Stable correction of a genetic deficiency in human cells by an episome carrying a 115 kb genomic transgene

Richard Wade-Martins[1,2]*, Robert E. White[1], Hiroshi Kimura[3], Peter R. Cook[3], and Michael R. James[1,4]*

*[1]Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, United Kingdom. [2]Current address: Department of Neurosurgery, Massachusetts General Hospital, Harvard Medical School, Room 6119, Building 149, 13th Street, Charlestown, MA 02129. [3]Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, United Kingdom. [4]Current address: Queensland Institute of Medical Research, 300 Herston Road, Q 4006, Brisbane, Australia. *Corresponding authors (michael.james@well.ox.ac.uk; michaelJ@qimr.edu.au; or wmartins@helix.mgh.harvard.edu).*

Received 14 March 2000; accepted 3 October 2000

Persistent expression of a transgene at therapeutic levels is required for successful gene therapy, but many small vectors with heterologous promoters are prone to vector loss and transcriptional silencing. The delivery of genomic DNA would enable genes to be transferred as complete loci, including regulatory sequences, introns, and native promoter elements. These elements may be critical to ensure prolonged, regulated, and tissue-specific transgene expression. Many studies point to considerable advantages to be gained by using complete genomic loci in gene expression[1–3]. Large-insert vectors incorporating elements of the bacterial artificial chromosome (BAC) cloning system[4], and the episomal maintenance mechanisms of Epstein–Barr virus (EBV), can shuttle between bacteria and mammalian cells, allowing large genomic loci to be manipulated conveniently[5]. We now demonstrate the potential utility of such vectors by stably correcting a human genetic deficiency in vitro. When the complete hypoxanthine phosphoribosyltransferase (*HPRT*) locus of 115 kilobases (kb) was introduced into deficient human cells, the transgene was both maintained as an episome and expressed stably for six months in rapidly dividing cell cultures. The results demonstrate for the first time that gene expression from an episomal genomic transgene can correct a cell culture disease phenotype for a prolonged period.

The flanking regulatory sequences, introns, and native promoter elements that make up a genomic locus, together with coding sequence, can span over 100 kb of chromosomal DNA[6], but most current gene delivery vectors are unable to accommodate such large inserts. A vector capable of routine manipulation of genomic DNA within human cells would be useful both for gene therapy and within the emerging field of functional genomics. Recently BACs containing complete genomic loci have been applied to functional gene expression studies in transgenic mice[7–9], and the use of a genomic locus provided appropriate levels of gene expression that was regulated in a tissue-specific manner. Episomes based on EBV (a 172 kb human herpesvirus) are an attractive vector system for gene expression[5,10–12]. Such episomes based on the BAC vector and incorporating the EBV latent origin of replication (*oriP*) elements necessary for nuclear retention have a large insert capacity (>100 kb), and are maintained stably in human cells in the presence of the viral protein EBV nuclear antigen-1 (EBNA-1)[5]. Such plasmids are shuttle vectors that can be routinely moved back and forth between human and bacterial cells. Rescue of the episome into bacterial cells from human cells allows any rearrangement to be easily assessed.

The aim of this study was to express the human HPRT enzyme from a large genomic DNA insert carried by an episomal vector. HPRT is a housekeeping enzyme expressed in all cells and catalyzes an early step in the purine salvage pathway in mammalian cells. The human *HPRT* locus lies at Xq26.1, and mutations within the gene cause the debilitating diseases of Lesch–Nyhan syndrome and gouty arthritis[13]. As a preliminary step, HP10, a *HPRT*$^-$ derivative of the Simian Virus 40 (SV40)-immortalized human male lung fibroblast cell line MRC-5V2, was transformed with pGEBNA (see Experimental Protocol), a small episomal vector designed to act as a helper plasmid with respect to expression of EBNA-1. This helper,

