

IV121

**Vybrané aplikace informatiky
v biologii**

Úvodní přednáška

David Šafránek, 4.3.2021

Outline

- Přehled obsahu kurzu
- Formální informace o kurzu
- Východiska bioinformatiky a systémové biologie
- Motivace k využití informatiky

Cíl předmětu

- Ukázat využití informatických metod a nástrojů v různých problémech řešených v biologii a biomedicínských oborech
- Poskytnout elementární praktickou znalost (hands-on v každém ze tří témat níže)
- 3 hlavní zaměření:
 - systémová biologie: modelování a analýza komplexních systémů
 - bioinformatika: modelování struktur (vektorové modelování v 3D)
 - zpracování dat: síťová reprezentace informace, dolování znalostí
- Možnost individuálního plánu pro studenty mimo FI

Osnova

- Úvod do problematiky, formální požadavky (*D. Šafránek*)
- Kvalitativní modely, jejich reprezentace a jejich analýza (*D. Šafránek*)
- Kvantitativní modely, jejich reprezentace a analýza (*D. Šafránek*)
- 3D geometrie, CSG (*M. Lexa*)
- Reprezentace informace prostřednictvím sítí (*V. Nováček*)
- Získání dat z literatury (*V. Nováček*)
- Integrace dat (*V. Nováček*)
- Aplikace metod umělé inteligence a strojového učení (*V. Nováček*)

Předpoklady

- Nejsou definovány formální prerekvizity
- Předmět poskytuje elementární vhled do probíraných témat (nikoliv detailní znalost)
- Určeno zejména k rozšíření (mezioborových) obzorů
- Nasměrování do relevantních oborů (zaměření budoucího studia)

Přednášející

- doc. David Šafránek, safranek@fi.muni.cz
 - systémová biologie (zejména modelování a aplikace formálních metod, vývoj a aplikace algoritmů)
- Dr. Matej Lexa, lexa@fi.muni.cz
 - bioinformatika (zejména analýza sekvencí a struktur, vývoj a aplikace algoritmů)
- Dr. Vít Nováček, novacek@fi.muni.cz
 - zpracování informací, umělá inteligence (aplikace v bioinformatice, medicíně i SB)

Klasifikace

- Hodnocení (celkem 100 bodů)
 - cvičení — 3 úkoly, každý 30 bodů, celkem 90 bodů
 - diskuse řešení cvičení na konci semestru (10 bodů)
 - dodatečné body navíc za aktivitu během kurzu
 - známka E: ≥ 50 bodů
- Deadlines
 - 15.6. odevzdání řešení cvičení
 - 16.6. — 7.7. diskuse řešení online

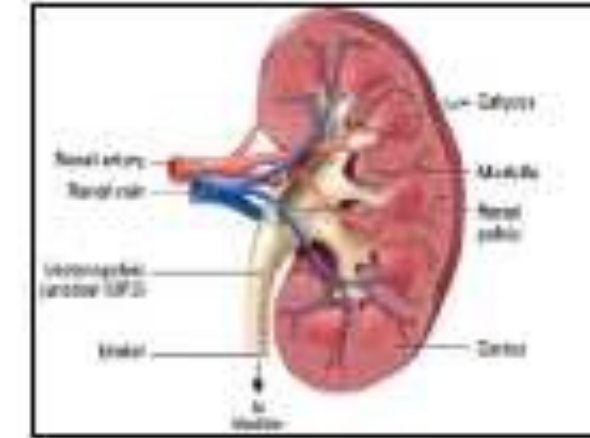
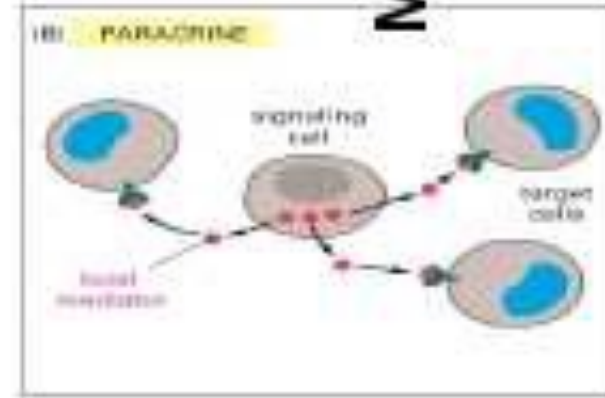
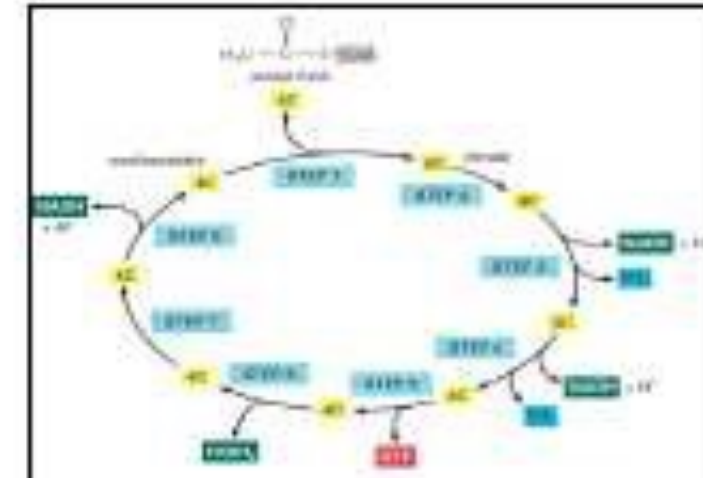
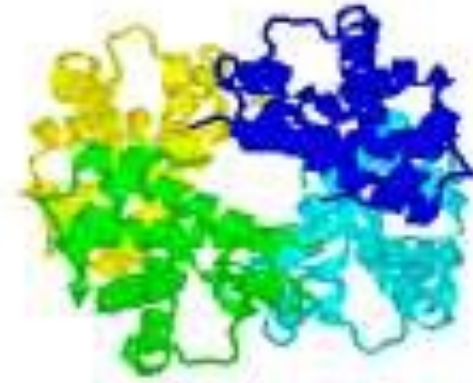
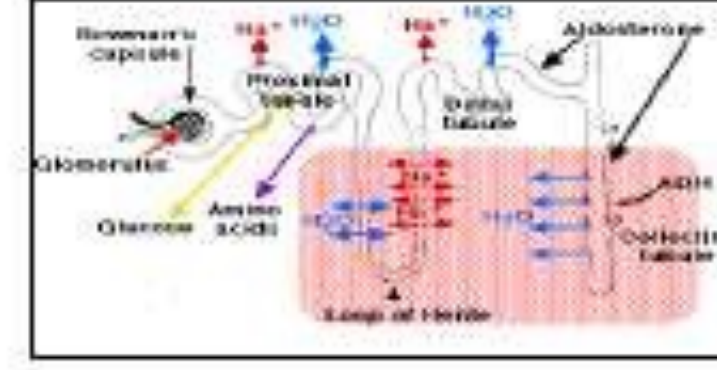
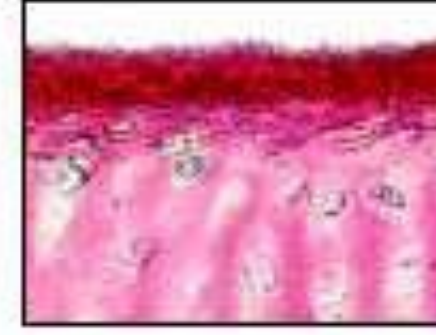
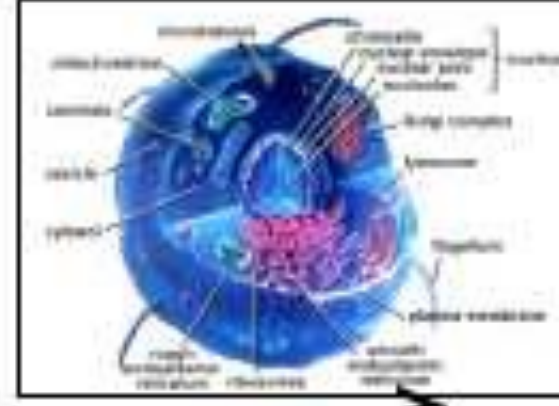
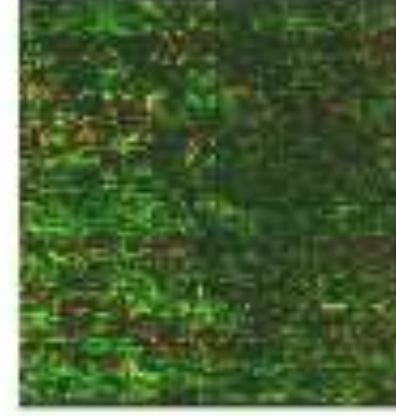
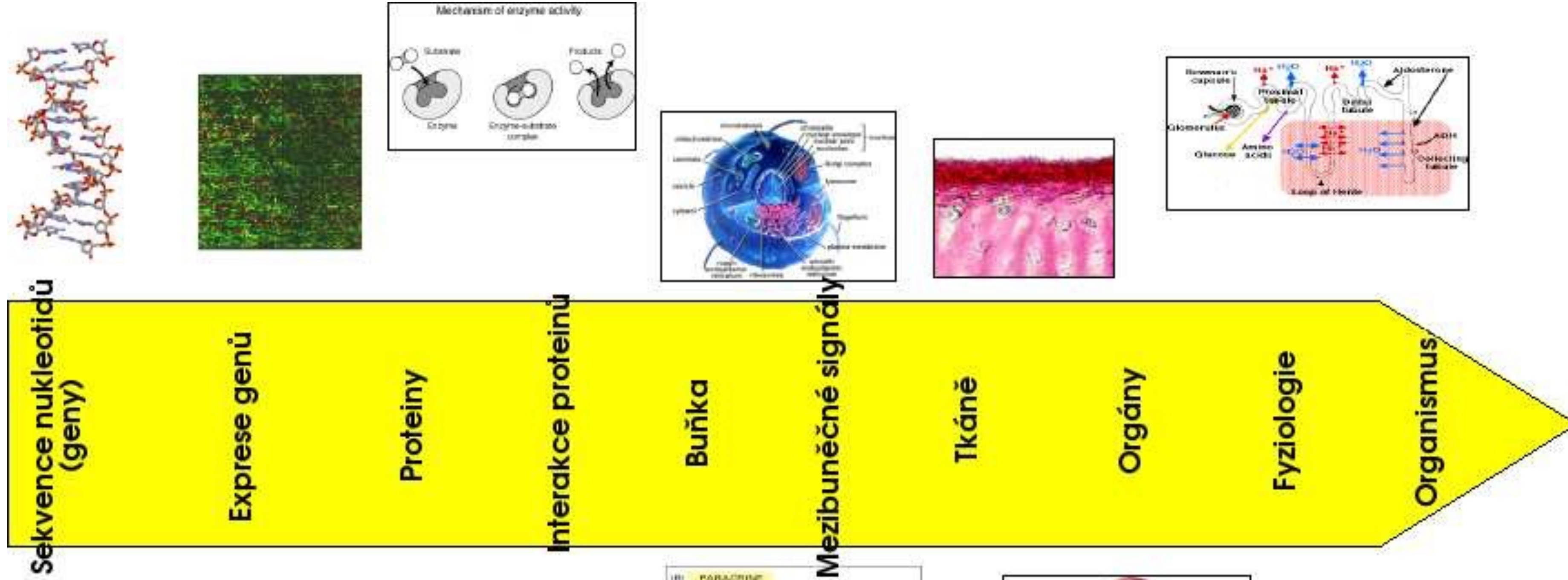
Východiska bioinformatiky a systémové biologie

