

Izvirni znanstveni članek ■

## Meta-analiza v tehnologiji genskih mikromrež

## Meta-Analysis in the Technology of DNA Microarrays

### Andrej Kastrin

**Izvleček.** Genske mikromreže so v zadnjih letih dobile primat rutinske metode za merjenje izraženosti genov. Kljub hitri rasti števila raziskav vprašanje ponovljivosti rezultatov posameznih poskusov ostaja odprto. S sistematično integracijo rezultatov primerljivih poskusov v obliki meta-analize povečamo velikost vzorca, kombinirana ocena velikosti učinka za proučevanje genski produkt pa je posledično bolj zanesljiva. Poveden je preprost način združevanja rezultatov mikromrežnih poskusov na osnovi statističnega modela s fiksnimi in slučajnimi vplivi. Uporaba modela je prikazana na integraciji treh različnih podatkovij izraženosti genov pri človeških zarodkih s trisomijo kromosoma 21. Uporaba, razvoj in implementacija metod za integracijo mikromrežnih podatkov je ključnega pomena za kakovost in posplošljivost rezultatov genomskih raziskav.

**Abstract.** With the explosion of microarray technology, an enormous amount of data is being generated. Systematic integration of gene expression data from different sources increases reliability of detecting differentially expressed genes. The challenge, however, is in designing and implementing efficient analytic methodologies for combination of data generated by different studies. We consider the meta-analysis of different microarray data sets using a fixed and random effect paradigm and demonstrate how relatively standard statistical approach yield promising results. We illustrated proposed method by integrating gene expression profiles from three different prenatal trisomy 21 studies. Our results show that this approach would serve as a plausible method for analyzing microarrays beyond the specific implications for trisomy 21.

---

Institucija avtorja: Inštitut za medicinsko genetiko, Univerzitetni klinični center Ljubljana.

Kontaktna oseba: Andrej Kastrin, Inštitut za medicinsko genetiko, Univerzitetni klinični center Ljubljana, Šlajmerjeva 4, 1000 Ljubljana. email: andrej.kastrin@guest.arnes.si.

■ **Infor Med Slov:** 2007; 12(2): 14-21

## Uvod

Z vprašanjem zanesljivosti se srečujemo v vseh znanstvenih disciplinah, ki poskušajo svoje raziskovalne domneve preveriti z empiričnimi izsledki. Zanesljivi rezultati so ključnega pomena za doseganje osnovnega cilja znanstvenega raziskovanja, t.j. ugotavljanja zakonitosti, ki nam omogočajo pojasnjevanje in napovedovanje opazovanih pojavov. Zanesljivost v širšem smislu pomeni, da bomo s ponavljanjem meritev istega pojava v istih okoliščinah dobili podobne rezultate. Zanesljivost merjenja je tem večja, čim bolj so razlike v izmerjenih vrednostih posledica dejanskih sprememb merjenega pojava in čim manjši je vpliv slučajnih dejavnikov.

Razmerje med eksplozivno uporabo tehnologije genskih mikromrež ter potrebo po implementaciji računskih modelov in statističnih metod za njihovo analizo se je v zadnjih letih ustalilo.<sup>1,2</sup> Kljub temu, da so v bioloških znanostih mikromreže dobile primat rutinske metode za merjenje izraženosti genov in njihovih produktov, vprašanje ponovljivosti rezultatov posameznih poskusov ostaja odprto. Na izmerjeno vrednost izraženosti genskega produkta namreč poleg dejanske, biološke vrednosti, vplivajo tudi sistematične napake (npr. način normalizacije), slučajne napake ter napake, ki so sicer slučajne glede na dejansko vrednost, a korelirajo med seboj pri ponovljenih meritvah.<sup>3</sup>

Smiseln odgovor na problem zanesljivosti posameznih poskusov ponuja njihova integracija v obliki meta-analize. Gre za statistično analizo, v kateri na sistematičen način združujemo rezultate posameznih med seboj (ne)odvisnih poskusov.<sup>4</sup> Ideja o združevanju podatkov večih neodvisnih raziskav je Fisherjeva, temelje moderne meta-analize pa je pred slabimi tremi desetletji postavil Glass, ko je na metodološko rigorozen način ovrzel smelo Eysenckovo tezo o ničnosti učinka psihoterapije.<sup>4</sup> Ustrezno izvedena meta-analiza (i) ponuja sistematične, hitre in zanesljive odgovore na raziskovalne domneve, (ii) zaradi večje količine podatkov povečuje statistično moč zaključevanja,

(iii) daje pregled nad metodologijo izvedbe posameznih poskusov ter nenazadnje (iv) omogoča velik prihranek sredstev na račun ponovitvenih poskusov.

