The cDNA Library

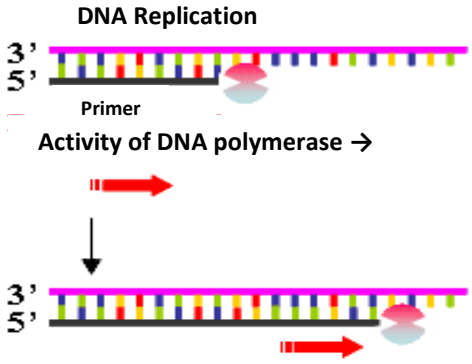
(NOTES)

Complementary DNA (**cDNA**) is double-stranded DNA synthesized from a single stranded RNA template (e.g. **messenger RNA** or **microRNA** template) in a reaction catalyzed by the enzyme reverse transcriptase.

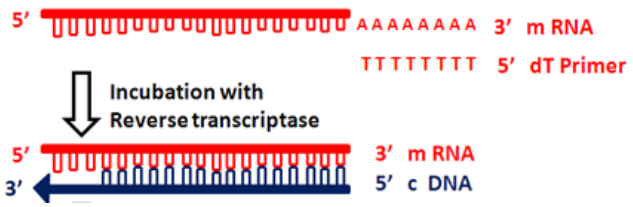
MicroRNA (**miRNA**) is a small non-coding RNA molecule (containing about 22 nucleotides) found in plants, animals and some viruses, which functions in post-transcriptional regulation of gene expression.

The **cDNA** is created from a mature mRNA (originated from a **eukaryotic cell**) with the use of **reverse transcriptase** enzyme. In eukaryotes, a **poly-(A) tail**, consisting of a long sequence of adenine nucleotides, facilitates to distinguish the mRNA from **tRNA** and **rRNA**; it can therefore be used as a **primer** site for reverse transcription.

Primer is a short strand of DNA or RNA that serves as a starting point for DNA synthesis.



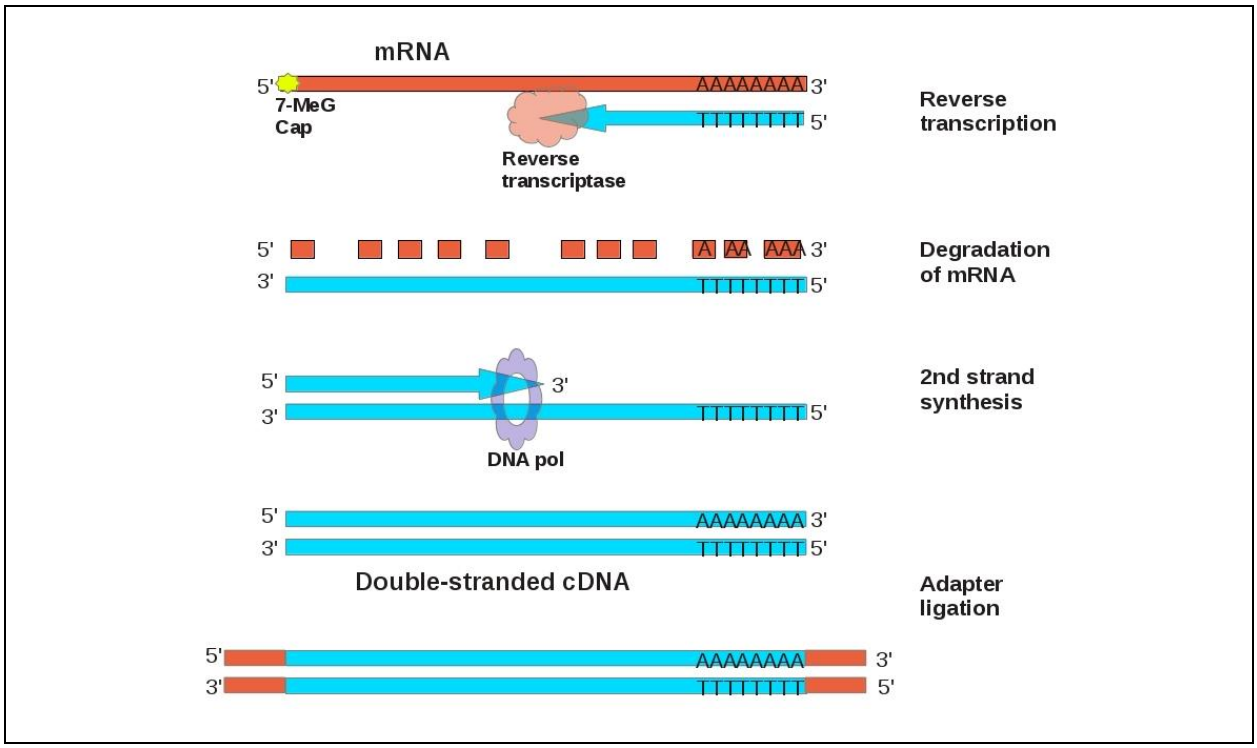
The primer is required for DNA replication because **DNA polymerase** (DNA replication enzymes) can only add new nucleotides to an existing strand of DNA (the primer). Various primers, with known nucleotide sequences, are used in biotechnology labs.



