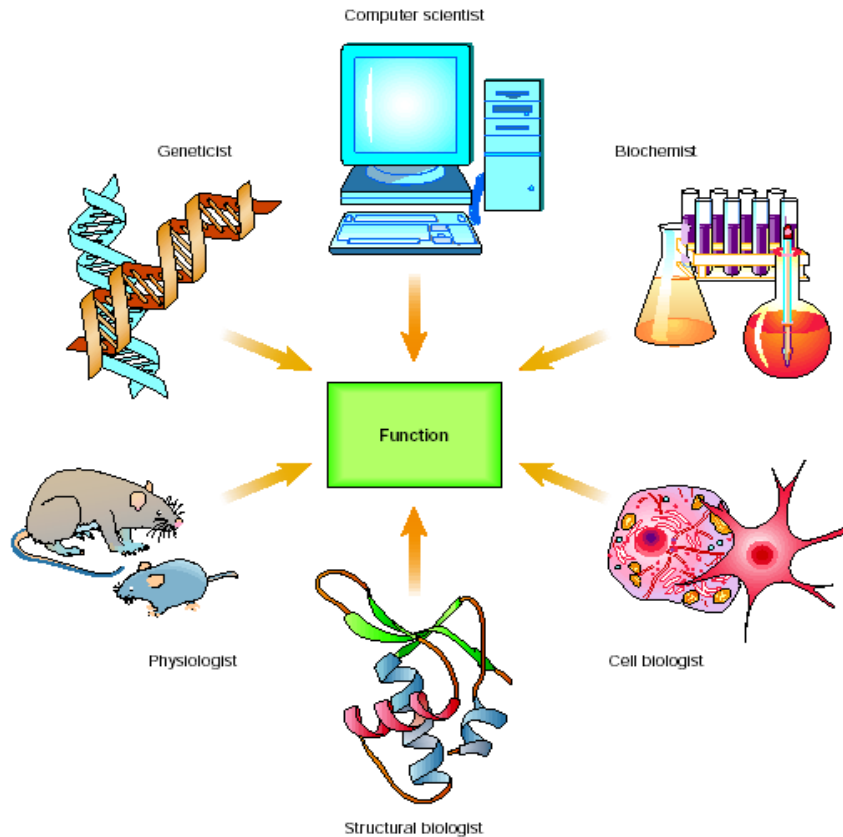


genom → transkriptom → proteom

Na osnovi nukleotidnih zaporedij identificiranih genov in povezave med zaporedji in fenotipi lahko dobimo uporabne informacije o delovanju genoma in o boleznih. Potrebno je sodelovanje strokovnjakov z več področij.



Genomika

Genomika je veda, ki preučuje gene in njihove funkcije. Spoznanja dopolnjujejo znanje o mehanizmih bolezni, vključno z vlogo dednosti in okolja za razvoj bolezni. Z odkrivanjem novih bioloških tarč bo prišlo do razvoja novih vrst zdravil, cepiv in diagnostičnih sredstev.

Kapaciteta mednarodnega konzorcija 16 skupin za določitev zaporedja človekovega genoma je bila ~2000 b/s (172 Mb/dan). Razvoj je šel v smeri dokončanja zaporedja posameznih kromosomov z visoko natančnostjo (8-10x pokritost), določanja zaporedja drugih genomov in analize dobljenih rezultatov – študij funkcije odkritih genov. Pri tem je v veliko pomoč tehnologija mikromrež, s katero lahko preučujemo izražanje posameznih genov. Poznavanje nukleotidnih zaporedij je osnova za študij kompleksih dednih procesov v organizmu.

Računali so, da naj bi razjasnitev nukleotidnega zaporedja pri človeku povečala število tarčnih molekul za zdravljenje za en red velikosti (s ~500 na 3000-5000).

Kot posebne panoge se pojavljajo **primerjalna genomika** (primerjava zaporedij med vrstami, med posamezniki), **funkcijska genomika** (povezava genotip/fenotip), **fiziološka genomika**, **nevrogenomika**, **strukturna genomika**...

Na veliko odprtih vprašanj bo mogoče odgovoriti šele z raziskovanjem proteinov → **proteomika**.

Primerjalna genomika



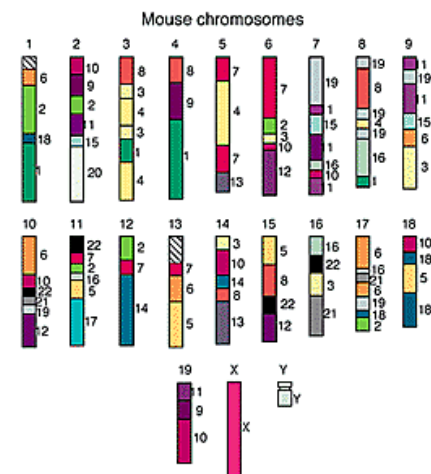
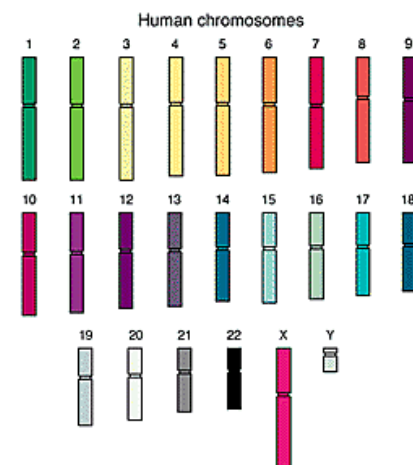
Lastnosti genov pri človeku lahko raziskujemo s pomočjo primerjalnih študij na drugih organizmih.

Miš je uporabna kot primerjalni organizem za človeka, ker je človeku genetsko dokaj podobna in ker obstaja veliko število transgenskih miši, ki nosijo posamezne odseke človeških kromosomov, vključno z nekaterimi okvarjenimi geni. Miš je lahko preiskovati z agresivnimi metodami, obstoj klonov omogoča opazovanje v paralelkah.

Za razumevanje nekaterih osnovnih procesov v molekularni biologiji so pomembni tudi podatki o genomih nižjih organizmov, vključno z bakterijami in kvasovkami.

Primerjalna genomika je uporabna tudi za določanje evlucijskih odnosov med vrstami in ugotavljanje evlucijskih poti po skupinah organizmov.

Na osnovi predhodnega poznavanja lastnosti proteinov, za katera zapisujejo določena zaporedja, lahko podobnim nukleotidnim zaporedjem pripišemo verjetno vlogo v nekem organizmu.



Primerjalna genomika: gobavost

Analiza SNP različnih izolatov *Mycobacterium leprae* je pokazala, da se je patogen razvil na enem mestu in se razširil s selitvami prebivalstva in kolonializmom.

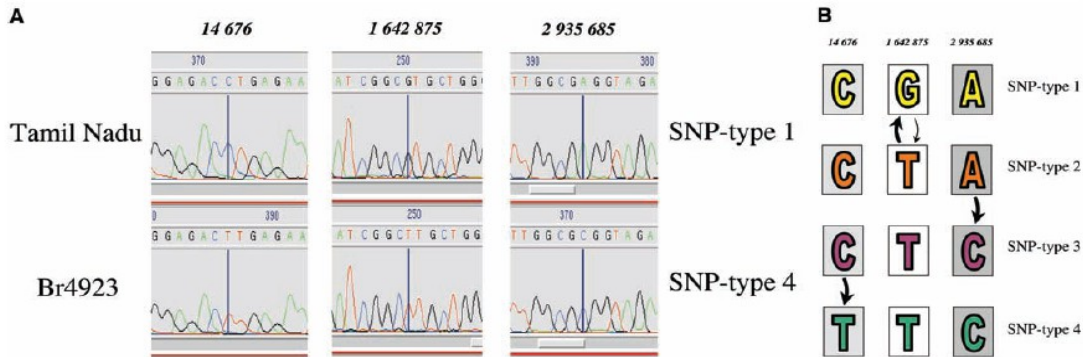


Fig. 1. SNP analysis of isolates of different geographical origin and parsimony. (A) Comparison of polymorphic sites in the genomes of the TN and Br4923 strains by automated DNA sequencing. Coordinates are the position in the genome of the TN strain, and the vertical bar indicates

the polymorphic base. (B) The most parsimonious route to account for the four SNP types. Bold arrows indicate the most likely direction, based on historical and geographic considerations; the faint arrow denotes an alternative route.