Termin meta-analiza morda v kontekstu analize mikromrežnih podatkovij nekoliko zavaja, saj v primerjavi s klasičnimi meta-analitičnimi študijami ne združujemo končnih rezultatov posameznih poskusov v skupno oceno, pač pa le to sestavimo na osnovi ponovne analize surovih rezultatov. Na ta način normaliziramo razlike med poskusi, ki so posledica uporabe različnih tipov mikromrež na genomih oz. transkriptomih enakih ali različnih organizmov, različnih protokolov hibridizacije, odčitavanja in analize rezultatov. Velikost vzorca in s tem statistična moč se povečata, kombinirana ocena velikosti učinka za proučevan genski produkt pa je bolj zanesljiva, kar posredno vpliva tudi na večjo klasifikacijsko točnost merskega instrumenta.<sup>3</sup> Glavna problema, na katera naletimo pri taki vrsti analize, sta (i) definiranje univerzuma genov, ki so skupni vsem poskusom ter (ii) opredelitev cenilke, ki ustrezno povzame informacijo posameznih poskusov v skupno oceno.

Kljub razmeroma dolgi tradiciji uporabe meta-analize v biomedicini je njena aplikacija na področju mikromrež še v povojih.<sup>5-20</sup> Prvi korak v tej smeri je naredil Rhodes,<sup>5</sup> ki je združil profile genske izraženosti pri bolnikih z rakom na osnovi Fisherjeve metode združevanja *p*-vrednosti. Surova statistična pomembnost izračunana za posamezen gen pa nam ničesar ne pove o njegovi praktični pomembnosti. Zato je bila kasneje izvedena serija praktičnih eksperimentov in teoretičnih simulacij z različnimi merami velikosti učinka.<sup>6,9</sup> Trenutno sta v središču pozornosti razvoj in implementacija metod, ki temeljijo na Bayesovih hierarhičnih modelih.<sup>19,21</sup>

V nadaljevanju prispevka obravnavamo preprosto in hitro Choievo<sup>6</sup> metodo združevanja rezultatov mikromrežnih poskusov na osnovi statističnega modela s fiksnimi in slučajnimi vplivi.

## Metoda

### Statistični model

V poskusu  $i=1,2,\dots,k$  obravnavamo proučevano ( $t$ ) in kontrolno ( $c$ ) skupino. Z  $n_{it}$  in  $n_{ic}$  označimo število vzorcev po skupinah v poskusu  $i$ . Kot mero razlike v izraženosti gena oz. zaporedja med proučevano in kontrolno skupino v poskusu  $i$  definiramo cenilko velikosti učinka:

$$d_i = \frac{\bar{X}_{it} - \bar{X}_{ic}}{S_{ip}}, \quad (1)$$

kjer sta v  $\bar{X}_{it}$  in  $\bar{X}_{ic}$  povprečni vrednosti izraženosti gena v proučevani in kontrolni skupini v poskusu  $i$ ,  $S_{ip}$  pa skupni standardni odklon. Pokazati se da, da je ta ocena na majhnih vzorcih pozitivno pristrana.<sup>4</sup> Zato uporabimo popravek za končnost populacije:

$$d'_i = d_i - \frac{3d_i}{4(n-2)-1}. \quad (2)$$

V nadaljevanju privzemimo, da z  $d_i$  označujemo nepristrano popravljeno oceno. Njeno varianco izračunamo po enačbi:

$$\hat{\sigma}_{d_i}^2 = \left( \frac{1}{n_{it}} + \frac{1}{n_{ic}} \right) + \frac{d_i^2}{2(n_{it} + n_{ic})}. \quad (3)$$

