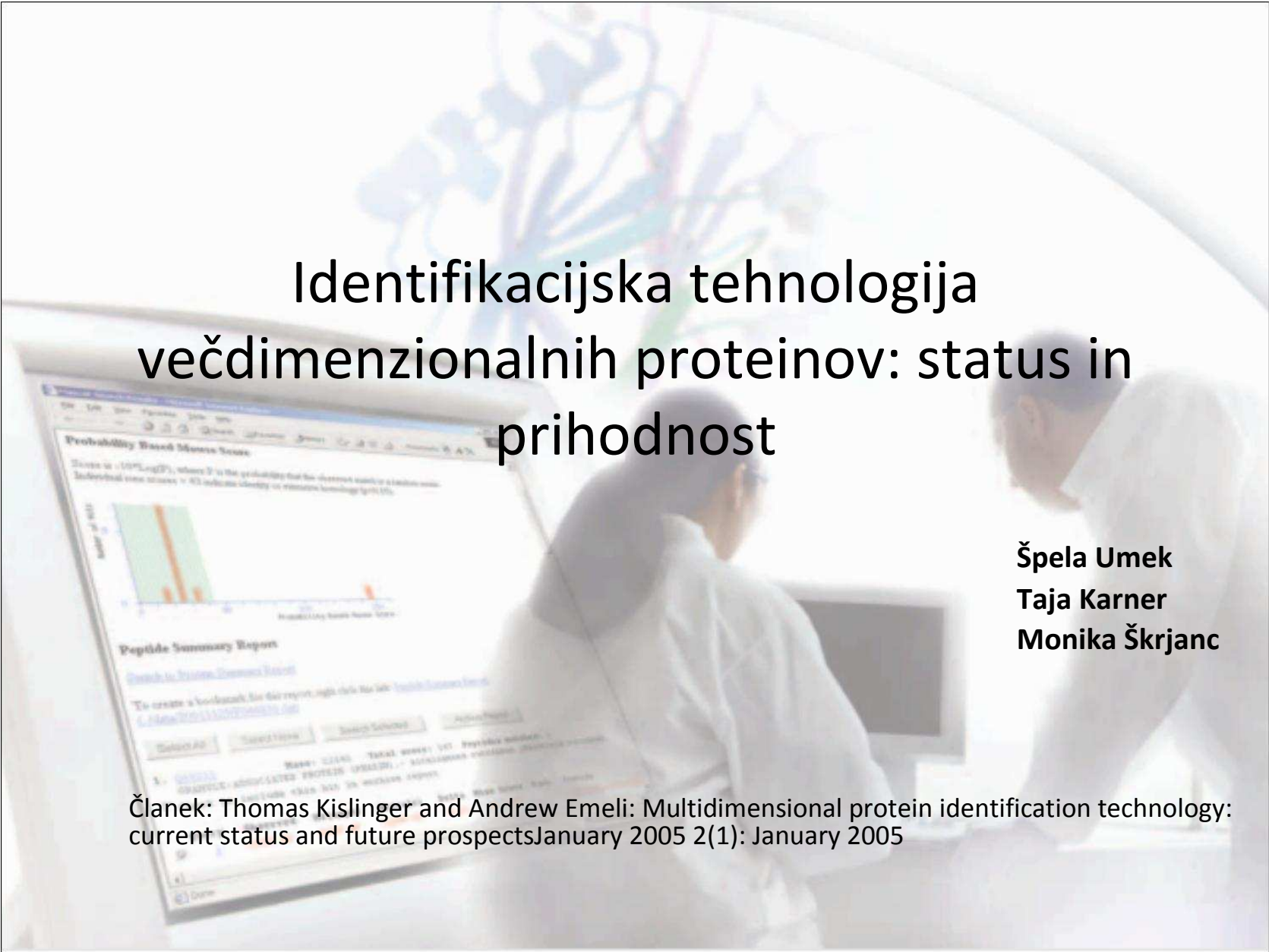




Identifikacijska tehnologija večdimenzionalnih proteinov: status in prihodnost

Špela Umek
Taja Karner
Monika Škrjanc

Članek: Thomas Kislinger and Andrew Emeli: Multidimensional protein identification technology: current status and future prospects January 2005 2(1): January 2005

- profiliranje proteinov z tandemsko masno spektrometrijo
- razumevanje razvojnih, fizioloških in bolezenskih procesov
- osnova za obsežno proteinsko profiliranje kompleksnih bioloških vzorcev
- šibka stran - analiza podatkov

