

# C7188 Úvod do molekulární medicíny 5/12



## Moderní metodické přístupy v molekulární medicíně III



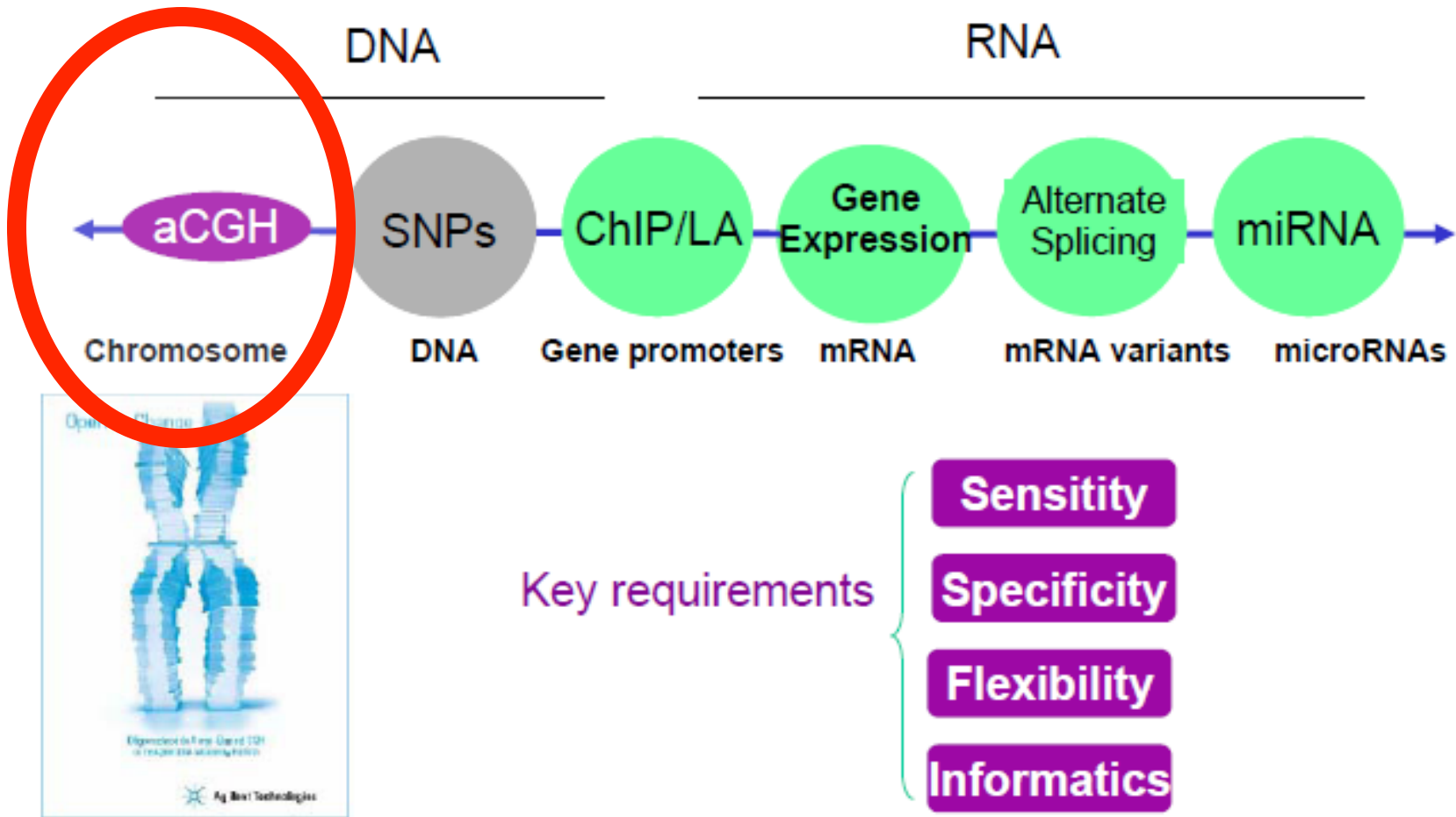
## GENOMIKA III, PROTEOMIKA



**Ondřej Slabý, Ph.D.**  
*Masarykův onkologický ústav  
CEITEC  
Lékařská fakulta Masarykovy univerzity*



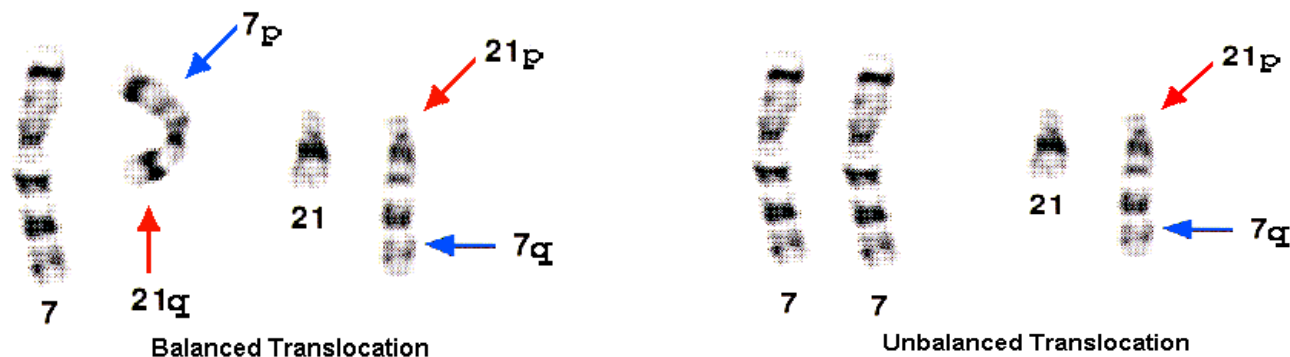
# Trend: More Microarray Applications



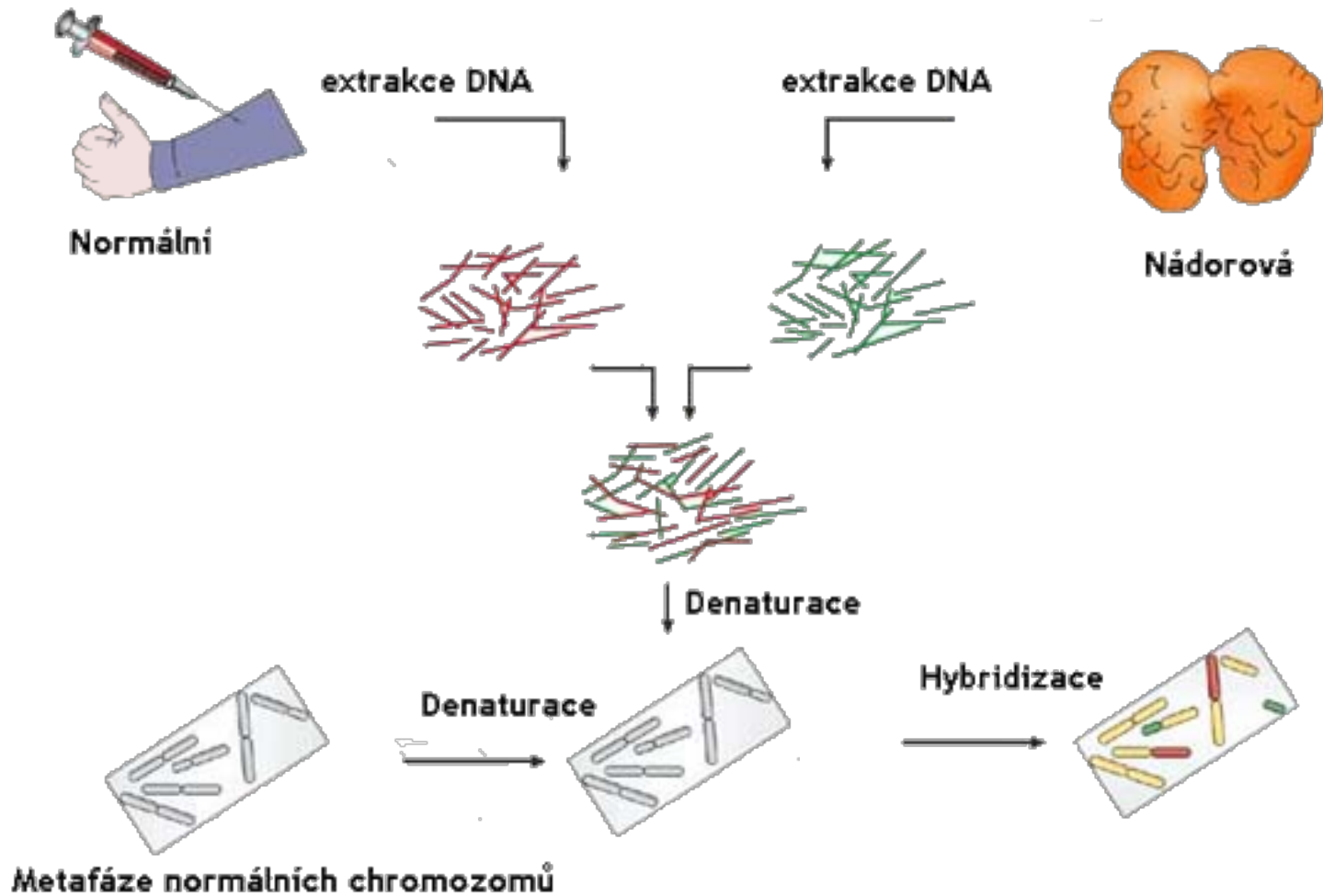
## Komparativní genomová hybridizace (CGH)

- Současná hybridizace dvou vzorků DNA značených odlišnými fluorochromy na normální metafázní chromozomy fixované na podložním skle
  - DNA pacienta (nádorová)  
značena **zeleně** (FITC)
  - kontrolní DNA (DNA zdravého jedince)  
značena **červeně** (TRITC)

Nelze detekovat balancované translokace a inverze, obecně změny při nichž nedochází ke kvantitativní změnám

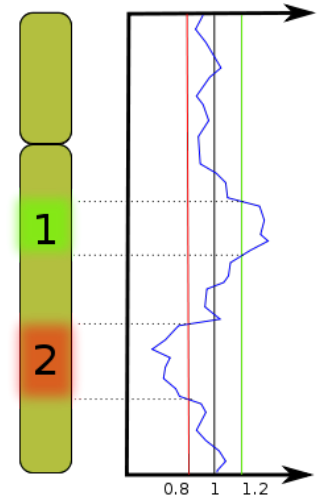
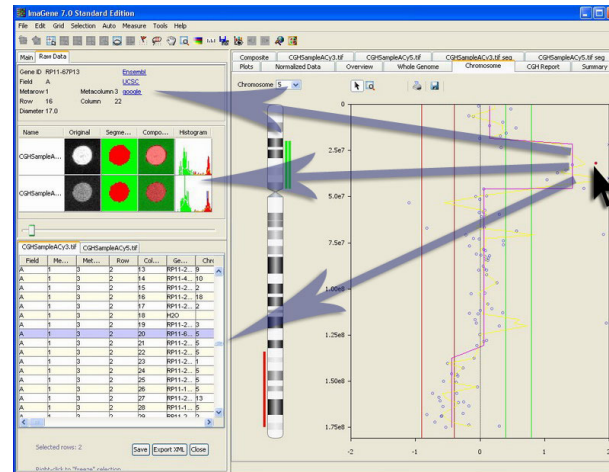
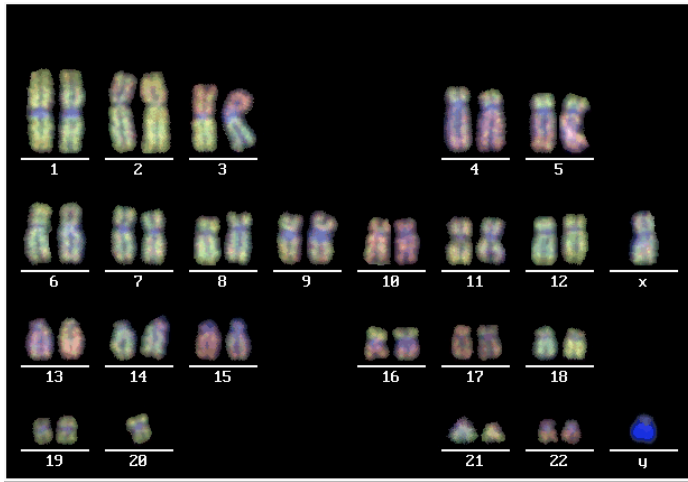


# Princip chromozomové CGH





# Výsledek chromozomové CGH



- vizualizován pomocí fluorescenčního mikroskopu
- pro každý fluorochrom je obraz snímán kamerou do počítače
- speciální softwarový program změří intenzitu obou fluorochromů podél každého chromozomu
- výsledkem analýzy: CGH profil - poměr intenzit signálů obou fluorochromů
  - poměr 1 = nedošlo k žádné kvantitativní změně
  - poměr  $<0,75$  = delece
  - poměr  $>1,25$  = duplikace, amplifikace

# Genomová array CGH (aCGH)

## Minulost – cDNA arrays for CGH

## Původ genomových sekvencí DNA pro aCGH

YAC [Yeast Artificial Chromosomes] (0,2-2 Mb)

BAC [Bacterial Artificial Chromosomes] (300 kb)

PAC [P1-derived Artificial chromosomes] (130-150 kb)

plazmidů (30-45 kb)

Rozlišovací schopnost array-CGH závisí právě na typu použitého vektoru, V každém případě je však řádově vyšší než u chromozomové CGH!!!

## Současnost – oligonukleotidové arrays CGH (oaCGH)

tiling arrays – překrývající se próby, 10 bp



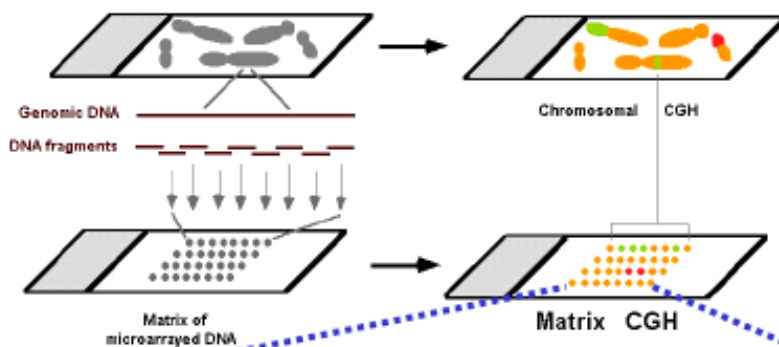
		Resolution	Coverage	
a	<b>Cytogenetics</b>			
		Karyotyping	> 10 Mb	Complete
		SKY	> 2 Mb	Complete
		Traditional CGH	> 2 Mb (cytoband)	Complete
		FISH (interphase)	≥ 20 Kb	Probe Specific
	FISH (metaphase)	≥ 100 Kb	Probe Specific	
b	<b>aCGH</b>			
		BAC	100 Kb (Spectral Genomics - 2 Mb)	Complete
		cDNA	2 Kb	Genes Only
	Oligo (60-mer)	0.06 Kb	Complete	

# Genomová array CGH (aCGH)

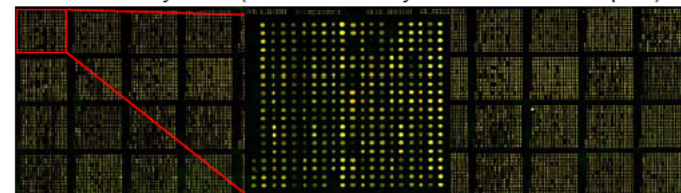
## Array-CGH analysis

Reference DNA  Test DNA 

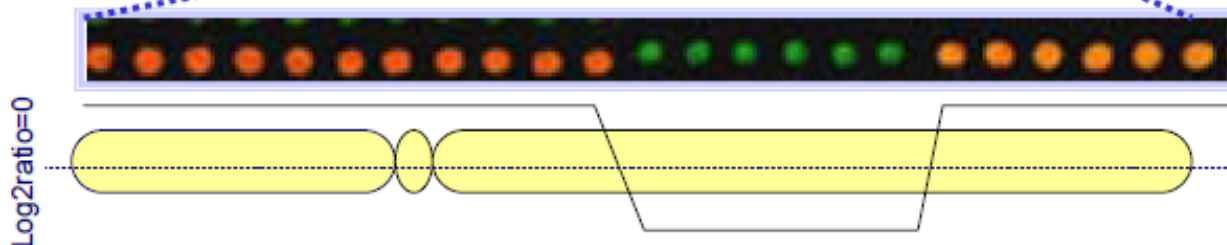
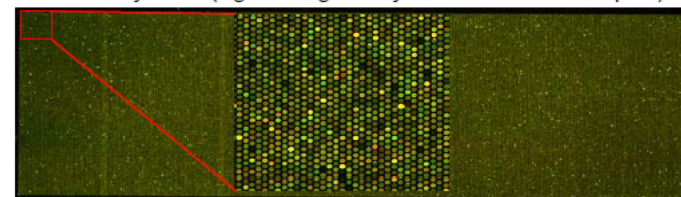
 = material gained in test genome  
 = material lost in test genome  
 = material balanced in test genome