Za gen  $g$  naj  $d_i$  označuje izmerjeno vrednost dejanskega učinka  $\theta_i$  v poskusu  $i$ ,  $\mu$  pa njegovo dejansko velikost učinka, t.j. verjetnostno limito, ki se ji približuje aritmetična sredina velikosti učinkov  $d_i$  za gen  $g$  med skupinama  $t$  in  $c$ , če število poskusov  $k$  narašča čez vse meje. Velikost učinka za gen  $g$  v poskusu  $i$  zapišemo v obliki dvostopenjskega hierarhičnega modela:<sup>6</sup>

$$\begin{cases} d_i = \theta_i + \varepsilon_i, & \varepsilon_i \sim N(0, s_i^2) \\ \theta_i = \mu + \delta_i, & \delta_i \sim N(0, \tau^2) \end{cases} \quad (4)$$

Prva stopnja opisuje fiksni, druga pa slučajni model učinka. V modelu je s  $\tau^2$  označena varianca

med poskusi,  $s_i^2$  pa označuje varianco znotraj poskusov, ki jo ocenimo po enačbi (3). V modelu s fiksnimi vplivi je homogenost med poskusi popolna, zato velja  $\theta_1=\theta_2=\dots,\theta_n$ . Komponento napake v celoti pojasnimo s slučajnimi napakami in zanemarimo variabilnost med poskusi, torej velja  $\tau^2=0$  in  $d_i \sim N(\mu, s_i^2)$ . V modelu s slučajnimi vplivi pa upoštevamo tudi variabilnost med poskusi, zato velja  $d_i \sim N(\theta_i, s_i^2)$  in  $\theta_i \sim N(\mu, \tau^2)$ . Veljavnost modela s fiksnimi vplivi preverimo s testiranjem domneve  $\tau^2=0$ . V ta namen uporabimo Cochranovo testno statistiko:<sup>21</sup>

$$Q = \sum w_i (d_i - \hat{\mu})^2, \quad (5)$$

$$\text{kjer sta } w_i = \frac{1}{s_i^2} \text{ in } \hat{\mu} = \frac{\sum w_i d_i}{\sum w_i}.$$

Ničelna porazdelitev testne statistike sledi  $\chi^2_{k-1}$  porazdelitvi. Če ničelno domnevo zavrnilo, ocenimo  $\tau^2$  po obrazcu, ki sta ga predlagala DerSimonian in Laird:<sup>22</sup>

$$\tau^2 = \max \left\{ 0, \frac{Q - (k-1)}{\sum w_i - \sum w_i^2 / \sum w_i} \right\}. \quad (6)$$

Oceno dejanske vrednosti velikosti učinka  $\mu$  izračunamo kot:

$$\hat{\mu} = \frac{\sum (s_i^2 + \tau^2)^{-1} d_i}{\sum (s_i^2 + \tau^2)^{-1}}, \quad (7)$$

njeno varianco pa po enačbi:

$$\text{Var}(\hat{\mu}) = \frac{1}{\sum (s_i^2 + \tau^2)^{-1}}. \quad (8)$$

Vrednosti statistik za model s fiksnimi učinki izračunamo enako, upoštevamo pa, da velja  $\tau^2=0$ . Oceno velikosti učinka za gen  $g$  v  $k$  poskusih standardiziramo in izrazimo v  $z$ -vrednostih:

$$z_g = \frac{\hat{\mu}}{\text{Var}(\hat{\mu})} \quad (9)$$

Standardizirane vrednosti se porazdeljujejo normalno, zato nivo statistične značilnosti gena  $g$  izpeljemo po standardnem postopku.

### Opis eksperimenta

Statistični model smo uporabili za integracijo rezultatov izraženosti genov pri človeških zarodkih s trisomijo kromosoma 21 (Downov sindrom). Ob pregledu spletnega podatkovnega skladišča mikromrežnih podatkov Gene Expression Omnibus<sup>23</sup> so vključitvenemu kriteriju ustrezali trije poskusi: GSE1397,<sup>24</sup> GSE1789<sup>25</sup> in GSE6283. Poskusi so bili opravljeni na celicah tkiv velikih in malih možganov, srca, amniona in horionskih resic ter v različnih razvojnih obdobjih zarodka. Z izjemo poskusa GSE1397 so bili vzorci znotraj poskusov med seboj neodvisni. Vsi poskusi so bili izvedeni v okolju Affymetrixovih genomskih biočipov (HG-U133A oz. HG-U133 Plus 2.0). Za normalizacijo surovih podatkov je bil uporabljen mas5 algoritem.<sup>26</sup>

