The Ultimate Guide to Building, Managing, and Analyzing Your Plasmid Library

Table of contents:

eLabNext
EPPENDORF GROUP

Plasmids are small, circular, double-stranded DNA molecules that can exist independently of chromosomal DNA in various host model organisms. Since their discovery and application in the life sciences using recombinant DNA technology, they've become an indispensable part of any R&D lab performing genetic engineering, designing and expressing recombinant proteins, developing new vaccines or therapies, and studying gene function and regulation.

On the clinical side, plasmids have become a pivotal starting point for vaccines and viral vector-based treatments, such as gene therapies. Through careful control over promoter or terminator sequences, coding regions, selectable markers, and the origin of replication, you can control how many copies of and where your gene-of-interest is expressed or translated, etc. In other words, plasmids give selectivity and control over the expression of the genes they carry.

Given their pivotal nature, keeping track of your collection of plasmids is essential for error-free, reproducible research. Slight changes in plasmid sequence can occur during the cloning process, passaging through a prokaryotic or eukaryotic host or optimizing sequences to improve expression and efficiency. Quality can also degrade over time due to improper storage or contamination. Additionally, as the number of plasmids in a library increases, it becomes increasingly difficult to maintain an organized and accessible system for tracking plasmids, including recording their location, properties, and usage history.

To address these challenges, many labs implement careful record-keeping and storage protocols, including using unique identifiers, periodic quality control checks, and regularly updating digital plasmid databases using Microsoft Excel trackers or a dedicated Laboratory Information Management System (LIMS). When using these systems, the utmost attention and care need to be exercised: Mistakes or errors in the plasmid backbone, antibiotic resistance, selection marker, or optimal bacterial cells to transform into can lead to confusion, errors, and wasted time and resources.

In this whitepaper, we'll introduce some of the common plasmids used in the life science space and provide some best practices for building, maintaining, managing, and storing a plasmid library.

Building a Plasmid Library for Your Application

The saying goes, "Pick the right tool for the job." Different plasmids are used in an R&D setting to perform different tasks. So it's very common to keep various parental plasmids and their derivatives in your freezer, ready to use based on your planned experiments.

For example, to study the role of a gene in a disease model, you can build a plasmid library consisting of various functional domains of the gene or variants with missing domains and targeted mutations. This way, you can decipher the biological role of that gene, the functional domain(s) of the gene, and critical amino acid residues involved in the functioning of that gene. In addition, you could create plasmid libraries based on different projects you are working on and processes used to prepare the plasmid DNA (research versus clinical grade material or simple DNA prep versus endotoxin-free plasmid prep for mammalian cells).
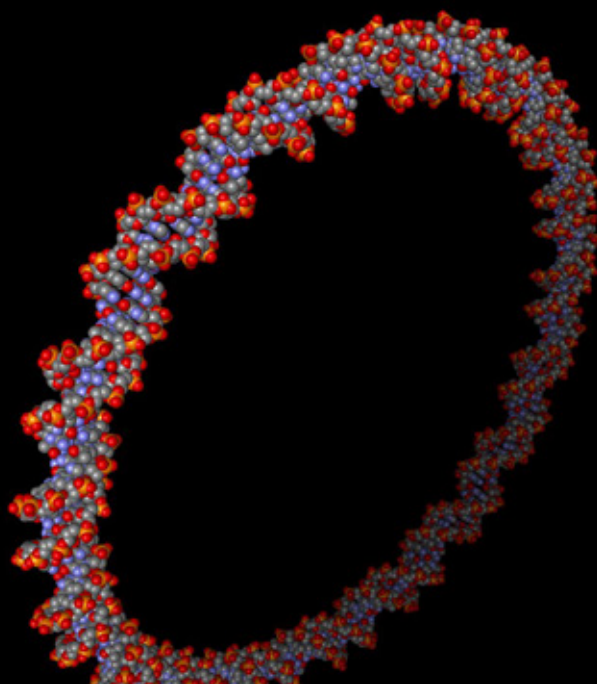
Keeping track of various plasmids with all the relevant information – including backbone, cloning strategy, purification strategy, etc. – becomes critical and essential to generating reliable and reproducible results.

The Most Widely Used Plasmids in R&D

Plasmids are employed in fields such as molecular biology, genetic engineering, biotechnology, and synthetic biology for various applications, such as studying gene function, gene expression/editing, creating genetically modified organisms, producing recombinant proteins, and developing new therapies.

For these applications, there are families of well-known, validated plasmids that are commonly deployed.

pUC19 vectors

This plasmid is commonly used for cloning and protein expression in E. coli. It contains an origin of replication, lac promoter, selectable markers, and restriction sites for inserting DNA sequences.

