

Integracija modelov metabolizma in genskega regulatornega omrežja

Poskusi, problemi in predlogi

Jernej Puc

IŠRM – UNI-LJ

Marec 2019

Obravnavana je integracija modelov metabolizma, ki povzema stehiometrijo okoli sinteze holesterola v mišjih jetrih, in razširjen model sesalske cirkadiane ure, ki je predstavljen s sistemom navadnih diferencialnih enačb.

Namen združevanja je omogočiti, da modela vplivata drug na drugega in tako skupaj prikažeta bolj realistično sliko opazovanega sistema, kot če delujeta ločeno.

Cirkadiana ura:

▶ Primer biokemijskega oscilatorja:

- Koncentracije sestavnih snovi nihajo s periodo, ki v konst. pogojih znaša približno 24.2 ure.
- Zunanji dražljaji (npr. svetloba) imajo vlogo pri ponastavitvi ure.

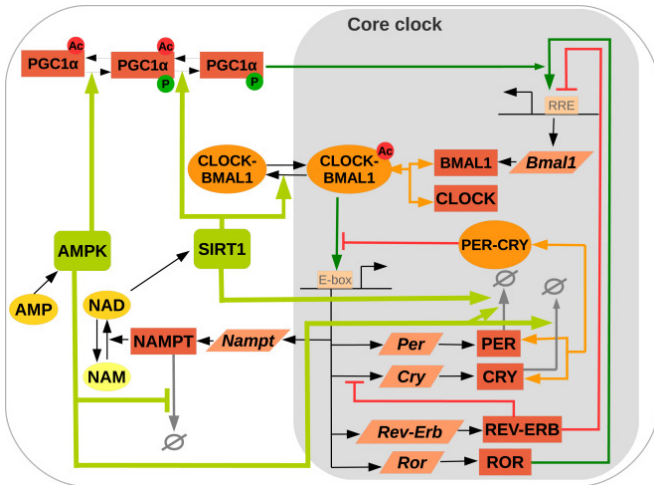
▶ Oscilatorji so lokalni:

- Glavna ura v možganih s signaliziranjem regulira in usklajuje periferne ure po drugih tkivih v telesu.

▶ Ima velik vpliv pri regulaciji bioloških procesov:

- Pri sesalcih naj bi bila transkripcija vsaj 43% genov odzivna na cirkadiane ritme.

Kontekst – GRO



Slika: Primer modela sesalske cirkadiane ure so zgradili Woller et al. (2016). Gre za ODE model, ki se osredotoča na vpliv dveh metabolitov (NAD in AMP), ki igrata vlogo v energijski pretvorbi.

Kontekst – razširjeni model cirkadiane ure

V nadaljevanju je izpostavljen model Magdevske (2019), ki se v svojem delu osredotoča na vpliv cirkadianih genov Ror in Rev-Erb na (bio)sintezo holesterola.

V ta namen izpelje sistem navadnih diferencialnih enačb, ki enosmerno povezuje modela Woller et al., 2016, in Pool et al., 2018 – ki v svojem delu s sistemom ODE modelirajo del ti. mevalonatne poti, tekom katere poteka največji delež sinteze holesterola – ter model nadalje poenostavi.

Tako razširjen ODE model omogoča več možnosti povezovanja z modelom metabolnega omrežja.

Kontekst – pomisleki o razširjenem modelu cirkadiane ure

- ▶ Nejasnosti v izvornih modelih:
 - Ni razloženo, od kod so bili pridobljeni nekateri parametri.
 - Nekatere zveze so izpeljane pomanjkljivo oz. niso izpeljane.
 - Začetne vrednosti, ki so ključne pri reševanju sistemov ODE, bodisi niso podane bodisi so postavljene na 0 (kar nima biološke osnove).
- ▶ Razširjeni model vključuje precej predpostavk:
 - *Predpostavimo, da traja 60 minut, da se translacija proteina INSIG začne.*
 - *Predpostavimo, da je v celicah insulina dovolj, da je razpolovni čas mRNA za INSIG štiri ure.*
 - *Predpostavimo, da je disociacijska konstanta za ROR enaka disociacijski konstanti za REV-ERB.*
 - *Predpostavimo, da je vrednost disociacijske konstante vezave SREBP na INSIG s holesterolom enaka vrednosti disociacijske konstante vezave holesterola in INSIG.*
 - *Empirično smo ugotovili, da je primerna vrednost ranga [...], izbrali smo vrednost [...].*
 - *Ena izmed vrednosti, pri kateri smo opazili, da je delovanje ustrezno, je [...].*

Kontekst – popravki razširjenega modela cirkadiane ure

Po prvih poskusih zagona se je izkazalo, da že znotraj razširjenega ODE modela samega povratni vpliv holesterola nima očitnega izraza.

