Systematic construction and validation of kinetic model from metabolic networks

PART-I

AKHIL KUMAR, Summer Intern

Insulin and Diabetes

- Insulin:
  - Pancreatic ß- cells secrete insulin
  - Functions of Insulin:
    - Muscle - immediate energy
    - Liver - medium term energy storage (as glycogen)
    - Fat cells - long term energy storage (fat deposition)

Insulin and Diabetes

- Diabetes:
  - Glucose is chronically elevated
  - Leads to blindness, kidney failure, limb amputation, cardiovascular disease, and death
  - Types:
    - Type I diabetes (“Juvenile”)-
      - absolute lack of insulin
      - Autoimmune destruction of ß- cells
    - Type II diabetes (“Adult onset”): relative lack of insulin-
      - Reason 1: insulin resistance
      - Reason 2: failure of ß- cells to produce enough insulin to compensate

Bringing Metabolic networks to life

- Translating a known metabolic network into a dynamic model requires rate laws for all chemical reactions
- Mathematical expressions depend on the underlying enzymatic mechanism; they can become quite involved and may contain a large number of parameters
- Rate laws and enzyme parameters are still unknown for most enzymes
- As a first attempt, all reactions could be described by versatile laws such as mass-action kinetics, generalised mass-action kinetics or linlog kinetics. However, these kinetic laws fail to describe enzyme saturation at high substrate concentrations, which is a common and relevant phenomenon**?**

Convenience Kinetics

- Convenience kinetics can be used to '*translate a biochemical network into a dynamical model with plausible biological properties*'
- It implements enzyme saturation and regulation by activators and inhibitors, covers all possible reaction stoichiometries, and can be specified by a small number of parameters
- Parameter estimates can be easily computed from a least-squares fit to:
  - Michaelis-Menten values,
  - turnover rates,
  - equilibrium constants, and
  - other quantities (routinely measured and stored)

Convenience Kinetics (contd.)

$$\alpha_1 A_1 + \alpha_2 A_2 + \ldots \leftrightarrow \beta_1 B_1 + \beta_2 B_2 + \ldots,$$

$$v(a,b) = E \frac{k_+^{cat} \prod_i \tilde{a}_i^{\alpha_i} - k_-^{cat} \prod_j \tilde{b}_j^{\beta_j}}{\prod_i (1+\tilde{a}_i+\ldots+\tilde{a}_i^{\alpha_i}) + \prod_j (1+\tilde{b}_j+\ldots+\tilde{b}_j^{\beta_j}) - 1}$$

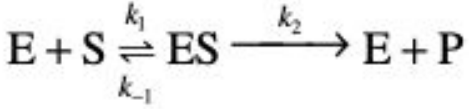
Reaction velocities don't only depend on reactant concentrations, but can also be controlled by **modifiers**. For each of them, we multiply the above equation by a prefactor,

Prefactor = 1 + (activator concentration/activation constant) , or

1 / (1 + (inhibitor concentration/inhibition constant))

Michaelis Menten Kinetics (or Saturation Kinetics)

- A mathematical model of the kinetics of single-substrate-enzyme catalyzed reactions was first developed by Michaelis-Menten
- At high substrate conc., enzyme is saturated



- Assumptions:
  - ES complex is established rather rapidly
  - Rate of the reverse reaction of the second step is negligible (only holds if product accumulation is negligible)

