

An Introduction To Systems Biology
Design Principles of Biological Circuits
Uri Alon

Lukáš Kuhajda

2019/2020

Kapitola 1

Transkripční síť

- transkripce (přepis) - výroba RNA dle informace z DNA
- transkripční síť - interakce mezi transkripčními faktory a geny
- buňka - tisíce proteinů - každý svůj úkol - vysoká přesnost
 - sledování prostředí, produkce co je třeba v potřebném množství - prováděno **transkripční sítí**

1.1 Kognitivní problémy buňky

- kognitivní - poznávací, vjemový
- rozeznávání tepla, tlaku, signály z jiných buněk, prospěšné živiny, škodlivé látky..., vnitřního stavu (poškození DNA, proteinů, membrány)
- **transkripční faktory** - reprezentace stavu prostředí
 - pro rychlý přechod mezi aktivními a pasivními molekulovými stavy
 - aktivní trans.fak. se může vázat na DNA a regulovat rychlost čtení cílových genů
 - čtení genů do mRNA → přeložení do proteinu (**produkt genu**) → působení na prostředí

1.2 Prvky transkripční sítě

- gen - úsek DNA - sekvence kóduje informaci pro produkci proteinu
 - transkripce genu - **RNA polymeráza** tvoří mRNA odpovídající kódovací sekvenci genu
 - **promotor** - kontrola rychlosti přepsání genu, množství mRNA za jednotku času
 - regulační oblast DNA předcházející genům
 - vazba RNAP s daným místem v promotoru - kvalita místa úměrná s transk. mírou
 - RNAP - jednání se všemi geny, změny ale prováděny z jednotlivých genech transk. faktory
 - vazbou transk.fak. s místem v promotoru cílového genu dochází k ovlivnění trans.míry
 - ovlivnění kdy RNAP zahájí transkripci genu
- transkripční faktor
 - **aktivátor** - zvyšuje míru transkripce genu
 - **represor** - nepřepisování genu v RNA, redukce míry transkripce
- **transkripční síť** - proteiny transk.fak. kódovány geny, které jsou regulovány jinými trans.fak.
 - popisuje všechny regulační transkripční interakce v buňce
 - **nody** - uzly v grafu, geny
 - **hrany** - transkripční regulace genu produktem proteinu jiného genu
 - **signály** - vstup do sítě
 - molekuly, modifikace proteinu - přímé ovlivnění aktivitu transk.fak.

- často způsobení fyzické změny tvaru proteinu transkr.fak.
 - převzetí aktivního molekulového stavu
- signál → změna aktivity transkr.fak. → změna produkce proteinů
- některé proteiny jsou transkr.fak., které aktivují další geny
- zbytek vykonává různé funkce v buňce - budování struktury, katalyzace

1.2.1 Oddělení časových intervalů

- vázání transkr.fak. k dané části DNA v rámci sekund
- transkripce a translace cílového genu v rámci minut
- hromadění proteinového produktu v rámci minut až hodin
- úroveň aktivity trans.fak. považována za stabilní v rámci rovnic popisujících dynamiku pro pomalý časový rozvrh změn úrovně proteinů
- velká škála mechanismů, jak transkr.fak. regulují geny
- při spojení transkr.fak. s DNA se doplní o RNAp pro kontrolu produkce mRNA
- transkripční síť - modularita komponentů
 - možnost vzít DNA z genu z jednoho do druhého organismu
 - síť tvárná - evoluce, snadné začlenění nových genů a regulací
- GFP - green fluorescent protein - světélkování u medúz
- hrany v síti se jeví, že se vyvíjí v rychlejších intervalech než kódování částí genů
 - myši, lidi - podobné geny, rozdíl transkripční regulaci genů
 - rozdíl v druzích ne tolik v genech, ale v hranách transkr.sítě

1.2.2 Označení hran: aktivátory a represory

- aktivátor - pozitivní kontrola, vzrůst míry transkripce, node obsahuje více + hran
- represor - negativní kontrola, redukce míry transkripce, node obsahuje více - hran
- síť většinou obsahuje více pozitivních hran - 60-80%
- většina aktivátorů funguje za určitých podmínek jako represory pro dané cílové geny
 - platí i naopak
- duální transkr.fak - při splnění podmínky působí na gen jako aktivátor, při jiné jako pasivátor
- transkr.fak. mají tendenci použít stejný mód regulace pro většinu jejich cílových genů
 - označení hran jdoucích do nodu (transkripční interakce regulující gen) nejsou tak korelované
- označení hran odcházejících bývaá spíše korelované, než označení hran příchozích

1.2.3 Čísla na hranách: vstupní funkce

- označení síly interakce
- **vstupní funkce** - síla účinku transkr.fak. na míru transkripce cílového genu
- $X \rightarrow Y$... *rate of production* $Y = f(X^*)$ - míra produkce proteinu Y za jednotku času
 - X^* - aktivní forma koncentrace X - $f(X^*)$ obvykle monotóní, tvar S - rostoucí když X aktivátor a naopak
- Hillova funkce - popisuje mnoho vstupních funkcí genů
 - pro aktivátor

$$f(X^*) = \frac{\beta X^{*n}}{K^n + X^{*n}} \quad (1.1)$$

- K ... aktivační koeficient - udává se v koncentraci X nutnou pro aktivaci výrazu
 - souvisí se slučivostí mezi X a jeho umístěním na promotoru, + další faktory
- β - maximální hodnota výrazu, $X^* \gg K$
 - při vysoké koncentraci se X^* váže na promotors vysokou ppstí a stimuluje RNAP k produkci velkého množství mRNA za jednotku času
- n - **Hillův koeficient** - řídí šikmost vstupní funkce
 - obvykle 1-4, čím vyšší tím šikmější
- pro represor

$$f(X^*) = \frac{\beta}{1 + (\frac{X^*}{K})^n} \quad (1.2)$$

- každá hrana tedy může nést 3 čísla, změny při evoluci
- spousta genů má nenulový počáteční stav hodnoty výrazu - **základní hodnota výrazu** - β_0

