

Strokovni članek ■

## Tehnologija DNA mikromrež in njena uporaba v medicini

**Peter Juvan, Damjana Rozman**

**Izvleček.** Tehnologija DNA mikromrež omogoča celostne študije genoma (DNA) in transkriptoma (RNA). Na ravni transkriptoma sledimo izražanju genov (ekspresijske mikromreže), na ravni genoma pa iščemo področja v genomu, kjer je prišlo do spremembe v številu kopij DNA zaporedij (primerjalna genomska hibridizacija), ugotavljamo razlike v DNA zaporedjih (čipi za ponovno sekvencioniranje, SNP čipi), ali pa iščemo regulatorna DNA zaporedja, na katera se vežejo izbrani regulatorni proteini (tehnologija čip-čip). Priprava in uporaba mikromrež vključuje veliko bioinformatičnega dela, in sicer pri izboru nukleotidnih zaporedij, ki bodo predstavljala posamezne gene, pri načrtovanju poskusov, ter pri zajemanju, urejanju, shranjevanju in analizi pridobljenih podatkov.

## DNA microarray technology and its applications in medicine

**Abstract.** DNA microarray technology enables large-scale genome (DNA) and transcriptome (RNA) studies. At transcriptome level we track expression levels of individual genes (expression microarrays), and at genome level we search for DNA copy number aberrations (comparative genomic hybridization), detect differences in DNA sequences (re-sequencing and SNP chips) and search for specific transcription factor binding sites (chip-on-chip technology). Microarray design and the application of that technology involves a lot of bioinformatic work, e.g. selection of nucleotide sequences representing individual genes, planning experiments, capturing, managing and storing data and their analysis.

■ **Infor Med Slov:** 2006; 11(1): 2-15

---

Instituciji avtorjev: Fakulteta za računalništvo in informatiko, Univerza v Ljubljani (PJ), Center za funkcijsko genomiko in bio-čipe, Medicinska fakulteta, Univerza v Ljubljani (DR).

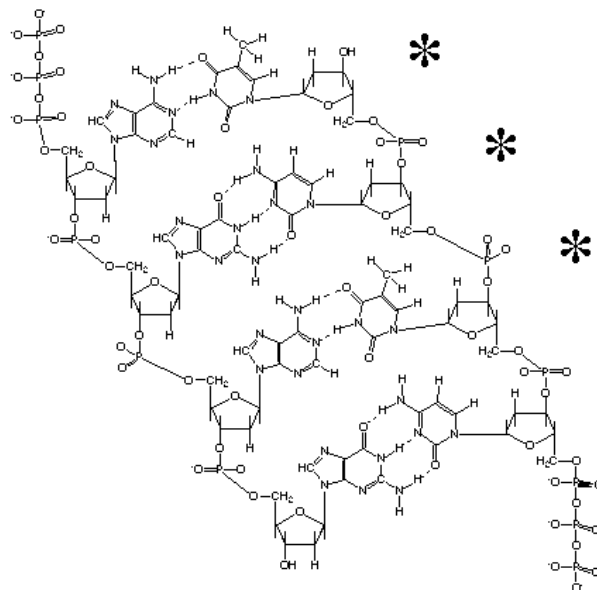
Kontaktna oseba: Peter Juvan, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, Tržaška 25, SI-1001 Ljubljana. email: peter.juvan@fri.uni-lj.si.

## Uvod

Organizem je kompleksen sistem, kjer tisoči genov in njihovih produktov (RNA in proteinov) usklajeno delujejo in ustvarjajo čudežnost življenja. Posamezen gen si lahko predstavljamo kot besedo v slovarju genoma, pristope, ki omogočajo celostno razumevanje izražanja genoma, pa kot orodja, ki nam pomagajo razumeti, kako se besede genoma povezujejo v smiselno besedilo.

Do nedavnega je v genomskih raziskavah prevladovalo pravilo "en gen – en poskus". Ta pristop podrobnega raziskovanja posameznih genov je od odkritja DNA leta 1953 pa do danes zaznamoval življenjske poti mnogih raziskovalcev, kamenčki posameznih odkritij pa so se sestavljali v mozaik skozi leta in desetletja. Vendar življenje celice ni sestavljeno iz izoliranih enot. Celica je splet signalnih, presnovnih, transportnih in drugih procesov, ki so tesno povezani med seboj in so v nenehnem stiku s svojim okoljem. Vsaka sprememba, ki jo povzročimo v celici, tudi če je ciljana na spremembo ene same podenote, na primer spremembe izražanja enega samega gena, se odraža v desetinah ali stotinah sprememb vzdolž različnih poti. Da bi čim bolj spoznali zakonitosti življenja, je smiselno uporabljati celostne (globalne) pristope, ki nam omogočajo sledenje mnogoterim lastnostim v enem samem poskusu. V zadnjih nekaj letih se je temu najbolj približala tehnologija DNA mikromrež, ki je eno od najučinkovitejših orodij za globalne študije izražanja genoma. Sprva se je pojem "DNA čip" uporabljal le takrat, ko je šlo za litografsko pripravljene čipe, s pojmom "DNA mikromreža" pa smo opisovali vse preostale tehnologije priprave. Zaradi napredka tehnike, ki z natančnimi roboti omogoča nanašanje molekul DNA z visoko gostoto, je ločnica med "čipom" in "mikromrežo" pogosto nejasna. Tako včasih govorimo o čipu tudi takrat, kadar so izbrane molekule DNA, ne glede na izvor in način priprave, nanešene na trdno podlago v visoki gostoti, na primer več tisoč skupin molekul na ploščici velikosti objektnega mikroskopskega stekelca, medtem ko mikromreža

vsebuje manjše število skupin molekul DNA. Uporabljajo se še pojmi genomski čip, mreža genov in bio-čip. Slednji izraz je najbolj splošen, saj predpona "bio" opisuje, da se na biološkem čipu lahko nahajajo biomolekule katerekoli vrste – DNA, RNA, proteini in druge.



**Slika 1** Podroben prikaz hibridizacije – komplementarnega parjenja dveh enojnih verig nukleinskih kislin (\* označuje fluorescentno označeno verigo).

## Vrste DNA čipov in mikromrež

DNA čipi so mikroskopske skupine tisočih DNA molekul z znanimi nukleotidnimi zaporedji, ki jih pritrdimo na podlago. Vsak gen je na mikromreži ali čipu zastopan z vsaj eno skupino identičnih DNA molekul. Organizirana razporeditev skupin DNA molekul predstavlja matrico, na kateri po hibridizaciji določimo razliko v izražanju genov med poskusnim vzorcem in kontrolo. Postopek hibridizacije temelji na komplementarnem parjenju baz A-T in G-C po modelu Watsona in Cricka (slika 1). Mikromrežo z organizirano razporeditvijo tarčnih DNA molekul izpostavimo fluorescentno ali radioaktivno označeni preizkusni snovi (imenovani proba ali sonda), ki jo pripravimo iz preiskovanih celic ali tkiv.

Hibridizacijski signal na določenem mestu matrice nam izraža identiteto nukleotidnega zaporedja, velikost signala pa je merilo za količino izraženega genskega produkta. Prvi DNA čipi so se uporabljali za spremljanje celostnega (globalnega) izražanja genov, vedno več pa je tudi uporabe za določanje sprememb na ravni genoma. Dandanes se DNA čipi najpogosteje uporabljajo za:

1. ugotavljanje količine izraženih genov (transkriptom ali ekspresijsko profiliranje) ter
2. določanje nukleotidnih zaporedij in sprememb na ravni genoma – sekvencioniranje, iskanje enojnih nukleotidnih polimorfizmov (SNP) in mutacij, primerjalno genomsko hibridizacijo (CGH) ter iskanje regulatornih DNA zaporedij (metoda čip-čip – kromatinska imunoprecipitacija z analizo čipov).