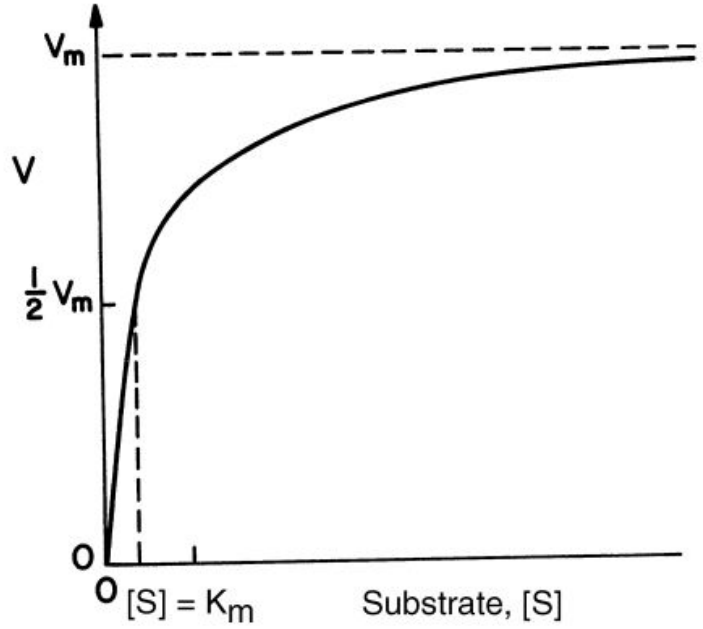
Michaelis Menten Kinetics (contd.)

- Major approaches to derivation:
  - Rapid-Equilibrium approach (used by Michaelis-Menten)
  - Quasi-Steady state approach

$$v = \frac{d[\mathrm{P}]}{dt} = k_2 \frac{[\mathrm{E}_0][\mathrm{S}]}{K'_m + [\mathrm{S}]} = \frac{V_m[\mathrm{S}]}{K'_m + [\mathrm{S}]}$$

- Vm : maximum forward velocity of reaction
- Michaelis constant:

$$K_M = \frac{k_2 + k_{-1}}{k_1}$$



Turnover Number

Maximum number of substrate molecules that can be converted into product molecules per catalytic site of a given concentration of enzyme per unit time

$$k_{cat} = \frac{V_{max}}{E_t}$$

Kinetic Models

- It consists of '*x*' **enzymatic reactions**, '*y*' **metabolic state variables**, '*z*' **parameters**
- The model integrates some subsystems. E.g.
  - Subsystems integrated in an exemplar model:
    - Glycolysis
    - TCA cycle
    - Respiratory chain
    - NADH shuttles
    - Pyruvate cycle
- It also takes into account compartmentalisation of reactions in various parts of the cell
  - Compartmentalisation into:
    - Cytoplasm and mitochondrial matrix

Kinetic Models(contd.)

- Model parameters exerting the most influence on model results are identified through a '**sensitivity analysis**':
  - Example: Parameter sensitivity analysis:
    - Critical parameters to system behavior as of the Glycolytic pathway than parameters in TCA cycle than in other pathways
    - Initial reactions of non-glycolytic pathways are important for behaviour of system

PART-II

AKHIL KUMAR,
Summer Intern

BACKGROUND

- The mechanistic description of enzyme kinetics in a dynamic model of metabolism requires specifying the numerical values of a large number of kinetic parameters

- **Challenge:** Dealing with the frequent cases where data to construct detailed kinetic models is lacking
- **Solution:** The parameterisation challenge is often addressed through the use of simplifying approximations to form reaction rate laws with reduced number of parameters

Can such simplified models reproduce the dynamic characteristics of the full system?

Models used

- A set of kinetic models of **Red Blood Cell metabolism** using various approximate rate laws were constructed
- The approximate rate laws used were:
  - Michaelis-Menten rate law with measured enzyme parameters
  - Michaelis-Menten rate law with approx. via Convenience kinetics
  - Thermodynamic rate law resulting from a metabolite saturation assumption
  - Pure chemical reaction mass action rate law that removes the role of enzyme from the reaction kinetics
- We utilized the *in vivo data* for human RBC to compare the effects of rate law choices against the backdrop of physiological flux and concentration differences

Robustness Check

- A straightforward way to estimate the similarity of the behavior between different rate laws is to simulate the response of each model to perturbation
- **Perturbation:** Here, it denotes the change of certain metabolite concentrations at time t=0, after which the system is allowed to simulate through a long enough time such that the original steady state is once again reached.
- Nine different perturbations were performed
    - For example, Perturbation 1: the concentrations of ATP, ADP, and Pi were perturbed at the same time to simulate the hydrolysis of ATP in the system

