

# SYNTHETIC BIOLOGY PRACTICAL

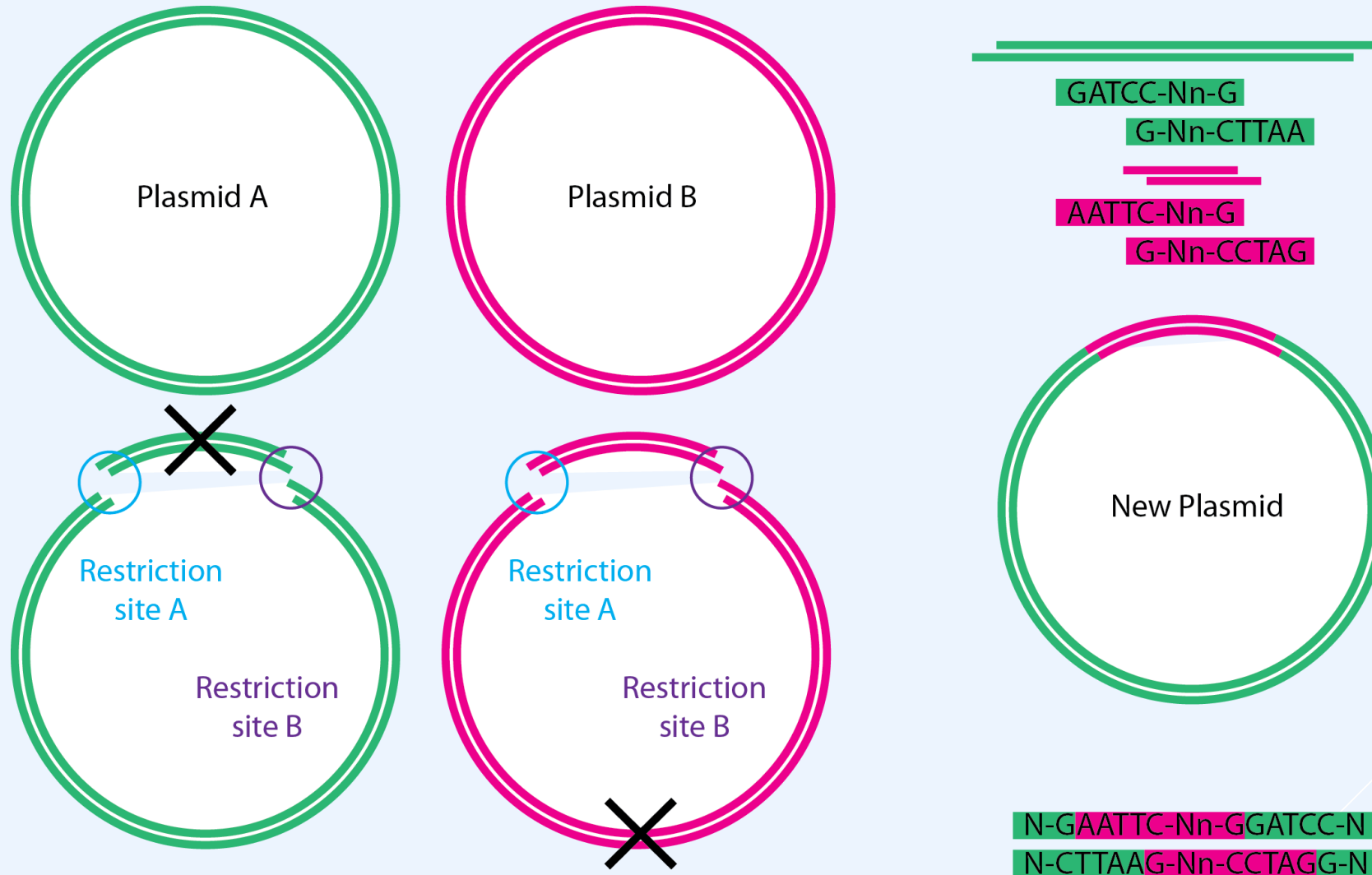
Biological Databases and tools for handling  
DNA sequences



# BACKGROUND

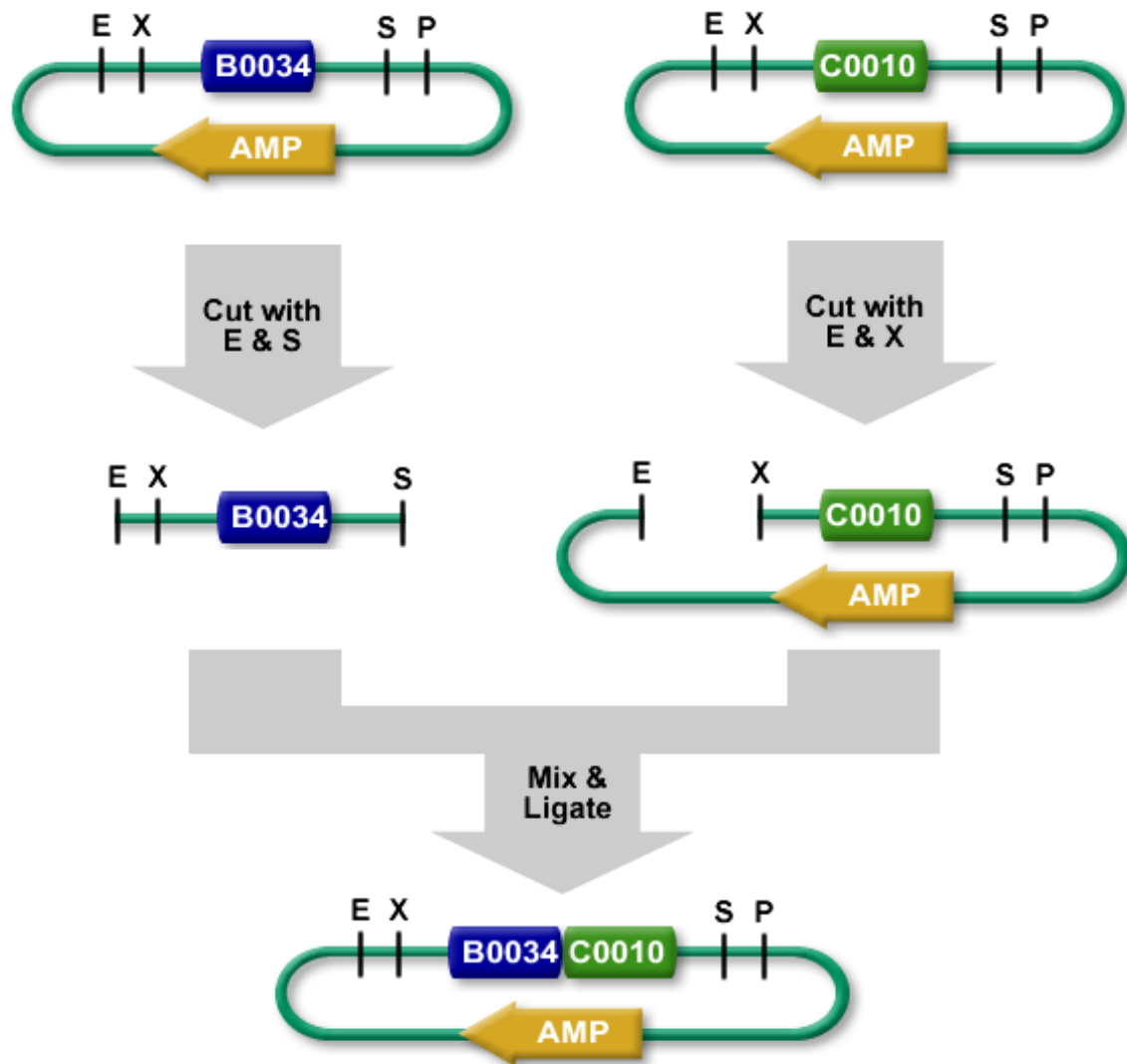
- ▶ Your boss has asked you to express the gene *mecA* in *E. coli* so that you can characterize the protein and told you that you can get a plasmid from a post doc in another lab.
- ▶ The post doc gives you a tube with the plasmid and a fasta file with the nucleotide sequence.

