

# Kapitola VII

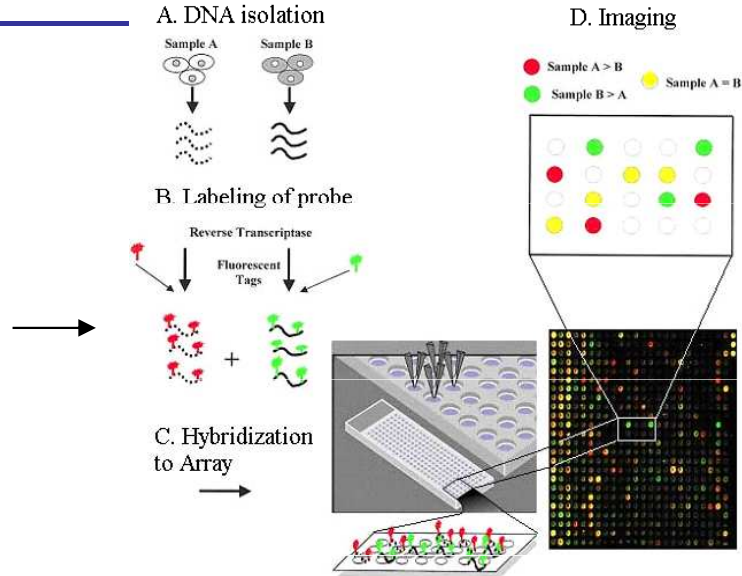
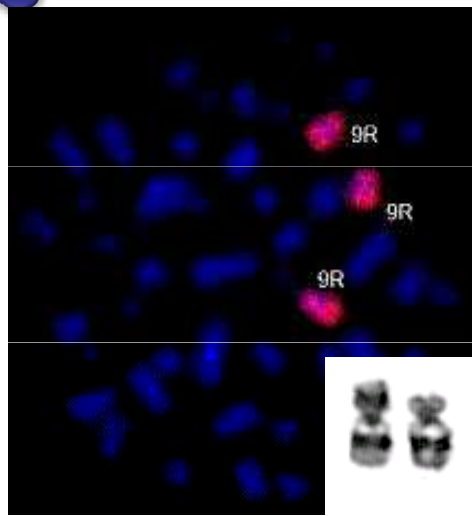
---

## Analýza array CGH

# arrayCGH

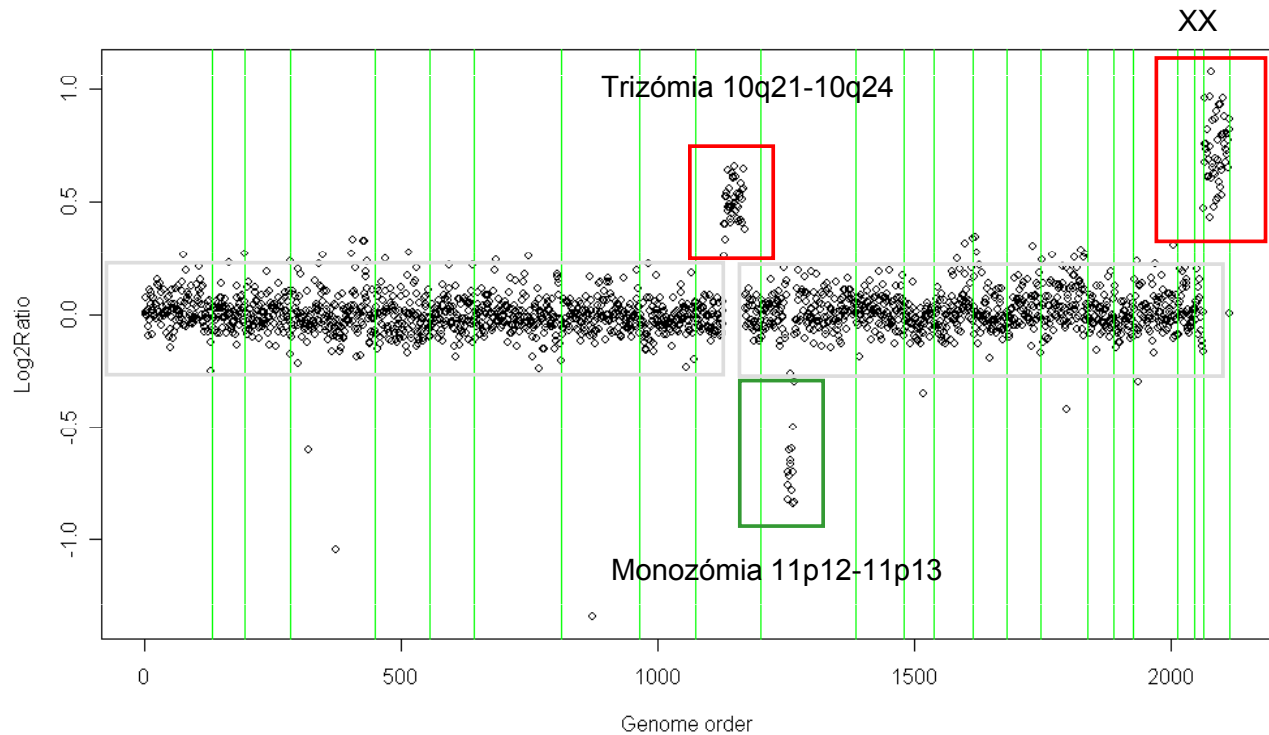
- Typ cDNA microarray
- Založený na CGH – comparatívnej genómovej hybridizácii
- Porovnáva DNA dvoch vzoriek
- Cieľ: zistiť zmeny v počte kópií génov medzi vzorkami

