J Biochem Technol (2024) 15(2): 18-22
https://doi.org/10.51847/73VjyOe9IQ
ISSN: 0974-2328

Genomic Libraries – An Overview and a Narrative Review

S. M. Sowbaraniya, Shilpa Syam*

Received: 20 February 2024 / Received in revised form: 27 April 2024, Accepted: 02 May 2024, Published online: 15 May 2024

Abstract

A gene library is a collection of different DNA sequences from an organism, which has been cloned into vectors for ease of purification, storage, and analysis. There are two types of gene libraries based on the availability of a source of DNA used. The genomic library contains DNA fragments that represent the entire genome of an organism, whereas CDNA libraries will contain only the complementary DNA molecules synthesized from the mRNA molecules in a cell. For the construction of gene libraries, we have to know the size of the gene and insert size or capacity of vectors. Vectors used for gene library construction are plasmid, lambda phage, Cosmid, bacterial artificial chromosome, and yeast artificial chromosome. The most commonly used vector type is lambda, but most genomic libraries are constructed using BAC and YAC vectors. For the construction of gene libraries, the steps followed are the isolation of bacteria followed by cloning and sequencing DNA. The main use of genetic libraries is to increase the likelihood that a particular portion of a DNA source will be found within the collection. It is used to collect and store information as a set of DNA molecules. Gene libraries are stored inside bacterial cells. Bacterial colonies have the ability to store different genes, similar to how books are stored in a library.

Keywords: Gene library, Gene, cDNA, Plasmid, Vectors, Bacteria

Introduction

Genetic libraries are collections of different DNA sequences from organisms that are cloned into vectors for ease of storage and analysis. Because this is a collection of cloned DNA, there is a high probability that a specific portion of the DNA source will be found within the collection. A genetic library is a collection of DNA fragments from varying biological systems. Gene libraries may contain the entire genome sequence or cDNA. However, the DNA sequence is formed from mRNA. The two major types of gene libraries are classified depending on the source of DNA used for construction. The types are Genomic libraries and cDNA libraries (Caucheteur *et al*., 2018). Genomic libraries have only one DNA fragment, which represents the entire genome of an organism. Genomic libraries are further divided into A) Nuclear Genomic library B) Organelle genomic library (Burman *et al*., 2020). The complementary DNA library will contain only complementary(c) DNA molecules, which are synthesized from mRNA gene molecules in a cell. This group comprises clones that contain solely the cDNA of genes that express specific cells or tissues at certain times.

Construction of gene libraries involves the size of the gene and inserting the size or capacity of the vector. Based on this, gene library size can be known and also ensure the high probability of insertion (Zabarovsky, 2006). For the construction of libraries, molecular tools like molecular scissors and pastes are required. It is used to ligate the DNA into the required size (Jayaraman *et al*., 2019). Vectors like plasmids, lambda phage, Cosmid, BAC, and YAC vectors are used to construct genomic libraries depending on genie size. Among these vectors, the most commonly used type is lambda because of its structural flexibility. Since these lambda phages have a narrow host range, BAC and YAC are used for gene library construction (Park *et al*., 2016).

The construction of gene libraries varies for genomic libraries and cDNA libraries. But the steps followed are more or less similar. Extract and purification of DNA and digest with restriction enzymes. Consequently, this creates fragments of similar size that contain one or more genes; fragments are created and should be inserted into vectors. So, these vectors are also cut using the same restriction enzyme to seal the DNA fragments, and ligase is used. Therefore, this method creates a large pool of recombinant molecules. Applications of gene libraries are genome-wide association studies that aim to find specific gene targets and polymorphisms within the human race (Choi, 2014). An article suggested that cinnamon extract shows better antibacterial activity. This is due to the expression of genes related to lipids. Chlorhexidine has a bisbiguanide compound with a structure consisting of two p chlorophenyl guanidine genes, which shows good oral prophylaxis. Genes such as ccpa and brpa present in streptococcus mutans help form a plaque and cause dental caries (Shahana & Muralidharan, 2016). The presence of virulence genes and several resistance phenotypes makes Enterococcus faecalis form plaque and causes several root canal medicaments to be less effective than without endomethasone. Acinetobacter baumannii has a plasmid-mediated blaOXA58 gene, which plays an important role in urinary tract infections. The CHM gene present in the orange peel helps in the streptococcus mutation and enterococcus. Laurus nobilis genes have modified genes and have greater anti-inflammatory action due to modified genes. Plasmid encoded bla Tem has different gene sequences and various properties, and it

S. M. Sowbaraniya, Shilpa Syam*
Department of Oral Medicine and Radiology, Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences (SIMATS) Saveetha University, Chennai-600077, Tamil Nadu, India.

