

Principy microarrays II

Pavla Gajdušková
Analytická cytometrie, 1. prosince 2009

Obsah přednášky

Technologie přípravy microarrays

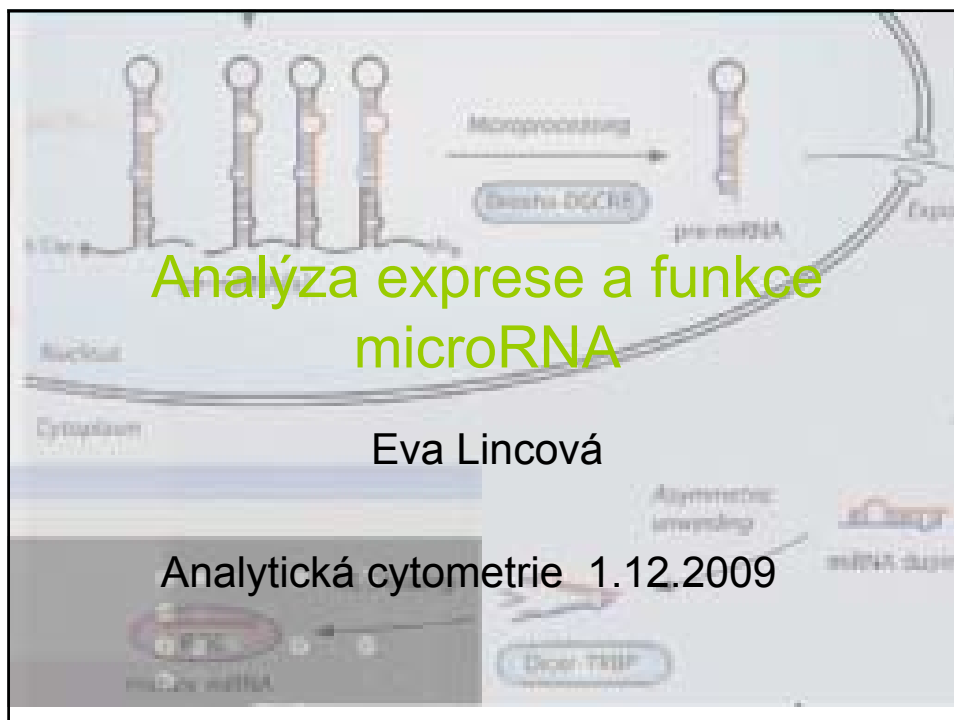
Oblasti použití microarrays v biologii

Úvod do statistického hodnocení dat

Příklady konkrétních aplikací z literatury

Oblasti použití microarrays v biologii

Typ array	Sondy na microarray	Co se fluorescenčně značí a hybridizuje	... analýza čeho
Expresní	DNA (cDNA, oligonucleotidy)	mRNA / cDNA	měření množství mRNA v bunkách, nádorech ...
miRNA	oligonukleotidy	miRNA	měření množství miRNA
CGH	DNA (BAC vektory, oligonukleotidy)	DNA	změny v genomu (zisk, ztráta chromozomů nebo jejich částí)
SNP	DNA (oligonukleotidy)	DNA	detekce „Single Nucleotid Polymorphisms“; změny v genomu
Metylace	DNA (CpG islands)	DNA (ovlivněná bisulfidem sodným)	míra metylace promotorových oblastí
Promoter	DNA (promotorové oblasti ~ 1kb)	DNA (ChiP obohacená)	místa vazby transkripčních faktorů, modifikace histonů
Tilling	DNA	všechno dříve zmíněné	všechno dříve zmíněné, sekvenování, anotace genů
Protein	protilátky	protein	exprese proteinů (ELISA)



Analýza exprese a funkce microRNA

1. Regulační RNA molekuly
2. Struktura, vznik a funkce miRNA
3. Databáze a predikce miRNA
4. Metody analýzy miRNA
 - microarrays
 - qPCR
 - bead-based methods
 - funkční analýzy



1. Regulační RNA molekuly

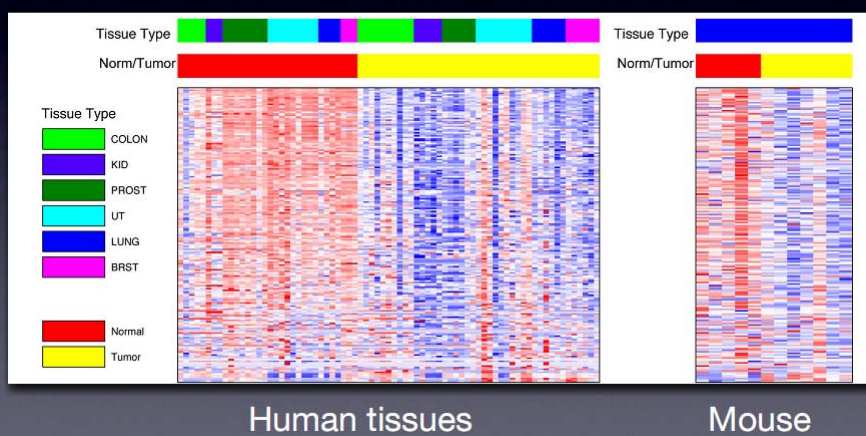
- Negativní regulátory translace a stability mRNA
- 1998 objev RNA interference (Andy Fire a Craig Mello)
- Význam RNAi:
 - Obrana proti virové infekci
 - Udržení genomové stability (umlčení transposonů)
 - Inhibice syntézy proteinů a regulace vývoje organismu
 - Zabránění transkripce – udržení chromatinu v kondenzovaném stavu
 - Experimentální umlčení genů
 - Genová terapie



