

*E. coli* Automatic Directed  
Evolution Machine (E.ADEM):  
A Universal Framework

姜浩

2009.4.18

# How to Engineer Genetically Engineered Machines?

Input

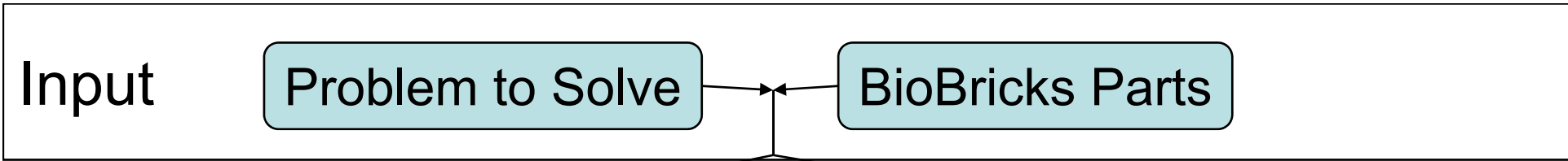
Problem to Solve

BioBricks Parts

Output

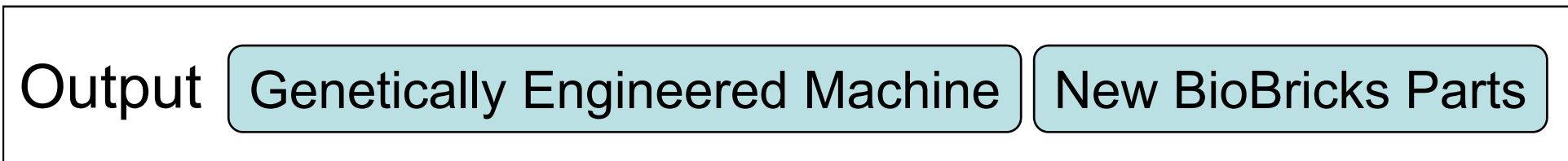
Genetically Engineered Machine

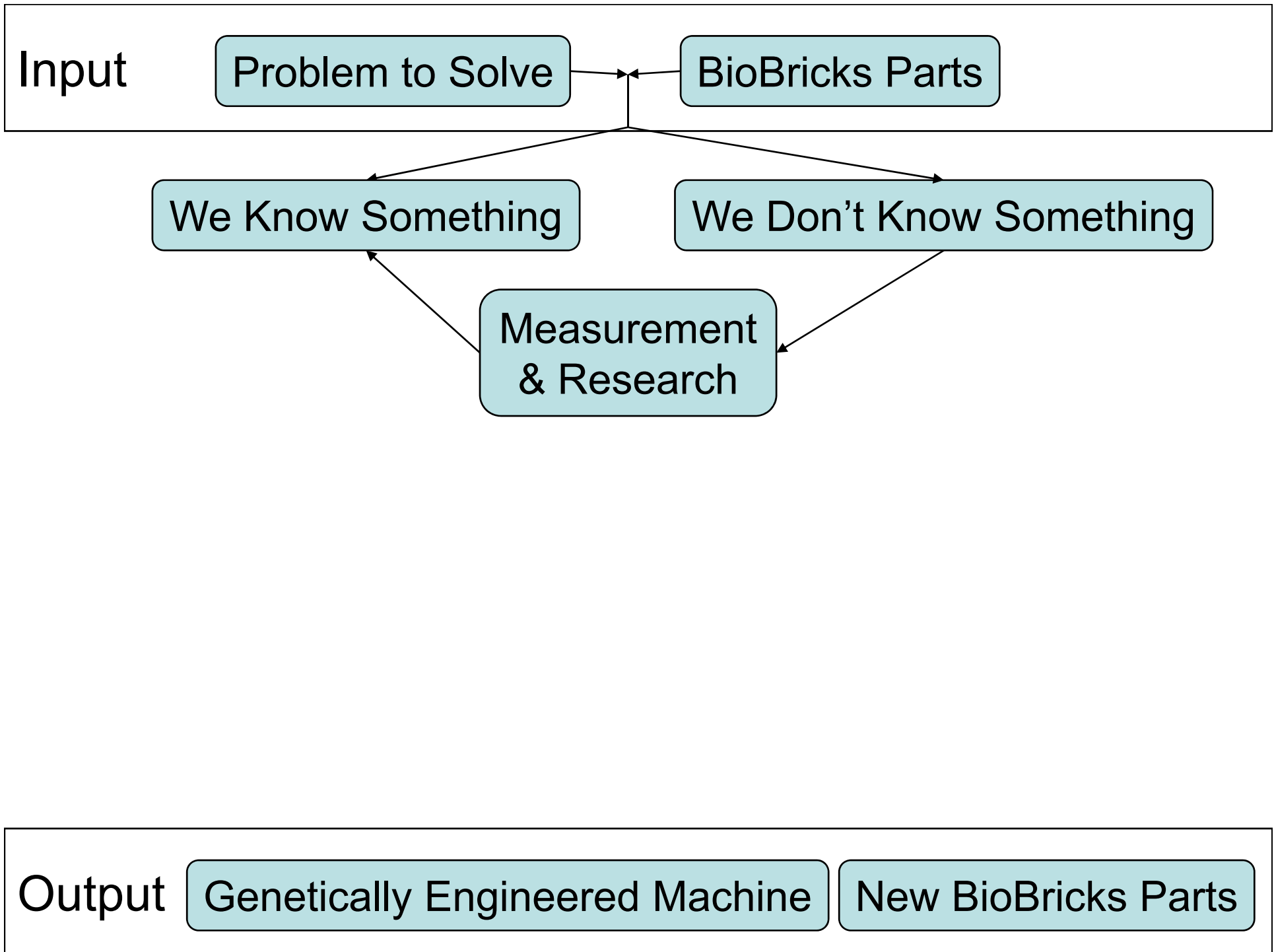
New BioBricks Parts

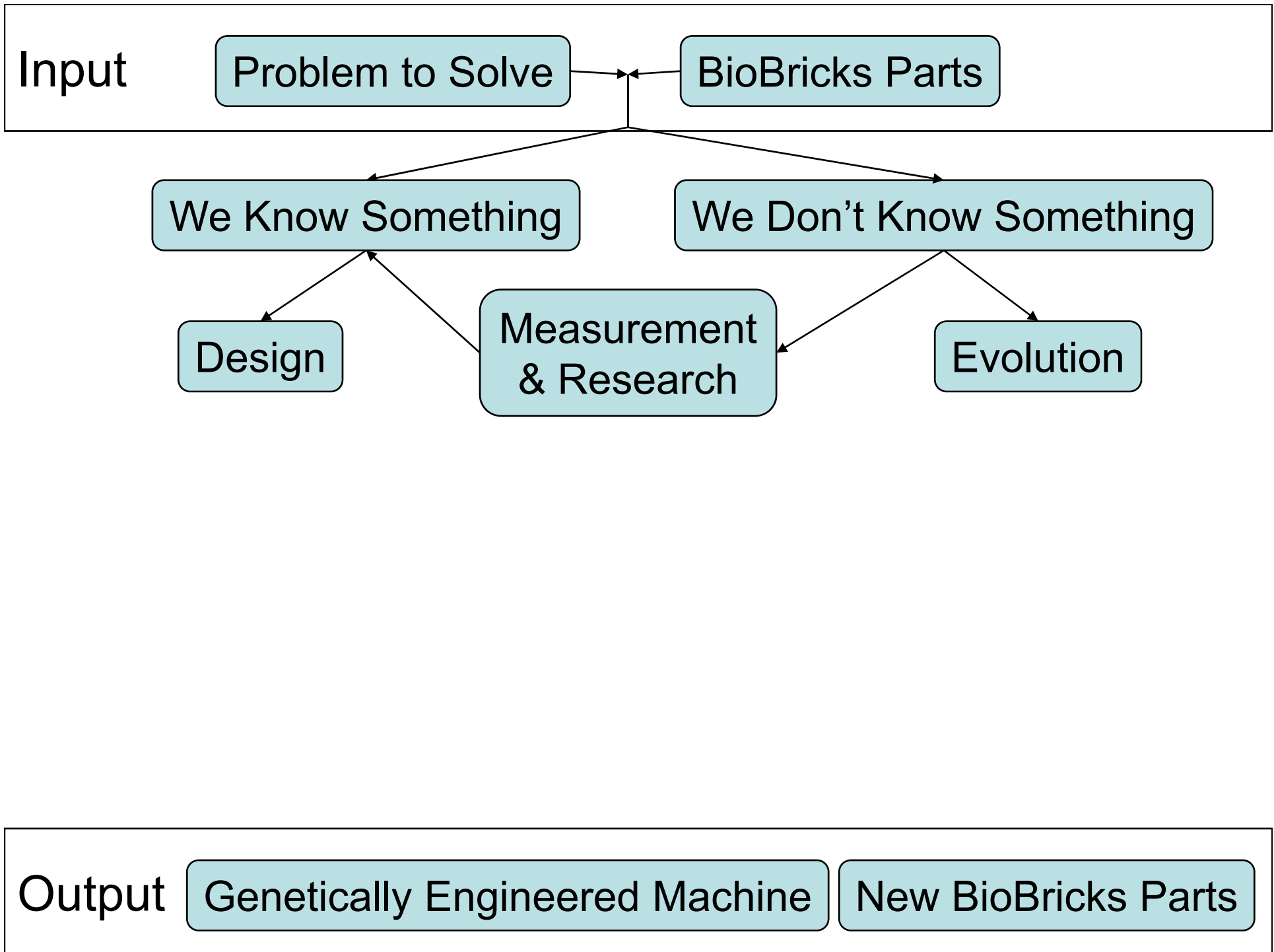


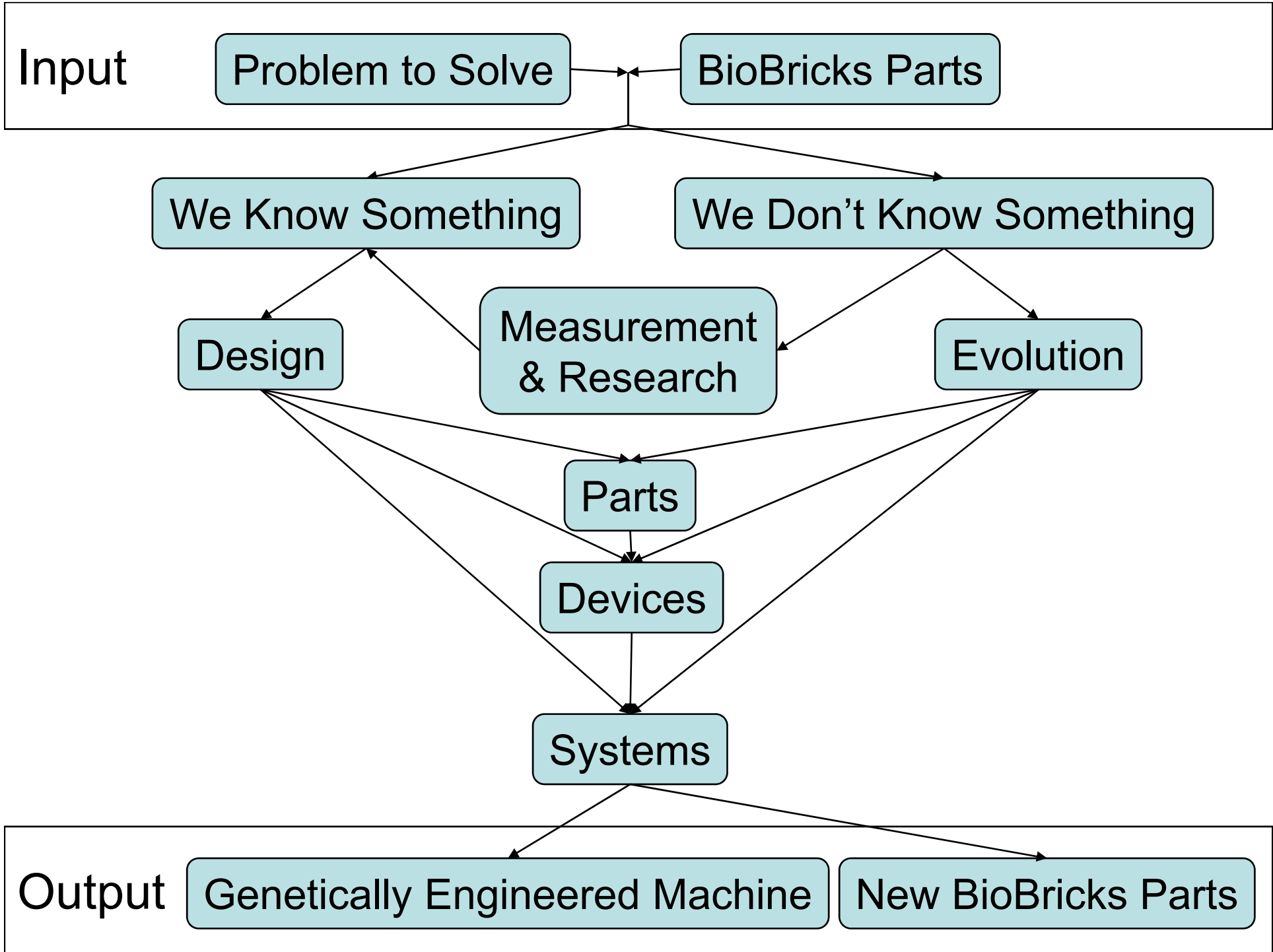
We Know Something

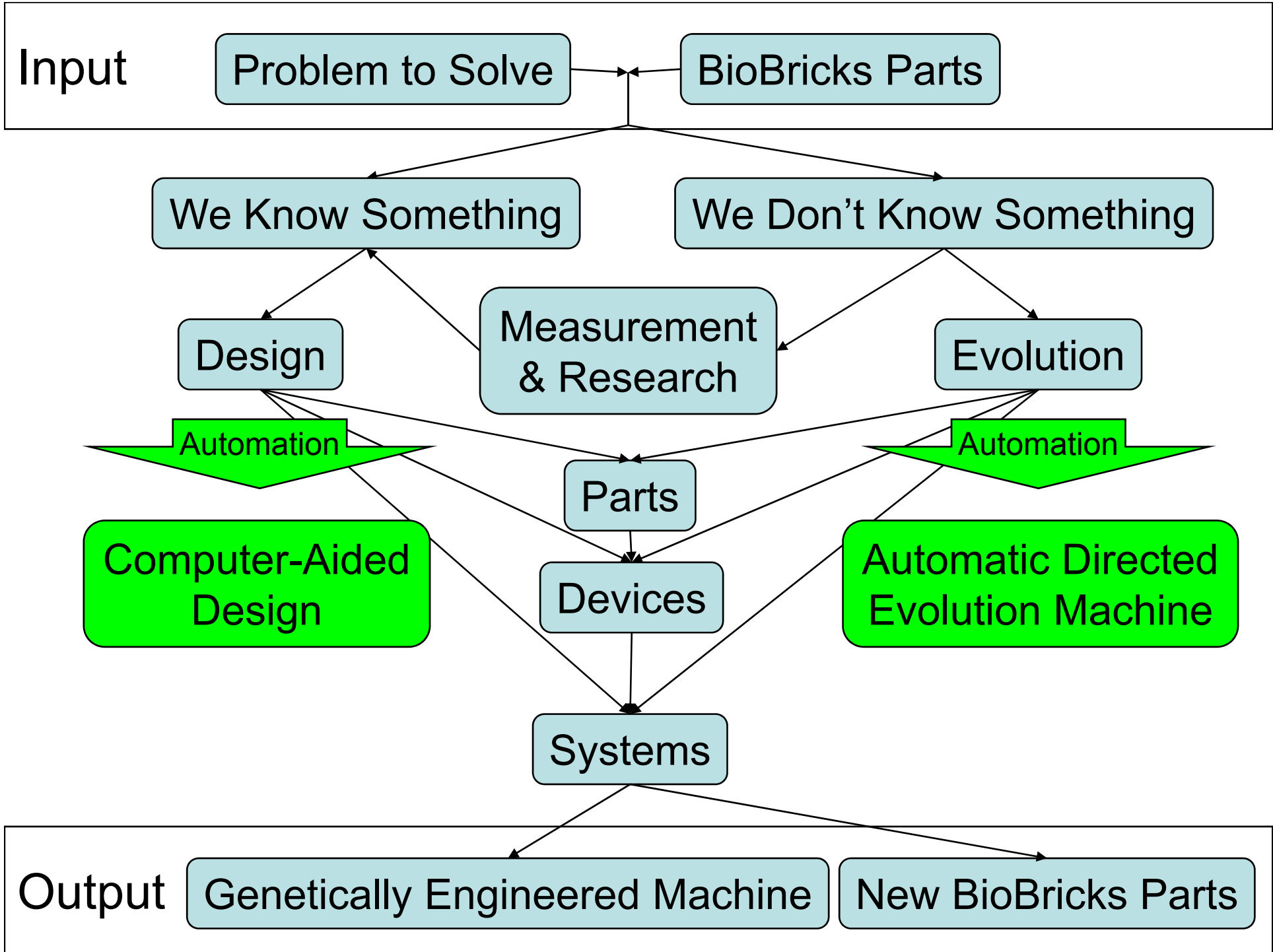
We Don't Know Something



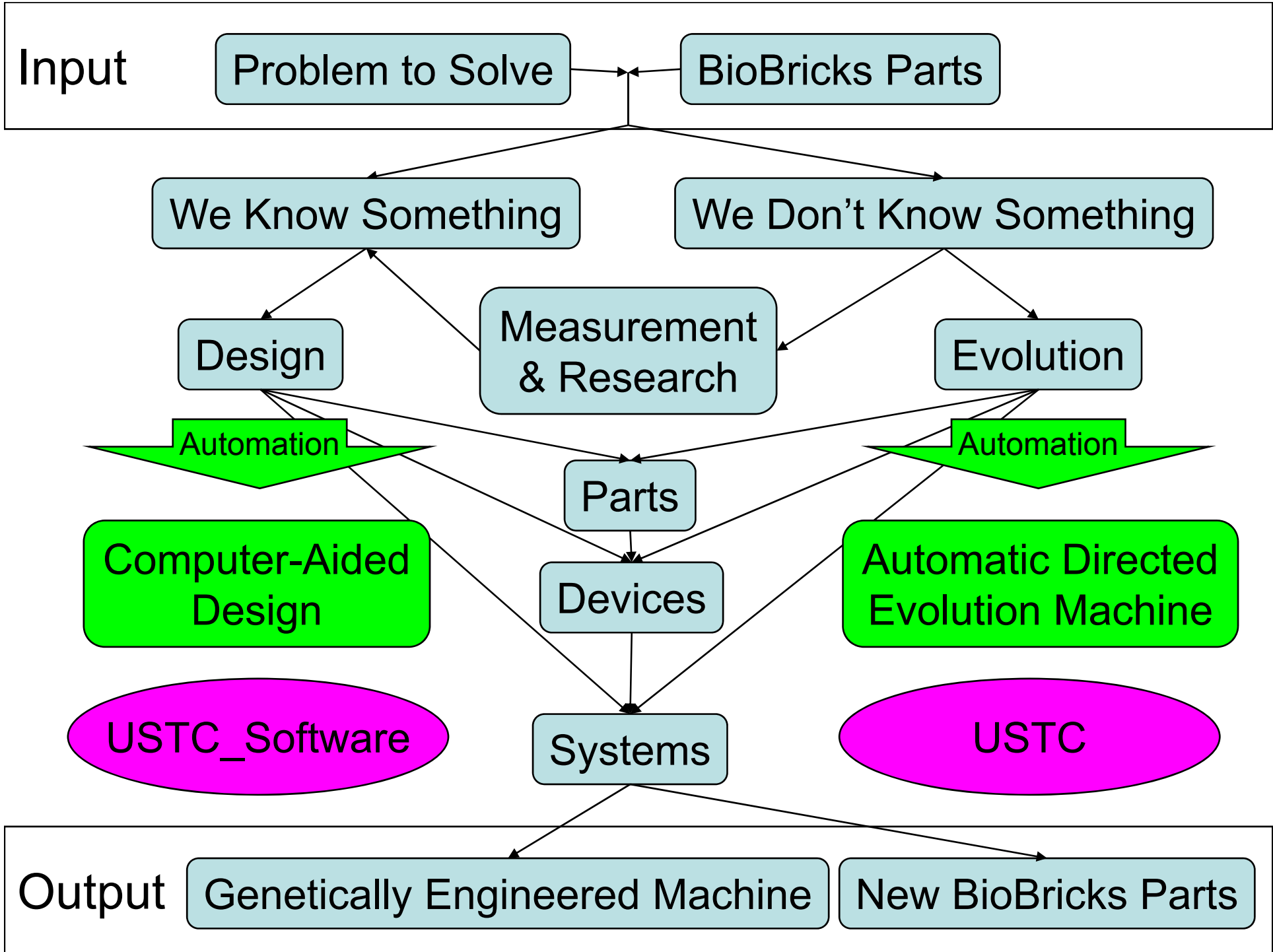






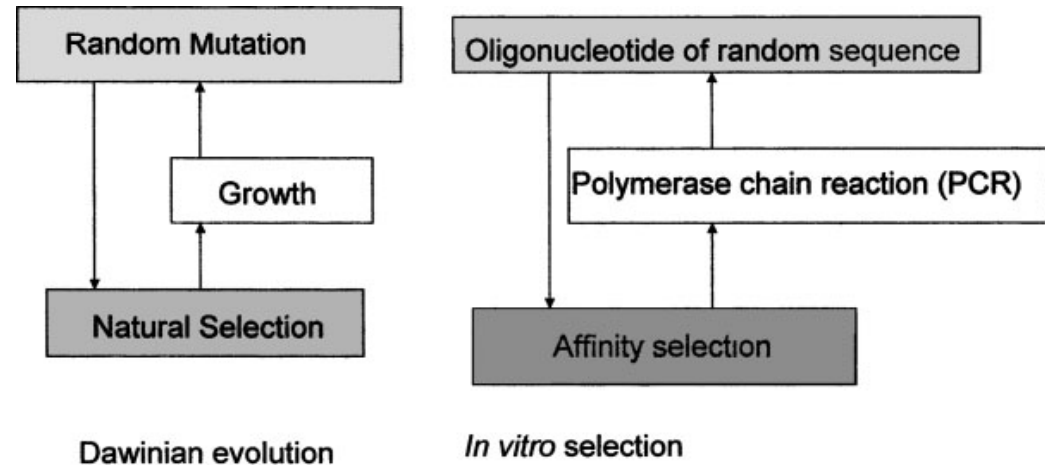




