**ISOLATION AND PURIFICATION OF RNA (RIBONUCELIC ACIDS)**

Obtaining pure RNA is an essential step in the analysis of patterns of gene expression and understanding the mechanism of gene expression. Thus, isolation of pure, intact RNA is one of the central techniques in molecular biology and represents an important step in Northern analysis, nuclease protection assays, RNA mapping, RT-PCR, cDNA library construction and in vitro translation experiments. Two strategies of RNA isolation are usually employed: isolation of total RNA and isolation of mRNA. A typical eukaryotic cell contains about 10to20 pg of RNA, most of which is localized in the cytoplasm, whereas a prokaryotic cell contains 0.02 to 0.05 pg of RNA. About 80 to 85 percent of eukaryotic RNA is ribosomal RNA, while 15 to 20 percent is composed of a variety of stable low molecular weight species such as transfer RNA and small nuclear RNA. Usually about 1 to 3% of the cell RNA is messenger RNA (mRNA) that is heterogeneous in size and base composition. Almost all eukaryotic mRNAs are monocystronic and contain a post transcriptionally added, poly-adenylic acid (poly A) tract at their 3' terminal. This 3' poly-A tail permits separation and isolation of mRNA from all other classes of RNA present in the cell. In prokaryotes many mRNAs are polycistronic and none contain poly A tracts. This adds to the difficulty of mRNA purification from prokaryotic cells.

**Purification of total RNA**

Isolation of total RNA is most frequently used when pure RNA is required for experiments. This is because these techniques are less laborious and require less time to perform than isolation of mRNA. Various techniques for purification of total RNA are now in use and their application depends on the nature of RNA required. For example, if RNA is going to be used for quantitative RT-PCR, intactness of the purified RNA is not critical, while intact RNA is required for cDNA library preparation or Northern blot analysis. Complete removal of DNA contamination is critical if RNA is to be used in RT-PCR but is not important in *in vitro* translation. The techniques described here yield pure intact RNA good for any applications.

The physical and chemical properties of RNA and DNA are very similar, thus, the basic procedures used in RNA purification are similar to those of DNA. All of the RNA purification methods incorporate the following steps:

a. disruption of cells or tissue,

b. effective denaturation of nucleoprotein complexes and removal of proteins,

c. concentration of RNA molecules and,

d. determination of purity and integrity of isolated RNA.

In addition, methods must include procedures that remove co-purified DNA from the preparation. In contrast to DNA purification, guarding against physical shearing of RNA molecules is not necessary because RNA molecules are much smaller and much more flexible than DNA molecules. Therefore, in RNA protocols, strong physical forces during cell and tissue breakage are frequently used and the use of wide-mouth pipettes is not required. However, RNA isolation is much more difficult than DNA purification largely due to the sensitivity of RNA to degradation by internal and external ribonucleases. These enzymes are omnipresent and are very stable molecules that do not require any cofactors for their function. A crucial aspect of any procedure for RNA purification is fast and irreversible inactivation of endogenous RNases and protection against contamination with exogenous RNase during the isolation procedure. To these ends, all extraction buffers include powerful RNase inhibitors and all solutions and equipment used are treated to remove exogenous RNases.

**Elimination of RNases**

The most commonly used inhibitors included in extraction buffers to inhibit endogenous RNase are:

* Strong protein denaturation agents *Guanidinium hydrochloride* and *Guanidinium isothiocyanate* used at a concentration of 4 M. These chaotropic agents can quickly inactivate endogenous RNases and contribute to denaturation of nucleoprotein complexes. To irreversibly denature RNase by these compounds, a high concentration of 2-mercapthoethanol is also included.
* Vanadyl-ribonucleoside complexes. Oxovanadium IV ions form complexes with any ribonucleoside and bind to most RNases inhibiting their activity.
* Aurintricarboxylic acid (ATA). This compound binds selectively to RNase and inhibits its activity. ATA is usually incorporated into extraction buffers used for bacterial RNA preparations. The inhibitor can affect certain enzymes and should not be used if RNA will be needed for primer extension or SI nuclease analysis.
* Macaloid. This naturally occurring clay (sodium magnesium lithofluorosilicate) being negatively charged, strongly absorbs all RNase. The macaloid and bound RNase are removed from the preparation by centrifugation
* Protein RNase inhibitors such as RNasin. A protein originally isolated from human placenta, inhibits RNase by non-competitive binding. It cannot be used in extraction buffers containing a strong denaturant.

