CH 3. CLONING OF NUCLEIC ACIDS

I. DEFINITION AND AIM

DNA cloning is a fundamental technique in recombinant DNA technology. It can be defined as a method of isolation and amplification of precise DNA fragments so that billions identical molecules are pure in a tube and can be used in subsequent experiments. It is not wrong to consider cloning of a gene as its purification. The aim of cloning is to allow the study of the cloned molecule in terms of sequence and function. It is also common to clone a nucleic acid molecule in order to use it as a probe. In other words, cloning is necessary to study all related gene aspects as purification of proteins is necessary to study their functions.

Cloning involves construction of hybrid DNA molecules (recombinant) that are able to self-replicate in a host cell, usually bacteria. This is accomplished by inserting DNA fragments into a plasmid or bacteriophage cloning vector, introducing the vector into bacterial cells where it is amplified by the bacterial replication machinery. The inserted DNA fragment, or insert, can be derived from any organism and obtained from genomic DNA, cDNA, previously cloned DNA, PCR or it can be synthesized *in vitro*.

The process of cloning relies on use of DNA modifying enzymes, the most important ones are: REs and ligases. Cloning is not a single step experiment[18], it involves library construction by the following tasks: nucleic acid extraction or preparing DNA fragment(s) to be cloned, preparing vector for cloning, joining the fragment(s) with the vector, introducing recombinant vector(s) into bacteria (transfection), screening and selecting for cells with vector.

II. BRIEF DESCRIPTION OF VECTORS

Replication property of DNA is the basis of cloning process. The replication machinery of host (bacteria) cells must be used during cloning since *in vitro* replication is not efficient especially when heterogeneous fragments are being amplified. In fact, DNA extraction yields DNA fragment of different lengths and most of them lack the replication origin. Therefore, DNA fragments are not capable of self-replication in bacterial cells unless they are integrated in an autonomously replicating bacterial element (cloning vectors). These are man-made replication units (recombinant) derived from extrachromosomal bacterial replicons such as plasmids, bacteriophages or viruses.

The most commonly utilized classes of vectors are: plasmids, bacteriophage genomes, cosmids, shuttles and YAC (yeast artificial chromosome). The choice of vector largely depends on size of fragments to be cloned, on the host cell and on the planed selection and screening strategies.

1. Plasmids

Plasmids are circular dsDNA (Figure 3.1) naturally present in certain bacteria as extrachromosomal elements (2 to 5 kb in size). Replication of plasmids may depend on that of the host's chromosome or not. They can be horizontally transmitted by conjugation. Plasmids can be experimentally introduced (transfection[19] by different ways) into bacteria which acquire the selectable marker carried by the plasmid. Plasmids are stable unless their size exceeds 10 kb. If a plasmid is more than 10 kb, its penetration inside a host cell becomes very difficult. Most plasmids can not accept the insertion of DNA molecules more than 9 kb in length.

Natural plasmids have been modified by addition and/or removal of particular sequences in order to make them more suitable to molecular biology use. Several features, listed below, are necessary to make a plasmid useful as vector.

a. Replicon

Replicon comprises a replication origin sequence and the genes encoding RNAs or proteins that are necessary for control of plasmid replication. High-copy-number plasmids are the most frequently used plasmids in cloning. They accumulate 20 or more copies per bacterium whereas low-copy-number plasmids have less than 20 copies per cell. Most recombinant plasmids contain either the ColE1 (e.g. pBluescript) or closely related pMB1 replicon (e.g. pBR322, pGEM and pUC).

[18] It is possible to clone DNA molecule by PCR (directly without a library) if sequence of specific primers is available.
[19] The procedure for introducing exogenous donor DNA into recipient cells is called transfection.

b. Selectable markers

Selectable markers are usually genes conferring resistance to antibiotics or metabolic advantage[20] (Figure 3.1). The selectable marker maintains the presence of the plasmid in cells. In molecular biology, a selectable marker helps distinguishing cells transformed with the vector from those that are not. Most plasmids carry a β-lactamase gene, the product of which inactivates penicillin and its derivative ampicillin. Other genes, *tet*^{R} and *cam*^{R} are for tetracycline resistance and chloramphenicol resistance, respectively. When transformed cells are grown in the presence of antibiotic, only cells carrying the antibiotic-resistant plasmid survive.

c. MCS

These are the cloning sites named Multiple Cloning Site (MCS, Figure 3.1) or polylinker cloning region where a number of RE cleavage sites are adjacent to each other. These sites are chosen to be unique in the vector sequences. Thus, recombinant fragments can be easily introduced into the circular vector by digesting the vector with one or two enzymes present only in the MCS and ligating the desired fragment into it.

d. Active promoter

Plasmids used in molecular biology can be divided into cloning vectors and expression vectors. The main difference between the two types is the presence of active promoter sequence(s) flanking the MCS such as promoter sequences for SP6, T7, or T3 RNA polymerases (Figure 3.1). These promoters can be used for *in vitro* or *in vivo* production of large quantities of RNA transcripts from DNA inserted into the polylinker. For example these sequences are present in pBluescript and pGEM plasmids but are absent from pUC plasmids.