Proč studovat miRNA

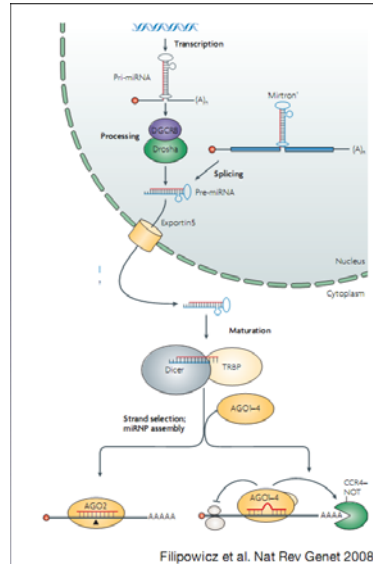
- Kontrola buněčného cyklu, apoptózy, vývojových a fyziologických procesů
(diferenciace kmenových buněk, hematopoéza, hypoxie, vývoj svalstva, neurogeneze, sekrece inzulinu, metabolismus cholesterolu, imunitní odpověď, replikace virů)
- Evolučně konzervovaný mechanismus
- Tkáňově specifická a časově ohraničená exprese během embryogeneze
- Změny v expresi při mnoha onemocněních

Many microRNAs are down-regulated in primary human tumours



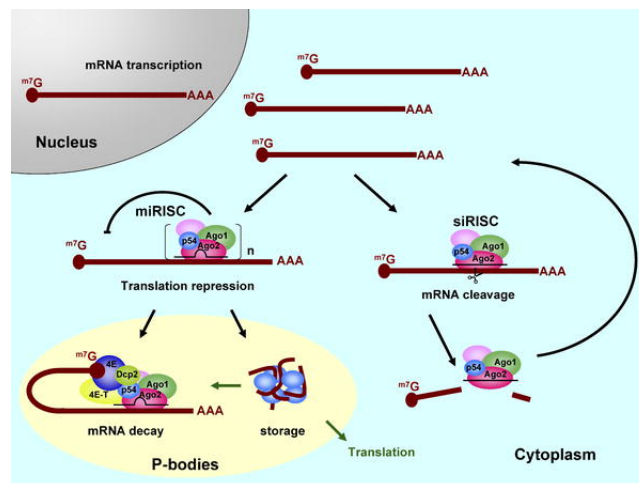
2. Biogeneze miRNA

1. Dlouhý primární transkript -> pri-miRNA (stovky nt-desítky kbp)
2. Sestřih - RNasa III Drosha + Pasha/DGCR8 -> pre-miRNA (70 nt)
3. Export (Exportin 5)
4. Maturace - RNasa III Dicer - začlenění jednořetězcové 19-23 nt dlouhé miRNA do komplexu RISC (tvořen proteiny Argonaut)
5. Negativní regulace mRNA - inhibice translace nebo degradace (P-bodies)



2. Funkce miRNA

- Vazba komplexu na 3'UTR
- Degradace mRNA
- Inhibice translace – iniciační nebo elongační fáze



3. Databáze a predikce miRNA

- miRBase (09/09 ver. 14.0)
- <http://www.mirbase.org/>



2/11/2009
miRNA count: more than 10000 entries
(762 human)

3. Predikce cílových molekul

- Každá miRNA může regulovat až 100 různých molekul
 - miRNA jsou krátké
 - 3' UTR jsou dlouhé a nepříliš kvalitně osekvenované a anotované
 - Dosud málo ověřených cílových molekul
 - Vazebná místa nejsou dokonale komplementární
 - Klíčová oblast specifity je krátká (6-8 nt)
- TargetScan
 - PicTar
 - miRBase Target



4. Metody analýzy miRNA

- ... aneb co brát v úvahu při výběru metody
- „Throughput“ – počet vzorků, miRNA
- Případné artefakty
- Cena (reagencie, materiál, instrumentace)
- Otevřená vs. uzavřená platforma (možnost vytvářet vlastní assays)
- Možnost standardizace

Příprava vzorků

- Celková RNA (Trizol)
- RNA obohacená o malé molekuly (komerční kity)
- Kontrola kvality