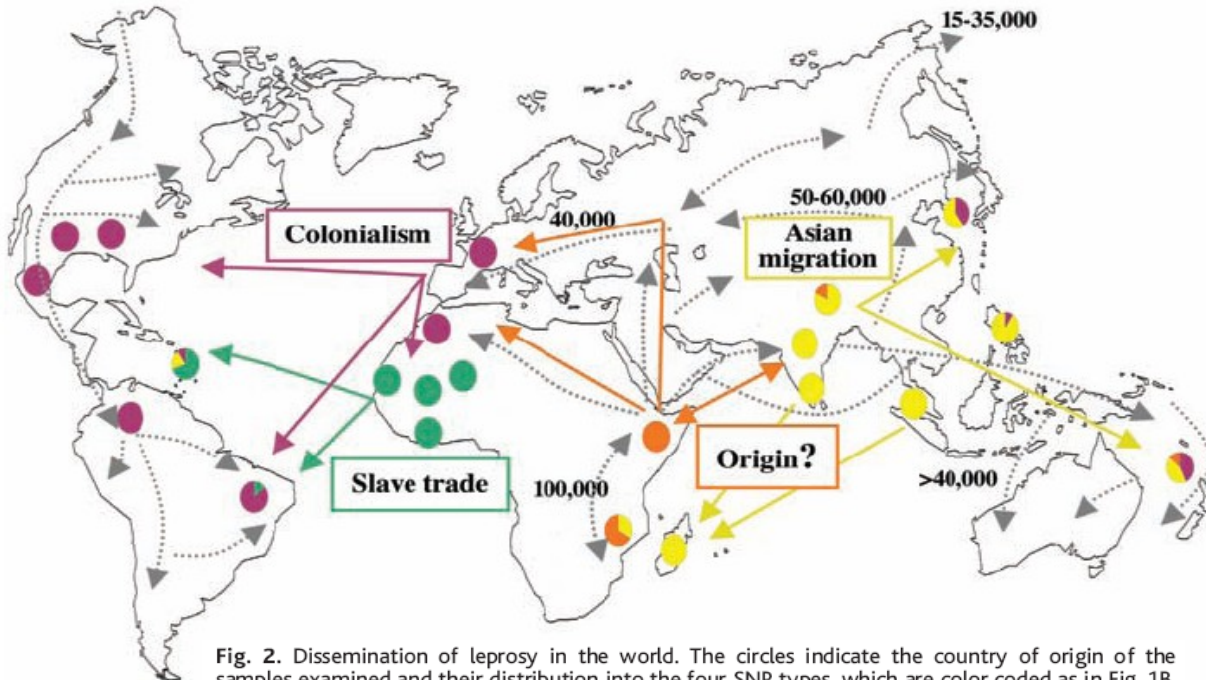


Fig. 2. Dissemination of leprosy in the world. The circles indicate the country of origin of the samples examined and their distribution into the four SNP types, which are color coded as in Fig. 1B. The colored arrows indicate the direction of human migrations predicted by, or inferred from, our SNP analysis; gray arrows correspond to the migration routes of humans derived from genetic, archaeological, and anthropological studies, with the estimated time of migration in years (27, 28).

Table 3. SNP analysis of *M. leprae* from different countries.

Country	SNP type 1	SNP type 2	SNP type 3	SNP type 4
New Caledonia	3	1	3	
Philippines	19		2	
Korea	3		2	
Thailand	1			
Nepal/North India	23	5		
South India	4			
Madagascar	6			
Ethiopia		2		
Malawi	4	6		
Mali				31
Ivory Coast				6
Guinea				1
Senegal				2
Morocco			2	
France			2	
Brazil			12	2
French West Indies	4		2	14
Venezuela			5	
Mexico			1	
United States			3	
WARM*			4	
Total	67	14	38	56

*Wild armadillos from Louisiana, USA.

Primerjalna genomika: človek / opice

Zakaj so razlike med človekom in šimpanzom ob >98 % identičnosti genomov tako velike?

Razlog je v tem, da so se v zadnjih 5 milijonih let zelo hitro spremenili geni za transkripcijske faktorje.

Z biočipi, na katerih so bile kot sonde različni geni, ki se izražajo v jetrih, so preverili njihovo izražanje pri različnih vrstah opic in pri človeku. Izločili so tiste, ki se pri človeku izražajo bistveno drugače kot pri ostalih primatih in ugotovili, da gre večinoma za transkripcijske faktorje.

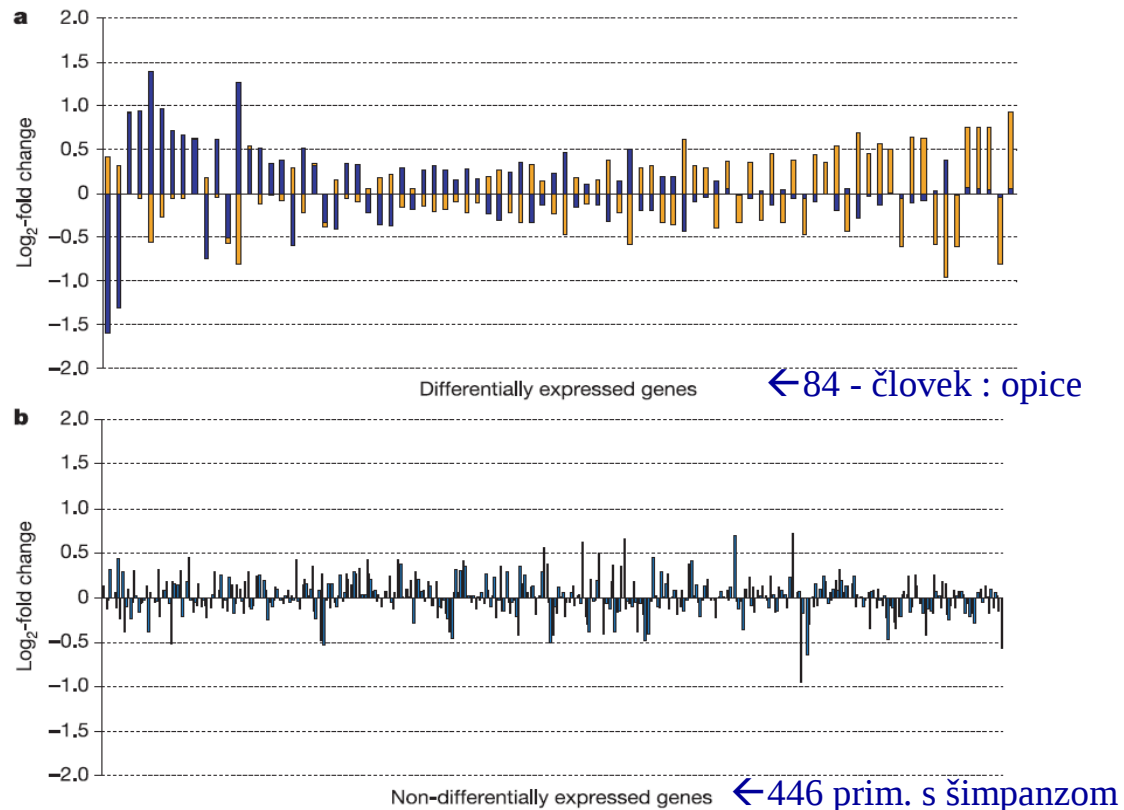
Zakaj so se ti evolucijsko spreminjali tako hitro? Zaradi prehoda na kuhano hrano?