Buněčná a molekulární biologie

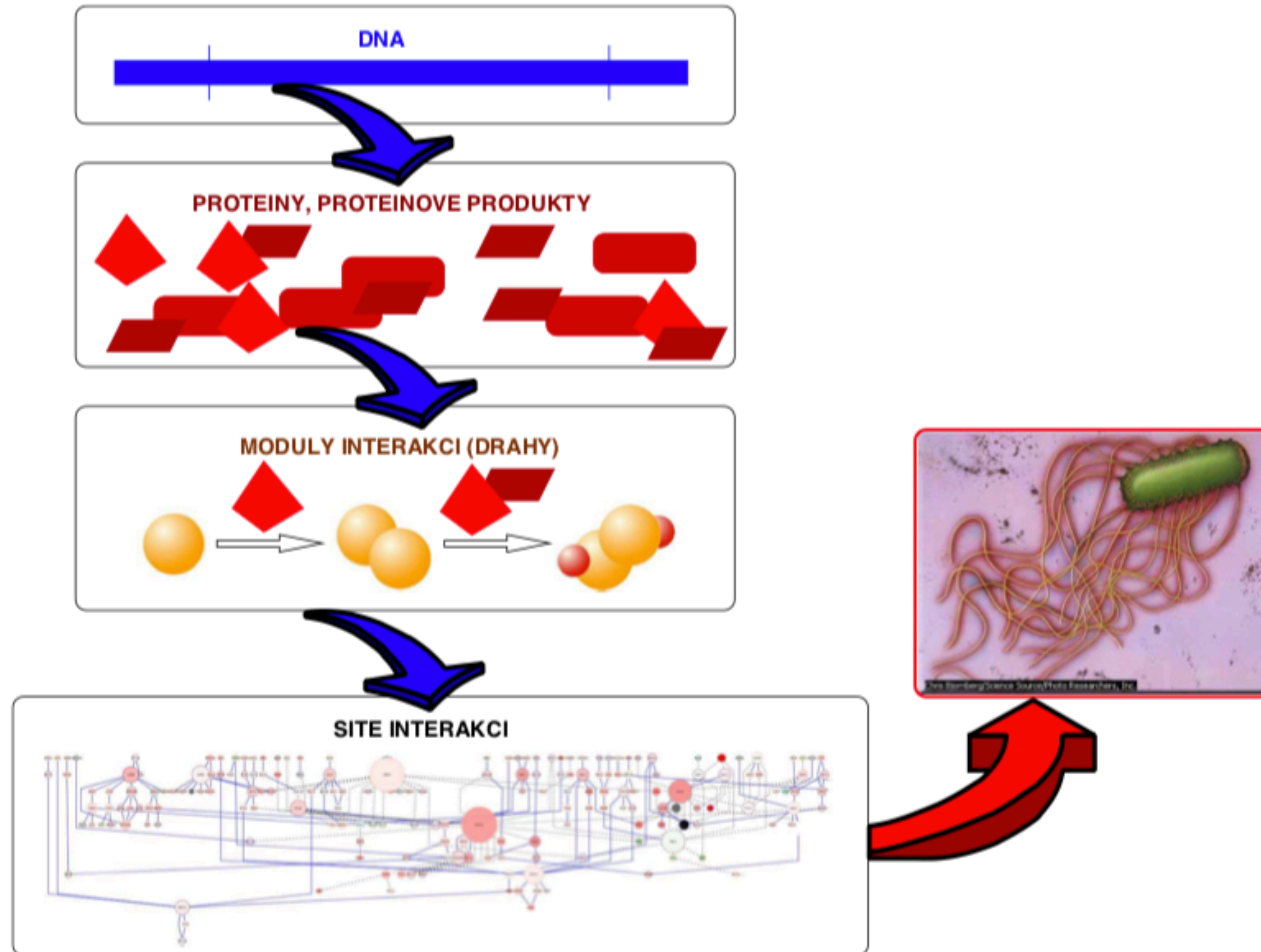


in vivo/in vitro

Expanze informace v biologii



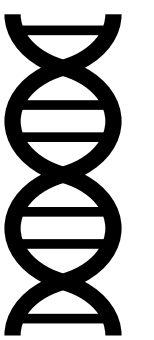
Expanze informace v biologii



Bioinformatika a systémová biologie

Revoluce v biologii I...

- **bioinformatika – “*inteligentní mikroskop*”**
 - uchopení informace v biologii (sekvence, struktury) a o biologii (ontologie)
 - strojové zpracování dat a inference znalostí
 - modelování struktur



Současné metody generují obrovské množství různě (ne)přesné informace o biologických objektech.

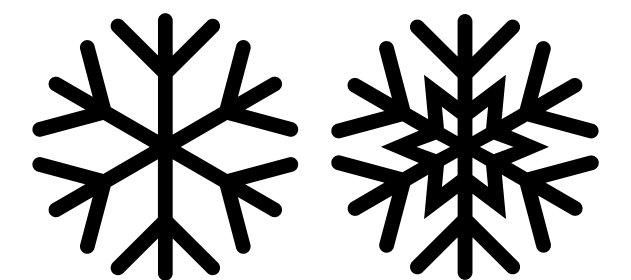
*Bioinformatika poskytuje efektivní techniky pro reprezentaci, zpracování a analýzu **dat**.*

Bioinformatika a systémová biologie

Revoluce v biologii II...

- **systémová biologie**

- pochopení časoprostorového chování biologické entity v kontextu známé informace
- modelování dynamiky biologické entity (procesu, živého organismu)
- vyvození emergentních vlastností a pochopení okolností při kterých vznikají
 - např. rakovinové bujení, cirkadiální rytmy, diferenciace buněk, apod.



Od integrace informací ke spustitelnému modelu biologické entity/procesu.

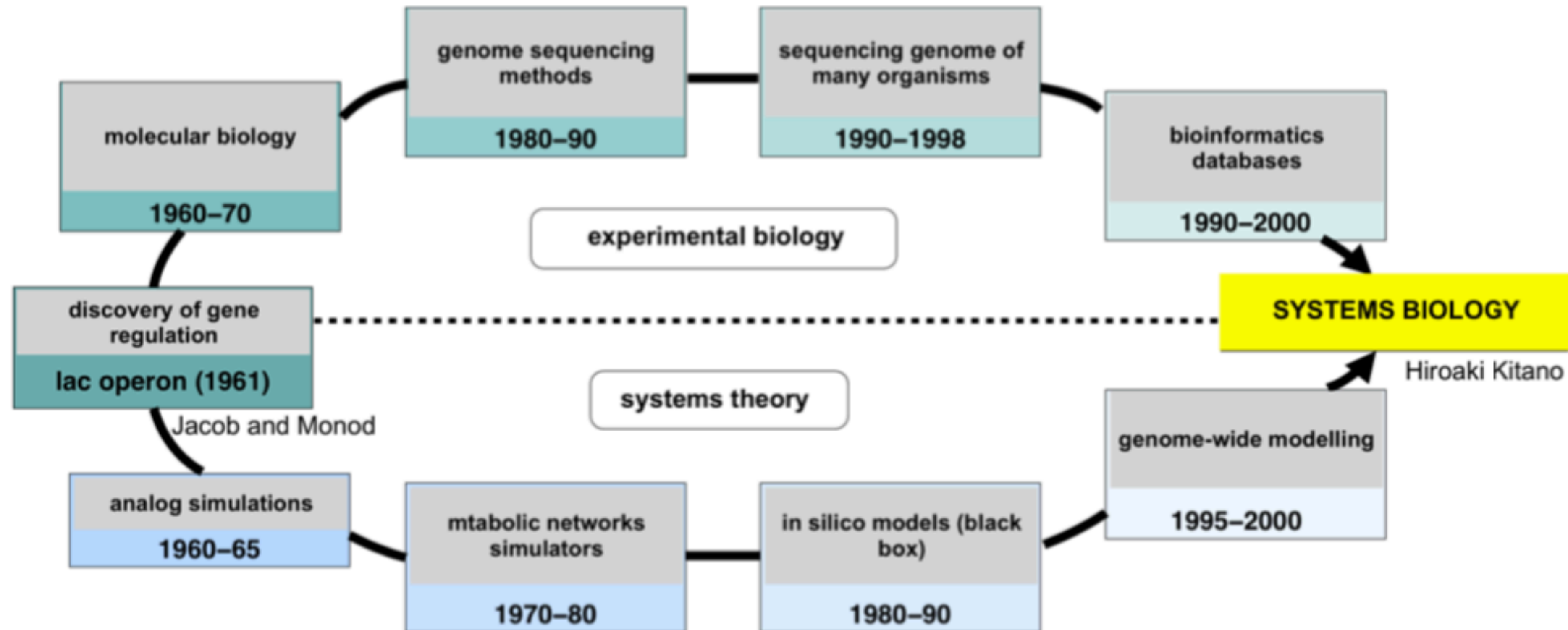
*SB poskytuje efektivní techniky pro reprezentaci, zpracování a analýzu **modelů**.*

Výpočetní systémová biologie



in silico

Systemová biologie



“Major importance need to be placed on the establishment of methodologies and techniques that enable us to understand biological systems as systems.”

