

# Method to Assemble Genomic DNA Fragments or Genes on Human Artificial Chromosome with Regulated Kinetochores Using a Multi-Integrase System

Nicholas C. O. Lee,<sup>†,‡,§</sup> Jung-Hyun Kim,<sup>†,‡</sup> Nikolai S. Petrov,<sup>†</sup> Hee-Sheung Lee,<sup>†</sup> Hiroshi Masumoto,<sup>‡</sup> William C. Earnshaw,<sup>§</sup> Vladimir Larionov,<sup>†</sup> and Natalay Kouprina<sup>\*,†</sup>

<sup>†</sup>Developmental Therapeutics Branch, National Cancer Institute, Bethesda, Maryland 20892, United States

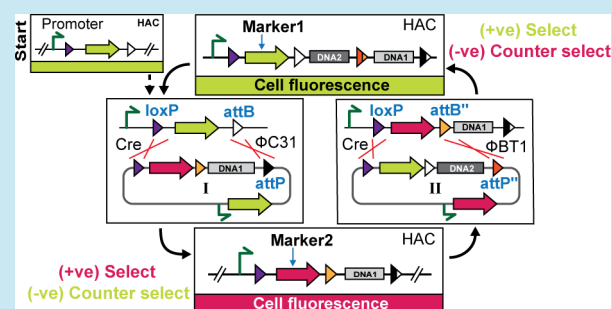
<sup>‡</sup>Laboratory of Cell Engineering, Department of Frontier Research, Kazusa DNA Research Institute, 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan

<sup>§</sup>Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3JR, Scotland

## Supporting Information

**ABSTRACT:** The production of cells capable of carrying multiple transgenes to Mb-size genomic loci has multiple applications in biomedicine and biotechnology. In order to achieve this goal, three key steps are required: (i) cloning of large genomic segments; (ii) insertion of multiple DNA blocks at a precise location and (iii) the capability to eliminate the assembled region from cells. In this study, we designed the iterative integration system (IIS) that utilizes recombinases Cre,  $\Phi$ C31 and  $\Phi$ BT1, and combined it with a human artificial chromosome (HAC) possessing a regulated kinetochores (alphoid<sup>tetO</sup>-HAC). We have demonstrated that the IIS-alphoid<sup>tetO</sup>-HAC system is a valuable genetic tool by reassembling a functional gene from multiple segments on the HAC. IIS-alphoid<sup>tetO</sup>-HAC has several notable advantages over other artificial chromosome-based systems. This includes the potential to assemble an unlimited number of genomic DNA segments; a DNA assembly process that leaves only a small insertion (<60 bp) scar between adjacent DNA, allowing genes reassembled from segments to be spliced correctly; a marker exchange system that also changes cell color, and counter-selection markers at each DNA insertion step, simplifying selection of correct clones; and presence of an error proofing mechanism to remove cells with misincorporated DNA segments, which improves the integrity of assembly. In addition, the IIS-alphoid<sup>tetO</sup>-HAC carrying a locus of interest is removable, offering the unique possibility to revert the cell line to its pretransformed state and compare the phenotypes of human cells with and without a functional copy of a gene(s). Thus, IIS-alphoid<sup>tetO</sup>-HAC allows investigation of complex biomedical pathways, gene(s) regulation, and has the potential to engineer synthetic chromosomes with a predetermined set of genes.

**KEYWORDS:** human artificial chromosome, HAC, iterative integration system, IIS, DNA assembly, synthetic biology



Since their development, HACs or human artificial chromosomes have been considered to be a promising system for gene delivery and expression, with the potential to overcome many problems caused by the use of virus-based gene transfer systems. Indeed, HACs are stably maintained as single copy episomes. Their use avoids the limited cloning capacity, the lack of copy number control and potential insertional mutagenesis due to integration into host chromosomes that has hampered the use of viral vectors.<sup>1–7</sup> There are two types of HACs in general use: “top down” HACs constructed by truncation of the natural chromosomes and “bottom up” HACs generated from BACs carrying 30–200 kb DNA from natural higher-order centromeric DNA repeats. Such BACs are substrates for HAC formation in a process that is accompanied by 20–30-fold multimerization of the input BAC DNA in human cells. Over the past 20 years, many groups have reported the successful generation of HACs of both types<sup>1–7</sup>

(and references therein), and importantly all HACs constructed so far have constitutive kinetochores.<sup>8,9</sup> In the past few years, the amount of research based on the HAC-vectors has significantly increased due to the engineering of HACs with a single loxP gene integration site<sup>3,10,11</sup> (and references therein). Such research has included studies on the potential of HACs for gene function analysis, cell reprogramming and animal transgenesis.<sup>1–7,12–22</sup>

Our group constructed the alphoid<sup>tetO</sup>-HAC<sup>23</sup> from a synthetic alpha-satellite (alphoid) DNA array with precisely defined DNA sequence starting from a 343 bp dimer amplified by the RCA-TAR method<sup>24</sup> up to 50 kb in size. This HAC includes approximately 6000 copies of the 42 bp tetracycline

Received: June 14, 2017

Published: August 11, 2017

operator (tetO) sequence incorporated into each alphoid dimer.<sup>25</sup> Because the tetO sequence is bound with high affinity and specificity by the Tet repressor (TetR), the 1.1 Mb-size alphoid array in this HAC can be targeted efficiently with TetR fusion proteins that allow the specific manipulation of chromatin within the functional kinetochore. This provides a highly versatile model system for the study of centromere and its impact on kinetochore structure and function in human cells.<sup>8,9</sup> In addition, this feature of alphoid<sup>tetO</sup>-HAC gives the unique possibility to eliminate the HAC from cells, allowing workers to compare the phenotypes of human cells with and without a functional copy of a target gene inserted into the HAC.<sup>12,13</sup> Such rigorous control is important for proper interpretation of gene complementation analysis and studies of new gene function. Recently the alphoid<sup>tetO</sup>-HAC was applied to investigate the problem of chromosome instability (CIN), involving the unequal distribution of chromosomes to daughter cells during mitosis which is observed in the majority of solid tumors. While CIN acts as a driver of cancer genome evolution and tumor progression, recent findings point to the existence of a threshold level beyond which CIN becomes a barrier to tumor growth and, therefore, can be exploited therapeutically.<sup>26,27</sup> Newly developed alphoid<sup>tetO</sup>-HAC-based assays allow for a quick and efficient screening of hundreds of drugs to identify those affecting chromosome mis-segregation and also to rank compounds with the same or similar mechanism of action based on their effect on the rate of chromosome loss.<sup>28–30</sup> The fact that the HAC is nonessential in the absence of Blasticidin selection greatly simplifies the interpretation of these assays. Similarly, the alphoid<sup>tetO</sup>-HAC-based approach was used to identify new human CIN genes,<sup>31</sup> mutations in which are thought to be an early event in tumor development predisposing cells to the accumulation of genetic changes leading to progression to a cancerous state.

A few years ago several laboratories suggested using artificial chromosomes (AC) to assemble entire genomic loci using a multi-integrase system.<sup>32–36</sup> Such an AC-based approach opens new horizons for synthetic biology, especially in higher eukaryotes. In this study, we designed an iterative integration system (IIS) in such a way that potentially renders it capable of accepting an unlimited number of DNA fragments. In general, IIS is a method to construct large segments of transgenic DNA in a vertebrate host genome. Our IIS method is based on three recombinase enzymes, *i.e.*, Cre,  $\Phi$ C31 and  $\Phi$ BT1. Cre is a bidirectional enzyme that catalyzes the recombination between two substrate loxP sites and generates two product loxP sites. In contrast, in the absence of an additional protein Xis, recombinases  $\Phi$ C31 and  $\Phi$ BT1 are unidirectional enzymes that recombine an attachment bacteria (*attB*) site and an attachment phage (*attP*) site to produce *attR* and *attL* sites that are not substrates for further reaction. A novel IIS system was combined with alphoid<sup>tetO</sup>-HAC.

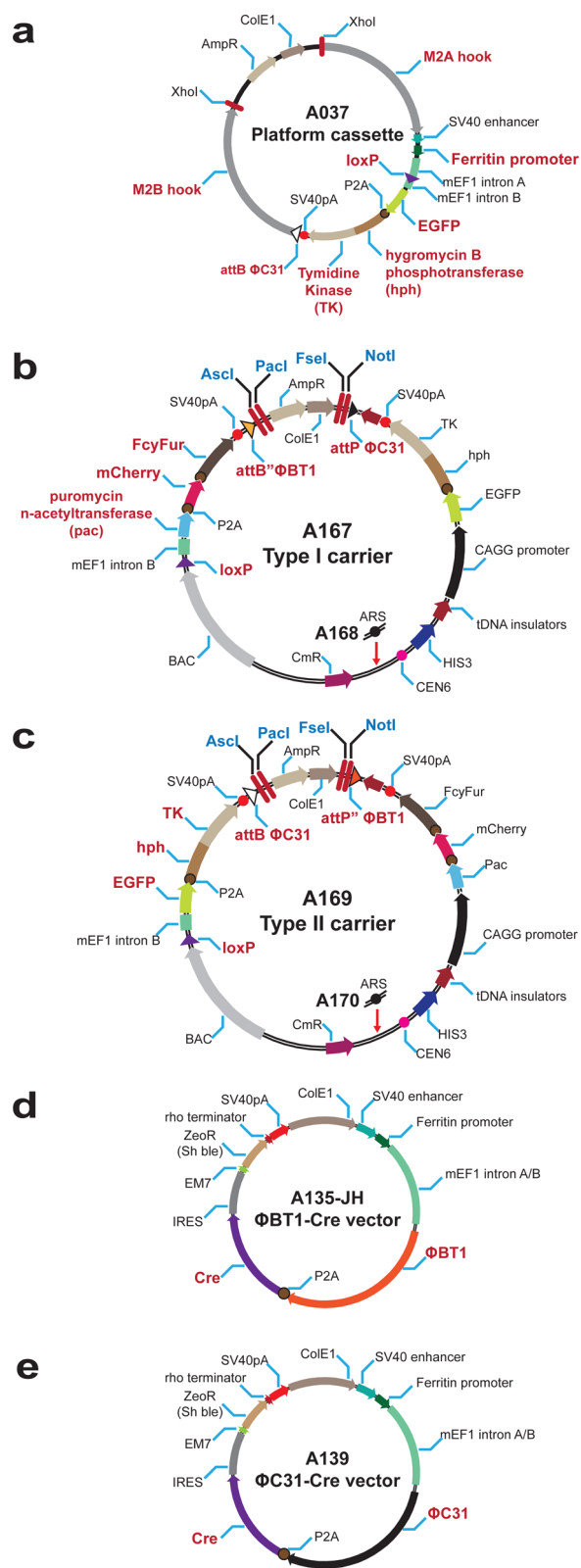
The IIS-alphoid<sup>tetO</sup>-HAC system can be used to assemble and deliver large DNA fragments that include groups of genes for functional studies. In future, a new HAC-based multi-integrase system carrying multiple genes may be a unique tool for treatment of multigene genetic disorders and for transgenesis experiments with the purpose to understand complex diseases. In addition, the IIS-alphoid<sup>tetO</sup>-HAC system provides a potentially valuable genetic tool to engineer synthetic chromosomes with a predetermined set of genes in order to investigate complex biomedical pathways and gene(s) regulation.