Variiranje reda parametrov, ki določata hitrost sinteze in degradacije holesterola, je nakazovalo, da sta bili njuni vrednosti v modelu neustrezno določeni.

Po preverjanju poglavja **5.4** o določitvi parametrov v delu Magdevske in parametrizacije v delu Pool et al. so bili narejeni naslednji popravki:

5.4.1 $\mu_{11} : 5.83 \rightarrow 5.97$

5.4.2 $\mu_{12} : 1.6 \cdot 10^{-8} \rightarrow 1.95 \cdot 10^{-4}$

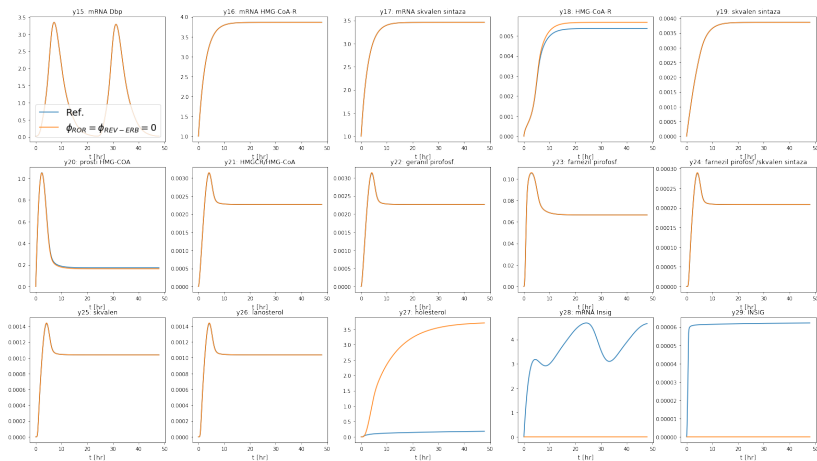
5.4.3 $\delta_{13} : 0.80 \rightarrow 0.40$

5.4.5 $K_{13} : 2.84 \cdot 10^{-2} \rightarrow 2.73 \cdot 10^{-1}$

5.4.6 $K_{12} : 2.84 \cdot 10^{-2} \rightarrow 2.73 \cdot 10^{-1}$

5.4.7 $g_{i0} : 1.29 \rightarrow 0.45$

Kontekst – razširjeni model cirkadiane ure



Slika: Kljub popravkom Poolov del (spr. y16 do y29) razširjenega modela izkazuje le delno odzivnost na spremembe v cirkadianem delu (y0-y15). Prikazani so simulirani poteki koncentracij v obdobju 48 ur pri aktualnih nastavitvah in ob izničenju povezav ϕ_{ROR} in $\phi_{REV-ERB}$.

Iz javno dostopnega GEM-a, ki opisuje metabolizem v mišjih jetrih, se je izoliral manjši model s poudarkom na sintezi holesterola.

Bolj natančno: iz celotnega modela se je vzelo skupino podsistemov “*cholesterol biosynthesis*” (vključno s “*Kandustch-Russel pathway*” in “*Bloch pathway*”) ter podsistem “*cholesterol metabolism*”.

V manjšem modelu so bili poleg tega zanemarjeni esencialni metaboliti, za katere je bilo predpostavljeno, da so dovolj razpoložljivi oz. da sami ne omejujejo pretokov v omrežju (npr. NAD).

Kontekst – GEM

Cirkadiani ritmi vplivajo na izražanje HMG-CoA-reduktaze (nadalje: HMGCR), ki vpliva na sintezo holesterola preko kompleksa HMGCR/HMG-CoA (HMG-CoA, na katerega je vezana HMGCR).

Ta zveza je sicer modelirana v razširjenem ODE modelu, vendar izsek GEM-a kompleksa HMGCR/HMG-CoA ne vključuje. Namesto tega opisuje reakcijo, ki jo HMGCR katalizira, in katere produkt je mevalonatna kislina oz. mevalonat.

Z namenom začetka s FBA pri mevalonatu se je v GEM-u zamenjalo vhodne/izhodne procese (transportne in izmenjevalne “reakcije”) z vhodnim procesom za mevalonat ter izhodom za holesterol.

Končni GEM je tako vseboval 56 metabolitov in 62 reakcij, optimizacijski kriterij v FBA pa je bil določen s pretoki reakcij, ki neposredno sintetizirajo holesterol.

