Commentary

YAC transgenes: Bigger is probably better

Bernard G. Forget

Departments of Internal Medicine and Genetics, Yale University School of Medicine, New Haven, CT 06520-8021

A number of recent articles have reported the successful generation of transgenic mouse lines carrying large fragments of DNA contained in yeast artificial chromosomes (YACs) (1–5). The ability to transfer such large segments of DNA provides a number of opportunities, including the analysis of the function of large genes or gene clusters in their natural DNA configuration, the study of the influence of normally remote control elements on gene expression, and the analysis of long-range effects of regulatory elements such as those involved in X chromosome inactivation and imprinting. Other applications of YAC transgenes include the study of large DNA fragments for their ability to complement or rescue the loss of function associated with genetic defects in mice, and thereby confirm the identity or nature of genes isolated by positional cloning (reviewed in ref. 6). The ability to modify small regions of DNA within a YAC by homologous recombination also permits the study of the effect of mutations, such as short deletions, insertions, or nucleotide substitutions, on the function of a large gene or gene cluster.

Two different approaches have been utilized to obtain successful transfer of YACs into the mouse genome: direct microinjection of DNA into fertilized eggs, using procedures to avoid shearing or fragmentation of the DNA; or initial transfer of YAC DNA into embryonal stem (ES) cells by lipofection or spheroplast fusion, followed by injections of the transfected ES cells into blastocysts and production of chimeric mice. The second approach is somewhat more complicated and time consuming, requiring transmission of the transgene through the germ line before nonchimeric heterozygous mouse lines can be established, but it has the advantage that the integrity and other features of the transgene can be thoroughly characterized in the cultured ES cell line prior to the production of transgenic mice.

In a recent issue of the *Proceedings*, Peterson *et al.* (1) reported the successful transfer into transgenic mice of a 248-kb YAC containing the 82-kb human β-globin gene cluster and 148 kb of its flanking DNA. This was accomplished by direct microinjection into fertilized eggs. Technical modifications to avoid shearing and denaturation included use of high-ionic-strength solutions as a protective agent and use of ultrafiltration as a means to concentrate the DNA solutions. Of the pups born, 11.5% were transgenic. Of 17 founders, 13 contained both human γ- and human β-globin genes and 10 of these contained both left and right vector arm sequences. Nevertheless, 3 founders that retained both γ and β genes expressed the β-globin mRNA at very low levels in adult erythroid cells: <1% of the level of the endogenous mouse α-globin mRNA. These results suggest the possibility of some internal rearrangement or disruption of important regulatory sequences. The level of β-globin mRNA expression in the remaining founders with apparently intact YAC transgenes varied between 6% and 80% of the endogenous mouse α mRNA. The developmental pattern of expression of the different genes in the human β cluster was studied in the F_1 and F_2 progeny of one of the expressing lines and found to be correct; i.e., the embryonic ε-globin gene was expressed only in yolk sac-derived erythroid cells, and expression of the adult β-globin gene followed the same pattern as that of the mouse β^{major} gene in fetal liver and adult erythroid cells. Expression of the human γ gene followed the same pattern as that of its murine evolutionary homolog, the βh1 gene, being present both during the yolk sac and early fetal liver stages of erythropoiesis. The tissue-specific and developmentally correct expression of the transferred globin genes in the presence of YAC vector sequences is noteworthy in view of earlier studies indicating an adverse effect of plasmid vector sequences on tissue-specific expression of globin transgenes.

The study of Peterson *et al.* (1) is significant because it examined the function of a normally remote genetic regulatory element in its natural context. The 5' locus control region (LCR) is a region of DNA located ≈20 kb 5' to the ε-globin gene in the β cluster that was initially identified by the presence of four developmentally stable, erythroid-specific DNase I-hypersensitive sites (7, 8). When linked to a β-globin gene, the LCR confers position-independent high levels of expression of the gene in erythroid cells of transgenic mice (9). A large number of studies have examined the role of the LCR in the regulation of the developmental switches that occur in globin gene expression (reviewed in refs. 10–13). Significant differences in the developmental pattern of expression of various LCR/globin transgenes were obtained between different laboratories and it was argued that the differences were possibly due to the different structure of the transferred DNA fragments, which consisted of varying portions of the β gene cluster that frequently lacked substantial amounts of gene-flanking DNA and were ligated together in ways that did not preserve the normal distance between LCR and globin genes. The previously most "physiologic" transgene to be studied consisted of a 70-kb fragment derived by ligating together two cosmid inserts (14). This 70-kb fragment contained the entire 5' LCR in its natural spatial configuration to the rest of the β cluster, and the various β-like globin genes were expressed in a developmentally correct manner similar to that observed in mice carrying the YAC transgene. Peterson *et al.* point out that the YAC transgene has the advantage of containing more of the natural cluster than the 70-kb cosmid transgene, including a developmentally stable 3' hypersensitive site and extensive downstream DNA that is normally part of the general DNase I sensitive domain of the β-globin gene cluster. The ability to manipulate the insert of the YAC by using recombination to generate various mutations, such as the deletion of individual hypersensitive sites in the LCR, will most probably provide a great deal of important new information in the near future on the role of specific cis-acting elements in the regulation of globin gene expression during development.