Obstaja več vrst DNA mikromrež, od raziskovalnega vprašanja oziroma namena uporabe pa je odvisno, katera je najprimernejša. Mikromreže s kratkimi oligonukleotidi, ki so sintetizirani na matrici in situ, so sprva imenovali DNA čipi, saj je tehnologija priprave podobna pripravi računalniških čipov. Vodilno vlogo pri pripravi teh čipov ima podjetje Affymetrix (<http://www.affymetrix.com>), ki je metodologijo tudi razvilo in patentiralo. Pri Affymetrixovi tehnologiji je vsak gen zastopan s preko 10 kratkimi nukleotidi, specifičnost hibridizacije posameznega nukleotida pa se določa na podlagi odsotnosti hibridizacije z nukleotidom, ki se od tarčnega razlikuje za eno samo bazo.

Obstajajo tudi mikromreže z dolgimi oligonukleotidnimi sondami, mikromreže s komplementarnimi DNA (cDNA) sondami in mikromreže z več 100 kb dolgim odseki genomske DNA. Razen pri tehnologiji Affymetrix so sonde sintetizirane po klasičnih molekularno-bioloških postopkih in so kasneje z nanašalnim robotom (angl. spotterjem) nanešene na podlago. Tehnologija dolgih oligonukleotidov (50 – 80 bp) uporablja eno do tri oligonukleotidne probe za posamezni gen. Skupina Patricka Browna s sodelavci z Univerze Stanford je razvila tako

imenovane klasične DNA mikromreže, kjer so na stekleno ploščico nanešene 300 do 500 bp komplementarne DNA (cDNA), od katerih vsaka predstavlja en gen. Le-te pridobimo iz celičnih informacijskih RNA (mRNA) s pomočjo gensko-specifičnih začetnih oligonukleotidov v reakciji obratnega prepisovanja in verižnega pomnoževanja s polimerazo (RT-PCR). Tudi zelo dolge odseke genomske DNA pridobimo s klasičnimi molekularno-biološkimi postopki in jih naknadno nanesemo na podlago.

Kljub temu, da so bili Affymetrixovi DNA čipi za ekspresijsko profiliranje na tržišču prvi, pa zaradi visoke cene in zaprtosti sistema (podatki o nukleotidnem zaporedju na čipu uporabljenih oligonukleotidov niso dostopni, uporabnik ne more spreminjati genov na čipu) za mnoge raziskovalne laboratorije niso dostopni. V prihodnosti bo sicer cena komercialno dostopnih čipov in mikromrež padala in postajala dostopnejša, vendar pa je uporaba klasičnih mikromrež na principu cDNA še vedno smiselna za usmerjene študije omejenega števila genov (čipi nizke gostote), kot je npr. določanje bolezenskih markerjev izbranega, že dobro definiranega obolenja.

## DNA čipi za ekspresijsko profiliranje

Za ekspresijsko profiliranje se uporabljajo tako oligonukleotidne kot cDNA mikromreže. Obstajata dve glavni smernici priprave ekspresijskih DNA mikromrež. V prvem primeru čip vsebuje gene, ki jim je skupna fizična lokacija (npr. človeški kromosom 22) ali kar vse gene določenega organizma. Na tržišču so že dostopni DNA čipi celotnega genoma kvasovke *S.cerevisiae*, mnogih bakterij, pa tudi celotnega človeškega, mišjega in podganjega genoma. Glede na sedanjo tehnologijo je na stekleno ploščico moč nanesti do 50.000 genov. Celotni človeški genom se pri čipih podjetij Affymetrix in Agilent (<http://www.agilent.com>) nahaja na eni stekleni ploščici. Pri drugem pristopu čip vsebuje izbrane

gene, ki so med seboj smiselno (tematsko) povezani. Načrtovanje takih čipov izhaja iz a priori biokemičnega znanja o funkciji genov in iz poznavanja fiziologije ter patofiziologije proučevanih obolenj. Tematski čipi raziskovalcem ponujajo neomejene možnosti tvorjenja in uresničevanja idej na temeljnem, aplikativnem in klinično-diagnostičnem področju. Na tržišču je vedno večje število usmerjenih čipov, ki se ukvarjajo z izražanjem genov med onkogenezo (onko-čipi) in tekom drugih pogostih bolezni. Mnogi se razvijajo v smislu diagnostičnih orodij, ki bi lahko našla pot tudi v klinično prakso.

### **Priprava čipov nizke gostote**

Ne glede na to, ali se odločimo za pripravo čipa, ki bo vseboval dolge oligonukleotide, ali za čip na podlagi cDNA, začetna faza načrtovanja zahteva veliko bioinformatičnega dela. Izbor in število genov sta odvisna od raziskovalnega vprašanja, ki ga želimo s tehnologijo DNA mikromrež reševati. Dostop do nukleotidnih zaporedij vseh genov, ki so bila določena v akademskih inštitucijah, je omogočen preko mnogih strežnikov, kot je npr. strežnik Nacionalnega centra za biotehnoške informacije v ZDA (NCBI, <http://www.ncbi.nih.gov>). Preprosteje je, če imamo dostop do katere od velikih urejenih cDNA knjižnic, ki vsebuje v klasične vektorje klonirane cDNA večjega števila (oziroma vseh) genov posameznih genomov ali tkiv.

Tudi pri izboru in pripravi nukleotidnih zaporedij, ki bodo predstavljala posamezen gen, imajo ključno vlogo orodja bioinformatike. Vsako nukleotidno zaporedje mora specifično predstavljati en sam gen in mora imeti čim manjšo homologijo z vsemi ostalimi nukleotidnimi zaporedji na čipu. Od pravilnega načrtovanja čipa in poznavanja lastnosti nukleotidnih zaporedij je odvisna končna interpretacija rezultatov. V primeru čipa na osnovi dolgih oligonukleotidov izberemo za vsak gen 2 – 3 od 50 do 80 baz dolga zaporedja, ki specifično predstavljajo en sam gen. V primeru cDNA čipov izberemo za vsak gen dva kratka oligonukleotida, s pomočjo katerih v verižni