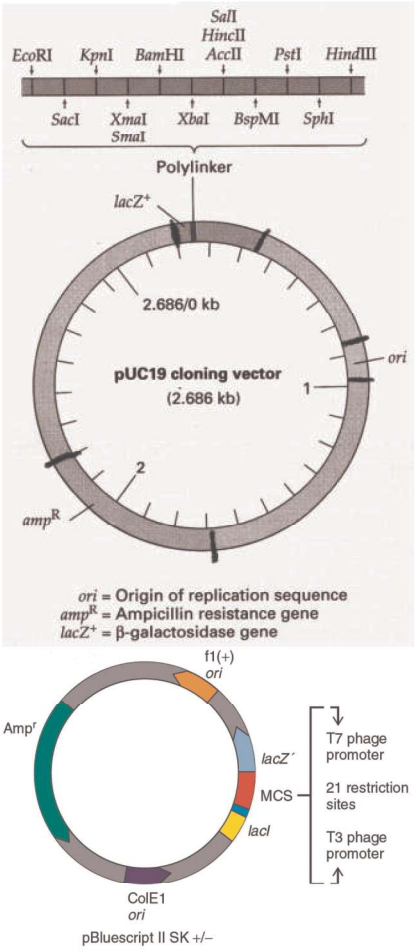

Figure 3.1: Examples of plasmids used as vectors.

e. Screening tool

The screening tool permits recognition of colonies transformed with vectors containing insert from those transformed with vector without insert. The most commonly used method is the insertional inactivation of β-galactosidase enzyme. The vectors have a MCS inserted into the lacZ gene sequence that encodes the first 146 amino acids of β-galactosidase. β-galactosidase metabolizes a chromogenic substrate (X-gal) into a blue product (blue colonies if enzyme is active). Insertion of DNA fragment into the polylinker region disrupts the open reading frame of β-galactosidase (Figure 3.1). Thus, cells carrying insert (cDNA) in the plasmid are deficient in that enzyme which makes them rapidly identifiable by the white color of colonies.

[20] For example genes for amino acid synthesis. Such vectors are used in auxotroph mutant hosts.

Other vectors are suitable for direct positive selection of plasmids containing inserts rather than screening for them. In these vectors, the polylinker (MCS) is introduced in a gene whose product causes cell toxicity and death. Insertional inactivation of the toxic protein (by insertion of a cloned DNA fragment) allows the cell to survive. This strategy enables positive selection for cells carrying recombinant plasmid (with insert).

2. Bacteriophage genome

This category of vectors does not have the properties of plasmids. The most commonly utilized vector is bacteriophage lambda genome. The vector consists of a modified version of the genome which always leads to a lytic pathway (lysogeny is inactivated). Native genome consists of 45 kb divided into three regions. The two lateral regions are essential for the lytic viral cycle whereas the central one, ~15 kb, is not (Figure 3.2). Removal of the central region allows the vector to receive an insert between 8 and 22 kb.

Bacteriophage lambda DNA may be linear or circular since it has cohesive *cos* sites at its extremities which mediate circularization and formation of concatemer[21] during the normal replication cycle. Moreover, proteins which help packing into capsids interact with *cos* sites.

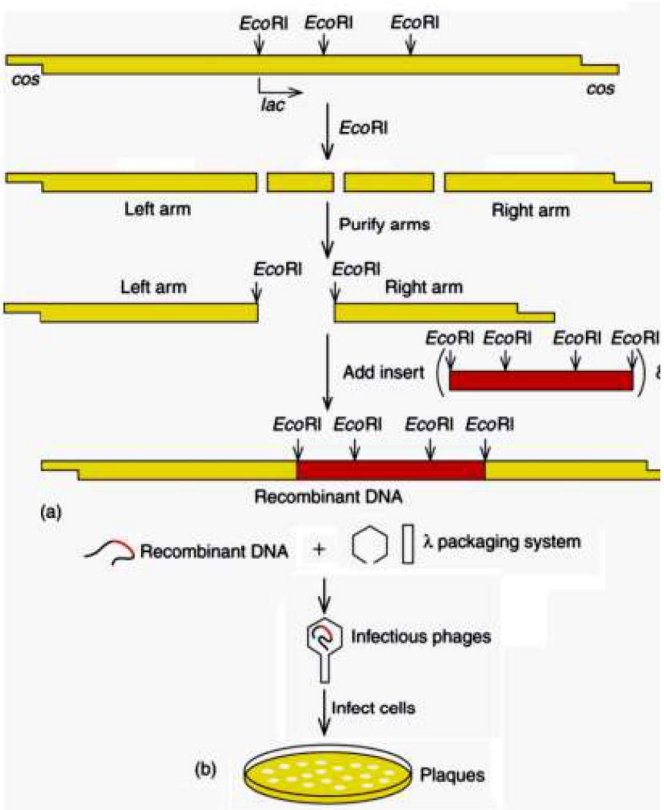

Figure 3.2: Lambda bacteriophage genome utilized as a vector in DNA library.