Hiroaki Kitano

Systemová biologie

komponentová biologie

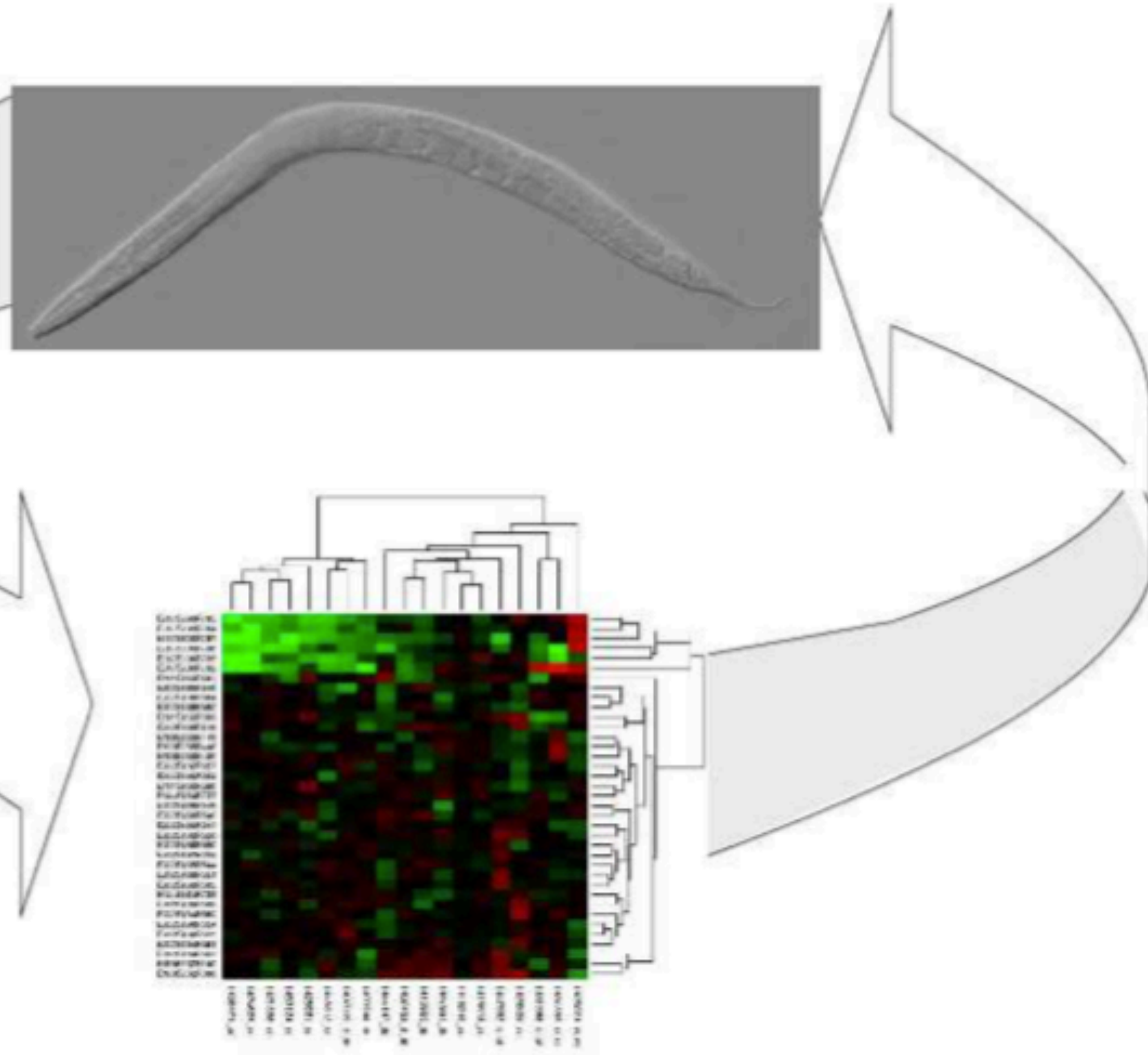
systemová biologie

high-throughput technologie

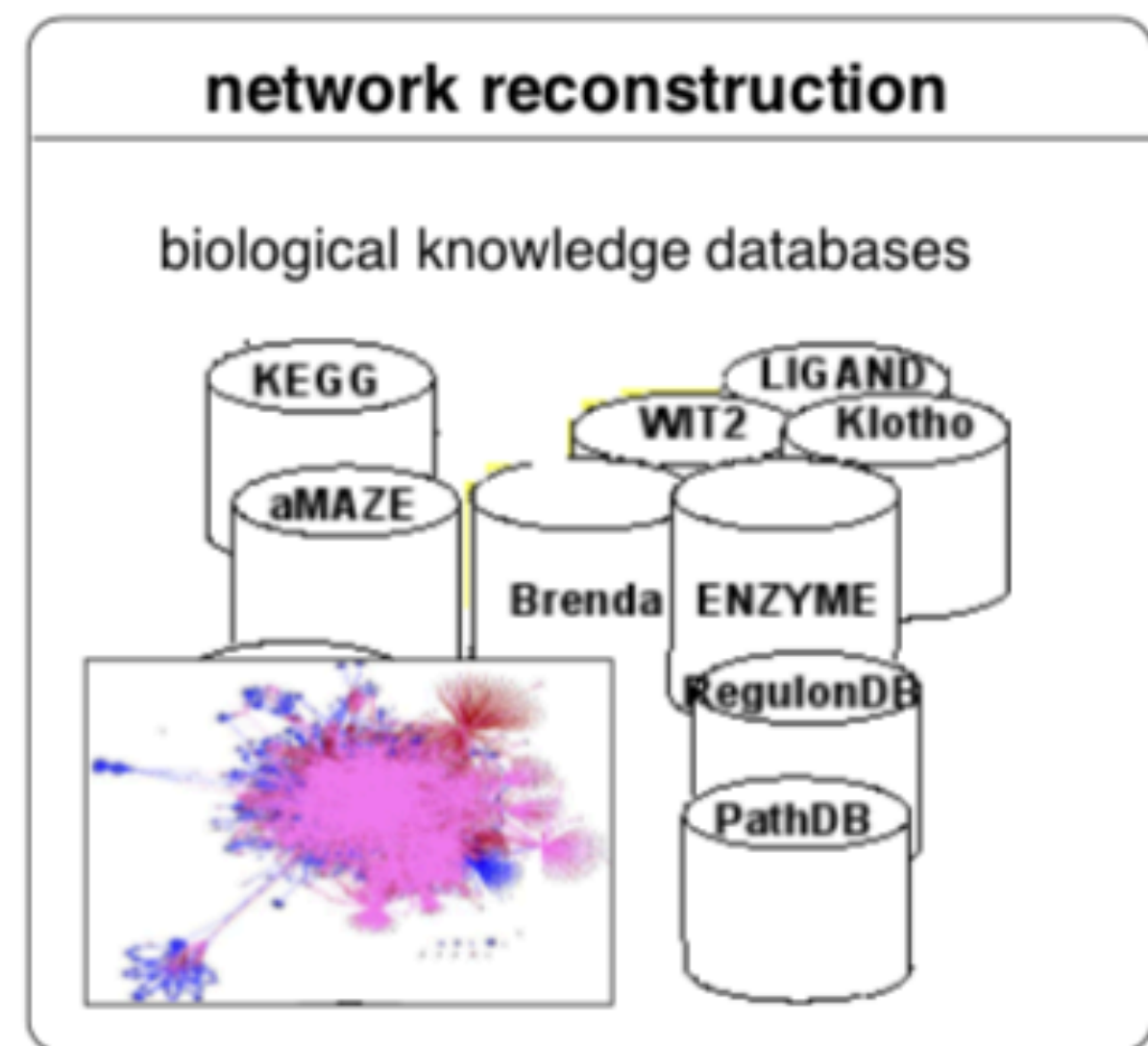
genomika
proteomika

integrativní analýza

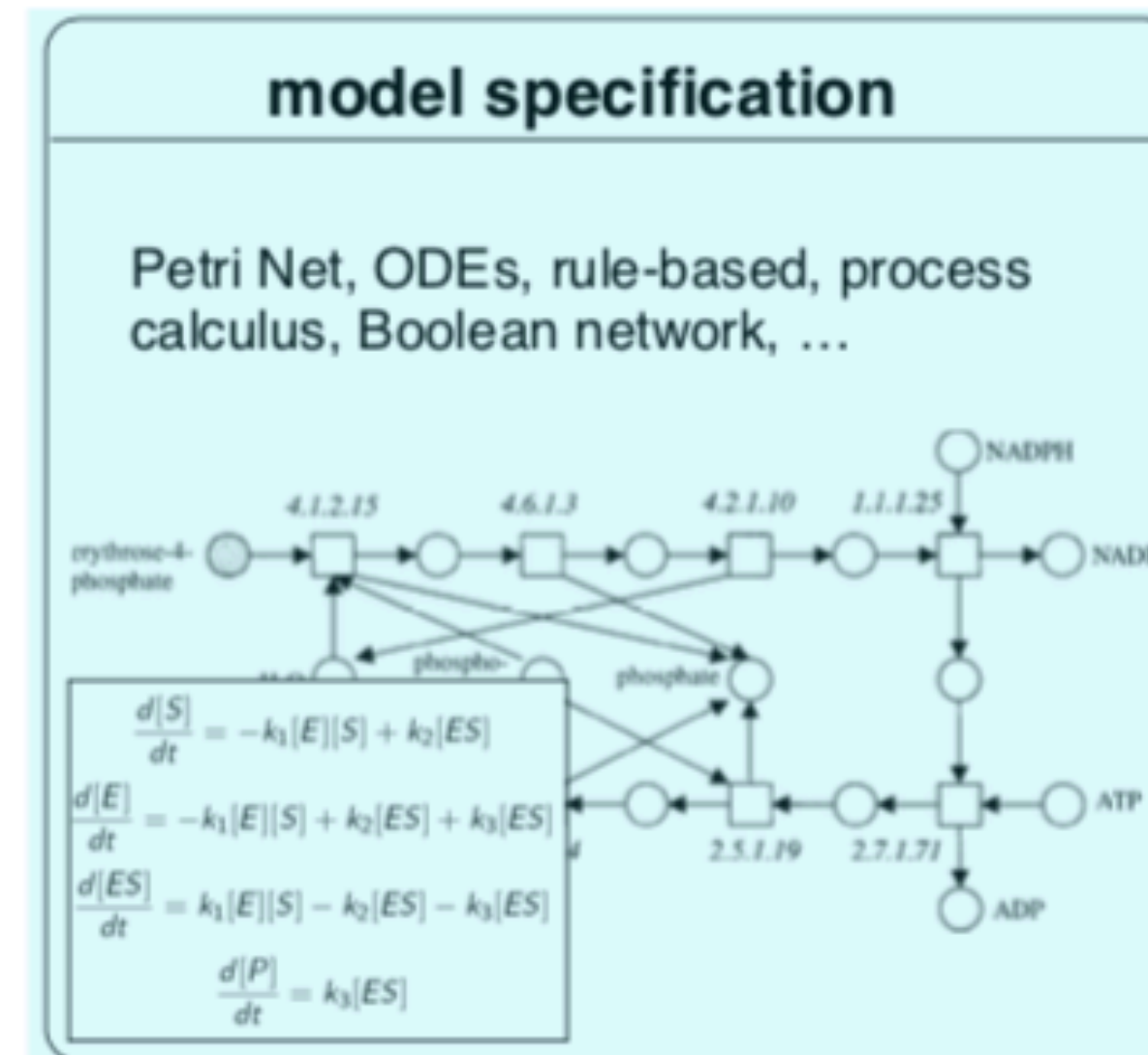
bionformatika
modely (in silico)
simulace