MassPep

File Options

Select Digest: Trypsin, CnBr, Pepsin pH 1.3, Pepsin > pH 2

Size Range: 1000, 1500, 2000, 3000

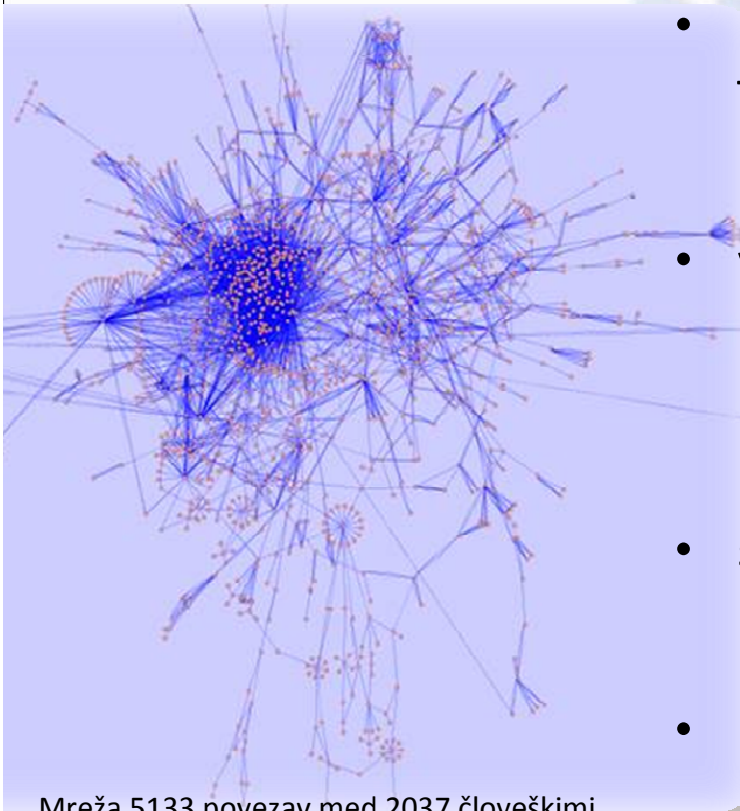
Epitope: 8, 9, 10, 11

Main

Sequence Manager

Organism	pl	Hydro. In.	Hydro.	Accession	Digest M.	Mass (D)	Sequence	Location	Selected	Descript.
Bos taur...	3.18		22.222	S07103	CnBr	1,746.603	ENGAC	31	☒	cOMP-g
Bos taur...	9.735		41.667	S07103	CnBr	2,975.516	EFFGR	280	☒	cOMP-g
Bos taur...	9.545		60.507	103	CnBr	2,040.067	NAFAE	402	☒	cOMP-g
Bos taur...	9.872		33.333	S07103	CnBr	1,085.549	HFRAVS	420	☒	cOMP-g
Bos taur...	3.907		41.667	S07103	CnBr	1,276.606	EALTEY	581	☒	cOMP-g
Bos taur...	6.578		40.507	103	CnBr	1,140.662	LEEKOK	593	☒	cOMP-g
Bos taur...	4.378		38.462	S07103	CnBr	2,794.459	VDGLD	603	☒	cOMP-g
Bos taur...	4.244		42.857	S07103	CnBr	2,410.225	SSVDL	629	☒	cOMP-g
Bos taur...	11.647		80.182	IRBP_B	CnBr	1,178.741	VRKWL	2	☒	Interphot.
Bos taur...	5.415		60.182	IRBP_B	CnBr	2,091.131	LLCOLT	12	☒	Interphot.
Bos taur...	3.576		53.333	IRBP_B	CnBr	1,721.82	AVVLLD	32	☒	Interphot.
Bos taur...	6.501		40.182	IRBP_B	CnBr	1,633.808	SDQWR	692	☒	Interphot.
Bos taur...	3.329		52.381	IRBP_B	CnBr	2,097.997	SASTOE	898	☒	Interphot.
Bos taur...	3.685		52.632	IRBP_B	CnBr	1,983.032	VLTSTL	1138	☒	Interphot.
Bos taur...	9.67		50.182	IRBP_E	CnBr	1,539.864	VVEQL	77	☒	Chain E...
Bos taur...	4.622		43.75	OBQ1_B	CnBr	1,838.027	PAVNEE	2	☒	Guanine...
Bos taur...	4.606		35.714	OBQ1_B	CnBr	1,668.912	EVDDLK	18	☒	Guanine...
Bos taur...	12.023		25.0	AAS5942	CnBr	1,800.944	HCSLR	192	☒	anyValyl...
Bos taur...	5.674		23.81	OCA1_B	CnBr	2,484.15	DGKSVL	6	☒	Guanylyl...
Bos taur...	3.721		40.0	OCA1_B	CnBr	1,840.806	FETDF	60	☒	Guanylyl...
Bos taur...	3.483		37.037	OCA1_B	CnBr	2,975.366	TAGEET	131	☒	Guanylyl...
Bos taur...	10.479		15.789	IAOR_P	CnBr	2,236.106	SSPOS	54	☒	Chain P...
Bos taur...	12.519		37.5	IB9X_A	CnBr	1,895.12	RTTRTL	46	☒	Chain A...
Bos taur...	6.314		53.333	IB9X_A	CnBr	1,648.85	AVATGS	326	☒	Chain A...
Bos taur...	3.895		84.4	AF43490	CnBr	2,632.499	TEECAL	31	☒	inwardly...
Bos taur...	11.475		40.4	AF43490	CnBr	1,160.653	FRVQL	92	☒	inwardly...
Bos taur...	10.387		33.333	AGCR	CnBr	2,388.102	KEMLRI	70	☒	Gamma...

Number of peptides generated 347



Mreža 5133 povezav med 2037 človeškimi proteini, povezanimi s transkripcijskim faktorjem NF- κ B, ki se aktivira ob vnetju. Povezave so predstavljene s programom Cytoscape 2.1.

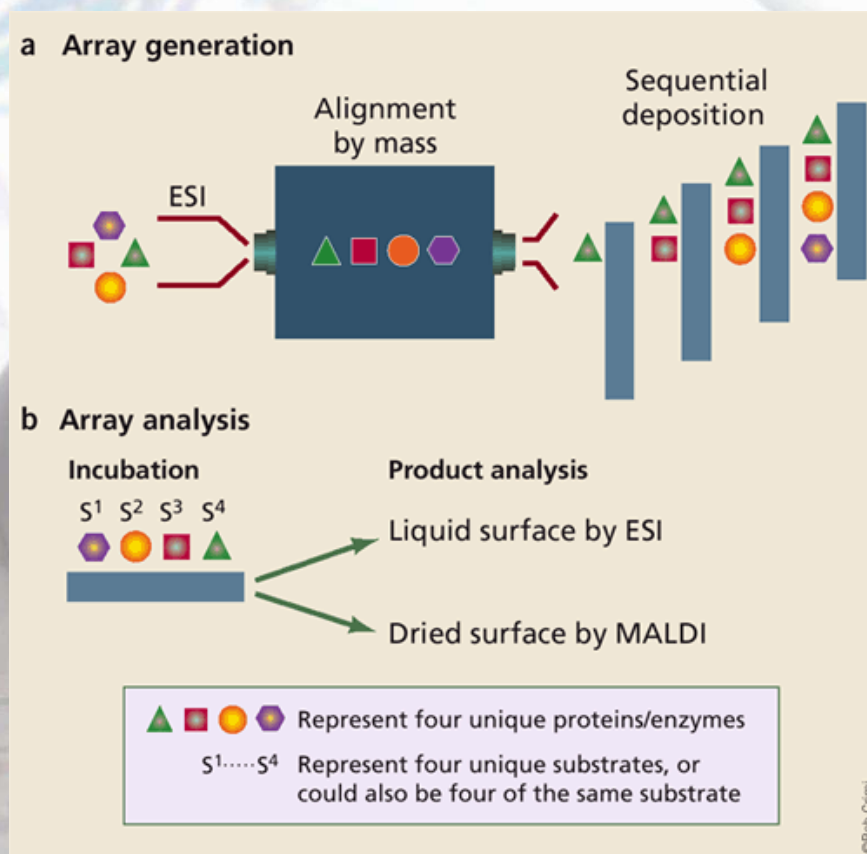
Vir:

http://www.kvarkadabra.net/images/articles/Biokemija-cez-10-let_3_original.jpg

- »post-genomna« era sistemske biologije – izziv: globalna analiza genske in proteinske ekspresije
- večina evkariontskih proteinov je podvržena posttranskripcijski modifikaciji in/ali posttranslacijskim modifikacijam
- spremenijo aktivnost, interakcije in lokalizacijo proteina v celici
- poslabšajo detekcijo z MS
- zelo problematično za detekcijo nizko koncentriranih proteinov v prisotnosti višje koncentriranih proteinov (npr. hišnih encimov)

Tekočinska kromatografija - MS

- uvedbo mehke ionizacijske tehnike v 80-ih
- razvoj ionizacije v matriksu z desorpcijo z laserjem (MALDI) in elektrosprejne ionizacije (ESI)
- ESI uporabljajo v negelski analizi kompleksih proteinskih mešanic, zraven katere se uporablja MudPIT metoda



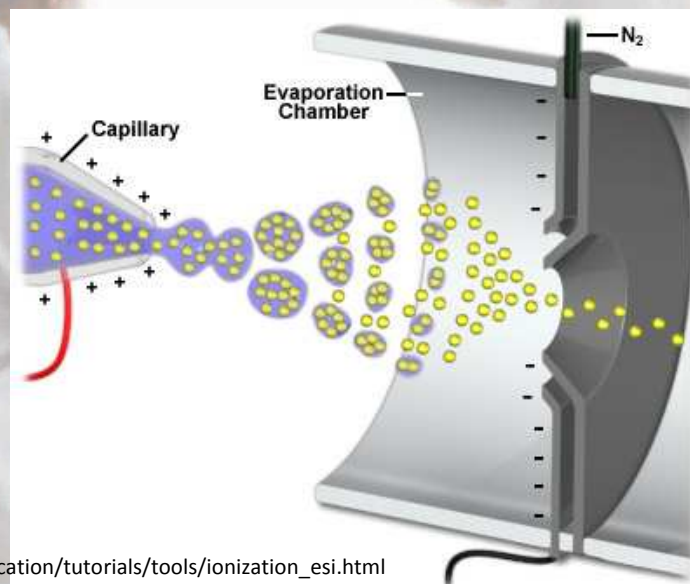
Vir: http://www.nature.com/nbt/journal/v21/n10/fig_tab/nbt1003-1156_F1.html

ESI-MS

- proteini v kisli raztopini (prebitek H^+)
- ionizirani z razprševanjem v prisotnosti močnega električnega polja ($\sim 2\text{-}5\text{ kV}$)
- tvorjenje disperznih pozitivnih ionov (nebulizacija)
- vakuumaska regija masnega analizatorja s specializirano ionsko optiko
- ESI zahteva, da ima molekula dokaj veliko maso
- inštrument ima majhen masni razpon



Vir: https://www.i-med.ac.at/iftz/zentrale_gruppen/proteinfacility/esi-ms.html



Vir: http://www.magnet.fsu.edu/education/tutorials/tools/ionization_esi.html

- spekter prikazan z razmerjem mase in naboja (m/z) na x-osi in relativno intenzivnostjo (%) posameznega vrha prikazano na y-osi

- izračun za določitev neznane mase M_r

- pogosto uporabljeni tandemski masni spektrometri

- zmožni hitrega izoliranja in fragmentacije posameznih peptidnih ionov

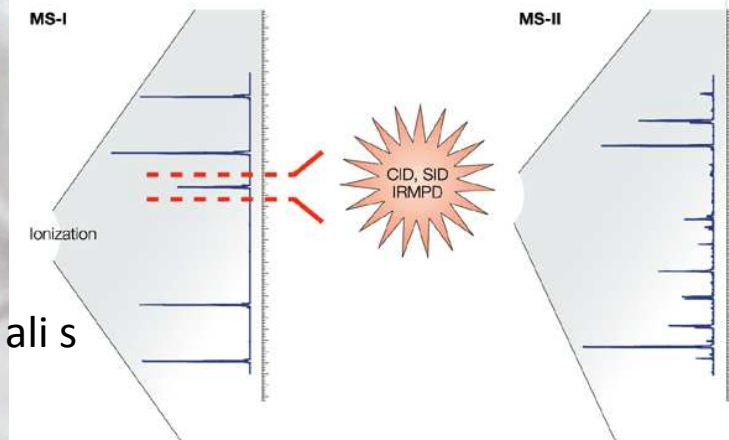
- s trkom povzročena disociacija (= CID)

- MS/MS spekter osnova za določevanje ustreznega AK zaporedja peptidov s klasičnim interpretiranjem fragm. ionov ali s pomočjo avtomatiziranih iskalnih algoritmov.

