Development of a targeted gene integration procedure for the production of biopharmaceutical proteins

汪, 雪

https://doi.org/10.15017/1866307

出版情報：九州大学, 2017, 博士（工学）, 課程博士
バージョン：
権利関係：

Development of a targeted gene integration procedure for the production of biopharmaceutical proteins

Xue Wang

July 2017

TABLE OF CONTENTS

Chapter 1

Introduction

1.1 A brief history, status quo, breakthroughs and setbacks

The biopharmaceuticals, also known as biological medical products or biologic(al)s, refer to pharmaceutical products generated in, extracted from, or semi-synthesized from biological resources [1]. Biopharmaceutical products include vaccines, blood or blood components, allergenic extracts, recombinant therapeutic protein, transplantable tissues or organs, and even nucleic acid-based genetic material, somatic cells or other cells used in cell therapy [1-3]. Unlike chemical synthesized drugs, biopharmaceuticals and their components or precursors are isolated from living sources like mammalian, poultry, plant, or microbial and so on. Culturing of the livings and purifying the end products from the low-yield raw products are difficult and high-cost, that is why biological medical products are often much more expensive than chemical synthesized medicine. At the same time, however, the demand of biopharmaceuticals has been increasing since the first FDA-approved biosynthetic human insulin benefited numerous patients from 1982 [4], and it cannot be replaced by conventional medicine due to the characteristic of better curative effect. It has the curative effect on terminal illnesses that are not cured by conventional medicines. For example, nucleic acid vaccines at the forefront of biomedical research may be applied to treatment for incurable diseases in the future [5].

As a major classification of biologics, recombinant mammalian protein is central of structural studies for new peptide-based drug development in biomedical fields and is also important to commercial biopharmaceuticals, particularly to the manufacture of antibodies.

Recombinant protein drugs include three types: (i) Substances (quite) identical to signal proteins formed in human cells, such as stimulating hormone represented by the first developed biosynthetic human insulin. (ii) Monoclonal antibodies (mAbs) using hybridoma technology or other methods, the first mAb product was invented by César Milstein and Georges J. F. Köhler in 1975 [6], since then, myriads of mAbs are used to fight off bacteria and viruses as they can recognize and target to specific cells or proteins. The mAb works similarly to the antibodies used in the human immune system, and they could be custom-designed as commercial individual products. (iii) Fusion proteins (receptor constructs) based on natural interactions between immunoglobulin frames and receptors. The immunoglobulin structure is stable and possesses useful features in terms of pharmacology, whereas the receptor dominates the detailed specificity. Recombinant mammalian protein added major better therapeutic options for the treatment of many diseases, including some for which previously therapies were inadequate. However, the advent of recombinant mammalian protein has also raised some complex issues, the cost is dramatically higher than conventional medications. Thus it is urgent to develop the manufacturing process of biopharmaceuticals while decreasing its cost.

Production yield and quality of recombinant proteins are critical parameters for manufacturing values, to meet the industrial requirement, the obtained products should have (i) a high titer that determined by the cell-specific production rate and the overall process duration, it is related to the generation of a good clone and optimization of culture conditions. (ii) An adequate efficacy for treatment, that is decided by the process of glycosylation in host cells, and always the matter of choosing a suitable host cell line for the target products and/or using human derived coding genes. From the generation of the producer clones and ending with the harvest of the target products, each producing process affects the final income, many methods developed for improving the manipulation processes were reported (shown in Fig. 1-

1). In this chapter, methods developed for benefiting industrial production of recombinant proteins are reviewed from different aspects.

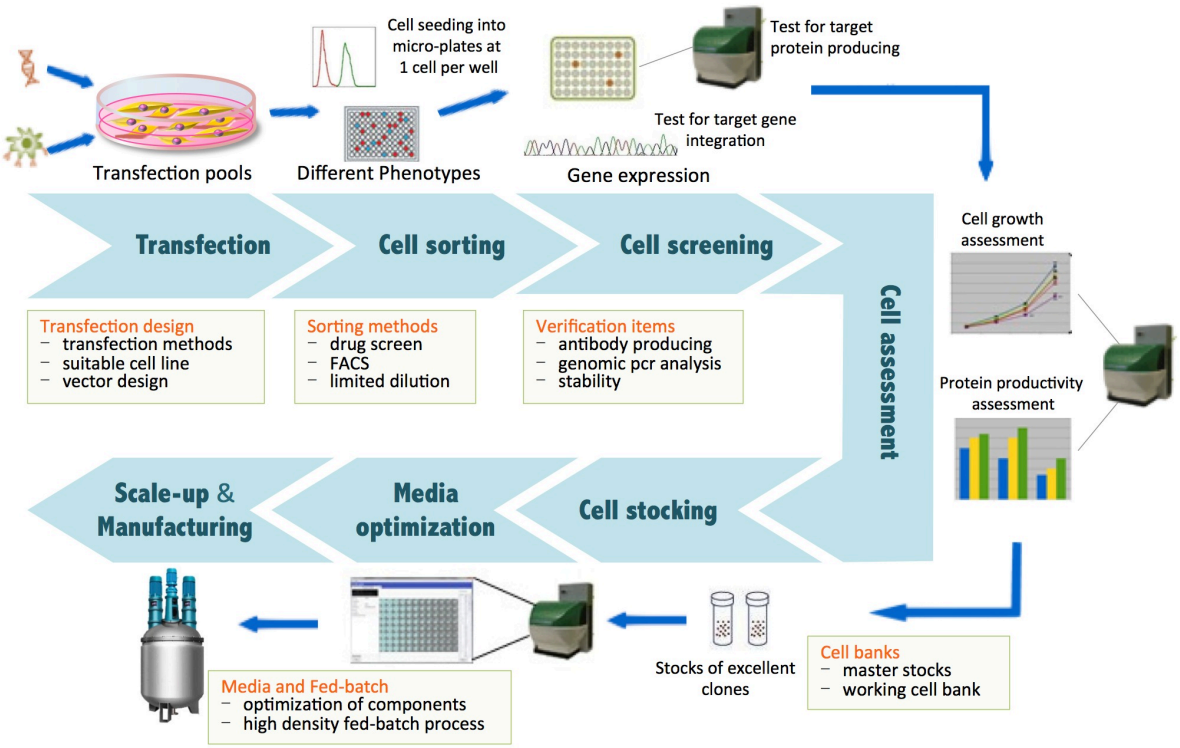


Fig. 1-1 Schematic diagram of recombinant protein production processes.

First of all, to choose an adequate cell line for a certain recombinant product is very important. Mammalian cells are the dominant production system for recombinant protein expression as the majority of recombinant proteins can not be obtained from other types of cells because most recombinant proteins encoded by human-derived gene only can be expressed and post-translational modified in mammalian cells with close genetic relationships. Nowadays, the majority of FDA-approved biopharmaceuticals are produced in mammalian cells [7]. The Chinese hamster ovary (CHO) cells were used as an expression system for the first FDA-approved recombinant protein drugs (t-PA; tissue plasminogen activator) in 1986. Since then, a variety of flexible cell expression system have emerged, including Sp2/0, NS0, HEK293, PER.C6 and so on, but approved recombinant drugs used for

human therapy are only produced in CHO, NS0, and Sp2/0 [8]. The CHO cells shared 70% market [9] are the preferred option for recombinant protein expression, and the highest product titers with ~1 g/L in batch and 1~10 g/L in fed-batch processes could be obtained. More importantly, CHO cells do not add non-human origin glycosylation modification on expressed recombinant proteins. NS0 and Sp2/0 cells take a few proportion of the market, due to their nonhuman-like glycosylation, they are only used to produce α-gal (galactose-α, 3-Gal) and Neu5Gc (N-glycolyl-neuraminic acid).

To meet the commercial application demands, various methods have been developed to increase the titer of recombinant proteins. They are typically central around two methods, one is increasing the cell producing ability in genomic modification scope, another is optimization of culture conditions for improving the yield and decreasing the cost. The mainstream gene amplification method is using an inhibitor to stimulate the increase of transgene copy number in transfected cells, such as well-known DHFR/MTX system works on CHO DG44 cell line [10]. Despite of fundamental gene amplification method, modifying the gene expression cassettes is another way to approach high expression level of a target gene. For example, an EF1α (elongation factor 1α gene) promoter able to accelerate recombinant proteins expression in mammalian cells was reported [11], HNRPA2B1/CBX3 locus derived CpG-island fragments were proved to enhance expression and reduce silencing of transgenes in mammalian cells [12]. Thus, any reported advantages in genomic modification scope could be flexibly applied to approach a desirable productivity. Glycosylation, which determines the correct function of most therapeutic proteins, could be predicted via primary sequences around the glycosylation site [13]. The therapeutic effect of recombinant proteins could also be improved through genetically modified intracellular metabolism since the whole-genome sequence of CHO-K1 has been published in 2011 [14], and the first available community-curated genome-scale metabolic reconstructions of CHO

was created by several research groups and became available from 2016 [15]. This model annotated 1766 genes and 4455 metabolites acting on 6663 reactions, and it is expected to contribute to application and understanding of CHO cells.

The optimization of culture conditions mainly focuses on the optimization of medium components, batch conditions, temperature and so on. Commonly, "Bottom-Up" and "Top-Down" strategies are used to adjust the medium components [16]. "Bottom-Up" strategy carefully inspects all effect elements one by one, while "Top-Down" investigates on a group of elements after divided components into subgroups. For instance, medium consisted of defined chemical components could be successfully developed by "Top-Down" strategy for culturing CHO cells or NS0 cells [17]. Besides, a particular milestone is widely adopted serum-free medium systems, which is possible to both reduce the cost and improve the product safety via avoiding the contamination of exogenous viruses. Additives, including insulin, transferrin and putrescine, used to improve serum-free medium to increase the target protein titer and/or target protein expression level were also investigated and developed. It is even suggested that the addition of a little lactate at the early culture stage favors the metabolic efficiency [18], and the relationship between protein glycosylation and high ammonia concentration was reported [19].

Based on all of above methods, the productivity of mammalian cell derived recombinant proteins in a fed-batch progress was increased from 0.05 g/l in 1986 [7] to 6 g/l (20~50 pg/cell/day) in 2015 [20]. To facilitate cell line development in the manufacture of mammalian recombinant protein therapeutics, an accumulative site-specific gene integration system (AGIS) using Cre recombinase and mutated *loxPs* was developed in our laboratory previously. This system enables repeated integration of transgenes into a predetermined cell genome, supposed to accelerate the cell line development on biopharmaceutical producing process.

1.2 Research purpose

A new accumulative site-specific gene integration system (AGIS) with some improvements was described in this thesis, which speeded up the establishment process of producer CHO cells. In newly designed AGIS, two donor plasmids were introduced into cells in a single transfection and recombinant cells were screened using a FACS device. A mutated *loxP* target site and a DsRed reporter gene were integrated into a genomic locus of hypoxanthine phosphoribosyltransferase in CHO cells by homologous recombination (CHO/R1). Donor plasmids containing a scFv-Fc expression cassette and fluorescent protein genes flanked by the compatible target sites were co-transfected into CHO/R1 cells with a Cre expression vector. Recombinant target cells showing a fluorescence shift could be screened using a cell sorter. The established clones exhibited similar profiles of fluorescent expression. The scFv-Fc productivity was improved by increasing the copy number of the transgenes.

Minicircle DNA vectors lacking the bacteria backbone sequences were applied as donor vectors for efficient transfection and expression of transgenes, the differences between transgenes with and without backbone in the same integration site are investigated, and two methods are explored for the deletion of the backbone of plasmid, one is using a minicircle DNA vector, which does not contain the backbone sequences; another is using Cre/*loxP* system to delete the bacteria sequences after integration of a plasmid vector. By comparing the targeted integration efficiency, this study investigated what kind of vector construction exhibits the best for introducing transgenes into cells, and what kind of recombinant system enables to get promising target clones for industrial production, providing instructions for generation of producer cell lines in recombinant protein manufacturing.

1.3 Thesis components

In Chapter 1, background, bottlenecks and current techniques in the fields of biopharmaceutical manufacture are introduced, particularly the recombinant protein produced in mammalian cells. In addition, the research purpose and strategy of the study in this thesis are briefly introduced.

In Chapter 2, techniques and mechanisms related to the study in this thesis are reviewed.

In Chapter 3, a study on the DNA transfection using minicircle vectors lacking bacteria-derived plasmid backbone sequences is demonstrated, which improved the target gene integration efficiency and the cells viability during cultivation.

In Chapter 4, an improved version of AGIS that facilitates and accelerates the establishment of high-producer Chinese hamster ovary (CHO) cells is described. This system enables speed-up and repeated multiple transgenes integration into a predetermined *hprt* locus in mammalian cells.

In Chapter 5, the contents of this thesis and future application are summarized.

Chapter 2

Background

2.1 Introduction

Gene amplification and target protein overexpression are the priorities among central interests in the fields of industrial protein production. To accomplish target protein overexpression in cultured cells, two essential elements are needed. The first one is to bring the target gene into fonder cell genome and increase integrated gene copy number. The most widely used method to achieve this is to use genome modification techniques combining inhibitor drug screening methods for gene integration and amplification. The increase of the target gene copy number allows the target product overexpression. This process is critical for achieving high-yield of various recombinant protein pharmaceuticals for satisfying the industrial production requirement. The second element is to achieve the high-producing recombinant proteins by ensuring amplified gene transcript, which requires that the transfected genes are inserted into an appropriate locus for high-level expression; Site specific recombination could be used to bring in the target gene into predicted site [21], and many genomic loci reported as genomic safe harbors for stable and/or high transgene expression could be applied [22-23]. In addition, an improper site could cause gene silencing, apoptotic cell death or decrease of biomedical product titer through affecting cell growth ability.

Based on these two crucial elements, this chapter has reviewed gene editing techniques, target gene amplification methods for antibody overexpression methods, and genome loci applicable to transgene expression.

2.2 Genome Editing Techniques

Taking the advantage of DNA editing techniques, it seems possible to give new phenotypes to living things or disable their abilities nowadays. We have plenty of options suitable for various situations. A literature review for the major gene editing techniques widely applied in both laboratory and commercial production has been conducted.

2.2.1 Gene Targeting by HR (Homologous Recombination)

Since the inchoate genomic edited mammalian cells emerged in last 1980s [24-26], biologists cheered for the purposeful gene editing based on HR (Homologous Recombination). HR is a genetic recombination method, in which nucleotide sequences are exchanged between two identical or analogous DNA fragments [27]. It happens when cells want to repair harmful DSB (double-strand DNA breaks). When DSB occurs, a stretch of DNA near the 5' ends of the break are cut off in resection step, then, in the strand invasion step, the overhanging 3' end of broken DNA will invades another undamaged identical or analogous DNA to form the D-loop. Afterwards, as show in Fig. 2-1, events like DSBR (double-strand break repair) or the SDSA (synthesis-dependent strand annealing) process for bringing in of foreigner DNA sequences.

However, the application of HR mediated target editing has been hampered by several reasons, such as the low transfection efficiency and adverse mutagenic effects, further more, as HR only emerge in prokaryotic DNA replication and cell cycle after S phase in eukaryotic cells [28]. Usually, the screening process for obtaining a proper clone by HR method takes a relatively long time.

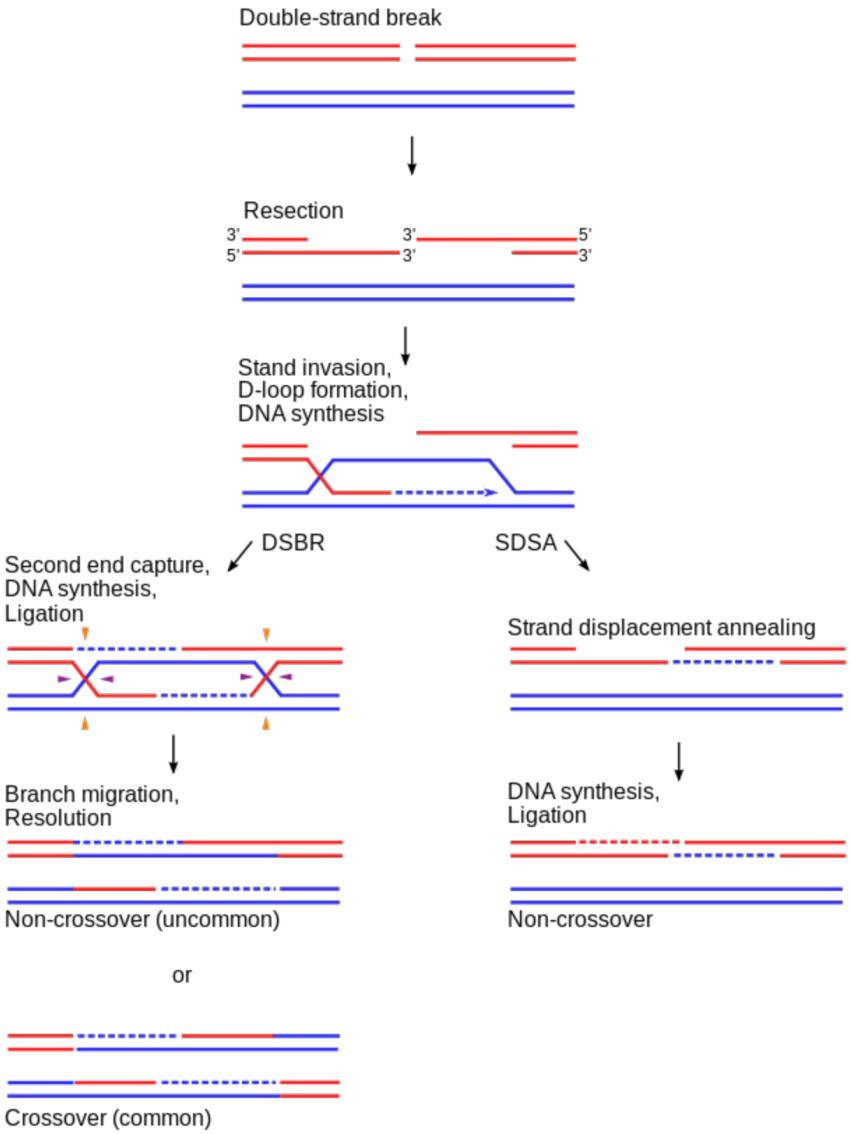


Fig. 2-1 Schematic diagram of two pathways followed the double strands break in homologous recombination. The DSBR pathway most often results in chromosomal crossover (bottom left), while SDSA always ends with non-crossover products (bottom right). Revised on reference [27].

2.2.2 Genome Editing by Engineered Endonucleases

According to the rapid development of genome sequencing and newly emerged genome editing techniques, investigators are able to manipulate a diverse range of genes in various types of cells and organisms. These versatile and commercialized techniques are represented by ZFNs (Zinc-finger nucleases), TALENs (transcription activator-like effector nucleases), and CRISPR/Cas9 (CRISPR/Cas-based RNA-guided DNA endonucleases). The primer cause of the commercialization of these nucleases is the available of programmable nucleases composed of sequence-specific DNA-binding domains that fused to a nonspecific DNA cleavage. By bringing in targeted DNA DSBs stimulating the cellular DNA repairing, including error-prone NHEJ (non-homologous end joining) and HDR (homology-directed repair) [29], these chimeric nucleases could be capable of modifying genome sequences precisely and efficiently at a certain site. Introduction of the basic theories and the application prospection of these commercialized gene editing techniques were reviewed in this section.

2.2.2.1 ZFNs (Zinc-finger nucleases)

ZFNs are composed of two domains. One is a DNA-binding domain, in which 3 to 6 individual zinc finger repeats are contained, each zinc finger can recognize 3 basepairs DNA nucleotides to generate a 3-finger array. As shown in Fig. 2-2A, with rightly selected artificial synthesized individual zinc fingers, DNA-binding domain can specifically target on desired sequences. Another is DNA-cleavage domain, a non-specific type II restriction endonuclease *Fok*I. *Fok*I must fuse to C-terminus of Zinc-finger nucleases to dimerize before making a DSB. As shown in Fig. 2-2B, in order to cleave both strains of genome DNA, two individual ZFNs were needed to bind opposite DNA strands with a certain distance. Commonly, 5' end of two binding sites were separated by 5 to 7 bp [30-31]. ZFNs ignited DBS could from targeted mutagenesis or facilitate the target gene insertion or alteration by coordinated with homologous donor DNA.

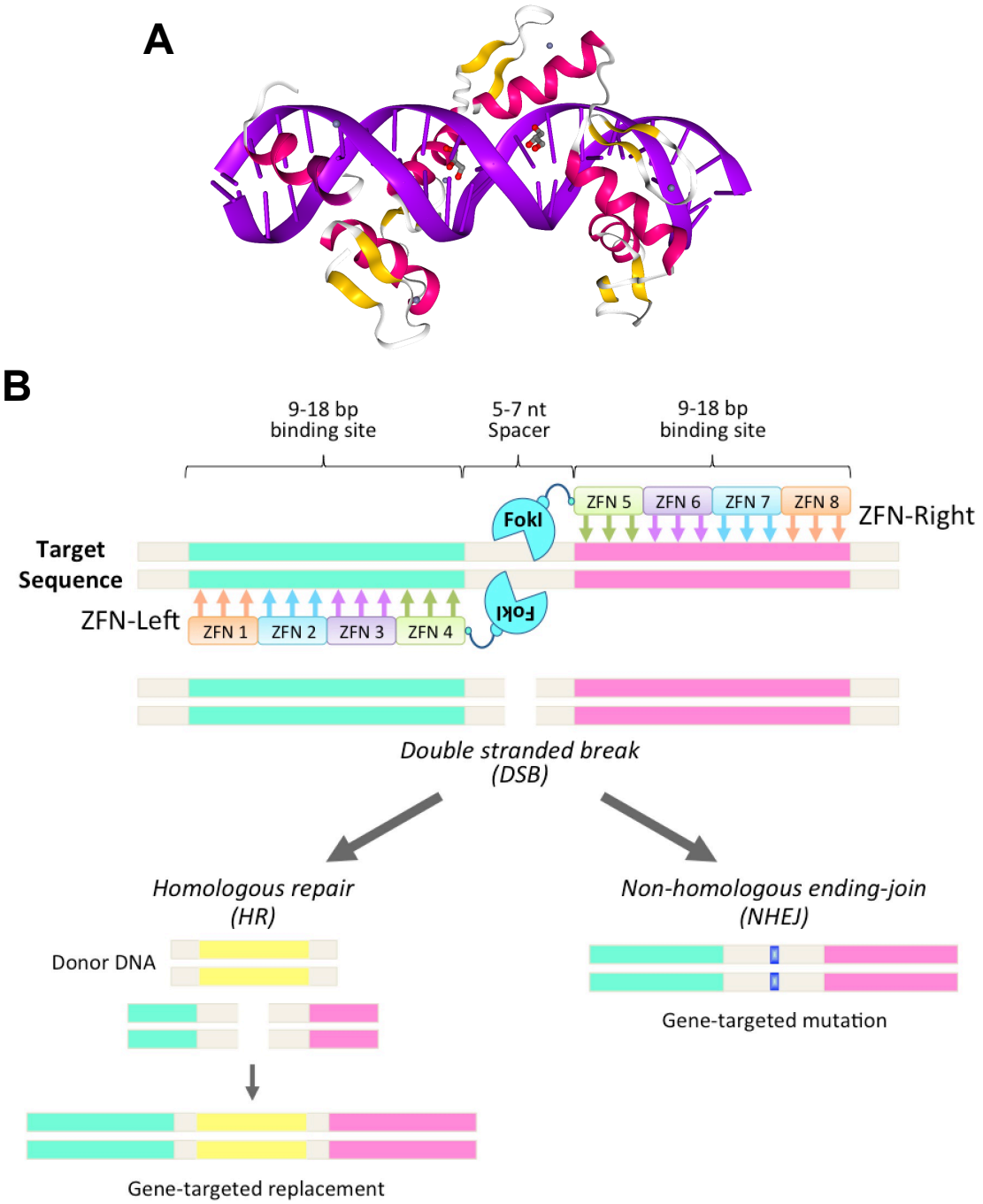


Fig. 2-2 Schematic diagram of ZFN components and genome-editing processes.

(A) The Crystal of Zinc-finger protein fused to target DNA sequence (PDB ID: 2I13);

(B) Repair modules of double strand break after ZFN nuclease cleavage. Homologous recombination repair is using the homologous donor DNA as template (left). Alternatively, the break can be repaired by non-homologous end joining, leading to mutations at the cleavage site (right). These may be deletions, insertions, and base substitutions, usually quite localized, but sometimes extending away from the break.

As the first widely applied innovative gene editing technology, ZFNs are available to manipulate the genomes of various mammalian cells, including human cells. Via ZFNs caused CCR5 gene disturbing on the CD4+ human T-cells, Zinc finger nucleases have already been used in clinical trial by Sangamo BioSciences (Richmond, CA) as a potential treatment for HIV/AIDS [32]. ZFNs are also used to model human diseases with isogenic human disease models due to its ability of achieving high genetic knock-out rate within a target endogenous gene. However, one significantly concerned drawback is that any off-target double strand breaks could trigger randomly off-target insertions or deletions in the genome.

2.2.2.2 TALENs (transcription activator-like effector nucleases)

TALENs, another developed protein-mediated gene-editing technique similar to ZFNs, are composed of the non-specific DNA-cleaving nuclease fused to a customizable DNA-binding domain and cleavage domain (shown in Fig. 2-3). The DNA-binding domain is composed of highly conserved repeats proteins secreted by *Xanthomonas* bacteria when they infect host plant [33], two divergent 12^{th} and 13^{th} amino acids in this nucleases referred to Repeat Variable Diresidue (RVD) are highly variable, and it could be designed to recognize specific nucleotide sequence or improve the cleavage specificity and/or activity in yeast assay [34], plant cells [35], and animal cells [36]. Additionally, TALEN can combine with other genome editing tools such as meganucleases to improve the recognition specificity. After cleavage, TALEN induced in situ double-strand breaks (DSB), which cells respond to repair mechanisms as same as ZFNs (as illustrated in Fig. 2-2B). Foreign DNA could be introduced into DSB via HR. Alternatively, targeted genetic mutation happened through NHEJ.

Unlike monopolized ZFNs, artificial TALE libraries are commercialized and offered by many companies, such as Cellectis Bioresearch (France), Transposagen Biopharmaceuticals (Lexington, KY, USA), Life Technologies (Grand Island, NY, USA) and so on. Besides,

TALEN restriction enzymes are easier to be engineered for cutting specific DNA sequences [37]. These current statuses impel the wider application of TALEN.

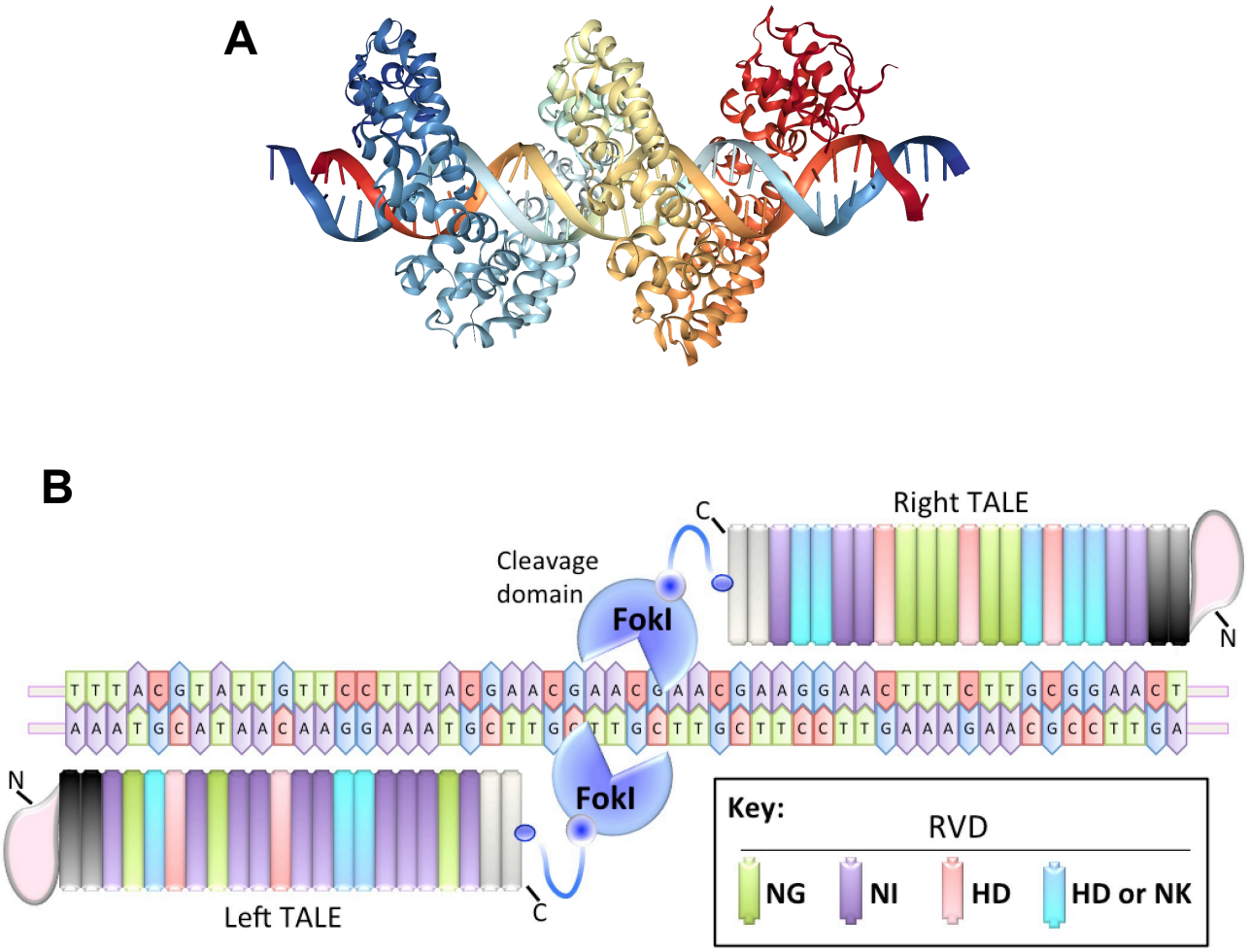


Fig. 2-3 Schematic diagram of TALENs components and genome-editing processes.
(A) The Crystal Structure of TALE protein bond to target DNA sequence (PDB ID: 3UGM);
(B) Schematic diagram illustrated TALE protein binding to target DNA sequence by 12th and 13th amino acids (high repeat variant RVD) and cleavage domain controlled by *Fok*I endonuclease, here just shows assembly of four kinds of RVD.

2.2.2.3 CRISPR（Clustered regularly interspaced short palindromic repeats）

Unlike protein-depended ZFNs and TALENs techniques, newly developing CRISPR relies on ribonucleotide complex formation. As n in Fig. 2-4A, the commonly applied CRISPR/Cas9 system typically contains three major components: cas genes, and a repeat-spacer array and tracRNA gene [36]. And the trans-activating (tracr) RNA will form a dsRNA (guide RNA or trac/crRNA) with CRISPR RNA (crRNA) in four-component system including Cas9 endonuclease and RNaseIII [38-39]. Further more, guild RNA will find and cut the target DNA to make DSB for genomic editing. For any targeted sequence in mammalian cells genome, corresponding Guild RNAs can be easy designed and cheaply produced, which makes CRISPR more efficient and wider used than TALENs and ZNFs nowadays.

