Plasmids 101:
A Desktop Resource

Created and Compiled by Addgene
September 2014 (1st Edition)
www.addgene.org

addgene
A better way to share plasmids

Plasmids 101: Introduction to Addgene's Resource

Any newcomer who joins a molecular biology lab will undoubtedly be asked to design, modify, or construct a plasmid. Although the newcomer likely knows that a plasmid is a small circular piece of DNA found in bacterial cells, she may need some extra guidance to understand the specific components that make up a plasmid and why each is important.

Our mission with this eBook, Plasmids 101: A Desktop Resource, is to curate a one-stop reference guide for plasmids. This resource is designed to educate all levels of scientists and plasmid lovers and serves as an introduction to plasmids, allowing you to spend less time researching plasmid basic features and spend more time developing the clever experiments and innovative solutions necessary for advancing the field.

Looking for more useful information from Addgene? Be sure to check out other Addgene resources:

www.addgene.org
www.facebook.com/Addgene
www.twitter.com/addgene

TABLE OF CONTENTS

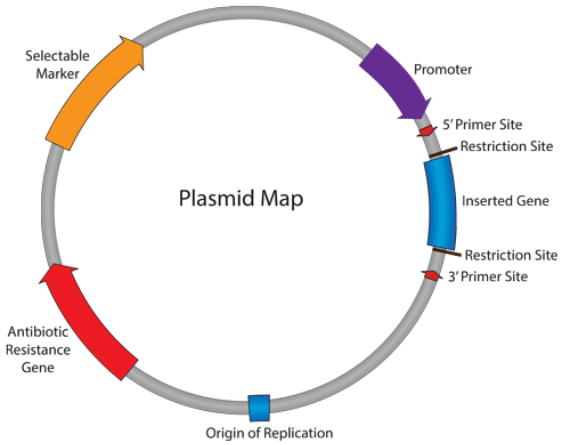
TABLE OF CONTENTS (CONT)

TABLE OF CONTENTS (CONT)

CHAPTER 1: WHAT IS A PLASMID

By Margo R. Monroe, Addgene with contributions from Marcy Patrick, Addgene | Jan 14, 2014

At their most basic level, plasmids are small circular pieces of DNA that replicate independently from the host's chromosomal DNA. They are mainly found in bacteria, but also exist naturally in archaea and eukaryotes such as yeast and plants. In nature, plasmids provide one or more functional benefits to the host such as resistance to antibiotics, degradative functions, and/or virulence. All natural plasmids contain an origin of replication (which controls the host range and copy number of the plasmid) and typically include a gene that is advantageous for survival, such as an antibiotic resistance gene. In contrast, plasmids utilized in the lab are usually artificial and designed to introduce foreign DNA into another cell. Minimally, lab-created plasmids have an origin of replication, selection marker, and cloning site. The ease of modifying plasmids and the ability of plasmids to self-replicate within a cell make them attractive tools for the life scientist or bioengineer.



Vector Element	Description
Origin of Replication (ORI)	DNA sequence which allows initiation of replication within a plasmid by recruiting transcriptional machinery proteins
Antibiotic Resistance Gene	Allows for selection of plasmid-containing bacteria.
Multiple CloningSite (MCS)	Short segment of DNA which contains several restriction sites allowing for the easy insertion of DNA. In expression plasmids, the MCS is often downstream from a promoter.
Insert	Gene, promoter or other DNA fragment cloned into the MCS for further study.
Promoter Region	Drives transcription of the target gene. Vital component for expression vectors: determines which cell types the gene is expressed in and amount of recombinant protein obtained.
Selectable Marker	The antibiotic resistance gene allows for selection in bacteria. However, many plasmids also have selectable markers for use in other cell types.
Primer Binding Site	A short single-stranded DNA sequence used as an initiation point for PCR amplification or sequencing. Primers can be exploited for sequence verification of plasmids.

The above plasmid map and table outline the common engineerable features of plasmids.

CHAPTER 1: WHAT IS A PLASMID? (CONT.)

How is a Plasmid Constructed in the Lab?

Due to their artificial nature, lab plasmids are commonly referred to as "vectors" or "constructs". To insert a gene of interest into a vector, scientists may utilize one of a variety of cloning methods (restriction enzyme, ligation independent, Gateway, Gibson, and more). The cloning method is ultimately chosen based on the plasmid you want to clone into. Regardless, once the cloning steps are complete, the vector containing the newly inserted gene is transformed into bacterial cells and selectively grown on antibiotic plates.

Addgene has compiled various educational resources to facilitate plasmid use in the lab. Addgene's Online Plasmid Guide includes information about molecular cloning, how to choose a plasmid vector, molecular biology tools and references, and how to maintain your plasmid stocks. The guide also contains multiple protocols and troubleshooting tips to make plasmid usage as simple and straightforward as possible.

How do Scientists Use Plasmids?

Generally, scientists use plasmids to manipulate gene expression in target cells. Characteristics such as flexibility, versatility, safety, and cost-effectiveness enable molecular biologists to broadly utilize plasmids across a wide range of applications. Some common plasmid types include: Cloning plasmids, Expression plasmids, Gene knock-down plasmids, Reporter plasmids, Viral plasmids, and Genome engineering plasmids.

To date, scientists around the world are extensively using these vectors for experiments encompassing fluorescent imaging, recombinant DNA technology, mass protein production, disease modeling, drug discovery, and genome editing (just to name a few).

Where Can I Find Additional Resources for Using Plasmids?

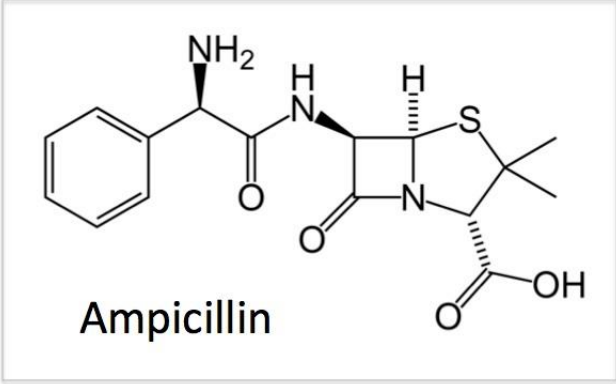
In addition to this eBook, Addgene has complied more details on the history, importance, and types of plasmids in the Addgene Molecular Biology Plasmid Reference Guide on our website.

In this chapter, we will cover the basics of various plasmid elements, including the antibiotic resistance gene, origin of replication, promoter, and more. We also have tables and charts for you to use as references at the lab bench and practical tips for your experiments.

ANTIBIOTIC RESISTANCE GENES

By Marcy Patrick, Addgene | Jan 30, 2014

Resistance to antibiotics is a widely used tool in molecular biology, yet scientists rarely stop to think about how much easier it makes our lives. Plasmid transformation into *E. coli* is a fairly inefficient process– just 1 out of 10,000 cells on average! Without some means of quickly determining which cells successfully received the correct plasmid, scientists would spend hours to days trying find their correct clones. Additionally, the presence of a plasmid is disadvantageous from the bacterium's perspective – a plasmid-containing cell must replicate the plasmid in addition to its own chromosomal DNA, costing additional resources to maintain the plasmid. Adding an antibiotic resistance gene to the plasmid solves both problems at once – it allows a scientist to easily detect plasmid-containing bacteria when the cells are grown on selective media, and provides those bacteria with a pressure to keep your plasmid. Viva la (bacterial) resistance!



Ampicillin

What are Antibiotics?

Antibiotics are generally defined as agents that kill bacteria, or inhibit their growth. Although originally sourced from natural products, many common antibiotics used in labs today are semi-synthetic or fully synthetic compounds. Antibiotics can be categorized based on whether they directly kill bacteria (bactericidal) or slow growth/prevent cell division (bacteriostatic); however, the distinction between the two categories may be a bit of a gray area as some bacteriostatic reagents can kill bacteria when used at high concentrations (and vice versa). Looking around the lab, you'll likely find many of the antibiotics listed in the table below. Note, in this post we'll focus primarily on antibiotics against Gram negative bacteria. In future posts, we'll detail selection in non-bacterial cells such as yeast or mammalian cells.

How Else Can Antibiotics Be Used in the Lab?

Historically, antibiotics have also been used to disrupt genes at the chromosomal level. Scientists introduce an antibiotic resistance cassette within the coding region of the gene they are trying to disrupt or delete, which both inactivates the gene and acts as a marker for the mutation. When designing these types of experiments it is best practice not to use the same resistance cassette for the mutation and for plasmid selection. Additionally, scientists can use the loss of resistance as a marker for successful cloning. In these instances, the cloning vector typically has two separate resistance cassettes and your gene of interest is cloned into/inactivates or completely removes (in the case of Gateway cloning) one cassette. Counter selection allows the scientist to select bacteria that are only resistant to the antibiotic that remains intact.

ANTIBIOTIC RESISTANCE GENES (CONT.)

Tips and Tricks from the Bench:

- Use fresh stocks. Most antibiotics are stable in powder form, but quickly breakdown in solution. Storing aliquots at -20°C and avoiding repeated freeze/thaw cycles will keep most antibiotics viable for at least 6 months.
- Ampicillin breaks down especially fast and plates should be used within 1 month for optimal efficiency. Beware of satellite colonies!
- Carbenicillin is more stable than Ampicillin and can be used in place of Ampicillin in most applications.
- Antibiotics vary in their sensitivity to heat and/or light – do not add them to media hotter than about 55°C and store plates/stocks wrapped in foil if a light-sensitive antibiotic like Tetracycline is used.
- Keep in mind that some *E. coli* strains have natural antibiotic resistances, so make sure your plasmid and *E. coli* strain are compatible! Check out this list of common *E. coli* genotypes and their natural resistances online from OpenWetWare.

COMMON ANTIBIOTICS TABLE

The table below lists antibiotics commonly found in the lab, their mechanism for killing bacteria, and their general working concentrations. For instructions on how to prepare antibiotic stocks, see Addgene's Online Reference Page.

Name	Class	Mode of Action*		Working Concentration**
Kanamycin	aminoglycoside	Binds 30S ribosomal subunit; causes mis-translation	Bactericidal	50-100 µg/mL
Spectinomycin	aminoglycoside	Binds 30S ribosomal subunit; interrupts protein synthesis	Bactericidal	7.5-50 µg/mL
Streptomycin	aminoglycoside	Inhibits initiation of protein synthesis	Bactericidal	25-100 µg/mL
Ampicillin	beta-lactam	Inhibits cell wall synthesis	Bactericidal	100-200 µg/mL
Carbenicillin	beta-lactam	Inhibits cell wall synthesis	Bactericidal	100 µg/mL
Bleomycin	glycopeptide	Induces DNA breaks	Bactericidal	5-100 µg/mL
Erythromycin	macrolide	Blocks 50S ribosomal subunit; inhibits aminoacyl translocation	Bacteriostatic	50-100 µg/mL in EtOH
Polymyxin B	polypeptide	Alters outer membrane permeability	Bactericidal	10-100 µg/mL
Tetracycline	tetracyclin	Binds 30S ribosomal subunit; inhibits protein synthesis (elongation step)	Bacteriostatic	10 µg/mL
Chloramphenicol		Binds 50S ribosomal subunit; inhibits peptidyl translocation	Bacteriostatic	5-25 µg/mL in EtOH

**In prokaryotes.* ***Dissolve in dH_2O and sterile filter unless otherwise specified.*

ORIGIN OF REPLICATION

By Kendall Morgan, Addgene with contributions from Marcy Patrick, Addgene | Feb 6, 2014

Let's consider another basic elements of any plasmid: the origin of replication/replicon. The replicon is comprised of the origin of replication (ORI) and all of its control elements. The ORI is the place where DNA replication begins, enabling a plasmid to reproduce itself as it must to survive within cells.

The replicons of plasmids are generally different from those used to replicate the host's chromosomal DNA, but they still rely on the host machinery to make additional copies. ORI sequences are generally high in As and Ts. Why, you ask? Well, A-T base pairs are held together with two hydrogen bonds not three as G-C pairs are. As a result, stretches of DNA that are rich in A-T pairs melt more readily at lower temperatures. When DNA melts, it gives the replication machinery room to come in and get busy making copies.

So Many Origins, So Little Time

There are lots of ORIs out there so, for simplicity's sake, we've ignored those used in eukaryotic cells and viruses and focused only on those found in bacteria. Some common ones you might see include ColE1, pMB1 (which comes in a few slightly different but well known derivatives), pSC101, R6K, and 15A. Not all origins of replication are created equal. Some will produce many plasmid copies and others produce just a few copies depending on how they are regulated. Generally, control of replication is referred to as "relaxed" or "stringent" depending on whether the ORI is positively regulated by RNA or proteins, respectively. A plasmid's copy number has to do with the balance between positive and negative regulation and can be manipulated with mutations in the replicon. For example, the pMB1 ORI maintains about 20 copies per cell, while pUC – which differs by only two mutations – will produce as many as 700 copies per cell.

So, how do you choose? Addgene Senior Scientist Marcy Patrick says researchers can ask themselves a few questions to get started: Will the plasmid be used exclusively in *E. coli*? Gram negative bacteria in general? Both Gram negatives and Gram positives? Will you have only one plasmid type in your cells at a time? Do you want to make a lot of your plasmid? Is the gene toxic in high amounts? It is always good to keep in mind that plasmids with low to medium copy numbers can still express massive amounts of protein given the proper promoter (stay tuned!) and growth conditions.

Choose Your Origin of Replication Wisely

The best choice of ORI depends on how many plasmid copies you want to maintain, which host or hosts you intend to use, and whether or not you need to consider your plasmid's compatibility with one or more other plasmids. Generally speaking, plasmids with the same ORIs are incompatible because they will compete for the same machinery, creating an unstable and unpredictable environment. As a rule, plasmids from the same

ORIGIN OF REPLICATION (CONT.)

group should not be co-transformed, so if you require two plasmids for an experiment, make sure they have "compatible" ORIs.

The table below defines common cloning vectors, their copy number, ORI, and incompatibility groups. Note the A and B compatibility grouping is an arbitrary designation, and plasmids from the same incompatibility group should not be co-transformed.

Common Vectors	Copy Number*	ORI	Incompatibility Group	Control
pUC	~500-700	pMB1 (derivative)	A	Relaxed
pBR322	~15-20	pMB1	A	Relaxed
pET	~15-20	pBR322	A	Relaxed
pGEX	~15-20	pBR322	A	Relaxed
pColE1	~15-20	ColE1	A	Relaxed
pR6K	~15-20	R6K*	B	Stringent
pACYC	~10	P15A	B	Relaxed
pSC101	~5	pSC101	B	Stringent
pBluescript	~300-500	ColE1 (derivative) and F1**	A	Relaxed
pGEM	~300-500	pUC and F1	A	Relaxed

[+]*Actual copy number varies. See below for additional considerations.*

**Requires pir gene for replication (reference).*

***F1 is a phage-derived ORI that allows for the replication and packaging of ssDNA into phage particles. Plasmids with phage-derived ORIs are referred to as phagemids.*

ORIGIN OF REPLICATION (CONT.)

Other Factors that Affect Copy Number

Although the sequence and regulation of the ORI dramatically affect the copy number of a plasmid, other external factors contribute as well. These considerations are especially useful to keep in mind if you are planning to purify your plasmid DNA:

The insert:

- Bacteria tend to maintain fewer copies of plasmids if they contain large inserts or genes that create a toxic product.

The E. coli strain:

- Most *E. coli* strains can be used to propagate plasmids, but endA- *E. coli* are best for high yields of plasmids.

Growth conditions:

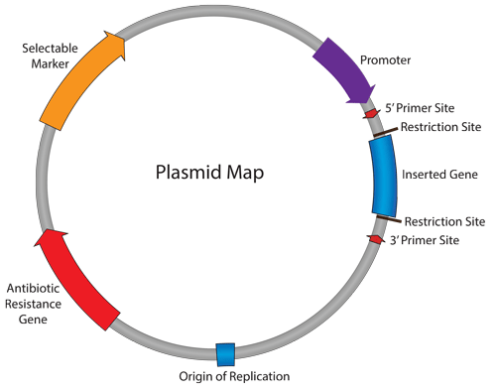
- The amount of aeration, temperature, culture volume, antibiotic, and medium can all affect copy number. Some ORIs are temperature sensitive; others ORIs can be "tricked" into amplifying more copies with the addition of Chloramphenicol – make sure your growth conditions aren't working against you!

The culture inoculum:

- Freshly streaked bacteria have higher copy numbers – for optimal results always pick a single colony and do not subculture directly from glycerol stocks, agar stabs, or liquid cultures.
- Incubation for 12-16 hours tends to give higher copy numbers since the bacteria have just reached stationary phase, but the cells have not started to die off.

THE PROMOTER REGION – LET'S GO!

By Kendall Morgan, Addgene with contributions from A. Max Juchheim, Addgene and Marcy Patrick, Addgene | Apr 3, 2014



Up to this point we can replicate our plasmid and make sure cells maintain it; the next step is getting the plasmid to express our gene of interest. Enter the promoter - the element responsible for initiating the transcription of your insert into RNA.

