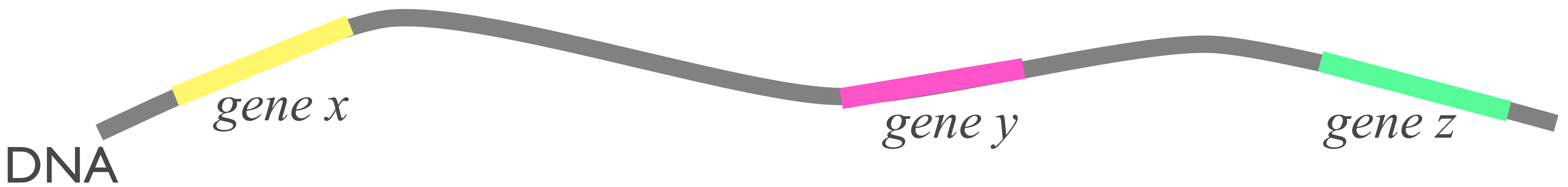
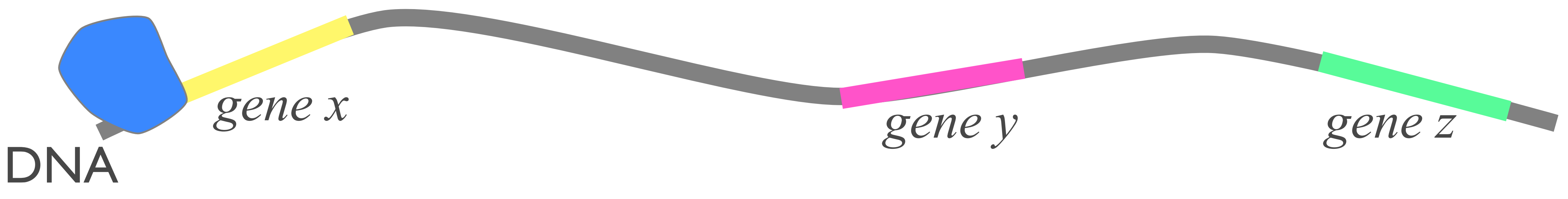
Basics on Metabolic network analysis