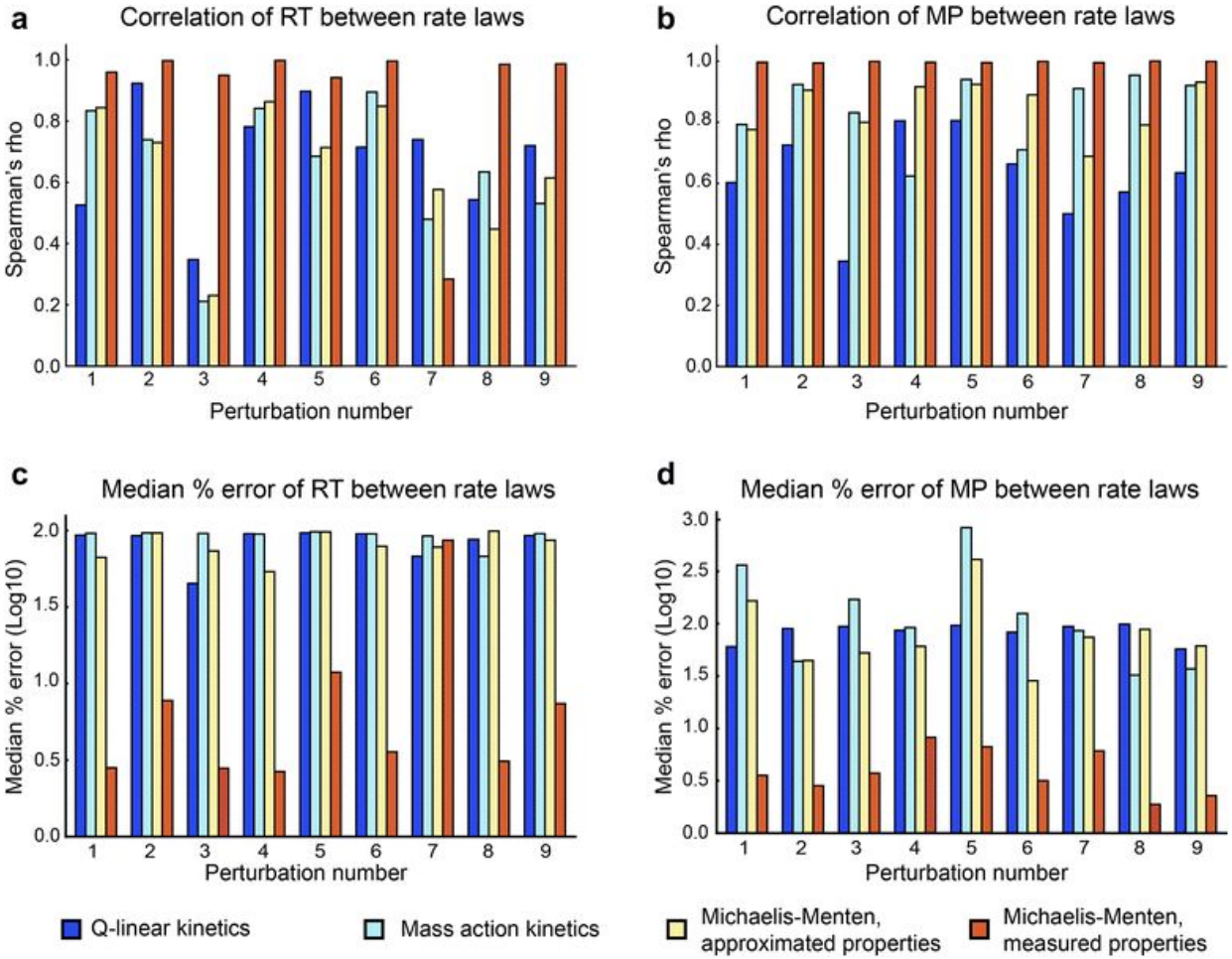
MP: Maximum perturbation

(Largest percent change in concentration compared to the steady state concentration that occurred during the simulation)

RT: Relaxation time

(to calculate the RT of a metabolite, we identify the last time point at which the deviation from the steady state concentration is at least 5% of the MP)

