npg

Open

ORIGINAL ARTICLE

Refined human artificial chromosome vectors for gene therapy and animal transgenesis

Y Kazuki[1,2], H Hoshiya[1], M Takiguchi[1], S Abe[1], Y Iida[1], M Osaki[1], M Katoh[3], M Hiratsuka[4], Y Shirayoshi[5], K Hiramatsu[1], E Ueno[1], N Kajitani[1], T Yoshino[6], K Kazuki[1], C Ishihara[1], S Takehara[1], S Tsuji[1], F Ejima[7,10], A Toyoda[7,10], Y Sakaki[7,11], V Larionov[8], N Kouprina[8] and M Oshimura[1,2,9]

Human artificial chromosomes (HACs) have several advantages as gene therapy vectors, including stable episomal maintenance, and the ability to carry large gene inserts. We previously developed HAC vectors from the normal human chromosomes using a chromosome engineering technique. However, endogenous genes were remained in these HACs, limiting their therapeutic applications. In this study, we refined a HAC vector without endogenous genes from human chromosome 21 in homologous recombination-proficient chicken DT40 cells. The HAC was physically characterized using a transformation-associated recombination (TAR) cloning strategy followed by sequencing of TAR-bacterial artificial chromosome clones. No endogenous genes were remained in the HAC. We demonstrated that any desired gene can be cloned into the HAC using the Cre-loxP system in Chinese hamster ovary cells, or a homologous recombination system in DT40 cells. The HAC can be efficiently transferred to other type of cells including mouse ES cells via microcell-mediated chromosome transfer. The transferred HAC was stably maintained *in vitro* and *in vivo*. Furthermore, tumor cells containing a HAC carrying the suicide gene, *herpes simplex virus thymidine kinase* (*HSV-TK*), were selectively killed by ganciclovir *in vitro* and *in vivo*. Thus, this novel HAC vector may be useful not only for gene and cell therapy, but also for animal transgenesis.
Gene Therapy (2011) **18**, 384–393; doi:10.1038/gt.2010.147; published online 18 November 2010

INTRODUCTION

Conventional gene transfer techniques using viruses, plasmids, P1 phage-derived artificial chromosomes (PACs), bacterial artificial chromosomes (BACs) and yeast artificial chromosomes (YACs), can insert DNA randomly into the host genome,[1] possibly causing cancer in humans and unpredicted expression in transgenic animals.[2–5] Human artificial chromosomes (HACs) exhibit several important characteristics desired for an ideal gene delivery vector, including stable episomal maintenance, and the capacity to carry large genomic loci with their regulatory elements, thus allowing physiological regulation of the introduced gene in a manner similar to that of the native chromosome.[6] HACs have been generated using either a 'top–down approach' (engineered chromosome), or a 'bottom–up' approach (*de novo* artificial chromosome).[7] Although several groups have reported functional analyses *in vitro* using *de novo* HACs, several factors limit the application of *de novo* generated HACs as gene delivery vehicles. The most critical problem is their undefined structure and the unpredictable relationship between the input DNA and resultant HAC, especially in terms of their size and composition.[8–10] On the other hand, several groups have reported the creation of engineered HACs by random segmentation or targeted telomere-associated chromosomal fragmentation in homologous recombination-proficient chicken DT40 cells.[11–14] We previously developed HAC vectors from normal human chromosome 14 (hChr.14) or 21 (hChr.21) by the engineering approach, and named the products SC20-HAC and 21ΔqHAC/21ΔpqHAC.[13,14] The SC20-HAC was transmittable through the germline and was rather stable in mice and cattle.[13,15–17] Both 21ΔqHAC and 21ΔpqHAC were very stable in human cell lines.[14,18,19] However, SC20-HAC and 21ΔqHAC/21ΔpqHAC contain several structurally undefined regions carrying many endogenous genes, which cause partial trisomy in cells propagating these HACs. This may affect physiological gene expression and normal development. Although these HACs showed significant potential for gene therapy and animal transgenesis, the ideal gene delivery vector should be structurally defined and should not contain endogenous genes from the original chromosome. In this study, we developed a novel HAC vector of known sequence containing no endogenous genes using the top–down approach. We also developed several gene insertion systems on the HAC for functional analysis of multiple genes, safe human gene therapy and efficient animal transgenesis.

RESULTS

Construction of 21HAC1

Previously, we developed a HAC vector from normal hChr.21 by a top–down approach using sequence information from hChr.21.[14,20] However, several transcripts were identified on the previously developed 21ΔqHAC/21ΔpqHAC.[21] Thus, we have attempted to construct a HAC containing no endogenous genes from hChr.21 (Figure 1). Truncation of hChr.21 and insertion of loxP into hChr.21 was carried out based on a new information on the structure of pericentromeric regions of the hChr.21.[21] We developed new vectors with targeting sequences from contigs (AP001657 and AL163201) that are the most proximal to the centromeric alphoid DNA array. Although the pericentromeric sequences of AP001657 and AL163201 are not repetitive and mostly unique, the nature of these pericentromeric sequences has not been reported. As we cannot confirm where the targeting construct was inserted if the repetitive alpha satellite sequence is used for the targeting, we utilized the known and unique sequence for the construction of the targeting vector. Such a strategy allowed us to construct a HAC lacking any endogenous genes. The 21ΔqHAC/21ΔpqHAC contained a 3′neo-loxP site for cloning a desired gene by reconstitution of the neo cassette. In this study, 5′HPRT-loxP was used to clone a desired gene by reconstitution of the HPRT cassette, because the neo gene frequently has been used for gene-targeting and chromosome-tagging in A9 cell libraries containing a single human chromosome.[22,23] As DT40 cells exhibit a high frequency of homologous recombination between exogenous DNA templates and their chromosomal counterparts, DT40 cells containing hChr.21 tagged with pSTneo were used for modification of hChr.21. A schematic diagram of the construction of the HAC and its map are shown in Figures 1a and f, respectively. With the targeting construct, 5′HPRT-loxP-Hyg-TK, for the cloning site (loxP) and negative selection (herpes simplex virus thymidine kinase (HSV-TK)), 46 of 237 drug-resistant clones selected in the presence of both hygromycin and

G418 were targeted correctly, as shown by PCR and Southern blot analyses (designated as hChr.21-loxP) (Supplementary Figures S1a and b). Fluorescence in situ hybridization (FISH) analyses showed that the targeting construct was integrated into the hChr.21 in DT40 cells (Figures 1b and c). With the targeting construct, pBS-TEL/Δp Puro, for the deletion of the p-arm of hChr.21, 3 of 206 drug-resistant clones selected in the presence of both puromycin and hygromycin were targeted correctly, as shown by PCR (designated as hChr.21-loxPΔp) (Supplementary Figure S1c). FISH analyses showed that the targeting construct was integrated into the hChr.21-loxP in DT40 cells (Figure 1d). With the targeting construct, pBS-TEL/Δq HisD, for the deletion of the q-arm of hChr.21, 2 of 95 drug-resistant clones selected in the presence of L-histidinol, puromycin and hygromycin were targeted correctly, as shown by PCR (designated as hChr.21-loxPΔpq or 21HAC1) (Supplementary Figure S1d). FISH analyses showed that the targeting construct was integrated into the hChr.21-loxPΔp in DT40 cells (Figure 1e). These data show that a HAC vector devoid of endogenous genes has been developed. Next, we tested whether desired gene(s) could be cloned into the HAC using several methods (Figure 2). Detailed genetic maps of the 21HAC1, 21HAC2, 21HAC3 and 21HAC4 in the following analyses, were shown in Figure 3.

Gene cloning into 21HAC1 by homologous recombination

To simplify further analysis of mitotic stability of the HAC, the fluorescent markers, enhanced green fluorescent protein (EGFP) and DsRed, were introduced into the 21HAC1 using homologous recombination in chicken DT40 cells. Homologous recombination in DT40 cells has been used for the modification of hChr.21 as described above, and the Cre-loxP system has been used for gene insertion into the 21ΔqHAC/21ΔpqHAC.[14] In this study, we used homologous recombination for sequential gene cloning into the 21HAC1 in DT40 cells, as
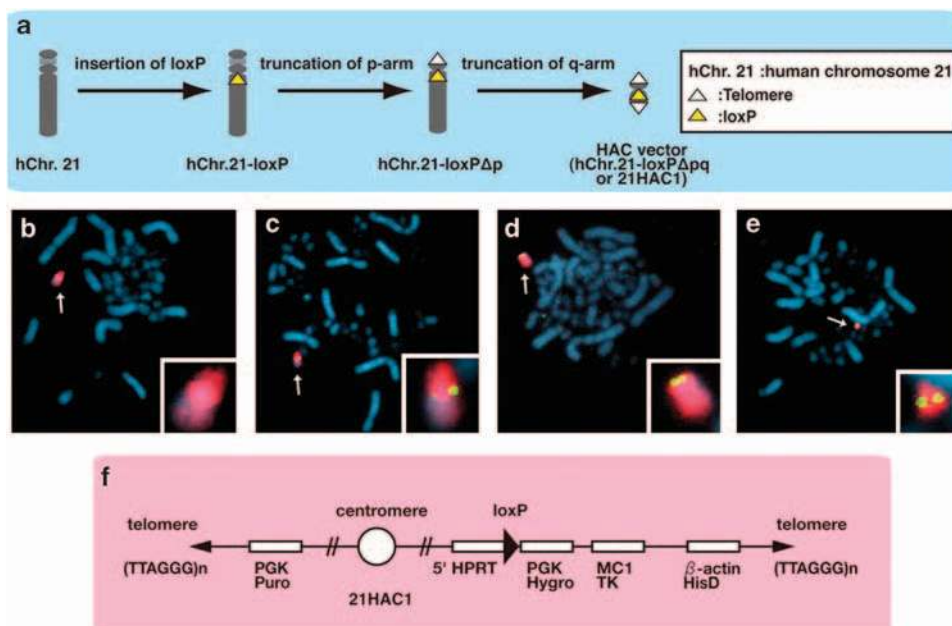
Figure 1 Construction of 21HAC1 from hChr.21. (a) Strategy for construction of 21HAC1 from hChr.21. (b–e) FISH analyses for DT40 (hChr.21) (b), DT40 (hChr.21-loxP) (c), DT40 (hChr.21-loxPΔp) (d) and DT40 (hChr.21-loxPΔpq/21HAC1) (e). Digoxigenin-labeled human COT-1 DNA (red) was used to detect the HAC. Biotin-labeled PGK-Hyg, PGK-Puro and β-actin-HisD (green) were used to detect the marker gene on the HAC in c, d and e, respectively. Chromosomal DNA was counterstained with 4,6-diamidino-2-phenylindole (DAPI). The insets show an enlarged image of the modified hChr.21 (HAC) (arrow). (f) Map of the 21HAC1 vector.