The high cleavage efficiency of CRISPR could be mentioned in the same breath to restriction enzymes. However, the target-off phenomenon troubles the researchers and hinders it from coming into service. To solve this problem, advanced CRISPRS are rapidly developed. Rudolf Jaenisch have injected the mice by using target RNA instead of target DNA, they verified the RNA injection had better accuracy in CRISPRS/Cas9 and established a quicker [40], cheaper and easier method for setting up animal model. Another problem is the CRISPR/Cas9 introduced DSB is a blunt end, which challenge the targeted-DNA exchange. Sergey Shmakov and his group have found CRISPR/Cpf1 and CRISPR/C2c1 could cleavage the target DNA into a sticky end, making the introduction of foreigner DNA available [41-42]. The C2c1 can produce a 7nt sticky end, which was supposed to improve the HR efficiency. Compared to Cas9, the smaller size of protein Cpf1 and C2c1 makes it easier to get into organs and cells, and Cpf1 is flexible as its cutting site is far from the recognition site. The difference structure and working processes are shown in Fig. 2-4B.

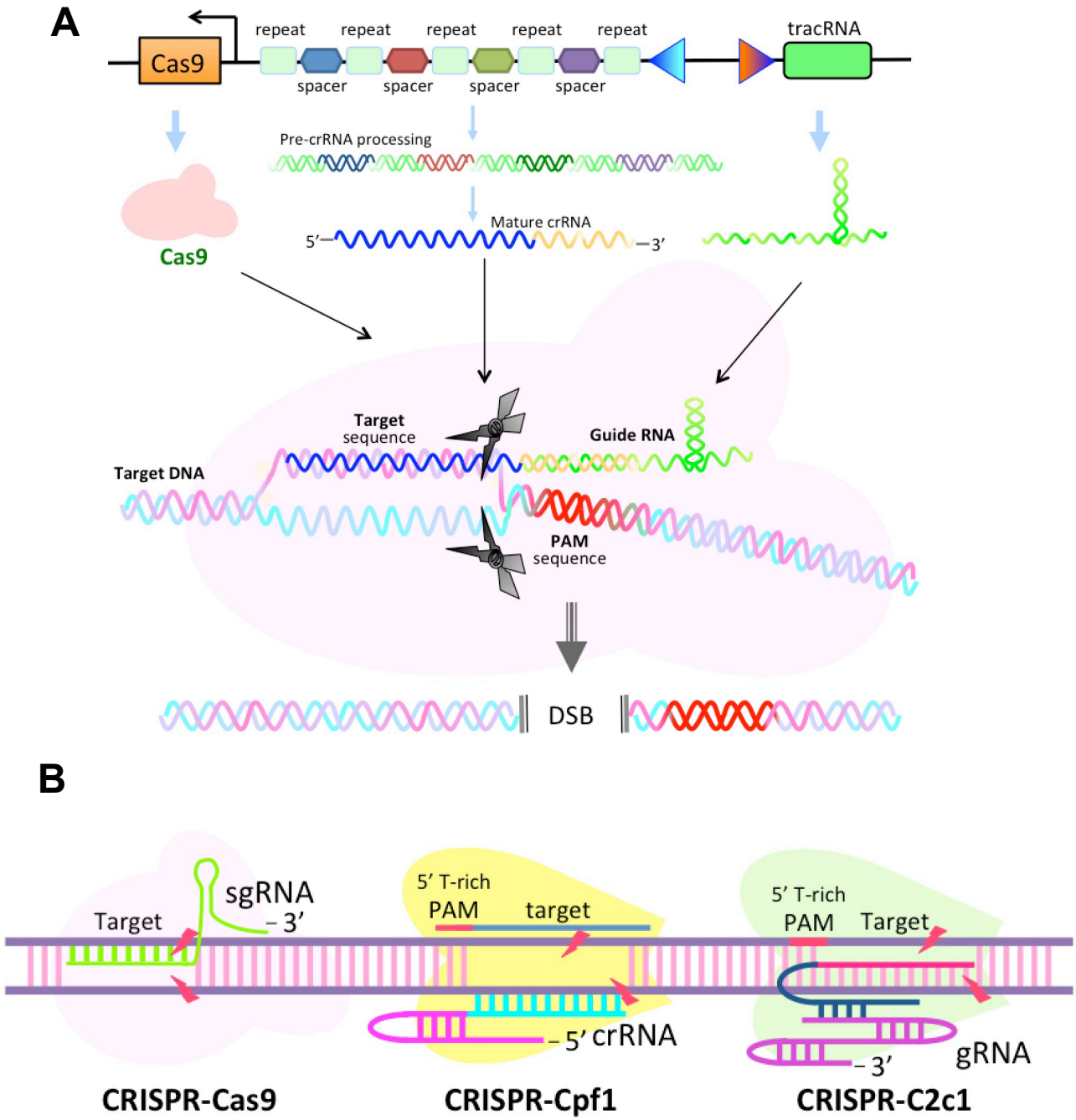


Fig. 2-4 Schematic diagram for DNA-cutting mechanisms of CRISPR systems.

(A) Simplified diagram of a CRISPR locus and cas9 working process. Three principal components of a CRISPR locus are shown: cas genes, a repeat-spacer array and leader trace sequence. The trace RNA linked to cas9 protein while the crRNA bound to target DNA sequence for inducing DSB.

(B) Simplified illustrating dramas of three promising CRISPR system.

2.2.3 SSRs (Site-Specific Recombinases)

Site-specific recombinations are genetic integration (INT), excision/resolution (RES), and inversion (INV) between two specific DNA segments, and they are catalyzed by site-specific recombinases (SSRs), which play a pivotal role in the life cycles of many microorganisms including bacteriophages and bacteria. Because SSRs only catalyze on some specific sequences, they are highly specific, efficient and fast, even when they are acting on complicated eukaryotic cells [43].

TABLE 2-1. Different types of SSRs catalyzed DNA recombinations. [44-46]

Family	**Sources:** Bacteria/Phages; Yeasts	**Recognition target sites**	**Symbol(s)**
Tyr-Recombinases	Cre, Dcre, VCre, SCre, Dre, Vika	Individual;	
	Flp, KD, B2B3	34/48 bp	
Tyr-Integrases	λ, HK022, HP1	attP/attB	
Ser-Resolv./Invertases	ΓδParA, Tn3, Gin	attP/attB	
Ser- Integrases	ϕC31, Bxb1, R4	attP/attB	

Based on which amino-acid residue acted as the catalyzer, SSRs are classified as tyrosine-type or serine-type recombinases. For example, the famous Holliday junction is catalyzed by the tyrosine-type recombinases respected by Cre and Flp, while the rejoining of cleaved substrate DNAs and 180° rotation are catalyzed by the serine-type recombinases. These classified SSRs could be further divided into two sub-families according to the reaction

directions. Table 2-1 introduced different types of SSRs and the reactions happening on different distinguished target site.

The integration/resolution and inversion (INT/RES and INV) modes depend on the orientation of recombinant target sites, as indicated in Fig. 2-5. Depending on the recombination regulars of different kind of recombinase, researchers can realize the introducing of the predictable target sequences into host cell genome.

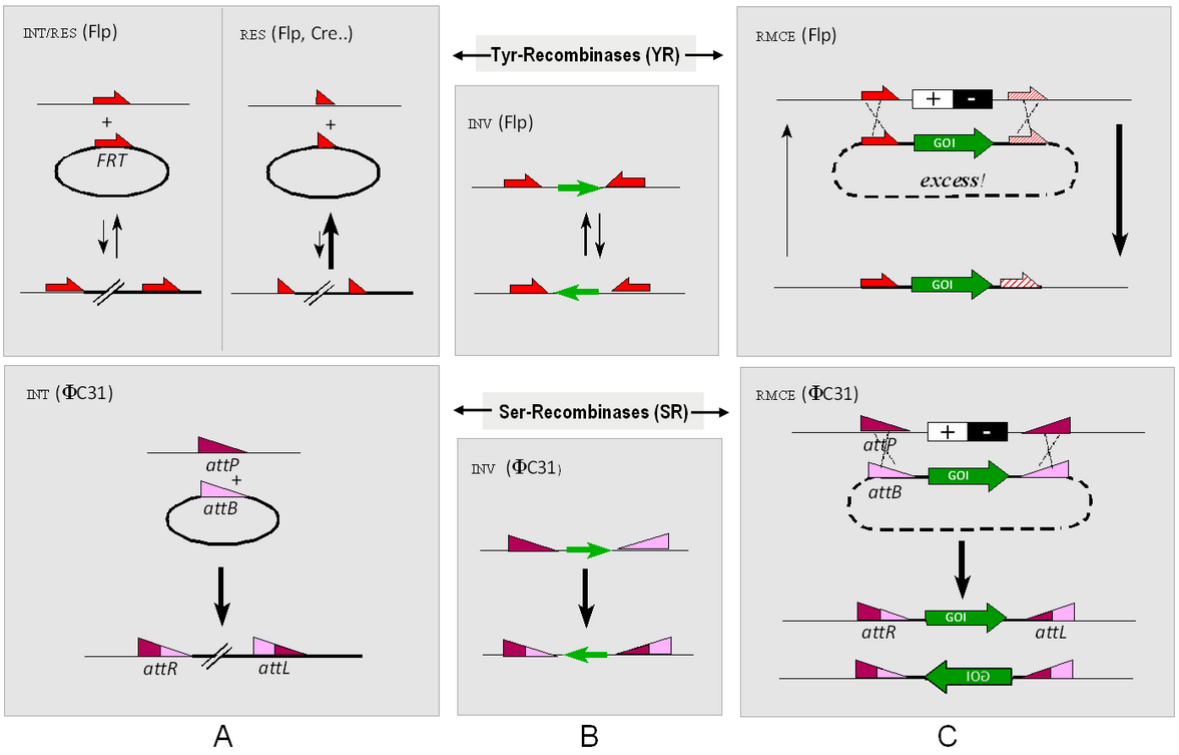


Fig. 2-5 Recombination regulars of different SSRs depend on target-site orientations. [44] Interaction of identical substrate sites (*loxPs* or *FRTs*) leads to products of the same composition, whereas recombination of two non-identical educts leads to different hybrid sites (*attP-attB* → *attR-attL*). GOI, "gene of interest"; [+/-], a positive-negative selective marker such as *hygtk*-fusion.

Nowadays, gene editing techniques like HR and SSRs can be combined into a highly efficient "tag-and-exchange technologies" [47]. In principle, HR can perform the specific genomic rearrangement, but it is compromised by an extremely low efficiency for

eukaryotes. Although ZFNs, TALENs and CRISPR are actual tools supporting HR, SSRs may be more handy and inexpensive. So, it is the utilizability of SSRs for introducing the reasonable construction into host genome after the foremost bring in of some specific sequences conducted by HR techniques.

2.3 Over-expression of Pharmaceutical Protein by Target Gene Amplification

Chinese Hamster Ovary (CHO) cells have become a workhorse for manufacturing recombinant proteins because they can produce recombinant proteins with "native human-like" modifications. The high yield of recombinant proteins in CHO cells can be achieved through gene amplification, which plays a pivotal role in generation of mammalian recombinant cell lines for industrial pharmaceutical producing. Because the chances of spontaneous transgene amplification occurring in mammalian cells is extremely low, various gene amplification methods have been developed to generate recombinant protein over-expressing cells with multiple copies of target genes. In this section, inchoate inhibitor drugs induced gene amplification methods and recently developed SSRs mediated target protein over-expression systems used in CHO cells are reviewed.

2.3.1 DHFR/MTX (Dihydrofolate reductase / Methotrexate)

Amplification of the dihydrofolate reductase gene (DHFR) under the selection pressure of methotrexate (MTX) is commonly used in generating high producer recombinant CHO cells. This most widely used high producer cell selective method involves amplification of the recombinant protein expression gene together with physically linked dihydrofolate reductase gene (DHFR) in CHO cell lines lacking the functional DHFR gene, under the selection pressure of increasing concentrations of methotrexate (MTX). The DHFR enzyme is able to converses folate to tetrahy drofolate, while the DHFR defective cells cannot live without hypoxanthine, thymidine and glycine. As the inhibitor of DHFR, MTX binds to

DHFR, there by inhibiting tetrahy drofolate production. Only cells with increased copies of DHFR genes are survived and selected. Meanwhile, the DHFR genes amplification drive the increase of adjacent target genes, therefore cause target genes amplification and overexpression of recombinant proteins. The Fig. 2-6 displayed the screen process of drug selected high producer cells.

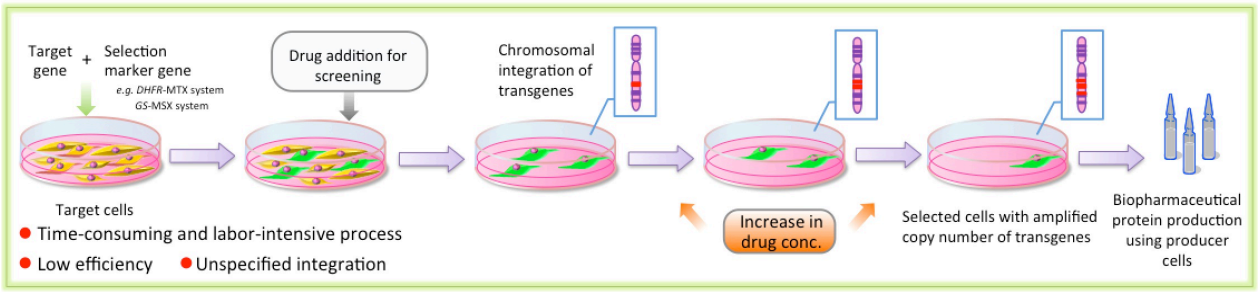


Fig. 2-6 Selective processes of high-level producer CHO cells under DHFR/ MTX or GS/MSX system.

DHFR/MTX method has been most commonly used in investigating of the mechanism of gene amplification [48] and has been extensively applied to approach industrial recombinant protein production. Although this method enables to obtain high-producer clones, it is labor-intensive and time-consuming, requiring repeated cycles of selection and sub-cloning [49]. In addition, high-producer clones obtained using this randomly inserted system are reported to exhibit instability like rapid decrease in target protein synthesis as culture time progresses [50].

2.3.2 GS/MSX (Glutamine Synthetase / Methionine Sulfoxamine)

The GS/MSX system is similar to the DHFR/MTX system (shown in Fig. 2-6). Glutamine synthetase (GS) is an enzyme that catalyzes the conversion of glutamate and

ammonia to glutamine. Methionine sulfoxamine (MSX) binds to the GS enzyme and prevents the glutamine production. In the GS/MSX system, in order to achieve top-producing cell lines, the balance between the concentration of MSX inhibitor and expression level of GS (the selection stringency) is foremost. There are two ways to achieve large amount of target proteins, target genes amplification and/or valid cells concentration increase caused by the gradually increased MSX selective pressure. Similar to DHFR/MTX, GS/MSX system has been widely applied to pharmaceutical manufacturing. However, this technique also has the same problems of time-consuming and labor-intensive, as a long-term process with multi-rounds of selection is needed for obtaining a stable and high yield cell line [51-53].

2.3.3 IR/MAR (Initiation Region/ Matrix Attachment Region)

IR/MAR method is based on a plasmid carrying a mammalian replication initiation region (IR) with a matrix attachment region MAR [54], which caused the spontaneously amplification of transgenes in transfected cells. IR/MAR generate extra-chromosomal double minutes (DMs) and/or chromosomal homogeneously staining regions for gene amplification [55-56], and the selection of target clone requires selective pressure caused by resistance drug such as blasticidin or neomycin. High efficiency of IR/MAR-induced gene amplification is detected in almost all transfectants, this quick amplification is initiated by formation of a large extra-chromosomal circular molecule containing multiple arranged repeated sequences of IR/MAR plasmids before chromosomal integration [54]. In addition, if the plasmids are extendedly amplified, the produced DNA appears DMs damaged extra-chromosomal circular DNA will be eliminated from the cells through the generation of micronuclei [55].

An IR/MAR-DHFR fusion method was reported of dramatically increased the expression the GFP, Fc receptor, and recombinant antibody in CHO cells efficiently [54]. In contrast,

most of the recombinant protein producer cells generated by the inhibitor drug selection method are highly unstable, and the recombinant proteins production rate decreased rapidly along with the culture time progresses.

The IR/MAR-DHFR combination method resulted in stable amplification and generated clonal cells that produced large amounts of antibody proteins over a long period of time. Thus, the IR/MAR method helps the generation and isolation of stable cells producing high-level recombinant proteins.

2.3.4 Site-specific Recombination Mediated Target Gene Amplification

In defiance of the breakthroughs in novel expression methods for industrial producer cell line engineering, unpredictable transgenes silencing and unstable proteins expression level as the results of randomly integrated locations in host genomes are still the bottlenecks in pharmaceutical manufacturing. To generate a suitable cell lines for expressing high amount of recombinant proteins using random integration linking genomic amplification, the conduction of several rounds of selection is necessary; however, it is laborious-intensive and time-consuming. Also, due to position effects, most of the transfected cells showed low and unpredictable yields of expression after long time culture [57].

Site-specific recombination with engineered vectors and high-throughput selection methods such as fluorescence-activated cell sorting associated the selection for high productivity and stable clones were developed and used. Site-specific recombination methods increase the recombinant protein production by specifically inserting a vector at a locus with specific expression trait. There are several commercially available systems for producing site-specific recombinant proteins, for instances, (i) tyrosine recombinases including Cre and Flp have been applied for site-specific integration in animal cells [58], and many attempts have been made to develop the specificity and efficiency of these two system [59-60]. (ii) Class II transposable elements such as the *Sleeping Beauty* (SB),

piggyBac (PB), and *Tol2* transposons, they move in the host genome via a "cut and paste" mechanism [21]. Use of the site-specific recombination techniques may be relatively expensive, but the screening for acceptable cell line is performed only once. Thus, site-specific recombination techniques are considered as the most promising tool of genetic modification for biopharmaceutical manufacturing, they deserve to be exploited.

2.4 GSHs (Genomic safe harbors) for recombinant proteins production

As previously mentioned in this thesis, cell lines possessing high-copy number of transgenes commonly obtained through gene amplification under inhibitor selective pressure while the target genes were randomly integrated into the host cells genome. Inevitably, the circumstances of integration sites lead to transgene silencing sometimes. Foreign genes could cause genetic interference, heterogeneous chromatin and even cell apoptosis. Although these disadvantages could be improved by avoiding hypernomic pressure selecting, the discovery of genomic safe harbor becomes quite necessary for producing recombinant proteins.

Genomic safe harbors (GSHs) are host genomic locus able to accommodate the integration of new genetic material in a manner that ensures newly inserted genetic elements: (a) function predictably and (b) do not cause alterations of the host genome posing a risk to the host [22]. Therefore, GSHs are ideal sites for transgene integration whose use can empower and stabilize recombinant protein production. Unfortunately, no absolutely validated GSHs exist in the mammalian genome currently. Here review for formerly proposed GSHs in therapeutic applications or recombinant protein producing was given. In view of recent advances in genome biology, gene targeting technologies and regenerative medicine, gene insertion into GSHs can potentially catalyze a better establishment of high-producer cell lines.

2.4.1 AAVS1 (Adeno-associated virus integration site 1)

The adeno-associated virus integration site 1 (*AAVS1*) locus refers to the region near the first exon and intron of the *PPP1R12C* (protein phosphatase 1, regulatory subunit 12C, chromosome 19) gene, which is ubiquitously expressed. It is considered as a GSH based on the observation that non-pathogenic adeno-associated virus 2 (AAV2) integrates at this site. Monoallelic disruption of the *PPP1R12C* gene does not appear adverse affect in the target cells [61].

The *AAVS1* site has an open chromatin conformation structure allows transgenes access to perform recombination, transcription, replication, and chromosome segregation [62]. The open chromatin structure at the *AAVS1* site is also associated with cis-acting insulators [63], an insulator-like structure acting as boundaries to the surrounding heterochromatin that silences the integrated genes have been reported. Indeed, transgenes integrated in the *AAVS1* site showed long term and stable expression in a various host cells including hiPSCs and hESCs [64-65]. Accordingly, the *AAVS1* locus could serves as a useful site for integration of foreigner genes.

2.4.2 CCR5 (C-C chemokine receptor type 5)

The *CCR5* protein is a G protein–coupled receptor functions as a chemokine receptor. The chemokine (C-C motif) receptor 5 (*CCR5,* chromosome 3) gene loci was identified as a GSH after the discovery of people naturally carrying the *CCR5* gene disruptions were resistant to HIV-1 infection and have no pathology phenomenon [66]. Genetic approach involving intra-bodies that block *CCR5* expression has been proposed as a treatment for HIV-1 infected individuals.

Relevant clinical trial was using zinc finger nuclease modified patient cells that carrying the *CCR5*-Δ32 trait, then the genetically modified cells were reintroduced into the patients' body, the reported results were promising [67]. The delivery and long-term expression of

shRNAs on *CCR5* locus in human primary peripheral blood mononuclear cells primary hematopoietic stem have also been reported [68], indicated the stable transcription of transgenes inserted in *CCR5* loci.

2.4.3 ROSA26 (Gt(ROSA)26Sor)

Gt(ROSA)26Sor, commonly refers to *ROSA26* locus, was first described as a gene trap which is ubiquitously expressed in mouse embryos cells [69]. *Rosa26* is highly expressed in most cells and tissues of the adult mice, even wild type mice do not have β-galactosidase insertion.

Gt(ROSA)26Sor can be readily modified by homologous recombination in ES cells to express transgenes ubiquitously (more or less) in embryonic and adult mice [69]. There are even ready-made targeting vectors that facilitate genetic editing on this locus, and the integrated transgenes were definitely located in same place on chromosome 6. In contrast, classical techniques cause randomly integration of plasmid sequences into the host genome. Furthermore, targeted integration in *Rosa26* generally does not affect the viability or fertility of the mouse, as no promoters is restricted in its expression in this locus by unfavorable chromatin configurations. Thus, *Rosa26* has been widely applied for expressing endogenous sequences as reporter genes as a GSH [70].

2.4.4 HPRT (Hypoxanthine phosphoribosyl transferase)

Hypoxanthine phosphoribosyl transferase (*hprt*, chromosome X), a unique locus in CHO cells chromosomes, is a housekeeping gene related to the conversion of hypoxanthine to inosinemonophosphate and guanine to guanosine monophosphate in the non-essential purine salvage pathway.

Hprt locus can provide an efficient alternative strategy to reproducibly generate animal models with tetracycline-induced transgene expression, which has been reported in

developing the advantages of mosaic mice, cell hybridization and gene transfer techniques [43]. In addition, long-term expression level of transgene was also reported [71], demonstrated that the *hprt* locus could provide an efficient and simple method to enrich target recombinant proteins in successfully targeted cells.

2.5 Conclusion

For industrial biopharmaceutical production, the construction of producer cell lines is one of the most important steps toward achieving stable and high productivity of proteins with acceptable quality for human use. Conventionally, producer cell lines are generated by transfection of a vector encoding a target gene expression unit, followed by the application of gene amplification methods such as dihydrofolate reductase-methotrexate and glutamine synthetase-methionine sulfoximine selection [72]. In these methods, transgenes are randomly integrated into the CHO cell genome, and hence producer cells are screened from a large number of cells with a variety of transgene integration sites, requiring laborious and time-consuming work.

Specific genomic loci for stable and/or high transgene expression have been identified as genomic safe harbors or hot-spot sites [22-23]. Targeted integration of transgenes into such locus can be achieved by using the site-specific recombinant system whereby the recombinant target site is pre-introduced into the desired target locus. These systems can be expected to guarantee fast and efficient selection of stable clones with high productivity.

Chapter 3

Improvement of Cre-mediated transgene integration using minicircle DNA vectors

3.1 Introduction

A plasmid vector for recombinant protein expression in mammalian cells mainly consists of two parts, a bacterial backbone and an eukaryotic transcription cassette. Commonly, the bacterial backbone contains necessary units for plasmid amplification in bacterial cells, such as the replication origin for amplification, the resistant gene for selecting the bacterial carrying target plasmid and ensuring stable plasmids replication, and so on. The eukaryotic expression cassette is used to produce target protein in transfected cells or organs, it includes the transgenes for genetic therapy or genome editing in eukaryotic cells, eukaryotic regulatory elements such as the eukaryotic promoter and the polyadenylation terminator required for transgene expression. The bacterial backbone usually spatially takes one half of the plasmid without any useful information for recombinant protein expression, and the commonly utilized antibiotic resistance markers could cause unstable and low-efficacy in target protein expression [73].

A smaller minicircle DNA vector lacking the bacterial backbone sequences has avoided the undesirable effects caused by bacterial backbone sequences. Minicircle DNA vectors could be produced by site-specific recombination occurred at the joint regions between backbone sequences and eukaryotic transcription cassette (Fig. 3-1). Comparing to the conventional plasmids carrying the same eukaryotic transcription cassette, the minicircle vectors showed a better transfection efficiency and a more stable transgene expression [74].

So far, most reports about the advantages of minicircle DNA vectors are focusing on gene therapy of animal models in vivo, but few studies are related to its merits on recombinant proteins producing in animal cells in vitro. In current study, targeted gene editing and recombinant protein producing by using a minicircle DNA vector was investigated and compared with the conventional plasmid.

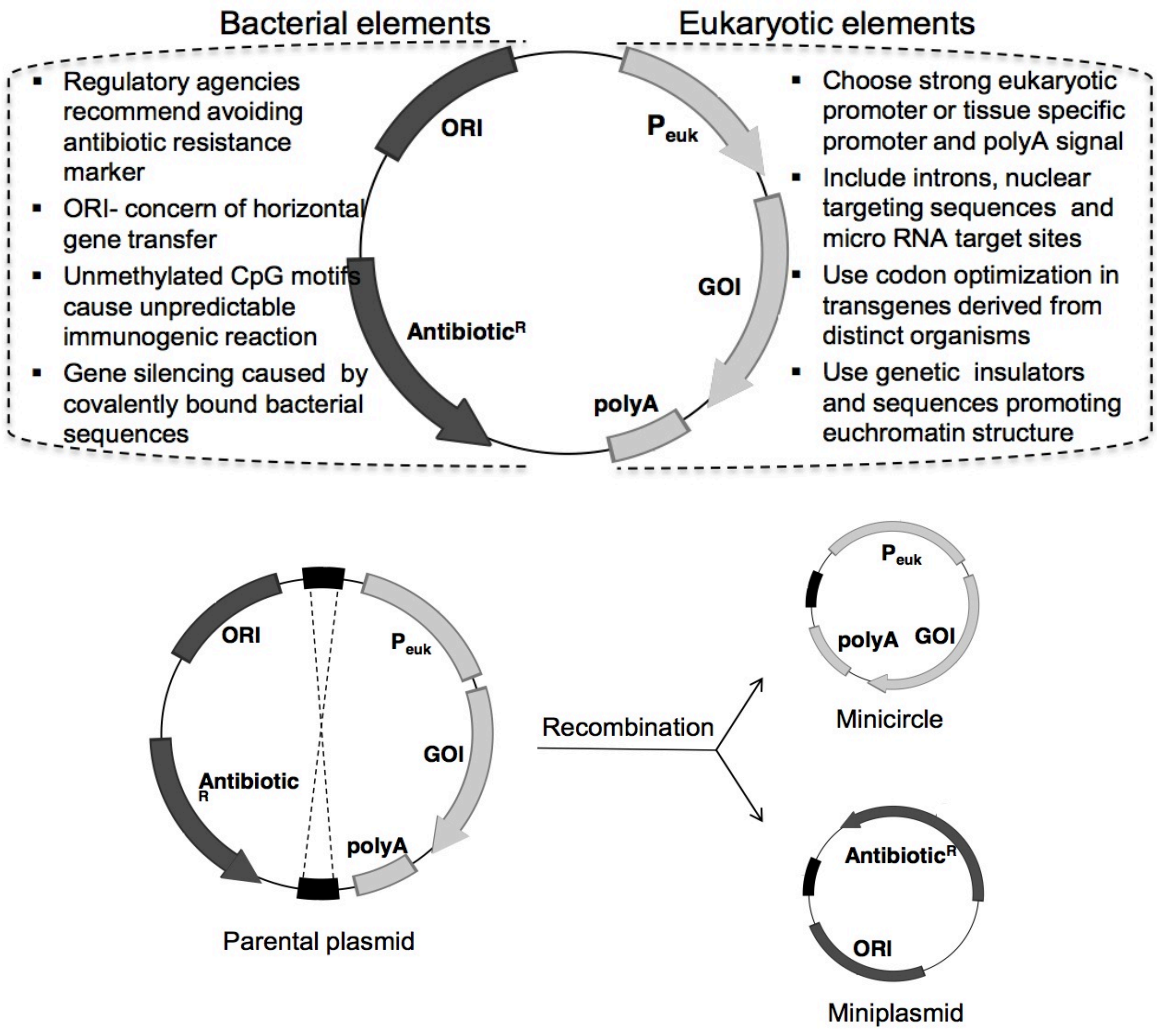


Fig. 3-1 Schematic diagrams for functions of each unit in a conventional plasmid and the mechanism of producing minicircle DNA vectors [75].

3.2 Improvement of Cre-mediated transgene integration using minicircle DNA vectors

3.2.1 Experimental Purpose

Bacterial backbone sequences of a vector have been reported to show many disadvantages in transgene expression and integration in vivo and/or in vitro. A minicircle DNA vector can improve the transfection efficiency as it has a much smaller size compared to conventional plasmid vectors, but few reports were studied on its affection on target site recombination accuracy and efficiency. In this study, the differences between vectors with and without bacterial backbone sequences were investigated, when a transgene in the vectors was integrated into the same chromosomal site of CHO cells using Cre/*loxP* system. Two methods were explored for the deletion of the bacterial backbone sequences of a plasmid; one was using a minicircle DNA vector, which did not contain the bacterial backbone sequences, and another was using Cre/*loxP* system to delete the bacterial backbone sequences after integration of plasmid vector. By comparing the targeted integration efficiency, this study investigated what kind of vector construction exhibits the best for introducing transgenes into cells, and what kind of recombinant system enables obtaining of target clones for industrial production, thus providing instructions for genomic modification in recombinant proteins manufacturing.

3.2.2 Materials and Methods

3.2.2.1 Cells and culture media

The founder cells used for transgene integration were CHO/R1 cells, in which a mutated *loxP* site (*loxP1*) and an expression cassette encoding the red fluorescent protein, DsRed, were introduced into the *hprt* locus. Cells were cultured under adherent conditions using Ham's F12 medium (Sigma-Aldrich, St. Louis, MO, USA) supplemented with 10% fetal bovine serum (Biowest, Nuaillé, France), 100 units/mL streptomycin sulfate and 90 μg/mL

penicillin G potassium (Wako Pure Chemical Industries, Osaka, Japan) at 37°C and 5% CO_2 in a humidified incubator.

3.2.2.2 Plasmid construction

A targeting vector harboring homologous arm regions containing exon 3 of the *hprt* gene was constructed as follows. Homologous DNA fragments were amplified by polymerase chain reaction (PCR) for the 5′ homologous arm region (HA1; 2.5×10^3 base pairs (bp); primers Pr1 and Pr2) and the 3′ homologous arm region (HA2; 2.5×10^3 bp; primers Pr3 and Pr4). A kanamycin/neomycin resistance gene (*Neo*) expression cassette was amplified by PCR from pQBI25 (Wako), using primers Pr5 and Pr6. A DNA fragment encoding a bacterial replication origin and an ampicillin resistance gene was amplified by PCR from pQBI25, using primers Pr7 and Pr8. These DNA fragments were ligated together to generate pHA1/SV40/*Neo*/polyA/HA2 for *hprt* gene targeting.