Da se izognemo nevšečnostim, ki nas doletijo pri uporabi  $p$ -vrednosti v analizi mikromrežnih podatkovij, smo izračunali še recipročne  $q$ -vrednosti, pri katerih rezultat izrazimo z deležem napačno pozitivnih zadetkov (*angl.* false discovery rate) med dejansko različno izraženimi zaporedji med proučevano in kontrolno skupino.<sup>27-29</sup>

Statistični model smo implementirali v okolju R.<sup>30</sup> Programska koda v Sweave zapisu je dostopna na spletnem naslovu <http://www2.arnes.si/~akastr1/>.

### Rezultati

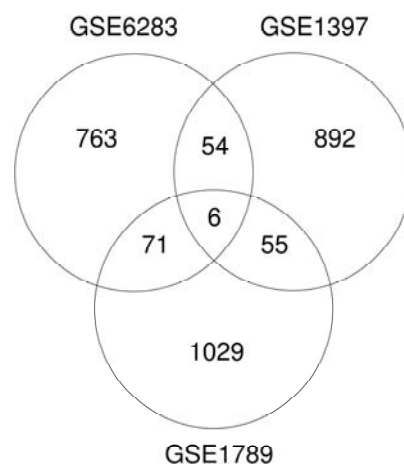
V meta-analizo je bilo vključenih 55 vzorcev; 27 zarodkov s trisomijo kromosoma 21 in 28 kontrolnih zarodkov (Tabela 1). Uporabljenim okoljem je bilo skupnih 22277 zaporedij.

**Tabela 1** Opisne statistike poskusov vključenih v meta-analizo.

GEO	Okolje	$n_t$	$n_c$
GSE1397	HG-U133A	11	14
GSE1789	HG-U133A	10	5
GSE6283	HG-U133A Plus 2.0	6	9

Pojasnilo: GEO - identifikator poskusa;  $n_t$  - število proučevanih zarodkov;  $n_c$  - število kontrolnih zarodkov.

Pred aplikacijo predlaganega statističnega modela smo najprej preverili prosto ujemanje statistično značilno izraženih zaporedij med poskusi. V ta namen smo znotraj vsakega poskusa zaporedjem priredili ocene velikosti učinka, jih izrazili v standardiziranih  $z$ -vrednostih, izračunali gostoto verjetnosti ter zaporedja uredili po stopnji statistične značilnosti. Vennov diagram (Slika 1) prikazuje število statistično značilno izraženih zaporedij znotraj poskusov in med poskusi. Na nivoju tveganja  $\alpha=0.05$  je v preseku šest skupnih zaporedij, na nivoju tveganja  $\alpha=0.01$  pa je presek prazna množica.



**Slika 1** Vennov diagram prikazuje število statistično značilno izraženih zaporedij znotraj poskusov ter njihovo ujemanje med poskusi ( $\alpha=0.05$ ).

Veljavnost modela s fiksnim učinkom najlaže preverimo s Q-Q grafikonom (Slika 2), ki prikazuje odnos med dejansko in pričakovano porazdelitvijo Cochranove testne statistike. Prednost takega prikaza pred uvedbo dodatne testne statistike je v tem, da laže vidimo, kje in v kolikšni meri dejanska porazdelitev odstopa od

pričakovane. V našem primeru se odstopanje med porazdelitvama povečuje z večanjem dejanskih  $Q$  vrednosti, zato veljavnost modela s fiksnimi učinki zavrnamo.

Standardiziranim vrednostim velikosti učinka za posamezna zaporedja priredimo gostote verjetnosti ( $p$ -vrednosti). V Tabeli 2 je povzeto število statistično značilnih zaporedij za izbrane stopnje tveganja  $\alpha$ . Na nivoju tveganja  $\alpha=0.01$  izločimo 231 zaporedij, za katere trdimo da se statistično značilno razlikujejo med proučevanimi in kontrolnimi zarodki. Ob bolj liberalnem kriteriju ( $\alpha=0.05$ ) je takih zaporedij 1071.