Marko Budinich, Jérémie Bourdon and Damien Eveillard
LINA, Université de Nantes

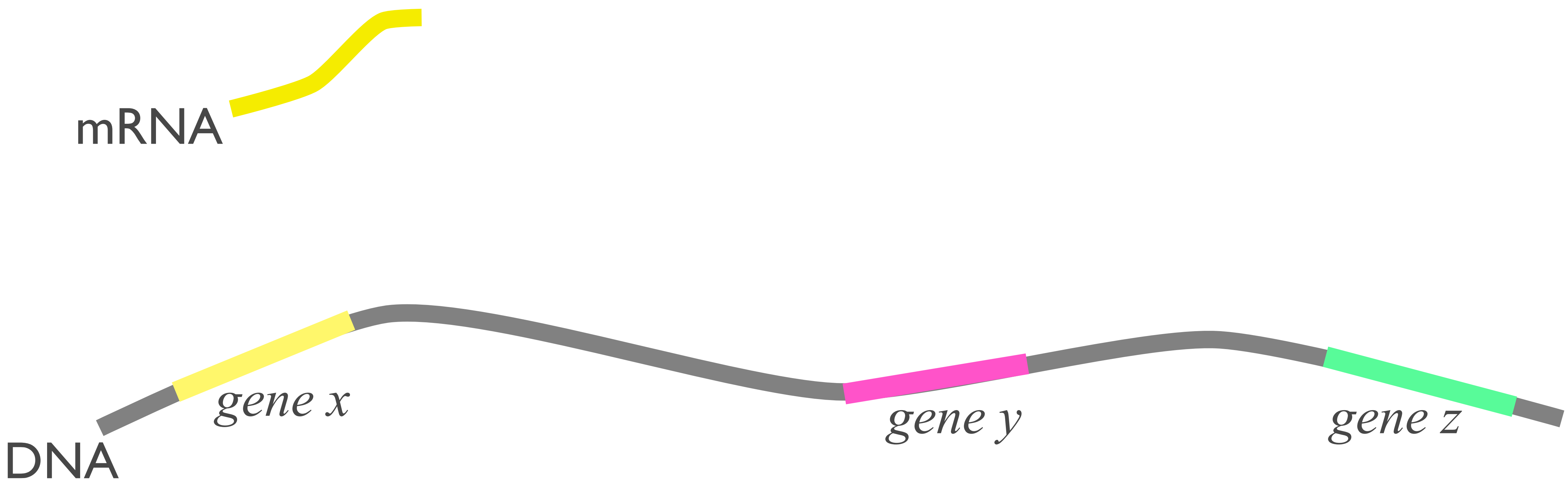
A simplistic illustration



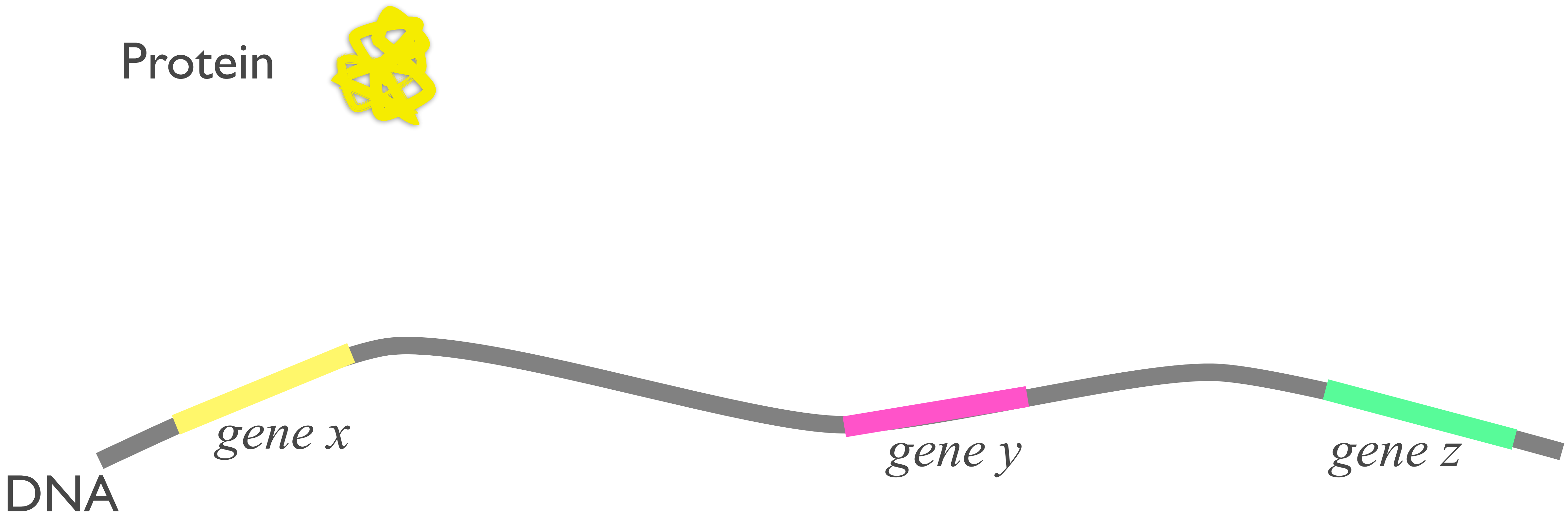
A simplistic illustration



A simplistic illustration

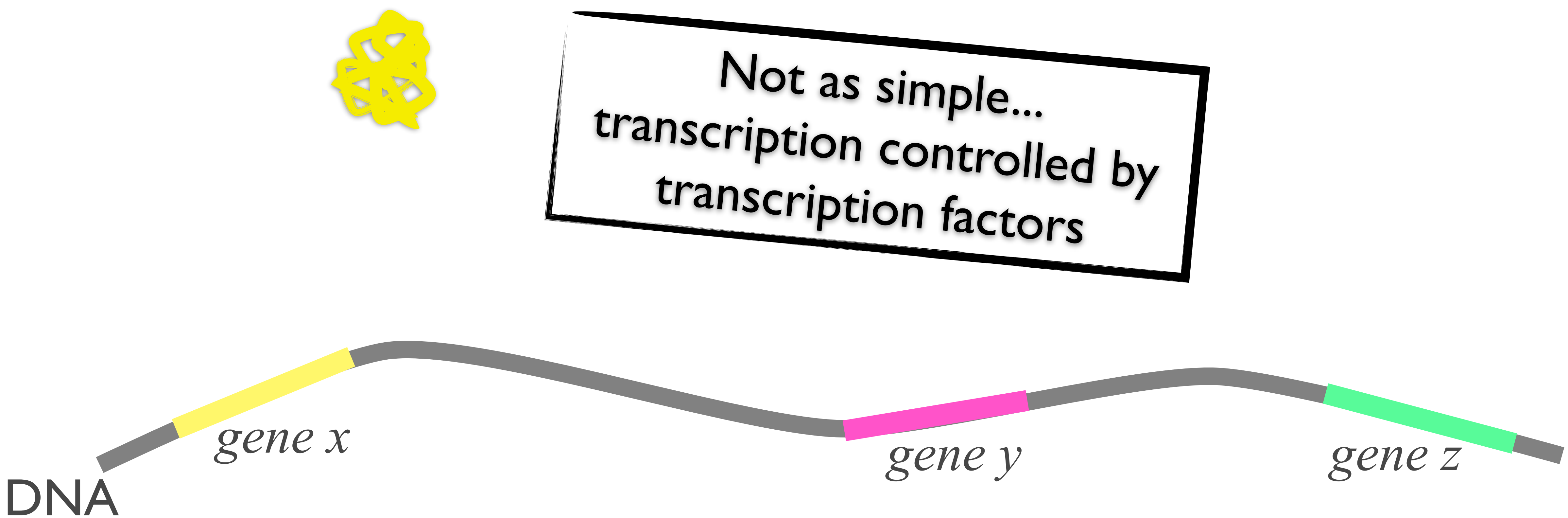


A simplistic illustration

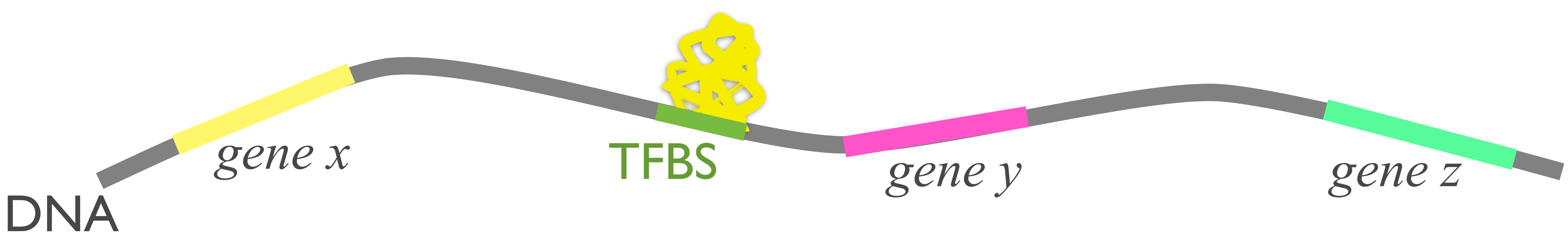


A simplistic illustration

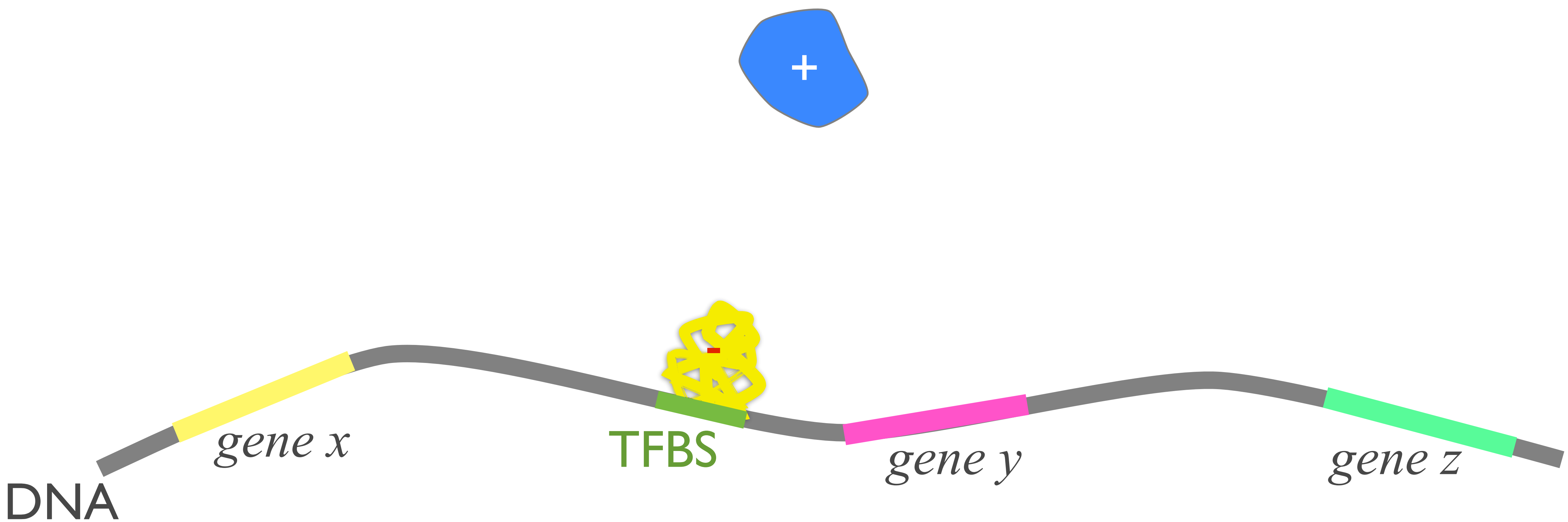
Not as simple...
transcription controlled by
transcription factors



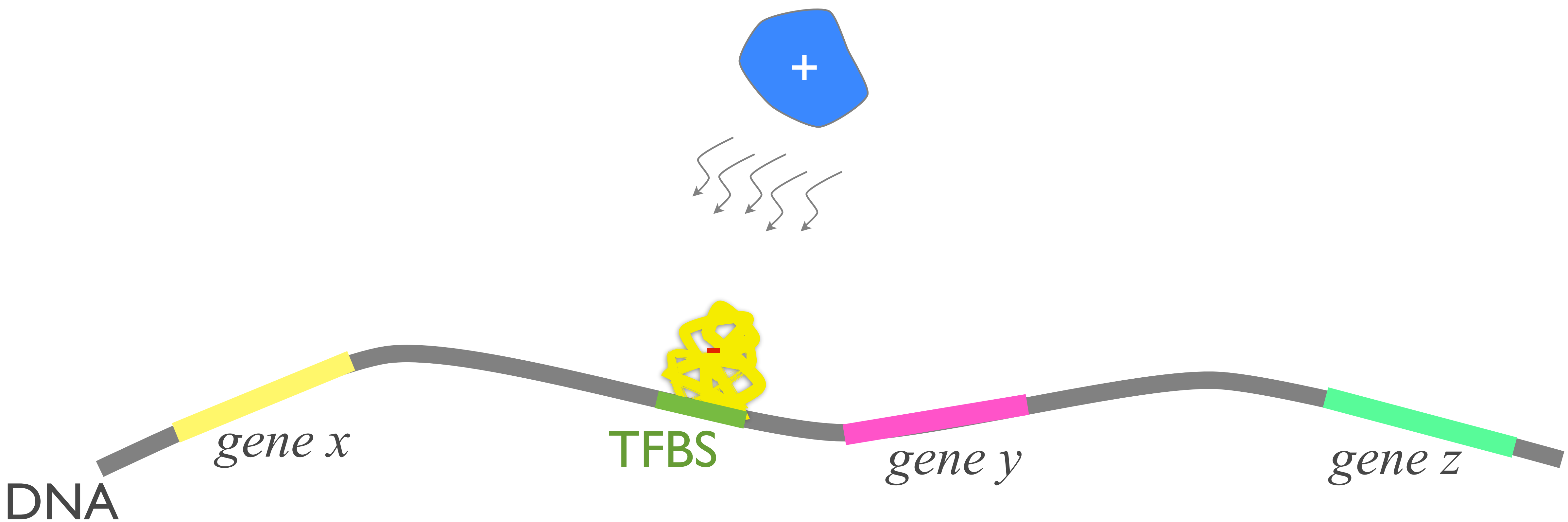
A simplistic illustration



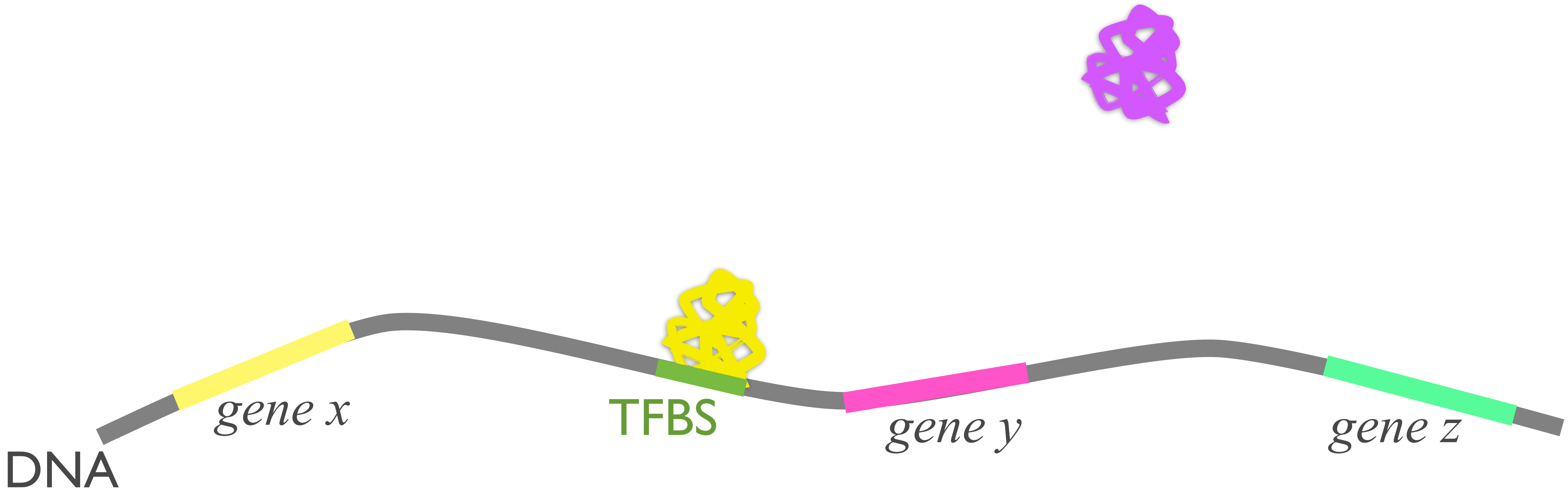
A simplistic illustration



A simplistic illustration



A simplistic illustration



A simplistic illustration

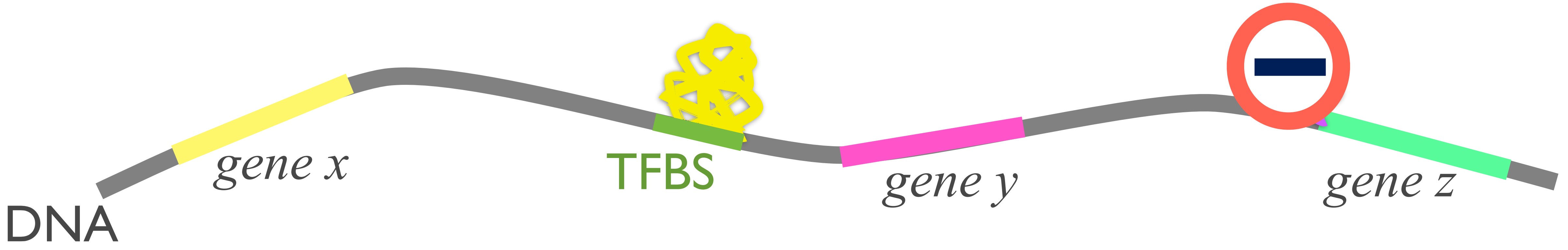
Major issue #1:

How to detect efficiently some potential and relevant targets patterns of TF on biological sequences ?

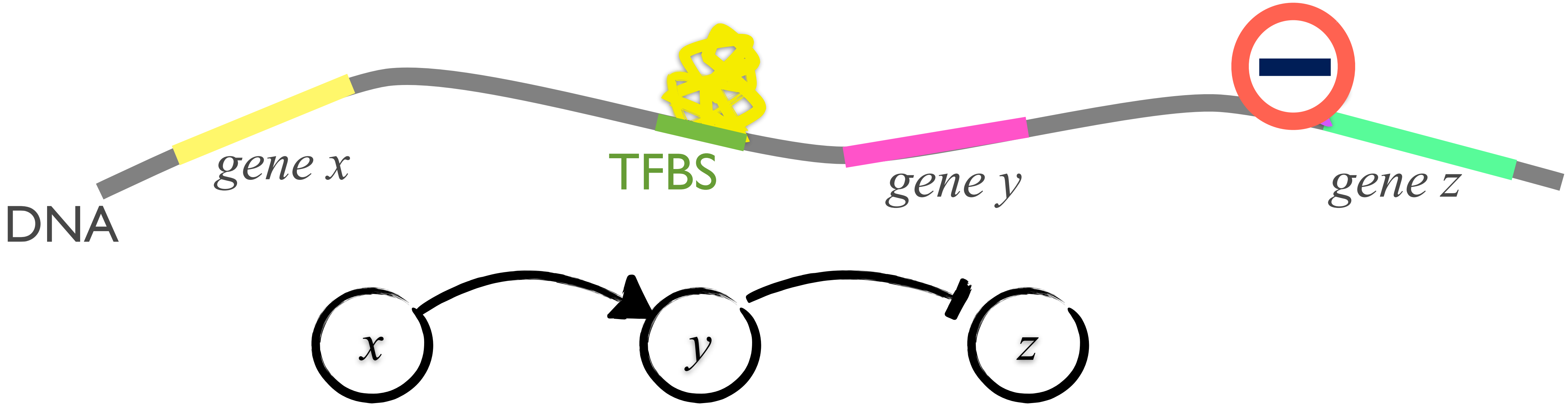
Genomics / metagenomic studies….