1 Mb array-CGH (NMC BAC-array with 18.000 DNA spots)

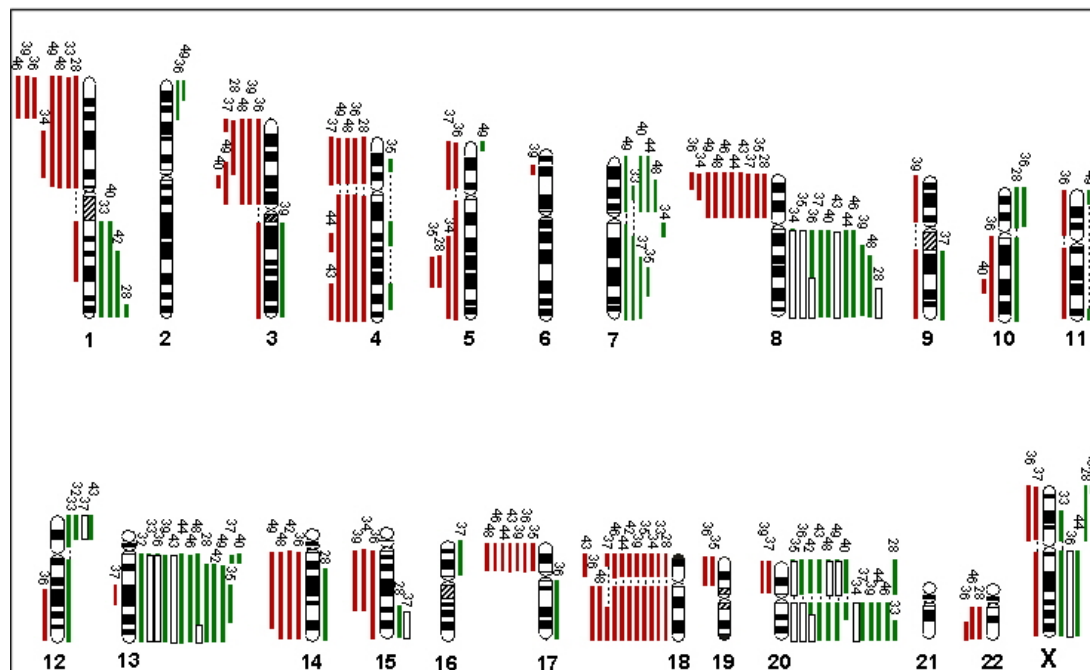


6 Kb array-CGH (Agilent oligo-array with 244.000 DNA spots)

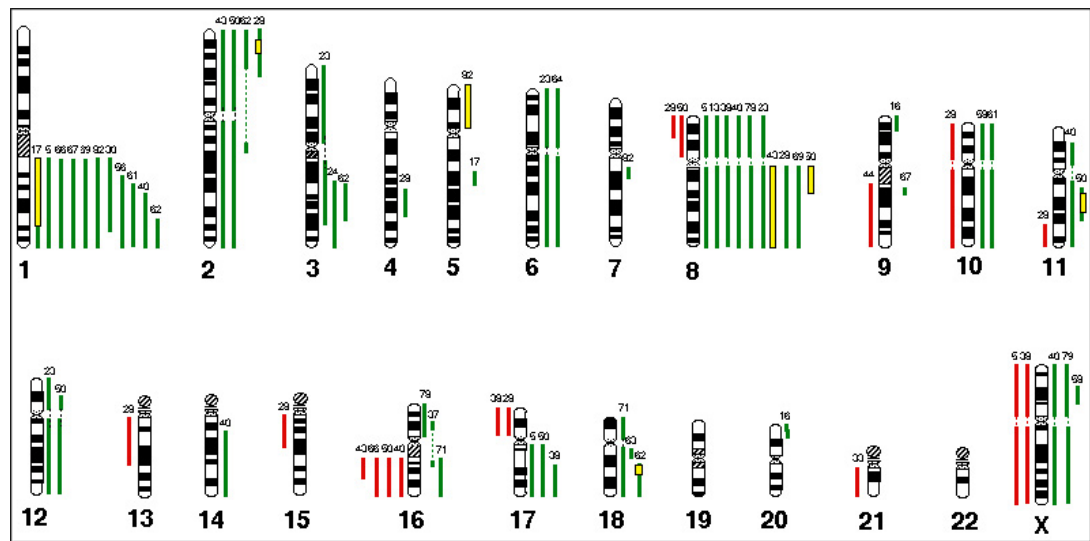


# CGH v onkologii

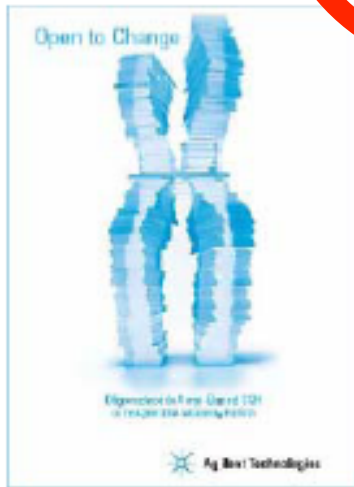
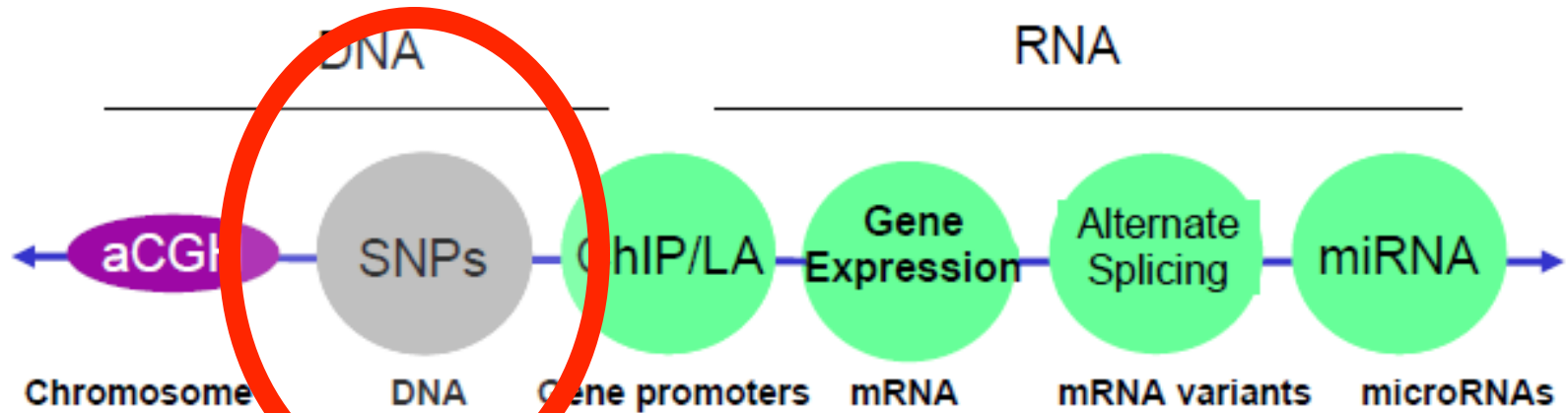
**kolorektální karcinom**  
20 primárních tumorů



**Karcinom hrdla děložního**  
20 primárních tumorů



# Trend: More Microarray Applications



Key requirements

- Sensitivity
- Specificity
- Flexibility
- Informatics

## SNP čipy

Za geneticky polymorfní je považován znak s nejméně dvěma geneticky podmíněnými variantami v jedné populaci, které se nachází v takových frekvencích, že i zřídka má frekvenci alespoň 1%.

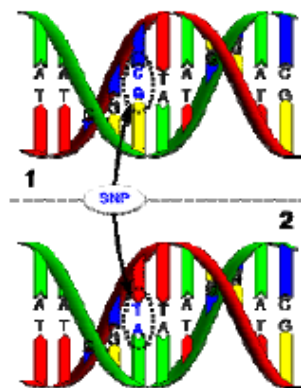
SNP = single nucleotide polymorphism, jsou jednonukleotidové polymorfní znaky  
Celogenomové mapy SNPs jsou dostupné ve webových databázích (~6 milionů)

Mezi lidmi je přibližně 99,9% shoda v sekvenci DNA

Zbývajících 0,1% nás činí jedinečnými (jak vypadáme, nemoci, kterými budeme trpět, ...)

Přibližně 1 SNP per 1.000 bp

90% genů obsahuje alespoň 1 SNP



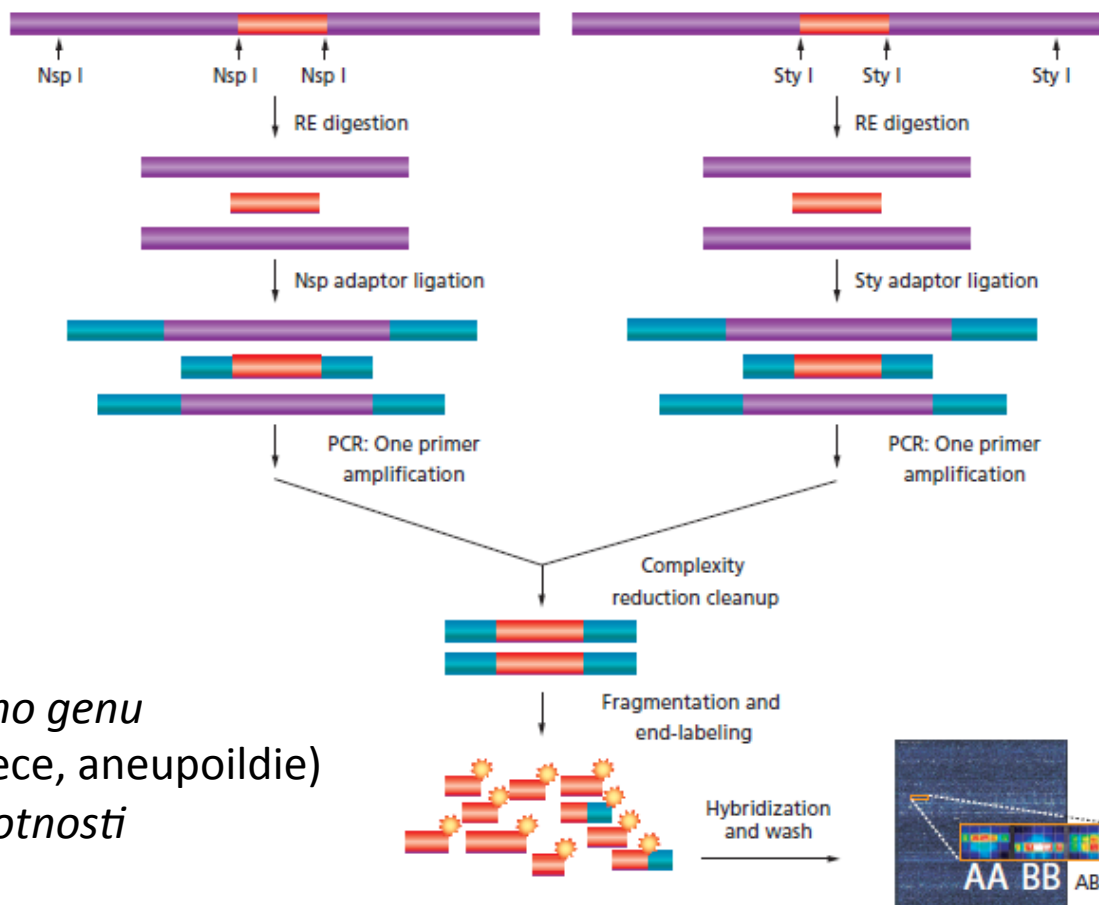
```

ATATGTATGTGGTATATAC-TCAAAATGTAATAAATTC-TAT
ATATGTATGTGGTATATAC-TCAAAATGTAATAAATTC-TAT
ATATGTATGTGGTATATAC-TCAAAATGTAATAAATTC-TAT
ATATGTATGTGGTATATAC-TCAAAATGTAATAAATTC-TAT
ATATGTATGTGTATATAC-TCAAAATGTAATAAATTC-TAT
ATATGTATGTGTATATAC-TCAAAATGTAATAAATTC-TAT
ATATGTATGTGGTATATAC-TCAAAATGTAATAAATTC-TAT
ATATGTATGTGGTATATAC-TCAAAATGTAATAAATTC-TAT
TATGTATGTGT-ATATAC-TCAAAATGTAATAAATTC-TAT
  
```

## Affymterix SNP čipy

Mapping 10K array => Mapping 100K array => Genome-wide Human SNP array 5.0 (500K) => Genome-wide Human SNP array 6.0 (1.8 million, SNP, CNV)

Figure 1: Overview of the Genome-Wide Human SNP Assay 5.0/6.0.



Umožňuje:

*Detekci SNP*

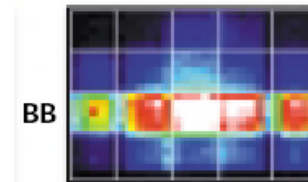
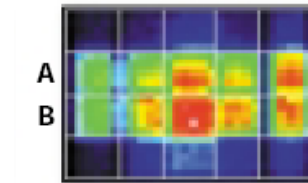
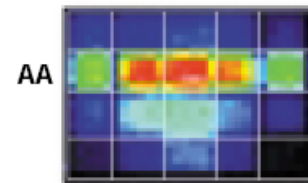
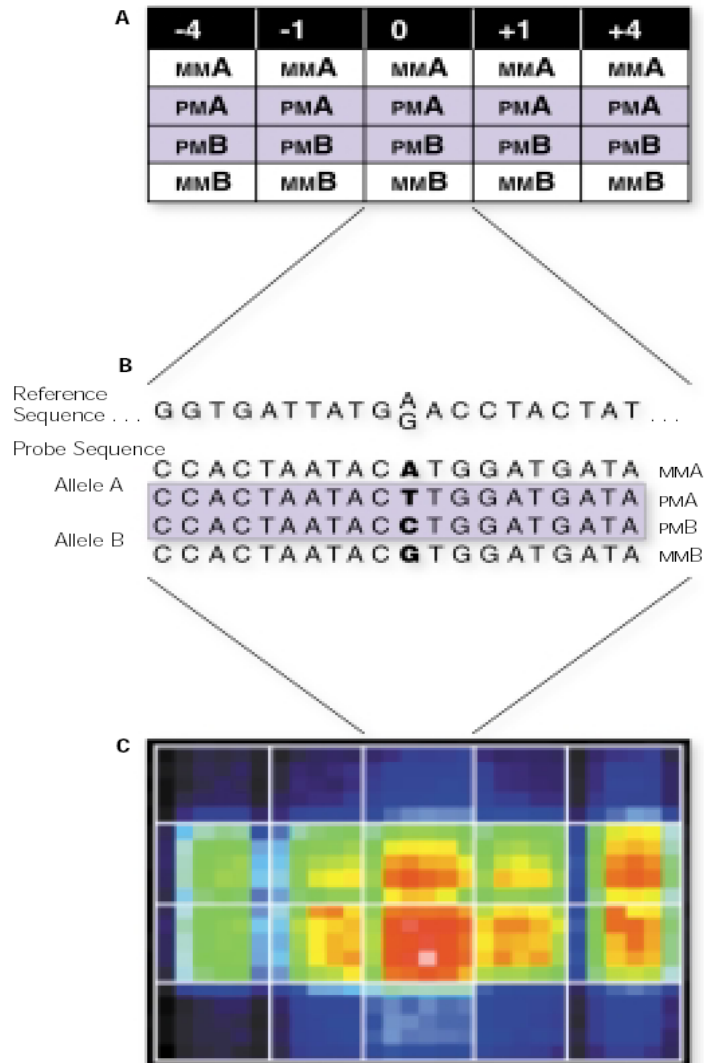
*Počet kopií daného genu*

*(amplifikace, delece, aneuploidie)*

*Ztráta heterozygotnosti*

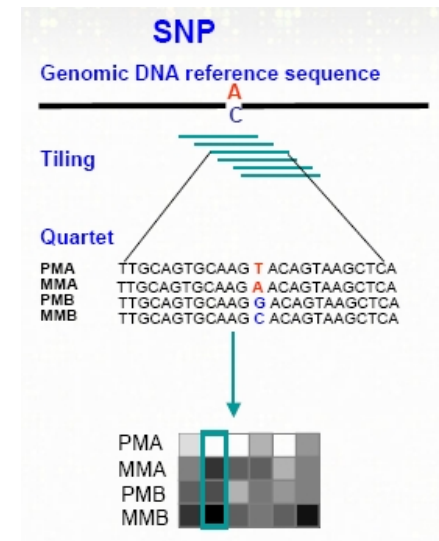
# Affymetrix SNP čipy

How the GeneChip® HuSNP™ Array Calls Genotypes



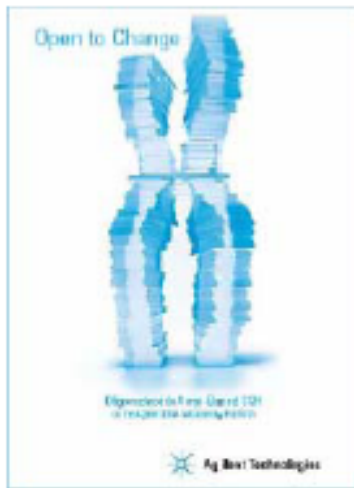
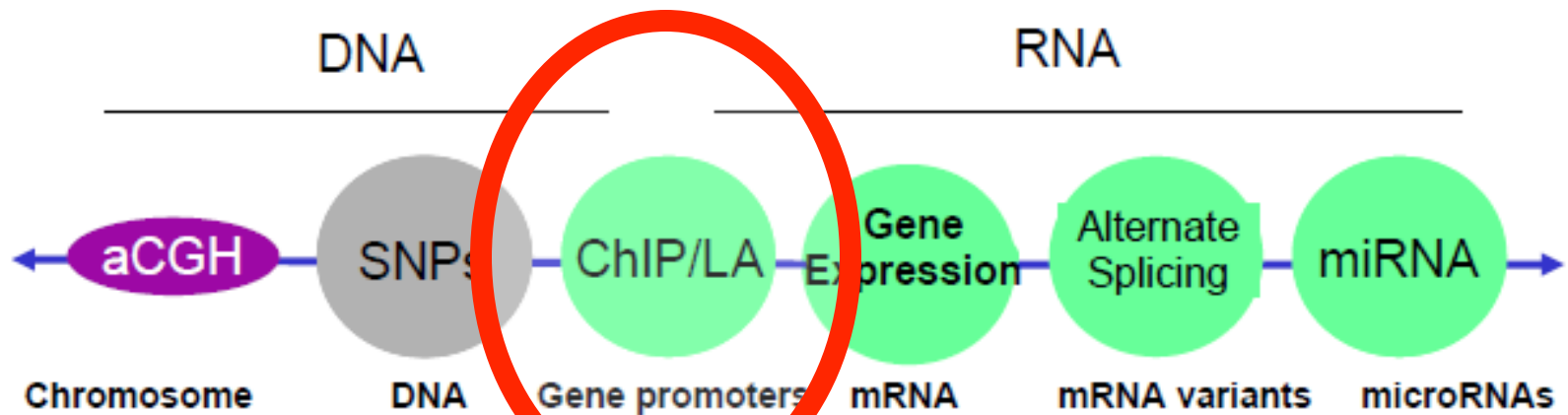
- HJ dané alely: AA, AB, BB
- Intenzita signálu: počet kopií