1.2.4 Logické vstupní funkce: jednoduchý framwork k pochopení dynamiky sítě

- Hill dobrý pro detailní modely, z matematického hlediska lepší jednodušší funkce pro zachycení základního chování
- **logická aproximace** - častá aproximace vstupní funkce v transkripční síti, skoková
 - pro jednoduché grafické řešení dynamických rovnic
 - gen buď zapnutý (ON - $f(X^*) = \beta$) / vypnutý (OFF - $f(X^*) = 0$), aktivační práh = K
 - pro aktivátor

$$f(X^*) = \beta \theta(X^* > K) \quad (1.3)$$

$\theta \dots 0/1$ - pro represor obrácené logické znaménko

1.2.5 Vícerozměrové vstupní funkce řídící geny s více vstupy

- sousta nodů má více než 1 vstupní hranu
- aktivita promotoru je vícerozměrné vstupní funkce různých transk.fak. - můžou být aproximovány logickými funkcemi
- geny co pro aktivaci potřebují oba aktivátory proteinu vázané na promotor

$$f(X^*, Y^*) = \beta \theta(X^* > K_x) \theta(Y^* > K_y) \sim X^* AND \cdot Y^* \quad (1.4)$$

- alespoň jeden

$$f(X^*, Y^*) = \beta \theta(X^* > K_x OR \cdot Y^* > K_y) \sim X^* OR \cdot Y^* \quad (1.5)$$

- některé jako suma vstupů

$$f(X^*, Y^*) = \beta_x X^* + \beta_y Y^* \quad (1.6)$$

- jednoduchá změna funkční formy vstupní funkce mutací v promotoru regulovaného genu
 - změna z AND na OR během pár mutací promotoru v *lac* (E.coli)

1.2.6 Prozatímní shrnutí

- transkripční síť popisuje transkripční regulaci genů
- node reprezentuje gen, hrana z $X \rightarrow Y$ - gen X se kóduje pro trans.fak proteinu, který se váže na promotor genu Y, upravuje míru transkripce
 - protein kódovaný genem X mění míru produkce proteinu kódovaného genem Y, Z pak zas může být transk.fak. genu Z - vytváření sítě
- většina nodů značí geny, které nejsou transk.fak, ale nesou různé funkce buňky
- vstupy do sítě jsou informace o prostředí, mění příslušné transkr.fak.
- aktivní transkr.fak. se vážou na daná místa DNA promotoru cílového genu - regulace míry transkripce
 - míra produkce produktu genu Y je funkce koncentrace aktivního transkr.fak. X^*
 - geny regulované více transkr.fak. ají vícerozměrovou vstupní funkci - Hill, logické
- hrany a vstupní funkce jsou pod tlakem výběru, nevyužívaná hrana zanikne mutacemi
 - změna jednoho nebo pár písmen v oblasti DNA, kde se váže X v promotoru Y k zrušení hrany $X \rightarrow Y$

1.3 Dynamika a časová odezva jednoduché regulace genu

- 1 hrana sítě, gen regulovaný 1 regulátorem bez dalších vstupů (nebo jen konstanty)
- $X \rightarrow Y$ - transk.fak. X reguluje gen Y
- bez vstupu X neaktivní a Y není produkováno
- objevení signálu S_x , okamžitá přeměna X do aktivní formy X^* a navázání na promotor genu Y
 - \rightarrow přepis genu Y, překlad mRNA \rightarrow hromadění proteinu Y
 - buňka produkuje protein Y v konstantní míře - β - jednotky koncentrace za jednotku času
- vyvážení produkce proteinu
 - degradace** - specifická destrukce speciálními proteiny v buňce
 - míra degradace - α_{deg}
 - ředění** - redukce koncentrace v důsledku zvýšení objemu buněk během růstu
 - míra ředění - α_{dil}
 - míra degradace/ředění

$$\alpha = \alpha_{deg} + \alpha_{dil} \quad (1.7)$$

- změna koncentrace Y v čase

$$dY/dt = \beta - \alpha Y \quad (1.8)$$

- v ustáleném stavu, konstantní koncentrace Y_{st} , řešení $dY = 0$

$$Y_{st} = \beta/\alpha \quad (1.9)$$

- odstranění signálu, konec produkce ... $\beta = 0$

$$Y(t) = Y_{st}e^{-\alpha t} \quad (1.10)$$

- **časová odezva** - měří rychlost změny Y, $T_{1/2}$
 - čas k dosažení poloviny počáteční a koncovou úrovní v dynamické procesu
 - z předchozí rovnice z Y_{st} do $Y = 0 \rightarrow Y(t) = Y_{st}/2$

$$T_{1/2} = \log(2)/\alpha \quad (1.11)$$

- některé proteiny mají vysokou degradační míru α , je tedy nutná vysoká produkce β
 - cyklus produkce a destrukce - rychlá reakce, když je potřeba změny

