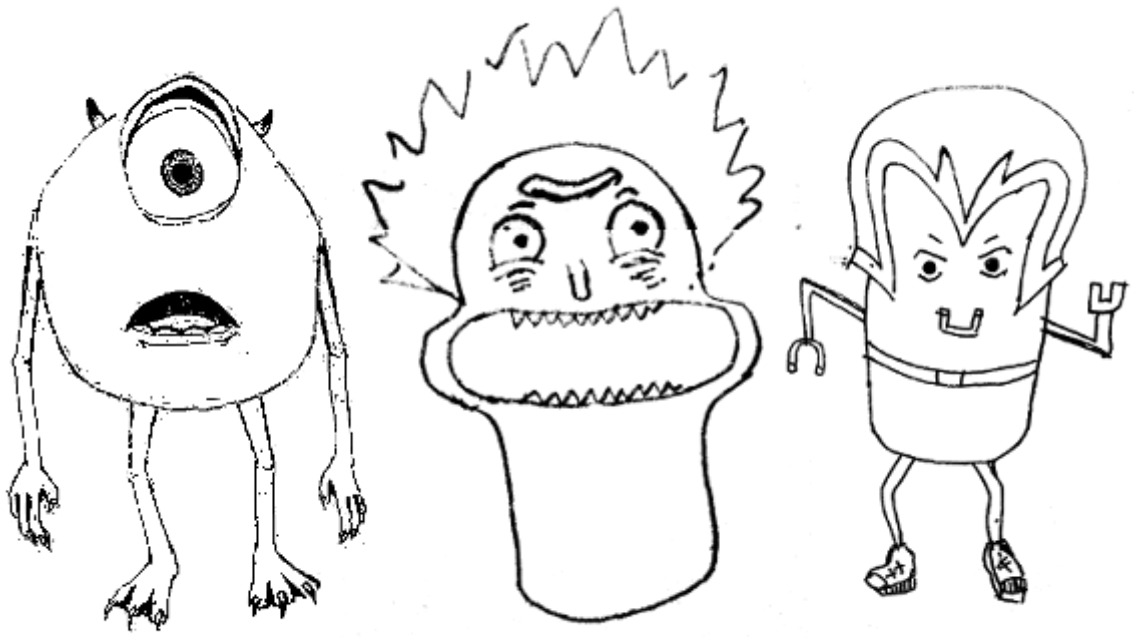


The iGEM-er's Guide to the Chassis(es)

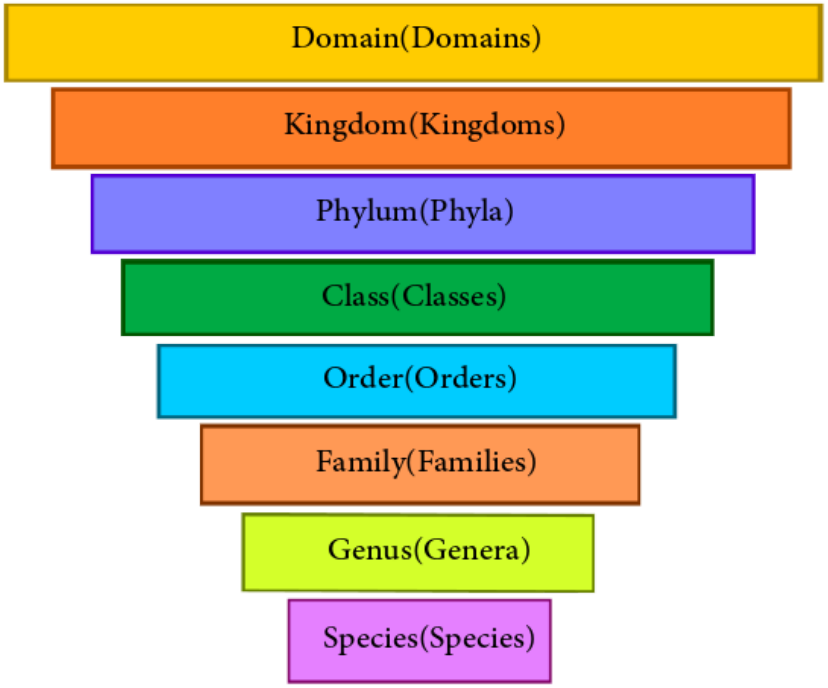
DON'T PANIC.