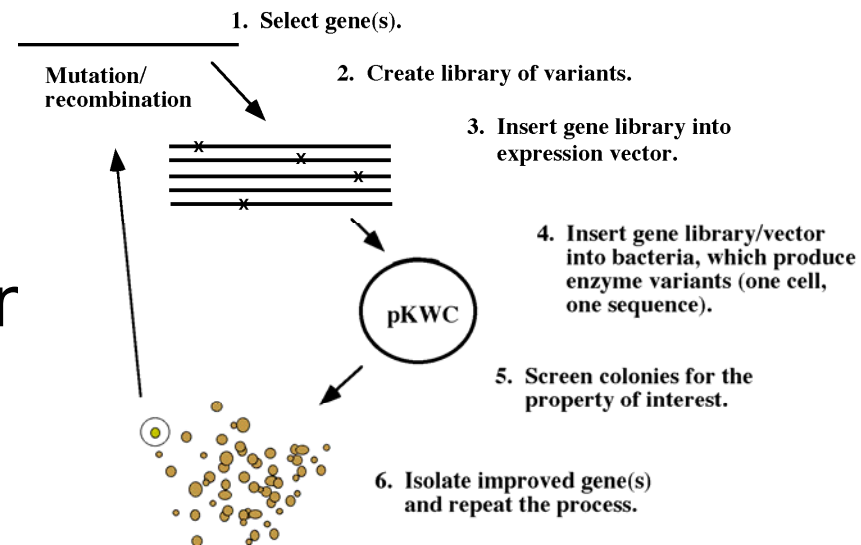


# Directed Evolution

- Target
  - Enzymes
  - Aptamers
  - Ribozymes
  - Synthetic genetic circuit components



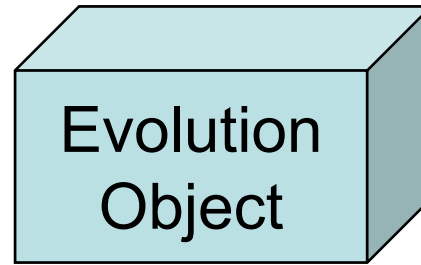
- Diversification
  - Error-prone PCR
  - DNA shuffling
- Screening or selection
  - *In vitro*
  - *In vivo*



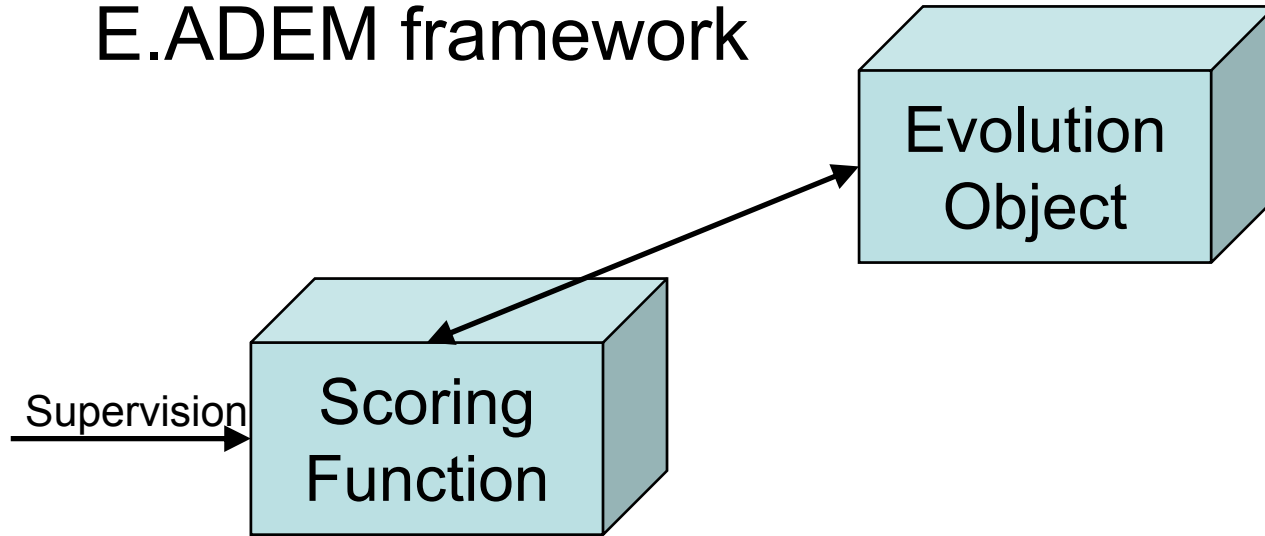
# Problems

- Labor-intensive
- Cost-intensive
- Limited library size
- Noncontinuous
- Conditions optimization needed
- ...
  
- Solve: automatic & *in vivo*

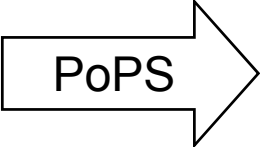
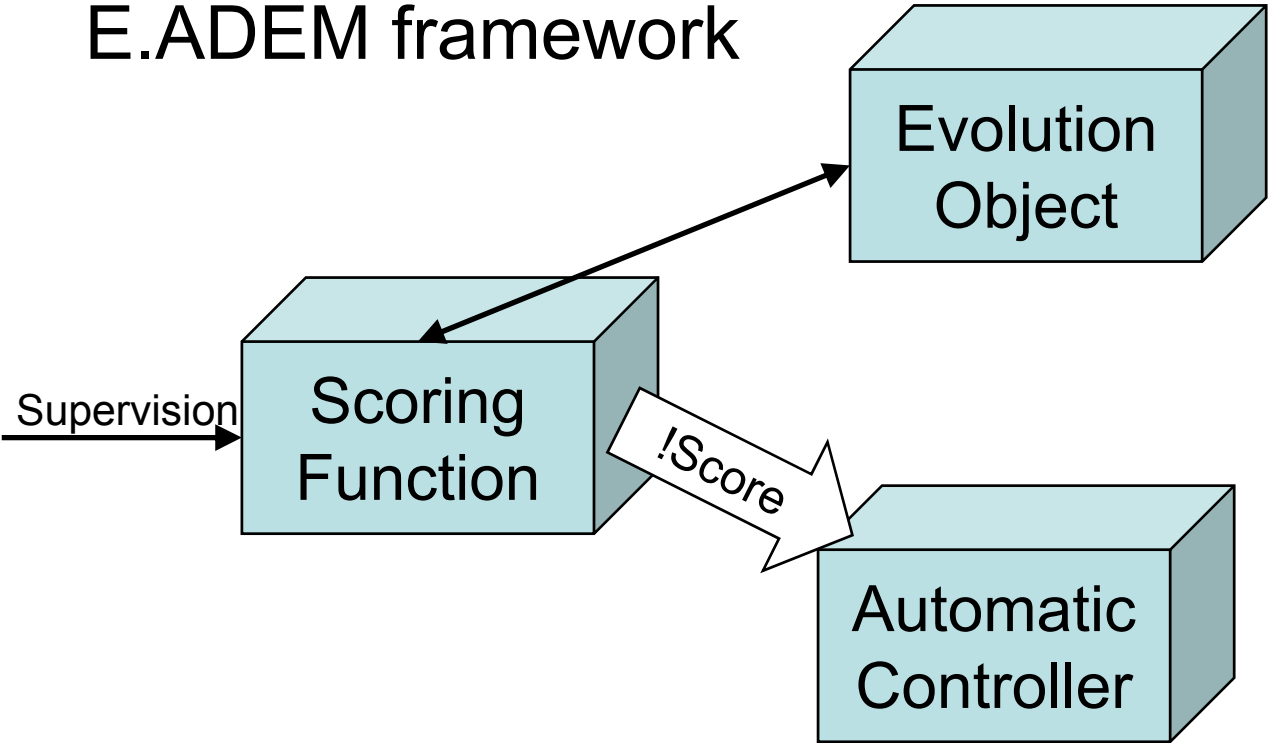
E.ADEM framework