In practice, the term "promoter" describes the combination of the promoter (RNA polymerase binding site) and operators (response elements). Promoters are about 100 to 1000 base pairs long and found upstream of their target genes. The sequence of the promoter region controls the binding of the RNA polymerase and transcription factors, therefore promoters play a large role in determining where and when your gene of interest will be expressed.

The RNA Polymerase(s)

RNA is transcribed from DNA using an RNA polymerase (RNAP). In bacteria this is done by a single enzyme; however, eukaryotes have multiple polymerases which are each responsible for a specific subset of RNAs. To gain this specificity, the eukaryotic RNAP can recognize and bind to specific promoter elements. This means that the promoter present in your plasmid backbone must to be compatible with the type of RNA that needs to be made: if you want mRNA (for gene expression) you need to use an RNAP II promoter, whereas small RNAs (such as shRNA) are transcribed from the RNAP III promoters. This post features promoters for general RNAP II and RNAP III transcription; however, using viral LTRs as RNAP II promoters is commonly employed in lentiviral and retroviral constructs and we will discuss these in a future post on viral vector parts.

Promoter Specificity

Aside from choosing a promoter based on type of RNA transcript, you will also need to make sure your plasmid has a promoter suited to working in your host organism. Because transcription machinery differs between cell types or organisms, promoters must be similarly variable. Bacterial promoters only work in prokaryotic cells and typically only in the same or closely related species from which they were derived. Similarly, the various eukaryotic cell types (mammalian, yeast, plants, etc.) require unique promoters and there is very little crossover. Generally speaking, promoters in bacteria are less diverse and complex, having fewer parts than those in eukaryotic cells. Some promoters are constitutively active and on all the time while others are more carefully controlled.

Regulated promoters might act only in certain tissues or at certain times in development or there may be ways to turn them on or off at will with a chemical, heat, or light. In the cell, promoters themselves are

THE PROMOTER REGION – LET'S GO! (CONT.)

controlled by still other regulatory factors: enhancers, boundary elements, insulators, and silencers; however, some "leakiness" of transcription may occur. This is normally not a big issue for cells, but it may confound research results or even kill your cells if your gene of interest is toxic. To combat this, scientists have created synthetic promoters, which typically include some combination of other promoter elements, and tend to be more tightly regulated.

Common Promoters for Eukaryotes and Prokaryotes

See the two reference tables below listing some of the most common bacterial and mammalian promoters. These lists are by no means exhaustive, but should be a good place to start when trying to pick your perfect promoter!

Eukaryotic Promoters

Promoter	Primarily used for	RNA transcript	Description	Expression	Additional considerations
CMV	General expression	mRNA	Strong mammalian expression promoter from the human cytomegalovirus	Constitutive	May contain an enhancer region. Can be silenced in some cell types.
EF1a	General expression	mRNA	Strong mammalian expression from human elongation factor 1 alpha	Constitutive	Tends to give consistent expression regardless of cell type or physiology.
SV40	General expression	mRNA	Mammalian expression promoter from the simian vacuolating virus 40	Constitutive	May include an enhancer.
PGK1 (human or mouse)	General expression	mRNA	Mammalian promoter from phosphoglycerate kinase gene.	Constitutive	Widespread expression, but may vary by cell type. Tends to resist promoter down regulation due to methylation or deacetylation.
Ubc	General expression	mRNA	Mammalian promoter from the human ubiquitin C gene	Constitutive	As the name implies, this promoter is ubiquitous.

THE PROMOTER REGION – LET'S GO! (CONT.)

Eukaryotic Promoters

Promoter	Primarily used for	RNA transcript	Description	Expression	Additional considerations
human beta actin	General expression	mRNA	Mammalian promoter from beta actin gene	Constitutive	Ubiquitous. Chicken version is commonly used in promoter hybrids.
CAG	General expression	mRNA	Strong hybrid mammalian promoter	Constitutive	Contains CMV enhancer, chicken beta actin promoter, and rabbit beta-globin splice acceptor.
TRE	General expression	mRNA	Tetracycline response element promoter	Inducible with Tetracyline or its derivatives.	Typically contains a minimal promoter with low basal activity and several tetracycline operators. Transcription can be turned on or off depending on what tet transactivator is used.
UAS	General expression	mRNA	Drosophila promoter conaining Gal4 binding sites	Specific	Requires the presence of Gal4 gene to activate promoter.
Ac5	General expression	mRNA	Strong insect promoter from Drosophila Actin 5c gene	Constitutive	Commonly used in expression systems for Drosophila.
Polyhedrin	General expression	mRNA	Strong insect promoter from baculovirus	Constitutive	Commonly used in expression systems for insect cells.
CaMKIIa	Gene expression for optogenetics	mRNA	Ca2+/calmodulin-dependent protein kinase II promoter	Specific	Used for neuronal/CNS expression. Modulated by calcium and calmodulin.
GAL1, 10	General expression	mRNA	Yeast adjacent, divergently transcribed promoters	Inducible with galactose; repressible with glucose	Can be used independently or together. Regulated by GAL4 and GAL 80.
TEF1	General expression	mRNA	Yeast transcription elongation factor promoter	Constitutive	Analogous to mammalian EF1a promoter.

THE PROMOTER REGION – LET'S GO! (CONT.)

Eukaryotic Promoters

Promoter	Primarily used for	RNA transcript	Description	Expression	Additional considerations
GDS	General expression	mRNA	Strong yeast expression promoter from glyceraldehyde 3-phosphage dehydrogenase	Constitutive	Very strong, also called TDH3 or GAPDH.
ADH1	General expression	mRNA	Yeast promoter for alcohol dehydrogenase I	Repressed by ethanol	Full length version is strong with high expression. Truncated promoters are constitutive with lower expression.
CaMV35S	General expression	mRNA	Strong plant promoter from the Cauliflower Mosaic Virus	Constitutive	Active in dicots, less active in monocots, with some activity in animal cells.
Ubi	General expression	mRNA	Plant promoter from maize ubiquitin gene	Constitutive	Gives high expression in plants.
H1	small RNA expression	shRNA	From the human polymerase III RNA promoter	Constitutive	May have slightly lower expression than U6. May have better expression in neuronal cells.
U6	small RNA expression	shRNA	From the human U6 small nuclear promoter	Constitutive	Murine U6 is also used, but may be less efficient.

THE PROMOTER REGION – LET'S GO! (CONT.)

Prokaryotic Promoters

Promoter	Primarily used for	Description	Expression	Additional considerations
T7	in vitro transcription/ general expression	Promoter from T7 bacteriophage	Constitutive, but requires T7 RNA polymerase.	Can be used for in vitro transcription only if 2 different phage promoters are present in opposite orientations to gene.
T7lac	High levels of gene expression	Promoter from T7 bacteriophage plus lac operators	Negligible basal expression when not induced. Requires T7 RNA polymerase, which is also conrolled by lac operator. Can be induced by IPTG.	Commonly found in pET vectors. Very tightly regulated by the lac operators. Good for modulating gene expression through varied inducer concentrations.
Sp6	in vitro transcription/ general expression	Promoter from Sp6 bacteriophage	Constitutive, but requires SP6 RNA polymerase.	Can be used for in vitro transcription only if 2 different phage promoters are present in opposite orientations to gene.
araBAD	General expression	Promoter of the arabinose metabolic operon	Inducible by arabinose and repressed catabolite repression in the presence of glucose or by competitive binding of the anti-inducer fucose	Weaker. Commonly found in pBAD vectors. Good for rapid regulation and low basal expression; however, not well-suited for modulating gene expression through varied inducer concentrations.
trp	High levels of gene expression	Promoter from E. coli tryptophan operon	Repressible	Gets turned off with high levels of cellular tryptophan.
lac	General expression	Promoter from lac operon	Constitutive in the absense of lac repressor (lacI or lacIq). Can be induced by IPTG or lactose.	Leaky promoter with somewhat weak expression. lacIq mutation increases expression of the repressor 10x, thus tightening regulation of lac promoter. Good for modulating gene expression through varied inducer concentrations.
Ptac	General expression	Hybrid promoter of lac and trp	Regulated like the lac promoter	Contains -35 region from trpB and -10 region from lac. Very tight regulation. Good for modulating gene expression through varied inducer concentrations. Generally better expression than lac alone.
pL	High levels of gene expression	Promoter from bacteriophage lambda	Can be temperature regulatable	Often paired with the temperature sensitive cI857 repressor.

CHAPTER 2: EUKARYOTIC EXPRESSION VECTORS

MAMMALIAN VECTORS

By Marcy Patrick, Addgene | Mar 25, 2014

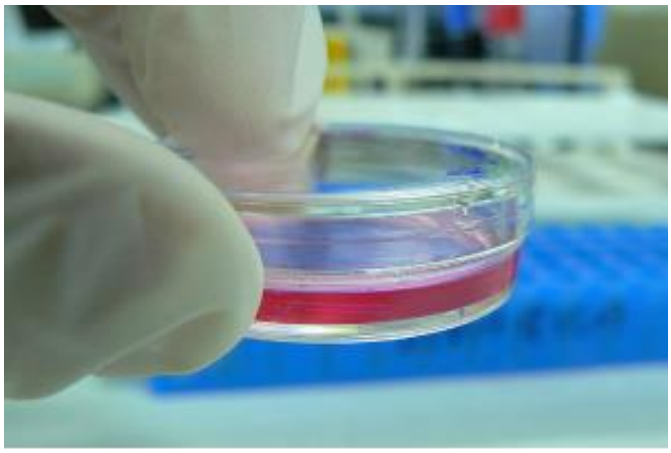


Image Source: Wikipedia, accessed March 14, 2014. Photo Author: kaibara87.

Although plasmids do not naturally exist in mammals, scientists can still reap the benefits of plasmid-based research using synthetic vectors and cultured mammalian cells. Of course, these mammalian vectors must be compatible with the cell type they are transfected into – a bacterial origin of replication (ORI) will not allow for plasmid replication in mammalian cells, for example, and a toxin that kills bacteria may not have any discernable effect on mammalian cells. So how do mammalian plasmids differ from their bacterial counterparts? How does replication occur? Is selection necessary for transfected cells?

What is Transfection?

Before getting into the mammalian plasmid components, it may be useful to describe the means of introducing genetic material (such as plasmids) into mammalian cells, a process called transfection. Transfection is somewhat comparable to bacterial transformation (the introduction of DNA into bacterial cells); however, the techniques and reagents vary. Plasmid transfection into mammalian cells is fairly straightforward and the resultant cells can either express the plasmid DNA transiently (similar to bacteria) or incorporate the genetic material directly into the genome to form a stable transfection. Unlike bacterial transformation, scientists do not "select" for cells that have taken up the plasmid in the same way. Selection methods, described below, are typically employed when creating stable cell lines and are not used for general plasmid selection. Instead, reporter genes are often employed to easily monitor transfection efficiencies and expression levels in the cells. Ideally, the chosen reporter is unique to the cell, is expressed from the plasmid, and can be assayed conveniently. A direct test for your gene of interest may be another method to assess transfection success. GFP is often used as a reporter. Click here to skip ahead in this eBook to fluorescent tagging.

MAMMALIAN VECTORS (CONT.)

Transient Transfection and the Elusive "Mammalian ORI"

For many experiments, it is sufficient for the transfected plasmid to be expressed transiently. Since the DNA introduced in the transfection process is not integrated into the nuclear genome, in the absence of plasmid replication, the foreign DNA will be degraded or diluted over time. This, however, may not be a problem depending on the duration or other parameters of your experiment. Mammalian cells double at a much slower rate than that of bacteria (~24 h vs 20 min, respectively). Therefore, it is not always mission critical to make sure the plasmid replicates in the cell, as many of these experiments are concluded within 48 h of transfection.

Of course, it is possible that you may not want the plasmid depleted, but still want to use transient transfection methods. Since there are no "natural" mammalian ORIs, scientists have usurped viral-based ORIs to fill the void. These ORIs, however, require additional components expressed in trans within the cell for effective replication. Cell lines expressing the Epstein–Barr virus (EBV) nuclear antigen 1 (EBNA1) or the SV40 large-T antigen (293E or 293T cells), allow for episomal amplification of plasmids containing the viral EBV or SV40 ORIs, respectively. The presence of these viral components greatly reduces the rate of plasmid dilution but does not guarantee 100% transfection efficiency.

Stable Transfection

A stable transfection is used to create a population of cells that have fully and successfully incorporated foreign genetic material into their genomes. Unlike plasmids used for expression in yeast and bacteria, plasmids used for stable transfections rarely contain an ORI since the integrated DNA will be replicated as part of the genome. Because the foreign DNA becomes a permanent addition to the host genome, the cells will continually express the genetic traits of the foreign material and will subsequently pass it on to future generations. Stably transfected cells may be considered an entirely new cell line from that of the original parental cells.

Positive Selection in Mammalian Cells

To achieve stable transfection, there should be a selective pressure to force cells to incorporate the plasmid DNA into the genome. For the purposes of this post, we will define positive selection as the means of picking up positive traits (i.e. the plasmid contains a cassette that will make cells resistant to a toxin), whereas negative selection would be the picking up of a negative trait (i.e. the plasmid contains a cassette that will make cells sensitive to a toxin). In the table below we focus on positive selection; however, negative selection techniques can be used in conjunction with positive selection to ensure your gene gets targeted to a specific location within the genome.

MAMMALIAN VECTORS (CONT.)

Positive selection in mammalian cells works similarly to that in bacteria and a table of the most commonly used selection markers are listed below:

Name	Gene Conferring Resistance	Cell Types*	Mode of Action**	Working Concentration***
Blasticidin	*bsd*	HeLa, NIH3T3, CHO, COS-1, 293HEK	Inhibits termination step of translation	2-10 ug/mL
G418/Geneticin	*neo*	HeLa, NIH3T3, CHO, 293HEK, Jurkat T cells	Blocks polypeptide synthesis at 80S; inhibits chain elongation	100-800 ug/mL
Hygromycin B	*hygB*	HeLa, NIH3T3, CHO, Jurkat T cells	Blocks polypeptide synthesis at 80S; inhibits chain elongation.	50-500 ug/mL
Puromycin	*pac*	HeLa, 293HEK, Jurkat T cells	Inhibits protein synthesis; premature chain termination	1-10 ug/mL
Zeocin	*Sh bla*	HeLa, NIH3T3, CHO, COS-1, 293HEK, Jurkat T cells	Complexes with DNA; causes strand scissions	100-400 ug/mL

*Not comprehensive.

** In eukaryotes.

***The concentration used for selection is typically more (double) than that used for maintenance of a transfected cell line.

Keep These Tips in Mind:

- There is not one recommended concentration for selection in mammalian cells. Before doing a transfection experiment, it is important to determine the proper concentration required for efficient selection. This is usually achieved by performing a "kill curve" (basically growing cells in various concentrations of the selection reagent). Cells should die within 3-5 days and resistant colonies appear in about 10-14 days depending on how quickly your cells divide.
- Gentamicin is often used as a supplement in mammalian cell culture to suppress bacterial growth, and is not appropriate for mammalian selection – do not confuse this with G418 (aka Geneticin).
- Neomycin should not be used for mammalian expression – instead use G418. This can be confusing since the neo/kan gene confers G418 resistance; however, like gentamicin, neomycin is typically used to suppress bacterial growth.

MAMMALIAN VECTORS (CONT.)

- The selection agent should not be added to culture media until 24-48 h post transfection when creating stable cell lines.

Looking for Mammalian Expression Plasmids? Check Out Addgene's Empty Backbone Resource:

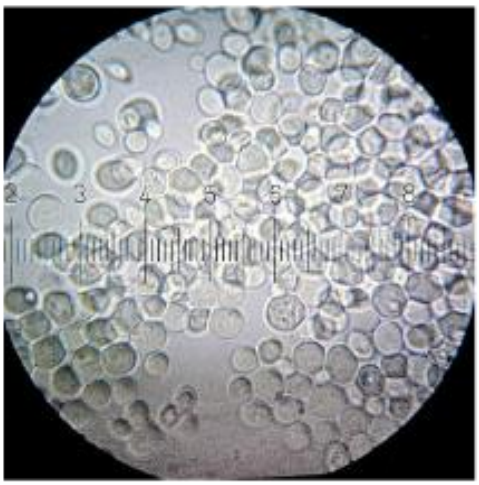
- Choosing Your Perfect Plasmid Backbone

YEAST VECTORS

By Marcy Patrick, Addgene | Feb 25, 2014

Why Do Scientists Use Yeast Vectors?

Yeast are eukaryotes and thus contain complex internal cell structures similar to those of plants and animals. Unlike bacteria, yeast can post-translationally modify proteins yet they still share many of the same technical advantages that come with working with prokaryotes. This includes but is not limited to: rapid growth, ease of replica plating and mutant isolation, a well-defined genetic system, and a highly versatile DNA transformation system. Unlike most other microorganisms, yeast have both a stable haploid and diploid state which is useful for genetic analysis, as well as an efficient mechanism of homologous recombination to facilitate simple gene replacement/mutation. Yeast expression plasmids used in the lab typically contain all the necessary components to allow shuttling between *E. coli* and yeast cells. To be useful in the lab, the vectors must contain a yeast-specific origin of replication (ORI) and a means of selection in yeast cells, in addition to the bacterial ORI and antibiotic selection markers.