Pro každou alelu  
 5 sond  
 +  
 5 párových sond  
 s nukleotidovou záměnou  
 (např. v pozici -4)





# Trend: More Microarray Applications

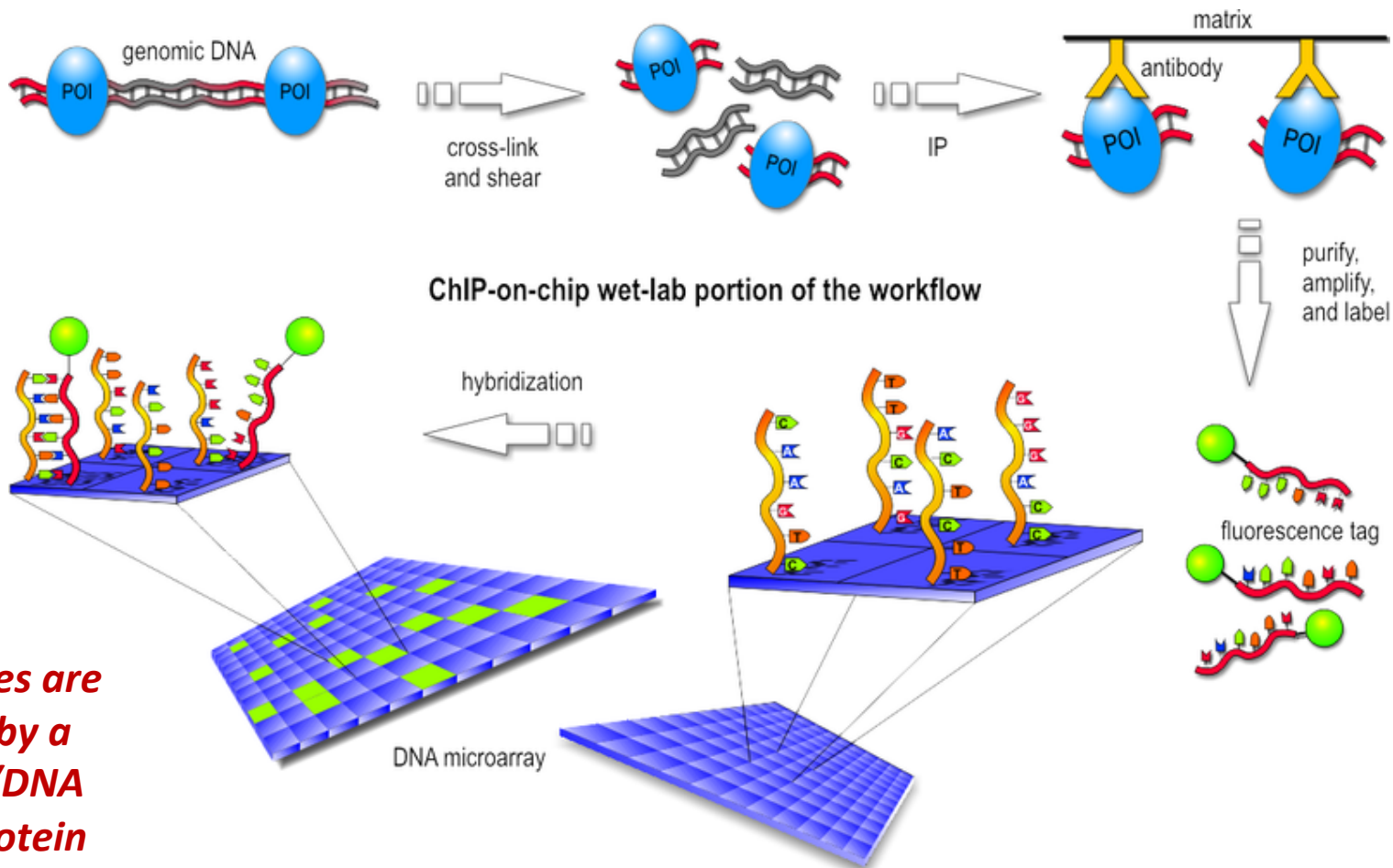


Key requirements

- Sensitivity
- Specificity
- Flexibility
- Informatics

## ChIP-on-chip technologie

kombinace chromatinové imunoprecipitace a čipové technologie

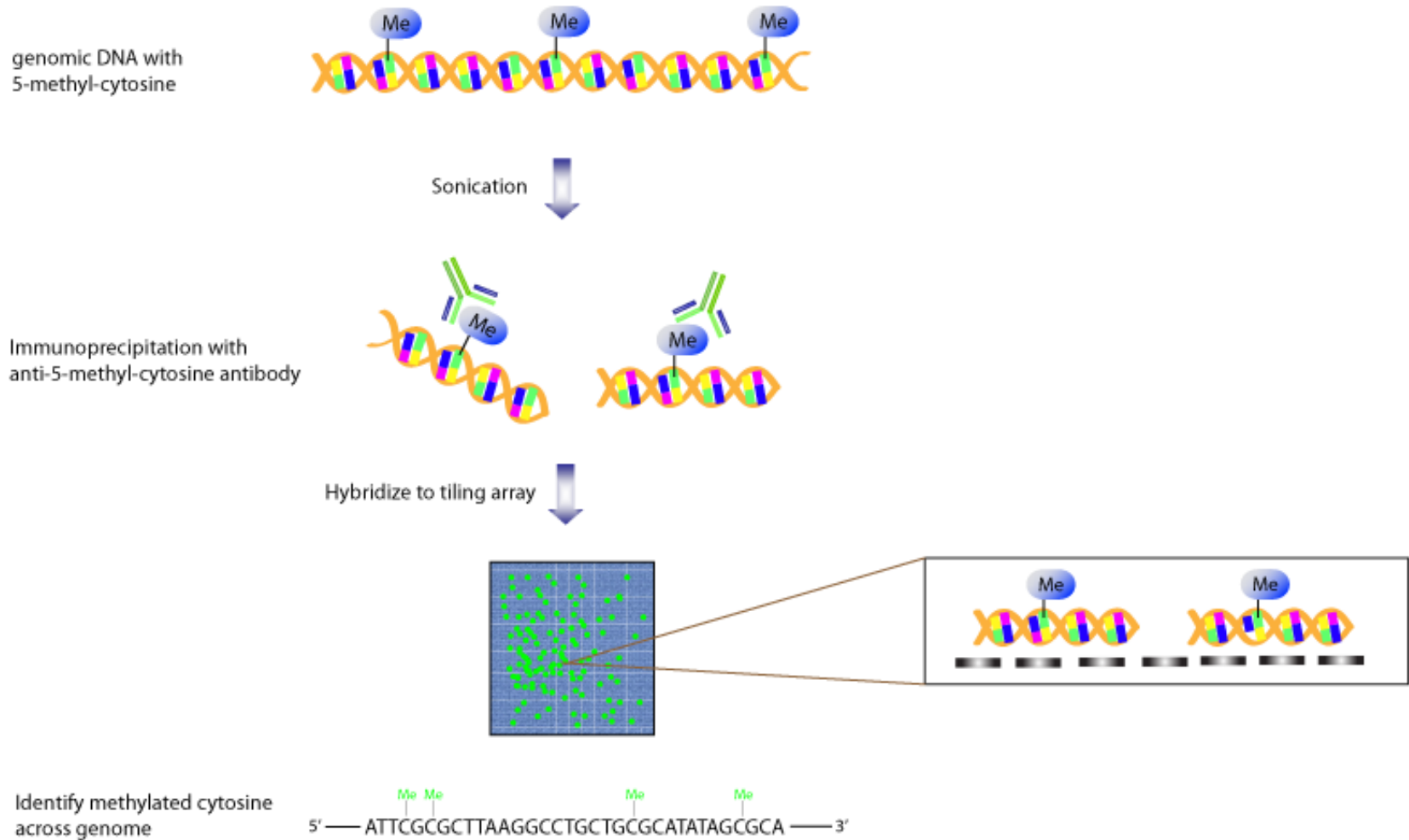


To answer:  
***which genes are regulated by a known TF/DNA binding protein***

**Discover protein/DNA interactions!!**

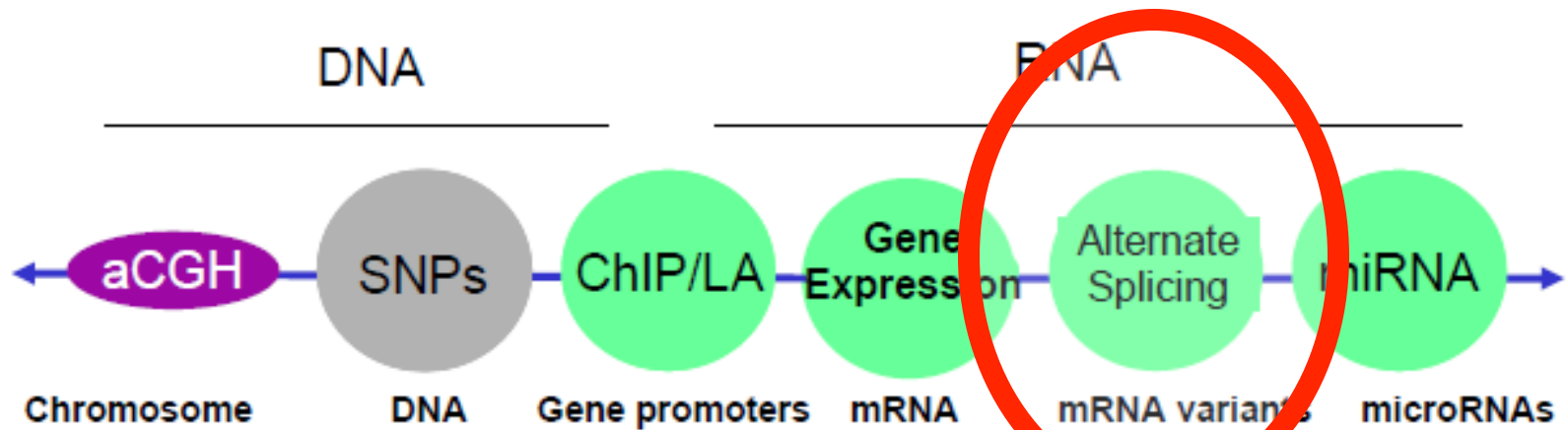
*Např. Agilent Technologies...*

## Technologie čipů k detekci methylace CpG oblastí



*Např. Agilent Technologies...*

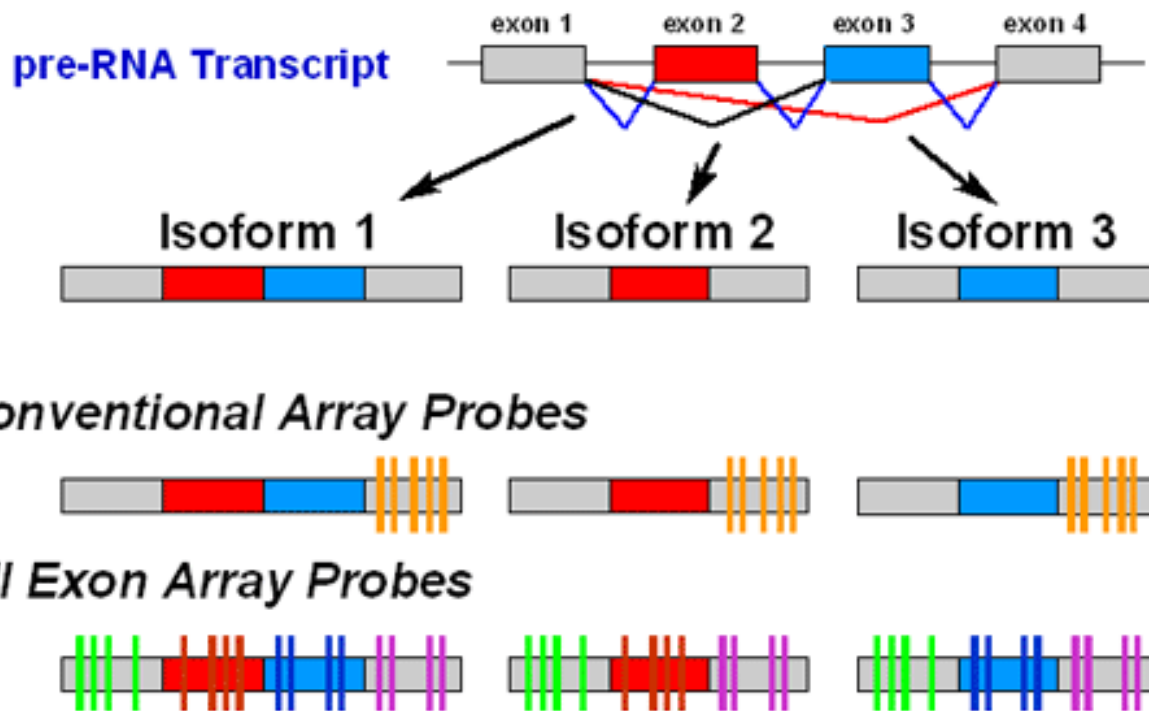
# Trend: More Microarray Applications



Key requirements

- Sensitivity
- Specificity
- Flexibility
- Informatics

## Technologie exonových čipů



***Umožňuje detekci, i nových, sestřihových variant známých genů***

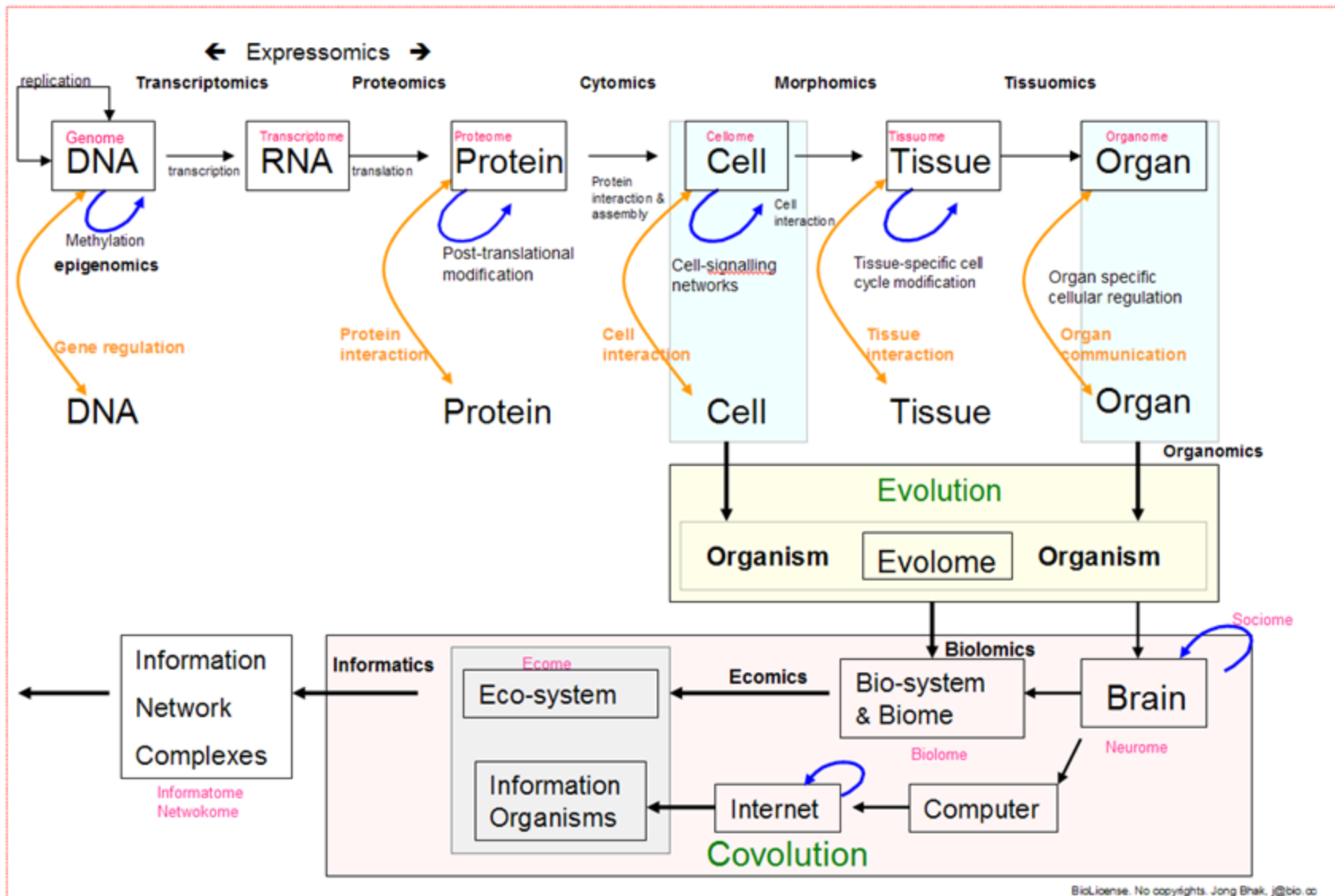
*Affymetrix GeneChip® Human Exon 1.0 ST Array*

## **-omics Mania**

biome, cellomics, chronomics, clinomics, complexome, crystallomics, cytomics, cytoskeleton, degradomics, diagnomics<sup>TM</sup>, enzymome, epigenome, expressome, fluxome, foldome, secretome, functome, functomics, **genomics**, glycomics, immunome, transcriptomics, integromics, interactome, kinome, ligandomics, lipoproteomics, localizome, phenomics, metabolome, pharmacometabonomics, methylome, microbiome, morphome, neurogenomics, nucleome, secretome, oncogenomics, operome, **transcriptomics**, ORFeome, parasitome, pathome, peptidome, pharmacogenome, pharmacomethylomics, phenomics, phylome, physiogenomics, postgenomics, predictome, promoterome, **proteomics**, pseudogenome, secretome, regulome, resistome, RNome, ribonome, ribonomics, riboproteomics, saccharomics, secretome, somatonome, systeome, toxicomics, transcriptome, translatome, secretome, unknome, vaccinome, variomics...