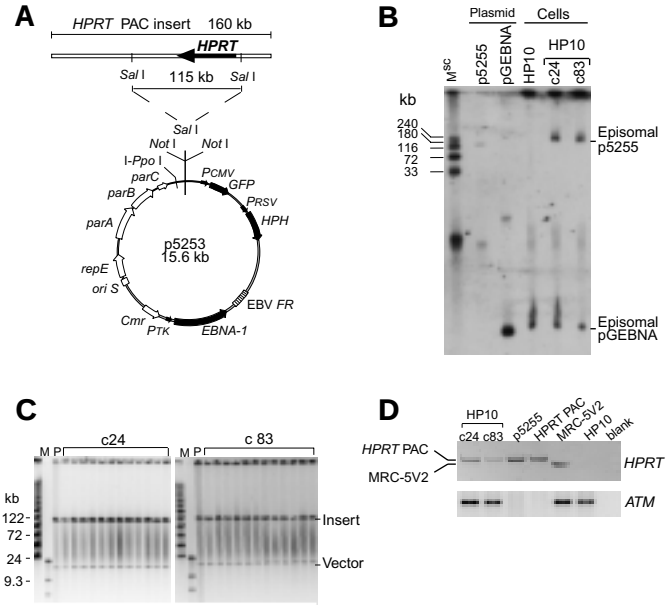

Figure 1. (A) Vector map. p5253 is a BAC-based vector incorporating two elements of EBV required for episomal maintenance: the family of repeats (*FR*) region from *oriP*, and an expression cassette for *EBNA-1*. A 115 kb *Sal*I fragment of genomic DNA was excised from PAC71G04 and inserted into p5253 at a unique *Sal*I site (flanked by *Not*I sites) to create p5255. (B–D) Genomic analysis of c24 and c83. (B) Analyzing intact episomes by electrophoresis and Southern blotting (using an *EBNA-1* probe). HP10 cells and the two transformants (c24 and c83) were lysed in situ in the gel, and supercoiled episomes resolved from linear chromosomal DNA. M^{SC}, size markers (supercoiled BACs bearing *EBNA-1*); p5255 and pGEBNA, vector and helper plasmid DNA. (C) Plasmid rescue. Episomal DNA from c24 or c83 was used to transform bacteria, and 12 of the resulting colonies selected were analyzed. After recovering plasmid DNA and cutting with *Not*I to release the insert, fragments were subjected to PFGE; the gel was then stained with ethidium bromide and photographed. Each lane contains DNA from a separate bacterial colony. M, Size markers; P, *Not*I fragments of p5255. (D) PCR genotype analysis. HUMHPRTB is a highly polymorphic tetranucleotide marker within the *HPRT* locus. Untransfected HP10 is negative for the marker, but the c24 and c83 have the genotype derived from the *HPRT* PAC locus, not of wild-type MRC-5V2. A STS marker from the ataxia–telangiectasia mutated (*ATM*) locus is a positive control for PCR amplification from the genomic DNA extracts.