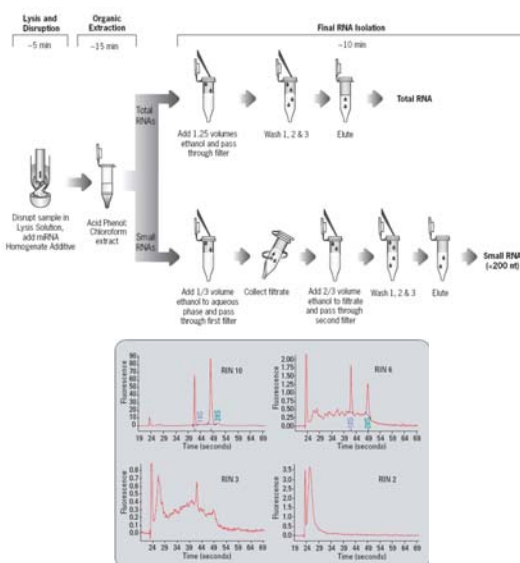
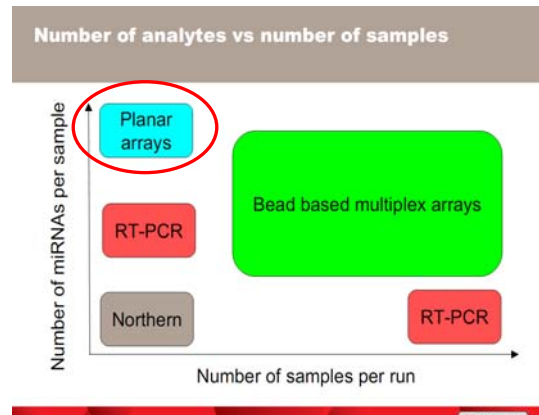


Figure 2. Sample electropherograms used to train the RNA Integrity Number (RIN) software. Samples come from either RIN 10 or degraded RIN 0.

4. Metody analýzy miRNA

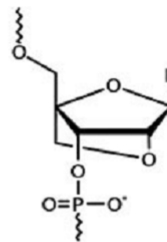
- microarrays
- qPCR
- bead-based methods
- funkční analýzy



locked nucleic acids for miRNA detection

Zvýšení afinity

	LNA™	RNA
T _m increase/monomer against DNA (°C)	2.0-6.0	-0.5-0.5
T _m increase/monomer against RNA (°C)	3.0-8.0	1.0-1.5
ΔT _m at single mismatch against DNA	LNA>>DNA	RNA>DNA
Compatible with standard molecular biology	Yes	Yes
Water solubility	High	High

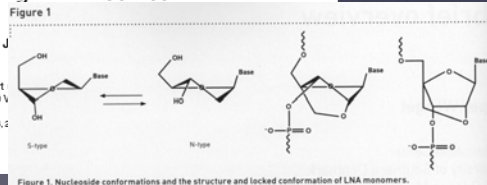


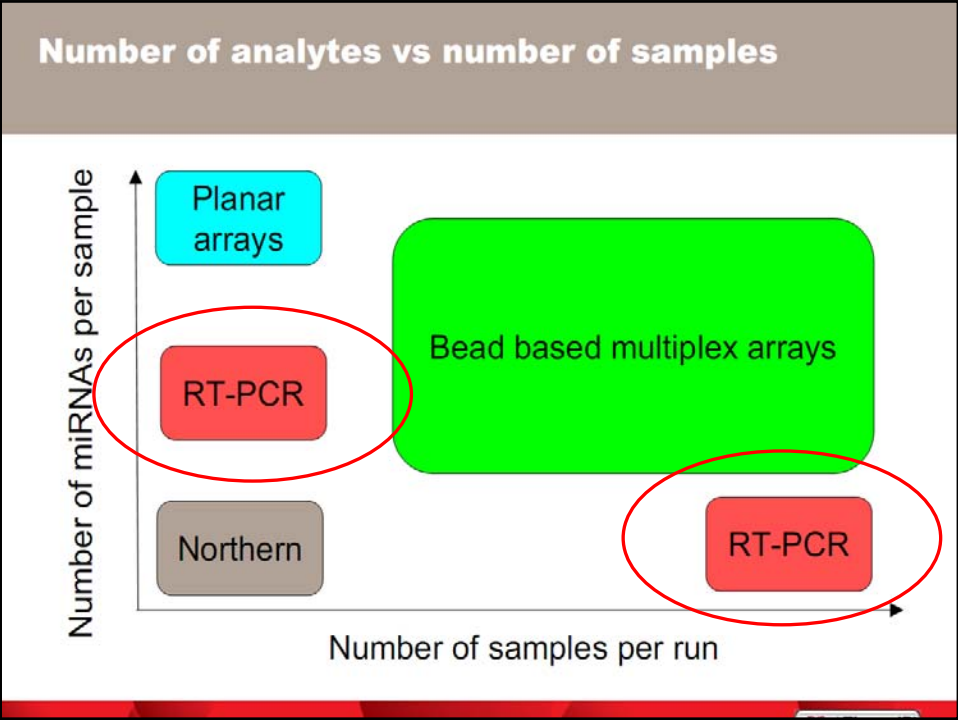
Sensitive and specific detection of microRNAs by northern blot analysis using LNA-modified oligonucleotide probes

Anna Válczi, Csaba Homiyik, Nóra Varga, J and Zoltán Havelda*

Agricultural Biotechnology Center, Szent-Györgyi Albert Functional Genomics, Eötvös, Byggestubben 9, DK-2950 V