There is no selectable marker in this vector and its presence is not useful since phages infect host cells leading to their lysis. The vector is encapsidated *in vitro* by incubation with capsid proteins purified from two complementary mutant lysogenic bacteria. Recombinant phage DNA molecules of more than 50 kb are unstable and can not be packed during *in vitro* encapsidation.

Different experimentally modified generations (by modification of the original genome) of this vector exist. Other phage genomes are used as vectors in DNA technology. For instance, the M13 phage genome which is 6.4 kb is characterized by existing in two forms during its reproduction cycle: the ds form and the ss one.

3. Cosmids

Cosmids are chimera cloning vectors (size ≈ 5 kb) experimentally designed by combination between phage lambda fragment (*cos* sites) with plasmid sequences (*ori* and *ampR*) and a MCS. The advantages of cosmids is their high capacity of receiving DNA (up to 30 kb) and their ability to be packed into capsids by *in vitro* encapsidation (because of their *cos* sites) and delivered to host cells by natural infection. Once inside a bacterium, they behave as plasmid and do not lead lytic pathway.

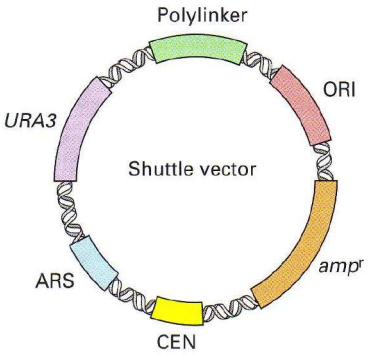

Figure 3.3: Shuttle vector.

4. Shuttle vectors

Shuttle vectors, unlike the previously described vectors which are restricted to E. coli, can be used in both bacteria and eukaryotic cells (yeast). A shuttle vector is circular dsDNA carrying *ori* sequence (replicable in E. coli), *ampR* (for selection in bacteria) as well as an autonomously replicating sequence (ARS), which functions as an origin for DNA replication in yeast, a yeast

[21] A concatemer consists of numerous copies of the DNA which are linked head to tail.

centromere (called CEN), which allows faithful segregation of the plasmid during yeast cell division, and a yeast gene encoding an enzyme for uracil synthesis (URA3), which serves as a selectable marker in an appropriate yeast mutant (auxotroph *ura3* mutant) and polylinker to perform insertion (Figure 3.3). Certain shuttle vectors are not provided with a replication origin in yeast, yet they replicate after their integration into the host chromosome.

5. YAC

Because of the limited cloning capacity of the other vectors (several tens of kb at most) YAC were developed to receive and clone up to 1000 kb (one Megabase) and even chromosome fragments. YAC are linear structures comprising yeast telomeres at both ends, yeast centromere sequence (CEN), selectable markers (URA3 or TP1 for tryptophan synthesis), an ARS (auto replicating sequence) and a multiple cloning site.

Other artificial chromosomes are available such as the bacterial artificial chromosome derived from E. coli plasmid and genome.

III. Library of cDNA

A cDNA (complementary DNA) library is a set of cDNA clones prepared from mRNAs isolated from a particular type of tissue[22].

Construction of a cDNA library is a highly sophisticated procedure. The term cDNA refers to DNA molecules which are complementary to mRNA molecules and therefore differ from their respective genes by lacking the promoter and all regulatory sequences (upstream and downstream) as well as the introns. Cloning of a cDNA may be accomplished by the steps described next.

