



**LÉKAŘSKÁ FAKULTA MASARYKOVY UNIVERSITY**  
**Interní hematoonkologická klinika LF MU a FN Brno**  
**Centrum molekulární biologie a genové terapie**



# Moderní metody analýzy genomu

## Aplikace III

5.5. 2011

Boris Tichý



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Tato prezentace je spolufinancována  
Evropským sociálním fondem  
a státním rozpočtem České republiky

# RNA-Seq

Transcriptome sequencing

Single-end – kvantifikace

Paired-end – struktura transkriptů

Tag sequencing

3' tagy, kvantifikace, bez informace o struktuře

Degradome sequencing

5' tagy, identifikace cílů microRNA

Small RNA sequencing

MicroRNA kvantifikace i de-novo identifikace

SAGE (Serial Analysis of Gene Expression)

Konkatemery tagů, původně Sanger sekvenace

RIP (RNA ImmunoPrecipitation)

Imunoprecipitace, RNA vázající proteiny

## RNA-Seq

Oligo-dT vs. hexamer priming reverzní transkripce

3', 5' end reprezentace

PolyA selekce vs. rRNA deplece

Zastoupení non-polyA (non-coding) RNA

Vysoce abundantní geny

Až 75% transkriptů ← 5% exprimovaných genů

Možnost detekce fúzních genů

Speciální algoritmy

Mapování readů na exon-exon rozhraní

## Aplikace nových technologií

### Epigenomika/epigenetika

In biology, and specifically genetics, epigenetics is the study of heritable changes in phenotype (appearance) or gene expression caused by mechanisms other than changes in the underlying DNA sequence, hence the name epi- (Greek: επί- over, above) -genetics. These changes may remain through cell divisions for the remainder of the cell's life and may also last for multiple generations. However, there is no change in the underlying DNA sequence of the organism;<sup>[1]</sup> instead, non-genetic factors cause the organism's genes to behave (or "express themselves") differently.

