

Construction of mammalian artificial chromosomes: prospects for defining an optimal centromere

Dirk Schindelhauer

Summary

Two reports have shown that mammalian artificial chromosomes (MAC) can be constructed from cloned human centromere DNA and telomere repeats, proving the principle that chromosomes can form from naked DNA molecules transfected into human cells. The MACs were mitotically stable, low copy number and bound antibodies associated with active centromeres. As a step toward second-generation MACs, yeast and bacterial cloning systems will have to be adapted to achieve large MAC constructs having a centromere, two telomeres, and genomic copies of mammalian genes. Available construction techniques are discussed along with a new P1 artificial chromosome (PAC)-derived telomere vector (pTAT) that can be joined to other PACs *in vitro*, avoiding a cloning step during which large repetitive arrays often rearrange. The PAC system can be used as a route to further define the optimal DNA elements required for efficient MAC formation, to investigate the expression of genes on MACs, and possibly to develop efficient MAC-delivery protocols. *BioEssays* 1999;21:76–83. © 1999 John Wiley & Sons, Inc.

Introduction

The systematic introduction of genes into mammalian cells has been an elusive goal. Initial tries were thwarted by the lack of cloning tools for very large DNA. Subsequent cDNA-based techniques proved useful in a variety of applications in cultured cells and for some applications in animals and humans; however, their lack of control over the exact level and duration of gene expression remained a limiting factor. To permit regulated expression, genes would need to be introduced in a form reflecting genes on natural chromosomes. Only recently have the first models to truly hold promise for genomic expression systems concept been constructed, in which yeast artificial chromosomes (YAC)—cloned copies of genomic genes (including all regulatory sequences)—were

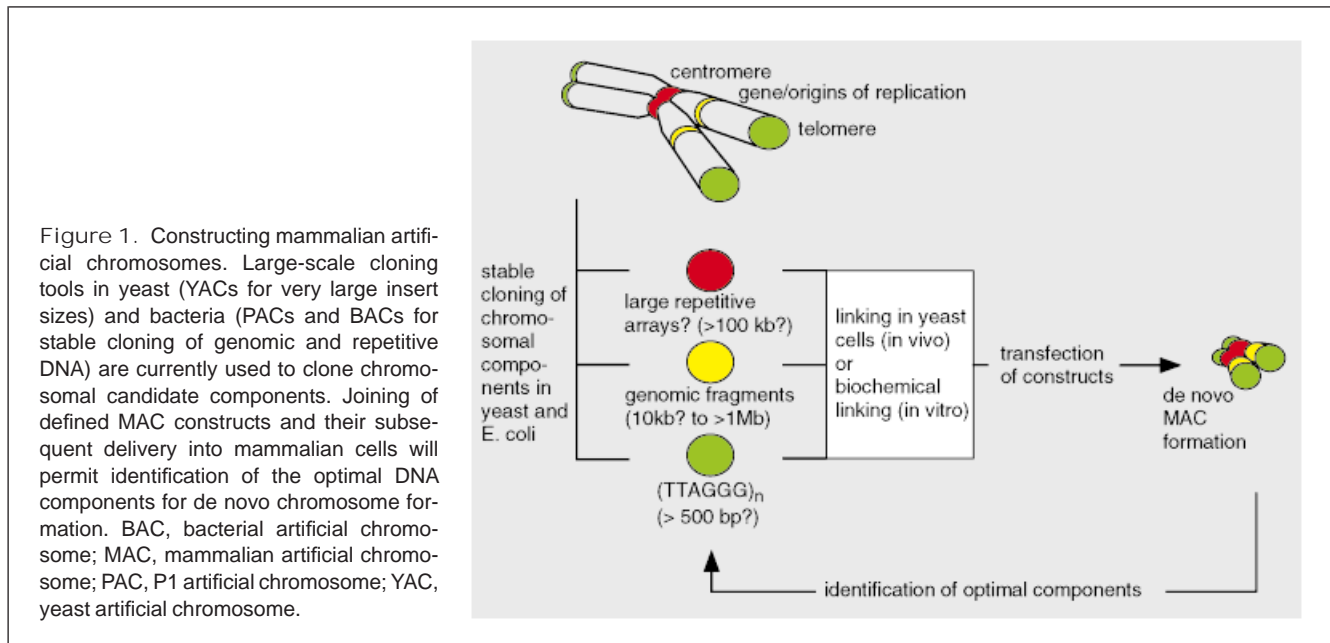
integrated at random positions into the host chromosomes of cell cultures and transgenic animals. Even so, the challenges of avoiding changes of the host genome while still maintaining stably segregating low copy number DNA elements during mitosis, but not meiosis (for germline exclusion in humans) were not met.

To fulfill these requirements, mammalian artificial chromosomes (MACs) are needed. We could do this as has been done with YACs in yeast—i.e., just put a gene, a centromere, and an origin of replication together and add telomeres to the ends. In contrast to YACs, MACs are very difficult to assemble. In mammals, not only are the telomere DNA ten- to hundredfold and the centromere DNA more than thousandfold larger, their high degree of repetition makes cloning difficult. The precise DNA requirements for the individual MAC components and ways to join large DNA to functional constructs still need further definement (Fig. 1). The research on MACs has already started to unravel the intricacies of how chromosomes work. Once technical problems have been solved, MACs will facilitate gene expression studies, will be used in cell culture and animals as biologic factories (i.e., in the udder of livestock) for the production of proteins for

Funding agencies: Medical Research Council (MRC); Deutsche Forschungsgemeinschaft (DFG).

Correspondence to: Dirk Schindelhauer, Department of Medical Genetics, Kinderpoliklinik, Ludwig Maximilians-Universitaet, Goethestrasse 29, D-80336. Muenchen, Germany;

E-mail: dirk@pedgen.med.uni-muenchen.de



therapeutic or other uses, and might allow safe and efficient long-term gene therapy. This article describes the recent progress in identifying functional centromeric DNA by engineering MACs from cloned human DNA.^(1,2) The scientific background and remaining questions are discussed along with proposed methods to overcome the difficulties in cloning, joining, and delivering large MAC molecules.