Expression profiling in primates reveals a rapid evolution of human transcription factors

NATURE | Vol 440 | 9 March 2006

Table 1 | Inter-species differentially expressed genes

	Chimpanzee	Orangutan	Rhesus macaque
Human	110	128	176
Chimpanzee	-	150	141
Orangutan	-	-	129



Funkcijska genomika

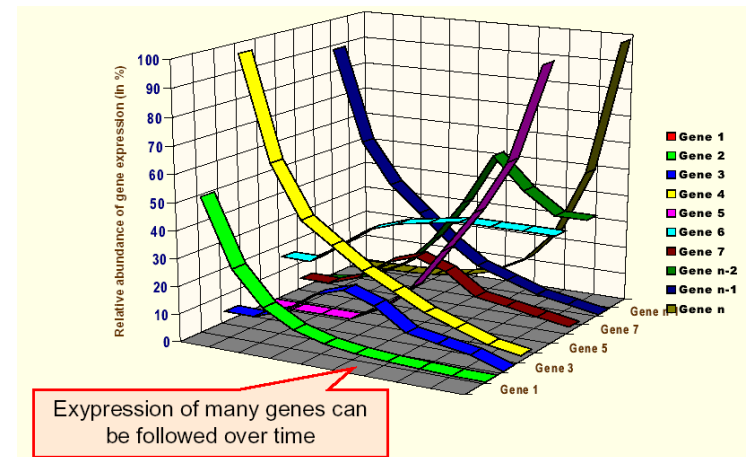
postgenomika = funkcijska genomika

Vsakemu genu v genomu želimo določiti funkcijo s sistemskim pristopom, na veliko.

Nekateri genomiko delijo samo na dva dela, na **funkcijsko** in **strukturno**. Strukturna poskuša predvideti terciarne strukture proteinov na osnovi nukleotidnih zaporedij oz. primarnih struktur prevedenih zapisov. **Funkcijsko** genomiko naj bi po tej delitvi sestavljali **proteomika** (katalogiziranje vseh proteinov, ki jih ima na voljo nek organizem) in **citomika** (dinamičnost žive celice: celični sistemi, njihova regulacija in spreminjanje komponent).

Orodja v funkcijski genomiki: ekspresijski profili, profili interakcij med proteini, analize proteoma, profili transgenskih fenotipskih mutant.

Analysing gene expression in the heart by *in situ* hybridization. **a, b**, *in situ* hybridization, showing expression of the mouse *Sh3bgr* gene, at two different stages of embryonic development. **c**, A tissue section of the mouse embryo, showing the expression of the *Sh3bgr* gene in developing heart.



Functional Genomics

“Transcriptomes”
(Expression Profiles)
DNA Micorarrays

Computational
Analyses of
Sequences

“Proteomics”
(protein structure, protein-protein
interactions, protein identification)
Affinity technologies, Mass
Spectrometry, protein
microarrays

“Genetic
Fingerprinting”
(Change in phenotype as
a result of mutation)

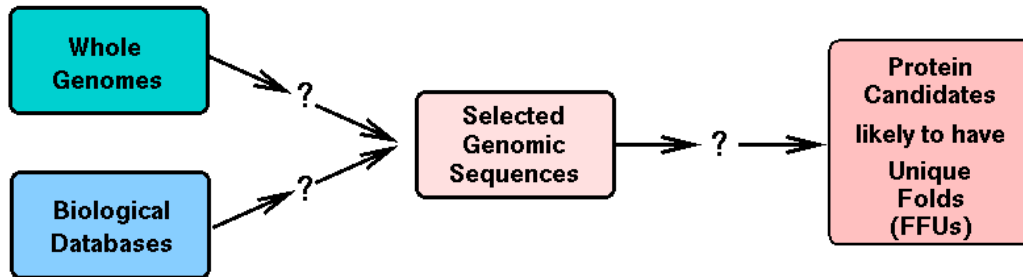
(Bioinformatics)

Strukturna genomika

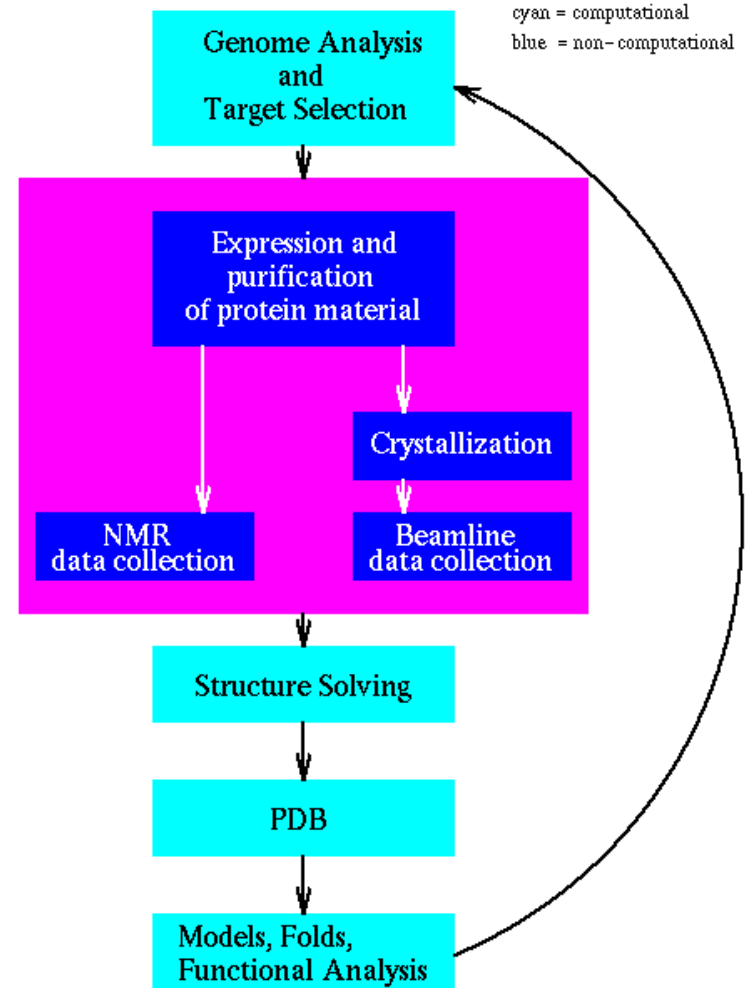
Cilj: vsakemu primarnemu zaporedju predvideti terciarno zgradbo.

Vmesna naloga: določiti 3D zgradbo po 1 proteinu iz vsake družine.