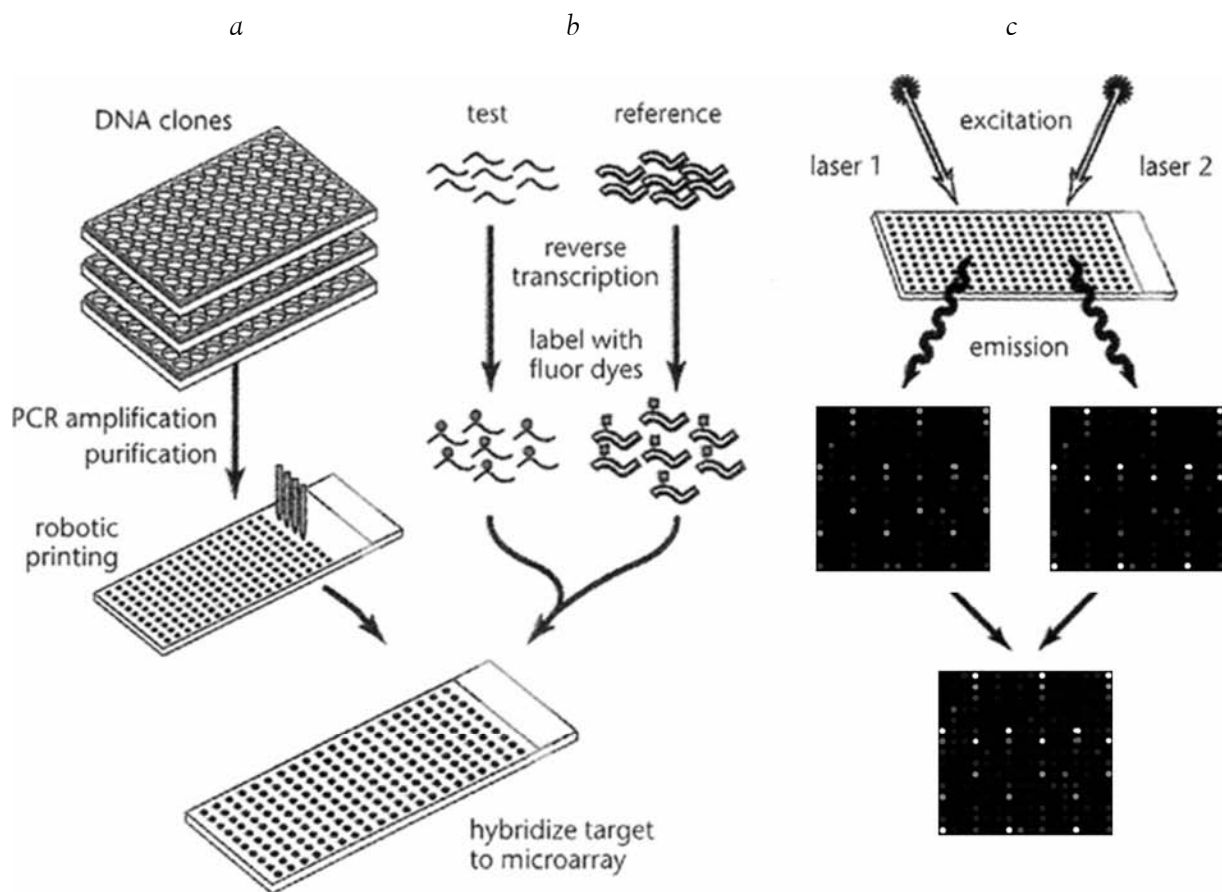
reakciji s polimerazo (PCR) pomnožimo odsek cDNA iz obratno prepisane RNA, ki je izolirana iz celice ali tkiva, kjer se izbrani gen izraža (slika 2a). Kot matrica za pomnoževanje posameznih genov so zelo primerne tudi urejene cDNA knjižnice, kjer je možno vse gene pomnožiti z dvema univerzalnima oligonukleotidoma, katerih prijemališči sta v vektorju, v katerem so cDNA klonirane, vendar je pri tem pristopu velika možnost križnih kontaminacij klonov.

Pred nanosom na trdno podlago je pri pomnoženih cDNA potrebno preveriti integriteto in koncentracijo nukleotidnih zaporedij. Za preverjanje integritete je primerna agarozna elektroforeza ali mikrokapilarna elektroforeza z elektroforeznim čipom (tehnologija Agilent Technologies). Preverjanje koncentracije in čistosti DNA poteka spektrofotometrično z merjenjem UV-absorpcije pri 260nm in 280 nm, pri čemer je za veliko število vzorcev neobhodno potreben spektrofotometer na mikrotiterske ploščice.

Po kontroli koncentracije in kakovosti DNA vzorce razredčimo z enim volumnom dvakrat koncentrirane nanašalne raztopine in jih v vnaprej določenem zaporedju razporedimo na 384-mestni mikrotiterski ploščici. Računalniški program nanašalnega robota mora prenesti motiv DNA vzorcev na mikrotiterski ploščici v nov, vnaprej izbrani motiv DNA vzorcev, kot se bodo pojavili na čipu (slika 2a). Zaradi statistične obdelave je priporočljivo, da je vsak DNA vzorec na čipu predstavljen vsaj v duplikatu. Nanašalni robot ima običajno večje število igel (tudi do 48), ki jih po želji odvezujemo ali dodajamo in s tem prilagajamo hitrost nanašanja in geometrijo čipa. Skupne lastnosti vseh nanašalnih robotov na tržišču so velika mehanska natančnost (natančnost nanašanja na 3 – 5 mikrometrov) in pripadajoča programska oprema za njihovo krmiljenje, ki praviloma omogoča, da so posamezne točke matrice medsebojno oddaljene največ 10 mikrometrov. Na tržišču je dostopnih več vrst nanašalnih robotov. Roboti lahko vsebujejo kapilare, ki posrkajo nekaj femtolitrov raztopine DNA in jo nato odložijo na trdno podlago čipa.

Prednost teh nanašalcev je majhna poraba vzorca DNA, slabost pa občutljivost kapilar, ki se zaradi majhnega premera pogosto mašijo, in možnost kontaminacije vzorcev, če sistem spiranja ni popoln. Druga vrsta nanašalnih robotov ima namesto kapilar igle, pri čemer za nanašanje raztopine DNA na trdno podlago čipa izkorišča pojav površinske napetosti kapljev. Njihova slaba lastnost je počasnost, saj je obisk

mikrotiterske ploščice z vzorci potreben pred vsakim nanosom posebej. Obstaja tudi kombinacija med kapilaro in iglo, tako imenovana razcepljena igla (angl. split pin), pri kateri količina raztopine DNA, ki se nabere v razpoki, zadostuje za večje število (tudi več kot 100) nanosov. Prednost tega sistema je velika hitrost nanašanja brez težav s spiranjem, ki so prisotne pri kapilarnem sistemu.



**Slika 2** Prikaz postopka priprave in analize DNA čipa po Patrick-Brownovi tehnologiji. Faze: *a* - priprava DNA čipa vključuje sintezo tarčnih DNA molekul in njihovo pritrditev na podlago z nanašalnim robotom; *b* - postopek hibridizacije s testnim in referenčnim fluorescentno označenim vzorcem nukleinskih kislin; *c* - analiza signala z optičnim laserskim čitalcem.

Podlago čipa navadno predstavlja kemično obdelano steklo (nanos oligonukleotidov in cDNA zahteva drugačne kemično obdelane steklene ploščice), lahko pa je to tudi najlonska ali nitrocelulozna membrana. V primeru obdelanega stekla pride do kemične reakcije med nukleotidi DNA in aktivnimi skupinami na premazu stekla.

DNA še dodatno pritrdimo z zamreženjem pod vplivom ultravijolične svetlobe. Sledi spiranje v več raztopinah detergenta in soli padajoče ionske jakosti ter hitro odstranjevanje vodnih kapljic z vakuumsko centrifugo. Slednje je ena izmed kritičnih stopenj postopka za pripravo kvalitetnih čipov, saj v primeru počasnega sušenja vodne

kapljice pustijo obris, ki je viden tudi po hibridizaciji in otežkoča kvantifikacijo signalov na področju motnje.

### Hibridizacija in odčitavanje

Po zgoraj opisanem postopku obdelani čip je pripravljen za hibridizacijo z izbranim fluorescentno označenim preiskovanim vzorcem. Preiskovani vzorec predstavlja iz celic ali tkiv izolirana RNA, ki jo s kompleti za označevanje prevedemo v fluorescentno označeno tkivno ali celično cDNA. Fluorescentno označeno cDNA skupaj s hibridizacijsko raztopino enakomerno porazdelimo po čipu (slika 2b). Hibridizacijo lahko izvajamo ročno, v vodni kopeli, ali na avtomatski hibridizacijski postaji, kjer je možno hibridizirati do nekaj deset čipov hkrati. S kontrolne plošče lahko uravnavamo temperaturni in časovni interval za vsak čip posebej, tako med hibridizacijo kot tudi med kasnejšim spiranjem. V primerjavi z ročno hibridizacijo in spiranjem daje aparat za avtomatsko hibridizacijo bolj ponovljive rezultate hibridizacije, hkrati pa se izognemo tudi mehanskim poškodbam čipa in motnjam, ki lahko nastanejo zaradi učinka vodnih kapljic med postopkom ročnega spiranja.

