

**Interní hematologická klinika LF MU a FN Brno  
Centrum molekulární biologie a genové terapie  
CEITEC MU**

# Moderní metody analýzy genomu

Jitka Malčíková

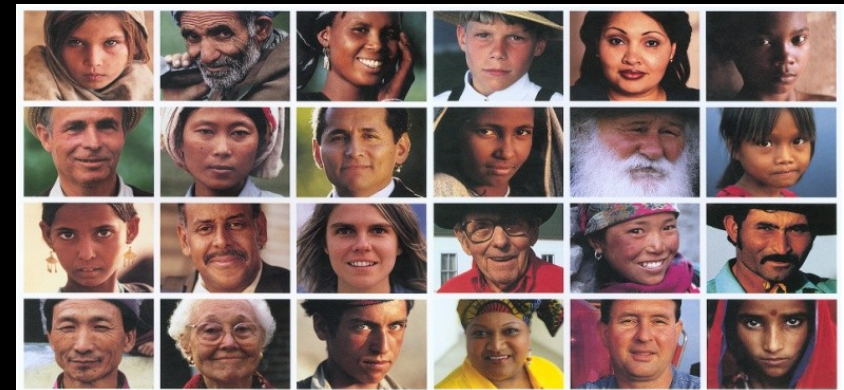
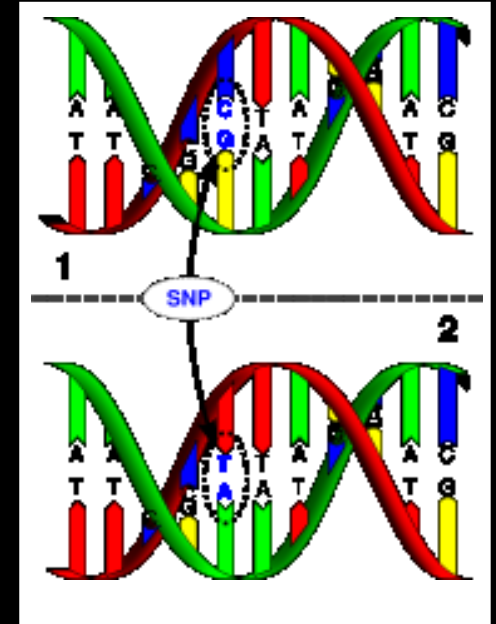
20.2.2013

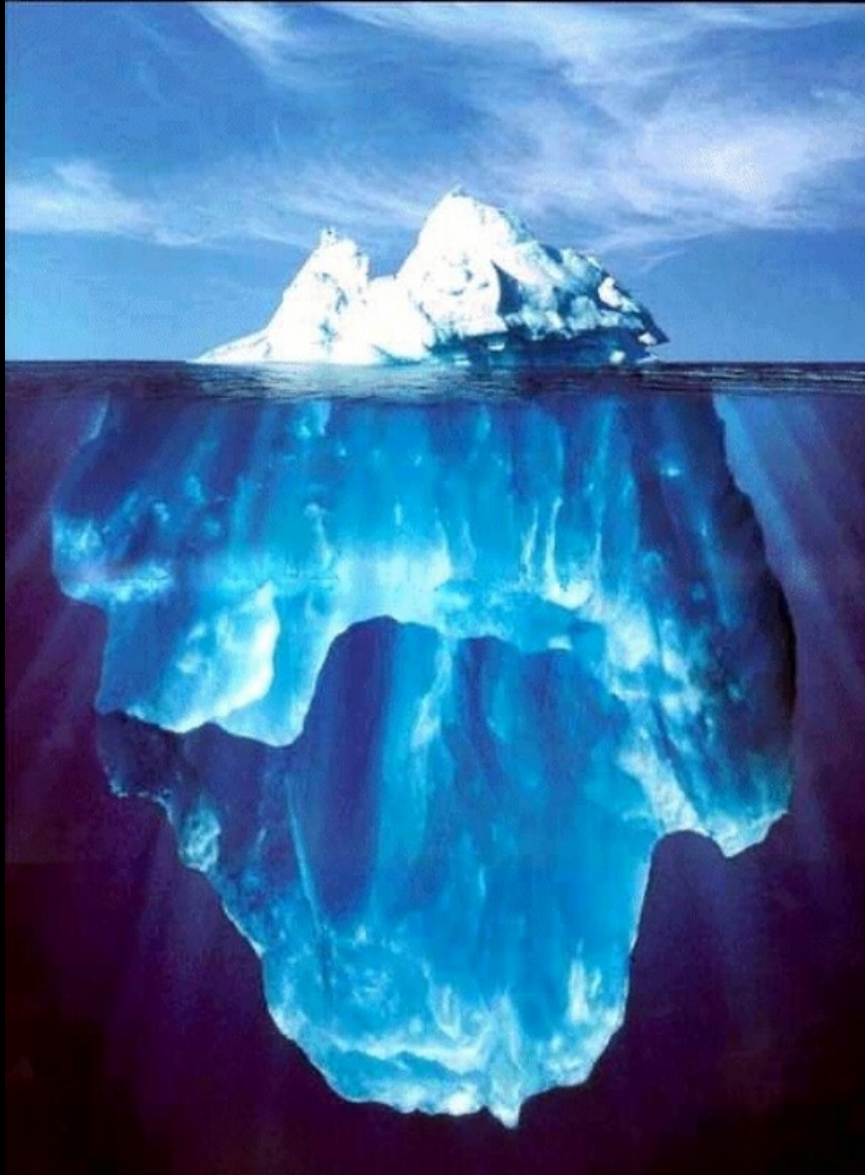
# Syllabus

- Úvod (Malčíková)
- Microarrays (Malčíková)
- Metody analýzy miRNA a dalších nekódujících RNA (Mráz)
- NGS (Tichý)
  - 454
  - SMRT
  - Sequencing by synthesis, sequencing by ligation – Illumina
  - IonTorrent
- Aplikace – Genomika, Transkriptomika, Epigenomika (Tichý)
- Analýza interakcí proteinů se specifickými sekvencemi genomu (Štros)
- Další technologie – Fluidigm, Sequenom, Luminex, Nanostring (Tichý)

# Variabilita genomu

- Jednouklettidové polymorfismy  
(Single Nucleotide Polymorphisms - SNP)
  - polymorfismus vs mutace
- Změny v počtu kopií genetického materiálu  
(Copy Number Variations – CNV)
- Velikost repetitivních sekvencí v genomu
- Epigenetické změny – imprinting





## Genom ( $3.2 \times 10^9$ bp, 20-25 tis genů)

- Transkripce
- Posttranskripční modifikace
- Alternativní sestřih

## Transkriptom

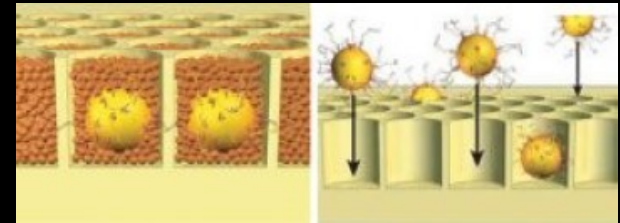
- Translace
- Posttranslační modifikace
- Alternativní konformace

## Proteom (miliony proteinů)

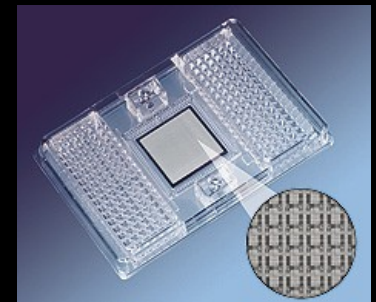
- Dynamický systém – odráží momentální stav buňky
- Hladina proteinu často nekoreluje s hladinou mRNA

# Moderní metody – analýza „omů“

- Čipové technologie
- Sekvenování nové generace



- Fluidigm – mikrofluidní kvantitativní PCR



- Hmotnostní spektrometrie

- Analýza proteomu, ale i genové exprese, SNP a metylací



# Moderní metody



„High-throughput“ „Large-scale“ „Ultra-deep“

- Snaha o získání co nejvíce informací
- Výsledkem je ohromné množství dat
- Význam biostatistiky



# Sekvenování genomů

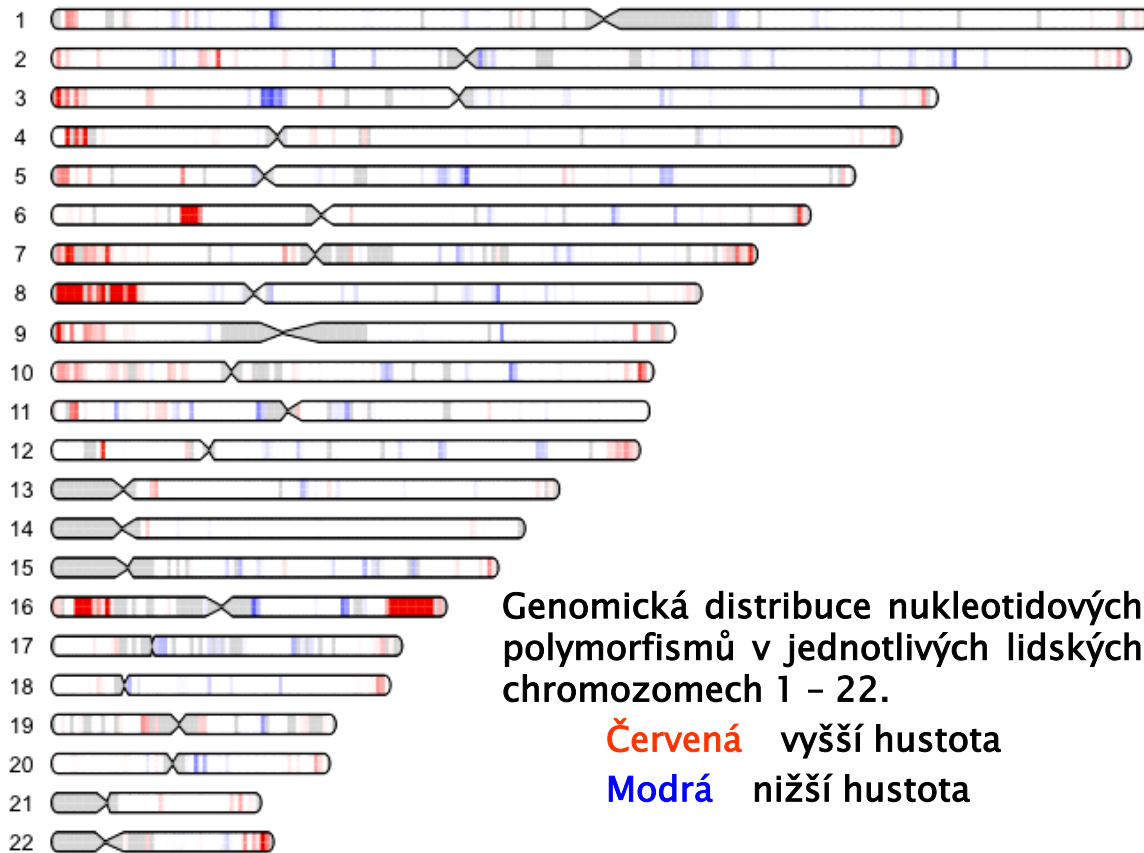
- První individuální genom (kompletní diploidní sekvence konkrétního člověka) – Institut J. C. Ventera
  - Použitá technologie: Sangerovo kapilární sekvenování
  - Výsledky: identifikováno 2,81 miliardy nukleotidů (7,5 nás. pokrytí), 4,1 milionu variant (44 % změn heterozygotních), z toho **3,2 mil. SNP**, 292 tis. malých inzercí/delecí

*(Levy, S., et al. PloS Biol 2007)*

- Druhý individuální genom - Baylor College, Texas
  - Použitá technologie: sekvenování nové generace
  - Výsledky: identifikováno **3,3 mil. SNP** (7,4 nás. pokrytí), z toho **10 tis. ovlivňuje záměnu aminokyseliny v proteinovém produktu genů**

*(Wheeler, D.A., et al. Nature, 2008)*

# Projekt 1000 genomů - „A deep catalog of human genetic variations“



Již bylo v lidském genomu  
Identifikováno:

15 milionů SNP

1 milion krátkých inzercí/delecí

20 tisíc strukturních variací

A map of human genome variation from population-scale sequencing. 1000 Genomes Project Consortium, Durbin RM, Abecasis GR, Altshuler DL, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA., *Nature*. 2010 Oct 28;467(7319):1061-73. (Sanger Institute, 179 human genomes)

Diversity of human copy number variation and multicopy genes. Sudmant PH, Kitzman JO, Antonacci F, Alkan C, Malig M, Tsalenko A, Sampas N, Bruhn L, Shendure J; 1000 Genomes Project, Eichler EE, Abecasis GR, Altshuler DL, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA., *Science*. 2010 Oct 29;330(6004):641-6. (University of Washington, Seattle, 159 human genomes)



# Úskalí

- Interpretace
- Komerční „direct-to-consumer“ (DTC) genetic testing



# Microarrays



# Čipové technologie

- 1995 - Mark Schena

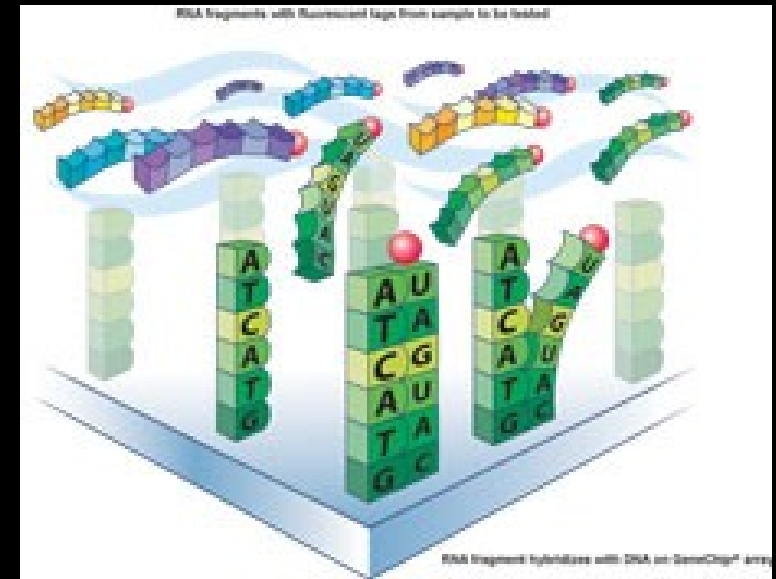
„Quantitative monitoring of gene expression patterns with a complementary DNA microarray“, Science (1995); 270(5235): 467-70.

