(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 February 2019 (28.02.2019)

WIPO | PCT

(10) International Publication Number
WO 2019/037099 A1

(72) Inventors; and
(71) Applicants: QIN, Wenning [US/US]; 33 Harnden Avenue, Watertown, MA Massachusetts 02472 (US). WANG, Haoyi [CN/CN]; Building 956, Unit 1, Room 202, Zhong Guan Cun Nan San Street, Haidian District, Beijing 100089 (CN).

(74) Agent: KING & WOOD MALLESONS; 20th Floor, East Tower, World Financial Centre, No.1 Dongsanhuan Zhonglu, Chaoyang District, Beijing 100020 (CN).

(54) Title: LARGE SCALE MODIFICATION OF EUKARYOTIC GENOME

(57) Abstract: Provided are methods and compositions/reagents for efficiently carrying out large scale modification of eukaryotic genome.

LARGE SCALE MODIFICATION OF EUKARYOTIC GENOME

BACKGROUND OF THE INVENTION

Although sharing common features, animals differ from humans in genetic composition and gene function. To improve the translational value of animal models, it is often necessary to

5   partially replace and alter its genome and reconstitute with human specific biology and function. Although there has been significant improvement over the last 30 years, current genome engineering technologies are still cumbersome and laborious, limited by size and confines of the donor genomic DNA fragment available.

To replace a gene or gene cluster in the recipient genome with its ortholog or syntenic

10   region from the donor genome, a shuttle vector must first be identified or constructed carrying genomic DNA fragment of the donor origin and transferring it to the recipient genome. Over the years, several systems have been invented serving this need, including plasmid, cosmid (Hohn and Murray, 1977), BAC (bacterial artificial chromosome) (Osoegawa et al., 2001), YAC (yeast artificial chromosome) (Burke et al., 1987), HAC (human artificial chromosome) (Hoshiya et al.,

15   2009), and "transchromosome" (Tomizuka et al., 1997).

Cosmid is a circular plasmid, with capacity for DNA insert of 45 kb (kilo base pairs). For smaller genes, this size may be sufficiently large to accommodate the entire gene, including the 5' and 3' regulatory sequences.

BAC vector is also a circular plasmid, with an average insert size of 150-350 kb. With its

20   fair size and stable insert, BAC has been the popular choice providing human and murine genomic DNA fragments for a variety of uses, including whole genome sequencing and genome engineering.

YAC is also circular and larger and can hold inserts up to 2,000 kb, with an average insert size of 250-400 kb. Although insert size is appealing, owing to its large size and

25   propagation in the highly recombinant yeast host, YAC clones are plagued with issues of artifact, including intramolecular rearrangement of the insert (insertion, deletion, and inversion, etc.), ligation of more than one insert into the same vector, and combination of more than one YAC into one YAC clone. As such, YAC library is of limited use.

In addition to these preexisting libraries of genomic DNA fragments, genomic DNA can

also be transferred directly from donor to recipient cells by fusion of the two cell types, in a process called microcell-mediated chromosome transfer. The sizes of these "transchromosomes" can be very large, up to the length of the entire chromosome. However, what is transferred is by chance and often fragmented and in addition, propagation to the next round of cell division is not

5   always guaranteed.

In addition to their sizes, another key limitation common to these vectors is that the insert carried in each vector was generated by fragmenting the linear chromosome into smaller pieces and as such, the start and end points of an insert can not be controlled. For Cosmid, BAC and YAC libraries, genomic DNA from the donor is digested with a restriction enzyme. For

10   "transchromosome," although transfer of an intact chromosome is possibly, chromosome is more likely fragmented during the microcell mediated transfer process.

One example is the RPCI-11 BAC library (Osoegawa et al., 2001) that carries the ensemble of human genomic DNA fragments in the pBACe3.6 vector. To generate the library, blood sample was collected from a human donor, genomic DNA isolated and partially digested

15   with EcoRI restriction enzyme. The fragmented DNA was then separated on agarose gel and size selected DNA fragments cloned into the pBACe3.6 vector between the EcoRI sites (for a subset of the clones, the same donor DNA was partially digested with MboI, size selected, and ligated into the pTARBAC1 cloning vector at the BamHI site). The ligation products were subsequently transformed into DH10B cells and individual clones arrayed into 384-well

20   microtiter plates. The entire collection of these plates are stored for archive and retrieval. For the human genomic DNA library, RPCI-11 consists of 1,440 plates and a total of 543,797 clones, covering the entire 3×10^9 base pairs of the human genome in multiple coverages. Depending on where EcoR1 enzyme cuts, it is possible that a gene may be carried on more than one BAC clone and a gene cluster larger than the carrying capacity of pBACe3.6 (150-350 kb) most definitely

25   will be housed in more than one BAC clone.

Genes within the same gene family often are located close by on a chromosomal region and these gene clusters vary in size but often are large. For example, the human immunoglobulin heavy chain (IGH) locus is mapped to the most distal band of the long arm of chromosome 14 (14q32.33) and it occupies 1.3 Mb (1.3×10^6 bps), out of a total of 107×10^6 (or 1.2%), of the

30   chromosome 14 space. The IGH gene cluster consists of 123-129 IGHV (could vary among individuals), 27 IGHD, 9 IGHJ and 11 IGHC genes, along with other genes that are colocalized

to this region (Giudicelli et al., 2006). Another example is the human major histocompatibility (MHC) gene cluster. The MHC gene cluster is located on the short arm of chromosome 6 at band 6p21.3 and spans about 4 Mb, out of a total of 171 Mb, or 2.3%, of the human chromosome 6 content. Genes of the MHC classes 1 and 2 are located within this region.

5   Mice carrying human immunoglobulin (IG) genes have been created and proven to be of tremendous pharmaceutical value (Green, 2014). The first generation mouse models were created carrying human IG transgenes on a murine knockout background (Bruggemann et al., 1989; Green et al., 1994; Lonberg et al., 1994; Taylor et al., 1992). Although capable of producing antibody of a high affinity when challenged, these mice have major limitations owing

10   to early transgenic technologies used to create these mice. They carry the BAC or YAC derived transgene and accordingly, partial representation of the IGV gene cluster. In addition, these transgenes randomly integrate into the mouse genome and are out of the genomic context in which they naturally reside. As a consequence, they may not have retained authentic transcriptional regulation offered by particularly the 3' flanking sequences which are known to

15   be important.

        Employing gene targeting or recombinase mediated cassette exchange (RMCE), two other IG mouse strains have also been created, which overcome the limitations found in a transgenic model. Regeneron Pharmaceutical, using conventional gene targeting with BAC based targeting vectors, created the VelocImmune mice with precise replacement of the murine

20   Ighv, d and j genes with its human ortholog (Macdonald et al., 2014). In another effort, the Kymab group resorted to RMCE and successfully generated a mouse model also with precise replacement of the murine Ighv, d, and j genes with their human syntenic region (Lee et al., 2014). Both efforts involved ferrying the human IGH genes, one BAC vector at a time, into the mouse embryonic stem cells, and the entire process took multiple rounds of transfer and years of

25   work. In addition, before used as a vector, some of the BAC clones needed to be "trimmed" to delete or add genes to them so they tile and cover the entire human IGH gene locus.

        Therefore, improved techniques are needed so that genomes can be engineered with a better efficiency.

SUMMARY OF THE INVENTION

30   The invention described herein provides compositions and methods to produce large

DNA fragments (*e.g.*, hundreds of Kb to million base pairs) containing a genomic region of interest (GOI) with defined start and end points. For most genome engineering needs, these large DNA fragments (termed "megacircle" herein) can be used to ferry genes or blocks of genes of donor origin to the recipient genome in one or multiple transfers. In one example, the

5   megacircle can be used to transfer a gene cluster from a first / donor organism (such as the human IGH gene cluster) to the cells of a second / recipient organism (such as a model organism like mouse), and replace a target (*e.g.*, the corresponding genomic) region in the second / recipient organism (*e.g.*, murine Igh locus) with a (such as its syntenic) region from the first / donor organism (*e.g.*, human). Any other gene clusters can be used in the invention, including,

10   but not limited to, those for TCR (T cell receptor) gene cluster, MHC classes 1 and 2, and cytochrome P450 genes. This method of the invention can be used to reduce large scale genome engineering to a few steps, including eliminating the need to amplifying large segments of DNA fragments in a prokaryotic (such as bacteria) or eukaryotic (such as yeast) host cell, thus improving efficiency and saving time and cost.

15   Briefly, the methods of the invention includes the production of megacircles, which may comprise the following steps: 1. release a genomic DNA comprising the GOI from a host genome; 2. ligate the genomic DNA into megacircle with or without another DNA fragment; 3. RCA amplification of megacircle, generating double strand concatemers; 4. Digest the concatemer to make monomers with compatible sticky overhangs (megastrand); and 5. Self-

20   ligating the megastrands to re-generate megacircles.

              The megacircle or megastrand of the invention can, for example, serve as donor template of homologous recombination (HR) if it contains genomic sequence homologous to a target genomic region in a recipient or host. It could also contain binding site for endonucleases such as CRISPR, TALEN and ZFN.

25   Thus one aspect of the invention provides a DNA fragment comprising a genomic region of interest (GOI, which can be a single complete or partial gene, or a cluster of related or unrelated genes, non-encoding sequence, or any genomic region of interest), wherein the GOI is at least about 20 kb (*e.g.*, 50 kb, 100 kb, 150 kb, 200 kb, 300 kb, 400 kb, 500 kb, 750 kb, 1 mb, 1.25 mb, 1.5 mb, 1.75 mb, 2 mb or more), and is flanked by a first unique enzyme cleavage site

30   at the proximal end of the genomic DNA fragment and a second unique enzyme cleavage site at the distal end of the genomic DNA fragment, wherein the DNA fragment further comprises an

exogenous sequence (such as a selection marker, a recombinase target site (RTS), or a sequence from a non-donor source) between the proximal end of the DNA fragment and the proximal end of the GOI, or between the distal end of the DNA fragment and the distal end of the GOI; and wherein the DNA fragment is devoid of a sequence required for self-replication in a host cell

5   (*e.g.*, ORI, pBR322, pUC, pSC101, or ARS).

As used herein, "unique" means that the enzyme cleavage site does not exist inside the GOI, so that when an enzyme specific for the unique enzyme cleavage site is used in digestion, the enzyme does not cleave within the GOI. However, the unique enzyme cleavage site may well exist in other parts of the same genome.

10   In certain embodiments, the first and the second unique enzyme cleavage sites have the same sequence, or are recognized by one (the same) cleavage enzyme, or have sticky ends compatible for ligation.

In certain embodiments, the DNA fragment is a double-stranded circle (*e.g.*, linked at the first and the second unique enzyme cleavage sites).

15   In certain embodiments, the DNA fragment further comprises a first positive selection marker cassette located between the RTS and an end of the DNA fragment, wherein the GOI and the positive selection marker cassette flank the RTS.

In certain embodiments, the DNA fragment further comprises a heterotypic recombinase target site (htRTS), wherein the RTS and the htRTS flank the GOI.

20   In certain embodiments, the DNA fragment is circular, with a linkage created by linking the first and the second unique enzyme cleavage sites, wherein the linkage is between the positive selection marker cassette and the htRTS (or the RTS).

In certain embodiments, the DNA fragment is amplified by rolling circle amplification (RCA).

25   In certain embodiments, the GOI is immunoglobulin heavy chain gene cluster, immunoglobulin kappa light chain gene cluster, immunoglobulin lambda light chain gene cluster, a TCR (T-cell receptor) gene cluster, an MHC class 1 gene, an MHC class 2 gene, or a cytochrome P450 gene, or any other genomic fragment.

In certain embodiments, the first and/or the second unique enzyme cleavage site(s) are

30   cleavable by a restriction endonuclease, a homing endonuclease (such as I-Sce1), a

CRISPR/Cas9 nuclease, a CRISPR/Cpf1 nuclease, a TALEN (transcription activator-like effector nuclease), a ZFN (Zinc Finger nuclease), or a Pyrococcus furiosus Argonaute (PfAgo) based artificial restriction enzyme (ARE).

In certain embodiments, one of the first and the second unique enzyme cleavage sites is a natural / pre-existing genomic sequence.

In certain embodiments, the RTS or the htRTS is locus of crossover in P1 (loxP), flippase recognition target (FRT), attP/attB, or mutants thereof, or the recognition site of other site-specific recombinases.

Another aspect of the invention provides a method of producing the DNA fragment of any one of the subject DNA fragments, the method comprising: (a) inserting, in a host genome comprising said GOI, said first unique enzyme cleavage site and said RTS proximal (or distal) to the GOI, wherein the GOI is proximal (or distal) to the second unique enzyme cleavage site, and wherein an optional first positive selection marker cassette is distal (or proximal) to said first unique enzyme cleavage site and proximal (or distal) to said RTS; (b) digesting the host genome with a first unique enzyme that cleaves said first unique enzyme cleavage site, and a second unique enzyme that cleaves said second unique enzyme cleavage site, to release the DNA fragment comprising the GOI from the host genome; (c) circularizing the released DNA fragment to form a megacircle, by ligating the ends of said released DNA fragment generated by cleavage by said first unique enzyme and said second unique enzyme, under conditions that promote intramolecular self-ligation; (d) amplifying said megacircle by rolling circle amplification (RCA) with a primer (or multiple primers) unique to said GOI.

In certain embodiments, steps (a) and (b) can be done simultaneously, or sequentially in either order.

In certain embodiments, the second unique enzyme cleavage site is naturally existing sequence. Note, however, in a similar aspect of the invention, a method is contemplated in which in step (a), two naturally existing unique enzyme cleavage sites flank the GOI.

In certain embodiments, step (a) further comprises inserting a 3rd unique enzyme cleavage site between the first unique enzyme cleavage site and the RTS.

In certain embodiments, the second unique enzyme cleavage site, together with a second positive selection marker cassette and htRTS, are inserted into said host genome distal (or

proximal) to GOI, preferably the second unique enzyme cleavage site is distal (or proximal) to the htRTS and proximal (or distal) to the second positive selection marker cassette.

Another aspect of the invention provides a method of producing the DNA fragment of any one of the subject DNA fragments, the method comprising: (a) removing / retrieving a

5   genomic DNA fragment comprising said GOI from a host genome, through digesting the host genome with a first enzyme specific for a first unique enzyme cleavage site and a second enzyme specific for a second unique enzyme cleavage site, wherein said first and said second unique enzyme cleavage sites are not compatible for ligation with each other; (b) ligating the retrieved genomic DNA fragment in step (a) with a linear recombinant DNA construct to produce a

10   megacircle, wherein said linear recombinant DNA construct comprises a positive selection marker cassette flanked by a pair of heterotypic RTSs or flanked by a RTS and a 3rd unique enzyme cleavage site, and wherein the ends of the linear recombinant DNA construct are defined by said first and said second unique enzyme cleavage sites, respectively; (c) amplifying said megacircle by rolling circle amplification (RCA) with a primer unique to said GOI.