Ujemanje med zaporedji izbranimi po povedenem statističnem modelu s slučajnimi vplivi ter posameznimi poskusi najlaže predstavimo grafično s pomočjo toplotnega grafikona (Slika 3). Vrstice v matriki odgovarjajo zaporedjem, ki smo jih opredelili kot statistično značilna znotraj vsaj enega poskusa. Na nivoju tveganja  $\alpha=0.01$  dobimo 711 takih zaporedij. Toplotni grafikon odkriva največje prekrivanje s poskusom GSE1789; ujemanje z ostalima poskusoma je manjše. Takšna primerjava je smiselna predvsem zaradi tega, ker je meta-analiza občutljiva na velikosti učinkov, ki so med poskusi stalni in ne nujno veliki. Praktično to pomeni, da bo v meta-analizi višji rezultat doseglo zaporedje, pri katerem je ocenjena velikost učinka približno enaka prek vseh treh poskusov, kot pa zaporedje, ki je močno izraženo le v enem poskusu.

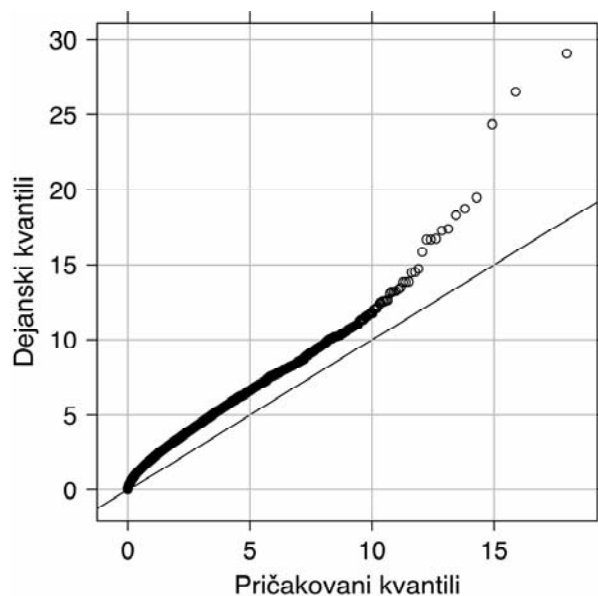
**Tabela 2** Število statistično značilno izraženih zaporedij za izbrane stopnje tveganja  $\alpha$ .

$\alpha$	0.001	0.01	0.025	0.05	0.1
$p$ -vrednost	48	231	559	1071	2201
$q$ -vrednost	0	3	5	14	21

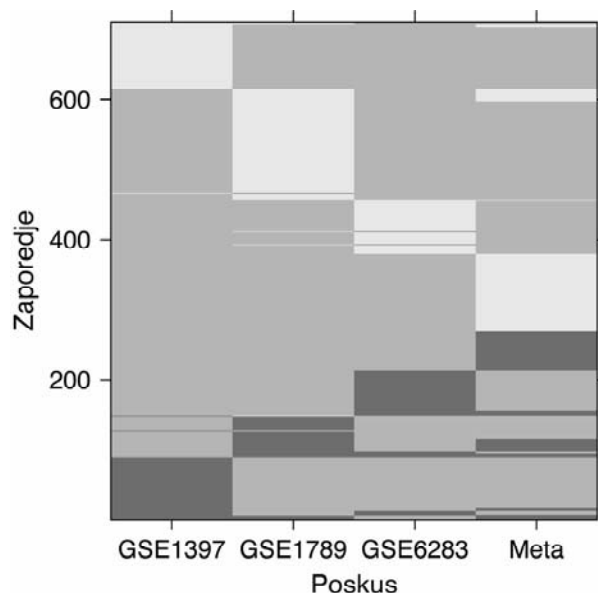
Pojasnilo: Za podrobnosti o  $p$  in  $q$ -vrednostih glej besedilo.

