Gene Therapy (2002) 9, 713–718
© 2002 Nature Publishing Group All rights reserved 0969-7128/02 $25.00
www.nature.com/gt

Chromosome engineering: prospects for gene therapy

BR Grimes[1], PE Warburton[2] and CJ Farr[3]
[1]*Department of Genetics, School of Medicine, Case Western Reserve University and University Hospitals of Cleveland, Cleveland, OH, USA;* [2]*Mount Sinai School of Medicine, New York, NY, USA; and* [3]*Department of Genetics, University of Cambridge, Cambridge, UK*

Recent advances in chromosome engineering and the potential for downstream applications in gene therapy were presented at the Artificial Chromosome Session of Genome Medicine: Gene Therapy for the Millennium in Rome, Italy in September 2001. This session concentrated primarily on the structure and function of human centromeres and the ongoing challenge of equipping human artificial chromosomes (HACs) with centromeres to ensure their mitotic stability. Advances in the 'bottom up' construction of HACs included the transfer into HT1080 cells of circular PACs containing alpha satellite DNA, and the correction of HPRT deficiency in cells using HACs. Advances in the 'top down' construction of HACs using telomere associated chromosome fragmentation in DT40 cells included the formation of HACs that are less than a megabase in size and transfer of HACs through the mouse germline. Significant progress has also been made in the use of human minichromosomes for stable trans-gene expression. While many obstacles remain towards the use of HACs for gene therapy, this session provided an optimistic outlook for future success.
Gene Therapy (2002) **9,** 713–718. DOI: 10.1038/sj/gt/3301763

Keywords: *engineered chromosomes; gene therapy vector*

Introduction

The Artificial Chromosome Session provided an illuminating update on recent advances in chromosome engineering technology and potential downstream applications in the area of gene therapy. A significant advantage of using engineered chromosomes to deliver genes into cells is that they can replicate and segregate independently of host chromosomes, thereby avoiding insertional mutagenesis problems that may arise with conventional vectors. Additionally, since fragments in the megabase size range can be introduced into an engineered chromosome, large genomic DNA fragments spanning genes and flanking regulatory regions could be accommodated which should permit, in principle, stable transgene expression.

An important consideration for the future development of engineered chromosomes is the need for a better understanding of normal chromosome function. There are three elements considered essential for linear chromosome transmission, namely telomeres (which cap and protect linear chromosome ends), replication origins (for DNA duplication) and a centromere that permits chromosome segregation. Telomeric DNA has been structurally defined and less than a kilobase (kb) of telomeric DNA can promote telomere formation in mammalian cells.[1,2] The extent to which origin activity is dependent on specific DNA elements is still unclear. However, origins occur frequently in the genome,[3] and our ignorance of origin requirements has not so far posed a problem in the development of artificial chromosomes. Critical, however, is the nature of the chromosomal region underlying the centromere/kinetochore structure and here also data remain somewhat equivocal. The relationship between alpha satellite DNA (the human centromere associated sequence),[4] and centromere function is underscored by its ability in a cell transfection assay to seed a *de novo* functional centromere,[5,6] under conditions where a random DNA sequence is ineffective.[7] However, other studies raise questions over the role of DNA sequence in centromere identity and implicate involvement of epigenetic factors (discussed below).

The approaches taken for the creation of engineered human chromosomes fall into two broad categories, depending on whether the starting material is naked DNA containing sequences capable of generating a *de novo* chromosome following human cell transfection ('bottom up') or a pre-existing chromosome ('top down'). Chromosomes formed from naked DNA are termed 'artificial chromosomes' (Figure 1). The first human artificial chromosomes (HACs) to be described were constructed from unlinked components comprising synthetic alpha satellite, telomere DNA and random genomic DNA fragments.[5] Transfection of this mixture into cultured human HT1080 cells resulted in *de novo* HAC formation in a proportion of clonal lines analyzed using fluorescence *in situ* hybridization (FISH). Following this study, several reports showed that pre-assembled YAC- or PAC-based vectors containing alpha satellite DNA could also form HACs (Figure 1).[6–10] In each case, HACs have been shown to be low in copy number, mitotically stable and to have recruited functionally important centromere proteins. Moreover, they appear to consist of multiple copies

