

# Vectors Used in Bacterial and Eukaryotic Host Cells

Tianyi Shi

2020-01-20

## Contents

<b>1</b>	<b>The structure of the vectors</b>	<b>2</b>
1.1	Simple plasmids . . . . .	2
1.2	Other bacterial vectors . . . . .	2
1.3	Other Eukaryotic Vectors . . . . .	4
<b>2</b>	<b>Vector Construction</b>	<b>5</b>
2.1	Simple plasmids (for both bacteria and eukaryotes) . . . . .	5
2.2	Other vectors for bacteria . . . . .	5
2.3	Viral vectors for eukaryotes . . . . .	5
<b>3</b>	<b>Transformation/Transfection</b>	<b>5</b>
3.1	Transformation of bacterial cells . . . . .	5
3.2	Transfection of Eukaryotic Cells . . . . .	7
	<b>References</b>	<b>9</b>

### What are the key similarities and differences between vectors used in bacterial and eukaryotic host cells? What can such vectors be used for?

There are various ways by which genes can be introduced into a host organism and expressed. The choice of source genetic material, vector, and the method of transformation/transfection for a particular experiment are largely dependent on the physiology of the host organism, the size of the transgene, and the purpose of the study. However, it is useful to categorise these strategies into bacteria-compatible ones and eukaryote-compatible ones.

Most molecular cloning experiments begin with the production of a recombinant plasmid containing the gene of interest (GOI). For bacterial hosts, the recombinant plasmid is directly introduced into the cells. Bacteria naturally possess the machinery for replicating plasmids, which allows inheritance of the GOI in their progenies. Eukaryotic cells cannot reliably replicate plasmids, therefore introducing plasmids into eukaryotic cells normally results in *transient transfection*, which means the GOI is expressed in successfully transfected cells but not in their progenies. In order to produce a *stable cell line*, the GOI must be integrated into the genome of the host cell (by homologous recombination), which is an extremely rare event in nature but can be induced artificially, for example by viral/bacterial infection or by designer nucleases such as TALEN, ZFN and CRISPR-Cas9. In order to be compatible with the different transformation/transfection strategies and the physiology of the host cells, the molecular structure of the plasmids (or vectors of other kinds) vary.

# 1 The structure of the vectors

## 1.1 Simple plasmids

Plasmids are extrachromosomal self-replicating cytoplasmic (usually circular) DNA elements found in prokaryotes and, less commonly, in eukaryotes. They come in various forms, from simple plasmids used for direct transformation and fosmids for phage transduction. Plasmids optimised for expressing recombinant proteins share several features, as illustrated in Fig. 1 and 2:

- An origin of replication (*ori*), which allows it to be duplicated in bacterial cells. The type of origin of replication affects the copy number. Eukaryotic vectors generally carry two *oris*, one bacterial (e.g. pUC *ori*) and one viral (e.g. SV40 *ori*). The bacterial *ori* is for amplifying the plasmid in bacteria before transfection, and the viral *ori* allows for episomal amplification of plasmids in the eukaryotic host.
- A cloning site where the GOI is inserted. Typically it is crowded with restriction enzyme cutting sites (compatible with RE/ligase dependent cloning), but it can also be made compatible with other cloning methods, such as the LIC-compatible pBLIC-puro<sup>1</sup> plasmid developed by Patel et al. (2012).
- A promoter and associated operator/enhancer. The promoter sequences for bacteria and eukaryotes are different (-35/-10 consensus sequence v.s. TATA box) For a bacterial vector, the promoter is usually controlled by an inducible operator so that protein expression is induced manually only when the optimal bacterial density is reached (e.g. *lac* operator induced by IPTG). In contrast, for a eukaryotic vector, the promoter is usually made as processive as possible, because the natural rate of transcription is generally low, and cell growth and division are not important (they do not serve to amplify the plasmid but to express the protein). For example, the pEF-BOS vector developed by Mizushima and Nagata (1990) contains the promoter of EF-1 $\alpha$  (EF-1 $\alpha$  is one of the most abundant proteins in eukaryotic cells, so its promoter is highly processive). Recently, promoters of viral origin (e.g. CMV, SV40) are gaining popularity.
- A transcription terminator/polyA signal. The mechanisms of transcription termination in bacteria and eukaryotes differ. In bacteria, the terminator sequence consists of a symmetric DNA sequence, and its transcribed product folds into a hairpin structure, releasing RNA polymerase (RNAP) from DNA. In eukaryotes, CstF and CPSF bound on RNAP C-terminal domain recognise 3'-end processing (polyadenylation) sequence (AAUAAA and GU-rich sequence) and induce cleavage and polyadenylation.
- A selectable marker. Usually an antibiotic resistance gene (bacteria and eukaryotes are sensitive to different antibiotics). Fluorophore genes are also common in eukaryotic vectors.
- Some plasmids possess a start codon and associated Shine-Dalgarno sequence (for bacteria) or Kozak sequence (for eukaryotes).
- Some plasmids have tag sequences near the start or end of the ORF for easy purification of the recombinant protein by affinity chromatography (e.g. hex-His, GST, HSV) and associated cleavage sequence (e.g. TEV) for easy removal of the tag.