Ob isti stopnji tveganja je med 14 izbranimi zaporedji 5% lažno pozitivnih zadetkov (Tabela 2). Izbrana množica zaporedij, odgovarjajoči geni, standardizirane velikosti učinkov ter natančne  $q$ -vrednosti so povzete v Tabeli 3. 14 zaporedij

ustreza 13 genom na kromosomu 21; vsi geni so nadpovprečno izraženi.



**Slika 2** Q-Q grafikon za preizkus veljavnosti modela s fiksnim učinkom.



**Slika 3** Toplotni grafikon. Črne proge se nanašajo na nadpovprečno izražena, bele pa na podpovprečno izražena zaporedja. Sive proge označujejo statistično neznačilna zaporedja. Za podrobnosti glej besedilo.

**Tabela 3** Množica genov izbrana na osnovi rangiranih  $q$ -vrednosti ( $\alpha=0.05$ ).

Zaporedje	Gen	$z$	$q$
201086_x_at	SON	-5.380	0.002
200642_at	SOD1	-5.204	0.005
214988_s_at	SON	-4.952	0.005
202671_s_at	PDXK	-4.723	0.013
202217_at	C21orf33	-4.580	0.021
200944_s_at	HMGNI	-4.431	0.030
219767_s_at	CRYZL1	-4.428	0.030
201644_at	TSTA3	-4.292	0.049
202749_at	WRB	-4.246	0.049
218386_x_at	USP16	-4.262	0.049
200740_s_at	SUMO3	-4.204	0.049
213000_at	MORC3	-4.192	0.049
211065_x_at	PFKL	-4.185	0.049
216954_x_at	ATP5O	-4.165	0.049

Pojasnilo: Oznaka zaporedja se nanaša na Affymetrixov identifikator.

## Razprava

Integracija empiričnih izsledkov je bistven del sodobne statistične analize mikromrežnih podatkov, s tem pa tudi raziskovalnega in strokovnega dela v biologiji in njenih sorodnih področjih. Prikazan primer uporabe meta-analize za integracijo mikromrežnih podatkov je preprosto izračunljiv, povedeni rezultati pa neposredno razložljivi, bodisi na nivoju genov, bodisi na nivoju ontologij (npr. Gene Ontology). Zaradi metodološke narave prispevka se v biološko interpretacijo rezultatov ne spuščamo.

Nujnost uporabe integrativnega pristopa implicirajo tako rezultati simulacijskih eksperimentov kot tudi klinične raziskave, ki pri preverjanju enakih raziskovalnih domnev in ob primerljivih sospremenljivkah prihajajo do nasprotujočih si ugotovitev. Ein-Dor<sup>31</sup> je npr. dokazal, da je stabilnost razvrstitve genov kandidatov znotraj enega poskusa močno odvisna od izbrane kombinacije proučevanih in kontrolnih vzorcev. Michiels<sup>32</sup> pa pred interpretacijo in pripisovanjem biološkega pomena takim rezultatom celo priporoča uporabo vzorčenja z vračanjem. Klinični primer zmede v raziskovalnih

izsledkih sta nedavno objavljena mikromrežna poskusa, ki sta ugotavljala možne bioznačevalce poteka razvoja Huntingtonove bolezni in v primerljivih pogojih generirala neprimerljive izsledke.<sup>33,34</sup>

Prosto dostopna skladišča mikromrežnih podatkov (npr. GEO,<sup>23</sup> ArrayExpress,<sup>35</sup> SMD,<sup>36</sup> CIBEX<sup>37</sup>) odpirajo nove možnosti integrativnemu pristopu v raziskovanju mikromrežnih podatkov. Kljub temu, da so najbolj vplivne revije (<http://www.mged.org/Workgroups/MIAME/journals.html>) že pred časom kot pogoj za objavo znanstvenega prispevka zahtevale dosledno upoštevanje MIAME standarda (*angl.* Minimum Information About a Microarray Experiment), pa je kvaliteta povzetrov o uporabljenih metodah obdelave podatkov pogosto nezadovoljiva.<sup>38</sup> Standard jasno zahteva objavo surovih podatkov (npr. CEL ali GPR datoteke), raziskovalne skupine pa pogosto objavljajo zgolj normalizirane rezultate, kar onemogoča kvalitetno integracijo podatkov prek različnih poskusov.

