

# 12 - Gene Cloning

**Gene cloning** involves removing a gene from the genome of an organism and then placing that isolated gene into the genome of a bacterial cell. The bacterial cell is then responsible for maintaining this foreign gene. For examples, the bacterial cell copies the foreign gene by DNA replication, the bacterial transcribes the foreign gene to make a messenger RNA (mRNA) molecule, and the bacterial cell translates the mRNA to make a protein product. Many important human proteins, including insulin (for diabetes patients) and factor VIII (for type A hemophilia patients), have been produced in large quantities by bacterial cells via this gene cloning technique.

## Overview

In the early 1970s, scientists isolated the DNA from two different organisms and covalently linked them together to form a hybrid DNA molecule in a test tube. This new hybrid DNA molecule, which contained DNA from two sources, is a **recombinant DNA molecule**. This experiment demonstrated the first use of **recombinant DNA techniques**, methods to manipulate DNA molecules outside of a living organism. Since then, many advances have made recombinant DNA techniques common practice among biologists.

Our discussion will focus on a process called **gene cloning (figure 12.1)**. Gene cloning involves isolating a particular gene of interest and inserting that gene into a **vector** DNA molecule. Commonly used vectors include the circular **plasmid** DNA molecules found in many bacteria. The resulting recombinant DNA molecule, composed partly of the gene of interest and partly the plasmid vector, is then introduced into a host bacterial cell by **transformation**. The host bacterial cell maintains the recombinant DNA molecule, so the gene of interest (i.e., the **cloned gene**) can be studied in more detail. Once gene cloning is complete, the cloned gene can be used in:

- **DNA sequencing experiments.** DNA sequencing provides the base pair sequence of the cloned gene.
- **Mutagenesis experiments.** Mutations can be introduced at desired locations within the cloned gene. The phenotypic consequences of these mutations can then be studied.
- **Protein expression studies.** Cloned genes can be expressed via transcription and translation to make a protein product. Protein expression allows scientists to examine the function of the protein product encoded by a cloned gene or the cloned gene can be expressed at high levels for medical purposes (i.e., to express the human insulin or factor VIII proteins).
- **Gene therapy.** Cloned genes can be introduced into human cells as a treatment for genetic diseases.