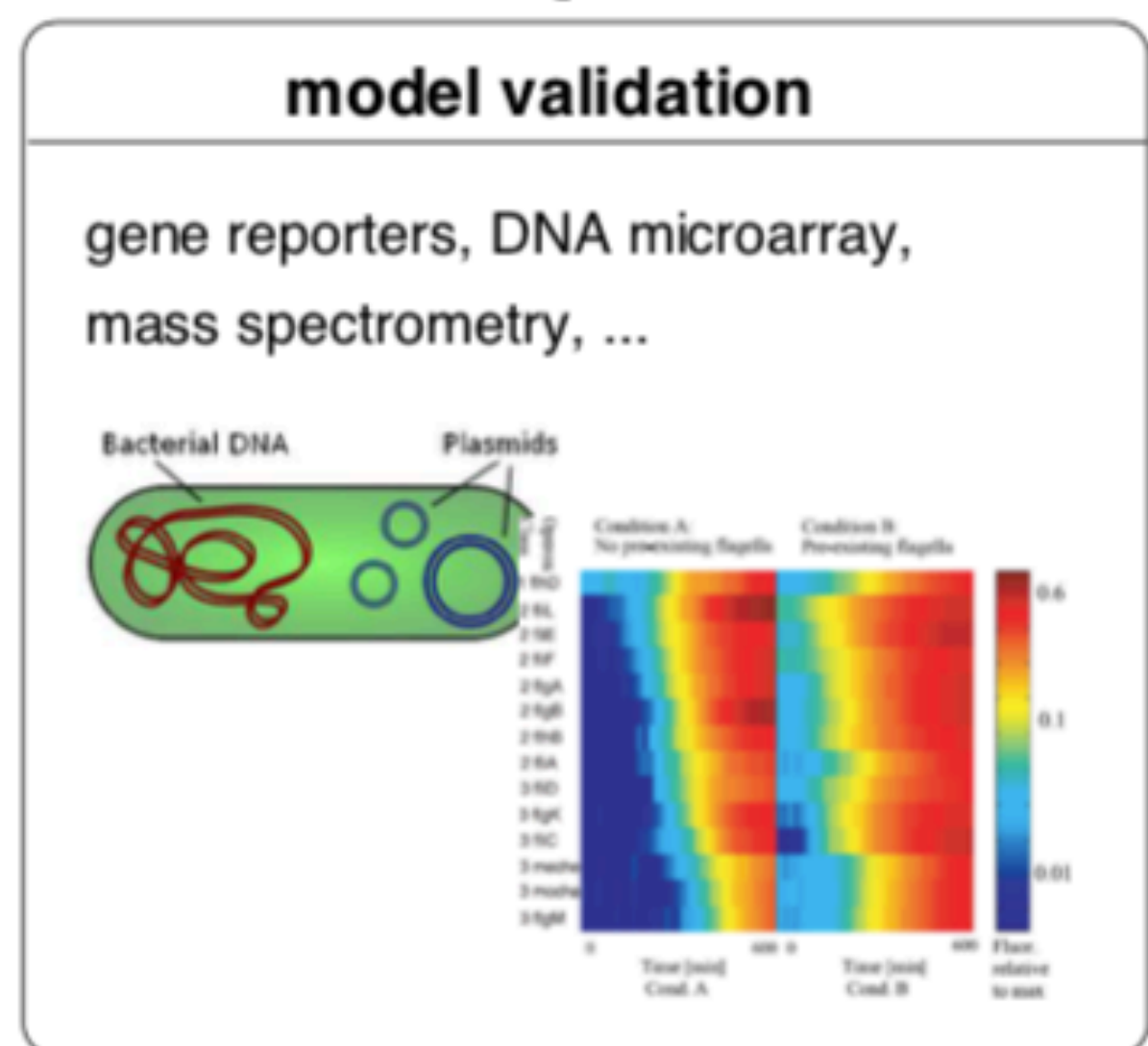
Systemová biologie



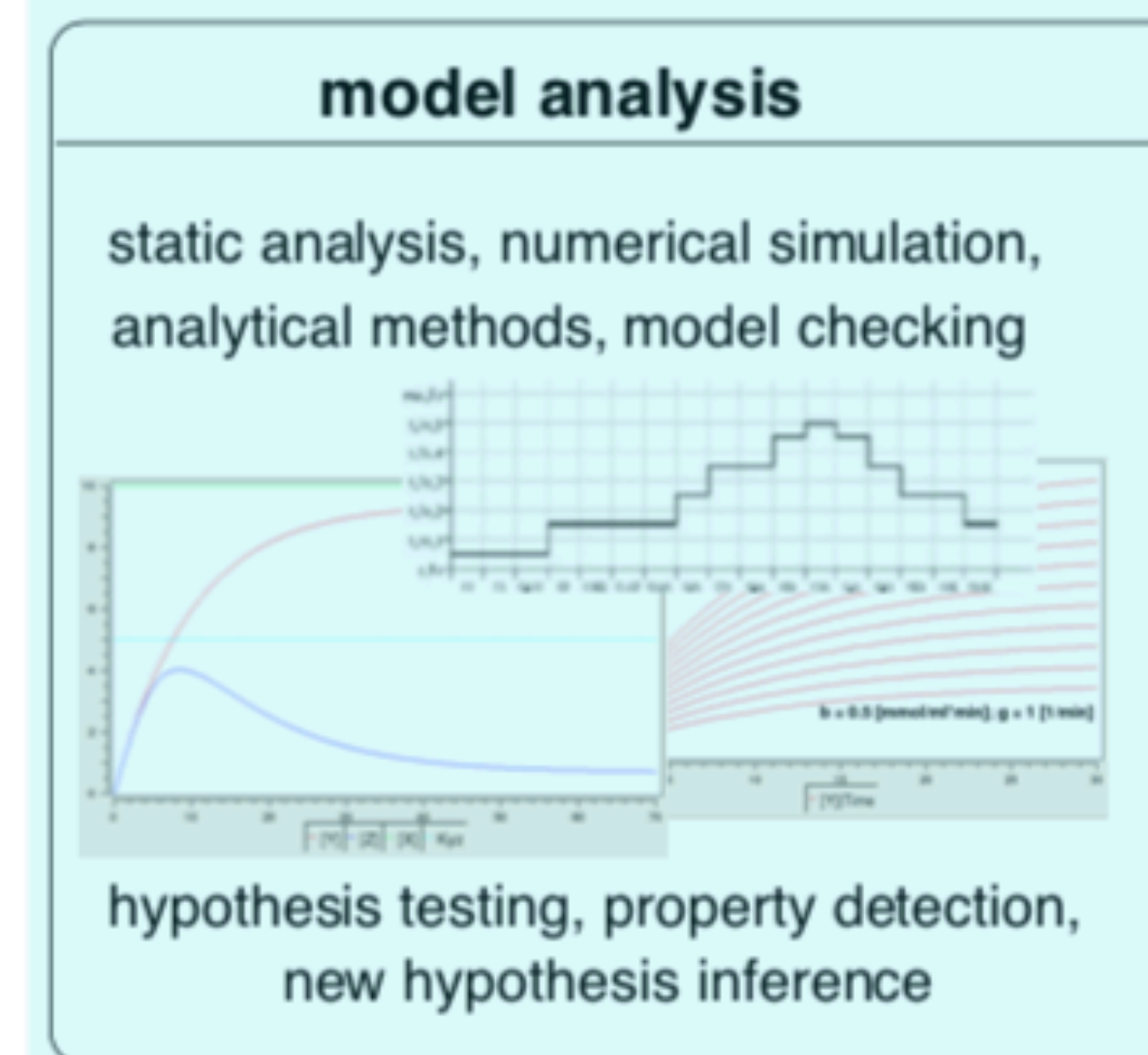
biological network
 hypothesis



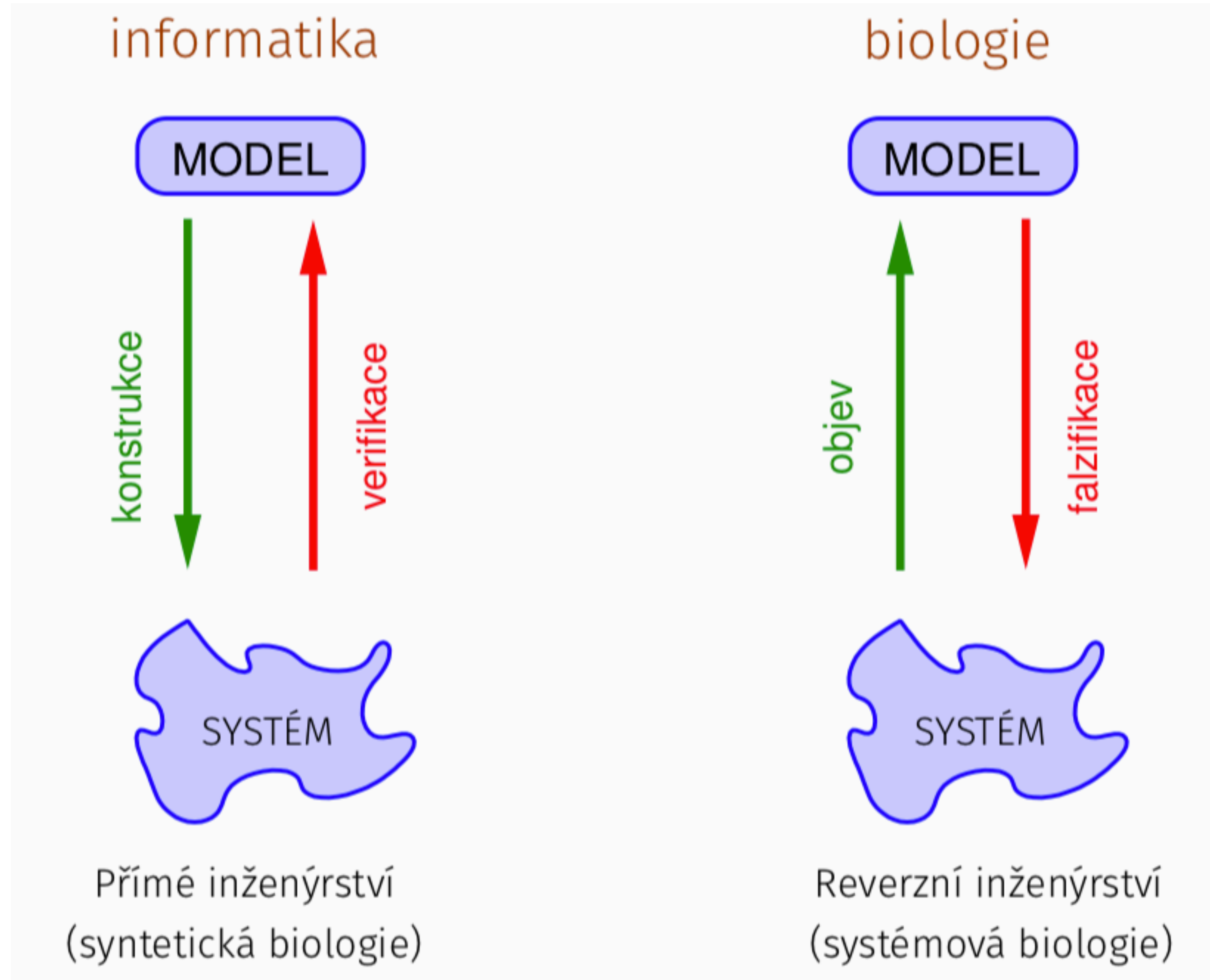
computer lab



emergent properties
 model questions



Využití modelu



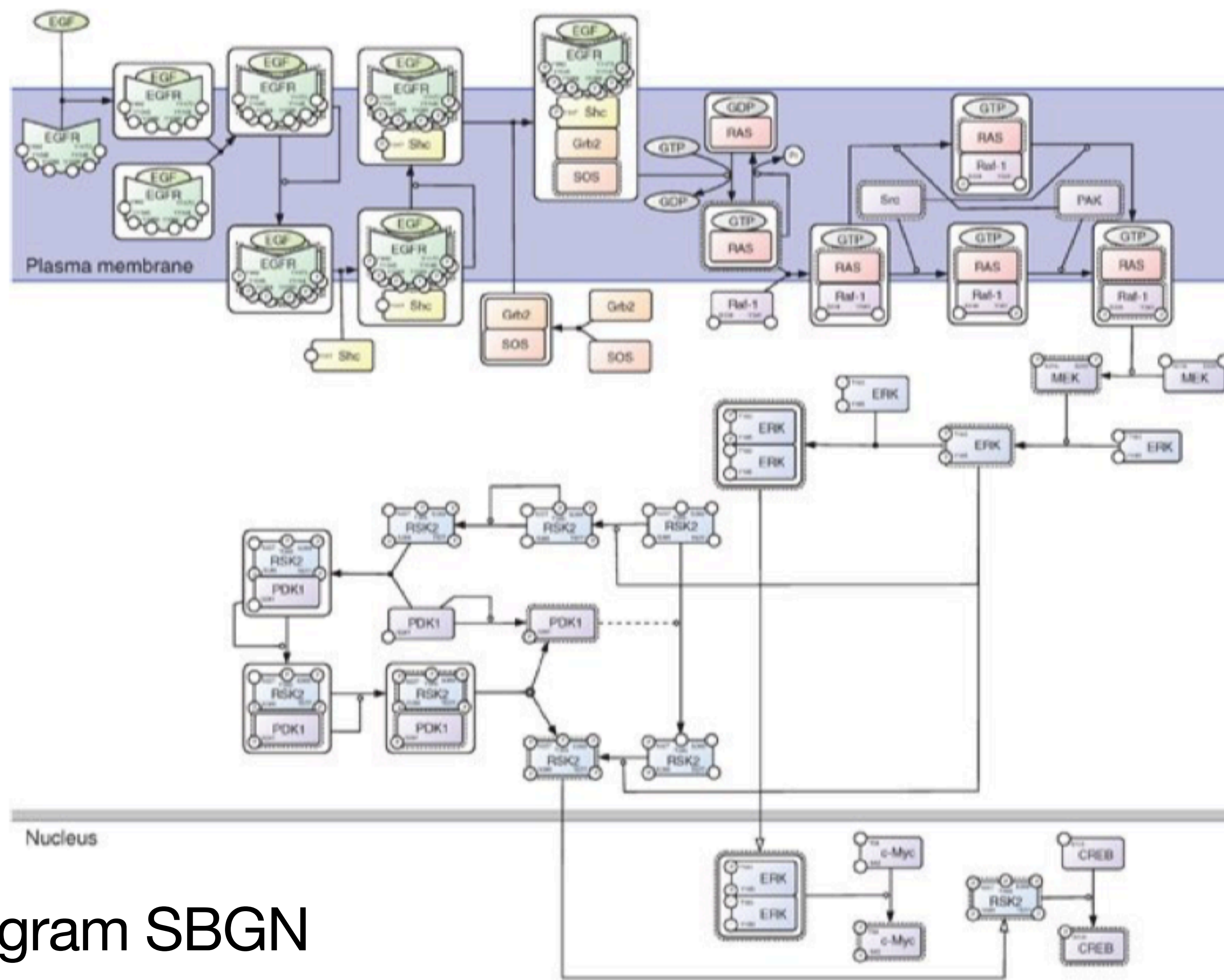
Typy modelů dle struktury

- reakční
- influenční