Recent successes—de novo chromosome formation

Several years ago, mammalian telomeric sequences randomly integrated into natural chromosomes were shown to seed new telomeres and result in telomere-associated chromosome fragmentation (TACF).^(3,4) While a few hundred basepairs of cloned (TTAGGG)_n tandem repeats appear sufficient for telomere function in cultured cells, the functional introduction of centromeres was missing, although some aspects of centromere function were associated with transfected satellite DNA.^(5–8) All normal centromeres of higher eukaryotes, including fission yeast (*Schizosaccharomyces pombe*), but not budding yeast (*Saccharomyces cerevisiae*), contain large regions of repetitive DNA. In humans (and other primates), alpha satellite DNA, with a basic tandem repeat unit of 170/171 bp organized in higher-order repeat structures, builds arrays with sizes of 200 kb to several Mb spanning the centromeres of all normal chromosomes.^(9,10)

Alpha satellite DNA was subsequently used by two groups to generate MACs from cloned DNA. Harrington et al.⁽¹⁾ synthetically ligated alpha satellite arrays of >100 kb in length from single characterized higher-order repeats of chromosomes 17 and Y, providing an elegant way to exclude

any additional sequences somewhere lying within the repetitive arrays as known of the essential core sequences of fission yeast centromeres⁽¹¹⁾ and to ensure high homology (identity) of the higher-order repeat units throughout the arrays. The bacterial artificial chromosome (BAC)—cloned centromere arrays combined with a selectable marker gene—were further ligated to larger sizes and cotransfected with PCR-propagated telomere repeats and total human genomic DNA, leading to randomly assembled microchromosomes. One of the four cell lines presented contained MACs that had not acquired host DNA. Ikeno et al.⁽²⁾ “retrofitted” yeast artificial chromosomes (YACs) containing alpha satellite arrays of approximately 80 kb from chromosome 21 with a selectable marker gene and telomeric repeats using a homologous recombination procedure and propagation of the construct in a recombination-deficient yeast strain. Interestingly, of two tested subtypes of chromosome 21 alpha satellite DNA, only the highly homologous array of subtype 21-I in construct “7C5,” reproducibly formed MACs that did not acquire host cell DNA. This was not the case for the diverged array 21-II, which also lacked CENP-B boxes,⁽¹²⁾ specific 17-bp binding sites for the centromere-binding protein B, present in subsets of alpha satellite DNA (except on the Y chromosome) of great apes, but not other primates.^(13–16) The repeated formation of MACs using this approach suggested that linking of specific chromosomal elements before transfection enabled MAC formation. Both groups showed that the de novo centromeres bound centromere-binding proteins (CENPs),⁽¹⁷⁾ including CENP-C and -E, which are associated with active centromeres.^(18,19)

Despite proving that transfected naked DNA acts as a template to form chromatin of functional chromosomes and showing that alpha satellite DNA is capable of forming de novo centromeres ensuring stable mitotic segregation without selection, there are still limitations in understanding chromosome function. No specific DNA elements acting as probable origins of replication⁽²⁰⁾ have yet been included. Moreover, the YAC-based approach did not even include “normal” (noncentromeric) genomic DNA, believed to statistically provide origin of replication function when exceeding several 10 kb. Nevertheless, once the mini- and microchromosomes were formed they replicated faithfully and retained a low copy number (1–3 per cell) during long-term culture. Telomeric sequences were consistently observed in the MACs by fluorescent in situ hybridization (FISH) analysis, although the small size of the MACs did not allow a clear separation of the two telomere signals.

The minimal size of a MAC construct

The minimal sequence elements for second-generation constructs necessary to achieve MACs containing one centromere, two telomeres, and genomic copies of genes in a predictable fashion remain to be defined. The structure of first-generation MACs did not simply reflect the structure and size of the transfected DNA molecules. The 6- to 10-Mb size of in vivo ligated microchromosomes⁽¹⁾ might be substantially larger than a putative minimal arrangement of functional input DNA. The YAC-based MACs formed by concatenation and/or amplification and were estimated cytogenetically to be 1–5 Mb in size.⁽²⁾ As mapping of several Mb of repetitive DNA (!) would have been required to delineate the exact structure of the first generation MACs, few data are available. Thus, it remains unclear whether concatenation/amplification, a common phenomenon observed at transgene loci,^(21,22) is a result of unspecific processes affecting transfected DNA or reflects an essential step during MAC formation. MAC-specific reasons for the observed large sizes could be (1) the minimal size of a stable human chromosome that could lie within the Mb range; (2) an additive effect of several single alpha satellite arrays needed for efficient centromere formation; (3) incomplete telomere healing on a fraction of molecules provided, leading to ligation/recombination of free DNA ends (more likely if cells were overloaded with MAC constructs); (4) impaired expression of the selectable marker gene in close proximity to the comparably short centromere template because of centromeric chromatin spreading into noncentromeric DNA, the so-called position-effect variegation or silencing, studied on *Drosophila* and fission yeast centromeres⁽²³⁾; and (5) uncontrolled initial replication followed by ligation before a stable chromatin state can form “specific epigenetic patterns” like methylation, histone acetylation, and nuclear compartmentalization. The reproducible formation of MACs in a significant proportion of transfected cells indicates that

rare events are not required for MAC formation using the YAC-based system.

Would it be possible to construct MACs which would maintain the DNA structure of the transfected molecules? Several questions need clarification before such constructs could be realized. How much telomeric sequence is required, first, to assemble protective telomeres on free DNA ends before unwanted recombinational repair allows progress through the cell cycle,⁽²⁴⁾ and second, to compensate for the stepwise loss of DNA during replication^(25–27)? How many tandem repeats of alpha satellite DNA are needed to assemble centromeric chromatin and kinetochore structures efficiently before mitosis in the transfected cell? Are there certain alpha satellite sequences or other regions like those found at neocentromeres in the absence of specific centromere sequences^(28–31) that would form MAC centromeres more efficiently? How much flanking DNA would facilitate efficient gene expression (including expression of selectable markers) within an active chromatin domain? What would the minimal size of a stable MAC be if all components would work optimal?