Simulation comparison of four simplified rate laws against a reference module

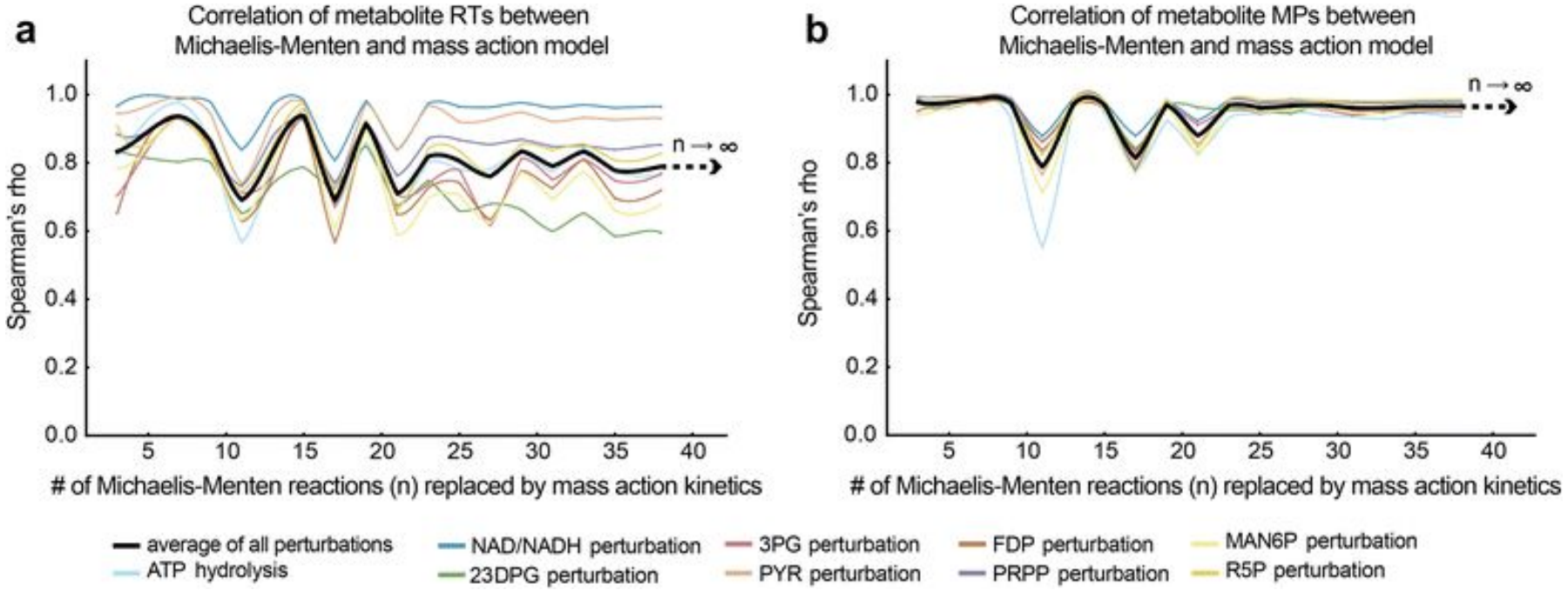


Results

- From Spearman correlation and Median percent error plots it was found that the **Michaelis-Menten kinetics with measured properties** behaved substantially better on both metrics compared to other rate laws
- These analyses were repeated on the previously published models of RBC, and the trends were **verified**
- Michaelis-Menten rate law with measured enzyme parameters yields an excellent approximation of the full system dynamics
- However, iteratively replacing mechanistic rate laws with approximations resulted in a model that retains a high correlation with the true model behavior

Results

- Since the simplified rate laws introduce noticeable discrepancies in the dynamic behavior, so, we want to determine
  - whether these discrepancies would continue to increase as simplified rate laws are applied to more reactions until the correlation completely disappears, or
  - whether the approximate model would stabilize at some positive correlation to the true model
- We found that the discrepancy ceases to grow after a certain point, and it appears likely that **models constructed with entirely simplified rate laws will be useful approximations of the real system**, at least for small perturbations