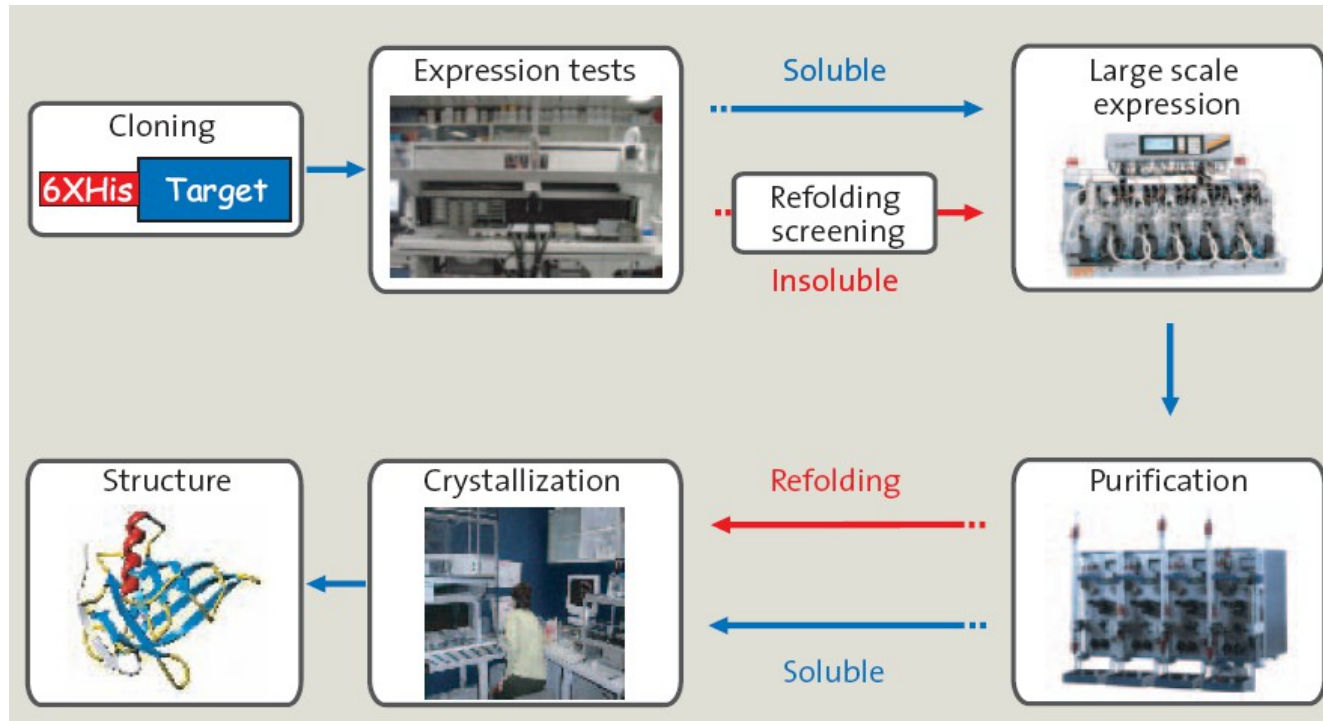
Veliki centri za strukturno biologijo izvajajo kristalizacijske poskuse in določajo prostorske zgradbe s pomočjo močnih virov žarkov.