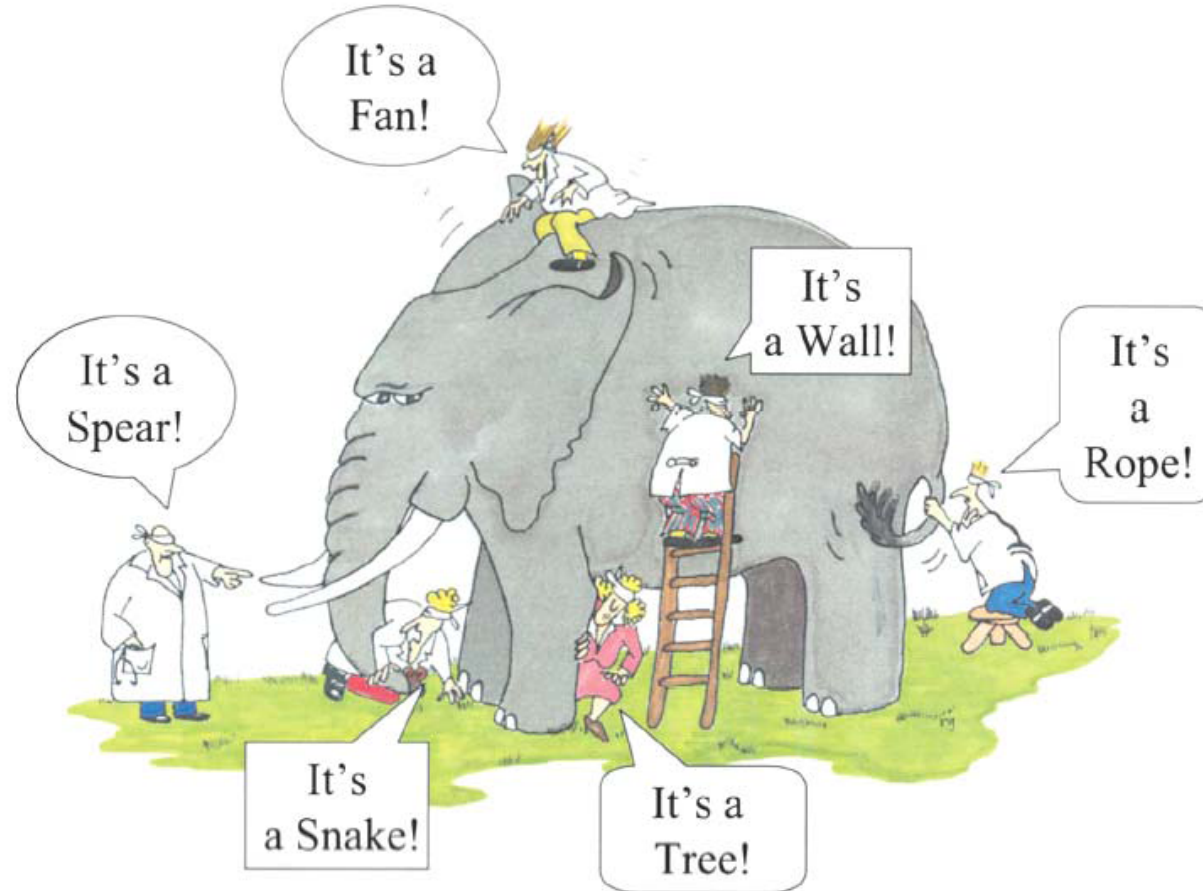
*An **omics** is a neologism referring to a broad field of study in biology, ending in the suffix "-omics" such as genomics, proteomics or Interactomics. The related neologism "'omes'" are the objects of study of the field such the genome or proteome, respectively ("omes" stems from the Greek for 'all', 'every', 'whole' or 'complete').*

# Omics Pathway version 1.0 20071231,



Biolicense. No copyrights. Jong Bhak. j@bio.ac

**Different -omics sciences describe many levels of biomolecular organization – but if used in isolation may give misleading inferences about the system!!!**



**Fig. 1.** *The blind men and the elephant.* Poem by John Godfrey Saxe (Cartoon originally copyrighted by the authors; G. Renee Guzlas, artist).

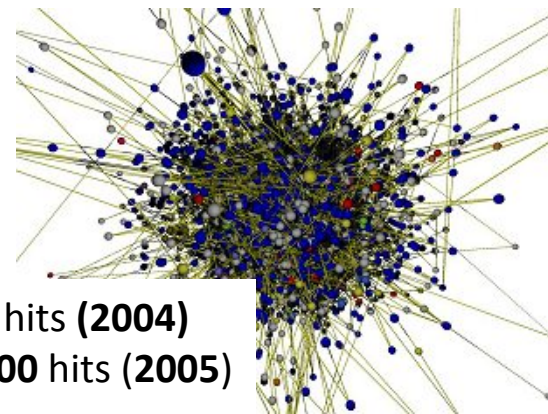


# JEDEN GENOM, DVA PROTEOMY



# PROTEOMIKA

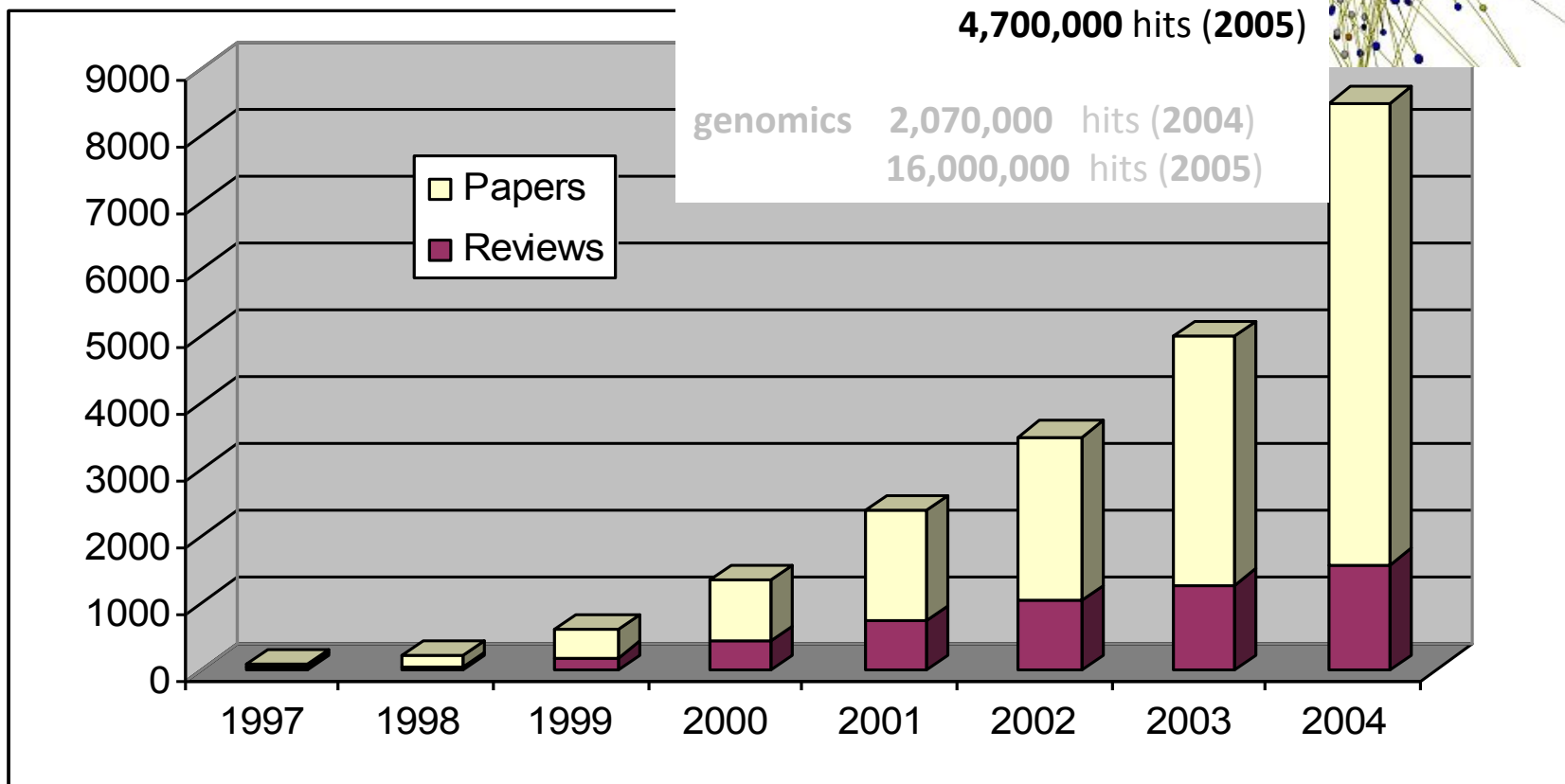
From **220** publications in the previous millennium ('94-'99)  
 To **21,350** (!!!) publications in this millennium ('00-'05)



**GENOMICS 81975**  
**PROTEOMICS 36494**

**proteomics 886,000 hits (2004)**  
**4,700,000 hits (2005)**

**genomics 2,070,000 hits (2004)**  
**16,000,000 hits (2005)**



## Co je to proteomika?

Proteomika je obor, který se zabývá systematickou analýzou proteinů z hlediska jejich identity, množství a funkcí

**Proteomika**

MarcWilkins, 1994

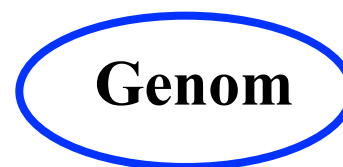


**Genomika**

PROTEiny exprimované genOMem



**Expresa**



+posttranslační modifikace

+alternativní sestřih

+alternativní zavinutí



➤ Souhrn všech proteinů v daném organismu

➤ Lidské tělo obsahuje miliony proteinů

➤ Expresa proteinů v rámci jednoho organismu se liší

➤ Souhrn všech genů v daném organismu

➤ Lidský genom obsahuje 20-25.000 genů

➤ Genom je konstantní celek

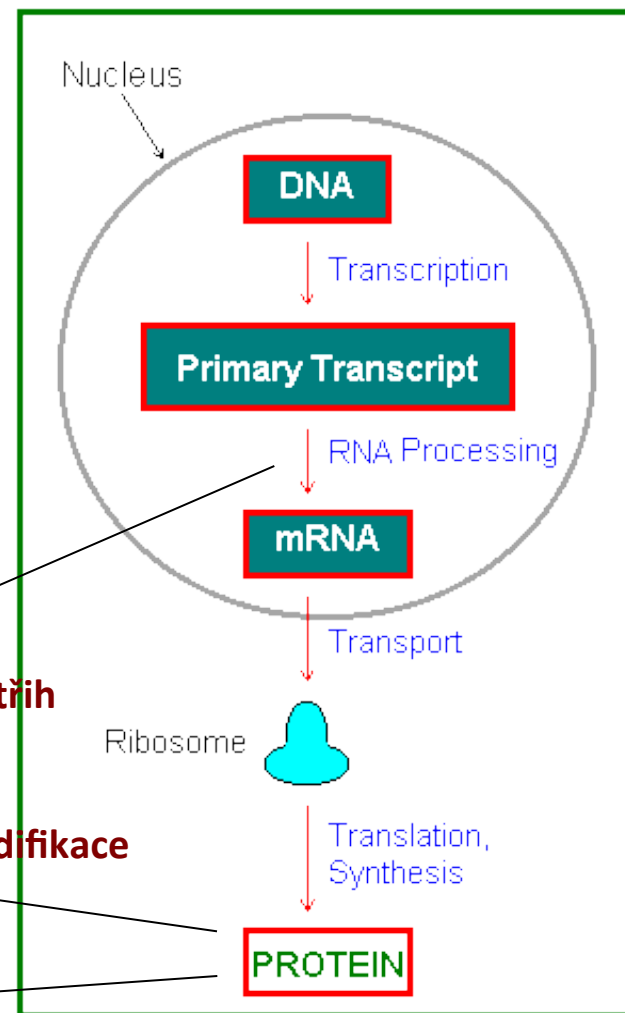
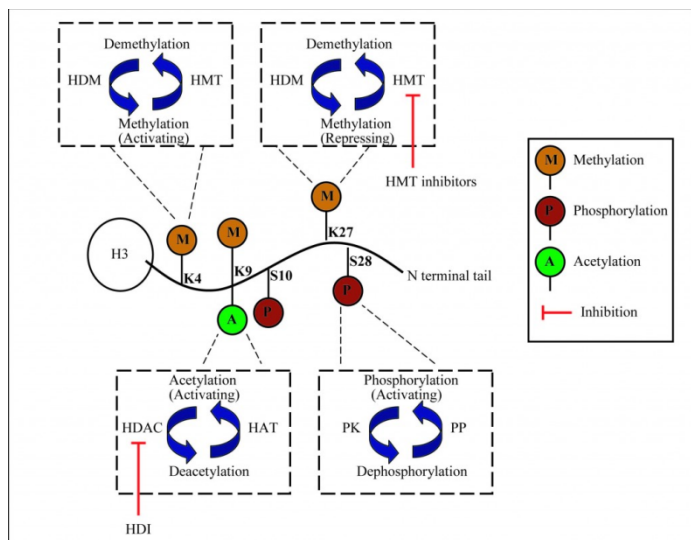
# Nárůst diverzity proteinů

## ➤ Posttranslační modifikace

1. Připojení funkčních skupin (acetát, fosfát, lipidy, cukry)
2. Modifikace amino skupin
3. Strukturální změny (tvorba disulfidických vazeb, proteolytické štěpení)

## ➤ Alternativní sestřih

## ➤ Alternativní zavinutí



Alternativní sestřih

Posttranslační modifikace

Alternativní zavinutí

## Přístupy k proteomickému studiu

Podle cíle:

*proteomika analytická, strukturní, funkční, ...*

Podle způsobu provedení:

*proteomika diferenciální, high-throughput...*

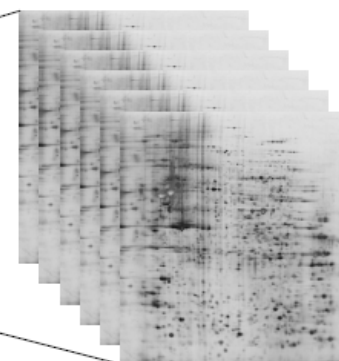
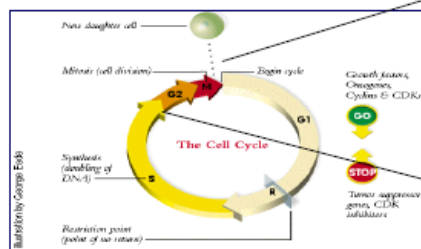
**High-throughput proteomika** je zaměřena na rychlé získávání údajů o přítomnosti bílkovin, což je důležité zejména pro screeningové účely ve zdravotnictví, zemědělství, kontrole potravin atd.

**High-coverage proteomika** se zabývá získáváním údajů o sekvenci aminokyselin včetně post-translačních modifikací bílkovin s cílem co nejvyššího pokrytí primární struktury (tj. stanovením pořadí co nejvyššího počtu aminokyselinových zbytků vbílkovině), což je důležité při studiu role bílkovin v životních procesech.