- $Y = 0$, spuštění signálu, začátek hromadění Y , růst k $Y_{st} = \beta/\alpha$

$$Y(t) = Y_{st}(1 - e^{-\alpha t}) \quad (1.12)$$

1.3.1 Časová odezva stabilních proteinů je jedna generace buněk

- stabilní proteiny - $\alpha_{deg} = 0$, nejsou aktivně degradovány v rostoucích buňkách
 - produkce hlídána ředěním, $\alpha = \alpha_{dil}$
- produkce proteinu, náhlý konec ($\beta = 0$), růst buňky do dvojnásobku, rozdělení ve dvě po **jednogenečním čase** τ - koncentrace klesá na 50%

$$T_{1/2} = \log(2)/\alpha_{dil} = \tau \quad (1.13)$$

- tato časová odezva může být limitujícím faktorem, který představuje omezení pro tvorbu efektivních genetických obvodů

Kapitola 2

Autoregulace: síťové motivy

- cíle

- **síťový motiv** - definovat vzor stavebních bloků v síti
- zkoumat nejjednodušší motiv v transkripční síti - **negativní autoregulace**
- ukázat, že tento motiv má užitečnou funkci - urychlení odezvy transk.interakce a stabilizace

2.1 Vzory, náhodně vytvořené sítě, síťové motivy

- E.coli - mnoho vzorů nodů a hran, hledání důležitých vzorů - statistika
 - porovnávání se souborem **náhodných sítí**
 - sítě se stejným počtem nodů a hran jako reálné, náhodné spojení mezi nody a hranami
- **síťové motivy** - vzory opakující se výrazně častěji než v náhodných sítích
 - motivy odolávající mutacím, které náhodně upravují hrany

2.1.1 Detekce motivů srovnáváním s náhodnými sítěmi

- ER model - Erdos and Renyi - nejjednodušší soubor náhodných sítí
- pro hodnotné srovnání je třeba náhodných sítí se základními rysy reálných sítí
 - stejný počet hran (edge - E) a nodů (N) jako reálná, náhodně vytvořené spoje mezi nody
 - $N(N-1)/2$ možných párů nodů, možné oba směry $\rightarrow N(N-1)$
 - hrana může do stejného nodu, ze kterého vycházela - N

počet možných hran:

$$N(N-1) + N = N^2 \quad (2.1)$$

- pravděpodobnost hrany $p = E/N^2$

2.2 Autoregulace: motiv sítě

- srovnávání transkripční sítě E.coli s náhodnými
- **vlastní hrany** - vrací se do stejného nodu, ze kterého vyšly - 40
 - transkr.fak. regulující transkripci vlastního genu
 - autogenní kontrola, **autoregulace**
 - 34 jsou zde represory - potlačení vlastní transkripce - **negativní autoregulace**
 - statisticky by v náhodně vytvořené síti měla být pouze 1 ± 1 vlastní hrana
 - vypočtená odpovídající odchylka $Z \sim 32$ je statisticky velmi významná

- vlastní hrany - autoregulace - je motiv sítě

2.2.1 Urychlení odezvy obvodu genu negativní autoregulací

- transkripční faktor X se naváže na vlastní promotor k potlačení produkce mRNA
- čím vyšší koncentrace, tím nižší míra produkce
- dynamika popsána mírou produkce $f(X)$ a mírou degradace/ředění - pro aproximaci využít Hill - X na začátku není, $t=0$
 - rychlá produkce, dosažení vysokého ustáleného stavu, konec produkce na $X=K$ (koef. represe)

$$T_{1/2} = \frac{K}{2\beta} \quad (2.2)$$

- čím silnější nerepresovaná aktivita β , tím rychlejší odezva
- negativní autoregulací rychlé nabytí počáteční produkce, autorepresí ukončení produkce
- evolucí jednoduché nastavení parametrů β a K nezávisle na sobě
 - K - například mutací v místě vázání X v promotoru
 - β - mutacemi v místě vázání RNAP v promotoru
- dosažení výrazně vyšší rychlosti odezvy než v případě bez negativní autoregulace
- negativní autoregulace
 - silný promotor dá rychlou produkci
 - vhodný koeficient represe poskytuje žádoucí ustáleného stav
 - v jednoduché regulaci dosažení značně vyššího ustáleného stavu, potom nežádoucí nadměrná exprese produktu genu

2.2.2 Negativní autoregulace podporuje robusnost v kolísání míry produkce

- další výhodou negativní autoregulace zvýšení robusnosti expresní úrovně ustáleného stavu ke kolísání produkční míry β
- kolísání metabolické kapacity v buňce, regulačního systému, náhodné efekty v produkci
 - buněčná dvojčata rozdílné míry produkce β u většiny proteinů
 - rozdíly i po dobu celé generace
 - represní práh K se většinou moc neliší
- lineární závislost regulace na kolísání produkční míry β - $X_{st} = \beta/\alpha$ - u negativní autoregulace závisí hodnota ustáleného stavu pouze na represním práhu K - $X_{st} = K$
 - K se v porovnání buňka-buňka tolik neliší → zvýšení robusnosti ustáleného stavu z hlediska kolísání míry produkce