***E-mail:** shilpasyam.sdc@saveetha.com

shows that long-coded non-coding RNA such as Xist present in N-N-methyladenosine present has activity against human disease (Priyadharsini *et al*., 2018). *Acinetobacter* spp. Belonging to genotypes 4 and 5 were found to harbor 6–10 and 2–8 potential drug-resistant genes makes it potential. In silico analysis of virulence genes in an emerging dental pathogen A. baumannii and related species. An article suggested that kaempferol, apigenin, and catechin from A. nilotica seem to possess a promising inhibitory effect against the wild and mutated ERG11 of C (Ushanthika *et al*., 2019).

Results and Discussion

Gene Library and Its Types

A library of genetic information is a vast collection of DNA fragments that are duplicated from an organism, tissue, or organelle. Therefore, it can contain the entire genome sequence or a cDNA sequence formed from mRNA. Gene libraries are classified into two types based on the source of DNA such as A. genomic libraries and B. cDNA libraries (Van Dijk *et al*., 2014; Ghule *et al*., 2023). Genomic libraries are made from genomic DNA in which molecules of DNA are very large, so it is necessary to fragment them into small pieces to insert into vectors. Genomic DNA libraries are further classified into a) nuclear genomic library b) organelle genomic library. Nuclear genomic libraries contain the total DNA content of the nucleus. In this type of library construction, we extract specifically nuclear DNA and use it to construct the library. Organelle genomic libraries are collections of cloned, restriction enzyme-digested DNA fragments that contain at least one copy of every DNA sequence in the genome. Complementary DNA libraries are made from cDNA, which is copies of every DNA from mRNA molecules. Consequently, to make cDNA, mRNA is isolated from a tissue or whole organism, and DNA is copied from the mRNA template using the enzyme reverse transcriptase. Thus, using this, it will be reverse genetics based on the principle of reverse genetics principles. The resulting cDNA molecules are used for the construction of gene libraries (Pirone-Davies *et al*., 2020; Mashhour *et al*., 2023).

Vectors and its Types

The fragments of DNA are inserted into vectors using DNA ligase. Vector is used for easy amplification and retrieval of specific clones from the library for analysis.

Fragmentation

Thus, the DNA will be commonly fragmented into one of three general sizes based on which vectors can be selected. Vectors are classified into small, medium, and large inserts.

Vectors for Small Inserts

Libraries with less than 10 kB insertions are almost invariably cloned into plasmid vectors. So, these plasmid vectors are maintained in E. coli at a range of approximately 309 copies per cell and show resistance to ampicillin. It is one of the oldest methods, and it is unreliable for cloning all DNA classes. An article suggested that insertion resulting in an AcrR frameshift and large chromosomal deletions in the essential oil improves the antibacterial activity and helps prevent caries, also performing similarly to chlorhexidine (Sabarathinam *et al*., 2017, Petronis *et al*., 2023).

Vectors for Medium Inserts

Inserts of 1 approximate size 10-40 kB are most stable when vectors with low copy numbers in the range of 10-20 copies per cell are used. The most common types of medium insert vectors are E plasmid, bacteriophage, and lambda. Because of the structural flexibility, lambda types are mostly used, but the disadvantage associated with it is a narrow host range (Rohweder *et al*., 2019; Zakaev *et al*., 2023).

Vectors for Large Inserts

EBAC vectors are developed to contain inserts greater than 300kb, but they also accept inserts in the 100-200 kB range. PAC can also be used for large inserts, but the major disadvantage is that it is used to construct libraries for lower organisms such as a mouse, drosophila, and non-related genes, which may lead to mishaps (Mukai *et al*., 2020; Jain, 2023).

Construction of Gene Libraries

Construction of Genomic Library

Genomic library construction begins with cleaning the genome into small pieces using restriction endonuclease. Those fragments are then cloned in vectors and introduced microbes or packed using phage particles. So, it's basically like a library here; each book is a clone in the library. Extract DNA from the organism and select the vector according to the size of inserts. The vector has to open up using restriction endonuclease enzymes and insert fragmented DNA, opening up the vector, which leads to the formation of recombinant DNA. Once rDNA is formed, it is introduced to bacteria to clone **(Figure 1)**.