The most frequent sources of exogenous RNase contamination are one's hands, and bacteria and fungi present on airborne dust particles. To remove exogenous RNase contamination, the most frequently used inhibitors are:

* Diethyl pyrocarbonate (DEPC). DEPC causes enzyme inactivation by denaturing proteins. Inactivation of RNase is irreversible. The compound is used for removing RNase from solutions and glassware used in RNA preparation. DEPC should be used with care because it is highly flammable and a suspected strong carcinogen.
* RNaseZap®or RNaseOff solutions. These commercially available reagents destroy RNases on contact very effectively. The decontamination solutions are not toxic and can be used to remove RNase from all surfaces and equipment. The compositions of these reagents are trade secrets.

Three methods of RNA isolation are: a guanidinium hot –phenol method, a high-salt lithium chloride method and a method using a commercially available reagent, TRI-Reagent™.

**TRI-Reagent™ based RAN isolation protocol:**

The procedure described here is a modification of the procedure recommended by Molecular Center Inc. The procedure is carried out in 1.5 ml microfuge tubes. It can be used with 10 mg of plant or animal tissue prepared by grinding in liquid nitrogen (see section Preparation of cell materials) or with 1x 106 eukaryotic cells or 1x 107 bacterial cells.

This procedure can be scaled up to prepare a large quantity of RNA.

1. Add 10 mg of powdered tissue to a microfuge tube. Let the liquid nitrogen evaporate for a few minutes. To isolate RNA from bacterial or yeast cells, the appropriate amount of cells should be collected by centrifugation.

2. Add 700 ul of TRI Reagent. Close the tube and mix by vortexing. Incubate for 5 minutes at room temperature. Note: The amount of TRI Reagent" should be no less than 9 volumes of reagent to one volume of material.

Note: When preparing RNA from yeast or bacteria, add an equal volume of acid-washed glass beads and continue vortexing for 1 minute to facilitate cell breakage. To prepare glass beads soak them for 1 hour in concentrated nitric acid, wash acid off with deionized water and dry beads by baking in 180 °C oven overnight.

3. Add 140 uI (0.2 volume) of chloroform (not CIA) and mix by vortexing for 15 seconds. Incubate at room temperature for 2 to 15 minutes.

4. Centrifuge for 15 minutes at room temperature. After centrifugation, the mixture will be separated into two phases, a bottom phase containing chloroform and an upper, aqueous phase containing RNA.

5. Transfer the aqueous phase to a fresh microfuge tube and precipitate RNA by the addition of 175 uIofisopropanol (0.25 volume of the original volume of TRI Reagent) and 175 ul of high salt precipitation solution. Mix by inverting the tube several times and incubate at room temperature for 5 to 10 minutes.

6. Place the tube into the centrifuge, orienting the attached end of the tube lid, away from the center of rotation. Centrifuge for 10 minutes at room temperature to collect precipitated RNA.

7. Remove the tube from the centrifuge. Remove supernatant using a P200 Pipetman.

8. Wash the pellet with 1 ml of 75% ethanol. Add ethanol to the tube and mix by inverting several times.

9. Place the tube into a microfuge and centrifuge for 5 minutes at room temperature. Remove ethanol with a P200 Pipetman. Repeat ethanol washing one more time.

10. Place the tube into the centrifuge, making sure that the side containing the pellet faces away from the center of rotation. Start the centrifuge until it reaches 500 rpm (1-2 seconds). This will collect ethanol from

the sides of the tube. Remove ethanol using a P200 Pipetman equipped with capillary tip.