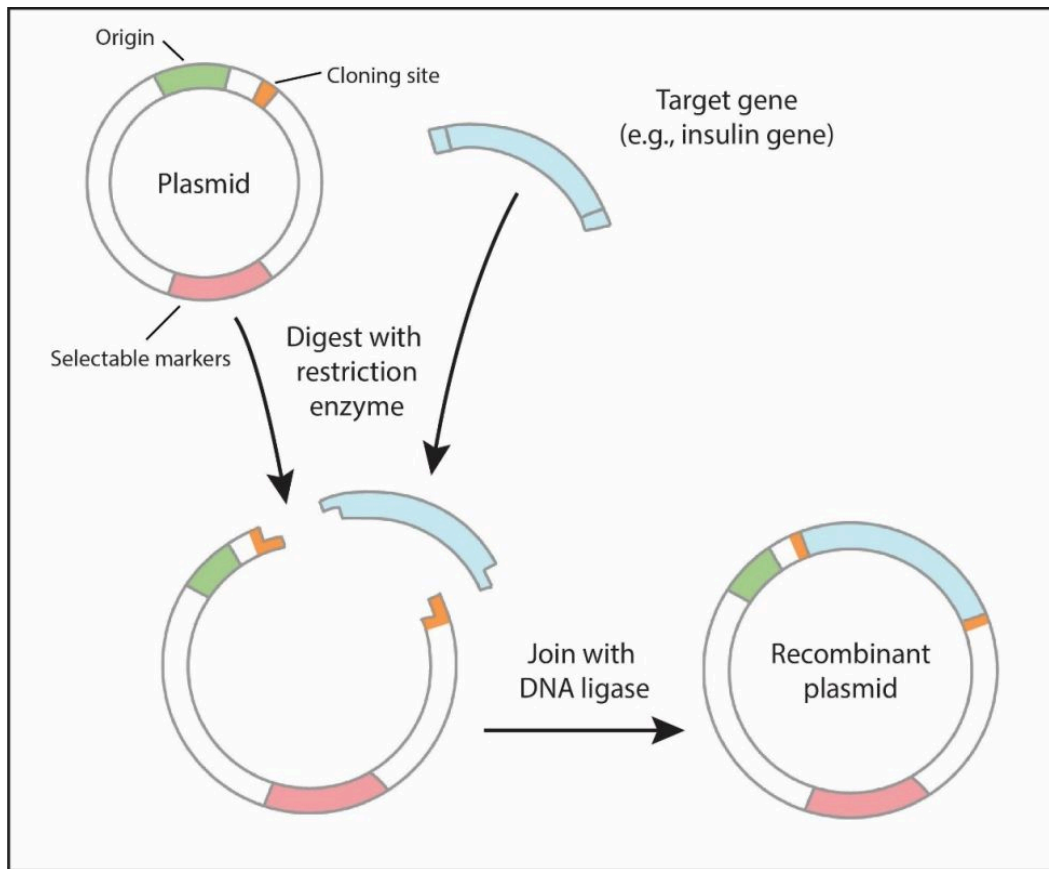


Figure 12.1 **Gene cloning overview** - Image by [Khan Academy](https://www.khanacademy.org), and modified by SL

## Key Questions

- What is a recombinant DNA molecule?
- What is meant by the term gene cloning?
- Why is it useful to clone a gene?

## Vectors

As mentioned earlier, gene cloning involves removing a gene of interest from an organism and inserting that gene into a **vector** DNA molecule (**figure 12.2**). The resulting recombinant DNA molecule is then introduced into a **host** organism for maintenance. The host organism is often the bacterium *Escherichia coli*.

From this point on, let us assume that we are interested in studying the insulin gene isolated from the human genome. The overall goal of our gene cloning experiment is to insert the human insulin gene into a vector DNA molecule and introduce the recombinant DNA molecule into the bacterium *E. coli*. In our scenario, the vector DNA molecule will function to:

- **Carry the insulin gene** (also called an **insert** or **cloned gene**). The insulin gene inserted into a vector can then be recognized and maintained by the host *E. coli* cell.
- **Copy the insulin gene**. Vectors are capable of efficient DNA replication within the host *E. coli* cell, and host *E. coli* cells often contain many copies of the vector. Thus, if the insulin gene is inserted in a vector, DNA replication of the vector will produce many copies of the insulin gene in each host *E. coli* cell.

**Plasmids**, small circular DNA molecules found in many bacteria, often serve as vector molecules (figures 12.1 and 12.2). These plasmids are not part of the bacterial chromosome. Plasmids contain:

- An **origin of replication**. The origin allows the plasmid to be replicated efficiently within the host cell. Some origins, such as *OriC*, allow replication of the plasmid within a particular host cell species (i.e., specifically the bacterium *E. coli*). Other origins allow efficient replication in different host species. The origin also determines the plasmid **copy number**. High copy number plasmids contain “strong” origins that allow frequent DNA replication to produce hundreds of plasmid DNA molecules per cell. Low copy number plasmids have less efficient origins, resulting in relatively few plasmid molecules per cell.
- **Unique gene cloning sites**. Gene cloning sites are plasmid DNA sequences that function as potential insertion sites for the cloned DNA sequence (i.e., the human insulin gene).
- **Selectable markers**. Plasmids contain selectable marker genes that confer an antibiotic resistant phenotype to the host bacterial cell. Thus, by growing host bacterial cells in the presence of an antibiotic, the researcher can ensure that the bacterial cells contain a plasmid. Common selectable marker genes include the *amp<sup>R</sup>* gene, which confers resistance to the antibiotic ampicillin, and the *tet<sup>R</sup>* gene, which confers resistance to the antibiotic tetracycline. In the case of the *amp<sup>R</sup>* gene, a bacterial cell that is resistant to ampicillin contains the plasmid; a bacterial cell that is sensitive to ampicillin does not contain the plasmid. Ampicillin resistant bacteria survive when grown on agar plates that contain ampicillin; ampicillin sensitive bacteria die when grown on agar plates containing ampicillin.

