

The background of the slide is a complex, abstract pattern of thin, white, intersecting lines that form a dense, three-dimensional lattice. Scattered throughout this lattice are several semi-transparent, blue, spherical objects of varying sizes, some of which appear to be glowing or reflecting light, giving the overall image a scientific and technical feel.

# DATA INTEGRATION OF STOICHIOMETRIC MODELS

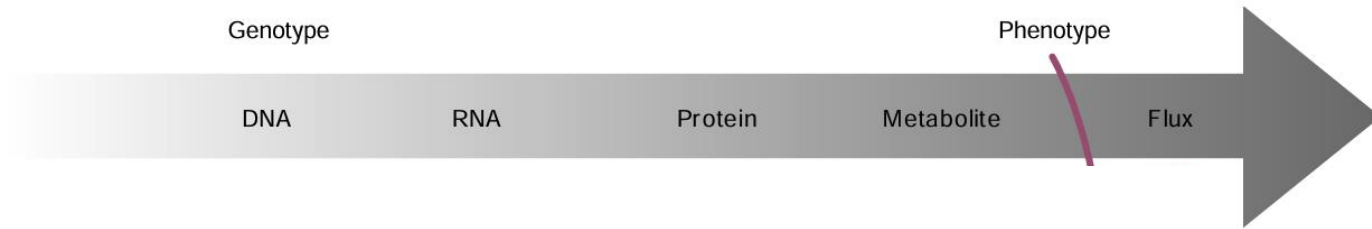
Principles and Applications of Systems Biology