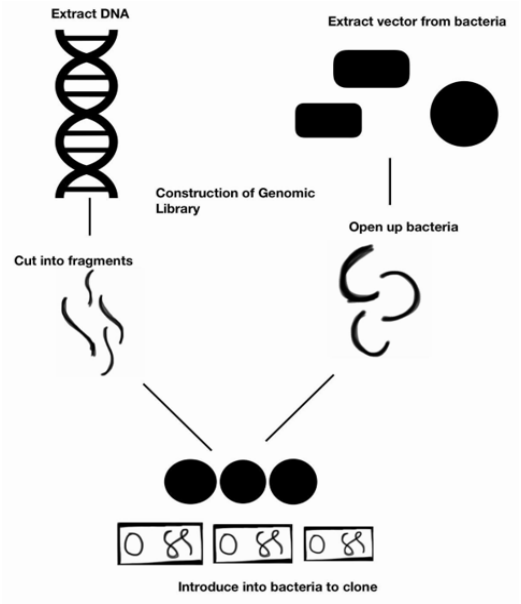

Figure 1. construction of the genomic library

Screening of Genomic Library

The screening is done to identify the gene of interest. Certain techniques are followed in screening, such as colony hybridization and plaque hybridization. So, colony hybridization is a screening of a library with a labeled probe, such as radioactive, which can be used to identify specific sequences of DNA, RNA, enzyme, protein, or antibody (Andriese, 1998). In contrast, plaque hybridization is based on the hybridization of oligonucleotide probes to target DNA. In this technique, petit dishes to filter the DNA are directly transferred.

Construction of cDNA Library

The construction of cDNA libraries is achieved by the extraction of mRNA. This mRNA was extracted and purified by triazole extraction and column purification. The mRNA is elated by using an elating buffer and some heat to separate the mRNA strands. Then, the mRNA is removed using RNase enzyme, leaving it as single-stranded cDNA. Once the cDNA is obtained, the same steps followed for the construction of the genomic library are followed.

Screening of cDNA Library

The identification of a specific DNA fragment allows for its isolation, subsequent amplification, and determination of its sequence. It is based on denature hybridization with a probe by performing autoradiography.

Sequencing of Gene Libraries

There are two types of gene libraries. Each of these has a different sequence and uses. The process of DNA cloning allows for the selection of a single copy of an entire DNA or RNA sequence from dozens of available sequences within cellular structures. The DNA sequence can be amplified after the fragmentation of DNA. Many articles documented that methicillin-resistant staphylococcus is the most common cause of hospital-acquired infection. Sequencing of VRSA shows unambiguously that each strain lineages well for acquiring antibiotic resistance in mixed infection (Misikir *et al*., 2016).

Uses of Sequences

The nucleotide sequence is used to predict the amino acid sequence of proteins. Though it has 6 different reading frames in which DNA sequences can be translated into protein, the correct one is generally recognized as the only one-stop frequent codon. Sequences that encode stretches of amino acids much longer than this are Exxon's. Selected DNA segments can be used for polymerase chain reactions (Langille *et al*., 2013).

Applications of Genomic Libraries and Its Uses

Application of Genomic Library

Genomic libraries play an important role in DNA sequencing projects. It also plays an important role in identifying novel pharmaceutically important genes. The identification of new genes that are not active in the host is greatly facilitated by genomic libraries and enhances our understanding of the complexity of the genome. Point mutation of the gene ERG11 is useful in cloning studies and the study of non-synonymous mutation in Candida species (Shahzan *et al*., 2019). Laurus nobilis gene studying is also used in various commercial products such as tea, a fragrance for food and cosmetic industries. Acinetobacter baumannii, inserting the GFP gene in a chromosome, is useful for genetic studies in important pathogens (Girija *et al*., 2018).

Uses of Genomic Libraries

A major role in transgenic animal production is through the genomic sequence. Creating cDNA libraries to establish what gene is expressed. Used to study the genetic mutation in cancer tissues. Major uses are shotgun sequencing and random shotgun sequencing. Researchers can explore the genome of an organism to learn more about genomic structure and function. Used to map the genome, identify the locations of specific genes, and for cloning segments of DNA. It is the first step in any DNA sequencing project.

Applications and Uses of cDNA Libraries and Its Uses

Application of cDNA Library

cDNA library is used to isolate homogeneous genes. It helps express eukaryotic genes in prokaryotes. The major application includes studying the expression of mRNA. It can also facilitate the generation of antibodies and monoclonal antibodies. It plays an important role in the discovery of novel genes. To elucidate gene function, to get high yields of recombinant cDNA. Commercial applications are the production of proteins and other biological molecules. Various applications include identifying carcinogens, studying alternative splicing, and obtaining pure samples of genes.