- reakčně-influenční
- agentové

Typy modelů dle sémantiky

- Diskrétní nebo spojité v čase, bez času
- Diskrétní nebo spojité v prostoru, bez prostorové informace
- Deterministické, nedeterministické, pravděpodobnostní
- Kvalitativní, kvantitativní

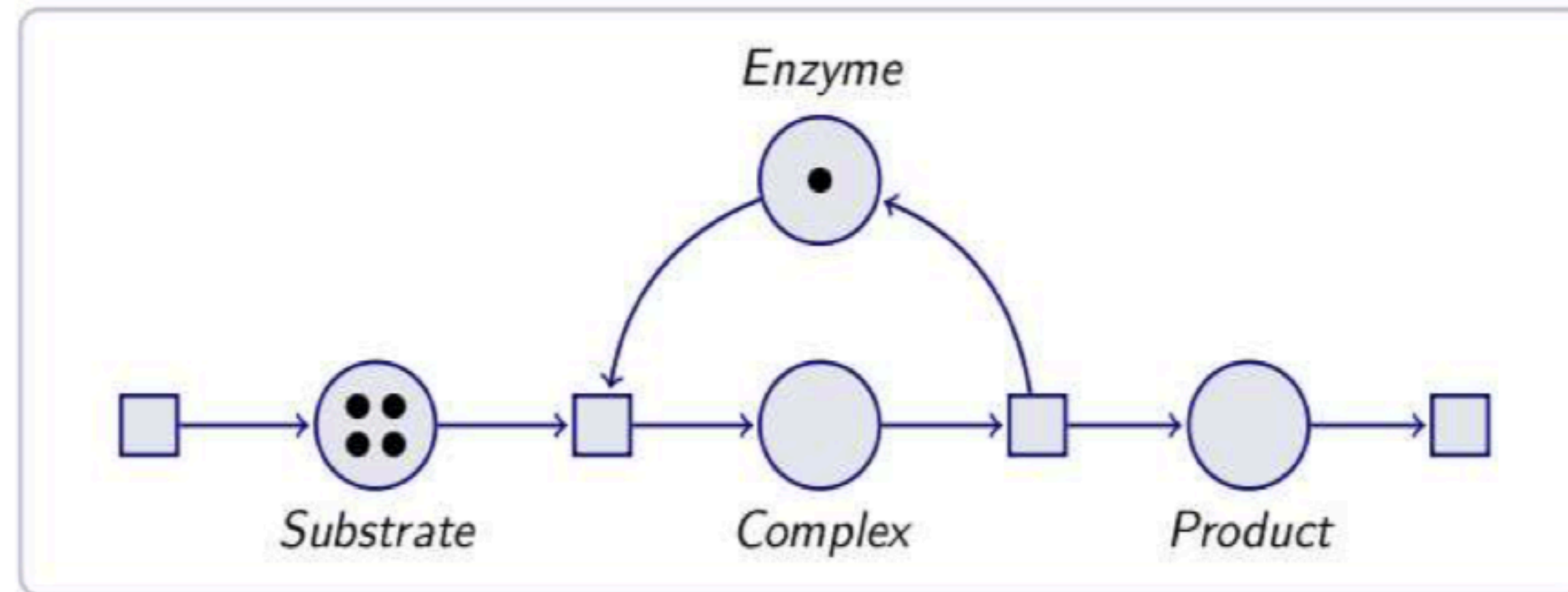
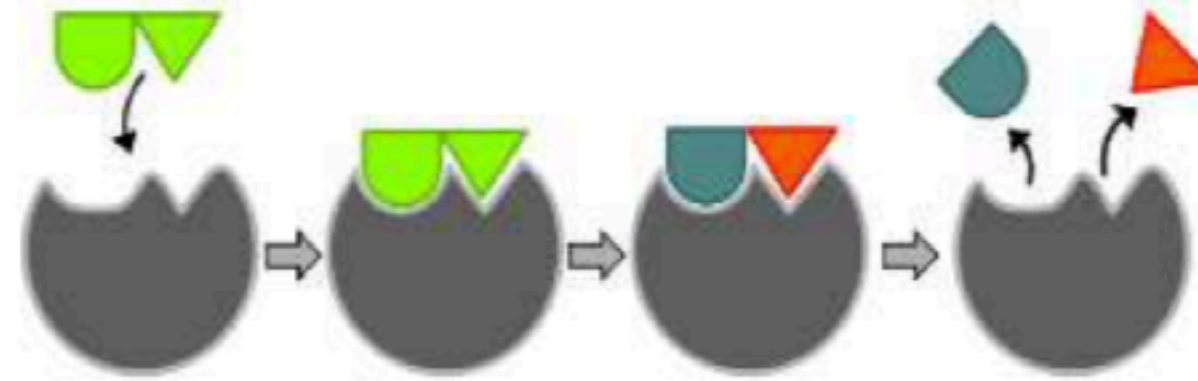
Typy modelů



