

# IV107 Bioinformatika I

## Přednáška 3

Katedra informačních technologií  
Masarykova Univerzita Brno

Podzim 2023



# Před týdnem

DNA nepřímo (přes proteiny) řídí většinu procesů v buňce. Její struktura vykazuje všechny potřebné rysy. Je nositelem genů.

## Exprese genů

- ▶ transkripce
- ▶ translace

## Struktura proteinů

- ▶ primární (sekvence aminokyselin)
- ▶ sekundární ( $\alpha$ ,  $\beta$ , etc.)
- ▶ terciární (3-D)
- ▶ kvartérní (makromolekulární komplexy)



### [Hemoglobin](#) Chime

An introductory presentation suitable for lectures or individual study.

[\(Deutsch, Português\)](#)



### [DNA Structure](#) Chime

An introductory level, nonlinear self-paced tutorial. [\(Deutsch, Español, Português\)](#)

'Win-IE6-Compatible'



### [Antibody](#) Chime

An introductory presentation suitable for lectures or individual study.

[\(Deutsch\)](#)



### [MHC](#) Chime

The Major Histocompatibility Complex presents peptides from foreign proteins to T lymphocytes, crucial to disease immunity.



### [Lipid Bilayers and Membrane Channel](#) Chime

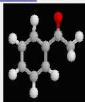
Introduces cholesterol and phospholipids, then proceeds to lipid bilayer and the gramicidin membrane channel embedded within the bilayer. Includes molecular dynamics simulations of both gel and fluid membrane states. By Eric Martz and Angel Herráez. [\(Español\)](#)



### [Molecular Vibrations: IR Spectrum](#) Chime

by Mołyka, Lahti & Lancashire.

'Win-IE6-Compatible'



### [Water](#) Chime

Theoretical simulation of 10 water molecules condensing into a hydrogen-bonded droplet. Includes challenge questions for students.

'Win-IE6-Compatible'



### [Lac repressor](#) Chime

bending the DNA operon as it goes from nonspecific to specific DNA binding.

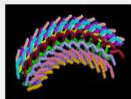
'Win-IE6-Compatible'



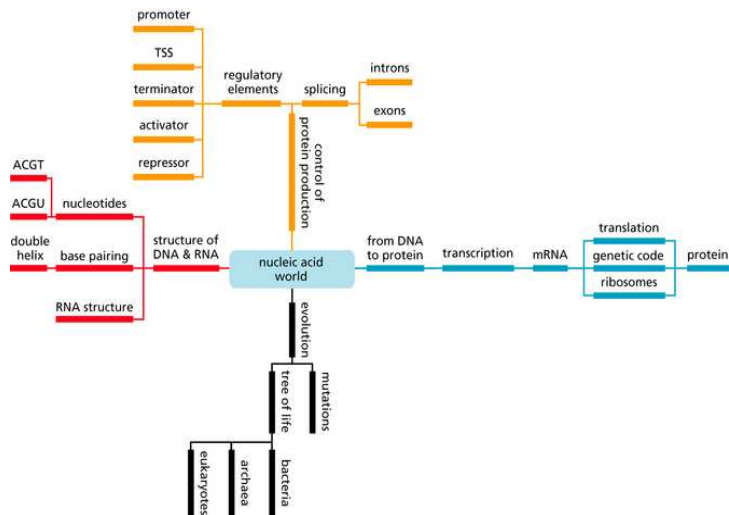
### [Bacterial Flagellar Hook](#) Chime

Bacterial flagella are whip-like organelles that bacteria use to swim about. The hook acts as a molecular universal joint. It's composed of over 100 protein subunits.

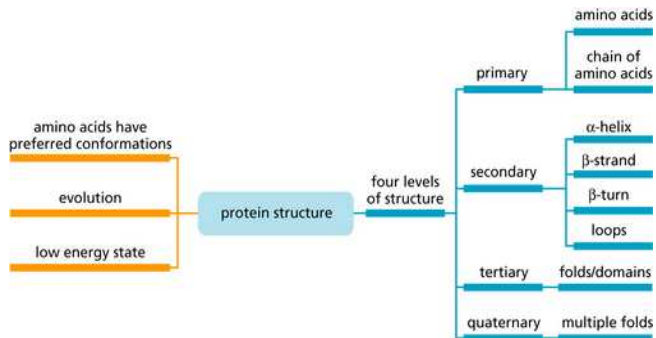
'Win-IE6-Compatible'