Final clues about the DNA requirements for MAC constructs are expected to come from experiments using MAC constructs; however, functional chromosomal elements can also be defined by reducing natural chromosomes. TACF can reduce chromosomes down to only minimal noncentromeric DNA and small parts of the naturally occurring centromere.⁽³²⁾ The shortest alphoid array on a mitotically stable, truncated Y chromosome spanned 140 kb.⁽³³⁾ Sandwiching of a selectable marker gene within 12 kb of a functional telomere and alpha satellite DNA of a truncated chromosome did not show impeded gene function⁽³⁴⁾ arguing against an important role of silencing at human centromeres, although putative suppressive chromatin, marking the actual active centromere site, could have moved away from the marker gene within >1 Mb of remaining alpha satellite DNA. The shortest minichromosomes generated by TACF are at present 2.5–4 Mb.⁽³⁵⁾ Attempts to further shorten a natural 1.3-Mb minichromosome in *Drosophila* at approximately <600 kb resulted in reduced meiotic and/or mitotic stability with complete chromosome loss at <220–400 kb, including a 220-kb centromere region.⁽³⁶⁾ Notably, minichromosomes of <300 kb that formed neocentromeres showed efficient mitotic and meiotic transmission.⁽³⁷⁾

The right construct in the right cell

Are different cell types capable of efficiently forming de novo chromosomes? Mammalian xenogeneic transgenes are generally regulated correctly, give rise to functional proteins and are replicated.^(38,39) Mammalian (and other vertebrate) telomeres have identical repeat sequences with varying interspecies length ranging from 2 to 15 kb in humans to ≤100 kb in mice,^(40–42) and TACF constructs function in a variety of

species. Interestingly, both successful de novo MAC approaches were carried out in HT1080 cells. This telomerase-positive pseudodiploid human fibrosarcoma cell line was known to allow telomere healing in TACF experiments. Telomere length declines throughout the life span in many somatic tissues in humans, and can be used as a marker to predict replicative capacity of human cultured fibroblasts (about 10 doublings per 1 kb).⁽²⁵⁾ Whether alternative lengthening of telomeres (ALT), observed in telomerase-negative tumours and cell lines,^(43–45) could also lead to healing of MAC telomeres remains to be determined. Despite gross changes in the centromeric satellite sequences in mammals during evolution^(46,47) human centromeres function in rodent cells during mitosis, as shown by interspecific cell hybrids. However, centromeres in hybrid cells do not form via protein-free DNA intermediates. For unknown, probably complex reasons, chromosomes of certain species are preferentially lost in interspecific hybrids.^(48–50) It is unclear whether the xenogeneic de novo formation of MAC centromeres would work efficiently. For the potential use of MACs in transgenic animal technology or as somatic gene therapy vehicle, it would be critical to know whether certain MAC constructs could be used for mitotic and/or meiotic transmission in various animal (model) systems.

Cloning of DNA in bacteria, yeast, and mammalian cells

Since the biochemical construction of artificial plasmids⁽⁵¹⁾ and their subsequent introduction into *Escherichia coli* during the 1970s, specific DNA elements have been used to promote the stability of the transferred DNA. One milestone in the development of bacterial cloning systems was the use of high copy number origins of replication based on the colicogenic factor E1 replicon (ColE1), which allowed high yields of plasmid DNA. A specific partition function was not needed, due to random segregation of the plasmids during cell division, which allowed the construction of small cloning vectors. As soon as larger insert sizes (>10 kb) could be transferred using phage based cell entry pathways, it became clear that the frequency of rearrangements of the cloned DNA decreased with lower copy numbers (cosmids, fosmids, P1). The use of electroporation allowed transfer of much larger plasmids (>100 kb). Fertility factor- and P1 phage-based unit copy replicons (2–3 copies per cell) facilitated stable cloning of large foreign DNA sequences in recombination-deficient *E. coli* strains with bacterial⁽⁵²⁾ and P1⁽⁵³⁾ artificial chromosomes (BACs and PACs, respectively). Both vectors have centromere-like sequences and encode for similar sets of proteins that promote faithful partitioning.⁽⁵⁴⁾

In budding yeast, essential chromosomal elements were identified by their ability to add stability to replicating high copy number plasmids. Stable segregation of low copy artificial chromosomes was achieved when linear vectors

were constructed that contained an autonomously replicating sequence (ars), a centromere sequence (cen), telomere sequences (TG_{1–3})_n, and, in addition to the few hundred base pair (bp)-sized specific chromosomal elements, ≥50 kb of unspecific DNA.⁽⁵⁵⁾ Moreover, large inserts (>1 Mb) are possible with YACs, useful in construction of large insert libraries for mapping of genomes and large-scale cloning.

Although genomic BAC/PAC libraries generally have smaller insert sizes, stable PAC clones containing homogeneous alpha satellite arrays of ≤200 kb, a length comparable to the upper size range of the total library used (unpublished observations),^(56,57) have been isolated. Technical progress in the treatment of large DNA molecules for library production, such as mild electroporation conditions described for the construction of a chicken genomic BAC library, led to an average clone size of 425 kb.⁽⁵⁸⁾ Despite the larger potential capacity of YACs their usefulness for the cloning and mapping of foreign DNA is limited by frequent instability, chimerism, and recombination.^(59–61) The repetitive sequences found in substantial parts of mammalian genomes are believed to be one reason for the instability. This theory is supported by the difficulties in cloning large arrays of tandemly repeated DNA, even in recombination deficient yeast strains^(62,63)

A system to clone circular plasmids several hundred kb in size in human cells is human artificial episomal chromosomes (HAEC). These rely on the high copy number latent origin of replication (oriP) and on the replication promoting protein EBNA-1 from human Epstein-Barr virus (EBV).⁽⁶⁴⁾ Alternatively, bacterial and yeast cloning systems for large DNA could be adapted to construct second-generation MACs to stably clone low copy number chromosomes in mammalian cells. Such constructs preferably should (1) be as large as possible to ensure potential minimal chromosome size without the need for amplification and/or concatenation of input DNA; (2) have the capacity for large, possibly more efficient, centromere arrays and genomic copies of genes; (3) be flexible in composition, allowing easy exchange of certain regions to help define optimal MAC elements; (4) be stably cloned; and (5) yield high amounts of intact DNA to enable the development of efficient MAC delivery systems.