DNA

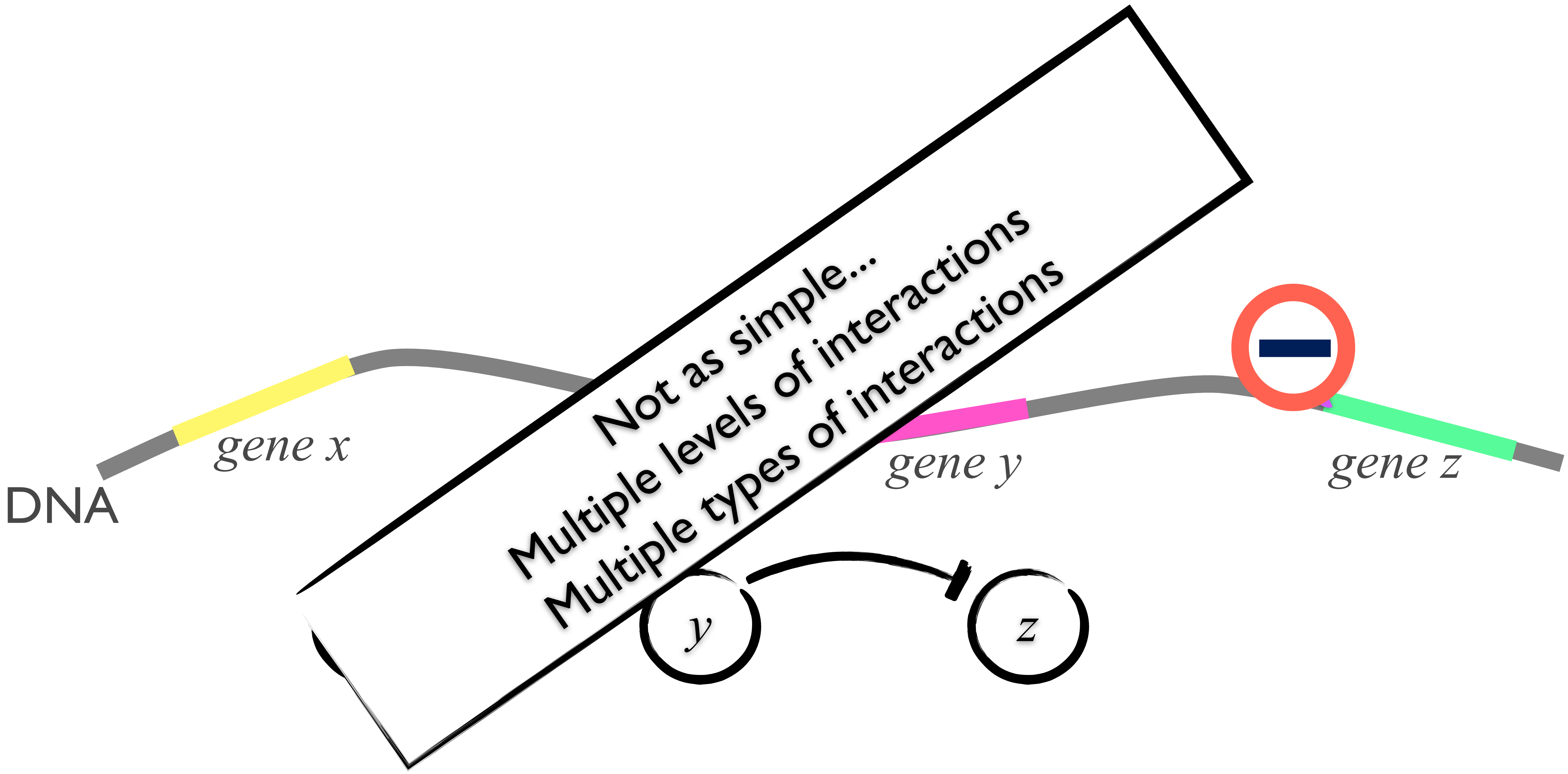
A simplistic illustration



A simplistic illustration



A simplistic illustration



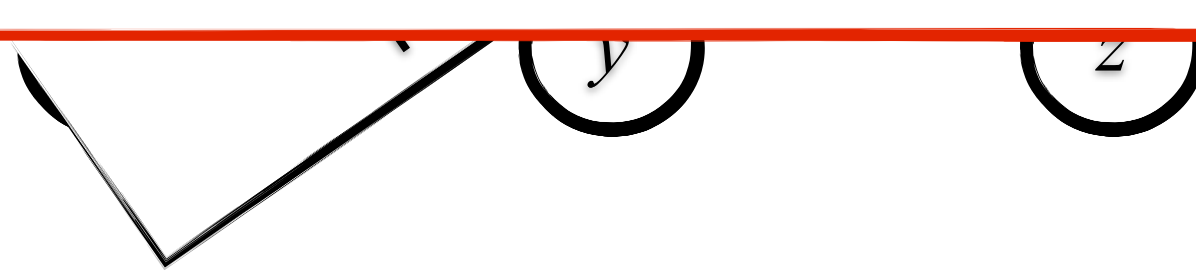
A simplistic illustration

Major issue #2:

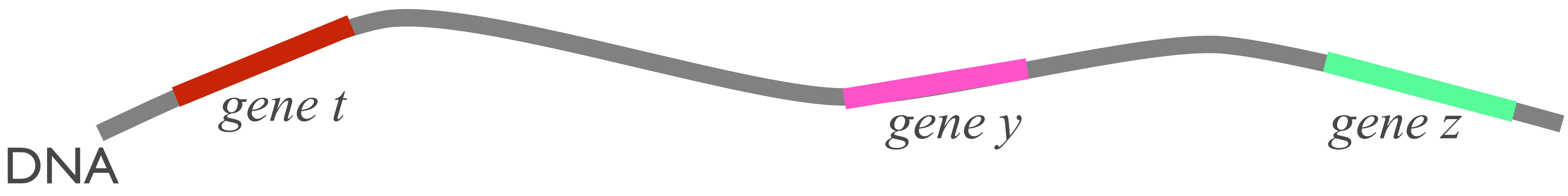
How to model a living system using (1) biological expert intuitions (individual level) and (2) experimental observations (population level) ?

Systems Biology : Regulatory Networks & data integration

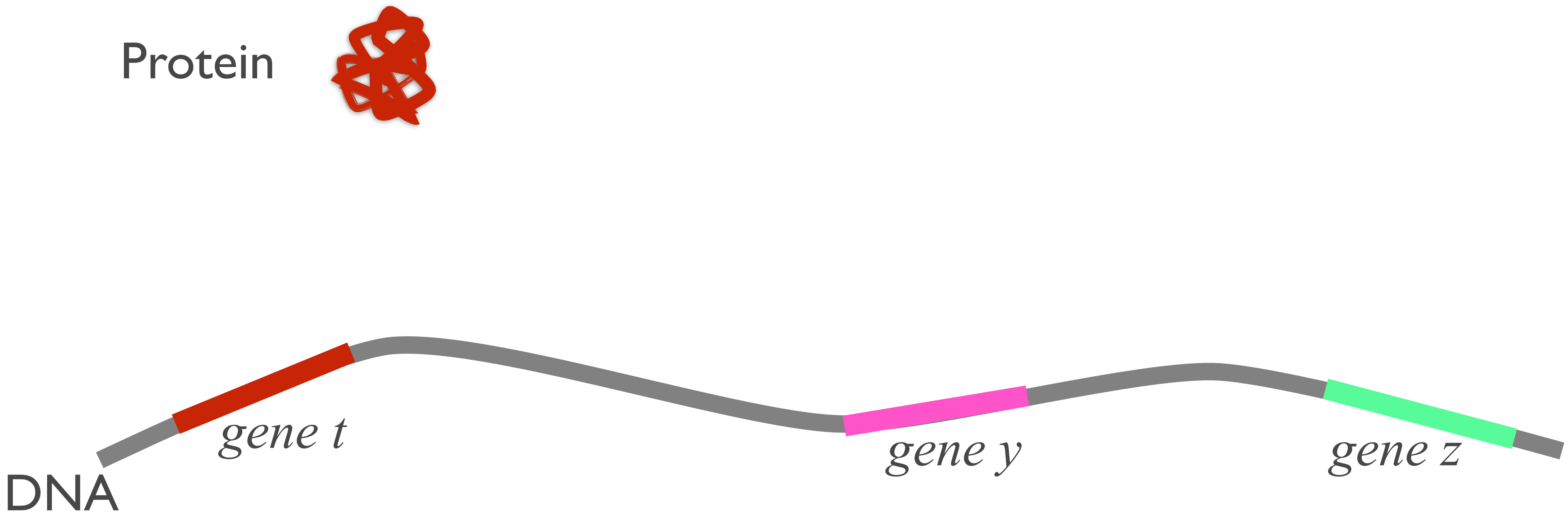
DN



A simplistic illustration



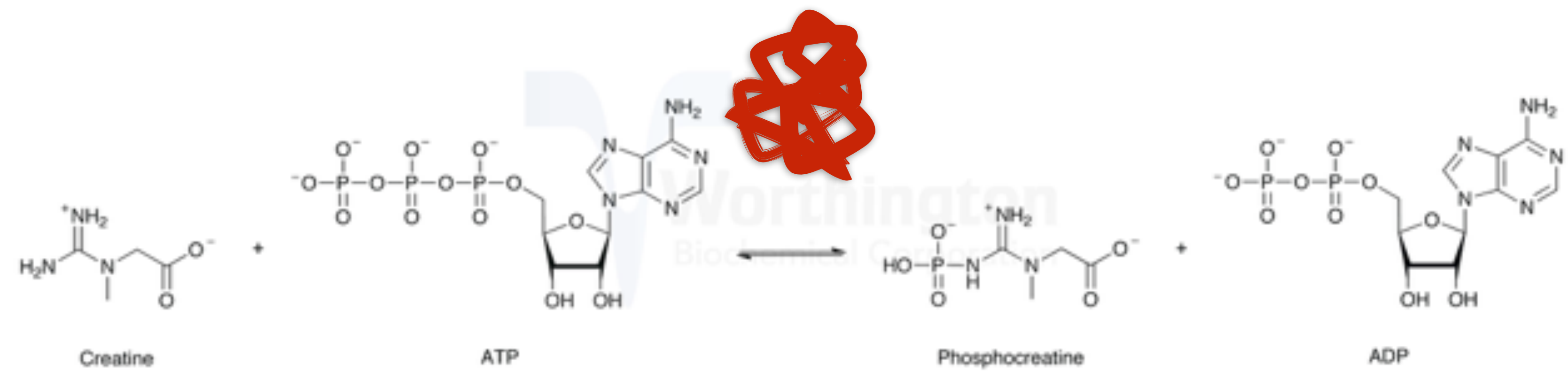
A simplistic illustration

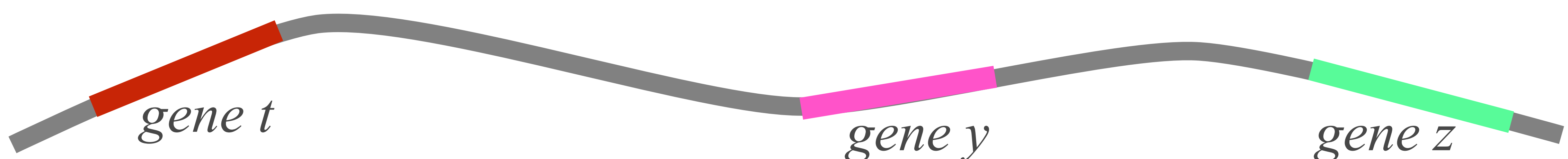


A simplistic illustration

Creatine kinase catabolyzes the chemical reaction

Protein

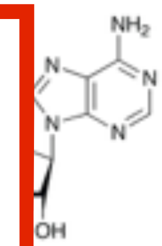




DNA

A simplistic illustration

Creatine kinase catabolyzes the chemical reaction



Major issue #3:

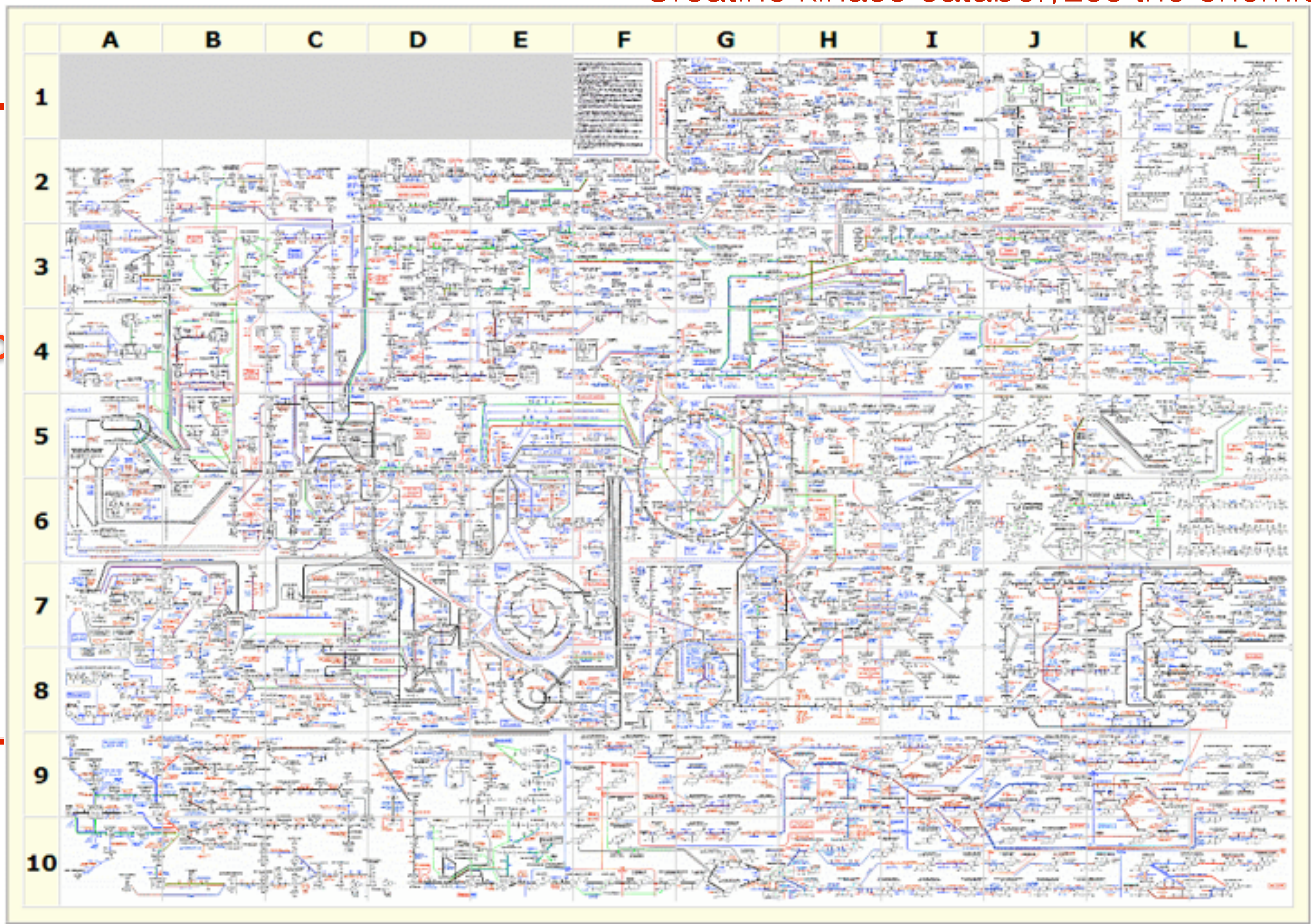
How to understand/predict/enhance the biochemical functioning/ productions of a cell/tissue/organism ?

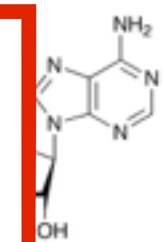
Systems Biology : Metabolic network analysis

A simplistic illustration

Creatine kinase catabolyzes the chemical reaction

Ho ng/





Mathematical modeling
a very simple task !

Hypothesis 0: from now, we consider the metabolic network composed by (and only by) the possible biochemical reactions (the enzyme is present or it is not required).

R1 : GLC + ATP <—> G6P
R2 : ATP + G6P —> 2 * G3P
R3 : ATP + GLC + G6P —> lactose
……

	R1	R2	R3
GLC	-1	0	-1
ATP	-1	-1	-1
G6P	+1	-1	-1
G3P	0	+2	0
lactose	0	0	+1

Stoichiometric matrix S

Mathematical modeling a very simple task !

R1 : GLC + ATP <—> G6P
R2 : ATP + G6P —> 2 * G3P
R3 : ATP + GLC + G6P —> lactose
……

	R1	R2	R3
GLC	-1	0	-1
ATP	-1	-1	-1
G6P	+1	-1	-1
G3P	0	+2	0
lactose	0	0	+1

Stoichiometric matrix S

The evolution of G6P concentration can be deduced by the speeds of reactions (i.e., fluxes) R1, R2 and R3

$$d[G6P](t)/dt = +1 * V_{R1} -1 * V_{R2} -1 * V_{R3}$$

Hypothesis 1: The metabolites can't accumulate in the cell.

$$0= d[G6P](t)/dt = +1 * V_{R1} -1 * V_{R2} -1 * V_{R3}$$

i.e.,: $V_{R1} = V_{R2} + V_{R3}$

Mathematical modeling
a very simple task !

R1 : GLC + ATP <—> G6P
R2 : ATP + G6P —> 2 * G3P
R3 : ATP + GLC + G6P —> lactose
……

	R1	R2	R3
GLC	-1	0	-1
ATP	-1	-1	-1
G6P	+1	-1	-1
G3P	0	+2	0
lactose	0	0	+1

Stoichiometric matrix S

The ev … ds of

Finding appropriate fluxes ≈ solving systems of linear equations !!!
—> quite easy !

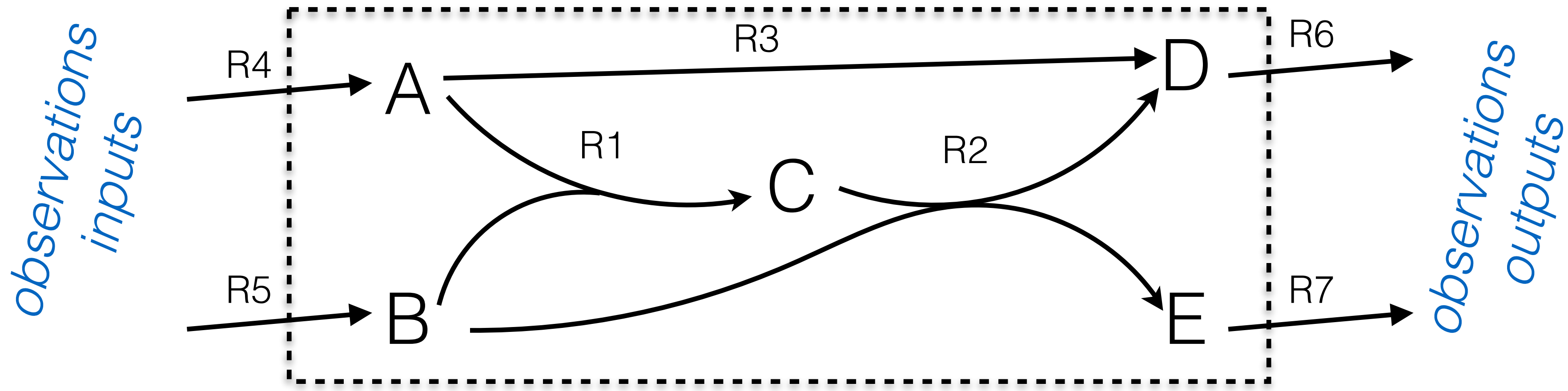
Hypothesis 1: The metabolites can't accumulate in the cell.

$$0 = d[G6P](t)/dt = +1 * V_{R1} - 1 * V_{R2} - 1 * V_{R3}$$

$$\text{i.e.,: } V_{R1} = V_{R2} + V_{R3}$$

Example

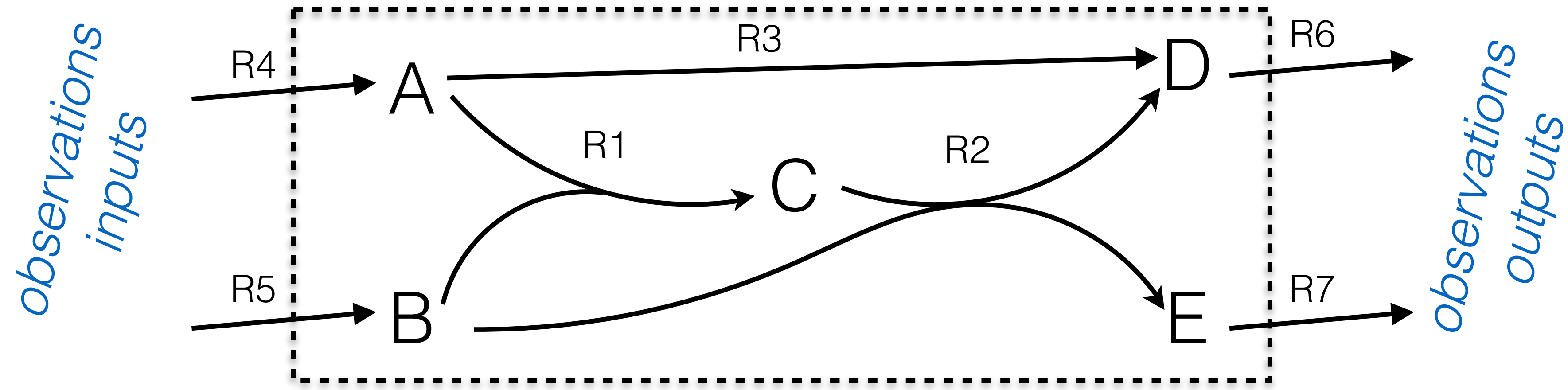
R1 : 2A + B —> C
R2 : B + C —> D + 2E
R3 : 4A —> D
R4 : —> A
R5 : —> B
R6 : D —>
R7 : E —>



Question 1 : Write the stoichiometric matrix
Question 2 : Write the linear system that has to be solved
Question 3 : Suppose that we observe V4 = 1 and V7 = 1 and solve the system
Question 4 : Suppose that we observe V4 = 1 and V7 = 2 and solve the system
Question 5 : All the reactions are irreversible. Plot the solution set. Suppose now that V4>1 and V7>1.
Question 6 : One has V4 < 6 and V7<4. What is the solution that maximizes V4+V7 ?