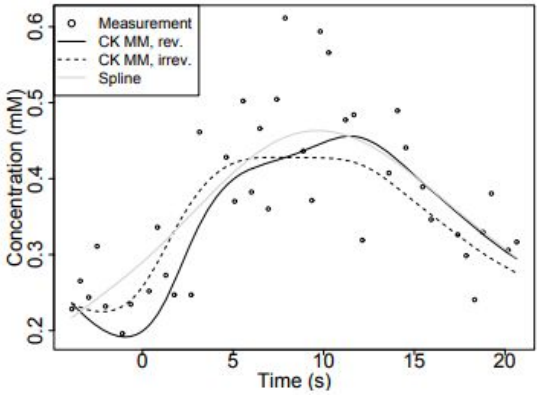
A simple test case with Michaelis-menten kinetics was set up, and then the Michaelis-Menten based reactions were iteratively replaced with mass action kinetics based reactions

The correlation of metabolite RT and MP between Michaelis-Menten and mass action kinetics fluctuated initially but gradually stabilized as more reactions were replaced with mass action kinetics

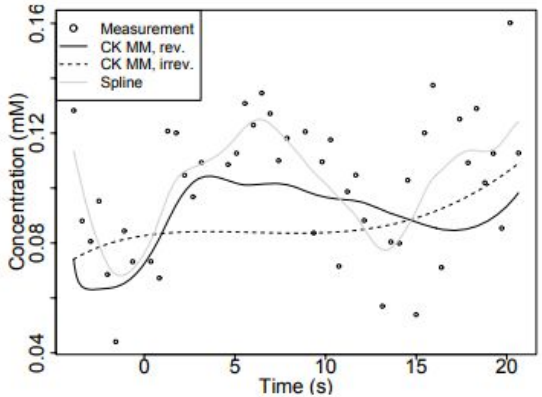
CK: Reversible/irreversible ??

- We have two types of Convenience Kinetics(CK): Reversible and Irreversible
- To decide which one is better and should be preferred,
  - two alternate differential equation models were constructed
  - Evolutionary algorithms were applied to both the models

- Evolutionary Algorithms are known to handle highly non-linear optimization problems



(a) Acetolactate (AcLac)



(d) 2-Ketoisocaproate (KIC)

CK: Reversible / Irreversible ?

4 Optimization Procedures out of 8 different optimization procedures tested showed **Outstanding performance** approximating the in vivo data

With regards to the more complex metabolic systems, it was shown that the **Convenience kinetics** is an appropriate standard formalism when exact knowledge of the underlying mechanism is not available

Reversible variant allowed good reconstruction of the experimental data while the irreversible alternative often produced implausible straight lines

References

1. Bin Du, Daniel C. Zielinski, Erol S. Kavvas, et al.

   Evaluation of rate law approximation in bottom-up kinetic models of metabolism. BMC Systems Biology

   https://bmcsystbiol.biomedcentral.com/articles/10.1186/s12918-016-0283-2#MOESM1

2. Andreas Dräger, Marcel Kronfeld, Jochen Supper, et al.

   Benchmarking Evolutionary algorithms on Convenience Kinetics models of Valine and Leucine biosynthesis in C. glutamicum. 2007 IEEE Congress on Evolutionary Computation

   http://www.cogsys.cs.uni-tuebingen.de/publikationen/2007/Draeger2007b.pdf

References(contd.)

3. Pedro Mendes, Natalie J. Stanford, Kieran Smallbone

Kinetic Modeling of large-scale metabolic networks

https://dl.acm.org/citation.cfm?id=2037511

Part-III

AKHIL KUMAR,
Summer Intern

Fundamental problem in Functional Genomics?

- mRNAs
- Proteins
- Metabolites
- Protein-protein interactions
- Protein-DNA interactions

How to integrate the informations obtained!?

Systems Biology is the field of research attempting to ***understand all of the interactions in a system*** as opposed to focusing on the individual parts.