- Analýza exprese 45 genů Arabidopsis

- 1996 - Affymetrix GeneChip

# Princip čipových technologií

- Princip – sondy vázané na nosný podklad



- Různé dělení podle
  - Aplikace (genotypizace; exprese RNA, proteinů; epigenetika...)
  - Technologie výroby (*in situ* syntéza; deponování)
  - Typu sond (umělé chromozómy – BAC, PAC; cDNA; oligonukleotidy)
  - Značení (fluorescence – jedno/více-kanálové; autoradiografie, chemiluminiscence)

# Nosný podklad

## ■ Sklo

- mikroskopické sklíčko (25 x 75 mm)
- speciální formáty
- různé úpravy povrchu – aminosilan, epoxy atd.

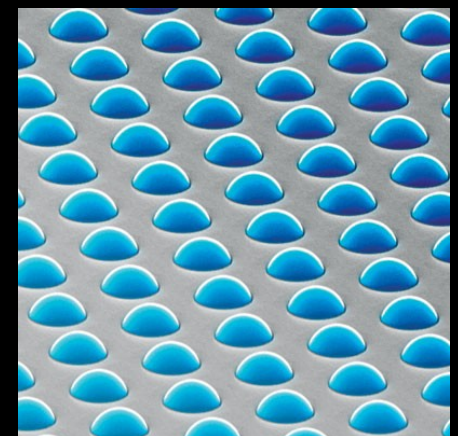
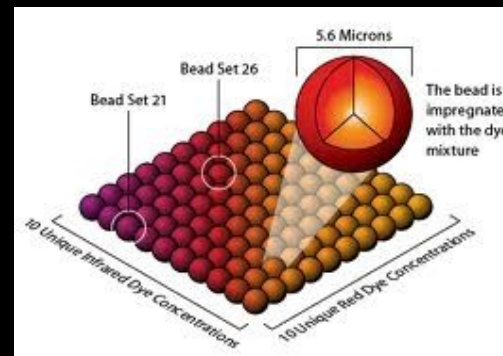


## ■ Nylonová nebo nitrocelulózová membrána

## ■ Plast, gel

## ■ Mikrokuličky

- BeadArrays® (Illumina)
- xMAP (Luminex)

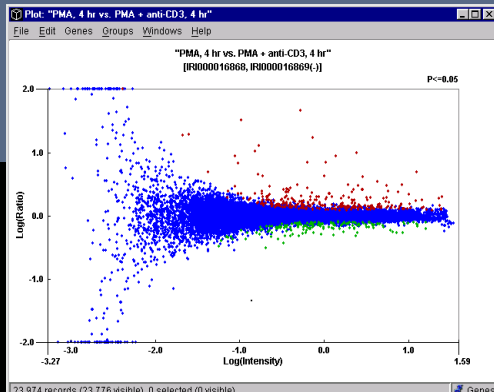




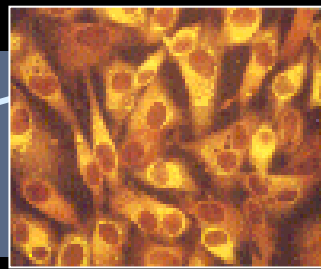
# Aplikace

- Analýza genomových aberací, genotypizace
  - CGH, SNP čipy
  - Resekvenační čipy
- Epigenomika
  - Mapování vazby proteinů na DNA (ChIP-on-chip)
  - Mapování metylované DNA (MeDIP-chip)
- Exprese
  - mRNA, miRNA
  - Proteiny
  - Buňky, tkáně





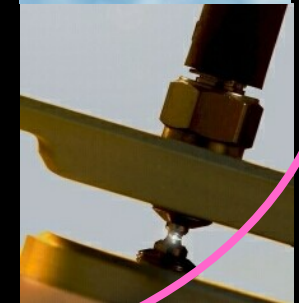
**Analýza dat,  
statistické  
zpracování**



**Buňky, tkáň...**



**Bioanalyser**



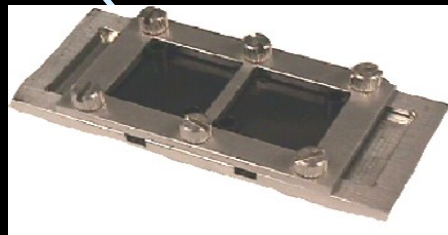
**Spektrofotometr  
NanoDrop**

**Izolace DNA,  
RNA,  
proteinů...**

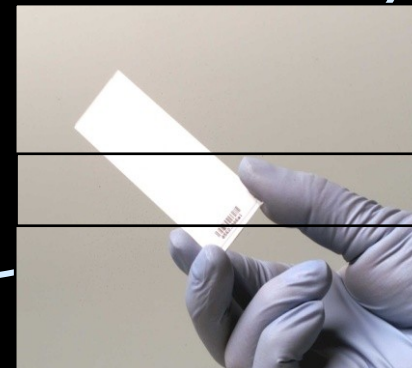
# Čipová analýza



**Skenování čipu**



**Hybridizace materiálu  
(RNA, proteinů...) s  
čipem**



**Příprava čipu**

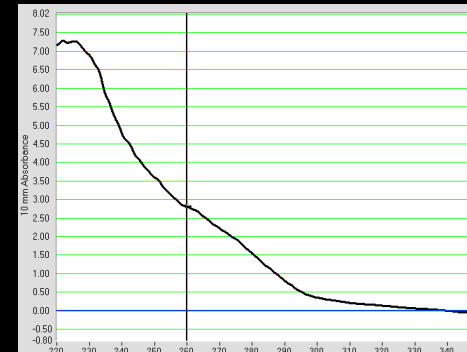
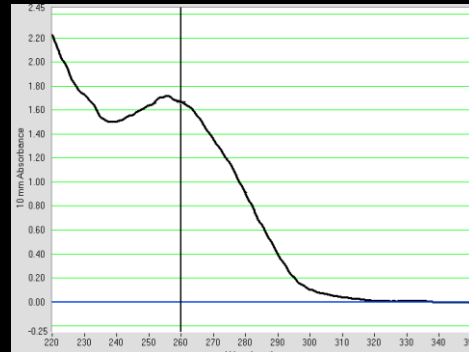
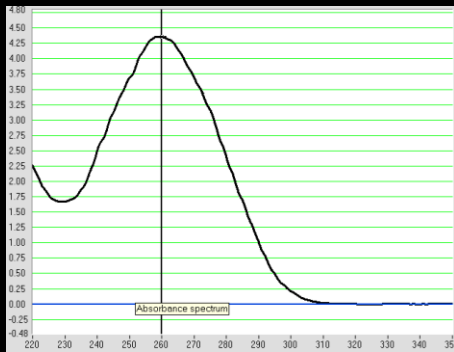


**Amplifikace a  
značení**

# Kontrola kvality

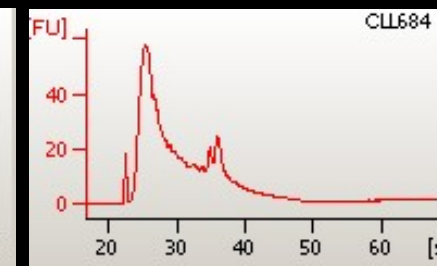
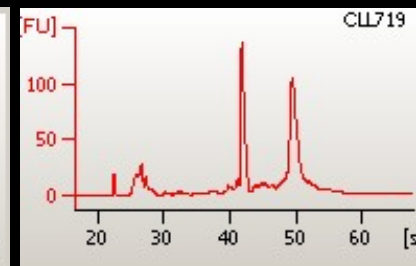
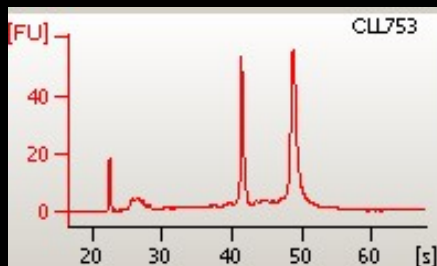
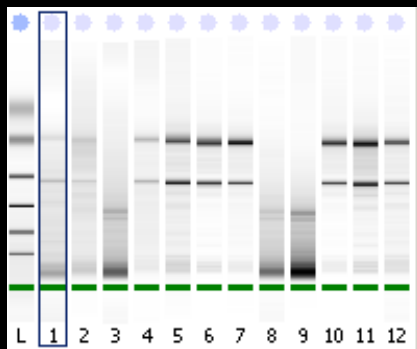
## ■ Nanodrop

- kvantifikace nukleových kyselin  $A_{260}$
- měření čistoty:  $A_{260/280}$  (kontaminace proteiny);  $A_{260/280}$  (kontaminace org. látkami)



## ■ Bioanalyser – „lab on a chip“

- Degradace RNA
- RIN – RNA integrity number – poměr 28S a 18S rRNA



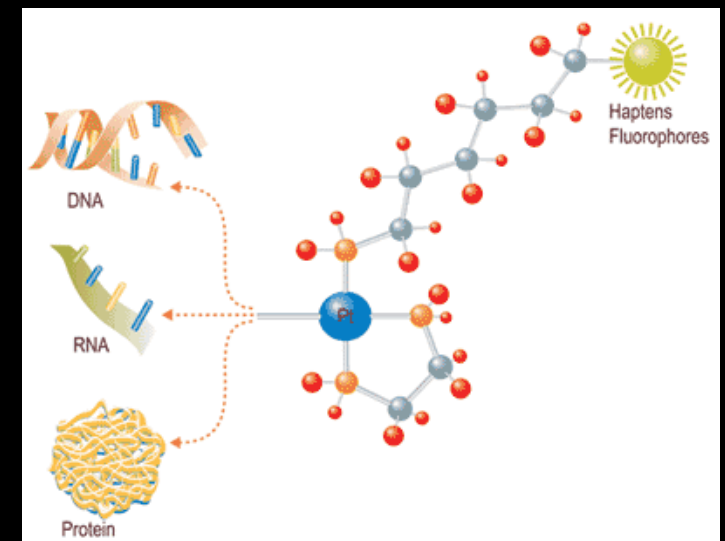
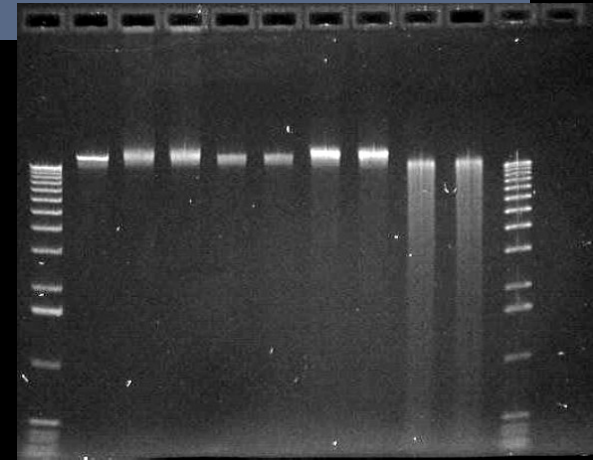


# Čipy pro analýzu genomu

- Výchozí materiál – genomová DNA
- Detekce nebalancovaných genomových změn
  - Amplifikace, delece – CNV – copy number variations
  - Nedetekuje balancované chromozomální translokace
  - Array CGH
- Detekce uniparentální dizomie, genotypizace
  - SNP čipy
- Detekce mutací, polymorfismů
  - Resekvenační čipy

# Postup

- Izolace DNA
- Kontrola kvality, stanovení kvantity
- Amplifikace – celogenomová, PCR
- Fragmentace – enzymaticky
- Značení - fluorescence
- Hybridizace, odmytí
- Skenování
- Analýza dat