11**.** Dissolve the pelleted RNA in 25-50 µI of RNase free 1 mM EDTA. To facilitate dissolving RNA, place the tube into a 65 °C water bath for 10 to 15 minutes. Store the RNA sample at -70 °C.

12. Determine the concentration of RNA by measuring absorbance at 260 nm. Initially use a 1:100 dilution of the sample in PBS. The absorbance reading should be in the range 0.1 to 1.5. Calculate the concentration of

RNA using the equation:

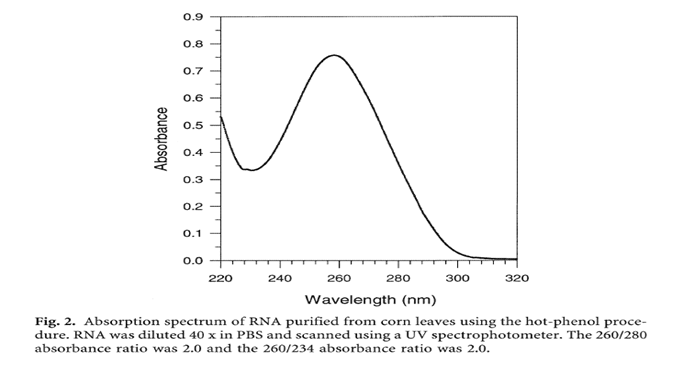
A260

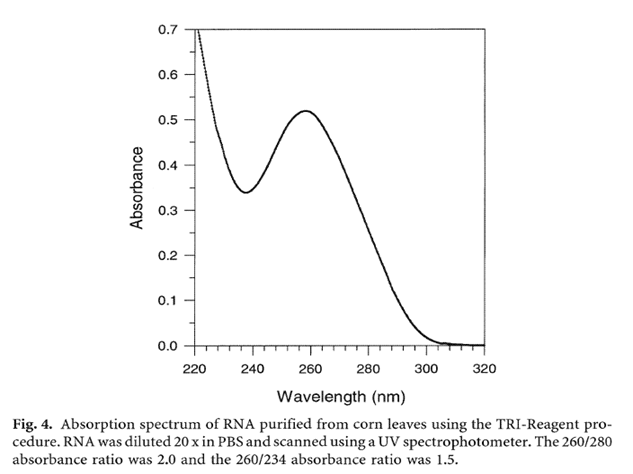
N = ------------

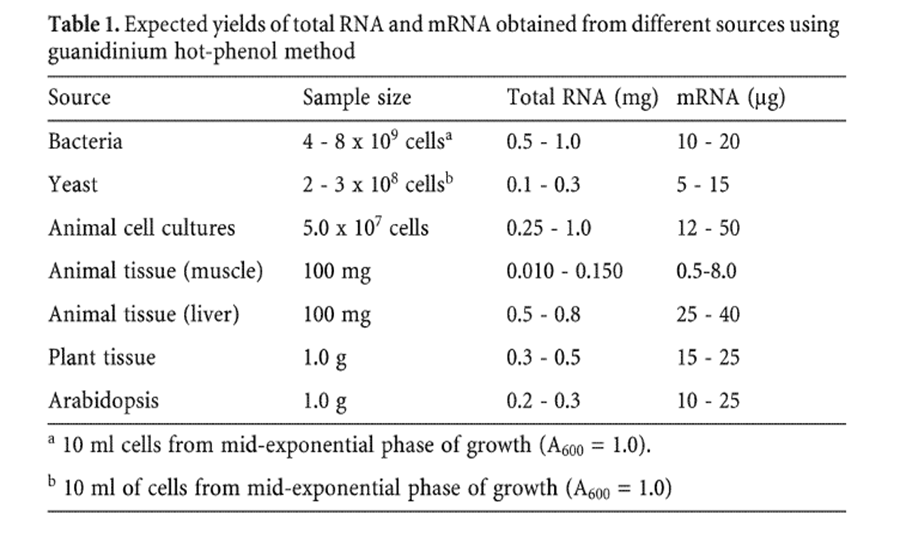
ε260

Where ε260 is the RNA extinction coefficient, N is RNA concentration in ug/ml and A260 is the absorbance reading (corrected for dilution). The absorption coefficient for total RNA is usually taken to be 0.025 ug-1 cm-1 giving a solution of 40 ug/ml an absorbance of1(e.g., 1/0.025=40 ug/ml).