# arrayCGH - princíp



$$\log_2 \text{ratio} = \log_2 \frac{\text{Cy5}}{\text{Cy3}}$$

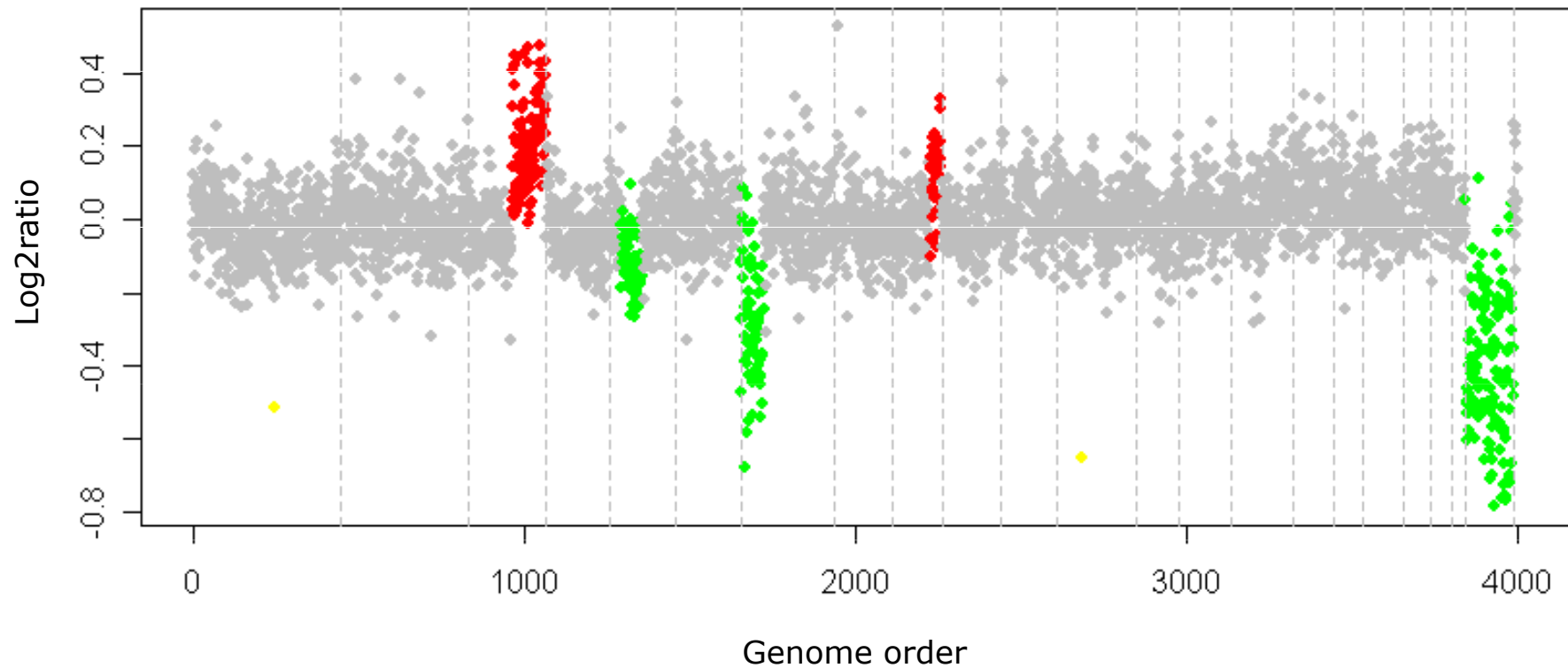
- Delecia ~ strata genetickej informácie (monozómia)
- Amplifikácia ~ zisk genetickej informácie (trizómia)
- SNPs – jednonukleotidové polymorfizmy