**EPFL**

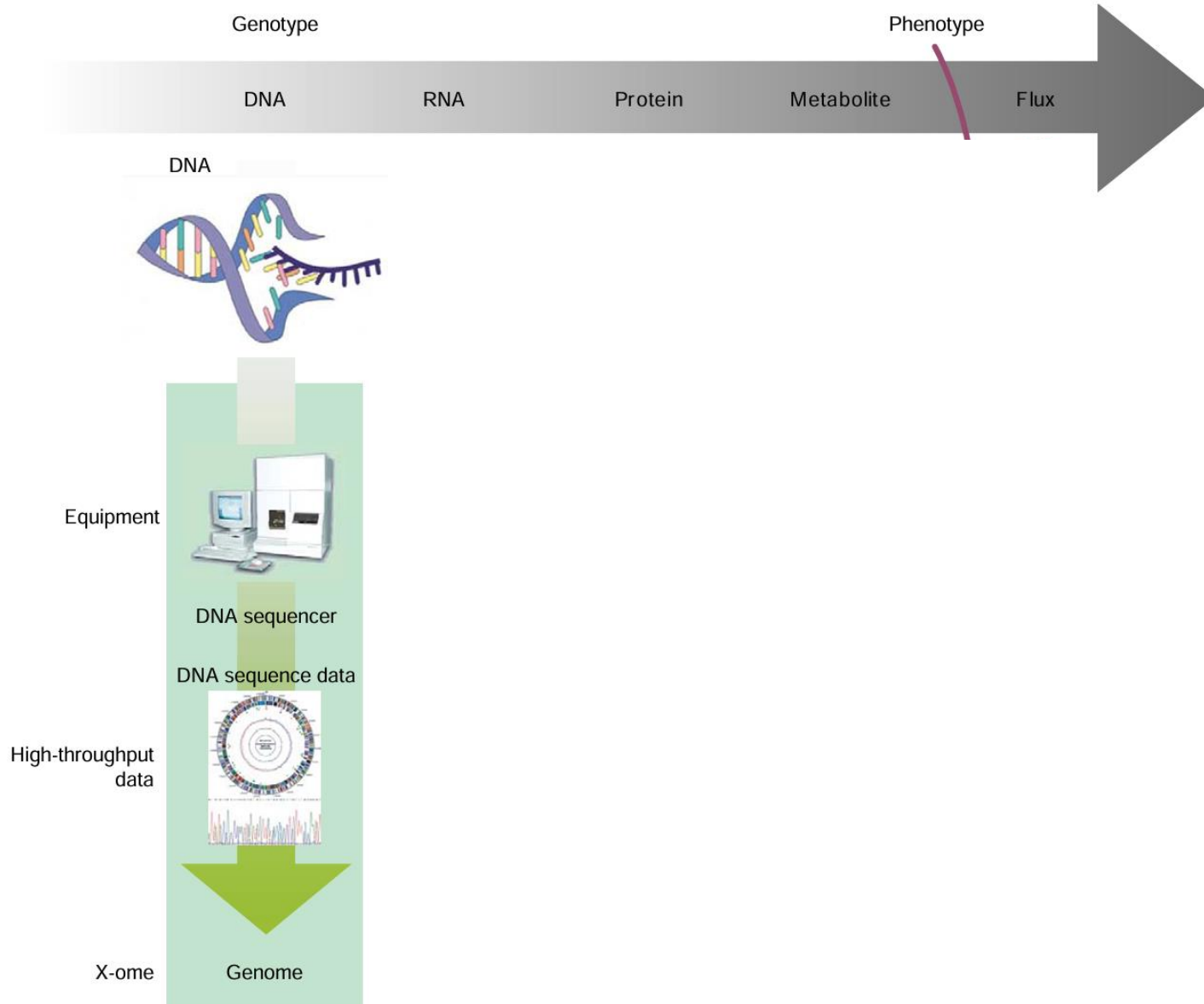
Vassily Hatzimanikatis

October 2025

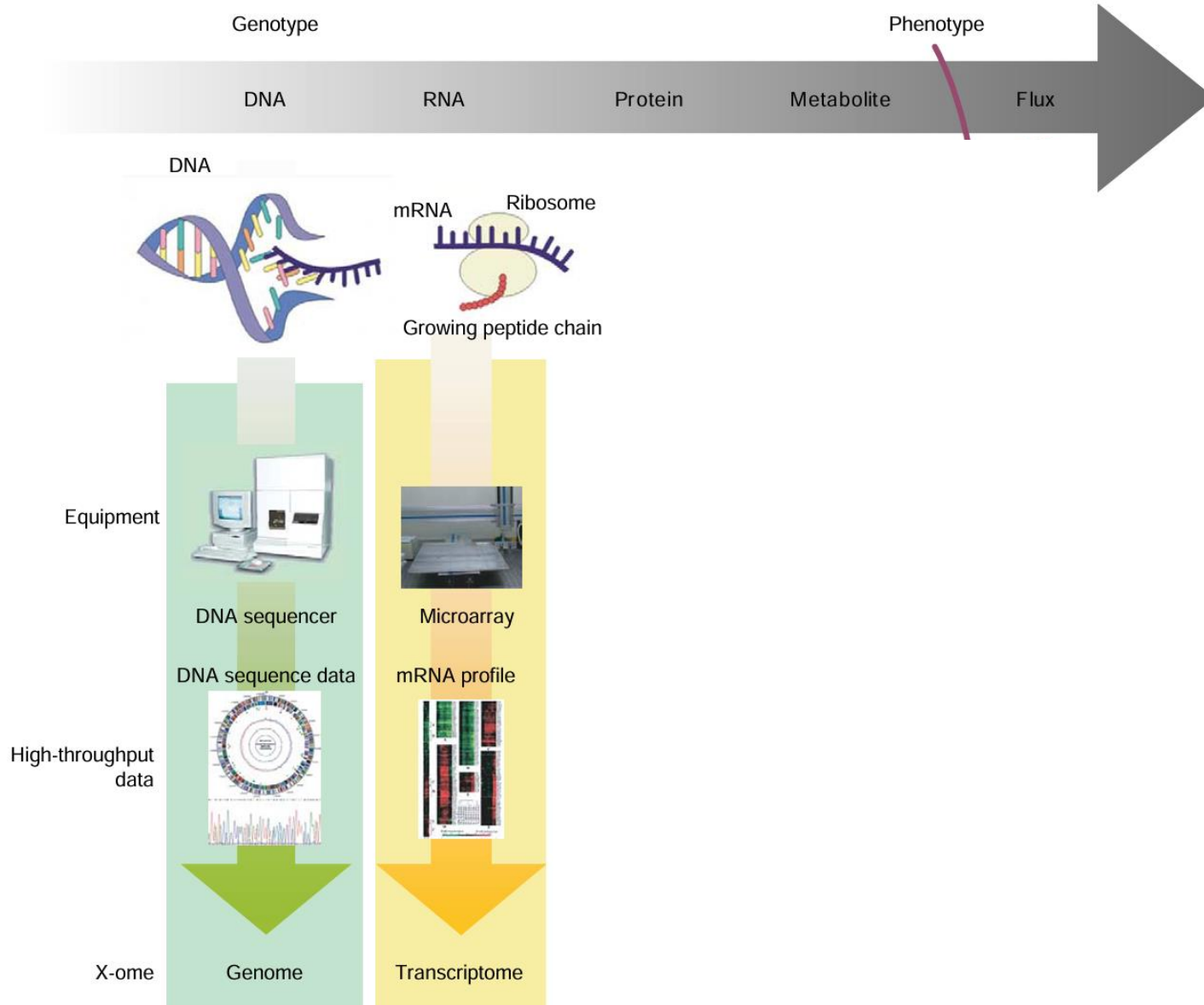
# From central dogma to omics



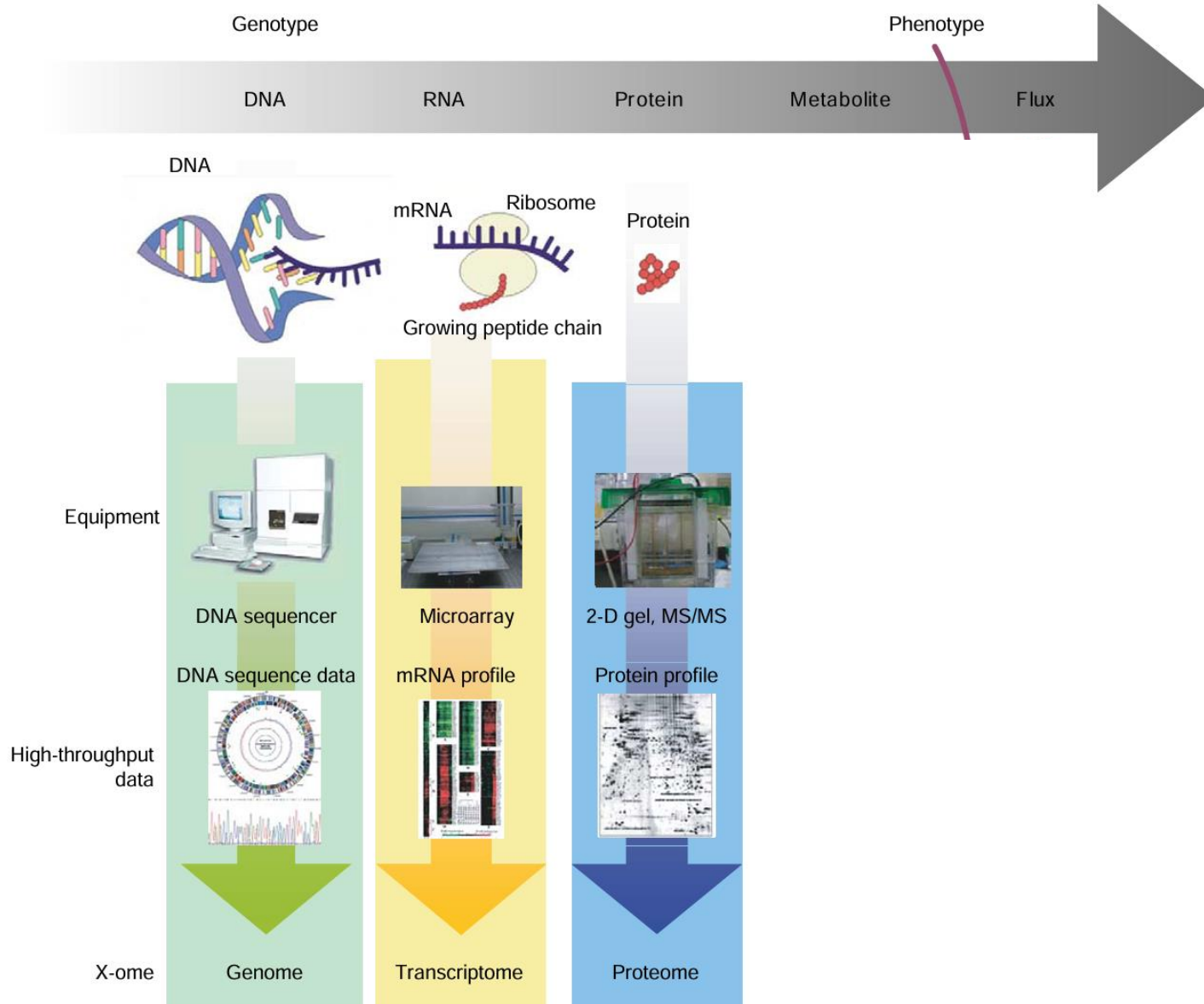
# From central dogma to omics



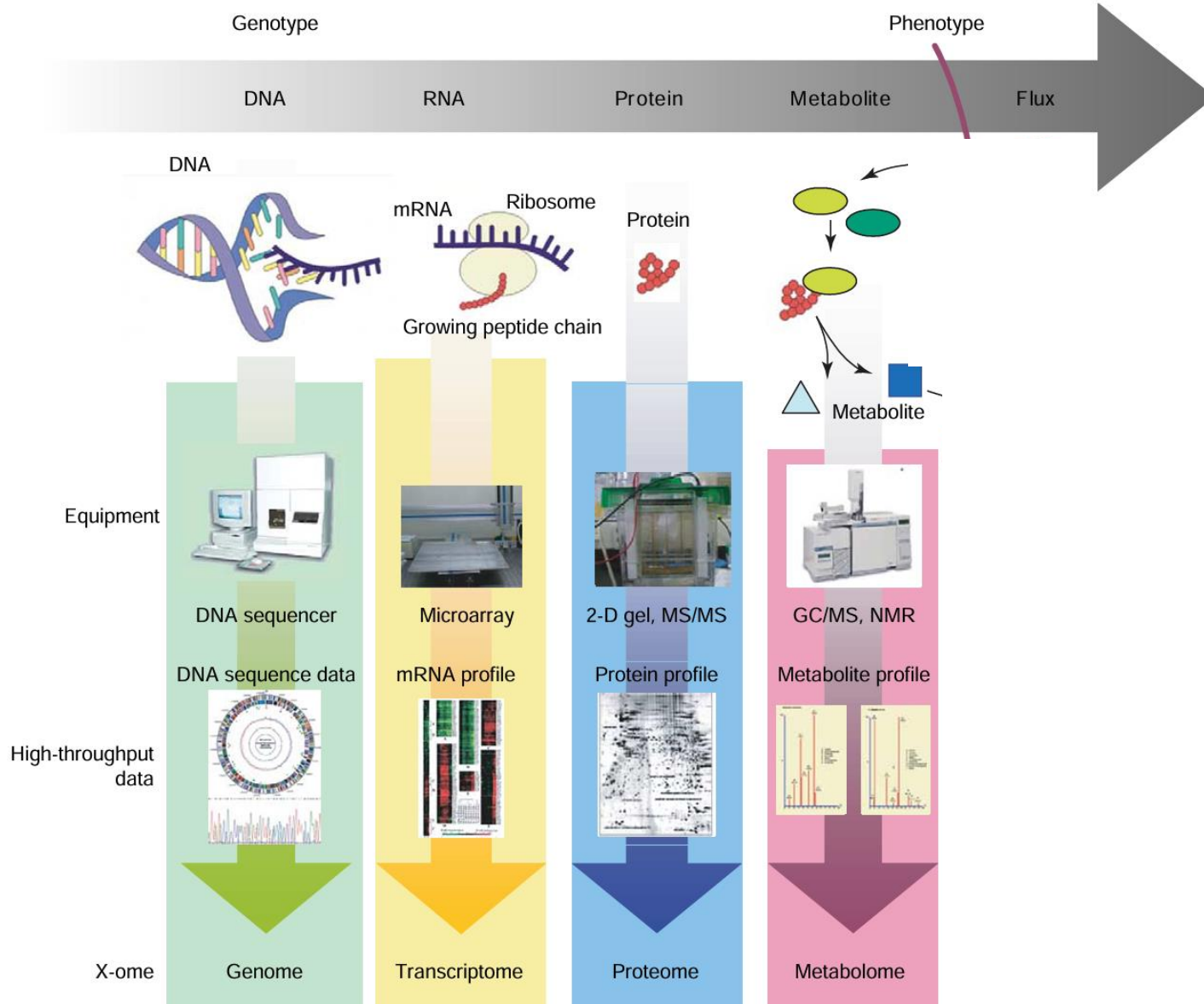
# From central dogma to omics



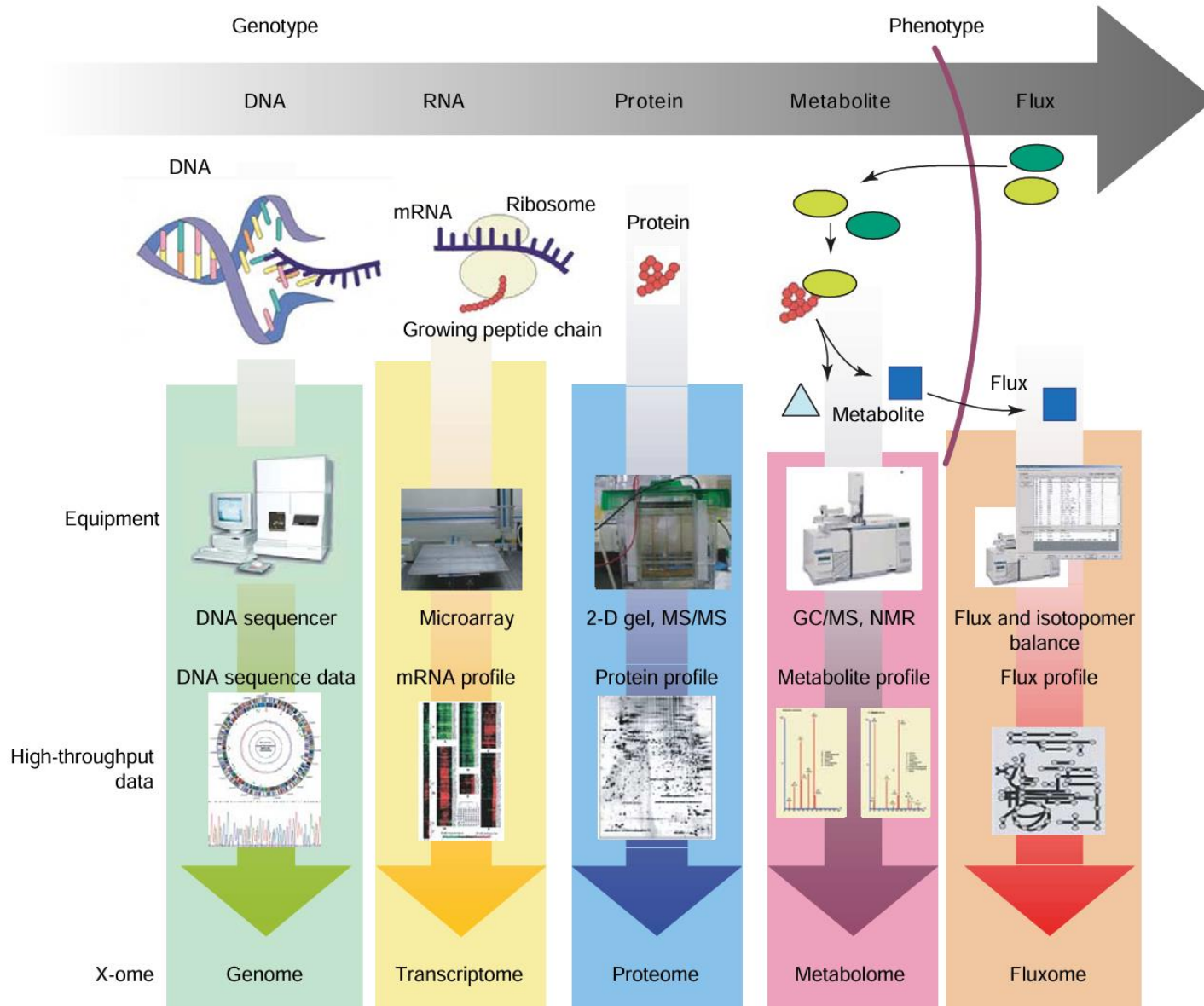
# From central dogma to omics



# From central dogma to omics



# From central dogma to omics



# From omics to GEMs

Genome

Transcriptome

Proteome

Metabolome

Fluxome

# From omics to GEMs

Genome

Transcriptome

Proteome

Metabolome

Fluxome