In vivo construction of second generation MACs using YACs

The potential capacity of YACs is one major advantage over current bacterial cloning systems. Using mitotic homologous recombination via specifically designed YAC linking vectors, two or more YACs can be joined in yeast cells.^(65,66) Attempts to reconstitute the human dystrophin gene by homologous recombination resulted in a 2.3-Mb YAC lacking only 100 kb (probably an unstable region).⁽⁶⁷⁾ Analogous to the recent series of YAC transgenic animal and cell culture experiments^(68–73) genomic regions carrying all sequence elements required for regulated and stable gene expression could be

combined with a YAC carrying a functional mammalian centromere. If amplification/concatenation of the YAC-based MAC construct 7C5 was needed to exceed the putative minimal size limit of human chromosomes, then joining large genomic sequences to 7C5 by YAC linking could produce molecules having similar dimensions. The free end of the joined YAC would have to be retrofitted with telomeric (TTAGGG)_n sequences before or after combination. Since retrofitting of telomeres and the joining reaction are carried out *in vivo* followed by subsequent growing steps, repeated structural analysis is necessary to ensure the integrity of the final construct. Now that sequencing of the YAC host *S. cerevisiae* has been completed, further analysis of gene products might well lead to improvements in YAC stability and cloning. One ongoing project is the creation of yeast cells having telomerase with a “humanized” RNA template that could eliminate the need to integrate (TTAGGG)_n sequences subtelomeric to YAC telomeres (M. Rosenfeld, personal communication).

Biochemical construction of MACs using PACs—implications for a centromere test

Middle-sized genes and large alpha satellite arrays can be stably and reliably cloned in PACs/BACs; however, problems arise with insert stability and size when they are cloned together into one large molecule. Recent progress in the isolation of large amounts (>5 mg/1 ml agarose plug) of PAC/BAC DNA, based on the lack of movement of large, intact circular DNA during pulsed-field gel electrophoresis, has allowed efficient *in vitro* joining of two large fragments carrying a lox site (specific 34-bp locus of crossing over),⁽⁷⁴⁾ using *in gel* site-specific recombination (IGSSR) mediated by purified Cre recombinase.⁽⁷⁵⁾ Since the large joined molecule can be isolated immediately, without an additional cloning step, no further analysis of DNA integrity is required. Many different substrate PACs to be joined can be prepared, analyzed and stored before the final composition is chosen for subsequent MAC constructs. The ease with which single parts of the construct may be exchanged without changing other regions allows rapid comparison of candidate components under controlled conditions.

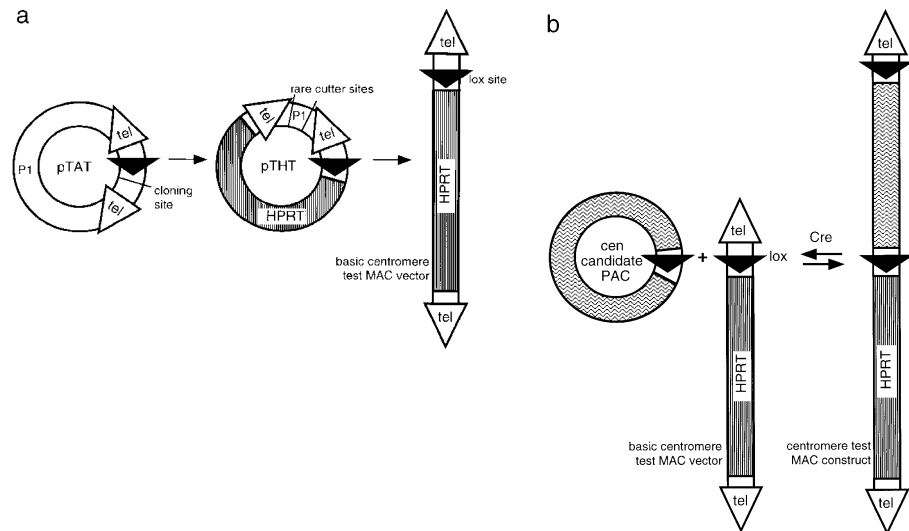
The newly developed ditelomeric PAC vector “pTAT,” currently under evaluation, could be used to stably clone a large centromere array and to subsequently release it as a linear fragment. This fragment would contain terminal telomeres and an internal lox site for comparatively rapid *in vitro* joining (insertion) of any unchanged circular PAC clone. To identify an optimal centromere, circular centromere candidate PACs could be switched into a linear ditelomeric MAC component containing genomic DNA with origin of replication function (statistically) and a selectable marker gene. Preferably, a large genomic region serving for both functions could be integrated between the telomeres. For this purpose, a

175-kb PAC containing the intact human hypoxanthine phosphoribosyl transferase (HPRT) gene was isolated for cloning into pTAT leading to a basic centromere test vector “pTHT” (Fig. 2). Transfection of a 95-kb genomic fragment containing the 42-kb HPRT primary transcript and flanking sequences showed that low copy numbers of 1–5, randomly integrated into the genome, were frequently observed in transfected HPRT(-) fibroblast cell lines under HAT selection (hypoxanthine, aminopterin, thymidine) (unpublished observations). Thus, the genomic HPRT gene could fulfil the requirements for MAC selection in the ideal case of one unamplified construct per cell.