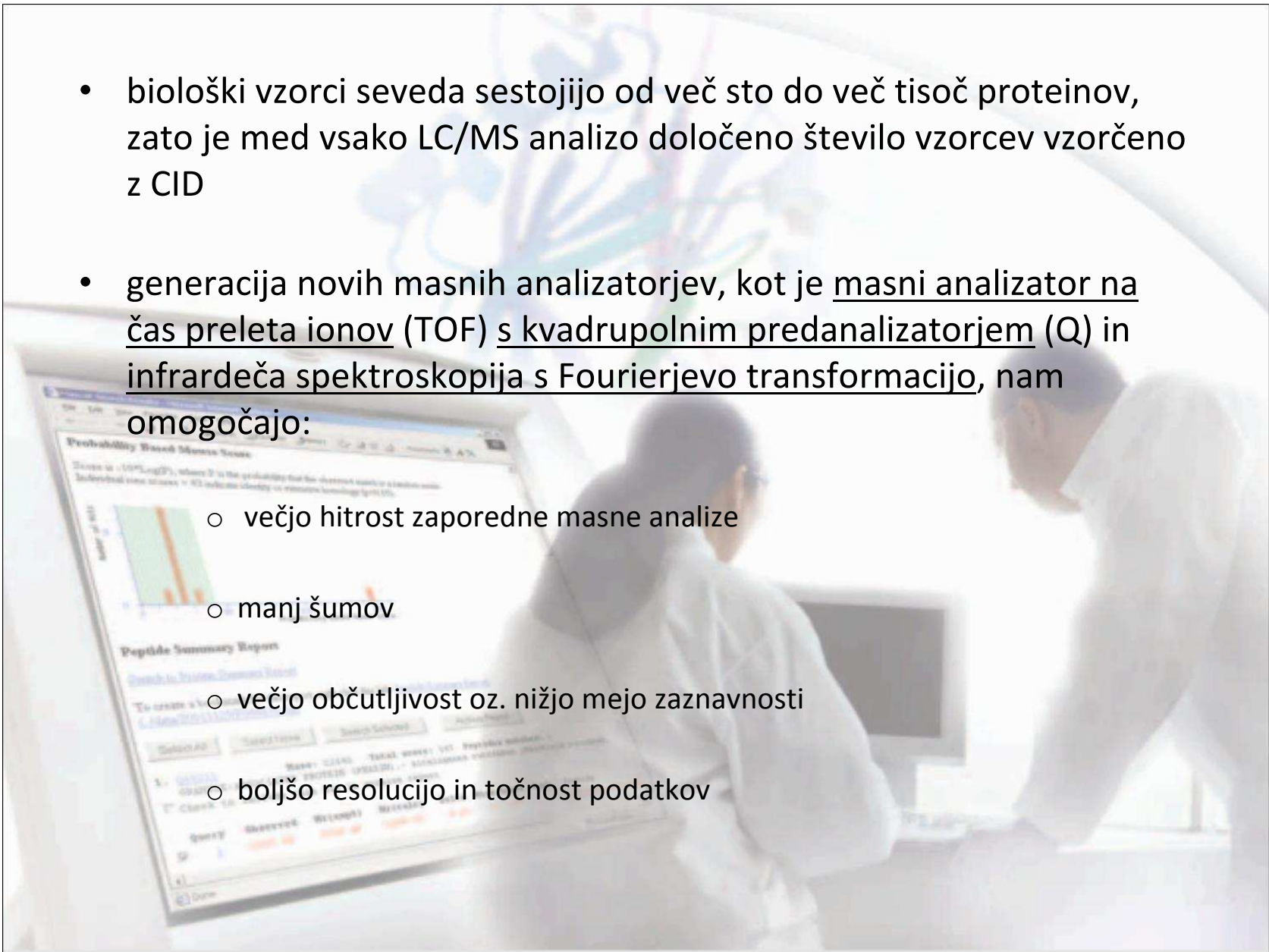
$$p = m / z$$

$$p = 1 (M_r + z_1) / z_1$$

$$p_2 = \{M_r + (z_1 - 1)\} / (z_1 - 1)$$



- biološki vzorci seveda sestojijo od več sto do več tisoč proteinov, zato je med vsako LC/MS analizo določeno število vzorcev vzorčeno z CID
- generacija novih masnih analizatorjev, kot je masni analizator na čas preleta ionov (TOF) s kvadрупolnim predanalizatorjem (Q) in infrardeča spektroskopija s Fourierjevo transformacijo, nam omogočajo:
 - večjo hitrost zaporedne masne analize
 - manj šumov
 - večjo občutljivost oz. nižjo mejo zaznavnosti
 - boljšo resolucijo in točnost podatkov

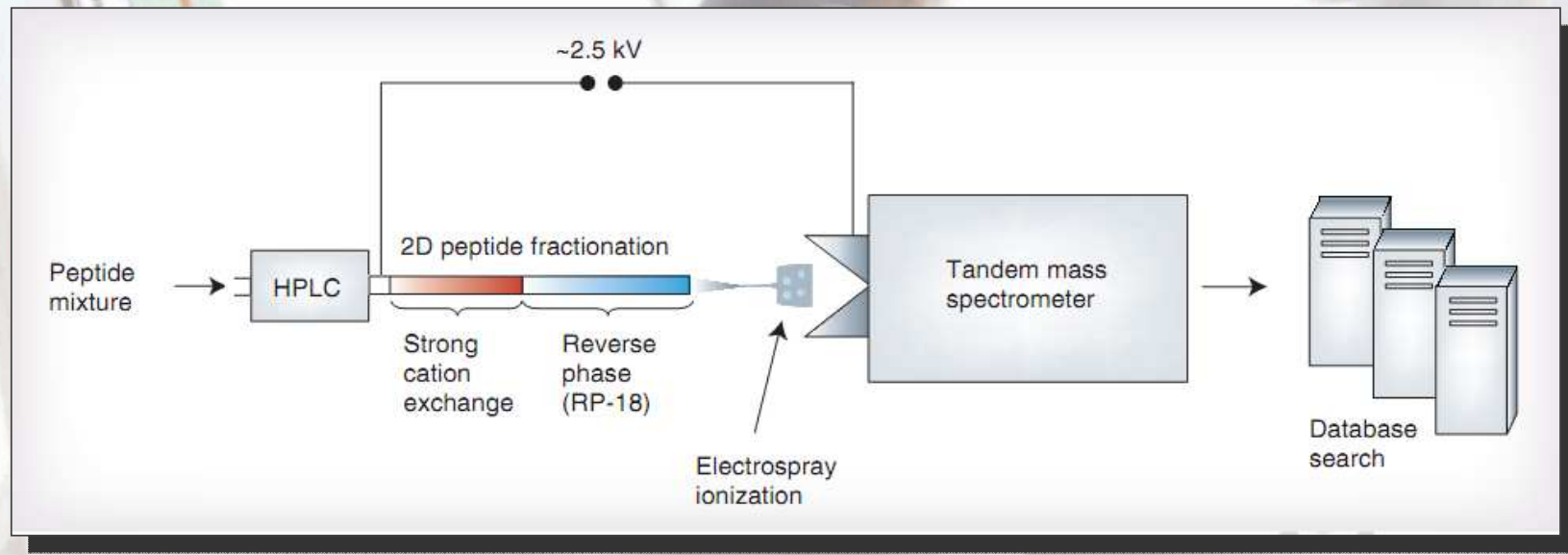


SHOTGUN PEPTIDE SEQUENCING

- interpretacija rezultatov MS/MS spektrov, ki jih dobimo z CID metodo
- sekvencioniranje *de novo* - časovno prezahtevno
- zahteva po avtomatizaciji
- SEQUEST - prva programska oprema
- določitev nabora kandidatov peptidnih sekvenc, ki bi jih bilo smiselno primerjati s spektrom
- vključitev le tistih, ki imajo maso blizu opazovani masi peptidnih ionov (uporaba navzkrižne korelacije).
- algoritem STATQUEST
- strožje ocene natančnosti