Example

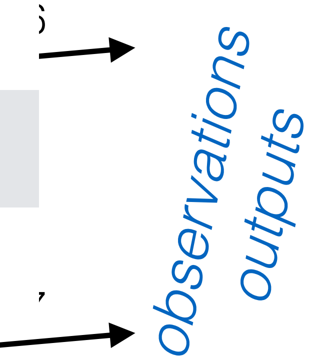
R1 : 2A + B —> C
R2 : B + C —> D + 2E
R3 : 4A —> D
R4 : —> A
R5 : —> B
R6 : D —>
R7 : E —>



Question 1 : Write the stoichiometric matrix

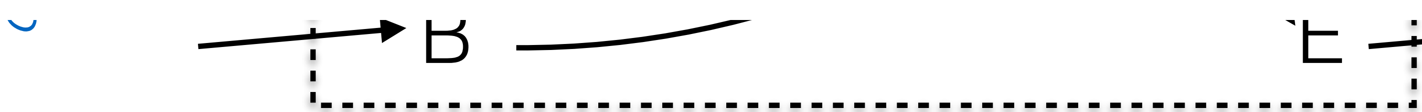
	R1	R2	R3	R4	R5	R6	R7
A	-2	0	-4	1	0	0	0
B	-1	-1	0	0	1	0	0
C	+1	-1	0	0	0	0	0
D	0	+1	1	0	0	-1	0
E	0	+2	0	0	0	0	-1

	R1	R2	R3	R4	R5	R6	R7
A	-2	0	-4	1	0	0	0
B	-1	-1	0	0	1	0	0
C	+1	-1	0	0	0	0	0
D	0	+1	1	0	0	-1	0
E	0	+2	0	0	0	0	-1



R6 : D —>

R7 : E —>



Question 1 : Write the stoichiometric matrix

Question 2 : Write the linear system that has to be solved

Question 3 : Suppose that we observe V4 = 1 and V7 = 1 and solve the system

Question 4 : Suppose that we observe V4 = 1 and V7 = 2 and solve the system

Question 5 : All the reactions are irreversible. Plot the solution set. Suppose now that V4>1 and V7>1.

Question 6 : One has V4 < 6 and V7<4. What is the solution that maximizes V4+V7 ?

	R1	R2	R3	R4	R5	R6	R7
A	-2	0	-4	1	0	0	0
B	-1	-1	0	0	1	0	0
C	+1	-1	0	0	0	0	0
D	0	+1	1	0	0	-1	0
E	0	+2	0	0	0	0	-1

observations
outputs

R6 : D —>

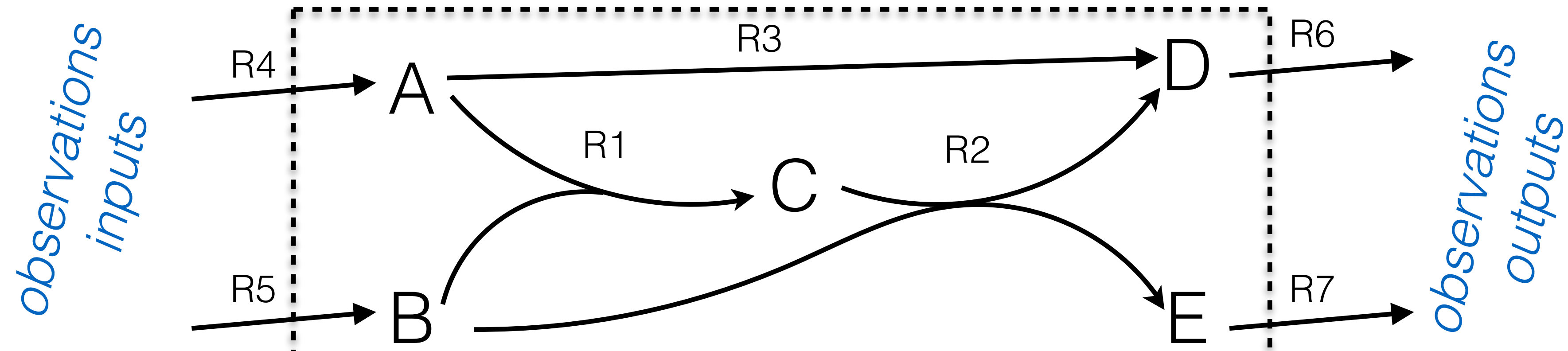
for A : V4 = 2 V1 + 4 V3
for B : V5 = V1 + V2
for C : V1 = V2
for D : V2 + V3 = V6
for E : 2 V2 = V7

V1 = V2, V5 = V7 = 2 V2, V6 = V2 + V3 and V4 = 2 V2 +4 V3

only 2 unknowns but we must observe one of V4 or V6 and one of V5 or V7.

Example

R1 : 2A + B —> C
R2 : B + C —> D + 2E
R3 : 4A —> D
R4 : —> A
R5 : —> B
R6 : D —>



for A : V4 = 2 V1 + 4 V3
for B : V5 = V1 + V2
for C : V1 = V2
for D : V2 + V3 = V6
for E : 2 V2 = V7

V1 = V2, V5 = V7 = 2 V2, V6 = V2 + V3 and V4 = 2 V2 +4 V3

only 2 unknowns but we must observe one of V4 or V6 and one of V5 or V7.

for A : $V4 = 2\ V1 + 4\ V3$
for B : $V5 = V1 + V2$
for C : $V1 = V2$
for D : $V2 + V3 = V6$
for E : $2\ V2 = V7$

$V1 = V2$, $V5 = V7 = 2\ V2$, $V6 = V2 + V3$ and $V4 = 2\ V2 + 4\ V3$

only 2 unknowns but we must observe one of V4 or V6 and one of V5 or V7.

observations
outputs

Question 1 : Write the stoichiometric matrix
Question 2 : Write the linear system that has to be solved
Question 3 : Suppose that we observe $V4 = 1$ and $V7 = 1$ and solve the system
Question 4 : Suppose that we observe $V4 = 1$ and $V7 = 2$ and solve the system
Question 5 : All the reactions are irreversible. Plot the solution set. Suppose now that $V4>1$ and $V7>1$.
Question 6 : One has $V4 < 6$ and $V7<4$. What is the solution that maximizes $V4+V7$?

for A : V4 = 2 V1 + 4 V3
for B : V5 = V1 + V2
for C : V1 = V2
for D : V2 + V3 = V6
for E : 2 V2 = V7

V1 = V2, V5 = V7 = 2 V2, V6 = V2 + V3 and V4 = 2 V2 +4 V3

only 2 unknowns but we must observe one of V4 or V6 and one of V5 or V7.

observations
outputs

Question 1 : Write the stoichiometric matrix
Question 2 : Write the linear system that has to be solved
Question 3 : Suppose that we observe V4 = 1 and V7 = 1 and solve the system
Question 4 : Suppose that we observe V4 = 1 and V7 = 2 and solve the system
Question 5 : All the reactions are irreversible. Plot the solution set. Suppose now that V4>1 and V7>1.
Question 6 : One has V4 < 6 and V7<4. What is the solution that maximizes V4+V7 ?

V1 = V2 = 0.5, V3 = 0, V4 = 1, V5 = 1, V6 = 0.5, V7 = 1

for A : V4 = 2 V1 + 4 V3
for B : V5 = V1 + V2
for C : V1 = V2
for D : V2 + V3 = V6
for E : 2 V2 = V7

V1 = V2, V5 = V7 = 2 V2, V6 = V2 + V3 and V4 = 2 V2 +4 V3

only 2 unknowns but we must observe one of V4 or V6 and one of V5 or V7.