A gene encoding ATG-deleted red fluorescence protein, (ATG–)*DsRed,* which was tagged with ATG and *loxP1*, was amplified by PCR from pIRES2-DsRedExpress (Clontech, Mountain view, CA, USA), using primers Pr9 and Pr10. The PCR product, tagged with *Bam*HI and *Hin*dIII sites, was digested with the relevant restriction enzymes and ligated into *Bam*HI- and *Hin*dIII-digested pBApo-EF1α Pur DNA (Takara Bio, Kusatsu, Japan) to generate pBApo/*DsRed*. A chEF1α promoter sequence was amplified by PCR from CHO-K1 cell DNA (cooperated by TOTO company). The PCR product, digested with *Eco*RI and *Bam*HI, was ligated into *Eco*RI- and *Bam*HI-digested pBApo/*DsRed* to generate pBApo/ chEF1α*DsRed*. A DNA fragment encoding a *DsRed* expression unit (chEF1α/ATG-*loxP1*-(ATG–)*DsRed*/polyA), comprising a chEF1α promoter, an ATG-*loxP1*-(ATG–)*DsRed* construct and a TK polyA signal region, was prepared from pBApo/chEF1α*DsRed* and was ligated into pHA1/SV40/Neo/polyA/HA2 to generate pHA1/chEF1α/ATG-*loxP1*-(ATG–) *DsRed*/polyA/SV40/Neo/polyA/HA2 (R1).

An *scFv-Fc* antibody gene was amplified by PCR using primers Pr11 and Pr12 from plasmid in previous research [76]. The PCR product was ligated together with a chEF1α promoter into pCR-BluntII-TOPO (Invitrogen, Carlsbad, CA, USA) to generate pchEF1α/*scFvFc*. A double-stranded DNA oligonucleotide, 5′-GAA TTC ATA ACT TCG TAT AAC CAT AAT TAT ACG AAC GGT AAC TAG TAA GAT ATC AAA TCG ATA ACT GCA GAA ACG CGT AAG CTA GCT ACC GTT CGT ATA AAG TAT CCT ATA CGA AGT TAT CCG GAT CCA ACT CGA GAT AAC TTC GTA TAA CCA TAA TTA TAC GAA GTT ATG CAT GC-3′, containing three mutated *loxP* sites, *loxP4*, *loxP2* and *loxP6*, was chemically synthesized (Medical & Biological Laboratories, Nagoya, Japan) and ligated into the plasmid, pIDTSMART-AMP (Medical & Biological Laboratories), to generate pIDT/*loxP4*/*loxP2*/*loxP6*. A blasticidin resistant gene (*Bla*) lacking the original ATG codon and the SV40 polyA signal region was amplified by PCR from pCEP4/Blar [76] using primers Pr13 and Pr14. The PCR product was ligated into *Bam*HI- and *Xho*I-digested pIDT/*loxP4*/*loxP2*/*loxP6* to generate pIDT/*loxP4*/*loxP2*/(ATG–)Bla/polyA/*loxP6*. A DNA fragment containing *loxP4*/*loxP2*/(ATG–)Bla/polyA/*loxP6*, prepared from pIDT/*loxP4*/*loxP2*/(ATG–)Bla/polyA/*loxP6*, was ligated into blunt-ended pBluescript (Stratagene, La Jolla, CA, USA), which was obtained by digestion with *Sac*I and *Xho*I, to generate pBlue/*loxP4*/*loxP2*/(ATG–)Bla/polyA/*loxP6*. A DNA fragment encoding an scFv-Fc expression unit (chEF1α/*scFv-Fc*/pA), comprising a chEF1α promoter, an *scFv-Fc* gene and the SV40 polyA signal region, was prepared from pchEF1α/scFvFc and was ligated into *Eco*RV- and *Nhe*I-digested pBlue/*loxP4*/*loxP2*/(ATG–)Bla/polyA/*loxP6* to generate pBlue/*loxP4*/chEF1α/*scFv-Fc*/pA/*loxP2*/(ATG–)Bla/polyA/*loxP6* (R2).

To construct the donor plasmid R2v2 with *loxP6* deleted, A DNA fragment containing *loxP2*, a blasticidin resistant gene (*Bla*) lacking the original ATG codon and the SV40 polyA signal region was amplified by PCR was amplified by PCR from pBlue/*loxP4*/chEF1α/*scFv-*

Fc/pA/*loxP2*/(ATG–)Bla/polyA/*loxP6* (R2), using primers Pr15 and Pr16. The PCR product was ligated into blunt-ended pBluescript (Stratagene, La Jolla, CA, USA), which was obtained by digestion with *EcoR*V, to generate pBlue/ *loxP2*/(ATG–)Bla/polyA. Then the DNA fragment encoding an *loxP2*, the (ATG–)Bla and the SV40 polyA signal region, was prepared from pBlue/ *loxP2*/(ATG–)Bla/polyA and was ligated into *Sph*I- and *Nhe*I-digested pBlue/*loxP4*/chEF1α/*scFvFc*/pA/*loxP2*/(ATG–)Bla/ polyA/*loxP6* to generate pBlue/*loxP4*/chEF1α/*scFv-Fc*/pA/*loxP2*/(ATG–)Bla/ polyA (R2v2).

To construct a parental plasmid for the production of a donor minicircle vector, chemically synthesized oligonucleotides (Pr17 and Pr18) incorporating cloning sites for restriction enzymes (*Xba*I, *Bam*HI, *Cla*I and *Sal*I) and a *loxP4* site were annealed to form a double-stranded DNA fragment with *Xba*I and *Sal*I sticky ends. This was ligated into *Spe*I- and *Sal*I-digested minicircle parental plasmid (pMC) (Cat. no. MN602A-1, SBI, Palo Alto, CA, USA) to generate pMC/*loxP4*. A DNA fragment encoding an scFv-Fc expression unit (chEF1α/*scFv-Fc*/pA), (ATG–)*Bla* with a polyA signal, and two mutated *loxP* sites, *loxP2* and *loxP6*, was prepared from R2 and was ligated into *Cla*I-digested pMC/*loxP4* to generate pMC-R2. To delete mutated *loxP*6, the DNA fragment encoding a portion of scFv-Fc, mutated *loxP2* and *ATG-deleted Bla*/polyA regions, was amplified by PCR from pMC/R2 using the primers Pr19 and Pr20. The PCR products were digested with *Sal*I and ligated into *Sal*I-digested pMC/R2 to generate pMC-R2v2.

The blunt-ended DNA fragments were prepared using a kit (DNA blunting kit, Takara). All PCR reactions were performed using KOD plus neo DNA polymerase (Toyobo, Tsuruga, Japan) according to the manufacturer's instructions. All DNA sequences derived from chemically synthesized oligonucleotides and PCR products were confirmed by DNA sequencing on a Prism 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

The construction flowchart for plasmid R1 used for fonder cells establishment was shown in Fig. 3-2, and this experiment was cooperated with TOTO Company.

The construction flowchart for R2, R2v2, pMCR2, pMCR2v2 plasmids used in current experiment are shown in Fig. 3-3, 3-4, 3-5 and 3-6.

The DNA sequences of the mutated *loxP* sites used in this study are shown in Table 3-1.

The serial numbers of *loxPs* are according to AGIS in this thesis.

The primer pairs used for vector construction are summarized in Table 3-2.

TABLE 3-1. Sequences of wild-type and mutated *loxP* sites.

Symbol	Sequence (5′→3′)		
	Left-arm region	Spacer region	Right-arm region
loxp	ATAACTTCGTATA	ATGTATGC	TATACGAAGTTAT
loxp1	ATAACTTCGTATA	AAGTATCC	TATACGAACGGTA
loxp2	TACCGTTCGTATA	AAGTATCC	TATACGAAGTTAT
loxp3	TACCGTTCGTATA	AAGTATCC	TATACGAACGGTA
loxp4	ATAACTTCGTATA	ACCATAAT	TATACGAACGGTA
loxp6	ATAACTTCGTATA	ACCATAAT	TATACGAAGTTAT
loxp8	ATAACTTCGTATA	AAGTATCC	TATACGAAGTTAT

Mutated sequences are underlined.

TABLE 3-2. Primers used for vectors construction.

Primer	Sequence (5'→3')
Pr1	TTC CTG CAG GTC GCG AGT CTG TGT GTA TGT TTG TGA TAG GC
Pr2	TTA CGC GTT GAT AAA ATC TAC AGT CAT GGG
Pr3	TTG GAT CCG ACT GAA GAG CTA CTG TGT A
Pr4	TTC CAT GGT CGC GAT GAA GGT TAT AGA GCA TAG GGG ACC
Pr5	CCT TTC TAG ACT TCT GAG GCG GAA AGA ACC
Pr6	CTT ATC GAT TCA CAC AAA AAA CCA ACA CAC AGA TGT AAT GAA AAT AAA GAT ATT TTA TTG TGG GCG AAG AAC TCC AGC A
Pr7	ACA TGT GAG CAA AAG GCC AGC AAA AGG CCA GGA AC
Pr8	CCT GCA GGG ACG TCA GGT GGC ACT TTT CGG GGA AAT GTG C
Pr9	AAG GAT CCA CCA TGA TAA CTT CGT ATA AAG TAT CCT ATA CGA ACG GTA GCG CCT CCT CCG AGG ACG TCA T
Pr10	TTT AAG CTT CTA CAG GAA CAG GTG GTG GC
Pr11	AAG AAT TCA AGC TTG CCG CCA CCA TGA GGT CTT TGC TAA TCT TGG
Pr12	CTA GCT AGC AAG CTT TCA TTT ACC CG
Pr13	AAG GAT CCG CCA AGC CTT TGT CTC AAG AAG AA
Pr14	AAA CTC GAG CAC CGT CAT CAC CGA AAC G
Pr15	AAG GAT CCT CTA GAG TGA GCA AGG GCG AGG AG
Pr16	GGA CTA GTG GGA TTT TGC CGA TTT CG
Pr17	CTA GAG GAT CCA TAA CTT CGT ATA ACC ATA ATT ATA CGA ACG GTA ATC GAT G
Pr18	TCG ACA TCG ATT ACC GTT CGT ATA ATT ATG GTT ATA CGA AGT TAT GGA TCC T
Pr19	CGG GAC CGA AGT CAT CGT G
Pr20	AAA AGT CGA CGC ACC GTC ATC ACC GAA AC

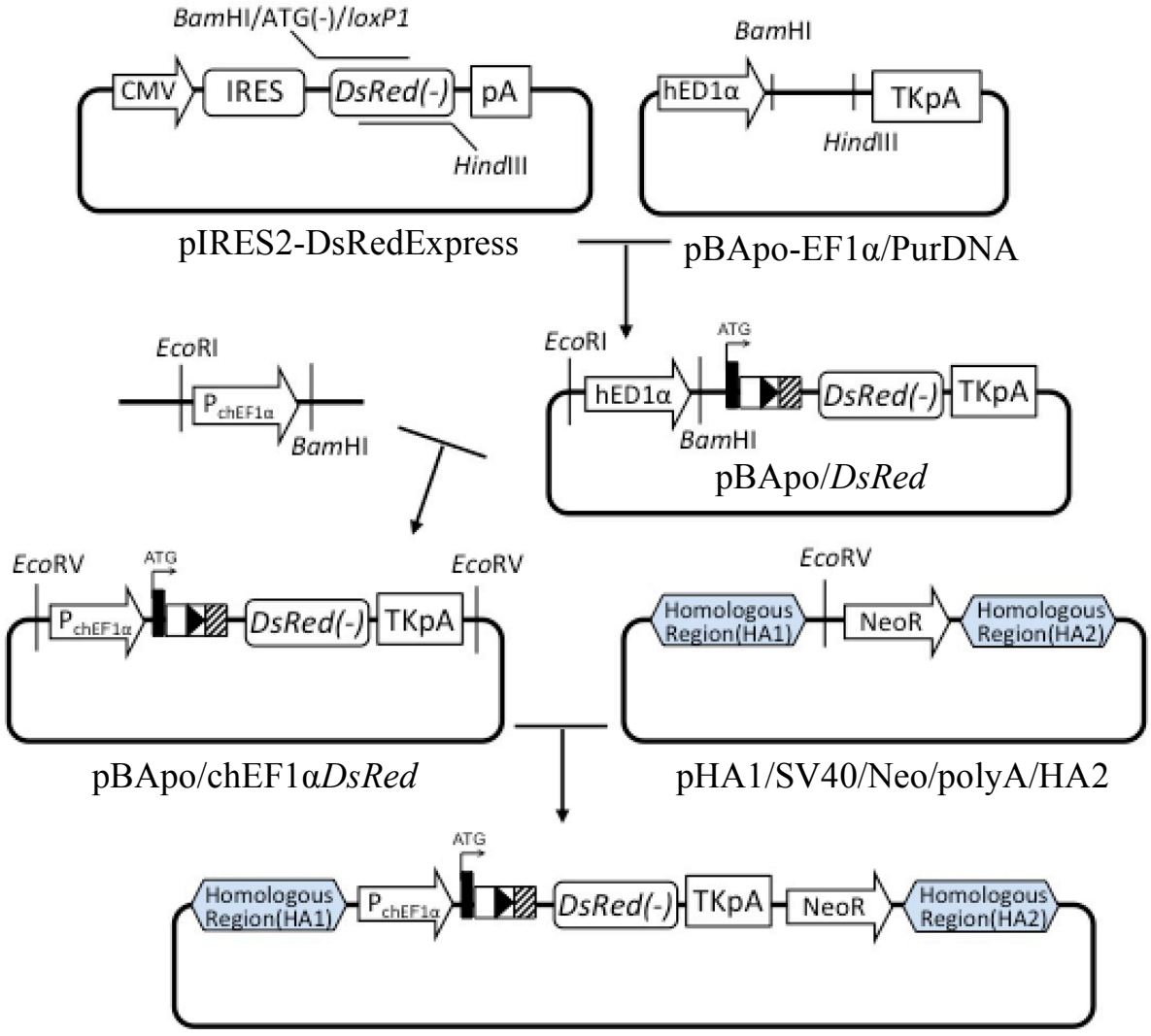


Fig. 3-2 Flowchart for the construction of R1 plasmid, which was cooperated with TOTO company.

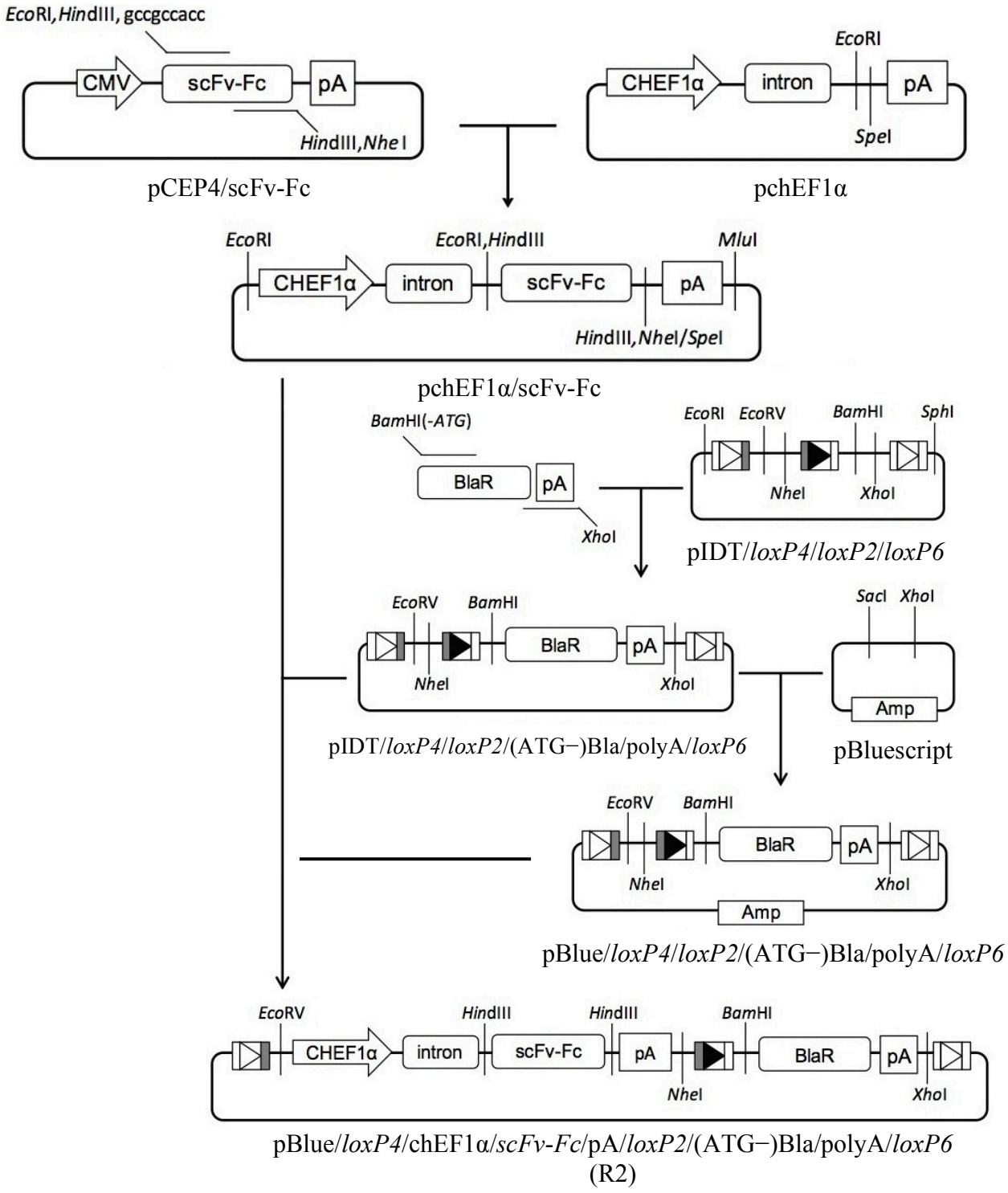


Fig. 3-3 Flowchart for the construction of R2 plasmid.

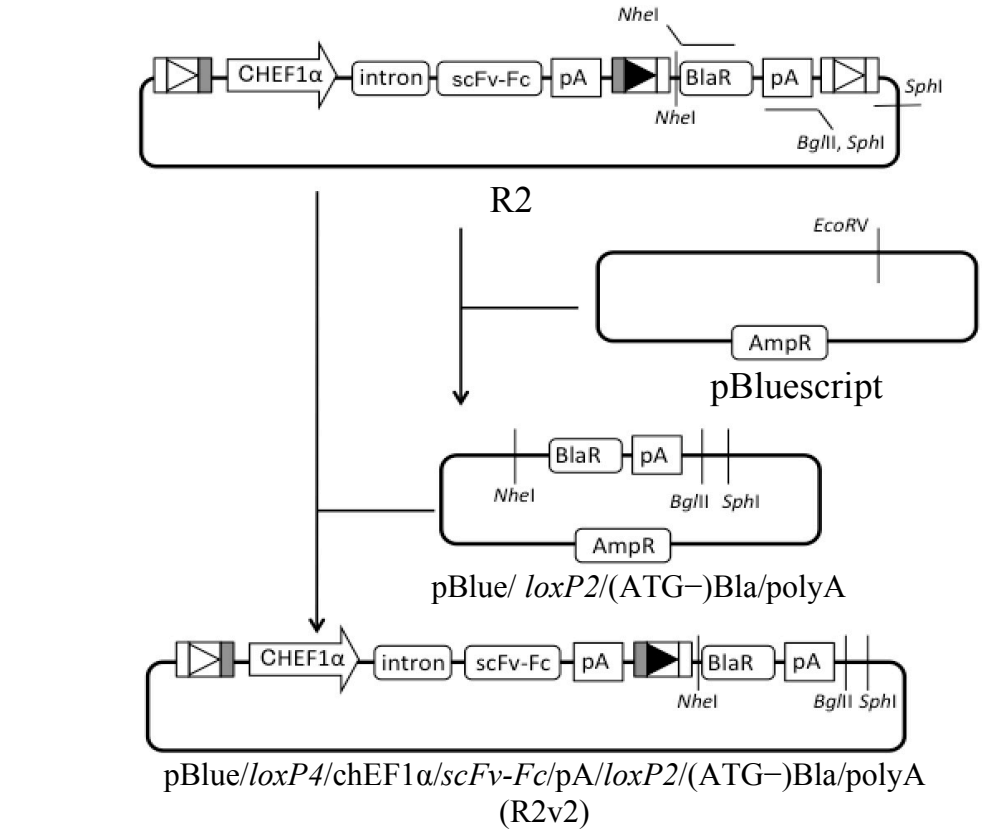


Fig. 3-4 Flowchart for the construction of R2v2 plasmid.

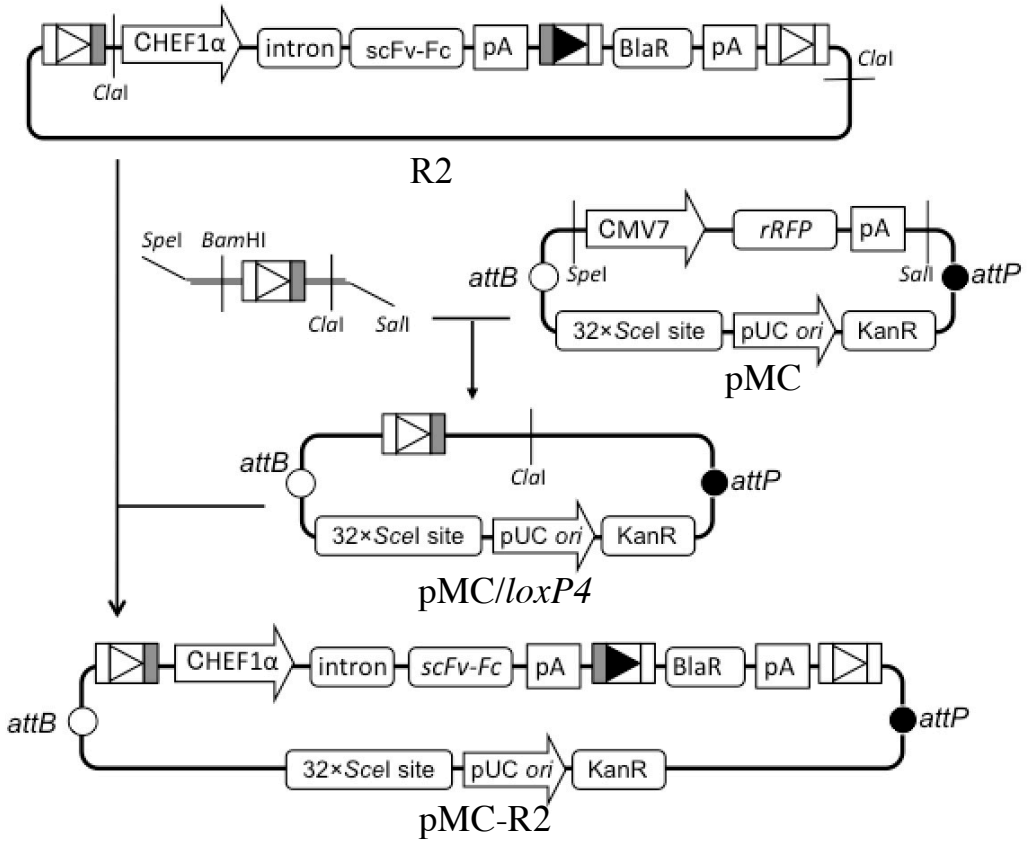


Fig. 3-5 Flowchart for the construction of pMC-R2 plasmid.

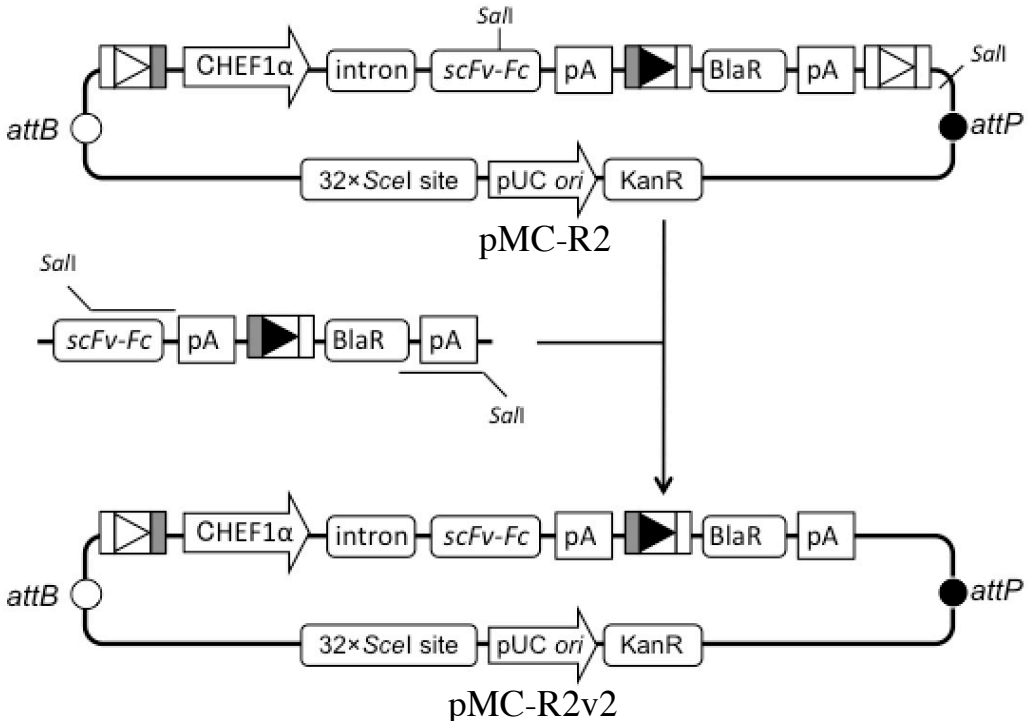


Fig. 3-6 Flowchart for the construction of pMC-R2v2 plasmid.

3.2.2.3 Minicircle preparation

Minicircles were prepared using a commercial kit (Cat. No. MN920A-1, SBI) according to the manufacturer's instructions, with some modifications. Briefly, pMC-R2 was transformed into a minicircle producer *E. coli* strain, ZYCY10P3S2T, and positive colonies were selected on LB plates containing 30 μg/mL kanamycin (Wako) after incubation overnight at 37°C. After pre-culture of clones in 2 mL of LB medium containing 30 μg/mL kanamycin at 30°C for 3 h with shaking (160 rpm), the cells were seeded into 25 mL growth medium and incubated at 30°C for 16–18 h with shaking. After the pH of culture broth was adjusted to 7.0 by adding 1 M NaOH solution, followed by mixing with an equal volume of induction medium containing arabinose, the cells were further cultured at 30°C for 5.5–6 h with shaking. A plasmid extraction kit (Qiagen, Hilden, Germany) was used to prepare the minicircle DNA, and contaminating genomic DNA and the parent plasmid were removed by digestion with a restriction enzyme (*Nde*I, Takara) and DNase (Plasmid-Safe™ ATP-Dependent DNase; Epicentre, Madison, WI, USA).

3.2.2.4 Cre-mediated transgene integration into the CHO cell genome

For Cre-mediated integration, recipient CHO cells containing (ATG−)*DsRed* and a *loxP* target site (*loxP1*) at the *hprt* locus were established as follows. The R1 plasmid was linearized by digesting with *Nru*I and transfected into CHO-T cells (derived from CHO-K1) using an electroporation device (Amaxa Nucleofecion system, Lonza, Basel, Switzerland) according to the manufacturer's protocol. The cells were selected in medium containing 500 μg/mL G418 (Thermo Fisher Scientific, Waltham, MA, USA) and 20 μM 6-thioguanine (Wako). DsRed expression was analyzed by flow cytometry (SH800 Cell Sorter, Sony, Tokyo, Japan). Integration of the target DNA fragment into the CHO cell genome by homologous recombination was confirmed by PCR. The DNA sequence of the *loxP1* site was confirmed by analyzing the PCR product (data not shown). The resulting cells, expressing DsRed, were designated as CHO/R1.

For the transgene integration, plasmids or minicircle DNA vectors were used as the donor vectors. The CHO/R1 fonder cells were transfected with donor vector using an electroporation device (Neon transfection system, Invitrogen). 1.2 μg donor vector and 0.1 μg Cre expression plasmid (pCEP4/NCre [76]) were co-transfected into 5×10^6 fonder cells, electric shocked cells were cultured and recovered in a 60-mm tissue culture dish (Thermo Fisher Scientific) for 5 days before isolating.

Target cells exhibiting the fluorescent shift from red (DsRed) to non-color were isolated using a cell sorter (SH800) and blasticidin resistance screen method.

The established clones were analyzed by quantitative real-time PCR to determine transgene copy number. scFv-Fc production was measured by ELISA, as described below.

3.2.2.5 Blasticidin resistance drug screening and limited dilution

On day 5 post-transfection, cells were seeded into 6-well plate (Thermo Fisher Scientific) with 1×10^3 cells per well and cultured with medium containing 5 μg/ml blasticidin resistant

drug (Invitrogen) for 2 weeks, then limited dilution was applied for selecting target clones, cells were seeded in 96-well plate (Thermo Fisher Scientific) with in 0.5 cell per well. After 1 more week culturing, cells were observed under microscope, single clones exhibiting extinguished fluorescence were scaled up for PCR analysis and ELISA measuring.

3.2.2.6 Genomic PCR analysis & real-time PCR analysis

Genomic DNA was extracted from cells using a commercially available kit (MagExtractor Genome; Toyobo). Regions of Cre-mediated recombination were amplified by PCR using genomic DNA (50 ng) as a template. PCR was initiated with DNA polymerase (G-Taq, Cosmo Genetech, Seoul, Korea) at 95°C for 2 min, followed by 35 cycles of amplification at 95°C for 30 s, 56–57°C for 40 s, 72°C for 15–18 s and 72°C for 5 min for final extension.

The long fragment (> 3000kb) was detected by another DNA polymerase (Tks Gflex™, Takara, Japan), initiated at 94°C for 1 min, followed by 30 cycles of amplification at 98°C for 10 s, 60°C for 15 s, 68°C for 105s for final extension. The genetic sequences of the amplicons were determined using a Prism 3130 Genetic Analyzer.

Copy numbers of the scFv-Fc expression unit were determined by real-time PCR (PikoReal96 Real-time PCR system, Thermo Fisher Scientific) as described in a previous report [77]. scFv-Fc producer CHO cells [76] possessing a single copy of the transgene in their genome, verified by Southern blotting, were used as a single copy control. The copy number values were expressed as means plus or minus the standard deviation.

The primers pairs (①-③) used are summarized in Table 3-3.

TABLE 3-3. Primers used for genomic PCR analysis.

Primer	Sequence (5′→3′)
Fw ①	TTT TGA GTT TGG AGC GAA GC
Rv ①	CCC AGC ACC ACG AGT TCT G
Fw ②	TGC TGC CCT CTG GTT ATG TG
Rv ②	CCA TCC CTG ACA CTT GTG AAA TC
Fw ③	CAG CAG GGG AAC GTC TTC TC
Rv ③	CAC CTT GAA GCG CAT GAA C

3.2.2.7 Measurement of cell growth and scFv-Fc production rates

The scFv-Fc production rate was measured as follow described method. Briefly, all the established CHO cell lines (2.5×10^4 cells/well) were seeded in 24-well tissue culture plates (Thermo Fisher Scientific) with 0.5 mL serum-containing F12 medium and cultured for 6 days. The medium was replaced with an equal volume of fresh medium every 24 h and the spent medium was retained for measuring scFv-Fc concentration. Viable cell density was determined by the trypan blue exclusion method. The IgG fraction of a rabbit anti-human IgG (Fc) (Rockland Immunochemicals, Philadelphia, PA, USA) and a rabbit peroxidase-conjugated anti-human IgG antibody (Rockland Immunochemicals) was used as primary and secondary antibodies, respectively. A human Fc fragment (Jackson Immuno Research, West Grove, PA, USA) or purified scFv-Fc [78] was used as standards to create dilution series for calibration curves. Samples were prepared in triplicate, and data were expressed as means plus or minus the standard deviation.