## RESULTS

**Construction of Plasmids for the IIS-alphoid<sup>tetO</sup>-HAC System.** Five basic plasmids were constructed for the IIS-alphoid<sup>tetO</sup>-HAC system: the platform cassette A037 on which DNA fragments/genes can be assembled, two carrier vectors used to deliver the transgenic DNA fragments containing promoterless compound markers called either PCF (Pac-mCherry-FcYFur) (Type I A167 plasmid) or GHT (eGFP-hph-TK) (Type II A169 plasmid) and two recombinase expression plasmids that express either Cre and  $\Phi$ BT1 (A135-JH) or Cre and  $\Phi$ C31 (A139) (Figure 1) (for construction details, see Methods and Figures S1–S5).

The platform cassette A037 consists of the SFM promoter (SV40 enhancer plus Feritin promoter), a split mouse elongation factor 1 (mEF1) intron containing a loxP site within, the GHT marker and the *attB* site of  $\Phi$ C31 (Figure 1a). The platform cassette was targeted into the unique human Ch13 genomic segment present within alphoid<sup>tetO</sup>-HAC<sup>12</sup> using homologous recombination in a highly recombinogenic chicken B-lymphoma DT40 cells (Figure 2a). Before targeting, the platform cassette was XhoI-linearized to release 5' M2A and 3' M2B hook sequences of 3.5 kb and 3.9 kb in size, respectively, which have homology to the Chr13 genomic segment (see Methods for details). Four pairs of diagnostic primers were used to verify correct insertion of the platform cassette into the Ch13 segment. The expected size products of 4.4 kb and 3.9 kb with B128/B124 and B129/B126 pairs of primers, correspondingly, and 4.6 kb and 4.4 kb with the B034/B126 and B034/B127 pairs of primers, correspondingly, (Table S1) confirms correct recombination between the M2A and M2B hook sequences and homologous sequences of the Ch13 segment (Figure 2b). After insertion of the platform cassette into the HAC, the cells became green (GFP+), resistant to Hygromycin and sensitive to Ganciclovir. Next, the alphoid<sup>tetO</sup>-HAC carrying the platform cassette was moved from chicken DT40 cells to hamster CHO cells by Microcell-Mediated Chromosome Transfer (MMCT) (Figure 2a). A FISH image of the alphoid<sup>tetO</sup>-HAC in CHO cells (Clone #CHO BH3:37) is shown in Figure 2c.

The Type I and Type II carrier vectors, A167 and A169, correspondingly, were constructed to deliver a transgenic DNA segment of interest into the platform cassette inserted into the HAC (Figures 1b, 1c). These carrier vectors are YAC-BAC shuttle vectors and, therefore, can propagate in a single molecule state within *E. coli* as a bacterial artificial chromosome (BAC) with chloramphenicol selection and in *S. cerevisiae* as a yeast artificial chromosome (YAC) with the yeast HIS3 gene as a selectable marker. Insertion of a transgenic DNA segment into the carrier vector can be done *via* DNA ligation or yeast-based transformation-associated recombination (TAR) cloning.<sup>37–41</sup> The Type I carrier vector A167 contains in 5'–3' order a loxP site, a promoterless PCF marker, an *attB*'  $\Phi$ BT1 site, a cloning site for DNA insertion, an *attP*  $\Phi$ C31 site, a GHT marker under a CAGG promoter flanked by tDNA insulators<sup>42,43</sup> and a YAC-BAC backbone (Figure 1b). The Type II carrier vector A169 contains a loxP site, a promoterless GHT marker, an *attB*  $\Phi$ C31 site, a cloning site for DNA insertion, an *attP*'  $\Phi$ BT1 site, a PCF marker under a CAGG promoter flanked by tDNA insulators and a YAC-BAC backbone (Figure 1c). For the purpose of TAR cloning<sup>37,44</sup> short mammalian genomic DNA segments that do not have yeast ARS-like sequences for a proper propagation in yeast



**Figure 1.** Scheme of plasmids used in the IIS-alphoid<sup>tet<sup>O</sup></sup>-HAC system. (a) The platform cassette A037. (b) Type I A167 and Type I-ARS A168 carrier plasmids to deliver a genomic DNA fragment. (c) Type II A169 and Type II-ARS A170 carrier plasmids to deliver a genomic fragment. (d) A135-JH plasmid expressing  $\Phi$ BT1 integrase and Cre recombinase. (e) A139 plasmids expressing  $\Phi$ C31 integrase and Cre recombinase. Restriction sites used in this work are marked in blue.

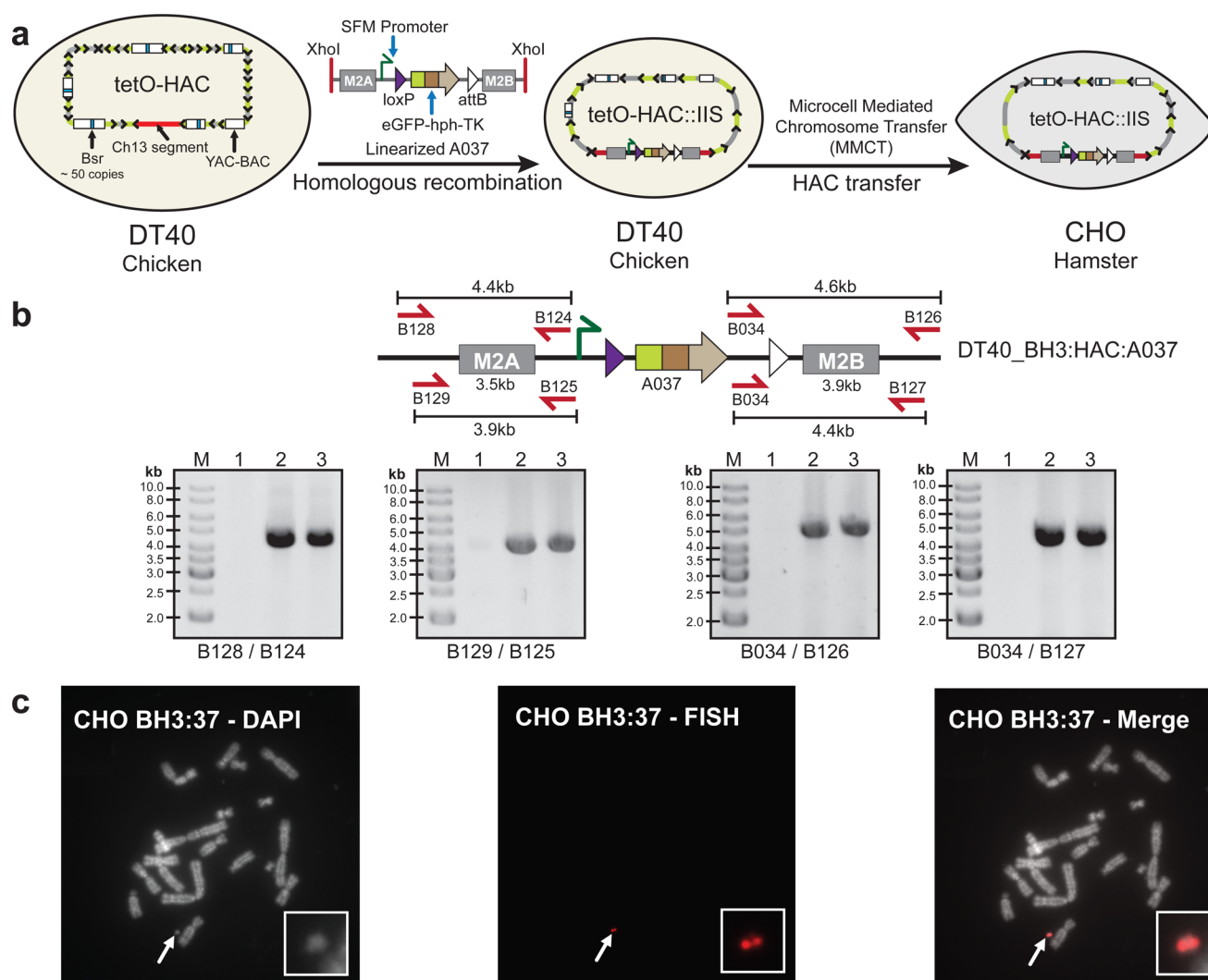
cells, a variant of each carrier vector was made containing yeast origin of replication (ARS), *i.e.*, Type I-ARS (A168) and Type II-ARS (A170) (Figures 1b, 1c and Figures S6, S7).

Two expression plasmids were constructed for the IIS-alphoid<sup>tet<sup>O</sup></sup>-HAC system. In these plasmids, a P2A self-cleaving peptide is used to translationally link the expression of two recombinases so that the two recombinases are expressed in equal ratio. The plasmid A135-JH expresses Cre recombinase and  $\Phi$ BT1 integrase (Figure 1d) while the plasmid A139 expresses Cre recombinase and  $\Phi$ C31 integrase (Figure 1e). These plasmids also contain a Zeocin marker that is transcriptionally linked to the recombinase expression *via* an internal ribosomal entry site (IRES), allowing selection of these plasmids if desired.

**Description of the IIS-alphoid<sup>tet<sup>O</sup></sup>-HAC System.** The IIS-alphoid<sup>tet<sup>O</sup></sup>-HAC system works as follows. It starts with CHO cells containing alphoid<sup>tet<sup>O</sup></sup>-HAC bearing the platform cassette A037 (Figure 3a). As the GHT marker is expressed, the cells are green (GFP), Hygromycin resistant (hph) and are killed upon exposure to Ganciclovir (TK). Next, these cells are cotransformed with two plasmids, *i.e.*, the A139 plasmid that expresses Cre recombinase and  $\Phi$ C31 integrase and the Type I carrier vector A167 that contains a transgenic DNA segment of interest (DNA1). Expression of Cre and  $\Phi$ C31 promotes two recombination events (loxP–loxP and  $\Phi$ C31attB–attP) between the Type I vector and the platform cassette (Figure 3b). The order of recombination is unimportant as the final product is identical. The SFM promoter within the platform cassette now drives the promoterless PCF marker and contains the adjacent attB''  $\Phi$ BT1 site from the Type I vector, as the original GHT marker is replaced with the PCF marker. This is accompanied by integration of the DNA segment of interest (DNA1) into the platform cassette within the HAC and deletion of all other components of the Type I vector. As a result, the cells that successfully complete both recombination reactions lose green fluorescence and sensitivity to Ganciclovir and gains red fluorescence, resistance to Puromycin and sensitivity to 5-Fluorocytosine.

