

Funkcijska genomika in sistemska biologija

Radovan Komel

Medicinski center za molekularno biologijo, Inštitut za biokemijo, Medicinska fakulteta UL, Vrazov trg 2, 1000 Ljubljana

Uvod

Razkritje primarne strukture človeškega genoma in genomov drugih organizmov, revolucija v informatiki in dramatičen napredek v nanotehnologiji so temelji in stebri novega, integrativnega znanstvenega področja, imenovanega funkcijska genomika. Število ugotovljenih nukleotidnih zaporedij genomov bakterij, gliv, rastlin in živali narašča iz leta v leto in znanost sedaj zanima, kakšni geni so v teh genomih in kakšna je njihova vloga v življenju celice in organizma. Če smo do nedavnega raziskovali samo strukture in delovanja posameznih genov ter produktov njihovega izražanja, funkcijska genomika v tem trenutku uvaja globalen pristop na ravneh transkriptoma, proteoma in metaboloma, ki jih povezuje v holističen pogled na celico, njeno strukturo in njeno vlogo.

Genomika

Februarja 2003 je preteklo petdeset let, odkar sta James Watson in Francis Crick v znanstveni reviji Nature objavila razkritje dvojnoverižne vijačne strukture DNA, osnovne komponente genomov vseh organizmov na Zemlji, in njene vloge kot informacijske molekule življenja. Malo pred tem, dve leti nazaj, sta mednarodni konzorcij 'Projekt človeški genom' in zasebna družba 'Celera Genomics' v revijah Nature in Science ločeno, vendar istočasno objavila osnutka nukleotidnega zaporedja 3 milijarde nukleotidnih parov velikega človeškega genoma in napovedala začetek obdobja 'funkcijske genomike', ko bomo raziskali vlogo in delovanje vseh 35.000 – 40.000 človeških genov in tako morda razumeli kako deluje celica, kako se je

razvilo življenje in kakšne so poti nadaljnje evolucije. Že pred več kot tridesetimi leti so znanstveniki odkrili način rezanja molekul DNA na manjše fragmente in njihovega združevanja v poljubna nova zaporedja in tudi vnašanje omenjenih fragmentov DNA v različne organizme, kjer so se v primeru, ko je šlo za gene, ti lahko izrazili v obliki svojih proteinskih produktov. Z genskim inženirstvom, ki je nepogrešljivi del sodobne biotehnologije, smo do danes pridobili že vrsto rekombinantnih beljakovin, ki jih medicina uporablja za zdravljenje številnih bolezni, pa tudi industrijskih encimov, ki jih uporabljamo v kemični, predelovalni in prehrabeni industriji. Z vnosom novih genov nam je uspelo ustvariti transgenske rastline in živali, na področju medicine je prišlo do prave revolucije v diagnostiki bolezni in pri napovedovanju nagnjenja k nekaterim boleznim, veliki pa so tudi obeti na področju tako imenovanega 'genskega zdravljenja'.

Tradicionalen pristop v genomiki je bil način 'korak za korakom' – odkrivanje genomskih označevalcev, njihovo povezovanje v karte kromosomov in genomov, nato ugotavljanje nukleotidnega zaporedja delov kromosomov in njihovo združevanje v nukleotidna zaporedja kromosomov in genomov, vmes pa seveda tudi preučevanje posameznih odkritih genov in celo uravnavanja izražanja povezanih genov. Odkrita je bila struktura prokariotskega in eukariotskega gena ter načini in uravnavanje njihovega izražanja. V biotehnologiji je ob genskem inženirstvu pravo revolucijo prineslo tako imenovano proteinsko inženirstvo, kjer na osnovi računalniških modelov lahko izdelujemo genske konstrukte in jih izražamo v obliki spremenjenih, vendar načrtovanih proteinskih produktov. To je biokemiji omogočilo poglobljeno raziskovanje odnosa struktura –



funkcija. Velik razcvet je doživela mutacijska analiza genov, kjer razlike in spremembe v genski strukturi vzporejamo s fenotipskimi lastnostmi in na teh osnovah gradimo pristope molekularne genetike oziroma molekulske diagnostike in genskega zdravljenja.

Med odkrivanjem zaporedij genomov različnih organizmov se je pojavila tako imenovana 'primerjalna genomika', ki z analizo podatkov išče razlike oziroma podobnosti med genomi in geni. Z njo dobivamo nova spoznanja o evoluciji bioloških molekul in organizmov ter o povezanosti oblik življenja. Prinesla je tudi spoznanje, da se posamezniki ene vrste po zaporedju nukleotidov lahko razlikujemo; razlike, poznane kot 'polimorfizmi enega nukleotida' pri človeku nastopajo vsakih 100-500 nukleotidnih parov in njihovo ugotavljanje v povezanosti s fenotipi kompleksnih večgenskih boleznih obeta prodor pri razjasnjevanju molekularskih osnov omenjenih boleznih. Številni označevalci genomske polimorfizmov se danes že uporabljajo v sodni medicini in kriminalistiki, za ugotavljanje razlik oziroma povezav med posamezniki, in za podobne namene tudi v biotehnologiji, pri selekciji in žlahtnjenju vrst.