3.2.3 Results and Discussion

3.2.3.1 Strategy for target integration mediated by Cre recombinase and mutated loxPs

Site-specific gene integration enables precise transgene insertion into predetermined genomic sites if the unique target site was introduced into a specific chromosomal locus. As one of the frequently applied site-specific recombination, irreversible knock-in of target gene in to host cell's chromosome mediated by Cre-catalyze *loxPs* was reported [79]. Our laboratory has previously study on a series of mutated *loxPs* to improve the integration efficiency [80], two *loxPs* pairs showed the highest efficiency were used to introduce the target gene and delete the backbone in current research. A CHO cell line (CHO/R1 cells) possessing a single arm-mutated *loxP* site for transgene integration was used as the host. It was obtained by introducing the mutated *loxP* transgene integration site into the *hprt* locus of the CHO cell genome via homologous recombination in a previous study, and though Cre mediated target gene integration using the high efficiency *loxPs* mentioned above, a satisfied target protein expression level of this locus and was observed [81]. Thus, this system can be used to evaluate and judge the transfection efficiencies of different vectors. In present study, target cells were selected using resistance screening associated with fluorescence loss resulting from Cre-mediated genomic recombination. Fig. 3-7 shows the schematic drawing of integrations of different transgenes. CHO/R1 cells carrying a *DsRed* expression unit at the *hprt* locus constitutively express the red fluorescent protein. After Cre mediated integration, the blasticidin resistance gene was integrated into the downstream of chEF1α promoter, hence the cell lost fluorescence and became resistant to blasticidin. Simultaneously, a scFv-Fc expression unit was introduced into the same locus for producing recombinant proteins. Unlike integration of R2v2 and mcR2v2, Cre-mediated knock-out events happened after the integrations of R2 and mcR2v2, resulting in partially integration of the vector and elimination of the backbone sequences.

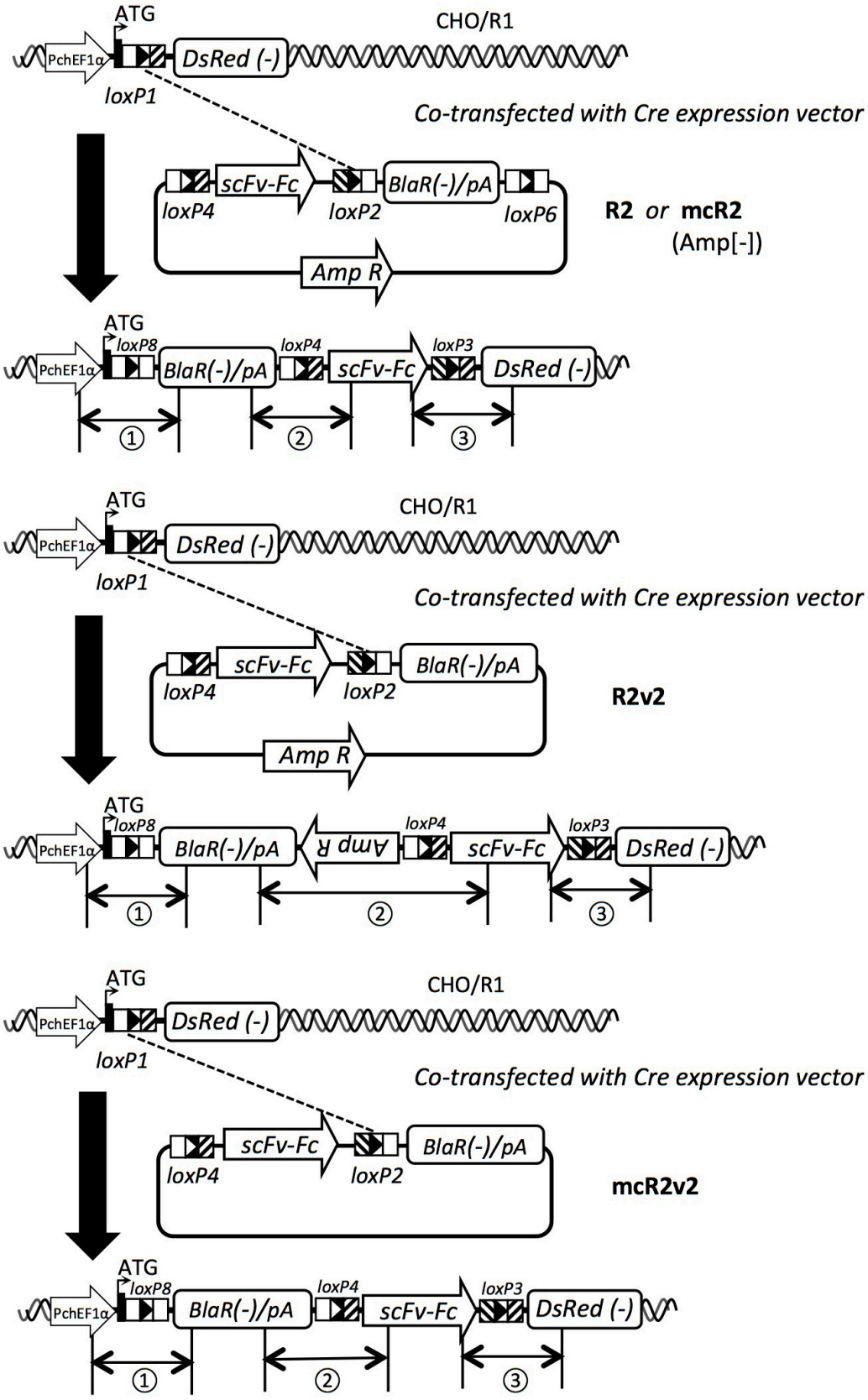


Fig. 3-7 Schematic drawing of integrations of different transgene.

3.2.3.2 Transgene integration into the predetermined target site

To obtain CHO/B[scFv-Fc] cells, the integration process was initiated by co-transfection of donor plasmids and the Cre expression plasmid into CHO/R1 cells (Fig. 3-7). At 5 days post-transfection, cells were analysis by FACS for checking the transfection efficiencies. Obviously, minicircle vectors showed about 3-fold higher efficiencies than conventional plasmids (Fig. 3-8A) due to the smaller molecular size (Fig. 3-8B). Simultaneously, post-transfected cells were seeded into 6-well tissue culture plates and cultured by medium containing blasticidin drug. After 2 weeks cultivation, the survived cells were isolated by limited dilution method, target clones lost red fluorescence and further analyzed by FACS (Fig. 3-9). Selected target clones were verified by genomic PCR analysis.

Genomic PCR was performed for detection of Cre-meditated site-specific integration (Fig. 3-10A) using various primer sets to validate the recombination reactions (annotated in Fig. 3-7). The expected sizes of DNA fragments were amplified for transfected CHO/R1 cells using the primer pairs ①-③ (Table. 3-3). Sequence analysis of the amplified DNA fragments confirmed that the transgenes were integrated at the target site pre-introduced in the *hprt* locus (Fig. 3-10B). These results indicated that Cre-recombination were processed as schematic drawing predicted, eukaryotic transcription cassettes from each vector have successfully integrated into the *hprt* locus.

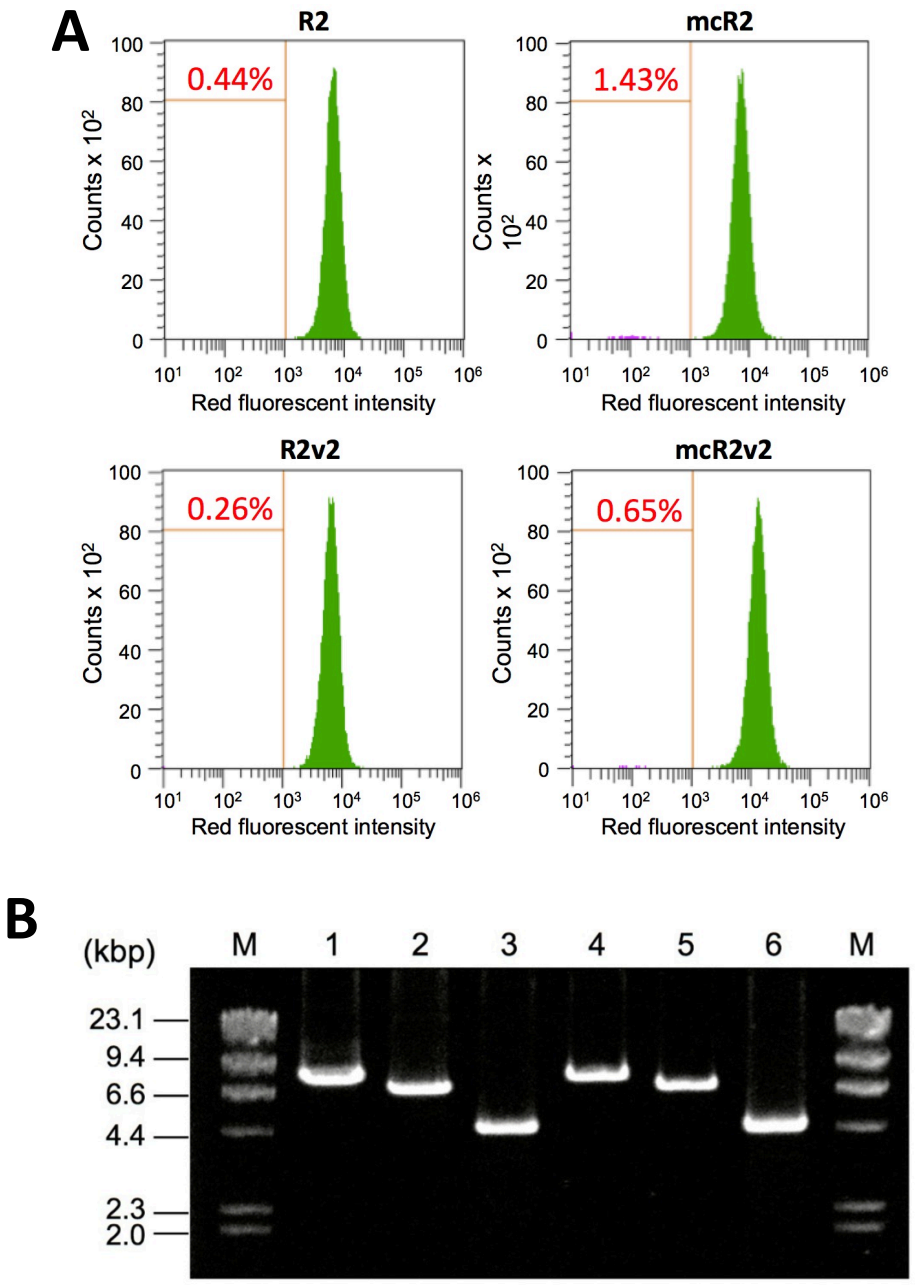


Fig. 3-8 Operation of different DNA vectors for Cre-mediated integrations.

(A) Electrophoresis analysis of the vectors. Linearized plasmid and minicircle vectors were electrophoresed on a 1% (w/v) agarose gel. Lane M, molecular weight markers (λ-HindIII digest); Lane 1, pMC-R2 (8,688 bp); Lane 2, R2 (7,540 bp); Lane 3, mcR2 (4,650 bp); Lane 4, pMC-R2v2 (8,624 bp); Lane 5, R2v2 (7,506 bp); Lane 6, mcR2v2 (4,586 bp).

(B) Comparison of the transfection efficiencies of R2, mcR2, R2v2 and mcR2v2 for the transgene integration using Cre recombination. The total numbers of cells with a color change from red to non-color were counted using FACS.

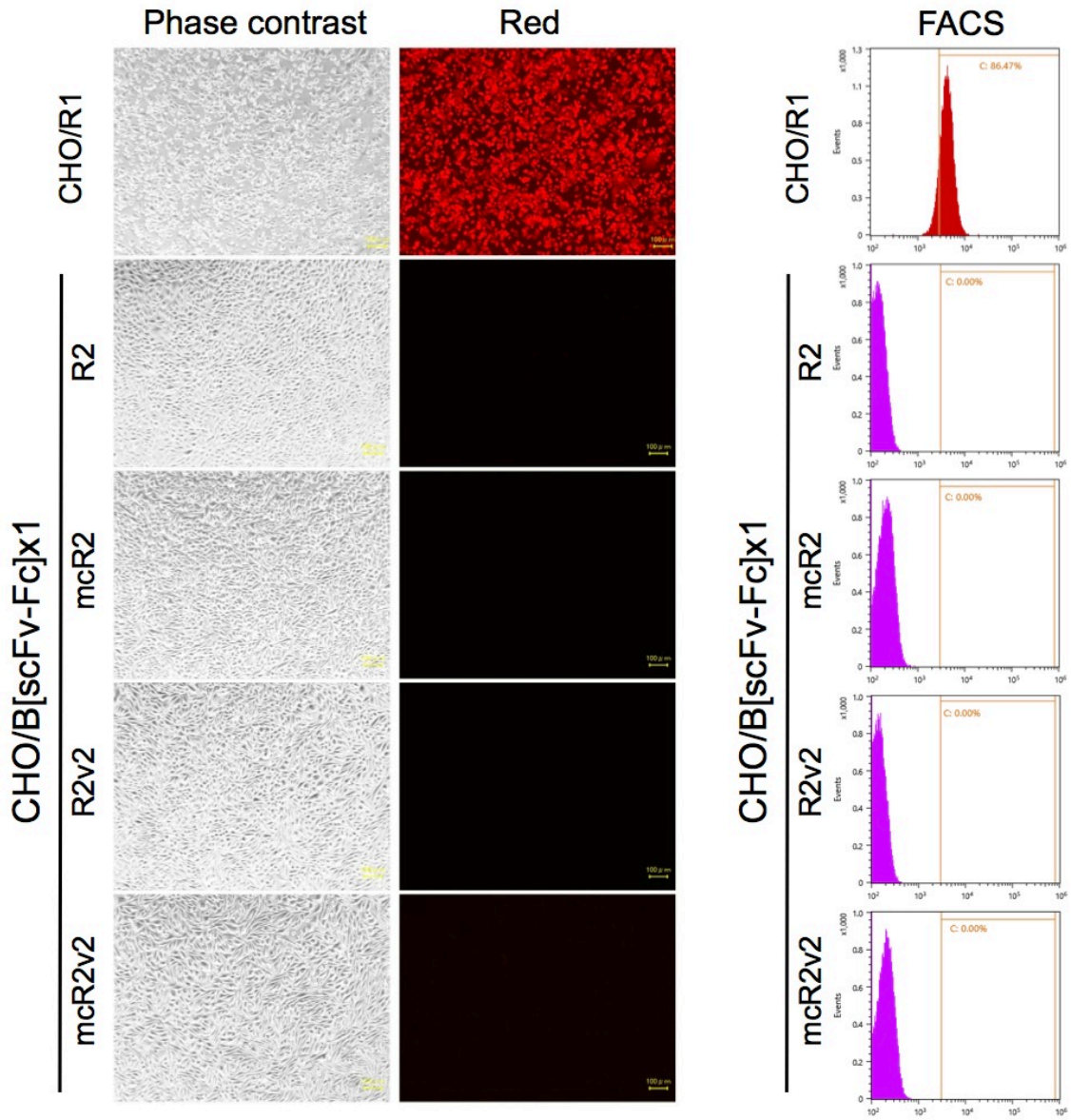


Fig. 3-9 Reporter fluorescence eliminated after target genes introduced into predetermined *loxP* sites in CHO cell clones. Scale bars indicate 100μm.

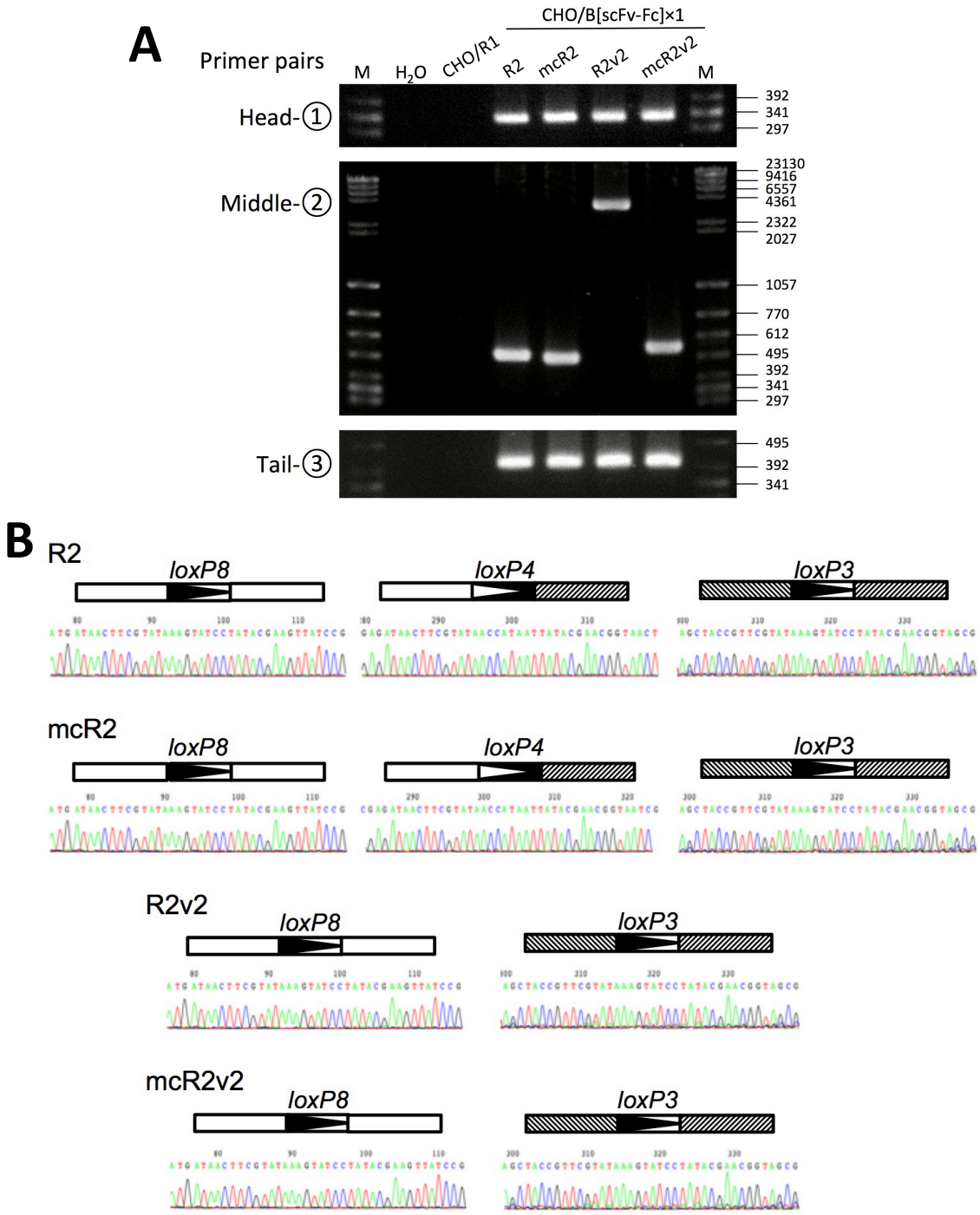


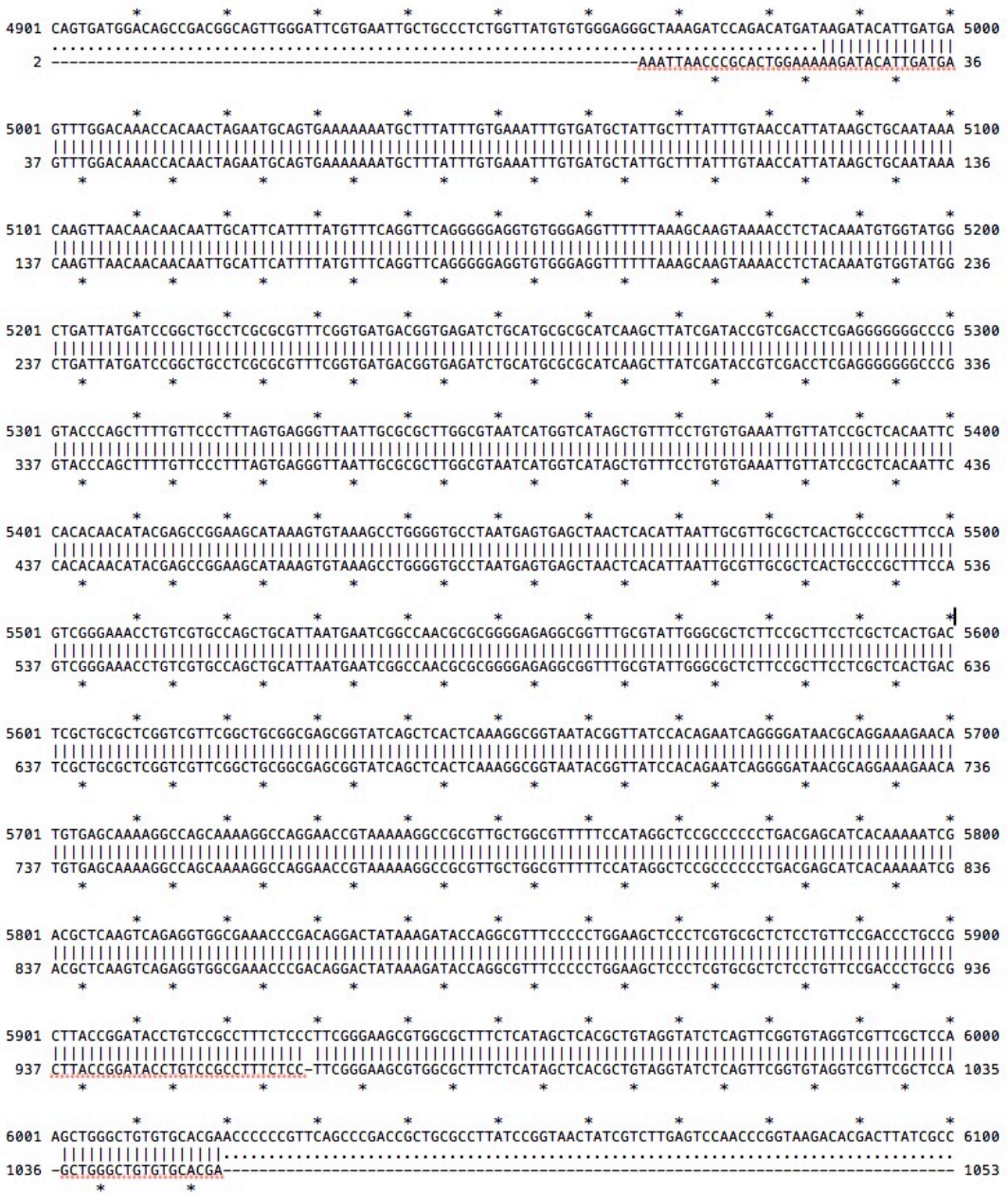
Fig. 3-10 Genome structure analyses of generated CHO cell clones.

(A) PCR analysis of the clones' genomic DNA using specific primers (Table 3-3). Annealing sites of the primers in the transgene are depicted by arrows in Fig. 3-7. Lane M, molecular weight standard markers (mix of λ-*Hin*dIII and ΦX174-*Hin*dII digests); Lane H_2O, negative control with water as the template.

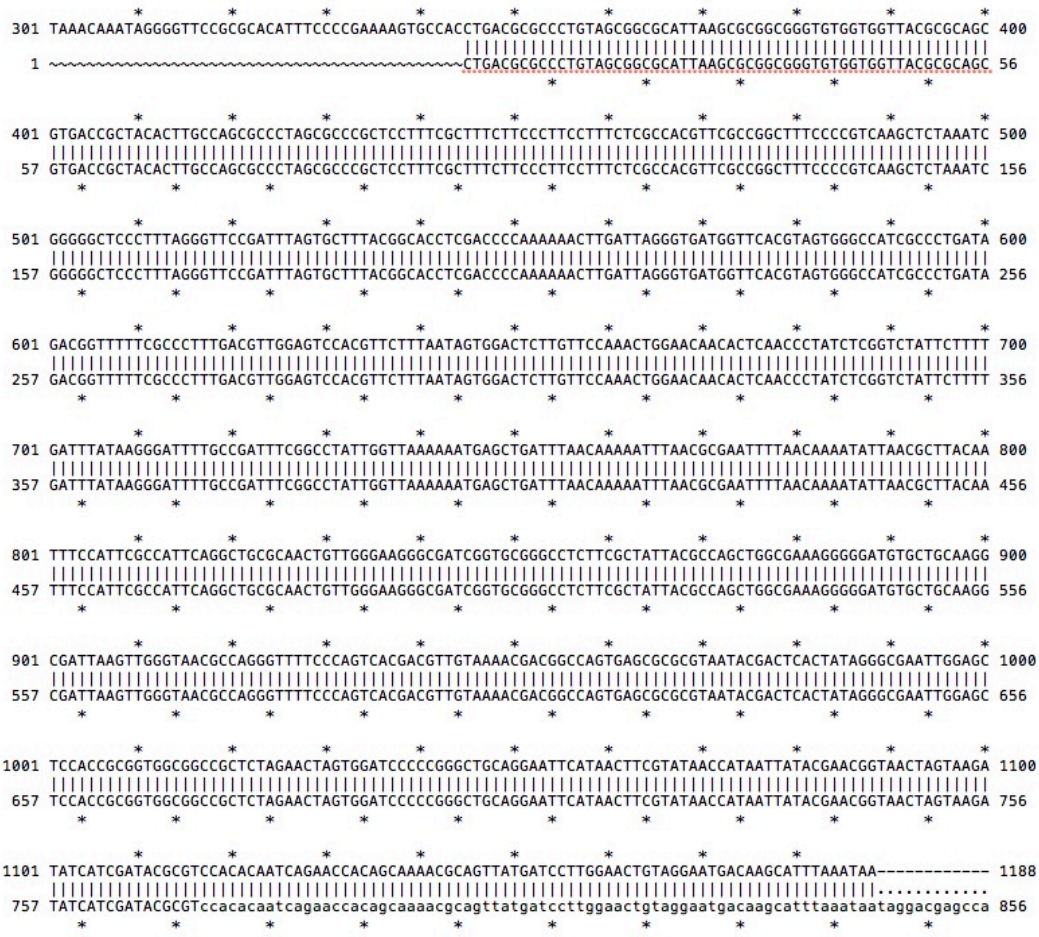
(B) DNA sequence analysis of recombined sites. Sequences of the PCR products shown in Fig. 3-10A were analyzed, and the sequences of the *loxP* sites associated with the recombination reactions are shown.

Cre recombination just happened once in the cases of R2v2 and mcR2v2 inserted clones. The regions surrounding *loxP4* (② in Fig. 3-7) were amplified by genomic PCR and aligned to vector R2v2 and mcR2v2, respectively. The following DNA sequence alignment maps showed good gene comparison results, which indicated the entire integration of these two vectors in obtained CHO/B[scFv-Fc]×1 cells.

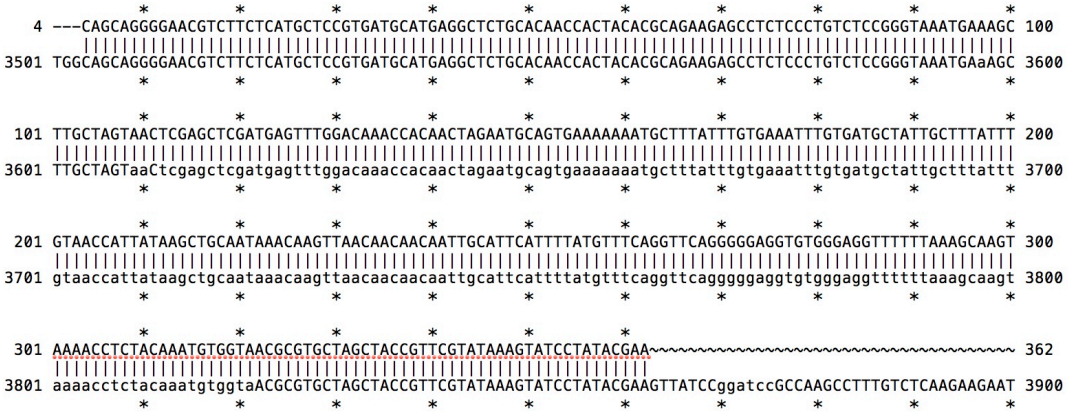
Forward sequences alignment of fragment② derived from R2v2 integrated CHO/B[scFv-Fc]×1 cells.



Reverse sequences alignment of fragment② derived from R2v2 integrated CHO/B[scFv-Fc]×1 cells.



Sequences alignment of fragment② derived from mcR2v2 integrated CHO/B[scFv-Fc]×1 cells.



3.2.3.3 Comparison of target genes integration efficiency and accuracy of CHO clones derived from different plasmids and minicircle vectors

After the limited dilution, 50 clones of each kind of target cells were selected for applying genomic PCR and checking the ability of producing scFv-Fc by ELISA method. The results of different proportion are shown in Table 3-4. Both types of minicircles vectors significantly showed better integration performance than their correspondingly plasmid vectors.

TABLE 3-4. Statistic data for integration results of four kinds of vectors

Vector type	scFv-Fc (+)	scFv-Fc(+) G-PCR(+)	scFv-Fc(+) G-PCR(-)	scFv-Fc(-) G-PCR(+)	scFv-Fc(-) G-PCR(-)	Dead/Alive clone No.
R2	28/50	22/44	3/44	0/44	19/44	6/44
mcR2	49/50	45/47	2/47	0/47	0/47	3/47
R2v2	25/50	21/43	0/43	4/43	18/43	7/43
mcR2v2	40/50	37/46	2/46	0/46	7/46	4/46

According to Table 3-4, almost all the mcR2 derived clones can produce target protein and 45 of 47 collected clones showed target proteins expression, while just half of the R2 derived clones showed target proteins producing and predetermined structures. This phenomenon was similar to mcR2v2 and R2v2, the appeared probability of targeted integration of minicircle DNA vectors are 2-fold of plasmids targeted integration efficiency. Thus, higher integration efficiencies of minicricle vectors favored the antibody transgene integration and expression. Minicircle vectors transfected cells can be used as recombinant producer cells without isolation and scale-up processes, which saved times and efforts.

Four clones derived from R2v2 showed right integration structure but without scFv-Fc producing, which is different from other types of clones. It is indicated that the inserted backbone sequences occasionally caused the silencing of neighboring genes. The gene silencing is commonly happened in transgene expression processes, in particular the involved

eukaryotic promoter, which caused by promoter interference. In addition, some reports discovered that the DNA length could affect the transgene expression [82], and covalent linkage of transcript unit to bacteria backbone could cause transgene silencing [83].

