The EMBO Journal vol.11 no.2 pp.417-422, 1992

Molecular complementation of a collagen mutation in mammalian cells using yeast artificial chromosomes

William M.Strauss and Rudolf Jaenisch

Whitehead Institute for Biomedical Research and Department of Biology, Massachusettes Institute of Technology, Nine Cambridge Center, Cambridge, MA 02142, USA

Communicated by R.Jaenisch

The cloning of large contiguous segments of mammalian DNA in *Saccharomyces cerevisiae* has become possible with the advent of Yeast Artificial Chromosomes (YACs). We are interested in extending the technique of genetic complementation analysis to the molecular level through the introduction of YACs into mammalian cells and the mammalian germline. We report the successful introduction of homogeneous DNA derived from a 150 kbp YAC spanning the murine *Col1a1* locus into murine fibroblasts carrying a mutation at this locus. The YAC DNA was fractionated by pulse field electrophoresis, condensed with polyamines, and introduced into mutant fibroblasts via DNA-lipid micelles. The DNA was integrated as a stable intact unit in 10% of the transfected clones conferring collagen RNA expression to the mutant cells.

Key words: collagen/genetic complementation/mouse/Mov-13/yeast artificial chromosomes

Introduction

One goal of the current Genome Initiatives (Coulson *et al.*, 1986, 1988; Olson *et al.*, 1986; Garza *et al.*, 1989; Roberts, 1989; Watson, 1990) is the establishment of physical maps of complex genomes of various organisms. Comparison of the physical map to genetic maps will be an important means to identify and isolate new genes. One possible way to relate the two maps to each other and to resolve discrepancies between them is through complementation of genetically defined mutations by physically defined segments of the genome. Strategies involving cDNAs cannot supply the regulatory sequences required for spatio-temporal developmental expression. Furthermore, many genes are known to be much larger than can be accommodated in current cloning vectors (Gitschier *et al.*, 1984; Koenig *et al.*, 1987). This impedes efforts to study the structure and function of such genes using experimental approaches that involve DNA mediated gene transfer into mammalian cells or animals.

A general scheme for molecular complementation of large genes and mammalian mutations *in vivo* requires the ability to introduce a complete transcriptional complex into a mammalian cell. The introduction of the requisite large pieces of cloned DNA into the target cell represents a technical hurdle both for DNA cloning and for transfection technology. With the advent of Yeast Artificial Chromosome (YAC) cloning technology (Murry and Szostak, 1983; Dawson *et al.*, 1986; Murry *et al.*, 1986; Burke *et al.*, 1987) the first of these barriers was overcome.

The second hurdle has initially been approached by efforts in which the entire YAC-containing yeast genome was introduced into mammalian cells, resulting in transfer of both YAC DNA and random fragments of the yeast genome. The whole yeast genome was introduced into mammalian cells either by spheroplast fusion (Ward *et al.*, 1986; Allshire *et al.*, 1987; Pachnis *et al.*, 1990; Pavan *et al.*, 1990) using high molecular weight precipitates (D'Urso *et al.*, 1990; Eliceiri *et al.*, 1991) or by using DNA-lipid micelles (Gnirke *et al.*, 1991) of crude, unpurified YAC-containing yeast DNA. In these studies large anonymous YACs were introduced, but not assayed for function. Alternatively direct selection of enzymatic function contained within the YAC was employed. In this report we present a general method for introducing homogeneous gel purified YAC DNA into mammalian cells. We show that a 150 kb YAC clone spanning the α1(I) collagen locus was introduced intact and RNA was expressed in murine fibroblasts.

As a well defined model system, we chose to complement mutant cell lines derived from a homozygous mutant mouse line, Mov13, defective in proα1(I) collagen expression (Jaenisch *et al.*, 1983; Schnieke *et al.*, 1983). This recessive mutation was induced by the insertion of a Moloney virus genome into the first intron of the collagen gene resulting in a defective collagen allele. The viral integration causes a block of collagen transcription and results in viable homozygous fibroblasts but perinatal death of homozygous embryos (Jaenisch *et al.*, 1983; Schnieke *et al.*, 1983).

Results

Structure of collagen YAC clone (Y22)

The aim of these experiments was to establish procedures which would allow the introduction of a YAC clone representing the *Col1a1* locus into mammalian cells. It was therefore crucial to verify the structure of the transfected DNA and its intactness by molecular analysis. The use of *Mus spretus* in interspecific mouse backcrosses to standard laboratory inbred lines has proven to be a powerful method for mapping genes (Copeland and Jenkins, 1991; Nadeau *et al.*, 1991) and we exploited this genetic divergence for our analysis as it allowed us to distinguish a wild-type *M.spretus* collagen transgene in a *Mus musculus* background by standard RFLP linkage analysis (Botstein *et al.*, 1980). A *M.spretus* DNA YAC library was generated from DNA digested to completion with *Sal*I and a 150 kb YAC clone (Y22) spanning the collagen gene was recovered (Strauss *et al.*, 1992). The left hand cloning arm (see Figure 1) of the YAC clone contains *TRP1/ARS1, CEN4* and *Pgk-Neo*, the latter marker permitting drug selection in mammalian cells. The right hand arm contains the *URA3* marker. The

5′ end of the collagen gene is located ~10 kb from the left cloning site and is transcribed in the opposite direction from the neomycin gene.