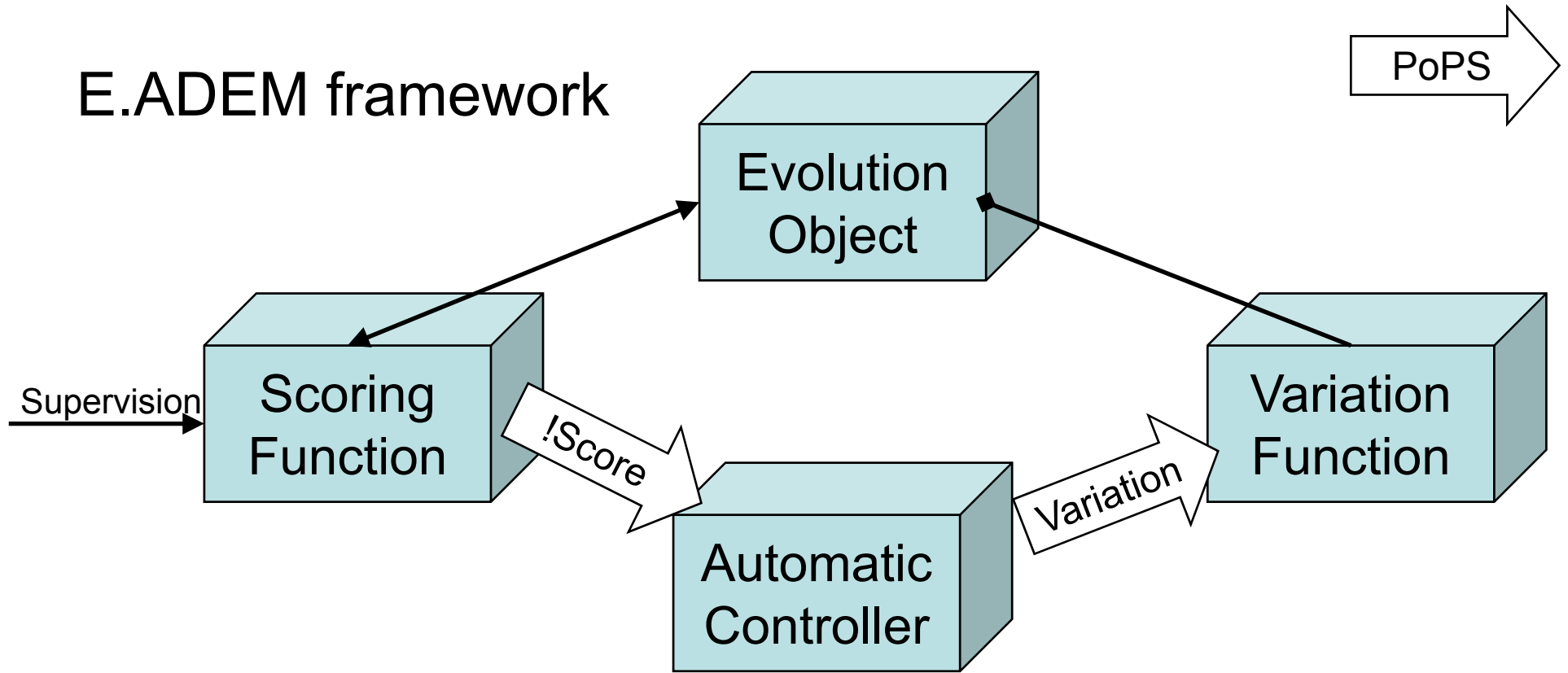
# E.ADEM framework



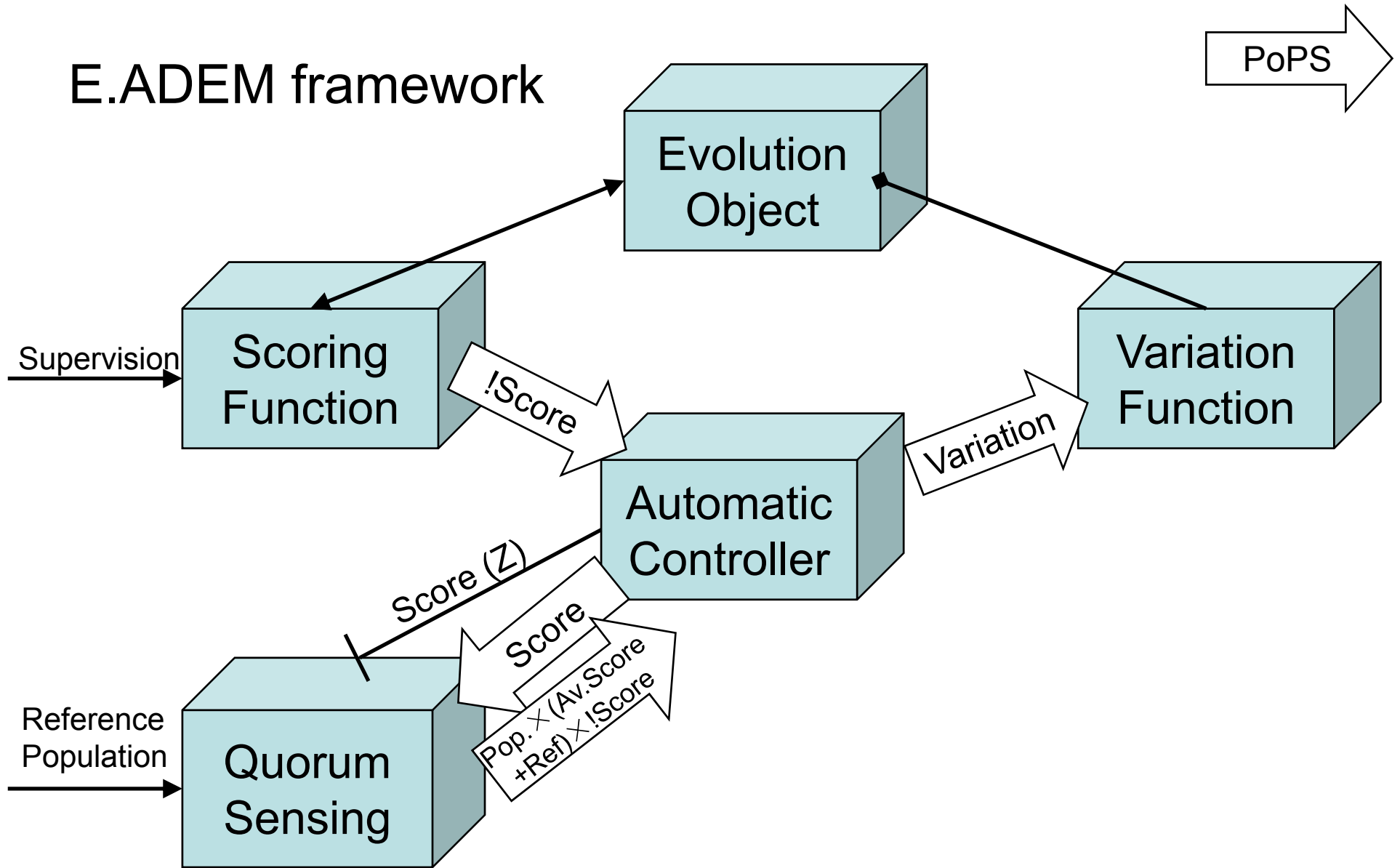
# E.ADEM framework



# E.ADEM framework

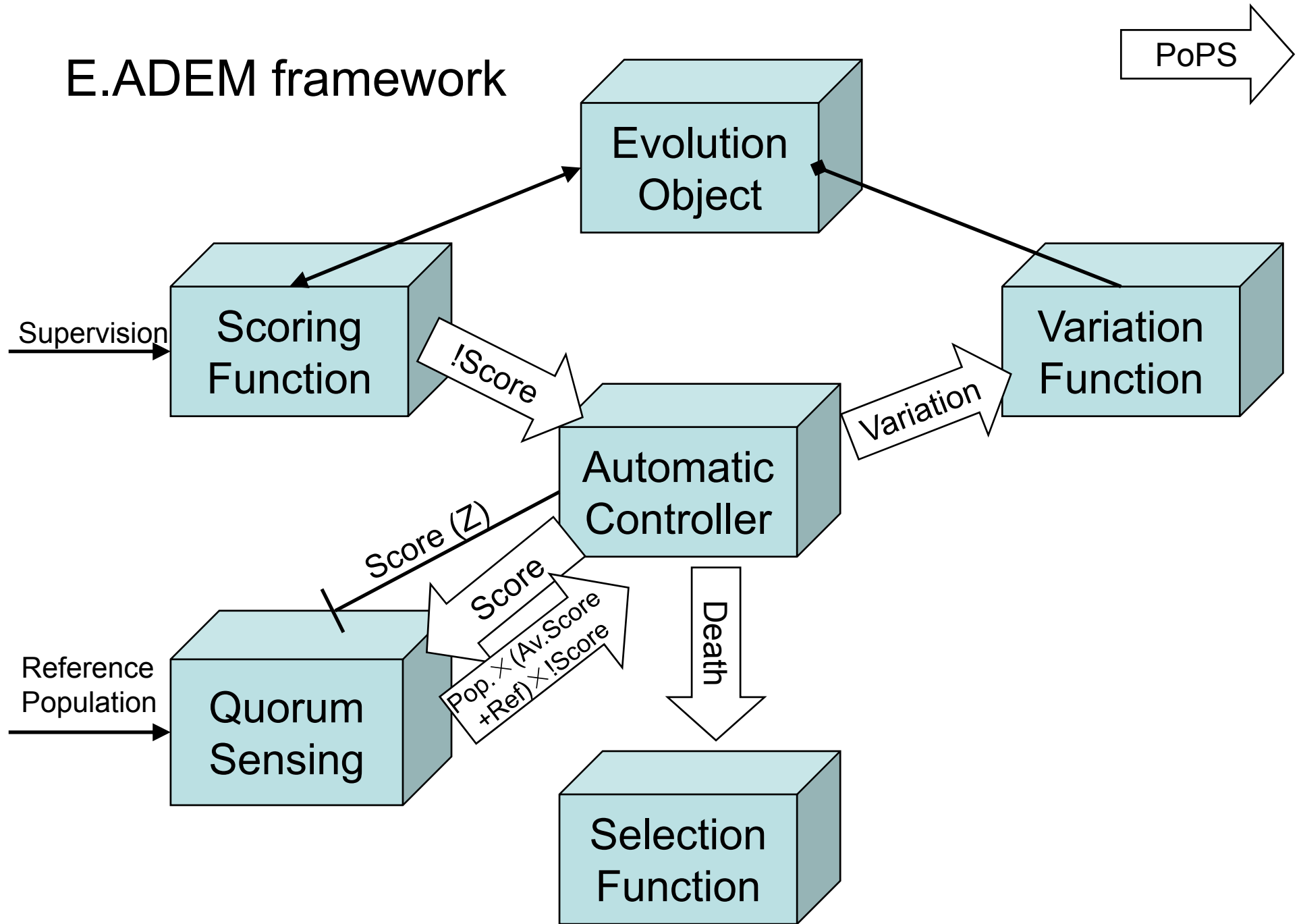


# E.ADEM framework

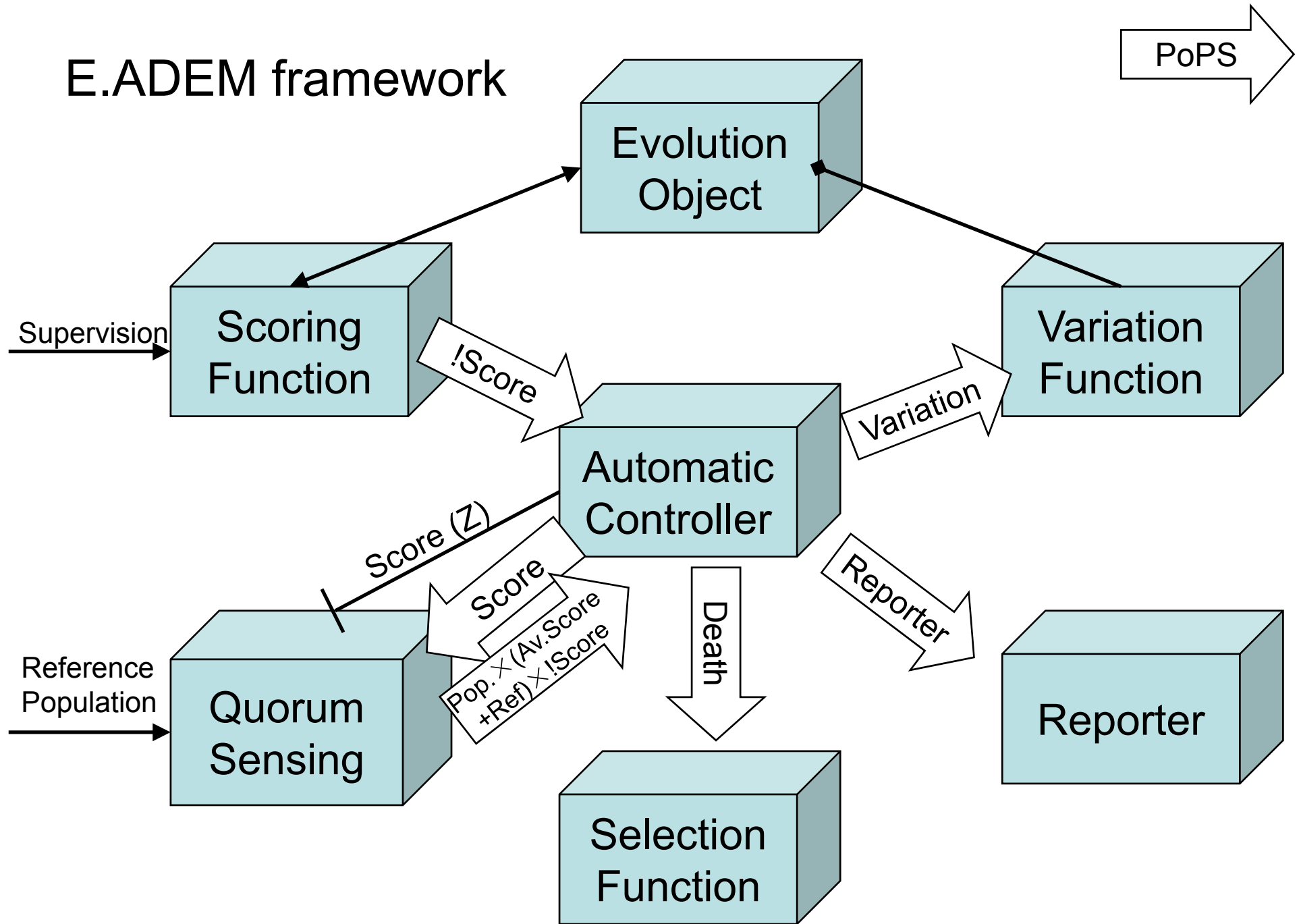




# E.ADEM framework



# E.ADEM framework



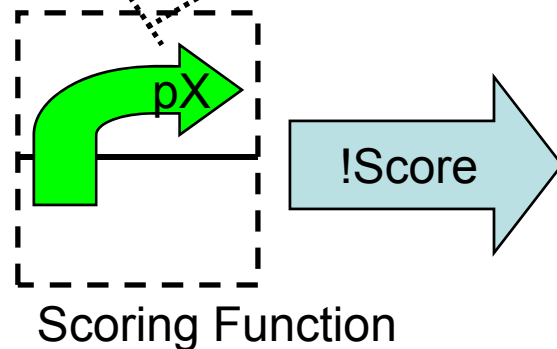
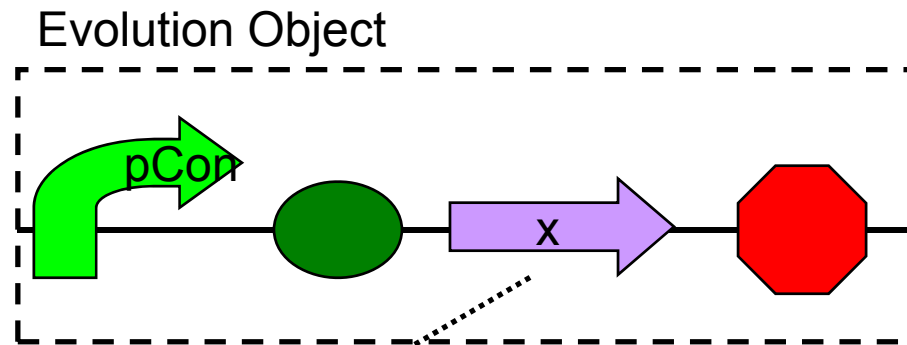
# Evolution Object

- Parts
  - Protein
    - Regulator
    - Receptor
    - Binding Partner
    - Enzyme
  - RNA
    - Regulator
    - Aptamer
    - Ribozyme
    - RBS
  - DNA
    - Promoter
    - Origami
- Devices
  - Sensor
  - Logic device
  - Signaling
  - Protein generator
  - Reporter
- Systems
  - Genome
  - Metabolic pathway
  - Signal transduction pathway
  - Logic circuit
  - Complex system simulation
  - Mathematic problem

# Scoring Function

- Genotype →  
Phenotype →  
Signal Transduction →  
Transcription Rate (PoPS) →  
Universal Interface
- Functional Object
- Response Object
  - e.g. Sensor, Logic Device
  - Supervised Learning

# Transcription Repressor



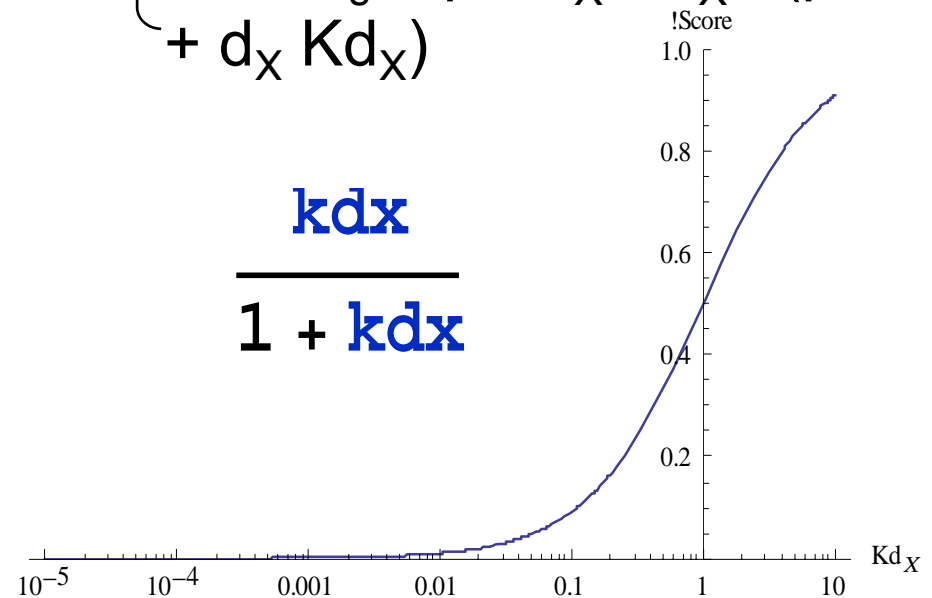
$$X' = pCon - d_x X$$

$$!Score = pX Kd_x / (X + Kd_x)$$

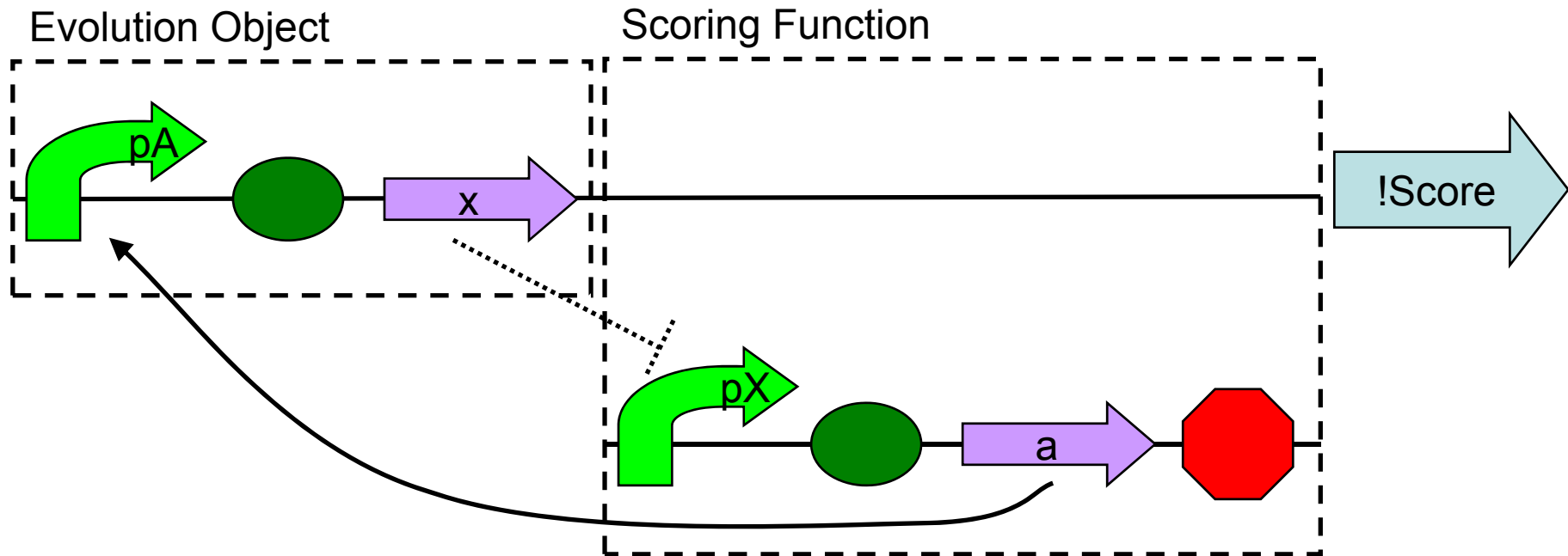
$$X_s = pCon / d_x$$

$$!Score_s = pX d_x Kd_x / (pCon + d_x Kd_x)$$

$$\frac{kdx}{1 + kdx}$$



# But it is not so simple...

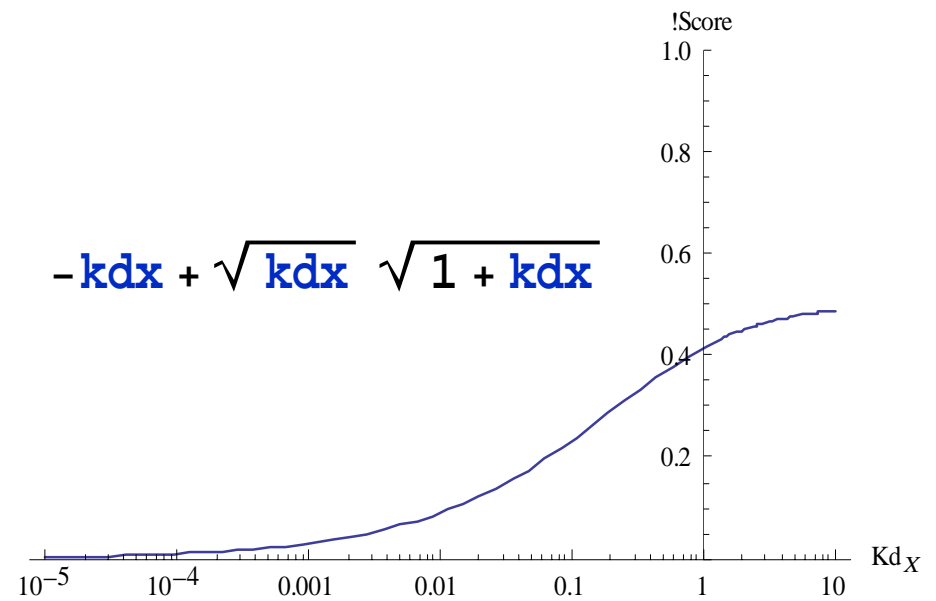
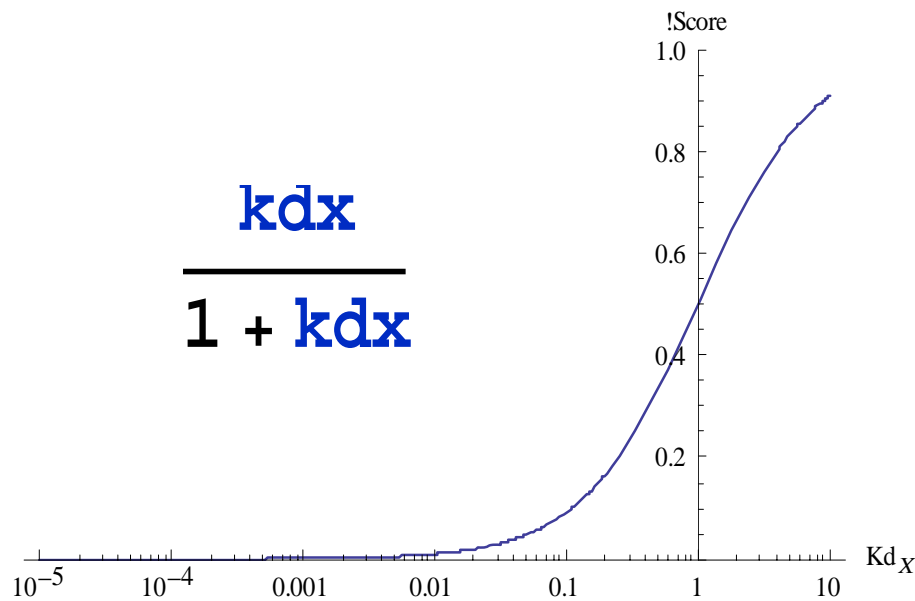


$$\begin{cases} X' = !Score - d_X X \\ A' = pX Kd_X / (X + Kd_X) - d_A A \\ !Score = pA A / (A + Kd_A) \end{cases}$$

$$nScore \rightarrow pa - \frac{2 da kda pa (dx kdx + pa)}{da kda (dx kdx + 2 pa) + dx kdx px + \sqrt{dx} \sqrt{kdx} \sqrt{da^2 dx kda^2 kdx + 2 da kda (dx kdx + 2 pa) px + dx kdx px^2}}$$

$$x \rightarrow \frac{pa - \frac{2 da kda pa (dx kdx + pa)}{da kda (dx kdx + 2 pa) + dx kdx px + \sqrt{dx} \sqrt{kdx} \sqrt{da^2 dx kda^2 kdx + 2 da kda (dx kdx + 2 pa) px + dx kdx px^2}}{dx}}$$