**Build context specific models**

**&**

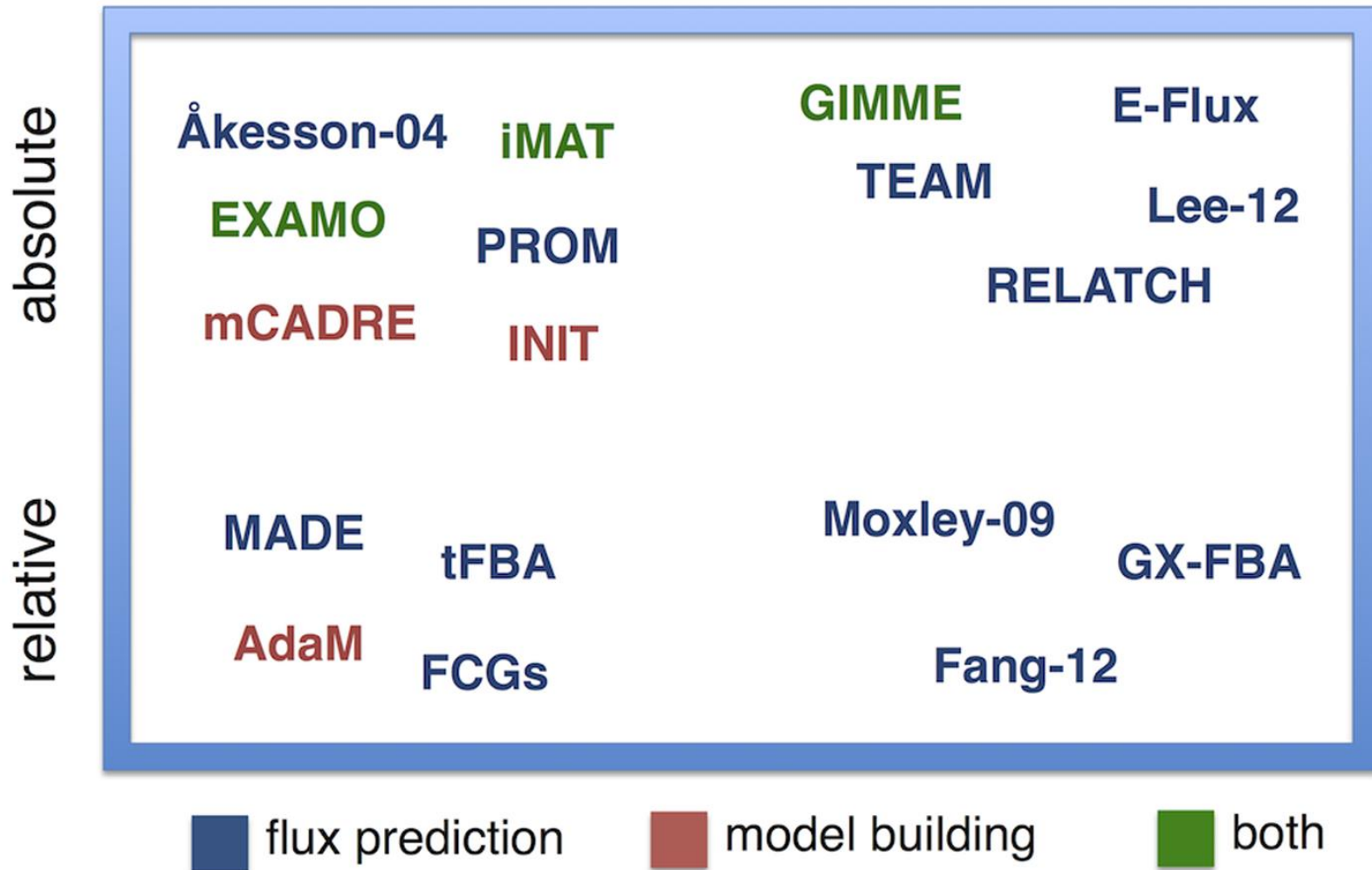
**Improve flux prediction**

# Data integration into GEMs

discrete

vs

continuous

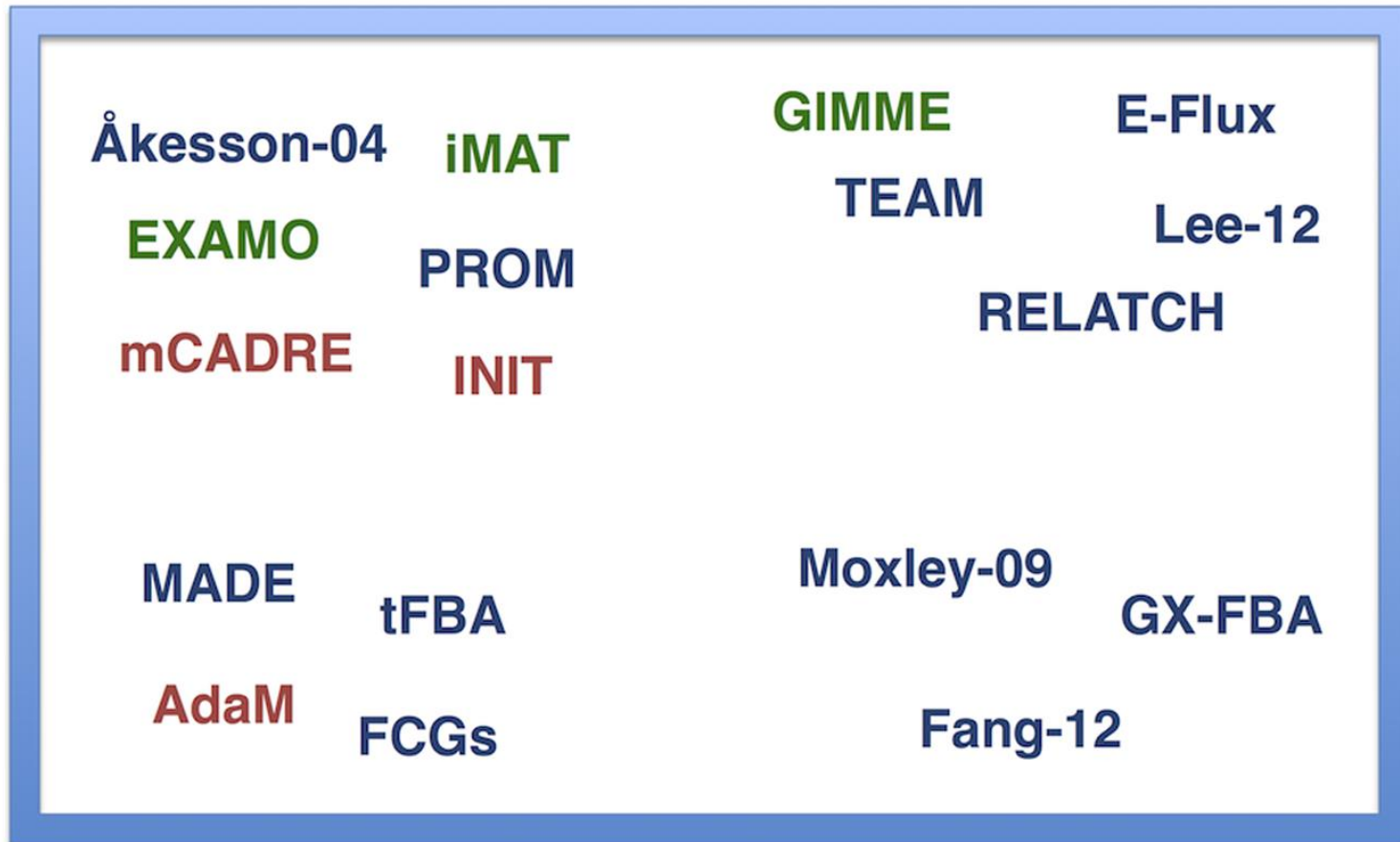


# Data integration into GEMs

discrete

continuous

absolute  
vs  
relative

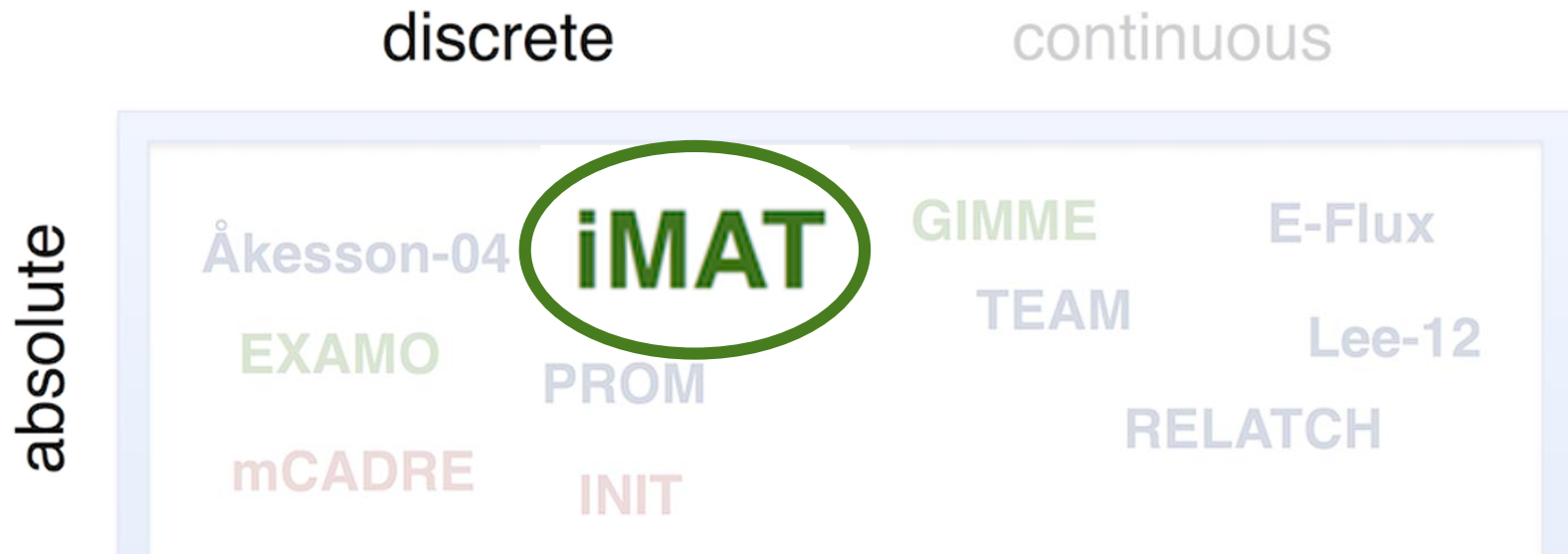


■ flux prediction

■ model building

■ both

# iMAT method



Network-based prediction of human tissue-specific metabolism

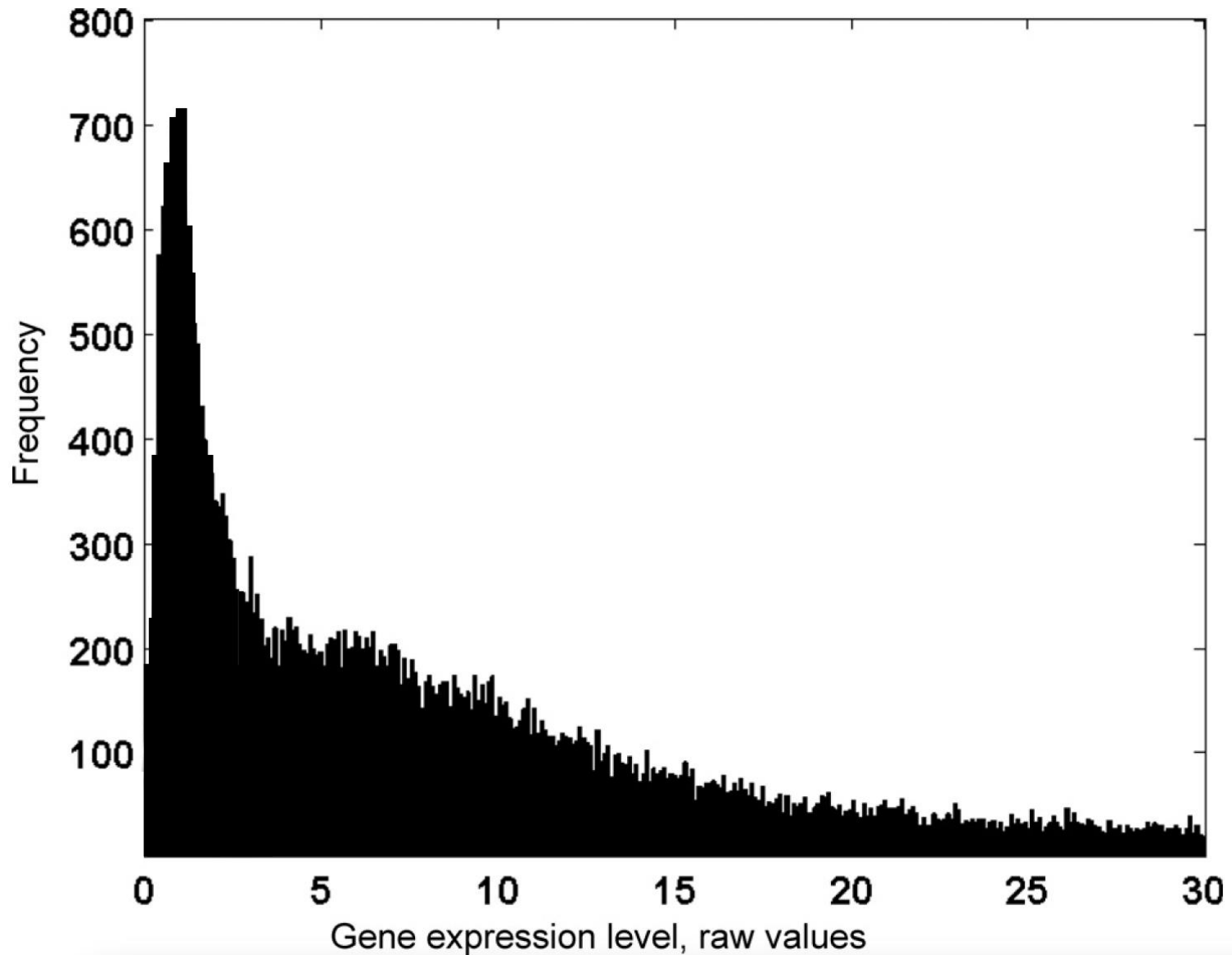
Tomer Shlomi<sup>1,4</sup>, Moran N Cabili<sup>1,4</sup>, Markus J Herrgård<sup>2</sup>, Bernhard Ø Palsson<sup>2</sup> & Eytan Ruppin<sup>1,3</sup>