15   In certain embodiments, said first and said second unique enzyme cleavage sites are the same, and said first and said second unique enzymes are the same.

In certain embodiments, the method further comprises, before amplification by RCA, treating the ligation mixture with an exonuclease to eliminate linear DNA fragments.

In certain embodiments, RCA is performed using phi29 (θ29) DNA polymerase, and one

20   or more primers each specific for the GOI.

In certain embodiments, RCA comprises multiple displacement amplification (MDA).

In certain embodiments, the method further comprises resolving the RCA concatemer product, preferably with said an enzyme specific for said 3rd unique enzyme cleavage site (if present), or with said unique first enzyme or said unique second enzyme.

25   In certain embodiments, the method further comprises self-ligating digested RCA concatemer product under conditions favoring the formation of intramolecular self-ligation (*e.g.*, low concentration).

In certain embodiments, the conditions comprise emulsifying the digested RCA concatemer product to form single oil droplets, each comprising no more than one linear DNA

30   fragment, to promote self-ligation.

In certain embodiments, the method further comprises treating ligation mixture of the digested RCA product with an exonuclease to eliminate linear DNA fragments.

In certain embodiments, step (a) is carried out by nuclease mediated homology directed repair (HDR), or homologous recombination (HR), or a combination thereof.

Another aspect of the invention provides a method of replacing a host (*e.g.*, mouse) genomic region in a host genome with a (syntenic) DNA fragment from a donor (*e.g.*, human) genome, the method comprising: (a) replacing the host genomic region with a pair of heterotypic RTSs (*e.g.*, via nuclease enabled HDR, or HR, or a combination thereof), wherein said pair of heterotypic RTSs flank an optional positive/negative selection cassette (such as hygroTK selection cassette); (b) providing the DNA fragment of any one of claims 1-11 from the donor genome, and allowing the DNA fragment from the donor genome to integrate into the host genome through one of said pair of heterotypic RTSs in the presence of site specific recombinase specific for said one of said pair of heterotypic RTSs; (c) optionally, allowing deletion of said positive/negative selection cassette (if present) through the other of said pair of homotypic RTSs.

Another aspect of the invention provides a method of replacing a host (*e.g.*, mouse) genomic region from a host genome with a (syntenic) DNA fragment from a donor (*e.g.*, human) genome, the method comprising: (a) inserting a single RTS, along with an optional positive/negative selection cassette (such as hygroTK selection cassette), proximal (or distal) to the host genomic region (*e.g.*, via nuclease enabled HDR, or HR, or a combination thereof); (b) providing the DNA fragment of any one of claims 1-11 from the donor genome, and allowing the DNA fragment from the donor genome to integrate into the host genome through said RTS in the presence of site specific recombinase specific for said RTS; (c) optionally, allowing deletion of said positive/negative selection cassette (if present) and the host genomic region through nuclease mediated NHEJ.

In certain embodiments, said host genomic region is deleted prior to step (b).

In certain embodiments, the method is carried out in a zygote, an oocyte, a sperm cell (spermatogonial stem cell line), or an ES cell of the host, preferably by microinjecting, electroporating, or transfecting exogenous components (*e.g.*, CRISPR/Cas9 and guide RNAs targeting said first and said second unique enzyme cleavage sites; megacircles; recombinase protein or coding sequence thereof).

In certain embodiments, the positive/negative selection marker cassette comprises a (neomycin/puromycin/hygromycin/blasticidin/zeocin) resistant / TK (or HPRT) gene under the expression control a eukaryotic promoter (such as the PGK promoter) and a polyA coding sequence.

5   Another aspect of the invention provides a mouse generated by any one of the subject methods.

Another aspect of the invention provides a mouse comprising in its genome an exogenous genomic DNA from a donor (e.g., a human, a mammal, or different mouse strain), wherein the exogenous genomic DNA comprises a polymorphism within the species of the donor, and

10   wherein the exogenous genomic DNA comprises any one of the subject DNA fragments.

In certain embodiments, the exogenous genomic DNA comprises an immunoglobulin heavy chain gene cluster, immunoglobulin kappa light chain gene cluster, immunoglobulin lambda chain gene cluster, a TCR, an MHC class 1 gene, an MHC class 2 gene, or a cytochrome P450 gene.

15   It should be understood that any one embodiment described herein, including those only described under one aspect of the invention, can be combined with any one or more other embodiments, unless explicitly disclaimed or improper.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a schematic drawing showing a first strategy ("Strategy 1") for generating one

20   embodiment of the megacircle of the invention by rolling cycle amplification (RCA). The generated megacircle can be used in RMCE. In Step 1: A pair of unique enzyme cleavage sites, including restriction enzyme (RE) sites or protospacer/PAM sequences for a nuclease, such as that for Cpf1, is identified as flanking the genomic region of interest / gene of interest (collectively "GOI"). Sometimes two unique enzyme cleavage sites can be identified for certain

25   GOI. Human genomic DNA is isolated and cleaved with the nuclease or restriction enzyme (RE) to release the genomic DNA fragment from its chromosomal location. The two sites are chosen such that they are incompatible with each other and as such, it precludes self ligation of the linear DNA fragment into a circle in a ligation reaction. Concurrently, a recombinant DNA construct is created to carry a positive selection marker cassette flanked by a pair of heterotypic recombinase

30   target sites (RTSs). In addition, it may carry a pair of protospacer/PAM sequences or restriction

sites matching those flanking the GOI as depicted. When cleaved with the corresponding nuclease or RE, the recombinant DNA construct carries incompatible sticky ends in cis, but in trans, the sticky ends perfectly match those flanking the GOI. These two species, the digested linear genomic DNA and the recombinant DNA construct, are brought into a ligation reaction to

5   be ligated into a "megacircle." Step 2: the megacircle DNA serving as a template is amplified through RCA so a large quantity of the megacircle could be generated. Upon completion of RCA, individual megacircle units are released from the linear concatemer by digestion with an enzyme specific for a unique enzyme cleavage site, such as an I-SceI site that had been incorporated into the recombinant DNA construct earlier. The resolved RCE products are then

10   self-ligated to re-form megacircles, which can be used in RMCE.

FIG. 2 is a schematic drawing showing a second strategy ("Strategy 2") for generating one embodiment of the megacircle of the invention by rolling cycle amplification (RCA). The generated megacircle can be used in RMCE. In Step 1: A positive selection cassette (+a), along with an RTS site and an I-SceI site, is inserted into the proximal end of the GOI by homologous

15   recombination (HR), nuclease mediated homology directed repair (HDR), or a combination of both. It is followed by a second targeting event, inserting the second selection cassette (+b), along with the heterotypic RTS site and a second I-SceI site, to the distal end of the GOI. Step 2: Genomic DNA is isolated and genomic DNA fragment carrying GOI flanked by the positive selection cassette and an RTS site at the proximal end and a heterotypic RTS site at the distal end

20   released from its chromosomal location by I-SceI digestion. This DNA fragment is self-ligated (compatible stick ends) to form the megacircle molecule. Step 3: the megacircle is amplified by RCA, using a primer (or multiple primers) unique to the GOI. Upon completion of the RCA, the RCA concatemer will be resolved to release individual "megastrand" units, each can self-ligate back into a megacircle for use in RMCE.

25   FIG. 3 a schematic drawing showing the RMCE Reaction using the subject megacircles described in FIGs. 1 and 2. Step 1: A pair of heterotypic RTS sites is inserted into a target chromosomal location, replacing the murine syntenic region (B), either by HR, nuclease mediated HDR, or a combination of both. An optional positive/negative selection cassette is incorporated to allow selection and enrichment of the successful recombinant. Step 2A: the

30   human GOI, carried in a megacircle, is integrated into the genomic locus, mediated by recombinase mediated integration. Integration could occur through either pair of the two pairs of

homotypic sites in trans, each generating a different intermediate product. Step 2B: Selection cassette is excised from the locus when the other pair of homotypic RTS sites in cis, as compared to the pair used in Step 2A, is employed, and mutant allele with GOI replacing the selection cassette generated. Note that in step 2B, although the other homotypic pair in cis could also in theory recombine and regenerate the two substrates that started the RMCE reaction, this reaction is not favored / efficient, as the two sites are separated by the large GOI.

FIG. 4A shows the outcome of the murine Ighv, d, j & c genes when RMCE is employed to generate mice carrying a complete repertoire of the human IGHV, D & J genes.

FIG. 4B shows the human IGHV, D & J genes included for humanization when RMCE is employed to generate mice carrying a complete repertoire of the human IGHV, D & J genes.

FIG. 4C shows the overall flow scheme when RMCE is employed (with Cre/Lox as an example) to generate mice carrying a complete repertoire of the human IGHV, D & J genes.

FIG. 5 is a schematic drawing showing a first strategy ("Strategy 1") for generating one embodiment of the megacircle of the invention by rolling cycle amplification (RCA). The generated megacircle can be used in recombinase mediated integration. This is similar to the flow scheme outlined in FIG. 1, but with only one RTS site incorporated into the megacircle.

FIG. 6 is a schematic drawing showing a first strategy ("Strategy 2") for generating one embodiment of the megacircle of the invention by rolling cycle amplification (RCA). The generated megacircle can be used in recombinase mediated integration. This is similar to flow scheme outlined in FIG. 2, but with only one RTS site incorporated into megacircle. In Step 1, only the proximal end of GOI is targeted, and an optional selection cassette, along with an RTS site and an I-Sce1 site, are incorporated.

FIG. 7 shows recombinase mediated integration. In Step 1: A single RTS site is inserted into the chromosomal location, along with a positive (+a)/negative(-) selection cassette, either by HR, nuclease mediated HDR, or a combination of both. Step 2: The human GOI, carried in a megacircle, is integrated into the genomic locus, mediated by the corresponding site specific recombinase (SSR). The positive selection cassette (+b) is different from the selection marker gene (+a) in Step 1. Step 3: The murine syntenic sequence (B), along with the two selection cassettes, are eliminated from the targeted allele by non homologous end joining (NHEJ). The mutant allele now carries the human GOI replacing murine syntenic sequence (B).

FIG. 8A shows the outcome of the murine Ighv, d, j & c genes when recombinase mediated integration is employed to generate mice carrying a complete repertoire of the human IGHV, D & J gene.

FIG. 8B shows the human IGHV, D & J genes included for humanization when recombinase mediated integration is employed to generate mice carrying a complete repertoire of the human IGHV, D & J genes.

FIG. 8C shows the overall flow scheme when recombinase mediated integration is employed (with Cre/Lox as an example) to generate mice carrying a complete repertoire of the human IGHV, D & J genes.

FIG. 9 shows overall activity for recombinase mediated integration approach.

DETAILED DESCRIPTION OF THE INVENTION

1.   Introduction

The invention described herein provides a donor DNA molecule ("linear or circular DNA fragment" also known as megastrand or megacircle, respectively) that satisfies several criteria: one, it can be large enough to accommodate the entire gene or gene cluster (in the hundreds Kb to mega base pair range), though it can be as small as encompassing a partial gene or certain non-coding sequence; second, it has a pre-determined start and end (*e.g.*, proximal and distal ends in relation to the centromere of the chromosomal from which the DNA fragment originates), marking precisely the boundary of the gene or gene cluster. In addition, such a donor DNA molecule contains additional exogenous sequence, such as a recombinase target sequence (RTS), and preferably, the DNA fragment is devoid of a sequence sufficient for or required for self-replication in a host cell (*e.g.*, the ORI sequence in plasmids etc. required for DNA self-replication in a host cell, pBR322, pUC, pSC101, or ARS).

The megastrand or megacircle of the invention can be used in numerous settings, including (but not limited to) humanize large sections of a mouse genome with donor DNA from a single individual (such as a single human), as opposed to (smaller) donor DNAs from multiple individuals.

Aspects of the invention are described in further details below.

2.   Megacircle Generation

The invention described herein provides a method that generates a large genomic DNA fragment called "megastrand," and when circularized, "megacircle," which typically has a size in the range of hundreds of Kb to million base pair range (FIGs. 1, 2, 5, & 6). According to this method, for any specific genomic DNA fragment of interest, megastrand or megacircle can be

5   generated in vitro in large quantity, free of a prokaryotic or eukaryotic vector backbone sequence, and with no need to propagate in bacterial, yeast, or other host cells.

As outlined in FIGs. 1 and 5, for the genomic DNA sequence identified, appropriate restriction sites flanking the fragment are identified. The two sites should be different such that they will preclude a head to tail self-ligation in a ligation reaction of the genomic fragment when

10   released from its genomic location. Genomic DNA is then isolated and cleaved with the corresponding RE to generate the genomic fragment. Alternatively, site specific nuclease, such as Cas9 and Cas9 nickase (Jinek et al., 2012, incorporated by reference), Cpf1 (Zetsche et al., 2015, incorporated by reference), or PfAgo (Enghiad and Zhao, 2017, ACS Synth. Biol., 2017, 6 (5), pp 752–757, incorporated herein by reference, which describes a Pyrococcus furiosus

15   Argonaute (PfAgo) based platform for generating artificial restriction enzymes (AREs) capable of recognizing and cleaving DNA sequences at virtually any arbitrary site and generating defined sticky ends of varying length) can be used to specifically release the target genomic DNA fragment from the chromosome. Concurrently, a recombinant DNA construct is synthesized, carrying a selection marker gene flanked by restriction sites or nuclease cleavage sites matching

20   those of the genomic fragment as depicted. The two species, the genomic DNA fragment and the recombinant DNA carrying matching ends, are then ligated to form a circle (megacircle). To enrich for the megacircle, the ligation mixture can be treated with exonuclease to eliminate all fragments that remain linear. While circles may be formed from all genomic fragments bearing compatible sites (in the case of using CRISPR or PfAgo, the circle should be more specific),

25   amplification of the target circle can be accomplished using RCA (Ali et al., 2014). If needed, megastrand DNA can also be enriched by affinity purification or separation on pulsed field gel electrophoresis before used in a ligation reaction.

In an RCA reaction, phi29 (Ø29) DNA polymerase (Blanco et al., 1989; Blanco and Salas, 1984) is commonly used, and one or multiple primers specific for the target genomic region

30   is/are provided in the reaction. Considering that the rate of synthesis of phi29 DNA polymerase is about 50 nucleotides per second, for a 500 kb fragment, it takes about 10,000 seconds (or 3

hours) to complete one round of synthesis for a single strand. However, in some embodiments, using multiple displacement amplification (MDA) strategy, about 100 or more oligonucleotides can be made to hybridize to one strand and to tile the entire length of the megacircle template. In the same reaction, another 100 or more oligonucleotides are provided targeting the other strand.