Correspondence: BR Grimes, Department of Genetics, School of Medicine, Case Western Reserve University and University Hospitals of Cleveland, Adelbert Road, Cleveland, OH 44106-4955, USA

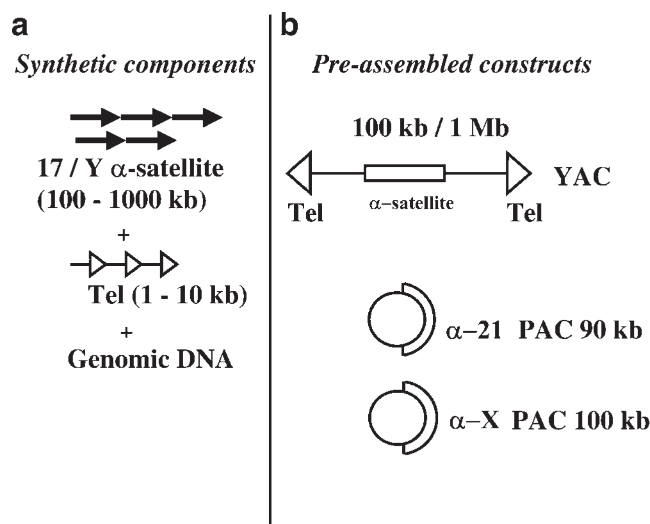

Figure 1 *Systems for creating de novo human artificial chromosomes ('bottom up'). Transfection of DNA containing alpha satellite DNA as a centromere seeding component can result in de novo centromere formation on an artificial chromosome. Clones resistant to the selectable marker gene present on the input DNA are harvested and analyzed cytogenetically to identify lines where a HAC has formed. HAC substrates can be either a transfection mixture containing synthetic components (a) or a pre-assembled vector (b). The first pre-assembled vector shown to form a HAC was a 100 kb telomere capped YAC containing cloned alpha satellite DNA from human chromosome 21.[6] Circular input forms containing alpha satellite DNA either from chromosome 21 or the X chromosome have also been shown to be HAC substrates.*

of the transfected DNA, assembled by the cell to yield chromosomes in the 1–10 Mb size range.

In the 'top down' approach, telomere-associated chromosome fragmentation is used to produce minimal human chromosomes (referred to as 'minichromosomes') which have lost most or all of their chromosome arm material. The existence of spontaneously arising or radiation-induced human marker chromosomes offers an alternative starting point for deriving a chromosome-based vector system.[11–17] A further chromosome engineering technology under development relies upon integration of exogenous DNA into the centromere region of a host mouse or human chromosome. Amplification of DNA at the integration site can result in chromosome breakage events, producing new large chromosomes typically in the 60–400 Mb size range.[18,19] Due to their size and high satellite DNA content, these can be purified in quantities sufficient for microinjection into mouse pronuclei,[20] or lipofection into a range of cell types.[21]

The following sections will review our current understanding of the centromere and the state of progress and possible future directions in chromosome engineering technology.

Epigenetics of human centromere formation

In *S. cerevesiae* introduction of a plasmid containing the 125 bp CEN DNA leads reproducibly to its mitotic stability. However, in mammals the rapid evolution and lack of sequence conservation of the large arrays of tandemly repeated DNA found at centromeres across species suggests that an equivalent mammalian CEN DNA may not exist.[22] Nevertheless, alpha satellite DNA is found at all normal human centromeres and is capable of *de novo* centromere formation in human cells.

The analysis of human variant centromeres also suggests that primary DNA sequence may not be a strict determinant of centromere function. There are examples of both naturally occurring and experimentally generated arrays of human alpha satellite DNA that do not form centromeres. Mitotically stable dicentric chromosomes resulting from chromosomal fusion events (eg human Robertsonian translocations) are observed, where one centromere has been inactivated and is no longer forming a kinetochore, yet contains apparently normal arrays of alpha satellite DNA. Furthermore, one of the fates of alpha satellite DNA when introduced into cells in 'bottom up' approaches is integration into chromosome arms, often in an amplified manner, in a non-centromeric configuration. Both of these types of ectopic alpha satellite DNA cell lines have lent considerable insight into the role of alpha satellite DNA in centromere formation and requirements for additional proteins and chromatin modifications associated with active centromeres.[23–25] One exciting avenue of research involves attempts to manipulate and reactivate such regions of nonfunctional alpha satellite DNA.