## 1.2 Other bacterial vectors

There are other forms of circular DNA, each with different desirable properties (larger insertion size, low copy number, phage-compatibility, etc.). They are sometimes called ‘high capacity vectors’ because their insertion sizes are greater than simple plasmids.

Bacterial artificial chromosomes (BACs) contain some regions derived from a special plasmid called the F (fertility) factor: the region containing the origin of replication as well as genes that ensured its precise segregation during bacterial cell division. A great advantage of BAC vectors is the large insertion size (100-200 kb). But the very large insertion can be a problem, in that it cannot be manipulated by restriction enzymes.

---

<sup>1</sup>pBLIC-puro was a gift from Priyamvada Rai (Addgene plasmid # 45197 ; <http://n2t.net/addgene:45197> ; RRID:Addgene\_45197)

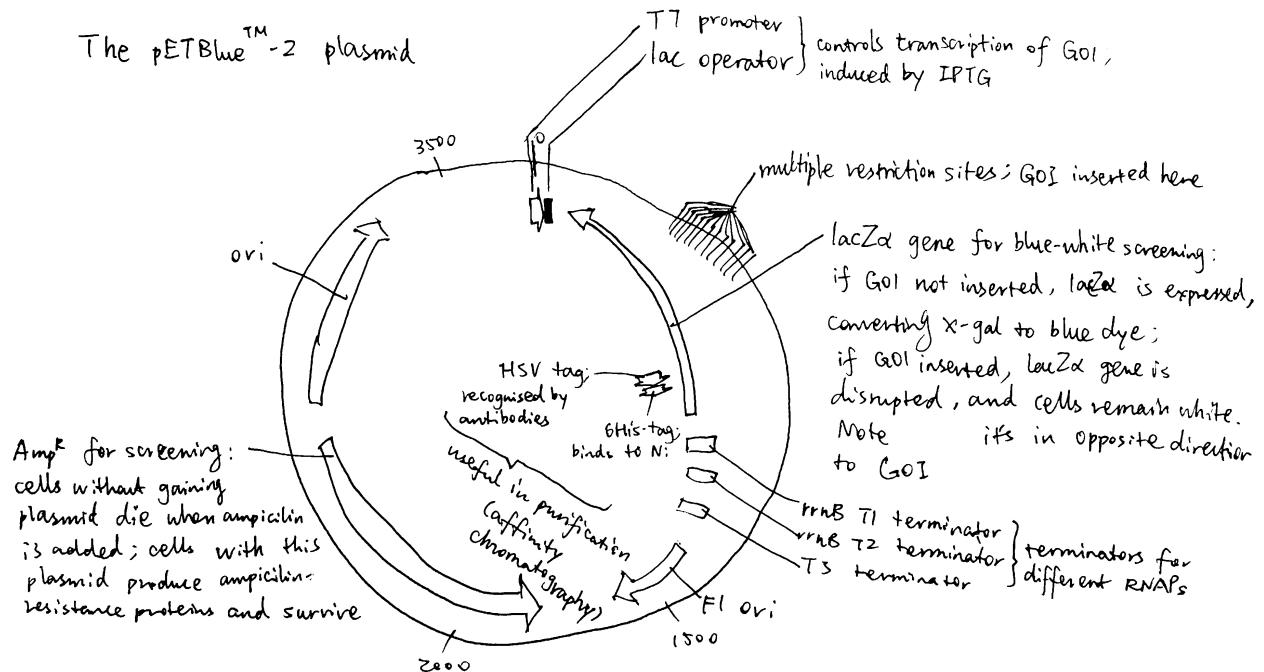


Figure 1: The bacterial expression vector pETBlue-2

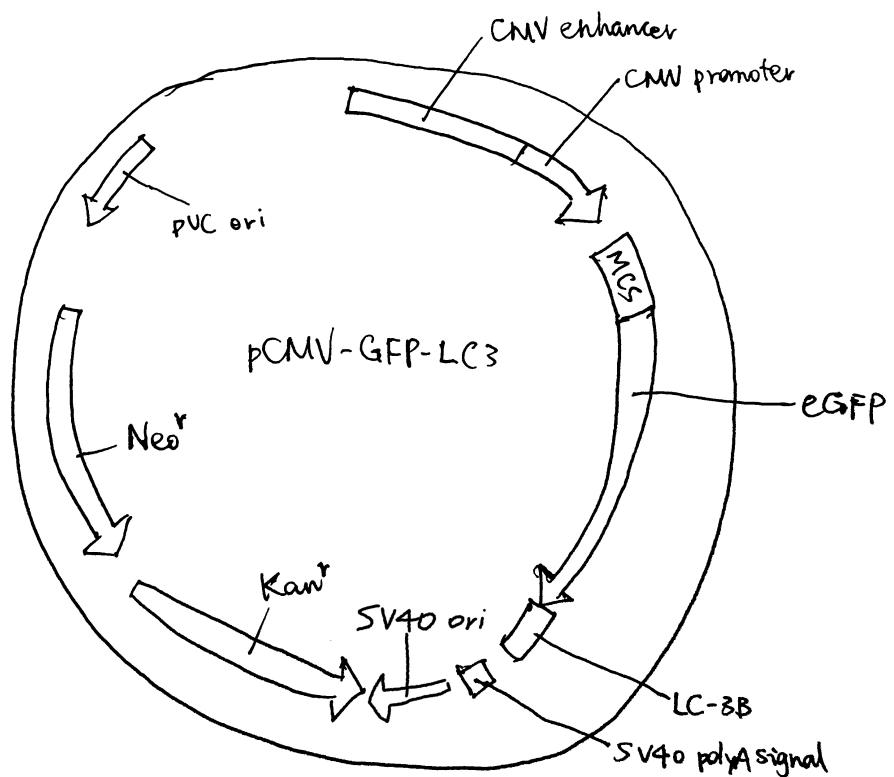


Figure 2: The eukaryotic expression vector pCMV-GFP-LC3

Instead, in vivo homologous recombination-based strategies are used to insert DNA fragments into BACs. (Fig. 3)

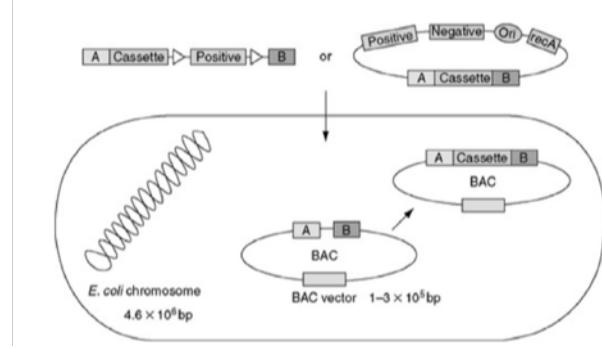


Figure 3: Making a BAC

Fosmids are hybrids of  $\lambda$  phage DNA and bacterial F plasmid DNA. Fosmids are packaged into  $\lambda$  phage particles, which deliver them into bacterial cells. Due to the presence of the F plasmid origins of replication, fosmids are maintained at a very low copy number (normally single copy).

Fig. 4 is a schematic of 6 high capacity vectors (Saraswathy and Ramalingam 2011).

