

Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů

# IV107 Bioinformatika I

## Přednáška 8

Katedra informačních technologií  
Masarykova Univerzita Brno

Podzim 2019

# Cesta k údajům o expresi

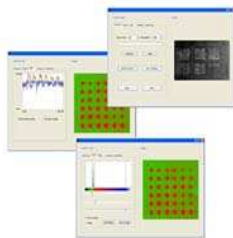
IV107 Bioinformatika I -  
Přednáška 8

Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů





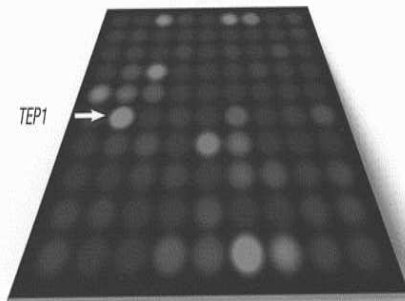
Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů

A scanner measures the fluorescence of each dye separately.



The higher ratio of red to green fluorescence at the spot labeled *TEP1*, for instance, shows that the *TEP1* gene is expressed more highly in sporulating than in budding yeast cells.

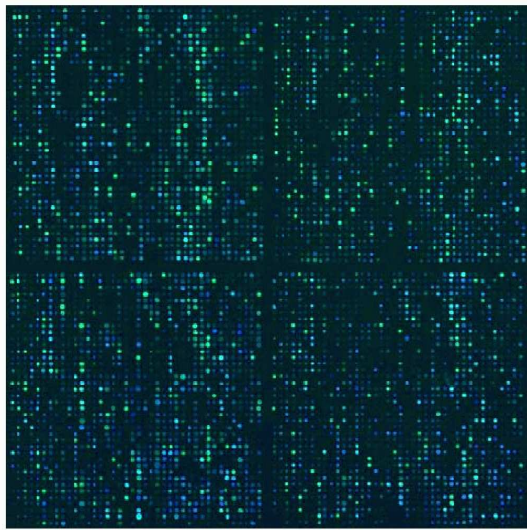
# Nasnímaný obraz z DNA čipu vypovídá o přítomnosti DNA určité sekvence ve vzorku

Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů





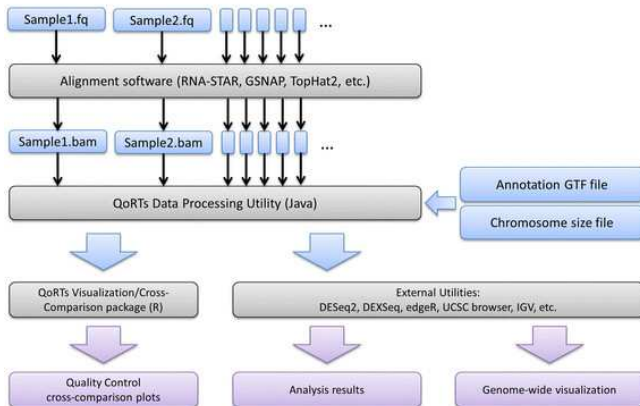
Získávání expresních dat

Uplatnění v medicíně

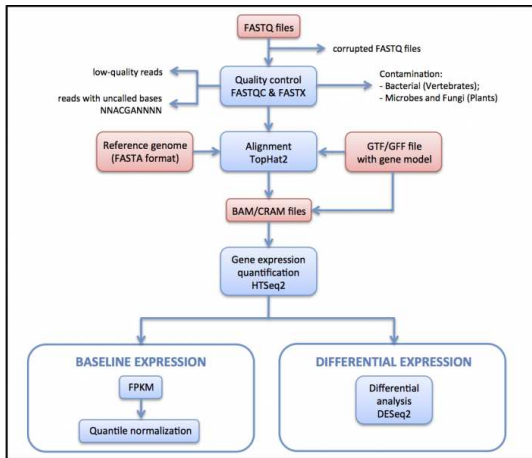
Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů

## Analysis Pipeline



# Postup při zpracování RNA-seq dat



Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů



Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů

**Rakovina** nekontrolované dělení buněk v důsledku mutace (onkogeny, supresory nádorů).

Měření genové exprese nachází uplatnění v diagnostice rakoviny.

- ▶ o jaký druh rakoviny se jedná ?
- ▶ jaká je prognóza pro danou diagnózu ?

Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů

- ▶ Myelogenní, lymfocytická, akutní, chronická ?
- ▶ Acute Myelogenous Leukemia (AML) Acute Lymphocytic Leukemia (ALL)
- ▶ Není produkováno dostatečné množství krvinek
- ▶ Anemie → slabost, infekce, snížená srážlivost krve
- ▶ AML: 10,000 případů ročně
- ▶ ALL: 3,500 dospělí / 2,400 novorozenci
- ▶ AML nebo ALL ?