Figure 2 Schematic diagram of cloning genes into the 21HAC vector. (**a**) Homologous recombination-type cloning (sequential gene insertion). The desired gene can be sequentially cloned into a specific site on the HAC in DT40 cells by homologous recombination. The 21HAC2 can be transferred to the hiMSC, mES and HT1080 for each experiment via CHO (hprt−/−)(21HAC2) cells. (**b**) Insertion-type cloning-1 (HPRT version). The 21HAC1 can be transferred to CHO (hprt−/−) 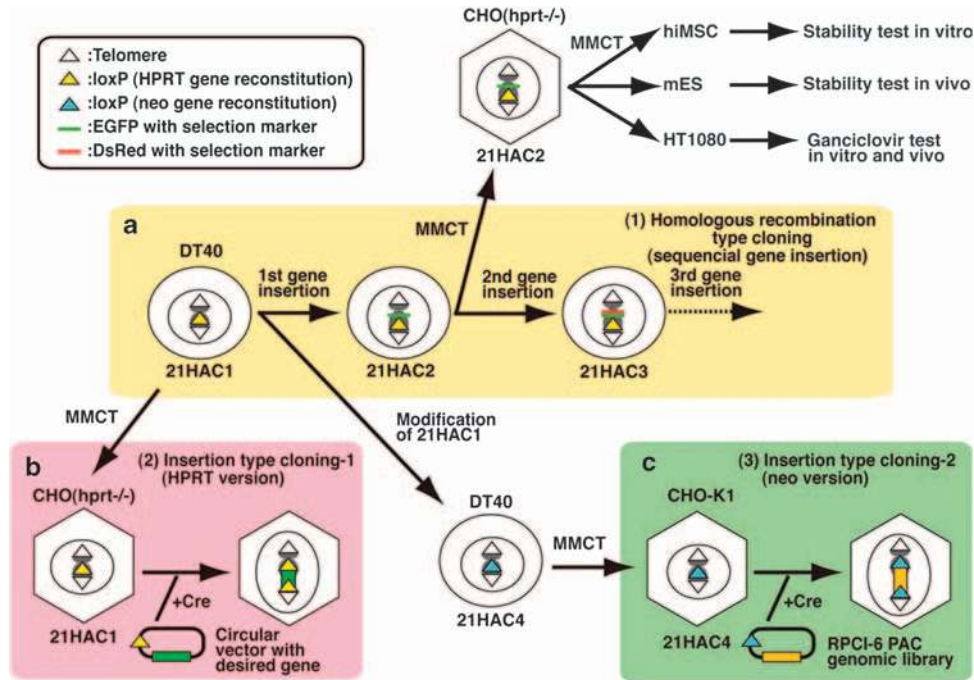cells. A circular vector can be cloned into the HAC in CHO (hprt−/−) cells by Cre-loxP-mediated gene insertion with reconstitution of the *HPRT* gene. (**c**) Insertion-type cloning-2 (neo version). The 21HAC4 modified from 21HAC1 can be transferred to CHO-K1 cells. A circular vector can be cloned into the HAC in CHO-K1 cells by Cre-loxP-mediated gene insertion with reconstitution of the neo gene.
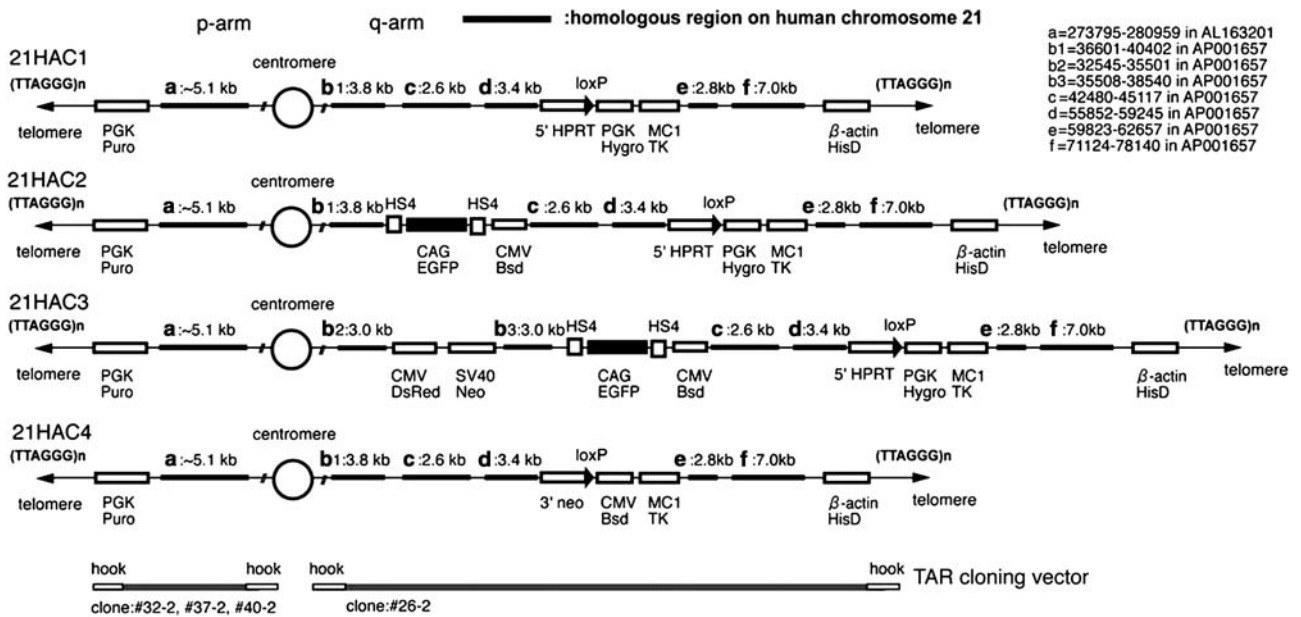
Figure 3 Detailed map of 21HACs. (**a–f**) shows homologous regions on hChr.21 for the targeting. The targeting vector, I-EGFP-I-Bsd, DsRed-neo and 3′neo-loxP-Bsd-TK were inserted into the 21HAC1, 21HAC2 and 21HAC1, respectively. The information of the targeting vectors was described in Supplementary Figure S1–S4. The hook for the TAR cloning vector is described in the loci of the 21HAC4.

well as for its modification (Figure 2a). First, to insert the *EGFP* gene into the 21HAC1 as a model, a targeting construct (designated as I-EGFP-I-Bsd) with homologous arms, the *EGFP* flanked by HS4 insulator and *blasticidin* genes, was inserted in the proximal region

of the targeting site of 5′HPRT-loxP-Hyg-TK (Supplementary Figure S2a). Since we previously succeeded tissue-specific gene expression on the 21ΔpqHAC using chicken-derived HS4 insulator to block the promoter interference, we also utilized the HS4 for gene expression on

the 21HAC1.[18] With the targeting construct, I-EGFP-I-Bsd, 11 of 47 drug-resistant clones selected in the presence of both blasticidin S and hygromycin were targeted correctly, as proven by PCR (this HAC was designated as 21HAC2). FISH analyses showed that the targeting construct was integrated into the 21HAC1 in DT40 cells (Figure 4a). Most of the cells were positive for EGFP, showing that the HAC was retained stably and the *EGFP* gene on the HAC was expressed stably in DT40 cells (Figure 4g). Second, to insert the *DsRed* gene into the 21HAC2 as a model, the targeting constructs (designated as DsRed-neo) with the homologous arms, *DsRed* gene and *neomycin* (*neo*) gene, were inserted in the proximal region of the targeting site of the I-EGFP-I-Bsd (Supplementary Figure S2b). With the targeting construct, DsRed-neo, 34 of 79 drug-resistant clones selected in the presence of both blasticidin S and G418 were targeted correctly, as shown by PCR (designated as 21HAC3). FISH analyses showed that the targeting construct was integrated into the 21HAC2 in DT40 cells (Figure 4b). Most of the cells were positive for DsRed, showing that the HAC was retained stably and the *DsRed* gene on the HAC expressed stably in DT40 cells (Figure 4g). These data suggest that multiple genes can be inserted into the HAC using the homologous recombination approach, and that the expression of these genes was stable. Next, we tested gene cloning into the HAC by Cre-loxP system (Figures 2b and c).