flux prediction

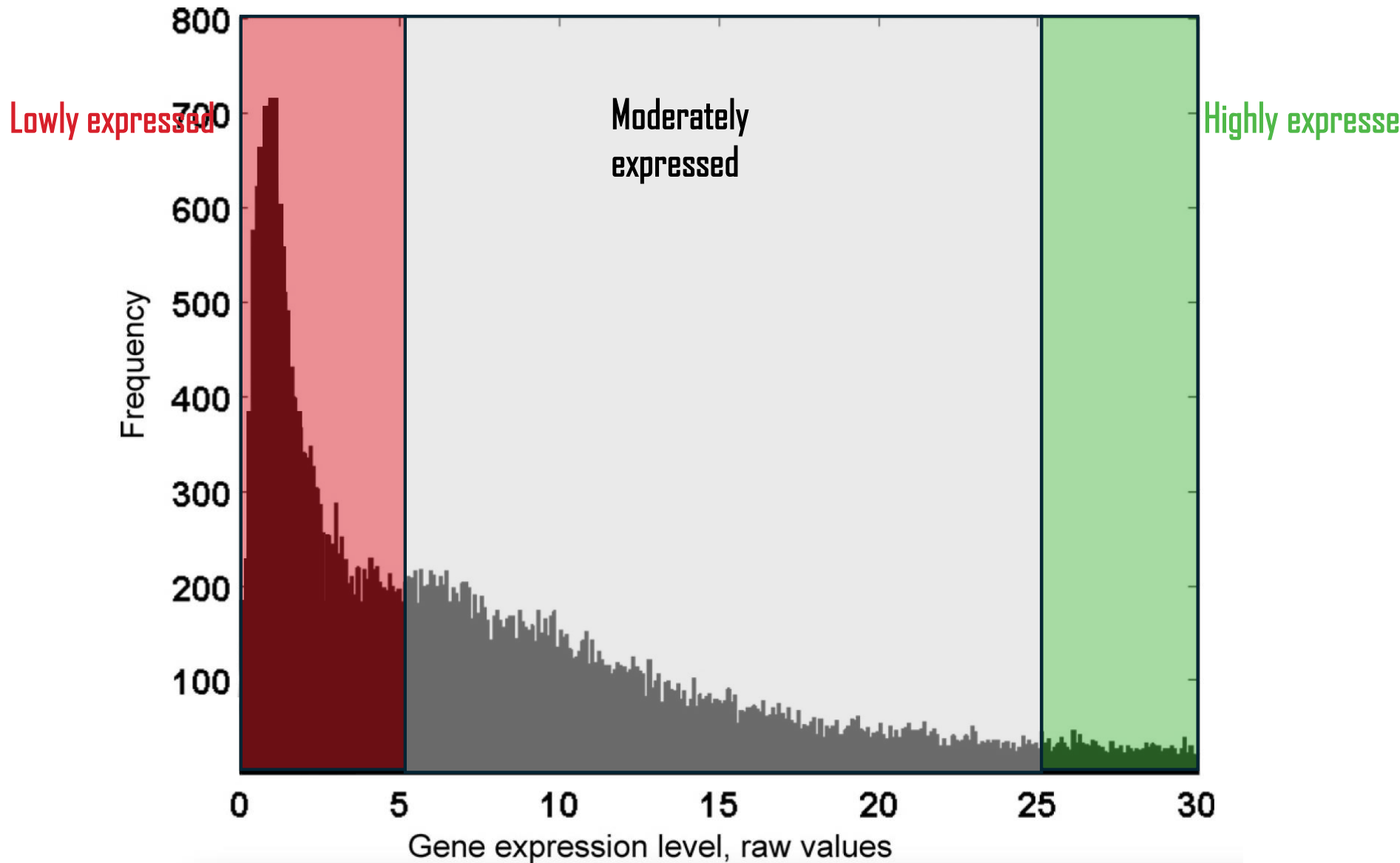
model building

both

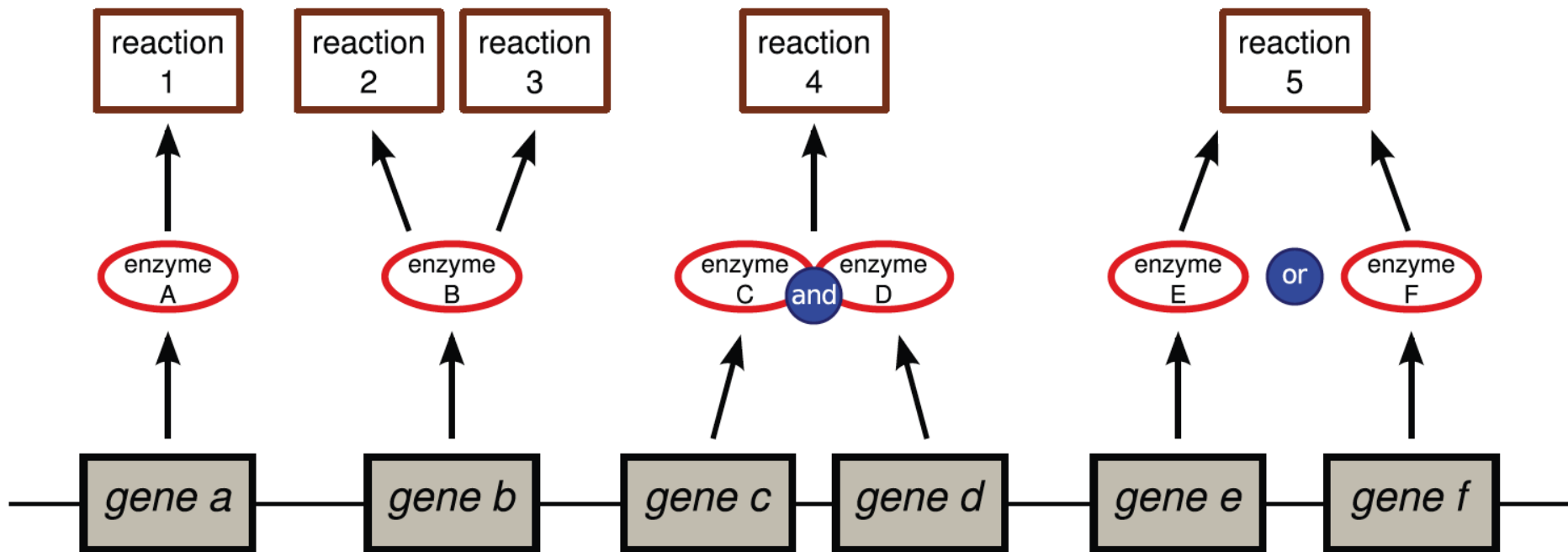
# Gene expression level distribution



# Gene expression level distribution



# Gene-Protein-Reaction (GPR) rules

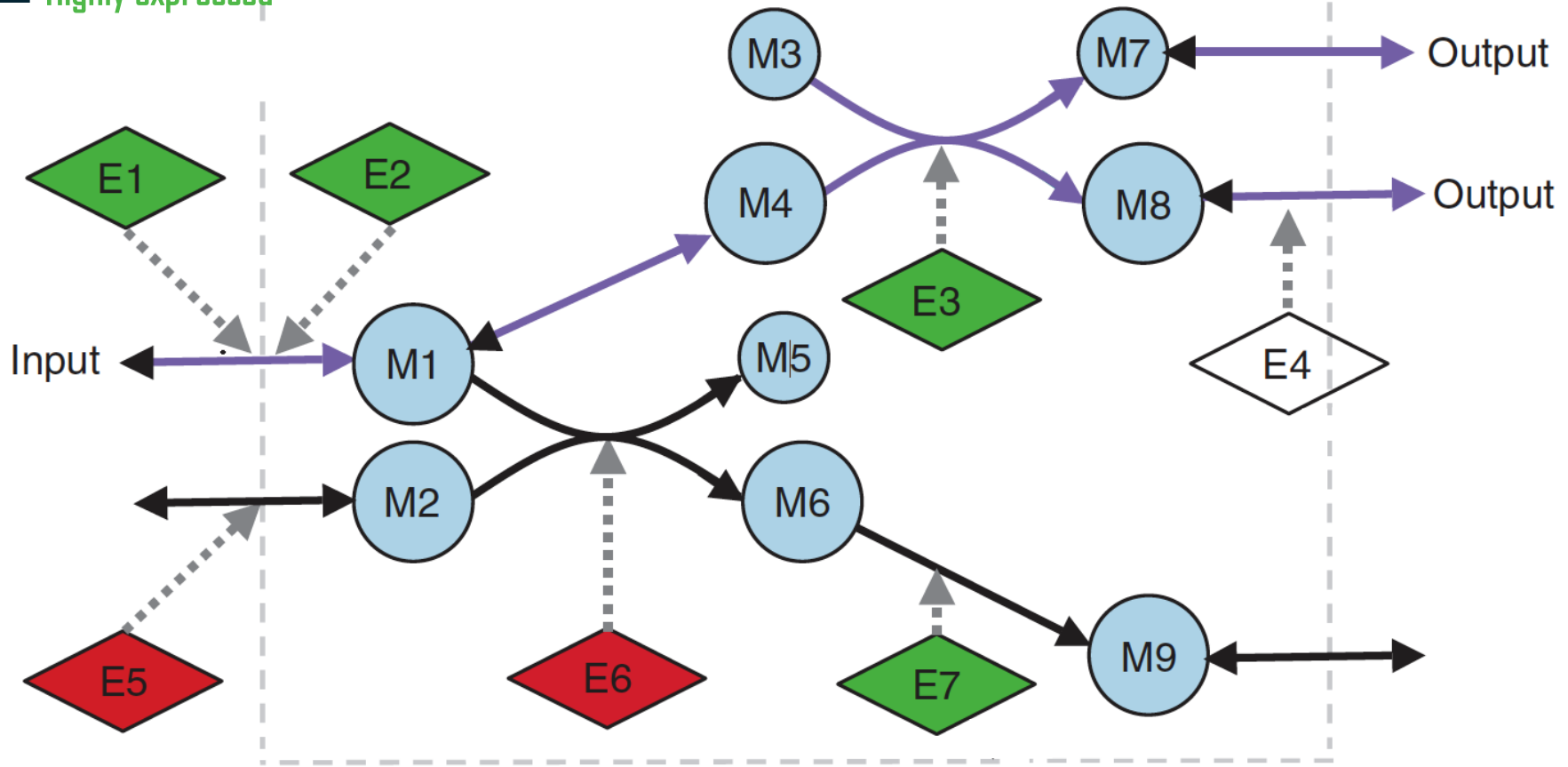


# iMAT method (example)

□ Moderately expressed

■ Lowly expressed

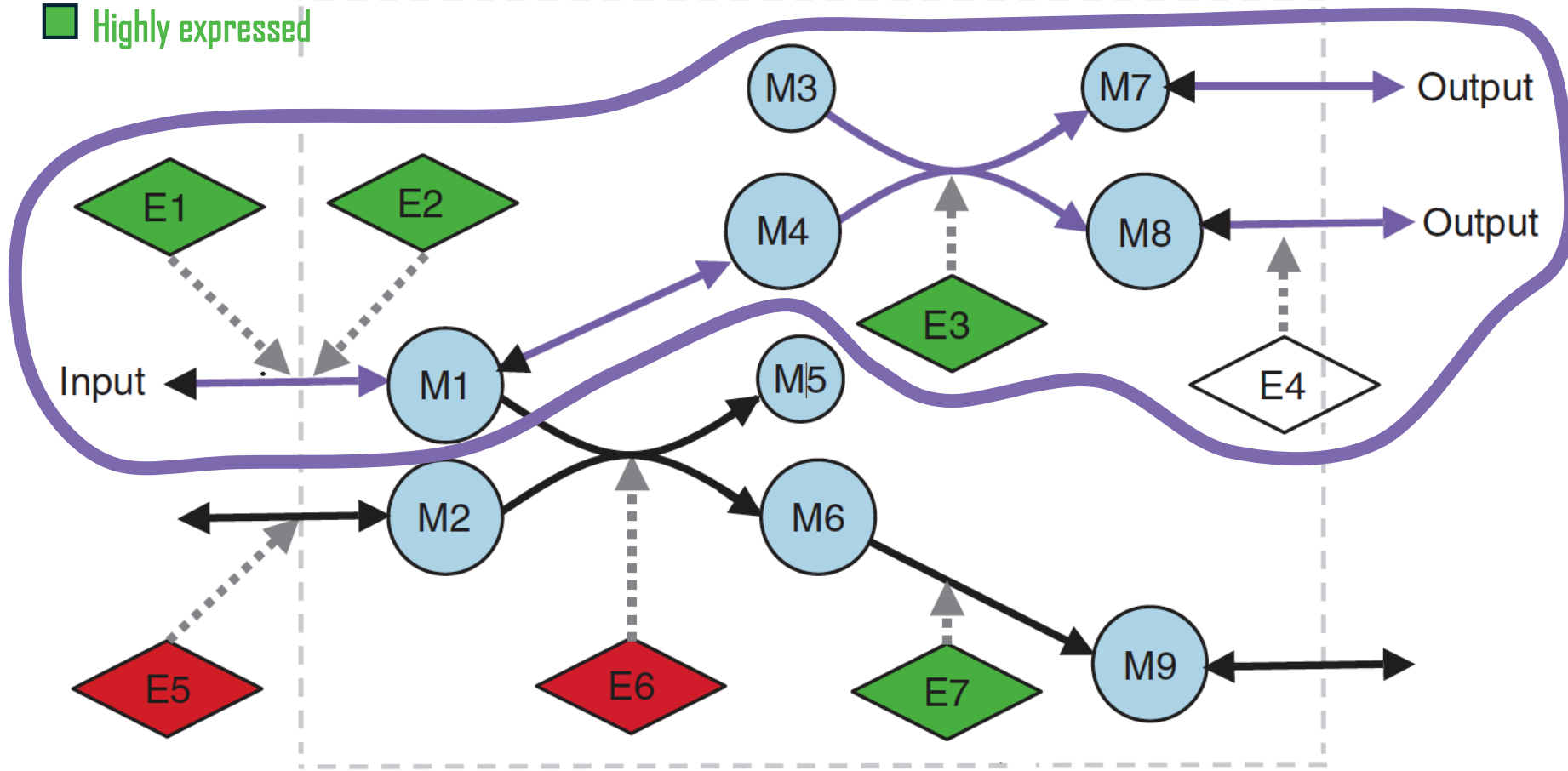
■ Highly expressed



# iMAT method (example)

- Moderately expressed
- Lowly expressed
- Highly expressed

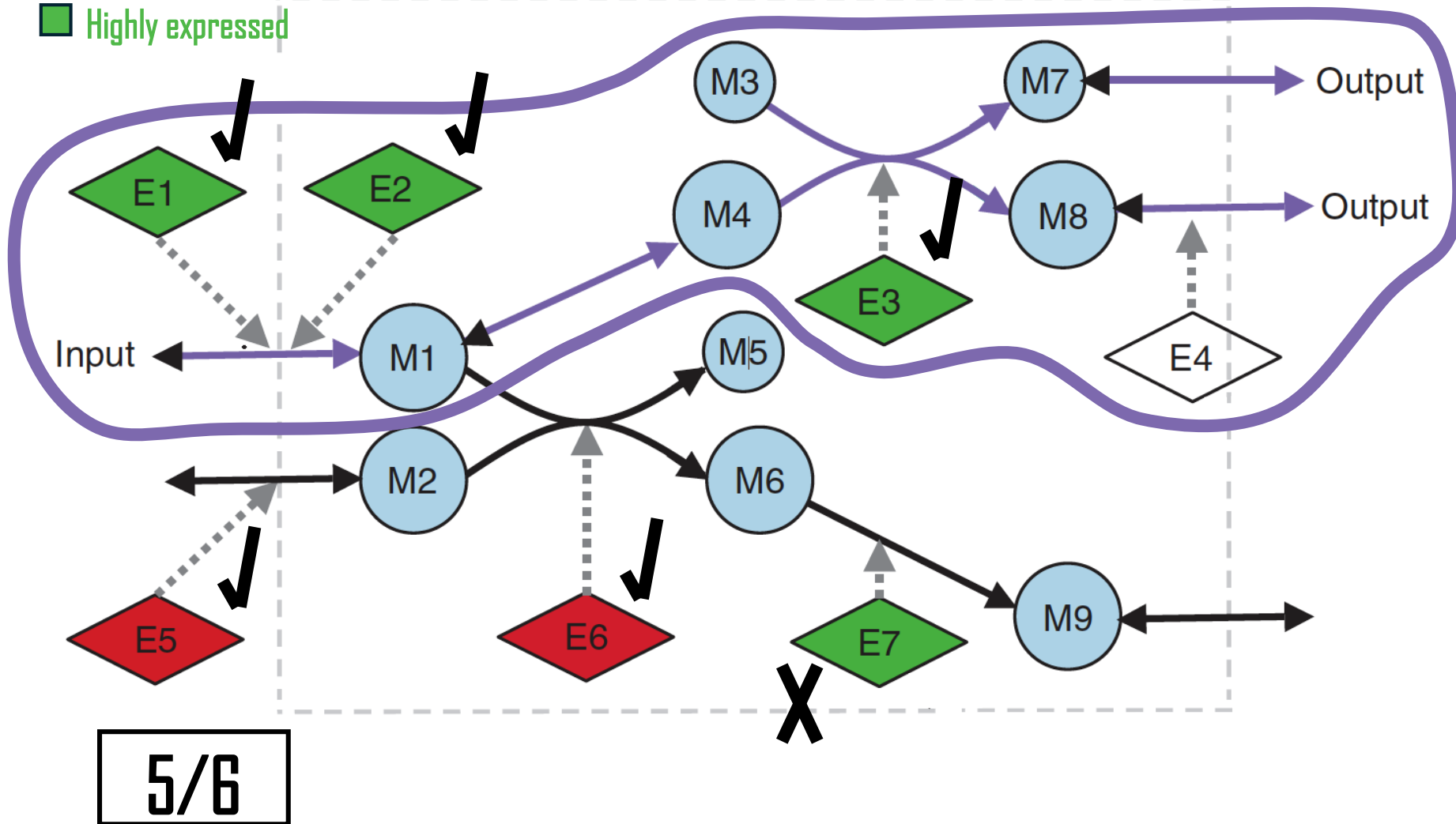
Maximize data integration score  
(E1, E2, E3, E5, E6, E7)



# iMAT method (example)

- Moderately expressed
- Lowly expressed
- Highly expressed

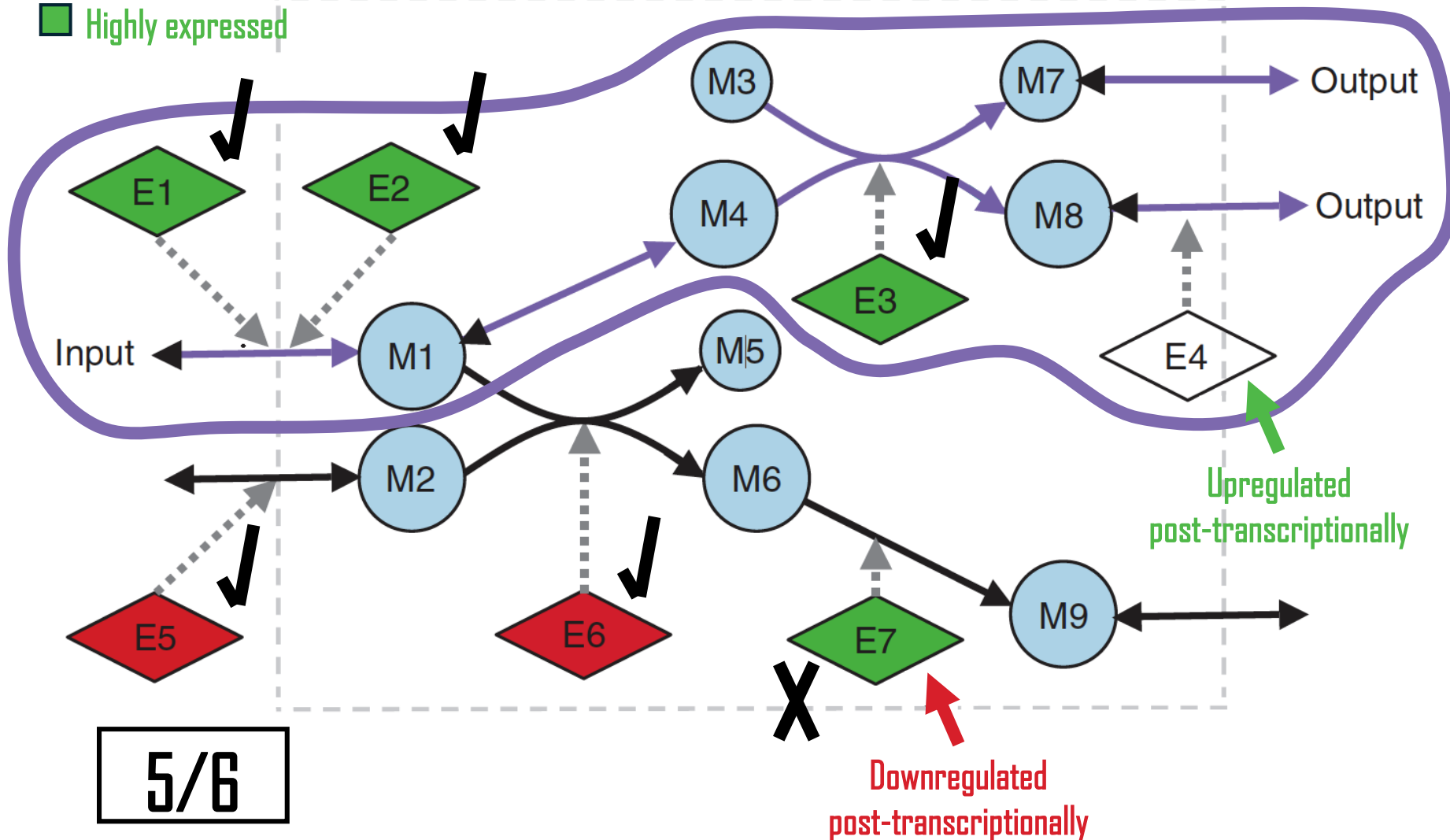
Maximize data integration score  
(E1, E2, E3, E5, E6, E7)



# iMAT method (example)

- Moderately expressed
- Lowly expressed
- Highly expressed

Maximize data integration score  
(E1, E2, E3, E5, E6, E7)



# iMAT method

## MILP formulation

$$\max_{v, y^+, y^-} \left( \sum_{i \in R_H} (y_i^+ + y_i^-) + \sum_{i \in R_L} y_i^+ \right)$$

s.t

$$S \cdot v = 0$$

$$v_{\min} \leq v \leq v_{\max}$$

$$v_i + y_i^+ (v_{\min, i} - \varepsilon) \geq v_{\min, i}, i \in R_H$$

$$v_i + y_i^- (v_{\max, i} + \varepsilon) \leq v_{\max, i}, i \in R_H$$

$$v_{\min, i} (1 - y_i^+) \leq v_i \leq v_{\max, i} (1 - y_i^+), i \in R_L$$

$$v \in R^m$$

$$y_i^+, y_i^- \in [0, 1]$$

$R_H$  : highly expressed reactions

$R_L$  : lowly expressed reactions

$y_i^+$  : reaction is active

$y_i^-$  : reaction is not active

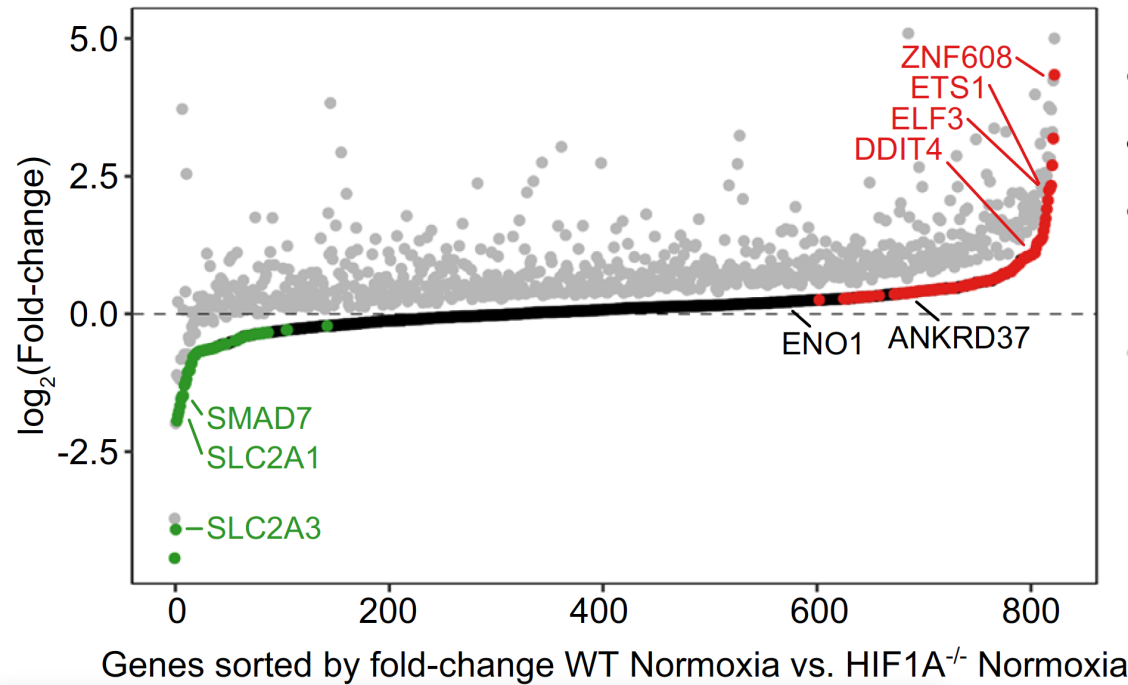
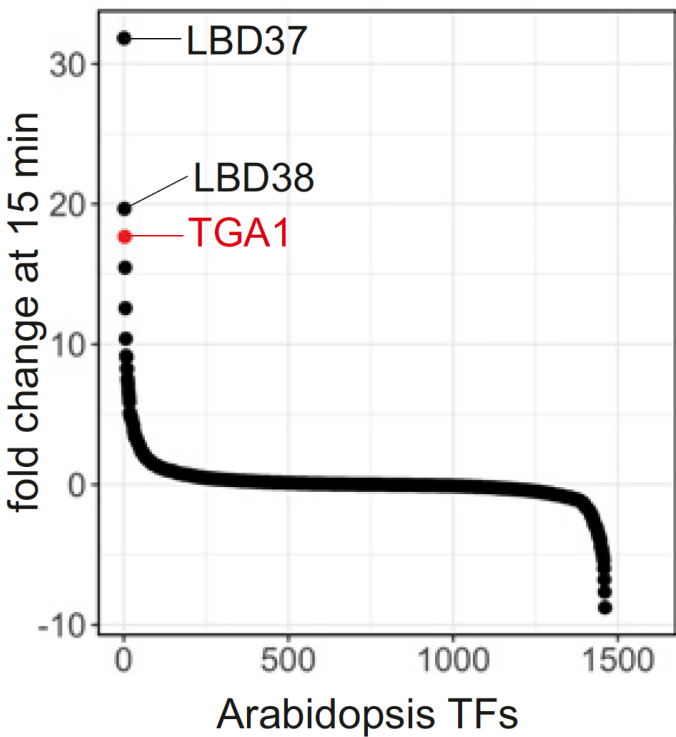
RESEARCH ARTICLE

## Enhanced flux prediction by integrating relative expression and relative metabolite abundance into thermodynamically consistent metabolic models