Raziskave genomov so seveda proizvedle ogromne količine podatkov, ki jih je bilo potrebno zbrati, primerjati in urediti v smiselne celote. Izgradile so se javne zbirke podatkov, iz katerih je v vsakem trenutku mogoče pridobiti informacijo o posameznem genu/proteinu in v najboljšem primeru tudi o njegovi povezanosti z ostalimi. Mogoče so primerjalne analize, pri katerih raziskovalec odkrito strukturo vzporeja s podatki o že raziskanih strukturah in na tej osnovi gradi strategijo svojih nadaljnjih raziskav. Izgradila so se tudi računalniška orodja, s katerimi raziskovalec načrtuje nove molekule in posege v genome organizmov. Bioinformatika je začela povezovati različna področja genomike in genomiko s proteomiko.

Funkcijska genomika

Večina raziskav do razkritja strukture človeškega genoma je bila usmerjenih v posamezne gene/proteine oziroma v razmeroma majhno število genov in drugih bioloških označevalcev. Tak pristop ni prinesel zadovoljstva pri raziskavah, ki so poskušale razkriti molekulske osnove in uravnavanje kompleksnejših bioloških fenomenov, fizioloških stanj in patologij. Post-genomsko obdobje, ki je nastalo iz rezultatov raziskovanja genomov, revolucije v informatiki in izrednega napredka v nanotehnologiji, je prineslo nov tehnološki val, nastanek do nedavnega še 'futurističnih' naprav, ki omogočajo globalne raziskave genomov in njihovega izražanja. Ob genomiki smo začeli govoriti še o drugih '-omikah': 'transkriptomiki', 'proteomiki', 'metabolomiki' in v zadnjem času še 'interaktomiki'. Vse te '-omike' smo nekako združili v izrazu 'funkcijska genomika', ki predstavlja novo, integrativno znanstveno področje in osnovo tako imenovane 'nove biotehnologije' in 'nove medicine'. Namen sedaj ni samo ugotoviti vse gene posameznih organizmov ali kar 'gene življenja' in njihove vloge v celici, organu in organizmu, temveč poiskati tudi vse medsebojne povezave in odvisnosti genov in njihovih produktov.

Transkriptom je celota izrazov genov nekega genoma, vse vrste in količina molekul RNA, ki so v izbranem trenutku oziroma fiziološkem stanju navzoče v celici. Proteom je njegov proteinski komplement in metabolom zastopanost vseh majhnih molekul, reaktantov oziroma produktov delovanja proteoma. Razvoj robustnih, globalnih in robotiziranih tehnoloških pristopov nam sedaj omogoča, da vzporedno in istočasno preučujemo veliko število genov in proteinov v izbranih celičnih razmerah kot npr. v neki razvojni stopnji, nekem celičnem tipu ali nekem patološkem stanju. Zbrane podatke lahko primerjamo, jih vzporejamo s podatki iz drugih laboratorijev in poskušamo razbrati funkcijski pomen posameznih genov oziroma proteinov. Tehnologije DNA- in proteinskih mikromrež, visokozmogljive dvo-dimenzionalne elektroforeze ter masne

spektrometrije so na široko odprle vrata v razumevanje vloge genov in prepletenosti njihovega delovanja. Ob tem se zavedamo, da je informacija o genih sicer pomemben del celovite slike, vendar so resnični 'igralci' v življenju celice proteini. Ko danes ocenjujemo, da ima človeški genom približno 35.000 – 40.000 različnih genov, verjetno ni pretirana ocena, da njegov proteom tega količinsko presega za en velikostni razred. Zanimajo nas možni načini izražanja posameznih genov, postranskripcijske in postranslacijske modifikacije njihovih produktov in njihove molekulske interakcije na ravneh proteini-nukleinske kisline, proteini-proteini in proteini-majhne molekule.

Delovanje proteinov je neločljivo povezano s strukturo. Z učinkovitimi tehnikami množičnega izražanja genov, kristalizacije proteinov in rentgenske kristalografije pridobivamo podatke, ki omogočajo izgradnjo računalniških programov za predvidevanja novih celovitih kot tudi delnih proteinskih prostorskih struktur in njihovih interakcij. To na široko odpira vrata v načrtovanje zelo specifičnih in natančno delujočih učinkovin, ki jih bomo po eni strani uporabljali za še bolj poglobljena preučevanja vloge posameznih proteinov v zapletenih mrežah molekulskih interakcij, po drugi strani pa kot 'ciljna zdravila', ki bodo delovala na mesto patološkega molekulskega dogodka, brez škodljivih stranskih oziroma sistemskih učinkov.

Z izničenjem genov v celicah in organizmih ter opazovanjem fenotipskih posledic smo dobili neposreden vpogled v vlogo posameznih genov oziroma njihovih proteinskih produktov v fiziologiji celice/organizma. Dogajanja in vivo na celični ravni lahko danes spremljamo tudi z mikroskopijo v realnem času in fluorescenčnim označevanjem proteinov, ki se izražajo v celici, kamor smo vnesli ustrezen genski konstrukt. Sodobna strukturalna biokemija (v omenjenem okviru včasih imenovana kar 'strukturalna genomika') in celična biologija sta tako nepogrešljiva dela funkcijske genomike.