The efficient formation of MACs is expected to allow a high relative rate of transfectants without integrating into the host genome. The percentage of integration events of resistant clones depends on both the functionality of the chromosomal components and the intactness of the delivered constructs. Additionally, it could be determined whether selection against HPRT gene function using 6-thioguanine (6TG) leads to HPRT gene inactivation (epigenetic gene silencing) or to loss of MACs in a manner dependent on centromere activity. To directly compare length and repeat structure influences on the efficiency of *de novo* centromere formation and mitotic stability, a selection of large alpha satellite PACs with varying features could be used. Further interesting candidates could include arrays of non-alphoid centromeric satellite sequences or neo-centromere regions. Finally, centromeric repeat arrays from other species (e.g., mouse minor satellite DNA) could be tested to develop animal MAC model systems.

The delivery of MACs has to be improved. Improved delivery of large constructs is required for potential applications of MACs. For the development of a test system to compare MAC constructs differing in single components (e.g., centromere) reproducible rates of intactly delivered DNA molecules would be preferable. One limiting step in developing tests for different MAC-construct/complex formulations is the low yield of MAC constructs. Improving the stability and reliability of the cloning system would further enhance testing, as large constructs are more sensitive to rare rearrangements or mutations during cloning that could alter MAC function. The proposed *in vitro* construction system is based on a new DNA isolation procedure allowing microgram quantities of purified, intact PAC/BAC DNA per agarose plug.⁽⁷⁵⁾ The factor >100 higher DNA yield from PACs/BACs compared to the DNA yield of a similar sized YAC results not from the copy number per cell, but the smaller volume of *E. coli* cells compared with *S. cerevisiae* (10²–10³ × less) and the avoidance of the “diluting” pulsed-field gel run needed for the separation of the host chromosome. Even if biochemical joining by IGSSR were only 10% efficient, there would still be a respectable construct yield.

Figure 2. A workable centromere test using the PAC-based MAC construction system. a: The ditelomeric vector pTAT, based on the P1-derived unit copy replicon (P1) of PCYPAC2N⁽⁵³⁾ and containing two homogeneous telomeric repeats (TTAGGG)_{n=135} stably cloned in a head-to-head configuration (tel), can be used to clone a genomic copy of the human HPRT gene. The resulting pTHT provides both, a selectable marker expressed in its "chromatin context," and origins of replication likely to be present in >100 kb of genomic DNA. Linearization of pTHT using a rare cutter (i.e., the extremely rare cutting intron-encoded restriction nuclease IScel) leads to a basic centromere test MAC vector containing terminal telomeric repeats, the genomic copy of the human HPRT gene, and a lox site to insert circular centromere sequences. b: Any centromere candidate sequence stably cloned in a PAC (cen candidate) can be readily inserted into the basic MAC vector in a single in vitro IGSSR reaction using purified Cre recombinase, which acts on the lox sequence present in both components. HPRT, hypoxanthine phosphoribosyl transferase; IGSSR, in gel site-specific recombination; MAC, mammalian artificial chromosome; PAC, P1 artificial chromosome.



MAC constructs have been transfected in cell culture using standard lipofection and microinjection,^(1,2,76) two of several delivery concepts presently under evaluation. Lipofection treats large numbers of cells simultaneously relying on high efficiency to transfect a major fraction, while microinjection uses single cells, technically demanding due to the needs both for gentle treatment of each cell and for automation to achieve high cell numbers. For transfection of cells in vitro and in vivo, methods have to be developed or improved with the ultimate aim of introducing intact MAC constructs into a large fraction of primary culture and tissue cells without the need for selection. Whether ballistic transfection or other "violent" procedures could be applied to deliver intact MAC constructs remains unclear and might depend on the development of gentle packaging procedures and stabilizing DNA carriers to prevent shearing of the large molecules. Shock-wave permeabilization, a recently introduced method based on a lithotripter with the potential to achieve DNA transfer to a restricted region of an organ or tissue in vivo, has been used in cell culture for molecules of ≤ 10 kb.⁽⁷⁷⁾ *E. coli* cells, generally difficult to penetrate due to their rigid cell wall, have been transformed with BACs of ≤ 725 kb using electroporation⁽⁵⁸⁾; however, intact delivery of very large fragments has not been reported for mammalian cells.^(78,79) A new Argon laser-based method for gene transfer called optoporation has not yet been used for large molecule transfection, however this mild treatment of single cells might become a useful method.^(80,81) Automated optical control of DNA uptake by

incorporation of a visible tracking system into the MAC constructs could be used for immediate laser-assisted elimination of nontransfected cells. The efficiency of a polyfection procedure was substantially improved using a viral entry mechanisms by incorporating psoralen inactivated adenovirus particles into polyplex formulations as a carrier to transfect large DNA molecules (BACs).⁽⁸²⁾

Conclusion

MACs have not yet lived up to their promise. A number of problems must be solved before they can be applied in biotechnology and medicine. The minimal chromosomal DNA elements must be further defined, functionality in different cell types and species still has to be shown, and efficient protocols for delivery into cells in culture and in situ have to be developed. One still poorly understood component is the centromere: which sequences are sufficient and which are efficient templates for kinetochore formation? MAC construction systems based on in vivo YAC linking or biochemical PAC joining have been proposed, allowing very large construct sizes on one hand and high accuracy and compositional flexibility, paired with high yields of constructs on the other. Availability of large amounts of reliable MAC constructs might aid the development of improved delivery systems. This should accelerate MAC research and lead to better understanding of the functional elements of mammalian chromosomes. Yeast and bacterial cloning systems adapted to construct MAC vectors containing a broad range of genes

can be envisaged. The new PAC-based construction system provides an important tool to achieve comparably rapid assembly of MACs differing in single chromosomal components. Since the DNA used for gene expression is cloned only once in a stable cloning system and no further subcloning after its structural and functional characterization is required, the construction system could represent a safe means for use in therapeutic protocols. To achieve this end, we face the difficult task of developing regulated long-term gene expression vectors based on genomic gene copies on MACs—individual mitotically stable genetic elements that neither integrate into the host genome nor rely on immunogenic viral proteins.