Metabolic Engineering

- **Challenge:** To build genome-scale network models that allow quantitative predictions of the cell's state along with time

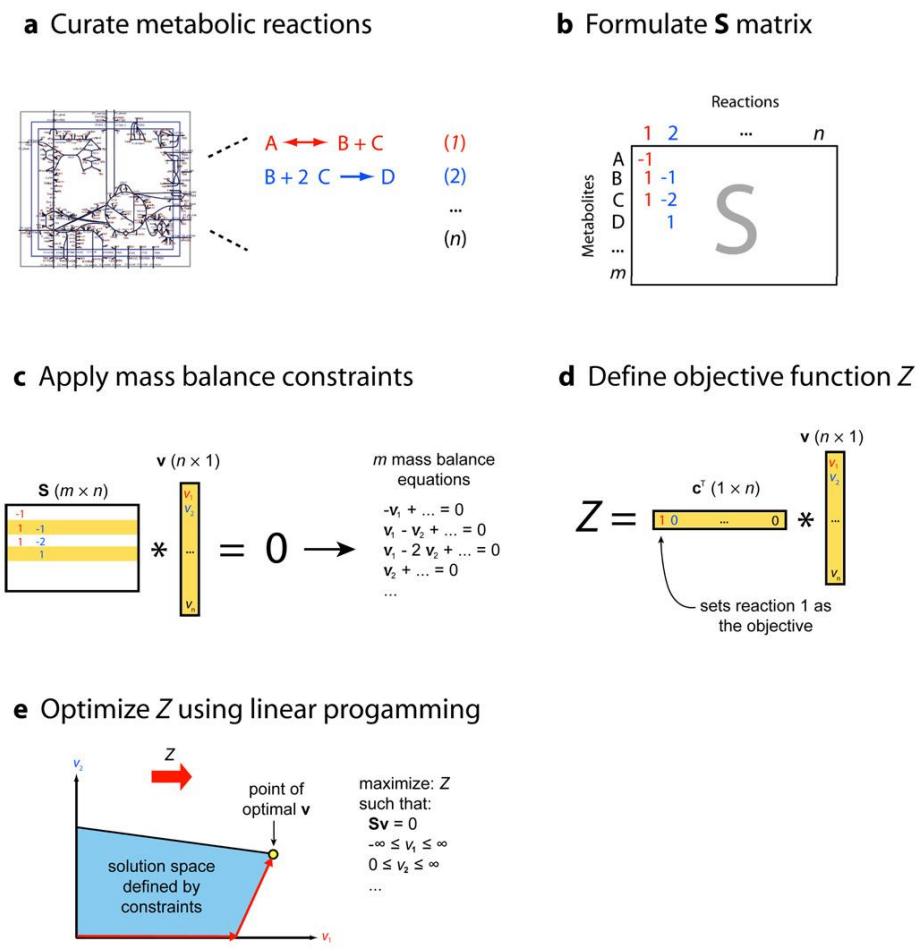
We try to tackle this issue through novel *in silico* approaches, such as reconstruction of dynamic models

- **Limitation:** The lack of available experimental information—which affects the accuracy and feasibility of solutions.
- Major Approaches:
  - Stationary Representation of the system(Constraint Based Modelling)
  - Dynamic Representation of the system (ODE based)
  - Hybrid Modelling

Flux Balance Analysis

For analyzing the flow of metabolites through a metabolic network and finding a ***relevant flux distribution*** [Orth et al., 2010]

- Constraints(Typically linear)
  - capacity/reversibility constraints imposed by bounds on the values of the fluxes, or
  - flux balance constraints imposed by stoichiometry
- Formulation of an optimization problem



Dynamic Representation

Dynamic modeling acknowledges the *changes metabolite concentrations suffer over time.*

Metabolic network modeling can be based on the:

- knowledge of enzyme mechanisms and
- experimental data to build a representation of a dynamic system,

It uses a ***system of ordinary differential equations*** (ODEs).

ODEs contain initial values for metabolite concentrations, reaction rate equations and kinetic parameters.