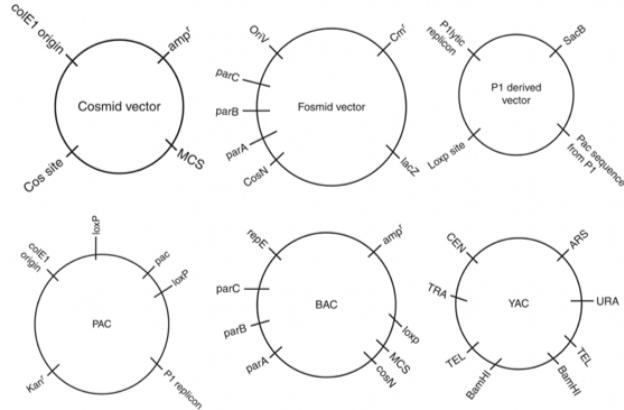


Figure 4: Schematic of 6 high capacity vectors.

### 1.3 Other Eukaryotic Vectors

Yeast artificial chromosome (YAC) has a cloning capacity up to 3000 kbp. It is introduced into the yeast cells by electroporation, then it is maintained as a linear DNA-like chromosome. It is replicated along with other chromosomes in yeast and its copy number of one is maintained after cell division.

Some viruses (e.g. Adenovirus, Lentivirus and Baculovirus) and bacteria (e.g. *Agrobacterium tumefaciens*) are reliable vectors for stable transfection of eukaryotic cells (see section 3.2.3)

## 2 Vector Construction

### 2.1 Simple plasmids (for both bacteria and eukaryotes)

Traditionally, restriction endonucleases and T4 ligase are used to insert GOI into plasmids. This method has several drawbacks:

- the association between the short 2-4 nt overhangs (sticky ends) is weak
- each DNA insert has to be inspected for any internal restriction sites
- due to the above two reasons, it has very limited capability of constructing multi-fragment plasmids
- incomplete DNA digestion and poor ligation yields
- unwanted amino acids can be introduced to the expressed protein
- the GOI can only be cloned in the vector position where the selected restriction site is present

Ligation-independent cloning (LIC) overcomes many of the problems described above. In a T4 DNA polymerase-dependent approach, the sticky sequences are made by the 3'-to-5' exonuclease activity of T4 polymerase. These sequences are long, allowing formation of stable recombinant plasmid without *in vitro* ligase treatment. The nicks in sugar-phosphate backbone are later fixed by the host's ligase. As T4 polymerase (exonuclease activity) always proceeds from 3' ends, any internal sticky sequences will not be disrupting and thus the custom sticky sequence can be used for any DNA inserts.

Fig. 5 compares the mechanisms of the traditional RE/ligase-based cloning and the T4 polymerase-dependent LIC approach.

There are also other alternatives to the traditional method:

- the Gateway cloning system exploits the site-specific recombination system utilized by bacteriophage lambda to shuttle sequences between plasmids bearing flanking compatible recombination attachment (att) sites
- Gibson assembly is a molecular cloning method which allows for ‘stitching’ multiple overlapping DNA fragments together in a single, isothermal reaction, using exonuclease, polymerase, and ligase.
- TOPO cloning exploits Taq polymerase’s feature that adds non-specific A to the 3’ end, and uses topoisomerase I to hold vectors open and to promote ligation

### 2.2 Other vectors for bacteria

BACs and fosmids: see section 1.2

### 2.3 Viral vectors for eukaryotes

See section 3.2.3.1

## 3 Transformation/Transfection

### 3.1 Transformation of bacterial cells

#### 3.1.1 Heat-shock transformation

Heat-shock transformation is used for small vectors such as plasmids. The general procedure is as follows:

- 1) Host cells are incubated in a solution containing divalent cations (typically  $\text{CaCl}_2$ ) on ice.
- 2)  $\text{CaCl}_2$  partially disrupts the cell membrane, which allows the recombinant DNA enter the host cell.  
Such cells are called competent cells.