# Schéma 1



# Schéma 2



# Outline



# Techniky manipulující DNA a RNA

- ▶ Rekombinace DNA a klonování
  - ▶ izolace z buněk
  - ▶ štěpení restrikcími endonukleázami
  - ▶ rekombinace (ligace)
  - ▶ transformace organismů rekombinantní DNA
- ▶ Syntéza, kopírování a amplifikace DNA
  - ▶ syntéza oligonukleotidů
  - ▶ transkripce in vitro
  - ▶ syntéza cDNA
  - ▶ amplifikace pomocí PCR
- ▶ Editace DNA
  - ▶ Cas9-CRISPR
  - ▶ TALEN
  - ▶ ZF nukleázy
- ▶ Detekce, analýza a sekvenace
  - ▶ elektroforéza (dělení podle velikosti)
  - ▶ hybridizace se značenými sondami
  - ▶ určování sekvence

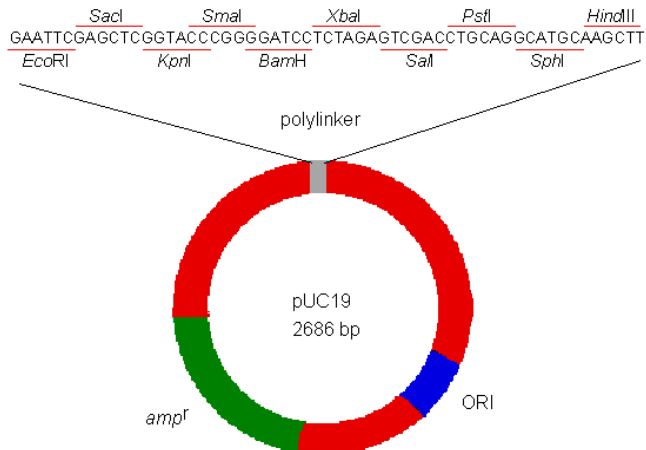
# Restrikční endonukleázy

enzym	zdroj	restrikční místo	fragment[kbp]
AluI	<i>Arthrobacter luteus</i>	AG↓CT TC↑GA	0.3
BamHI	<i>Bacillus amyloliquefaciens</i> H	G↓GATC C C CTAG↑G	7.0
EcoRI	<i>Escherischia coli</i> R	G↓AATT C C TTAA↑G	3.1
HaeIII	<i>Haemophilus aegyptus</i>	GG↓CC CC↑GG	0.6
NotI	<i>Norcadia otitidis-caviarum</i>	GC↓GGCC GC CG CCGG↑CG	<9700
PstI	<i>Providencia stuartii</i>	C TGCA↓G G↑ACGT C	7.0
TaqI	<i>Thermus aquaticus</i>	T↓CG A A GC↑T	1.4
HindIII	<i>Haemophilus influenzae</i> Rd	A↓AGCT T T TCGA↓A	3.1