# Analýza arrayCGH dát I.

amplifikácia =  $\log_2(3/2)=0.58$

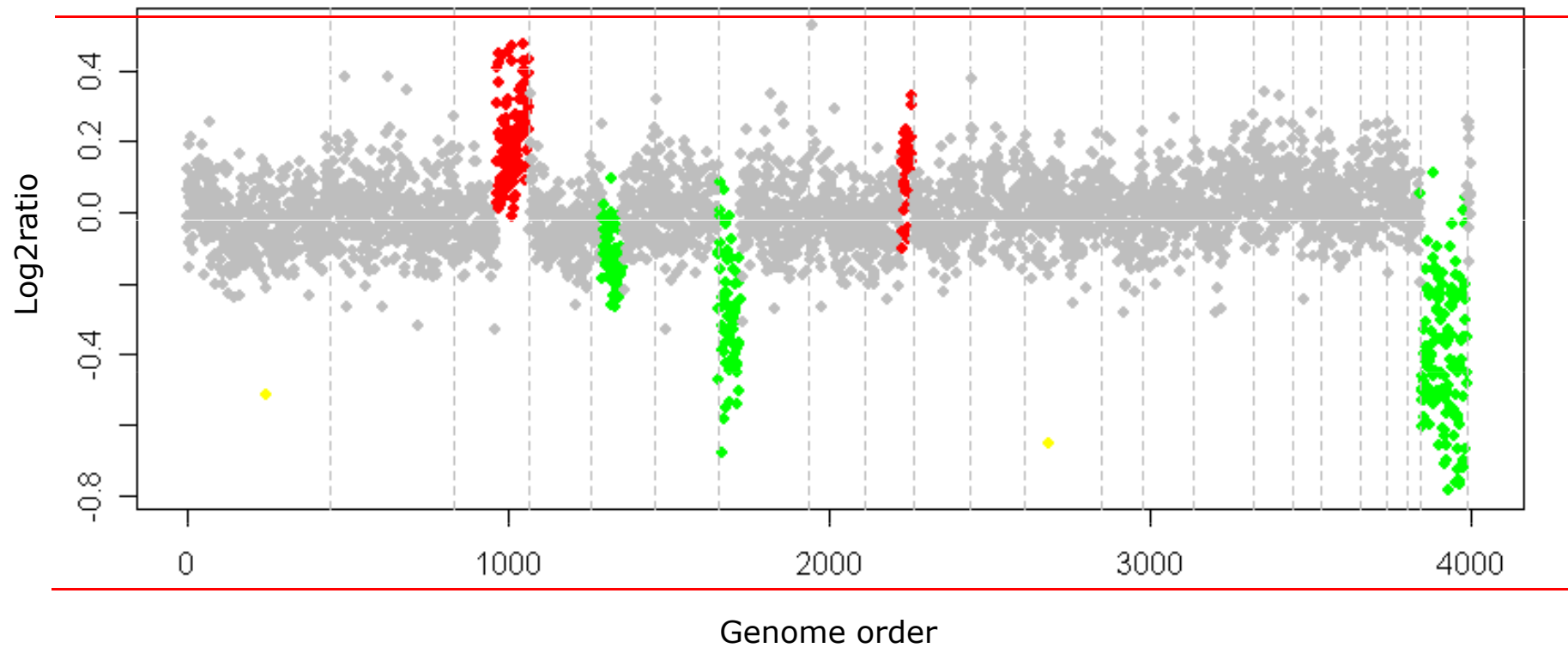
delecia =  $\log_2(1/2)=-1$



# Analýza arrayCGH dát I.

amplifikácia =  $\log_2(3/2)=0.58$

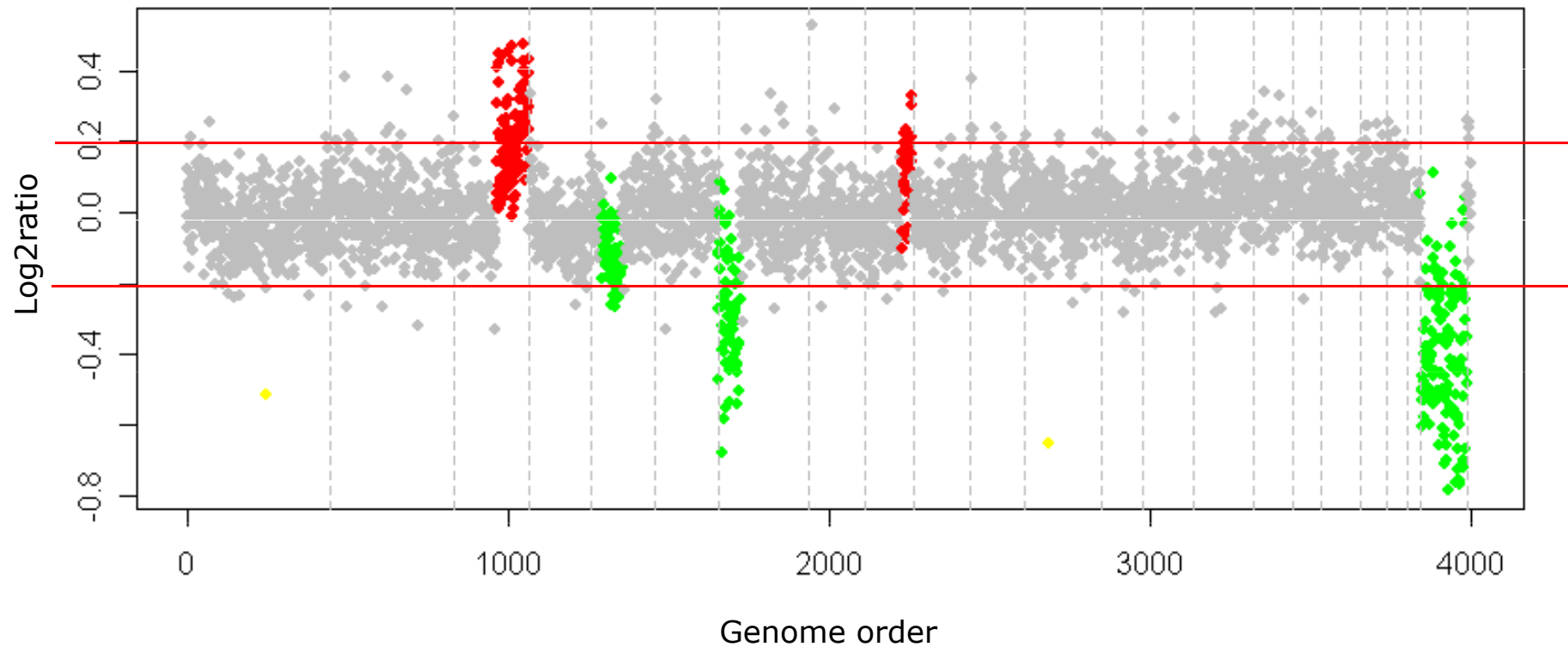
delecia =  $\log_2(1/2)=-1$



# Analýza arrayCGH dát I.

amplifikácia =  $\log_2(3/2)=0.58$

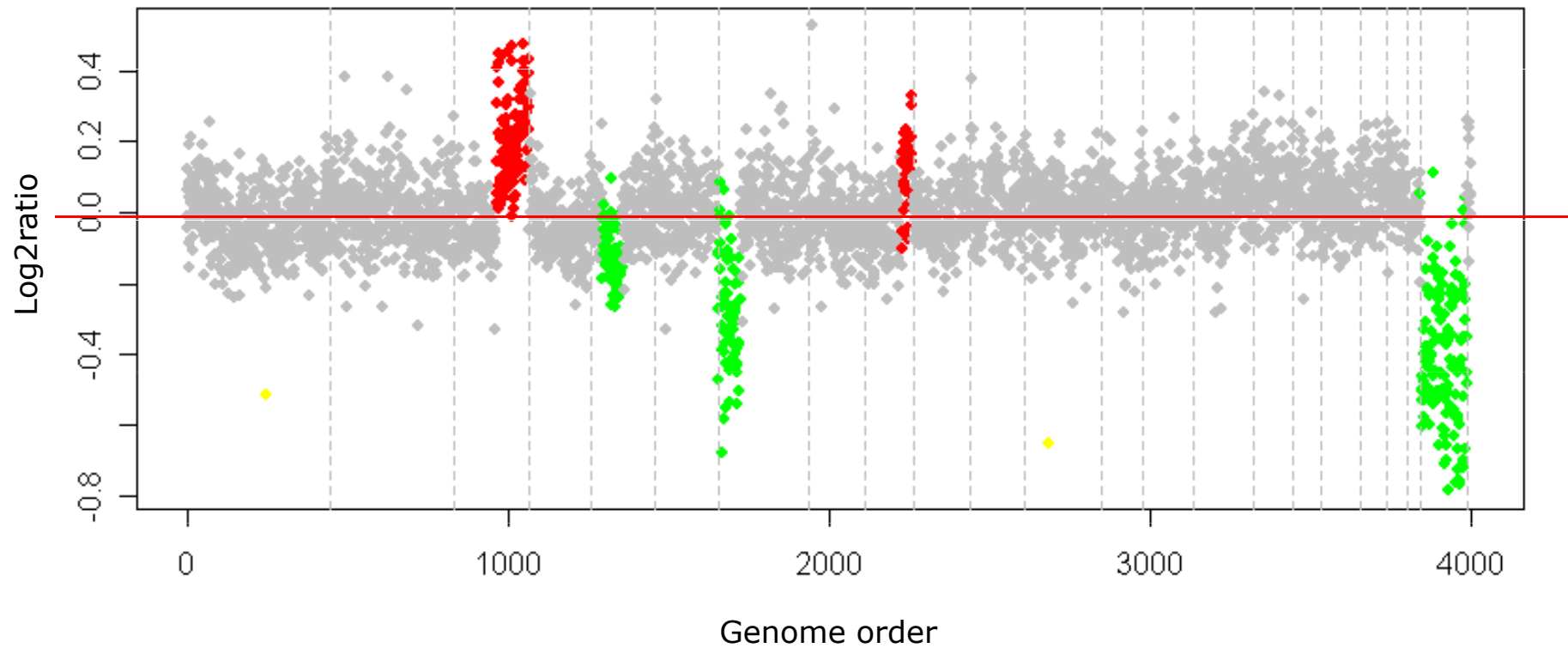
delecia =  $\log_2(1/2)=-1$



