

Clotho 3.0 Use Cases



Nicholas Roehner and Prashant Vaidyanathan
“Clotho 3.0 Launch Party”
February 26th, 2015

Rafik B. Hariri Institute for Computing and Computational Science & Engineering



Problem Statement

The representation of design data in synthetic biology is spread across

- Legacy file formats that are not suited to the purpose of engineering.
- Generic file formats that are difficult to reason over using machines.

Problem Statement

The representation of design data in synthetic biology is spread across

- Legacy file formats that are not suited to the purpose of engineering.
- Generic file formats that are difficult to reason over using machines.

These facts make it difficult to reuse application code and data across projects.

Clotho 3.0 Schemas

Clotho 3.0 comes with schemas for biological designs and other data in order to

- Provide a basic object model that developers can build from as needed.
- Enable developers to easily persist object-oriented representations of their data.
- Aid in the creation of converters between different data formats.

Use Cases

- Sequence Editing and Optimization
- Modeling and Experiment Planning
- Statistical Design of Experiments
- Project Documentation and Networking

Use Cases

- Sequence Editing and Optimization: Merlin
- Modeling and Experiment Planning
- Statistical Design of Experiments
- Project Documentation and Networking

Use Cases

- Sequence Editing and Optimization: Merlin
- Modeling and Experiment Planning: Phoenix
- Statistical Design of Experiments
- Project Documentation and Networking

Use Cases

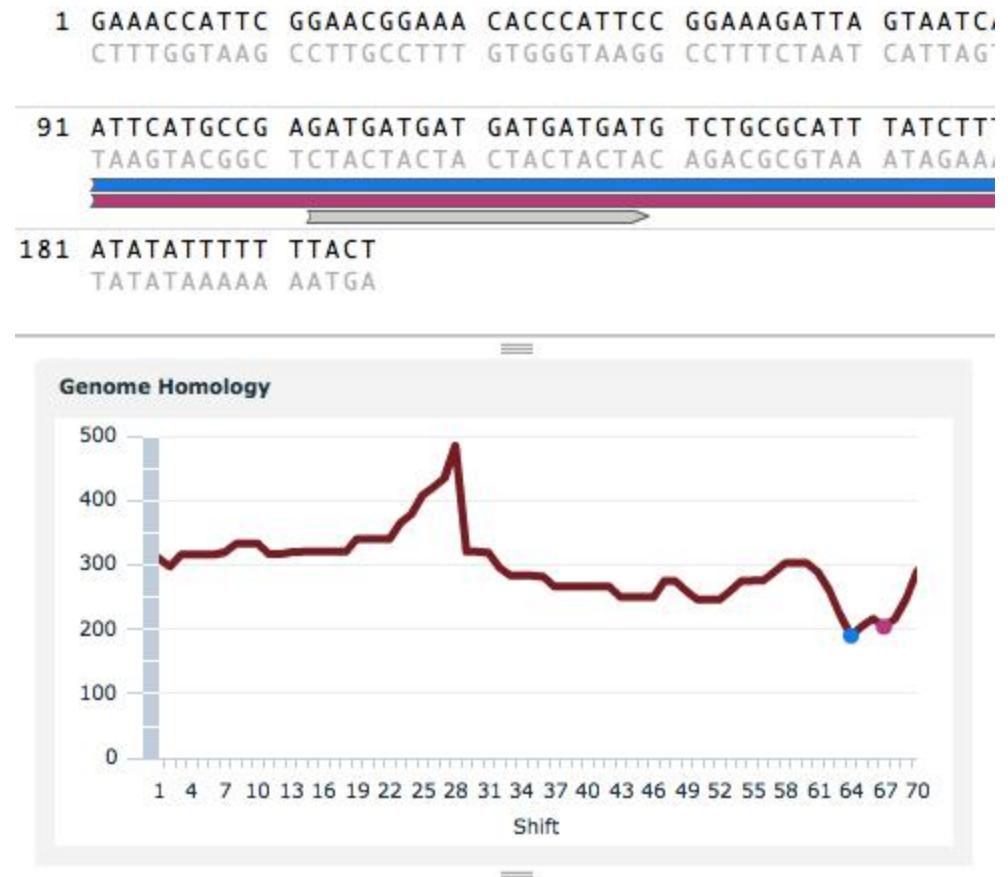
- Sequence Editing and Optimization: Merlin
- Modeling and Experiment Planning: Phoenix
- Statistical Design of Experiments:
Double Dutch
- Project Documentation and Networking

Use Cases

- Sequence Editing and Optimization: Merlin
- Modeling and Experiment Planning: Phoenix
- Statistical Design of Experiments:
Double Dutch
- Project Documentation and Networking:
Phagebook

Merlin

- Sequence viewing
- Oligo optimization
- Genome-to-oligo
- Oligo-to-oligo



GenBank

The screenshot shows the NCBI GenBank Overview page. The search bar contains "eat-4 elegans" and the "Go" button is highlighted in red. The main content area displays information about GenBank, mentioning its status as the NIH genetic sequence database and its size.

NCBI

SITE MAP

Submit to GenBank

BankIt

Sequin

GenBank Overview

PubMed Entrez BLAST OMIM Books Taxonomy Structure

Search for **eat-4 elegans** Go

What is GenBank?

GenBank® is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences ([Nucleic Acids Research 2004 Jan 1;32\(1\):23-6](#)). There are approximately 37,893,844,733 bases in 32,549,400 sequence records as of February 2004 (see [GenBank growth statistics](#)). As an example, you may view the [record](#) for a *Saccharomyces cerevisiae* gene. The complete [release notes](#) for the current version of GenBank are available. A news release is made available.

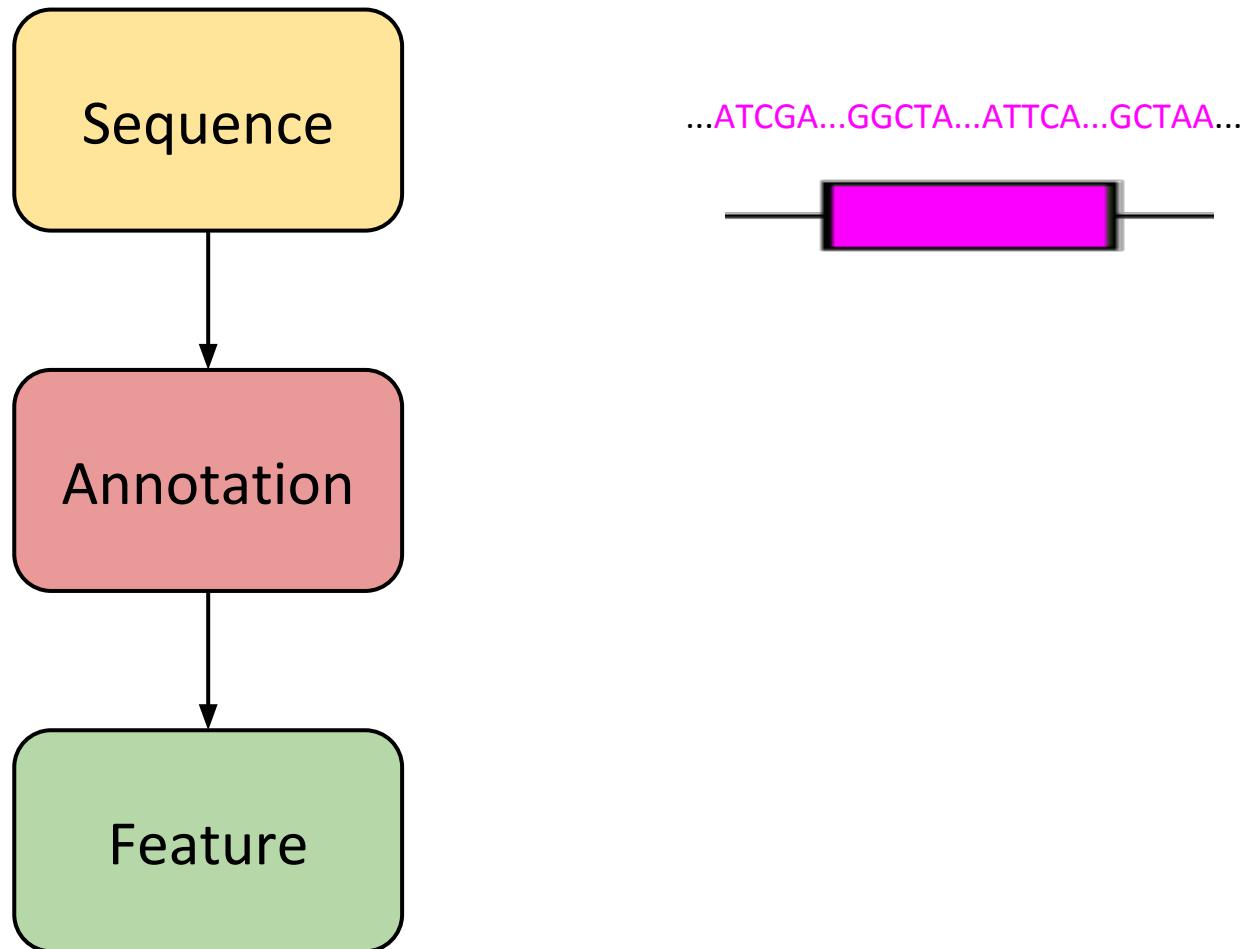
Representing Sequences

Sequence

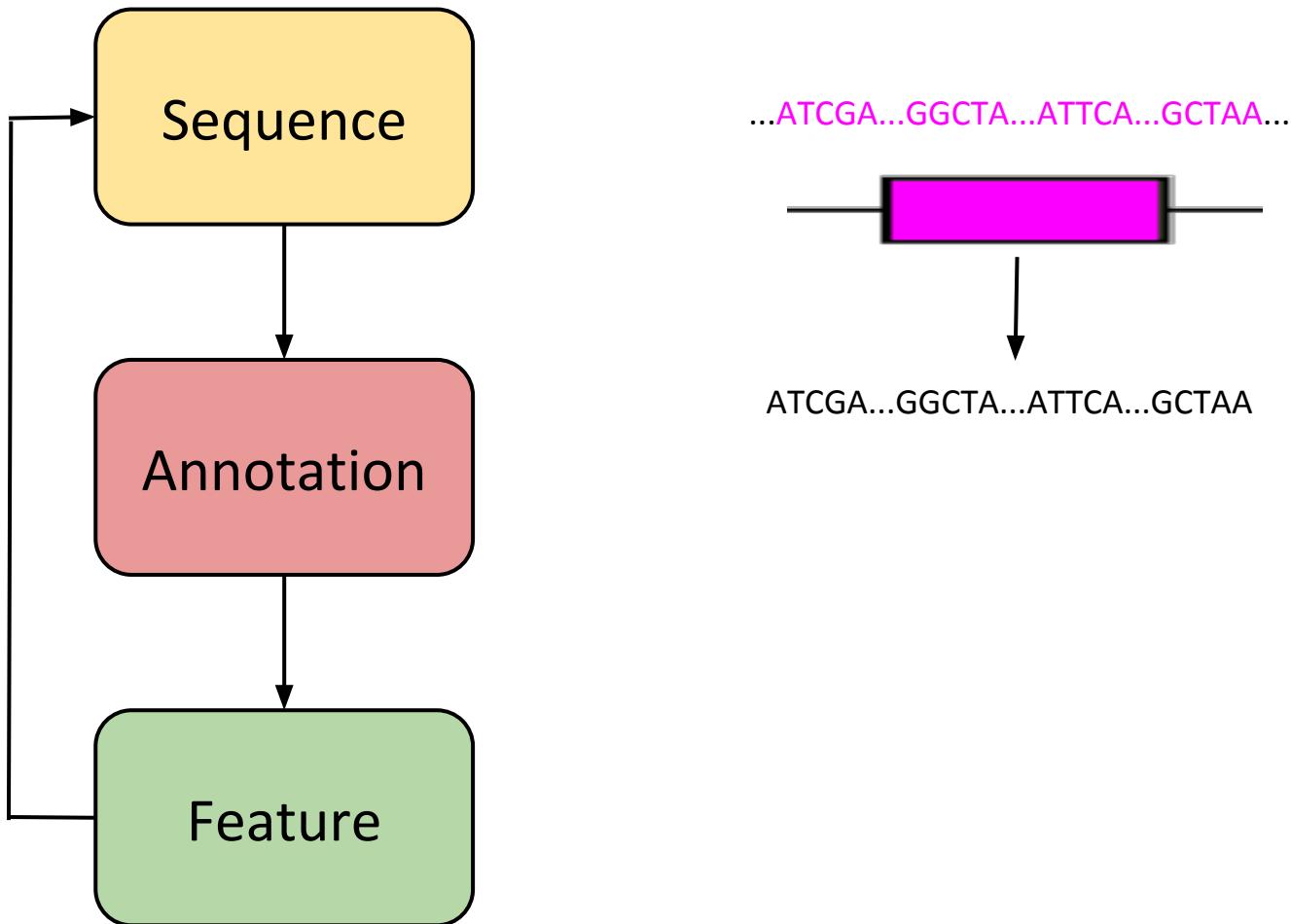
...ATCGA...GGCTA...ATTCA...GCTAA...



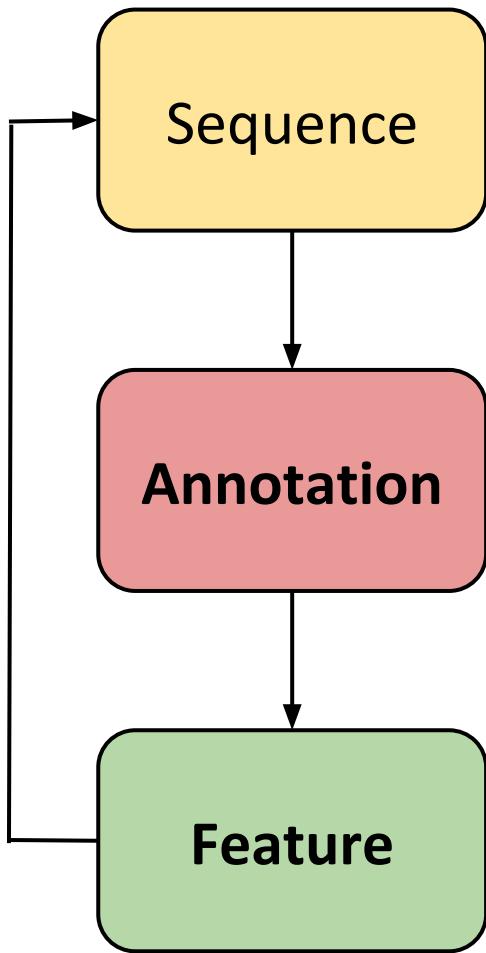
Annotating with Features



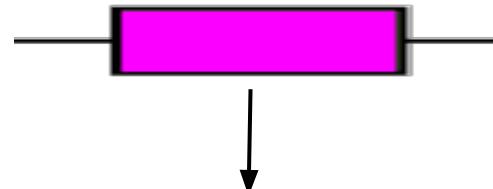
Hierarchical, Modular Annotation



Hierarchical, Modular Annotation



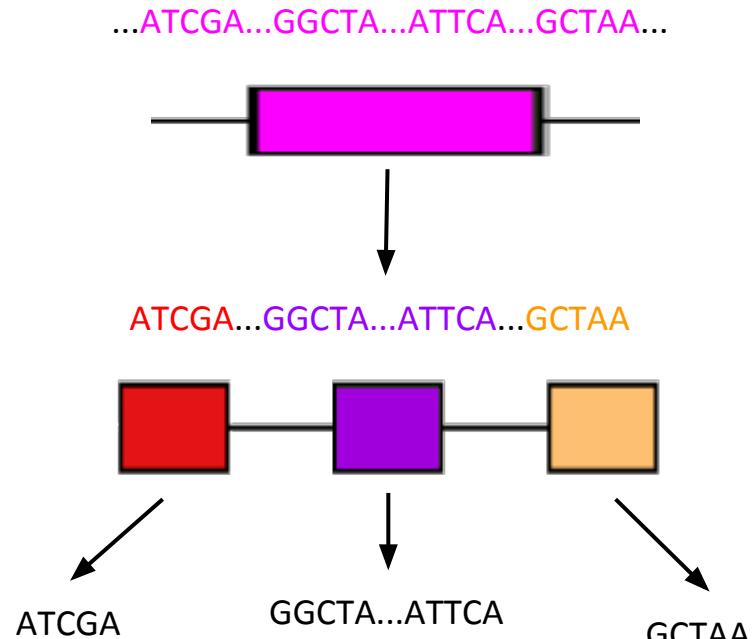
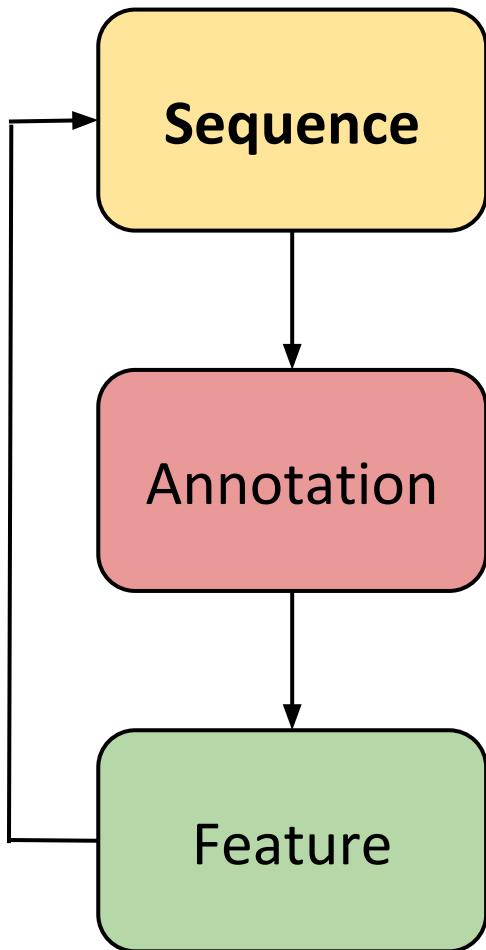
...ATCGA...GGCTA...ATTCA...GCTAA...