Index

Taxonomy

Taxonomy refers to the science of giving all organisms a scientific name, and is a part of an elaborate system of classification which gives each organism a unique name by which it can be identified in the scientific community, while also giving us information about how it is related to other organisms.

Animals are classified in groups which are subsets of each other, as follows:



The last two groups *Genus* and *Species* together form the organism's scientific name. Taking the example of the northern cardinal, a North American bird:

The scientific name of the bird is ***Cardinalis*** *(genus)* ***cardinalis*** *(species).*

When one starts working with an organism, it is important to find out all taxonomical details of it. Since we are working with it in the laboratory, checking specifications such as strain, which takes into account minor specifications, since we need to know the exact behaviour of the organism.

Hence, when you find the organism with the relevant biochemical processes for your experiment, finding out the taxonomy and strain before ordering is important.

Biosafety Levels

Biological Safety Levels or Biosafety levels are protocols which ensure biocontainment of infectious agents i. e. they are precautionary measures to ensure safety while working with possible biohazards.

There are four biosafety levels for the handling of infectious agents and laboratory animals, in increasing order of protection required by the personnel, the environment, and the community.

https://www.cdc.gov/biosafety/publications/bmbl5/bmbl5_sect_iv.pdf

The information about the biosafety level of your organism will be available while purchasing it from any collection centre; if brought from a laboratory, this information can be got from the. Regardless, before working with an organism, it is important to know its biosafety level, to check if the lab is equipped to implement the required protocols as detailed in the manual above; and ensure that at all times,all safety measures are in place and being followed.

Biosafety Level Table: (As an example)

Organism	Biosafety Level
Bacillus thuringiensis	1
Bacillus subtilis	1
Bacillus Megaterium	1
Pseudomonas aeruginosa	2
Pseudomonas putida	1
Magnetospirillum magneticum	1
Haloferax volcanii	1
Azotobacter vinelandii	1
Escherichia coli	1

Growth Characteristics

While working with an organism, in addition to attending to the biochemical pathway which will return the desired product, we need to ensure healthy growth and care of the organism. Firstly, noting whether the organism is an obligate aerobe, or anaerobe (requires oxygen or is poisoned by oxygen), or a facultative anaerobe (can survive with or without oxygen), we can determine the oxygenation requirements of the medium. Catering to its nutrient requirements for survival and noting the organism's growth, the functioning of all essential pathways and overall health allows to select the optimal conditions for the growth of the organism, which by extension allows most efficient functioning of the pathway in question.

The growth curve of the organism with different compositions of nutrient media, with and without the relevant substrate, is plotted, and conclusions are drawn as to what media composition is most suitable to the experiment. Existing literature can be referred to, for an idea of where to start; allowing us to draw guidelines based on preexisting work.

An example of a growth rate assay for E coli, under different experimental conditions, from http://parts.igem.org/Part:BBa_K1225000 .

Growth Rate Assay

BASIC INFORMATION

Purpose: To assess what effect, if any, our genetic parts have on the growth rate of E.coli.
Chassis: E.coli
Strain: BL21
Protocols: Purdue iGEM Growth Curve Protocol
Date: 09/25/13