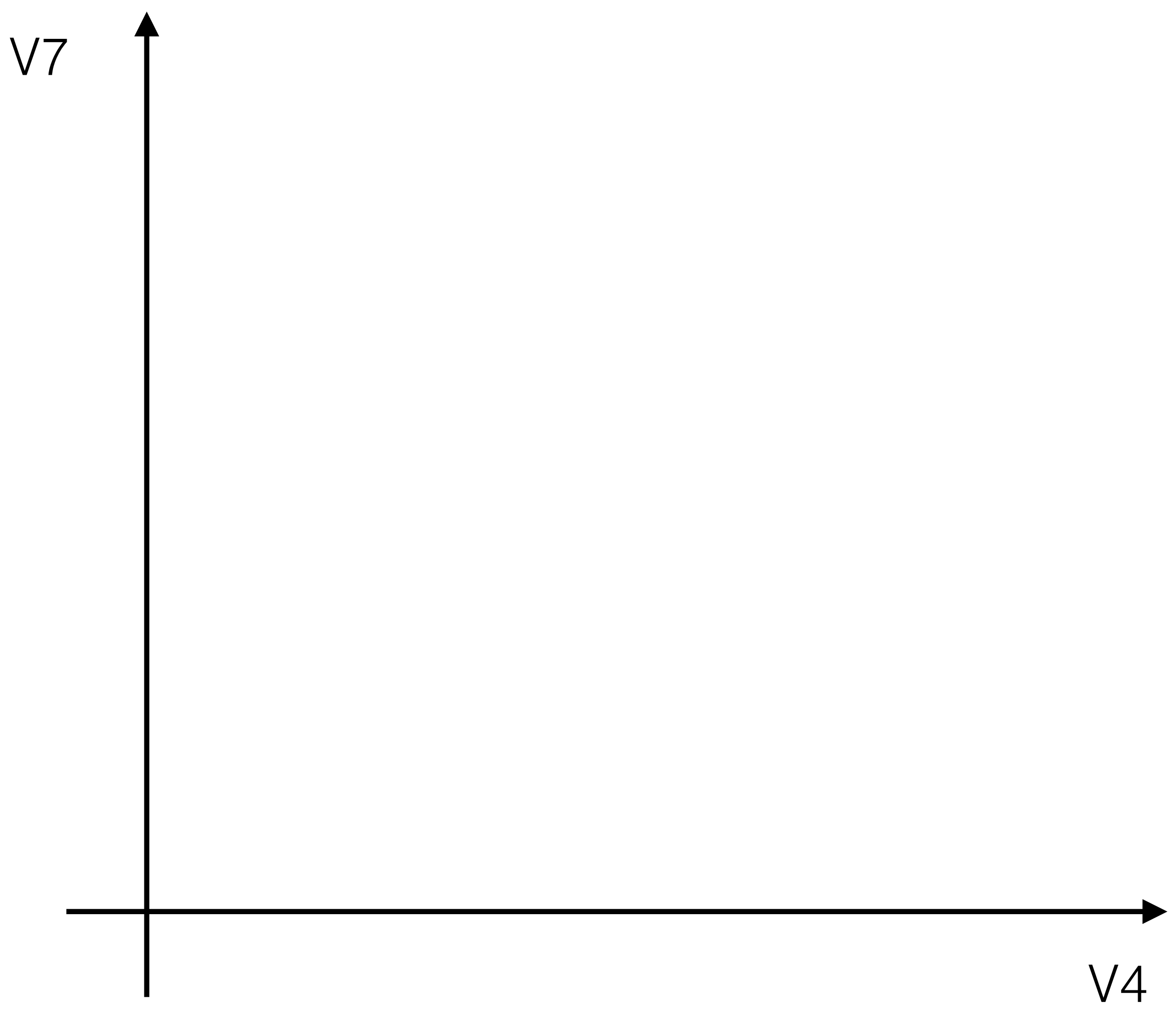
observations
outputs

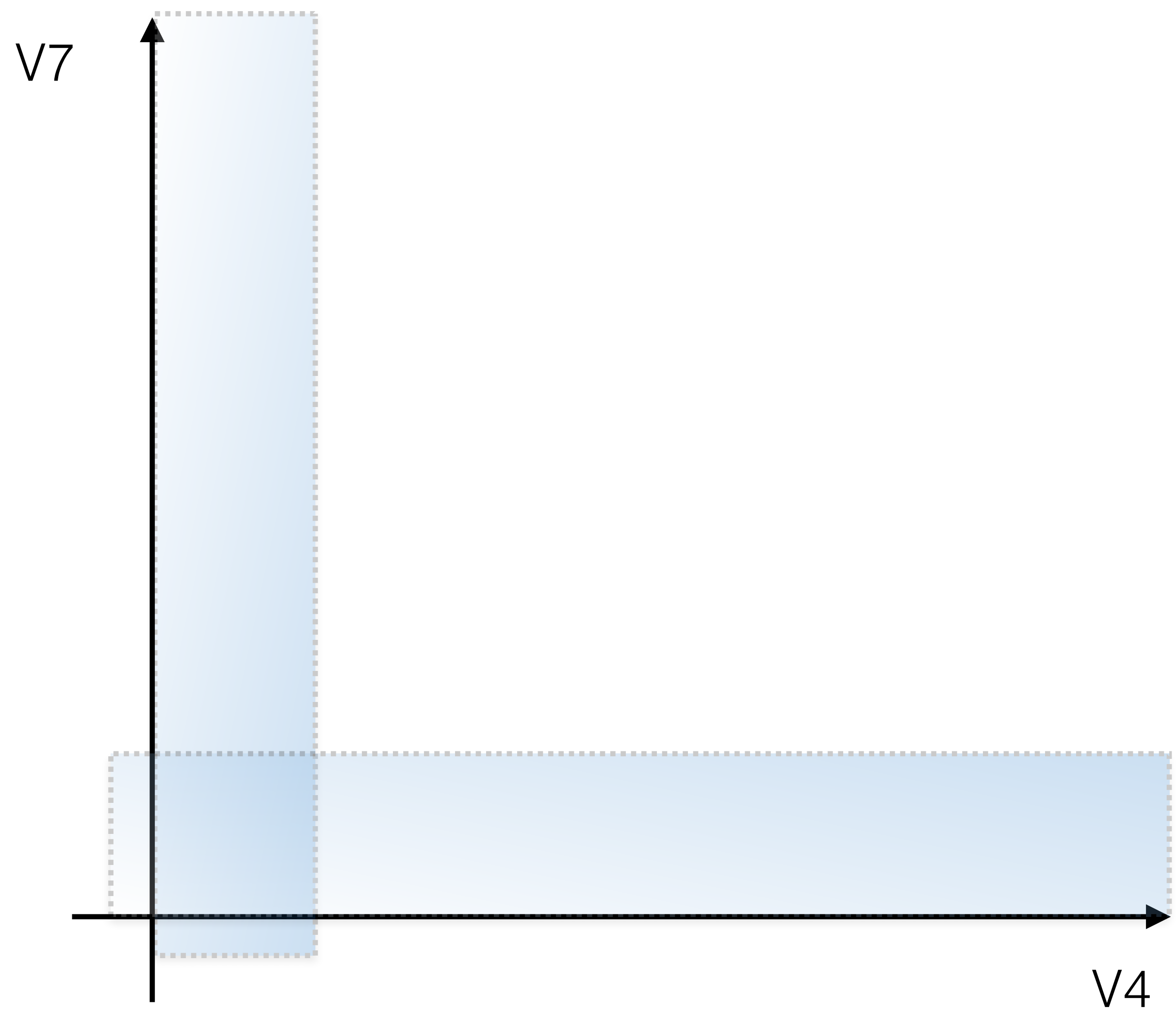
Question 1 : Write the stoichiometric matrix
Question 2 : Write the linear system that has to be solved
Question 3 : Suppose that we observe V4 = 1 and V7 = 1 and solve the system
Question 4 : Suppose that we observe V4 = 1 and V7 = 2 and solve the system
Question 5 : All the reactions are irreversible. Plot the solution set. Suppose now that V4>1 and V7>1.
Question 6 : One has V4 < 6 and V7<4. What is the solution that maximizes V4+V7 ?

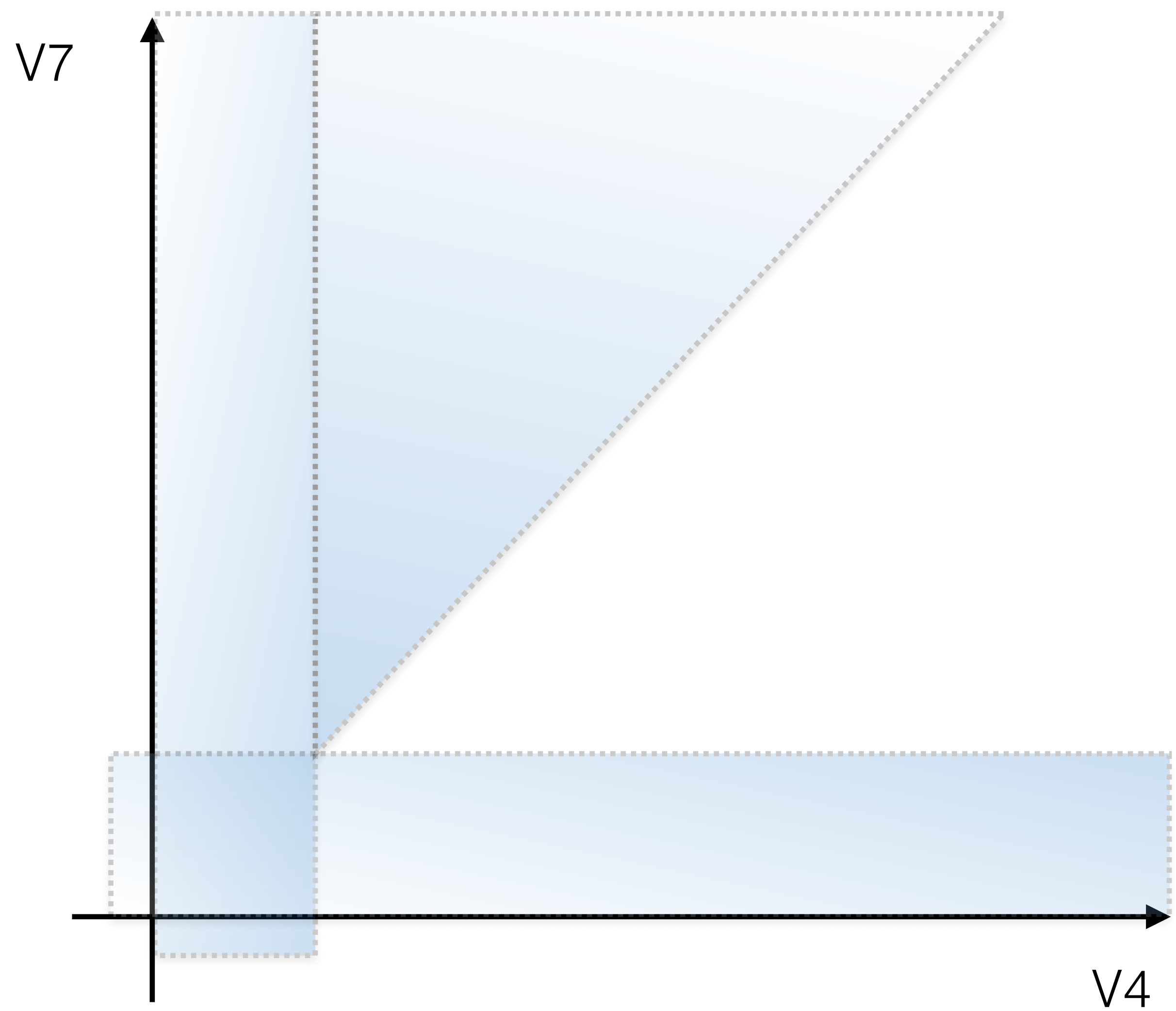
V1 = V2 = 0.5, V3 = 0, V4 = 1, V5 = 1, V6 = 0.5, V7 = 1