Here are some examples of the applications where pUC19 plasmids have been used:

- **Development of DNA sequencing methods:**
In 1985, Sanger and colleagues published a paper describing a method for sequencing DNA using modified versions of pUC19 and other plasmids. This method revolutionized DNA sequencing and helped pave the way for the Human Genome Project.

- **Production of recombinant proteins:**
pUC19 is commonly used as a vector for expressing recombinant proteins in bacterial cells. For example, in a 1998 study published in the Journal of Immunology, researchers used pUC19-based vectors to express and purify a recombinant protein that was used to generate antibodies against the influenza virus.

- **Genetic engineering of crops:**
In a 2003 study published in the Proceedings of the National Academy of Sciences, researchers used pUC19-based vectors to create transgenic tomato plants that expressed a gene from a bacterium that conferred resistance to the herbicide glyphosate.

eLabNext
EPPENDORF GROUP

- **Study of bacterial genetics:**

pUC19 is often used in studies of bacterial genetics and gene regulation. For example, in a 1994 study published in Molecular Microbiology, researchers used pUC19-based vectors to study the regulation of a gene involved in the flagellar synthesis.

pET vectors

These vectors are used for high-level protein expression in E. coli. The name "pET" stands for "plasmid for expression of T7 RNA polymerase," which refers to the fact that these vectors use the T7 promoter to drive the expression of cloned genes in E. coli cells.

Key features that make pET vectors useful for protein expression studies are:

- **T7 promoter:**

A strong and efficient promoter that drives transcription of cloned genes in E. coli cells.

- **Selection markers:**
Typically contain selection markers for antibiotic resistance, allowing researchers to select for cells that have taken up the plasmid.

- **Multiple cloning sites:**
Allows researchers to insert genes-of-interest downstream of the T7 promoter.

- **Fusion tags:**
Often include fusion tags, such as hexahistidine tags or GST tags, used to purify the expressed protein from bacterial lysates.

- **Inducible expression:**
Inducible promoters, such as the lac promoter, allow researchers to control the timing and level of protein expression.

- **Diverse applications:**
Can express/study proteins-of-interest such as enzymes, antibodies, and other biologically active molecules.

pGEX vectors

These plasmids are used to express and purify recombinant proteins fused with glutathione S-transferase (GST) in E. coli. The GST tag allows for easy purification of the fusion protein, and the tag can be removed using proteases.

Common applications include:

- **Protein purification:**
The GST tag allows easy protein purification using glutathione-affinity chromatography.

- **Protein-protein interaction studies:**
Such recombinant proteins can be used in protein-protein interaction studies.

- **Enzyme assays:**
Such recombinant proteins can be used in enzyme assays to study a protein's enzymatic activity.

- **Antibody production:**
Such recombinant proteins can be used to generate antibodies for use in various applications.

pBABE vectors

These are widely used for retroviral gene transfer and stable gene expression in mammalian cells. They contain selectable markers and viral elements required for packaging and infecting target cells. The pBABE vectors allow for the introduction of exogenous DNA into target cells, where it integrates into the host genome and is stably expressed.

Some of the common applications include:

- **Gene expression studies:**
Used to overexpress genes-of-interest in target cells to study their function.

- **Functional assays:**
Used to express proteins or other molecules in target cells for functional assays, such as cell proliferation or apoptosis assays.

- **Gene therapy:**
Used to deliver therapeutic genes to target cells for the treatment of genetic disorders or other diseases.

- **Stem cell research:**
Used to introduce genes into stem cells to study their differentiation or to generate transgenic animals.

eLabNext
EPPENDORF GROUP

Lentiviral vectors

These popular viral vectors are used for gene transfer and gene therapy in mammalian cells.

They are related to the human immunodeficiency virus (HIV) but are modified to be replication-defective and safe for use in laboratory and clinical settings. They can be used to express genes in various cell types, including difficult-to-transfect cells.
Some of the most common uses include:

- **Gene delivery:**
Efficiently deliver exogenous DNA into various cell types, including both dividing and non-dividing cells, allowing for stable gene expression.

- **Gene editing:**
Used to deliver gene-editing tools such as CRISPR/Cas9 into target cells for gene editing applications.

- **Cancer therapy:**
Used to deliver therapeutic genes or anti-cancer drugs to cancer cells to treat cancer.

- **Stem cell research:**
Used to introduce genes into stem cells to study their differentiation or to generate transgenic animals.

- **Vaccine development:**
Used to develop vaccines against a variety of infectious diseases.