5   After 30 hours, minimally, the target fragment will have been amplified 1,000 times, as linear double-stranded concatemers. The RCA product can then be digested with an enzyme, such as a restriction enzyme which cuts once per unit of the concatemer, and the individual unit ligated under low substrate concentration to form the circle. With the high fidelity of phi29 DNA polymerase (error rate $1/10^6$ - $1/10^7$), this process potentially can be repeated multiple times.

10   When the target circle is sufficiently enriched, random hexamer can also be used for additional amplification with higher efficiency. To improve efficiency, exonuclease-resistant primer or random-hexamer primers with thiophosphate linkages for the two 3' terminal nucleotides can be used in RCA reaction.

Alternatively, GOI can also be flanked with I-Sce1 sites (FIG. 2) or with an engineered

15   nuclease cleavage site incorporated into one end of the GOI, matching the naturally occurring site at the other end of GOI (FIG. 6). These strategies require the use of HR, nuclease mediated HDR, or a combination of both to insert the selection cassette into an endogenous locus. However, the advantage is that it allows a head to tail self-ligation of the megastrand DNA in cis to form a circle, as compared with the strategies outlined in FIGs. 1 and 5, which require ligation

20   between two molecular species. After megacircle is obtained and multiple rounds of amplification from RCA, the final product is digested and self ligated to form a megacircle. In the case that the self-ligation efficiency is low, emulsion can be used to separate each molecule into a single oil droplet to promote intramolecular ligation efficiency. Linear fragment that remains will be eliminated by exonuclease treatment, improving the relative abundance of

25   megacircle.

In the case that the final concentration of megacircle is low, to support recombinase mediated integration or RMCE, the megacircle preparation can be injected directly into the nucleus of mouse embryonic stem cells, as compared to delivery through electroporation, which may require a higher concentration. If mouse strains carrying the landing pad would have been

30   created and zygotes available, megacircle could be delivered directly into the zygotes, along with the recombinase, to elicit RMCE or recombinase mediated integration.

The strategies outlined in FIGs. 1 and 5 are particularly appealing, as they do not require targeting of the human chromosome in human ES or iPS cells - the two human cell types are not known for high targeting efficiency and may be challenging to work with.

In addition to generating megacircle through RCA, it may also be assembled by Gibson
5 assembly or synthesized de novo.

3.   *Employing RMCE for Large Scale Genome Engineering*

HR is a naturally occurring mechanism that repairs DNA damage and exchanges DNA content between sister chromatids during mitosis and meiosis. Exploiting into this mechanism, precise changes can be introduced into the genome of eukaryotic cells (Doetschman et al., 1987;
10 Thomas and Capecchi, 1987). Although powerful, it is a laborious and time consuming process, more so when it comes to large scale genome engineering. For example, to replace the murine Igh locus with its human syntenic region (940 kb), using BAC targeting vectors, it took 9 rounds of manipulation in mouse ESCs and years of work to complete the project (Macdonald et al., 2014).

15   Over the years, another line of work explored into RMCE to complement HR and to facilitate large scale genome engineering (Hasegawa et al., 2011; Wallace et al., 2007). By using HR to insert a "landing pad" into the Igh locus and employing iterative rounds of RMCE, the Kymab group was able to accomplish a humanization effort similar to the Regeneron VelocImmune mice but reducing the task of mouse ESC manipulations to 5 rounds (Lee et al.,
20 2014). The added advantage for RMCE based approach is the much high targeting efficiency (35% overall for RMCE versus less than 1% for HR, using BAC based targeting vector), which should translate into less screening work and a better timeline. The limiting factor, in the case of RMCE, as described earlier, is the BAC vector employed. With a carrying capacity of 150-350 kb per targeting vector, compared with 1.0 Mb of the human IGHV, D, and J genes that need to
25 be transferred, it necessitates multiple rounds of transfer.

SSR binds to a short stretch of sequence, RTS, and promotes strand exchange between the two substrates to form two strand-exchanged products (O'Gorman et al., 1991; Sauer and Henderson, 1988; Thorpe and Smith, 1998; Thyagarajan et al., 2001). For Cre recombinase, the target site is 34 bps in length (LoxP), consisting of two 13 bps of palindromic repeats separated
30 by an 8 bps spacer sequence. One recombinase molecule binds to one 13 bps repeat sequence

and as such, each 13 bps repeat constitutes a recombinase binding element (RBE). For Flp, the minimal target site (FRT) is similar in structure to the LoxP site for Cre recombinase, although the 13 bps repeats are not strictly palindromic but differ by one base pair. Also the full FRT site has an additional 13 bps of repeat, located 5' to and separated from the minimal sequence by one additional nucleotide, for a total of 48 bps. This segment is not needed for excision but essential for integration, including RMCE. For θC31, the target sites are called attP and attB. Each site consists of two 18 bps binding elements separated by two nucleotides of spacer sequence and the two 18 bps binding elements are not particularly palindromic. For the two RTS sites that are brought together to recombine, there are 4 recombinase molecules involved, each binding to one RBE, forming a synapse (Van Duyne, 2001).

RMCE was invented 35 years ago (Schlake and Bode, 1994) (Figure 3). To start the process, two RTS sites that do not recombine between themselves (heterotypic), often because of the mutations introduced into the spacer sequence, and made to flank a selection cassette, are integrated into the genomic locus often by gene targeting. This serves as the "landing pad" onto which other cassette can land and swap with. To swap, a donor plasmid, or megacircle in our case, carrying a matching set of heterotypic RTS sites will then be introduced. In this scheme, although the two RTS sites in cis are heterotypic, either of the two pairs of RTS sites in trans is homotypic (the same) to each other. Upon exposure to the corresponding recombinase, the two homotypic pairs of RTS sites will synapse and strand exchange occur such that the selection cassette located in the landing pad will be replaced by the GOI from the donor plasmid. Employing RMCE, a transgene can be inserted into any genomic locus that has been engineered to carry the RTS sites. Compared with conventional transgenesis, insertion of a transgene into a safe harbor site is preferred, as performance of the transgene is more predictable. This is because the transgene can be inserted into a genomic locus that is known to support expression of the transgene and it can be done such that only one copy of the transgene is incorporated. It is particularly convenient if a strain that carries a landing pad in a safe harbor is available, and for new needs, swapping can be arranged for incorporation of the GOI. The other advantage is the higher targeting efficiency for RMCE, as compared to a direct gene targeting effort, particularly for larger GOI, as mentioned earlier.

RMCE is a two-step process (Malchin et al., 2010; Schlake and Bode, 1994; Takata et al., 2011) (Figure 3). The first step, by nature, is recombinase mediated integration. It occurs in

trans, between the pair of homotypic RTS sites. In this process, the circular donor DNA is integrated into the recipient site, creating an intermediate product. Depending on the pair of homotypic sites involved, two different intermediate products are created which will then serve as the substrates for the next step. The second step in an RMCE reaction is recombinase

5   mediated excision. It occurs in cis, between the two homotypic recognition sites now brought in cis. If it is the same pair as used in step 1 (integration), it reproduces the two substrates that started the RMCE reaction. However, if it is the other pair, two products will be produced, with one being the desired product, now with the selection marker gene replaced with the GOI. As the enzyme is the same for the two steps, RMCE should progress seamlessly from step 1 to step

10   2, or in the reverse direction, reaching equilibrium as defined by the reaction.

Based on this understanding, in the scenario of replacing a murine genomic region with its human syntenic sequence by RMCE or recombinase mediated integration, the donor needs to be a circular molecule so it can be integrated into the genomic locus and it has to be large enough to accommodate a gene or gene cluster. These two criteria can be met with our megacircle as

15   described above.

For other utilities, a megastrand can be produced by digesting the concatemer produced from an RCA reaction with the appropriate restriction enzyme (I-Sce1 in FIGs. 1, 2, 5 & 6). It can then be used directly as a large transgene or modified further for gene targeting in an HR reaction or nuclease mediated HDR.

20   *4.   Humanization of the Immunoglobulin Heavy Chain Locus*

Two different ways, RMCE and recombinase mediated integration, can be used to replace the murine Ighv, d, and j genes with their human ortholog genes, using megacircle as the vehicle to transfer human syntenic sequence.

FIGs. 4A and 4B depict fate of the murine Igh and human IGH genes (Walter et al.,

25   1990), alleles, and technical elements incorporated in the process to support a RMCE scheme. As illustrated in FIG. 4A, for convenience, the murine Igh locus could be divided into 4 blocks. Block 1 is 2.5 Mb and marked by Ighv1-86 at the very 5' end and Igv5-1 gene at the 3' end. This block is deleted employing, for example, CRISPR mediated NHEJ mutations; block 2 is 91 kb in size and includes Adam6a and Adam6b genes that is retained, as these two genes are

30   important for fertility of the sperm; block 3 is 54 kb in size and starts with Ighd1-1 and ends with

Ighj4 genes. This region is deleted and replaced with the PGK/HygroTK type +/- cassette,

bringing along the heterotypic LoxP/Lox5717 sites in preparation for RMCE; block 4 carries the

Igh enhancer sequence, switch region, and constant region genes in their natural figuration and is

left unperturbed. To modify the murine Igh allele in preparation for acceptance of the human

5   IGH genes, a positive/negative selection cassette (Hygro/TK), along with the LoxP/Lox5171

sites, are inserted into the murine Igh locus, replacing 54 kb of the murine Igh genes from Ighd1-

1 to Ighj4. This can be accomplished by HR, nuclease mediated HDR, or a combination of both,

in mouse embryonic stem cells.

      For convenience, the human IGH locus can also be divided into 3 blocks (FIG. 4B).

10   Block 1 starts with IGHV(III)-82 gene and ends with IGHV(II)-74-1 gene; block 2 is 948 kb in

size and marked by IGHV3-74 at the 5' end and IGHJ6 at the 3' end. Genes in this block are

transferred to and replace the murine Igh locus; block 3 starts with the enhancer sequence and

extends to the rest of the human chromosome 14 in its natural configuration. To modify the

allele in preparation for megacircle generation, gene targeting is performed to insert a LoxP site

15   and an I-Sce1 site between blocks 1 and 2, along with a positive selection cassette. When

targeted clone(s) is identified, they can be driven to homozygocity by a "loss of heterozygocity"

protocol, when cultured in a high concentration of G418. After homozygote clones are obtained,

they can be modified further to carry the Lox5171 site, along with an I-Sce1 site and a positive

selection cassette, between blocks 2 and 3, by HR, nuclease mediated HDR, or a combination of

20   both. The IGH allele with both LoxP and Lox5171 sites incorporated and flanked by the I-Sce1

sites can then be digested with I-Sce1 enzyme and megastrand released. Upon circularization,

the megacircle can be used as the template for megacircle amplification as outlined in FIG. 2.

Alternatively, a megacircle can be prepared as outlined in FIG. 1, without gene targeting in

human cells.

25   FIG. 4C illustrates how this scheme is carried out. When a sufficient quantity of

megacircles are obtained, they can be electroporated into mouse ESCs that have been engineered

to carry the LoxP/Lox5171 sites at the murine Igh locus, along with Cre recombinase, to initiate

RMCE. By RMCE, the human IGH sequence is incorporated into the murine Igh locus. If

desired, the 3' Lox5171 site can be flanked by PiggyBac transposon sequences and, when

30   exposed to PiggyBac, the transposon is eliminated, along with the Lox517 site. This leaves no

"footprint" of genome engineering in the region between human IGH and murine constant region

genes. At the 5' end of the recombined allele, there is a single LoxP site remaining between the murine Adam6b and the IGHV3-74 genes.

Megacircle can also be engineered to support recombinase mediated integration as outlined in FIG. 7. To use this scheme to incorporate the human IGH genes into the murine Igh locus, a detailed plan is illustrated in FIGs. 8A, 8B, and 8C. This approach may be technically less challenging, as it requires only one integration event. The murine Igh locus is modified to carry the single RTS (FIG. 8A). To release the IGH genomic fragment from human chromosome 14, a Cpf1 protospacer/PAM sequence is first identified marking the 3' end of IGH block 2. This same Cpf1 protospacer/PAM sequence is then incorporated into the targeting vector and inserted into the region between blocks 1 and 2, as illustrated in FIG. 8B. If possible, restriction sites flanking block 2 can also be identified and used. As shown in FIG. 8C, the product from recombinase mediated integration carries the two selection cassettes 3' to the IGH genes and can be deleted using nuclease mediated NHEJ mutations flanking the region. As an alternative, megacircle prepared according to FIG. 5 can also be used for recombinase mediated integration to insert the human IGH genes into the murine locus. In this case, a megacircle is generated by ligating the human IGH genomic fragment and a recombinant DNA construct carrying compatible ends, without the need to engineer the human IGH genes in a human cell line (ESC or iPSCs).

To accomplish these tasks, multiple technical elements are employed and incorporated, including homotypic and heterotypic Lox sites, I-Sce1 site (or restriction sites for other rare cutter enzymes, such as other homing-endonucleases), and Cpf1 protospacer/PAM sequences, based on well-known molecular biology mechanisms, including HR, RMCE, recombinase mediated integration, nuclease mediated NHEJ, nuclease mediated HDR, and rolling cycle amplification. All elements can be tested individually and in a flow scheme, such as the one illustrated in FIG. 9. Of particular interest is the minicircle, which shares all technical elements as the megacircle but does not carry the large "GOI." Successful recombinase mediated integration or RMCE by this minicircle demonstrates that the system performs as expected.

References

1. Ali, M.M., Li, F., Zhang, Z., Zhang, K., Kang, D.K., Ankrum, J.A., Le, X.C., Zhao, W., 2014. Rolling circle amplification: a versatile tool for chemical biology, materials science and medicine. Chemical Society reviews 43, 3324-3341.

2. Blanco, L., Bernad, A., Lazaro, J.M., Martin, G., Garmendia, C., Salas, M., 1989. Highly efficient DNA synthesis by the phage phi 29 DNA polymerase. Symmetrical mode of DNA replication. The Journal of biological chemistry 264, 8935-8940.

3. Blanco, L., Salas, M., 1984. Characterization and purification of a phage phi 29-encoded DNA polymerase required for the initiation of replication. Proceedings of the National Academy of Sciences of the United States of America 81, 5325-5329.

4. Bruggemann, M., Caskey, H.M., Teale, C., Waldmann, H., Williams, G.T., Surani, M.A., Neuberger, M.S., 1989. A repertoire of monoclonal antibodies with human heavy chains from transgenic mice. Proceedings of the National Academy of Sciences of the United States of America 86, 6709-6713.

5. Burke, D.T., Carle, G.F., Olson, M.V., 1987. Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science 236, 806-812.

6. Doetschman, T., Gregg, R.G., Maeda, N., Hooper, M.L., Melton, D.W., Thompson, S., Smithies, O., 1987. Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. Nature 330, 576-578.

7. Enghiad, B., Zhao, H., 2017. Programmable DNA-Guided Artificial Restriction Enzymes. ACS synthetic biology 6, 752-757.