Integracija – možne stične točke

Število skupnih metabolitov omejuje možne interakcije med modeloma.

Razširjeni model cirkadiane ure:

- ▶ ima zunanji vhod v AMPK (povezan s hranjenjem organizma),
- ▶ omogoča povezavo preko modeliranja substrata (HMG-CoA) in encima (HMGCR) v reakciji sinteze mevalonata,
- ▶ omogoča povezavo preko **vpliva Ror agonistov** (v GEM-u: zimosterol, dezmosterol, lanosterol, provitamin D3),
- ▶ omogoča povezavo preko **vpliva sirtuina na acetil-CoA**, ki je predhodnik HMG-CoA (čeprav manjka v izseku GEM-a).

Sicer so v obeh modelih prisotni naslednji metaboliti:

- HMG-CoA
- geranil pirofosfat
- farnezil pirofosfat
- skvalen
- lanosterol
- holesterol
- NAD (v okviru večjega GEM-a)

Integracija – problem različnosti modelov

V FBA se lahko določi omejitev pretoka z nastavitvijo zgornje in spodnje meje na isto vrednost glede na predhodne rešitve sistema ODE.

Problem pri združevanju modelov leži v tem, da ODE model razpolaga s koncentracijami, FBA pa z metabolnimi pretoki – povezave med temi količinami niso nujno očitne.

Integracija – problem različnosti modelov

Nielsen (2005) definira metabolni pretok kot funkcijo:

- ▶ koncentracije encima (posledica genskega izražanja),
- ▶ lastnosti encima (dovzetnost za substrate),
- ▶ koncentracije drugih metabolitov, ki v reakciji bodisi sodelujejo kot reaktanti ali produkti bodisi drugače vplivajo na aktivnost encima.

oz. formalno, če se privzame, da so lastnosti encima v opazovanem okolju konstantne, je pretok i -te reakcije v_i , v kateri sodelujejo metaboliti, encimi idr. s koncentracijami $[M_1], \dots, [M_n]$ po zgornjem opisu:

$$v_i(t) = f([M_1], \dots, [M_n], t)$$

Če je pretok v_i funkcija več koncentracij, lahko različne kombinacije koncentracij vodijo do istega pretoka (funkcija torej nima inverza). Na podlagi te definicije se tako pojavi implikacija, da se ne more vedno sklepati, da se lahko isto pretvorbo uporabi v obe smeri.

Integracija – koncentracija → pretok

Po **zakonu o delovanju mase** je hitrost kemijske reakcije premo sorazmerna s produktom koncentracij reaktantov oz. za produkt P , encim E in substrat S :

$$\frac{d[P]}{dt} = k_{\text{kat.}}[E][S]$$

Po **Michaelis-Mentenini predpostavki** sta v sorazmerju tudi produkt koncentracij in koncentracija kompleksa $[E][S] = \frac{1}{K_d}[ES]$. V primeru, da model simulira tako $[E]$ in $[S]$ kot $[ES]$, je pretok $v = \frac{d[P]}{dt}$ torej lahko določen bodisi s produktom ali kompleksom.

Opomba: Michaelis-Mentenina enačba, izpeljana iz

$$[ES] = \frac{[E]_0[S]}{K_d + [S]}$$

v dinamičnem primeru ni uporabna, saj se “začetna” koncentracija $[E]_0$ na vsakem časovnem koraku spreminja in velja $[E]_0 = [E] + [ES]$. Tako je enostavneje uporabiti neposredno sorazmerno zvezo.

Integracija – koncentracija → pretok

Za encim HMGCR, substrat HMG-CoA in kompleks pa HMGCR/HMG-CoA, je bil pretok, ki pritiče sintezni reakciji mevalonata $v_{\text{mev.}}$ določen z:

$$v_{\text{mev.}}(t) = \frac{[\text{HMGCR/HMG} - \text{CoA}](t)}{[\text{HMGCR/HMG} - \text{CoA}]^{\text{max}}} v_{\text{mev.}}^{\text{max}}$$

Pri tem sta vrednosti $[\text{HMGCR/HMG} - \text{CoA}]^{\text{max}}$ in $v_{\text{mev.}}^{\text{max}}$ največji, ki jih lahko zavzameta $[\text{HMGCR/HMG} - \text{CoA}]$ in $v_{\text{mev.}}$, ko sta modela obravnavana neodvisno drug od drugega (pri slednjem je uporabna FVA – flux variability analysis).