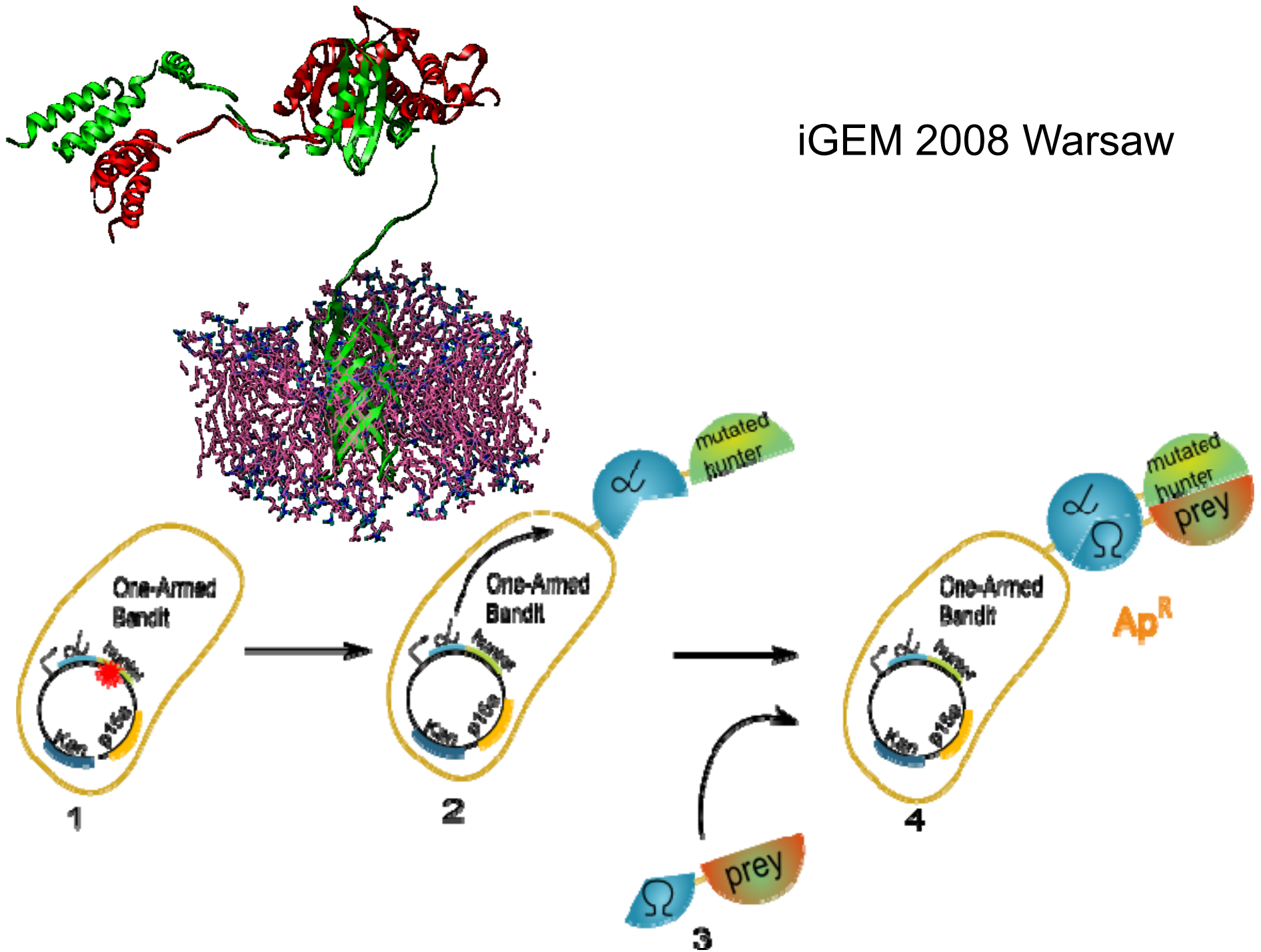
$$a \rightarrow \frac{-da dx kda kdx + dx kdx px + \sqrt{dx} \sqrt{kdx} \sqrt{da^2 dx kda^2 kdx + 2 da kda (dx kdx + 2 pa) px + dx kdx px^2}}{2 da (dx kdx + pa)}$$



- Promoter
  - Similar to Transcription Repressor
- Receptor
  - Signal Transduction
- Enzyme
  - Product or Substrate
  - Signal Transduction
- Binding Partner
  - Yeast Two-Hybrid, Phage Display, Protein Fragment Complementation
  - New Approaches

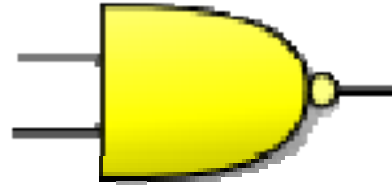


iGEM 2008 Warsaw

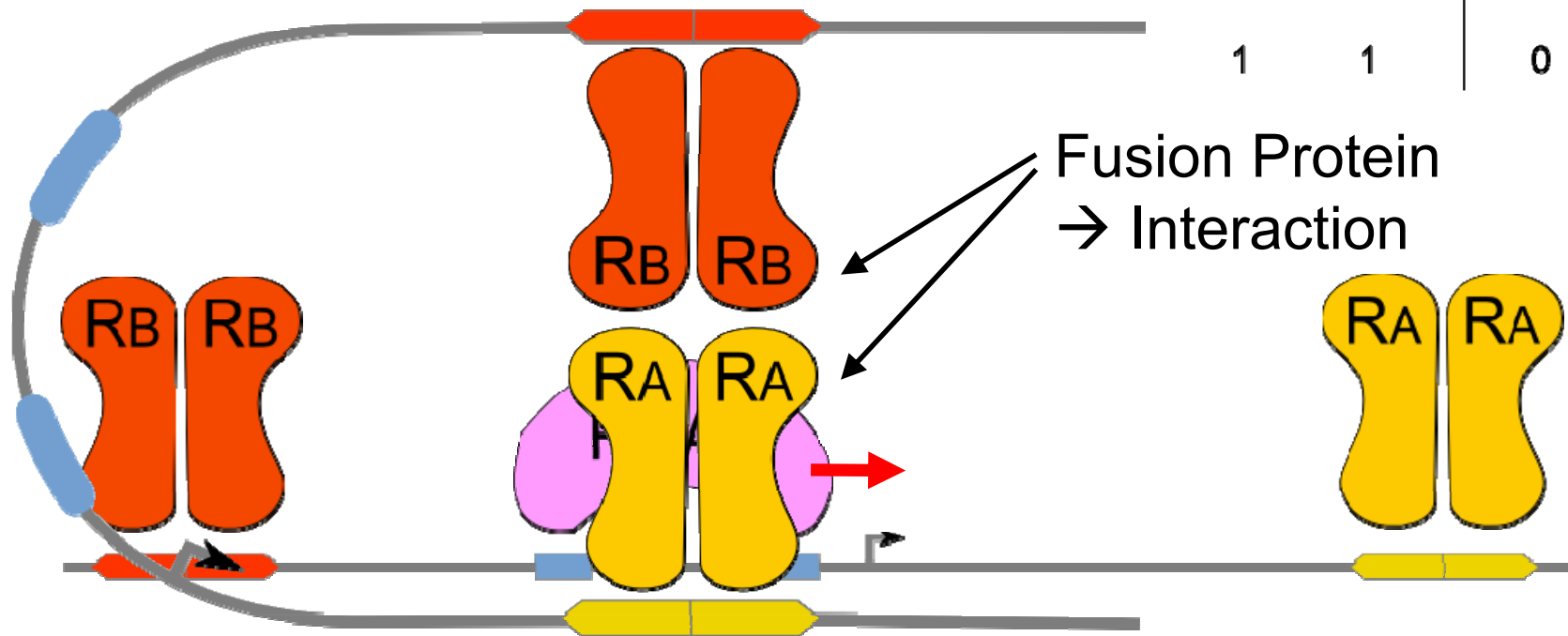


# Cis-acting Logic Promoters

NAND Gate



A	B	O
0	0	1
0	1	1
1	0	1
1	1	0



## An RNA Aptamer that Induces Transcription

Anke Hunsicker,<sup>1,5</sup> Markus Steber,<sup>1,2,5</sup> Günter Mayer,<sup>3,4</sup> Johannes Meiert,<sup>1</sup> Marcus Klotzsche,<sup>1</sup> Michael Blind,<sup>3</sup> Wolfgang Hillen,<sup>1</sup> Christian Berens,<sup>1</sup> and Beatrix Süss<sup>2,\*</sup>

<sup>1</sup>Lehrstuhl für Mikrobiologie, Friedrich-Alexander-Universität Erlangen-Nürnberg, Staudtstr. 5, 91058 Erlangen, Germany

<sup>2</sup>Institut für Molekulare Biowissenschaften, Johann Wolfgang Goethe-Universität Frankfurt, Max-von-Laue-Str. 9, 60438 Frankfurt (Main), Germany

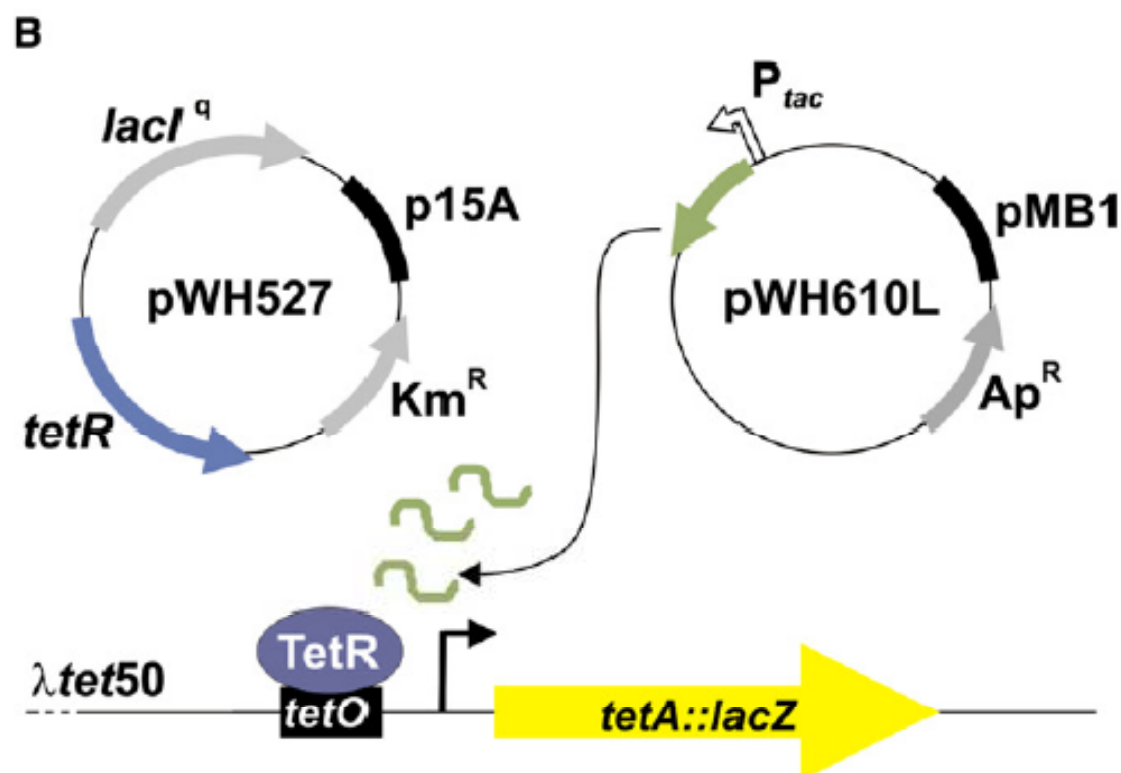
<sup>3</sup>NascaCell Technologies AG, Max-Lebsche-Platz 31, 81377 München, Germany

<sup>4</sup>Life and Medical Sciences Bonn, Program Unit Chemical Biology, c/o Kekulé-Institute for Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Str. 1, 53121 Bonn, Germany

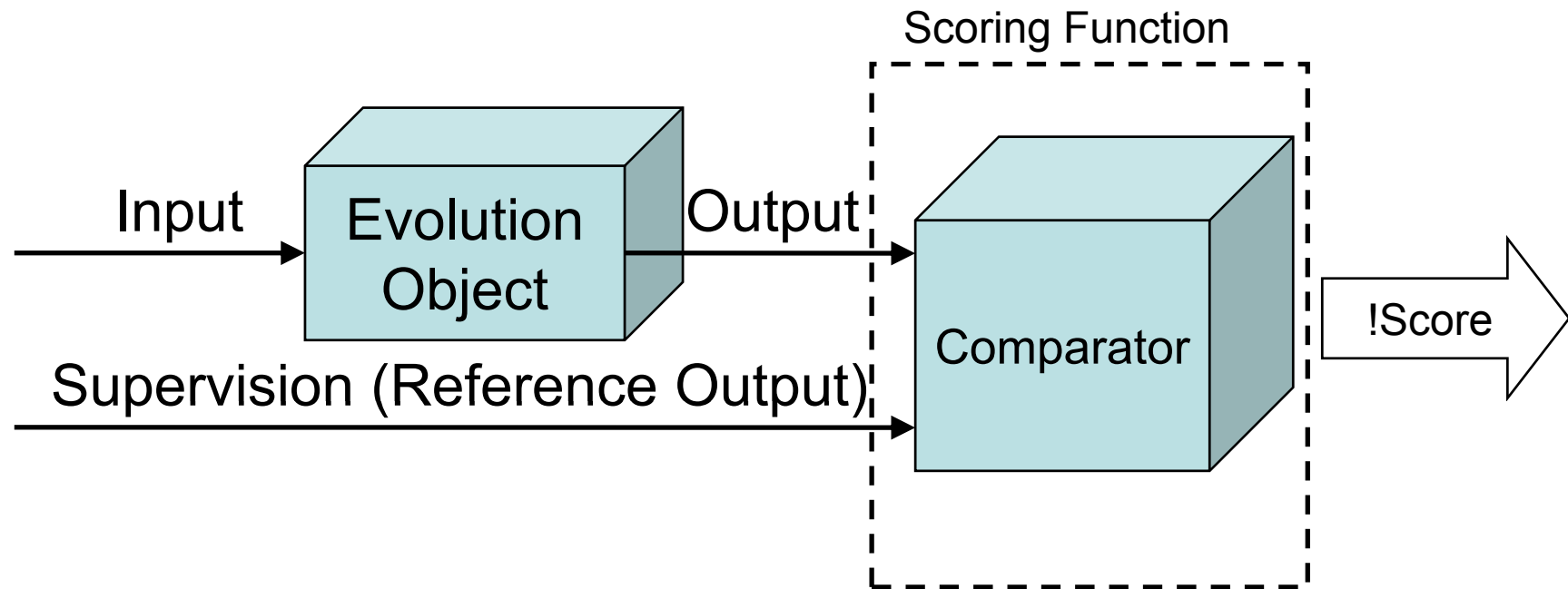
<sup>5</sup>These authors contributed equally to this work.

\*Correspondence: [suess@bio.uni-frankfurt.de](mailto:suess@bio.uni-frankfurt.de)

DOI 10.1016/j.chembiol.2008.12.008

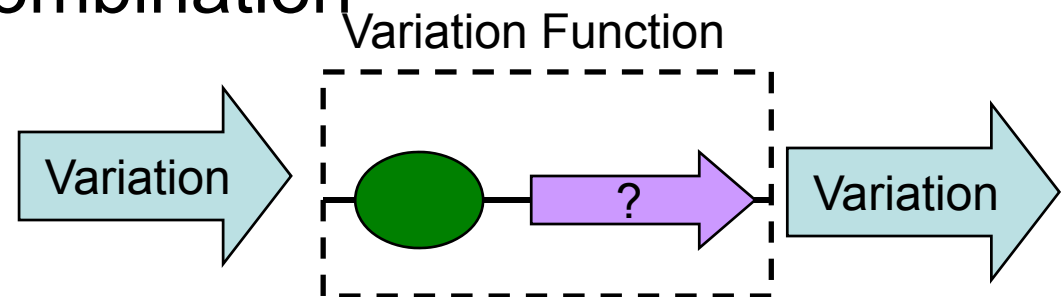


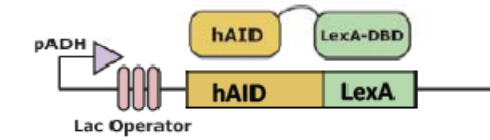
# Supervised Learning



# Variation Function

- Targeted Mutagenesis
  - Activation induced cytidine deaminase (AID)
  - Error-prone DNA polymerase I
  - Bacteriophage ?
  - Error-prone reverse transcription ?
- Recombination
  - Site-specific recombination
  - Homologous recombination
- Gene
- Network

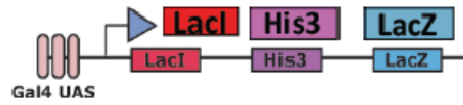




Mutator

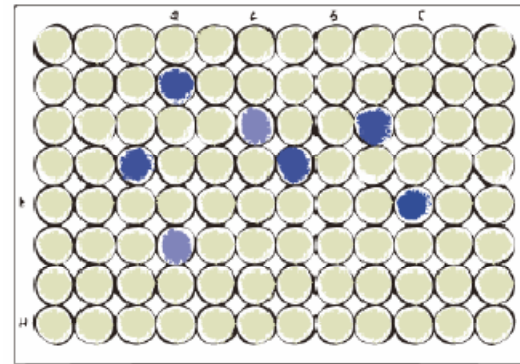


Target

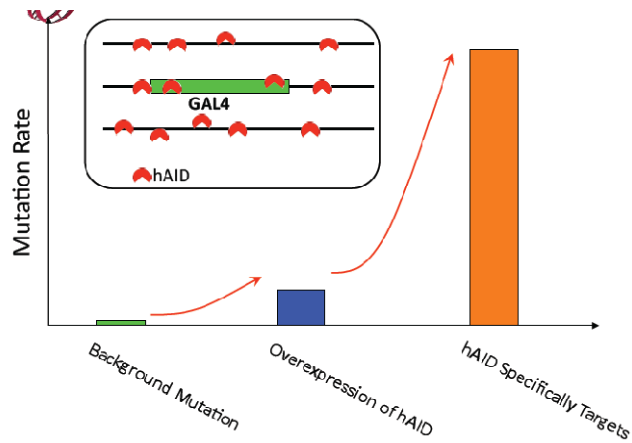


Brake

From Qualitative to Quantitative.