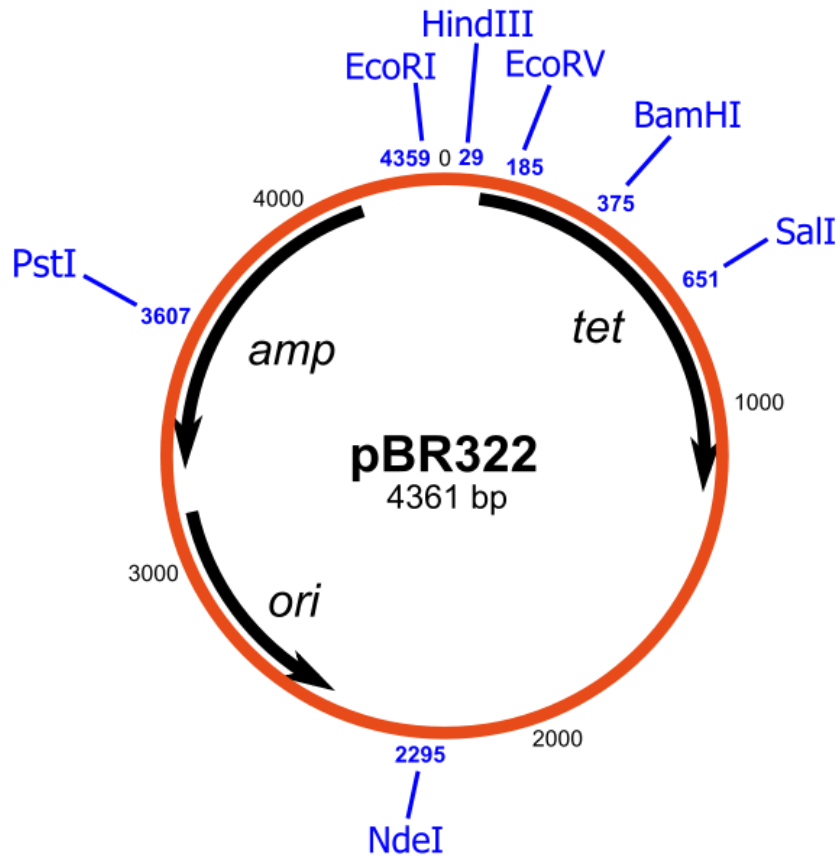


Figure 12.2 **The bacterial plasmid vector pBR322**. Important DNA sequences within the plasmid include the origin of replication (*ori*), the selectable marker genes (*amp*, *tet*) and gene cloning sites (*EcoRI*, *HindIII*, etc.).

## Key Questions

- What is a vector?
- Identify two types of molecules that can serve as vectors.
- Describe the functions of three types of DNA sequences that allow plasmids to serve as vectors.
- What is a host cell?

## Restriction Enzymes

How do we take the insulin gene from a human chromosome and insert it into a plasmid vector DNA molecule?

**Restriction enzymes** are important molecule tools used in the gene cloning process. Restriction enzymes are endonucleases that recognize specific DNA sequences (**restriction enzyme sites**) and cleave the phosphodiester bonds within both DNA strands. The restriction enzyme site is typically a **palindrome** DNA sequence. For example, the restriction enzyme **EcoRI** isolated from the bacterium *E. coli* cuts the DNA sequence 5'-GAATTC-3'. The complementary strand is 3'-CTTAAG-5', which is identical to the original restriction enzyme sequence but in the reverse orientation (i.e., a palindrome). *EcoRI* cuts both DNA strands within this palindromic DNA sequence between the G and A nitrogenous bases.