Integracija – pretok → koncentracija

Po predpostavki FBA so vsi “vmesni” pretoki (ki igrajo vlogo tako substratov oz. reaktantov kot produktov) v ravnovesju oz.

$$\frac{d\mathbf{P}}{dt} = \mathbf{N}\mathbf{v} = \mathbf{0}$$

V okviru teh metabolnih procesov se torej ne pričakuje spremembe v koncentracijah. Preostane “izhodni” pretok, v obravnavanem primeru pretok sinteze holesterola.

V bistvu pa se celo za tega ne more sklepati, da se ne izniči – model je preprosto pomanjkljiv. Če se iz izseka GEM-a ne bi odstranile nadaljnje reakcije, bi bilo privzeto, da je v ravnovesju tudi koncentracija holesterola.

Integracija – pretok → koncentracija

ODE model sicer simulira koncentracijo holesterola že sam, vendar ni očitno – če bi pretvorba iz pretoka v koncentracijo obstajala – da se jo lahko z rezultatom FBA enostavno preglasi.

Kot rečeno so iz GEM-a izpuščene reakcije, ki holesterol porabljajo, zato bi bilo ODE modelu prepuščeno vsaj “breme” simuliranja porabe.

Hkrati se rezultat FBA strogo nanaša na stanje, ko se koncentracije opazovanih metabolitov ne spreminjajo, in predpostavko, da so kemijske reakcije edine, ki k temu pripomorejo, zato se niti ne more sklepati, da so lahko isti pretoki uporabljeni tudi v dinamičnem scenariju.

Implementacija

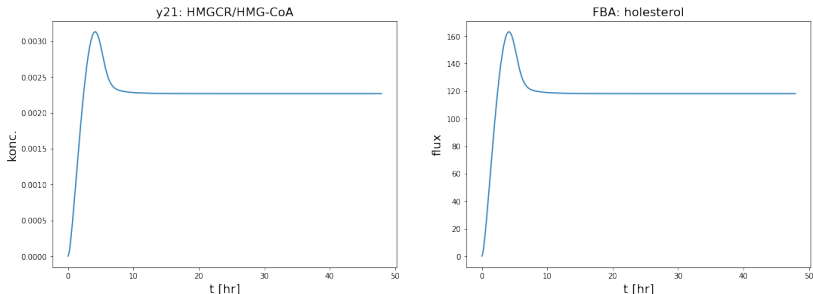
Zaradi ključnih predpostavk FBA se torej neizogibno izgubijo informacije, kar otežuje vpeljavo povratne informacije pri integraciji modelov.

Tako ostane enosmerna implementacija, ki je že uveljavljena in prepoznavna pod imenom **dFBA** oz. *dynamic FBA*.

Slednja je razširitev FBA na osnovi, da metabolizem deluje drugače pod različnimi pogoji, ki so podani v obliki časovnega zaporedja: na vsakem časovnem koraku se glede na rezultate ODE modela posodobijo meje pretokov, ponovno izvede FBA in vrne optimalno porazdelitev pretokov ob danem času.

Predpostavke FBA pri tem niso ovržene – nasprotno, poskrbeti je treba, da je časovni korak dovolj velik, da je predpostavkam zadoščeno (in dovolj majhen, da ni prevelike izgube v natančnosti numerične integracije pri reševanju sistema ODE).

Implementacija



Slika: Zaradi preprostosti povezave med modeloma je rezultat FBA linearna preslikava vhodne koncentracije. Časovni interval med posameznimi koraki je bil sicer nastavljen na 6 minut (desetina ure), začetne vrednosti večine opazovanih koncentracij pa na 0 z namenom, da bi bile karakteristike bolj očitne – v praksi bo koncentracija nekaterih metabolitov ostajala približno konstantna, kot je tudi razvidno iz konvergence opazovanih količin.

Očitno je, da integracijski model trenutno ni uporaben.

Tudi če bi bil rezultat FBA netrivialen, se lahko zaradi ene same povezave do tega pride tudi z ločeno obravnavo podmodelov:

1. ODE model simulira vpliv sprememb v cirkadiani uri na HMGCR/HMG-CoA.
2. Potek HMGCR/HMG-CoA določa časovno zaporedje mej pretoka za sintezo mevalonata.
3. FBA izpostavlja sintezo mevalonata kot ključni omejujoči element na poti do sinteze holesterola.

Predlogi – ODE \rightarrow FBA

Da bi bil model bolj relevanten, ga je potrebno razširiti oz. vsaj bolje izkoristiti možnosti povezovanja, npr.:

- ▶ Vključitev koncentracij intermediatov sinteze holesterola, ki so prisotni v obeh modelih, v določitev mej pretokov pripadajočih sinteznih reakcij.
- ▶ Vključitev NAD, ki oscilira v modelu cirkadiane ure, v določitev mej pretokov večjega števila reakcij pri večjem GEM-u, namesto da je iz tega povsem izvzet.
- ▶ Vključitev vpliva sirtuina na sintezo acetil-CoA (del večjega GEM-a).
- ▶ Dodatna razširitev GRO, ki bi vnesla nove priložnosti za povezave.