As shown in the table, some clones died during scale-up due to cell apoptosis. Transfection of vectors contain backbone sequences have a higher probability to undergo apoptosis, 6 and 7 clones were died when using R2 and R2v2 plasmids respectively, while 3 and 4 in the case of mcR2 and mcR2v2 vectors. In prokaryotes genome, the CpG motifs occur 10-20 times more than in mammalian cells genome, DNA sequence containing unbalanced unmethylated CpG motifs could initiate immune response and inflammatory reactions, results in apoptosis-mediated cell death after transduction of transgenes [84-85], that is why plasmids containing bacteria backbone sequences caused a relatively high probability of apoptosis in current study.

3.2.3.4 Analysis of scFv-Fc production by Cre recombinant CHO cells

Transgene copy number and scFv-Fc productivity were measured for 15 randomly selected target clones generated in each kind of integration. As determined by quantitative real-time PCR, most of the clones showed single copy of scFv-Fc transgene as expected (Fig. 3-11A).

The scFv-Fc producer cells and the founder cells (CHO/R1) were cultured for 6 days to analyze cell growth and scFv-Fc productivity. There were no significant differences in target gene expression level (Fig. 3-11C), the scFv-Fc productivity of all types of CHO/B[scFv-Fc]×1 cells are about 10 pg/(cell·day), these results indicated that the scFv-Fc productivity in *hprt* locus is stable. Although the productivity was not significant different between each types of cells, the clones established by minicircle vectors showed a better growth rate (Fig. 3-11B), it is possibly due to promoter interference [89]. Thus, the clones established using minicircles showed a better scFv-Fc titer (Fig. 3-11D).

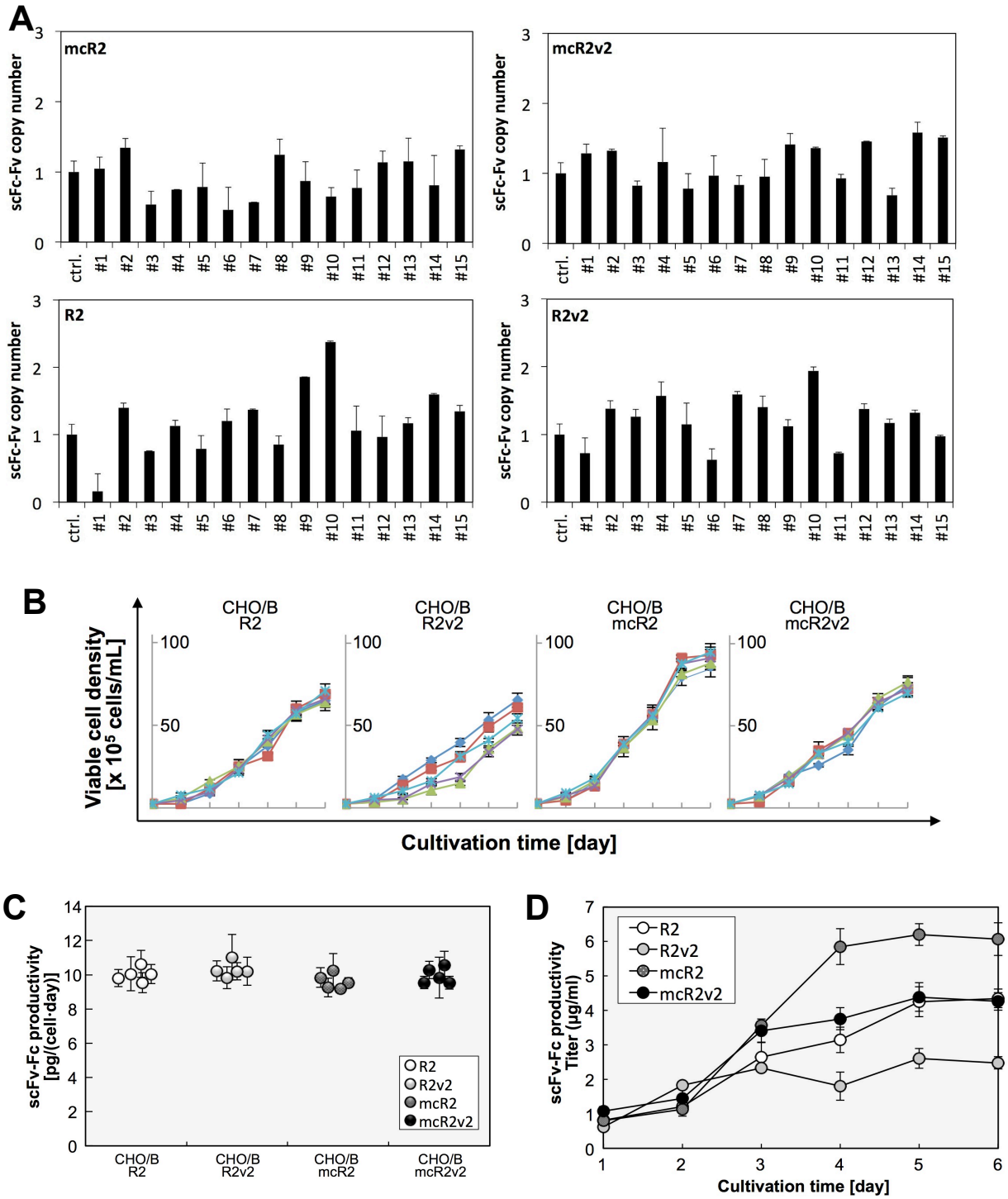


Fig. 3-11 Analysis of CHO cell clones generated from different vectors. Data are the means ± standard deviations of triplicate experiments.

(A) Relative scFv-Fc copy number.

(B) Cell growth trajectories.

(C) Specific scFv-Fc productivity.

(D) scFv-Fc titer during cultivation.

3.3 Conclusion

In the current study, four plasmid and minicircle vectors were used to introduce a blasticidin gene and a scFv-Fc expression unit into founder CHO/R1 cell genome. After targeted integration, the target cells were selected by blasticidin resistance screening facilitated by microscope observing of faded fluorescence. Half of obtained clones from plasmids transfected clones showed unexpected structures and transgene silencing, while almost all the minicircle transfected target clones showed expected structure and expressed antibody scFv-Fc. Thus, bacterial backbone could interfere the recombinant efficiency, no matter it is integrated into the host genome or not. The transfection efficiency was improved 3-fold using minicircle DNA vector, when detected by FACS device.

Target clones generated Cre-mediated integration of different vectors showed the similar specific productivity about 10 pg/(cell·day), indicated that the *hprt* locus could be a selectable hot spot sit for transgenes expression. However, the cells grow rates were affected by bacterial backbone sequences significantly, the cells showed slow growth rates when carrying the bacterial backbone sequences. The clones generated using minicircle vectors showed significantly higher growth rate, the antibody titers of these clones were also better.

3.4 Summary

The best specific productivity of recombinant antibodies under optimized culture conditions is currently 50–90 pg/(cell·day) in high-producer cells with high copies of transgenes generated by the conventional gene amplification procedure [72]. In this study, scFv-Fc productivity reached 10 pg/(cell·day) possessing only one transgene expression unit, which proved that *hprt* locus is available for commercial application. Moreover, the discovery of minicircle DNA improved target cells productivity base on site-specific integration favored the manufacture of recombinant proteins.

3.5 Prospect

According to the results, minicricle DNA vectors lacking the bacterial backbone sequences improved the targeted integration efficiency, and favored the growth rate of selected clone. This form of vector will be applied to newly designed AGIS system described in this thesis. It is recommendable to use minicircle DNA vectors for generating producer cell lines of pharmaceutical recombinant proteins.

Chapter 4

Accumulative scFv-Fc antibody gene integration into the hprt chromosomal locus of CHO cells

4.1 Introduction

Previously, an accumulative site-specific gene integration system (AGIS) based on the Cre-recombinase/*loxP* system, using mutated *loxP* sites was developed in our laboratory [86]. AGIS can provide a simple and efficient method for repeated integration of transgenes into a predetermined chromosomal locus. This method has been applied for the generation of recombinant CHO cells for producing antibodies [76]. Within the locus of the house-keeping gene, *hypoxanthine phosphoribosyltransferase (hprt)*. *Hprt* is known to be selectable hot spot for transfect gene expression [87]. Koyama et al. reported that stable transgene expression was observed throughout 129 days' cultivation when a transgene was integrated into *hprt* locus in mammalian cells [71].

In this thesis, a CHO cell line was generated, in which a *loxP* target site for AGIS was introduced into the *hprt* locus by homologous recombination. This CHO cell line was used as a founder for AGIS. To facilitate and accelerate the generation of high-producer cells, two copies of transgenes were introduced into the *hprt* locus in single transfection procedure. To improve the screening process, CHO cells in which the transgene copies were integrated at the *hprt* locus were separated based on a shift in the color of their fluorescence that was mediated by a change in the screening marker they expressed. Furthermore, minicircle DNA vectors lacking a bacterial backbone were used to improve transfection efficiency. These improvements in AGIS are effective for generating recombinant CHO cells with multiple

transgenes at a predetermined locus, and that AGIS is a useful tool for establishing producer cells for biopharmaceutical proteins.

Mutated *loxP* pairs showing high integration efficiency have been screened for AGIS [80]. However, the process of establishing cells with multiple integrated copies of the transgene is still time-consuming. This chapter described an improved version of AGIS that facilitates and accelerates the establishment of high-producer Chinese hamster ovary (CHO) cells. Two donor vectors were simultaneously introduced into the cells in a single transfection. Cells with successfully targeted transgene integration were screened based on a change in the color of the reporter fluorescent protein that they express. Repeated rounds of integration allowed the transgene copy number to be increased. As a model, an scFv-Fc antibody gene was integrated into the *hprt* locus of the CHO cell genome. After three rounds of integration, a high-producer CHO cell clone with six copies of the scFv-Fc gene was successfully established. scFv-Fc productivity was approximately four-fold greater than a control cell line harboring a single copy of the transgene. This newly designed AGIS procedure should facilitate the development of producer cells suitable for biopharmaceutical protein production.

4.2 Strategy for accelerated integration and screening processes in AGIS

In previously developed AGIS [88], a CHO cell line possessing a single arm-mutated *loxP* site for transgene integration was used as the founder (as shown in Fig. 4-1). In the present study, it was further modified by introducing a mutated *loxP* transgene integration site into the *hprt* locus of the CHO cell genome by homologous recombination (CHO/R1 cells). Thus, targeted multiple transgenes integration in *hprt* locus followed by high-level transgene expression using this improved version of AGIS is expected. To facilitate and accelerate the integration process, two transgene donor vectors were introduced into cells by a single transfection, whereas only a single transgene was sequentially introduced using the previous AGIS procedure [88]. Targeted cells were screened using a fluorescence-activated cell-

sorting (FACS) device, based on the color change of reporter fluorescent proteins. This accelerated the screening process compared with drug resistance-mediated screening.

Fig. 4-2 shows a schematic presentation of each round of transgene integration. scFv-Fc antibody expression units were repeatedly introduced into the mutated *loxP* site located between the chEF1α promoter and the *DsRed* or *EGFP* gene. CHO/R1 cells carrying a *DsRed* expression unit at the *hprt* locus constitutively express the red fluorescent protein. For the first integration round, two donor plasmids carrying scFv-Fc expression units, R2 and R3-Green, were co-transfected into CHO/R1 cells together with the Cre expression vector. After targeted integration of the donor vectors, the *EGFP* gene was integrated downstream the chEF1α promoter, and hence the cell color changed from red to green. In the resultant cells (CHO/G[scFv-Fc]×2), two scFv-Fc expression units had been integrated into the genome in a single transfection (Fig. 4-2). For the second integration round, another two donor vectors, mcR2 (or R2) and R3-Red, were similarly co-transfected into CHO/G[scFv-Fc]×2 cells together with the Cre expression vector. The targeted cells (CHO/R[scFv-Fc]×4), in which four scFv-Fc expression units were integrated into the *hprt* locus, could be screened by a color change from green to red. The integration process was repeated three times in this study, producing CHO cells with six scFv-Fc expression units integrated into the *hprt* locus (CHO/G[scFv-Fc]×6). This integration and screening strategy greatly reduced the time required to establish cells with multiple transgene copies, compared with earlier versions of AGIS.

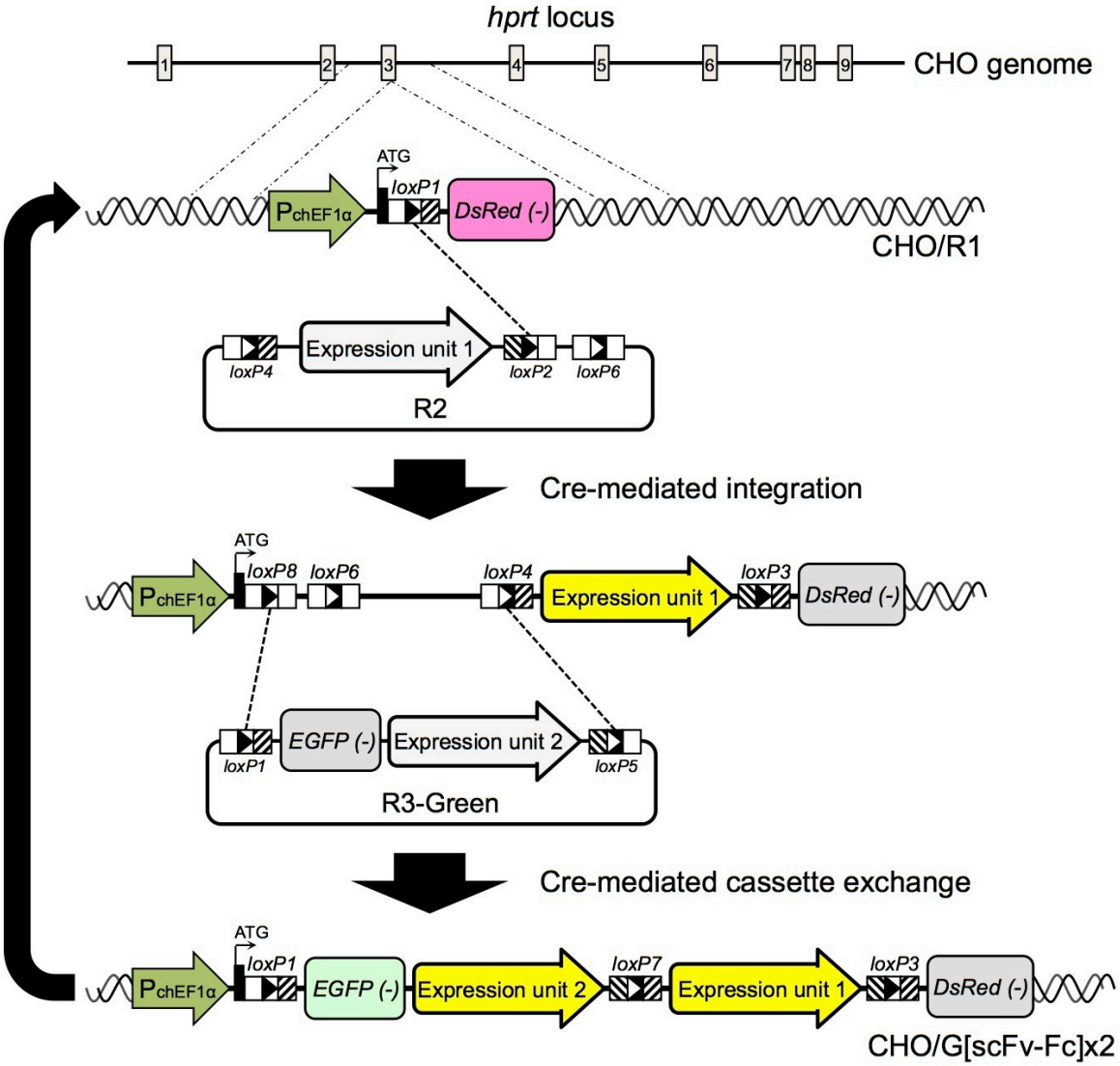


Fig. 4-1 Schematic drawing of expected Cre-mediated transgene integration processes using two donor plasmids in the first integration round.

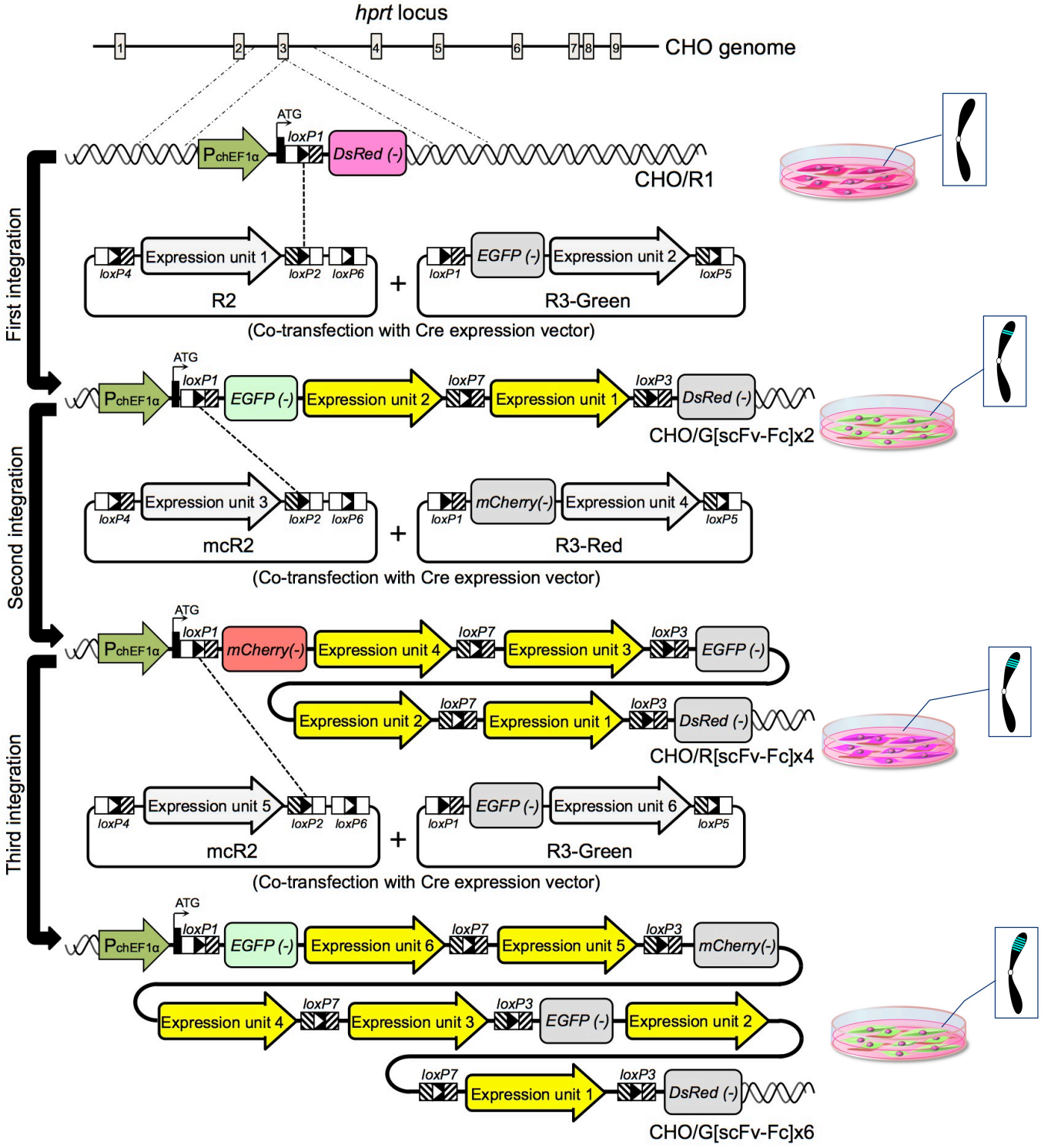


Fig. 4-2 Schematic drawing of the three transgenes integration rounds included in Cre-mediated AGIS.

4.3 First round of transgene integration using two donor vectors

4.3.1 Experimental Purpose

In the AGIS system, two donor plasmids were transfected in to the fonder cells with cre expression vector only in one step. Thus, to ensure the target genes integration is very important. In these experiment, the first round of transfection was applied to confirm the practicability of the co-transfection, the predetermined integration of two copies of scFv-Fc expression units mediated by mutated Cre/*loxP* recombination were verified.

For antibody expression cells, the target gene expression level and cells growth ability are most important of all, single cell expression level and cells growth status could decide the products yield by affect the antibody titer in the collected raw medium, which is related to the industrial economic benefit. In this section, the expression level of scFv-Fc also mediated to confirm the antibody producing ability of the clones established by using first round of AGIS, and the cells growth curve also observed.

4.3.2 Materials and Methods

4.3.2.1 Cells and culture media

The founder cells used for transgene integration were CHO/R1 cells, in which a mutated *loxP* site (*loxP1*) and an expression cassette encoding the red fluorescent protein, DsRed, were introduced into the *hprt* locus (described in Chapter 3). Cells were cultured under adherent conditions using Ham's F12 medium (Sigma-Aldrich, St. Louis, MO, USA) supplemented with 10% fetal bovine serum (Biowest, Nuaillé, France), 100 units/mL streptomycin sulfate and 90 μg/mL penicillin G potassium (Wako Pure Chemical Industries, Osaka, Japan) at 37°C and 5% CO_2 in a humidified incubator.

4.3.2.2 Plasmid construction

A double-stranded DNA oligonucleotide, 5′-GAG CTC ATA ACT TCG TAT AAA GTA TCC TAT ACG AAC GGT AGC GGA TCC AAG CTT ACT AGT GAT ATC ATC GAT ACG CGT GCT AGC TAC CGT TCG TAT AAC CAT AAT TAT ACG AAG TTA TCT CGA G-3′, containing two mutated *loxP* sites, *loxP1* and *loxP5*, was chemically synthesized (Medical & Biological Laboratories) and ligated into the plasmid, pIDTSMART-AMP, to generate pIDT/*loxP1*/*loxP5*. A DNA fragment incorporating an enhanced green fluorescence protein (EGFP) gene lacking the original ATG codon and the SV40 polyA signal region was amplified by PCR from pIRES-EGFP [80], using primers Pr1 and Pr2. The PCR product was digested with the relevant restriction enzymes and ligated into *Bam*HI- and *Spe*I-digested pIDT/*loxP1*/*loxP5* to generate pIDT/*loxP1*/(ATG–) *EGFP*/polyA/*loxP5*. A DNA fragment containing (ATG–)*EGFP* and a poly A signal flanked by *loxP1* and *loxP5*, obtained by digestion with *Sac*I and *Xho*I from pIDT/*loxP1*/(ATG–)*EGFP*/polyA/*loxP5*, was ligated into *Sac*I- and *Xho*I-digested pBluescript to generate pBlue/*loxP1*/(ATG–)*EGFP*/polyA/*loxP5*. A DNA fragment encoding an scFv-Fc expression unit (chEF1α/*scFv-Fc*/pA), prepared from pchEF1α/scFvFc, was ligated into *Eco*RV- and *Nhe*I-digested pBlue/*loxP1*/(ATG–)*EGFP*/ polyA/*loxP5* to generate pBlue/*loxP1*/(ATG–) *EGFP*/polyA/chEF1α/*scFv-Fc*/pA/*loxP5* (R3-Green).

The blunt-ended DNA fragments were prepared using a kit (DNA blunting kit, Takara). All PCR reactions were performed using KOD plus neo DNA polymerase (Toyobo, Tsuruga, Japan) according to the manufacturer's instructions. All DNA sequences derived from chemically synthesized oligonucleotides and PCR products were confirmed by DNA sequencing on a Prism 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

Construction flowchart for plasmid R2 was shown in materials and method part of Chapter 3 in this thesis.

The primer pairs used for vector construction are summarized in Table 4-1.

The flowchart for the construction of R3-Green is given in Fig. 4-3.

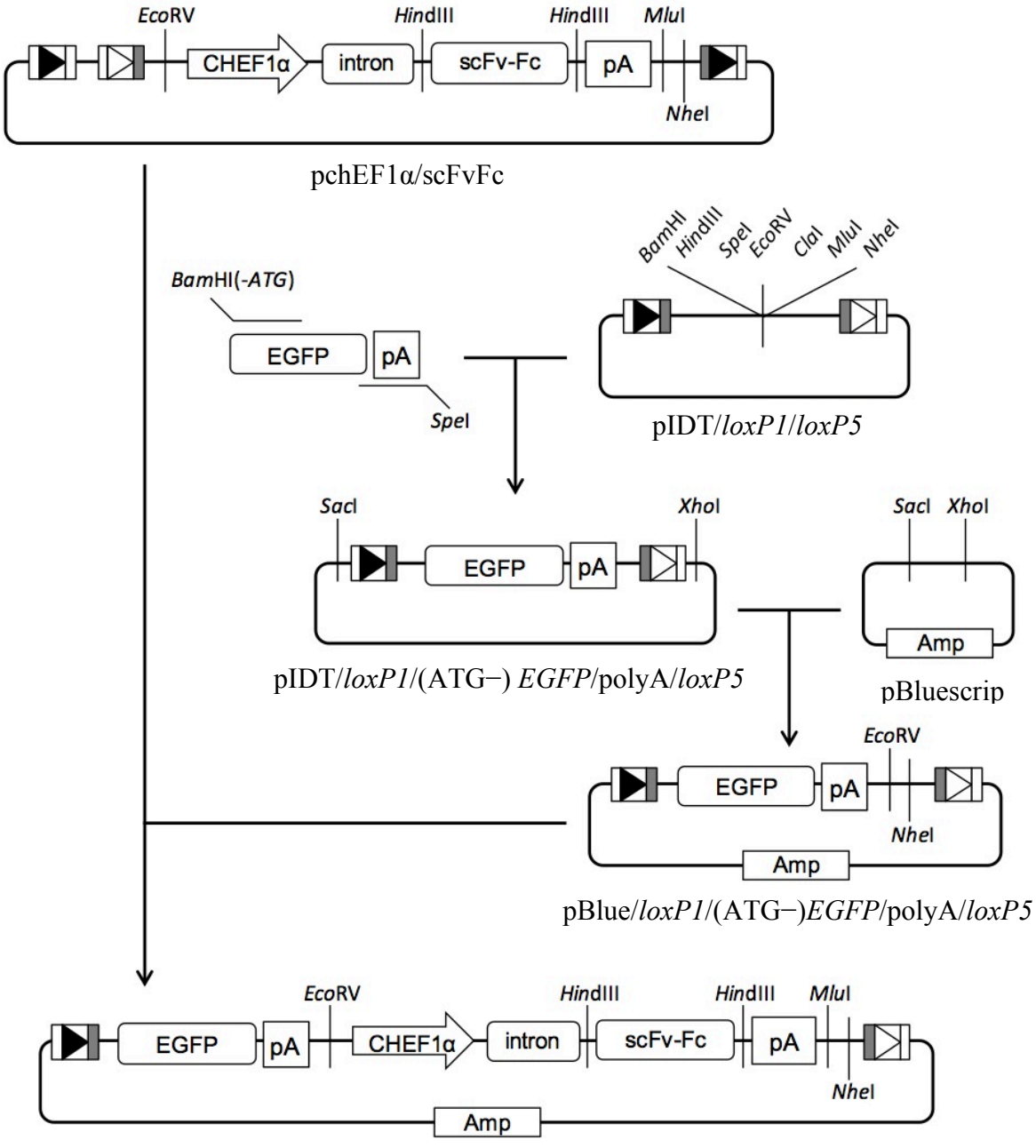


Fig. 4-3 Flowchart for the construction of R3-Green plasmid.

TABLE 4-1. Primers for vectors construction.

Primer	Sequence (5'→3')
Pr1	AAG GAT CCT CTA GAG TGA GCA AGG GCG AGG AG
Pr2	GGA CTA GTG GGA TTT TGC CGA TTT CG

4.3.2.3 *LoxP* sequence used in first round of integration

The DNA sequences of the mutated *loxp* sites used in this experiment are shown in Table 4-2.

TABLE 4-2. Sequences of wild-type and mutated *loxP* sites.

Symbol	Sequence (5′→3′)		
	Left-arm region	Spacer region	Right-arm region
loxP	ATAACTTCGTATA	ATGTATGC	TATACGAAGTTAT
loxP1	ATAACTTCGTATA	AAGTATCC	TATACGAACGGTA
loxP2	TACCGTTCGTATA	AAGTATCC	TATACGAAGTTAT
loxP3	TACCGTTCGTATA	AAGTATCC	TATACGAACGGTA
loxP4	ATAACTTCGTATA	ACCATAAT	TATACGAACGGTA
loxP5	TACCGTTCGTATA	ACCATAAT	TATACGAAGTTAT
loxP6	ATAACTTCGTATA	ACCATAAT	TATACGAAGTTAT
loxP7	TACCGTTCGTATA	ACCATAAT	TATACGAACGGTA
loxP8	ATAACTTCGTATA	AAGTATCC	TATACGAAGTTAT

Mutated sequences are underlined.

4.3.2.4 Cre-mediated accumulative gene integration into the CHO cell genome

For Cre-mediated AGIS, recipient CHO cells containing (ATG–)*DsRed* and a *loxP* target site (*loxP1*) at the *hprt* locus were established as follows. The R1 plasmid was linearized by digesting with *Nru*I and transfected into CHO-T cells (derived from CHO-K1) using an electroporation device (Amaxa Nucleofecion system, Lonza, Basel, Switzerland) according to the manufacturer's protocol. The cells were selected in medium containing 500 μg/mL G418 (Thermo Fisher Scientific, Waltham, MA, USA) and 20 μM 6-thioguanine (Wako). DsRed expression was analyzed by flow cytometry (SH800 Cell Sorter, Sony, Tokyo, Japan). Integration of the target DNA fragment into the CHO cell genome by homologous recombination was confirmed by PCR. The DNA sequence of the *loxP1* site was confirmed by analyzing the PCR product (data not shown). The resulting cells, expressing DsRed, were designated as CHO/R1.

For the first round of transgene integration using Cre-mediated AGIS, CHO/R1 cells were seeded at a density of 1.2×10^6 cells in a 60-mm culture dish (Thermo Fisher Scientific). On the next day, 4 μg of donor plasmid R2 was premixed with 4 μg of donor plasmid R3-Green and co-transfected into the CHO/R1 cells together with 0.2 μg of Cre expression plasmid (pCEP4/NCre) (7) using lipofection reagent (Lipofectamine2000, Invitrogen). After 48 h, recombinant targeted cells exhibiting a fluorescence shift from red (DsRed) to green (EGFP) were sorted into 96-well tissue culture plates (Thermo Fisher Scientific) using a cell sorter (SH800), for single cell cloning. The transgene integration site and genomic structure of screened cells were confirmed by PCR, and the clones with double copies of the scFv-Fc expression unit at the *hprt* locus were designated as CHO/G[scFv-Fc]×2. The cells harboring a single copy of the scFv-Fc expression unit using R2 as a donor plasmid (CHO/B[scFv-Fc]×1) was also cloned.

4.3.2.5 Genomic PCR analysis & real-time PCR analysis

Genomic DNA was extracted from cells using a commercially available kit (MagExtractor Genome; Toyobo). Regions of Cre-mediated recombination were amplified by PCR using genomic DNA (50 ng) as a template. PCR was initiated with DNA polymerase (G-Taq, Cosmo Genetech, Seoul, Korea) at 95°C for 2 min, followed by 35 cycles of amplification at 95°C for 30 s, 56–57°C for 40 s, 72°C for 15–45 s and 72°C for 5 min for final extension. The genetic sequences of the amplicons were determined using a Prism 3130 Genetic Analyzer.

Copy numbers of the scFv-Fc expression unit were determined by real-time PCR (PikoReal96 Real-time PCR system, Thermo Fisher Scientific) as described in a previous report [77]. scFv-Fc producer CHO cells (7) possessing a single copy of the transgene in their genome, verified by Southern blotting, were used as a single copy control. The copy number values were expressed as means plus or minus the standard deviation.