Coming back to the Fundamental Problem

Existing Literature on Multi-omics data integration is mainly in FBA approach

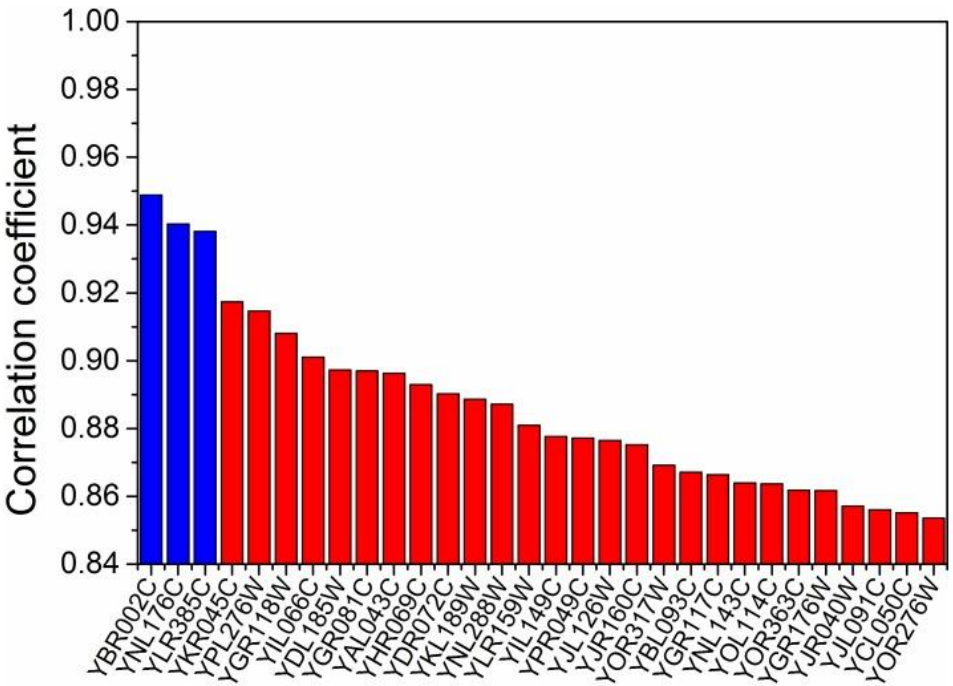
Key points:

- There are two types of algorithms:
  - Switch approach (GIMME, iMAT)
  - Valve approach (E-flux, PROM)
- **Assumption: “**mRNA transcript levels are a strong indicator of the level of protein activity”
- **Limitation:** Many recent studies have shown that there’s a **positive correlation** between mRNA abundance and protein abundance but from **0.21 to 0.61**.

FBA derived algorithm

Problem: The objective function used in FBA approach is **arbitrary** and sometimes leads to inaccurate phenotypic predictions

Solution: Deriving an ***omics guided Objective function*** to develop a novel FBA algorithm(omFBA) to correlate the genotype with the phenotype



Correlation between phenotype-matched weighting factors and gene expressions

Transcriptomics data vs Proteomics data

- ❏ Capture percentage
- ❏ Scalability
- ❏ Reference availability
- ❏ Uniformity
- ❏ Technical Bias
- Do we really need integration of data?

Machado and Herrgård claim that in many cases integration of data doesn't improves model predictions and in some cases simple approach gives better results

- Convenience Kinetics is safe to use if we're perturbing well within physiological range

References

1. What is Flux Balance Analysis(FBA)?
2. It is all about Metabolic Fluxes
3. OM-FBA: Integrate Transcriptomics Data with Flux Balance Analysis to Decipher the Cell Metabolism (using a omics based objective function to create a new algorithm for FBA)
4. Flux balance analysis of biological systems: applications and challenges

   {This tells in section “Analysis of perturbations” to perform robustness check by deleting one or more genes from the system by constraining the reaction fluxes corresponding to genes and therefore to corresponding proteins to 0}

5. Integrated analysis of Transcriptomic and Proteomic data
6. A Review of Dynamic Modeling Approaches and Their Application in Computational Strain Optimization for Metabolic Engineering
7. Genome, transcriptome and proteome: the rise of omics data and their integration in biomedical sciences
8. Integrated Analysis of Transcriptomic and Proteomic Data