# CGH - komparativní genomová hybridizace

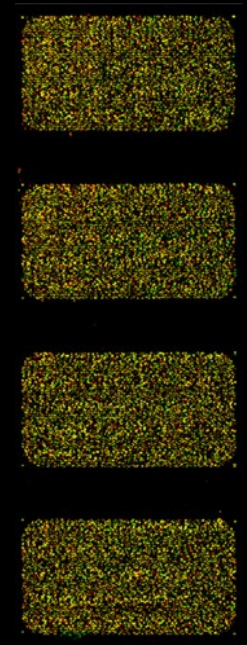
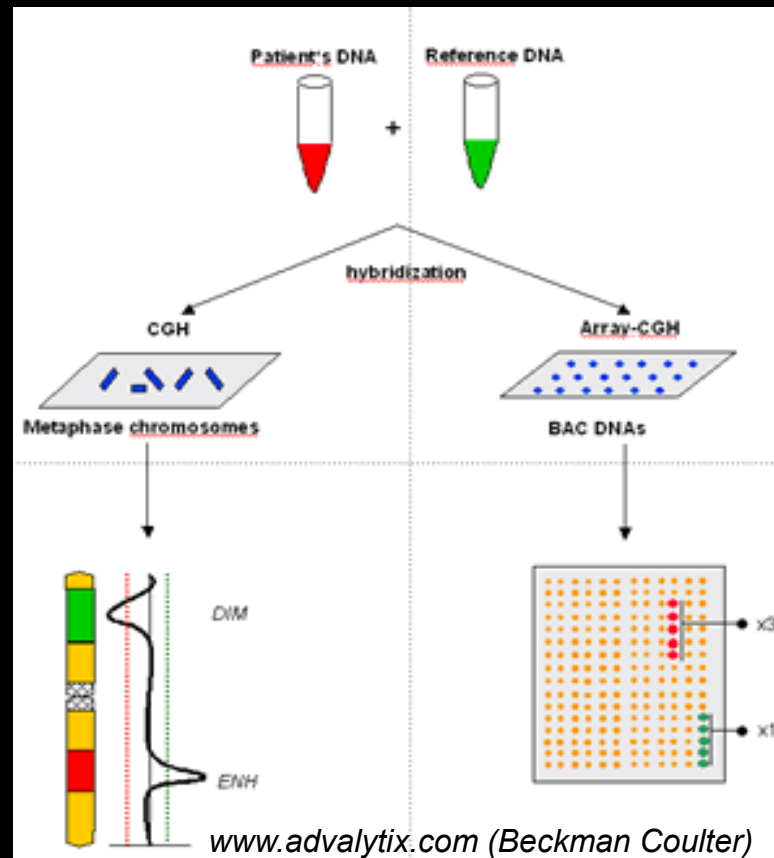
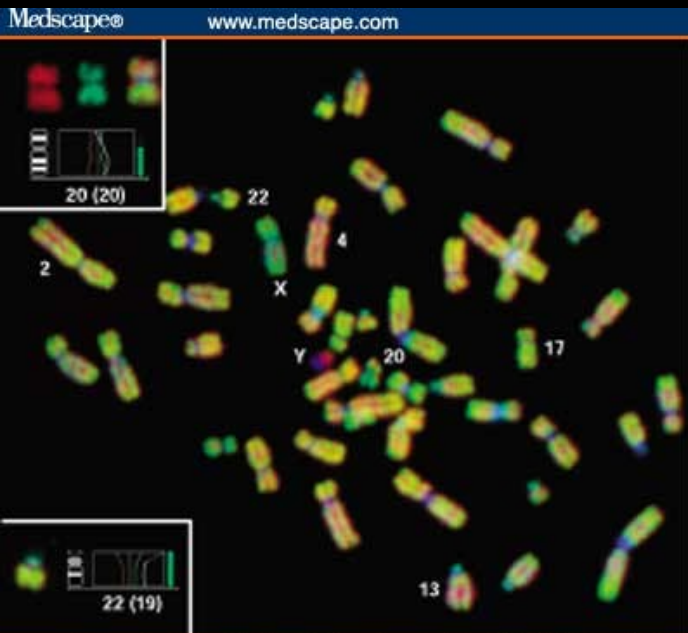
Kohybridizace značené **DNA vzorku** a **kontrolní DNA**

- na normální metafázní chromozomy

- klasická CGH, HR-CGH
- Metoda molekulární cytogenetiky

- na sondy navázané na sklíčku

- array CGH – čipová technologie



# Rozdělení aCGH podle typu sond

- Délka a hustota pokrytí DNA sond určuje rozlišení čipu
- Úseky genomové DNA vložené do vektorů
  - YAC (Yeast Artificial chromosome) - 200 -1000 kb
  - BAC (Bacterial Artificial chromosome) - 50–200 kb
  - PAC (Phage Artificial chromosome) - 75-200 kb
  - Kosmidy – 30-40 kb
  - Rozlišení ~ 1Mb
- cDNA - 1-2 kb
  - Pouze genové oblasti
  - Horší schopnost detekce jednokopiových změn
  - Rozlišení - 1-2 kb
- Oligonukleotidy - 25-85 bází
  - Rozlišení - pod 1 kb



# Platformy aCGH

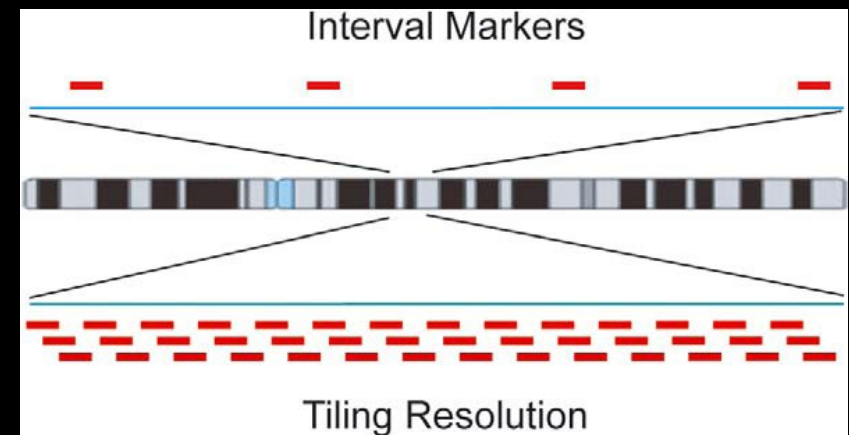
## ■ Cílené

- Analýza vybraných „hot spot“ oblastí spojených s konkrétním onemocněním

## ■ Chromozomové

## ■ Celogenomové

- Sondy rozmístěné
  - rovnoměrně po genomu
  - v určitých intervalech
  - Tiling arrays – sondy se překrývají – vysoké rozlišení

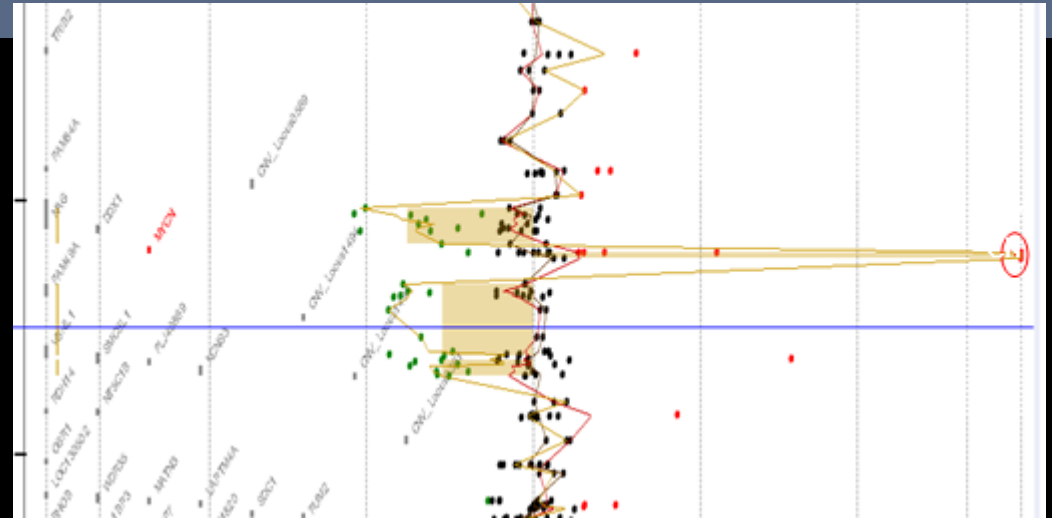
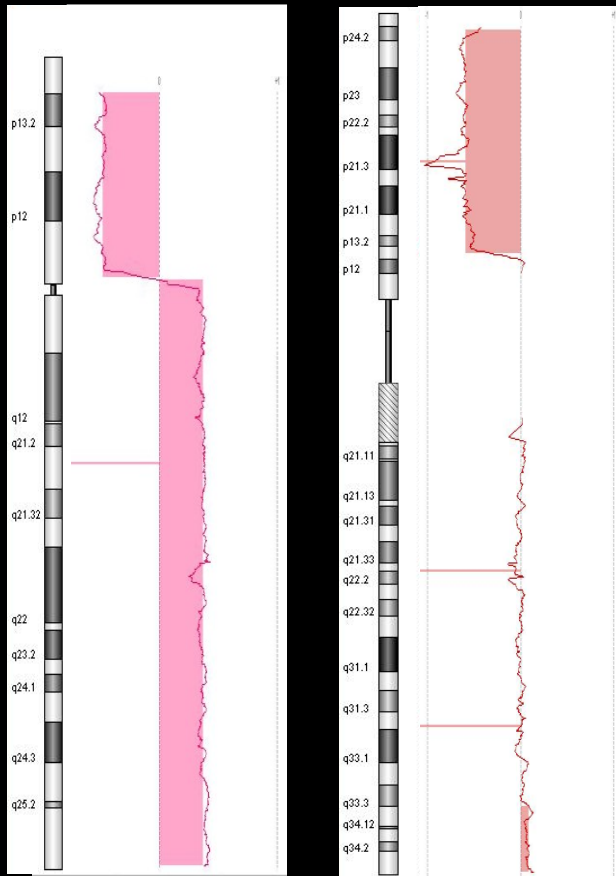