Tasks for the Structural Genomics Process



Strukturna genomika: potek dela



The AFMB high throughput pipeline, from cloning to solving protein structures

AFMB = Architecture et Fonction des Macromolécules Biologiques, Marseille

Iz: Tecan Journal 02/2005

Strukturna genomika: rekombin. DNA

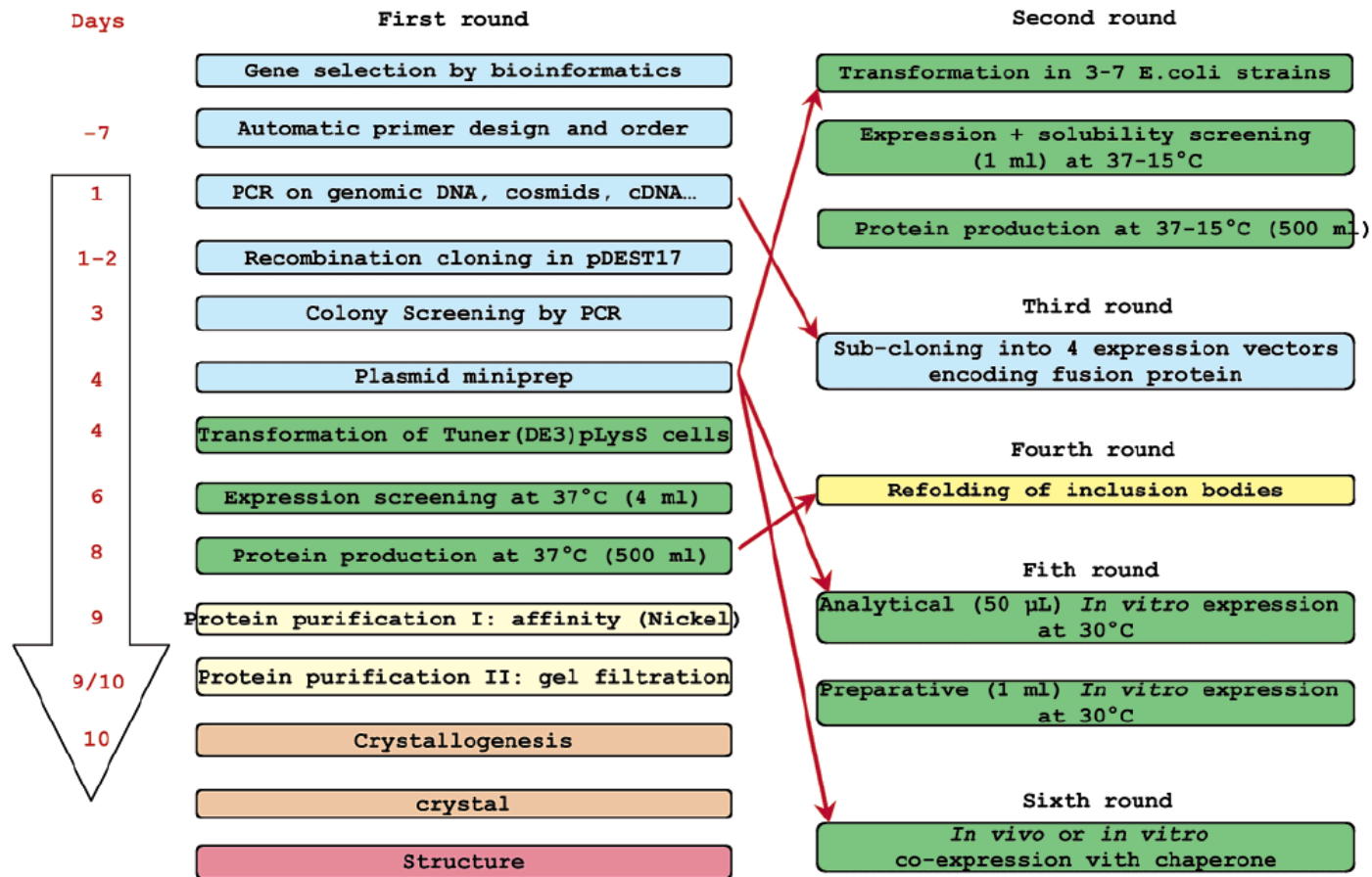


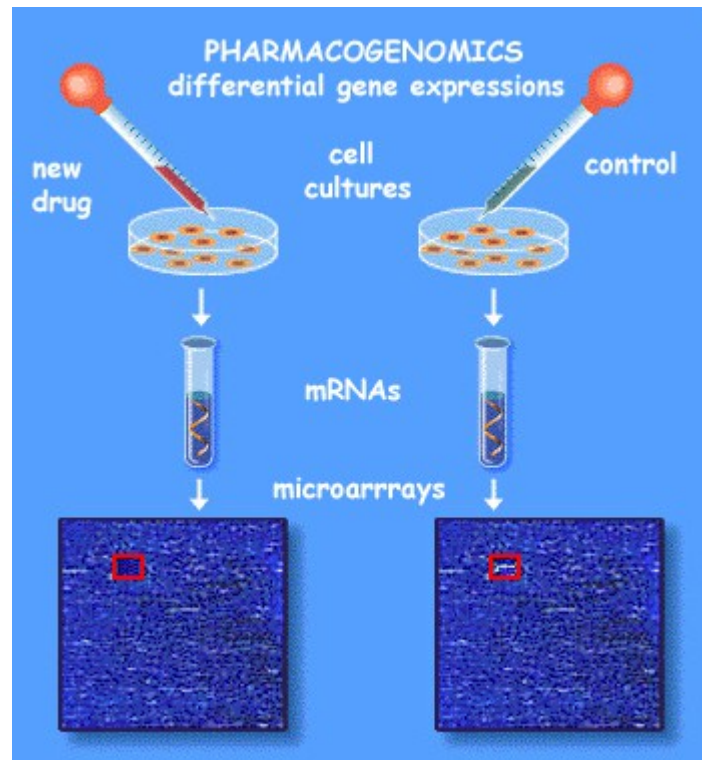
FIGURE 2. Evolution of the work scheme. The 108 target genes were processed in a simple first screening round. The targets that led to negative results (i.e., no expression or insoluble protein) underwent a second screening round, and so on. To date, rounds 1–4 but not 5 and 6 have been effectively tried in SG programs.

Farmakogenomika

Farmakogenetika: študij variant posameznih genov, ki bi lahko vplivali na potek zdravljenja.

Farmakogenomika: študij več genov (genoma) hkrati, saj je odziv organizma na neko zdravilo odvisen od več proteinov hkrati. Hkrati odkriva možne tarče za razvoj novih zdravil.

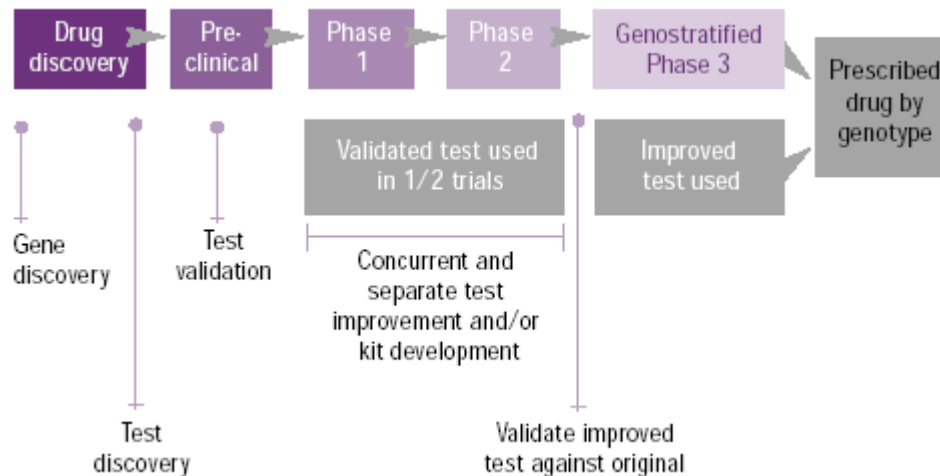
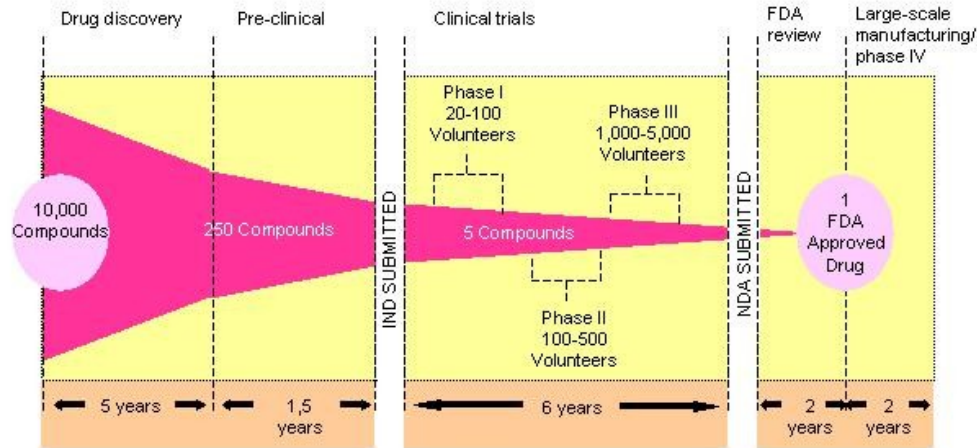
Ko bodo znane genske lastnosti posameznika, se bodo odprla vrata za poosebljeno zdravljenje (*personalized medicine*); doslej namreč zdravimo na splošno, v prihodnosti pa bi zdravili posamezne subpopulacije ločeno. Že doslej je bilo namreč jasno, da so nekatere vrste odziva na zdravila genetsko pogojene – odzive je treba povezati s posameznimi genskimi značilnostmi (SNP) in zdraviti ločeno.



<http://porpax.bio.miami.edu/~cmallery/150/gene/pharmacogenomics.gif>

Farmakogenomika /2

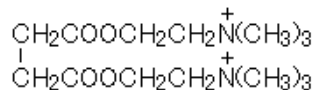
Proces razvoja in validacije zdravil bo v bodoče drugačen, kot je bil. Farmakogenomika (toksikogenomika) bo v fazi testiranja predvidela odzive posameznikov, zato bodo zdravila testirali le na ljudeh, za katere predvidevajo, da se bodo pozitivno odzivali, seveda pa bodo potem odobrena zdravila tudi predpisovali ločeno, glede na genske lastnosti pacienta. Predvideti bo mogoče najboljšo vrsto zdravila, pa tudi optimalne doze.



Farmakogenomika: problemi, primeri

Nature **425**, 760 - 762 (2003)

~1955: anestetik sukcinilholin (sprošanje mišic):
nekateri pacienti med operacijo prenehajo dihati.

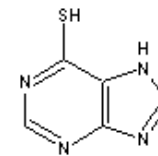


Biokemijska osnova: okvarjena holinesteraza

(1/2500 homozigotov z mutiranim zapisom ne razgrajuje tega substrata.)

~1955: merkaptopurin za zdravljenje levkemije.

Zelo toksičen, smiselno doziranje le v zelo ozkem območju,
vendar so za nekatere paciente te doze bile izrazito toksične,
za druge pa ne. Zdravilo presnavlja tiopurin-metiltransferaza (TMaza).



~1995: ugotavljanje mutacij v genu za TMazo: pri pacientih, ki so imeli hude zastrupitve, je vsaj 1 neobičajen SNP (identificirali 3 mesta). Zato paciente testirajo, preden se odločijo za terapijo.

Delovanje ~20 % zdravil naj bi bilo povezanih s polimorfizmom v posameznih genih, nadaljnjih 20 % pa s polimorfizmom v več genih, ki so pomembni za presnovo zdravil. Upoštevati pa je treba tudi transport učinkovin, interakcijo z receptorji, imunske dejavnike itd.

Med encimi, ki so pomembni za presnovo zdravil, je na prvem mestu družina P450 (oksidirajo učinkovine v jetrih). Od 57 genov so 3 še posebej pomembni (zapisujejo za proteine, ki presnavljajo ~75 % vseh zdravil) in 2 od teh sta polimorfna. CYP2D6 pretvarja kodein (sredstvo proti bolečinam) v morfin. Pri ~10 % ljudi ta pretvorba ne deluje zaradi mutacije v genu za CYP2D6, ki je pomemben tudi za presnovo antidepressivov in beta-blokatorjev. Več biotehnoloških podjetij že prodaja DNA-mikromreže za določanje SNP, ki vplivajo na aktivnost encimov P450.