2.2.3 Pozitivní regulace zpomaluje odezvu a může vést ke dvojí stabilitě

- aktivace vlastní transkripce, 10% transkripční faktor v E.coli, pomalá dynamika
- s růstem úrovně X roste míra produkce
- dosahuje polovičního zpoždění ve srovnání s jednoduchou regulací
- vhodné pro procesy trvající dlouhou dobu - vývojové procesy
- dvojí stabilita - míra pozitivní autoregulace je silná v porovnání míry degradace/ředění
- při aktivaci genu zůstane aktivní i po zmizení vstupního signálu
- využití ve vývojových transkripčních sítích, kde je potřeba stanovit osud buňky

Kapitola 3

Motiv sítě dopředné smyčky

- mnoho vzorů, jen ty nalézané významně jsou motivy
- motivy mají definovanou funkčnost zpracování informace
 - jejich benefity vysvětlují, proč evolucí znikají v různých organismech
- zde zaměření na motivy se 3 nody
 - 13 možností, pouze jedna je motiv sítě (feed-forward loop FFL)
 - důležité pochopit regulace jednotlivých hran - každá může být buď aktivátor nebo represor
 - 8 FFL typů, poze 2 se vyskytují častěji v sítích
 - zde zabývání se dynamikou, běžné typy filtrují šum, pulsy a odezvové zrychlení

3.1 Výskyt subgrafu v náhodných sítích

- **subgraf** - vzor z více nodů
- **dopředná smyčka** - 3 nody, nevrací se do počátečního
- výskyt trojúhelníkových subgrafů je náhodných sítích velice vzácný
- při výskytu v reálné síti lze považovat za motiv

3.2 Dopředná smyčka je motiv sítě

- v E.coli 42 dopředných smyček, žádná se zpětnou vazbou (3 nody spojeny dokola, často jako anti-motiv)
 - FFL je silný motiv, mnohem častěji než v náhodných
 - FFL jediný motiv ze všech 13 3-nodových možností

3.3 Struktura dopředné smyčky genetického obvodu

- transkr.fak. X reguluje transkr.fak. Y a oba regulují transkr.fak. Z, 2 paralelní regulační cesty, 8 typů
 - **koherentní** - nepřímá cesta má stejné celkové znaménko jako přímá cesta
 - **nekoherentní** - nepřímá cesta má celkově opačné znaménko co přímá
- nejčastěji C1-FFL (coherent 1) - všechny aktivátory - $+ / ++$
- 2. nejčastější je I1-FFL (incoherent 1) - $+ / +-$
- všechny ostatní se objevují výrazně méně často
- integrace X a Y do promotoru Z - aktivace - AND/OR
- X a Y často reagují na vnější podněty - např. molekuly vázající se na transkr.fak. - podnět z venku

většinou zpracováván výrazně rychleji než vnitřní transkripční interakce v FFL

- příchod S_x , aktivace $X \rightarrow X^*$, navázání na dané místo DNA promotoru genu Y a Z během pár s, také změna míry transkripce \rightarrow koncentrace Z se mění v intervalu od minut do hodin

3.4 Dynamika C1-FFL s logikou AND

- signál S_x , aktivace na X^* (**kroková stimulace** X), navázání k promotoru Y, produkce proteinu Y \rightarrow druhý transrk.fak. v FFL
- paralelně se X^* váže na promotor genu Z - logika AND - nestačí k aktivaci
 - aktivace po překročení koncentrace Y **aktivačního práhu** genu Z - K_{yz}
 - aktivace vyžaduje i 2. vstupní signál S_y - Y v aktivní formě Y^*
- příchod S_x , čekání na akumulaci Y pro aktivaci Z - **zpoždění**

3.5 C1-FFL je prvek zpoždění citlivý na znaménka

- zpoždění s ON krokem S_x , ne při OFF kroku - zpoždění citlivé na znaménko
- může být bráno jako druh asymetrického filtru
 - signál S_x s kratší dobou trvání než odezva nijak dál nepůsobí na Z - **detektor trvání** na ON
 - okamžitá odezva na OFF pulsy
- inženýrské použití - když náklady na chybu jsou nesymetrické
- transkripční sítě - spíše ochranná funkce, filtrace kolísání
- případ OR - odpadá zpoždění při ON, vzniká při OFF - buzení Z i když zrovna vypadnul signál
 - zpoždění cca o hodinu

3.6 I1-FFL

- X aktivátorem Y i Z, Y represorem Z
- $S_x \rightarrow$ aktivace na X^* , vazba k Z, inicializace transkripce - produkce proteinu
 - paralelně aktivace produkce Y, nahromadění nad práh, tlumení Z
- I1-FFL tedy může generovat pulsy produkce Z
- urychlení odezvy - nejdříve vysoká produkční míra, se zpožděním represor na snížení produkční míry
 - udržení žádoucí úrovně ustáleného stavu
- při zmizení signálu S_x žádná akcelerace nebo zpoždění - okamžité ukončení produkce Z
 - kvůli AND logice promotoru Z
 - s logikou OR funguje stejně, akorát je odezva při OFF a ne při ON
 - po ukončení exponenciální pokles Z podle míry degradace/ředění
 - I1-FFL je prvek zpoždění citlivý na znaménka