Gene cloning into 21HAC1 by Cre-loxP system (HPRT version)

In this study, Cre-loxP-mediated gene insertion into the 21HAC1 in DT40 cells using *HPRT* gene reconstitution could not be used, because we used hprt-positive DT40 cell lines. Therefore, the 21HAC1 was transferred to hprt-deficient Chinese hamster ovary (CHO) (hprt−/−) cells via microcell-mediated chromosome transfer (MMCT). PCR analyses with 21HAC1-specific primers showed that 22 out of the 38 drug-resistant clones were positive for 21HAC1 (data not shown). FISH analysis showed that the 21HAC1 was successfully transferred into the CHO (hprt−/−) cells (Figure 4c). To test whether a circular vector with a desired gene can be cloned into the 21HAC1 in CHO (hprt−/−), a Cre-expression vector and a plasmid vector containing the *EGFP* flanked by HS4 insulator, loxP and 3′HPRT (I-EGFP-I-loxP-3′HPRT), were co-transfected into CHO (hprt−/−) cells containing the 21HAC1, and recombinant clones were selected using hypoxanthine - aminopterin - thymidine (HAT) over a period of 10 days (Supplementary Figure S3). In all, 4 out of 12 drug-resistant clones were positive by PCR with GFP-specific primers and were positive for GFP (Figure 4h), suggesting that the remaining 8 clones have some rearrangements or deletions for the GFP expression. FISH analyses showed that the construct was inserted into the 21HAC1 in CHO (hprt−/−) cells (Figure 4d). This ratio (4 of 12, or 33%) of circular vector insertion into the 5′HPRT-loxP site on the 21HAC1 by *HPRT*
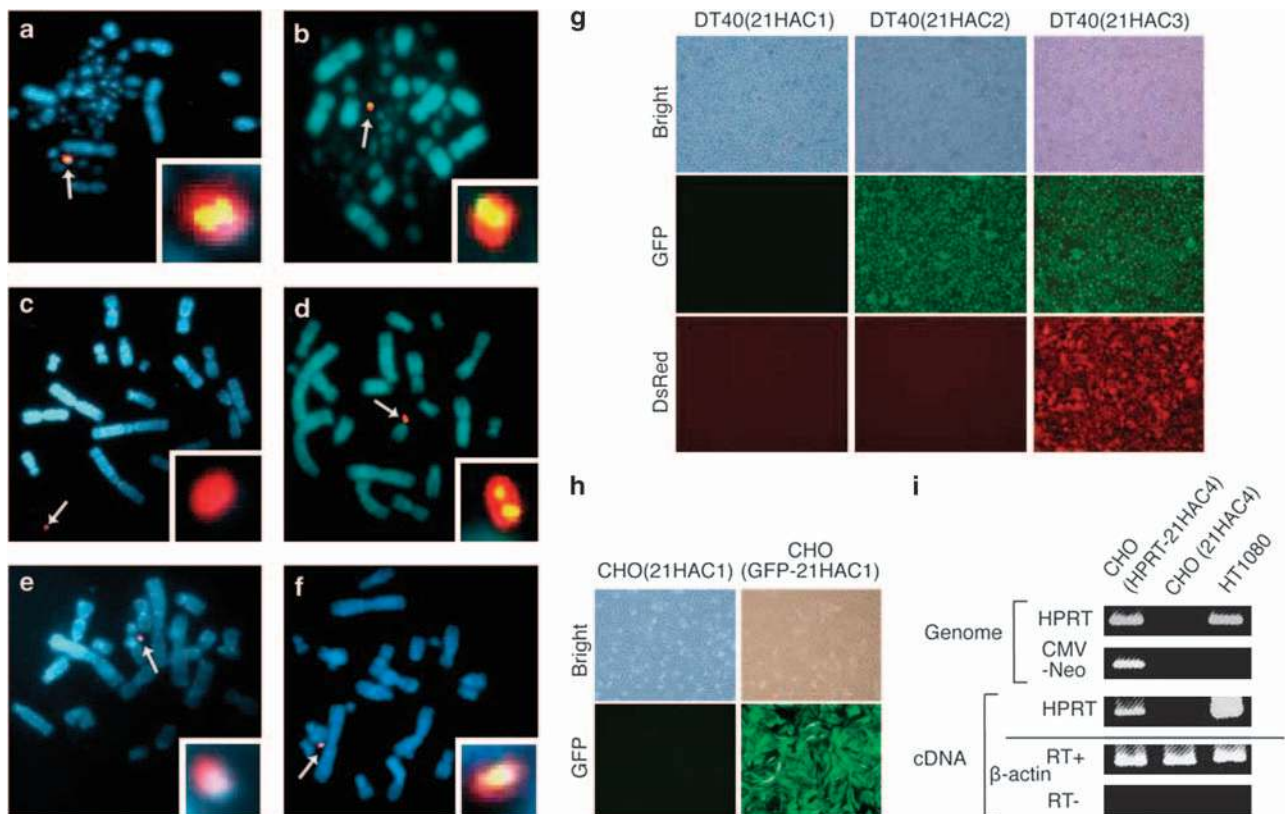
Figure 4 Analysis of a HAC containing desired genes. (a–f) FISH analyses for DT40 (21HAC2) (a), DT40 (21HAC3) (b), CHO (21HAC1) (c), CHO (GFP-21HAC1) (d), CHO (21HAC4) (e) and CHO (HPRT-21HAC4) (f). Digoxigenin-labeled human COT-1 DNA (red) was used to detect the HAC. Biotin-labeled CAG–EGFP, CMV-DsRed and RP6-127C8 (green) were used to detect the marker gene on the HAC in (a and d), (b) and (f), respectively. Chromosomal DNA was counterstained with 4,6-diamidino-2-phenylindole (DAPI). The inset shows an enlarged image of the HAC (arrow). (g) Expression of the fluorescent gene on the HAC in DT40 cells. Phase-contrast (top panel), GFP-fluorescence (middle panel) and DsRed-fluorescence (bottom panel) micrographs are shown. (h) Expression of the fluorescent gene on the 21HAC1 in CHO cells. (i) Representative genomic PCR and reverse transcriptase (RT)-PCR data for detecting HPRT-21HAC4 in CHO cells. HT1080 and CHO (21HAC4) cells were used as positive and negative controls, respectively. β-actin was used as an internal control.

gene reconstitution was comparable to that into the 3′neo-loxP site on the 21ΔqHAC/21ΔpqHAC by *neo* gene reconstitution.[14] These data suggest that a desired gene can be cloned into the 21HAC1 by the Cre-loxP system without utilizing a *neo* gene.

Gene cloning into 21HAC4 by Cre-loxP system (neo version)

In the Cre-loxP system previously developed on the 21ΔqHAC/21ΔpqHAC, desired genes were cloned using the *neo* gene reconstitution. Furthermore, one of the human BAC/PAC libraries constructed for the human genome sequencing project,[24] RPCI-6, has a 5′neo-loxP site, thus allowing the BAC/PAC inserts to transfer readily into the 21ΔqHAC/21ΔpqHAC vector with 3′neo-loxP-Bsd. To load previously developed vectors or the RPCI-6 library into the novel 21HAC1 vector developed in this study, the 3′neo-loxP-Bsd cassette was inserted into the 21HAC1 in DT40 cells instead of the 5′HPRT-loxP-Hyg cassette. With the targeting construct, 3′neo-loxP-Bsd-TK for the cloning site (loxP) and negative selection (*HSV-TK*), 18 of 58 drug-resistant clones selected in the presence of blasticidin S were targeted correctly, as shown by PCR (designated as 21HAC4) (Supplementary Figure S4a). FISH analyses showed that the targeting construct was integrated into the correct locus on the 21HAC1 in DT40 cells (data not shown). Then, the 21HAC4 was transferred to CHO-K1 cells from DT40 cells via MMCT (Figure 4e). To test whether a desired gene can be cloned into the 21HAC4 by neo gene reconstitution, a PAC containing the *HPRT* gene and a 5′neo-loxP site, RP6-127C8, which was previously screened,[19] as well as a Cre-expression vector were co-transfected into CHO-K1 cells containing 21HAC4, and the recombinant clones were selected using G418 over a period of 7 days (Figure 2c and Supplementary Figure S4b). In all, 2 out of 10 drug-resistant clones were positive by PCR with HPRT-specific primers and by FISH analysis, and expressed the human *HPRT* gene by reverse transcriptase-PCR analysis (Figures 4f and i). This ratio (2 of 10, or 20%) of the PAC vector insertion into the 3′neo-loxP site on the 21HAC4 by neo gene reconstitution was comparable to the previous study.[19] These data show that the previously developed plasmid vector for cloning into the 21ΔqHAC/21ΔpqHAC or the screened RPCI-6 vector can be used for cloning into the 21HAC4 using the *neo* gene reconstitution.