Hibridizacijski signal odčitamo z večlaserskim optičnim čitalcem (angl. scanner), ki omogoča sledenje fluoroforom različnih valovnih dolžin (slika 2c). Navadno imamo na istem čipu le dva različno obarvana fluorescentna signala. Ker pa je na tržišču dostopnih vse več različnih fluoroforov, se priporoča čitalec z vsaj dvema laserjema, ki skupaj s filtri omogoča zaznavanje 4 – 6 različnih fluoroforov. Čitalec mora imeti dovolj visoko ločljivost, da zazna točke, ki so medsebojno oddaljene le nekaj mikrometrov. Rezultat odčitavanja je računalniška slika, kjer je jakost hibridizacijskega signala ponazorjena z intenziteto slikovnih pik. Na tržišču je cel spekter dvo- in tri-laserskih čitalcev, ki imajo pri enakem številu laserjev medsebojno primerljive cene.

### Litografsko pripravljene čipi visoke gostote

Zaradi tehnične zahtevnosti in težavne standardizacije v raziskovalnih laboratorijih običajno ni mogoče pripravljati litografskih čipov, so pa ti čipi dostopni komercialno. To tehnologijo je vpeljala in najdlje razvila ameriška firma Affymetrix, po kateri čipi tudi nosijo ime. Vsak gen je na Affymetriksovem čipu predstavljen z najmanj desetimi oligonukleotidi, dolgimi do 20 bp. Ker so tako kratka nukleotidna zaporedja podvržena nespecifični hibridizaciji, še posebej pri sorodnih genih, ima vsak oligonukleotid na čipu tudi kontrolni oligonukleotid, ki se od pravega razlikuje za en sam nukleotid. Po hibridizaciji in odčitavanju signala se pri določanju izražanja posameznih genov upošteva le tiste oligonukleotide, kjer je hibridizacijski signal močan pri pravi probi in odsoten pri kontrolni probi. Zaradi popolnoma drugačne tehnologije priprave čipa je tudi postopek hibridizacije s tehničnega stališča drugačen, kot je opisano v prejšnjem razdelku. Prav tako je potreben čitalec večje občutljivosti, saj so in situ sintetizirani oligonukleotidi medsebojno oddaljeni manj kot naknadno na trdno podlago nanešene molekule DNA. Potrebna je tudi posebna programska oprema (GCOS - GeneChip Operating Software), ki s pomočjo ustrezne podatkovne zbirke vsakega od kratkih oligonukleotidov poveže z odgovarjajočim genom. Kljub "zaprtosti" sistema, ki zahteva popolnoma nov sklop opreme (hibridizacijska komora, laserski čitalec in programska oprema), zaradi velike standardizacije uporabniku zagotavlja visoko raven ponovljivosti meritev in možnost študije izražanja genov na ravni celotnih genomov (človeški genom je trenutno z več kot 47.000 transkripti dostopen na enem Affymetriksovem čipu). Visoke cene teh čipov (med 500 in 1000 EUR/čip) zaenkrat onemogočajo rutinsko uporabo malim uporabnikom in dajejo prednost dostop velikim industrijskim in kliničnim partnerjem. S stališča manjših uporabnikov so komercialni čipi visoke gostote pomembni predvsem kot prva stopnja raziskav, ki se bodo kasneje osredotočile na omejeno število genov z uporabo mnogo cenejših DNA čipov nizke gostote. Litografski čipi visoke

gostote so perspektivni tudi za uporabo v farmacevtski industriji, predvsem pri testiranju zdravilnih učinkovin na izražanje celotnega človeškega genoma.

## **DNA čipi za določanje nukleotidnih zaporedij in sprememb na ravni genoma**

Za drugo vrsto DNA mikromrež je značilno, da ne spremljamo ravni izražanja genov (raven mRNA), temveč določamo nukleotidno zaporedje na izbranih delih genoma (raven DNA). Za namen sekvencioniranja se razvijajo "prekrivajoče" mikromreže (angl. tiling microarrays), kjer so kromosomi po celotni dolžini pokriti s prekrivajočimi se 20 bp dolgimi oligonukleotidi. S tovrstnimi čipi je možno določiti genotip posameznika, tako nukleotidno zaporedje normalnih alelov kot tudi okvarjenih bolezenskih alelov. Mikromreže za sekvencioniranje celotnih genomov nekateri preprostejših organizmov (npr. različni soji kvasovk *S.cerevisiae*) vsebujejo prekrivajoča se zaporedja celotnega genoma kvasovke.<sup>1</sup> Pri pripravi takih mrež je zelo pomembna bioinformatika, saj je potrebno zadostiti predpostavki, da lahko vsako mesto v genomu teoretično vsebuje katerokoli od štirih baz: A,C,G ali T.

Podoben princip (vsako mesto v genomu lahko vsebuje katerokoli od štirih baz) velja tudi pri iskanju enojnih nukleotidnih polimorfizmov (SNP) in/ali mutacij. Razlika z mikromrežami za sekvencioniranje je v obsegu DNA zaporedja na čipu. Mikromreže za sekvencioniranje morajo vsebovati celotno področje genoma, katerega nukleotidno zaporedje želimo preveriti, SNP čipi pa so lahko omejeni npr. na posamezni polimorfni gen, ki je odgovoren za določeno obolenje. Posamezniki se namreč v svojem genskem zapisu med seboj razlikujemo tudi po polimorfnih nukleotidnih zaporedjih, med katere spadajo tudi SNP. Analiza DNA iz različnih osebkov bo na mikromreži pokazala prisotnost ali odsotnost signala, ki potrjuje prisotnost oziroma odsotnost

določenega SNP, z boleznijo povezanega. Tako lahko s čipom, ki vsebuje vse doslej poznane mutacije gena za cistično fibrozo (eno izmed najpogostejših monogenških obolenj), v enem samem poskusu ugotovimo, katere mutacije tega gena so prisotne pri preiskovanem osebkku. SNP čipi že nadomeščajo dolgotrajnejše metode, kjer se je vsaka mutacija preiskovala ločeno.

Primerjalna genomska hibridizacija (CGH) na čipu je visoko ločljivostna metoda za preiskovanje kvantitativnih sprememb na ravni genoma.<sup>2</sup> Kvantitativne spremembe vključujejo delecije (primanjkljaj oziroma izrez določenega odseka genoma), insercije (vključitev nukleotidnih zaporedij v genom), amplifikacije (pomnožitve določenega odseka DNA) in kromosomske prerazporeditve, kjer se deli dednine med posameznimi kromosomi nepravilno izmenjajo. Klasične citogenetske metode, s katerimi so preiskovali te spremembe (kariotipizacija, proganje kromosomov, fluorescentna in situ hibridizacija – FISH, idr.) imajo pomanjkljivo ločljivost in zahtevajo deleče se celice. Na CGH čipu so nanešeni klonirani poljubno dolgi (tudi do nekaj 100 kb) odseki DNA točno določene lokacije na kromosomu. To omogoča zelo natančno določevanje sprememb in njihovo preslikavo na določeno nukleotidno zaporedje. Resolucija CGH čipov je odvisna od razdalje med posameznimi področji DNA na kromosomu. Nekatera podjetja že razvijajo prekrivajoče se CGH čipe, ki lahko odčitajo spremembe na 1 bp natančno. Metoda se v največji meri uporablja v onkologiji, in sicer v diagnostične namene (prepoznavanje vrste raka) in s tem povezano prognozo, uporabna pa je tudi za določevanje kromosomskih aberacij različnih genetskih obolenj in za prenatalno diagnostiko različnih kromosomskih abnormalnosti, npr. Downovega sindroma (trisomija kromosoma 21). Analiza DNA različnih osebkov na mikromreži pokaže prisotnost ali odsotnost signala, kar potrjuje prisotnost oziroma odsotnost določene kromosomske spremembe, povezane z boleznijo.