## Přístupy k proteomickému studiu

### *Funkční proteomika*

se zabývá nejen funkcí bílkovin, ale zejména studiem komplexních biologických procesů, jejichž pochopení má zásadní význam pro studium vývoje organismu či mechanismu a léčby nemocí.



**Shotgun proteomika** je obecný postup pro identifikace bílkovin, kdy je neseparovaná směs bílkovin nejprve enzymově rozštěpena na směs peptidů, které jsou následně separovány vhodnou metodou (nejčastěji HPLC) a potom sekvenovány tandemovou MS (často jsou využívány i nespécifické enzymy např. Proteinasa K).

## Obecné schéma klasického proteomického experimentu

Separace směsi bílkovin

2D-elfo, chromatografie LC, a jejich kombinace

Výběr proteinů  
pro indentifikaci

Štěpení vybraných roteinů  
a čištění peptidů

Enzymatické či chemické  
štěpení spotů nebo LC frakcí,  
čištění

Měření přesné hmotnosti/náboje peptidů  
případně jejich fragmentů

Získání hmotnostních spekter

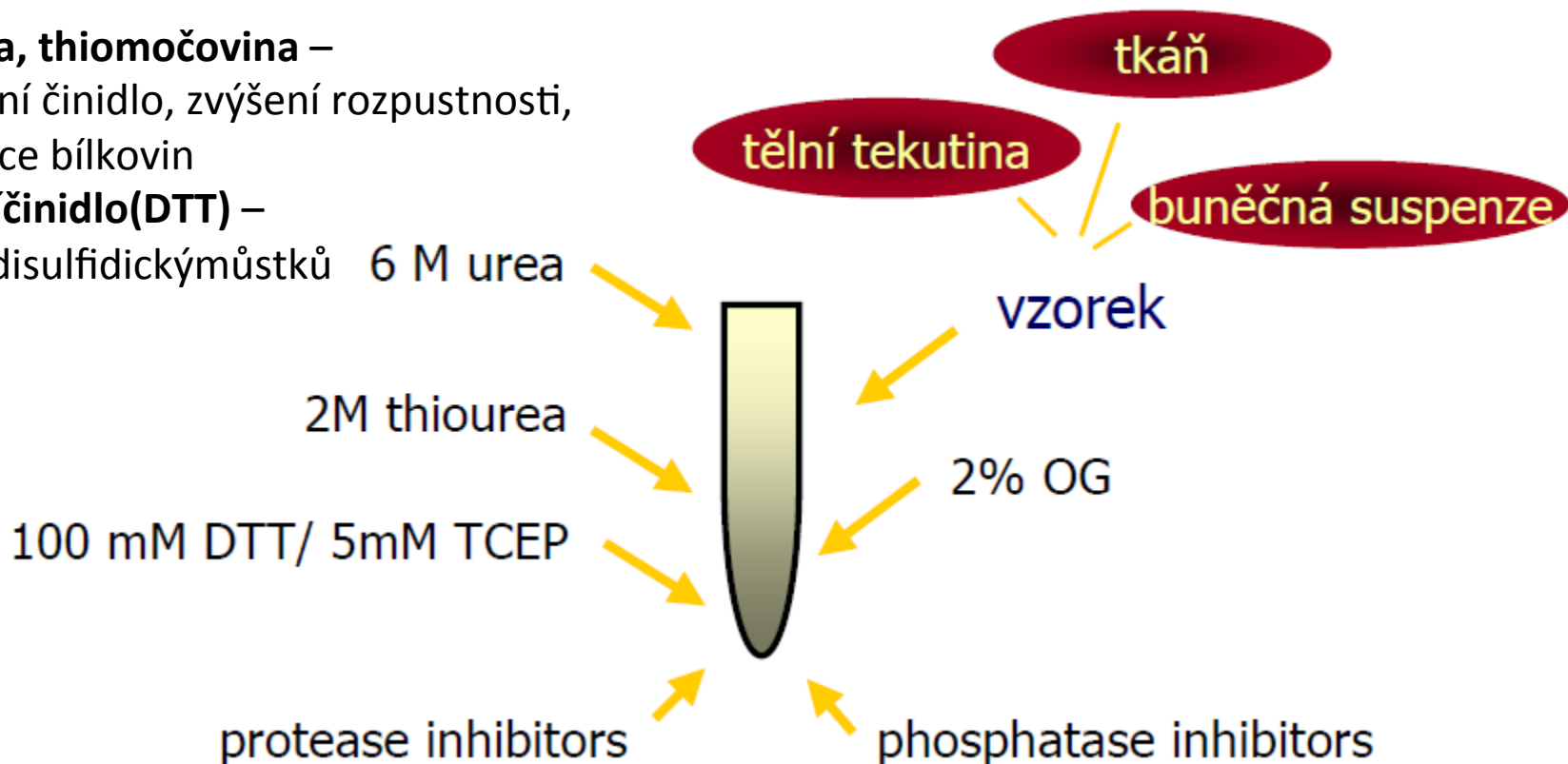
Identifikace proteinů

Porovnání hmotnostní sady peptidů s údaji dostupnými v databázích.

## Příprava vzorku pro proteomické analýzy

**Močovina, thiomčovina** –  
chaotropní činidlo, zvýšení rozpustnosti,  
denaturace bílkovin

**Redukční činidlo (DTT)** –  
redukce disulfidických můstků



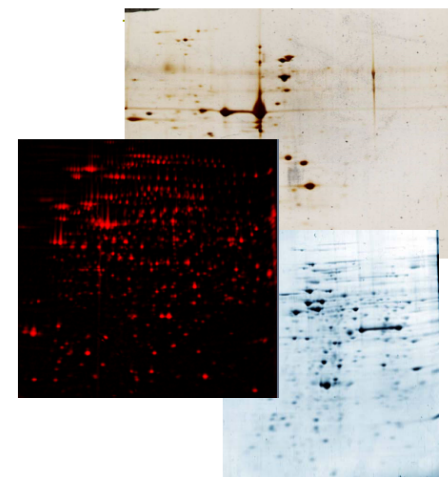
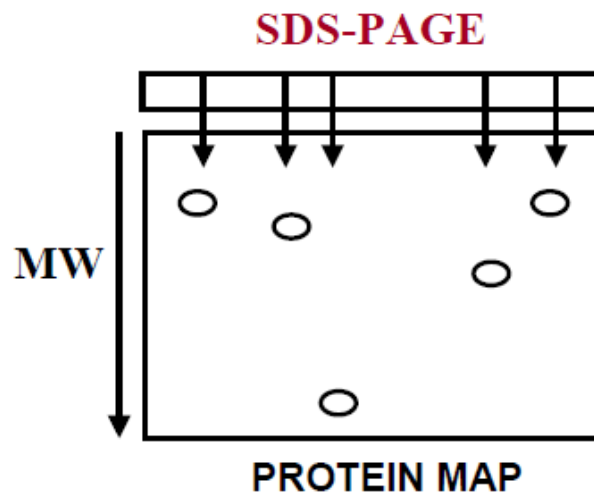
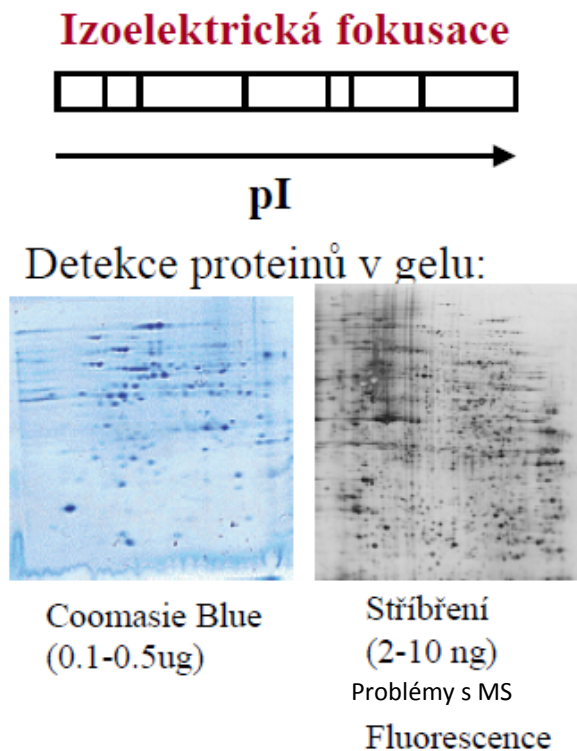
Analýza „subproteomů“

cytoplasma, ribosomy: diferenciální centrifugace

membránová frakce: diferenciální centrifugace, extrakce detergenty  
(TX-114) nebo extrakce  $\text{Na}_2\text{CO}_3$



# Dvojrozměrná gelová elektroforéza (2D-ELFO) – SEPARACE



## SDS-PAGE (Sodiumdodecylsulfat–polyakrylamidová elektroforéza elektroforéza)

Bílkoviny se rozdělují na základě jejich MW

Záporně nabitý SDS tvoří komplexy s bílkovinou a stíní náboje proteinu (1,4g SDS :1g proteinu). SDS uděluje proteinům uniformní náboj na jednotku hmotnosti

## IEF

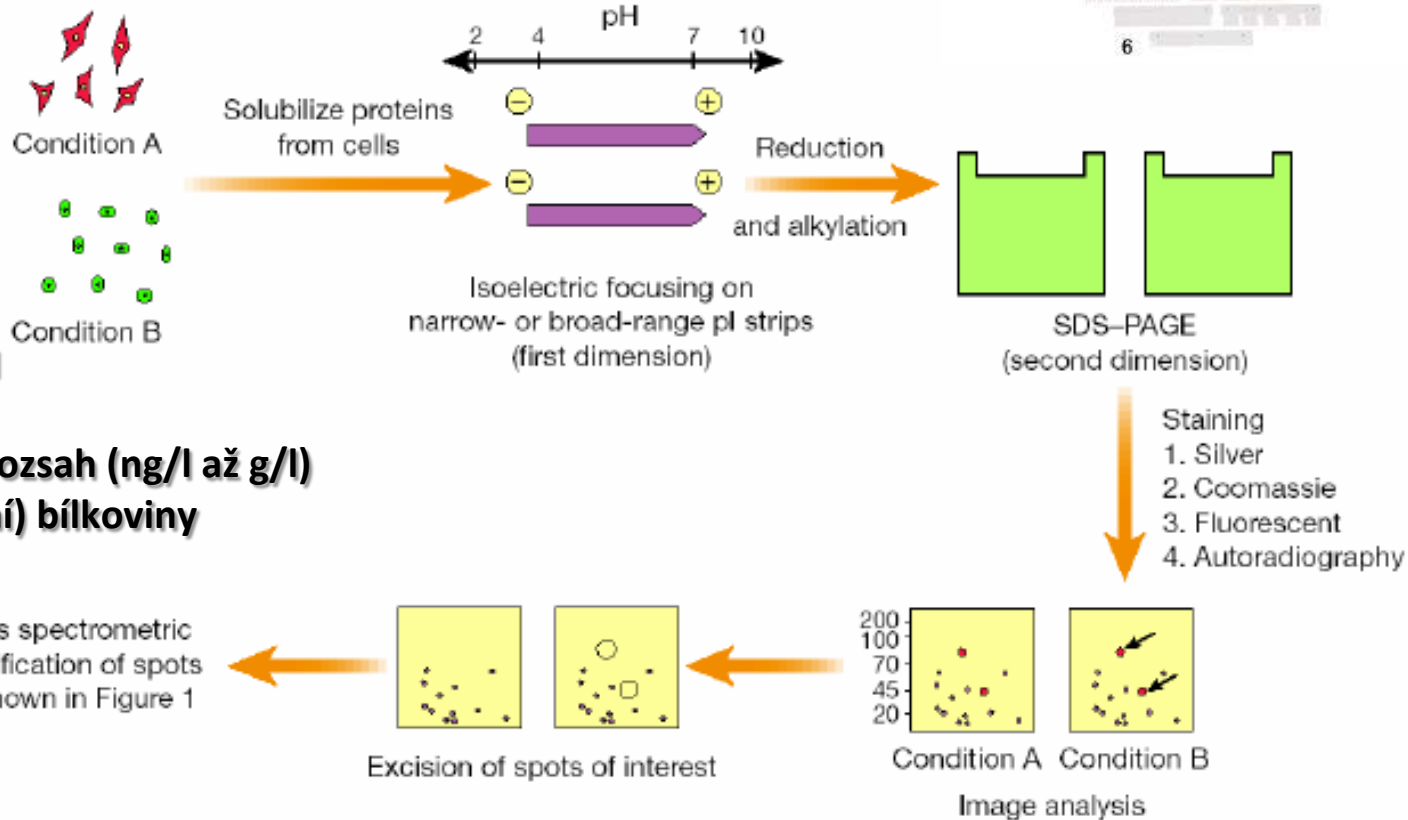
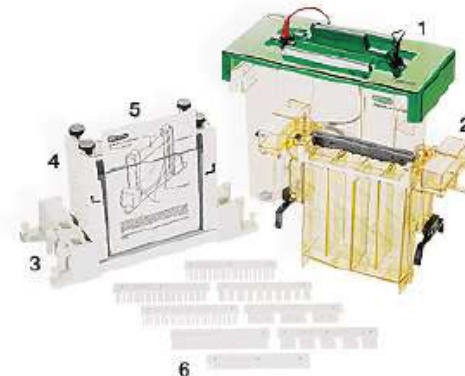
**Celkový náboj proteinu**(net charge) je součtem všech jeho negativních i pozitivních nábojů.

**Kyselé a zásadité skupiny** jsou protonovány a deprotonovány v závislosti na pH okolí.

**Amfoterní molekula** (bílkovina) v elektrickém poli v gradientu pH migruje do bodu, kde je její celkový náboj rovný nule.

pH tohoto bodu odpovídá izoelektrickému bodu (pI) dané bílkoviny.