Source: Wikipedia, accessed 2/23/2014.
Author: Bob Blaylock

The Yeast Origin of Replication

(**Please note:** This first section primarily pertains to ORIs in budding yeast, *Saccharomyces cerevisiae*; however, we've also noted some features required for the replication of fission yeast, *Schizosaccharomyces pombe,* vectors at the end.)

We have already covered how the regulation of bacterial ORIs determines plasmid copy number within the bacterial cell. Similarly, the specific ORI elements included within a yeast vector determine how the plasmid is replicated and maintained within the yeast cell. These elements control not only the number of plasmids found in each cell, but also whether the plasmid gets integrated into the host DNA or is independently replicated as an episome.

YEAST VECTORS (CONT.)

The four main types of yeast plasmids are defined below:

- ***Yeast Integrating plasmids (YIp)***: These plasmids lack an ORI and must be integrated directly into the host chromosome via homologous recombination.
- ***Yeast Replicating plasmids (YRp)***: These vectors contain an Autonomously Replicating Sequence (ARS) derived from the yeast chromosome. As the name suggests, these vectors can replicate independently of the yeast chromosome; however, they tend to be unstable and may be lost during budding.
- ***Yeast Centromere plasmids (YCp)***: These are considered low copy vectors and incorporate part of an ARS along with part of a centromere sequence (CEN). These vectors replicate as though they are small independent chromosomes and are thus typically found as a single copy. Unlike the ARS vectors, CEN vectors are stable without integration.
- ***Yeast Episomal plasmids (YEp)***: These are most similar to bacterial plasmids and are considered "high copy". A fragment from the 2 micron circle (a natural yeast plasmid) allows for 50+ copies to stably propogate per cell. The copy number of these vectors can also be controlled if specific regulatable elements are included (reviewed here).

Plasmids for use in *S. pombe*, on the other hand, do not require a well-defined ORI. Instead, the size and A-T content of the DNA (apparently independent of a known specific sequence) dictate the replication of these vectors. *S. pombe* plasmids oftentimes utilize an ARS to aid in high transformation efficiency; however, this region does not necessarily promote replication.

Selection Markers for Yeast

Historically, scientists have utilized auxotrophic selection rather than antibiotic selection when working with yeast, due to high rates of spontaneously occurring resistant mutants and the insensitivity of yeast strains to some antibiotics. Auxotrophy is defined as the inability of an organism to synthesize a particular organic compound required for its growth. Many auxotrophic strains of yeast exist which can be easily maintained when grown on media containing the missing nutrient. Scientists can exploit these host mutations by including a copy of a functional gene which complements the host's auxotrophy. When grown on media NOT containing the nutrient, the host cells will die unless they have incorporated the plasmid carrying the required gene.

YEAST VECTORS (CONT.)

The table below lists some of the most commonly used selection markers in yeast and provides the element needed to overcome the auxotrophy as well as additional uses for said element. This link provides a more extensive list of yeast auxotrophic markers and includes the associated references.

Yeast selection marker	What does the marker help synthesize?	Counterselection? (growth-based positive selection for the loss of the marker gene)	For use in:	Can this be used in other yeast species?	Can this be used for auxotrophic selection in E. coli?	Additional considerations
HIS3	L-hisitidine	no	*S. cerevisiae*	no	yes	
URA3	pyrimidine (uracil)	yes - Grow with 5-FOA.	*S. cerevisiae*	yes - This can complement *ura4* - *S. pombe*, but the complementation is weak.	yes	
LYS2	L-lysine	yes - Grow on plates containing alpha-aminoadipate in the absence of a nitrogen source.	*S. cerevisiae*	no	no	
LEU2	L-leucine	no	*S. cerevisiae*	yes - This can complement leu1- *S. pombe*, but mulitple copies are required.	yes	
TRP1	L-tryptophan	yes - Grow with 5-FAA.	*S. cerevisiae*	no	no	*TRP1* alters some yeast phenotypes. This marker should not be used in gene disruption experiments.
MET15	L-methionine and overproduces hydrosulfide ions	yes - Grow with methylmercury.	*S. cerevisiae*	no	no	Can be used for color and growth selection if divalent lead ions are used in the growth media.
ura4+	pyrimidine (uracil)	yes - Grow with 5-FOA*.	*S. pombe*	no	no	*FOA in fission yeast induces mutation in the *ura5+* gene in addition to *ura4+*.
leu1+	L-leucine		*S. pombe*	no	no	
ade6+	purine (adenine)		*S. pombe*	no	no	

YEAST VECTORS (CONT.)

Considerations When Using Auxotrophic Selection

Of course, there are some drawbacks to using auxotrophic markers as a means of selection:

1. A specific selection marker needs to be used with a yeast strain deficient in that compound. Therefore known auxotrophic strain/ selection element pairs must be utilized or a new combination needs to be created in advance of the experiment.

2. The marker provided by the plasmid may be expressed at higher than normal physiological levels due to high copy numbers. This creates a potential metabolic burden on the yeast cells.

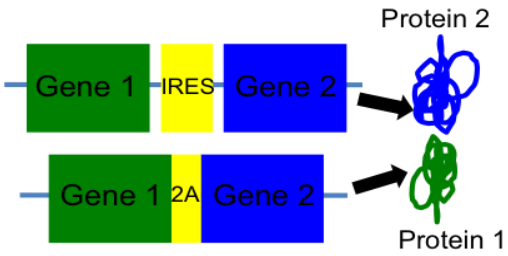
3. Some phenotypes may be altered due to the presence of the selection marker at non-physiological levels.

Scientists have tried varied approaches to combat these issues. One method to reduce the amount of marker gene expression is to use a partially defective promoter to drive expression of the selection marker. This reduces the amount of gene product present in the cell, thus allowing the yeast to maintain higher copy numbers. Additionally, improvements in antibiotic selection have made utilizing the more traditional drug selection methods feasible in yeast as a complement or alternative to using auxotrophic markers. The MX series of antibiotic resistance cassettes is most commonly employed, with the KanMX being the most prevalent due to its versatility and ability to be used in bacterial, yeast, and mammalian cells.

MULTICISTRONIC VECTORS

By Melina Fan, Addgene with contributions from Marcy Patrick, Addgene | Sep 8, 2014

Co-expression of multiple genes is valuable in many experimental settings. To achieve this, scientists use a multitude of techniques including co-transfection of two or more plasmids, the use of multiple or bidirectional promoters, or the creation of bicistronic or multicistronic vectors. Unlike promoters which will create unique mRNA transcripts for each gene that is expressed, multicistronic vectors simultaneously express two or more separate proteins from the same mRNA. We've discussed promoters before so in this blog post we'll cover basics of multicistronic vectors: why they are useful, how they work, and how to get started with them.



Why Use Multicistronic Vectors?

Detecting cells that are expressing your gene, especially if you are studying a novel gene, is not always a straightforward process. Rather than try to directly detect your gene of interest, scientists have instead developed novel methods to co-express your gene along with a reporter, such as a fluorescence gene or a resistance gene. These reporters allow you to easily screen or select for cells that are expressing your gene of interest at high levels. Unlike vectors that express screenable or selectable markers from a unique promoter, multicistronic plasmids ensure that any cells that are positive for your marker should also be expressing your gene as they are both derived from the same transcript.

Of course multicistronic vectors do not have to exclusively be used as a means of detection; they are useful almost anytime you want to express multiple genes in the same cell. Although it is possible to drive co-expression by using a plasmid with multiple, individual expression cassettes, having the genes expressed from the same cassette is sometimes advantageous, particularly when only a portion of the plasmid is packaged for viral delivery, or the relative expression levels between two or more genes is important.

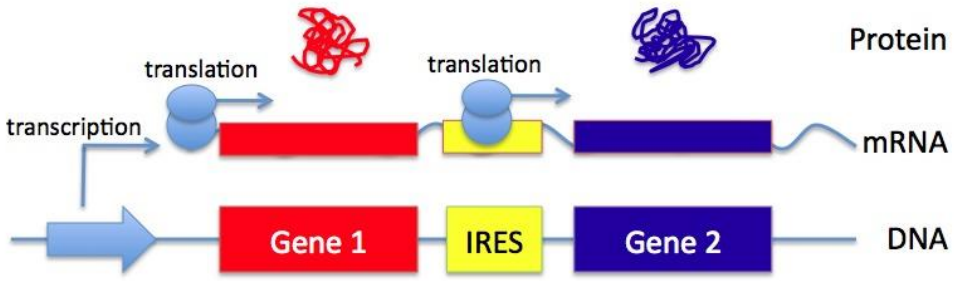
How Do Multicistronic Vectors Work?

Scientists have "borrowed" some tricks discovered in positive single-stranded RNA viruses to allow for the efficient translation of multiple genes from a single transcript. The two strategies most widely incorporated into plasmids for research purposes are described below.

MULTICISTRONIC VECTORS (CONT.)

IRES Elements:

Translation in eukaryotes usually begins at the 5' cap so that only a single translation event occurs for each mRNA. However, some bicistronic vectors take advantage of an element called an Internal Ribosome Entry Site (IRES) to allow for initiation of translation from an internal region of the mRNA.



In the figure above, you can see that the IRES element acts as another ribosome recruitment site, thereby resulting in co-expression of two proteins from a single mRNA.

IRES was originally discovered in poliovirus RNA, where it promotes translation of the viral genome in eukaryotic cells.[1,2] Since then, a variety of IRES sequences have been discovered - many from viruses, but also some from cellular mRNAs. What they all have in common is the ability to spark translation initiation independent of the 5' cap.

IRES elements are very useful and commonly found in bicistronic vectors; however, they do have some disadvantages. These elements are quite large (500-600 bp) and may take up precious space in viral transfer vectors with limited packaging capacity. Additionally, it may not be feasible to express more than two genes at a time using IRES elements. Further, scientists have reported lower expression of the downstream cistron due to factors such as the experimental cell type and the specific genes cloned into the vector.[3]

2A Peptides:

To overcome some of the disadvantages of the IRES element, scientists have adapted "self-cleaving" 2A peptides into their multicistronic vectors. These peptides, first discovered in picornaviruses, are short (about 20 amino acids) and produce equimolar levels of mulitple genes from the same mRNA. The term "self-cleaving" is not entirely accurate, as these peptides are thought to function by making the ribosome skip the synthesis of a peptide bond at the C-terminus of a 2A element, leading to separation between the end of the 2A sequence and the next peptide downstream.[4] The "cleavage" occurs between the Glycine and Proline

MULTICISTRONIC VECTORS (CONT.)

residues found on the C-terminus meaning the upstream cistron will have a few additional residues added to the end, while the downstream cistron will start with the Proline.

The table below lists the four common 2A peptides employed by scientists. 2A cleavage is universal in eukaryotic cells, and, although some scientists report close to 100% cleavage with some of these peptides, no consensus has been reached on which peptide works best. Likely the choice of specific 2A peptide will ultimately depend on a number of factors such as cell type or experimental conditions.

Peptide	Amino acid sequence*
T2A:	(GSG) E G R G S L L T C G D V E E N P G P
P2A:	(GSG) A T N F S L L K Q A G D V E E N P G P
E2A:	(GSG) Q C T N Y A L L K L A G D V E S N P G P
F2A:	(GSG) V K Q T L N F D L L K L A G D V E S N P G P

***** (GSG) residues can be added to the 5' end of the peptide to improve cleavage efficiency.

How Do I Get Started?

If you are looking to co-express your gene of interest along with a fluorescent protein or selectable marker, it is easiest to start with a plasmid that already has the multicistronic element and reporter cloned in. In these plasmids you would simply clone your gene of interest into the multiple cloning site up or down stream of the IRES or 2A element (depending on the placement of the reporter gene).

MULTICISTRONIC VECTORS (CONT.)

Addgene's collection offers a variety of plasmids to express two or more genes, some of which are listed below. We should note that these vectors are primarily designed for bicistronic expression; however, many could be easily manipulated to express more than two genes.

Plasmid Name	Multicistronic Element	Expression Type
MSCV-IRES-EGFP	IRES	Retroviral
pMSCV-pBabeMCS-IRES-RFP	IRES	Retroviral
pMSCV-IRES-YFP II	IRES	Retroviral
pCMMP-MCS-IRES-Puro	IRES	Retroviral
pEF1a-IRES-Neo	IRES	Mammalian
MSCV-IRES-Luciferase	IRES	Retroviral
pWPI	IRES	Lentiviral
AmCyan-P2A-mCherry	P2A	Mammalian
pC5Kan-P2A	P2A	Insect
pUltra	P2A and T2A	Lentiviral
Ac5-STABLE2-neo	T2A	Insect

For the co-expression of multiple unique genes, you can start with a plasmid that has multiple cloning sites flanking the multicistronic element(s), or you could replace one of the reporter genes above with your gene or genes of interest. Some of the plasmids listed in the table above (and their related plasmids) are designed to have one or more of the genes replaced.

Additionally, 2A peptides could be PCR-cloned between your genes of interest and you can then insert the whole multicistronic cassette into a backbone as a single unit. Although it is recommended to use the 2A peptides instead of an IRES when stoichiometrically equivalent levels of expression are required, we should also note that IRES and 2A peptides are not mutually exclusive elements. Labs have successfully utilized 2A and IRES elements within the same multicistronic vector, effectively making a construct that expresses multiple unique genes at equivalent levels upstream of an IRES fluorescent reporter for easy detection.[5]

MULTICISTRONIC VECTORS (CONT.)

References:

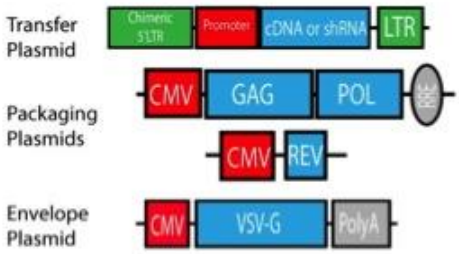
1. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. Pelletier et al (Nature. 1988 Jul 28;334(6180):320-5.) PubMed.
2. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. Jang et al. (J Virol. 1988 Aug;62(8):2636-43.) PubMed.
3. Highly Efficient Multicistronic Lentiviral Vectorswith Peptide 2A Sequences. Ibrahimi et al. (Hum Gene Ther. 2009 Aug;20(8):845-60. doi: 10.1089/hum.2008.188.) PubMed.
4. High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. Kim et al (PLoS One. 2011;6(4):e18556. doi: 10.1371/journal.pone.0018556. Epub 2011 Apr 29.) PubMed.
5. Scalable signaling mediated by T cell antigen receptor-CD3 ITAMs ensures effective negative selection and prevents autoimmunity. Holst et al (Nat Immunol. 2008 Jun;9(6):658-66. doi: 10.1038/ni.1611. Epub 2008 May 11.) PubMed.
6. *Read more Plasmids 101 posts or browse Addgene's Plasmid Guide for more molecular biology and cloning information.*

CHAPTER 3: VIRAL EXPRESSION VECTORS

VIRAL VECTOR ELEMENTS

By Marcy Patrick, Addgene | Jul 17, 2014

The use of viral vectors in research is beneficial for a number of reasons, including but not limited to: helping to get difficult-to-deliver DNA into mammalian cells, increasing the efficiency of gene transduction, allowing for control over which cells are infected through viral pseudotyping, and ease of vector cloning and modification. At the most basic level, viral vectors consist of a viral genome that has been adapted into a plasmid-based technology and modified for safety through the removal of many essential genes and the separation of the viral components.



Of the many viruses out there, only a subset are commonly used in the lab: gamma-retrovirus, lentivirus, and adeno-associated virus.

Adeno-associated Virus (AAV)

Adeno-associated virus or AAV is a small parvovirus that infects humans and some other primate species, but is not currently known to cause disease. If AAV is exclusively used to infect human cells, it stably integrates at a specific chromosomal site as part of its lysogenic life cycle; however, the addition of a helper virus such as adenovirus or a helper plasmid containing the specific viral proteins E1A, E1B, E2A, E4, and VA will cause AAV to enter a lytic cycle and replicate. There are at least eleven different serotypes of AAV, with likely more to be discovered. The most studied and experimentally used serotype is AAV2; however, many factors including species and cell type may make other serotypes more desirable.

Gamma-retrovirus and Lentivirus

A retrovirus is an RNA virus that uses reverse transcriptase to make a DNA provirus which can be incorporated into the host's genome. There are two common genera of retrovirus used by scientists: Gamma-retrovirus (many times shortened to just "retrovirus") and lentivirus. From these two genera, the most studied types are Murine Leukemia virus (MLV) and human immunodeficiency virus-1 (HIV-1), respectively. Since these viruses are closely related, their life cycle and the required components are basically the same, although differences do exist at the sequence level. This means that plasmids containing virus-specific elements such as the LTRs or structural proteins are not interchangeable; however, other, more general, viral components such as a heterologous envelope or a post-transcriptional regulatory element can be used across either system. The most notable difference between gamma-retrovirus and lentivirus is the fact that lentivirus can infect both dividing and non-dividing cells, whereas gamma-retrovirus is restricted to dividing cells only.