Oligo-dT (5' dt Primer) **is** a short chain of thymine deoxynucleotides; it can be used as a **primer** during **cDNA** synthesis, because adenine deoxynucleotides (A A A A A A) of **poly-(A)** tail can base-pair with thymine deoxynucleotides (T T T T T T) of **Oligo-dT**.

The enzyme reverse transcriptase can be used, along with an **oligo-dT** primer (that is complementary to the **poly-(A) tail** of mRNA) to synthesize a complementary DNA

(**cDNA**) molecule. The cDNA can then be cloned into a plasmid or it may be amplified by **PCR** by ligating (joining) adapters that contain **restriction endonuclease** cleavage sites or PCR primer-sequences.



The **cDNA library** is a combination of cloned fragments of cDNA inserted into a collection of host cells. When all messenger RNA molecules in one cell or a population of cells are converted into cDNA clones, it is known as cDNA library. The tissue-specific **cDNA libraries** can also be produced, using cDNA clones prepared from cell-genome of the tissue concerned.

Molecular cloning of eukaryotic genes is often impracticable because they contain numerous large **introns**. Eukaryotic genes (with large introns) cannot be cloned in plasmid because **plasmid vectors** have a practical size limit of less than 10 kilo-base pairs (kbp), and PCR is also difficult beyond 10 kb.

The cDNA is produced in the nucleus from fully transcribed mRNA (with no introns) and therefore contains only the expressed genes of an organism. In fact, the mRNA, lacking introns, is a compact version of a eukaryotic gene that retains

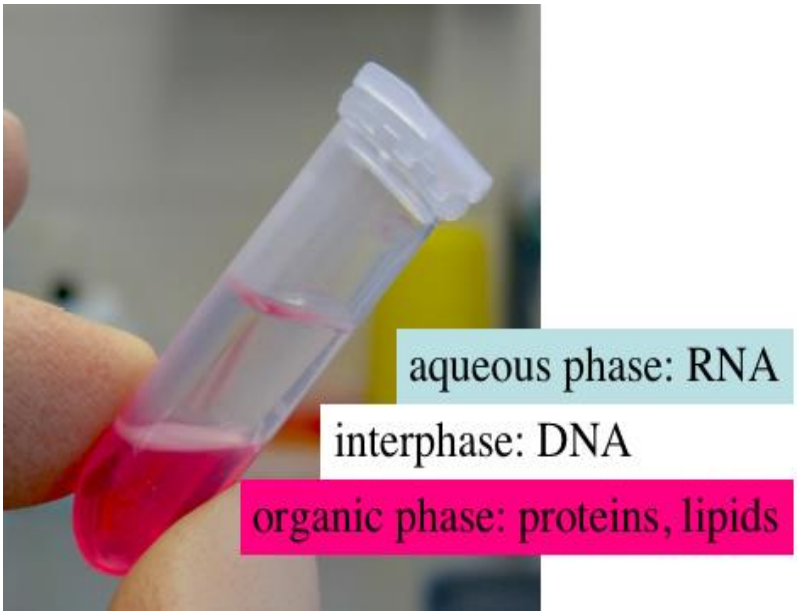
all of the protein-coding information and, hence, can be readily expressed in a bacterial cell.