*Davies et al., 2005*

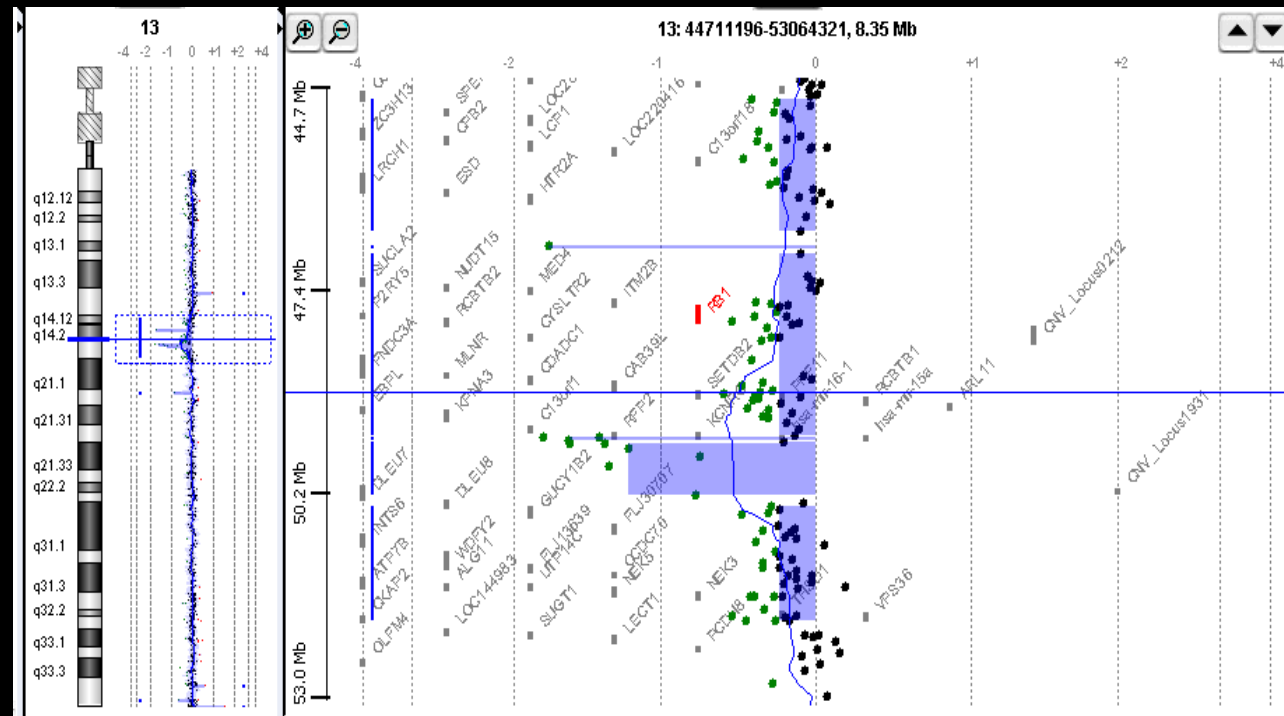
# Výsledky aCGH

Delece + amplifikace

izochromozom 17 delece 9p

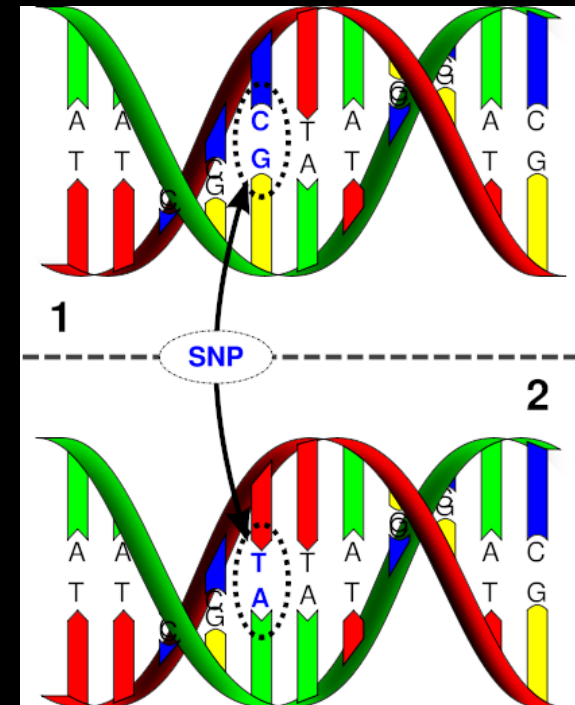


Krátká delece mono a bialecká



# SNP čipy

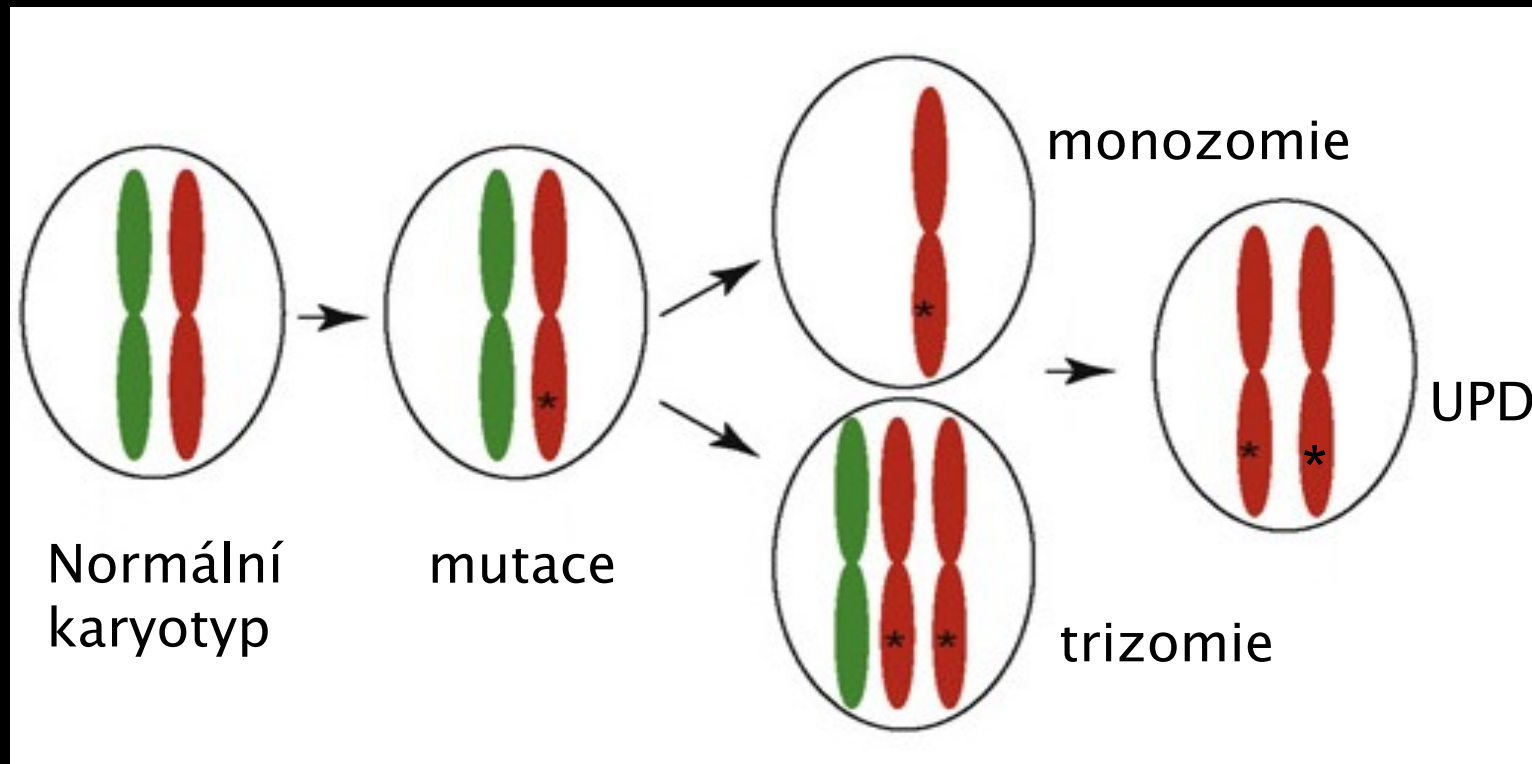
- Čipy umožňující rozlišit nejen CNV, ale i SNP
  - Sondy pro detekci CNV jako u CGH čipů + sondy speciálně navržené pro detekci SNP
- SNP – jednonukleotidové polymorfismy
- Určení alelového statusu, haplotypu
  - Genotypizace, asociační studie
- Detekce uniparentální dizomie



# Uniparentální dizomie (UPD)

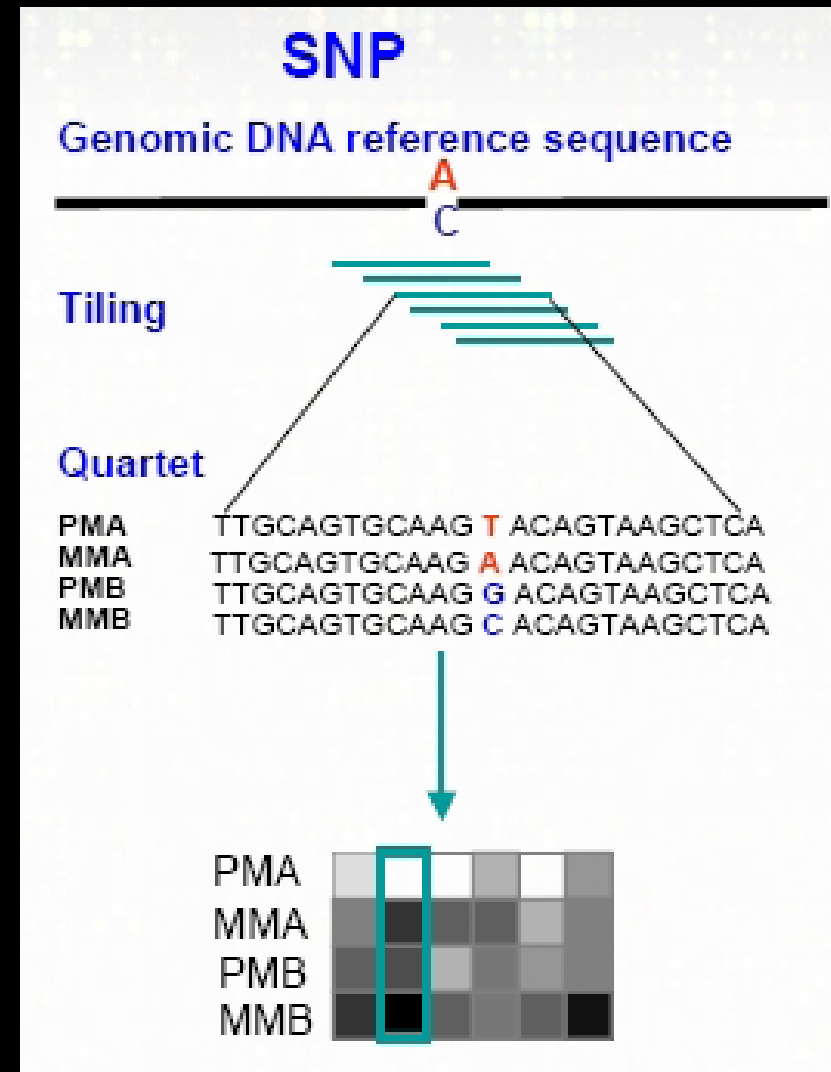
CNN LOH – copy number neutral loss of heterozygosity

- Vrozená, získaná
- Heterodisomie vs isodisomie
- Vyskytuje se u řady onemocnění



# Princip SNP čipů - Affymetrix

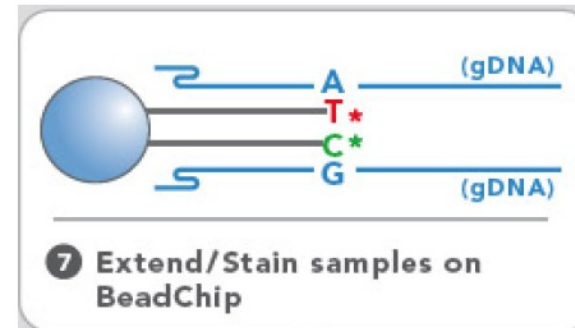
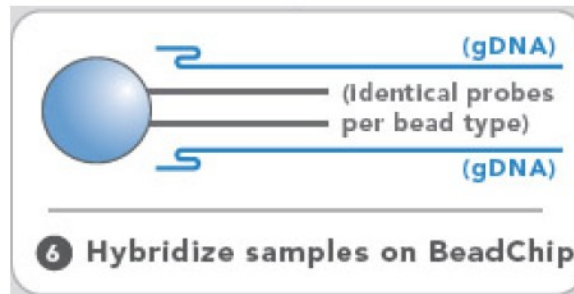
- 25b sondy, záměna 13. nukleotidu - největší vliv na sílu vazby při neshodě
- Jednobarevná detekce
  - na čip se hybridizuje pouze označená testovaná DNA



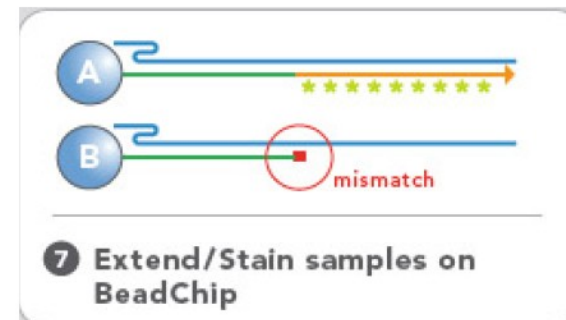
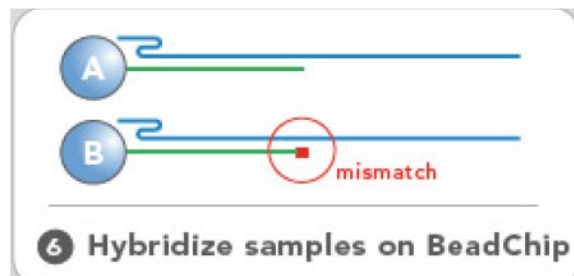
# Princip SNP čipů - Illumina

- 1x 50b, single base extension; pouze A/G, A/C, T/C, T/G (dvoubarevná metoda)
- 2X 50b, allele specific extension