Mnoga obolenja nastanejo tudi zaradi napačnega uravnavanja izražanja genov, ki nastane zaradi napačne vezave regulatornih proteinov

(transkripcijskih faktorjev) na regulatorne odseke genov (promotorje). Obstaja mnogo klasičnih metod za iskanje interakcij med regulatornimi odseki DNA in proteini, v zadnjih letih pa je v porastu metoda čip-čip. Gre za oligonukleotidne čipe, ki lahko vsebujejo zaporedja celotnega genoma (*S.cerevisiae*), pri kompleksnejših organizmih pa vsebujejo le posamezne dele genoma (posamezni kromosom ali neprevedene – regulatorne regije kromosoma). Metoda se zaenkrat uporablja bolj v raziskovalne namene, saj ima tehnične omejitve pri pripravi vzorca za procesiranje na čipu. Za analizo moramo proteine, ki so v celici vezani na regulatorne odseke DNA, trajno pritrčiti na DNA, nato pa to DNA izločiti iz zmesi in jo analizirati na čipu. Prvega dela postopka ne moremo izvajati na celem organizmu, ampak le na izoliranih celicah.<sup>3</sup>

## Vloga informatike pri tehnologiji DNA mikromrež

Tehnologija DNA mikromrež je prinesla nove izzive tudi na področju računalniških znanosti in statistike. Prednost tehnologije DNA mikromrež ni v samem načinu merjenja izraženosti genov, pač pa v zmožnosti merjenja več tisoč genov naenkrat. Preskok v obsegu meritev iz posameznih genov na več tisoč genov je zahteval razvoj novih standardov in metod za upravljanje, analizo in vizualizacijo podatkov ter njihovo algoritmično implementacijo na računalnikih, zaradi česar je področje analize podatkov DNA mikromrež v zadnjih letih postalo eno izmed najbolj obetavnih in donosnih področij za raziskovalce s področja računalniških znanosti in statistike.

Razvoj tehnologija DNA mikromrež je v veliki meri prehitel razvoj ustreznih analitičnih metod. V praksi pogosto prihaja do primera, da se podatki analizirajo naknadno, torej v času, ko so meritve že opravljene oziroma poskusa ni več možno spreminjati. Pri analizi se pogosto pokažejo pomanjkljivosti, ki so plod neoptimalne zasnove poskusa. Zato je sodelovanje biologov in analitikov pomembno že v času načrtovanja poskusa. Za

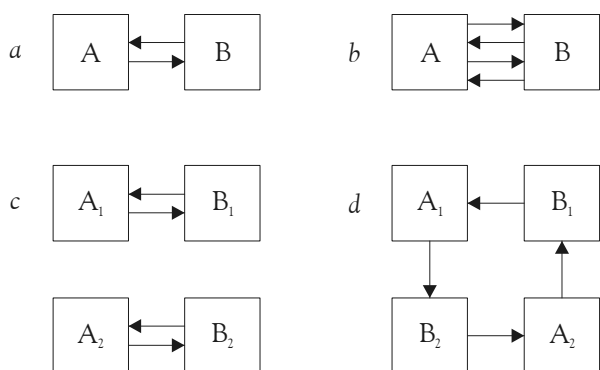
uspešno izvedbo biološkega poskusa je pomembno, da analitik dobro razume njegov namen, da sodeluje tako pri njegovi zasnovi kot tudi pri analizi podatkov in ovrednotenju rezultatov, in da se biolog zaveda nevarnosti in omejitev povezanih s tehnologijo DNA mikromrež.

Področje bioinformatike, ki se najširšem smislu nanaša na zajemanje, urejanje, shranjevanje in analizo raznovrstnih bioloških podatkov, zahteva od raziskovalcev hkratno poznavanje metod bioloških, računalniških in statističnih znanosti ter spretnost pri uporabi računalnika kot eksperimentalnega orodja. Čeprav stranskega pomena v primerjavi z analizo, urejanje, poimenovanje in shranjevanje podatkov trenutno zavzema večino časa bioinformatikov, saj je sistematičnost na tem področju predpogoj za uspešno aplikacijo analitičnih metod. Mnogo objav s področja bioinformatike predstavlja nove oziroma izboljšane analitične metode in algoritme, zlasti s področja strojnega učenja, prilagojene za analizo podatkov DNA mikromrež. Pogosto se zastavlja vprašanje, katera izmed teh metoda je trenutno "najboljša". Zavedati se moramo, da izbira analitične metode ni tako pomembna kot dobra zasnova biološkega poskusa, saj nobena metoda za analizo ne more odtehtati pomanjkljivosti pri njegovi zasnovi.

### Načrtovanje poskusa

Zaradi visokih stroškov, ki so povezani z uporabo tehnologije DNA mikromrež, se načrtovanje poskusa v nasprotju s tipičnim biomedicinskim poskusom pogosto prične s stališča denarnih sredstev. Osnovno vodilo na ta način zasnovanega poskusa ni testiranje določene hipoteze, pač pa tvorjenje hipotez oziroma upanje, da pridemo do smernic za nadaljnje raziskave. Tak pristop je zgrešen že v osnovi. Poskus mora biti osnovan na biološkem vprašanju, na katerega želimo odgovoriti, vprašanje pa mora biti osredotočeno in zastavljeno dovolj ozko, da lahko nanj odgovorimo v okviru denarnih sredstev, ki so nam na voljo.





**Slika 3** Različni načrti poskusa z dvobarvnimi DNA mikromrežami, ki vključujejo dva tretmaja (A in B). Načrta *a* in *b* predvidevata dve oziroma štiri tehnične ponovitve z zamenjavo barv (angl. dye swap); *c* in *d* sta osnovana na dveh neodvisnih bioloških ponovitvah tretmajev (ponazorjeno z indeksom pri oznaki tretmaja); *c* predvideva biološko ponovitev načrta *a*; *d* prikazuje enostaven krožni načrt.

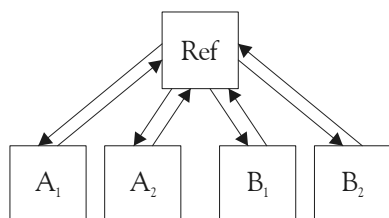
Pri načrtovanju poskusa moramo upoštevati različne komponente variance, ki je prisotna na posameznih korakih poskusa. V grobem jo lahko razdelimo na biološko varianco, ki je prisotna zaradi genetskih razlik med organizmi in vpliva okolja, tehnično varianco, do katere pride med postopkom izolacije, označevanja in hibridizacije, ter meritveno napako laserskega čitalca. Posamezne komponente variance zajamemo z ustreznimi ponovitvami, in sicer s ponovljenimi biološkimi vzorci, tehničnimi ponovitvami poskusov, in reproduciranimi probami znotraj posameznih čipov. Statistične teste lahko osujemo na podlagi katerekoli izmed zajetih komponent variance, razlika je le v interpretaciji njihovih rezultatov. Če nas zanima učinek določenega tretmaja, moramo teste osnovati na biološki varianci. Statistični testi na podlagi tehnične variance nam pokažejo razlike znotraj posameznih skupin. Prispevek posamezne komponente variance lahko ocenimo s korelacijo; pri ponovitvah znotraj posameznega čipa je ponavadi višja od 95%, pri tehničnih ponovitvah pade tipično na 60 do 80%, pri bioloških ponovitvah pa je lahko samo še 30%.<sup>4</sup> Ponovljivost poskusa bo veliko višja, če se izognemo biološkim ponovitvam in opravimo samo tehnične, vendar so rezultati takega poskusa varljivi – statistično

značilne razlike med skupinami lahko odražajo zgolj naključne razlike med posameznimi osebki.