Uses of cDNA

A useful tool in the area of biotechnology. Useful In reverse genetics where additional genomic information is useful. Used for screening genomic libraries. Used in isolating the genes that code for particular mRNA and used for cloning of full-length cDNA molecules for in vitro study of gene function. Study of alternative splicing in different cells or tissues and for the repertoire of mRNA expressed in different cells or tissues and discovery of novel genes. Mainly used to reproduce eukaryotic genomes as the amount of information is reduced to remove the large numbers of non-coding regions from the library. Used to express eukaryotic genes in prokaryotes.

Disadvantages

Disadvantages of cDNA Library

- Contains only sequences that are present in mRNA.
- Introns and other sequences are altered after transcription is not present.
- Frequency of a particular DNA sequence in a cDNA library depends on the abundance of the corresponding mRNA.

Disadvantages of the Genomic Library

- The mishap in construction of the gene library.
- Errors can occur in genetic maps.
- Reference genome may be incomplete.

- Shotgun sequencing requires vast amounts of computing power and sophisticated software.

Global Platforms for Human Gene Sequencing and Mapping

International Breakpoint Mapping Consortium (IBMC) seeks to create a saturated map of balanced chromosomal rearrangements as a way to gain functional knowledge of the human genome. The library will serve as a biomedical platform. It emphasizes promoting international collaboration when initiating new partnerships for IBMC.

Conclusion

Genomic libraries help establish accurate genomic libraries and study specific gene targets and polymorphism in the human race. It also helps in future human cell line studies. There are several tools to avoid mishaps during the sequencing of gene libraries, and it's a construction that allows the establishment of qualitative gene sequencing.

Future Use

- Human cell line development for biomedical research.
- Construction of a genomic library of endangered species, which plays an important role in species conservative strategy.

Acknowledgments: We would like to thank the research department of the college for helping us conduct the study.

Conflict of interest: None

Financial support: None

Ethics statement: None

References

Andriese, T. S. (1998). *Stratagene's pSCOS1 as a cloning vector in the construction of a genomic library*. San Jose State University. doi:10.31979/etd.5use-3cge

Burman, A., Garcia-Milian, R., & Whirledge, S. (2020). Gene X environment: The cellular environment governs the transcriptional response to environmental chemicals. *Human Genomics*, *14*(1), 1-14. doi:10.1186/s40246-020-00269-1

Caucheteur, D., Robin, G., Parez, V., & Martineau, P. (2018). Construction of a synthetic antibody gene library for the selection of intrabodies and antibodies. *Phage Display: Methods and Protocols*, *1701*, 239-253. doi:10.1007/978-1-4939-7447-4_12

Choi, N. (2014). The application profiles and development characteristics of library open source software projects. *Library Hi Tech*, *32*(2), 260-275. doi:10.1108/lht-09-2013-0127

Ghule, V., Deshpande, A., Gurwale, S., Kambale, T., & Iqbal, B. (2023). Chondroblastoma in a distal phalanx of the great Toe--A rare case report. *Clinical Cancer Investigation Journal*, *12*(1), 14-16. doi:10.51847/OgnfaLJLxq

Girija, S. A., Jayaseelan, V. P., & Arumugam, P. (2018). Prevalence of VIM-and GIM-producing Acinetobacter baumannii from patients with severe urinary tract infection. *Acta Microbiologica et Immunologica Hungarica*, *65*(4), 539-550. doi:10.1556/030.65.2018.038

Jain, C. (2023). Evaluation of chromogenic agar media for isolation, identification and direct antibiotic susceptibility testing of uropathogens. *International Journal of Pharmaceutical Research & Allied Sciences*, *12*(2), 7-12. doi:10.51847/Kd4vp42v9B

Jayaraman, J., Halane, M. K., Choi, S., McCann, H. C., & Sohn, K. H. (2019). Using bioinformatics and molecular biology to streamline construction of effector libraries for phytopathogenic Pseudomonas syringae strains. *Plant Innate Immunity: Methods and Protocols*, 1991, 1-12. doi:10.1007/978-1-4939-9458-8_1

Langille, M. G., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., Clemente, J. C., Burkepile, D. E., Vega Thurber, R. L., Knight, R., et al. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology*, *31*(9), 814-821. doi:10.1038/nbt.2676

Mashhour, K., Saad, E., Abdelghany, H., & Hashem, W. (2023). 3D-CRT versus SIB IMRT acute toxicity outcomes in preoperative concurrent chemo-radiotherapy for locally advanced Rectal cancer. *Clinical Cancer Investigation Journal*, *12*(1), 36-42. doi:10.51847/uBAn5N4CCd

Misikir, H. M., Bardossy, A. C., Hartman, P., Moreno, D., Suleyman, G., & Perri, M. (2016). Risk factors associated with vancomycin-resistant enterococcus (VRE) and methicillin-resistant Staphylococcus aureus (MRSA) co-infection. *Clinical Microbiology and Infectious Diseases*, *1*(2), 42-44. doi:10.15761/cmid.1000109