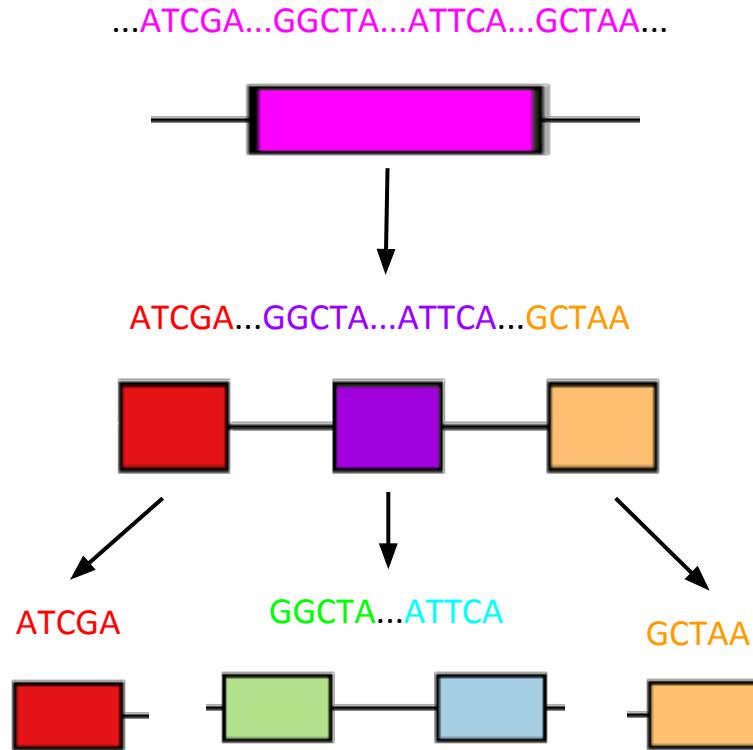
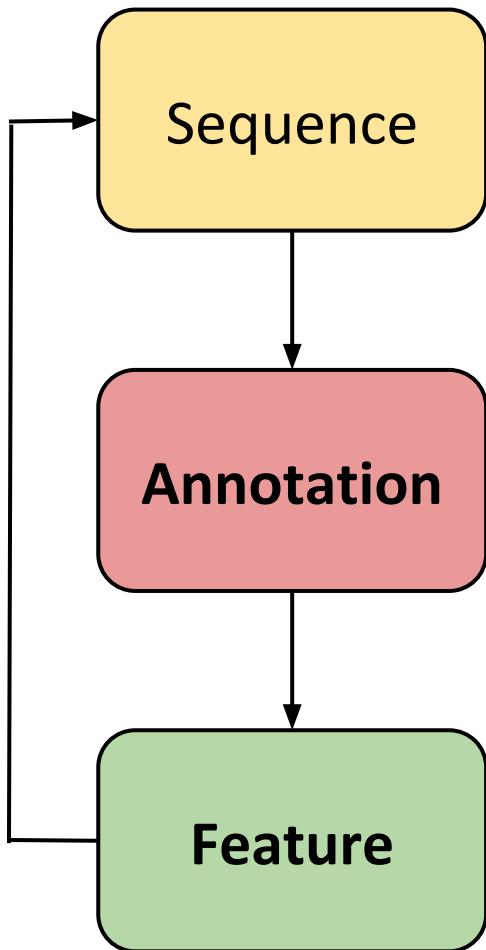
ATCGA...GGCTA...ATTCA...GCTAA



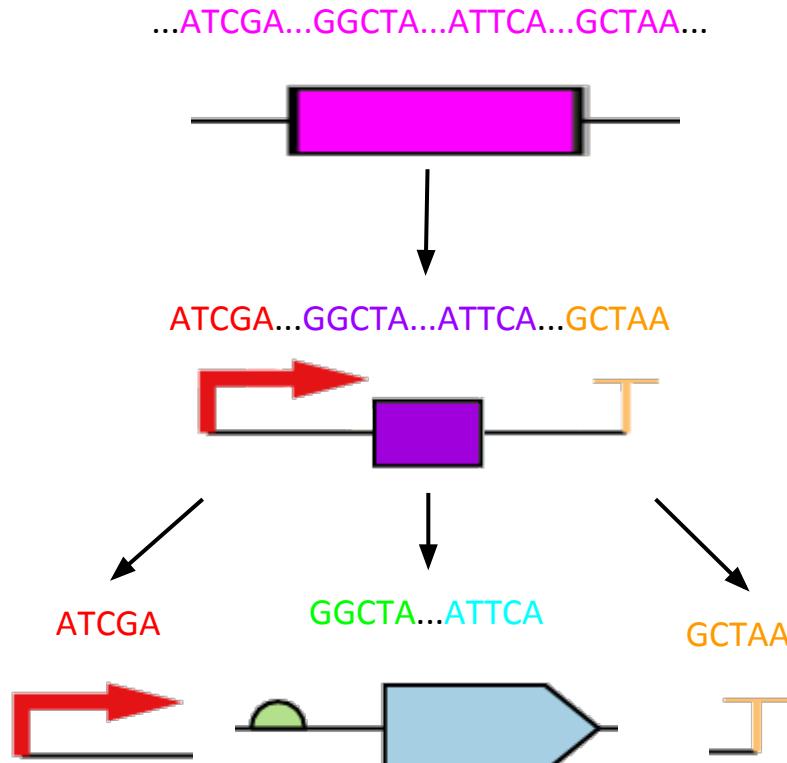
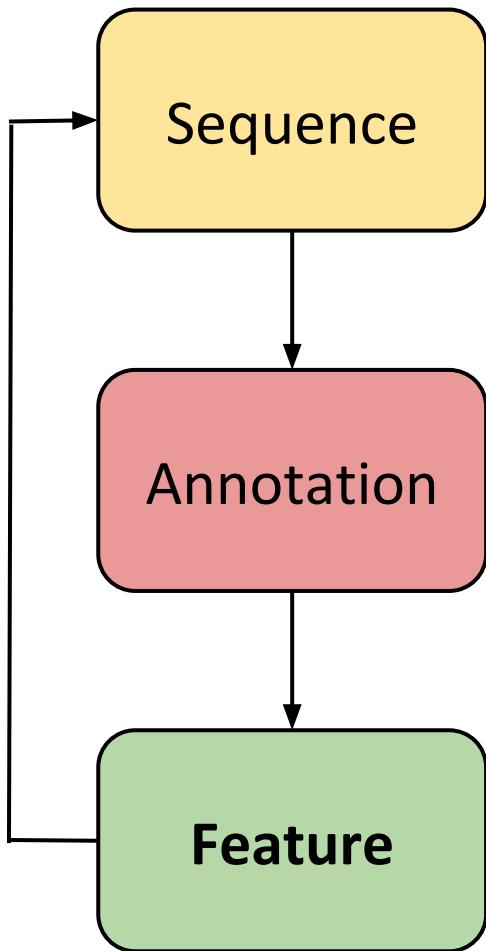
Hierarchical, Modular Annotation



Hierarchical, Modular Annotation



Visualizing Feature Roles



Non-Primary Structure

Sequence

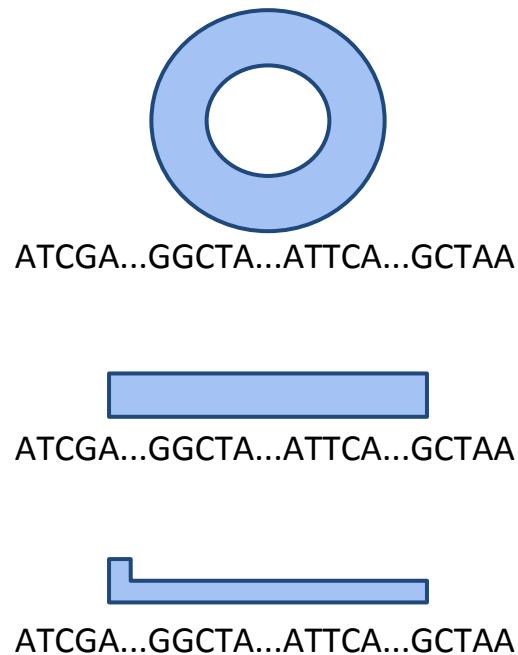
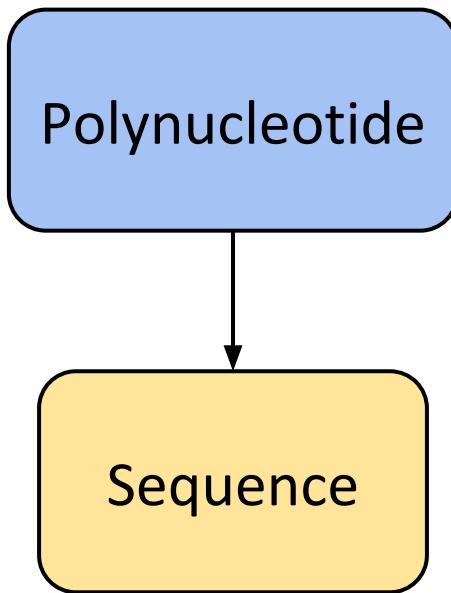
ATCGA...GGCTA...ATTCA...GCTAA

ATCGA...GGCTA...ATTCA...GCTAA

ATCGA...GGCTA...ATTCA...GCTAA



Non-Primary Structure



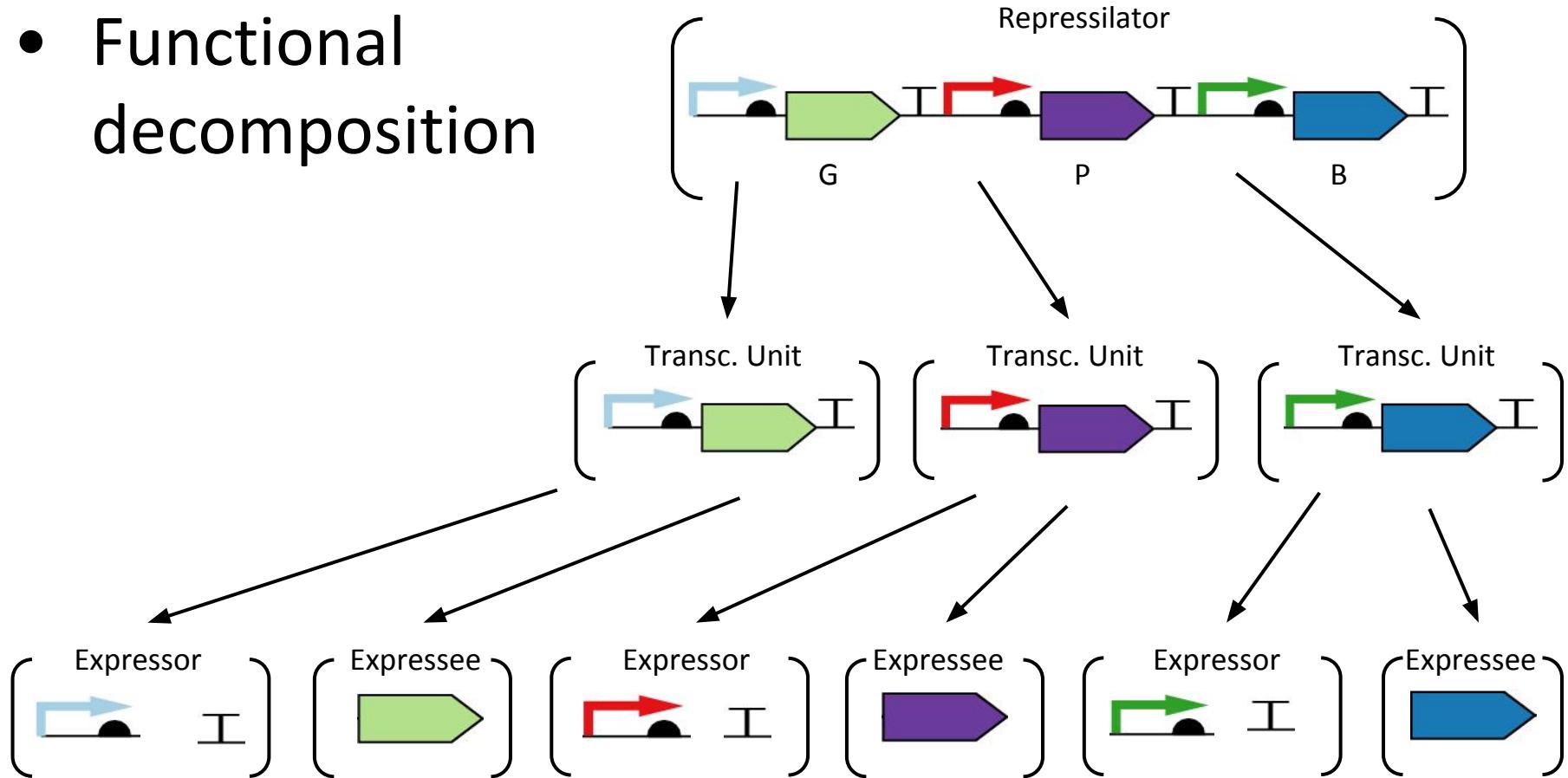
Sequence Editing and Optimization

Clotho provides schemas for representing

- Sequences
- Hierarchical, modular annotation
- Features and their functional roles
- Non-primary sequence structure