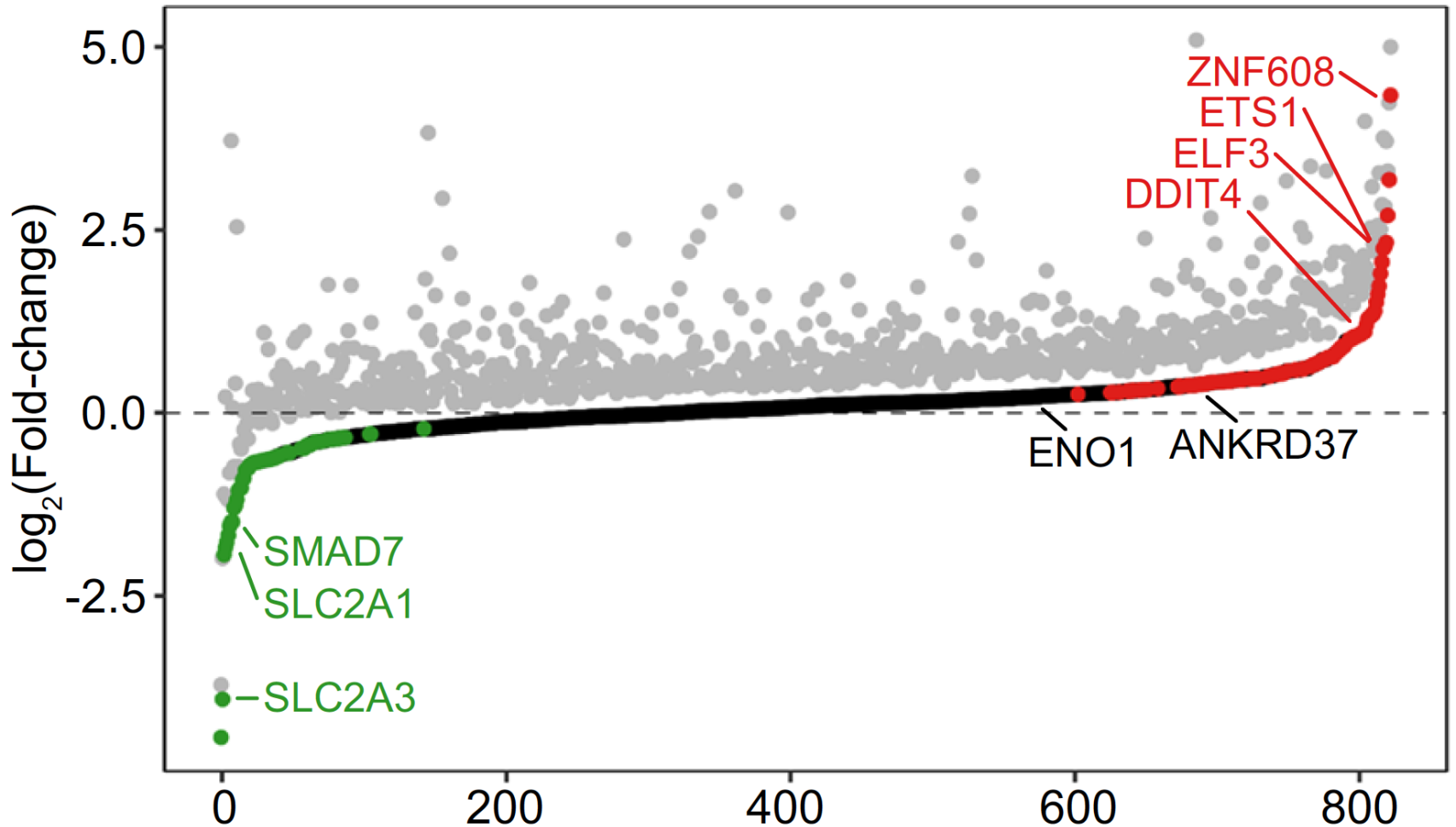
**Vikash Pandey, Noushin Hadadi , Vassily Hatzimanikatis \***

Laboratory of Computational Systems Biotechnology, EPFL, Lausanne, Switzerland

# Expression deregulation distribution

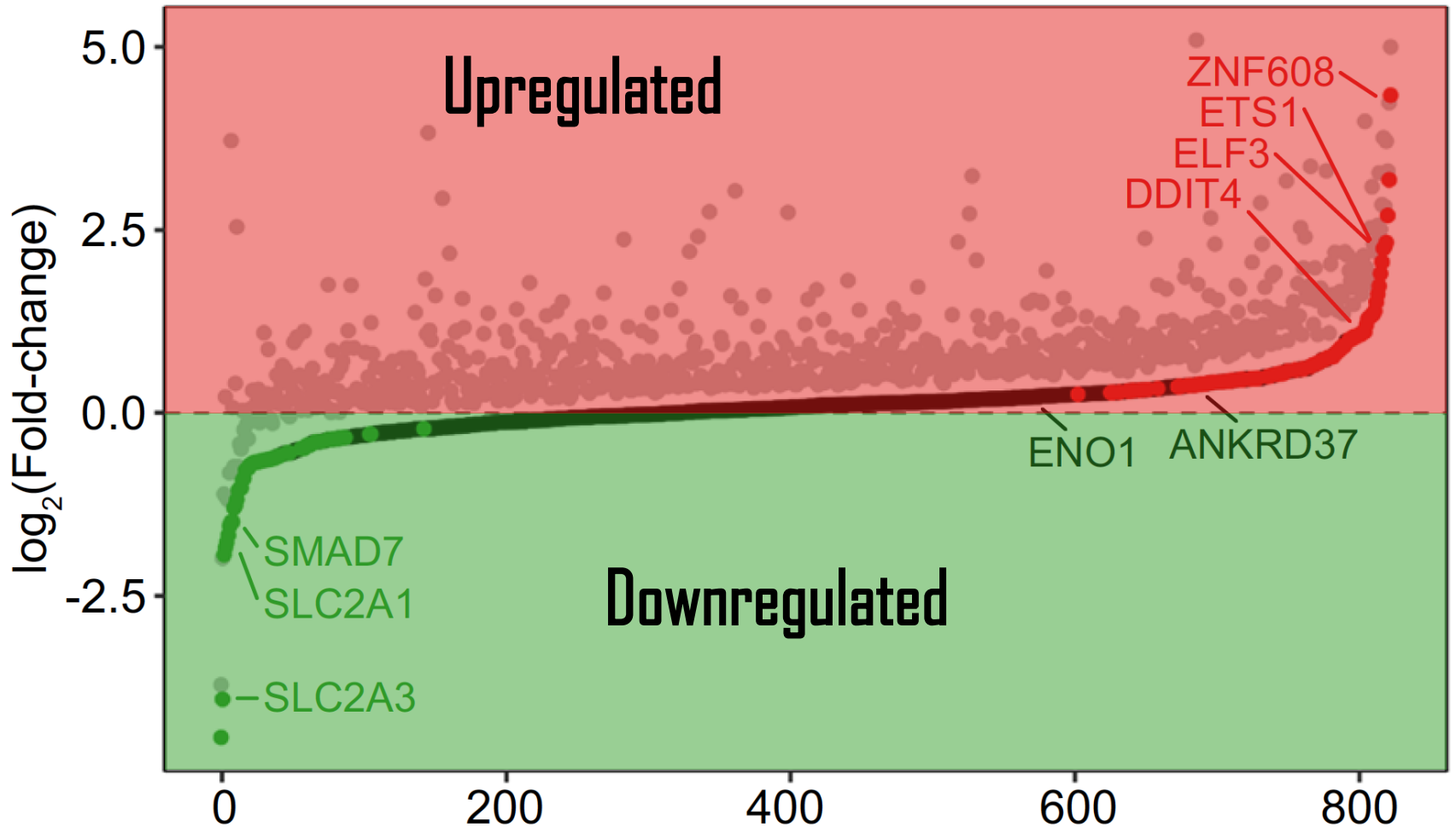


# Expression deregulation distribution



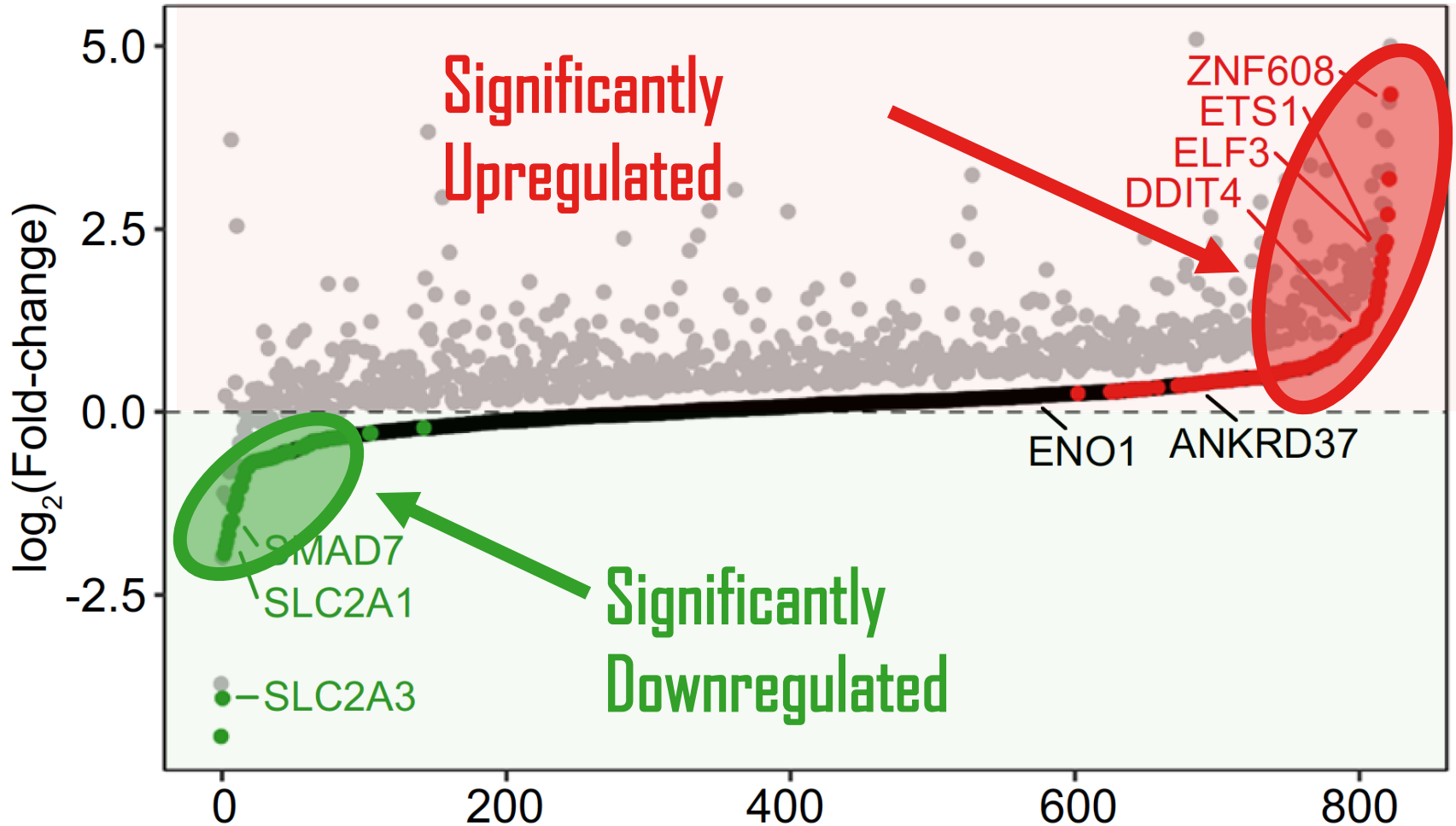
Genes sorted by fold-change WT Normoxia vs. HIF1A<sup>-/-</sup> Normoxia

# Expression deregulation distribution



Genes sorted by fold-change WT Normoxia vs. HIF1A<sup>-/-</sup> Normoxia

# Expression deregulation distribution



Genes sorted by fold-change WT Normoxia vs. HIF1A<sup>-/-</sup> Normoxia

# REMI workflow for expression data integration

## Inputs

FBA / TFA  
model

+

①



Transcriptomic data for  
*reference* condition

②

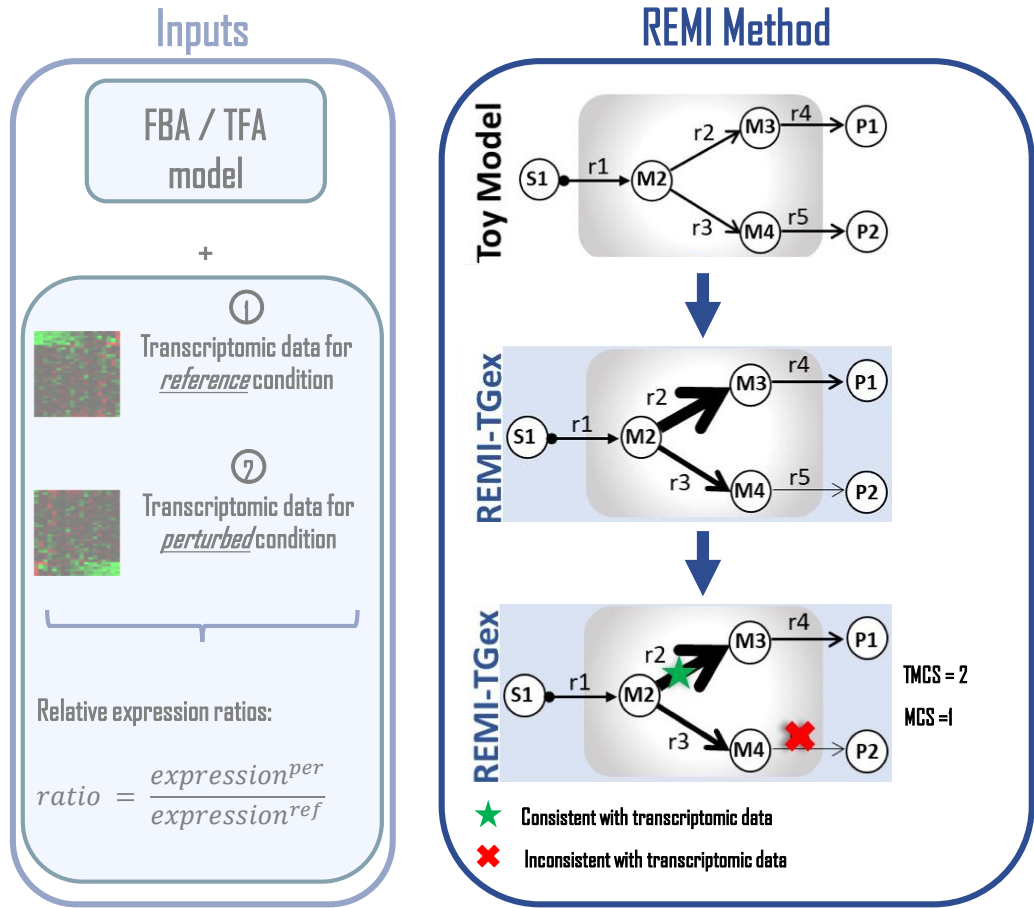


Transcriptomic data for  
*perturbed* condition

Relative expression ratios:

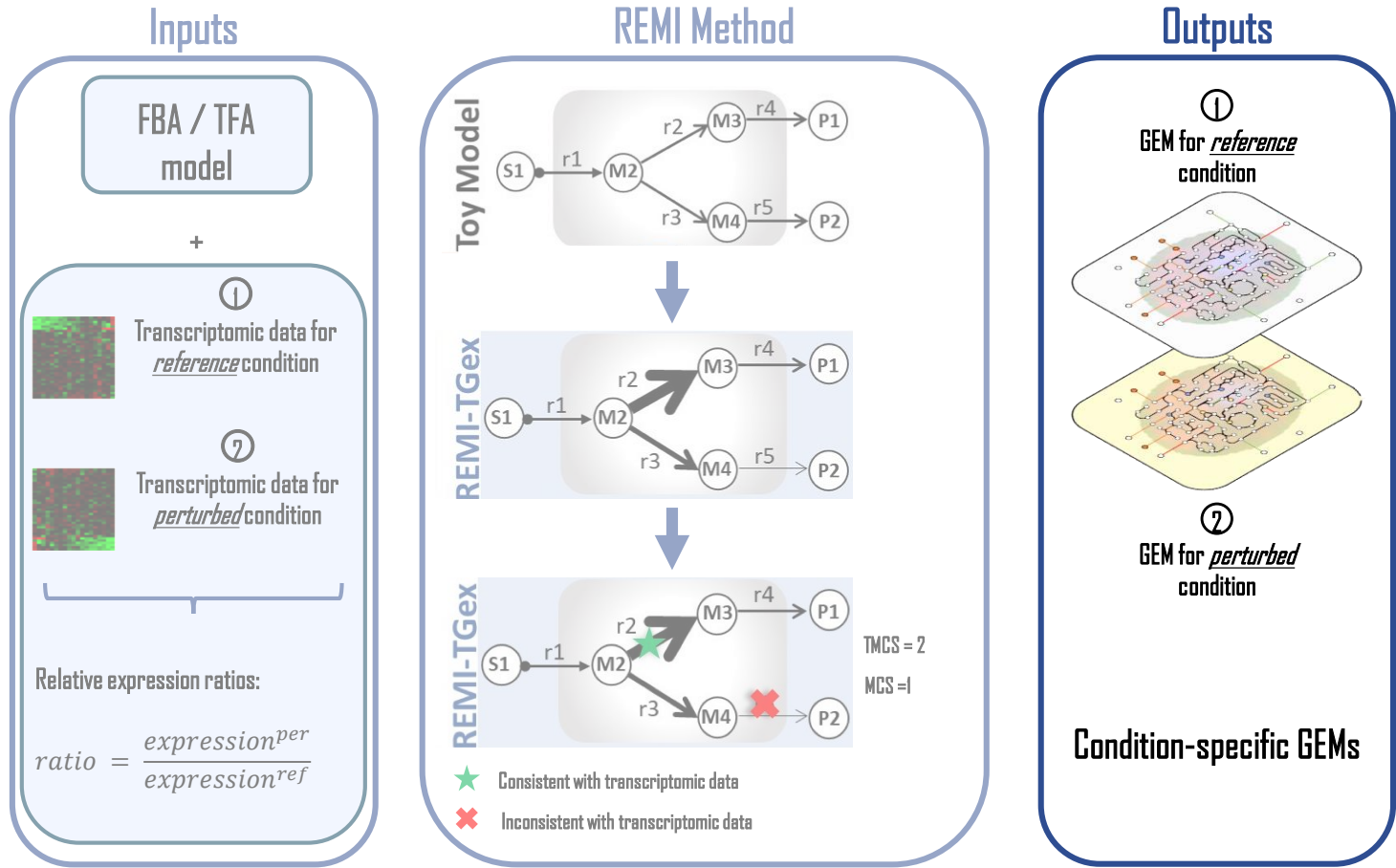
$$\text{ratio} = \frac{\text{expression}^{per}}{\text{expression}^{ref}}$$