To determine the purity of the RNA, measure absorbance at 260 nm, 280 nm and 234 nm and calculate the 260/280 and 260/234 ratios.







**Complementary DNA (cDNA) preparation**

Complementary DNA libraries reflect gene expression at certain times for specific cells, whereas genomic DNA libraries represent all genetic information in somatic cells. The complexity of cellular organization reflects a genetic program that encodes a collection of genes and the means to use them by manufacturing proteins for cellular structures, functional activities, and reproduction of cells themselves. The essential aspect of this process is protein synthesis based on the information stored in the sequence of nucleotides that make up a gene (a transcribable segment of a DNA molecule) as the blueprint. The information is transcribed as a complementary sequence of the nucleotides (mRNA or the transcript) that carries the genetic information from the nucleus to the protein-synthesizing machinery in the cytoplasm. Then, mRNA is translated into the sequence of amino acids that make up a protein. The basis of the widely used novel strategies for the generation of cDNA libraries are base pair complementarities, reverse transcription, and polymerase chain reactions.

**Base Pair Complementarities**

Nucleic acids exhibit base pair complementarities that faithfully convert one strand of RNA/DNA to a complementary one. Although all genetic information in the somatic cells of a specific organism can be expressed as a transcript, many DNA sequences are not transcribed. These segments of DNA are the coding exons and the noncoding introns. Basically, the genetic information is stored as a strand of a DNA molecule consisting of four bases: adenine, thymine, guanine, and cytosine. A second complementary strand of DNA can be formed by DNA polymerase. Polymerases, enzymes that function in DNA replication and RNA transcription, synthesize a nucleic acid from the genetic information encoded by the template strand. The polymerases are unique because they take direction from another nucleic acid template, which is either DNA or RNA. During the formation of a second strand of DNA, bases are generated according to the Watson– Crick base-pairing pattern. That is to say, every cytosine is replaced by a guanine, every guanine by a cytosine, every adenine by a thymine, and every thymine by an adenine. In this way, information in DNA is correctly transcribed into RNA.

**Probe Hybridization**

Another unique feature of the base pair complementarity is probe hybridization. The findings of Gillespie and Spiegelman that viral genomic DNA and RNA in infected cells showed a base pair complementarity opened an avenue for specific hybridization between a gene and its transcript as a DNA–RNA hybrid. Subsequently, the DNA–DNA or DNA–RNA hybrids have been employed in a large number of powerful techniques like cDNA preparation and hybridization of nucleic acids.

**Polymerases Are Essential for DNA Synthesis**

Polymerases that use RNA as a template to form a complementary DNA are RNA-direct DNA polymerases. One of these enzymes is reverse transcriptase, usually observed as a part of the viral particle, during the life cycle of retroviruses and other retrotransposable elements. Purified reverse transcriptase is used to generate complementary DNA from polyadenylated mRNAs; therefore, double-stranded DNA molecules can be formed from the single-stranded RNA templates. The synthesis of DNA on an RNA template mediated by the enzyme reverse transcriptase is known as reverse transcription.

**A Primer is Required for Reverse Transcription**

Although polymerases copy genetic information from one nucleotide into another, including copying a mRNA to generate a complementary DNA strand in the presence of reverse transcriptase, they do need a “start signal” to tell them where to begin making the complementary copy. The short piece of DNA that is annealed to the template and serves as a signal to initiate the copying process is the primer. The primer is annealed to the template by base pairing so that its 3'-terminus possesses a free 3'-OH group and chain growth is exclusively from 5' end to the 3' end for polymerization. Wherever such as primer–template pair is found, DNA polymerase will begin adding bases to the primer to create a complementary copy of the template.