Acknowledgments

I thank Brenda Grimes, Peter Warburton, Rosemary Bayne, and Howard Cooke for many helpful discussions during my work as a guest researcher in the Chromosome Biology Section at the MRC Human Genetics Unit, Edinburgh; Titia de Lange for providing cloned telomeric repeats to the group and Rosemary Bayne for an IScel-capped version; Brenda Grimes for initial help in preparing the manuscript; Kerry Baldwin Jedele for review of the manuscript; and Thomas Meitinger for his support.

References

- Harrington JJ, Van Bokkelen G, Mays RW, Willard HF. Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. *Nature Genet* 1997;15:345–355.
- Ikeno M, Grimes B, Okazaki T, Nakano M, Saitoh K, Hoshino H, McGill NI, Cooke HJ, Masumoto H. Construction of YAC-based mammalian artificial chromosomes. *Nature Biotechnol* 1998;16:431–439.
- Farr C, Fantes J, Goodfellow P, Cooke HJ. Functional reintroduction of human telomeres into mammalian cells. *Proc Natl Acad Sci USA* 1991;88:7006–7010.
- Barnett MA, Buckle VJ, Evans EP, Porter AC, Rout D, Smith AG, Brown WR. Telomere directed fragmentation of mammalian chromosomes. *Nucleic Acids Res* 1993;21:27–36.
- Haaf T, Warburton PE, Willard HF. Integration of human alpha-satellite DNA into simian chromosomes: centromere protein binding and disruption of normal chromosome segregation. *Cell* 1992;70:681–696.
- Larin Z, Fricker MD, Tyler-Smith C. De novo formation of several features of a centromere following introduction of a Y alphoid YAC into mammalian cells. *Hum Mol Genet* 1994;3:689–695.
- Taylor SS, Larin Z, Tyler-Smith C. Analysis of extrachromosomal structures containing human centromeric alphoid satellite DNA sequences in mouse cells. *Chromosoma* 1996;105:70–81.
- 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.
- Wevrick R, Willard HF. Long-range organization of tandem arrays of alpha-satellite DNA at the centromeres of human chromosomes: high-frequency array-length polymorphism and meiotic stability. *Proc Natl Acad Sci USA* 1989;86:9394–9398.
- Abruzzo MA, Griffin DK, Millie EA, Sheehan LA, Hassold TJ. The effect of Y-chromosome alpha-satellite array length on the rate of sex chromosome disomy in human sperm. *Hum Genet* 1996;97:819–823.
- Steiner NC, Hahnenberger KM, Clarke L. Centromeres of the fission yeast *Schizosaccharomyces pombe* are highly variable genetic loci. *Mol Cell Biol* 1993;13:4578–4587.
- Ikeno M, Masumoto H, Okazaki T. Distribution of CENP-B boxes reflected in CREST centromere antigenic sites on long-range alpha-satellite DNA arrays of human chromosome 21. *Hum Mol Genet* 1994;3:1245–1257.
- Masumoto H, Masukata H, Muro Y, Nozaki N, Okazaki T. A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. *J Cell Biol* 1989;109:1963–1973.
- Haaf T, Mater AG, Wienberg J, Ward DC. Presence and abundance of CENP-B box sequences in great ape subsets of primate-specific alpha-satellite DNA. *J Mol Evol* 1995;41:487–491.
- Goldberg IG, Sawhney H, Pluta AF, Warburton PE, Earnshaw WC. Surprising deficiency of CENP-B binding sites in African green monkey alpha-satellite DNA: implications for CENP-B function at centromeres. *Mol Cell Biol* 1996;16:5156–5168.
- Yoda K, Nakamura T, Masumoto H, Suzuki N, Kitagawa K, Nakano M, Shinjo A, Okazaki T. Centromere protein B of African green monkey cells: gene structure, cellular expression, and centromeric localization. *Mol Cell Biol* 1996;16:5169–5177.
- Earnshaw WC, Rothfield N. Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma* 1985;91:313–321.
- Pluta AF, Earnshaw WC. Specific interaction between human kinetochore protein CENP-C and a nucleolar transcriptional regulator. *J Biol Chem* 1995;271:18767–18774.
- 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;4:2189–2197.
- Pelizon C, Diviacco S, Falaschi A, Giacca M. High-resolution mapping of the origin of DNA replication in the hamster dihydrofolate reductase gene domain by competitive PCR. *Mol Cell Biol* 1996;16:5358–5364.
- Jaenisch R. Transgenic animals. *Science* 1988;240:1468–1474.
- Garrick D, Fiering S, Martin DIK, Whitelaw E. Repeat-induced gene silencing in mammals. *Nature Genet* 1998;18:56–59.
- Karpen GH, Allshire RC. The case for epigenetic effects on centromere identity and function. *Trends Genet* 1997;13:489–496.
- Sandell LS, Zakian VA. Loss of a yeast telomere: arrest, recovery, and chromosome loss. *Cell* 1993;75:729–739.
- Allsopp RC, Vaziri H, Patterson C, Goldstein S, Younglai EV, Futcher AB, Greider CW, Harley CB. Telomere length predicts replicative capacity of human fibroblasts. *Proc Natl Acad Sci USA* 1992;89:10114–10118.
- Melek M, Shippen DE. Chromosome healing: spontaneous and programmed de novo telomere formation by telomerase. *BioEssays* 1996;18:301–308.
- van Steensel B, Smogorzewska A, deLange T. TRF2. protects human telomeres from end-to-end fusions. *Cell* 1998;92:401–413.
- du Sart D, Cancilla MR, Earle E, Mao J-i, Saffery R, Tainton KM, Kalitsis P, Martyn J, Barry AE, Choo AKH. A functional neo-centromere formed through activation of a latent human centromere and consisting of non-alpha-satellite DNA. *Nature Genet* 1997;16:144–153.
- Depinet TW, Zackowski JL, Earnshaw WC, Kaffe S, Sekhon GS, Stallard R, Sullivan BA, Vance GH, Van Dyke DL, Willard HF, Zinn AB, Schwartz S. Characterisation of neo-centromeres in marker chromosomes lacking detectable alpha-satellite DNA. *Hum Mol Genet* 1997;6:1195–1204.
- Choo KH. Centromere DNA dynamics: latent centromeres and neocentromere formation. *Am J Hum Genet* 1997;61:1225–1233.
- Choo KHA. Turning on the centromere. *Nature Genet* 1998;18:3–4.
- Farr CJ, Bayne RA, Kipling D, Mills W, Critcher R, Cooke HJ. Generation of a human X-derived minichromosome using telomere-associated chromosome fragmentation. *EMBO J* 1995;14:5444–5454.
- Heller R, Brown K, Burgtorf C, Brown W. Mini-chromosomes derived from the Y chromosome by telomere directed chromosome breakage. *Proc Natl Acad Sci USA* 1996;93:7125–7130.
- Bayne RA, Broccoli D, Taggart MH, Thomson EJ, Farr CJ, Cooke HJ. Sandwiching of a gene within 12 kb of a functional telomere and alpha satellite does not result in silencing. *Hum Mol Genet* 1994;3:539–546.
- Shen MH, Yang J, Loupart M-L, Smith A, Brown W. Human minichromosomes in mouse embryonal stem cells. *Hum Mol Genet* 1997;6:1375–1382.
- Murphy TD, Karpen GH. Localization of centromere function in a *Drosophila* minichromosome. *Cell* 1995;82:599–609.
- Williams BC, Murphy TD, Goldberg ML, Karpen GH. Neocentromere activity of structurally acentric mini-chromosomes in *Drosophila*. *Nature Genet* 1998;18:30–37.