An error detection mechanism was built into the IIS-alphoid<sup>tet<sup>O</sup></sup>-HAC system because the recombinase-mediated reactions can fail to go to completion. Furthermore, growth, maintenance, screening and storage of candidate vertebrate colonies are far slower, more labor intensive and space limited than either in yeast or bacteria. Thus, it was desirable to remove as many faulty colonies as possible by drug selection and keep only the best candidates for subsequent detailed characterization.

If either of the two recombination reactions fails, this failure event can be selected against and screened out by the error proofing design of the IIS-alphoid<sup>tet<sup>O</sup></sup>-HAC system (Figure 4). As illustrated, the backbone of each carrier vector has its own constitutively active compound marker. Hence, if recombination by Cre fails but  $\Phi$ C31 occurs (Figure 4a), the A167 Type I carrier vector will integrate into the platform cassette but the PCF marker it carries will remain promoterless. Cells carrying this error are removed by Puromycin selection and counter-selection with Ganciclovir. Alternatively, if recombination by  $\Phi$ C31 fails but Cre occurs (Figure 4b), the Type I vector A167 will integrate into the construction platform and the SFM promoter will capture the PCF marker, leaving the original GHT marker promoterless. However, the backbone of the Type I vector is retained, and a fully expressed GHT marker under the CAGG promoter remains. Hence, cells generated by

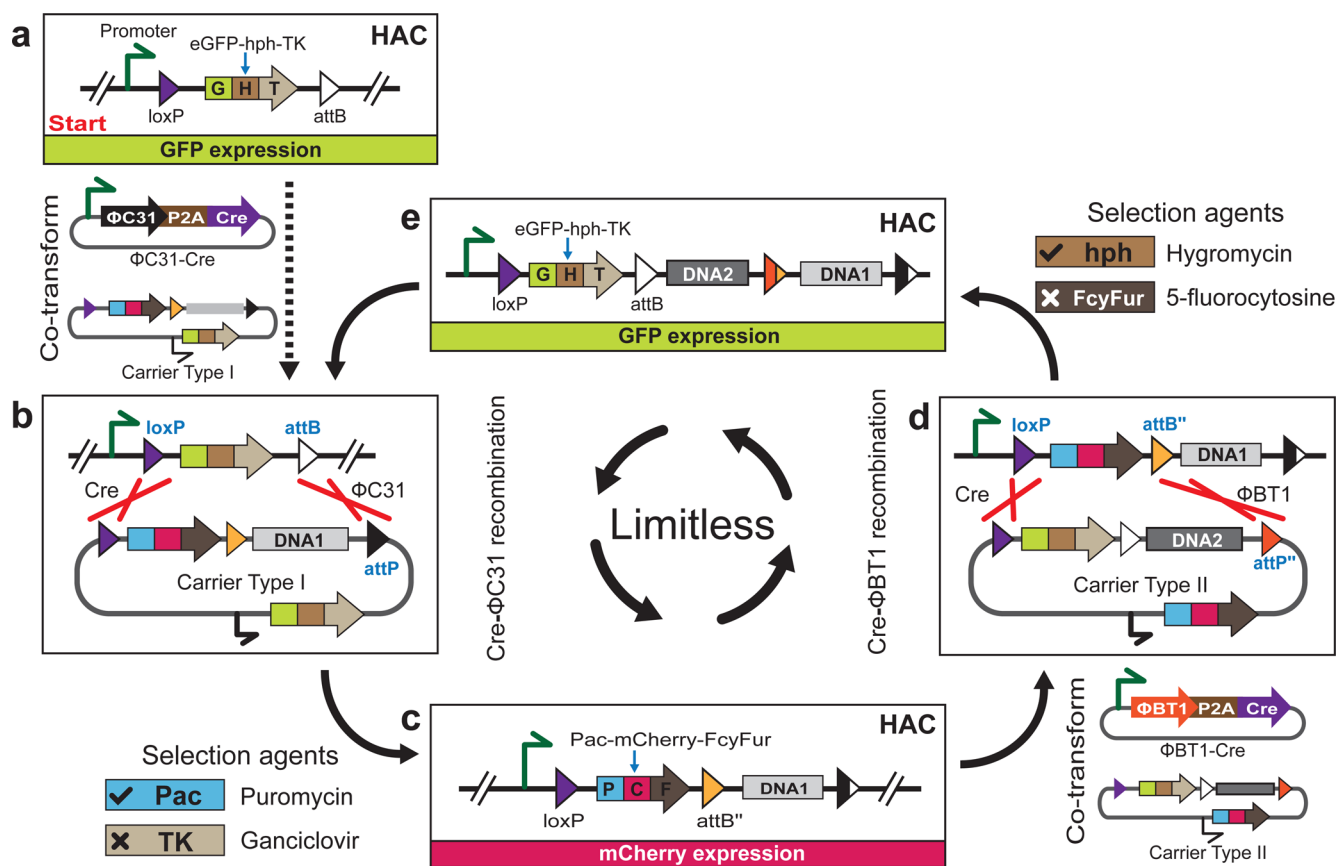


**Figure 2.** Insertion of the platform cassette into  $\text{alphoid}^{\text{tetO}}$ -HAC in hamster CHO cells. (a) The XhoI-linearized platform cassette A037 was inserted into the Ch13 genomic segment present within  $\text{alphoid}^{\text{tetO}}$ -HAC by homologous recombination in DT40 cells. Then the  $\text{alphoid}^{\text{tetO}}$ -HAC carrying the platform cassette was MMCT transferred from chicken DT40 cells to hamster CHO cells. (b) Diagnostic PCRs to verify correct targeting of the platform cassette A037 into the Ch13 region present within the HAC. Two pairs of diagnostic primers, B128/B124 and B129/B126, amplified the expected products of 4.4 kb and 3.9 kb in size, respectively, confirming correct recombination between the M2A hook sequence of A037 and a homologous sequence of the Ch13 segment. Two pairs of diagnostic primers, B034/B126 and B034/B127, amplified the expected products of 4.6 kb and 4.4 kb in size, respectively. This confirmed the correct recombination between the M2B hook sequence of A037 and a homologous sequence of the Ch13 segment. M, a GeneRuler DNA ladder mix (Fermentas). Lane 1, DT40 cells carrying  $\text{alphoid}^{\text{tetO}}$ -HAC without A037 insertion (a negative control); Lane 2, targeted DT40: BH3:A037 clone #48; Lane 3, targeted DT40: BH3:A037 clone #49. (c) FISH analysis of the  $\text{alphoid}^{\text{tetO}}$ -HAC carrying the platform cassette in CHO cells. Arrows indicate to the HAC visualized with the BAC specific probe (in red).

such a failure event are Puromycin resistant and have both red and green fluorescence. These types of cells can be removed by counter selection using Ganciclovir against the thymidine kinase component of the GHT marker. In order to avoid the loss of TK gene activity by silencing, the GHT marker is protected by flanking murine tDNA insulators.

Once the first round of recombination is completed (DNA1), the second round of DNA integration (DNA2) can be started. The platform cassette in the HAC now contains a loxP site, an expressed PCF marker and an  $\text{attB}''$   $\Phi$ BT1 site (Figure 3c). Expression from the PCF marker gives the cells red fluorescence, resistance to Puromycin and sensitivity to 5-Fluorocytosine. Next, the cells are cotransformed with plasmid A135-JH expressing Cre recombinase and  $\Phi$ BT1 integrase and the Type II carrier vector A169 containing a second transgenic DNA segment of interest (DNA2). Cre and  $\Phi$ BT1 expression

causes two recombination events ( $\text{loxP}$ – $\text{loxP}$  and  $\Phi$ BT1  $\text{attB}''$ – $\text{attP}''$ ) between the Type II vector and the platform cassette (Figure 3d). This leads to the replacement of the PCF marker by the GHT marker and an  $\text{attB}''$   $\Phi$ BT1 site, followed by the insertion of the second DNA segment of interest (DNA2) from the Type1 carrier vector. As a result, the platform cassette in the HAC will contain a loxP site, an expressed GHT marker and an  $\text{attB}''$   $\Phi$ BT1 site (Figure 3e). A small insertion (<60 bp) scar comprising the recombination product of  $\text{attB}/\text{attP}$ , the  $\text{attR}$  site is left between the first (DNA1) and the second (DNA2) DNA segments of interest. Selection with Hygromycin and counter selection with 5-Fluorocytosine ensures that only cells that have correctly undergone the second round of assembly will survive (Figures 4c, 4d). Untransformed parental cells and cells with incomplete recombination are killed by this double selection.



**Figure 3.** Scheme of DNA segment integration by the iterative integration system (IIS). (a) The starting platform cassette. Cells express the GHT marker, *i.e.*, a green fluorescence protein (GFP). Also, they are hygromycin resistant (hph) and Ganciclovir sensitive (TK). (b) Recombination between a Type I carrier vector and a platform cassette by Cre recombinase and ΦC31 integrase. The GHT marker is replaced by the PCF marker and the first DNA segment of interest is integrated into the platform cassette (DNA1). The integration event is selected using Puromycin and Ganciclovir. (c) A structure of the platform cassette after the 1st round of integration. The PCF marker is expressed. Therefore, the cells have red fluorescence (mCherry), Puromycin resistance (Pac) and 5-Fluorocytosin sensitivity (FcyFur). (d) Recombination between a Type II carrier vector and a platform cassette by Cre recombinase and ΦBT1 integrase. The PCF marker is replaced by the GHT marker and the second DNA segment of interest is integrated into the platform cassette (DNA2). The integration event is selected using Hygromycin and 5-Fluorocytosine. (e) A structure of the platform cassette after the second round of recombination. The cells express the GHT marker, *i.e.*, a green fluorescence protein (GFP). They become again Hygromycin resistant (hph) and Ganciclovir sensitive (TK). This structure is identical to the starting cassette aside from the integration of DNA segments of interest, DNA1 and DNA2.