Bioinformatika se je izkazala kot odlično orodje pri zbiranju podatkov iz raziskav izražanja genov, določevanja struktur proteinov in raziskovanja molekulskih interakcij. In silico metode in programi so bili bistveni za projekte določanja nukleotidnih zaporedij genomov, široko pa se tudi uporabljajo pri napovedovanju struktur in iz njih izhajajoče funkcije. Funkcijska genomika sedaj bioinformatiki odpira nove izzive: primerjave genomov, iskanje mrež proteinskih interakcij in iskanje mrežnih povezav med izražanjem genov na ravni transkriptoma in proteoma.

Zanimivo razlago oziroma razčlenitev funkcijske genomike je predstavil program Evropske znanstvene fundacije 'Integrirani pristopi funkcijske genomike'. Novo vedo razčlenjuje v vertikalna področja, ki predstavljajo različne metodološke oziroma tehnološke pristope, in horizontalna področja, ki pomenijo povezovanje podatkov iz vertikalnih področij - za oblikovanje splošno uporabnih bioinformacijskih orodij.

Vertikalna področja so:

- Analiza fenotipskih sprememb, ki so posledice mutageneze in izničenja genov.
- Mikromreže za ugotavljanje izražanja genov (ekspresijsko profiliranje) in za ugotavljanje mutacij.
- Proteomika: identifikacija in karakterizacija proteinov, njihovo izražanje in interakcije.
- Strukturna genomika: določanje strukture proteinov, njihova klasifikacija, modeliranje in primerjalne analize.
- *In silico* metode za opisovanje celičnih sistemov (iz podatkov v literaturi in podatkovnih zbirkah), računalniško napovedovanje in simulacija.

Horizontalna področja pa so:

- Primerjava različnih eksperimentalnih sistemov, standardizacija in laboratorijsko prilagajanje.



- Urejanje in upravljanje podatkov: podatkovne zbirke, povezovanje med njimi, ontologije.
- Kombiniranje in integracija podatkov iz funkcijske genomike za nova spoznanja v biologiji.

Zaključek

Raziskave v funkcijski genomiki, ki je že sama po sebi povezujoča veda, so privedle do spoznanja, da celico lahko obravnavamo kot sistem sočasnosti molekulskih dogodkov. Zanima nas sočasno izražanje vseh genov v izbranih fizioloških razmerah, vendar tudi to, kako ti s tem vplivajo en na drugega. Zato se celični transkriptom kot sistem prepleta s celičnim proteomom in ta s celičnim metabolomom. Začenjamo govoriti o 'interaktomu' in o 'sistemski biologiji', ki naj bi predstavljala nadgradnjo funkcijske genomike, saj bi bil njen namen zbrati in povezati mreže različnih '-omik' z metaboličnimi potmi in tako omogočiti razumevanje biologije celice, organov in organizmov na višji ravni. Izraz 'sistemska biologija' je nov in ni še soglasja, kaj pravzaprav natančno pomeni. Tudi, če privzamemo, da bomo povezovali transkriptomiko, proteomiko in metabolomiko, pa še kakšno drugo '-omiko', si lahko zamislimo, da gre za povezovanje podsistemov enega sistema, kar je celica v nekem fiziološkem stanju. Kaj pa celica v drugih fizioloških stanjih? In tkivo oziroma organ? In organizem? In skupnost organizmov? Torej bomo povezovali številne sisteme, ki so

sestavljani iz različnih podsistemov. Zato morda tudi skromnejši neposreden prevod izraza 'systems biology' v slovenski izraz 'biologija sistemov', vsaj do končnega spoznanja 'življenja kot sistema', ne bi bil najslabši!? Kakorkoli že, prihodnost bioloških oziroma biomedicinskih ved ali kar 'ved o življenju' se prikazuje kot zanimiv izziv!

Literatura

1. Human Genome Consortium: The human genome. Nature 15 Febr. 2001
2. Celera Genomics: The human genome. Science 2001; 295(5507): 1145-1434.
3. Berg P, Cooper NG (ed.): The Human Genome Project – Deciphering the Blueprint of Heredity. Mill Valley, California 1994: University Science Books.
4. Cox TM, Sinclair J: Molecular Biology in Medicine. Oxford 1997: Blackwell Science Ltd.
5. Brown SM, Hay JG, Ostrer H: Essentials of Medical Genomics. New Jersey 2003: John Wiley & Sons, Inc.
6. European Science Foundation: Integrative approaches for functional genomics – an ESF scientific program. <http://www.esf.org/> (http://www.esf.org/esf_article.php?language=0&article=40&domain=2&activity=1)
7. National Human Genome Research Institute: Methods for Discovering and scoring single nucleotide polymorphisms. NIH Guide Jan 9, 1998. <http://www.genome.gov/>
8. Henry CM, Washington C: Systems biology. Chem Eng N 2003 ; 81(2): 45-55.