An additional type of variant centromere, the neocentromere, is found on small rearranged chromosome fragments that have separated from endogenous centromeres and do not contain alpha satellite DNA. Nevertheless, neocentromeres form fully functional kinetochores and provide mitotic stability to chromosome fragments that would normally be acentric and rapidly lost. The existence of neocentromeres shows that the presence of alpha satellite DNA is not an absolute requirement for human centromere formation. Crucially these unusual centromeres provide an opportunity to examine centromere structure in the absence of the vast repetitive DNA arrays normally present. Over 50 examples of clinically obtained neocentromeres have been described thus far, several of which have been localized to regions as small as several hundred kb of genomic DNA.[26,27]

The two types of centromere variants described here suggest that human centromere formation is epigenetically determined.[28] Thus, from an epigenetic perspective, the inactive arrays of alpha satellite DNA in dicentric chromosomes have lost the centromere-determining epigenetic mark, while neocentromeres have gained the mark in a previously noncentromeric location. In this light, the top down approach for engineered chromosome formation will permit retention of this epigenetic mark, as the procedure is performed *in vivo* in tissue culture cells and never goes through a naked DNA stage where the mark would be erased. The relatively low efficiency of *de novo* centromere formation using bottom up approaches may be explained by the requirement of the introduced DNA constructs to obtain this epigenetic mark before they become integrated or degraded. Nonetheless, *de novo* centromere formation using human alpha satellite DNA, as well as chromosomal regions with a relatively high incidence of observed neocentromeres,[26,27] suggests that particular DNA sequences may have an increased propensity for centromere formation over the bulk of genomic DNA. This may be attributed to an increased affinity for CENP-A, the centromere-specific histone H3 variant, and formation of a specialized centromeric chromatin structure. Current models of epigen-

etic centromere determination propose that once this CENP-A containing chromatin is in place, it can be propagated through repeated cell divisions by a self-templating mechanism that would not be strictly dependent on underlying DNA sequence.[29,30] Future strategies for artificial chromosome formation may include attempts to initially provide for this epigenetic mark, such as introduction into cells of DNA partially reconstituted with CENP-A-containing nucleosomes.

De novo artificial chromosomes and gene expression

Since the publication of the first two landmark papers describing the formation of HACs from either unlinked synthetic components,[5] or a pre-assembled YAC containing cloned alpha satellite DNA,[6] there has been progress in the development of new PAC- and BAC-based systems for constructing *de novo* HACs.[7,10,31,32] The recent demonstration that circular HACs can be formed from circular PAC DNA lacking telomere repeats and containing 70 kb of alpha satellite DNA, represents a step forward in determining the minimal requirements for generating a mitotically stable HAC.[7] Previous vectors included telomere repeats, the aim being to generate a linear HAC reflecting host chromosome structure. However, obtaining telomere function on *de novo* HACs has not been as straightforward as anticipated and only a few HACs have proven to be linear in structure,[5] whilst others remain uncharacterized in this respect. Nevertheless, results from FISH analysis of human cell lines containing circular HACs showed that they can be highly mitotically stable, suggesting that circular HACs may be a viable alternative. An added advantage is the relative ease of preparation of microgram quantities of PAC and BAC DNA compared with YAC-based vectors.

The simplicity of the circular input assay system should allow optimization of the centromere component in *de novo* HAC vectors. At the current meeting, results were reported from the Willard, Masumoto and Larin laboratories, using circular input DNA which demonstrated some alpha satellite DNA sequences can form HACs with much higher efficiency than others. A critical feature of the experiments carried out by the Willard and Masumoto groups was that synthetic alpha satellite arrays were used in the input DNA sequence, offering the significant advantage that the input DNA is completely defined, an important parameter in the design of any future gene therapy vector.