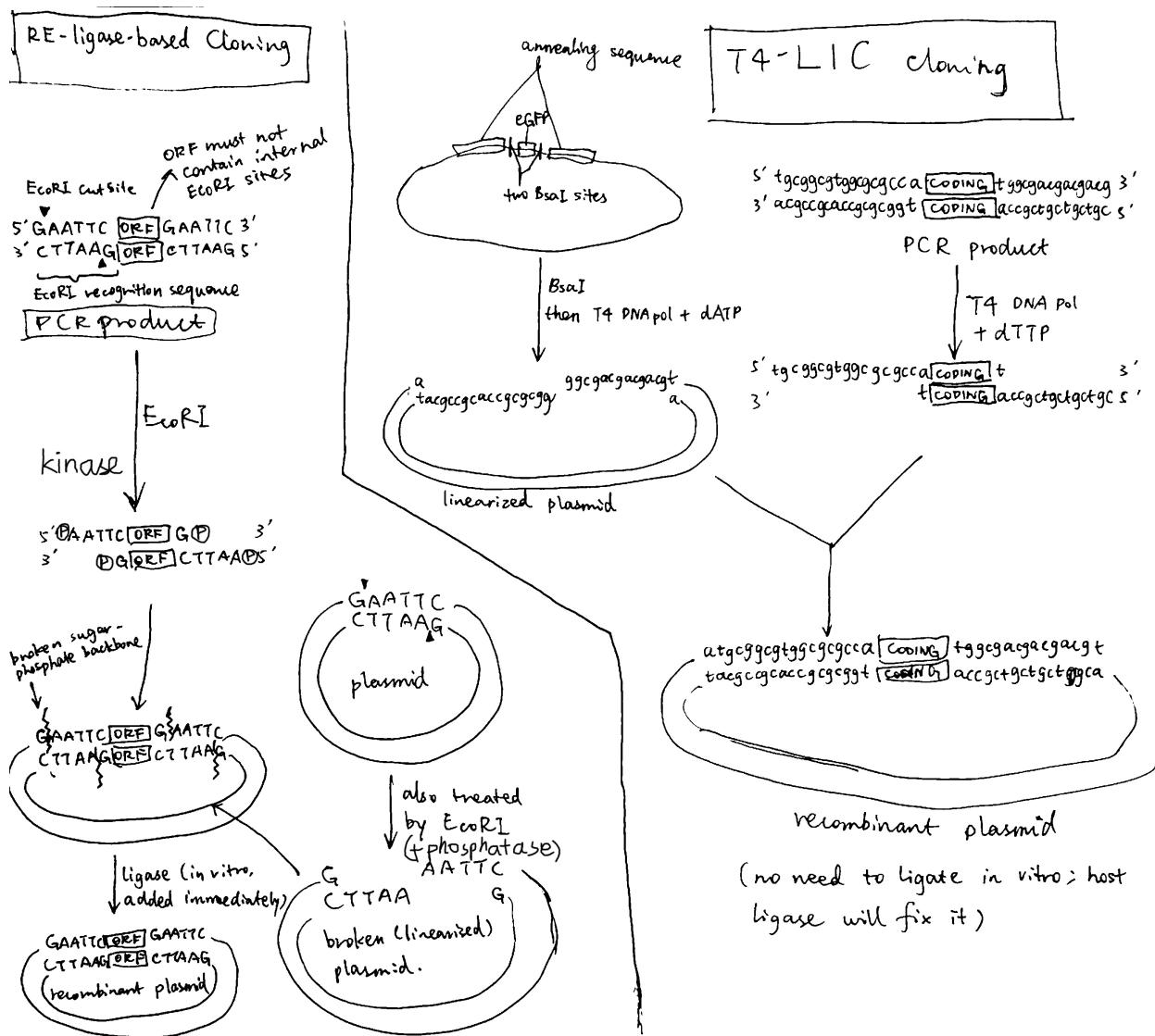


Figure 5: Left, traditional cloning based on restriction endonuclease and ligase; right, ligase-independent cloning based on T4 polymerase

- 3) Cells are exposed to a heat pulse (heat shock), and the thermal imbalance causes the entry of DNA through disrupted plasma membrane.

### 3.1.2 Electroporation

Electroporation can be used for larger vectors such as BAC and PAC<sup>2</sup>. The general procedure is as follows:

- 1) Host cells are placed into a cuvette, together with the vector. The cuvette is connected to electrodes.
- 2) The cuvette containing the mixture is subjected to intense electric pulses (2500 V/cm for bacteria, lower for animal and plant cells) each lasting for only a few milliseconds
- 3) Most cells would die under such treatment, but for those survived, their membranes are polarised by the electric field and are disrupted, and DNA enters through the pores. Finally, the membrane reseals after the treatment.