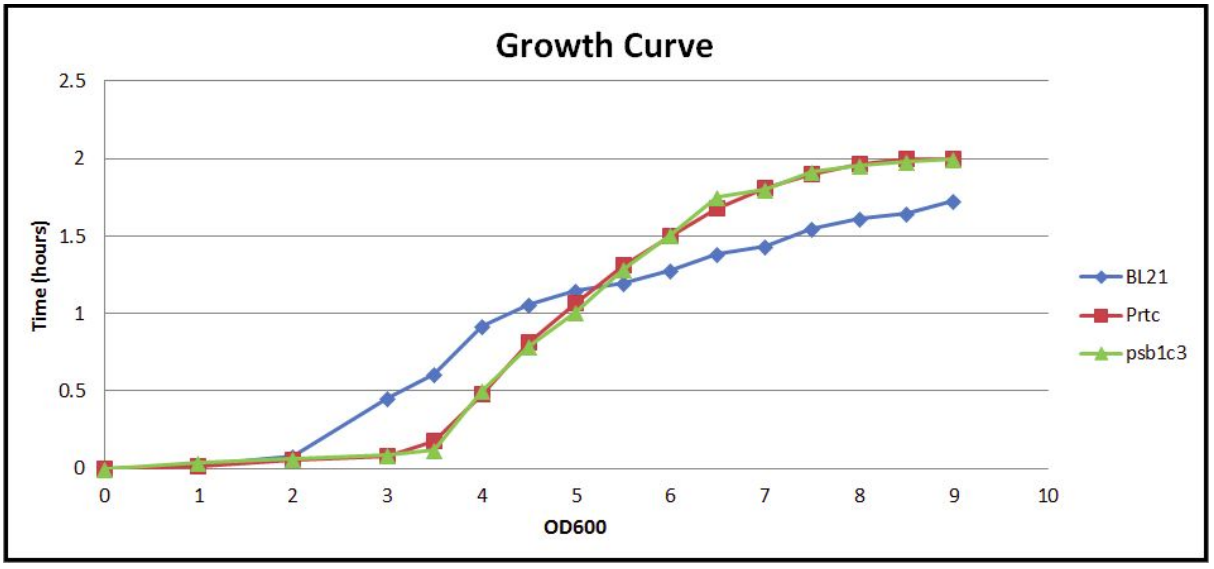
GROWTH CONDITIONS

Media Type: Luria Broth (LB)
Vessel: 10mL Culture Tube
Volume: 5mL
Incubation: 37 C, 250 rpm

MEASUREMENT INFORMATION

Data Type: Growth Curve (OD vs Time)
Location: Bindley Bioscience Center
Machine Name: N/A
Time Interval: 30min
Total Time: 270min
Notes: The growth curve on the right contains three curves: control BL21 cells, BL21 cells with just pSB1C3, and BL21 cells with pSB1C3 with the prtc* promoter inserted. We

believe that the strain of producing the antibiotic resistance delayed the pSB1C3 and prtc* cultures from entering exponential phase at the same time the control BL21 cells did. Because the control cells entered the phase quicker, they also reached the stationary phase faster. Overall, we determined that the prtc* promoter has no significant effect on the growth of BL21 cells that would be problematic for future assays.



Culture Sources

Primarily, there are two ways of obtaining a culture of the organisms you wish to work with – by ordering it from a culture centre, or by approaching a laboratory which is working with it. The third option of isolating the organism directly from natural sources is not recommended as – aside from the difficulty in isolation itself – characterization of the organism to find out which strain it is and what mutations it possesses is a tedious process.

To order from a culture centre, the information you will be required to provide varies from one culture centre to another – including a letter from the Head of Department or any other authority testifying to your requirement for the organism. All the details of the strain will be available on the website. Obtaining the required organism from the laboratory may be more tedious, as transportation of the organism may be difficult; additionally, the laboratory will have to provide all the information about the strain. On the plus side, however, they can provide experience-based tips about using the strain.