VIRAL VECTOR ELEMENTS (CONT.)

Common Viral Vector Elements

When developing viral vectors, scientists strive for a number of features: low risk, high expression, large payload capacity, ability to infect target cells, no immune response from the host, and easy to develop/use in the lab. In virtually all viral expression systems employed by scientists, non-essential components are stripped away and the remaining native genes are spread over multiple plasmids to ensure safety. The choice for how the elements are divided up is dependent on whether the component needs to be provided *in cis* (on the same plasmid) or *in trans* (on a separate plasmid) as your insert. The tables below list some components most commonly found in viral vector systems.

AAV Elements

Element	Plasmid type	Provided?	Purpose
ITR	Cloning	*in cis*	Inverted terminal repeat; 145 bases each. Symmetry of ITRs is required for efficient multiplication of the AAV genome. Forms a T-shaped hairpin that serves as the origin of viral DNA replication. Contains D region required for packaging. Cloning capacity between the ITRs is ~4kb.
Rep	Packaging	*in trans*	Packaging proteins with four possible variants: Rep78, Rep68, Rep52, and Rep40; Required for genome replication and necessary for integration. Rep proteins from most serotypes can be interchangeably used with any ITR serotype*
Cap	Packaging	*in trans*	Structural capsid proteins with three variants: VP1, VP2, and VP3; VP1 possesses phospholipase A2 activity, which is likely necessary to release the AAV particles from late endosomes. VP2 and VP3 are crucial for correct virion assembly. Determines the serotype/ viral tropism.

**Exception is AAV5, which requires AAV5 Rep and AAV5 ITRs for packaging.*

VIRAL VECTOR ELEMENTS (CONT.)

Gamma-retroviral and Lentiviral Elements

Element	Plasmid type	Provided?	Purpose
LTR	Transfer	*in cis*	Long terminal repeats; U3-R-U5 regions found on either side of a retroviral provirus (see below). Cloning capacity between the LTRs is ~8.5kb, but inserts bigger than ~3kb are packaged less efficiently.
U3 (subelement of LTR)	Transfer	*in cis*	Unique 3'; region at the 3' end of viral genomic RNA (but found at both the 5' and 3' ends of the provirus). Contains sequences necessary for activation of viral genomic RNA transcription.
R (subelement of LTR)	Transfer	*in cis*	Repeat region found within both the 5'and 3' LTRs of retro/lentiviral vectors. Tat binds to this region.
U5 (subelement of LTR)	Transfer	*in cis*	Unique 5'; region at the 5' end of the viral genomic RNA (but found at both the 5' and 3' ends of the provirus).
5' LTR	Transfer	*in cis*	Acts as an RNA pol II promoter. The transcript begins, by definition, at the beginning of R, is capped, and proceeds through U5 and the rest of the provirus. Third generation vectors use a hybrid 5'LTR with a constitutive promoter such as CMV or RSV.
TAR	Transfer (2nd generation only)	*in cis*	Trans-activating response element; located in the R region of the LTR and acts as a binding site for Tat.
3' LTR	Transfer	*in cis*	Terminates transcription started by 5' LTR by the addition of a poly A tract just after the R sequence.
cPPT	Transfer	*in cis*	Central polypurine tract; recognition site for proviral DNA synthesis. Increases transduction efficiency and transgene expression.
Psi (Ψ)	Transfer	*in cis*	RNA target site for packaging by Nucleocapsid.
RRE	Transfer	*in cis*	Rev Response Element; sequence to which the Rev protein binds.
WPRE	Transfer	*in cis*	Woodchuck hepatitis virus post-transcriptional regulatory element; sequence that stimulates the expression of transgenes via increased nuclear export.
Gag	Packaging	*in trans*	Precursor structural protein of the lentiviral particle containing Matrix, Capsid, and Nucleocapsid components.
Pol	Packaging	*in trans*	Precursor protein containing Reverse Transcriptase and Integrase components.
Rev	Packaging (on separate plasmid from Gag/Pol in third generation systems)	*in trans*	Binds to the Rev Response Element (RRE) within unspliced and partially spliced transcripts to facilitate nuclear export.
Tat	Packaging (second generation only)	*in trans*	Trans-activator; binds TAR to activate transcription from the LTR promoter.
VSVG	Envelope	*in trans*	Vesicular somatitis virus G glycoprotein; Broad tropism envelope protein used to psuedotype most lentiviral vectors.

VIRAL VECTOR ELEMENTS (CONT.)

Looking for Viral Vectors? Check Out Viral Vectors at Addgene:

- Addgene's most popular lentiviral plasmids
- Addgene's most popular retroviral plasmids
- Addgene's collection of AAV plasmids

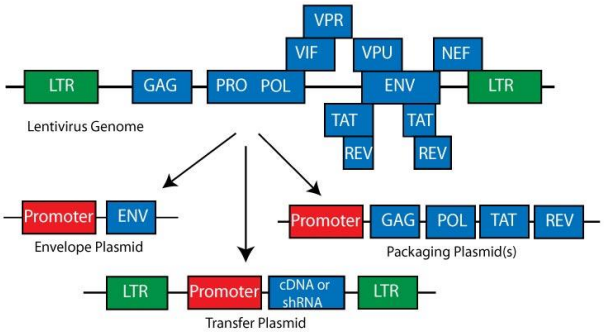
Learn More:

- Overview of Lentiviral Packaging Plasmids
- Lentiviral Protocols and Resources
- More Lentiviral FAQs Answered by Addgene
- Didier Trono Lab: Lentivectors Toolbox

LENTIVIRAL PLASMID FAQS ANSWERED

By Kendall Morgan | April 23, 2014

Lentiviruses are a useful and efficient tool to introduce your gene of interest into cells. Unlike gamma-retroviruses that can only infect dividing cells, lentiviruses can infect dividing and non-dividing cells.



Addgene has an extensive collection of lentiviral plasmids created for a variety of applications including cDNA expression, shRNA-mediated knockdown, Tet and Cre-regulated expression, CRISPR genome editing, and more. Not surprisingly, Addgene receives many questions from scientists all over the world looking for some additional information or clarification on lentiviral vectors.

Frequently Asked Questions

How are lentiviral particles produced?

This useful resource from Addgene Depositor (and lentiviral expert) Didier Trono's lab webpage explains how lentiviral particles are produced by co-transfecting three essential ingredients: the lentiviral packaging vector, a transfer vector including the gene of interest, and the plasmid coding for an envelope. Three generations of lentiviral packaging systems have been developed over the years based on HIV-1; however, the 1st-generation vectors are never used as they pose too much of a biosafety risk to scientists. Second-generation vectors are likely sufficient for most experiments; however, the 3rd-generation packaging systems offer the maximal biosafety available for the technology. The minor limitation in using the 3rd-generation system is that it involves transfecting four different plasmids into the producer cells (two packaging plasmids, an envelope plasmid, and the lentiviral transfer vector), rather than the three used for 2nd-generation systems.

Are retroviruses and lentiviruses the same thing? Can I package lentivirus with retroviral packaging plasmids (and vice versa)?

Lentiviruses are part of the retroviral family. What scientists often refer to as "retroviruses" are technically gamma-retroviruses - another, separate member of the retroviral family. While both lentiviruses and gamma-retroviruses use the same genes for packaging (that's gag, pol, and env), the isoforms of these proteins, as well as the viral long terminal repeats (LTRs) are different. As a result, lentiviral and retroviral packaging vectors are not interchangeable.

LENTIVIRAL PLASMID FAQS ANSWERED (CONT.)

Can I use lentiviral transfer vectors for transient transfections?

It is not recommended to use viral plasmids transiently. This is especially true if the transfer vector does not contain an independent promoter and instead exclusively uses the viral LTR promoter to drive expression, since the LTR is not strong enough to drive expression of the insert transiently. Virus must be produced and your target cells infected in order to see adequate expression.

2nd vs. 3rd Generation Lentiviral Plasmids

How can you tell 2nd- and 3rd- generation transfer vectors apart?

Addgene defines 2nd- or 3rd-generation transfer vectors based on whether or not they have a chimeric 5'LTR. Second-generation transfer plasmids have a wildtype 5'LTR, which requires the presense of the Tat protein to work, while 3rd-generation transfer plasmids have a chimeric 5'LTR that includes a CMV or RSV promoter as well as a portion of the 5'LTR. Including a chimeric 5'LTR removes the requirement for the HIV Tat protein, thus decreasing the probability of creating replication-competent lentivirus in your target cells.

How do the packaging systems differ between 2nd- and 3rd- generation vectors? What does this mean for transfer vectors?

The lentiviral packaging systems differ in two ways: 1) Second-generation systems consist of two plasmids while 3rd-generation systems require three plasmids. 2) Second-generation systems express HIV Tat; 3rd-generation systems do not.

Second-generation transfer plasmids MUST be packaged with a 2nd-generation system because the wildtype 5'LTR promoter requires Tat to function. Third-generation plasmids can be packaged with either system. Please note that the generation used to package the virus does not change the generation of the transfer vector.

What envelope glycoprotein will be expressed on virus particles?

This depends on the envelope plasmid you use; however, the choice of envelope dictates the tropism of the virus. The VSV-G is very common due to its broad host range, but viruses can be easily pseudotyped by utilizing an alternative envelope plasmid as described here.

Are transfer plasmids replication competent?

Most (if not all) Addgene transfer plasmids are replication-deficient meaning they can be used to create virus that is capable of infecting target cells, but cannot produce any new viral particles after the initial infection.

LENTIVIRAL PLASMID FAQS ANSWERED (CONT.)

What is SIN?

SIN is shorthand for self-inactivating, which is achieved by deleting a large portion of the 3'LTR. This results in short, inactivated 5' and 3' LTRs after integration and further reduces the possibility of creating replication-competent virus. Both 2nd- and 3rd-generation vectors can be SIN.

Which genes are deleted or modified to result in replication deficiency? What percent of the vector components are HIV-based?

No single plasmid contains all the components necessary to produce viral particles. The components are divided as follows:

- Transfer vectors contain minimal cis-acting HIV components: LTRs, PPT, RRE, and psi packaging signal. Viral components typically total <1.5kb, which is rarely more than 30% of the transfer plasmid. Many transfer vectors are SIN.
- Packaging plasmid(s) contain the minimal number of HIV genes required for virus production (3 or 4). Third-generation vectors contain gag, pol, and rev. Second-generation contains those three genes plus tat.
- The envelope plasmid provides a heterologous envelope for pseudotyping and is not HIV-derived.

LENTIVIRAL PLASMID FAQS ANSWERED (CONT.)

Concerning Safety of Lentiviral Transduction

You mentioned biosafety. What are the major risks when it comes to lentiviruses in terms of biosafety?

According to the American Biological Safety Association, "the two major risks of lentiviral vectors are: 1) the potential generation of replication competent virus [usually HIV-1]; and 2) the potential for oncogenesis through insertional mutagenesis. These risks are largely based upon the vector system used and the transgene insert encoded by the vector."

Biosafety should always be considered with respect to the specific experiments being performed and you should understand the guidelines for using such reagents as outlined by your institution or country.

The table below breaks down the biosafety concerns associated with various aspects of lentiviral transduction along with suggestions on how to lower the risk.

Biosafety Concern	High Risk	Lower Risk
Vectors	Two plasmids (1st gen) used to produce virus; expression of viral genes	Three (2nd gen) or four (3rd gen) plasmids used to produce virus; removal of viral genes
Insert/transgene	Oncogenic, apoptotic, or toxic	Non-oncogenic, nonapoptotic, non-toxic
Envelope/host	Permissive/amphotrophic	Ecotropic/non-permissive
Propagation	Large scale	Lab scale (less than 100mL)

CHAPTER 4: PLASMIDS THAT GLOW

HOW DO FLUORESCENT PROTEINS GLOW?

By Gal Haimovich of greenfluorescentblog | May 20, 2014

Fluorescent Proteins (FPs) were first discovered over 50 years ago, with the discovery of the Green Fluorescent Protein (GFP), a protein from the jellyfish Aequorea Victoria. Since that discovery, the family of FPs just keeps getting larger with hundreds of variants available. Additionally, many other FPs have been found in other organisms.

Source Wikipedia, Accessed 5/14/2014
Author: www.glofish.com

Fluorescence is the emission of light by a substance that has absorbed light. The emitted light is at a longer wavelength than the exciting wavelength. Thus, FPs are proteins with this unique capacity.

Many of these FPs are fluorescent when ectopically expressed in most organisms. Furthermore, fusing FPs to another protein usually does not affect its fluorescence. Therefore, FPs are used to study many biological questions.

The two most common uses for fluorescent proteins are: 1) to test the expression level in a specific system (by measuring the fluorescence intensity); and 2) to visualize the localization of the FP (fused to the protein of interest), thus tracking the localization of that biomolecule inside living cells.

GREEN FLUORESCENT PROTEIN

By Marcy Patrick, Addgene with contributions from A. Max Juchheim, Addgene | May 15, 2014

Bioluminescence and fluorescence from proteins such as Green Fluorescent Protein (GFP) has likely existed in creatures such as jellyfish for millions of years; however, it took until the 1960s for scientists to begin to study GFP and deduce its biochemical properties. Now GFP and its fluorescent derivatives are a staple in the lab. GFP is used in research across a vast array of biological disciplines and scientists employ GFP for a wide number of functions, including: tagging genes for elucidating their expression or localization profiles, acting as a biosensor or cell marker, studying protein-protein interactions, visualizing promoter activity, and much more.

GREEN FLUORESCENT PROTEIN (CONT.)

Why Green Fluorescent Protein?

GFP is a ~27 kDa protein consisting of 238 amino acids derived from the crystal jellyfish *Aequorea victoria*. It has a fluorescent emission wavelength in the green portion of the visible spectrum (hence the name), which is due to a chromophore formed from a maturation reaction of three specific amino acids at the center of the protein (Ser65, Tyr66, and Gly67). When first discovered, one of the most surprising aspects of GFP was the fact that the chromophore forms spontaneously and without additional co-factors, substrates, or enzymatic activity – it only requires the presence of oxygen during maturation. This meant that the protein could be taken directly from *A. Victoria* and expressed in any organism while still maintaining fluorescence.

Image source: Wikimedia Commons, accessed: 5/15/2014. Photo author: Ingrid Moen, et al.

The protein structure, first reported in 1996, is an eleven β-sheet-containing "barrel" shape, with the chromophore concealed at the center of the structure, shielded from quenching by aqueous solvent. This tightly-packed structure explains the importance of the entire GFP protein, which is almost completely required to maintain fluorescent activity; very little truncation is tolerated, however, point mutations are acceptable. GFP's main advantage over conventional fluorescent dyes of the time was the fact that it was non-toxic and could be expressed in living cells, enabling the study of dynamic, physiological processes.

Re-engineering GFP to Increase its Range of Color and Application

Almost as soon as its sequence was elucidated, scientists began engineering new versions of GFP through mutagenesis in order to improve its physical and biochemical properties. In 1995, Roger Y. Tsien described an S65T point mutation that increased the fluorescence intensity and photostability of GFP. This also shifted its major excitation peak from 395 nm to 488 nm, effectively ameliorating the deficiencies found in the wildtype protein and facilitating its widespread use in research. Many other mutations have since been introduced to GFP and new iterations of fluorophores are constantly being engineered.

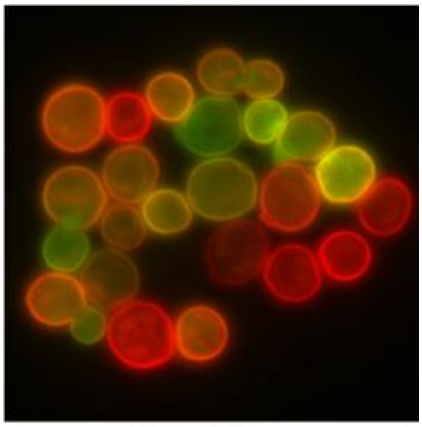


Image source: Wikimedia Commons, accessed: 5/15/2014. Photo author: Masur.

GREEN FLUORESCENT PROTEIN (CONT.)

Table 1 below lists a few common fluorescent proteins and their mutations relative to wildtype GFP. Although not listed here, many permeations within each color also exist with only slight variations separating them.

Table 1: *The specific mutations comprising common fluorophores*

Fluorescent Protein	Mutations Relative to Wildtype GFP
EGFP	F64L; S65T
EYFP	S65G; V68L; S72A; T203Y
mYFP	S65G; V68L; Q69K; S72A; T203Y; A206K
Citrine	S65G; V68L; Q69K; S72A; T203Y
ECFP	F64L; S65T, Y66W; N149I; M153T; V163A
mCFP	F64L; S65T, Y66W; N149I; M153T; V163A; A206K
Cerulean	F64L, S65T, Y66W, S72A, Y145A, H148D, N149I, M153T, V163A
EBFP	F64L, S65T, Y66H, Y145F

Please note that many fluorescent proteins found on the red side of the spectrum are not GFP derivatives, but are instead related to the dsRed protein isolated from *Discosoma sp.* Similar work has been done to expand the red-fluorescent protein repertoire; however, these proteins are unique from GFP and the mutation definitions found in **Table 2 below** may not apply.