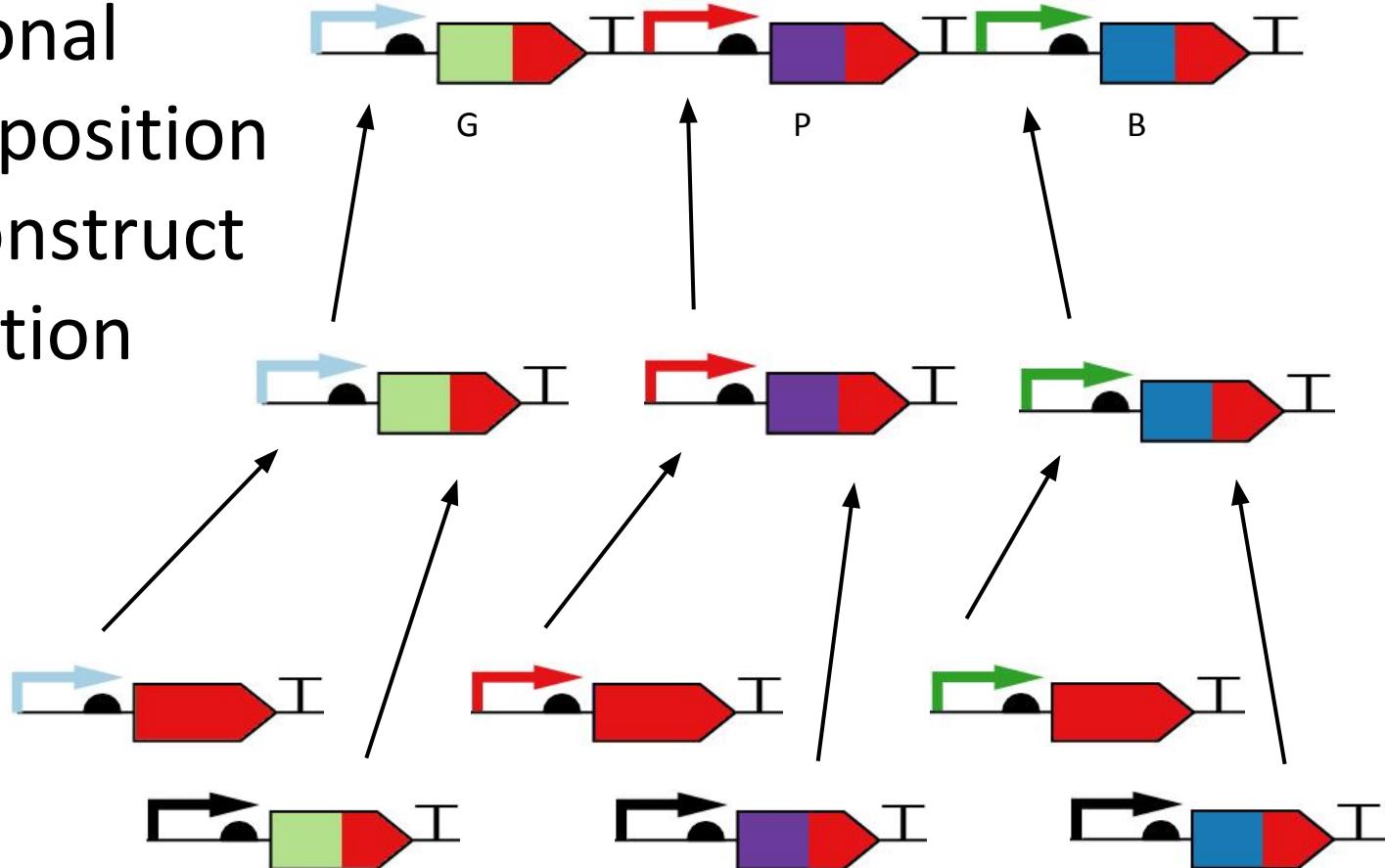
Phoenix

- Functional decomposition



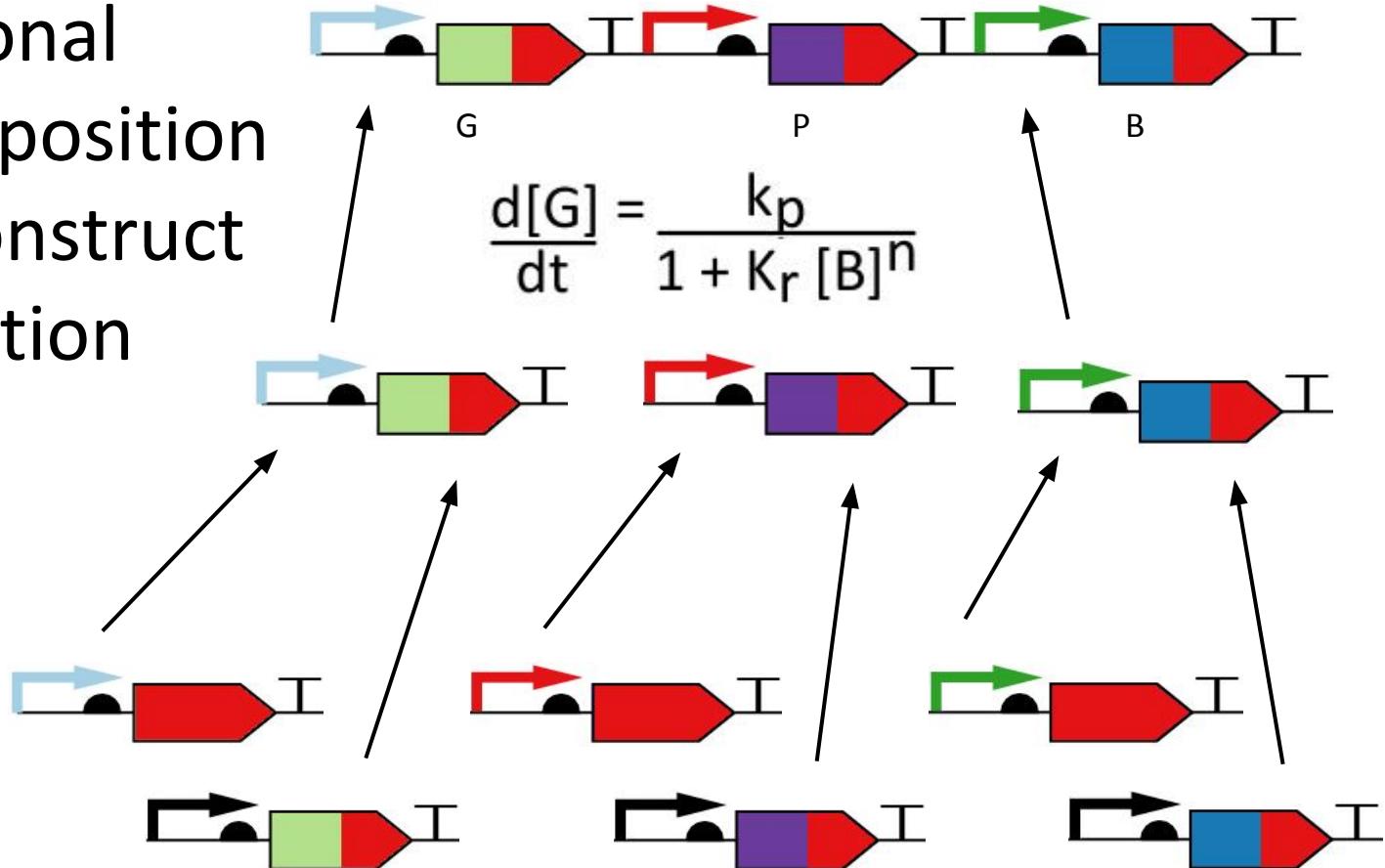
Phoenix

- Functional decomposition
- Test construct generation

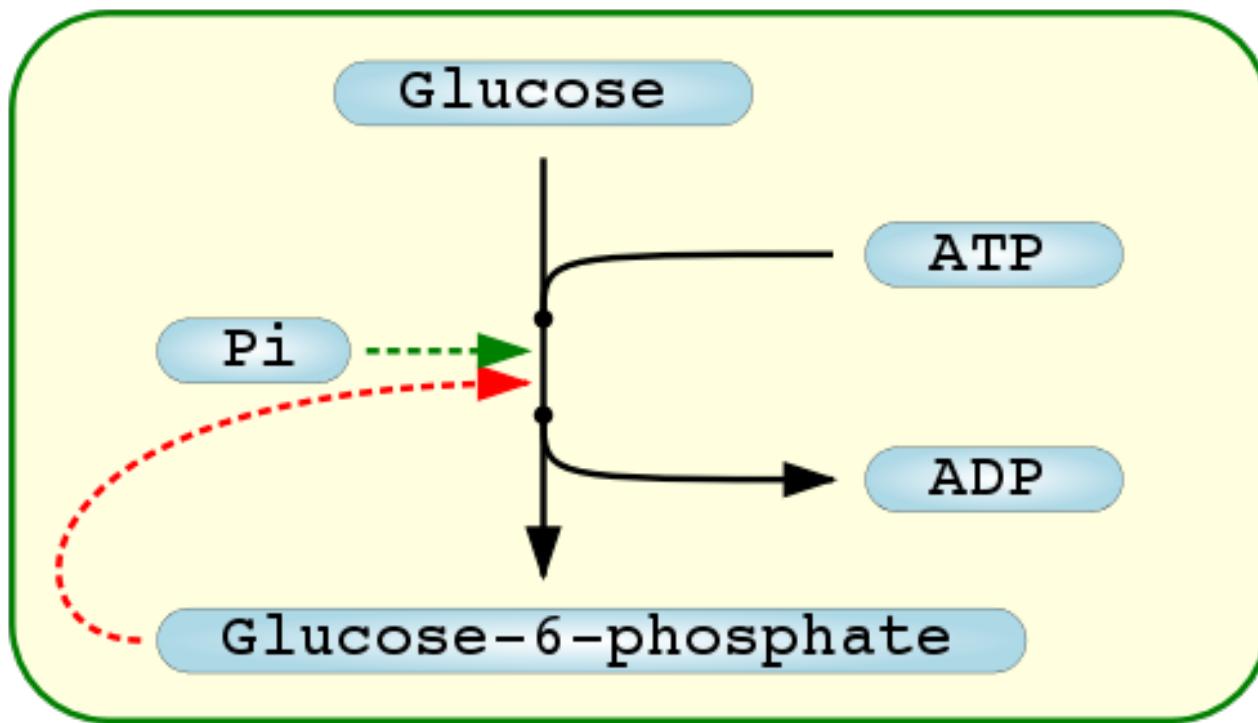


Phoenix

- Functional decomposition
- Test construct generation
- Model fitting



Systems Biology Markup Language

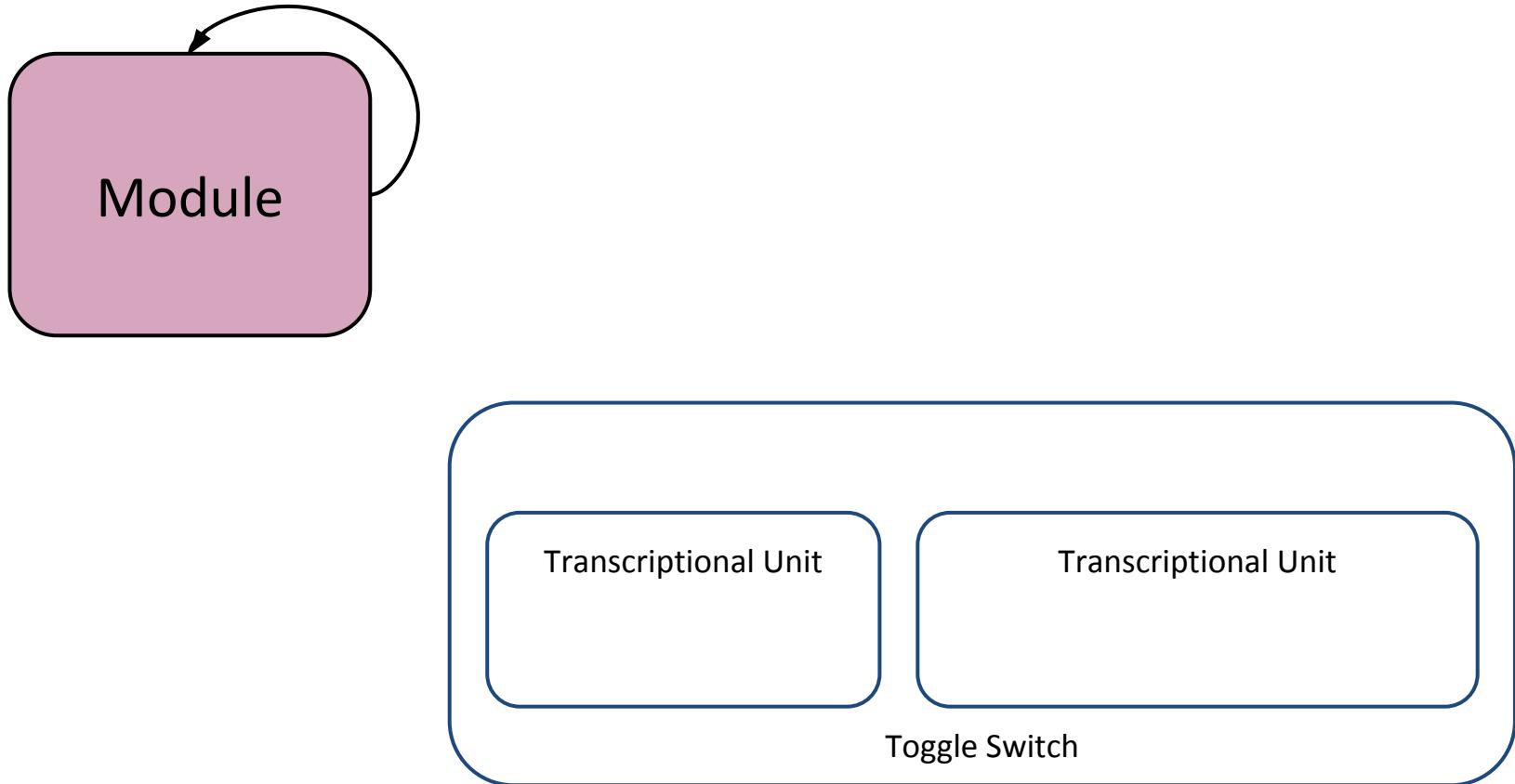


Functional Modules

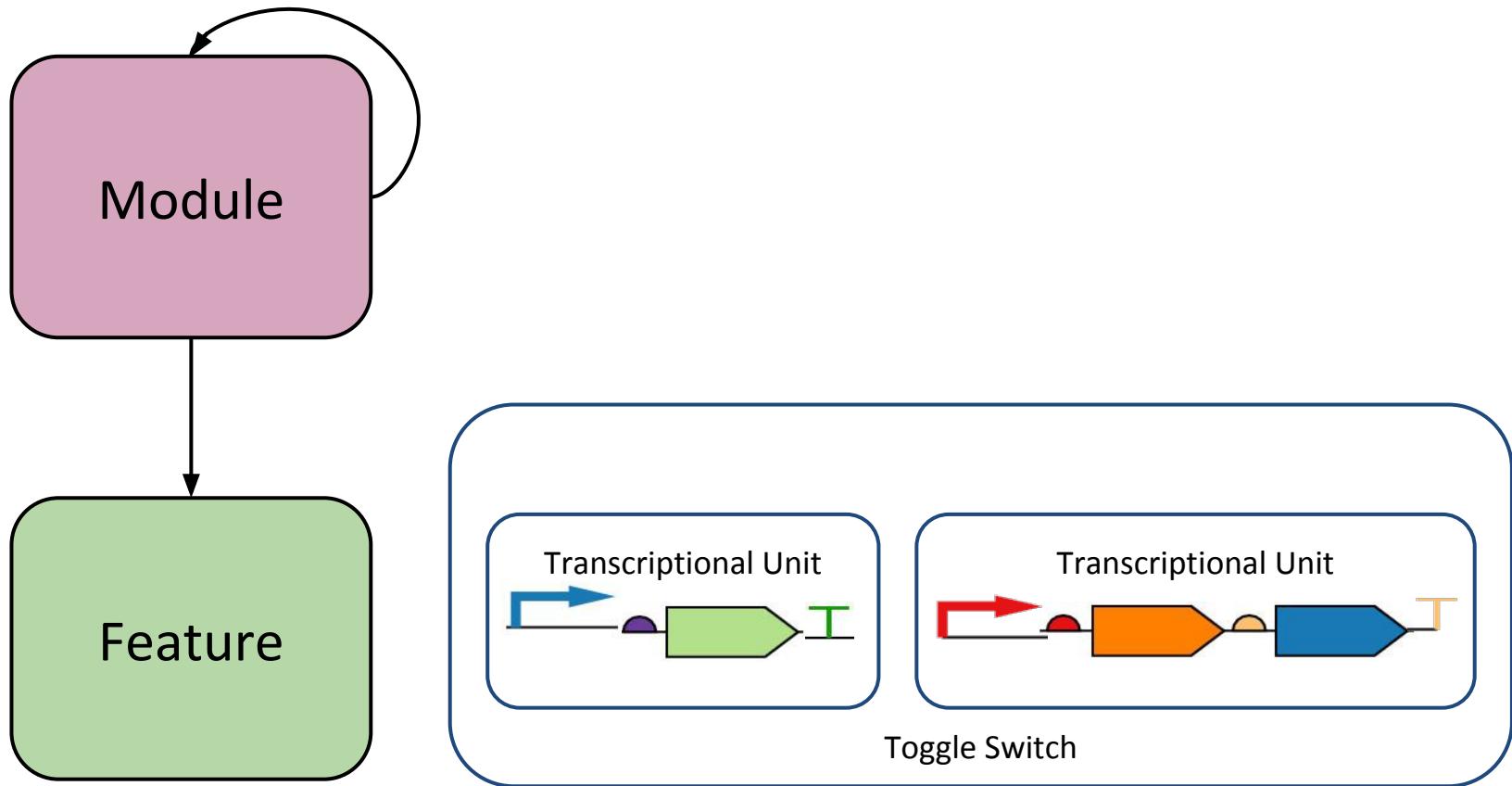
Module

Toggle Switch

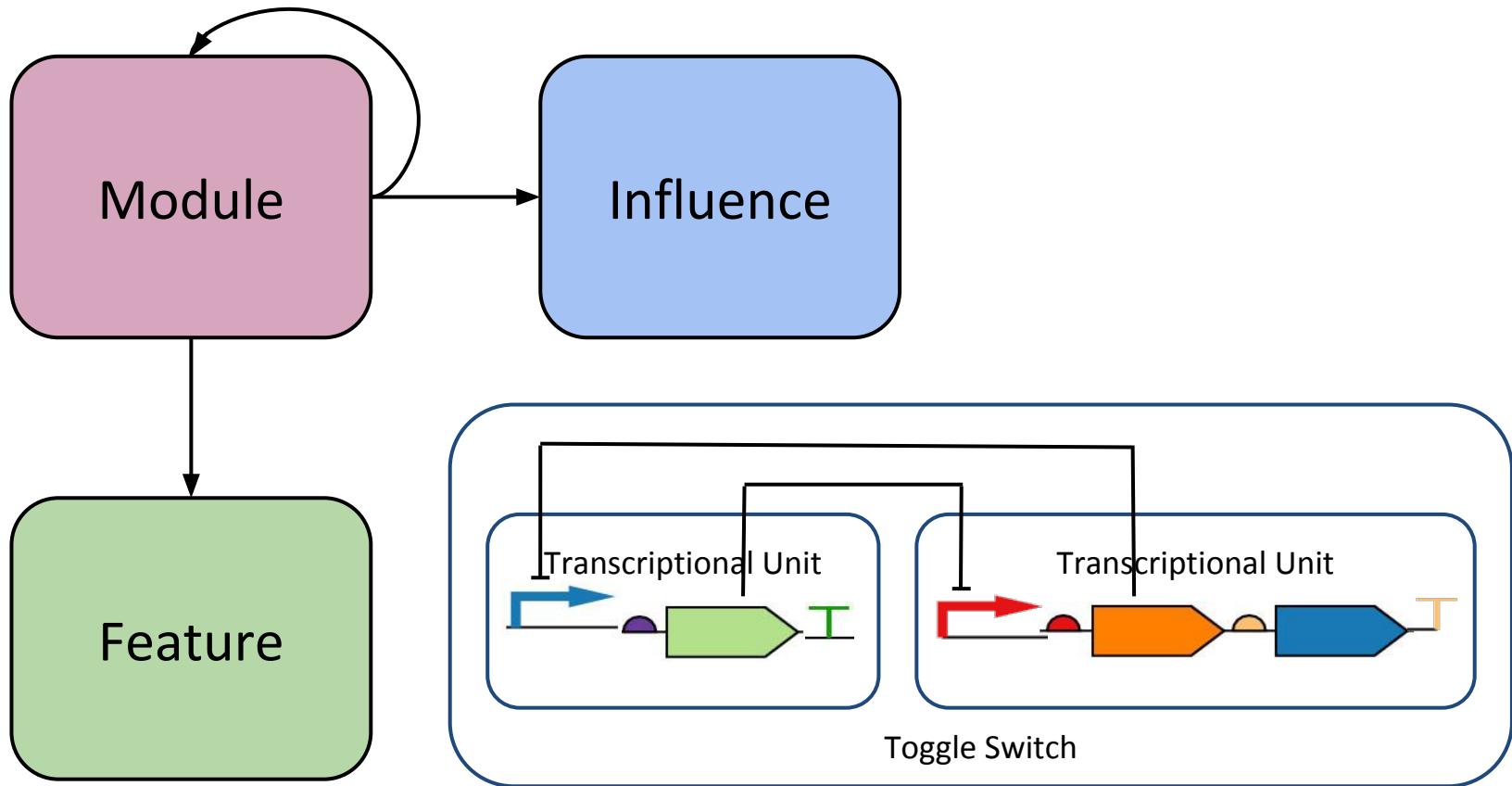
Functional Hierarchy



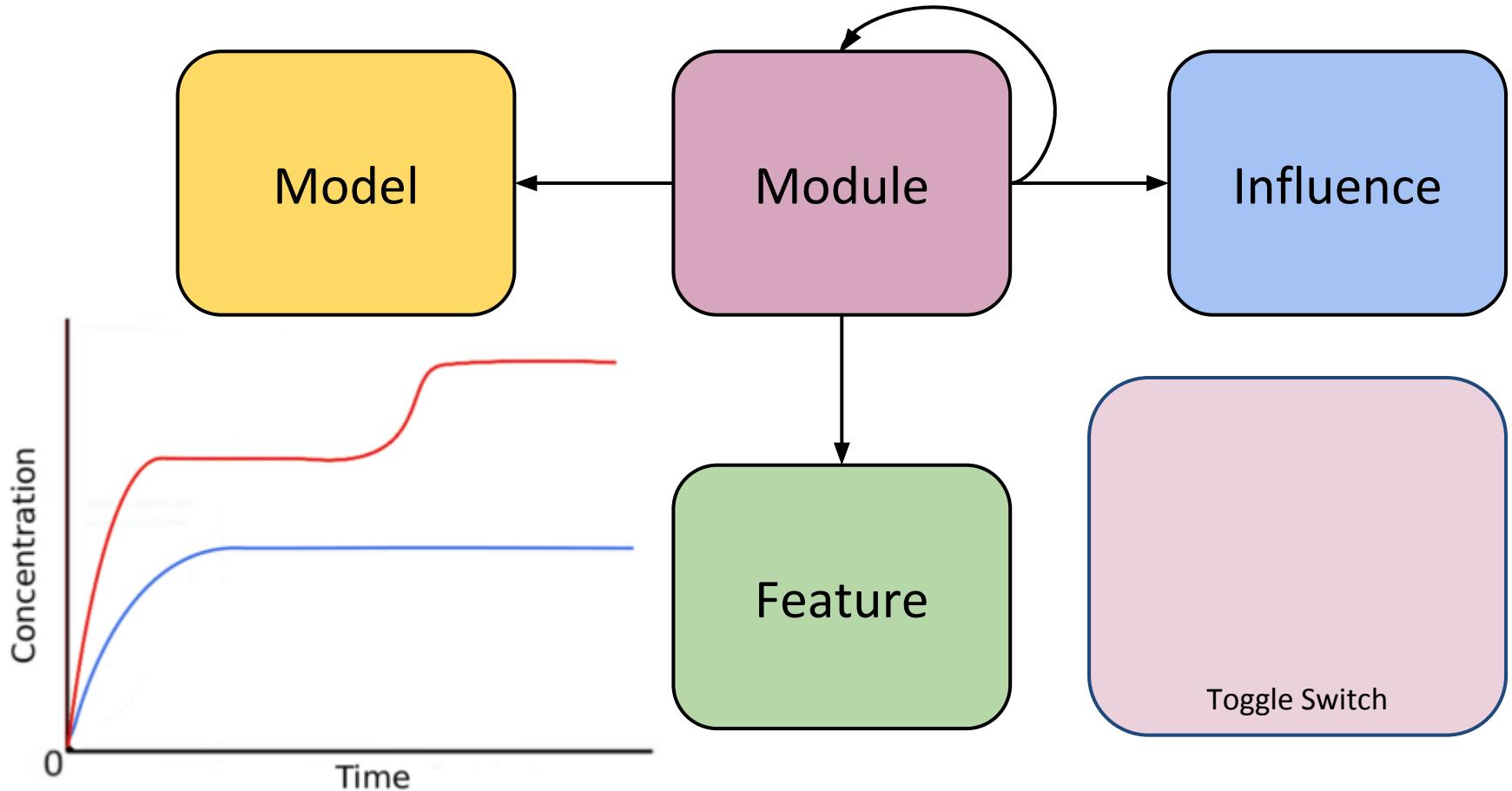
Grouping Features into Modules



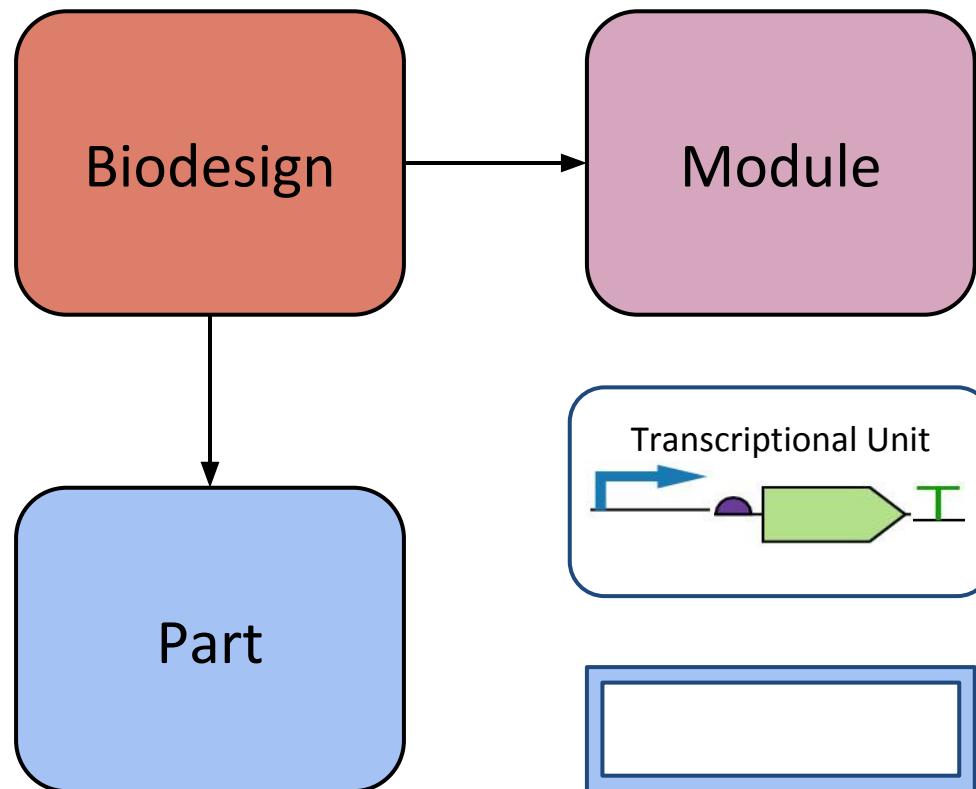
Influences Between Features



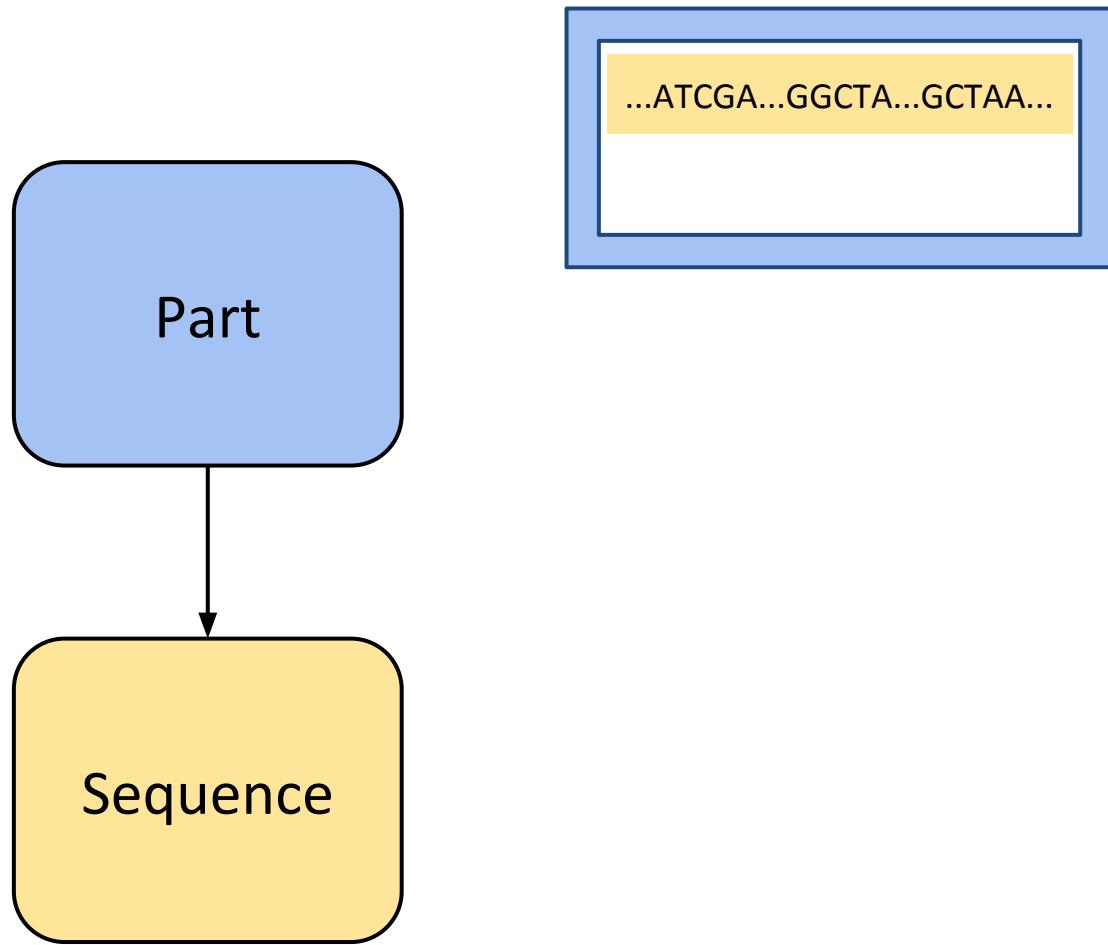
Associating Mathematical Models



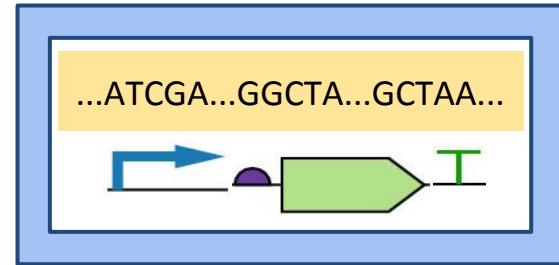
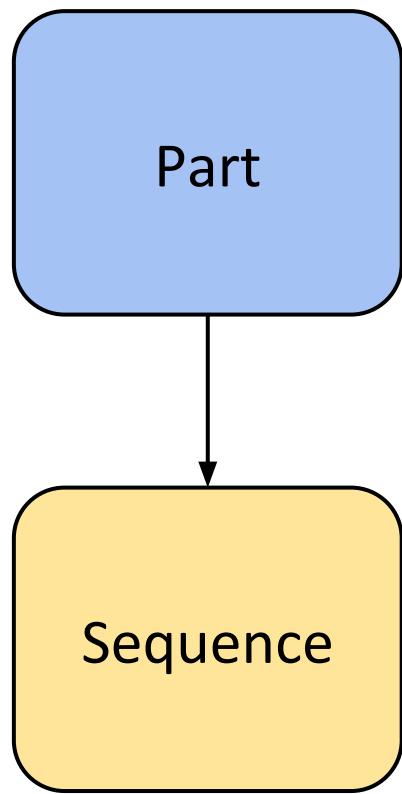
Linking Function and Assembly