Culture Centre	
ATCC	Human, animal, insect, fish and stem cell lines and numerous bacterial, fungus and virus (animal and plants) strains

MTCC	Around 12000 microbial strains available
ECACC-Sigma Aldrich	Cell line of over 45 species, numerous fungus, bacterial and virus strains, HpiSci,
DGRC	Over 159 cell lines representing Drosophila melanogaster and closely related species
DSMZ	Numerous cultures of microorganisms, plant cell cultures, plant viruses andhuman and animal cell lines
CLS- CELL LINE SERVICE	Human and animal cell lines
IPD-ESTDAB	Melanoma Cell lines
NCCS	Cultures of archaea, bacteria and fungi
Genscript	GPCR & ion channel stable cell lines, Fc receptor cell lines, immune checkpoint stable cell lines
JCRB cell bank	General cells, Luciferase-expressing cancer cells, Mouse homozygous mutant ES cells, Immortalized Mesenchymal stem cells, Genetically-modified cells
Riken cell bank	General cell lines, human and animal iPS cells, animal ES cell lines, cord blood stem cells, etc
BGSC	Bacillus strains
NCTC	Bacterial cultures
Coriell Institute	Fibroblasts, endothelial, smooth muscle, and epithelial cell cultures from biopsies, EBV-transformed lymphoblast lines
Flinn Scientific	Protozoa and microbial cells

MTA or Material Transfer Agreements are contracts drawn up to regulate the acquisition of biological and research materials or occasionally, data. Industry may want to assert ownership of any inventions made with those materials, viewing their materials as property, or restrict publication of unfavorable results. Universities may not allow changes in their materials to be published as research results, and may want full dissemination of research results. When acquiring biological material like cells from a laboratory, it may be necessary to inform the culture centre from which it was originally acquired. In any case, it is important to follow MTAs for any transaction that takes place. An example of an MTA is given below:

https://www.autm.net/resources-surveys/material-transfer-agreements/mta-guiding-principles/

Maintenance protocols

After initially culturing the organism, the organism needs to be maintained for many generations. Therefore, protocol for subculturing the organism (different methods of plating - slants, agar plates, stabs, liquid media etc), for storing the organism (glycerol

stocks, lyophilization etc) needs to be noted, to allow multiple experiments to be carried out over time.

Transformation protocols

While carrying out experiments with an organism, very often, you will be expressing modified genes which are not already present in the organism, even while using pathways already present in the organism. To introduce these modified genes in the organism, various transformation processes can be used; transformation is the process of introducing foreign DNA into the host organism.

Methods for genetic Transformation of cells: (Reference: *www.omicsonline.org* **)**

Technique	Procedure	Most important parameters involved	Advantages	Drawbacks
Electroporation	DNA is inserted through pores due to permeabilization of the cell membrane induced by strong electrical pulses.	Pulse length, energy and duration of the electrical field, extent and duration of membrane permeation, mode and duration of molecular flow, DNA concentration, tolerance of cells to membrane permeation.	Simple, fast, low cost.	Low efficiency, requires laborious protocols, and transforms mainly protoplasts.
Biolistics	High density carrier particles covered with genes are accelerated through the cells leaving the DNA inside by an adsorption mechanism.	Kinetic energy of the bombarding particles, temperature, the amount of cells, their ability to regenerate,susceptibility of the tissue, the number of DNA-coated particles, as well as the amount of DNA that covers each particle.	Simple, no need to treat the cell wall, allows transformation of different cells, independent of the physiological properties of the cell, allows the use of multiple transgenes.	High cost, low efficiency Transformation parameters must be optimized to each biological target employed, there is a risk of multiple copies of the introduced genes, DNA and cells can be damage.