TAR cloning of p- and q-arms of the 21HAC4 and their sequencing

Previously, a transformation-associated recombination (TAR) cloning strategy[25] was successfully used to characterize neocentromeres.[26] In this work, we performed TAR cloning and sequencing of the 21HAC to compare the informative sequences and to confirm that the correct targeting events were made during the modifications in DT40 cells. TAR cloning was used to isolate fragments containing p- and q-arms of the 21HAC4 propagated in DT40 cells (Figure 3). The p- and q-arm regions were selectively cloned as circular YACs in yeast using TAR vectors carrying an arm-specific unique targeting sequence and an alphoid DNA repeat as a second targeting sequence (see Materials and methods section). Transformation experiments were carried out with freshly prepared yeast spheroplasts and a linearized TAR-specific vector as described in Materials and methods section. For each vector, from three to five positive clones were identified using a set of diagnostic primers. The size of the positive YACs for the p-arm region was ~20 kb; the size of the positive YACs for the q-arm region was ~80 kb. The YAC clones are relatively stable during propagation in yeast. For further analysis, circular YACs were retrofitted into BACs by homologous recombination in yeast and transformed into *Escherichia coli* (*E. coli*). Most of BACs rescued in *E. coli* underwent deletions, suggesting that the inserts are unstable in bacterial cells. However, after screening the *E. coli* transformants, we succeeded in finding BAC

clones with no detectable rearrangements, when transformation and subsequent growth of *E. coli* cells was performed at 30 °C. Lack of rearrangements in the BAC inserts was confirmed by comparison of original YAC isolates, retrofitted YAC/BACs and BACs (data not shown). The #26-2 (q-arm), and #32-2, #37-2 and #40-2 (p-arm) BACs were chosen for sequencing analysis, and homology search of the determined sequence of each BAC clone was carried out against current public databases. The #26-2 is isolate for q-arm, whose size was about 81 kb, and the #32-2, #37-2 and #40-2 were duplicate isolates for the p-arm, whose size was about 13 kb. The sequences are available from DDBJ/GenBank/EMBL under accession number AB553833–AB553834. Thus, the sequence data of TAR isolates show that chromosome truncation during construction of the HAC was driven exclusively by homologous recombination, and that the 21HAC1 (as well as its derivatives) does not contain any endogenous gene, and that correct targeting event occurred in loxP site, p-arm and q-arm. However, as the complete sequence of ~100 kb pericentromeric regions remained in the 21HAC (Figure 3), unknown and unreported genes or pseudogenes may exist on the 21HAC.

Pulsed-field gel electrophoresis and Southern blotting in 21HACs

To investigate the structural integrity of the 21HACs during the modifications in DT40 cells, Southern blot analyses were performed with genomic DNA isolated from DT40 containing the 21HACs, digested with *Bam*HI. The restriction fragments were size-separated by pulsed-field gel electrophoresis, Southern blotted and hybridized with a probe to α-satellite derived from human chromosome 21.[27] The restriction fragments from the 21HAC1, 21HAC2 and 21HAC4 were consistent with those from the parental hChr.21 (Supplementary Figure S5). The total size of the 21HACs was estimated to be about 5 Mb from the pulsed-field gel electrophoresis, Southern blotting and sequence analyses. These data suggest that the gross structure of the 21HACs was stably maintained during the modifications in DT40 cells.

Stability of 21HAC2 in vitro

As the 21HAC2 contains *EGFP* gene for monitoring, we used this HAC for the following experiments. To investigate the stability of a HAC in human cells, the 21HAC2 was transferred to human immortalized mesenchymal stem cell (hiMSC) line, which was generated and characterized previously,[18] by MMCT. Ten GFP-positive clones were selected and examined in the following experiments (Figure 5a). PCR analyses showed that 4 out of the 10 drug-resistant clones were positive for the markers on the 21HAC2 (data not shown). FISH analyses showed that the 21HAC2 was present as an individual chromosome in the hiMSCs (Figure 5b). After the hiMSC (21HAC2) cells were cultured for about 3 months without selection, the stability of the 21HAC2 and GFP expression on the 21HAC2 were tested. FISH analyses revealed that the 21HAC2 was independently and stably maintained in hiMSCs (Figure 5c), and most of the hiMSC (21HAC2) cells were GFP-positive (data not shown). These data suggest that the 21HAC2 can be maintained stably and the *EGFP* gene on the 21HAC2 can be expressed stably even after long-term culture *in vitro*.

Stability of 21HAC2 in vivo

Next, to demonstrate that this HAC can be used for the generation of trans-chromosomic mice for the functional analysis of a desired gene or genes in mice, the 21HAC2 was introduced into the mouse ES cell lines, TT2F and E14, by MMCT. Ten TT2F lines and seven E14 lines were GFP-positive clones and so were examined in detail (Figure 5d).
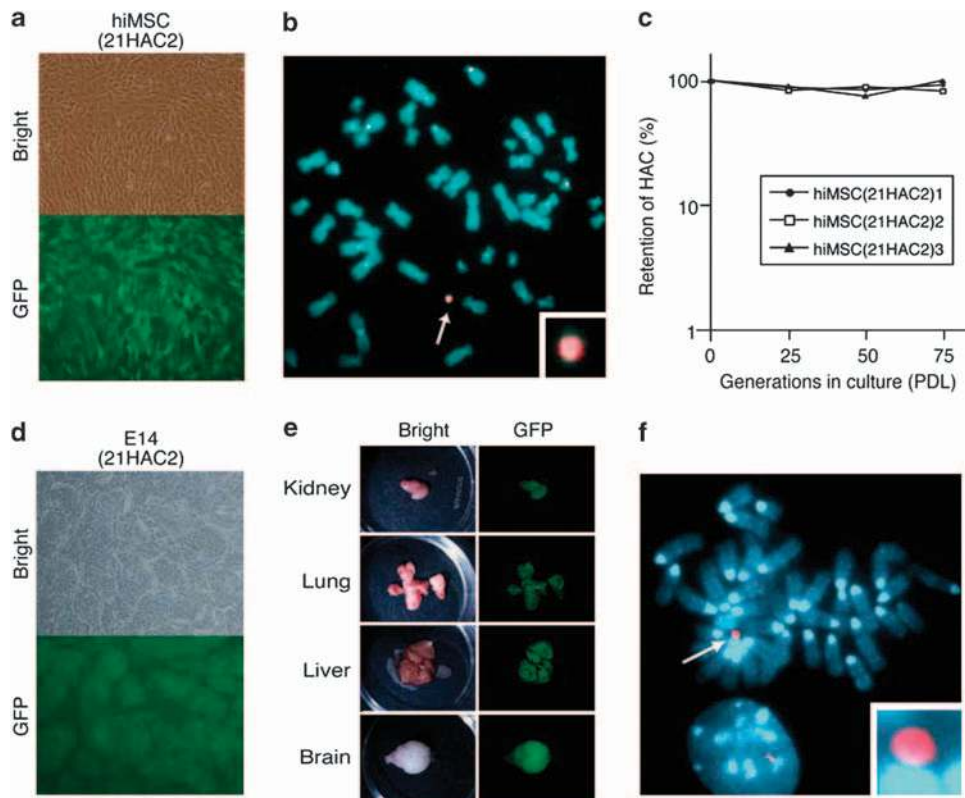
Figure 5 Stability of 21HAC2 in vitro and in vivo. (**a**) Morphology of hiMSC (21HAC2) cells. Phase-contrast (top panel) and fluorescence (bottom panel) micrographs are shown. (**b**) FISH analyses for hiMSC (21HAC2) cells. An arrow indicates the 21HAC2 and the inset shows an enlarged image of the 21HAC2. Digoxigenin-labeled p11-4, derived from hChr.21 (red) was used to detect the HAC. (**c**) Mitotic stability of 21HAC2 in hiMSC (21HAC2) cells as a function of population doubling (PDL). (**d**) Morphology of E14 (21HAC2) cells. Phase-contrast (top panel) and fluorescence (bottom panel) micrographs are shown. (**e**) GFP expression in various tissues of trans-chromosomic (Tc) mouse containing 21HAC2. Bright (left panel) and fluorescence (right panel) micrographs are shown. (**f**) FISH analyses for Tc tail-fibroblasts containing the 21HAC2. Digoxigenin-labeled human COT-1 DNA (red) was used to detect the HAC. An arrow indicates the 21HAC2 and the inset shows an enlarged image of the 21HAC2.