The following restriction digests were used to analyze the structure of the transfected DNA: a *Pvu*II polymorphism at the 5′ end of the *Col1a1* gene and a *Bam*HI polymorphism at the 3′ end distinguished the *M.spretus Col1a1* allele from the endogenous *Col1a1* allele when probed with a collagen

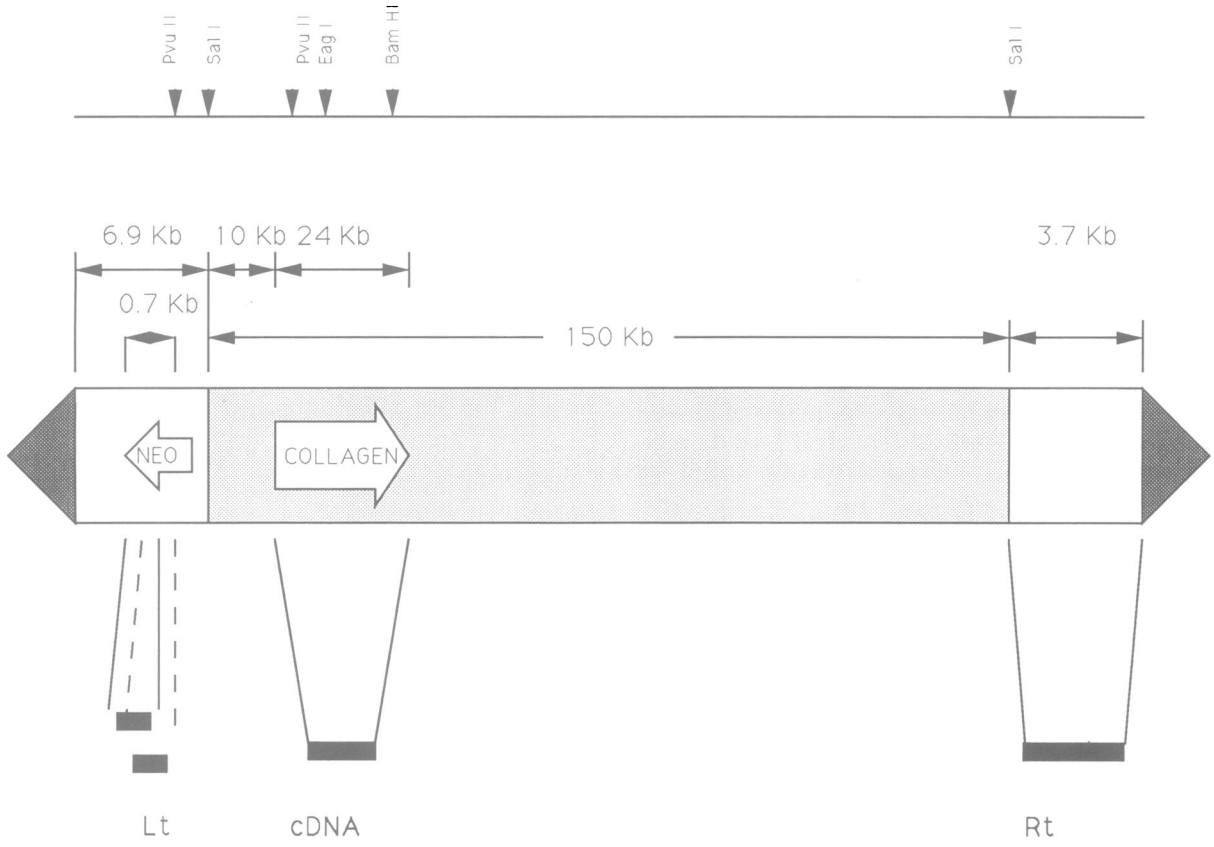

Fig. 1. A diagram of YAC clone 22 which contains the *M.spretus* α1(I) collagen allele. The YAC contains a unique *Eag*I site within the collagen transcription unit. The vector arms contained a single *Pvu*II site within the *Neo* gene. The collagen transcription unit is contained within the 150 kb clone and the start site for transcription is ~10 kb from the left hand cloning arm. The left arm contains the *TRP1/ARS1* and *CEN4* functionalities. The right arm contains the *URA3* gene. The location of probes Lt, collagen cDNA and Rt which were used for hybridization is indicated.