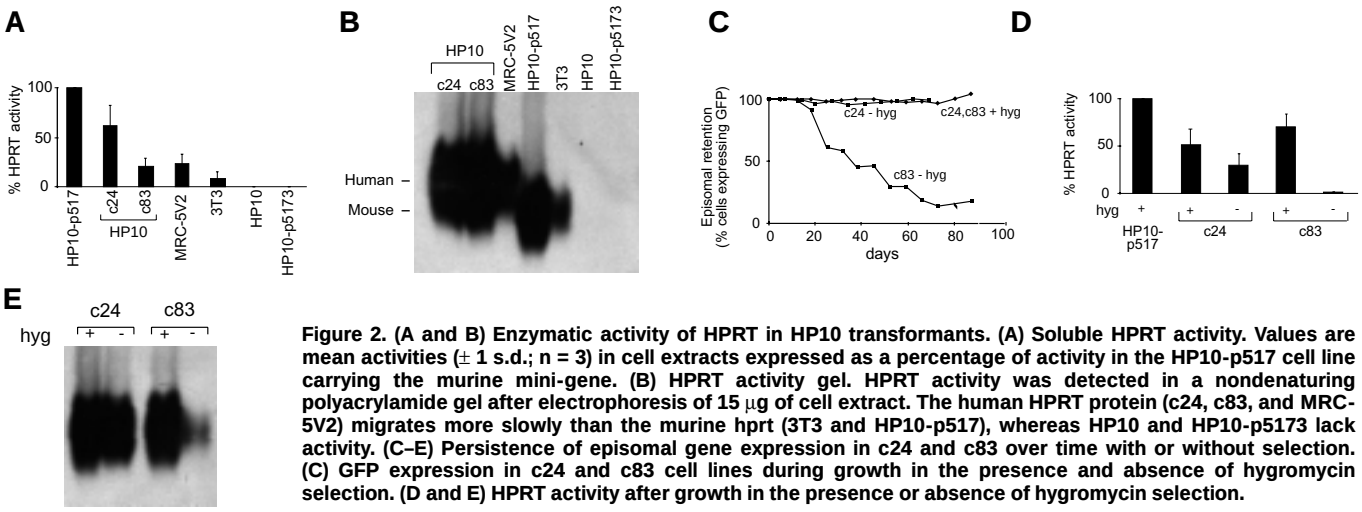

Figure 2. (A and B) Enzymatic activity of HPRT in HP10 transformants. (A) Soluble HPRT activity. Values are mean activities (± 1 s.d.; n = 3) in cell extracts expressed as a percentage of activity in the HP10-p517 cell line carrying the murine mini-gene. (B) HPRT activity gel. HPRT activity was detected in a nondenaturing polyacrylamide gel after electrophoresis of 15 μg of cell extract. The human HPRT protein (c24, c83, and MRC-5V2) migrates more slowly than the murine hprt (3T3 and HP10-p517), whereas HP10 and HP10-p5173 lack activity. (C–E) Persistence of episomal gene expression in c24 and c83 over time with or without selection. (C) GFP expression in c24 and c83 cell lines during growth in the presence and absence of hygromycin selection. (D and E) HPRT activity after growth in the presence or absence of hygromycin selection.

though not strictly necessary, boosts the efficiency of obtaining subsequent transformation by the large genomic episomes. The HP10 (EBNA-1$^+$) cells (hereafter referred to as HP10) were transfected with the vector containing the 115 kb *HPRT* genomic insert (p5255; Fig. 1A) using lipofectin/integrin-binding peptide/DNA (LID) complexes[14]. Green fluorescent protein (GFP) expression revealed a transfection efficiency of ~10–15%. Limiting-dilution cloning under combined hygromycin and G418 selection was used to obtain HP10-p5255 clonal cell lines. We used the Gardella gel technique[15] as a convenient method of rapidly screening a large number of clonal cell lines to identify those carrying episomal elements. Cells were lysed in situ in the slots of an agarose gel, and electrophoresis then resolved episomal DNA from chromosomal DNA (Fig. 1B). Two cell lines were identified that contained p5255 as an unrearranged intact episomal element, and named HP10 clones 24 and 83 (c24 and c83).

Approximately three months after the transfection, plasmid rescue from c24 and c83 recovered only unrearranged, intact p5255 episome (Fig. 1C). Genotype analysis using the HUMHPRTB marker (Fig. 1D), a highly polymorphic tetranucleotide repeat marker within the human *HPRT* locus[16], revealed that first, HP10 cells are negative for this marker, and second, the HUMHPRTB allele present in the c24 and c83 cell lines matches that found in *HPRT* PAC and in p5255 and is different from that of wild-type MRC-5V2 cells. The assay confirms that a reversion to the wild-type MRC-5V2 *HPRT* has not occurred in c24 and c83 cells. Total genomic DNA was prepared from c24 and c83 and digested by the restriction enzyme *I-Ppo*I, which linearizes p5255. Pulsed-field gel electrophoresis (PFGE) and Southern blotting studies revealed the 130 kb band expected for linearized episome (data not shown). The copy numbers of p5255 and pGEBNA were determined as 2 and 15, respectively, for c24; and 3 and 8, respectively, for c83 by quantitative Southern blotting (data not shown). As positive and negative controls for the *HPRT*$^-$ phenotype correction assays, two further transformants of HP10 were derived using p517, an episomal vector that carries the murine *hprt* mini-gene, and with p5173, an episomal vector that expresses the *HPH* and *GFP* genes only[5].