38. Featherstone T, Huxley C. Extrachromosomal maintenance and amplification of yeast artificial chromosome DNA in mouse cells. *Genomics* 1993;17:267–278.
39. Tomizuka K, Yoshida H, Uejima H, Kugoh H, Sato K, Ohguma A, Hayasaka M, Hanaoka K, Oshimura M, Ishida I. Functional expression and germline transmission of a human chromosome fragment in chimaeric mice. *Nature Genet* 1997;16:133–143.
40. Meyne J, Ratliff RL, Moyzis RK. Conservation of the human telomere sequence (TTAGGG)_n among vertebrates. *Proc Natl Acad Sci USA* 1989;86:7049–7053.
41. de Lange T, Shiue L, Myers RM, Cox DR, Naylor SL, Killery AM, Varmus HE. Structure and variability of human chromosome ends. *Mol Cell Biol* 1990;10:518–527.
42. Kipling D, Cooke HJ. Hypervariable ultra-long telomeres in mice. *Nature* 1990;347:400–402.
43. Small MB, Hubbard K, Pardinas JR., Marcus AM, Dhanaraj SN, Sethi KA. Maintenance of telomeres in SV40-transformed pre-immortal and immortal human fibroblasts. *J Cell Physiol* 1996;168:727–736.
44. Bryan TM, Englezou A, Dalla-Pozza L, Dunham MA, Reddel RR. Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. *Nature Med* 1997;3:1271–1274.
45. Morii K, Tanaka R, Onda K, Tsumanuma I, Yoshimura J. Expression of telomerase RNA, telomerase activity, and telomere length in human gliomas. *Biochem Biophys Res Commun* 1997;239:830–834.
46. Marçais B, Charliou JP, Allain B, Brun E, Bellis M, Roizes G. On the mode of evolution of alpha satellite DNA in human populations. *J Mol Evol* 1991;33:42–48.
47. Willard HF. Evolution of alpha satellite. *Curr Opin Genet Dev* 1991;1:509–514.
48. Zelesco PA, Graves JA. Chromosome segregation from cell hybrids. IV. Movement and position of segregant set chromosomes in early-phase interspecific cell hybrids. *J Cell Sci* 1988;89:49–56.
49. Graves JA, Barbieri I. Chromosome segregation from cell hybrids. VII. Reverse segregation from karyoplast hybrids suggests control by cytoplasmic factors. *Genome* 1992;35:537–540.
50. Abbott C, Povey S. Somatic cell hybrids. The basics. Fate of hybrid cells and expansion of the culture. Beynon R, Brown TA, Howe C, editors. Oxford: IRL Press; 1995. p 16–19.
51. Jackson DA, Symons RH, Berg P. Biochemical method for inserting new genetic information into DNA of Simian Virus 40: circular SV40. DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli*. *Proc Natl Acad Sci USA* 1972;69:2904–2909.
52. Shizuya H, Birren B, Kim U-J, Mancino V, Slepak T, Tachiiri Y, Simon M. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc Natl Acad Sci USA* 1992;89:8794–8797.
53. Ioannou PA, Amemiya CT, Garnes J, Kroisel PM, Shizuya H, Chen C, Batzer MA, de Jong PJ. A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nature Genet* 1994;6:84–89.
54. Austin SJ. Bacterial plasmids that carry two functional centromere analogs are stable and are partitioned faithfully. *J Bacteriol* 1984;158:742–745.
55. Murray AW, Szostak JW. Construction of artificial chromosomes in yeast. *Nature* 1983;305:189–193.
56. Mejia JE, Monaco AP. Retrofitting vectors for *Escherichia coli*-based artificial chromosomes (PACs and BACs) with markers for transfection studies. *Genome Res* 1997;7:179–186.
57. Patsalis PC, Rose N, Georgioui A, Yurov Y, Vorsanova S, deJong PJ, Ioannou P. Identification and characterization of a large subset of human PAC clones with centromeric and telomeric probes. *Cytogenet Cell Genet* 1997;77:27.
58. Zimmer R, Verrinder Gibbins AM. Construction and characterization of a large-fragment chicken bacterial artificial chromosome library. *Genomics* 1997;42:217–226.
59. Ried K, Mertz A, Nagaraja R, Trusgnich M, Riley JH, Anand R, Lehrach H, Page D, Ellison JW, Rappold G. Characterization of a YAC contig spanning the pseudoautosomal region. *Genomics* 1995;29:787–792.
60. Schindelbauer D, Hellebrand H, Grimm L, Bader I, Meitinger T, Wehnert M, Ross M, Meindl A. Long-range map of a 3.5-Mb region in Xp11.23–22. with a sequence-ready map from a 1.1-Mb gene-rich interval. *Genome Res* 1996;6:1056–1069.
61. Courseaux A, Szepietowski P, Fernandes M, Serizet C, Kawaguchi Y, Grosgeorge J, Perucca-Lostanlen D, Shows TB, Todd JA, Nowak NJ, Gaudray P. Framework YAC contig anchored into a 3.2-Mb high-resolution physical map in proximal 11q13. *Genomics* 1997;40:13–23.