3.7 Proč jsou jiné FFL typy tak vzácné?

- některé postrádají citlivost k jednomu z jejich dvou vstupů
- nepřítomnost signálu S_y u I1-FFL - Z není represován, velký vliv na úroveň ustáleného stavu Z
- I4-FFL - $+/-+$ - cca 5% z FFL
 - urychlovač s citlivostí na znaménko i pulsní generátor
 - narozdíl od I1-FFL nezávisí na S_y , při nepřítomnosti S_x je nesplněna logika AND z X do Z

- nepřítomnost citlivosti vůči jednomu vstupu je důvod, proč se nevyskytuje moc často
 - stejný důvod i pro I3-FFL
 - podobný důvod i u C3-FFL a C4-FFL
 - u ostatních případů a OR formy je složitější vysvětlování

3.8 Konvergentní evoluce FFLs

- základní V-tvar kdy X a Y regulují Z, od jedné do pár mutací se může přidat i regulace $X \rightarrow Y$
 - pokud nepomůže nebo ještě zhorší - rychle evolucí odstraněna
 - u C1-FFL a I1-FFL je funkce jasná
- **homologní geny** - geny z jednoho společného, značný stupeň podobnosti
 - možný podobný vznik u FFLs - většinou však ne
- vznik spíše nezávislý na stejných obvodech ve spoustě případech

Kapitola 4

Časové programy a globální struktura transkripčních sítí

- zde kompletace průzkumu smyslových motivů sítě - předtím autoregulace a dopředné smyčky
- další 2 motivy
 - SIM - single-input module
 - 1 regulátor kontroluje skupinu genů
 - vytváří časové programy exprese - geny jsou zapnuty postupně jeden po druhém ve stanoveném pořadí
 - "just-when-needed" strategy - není produkce proteinu, dokud není potřeba
 - optimální pro rychlou produkci systému, který se skládá z různých typů proteinů
 - s omezeným zdrojem pro produkci těchto proteinů
 - DORs - dense overlapping regulons - hustě se překrývající regulony
 - husté pole regulátorů kombinatoricky kontrolující výstupní geny
 - můžou nést výpočty důležité pro rozhodování
- nakonec, jak motivy zapadají do sebe, aby vytvořily síť

4.1 Síťový motiv SIM

- větší motiv - více vzorů pohromadě s typickým vzorem architektury - 1. rodina motivů je SIM
- hlavní transk.fak. X kontroluje skupinu cílových genů - každý jeden vstup
 - znaménko (aktivace/represe) pro všechny kontrolované nody stejné
- X bývá často autoregulované
- silný motiv oproti náhodným sítím
- geny mají běžné biologické funkce - geny se účastní v nějaké metabolické cestě
- geny pracují sekvenčně pro nashromáždění žádoucí molekuly atom po atomu
- jiné kontrolují geny reagující na stresové podněty - poškození DNA, teplotní šok...
 - produkce proteinů opravujících škody
 - podskupiny genů specializovaných na dané aspekty

4.2 SIM umí generovat časové programy exprese

- aktivace genů v daném pořadí
- různé práhy X pro každý cílový gen
 - práh pro každý promotor závisí na místě jeho vázání s X v jeho promotoru
- postupná aktivace genů, při úbytku zas postupná deaktivace z druhé strany - LIFO

- v E.Coli systém argininu, SOS DNA poničení - zpoždění v pořadí o 0.1 generaci genu (5-10 min)
- ekonomický design - produkce až ve chvíli potřeby
- přesné **časový řád** může být pozměněn mutacemi
 - při opravách rychlá produkce opravných prostředků, po ukončení různě rychlý pokles
 - geny řešící nejmenší problémy se vypnou první, ty pro rozsáhlejší se vypínají později
 - řeší celou řadu globálních buněčných odezev - geny časované podle buněčného cyklu v bakterii
 - geny často regulované hlavním regulátorem + přídatnými regulátory pro subsystémy
 - nemusí být tedy SIM
 - více regulátorů, jeden má konstatní aktivitu během sledovaného intervalu
- vznik SIM evoluční konvergencí ke stejnému regulačnímu vzoru u různých organismů - jako u FFL
- uděluje se vůči mutacím, protože je dostatečně užitečný
- někdy proces aktivace-deaktivace jako FIFO
 - některé části procesu potřebné dříve jiné později, ale jen jednou

4.3 Topologická generalizace síťových motivů

- zatím jednoduché - už 4 nody 199 možných vzorů, přes 9000 pro 5
- **topologická generalizace motivů** - rozdělení motivů do rodin dle jejich funkcionality
 - když je FFL motiv, každý z těch 3 vzorů může být motiv - pouze 1 je ale opravdu motiv
 - **multi-output FFL** - může generovat časový FIFO program