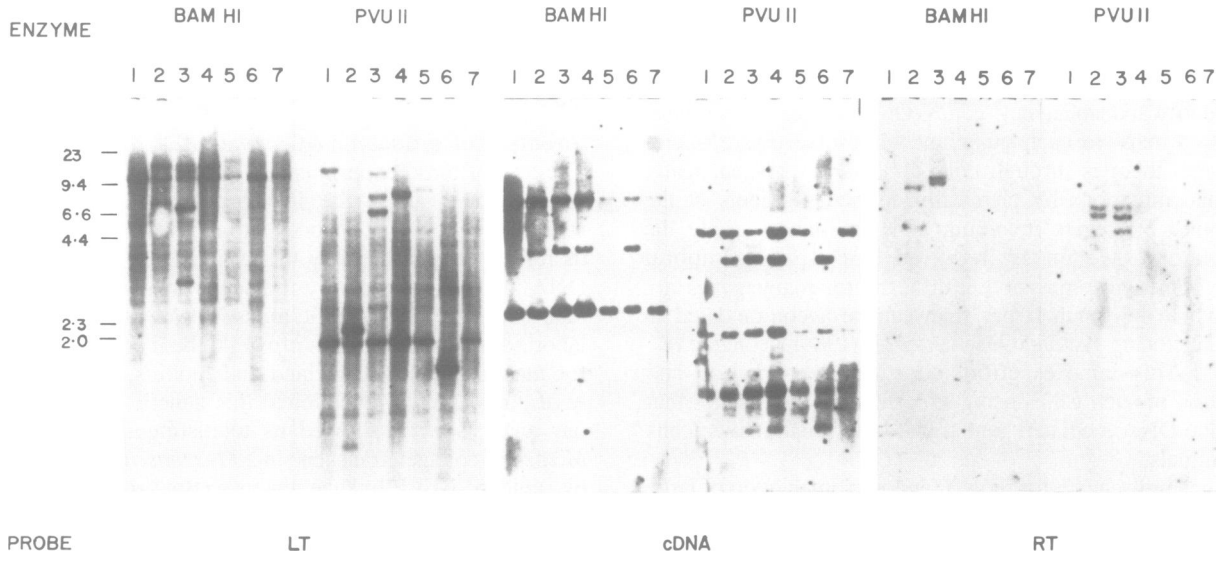

Fig. 2. Southern analysis of YAC-Y22 transfected clones. The DNA was digested with *Bam*HI or *Pvu*II, the filters sequentially hybridized to probes Lt, *Col1a1* cDNA and Rt. In this experiment the Lt probe contained *PGK* and *NEO* sequences and therefore cross hybridized to endogenous *PGK* elements. Lanes 1–7 contained 10 μg of genomic DNA digested with the indicated restriction enzymes. The numerically marked lanes correspond to genomic DNAs from clone MC16 (lane 1), MC15 (lane 2), MC44 (lane 3), MC40 (lane 4), MC17 (lane 5) and controls *M.spretus* (lane 6) and C57BL/6 (lane 7). Molecular size is indicated on the far left in kbp.

cDNA (see below Figure 2, *M.spretus*, lane 6, and C57BL/6, lane 7. The probes Lt and Rt allowed the visualization of the left or right hand sides of the cloned DNA when digested with non-cutting enzymes *Not*I, *Mlu*I, single cutter *Eag*I or multiple cutter *Bam*HI or *Pvu*II (see Figure 1).

Isolation of deproteinized YAC DNA

In order to recover the largest quantity of DNA possible for the transfection studies the saturation density for yeast embedded in agarose plus was determined. The maximum concentration of yeast cells in agarose plugs was determined to be 4×10^9 yeast cells/ml. At this density it was estimated that ~ 1 μg of a 150 kb YAC could be recovered from each milliliter of PFG gel. The high cell density in the agarose plugs enabled the recovery of 1–5 μg of YAC DNA per milliliter of PFG slice, which corresponds to 7.25–36.25 fmol for a 150 kb molecule. In practice, 100–500 ng of DNA were used per transfection.

To increase the stability of YAC DNA in the transfection protocol the effect of DNA condensing agents was analyzed. For this, the effect of increasing concentrations of spermine on stability and solubility of T-even phage DNA or *Saccharomyces cerevisiae* chromosomes was studied. We found that at low ionic strength a concentration of 1 mM spermine would aggregate and precipitate condensed DNA. Concentrations between 50 and 100 μM spermine in 20 mM Tris pH 7.6, 1 mM EDTA, however, protected the high molecular weight DNA against fragmentation, as tested by PFGE, while still keeping the DNA in solution (data not shown).

The use of spermine was essential to protect the YAC against breakage during handling. Transfections which did not include spermine contained no intact clones despite yielding similar numbers of G418 resistant clones. The optimal concentration for obtaining G418 resistant clones carrying an intact YAC clone was determined to be 50 μM Higher concentrations of polyamine did not improve integrity of YAC DNA in our titration assays, but caused massive aggregation and precipitation of the YAC DNA in the low ionic strength buffers used for transfection and inhibited transfection frequency.

After recovery of the PFG gel slice containing the YAC DNA, and prior to further manipulation, the gel was dialyzed in buffer containing 50 μM spermine–HCl for 36–48 h at 4°C. This procedure proved to be adequate to preserve integrity of high molecular weight DNA in a soluble form during subsequent manipulations.

Finally the use of DOTMA to form DNA–lipid micelles was effective in the delivery of YAC DNA to cellular targets. Formation of high molecular weight precipitates could be avoided by adjusting the concentration of lipid during the micelle formation step. The ratio of spermine (polycation condensing agent) to DOTMA (cationic lipid) was calibrated so as to minimize their competition for DNA.