The primers (α–ι) used are summarized in Table 4-3.

TABLE 4-3. Primers used for genomic PCR analysis.

Primer	Sequence (5′→3′)
β	CAC CTT GAA GCG CAT GAA C
γ	TTT TGA GTT TGG AGC GAA GC
δ	CCC AGC ACC ACG AGT TCT G
ε	TGC TGC CCT CTG GTT ATG TG
ζ	CCA TCC CTG ACA CTT GTG AAA TC
η	CAG CAG GGG AAC GTC TTC TC

4.3.2.6 Measurement of cell growth and scFv-Fc production rates

The scFv-Fc production rate was measured as follow described. Briefly, all the established CHO cell lines (2.5×10^4 cells/well) were seeded in 24-well tissue culture plates (Thermo Fisher Scientific) with 0.5 mL serum-containing F12 medium and cultured for 6 days. The medium was replaced with an equal volume of fresh medium every 24 h and the spent medium was retained for measuring scFv-Fc concentration. Viable cell density was determined by the trypan blue exclusion method. The IgG fraction of a rabbit anti-human IgG (Fc) (Rockland Immunochemicals, Philadelphia, PA, USA) and a rabbit peroxidase-conjugated anti-human IgG antibody (Rockland Immunochemicals) were used as primary and secondary antibodies, respectively. A human Fc fragment (Jackson Immuno Research, West Grove, PA, USA) or purified scFv-Fc [78] was used as standards to create dilution series for calibration curves. Samples were prepared in triplicate, and data were expressed as means plus or minus the standard deviation.

4.3.3 Results and Discussion

4.3.3.1 First round of transgenes integration using two donor vectors

To obtain CHO/G[scFv-Fc]×2 cells, the first integration process was initiated by co-transfection of donor plasmids, R2 and R3-Green, and the Cre expression plasmid into CHO/R1 cells (Fig. 4-2). At 48h post-transfection, EGFP-expressing cells were collected by FACS and dispensed into 96-well plates at a single cell per well (Fig. 4-4A), no target cells were observed with the absence of Cre expression plasmid, ensuring that integration of target gene into predetermined *hprt* locus when Cre mediated recombination occurs. After 5–8 days of culture, target cells were observed by fluorescence microscopy (Fig. 4-5) and scaled up. 5 CHO/G[scFv-Fc]×2 clones showing a color shift from red to green (Fig. 4-4B, analyzed by FACS) were established.

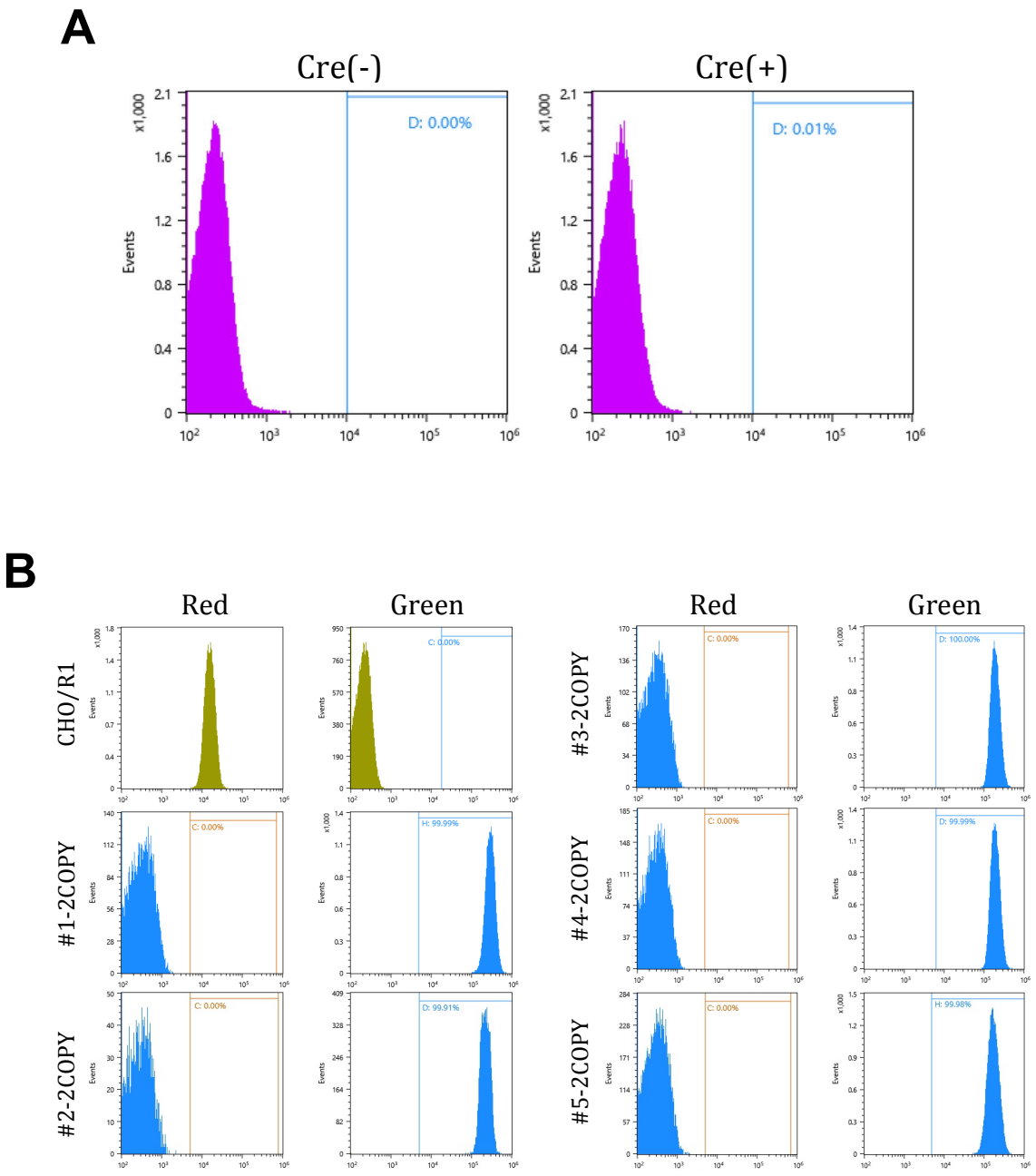


Fig. 4-4 Target cells shifed color after first round transfection.

(A) FACS analysis of donor plasmids transfected fonder cells. Left figure, without Cre expression plasmids; Right figure, transfection donor plasmids with Cre expression plasmids.

(B) FACS analysis of established target clones.

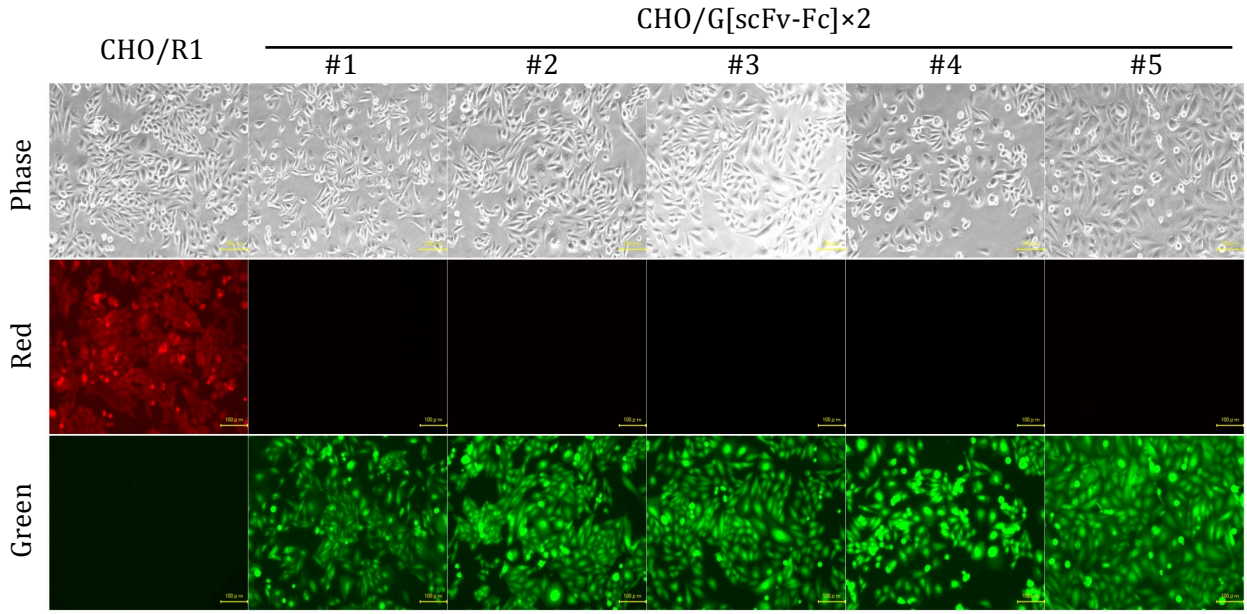


Fig. 4-5 Phase contrast and fluorescence microscopic observation of clones obtained from first round of AGIS integration.

4.3.3.2 Genomic PCR analysis for first round selected clones

Genomic PCR was performed for the detection of Cre-meditated site-specific integration, using various primer sets to validate the recombination reactions (Fig. 4-6A). The expected sizes of DNA fragments were amplified for some established clones using the primer pairs γ and θ, η and β, and η and ζ (Fig. 4-6B). Sequence analysis of the amplified DNA fragments confirmed that the transgenes were integrated at the target site pre-introduced into the *hprt* locus (Fig. 4-7), and that a new *loxP1*, the same as the initial integration site in CHO/R1 cells, was introduced into the *hprt* locus of the CHO/G[scFv-Fc]×2 cells. These results indicated that AGIS incorporating a single transfection with two donor vectors had successfully produced cells with two integrated transgenes.

Fig. 4-6 Genome structure analyses of CHO cell clones established by first AGIS.

(A) Expected genome structures of targeted recombinant CHO/G[scFv-Fc]×2 clones, and demonstration of primer pairs used in genomic PCR.

(B) PCR analysis of the clones' genomic DNA using specific primers. Lane M, molecular weight standard markers (mix of λ-*Hin*dIII and ΦX174-*Hin*dII digests); Lane W, negative control with water as the template; Lane F, Founder CHO/R1; Lane 1~5, CHO/G[scFv-Fc]×2 clones.

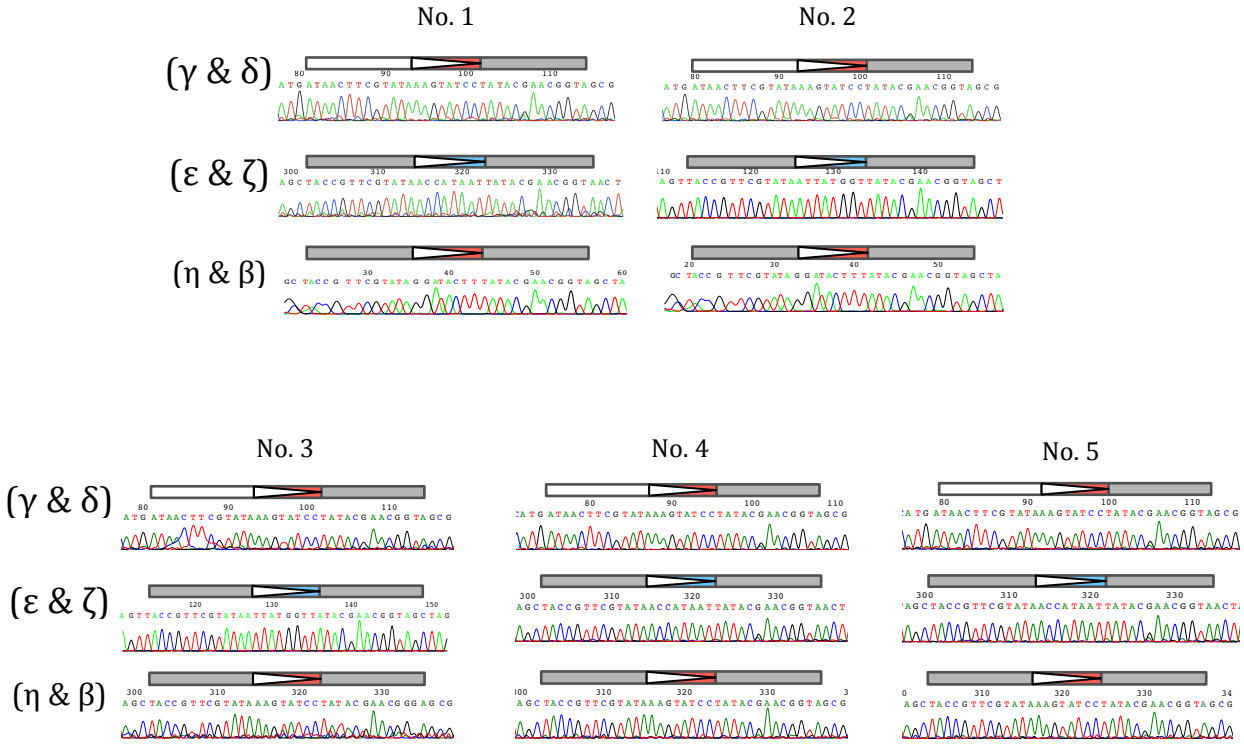


Fig. 4-7 DNA sequence analysis of recombined sites. Sequences of the PCR products shown in Fig. 4-6 were analyzed, and the sequences of the *loxP* sites associated with the recombination reactions are shown.

4.3.3.3 Evaluation of CHO cell clones generated by the first round of transgene integration

For comparing the scFv-Fc expression levels of each clone, a CHO cell clone possessing a single copy of the scFv-Fc expression unit at the *hprt* locus (CHO/B[scFv-Fc]×1 cells) was generated, which just express one copy of scFc-Fc antibody gene. By co-transfection of the donor plasmid, R2, and the Cre expression vector into CHO/R1 cells.

After the first round of transgene integration using AGIS, five clones confirmed by PCR were selected, and their scFv-Fc productivity was measured. CHO/B[scFv-Fc]×1 cells were used as a control. The scFv-Fc productivity for three clones (#1–3) was two-fold higher than that of the CHO/B[scFv-Fc]×1 cells (Fig. 4-8A). When the transgene copy numbers were determined by quantitative real-time PCR, these clones showed two copies of the scFv-Fc gene (Fig. 4-8B), indicating that scFv-Fc productivity corresponded to copy number of the transgene. On the other hand, clone #5 exhibited the highest scFv-Fc productivity. Quantitative real-time PCR analysis revealed that clone #5 possessed six copies of the scFv-Fc gene. These results indicated that the additional transgenes were integrated into off-target genomic sites, or that the *hprt* locus of clone #5 was triplicated during the cell screening process. Since clones showing unexpected behavior should be excluded, clone #1 was used for the next round of transgene integration. There were no significant differences in growth rates among the cells (Fig. 4-9).

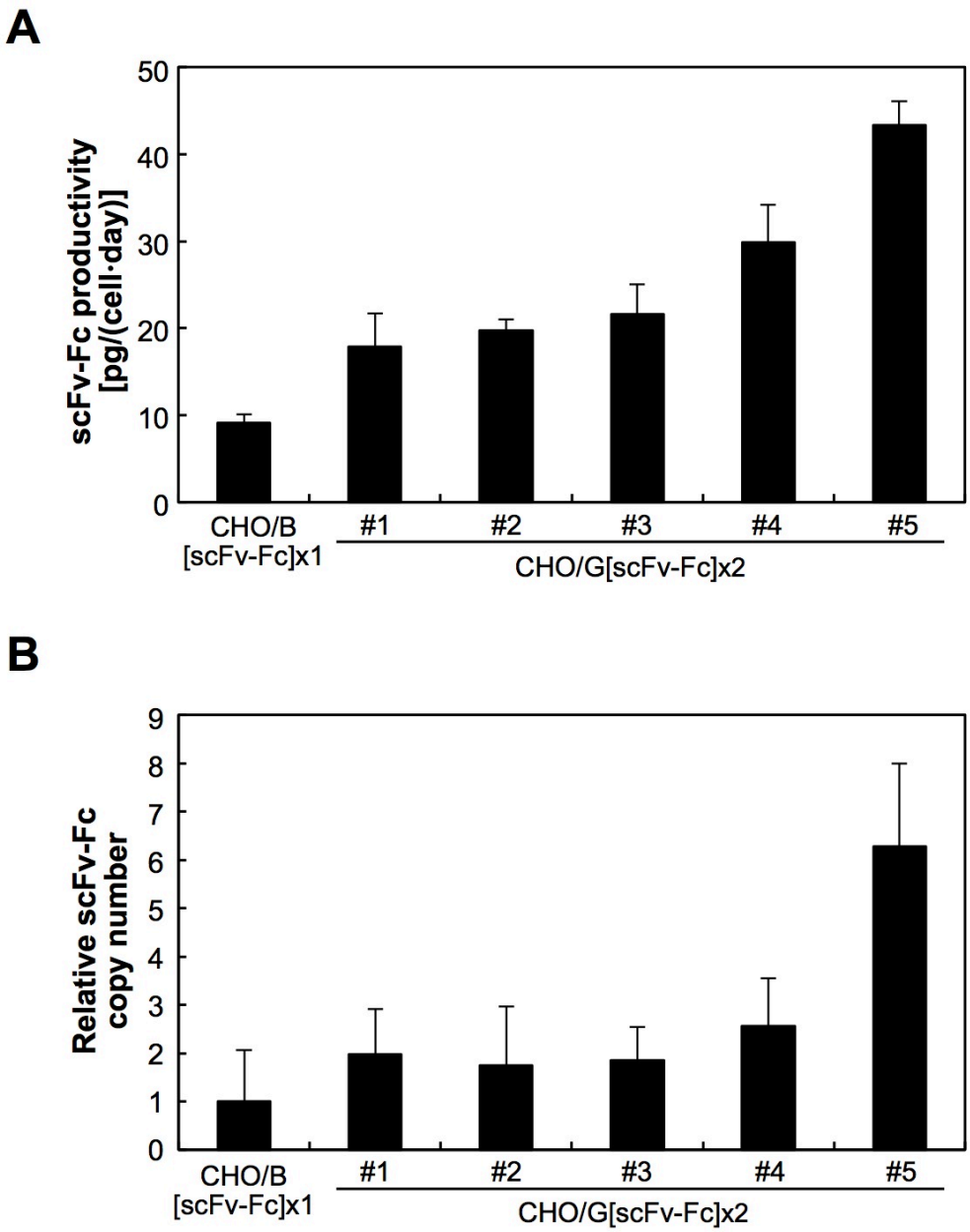


Fig. 4-8 Evaluation of CHO cell clones generated from the first round of integration using AGIS.

(A) Specific productivity of scFv-Fc by five CHO/G[scFv-Fc]×2 clones. The single-copy clone, CHO/B[scFv-Fc]×1 was used as a control.

(B) Relative scFv-Fc copy numbers of five CHO/G[scFv-Fc]×2 clones. Data are the means ± standard deviations of triplicate experiments.

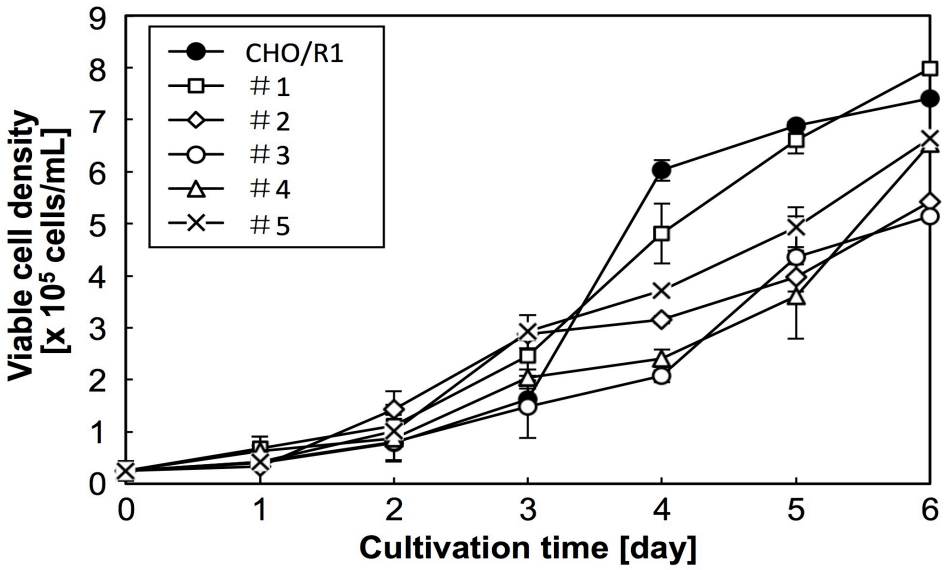


Fig. 4-9 Evaluation of obtained CHO cell growth trajectories.

4.3.4 Conclusion

In this experiment, target cells shifted fluorescent from red to green after first round of transfection, that accelerated the selection of target cells and 5 clones were obtained after FACS. The genomic structure at target recombination site of these clones were verified by genomic PCR analysis, proved that Cre-recombination events progressed as same as predicted. Obtained CHO/G[scFv-Fc]×2 cells showed at least twice fold scFv-Fc expression level of CHO/B[scFv-Fc]×1 cells, and Quantitative real-time PCR results showed that the scFv-Fc copy numbers of obtained cells are in proportion to antibody expression levels, proved that no target gene silencing happened in selected clones, which is very different target gene randomly integrated cells selected from inhibitor drug-screen methods.

In conclusion, in first round AGIS transfection, two copies of target expression units from two donor plasmids could be brought into the target locus by designed Cre recombination and mutated *loxP*s, in addition, transfected cells showed reasonable expression levels indicated favorable transcription of transgenes at *hprt* locus.

4.4 Second and third rounds of transgene integration in AGIS

4.4.1 Experimental Purpose

In the AGIS system, donor plasmids were repeatedly transfected in to the acceptable *loxP* site by Cre-recombination. Thus, to ensure the repeated target site integration is very important. In this experiment, the CHO/G[scFv-Fc]×2 got from first round of integration was applied to second round of transfection to obtain CHO cells possessing 4 copies of scFv-Fc express unit, and the newly obtained cells were used to confirm third round of AGIS integration.

In this section, the target is to establish high producer cells by using AGIS system, and one of the important points is that cells productivity should be corresponding to the antibody genes' copy number, which is different from random inserted recombinant cells. Thus, the target gene expression level and growth ability of obtained cells from each round of AGIS are confirm, and the copy number of target antibody expression units also measured, for verifying the positive correlation to antibody producing ability.

Finally, in order to evaluate whether AGIS established high producer cells is enough stable for industrial producing, stability of obtained cells carrying six copies of antibody expression units were check by long-term cultivation.

4.4.2 Materials and Methods

4.4.2.1 Cells and culture media

The founder cells used for transgene integration were CHO/R1 cells, in which a mutated *loxP* site (*loxP1*) and an expression cassette encoding the red fluorescent protein, DsRed, were introduced into the *hprt* locus. Cells were cultured under adherent conditions using Ham's F12 medium (Sigma-Aldrich, St. Louis, MO, USA) supplemented with 10% fetal bovine serum (Biowest, Nuaillé, France), 100 units/mL streptomycin sulfate and 90 μg/mL

penicillin G potassium (Wako Pure Chemical Industries, Osaka, Japan) at 37°C and 5% CO_2 in a humidified incubator.

4.4.2.2 Plasmid construction

A gene encoding another red fluorescent protein, mCherry, lacking the original ATG codon, was amplified by PCR from pmCherry (Takara), using primers Pr1 and Pr2. The PCR product was ligated into *Bam*HI- and *Xba*I-digested pBlue/*loxP1*/(ATG–) *EGFP*/polyA/*loxP5* to generate pBlue/*loxP1*/ (ATG–)*mCherry*/polyA/*loxP5*. A DNA fragment encoding (ATG–)*mCherry* and a polyA signal flanked by *loxP1* and *loxP5* was amplified by PCR from pBlue/*loxP1*/(ATG–)*mCherry*/ polyA/*loxP5*, using primers Pr3 and Pr4 . The PCR product was ligated into *Bgl*II- and *Xho*I-digested pcDNA4/TO/myc-HisA (Invitrogen) to generate pcDNA4/*loxP1*/(ATG–) *mCherry*/polyA/*loxP5*. A DNA fragment encoding an scFv-Fc expression unit (chEF1α/*scFv-Fc*/pA), prepared from pchEF1α/scFvFc, was ligated into *Eco*RV- and *Nhe*I-digested pcDNA4/*loxP1*/(ATG–)*mCherry*/polyA/*loxP5* to generate pcDNA4/*loxP1*/ (ATG–)*mCherry*/polyA/chEF1α/*scFv-Fc* /polyA/*loxP5* (R3-Red).

The blunt-ended DNA fragments were prepared using a kit (DNA blunting kit, Takara). All PCR reactions were performed using KOD plus neo DNA polymerase (Toyobo, Tsuruga, Japan) according to the manufacturer's instructions. All DNA sequences derived from chemically synthesized oligonucleotides and PCR products were confirmed by DNA sequencing on a Prism 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

The flowchart for the construction of R3-Red is given in Fig. 4-10.

Construction flowchart for plasmid R2, pMC-R2 was shown in materials and method part of Chapter 3 in this thesis. Construction flowchart of plasmid R3-Green used in this section was shown in Fig. 4-3.

The primer pairs used for vector construction are summarized in Table 4-4.

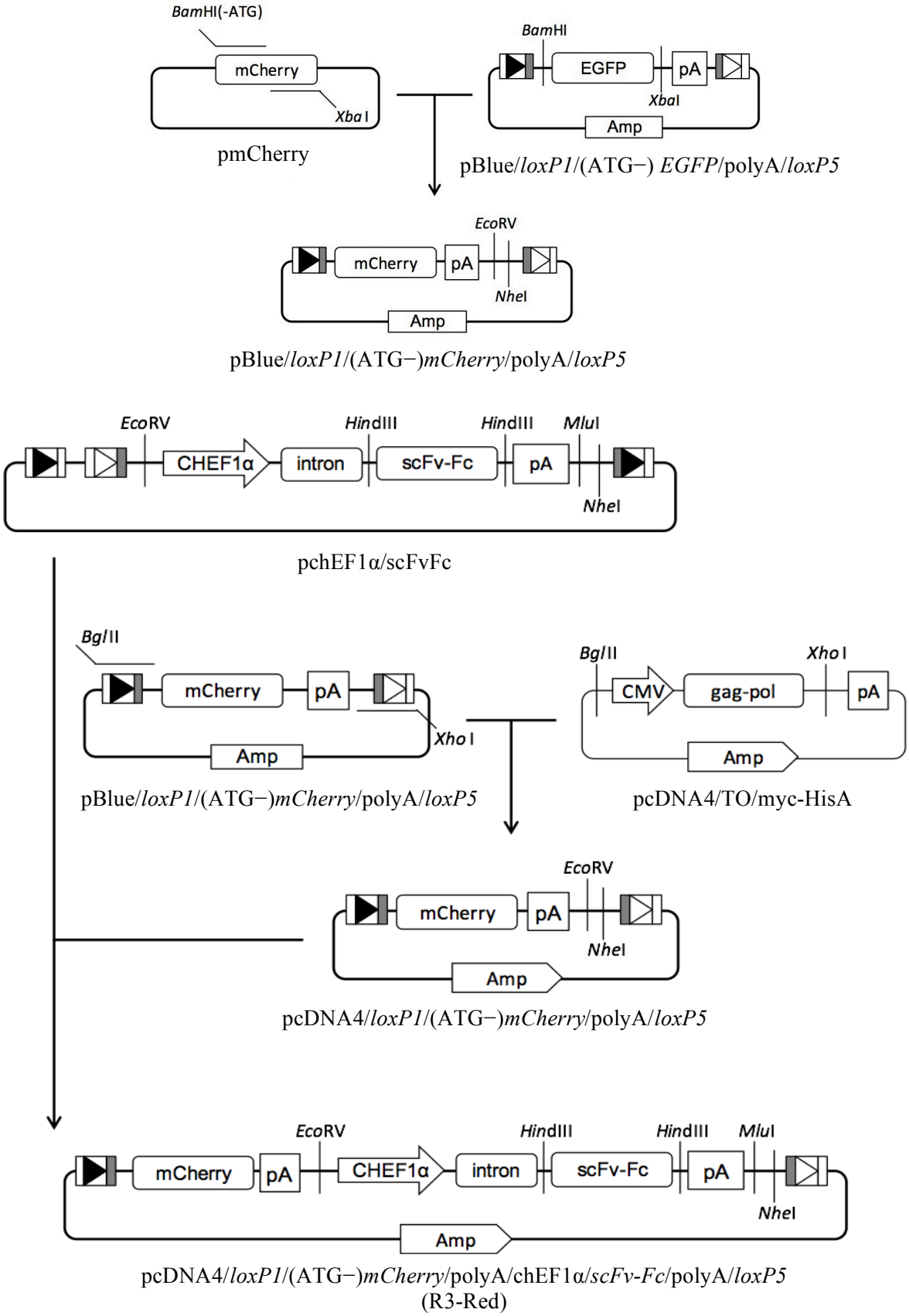


Fig. 4-10 Flowchart for the construction of R3-red plasmid.

TABLE 4-4. Primers for vectors construction.

Primer	Sequence (5'→3')
Pr1	AAG GAT CCG TGA GCA AGG GCG AGG AG
Pr2	AAT CTA GAC TAC TTG TAC AGC TCG TCC ATG C
Pr3	AAA GAT CTA TCG ATA TAA CTT CGT ATA AAG TAT CCT ATA CGA ACG GTA GCG
Pr4	AAA CTC GAG ATA ACT TCG TAT AAT TAT GGT TAT ACG AAC GGT AGC TAG

4.4.2.3 *LoxP* sequence used in first round of integration

The DNA sequences of the mutated *loxP* sites used in this experiment are shown in Table 4-5.

TABLE 4-5. Sequences of wild-type and mutated *loxP* sites.

Symbol	Sequence (5′→3′)		
	Left-arm region	Spacer region	Right-arm region
loxP	ATAACTTCGTATA	ATGTATGC	TATACGAAGTTAT
loxP1	ATAACTTCGTATA	AAGTATCC	TATACGAACGGTA
loxP2	TACCGTTCGTATA	AAGTATCC	TATACGAAGTTAT
loxP3	TACCGTTCGTATA	AAGTATCC	TATACGAACGGTA
loxP4	ATAACTTCGTATA	ACCATAAT	TATACGAACGGTA
loxP5	TACCGTTCGTATA	ACCATAAT	TATACGAAGTTAT
loxP6	ATAACTTCGTATA	ACCATAAT	TATACGAAGTTAT
loxP7	TACCGTTCGTATA	ACCATAAT	TATACGAACGGTA
loxP8	ATAACTTCGTATA	AAGTATCC	TATACGAAGTTAT

Mutated sequences are underlined.

4.4.2.4 Minicircle preparation

Minicircles were prepared using a commercial kit (Cat. No. MN920A-1, SBI) according to the manufacturer's instructions, with some modification. Briefly, pMC-R2 was transformed into a minicircle producer *E. coli* strain, ZYCY10P3S2T, and positive colonies were selected on LB plates containing 30 μg/mL kanamycin (Wako) after incubation overnight at 37°C. After pre-culture of clones in 2 mL of LB medium containing 30 μg/mL kanamycin at 30°C for 3 h with shaking (160 rpm), the cells were seeded into 25 mL growth medium and incubated at 30°C for 16–18 h with shaking. After the pH of culture broth was adjusted to 7.0 by adding 1 M NaOH solution, followed by mixing with an equal volume of induction medium containing arabinose, the cells were further cultured at 30°C for 5.5–6 h with shaking. A plasmid extraction kit (Qiagen, Hilden, Germany) was used to prepare the minicircle DNA, and contaminating genomic DNA and the parent plasmid were removed by digestion with a restriction enzyme (*Nde*I, Takara) and DNase (Plasmid-Safe™ ATP-Dependent DNase; Epicentre, Madison, WI, USA).