The natural function of a restriction enzyme is to protect the bacterial cell from foreign DNA, particularly bacteriophage DNA injected into the bacterial cell during a phage infection. Several hundred restriction enzymes have been isolated from bacteria and are available commercially for purchase.

## Key Questions

- What is a restriction enzyme and a restriction enzyme site?
- In terms of DNA sequences, what is meant by the term sequence palindrome?

## Producing Recombinant DNA Molecules

How can we use the restriction enzyme **EcoRI** to clone the human insulin gene into a plasmid vector DNA molecule?

Suppose that the restriction enzyme *EcoRI* recognizes the restriction enzyme sites shown below in both the insulin gene and in the plasmid DNA molecule (see **figure 12.3**) Note that the plasmid is cut by *EcoRI* at a single site, while *EcoRI* cuts on both ends of the human insulin gene (the target gene in the figure 12.3).

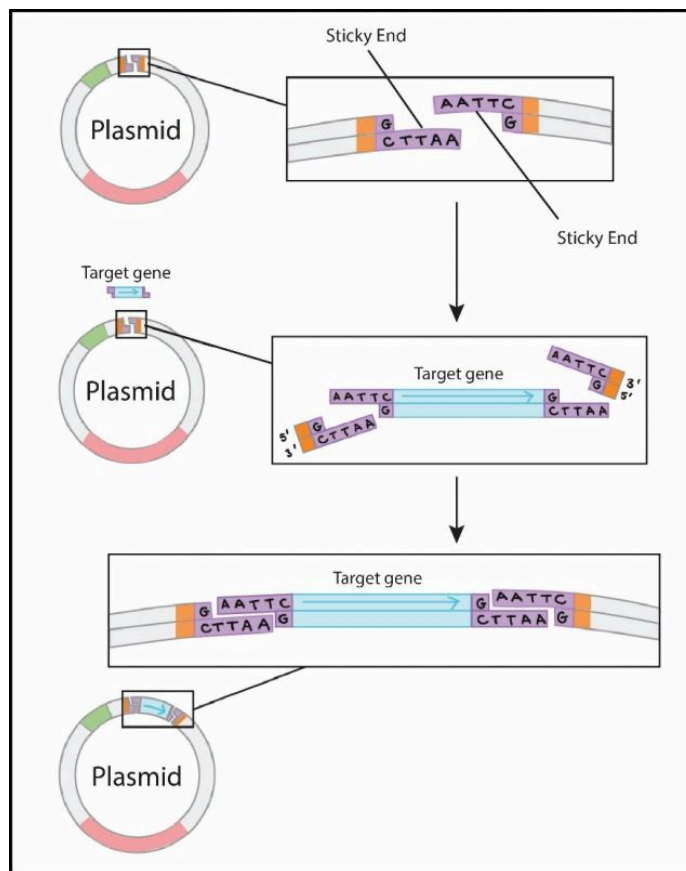


Figure 12.3 **Cut Site and Sticky Ends** - Image from Khan Academy and modified by SL

*EcoRI* cleaves the cloning site within the plasmid and the insulin gene to produce complementary single-stranded regions called **sticky ends**. When mixed, the sticky ends of the insulin DNA form hydrogen bonds with the sticky ends from the vector DNA. **DNA ligase** then forms the final covalent bonds in each DNA strand, covalently linking the insulin gene into the plasmid vector.

## Key Questions

- What are sticky ends?
- What is the function of DNA ligase in a gene cloning experiment?