Transfection of YAC DNA into cells

During the transfection procedure spermine was maintained at 50 μM prior to and during micelle formation. As the DNA–micelle mixture is toxic, fibroblasts were exposed for no longer than 8–12 h. The density of the monolayer was found to be important for maintaining the viability of a majority of the cells. The cells were kept at >90% confluence during transfection steps and subsequently split to low cell density prior to selection (<30%).

After transfections of 8×10^7 M13-5 target cells, 18 Mov13 derived G418 resistant clones were obtained and analyzed. To further characterize these clones, DNA was isolated for examination by Southern analysis. The results of these experiments are presented in Figure 2. The collagen gene RFLPs were detected when control *M.spretus* DNA (lanes 6) or C57BL/6 DNA (lanes 7) were hybridized to the collagen cDNA probe (middle panel). *Bam*HI digestion produced a 3.5 kb band in *M.spretus* DNA which was absent in C57BL/6 while *Pvu*II digestion generated a 3 kb fragment in *M.spretus* DNA and a 4.4 kb fragment in C57BL/6 DNA. Lanes 1–5 contain DNA isolated from G418 resistant clones. Three out of a total of 18 drug resistant clones revealed the presence of the collagen gene RFLPs as shown in MC15 (lane 2), MC44 (lane 3) and MC40 (lane 4), indicating that these clones contained donor *Col1a1* gene sequences. DNA from the other clones, (for example in lanes 1 and 5), failed to reveal the *M.spretus* specific *Col1a1* gene from the transfected Y22 DNA.

When the same blots were hybridized to the *Pgk–Neo* probe (Lt in Figure 1), a number of bands were detected in all lanes corresponding to the endogenous *Pgk* genes. In all neo resistant clones, however, one or more additional fragments were detected (lanes 1–5). Because the left arm of the vector has no *Bam*HI or *Pvu*II sites distal to the probe, these fragments reveal that one or several (lanes 2 and 3)

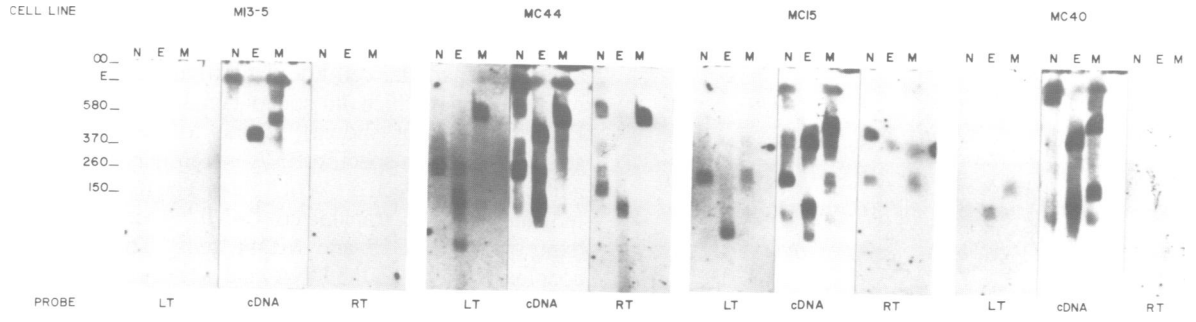

Fig. 3. PFG analysis of transfected clones. PFGE was performed on high molecular weight DNA after digestion with the restriction enzymes *Not*I, *Eag*I and *Mlu*I. The filter was sequentially hybridized to probes Lt, *Col1a1* cDNA and Rt. Probe Lt only contained *Neo* sequences and therefore does not cross hybridize to endogenous elements. Cell lines are indicated on top, probes on the bottom, molecular size in kbp is shown on left, the symbols ∞ and E__ refer to the origin and the exclusion zone on the PFG. The letters N, E, M refer to the restriction enzymes *Not*I, *Eag*I, and *Mlu*I.

copies of the left arm vector have integrated in the transfected cells rendering the cells G418 resistant. In contrast, when the blots were probed with the right vector arm specific Rt fragment which is 150 kb distant from the *neo* gene (compare Figure 1), only the DNA from MC15 and MC44 (lanes 2 and 3) showed several hybridizing bands. The results in Figure 2 therefore indicated that the latter two clones (clones #15 and #44) contained two or three copies respectively of the left and right vector arms as well as the *M.spretus* specific *Col1a1* RFLP. The experiments described in the next section were carried out to analyze the overall structure of the transfected YAC DNA.

Organization of transfected DNA

Pulse field electrophoresis was performed to study whether the three markers Lt, collagen and Rt would colocalize on the same high molecular weight restriction fragment. DNA from parental M13-5 cells (left panel of Figure 3), from clones MC44 and MC15 (middle panels) and from clone MC40 (right panel) were digested with *Not*I (N), *Eag*I (E) and *Mlu*I (M), and probed with the three different probes as indicated in the figure. *Not*I and *Mlu*I do not cut in the YAC clone while *Eag*I has a single recognition site in the *Col1a1* sequences (compare Figure 1). As expected, DNA from the parental M13-5 cells showed no hybridization to the Rt and Lt probes. Digestion with *Eag*I, which has a single recognition site in the *Col1a1* gene, generated only a single collagen specific fragment. It is possible that the *Eag*I site in the endogenous collagen gene is methylated or that the band at 370 represents a doublet. Digestion with *Mlu*I produced two fragments of 500 kb and 650 kb, the upper band possibly being due to a partial digest. Digestion with *Not*I did not resolve a collagen specific band.