# CLONING

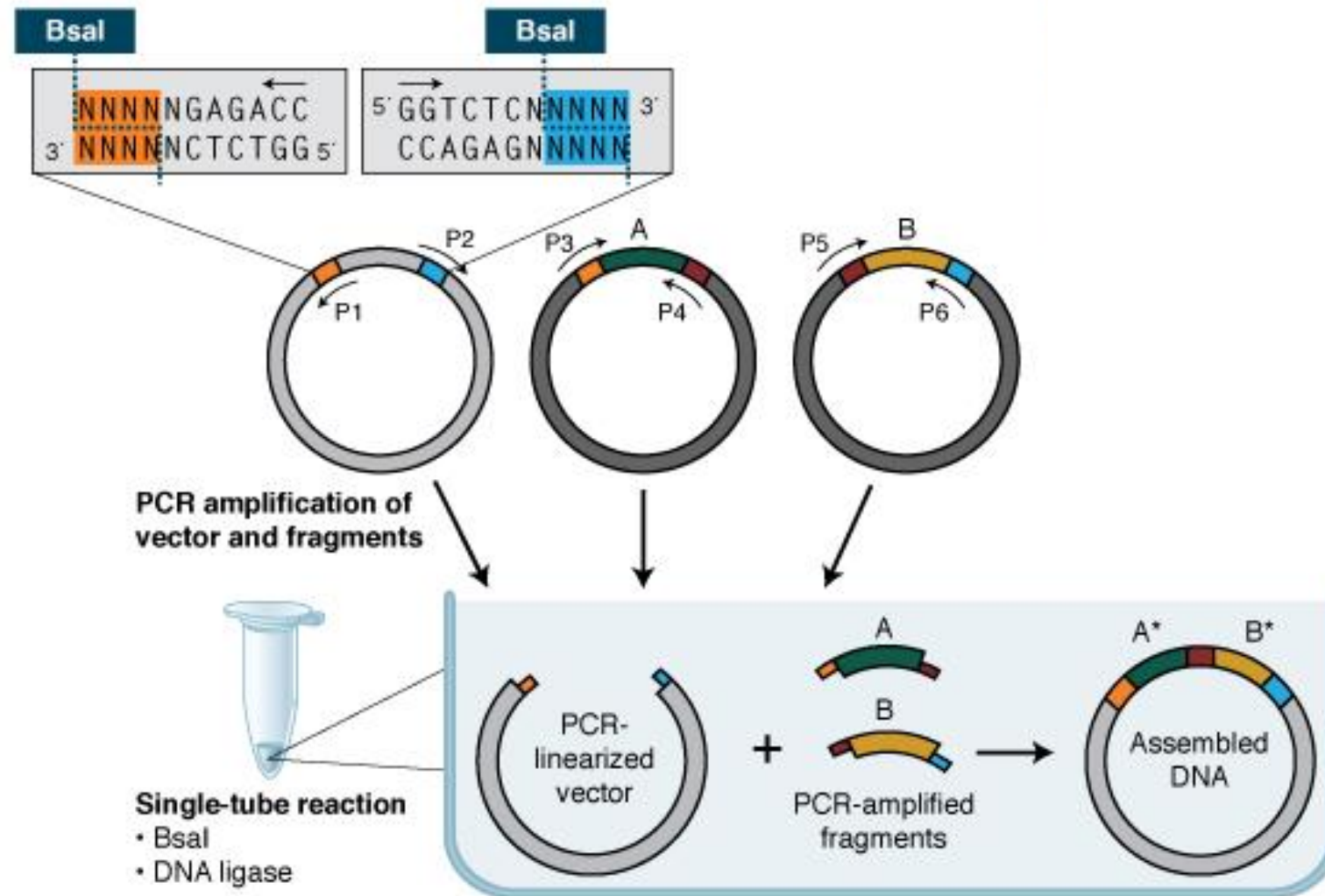


# BIOBRICK CLONING

E = EcoRI  
X = XbaI  
S = SpeI  
P = PstI

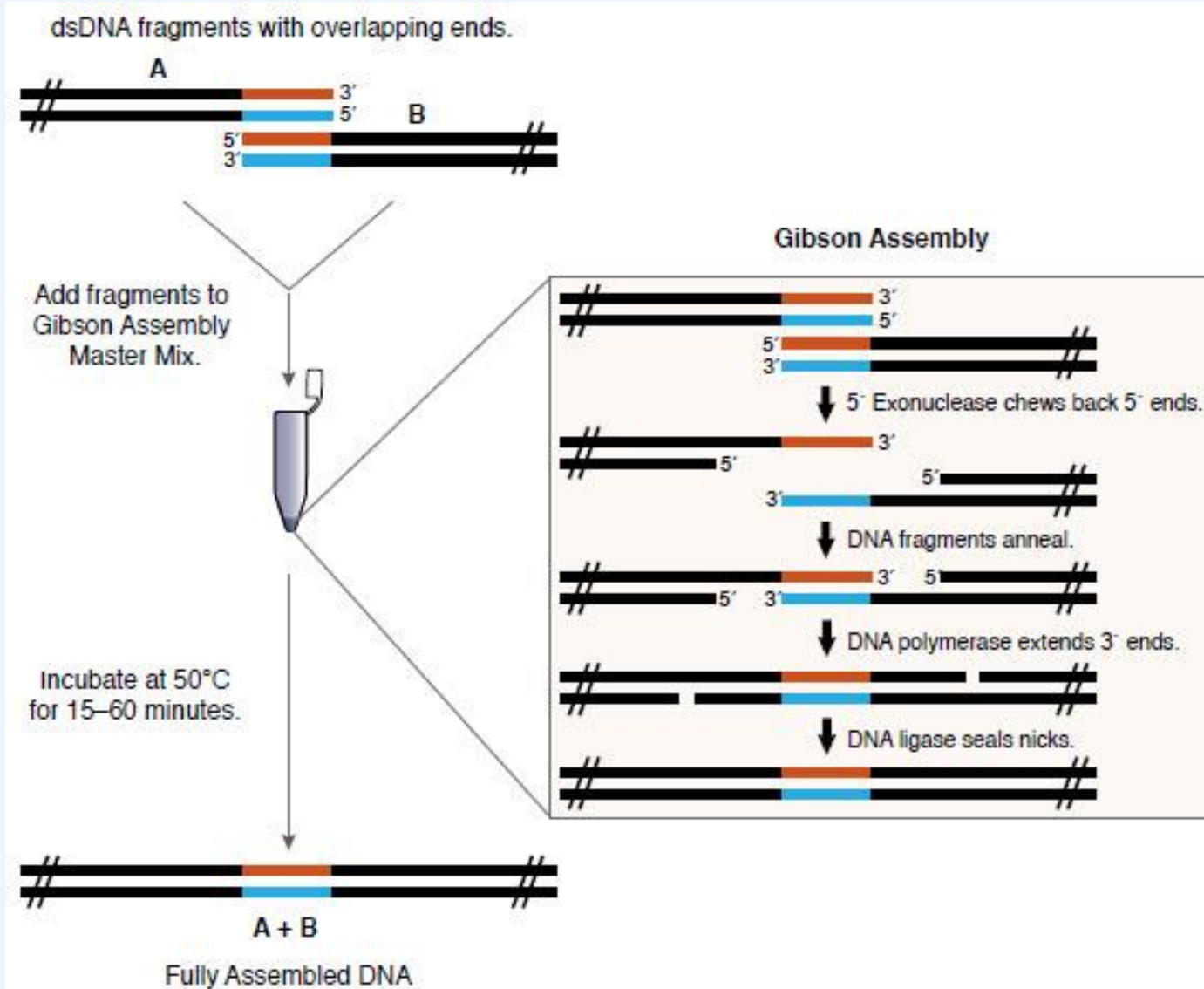


# GOLDEN GATE CLONING



\* While A and B insert sequences involved in 4-base overlaps are shown in separate colors for clarity, the actual assembly is seamless; 4-base overlaps are insert derived.