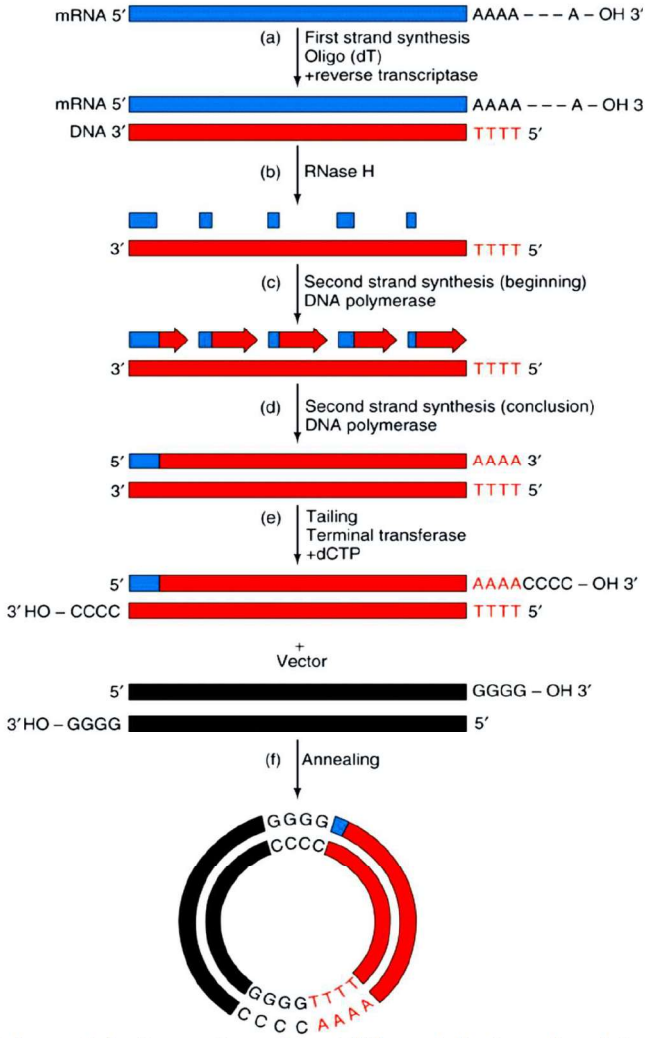

Figure 3.4: Generation of a cDNA and its insertion into vector.

1. cDNA preparation

This step is started by preparing a pure sample of mRNA which contains a heterogeneous population of mRNA corresponding to all expressed genes[23] in the utilized cell type or tissue. The diverse mRNA molecules are not equally abundant since they correspond to genes not expressed at the same rate. cDNA are prepared by retrotranscription of all mRNA in the sample using an oligo-dT primer and the viral reverse transcriptase (Figure 3.4). The hybrid molecules produced are mildly treated with RNase H which digests RNA molecules involved in ds form. Consequently, certain RNA fragments are kept and serve as primers for the activity of a DNA polymerase. A ligase may be used to join the fragments.

Certain reverse transcriptases are able to polymerize nucleotides at the 3'end without template. The added nucleotides are long enough to form a loop that serves as primer for polymerization of the second cDNA strand. In such a case, digestion with RNase H may be complete and the resulting cDNA should be treated with Nuclease S1 in order to remove the loop.

[22] The cDNA libraries prepared from different tissues, different cell types or even from the same cell type but placed in different conditions, will be significantly different, except for cDNA of house keeping genes that are found in any library prepared from any tissue. However, the genes whose expression is tissue-specific will differ.

[23] Protein-coding genes, except histones mRNA that are not polyadenylated.

2. Insertion of cDNA into vector

There are two types of ligation procedures: ligation of fragments with compatible cohesive ends and ligation of fragments with blunt termini. Since ligation efficiency of blunt ended DNA molecules is low, it is therefore better to generate cohesive ends. Generation of ligation-compatible cohesive ends in both plasmid and DNA fragments can be performed by cleaving both vector and fragment DNA with the same RE (or compatible). This approach is not suitable since we deal with a heterogeneous population of cDNA except in the poly-dA/poly-dT region[24].

It is therefore indispensable to add at the extremities of every cDNA oligonucleotide adaptors named linkers which contain suitable restriction site. The chosen linker must carry the same restriction site as that utilized for vector linearization. Addition of linkers is possible although cDNA extremities are blunt end similarly to those of linkers. In fact, the ligase can add linkers (present in high concentration) to both ends. Ligation of linkers is followed by their digestion[25] with the appropriate RE, purification of cDNA and then their ligation with the linearized vector (Figure 3.5).

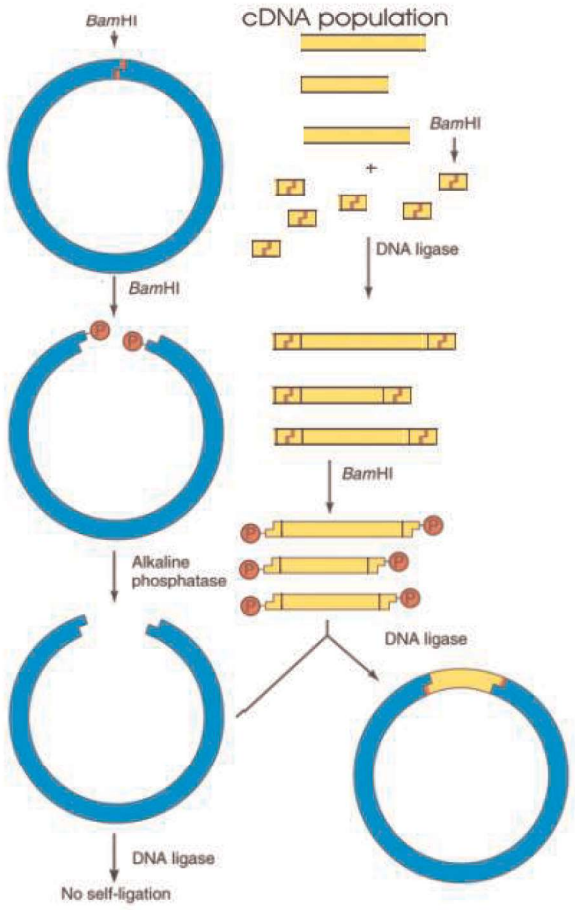

Figure 3.5: Principle of insertion in a plasmid vector by use of linkers.