Ustreznost poskusa lahko na preprost način ocenimo s številom prostostnih stopenj (angl. degrees of freedom), ki ustreza številu neodvisnih enot poskusa, od katerega odštejemo število različnih tretmajev. Neodvisne enote poskusa ustrezajo osebkom, ki bi teoretično lahko bili podvrženi poljubnemu tretmaju, in katerih vzorci so bili tretirani neodvisno drug od drugega v vseh fazah poskusa. Če pridemo do števila 5 ali več, smo na dobri poti. V določenih okoliščinah lahko število osebkov presega maksimalno število hibridizacij, ki jih lahko opravimo, ali pa količina RNA posameznega osebka ne zadošča za predvideno število hibridizacij. V takih primerih lahko osebkke združimo v skupine (angl. pools), pri čemer skupine predstavljajo nove neodvisne enote poskusa. Z zmanjšanjem števila neodvisnih enot poskusa se zmanjša komponenta biološke variance, ne pa tudi tehnična komponenta, kar poveča verjetnost napačnih zaključkov statističnih testov. V splošnem je bolje narediti večje število manjših skupin oziroma več bioloških in manj tehničnih ponovitev.

Pomemben korak pri načrtovanju poskusa je določiti število tehničnih ponovitev, kar je pri dvobarvnih čipih tesno povezano z odločitvijo, katere pare vzorcev bomo hibridizirali na istem čipu. Do učinkovite zasnove poskusa lahko pridemo z upoštevanjem nekaj enostavnih pravil.<sup>5</sup> Načrt predstavimo z usmerjenim grafom, kjer vozlišča predstavljajo biološke vzorce (neodvisne enote poskusa), usmerjene povezave pa hibridizacije, pri katerih vzorec na začetku povezave obarvamo z rdečim (Cy5) in vzorec na koncu povezave z zelenim (Cy3) barvilom. Slika 3 prikazuje različne zasnove poskusa za neposredno primerjavo vzorcev A in B. Načrta *a* in *b* predvidevata samo tehnične ponovitve (na podlagi zamenjave barv), *c* in *d* pa tudi biološke. Učinkovitost posamezne primerjave je odvisna od dolžine in števila poti med primerjanima vzorcema.<sup>6</sup> Najbolj učinkovito je skupaj hibridizirati tiste vzorce, katerih primerjava je najbolj zanimiva. Primerjati je možno tudi vzorce,

ki niso hibridizirani skupaj, pod pogojem da med njimi obstaja pot v grafu. Načrti *b*, *c* in *d* na sliki 3 predvidevajo enako število hibridizacij (štiri), med seboj pa se razlikujejo v učinkovitosti posameznih primerjav. Krožna zasnova (načrt *d*) je učinkovita predvsem pri manjšem številu vzorcev, vendar se njena učinkovitost močno zmanjša v primeru neuspeha katerekoli od hibridizacij. Robustnost krožne zasnove lahko povečamo s prepletanjem. Potencialno pristranost primerjav minimiziramo z uravnoteženim načrtom, kjer iz vsakega vzorca naredimo sodo število tehničnih ponovitev, od katerih polovico označimo z enim, polovico pa z drugim barvilom.



**Slika 4** Načrt poskusa z dvobarvnimi DNA mikromrežami, kjer vzorce ( $A_1 \dots B_2$ ) primerjamo preko skupne reference (Ref).

Vzorci lahko med seboj primerjamo tudi posredno preko skupne reference (slika 4). Večina današnjih poskusov je zasnovanih na ta način, saj ima kljub očitni neučinkovitosti (kar polovica meritev se nanaša na skupno referenco) mnogo prednosti, kot so npr. enaka učinkovitost vseh primerjav, možnost razširitve z novimi vzorci in manjša možnost napake pri laboratorijskem delu. Pri uporabi skupne reference je pomembna izbira reference, ki bo "prižgala" vse probe na naši mikromreži, in je hkrati homogena, stabilna in prisotna v zadostni količini. Odločimo se lahko med nakupom univerzalne reference in lastno referenco, ki jo ustvarimo z združitvijo RNA iz vseh vzorcev, ki jih bomo analizirali. Prednost slednje je predvsem v količini RNA, ki je podobna kot pri naših vzorcih.

Dober načrt mora upoštevati dejstvo, da je za veljavnost statističnih testov potrebno pri vseh korakih elemente poskusa izbirati naključno. Najbolj pomembna je naključna izbira osebkov in

tretmajev. Če na to izbiro ne moremo vplivati, moramo zagotoviti dovolj velik vzorec, da je raznolikost populacije dobro predstavljena. Naključna mora biti tudi izbira vzorcev za hibridizacije, ki jih bomo opravili v istem dnevu, in izbira barvila pri tehničnih ponovitvah poskusa. Čipi so ponavadi tiskani v serijah, ki se med seboj lahko močno razlikujejo v kakovosti. Pogosto se razlike v kakovosti pojavijo tudi glede na položaj čipa pri tiskanju in celo glede na pozicijo probe na čipu.<sup>7</sup> Zato je pomembno, da tudi čipe izbiramo naključno, in da so probe na čipu razporejene v naključnem vrstnem redu (idealno bi bilo, če bi bila razporeditev prob na vsakem čipu drugačna, vendar bi to močno otežilo njihovo izdelavo in analizo).

### Shranjevanje podatkov, standardi in ontologije

Podatki, ki jih pridobimo z uporabo DNA mikromrež, so v veliki meri odvisni od pogojev, pri katerih je bil izveden poskus. Za njihovo pravilno interpretacijo moramo poznati podrobnosti izdelave čipa (pozicije in zaporedja transkriptov), pripravo vzorcev in z njimi povezane tretmaje, korake pri izvedbi poskusa, nastavitve laserskega čitalca in postopek normalizacije in transformacije podatkov. Velika količina podatkov in njihova raznolikost, ki je povezana z uporabo tehnologije DNA mikromrež, je zahtevala razvoj standardov za njihovo upravljanje, shranjevanje in izmenjavo. Pobuda je na tem področju prišla od mednarodne organizacije biologov, računalnikarjev in analitikov, imenovane Microarray Gene Expression Data (MGED) Society (<http://www.mged.org>). Organizacija združuje šest projektov s področja standardizacije, od katerih so najpomembnejši trije:

- Minimum Information About a Microarray Experiment (MIAME) je pobuda za standardizacijo opisa poskusa z uporabo tehnologije DNA mikromrež;<sup>8</sup>
- MicroArray and Gene Expression (MAGE) združuje podatkovni model (MAGE-OM) in

jezik (MAGE-ML) za predstavitev podatkov DNA mikromrež;<sup>9</sup>

- MGED Ontology (MO) je ontologija za označevanje poskusov z DNA mikromrežami, ki predpisuje termine z vidika načrtovanja poskusa, zgradbe čipa, priprave vzorcev, hibridizacijskih protokolov in analize podatkov.<sup>10</sup>

Standardi so pomembni za dvig kakovosti podatkov, izmenjavo in pravilno interpretacijo podatkov, primerjavo podatkov različnih raziskovalnih skupin in različnih bioloških sistemov ter možnost reprodukcije poskusov. Standard MAGE omogoča izmenjavo podatkov med različnimi sistemi in njihovo medsebojno primerjavo. Tehtnost primerjave je odvisna od poznavanja parametrov poskusa, zaradi česar je pomembna standardizacija minimalnega nabora parametrov za opis poskusa oziroma standard MIAME. Računalniška avtomatizacija izmenjave podatkov in njihova medsebojna primerjava je možna le ob uporabi standardnega nabora terminov oziroma ob uporabi ontologij, pri katerih poleg terminov definiramo tudi logične opise uporabljenih terminov in njihove medsebojne relacije. Namen ontologija MO ni združiti termine s tako širokega področja, kot ga pokriva tehnologija DNA mikromrež, pač pa določiti ogrodje za uporabo terminov iz obstoječih ontologij. Ontologija MO vključuje mnogo terminov, ki se ne nanašajo izključno na tehnologijo DNA mikromrež (npr. termini za opis vzorcev, načrta poskusa ipd.), pač pa jih je možno aplicirati tudi na druge tehnologije funkcijske genomike (masno spektrometrijo, in situ hibridizacijo), zaradi česar je v prihodnosti predvidena njena razširitev z ontologijo FuGO (Functional Genomics Investigation Ontology), ki bo omogočala skladno označevanje poskusov s področja funkcijske genomike neodvisno od uporabljenih tehnologij.