In parallel with the development of improved HAC vectors, two groups have now reported the first proof-of-principle experiments to show that a large *HPRT* human gene can be incorporated into and expressed from a HAC in human *HPRT*-deficient HT1080 cells. In one approach, a BAC-based vector was constructed that was linearized to generate a 404 kb telomere capped construct containing 220 kb of alpha satellite DNA and a 160 kb fragment containing a human *HPRT* gene (Figure 2a).[32] In the second approach, a co-transfection strategy was used combining a PAC containing 70 kb of alpha satellite DNA with a second (160 kb) genomic PAC spanning a human *HPRT* gene (Figure 2b).[33] Lines were developed containing an *HPRT*-expressing HAC as the sole fate of the transfected DNA (Figure 2c). Furthermore, Northern analysis demonstrated that *HPRT* transgene-derived

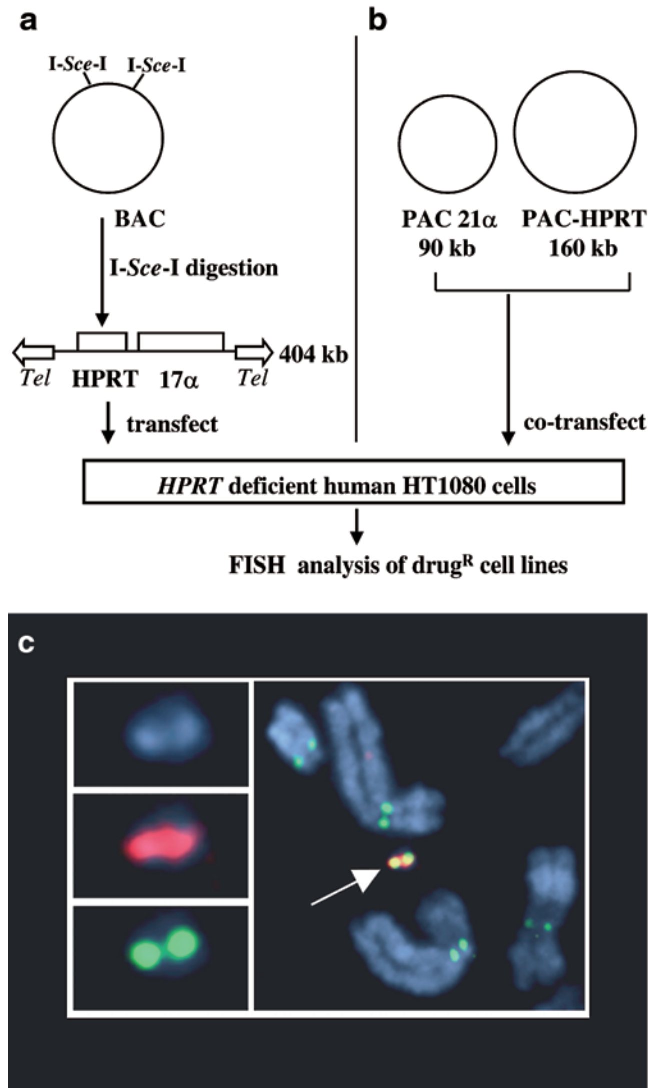

Figure 2 Gene expression on de novo HACs. Two methods were used to incorporate a large genomic DNA fragment containing a human HPRT gene into a HAC. Either (a) a pre-formed BAC construct or (b) a co-transfection mixture containing two PACs were transfected into HPRT deficient HT1080 cells. Drug-resistant clones were analyzed using FISH for the presence of a HAC. (c) Human metaphase spread containing an HPRT -expressing HAC (arrow) formed by the co-transfection strategy outlined in (b). The functionally important centromere protein, CENP-A, is present at host centromeres and on the HAC (green double dot signals; HAC indicated by arrow). An HPRT probe (red) was used to identify the HAC using FISH. Chromosomes were counterstained with DAPI (blue). Overlapping signals are yellow. Merged image (main panel). Inserts show separated HAC images: DAPI (top), HPRT /DAPI (middle) and CENP-A/DAPI (bottom).