Agitation with glass beads	Rapid agitation with glass beads allows the penetration of the plasmid DNA.	DNA and concentration, sensitivity of cells to membrane permeation, amount of cells and their ability to regenerate.	Fast, simple, low cost. Does not need sophisticated devices, chemical treatments or enzymatic cocktails.	Low efficiency because DNA get damaged.
Vacuum infiltration	Vacuum application generates a negative pressure that increases inter-cell spaces allowing the infiltration of *Agrobacterium*.	Duration and intensity of the vacuum, temperature, pH and time of induction of virulence genes.	Simple, fast, medium efficiency, with low somaclonal variation and many independent cells transformed.	Some strains of *Agrobacterium* are unable to infect certain cell types, risk of multiple copies of the introduced genes.
Silicon carbide whisker	Silicon carbide fibers are mixed in a vortex with a suspension of tissue and DNA allowing introduction by abrasion.	Fiber size, vortex parameters (type, duration and speed of agitation), vessel shape, thickness of the cell wall and cell's ability to regenerate.	Simple, fast, low cost and can be used indifferent cell types.	Very low efficiency. Cells can be damaged affecting regeneration capabilities. Could be hazardous to technicians due to fibers' inhalation.
Laser microbeams	A laser microbeam punctures self-healing	Laser characteristics to be used as optical tweezers coupled to	Allows precise and gentle treatment of cells, subcellular	High cost (expensive equipment

	holes into the cell wall allowing DNA penetration.	the appropriate microscope.	structures, and even individual DNA molecules.	required), and laborious.
Ultrasound	Introduces DNA molecules into cells via acoustic cavitation that temporarily changes the permeability of the cell membrane.	Intensity, exposure time, central frequency, type of application (continuous or pulsed), pulse repetition frequency, and duty cycle.	High efficiency, medium cost and can be used in different cell's types.	May damage the cells by breaking their membrane.
Shock waves	Cell permeabilization occurs due to shock wave–induced cavitation.	Frequency, energy, voltage, shock wave profile and number of shock waves.	Fast, easy to perform, reproducible with high efficiency, no need of enzymatic cocktails, can be used to transform several cell types.	Shock wave generators for this purpose are not on the market yet and experimental devices are relatively expensive.

Vectors

One of these transformation processes, is vector–mediated transfer. A vector carries the foreign DNA into the cell. It is a strand of DNA such as a plasmid which contains the required modified gene (foreign DNA) which is taken up by the host organism. Depending on a number of factors, such as whether the organism is a prokaryote or a eukaryote, a number of vectors exist. These vectors can be purchased from companies such as Thermo

Fisher, NEB etc. Each vector contains other genes which express traits which will allow us to select the transformed cells, such as antibiotic resistance, and blue-white selection. While selecting a vector to work with, close attention has to be paid to the selection technique to isolate transformed cells, to check if it is compatible with the given strain.

Some vectors which can be used are as follows:

1. Plasmid vectors:
    a. Plasmids are extrachromosomal, self-replicating, usually circular, double-stranded DNA found naturally in bacteria and some eukaryotes.
    b. Plasmid vectors contain an origin of replication (ORI) for individual replication, antibiotic resistance markers or other selection markers(lacZ gene for blue-white selection) and one or more restriction enzyme sites, allowing insertion of foreign DNA.
    c. Plasmid vectors for E. coli include pBR322, pUC19. Shuttle vectors are vectors that function in two or more organisms, succh as yeast plasmid, which functions in yeast and E. coli. They have two ORIs and sets of selection markers, one to function in each organism,
2. M13 Bacteriophage:
    a. M13 is a filamentous phage that infects E. coli containing the F pilus.
    b. It is a single stranded molecule with a genome of 6407 bp, which becomes double-stranded during replication.
    c. Can be used to clone single-stranded DNA, such as in site-directed mutagenesis.
    d. Due to its small genome, it can be easily manipulated; the double-stranded replicative form can be easily manipulated and purified like a plasmid.
3. Bacteriophage lambda:
    a. Linear, double stranded DNA, 48,514 bp long, with 12 unpaired bases on each end.
    b. These unpaired ends are called cohesive sites or cos sites; are uselful for packaging the DNA into phage head.
    c. Large portion of the central DNA genome is not required for infection, and thus can be removed and replaced with foreign DNA.
4. Cosmids:
    a. Vector combining certain features of plasmids and cos sites from Bacteriophage lambda.
    b. Simple cosmids contain an ORI, restriction sites, selection markers, and lambda cos sites.
5. YAC vectors:
    a. Yeast Artificial Chromosomes or YAC is used to clone large genomes such as those in the Human Genome Project.
    b. Contain a telomeric repeating sequence, the centromere and an autonomously repeating sequence.
    c. Also contain restriction sites and and selectable marker genes.
6. BAC vectors:
    a. Bacterial Artificial Chromosomes or BAC, are based on the natural extrachromosomal plasmid of yeast – the fertility or F plasmid.

b. They contain genes for the maintenance and replication of the F plasmid, a selectable marker and restriction enzyme sites.