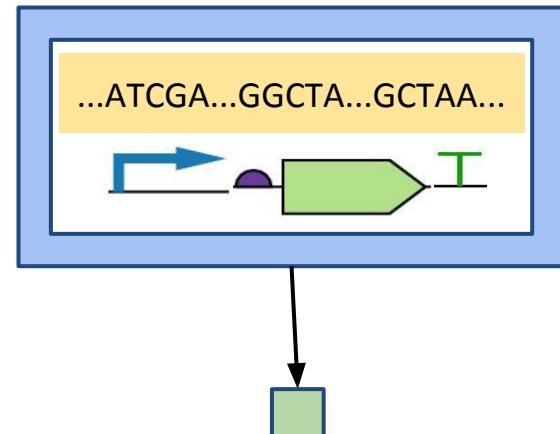
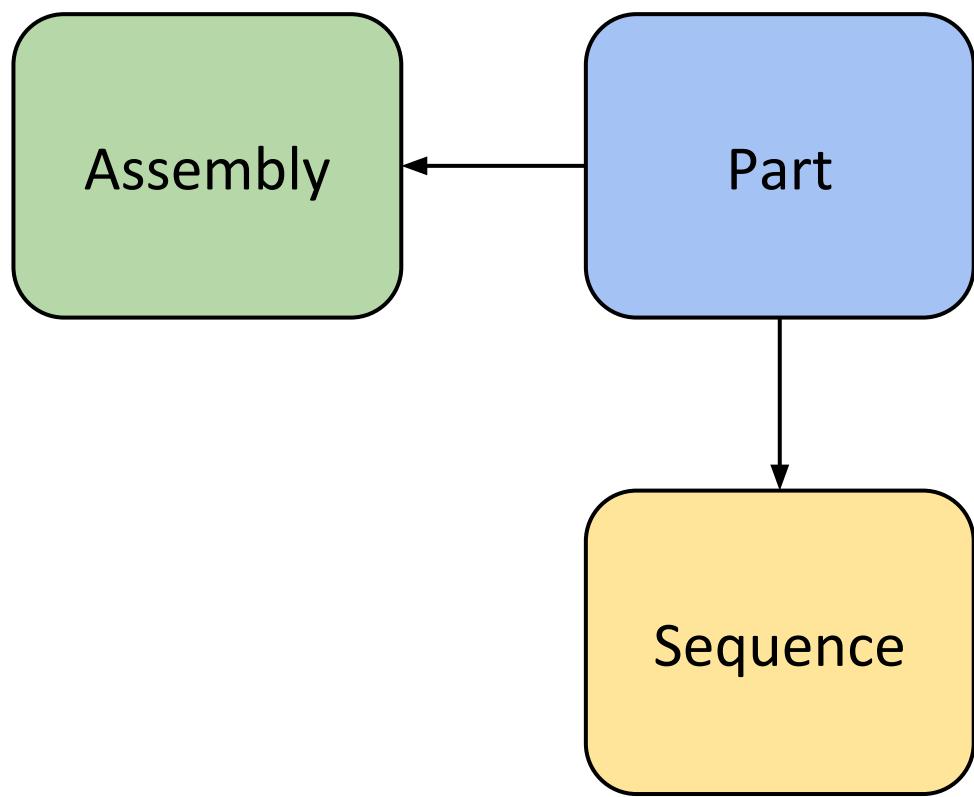
Parts Put Sequences Together



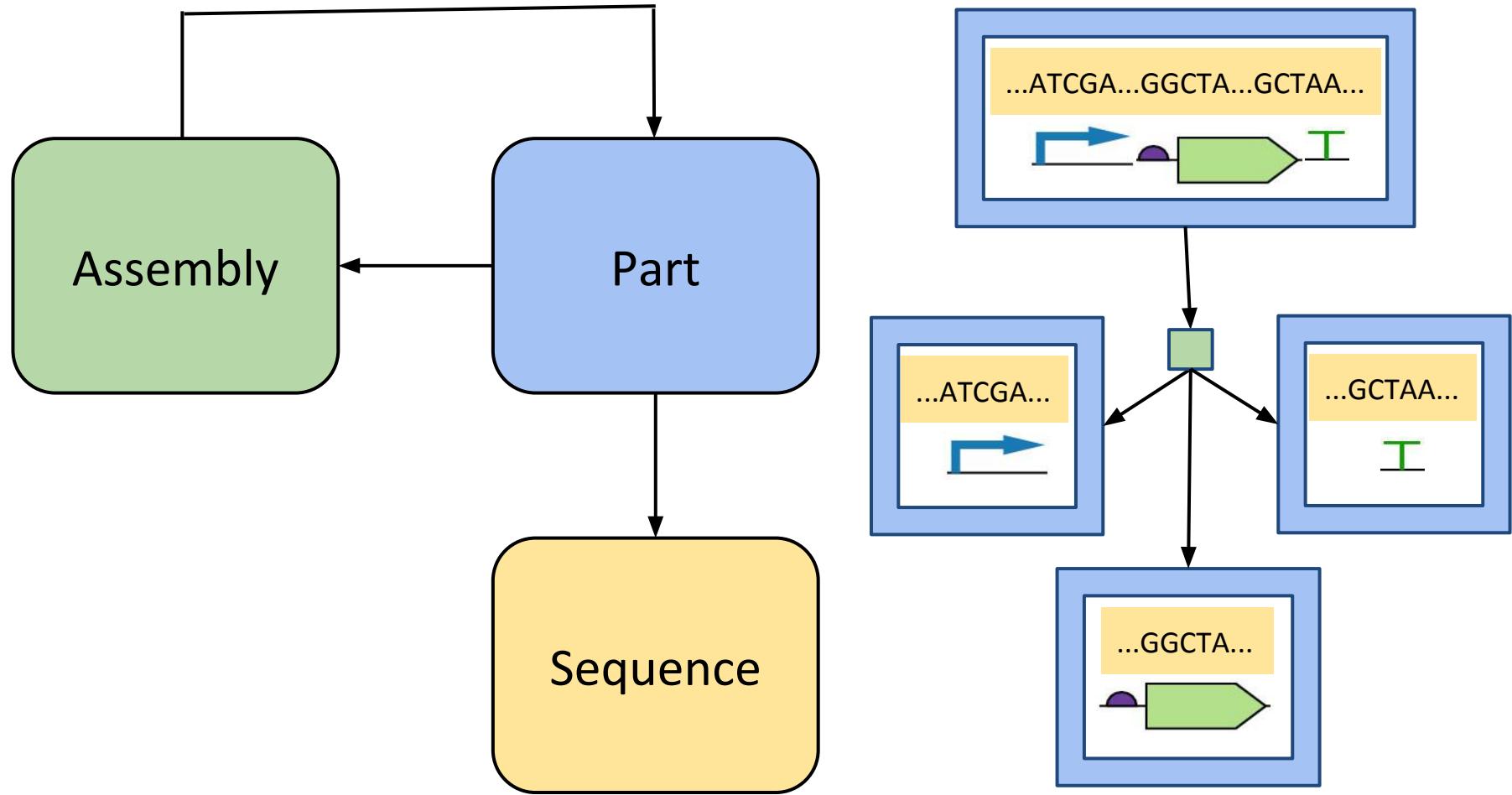
Parts Lack Explicit Function



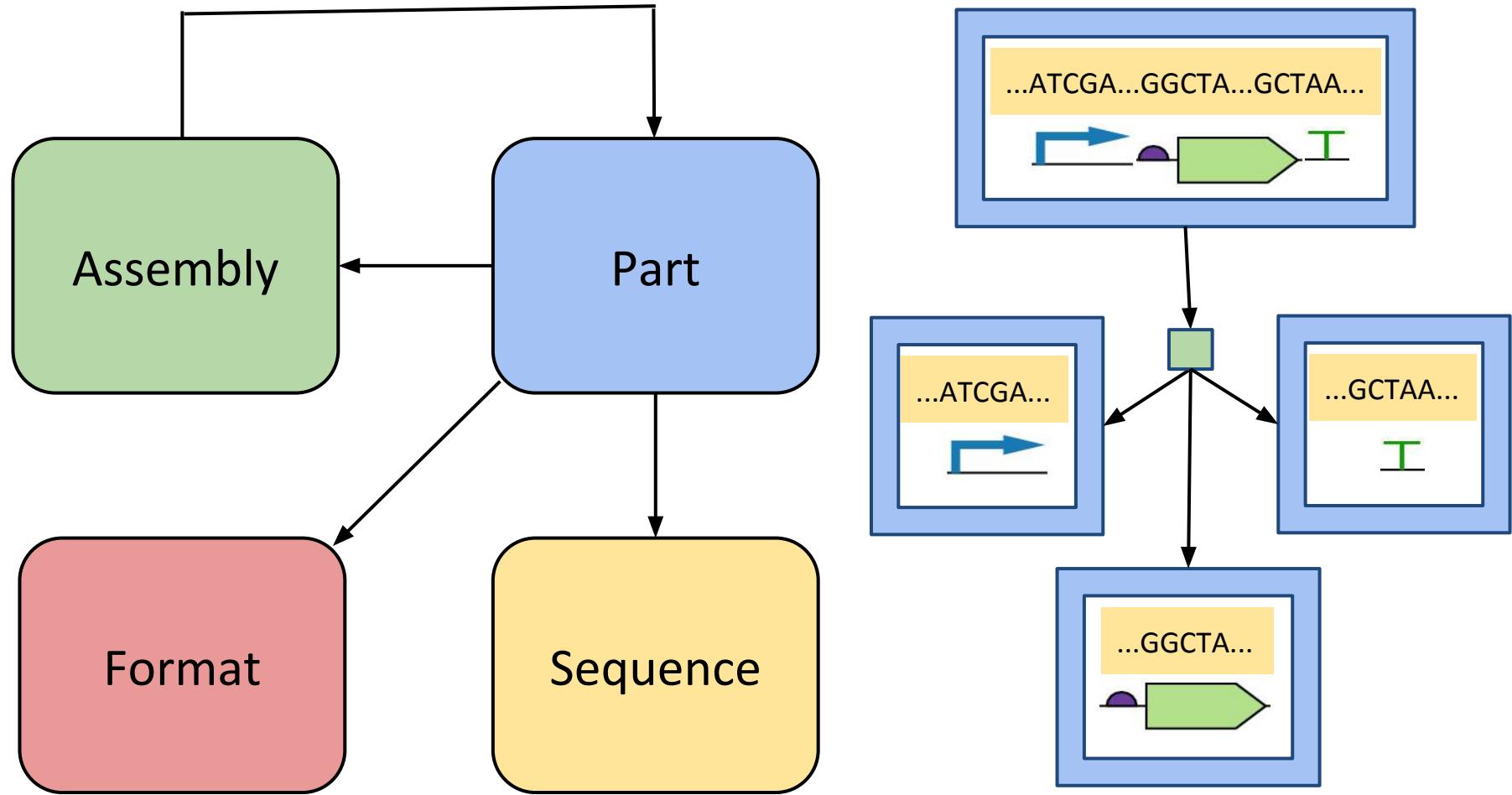
Assembling Composite Part



Assembling Composite Part

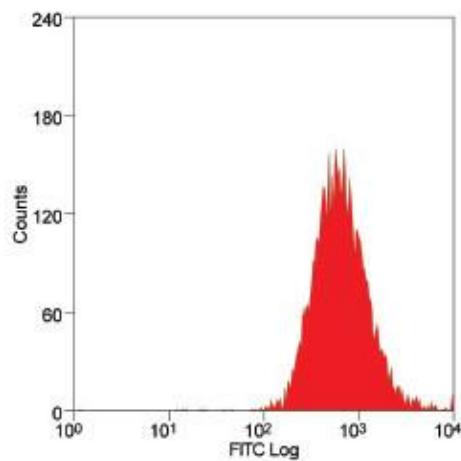


Part Formats

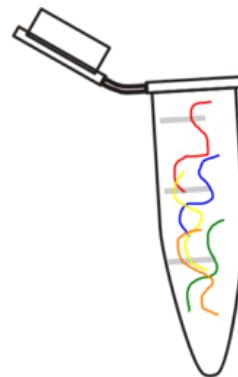
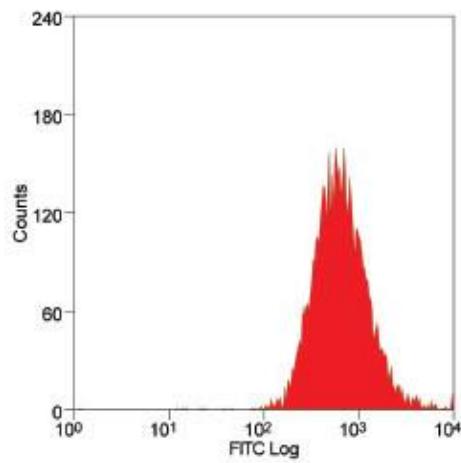
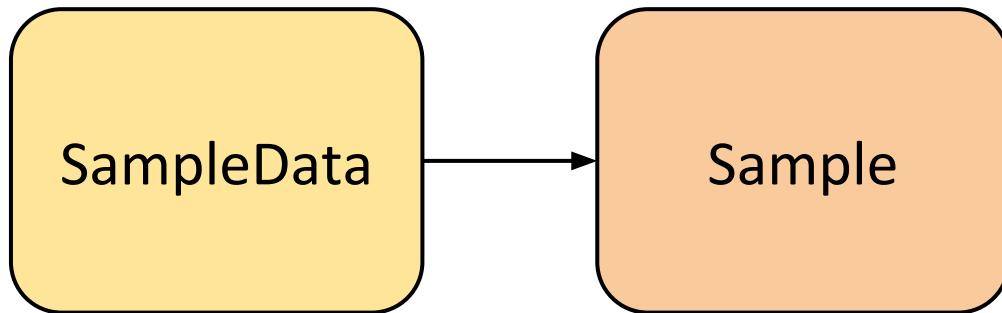


Storing Sample Data

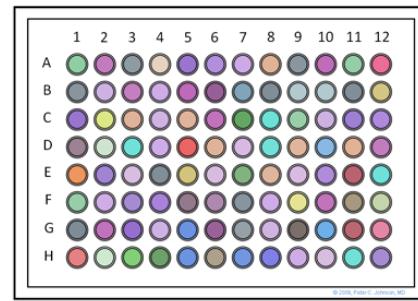
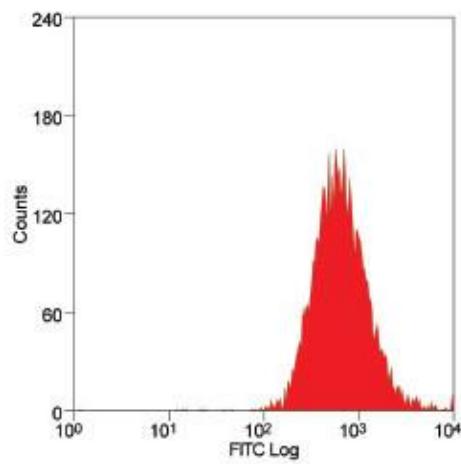
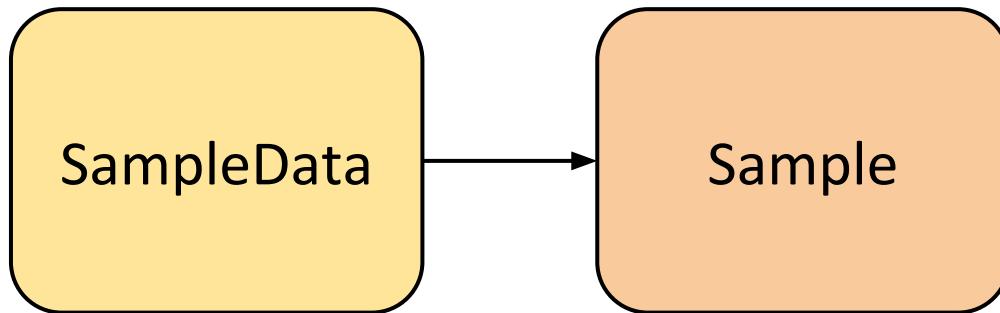
SampleData



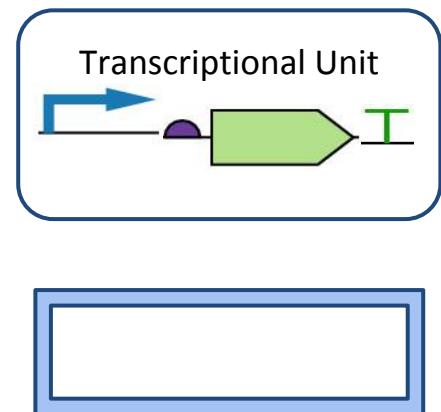
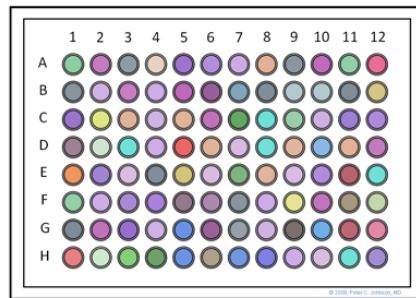
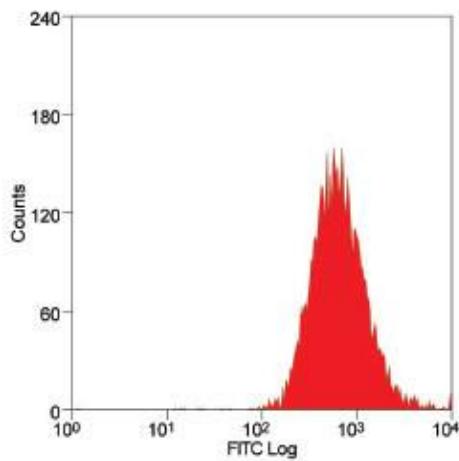
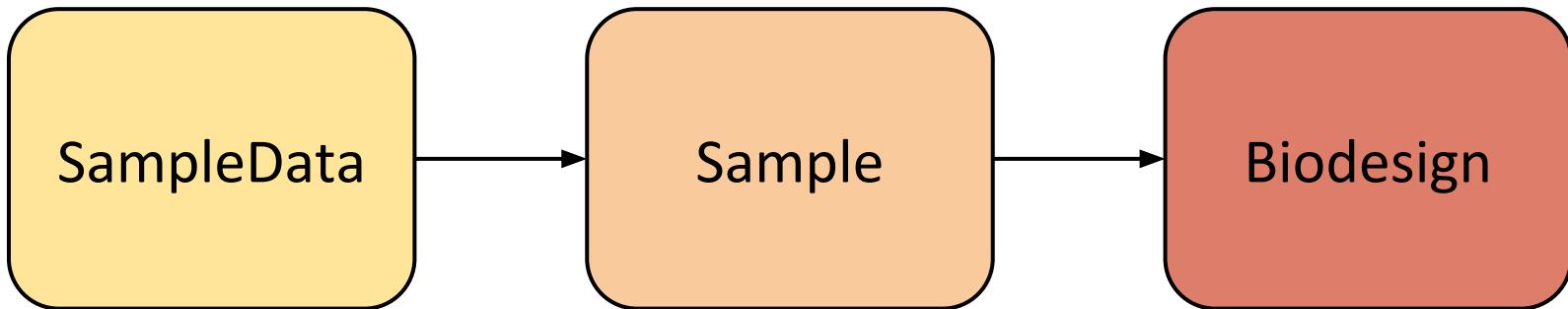
Tracking Samples