HPRT mRNA levels remained constant after 60 days under non-selective conditions.[33] Correction of the *HPRT* deficiency was demonstrated by growth on HAT selection in each study. The Larin and Schindelhauer laboratories reported that efforts are now underway to introduce a range of human genes into HACs.

The generation of minimal chromosomes by telomere associated chromosome fragmentation

The general strategy has been to manipulate a non-essential chromosome, retained in a somatic cell hybrid, by

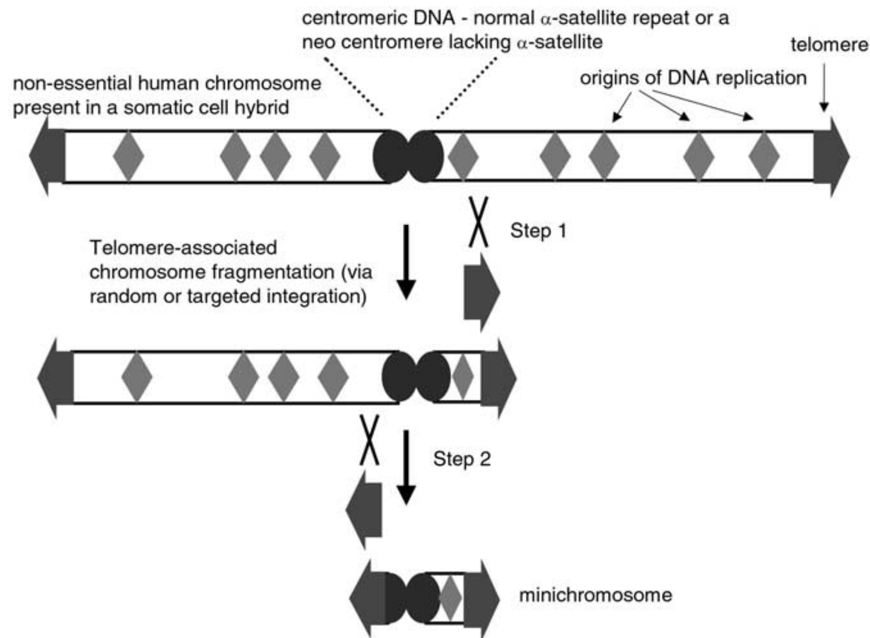

Figure 3 The generation of α-satellite and neocentromere-based minichromosomes by telomere-associated chromosome truncation ('top down'). This approach has been used to generate minichromosomes based on both the human X and Y centromere regions, as well as from a mardel(10) marker chromosome, which contains a neocentromere. The (TTAGGG)n-containing constructs used to fragment the chromosome, with accompanying removal of the bulk of chromosome arm sequences, have been introduced both through random integration and by homologous recombination events.

using cloned telomeric DNA to induce chromosome breakage, and allow the recovery of stable truncated chromosomes and minichromosomes (Figure 3). Early attempts relied on random integration. More recently, human chromosomes have been transferred into the recombination-proficient chicken cell line DT40, making it practical to target telomere seeding events and generate products of defined structure.[34–36]

Successive rounds of telomere seeding and chromosome truncation have been used to generate a series of human Y-derived linear minichromosomes ranging in size from approximately 9 Mb down to 1.7 Mb.[35,37,38] While stable in human HT1080 and chicken DT40 cells, when introduced into mouse cells (ES and/or LA-9) various human Y-based minichromosomes were found to be unstable.[37,39] One of the minichromosomes rearranged and was stabilized through the acquisition of mouse centromeric DNA. This 4.5 Mb minichromosome was found to be stably maintained in mouse somatic cells and could be transmitted through the germline. Shen and colleagues[39] reported that this minichromosome is stably transmitted in three different mouse strains.

A similar telomere seeding and chromosome fragmentation strategy has been used by Mills and colleagues to generate a series of human X centromere-based minichromosomes (size range approximately 3 Mb–<1 Mb),[34] while Saffery *et al* have generated a series of minichromosomes based on the neocentromere of a marker chromosome derived from human chromosome 10.[40] Neocentromere-based minichromosomes (size range: 0.7 Mb–1.8 Mb) offer potential advantages over those retaining normal repetitive centromere DNA: they may provide a more permissive chromatin environment for full and appropriate gene expression, while their single-copy DNA make-up means that they can be fully characterized at the molecular level.