The HP10 derivative of the human fibroblast cell line MRC-5V2 used in this study has a deletion in the *HPRT* gene, resulting in the cells having no detectable HPRT enzymatic activity. Three different assays demonstrated the correction of the HPRT$^-$ cell phenotype in c24 and c83 cells. First, the two clonal lines were shown to survive continuous HAT (hypoxanthine/aminopterin/thymidine) selection for at least one month, whereas untransfected HP10 and HP10-p5173 cells died within 10 days of selection (data not shown). Second, strong HPRT activity (assayed by conversion of [^{3}H]hypoxanthine to [^{3}H]IMP) was found in protein extracts prepared from c24 and c83 cell but not in HP10 or HP10-p5173 cell lines (Fig. 2A). The levels of HPRT activity from the episomal genomic transgene do not differ greatly from the case of the native chromosomal locus. No HPRT activity was detected in the HP10 or HP10-p5173 cell lines. Third, an in situ gel activity assay confirmed that the HPRT activity was due to the expression of human HPRT (Fig. 2B). Human HPRT protein migrates slower than the murine hprt protein present in mouse 3T3 cells, and encoded by p517. Strong human HPRT activity was observed in the c24 and c83 cell lines, with no activity in the negative control HP10 or HP10-p5173 lines.

A characteristic feature of EBV-based episomes is that the episome is lost in exponentially dividing cells grown in the absence of selection at ~2–10% per cell generation[5,10]. Cells were grown under hygromycin selection for approximately three months following transfection, during which time all cells expressed GFP. Cells were then grown with or without hygromycin and the fraction of cells expressing GFP was monitored (Fig. 2C). The loss rate of p5255 for c83 was 4% per cell generation (using a first-order rate of loss model), in close agreement with our previous study[5]. At the end of the relaxation experiment when 10% of c83 cells retained episome according to GFP fluorescence, the HPRT activity had dropped sharply compared to the cell line grown continuously under selection (Fig. 2D, E). Southern blotting also showed that the p5255 copy number in c83 had fallen to 0.1 copies per cell (data not shown) after 90 days of growth in the absence of selection. This close correlation of episome retention and HPRT expression in c83 is consistent with phenotype correction solely by expression from the episomal genomic transgene. In contrast, all c24 cells expressed GFP after 70 days of growth in the absence of selection (Fig. 2C), and HPRT expression remained high (Fig. 2D, E). These data indicated that after prolonged continuous selection, cells containing an integrated copy of the transforming DNA may predominate within c24.

To investigate further the episomal status of p5255 in c24 and c83, we carried out fluorescent in situ hybridization (FISH) (Fig. 3). To overcome the problem of poor signal from closed circular episomal DNA[11,17], two probes were used in co-localization experiments. The *HPRT* PAC (green signal) was used as a positive control to identify the endogenous *HPRT* locus in HP10. (The precise X chromosome deletion in the HP10 cell line is not known, but enough flanking DNA is retained to hybridize to the PAC insert.) A vector-specific probe (pBAC108L; red signal) was used to identify vector DNA. The presence of p5255 was only recorded when both red and green signals overlapped. No co-localized signals were seen in 20 out of 21 metaphase spreads of the negative control HP10 cell line (Fig. 3A). In both HP10 p5255 clonal lines the copy number determined by FISH (Fig. 3B–E) was in close correlation with that found by quantitative Southern blotting.

Within both HP10-p5255 clonal lines we observe four types of red and green co-localized signal: (1) a single signal not associated with chromosomes; (2) a single signal closely associated with a host cell