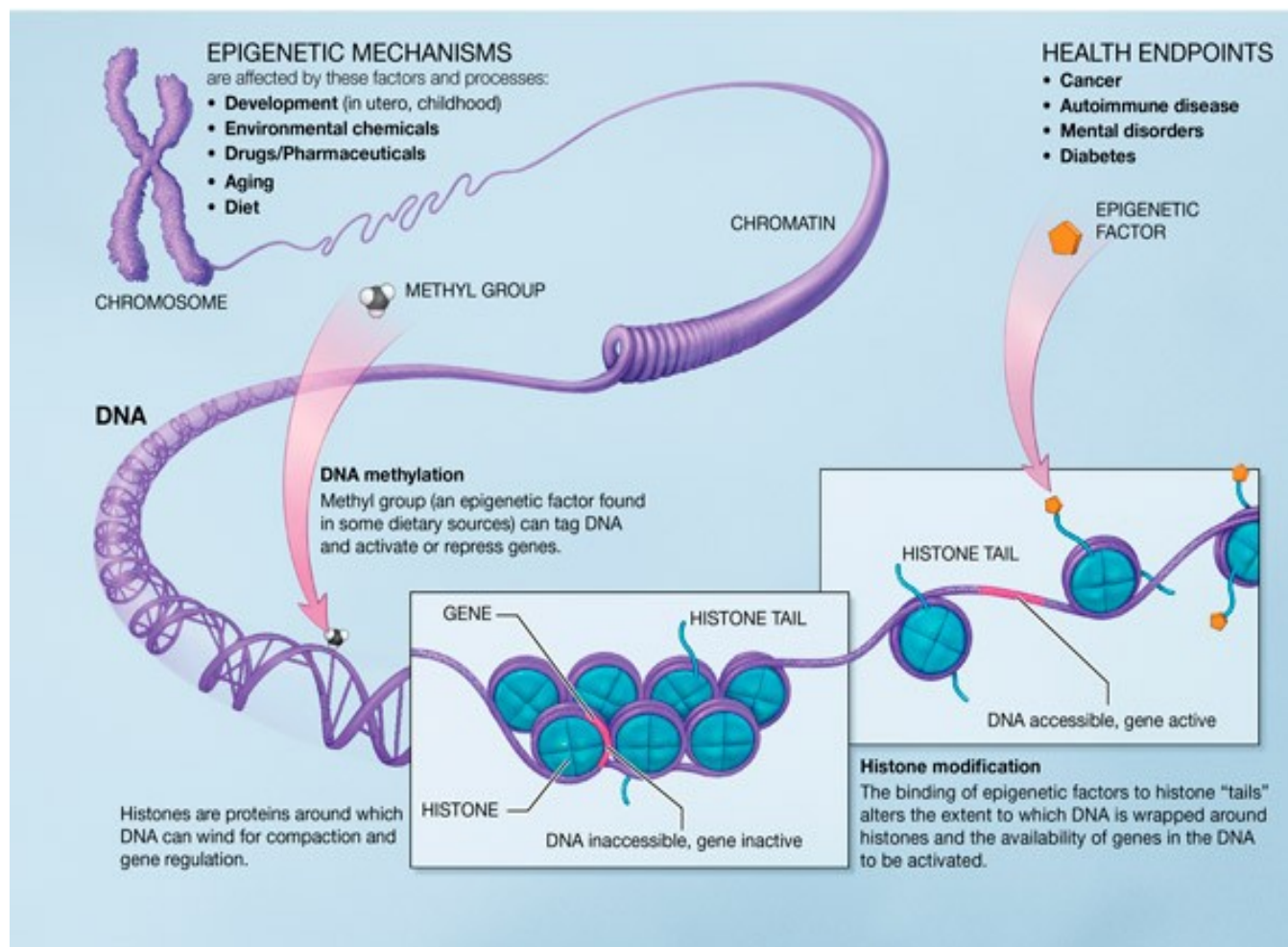
# Epigenetika

DNA metylace

C → Met-C, snížená exprese

Modifikace histonů

Aktivní | neaktivní chromatin



# Chromatinová imunoprecipitace

Modifikované histony

Acetylované, metylované

Další DNA vázající proteiny

Transkripční faktory

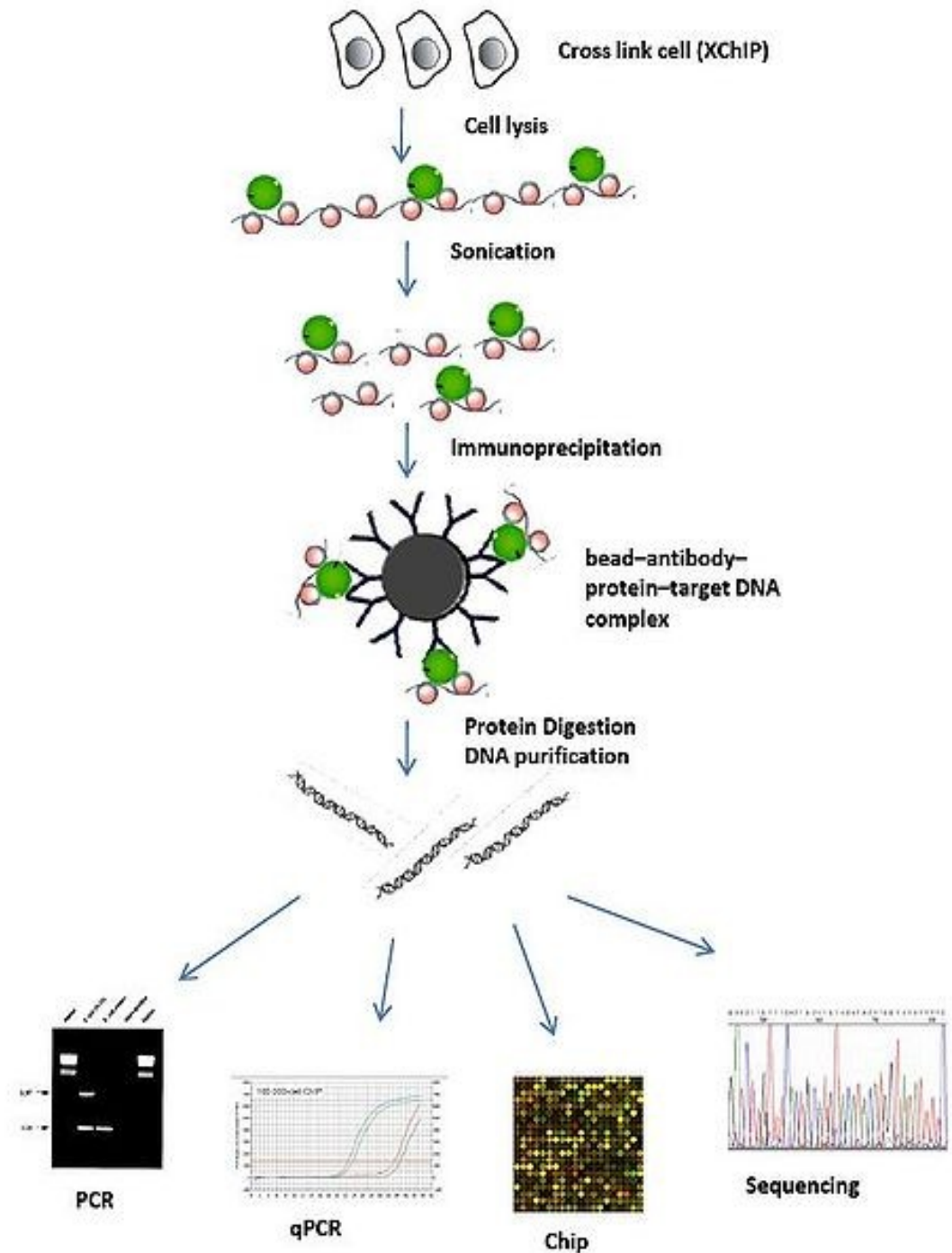
RNA polymerázy

Metylovaná DNA

Kvalita protilátky

Změna epitopu

formaldehyd



## Methylace DNA

MeDIP - Immunoprecipitace metylované DNA

Protilátka rozpoznávající Met-C

Pozice metylovaných úseků DNA

Bisulfite treatment

Konverze C → U, Met-C se nemění

Přesná identifikace jednotlivých metylovaných bazí

Některé dědičné choroby, nádory

## Metagenomika

Sekvenace mikrobiálních populací

Informace I o nekultivovatelných organismech

Identifikace nových genů, nových vlastností

Půdní, vodní, střevní bakterie apod.



# Analýza NGS dat

Assembling – vytvoření kontigů

De-novo

Mapování na referenční sekvenci

Identifikace variant

SNP (SNV)

In/del

Strukturní aberace – chromosomy, transkripty

Kvantifikace

DNA – amplifikace/delece, místa vazby transkripčních faktorů

RNA – tagy, exony, transkripty

# Analýza NGS dat

## Integrace dat

### Databáze

Protein-protein interakce

Ontologie – GO (GeneOntology), MeSH terms

Transkripční faktory, microRNA targets

Expresní profily – GSEA, profily chemických látek

### Různé typy dat

ChIP + RNA + metylace + DNA copy number

# Analýza NGS dat

## Statistická analýza

“klasické” statistické metody

T-test, ANOVA, neparametrické testy

Vícerozměrné analýzy

Clustering (HCL, k-means), PCA

Klasifikační algoritmy

Lineární (LDA), nelineární (KNN)

Vazba s dalšími informacemi (analýza anotací)

Over-reprezentace

Sítě