procesní diagram SBGN

CellDesigner

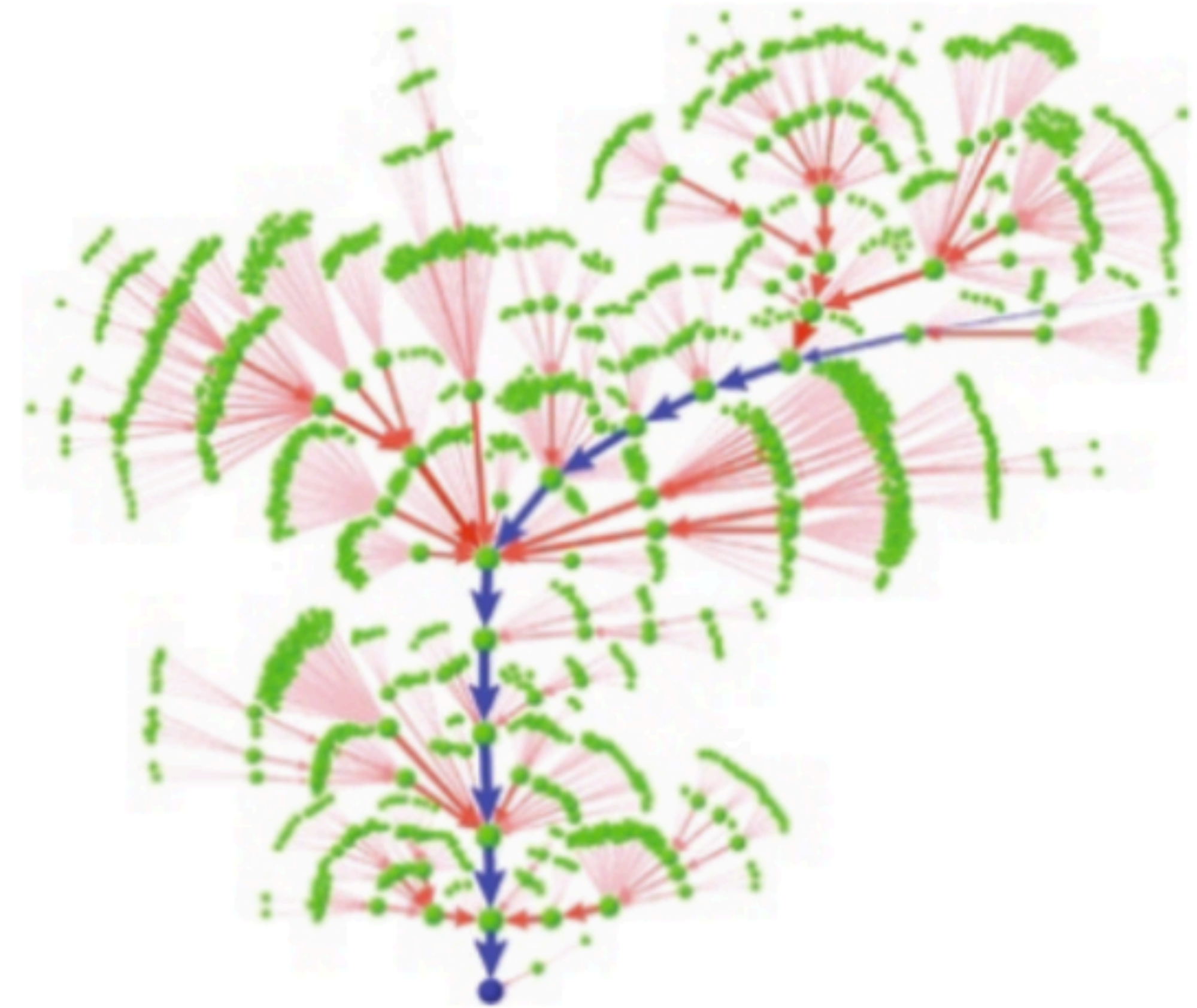
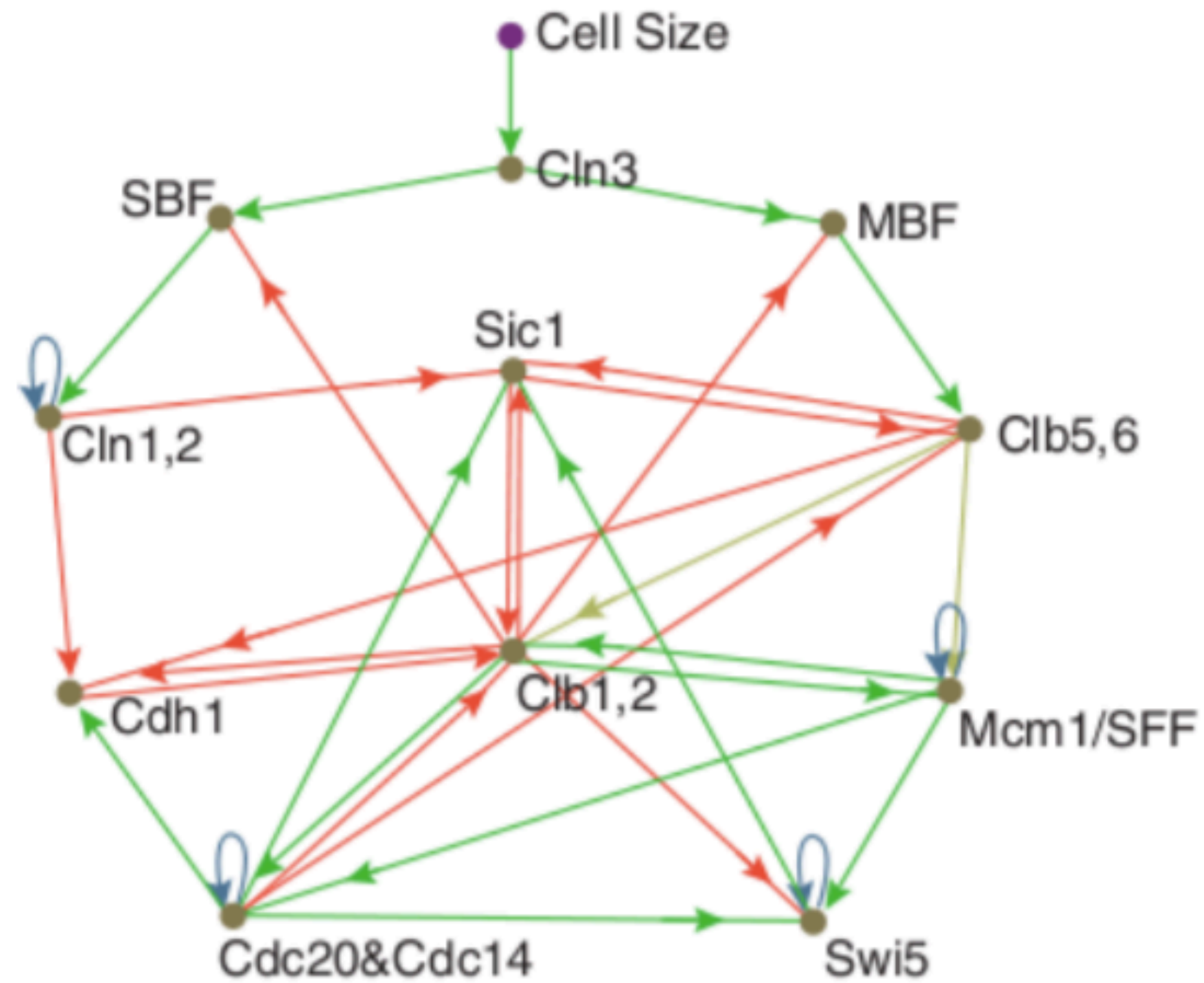
Typy modelů



