain in a new position on the HAC DNA. In accordance with this hypothesis, the centromere of murine HAC 2-M changed position, following its transfer back into a human background, to generate HAC 2-MH. In cell lines where the centromeric region markings were not recovered, the lack of a definite centromere could have lead to the HAC DNA becoming integrated into the host chromosome as in HAC 3-M. If the choice between the two fates is random, the larger the HAC size the higher the chance that a sequence could be recruited to form a centromere. This may explain why the smallest human HAC, 3-H, showed the highest incidence of integrated DNA into the murine host genome following transfer by MMCT. This finding is also supported by recent work from Jansen *et al*,²¹ who demonstrated that loading of CENP-A on newly replicated DNA occurs in G1 phase, and requires passage through, and exit from mitosis. In all the HAC we analysed, the aliphoid DNA was always part of the centromere, while the amount of non-aliphoid DNA associated with CENP A was variable between different clones. This suggested that the aliphoid DNA is a preferential sequence for centromere formation, either because it retains its centromere markings more efficiently, or because it is targeted by the centromere assembly system, even in the murine background where no aliphoid DNA is normally present.

Finally, we showed that HAC displaying high level of instability in murine cells are fully stable once transferred back into human background. This data suggest that the cause of the instability is not a structural rearrangement that inactivated the centromere. It is likely that to achieve full stability and centromere formation in murine cells, artificial chromosomes require the presence of murine specific centromeric sequences (minor satellite), as reported

previously,²² to produce a mammalian artificial chromosome.²³ It would also be interesting, however, to investigate how to stabilise human chromosome in murine cells, possibly using an approach recently developed that allows targeted epigenetic modifications to be introduced into the HAC centromere,^{6,24} Alternatively, the effect of fusing microcells to cells synchronized at a specific cell-cycle stage could also be investigated, to understand whether the differential stability displayed by HAC in murine cells is linked to the loss of centromere markings.

In conclusion, our data demonstrate that the HAC in human cells are generally organised and behave like the endogenous host chromosomes, and there is no difference in chromatin organisation and binding of centromeric proteins between the HAC generated in human or murine cells. However, the data indicated that the HAC centromere region may have changed in size and region following transfer to murine cells. Overall, this work demonstrated that the HAC are a major resource tool for investigating different areas of cell biology and comparative aspects of chromosome epigenetics, owing to their small size and low complexity.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Andy Choo and Paul Kalitsis for the kind gift of the anti-mouse CENP A antibody. This work was supported by the Wellcome Trust grant [090532/Z/09/Z].

- 1 Harrington JJ, Van Bokkelen G, Mays RW *et al*: Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. *Nat Genet* 1997; **15**: 345–355.
- 2 Grimes BR, Larin Monaco Z: Artificial and engineered chromosomes: developments and prospects for gene therapy. *Chromosoma* 2005; **114**: 230–241.
- 3 Okada T, Ohzeki J, Nakano M *et al*: CENP-B controls centromere formation depending on the chromatin context. *Cell* 2007; **131**: 1287–1300.
- 4 Lam AL, Boivin CD, Bonney CF *et al*: Human centromeric chromatin is a dynamic chromosomal domain that can spread over noncentromeric DNA. *Proc Natl Acad Sci USA* 2006; **103**: 4186–4191.
- 5 Nakano M, Cardinale S, Noskov VN *et al*: Inactivation of a human kinetochore by specific targeting of chromatin modifiers. *Dev Cell* 2008; **14**: 507–2.
- 6 Bergmann JH, Martins NM, Larionov V *et al*: HACKing the centromere chromatin code: insights from human artificial chromosomes. *Chromosome Res* 2012; **20**: 505–519.
- 7 Ohzeki J, Bergmann JH, Kouprina N *et al*: Breaking the HAC Barrier: histone H3K9 acetyl/methyl balance regulates CENP-A assembly. *EMBO J* 2012; **31**: 2391–2402.
- 8 Mejía JE, Willmott A, Levy E *et al*: Functional complementation of a genetic deficiency with human artificial chromosomes. *Am J Hum Genet* 2001; **69**: 315–326.
- 9 Moralli D, Chan DY, Jefferson A *et al*: HAC stability in murine cells is influenced by nuclear localization and chromatin organization. *BMC Cell Biol* 2009; **10**: 18.
- 10 Mejía JE, Alazami A, Willmott A *et al*: Efficiency of de novo centromere formation in human artificial chromosomes. *Genomics* 2002; **79**: 297–304.
- 11 Alazami AM, Mejía JE, Monaco ZL: Human artificial chromosomes containing chromosome 17 aliphoid DNA maintain an active centromere in murine cells but are not stable. *Genomics* 2004; **83**: 844–851.
- 12 Moralli D, Simpson KM, Wade-Martins R *et al*: A novel human artificial chromosome gene expression system using herpes simplex virus type 1 vectors. *EMBO Rep* 2006; **7**: 911–918.
- 13 Kalitsis P, Fowler K, Earle E *et al*: Targeted disruption of mouse centromere protein C gene leads to mitotic disarray and early embryo death. *Proc Natl Acad Sci USA* 1998; **95**: 1136–1141.
- 14 Sullivan BA: Optical mapping of protein-DNA complexes on chromatin fibers; in Bridger JM, Volpi EVW (eds) *Methods in Molecular Biology*. New York, 2010 vol **659**: pp 99–115.
- 15 Wade-Martins R, Smith ER, Tyminski E *et al*: An infectious transfer and expression system for genomic DNA loci in human and mouse cells. *Nat Biotechnol* 2001; **19**: 1067–1070.
- 16 Sullivan BA, Karpen GH: Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat Struct Mol Biol* 2004; **11**: 1076–1083.

- 17 Grimes BR, Babcock J, Rudd MK *et al*: Assembly and characterization of heterochromatin and euchromatin on human artificial chromosomes. *Genome Biol* 2004; **5**: 14911–14918.
- 18 Goto H, Yasui Y, Nigg EA *et al*: Aurora-B phosphorylates histone H3 at serine28 with regard to the mitotic chromosome condensation. *Genes Cells*, **7**: 11–17.
- 19 Spence JM, Mills W, Mann K *et al*: Increased missegregation and chromosome loss with decreasing chromosome size in vertebrate cells. *Chromosoma* 2006; **115**: 60–74.
- 20 Hirano T: At the heart of the chromosome: SMC proteins in action. *Nat Rev Mol Cell Biol* 2006; **7**: 311–312.
- 21 Jansen LE, Black BE, Foltz DR *et al*: Propagation of centromeric chromatin requires exit from mitosis. *J Cell Biol* 2007; **176**: 795–805.
- 22 Shen MH, Mee PJ, Nichols J *et al*: A structurally defined mini-chromosome vector for the mouse germ line. *Curr Biol* 2000; **10**: 31–33.
- 23 Farr CJ, Spence JM: Mammalian artificial chromosomes (MACs). *eLS* 2006; 10.1038/npg.els.0005671.
- 24 Ohzeki J, Bergmann JH, Kouprina N *et al*: Breaking the HAC Barrier: histone H3K9 acetyl/methyl balance regulates CENP-A assembly. *EMBO J* 2012; **31**: 2391–2402.

Supplementary Information accompanies this paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)