PCR analyses using primers for the detection of the 21HAC2 showed that 9 of the 10 TT2F clones and 6 of the 7 E14 clones contained the 21HAC2 (data not shown). FISH analyses showed that the 21HAC2 was present as an individual chromosome in these mouse ES cells (Supplementary Figure S6a). In order to determine the developmental potential of these cells, as well as the expression of the transferred EGFP in various tissues, chimeric mice were produced from TT2F (21HAC2) and E14 (21HAC2). Chimeras with various forms of coat-color chimerism were successfully obtained and GFP-positive chimeric mice were analyzed. The five chimeric mice with high coat-color chimerism were mated, and the 21HAC2 was transmitted through the germline in three out of the five chimeric mice. The GFP-positive trans-chromosomic F1 mice were used for the following analyses. The EGFP gene driven by the CMV early enhancer/chicken β actin (CAG) promoter on the 21HAC2 was expressed in all tissues examined (Figure 5e), suggesting that the 21HAC2 was stably maintained in vivo. FISH analyses showed that the 21HAC2 was present as an individual chromosome in trans-chromosomic tail-fibroblasts (Figure 5f), and that the 21HAC2 was detected in at least 80% of kidney, lung, liver and brain cells (data not shown), similar to the previous reports.[15,19] These data suggest that the 21HAC2 does not interfere with normal development and is stably maintained in mice.

Growth suppression of cells with a HAC in vitro and in vivo
To construct a HAC for safe gene therapy, a suicide gene, HSV-TK, was inserted into the 21HAC series developed in this study. To assess whether the HSV-TK on the HAC is functional in vitro and in vivo, the 21HAC2 was transferred to the human fibrosarcoma cell line, HT1080, by MMCT. The HT1080 (21HAC2) cells were characterized by GFP-fluorescence, PCR and FISH analyses as described above (Figures 6a and b), and treated with ganciclovir (GCV) in vitro and in vivo. When the total cell numbers were examined, the in vitro growth rate of HT1080 (21HAC2) cells was comparable to that of the parental HT1080 cells without selection (Figure 6c). On the other hand, the growth rate of HT1080 (21HAC2) cells was remarkably suppressed by GCV when compared with the growth of the parent cell line, that is, 10^5 vs 10^7 on sixth day (Figure 6d). Thus, the cells containing the HAC are sensitive to the treatment with GCV.

Next, to determine the effect of GCV in vivo, HT1080 (21HAC2) and parental HT1080 cells were subcutaneously injected into nude mice, then GCV was injected to the mice. The tumor weights from mice harboring HT1080 (21HAC2) cells were significantly lower than those from the parental HT1080 for the concentration of GCV, compared with the controls (phosphate-buffered saline treatment) ($P<0.01$ vs $P=0.923$) (Figures 6e and f). These data suggest that HSV-TK on 21HAC2 is biologically functional in vivo and in vitro.

Purification of ES cells without HAC in vitro
To eliminate the HAC from mouse ES cells, E14 cells carrying the 21HAC2 were selected on GCV-containing medium. In all, 65 out of the 70 drug-resistant clones were GFP-negative. Six of these GFP-negative clones were analyzed by FISH. The 21HAC2 was not
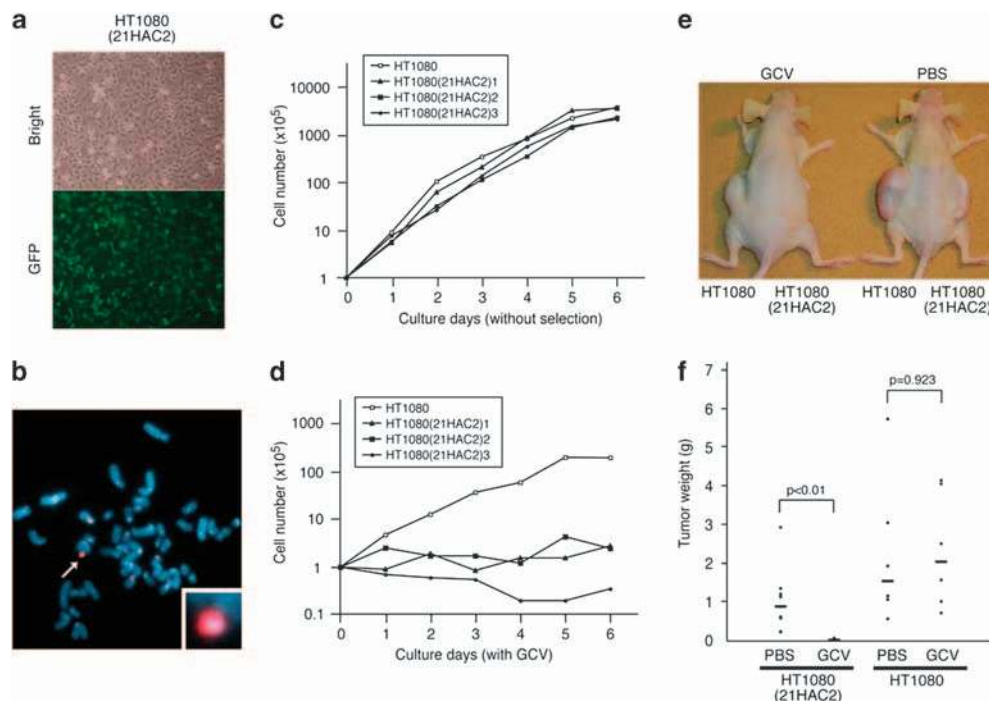
Figure 6 Effect of GCV on HT1080 (21HAC2) with *HSV-TK in vitro* and *in vivo*. (**a**) Morphology of HT1080 (21HAC2) cells. Phase-contrast (top panel) and fluorescence (bottom panel) micrographs are shown. (**b**) FISH analyses for HT1080 (21HAC2) cells. An arrow indicates the 21HAC2 and the inset shows an enlarged image of the 21HAC2. Digoxigenin-labeled p11-4, derived from hChr.21 (red), was used to detect the HAC. (**c** and **d**) Growth curve of HT1080 and HT1080 (21HAC2) cells with **d** or without **c** GCV selection *in vitro*. (**e**) Tumors derived from HT1080 and HT1080 (21HAC2) cells with (left) or without (right) GCV treatment. (**f**) Tumor weight in HT1080- and HT1080 (21HAC2)-derived tissues with phosphate-buffered saline (PBS) or GCV treatment. A statistical analysis was performed using a two-tailed Student's *t*-test.

detected in any of them, and the karyotype of all these clones was normal, showing that GCV does not induce chromosomal abnormalities (Supplementary Figure S6b and c).

DISCUSSION

In this work, a novel HAC with a predefined sequence and its derivatives are described (21HAC1, 21HAC2, 21HAC3 and 21HAC4). As all the 21HAC-derivatives contain the *HSV-TK*, they can be used for negative selection *in vitro* and *in vivo*. As the 21HAC1 does not contain fluorescent gene, the tissue-specific promoter-driven fluorescent gene can be inserted to the 21HAC1 for tissue-specific monitoring. As fluorescent marker gene was inserted into the 21HAC2 and 21HAC3, cells containing the 21HAC can be easily monitored after MMCT. As the 21HAC4 contains the 3′neo-loxP, the RPCI-6 PAC library containing 5′neo-loxP site can be transferred into the 21HAC4. Thus, each HAC derivative may be used for different purposes, including gene cloning/expression experiments, cell reprogramming and analysis of mitotic segregation of HAC.

Previously, human immunoglobulin genes (1.5–2 Mb) were cloned into the SC20-HAC, and the human *DMD* gene (2.4 Mb) was cloned into the 21HAC2 developed in this study, by a combination of Cre-loxP-mediated chromosome translocation and telomere-directed truncation in homologous recombination-proficient chicken DT40 cells.[13,28] Inserts with a desired gene from the PAC library, RPCI-6, can be directly cloned into 21ΔqHAC/21ΔpqHAC and 21HAC4 without modification of a PAC clone.[19] However, only one chromosome region (or gene) can be cloned into a HAC using the Cre-loxP system. In this study, two different vectors containing a desired gene were

inserted sequentially into the 21HAC1 by homologous recombination in DT40 cells. Theoretically, two or more vectors containing a desired gene can be inserted sequentially into the HAC. Therefore, any combination of genes, including a full-length genomic DNA, can in theory be cloned into the HAC derivatives by a combination of these cloning systems and transferred into a desired cell type using the HAC. The HAC with a gene of interest can be also transferred to mouse ES cells, and the chimeric mice or their offspring will be useful either for functional analyses *in vivo* or as a disease model.[29]

During the investigation of the HAC integrity by Southern blotting with a human chromosome 21-derived alphoid satellite probe, we found no rearrangement of the restriction fragments of the HAC centromere in DT40 clones. For the future application of the HAC to the gene therapy protocol, any rearrangements that may lead to safety concerns are not desirable even if they are inert. Although several HAC vectors have proven to be very useful tools,[6] previously developed HAC vectors were structurally undefined, making them unsuitable for gene and cell therapies in human. In this study, using a TAR cloning strategy and following sequencing of p- and q-arm regions of the HAC, we proved that our novel HAC consists of exclusively repetitive elements representing a functional centromere and does not contain any endogenous genes. Thus, this HAC has multiple applications for human therapeutic goals. However, it is important to examine that the HAC does not induce the malignant transformation. If cells containing this HAC are transformed to tumor cells, most of the tumor cells, if not all, can be eliminated by GCV injection *in vivo*.