Flowchart for construction of pMC-R2 was shown in Chapter 3 in this thesis.

4.4.2.5 Cre-mediated accumulative gene integration into the CHO cell genome

For the second round of transgene integration, a minicircle DNA vector (mcR2) was used as one of the donor vectors. The CHO/G[scFv-Fc]×2 cells (5.0×10^6) were co-transfected with mcR2 (2.5 μg), R3-Red (2.5 μg) and pCEP4/NCre (1 μg) using an electroporation device (Neon transfection system, Invitrogen). After 5 days' culture in a 60-mm tissue culture dish (Thermo Fisher Scientific), the cells exhibiting a fluorescent shift from green (EGFP) to red (mCherry) were isolated using a cell sorter. The transgene integration site and genomic structure of isolated cells were confirmed by PCR, and the clones with four copies of the scFv-Fc expression unit at the *hprt* locus were designated as CHO/R[scFv-Fc]×4.

Similarly, the third round of integration was performed using CHO/R[scFv-Fc]×4 cells by transfecting the plasmids, mcR2, R3-Green and pCEP4/NCre. The cells exhibiting a fluorescent shift from red (mCherry) to green (EGFP) were isolated using a cell sorter. Clones with six copies of the scFv-Fc expression unit at the *hprt* locus were designated as CHO/G[scFv-Fc]×6.

The established clones were analyzed by quantitative real-time PCR to determine transgene copy number. scFv-Fc production was measured by ELISA, as described below.

4.4.2.6 Genomic PCR analysis & real-time PCR analysis

Genomic DNA was extracted from cells using a commercially available kit (MagExtractor Genome; Toyobo). Regions of Cre-mediated recombination were amplified by PCR using genomic DNA (50 ng) as a template. PCR was initiated with DNA polymerase (G-Taq, Cosmo Genetech, Seoul, Korea) at 95°C for 2 min, followed by 35 cycles of amplification at 95°C for 30 s, 56–57°C for 40 s, 72°C for 15–45 s and 72°C for 5 min for final extension. The genetic sequences of the amplicons were determined using a Prism 3130 Genetic Analyzer.

Copy numbers of the scFv-Fc expression unit were determined by real-time PCR (PikoReal96 Real-time PCR system, Thermo Fisher Scientific) as described in a previous report [77]. scFv-Fc producer CHO cells [76] possessing a single copy of the transgene in their genome, verified by Southern blotting, were used as a single copy control. The copy number values were expressed as means plus or minus the standard deviation.

The primers (α–ι) used are summarized in Table 4-6.

TABLE 4-6. Primers used for genomic PCR analysis.

Primer	Sequence (5′→3′)
α	TGC TCA CAT TTG GTG GGT GG
β	CAC CTT GAA GCG CAT GAA C
γ	TTT TGA GTT TGG AGC GAA GC
δ	CCC AGC ACC ACG AGT TCT G
ε	TGC TGC CCT CTG GTT ATG TG
ζ	CCA TCC CTG ACA CTT GTG AAA TC
η	CAG CAG GGG AAC GTC TTC TC
θ	ACC ACC CCG GTG AAC AG
ι	CAG ATG AAC TTC AGG GTC AGC TT

4.4.2.7 Measurement of cell growth and scFv-Fc production rates

The scFv-Fc production rate was measured as follow described. Briefly, all the established CHO cell lines (2.5×10^4 cells/well) were seeded in 24-well tissue culture plates (Thermo Fisher Scientific) with 0.5 mL serum-containing F12 medium and cultured for 6 days. The medium was replaced with an equal volume of fresh medium every 24 h and the spent medium was retained for measuring scFv-Fc concentration. Viable cell density was determined by the trypan blue exclusion method. The IgG fraction of a rabbit anti-human IgG (Fc) (Rockland Immunochemicals, Philadelphia, PA, USA) and a rabbit peroxidase-conjugated anti-human IgG antibody (Rockland Immunochemicals) was used as primary and secondary antibodies, respectively. A human Fc fragment (Jackson Immuno Research, West Grove, PA, USA) or purified scFv-Fc [78] was used as standards to create dilution series for calibration curves. Samples were prepared in triplicate, and data were expressed as means plus or minus the standard deviation.

4.4.2.8 Long-term culture for observing scFv-Fc producing cells' stability

For long-term culture, the cells were seeded in 6-well plates at a density of 3.0×10^5 cells/well. The culture medium was replaced with fresh one every other day, and the cells were re-seeded at the same density. The culture was repeated for 48 days.

4.4.3 Results and Discussion

4.4.3.1 Second and third rounds of transgene integration

CHO/G[scFv-Fc]×2 (clone #1) cells were used for the second round of transgene integration with AGIS. In the first approach, R2 and R3-Red were used as donor vectors. Transfected clones were isolated, amplified and analyzed for transgene integration. Results from genomic PCR analysis showed that cells possessing the expected transgene structure were not obtained. In this protocol, three vectors (R2, R3-Red and the Cre expression vector) have to be introduced into the cells, and four Cre-mediated recombination reactions must occur at the proper sites to generate the targeted cells. Therefore, transfection efficiency is a critical factor; however, transfection efficiency was very low, based on the color change that was seen.

The second approach aimed to improve transfection efficiency. A minicircle DNA vector, mcR2 was prepared (Fig. 4-11A, B) and used for transfection instead of R2. mcR2 lacks the bacteria-derived plasmid backbone of R2. Thus, the reduction in vector size could be expected to increase the frequency of targeted cell generation through improved transfection efficiency. In practice, the frequency of color-shifted cells using mcR2 was 19-fold higher than that using R2 (Fig. 4-11C). As a result, CHO/R[scFv-Fc]×4 clones showing the color change by FACS was established (Fig. 4-12A, B).

Since the CHO/R[scFv-Fc]×4 cells possessed the *loxP1* site (Fig. 4-2) required for another round of transgene integration by AGIS, the cells were used for the third round of integration. mcR2 and R3-Green were used as donor vectors. EGFP-positive cells were isolated and then screened for increased scFv-Fc production. A targeted CHO/G[scFv-Fc]×6 clone shifter color again was obtained (Fig. 4-12A, B).

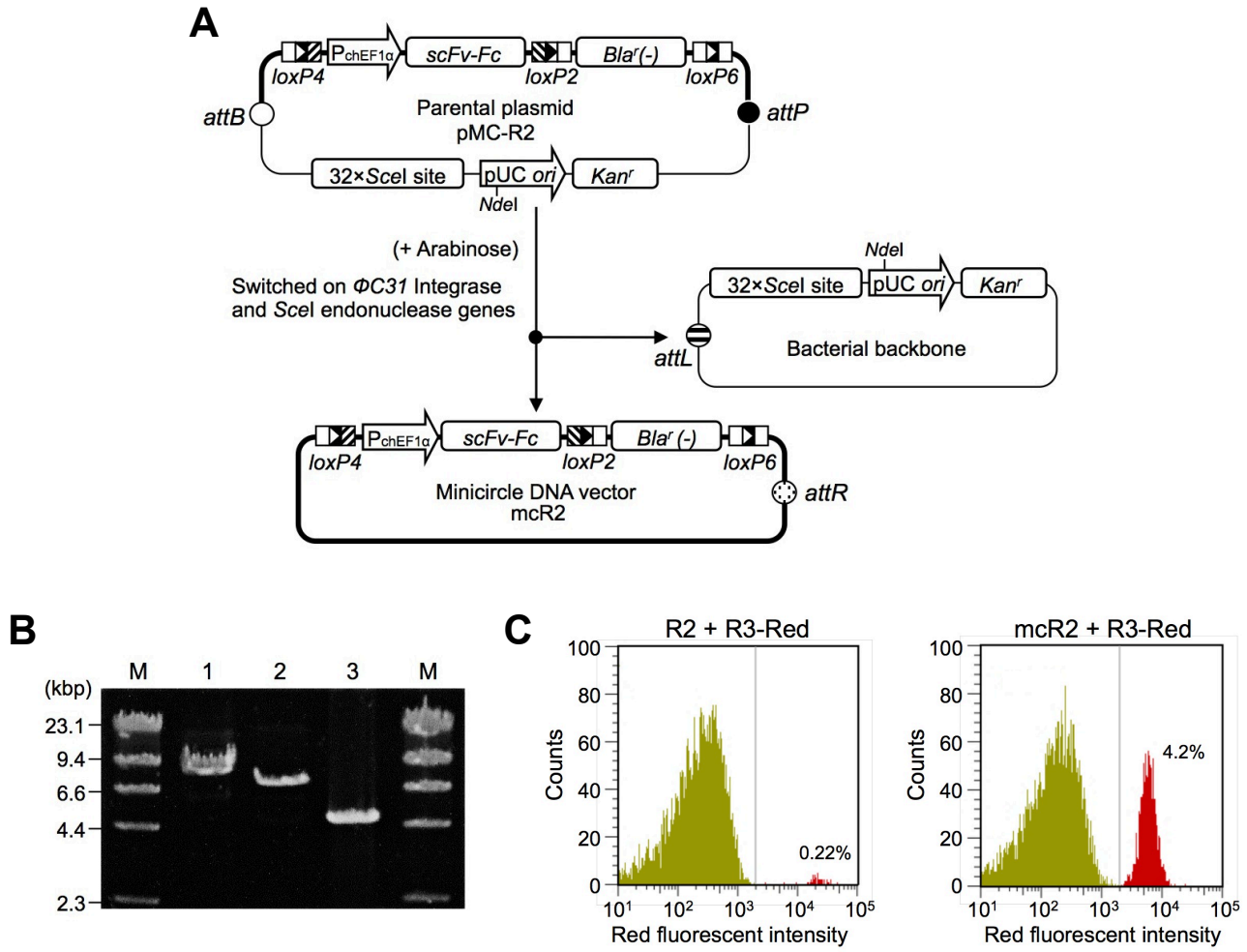


Fig. 4-11. Preparation of a minicircle DNA vector for AGIS.

(A) Schematic presentation of the generation of a minicircle DNA vector, mcR2, from the plasmid, pMC-R2, in *E. coli* cells. The parental pMC-R2 plasmid contained a pUC *ori* and a kanamycin resistance gene flanked by *attP* and *attB* sites, which can be recognized by ΦC31 integrase. In the minicircle producer strain, ZYCY10P3S2T *E. coli* cells, ΦC31 integrase and *Sce*I endonuclease genes are activated upon the addition of arabinose. Thus, pMC-R2 produced fragments of circular DNA, mcR2, and the bacterial backbone. The untransformed and backbone plasmids were linearized by *Nde*I digestion, then further degraded by endogenous bacterial DNase to leave the minicircle DNA vector, mcR2.

(B) Electrophoresis analysis of the vectors. Linearized plasmid and minicircle vectors were electrophoresed on a 1% (w/v) agarose gel. Lane M, molecular weight markers (λ-*Hin*dIII digest); Lane 1, pMC-R2 (8,688 bp); Lane 2, R2 (7,540 bp); Lane 3, mcR2 (4,650 bp).

(C) Comparison of the transfection efficiencies of R2+R3-Red and mcR2+R3-Red for the second round of transgene integration using AGIS. The total numbers of cells with a color change from green to red were counted using FACS.

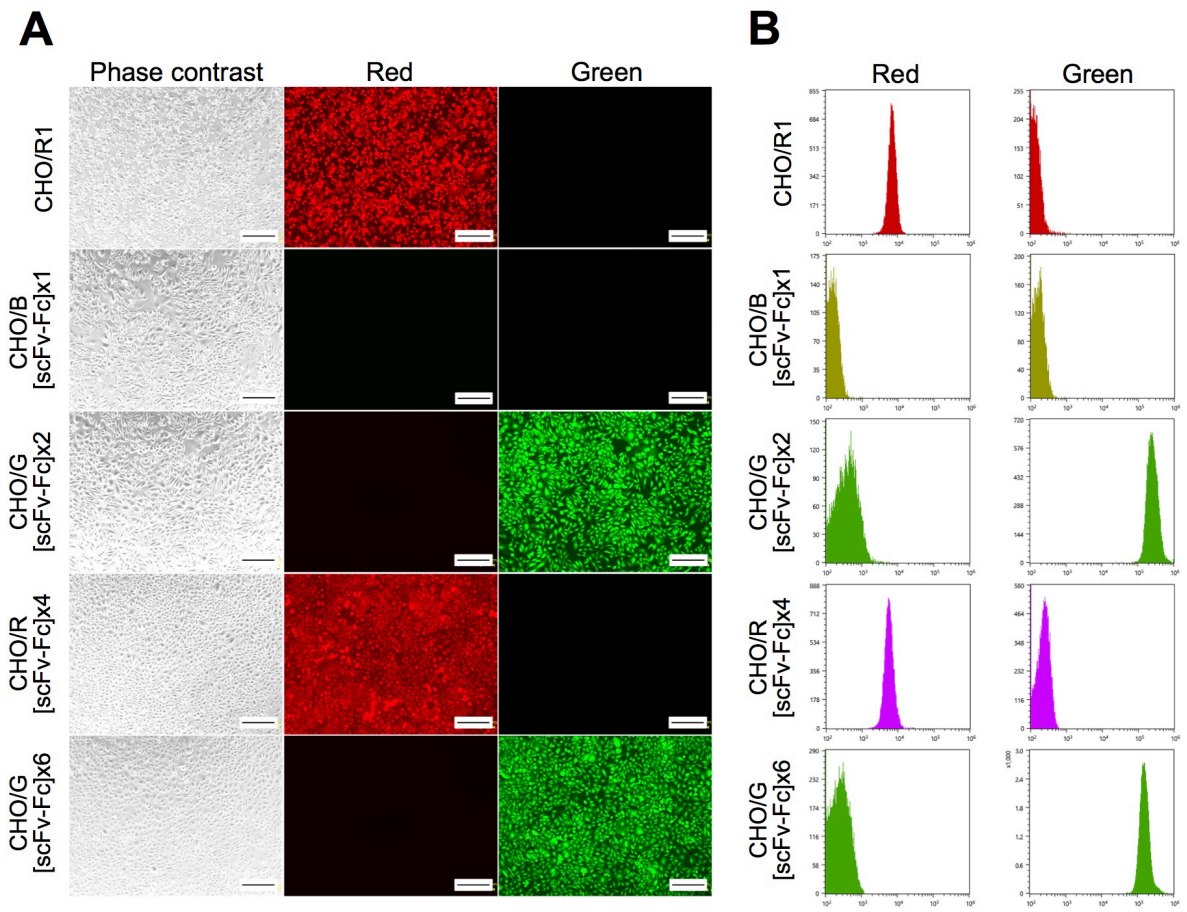


Fig. 4-12 Reporter fluorescent protein gene expression in CHO cell clones established by AGIS.
(A) Phase contrast and fluorescence microscopic observation of clones obtained from each integration round. Scale bars indicate 200μm.
(B) FACS analysis of clones established from each integration round.

4.4.3.2 Genomic PCR analysis for obtained clones in AGIS

The obtained targeted CHO/B[scFv-Fc]×1, CHO/G[scFv-Fc]×2, CHO/R[scFv-Fc]×4 and CHO/G[scFv-Fc]×6 clone possessing the proper transgene structure, as confirmed by PCR amplification of the junction regions of Cre-meditated integration (Fig. 4-13) and subsequent sequence analysis of the amplicons (Fig. 4-14).

All these data verified Cre mediated mutated-*loxPs* recombination in AGIS progressed as designed, after each round of transfection, transgenes could be integrated into the predicted target site.

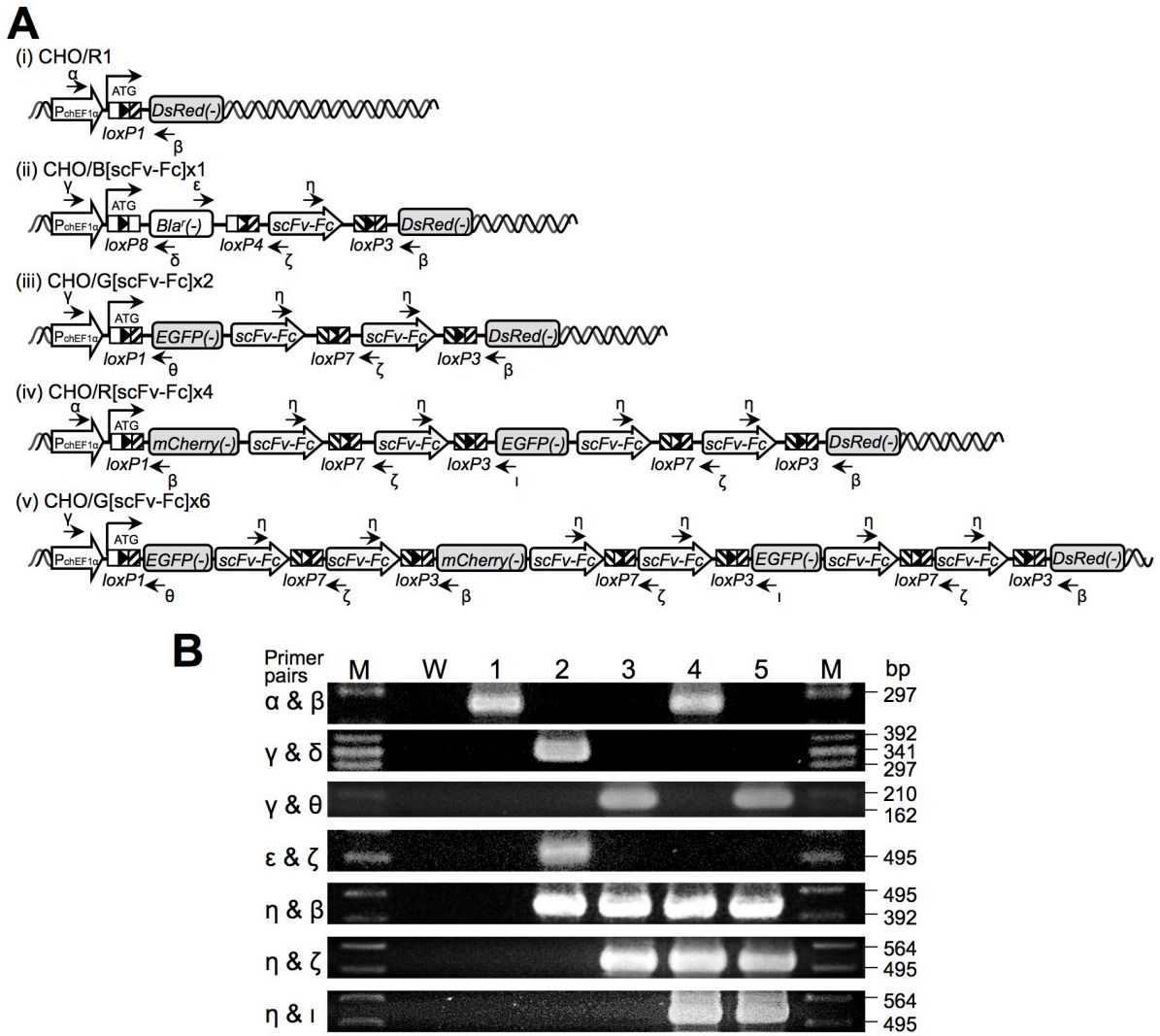


Fig. 4-13 Genome structure analyses of CHO cell clones established by AGIS.

(A) Expected genome structures of targeted recombinant CHO clones. (i) Engineered founder cells, CHO/R1, (ii) cells with one transgene copy, CHO/B[scFv-Fc]×1, obtained after one round of integration using an earlier version of AGIS, (iii) cells with two transgene copies, CHO/G[scFv-Fc]×2, obtained after one round of integration using the improved AGIS, (iv) cells with four transgene copies, CHO/R[scFv-Fc]×4, obtained after a second round of integration, and (v) cells with six transgene copies, CHO/G[scFv-Fc]×6, obtained after the third round of integration.

(B) PCR analysis of the clones' genomic DNA using specific primers (Table 4-6). Annealing sites of the primers in the transgene are depicted by arrows in Fig. 4-13A. Lane M, molecular weight standard markers (mix of λ-*Hin*dIII and ΦX174-*Hin*dII digests); Lane W, negative control with water as the template; Lane 1, Founder CHO/R1; Lane 2, CHO/B[scFv-Fc]×1; Lane 3, CHO/G[scFv-Fc]×2; Lane 4, CHO/R[scFv-Fc]×4; Lane 5, CHO/G[scFv-Fc]×6.

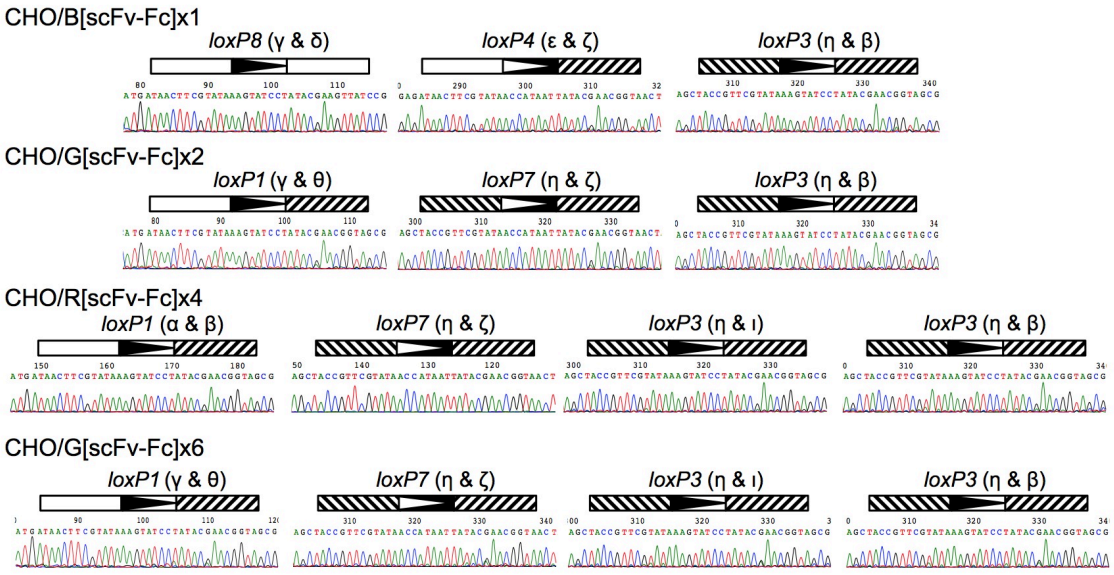


Fig. 4-14 DNA sequence analysis of recombined sites. Sequences of the PCR products shown in Fig. 4-13B were analyzed, and the sequences of the *loxP* sites associated with the recombination reactions are shown.

4.4.3.3 Analysis of scFv-Fc production by AGIS-generated CHO cells

Transgene copy number and scFv-Fc productivity were measured for CHO cells generated in each round of transgene integration. As determined by quantitative real-time PCR, the transgene copy numbers of CHO/G[scFv-Fc]×2, CHO/R[scFv-Fc]×4 and CHO/G[scFv-Fc]×6 cells were 2.0 ± 0.2, 4.0 ± 0.2 and 5.9 ± 0.6, respectively (Fig. 4-15A), indicating that copy number corresponded to the number of integration rounds as expected.

The scFv-Fc producer cells and the founder cells (CHO/R1) were cultured for 6 days to analyze cell growth and scFv-Fc productivity. There were no significant differences in growth rates among the cells (Fig. 4-15B). The scFv-Fc productivity of CHO/B[scFv-Fc]×1, CHO/G[scFv-Fc]×2, CHO/R[scFv-Fc]×4 and CHO/G[scFv-Fc]×6 cells was 10.8 ± 0.5, 20.1 ± 1.2, 35.7 ± 0.7 and 44.4 ± 1.3 pg/(cell·day), respectively (Fig. 4-15C). These results indicated that the scFv-Fc productivity increased corresponding to the transgene copy number, although the productivity was not exactly proportional to the copy number, possibly due to promoter interference [89]. Nonetheless, transcriptional enhancement via increased number of expression units was effective in improving scFv-Fc productivity.

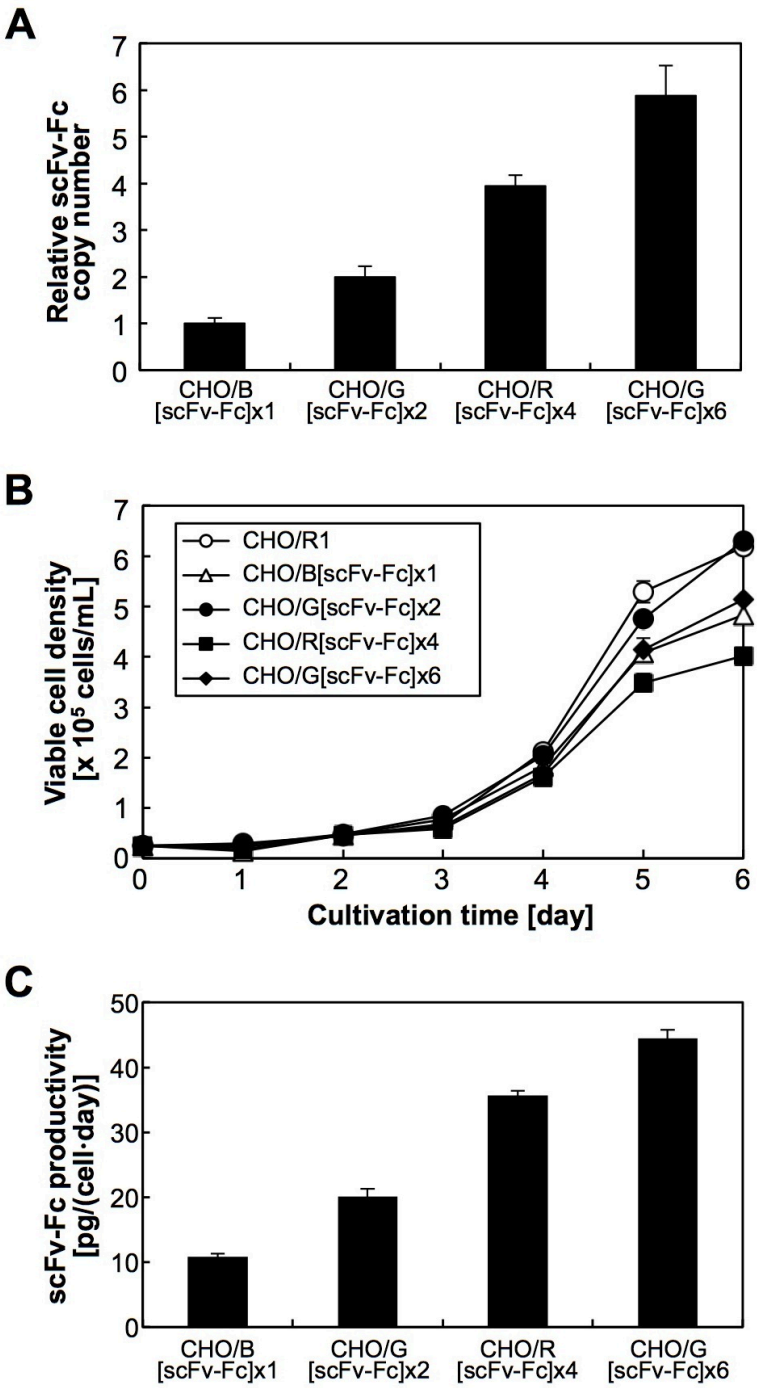


Fig. 4-15 Analysis of CHO cell clones established from each round of transgene integration using AGIS.

(A) Relative scFv-Fc copy number.

(B) Cell growth trajectories.

(C) Specific scFv-Fc productivity. Data are the means ± standard deviations of triplicate experiments.

4.4.3.4 Analysis of scFv-Fc production stability by AGIS-generated CHO cells

Furthermore, the stability in growth and productivity was evaluated using CHO/G[scFv-Fc]×6 cells (Fig. 4-16). The cells were cultured for 48 days (population doubling level; PDL=42). The specific growth rate was not changed during the culture, and the scFv-Fc productivity was remained more than 80%. Thus, the cells were relatively stable in growth and productivity for long-term culture. These results suggest that the *hprt* locus could be beneficial for high and stable transgene expression.

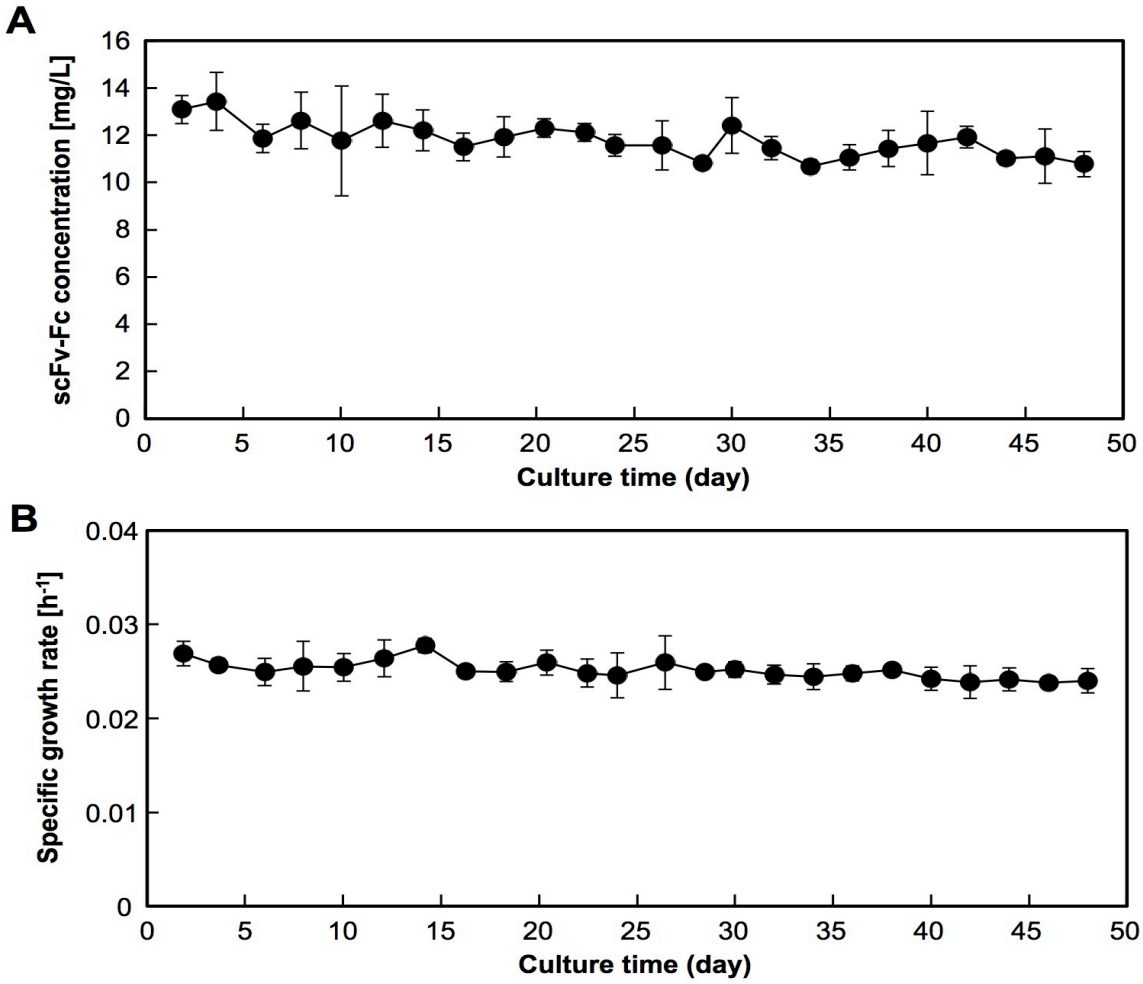


Fig. 4-16 Long-term culture of CHO/G[scFv-Fc]×6 cells.