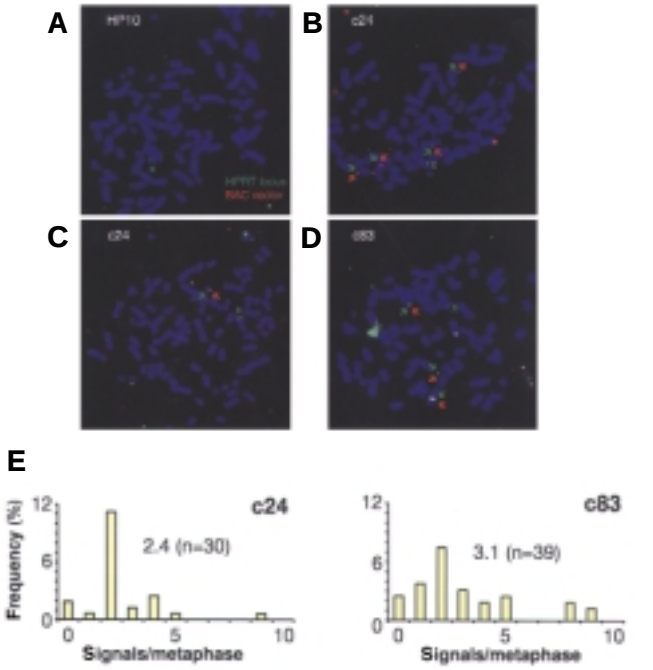

Figure 3. Fluorescent in situ hybridization analysis. Metaphase chromosome spreads were probed with *HPRT* PAC (green) and vector (pBAC108L; red), and DNA counterstained with TOTO-3 (blue). (A) HP10 (B and C) c24. (D) c83. Green X, genomic *HPRT* locus identified as a symmetrical green doublet on chromosome X long arm. Green and red arrows, p5255 identified as yellow spots (*HPRT*-BAC (red/green)-double positives). (E) p5255 copy number in c24 and c83 metaphase spreads (average number of signals per metaphase and the number of metaphases analyzed is shown).

chromosome; (3) two or more signals occurring in an asymmetrical pattern associated with the same chromosome; and (4) paired signals occurring in a symmetrical pattern. The first two forms clearly represent a single episomal element. The third and fourth forms are best interpreted as representing episomal constructs that have recently replicated, but not yet segregated at cell division[17]. In c83 the episome was seen only as an extrachromosomal element (Fig. 3D). In c83, paired signals occurring in a type 4 symmetrical pattern were very rarely seen at a similar chromosomal site, a pattern consistent with extrachromosomal status. In c24, in 12 of 30 metaphases, only episomal elements of types 1, 2, and 3 were seen (Fig. 3B). In 12 of the remaining metaphases, a paired symmetrical signal was seen at the end of a small chromosome (Fig. 3C). These data suggest that c24 has diverged into a mixed population of cells containing either solely episomes or solely integrated vector, although a low incidence of cells containing both were observed. With prolonged culture time under drug selection the cells with the integrated form predominate (Fig. 2C). Notably, however, HPRT expression from the integrated and the episomal transgenes are broadly comparable (Fig. 2D, E).

The vector described here has several advantages. It has a high capacity, it can shuttle between mammalian cells and bacteria, and it is maintained stably as an episome in mammalian cells. We used this episomal vector to introduce the native human *HPRT* locus into human HPRT$^-$ fibroblasts, and expression was demonstrated in three different ways. The episome copy number was low (≈2 copies per cell) and roughly equivalent to the diploid state. Similarly, enzyme levels were within physiological levels. This provides the first demonstration that a genomic locus carried on an episome can correct a genetic defect in vitro through production of sufficient levels of the required protein.

This episomal system has considerable potential. The vector can be manipulated in bacteria using current protocols, including those for mutagenesis and homologous recombination[7,8,18]. Microgram quantities of DNA can be prepared easily from the bacterial host, again using standard methods. Transfection using LID complexes can be optimized to obtain high transfection efficiencies of BAC-based constructs of up to 240 kb in size (R.E. White et al., unpublished data). Coupled with the stability of the episome, this means that many successful transformants are produced, and these can be easily identified by expression of GFP. Moreover, plasmids can be rescued in bacteria, so allowing the integrity of the transgene to be monitored.

We believe that our results show proof of principle for the use of the native form of a large gene to complement a human genetic defect, and that this may highlight an increasing appreciation of the benefits of such an approach. Furthermore, the EBV system has the potential for in vivo delivery of large genomic DNA inserts by packaging as infectious virions using helper-virus systems[19,20]. We have preliminary data showing packaging and infectious delivery of vectors with inserts of 60–150 kb of genomic DNA (R.E.White, R.Wade-Martins & M.R. James, unpublished data). Although EBV (a lymphotropic herpesvirus) has particular advantages, it acts as a paradigm for the wider herpesvirus family, which could provide a broader host cell tropism. Unlike other viral vectors, including adenovirus, the capacity of herpesvirus vectors makes the majority of genomic loci amenable to this approach. However, any safety concerns regarding the use of EBNA-1 (ref. 21) will need to be rigorously assessed before such episomal vectors can find clinical application.

Finally, this system offers a simple and rapid first approach to test the functional significance of the large number of gene-associated single-nucleotide polymorphisms[22] by site-specific recombination in bacteria followed by transfer to appropriate in vitro cell culture models. With the recent completion of the first draft of the human genome sequence, all genes will be available in BAC format, which will make this approach even more accessible.