Mukai, T., Yoneji, T., Yamada, K., Fujita, H., Nara, S., & Su'etsugu, M. (2020). Overcoming the challenges of megabase-sized plasmid construction in Escherichia coli. *ACS Synthetic Biology*, *9*(6), 1315-1327. doi:10.1021/acssynbio.0c00008

Park, S. E., Youn, Y. N. N., Yu, Y. M., & Kim, J. K. (2016). cDNA library construction and gene screening of Harmonia axyridis (Coleoptera: Coccinellidae) for gene functional analysis. In *2016 International Congress of Entomology*. ESA. doi:10.1603/ice.2016.112244

Petronis, Z., Zigmantavicius, J., & Januzis, G. (2023). Histologic and Histomorphometric analysis of sinus floor elevation using Calcium Phosphate materials: A systematic review. *Annals of Dental Specialty*, *11*(2), 7-14. doi:10.51847/yqkoQ0Tw5p

Pirone-Davies, C., McFarland, M. A., Parker, C. H., Adachi, Y., & Croley, T. R. (2020). The utility of genomic and transcriptomic data in the construction of proxy protein sequence databases for unsequenced tree nuts. *Biology*, *9*(5), 104. doi:10.3390/biology9050104

Priyadharsini, J. V., Girija, A. S., & Paramasivam, A. (2018). An insight into the emergence of Acinetobacter baumannii as an oro-dental pathogen and its drug resistance gene profile–An in silico approach. *Heliyon*, *4*(12), e01051. doi:10.1016/j.heliyon.2018.e01051

Rohweder, B., Lehmann, G., Eichner, N., Polen, T., Rajendran, C.,

Ruperti, F., Linde, M., Treiber, T., Jung, O., Dettmer, K., et al. (2019). Library selection with a randomized repertoire of (βα) 8-barrel enzymes results in unexpected induction of gene expression. *Biochemistry*, *58*(41), 4207-4217. doi:10.1021/acs.biochem.9b00579

Sabarathinam, J. E. M. B. U. L. I. N. G. A. M., Nallaswamy, D., & NP, M. (2017). Comparison of herbal mouthwash with commercially available 0.2% Chlorhexidine and 2% betadine mouthwashes in patients after stage-1 implant surgery. *Asian Journal of Pharmaceutical and Clinical Research*, 56. doi:10.22159/ajpcr.2017.v10i12.17786

Shahana, R. Y., & Muralidharan, N. P. (2016). Efficacy of mouth rinse in maintaining oral health of patients attending orthodontic clinics. *Research Journal of Pharmacy and Technology*, *9*(11), 1991-1993. doi:10.5958/0974-360x.2016.00406.6

Shahzan, M. S., Girija, A. S., & Priyadharsini, J. V. (2019). A computational study targeting the mutated L321F of ERG11 gene in C. albicans, associated with fluconazole resistance with bioactive compounds from Acacia nilotica. *Journal de Mycologie Médicale*, *29*(4), 303-309. doi:10.1016/j.mycmed.2019.100899

Ushanthika, T., Smiline Girija, A. S., Paramasivam, A., & Priyadharsini, J. V. (2019). An in silico approach towards identification of virulence factors in red complex pathogens targeted by reserpine. *Natural Product Research*, *35*(22), 1-6. doi:10.1080/14786419.2019.1641811

Van Dijk, E. L., Jaszczyszyn, Y., & Thermes, C. (2014). Library preparation methods for next-generation sequencing: Tone down the bias. *Experimental Cell Research*, *322*(1), 12-20. doi:10.1016/j.yexcr.2014.01.008

Yeslam, H. E., Aljadaani, A. K., Almalky, A. M., Zahran, M. M., & Hasanain, F. A. (2023). Effect of luting agent on the load-bearing capacity of milled hybrid ceramic single-tooth restoration. *Annals of Dental Specialty*, *11*(3), 68-76. doi:10.51847/l0X0mVhoeA

Zabarovsky, E. R. (2006). Genomic DNA libraries, construction and applications. *Reviews in Cell Biology and Molecular Medicine*. doi:10.1002/3527600906.mcb.200300064

Zakaev, T. T., Bakrieva, M. V., Alkhazova, R. T., Girkina, D. B., Chagarova, A. Y., & Polyanskaya, A. A. (2023). Cardiovascular safety in the treatment of chronic Rheumatic pathologies. *International Journal of Pharmaceutical Research and Allied Sciences*, *12*(2), 54-57. doi:10.51847/wXAjWjcJY7