# REMI workflow for expression data integration



\* TMCS = Theoretical Maximum Consistency Score  
MCS = Maximum Consistency Score

# REMI workflow for expression data integration



# REMI workflow for expression data integration

### Inputs

FBA / TFA model

+

① Transcriptomic data for *reference* condition

② Transcriptomic data for *perturbed* condition

Relative expression ratios:

$$ratio = \frac{expression^{per}}{expression^{ref}}$$

### REMI Method

**Toy Model**

**REMI-TGex**

**REMI-TGex**

TMGS = 2  
MGS = 1

- ★ Consistent with transcriptomic data
- ✗ Inconsistent with transcriptomic data

### Outputs

① GEM for *reference* condition

② GEM for *perturbed* condition

Condition-specific GEMs

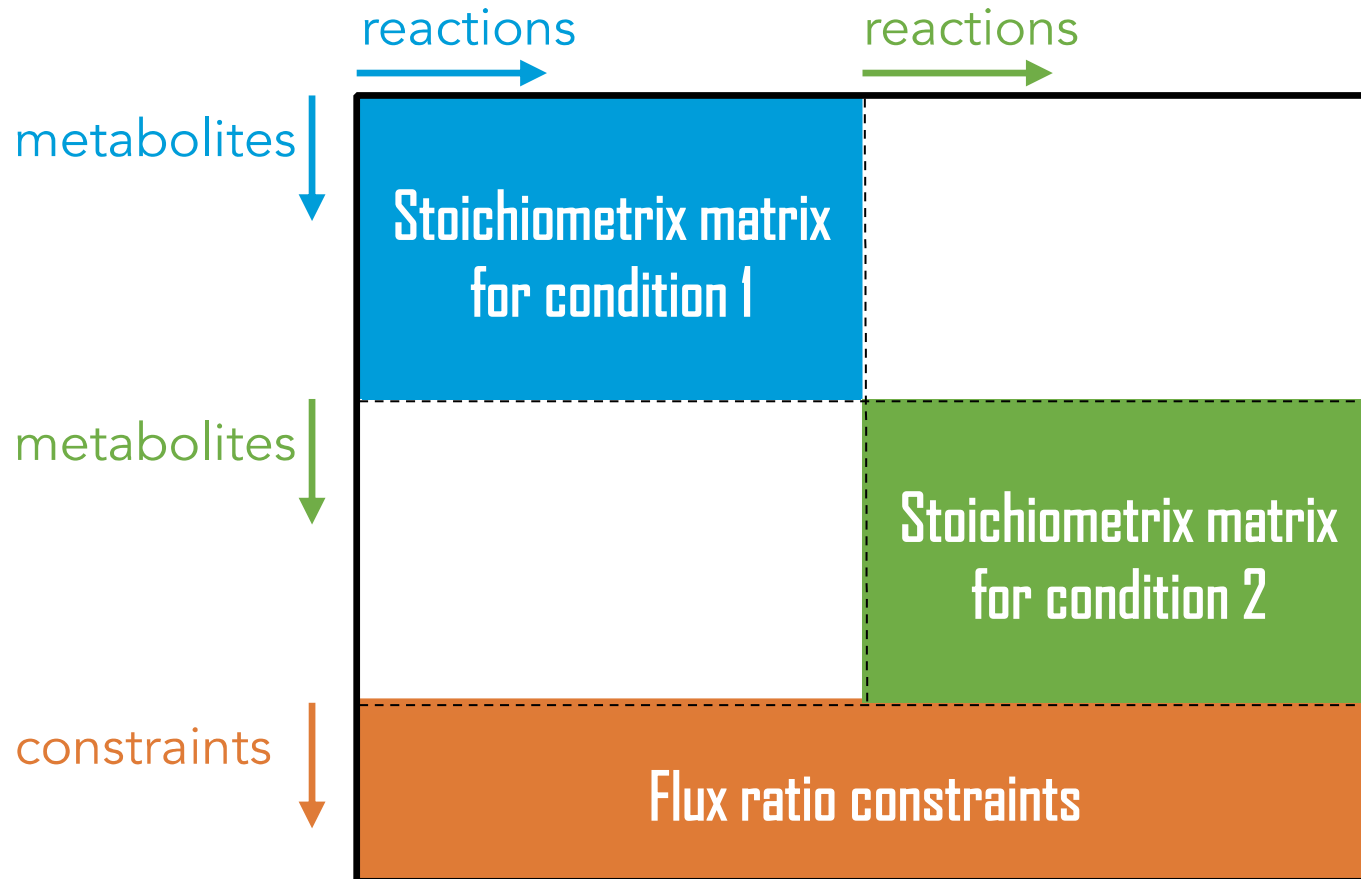
### Analysis

Fluxes deregulation

Gene essentiality

# REMI method

## Master S matrix



# REMI method

## MILP formulation

$$\text{Consistency Score} = \text{Maximize} \sum_{i=1}^n B_i$$

$$n = u + a$$

$$v_i^{\text{mutant}} \geq \varepsilon * B_i \quad [i = 1, \dots, n \text{ for both upregulated and downregulated reactions}]$$

$$v_i^{\text{wild}} \geq \varepsilon * B_i \quad [i = 1, \dots, n \text{ for both upregulated and downregulated reactions}]$$

$$v_i^{\text{mutant}} \geq p * v_r^{\text{wild}} - \sigma_i \quad [i = 1, \dots, u \text{ for upregulated reactions}]$$

$$v_i^{\text{mutant}} \leq p * v_r^{\text{wild}} + \sigma_i \quad [i = 1, \dots, d \text{ for downregulated reactions}]$$

$$\varepsilon * (1 - B_i) \leq \sigma_i \leq \varepsilon + (1 - B_i) * M' \quad i = 1, \dots, n$$

# REMI method

## MILP formulation

$$\text{Consistency Score} = \text{Maximize } \sum_{i=1}^n B_i$$

$$n = u + a$$

$$v_i^{\text{mutant}} \geq \varepsilon * B_i \quad [i = 1, \dots, n \text{ for both upregulated and downregulated reactions}]$$

$$v_i^{\text{wild}} \geq \varepsilon * B_i \quad [i = 1, \dots, n \text{ for both upregulated and downregulated reactions}]$$

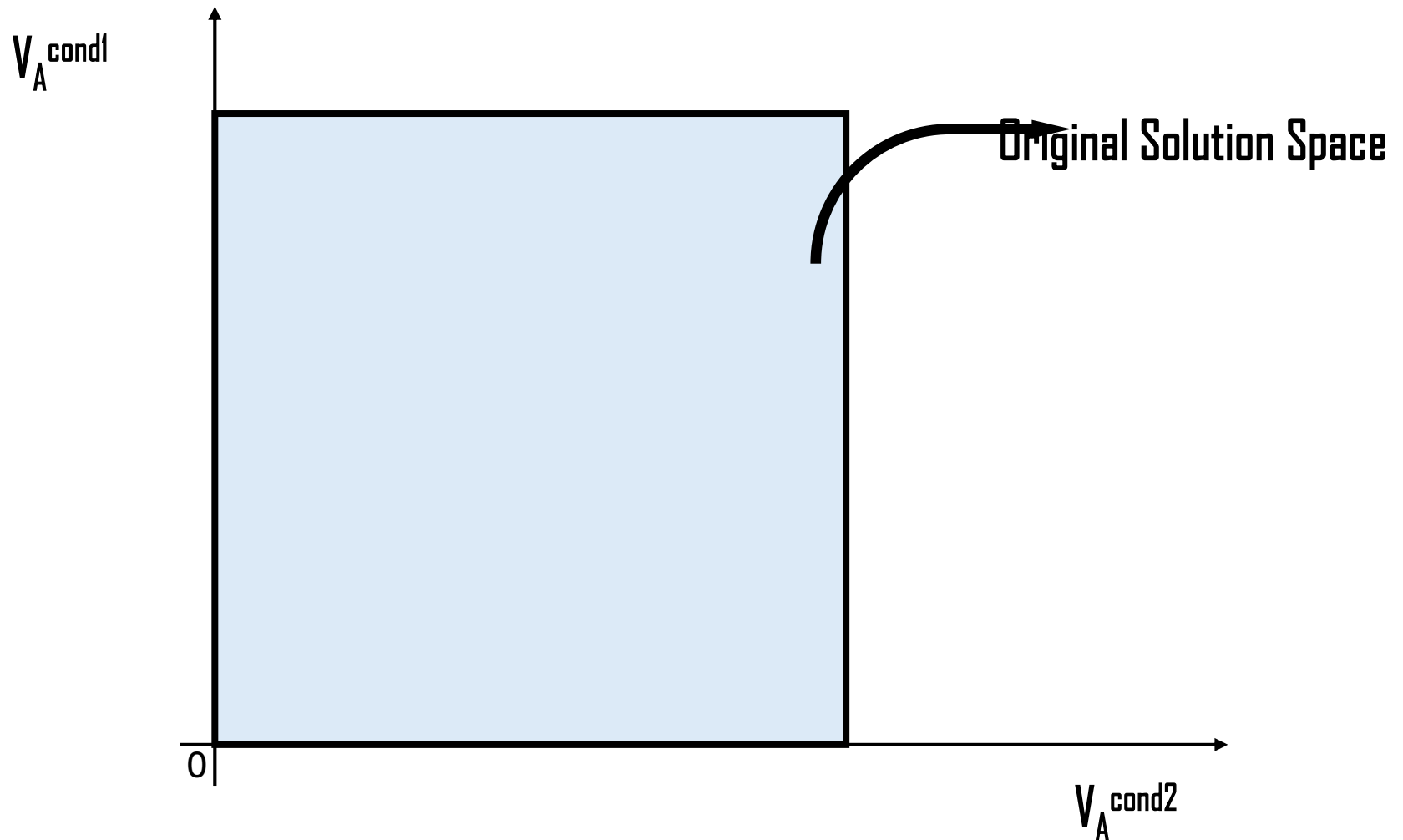
### Fold change ratios

$$v_i^{\text{mutant}} \geq p * v_r^{\text{wild}} - \sigma_i \quad [i = 1, \dots, u \text{ for upregulated reactions}]$$

$$v_i^{\text{mutant}} \leq p * v_r^{\text{wild}} + \sigma_i \quad [i = 1, \dots, d \text{ for downregulated reactions}]$$

$$\varepsilon * (1 - B_i) \leq \sigma_i \leq \varepsilon + (1 - B_i) * M' \quad i = 1, \dots, n$$

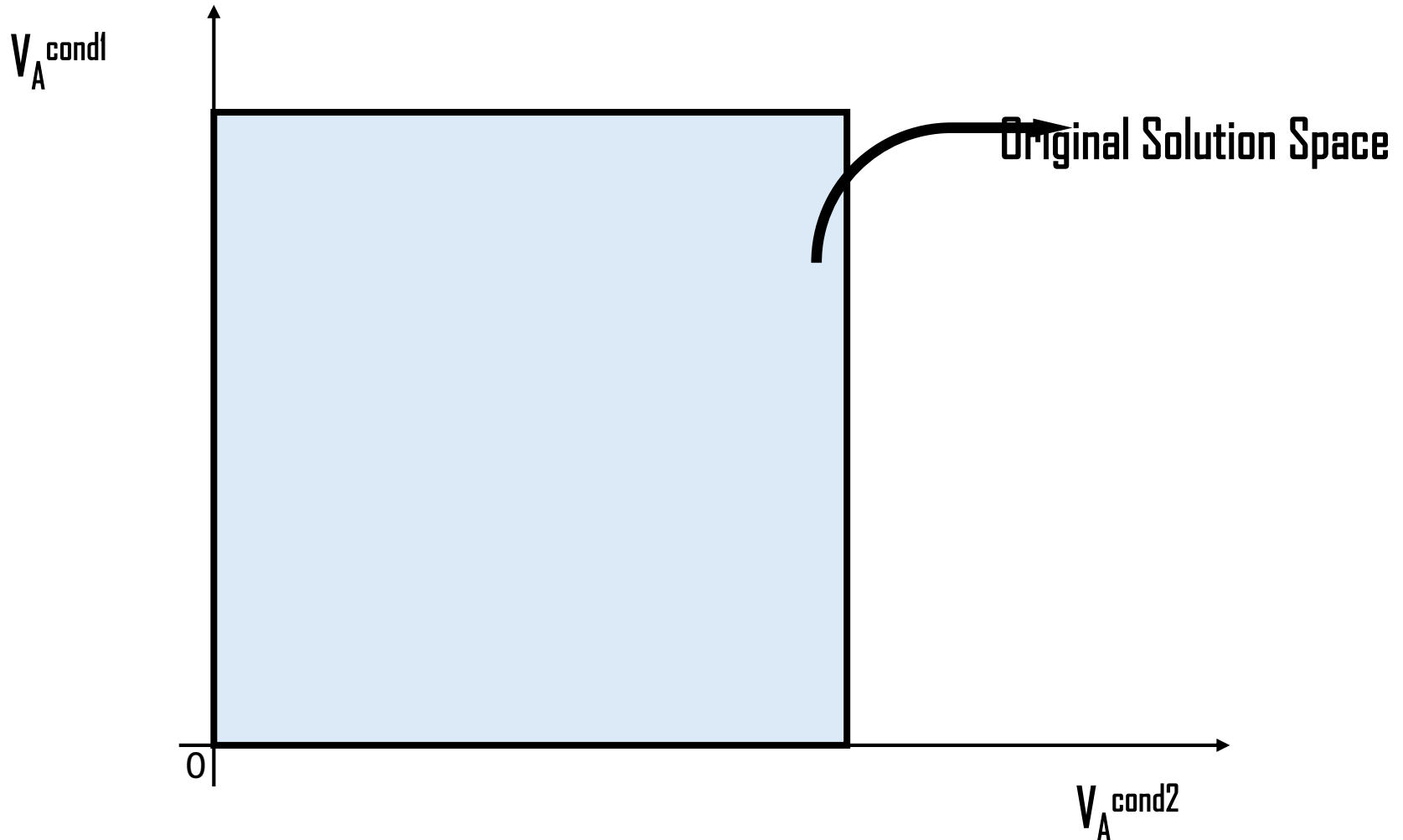
# REMI method (example)



# REMI method (example)

Flux ratio constraint:

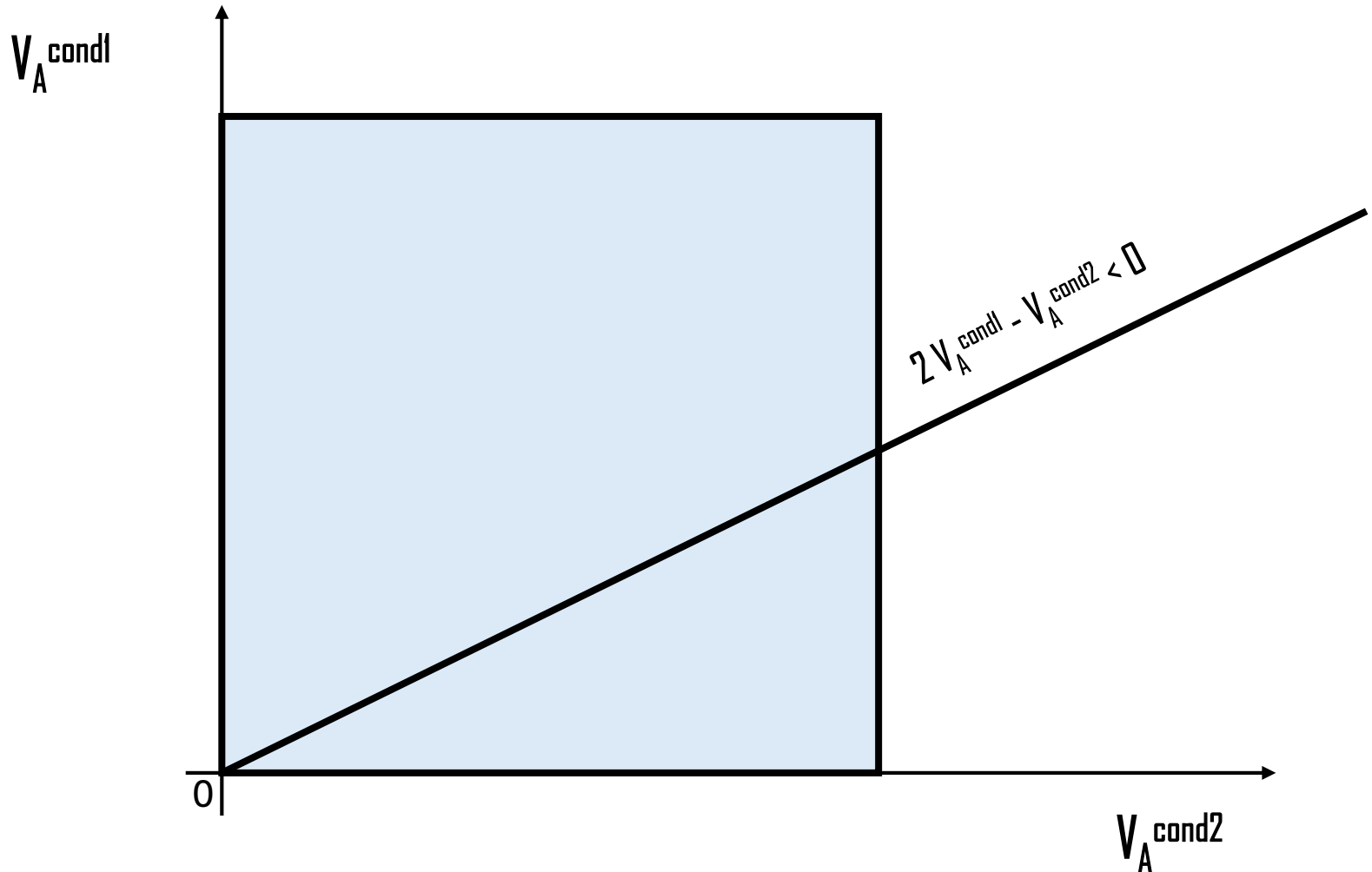
$$V_{A \text{ cond}2} > 2 V_{A \text{ cond}1} \Rightarrow 2 V_{A \text{ cond}1} - V_{A \text{ cond}2} < 0$$



# REMI method (example)

Flux ratio constraint:

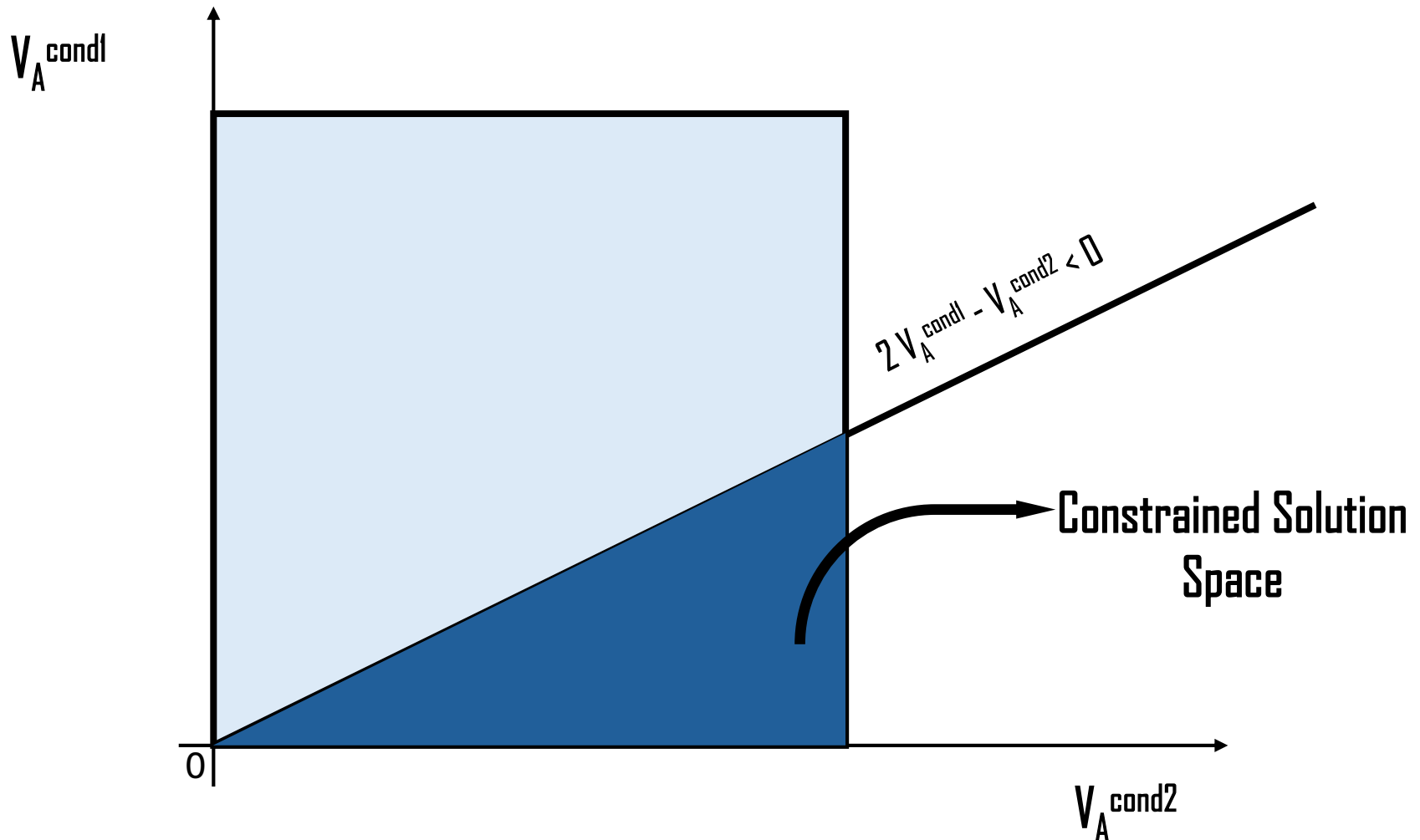
$$V_{A \text{ cond}2} > 2 V_{A \text{ cond}1} \Rightarrow 2 V_{A \text{ cond}1} - V_{A \text{ cond}2} < 0$$



# REMI method (example)

Flux ratio constraint:

$$V_{A \text{ cond}2} > 2 V_{A \text{ cond}1} \Rightarrow 2 V_{A \text{ cond}1} - V_{A \text{ cond}2} < 0$$



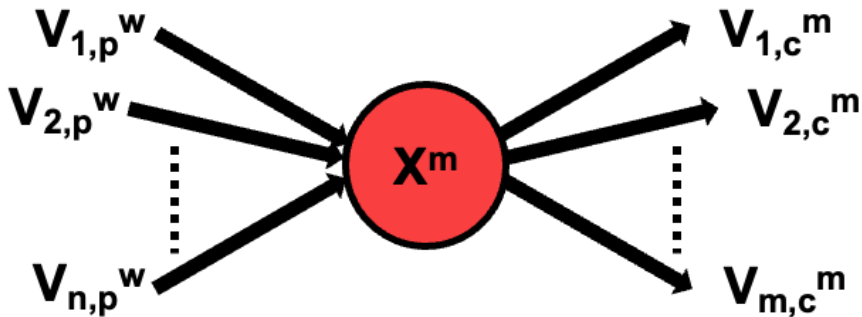
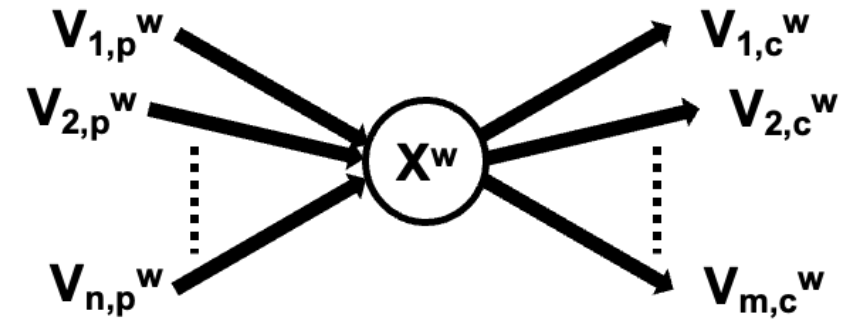
# REMI method – Metabolomics

From metabolomics to fluxes

# REMI method - Metabolomics

From metabolomics to fluxes

- Upregulated metabolite



$$\sum_{i=1}^n V_{i,p}^m \geq \frac{X^m}{X^w} \sum_{i=1}^n V_{i,p}^w$$

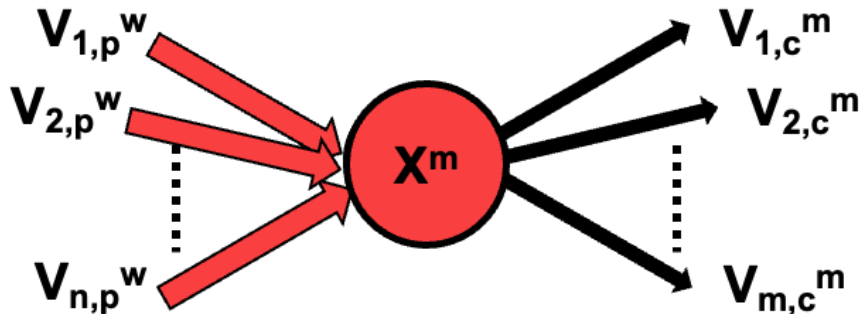
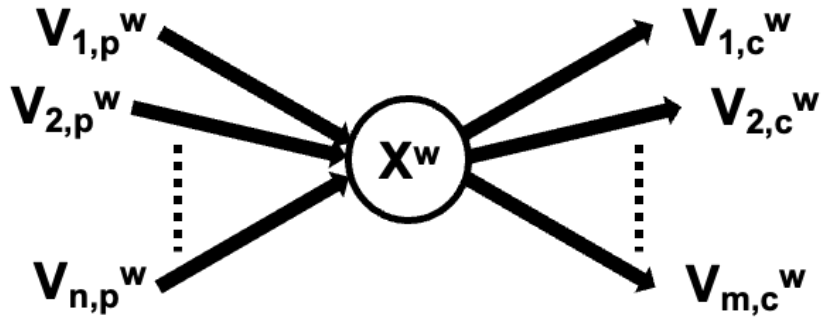
or

$$\sum_{j=1}^m V_{j,c}^m \leq \frac{X^w}{X^m} \sum_{j=1}^m V_{j,c}^w$$

# REMI method – Metabolomics

From metabolomics to fluxes

- Upregulated metabolite (**Increased Production**)



$$\sum_{i=1}^n V_{i,p}^m \geq \frac{X^m}{X^w} \sum_{i=1}^n V_{i,p}^w$$

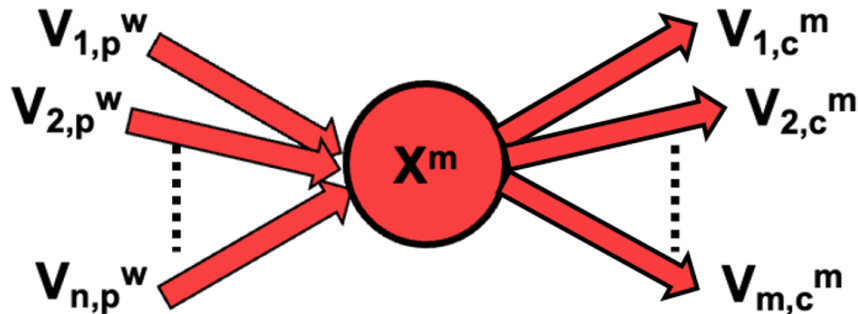
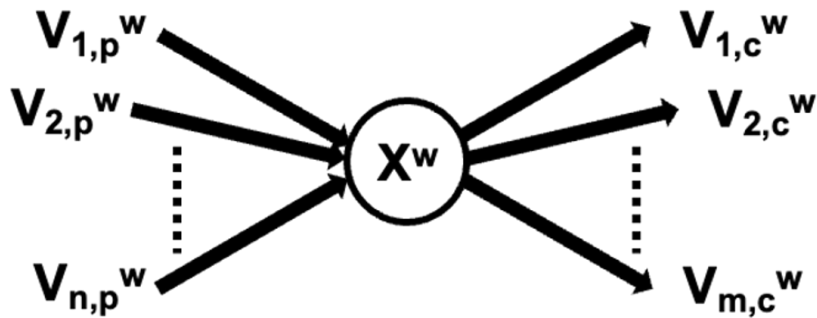
or

$$\sum_{j=1}^m V_{j,c}^m \leq \frac{X^w}{X^m} \sum_{j=1}^m V_{j,c}^w$$

# REMI method – Metabolomics

From metabolomics to fluxes

- Upregulated metabolite (**Increased Production**)



$$\sum_{i=1}^n V_{i,p}^m \geq \frac{X^m}{X^w} \sum_{i=1}^n V_{i,p}^w$$

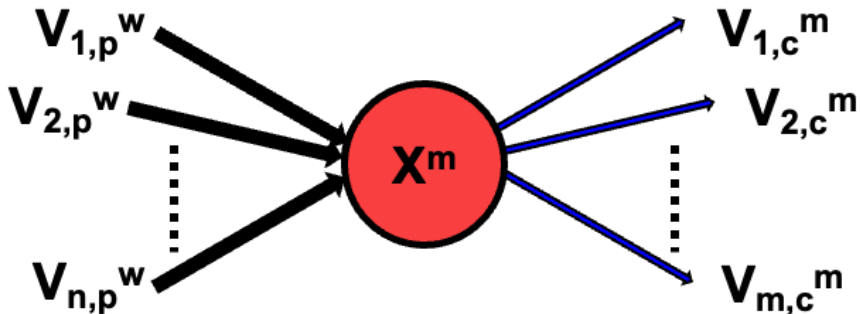
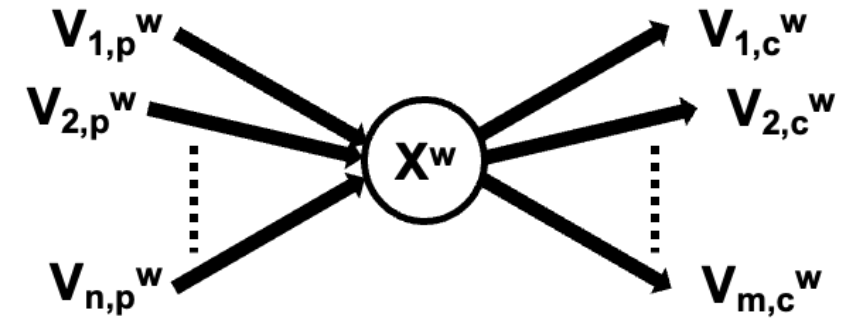
or

$$\sum_{j=1}^m V_{j,c}^m \leq \frac{X^w}{X^m} \sum_{j=1}^m V_{j,c}^w$$

# REMI method - Metabolomics

From metabolomics to fluxes

- Upregulated metabolite (Decreased Consumption)



$$\sum_{i=1}^n V_{i,p}^m \geq \frac{X^m}{X^w} \sum_{i=1}^n V_{i,p}^w$$

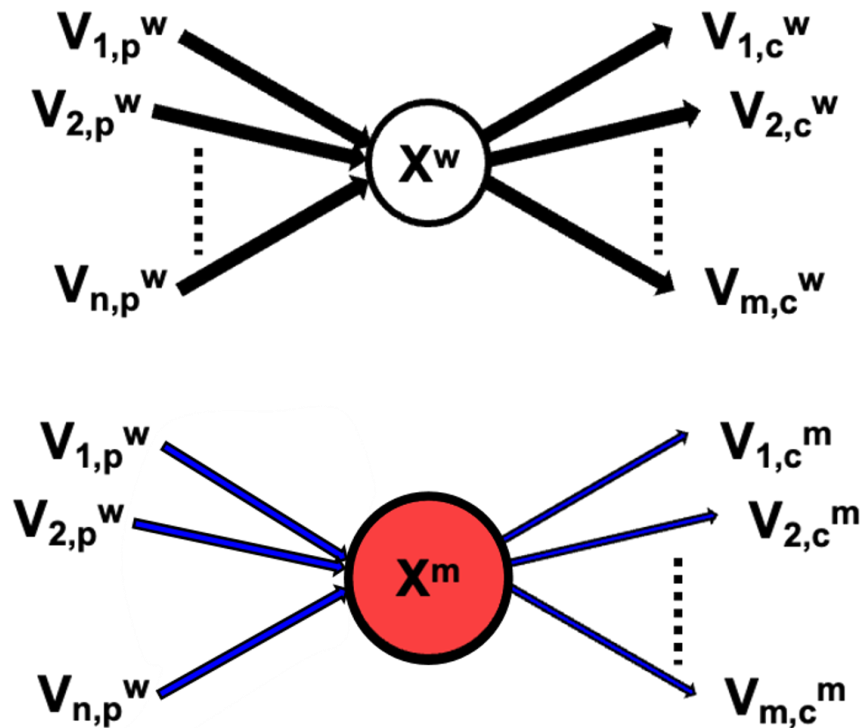
or

$$\sum_{j=1}^m V_{j,c}^m \leq \frac{X^w}{X^m} \sum_{j=1}^m V_{j,c}^w$$

# REMI method – Metabolomics

From metabolomics to fluxes

- Upregulated metabolite (**Decreased Consumption**)