Predlogi – FBA → ODE

Ostaja problem, da model ne vsebuje legitimne povratne povezave – specifično, od metabolizma do same cirkadiane ure – kar bi se od integracijskega modela pričakovalo.

Priložnost je moč najti v vplivu intermediatov sinteze holesterola na Ror. Posebej zimosterol in dezmosterol sta bila izpostavljena kot izrazita agonista, lanosterol kot šibki agonist in provitamin D3 kot šibki inverzni agonist.

Pri tem je problem, da [vir](#) ne navaja jasne formulacije. V internem poskusu je bil parameter transkripcije za Ror določen sorazmerno z vrednostmi pretokov sinteze agonistov (z arbitrarnimi parametri), vendar se to ne izogne izpostavljenim pomislekom – korektna in ustrezna povezava ni bila dognana.

Validacija

Če bi bil integracijski model implementiran brez trdne teoretične osnove, bi bil lahko še vedno uporaben. Kot se je izrazil statistik George Box:

“Essentially, all models are wrong, but some are useful.”

Vseeno pa je pomanjkanje konkretnih eksperimentalnih podatkov velika ovira tudi pri preverjanju uporabnosti oz. validaciji modelov.

V literaturi se namreč najdejo opisi, kot so *“povečanje X vodi do povečanja Y”*, *“X follows a circadian pattern”* in neobrazloženi grafi, kar pušča veliko prostora za interpretacijo (in implementacijo).

- * Na ta način je bil validiran model Magdevske (*“opazimo, da koncentracija holesterola naraste ob izbitju gena za ROR in pade ob izbitju gene za REV-ERB, kar se sklada s pričakovanimi rezultati”*, medtem ko model Pool et al. niti ni bil validiran. Woller et al. so vpliv hranjenja na cirkadiano uro validirali na podlagi konkretnih podatkov.

Brez teoretične podlage in zadostne eksperimentalne potrditve pa je reprezentativnost pravega biološkega sistema torej brezpredmetno trditi.

Tako smiselna validacija kot že sama postavitev ustreznega modela postavljata večjo pristojnost relevantnemu biološkemu znanju kot računalniškim veščinam. Zato se zdi smiselno, da je “breme” proaktivnosti prepuščeno ustrezni stroki.

Zaključki

V retrospektivi so bile ovire že na daleč izpostavljene:

► **Imam et al., 2015:**

- Poudarek, da je potrebno natančno pridobivanje podatkov za posamezne modele.
- Izzivi v konstrukciji in validaciji integracijskih modelov.
- Omejitve pri združevanju.

► **Klipp et al., učb.:**

- *"The diversity of modeling approaches makes it also very difficult to merge established models (e.g., for individual metabolic pathways) into larger supermodels (e.g., models of complete cell metabolism)."*
- *"Models gain their reference to reality from comparison with experiments, and their benefits therefore depend on the quality of the experiments used."*

► Dejstvo, da se v literaturi ne pojavljajo navidez intuitivni pristopi.

Poleg tega je bil model grajen brez konkretne predstave o njegovi eventuelni uporabi – zanašalo se je na preverjanje hipotez, ki niso bile nikoli zastavljene, brez zagotovila, da bodo relevantni vidiki v času potrebe zares pokriti.

Jasnost končnega cilja pa bi se verjetno odražala tudi v bolj jasnem delu – raziskovanje na silo namreč ne zagotavlja samodejnih rezultatov.

- ▶ L. Magdevska. *Matematični model vpliva cirkadiane ure na biosintezo holesterola*. Magistrsko delo, Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, 2019.
- ▶ A. Woller, H. Duez, B. Staels, M. Lefranc, *A mathematical model of the liver circadian clock linking feeding and fasting cycles to clock function*. Cell Reports, vol. 17, pp. 1087–1097, 2016.
- ▶ F. Pool, R. Currie, P. K. Sweby, J. D. Salazar, M. J. Tindall. *A mathematical model of the mevalonate cholesterol biosynthesis pathway*. Journal of Theoretical Biology, vol. 443, pp. 157–176, 2018.
- ▶ Edda Klipp, Wolfram Liebermeister, Christoph Wierling, Axel Kowald. *Systems Biology: A Textbook*. John Wiley & Sons, 2016.