Infinium II



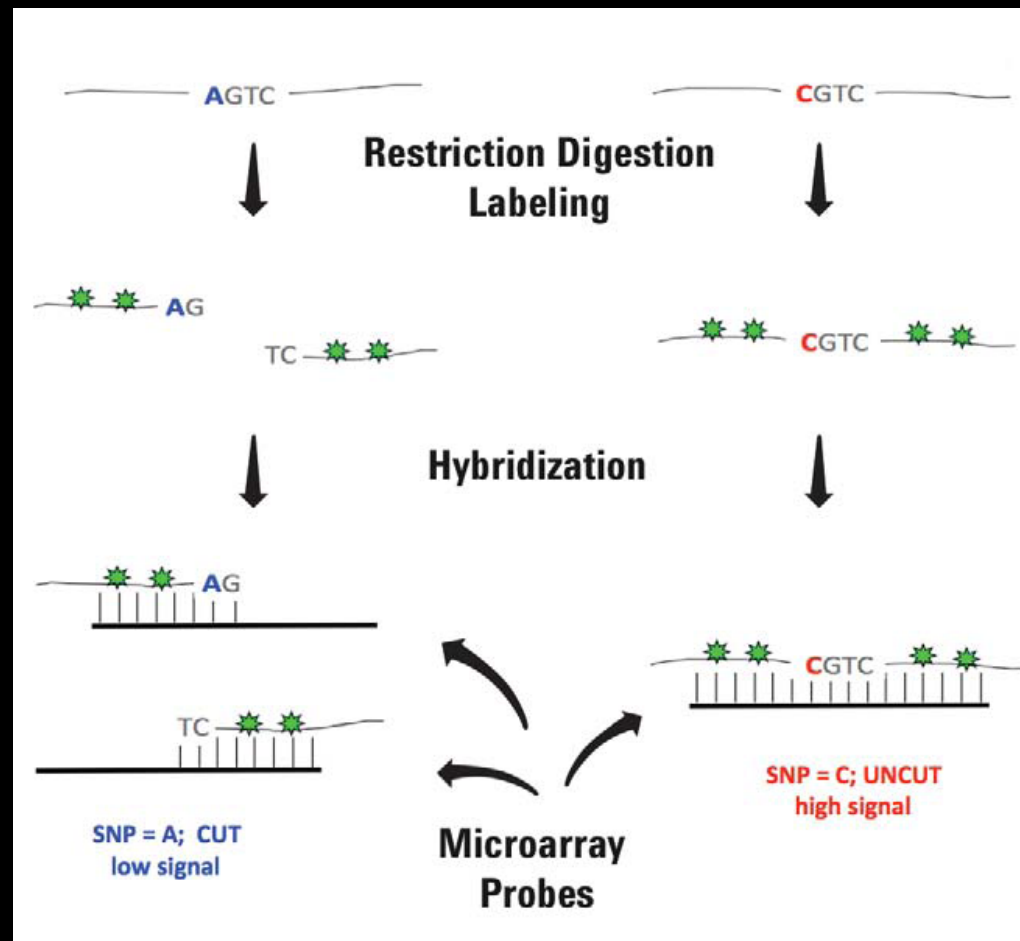
Infinium I





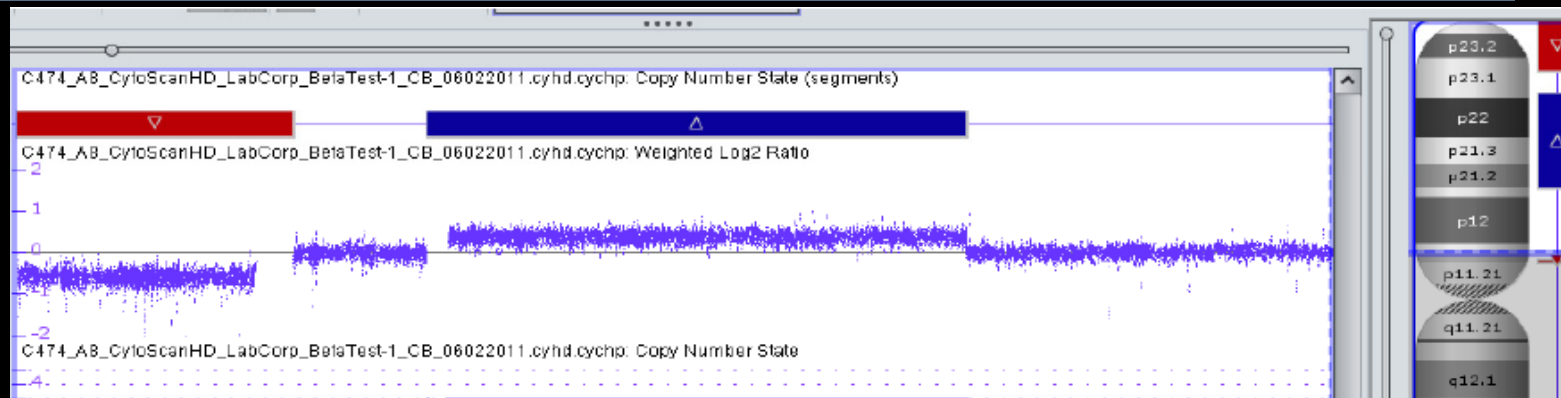
# Princip SNP čipů - Agilent

- 60b, SNPs pouze v místech rozpoznávaných restričními endonukleázami

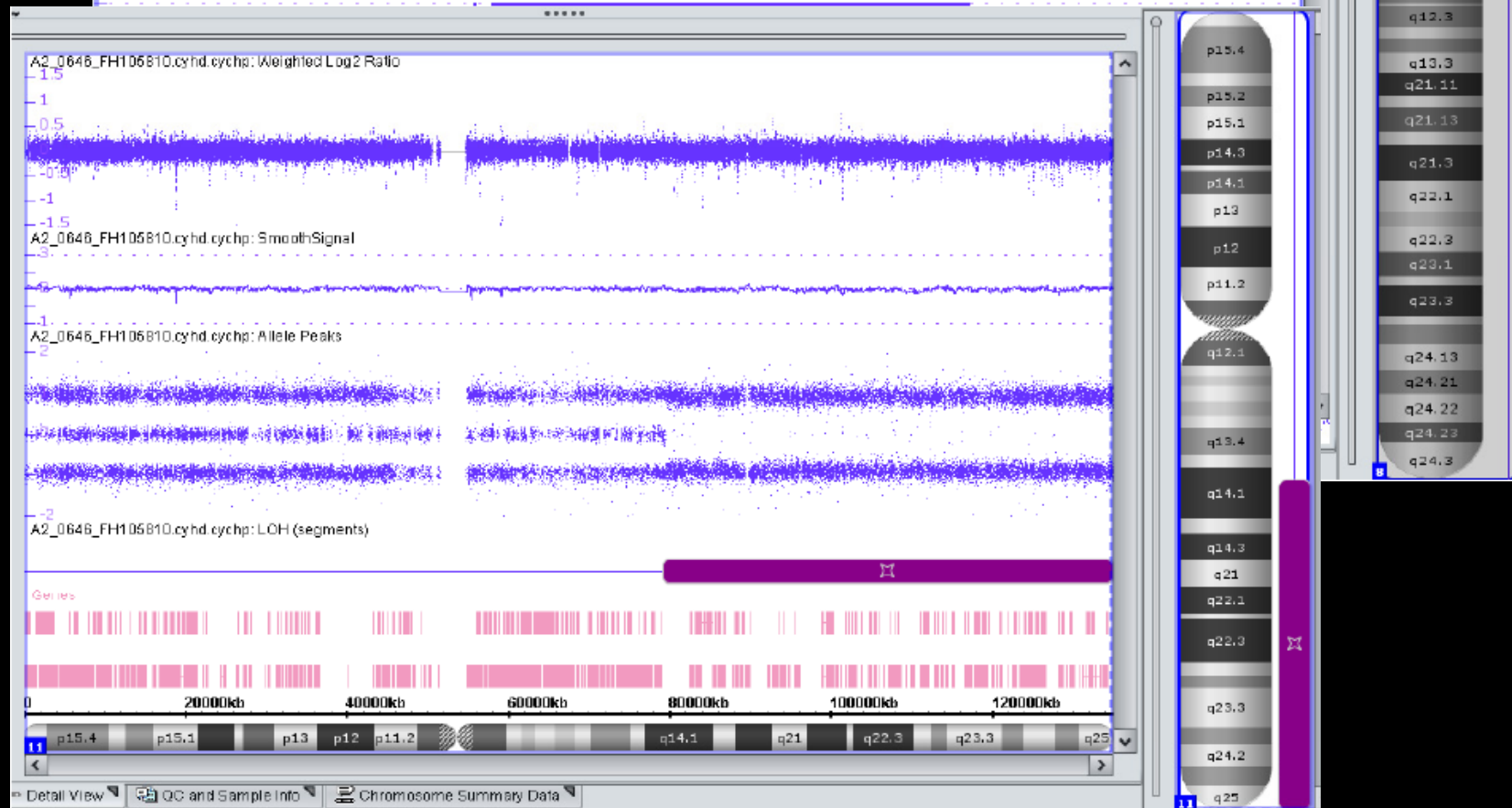


# Výsledky SNP čipů

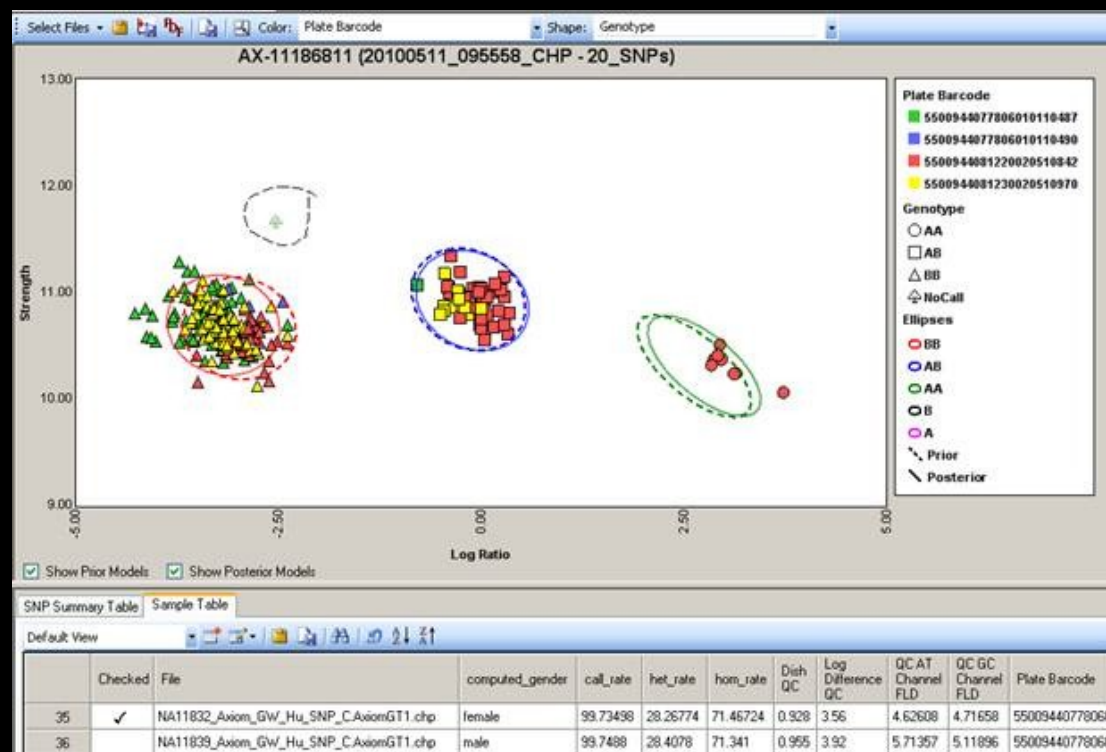
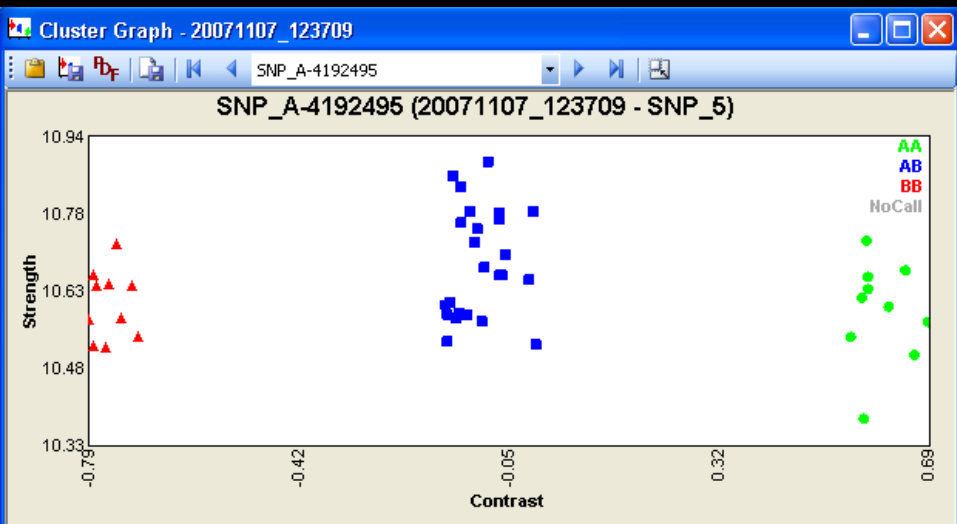
Amplifikace + delece



Uniparentální  
dizomie



# Výsledky SNP čipů - genotypizace





# Využití CGH, SNP čipů

- **Nádorová genomika**
  - V nádorových buňkách často změny genomu – primární i sekundární
- **Klinická genetika**
  - Přesnější charakterizace mikródelečních syndromů
  - Identifikace genomových změn u pacientů s mentální retardací i jinými vrozenými postiženími
- **Prenatální a preimplantační diagnostika**
- **Asociační studie – asociace SNP i CNV s řadou onemocnění (revmatoidní artritida, diabetes, nádorová onemocnění, psychiatrická onemocnění)**
- **Evoluční studie**

# Resekvenační čipy

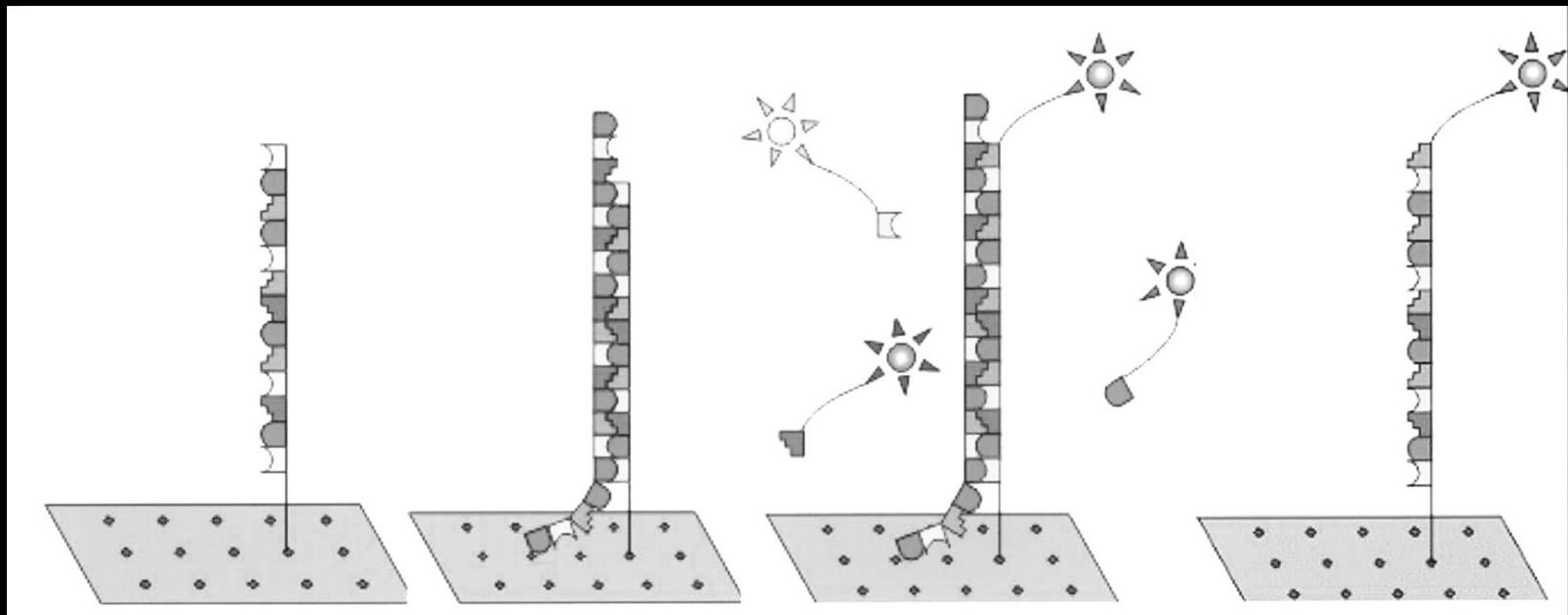
- Affymetrix

- Stejný princip jako detekce SNP
- SNP = testovány 2 varianty X resekvenování = testovány 4 varianty