As shown in the middle two panels, all three probes detected additional common fragments in DNA isolated from clones MC44 and MC15 and digested with the three different enzymes. *Not*I resolved several additional collagen specific bands in clones MC44 and MC15, one of which (260 kb in MC44 and 240 kb in MC15) hybridized also with both vector arms. Similarly, *Mlu*I resolved a 500 kb band in MC44 and a 200 kb fragment in MC15 which hybridized with all three probes. The single cutter *Eag*I generated two collagen specific bands of 260 and 150 kb in DNA from MC44 which also hybridized to either Lt probe only (260 kb) or to both the Lt and Rt probes (150 kb).

Whether this latter band represented a single fragment or a doublet was not resolved. Digestion of clone MC15 DNA with *Eag*I produced a 70 kb band which hybridized with the Lt and collagen probes and a 370 kb band which hybridized with the Rt and the collagen probes. Additional bands were present which hybridized to one or the other probe but not jointly to all three probes. In contrast to the clones MC15 and 44, no signal was detected in MC40 DNA with the Rt probe whereas the Lt and the collagen probes detected some common bands.

The PFG analyses in Figure 3 are consistent with the results described in Figure 2 which suggested that all three markers, the *Col1a1* RFLP and the right and the left vector arms were carried in clones 15 and 44, but that only the collagen gene and the left arm were present in clone 40. Importantly, the results strongly suggest that at least one intact copy of the transfected YAC clone had stably integrated in the genomes of clones MC15 and MC44, while MC40 carried a partially deleted copy. It is likely that all clones carried additional partially deleted DNA copies, the structure of which was not resolved.

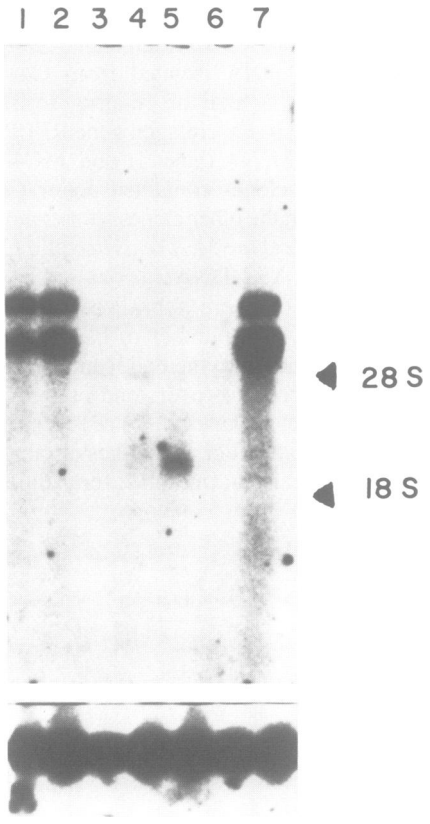

Fig. 4. Northern blot of Y22 transfected clones. The filter was hybridized to the *Col1a1* cDNA probe. Lanes 1–5 contain DNA from clones MC15, MC44, MC16, MC17 and MC18, and lanes 6 and 7 from M13-5 and 3T3 cells, respectively. The positions of 28S and 18S rRNAs are indicated on the right for reference. The lower panel is a reprobing of the same blot with a rat tubulin probe, to show equal loading of the lanes.

Expression of the transfected Col1a1 gene

To assess whether the transfected *Col1a1* locus was expressed, RNA was isolated from several G418 resistant clones and examined by Northern blot analysis. Figure 4 shows that both clones MC15 and MC44 (lanes 1 and 2) as well as the fibroblast cell line 3T3 (lane 7) transcribed the expected two proα1(I) collagen specific mRNA species (Schnieke *et al.*, 1987). This RNA was not synthesized by the M13-5 parental cell line (lane 6) and three other G418 resistant clones which did not contain the *M.spretus* collagen RFLPs (data not shown). Clone MC18 transcribed an aberrant collagen specific RNA which may originate from a truncated transfected *Col1a1* gene. The results in Figure 4 indicate that the transfected *Col1a1* gene was normally expressed in MC15 and MC44 cells. This supports the conclusion based on the DNA analyses described above that these two clones carry an intact copy of the transfected Y22 clone.

Discussion

We are interested in developing a complementation strategy where defined YACs are introduced into the murine germline

through transfection of embryonic stem cells (Robertson, 1986). Development of molecular complementation in the mammal could facilitate positional cloning, analysis of very large or complex genes, and elucidation of the role of higher order chromosome organization in cellular and organismal development. As a first step we have devised a method for transfecting murine fibroblasts with highly purified YAC DNA. The main points of this work were the development of procedures for isolating intact homogeneous YAC DNA, introduction of this DNA into mammalian cells and the demonstration of its biological activity.