Standardi ne predpisujejo strukture podatkovnih zbirk, zaradi česar so različne raziskovalne skupine glede na lastne potrebe in omejitve razvile zbirke, ki se med seboj močno razlikujejo tako po strukturi

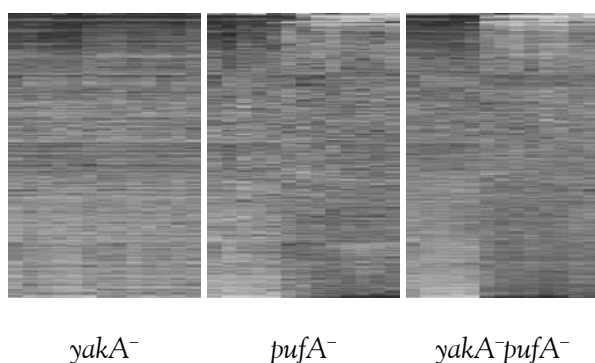
kot tudi po funkcionalnosti. Gardiner-Garden in soavtorji<sup>11</sup> so pripravili njihov pregled in primerjavo, Anderle in soavtorji<sup>12</sup> pa so opisali potek tipičnega poskusa z DNA mrežami s stališča upravljanja s podatki in njihovega shranjevanja. Upoštevanje priporočil MGED in javna dostopnost podatkov je danes pri večini vplivnejših revij pogoj za objavo prispevka s področja DNA mikromrež. Javna podatkovna skladišča, med katerimi so največja evropski ArrayExpress Evropskega instituta za bioinformatiko (EBI),<sup>13</sup> ameriški GEO Nacionalnega centra za biotehnoške informacije (NCBI)<sup>14</sup> in japonski CIBEX Nacionalnega instituta za genetiko (NIG),<sup>15</sup> odpirajo možnost za nove ideje in primerjave, ki sicer ne bi bile možne znotraj posameznih institucij.

### **Analiza in odkrivanje zakonitosti v podatkih DNA mikromrež**

Na prvi pogled se zdi, da lahko v namen analize podatkov DNA mikromrež uporabimo standardne statistične prijeme za določitev relacij med spremenljivkami, kjer posamezne probe (geni) na čipu predstavljajo neodvisne spremenljivke. Kljub temu je dandanes večina analiz opravljenih s pomočjo metod strojnega učenja. Glavni razlog je v naravi podatkov. Standardne statistične metode so prilagojene za podatke, kjer je število meritev vsaj za red velikosti večje od števila spremenljivk. Pri podatkih DNA mikromrež je slika ravno obrnjena: zaradi visoke cene meritev ali pomanjkanja primernih bioloških vzorcev imamo običajno opravka z relativno majhnim številom meritev (tipično med 10 in 100) in velikim številom spremenljivk (tipično reda velikosti od 1.000 do 10.000). Področje strojnega učenja je uspešno prodrlo na področje analize podatkov DNA mikromrež zaradi svojih bogatih izkušenj z analizo slabo determiniranih sistemov visokih dimenzij (npr. s področja avtomatskega razpoznavanja človeških obrazov). V resnici tudi na področju statistike poteka mnogo raziskav s področja analize takih sistemov, le da do nedavnega žal niso prodrle na področje analize podatkov DNA mikromrež. Podobno kot na drugih področjih znanstvenega raziskovanja je tudi

pri analizi podatkov DNA mikromrež za verodostojnost zaključkov namreč potrebna ocena njihove statistične značilnosti.

Kot primer praktične uporabe tehnologije DNA mikromrež naj povzamemo študijo s področja funkcijske genomike, katere glavni cilj je določiti biološke funkcije genom, skupinam genov in predvsem interakcijam med geni. Zmožnost hkratnega merjenja izraženosti več tisoč genov, ki jo omogoča tehnologija DNA mikromrež, ne ponuja samo mehanizma za tvorjenje hipotez oziroma obetavnih smernic za nadaljnje raziskave, pač pa ob ustrezni aplikaciji tudi orodje za potrjevanje njihove veljavnosti. Ekspresijsko profiliranje organizma *D.discoideum*<sup>16</sup> je pokazalo, da je rekonstrukcija regulatorne poti, ki uravnava prehod med rastjo in razvojem tega organizma, možna zgolj s pomočjo tehnologije DNA mikromrež in brez uporabe biološkega predznanja oziroma pristranosti. Slika 5 prikazuje ekspresijske profile dveh enojnih in ustrezne dvojne mutacije tega organizma, ki služijo kot generičen fenotip za določitev zaporedja vplivov mutiranih genov. Na podlagi podobnosti med fenotipi lahko s pomočjo klasične analize epistaze<sup>17</sup> in ocene statistične značilnosti<sup>18</sup> tudi na ravni transkriptoma potrdimo hipotezo, da gen *yakA* vpliva na gen *pufA* in ne obratno.



**Slika 5** Ekspresijski profili enojnih mutantov *yakA*<sup>-</sup> in *pufA*<sup>-</sup> ter dvojnega mutantu *yakA*<sup>-</sup>*pufA*<sup>-</sup> organizma *D.discoideum*.

## Uporaba in obeti tehnologije DNA mikromrež na področju medicine in farmakogenomike

Tehnologija DNA čipov in mikromrež ima veliko uporabnost pri odkrivanju genov, vključenih v bolezenske fenotipe, diagnosticiranju obolenj in odkrivanju novih učinkovin na področju farmakogenomike in toksikologije. Mnoga genetska obolenja so poligenska, kar pomeni, da ne nastanejo zaradi napake v enem samem genu, temveč je v razvoj bolezni vključenih več okvarjenih genov. Pri mnogih od teh obolenj je do sedaj poznan le en glavni "gen krivec", tehnologija DNA mikromrež pa omogočajo odkritje preostalih udeleženih genov.

Na področju medicine ponuja tehnologija DNA mikromrež možnost za določitev osnovnih vzrokov že znanih in za odkritje vzrokov še neznanih obolenj, nove strategije za iskanje t.i. bolezenskih markerjev in razvoj novih diagnostičnih orodij, s tem pa tudi izboljšano možnost za ustrezno preventivo in izbiro zdravljenja. Celostne študije izražanja genov z ekspresijskim profiliranjem močno spreminjajo naše vedenje o obolenjih in njihovih kompleksnostih. Vsaka sprememba v celici povzroči niz sprememb, katerim je bilo pred tehnologijo mikromrež praktično nemogoče slediti.

V onkologiji mikromreže obetajo možnost natančnejše diagnoze in prognoze rakastih obolenj. Golub in soavtorji<sup>19</sup> so prvi pokazali, da lahko zgolj na podlagi profilov izražanja genov ločimo med različnimi vrstami raka in tako avtomatsko, brez biološkega predznanja, identificiramo nove vrste raka, hkrati pa tudi napovemo uspešnost kemoterapije. Mejniki za prenos tehnologije DNA mikromrež v klinično prakso predstavlja tri leta kasneje objavljena strategija za terapijo raka dojke, oblikovana po meri pacienta.<sup>20</sup> Na tržišču se že pojavljajo čipi nizke gostote s področja onkologije, s katerimi bi bilo moč ugotavljati vrste tumorjev. Številne raziskave v smeri oblikovanja diagnostičnih čipov za sledenje pogostih obolenj pri človeku so v polnem razmahu v različnih laboratorijih po svetu. Problem hitrega prenosa

diagnostičnih čipov v širšo klinično prakso poleg cene predstavljata še standardizacija postopkov in aparaturo za odčitavanje signala.