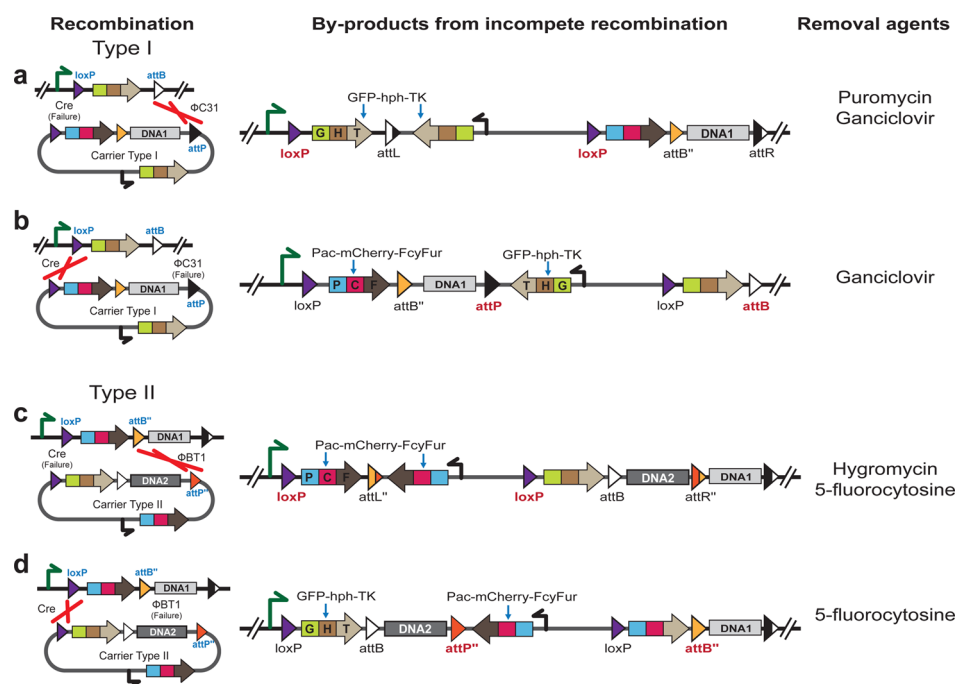
After two rounds of recombination, the construction platform is once again where it started, with the exception that two DNA segments (DNA1 and DNA2) of interest have now been integrated into the HAC (Figure 3e). The GHT marker is expressed and the cells have once again green fluorescence, Hygromycin resistance and sensitivity to Ganciclovir. Further rounds of DNA fragment insertions can be repeated indefinitely as required.

**Proof of Site-Specific Recombination Using the IIS- $\alpha$ phoid<sup>tetO</sup>-HAC System.** First, we tested whether the ΦBT1 and ΦC31 integrases and Cre recombinase were active and could mediate proper integration into  $\alpha$ phoid<sup>tetO</sup>-HAC carrying the platform plasmid. For this purpose, control experiments were performed using empty vectors. The recombinant assay plasmids were transfected into hamster CHO cells containing the HAC in the combinations A167 plus A139 or A169 plus A135-JH. Each step of insertion was confirmed by PCR.

At the beginning of the experiments, hamster CHO cells carrying the  $\alpha$ phoid<sup>tetO</sup>-HAC bearing the platform cassette were green (GFP+) (Clone #CHO BH3:37). PCR of genomic DNA isolated from the initial CHO cells carrying the HAC gave the expected 778 bp product with the forward primer for

the thymidine kinase (TK) gene (B075) and reverse primer for the M2B hook sequence (B681) and no product with the forward primer for the cytidine deaminase gene (FcyFur) (B485) and the reverse primer for the hook M2B sequence (B681) (Figures S8a, S8b; Lane 1) (Table S1).

For the first round of insertion, CHO cells carrying the  $\alpha$ phoid<sup>tetO</sup>-HAC bearing the platform cassette were cotransfected by Type I carrier plasmid A167 and plasmid A139 expressing ΦC31 integrase and Cre recombinase. The cells were cultured in Puromycin/Blasticidin S media. After 10 days of selection, five colonies possessing red but not green fluorescence were picked up and cultured in individual wells in Puromycin/Ganciclovir/Blasticidin S media. In three colonies, insertion of the 2.7 kb Ampicillin resistant gene plus pBR322 DNA fragment from the plasmid A167 into the platform cassette was confirmed by PCR with a pair of corresponding primers: forward B485 and reverse B681 (Table S1). Cells bearing this insertion gave the expected 4.1 kb product with B485/B681 primers but no product with B075/B681 primers (Figures S8a, 8b; Lanes 2). Clone #E10 was chosen for further experiments.



**Figure 4.** Error proofing design of the IIS-alphoid<sup>tet<sup>O</sup></sup>-HAC. (a,b) Products of incomplete recombination between Type I carrier plasmid with an active GHT cassette. (c,d) Products of incomplete recombination between Type II carrier plasmid with an active PCF cassette. The selection agents to remove each misassembled product are listed.

For the second round of insertion, the cells were cotransfected by Type II carrier plasmid A169 and A135-JH plasmid expressing  $\Phi$ BT1 integrase and Cre recombinase. The cells were cultured in Hygromycin/Blasticidin S media. After 10 days of selection, ten colonies expressing green but not red fluorescence were picked up and expanded in media containing Hygromycin/5-Fluorocytidine/Blasticidin S. For nine colonies, PCR of genomic DNA isolated from these cells gave the expected 6.2 kb PCR product with B075/B681 primers but no product with B485/B681 primers. This confirmed the correct insertion of the 2.7 kb fragment carried by plasmid A169 (Figures S8a, S8b; Lanes 3). Clone #E10-7 was chosen for further experiments.

For the third round of insertion, the cells were cotransfected and selected for as in the first round. A total number of ten clones were selected. PCR of genomic DNA isolated from two clones, Clone #E10-7-3 and Clone #E10-7-1, with B485/B681 primers gave the expected 9.6 kb product while there was no product with B075/B681 primers. This confirmed correct insertion of the 2.7 kb fragment carried by the A167 plasmid. (Figures S8a, S8b; Lanes 4 and 5).

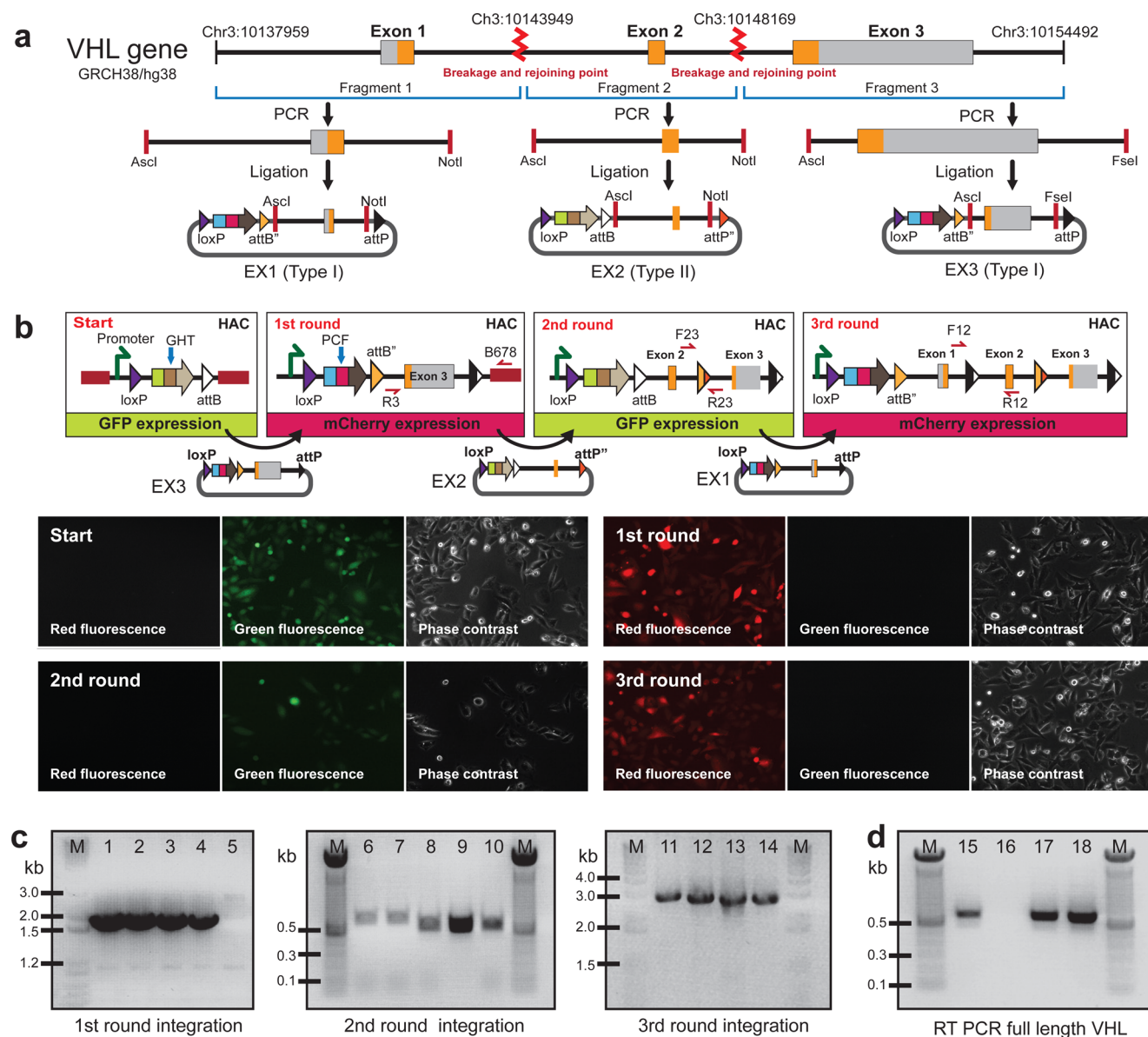
These results showed that the recombinases are functional and that site-specific insertions into IIS-alphoid<sup>tet<sup>O</sup></sup>-HAC could be achieved.

**VHL Gene Reconstruction Using IIS-alphoid<sup>tet<sup>O</sup></sup>-HAC System.** As a proof of principle, we applied the IIS-alphoid<sup>tet<sup>O</sup></sup>-HAC system to reconstruct the VHL gene. Briefly, the human VHL gene is ~17 kb in size and contains three exons. It is located on chromosome 3 (positions 10137959–10154492; GHCH38/hg38). Mutations in the gene are associated with Von Hippel-Lindau (VHL) syndrome that is a dominantly inherited hereditary cancer syndrome predisposing to a variety of malignant and benign tumors of the eye, brain, spinal cord, kidney, pancreas and adrenal glands.

For the experiments, we used the full-length VHL gene TAR-isolated previously from the total human genomic DNA as a ~25 kb YAC/BAC molecule.<sup>13</sup> Three fragments containing exon 1, exon 2 or exon 3 of the VHL gene were PCR-amplified from the TAR/YAC/BAC clone using specific primers (Table S1). The first *Asc*I-*Not*I fragment of 5990 bp in size (positions Chr3:10137959–10143949; GRCH38/hg38) containing exon 1 along with the VHL promoter and the third *Asc*I-*Fse*I fragment of 6323 bp in size (positions Chr3:10148169–10154492) containing exon 3 were ligated with the Type I carrier vector A168 (Pac-mCherry-FcyFur) that had been digested by *Asc*I/*Not*I or *Asc*I/*Fse*I, correspondingly. The second *Asc*I-*Not*I fragment of 4221 bp in size (positions Chr3:10143949–10148169; GRCH38/hg38) containing exon 2 was ligated with the Type II carrier vector A170 (GFP-hph-TK) that was digested by *Asc*I/*Not*I (Figure 5a). Three rounds of insertion were performed to assemble the full-length VHL gene using the IIS-alphoid<sup>tet<sup>O</sup></sup>-HAC system.