# Analýza arrayCGH dát I.

amplifikácia =  $\log_2(3/2)=0.58$

delecia =  $\log_2(1/2)=-1$



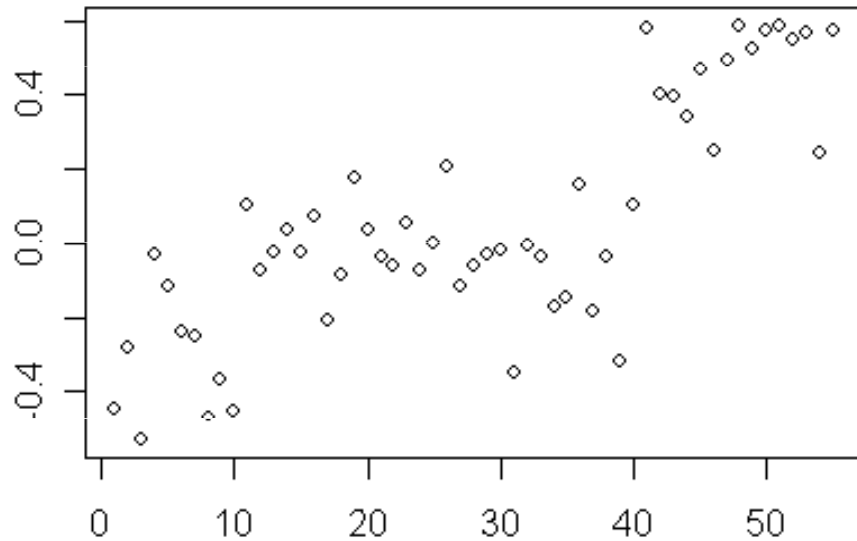
# Rozdelenie metód pre analýzu arrayCGH

- Väčšina metód pre analýzu arrayCGH využíva fakt, že v prípade delecií alebo amplifikácií sú obvykle zasiahnuté väčšie časti genómu (hlavne u nádorov) – závislosť na poradí
- Sú založené na
  - *Segmentácii* (detekcia zlomov)
    - Vytvárajú dobre odlišiteľné segmenty s odlišnými priemernými  $\log_2$ ratio
    - Klony v danom segmente zdieľajú rovnaký počet kópií
  - *Zhlukovaní*
    - Priame zhlukovanie segmentov do skupín (informácia o poradí inkorporovaná do metriky vzdialeností)
  - *Skrytých Markovových reťazcoch (HMM)*
    - Skryté stavy: úrovne zmeny
  - *Vyhladzovaní*
    - Regresné metódy, snažia sa preložiť grafom čo najviac optimálnu krivku