Received September 23, 2004; Revised and Accepted November 16, 2004





b. qRT-PCR – SYBR Green

miRCURY™ LNA microRNA PCR system

Step 1:
First-strand cDNA synthesis

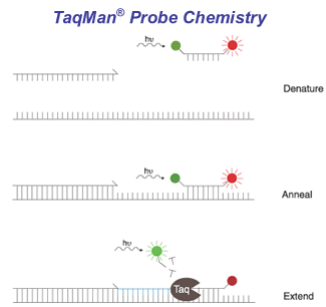
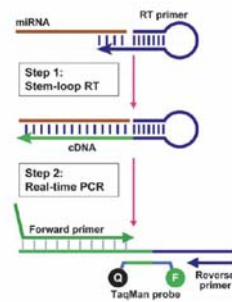


Step 2:
Real-time PCR amplification

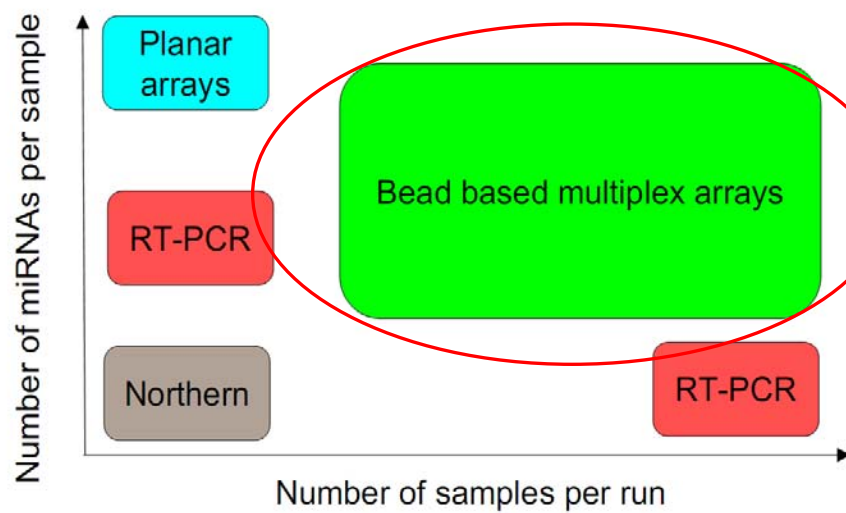


b. qRT-PCR - TaqMan

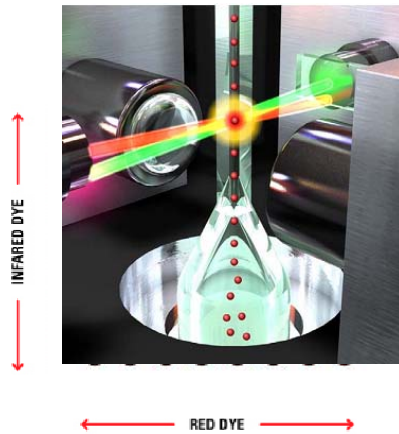
Chen *et al.* 2005 (ABI)



Number of analytes vs number of samples



c. Bead-based methods



100 distinct bead sets

Uniform polystyrene beads

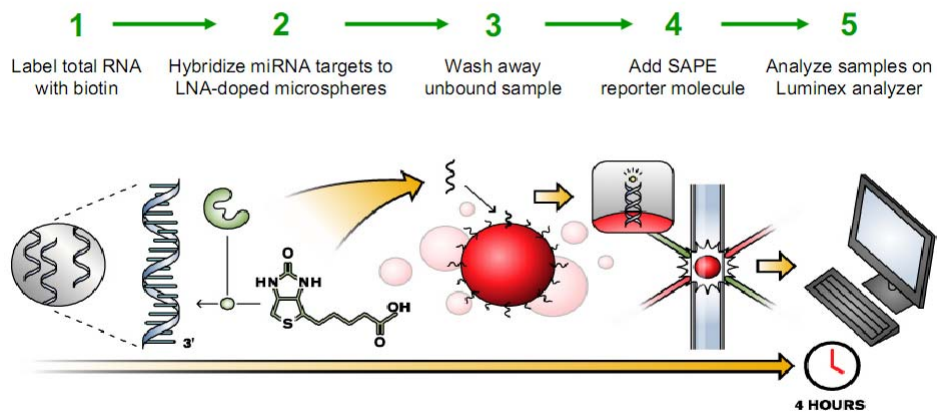
5.6 microns diam

Internal labeling with

- 2 dyes

- Each at 10 distinct levels

FlexmiR Workflow



d. Funkční analýzy

- Luciferázové reportéry
- Knockdown pomocí modifikovaných prób
- Overexprese miRNA