62. Neil D, Villasante A, Fisher RB, Vetrie D, Cox B, Tyler-Smith C. Structural instability of human tandemly repeated DNA sequences cloned in yeast artificial chromosome vectors. *Nucleic Acids Res* 1990;18:1421–1428.
63. Kohno K, Oshiro T, Kishine H, Wada M, Takeda H, Ihara N, Imamoto F, Kano Y, Schlessinger D. Construction and characterization of a rad51rad52 double mutant as a host for YAC libraries. *Gene* 1997;188:175–181.
64. Sun T-Q, Fenstermacher DA, Vos J-M. Human artificial episomal chromosomes for cloning large DNA fragments in human cells. *Nature Genet* 1994;8:33–41.
65. Larin Z, Taylor SS, Tyler-Smith C. A method for linking yeast artificial chromosomes. *Nucleic Acids Res* 1996;24:4192–4196.
66. Larin Z. Development of mammalian artificial chromosome vectors: prospects for somatic gene transfer. In: Lemoine NR, Cooper DN, editors. *Gene therapy*. BIOS Scientific Publishers Ltd; Oxford UK 1996. p 113–125.
67. Den Dunnen JT, Grootsholten PM, Dauwerse JG, Walker AP, Monaco AP, Butler R, Anand R, Coffey AJ, Bentley DR, Steensma HY, Van Ommen GJB. Reconstruction of the 2.4 Mb human DMD-gene by homologous YAC recombination. *Hum Mol Genet* 1992;1:19–28.
68. Hodgson JG, Smith DJ, McCutcheon K, Koide HB, Nishijima K, Dinulos MB, Stevens ME, Bissada N, Nasir J, Kanazawa I, Distech CM, Rubin EM, Hayden MR. Human huntingtin derived from YAC transgenes compensates for loss of murine huntingtin by rescue of the embryonic lethal phenotype. *Hum Mol Genet* 1996;5:1875–1885.
69. Montoliu L, Umland T, Schutz G. A locus control region at –12 kb of the tyrosinase gene. *EMBO J* 1996;15:6026–6034.
70. Schedl A, Ross A, Lee M, Engelkamp D, Rashbass P, van Heyningen V, Hastie ND. Influence of PAX6 gene dosage on development: overexpression causes severe eye abnormalities. *Cell* 86:71–82.
71. Vassaux G, Manson AL, Huxley C. 1997. Copy number-dependent expression of a YAC-cloned human CFTR gene in a human epithelial cell line. *Gene Ther* 1996;4:618–623.
72. Fujiwara Y, Miwa M, Takahashi R, Hirabayashi M, Suzuki T, Ueda M. Position-independent and high-level expression of human alpha-lactalbumin in the milk of transgenic rats carrying a 210-kb YAC DNA. *Mol Reprod Dev* 1997;47:157–163.
73. Manson AL, Trezise AE, MacVinish LJ, Kasschau KD, Birchall N, Episkopou V, Vassaux G, Evans MJ, Colledge WH, Cuthbert AW, Huxley C. Complementation of null CF mice with a human CFTR YAC transgene. *EMBO J* 1997;16:4238–4249.
74. Abremski K, Hoess R, Sternberg N. Studies on the properties of P1 site-specific recombination: evidence for topologically unlinked products following recombination. *Cell* 1983;32:1301–1311.
75. Schindelbauer D, Cooke HJ. Efficient combination of large DNA in vitro: in gel site specific recombination (IGSSR) of PAC fragments containing alpha satellite DNA and the human HPRT gene locus. *Nucleic Acids Res* 1997;25:2241–2243.
76. Schedl A, Grimes B, Montoliu L. YAC transfer by microinjection. *Methods Mol Biol* 1996;54:293–306.
77. Lauer U, Bürgelt E, Squire Z, Messmer K, Hofschneider PH, Gregor M, Delius M. Shock wave permeabilization as a new gene transfer method. *Gene Ther* 1997;4:710–715.
78. Wada M, Ihara Y, Tatsuka M, Mitsui H, Kohno K, Kuwano M, Schlessinger D. HPRT yeast artificial chromosome transfer into human cells by four methods and an involvement of homologous recombination. *Biochem Biophys Res Commun* 1994;200:1693–1700.
79. Hejna JA, Johnstone PL, Kohler SL, Bruun DA, Reifsteck CA, Olson SB, Moses RE. Functional complementation by electroporation of human BACs into mammalian fibroblast cells. *Nucleic Acids Res* 1998;26:1124–1125.
80. Turvets I, Lewis A, Palanker D, Gilo H, Vilenz A, Broder JC, Lewis S. Permeabilizing millions of cells with single pulses of an excimer laser. *Biotechniques* 1993;15:1022–1024.
81. Palumbo G, Caruso M, Crescenzi E, Tecce MF, Roberti G, Colasanti A. Targeted gene transfer in eucaryotic cells by dye-assisted laser optoporation. *J Photochem Photobiol B* 1996;36:41–46.
82. Baker A, Cotten M. Delivery of bacterial artificial chromosomes into mammalian cells with psoralen-inactivated adenovirus carrier. *Nucleic Acids Res* 1997;25:1950–1956.