Experimental protocol

Vector construction. The plasmid p5253 is based on the cloning vector pBAC108L (ref. 4) with the addition of four features for mammalian cell culture. First, the EBV *FR* region from the vector pHEBo (ref. 23); second, the hygromycin-B phosphotransferase (*HPH*) gene under the control of the Rous sarcoma expression (*PRSV*); third, the *EBNA-1* gene under the control of the HSV thymidine kinase promoter (*PTK*); and fourth, the green fluorescent protein (*GFP*) open reading frame from the plasmid pEGFP-N1 (Clontech Laboratories, Palo Alto, CA) driven by the cytomegalovirus immediate early promoter (*PCMV*). Large-insert subcloning was performed as described[5]. pGEBNA (data not shown) is based on p205 (ref. 23), with the *HPH* gene replaced by the Tn5 transposon G418 resistance gene; it acts as a helper plasmid to provide strong EBNA-1 expression to increase the efficiency of obtaining clonal cell lines. p517 is an episomal vector (M.R. James, unpublished data) expressing the murine *hprt* mini-gene[24]. p5173 is based on pHEBo (ref. 23) with the addition of the *PCMV-EGFP* cassette.

Tissue culture. HP10, a derivative of MRC-5V2 (an SV40-immortalized human male lung fibroblast cell line), contains a deletion of the *HPRT* locus on Xq26.1 (G. Dorado and A.R. Lehmann, personal communication). Cells were cultured at 37°C in 6% (vol/vol) CO_2 in Dulbecco's modified Eagle medium (Sigma, St. Louis, MO) supplemented with 10% fetal calf serum. A population of HP10 cells previously transformed with pGEBNA and maintained in 1 mg/ml G418 (Calbiochem, La Jolla, CA) was transfected with p5255 (purified on Qiagen Tip-100 columns; Qiagen USA, Valencia, CA) using LID complexes[14] as optimized for large DNA (R.E. White et al., unpublished data). Cells were transfected in a 6 cm dish with 2 μg of p5255, using the ratio 1 μg DNA : 0.75 μl lipofectin : 40 μl of Peptide 6. Clonal cell lines were obtained by limited-dilution cloning in the presence of hygromycin-B (Roche, East Sussex, UK) at 125 μg/ ml in combination with 1 mg/ml G418. Cell lines containing episomal p5255 were identified by episomal gel electrophoresis[15].

Plasmid rescue. Genomic DNA prepared from 5×10^6 cells was resuspended in 200 μl Tris–EDTA, and 1 μl of extract electroporated into DH-10B ElectroMax *Escherichia coli* cells (Gibco BRL, Grand Island, NY), which were then plated onto Luria–Bertani agar containing chloramphenicol (12.5 μg/ml). Plasmid DNA was prepared from the resulting bacterial colonies.

HPRT activity assays. A HPRT radioassay was employed[24]. 10^6 cells were

scraped into 1 ml of PBS without Mg^{2+} or Ca^{2+}, centrifuged at 5,000 *g* for 4 min, resuspended in 100 µl of lysis buffer (102 mM $Na_2HPO_4{\cdot}7H_2O$, 102 mM KH_2PO_4, pH 7.4) and lysed by four freeze-thawing cycles, centrifuged at 12,000 *g* for 10 min at 4°C, and the supernatant retained for activity assays. Reactions (50 µl) contained 0.625 mM $MgCl_2$, 13.9 mM NaCl, 0.1 mM EDTA, 7 mM Tris-HCl (pH 7.4), 1 mM phosphoribosyl pyrophosphate (PRPP), and 13 µM [8-^{3}H]hypoxanthine (2.3 Ci/mmol). At various times 10 µl was removed and added to 5 µl 100 mM EDTA on ice to halt the reaction and spotted onto a Whatman DE-81 filter disk. The disks were dried, washed twice in 10 mM ammonium formate (10 ml/disk) to precipitate the [^{3}H]IMP, and twice in absolute ethanol (10 ml/disk). The disks were dried, and tritium incorporation was measured by liquid scintillation. Enzymatic activity of HPRT was measured by fluorography in polyacrylamide gels as described[25].