# GIBSON ASSEMBLY



# OBJECTIVES

- ▶ Reverse Engineer a plasmid
  - ▶ Identify Promoter, RBS, CDS, Terminator, Restriction sites
- ▶ Insert a new gene into plasmid
  - ▶ Use Registry of Biological Parts or biocyc to find new DNA sequences
  - ▶ Optimize expression with RBS calculator
  - ▶ Codon Optimize CDS sequence for expression in desired chassis

# CLONING SOFTWARE

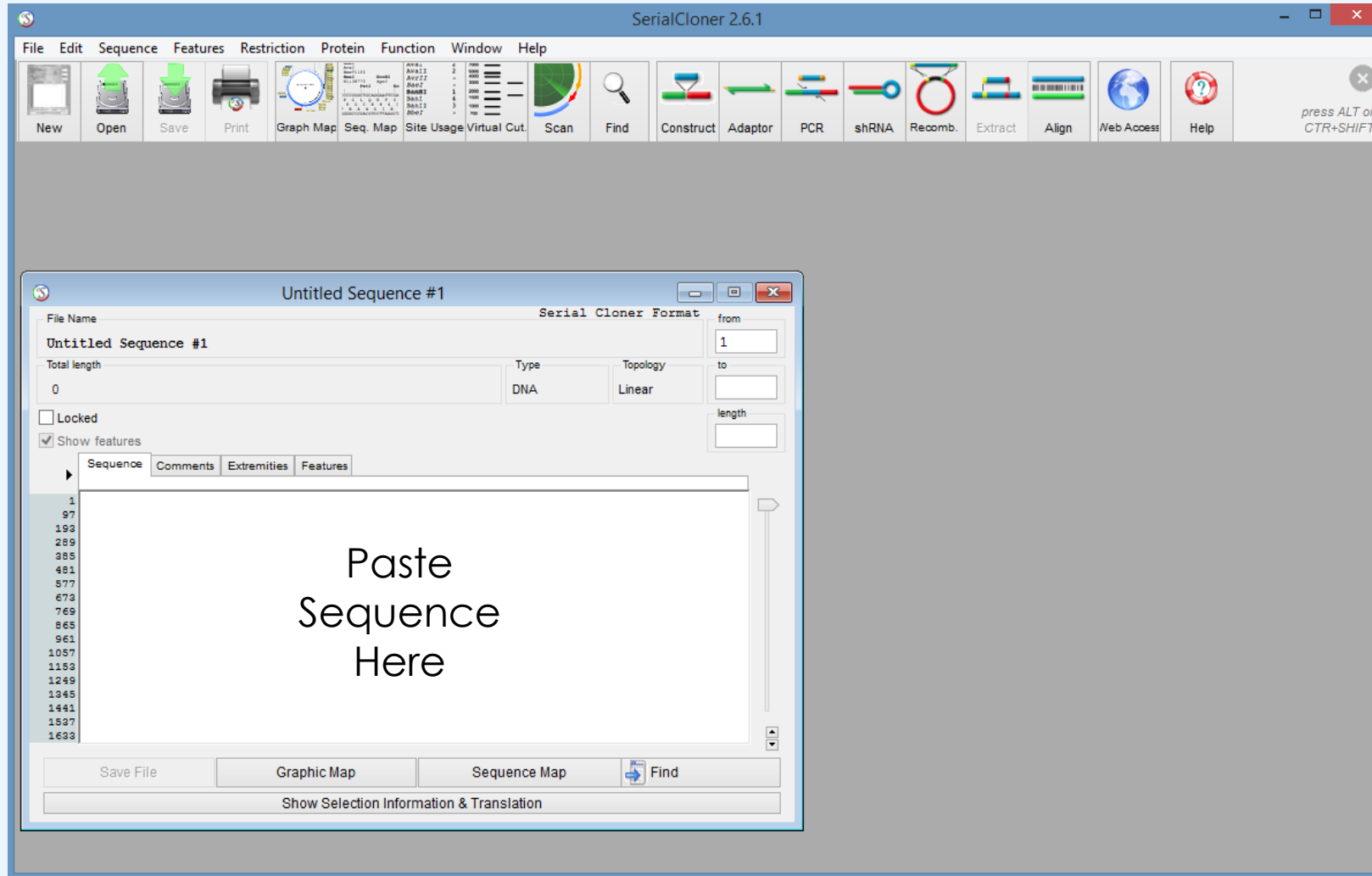
- ▶ Useful for Annotating DNA sequences, Primer Design, In silico Molecular biology (PCR, Restriction digestion, ligation)
- ▶ Free options
  - ▶ Serial Cloner, Benchling
- ▶ Paid options
  - ▶ Geneious, Clone Manager, Snap Gene



# NON-ANNOTATED DNA SEQUENCE

- ▶ Download sequence from [https://drive.google.com/file/d/0B2ciXonm2r\\_jWWZld09NMnB5ZVk/view?usp=sharing](https://drive.google.com/file/d/0B2ciXonm2r_jWWZld09NMnB5ZVk/view?usp=sharing)
- ▶ Load sequence into SerialCloner. Annotate the features of the sequence and determine what it does.

# SERIAL CLONER



# NCBI - BLAST

- ▶ First we identify the backbone
  - ▶ Go to the NCBI – Blast website  
(<http://blast.ncbi.nlm.nih.gov/Blast.cgi>)
  - ▶ NCBI Blast will search a Nucleotide or Protein database for matches to a given query sequence, or can align two sequences.
  - ▶ Click Nucleotide Blast
  - ▶ Paste DNA sequence and hit BLAST

BLAST® Basic Local Alignment Search Tool

Home Recent Results Saved Strategies Help

My NCBI [Sign In] [Register]

NCBI/BLAST Home

BLAST finds regions of similarity between biological sequences. [more...](#)

New DELTA-BLAST, a more sensitive protein-protein search [Go](#)

**BLAST Assembled Genomes**

Find Genomic BLAST pages:

Enter organism name or id-completions will be suggested [GO](#)

- [Human](#)
- [Mouse](#)
- [Rat](#)
- [Cow](#)
- [Pig](#)
- [Dog](#)
- [Rabbit](#)
- [Chimp](#)
- [Guinea pig](#)
- [Fruit fly](#)
- [Honey bee](#)
- [Chicken](#)
- [Zebrafish](#)
- [Clawed frog](#)
- [Arabidopsis](#)
- [Rice](#)
- [Yeast](#)
- [Microbes](#)

**Basic BLAST**

Choose a BLAST program to run.

<a href="#">nucleotide blast</a>	Search a nucleotide database using a nucleotide query <i>Algorithms: blastn, megablast, discontinuous megablast</i>
<a href="#">protein blast</a>	Search protein database using a protein query <i>Algorithms: blastp, psi-blast, phi-blast, delta-blast</i>
<a href="#">blastx</a>	Search protein database using a translated nucleotide query
<a href="#">tblastn</a>	Search translated nucleotide database using a protein query
<a href="#">tblastx</a>	Search translated nucleotide database using a translated nucleotide query

**Specialized BLAST**

Choose a type of specialized search (or database name in parentheses.)

- Make specific primers with [Primer-BLAST](#)
- Cluster multiple sequences together with their database neighbors using [MOLE-BLAST](#)
- Find [conserved domains](#) in your sequence (cds)
- Find sequences with similar [conserved domain architecture](#) (cdart)
- Search sequences that have [gene expression profiles](#) (GEO)

**Your Recent Results** [New](#)

[All Recent results...](#)

**News**

[PDB and Swiss-Prot on the FTP site](#)

The NCBI is now distributing the BLAST databases for protein PDB (pdbaa) and Swiss-Prot (swissprot) as stand-alone BLAST databases, rather than as subsets of the non-redundant (nr) database.