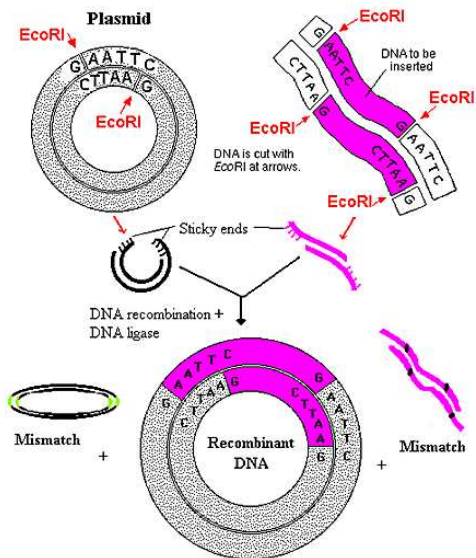




# Rekombinace a klonování

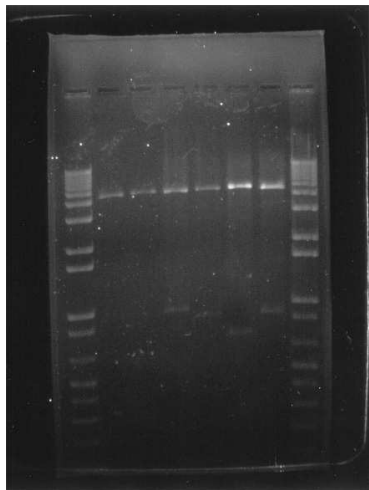


# Rekombinace DNA a klonování



Inserting a DNA Sample into a Plasmid

# Identifikace DNA na agaróзовém gelu



# V laboratoři



A

<http://www.ct.gov/dps/cwp/view.asp?a=2155&q=314998>



B

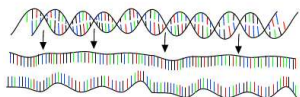
<http://www.sme.sk/c/3753490/Genetika-nadej-pre-pacientov.html>



# PCR (polymerázová řetězová reakce)

## PCR : Polymerase Chain Reaction

30 - 40 cycles of 3 steps :



Step 1 : denaturation

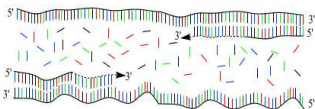
1 minut 94 °C



Step 2 : annealing

45 seconds 54 °C

forward and reverse primers !!!



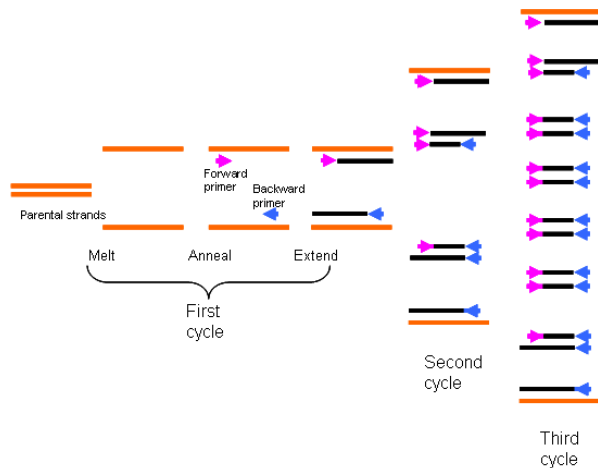
Step 3 : extension

2 minutes 72 °C  
only dNTP's

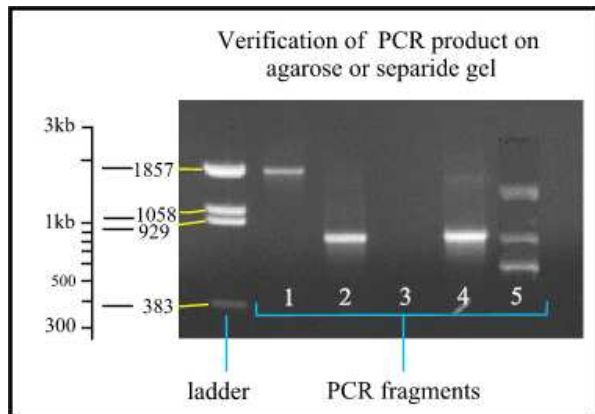
(Andy Veenstra 1999)



# PCR (polymerázová řetězová reakce)



# PCR (polymerázová řetězová reakce)

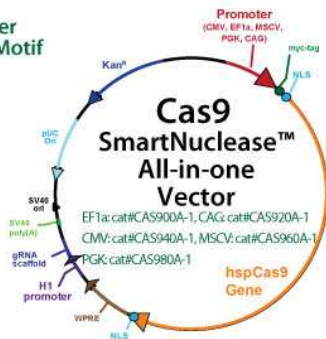
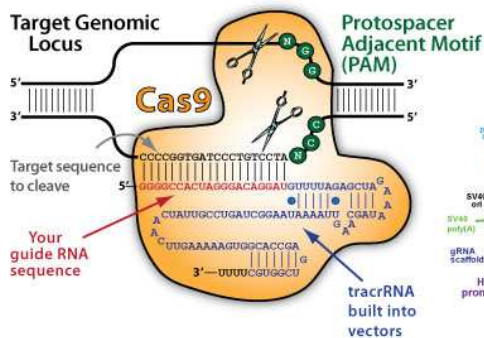