Recently, we succeeded in complete genetic correction of induced pluripotent stem cells from patients with Duchenne muscular dystrophy using the 21HAC2 containing the entire sequence for

human dystrophin.[30] Furthermore, we also showed a possibility for the treatment of glioma using hiMSC containing the 21HAC2 with HSV-TK.[31] However, detail information of construction of the 21HACs had not been reported previously. Thus, the data in this study are crucial information for the future gene delivery. As mouse ES cells that have lost the HAC could be selected by GCV, the HAC, containing defined factors, such as OCT4, SOX2, KLF4, cMYC and LIN28, may be useful for generating induced pluripotent stem cells avoiding the problem of insertional mutagenesis.[32,33] As the HAC could be transmitted through the germline and stably maintained *in vivo*, it will be useful to generate model animals containing a megabase-sized gene, such as a HLA cluster (3 Mb), CYP3A cluster (~1 Mb) or UGT2 cluster (1.5 Mb). Thus, the HACs developed in this study will be useful not only for gene and cell therapies, but also for the generation of humanized model animals. Thus, the HACs may be designated as 'multipotent vectors'.

MATERIALS AND METHODS

Plasmid construction
The region a to f in Figure 3 shows homologous region on the hChr.21 for the following targeting vectors. Targeting vectors were constructed as follows. For 5′HPRT-loxP-Hyg-TK: two 3.4- and 2.8-kb fragments for the homologous arms, corresponding to AP001657, were amplified by PCR using primers #21CEN < 1 > 1L/#21CEN < 1 > 1R (3.4 kb, region d), 5′-ACCTGGAATTTCC TACCATCCCCCATAA-3′ and 5′-ATCTCTCCAGAGGGACAGCATCATACCC-3′; #21CEN < 2 > 1L/#21CEN < 2 > 1R (2.8 kb, region e), 5′- CCTGCAAGTTAT GACCACTGGGGATTTT-3′ and 5′-CTGCAGTGAGCCGAGATCATACCACT GT-3′. The PCR products were digested with either *Eco*RI/*Bam*HI or *Xho*II and sub-cloned into *Eco*RI/*Bam*HI or *Bgl*II of pKO Scrambler V901 backbone vector (Lexicon Genetics, Woodlands, TX, USA), which contained the following three selection/insertion cassettes between the *Asc*I and *Kpn*I sites: a 2-kb fragment (MC1-TK) from pKO SelectTK V830 (Lexicon Genetics) containing the *HSV-TK*, a 1.4 kb *Cla*I/*Asc*I (5′HPRT-loxP) fragment from pKO SelectHPRT V820 (Lexicon Genetics) containing part of the human *HPRT* gene with the loxP site inserted at the *Xba*I site of intron B, and a 1.8 kb *Cla*I/ *Kpn*I fragment (PGK-Hyg) from PGKhygrodetlaLT20 encoding the hygromycin gene. For pBS-TEL/Δp Puro: a PCR product from AL163201 was amplified by PCR with the following primers #21p-BsrGI-*Bam*HI/21p-loxP 3Rb (5.1 kb, region a), 5′-ACGGATCCTGTACAATACACAGTTGACTGTCTCAGTGTG-3′ and 5′-ACGGATCCAAGCCCACATATTATTATGCTGTGCTTC-3′, digested with *Bam*HI (5.1 kb) and sub-cloned into the *Bam*HI sites of pBS-TEL/Puro vector.[13] For pBS-TEL/Δq HisD: a PCR product from AP001657 was amplified by PCR with the following primers q1L/q1R (7 kb), 5′-GGAGCAACAGGACCT CTCATTCCTTGTT-3′ and 5′-CCAATGTCAGGCACTCCTGCTCTAAATG-3′, digested with *Bam*HI (7 kb, region f) and sub-cloned into the *Bam*HI sites of pBS-TEL/HisD vector.[13] For I-EGFP-I-Bsd: two 3.8- and 2.6-kb fragments for homologous arms corresponding to AP001657 were amplified by PCR using primers #21cenG1L/#21cenG1R (3.8 kb, region b1), 5′-GTGAAGGCATTGCCA GTGTTTTCTTCTG-3′ and 5′-TTTGCCGTAGTCAGTGGGTGAATCATCT-3′; #21cenG3L/#21cenG 3R (2.8 kb, region c), 5′-GTTTGAGAGGACATGCAACAC CCTGAAT-3′ and 5′-TTTAGATGCAGGGGCATACTGTGAGCAT-3′, digested with either *Apa*I or *Xba*I, and sub-cloned into the equivalent sites of pCMV/ Bsd (Invitrogen). *Sal*I/*Hind*III fragments of CAG–EGFP from pCX-EGFP (a gift from Dr M Okabe) were cloned into pJC5-4 (a gift from Dr G Felsenfeld) including HS4 insulators, then the CAG–EGFP fragments flanked with the HS4 insulator were cloned into the *Not*I/*Xho*I sites of pCMV/ Bsd with two homologous arms. For DsRed-neo: two 3.0- and 3.0-kb fragments for homologous arms corresponding to AP001657 were amplified by PCR using primers #21cenF (*Pci*I *Apa*LI *Asc*I)/#21cenR(*Nco*I) (3.0 kb, region b2), 5′-CCCGTGCACGGCGCGCCACTAGCCAGTTTTCCCAGCAC-3′ and 5′-CC CCATGGAAAATATTTTAACAAGAGTTATTGTTTTGC-3′; #21cenF (*Nco*I-1)/ #21cenR(*Asc*I) (3.0 kb, region b3), 5′-CCCATGGCATATTGTAAATAAGACA CCCTCCGT-3′ and 5′-CCCGGCGCGCCGAGTATCCATGAAGAAATGCACA-3′, digested with either *Pci*I or *Pci*I/*Asc*I and sub-cloned into the equivalent sites of pDsRed2-Express-N1 (Takara, Ostu, Japan). For I-EGFP-I-loxP-3′HPRT:

a 2.3 kb *Eco*RI/*Asc*I fragment (3′HPRT-loxP) from pKO SelectHPRT V820 (Lexicon Genetics) containing part of the human *HPRT* gene with a loxP site inserted at the *Xba*I site of intron B, was cloned into the *Eco*RI/*Asc*I sites of V901 (V901-3′HPRT-loxP). Next, the CAG–EGFP constructs flanked with the HS4 insulator were cloned into V901-3′HPRT-loxP. For 3′neo-loxP-Bsd-TK: *Bam*HI fragment from pSF1 with CMV-Bsd[14] was cloned into the *Bam*HI site of V901 (V901-3′neo-loxP-Bsd). The *Asc*I/*Sal*I fragments from V901-3′neo-loxP-Bsd were cloned into the equivalent sites of V901, inserting the three fragments #21CEN < 1 > 1L/#21CEN < 1 > 1R (3.4 kb), MC1-TK (2 kb) and #21CEN < 2 > 1L/#21CEN < 2 > 1R (2.8 kb).

Cell culture
Chicken DT40 cells containing hChr.21 were generated previously[34] and maintained at 40 °C in RPMI 1640 medium supplemented with 10% fetal bovine serum, 1% chicken serum, 50 μM 2-mercaptoethanol and 1.5 mg ml⁻¹ G418. CHO-K1 or hprt-deficient CHO (CHO (hprt−/−)) cells used as fusion recipients for chromosome transfer were maintained at 37 °C in Ham's F-12 nutrient mixture (Invitrogen, Carlsbad, CA, USA) supplemented with 10% bovine calf serum. Mouse embryonic fibroblasts were isolated from 13.5 days *post coitum* embryos. The tail-fibroblasts, mouse embryonic fibroblasts, hiMSC and human fibrosarcoma (HT1080) cells were grown in Dulbecco's modified Eagle's medium (Sigma, St Louis, MO, USA) plus 10% fetal bovine serum. The parental mouse ES cell lines, TT2F and E14, and the microcell hybrid clones, were maintained on mitomycin C (Sigma)-treated Jcl:ICR (CLEA Japan, Tokyo, Japan) mouse embryonic fibroblasts as feeder layers in Dulbecco's modified Eagle's medium with 18% fetal bovine serum (Thermo Scientific Hyclone, Yokohama, Japan), 1 mM sodium pyruvate (Invitrogen), 0.1 mM non-essential amino acids (Invitrogen), 0.1 mM 2-mercaptoethanol (Sigma), 2 mM L-glutamine (Invitrogen) and 1000 U ml⁻¹ leukemia inhibitory factor (LIF) (Funakoshi, Tokyo, Japan).

Modification of hChr.21 and HAC in DT40 cells
Homologous recombination-proficient chicken DT40 cells (1×10⁷) were collected in 0.5 ml RPMI with 25 μg of linearized targeting vector and electroporated at 550 V and 25 μF using a Gene Pulser apparatus (Bio-Rad, Tokyo, Japan). Drug-resistant DT40 clones were then selected in 1.5 mg ml⁻¹ G418, 1.5 mg ml⁻¹ hygromycin, 0.3 μg ml⁻¹ puromycin, 15 μg ml⁻¹ blasticidin S or 0.5 mg ml⁻¹ L-histidinol. Homologous recombination in DT40 hybrid clones was identified by PCR as follows. The modified chromosomes were transferred to CHO cells via MMCT.

MMCT
MMCT was performed as described previously.[23] DT40 containing the HAC were transferred to CHO cells by MMCT, then CHO cells containing the HAC were used as donor microcell hybrids. Briefly, mouse ES and hiMSC and HT1080 cells were fused with microcells prepared from donor hybrid CHO (21HAC2) cells, and selected with blasticidin S (3–8 μg ml⁻¹). As a fluorescent marker gene, such as *EGFP*, was inserted into the 21HAC2 and 21HAC3, we can easily monitor which cells contain the 21HAC after MMCT. The transferred HAC in each line was characterized by PCR, reverse transcription-PCR and FISH analyses.