Tue, 25 Nov 2014 16:00:00 EST

[More BLAST news...](#)

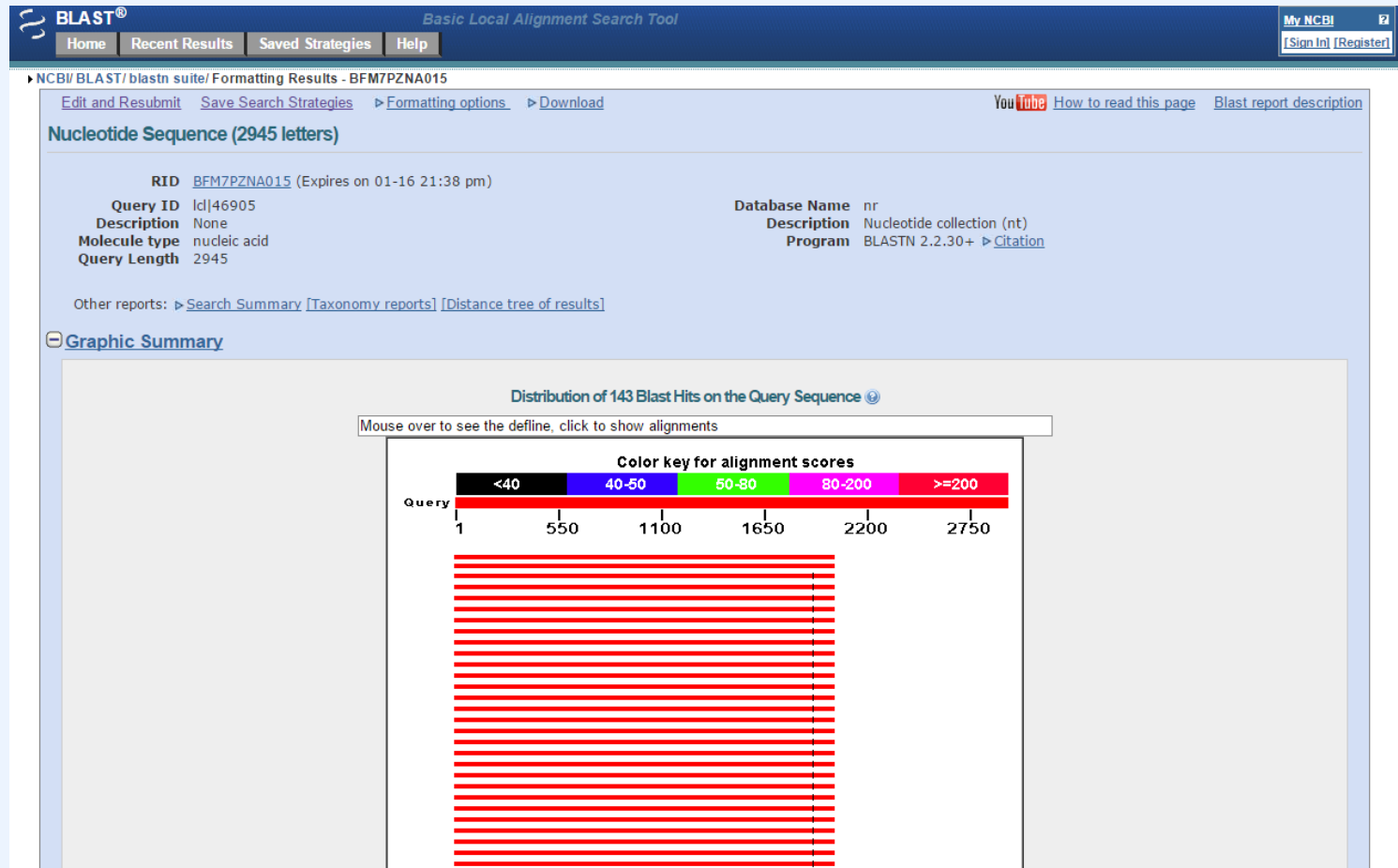
**Tip of the Day**

[How to save custom search pages.](#)

So you have made a few BLAST searches and after adjusting the database, organism limits and maybe a few Algorithm Parameters you arrive at what you think is a good search strategy.

[More tips...](#)

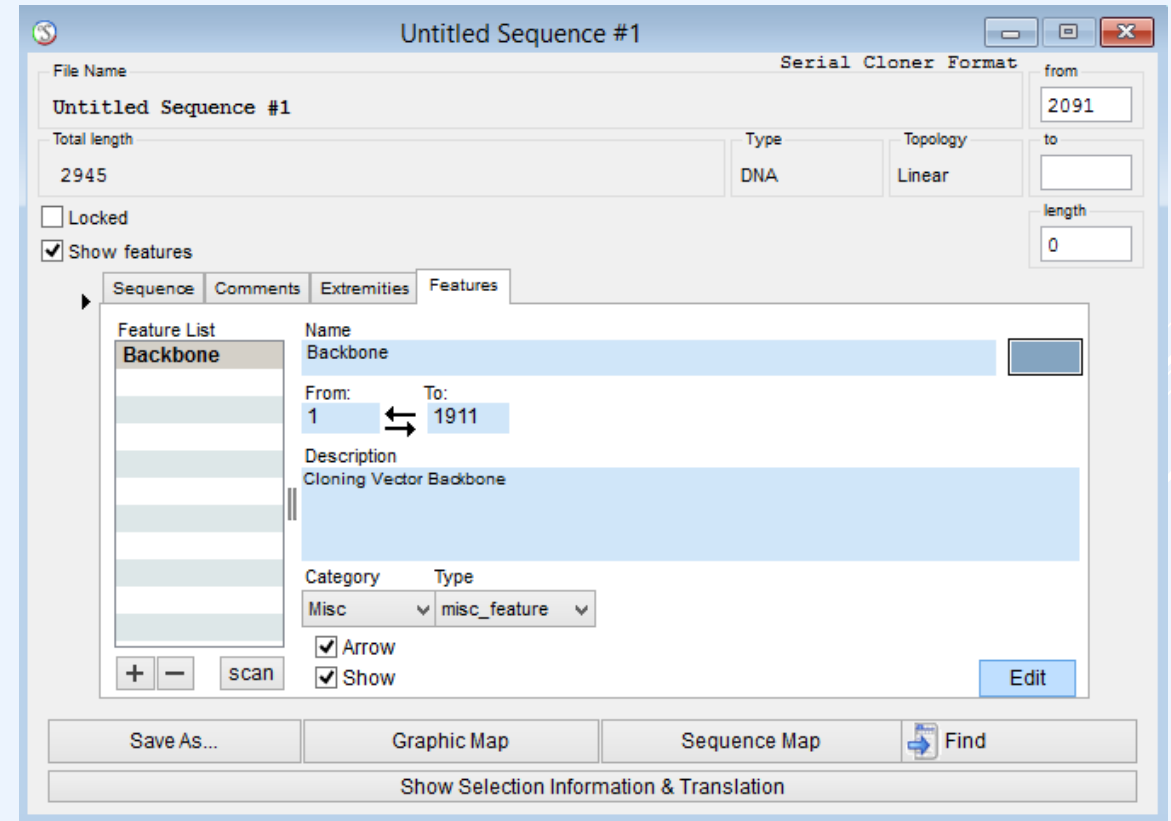
# IDENTIFYING THE VECTOR BACKBONE



- ▶ The red bars show us regions with high alignment scores for sequences in the nucleotide database. A list with all those sequences and alignments is below. As many of the hits for this region are for different cloning vectors, it is likely this is a common vector backbone.

# ANNOTATING THE BACKBONE

- ▶ In Serial Cloner
  - ▶ Click the Features tab of your DNA sequence
  - ▶ Click the + sign below the features list to add a new feature
  - ▶ Name the feature and annotate where it covers your DNA sequence using the alignment from NCBI – Blast.