Petriho síť

Snoopy

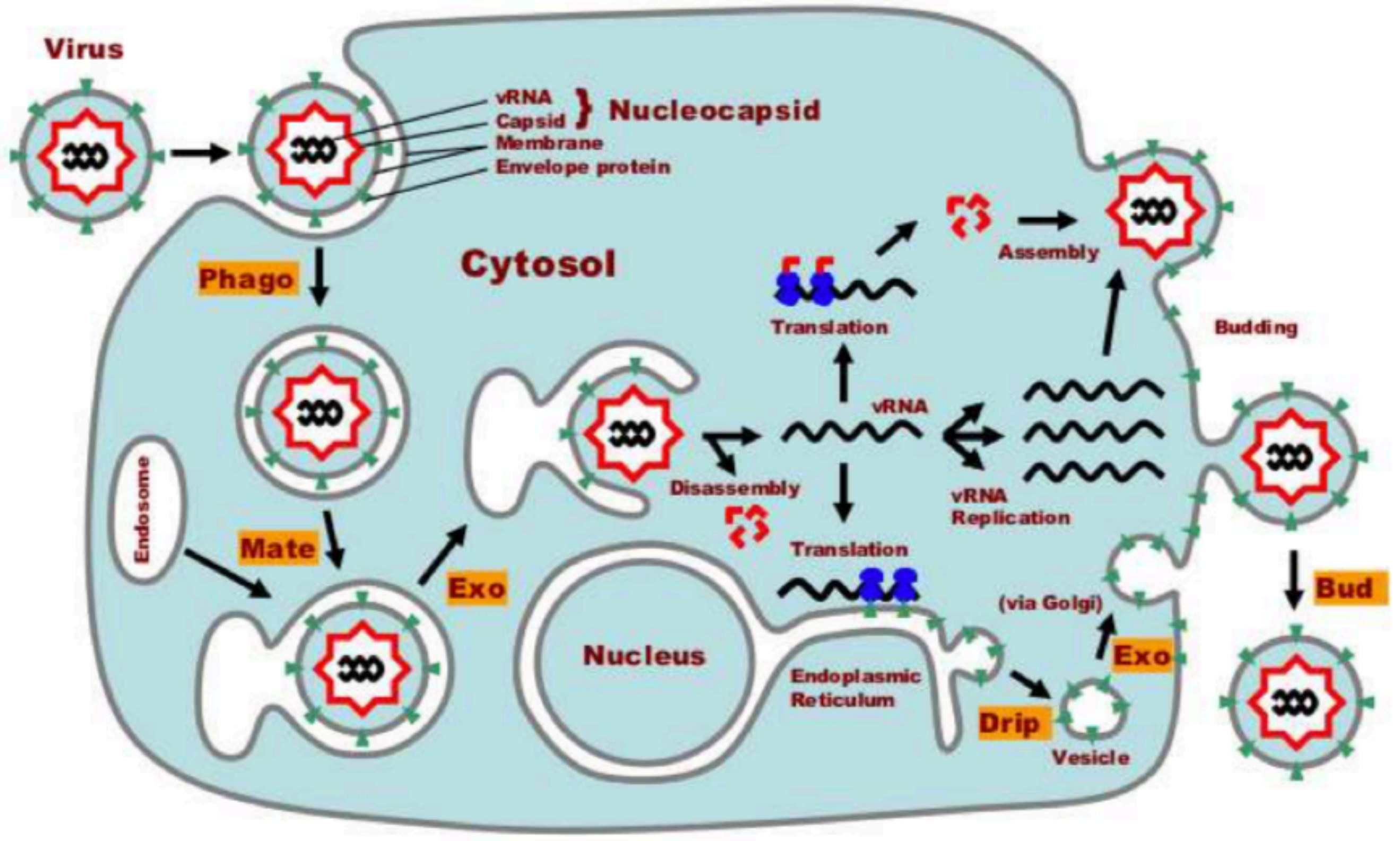
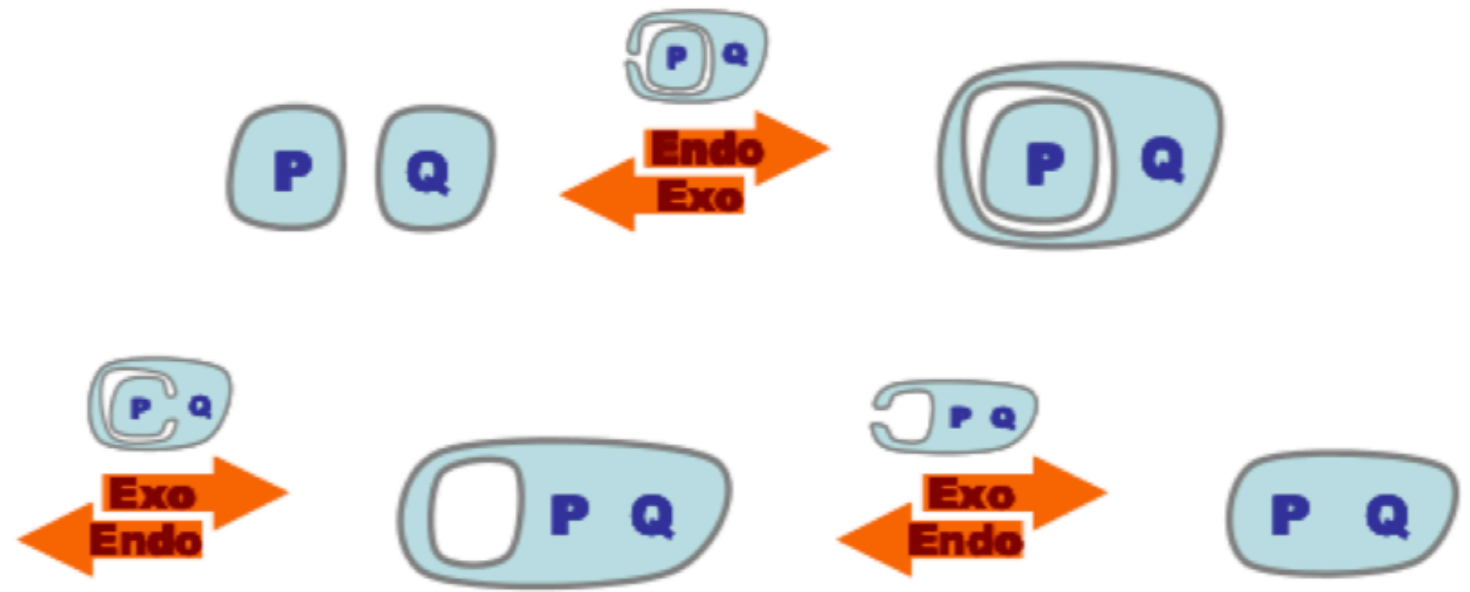
Typy modelů



Boolovská (regulační) síť

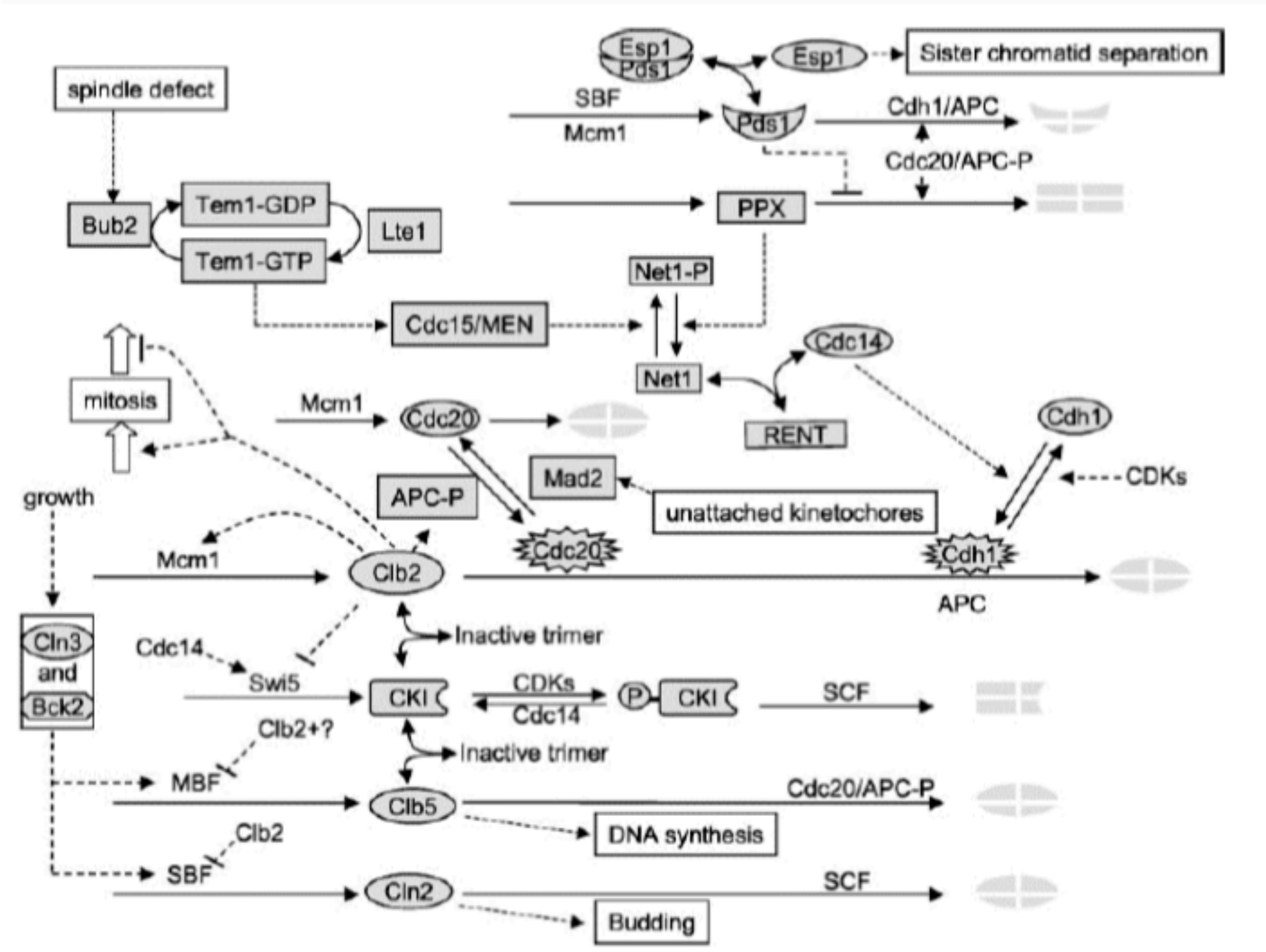
CellCollective

Typy modelů



Brane calculus

Typy modelů



$$\frac{d[\text{mass}]}{dt} = k_g \cdot [\text{mass}]$$

$$\frac{d[\text{Cln2}]}{dt} = (k'_{s2} + k''_{s2} \cdot [\text{SBF}]) \cdot [\text{mass}] - k_{d2} \cdot [\text{Cln2}]$$