These well-defined, linear minichromosomes are approaching a size and structure where their use as first generation chromosome vectors is being considered. One route for retrofitting such structures with DNA of interest is through Cre/*lox*P-mediated site-specific recombination.[41] An elegant application of this technology has recently been described by Kuroiwa and colleagues for the generation of an engineered chromosome carrying the immunoglobulin lambda-light chain locus (Figure 4).[16] Many of the tools are therefore in place, although questions over species differences in centromere function (and possibly in other areas of chromosome biology) remain to be fully investigated.

Engineering human marker chromosomes and mouse models

Use of the Cre/*lox*P site-specific recombination technology for introducing genes into marker chromosomes was reported by the Raimondi and Marynen laboratories. Of particular relevance for potential gene therapy applications, Ascenzioni and colleagues introduced a human *CFTR* gene contained within a 310 kb YAC (using a co-transfection strategy), into a structurally characterized 5.5 Mb human minichromosome derived from chromosome 1 by gamma-irradiation,[13] and demonstrated stable transgene expression in cell lines.

Recent advances in chromosome transfer technologies have permitted the introduction of human marker chromosomes into mice that can be transmitted through the mouse germline.[14,15,17] The production of mice harboring a human chromosome(s) showing correct tissue-specific human immunoglobulin gene expression provides the first mouse model addressing the capacity of a human marker chromosome to deliver tissue specific transgene expression. The Ishida and Marynen laboratories have

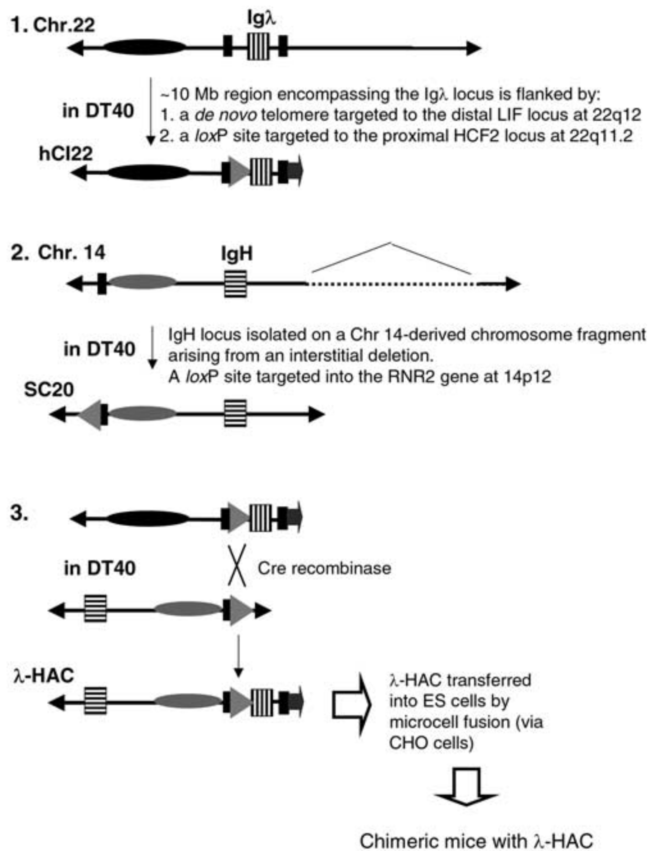

Figure 4 *Strategy for generating a stable chromosome vector carrying the human IgH and Igλ loci (after Kuroiwa et al[16]). A combination of retrofitting by homologous recombination and targeted de novo telomere formation in the recombination-proficient chicken cell line DT40, together with site-specific recombination, was used to clone two large genomic regions encompassing Ig loci on to a single, mitotically and structurally stable, minichromosome. This chromosome-based vector was then transferred into mice via ES cells and cell fusion techniques.*

now taken this technology one step further by modifying human marker chromosomes using the Cre/*lox*P system to introduce genes that were stably expressed following transfer of the modified marker chromosome into mice.[16,17]

Concluding remarks and future perspectives

For gene therapy applications, a chromosome-based vector should have a defined molecular structure, be mitotically and structurally stable, and capable of introducing a transgene(s) with reliable and appropriate expression patterns. While substantial progress towards the production of a chromosome vector has been made in the last 4 years, an optimal structure has yet to be fully defined and major breakthroughs in the technology of HAC manufacture and delivery are essential before application in the gene therapy arena can be realized.[42,43] Nevertheless, chromosome-based vectors for mammalian cells are now a reality and the results presented in the Artificial Chromosome Session warrant an optimistic outlook.