Glavna pomanjkljivost študije je omejena posplošljivost rezultatov, ki izhajajo iz aplikacije modela s slučajnimi napakami na izbranem vzorcu poskusov.<sup>4,6,39</sup> Kljub omejitvam aparata statističnega sklepanja zaradi majhnega števila poskusov in naivni predpostavki neodvisnosti vzorcev v enem od poskusov,<sup>24</sup> smo močno skrčili množico genov kandidatov, aktivnih v razvoju humanega zarodka s trisomijo kromosoma 21. Izkopano znanje in predstavljen metodološki okvir lahko služita kot odskočna deska za nadaljno raziskovalno delo na področju genomike slučajnih kromosomskih nepravilnosti združljivih z življenjem.

Statistika se kljub oviram in težavam, ki izhajajo iz njenega izrazito interdisciplinarnega področja delovanja, čedalje bolj uspešno integrira tako v teoretične kot aplikativne biološke vede. Uporaba, nadaljnji razvoj in implementacija metod za integracijo mikromrežnih podatkov lahko bistveno pripomorejo k večji kakovosti, razložljivosti in posplošljivosti rezultatov genomskih raziskav.

## Literatura

1. He YD: Genomic approach to biomarker identification and its recent applications. *Cancer Biomark* 2006; 2(3-4): 103-133.
2. Phan JH, Quo CF, Wang MD: Functional genomics and proteomics in the clinical neurosciences: data mining and bioinformatics. *Prog Brain Res* 2006; 158: 83-108.
3. He W, Bull SB, Gokgoz N, Andrulis I, Wunder J: Application of reliability coefficients in cDNA microarray data analysis. *Stat Med* 2006; 25(6): 1051-1066.
4. Whitehead A: *Meta-Analysis of Controlled Clinical Trials*. Chichester 2002: Wiley.
5. Rhodes DR, Barrette TR, Rubin MA, Ghosh D, Chinnaiyan AM: Meta-analysis of microarrays: interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. *Cancer Res* 2002; 62(15): 4427-4433.
6. Choi JK, Yu U, Kim S, Yoo OJ: Combining multiple microarray studies and modeling interstudy variation. *Bioinformatics* 2003; 19 Suppl 1: i84-i90.
7. Ghosh D, Barrette TR, Rhodes D, Chinnaiyan AM: Statistical issues and methods for meta-analysis of microarray data: a case study in prostate cancer. *Funct Integr Genomics* 2003; 3(4): 180-188.
8. Jiang H, Deng Y, Chen H, Tao L, Sha Q, Chen J, Tsai C, Zhang S: Joint analysis of two microarray gene-expression data sets to select lung adenocarcinoma marker genes. *BMC Bioinformatics* 2004; 5: 81.
9. Parmigiani G, Garrett-Mayer ES, Anbazhagan R, Gabrielson E: A cross-study comparison of gene expression studies for the molecular classification of lung cancer. *Clin Cancer Res* 2004; 10(9): 2922-2927.
10. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, Pandey A, Chinnaiyan AM: Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. *Proc Natl Acad Sci U S A* 2004; 101: 9309-9314.
11. Shen R, Ghosh D, Chinnaiyan AM: Prognostic meta-signature of breast cancer development by two-stage mixture modeling of microarray data. *BMC Genomics* 2004; 5(1): 94.
12. Wang J, Coombes KR, Highsmith WE, Keating MJ, Abruzzo LV: Differences in gene expression between B-cell chronic lymphocytic leukemia and normal B cells: a meta-analysis of three microarray studies. *Bioinformatics* 2004; 20(17): 3166-3178.
13. Hu P, Greenwood CMT, Beyene J: Integrative analysis of multiple gene expression profiles with quality-adjusted effect size models. *BMC Bioinformatics* 2005; 6: 128.
14. Stevens JR, Doerge RW: Combining Affymetrix microarray results. *BMC Bioinformatics* 2005; 6: 57.
15. Warnat P, Eils R, Brors B: Cross-platform analysis of cancer microarray data improves gene expression based classification of phenotypes. *BMC Bioinformatics* 2005; 6: 265.
16. Xu L, Tan AC, Naiman DQ, Geman D, Winslow RL: Robust prostate cancer marker genes emerge from direct integration of inter-study microarray data. *Bioinformatics* 2005; 21(20): 3905-3911.
17. Conlon EM, Song JJ, Liu JS: Bayesian models for pooling microarray studies with multiple sources of replications. *BMC Bioinformatics* 2006; 7: 247.
18. Park T, Yi SG, Shin YK, Lee S: Combining multiple microarrays in the presence of controlling variables. *Bioinformatics* 2006; 22: 1682-1689.
19. Conlon EM, Song JJ, Liu A: Bayesian meta-analysis models for microarray data: a comparative study. *BMC Bioinformatics* 2007; 8: 80.
20. Fishel I, Kaufman A, Ruppin E: Meta-analysis of gene expression data: a predictor-based approach. *Bioinformatics* 2007; 23(13): 1599-1606.
21. Cochran BG: The combination of estimates from different experiments. *Biometrics* 1954; 10: 101-129.
22. DerSimonian R, Laird N: Meta-analysis in clinical trials. *Control Clin Trials* 1986; 7(3): 177-188.
23. Barrett T, Troup DB, Wilhite SE, et al.: NCBI GEO: mining tens of millions of expression profiles—database and tools update. *Nucleic Acids Res* 2007; 35(Database issue): D760-D765.
24. Mao R, Wang X, Spitznagel EL, et al.: Primary and secondary transcriptional effects in the developing human Down syndrome brain and heart. *Genome Biol* 2005; 6(13): R107.
25. Conti A, Fabbrini F, D'Agostino P, et al.: Altered expression of mitochondrial and extracellular matrix genes in the heart of human fetuses with chromosome 21 trisomy. *BMC Genomics* 2007; 8: 268.
26. Irizarry RA, Hobbs B, Collin F, et al.: Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003; 4(2): 249-264.