V1 = V2 = 1, V3 = **-0.25**, V4 = 1, V5 = 2, V6 = 0.75, V7 = 1

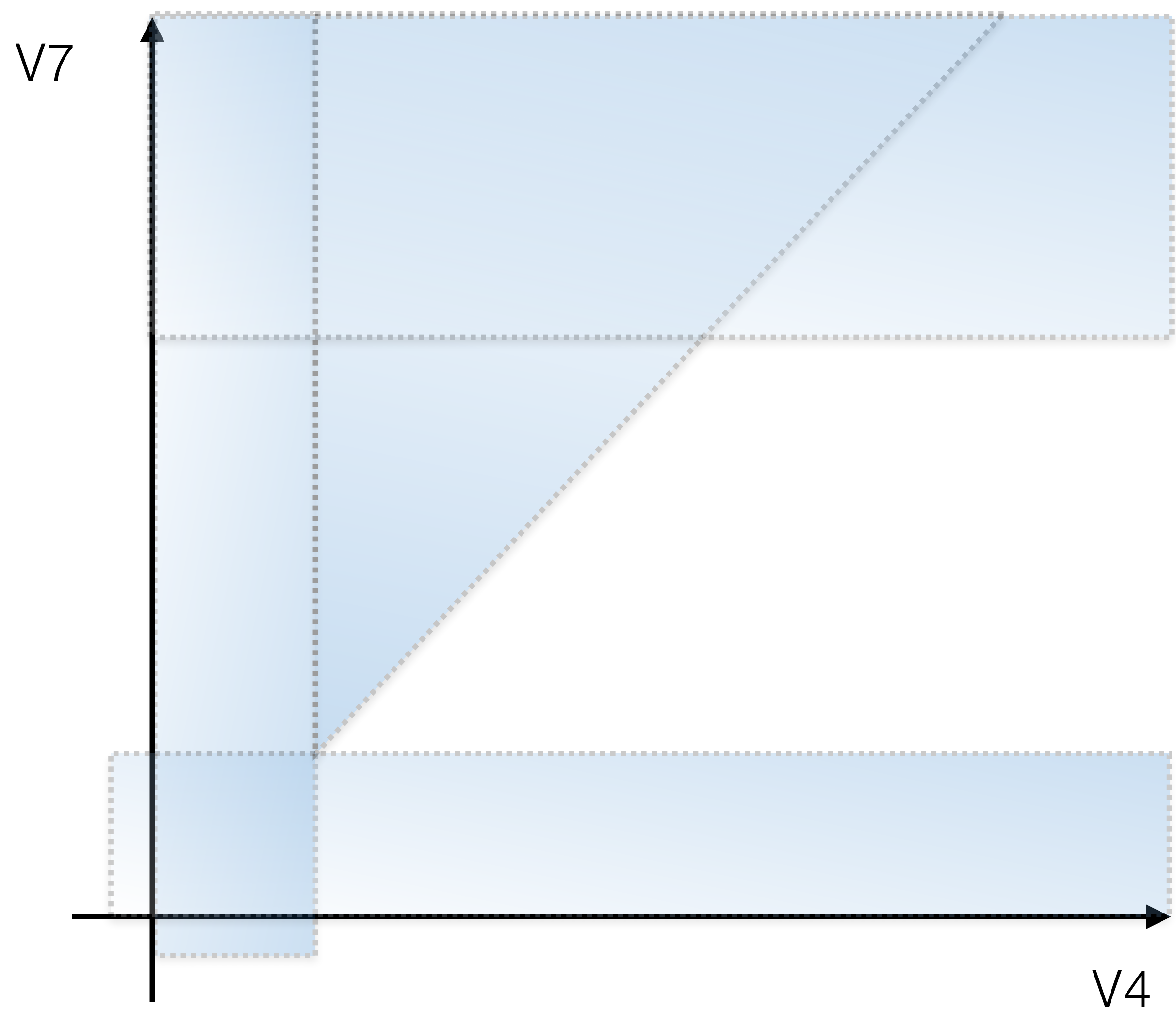
The flux of an irreversible reaction is necessarily positive !!

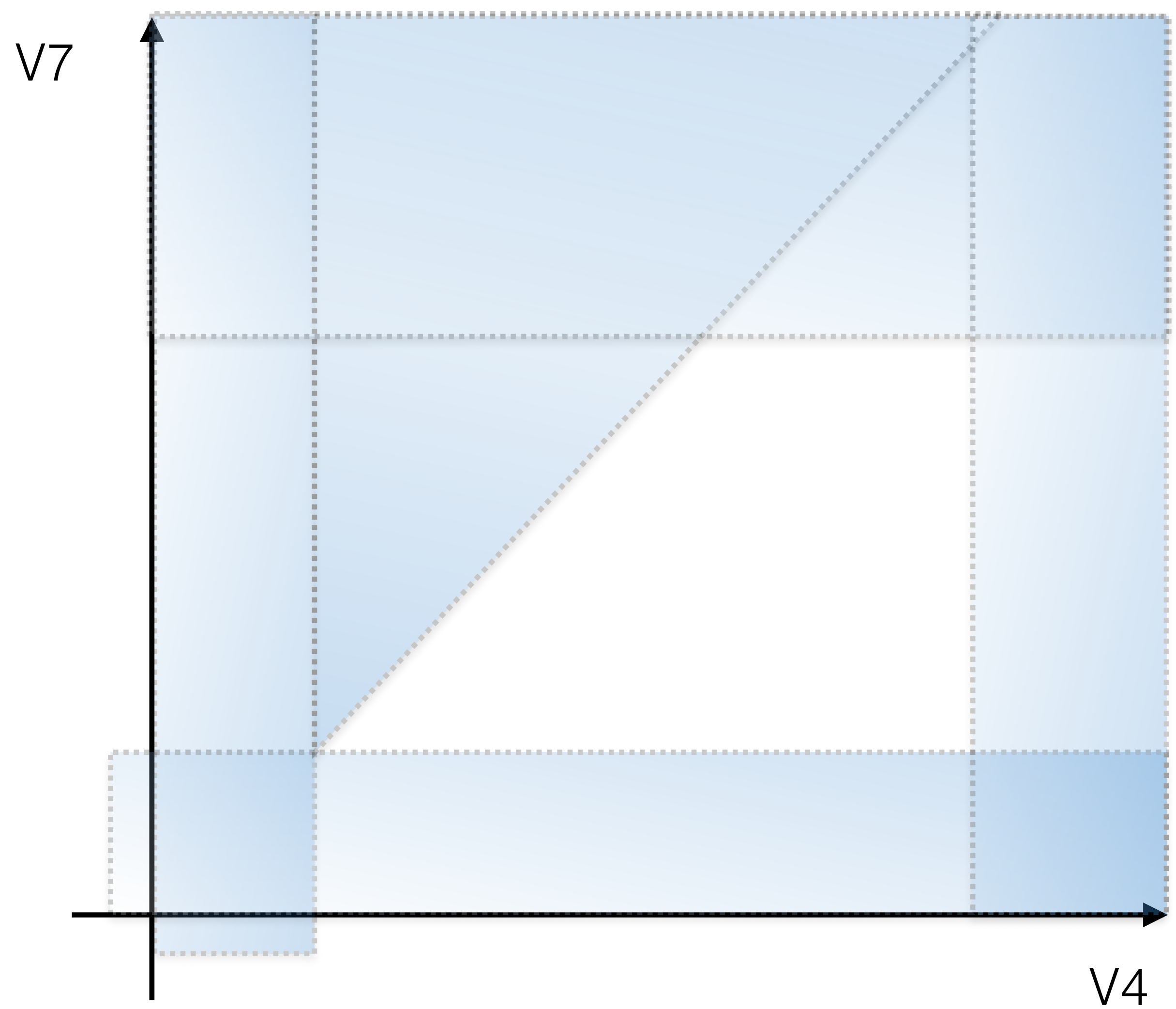

V7
V4

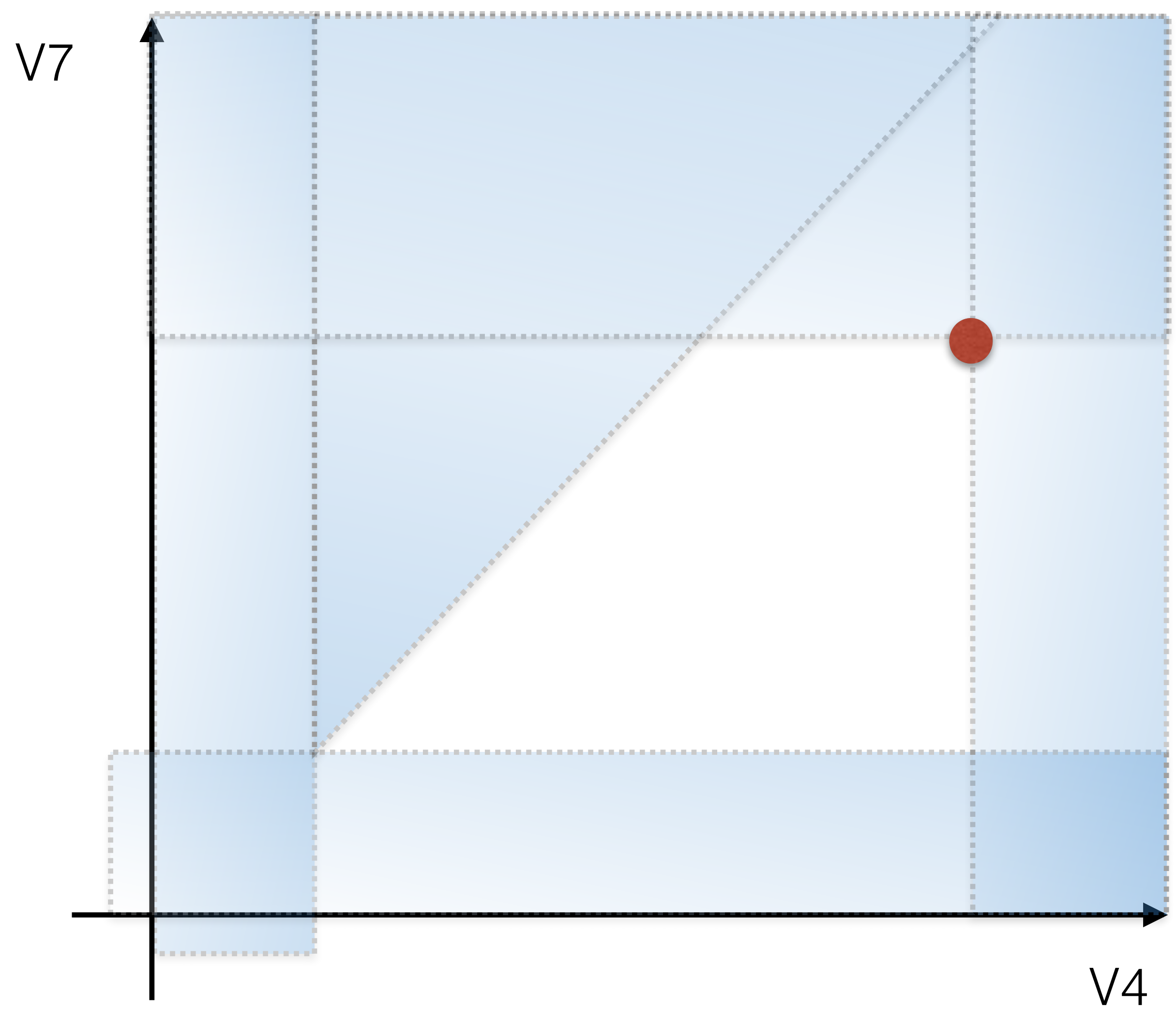

V7
V4

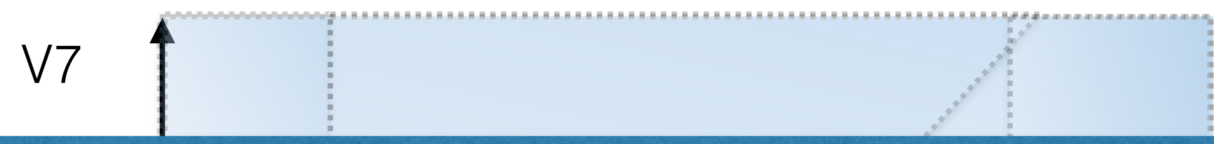


V7
V4


V7
V4


V7
V4


V7
V4



In higher dimensions,

1. The figure defined by inequalities on linear expressions is always convex (it is a convex polyedra also known as a simplex)