**Round 1.** A168 vector containing exon 3 and A139 vector expressing  $\Phi$ C31 integrase and Cre recombinase were cotransfected into hamster CHO cells propagating the alphoid<sup>tet<sup>O</sup></sup>-HAC bearing the platform cassette A037. Cells with this HAC are originally green (GFP+). After the first round, the cells switched to red (mCherry+), Puromycin and 5-Fluorocytidine resistant and carried a modified alphoid<sup>tet<sup>O</sup></sup>-HAC bearing the inserted exon 3 (Figure 5b). A total of 14 clones were obtained after positive selection, with 4 remaining after negative selection: Clone #1, Clone #11, Clone #12 and Clone #14. Clone #14 was chosen for the second round of insertion.

**Round 2.** The A170 vector containing exon 2 and the A135-JH vector expressing  $\Phi$ BT1 integrase and Cre recombinase were cotransfected into Clone #14e CHO cells carrying the alphoid<sup>tet<sup>O</sup></sup>-HAC bearing the inserted exon 3. After the second round, the cells switched back to green (GFP+), Hygromycin



**Figure 5.** VHL gene reconstruction using the IIS-alphoid<sup>tet<sup>O</sup></sup>-HAC system. (a) Construction of Type I A168 and Type II A170 carrier plasmids containing fragments with exon1 or exon 2 or exon 3 of the VHL gene. Positions of the VHL fragments in the human genome (GRCH38/hg38) are indicated. (b) Three rounds of insertion of the VHL fragments into alphoid<sup>tet<sup>O</sup></sup>-HAC carrying the platform cassette are shown. Representative images of changing cell color after each round are shown. (c) Diagnostic PCRs for insertion of the VHL fragments after each round of recombination. *Round 1*: Lanes 1 to 5, diagnostic PCR for the fragment 3 with the expected size 1.7 kb. Lane 1, Clone #1; Lane 2, Clone #11; Lane 3, Clone #12; Lane 4, Clone #14; Lane 5, genomic DNA from CHO cells as a negative control. *Round 2*: Lanes 6–10, diagnostic PCR for junction between fragments 2 and 3. The expected size is 688 bp. The expected size for a negative control is 608 bp. Lane 6, Clone #14–4; Lane 7, Clone #14–12; Lane 8, alphoid<sup>tet<sup>O</sup></sup>-HAC carrying the full-length human VHL gene in CHO cells; Lane 9, a TAR/BAC clone containing the VHL gene; Lane 10, human genomic DNA as a positive control. *Round 3*: Lanes 11–14, diagnostic PCR for junction between fragments 1 and 2. The expected size is 2788 bp. The expected size for a negative control is 2716 bp. Lane 11, Clone #14–12–3; Lane 12, alphoid<sup>tet<sup>O</sup></sup>-HAC carrying the full-length VHL gene in CHO cells; Lane 13, a TAR/BAC clone containing the VHL gene; Lane 14, human genomic DNA as a positive control. (d) RT-PCR for mRNA of VHL. The expected product size is 642 bp. Lane 15, Clone #14–12–3; Lane 16, genomic DNA from CHO cells as a negative control; Lane 17, human genomic DNA from HeLa cell line as a positive control; Lane 18, human genomic DNA from HT1080 cell line as a positive control.

and 5-Fluorocytidine resistant and carried the alphoid<sup>tet<sup>O</sup></sup>-HAC with inserted exon 3 and exon 2 (Figure 5b). After positive selection, 20 clones were obtained, of which 2 were left after negative selection: Clone #14–4 and Clone #14–12. Clone #14–12 was chosen for the third round of insertion.

**Round 3.** The A168 vector containing exon 1 and the A139 vector expressing ΦC31 integrase and Cre recombinase were cotransfected into CHO cells carrying the alphoid<sup>tet<sup>O</sup></sup>-HAC bearing exons 2 and 3. After the third round, the cells switched

to red again (mCherry+), Puromycin, Gancyclovir and 5-Fluorocytidine resistant. These cells carry an alphoid<sup>tet<sup>O</sup></sup>-HAC with all three exons inserted (Figure 5b). After positive selection, 3 clones were obtained, with 2 remaining after negative selection: Clone #14–12–2 and Clone #14–12–3. Insertion after each round was confirmed by PCR (Figure 5c) with corresponding primers (first round R3/B678; second round F23/R23; third round F12/R12) (Table S1).

Reconstruction of a functional human VHL gene was confirmed by RT-PCR of the reconstituted RNA transcript from Clone #14–12–3 (Figure 5d) with the primers VHLstart-F1/VHLEXON-3R (Table S1). The PCR product was sequenced and found to match the correct human VHL sequence (Figure S9). Note that in order for this mRNA reassembly to work, the inserted DNA segments were designed so that the ~60 bp recombination sites were located within introns, and were therefore spliced out of the mature RNA transcript.

## DISCUSSION

At present, there are many methods to produce transgenic cells for functional studies of genes. The most common methods rely on either transfection of BAC DNA carrying a gene of interest into the host cells or transduction with viruses. However, these methods lead to random integrations into host chromosomes. As a result of that, the expression level of genes varies greatly due to position effects and the number of copies integrated. In addition, the use of viruses limits the size of a gene that can be successfully transduced. Another popular approach is based on integration of a gene into a “hot spot” of a mammalian genome. In this case, homologous recombination for targeting gene integration is very specific due to the usage of a bacteriophage P1-derived Cre recombinase or  $\Phi$ C31 integrase. However, the efficiency of gene integration remains low. Furthermore, integration of several genes into the same “hot spot” is very difficult, if at all possible.

Several years ago five groups developed a principally new approach to produce transgenic cells.<sup>32–36</sup> Their approach was to combine the usage of artificial chromosome-based vectors with a multi-integrase system. This allowed homogeneous gene expression without integration of the vector carrying the target genes into host chromosomes. In addition, artificial chromosomes may be transferred from one cell line to another cell line. Hence, once a chromosome vector expressing a gene of interest was built, it could be moved and used in many different cell lines.

In this study, we developed a novel human artificial chromosome-based system, the IIS-alphoid<sup>tetO</sup>-HAC. This system utilizes two compound markers termed GHT and PCF. Each compound marker is composed of a positive selection marker, a counter selection marker and a fluorescence marker. The GHT marker is composed of sequences encoding green fluorescence protein (GFP), hygromycin phosphotransferase (hph) and herpesvirus thymidine kinase (Tk). The PCF marker is composed of sequences encoding puromycin N-acetyl-transferase (Pac), red fluorescence protein (mCherry), and a fusion protein of cytosine deaminase and uracil phosphoribosyl transferase (FcyFur). These compound markers can be distinguished visually and can be selected for or counter-selected against using the appropriate drug. The IIS-alphoid<sup>tetO</sup>-HAC system includes three enzymes, Cre,  $\Phi$ C31 and  $\Phi$ BT1. Binding sites of these enzymes are arranged in a manner that allows the IIS-alphoid<sup>tetO</sup>-HAC to use a promoter capture and a marker exchange strategy to assemble any desired number of genomic DNA segments. In this strategy, the compound marker that is expressed switches each time a new DNA segment is added to the platform cassette of the HAC. The compound markers allow positive selection for cells where a new DNA segment is added to the platform cassette and counter selection against the cells when this event does not happen. Efficiency and accuracy of the IIS-alphoid<sup>tetO</sup>-HAC

system has been demonstrated by assembling a functional copy of the VHL gene from multiple DNA segments.

The reassembly of genes from DNA segments by IIS requires breakage and rejoining junctions to be made to the target gene. The optimal placement of these junctions needs to fulfill several requirements. First, the breakage-rejoining junctions must be placed within introns of the target gene. This is necessary as spent integration sites (attR) (<60 bp) are left between adjacent DNA segment (see Figure 3) during the gene assembly process. Placement within introns, allows these integration sites to be spliced out from mRNA and so avoid disruption to the exonic coding sequence.

Second, the placement of breakage-rejoining junctions must not disrupt the splicing efficiency of its host intron. Hence, the breakage and joining junction should not be near any features critical to intron function such as the splice donor site, branchpoint sequence, poly pyrimidine tract and the splice donor site. The spliced donor is on the 5' end of the intron and this is easily avoided by placing the junction no closer than several hundred bp from it. The branchpoint sequence and poly pyrimidine tract are typically 20–50 bp upstream from the splice acceptor site (3' end of intron). These sequences can be avoided by placing the breakage-rejoining junction several hundred bp upstream of the splice acceptor sequence. Splice site prediction software can be used to identify the branch sequence (i.e., [http://regulatorygenomics.upf.edu/Software/SVM\\_BP/](http://regulatorygenomics.upf.edu/Software/SVM_BP/) or <http://www.umd.be/HSF3/>).

Third, insertion of the integrase sites at the breakage-rejoining junctions should not generate a cryptic splice donor or acceptor site. Splice prediction software such as <http://www.umd.be/HSF3/> can help predict the creation of cryptic splice site.

Fourthly, it is envisioned that TAR cloning would be the preferred method used to clone the gene segments. Hence, the breakage-rejoining junctions must avoid repeat sequences. Repeat sequences can be identified using the “Repeat mask” option in UCSC DNA download window.

Although several laboratories made a significant progress for synthetic biology by constructing multi-integrase systems on different HACs and MACs, a new IIS-alphoid<sup>tetO</sup>-HAC system has some notable advantages that set it apart from other similar artificial chromosome-based systems<sup>32–36</sup> (Figure S10). First, the maximum number of DNA segments that can be added to the IIS-alphoid<sup>tetO</sup>-HAC is only limited by carrying capacity of a human chromosome, which is several hundred Mbs. In effect the IIS-alphoid<sup>tetO</sup>-HAC can insert any desired number of genomic DNA segments into the construction platform located within the HAC. Second, each step of insertion is accompanied by a change in cell color that simplifies the selection of correct clones. Third, in this system the insertion “scar” between adjacent DNA segments is greatly reduced (<60 bp), consisting of a single recombined integration sites (attR) that could, as shown here, be incorporated within introns to allow reassembly a gene from multiple parts. Fourthly, IIS-alphoid<sup>tetO</sup>-HAC has an error proofing mechanism to remove mis-incorporated DNA segments and improve the integrity of assembly. Lastly, as we have previously shown<sup>12,13</sup> alphoid<sup>tetO</sup>-HAC carrying a gene(s) of interest can be removed from the cells in culture by targeting with TetR fusion proteins that allow a unique possibility to compare the phenotypes of human cells with and without a functional copy of a gene(s) inserted into the HAC.