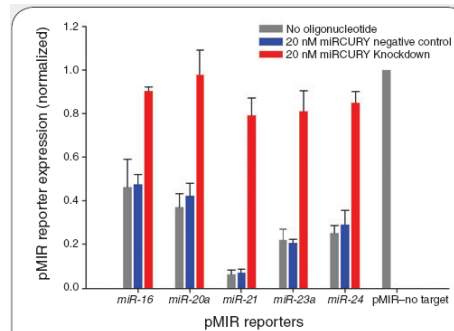


Figure 1 | Representative knockdown of microRNAs by miRCURY LNA Knockdown probes. Expression levels of the indicated miR reporter constructs carrying a miR complementary region. The no-target control is the parental vector without a miR target. HeLa cells were cotransfected with miR reporter (firefly luciferase), pGL4.13-Renilla (for normalization) and miRCURY LNA Knockdown probes against the named *hsa-miR* targets. Expression levels were measured using Dual-Glo™ Luciferase (Promega). Each experiment was performed at least three times, in triplicate, and a representative example is shown.

Shrnutí + „take home message“

- miRNA mohou regulovat až 30% genů interakcí s 3'UTR na mRNA
- Metoda analýzy závisí na počtu detekovaných miRNA a počtu vzorků
- Profil exprese miRNA se mění při celé řadě onemocnění – diagnostický či terapeutický potenciál

Oblasti použití microarrays v biologii

Typ array	Sondy na microarray	Co se fluorescenčně značí a hybridizuje	... analýza čeho
Expresní	DNA (cDNA, oligonucleotidy)	mRNA / cDNA	měření množství mRNA v bunkách, nádorech ...
miRNA	oligonukleotidy	miRNA	měření množství miRNA
CGH	DNA (BAC vektory, oligonukleotidy)	DNA	změny v genomu (zisk, ztráta chromozomů nebo jejich částí)
SNP	DNA (oligonukleotidy)	DNA	detekce „Single Nucleotid Polymorphisms“; změny v genomu
Metylace	DNA (CpG islands)	DNA (ovlivněná bisulfidem sodným)	míra metylace promotorových oblastí
Promoter	DNA (promotorové oblasti ~ 1kb)	DNA (ChIP obohacená)	místa vazby transkripčních faktorů, modifikace histonů
Tilling	DNA	všechno dříve zmíněné	všechno dříve zmíněné, sekvenování, anotace genů
Protein	protilátky	protein	exprese proteinů (ELISA)

Použití microarrays ke studiu DNA

Komparativní genomická hybridizace

- BAC arrays
- oligo arrays
- SNP arrays
- tilling arrays (BAC a oligonukleotidy)
- exon-specific arrays
- (dříve i cDNA arrays používané pro expresi)

Genotypování

- SNP arrays

Sekvenování

- Re-Sequencing arrays

ChIP-Chip experimenty

- tilling arrays (oligonukleotidy)

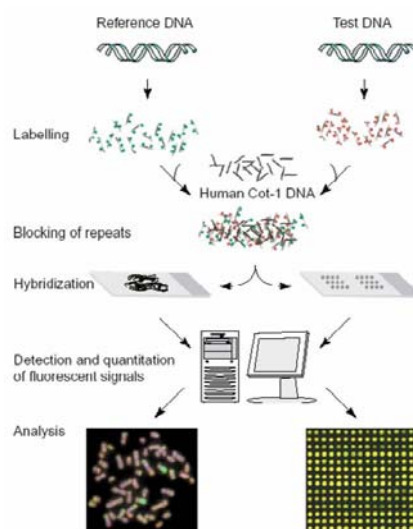
Komparativní genomická hybridizace (CGH)

molekulárně cytogenetická metoda, která slouží k analýze změn obsahu DNA v živých organismech

(delece, zisk, amplifikace různých oblastí genomu)

porovnávání intenzity fluorescence zkoumaného vzorku DNA a normálního diploidního vzorku DNA v různých místech genomu

Komparativní genomická hybridizace (CGH)



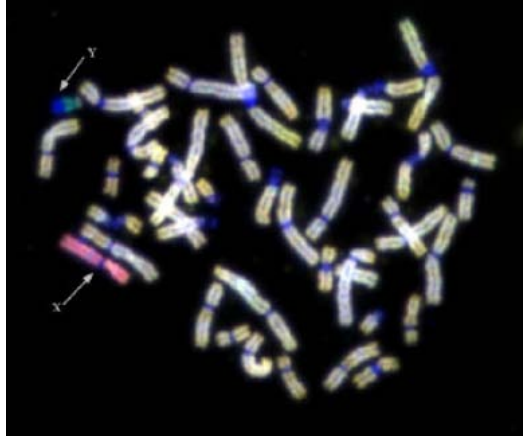
Mantripragada et al. Trends in Genetics 2003