## A Typical Gene Cloning Experiment

The entire procedure used to clone the insulin gene into a plasmid vector is outlined below (**figure 12.4**). For the purposes of this hypothetical experiment, let us assume that the plasmid DNA molecule contains a restriction enzyme site (cloning site), the *amp<sup>R</sup>* gene as a selectable marker, *OriC*, and the *lacZ* gene. The function of the *lacZ* gene will be described below.

1. **The plasmid vector and the chromosome containing the human insulin gene are cut with the same restriction enzyme.** A restriction enzyme cuts the chromosomal DNA into many small pieces; one of these chromosome pieces contains the insulin gene. The same restriction enzyme is used to cut the plasmid DNA within the cloning site. The DNA fragment containing the insulin gene and the plasmid DNA molecule have complementary sticky ends.
2. **The cleaved plasmid and the insulin gene are mixed.** Three different events can occur:
  - The plasmid sticky ends hydrogen bond with each other. This situation produces an intact plasmid molecule that does not include the insulin insert.
  - Chromosomal DNA fragments that do not include the insulin gene form hydrogen bonds with the plasmid sticky ends. The recombinant DNA molecule produced will contain the wrong insert.
  - The insulin gene fragment forms hydrogen bonds with the plasmid sticky ends.
3. **DNA ligase is added.** DNA ligase catalyzes the covalent linkage of the insert DNA fragment to the plasmid DNA molecule. Some of these recombinant DNA molecules contain the desired insulin gene insert.
4. **Transformation of host *E. coli* cells.** The recombinant DNA molecules are now introduced into *E. coli* host cells for maintenance. When *E. coli* cells are treated with calcium, the bacteria become **competent** to take up DNA from the environment. When the recombinant DNA molecule is added to these competent bacterial cells, and the bacteria are shocked by a brief heat treatment, the recombinant DNA molecule is taken into the cytoplasm of the *E. coli* (**transformation**).
5. **Host bacteria are grown on agar plates that contains ampicillin.** Two scenarios are possible after transformation:
  - **Some *E. coli* cells were not transformed** (i.e., do not contain a recombinant DNA molecule). These bacterial cells cannot grow on agar plates containing ampicillin because the bacteria lack the *amp<sup>R</sup>* gene.
  - **Some *E. coli* cells were transformed.** Since the plasmid vector contains the *amp<sup>R</sup>* gene, transformed cells are now resistant to ampicillin. As a result, these transformed cells grow on agar plates that contain ampicillin. It is important to note that this population of growing bacteria contains three types of plasmids:
    - Some plasmids lack inserts altogether.
    - Some plasmids contain other chromosomal DNA fragments as the insert (i.e., the wrong insert).
    - Some plasmids contain the insulin gene as the insert (i.e., the correct insert).
6. **Identify colonies that contain an insert by the blue-white screening method.** Many cloning experiments are designed so that the restriction enzyme cutting site within the plasmid is located the *lacZ* DNA sequence. The *lacZ* gene produces the enzyme  **$\beta$ -galactosidase**. Cloning the insert disrupts the *lacZ* gene, preventing  **$\beta$ -galactosidase** production. **Disruption of the *lacZ* gene** allows researchers to distinguish bacteria that contain an insert versus bacteria that do not contain an insert.
  - **Recombinant plasmid vector without an insert.** In this case, the plasmid contains an intact *lacZ* gene. The plasmid *lacZ* gene produces  $\beta$ -galactosidase.
  - **Recombinant plasmid vector that contains an insert.** Since the presence of an insert disrupts the *lacZ* gene, no  $\beta$ -galactosidase is produced.

How do we determine if  $\beta$ -galactosidase is produced? This is done by plating the bacterial cells on an agar plate that not only includes ampicillin but also **IPTG**, a chemical that activates the *lacZ* gene to produce the  **$\beta$ -galactosidase** protein. The agar plate also contains **x-Gal**, a chemical substrate for the  $\beta$ -galactosidase enzyme.