Another successful YAC transfer into mice by direct microinjection of fertilized eggs was reported by Schedl *et al.* (2) using a 250-kb YAC containing the 80-kb mouse tyrosinase gene and 155 kb of its 5' flanking DNA. Modifications to standard DNA preparation procedures included the use of solutions containing high ionic strength, spermidine, and spermine, as well as dialysis using a floating dialysis filter.

Transgenic mouse lines were identified by rescue of the albino phenotype in

mutant NMRI recipient mice. Five of 24 newborns (21%) were pigmented, and only these 5 were positive for the presence of the transgene, including 4 that showed hybridization to probes from both vector arms. Two founders, containing one and two copies of the YAC, respectively, transmitted the transgene to F_1 offspring as a mendelian trait. A third founder containing eight copies of the YAC, including some apparent rearrangement, was a mosaic but did transmit the transgene to offspring. In the F_1 offspring of these 3 lines, a copy-dependent, position-independent level of normal tyrosinase gene expression was obtained that was comparable, by reverse transcription–PCR, to the level of endogenous mutant tyrosinase mRNA. This result suggests the presence in the transgene of regulatory elements analogous to the globin LCR. Previous experiments using tyrosinase minigenes have also resulted in rescue of the albino phenotype in transgenic mice, but with widely varying levels of tyrosinase gene expression, possibly due to position effects (15–17). In particular, a smaller, 35-kb YAC transgene containing a tyrosinase minigene with ≈5 kb of the tyrosinase gene 5′ flanking DNA was expressed less efficiently than in transgenic mice bearing the same DNA fragment that had been microinjected in the absence of vector sequences (16, 17).

Three other studies (3–5) report the production of mouse lines carrying functional YAC transgenes, by initial transfer of the YAC into ES cells by lipofection or spheroplast fusion followed by injection of selected transfected ES cells into blastocysts to generate chimeric mice. In two of the studies (3, 4), the YAC containing the inserted DNA was initially purified and separated from other yeast genomic DNA by pulsed-field gel electrophoresis, then transferred by lipofection. In the third study (5), spheroplast fusion was used resulting in the transfer of varying amounts of yeast genomic DNA, in addition to the YAC, into different transfected ES cells. Spheroplast fusion had been previously used to successfully transfer YACs and yeast DNA into mammalian tissue culture cells that expressed the gene incorporated into the YAC, but it was not previously known whether the yeast DNA would adversely affect the pluripotency of ES cells or the ability of the transferred DNA to be transmitted through the germ line.

Strauss et al. (3) used lipofection to transfer a 150-kb YAC containing the murine α1 collagen (Col1a1) gene and a selectable marker inserted into one of the YAC arms. The DNA was isolated in the presence of spermine, and poly(L-lysine) was added prior to cationic lipid. ES cells were exposed to the DNA–lipid complex in suspension rather than as a monolayer

in order to increase transfection efficiency, which was ≈30 drug-resistant clones per 10^8 transfected cells, a 30-fold improvement compared with previous protocols. Of 35 drug-resistant clones, 7 carried markers for the Col1a1 insert as well as for both YAC arms, 5 contained the Col1a1 insert and only the left YAC arm, and 23 contained only the left YAC arm bearing the selectable marker. Thus the majority of drug-resistant transfectants contained partially deleted YAC DNA. The 3 transfectants with nonrearranged YAC DNA were used to generate chimeric transgenic mice. Germ-line transmission was obtained from chimeric mice derived from 2 of the 3 cell clones. Quantitative analysis of collagen mRNA by RNase protection assays indicated that the level of transgene expression was similar to that of the endogenous gene in both mouse lines carrying the transgene at different chromosomal sites.

Choi et al. (4) also used lipofection to transfer into ES cells a YAC containing an 85-kb fragment of the human immunoglobulin heavy-chain gene containing at least one of each element required for correct rearrangement and expression of a human IgM molecule. In this study, the selectable marker was cotransfected on a separate DNA fragment. The efficiency of transfer was lower than in the study of Strauss et al. (3), which used a linked selectable marker: only 15 of 1221 drug-resistant cell clones contained the diagnostic heavy-chain gene fragments and only 3 of the 15 clones carried the intact YAC. The three ES cell clones with intact YAC DNA were injected into blastocysts to generate chimeric mice. Germ-line transmission was obtained from one of the founders, a 40% chimeric male who stably transmitted the ES cell genotype to 20 of 73 offspring. Transgene expression was analyzed by quantitation of human μ chain in serum by ELISA and was found to be relatively low in comparison to endogenous mouse IgM, perhaps due in part to competition from the endogenous gene: 10-fold higher levels of expression were obtained when the transgene was introduced into a mouse strain in which the endogenous heavy chain genes were inactivated. Other possible factors responsible for the relatively low level of expression include the lack of distant regulatory elements, the relatively small size of the transgene providing less of a "buffer" against position effects, and the choice of the single V_H6 segment in the transgene. Larger YACs containing the complete human immunoglobulin gene loci in the germ-line configuration could theoretically be used to generate transgenic mice capable of producing a large repertoire of human antibodies.