MudPIT

- zmanjšati kompleksnost vzorcev
- Yates s sodelavci
- frakcioniranje peptidov z večdimenzionalno kapilarno-lestvično kromatografijo, ki so jo izvedli pred LC/MS analizo
- Osnovni MudPIT postopek



MudPIT

- glavna prednost: kompleksnost vzorca občutno zmanjšana
- detekcija manj koncentriranih ali slabo ioniziranih peptidov se občutno izboljša
- MudPIT lahko identificira več kot 1000 proteinov v 24 urah
- detektira lahko majhne, pozitivno nabite proteine in tudi hidrofobne transmembranske proteine
- odkrivanje več kemičnih in izotopnih označevalcev za večjo pravilnost določevanja koncentracije proteinov

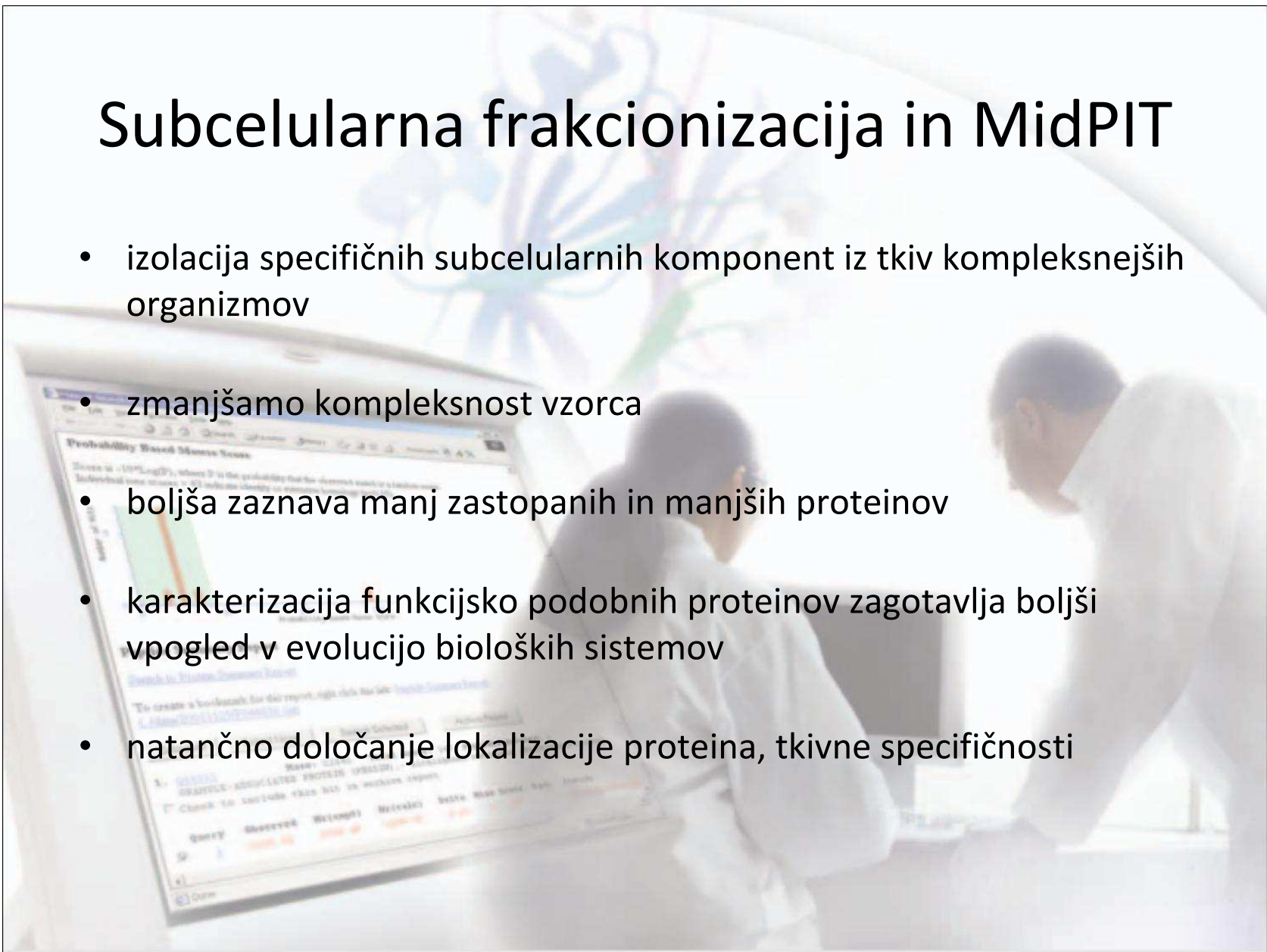
MudPIT

- analiza posttranslacijskih modifikacij s kombinacijo različnih proteoliz
- pomembna v funkcijski proteomiki
- analiziranje enakega vzorca:
 - MudPIT: 2363 proteinov
 - 2D elektroforeza: 556 proteinov
- določevanje tkivno specifičnih proteinov
- pomoč pri razvijanju učinkovitih cepiv proti parazitom, ki napadajo ljudi in povzročajo različne bolezni (npr. malarija)



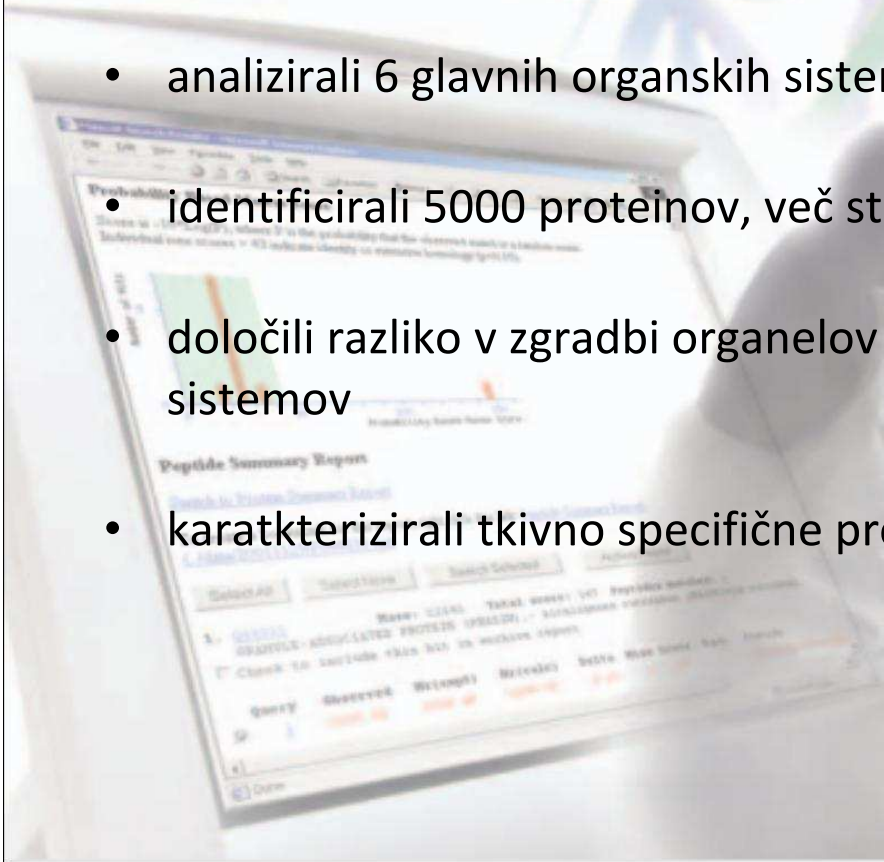
Subcelularna frakcionizacija in MidPIT

- izolacija specifičnih subcelularnih komponent iz tkiv kompleksnejših organizmov
- zmanjšamo kompleksnost vzorca
- boljša zaznava manj zastopanih in manjših proteinov
- karakterizacija funkcijsko podobnih proteinov zagotavlja boljši vpogled v evolucijo bioloških sistemov
- natančno določanje lokalizacije proteina, tkivne specifičnosti