The cDNA is often used to clone eukaryotic genes in **prokaryotes**. When scientists want to express a specific protein in a cell that does not normally produce that protein, the scientists insert the cDNA in a vector and transfer it as recombinant DNA molecule into the host cell. The desired gene, carried by the cDNA, then gets expressed in the recipient prokaryotic cell to produce the specific protein.

Information in cDNA libraries is a powerful and useful tool since gene products are easily identified, but the libraries lack information about enhancers, introns, and other regulatory elements that are found in a **genomic DNA library**.

Synthesis of cDNA molecule

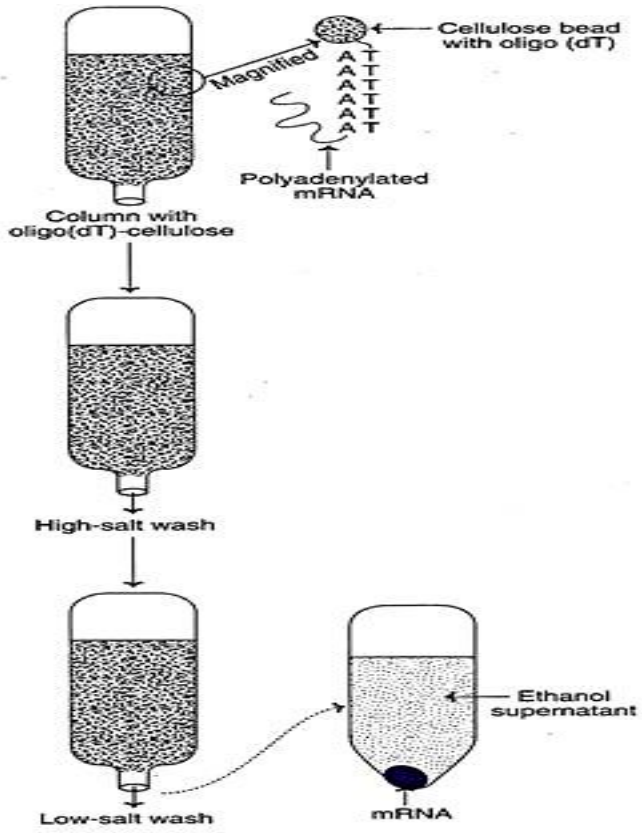
Although there are several methods for cDNA synthesis, cDNA is most often synthesized from mature mRNA (with no introns) using the enzyme **reverse transcriptase**. This enzyme operates on a single strand of mRNA, generating its complementary DNA based on the pairing of RNA nucleotides (**A**, **U**, **G** and **C**) to their DNA complements (**T**, **A**, **C** and **G**, respectively).



STEPS:

(1) The mRNA extraction: The mature mRNA strands are extracted from the cell. The mRNA is obtained and purified from the rest of the RNAs. Several methods exist for purifying RNA such as **trizol extraction** and **column purification**.