Komparativní genomická hybridizace (CGH)

metafázní chromozomy
- dárce s normálním
diploidním karyotypem

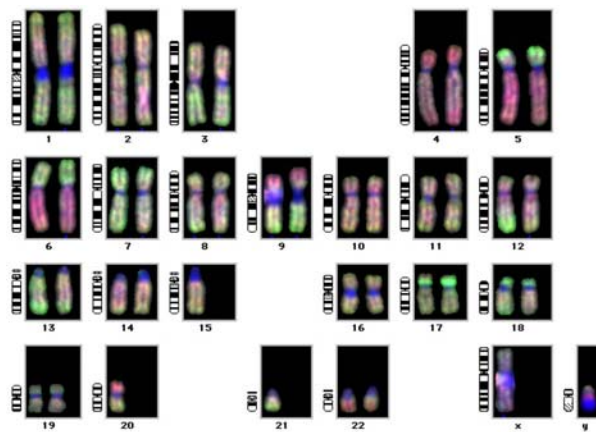
DNA:
cy3 (červená)
zkoumaný vzorek
cy5 (zelená)
referenční DNA – 2n

rozlišení ~ 20MB



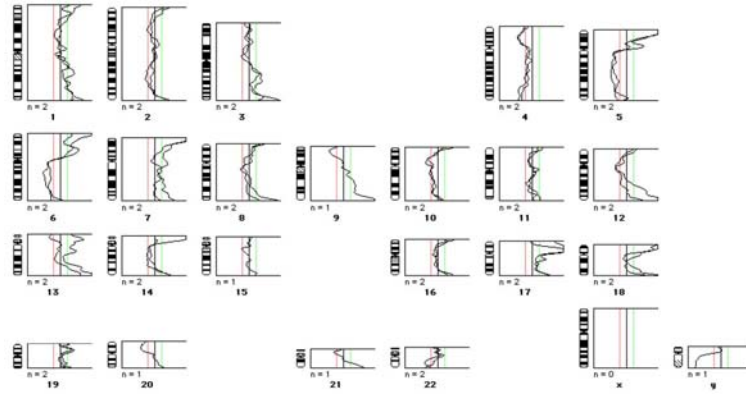
From Szu Hai K. presentation: Determination of Genomic Imbalances by Genome-wide Screening Approaches

Komparativní genomická hybridizace (CGH)



From Szu Hai K. presentation: Determination of Genomic Imbalances by Genome-wide Screening Approaches

Komparativní genomická hybridizace (CGH)



From Szu Hai K. presentation: Determination of Genomic Imbalances by Genome-wide Screening Approaches

„Array“ komparativní genomická hybridizace (Array CGH)

chromosomy nahrazeny body na mikroskopickém sklíčku,
které obsahují specifické DNA sekvence

Typy sond natištěných na microarray sklíčku

BAC klony až 32 000 BAC klonů na jednom sklíčku
~ 160 kb dlouhé úseky DNA

Oligonukleotidy 25 – 80 párů dlouhé oligonukleotidy
mohou pokrývat i celý genom (repetitivní
sekvence jsou vynechány)

známe polohu a pořadí všech sond v lidském genomu

Knihovny BAC klonů pro array CGH

BACPAC resources (CHORI)

<http://bacpac.chori.org>

Research Genetics (Invitrogen)

<http://www.resgen.com/resources/index.php3>

The Sanger Centre

<http://www.geneservice.co.uk/home/>

Cheung V. G. et al., Integration of cytogenetic landmarks into the draft sequence of the human genome. *Nature* 409: 953 – 958, 2001.

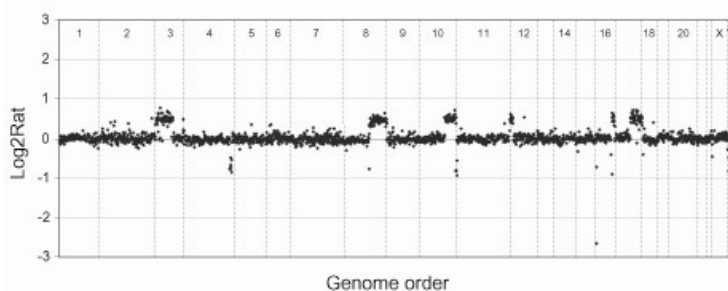
Greshock J. et al., 1-Mb Resolution Array-Based Comparative Genomic Hybridization Using a BAC Clone Set Optimized for Cancer Gene Analysis. *Genome Res* 14: 179-187, 2004.

Krzywinski M. et al., A set of BAC clones spanning the human genome. *Nucleic Acids Res* 32: 3651-3660, 2004.

Array CGH s použitím BAC klonů

$$\text{Log}_2\text{Rat} = \text{Log}_2 R/G$$

$\text{Log}_2\text{Rat} = 0$	2 kopie	$\text{Log}_2\text{Rat} = -1$	1 kopie ("loss")
$\text{Log}_2\text{Rat} = 0.5$	3 kopie ("gain")	$\text{Log}_2\text{Rat} < -1$	homozygotní delece
$\text{Log}_2\text{Rat} = 1$	4 kopie ("gain")		
$\text{Log}_2\text{Rat} = 2$	8 kopií ("amplification")		



2464 BAC klonů

UCSF HumArray3.1

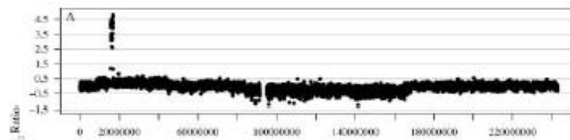
Typy sond natištěných na microarray sklíčku