27. Storey JD, Tibshirani R: Statistical significance for genomewide studies. *Proc Natl Acad Sci U S A* 2003; 100(16): 9440-9445.
28. Pyne S, Futcher B, Skiena S. Meta-analysis on control of false discovery rate: combining yeast ChIP-chip datasets. *Bioinformatics* 2006; 22(20): 2516-2522.
29. Pawitan Y, Michiels S, Koscielny S, Gusnanto A, Ploner A: False discovery rate, sensitivity and sample size for microarray studies. *Bioinformatics* 2005; 21(13): 3017-3024.
30. R Development Core Team: *R: A language and environment for statistical computing*. Vienna, Austria 2007: R Foundation for Statistical Computing.
31. Ein-Dor L, Kela I, Getz G, Givol D, Domany E: Outcome signature genes in breast cancer: is there a unique set? *Bioinformatics* 2005; 21(2): 171-178.
32. Michiels S, Koscielny S, Hill C: Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet* 2005; 365(9458): 488-492.
33. Borovecki F, Lovrecic L, Zhou J, et al.: Genome-wide expression profiling of human blood reveals biomarkers for Huntington's disease. *Proc Natl Acad Sci U S A* 2005; 102(31): 11023-11028.
34. Runne H, Kuhn A, Wild EJ, et al.: Analysis of potential transcriptomic biomarkers for Huntington's disease in peripheral blood. *Proc Natl Acad Sci U S A* 2007; 104(36): 14424-14429.
35. Parkinson H, Kapushesky M, Shojatalab M, et al.: ArrayExpress—a public database of microarray experiments and gene expression profiles. *Nucleic Acids Res* 2007; 35(Database issue): D747-D750.
36. Ball CA, Awad IAB, Demeter J, et al.: The Stanford Microarray Database accommodates additional microarray platforms and data formats. *Nucleic Acids Res* 2005; 33(Database issue): D580-D582.
37. Ikeo K, Ishi-i J, Tamura T, Gojobori T, Tateno Y: CIBEX: center for information biology gene expression database. *C R Biol* 2003; 326(10-11): 1079-1082.
38. Imbeaud S, Auffray C: 'The 39 steps' in gene expression profiling: critical issues and proposed best practices for microarray experiments. *Drug Discov Today* 2005; 10(17): 1175-1182.
39. Ades AE, Lu G, Higgins JPT: The interpretation of random-effects meta-analysis in decision models. *Med Decis Making* 2005; 25(6): 646-654.