# 2D-ELFO



**LIMITACE 2D-ELFO!!!!**

**Bílkoviny s extrémním pI**

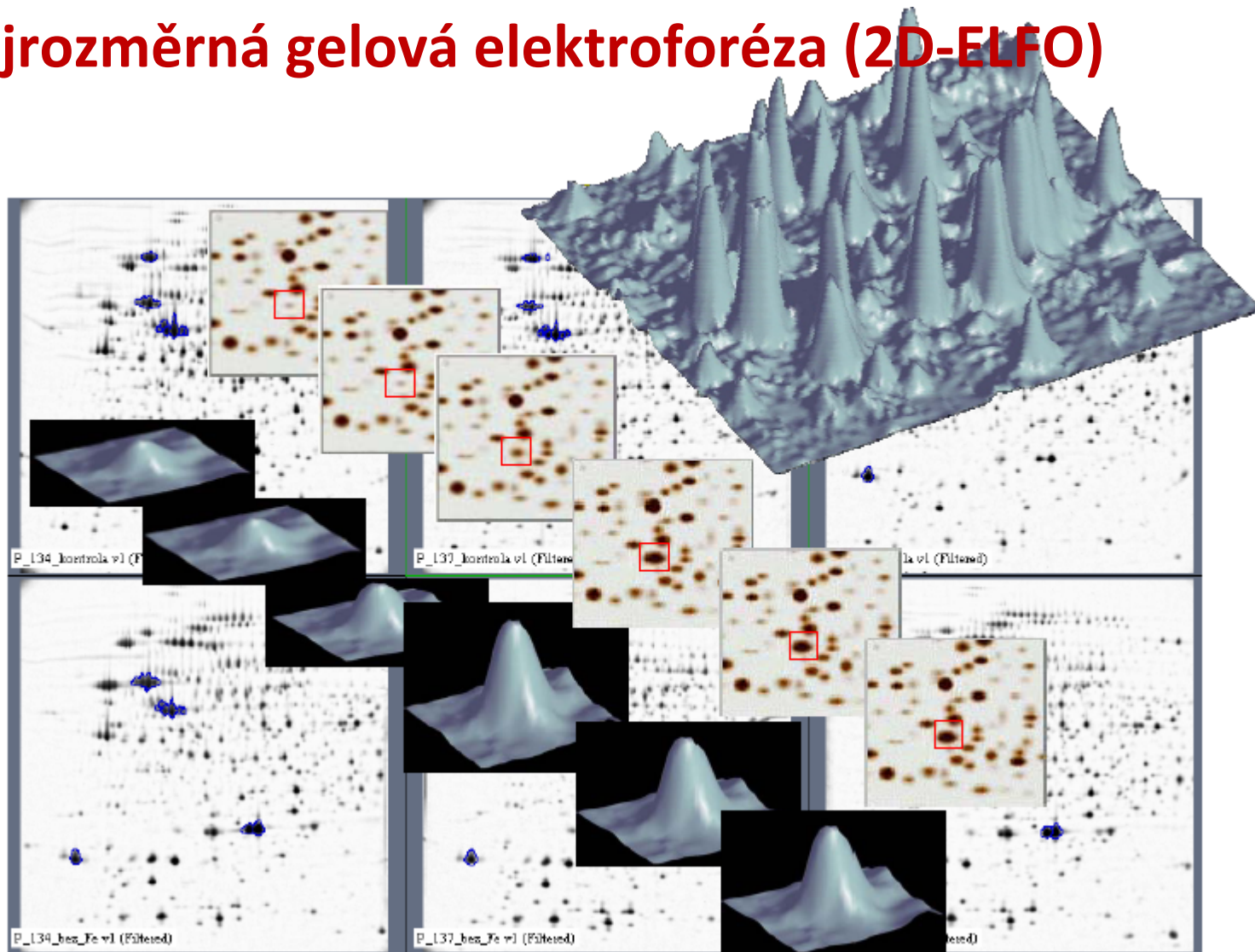
**Bílkoviny nad 150 kDa**

**Ohromný koncentrační rozsah (ng/l až g/l)**

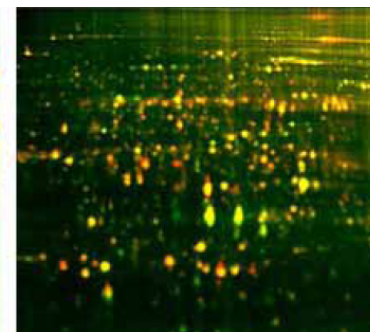
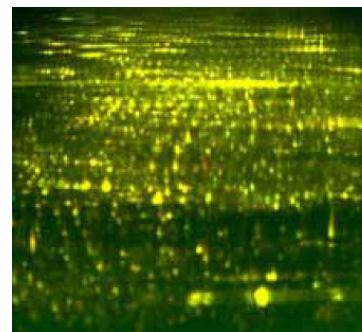
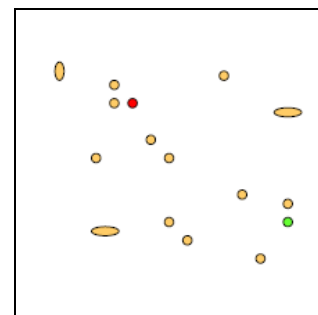
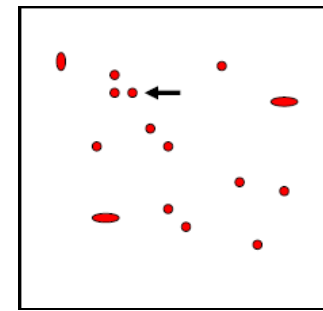
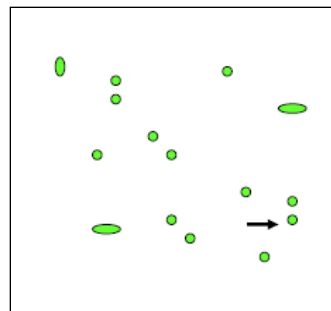
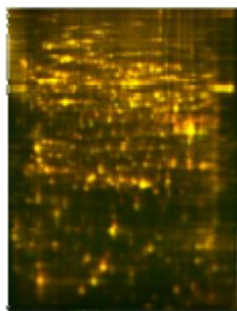
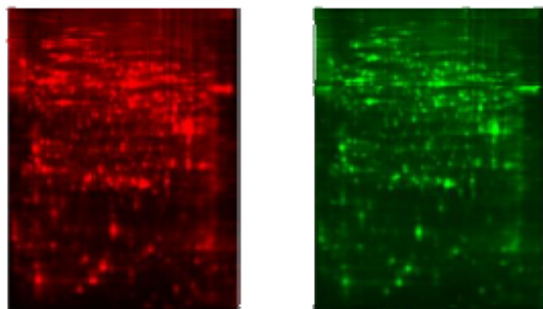
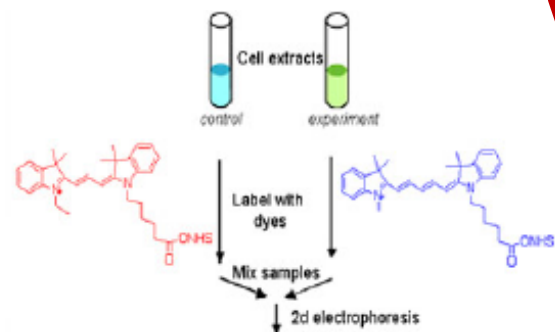
**Membránové(hydrofobní) bílkoviny**

Pandey & Mann Proteomics to study genes and genomes  
 NATURE | VOL 405 | 15 JUNE 2000 |

# Dvozměrná gelová elektroforéza (2D-ELFO)

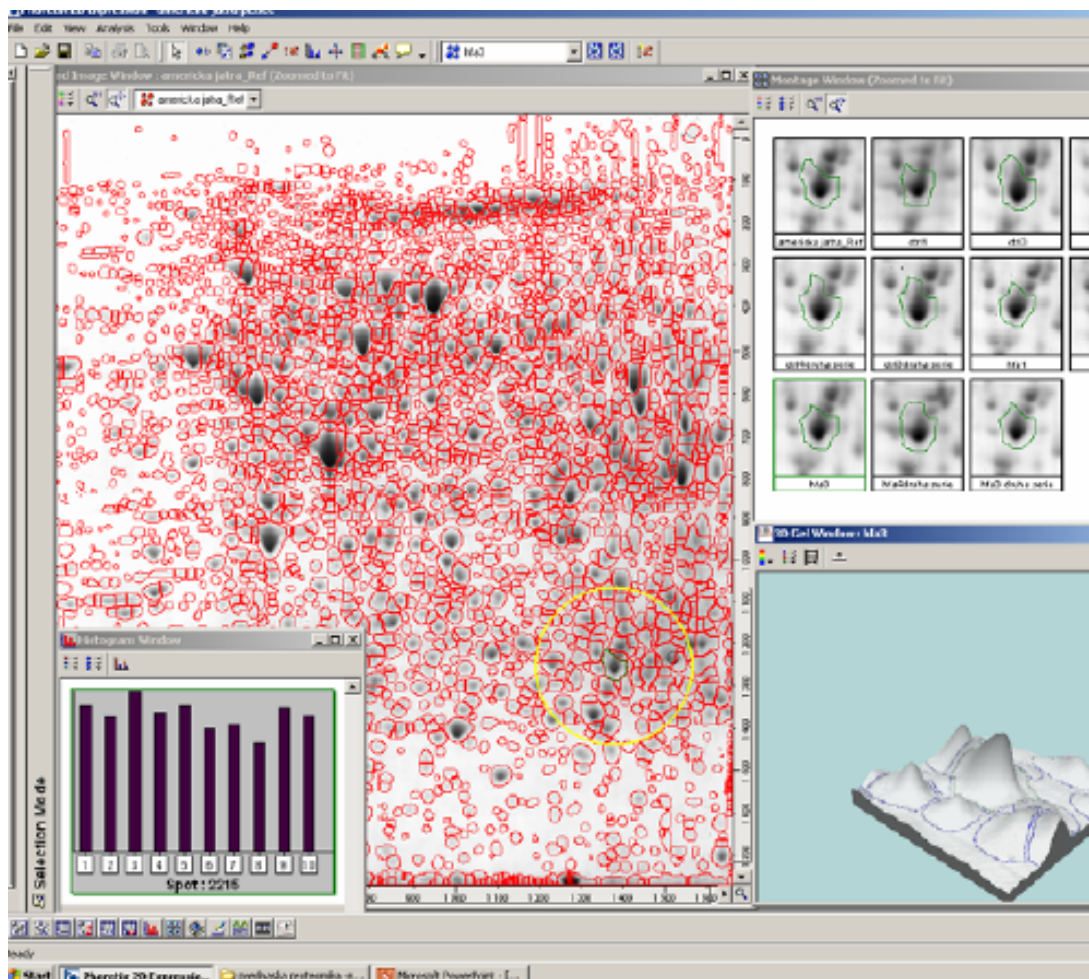


# Diferenciální dvojrozměrná gelová elektroforéza (2D-DIGE) VÝBĚR PROTEINŮ





# Softwarové vyhodnocení 2D-gelů



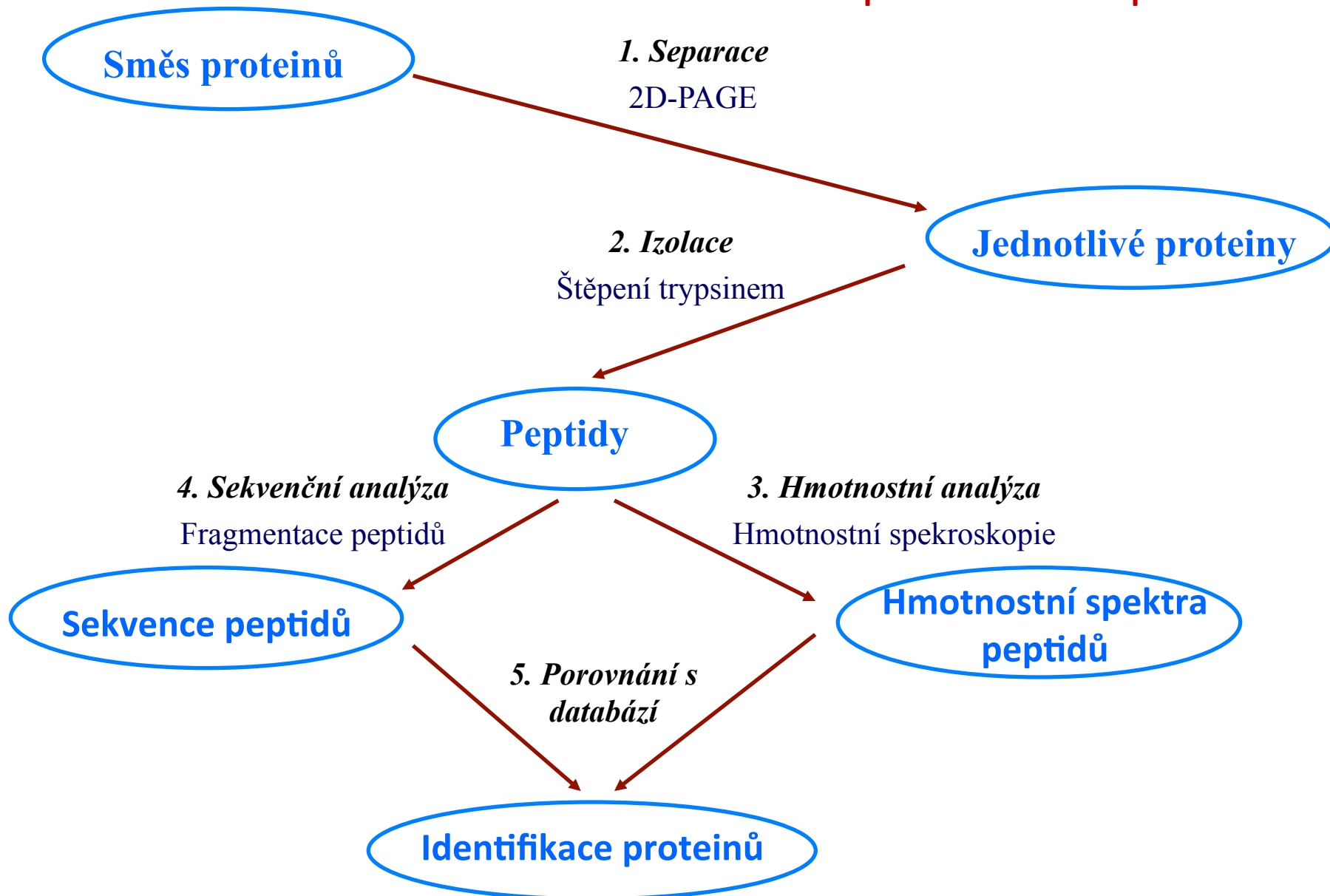
Komerční:

- PDQuest
- Phoretix
- Melanie

2D gels



"You've got one protein missing..."  
"No, you've got one extra protein!"

**Schéma – shrnutí klasického proteomického experimentu**

# Hmotnostní spektrometrie (Mass Spectrometry - MS)

## IDENTIFIKACE

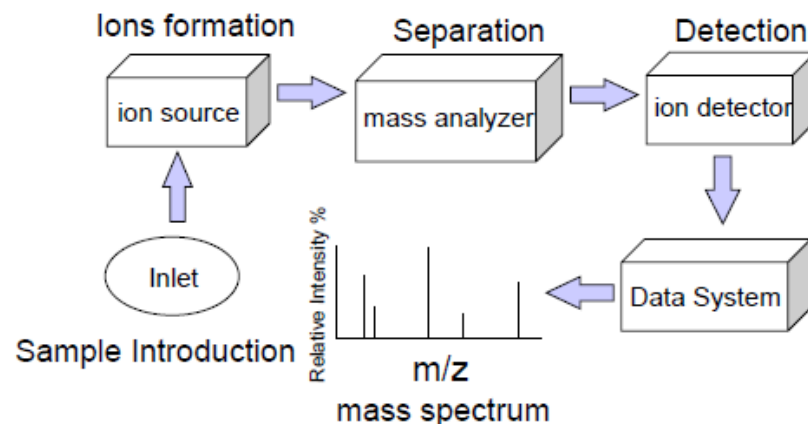
Separace látek podle rozdílů hmotnosti ( $m$ ) a náboje ( $z$ ) s využitím elektrického / magnetického pole.

Určovanou fyzikální veličinou je podíl hmoty a náboje ( $m/z$ ), při znalosti náboje umožňuje určit molekulovou hmotnost.

Výsledné hmotnostní spektrum → grafické znázornění četnosti iontů na hodnotě  $m/z$

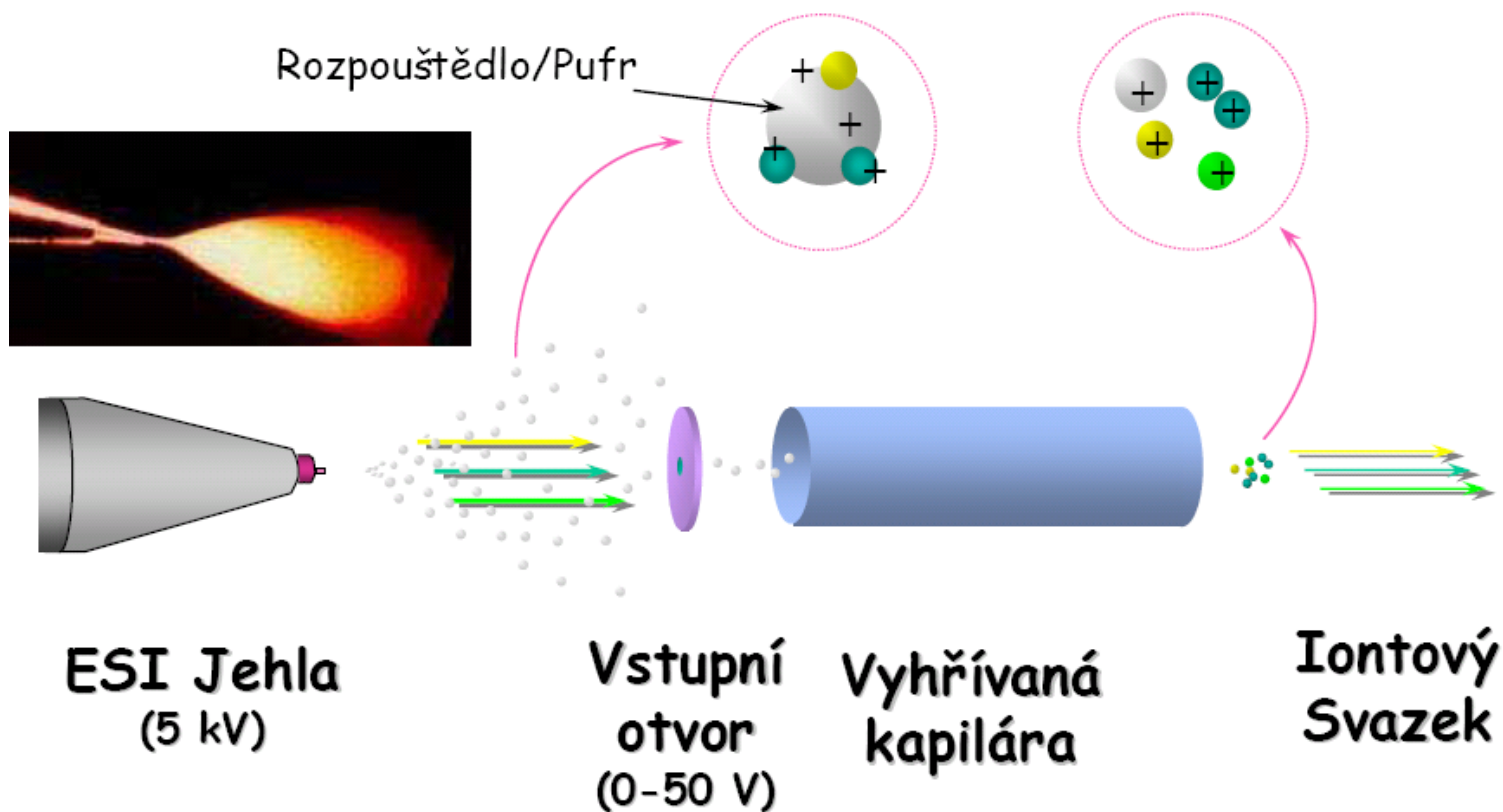
### IONTOVÉ ZDROJE

Potřeba vysušení a ionizace analyzovaných molekul - **měkké ionizační techniky (ESI, MALDI)** → měření makromolekulárních látek (**proteiny**, lipidové komplexy, polysacharidy)



## Techniky ionizace - elektropray (ESI)

jemná technika ionizace, nezpůsobuje fragmentaci analytu  
 vzorek rozpuštěn v těkavém organickém rozpouštědle  
 rozprašován pomocí nabité mikrostríkačky  
 odpařování rozpouštědla → molekulární ionty vstupují do spektrometru