8. Giudicelli, V., Duroux, P., Ginestoux, C., Folch, G., Jabado-Michaloud, J., Chaume, D., Lefranc, M.P., 2006. IMGT/LIGM-DB, the IMGT comprehensive database of immunoglobulin and T cell receptor nucleotide sequences. Nucleic acids research 34, D781-784.

9. Green, L.L., 2014. Transgenic mouse strains as platforms for the successful discovery and development of human therapeutic monoclonal antibodies. Current drug discovery technologies 11, 74-84.

10. Green, L.L., Hardy, M.C., Maynard-Currie, C.E., Tsuda, H., Louie, D.M., Mendez, M.J., Abderrahim, H., Noguchi, M., Smith, D.H., Zeng, Y., David, N.E., Sasai, H., Garza, D., Brenner, D.G., Hales, J.F., McGuinness, R.P., Capon, D.J., Klapholz, S., Jakobovits, A.,

1994. Antigen-specific human monoclonal antibodies from mice engineered with human Ig heavy and light chain YACs. Nature genetics 7, 13-21.

11. Hasegawa, M., Kapelyukh, Y., Tahara, H., Seibler, J., Rode, A., Krueger, S., Lee, D.N., Wolf, C.R., Scheer, N., 2011. Quantitative prediction of human pregnane X receptor and cytochrome P450 3A4 mediated drug-drug interaction in a novel multiple humanized mouse line. Molecular pharmacology 80, 518-528.

12. Hohn, B., Murray, K., 1977. Packaging recombinant DNA molecules into bacteriophage particles in vitro. Proceedings of the National Academy of Sciences of the United States of America 74, 3259-3263.

13. Hoshiya, H., Kazuki, Y., Abe, S., Takiguchi, M., Kajitani, N., Watanabe, Y., Yoshino, T., Shirayoshi, Y., Higaki, K., Messina, G., Cossu, G., Oshimura, M., 2009. A highly stable and nonintegrated human artificial chromosome (HAC) containing the 2.4 Mb entire human dystrophin gene. Molecular therapy : the journal of the American Society of Gene Therapy 17, 309-317.

14. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., Charpentier, E., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816-821.

15. Lee, E.C., Liang, Q., Ali, H., Bayliss, L., Beasley, A., Bloomfield-Gerdes, T., Bonoli, L., Brown, R., Campbell, J., Carpenter, A., Chalk, S., Davis, A., England, N., Fane-Dremucheva, A., Franz, B., Germaschewski, V., Holmes, H., Holmes, S., Kirby, I., Kosmac, M., Legent, A., Lui, H., Manin, A., O'Leary, S., Paterson, J., Sciarrillo, R., Speak, A., Spensberger, D., Tuffery, L., Waddell, N., Wang, W., Wells, S., Wong, V., Wood, A., Owen, M.J., Friedrich, G.A., Bradley, A., 2014. Complete humanization of the mouse immunoglobulin loci enables efficient therapeutic antibody discovery. Nat Biotechnol 32, 356-363.

16. Lonberg, N., Taylor, L.D., Harding, F.A., Trounstine, M., Higgins, K.M., Schramm, S.R., Kuo, C.C., Mashayekh, R., Wymore, K., McCabe, J.G., et al., 1994. Antigen-specific human antibodies from mice comprising four distinct genetic modifications. Nature 368, 856-859.

17. Macdonald, L.E., Karow, M., Stevens, S., Auerbach, W., Poueymirou, W.T., Yasenchak, J., Frendewey, D., Valenzuela, D.M., Giallourakis, C.C., Alt, F.W., Yancopoulos, G.D.,

Murphy, A.J., 2014. Precise and in situ genetic humanization of 6 Mb of mouse immunoglobulin genes. Proceedings of the National Academy of Sciences of the United States of America 111, 5147-5152.

18. Malchin, N., Molotsky, T., Borovok, I., Voziyanov, Y., Kotlyar, A.B., Yagil, E., Kolot, M., 2010. High efficiency of a sequential recombinase-mediated cassette exchange reaction in Escherichia coli. Journal of molecular microbiology and biotechnology 19, 117-122.

19. O'Gorman, S., Fox, D.T., Wahl, G.M., 1991. Recombinase-mediated gene activation and site-specific integration in mammalian cells. Science 251, 1351-1355.

20. Osoegawa, K., Mammoser, A.G., Wu, C., Frengen, E., Zeng, C., Catanese, J.J., de Jong, P.J., 2001. A bacterial artificial chromosome library for sequencing the complete human genome. Genome research 11, 483-496.

21. Sauer, B., Henderson, N., 1988. Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. Proceedings of the National Academy of Sciences of the United States of America 85, 5166-5170.

22. Schlake, T., Bode, J., 1994. Use of mutated FLP recognition target (FRT) sites for the exchange of expression cassettes at defined chromosomal loci. Biochemistry 33, 12746-12751.

23. Takata, Y., Kondo, S., Goda, N., Kanegae, Y., Saito, I., 2011. Comparison of efficiency between FLPe and Cre for recombinase-mediated cassette exchange in vitro and in adenovirus vector production. Genes to cells : devoted to molecular & cellular mechanisms 16, 765-777.

24. Taylor, L.D., Carmack, C.E., Schramm, S.R., Mashayekh, R., Higgins, K.M., Kuo, C.C., Woodhouse, C., Kay, R.M., Lonberg, N., 1992. A transgenic mouse that expresses a diversity of human sequence heavy and light chain immunoglobulins. Nucleic acids research 20, 6287-6295.

25. Thomas, K.R., Capecchi, M.R., 1987. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell 51, 503-512.

26. Thorpe, H.M., Smith, M.C., 1998. In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. Proceedings of the National Academy of Sciences of the United States of America 95, 5505-5510.

27. Thyagarajan, B., Olivares, E.C., Hollis, R.P., Ginsburg, D.S., Calos, M.P., 2001. Site-specific genomic integration in mammalian cells mediated by phage phiC31 integrase. Molecular and cellular biology 21, 3926-3934.

28. Tomizuka, K., Yoshida, H., Uejima, H., Kugoh, H., Sato, K., Ohguma, A., Hayasaka, M., Hanaoka, K., Oshimura, M., Ishida, I., 1997. Functional expression and germline transmission of a human chromosome fragment in chimaeric mice. Nature genetics 16, 133-143.

29. Van Duyne, G.D., 2001. A structural view of cre-loxp site-specific recombination. Annual review of biophysics and biomolecular structure 30, 87-104.

30. Wallace, H.A., Marques-Kranc, F., Richardson, M., Luna-Crespo, F., Sharpe, J.A., Hughes, J., Wood, W.G., Higgs, D.R., Smith, A.J., 2007. Manipulating the mouse genome to engineer precise functional syntenic replacements with human sequence. Cell 128, 197-209.

31. Walter, M.A., Surti, U., Hofker, M.H., Cox, D.W., 1990. The physical organization of the human immunoglobulin heavy chain gene complex. The EMBO journal 9, 3303-3313.

32. Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., Koonin, E.V., Zhang, F., 2015. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell 163, 759-771.

All references listed herein are incorporated by reference, including incorporation by reference at the instances where the specific reference is cited.

ABBREVIATIONS

GOI	genomic region of interest
HDR	homology directed repair
HR	homologous recombination
IG	immunoglobulin
Igh	murine immunoglobulin heavy chain gene
IGH	human immunoglobulin heavy chain gene
MHC	major histocompatibility complex
NHEJ	non homologous end joining
RBE	recombinase binding element
RCA	rolling circle amplification
RE	restriction enzyme
RMCE	recombinase mediated cassette exchange
RTS	recombinase target site
SSR	site specific recombinase

WHAT IS CLAIMED IS:

1. A DNA fragment comprising a genomic region of interest (GOI, which can be a single complete or partial gene, or a cluster of related or unrelated genes, non-encoding sequence, or any genomic region of interest), wherein the GOI is at least about 20 kb (*e.g.*, 50 kb, 100 kb, 150 kb, 200 kb, 300 kb, 400 kb, 500 kb, 750 kb, 1 mb, 1.25 mb, 1.5 mb, 1.75 mb, 2 mb or more), and is flanked by a first unique enzyme cleavage site at the proximal end of the genomic DNA fragment and a second unique enzyme cleavage site at the distal end of the genomic DNA fragment, wherein the DNA fragment further comprises an exogenous sequence (such as a selection marker, a recombinase target site (RTS), or a sequence from a non-donor source) between the proximal end of the DNA fragment and the proximal end of the GOI, or between the distal end of the DNA fragment and the distal end of the GOI; and wherein the DNA fragment is devoid of a sequence required for self-replication in a host cell (*e.g.*, ORI, pBR322, pUC, pSC101, or ARS).

2. The DNA fragment of claim 1, wherein the first and the second unique enzyme cleavage sites have the same sequence, or are recognized by one (the same) cleavage enzyme, or have sticky ends compatible for ligation.

3. The DNA fragment of claim 1 or 2, which is a double-stranded circle (*e.g.*, linked at the first and the second unique enzyme cleavage sites).

4. The DNA fragment of any one of claims 1-3, further comprising a first positive selection marker cassette located between the RTS and an end of the DNA fragment, wherein the GOI and the positive selection marker cassette flank the RTS.

5. The DNA fragment of claim 4, further comprising a heterotypic recombinase target site (htRTS), wherein the RTS and the htRTS flank the GOI.

6. The DNA fragment of claim 5, which is circular, with a linkage created by linking the first and the second unique enzyme cleavage sites, wherein the linkage is between the positive selection marker cassette and the htRTS (or the RTS).

7. The DNA fragment of any one of claims 2-6, wherein the DNA fragment is amplified by rolling circle amplification (RCA).

8. The DNA fragment of any one of claims 1-7, wherein the GOI is immunoglobulin heavy chain gene cluster, immunoglobulin kappa light chain gene cluster, immunoglobulin

lambda light chain gene cluster, a TCR (T-cell receptor) gene cluster, an MHC class 1 gene, an MHC class 2 gene, or a cytochrome P450 gene.

9. The DNA fragment of any one of claims 1-8, wherein the first and/or the second unique enzyme cleavage site(s) are cleavable by a restriction endonuclease, a homing endonuclease (such as I-Sce1), a CRISPR/Cas9 nuclease, a CRISPR/Cpf1 nuclease, a TALEN (transcription activator-like effector nuclease), a ZFN (Zinc Finger nuclease), or a *Pyrococcus furiosus* Argonaute (PfAgo) based artificial restriction enzyme (ARE).

10. The DNA fragment of any one of claims 1-9, wherein one of the first and the second unique enzyme cleavage sites is a natural / pre-existing genomic sequence.

11. The DNA fragment of any one of claims 1-10, wherein the RTS or the htRTS is locus of crossover in P1 (loxP), flippase recognition target (FRT), or attP/attB.

12. A method of producing the DNA fragment of any one of claims 1-11, the method comprising:

(a) inserting, in a host genome comprising said GOI, said first unique enzyme cleavage site and said RTS proximal (or distal) to the GOI, wherein the GOI is proximal (or distal) to the second unique enzyme cleavage site, and wherein an optional first positive selection marker cassette is distal (or proximal) to said first unique enzyme cleavage site and proximal (or distal) to said RTS;

(b) digesting the host genome with a first unique enzyme that cleaves said first unique enzyme cleavage site, and a second unique enzyme that cleaves said second unique enzyme cleavage site, to release the DNA fragment comprising the GOI from the host genome;

(c) circularizing the released DNA fragment to form a megacircle, by ligating the ends of said released DNA fragment generated by cleavage by said first unique enzyme and said second unique enzyme, under conditions that promote intramolecular self-ligation;

(d) amplifying said megacircle by rolling circle amplification (RCA) with a primer (or multiple primers) unique to said GOI.

13. The method of claim 12, wherein the second unique enzyme cleavage site is naturally existing sequence.

14. The method of claim 12 or 13, wherein step (a) further comprises inserting a 3rd unique

enzyme cleavage site between the first unique enzyme cleavage site and the RTS.

15.   The method of claim 12, wherein the second unique enzyme cleavage site, together with a second positive selection marker cassette and htRTS, are inserted into said host genome distal (or proximal) to GOI, preferably the second unique enzyme cleavage site is distal (or proximal) to the htRTS and proximal (or distal) to the second positive selection marker cassette.

16.   A method of producing the DNA fragment of any one of claims 1-11, the method comprising:

   (a)   removing / retrieving a genomic DNA fragment comprising said GOI from a host genome, through digesting the host genome with a first enzyme specific for a first unique enzyme cleavage site and a second enzyme specific for a second unique enzyme cleavage site, wherein said first and said second unique enzyme cleavage sites are not compatible for ligation with each other;

   (b)   ligating the retrieved genomic DNA fragment in step (a) with a linear recombinant DNA construct to produce a megacircle, wherein said linear recombinant DNA construct comprises a positive selection marker cassette flanked by a pair of heterotypic RTSs or flanked by a RTS and a 3rd unique enzyme cleavage site, and wherein the ends of the linear recombinant DNA construct are defined by said first and said second unique enzyme cleavage sites, respectively;

   (c)   amplifying said megacircle by rolling circle amplification (RCA) with a primer unique to said GOI.

17.   The method of any one of claims 12-16, wherein said first and said second unique enzyme cleavage sites are the same, and said first and said second unique enzymes are the same.

18.   The method of any one of claims 12-17, further comprising, before amplification by RCA, treating the ligation mixture with an exonuclease to eliminate linear DNA fragments.

19.   The method of any one of claims 12-18, wherein RCA is performed using phi29 (θ29) DNA polymerase, and one or more primers each specific for the GOI.

20.   The method of any one of claims 12-19, wherein RCA comprises multiple displacement amplification (MDA).

21. The method of any one of claims 12-20, further comprising resolving the RCA concatemer product, preferably with said an enzyme specific for said 3rd unique enzyme cleavage site (if present), or with said unique first enzyme or said unique second enzyme.

22. The method of claim 21, further comprising self-ligating digested RCA concatemer product under conditions favoring the formation of intramolecular self-ligation (*e.g.*, low concentration).

23. The method of claim 22, wherein said conditions comprise emulsifying the digested RCA concatemer product to form single oil droplets, each comprising no more than one linear DNA fragment, to promote self-ligation.

24. The method of claim 22 or 23, further comprising treating ligation mixture of the digested RCA product with an exonuclease to eliminate linear DNA fragments.

25. The method of any one of claims 12-24, wherein step (a) is carried out by nuclease mediated homology directed repair (HDR), or homologous recombination (HR), or a combination thereof.