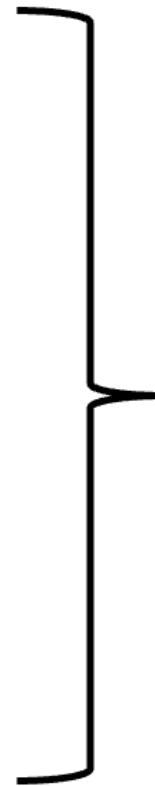
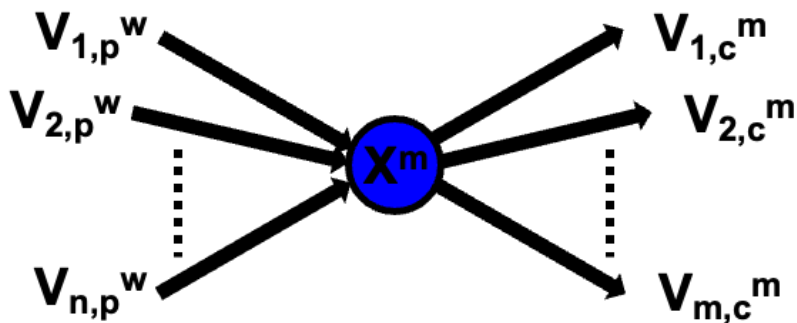
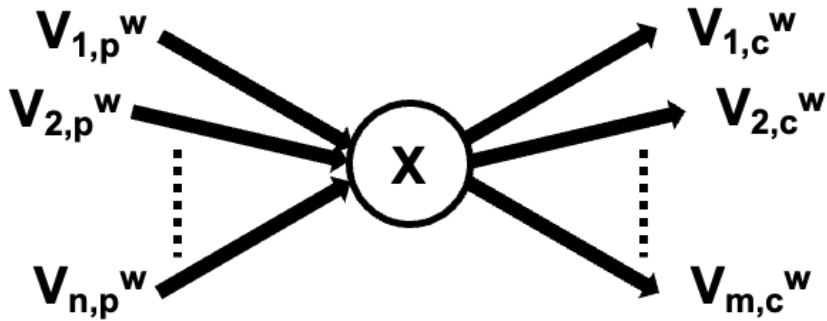
$$\sum_{i=1}^n V_{i,p}^m \geq \frac{X^m}{X^w} \sum_{i=1}^n V_{i,p}^w$$

or

$$\sum_{j=1}^m V_{j,c}^m \leq \frac{X^w}{X^m} \sum_{j=1}^m V_{j,c}^w$$

# REMI method - Metabolomics

From metabolomics to fluxes  
- Downregulated metabolite



$$\sum_{i=1}^n V_{i,p}^m \leq \frac{X^m}{X^w} \sum_{i=1}^n V_{i,p}^w$$

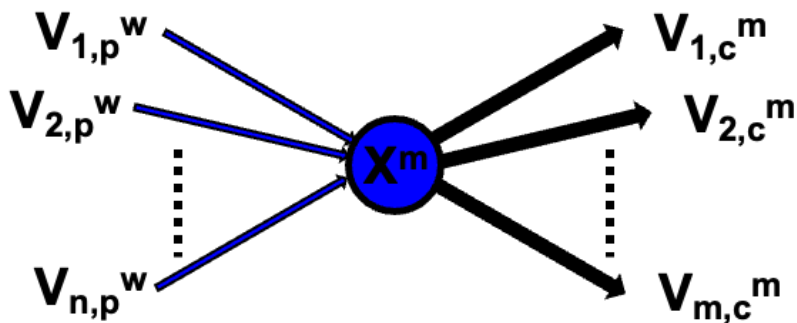
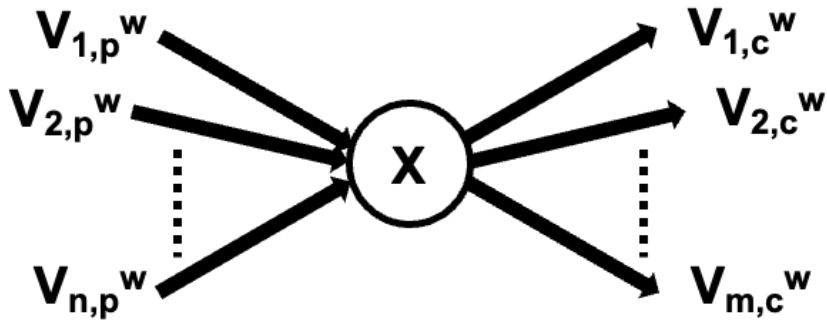
or

$$\sum_{j=1}^m V_{j,c}^m \geq \frac{X^w}{X^m} \sum_{j=1}^m V_{j,c}^w$$

# REMI method - Metabolomics

From metabolomics to fluxes

- Downregulated metabolite (**Decreased Production**)



$$\sum_{i=1}^n V_{i,p}^m \leq \frac{X^m}{X^w} \sum_{i=1}^n V_{i,p}^w$$

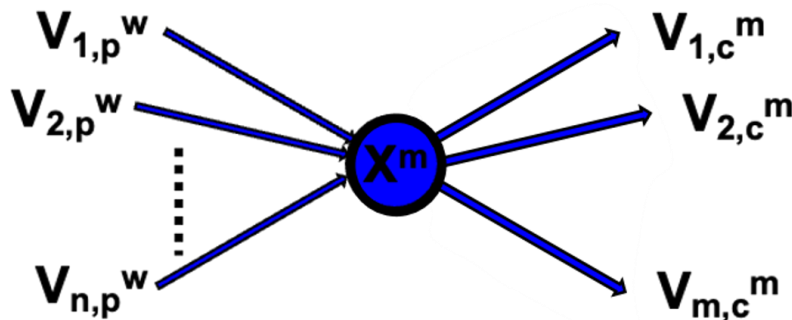
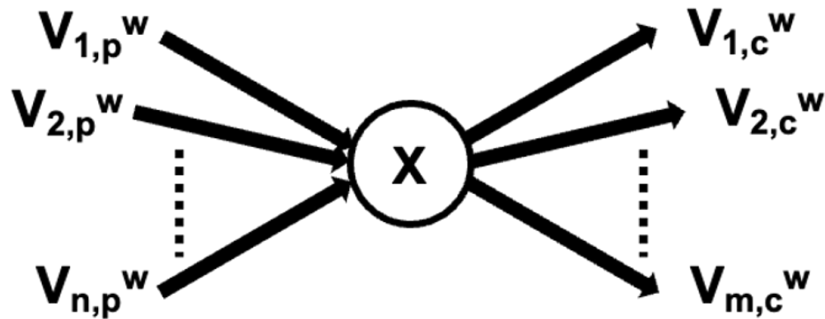
or

$$\sum_{j=1}^m V_{j,c}^m \geq \frac{X^w}{X^m} \sum_{j=1}^m V_{j,c}^w$$

# REMI method - Metabolomics

From metabolomics to fluxes

- Downregulated metabolite (**Decreased Production**)



$$\sum_{i=1}^n V_{i,p}^m \leq \frac{X^m}{X^w} \sum_{i=1}^n V_{i,p}^w$$

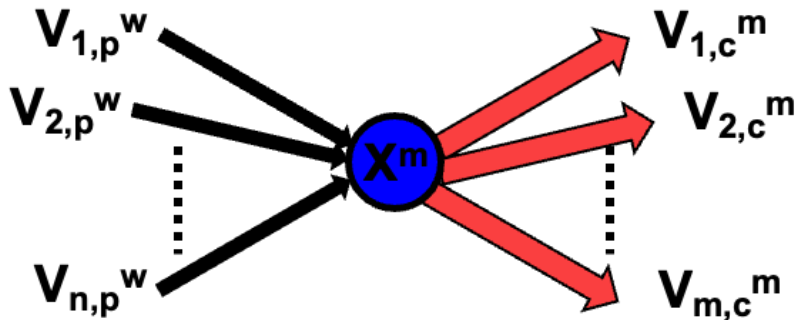
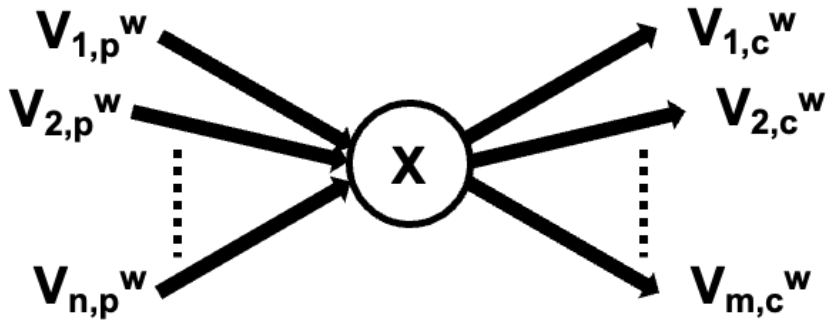
or

$$\sum_{j=1}^m V_{j,c}^m \geq \frac{X^w}{X^m} \sum_{j=1}^m V_{j,c}^w$$

# REMI method - Metabolomics

From metabolomics to fluxes

- Downregulated metabolite (**Increased Consumption**)



$$\sum_{i=1}^n V_{i,p}^m \leq \frac{X^m}{X^w} \sum_{i=1}^n V_{i,p}^w$$

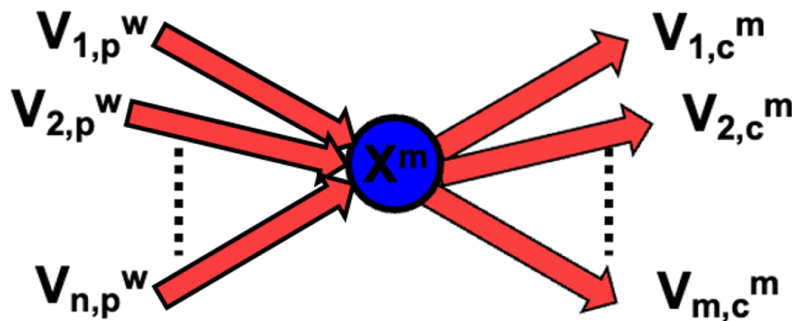
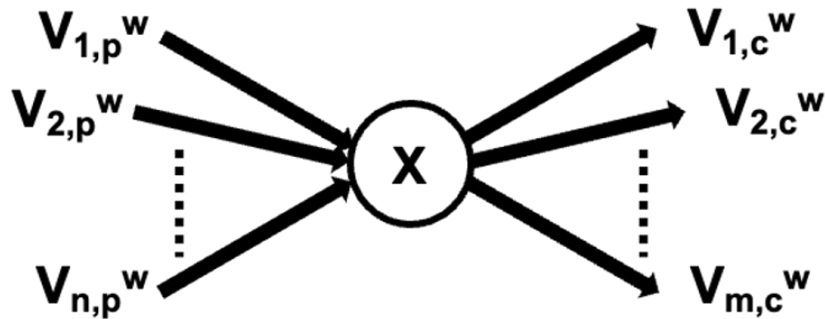
or

$$\sum_{j=1}^m V_{j,c}^m \geq \frac{X^w}{X^m} \sum_{j=1}^m V_{j,c}^w$$

# REMI method – Metabolomics

From metabolomics to fluxes

- Downregulated metabolite (**Increased Consumption**)



$$\sum_{i=1}^n V_{i,p}^m \leq \frac{X^m}{X^w} \sum_{i=1}^n V_{i,p}^w$$

or

$$\sum_{j=1}^m V_{j,c}^m \geq \frac{X^w}{X^m} \sum_{j=1}^m V_{j,c}^w$$

# And many more methods...

Algorithm	Year	Programming language
Covert-01 <sup>3</sup>	2001	NA
Covert-02 <sup>26</sup>	2002	NA
Akesson-04 <sup>27</sup>	2004	NA
SR-FBA <sup>28</sup>	2007	NA
Shlomi-08 <sup>17</sup>	2008	NA
GIMME <sup>9</sup>	2008	MATLAB + COBRA
E-Flux <sup>29</sup>	2009	MATLAB + COBRA
Moxley <sup>30</sup>	2009	NA
MBA <sup>31</sup>	2010	MATLAB + COBRA
MADE <sup>32</sup>	2011	MATLAB + COBRA
tFBA <sup>33</sup>	2011	NA
RELATCH <sup>34</sup>	2012	MATLAB + COBRA
INIT <sup>16</sup>	2012	MATLAB + COBRA + RAVEI
mCADRE <sup>35</sup>	2012	MATLAB + COBRA
AdaM <sup>36</sup>	2012	NA
Lee-12 <sup>37</sup>	2012	MATLAB + COBRA
Fang-12 <sup>38</sup>	2012	MATLAB + COBRA
GX-FBA <sup>39</sup>	2012	MATLAB + COBRA

Algorithm	Year	Programming language
TEAM <sup>40</sup>	2012	NA
GIM3E <sup>41</sup>	2013	Python+COBRApy
EXAMO <sup>42</sup>	2013	Python standalone
MTA <sup>43</sup>	2013	NA
FASTCORE <sup>44</sup>	2014	MATLAB + COBRA
tINIT <sup>45</sup>	2014	NA
E-Fmin <sup>46</sup>	2014	MATLAB
METRADE <sup>47</sup>	2015	MATLAB
Lsei-FBA <sup>48</sup>	2015	R-package
FASTCORMICS <sup>49</sup>	2015	MATLAB + COBRA
TREM-Flux <sup>50</sup>	2015	MATLAB
RegrEx <sup>51</sup>	2015	MATLAB + COBRA
CORDA <sup>52</sup>	2016	MATLAB + COBRA
OM-FBA <sup>53</sup>	2016	MATLAB
E-Flux2 <sup>54</sup>	2016	MATLAB+Java+ (MOST)
SPOT <sup>54</sup>	2016	MATLAB+Java+ (MOST)
metaboGSE <sup>55</sup>	2018	R-package
Benchmark-driven <sup>56</sup>	2019	MATLAB