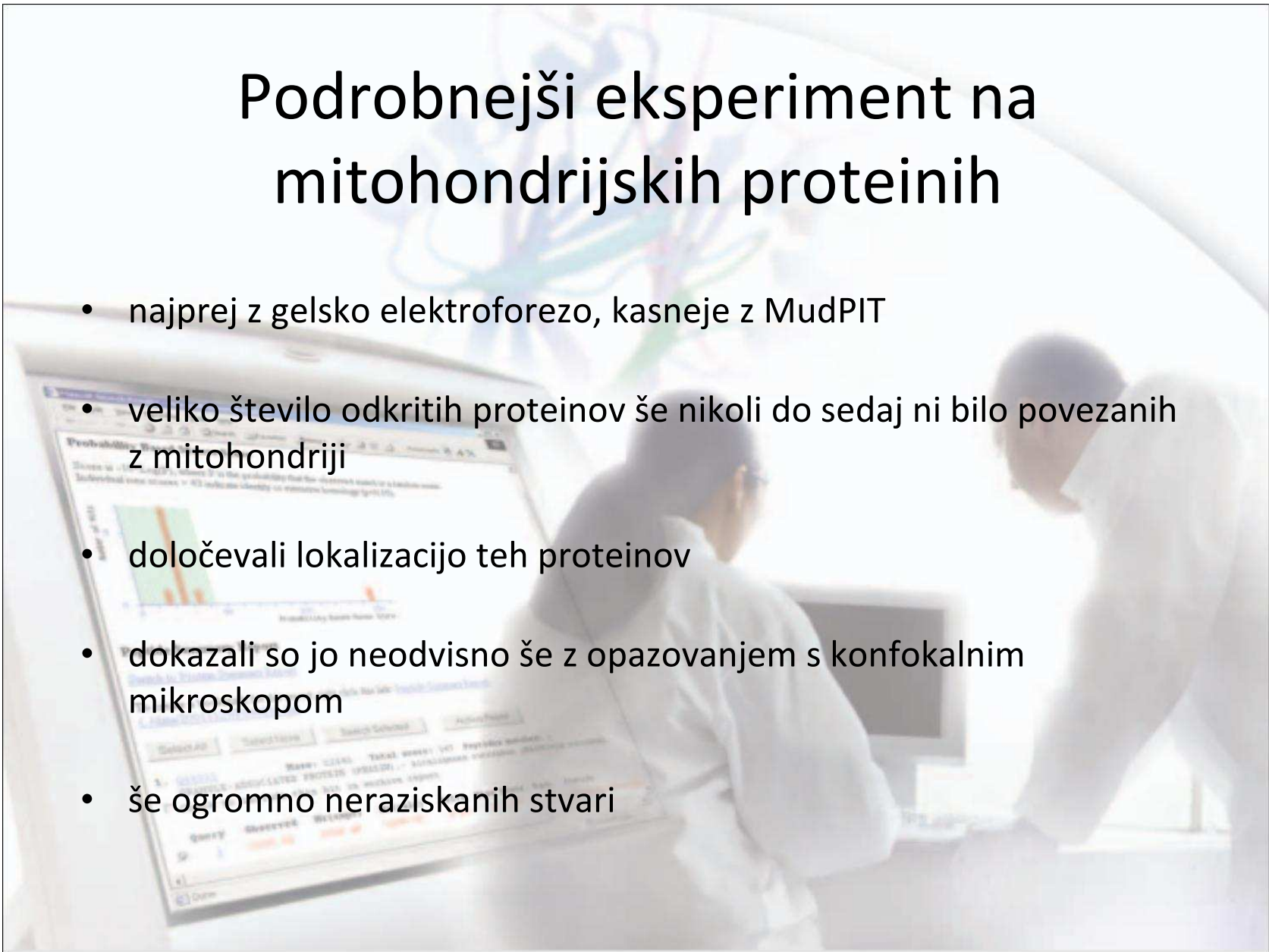
Analiza mišjih tkiv z MudPIT

- za lažje obravnavanje motenj v razvoju in bolezni miši
- analizirali 6 glavnih organskih sistemov (možgani, srce, pljuča...)
- identificirali 5000 proteinov, več sto novih
- določili razliko v zgradbi organelov v celicah različnih organskih sistemov
- karatkterizirali tkivno specifične proteine za določene predele



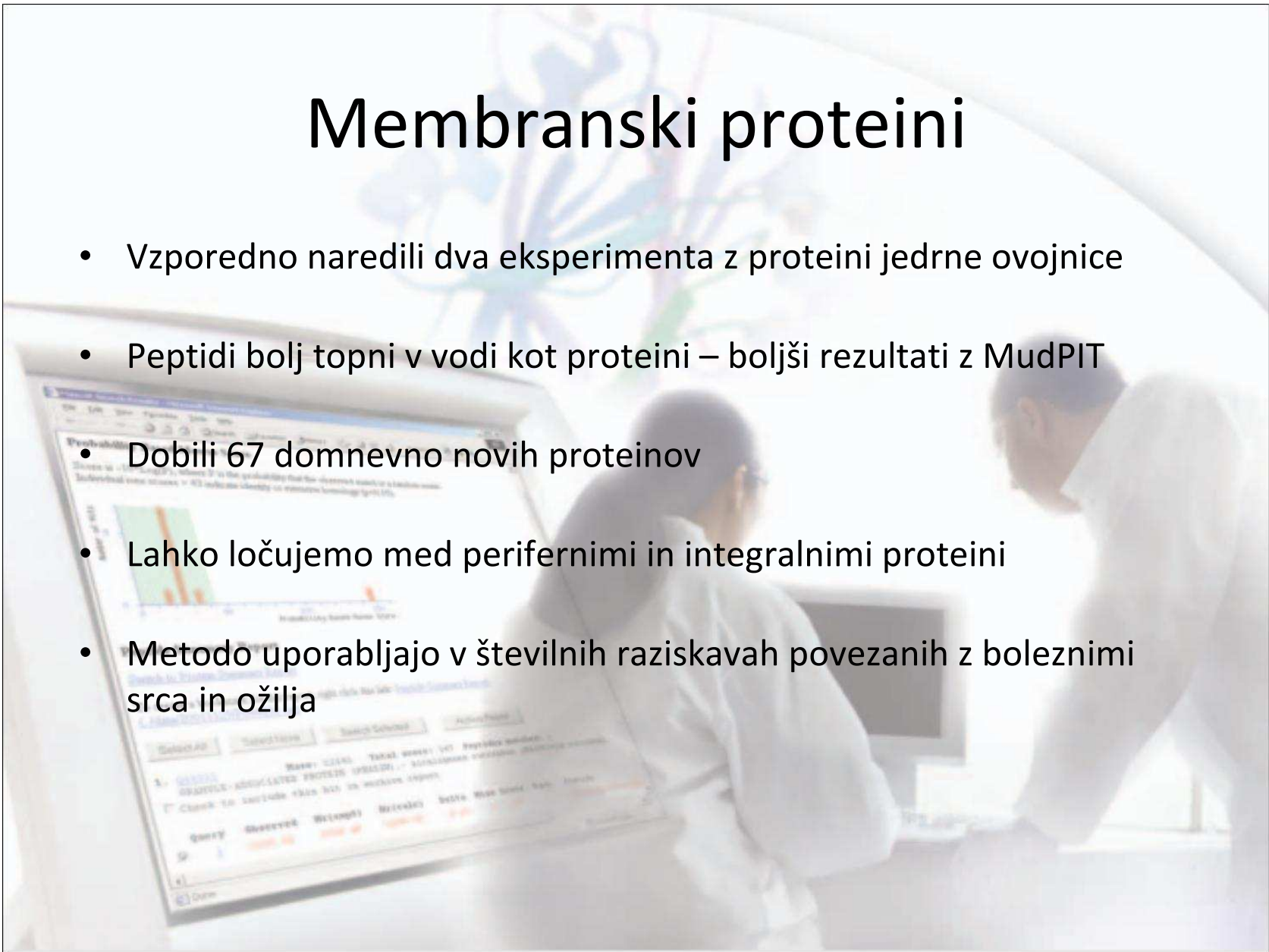
Podrobnejši eksperiment na mitohondrijskih proteinih