TCGGTAGCCATGAATGAGTTACTAC	<i>Probe 1 Forward</i>
TCGGTAGCCATGCATGAGTTACTAC	<i>Probe 2 Forward</i>
TCGGTAGCCATGGATGAGTTACTAC	<i>Probe 3 Forward</i>
TCGGTAGCCATGTATGAGTTACTAC	<i>Probe 4 Forward</i>
ATCGGTAGCCATGCATGAGTTACTACAGCT	<i>Genomic Sequence of interest</i>
TAGCCATCGGTACGTACTCAATGATGTCGA	
AGCCATCGGTAGATACTCAATGATG	<i>Probe 1 Reverse</i>
AGCCATCGGTAGCTACTCAATGATG	<i>Probe 2 Reverse</i>
AGCCATCGGTAGGTACTCAATGATG	<i>Probe 3 Reverse</i>
AGCCATCGGTAGTTACTCAATGATG	<i>Probe 4 Reverse</i>

# Resekvenační čipy

- APEX
  - Arrayed Primer EXtension
  - Prodloužení sondy o jeden nukleotid komplementární ke vzorku
  - 4-barevná technologie – nutné speciální vybavení



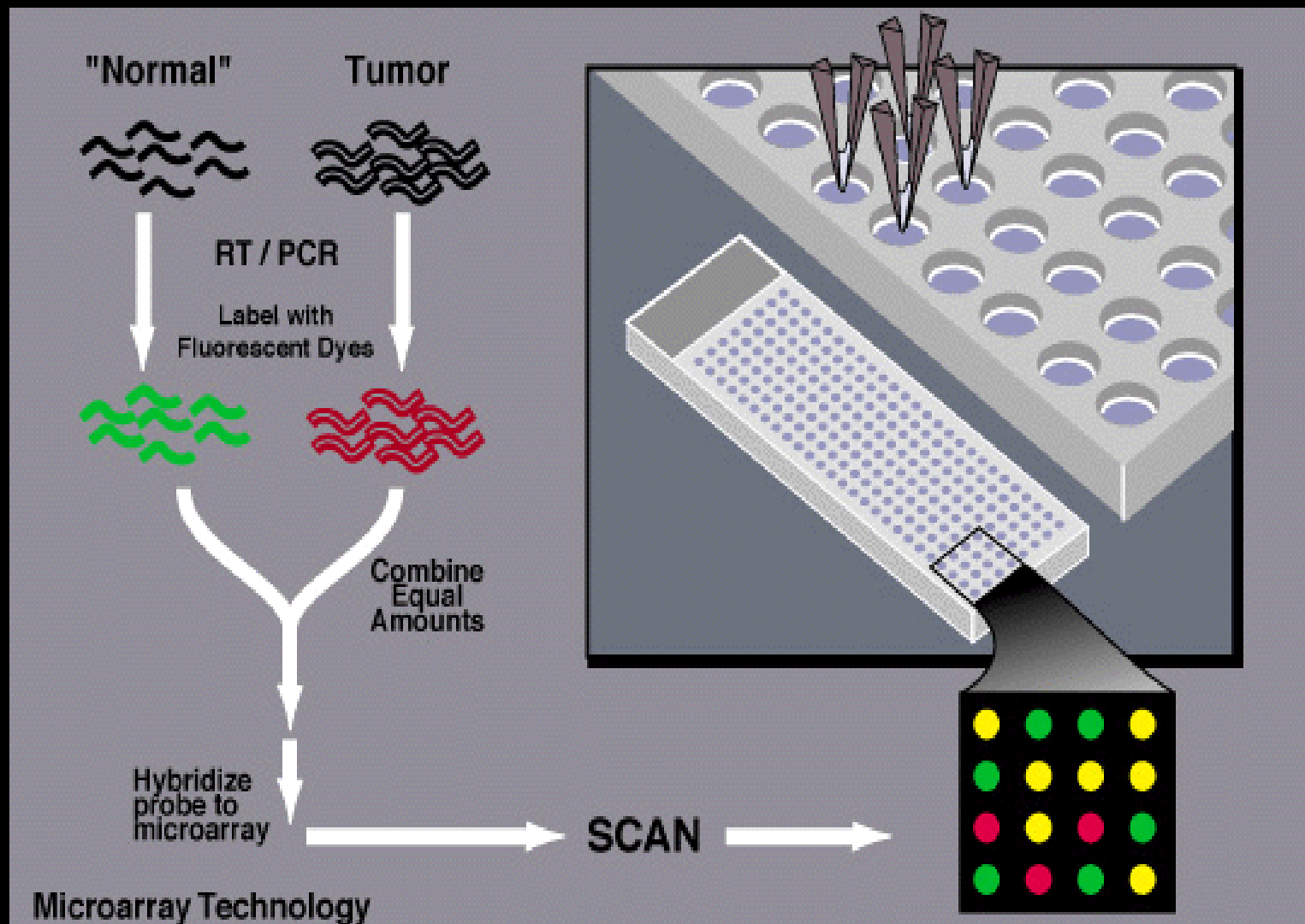


# Resekvenační čipy – hudba minulosti

- Paralelní sekvenace více oblastí
- Detekce mutací, na které je čip navržen
- Screeningová metodika
- Custom arrays – nutno ověřit Sangerovým sekvenováním
- Komerčně dostupné „diagnostické“ čipy
  - Roche – platforma Affymetrix
    - Cytochrom P450 – FDA approval, CE-IVD
    - P53 v testování
    - Není nutné ověřovat výsledek
    - Jen 1 gen , ale velmi rychlé, „user friendly“ (protokol max 2 dny)
  - APEX
    - Dostupné čipy pro konkrétní geny + možný vlastní design
    - Nutné ověřit sekvenací
    - Research use only

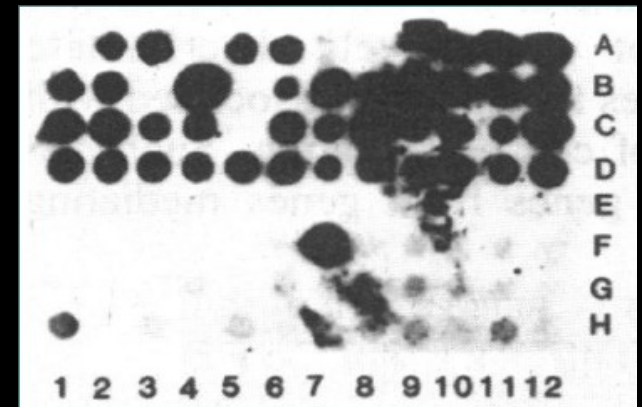
# Expresní čipy - princip

 Buňky kontrolní     Buňky sledované



# Historie

- 1987 Kulesh et al. - cDNA knihovna na papírovém filtru
- Přelom 80./90. - E. Southern – in-situ syntéza oligonukleotidů
- 1991 – Fodor et al. - světlem řízená paralelní syntéza
- 1995 – Schena et al. - cDNA expresní microarray
- 1996 – lidská cDNA microarray
- 1997 – celogenomová (kvasinka)
  - cDNA microarray
- 1997 – celogenomová (kvasinka)
  - oligonukleotidová microarray



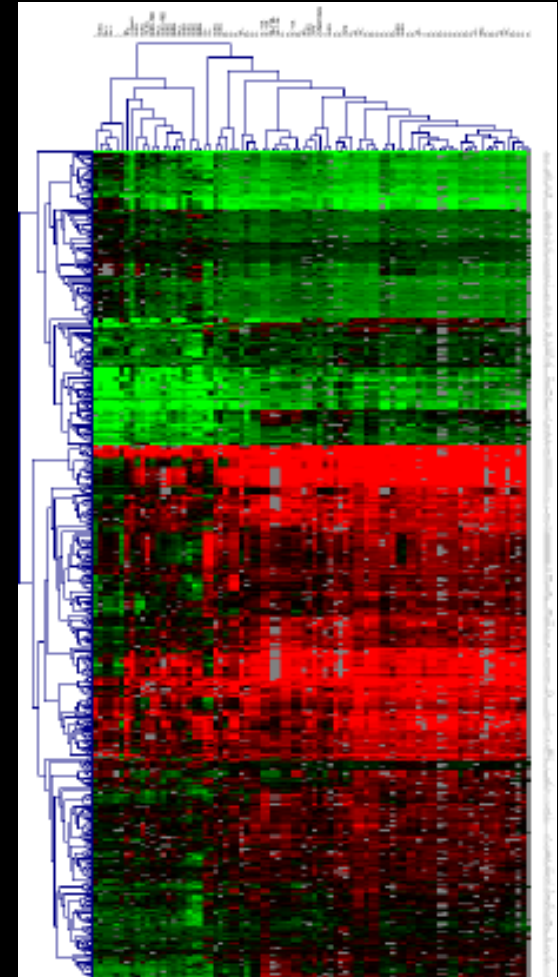


# Expresní čipy - typy

- Studium exprese mRNA (miRNA)
  - Celogenomové
  - Exonové (whole transcript)
  - Cílené – konkrétní dráhy, diagnózy
- High- i low- density
- Sondy – krátké i dlouhé oligonukleotidy, cDNA
- Sklíčka, cartridge, mikrokuličky, membrány
- Jedno- a dvou- kanálové

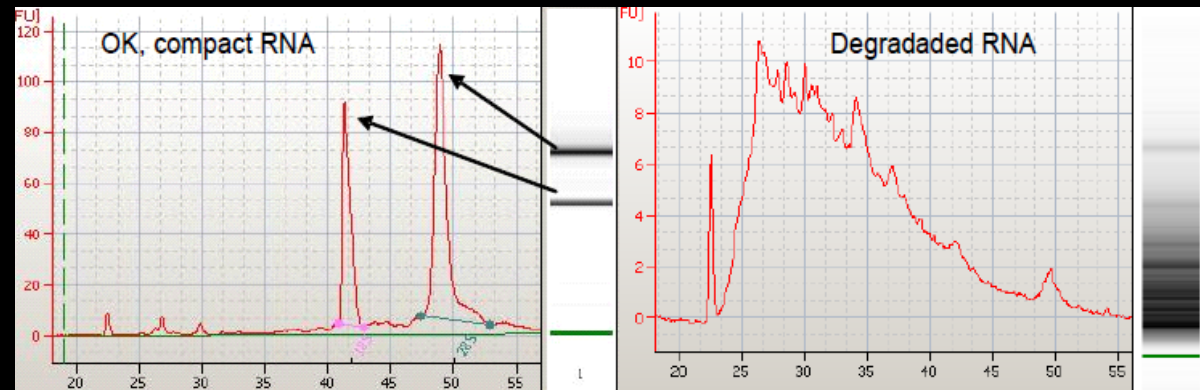
# Aplikace

- Studium efektů stimulů na genovou expresi
  - Chemikálie (např. léčiva)
  - Fyzikální faktory (záření)
  - Gene knock-out (siRNA)
  - Gene transfection (např. transkripční faktory, mutantní alely)
- Studium rozdílů mezi vzorky
  - Tkáně a orgány
  - Diagnózy



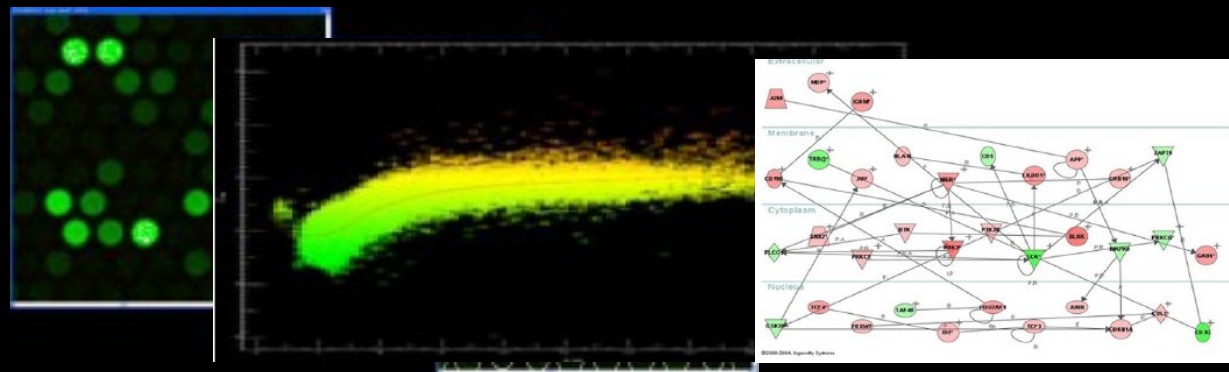
# Postup

- Izolace RNA
- Kvantita, kontrola kvality
- Přepis RNA
  - Oligo-dT priming
    - Kvalitní RNA
    - 3' biased sondy
  - Random priming – hexa-, okta-, nona-mery
    - I degradovaná RNA
    - Celý transkript