Top 10 Plasmid Providers

With the widespread use of plasmids in R&D, many commercial vendors have sprung up to provide high-quality plasmids to researchers. You can choose where you source your plasmids depending on your needs, affiliations, and budget (Addgene and DNASU are non-profit, while the others listed below are commercial operations).

Some of the most common uses include:

- **Addgene**
- **DNASU (Arizona State University)**
- **Genscript**
- **Thermo Fisher Scientific**
- **ATUM (formerly DNA2.0)**
- **OriGene Technologies**
- **VectorBuilder**
- **Genewiz**
- **Source BioScience**

Molecular Biology Techniques for Working with Plasmids

There are a variety of techniques used in the wet lab for building and maintaining plasmid libraries. The techniques used depend on the specific needs of the project.

Here are some of the most common techniques used when working with plasmids:

- **PCR amplification:**

PCR is a widely used technique for amplifying DNA sequences such as genes or their fragments from a plasmid.

- **Restriction enzyme digestion:**

Restriction enzymes are used to excise DNA at specific recognition sites, allowing for the isolation of specific DNA fragments, such as genes or their fragments from a plasmid.

- **Ligation:**

Ligation can be used to insert DNA fragments into plasmids for library construction. These DNA fragments could be products of PCR amplification or restriction digestion.

- **Transformation:**
Transformation is the process of introducing plasmids into bacterial cells, allowing for plasmid replication and maintenance.

- **Selection:**
The use of antibiotics or fluorescence-based selection is essential to select and maintain cells with plasmids post-transformation.

- **DNA extraction and purification:**

This involves isolating DNA from bacterial cells using various methods such as alkaline lysis, precipitation, and column- or bead-based purification.

- **Specialized cloning techniques:**

Novel cloning techniques such as Gibson assembly or Golden Gate assembly can also be employed for stitching plasmids together.

Overall, the best techniques for plasmid library construction and maintenance depend on the project's specific needs, including the library's size, the type of plasmids used, and the downstream applications.

eLabNext
EPPENDORF GROUP

How to Make Your Plasmid Library Sustainable

To have a sustainable supply of the plasmids in your library means you have access to an established, reliable source of high-quality reagents. This is crucial as plasmids are the building blocks for most R&D projects and therapies. Without pure and validated plasmids, keeping research (particularly in the fast-paced biotech industry) going would be impossible.

Follow the steps below to ensure that your plasmid library stays sustainable:

1 **Select the plasmids that are right for you:**

The first step in building a sustainable plasmid library is to select plasmids relevant to the organization's research needs. This may involve identifying plasmids that have been used in published studies or that are associated with particular research areas.

2 **Implement quality control:**

Once plasmids have been selected, they should undergo quality control (QC) measures to ensure they are pure, intact, and function as expected. This may involve sequencing plasmids to confirm identity, testing them for activity, and storing them under appropriate conditions to prevent degradation or contamination. As per a pre-determined schedule, periodic QC checks will need to be performed to ensure the library stays in high-quality condition.

3 **Standardize plasmid preparation protocols:**

Develop standardized (ideally, automated) protocols for preparing plasmids from bacterial cultures, including methods for isolating plasmid DNA and verifying plasmid integrity and purity.

4 **Catalog key information:**

Careful organization and management of the library are critical to ensure that plasmids are easily accessible and can be efficiently tracked and cataloged. This will involve using or developing a database to track plasmids, assigning each a unique identifier, and storing plasmids in a consistent and accessible manner.

5 **Share and collaborate:**

A sustainable plasmid library should encourage research sharing and collaboration. This may involve providing access to plasmids to researchers within or outside the organization (relevant to academic/educational institutions) through publicly accessible repositories.

6 **Use a storage and tracking system:**

The plasmids should be stored to maintain their stability and integrity, such as in glycerol stocks at -80°C. A reliable tracking system should also be implemented to keep track of plasmid identities, locations, and related information. Establishing a system for tracking the plasmids in your library is important, including information about the plasmid name, source, sequence, and any relevant annotations or notes.

7 **Manage data:**

Once plasmids have been selected, they should undergo quality control (QC) measures to ensure they are pure, intact, and function as expected. This may involve sequencing plasmids to confirm identity, testing them for activity, and storing them under appropriate conditions to prevent degradation or contamination. As per a pre-determined schedule, periodic QC checks will need to be performed to ensure the library stays in high-quality condition.

eLabNext
EPPENDORF GROUP

Following the above steps, organizations can build a sustainable plasmid library supporting ongoing research and scientific discovery.

Best Practices for Maintaining a Plasmid Library

Managing 1,000 plasmids may be easy for the average R&D lab, but what if you had to build a biobank of 20,000 or more? Establishing and following standardized and uniform naming conventions is the key to success.