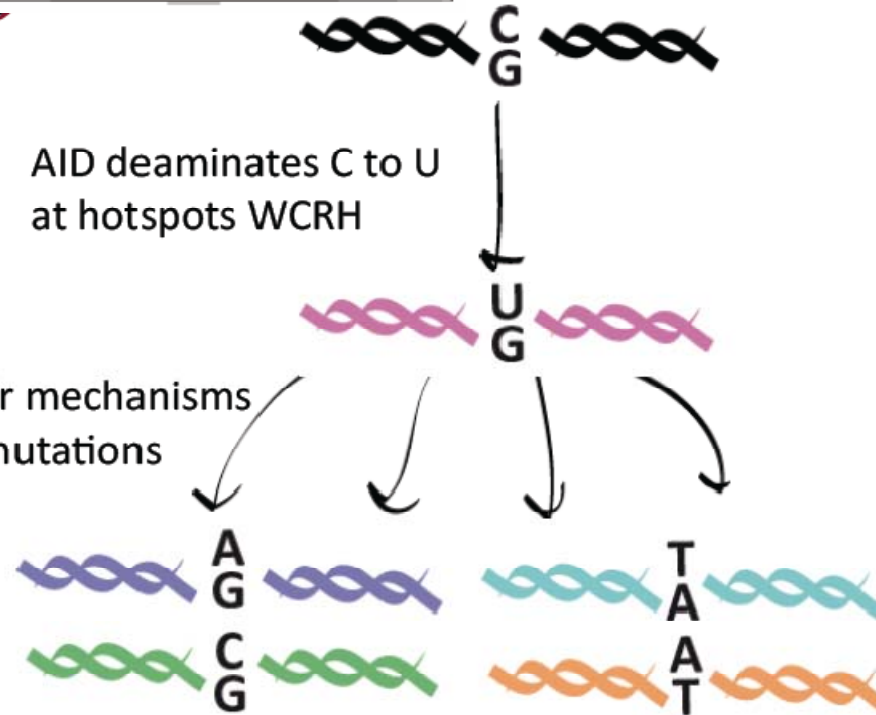


## Mutator: hAID



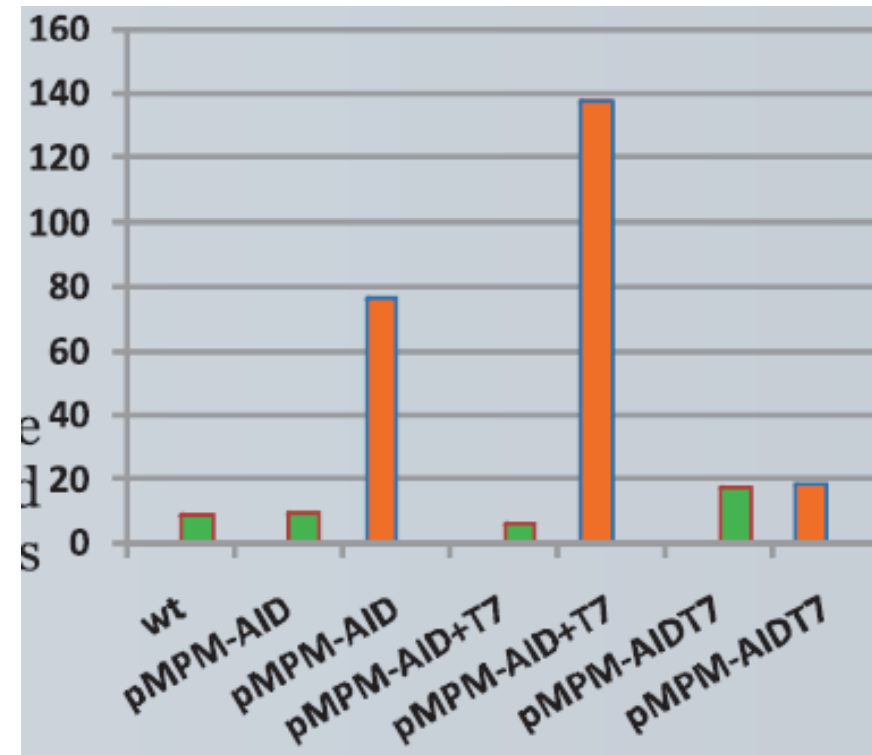
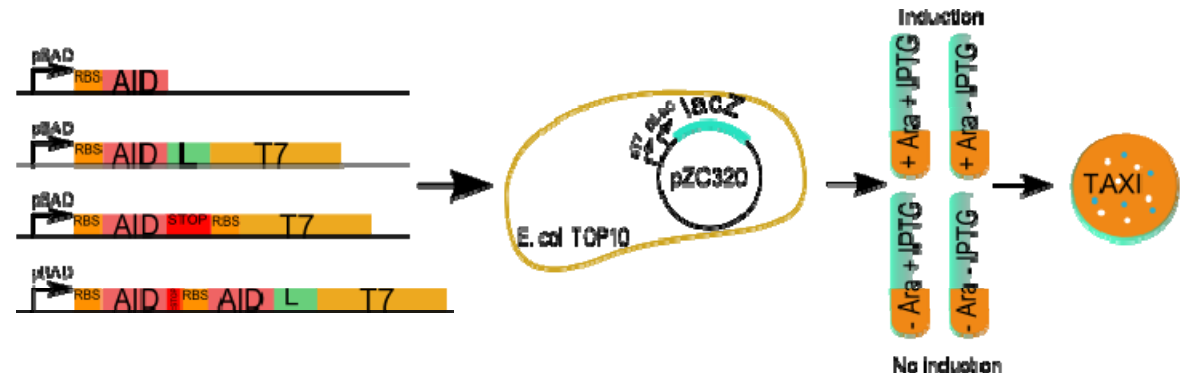
AID deaminates C to U at hotspots WCRH

DNA repair mechanisms diversify mutations





**Fig1.** Predicted structure of the AID (red) + polymerase T7 (green, linker) fusion. Yellow is a glycine-serine linker



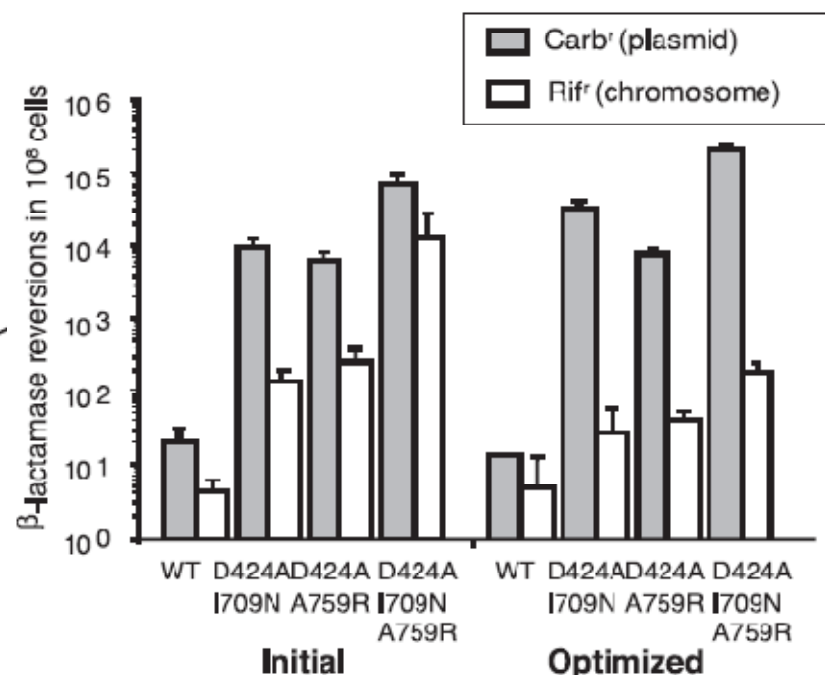
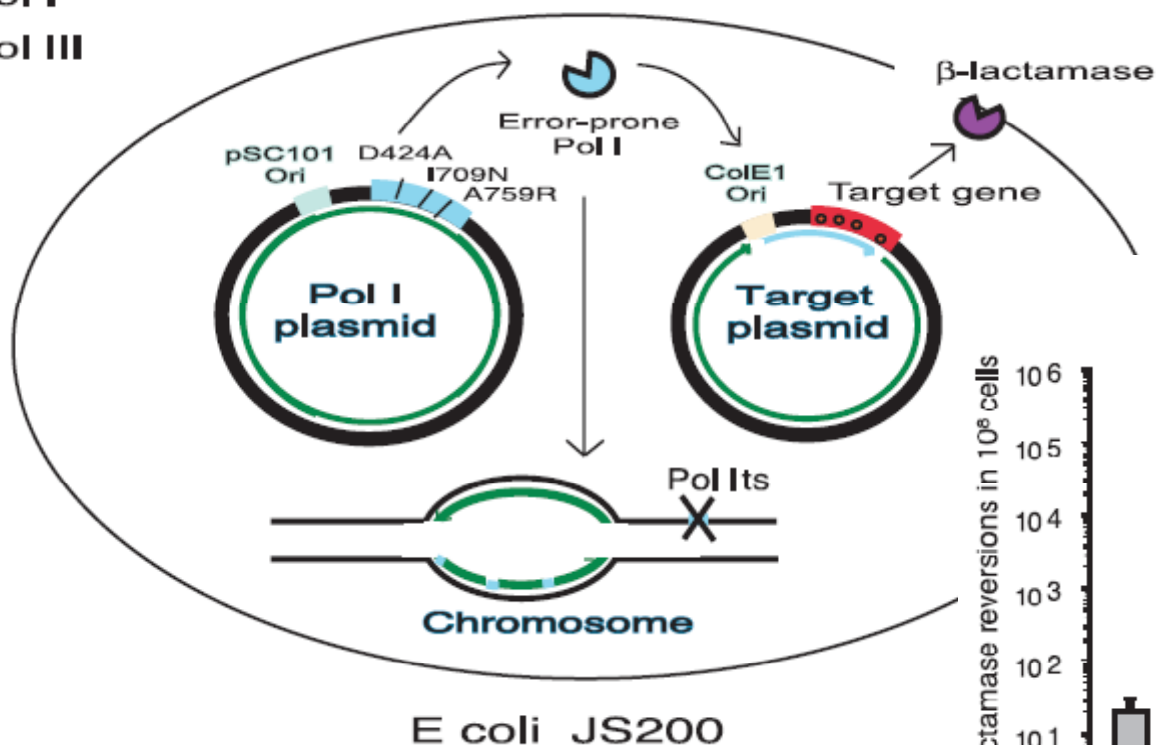
# Targeted gene evolution in *Escherichia coli* using a highly error-prone DNA polymerase I

Manel Camps, Jussi Naukkarinen, Ben P. Johnson, and Lawrence A. Loeb\*

The Joseph Gottstein Memorial Cancer Research Laboratory, Department of Pathology, University of Washington, Seattle, WA 98195-7705

Communicated by Gerald F. Joyce, The Scripps Research Institute, La Jolla, CA, June 25, 2003 (received for review April 20, 2003)

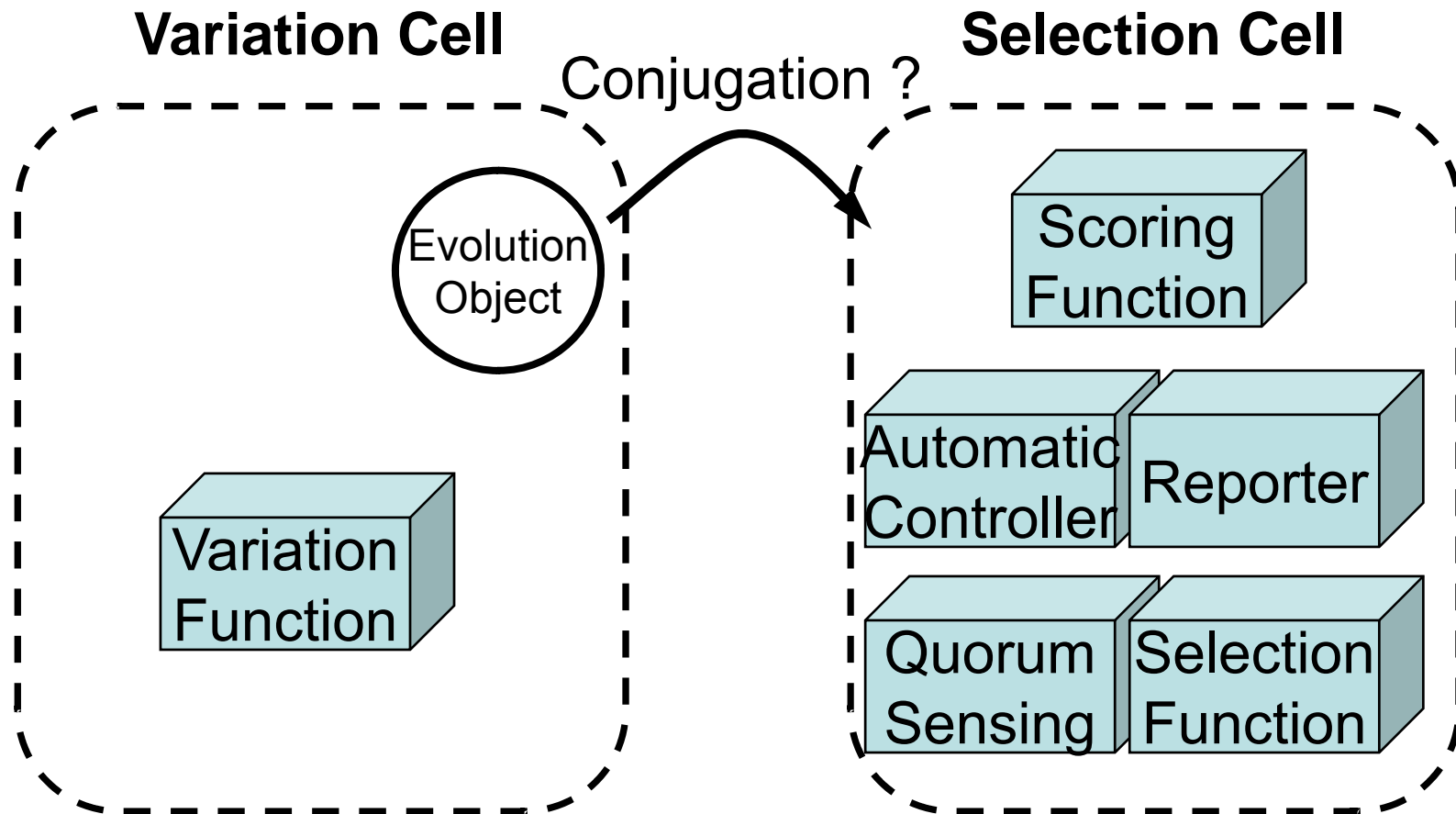
— Pol I  
— Pol III



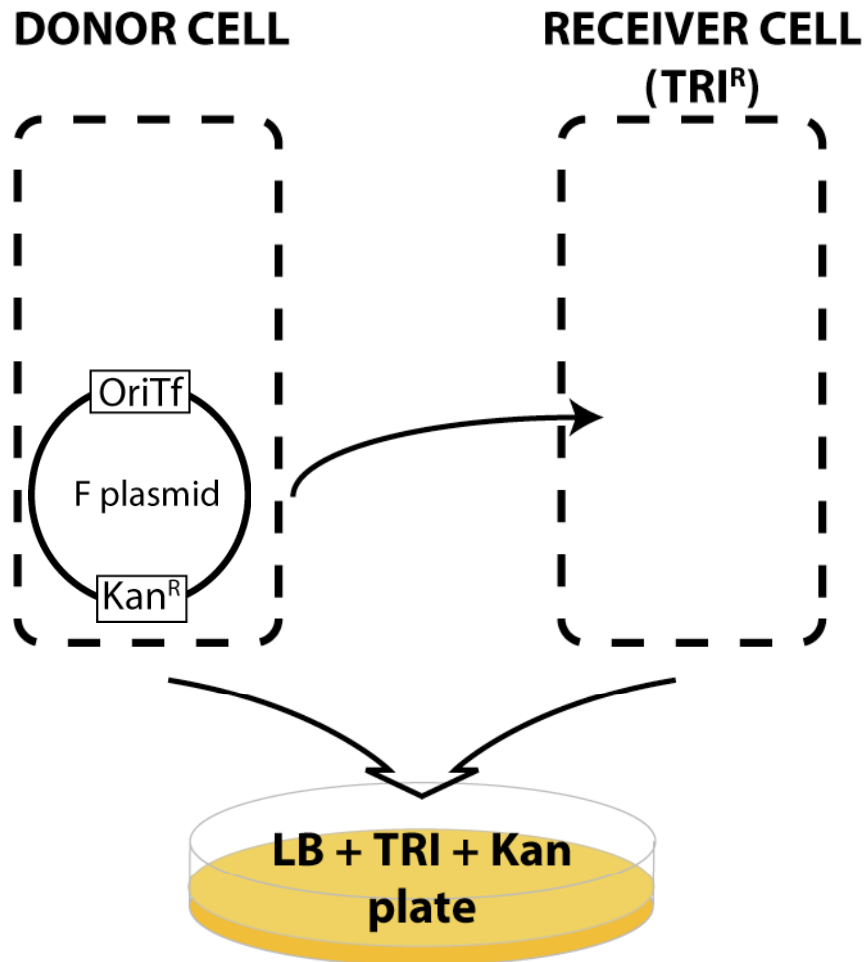


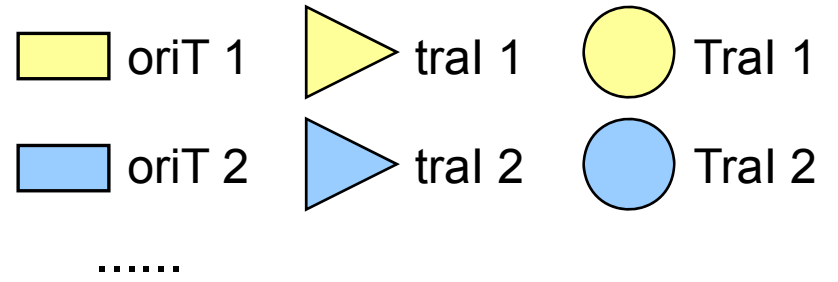
# Avoiding Non-Specific Mutation

- An immature idea:

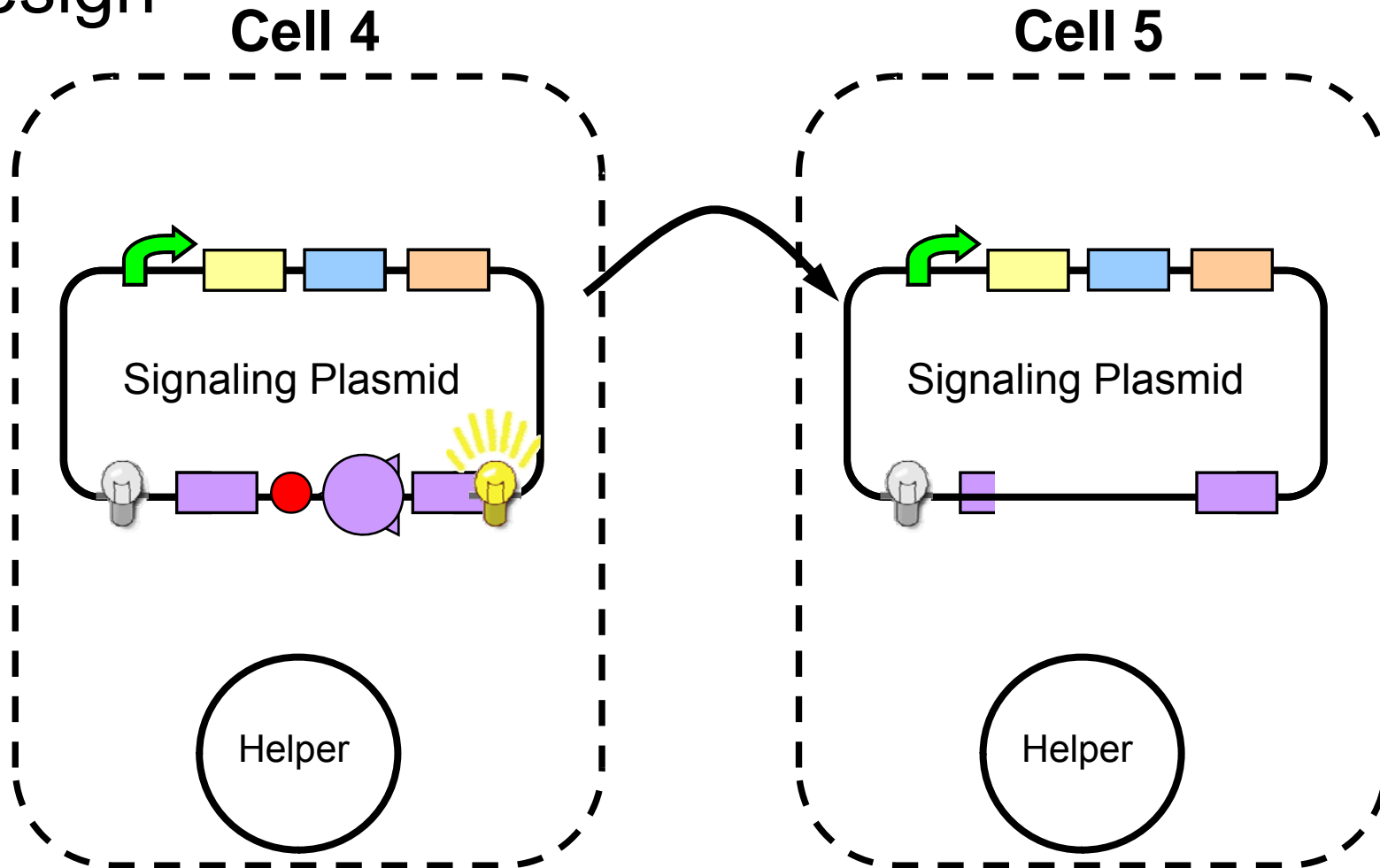


# Bacterial Conjugation

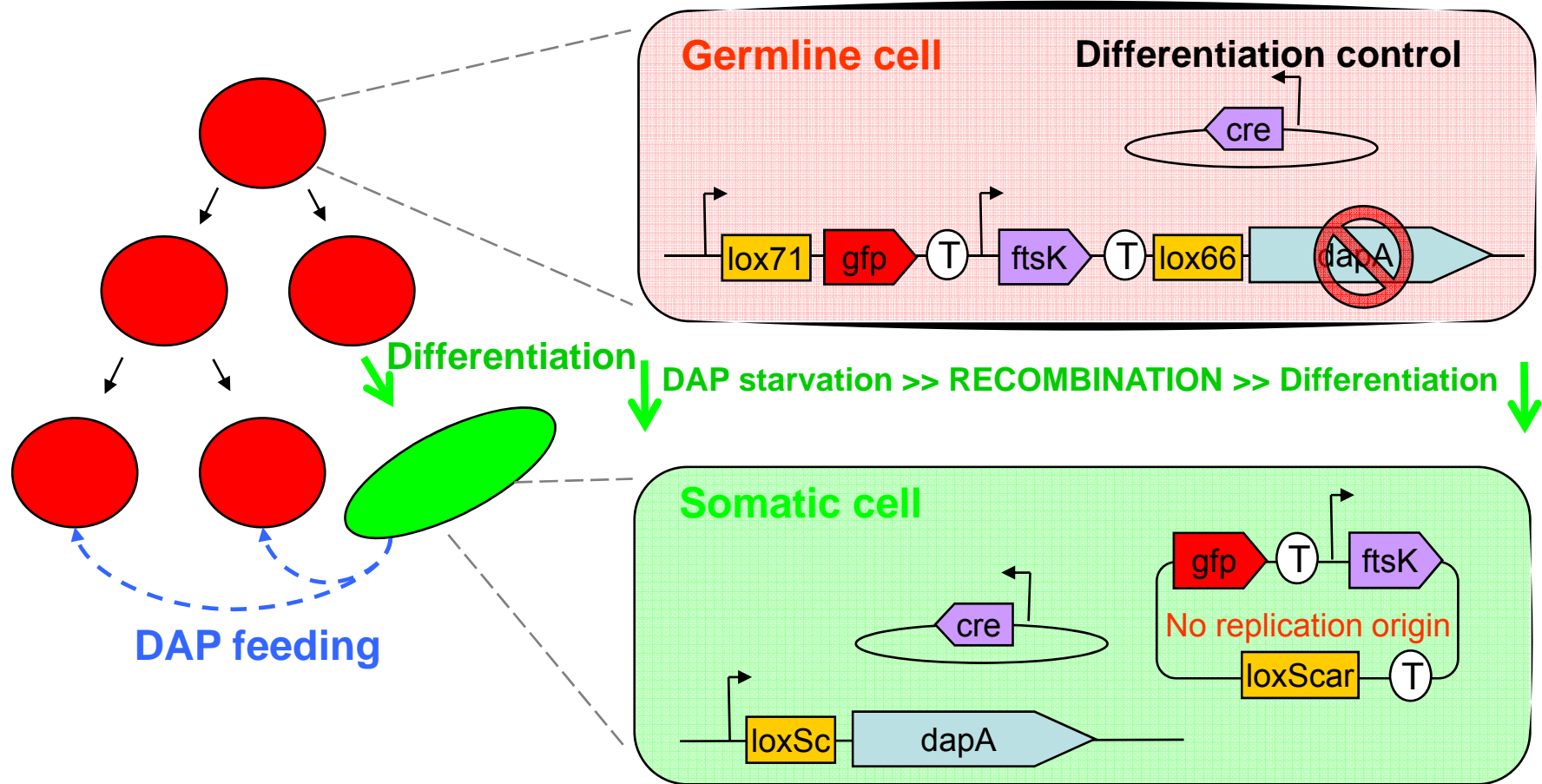




- Design



# iGEM 2006 Paris



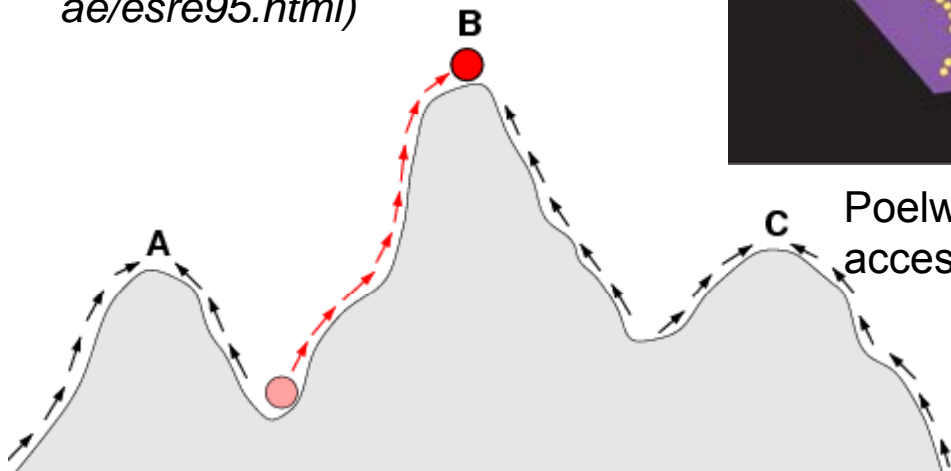
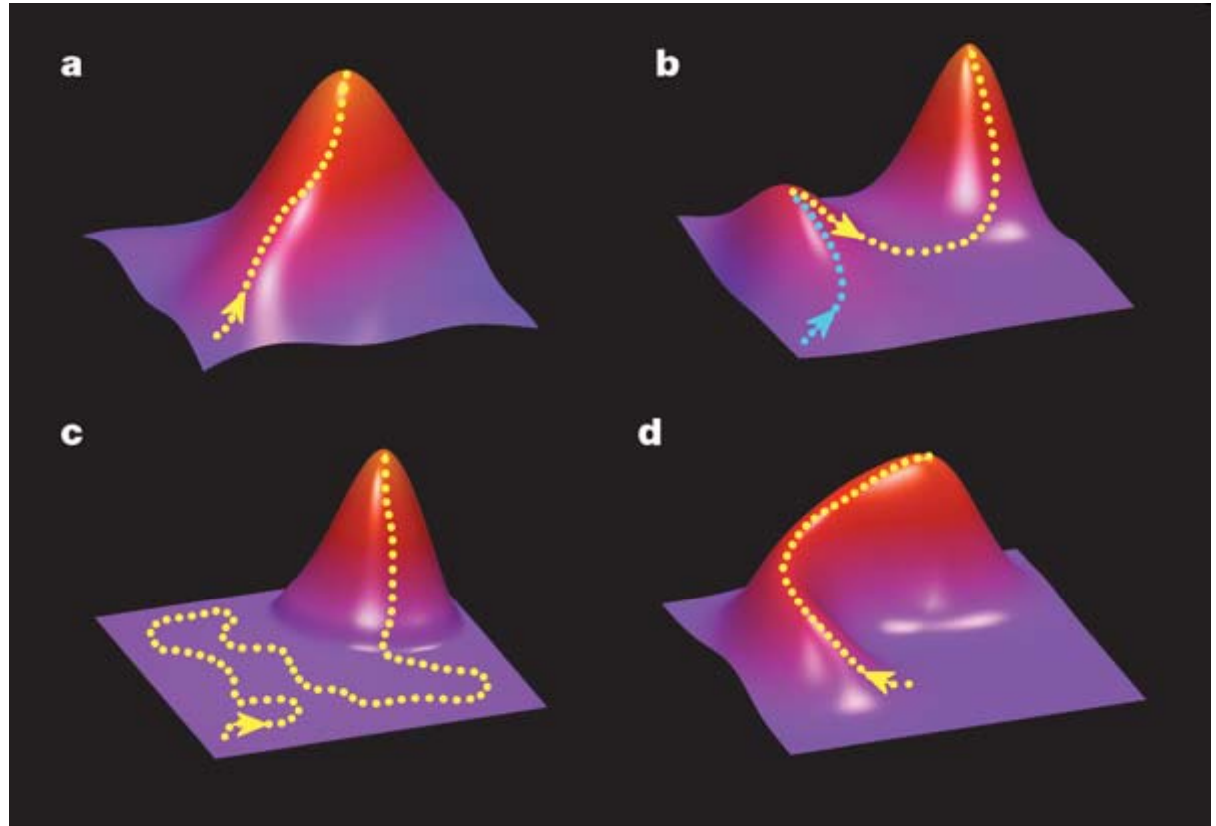
# Automatic Controller

- Nobody can really understand evolution.
  - The evolution of evolution itself
- Problems
  - Variation rate
    - Slow vs. Misconvergence
  - Selection pressure
    - Nondirectional vs. Die out
- Theories
  - Evolutionary Biology & Population Genetics
  - Evolutionary Algorithm

# Fitness Landscape

"**Evolution** is sometimes characterized by biologists as a metaphorical **uphill struggle** across a "**fitness landscape**" in which mountain **peaks** represent high "fitness," or ability to survive, and **valleys** represent low fitness. As evolution proceeds, a population of organisms in effect takes an "**adaptive walk**" across such a landscape."

(<http://gemini.tntech.edu/~mwmcr ae/esre95.html>)



Poelwijk et al. 2007. Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* **445**, 383-386.

[http://en.wikipedia.org/wiki/Fitness\\_landscape](http://en.wikipedia.org/wiki/Fitness_landscape)

Population	0	0	0	0	0	0	0	0	0
Average Score	0	0	0	0.5	0.5	0.5	1	1	1
Score	0	0.5	1	0	0.5	1	0	0.5	1
Death	0	0	0	0	0	0	0	0	0

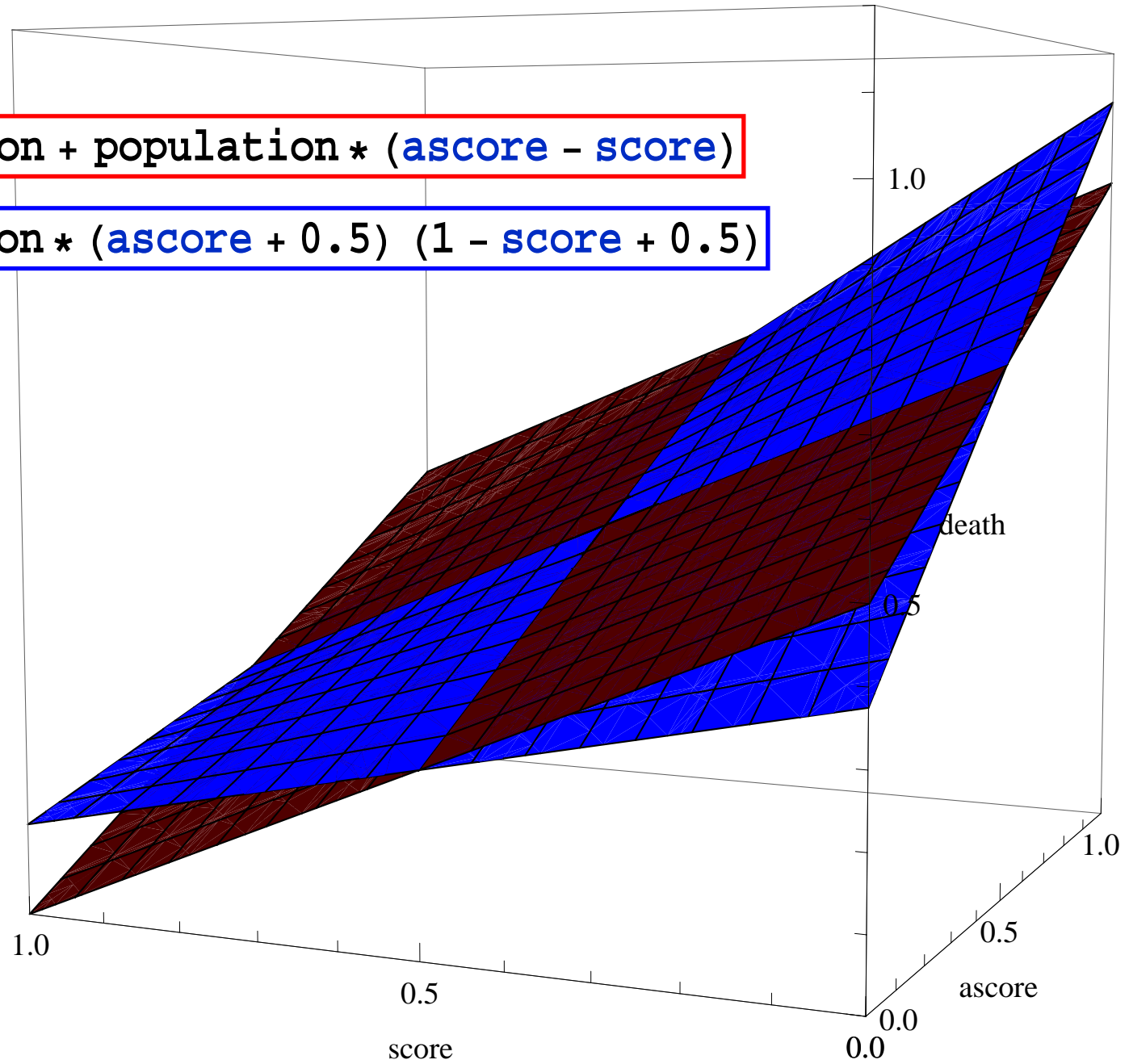
Population	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
Average Score	0	0	0	0.5	0.5	0.5	1	1	1
Score	0	0.5	1	0	0.5	1	0	0.5	1
Death	0.5	0.25	0	0.75	0.5	0.25	1	0.75	0.5