Tracking Samples



Defining Sample Contents



Model Fitting and Exp. Planning

Clotho provides schemas for representing

- Functional hierarchies of modules, features, and influences
- Parts to assemble sequences
- Biodesigns to link separate concerns
- Samples and data collected on them

Statistical Design of Experiments

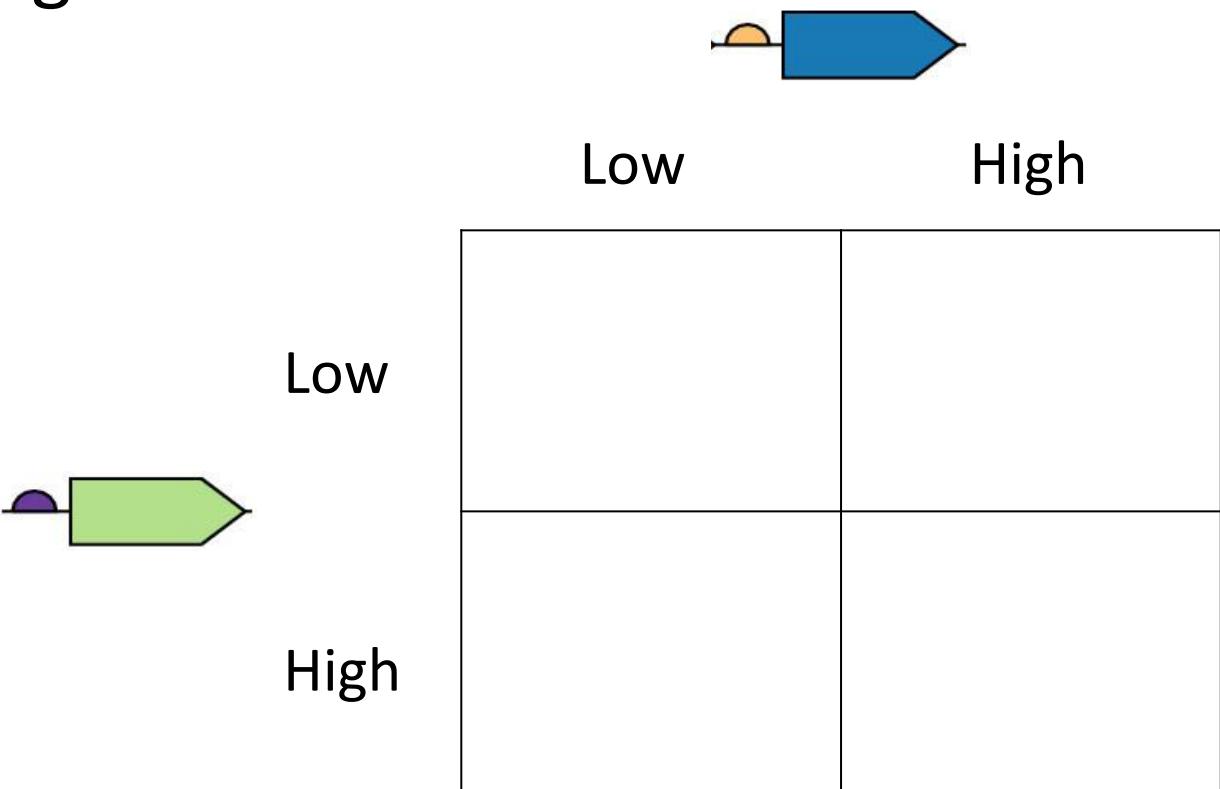
		Temperature	
		Low	High
pH	Low	Low pH, Low T	Low pH, High T
	High	High pH, Low T	High pH, High T

Statistical Design of Experiments

		Temperature	
		Low	High
pH	Low		Low pH, High T
	High	High pH, Low T	

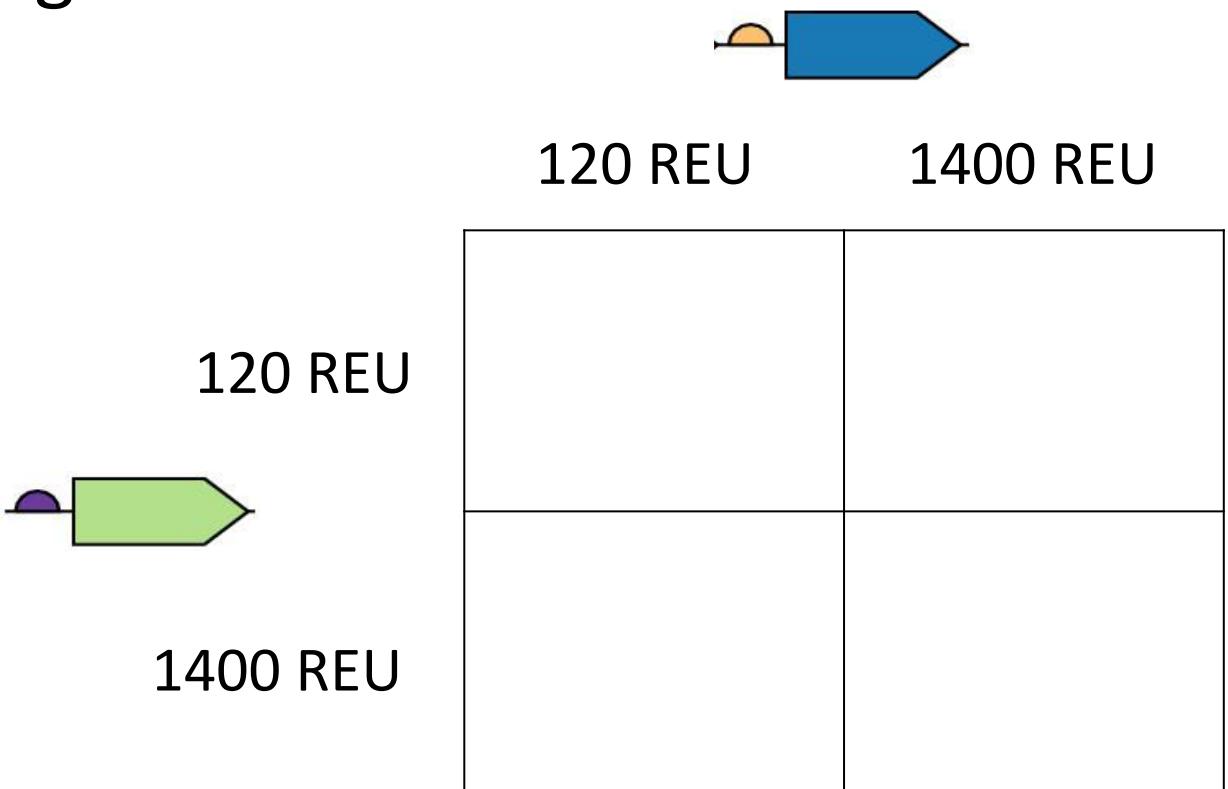
Double Dutch

- Level clustering



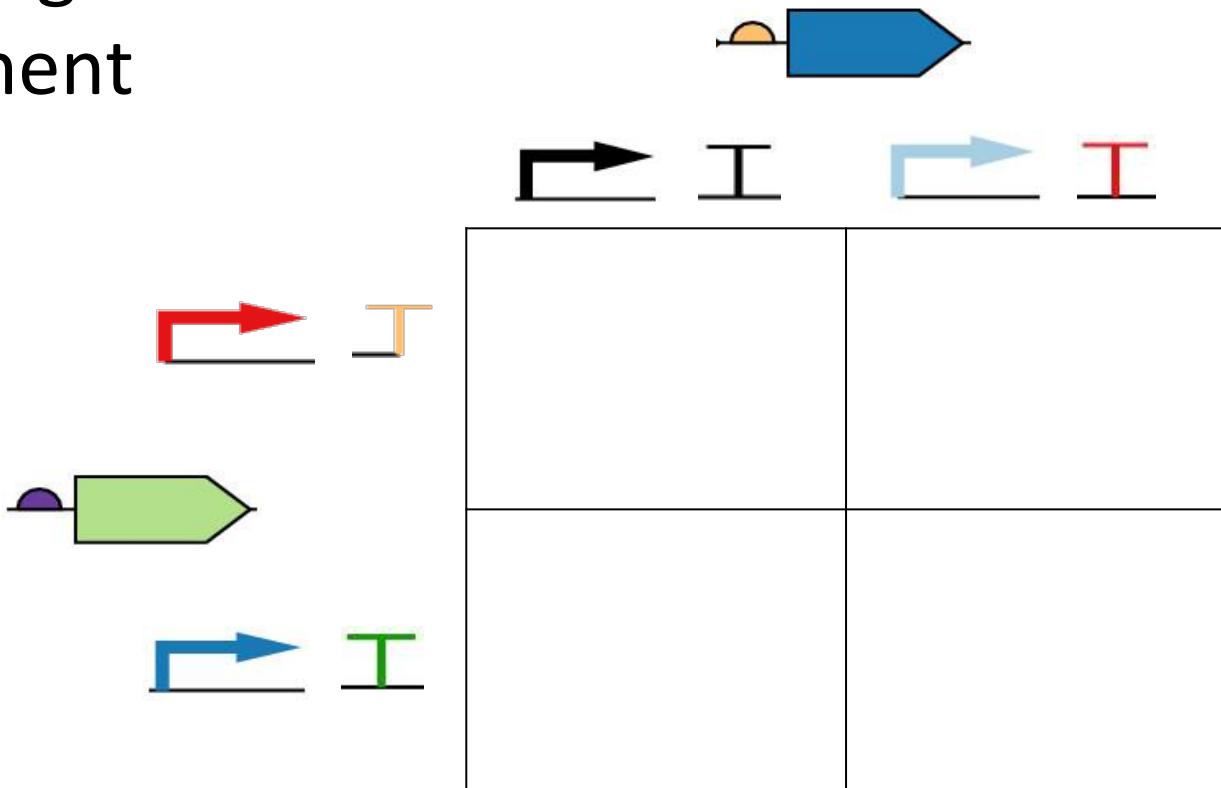
Double Dutch

- Level clustering



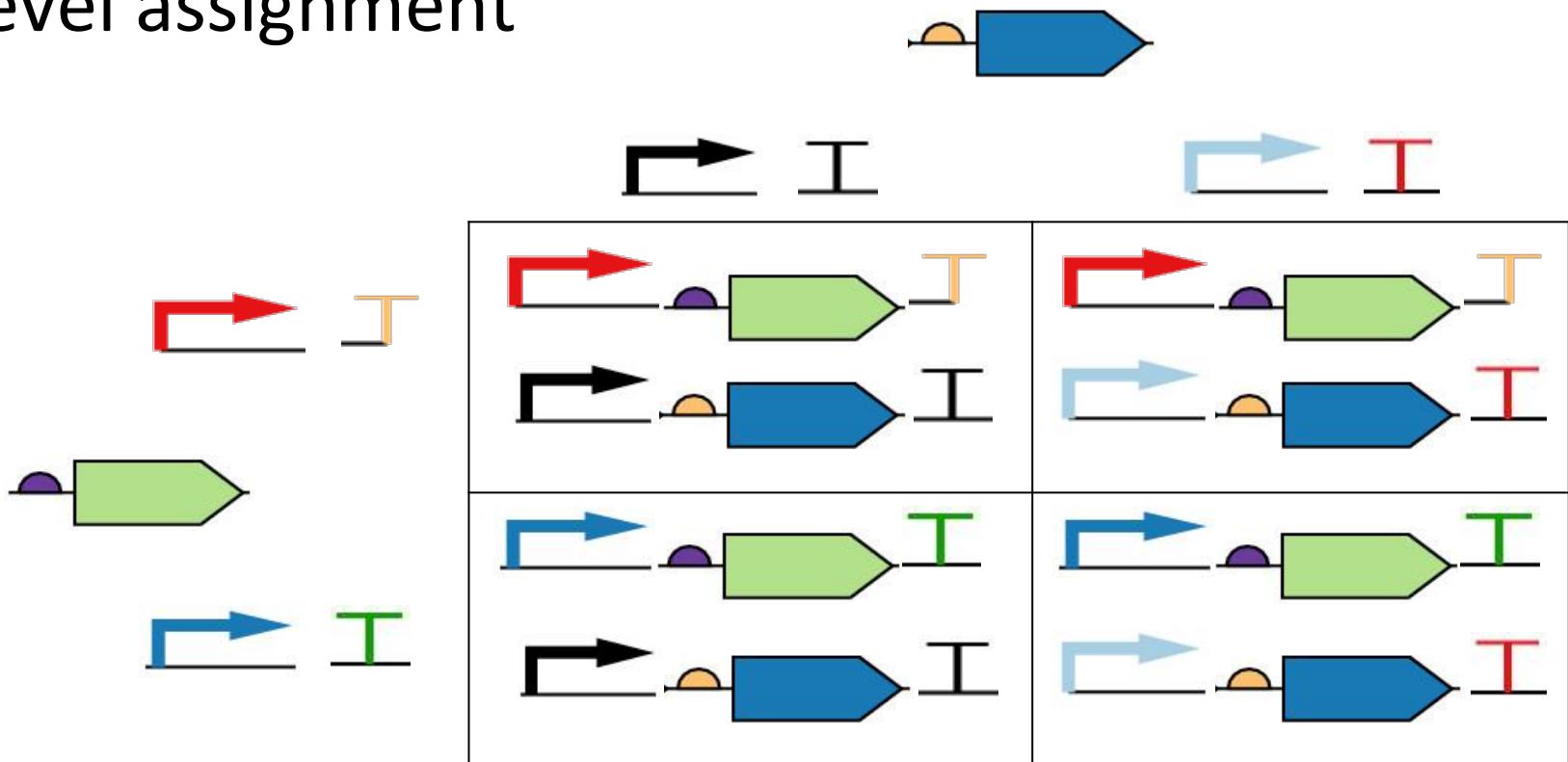
Double Dutch

- Level clustering
- Level assignment



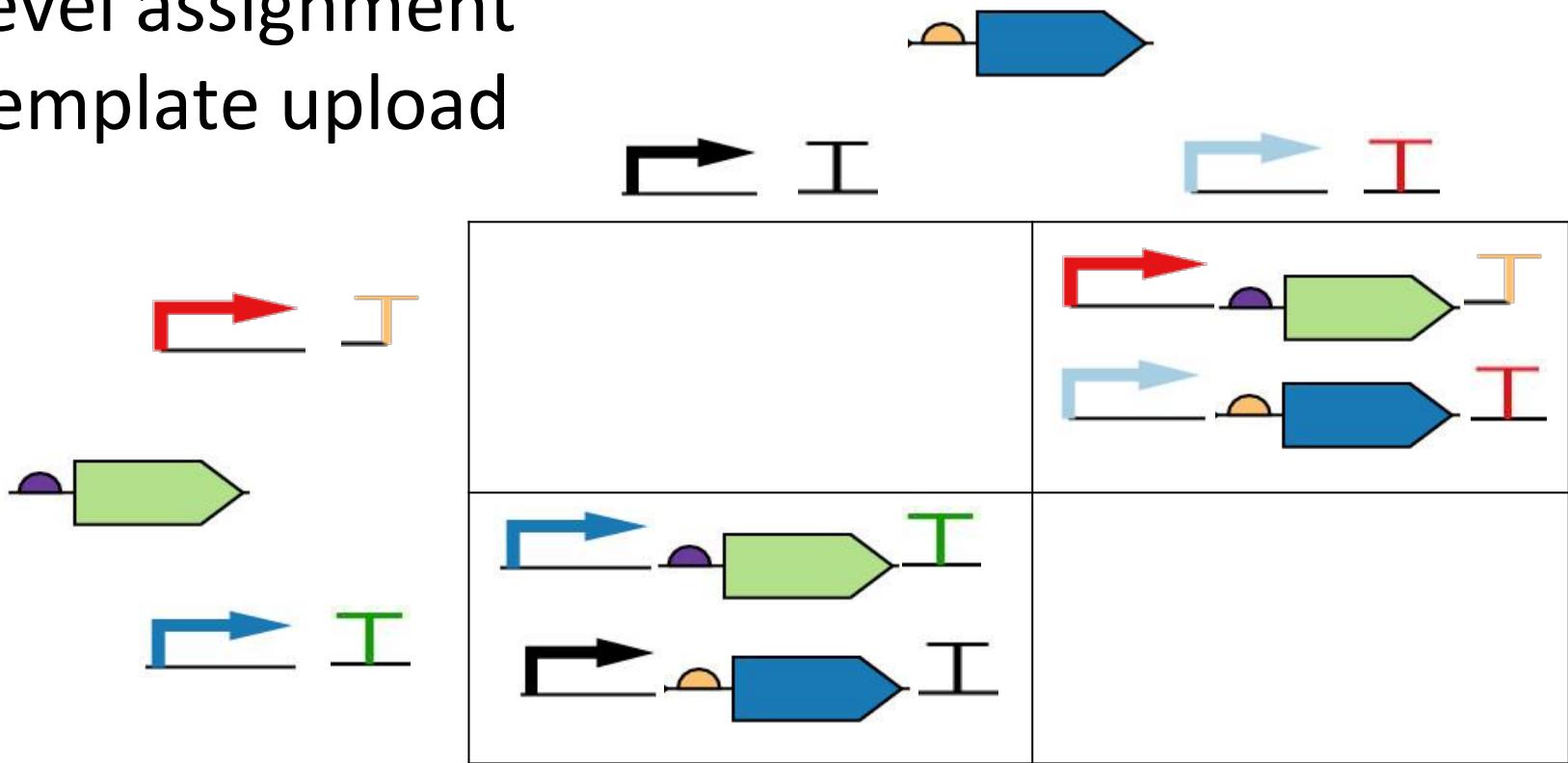
Double Dutch

- Level clustering
- Level assignment

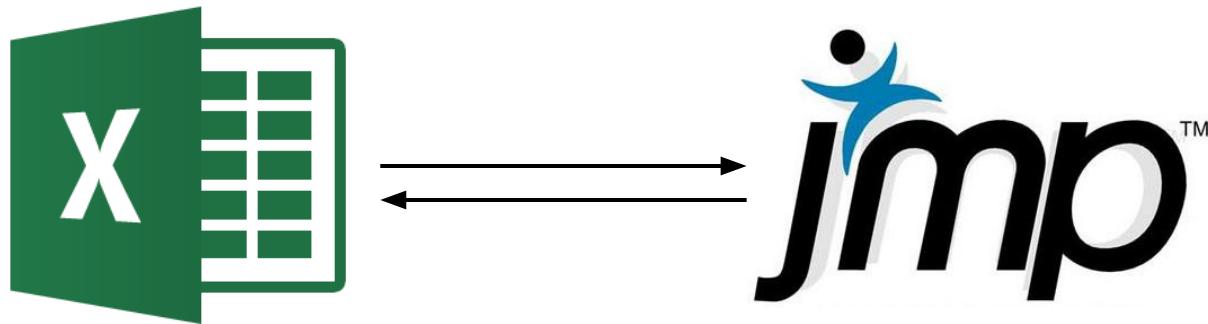


Double Dutch

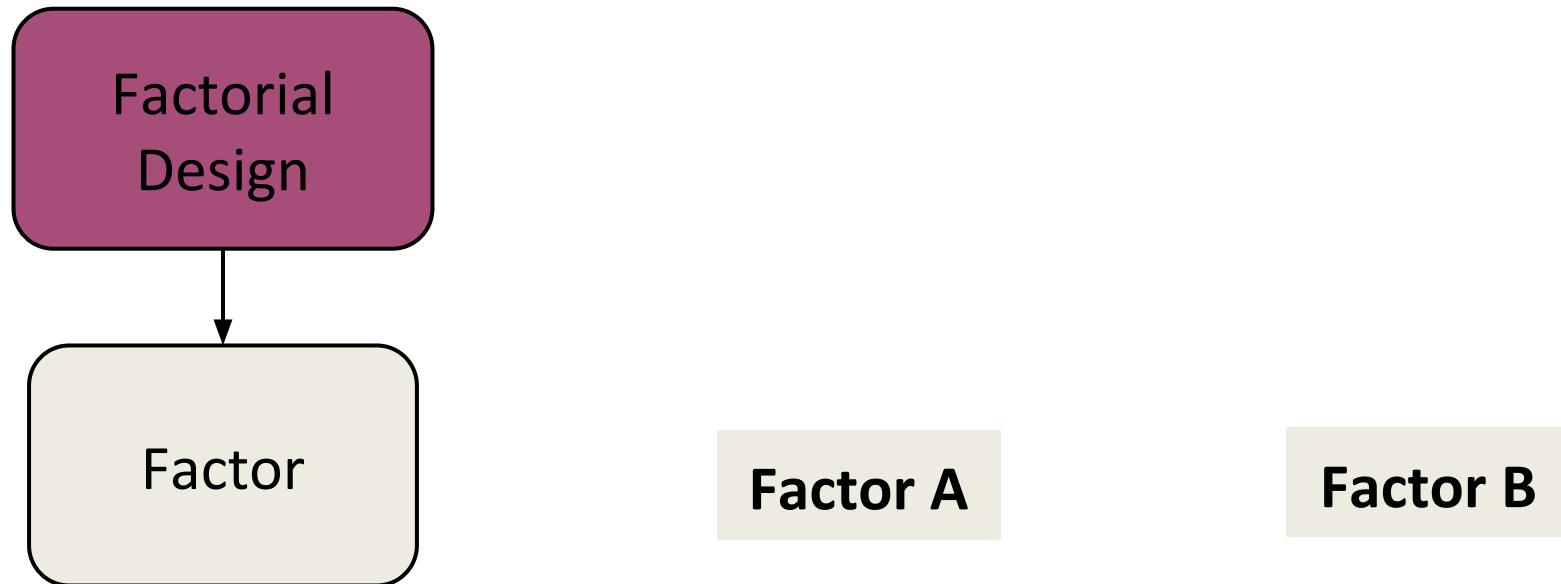
- Level clustering
- Level assignment
- Template upload