### 3.1.3 Other methods

Cosmid, fosmid and PAC vectors are introduced into bacterial cells via bacteriophages. (See section 1.2)

## 3.2 Transfection of Eukaryotic Cells

Transfection of eukaryotic cells can be transient or stable. Transient gene expression is generally used in academic settings, when the product of the GOI (usually a protein) is used for short-term research purposes. Stable transfection is usually used in industries, but it is also used in academia for specific purposes.

### 3.2.1 Physical methods

In general, physical methods work by briefly pushing holes through the membrane. The three specific methods are:

- electroporation, which is described above as a transformation method for bacteria, can also be applied to eukaryotic cells. The mechanism is the same but the voltage used for eukaryotic cells are lower.
- microinjection entails injection of recombinant DNA (or RNA) into the cytoplasm or nucleus. This is a method particularly convenient for large eukaryotic cells, such as the syncytial gonadal cell of *Caenorhabditis elegans* and the fertilised egg of *Mus musculus*.
- biolistic delivery of tungsten- or gold-coated DNA (or RNA) by gene gun is mostly used with plant cells.

### 3.2.2 Chemical methods

In general, chemical methods work by shedding the negative charge on DNA so as to approach and cross the plasma membrane, which is also negatively charged.

- CaPO<sub>4</sub> co-precipitation involves mixing DNA with CaCl<sub>2</sub> in a phosphate solution to generate a CaPO<sub>4</sub>-DNA co-precipitate, which is then added to cultured cells. DNA enters the cell by endocytosis.
- the cationic diethylaminoethyl(DEAE)-dextran works similarly by associating and neutralising negatively charged DNA
- cationic lipids work similarly by forming DNA-lipid complexes (DNA is not encapsulated in liposomes)
- By far, the most cost-effective chemical carrier is polyethylenimine (PEI), which works similarly by forming DNA-PEI complexes (L'Abbé et al. 2018)

---

<sup>2</sup>P1-derived artificial chromosome

### 3.2.3 Biological methods

The physical and chemical methods described above are generally used for transient transfection. Although they may produce stably transfected cells (when the rare homologous recombination events occur), the probability is very low and the GOI is inserted randomly and is thus not expressed with greatest efficiency (the chances are low that the GOI is inserted into a highly ‘transcriptionally open’ region).

Biological methods significantly improve the chance of integration of GOI into the host’s genome, and can even direct the site of insertion.

**3.2.3.1 Lentiviral transduction as an example of viral transfection (Merten, Hebben, and Bovolenta 2016)** Lentiviruses (LV) carry a genome made of (+) strand RNA. Upon infection, the viral RNA genome is reverse-transcribed and the cDNA is stably integrated into the host’s genome. Unlike  $\gamma$ -retroviruses, LVs exploit active nuclear transport and thus do not rely on cell division to access the nucleus. Fig. 6 is an example of the third generation HIV-1 vector system.