DNA mikromreže so nepogrešljive tudi v farmakogenomiki, vedi, ki proučuje, kako dedni zapis (genetski faktorji) posameznika vpliva na odziv organizma na zdravila. Je nekakšen hibrid med farmakologijo in funkcijsko genomiko. Farmakogenomika vključuje razvijanje novih zdravil, ki bodo ciljale le eno tarčo in tako zmanjšala verjetnost nezaželenih stranskih učinkov. Mikromreže omogočajo prikaz celostnega učinka učinkovine na genom in s tem določitev zaželenih (zdravilni učinek) in nezaželenih tarč (stranski učinek).

Zelo hitro si v klinično prakso utirajo pot čipi za določanje zaporedja in sprememb v genomu, predvsem SNP čipi. Pogosto se srečujemo z vprašanjem, zakaj določeno zdravilo pri vseh pacientih ne deluje enako in zakaj so nekatera zdravila lahko za določeno skupino ljudi celo toksična. Vzrok je v enojnih nukleotidnih polimorfizmih (SNP), miniaturnih spremembah med posameznimi genomi, ki naredijo vsak osebek genetsko edinstven. Eden izmed ciljev farmakogenomike je poiskati povezave med terapevtskimi odzivi na učinkovino in genetskim profilom posameznika. S pomočjo tehnologije mikromrež lahko določimo, ali bo osebek npr. hitro ali počasi presnavljal določena zdravila. To obeta, da bo nekoč moč doseči osebno terapijo, kjer bodo vrste in doze zdravil prilagojene genetskemu zapisu posameznika. Okolje, prehrana, starost, način življenja, splošno zdravstveno stanje ipd. sicer vplivajo na odgovor posameznika na zdravila, vendar je ključ do osebne terapije poznavanje razlik v zapisu genov, ki so odgovorni za presnovo zdravil.

SNP so tako postali eden izmed najpomembnejših tarč medicinskih in farmakogenomskih raziskav, saj lahko razkrijejo poti do novih tarč zdravilnih učinkovin. SNP in ostale diagnostične čipe pospešeno razvija firma Roche (<http://www.roche.com/home.html>). Zelo uporabni so SNP čipi, ki vsebujejo različice genov

naddružine citokromov P450, vključenih v presnovo zdravil. Z njimi lahko napovemo, ali bo posameznik hitro ali počasi presnavljal izbrano zdravilno učinkovino. Podatki o mutacijah v genih, ki so odgovorni za presnovo zdravil, so neprecenljivega pomena za kliniko, ki se dnevno srečuje s problemom različne učinkovitosti zdravil pri različnih posameznikih. Navzkrižno učinkovanje zdravil je za posameznika škodljivo in ima lahko usodne posledice.

V klinično prakso si utirajo pot tudi čipi za primerjalno genomsko hibridizacijo (CGH). Njihova uporabnost je usmerjena v detekcijo genomskih anomalij pri genetskih boleznih (kot so Downov sindrom, sindrom Prader-Willi in Angelman sindrom) in raku. Zaradi visoke robustnosti, občutljivosti in hitrosti metode lahko kmalu pričakujemo razvoj komercialno dostopnih diagnostičnih orodij in njihovo rutinsko uporabo v diagnostične namene.

### **Stanje tehnologije DNA čipov v Sloveniji**

Glede na naraščajoč interes in potrebe je bil junija 2001 ustanovljen Slovenski konzorcij za bio-čipe, h kateremu so pristopile slovenske akademske ustanove, klinične ustanove in farmacevtska industrija. Oprema za pripravo in analizo bio-čipov nizke gostote, kot tudi oprema za hibridizacijo čipov visoke gostote (tehnologija Affymetrix), ki je v lasti Slovenskega konzorcija za bio-čipe, je sedaj dostopna na Centru za funkcijsko genomiko in bio-čipe (CFGBC) na Medicinski fakulteti Univerze v Ljubljani (<http://cfgbc.mf.uni-lj.si>). Center je bil ustanovljen junija 2005 z namenom združiti infrastrukturo kot tudi interdisciplinarni kader (biološka in matematična znanja) za uporabo in kvaliteten razvoj tehnologije bio-čipov v Sloveniji.

V CFGBC smo slovenski raziskovalci Medicinske fakultete Univerze v Ljubljani, skupaj z raziskovalci Lek, d.d. in tujimi partnerji v okviru EU projekta STEROLTALK pripravili tematski cDNA čip Steroltalk, ki je usmerjen v študije homeostaze holesterola in presnove zdravil. Čip je

pripravljen v dveh različicah, in sicer za študije izražanja genov pri laboratorijski miški in pri človeku. Vsaka različica vsebuje 300 prob dolžine 300 do 500 bp, natisnjenih na stekleno ploščico. V prihodnosti naj bi se razvijal tudi v smeri diagnostike bolezni srca in ožilja.

## Literatura

1. Gresham D, Ruderfer DM, Pratt SC, et al.: Genome-wide detection of polymorphisms at nucleotide resolution with a single DNA microarray. *Science* 2006; 311(5769): 1932-1936.
2. Oostlander AE, Meijer GA, Ylstra B: Microarray-based comparative genomic hybridization and its applications in human genetics. *Clin Genet* 2004; 66(6): 488-495.
3. Buck MJ, Lieb JD: CHIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. *Genomics* 2004; 83(3): 349-360.
4. Churchill GA: Fundamentals of experimental design for cDNA microarrays. *Nat Genet* 2002; 32 Suppl: 490-495.
5. Kerr MK, Churchill GA: Experimental design for gene expression microarrays. *Biostatistics* 2001; 2(2): 183-201.
6. Yang YH, Speed T: Design issues for cDNA microarray experiments. *Nat Rev Genet* 2002; 3(8): 579-588.
7. Lee ML, Kuo FC, Whitmore GA, et al.: Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations. *Proc Natl Acad Sci U S A* 2000; 97(18): 9834-9839.
8. Brazma A, Hingamp P, Quackenbush J, et al.: Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet* 2001; 29(4): 365-371.
9. Spellman PT, Miller M, Stewart J, et al.: Design and implementation of microarray gene expression markup language (MAGE-ML). *Genome Biol* 2002; 3(9).
10. Whetzel PL, Parkinson H, Causton HC, et al.: The MGED Ontology: a resource for semantics-based description of microarray experiments. *Bioinformatics* 2006; 22(7): 866-873.
11. Gardiner-Garden M, Littlejohn TG: A comparison of microarray databases. *Brief Bioinform* 2001; 2(2): 143-158.
12. Anderle P, Duval M, Draghici S, et al.: Gene expression databases and data mining. *Biotechniques* 2003; Suppl: 36-44.
13. Parkinson H, Sarkans U, Shojatalab M, et al.: ArrayExpress--a public repository for microarray gene expression data at the EBI. *Nucleic Acids Res* 2005; 33(Database issue): D553-555.
14. Barrett T, Suzek TO, Troup DB, et al.: NCBI GEO: mining millions of expression profiles--database and tools. *Nucleic Acids Res* 2005; 33(Database issue): D562-566.
15. Ikeo K, Ishi-i J, Tamura T, et al.: CIBEX: center for information biology gene expression database. *C R Biol* 2003; 326(10-11): 1079-1082.
16. Van Driessche N, Demšar J, Booth EO, et al.: Epistasis analysis with global transcriptional phenotypes. *Nat Genet* 2005; 37(5): 471-477.
17. Avery L, Wasserman S: Ordering gene function: the interpretation of epistasis in regulatory hierarchies. *Trends Genet* 1992; 8(9): 312-316.
18. Juvan P: Artificial intelligence methods for discovery of relationships in genetic data. PhD thesis. Faculty of Computer and Information Science, University of Ljubljana. Ljubljana, Slovenia, 2005.
19. Golub TR, Slonim DK, Tamayo P, et al.: Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999; 286(5439): 531-537.
20. van 't Veer LJ, Dai H, van de Vijver MJ, et al.: Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; 415(6871): 530-536.