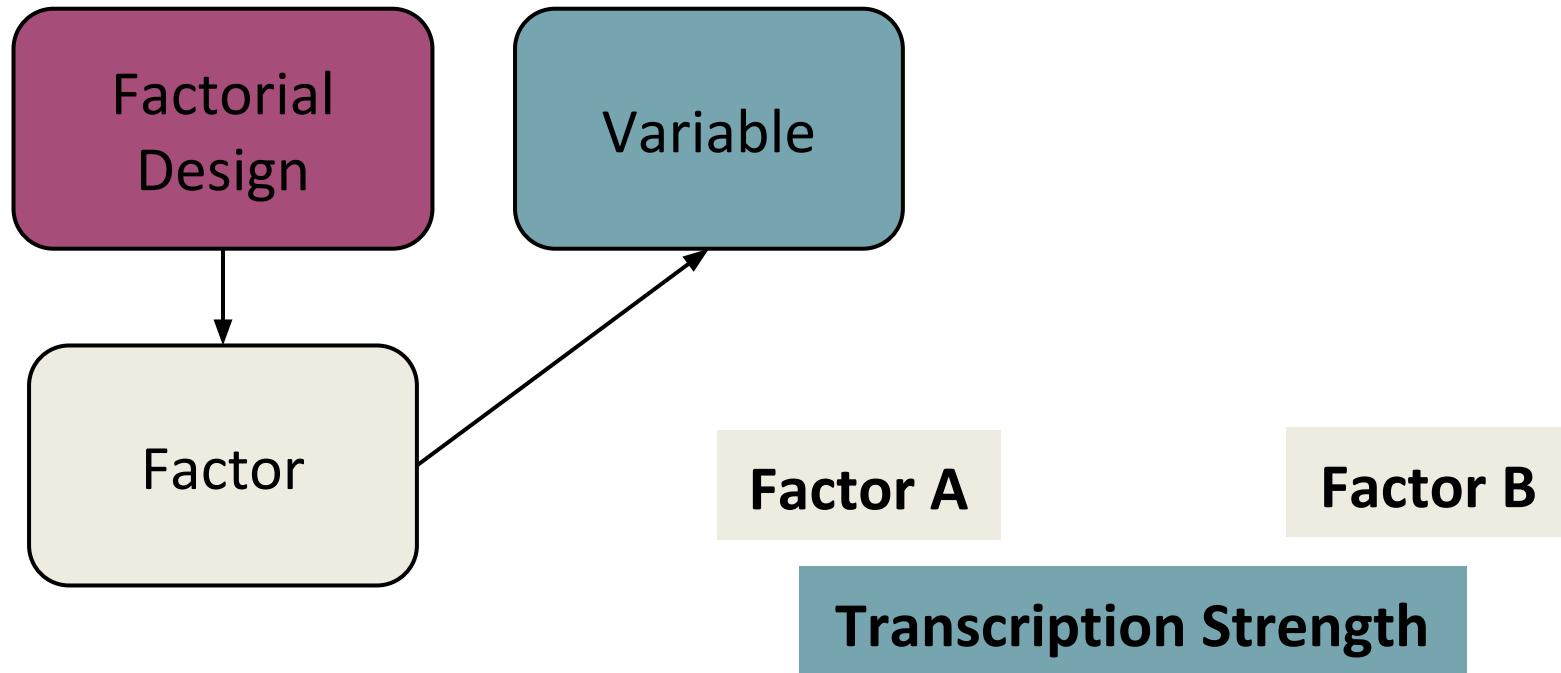
Excel and JMP



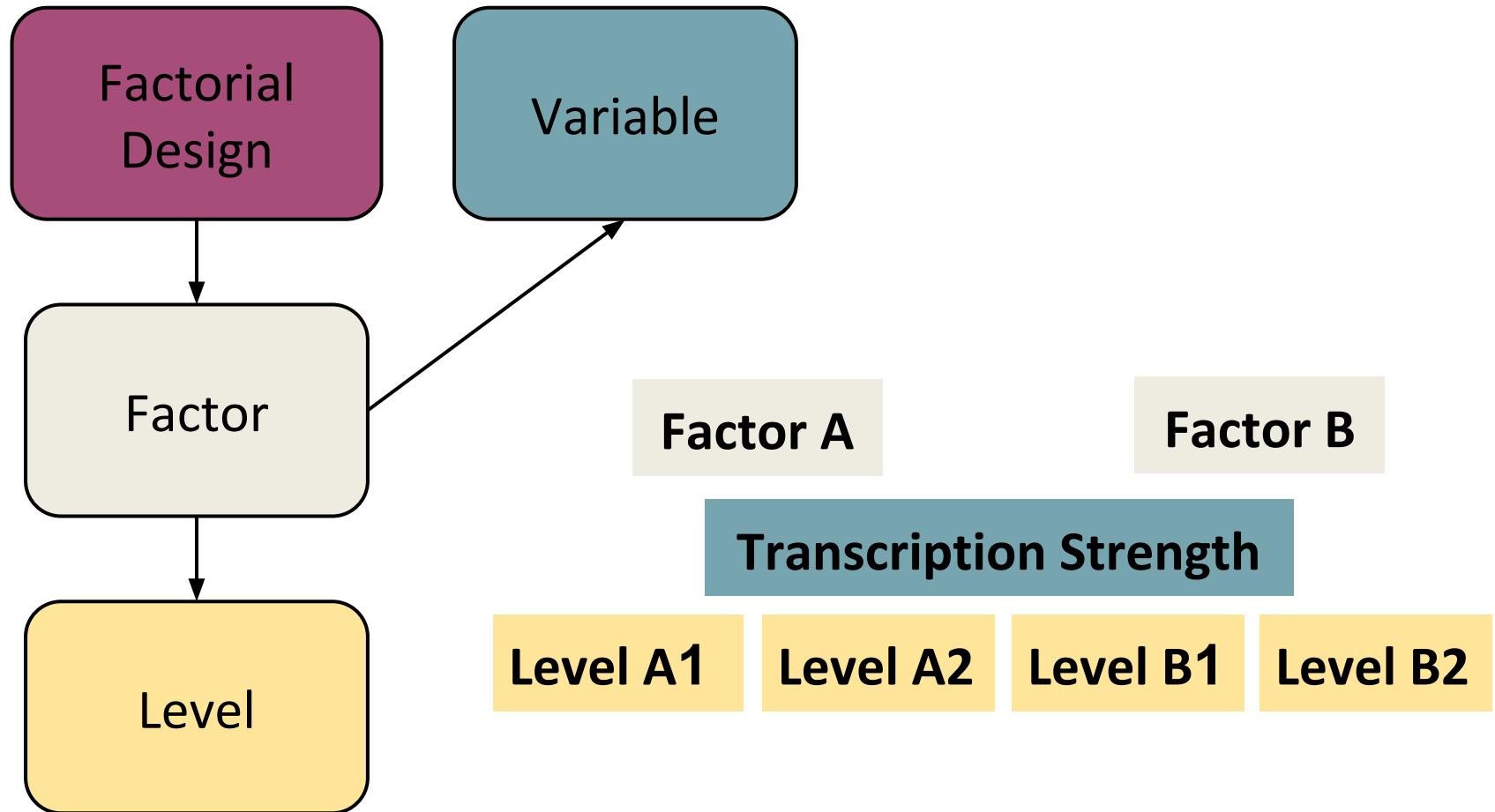
Factorial Design



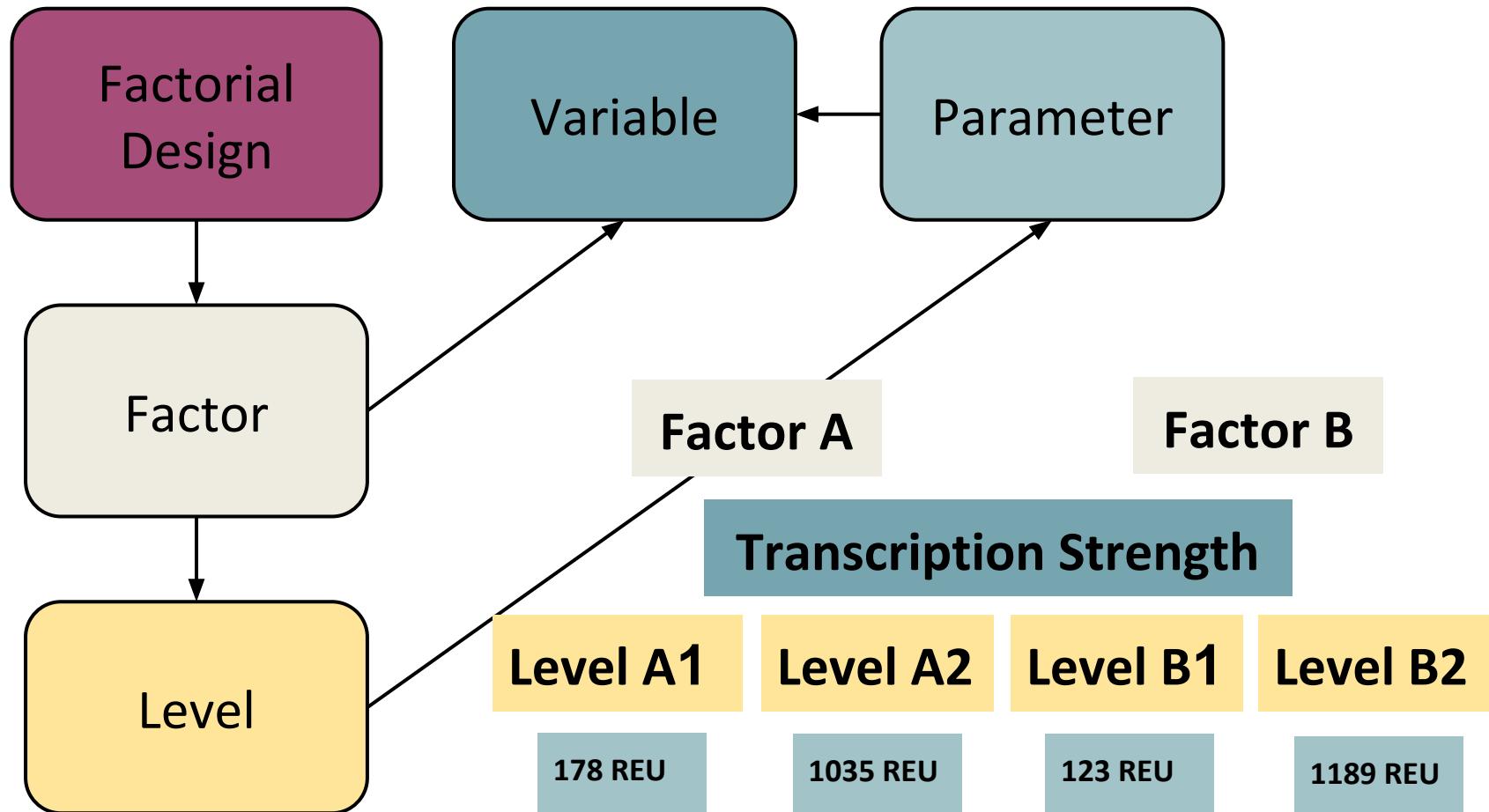
Factorial Design



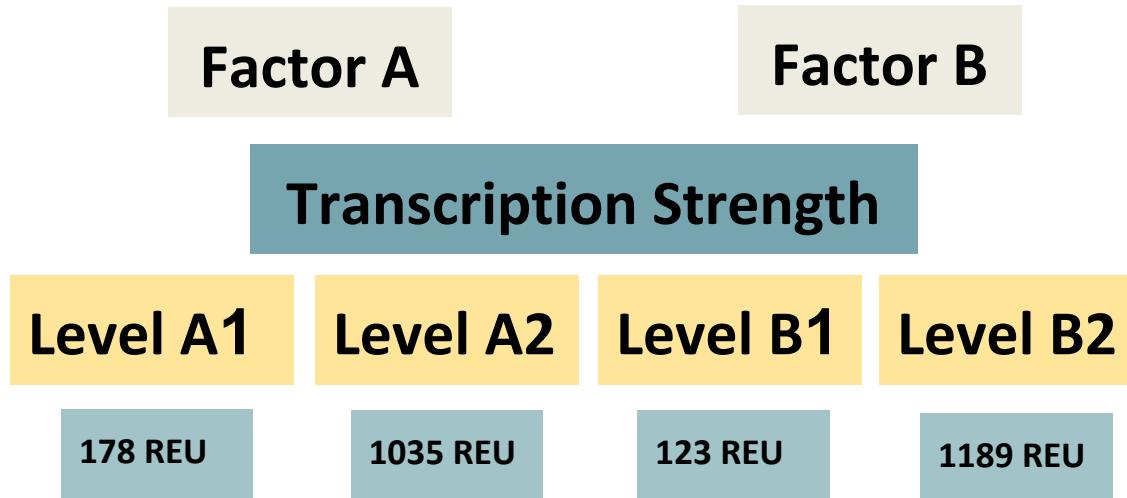
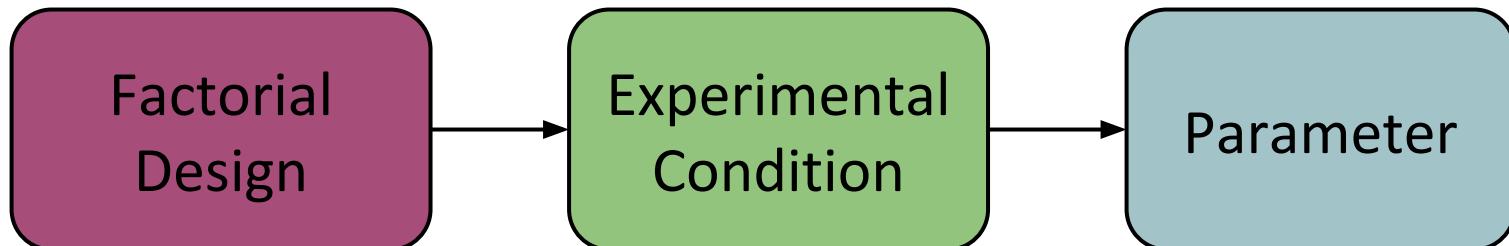
Factorial Design



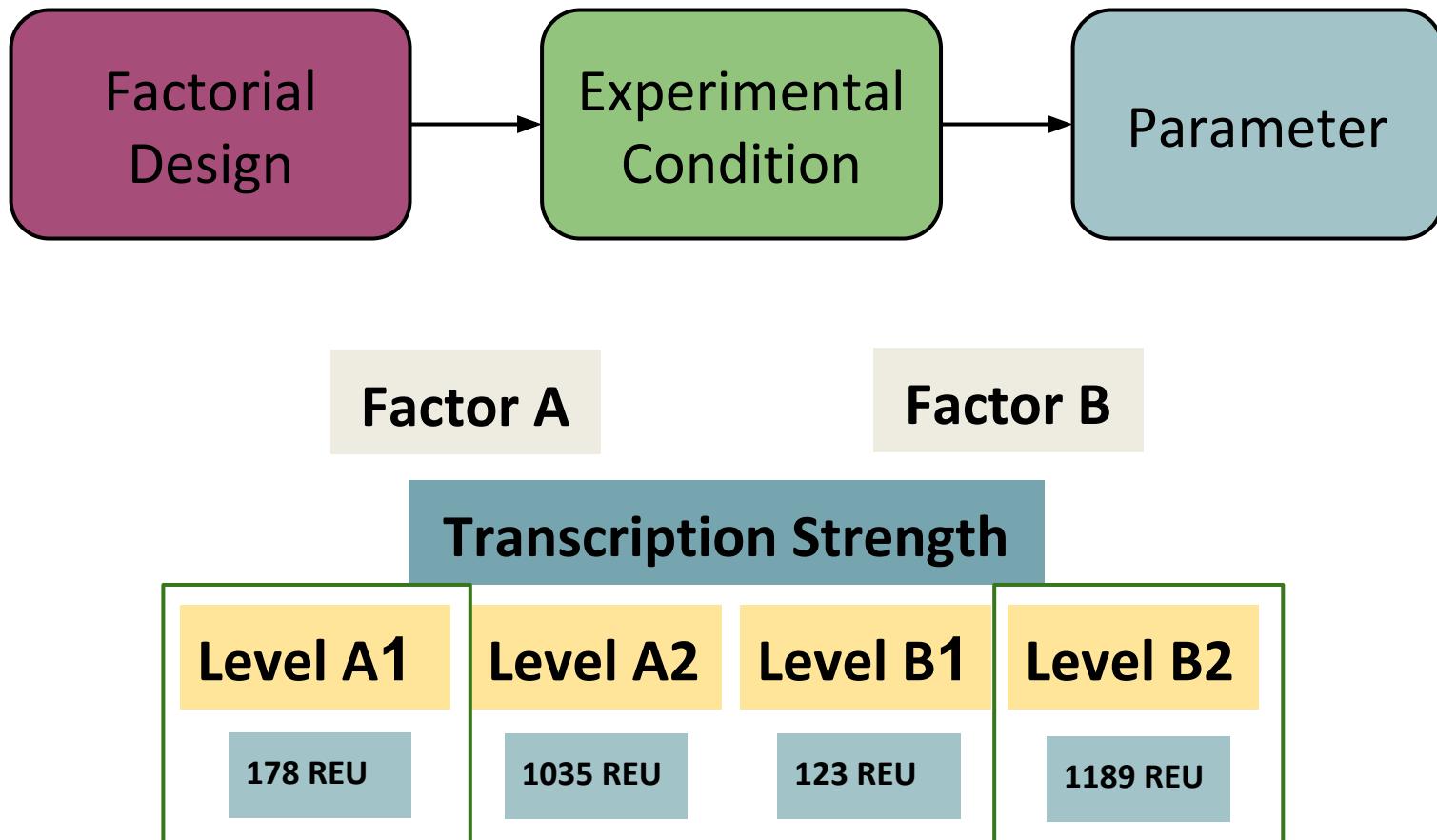
Factorial Design



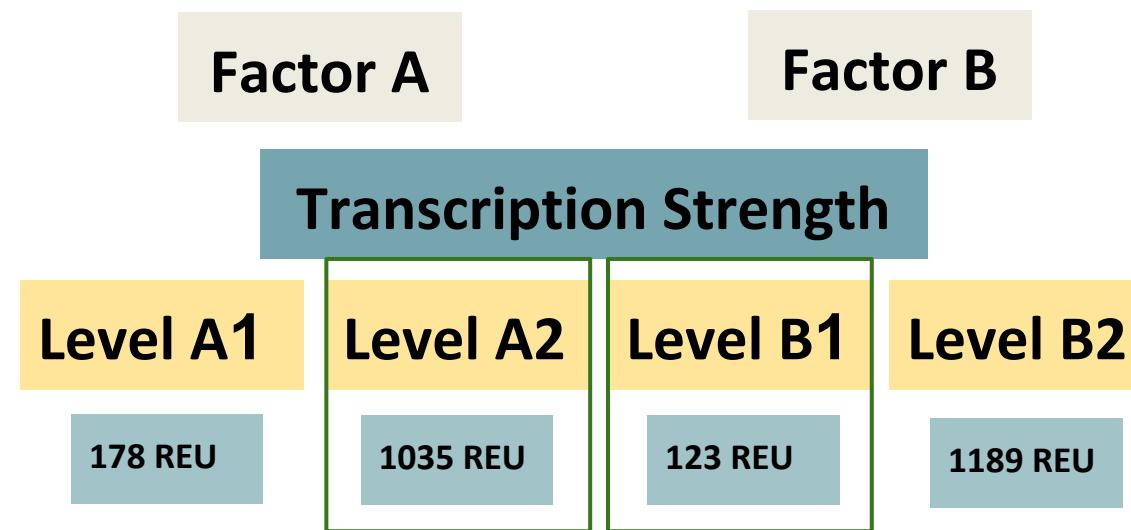
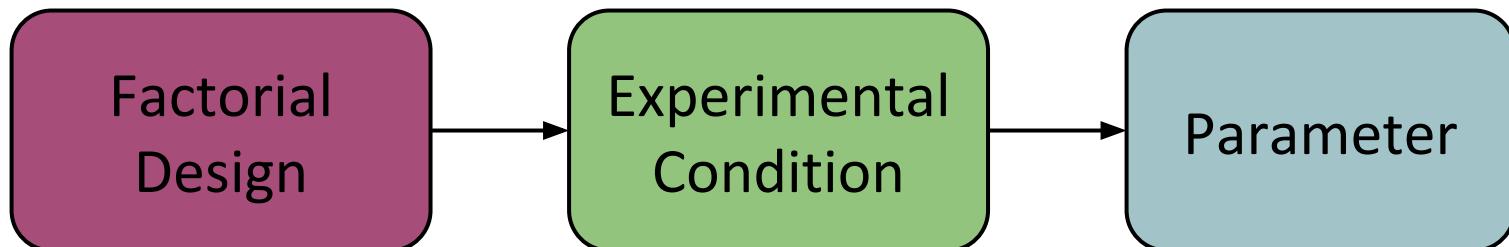
Fractional Factorial Design