During ligation, DNA ligase catalyzes formation of phosphodiester bonds between adjacent nucleotides only if phosphate is present at the 5'end of DNA molecules. Removing the 5'phosphates (by bacterial alkaline phosphatase or calf intestinal phosphatase) from both ends of the linearized plasmid DNA and leaving them on the fragment to be cloned can therefore minimize recircularization of the plasmid without insert.

Another procedure of insertion consists of the addition of poly-dC on the cDNA population and poly-dG on the linearized vector (Figure 3.4). Such a polymerization is catalyzed by the terminal deoxytransferase (TdT) which works without template. Incubation of the cDNA population with the prepared vector in presence of ligase leads to insertion in the vector.

It is obvious that insertion of the cDNA by the previous methods occurs in two possible orientations. In case one wants the cDNA to be expressed into protein (expression vector with a promoter upstream the MCS), the insert must be correctly oriented, otherwise transcription will result in an antisense RNA.

By the end of insertion (regardless of its method) we get vectors which are different since they carry different cDNA types. An abundant mRNA is reflected by an abundant cDNA and therefore numerous vectors carrying it. Obviously, not all vectors succeed in inserting a cDNA.

3. Transfection of bacteria (first amplification)

The recombinant vectors must be introduced into host cells (E. coli) in order to be amplified by replication. Experimental introduction of nucleic acids into cells is known as transfection. Competent[26] bacteria without plasmid are transfected by different methods such as electroporation and chemical transformation-heat chock. Electroporation involves a brief application of high voltage electric field to the cells at low temperature resulting in the formation of transient holes in the cell

[24] It is impossible to digest the cDNA themselves to generate cohesive ends since this leads to partial loss of sequence. Moreover, their sequences and restriction maps are unknown.
[25] It is important to have inactivated the potential restriction sites present in the cDNAs by methylating them (methylases).
[26] E. coli cells used for cloning do not possess a natural mechanism for DNA uptake (not competent). Competence can be induced by chemical means.

envelope through which plasmid DNA can enter the cell. Quick restoration of membrane fluidity and closing pores is crucial for cell survival after the pulse.

Chemical transformation involves exposure of competent bacteria to hypotonic solution of calcium chloride at 4°C and a brief heat shock at 42°C. Cations neutralize the negatively charged phosphates of lipids in the membrane whose fluidity is decreased (because of low temperature) which stabilizes interaction with cations. Cations form complex also with negatively charged phosphates of DNA (added to bacteria) thereby mediating adhesion between DNA and bacterial envelope. During the heat shock (brief incubation at 42°C) and the return to low temperature, an inward flow through the fragile zones of the envelope causes entrance of plasmids into the cell. Note that some chemical transformation protocols do not use heat shock step.

Since the number of real plasmid molecule is not in large excess, one plasmid type enters a cell and replicates. However, not all bacteria receive a plasmid and therefore a selection step is necessary. Transfected bacteria are then spread on a semi-solid agarose culture medium and incubated at 37°C (many culture dishes are utilized). Spreading must be efficient in order to separate the bacteria from one another so that they form independent colonies. The colonies present on the culture dishes are supposed to contain different recombinant vectors, each containing a specific cDNA. The set of colonies on the culture dishes is the cDNA library.

A complete cDNA library theoretically contains all cDNA clones corresponding to all mRNAs expressed in a cell or tissue type. An expression cDNA library refers to one in which all cDNAs are cloned in the sense orientation so that all the cDNA clones in the library can be induced to express their mRNAs and proteins.

4. Screening of the library and selection

There are three levels of selection and screening which must be performed. The first one is the selection of bacteria that have acquired plasmids. Second, it is the selection of bacteria hosting a plasmid with an inserted cDNA. Third, one must screen for the colony carrying the cDNA of interest.

a. First level

Selection procedure is actually planned at the moment of vector choice. One decides about the vector to use and the site of cDNA insertion so that selection and screening are possible. The plasmid contains a gene for resistance to antibiotics and therefore addition of the appropriate antibiotic to the gelose culture medium ensures the selection of bacteria having acquired a plasmid. Only those bacteria can grow and form colonies, the others can not divide.

b. Second level

The second selection level is dependent on the site of insertion and in known as insertional inactivation. The MCS is found in a gene whose product is easily seen so that one can visually distinguish the colonies carrying plasmids with cDNA. The most widely used is lacZ gene (Figure 3.1, see II. 1. e.).The second example is that of insertional inactivation of a gene carried by the plasmid and encoding a protein that causes cell death (see II. 1. e.).

c. Third level

The third selection level consists of picking up among the thousands of colonies the one whose vector carries the cDNA we seek to isolate and clone. This selection step requires a specific tool such as an oligonucleotide probe specific to the cDNA in question or an antibody specific to its corresponding protein. In fact, such tools may be available if the protein is already purified. Usually, one proceeds to cDNA cloning after having enough information about the protein. It means that purification of the protein precedes cloning of the corresponding cDNA[27]. A pure protein can be partially sequenced and can be used to prepare specific antibodies.