FISH analyses
FISH analyses were performed using either fixed metaphase or interphase spreads of each cell hybrid using digoxigenin-labeled (Roche, Basel, Switzerland) human COT-1 DNA (Invitrogen), digoxigenin-labeled p11-4 (hChr.21-derived α-satellite clone)[27] and biotin-labeled DNA (RP6-127C8, PGK-Hygro, PGK-Puro, β-actin-HisD, CAG–EGFP and CMV-DsRed), essentially as described previously.[23] Chromosomal DNA was counterstained with 4,6-diamidino-2-phenylindole (Sigma). Images were captured using the NIS-Elements system (Nikon, Tokyo, Japan) or the Argus system (Hamamatsu Photonics, Hamamatsu, Japan). The 20 metaphases and 100 interphases were counted for the chromosomal analyses.

Genomic PCR analyses

Genomic DNA was extracted from cell lines using a genomic extraction kit (Gentra System, Minneapolis, MN, USA), and PCR was performed using standard methods as described previously.[19] The primer pairs for detecting the targeting were described in Supplementary Table S2. DT40, CHO (hprt−/−), CHO-K1, hiMSC, HT1080, E14 and TT2F cells were used as negative controls.

Southern blot analyses

DNA samples were digested with SphI, and the digested samples were separated on 0.8% agarose gel electrophoresis and transferred to a Hybond N+ membrane (Amersham, Sunnyvale, CA, USA) in an alkaline solution. The membrane was hybridized overnight at 65 °C with 32P-labeled probes and washed three times at 65 °C in 0.1% sodium cloride- sodium citrate (SSC) for 30 min, 0.1% sodium dodecyl sulfate for 30 min. Radioactivity was visualized by the imaging film (KODAK, Tokyo, Japan). The primer pairs for probes to detect the targeting of 5′HPRT-loxP-Hyg-TK at the AP001657 locus were: 21A5-1L/21A5-1R (649 bp, 5′ probe), 5′-CAGGCAACTGTAACACAGTGGTAGGTA-3′ and 5′-AACAGTAGAGCAATTTCAGGCAGGTC-3′; 21qA3-3L/21qA3-3R (626 bp, 3′ probe), 5′-CGCAGCTTTTAGCTGAACTAAGGAGA-3′ and 5′-GTGACA CAGGGATACTCTGTCCAAAA-3′.

TAR cloning and sequence analyses

The TAR cloning vectors, pVC-parm/Sat+ and pVC-qarm/Sat−, containing one unique targeting sequence and an alphoid DNA repeat as the second targeting sequence, were constructed as follows. To construct pVC-parm/Sat+ vector, an 85-bp alphoid DNA sequence[35,36] was inserted into BamHI and XbaI polylinker sites of pVC604 (CEN6-HIS3).[37] A unique hook, corresponding to the open reading frame of Puro from the plasmid pPGKPuro (AF090453), was also inserted into the polylinker as a 546 bp ApaI–BamHI fragment. To construct the pVC-qarm/Sat− vector, an 85-bp alphoid DNA sequence was inserted into the ApaI and BamHI polylinker sites of pVC604. As a unique hook for pVC-qarm/Sat− vector, a 186-bp fragment of pBluescript was chosen (the pBluescript sequence was introduced into the 21HAC during truncation of the q-arm.). A unique hook was inserted into the BamHI and XbaI polylinker sites of the TAR vector. All targeting sequences were PCR amplified using the pairs of primers presented in Supplementary Table S1. The vectors were cut with BamHI (this site is located between the alphoid repeat and a unique sequence) before yeast transformation. The highly transformable Saccharomyces cerevisiae strain VL6-48 was used for TAR cloning.[37] Agarose plugs containing high molecular weight genomic DNA were prepared from chicken DT40 cells carrying 21HAC4, and were used for the TAR cloning experiments. Transformation experiments were carried out with freshly prepared yeast spheroplast as previously described.[37] In each experiment, approximately 2 μg of genomic DNA isolated from DT40 cells carrying 21HAC4, 1 μg of the linearized vector and 8×10^8 spheroplasts were used. Between 50 and 150 His+ transformants were then screened by diagnostic primers (Supplementary Table S1) to identify transformants positive for p- or q-arm-specific sequences. The diagnostic primers for centromeric contigs of human chromosome 21 were developed based on the available human genome sequence.[20] Transformants were first combined into pools and examined by diagnostic primers. Then, individual colonies containing either the p-arm or q-arm sequences were isolated from each positive pool by a second round of PCR screening. The PCR products were sequenced to verify that they matched the predicted sequences of hChr.21.

Retrofitting of YACs into YACs/BACs, electroporation of YAC/BACs into E. coli cells, and BAC DNA isolation were carried out as previously described,[37] except that electroporation and bacterial culture growth was carried out at 30 °C to stabilize inserts consisting of pericentromeric repeats. The size of BAC DNA was determined after digestion with NotI and further analysis by clamped homogenous electric fields (CHEF) gel electrophoresis.

Sequencing of each BAC clone was accomplished using an automated robotic system and standard procedures. Briefly, BAC DNA was fragmented for plasmid library construction with an average insert size of 2–4 kb using a HydroShear DNA Shearing Device (Genemachines, San Carlos, CA, USA). Amplification of each plasmid DNA was performed by the rolling circle amplification method (Templiphi DNA Amplification Kit, GE Healthcare, Buckinghamshire, UK). Sequencing of rolling circle amplification products was then carried out using a BigDye terminator v3.1 cycle sequencing kit (Applied Biosystems, Foster city, CA, USA). Sequencing products cleaned up by ethanol precipitation were run on automated ABI 3730 capillary sequencers (Applied Biosystems). The sequence data were assembled using the Phrap/Consed system.[38] Gap closure and re-sequencing of low-quality regions in the assembled data were done by primer walking on selected clones, the nested deletion method,[39] and construction of short-insert libraries.[40]

Pulsed-field gel electrophoresis and Southern blotting

A high molecular weight DNA for long-range mapping was prepared in agarose gel plugs as described in Trowell et al.[41] The plugs were sliced into segments, each containing 0.4×10^6 cells. Agarose-DNA segments were then digested for 18 h with BamHI (Takara) (100 U per segment). The DNA fragments were size-separated on a CHEF DR-II apparatus (Bio-Rad Laboratories). The run time was 24 h at 6 V cm⁻¹ with a 60–120 s switch time ramp. hChr.21-derived α-satellite clone p11-4[27] was labeled with α-32P-dCTP by the random priming method (Amersham). Hybridization was carried out at 65 °C in 0.5 M phosphate buffer (0.5 M Na_2PO_4, 7% sodium dodecyl sulfate, 1 mM EDTA, pH 7.2). Following hybridization, the filters were washed in 40 mM phosphate buffer/1% sodium dodecyl sulfate at 65 °C. The autoradiographs were obtained by an imaging system BAS2500 (Fuji Film, Tokyo, Japan).

Reverse transcriptase-PCR analyses

Total RNA was prepared from cultured cells using RNeasy columns, according to the manufacturer's instructions (Qiagen, Hilden, Germany), and then was treated with RNase-free DNase I (Wako, Osaka, Japan). First-strand complementary DNA synthesis was performed using random hexamers and Super-Script III reverse transcriptase (Invitrogen). PCR was performed with the complementary DNA using AmpliTaq Gold (Perkin Elmer, Waltham, MA, USA). Amplifications were performed at an annealing temperature of 58 °C for 30 cycles, and then the amplified fragments were resolved on a 2% agarose gel, followed by staining with ethidium bromide. The primer sequences were as follows: for HPRT (200 bp), 5′-TCCTCCTCCTGAGCAGTCA-3′ and 5′-CA CTCTCGAGCAAGACGTTCA-3′; for β-actin (165 bp), 5′-TGTTACCAACTGGG ACGACA-3′ and 5′-GGGGTGTTGAAGGTCTCAAA-3′.

Generation of chimeric mice

Chimeric mice were produced from the two TT2F (21HAC2) and two E14 (21HAC2) cell lines. Chimera production was performed as described previously.[23] Briefly, the mouse ES cells were injected into eight cell- or blastocyst-stage embryos derived from ICR mice (CLEA, Tokyo, Japan), and then transferred into pseudopregnant ICR females. Three chimeric mice showing 100% coat-color chimerism were used for mating. All animal experiments were approved by the Institutional Animal Care and Use Committee of Tottori University.

GCV treatment in vitro

HT1080 (21HAC2) cells that express HSV-TK can initially phosphorylate GCV, leading to an accumulation of its cytotoxic metabolite, GCV triphosphate, and subsequent incorporation into DNA and apoptosis. To assess the effect of GCV on the growth of HT1080 (21HAC2) cells in vitro, 1×10^5 HT1080 (21HAC2) and HT1080 cells were plated with or without 6 μM GCV, then the cells were counted each day for 6 days. Mouse ES cells containing the 21HAC2, E14 (21HAC2) cells, were treated with 6 μM GCV, then after 7 days, GFP-negative clones were analyzed.