**Formation of cDNA**

Generally, the cDNA of cells can be formed according to the following steps:

1. Isolation of the mRNA template: The source mRNAs can be enriched by increasing the abundance of specific classes of rare mRNAs via one of the following approaches: (1) antibody precipitation of the protein of interest that is synthesized in cell lines, (2) increasing the concentrations of relevant RNAs by drug-induced over expression of genes of interest, and (3) inhibition of protein synthesis by inhibitors, resulting in extended transcription of the early genes of mammalian DNA virus. The integrity of the mRNA is essential for the quality of cDNA generation. The size of mRNAs isolated should range from 500 bp to 8.0 kb, and the sequence should retain the capability of synthesizing the polypeptide of interest in vitro, such as in cell-free reticulocytes. When fractionated by electrophoresis and stained with ethidium bromide, a good preparation of mRNA should appear as a smear from 500 bp to 8 kb.

2. A short oligo(dT) primer is bound to the poly(A) of each mRNA at the 3' end.

3. The mRNA is transcribed by reverse transcriptase (the primer is needed to initiate DNA synthesis) to form the first strand of DNA, usually in the presence of a reagent to denature any regions of the secondary structure. RNAse is used to prevent RNA degradation.

4. DNA–RNA hybrids are formed.

5. The RNA is nicked by treatment with RNAse H to generate the free 3'-OH groups.

6. DNA polymerase I is added to digest the RNA, using the RNA fragments as primers, and replace the RNA with DNA. In some cases, a primer–adapter method is carried out as follows:

(1) terminal transferase is added to the first strand cDNA [add (dC) to provide free 3' hydroxyl groups]; (2) the tail of hybrided cDNA with oligo(dG) serves as the primer.

7. Double-stranded cDNA is formed.

**Why cDNA Libraries?**

Complementary DNA libraries are preferalbe to mRNA libraries for the following reasons:

1. cDNA can represent the gene that is expressed as mRNA in a specific tissue or specific cells at a specific time; therefore, the mRNAs in two different types of cell or the same type of cell with different treatments may vary because the expression of genes varies.

2. cDNA libraries usually provide reading frames encoded within the DNA insert after the noncoding intervening sequences are removed; therefore, cDNA reflects both a mRNA transcript and a protein translation product. cDNAs can be used as probes for screening the mRNA transcript as well as in the rapid identification of amino acid sequences of polypeptides or proteins. Because there are no introns in a cDNA molecule, they are frequently used in protein synthesis in vitro.

3. The protein-encoding mRNA may not be present in all cells showing the specific protein because the mRNA is easily degraded and the protein formed in the cell could be present as a stable form from an earlier expression of the mRNA.

4. Because different numbers of copies of different mRNAs are present in a cell (low, middle, and high abundance), a desirable characteristic of cDNA libraries is that they increase the number of the less abundant species and reduce the relative number of high and middle abundant species. By manipulating the rate of strand reannealing in a denatured cDNA preparation, the high and middle abundance species of mRNA can be removed. The resulting cDNA generated is representative of the rarer species. Other modifications can be used to achieve the enrichment of cell-, tissue-, or stage-specific mRNA species in the preparation of cDNA libraries.

5. Messenger RNA are difficult to maintain, clone, and amplify; therefore, they are converted to more stable cDNA, which is less susceptible than mRNA to degradation by contaminating molecules. For the above-mentioned reasons, cDNA libraries are preferred over mRNA libraries for genetic manipulations.