BAC klony až 32 000 BAC klonů na jednom sklíčku
~ 160 kb dlouhé úseky DNA

Oligonukleotidy 25 – 80 párů dlouhé oligonukleotidy
mohou pokrývat i celý genom (repetitivní
sekvence jsou vynechány)

známe polohu a pořadí všech sond v lidském genomu

Array CGH - oligonukleotidy (NimbleGen)



6-kb median
probe spacing

Selzer RR et al. Genes Chromosomes Cancer, 2005

SNPs

SNP = single nucleotide polymorphism

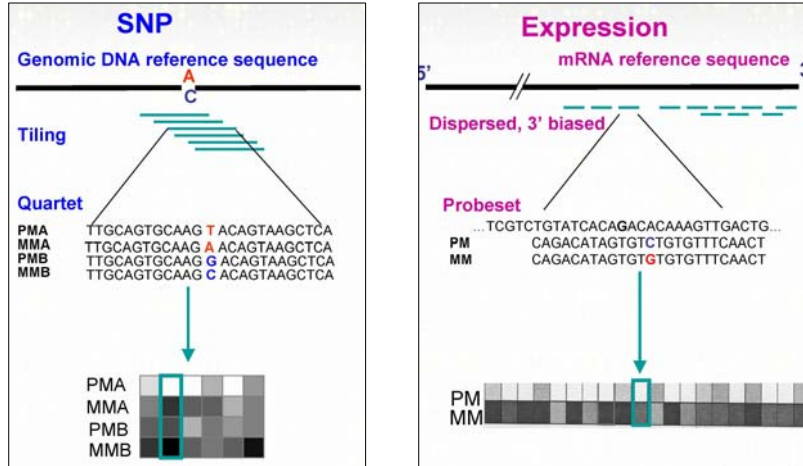
jednonukleotidové variace, které jsou náhodně rozmístěny v genomu (bodové mutace rozšířené v populaci)

nukleotidová variace, která se vyskytuje alespoň u 1% jedinců v populaci

předpokládaný počet SNPs: 10 milionů

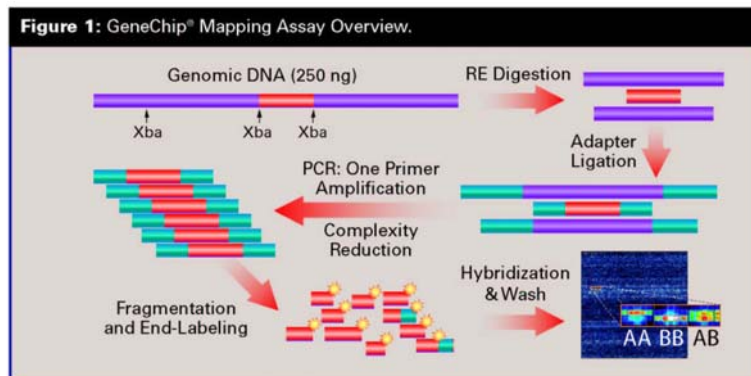
výskyt specifických SNP spojen s predispozicí k určitým chorobám

SNP arrays x expression arrays



From Xiao Y. presentation: Exploration and Analysis of Affymetrix SNP Arrays. Center for Bioinformatics & Molecular Biostatistics, UCSF Division of Biostatistics

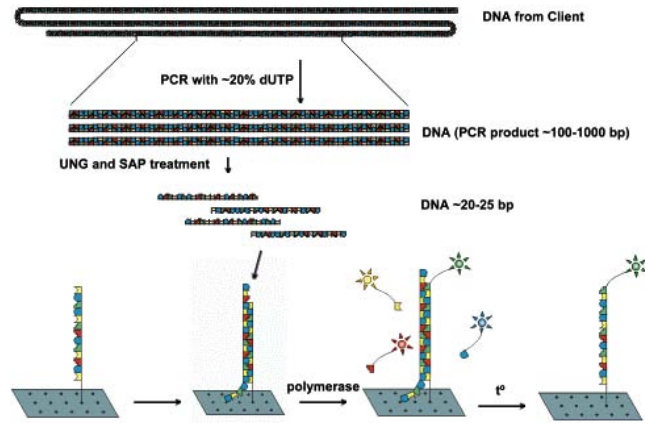
SNP Arrays - labeling



From Xiao Y. presentation: Exploration and Analysis of Affymetrix SNP Arrays. Center for Bioinformatics & Molecular Biostatistics, UCSF Division of Biostatistics

SNP Arrays - APEX technologie

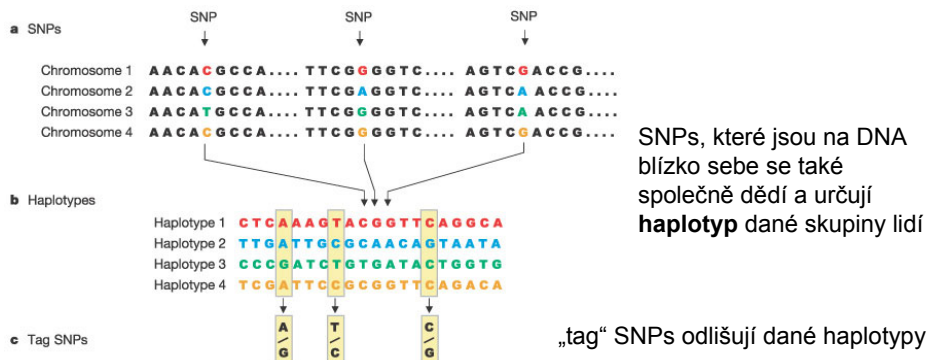
APEX = Arrayed Primer Extension



Kurg A. et al., Arrayed primer extension: solid-phase four-color DNA resequencing and mutation detection technology. Genet Test 4:1-7, 2000.