26. A method of replacing a host (*e.g.*, mouse) genomic region in a host genome with a (syntenic) DNA fragment from a donor (*e.g.*, human) genome, the method comprising:

(a) replacing the host genomic region with a pair of heterotypic RTSs (*e.g.*, via nuclease enabled HDR, or HR, or a combination thereof), wherein said pair of heterotypic RTSs flank an optional positive/negative selection cassette (such as hygroTK selection cassette);

(b) providing the DNA fragment of any one of claims 1-11 from the donor genome, and allowing the DNA fragment from the donor genome to integrate into the host genome through one of said pair of heterotypic RTSs in the presence of site specific recombinase specific for said one of said pair of heterotypic RTSs;

(c) optionally, allowing deletion of said positive/negative selection cassette (if present) through the other of said pair of homotypic RTSs.

27. A method of replacing a host (*e.g.*, mouse) genomic region from a host genome with a (syntenic) DNA fragment from a donor (*e.g.*, human) genome, the method comprising:

(a) inserting a single RTS, along with an optional positive/negative selection cassette (such as hygroTK selection cassette), proximal (or distal) to the host genomic region (*e.g.*, via nuclease enabled HDR, or HR, or a combination thereof);

(b)   providing the DNA fragment of any one of claims 1-11 from the donor genome, and allowing the DNA fragment from the donor genome to integrate into the host genome through said RTS in the presence of site specific recombinase specific for said RTS;

(c)   optionally, allowing deletion of said positive/negative selection cassette (if present) and the host genomic region through nuclease mediated NHEJ.

28.   The method of claim 27, wherein said host genomic region is deleted prior to step (b).

29.   The method of any one of claims 26-28, which is carried out in a zygote, an oocyte, a sperm cell (spermatogonial stem cell line), or an ES cell of the host, preferably by microinjecting, electroporating, or transfecting exogenous components (*e.g.*, CRISPR/Cas9 and guide RNAs targeting said first and said second unique enzyme cleavage sites; megacircles; recombinase protein or coding sequence thereof).

30.   The genomic DNA fragment of any one of claims 1-11, or the method of any one of claims 12-29, wherein the positive/negative selection marker cassette comprises a (neomycin/puromycin/hygromycin/blasticidin/zeocin) resistant / TK (or HPRT) gene under the expression control a eukaryotic promoter (such as the PGK promoter) and a polyA coding sequence.

31.   A mouse generated by any one of the method of claims 26-29.

32.   A mouse comprising in its genome an exogenous genomic DNA from a donor (*e.g.*, a human, a mammal, or different mouse strain), wherein the exogenous genomic DNA comprises a polymorphism within the species of the donor, and wherein the exogenous genomic DNA comprises the DNA fragment of any one of claims 1-11.

33.   The mouse of claim 32, wherein the exogenous genomic DNA comprises an immunoglobulin heavy chain gene cluster, immunoglobulin kappa light chain gene cluster, immunoglobulin lambda chain gene cluster, a TCR, an MHC class 1 gene, an MHC class 2 gene, or a cytochrome P450 gene.

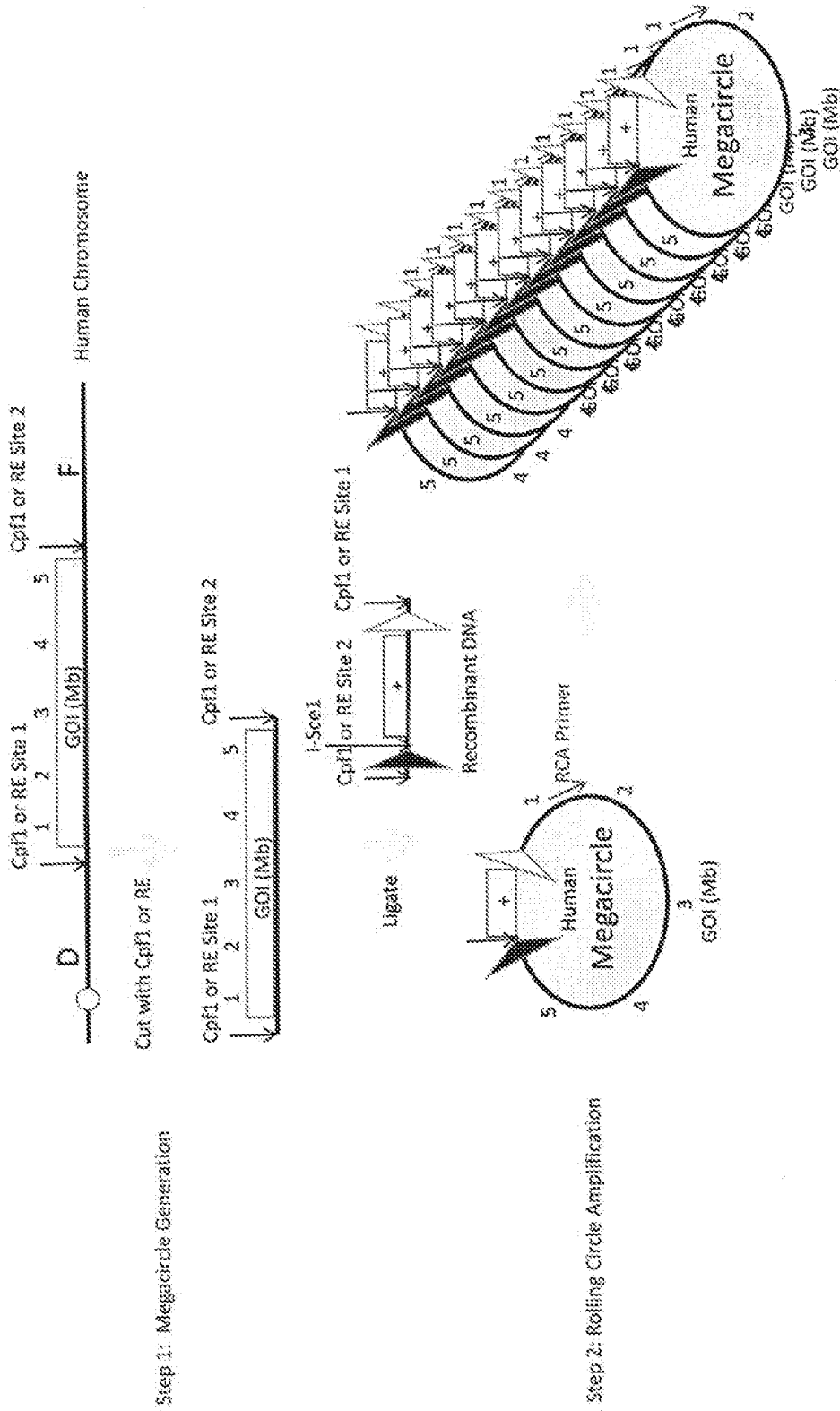Figure 1 Megacircle Generation for RMCE. Strategy 1
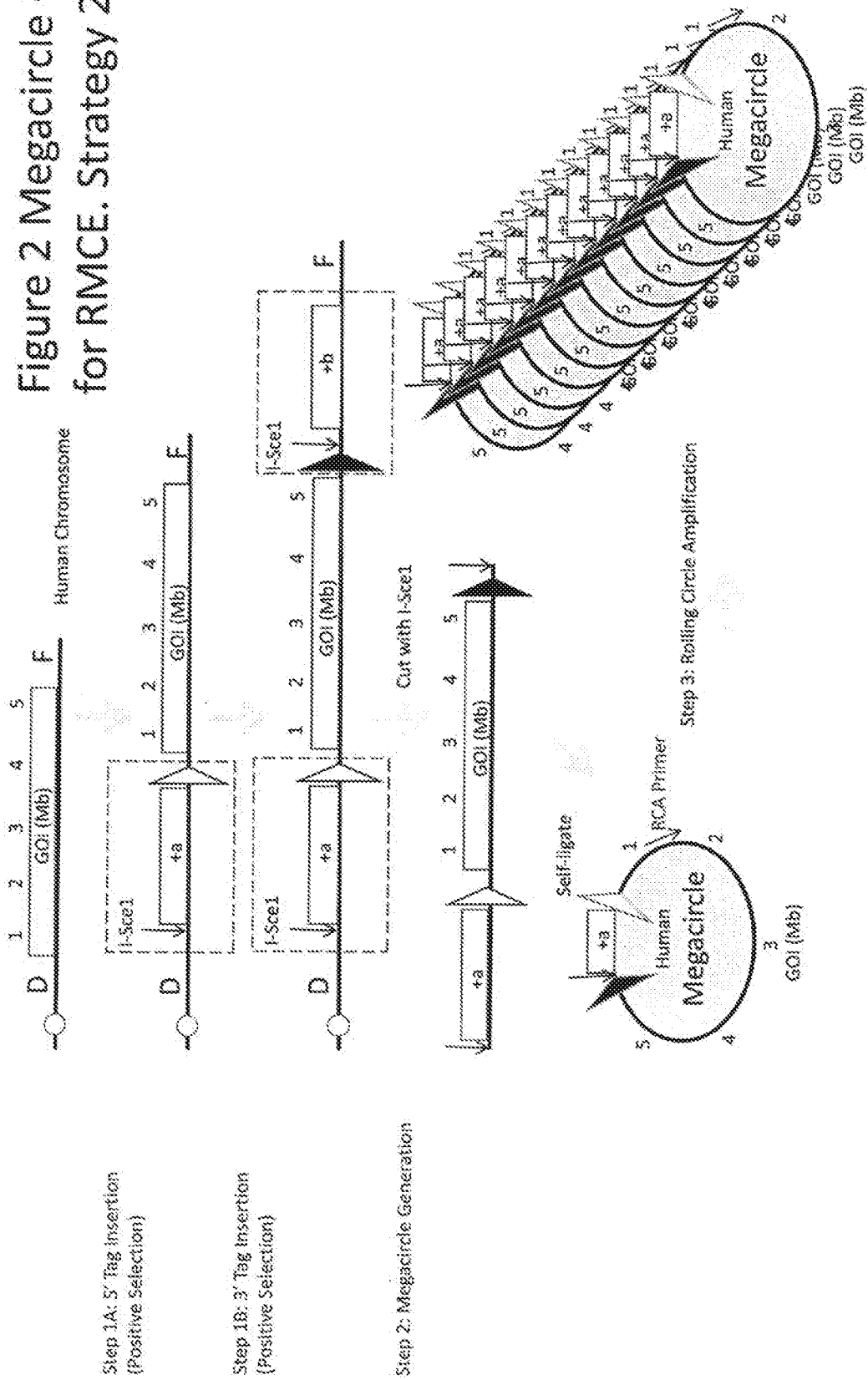
Figure 2 Megacircle Generation for RMCE. Strategy 2
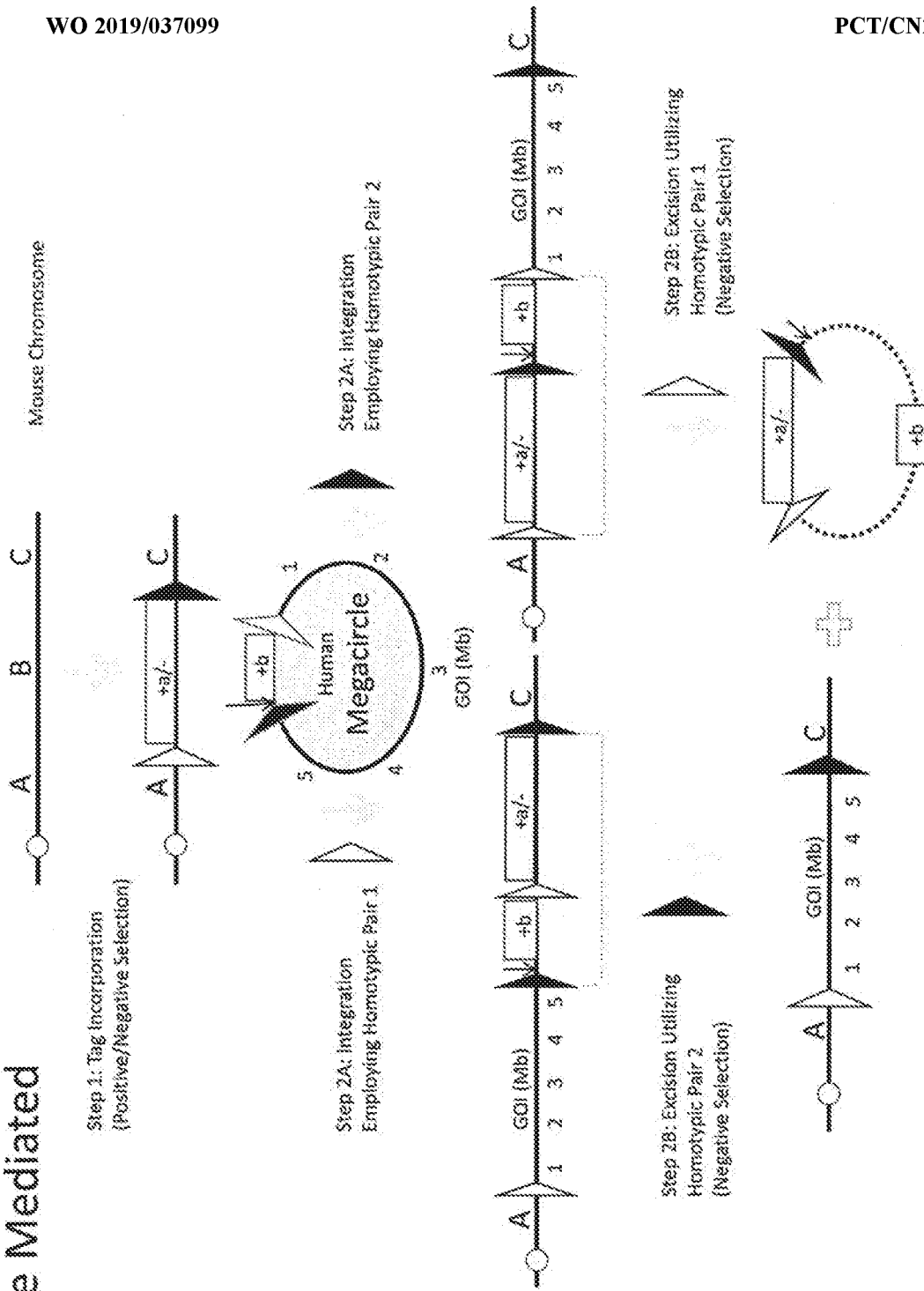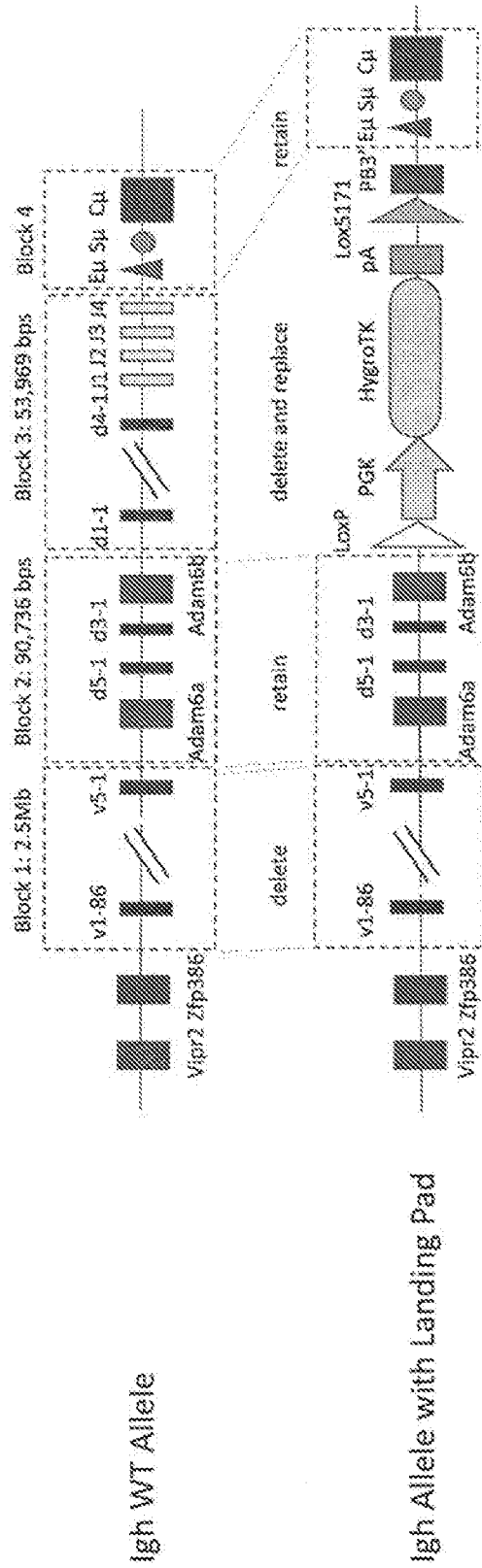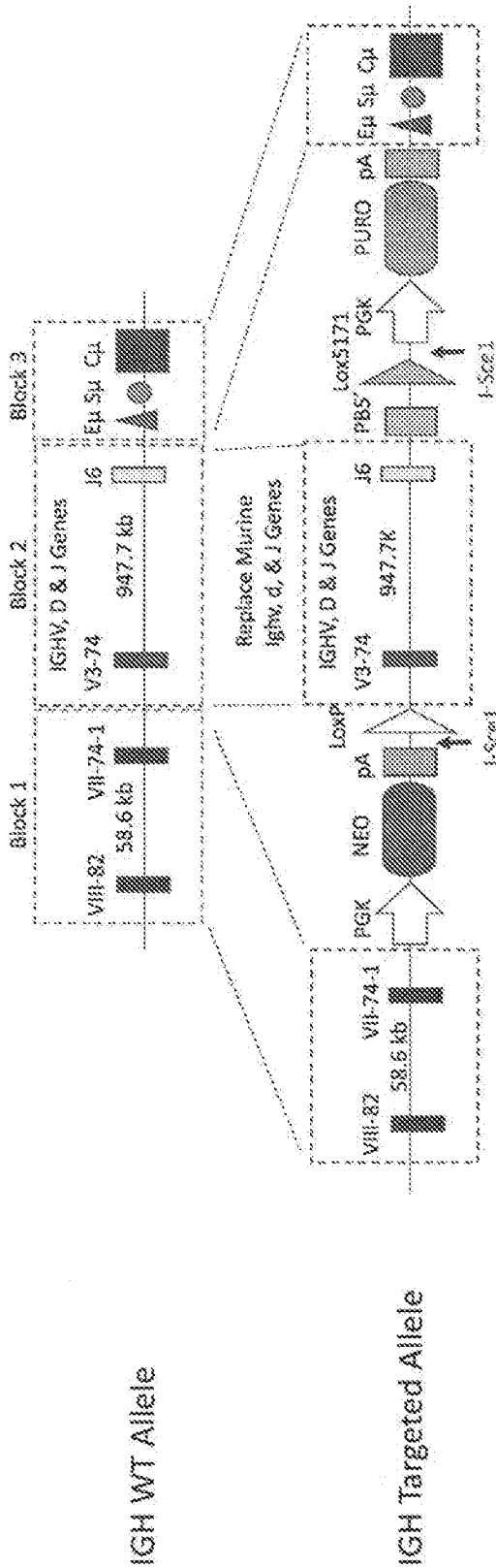
Figure 3 Recombinase Mediated Cassette Exchange