Fluorescent in situ hybridization. *HPRT* PAC71G04 and pBAC108L (ref. 4) were labeled with biotin using BioNick labeling system (Gibco BRL) and digoxigenin using Dig-Nick Translation Mix (Roche), respectively. After hybridization and washing, the glass slide was incubated with blocking buffer (PBST containing 1% BSA, 0.2% fish gelatin, and 25 mM glycine; pH 8) at room temperature for 30 min, and with 2 µg/ml Cy3-conjugated mouse anti-dig antibody (Jackson ImmunoResearch, West Grove, PA) and 10 µg/ml Alexa 488-conjugated NeutrAvidin (Molecular Probes, Eugene, OR) for 1 h and then washed 4× with PBST over 1 h. Nucleic acids were stained with 50 nM TOTO-3 iodide (Molecular Probes) in PBS at room temperature for 5 min. After final washing with PBS for 5 min, the sample was mounted in Vectashield (Vector Laboratories, Burlingame, CA). Images were obtained using a Bio-Rad MRC1000 confocal laser scanning microscope (Bio-Rad, Hercules, CA) equipped with an Argon/Krypton laser and coupled to a Nikon Diaphot 200 inverted microscope.

Acknowledgments

We would like to thank Drs José Mejía and Zoia Larin for providing and analyzing PAC71G04, Dr. Jon Frampton for sharing his flow cytometery expertise, and Dr. Steve Hart for his assistance with the LID complex transfection. We thank the Wellcome Trust for support. R.W-M. is a Wellcome Trust Prize Fellow; H.K. is supported by the E.P. Abraham Trust and the Wellcome Trust.

1. Manson, A.L. et al. Complementation of null CF mice with a human *CFTR* YAC transgene. *EMBO J.* **16**, 4238–4249 (1997).
2. Schedl, A. et al. Influence of *PAX6* gene dosage on development: overexpression causes severe eye abnormalities. *Cell* **86**, 71–82 (1996).
3. Schiedner, G. et al. Genomic DNA transfer with a high-capacity adenovirus vector results in improved *in vivo* gene expression and decreased toxicity. *Nat. Genet.* **18**, 180–183 (1998).
4. Shizuya, H. et al. 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* **89**, 8794–8797 (1992).
5. Wade-Martins, R., Frampton, J. & James, M.R. Long-term stability of large insert genomic DNA episomal shuttle vectors in human cells. *Nucleic Acids Res.* **27**, 1674–1682 (1999).
6. Li, Q., Harju, S. & Peterson, K.R. Locus control regions—coming of age at a decade plus. *Trends Genet.* **15**, 403–408 (1999).
7. Yang, X.W., Model, P. & Heintz, N. Homologous recombination based modification in *Escherichia coli* and germline transmission in transgenic mice of a bacterial artificial chromosome. *Nat. Biotechnol.* **15**, 859–865 (1997).
8. Yang, X.W., Wynder, C., Doughty, M.L. & Heintz, N. BAC-mediated gene-dosage analysis reveals a role for Zipro1 (Ru49/Zfp38) in progenitor cell proliferation in cerebellum and skin. *Nat. Genet.* **22**, 327–335 (1999).
9. Antoch, M.P. et al. Functional identification of the mouse circadian *clock* gene by transgenic BAC rescue. *Cell* **89**, 655–667 (1997).
10. Wohlgemuth, J.G., Kang, S.H., Bulboaca, G.H., Nawotka, K.A. & Calos, M.P. Long-term gene expression from autonomously replicating vectors in mammalian cells. *Gene Ther.* **3**, 503–512 (1996).
11. Simpson, K., McGuigan, A. & Huxley, C. Stable episomal maintenance of yeast artificial chromosomes in human cells. *Mol. Cell. Biol.* **16**, 5117–5126 (1996).
12. Calos, M.P. The potential of extrachromosomal replicating vectors for gene therapy. *Trends Genet.* **12**, 463–466 (1996).
13. Caskey, C.T. & Kruh, G.D. The HPRT locus. *Cell* **16**, 1–9 (1979).
14. Hart, S.L. et al. Lipid-mediated enhancement of transfection by a nonviral integrin targeting vector. *Hum. Gene Ther.* **9**, 575–585 (1998).
15. Gardella, T., Medveczky, P., Sairenji, T. & Mulder, C. Detection of circular and linear herpesvirus DNA molecules in mammalian cells by gel electrophoresis. *J. Virol.* **50**, 248–254 (1984).
16. Edwards, A., Hammond, H.A., Jin, L., Caskey, T.C. & Chakraborty, R. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics* **12**, 241–253 (1992).
17. Delecluse, H.J., Bartnizke, S., Hammershmidt, W., Bullerdiek, J. & Bornkamm, G.W. Episomal and integrated copies of Epstein–Barr virus coexist in Burkitt Lymphoma cell lines. *J. Virol.* **67**, 1292–1299 (1993).
18. Narayanan, K., Williamson, R., Zhang, Y., Stewart, A.F. & Ioannou, P.A. Efficient and precise engineering of a 200 kb β-globin human/bacterial artificial chromosome in *E. coli* DH10B using an inducible homologous recombination system. *Gene Ther.* **6**, 442–447. (1999).
19. Banerjee, S., Livanos, E. & Vos, J.-M.H. Therapeutic gene delivery in human B-lymphoblastoid cells by engineered non-transformng infectious Epstein–Barr virus. *Nat. Med.* **1**, 1303–1308 (1995).
20. Hammerschmidt, W. & Sugden, B. Identification and characterization of *oriLyt*, a lytic origin of DNA replication of Epstein–Barr virus. *Cell* **55**, 427–433 (1988).
21. Wilson, J.B., Bell, J.L. & Levine, A.J. Expression of Epstein–Barr virus nuclear antigen-1 induces neoplasia in transgenic mice. *EMBO J.* **15**, 3117–3126 (1996).
22. Glaser, V. IBM joins SNP consortium. *Nat. Biotechnol.* **18**, 245 (2000).
23. Yates, J.L., Warren, N. & Sugden, B. Stable replication of plasmids derived from Epstein–Barr virus in various mammalian cells. *Nature* **313**, 812–815 (1985).
24. Melton, D.W., McEwan, C., McKie, A.B. & Reid, A.M. Expression of the mouse HPRT gene: deletional analysis of the promoter region of an X-chromosome linked housekeeping gene. *Cell* **44**, 319–328 (1986).
25. Yee, J.K. et al. Epitope insertion into the human hypoxanthine phosphoribosyltransferase protein and detection of the mutant protein by an anti-peptide antibody. *Gene* **53**, 97–104 (1987).