Vector type	Insert size (kb)
Plasmid Bacteriophage lambda Cosmid YAC BAC	0.5–8 9–23 30–40 50–500 250–1000

Miscellaneous

RBS calculator: (Salis et al):

The Ribosome Binding Site (RBS) Calculator is a design method for predicting and controlling translation initiation and protein expression in bacteria. This calculator is based on a thermodynamics model. The method can predict the rate of translation initiation for every start codon in an mRNA transcript. The method may also optimize a synthetic RBS sequence to achieve a targeted translation initiation rate. Using the RBS Calculator, a protein coding sequence's translation rate may be rationally controlled across a 'given' fold range.

Link: https://salislab.net/software/

Uses:

- Optimization of synthetic metabolic pathways
- Designing and connecting genetic circuits.
- A robust calculator (Evolutionary context)
- Predicting translation initiation rates across a genome

Genome Mining 101:
Evolution of Genome Mining : (The phylogenetic tree is pretty informative)
http://pubs.rsc.org/en/content/articlehtml/2016/np/c6np00025h

Processing raw information: The experimentally determined sequence (raw information) is
processed using bioinformatics tools into genes, the proteins encoded and their function, the
regulatory sequences, and inferring phylogenetic relationships.

Genes: Gene prediction can be done by using computer programs like GeneMark for bacterial genomes and GENSCAN for eukaryotes.

Proteins: Protein sequences can be inferred from the predicted genes by using simple computer
Programs.

A regulatory sequence is a segment of a nucleic acid molecule which is capable of increasing or decreasing the expression of specific genes within an organism. Regulation of gene expression is an essential feature of all living organisms and viruses.Regulatory sequences can also be identified and analysed by using bioinformatics tools.
Information regarding the relationships between organisms can be obtained by aligning multiple sequences, calculating evolutionary distance and constructing phylogenetic trees. All of this is possible through genome mining.

Transcriptomics and its use:

The transcriptome is the complete set of transcripts in a cell, and their quantity, for a specific developmental stage or physiological condition. Understanding the transcriptome is essential for interpreting the functional elements of the genome and revealing the molecular constituents of cells. The key aims of transcriptomics are: to catalogue all species of transcript, including mRNAs, non-coding RNAs and small RNAs; to determine the transcriptional structure of genes, in terms of their start sites, 5' and 3' ends, splicing patterns and other post-transcriptional modifications; and to quantify the changing expression levels of each transcript during development and under different conditions. Transcriptomics data is fairly easy, reproducible, can assay thousands of endpoints at once, tells a snapshot of a specific time period in the cell/tissue/organism. GO/KEGG is advancing so there is more identity than there used to be, very sensitive, also specific. Using differential transcriptomics-guided genome-wide computational analysis of relevant promoters one can realize strong and weak promoters in an organism's genome.

A broad description of how one can put transcriptomics data to use can be looked up in the paper cited .(Cheng, J. K. and Alper, H. S. (2016) 'Transcriptomics-Guided Design of

Synthetic Promoters for a Mammalian System', ACS Synthetic Biology, 5(12), pp. 1455–1465. doi: 10.1021/acssynbio.6b00075.)

The usual pipeline used to obtain information about any regulatory element say a promoter is; obtaining the gene expression data during a given set of conditions and then processing and analyzing that data with a statistical model. This is then followed by identifying genes from different expression sets and annotating transcription factor binding sites upstream of the transcribed regions.
Then one attempts to build a synthetic promoter based on the transcription site binding frequencies and referring back to the annotated data. Then we start testing the designed promoters.