- najprej z gelsko elektroforezo, kasneje z MudPIT
- veliko število odkritih proteinov še nikoli do sedaj ni bilo povezanih z mitohondriji
- določevali lokalizacijo teh proteinov
- dokazali so jo neodvisno še z opazovanjem s konfokalnim mikroskopom
- še ogromno neraziskanih stvari



Membranski proteini

- Vzporedno naredili dva eksperimenta z proteini jedrne ovojnice
- Peptidi bolj topni v vodi kot proteini – boljši rezultati z MudPIT
- Dobili 67 domnevno novih proteinov
- Lahko ločujemo med perifernimi in integralnimi proteini
- Metodo uporabljajo v številnih raziskavah povezanih z boleznimi srca in ožilja



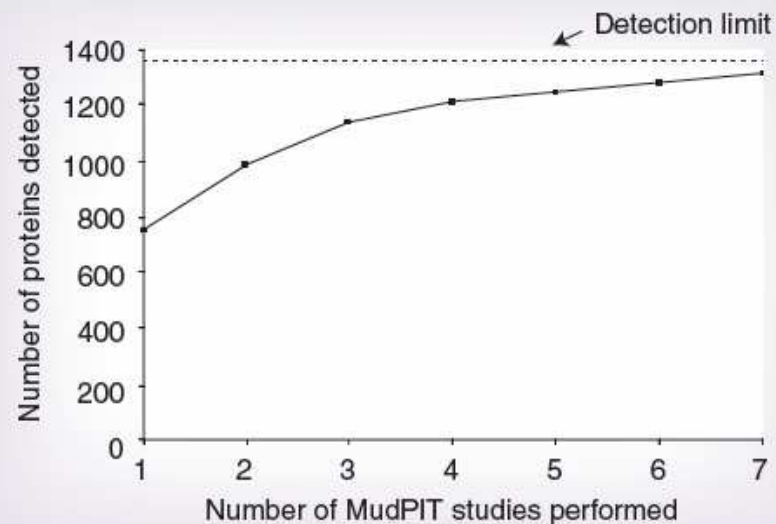
Meje MudPIT

- MS instrumenti-obratovani cikel ostaja isti
- Vir ionizacije-prednost manjše molekule



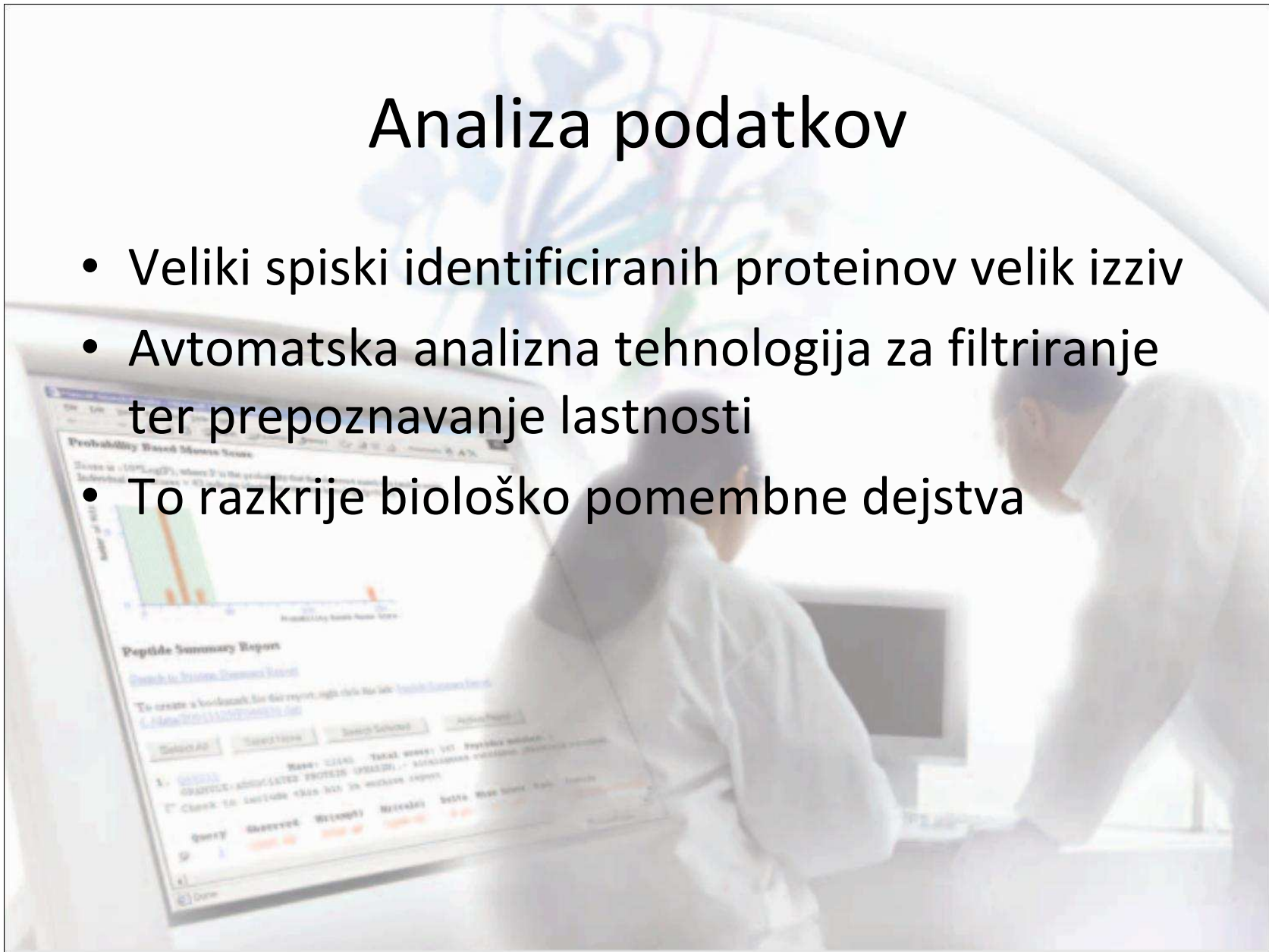
Izboljšave

- MS linearne pasti, hitrejši obratovalni cikel
- Večkrat ponovljen MudPIT-statistični modeli
- Poenostavljena kompleksnost tako da je dodan še en krog fragmentacije (FPLCinHPLC) pred MudPIT