Acknowledgements

We would like to thank meeting participants Hunt Willard, Bala Balakumaran, Katie Rudd and Roger Slee for valuable discussions.

References

1 Farr C *et al*. Functional reintroduction of human telomeres into mammalian cells. *Proc Natl Acad Sci USA* 1991; **88**: 7006–7010.
2 Hanish JP, Yanowitz JL, de Lange T. Stringent sequence requirements for the formation of human telomeres. *Proc Natl Acad Sci USA* 1994; **91**: 8861–8865.
3 Gilbert DM. Making sense of eukaryotic DNA replication origins. *Science* 2001; **294**: 96–100.
4 Willard HF. Centromeres: the missing link in the development of human artificial chromosomes. *Curr Opin Genet Dev* 1998; **8**: 219–225.
5 Harrington JJ *et al*. Formation of *de novo* centromeres and construction of first-generation human artificial microchromosomes. *Nat Genet* 1997; **15**: 345–355.
6 Ikeno M *et al*. Construction of YAC-based mammalian artificial chromosomes. *Nat Biotechnol* 1998; **16**: 431–439.
7 Ebersole TA *et al*. Mammalian artificial chromosome formation from circular alphoid input DNA does not require telomere repeats. *Hum Mol Genet* 2000; **9**: 1623–1631.
8 Masumoto H *et al*. Assay of centromere function using a human artificial chromosome. *Chromosoma* 1998; **107**: 406–416.
9 Henning KA *et al*. Human artificial chromosomes generated by modification of a yeast artificial chromosome containing both human alpha satellite and single- copy DNA sequences. *Proc Natl Acad Sci USA* 1999; **96**: 592–597.
10 Schueler MG *et al*. Genomic and genetic definition of a functional human centromere. *Science* 2001; **294**: 109–115.
11 Raimondi E *et al*. Gene targeting to the centromeric DNA of a human minichromosome. *Hum Gene Ther* 1996; **7**: 1103–1109.
12 Guiducci C *et al*. Use of a human minichromosome as a cloning and expression vector for mammalian cells. *Hum Mol Genet* 1999; **8**: 1417–1424.
13 Auriche C, Donini P, Ascenzioni F. Molecular and cytological analysis of a 5.5 Mb minichromosome. *EMBO Rep* 2001; **2**: 102–107.
14 Tomizuka K *et al*. Functional expression and germline transmission of a human chromosome fragment in chimaeric mice. *Nat Genet* 1997; **16**: 133–143.
15 Tomizuka K *et al*. Double trans-chromosomic mice: maintenance of two individual human chromosome fragments containing Ig heavy and kappa loci and expression of fully human antibodies. *Proc Natl Acad Sci USA* 2000; **97**: 722–727.
16 Kuroiwa Y *et al*. Manipulation of human minichromosomes to carry greater than megabase-sized chromosome inserts. *Nat Biotechnol* 2000; **18**: 1086–1090.
17 Voet T *et al*. Efficient male and female germline transmission of a human chromosomal vector in mice. *Genome Res* 2001; **11**: 124–136.
18 Kereso J *et al*. *De novo* chromosome formations by large-scale amplification of the centromeric region of mouse chromosomes. *Chromosome Res* 1996; **4**: 226–239.
19 Csonka E *et al*. Novel generation of human satellite DNA-based artificial chromosomes in mammalian cells. *J Cell Sci* 2000; **113**: 3207–3216.
20 Co DO *et al*. Generation of transgenic mice and germline transmission of a mammalian artificial chromosome introduced into embryos by pronuclear microinjection. *Chromosome Res* 2000; **8**: 183–191.
21 de Jong G *et al*. Efficient *in vitro* transfer of a 60-Mb mammalian artificial chromosome into murine and hamster cells using cationic lipids and dendrimers. *Chromosome Res* 2001; **9**: 475–485.
22 Henikoff S, Ahmad K, Malik HS. The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* 2001; **293**: 1098–1102.