Check out some tips for developing an effective plasmid naming convention:

- **Use a consistent format:**

Establish a consistent format for plasmid names that include key information, such as the plasmid backbone, selection markers, and relevant inserts or modifications.

- **Include a unique identifier:**

Assign a unique identifier to each plasmid in the library, such as a numerical or alpha-numeric code, to help distinguish it from other plasmids in the collection.

- **Avoid special characters and spaces:**

Avoid using special characters or spaces in plasmid names, as these can create issues with file naming, data management, and printing labels.

- **Keep names short and descriptive:**

Use descriptive names that are easy to remember and relate to the plasmid's key features or purpose, but keep names relatively short to avoid confusion.

- **Document the naming convention:**

Document the plasmid naming convention and make it easily accessible to all researchers who will be using the plasmid library.

- **Future-proof:**

Consider if your naming convention is compatible with other systems you may design/have in the future. So use a standardized naming convention that is widely recognized in the field.

By establishing and following these guidelines, you can ensure your plasmid libraries are well-organized and easy to manage.

Best Practices for Storing a Plasmid Library

Proper storage practices are essential for maintaining the quality and integrity of a plasmid library. Here are some of the best practices:

- **Use proper storage containers:** Plasmids should be stored in sterile, airtight containers that are designed for long-term storage. Plastic tubes or plates made of materials such as polypropylene are commonly used.

- **Store plasmids at the correct temperature:** Plasmids should be stored at -80°C for long-term and -20°C for short-term storage needs.

- **Label samples:** Proper labeling of containers (both the cap and sides) is crucial. At the minimum, labels should include the plasmid name, vector type, antibiotic resistance, and the date the plasmid was added to the library.

- **Backup your library:** Maintaining a backup library of plasmids is important in case of loss or damage to the primary library. This backup should be stored in a separate freezer.

- **Maintain a master stock:** Maintain a master stock of the plasmid library, which should be stored at -80°C. This should not be working stock; this separation ensures that the plasmids can be retrieved in case of any loss or contamination.

- **Perform periodic QC checks:** Plasmids should be periodically checked for quality and stability over time. This can be done by performing restriction enzyme digestions or sequencing and regularly thawing and retransforming plasmids to ensure viability.

- **Regularly check for contamination:** Periodically check plasmid libraries for contamination by performing colony PCR, restriction enzyme digestion, or sequencing to ensure the library remains free of unwanted bacterial strains.

- **Use appropriate buffers:** Store plasmid libraries in appropriate buffers to maintain plasmid stability. Common buffers include TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0) or SE buffer (10 mM Tris, 1 mM EDTA, 100 mM NaCl, pH 7.5).

Tools for In Silico Plasmid Library and Sequence Management

Many software tools are available for molecular biology analysis, and the best tools depend on the analysis being performed.
Popular software tools include:

- **CloneAssist:** This is an eLabNext Electronic Lab Notebook (ELN) software feature that helps researchers track and manage their cloning workflows.

- **SnapGene**: A user-friendly software for DNA sequence analysis, primer design, and plasmid mapping.

eLabNext
EPPENDORF GROUP

- **Geneious:** Geneious is a comprehensive software platform for molecular biology research, including sequence alignment, analysis, and annotation. It supports a wide range of sequencing data types and offers advanced features such as phylogenetic analysis and primer design.

- **CLC Genomics Workbench:** CLC Genomics Workbench is a software suite for analyzing high-throughput sequencing data, including DNA, RNA, and protein sequences. It offers a range of tools for data visualization, alignment, assembly, annotation, statistical analysis, and data mining.

- **PyMOL:** PyMOL is a molecular visualization system that allows users to create high-quality 3D images of proteins, nucleic acids, and other macromolecules. It offers advanced molecular modeling, animation, and molecular dynamics simulations.

- **MEGA:** MEGA (Molecular Evolutionary Genetics Analysis) is a software package for phylogenetic analysis and molecular evolution studies. It offers a range of tools for sequence alignment, tree building, and molecular evolution analysis.

- **GROMACS:** GROMACS (GROningen MAchine for Chemical Simulations) is a software package for molecular dynamics simulations of proteins, lipids, and nucleic acids. It offers a range of tools for simulation setup, analysis, and visualization and support for GPU acceleration.

- **NCBI BLAST:** BLAST (Basic Local Alignment Search Tool) is a widely used bioinformatics tool for sequence similarity searching. It allows users to search for sequences in large sequence databases and identify homologous sequences.

- **R:** R is a programming language and software environment for statistical analysis, data visualization, and machine learning. It is widely used in bioinformatics for data analysis and visualization.