$$\frac{d[\text{Clb5}]}{dt} = (k'_{s5} + k''_{s5} \cdot [\text{MBF}]) \cdot [\text{mass}] + (k_{d2c1} \cdot [\text{C5P}] + k_{d5} \cdot [\text{C5}] + (k_{d5f} \cdot [\text{F5P}] + k_{d5} \cdot [\text{F5}]) - (V_{d5} + k_{m5} \cdot [\text{Sic1}] + k_{m5} \cdot [\text{Cdc6}]) \cdot [\text{Clb5}]$$

$$\frac{d[\text{Clb2}]}{dt} = (k'_{s2} + k''_{s2} \cdot [\text{Mcm1}]) \cdot [\text{mass}] + (k_{d2c1} \cdot [\text{C2P}] + k_{d2} \cdot [\text{C2}] + (k_{d2f} \cdot [\text{F2P}] + k_{d2} \cdot [\text{F2}]) - (V_{d2} + k_{m2} \cdot [\text{Sic1}] + k_{m2} \cdot [\text{Cdc6}]) \cdot [\text{Clb2}]$$

$$\frac{d[\text{Sic1}]}{dt} = (k'_{s1} + k''_{s1} \cdot [\text{Swi5}]) + (V_{d2} + k_{d2}) \cdot [\text{C2}] + (V_{d5} + k_{d5}) \cdot [\text{C5}] + k_{pp1} \cdot [\text{Cdc14}] \cdot [\text{Sic1P}] - (k_{m2} \cdot [\text{Clb2}] + k_{m5} \cdot [\text{Clb5}] + V_{kp1}) \cdot [\text{Sic1}]$$

$$\frac{d[\text{Sic1P}]}{dt} = V_{kp1} \cdot [\text{Sic1}] - (k_{pp1} \cdot [\text{Cdc14}] + k_{d1}) \cdot [\text{Sic1P}] + V_{d2} \cdot [\text{C2P}] + V_{d5} \cdot [\text{C5P}]$$

$$\frac{d[\text{C2}]}{dt} = k_{m2} \cdot [\text{Clb2}] \cdot [\text{Sic1}] + k_{pp1} \cdot [\text{Cdc14}] \cdot [\text{C2P}] - (k_{d2} + V_{d2} + V_{kp1}) \cdot [\text{C2}]$$

$$\frac{d[\text{C5}]}{dt} = k_{m5} \cdot [\text{Clb5}] \cdot [\text{Sic1}] + k_{pp1} \cdot [\text{Cdc14}] \cdot [\text{C5P}] - (k_{d5} + V_{d5} + V_{kp1}) \cdot [\text{C5}]$$

$$\frac{d[\text{C2P}]}{dt} = V_{kp1} \cdot [\text{C2}] - (k_{pp1} \cdot [\text{Cdc14}] + k_{d2c1} + V_{d2}) \cdot [\text{C2P}]$$

$$\frac{d[\text{C5P}]}{dt} = V_{kp1} \cdot [\text{C5}] - (k_{pp1} \cdot [\text{Cdc14}] + k_{d5c1} + V_{d5}) \cdot [\text{C5P}]$$

$$\frac{d[\text{Cdc6}]}{dt} = (k'_{s6} + k''_{s6} \cdot [\text{Swi5}] + k''_{s6} \cdot [\text{SBF}]) + (V_{d2} + k_{d2}) \cdot [\text{F2}] + (V_{d5} + k_{d5}) \cdot [\text{F5}] + k_{pp1} \cdot [\text{Cdc14}] \cdot [\text{Cdc6P}] - (k_{m2} \cdot [\text{Clb2}] + k_{m5} \cdot [\text{Clb5}] + V_{kp1}) \cdot [\text{Cdc6}]$$

$$\frac{d[\text{Cdc6P}]}{dt} = V_{kp1} \cdot [\text{Cdc6}] - (k_{pp1} \cdot [\text{Cdc14}] + k_{d6}) \cdot [\text{Cdc6P}] + V_{d2} \cdot [\text{F2P}] + V_{d5} \cdot [\text{F5P}]$$

$$\frac{d[\text{F2}]}{dt} = k_{m2} \cdot [\text{Clb2}] \cdot [\text{Cdc6}] + k_{pp1} \cdot [\text{Cdc14}] \cdot [\text{F2P}] - (k_{d2} + V_{d2} + V_{kp1}) \cdot [\text{F2}]$$

$$\frac{d[\text{F5}]}{dt} = k_{m5} \cdot [\text{Clb5}] \cdot [\text{Cdc6}] + k_{pp1} \cdot [\text{Cdc14}] \cdot [\text{F5P}] - (k_{d5} + V_{d5} + V_{kp1}) \cdot [\text{F5}]$$

$$\frac{d[\text{F2P}]}{dt} = V_{kp1} \cdot [\text{F2}] - (k_{pp1} \cdot [\text{Cdc14}] + k_{d2f} + V_{d2}) \cdot [\text{F2P}]$$

$$\frac{d[\text{F5P}]}{dt} = V_{kp1} \cdot [\text{F5}] - (k_{pp1} \cdot [\text{Cdc14}] + k_{d5f} + V_{d5}) \cdot [\text{F5P}]$$

$$\frac{d[\text{Swi5}]_f}{dt} = k_{dsw1} + k'_{dsw1} \cdot [\text{Mcm1}] - k_{dsw1} \cdot [\text{Swi5}]_f$$

$$\frac{d[\text{Swi5}]}{dt} = k_{dsw1} + k'_{dsw1} \cdot [\text{Mcm1}] + k_{dsw1} \cdot [\text{Cdc14}] \cdot ([\text{Swi5}]_f - [\text{Swi5}]) - (k_{dsw1} + k_{m1} \cdot [\text{Clb2}]) \cdot [\text{Swi5}]$$

$$\frac{d[\text{APC-P}]}{dt} = \frac{k_{dapc} \cdot [\text{Clb2}] \cdot (1 - [\text{APC-P}])}{J_{dapc} + 1 - [\text{APC-P}]} - \frac{k_{dapc} \cdot [\text{APC-P}]}{J_{dapc} + [\text{APC-P}]}$$

$$\frac{d[\text{Cdc20}]_f}{dt} = k'_{d20} + k''_{d20} \cdot [\text{Mcm1}] - k_{d20} \cdot [\text{Cdc20}]_f$$

$$\frac{d[\text{Cdc20}]_a}{dt} = (k'_{d20} + k''_{d20} \cdot [\text{APC-P}]) \cdot ([\text{Cdc20}]_f - [\text{Cdc20}]_a) - (k_{mad2} + k_{d20}) \cdot [\text{Cdc20}]_a$$

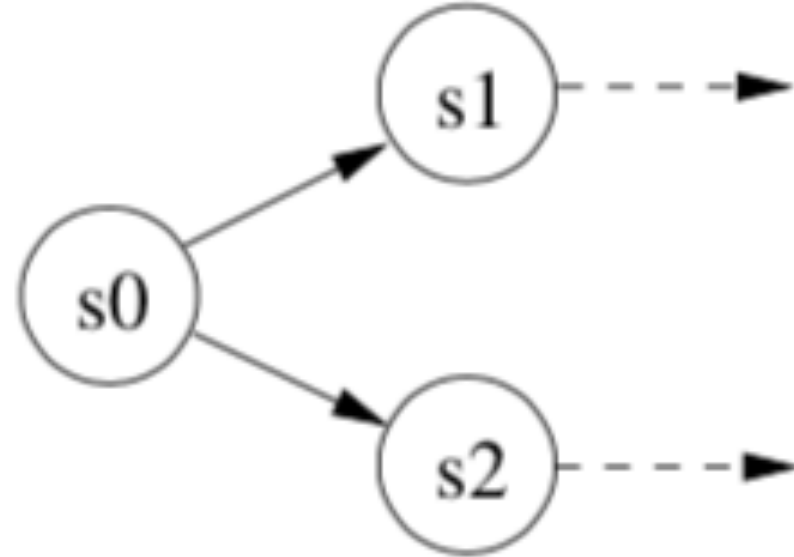
ODE

Copasi, Matlab

Typy modelů

state-transition systems

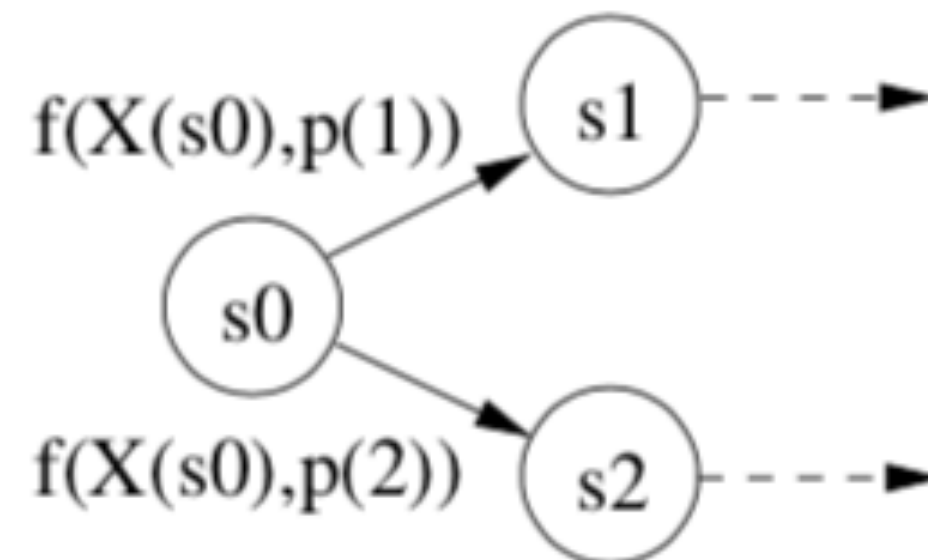
states: discrete molecule numbers or qualities (on/off)



qualitative model

stochastic model

Continuous-Time Markov Chains
states: discrete molecule numbers



continuous model

Ordinary Differential Equations
states: continuous concentrations

$$\dot{x} = f(x, p)$$

abstracted
time
modeled

states
← discrete continuous →

Osnova I. části (SB)

- Kvalitativní modely
 - Petriho síť
 - Boolovské síť
- Kvantitativní modely
 - ODE
 - stochastické Petriho síť