4.4 multi-output FFL umí generovat časové FIFO pořadí

- genetický systém kontrolující produkci flagel (bičků) - mototry E.coli
- když je buňka fajn, nikam se nepřemisťuje, když špatně nechá si narůst motýrky, vygeneruje navigační systém
- bičík cca 50nm, z 30 typů proteinů, 10x delší než tělo, 30 mikronů/s
 - skládá se dohromady po fázích - lego
 - uvnitř kanálek, kterým se další proteiny dostávají dál
 - informace o proteinech kódovaná v 6 operonech
 - skupina genů transkribována stejným kusem mRNA
 - operony regulovány 2 transkr.fak. - aktivátory
 - hlavní aktivátor X aktivuje Y, oba aktivují každý ze 6 operonů
 - **multi-output FFL**
 - každý operon může být aktivován pouze pomocí X nebo Y - podobné OR branám
 - just-when-needed produkce
 - + oproti SIM - FIFO - X najede, postupně překročí hranice aktivace promotorů
 - klesáním zpět by bylo opět LIFO - X klesne, ale je furt Y
 - Y má vlastní hranice pro každý gen - jiné pořadí na OFF než na ON
 - dále všechny funkce FFL - vytváření zpoždění při vypínání Z
 - deaktivace po určitém čase - detektor zpoždění pro každý výstup
 - při krátké ztrátě X se díky OR nemusí přerušit vstup
 - multi-output je generalizace FFL, která se vyskytuje nejčastěji v transkr. sítích

4.5 Integrace signálu a kombinatorická kontrola: bi-fan a hustě se překrývající regulony

- 4-nodové vzory - ze 199 možností jsou známy 2 motivy - two-output FFL a **bi-fan**
bi-fan - X_1 a X_2 oba regulují Z_1 Z_2
 - patří do skupiny **hustě se překrývajících regulonů - DORs**
- DOR - řada transkr.fak. regulujících sadu výstupních genů v hustě se překrývající směru
 - většinou ne všechny vstupy regulují všechny výstupy
 - kombinatorický rozhodovací systém
 - E.coli a kvasnice pár DORs - regulace několika set genů
 - sdílená společná globální funkce - reakce na stres, metabolismus živin...

4.6 Síťové motivy a globální struktury smyslových transkripčních sítí

- autoregulace, FFL, SIM, DOR - jak jsou vůči sobě, jak se překrývají?
- procedura hrubého zrna - substituce motivů za tvary - zjednodušení grafického znázornění sítě
- všechny geny jsou pokryty nějakým z rodiny motivů
- poměrně vzácné dlouhé kaskády ve smyslových transkr.sítích - časové nároky
 - jiné biologické sítě je často obsahují

Kapitola 5

Síťové motivy ve vývojových, transdukčních a neronových sítích

- vše předtím ve smyslových sítích - potřeba rychlých reakcí
- byly tvořeny malými sadami těchto motivů
- tedy - síťové motivy v jiných typech bio sítí
- **vývojová transkripční síť** - transkr. síť řídící osudy buněk - z vajíčka do mnohobuněčného org
 - difference buňky na jiný typ buňky
 - od smyslových rozdíl v časových intervalech a reverzibilitě
 - pomalejší nevratná rozhodování - vznik nových motivů sítě
- kromě transkr. sítí používá buňka další interakční sítě
 - interakční síť protein-protein, signálová transdukční síť, metabolické sítě
 - graficky různé barvy hran pro rozeznání různých sítí
 - rozdíly v časových intervalech
 - transkr.sítě v hodinách, signálová transdukční síť funguje v sekundách až minutách

5.1 Síťové motivy ve vývojových transkripčních sítích

- **senzorová transkr.síť** - reakce na vnější změny, skoro ve všech buňkách
- **vývojové transkripční sítě** - diferenciační procesy, vývoj, vajíčko → vícebuněčný organismus
 - rozdělení diferencují do jiné tkáně/pletiva
 - aby se staly součástí nové tkáně, musí vyjádřit specifickou sadu proteinů
 - sada určuje, jestli bude tkáň svalová nebo neurální
 - zkoumání na octomilkách, červech, mořských jezcích a lidech - pár silných motivů
 - C1-FFL a I1-FFL jako u senzorových sítí, taky autoregulační motivy a SIMs
 - dále pak jiné, co u senzorových sítí nebyly

5.1.1 Dvounodová pozitivní zpětná smyčka pro dělání rozhodování

- dva transkr.fak. regulují samy sebe - ve vývojových sítích spíše pozitivní
 - 2 pozitivní interakce - double-positive feedback - transkr.fak. se navzájem aktivují
 - 2 ustálené stavy - X a Y oba ON / oba OFF
 - signál pro produkci X/Y může nereverzivně uzamknout oba proteiny do ON
 - vzájemná aktivace
 - **blokovací mechanismus**
 - neujčinější, když geny regulované X a Y kódují proteiny patřící stejné tkáni

- 2 negativní interakce - double-negative feedback - vzájemná represe
 - také 2 stabilní ustálené stavy - X ON a Y OFF / X OFF a Y ON
 - vyjádření buď X nebo Y
 - vhodné, když geny regulované X patří buňkám s jiným osudem než geny regulované Y
- často ještě u obou nodů pozitivní autoregulace
 - zvyšování produkce při dosažení určité hranice
 - tím také stabilizace ON ustálený stav transkr.faktoru
- bistabilní přirozenost těchto motivů dovoluje buňkám dělat nereverzibilní rozhodnutí a přiřadit osudy, kdy specifická sada genů se vyjadřuje a další je utlumená