The successful transfection of gel purified YAC cloned DNA was influenced by three conditions: (i) quantity of DNA, (ii) preservation of DNA structure with polyamines, and (iii) use of DNA−lipid micelles for delivery of YAC DNA to fibroblast targets. We optimized recovery of YAC DNA from the PFG gel by increasing the loading capacity of the agarose plugs. The quantity of spermine−HCl was carefully adjusted to preserve DNA structure in solution, yet not to compete with the DOTMA. Furthermore, this method for DNA handling and condensation was utilized instead of other methods (Couto *et al.*, 1989) as it yielded superior results in our hands. Using a micellar solution maintained the DNA in a soluble state, avoiding high molecular weight precipitates, and resulted in the introduction of intact DNA molecules into the cells.

It was important to verify that an intact copy of the 150 kb YAC clone could be introduced into mammalian cells and to determine the frequency of physical damage to the transfected DNA inflicted by the transfection procedure. However, assaying 150 kb of genomic DNA for all possible alterations in native structure is technically difficult. This problem is further accentuated by the requirement to use cloned mouse DNA allelic to a mouse mutation. Distinguishing between the two alleles can be impossible with the existing molecular technologies. In an effort to overcome these difficulties we chose to use a 150 kb *Sal*I fragment spanning the α(I) collagen locus from *M.spretus*. The use of RFLPs between *M.spretus* and *M.musculus* allowed us to distinguish between endogenous DNA and transgene. Standard Southern analysis indicated that the markers derived from the vector arms as well as the *Col1a1* marker were present in the DNA of two out of a total of 18 clones examined. The PFG analyses in Figure 3 indicated that at least one intact copy of Y22 was stably integrated in the genome of each of these clones. This conclusion was based upon the colocalization of all three markers, either derived from the ends of Y22 (probes Lt or Rt) or from the internally located *Col1a1* gene, to the same high molecular weight fragment. Other bands which did not hybridize with all three markers suggested that the clones carried, in addition to an intact Y22 copy, partially deleted or rearranged copies. As expected, all G418 resistant clones carried the left end arm of Y22 which contains the *Pgk-Neo* gene used for drug selection. Finally, the transfected DNA was stable even after long term passage in the absence of drug selection (data not shown). This is consistent with stable integration and faithful mitotic segregation of the transfected YAC clones.

The use of the Mov13 mutant cell line, which is defective in collagen I synthesis, as recipient cell was advantageous as it permitted us to analyse the expression of the transfected clones. As predicted from the DNA analyses, the two clones MC15 and MC44 synthesized normal amounts of the two proα1(I) mRNA species. These results indicate that a functional transcriptional unit for the *Col1a1* gene was present in the transfected clones and corroborate the conclusion that the transfection protocol allows the introduction and stable propagation of intact copies of YAC DNA in the chromosomes of mammalian cells.

A constraint of this protocol for efficient transfer of large DNA fragments into mammalian cells appears to be the low frequency of transfected cells which have taken up DNA, only a fraction of which carries an intact copy of the YAC DNA. The principal limitation of achieving a larger proportion of drug resistant cells upon transfection is the small quantity of YAC DNA which can be purified from transformed yeast cells. Amplification vectors such as those described by Smith *et al.* (1990), may be favorable as they would enrich the YAC DNA over the yeast chromosomes, permitting the use of a more optimal concentration of DNA in the transfection protocol. Our present procedure yields a fraction of 10% of the successfully transfected cells carrying an intact copy of a 150 kb large piece of DNA. We do not know whether this fraction would be lower with larger YAC DNAs.

The methodology described in this work should be suitable to a wide variety of applications aimed at genetically manipulating mammalian cells or organisms. Of particular interest will be the introduction of YAC-sized DNA clones into embryonic stem cells allowing the transfer of large segments of mammalian DNA into the germline of animals. Previous efforts at introduction of YAC clones have utilized total yeast DNA. In this complex mixture the YAC of interest represents a small fraction of the DNA available for transfection. Not surprisingly, several investigators (D'Urso *et al.*, 1990; Pavan *et al.*, 1990; Eliceiri *et al.*, 1991; Gnirke *et al.*, 1991) found large portions of the yeast genome present in their stable transfectants. This contaminating material could be mutagenic, and thus deleterious to the capacity of an ES cell to contribute to the germline of a chimeric mouse. It was for this reason that our efforts were directed at developing a transfection protocol utilizing highly purified YAC DNA. Employing this protocol, we are testing ES cells, containing intact YAC DNA, for their contribution to the germline. This approach will allow the molecular complementation of mutations with clones covering large segments of physically defined chromosomal regions. At present we are using the Mov13 mouse as a model system in an effort to establish the procedures for correcting a defined genetic defect by introducing a YAC clone covering the mutated gene into the germline. Preliminary experiments of transferring Y22 into embryonic stem cells have yielded encouraging results.