Figure 4A Mice Carrying a Complete Repertoire of the Human IGHV, D & J Genes: Outcome of the Murine Ighv, d, j & c Genes

We could divide the murine Igh locus into 4 blocks, based on outcome of humanization:

❖ Block 1: Genes from v1-86 to v5-1 should be deleted with CRISPR-Cas (2.5 Mb)
  o The most 5′ ighv is v1-86, which is a pseudogene
  o The most 3′ ighv is 5-1, which is also a pseudogene
❖ Block 2: Adam6a/d5-1/d3-1/Adam6b should be retained (90,736 bps)
  o From 5′ to 3′, the order of the 4 genes is Adam6a/d5-1 (a pseudogene)/d3-1 (a functional gene)/Adam6b
  o Both Adam6a and Adam6b are in reverse orientation, as compared with the v, d, j & c genes
  o We will have 7,374 bps between Adam6b and ighd1-1, which carry the promoter for Adam6a
  o We will have 26,516 bps between ighv5-1 and the 3′ end of Adam6a, which is the 3′ flanking region for Adam6a
❖ Block 3: Genes from d1-1 to j4 will be deleted (53,969 bps) and replaced with the landing pad (3,159 bps)
  o This includes all d genes from d1-1 to d4-1 and all j genes (j1 to j4)
❖ Block 4: From Eμ (as defined by the Xba1 fragment) and the rest of the Igh 3′ constant region genes
  o Also Ighj4 (the last J gene) is separated from ighm (the first constant gene) by 5,783 bps. Within this region lies the enhance and switch sequences.
  o This region will not be touched and will be kept in situ

Figure 4B Mice Carrying a Complete Repertoire of the Human IGHV, D & J Genes: the Human IGHV, D & J Genes Included for Humanization

❖ Block 1: from V(III)-82 to V(II)-74-1 (58.6 kb)
  o This region carries 9 IGHV pseudogenes, one IGHV gene with ORF (7-81) and one not sure (7-77)
  o We will not include this region in the humanization effort
❖ Block 2: from V3-74 to J6 (947.7 kb)
  o This region carries the rest of the IGHV, all D, & all J genes, which will be incorporated into the mouse Igh locus
  o V3-74 is separated from V(II)-74-1 by 10,346 bps, which should include the V3-74 promoter. The Yang Lab design includes 136 bps of the sequence 5' from V3-74, which should be sufficient, as the first 100 bps carry necessary regulatory elements. Notice the Kymab design takes 49,314 bps sequence 5' to V3-74, which includes 6 additional IGHV genes (5-78, 7-77, (II)-76-1,3-36, V3-75, and (II)-74-1), with all being pseudogenes, with the exception of 7-77 (not sure);
  o J6 is separated from IGHM by 6,776 bps. The Yang Lab design takes 202 bps 3' flanking sequence of J6, while Kymab takes 455 bps of 3' flanking sequence of J6.
❖ Block 3: from the intergenic region of IGHJ6 and including all other 3' constant region genes
  o We will not include this region in the humanization effort
❖ I-SceI Sites incorporated to facilitate megacircle generation by Rolling Cycle Amplification.
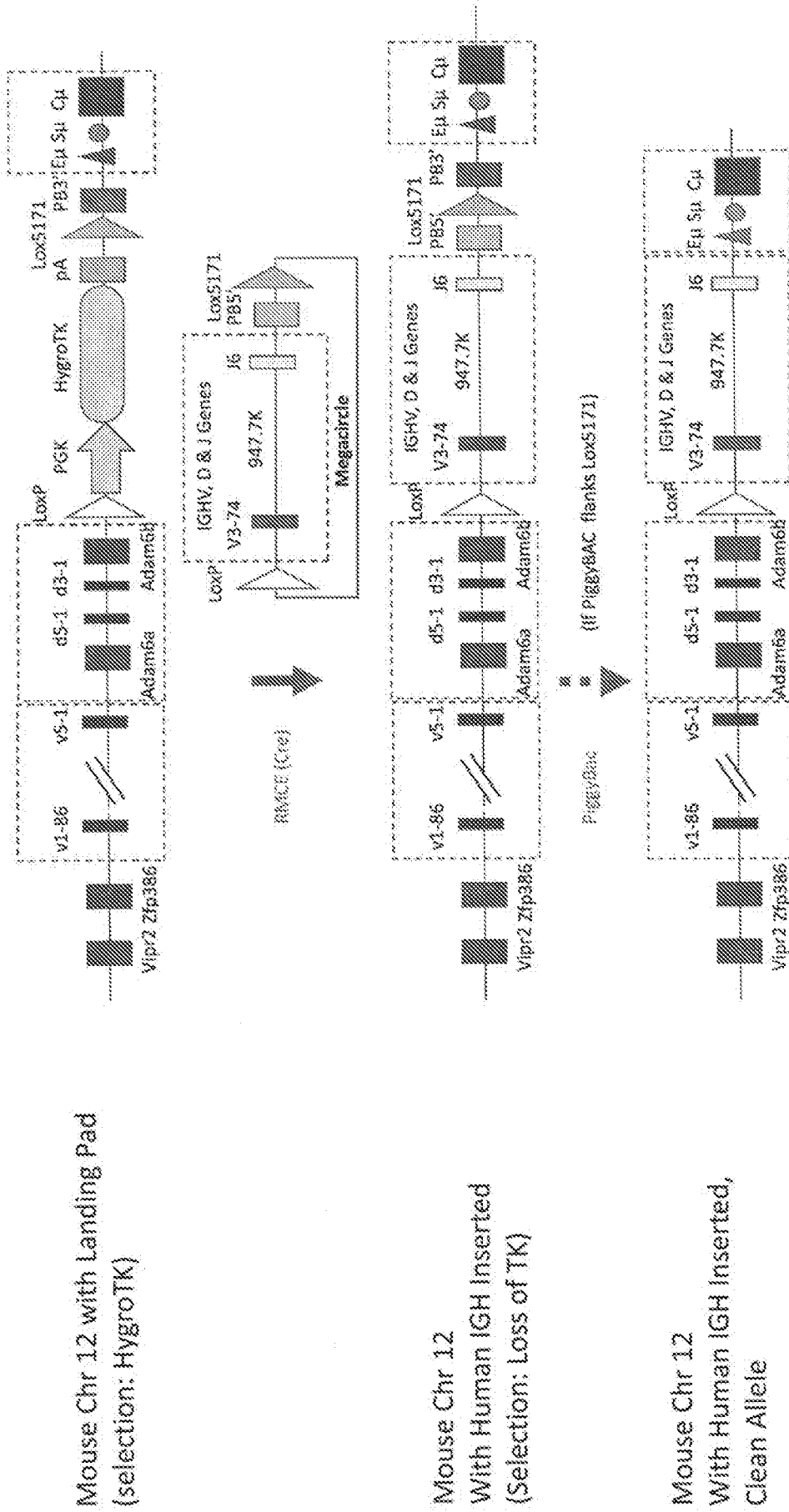
Figure 4C Overall Flowscheme (Cre/loxP/lox5171)

Mouse Chr 12 with Landing Pad
(selection: HygroTK)

Mouse Chr 12
With Human IGH Inserted
(Selection: Loss of TK)

Mouse Chr 12
With Human IGH Inserted,
Clean Allele

Figure 5 Megacircle Generation for Recombinase Mediated Integration. Strategy 1

Step 1: Megacircle Generation

Cpf1 or RE Site 1   Cpf1 or RE Site 2

D   1   2   3   GOI (Mb)   4   5   E   Human Chromosome

Cut with Cpf1 or RE

Cpf1 or RE Site 1   Cpf1 or RE Site 2

1   2   3   GOI (Mb)   4   5

Cpf1 or RE Site 1   I-Sce1

1   2   3   GOI (Mb)   4   5   Cpf1 or RE Site 2   Cpf1 or RE Site 1

+   Recombinant DNA

Ligate

Human Megacircle   1   2   3   GOI (Mb)   4   5   +

Step 2: Rolling Circle Amplification

RCA Primer

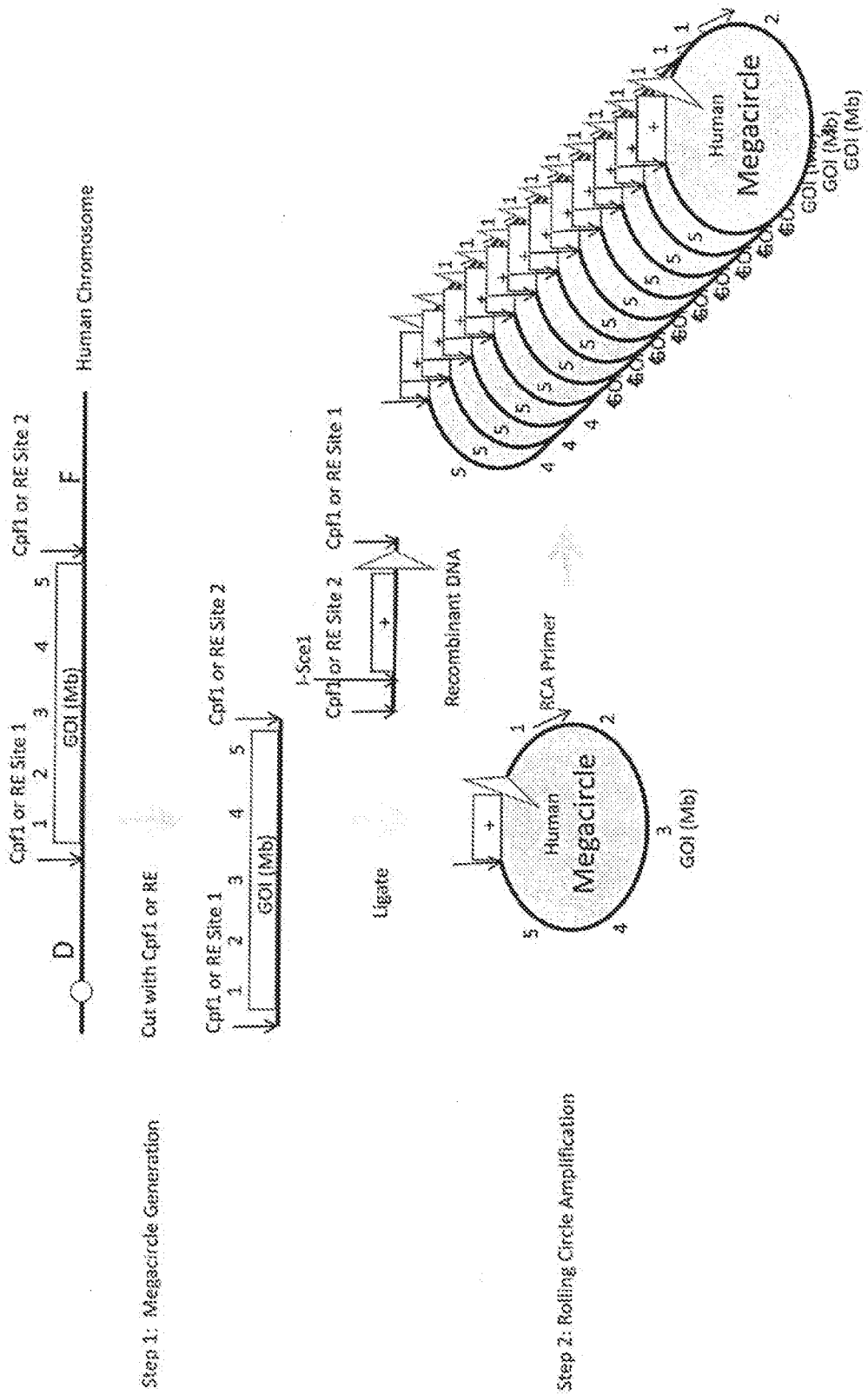Human Megacircle   GOI (Mb)   GOI (Mb)   GOI (Mb)