5.1.2 Regulující zpětná vazba a regulovaná zpětná vazba

- 2 hlavní 3-nodové motivy - obsahují zpětnou vazbu
 - trojúhelníkový motiv, kde X a Y vzájemně regulují Z - regulující zpětná vazba -
- 10 možných kombinací znamének
 - double-positive - stejné znaménko z obou do Z
 - double-negative - často rozdílné znaménko do Z
 - také 2-nodová smyčka, kde jsou oba nody společně regulovány zezhora 1 transkr.fak.
 - regulovaná zpětná vazba
 - prvek paměti - Z může řídit zapnutí/vypnutí smyčky
 - smyčka si pamatuje, jestli bylo Z aktivní nebo ne
 - zapamatování si osudu buněk, i když už zmizel původní signál

5.1.3 Dlouhé transkripční kaskády a vývojové časování

- dlouhé transkripční kaskády vzácné v senzorových sítích, ve vývojových motivy
- řetězy interakcí - $X \rightarrow Y \rightarrow Z \rightarrow \dots$
- čas odezvy každého článku je daný mírou degradace/ředění v $T_{1/2}$ - odezva kolem jedné generace
 - u vývojových sítí tato odezva sedí líp - lepší řízení vývojových procesů

5.1.4 Propojená dopředná smyčka v B.subtilis sporulační síti

- sporulace - proces tvorby spor - buňka sloužící k dlouhodobému přežití bakterie v nepříznivých podmínkách
- b.subtilis - bacil senný
- ve vývojových sítích FFL často jen část větších a složitějších obvodů, než u senzor.sítí
- při hladovění se buňky přestanou dělit a diferencuje do odolných sporů
 - spory obsahují mnoho proteinů, co se v bakterii při růstu nevyskytují
 - odpočívající buňka, skoro úplně dehydrovaná - přežije dlouho, spící stav
 - při správných podmínkách se transformuje zpět na normální bakterii
- při tvorbě sporu je třeba změnit tvorbu sady proteinů - sporulace
 - potřeba stovek genů - spínání ON a OFF ve vlnách - každá specifický vliv na formaci sporu
 - síť tvořena několika transkr.fak. seřazenými ve spojeném C1 a I1-FFL
 - nakomibování FFLs pro využití jejich zpoždění a generování pulsů - časový program genetické exprese
 - kaskádově seřazené FFLs generují 2 pulsy genů následované třetí pozdním

5.2 Síťové motivy v signálových transdukčních sítích

- **signálová transdukční síť** - mnohem rychlejší zpracovávání informací než 1 generace
- interakce mezi signálovými proteiny
- vycítit informaci z okolí, zpracovat, regulovat aktivitu transkr.fak
- vstupy obvykle detekovány **receptorovými proteiny**
 - jeden konec mimo buňku, druhý uvnitř v cytoplasmě
 - mimobuněčný konec detekuje molekuly zvané **ligandy**
 - navázáním konformační změna v receptoru, vnitřní část se stane aktivní a katalyzuje dané chemické modifikace do difusního messenger proteinu v buňce
 - posílání 1 bitu informace z receptoru do messengeru
 - modifikace funguje v rámci sekund až minut

5.3 Informace zpracovávána pomocí vícevrstvých perceptronů

- v signálových transdukčních sítích jsou nody signálové proteiny a hrany interakce
- silné 2-4-nodové motivy - bi-fan, diamant
- diamant - vícevrstvý vzor, podobný strukturám DOR v kaskádě s DOR, který dostává informaci z nadřazeného DORu
 - vzory DOR se ale normálně nevyskytují v kaskádě
- **vícevrstvé perceptrony** - podobná struktura využívána v umělé inteligenci a neuronových sítích

5.3.1 Tréninkový model pro perceptrony proteinové kinázy

- **kináza** - enzym, který přenáší fosfátovou skupinu z vysokoenergetické donorové molekuly (např. ATP) na určitou cílovou molekulu (substrát)
- **kaskáda proteinové kinázy** - cesty zpracování informací nalezené ve většině eukaryotických organismů
- aktivace kaskády, když se receptor naváže na ligand a aktivuje první kinázu $X \rightarrow$
 - $\rightarrow X$ fosforylace kinázy Y na dvou daných místech $\rightarrow Y$ dvojně zfosforylovaná \rightarrow
 - \rightarrow začne fosforylovat $Z \rightarrow$ po dvojně fosforylaci fosforyluje transkr.fak. \rightarrow exprese genu
- fosfatázy - enzymy proteinů postupně defosforylující kinázy - také obsažené v kaskádách
- často využito lešenářské proteiny držící kinázy u sebe
- adaptorové proteiny můžou spojit danou kaskádu k rozdílným vstupním receptorům v odlišném typu buňky
 - kaskády tedy znovuvyužitelné moduly
 - stejná kaskáda transdukuje rozdílný signál v rozdílné tkáni
- kaskády většinou ve vrstvách, často 3
 - v 1. řadě kinázy $X_1, X_2 \dots$ aktivují další řadu kináz $Y_1, Y_2 \dots$
 - formování vícevrstvého perceptronu který může integrovat vstupy z více receptorů
- **kinetika prvního řádu** - nejjednodušší kinetika pro kinázy
 - míra fosforylace Y z X je proporcí vůči koncentraci aktivního X násobené koncentrací jeho substrátu - nefosforylované Y , značení Y_0

$$rate.of.phosphorylation = vXY_0 \quad (5.1)$$

míra v kinázy X - kináza Y fosforylovaná 2 rozdílnými vstupními kinázami - X_1, X_2 - fosforylovaná forma Y_p , nefosforylovaná forma Y_0