2002: Nekateri pacienti z AIDS-om imajo močnejše imunske reakcije na zdravilo abakavir. Primerjava več sto SNP med pacienti, ki z zdravilom niso imeli težav in tistimi, ki so imeli reakcije, je pokazala, da so občutljivi pacienti imeli razlike v 3 nukleotidih (SNP).

Adverse Drug Reactions

ADRs, WHO definition, 1966-96, 39 prospective studies

3.1 billion prescriptions per year
33 million hospital admissions
2.1 million serious ADRs (6.7%)
3 billion \$ hospitalization costs
106,000 fatal ADRs (0.32%)
5th leading cause of death

Metagenomika = Okoljska genomika = Ekogenomika

Vzorke za določanje genomskih zaporedij zberejo v okolju in hkrati določajo veliko število različnih (predvsem mikrobnih) genomov, ne da bi jih nujno med seboj ločili.

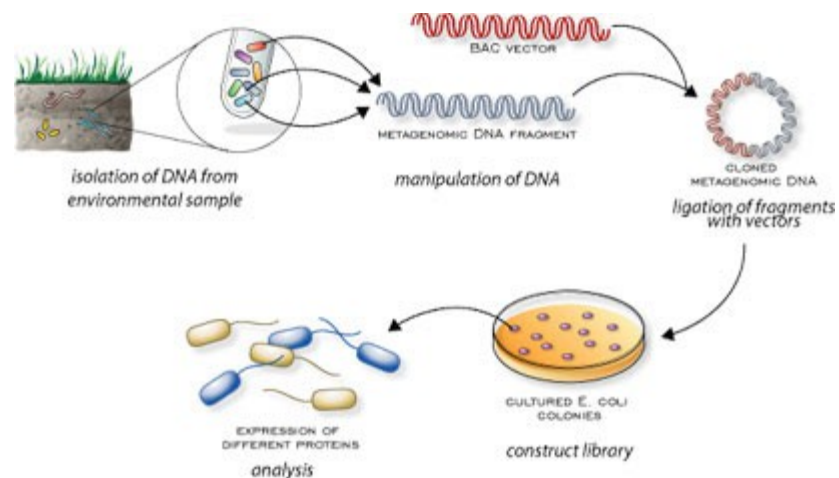
2002: v 200 litrih morske vode so našli >5000 različnih virusov

2003: v človeškem blatu je >1000 različnih virusov

2004: v 1 kg morskega sedimenta je ~1 milijon različnih virusov;

analiza voda iz Sargaškega morja – DNA kaže na prisotnost ~2000 vrst, od tega 148 tipov dotlej neznanih bakterij

→ ... podatki o mikrobni raznovrstnosti, sestavi ekosistemov, možnih metaboličnih procesih v mikrobni združbi



More than 9,000,000 Unique Genes in Human Gut Bacterial Community: Estimating Gene Numbers Inside a Human Body

Xing Yang^{1,2}, Lu Xie¹, Yixue Li^{1,3,4*}, Chaochun Wei^{1,3,5*}

Človek je 'sestavljeno organizem' s 23.000 lastnimi geni in več milijoni mikrobnih genov. Analizirali so genome 202 bakterij iz človeških prebavil in našli 552.700 različnih genov za proteine. V celoti naj bi človekova črevesna flora vsebovala ~9 milijonov različnih genov.

Človeka naseljuje ~100 trilijonov mikrobnih celic, kar je 10x več kot je človeških celic.

Table 2. Estimated total gene numbers for human gut bacterial community.

Similarity	0.97	0.90	0.80	0.60	0.30
Genes	8,988,806	6,533,896	5,799,165	4,071,772	2,932,368

The strain-to-species and species-to-genus gene counting used the same similarity at 0.97, 0.90, 0.80, 0.60 and 0.30 in the estimation. Detail of the gene counting model can be found in the gene counting model part of the Result section.

Transkriptomika

Transkriptomika se ukvarja z nabori prepisov (mRNA) v celicah. Transkriptom je med celicami v različnih fazah rasti, različnih bolezenskih stanjih ipd., različen.

Ugotovili so tudi, da se kot transkripti pojavljajo v veliki meri protismerna zaporedja, ki so pomembna pri uravnavanju izražanja.

Transkriptomika izvornih celic → vpogled v procese, pomembne za diferenciacijo celic.

Transkriptomika rakavih celic → razumevanje kancerogeneze.

Proteomika

Proteomika je veda, ki poskuša zaznati in opredeliti vlogo proteinov, ki med sabo interagirajo *in vivo*. Ob predpostavki, da vsak protein interagira s 5-50 drugimi proteini v celici, je število interakcij v celici kvasovke 30.000-300.000.

Proteini so po zgradbi bistveno bolj kompleksni od DNA, kar otežuje avtomatizacijo postopkov. Poznavanja nukleotidnega zaporedja predstavlja nepopolno informacijo, ker:

- raven mRNA pogosto ni merilo koncentracije proteina
- en gen zaradi alternativnega izrezovanja mRNA lahko zapisuje več proteinov
- proteini se posttranslacijsko modificirajo
- proteini imajo različne razpolovne čase
- proteini so lahko locirani v različnih razdelkih, kar vpliva na njihovo aktivnost
- nekateri proteini so aktivni šele, ko se sestavijo v kvartarne komplekse

Glavno orodje v proteomiki je **2D elektroforeza**, ki omogoča analizo kompleksnih zmesi proteinov, vendar ima veliko pomanjkljivosti (slaba topnost membranskih proteinov, območje detekcije proteinov je ozko, identifikacija proteinov v lisah je zahtevna). Razvoj gre v optimizacijo tehnike v povezavi z **masno spektroskopijo** za identifikacijo proteinov po končani ločitvi. Primerljivo zmogljivost kot 2D-elektroforeza imajo tudi predstavitev na fagih/ribosomih, dvohibridni sistemi ter proteinske mikromreže.

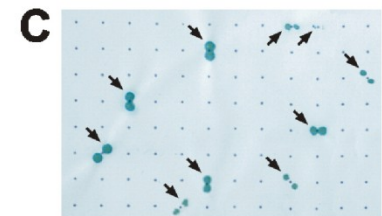
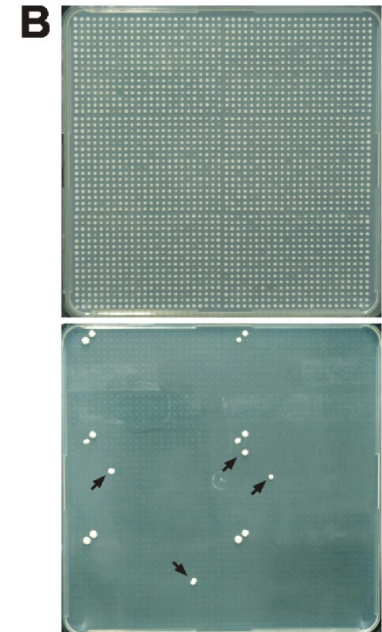
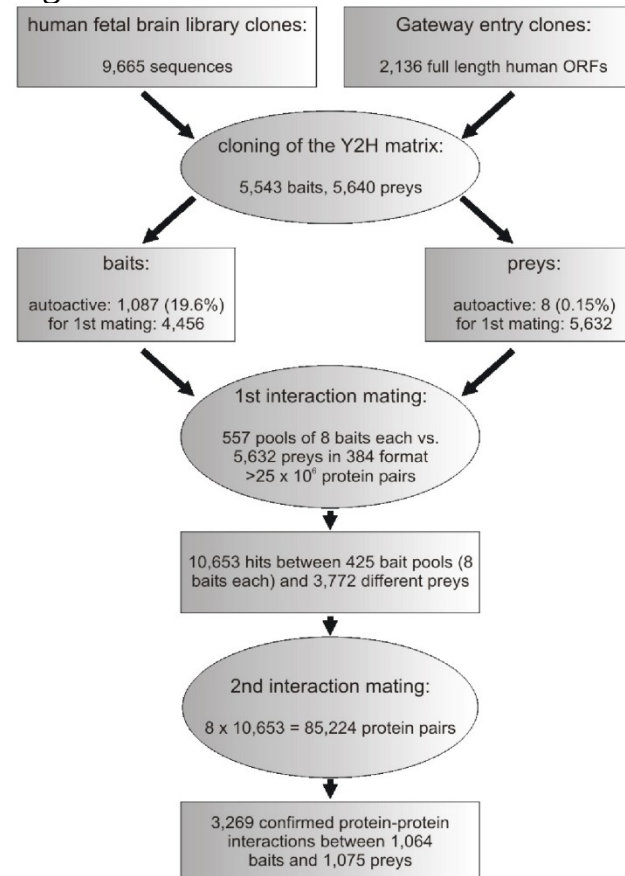
Izboljšani algoritmi za predvidevanje 3D zgradbe proteinov bodo omogočili širšo uporabo pri direktni pretvorbi podatkov z ravni DNA na raven proteinov. Vseeno pa bo veliko proteinov še vedno treba pripraviti s tehnikami rekombinantne DNA. Uveljavili se bodo verjetno vektorji, ki omogočajo učinkovito izražanje vnesenih genov v več različnih gostiteljskih sistemih.