**Potential Applications**

The most common application of mRNA/cDNA libraries is the identification of genes of interest. They are also used for other mRNA/cDNA manipulations to determine differentially expressed gene levels associated with structural and functional changes that are of high relevance to disease controls or pathways of specific molecule modulations. In the past, one rate-limiting step in this type of study is the lack of high-quality human mRNA generated from limited and heterogenous pathological specimens. The RNA-PCR method and other methods for generating highquality mRNAs will solve this problem. Coupled with microarray technologies and microdissected single cells, mRNA/cDNA libraries so generated can be used to monitor a large number of genes and provide a powerful tool for assessing differential mRNA expression levels for the identification of disease-associated genes. With the antisense knockout techniques, double-stranded mRNA silencing of posttranscriptional gene expression, and a newly developed cDNA–mRNA hybrid interference of gene expression, the function of an overexpressed genes can be examined. After generation of stage-specific cDNA libraries, one may examine other genes of interests and determine whether these genes are differentially expressed. Altered gene expression of certain molecules and their related receptors may shed light on the developmental, physiological, and pathological significance of these molecules. In addition, one can examine differential gene expression of other genes in the presence and absence of the gene of interest by modulating the levels of each gene of interest. In this manner, each marker, growth factor and/or its receptors, or genes associated with a physiological or pathological phenomenon could be thoroughly monitored for altered levels of expression.

**Expressed Sequence Tags (ESTs)**

Nowadays, genome analysis has employed a rapid analysis tool known as expressed sequence tags. In 1983, SD Putney for the first time demonstrated the use of cDNA in identification of genome. The term expressed sequence tags (ESTs) was coined by Anthony Kerlavage at the Institute for Genomic Research. In 1991, Mark Adams used EST in relation to gene discovery and Human Genome project.

**Expressed Sequence Tags or ESTs**, as the name suggests, are the new generation tools providing new dimension to transcriptome analysis. They are the tiny sequences of cistron randomly selected from genome library and can be used to identify and map the whole genome of any particular species. ESTs are usually 200 to 500 nucleotides long and are generated by sequencing the ends of DNA. ESTs can be obtained without much expenditure and are quite fast in genomic analysis. The EST sequences can be used to search the homologous organisms in different databases such as NCBI (National Centre for Biotechnology Information). Thus we can collect information on expression patterns of different species. Therefore, they play vital role in discovery of gene and genome analysis.

***Generation of Expressed Sequence Tags:*** The presence of introns makes gene identification quite difficult. The DNA is firstly transcribed to mRNA which is the key for synthesis of building blocks i.e. proteins by a process called translation. Interestingly, mRNAs do not contain the sequence transcribed from introns. Thus mRNA isolation is the key for ESTs construction. mRNA is quite unstable outside the cell, therefore reverse transcription is performed to convert it to cDNA, which is comparatively stable. Expressed Sequence tags are generated from cDNA cloned from mRNA of any particular species. As the cDNA used is complementary to mRNA, the ESTs represent portions of expressed genes.

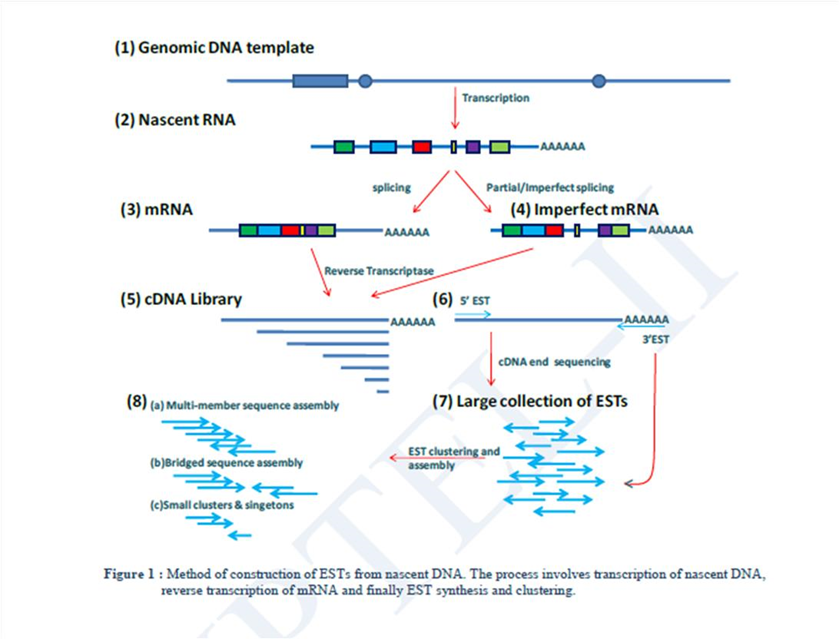
**The ESTs can be generated by following steps:**