$$Y = Y_o + Y_p \quad (5.2)$$

$$dY_p/dt = v_1 X_1 Y_o + v_2 X_2 Y_o + \alpha Y_p \quad (5.3)$$

- stabilní stav - $dY_p/dt = 0$
- řešení

$$Y_p/Y = f(w_1 X_1 + w_2 X_2) \quad (5.4)$$

f - rostoucí saturující funkce

$$f(u) = \frac{u}{1+u} \quad (5.5)$$

w - váhy vstupů

$$w_1 = v_1/\alpha \quad (5.6)$$

- koncentrace fosforylovaného Y je rostoucí funkce vážených sum dvou aktivit kinázy
- když je Y kináza, která musí být fosforylována na dvou místech aby byla aktivní, vstupní funkce je strmější

$$f(u) = \frac{u^2}{1+u+u^2} \quad (5.7)$$

- S-shaped vstupní funkce vedou k aktivaci Y pouze pokud je suma vstupů větší než mezní hranice - tedy 1
- **hranice aktivace** - $w_1 X_1 + w_2 X_2 = 1$
 - přímka v ploše vymezené vstupními aktivitami

5.3.2 Vícevrstvé perceptronu mohou provést detailní výpočty

- jednovrstvý perceptron X_1 - X_2 rozdělí plochu přímkou do 2 částí - přidáním vstvy možnost složitějších výpočtů
- 2 vstvy - Y_1 a Y_2 mají vlastní sady mezí ze 2 vstupů X_1 , X_2 - rozdělení v půlce
 - oblasti nízké a vysoké aktivity
- kinázy Y_1 a Y_2 můžou fosforylovat kinázu Z - pouze pokud jsou fosforylované

$$Z_p/Z = f(w_{z1} Y_1 + w_{z2} Y_2) \quad (5.8)$$

- když jsou váhy w malé, Y musí být fosforylované pro překročení hranice
 - Z je fosforylované pouze na ploše X_1 - X_2 , když jsou Y ny aktivní
 - oblast dána průsečíky dvou regionů aktivity Y_1 a Y_2
 - aktivace Z v regionu vyhrazeném 2 úsečkami
- když v prostřední vrstvě místo kinázy specifická fosfatáza - fosfatáza odstraní fosforylní úpravu - negativní váha
 - složitější aktivační regiony
- **diskriminace** - schopnost rozeznat jistý stimulující vzor
 - nastavením vah možnost rozeznat od sebe velmi podobné stimulující vzory
- **generalizace** - "vyplňování mezer" v částečných stimulujících vzorech
 - vpuštěním neúplného vzoru obdov reaguje jako by byl celý, pokud je mu daná část podobná
- **pozvolná degradace** - vadou na prvcích sítě nedojde k úplnému ukončení
 - výkon se zhoršuje úměrně poškození
- tyto 3 fenomény můžou charakterizovat funkcionalitu signálové transdukční sítě v buňce

5.4 Kompozitní síťové motivy: záporná zpětná vazba a oscilátorové motivy

- zatím proteinové signálové sítě a transkripční sítě zvlášť - v buňce ale pracují pospolu
 - často výstup signálové transduční cesty je transkr.fak
- spojená síť se 2 barvami hran - transkripční a protein-protein interakce
- motivy
 - 3 nody - transkr.fak. X transkripčně reguluje 2 geny Y a Z, protein.prod. pak přímo interagují
 - např. Y fosforyluje Z
 - kompozitní zpětnovazební smyčka ze 2 proteinů, které spolu interagují různými barvami
 - X transkr.fak aktivující gen Y, produkt Y interaguje s X na proteinové úrovni
 - často opačným způsobem
 - Y naváže na X a potlačuje jeho aktivitu transkr.faktoru bráněním přístupu k DNA
 - nejčastěji se objevuje v genetickém systému bakterie-člověk
 - pomalá aktivace Y, rychlé potlačování X - negativní stabilizace
 - stabilnější, než 2 pomalé interakce
- homeostáza - stabilita okolo fixovaného stavu - v mnoha systémech žádoucí
- oscilační dynamika - chování u některých systémů
 - cyklus buňky, kdy se periodicky duplikují genomy
 - cirkadiánské hodiny - pozoruhodně přesné biochemické obvody produkující oscilace na škále jednoho dne
- ...
- typický charakter - časování výrazně přesnější než amplituda
 - variace amplitudy způsobená měněním se vnitřním šumem z produkce proteinů
- mnoho oscilátorů bývá implementováno pomocí 2-barevného motivu kompozitní negativní zpětnovazební smyčky
 - transkr.fak. X má pozitivní autoregulaci
 - robustní časování i přes fluktuace biochemických parametrů komponenty
 - rodina relaxačních oscilátorů
 - zpětnovazební smyčka složená z regulátorů poskládaných dokola tvořící negativní zpětnovazební smyčku
 - 3 represory za sebou - represilátor
 - rodina oscilátorů zpoždění - méně precizní časování

5.5 Síťové motivy v neuronové síti C.elegans

- Ceanorhabditis elegans - háďátka obecné - půdní červ - cca 1000 buněk
 - nesouvisející sítě mají shodné motivy, anti-motivy a častěji se opakující různé vzory
- nody - neurony
- $X \rightarrow Y$ - X má synaptický spoj s Y
- podobné motivy jako v biochemických interakčních sítích - i přes odlišné prostorové a časové vazby