(A) The scFv-Fc concentration.

(B) Specific growth rate. Data are the means ± standard deviations of triplicate experiments and measured during the culture.

The best specific productivity of recombinant antibodies under optimized culture conditions is currently 50–90 pg/(cell·day) in high-producer cells generated by the conventional gene amplification procedure [72]. In this study, scFv-Fc productivity reached 44 pg/(cell·day) with six copies of the expression unit integrated into the *hprt* locus. Optimizing transgene structure by incorporating insulator elements and by arranging the orientation of expression units in tandem repeats may further increase the specific productivity [77].

4.4.4 Conclusion

In conclusion, AGIS was used to introduce an scFv-Fc antibody gene into the *hprt* locus of CHO cells. Accumulative transgene integration into this locus effectively generated high-producer CHO cells. Current designed a new strategy for AGIS to facilitate and accelerate the integration and screening processes.

Simultaneous transfection of two donor vectors for integration and screening of targeted cells by color change allows efficient generation of producer cells. Furthermore, the use of a minicircle DNA vector as a donor improved the frequency of targeted cell generation. Incorporating this integration and screening strategy into AGIS greatly reduced the time required to establish cells with multiple copies of the transgene.

4.5 Summary

In the previously developed AGIS, a CHO cell line possessing a single arm-mutated *loxP* site for transgene integration was used as the founder. AGIS enables repeated transgene integration into a predetermined chromosomal site in mammalian cells. However, the process of establishing cells with multiple integrated copies of the transgene is still time-consuming. In the present study, the system was further modified by introducing a mutated *loxP* transgene integration site into the *hprt* locus of the CHO cell genome by homologous

recombination (CHO/R1 cells). Two donor vectors were simultaneously introduced into the cells in a single transfection. Cells with successfully targeted transgene integration were screened based on a change in the color of the reporter fluorescent protein that they express. Repeated rounds of integration allowed the transgene copy number to be increased. As a model, an scFv-Fc antibody gene was integrated into the *hprt* locus of the CHO cell genome. After three rounds of integration, a high-producer CHO cell clone with six copies of the scFv-Fc gene was successfully established. scFv-Fc productivity was approximately four-fold greater than a control cell line harboring a single copy of the transgene. This newly designed AGIS procedure should facilitate the generation of proper producer cells for biopharmaceutical protein production.

4.6 Prospect

Currently, common methods for generating high producer cells are inhibitor drug caused amplification system such as DHFR/MTX and GS/MSX, long-term screening and instability of producer cells are the bottlenecks of this conventional method. The improved AGIS system could ameliorate the selection process of stable high producer cell lines, although the target gene expression level is repressed after third round integration, the protein expression level is still very high and close to the reported highest level, further more, the repress situation expected to be change by other methods such as incorporating insulator elements, and our laboratory is working on this issue.

Chapter 5

Conclusion and Perspectives

5.1 Summary of the thesis

As a major class of biologics, recombinant mammalian protein plays an important role in manufacturing biopharmaceuticals. Productivity and quality are critical parameters values for industrial production of recombinant proteins. The productivity is related to the generation of high producer cells and the optimization of culture conditions, while the quality related to therapeutic effect and safety is decided by the post-translational modifications such as folding and glycosylation in the host cells. Chinese hamster ovary (CHO) cells are the most commonly used mammalian hosts for producing recombinant proteins due to the characteristic of human-like post-translational modification. The mainstream method for bringing on multiple copies of target genes in CHO cells genome is to use inhibitor for stimulating the amplification of the target genes, such as well-known DHFR/MTX system and GS/MSX system. These conventional systems are able to obtain high producer cells, but they are labor-intensive and time-consuming, and the obtained high producer cells sometimes cause problems such as instability of target protein synthesis as culture time progresses [90].

In our laboratory, an accumulative site-specific gene integration system (AGIS) using Cre-recombinase and mutated *loxP* sites was previously developed. AGIS enables repeated transgene integration into a predetermined chromosomal site in mammalian cells. However, the process of establishing cells with multiple integrated copies of the transgene is still time-consuming. In the present study of this thesis, an improved version of AGIS that

facilitates and accelerates the establishment of high-producer Chinese hamster ovary (CHO) cells is demonstrated.

In Chapter 1, background, bottlenecks and current techniques in the field of biopharmaceutical manufacturing, particularly for recombinant protein production using the mammalian cells, are described as an introduction. And these research statuses reflexed breakthroughs of the study in this thesis.

In Chapter 2, a literature review for the techniques and mechanisms related to the generation of high recombinant protein producer cells was conducted. It is including some techniques and mechanisms theoretically support the study in this thesis, such as adoption of *hprt* locus as a transgene integration site, and irreversible site-specific integration conducted by Cre-mediated mutated-*loxPs* recombination.

In Chapter 3, a study on the DNA transfection using minicircle vectors lacking the bacteria-derived plasmid backbone sequences is demonstrated, in which minicircle vectors improved the target gene integration efficiency and the cells viability during cultivation. In this study, two types of recombination were applied to introduce a scFv-Fc encoding gene into CHO/R1 cells genome with a blasticidin gene, one is for simple integration (R2v2 and mcR2v2), another is for integration accompanied by Cre-mediated deletion (R2 and mcR2). For both types of integration, conventional plasmid and corresponding minicircle vectors were used. After targeted integration, 3-fold higher transfection efficiency was observed for using minicircle DNA vectors with a smaller size. The target cells were selected based on blasticidin resistance followed by microscopic observation of faded fluorescence. Half clones obtained from plasmids transfection showed unexpected structures and transgene silencing, while almost all minicircle transfected target clones showed expected structure and expression of scFv- Fc. Thus, the bacterial backbone could interfere the recombination process to decrease the efficiency of targeted recombination. In addition, some of cells

transfected using the R2v2 (backbone inserted) showed expected structure but unable to produce recombinant protein. These results suggested that adjacent integrated bacterial backbone caused a transgene silencing. All the generated clones with proper structures from different integrations showed similar productivity, indicating that *hprt* locus is available for stable transgene expression. However, the cell growth was affected significantly, the cells harboring the bacterial backbone sequence decreased slow growth rate, and the minicircle transfected clones showed absolutely higher growth rate, thus the accumulated antibody amount during the culture is very different. These findings exhibited that minicircle DNA vectors are recommended to use for generating producer cells.

In Chapter 4, an improved version of AGIS that facilitates and accelerates the establishment of high-producer Chinese hamster ovary (CHO) cells is described. This system enables speed-up and repeated multiple transgenes integration into a predetermined chromosomal site in mammalian cells. In improved AGIS, CHO/R1 cells carrying a DsRed expression unit at the *hprt* locus constitutively express the red fluorescent protein. To facilitate and accelerate the integration process, two transgene donor vectors were introduced into cells by a single transfection, whereas only a single transgene was sequentially introduced using the previous AGIS procedure. Targeted cells were screened using a fluorescence-activated cell-sorting (FACS) device, based on the color change of reporter fluorescent proteins. This accelerated the screening process compared with inhibitor drug mediated screening. Currently, the best specific productivity of recombinant antibodies under optimized culture conditions is 50–90 pg/(cell·day) in high-producer cells generated by the conventional gene amplification procedure [72]. In this study, scFv-Fc productivity reached 44 pg/(cell·day) with six copies of the expression unit integrated into the *hprt* locus. The stability in growth and productivity was evaluated using cells containing six scFv-Fc expression units. The cells were cultured for 48 days (population doubling level equals 42).

The specific growth rate was not changed during the culture, and the scFv-Fc productivity was remained more than 80%. These results suggested that *hprt* locus could be beneficial for high and stable transgene expression. This newly designed AGIS procedure should facilitate the development of producer cells suitable for biopharmaceutical protein production.

In Chapter 5, the thesis study is summarized and prospects for future applications are described.

5.2 Prospect

Gene amplification and target protein overexpression are the priorities among central of interests in the field of industrial recombinant protein production. Conventional methods for introduce multiple transgenes based on inhibitor drug screening are time-consuming and unstable, due to the uncontrolled integration site, studies in this thesis have tried to solve these problems. A newly designed AGIS method ameliorated the selection process of stable high producer cell lines was introduced in this thesis, and a CHO cell clone with six copies of the scFv-Fc gene was successfully established. This clone exhibits a high protein expression level close to the reported highest level, at the same time, it is stable in both productivity and reproduction during long-term cultivation, which also proved that *hprt* locus is available for recombinant protein production as a transgene integration site. In addition, the advantages of minicircle DNA vectors were confirmed in the study of this thesis, improved targeted transfection efficiency and recombinant cells growth rate were observed when using the minicircle DNA vectors. These newly designed methods are effective to speed up the selection process of suitable producer cells. Furthermore, they should facilitate the development of producer cells suitable for biopharmaceutical protein production.

References

[1] Walsh G. “Biopharmaceuticals: biochemistry and biotechnology.” John Wiley & Sons, 2013.

[2] Rader RA. “(Re) defining biopharmaceutical.” Nature biotechnology 26.7 (2008): 743-751.

[3] Dessain S, Fishman SE. “Preserving the Promise: Improving the Culture of Biotech Investment.” Academic Press, 2016.

[4] Williams DC, Van Frank RM, Muth WL, Burnett JP. “Cytoplasmic inclusion bodies in Escherichia coli producing biosynthetic human insulin proteins.” Science 215.4533 (1982): 687-689.

[5] Deering RP, Kommareddy S, Ulmer JB, Brito LA, Geall AJ. “Nucleic acid vaccines: prospects for non-viral delivery of mRNA vaccines.” Expert opinion on drug delivery 11.6 (2014): 885-899.

[6] Milstein C. “The hybridoma revolution: an offshoot of basic research.” Bioessays 21.11 (1999): 966-973.

[7] Wurm FM. “Production of recombinant protein therapeutics in cultivated mammalian cells.” Nature biotechnology 22.11 (2004): 1393-1398.

[8] Dumont J, Euwart D, Mei B, Estes S, Kshirsagar R. “Human cell lines for biopharmaceutical manufacturing: history, status, and future perspectives.” Critical reviews in biotechnology 36.6 (2016): 1110-1122.

[9] Kunert R, Reinhart D. “Advances in recombinant antibody manufacturing.” Applied microbiology and biotechnology 100.8 (2016): 3451-3461.

[10] Estes S, Melville M. “Mammalian cell line developments in speed and efficiency.” Mammalian Cell Cultures for Biologics Manufacturing. Springer Berlin Heidelberg, 2013. 11-33.

[11] Running Deer J, Allison DS. “High-Level Expression of Proteins in Mammalian Cells Using Transcription Regulatory Sequences from the Chinese Hamster EF1α Gene.” Biotechnology progress 20.3 (2004): 880-889.

[12] Williams S, Mustoe T, Mulcahy T, Griffiths M, Simpson D, Antoniou M, Irvine A, Mountain A, Crombie R. “CpG-island fragments from the HNRPA2B1/CBX3 genomic locus reduce silencing and enhance transgene expression from the hCMV promoter/enhancer in mammalian cells.” BMC biotechnology 5.1 (2005): 17.

[13] Senger RS, Karim MN. “Variable Site-Occupancy Classification of N-Linked Glycosylation Using Artificial Neural Networks.” Biotechnology progress 21.6 (2005): 1653-1662.

[14] Xu X, Nagarajan H, Lewis NE, Pan S, Cai Z, Liu X, Chen W, Xie M, Wang W, Hammond S, Andersen MR, Neff N, Passarelli B, Koh W, Fan HC, Wang J, Gui Y, Lee KH, Betenbaugh MJ, Quake SR, Famili I, Palsson BO, Wang J. “The genomic sequence of the Chinese hamster ovary (CHO)-K1 cell line.” Nature biotechnology 29.8 (2011): 735-741.

[15] Hefzi H, Ang KS, Hanscho M, Bordbar A, Ruckerbauer D, Lakshmanan M, Orellana CA, Baycin-Hizal D, Huang Y, Ley D, Martinez VS, Kyriakopoulos S, Jiménez NE, Zielinski DC, Quek LE, Wulff T, Arnsdorf J, Li S, Lee JS, Paglia G, Loira N, Spahn PN, Pedersen LE, Gutierrez JM, King ZA, Lund AM, Nagarajan H, Thomas A, Abdel-Haleem AM, Zanghellini J, Kildegaard HF, Voldborg BG, Gerdtzen ZP, Betenbaugh MJ, Palsson BO, Andersen MR, Nielsen LK, Borth N, Lee DY, Lewis NE. “A Consensus Genome-scale Reconstruction of Chinese Hamster Ovary Cell Metabolism.” Cell Systems 3.5 (2016): 434-443.

[16] Ozturk S, Hu WS. “Cell culture technology for pharmaceutical and cell-based therapies.” CRC Press, 2005.

[17] Ma N, Ellet J, Okediadi C, Hermes P, McCormick E, Casnocha S. “A single nutrient feed supports both chemically defined NS0 and CHO fed-batch processes: Improved productivity and lactate metabolism.” Biotechnology progress 25.5 (2009): 1353-1363.

[18] Chen N, Bennett MH, Kontoravdi C. “Analysis of Chinese hamster ovary cell metabolism through a combined computational and experimental approach.” Cytotechnology 66.6 (2014): 945-966.

[19] Yang M, Butler M. “Effects of ammonia on CHO cell growth, erythropoietin production, and glycosylation.” Biotechnology and bioengineering 68.4 (2000): 370-380.

[20] Reinhart D, Damjanovic L, Kaisermayer C, Kunert R. “Benchmarking of commercially available CHO cell culture media for antibody production.” Applied microbiology and biotechnology 99.11 (2015): 4645-4657.

[21] Ahmadi M, Damavandi N, Akbari Eidgahi MR, Davami F. “Utilization of Site-Specific Recombination in Biopharmaceutical Production.” Iranian biomedical journal 20.2 (2016): 68.

[22] Papapetrou EP, Schambach A. "Gene insertion into genomic safe harbors for human gene therapy." Molecular Therapy 24.4 (2016): 678-684.

[23] Cukuroglu E, Engin HB, Gursoy A, Keskin O. "Hot spots in protein–protein interfaces: towards drug discovery." Progress in biophysics and molecular biology 116.2 (2014): 165-173.

[24] Gordon JW, Scangos GA, Plotkin DJ, Barbosa JA, Ruddle FH. "Genetic transformation of mouse embryos by microinjection of purified DNA." Proceedings of the National Academy of Sciences 77.12 (1980): 7380-7384.

[25] Brinster RL, Chen HY, Messing A, van Dyke T, Levine AJ, Palmiter RD. "Transgenic mice harboring SV40 T-antigen genes develop characteristic brain tumors." Cell 37.2 (1984): 367-379.

[26] Palmiter RD, Brinster RL, Hammer RE, Trumbauer ME, Rosenfeld MG, Birnberg NC, Evans RM. "Dramatic growth of mice that develop from eggs microinjected with metallothionein-growth hormone fusion genes." Nature 300.5893 (1982): 611.

[27] Sung P, Klein H. "Mechanism of homologous recombination: mediators and helicases take on regulatory functions." Nature Reviews Molecular Cell Biology 7.10 (2006): 739-750.

[28] Woo RA, Poon RY. "Cyclin-dependent kinases and S phase control in mammalian cells." Cell cycle 2.4 (2003): 315-323.

[29] Wyman C, Kanaar R. "DNA double-strand break repair: all's well that ends well." Annu. Rev. Genet. 40 (2006): 363-383.

[30] Eom KS, Cheong JS, Lee SJ. "Structural analyses of zinc finger domains for specific interactions with DNA." J Microbiol Biotechnol 26.12 (2016): 2019-2029.

[31] Cathomen T, Keith Joung J. "Zinc-finger nucleases: the next generation emerges." Molecular Therapy 16.7 (2008): 1200-1207.

[32] Tebas P, Stein D, Tang WW, Frank I, Wang SQ, Lee G, Spratt SK, Surosky RT, Giedlin MA, Nichol G, Holmes MC, Gregory PD, Ando DG, Kalos M, Collman RG, Binder-Scholl G, Plesa G, Hwang WT, Levine BL, June CH. "Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV." New England Journal of Medicine 370.10 (2014): 901-910.

[33] Joung JK, Sander JD. "TALENs: a widely applicable technology for targeted genome editing." Nature reviews Molecular cell biology 14.1 (2013): 49-55.

[34] Sprink T, Metje J, Hartung F. “Plant genome editing by novel tools: TALEN and other sequence specific nucleases.” Current opinion in biotechnology 32 (2015): 47-53.

[35] Matsubara Y, Kato T, Kashimada K, Tanaka H, Zhi Z, Ichinose S, Mizutani S, Morio T, Chiba T, Ito Y, Saga Y, Takada S, Asahara H. “TALEN-mediated gene disruption on Y chromosome reveals critical role of EIF2S3Y in mouse spermatogenesis.” Stem cells and development 24.10 (2015): 1164-1170.

[36] Taheri-Ghahfarokhi A, Malaver-Ortega LF, Sumer H. “Genome Modification of Pluripotent Cells by Using Transcription Activator-Like Effector Nucleases (TALENs).” Cell Reprogramming: Methods and Protocols (2015): 253-267.

[37] Mahfouz MM, Li L, Shamimuzzaman M, Wibowo A, Fang X, Zhu JK. “De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks.” Proceedings of the National Academy of Sciences 108.6 (2011): 2623-2628.

[38] Barrangou R. “Diversity of CRISPR-Cas immune systems and molecular machines.” Genome biology 16.1 (2015): 247.

[39] Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. “A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.” Science 337.6096 (2012): 816-821.

[40] Cheng AW, Wang H, Yang H, Shi L, Katz Y, Theunissen TW, Rangarajan S, Shivalila CS, Dadon DB, Jaenisch R. “Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system.” Cell research 23.10 (2013): 1163-1171.

[41] Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, Volz SE, Joung J, van der Oost J, Regev A, Koonin EV, Zhang F. “Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system.” Cell 163.3 (2015): 759-771.

[42] Shmakov S, Abudayyeh OO, Makarova KS, Wolf YI, Gootenberg JS, Semenova E, Minakhin L, Joung J, Konermann S, Severinov K, Zhang F, Koonin EV. “Discovery and functional characterization of diverse class 2 CRISPR-Cas systems.” Molecular cell 60.3 (2015): 385-397.

[43] Caskey CT, Kruh GD. “The HPRT locus.” Cell 16.1 (1979): 1-9.

[44] Turan S, Bode J. “Site-specific recombinases: from tag-and-target-to tag-and-exchange-based genomic modifications.” The FASEB Journal 25.12 (2011): 4088-4107.

[45] Kishimoto K, Nakayama M, Kinoshita M. "In vivo recombination efficiency of two site-specific recombination systems, VCre/VloxP and SCre/SloxP, in medaka (Oryzias latipes)." Development, growth & differentiation 58.6 (2016): 516-521.

[46] Karimova M, Abi-Ghanem J, Berger N, Surendranath V, Pisabarro MT, Buchholz F. "Vika/vox, a novel efficient and specific Cre/loxP-like site-specific recombination system." Nucleic acids research (2012): gks1037.

[47] Branda CS, Dymecki SM. "Talking about a revolution: The impact of site-specific recombinases on genetic analyses in mice." Developmental cell 6.1 (2004): 7-28.

[48] Mondello C, Smirnova A, Giulotto E. "Gene amplification, radiation sensitivity and DNA double-strand breaks." Mutation Research/Reviews in Mutation Research 704.1 (2010): 29-37.

[49] Jun SC, Kim MS, Baik JY, Hwang SO, Lee GM. "Selection strategies for the establishment of recombinant Chinese hamster ovary cell line with dihydrofolate reductase-mediated gene amplification." Applied microbiology and biotechnology 69.2 (2005): 162.

[50] Chusainow J, Yang YS, Yeo JH, Toh PC, Asvadi P, Wong NS, Yap MG. "A study of monoclonal antibody-producing CHO cell lines: What makes a stable high producer?" Biotechnology and bioengineering 102.4 (2009): 1182-1196.

[51] Kim M, O'Callaghan PM, Droms KA, James DC. "A mechanistic understanding of production instability in CHO cell lines expressing recombinant monoclonal antibodies." Biotechnology and bioengineering 108.10 (2011): 2434-2446.

[52] Dorai H, Corisdeo S, Ellis D, Kinney C, Chomo M, Hawley-Nelson P, Moore G, Betenbaugh MJ, Ganguly S. "Early prediction of instability of chinese hamster ovary cell lines expressing recombinant antibodies and antibody-fusion proteins." Biotechnology and bioengineering 109.4 (2012): 1016-1030.

[53] Jun SC, Kim MS, Hong HJ, Lee GM. "Limitations to the Development of Humanized Antibody Producing Chinese Hamster Ovary Cells Using Glutamine Synthetase-Mediated Gene Amplification." Biotechnology progress 22.3 (2006): 770-780.

[54] Noguchi C, Araki Y, Miki D, Shimizu N. "Fusion of the Dhfr/Mtx and IR/MAR gene amplification methods produces a rapid and efficient method for stable recombinant protein production." PLoS One 7.12 (2012): e52990.

[55] Shimizu N, Miura Y, Sakamoto Y, Tsutsui K. “Plasmids with a mammalian replication origin and a matrix attachment region initiate the event similar to gene amplification.” Cancer research 61.19 (2001): 6987-6990.

[56] Shimizu N, Hashizume T, Shingaki K, Kawamoto JK. “Amplification of plasmids containing a mammalian replication initiation region is mediated by controllable conflict between replication and transcription.” Cancer research 63.17 (2003): 5281-5290.

[57] Davami F, Sardari S, Majidzadeh-A K, Hemayatkar M, Barkhrdari F, Omidi M, Azami M, Adeli A, Davoudi N, Mahboudi F. “Expression of a novel chimeric truncated t-PA in CHO cells based on in silico experiments.” BioMed Research International 2010 (2010).

[58] Turan S, Galla M, Ernst E, Qiao J, Voelkel C, Schiedlmeier B, Zehe C, Bode J. “Recombinase-mediated cassette exchange (RMCE): traditional concepts and current challenges.” Journal of molecular biology 407.2 (2011): 193-221.

[59] Voziyanov Y, Konieczka JH, Stewart AF, Jayaram M. “Stepwise manipulation of DNA specificity in Flp recombinase: progressively adapting Flp to individual and combinatorial mutations in its target site.” Journal of molecular biology 326.1 (2003): 65-76.

[60] Santoro SW, Schultz PG. “Directed evolution of the site specificity of Cre recombinase.” Proceedings of the National Academy of Sciences 99.7 (2002): 4185-4190.

[61] Zou J, Sweeney CL, Chou BK, Choi U, Pan J, Wang H, Dowey SN, Cheng L, Malech HL. “Oxidase-deficient neutrophils from X-linked chronic granulomatous disease iPS cells: functional correction by zinc finger nuclease–mediated safe harbor targeting.” Blood 117.21 (2011): 5561-5572.

[62] Lamartina S, Sporeno E, Fattori E, Toniatti C. “Characteristics of the adeno-associated virus preintegration site in human chromosome 19: open chromatin conformation and transcription-competent environment.” Journal of virology 74.16 (2000): 7671-7677.

[63] Ogata T, Kozuka T, Kanda T. “Identification of an insulator in AAVS1, a preferred region for integration of adeno-associated virus DNA.” Journal of virology 77.16 (2003): 9000-9007.

[64] Luo Y, Rao M, Zou J. “Generation of GFP Reporter Human Induced Pluripotent Stem Cells Using AAVS1 Safe Harbor Transcription Activator Like Effector Nuclease.” Current protocols in stem cell biology (2014): 5A-7.

[65] Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, Cost GJ, Zhang L, Santiago Y, Miller JC, Zeitler B, Cherone JM, Meng X, Hinkley SJ, Rebar EJ, Gregory PD, Urnov FD, Jaenisch R. "Genetic engineering of human pluripotent cells using TALE nucleases." Nature biotechnology 29.8 (2011): 731-734.

[66] Liu R, Paxton WA, Choe S, Ceradini D, Martin SR, Horuk R, MacDonald ME, Stuhlmann H, Koup RA, Landau NR. "Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection." Cell 86.3 (1996): 367-377.

[67] Samson M, Labbe O, Mollereau C, Vassart G, Parmentier M. "Molecular cloning and functional expression of a new human CC-chemokine receptor gene." Biochemistry 35.11 (1996): 3362-3367.

[68] Shimizu S, Yadav SS, An DS. "Stable delivery of CCR5-directed shRNA into human primary peripheral blood mononuclear cells and hematopoietic stem/progenitor cells via a lentiviral vector." SiRNA Delivery Methods: Methods and Protocols (2016): 235-248.

[69] Strathdee D, Ibbotson H, Grant SG. "Expression of transgenes targeted to the Gt (ROSA) 26Sor locus is orientation dependent." PLoS One 1.1 (2006): e4.

[70] Mao X, Fujiwara Y, Chapdelaine A, Yang H, Orkin SH. "Activation of EGFP expression by Cre-mediated excision in a new ROSA26 reporter mouse strain." Blood 97.1 (2001): 324-326.

[71] Koyama Y, Banzai T, Sonezaki S, Kusano K. "Stable expression of a heterogeneous gene introduced via gene targeting into the HPRT locus of human fibrosarcoma cells." Biotechnology and bioengineering 95.6 (2006): 1052-1060.

[72] Hacker DL, De Jesus M, Wurm FM. "25 years of recombinant proteins from reactor-grown cells-where do we go from here?" Biotechnology advances 27.6 (2009): 1023-1027.

[73] Lara AR, Ramírez OT, Wunderlich M. "Plasmid DNA production for therapeutic applications." Recombinant Gene Expression (2012): 271-303.

[74] Dietz WM, Skinner NE, Hamilton SE, Jund MD, Heitfeld SM, Litterman AJ, Hwu P, Chen ZY, Salazar AM, Ohlfest JR, Blazar BR, Pennell CA, Osborn MJ. "Minicircle DNA is superior to plasmid DNA in eliciting antigen-specific CD8+ T-cell responses." Molecular Therapy 21.8 (2013): 1526-1535.

[75] Šimčíková M, Prather KL, Prazeres DM, Monteiro GA. “Towards effective non-viral gene delivery vector.” Biotechnology and Genetic Engineering Reviews 31.1-2 (2015): 82-107.

[76] Kawabe Y, Makitsubo H, Kameyama Y, Huang S, Ito A, Kamihira M. “Repeated integration of antibody genes into a pre-selected chromosomal locus of CHO cells using an accumulative site-specific gene integration system.” Cytotechnology 64.3 (2012): 267-279.

[77] Kawabe Y, Inao T, Komatsu S, Huang G, Ito A, Omasa T, Kamihira M. “Improved recombinant antibody production by CHO cells using a production enhancer DNA element with repeated transgene integration at a predetermined chromosomal site.” Journal of bioscience and bioengineering 123.3 (2017): 390-397.

[78] Kamihira M, Ono K, Esaka K, Nishijima K, Kigaku R, Komatsu H, Yamashita T, Kyogoku K, Iijima S. “High-level expression of single-chain Fv-Fc fusion protein in serum and egg white of genetically manipulated chickens by using a retroviral vector.” Journal of virology 79.17 (2005): 10864-10874.

[79] Araki K, Araki M, Yamamura K. “Targeted integration of DNA using mutant lox sites in embryonic stem cells.” Nucleic acids research 25.4 (1997): 868-872.

[80] Inao T, Kawabe Y, Yamashiro T, Kameyama Y, Wang X, Ito A, Kamihira M. “Improved transgene integration into the Chinese hamster ovary cell genome using the Cre-loxP system.” Journal of bioscience and bioengineering 120.1 (2015): 99-106.

[81] Obayashi H, Kawabe Y, Makitsubo H, Watanabe R, Kameyama Y, Huang S, Takenouchi Y, Ito A, Kamihira M. “Accumulative gene integration into a pre-determined site using Cre/loxP.” Journal of bioscience and bioengineering 113.3 (2012): 381-388.

[82] Lu J, Zhang F, Xu S, Fire AZ, Kay MA. “The extragenic spacer length between the 5’ and 3’ ends of the transgene expression cassette affects transgene silencing from plasmid-based vectors.” Molecular Therapy 20.11 (2012): 2111-2119.

[83] Chen ZY, He CY, Meuse L, Kay MA. “Silencing of episomal transgene expression by plasmid bacterial DNA elements in vivo.” Gene therapy 11.10 (2004): 856-864.

[84] Klinman DM. “Immunotherapeutic uses of CpG oligodeoxynucleotides.” Nature Reviews Immunology 4.4 (2004): 249-259.

[85] Zhou HS, Liu DP, Liang CC. “Challenges and strategies: the immune responses in gene therapy.” Medicinal research reviews 24.6 (2004): 748-761.

[86] Kameyama Y, Kawabe Y, Ito A, Kamihira M. "An accumulative site-specific gene integration system using cre recombinase-mediated cassette exchange." Biotechnology and bioengineering 105.6 (2010): 1106-1114.

[87] Johnson GE. "Mammalian cell HPRT gene mutation assay: test methods." Genetic Toxicology: Principles and Methods (2012): 55-67.

[88] Obayashi H, Kawabe Y, Makitsubo H, Watanabe R, Kameyama Y, Huang S, Takenouchi Y, Ito A, Kamihira M. "Accumulative gene integration into a pre-determined site using Cre/loxP." Journal of bioscience and bioengineering 113.3 (2012): 381-388.

[89] Hasegawa K, Nakatsuji N. "Insulators prevent transcriptional interference between two promoters in a double gene construct for transgenesis." FEBS letters 520.1-3 (2002): 47-52.

[90] Noguchi C, Araki Y, Miki D, Shimizu N. "Fusion of the Dhfr/Mtx and IR/MAR gene amplification methods produces a rapid and efficient method for stable recombinant protein production." PLoS One 7.12 (2012): e52990.

Acknowledgements

First and foremost, I gratefully acknowledge the help of my supervisor, Prof. Masamichi Kamihira, who has offered me valuable suggestions and whose useful suggestions, incisive comments and constructive criticism have contributed greatly to the completion of this thesis. I do appreciate his patience, encouragement, and professional instructions during my academic studies.

I would like to further extend my thanks to my teachers, Dr. Yoshinori Kawabe and Assoc. Prof. Akira Ito, who have helped me directly and indirectly in my studies, from whose devoted teaching and enlightening lectures I have benefited a lot and overcame difficulties in experiments.

I am also grateful to my thesis committee members, Prof. Hiroyuki Ijima and Prof. Yoshiko Miura, for their critical reading of the thesis and their guidance.

Special thanks also go to all the lab members, past and present. They not only helped me with my research, but also let me know this country better. They are and will always be my friends. I also want to thanks my group members, for their supporting and contribution during my research in Japan.

Last but not the least, my gratitude also extends to my family who have been assisting, supporting and caring for me all of my life. Special thanks should go to my friends who have put considerable time and effort for my study.