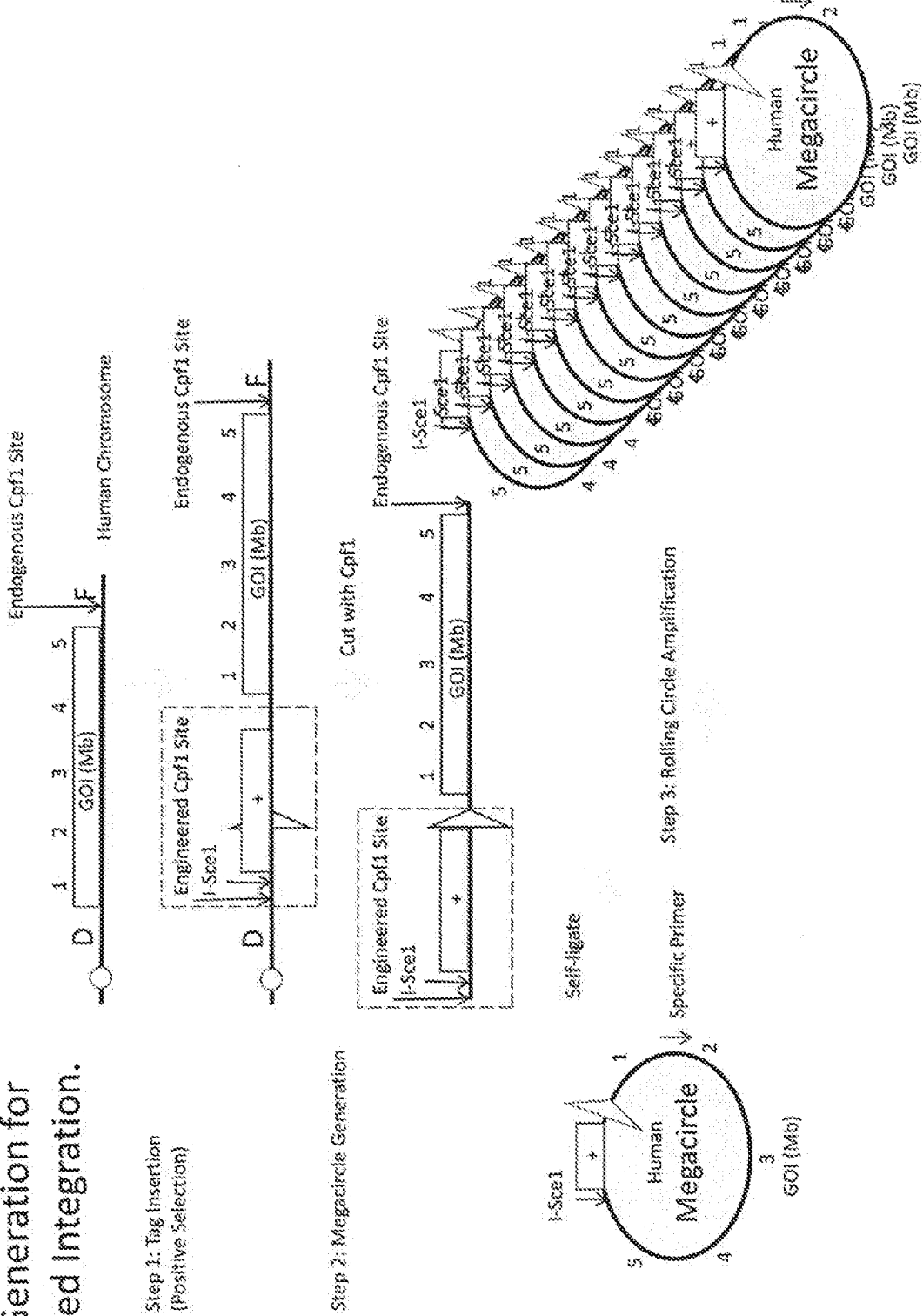Figure 6 Megacircle Generation for Recombinase Mediated Integration. Strategy 2
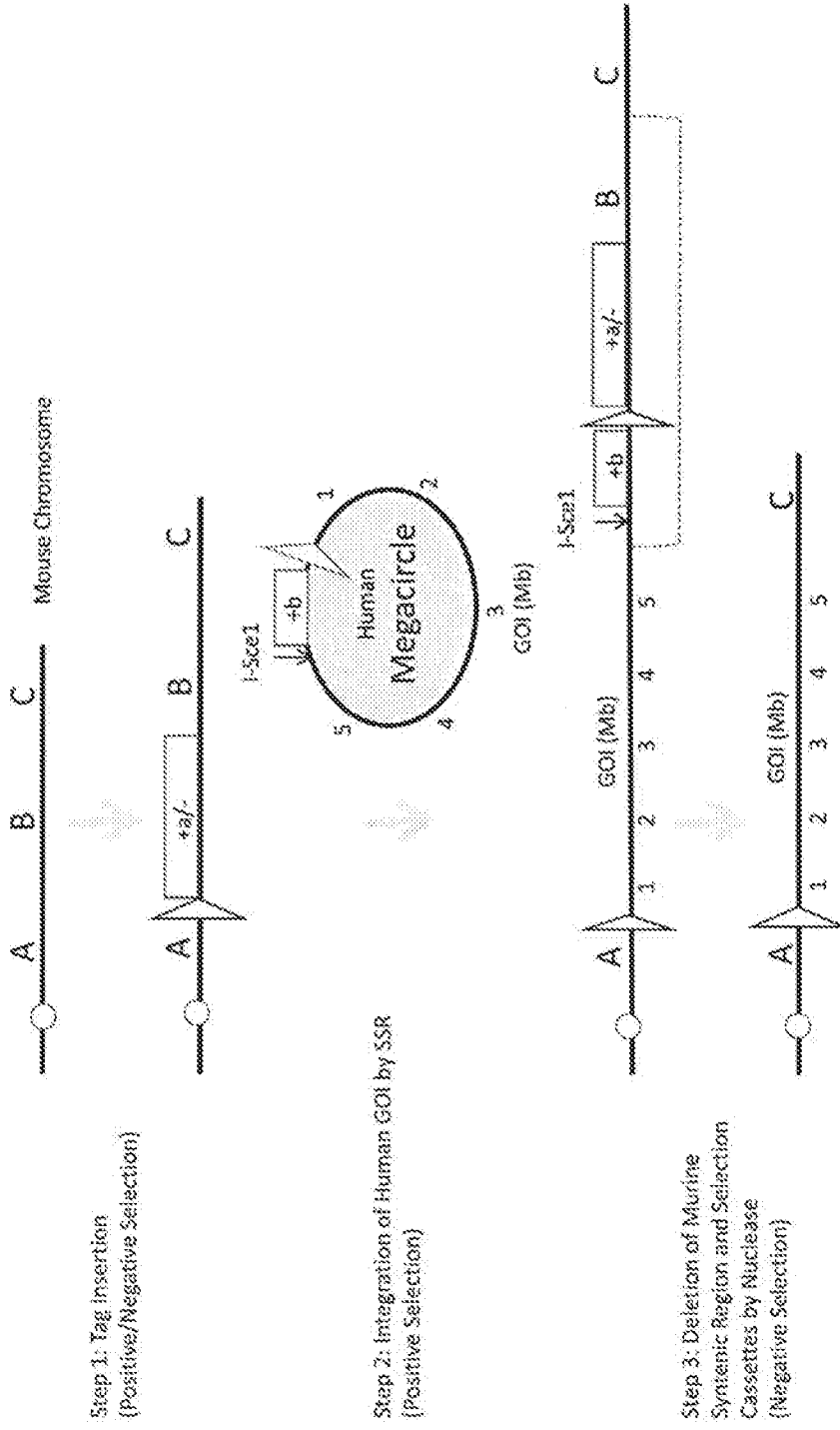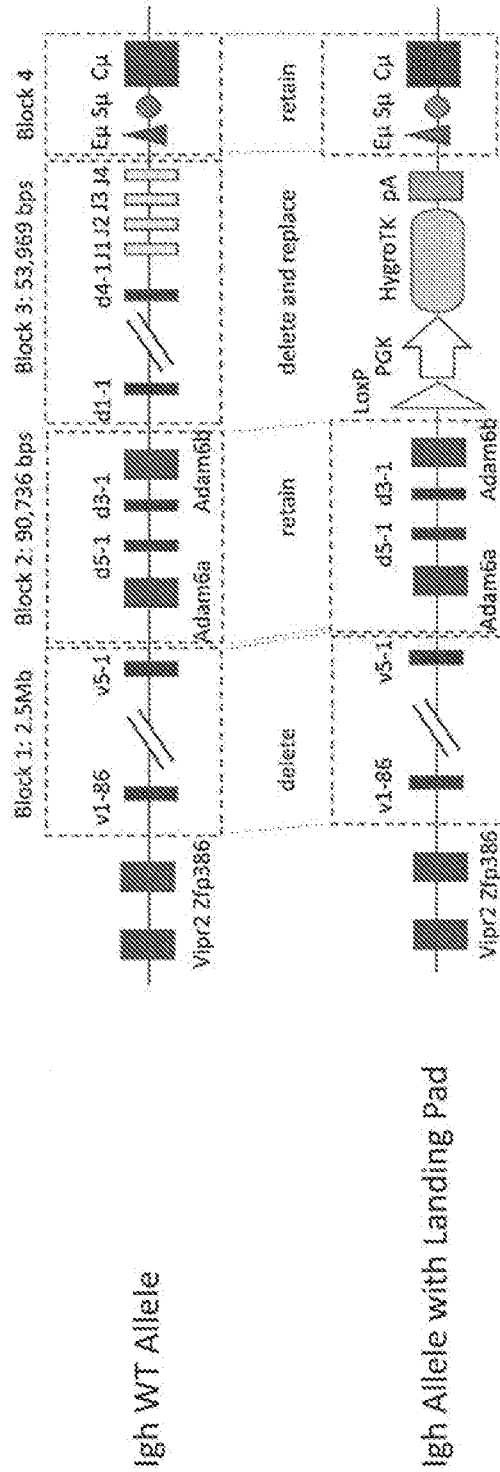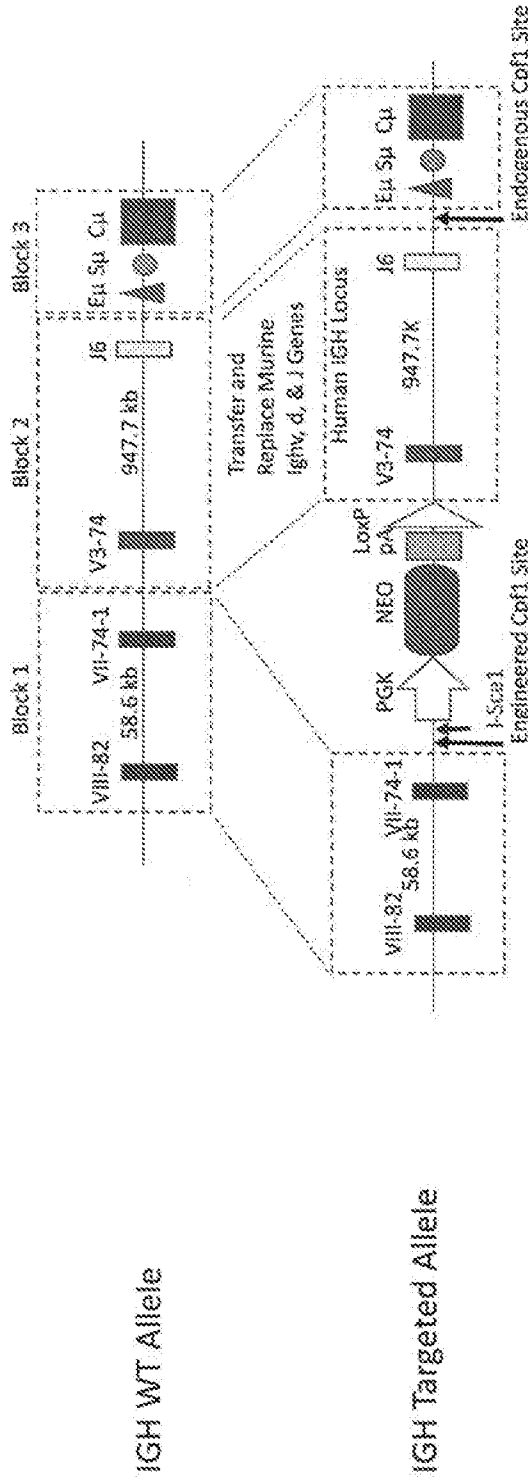
Figure 7 Recombinase Mediated Integration

Figure 8A Mice Carrying a Complete Repertoire of the Human IGHV, D & J Genes: Outcome of the Murine Ighv, d, j & c Genes

Figure 8B Mice Carrying a Complete Repertoire of the Human IGHV, D & J Genes Included for Humanization: the Human IGHV, D & J Genes Included for Humanization

❖ Block 1: from V(III)-82 to V(II)-74-1 (58.6 kb)
  ○ This region carries 9 IGHV pseudogenes, one IGHV gene with ORF (7-81) and one not sure (7-77)
  ○ We will not include this region in the humanization effort

❖ Block 2: from V3-74 to J6 (947.7 kb)
  ○ This region carries the rest of the IGHV, all D, & all J genes, which will be incorporated into the mouse Igh locus
  ○ V3-74 is separated from V(II)-74-1 by 10,346 bps, which should include the V3-74 promoter. The Yang Lab design includes 136 bps of the sequence 5' from V3-74, which should be sufficient, as the first 100 bps carry necessary regulatory elements. Notice the Kymab design takes 49,314 bps sequence 5' to V3-74, which includes 6 additional IGHV genes (5-78, 7-77, III-76-1,3-36, V3-75, and (II)-74-1), with all being pseudogenes, with the exception of 7-77 (not sure).
  ○ J6 is separated from IGHM by 6,776 bps. The Yang Lab design takes 202 bps 3' flanking sequence of J6, while Kymab takes 455 bps of 3' flanking sequence of J6.