Jakobovits et al. (5) used spheroplast fusion to transfer a 670-kb YAC contain-

ing the human hypoxanthine phosphoribosyltransferase (HPRT) gene into HPRT-deficient ES cells. All of the ES cell clones selected by resistance to hypoxanthine/aminopterin/thymidine contained the human HPRT gene. Eight of 20 clones analyzed contained complete Alu profiles of the human DNA as well as both intact YAC vector arms. Due to the use of spheroplast fusion, variable amounts of yeast genomic DNA, in addition to the recombinant YAC, were transferred to the ES cells. Some ES cell clones contained a full complement of yeast DNA whereas others retained 20% or less of the yeast genome. In the one clone with a full complement of yeast DNA that was examined, all of the yeast DNA was integrated at a single site. The amount of integrated yeast DNA did not affect the in vitro differentiation pattern of the ES cells, and the YAC and yeast DNA were stably retained in the absence of selective pressure during 40 days in culture. The human HPRT gene remained functional as evidenced by the ability of the ES clones to grow and differentiate in the presence of hypoxanthine/aminopterin/thymidine. Eighteen chimeric mice were derived from three different ES cell clones. Chimeric males generated from an ES cell with the complete yeast genome faithfully transmitted the integrated DNA fragments to offspring which expressed human HPRT mRNA in multiple tissues, including the liver, where expression levels were comparable to that of the endogenous gene. Thus, the integration of large amounts of yeast DNA, in addition to the YAC itself, did not appear to adversely affect the pluripotency of the transfected ES cells or the functional properties of the transfected gene, in either ES cells or transgenic mice.

In summary, these various studies clearly demonstrate that it is now feasible to faithfully transfer very large segments of DNA into transgenic mice and that the transferred genes will usually be expressed at levels comparable to that of the corresponding endogenous gene, as well as in a tissue-specific and developmentally correct manner. The efficiency of transfer varies considerably depending on the technique used and the incidence of initial rearrangements following transfer appears to be high, requiring careful screening of the ES cell and mouse lines to confirm the structural integrity of the transferred DNA. Nevertheless, the ability to now study large genes or large gene clusters in transgenic mice without altering their natural DNA configuration should provide exciting insights into the regulation of gene expression in complex systems.

1. Peterson, K. R., Clegg, C. H., Huxley, C., Josephson, B. M., Haugen, H. S.,

Furukawa, T. & Stamatoyannopoulos, G. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 7593–7597.

2. Schedl, A., Montoliu, L., Kelsey, G. & Schütz, G. (1993) *Nature (London)* **362,** 258–261.

3. Strauss, W. M., Dausman, J., Beard, C., Johnson, C., Lawrence, J. B. & Jaenisch, R. (1993) *Science* **259,** 1904–1907.

4. Choi, T. K., Hollenbach, P. W., Pearson, B. E., Ueda, R. M., Weddell, G. N., Kurahara, C. G., Woodhouse, C. S., Kay, R. M. & Loring, J. F. (1993) *Nature Genet.* **4,** 117–123.

5. Jakobovits, A., Moore, A. L., Green, L. L., Vergara, G. J., Maynard-Currie, C. E., Austin, H. A. & Klapholz, S. (1993) *Nature (London)* **362,** 255–258.

6. Capecchi, M. R. (1993) *Nature (London)* **362,** 205–206.

7. Tuan, D., Solomon, W., Li, Q. & London, J. M. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 6384–6388.

8. Forrester, W. C., Thompson, C., Elder, J. T. & Groudine, M. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 1359–1363.

9. Grosveld, F., van Assendelft, G. B., Greaves, D. R. & Kollias, G. (1987) *Cell* **51,** 975–985.

10. Townes, T. M. & Behringer, R. R. (1990) *Trends Genet.* **6,** 219–223.

11. Stamatoyannopoulos, G. (1991) *Science* **252,** 383.

12. Epner, E., Kim, C. G. & Groudine, M. (1992) *Curr. Biol.* **2,** 262–264.

13. Dillon, N. & Grosveld, F. (1993) *Trends Genet.* **9,** 134–137.

14. Stouboulis, J., Dillon, N. & Grosveld, F. (1992) *Genes Dev.* **6,** 1857–1864.

15. Yokoyama, T., Silversides, D. W., Waymire, K. G., Kwon, B. S., Takeuchi, T. & Overbeek, P. A. (1990) *Nucleic Acids Res.* **18,** 7293–7298.

16. Beermann, F., Ruppert, S., Hummler, E., Bosch, F. X., Müller, G., Rüther, U. & Schütz, G. (1990) *EMBO J.* **9,** 2819–2826.

17. Schedl, A., Beermann, F., Thies, E., Montoliu, L., Kelsey, G. & Schütz, G. (1992) *Nucleic Acids Res.* **20,** 3073–3077.