DNA cloning by homologous recombination in *Escherichia coli*

Youming Zhang, Joep P.P. Muyrers, Giuseppe Testa, and A. Francis Stewart*

Gene Expression Program, European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69117 Heidelberg, Germany.
**Corresponding author (stewart@embl-heidelberg.de).*

Received 27 July 2000; accepted 10 October 2000

The cloning of foreign DNA in *Escherichia coli* episomes is a cornerstone of molecular biology. The pioneering work in the early 1970s, using DNA ligases to paste DNA into episomal vectors, is still the most widely used approach. Here we describe a different principle, using ET recombination[1,2], for directed cloning and subcloning, which offers a variety of advantages. Most prominently, a chosen DNA region can be cloned from a complex mixture without prior isolation. Hence cloning by ET recombination resembles PCR in that both involve the amplification of a DNA region between two chosen points. We apply the strategy to subclone chosen DNA regions from several target molecules resident in *E. coli* hosts, and to clone chosen DNA regions from genomic DNA preparations. Here we analyze basic aspects of the approach and present several examples that illustrate its simplicity, flexibility, and remarkable efficiency.

Recently we described a new way to use homologous recombination for DNA engineering in *E. coli*. We termed the approach "ET recombination" because we first uncovered it using the Rac phage protein pair, RecE/RecT, and then later showed that the equivalent lambda phage protein pair, Redα/Redβ, also worked[1,2]. Our original goal was to develop a simple method to engineer bacterial artificial chromosomes (BACs)[1–4]. However, the approach worked so well that other applications became apparent, such as rapid creation of new *E. coli* strains through direct targeting of the *E. coli* chromosome[1], rapid generation of gene targeting constructs for use in mouse embryonic stem (ES) cells[5], and a new way to accomplish site-directed mutagenesis in plasmids[1] and BACs[6,7]. Recently three other groups, aiming to develop a better way to directly target the *E. coli* chromosome, have also demonstrated that ET recombination initiat-