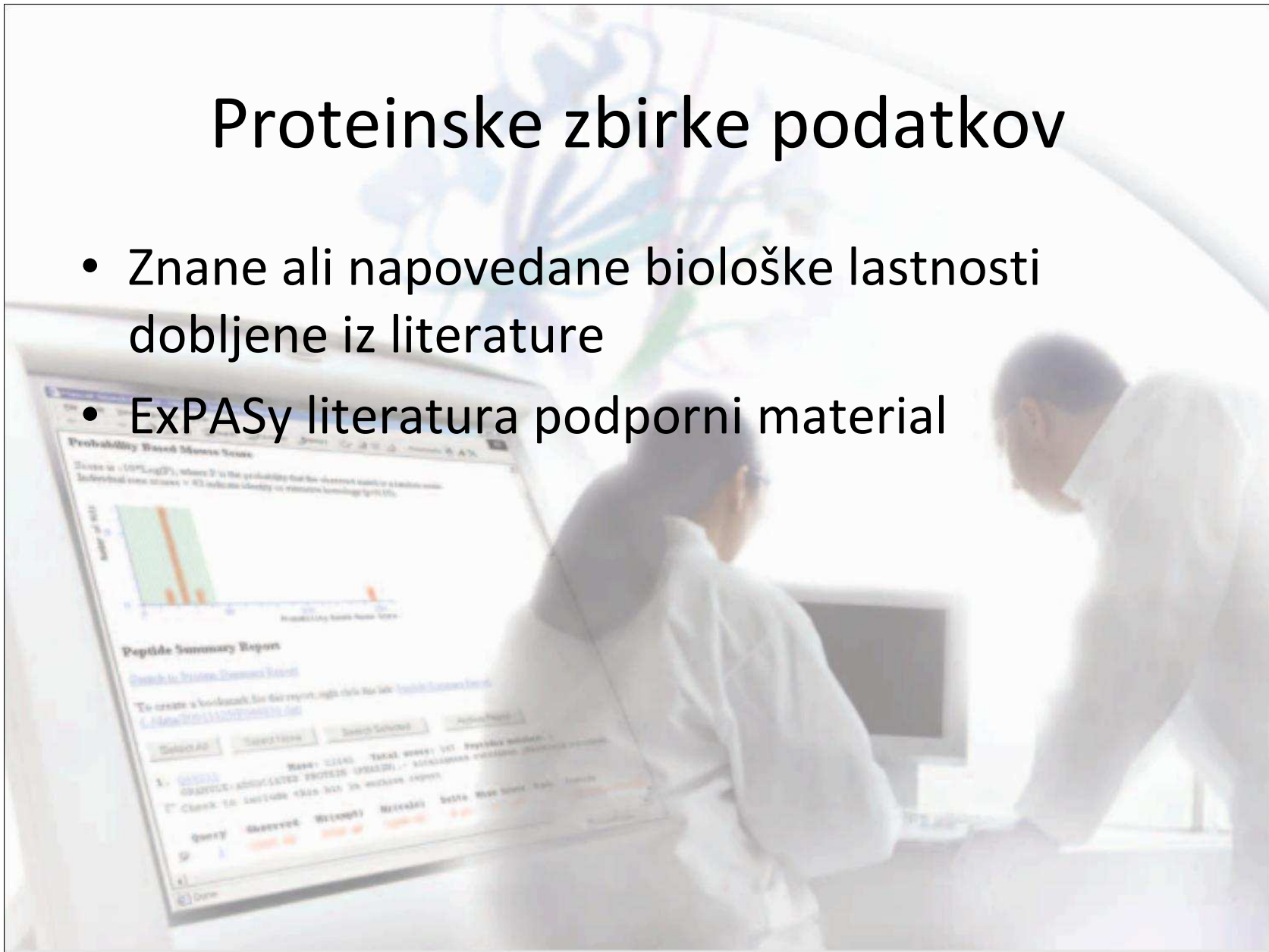
Analiza podatkov

- Veliki spiski identificiranih proteinov velik izziv
- Avtomatska analizna tehnologija za filtriranje ter prepoznavanje lastnosti
- To razkrije biološko pomembne dejstva



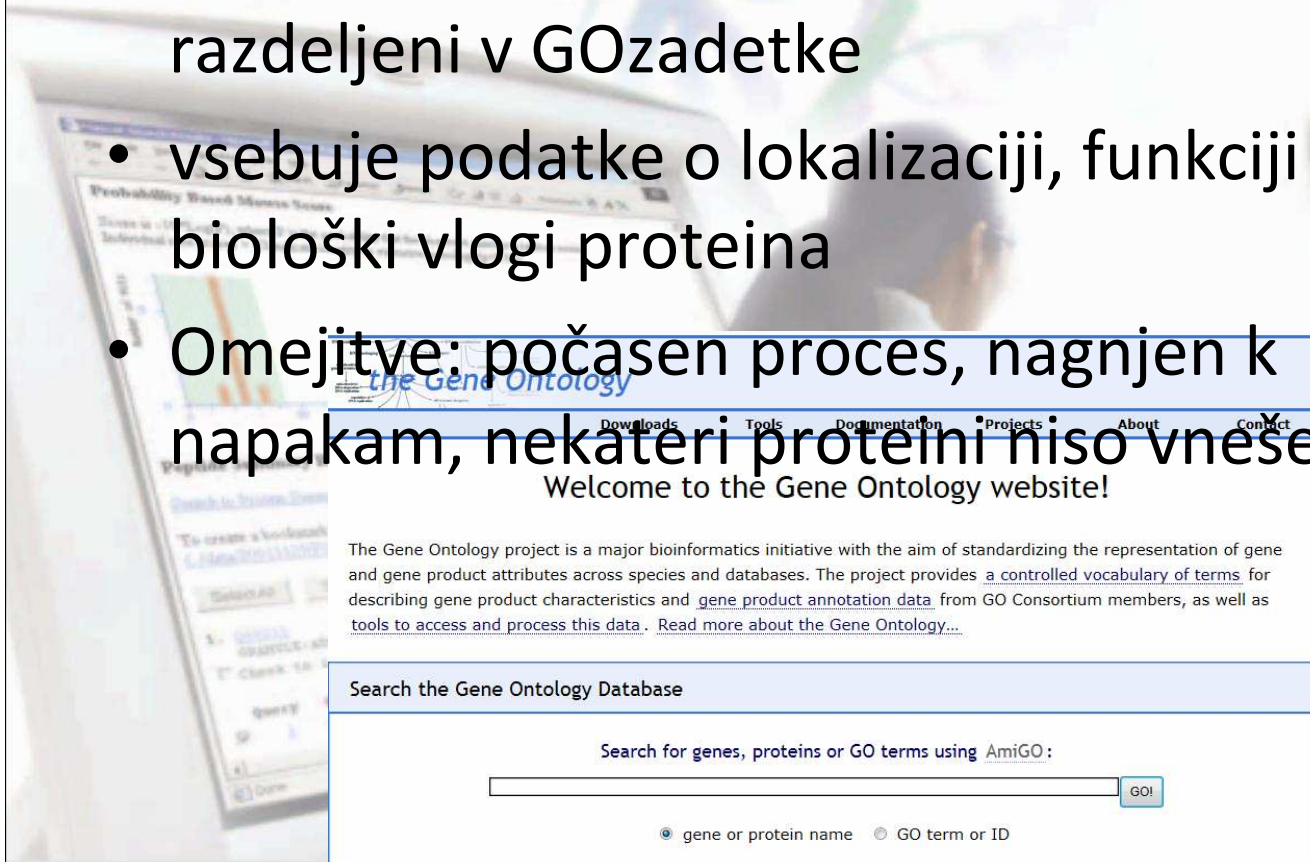
Proteinske zbirke podatkov

- Znane ali napovedane biološke lastnosti dobljene iz literature
- ExPASy literatura podporni material



Gene Ontology Consortium

- proteini identificirani z LC/MS ali MudPIT razdeljeni v GOzadetke
- vsebuje podatke o lokalizaciji, funkciji ter biološki vlogi proteina
- Omejitve: počasen proces, nagnjen k napakam, nekateri proteini niso vnešeni