For comparison, previously developed artificial chromosome-based systems<sup>32,33</sup> are limited to the number of efficient



recombinases known to work in mammalian cells (see Figure S10a). As a result, the maximum number of fragments that may be inserted into an artificial chromosome is low (five for the systems published and would be approximately 10 if the systems were modified to use all recombinases known to work in mammalian cells). In response, more advanced systems were recently developed by the same group that could perform unlimited multiple insertions<sup>34</sup> (see Figures S10c and S10d). However, all these systems mentioned above retain and integrate the entire plasmid backbone between DNA segments inserted into an artificial chromosome.<sup>32–34</sup> Therefore, they may experience problems assembling functional genes from gene segments due to the presence of cryptic splice sites or abnormal pausing of the splicing machinery that may change splice isoform generation. Similar problems exist in another artificial chromosome-based system<sup>35</sup> (see Figure S10b). Moreover, as this system uses multiple (<50) identical platform cassettes integrated in the chromosome-vector, the relative position of each new insert relative to one another is uncontrolled. The maximum number of insertions that can be added is limited by the number of platform cassettes, and is further reduced because each round adds an additional loxP site that may recombine and lead to destabilization of the chromosome upon Cre exposure. Another system is based on the principles similar to our system<sup>36</sup> (Figure S10e). However, it has several significant distinctions: it does not include color markers that help to distinguish one round of insertion from another; it does not have an error proofing mechanism; the HAC cannot be specifically and efficiently removed from the host cell cultures.

In summary, the IIS-alphoid<sup>tet<sup>o</sup></sup>-HAC system is able to efficiently and precisely carry out recombination in mammalian cells that allows the investigator to potentially insert any desired number of genomic fragments, leading to assembly of a functional copy of a gene and even more complex loci. The IIS-alphoid<sup>tet<sup>o</sup></sup>-HAC is a valuable unique genetic tool for investigating gene(s) function, complex biochemical pathways and has a great potential for animal transgenesis, development of therapeutic applications for complex diseases, and synthetic biology.

## METHODS

**Cell Lines and Culture.** Chicken B-lymphoma DT40 cells containing the alphoid<sup>tet<sup>o</sup></sup>-HAC were cultured in Roswell Park Memorial Institute medium 1640 (RPMI-1640, GIBCO) supplemented with 10% fetal bovine serum (Biowest, Nuaille, France), 1% chicken serum (Invitrogen), 50  $\mu$ M 2-mercaptoethanol (SIGMA) and 15  $\mu$ g/mL Blastidicin S (Funakoshi, Tokyo, Japan) at 40 °C in 10% CO<sub>2</sub>. Hypoxanthine phosphoribosyltransferase (HPRT)-deficient Chinese hamster ovary (CHO) cells (JCRB0218) carrying the alphoid<sup>tet<sup>o</sup></sup>-HAC were maintained in Ham's F-12 nutrient mixture (Invitrogen, USA) plus 10% fetal bovine serum (FBS) with 8  $\mu$ g/mL Blastidicin S Hydrochloride (Funakoshi, Japan). Blastidicin S is required for stable propagation of alphoid<sup>tet<sup>o</sup></sup>-HAC in the medium.

**Construction of Plasmids for the IIS-alphoid<sup>tet<sup>o</sup></sup>-HAC System.** A037: The platform cassette plasmid was constructed in two parts. The promoter and GHT marker components were amplified from multiple plasmid sources while the recombinase recognition sites, loxP and attB,  $\Phi$ C31 were added by PCR using long oligomers. The GHT compound marker is composed of a fusion of Green Fluorescence Protein (GFP),

P2A self-cleaving peptide, Hygromycin phosphotransferase (hph) and viral Thymidine Kinase (TK). The Chr13 targeting hooks (M2A and M2B) were PCR amplified from human genomic DNA and added to the construct. The construction steps and primer sequences used are depicted in Figure S1 and Table S1, respectively.

A167, A168, A169 and A170: The carrier plasmids were constructed in three sections: (i) the integration cassette that comprises of a promoterless compound marker with appropriate recombinase recognition sites; (ii) the error proofing cassette, which is composed of an expressed compound marker opposite to promoterless marker; (iii) the YAC-BAC shuttle vector. Two compound markers were built, the GHT and PCF markers. The GHT marker is as previously described while the PCF marker is a fusion of Puromycin N-acetyl-transferase (Pac), 2A self-cleaving peptide, mCherry fluorescence protein, a second 2A self-cleaving peptide, and cytosine deaminase fused to uracil phosphoribosyl transferase (FcyFur). The construction steps of the Type I carrier plasmid without ARS (A167) and with ARS (A168) are shown in Figure S2 and Figure S6, respectively. The construction steps of the Type II carrier plasmids without ARS (A169) and with ARS (A170) are shown in Figure S3 and Figure S7, respectively. All construction primers used are listed in Table S1

A135-JH and A139: These are the vectors that express recombinases  $\Phi$ BT1 and Cre,  $\Phi$ C31 and Cre, respectively. Each vector expresses two recombinases as a single peptide, linked by a 2A-self-cleaving peptide. The construction steps of A135-JH and A139 are shown in Figure S4 and Figure S5, respectively. All construction primers used are listed in Table S1.

**Insertion of the Platform Cassette A037 into the Alphoid<sup>tet<sup>o</sup></sup>-HAC in Chicken DT40 Cells.** The platform cassette A037 was targeted into the human Ch13 segment present within alphoid<sup>tet<sup>o</sup></sup>-HAC using homologous recombination in a highly recombinogenic chicken B-lymphoma DT40 cell line. Before targeting, the platform cassette was digested and linearized by XhoI to release M2A (positions Ch13:69420033–69523541; GRCH38/hg38) and M2B (positions Ch13:69523556–69527457; GRCH38/hg38) hook sequences that have homology to the Ch13 segment. 90 clones were obtained under Hygromycin selection. PCR analysis of genomic DNA from these clones using specific primers (Table S1) confirmed insertion of the platform cassette into the HAC in 14 of the clones.

**One Platform Cassette Per Alphoid<sup>tet<sup>o</sup></sup>-HAC, Per Cell.** In order for the IIS-HAC system to stably work, only one platform cassette can be present within the HAC and only one cassette within the cell in total. Consequently several strategies and assays were conducted to ensure there was only one platform cassette within the HAC and cell. First, the platform cassette was inserted into a position within the HAC that was previously found to be unique, using homologous recombination in chicken DT40 cell. Targeting was confirmed by PCRs. Hence, there should only be one copy of the platform cassette in the HAC. Second, the HAC containing the platform cassette was then transferred from chicken DT40 to hamster CHO cells by MMCT, thereby removing any copies of the platform cassette that may have integrated into the genome of the DT40 cell. Third, we also conducted fluorescence in situ hybridization (FISH) in conjunction with colony subcloning to ensure that the CHO cells we worked with had only a single HAC and thus one copy of the platform cassette. Fourthly, the double

selection system of IIS innately selects against the presents of more than one platform cassette. Hamster CHO cells, like many cancer cell line, experience chromosome instability. Hence the copy number of the HAC has been observed to spontaneously increase in a small percentage of cells during culturing. This issue is addressed in the IIS-HAC-based system by its compound markers that allow both positive selection and negative counter-selection which are conducted at each round of integration. More specifically, if there were two platform cassettes expressing the GHT marker (GFP-hph-TK) in the HAC, the integration of a DNA fragment with a promoterless PCF (Pac-mCherry-FcyFur) marker into one cassette would result in a cell expressing the PCF marker from one cassette and the GHT marker from the second cassette. The double selection with both puromycin (positive selection for Pac) and ganciclovir (counter-selection against TK) would remove such cells with one or more platform cassettes. Only cells with one platform cassette with the right compound marker can survive rounds of positive selection and negative counter-selection. Lastly and more importantly, PCR analysis of colonies from the integration of "empty vectors" (Figure S8) yielded single bands, indicating the presence of only one platform cassette. If there had been more than one platform cassette within the HAC, after each integration event, we should obtain more than one band in the PCR analysis, *i.e.*, one band for the platform cassette with the inset and one without. Thus, we can conclude that our IIS-HAC-based integration system has only one copy of inserted platform cassette.

**Microcell-Mediated Chromosome Transfer (MMCT) Technique.** The alphoid<sup>tetO</sup>-HAC containing platform cassette A037 was moved from chicken DT40 cells to hamster CHO cells using an improved microcell-mediated chromosome transfer technique.<sup>44</sup> After MMCT transfer six clones were obtained. PCR of genomic DNA isolated from these clones and FISH analysis confirmed the presence of an autonomous HAC in five clones. Clone #CHO BH3:37 was selected for further experiments.

**Fluorescence *In Situ* Hybridization (FISH).** FISH analysis was performed as following. Hamster CHO cells carrying the alphoid<sup>tetO</sup>-HAC bearing the platform cassette A037 were cultured in F12 medium with 10  $\mu\text{g}/\text{mL}$  of colcemid (Invitrogen) overnight at 37  $^{\circ}\text{C}$ . Metaphase cells were trypsinized and centrifugated for 4 min at 172g, treated in 10 mL of 50 mM KCl hypotonic solution for 20 min at 37  $^{\circ}\text{C}$  and washed three times in methanol:acetic acid (3:1) solution with a 4 min centrifugation at 172g between each wash. Cells were diluted to the appropriate density with fixative solution, spread onto precleaned slides (Thermo Fisher Scientific, Waltham, MA, USA) above steam (boiling water), and allowed to age 2 days at room temperature. For BAC probing, CHO metaphase slides were washed in 70% formamide in 2 $\times$  SSC for 2 min at 72  $^{\circ}\text{C}$ . Samples were dehydrated through a 70, 90, and 100% ethanol series for 4 min each and left to air-dry. The probe used for FISH was BAC32-2-mer(tetO) DNA containing 40 kb of alphoid-tetO array cloned into a BAC vector as described previously.<sup>11</sup> BAC DNA was labeled using a nick-translation kit with Orange 552 dUTP (5-TAMRA-dUTP) (Abbott Molecular). The probe was denatured in hybridization solution at 78  $^{\circ}\text{C}$  for 10 min and left at 37  $^{\circ}\text{C}$  for 30 min. The hybridization mix probe was applied to the sample and incubated at 37  $^{\circ}\text{C}$  overnight. Slides were washed with 0.4 $\times$  SSC, 0.3% Tween 20 for 2 min at 72  $^{\circ}\text{C}$ , briefly rinsed with 2 $\times$  SSC, 0.1% Tween 20 (10 s) and air-dried in darkness. The samples were counter-

stained with VECTASHIELD mounting medium containing DAPI (Vector Laboratories, Burlingame, CA, USA). Slides were analyzed by fluorescence microscopy. Images were captured using a DeltaVision imaging system in the CRC, LRBGE Fluorescence Imaging Facility (NIH) and analyzed using ImageJ software (NIH).