Purification of mRNA by affinity chromatography with oligo(dt) – cellulose

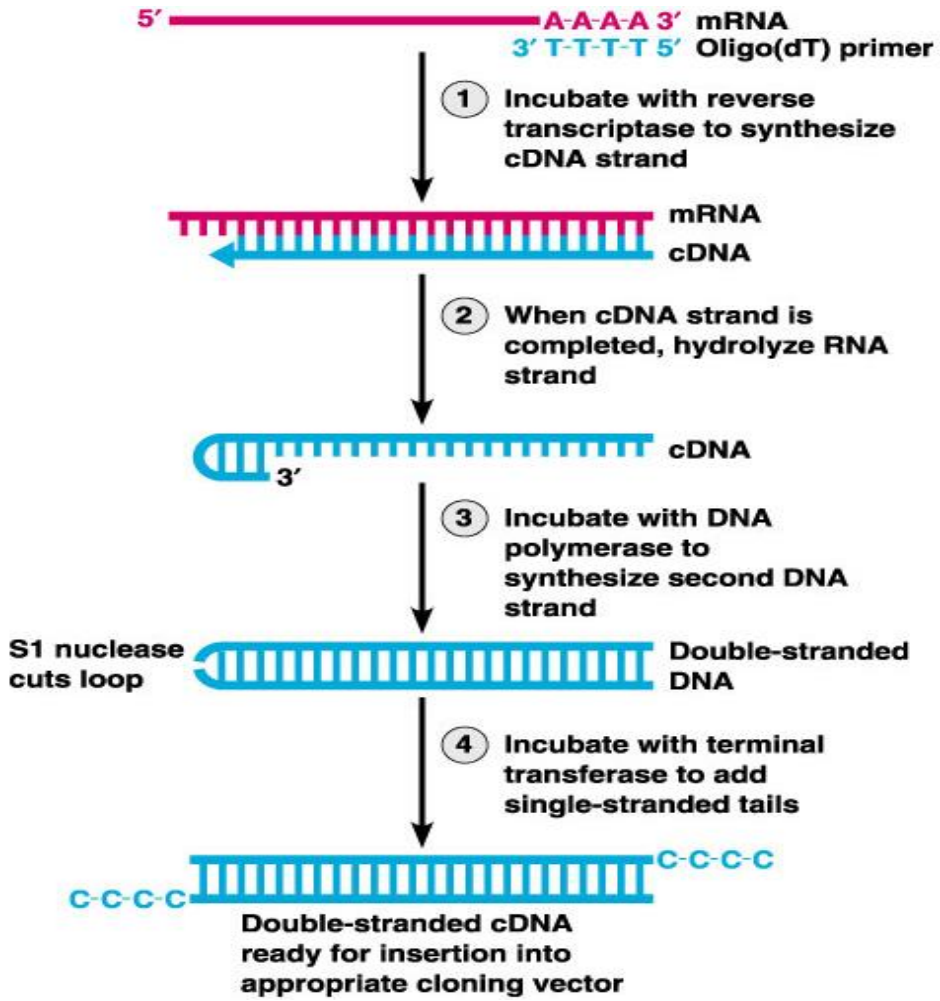


(2) Enzyme **Reverse transcriptase** is added, along with deoxyribonucleotide triphosphates (**A**, **T**, **G**, **C**). This synthesizes one complementary strand of DNA hybridized to the original mRNA strand.

(3) The resulting mRNA-cDNA **hybrid** is treated with **alkali** or an enzyme (**RNase H.**) to hydrolyze the RNA, leaving the single-stranded **cDNA**.

(4) After digestion of the mRNA, a single stranded DNA (**ssDNA**) is left, and because single stranded nucleic acids are hydrophobic, the ssDNA tends to loop around itself. As a result, a hairpin loop is formed in the ssDNA at the 3'-end.

Generation of double-stranded cDNA



(5) DNA polymerase can now synthesize the complementary DNA strand. For this, the looped-around **3'-end** of the first DNA strand can often be used as a **primer**. An enzyme called **S1 nuclease** is then used to cleave the loop.

(6) The double stranded cDNA is lastly incubated with the enzyme **terminal transferase** to add single stranded tails to the cDNA strands at 3'-end. Now, the

double-stranded **cDNA molecule** is ready for insertion into an appropriate cloning vector for amplification.