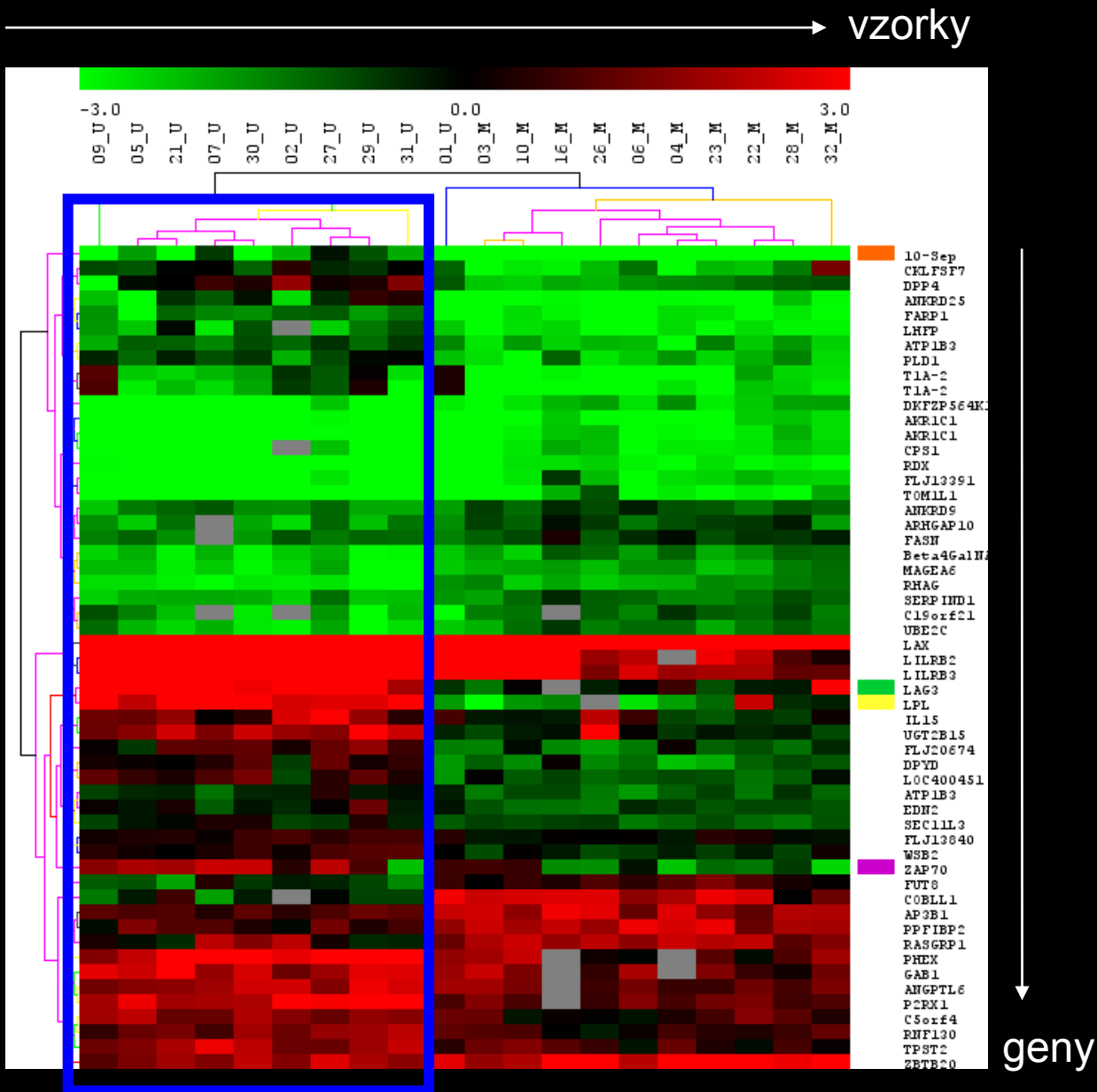


- Značení - fluorescence
- Hybridizace, odmytí
- Skenování
- Analýza dat

- Zpracování obrazu
- Normalizace
- Klastrová analýza (supervised, unsupervised)
- Integrace s dalšími daty – gene ontology, genové interakce, funkční vztahy

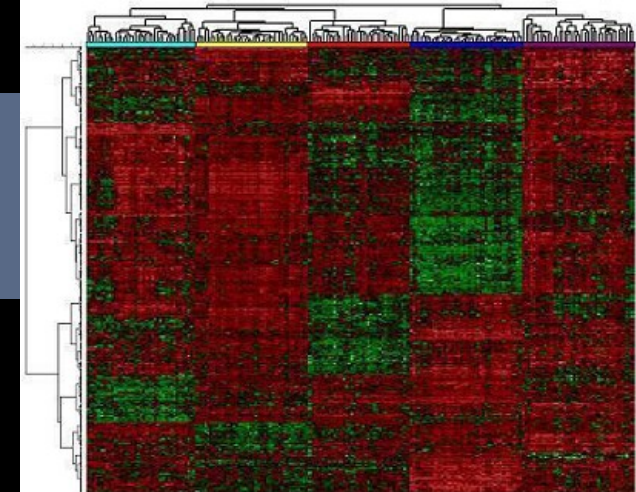


# Výsledky



- Identifikace genů rozdílně exprimovaných u konkrétních skupin vzorků
- Ověření pomocí real-time PCR

# Využití



- Pochopení molekulárních mechanismů
- Studium biomarkerů – diagnostické, prognostické, prediktivní
- Komerčně dostupné pro *in vitro* diagnostiku
  - FDA cleared / IVDMA (In Vitro Diagnostic Multivariate Index Assays)
    - MammaPrint®
      - Stanovení rizika metastázování po chirurgickém odstranění nádoru prsu u pacientek v časných stádiích
        - nízké riziko – hormonální terapie X vysoké riziko – agresivnější léčba (chemoterapie)
      - 70 genů
    - Tissue of Origin Test
      - Identifikace původu nádorů – metastáze, nediferencované nebo málo diferencované nádory
      - > 2000 genů
    - Řada dalších bez certifikace pro IVD (ColoPrint....)



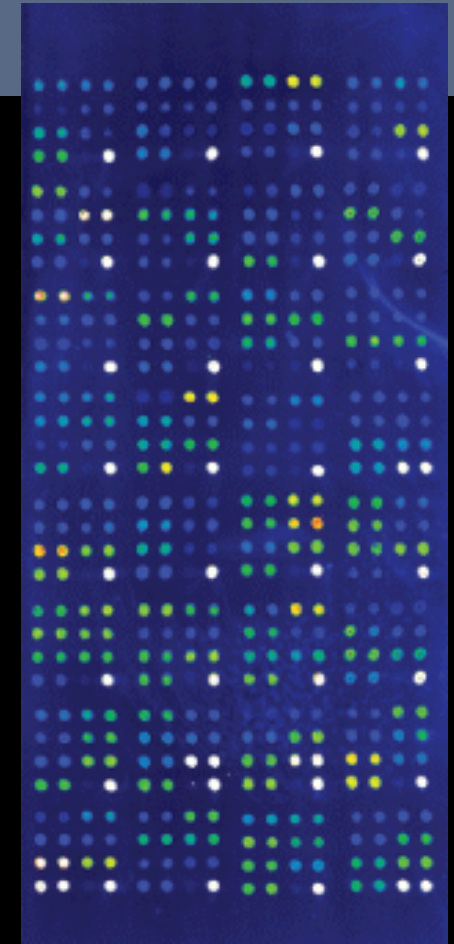
# Proteinové čipy

## ■ Expresní

- stanovení přítomnosti a koncentrace proteinů v komplexních vzorcích
  - Protilátkové čipy
  - Lyzátové čipy
  - Antigen čipy

## ■ Funkční

- studium interakce proteinů s jinými molekulami
  - proteiny, peptidy, nízkomolekulární látky, oligosacharidy, DNA



# Princip proteinových čipů



- Vazba protilátka-antigen - mikrosopot ELISA (Enzyme-Linked ImmunoSorbent Assay)
  - Systém využívající malé množství vázané protilátky a malého objemu vzorku je citlivější než systém se stonásobně větším objemem materiálu
  - Citlivost řádově ve femtomolech -  $10^6$  molekul/ml
  - Stanovení koncentrace analytu ve vzorku
    - Koncentraci přímo odpovídá množství vázaného analytu

*(Ekins RP, J Pharm Biomed Anal. 1989)*

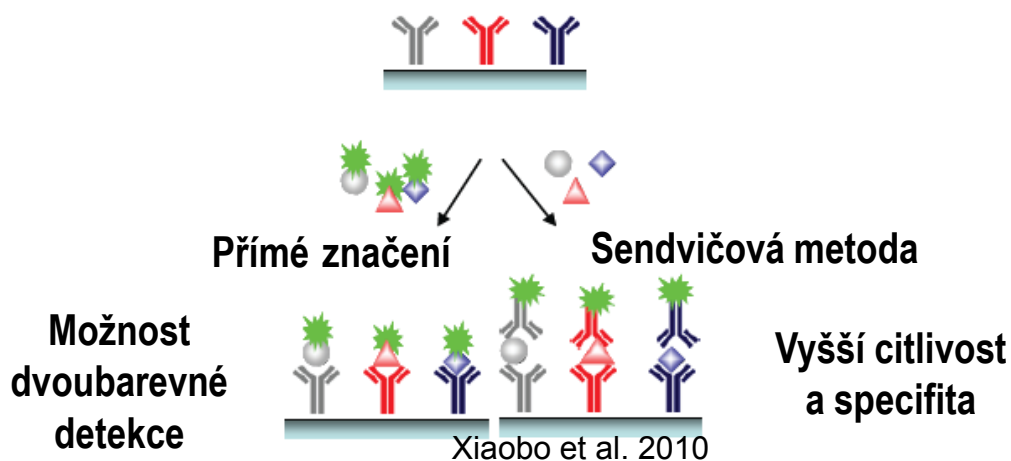
# Princip proteinových čipů

- Stovky – tisíce proteinů imobilizovaných ve formě mikrospotů na pevném povrchu
  - membrány (polystyrenové, PVDF - polyvinyliden fluorid, nitrocelulosové)
  - standardní mikroskopická sklíčka
    - s chemicky modifikovaným povrchem (poly-lysin, aldehydické skupiny)
    - potažená membránou
- Detekční metody
  - Nejčastěji fluorescence
  - Enzymaticky – značení alkalickou fosfatázou, křenovou peroxidázou
  - Chemiluminiscence
  - Radioaktivita
  - Zvýšení signálu – značení biotinem – vazba na značený streptavidin
  - Hmotnostní spektrometrie

# Čipy pro stanovení antigenů

## Protilátkové čipy

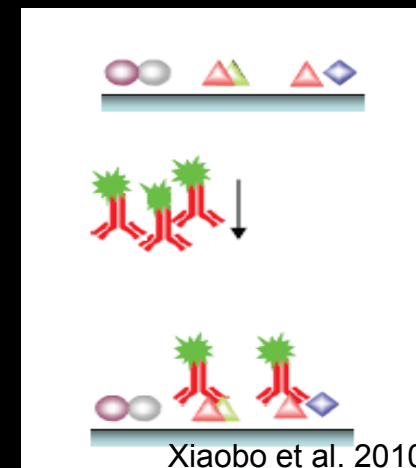
- Přímý formát (FPA- forward phase arrays)
- Na povrchu spotované protilátky
- Inkubace se vzorkem



- Detekce stovek antigenů ve vzorku
- Expresní profilování – „large-scale“ monitorování proteinové exprese
- Cílené na určité buněčné procesy (buněčný cyklus, cytokiny...)

## Lyzátové čipy

- Zpětný formát (RPA – reverse phase arrays)
- Na povrchu spotované vzorky (proteinové, tkáňové lyzáty, sérum)
- Inkubace se specifickými značenými protilátkami

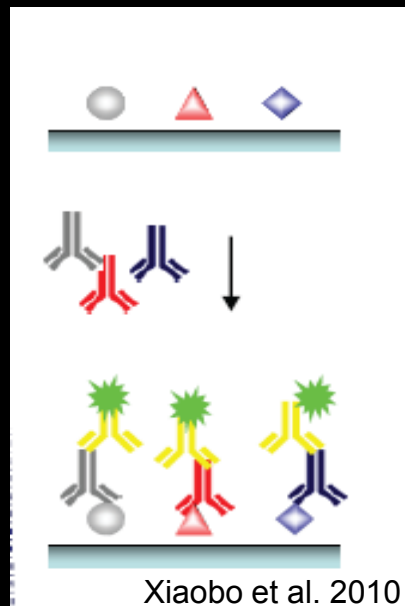


- Srovnání exprese až desítek antigenů u stovek různých vzorků
- Nevýhoda – malá hustota studovaných molekul ve spotu – pre-frafrakcionace pomocí 2D kapalinové chromatografie

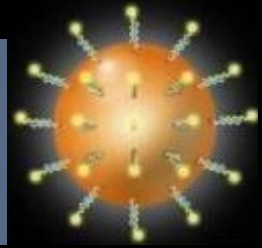
# Čipy pro stanovení protilátek

## Antigen čipy

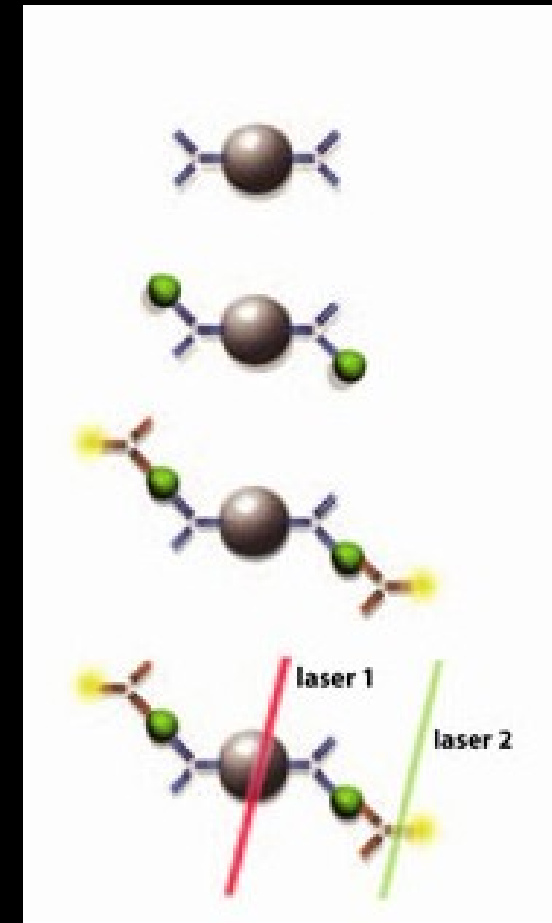
- Na povrchu spotované známé antigeny
  - syntetické proteiny, peptidy
- Inkubace se sérem obsahujícím studované protilátky
- Detekce přítomnosti protilátek ve vzorku, nejčastěji v séru