Tumor formation and GCV treatment in vivo

To produce tumors, 2×10^6 HT1080 (21HAC2) and HT1080 cells were subcutaneously injected into CD-1 (ICR)-nu mice (Charles River, Yokohama, Japan). After 10 days, the nude mice (n=6) were treated with an intraperitoneal injection of GCV (30 mg kg⁻¹) or phosphate-buffered saline every day. After 4 weeks, the mice were dissected and the tumors weighed.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

1 O'Connor TP, Crystal RG. Genetic medicines: treatment strategies for hereditary disorders. Nat Rev Genet 2006; 7: 261–276.
2 Hacein-Bey-Abina S, Von Kalle C, Schmidt M, McCormack MP, Wulffraat N, Leboulch P et al. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. Science 2003; 302: 415–419.
3 Costantini F, Radice G, Lee JL, Chada KK, Perry W, Son HJ. Insertional mutations in transgenic mice. Prog Nucleic Acid Res Mol Biol 1989; 36: 159–169.
4 Palmiter RD, Brinster RL. Germ-line transformation of mice. Annu Rev Genet 1986; 20: 465–499.
5 Soriano P, Gridley T, Jaenisch R. Retroviruses and insertional mutagenesis in mice: proviral integration at the Mov 34 locus leads to early embryonic death. Genes Dev 1987; 1: 366–375.
6 Oshimura M, Katoh M. Transfer of human artificial chromosome vectors into stem cells. Reprod Biomed Online 2008; 16: 57–69.
7 Basu J, Willard HF. Artificial and engineered chromosomes: non-integrating vectors for gene therapy. Trends Mol Med 2005; 11: 251–258.
8 Ikeno M, Inagaki H, Nagata K, Morita M, Ichinose H, Okazaki T. Generation of human artificial chromosomes expressing naturally controlled guanosine triphosphate cyclohydrolase I gene. Genes Cells 2002; 7: 1021–1032.
9 Mejia JE, Alazami A, Willmott A, Marschall P, Levy E, Earnshaw WC et al. Efficiency of de novo centromere formation in human artificial chromosomes. Genomics 2002; 79: 297–304.
10 Basu J, Compitello G, Stromberg G, Willard HF, Van Bokkelen G. Efficient assembly of de novo human artificial chromosomes from large genomic loci. BMC Biotechnol 2005; 5: 21.
11 Heller R, Brown KE, Burgtorf C, Brown WR. Mini-chromosomes derived from the human Y chromosome by telomere directed chromosome breakage. Proc Natl Acad Sci USA 1996; 93: 7125–7130.
12 Mills W, Critcher R, Lee C, Farr CJ. Generation of an approximately 2.4 Mb human X centromere-based minichromosome by targeted telomere-associated chromosome fragmentation in DT40. Hum Mol Genet 1999; 8: 751–761.
13 Kuroiwa Y, Tomizuka K, Shinohara T, Kazuki Y, Yoshida H, Ohguma A et al. Manipulation of human minichromosomes to carry greater than megabase-sized chromosome inserts. Nat Biotechnol 2000; 18: 1086–1090.
14 Katoh M, Ayabe F, Norikane S, Okada T, Masumoto H, Horike S et al. Construction of a novel human artificial chromosome vector for gene delivery. Biochem Biophys Res Commun 2004; 321: 280–290.
15 Shinohara T, Tomizuka K, Takehara S, Yamauchi K, Katoh M, Ohguma A et al. Stability of transferred human chromosome fragments in cultured cells and in mice. Chromosome Res 2000; 8: 713–725.
16 Tomizuka K, Shinohara T, Yoshida H, Uejima H, Ohguma A, Tanaka S et al. Double trans-chromosomic mice: maintenance of two individual human chromosome fragments containing Ig heavy and kappa loci and expression of fully human antibodies. Proc Natl Acad Sci USA 2000; 97: 722–727.
17 Kuroiwa Y, Kasinathan P, Choi YJ, Naeem R, Tomizuka K, Sullivan EJ et al. Cloned transchromosomic calves producing human immunoglobulin. Nat Biotechnol 2002; 20: 889–894.
18 Ren X, Katoh M, Hoshiya H, Kurimasa A, Inoue T, Ayabe F et al. A novel human artificial chromosome vector provides effective cell lineage-specific transgene expression in human mesenchymal stem cells. Stem Cells 2005; 23: 1608–1616.
19 Kazuki Y, Hoshiya H, Kai Y, Abe S, Takiguchi M, Osaki M et al. Correction of a genetic defect in multipotent germline stem cells using a human artificial chromosome. Gene Ther 2008; 15: 617–624.
20 Hattori M, Fujiyama A, Taylor TD, Watanabe H, Yada T, Park HS et al. The DNA sequence of human chromosome 21. Nature 2000; 405: 311–319.
21 Reymond A, Camargo AA, Deutsch S, Stevenson BJ, Parmigiani RB, Ucla C et al. Nineteen additional unpredicted transcripts from human chromosome 21. Genomics 2002; 79: 824–832.
22 Koi M, Shimizu M, Morita H, Yamada H, Oshimura M. Construction of mouse A9 clones containing a single human chromosome tagged with neomycin-resistance gene via microcell fusion. Jpn J Cancer Res 1989; 80: 413–418.
23 Tomizuka K, Yoshida H, Uejima H, Kugoh H, Sato K, Ohguma A et al. Functional expression and germline transmission of a human chromosome fragment in chimaeric mice. Nat Genet 1997; 16: 133–143.
24 Frengen E, Zhao B, Howe S, Weichenhan D, Osoegawa K, Gjernes E et al. Modular bacterial artificial chromosome vectors for transfer of large inserts into mammalian cells. Genomics 2000; 68: 118–126.
25 Kouprina N, Larionov V. TAR cloning: insights into gene function, long-range haplotypes and genome structure and evolution. Nat Rev Genet 2006; 7: 805–812.
26 Cancilla MR, Tainton KM, Barry AE, Larionov V, Kouprina N, Resnick MA et al. Direct cloning of human 10q25 neocentromere DNA using transformation-associated recombination (TAR) in yeast. Genomics 1998; 47: 399–404.
27 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.
28 Hoshiya H, Kazuki Y, Abe S, Takiguchi M, Kajitani N, Watanabe Y et al. A highly stable and nonintegrated human artificial chromosome (HAC) containing the 2.4 Mb entire human dystrophin gene. Mol Ther 2009; 17: 309–317.
29 Wallace HA, Marques-Kranc F, Richardson M, Luna-Crespo F, Sharpe JA, Hughes J et al. Manipulating the mouse genome to engineer precise functional syntenic replacements with human sequence. Cell 2007; 128: 197–209.
30 Kazuki Y, Hiratsuka M, Takiguchi M, Osaki M, Kajitani N, Hoshiya H et al. Complete genetic correction of ips cells from Duchenne muscular dystrophy. Mol Ther 2010; 18: 386–393.
31 Kinoshita Y, Kamitani H, Mamun MH, Wasita B, Kazuki Y, Hiratsuka M et al. A gene delivery system with a human artificial chromosome vector based on migration of mesenchymal stem cells towards human glioblastoma HTB14 cells. Neurol Res 2010; 32: 429–437.
32 Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 2006; 126: 663–676.
33 Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, Tomoda K et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 2007; 131: 861–872.
34 Kazuki Y, Kimura M, Nishigaki R, Kai Y, Abe S, Okita C et al. Human chromosome 21q22.2-qter carries a gene(s) responsible for downregulation of mlc2a and PEBP in Down syndrome model mice. Biochem Biophys Res Commun 2004; 317: 491–499.
35 Kouprina N, Ebersole T, Koriabine M, Pak E, Rogozin IB, Katoh M et al. Cloning of human centromeres by transformation-associated recombination in yeast and generation of functional human artificial chromosomes. Nucleic Acids Res 2003; 31: 922–934.
36 Ebersole T, Okamoto Y, Noskov VN, Kouprina N, Kim JH, Leem SH et al. Rapid generation of long synthetic tandem repeats and its application for analysis in human artificial chromosome formation. Nucleic Acids Res 2005; 33: e130.
37 Kouprina N, Larionov V. Selective isolation of genomic loci from complex genomes by transformation-associated recombination cloning in the yeast Saccharomyces cerevisiae. Nat Protoc 2008; 3: 371–377.
38 Gordon D, Abajian C, Green P. Consed: a graphical tool for sequence finishing. Genome Res 1998; 8: 195–202.
39 Hattori M, Tsukahara F, Furuhata Y, Tanahashi H, Hirose M, Saito M et al. A novel method for making nested deletions and its application for sequencing of a 300 kb region of human APP locus. Nucleic Acids Res 1997; 25: 1802–1808.
40 McMurray AA, Sulston JE, Quail MA. Short-insert libraries as a method of problem solving in genome sequencing. Genome Res 1998; 8: 562–566.
41 Trowell HE, Nagy A, Vissel B, Choo KH. Long-range analyses of the centromeric regions of human chromosomes 13, 14 and 21: identification of a narrow domain containing two key centromeric DNA elements. Hum Mol Genet 1993; 2: 1639–1649.

Supplementary Information accompanies the paper on Gene Therapy website (http://www.nature.com/gt)