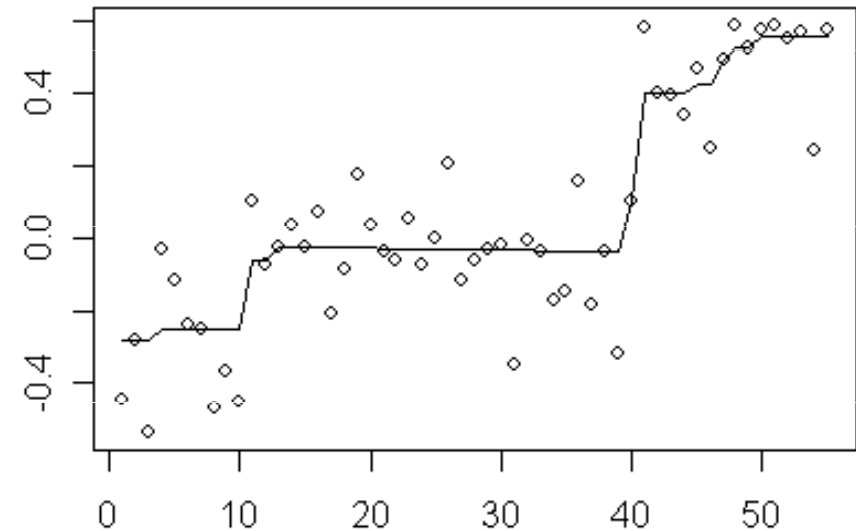


# Princíp segmentačných metód

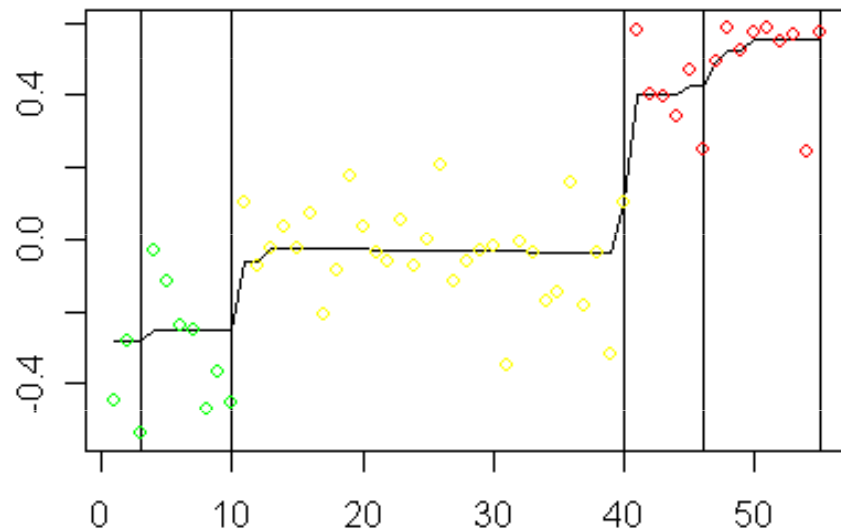
1. DATASET



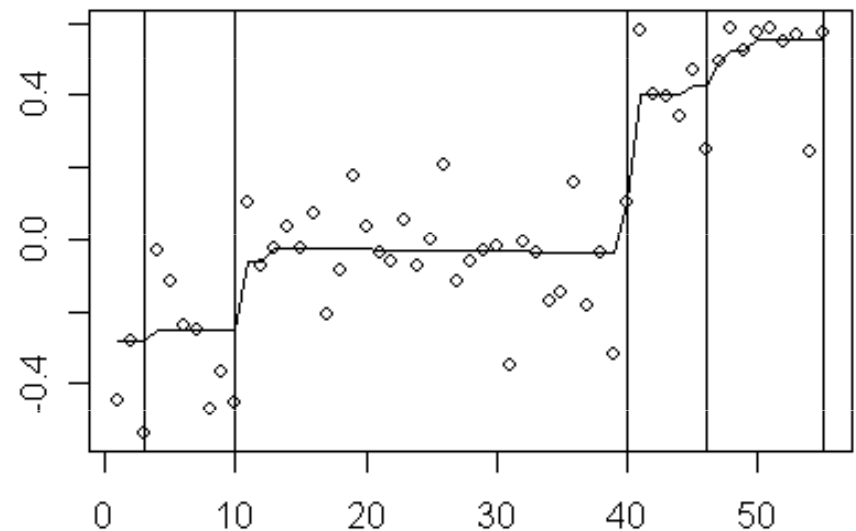
2. FILTERING, SMOOTHING



4. REGION ASSESSMENT



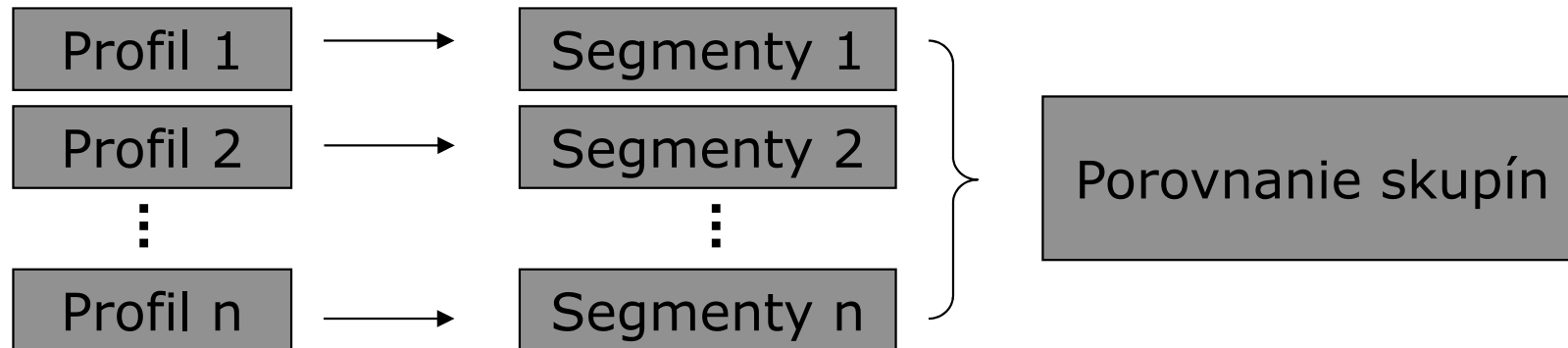
3. BREAKPOINTS DETECTION



# Prístupy k analýze

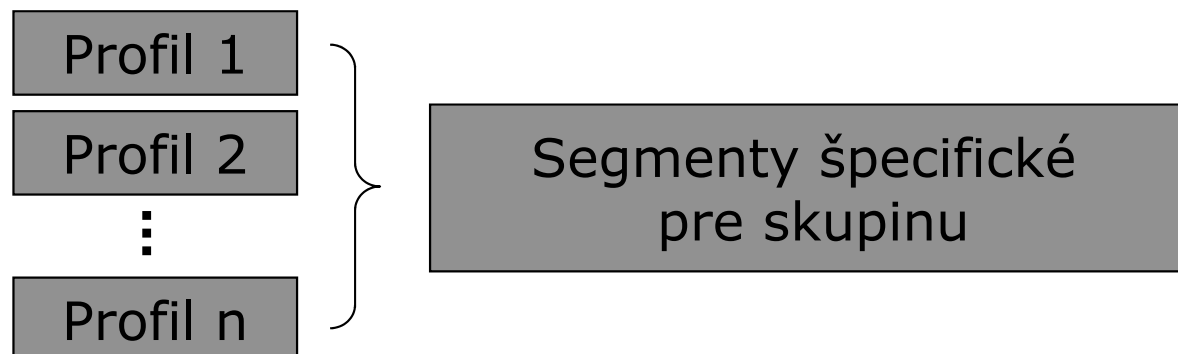
- Jednoprofilové

- Aberácie sa zisťujú u každej vzorky individuálne



- Viacprofilové

- Zisťujeme aberácie ktoré sú špecifické pre danú skupinu



# “Ideálna” metóda pre analýzu arrayCGH

- Celková dobrá výkonnosť
- Dobrá výkonnosť i na zašumených dátach