[27] This rule has many exceptions, especially in species whose genomes are entirely sequenced.

Antibodies can be used in screening the library by immunoblotting. Briefly, replica of all library culture dishes are prepared, that is colonies are partially and independently transferred onto a disc of nitrocellulose (Figure 3.6), lysis of the cells is then carried out and the nitrocellulose discs are incubated with the antibody. Revelation of the antibody is performed (by incubation with a labeled antibody against the first one) and the colonies expressing the protein are identified on the library dishes. Obviously, the cDNA must have been inserted in an expression vector in the correct orientation in order to allow its expression into protein. Moreover, the expressed proteins must fold properly[28] in order to acquire the same antigenic properties as the native protein.

When an oligonucleotide probe is available the library is screened by in situ plaque hybridization (Figure 3.6). The sequence of the oligonucleotide probe is designed based on the partial polypeptide sequence of the protein in question by reverse translation. Moreover, oligonucleotide probes may be taken from orthologous genes.

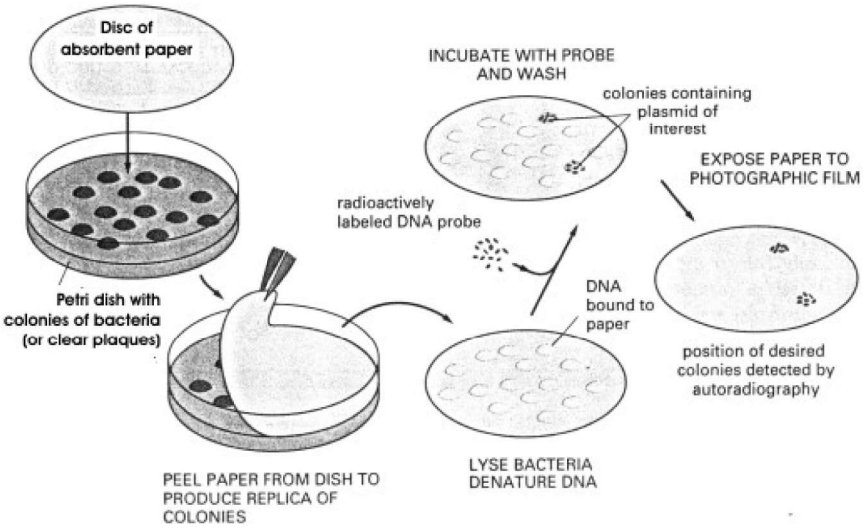

Figure 3.6: Preparation of replica of culture dishes (colonies or clear plaques) of a library and screening for the desired clone.

Replica of all library dishes are prepared and cells are broken and their nucleic acid content denatured by alkaline treatment. Chemically synthesized probe(s) is (are) radioactively labeled. The discs are hybridized with the labeled probes. Autoradiography helps identifying the colonies which contain the cDNA matched by the probe (Figure 3.6).

Note: Size of the oligonucleotide probe depends on the genome size of the species in which investigation is conducted. It is usually at least 15 or 18 nucleotides in order to exclude the possibility of hybridization with DNA other than the cDNA in question. In fact, a sequence of 18 nucleotides has one chance over 4^{18} to be repeated in a genome which is 4^{18} (more than 68 10^9) bp in size. A sequence of 15 nucleotides has one chance over 4^{15} to be repeated in a genome which is 4^{15} (about 10^9) bp in size. Therefore, a probe of 18 nucleotides is highly specific to screen a library of any species whose genome size is 4^{18} bp or less.

5. Amplification

The identified positive colony(ies) is (are) picked up and cultured in a liquid medium in the presence of the selective antibiotic. Cells are then harvested and their plasmid extracted and purified as already presented. Thus, a unique pure vector carrying one type of cDNA is found in the tube. The cDNA can be released from the vector by digestion with the appropriate RE (used for insertion) and can be separated from the vector by means of electrophoresis. Further analyses are required in order to confirm identity of the cloned cDNA. Actually, in addition to determination of size and restriction map, one should determine the cDNA sequence in order to be sure about its relationship with the protein and to compare it to previously cloned and published sequences in the diverse databases (e.g. Genbank). Comparison is necessary in order to see whether other researchers have already cloned the same DNA and also to determine homology of the cloned cDNA with previously cloned ones.

[28] Eukaryotic proteins expressed in prokaryotic cells do not necessarily undergo correct maturation and folding because prokaryotes have no glycosylation enzymes which play a crucial role in protein maturation and antigen formation.