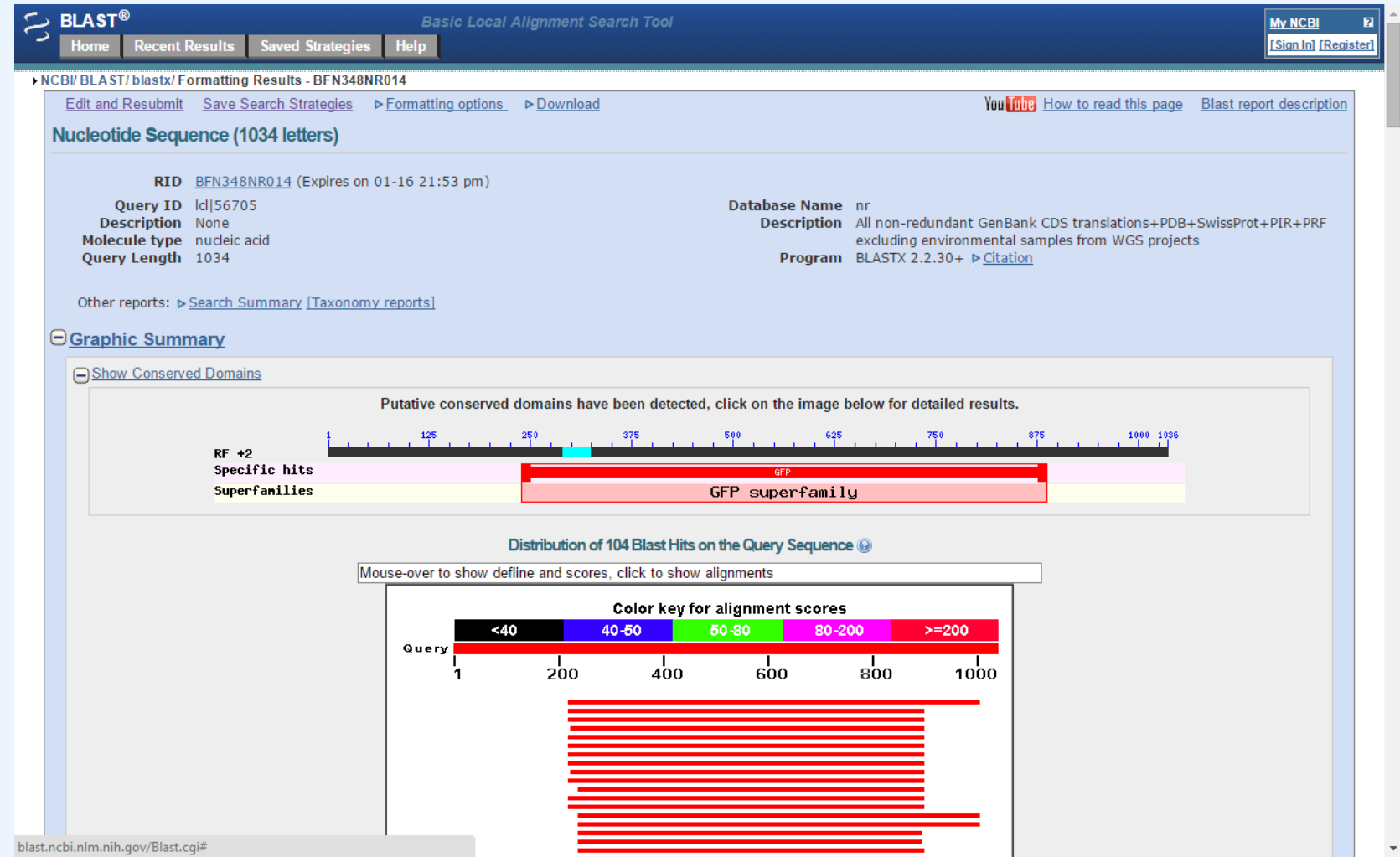


# ANNOTATING THE CDS

- ▶ Some software will predict ORF within a sequence for you, however you can also find the CDS with NCBI – Blast
- ▶ Select the remaining DNA sequence that was **not** part of the backbone
- ▶ Go to the NCBI – Blast website and this time select blastx
- ▶ BlastX will attempt to translate a nucleotide sequence to match to a protein database. This is useful due to the degeneracy in the genetic code

# BLASTX

- ▶ Blastx will give alignments similar to blastn, but will also tell you what protein super families it finds in a given sequence
- ▶ Blastx tells you that the sequence relates to the GFP family of proteins



# BLASTX ALIGNMENTS

- Scroll down to look at the specific alignments. The Blast results show that this is monomeric RFP. Use this alignment to annotate your sequence in Serial Cloner.

[Download](#) ▾ [GenPept](#) [Graphics](#)

▼ Next ▲ Previous ▲ Descriptions

monomeric red fluorescent protein [synthetic construct]  
Sequence ID: [gb|AAM54544.1|AF506027\\_1](#) Length: 225 Number of Matches: 1  
► [See 39 more title\(s\)](#)

Range 1: 1 to 225 [GenPept](#) [Graphics](#) ▼ Next Match ▲ Previous Match

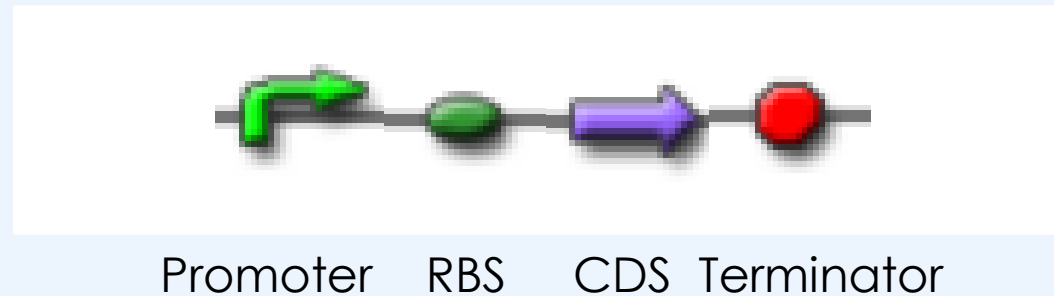
Score	Expect	Method	Identities	Positives	Gaps	Frame
436 bits(1122)	4e-151	Compositional matrix adjust.	225/225(100%)	225/225(100%)	0/225(0%)	+2
Query 221	MASSEDDVIKEFMRFKVRMEGSVNghefeiegegegRPYEGTQTAKLKVTKGGPLPFAWDI					400
	MASSEDDVIKEFMRFKVRMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDI					
Sbjct 1	MASSEDDVIKEFMRFKVRMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDI					60
Query 401	LSPQFQYGSKAYVKHPADIPDYLLKLSFPEGFKWERMNFEDGGVVTVTQDSSLQDGEFIY					580
	LSPQFQYGSKAYVKHPADIPDYLLKLSFPEGFKWERMNFEDGGVVTVTQDSSLQDGEFIY					
Sbjct 61	LSPQFQYGSKAYVKHPADIPDYLLKLSFPEGFKWERMNFEDGGVVTVTQDSSLQDGEFIY					120
Query 581	KVKL RGTNFP SDGPVMQKKT MGWEASTERMYPEDGALKGEIKMRLKLDGGHYDAEVKTT					760
	KVKL RGTNFP SDGPVMQKKT MGWEASTERMYPEDGALKGEIKMRLKLDGGHYDAEVKTT					
Sbjct 121	KVKL RGTNFP SDGPVMQKKT MGWEASTERMYPEDGALKGEIKMRLKLDGGHYDAEVKTT					180
Query 761	YMAKKPVQLPGAYKTDIKLDITSHNEDYTIVEQYERAEGRHSTGA					895
	YMAKKPVQLPGAYKTDIKLDITSHNEDYTIVEQYERAEGRHSTGA					
Sbjct 181	YMAKKPVQLPGAYKTDIKLDITSHNEDYTIVEQYERAEGRHSTGA					225

**Related Information**  
[Identical Proteins](#) - Proteins identical to the subject



# A TYPICAL GENETIC CASSETTE

- ▶ Most functional genetic units are made up of a Promoter, an RBS, a CDS, and finally a Terminator



- ▶ This can help you look for the other parts of the cassette

# IDENTIFYING THE PROMOTER AND TERMINATOR

- ▶ Promoter and terminator boundaries are often poorly characterised or annotated but there are still a few tricks to do
  - ▶ Blast the sequence upstream and downstream of the CDS to find the promoter
  - ▶ Check genbank entries of alignments to find annotations
  - ▶ Look for cloning sites/scars to determine the boundaries

[Display Settings:](#) ☒ GenBank [Send:](#) ☒

## Dual controller plasmid for Boolean Integrase Logic, complete sequence

GenBank: KC529324.1  
[FASTA](#) [Graphics](#)

---

[Go to:](#) ☐

LOCUS KC529324 75 bp DNA linear SYN 12-MAY-2013  
DEFINITION Dual controller plasmid for Boolean Integrase Logic, complete sequence.  
ACCESSION [KC529324](#) REGION: 2..76  
VERSION KC529324.1 GI:490341930  
KEYWORDS .  
SOURCE Dual controller plasmid for Boolean Integrase Logic  
ORGANISM [Dual controller plasmid for Boolean Integrase Logic](#)  
other sequences; artificial sequences; vectors.  
REFERENCE 1 (bases 1 to 75)  
AUTHORS Bonnet,J., Yin,P., Ortiz,M.E., Subsoontorn,P. and Endy,D.  
TITLE Amplifying genetic logic gates  
JOURNAL Science 340 (6132), 599-603 (2013)  
PUBMED [23539178](#)  
REFERENCE 2 (bases 1 to 75)  
AUTHORS Bonnet,J.  
TITLE Direct Submission  
JOURNAL Submitted (25-JAN-2013) Bioengineering, Stanford University, 473 Via Ortega, Stanford, CA 94305, USA  
FEATURES Location/Qualifiers  
source 1..75  
/organism="Dual controller plasmid for Boolean Integrase Logic"  
/mol\_type="other DNA"  
/db\_xref="taxon:[1330438](#)"  
misc\_feature <1..21  
/note="prefix Bba"  
regulatory 22..75  
/regulatory\_class="promoter"  
/note="pTetO"  
ORIGIN  
1 aattcgcggc cgcttctaga gtccctatca gtgatagaga ttgacatccc tatcagtgat  
61 agagatactg agcac  
//