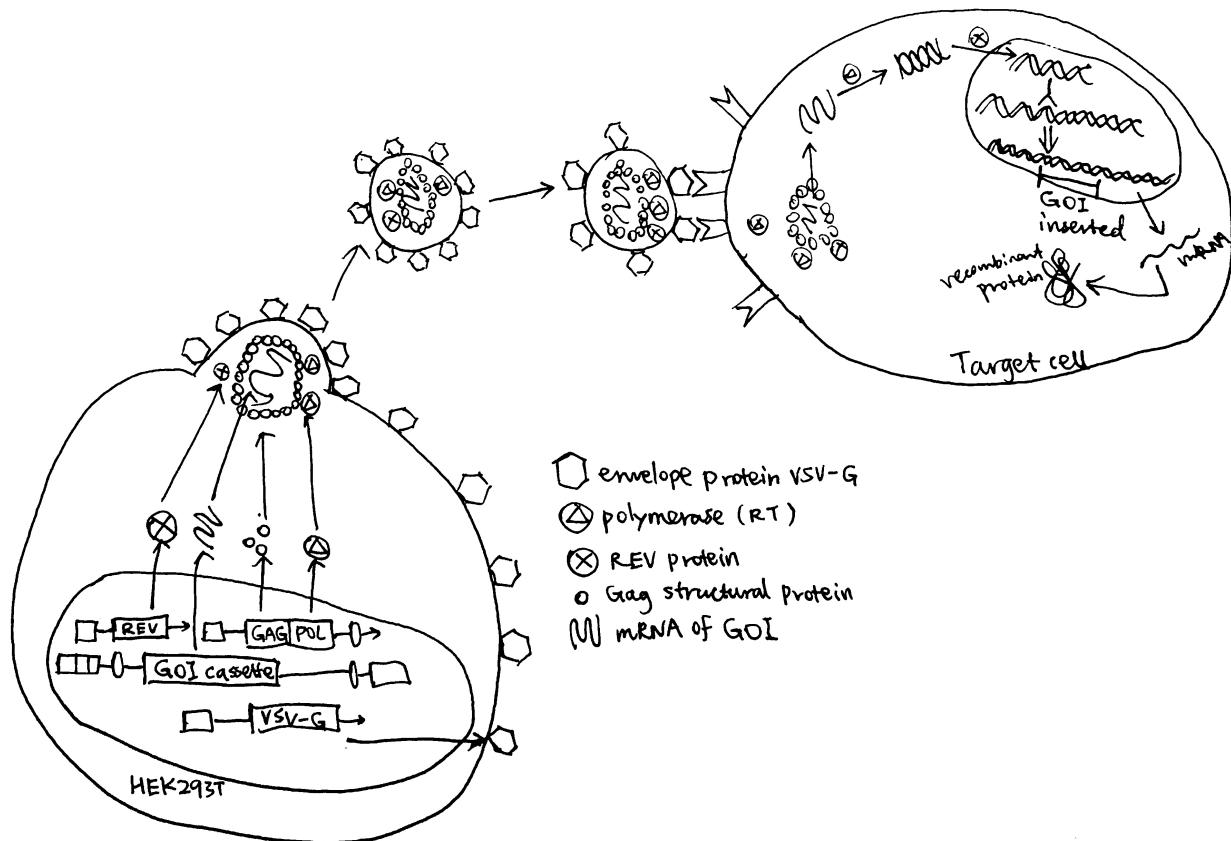


Figure 6: Production of recombinant LV particles and their infection. Production of recombinant LV particles is usually accomplished by transfection of HEK293T cells with a set of three helper plasmids and additionally a LV vector plasmid that encodes the GOI as well as optional reporter (e.g resistance, fluorophore) genes. LV particles released from HEK helper cells infect target cells and integrate the GOI into their genome.

**3.2.3.2 Ti plasmid and *Agrobacterium tumefaciens* infection** *Agrobacterium tumefaciens* is the bacterium that causes crown gall disease in plants. When it infects a plant cell, a part of its Ti plasmid, called T-DNA, is transferred and inserted (at random) into the genome of the plant cell. Thus, the T-DNA can be replaced with any GOI for stable transfection of plant cells.

**3.2.3.3 CRISPR-Cas9 Transfection** The CRISPR-Cas9 system comprises a short guide RNA (gRNA) and the Cas9 nuclease. The gRNA guides Cas9 to a specific gene locus that is complementary to its crRNA portion, then Cas9 makes a double strand break, which is dependent on the PAM (protospacer adjacent motif) sequence immediately downstream of the recognition sequence.

Following DNA cleavage, the break is repaired by either non-homologous end joining (NHEJ) or homology directed repair (HDR). It is the HDR mechanism that may cause integration of GOI into the cleavage site.