IV. GENOMIC LIBRARY

Unlike a cDNA library, a genomic library comprises all genome sequences and therefore any part of the genome can be cloned. A genomic library is a set of clones carrying restriction fragments produced by cleavage of the entire genome. In order to prepare a genomic library one follows the same principles as for cDNA library. However, it does not start from RNA and the tools are slightly different. The steps are presented below.

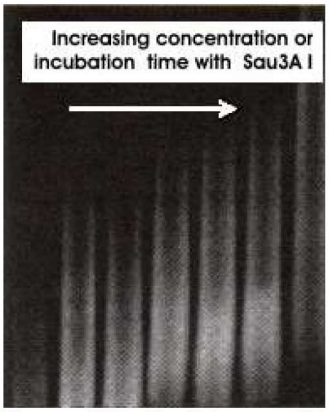

Figure 3.7: Partial or mild digestion of genomic DNA.

1. Mild digestion and insertion

Genomic high molecular weight DNA is extracted in a soft way in order to avoid its shearing[29]. The extracted DNA is mildly digested with the same RE[30] as that for preparation of the vector. Mild digestion produces overlapping DNA fragments having different sizes. It is performed by digesting with low enzyme amount which is determined after trial of several concentrations. Among them one chooses the one that gives the best range of DNA fragment size (Figure 3.7, Figure 3.8). For instance, if the gene to be cloned is 20 kb in size, one should perform mild digestion so that fragment sizes range between 15 and 25 kb. This range ensures that the gene in question will be complete on a single fragment.

The digested fragments are inserted at the place of the middle region of phage lambda genome (or in another vector) as shown in Figure 3.2. Ligase activity joins the inserted fragments with the vector which forms concatemer. The resulting vectors with inserts are named recombinant vector. Each copy of vector carries a piece of the digested genomic DNA.

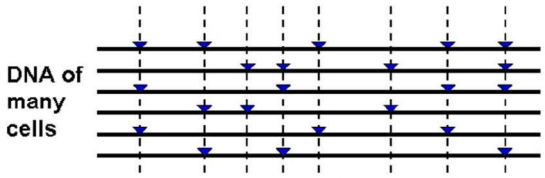

Figure 3.8: Principle of mild digestion. Horizontal lines refer to DNA of many cells. Vertical dashed lines refer to the RE sites. Heads of arrow refer to sites that are digested (mild digestion).

2. *In vitro* encapsidation

Incubation of the recombinant vectors with viral capsid proteins derived from two differently mutant lysogenic bacteria will results in packing of the genomic fragments and their vectors into viral particles. This is mediated by the interaction between *cos* sites at the phage DNA extremity and capsid proteins. Capsid will help natural delivery, by infection, of the vector into host cells for amplification.

3. Clear plaques

The encapsidated viral particles carrying the diverse genomic fragments are spread on many dishes containing bacterial lawns. A bacterial lawn is a semi-solid culture medium covered with a uniform layer of host bacteria (e.g. E. coli) which will be infected by the viral particles that reproduce and lead a lytic cycle. Because the medium is semi-solid, the released viral particles infect just the neighboring bacteria so that a clear plaque appears at the place of every viral particle (Figure 3.9). A clear plaque corresponds to broken cells and a colony of phages carrying exactly the same recombinant DNA. The library is the thousands of clear plaques present on all dishes.

[29] Shaking and pipeting should be avoided since they cause DNA fracture.

[30] Choice of the RE to perform mild digestion is based on the abundance of its site. RE that cut frequently are not preferred since they produce short fragments. An estimation of the number of cleavage sites for a RE within a given stretch of DNA, assuming even distribution of bases, is described by the equation: site number=$N/4^n$, where N is the number of base pairs in the DNA fragment and n is the number of bases in the RE site.

4. Selection and screening

A genomic library is screened according to the same principle as cDNA library screening. It is based on use of nucleic acid probe or specific antibody. If the cDNA has already been cloned, it will serve as a probe. Otherwise, an oligonucleotide probe can be designed after the partial polypeptide sequence and reverse translation. The process is similar to that of cDNA library. Replica of all dishes (all clear plaques) of the library are made and treated to denature the nucleic acids. After hybridization with the probe autoradiography reveals the positive clear plaques (those containing phages carrying the gene in question). The phages are then picked up and amplified by performing infection of large bacterial culture. The recombinant bacteriophages proliferate and cause host cell lysis. The bacterial culture eventually turns to phages only because of their lysis. The phages are harvested by centrifugation and their DNA is extracted and the gene released from the vector by digestion with the appropriate RE.