Population	1	1	1	1	1	1	1	1	1
Average Score	0	0	0	0.5	0.5	0.5	1	1	1
Score	0	0.5	1	0	0.5	1	0	0.5	1
Death	1	0.5	0	1.5	1	0.5	2	1.5	1

$$\text{Death} = \text{Population} + \text{Population} \times (\text{AverageScore} - \text{Score})$$

$$\text{population} + \text{population} * (\text{ascore} - \text{score})$$

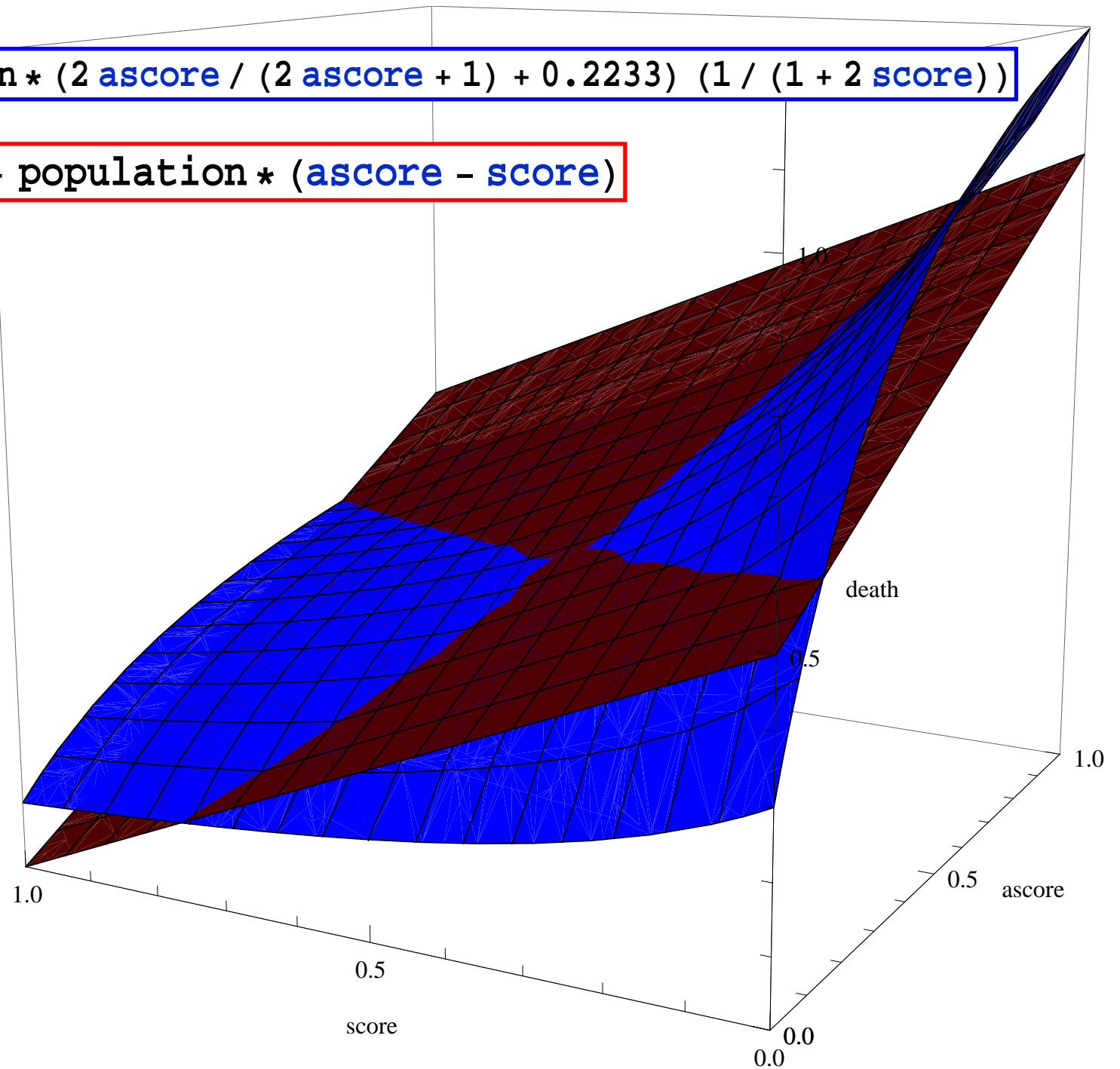
$$\text{population} * (\text{ascore} + 0.5) (1 - \text{score} + 0.5)$$





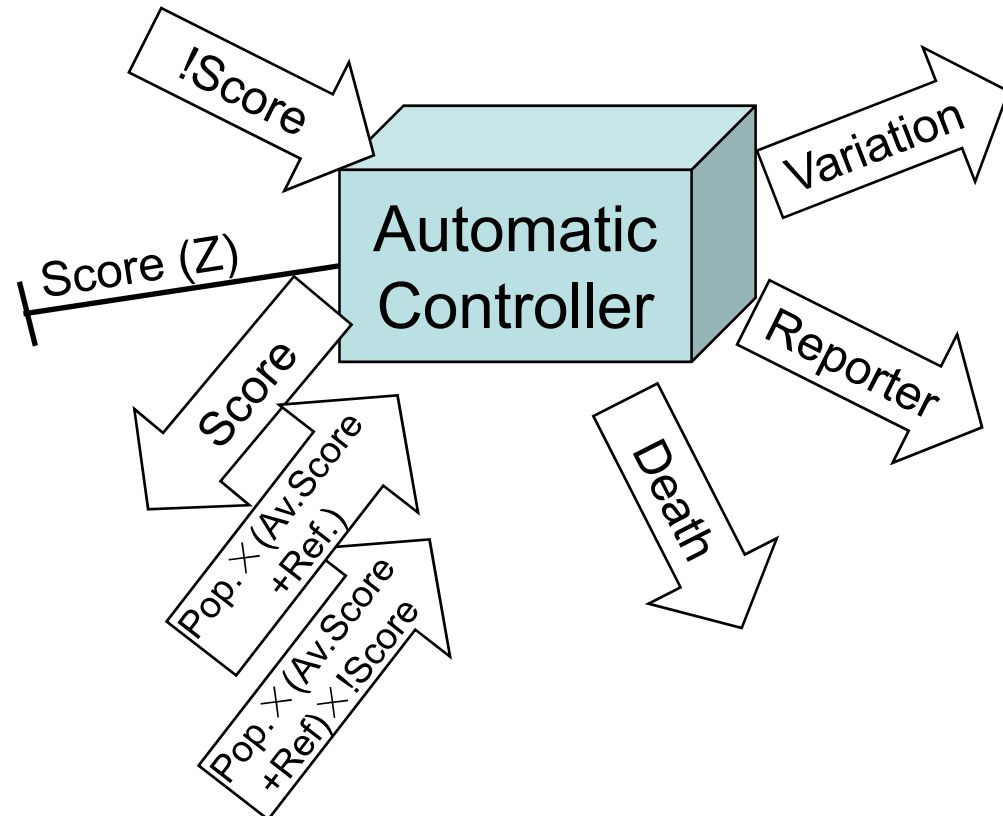
$$2.6898 \text{ population} * (2 \text{ ascore} / (2 \text{ ascore} + 1) + 0.2233) (1 / (1 + 2 \text{ score}))$$

$$\text{population} + \text{population} * (\text{ascore} - \text{score})$$

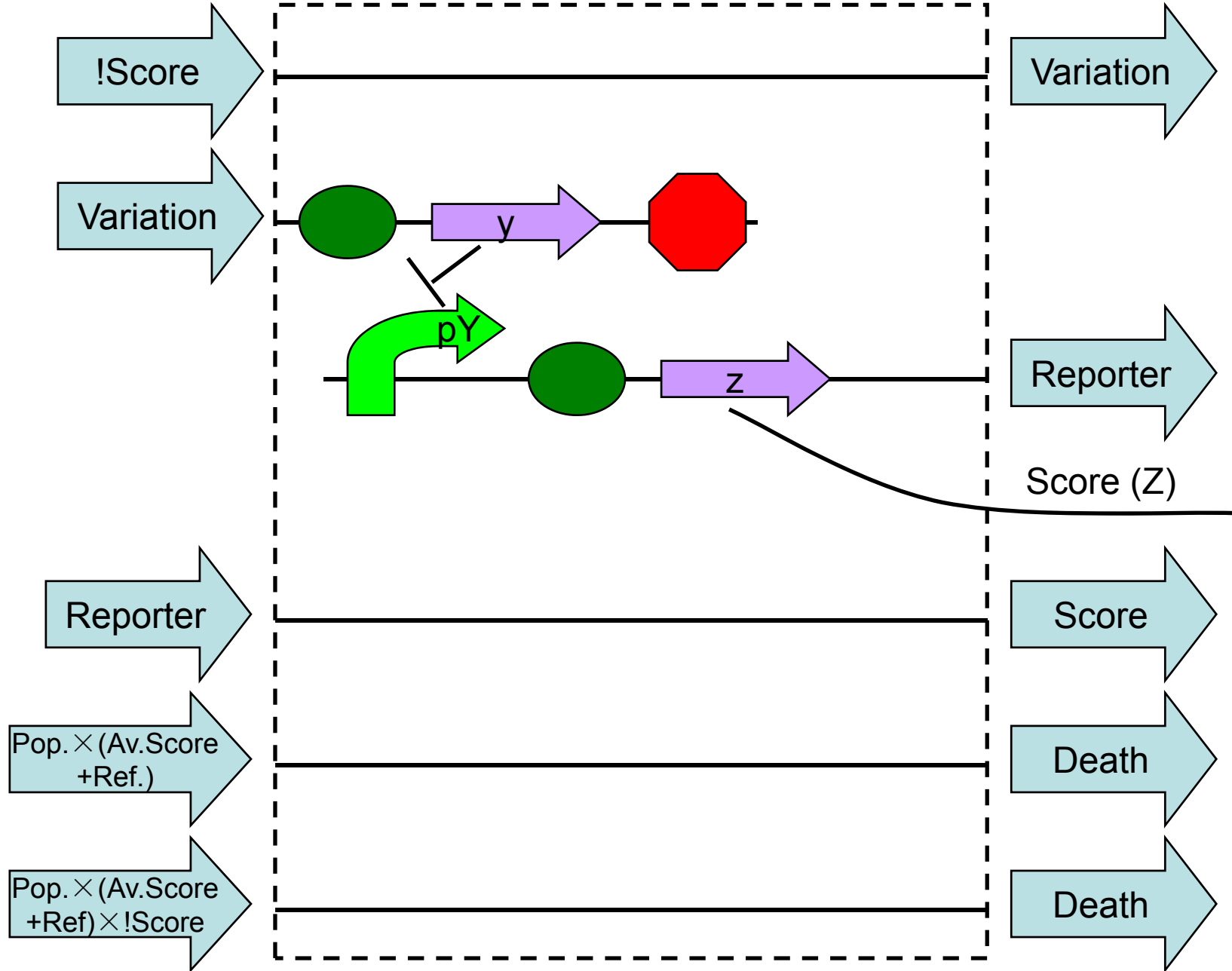


# Automatic Controller

- $\text{Death} = \text{Population} + \text{Population} (\text{AverageScore} - \text{Score})$   
 $\approx \text{Population} \times (\text{AverageScore} + C) \times !\text{Score}$
- $\text{Variation} = !\text{Score}$
- $\text{Score} = ! !\text{Score}$
- $\text{Reporter} = \text{Score}$
- $Z = \text{Score}$

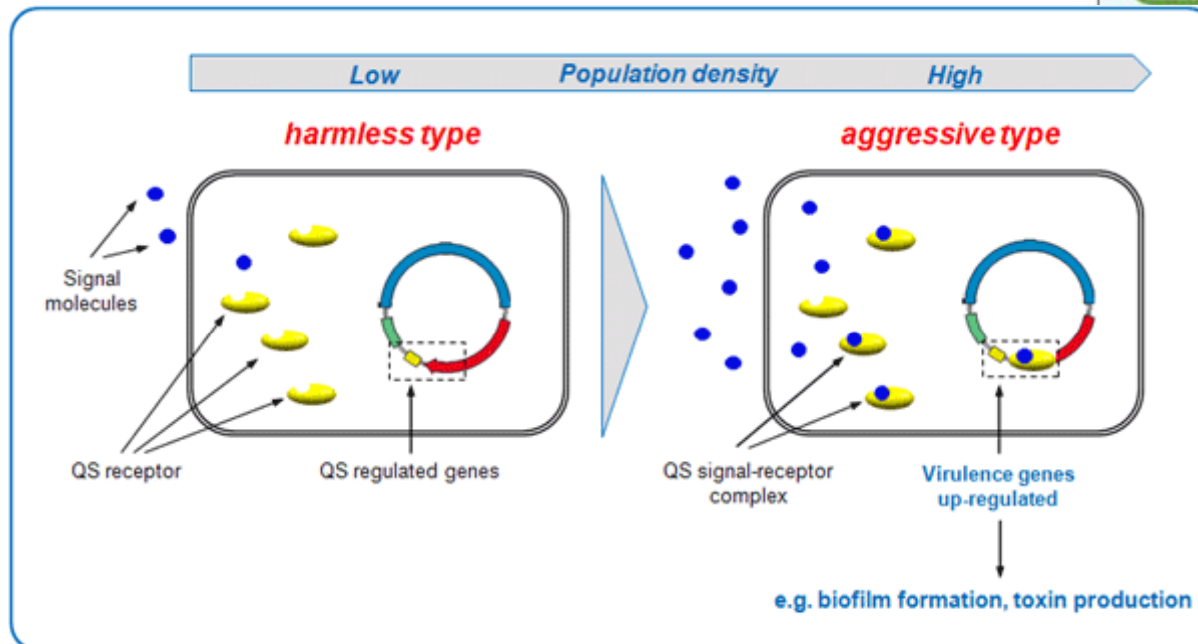
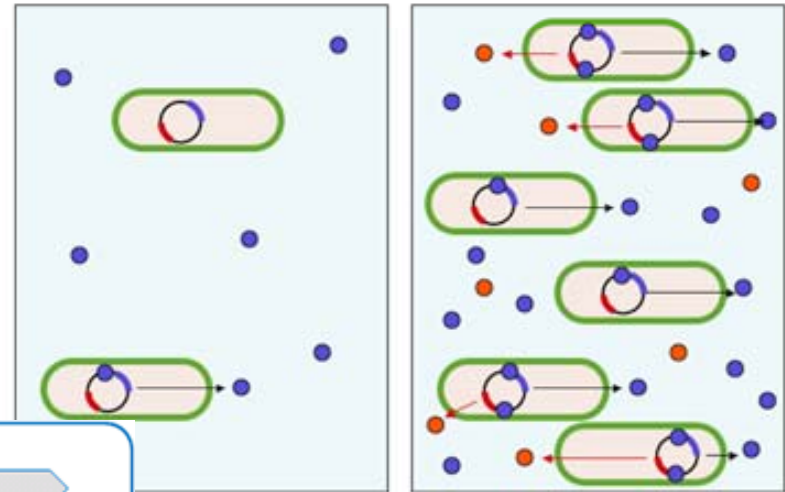