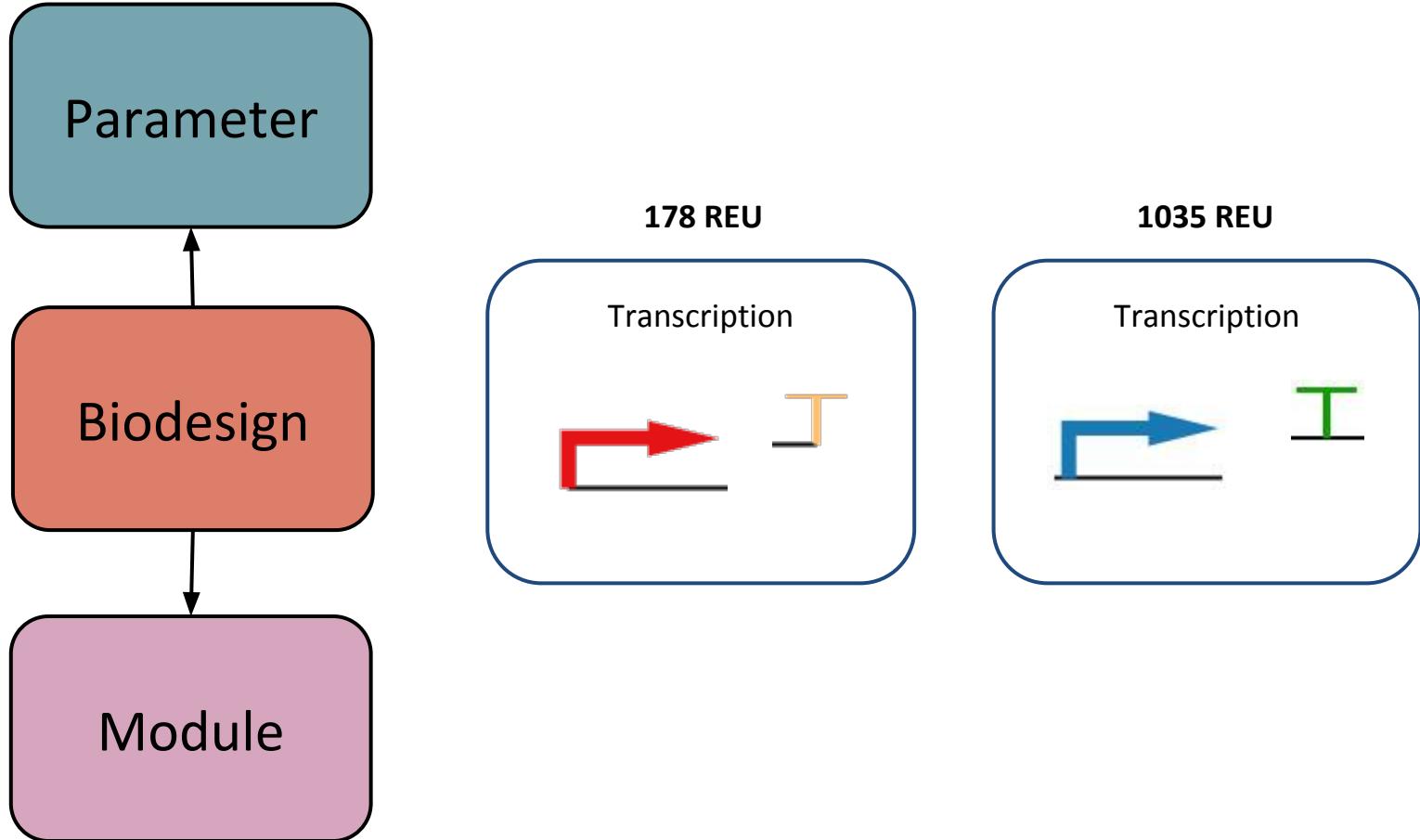
Fractional Factorial Design



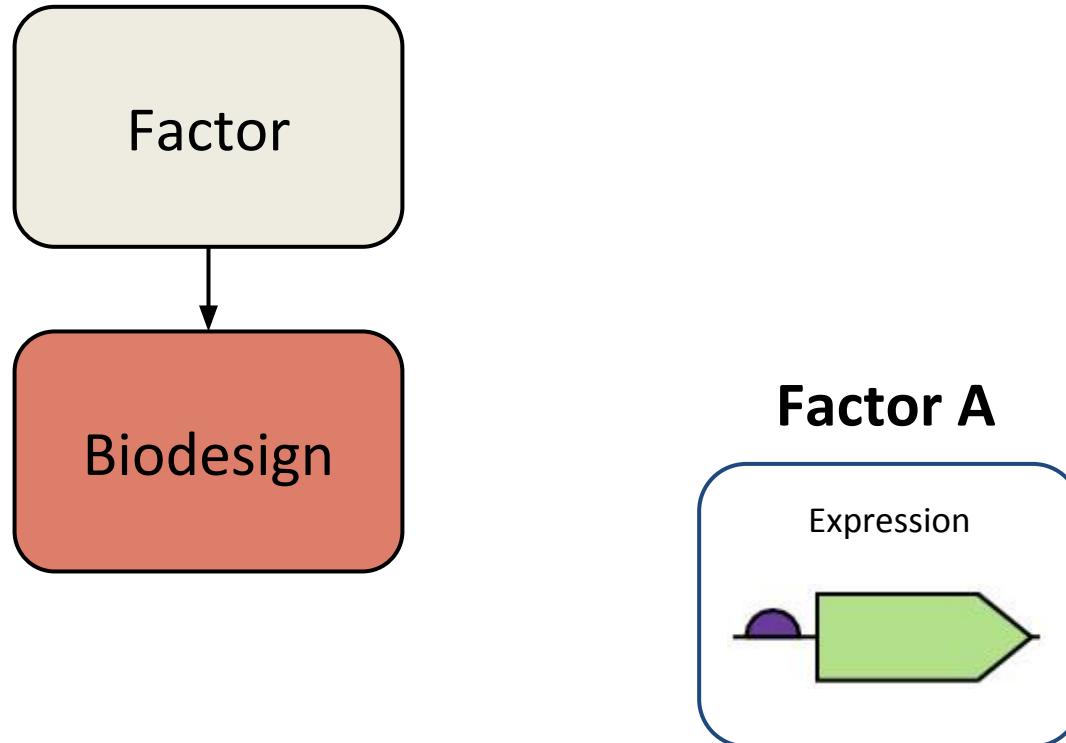
Fractional Factorial Design



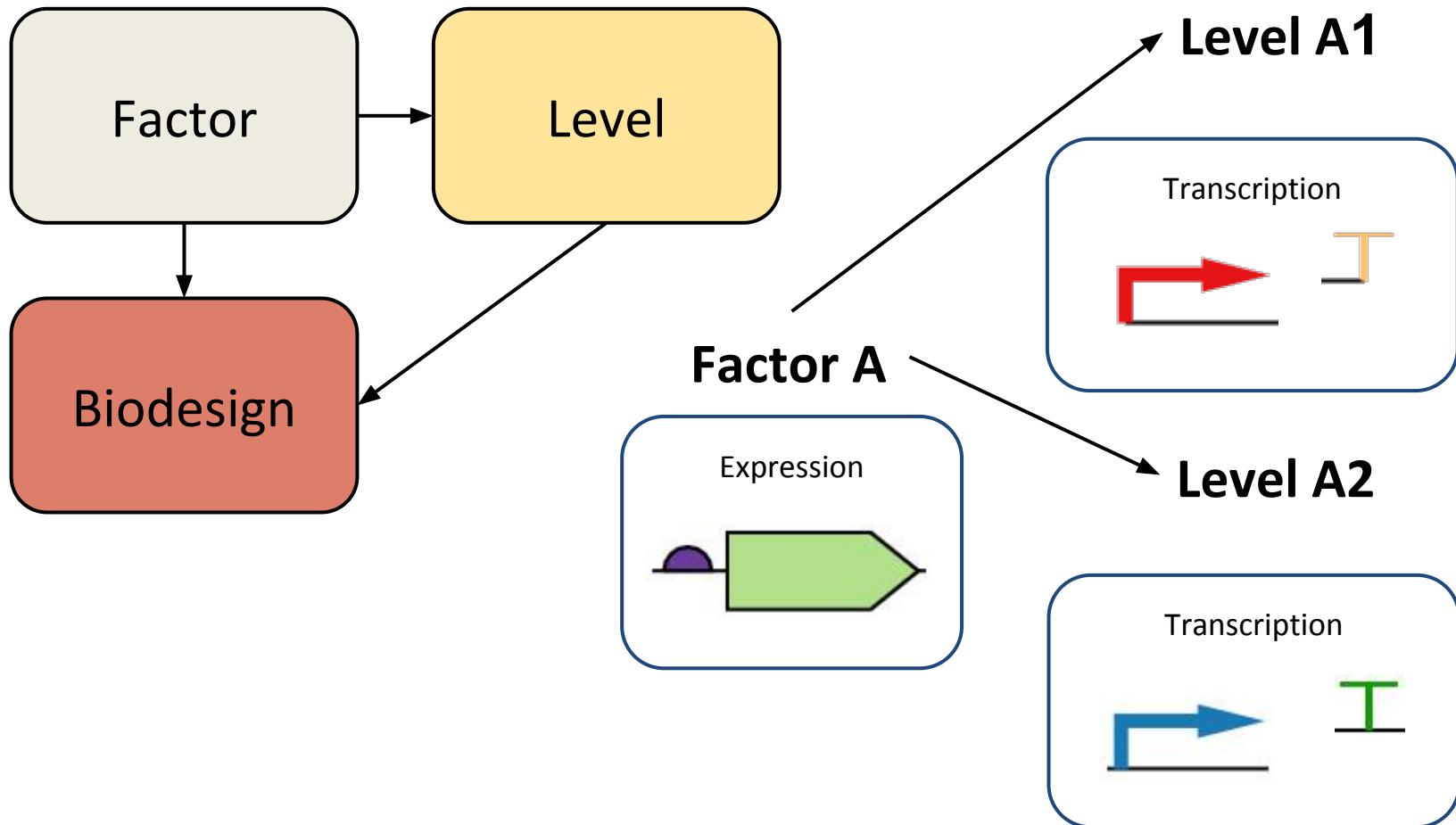
Qual. Function and Performance



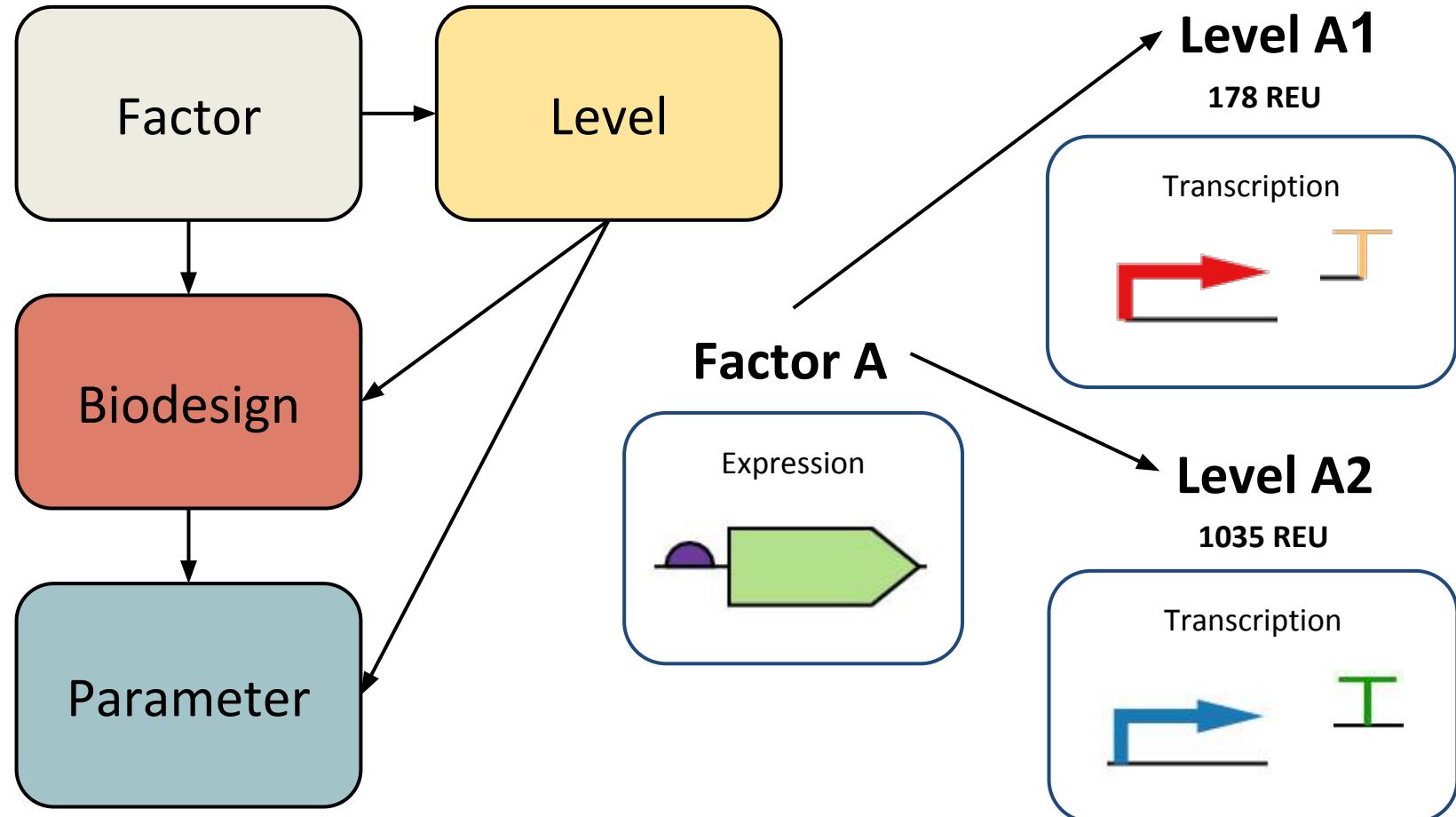
Mapping Exp. Design to Biodesign



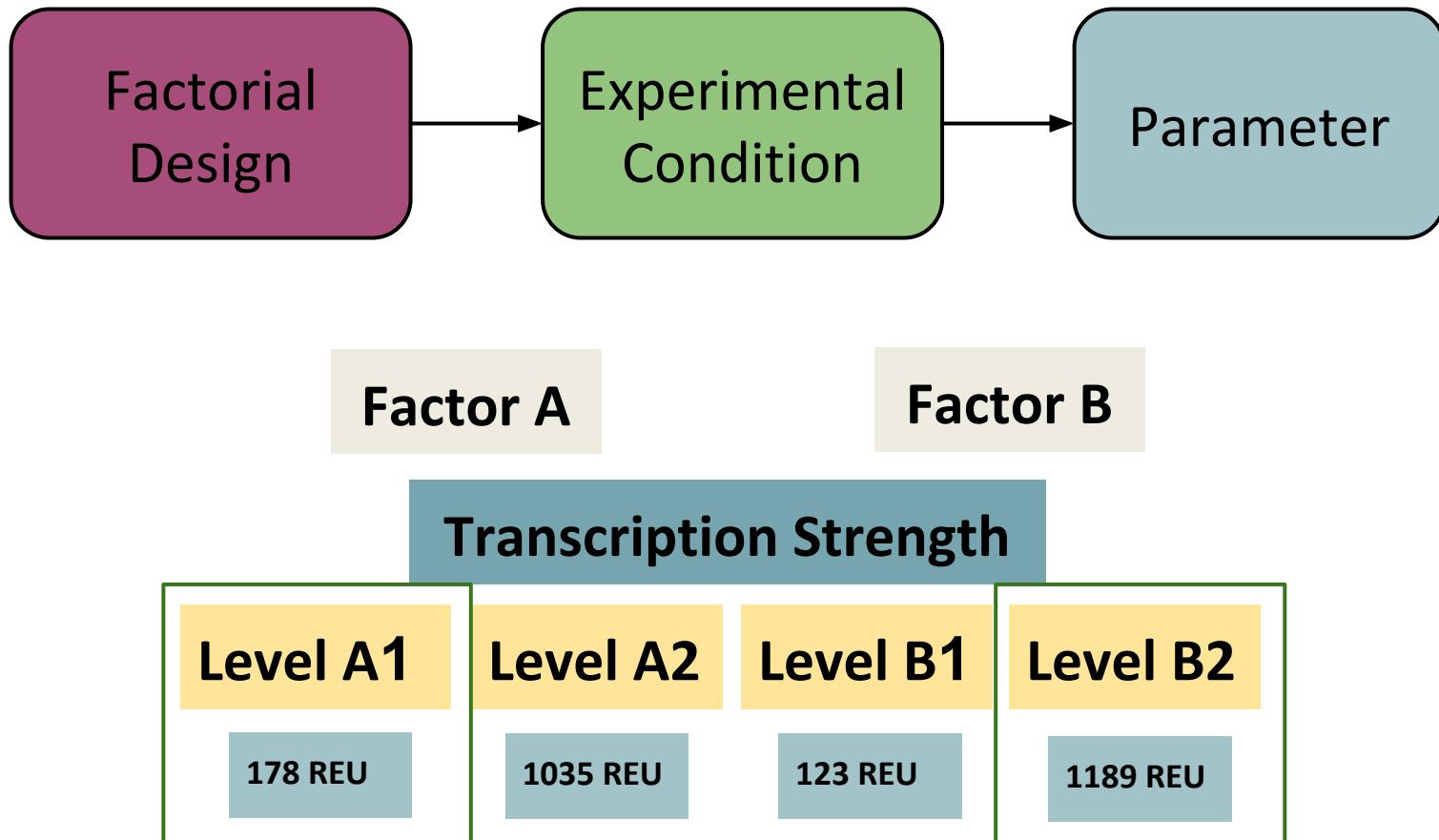
Mapping Exp. Design to Biodesign



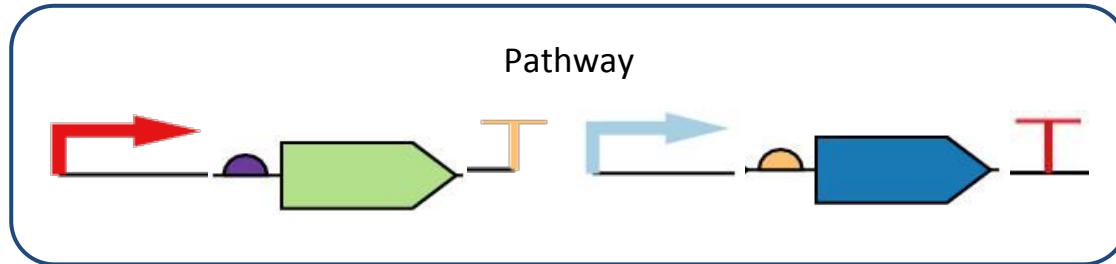
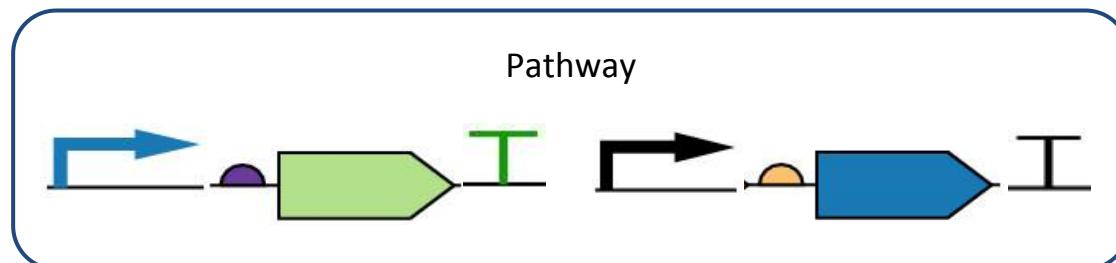
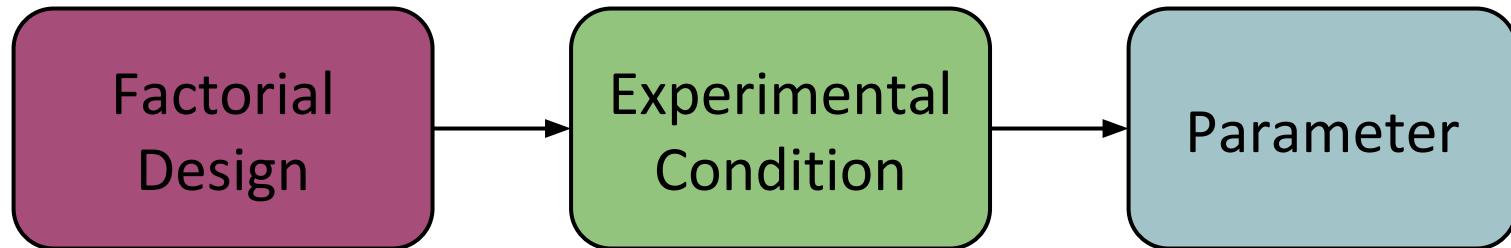
Validation via Common Parameters