**Plasmid DNA Transfection and Loading into Alphoid<sup>tetO</sup>-HAC.** CHO cells with the alphoid<sup>tetO</sup>-HAC were cotransfected with Type I carrier plasmid A167 and A139 plasmid expressing  $\Phi\text{C31}$  integrase and Cre recombinase for the first and third rounds of insertion or with Type II carrier plasmid A169 and A135-JH plasmid expressing  $\Phi\text{BT1}$  integrase and Cre recombinase for the second round of insertion. Briefly  $1 \times 10^5$  CHO cells were seeded in one well of a 6-well plate in growth media without selection antibiotics. Next day, 200  $\mu\text{L}$  of Opti-MEM (Gibco), 6  $\mu\text{L}$  of X-tremeGENE 9 DNA Transfection Reagent (Roche) 1.8  $\mu\text{g}$  of A167 and 0.2  $\mu\text{g}$  of A139 or 1.8  $\mu\text{g}$  A169 and 0.2  $\mu\text{g}$  A135-JH plasmids were mixed, incubated 20 min at room temperature and added to the cells dropwise. Next day, the cells were trypsinized and transferred to two 100 mm culture dishes with growth media containing 5  $\mu\text{g}/\text{mL}$  Blasticidin S and 5  $\mu\text{g}/\text{mL}$  Puromycin in case of the first and third rounds of insertion or 5  $\mu\text{g}/\text{mL}$  Blasticidin S and 200  $\mu\text{g}/\text{mL}$  Hygromycin B in the case of the second round. The cells were cultured for 7–10 days until colonies of about  $1 \times 10^3$  cells formed. Individual colonies were transferred to a 24-well plate and cultured in media with 5  $\mu\text{g}/\text{mL}$  Blasticidin S, 5  $\mu\text{g}/\text{mL}$  Puromycin and 10  $\mu\text{g}/\text{mL}$  Ganciclovir or 5  $\mu\text{g}/\text{mL}$  Blasticidin S, 200  $\mu\text{g}/\text{mL}$  Hygromycin B and 160  $\mu\text{g}/\text{mL}$  5-Fluorocytidine.

**PCR Reaction.** Genomic DNA was extracted from CHO cultured cells containing the alphoid<sup>tetO</sup>-HAC with DNeasy Blood & Tissue Kit (Qiagen). DNA concentration was measured by NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific). The PCR mixture included 0.5  $\mu\text{L}$  TaKaRa LA Taq (5 units/ $\mu\text{L}$ , Clontech, #RR002M, Lot KA3101CA), 5  $\mu\text{L}$  10 $\times$  LA PCR Buffer, 8  $\mu\text{L}$  dNTP mixture (2.5 mM each), 300 ng of genomic DNA, 2  $\mu\text{L}$  forward and 2  $\mu\text{L}$  reverse primers (10  $\mu\text{M}$ ), sterile distilled water up to 50  $\mu\text{L}$ . PCR products were resolved on 1% or 2% agarose gels.

**Construction of VHL Gene Fragment-Containing Plasmids and Loading into Alphoid<sup>tetO</sup>-HAC.** Three fragments containing either exon 1, exon 2 or exon 3 of the VHL gene were PCR-amplified from the TAR-isolated YAC/BAC clone containing a full-length genomic copy of the VHL gene<sup>13</sup> using specific primers (Table S1). The fragments were cloned into the appropriate plasmids as follows: a 5990 bp *AscI*-*FseI* fragment 1 containing exon 1 along with the VHL promoter was inserted into Type 1 carrier vector A168 digested by *AscI*/*FseI* restriction enzymes; a 4221 bp *AscI*-*NotI* fragment 2 containing exon 2 was inserted into Type 2 carrier vector A170 digested by *AscI*/*NotI* restriction enzymes; a 6323 bp *AscI*-*NotI* fragment 3 containing exon 3 was inserted into Type 1 carrier vector A168 digested by *AscI*/*NotI* restriction enzymes. Three rounds of insertion were performed to assemble a full-length VHL gene on the HAC. For the first round of insertion, 3  $\mu\text{g}$  of the exon 3-carrying vector and 2  $\mu\text{g}$  of  $\Phi\text{C31}$ -Cre-expressing A139 vector were cotransfected into CHO cells carrying the alphoid<sup>tetO</sup>-HAC bearing the previously inserted platform cassette A037 using the appropriate transfection reagents (MTI-GlobalStem). Four  $\mu\text{g}/\text{mL}$  Puromycin plus 4  $\mu\text{g}/\text{mL}$  Blasticidin S were used as a positive selection for 5 days and 4  $\mu\text{g}/\text{mL}$  Puromycin, 4  $\mu\text{g}/\text{mL}$  Blasticidin S and 5

$\mu\text{g/mL}$  Gancyclovir were used as a positive/negative selection for 1 week to select for correct insertion of exon 3. For the second round of insertion, 3  $\mu\text{g}$  of exon 2-carrying vector and 2  $\mu\text{g}$  of  $\Phi\text{BT1-Cre}$ -expressing A135-JH vector were cotransfected into CHO cells. 100  $\mu\text{g/mL}$  Hygromycin B plus 4  $\mu\text{g/mL}$  Blastidicin S were used as a positive selection for 5 days. 100  $\mu\text{g/mL}$  Hygromycin B, 4  $\mu\text{g/mL}$  Blastidicin S and 100  $\mu\text{g/mL}$  5-Fluorocytidine were used as a positive/negative selection for 1 week to select for correct insertion of exon 2. For the third round of insertion, 3  $\mu\text{g}$  of the exon 1-carrying vector and 2  $\mu\text{g}$  of  $\Phi\text{C31-Cre}$ -expressing A139 vector were cotransfected into hamster CHO cells. Four  $\mu\text{g/mL}$  Puromycin plus 4  $\mu\text{g/mL}$  Blastidicin S were used as a positive selection for 5 days and 4  $\mu\text{g/mL}$  Puromycin, 4  $\mu\text{g/mL}$  Blastidicin S and 5  $\mu\text{g/mL}$  Gancyclovir were used as a positive/negative selection for 1 week to select for correct insertion of exon 1. All insertions after each round were confirmed by PCR with the corresponding primers (Table S1) that diagnose a correctly assembled VHL gene.

**RT-PCR Reaction.** Transcription of the VHL gene from  $\text{alphoid}^{\text{tetO}}$ -HAC/VHL in hamster CHO cells was detected by RT-PCR by using specific primers described in Table S1.

## ■ ASSOCIATED CONTENT

### 📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acssynbio.7b00209](https://doi.org/10.1021/acssynbio.7b00209).

Construction history of vectors used in the Iterative Integration System (IIS) (Figures S1–S7); Initial test of IIS and PCR analysis (Figure S8); cDNA sequence comparison between human VHL reconstructed by IIS to native human and native CHO VHL (Figure S9); Schematic diagram of previous integration systems (Figure S10); Primers used in this study (Table S1) (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*Tel: +1 301 496 7941. Fax: +1 301 480 2772. E-mail: [kouprinn@mail.nih.gov](mailto:kouprinn@mail.nih.gov).

### ORCID

Nicholas C. O. Lee: [0000-0003-2628-6599](https://orcid.org/0000-0003-2628-6599)

### Author Contributions

<sup>#</sup>NCOL and JHK contributed equally to this paper. NCOL initiated and conceived the experiment. NCOL designed and built the integration system. NCOL, JHK, NSP and HSL performed the experiments. NK, HM, WCE and VL analyzed the data. NK and NCOL wrote the paper. WCE and NCO edited the manuscript.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors would like to thank the CRC, LRBGE Fluorescence Imaging Facility (NIH) and personally Drs. Karpova for instructions, consultations and help with the usage of a DeltaVision microscopy imaging system. This work was supported by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research, USA (VL and NK), a Wellcome Trust Principal Research Fellowship [grant number 107022] (WCE), MEXT KAKENHI grants

[16H04747 and 16H01414] (HM) and the Kazusa DNA Research Institute Foundation (HM).

## ■ REFERENCES

- (1) Kazuki, Y., and Oshimura, M. (2011) Human artificial chromosomes for gene delivery and the development of animal models. *Mol. Ther.* 19, 1591–1601.
- (2) Kouprina, N., Earnshaw, W. C., Masumoto, H., and Larionov, V. (2013) A new generation of human artificial chromosomes for functional genomics and gene therapy. *Cell. Mol. Life Sci.* 70, 1135–1148.
- (3) Oshimura, M., Kazuki, Y., Iida, Y., and Uno, N. (2013) New Vectors for Gene Delivery: Human and Mouse Artificial Chromosomes, In *eLS*, John Wiley & Sons, Ltd.
- (4) Kouprina, N., Tomilin, A. N., Masumoto, H., Earnshaw, W. C., and Larionov, V. (2014) Human artificial chromosome-based gene delivery vectors for biomedicine and biotechnology. *Expert Opin. Drug Delivery* 11, 517–535.
- (5) Katona, R. L. (2015) De novo formed satellite DNA-based mammalian artificial chromosomes and their possible applications. *Chromosome Res.* 23, 143–157.
- (6) Moralli, D., and Monaco, Z. L. (2015) Developing de novo human artificial chromosomes in embryonic stem cells using HSV-1 amplicon technology. *Chromosome Res.* 23, 105–110.
- (7) Kouprina, N., and Larionov, V. (2015) Recent advances in chromosome engineering. *Chromosome Res.* 23, 1–8.
- (8) Bergmann, J. H., Martins, N. M. C., Larionov, V., Masumoto, H., and Earnshaw, W. C. (2012) HAcKING the centromere chromatin code: insights from human artificial chromosomes. *Chromosome Res.* 20, 505–519.
- (9) Ohzeki, J., Larionov, V., Earnshaw, W. C., and Masumoto, H. (2015) Genetic and epigenetic regulation of centromeres: a look at HAC formation. *Chromosome Res.* 23, 87–103.
- (10) Kazuki, Y., Hoshiya, H., Takiguchi, M., Abe, S., Iida, Y., Osaki, M., Katoh, M., Hiratsuka, M., Shirayoshi, Y., Hiramatsu, K., Ueno, E., Kajitani, N., Yoshino, T., Kazuki, K., Ishihara, C., Takehara, S., Tsuji, S., Ejima, F., Toyoda, A., Sakaki, Y., Larionov, V., Kouprina, N., and Oshimura, M. (2011) Refined human artificial chromosome vectors for gene therapy and animal transgenesis. *Gene Ther.* 18, 384–393.
- (11) Iida, Y., Kim, J. H., Kazuki, Y., Hoshiya, H., Takiguchi, M., Hayashi, M., Erliandri, I., Lee, H. S., Samoshkin, A., Masumoto, H., Earnshaw, W. C., Kouprina, N., Larionov, V., and Oshimura, M. (2010) Human Artificial Chromosome with a Conditional Centromere for Gene Delivery and Gene Expression. *DNA Res.* 17, 293–301.
- (12) Kim, J. H., Kononenko, A., Erliandri, I., Kim, T. A., Nakano, M., Iida, Y., Barrett, J. C., Oshimura, M., Masumoto, H., Earnshaw, W. C., Larionov, V., and Kouprina, N. (2011) Human artificial chromosome (HAC) vector with a conditional centromere for correction of genetic deficiencies in human cells. *Proc. Natl. Acad. Sci. U. S. A.* 108, 20048–20053.
- (13) Kononenko, A. V., Bansal, R., Lee, N. C., Grimes, B. R., Masumoto, H., Earnshaw, W. C., Larionov, V., and Kouprina, N. (2014) A portable BRCA1-HAC (human artificial chromosome) module for analysis of BRCA1 tumor suppressor function. *Nucleic Acids Res.* 42 (21), e164.
- (14) Tedesco, F. S. (2015) Human artificial chromosomes for Duchenne muscular dystrophy and beyond: challenges and hopes. *Chromosome Res.* 23, 135–141.
- (15) Hasegawa, Y., Ishikura, T., Hasegawa, T., Watanabe, T., Suzuki, J., Nakayama, M., Okamura, Y., Okazaki, T., Koseki, H., Ohara, O., Ikeno, M., and Masumoto, H. (2015) Generating a transgenic mouse line stably expressing human MHC surface antigen from a HAC carrying multiple genomic BACs. *Chromosoma* 124, 107–118.
- (16) Liskovych, M., Ponomartsev, S., Popova, E., Bader, M., Kouprina, N., Larionov, V., Alenina, N., and Tomilin, A. (2015) Stable maintenance of de novo assembled human artificial chromosomes in embryonic stem cells and their differentiated progeny in mice. *Cell Cycle* 14, 1268–1273.