- **ClustalW**: A tool for multiple sequence alignment of nucleotide or protein sequences.

- **Primer3:** A tool for designing primers for PCR and other applications based on user-defined parameters.

- **Bowtie:** A fast and memory-efficient tool for aligning short DNA sequences to a large reference genome.

- **GATK:** A software suite for genomic analysis that includes tools for variant discovery, genotyping, and more.

- **SPAdes:** A genome assembly tool that can assemble genomes from Illumina and PacBio sequencing data.

eLabNext
EPPENDORF GROUP

Paying Tribute to the Workhorses of Molecular Biology: SnapGene and GenBank

At eLabNext, we hold a special place in our heart for SnapGene and consider it the godfather of modern-day molecular biology.

Listed below are some key features:

- **Sequence annotation:**
SnapGene allows users to easily annotate and visualize DNA sequences, including genes, promoters, and restriction sites.

- **Plasmid mapping:**
SnapGene can generate maps of plasmids and other DNA sequences, annotating genes, restriction sites, and other features.

- **DNA sequence alignment:**
SnapGene includes a powerful sequence alignment tool that can be used to align multiple DNA sequences and compare them side-by-side.

- **Primer design:**
Snapgene includes a primer design tool that automatically designs primers for PCR, cloning, and other applications based on user-defined parameters.

- **Cloning simulation:**
SnapGene can simulate cloning experiments and predict the outcome of cloning procedures based on the user's input.

- **DNA sequence editing:**
SnapGene provides a user-friendly interface for editing DNA sequences, including features for adding and deleting nucleotides, moving and copying sequences, and more.

- **Visualization tools:**
The software includes various visualization tools, including the ability to create interactive maps of plasmids, visualize DNA sequences in 3D, and more.

- **Protein visualization:**

The software can display translations of DNA sequences in amino acid sequences and provide options for visualizing protein domains and motifs.

- **Collaboration:**
SnapGene includes features that facilitate collaboration among team members, including the ability to share plasmid designs, comments, and annotations.

- **Restriction analysis:**
SnapGene includes a database of common restriction enzymes and can simulate restriction enzyme digestions of DNA sequences, with the ability to specify multiple enzymes and generates of the results.

SnapGene is a powerful molecular biology research tool compatible with GenBank, a widely used resource in research and development.

Here are some of its uses:

- **Genome sequencing:**

GenBank provides a vast repository of genomic sequence data for various organisms. Researchers can use this data to identify genes, analyze their function, and compare genomes across different species.

- **Drug discovery:**

GenBank can be used to identify genes that are involved in diseases, such as cancer or infectious diseases. This information can be used to develop new drugs or therapies that target these genes.

- **Biotechnology applications:**

GenBank data can be used to design and engineer new biological systems, such as genetically modified crops or bacteria that can produce useful chemicals.

- **Evolutionary studies:**

GenBank provides a wealth of information about the genetic makeup of different organisms, which can be used to study the evolution of different species and to investigate the relationships between different groups of organisms.

- **Diagnostic applications:**

GenBank data can be used to develop diagnostic tests for genetic diseases or for identifying specific microorganisms in clinical samples.

GenBank is a critical resource for researchers and developers in the field, providing a vast repository of genetic sequence data that can be used to advance scientific understanding, develop new technologies, and improve human health.

eLabNext
EPPENDORF GROUP

Conclusion

Whether you are just starting out or already have a substantial biobank/plasmid library, it is essential to follow strict guidelines, to maintain sustainable lab operations and workflow. The key is not to decentralize your data across many different platforms but to have **one cohesive digital lab platform** that can interface and collaborate with your favorite software, instruments, and storage solutions.

Whether using SnapGene for sequence analysis and a Thermo freezer for sample storage, always give your operations proper planning for a holistic solution!

Learn more about how eLabNext can enable proper plasmid library management and storage solutions **here** or **contact us** directly to guide you in your Lab Digital Transformation journey!

Author: Zareh Zurabyan

eLabJournal

About eLabJournal

eLabJournal offers an intuitive and flexible solution to manage information in your lab. The all-in-one Electronic Lab Notebook also includes modules for sample tracking and protocol management. eLabJournal improves efficiency when documenting, organizing, searching and archiving data, samples and protocols. The software is suitable for any lab ranging from small academic laboratories and strart-up companies to large academic institutes and globally operating companies.

eLabNext
EPPENDORF GROUP

✉ enquiries@elabnext.com
📞 +31 50 720 00 55

All of our product specialists have a scientific background and are happy to discuss your needs. Schedule a demo for a free, no-obligation product demonstration.

CONTACT US

www.elabnext.com