❖ Block 3: from the intergenic region of IGHJ6 and including all other 3' constant region genes
  ○ We will not include this region in the humanization effort

❖ Engineered nuclease protospacer/PAM and J-sce1 Sites incorporated to facilitate megacircle generation by Rolling Cycle Amplification.

Figure 8C Overall Flow Scheme (Cre/loxP)
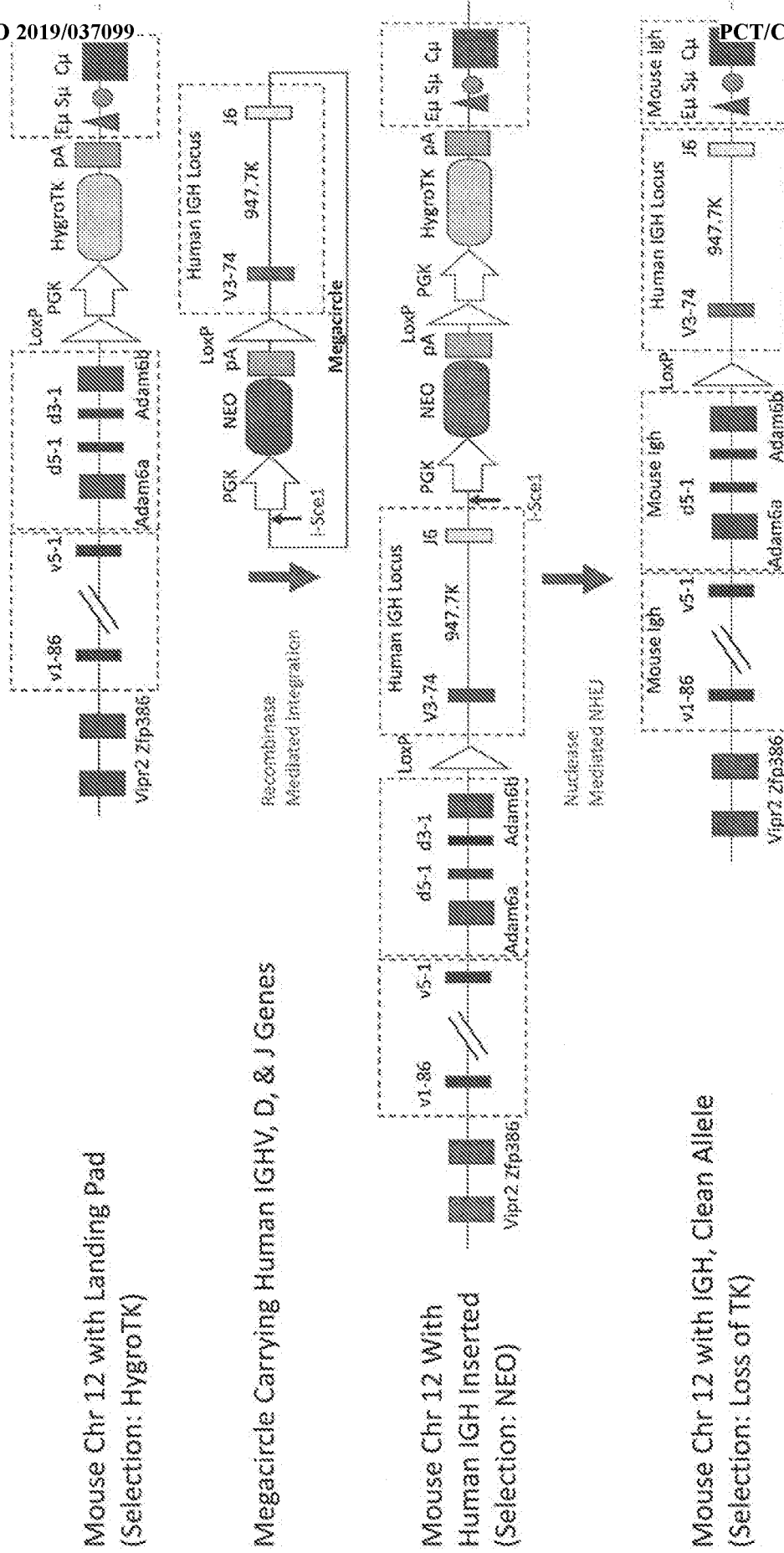
Figure 9 Overall Activity

Experiment 1: Test Igh KO Strategy (2.5 Mb)

Mouse Zygotes: Mouse Igh WT
↓ Inject CRISPR RNP with 4 Guides
Live Mice: Mouse Igh with v1.86 to v5-1 deleted (2.5 Mb)
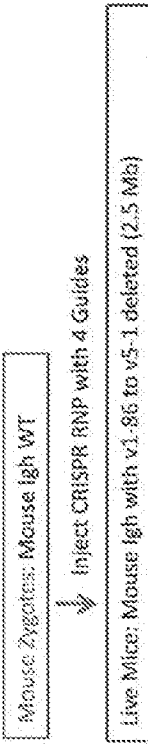
Experiment 2: Insert LoxP into IGH Locus

Human ESCs: Human IGH WT
↓ Electroporate CRISPR RNP & Donor
Human ESCs: Human IGHV Genes with LoxP Inserted
↓ Prepare Megacircle
Megacircle: Human IGHV Gene with LoxP Inserted
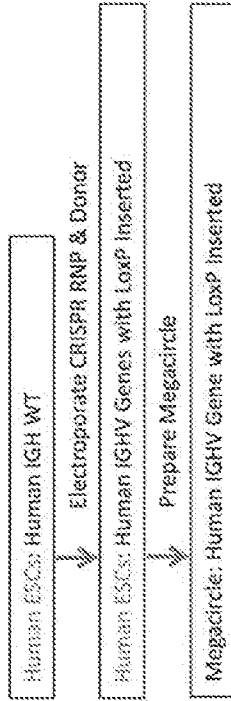
Experiment 3: Test All Elements

Mouse ESCs: Mouse Igh WT
↓ Electroporate CRISPR RNP & Donor
Mouse ESCs: Mouse Igh with Landing Pad
↓ Electroporate RMCE Minicircle & Cre Protein
Mouse ESCs: Mouse Igh with Successful RMCE
↓ Electroporate CRISPR RNP
Mouse ESCs: Mouse Igh Allele Cleaned up
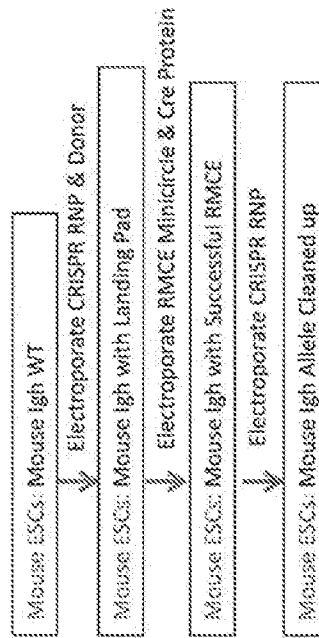
Experiment 4: Humanize Igh

Mouse ESCs: Murine Igh WT
↓ Electroporate CRISPR RNP & Landing Pad Donor
Mouse ESCs: Murine Igh with Landing Pad
↓ Electroporate Megacircle/Cre (RMCE)
Mouse ESCs: Murine Igh Locus Carrying Human IGHV, D, & J Genes, Targeted Allele
↓ Electroporate CRISPR RNP for Igh KO
Mouse ESCs: Murine Igh KO & Human IGHV, D, & J Gene KI, Targeted Allele
↓ Electroporate CRISPR RNP for Selection Cassette Removal
Mouse ESCs: Murine Igh KO & Human IGHV, D, & J Gene KI, Clean Allele
↓ Blastocyst Injection
Live Mice: Murine Igh KO & Human IGHV, D, & J Gene KI
↓ Breed to Kappa Light Chain Humanized Mice
Live Mice: HK Mice (Carrying Heavy and Kappa Light Chain Humanized Alleles)
↓ Breed to Lambda Light Chain Humanized Mice
Live Mice: HKL Mice (Carrying Heavy, Kappa, & Lambda Light Chain Humanized Alleles)

A.   CLASSIFICATION OF SUBJECT MATTER

C12N 15/09(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B.   FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, VEN, CNTXT, USTXT, WOTXT, CNKI, PubMed, Google: GOI, genomic region of interest, HDR, RMCE, recombinase mediated cassette exchange, IGHV, Megacircle Generation, Recombinase Mediated Integration, RTS, recombinase target site, genomic DNA, rolling circle amplification, RCA, immunoglobulin, TCR, MHC, P450, I-Sce1, CRISPR, CRISPR/Cpfl, Cpfl, ZFN, ARE, PfAgo, Cas9, megacircle, Cosmid, loxP, attP, attB

C.   DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 104334732 A (KYMAB LIMITED) 04 February 2015 (2015-02-04) abstract，figures 1-69	1-33
A	WO 2013041844 A2 (KYMAB LIMITED) 28 March 2013 (2013-03-28) the whole document	1-33
A	WO 0200875 A2 (PROTEMATION，INC.) 03 January 2002 (2002-01-03) the whole document	1-33
A	WO 02088353 A2 (ASSOCIATION POUR LE DEVELOPPEMENT DE LA RECHERCHE EN GENETIQUE MOLECULAIRE ADEREGEM) 07 November 2002 (2002-11-07) the whole document	1-33
A	WO 2004044150 A2 (THE UNITED STATES OF AMERICA，AS REPRESENTED BY THE SECRETARY OF AGRICULTURE) 27 May 2004 (2004-05-27) the whole document	1-33

☑ Further documents are listed in the continuation of Box C.   ☑ See patent family annex.

*	Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	
"E"	earlier application or patent but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P"	document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search	Date of mailing of the international search report
05 April 2018	**04 May 2018**

Name and mailing address of the ISA/CN	Authorized officer
STATE INTELLECTUAL PROPERTY OFFICE OF THE P.R.CHINA **6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing 100088** **China**	**YAO,Jinxiao**
Facsimile No. **(86-10)62019451**	Telephone No. **86-(10)-53961955**

Form PCT/ISA/210 (second sheet) (January 2015)

C.   DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2008101216 A2 (THE GOVT. OF THE USA AS REPRESENTED BY THE SECRETARY OF THE DEPT OF HEALTH AND HUMAN SERVICES) 21 August 2008 (2008-08-21) the whole document	1-33
A	WO 2015166272 A2 (IONTAS LIMITED) 05 November 2015 (2015-11-05) the whole document	1-33

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	104334732	A	04 February 2015	EP	2831244	A2	04 February 2015
				WO	2013061098	A2	02 May 2013
				SG	11201405058W	A	26 September 2014
				WO	2013061098	A3	20 June 2013
				AU	2013239501	A1	18 September 2014
				NZ	629202	A	26 August 2016
				EP	2831244	B1	31 January 2018
				EP	2785845	A2	08 October 2014
				HK	1200872	A1	14 August 2015
				JP	2015512634	A	30 April 2015
				US	2017096498	A1	06 April 2017
				IN	1879MUN2014	A	03 July 2015
				WO	2013144566	A2	03 October 2013
				CA	2867530	A1	03 October 2013
				WO	2013144566	A3	21 November 2013
				US	2013263292	A1	03 October 2013
				US	2016219846	A1	04 August 2016
				US	9253965	B2	09 February 2016
				US	2017101483	A1	13 April 2017
WO	2013041844	A2	28 March 2013	WO	2013041844	A9	27 November 2014
				DK	2757875	T3	08 January 2018
				AU	2017254949	A1	23 November 2017
				JP	2017079795	A	18 May 2017
				AU	2012311288	A1	01 May 2014
				AU	2012311286	A1	01 May 2014
				NZ	623756	A	24 June 2016
				CN	103945689	B	19 October 2016
				ES	2612935	T3	19 May 2017
				US	2014212416	A1	31 July 2014
				HK	1194757	A1	18 August 2017
				EP	2757875	A2	30 July 2014
				WO	2013041846	A3	01 August 2013
				EP	2758535	A2	30 July 2014
				CN	103945689	A	23 July 2014
				CN	104080916	A	01 October 2014
				CN	106995822	A	01 August 2017
				EP	3128009	A1	08 February 2017
				BR	112014006394	A2	28 March 2017
				AU	2012311288	B2	17 August 2017
				WO	2013041846	A2	28 March 2013
				WO	2013041844	A3	11 July 2013
				US	2014041067	A1	06 February 2014
				CA	2846319	A1	28 March 2013
				CN	107056936	A	18 August 2017
				CA	2846322	A1	28 March 2013
				NZ	623745	A	24 June 2016
				JP	2014533930	A	18 December 2014
				JP	5813880	B2	17 November 2015
				JP	2016002084	A	12 January 2016
				BR	112014006390	A2	28 March 2017

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
				US	2018030121	A1	01 February 2018
				JP	2014529998	A	17 November 2014
				EP	2757875	B1	29 November 2017
				EP	2758535	B1	09 November 2016
WO	0200875	A2	03 January 2002	US	2004005591	A1	08 January 2004
				US	6551828	B1	22 April 2003
				AU	6862701	A	08 January 2002
				US	2006019291	A1	26 January 2006
				US	6953689	B2	11 October 2005
				WO	0200875	A3	02 May 2002
WO	02088353	A2	07 November 2002	DE	60208014	D1	19 January 2006
				EP	1383891	A2	28 January 2004
				DE	60208014	T2	17 August 2006
				WO	02088353	A3	20 March 2003
				US	2004244071	A1	02 December 2004
				US	2003159160	A1	21 August 2003
				AT	312912	T	15 December 2005
				AU	2002304258	A1	11 November 2002
				US	7074611	B2	11 July 2006
				EP	1383891	B1	14 December 2005
WO	2004044150	A2	27 May 2004	WO	2004044150	A3	08 July 2004
				AU	2003290660	A1	03 June 2004
				DE	10251918	A1	19 May 2004
				AU	2003290660	A8	03 June 2004
WO	2008101216	A2	21 August 2008	US	2010022006	A1	28 January 2010
				US	9018011	B2	28 April 2015
				WO	2008101216	A3	31 December 2008
WO	2015166272	A2	05 November 2015	EP	3137602	A2	08 March 2017
				CN	106574264	A	19 April 2017
				CA	2947513	A1	05 November 2015
				WO	2015166272	A3	30 December 2015
				JP	2017518362	A	06 July 2017
				SG	11201609123R	A	29 November 2016
				GB	201407852	D0	18 June 2014
				US	2017073664	A1	16 March 2017
				EA	201691990	A1	31 May 2017
				AU	2015255013	A1	17 November 2016