Proteomika /2

Naloge proteomike:

- izdelava transkripcijskih profilov (kateri geni se prepisujejo v določenem tipu celic, razvojnem ali bolezenskem stadiju)
- razvoj postopkov za učinkovito izražanje in izolacijo proteinov
- metode za izdelavo proteinskih profilov v celicah (2D-PAGE/MS)
- študij interakcije med proteini (kateri proteini sodelujejo) – dvohibridni sistem kvasovk
- analiza biokemijskih poti – razumevanje signalnih in drugih kompleksnih procesov v celici
- obsežne študije zvijanja proteinov in določanja 3D-zgradbe
- bioinformatike analize proteomskih podatkov

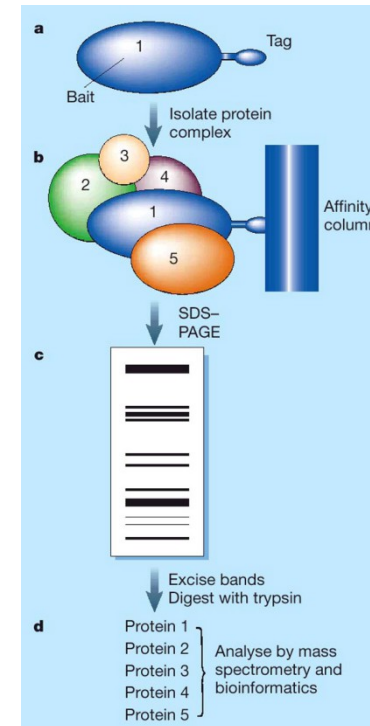
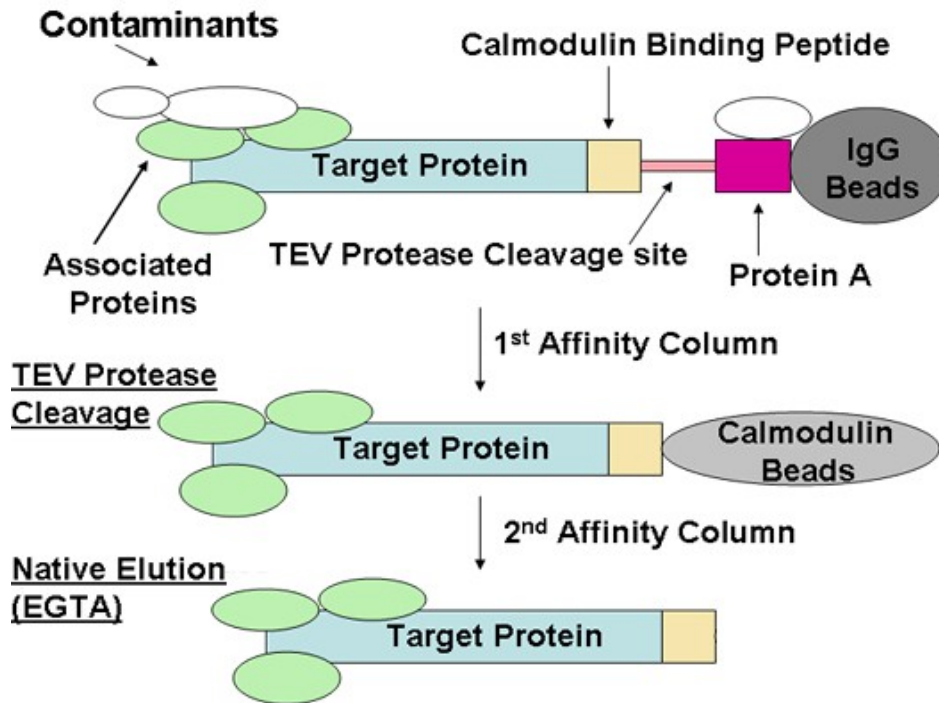
Določitev interakcij med proteini pri človeku (Cell 122, 2005):



Tandemsko afinitetno čiščenje (TAP)

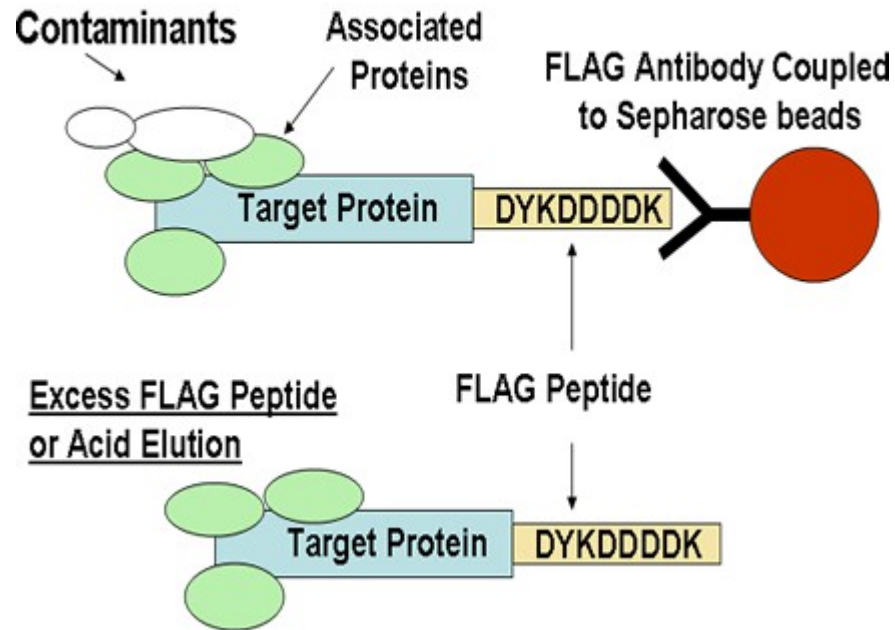
Konstrukti so pripravljani tako, da izrazimo rekombinantne fuzijske proteine, pri katerih je fuzijski del sestavljen iz dveh komponent: proteina, ki se veže na kalmodulin (CBP) in IgG-vezavne enote proteina A (ProtA) iz bakterije *Staphylococcus aureus*, med njima pa je cepitveno mesto za proteinazo TEV.