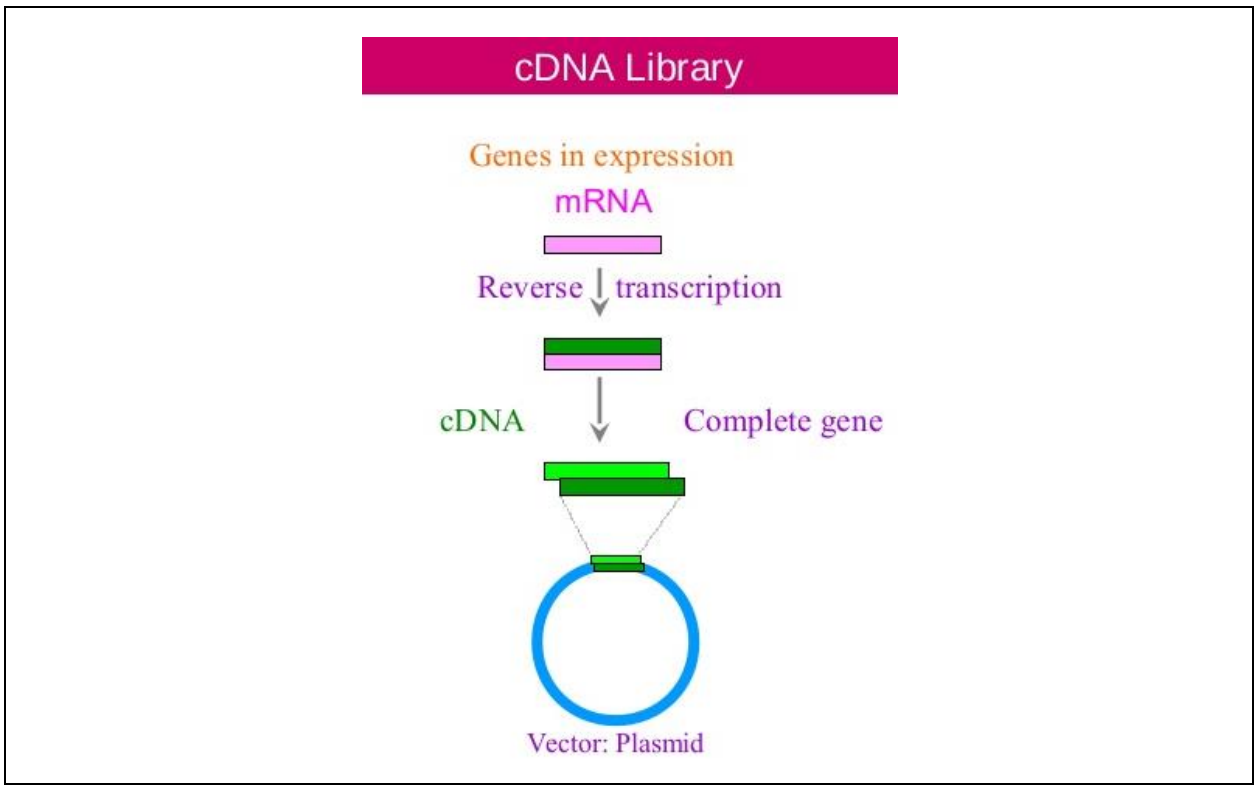
The construction of **cDNA library** involves following steps:

1. Isolation of mRNA
2. Synthesis of first and second strand of cDNA
3. Incorporation of cDNA into a vector
4. Cloning of cDNAs

Steps 1 and 2 have been explained above. The third and fourth steps are as follows:

Incorporation of cDNA into a vector and its cloning

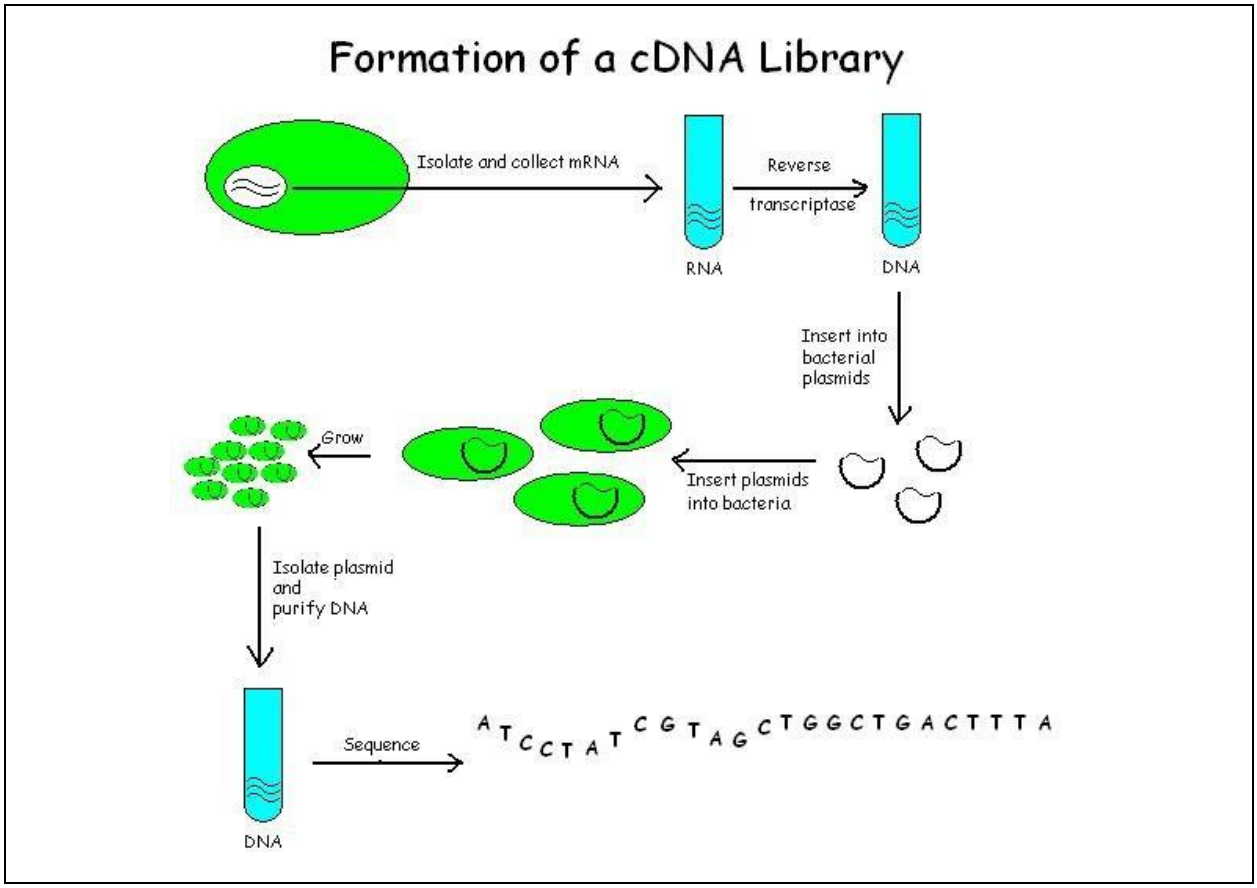
Restriction endonucleases and DNA ligase are used to clone the plasmid-confined cDNA inside the host cell.



The cloned bacteria are then selected, commonly through the use of antibiotic selection. Once selected, stocks of the bacteria are created, which can later be grown and sequenced to compile the **cDNA library**.

Applications of cDNA

Complementary DNA is often used in gene cloning or as gene probes or in the creation of a cDNA library. When scientists transfer a gene from one cell into another cell in order to express the new genetic material as a protein in the recipient cell, the cDNA will be added to the recipient cell (rather than the entire gene), because the DNA for an entire gene may include DNA that does not code for the protein or that interrupts the coding sequence of the protein (e.g., introns/enhancers).