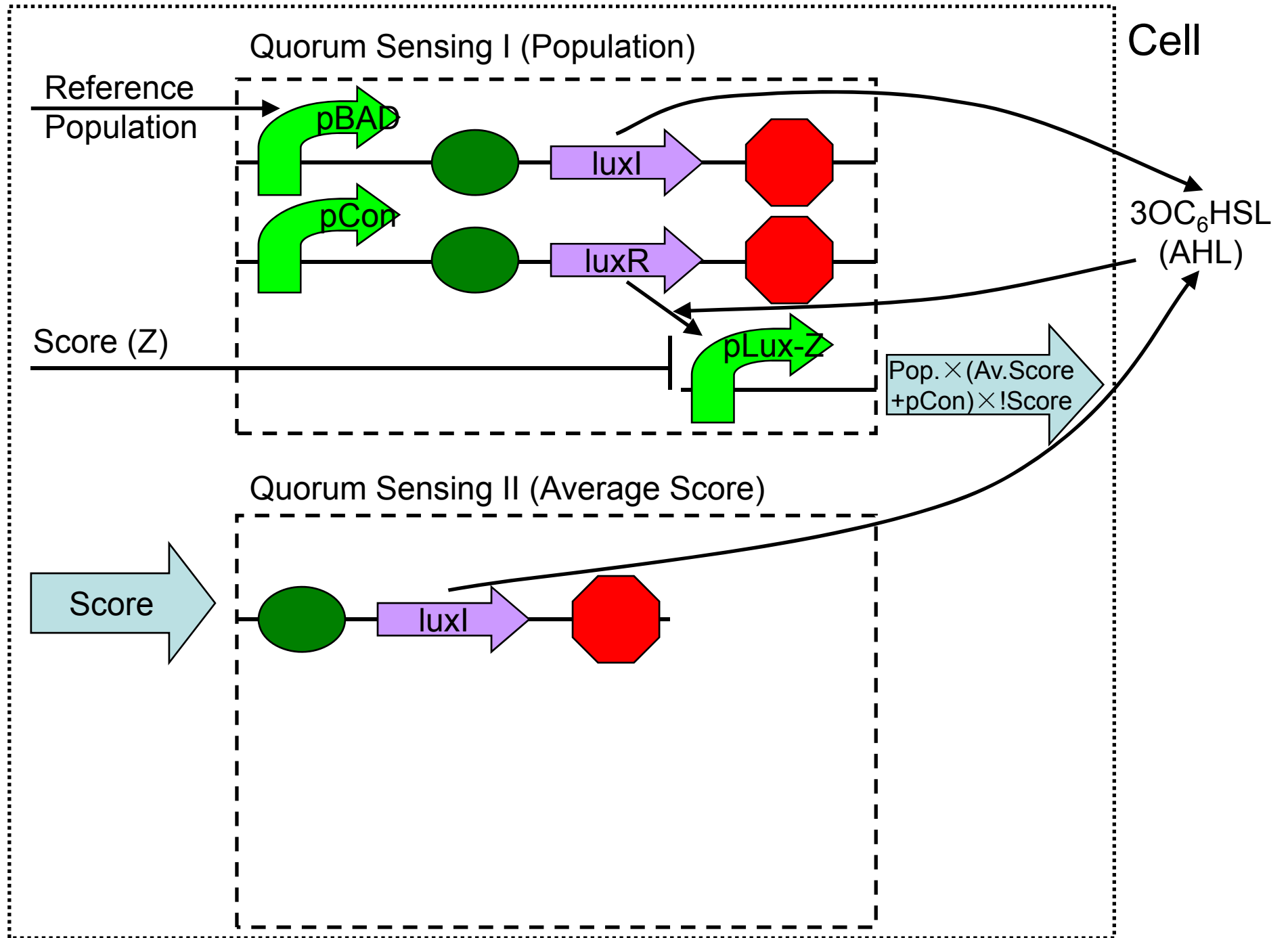
# Automatic Controller

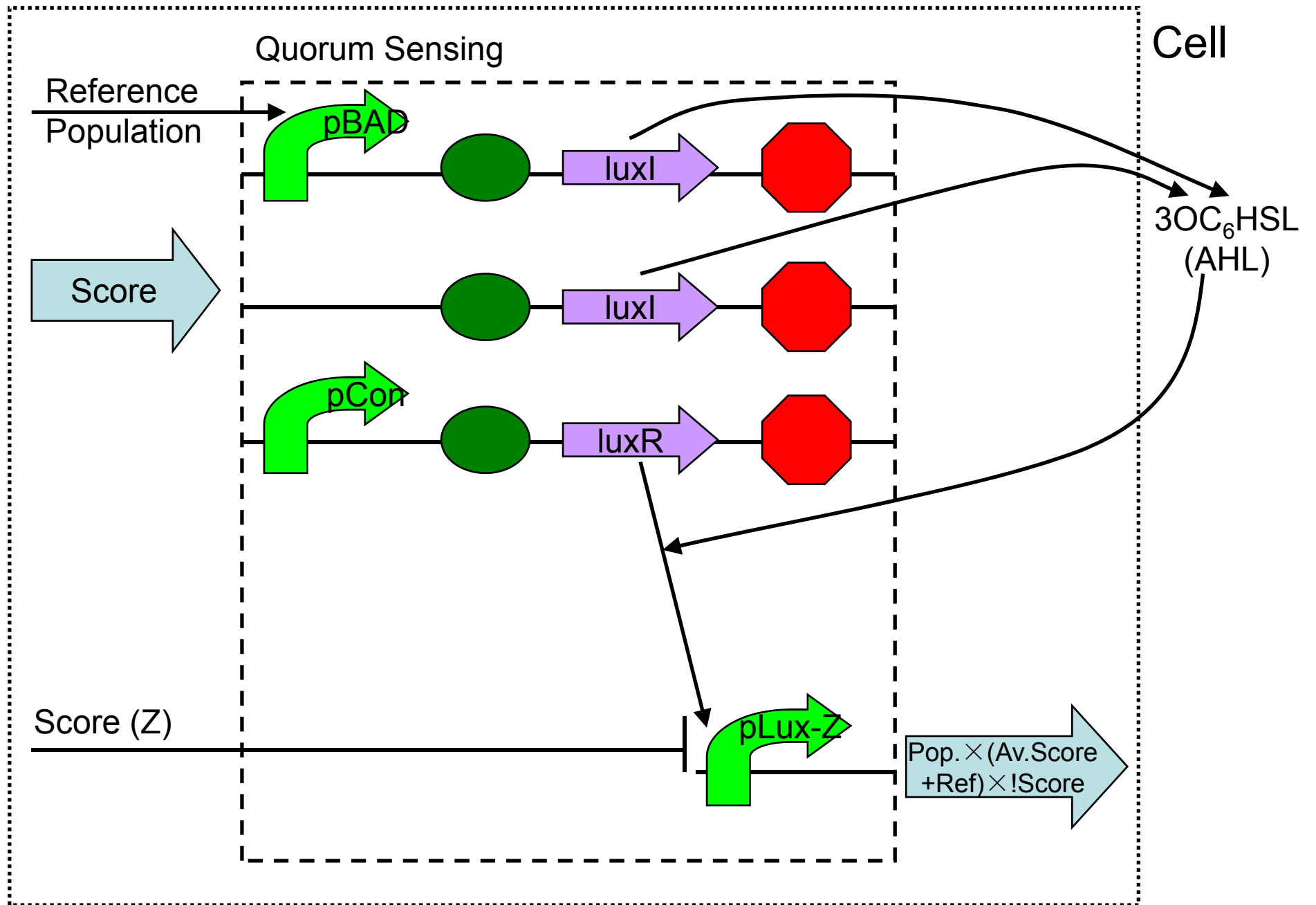


# Quorum Sensing

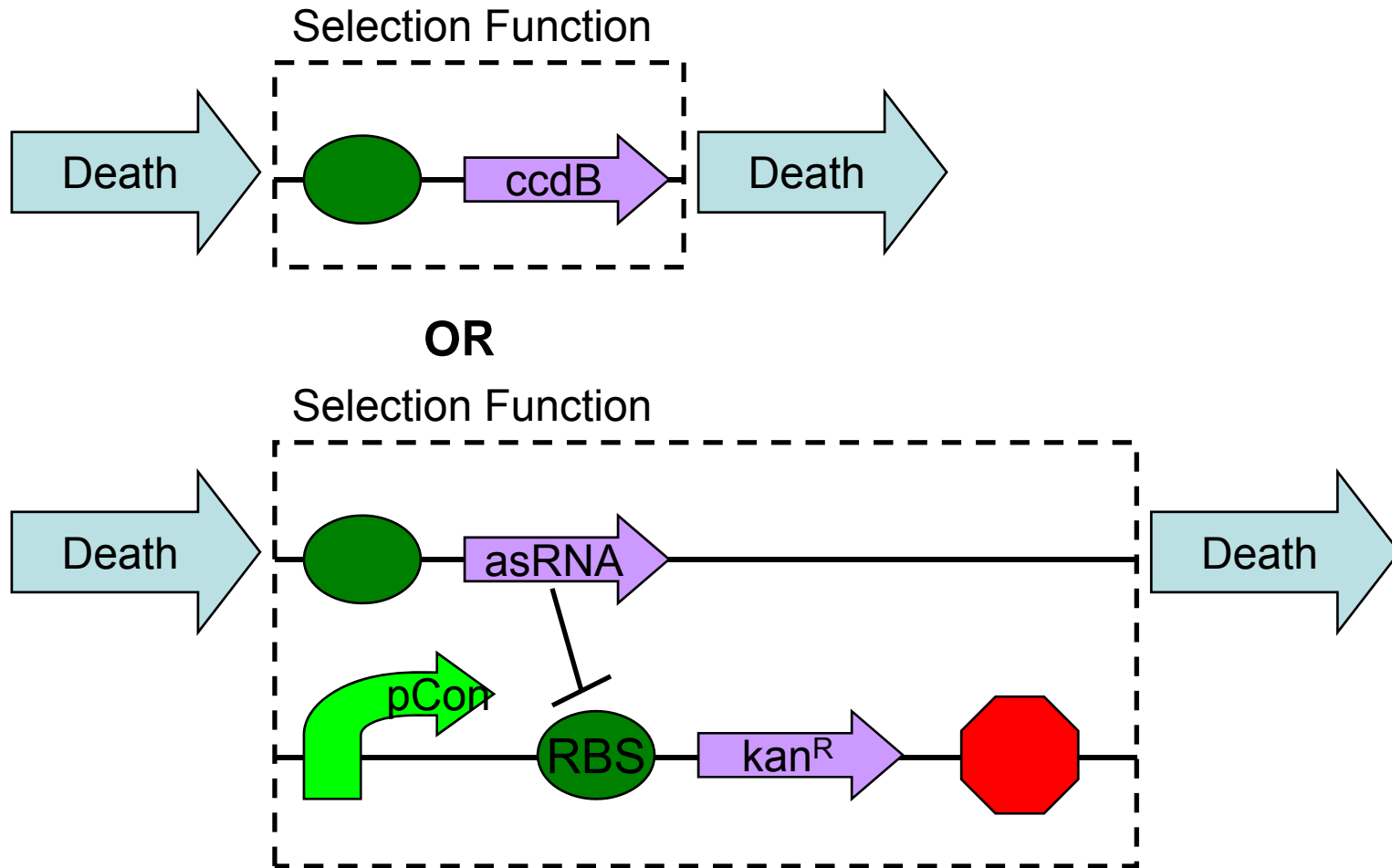
- Bacteria & social insects
- Decision-making by local density of population





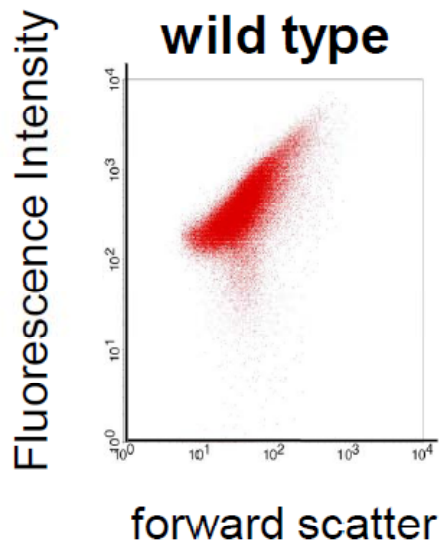
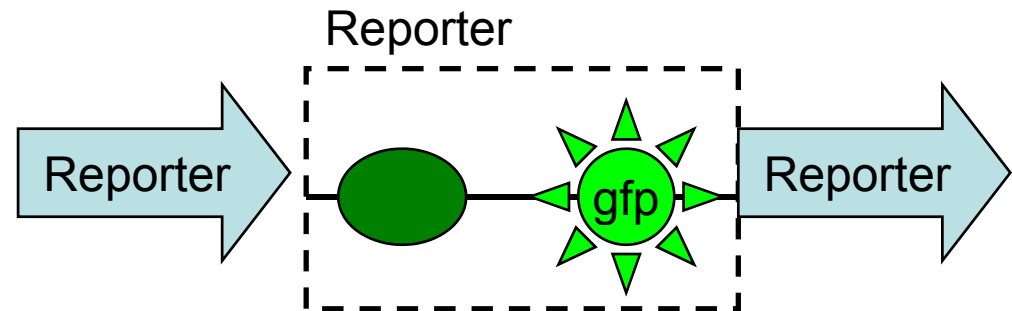


# Selection Function

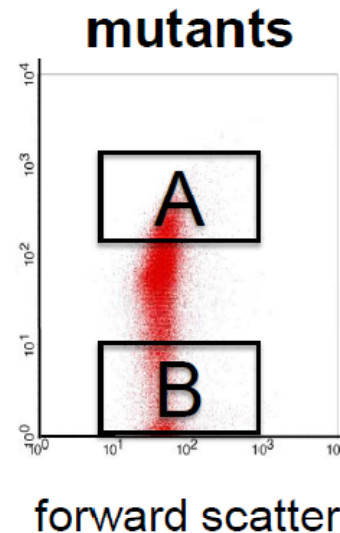


# Reporter

- Real-time evolution dynamics by FACS
  - Monitoring
  - Further screening



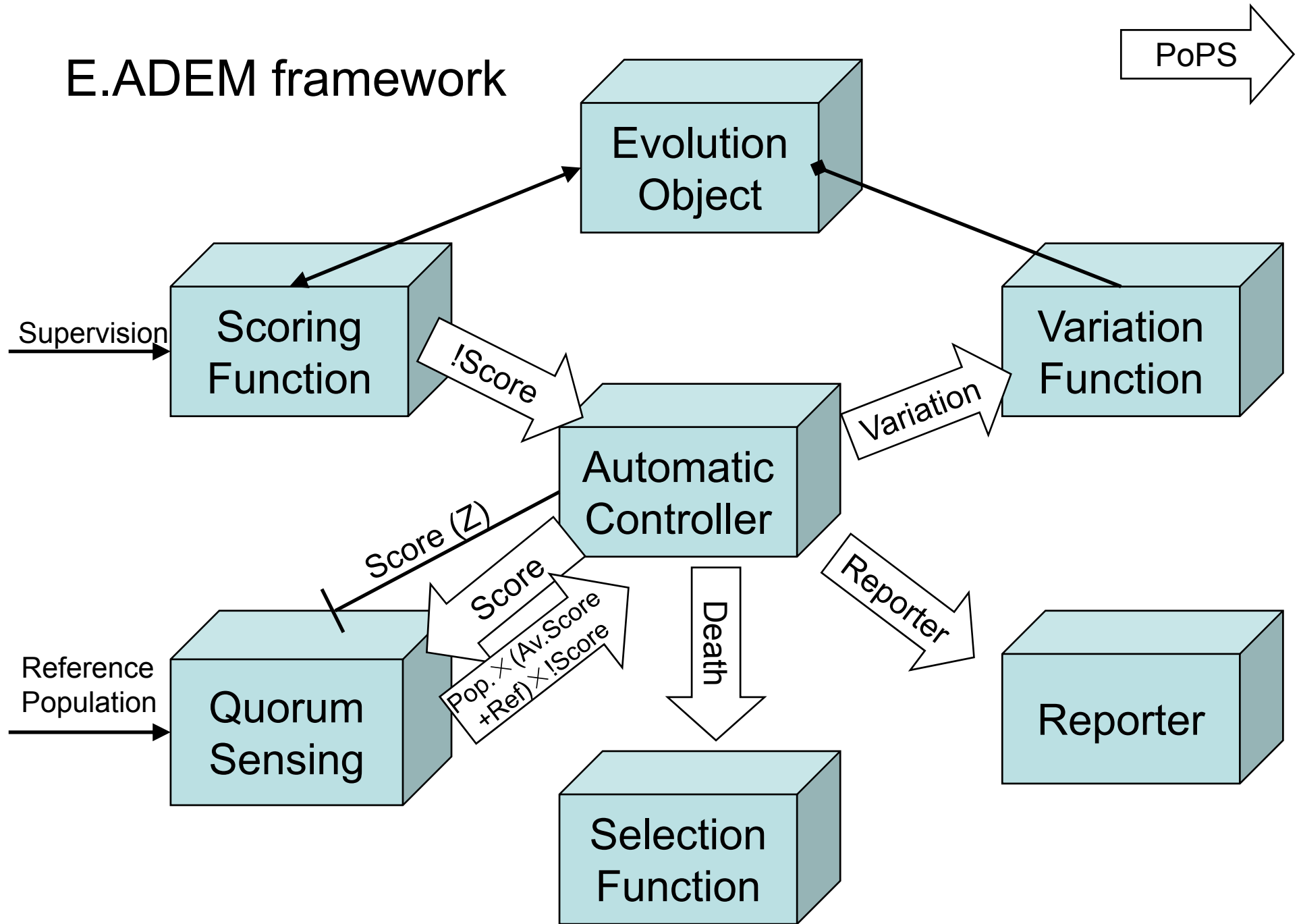
*mutagenesis*

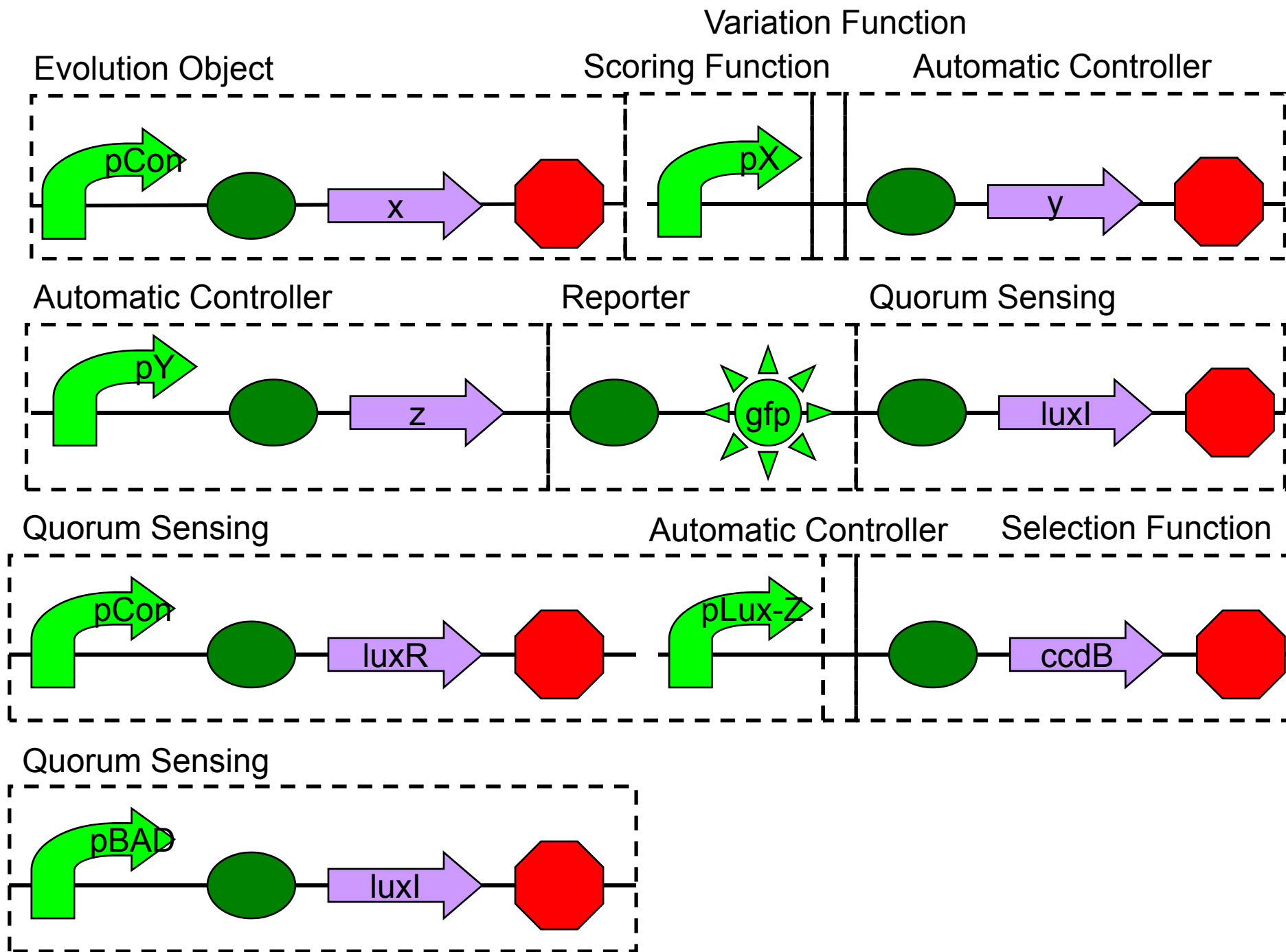


← iGEM 2008  
Tokyo\_Tech



# E.ADEM framework





- 6 promoters and 6 terminators
- 8 RBSs and 8 coding sequences
  - (2 luxI + luxR + ccdB + gfp + x + y + z)
- Interestingly, most of the coding sequences are used in iGEM 2008 USTC team.
  - Only coding sequences in Evolution Object, Scoring Function, and Selection Function are new.

# Future Plan

- Automatic Controller & Quorum Sensing
  - Modeling
  - Quantitative Analysis
  - BioBrick Parts
- Selection Function
- Variation Function
- Scoring Function

# Grouping

- Scoring
  - Evolution Object + Scoring Function + Automatic Controller + Reporter
- Population
  - Quorum Sensing + Selection Function
- Variation
  - Variation Function + Conjugation ?
- Light Sensor ?
- Wiki ?

150 years ago, Darwin told us where we came from using evolution.

Today, let's find out where we can go using evolution!

Thank You!