Materials and methods

YAC clones

Mus spretus was chosen as DNA source for construction of a *Sal*I complete digest YAC library from which a 150 kb clone was recovered spanning the murine α1(I) collagen locus (Strauss *et al.*, 1991). The YAC vector used for library construction was a modification of pYAC3 (12) called pYAC151, which has a neomycin resistance gene as a selectable marker in the *TRP1/ARS1* arm. This YAC clone (Y22) was recovered from a 32 000 clone library, and in detailed analysis showed excellent fidelity to *in vivo* organization (Strauss *et al.*, 1991).

Cell lines

Two fibroblastic cell lines were used for experiments: M13-5 is a line derived from the Mov13 strain and is homozygous for the defective α(1)I collagen

allele on the C57BL/6 background, it does not transcribe any detectable collagen (Schnieke *et al.*, 1983). The second line utilized was the fibroblastic cell line, NIH3T3.

Pulse field gel electrophoresis (PFGE)

All analytical and preparative electrophoresis was performed using a Bio-Rad CHEF (Chu *et al.*, 1986) DR II apparatus. Analytical gels utilized switching times from 10 to 70 s. Preparative gels used switching times from 10 to 40 s. The preparative gels were 0.6% Sea Plaque GTG grade agarose (FMC) in 0.5 × Tris−borate buffer.

YAC DNA isolation

S.cerevisiae containing the 150 kb collagen YAC was grown to saturation using uracil and tryptophan deficient liquid media (Sherman *et al.*, 1989). The saturated culture was washed in 1 M sorbitol−20 mM Tris pH 7.6, spheroplasted and then embedded in 1% Sea Plaque agarose at 40× the saturated growth density. The resulting plugs were lysed overnight according to previously published methods (Schwartz and Cantor, 1984; Strauss *et al.*, 1987). Plugs were aligned in a slot trough in an agarose gel and electrophoresed (see PFGE section). After electrophoresis, strips from either side of the gel were sliced and stained with EtBr to visualize the desired YAC clone. The YAC-containing gel strip was excised from the gel and dialyzed extensively for 36−48 h with 20 mM Tris−HCl/1 mM EDTA/ 50 μM spermine−HCl pH 7.6. The DNA gel strip was cut into sections weighing 0.2−0.6 g. The individual sections were loaded into a sterile 1.5 ml screw cap tube and heated to 65°C for 10 min. The liquid DNA gel mixture was equilibrated to 37°C and 15−30 U of agarase (Calbiochem) were added and mixed very gently. Digestion was complete in 1 h. This liquid YAC DNA was used immediately.

Transfection of mammalian cells

The liquid YAC DNA from a single 1.5 ml tube (0.2−0.6 g gel strip) was transferred to a 15 ml polystyrene tube with a large bore pipette, and 30 μg of DOTMA, supplied as the commercial product Lipofectin (Felgner *et al.*, 1987) was added to this mixture. After incubation at room temperature in a sterile environment for 15−30 min, 1.5−2 ml of serum free media was added and the whole mixture was poured carefully over a 6 cm tissue culture grade Petri dish containing a monolayer of M13-5 at 90−95% confluence. The transfection was allowed to continue for 8−12 h at 37°C 5% CO_2. At the end of the transfection course the supernatant was aspirated and replaced by 10% fetal calf serum/DMEM for another 12 h. The cells were then trypsinized and split to three 10 cm dishes. After 48 h selection was initiated with the addition of G418 to the culturing medium. Selection was continued for 14 days and all viable colonies were individually transferred into 12 well trays, and grown under selection. A single frozen stock was made from each well and analysis of clones was then initiated.

Isolation of DNA, RNA for analytic procedures

All methods were derived from previously published methods (Strauss *et al.*, 1987; Ausubel *et al.*, 1987, 1988, 1989).

DNA probes

Five probes were used in this analysis and are indicated in Figure 1. Lt corresponds to the 0.7 kb *Pvu*II−*Bam*HI fragment of pKJ-1, and contained half the *Neo* gene and the 3′ end of the *Pgk* gene (Adra *et al.*, 1987), or, for some experiments, only *Neo* gene sequences. The cDNA corresponds to a 4.8 kb full length transcript of the murine α1(I) collagen gene (unpublished results). Rt corresponds to the 1.7 kb *Bam*HI−*Pvu*II fragment of pBR322 which spans the *URA3*-containing side of the YAC vector called pYAC 151 (Strauss *et al.*, 1991). The rat α_1−tubulin full length cDNA probe served as an internal control (Lemischka *et al.*, 1981).

Acknowledgements

The authors would like to thank Drs D.Vollrath, P.Laird and J.Kriedberg for critical reading of the manuscript. This work was supported by NIH fellowship 5 F32 GM13756-02 and NIH grants 5 R35 CA44339-05 and HG000198-01.