Table 2: *Functional role of specific mutations in GFP derivatives*

Mutation	Known Function
S65T	Increased fluorescence, photostability, and a shift of the major excitation peak to 488 nm
F64L	Increased folding efficiency at 37C
Y66W	Causes the chromophore to form with an indole rather than phenol component (cyan derivatives)
Y66H	Blue-shifts the wavelength (blue derivatives)
Y145F	Increases quantum yield for BFP
Y145A and H148D	Stabilizes the structure of Cerulean derivatives
F99S, M153T, V163A	Improves folding at 37C, reduces aggregation at high concentrations
T203Y	Red-shifts the wavelength (yellow derivatives)
A206K	Interferes with dimer interface (monomeric derivatives)
K26R, Q80R, N146H, H231L, (and probably others)	Neutral mutations

GREEN FLUORESCENT PROTEIN (CONT.)

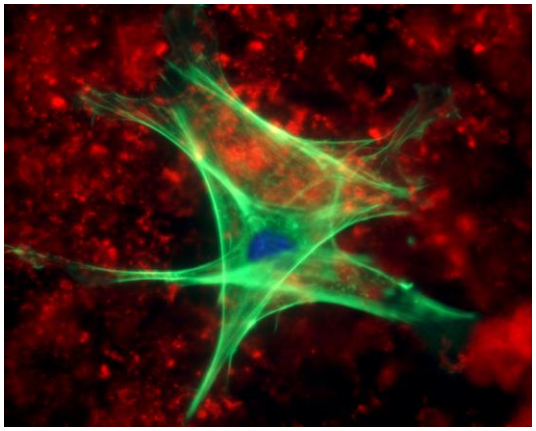
A Multitude of Applications

Due to its size and ease of use, GFP and other fluorescent proteins have become a mainstay in molecular biology. Scientists can easily utilize GFP-containing plasmids as a means to many functional ends. We've listed our favorites below, but many other uses currently exist, and new GFP technology is constantly being developed!

- **Fusion tagging:** One of the most common uses, GFP can be fused to the N- or C-terminus of a protein, which allows the scientist to visualize when and where the gene is expressed. Click here to view Addgene's collection of empty backbones for constructing fluorescent fusions.
- **Transcription Reporter:** Placing GFP under the control of a promoter of interest can be used to effectively monitor gene expression from that promoter in a given cell type. This type of transcription reporting was among the earliest uses of GFP.
- **Förster resonance energy transfer (FRET):** This is used to study the interactions between two proteins or between two domains of a protein that undergoes conformational change. Typcially two fluorescent proteins with overlapping excitation/emission spectra are used; one fused to each protein or domain being tested. Find FRET plasmids here.
- **Split EGFP:** An alternative to FRET, split EGFP has also been used to study protein-protein interactions. In this case, two portions of EGFP are fused to the proteins of interest, and when they come into close proximity, the two halves of EGFP undergo folding, maturation, and fluorescence.
- **Biosensors:** A wide array of GFP-based fluorescent biosensors has been designed to detect a variety of intracellular conditions, including ion (such as Ca^{2+}) concentrations and pH, using a range of strategies such as FRET, calmodulin, and others. Review Addgene's collection of fluorescent biosensors here.
- **Optogenetics:** Scientists can use light to detect, measure, and control molecular signals, cells, and groups of cells in order to understand their activity and visualize the effects of alterations to this activity. Learn more about optogenetics at OpenOptogenetics and find optogeneic actuators and sensors at Addgene.
- **Cell marking/selection:** Expression constructs like plasmids often include GFP as a marker to help identify which cells have successfully taken up the plasmid. This can serve as an alternative to antibiotic selection. Plasmids of this type may have the GFP under the control of an additional promoter from that of the gene of interest, or expressed from the same transcript as the gene of interest, but after an internal ribosome entry site (IRES). This is oftentimes used in conjuction with FACS (see below).
- **Fluorescence-activated cell sorting (FACS):** This is a type of flow cytometry that separates mixtures of cells into distinct populations based on fluorescent signal. Thus, FACS can be used to separate cells expressing GFP from cells that are not.

GREEN FLUORESCENT PROTEIN (CONT.)

- **Developmental/transgenic uses:** Because of its stability, GFP can be used in lineage tracking capacities in cell fate studies. It can also be used, when put under control of promoters of interest, to visualize the developmental stage at which these promoters are active. Further, GFP can label transgenically modified ES cells, which can then be used for implantation and generation of transgenic mice.
- **Purification:** GFP can be used as a general epitope tag for protein purification and a number of commercial antibodies to GFP are available.
- **Others:** We've really just scratched the surface of the potential applications for GFP. It has also been used to identify particular cell populations in drug screens, to visualize micrometastases in nude mice in cancer studies, act as a reporter for DNA double strand break repair, and to label pathogenic intracellular microbes to visualize host/pathogen interactions.



WHICH FLUORESCENT PROTEIN SHOULD I CHOOSE?

By Gal Haimovich of greenfluorescentblog | May 20, 2014

FPs Classified by the Emission Color (Emission Wavelength Range)

FPs are usually classified by emission color as outlined below (or emission wavelength range). By mutating GFP, the variants **blue FP** (BFP), **cyan FP** (CFP), and **yellow FP** (YFP) were derived.

Blue	424 - 467 nm	**Orange**	559 - 572 nm
Cyan	474 - 492 nm	**Red**	574 - 610 nm
Green	499 - 519 nm	**Far-Red**	625 - 659 nm
Yellow	524 - 538 nm	**Infra-Red**	≥ 670 nm

Unique Categories of FPs

In addition to emission wavelength range, there are other traits that need to be considered when choosing an FP:

- **Photoactivatable / Photoconvertible:** These proteins can switch their color when activated by a specific excitation wavelength. This means that the emission wavelength can change. In a few cases, the initial state of the protein is non-fluorescent, thus allowing very low background level of fluorescence. Examples for such photoactivatable or photoconvertible proteins are PA-GFP, Dendra2, and the mEOS proteins. Some proteins are reversibly switchable (e.g. rsEGFP, Dreiklang).
- **Fluorescent Timers (FT):** These proteins change their color over time. Therefore, these can be used as "timers" for cellular processes following their activation. The four main FTs are called Slow-FT, Medium-FT, Fast-FT, and mK-GO.
- **Large Stokes Shift (LSS):** Stokes shift (named after George G. Stokes) is the shift in wavelength from excitation to emission. For most FPs, Stokes shift is less than 50nm (often much less). For LSS proteins, the Stokes shift is ≥ 100nm. Specifically, these proteins are excited by UV light or blue light and their emission is green or red light. For example, T-Sapphire, LSSmOrange, and LSSmKate.
- **Fluorescent Sensors:** These FPs change their excitation/emission behavior upon environmental changes (e.g. pH, Ca^{2+} flux, etc). The most commonly used are GECIs - genetically encoded calcium indicators (e.g. GCaMP). Others include: pHluorin & pHTomato (pH sensors), HyPer (H_2O_2 sensor), ArcLight (voltage sensor), and iGluSnFr (glutamate sensor). More examples of these biosensors can be found at Addgene.

WHICH FLUORESCENT PROTEIN SHOULD I CHOOSE? (CONT.)

By Gal Haimovich of greenfluorescentblog | May 20, 2014

8 Points to Keep in Mind When Choosing an FP

1. **Excitation & Emission (ex/em):**
   - Each FP has its unique ex/em peak. Therefore, choose FPs that your system can excite, and detect the emission. For example, if your microscope has only two lasers, at 488nm and 561nm, you will not be able to use far red-FPs. If you do not have a filter that will pass blue light to the detector/camera, then BFPs are of no use to you.
   - When using more than one FP, make sure their emission light does not overlap in wavelength. In many microscopes the filters are not narrow enough to distinguish between closely related colors. Furthermore, most FPs have a broad range of emission which will be detected by longer-wavelength filters (e.g. GFP also emits yellow light).

2. **Oligomerization:** The first generations of FPs were prone to oligomerize. This may affect the biological function of the FP-fusion protein. Therefore, it is recommended to use monomeric FPs (usually denoted by a "m" as the first letter in the protein name, e.g. mCherry).

3. **Oxygen:** The maturation of the chromophore on many FPs (particularly those derived from GFP) requires oxygen. Therefore, these FPs cannot be used in oxygen deprived environment. Recently, a new GFP isolated from the Unagi eel was shown to mature independently of oxygen, making suitable for use in anaerobic conditions.

4. **Maturation Time:** Maturation time is the time it takes the FP to correctly fold and create the chromophore. This can be from a few minutes after it is translated to a few hours. For example, superfolder GFP (sfGFP) and mNeonGFP can fold in <10min at 37°C, mCherry takes ~15min, TagRFP ~100min and DsRed ~10hours.

5. **Temperature:** FPs maturation times and fluorescent intensity can be affected by the temperature. For instance, enhanced GFP (EGFP) was optimized for 37°C, and is therefore most suited for mammalian or bacteria studies, whereas GFP^{S65T} is better suited for yeast studies (24-30°C).

6. **Brightness:** The brightness is a measure of how bright is the emission. Brightness is calculated as the product of extinction coefficient and quantum yield of the protein, divided by 1000. In many cases the brightness is compared to that of EGFP which is set as 1. Some proteins are very dim (e.g. TagRFP657, which has a brightness of 0.1) and this should be taken into account.

7. **Photostability:** Fluorescent molecules gets bleached (i.e. lose the ability to emit light) after prolonged exposure to excitation light. Photostability can be as short as 100ms (EBFP) or as long as 1 hour (mAmetrine1.2). However, for most FPs it is a few seconds to a few minutes.

8. **pH Stability:** This parameter is important if you are planning to express the FP in acidic environments (e.g. yeast cytosol, which is slightly acidic, or synaptic vesicles). Some FPs have different ex/em spectra (e.g. mKeima) or change fluorescent intensity upon pH changes (e.g. pHluorin, pHTomato).

WHICH FLUORESCENT PROTEIN SHOULD I CHOOSE? (CONT.)

By Gal Haimovich of greenfluorescentblog | May 20, 2014

Keep this list handy as you plan your next experiment or to hand to the next labmate who asks you, "Which fluorescent protein should I use?" And if you're looking for more fluorescence microscopy tools and techniques to aid your work, head over to greenfluorescentblog.

Thank You to Our Guest Contributor!

Gal Haimovich, PhD, is a research fellow in the lab of Prof. Robert Singer at Albert Einstein College of Medicine. He is interested in everything related to gene expression, particularly at the RNA level. He maintains the **greenfluorescentblog.wordpress.com**.

More Helpful Websites & Resources:

- FP guide at Addgene
- Fluorescence Spectrum Viewer from BD bioscience.
- Fluorescence SpectraViewer at Invitrogen (Life technologies) site
- Evrogen Spectra Viewer at Evrogen
- ilovegfp – a site with very comprehensive data sheets on many FP variants

Three Review Articles on the Different Types of FPs:

- Stepanenko et. al. (2008) "Fluorescent proteins as biomarkers and Biosensors: Throwing color lights on molecular and cellular processes" *Curr. Protein. Pept. Sci.* 9(4):338.
- Chudakov et. al. (2010) "Fluorescent proteins and their applications in imaging living cells and tissues" *Physiol. Rev*. 90:1103.
- Wu et. al. (2011) "Modern fluorescent proteins and imaging technologies to study gene expression, nuclear localization, and dynamics" *Curr. Opin. Cell. Biol*. 23:310.

LUCIFERASE

By Jason Niehaus, Addgene | Jun 24, 2014

Luciferases are a class of enzymes capable of catalyzing chemical reactions in living organisms resulting in the emission of photons. The most familiar bioluminescent organism for most people is the firefly (*Photinus pyralis*) and perhaps not surprisingly it is also the most commonly used bioluminescent reporter. This beetle emits a yellow-green light with a peak emission at 560nm. Shortly after the initial article describing the cloning of firefly luciferase was published in 1985, several studies utilized luciferase as a genetic reporter in plant and mammalian cells. Luciferase assays have since become a gold standard in gene expression analysis and a luciferase gene (one of many available to choose from) is now a common feature in reporter plasmids.

Reporter Gene Assays

A reporter gene, such as luciferase, usually serves as an indicator of transcription within cells, where detection of the reporter protein or its enzymatic activity is measured. The effect of promoters or enhancer regions on gene expression can be determined by detection of the reporter in a specific assay, which ideally would have low background signal, high sensitivity and of course be quick, accurate and safe. In the case of a luciferase assay, photon emission is measured resulting from the catalysis of a chemical reaction requiring luciferin, ATP and oxygen as substrates. Production of photons by this bioluminescent reporter occurs slower than fluorescent-based methods, such as excitation of Green Fluorescent Protein (GFP) because of the nature of the chemical reaction compared to using a high-intensity laser to rapidly excite GFP. As a result of the different mechanisms to produce photons, chemiluminescent reporters are generally less bright than fluorescent proteins, but have the advantage of lower background levels and improved signal sensitivity since photons are simply measured – they are not required to initiate the reaction.

Biological Sources of Luciferase

Although luciferase isolated from the firefly beetle (*Photinus pyralis*) is the most frequently used bioluminescent reporter, other luciferases have been identified in various species with differing kinetics, substrate requirements and photon emission wavelengths. *Renilla* luciferase from the sea pansy (*Renilla reniformis*) emits a blue light with a peak emission at 480nm upon catalyzing a chemical reaction with the substrate coelenterazine and is likely the second most commonly used luciferase. Other organisms from which luciferases have been cloned from include click beetles (*Pyrophorus plagiophthalmus*), railroad worm

LUCIFERASE (CONT.)

(*Phrixothrix hirtus*), and a variety of marine crustaceans. Many of these luciferases have been subjected to further engineering to improve upon the properties of the native enzyme.

Table 1 below compares some general advantages and disadvantages of luminescence versus fluorescence.

Table 1: Properties of Luminescence versus Fluorescence

	Luminescence	Fluorescence
Source of Emitted Photons	Chemical reaction	High-energy photons
Kinetics of Photon Generation	Slower	Faster
Cofactors/Substrates	Required	Not Required
Signal Strength	Lower	Higher
Sensitivity	Higher	Lower
Background	Lower	Higher
Post-translational Modification	Not Required	Required
Photobleaching/Phototoxicity	Not susceptible	Susceptible
Subcellular Imaging	Improving	Well-established
High-throughput assays	Improving	Well-established

Table 2 below summarizes the native enzymes of luciferases.

Table 2: Properties of native luciferases

Organism	Peak emission (nm)	Size (kDa)	Substrate	Cofactors	Secreted
Luminous shrimp (***Oplophorus gracilirostris***)	455	19	Coelenterazine	None	Yes
Luminous Japanese ostracod (***Cypridina noctiluca***) **Sea firefly** (***Cypridina hilgendorfii***)	465	62	*Cypridina (Vargulin)*luciferin	None	Yes
Sea pansy (***Renilla reniformis***)	480	36	Coelenterazine	None	No
Deep sea coceopod (***Gaussia princeps***)	480	20	Coelenterazine	None	Yes
Metridia longa	480	24	Coelenterazine	None	Yes
North American firefly (***Photinus pyralis***)	560	61	Luciferin	ATP	No
Japanese firefly (***Luciola cruciate/lateralis***) **Italian firefly** (***Luciola italica***) **Southern Russian firefly** (***Luciola mingrelica***)	562	64	Luciferin	ATP	No
Jamaican click beetle (***Pyrophorus plagiophthalamus***)	544-593	64	Luciferin	ATP	No
Railroad worms (***Phrixothrix hirtus***)	630	64	Luciferin	ATP	No

LUCIFERASE (CONT.)

Common Uses for Plasmids Expressing Luciferase

Luciferase Reporter Assay

The typical luciferase-containing plasmid is most often used to investigate the effect of regulatory elements, such as promoters, enhancers and untranslated regions, or the effect of mutations of these regulatory elements on gene expression. The regulatory element (wild-type or mutant) is cloned upstream of the luciferase gene in the plasmid, which is then introduced into the desired mammalian, plant or bacterial cells. Expression of luciferase as the reporter gene is measured via its activity to produce light, often with a luminometer or scintillation counter, and the activity measurement can then be used to quantify the effect of the regulatory element. Depending on the specific luciferase used, a compatible substrate must be provided for the enzyme to catalyze the light-producing reaction. Many luciferase assays requires cell lysis as the most efficient means to disrupt the cell membrane and deliver the substrate; however, secreted luciferase or the use of alternative substrate reagents can permit measurements of luminescence from live cells.

Monitor miRNA Knockdown

Expression of luciferase can be monitored as a downstream measurement of miRNA efficiency on a target sequence. The gene of interest is cloned into a plasmid downstream of luciferase containing a stop codon, which is then transfected into the desired cells. A hybrid mRNA transcript containing luciferase and the gene of interest permits translation of functional luciferase, until miRNA target and degrade the mRNA sequence, resulting in a decrease in luciferase activity. This experimental protocol can readily screen miRNAs for effectiveness.