2. The optimal value of a linear expression is always on a vertex of the simplex (or on a face for very particular objectives)

3. Very efficient algorithms (such as the simplex algorithm) can find these optima

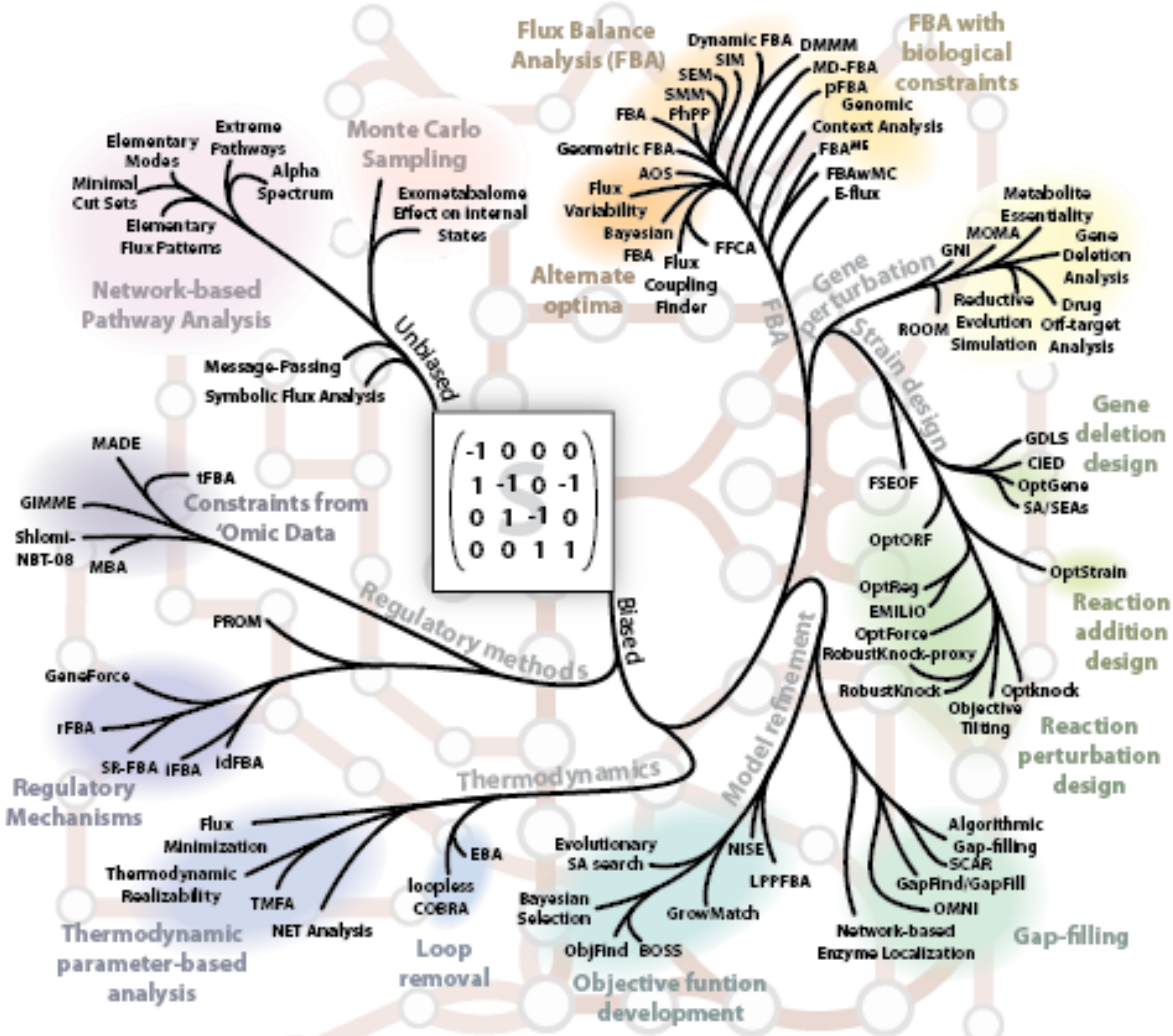
Why do we need a computer ?

- Current human metabolism has
  - 1,789 enzyme-encoding genes,
  - 7,440 reactions
  - 2,626 unique metabolites

- Study of the flux variations => 7440 * 2 * FBA
- Study of single gene knock out => 1789 * FBA
- Study of double gene knock out => 1789^2 * FBA

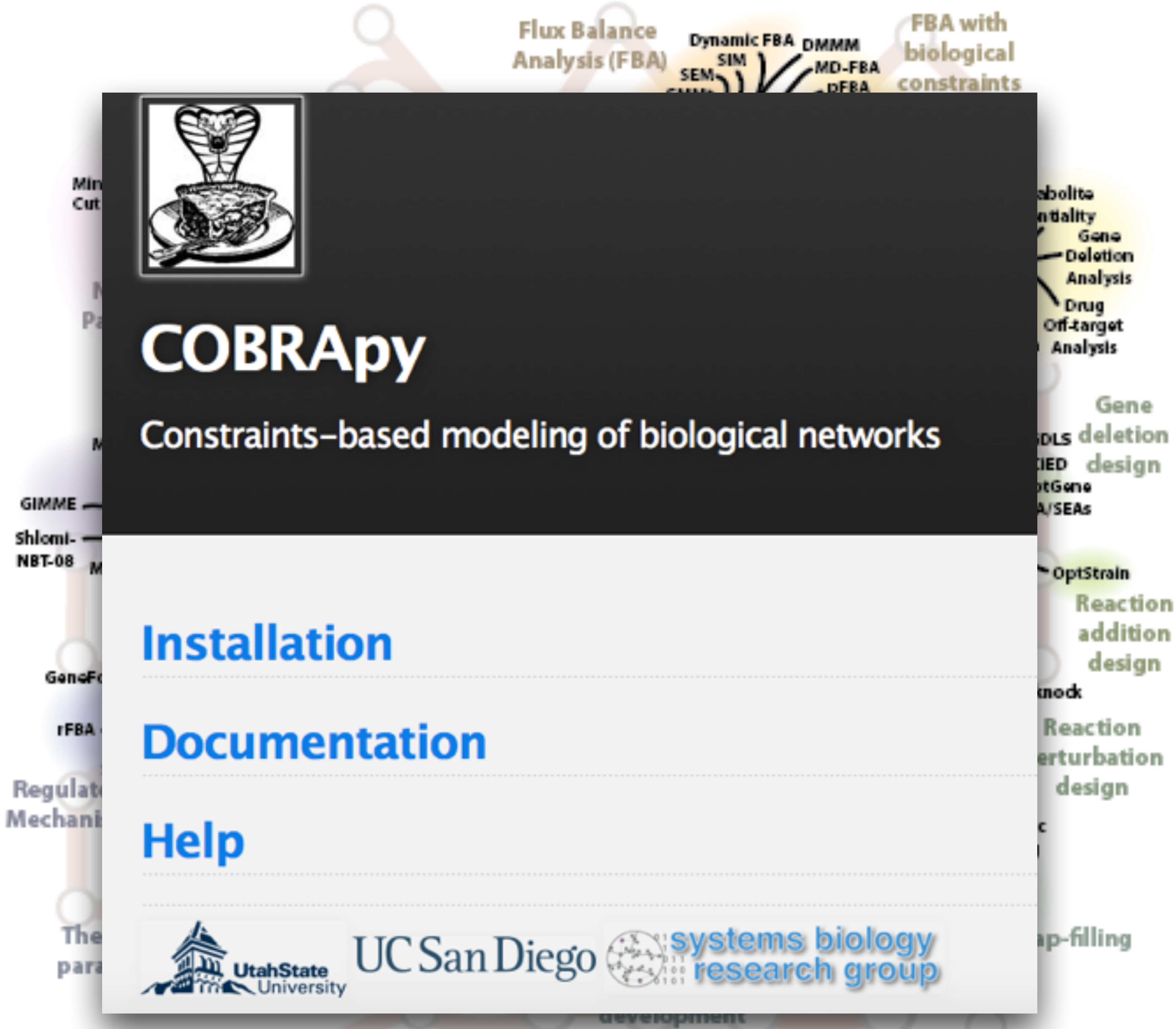
- Many other kind of analysis => A. Bockmayr on friday!

Many tools



Flux Balance Analysis (FBA)
FBA with biological constraints
Dynamic FBA
DMMM
SIM
MD-FBA
SEM
pFBA
SMM
PhPP
Genomic Context Analysis
FBA
Geometric FBA
AOS
Flux Variability
Bayesian FBA
FFCA
Flux Coupling Finder
FBAwMC
E-flux
Alternate optima
Monte Carlo Sampling
Exometabalome Effect on internal States
Elementary Modes
Extreme Pathways
Alpha Spectrum
Minimal Cut Sets
Elementary Flux Patterns
Network-based Pathway Analysis
Unbiased
Message-Passing
Symbolic Flux Analysis
Metabolite Essentiality
MOMA
GNI
Gene Deletion Analysis
Reductive Evolution Simulation
ROOM
Drug Off-target Analysis
Gene perturbation
FBA
Strain design
Gene deletion design
GDLS
CIED
OptGene
SA/SEAs
FSEOF
OptORF
OptStrain
Reaction addition design
OptReg
EMILiO
OptForce
RobustKnock-proxy
RobustKnock
Optknock
Objective Tilting
Reaction perturbation design
MADE
tFBA
GIMME
Constraints from 'Omic Data
Shlomi-NBT-08
MBA
Regulatory methods
PROM
Biased
GeneForce
rFBA
SR-FBA
iFBA
idFBA
Regulatory Mechanisms
Thermodynamics
Model refinement
Flux Minimization
Thermodynamic Realizability
TMFA
NET Analysis
Thermodynamic parameter-based analysis
EBA
loopless COBRA
Loop removal
Evolutionary SA search
NISE
LPPFBA
Bayesian Selection
GrowMatch
ObjFind
BOSS
Objective funtion development
Algorithmic Gap-filling
SCAR
GapFind/GapFill
OMNI
Network-based Enzyme Localization
Gap-filling

Many tools


Flux Balance Analysis (FBA)
Dynamic FBA
SIM
DMMM
MD-FBA
SEM
FBA with biological constraints
COBRApy
Constraints-based modeling of biological networks
Installation
Documentation
Help
UtahState University
UC San Diego
systems biology research group
Gene Deletion Analysis
Drug Off-target Analysis
Gene deletion design
GIMME
Shlomi-NBT-08
OptStrain
Reaction addition design
Reaction perturbation design
rFBA