Proteine izrazijo v celicah, ki jih preučujejo. Celice lizirajo in fuzijski protein očistijo najprej preko imobiliziranih IgG, cepijo s proteinazo TEV, potem pa prečistijo preko imobiliziranega kalmodulina (elucija z EGTA). Eluirajo se rekomb. proteini, povezani s tistimi celičnimi komponentami, s katerimi interagirajo v celici.



FLAG – čiščenje

Konstrukti imajo kot oznako vezan peptid FLAG, proti kateremu so na voljo komercialna protitelesa. Izvedemo imunoafinitetno kromatografijo. Eluirajo se rekomb. proteini, povezani s tistimi celičnimi komponentami, s katerimi interagirajo v celici.



Proteomika: Celera



Cilj: odkriti diagnostične markerje in tarče za zdravljenje bolezni.

Osnovne tehnike: tekočinska kromatografija in MS

(ne pa 2D-elektroforeza) zaradi lažje robotizacije.

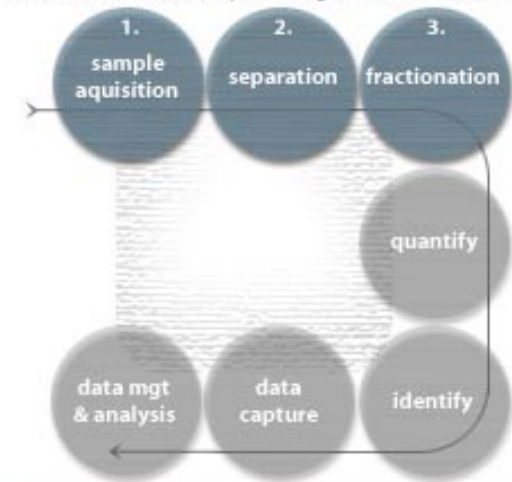
Uporabljajo aparature Applied Biosystems (LC-MS ESI-TOF in MALDI TOF-TOF);

določajo koncentracije posameznih peptidov in jih identificirajo.

Razvili so postopke za pripravo vzorcev, med njimi sisteme za sortiranje celic (do 50.000 celic/s), postopke za lovljenje specifičnih proteinov (tarč). Najti želijo proteine, ki se diferencialno izražajo v obolelih tkivih. Cilji so:

- raziskati vzorce izražanja proteinov v telesnih tekočinah za identifikacijo bolezenskih markerjev,
- identificirati proteine na površini celic, da bi našli tarče za terapevtska protitelesa,
- identificirati diferencialno izražene proteine v celičnih razdelkih, ki bi bili uporabni kot tarče za celično imunoterapijo,
- identificirati diferencialno izražene aktivne proteine, proti katerim bi lahko razvili nizkomolekularne učinkovine.

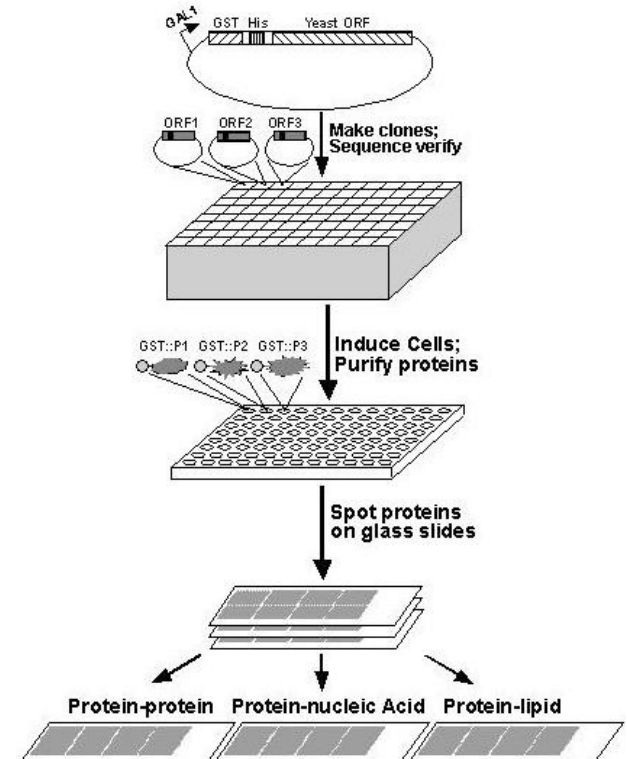
Proteomics Factory Configuration & Methods



1. Serum, tissue samples & related clinical information obtained from various sources
2. Separation to target specific proteins
Serum-deplete abundant proteins
Tissue-segregate desired cells (cell sorting), cell compartments
Biochemically enrich for targeted proteins such as cell surface membrane proteins
3. Enzymatic digestion of proteins into identifiable peptides for analysis

Proteomika: cilji, metode

- **Identifikacija proteinov:** 1D in 2D gelska elektroforeza
- **Izolacija in karakterizacija kompleksov:** Y2H, izol. označenih kompleksov
- **Aktivnost proteinov:** proteinske mikromreže



- Human Proteome Organization (HUPO, 2001), Plasma Proteome Institute (PPI, 2002): priprava proteoma plazme (vsi proteini v plazmi, tudi tisti, ki tja pridejo zaradi odmrtja nektrvnihih celic). *Razpon koncentracij posameznih proteinov v plazmi je 10 velikostnih razredov.* Obstajajo posamezni testi za 150 plazemskih proteinov, a je kombinacija testov predruga za rutinsko uporabo (17.000 USD) ⇒ pripraviti prstni odtis plazme in ga povezati z bolezenskimi stanji.
- Stanje tehnike trenutno še ne omogoča celovitih proteomskih analiz.

Metabolomika

Metabolomika poskuša razumeti metabolično dogajanje preko identifikacije metaboličnih produktov. Običajno ne gre za proteine, niti ne analizirajo kodirajočih ali regulatornih nukleinskih kislin, pač pa za nizkomolekularne metabolite.

2007: osnutek človeškega metaboloma – katalog 2500 metabolitov, 1200 zdravil in 3500 sestavin hrane, ki jih lahko najdemo v človeškem telesu. [<http://www.hmdb.ca/>]

Metabonomika: odgovor metabolizma na patofiziološke, okoljske ipd. spremembe.



Human Metabolome Database



HMDB Biofluid Browse

Search Biofluid for:

Search Biofluid Concentration: from to

Select Biofluid Type: Sorted by:

Metabolites Found in Plasma

(sorted by ascending common name)

Metabolite	Normal Concentration Range (μmol/L)	Abnormal Concentration Range (μmol/L)	Condition Associated with Abnormal Concentration Range or Specificity of Normal Range (Condition: Age (Gender))	Reference
(S)-3,4-Dihydroxybutyric acid MetaboCard	1.7 (1.2-2.2)		Normal: Adult (Both)	1167165
	18.0 (0.0-54.0)		Normal: Adult (Both)	8412012
(S)-3-Hydroxyisobutyric acid MetaboCard	20.0 (4.0-48.0)		Normal: Adult (Both)	8412012
1-alpha,25-Dihydroxyvitamin D2 MetaboCard	0.00012 +/- 0.000013		Normal: Adult (Both)	15702407
		0.00013 +/- 0.000026	Smith-Lemli-Opitz syndrome	15702407
1-Methylhistamine	0.00034 (0.00021-0.00047)		Normal: Adult (Both)	9234842

Biologija sistemov (sistemska biologija)

Integracija rezultatov genomike, proteomike, metabolomike: kompleksne interakcije & celovitost procesov v bioloških sistemih.