# Bead array - mikrosféry



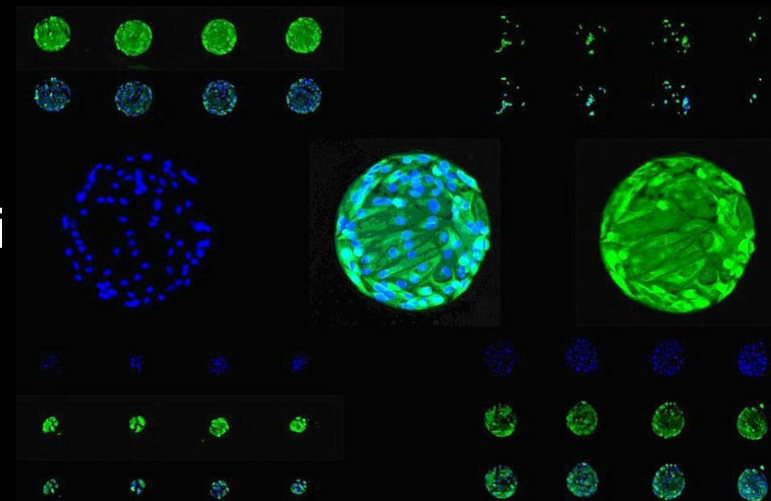
- průměr ~ 10  $\mu\text{m}$  (velikost lymfocytu)
- polystyrenové nebo latexové kuličky
- „kódovány“ rozdílnými barvami nebo velikostmi
- Detekce - např. flow cytometricky
  - rozpoznávání rozdílně kódovaných mikrosfér
  - identifikace a kvantifikace proteinů ve vzorku
- Množství stanovovaných proteinů je limitováno počtem druhů mikrosfér (~100)
- Možnost automatizace
- Využití – cílené – zejména detekce cytokinů, ale např. i detekce fúzních proteinů a onkoproteinů



[www.lucerna.chem.ch](http://www.lucerna.chem.ch)

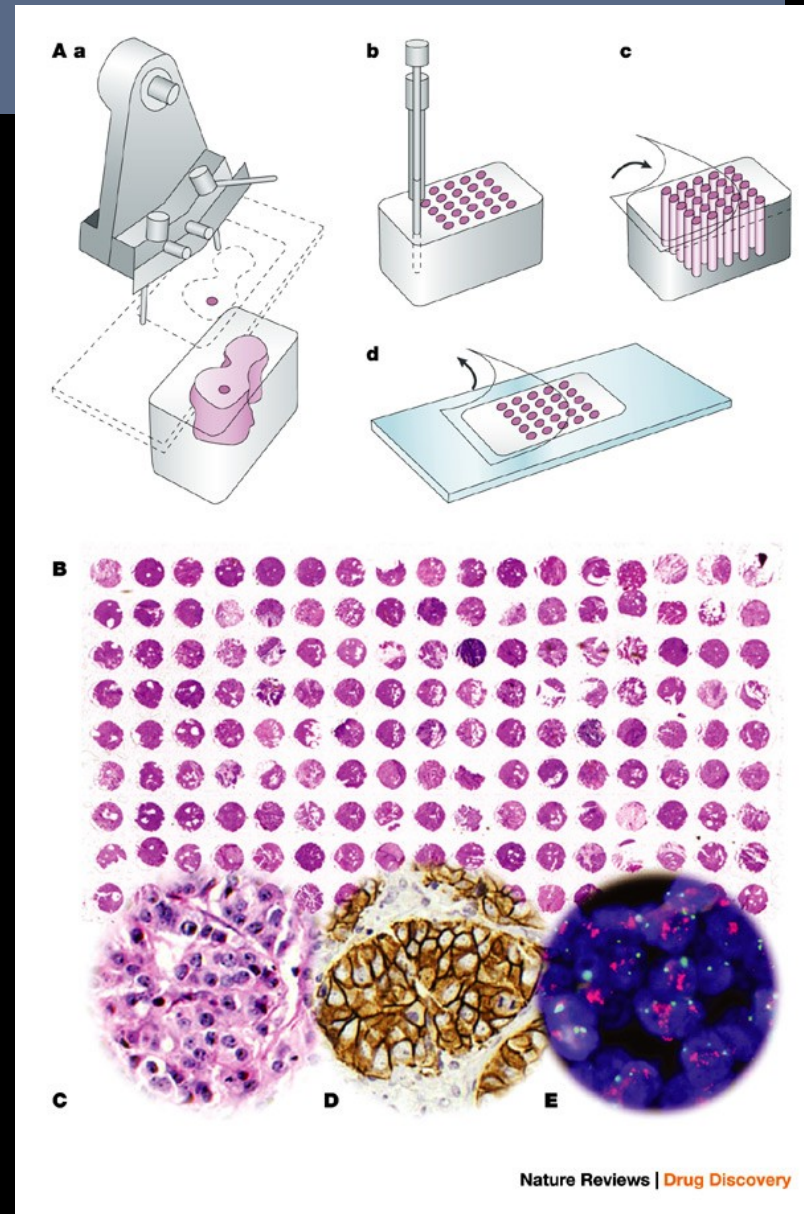
# Buněčné čipy

- Přímý formát
  - Spotované protilátky
  - Inkubace s buněčnou suspenzí
- „Transfekční čipy“ - buňky rostou na povrchu čipu s naspotovanou cDNA
  - Během růstu přijmou cDNA
  - Čip se spoty buněk exprimujících různé cizorodé proteiny
- Detekce
  - Přímou mikroskopicky na sklíčku
  - Další charakterizace značenými protilátkami



# Tkáňové čipy

- Varianta RPA čipů (zpětný formát)
- Spotují se celé vzorky tkání např. bioptických
- Inkubace se značenými protilátkami
- Velikost spotu 0,6 - 2mm





# Funkční čipy

- Studium interakce s jinými molekulami
  - Inkubace s jinými proteiny, DNA, malými molekulami (oligosacharidy, terapeutika, ...)
- Studium posttranslačních modifikací
- Studium enzymatické aktivity
- Studium kofaktorů a inhibitorů
- Studium interakcí ligand-receptor



protein - DNA



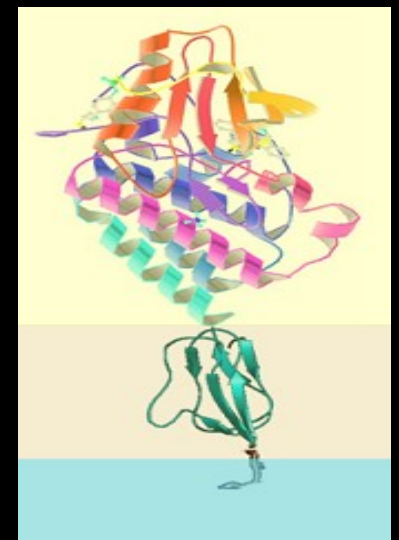
protein - protein



enzym - substrát

# Spotovány nativní proteiny

- Potřeba zachování struktury a biologické aktivity
- Nespecifická vazba přímo na povrch
  - Adsorpce
  - Kovalentní vazba přirozených chemických skupin proteinu na upravený povrch
  - Aktivní část proteinu může být schovaná, problém zachování konformace
- Chemoselektivní vazba – připojení chemické skupiny na definovanou pozici proteinu → specifická reakce s komplementární chemickou skupinou na povrchu sklíčka
  - Orientovaná pozice na sklíčku
- Imobilizace přes specifický rekombinantní tag
  - Zachování orientace, konformace, funguje jako „spacer“



## Další typy funkčních čipů

- Doménové čipy
  - Spotovány pouze jednotlivé domény proteinu
  - Pochopení protein-proteinových interakcí
- Peptidové čipy
  - Inkubace s proteiny, DNA, malými molekulami (oligosacharidy, terapeutika, ...)
- Čipy s chemickými knihovnami
  - Inkubace s proteiny (enzymy, receptory...)
  - Studium inhibitorů, studium ligand-receptorových interakcí-  
vyhledávání terapeutik

Typ čipu		Molekuly imobilizované na čipu	Stanovované molekuly	Inkubace	Počet spotů	Použití
<b>expresní</b>	protilátkové (přímý formát)	známé protilátky	antigeny	neznámý vzorek - proteinový lyzát	1000	proteinové profilování
	lyzátové (zpětný formát)	neznámý vzorek - proteinový lyzát	antigeny	známé protilátky	1000	proteinové profilování
	antigen čipy	známé antigeny	protilátky	neznámý vzorek – protilátky v séru	1000	stanovení přítomnosti protilátek v séru
	<b>funkční</b>	protein- protein protein- DNA, RNA protein- nízkomolek ulární látky Enzym- substrát	známé proteiny v nativním, funkčním stavu nebo peptidy,  nebo  chemické látky	partnerské molekuly interagující s proteiny  nebo  proteiny	100	studium interakcí proteinů s partnerským i molekulami

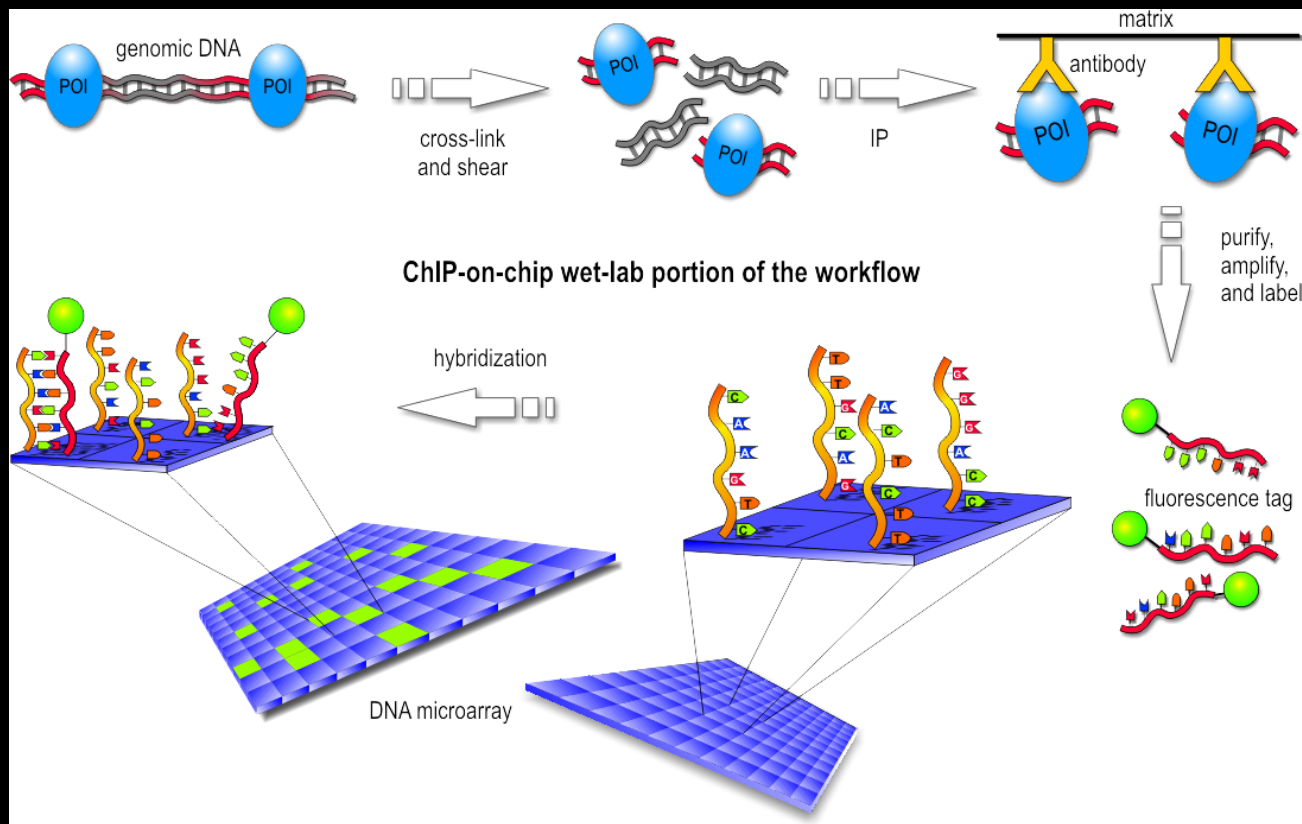


# Využití

- Studium biomarkerů – diagnostické, prognostické, prediktivní
- Identifikace terapeutických cílů
- Detekce autoprotilátek, studium imunitní odpovědi
  
- *In vitro* diagnostika – více než u ostatních typů čipů
  - Nejčastěji diagnostika autoimunitních onemocnění
  - Diagnostika infekčních onemocnění

# ChIP-on-chip (Chromatin immunoprecipitation)

- Mapování vazby proteinů na DNA
  - Např. transkripční faktory



## MeDIP-chip (Methyl-DNA immunoprecipitation)

- Stejný princip – imunoprecipitace s protilátkou proti 5-metyl-cytosinu