Analyze Cell Signaling Pathways

Known cellular signaling pathways can be investigated in a high-throughput screen using luciferase activity as a measure of signaling activity. Multiple response elements are cloned into a plasmid upstream of a minimal promoter linked to luciferase. Activation of the pathway will result in luciferase expression and more complex cell treatments can be designed to further explore the pathway.

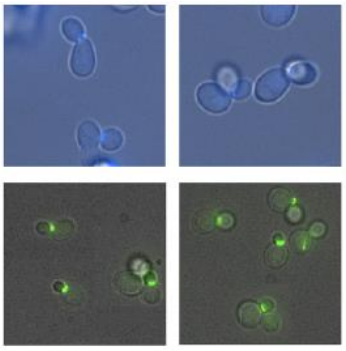
Browse Plasmids

- Browse plasmids that contain luciferase

CHAPTER 5: PLASMID TAGS

TAG YOUR FAVORITE YEAST GENES WITH EASE

By Julian Taylor-Parker, Addgene with contributions from Marcy Patrick, Addgene | Nov 19, 2013



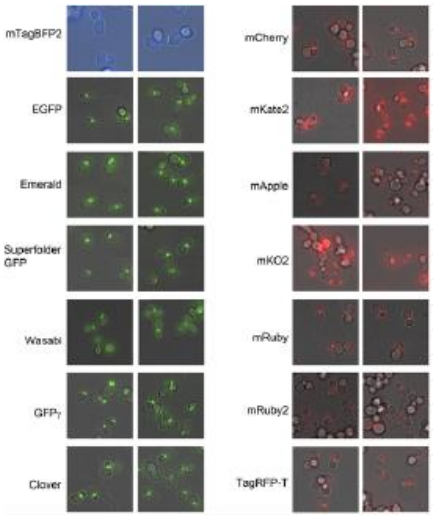
Modified from Lee S, et al. (2013) PLoS ONE 8(7): e67902.

Homologous recombination is the process by which nearly all domains of life repair genomic damage, specifically double strand breaks**. *Researchers have long taken advantage of this natural process to* integrate protein tags into the genomes of *S. cerevisiae* and *S. pombe*.** The protocol is surprisingly simple, requiring only a PCR product containing the modifying sequence flanked by approximately 50 base pairs of sequence homologous to the chromosomal site of insertion. The linear PCR product is introduced into the cell by direct transformation. A given insert will typically contain both a protein modification sequence and a selectable gene product for isolation of successful transformants.

Addgene distributes several ready-to-use, modular plasmids, combining fluorescent tags, epitope tags, protease sites, and selection markers. These are especially useful in protein complex studies where tagging of multiple protein products is desired, as multiple selection markers can ensure that all desired tags have been integrated. Simply design your amplification primers with the desired targeting homology—in frame, of course—and start tagging!

Yeast-Optimized Fluorophores for Imaging

Many imaging studies rely on **direct fusion of fluorescent proteins** (FPs) to a yeast gene of interest. These fluorescently tagged genes are expressed under native conditions and allow scientists to not only track the abundance, movement, and localization of individual proteins, but also investigate protein-protein interactions via FRET. Sidae Lee, Wendell Lim, and Kurt Thorn at UCSF have recently developed a series of blue, green, and red FPs that are codon optimized specifically for expression in yeast. These tagging vectors are based on previously described pFA6a-link vectors and include a Kan, SpHIS5, or CaURA3 selection marker. *Lee et al.(1)* assessed many of these fluorescent tags in *S. cerevisiae*, looking at their performance in categories such as brightness, stability, and disruption of the



TAG YOUR FAVORITE YEAST GENES WITH EASE (CONT.)

tagged protein. Based on their findings, the authors recommend optimal FP combinations for use in yeast imaging, categorized by specific filter sets and experimental output requirements. Select from these yeast-optimized fluorophore tagging vectors for your single or multi-color imaging experiments.

If you're looking for a great resource about imaging techniques, check out Kurt Thorn's microscopy blog.

Interested in Epitope Tags?

Others may be interested in **attaching epitope tags** to their genes of interest, allowing for easy capture and detection of proteins and complexes, without the artifacts sometimes associated with plasmid-based overexpression. Tim Formosa, at the University of Utah, has built a complete collection of yeast tagging modules with each possible combination of protease site (TEV or PreScission), epitope tag (12xHis, 2xStrep, 3xFlag, Protein A, or V5), and selection marker (KanMX, HphMX, or His3MX). Each PCR product from this collection will yield an insert with the format (protease site)-6xGly linker-(epitope tag)-ADH1 terminator-(selection marker). Additionally, Dr. Formosa has deposited six plasmids with a multiple cloning site in place of the epitope tag for creation of your own unique protein fusions. This collection is ideal for **tandem affinity purification** of protein complexes.

John Pringle and Jürg Bähler have deposited a large collection of plasmids for genome modification in yeast developed by Dr. Pringle's former lab at UNC Chapel Hill. *Bähler et al.(2)* describe a modular collection of plasmids for a wide variety of genome modifications in *S. Pombe*, including full and partial gene **deletion**, **overexpression** (by promoter substitution), and tagging at either the N- or C-terminus (3xHA, 13xMyc, GST, or GFP). *Longtine et al.(3)* describe a complimentary set of plasmids for use in *S. cerevisiae*, with the additional benefit of multiple selection markers for combining modifications within a single strain.

In addition to the collections featured above, many other modular yeast tagging systems have been developed in the labs of Anne Robinson, Eishi Noguchi, and Melissa Moore, to name a few.

References:

1. *Lee S, Lim WA, Thorn KS. PLoS One. 2013 Jul 2;8(7):e67902.*
2. *Bähler J, Wu JQ, Longtine MS, Shah NG, McKenzie A 3rd, Steever AB, Wach A, Philippsen P, Pringle JR. Yeast. 1998 Jul;14(10):943-51.*
3. *Longtine MS, McKenzie A 3rd, Demarini DJ, Shah NG, Wach A, Brachat A, Philippsen P, Pringle JR. Yeast. 1998 Jul;14(10):953-61.*

CHAPTER 6: YOU'VE MADE A PLASMID… .NOW WHAT?

HOW TO NAME YOUR PLASMID IN 3 EASY STEPS

By Matthew Ferenc, Addgene | Feb 13, 2014

There are no universal rules for naming plasmids but here are some good guidelines to follow in order to ensure that people can quickly and easily identify what your plasmid contains and other important information.

Tip: A lowercase "p" is often used as the first letter of a plasmid name and simply denotes that the object is a 'plasmid'. "p" is for plasmid.
pXXXXX-XXXX

Step 1: Backbone Name

Include the empty backbone name in your plasmid name. This simple piece of information can often convey many important details. Once you know the backbone a plasmid is based on, you can usually derive: a) the bacterial antibiotic resistance, b) the promoter that drives the insert, and c) any other selection markers (for use in other cell types, e.g. eukaryotic cells).
pBACKBONE-XXXXX

Tip: For a catalog of published and commercially available empty backbones, visit Vector Database.

Step 2: Insert Name

Include information about the insert in your plasmid name. This is often a 3-6 letter representation of the gene (or DNA sequence).

pBACKBONE-hGene

Tip: Often researchers will add a lower case letter to the beginning of their insert abbreviation to specify what species it is. Example: 'h' is for Human (*homo sapiens*), 'm' is for mouse (*mus musculus*), 'r' is for rat (*rattus rattus* or *rattus norvegicus*), etc.

Step 3: Add Your Tags

Add any tags or fusions that are on your insert. Typically you would list any tag or fusion protein in the order they appear in the plasmid and their relative position to the insert. Example, if you have a Flag tag on the N-terminal of your insert, you would list it first.

pBACKBONE-Flag-hGene

If there was also an EGFP fused to the C-terminal of your insert you would list it after the insert.

pBACKBONE-Flag-hGene-EGFP

How to Name Your Plasmid in 3 Easy Steps (cont.)

Other Considerations:

If your insert contains a mutation or modification, this should be included in the plasmid name. Mutations are generally listed as the amino acid change and not a nucleotide change. The proper way to denote an amino acid mutation is to list the one letter abbreviation of the wild type amino acid immediately followed by its position (number) relative to the start Methionine (Met) followed by the one letter abbreviation of the mutated amino acid currently at that position.

In contrast, unmutated or Wild Type (wt) versions of the insert are often denoted by "wt" either directly before or directly after the insert name.

Example: If the Glutamine at position 295 was mutated to an Alanine, Q295A.

Mutant version: *pBACKBONE-Flag-hGene(Q295A)-EGFP*

Wildtype version: *pBACKBONE-Flag-hGene-EGFP or pBACKBONE-Flag-hGene(wt)-EGFP*

These simple rules allow any scientist to know what is in a plasmid and often how it can be used just by reading the name.

How to Verify Your Plasmid

By Lianna Swanson, Addgene | Aug 28, 2014

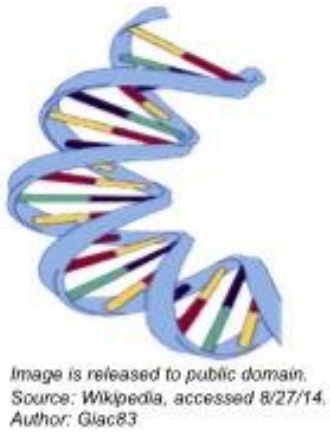


Image is released to public domain. Source: Wikipedia, accessed 8/27/14. Author: Glac83

Congratulations, you have a plasmid expressing your gene of interest (YGOI) and are ready to dive into your functional experiments! Whether you've cloned the plasmid yourself or obtained it from a colleague down the hall, it is always a good idea to take some time to confirm that you are working with the correct construct, and verify that the plasmid you received matches the expected sequence. Here at Addgene, we process all of the plasmids we distribute for quality control purposes in order to confirm the integrity of the DNA.

These are Addgene's two recommended methods for plasmid DNA verification: sequencing and diagnostic restriction digest.

Sequencing

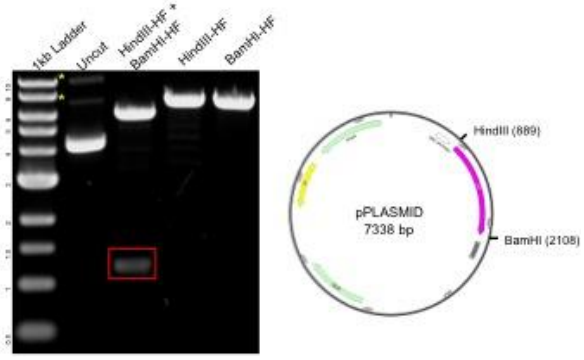
Sequencing determines the precise order of nucleotides within the DNA molecule, in this case a plasmid. To get started, you will first need to design and synthesize primers that perfectly compliment your plasmid sequence. We recommend starting with a backbone-specific primer that will sequence over the Multiple Cloning Site (MCS) and into YGOI. This way you can avoid designing multiple primers to verify unique genes inserted into the same backbone. Addgene has curated a comprehensive vector database that will help you find reference sequence for many commonly used backbones, as well as the specific primers used to confirm their integrity. You can also find a list of our most commonly used sequencing primers at the following link: http://www.addgene.org/mol_bio_reference/sequencing_primers/. It usually takes a couple of days to receive results after submitting your sample to a sequencing core (depending on the core facility and services available at your institution); however, it will save you time in the long run knowing that you are working with the correct plasmid.

Diagnostic Restriction Digest

Diagnostic digests can be used to confirm the relative structure of the plasmid based on the predicted sizes and organization of different features within the plasmid. One benefit of restriction analysis is that it can be used successfully without actually having full plasmid sequence available to you. This method is relatively quick and can be done right in your lab in less than a day (as long as you have purified DNA). Diagnostic restriction digests are comprised of 2 separate steps: 1) incubating your DNA with the selected endonucleases which cleave the DNA molecules at specific sites; and 2) running the reaction on an agarose gel to determine the relative sizes of the resulting DNA fragments.

How to Verify Your Plasmid (cont.)

The most common way of utilizing a restriction digest is to confirm the presence of an insert in a particular vector by excising it from the backbone. This is accomplished by using a combination of specific endonucleases that flank the insert. You will need to know both the approximate size of the vector backbone as well as the predicted size of the insert. You can search NCBI for YGOI to find the particular reference sequence if necessary.



The example plasmid on the right has a total size of 7.3kb, with the insert comprising 1.2kb of the total. The plasmid was digested with 2 unique enzymes (HindIII and BamHI) and run on an agarose gel. The resulting gel image includes a 1kb ladder (lane 1) that has bands ranging from about 500bp to 10kb, with the 3.0kb fragment having increased intensity to serve as a reference band. The uncut DNA (lane 2) shows 3 possible plasmid conformations, with relaxed and nicked marked with asterisks (*). The digested samples in the last 3 lanes include HindIII and BamHI alone resulting in a single band of the full size of the plasmid, ~7.3kb, and one double digest with HindIII and BamHI together, matching the backbone size of about 6kb and the released insert at about 1.2kb (red box). The results on the gel match the predicted sizes inferred from the plasmid information.

Restriction Digest Tips and Tricks:

The following tips for your digest and for your gel will make it easier for you to obtain a useful and informative diagnostic restriction digest.

For your digest:

- **Try choosing unique enzymes.** Enzymes that only cut once allow you to more easily and accurately visualize the full size of your construct.
- **Consider buffer and temperature compatibility when digesting with more than one enzyme.** Consult the manufacturer's manual for the optimal working conditions for each enzyme.
- **Watch out for methylation issues.** Enzymes like XbaI and ClaI are sensitive to methylation and their activity may be blocked. If you have to use these enzymes for your digest, you will need to purify your DNA from a dcm or dam methylation-deficient bacterial strain such as JM110 or INV110.

How to Verify Your Plasmid (cont.)

- **Avoid star activity.** Some endonucleases (for example BamHI) are capable of cleaving sequences which are similar, but not identical, to their defined recognition sequence. Most enzyme manufacturers make High Fidelity versions of the endonucleases and/or supply custom buffers as means to avoid this issue.

For your gel:

- **Add ethidium bromide (EtBr) to your gel before pouring it.** EtBr binds to the DNA and allows you to visualize the DNA under ultraviolet (UV) light, and adding it to your gel will save time!

- **Don't forget to add loading buffer to you digest reactions before loading them.** The glycerol in the buffer will make sure your sample settles in the gel well and the dyes provide a visual reference point so you can easily assess how far the gel has run. *Bonus:* The dyes also run at predicted sizes so you can estimate how far down the gel your bands have travelled based on the dye!

- **Always run a ladder.** Ladders allow you to interpret the bands that you get in your sample lanes. Choose your ladder based on the expected band sizes.

- **Always run control uncut DNA to ensure your enzymes are working.** When uncut plasmid DNA is isolated and run on an agarose gel, you are likely to see 3 bands. This is due to the fact that the circular DNA takes on several conformations the most abundant being: supercoiled, relaxed and nicked. If your digest lanes look like your uncut lane then there is something wrong!

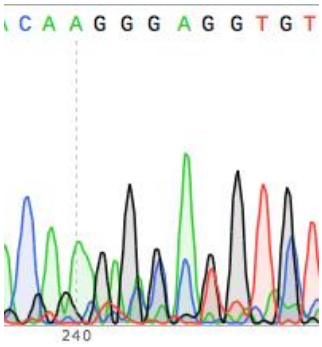
- **Quantify your DNA**. Loading too much DNA will make it difficult to obtain crisp bands and analyze the results. *Bonus:* knowing how much DNA you have loaded in each well will allow you to approximate the DNA mass of comparably intense samples of similar size.

- **Run the gel at 80-150V until you have good separation between your bands.** Stopping the gel when the bromophenol blue dye line is approximately 75-80% of the way down the gel will ensure you keep smaller bands from running off; however, you may need to run the gel for longer to achieve good separation of larger DNA fragments.

We hope these tips demonstrate that plasmid verification is not just necessary but also an easy process. Please visit Addgene's resource for plasmid verification to find additional tips and detailed protocols on topics such as how to set up your digests and pouring/running a DNA gel.

6 Tips for Analyzing and Troubleshooting DNA Sequencing Results

This blog was originally published on BitesizeBio here. *By Lianna Swanson, Addgene | Jun 26, 2014*



Addgene scientists analyze 50-100 sequencing reactions a week. Here are some good habits that we wanted to pass on to you to make sure you are getting the most out of the data you get back from your sequencing runs.

When you run a restriction digest on a gel you always include proper controls like uncut DNA and the proper ladder. These controls help you properly visualize your results. The most important of those is to always look closely at the trace file (or chromatogram) of the sequencing results you get back from your favorite sequencing facility.

When it comes to DNA sequencing the chromatogram is your visual control. And, like all controls, missing out is a big mistake.

An example of where the chromatogram can come to your rescue for analyzing and troubleshooting your DNA sequencing results

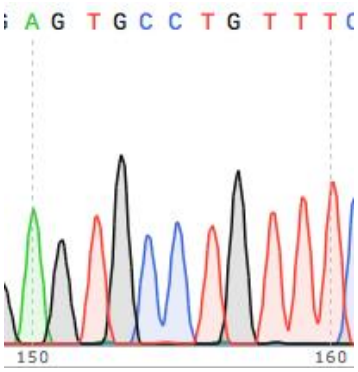
Above (top left corner of this section) is an example of a seemingly clean DNA sequence (no Ns in sight). If you never looked at the trace you would be happy.