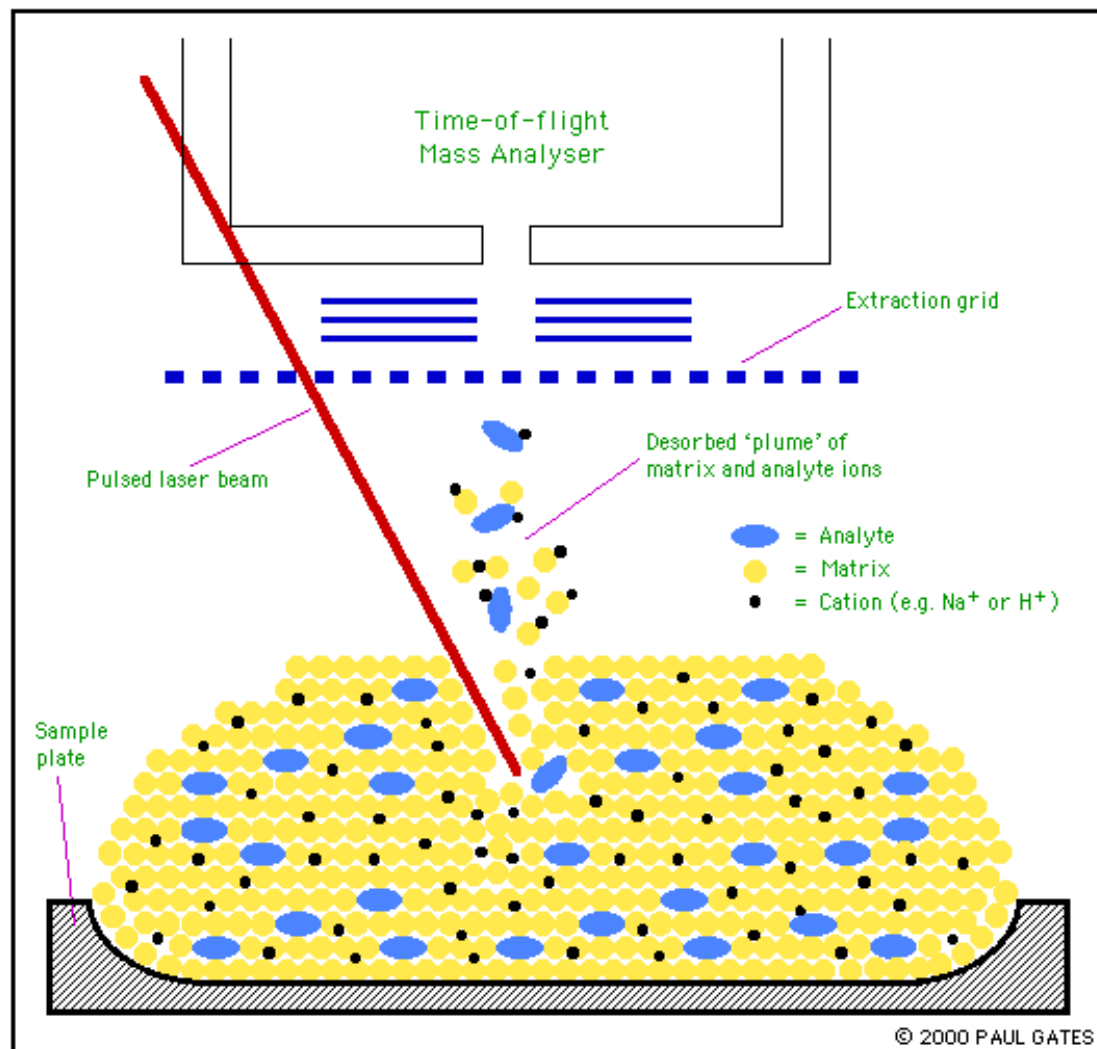




# Techniky ionizace–Matrix-Assisted Laser Desorption Ionization (MALDI)

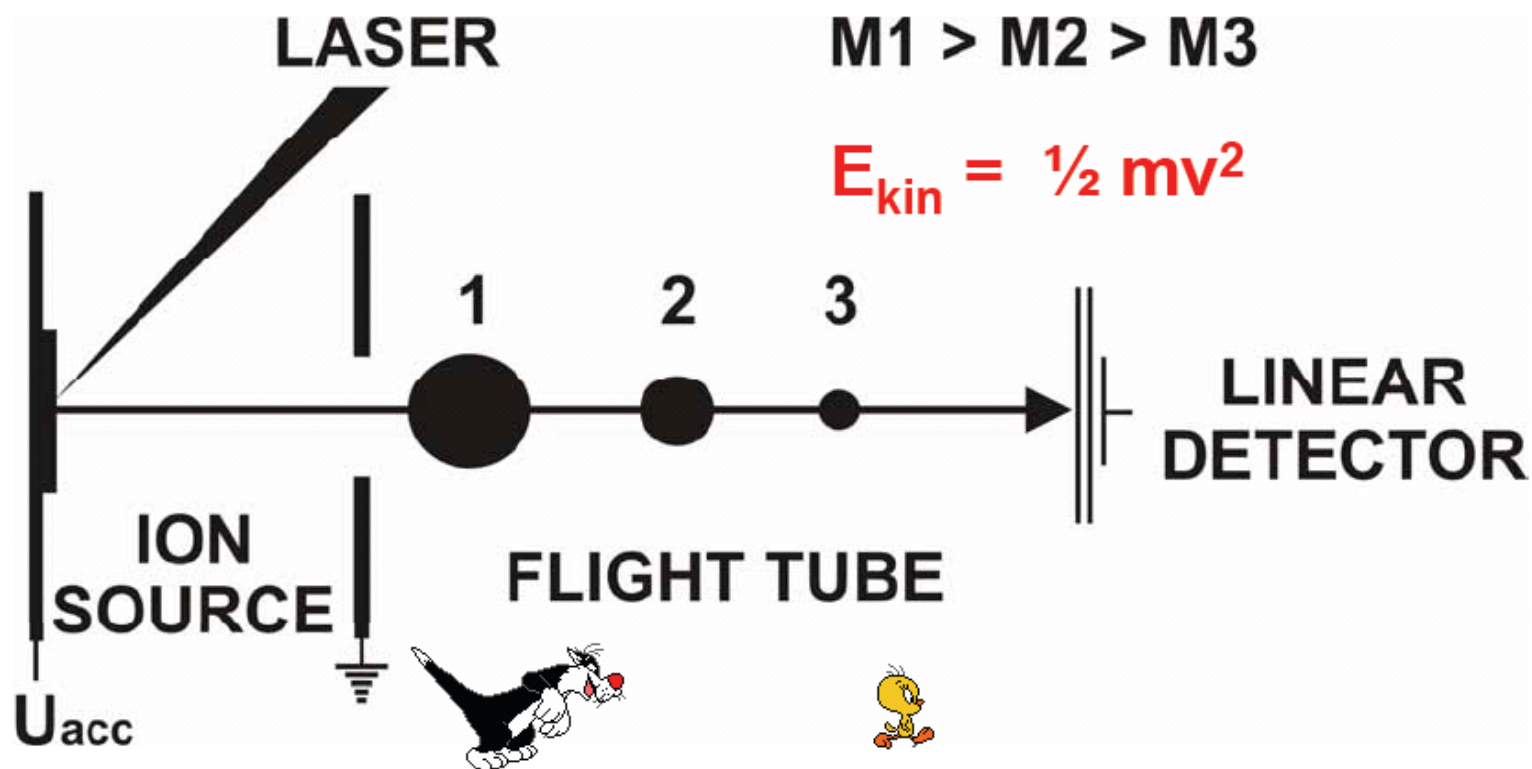
Smíchání vzorku s matricí  
 vysušení na kovové destičce  
**Matrice** (kyselina nikotinová,  
 dihydroxybenzoová)

- absorbuje energii laseru
- usnadňuje odpaření
- předává náboj analytu



## Typy analyzátorů MS – nejjednodušší Time of Flight (TOF)

Urychlení iontů v elektrickém poli o definovaných vlastnostech  $m/z$  lze určit z **doby letu** iontu trubicí analyzátoru



*Další typy analyzátorů – IT (ion trap-iontová past) Q3 (trojitý kvadrupól)*

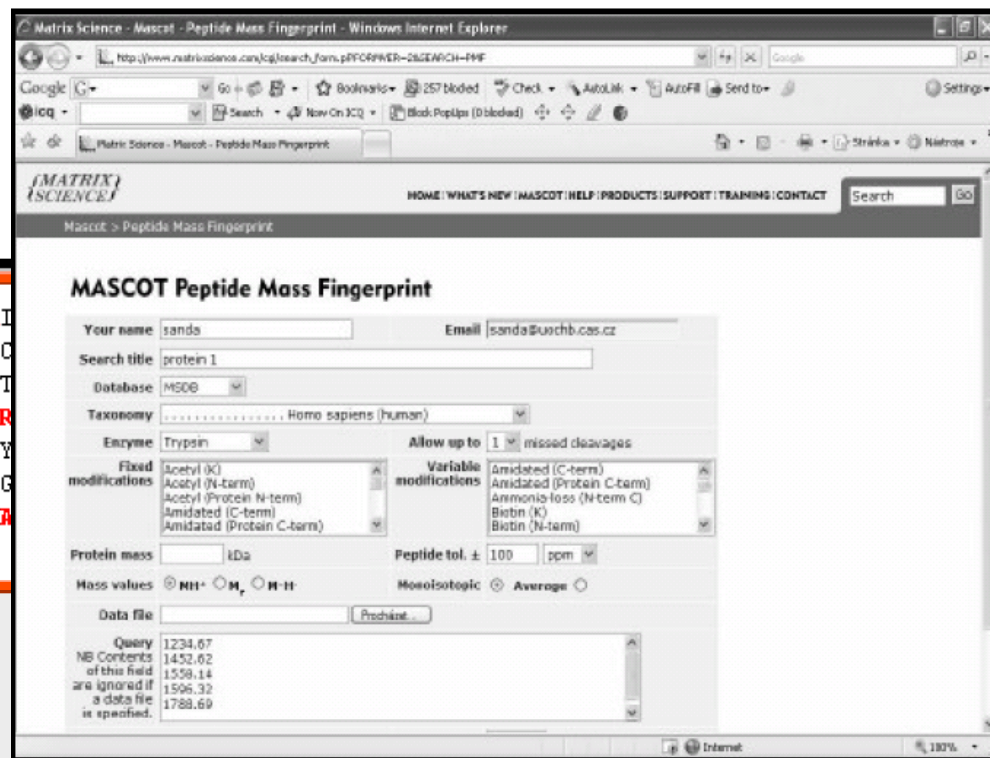
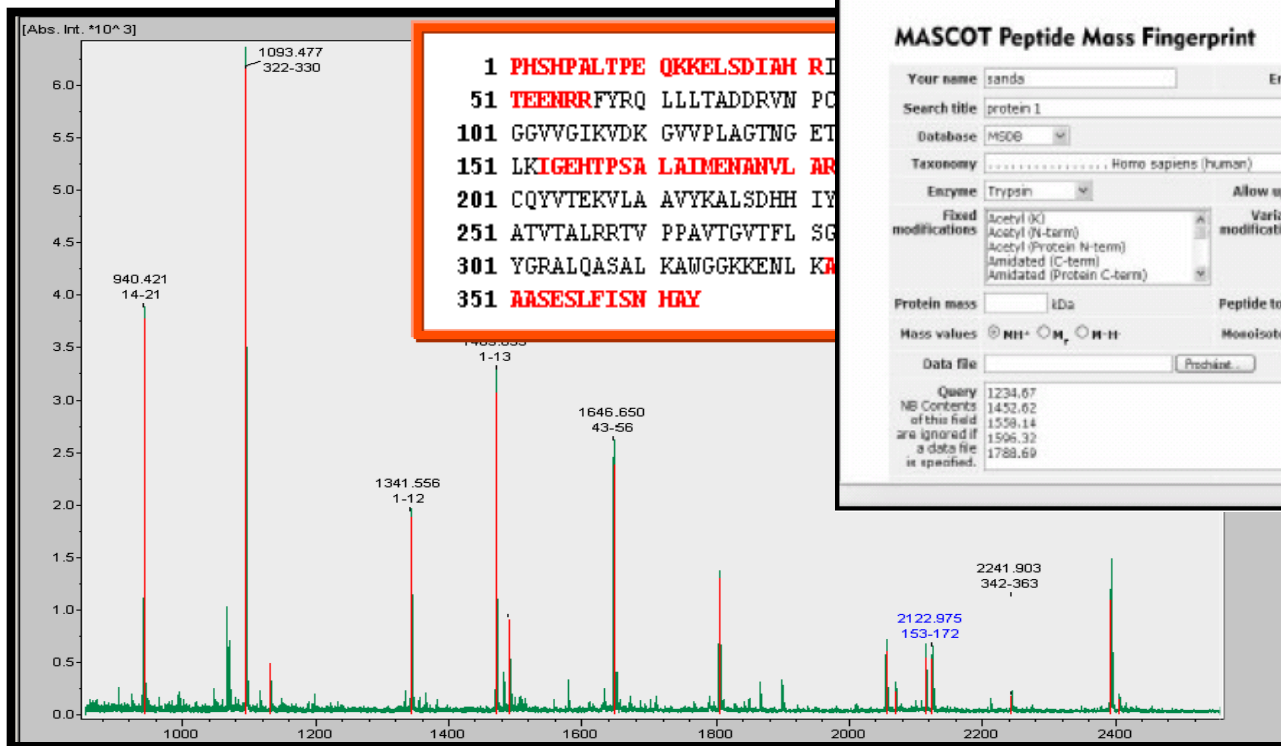
# Peptidové mapování (Peptide Mass Fingerprinting – PMF, MS a MS/MS) Identifikace proteinů z roztoku a gelu

Srovnání MS peptidového spektra s informacemi v databázích (AMK sekvence → teoretické štěpy)

## MALDI-TOF

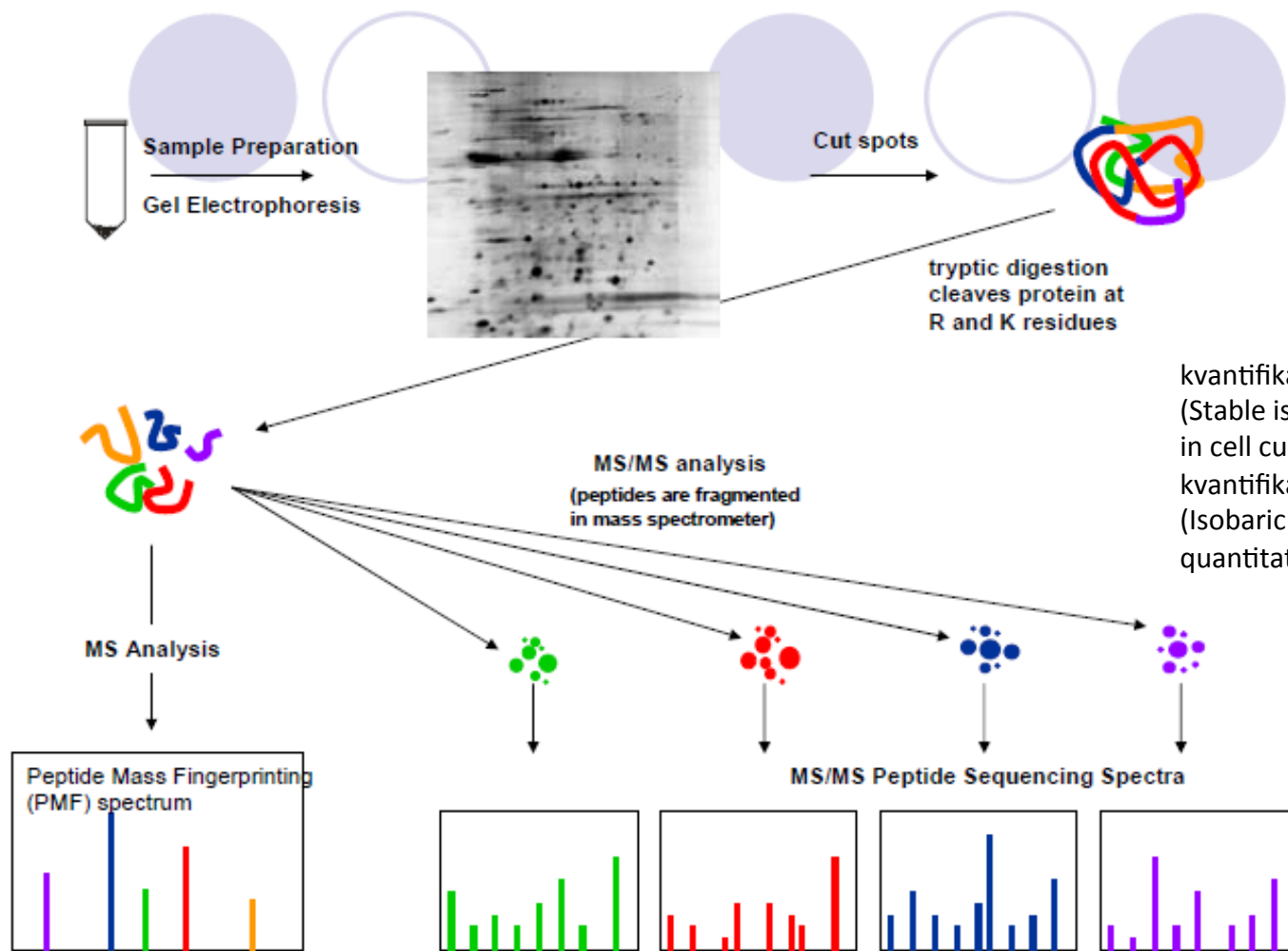
Interpretace MS dat

Programy : Mascot,  
ProteinProspector



Databáze:  
NCBI  
Swissprot

# Tandemová hmotnostní spektrometrie sekvenování peptidových fragmentů



kvantifikace peptidů v MS módu: **SILAC**  
(Stable isotope labeling by amino acids  
in cell culture)

kvantifikace peptidů v MS/MS módu: **iTRAQ**  
(Isobaric tags for relative and absolute  
quantitation)