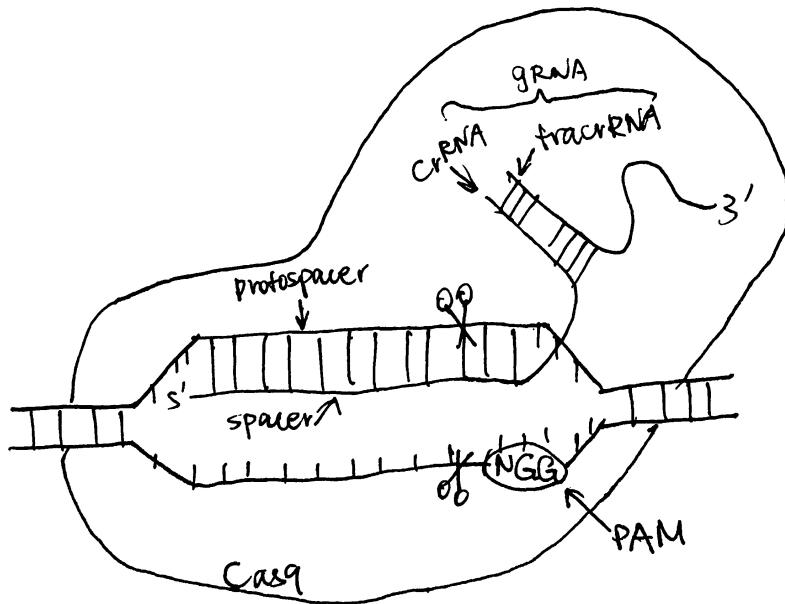


Figure 7: gRNA directed DNA cleavage by Cas9

## References

- Alberts, Bruce, Alexander Johnson, Julian Lewis, David Morgan, Martin Raff, Keith Roberts, and Peter Walter. 2014. *Molecular Biology of the Cell*. Book. 6th ed. Garland Science.
- Balbás, Paulina, and Argelia Lorence. 2012. *Recombinant Gene Expression: Reviews and Protocols*. Book. 2nd ed. Vol. 267. Methods in Molecular Biology. Springer.
- Griffiths, Anthony J. F., Susan R. Wessler, Sean B. Carroll, and John Doebley. 2015. “Gene Isolation and Manipulation.” In *Introduction to Genetic Analysis*, 11th ed. W.H. Freeman.
- Hacker, David L. 2018. *Recombinant Protein Expression in Mammalian Cells: Methods and Protocols*. Book. Vol. 1850. Methods in Molecularbiology. Humana Press. <https://doi.org/10.1007/978-1-4939-8730-6>.
- L’Abbé, Denis, Louis Bisson, Christian Gervais, Eric Grazzini, and Yves Durocher. 2018. “Transient Gene Expression in Suspension Hek293-Ebna1 Cells.” In *Recombinant Protein Expression in Mammalian Cells: Methods and Protocols*, edited by David L. Hacker. Vol. 1850. Humana Press.
- Merten, Otto-Wilhelm, Matthias Heben, and Chiara Bovolenta. 2016. “Production of Lentiviral Vectors.” *Molecular Therapy - Methods & Clinical Development* 3: 16017. <https://doi.org/https://doi.org/10.1038/mtm.2016.17>.
- Mizushima, S, and S Nagata. 1990. “PEF-Bos, a Powerful Mammalian Expression Vector.” *Nucleic Acids Res* 18 (17): 5322. <https://doi.org/10.1093/nar/18.17.5322>.

Patel, Asmita, Anisleidys Muñoz, Katherine Halvorsen, and Priyamvada Rai. 2012. "Creation and Validation of a Ligation-Independent Cloning (Lic) Retroviral Vector for Stable Gene Transduction in Mammalian Cells." *BMC Biotechnology* 12 (1): 3. <https://doi.org/10.1186/1472-6750-12-3>.

Saraswathy, Nachimuthu, and Ponnusamy Ramalingam. 2011. "High Capacity Vectors." In *Concepts and Techniques in Genomics and Proteomics*, 49–56. Woodhead Publishing. <https://doi.org/https://doi.org/10.1533/9781908818058.49>.

Singha, Tapan Kumar, Pooja Gulati, Aparajita Mohanty, Yogender Pal Khasa, Rajeev Kumar Kapoor, and Sanjay Kumar. 2017. "Efficient Genetic Approaches for Improvement of Plasmid Based Expression of Recombinant Protein in Escherichia Coli: A Review." Journal Article. *Process Biochemistry* 55: 17–31. <https://doi.org/https://doi.org/10.1016/j.procbio.2017.01.026>.

Valla, Svein, and Rahmi Lale. 2014. *DNA Cloning and Assembly Methods*. Book. Vol. 1116. Methods in Molecular Biology. Humana Press. <https://doi.org/10.1007/978-1-62703-764-8>.