- (17) Oshimura, M., and Katoh, M. (2008) Transfer of human artificial chromosome vectors into stem cells. *Reprod. BioMed. Online* 16, 57–69.
- (18) Hiratsuka, M., Uno, N., Ueda, K., Kurosaki, H., Imaoka, N., Kazuki, K., Ueno, E., Akakura, Y., Katoh, M., Osaki, M., Kazuki, Y., Nakagawa, M., Yamanaka, S., and Oshimura, M. (2011) Integration-Free iPS Cells Engineered Using Human Artificial Chromosome Vectors. *PLoS One* 6, e17267.
- (19) Tedesco, F. S., and Cossu, G. (2012) Stem cell therapies for muscle disorders. *Curr. Opin. Neurol.* 25, 597–603.
- (20) Tedesco, F. S., Hoshiya, H., D'Antona, G., Gerli, M. F., Messina, G., Antonini, S., Tonlorenzi, R., Benedetti, S., Berghella, L., Torrente, Y., Kazuki, Y., Bottinelli, R., Oshimura, M., and Cossu, G. (2011) Stem cell-mediated transfer of a human artificial chromosome ameliorates muscular dystrophy. *Sci. Transl. Med.* 3, 96ra78.
- (21) Yakura, Y., Ishihara, C., Kurosaki, H., Kazuki, Y., Komatsu, N., Okada, Y., Doi, T., Takeya, H., and Oshimura, M. (2013) An induced pluripotent stem cell-mediated and integration-free factor VIII expression system. *Biochem. Biophys. Res. Commun.* 431, 336–341.
- (22) Toth, A., Fodor, K. L., Blazso, P., Cserpan, I., Praznovszky, T., Tubak, V., Udvardy, A., Hadlaczky, G., and Katona, R. L. (2014) Generation of Induced Pluripotent Stem Cells by Using a Mammalian Artificial Chromosome Expression System. *Acta Biol. Hung.* 65, 331–345.
- (23) Nakano, M., Cardinale, S., Noskov, V. N., Gassmann, R., Vagnarelli, P., Kandels-Lewis, S., Larionov, V., Earnshaw, W. C., and Masumoto, H. (2008) Inactivation of a human kinetochore by specific targeting of chromatin modifiers. *Dev. Cell* 14, 507–522.
- (24) Ebersole, T., Okamoto, Y., Noskov, V. N., Kouprina, N., Kim, J. H., Leem, S. H., Barrett, J. C., Masumoto, H., and Larionov, V. (2005) Rapid generation of long synthetic tandem repeats and its application for analysis in human artificial chromosome formation. *Nucleic Acids Res.* 33, e130.
- (25) Kouprina, N., Samoshkin, A., Erliandri, I., Nakano, M., Lee, H. S., Fu, H. G., Iida, Y., Aladjem, M., Oshimura, M., Masumoto, H., Earnshaw, W. C., and Larionov, V. (2012) Organization of Synthetic Alphoid DNA Array in Human Artificial Chromosome (HAC) with a Conditional Centromere. *ACS Synth. Biol.* 1, 590–601.
- (26) Colombo, R., and Moll, J. (2011) Targeting aneuploid cancer cells. *Expert Opin. Ther. Targets* 15, 595–608.
- (27) Janssen, A., Kops, G. J. P. L., and Medema, R. H. (2009) Elevating the frequency of chromosome mis-segregation as a strategy to kill tumor cells. *Proc. Natl. Acad. Sci. U. S. A.* 106, 19108–19113.
- (28) Lee, H. S., Lee, N. C. O., Grimes, B. R., Samoshkin, A., Kononenko, A. V., Bansal, R., Masumoto, H., Earnshaw, W. C., Kouprina, N., and Larionov, V. (2013) A new assay for measuring chromosome instability (CIN) and identification of drugs that elevate CIN in cancer cells. *BMC Cancer* 13, 252.
- (29) Lee, H. S., Lee, N. C. O., Kouprina, N., Kim, J. H., Kagansky, A., Bates, S., Trepel, J. B., Pommier, Y., Sackett, D., and Larionov, V. (2016) Effects of Anticancer Drugs on Chromosome Instability and New Clinical Implications for Tumor-Suppressing Therapies. *Cancer Res.* 76, 902–911.
- (30) Kim, J. H., Lee, H. S., Lee, N. C. O., Goncharov, N. V., Kumeiko, V., Masumoto, H., Earnshaw, W. C., Kouprina, N., and Larionov, V. (2016) Development of a novel HAC-based "gain of signal" quantitative assay for measuring chromosome instability (CIN) in cancer cells. *Oncotarget* 7, 14841–14856.
- (31) Duffy, S., Fam, H. K., Wang, Y. K., Styles, E. B., Kim, J. H., Ang, J. S., Singh, T., Larionov, V., Shah, S. P., Andrews, B., Boerkoel, C. F., and Hieter, P. (2016) Overexpression screens identify conserved dosage chromosome instability genes in yeast and human cancer. *Proc. Natl. Acad. Sci. U. S. A.* 113, 9967–9976.
- (32) Yamaguchi, S., Kazuki, Y., Nakayama, Y., Nanba, E., Oshimura, M., and Ohbayashi, T. (2011) A Method for Producing Transgenic Cells Using a Multi-Integrase System on a Human Artificial Chromosome Vector. *PLoS One* 6, e17267.
- (33) Yoshimura, Y., Nakamura, K., Endo, T., Kajitani, N., Kazuki, K., Kazuki, Y., Kugoh, H., Oshimura, M., and Ohbayashi, T. (2015) Mouse embryonic stem cells with a multi-integrase mouse artificial chromosome for transchromosomal mouse generation. *Transgenic Res.* 24, 717–727.
- (34) Suzuki, T., Kazuki, Y., Oshimura, M., and Hara, T. (2014) A Novel System for Simultaneous or Sequential Integration of Multiple Gene-Loading Vectors into a Defined Site of a Human Artificial Chromosome. *PLoS One* 9, e110404.
- (35) Toth, A., Fodor, K., Praznovszky, T., Tubak, V., Udvardy, A., Hadlaczky, G., and Katona, R. L. (2014) Novel Method to Load Multiple Genes onto a Mammalian Artificial Chromosome. *PLoS One* 9, e85565.
- (36) Xu, Z., Lee, N. C. O., Dafhnis-Calas, F., Malla, S., Smith, M. C. M., and Brown, W. R. A. (2008) Site-specific recombination in *Schizosaccharomyces pombe* and systematic assembly of a 400kb transgene array in mammalian cells using the integrase of *Streptomyces* phage phi BT1. *Nucleic Acids Res.* 36, e9.
- (37) Larionov, V., Kouprina, N., Graves, J., Chen, X. N., Korenberg, J. R., and Resnick, M. A. (1996) Specific cloning of human DNA as yeast artificial chromosomes by transformation-associated recombination. *Proc. Natl. Acad. Sci. U. S. A.* 93, 491–496.
- (38) Kouprina, N., and Larionov, V. (2006) Innovation - TAR cloning: insights into gene function, long-range haplotypes and genome structure and evolution. *Nat. Rev. Genet.* 7, 805–812.
- (39) Kouprina, N., and Larionov, V. (2008) Selective isolation of genomic loci from complex genomes by transformation-associated recombination cloning in the yeast *Saccharomyces cerevisiae*. *Protoc.* 3, 371–377.
- (40) Lee, N. C. O., Larionov, V., and Kouprina, N. (2015) Highly efficient CRISPR/Cas9-mediated TAR cloning of genes and chromosomal loci from complex genomes in yeast. *Nucleic Acids Res.* 43, e55.
- (41) Kouprina, N., and Larionov, V. (2016) Transformation-associated recombination (TAR) cloning for genomics studies and synthetic biology. *Chromosoma* 125, 621–632.
- (42) Ebersole, T., Kim, J. H., Samoshkin, A., Kouprina, N., Pavlicek, A., White, R. J., and Larionov, V. (2011) tRNA genes protect a reporter gene from epigenetic silencing in mouse cells. *Cell Cycle* 10, 2779–2791.
- (43) Raab, J. R., Chiu, J., Zhu, J. C., Katzman, S., Kurukuti, S., Wade, P. A., Haussler, D., and Kamakaka, R. T. (2012) Human tRNA genes function as chromatin insulators. *EMBO J.* 31, 330–350.
- (44) Liskovych, M., Lee, N. C., Larionov, V., and Kouprina, N. (2016) Moving toward a higher efficiency of microcell-mediated chromosome transfer. *Mol. Ther.–Methods Clin. Dev.* 3, 16043.