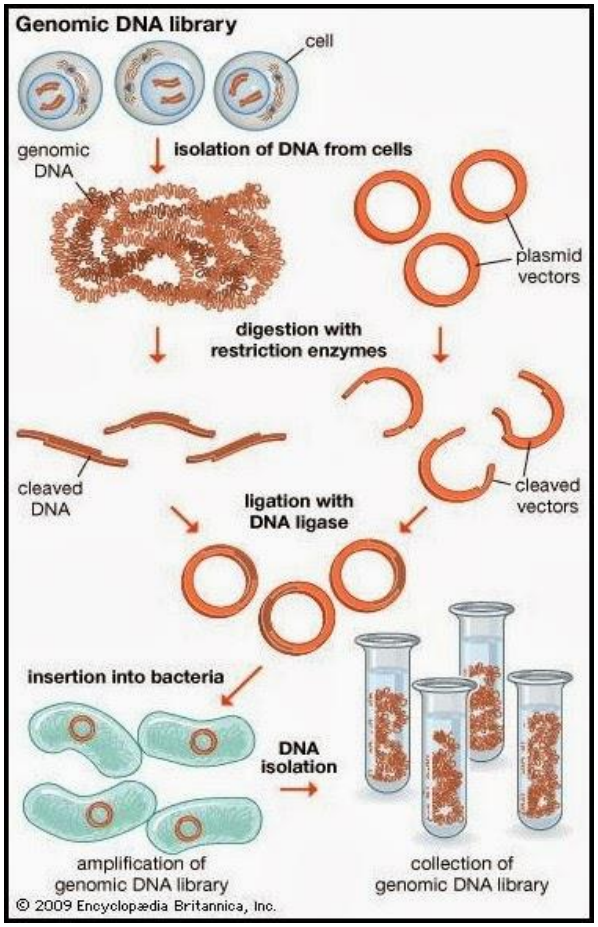


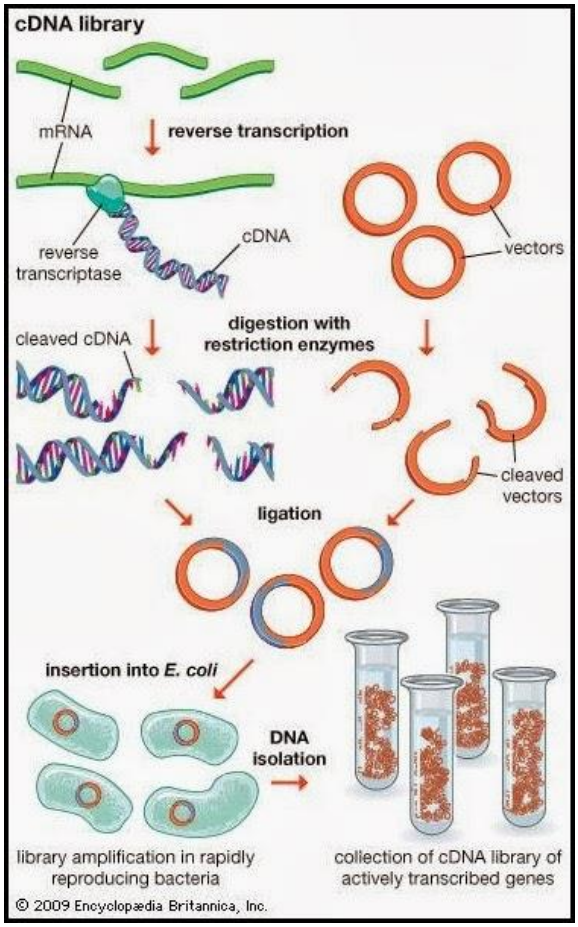
There are following applications of cDNA library:

1. The cDNA library is a powerful and useful tool in the area of biotechnology.
2. It is helpful in expressing eukaryotic genes in prokaryotes, which helps in the transcription process of prokaryotes.
3. It is used to isolate DNA sequences to code mRNA.

4. Advantage of cDNA library is to isolate homologous genes.
5. It is also used for the screening genomic libraries to isolate specific cDNA.
6. The cDNA of proteins can facilitate to generate antibodies and monoclonal antibodies.
7. The most important application of cDNA library is to study expression of mRNA.

Difference between genomic and library cDNA library





Difference between genomic library and cDNA library		
S. No.	**Genomic Library**	**cDNA Library**
1.	Genomic library includes all possible fragments of DNA from a given cell or organism.	The cDNA library carries only expressed gene sequences.
2.	It is larger in size.	It is smaller in size.
3.	It represents entire genome of an organism having both coding and non-coding regions.	It represents only the expressed part of the genome and contain only coding sequences called ESTs (expressed sequence tags).
4.	Expression of genes taken from genomic library is difficult in prokaryotic systems like bacteria due to absence of splicing mechanism in prokaryotes.	The cDNA carries only coding sequences, which can be directly expressed in eukaryotic systems.
5.	Vectors used in genomic library include plasmid, cosmid, lambda phage, BAC (Bacterial artificial chromosomes) and YAC (Yeast artificial chromosome) in order to accommodate large fragments of DNA.	Vectors used in cDNA library include plasmid, phagemid, lambda phage, etc. to accommodate small DNA fragments as cDNA contains no introns.
6.	Genomic library contains elaborated sequence information, representing total genome of the cell or organism because it is prepared from total genomic DNA.	The cDNA library contains limited sequence information because it is prepared from coding sequences only.