- Schopná detekovať krátke aberácie
- Výpočtovo nenáročná
- Zrozumiteľny algoritmus (dôležité pre nastavenie parametrov)

# Porovnanie metód I.

- Lai *et al.* (2005), Willenbrock & Fridlyand (2007),...

## 1. Segmentačné metódy

+ Celková dobrá výkonnosť

- Nie pre malé regióny a zašumené dáta

- Detekované segmenty sa musia ešte rozdeliť do biologicky relevantných regiónov

## 2. Vyhľadzovacie metódy

+ Dobrý výkon na zašumených dátach a malých regiónoch

- Výsledkom je len vyhladená krivka

- Nevhodné pre detekciu väčších zmien

# Porovnanie metód II.

## 3. Metódy založené na skrytých Markovových reťazcoch

- + Dobrá výkonnosť len na nezašumených dátach
- Nie pre malé regióny a zašumené dáta
- Výpočtovo náročné

## 4. Zhukovacie metódy

- + Jednoduché na pochopenie
- Neefektívne

# Analýza arrayCGH v R

- aCGH  
(<http://www.bioconductor.org/packages/devel/bioc/html/aCGH.html>)
- GLAD
- DNACopy
- cghMCR
- BioHMM

# Úloha [1 bod]

Priradte k detekovaným segmentom vzorky “05296” z dat coriell oblasti podľa cytogenetického pruhovania podľa dát cytoBand. Uvedte všetky pruhy, do ktorých segment zasahuje v tvare CHR<číslo chromozomu>\_označenie pruhu.