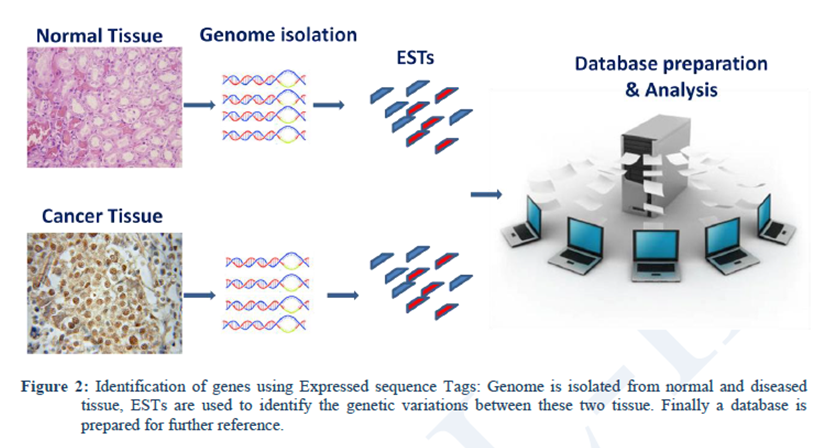
***Transcription of Genomic DNA:*** Genomic DNA is first transcribed to generate Nascent mRNA followed by splicing of synthesize perfect mRNA.

***Reverse transcription of mRNA:*** mRNA can also be directly isolated from the species by using different kits (e.g. RNAgent Promega). mRNA synthesized undergoes reverse transcription to form cDNA library.

***Generation of ESTs:*** From the cDNA library 5’ or 3’-ESTs are generated by cDNA end sequencing. 5’ EST is formed from a region of transcript which forms protein whereas the ending portion of cDNA forms 3’EST.

***Assembly and organization of ESTs:*** The constructed ESTs can then be assembled separately in multimember sequence assembly, Bridged sequence assembly and small clusters on the basis of size of ESTs.





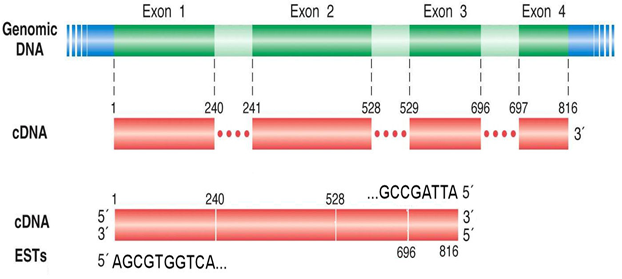


Fig 2: Comparative principle of Genomic DNA library, cDNA library and ESTs preparation

**Applications of ESTs**

ESTs can be used as functional DNA arrays for the analysis of whole genome of a species.

***ESTs as Genome searching tool:***

Scientists use the genome map to travel through the billions of nucleotide in the genome of a species. ESTs work as landmarks for the genome mapping. Currently, the most powerful technique for genome mapping is Sequence Tagged Site (STS) mapping. 3'-ESTs as a commonly used as a source of STSs because of their uniqueness in an individual’s genome.

***Identification of position of genes:***

ESTs are widely and efficiently used to locate an already known gene in the genome of a species. Most commonly used technique for this purpose is called as Sequence Tagged Site (STS) mapping.

***To find out gene responsible for disease:***

This method allows us to identify the genes responsible for any deformity or disease such as Alzheimer’s and colon cancer have already been investigated. ***ESTs and Human Genome Project:*** Existence of thousands of genes have been identified in Human Genome Project solely on the basis of ESTs.

***Use of ESTs in similarity searches:*** Due to their putative and fast behavior, NCBI has included millions of ESTs databases for several species. Scientists as well as genome sequencing centers are widely using these ESTs for similarity searching between different species.

------------------------------------------------------------------------------------------------------------------------------------------