Inference of Composite Biodesigns



Inference of Composite Biodesigns

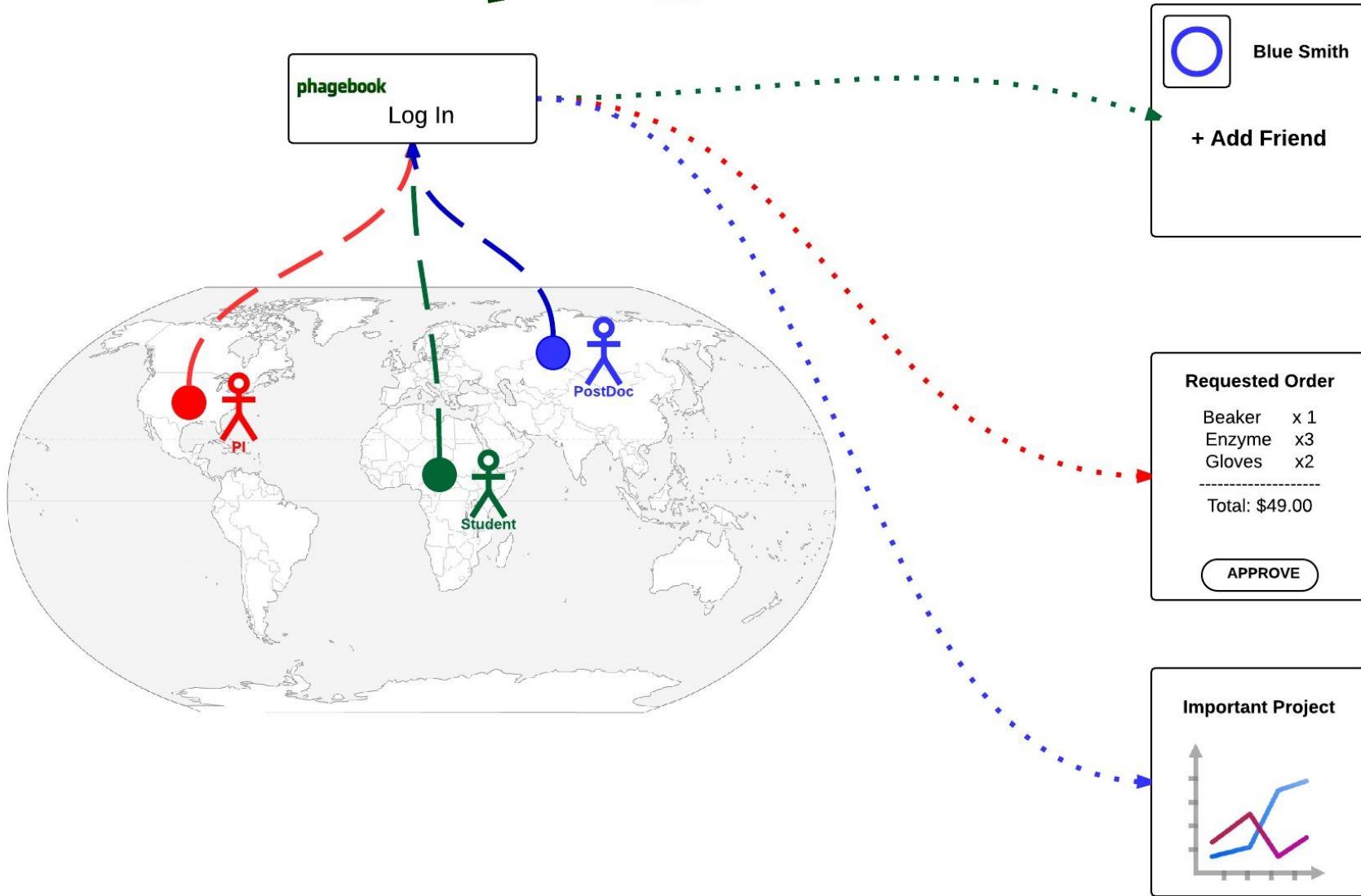


Statistical Design of Experiments

Clotho provides schemas for representing

- Factors and levels of factorial designs
- Variables, parameters, and experimental conditions

phagebook

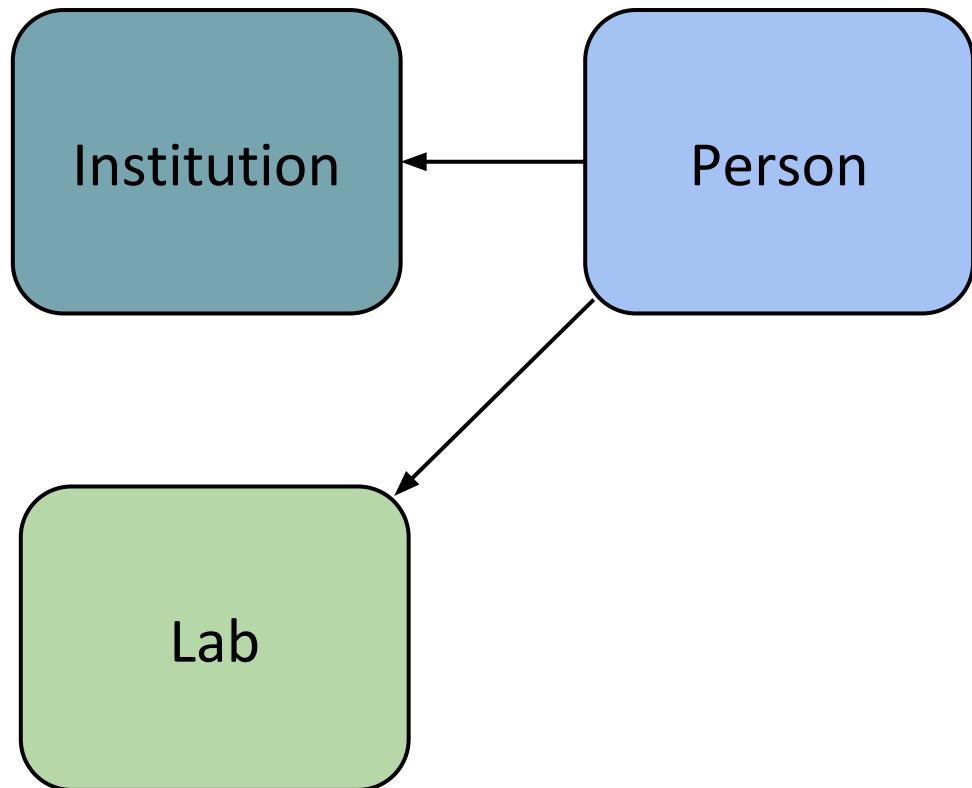


Project Docs and Networking

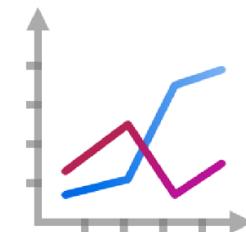
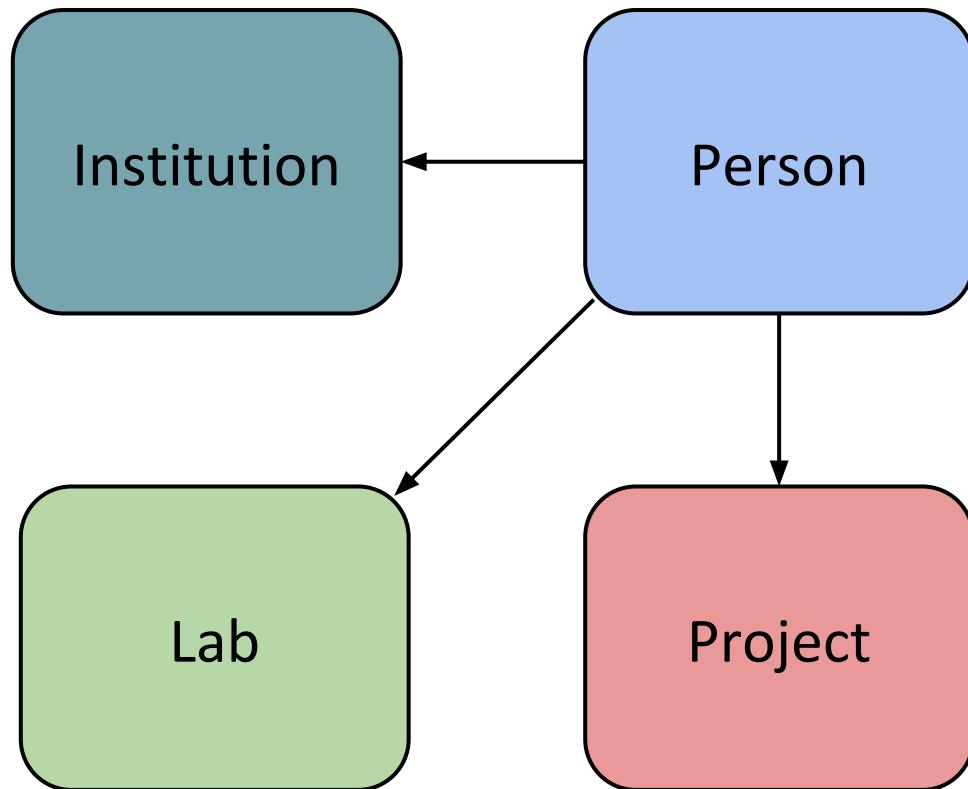
Person



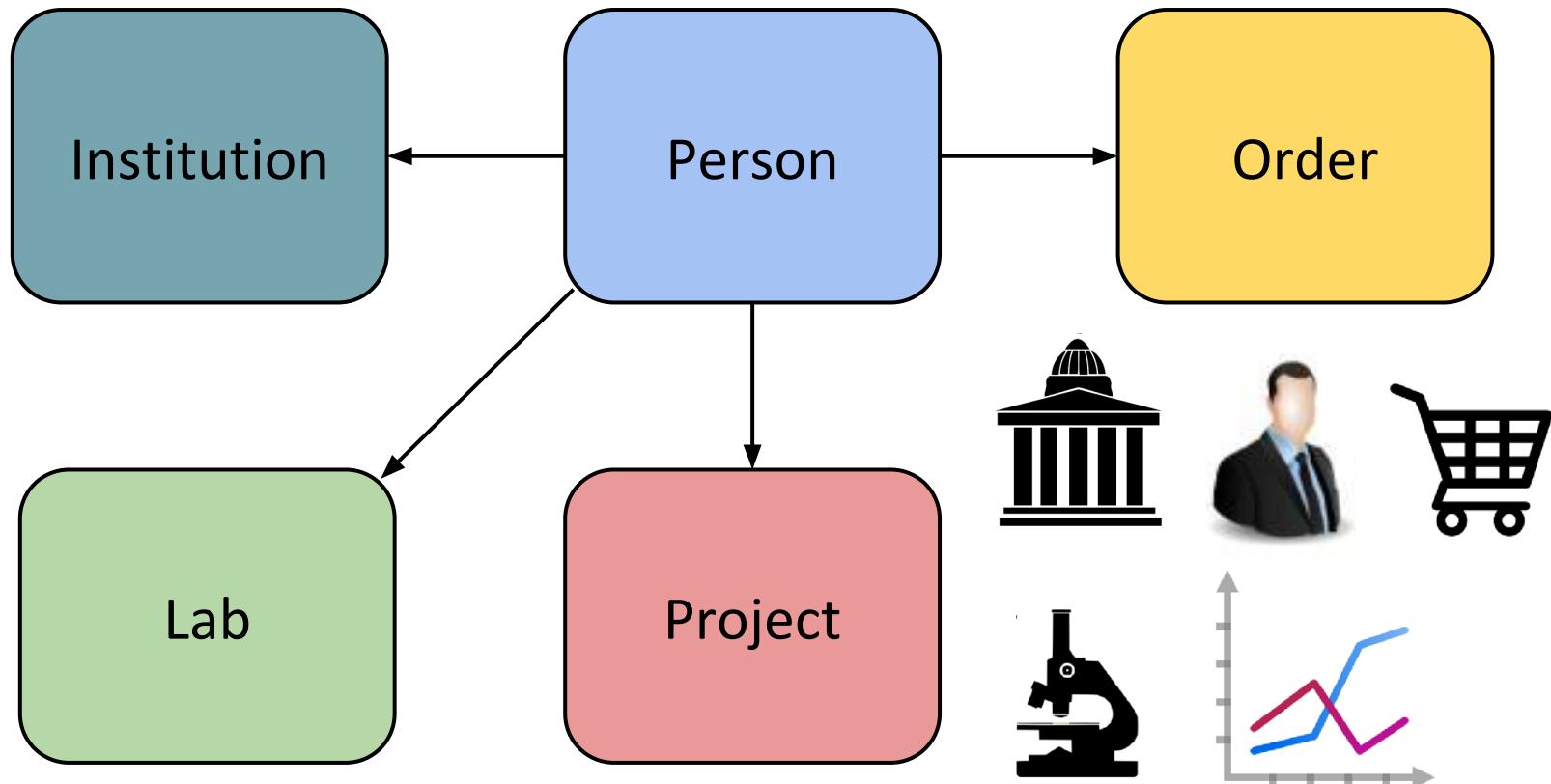
Project Docs and Networking



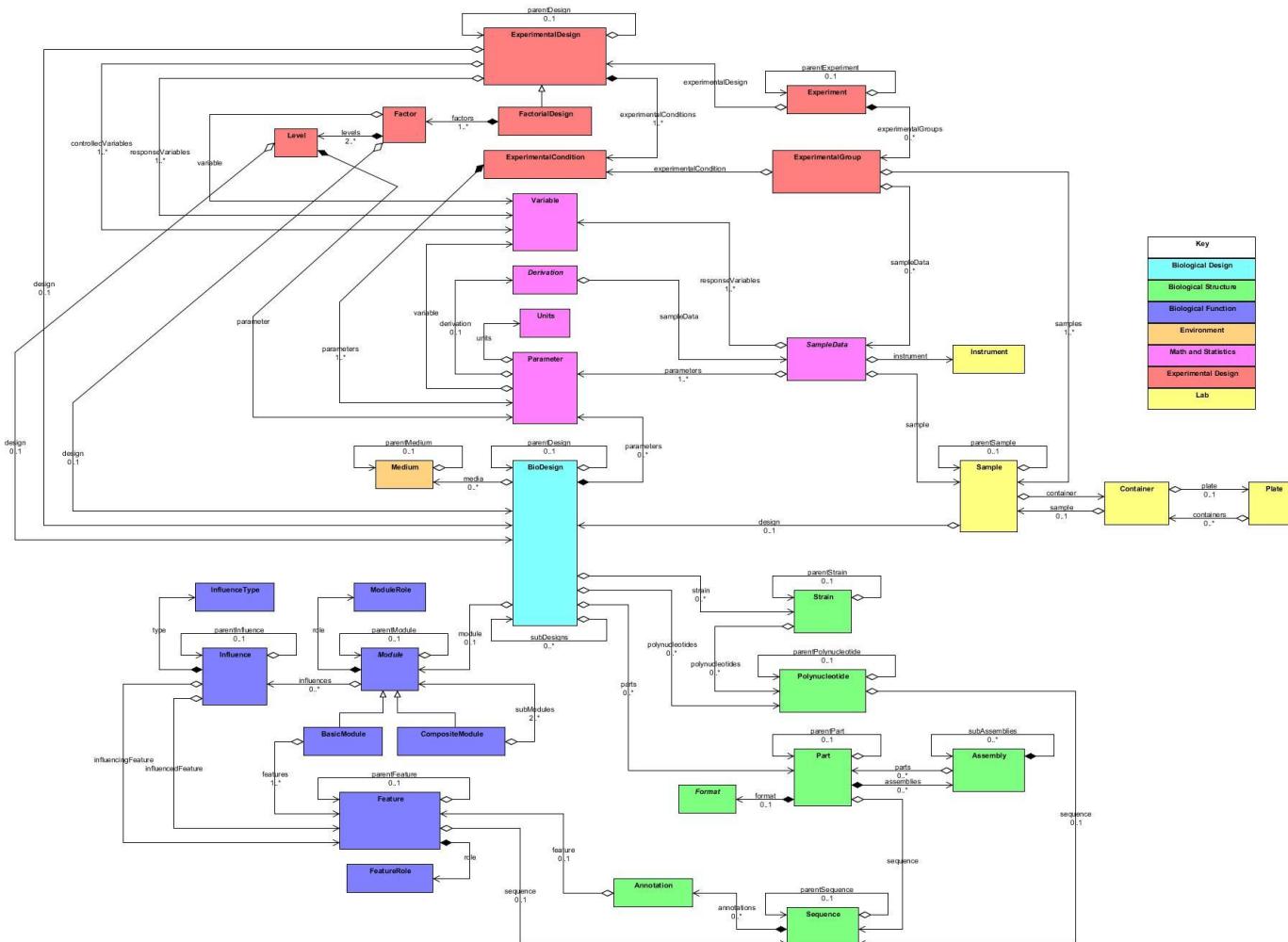
Project Docs and Networking



Project Docs and Networking



Clotho Object Model



Summary

Clotho provides schemas for representing

- Structure and function of biological designs
- Lab records
- Experimental designs
- Math and statistics

Summary

Clotho schemas can applied to a variety of applications, including

- Sequence editing and optimization
- Model fitting and experimental planning
- Statistical design of experiments
- Project documentation and networking

Conclusion

Clotho schemas provide a basic, extensible object model to kickstart the development of applications for synthetic biology and make it easier to persist data across different projects.

Acknowledgements

- National Science Foundation
Grant #1147158
- Rafik B. Hariri Institute for
Computing and Computational
Science & Engineering
 - Azer Bestavros
 - Linda Grosser
 - Lauren Dupree

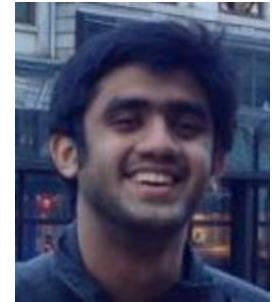
- Densmore Lab



Douglas
Densmore



Stephanie
Paige



Prashant
Vaidyanathan

- Anderson Lab



J. Christopher
Anderson



Bill
Cao



Mina
Li

- Phagebook Team