# IDENTIFYING THE RBS

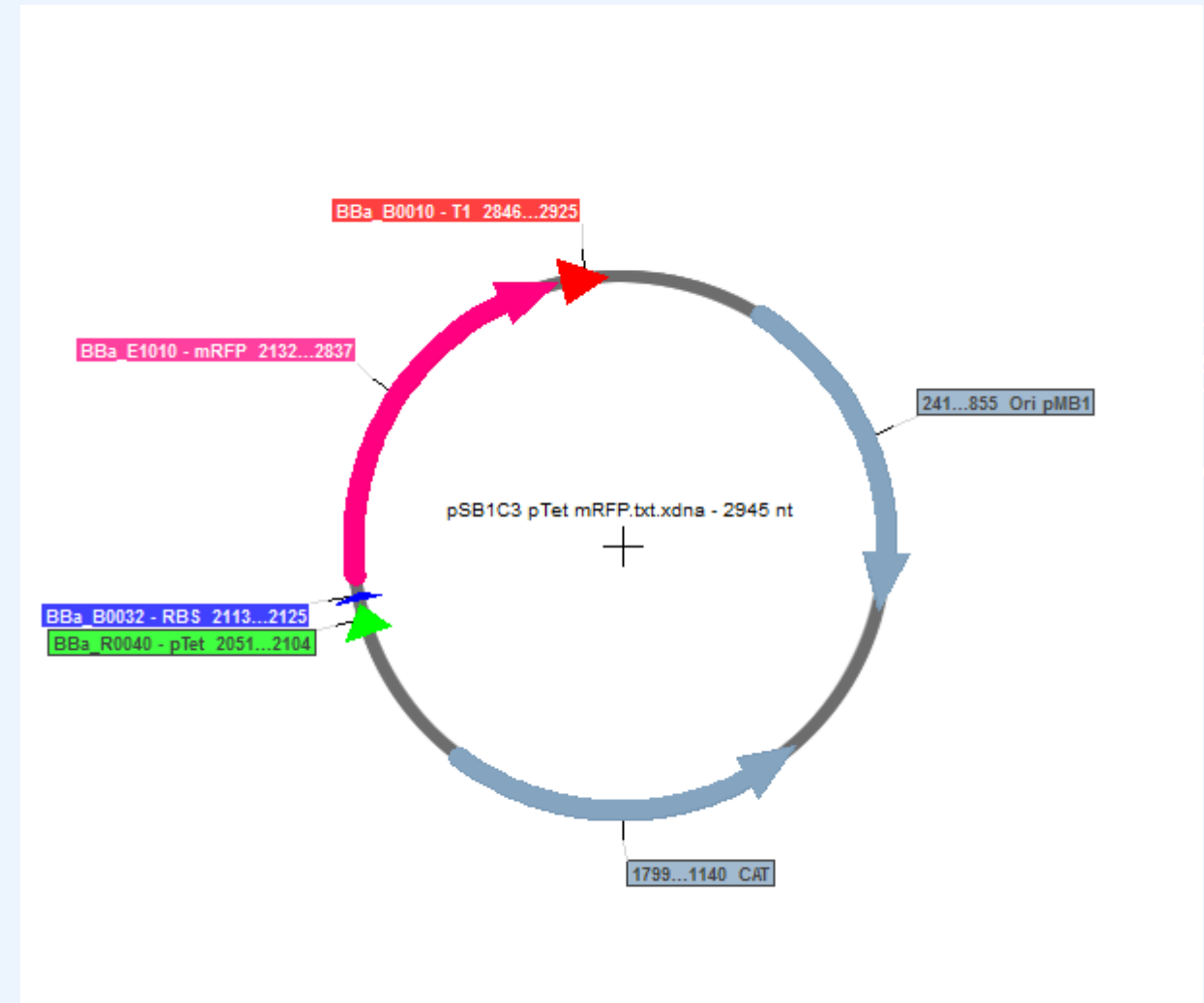
- ▶ The easiest part of the RBS to recognize is the Shine Dalgarno sequence.
  - ▶ The Consensus sequence in E. coli is AGGAGGU
  - ▶ Synthetic RBS will not necessarily follow this consensus
- ▶ If you fail to find a sequence similar to the Shine Dalgarno sequence between your promoter and CDS, search for Biobrick Scar sites or cloning sites which may indicate the RBS location.
- ▶ Compare to Synthetic RBS on [parts.igem.org](https://parts.igem.org)
- ▶ Reverse engineer mRNA with Salis RBS calculator at <https://salis.psu.edu>

# FINDING CLONING ARTIFACTS

- ▶ A biobrick scar is the sequence actaga and results from the ligation of a XbaI cut site with a SpeI cut site.
- ▶ Finding this sequence between the promoter, RBS, CDS, and terminator is a clear indication of the boundaries of each part.
- ▶ Additionally, Common Restriction sites can indicate restriction sites used to construct the Plasmid (EcoRI, BamHI, PstI, XbaI, HindIII, KpnI, NcoI, SacI, XhoI)
- ▶ Many modern cloning techniques (golden gate, gibson assembly) are called scarless cloning techniques because they avoid these artifacts, making them more difficult to reverse engineer.

# THE ANNOTATED SEQUENCE

- ▶ Serial cloner (and most forms of in silico cloning software) will allow you to visualize your annotated plasmids.
- ▶ Click on Graphic Map



# CREATING A NEW VECTOR

- ▶ Modify the vector provided to express *mecA* (which confers methicillin resistance in *Staphylococcus aureus*)

You will need the following tools:

- ▶ Biocyc / Kegg
- ▶ RBS calculator (<https://salis.psu.edu>)
- ▶ Codon optimization ([www.jcat.de/](http://www.jcat.de/))

# BIOCYC.ORG



Pathway Tools Tutorial Date  
Poll

[LOGIN](#) | [Why Login?](#) | [Create New Account](#)

Enter a gene, protein, metabolite or pathway...

[Quick Search](#)

[Gene Search](#)

Searching *Staphylococcus aureus aureus* N315 [change organism database](#)

[Sites](#) ▾ [Search](#) ▾ [Genome](#) ▾ [Metabolism](#) ▾ [Analysis](#) ▾ [SmartTables](#) ▾ [Help](#) ▾

## BioCyc Database Collection

BioCyc is a collection of 7667 Pathway/Genome Databases (PGDBs), plus software tools for understanding their data.

### Getting Started

New to BioCyc? Typical usage is:

- Select one or more databases (genomes) to search. To do so, click "change organism database" in the box in the top right of every page. By default, BioCyc searches *Escherichia coli* K-12 substr. MG1655.
- Search for a gene or pathway using the Quick Search, or see the Search menu for more options.

[New User Guide >>>](#)

### Tools

BioCyc provides tools for navigating, visualizing, and analyzing the

## FROM CARD CATALOG TO THE BOOK ON THE SHELF

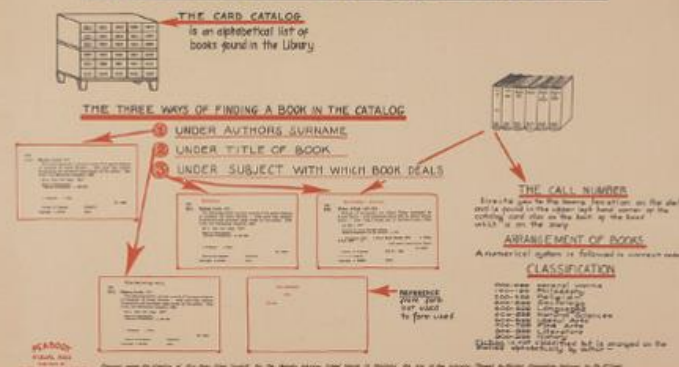


Image of Peabody Visual Aids poster by [char booth](#). Used under creative commons license.