Velké studie SNP

HapMap projekt: mezinárodní projekt, jehož cílem je identifikovat a katalogizovat SNPs v lidské populaci a vybrat z nich „tag“ SNPs, kterými se skupiny lidí odlišují



HapMap projekt

<http://www.hapmap.org/index.html.en>

HapMap kolekce lidské DNA 270 vzorků DNA

populace:	Nigerie	30 trojic vzorků (matka, otec, dítě)
	Japonsko	45 nepříbuzných vzorků
	Čína	45 nepříbuzných vzorků
	USA	30 trojic vzorků (matka, otec, dítě)

The International HapMap Consortium. **A second generation human haplotype map of over 3.1 million SNPs.** *Nature* 449, 851-861. 2007.

The International HapMap Consortium. **A Haplotype Map of the Human Genome.** *Nature* 437, 1299-1320. 2005.

Velké studie SNP

3000 zdravých jedinců

2000 pacientů	bipolar disorder	(1 SNP)	
„	coronary artery disease	(1 SNP)	
„	Crohn's disease	(9 SNPs)	
„	hypertension		
„	rheumatoid arthritis	(3 SNPs)	
„	type 1 diabetes	(1 SNP)	
„	type 2 diabetes	(3 SNPs)	P value < 5x10 ⁻⁷

Studovali 500 000 SNPs pomocí Affymetrix microarrays

Wellcome Trust Case control Consortium. **Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls.** *Nature*. 2007 Jun 7;447(7145):661-78.

Odchytky od referenčního genomu větší než 1kb

ještě v roce 2003 se myslelo, že většina „zdravých“ lidí se od referenčního genomu liší velmi nepatrně (SNPs, mikrosatelity)

array komparativní genomická hybridizace odhalila mnoho větších oblastí DNA, které se u zdravých lidí vyskytují v různém počtu

Copy number variation

DNA segment (většinou větší než 1 kb), který se u daného jedince vyskytuje v jiném počtu kopií než v referenčním lidském genomu

existuje mnoho takových oblastí v genomu (řádově tisíce)

“Database of Genomic Variants”

<http://projects.tcag.ca/variation/>

Copy number polymorphism – výskyt u více než 1% jedinců dané populace

Využití HapMap kolekce ke studiu copy number variant
všichni jedinci v této kolekci byli zdraví, přesto se našlo velké množství oblastí DNA (12% genomu), které se u těchto lidí nacházejí v různém počtu kopií

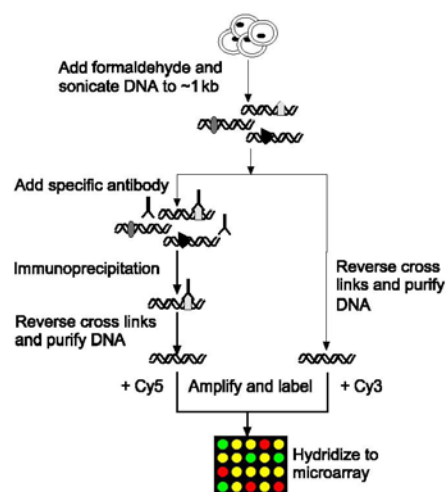
Copy number variation

hledání fenotypových projevů CNV („neškodná“ genomová varianta nebo příčina nemoci???)

CNV: pathogenic x benign x unknown clinical significance

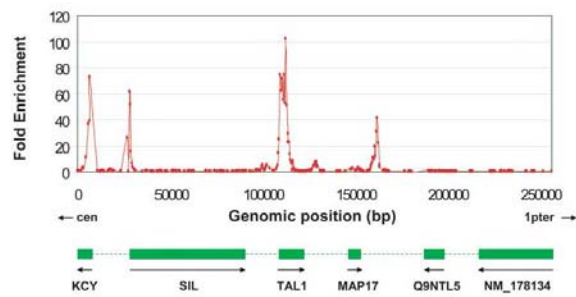
vnáší „zmatek“ do experimentů, které např. hledají příčinu vrozených genetických poruch (mentální opožděnost, vývojové odchylky)

Chromatin Immunoprecipitation on chip ChIP-Chip



Nalezení vazebného místa

256 kb oblast 1p32 pokrytá překrývajícími se PCR produkty (~400 bp)
protilátka: trimethylace histonu H3 Lys4



Carter and Vetrie 2004 Human Mol Genet