23 Larin Z, Fricker MD, Tyler-Smith C. *De novo* centromere formation following introduction of a Y alphoid YAC into mammalian cells. *Hum Mol Genet* 1994; **3**: 689–695.
24 Warburton PE, Cooke HJ. Hamster chromosomes containing amplified human alpha-satellite DNA show delayed sister chromatid separation in the absence of *de novo* kinetochore formation. *Chromosoma* 1997; **106**: 149–159.
25 Sullivan BA, Schwartz S. Identification of centromeric antigens in dicentric Robertsonian translocations: CENP-C and CENP-E are necessary components of functional centromeres. *Hum Mol Genet* 1995; **5**: 2189–2198.
26 Warburton PE *et al*. Molecular cytogenetic analysis of eight inversion duplications of human chromosome 13q that each contain a neocentromere. *Am J Hum Genet* 2000; **66**: 1794–1806.
27 Choo KHA. Domain rganization at the centromere and neocentromere. *Dev Cell* 2001; **1**: 165–177.
28 Warburton PE. Epigenetic analysis of kinetochore assembly on variant human centromeres. *Trends Genet* 2001; **17**: 243–247.
29 Sullivan BA, Blower MD, Karpen GH. Determining centromere identity: cyclical stories and forking paths. *Nat Rev Genet* 2001; **2**: 584–596.
30 Sullivan KF. A solid foundation: functional specialization of centromeric chromatin. *Curr Opin Genet Dev* 2001; **11**: 182–188.
31 Mejia JE, Larin Z. The assembly of large BACs by in vivo recombination. *Genomics* 2000; **70**: 165–170.
32 Mejia JE *et al*. Functional complementation of a genetic deficiency with human artificial chromosomes. *Am J Hum Genet* 2001; **69**: 315–326.
33 Grimes BR *et al*. Stable gene expression from a mammalian artificial chromosome. *EMBO Rep* 2001; **2**: 910–914.
34 Mills W *et al*. Generation of an approximately 2.4 Mb human X centromere-based minichromosome by targeted telomere-associated chromosome fragmentation in DT40. *Hum Mol Genet* 1999; **8**: 751–761.
35 Yang JW *et al*. Human mini-chromosomes with minimal centromeres. *Hum Mol Genet* 2000; **9**: 1891–1902.
36 Kuroiwa Y *et al*. Efficient modification of a human chromosome by telomere-directed truncation in high homologous recombination-proficient chicken DT40 cells. *Nucleic Acids Res* 1998; **26**: 3447–3448.
37 Shen MH *et al*. Human mini-chromosomes in mouse embryonal stem cells. *Hum Mol Genet* 1997; **6**: 1375–1382.
38 Shen MH *et al*. The accuracy of segregation of human mini-chromosomes varies in different vertebrate cell lines, correlates with the extent of centromere formation and provides evidence for a trans-acting centromere maintenance activity. *Chromosoma* 2001; **109**: 524–535.
39 Shen MH *et al*. A structurally defined mini-chromosome vector for the mouse germ line. *Curr Biol* 2000; **10**: 31–34.
40 Saffery R *et al*. Construction of neocentromere-based human minichromosomes by telomere-associated chromosomal truncation. *Proc Natl Acad Sci USA* 2001; **98**: 5705–5710.
41 Smith AJH *et al*. A site-directed chromosomal translocation induced in embryonic stem cells by Cre-*lox*P recombination. *Nat Genet* 1995; **9**: 376–385.
42 Willard HF. Genomics and gene therapy. Artificial chromosomes coming to life. *Science* 2000; **290**: 1308–1309.
43 Choo KH. Engineering human chromosomes for gene therapy studies. *Trends Mol Med* 2001; **7**: 235–237.