But look closer, the overlapping peaks in the chromatogram suggest the results are not as certain as the sequence may suggest. In fact this is so ambiguous that the DNA sequencing reaction should be repeated.

Here are a few guidelines to help with DNA sequencing troubleshooting and analysis:

1. You can use any of the following programs to view your .ab1 chromatogram file

- 4Peaks (Mac)
- SnapGene Viewer (Mac/PC)
- FinchTV(Mac/PC)
- Sequence Scanner (PC)
- Chromas (PC)



2. You should see individual, sharp and evenly spaced peaks

…like these (to the right)….

3. Expect to get 500-700 bases of clean reliable DNA sequence

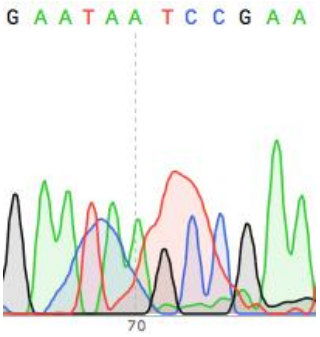
Anything less and you might suspect contamination in your sample or consider asking your sequencing facility to apply a special protocol for a difficult template. Anything more and you're venturing into the uncertain terrain.

6 Tips for Analyzing and Troubleshooting DNA Sequencing Results (cont.)

4. Never trust the first 20-30 bases of a DNA sequencing read

The peaks here are usually unresolved and small, so I suggest designing your primer at least 50bp upstream of the sequence of interest.

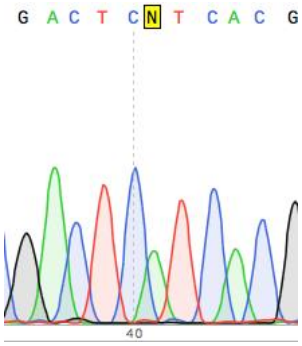
5. Use a silica spin column for purification of the samples you send for DNA sequencing



If your sequencing facility requires you to perform your own Big Dye PCR amplification reaction (as opposed to using the all-inclusive service some companies offer), you can purify the product either via the Sodium Acetate/isopropanol precipitation method or using a silica spin column available from several vendors. The precipitation method has an unfortunate side effect of messing up the reaction around base 70-75 of the read (see image to the left), so I would strongly recommend using a silica spin column. They can be pricey, but well worth it.

6. Edit your DNA sequence

Finally, when you do see a miscalled peak, don't be shy. Feel free to edit it. Most chromatogram viewing programs (even the free ones) allow you to edit the sequence.



We hope these tips will help you get the most out of your DNA sequencing results and to troubleshoot any problems that come up. Good luck analyzing your sequences!

More DNA Sequencing Resources:

- Tips for Using BLAST to Verify Plasmids

Acknowledgements:

- Thank you BitesizeBio for originally publishing this and allowing us to share it with our readers!
- The chromatograms in this article were created with SnapGene.

TIPS FOR USING BLAST TO VERIFY PLASMIDS

By Jason Niehaus, Addgene | Jun 24, 2014

At Addgene, we continually use the Basic Local Alignment Search Tool (BLAST) provided by NCBI. BLAST helps us compare the sequencing results of the plasmids in our repository with known reference sequences, such as full plasmid sequences provided by the laboratories that deposit their plasmids with us or other entries in NCBI's numerous databases.

As our repository has grown over the years (we now have over 33,000 plasmids!), the number of sequencing results we analyze as part of our quality control process has steadily grown. On a busy week, we may need to analyze more than 100 sequences each day as part of our quality control process. Consequently our team has refined our use of the BLAST web browser interface as efficiently as possible.
If you find yourself frequently on the BLAST website to verify plasmids or validating your new clones, try these tips to make the most of your time and sequence!

Choosing a BLAST Program

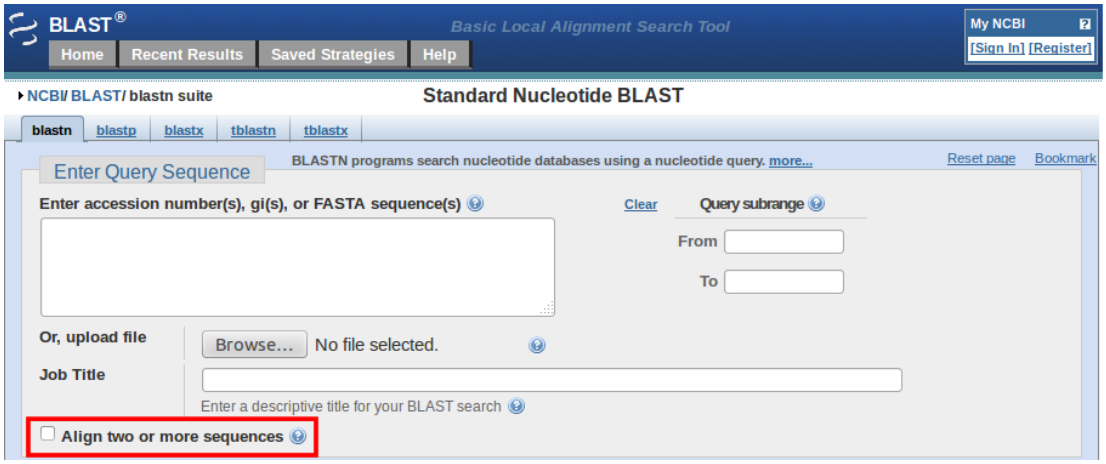
Of the five BLAST programs available, we primarily use Standard Nucleotide BLAST (blastn), Standard Protein BLAST (blastp), and Translated BLAST (blastx), depending on the plasmid region sequenced. NCBI has a terrific getting-started guide for BLAST, which includes a simple explanation of the different BLAST programs, databases, and elements of the BLAST search pages.

At Addgene, we use blastn to compare our sequencing results with empty vector or gene sequences to first determine how much of the result is trustworthy. Then we identify any discrepancies in the nucleotide sequence, such as mismatches, deletions, or insertions. We use blastp or blastx to compare our sequencing results to protein sequences to check open reading frames (ORFs) and determine the potential effect of any nucleotide discrepancies. The blastp and blastx programs are optimized differently and you may want to select one (or both) depending on the information you want to verify. We will delve into these differences below.

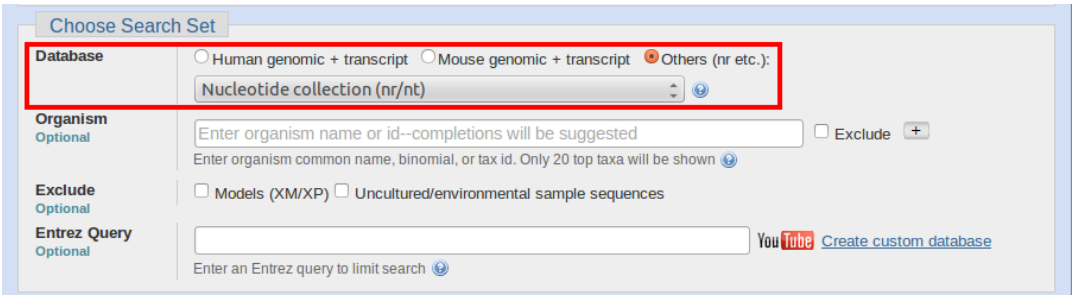
Optimizing *blastn* Searches

On the Standard Nucleotide BLAST page, ***the first decision we make is*** whether to compare our sequencing result to a single known reference sequence or to a BLAST sequence database. If you know the expected nucleotide sequence, check the "Align two or more sequences" checkbox and paste your reference sequence into the Subject Sequence box that appears. Aligning two nucleotide sequences is probably the fastest BLAST search to perform and will save you time compared to other types of BLAST searches.

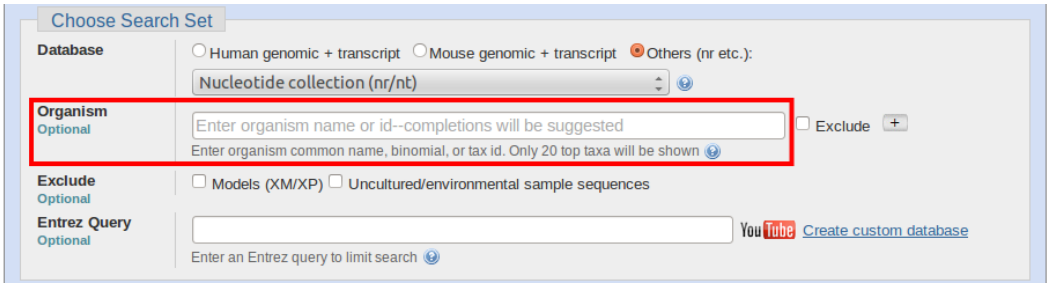
TIPS FOR USING BLAST TO VERIFY PLASMIDS (CONT.)



If you do not know the exact reference sequence for your result, choose one of the BLAST sequence databases from the dropdown menu. Typically, we use the default nucleotide database "Nucleotide collection (nr/nt)" as it contains a composite of GenBank, EMBL, DDBJ, and PDB sequences and may be the most comprehensive for searching.



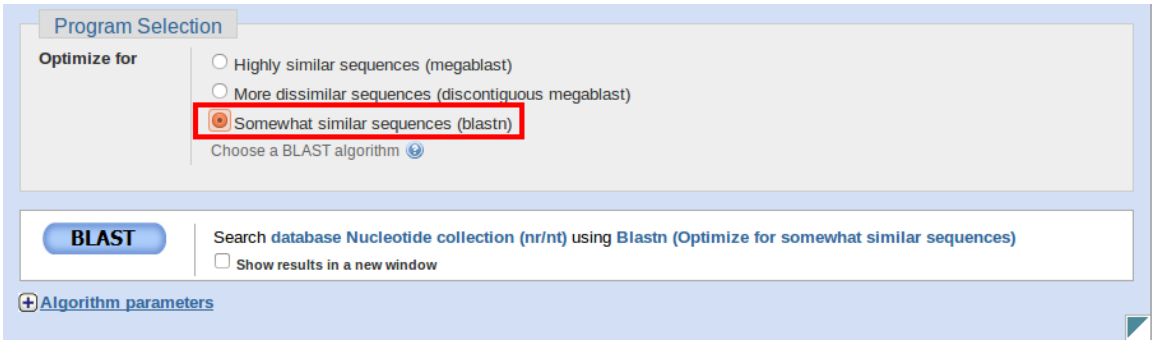
Timesaving Tip #1: *If you know the species that your sequencing result should match, enter the common or scientific name into the Organism box. This small piece of information can significantly reduce your wait time for blastn, blastp, and blastx searches!*



TIPS FOR USING BLAST TO VERIFY PLASMIDS (CONT.)

Now, before you click the BLAST button, consider the Program Selection parameter, as this will affect the amount of time to perform the search as well as the overall alignment results. The default setting is "Optimize for Highly similar sequences (megablast)", which is very fast and works best when the identity between your sequence and the reference/database sequence is ≥ 95%. [Our QC process would be trouble-free and much faster if 95% of our results were always correct!]

Since sequencing reactions are imperfect and sequence near the beginning or end of a reaction is often unreliable, we routinely select the "Somewhat similar sequences (blastn)" program for blastn so that we can extract practically every single, reliable basepair from our results.

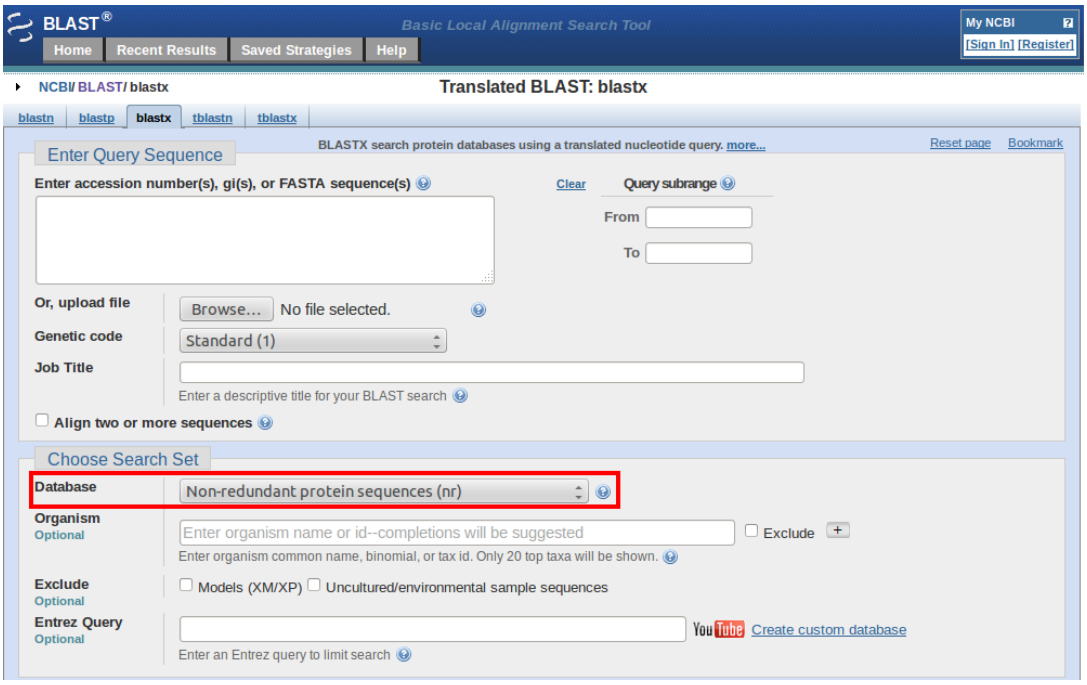


This option is not as fast as megablast, but can return longer alignments to compare with your sequencing trace file. Unlike megablast, the regular blastn program uses a smaller word size and lower scoring penalties for mismatches and gaps in the alignment. If you are curious about the differences in the blastn programs, check out the BLAST Help webpage.

Optimizing *blastx* Searches

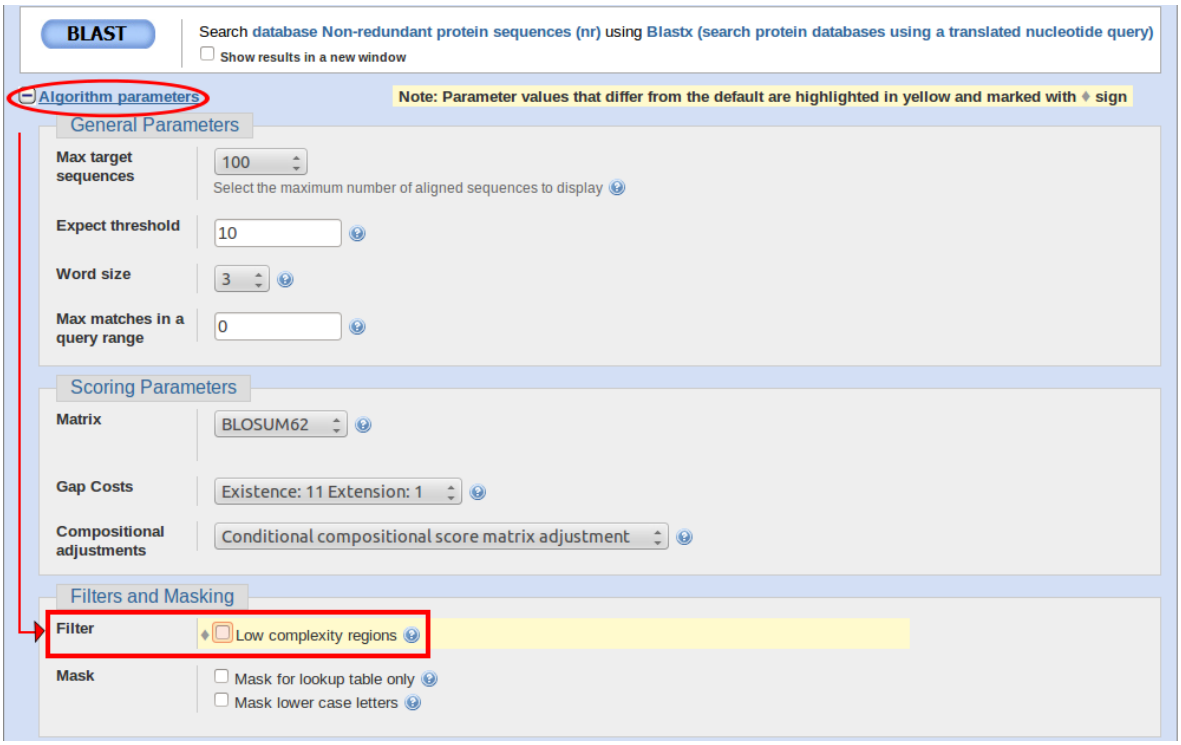
Once we have used blastn to determine the reliable portion of a sequencing result and noted any potential mismatches or gaps in the nucleotide sequence, we typically run a Translated BLAST (blastx) search to check for expected ORFs, mutations or truncations. A primary advantage of blastx is that you do not have to decide on a reading frame for your sequencing result – blastx checks all six possible frames against the database. Another benefit is that a frame shift mutation present in the ORF is readily apparent when viewing blastx results.