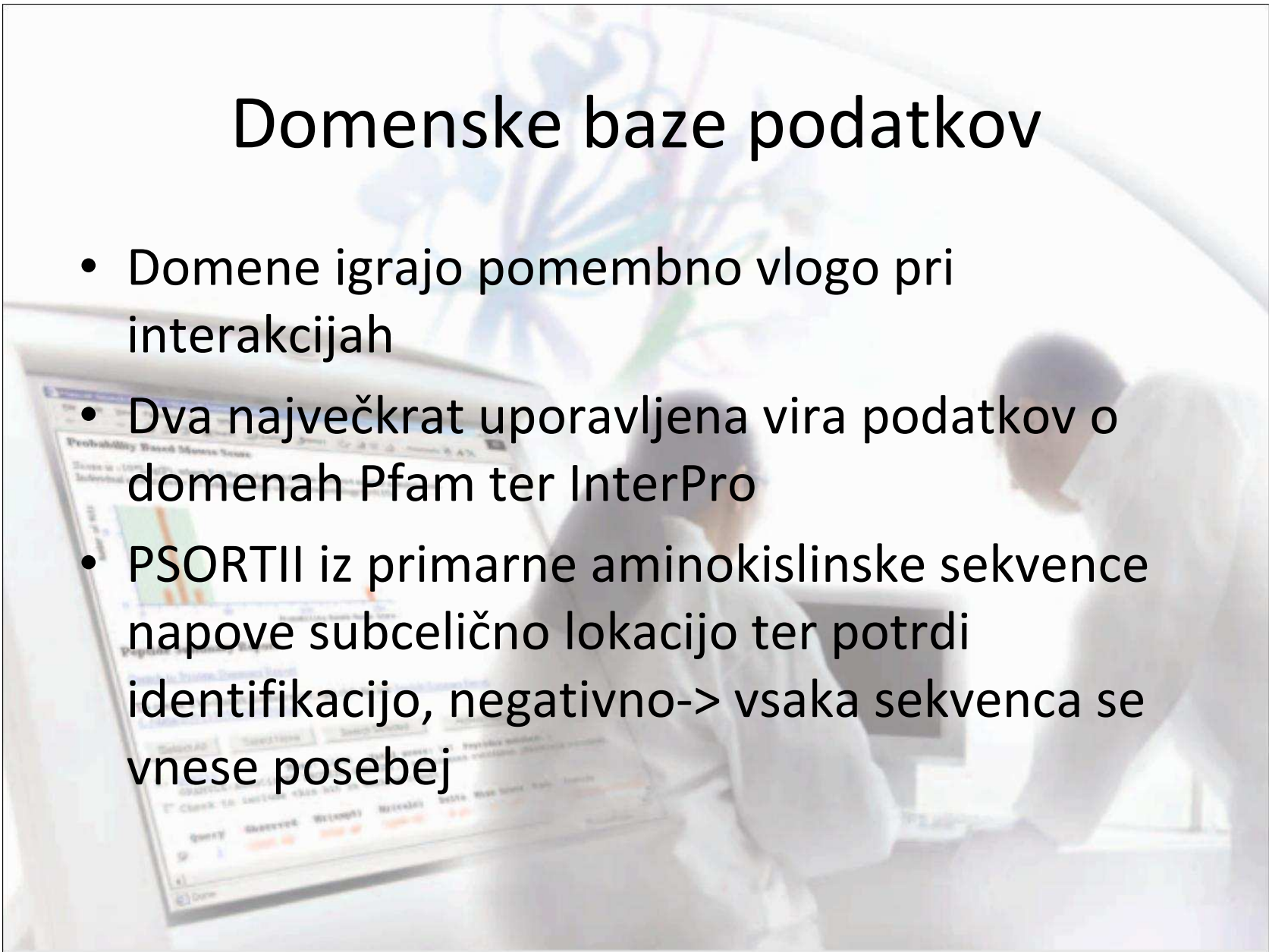


Iskanje z GO

- Iskanje z peptidom po bazah podatkov z sekvencami proteinov
- GOCluster, Perl program , ki avtomatsko razdelita seznam proteinov identificiranih z MudPIT v GOzadetke
- GOMiner poleg tega primerja tudi dva seta podatkov da obogati Gozadetke
- MouseSpec dobi seznam podatkov poda GOzadetke obogatene s statističnimi podatki

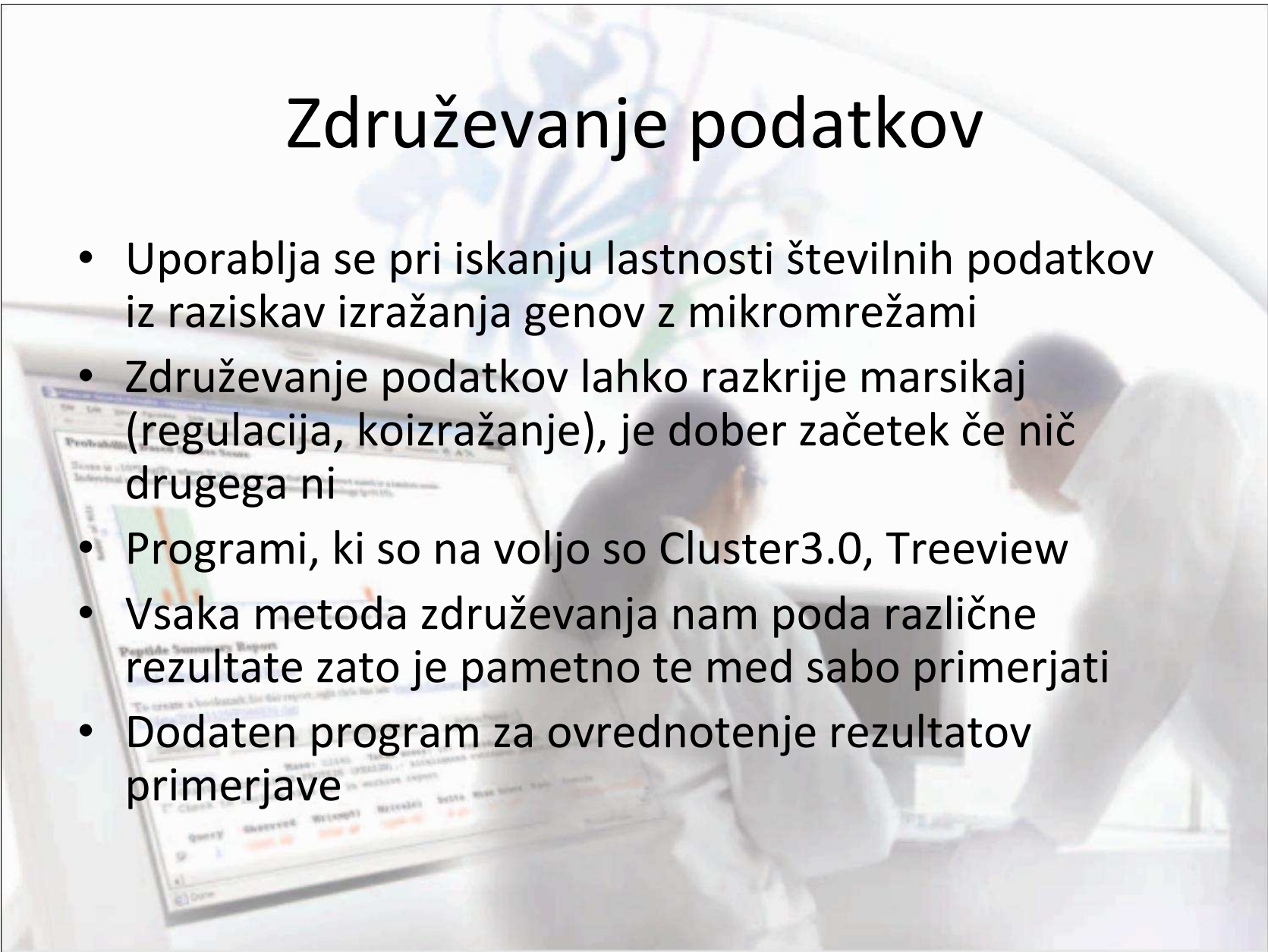
Domenske baze podatkov

- Domene igrajo pomembno vlogo pri interakcijah
- Dva največkrat uporabljena vira podatkov o domenah Pfam ter InterPro
- PSORTII iz primarne aminokislinske sekvence napove subcelično lokacijo ter potrdi identifikacijo, negativno-> vsaka sekvenca se vnese posebej



Združevanje podatkov

- Uporablja se pri iskanju lastnosti številnih podatkov iz raziskav izražanja genov z mikromrežami
- Združevanje podatkov lahko razkrije marsikaj (regulacija, koizražanje), je dober začetek če nič drugega ni
- Programi, ki so na voljo so Cluster3.0, Treeview
- Vsaka metoda združevanja nam poda različne rezultate zato je pametno te med sabo primerjati
- Dodaten program za ovrednotenje rezultatov primerjave



Ostali uporabni programi

- Spremembe količine posamezne komponente v poteh in funkcionalnih enotah
- GenMAPP nam to grafično prikaže
- Primerjava glavnih lastnosti (MudPIT) in nivo genskega traskripta (mikromreže)
- Za to se uporabljajo: Affymetrix GeneChip, BLAST