- **$\beta$ -galactosidase is expressed (no insert).** The bacterial colony is blue as  $\beta$ -galactosidase converts X-Gal, which is colorless, into a blue product.
- **$\beta$ -galactosidase is not expressed (insert present).** The colony is white because  $\beta$ -galactosidase is nonfunctional and thus cannot convert X-Gal into a blue product.

7. **Identify colonies that contain the insulin insert.** All the white colonies contain an insert; however, at this point, we cannot distinguish the white colonies that contain the insulin insert from the white colonies that have other DNA fragments as inserts. Several techniques can be used to identify which bacterial cells contain a recombinant vector with an insulin gene. For example, vectors isolated from white colonies can be cut with the same restriction enzyme used at the beginning of the cloning experiment to release the insert. The digested DNA can then be analyzed by agarose gel electrophoresis. The presence of a DNA fragment of the appropriate size for the insulin gene indicates that the chosen colony likely contains the insulin insert. Alternatively, the polymerase chain reaction (PCR), using primers specific for the insulin gene can be used to amplify the insulin insert. If a PCR product is produced using these insulin primers, then the insulin gene was cloned successfully. Finally, determining the nucleotide sequence of the cloned insert by DNA sequencing will determine if the recombinant molecule contains the insulin gene.

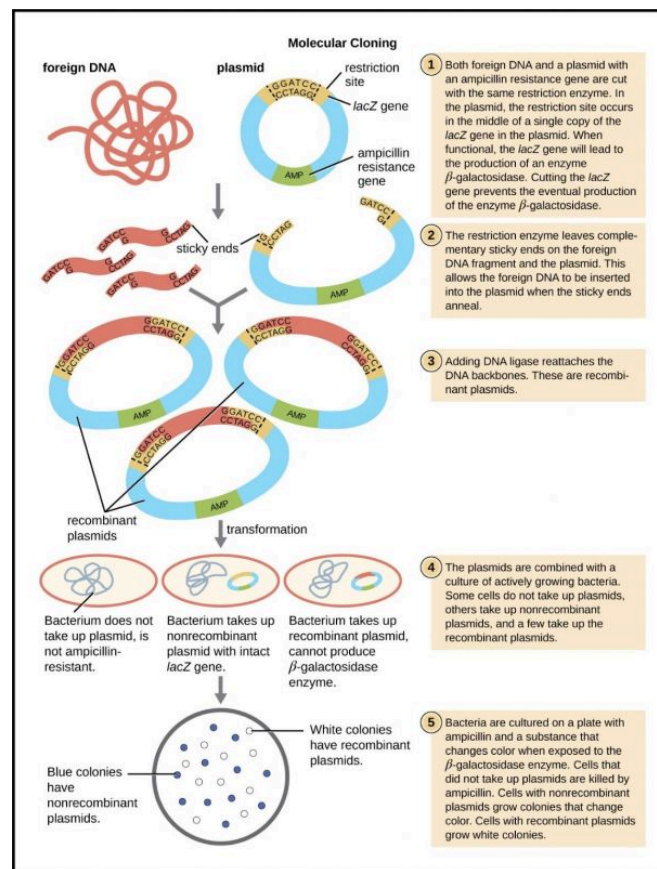


Figure 12.4 **Gene cloning overview** --- This image is used from OpenStax (access for free at <https://openstax.org/books/biology-2e/pages/1-introduction>)

## Key Questions

- Describe the seven steps involved in cloning the insulin gene into a bacterial plasmid.
- How are bacterial cells that were not transformed eliminated in the cloning experiment?
- How can blue-white screening be used to identify recombinant DNA molecules that contain an insert?
- How can you identify the recombinant DNA molecule that contains the insulin insert?



This content is provided to you freely by BYU-I Books.

Access it online or download it at

[https://books.byui.edu/genetics\\_and\\_molecul/23\\_\\_gene\\_cloning](https://books.byui.edu/genetics_and_molecul/23__gene_cloning).