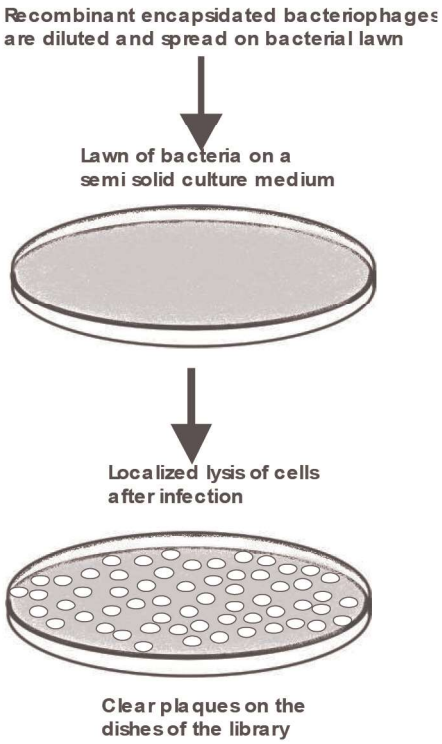

Figure 3.9: Formation of clear plaques by phage vector in the DNA library.

V. Selection by Functional Complementation

In some cases a DNA library can be screened for the ability to express a functional protein that complements a recessive mutation in the host cell. The process is referred to as functional complementation. This selection strategy is an efficient way to isolate a cloned gene without the heavy procedures of screening and selection. This approach is possible only if the protein, whose gene is being cloned, has a homologous counterpart in the host cell of the library. Another condition is the availability of a mutant strain (of host cell) for the gene in question and which can not grow in specific medium unless the gene being cloned is supplied by transfection and expressed.

To illustrate this method, we take the example of cdc28 cloning. Yeast genomic library inserted in a shuttle vector and introduced into mutant yeast cells. In fact, the yeast host cell is a double mutant strain that requires uracil for growth due to ura3 mutation and carries cdc28 thermo-sensitive mutation. Since the vector carries a copy of URA3 gene, only transformed (transfected) cells will grow in the absence of uracil. To select the colonies containing vectors with CDC28 gene, cells are passed from the permissive temperature 23°C (permits growth of the cdc28 mutant) to the nonpermissive temperature 36°C (no growth of cdc28 mutants). Only yeast colonies that carry recombinant plasmids expressing a wild-type copy of the CDC28 gene will be able to grow at 36°C, the remaining cells die. Temperature-resistant yeast colonies are identified and their plasmid extracted (after large scale culture) and further analyzed by subcloning[31] and DNA sequencing.

VI. Characterizing and using cloned DNA

Now that we have described the basic techniques for cloning DNA, we consider how cloned DNA are further characterized and the various ways in which they can be used. We start here with several widely used general techniques. More specific applications are given in the following chapters.

In order to manipulate or sequence a cloned DNA fragment, first it must be separated from the vector DNA. This can be accomplished by cutting the recombinant DNA clone with the same RE originally used to produce the recombinant vectors. The vector and its insert are sorted by gel electrophoresis. If desired, the band may be excised from the gel, nucleic acids purified and used in various manipulations (sequencing, directed mutation,...).

Once a cloned DNA fragment, especially a long one, has been separated from the vector DNA, it is often treated with various REs to yield smaller fragments. After separation by gel electrophoresis, all or some of these subfragments can be ligated individually into vector and introduced into host cell

[31] Cloning of smaller fragments of the cloned gene.

(subcloning). Subcloning is necessary to investigate the cloned gene. For instance, fragments of an appropriate length (about 500 nucleotides) are necessary for sequencing. Many other techniques use the cloned DNA and will be considered in the coming chapters.

VII. RNA CLONING

Obviously, RNA molecules can be cloned only after cloning of their corresponding DNA. A powerful approach for the analysis of gene structure and function has become possible with the development of RNA synthesis *in vitro* from cloned DNA templates, a process known as run-off *in vitro* transcription (Figure 3.10). The RNA polymerases encoded by the Salmonella typhimurium, bacteriophage SP6 and E. coli phage T7 are very efficient and have characteristics suitable for *in vitro* transcription.

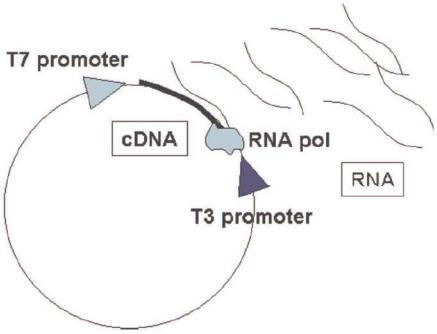

Figure 3.10: RNA cloning by run-off in vitro transcription.

The DNA is inserted in an expression vector such as pGem (Figure 3.1) in front of an active promoter such as the T7 bacteriophage promoter. The expression vector is then incubated with the appropriate RNA polymerase and all necessary factors (NTP, buffer,…) and *in vitro* transcription is performed.

The produced RNA molecules are then extracted and utilized for different purposes such as *in vitro* translation using lysate of reticulocytes, investigation of splicing and processing, RNA foot printing (RNA-protein interaction), sequencing, analysis of catalytic RNAs and design of probes. The use of transcription products (ssRNA) as asymmetric probes in Southern and Northern blotting is more sensitive than use of the corresponding cDNA. Moreover, they can be used as antisense RNA to affect expression of a specific target gene.