### Search BioCyc by Locus ID


You can now enter locus IDs into the BioCyc quick search box. There is no need to first select the correct BioCyc organism database - if the organism is part of BioCyc, the unique locus ID will automatically find it. Some example locus IDs look like this: MSM\_0046, MXAN\_1061.

1 2 3 4 5 6 7 8 9 10 11 12

### BioCyc Databases

The BioCyc databases are divided into three tiers, based on their quality.

# BIOCYC.ORG



Pathway Tools Tutorial Date Poll

LOGIN | Why Login? | Create New Account

Enter a gene, protein, metabolite or pathway...  
Searching *Staphylococcus aureus aureus* N315 [change organism database](#)

Quick Search | Gene Search

Sites ▾ | Search ▾ | Genome ▾ | Metabolism ▾ | Analysis ▾ | SmartTables ▾ | Help ▾

gene  
**mecA**

polypeptide  
**penicillin binding protein 2 prime**

[log in to add to SmartTable.](#)

[Staphylococcus aureus aureus](#) N315

Accession IDs	GJCB-41 (Saur158879Cyc) SA0038	Length	2007 bp / 668 aa
		Map Position	[45,031 <- 47,037] (1.6 centisomes, 6°) on NC_002745 chromosome

View in Genome Browser

Summary

Operons

Show All

Molecular Weight of Polypeptide

76.102 kD (from nucleotide sequence)

Unification Links

Entrez	15925745
Entrez-gene	1122812
NCBI-Protein	NP_373278.1
UniProt	NP_373278.1

Report Errors or Provide Feedback

Page generated by Pathway Tools version 19.5 (software by SRI International) on Mon Jan 18, 2016, biocyc14.  
saur158879cyc version 19.0.

hide

OPERATIONS

Sequences

- Protein Sequence
- Nucleotide Sequence
- Save Nucleotide Sequence to file
- Save Protein Sequence to file


Comparison Operations

- Show this gene in another database
- Change organisms/databases for comparison operations
- Search for this gene in other databases
- Show orthologs (with operon diagrams) in multiple databases
- Align in Multi-Genome Browser
- Align gene nucleotide sequence with orthologs
- Align gene product amino acid sequence with orthologs



http://www.genome.jp/kegg/

# KEGG



KEGG   [Help](#)

[» Japanese](#)

---

### KEGG Home

- [Release notes](#)
- [Current statistics](#)
- [Plea from KEGG](#)

### KEGG Database

- [KEGG overview](#)
- [Searching KEGG](#)
- [KEGG mapping](#)
- [Color codes](#)

### KEGG Objects

- [Pathway maps](#)
- [Brite hierarchies](#)

### KEGG Software

- [KegTools](#)
- [KEGG API](#)
- [KGML](#)

### KEGG FTP

- [Subscription](#)

[GenomeNet](#)

[DBGET/LinkDB](#)

[Feedback](#)

[Kanehisa Labs](#)

## KEGG: Kyoto Encyclopedia of Genes and Genomes

KEGG is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies (See [Release notes](#) for new and updated features).

### New articles

- [KEGG as a reference resource for gene and protein annotation](#)
- [BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences](#)

### Main entry point to the KEGG web service

[KEGG2](#)      [KEGG Table of Contents](#)      [Update notes](#)

### Data-oriented entry points

<a href="#">KEGG PATHWAY</a>	<a href="#">KEGG pathway maps</a> [ <a href="#">Pathway list</a> ]
<a href="#">KEGG BRITE</a>	<a href="#">BRITE functional hierarchies</a> [ <a href="#">Brite list</a> ]
<a href="#">KEGG MODULE</a>	<a href="#">KEGG modules</a> [ <a href="#">Module list</a>   <a href="#">Statistics</a> ]
<a href="#">KEGG ORTHOLOGY</a>	<a href="#">Ortholog groups</a> [ <a href="#">KO system</a>   <a href="#">Annotation</a> ]
<a href="#">KEGG GENOME</a>	<a href="#">Genomes</a> [ <a href="#">KEGG organisms</a> ]
<a href="#">KEGG GENES</a>	<a href="#">Genes and proteins</a> [ <a href="#">Release history</a> ]
<a href="#">KEGG COMPOUND</a>	<a href="#">Small molecules</a> [ <a href="#">Compound classification</a> ]
<a href="#">KEGG GLYCAN</a>	<a href="#">Glycans</a> [ <a href="#">Monosaccharide codes</a> ]
<a href="#">KEGG REACTION</a>	<a href="#">Biochemical reactions</a> [ <a href="#">Reaction modules</a> ]
<a href="#">KEGG DISEASE</a>	<a href="#">Human diseases</a> [ <a href="#">Cancer</a>   <a href="#">Pathogen</a> ]
<a href="#">KEGG DRUG</a>	<a href="#">Drugs</a> [ <a href="#">ATC drug classification</a> ]
<a href="#">KEGG MEDICUS</a>	<a href="#">Health information resource</a> [ <a href="#">Drug labels search</a> ]

### Organism-specific entry points

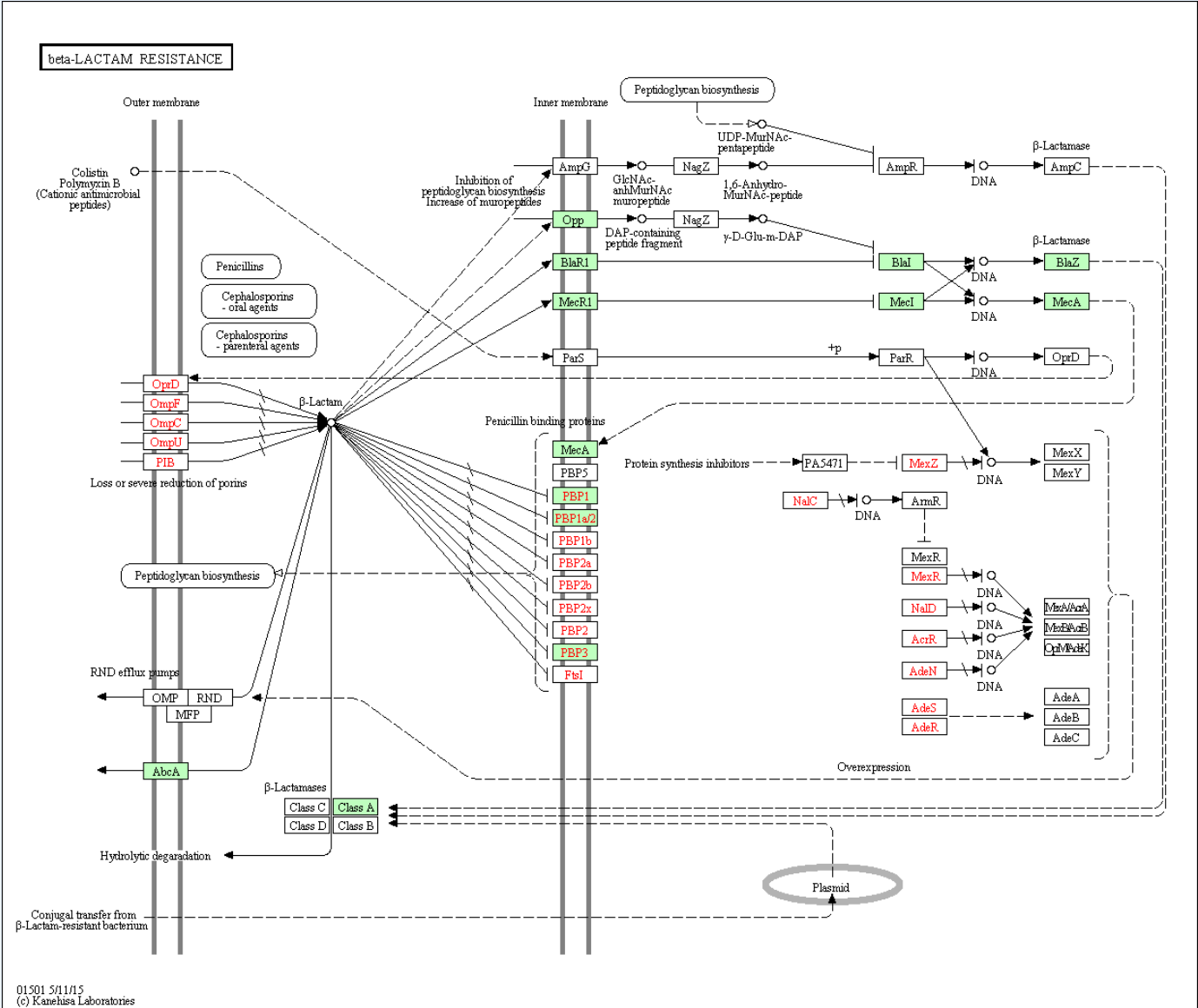
[KEGG Organisms](#)    Enter org code(s)      [hsa](#)    [hsa eco](#)