TIPS FOR USING BLAST TO VERIFY PLASMIDS (CONT.)



When using blastx at Addgene, we use the default "Non-redundant protein sequences (nr)" database as it contains the largest number of protein sequences. Just below the BLAST button, you may have noticed the "Algorithm parameters" link. Click on this link to view advanced BLAST options and for our suggested blastx customization. Similar to nucleotide sequences, proteins often have repeated or highly homologous regions, which by default are ignored in a standard blastx search. An alignment omitting repeated regions can be confusing, such as when you attempt to verify the starting methionine of a gene but the blastx results start the alignment at a more distal amino acid. We consistently run our blastx searches with the "Low complexity regions" filter unchecked so that these regions are included in the search to maximize the alignment length. While this recommendation is not infallible, we have found it saves analysis time to remove this default setting.

TIPS FOR USING BLAST TO VERIFY PLASMIDS (CONT.)

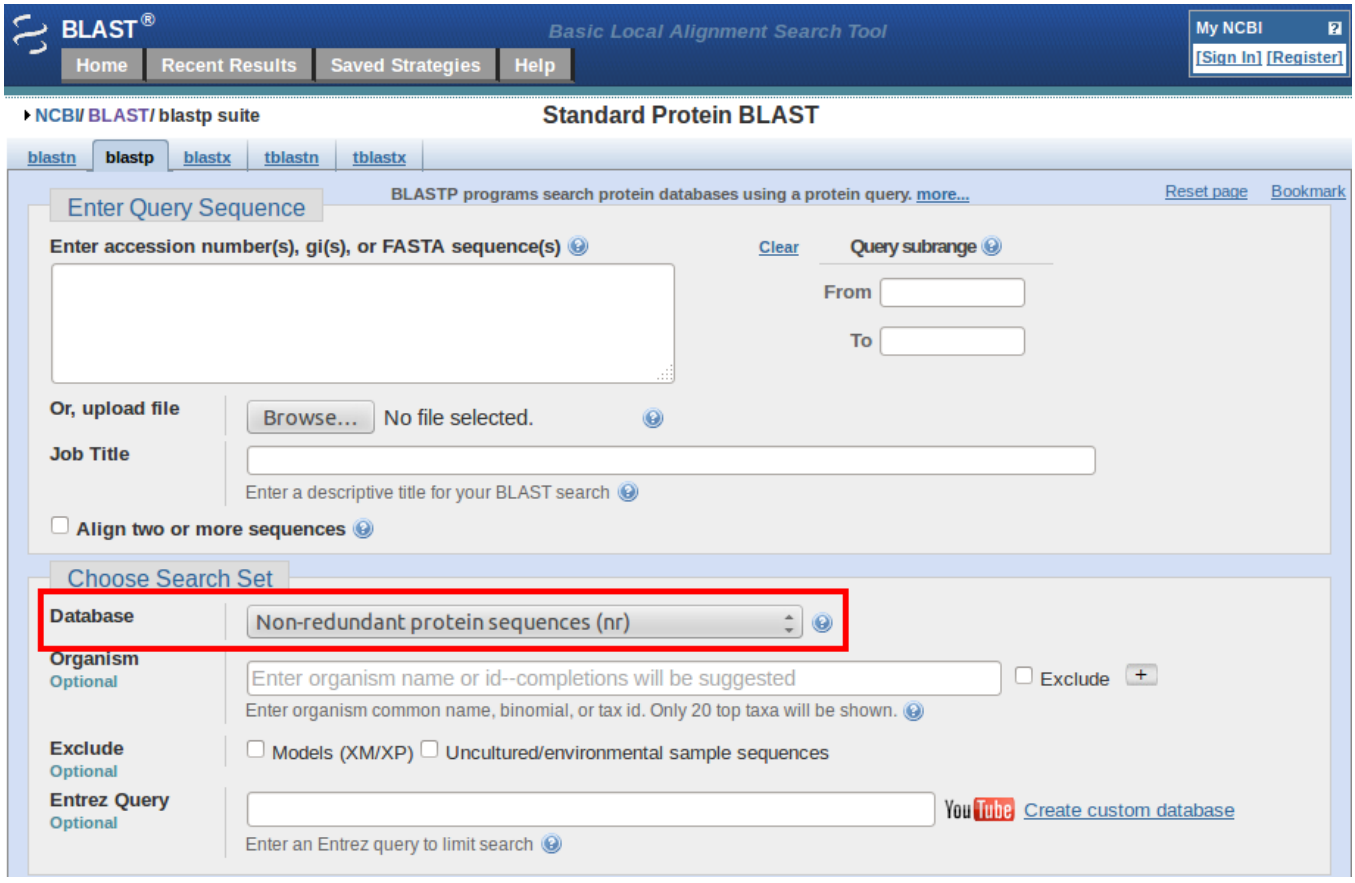


Timesaving Tip#2: *blastx searches are inherently slower than blastn or blastp, due to the additional searches involved in translating the nucleotide sequence into all six possible reading frames. If you know the expected protein sequence, use the "Align two or more sequences" option to drastically reduce waiting time for search results.*

TIPS FOR USING BLAST TO VERIFY PLASMIDS (CONT.)

Optimizing *blastp* Searches

Depending on the sequencing result, we often choose between a Standard Protein BLAST (blastp) and blastx search to verify expected protein sequence in a plasmid. If you know which reading frame to choose for your sequencing result and can easily translate it, we recommend using blastp over blastx. The primary advantage is time savings but an added benefit is that blastp searches do not filter low complexity regions by default, meaning that you do not have to remember to adjust any blastp algorithm parameters. We use the default scoring matrix BLOSUM62, but you may want to check the description of the other matrices to see if another would be more advantageous for your search.



Timesaving Tip #3: *Note that protein databases available are unlikely to have an exact entry for your favorite gene fused to an epitope tag or fusion protein. If your sequencing primer was chosen to confirm a tag or fusion protein is in-frame, we recommend using blastx with the "Align two or more sequences" option and pasting your expected protein sequence into the Subject Sequence box.*

TIPS FOR USING BLAST TO VERIFY PLASMIDS (CONT.)

BLAST Alternatives

Depending on your sequencing result and desired analysis, BLAST may not always be your optimal choice. For difficult sequence alignments that BLAST is unable to handle, Clustal is our frequent choice for pairwise or multiple sequence alignments of nucleotide or protein sequences. We also use COBALT for aligning multiple protein sequences, particularly for comparing different isoforms. In addition to our favorites, there are a number of sequence alignment tools available.

Try these resources for lists of alternatives to BLAST:

- ExPASy - http://www.expasy.org/genomics/sequence_alignment
- EMBL-EBI - http://www.ebi.ac.uk/services
  - http://www.ebi.ac.uk/Tools/webservices/#multiple_sequence_alignment_msa
  - http://www.ebi.ac.uk/Tools/webservices/#pairwise_sequence_alignment_psa

Happy BLASTing!

Do you have any tips for using BLAST to confirm your plasmid sequencing results or comments on our suggestions? Share your thoughts here to help other labs speed up their plasmid and cloning verification steps and free up more time for using your plasmids instead!

More Resources:

- Sequence Analysis of your Addgene Plasmid
- Addgene Plasmid Guide
- Sequence Analyzer by Addgene

Images:

- All BLAST images are modified screen shots from the NCBI BLAST website

CHAPTER 7: DEPOSITING YOUR PLASMIDS WITH ADDGENE

BENEFITS OF DEPOSITING

Addgene is a global, nonprofit plasmid repository, dedicated to making it easier for scientists to share plasmids. Scientists from labs like yours deposit plasmids to Addgene's collection. Deposited plasmids are stored and distributed for use in research and discovery worldwide.

Stay Organized…

Addgene manages archived samples for your lab, storing multiple copies at our site in Cambridge, MA and a copy at a partner facility offsite. All of your lab's deposited plasmids and cloning information are curated online for easy access.

Save Time…

You can easily direct requests for your plasmids to your personalized Addgene page. When your plasmids are requested, Addgene process and ships your plasmids for you.

Increase Citations…

Scientists who request your plasmids cite your publication and acknowledge your lab in their published work.

Track Reagents…

Scientists who request your plasmids cite your publication and acknowledge your lab in their published work.

Improve Community Resources…

Contributed materials are available to the global research community. Addgene's online plasmid database helps get your lab exposure for your plasmids beyond your immediate field.

Earn Free Plasmids…

As your plasmids are requested from Addgene, you earn reward points to use toward free plasmids.

Sign Your Last MTA…

Addgene's one-time deposit agreement means once your plasmids are online, you never have to sign another MTA for the distribution of your plasmids.

Manage Support Queries…

Addgene's knowledgeable and friendly support team answers questions about your deposited plasmids.

How To Deposit Your Plasmids

Ready to start depositing your plasmids with Addgene?

Addgene will archive your samples and distribute your plasmids to scientists at academic and nonprofit institutions worldwide. Depositing plasmids at Addgene is free. Please follow the steps below to submit your deposit:

1. **Choose the plasmids you want to deposit.** Laboratories commonly begin by selecting the 5-10 plasmids that are most frequently requested or most useful to scientists in their field.

2. **Provide the cloning data.** You have several options:
   - Online: Enter plasmid information on our website www.addgene.org/deposit/
   - On Paper: Fill out the enclosed General Information Form and Plasmid Submission Forms on paper. If you have plasmid maps or sequence data, email them to deposit@addgene.org
   - Share files: You can choose to let Addgene assemble the plasmid data for you. Send us a plasmid list, article references (if applicable), and any information you want included on your plasmid datasheets. You will still need to complete the General Information Form.

3. **Prepare your samples.** Addgene accepts plasmids in the form of either DNA or bacteria.
   - DNA: If depositing DNA, please send 10 µL of each plasmid (concentration: 0.2 to 1 µg/µL) in a clearly labeled 1.5 mL microcentrifuge tube. Parafilm the cap of each tube, and ship using the provided shipping materials.
   - Bacteria: If depositing bacteria, please streak only one strain per plate. Clearly label the agar side of the plate (i.e. not the lid) and use parafilm to seal. If you are sending bacteria, please use your own shipping materials that comply with safety regulations.

4. **Ship to Addgene.** Return the submission forms and plasmid samples to:
   Addgene
   One Kendall Square
   Suite B7102
   Cambridge, MA 02139
   USA

Questions? See the FAQ link at www.addgene.org/faq/ for answers to commonly asked questions. You are also welcome to email deposit@addgene.org or call +1-617-225-9000 for assistance.

How To Deposit Your Plasmids (cont.)

addgene

One Kendall Square Ste B7102, Cambridge, MA 02139
info@addgene.org • Tel: 617-225-9000 • Fax: 888-734-0533

General Information Form

You must submit this form in order to complete your submission.

Account Information

Your name: _____

Email address: _____

Host institution: _____

Principal investigator's name (if different): _____

Please answer the following questions to the best of your knowledge:

Is there any patent that would prevent the distribution of your plasmids to academic laboratories? Or is the distribution of this material governed by any license, corporate sponsored research, or MTA with a third party? ☐ Yes ☐ No

If YES, please explain: _____

Terms of Submission

You must agree to the following acknowledgements for Addgene to accept your materials:

I hereby certify that I am the individual stated in the Account Information Section and that all of the information provided in this submission is true and accurate to the best of my knowledge.

I hereby certify that I am affiliated with the academic or non-profit institution named in the Account Information Section.

I hereby certify that the material I am submitting to Addgene qualifies as BL1 or BL2 as defined by NIH guidelines.

As the Provider Scientist, I hereby agree to the distribution of the Material to other scientists under the terms of the Uniform Biological Material Transfer Agreement (UBMTA) and any applicable Ancillary Agreements, or alternate material transfer agreements, if required by my Institution.

Print Name: _____ **Date:** _____

Signature: _____

How To Deposit Your Plasmids (cont.)

addgene

One Kendall Square, Ste B7102, Cambridge, MA 02139
deposit@addgene.org • Tel: +1-617-225-9000

Plasmid Submission Form

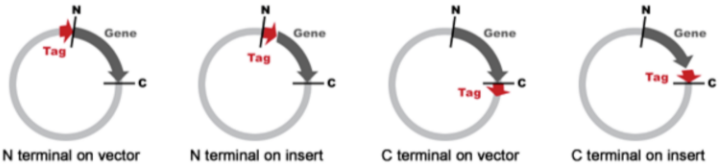
Plasmid Name: _____

Plasmid Purpose: _____
(For example: Inducible expression of p53 in mammalian cells)

Reference (or PubMed ID #): _____

Insert and Features

1) Gene/insert name (ex: MAPKK1): _____
Alternative gene names: _____
GenBank/Entrez ID of gene: _____
2) Insert size: _____
3) Species of gene (ex: human, mouse): _____
4) Relevant mutations/deletions (ex: S320A, Δ1-200): _____
5) Fusion proteins or tags (ex: GFP, GST, HA, Flag): See diagrams below.
Fusion/Tag:_____ ❒ N-vector ❒ N-insert ❒ C-vector ❒ C-insert
Fusion/Tag:_____ ❒ N-vector ❒ N-insert ❒ C-vector ❒ C-insert



Cloning Information

1) Vector backbone (ex: pBluescript II KS+): _____
Backbone manufacturer (ex: Stratagene): _____
2) Type of vector (check all that apply):

❒ Adenoviral ❒ RNAi ❒ Yeast Expression ❒ Mammalian Expression
❒ Retroviral ❒ Luciferase ❒ Insect Expression ❒ Bacterial Expression
❒ Lentiviral ❒ β-gal ❒ Worm Expression ❒ CRE/LOX
❒ Other: _____

3) Backbone size (without insert): _____
4) 5' cloning site (ex: EcoRI): _____ 3' cloning site: _____
Was either restriction site destroyed during cloning? _____
5) Promoter: _____
6) Cloning Method: ❒ Restriction Enzyme ❒ Gateway ❒ TOPO ❒ LIC
Other:_____
7) 5' sequencing primer (ex:T7): _____ 3' sequencing primer: _____

(Plasmid Submission Form - page 1 of 2)

How To Deposit Your Plasmids (cont.)

addgene

One Kendall Square, Ste B7102, Cambridge, MA 02139
deposit@addgene.org • Tel: +1-617-225-9000

Growth and Distribution

1) Bacterial resistance: (Please specify if the plasmid contains multiple antibiotic resistance genes)

❒ Ampicillin ❒ Kanamycin ❒ Tetracycline ❒ Chloramphenicol

❒ Other: _____

2) In bacteria, is this plasmid:

❒ High copy number ❒ Low copy number ❒ Unknown

3) Can this plasmid be grown in standard *E.coli* at 37°C?

❒ Yes ❒ No

If NO, please specify preferred bacterial strain for growth and preferred growth conditions:

_____

4) Selectable markers (non-bacterial): ❒ None

❒ Neomycin	❒ Zeocin	❒ TRP1	❒ HIS3
❒ Puromycin	❒ Blasticidin	❒ LEU2	❒ Basta
❒ Hygromycin	❒ Gentamicin	❒ URA3	❒ Other: _____

5) When expressed in bacteria, will this plasmid produce anything hazardous to humans or animals?

❒ Yes ❒ No

Please note: Addgene does NOT accept plasmids that require a BL3 or BL4 environment or plasmids from the list of *Select Agents* described in the Public Health Security and Bioterrorism Preparedness and Response Act of 2002, Public Law 107-188. The *Select Agents* list can be found at www.addgene.org/selectagents.html.

6) Is there any patent that would prevent the distribution of this plasmid to academic laboratories? Or is the distribution governed by any license, corporate sponsored research, or MTA with a third party?

❒ Yes ❒ No

If YES, please explain: _____

7) Was this plasmid created using funding from Howard Hughes Medical Institute?

❒ Yes ❒ No

8) If you did not originally clone this gene, from whom and where did you receive the plasmid used to derive this plasmid?

_____

9) Other information or comments:

_____

_____

_____

_____

NOTE: Please staple pages 1 and 2 together, or print this form double-sided.

If you have plasmid maps or any additional data, please include them in your shipment.
Email deposit@addgene.org or call +1-617-225-9000 if you need any assistance.

(Plasmid Submission Form - page 2 of 2)

ACKNOWLEDGEMENTS AND FINAL WORDS

Special thanks to Gal Haimovich of *greenfluorescentblog* for helping us explain why things glow.

Contributions to this eBook have been made by the following Addgenies:

Melina Fan

Matthew Ferenc

A. Max Juchheim

Caroline LaManna

Margo Monroe

Kendall Morgan

Jason Niehaus

Marcy Patrick

Lianna Swanson

Julian Taylor-Parker

If you have any questions, comments, or suggestions about how Addgene can improve its educational content, please contact us at blog@addgene.org.

Looking for more information about Addgene?
Connect with us!

www.addgene.org
www.facebook.com/Addgene
www.twitter.com/addgene
www.linkedin.com/company/addgene