# Tandemová hmotnostní spektrometrie sekvenování peptidových fragmentů

vícekroková separace na základě  $m/z$  (multiple MS)

-fragmentace určitého iontu: Post Source Decay (PSD), Collision Induced

Disociation (CID) - fragmentace při kontaktu s inertním plynem (dusík, helium) v kolizní cele

## Post Source Decay (PSD)

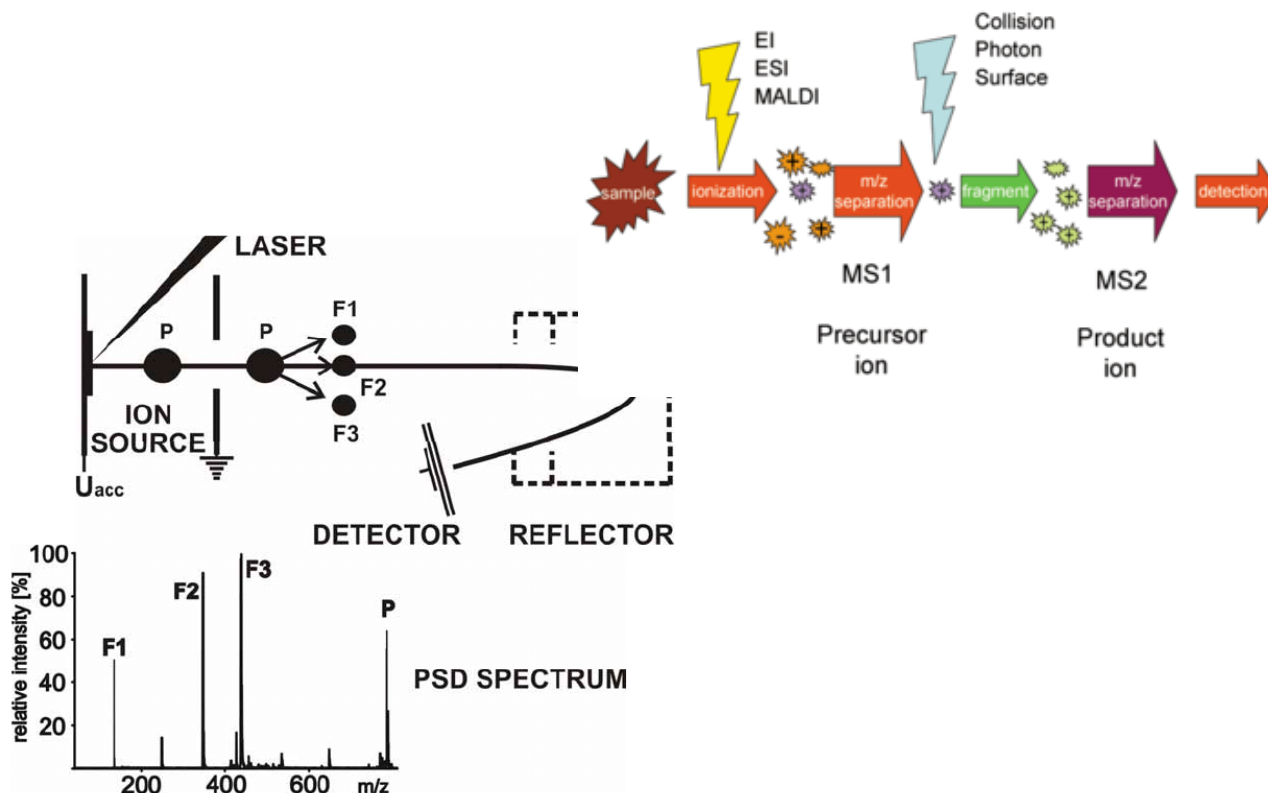
Při vyšší energii laseru →

fragmentace peptidu

(kolize mezi ionty analytu)



Peptide fragment  
Sequencing- metoda  
pro sekvenování  
peptidů



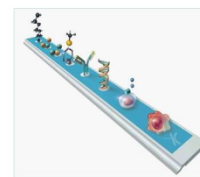
# ProteinChips and SELDI (Surface Enhanced Laser Desorption/Ionization) Kvantifikace proteinů

byla definována na počátku devadesátých let

- 1) je rozšířením MALDI-TOF-MS
- 2) metoda desorpce a ionizace netěkavých látek
- 3) měření proteinů v lineárním módu aktivní účast povrchu čipu při extrakci, prezentaci a strukturní modifikaci daného vzorku



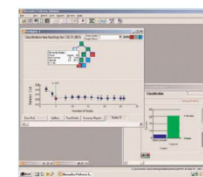
*The ProteinChip® Company*



**ProteinChip Arrays**  
» Separation



**ProteinChip Reader**  
» Detection



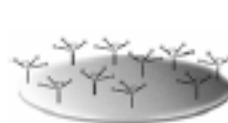
**ProteinChip Software**  
» Analysis



**hydrophobic**



**anionic**



**cationic**



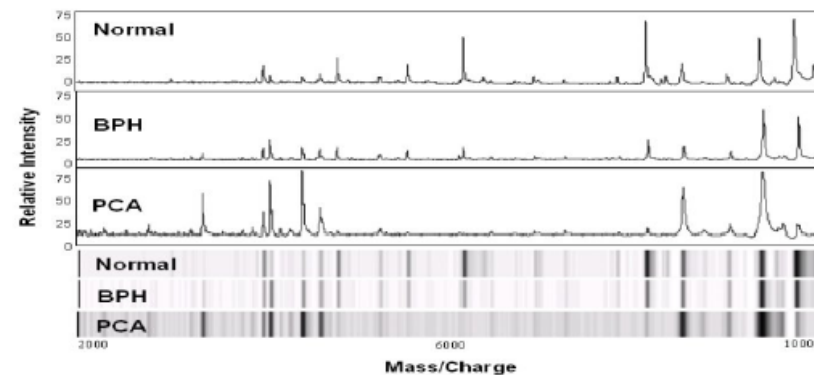
**IMAC**



**normal phase**

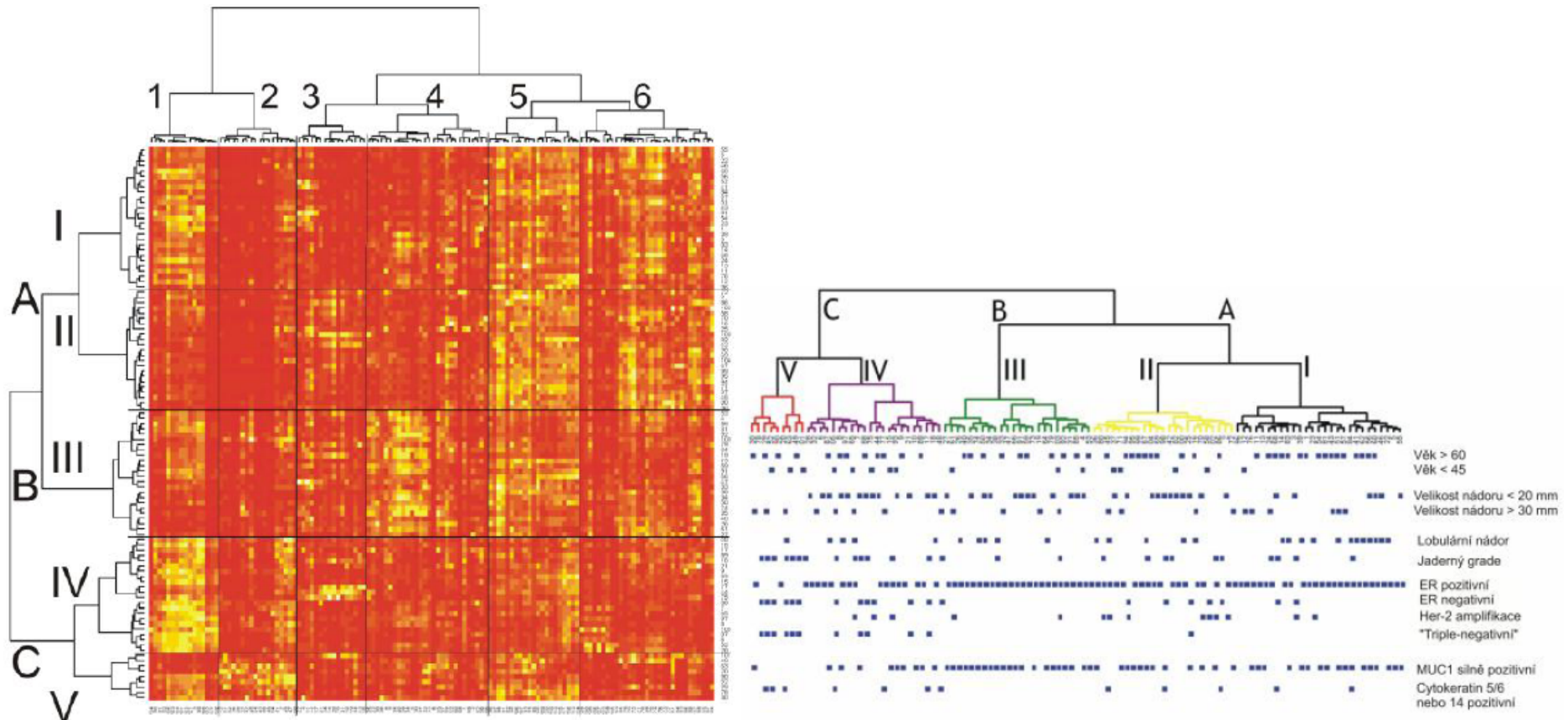
Rozdíly mezi metodami MALDI a SELDI

- 1) v prvotním zpracování samotného vzorku (inertní versus aktivní povrchy)
- 2) softwarové vyhodnocení získaných dat (hledání biomarkerů)





## Proteinové profily získané metodou SELDI umožňují rozdělit pacientky s karcinomem prsu do skupin analogicky jako u DNA čipů

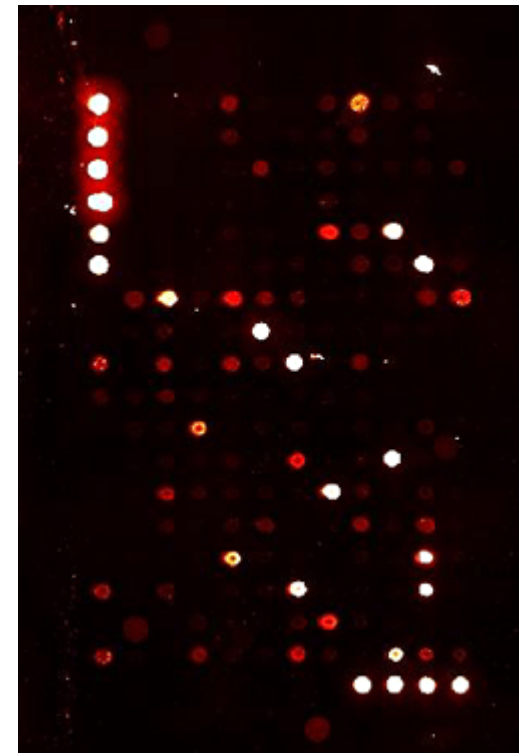
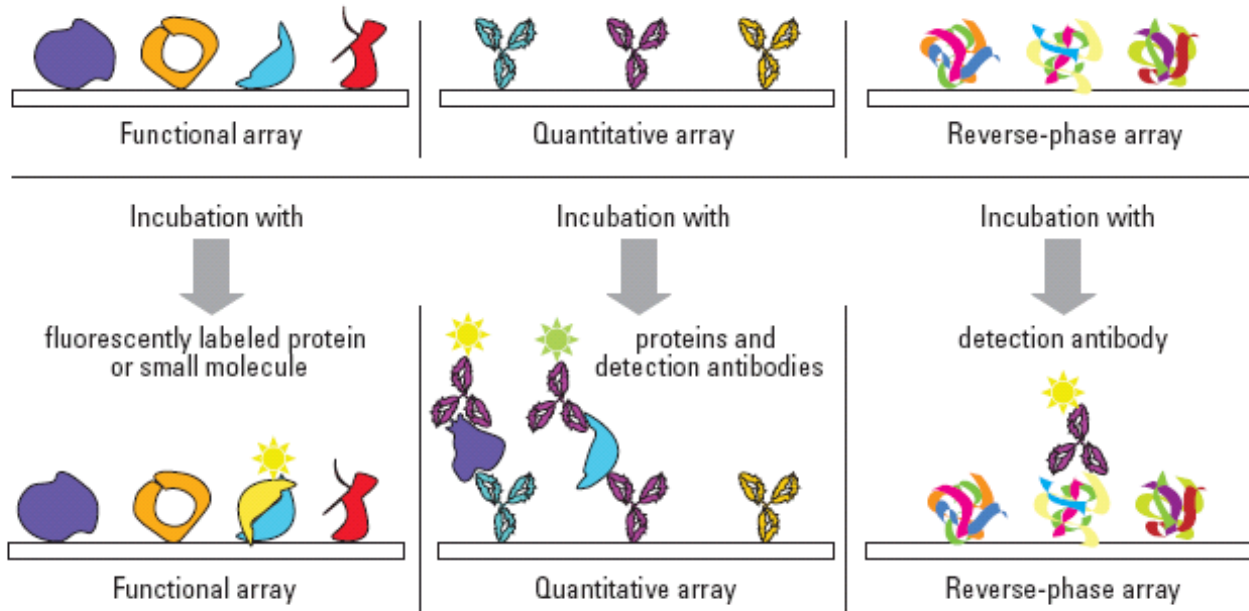


Brožková et al., 2008



## Proteinové arrays

- **Protilátkové microarrays** : kvantitativní array vazba proteinu na imobilizovanou protilátku → detekce sendvičovou metodou (fluorescenčně značená protilátka)
- **Reverzní array** : spotované lyzáty probované vždy jednou protilátkou
- Funkční array** – spotované purifikované proteiny probované fluorescenčně značenou látkou (protein, DNA, jiné biologicky aktivní molekuly)



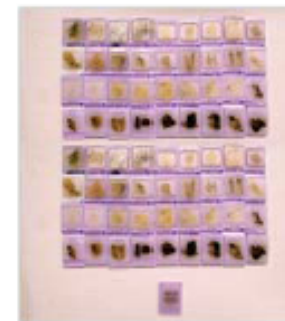
# Tkáňové arrays (Tissue MicroArrays - TMA)



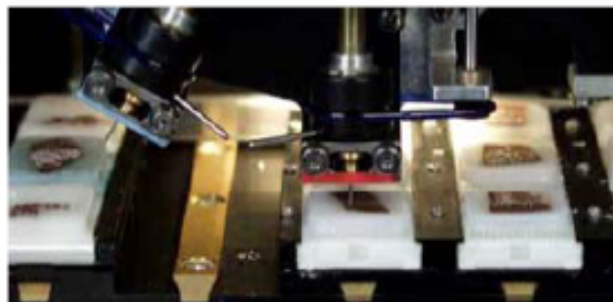
Collection of tissues



Annotation



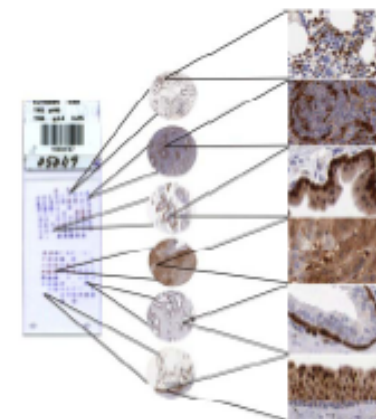
TMA design



TMA construction



TMA sectioning



Analysis

## Take home

Komparativní genomová hybridizace (CGH)

Genomová array CGH (aCGH)

SNP čipy

ChIP-on-chip technologie

Technologie čipů k detekci methylace CpG oblastí

Technologie exonových čipů

Proteomika

Obecné schéma klasického proteomického experimentu

Dvozměrná gelová elektroforéza (2D-ELFO)

Hmotnostní spektrometrie (Mass Spectrometry – MS, ISE, MALDI, SELDI, TOF, tandemová MS)

Proteinové arrays

Tkáňové arrays



## Náplň příští přednášky

Molekulární epidemiologie – definice a vymezení oboru, identifikace molekulárních rizikových faktorů vzniku a rozvoje onemocnění, analýza vztahu molekulárních faktorů a vlivů prostředí na rozvoj nádorového onemocnění, význam molekulární epidemiologie u karcinomu plic a kolorektálního karcinomu

Molekulární farmakologie I – cílená léčba – vývoj nových léčiv = identifikace nových molekulárních cílů, vysokovýkonný screening, tkáňové kultury, transgenní zvířecí modely, poměr rizik a prospěchu, ekonomická a etická hlediska při výběru identifikovaných cílů a vývoji nových léčiv

## Dotazy?