### Analysis tools

# KEGG

[illegible]

# KEGG



# THE REGISTRY OF BIOLOGICAL PARTS

The screenshot shows the top navigation bar of the Registry of Standard Biological Parts website. It includes links for 'tools', 'catalog', 'repository', 'assembly', 'protocols', 'help', 'search', a search input field with 'BBa' entered, and a 'login' button. Below the navigation bar is the title 'Registry of Standard Biological Parts'. The main content area is divided into two columns. The left column, titled 'Browse parts by type', lists various biological parts with icons and descriptions: Promoters, Ribosome Binding Site/about, Protein domains, Protein coding sequences, Translational units, Terminators, DNA, Plasmid backbones, Plasmids, Primers, and Composite parts. The right column, titled 'Categories', lists various categories with counts: biosafety, cds, chassis, classic, collections, direction, dna, function, legal, plasmid, plasmidbackbone, primer, promoter, proteindomain, rbs, regulation, ribosome, rnap, t3, terminator, test, test1, and viral\_vectors. At the bottom, there is a section titled 'Browse devices by type' with a note about developing new support for device specification and a link to 'Protein generators'.

tools catalog repository assembly protocols help search BBa login

## Registry of Standard Biological Parts

### Browse parts by type

Catalog	List
	<b>Promoters (?)</b> : A promoter is a DNA sequence that tends to recruit transcriptional machinery and lead to transcription of the downstream DNA sequence.
	<b>Ribosome Binding Site/about (?)</b> : A ribosome binding site (RBS) is an RNA sequence found in mRNA to which ribosomes can bind and initiate translation.
	<b>Protein domains (?)</b> : Protein domains are portions of proteins cloned in frame with other proteins domains to make up a protein coding sequence. Some protein domains might change the protein's location, alter its degradation rate, target the protein for cleavage, or enable it to be readily purified.
	<b>Protein coding sequences (?)</b> : Protein coding sequences encode the amino acid sequence of a particular protein. Note that some protein coding sequences only encode a protein domain or half a protein. Others encode a full-length protein from start codon to stop codon. Coding sequences for gene expression reporters such as LacZ and GFP are also included here.
	<b>Translational units (?)</b> : Translational units are composed of a ribosome binding site and a protein coding sequence. They begin at the site of translational initiation, the RBS, and end at the site of translational termination, the stop codon.
	<b>Terminators (?)</b> : A terminator is an RNA sequence that usually occurs at the end of a gene or operon mRNA and causes transcription to stop.
	<b>DNA (?)</b> : DNA parts provide functionality to the DNA itself. DNA parts include cloning sites, scars, primer binding sites, spacers, recombination sites, conjugative transfer elements, transposons, origami, and aptamers.
	<b>Plasmid backbones (?)</b> : A plasmid is a circular, double-stranded DNA molecules typically containing a few thousand base pairs that replicate within the cell independently of the chromosomal DNA. A plasmid backbone is defined as the plasmid sequence beginning with the BioBrick suffix, including the replication origin and antibiotic resistance marker, and ending with the BioBrick prefix.
	<b>Plasmids (?)</b> : A plasmid is a circular, double-stranded DNA molecules typically containing a few thousand base pairs that replicate within the cell independently of the chromosomal DNA. If you're looking for a plasmid or vector to propagate or assemble plasmid backbones, please see the set of <a href="#">plasmid backbones</a> . There are a few parts in the Registry that are only available as circular plasmids, not as parts in a plasmid backbone, you can find them here. Note that these plasmids largely do not conform to the BioBrick standard.
	<b>Primers (?)</b> : A primer is a short single-stranded DNA sequences used as a starting point for PCR amplification or sequencing. Although primers are not actually available via the Registry distribution, we include commonly used primer sequences here.
	<b>Composite parts (?)</b> : Composite parts are combinations of two or more BioBrick parts.

### Browse devices by type

We're in the process of developing new support for the specification of devices in the Registry. For the time being, please see the existing device tables below.

- Protein generators (?)**

- ▶ <http://parts.igem.org/>
- ▶ Good resource for finding DNA sequences for specific parts
- ▶ Characterization quality varies

# CODON OPTIMIZATION

Codon-Adaptation

---

1. Type/paste sequences below:

Standard genetic code is used for the input sequence. Click [here](#) to change!

---

2. Specify the pasted Sequence:

☐ DNA/RNA Sequence  
☐ Protein Sequence

---

3. Select organism:

----- Eukaryotes ----- ▼

---

4. Additional Options:

☐ Avoid rho-independent transcription terminators.  
☐ Avoid prokaryotic ribosome binding sites.  
☐ Avoid Cleavage Sites of Restriction Enzymes:

AatII  
AccI  
Acc65I  
AclI  
AfeI

☐ Only partly optimization in order to apply site directed mutagenesis.

---

Submit Reset

- ▶ <http://www.jcat.de/>
- ▶ Alternatives from Lifetechnologies, genscript, IDT and other synthesis companies
- ▶ Optimize expression for genes when transferring between species

# RBS CALCULATOR

- ▶ Translation rate depends on RBS
- ▶ RBS are not completely modular, the translation initiation rate will vary with the protein coding sequence
- ▶ Can design RBS for targeted translation initiation rate
- ▶ Can reverse engineer mRNA to determine relative translation initiation rate

## RBS Calculator<sub>v1.1</sub>

tunable control of the translation initiation rate

Title

Pre-Sequence [?]

Protein Coding Sequence [?]

Target Translation Initiation Rate [?]

Proportional scale (0 to 100,000+) ☐ Goal: Maximize

Organism or (16S rRNA) [?] (start typing)

Submit Job

For Non-Commercial Use Only. Click [here](#) for commercial usage.

Design Jobs: 4 queued, 8 currently running

UPDATES & TIPS

The next-generation RBS Calculator (v2.0) better predicts the translation initiation rates of mRNAs with long, highly structured 5' UTRs, structured protein coding sequences, and non-canonical Shine Dalgarno sequences. Read Espah Borujeni et. al., Nucleic Acid Research, 2013 for details.

# FINISHING UP YOUR NEW INSERT

- ▶ Add restriction sites (for Traditional, Biobrick, or Goldengate cloning) to your sequence to prepare it for synthesis. Add a few (3-5) bases on the ends of your sequence to ensure good binding of restriction enzymes.
- ▶ **or**
- ▶ Add homology regions for gibson assembly.
- ▶ The sequence can be synthesized by DNA synthesis companies such as IDT.

# PRIMERS FOR YOUR BACKBONE

Primers can be designed with primer3 ([bioinfo.ut.ee/primer3/](http://bioinfo.ut.ee/primer3/)) or using cloning software. You can use the sequences below for binding to your backbone.

- ▶ pSB1C3 pTet F = TACTAGAGCCAGGCATCAA
- ▶ pSB1C3 pTet R = AGTAGTGCTCAGTATCTCTATC

Add restriction sites (for Traditional, Biobrick, or Goldengate cloning) to your primers on the 5' end to prepare them for synthesis. Add a few (3-5) bases on the 5' of your sequence to ensure good binding of restriction enzymes.

**or**

Add homology regions for gibson assembly.



# IN SILICO CLONING

You can do in silico PCR by clicking on 'PCR' button in serial cloner.

Select your template (your now-annotated vector) and input the primers you designed.

You can do in silico cloning by clicking on 'Construct' button in serial cloner. Select your PCR product, and your new insert, and the restriction enzymes you want to use for the cloning.

Save your constructed sequence and check that it is what you expect.