- ▶ **Diagnóza ve specializované laboratoři**
  - ▶ ALL: 58
  - ▶ AML: 14
- ▶ **Léčba: chemoterapie, transplantace kostní dřeně**
  - ▶ ALL: corticosteroidy, vincristin, methotrexat, L-asparaginasa
  - ▶ AML: daunorubicin, cytarabin
- ▶ **Správná diagnóza je velice důležitá pro nasazení správné léčby**
- ▶ **Microarray může podporovat rozhodování**

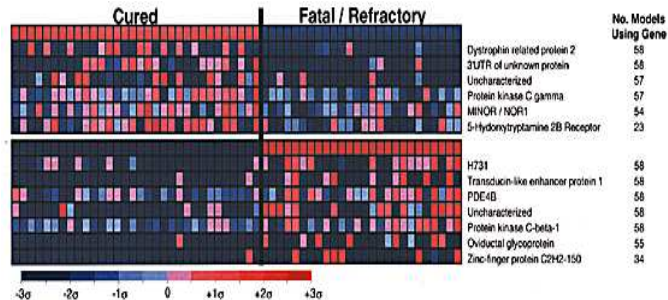
# Příklad sady genů, které korelují s prognózou

Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů



- ▶ měření s nízkým počtem opakování
- ▶ rozpoznávání obrazu
- ▶ převod na matici číselných hodnot
- ▶ normalizace
- ▶ statistika jednotlivých pozic a genů (průměrná hodnota, rozptyl, eliminace extrémních hodnot, atd.)
- ▶ kontrast: dvě sady experimentálních podmínek
- ▶ specifika:  $n \gg p$
- ▶ PCA, LDA, shluková analýza (clustering), strojové učení
- ▶ analýza v kontextu funkce (GO, KEGG)
- ▶ analýza v kontextu sekvence (blízké motivy a geny, zvláště promotor)
- ▶ analýza časových a prostorových řad (koregulované geny)

Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů

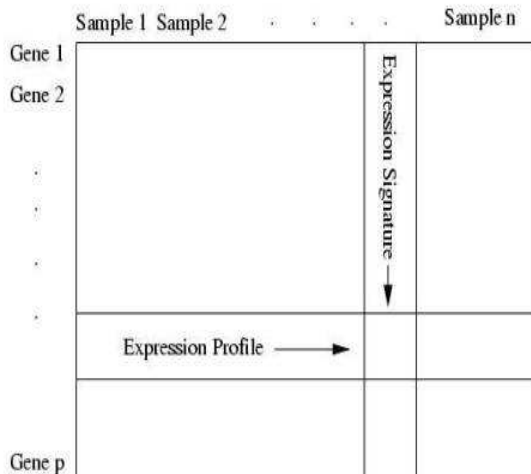
# Co je to expresní profil?

Získávání expresních dat

Uplatnění v medicíně

**Analýza expresních dat**

Příklad použití microarray dat  
v klasifikaci nádorů

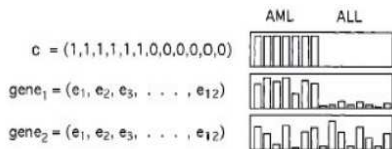


# Jak zjišťujeme podobnost dvou expresních profilů či jejich schopnost klasifikovat vzorek?

Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů

Expresní profil je chápán jako množina bodů v multidimenzionálním prostoru, popř. jako sada postupných výběrů z populace genů či zdroj informací o vzorku

$c \cdot gene_1$  skalární součin (cos uhlu, který zvírají)

$d(c, gene)$  euklidovská vzdálenost (postihuje i absolutní hodnoty exprese)

$corr(c, gene)$  korelace

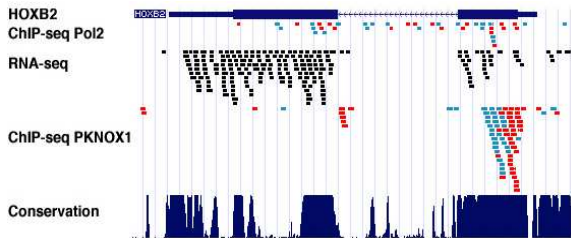
$MI(c, gene)$  vzájemná informace

Získávání expresních dat

Uplatnění v medicíně

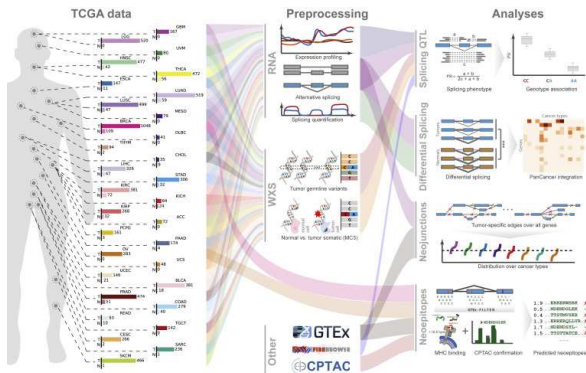
Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů



ChIP-seq and RNA-seq data exemplified at the HOXB2 gene



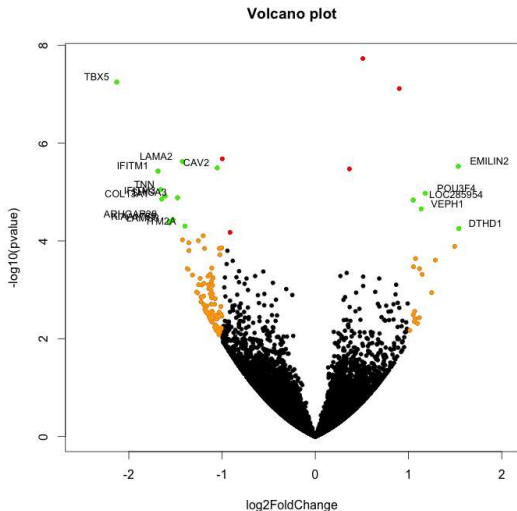


Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů

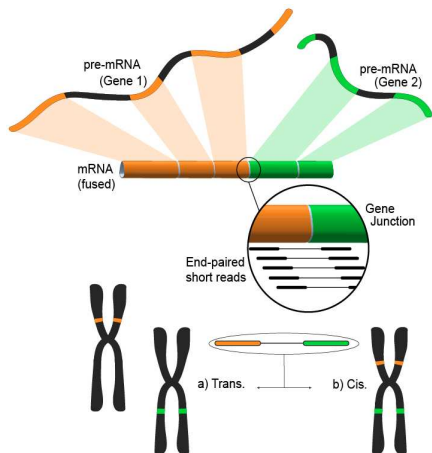


Získávání expresních dat

Uplatnění v medicíně

**Analýza expresních dat**

Příklad použití microarray dat  
v klasifikaci nádorů



Získávání expresních dat

Uplatnění v medicíně

**Analýza expresních dat**

Příklad použití microarray dat  
v klasifikaci nádorů

# In-situ hybridization



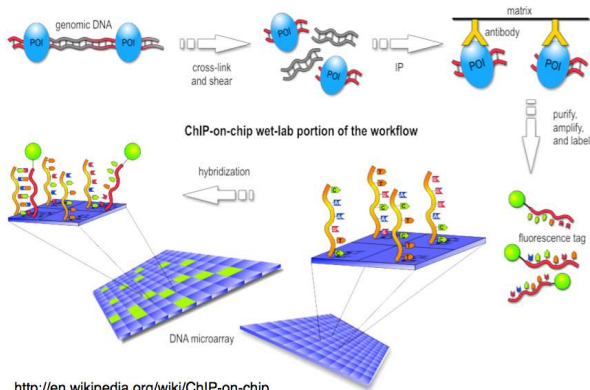
**In situ hybridization of  
Acs15 in mouse E14.5**

Získávání expresních dat

Uplatnění v medicíně

**Analýza expresních dat**

Příklad použití microarray dat  
v klasifikaci nádorů



<http://en.wikipedia.org/wiki/ChIP-on-chip>

Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

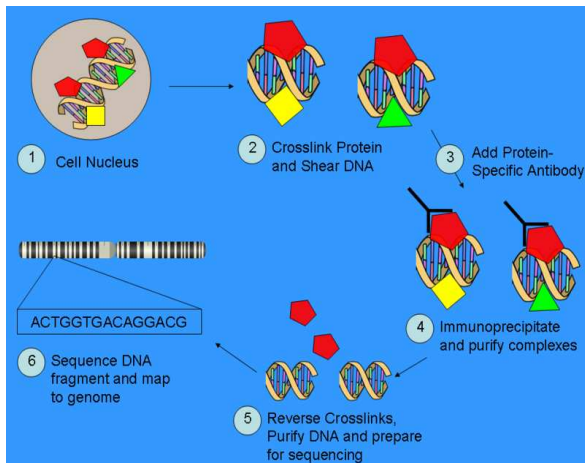
Příklad použití microarray dat  
v klasifikaci nádorů

Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů



# Multiclass cancer diagnosis using tumor gene expression signatures (Ramdaswamy et al., 2001)

Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů

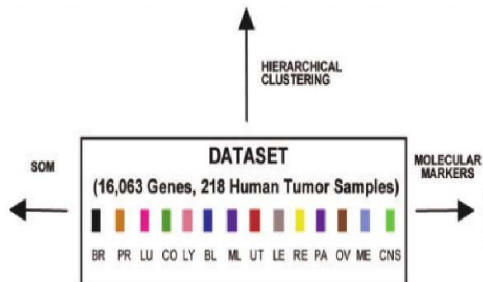
- ▶ V praxi se k diagnostice používá soubor klinických a histologických vyšetření.
- ▶ Je možné je nahradit molekulárním vyšetřením ?
  - ▶ 218 nádorů a 90 zdravých tkání
  - ▶ microarray (exprese genů)
  - ▶ 5 000 000 zjištěných hodnot
  - ▶ metody strojového učení (CLUSTER, SOM, SVM)
  - ▶ klasifikace (OVA = one v. all)
  - ▶ úspěšnost 78%

Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů





# Hierarichické aglomerativí zhlukování

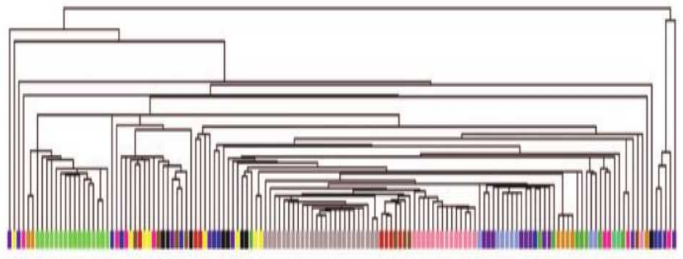
IV107 Bioinformatika I -  
Přednáška 8

Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

**Příklad použití microarray dat  
v klasifikaci nádorů**



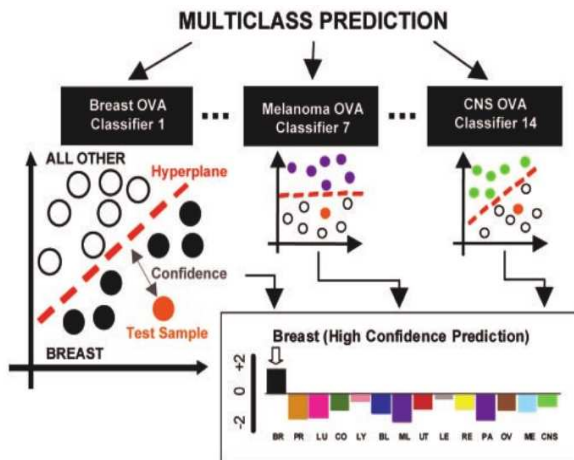
# Jak funguje klasifikace do více tříd

Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů



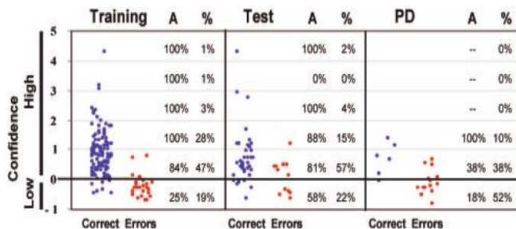
# Kvalita predikcí s různými stupni spolehlivosti

Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů



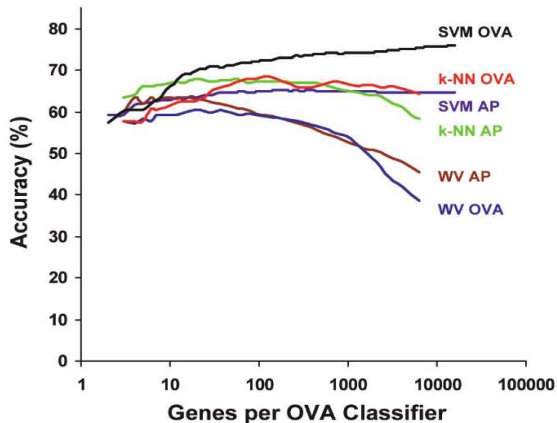
# Počet genů potřebný k úspěšné klasifikaci

Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů



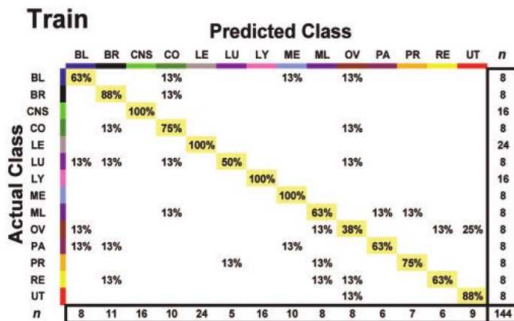
# Spolehlivost klasifikace

Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

Příklad použití microarray dat  
v klasifikaci nádorů



Získávání expresních dat

Uplatnění v medicíně

Analýza expresních dat

**Příklad použití microarray dat  
v klasifikaci nádorů**