References

Adra,C., Boer,P. and McBurney,M. (1987) Gene, **60**, 65−74.
Allshire,R.C., Cranston,G., Gosden,J.R., Maule,J.C., Hastie,N.D. and Fantes,P.A. (1987) *Cell*, **50**, 391−403.
Ausubel,F.M., Brent,R., Kingston,R.E., Moore,D.D., Seidman,J.G., Smith,J.A. and Struhl,K. (1987, 1988, 1989) *Current Protocols in Molecular Biology*.Greene-Wiley Interscience, New York.
Botstein,D., White,R.L., Skolnick,M. and Davis,R.R. (1980) *Am. J. Hum. Genet.*, **32**, 314−331.
Burke,D.T., Carle,G.F. and Olson,M.V. (1987) *Science*, **236**, 806−812.
Chu,G., Vollrath,D. and Davis,R.W. (1986) *Science*, **234**, 1582−1585.
Copeland,N.G. and Jenkins,N.A. (1991) *Trends Genet.*, **7**, 113−118.
Coulson,A., Sulston,J., Brenner,S. and Karn,J. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 7821−7825.
Coulson,A., Waterson,R., Kiff,J., Sulston,J. and Kohara,Y. (1988) *Nature*, **335**, 184−186.
Couto,L.B., Spangler,E.A. and Rubin,E.M. (1989) *Nucleic Acids Res.*, **17**, 8010−8013.
Dawson,D.S., Murray,A.W. and Szostak,J.W. (1986) *Science*, **234**, 713−717.
D'Urso,M., Zucchi,I., Ciccodicola,A., Palmieri,G., Abidi,F.E. and Schlessinger,D. (1990) *Genomics*, **7**, 531−534.
Eliceiri,B., Labella,T., Hagino,Y., Srivastava,A., Schlessinger,D., Pilia,G., Palmieri,G. and D'Urso,M. (1991) *Proc. Natl. Acad. Sci. USA*, **88**, 2179−3183.
Felgner,P.L., Gadek,T.R., Holm,M., Roman,R., Chan,H.W., Wenz,M., Northrop,J.P., Ringold,G.M. and Danielsen,M. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 7413−7417.
Garza,D., Ajioka,J.W., Burke,D.T. and Hartl,D.L. (1989) *Science*, **246**, 641−646.
Gitschier,J., Wood,W.I., Goralka,T.M., Wion,K.L., Chen,E.Y., Eaton,D.H., Vehar,G.A., Capan,D.J. and Lawn,R.M. (1984) *Nature*, **312**, 326−330.
Gnirke,A., Barnes,T.S., Patterson,D., Schild,D., Featherstone,T. and Olson,M.V. (1991) *EMBO J.*, **10**, 1629−1634.
Jaenisch,R., Harbers,K., Schnieke,A., Lohler,J., Chumakov,I., Jahner,D., Grotkopp,D. and Hoffman,E. (1983) *Cell*, **32**, 209−216.
Koenig,M., Hoffmann,E.P., Bertelson,C.J., Monaco,A.P. and Kunkle,L.M. (1987) *Cell*, **50**, 509−517.
Lemischka,I.R., Farmer,S., Racaniello,V.R. and Sharp,P.A. (1981) *J. Mol. Biol.*, **151**, 101−120.
Murry,A.W. and Szostak,J.W. (1983) *Nature*, **305**, 189−193.
Murry,A.W., Schultes,N.P. and Szostak,J.W. (1986) *Cell*, **45**, 529−536.
Nadeau,J.H., Herrmann,B., Bucan,M., Burkart,D., Crosby,J.L., Erhart,M.A., Kosowsky,M., Kraus,J.P., Michiels,F., Schnattinger,A., Tchetgen,M.B., Varnum,D., Willison,K., Lehrach,H. and Barlow,D. (1991) *Genomics*, **9**, 78−89.
Olson,M.V., Dutchik,J.E., Graham,M.Y., Brodeur,G.M., Helms,C., Frank,M., MacCollin,M., Scheinman,M.R. and Frank,T. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 7826−7830.
Pachnis,V., Pevny,L., Rothstein,R. and Costantini,F. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 5109−5113.
Pavan,W.J., Hieter,P. and Reeves,R.H. (1990) *Mol. Cell. Biol.*, **10**, 4163−4169.
Roberts,L. (1989) *Science*, **245**, 1438−1440.
Robertson,E.J. (1986) *Trends Genet.*, **2**, 9−13.
Schnieke,A., Harbers,K. and Jaenisch,R. (1983) *Nature*, **304**, 315−320.
Schnieke,A., Dziadek,M., Bateman,J.F., Mascara,T., Harbers,K., Gelinas,R. and Jaenisch,R. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 764−768.
Schwartz,D.C. and Cantor,C. (1984) *Cell*, **37**, 67−75.
Sherman,F., Fink,G.R. and Hicks,J.B. (1986) *Methods in Yeast Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Smith,D.R., Smyth,A.P. and Moir,D.P. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 8242−8246.
Strauss,W.M., Quertermous,T. and Seidman,J.G. (1987) *Science*, **237**, 1217−1219.
Strauss,W.M., Jaenisch,E. and Jaenisch,R. (1992) *Mammalian Genome*, in press.
Ward,M., Scott,R.J., Davey,M.R., Clothier,R.H., Cocking,E.C. and Balls,M. (1986) *Somat. Cell Mol. Genet.*, **12**, 101−109.
Watson,J.D. (1990) *Science*, **248**, 44−48.

Received on October 10, 1991; revised on November 5, 1991