# PCR (polymerázová řetězová reakce)

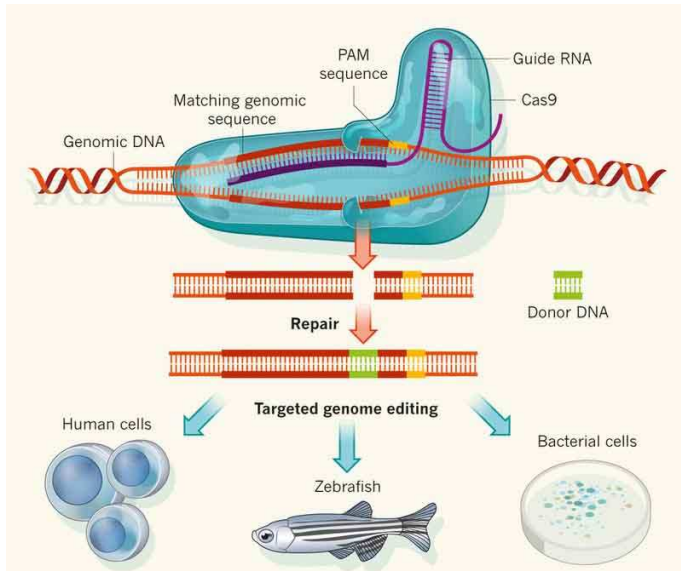




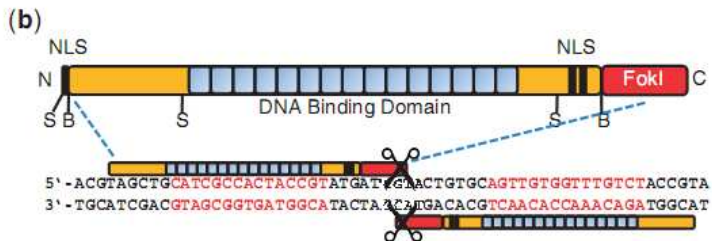
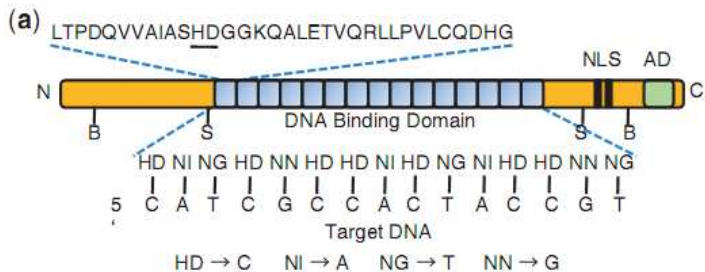
# Cas9-CRISPR



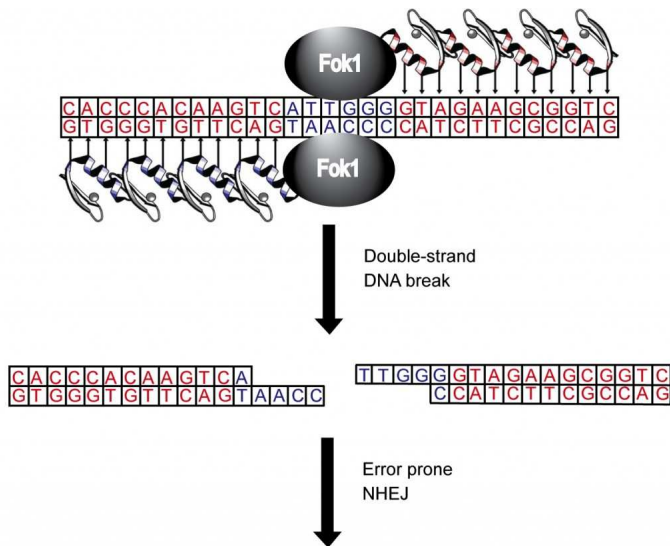
# Cas9-CRISPR



# TALEN



# ZF nukleázy



# Sekvence DNA

- ▶ Pořadí nukleotidů ve směru 5' → 3' uloženo pro budoucí generace jako dlouhý řetězec symbolů A, C, G a T
- ▶ Dobrá zpráva pro budoucnost bioinformatiků: neutichá zájem o sekvence
  - ▶ Personal Genomics Project (PGP) - několik tisíc genomů  
<https://www.personalgenomes.org/>
  - ▶ ERGA - genomy organismů, soustředění na ohrožené druhy <https://www.erga-biodiversity.eu/>
  - ▶ UK Biobank - genotyp 450000 osob  
<http://geneatlas.roslin.ed.ac.uk/>

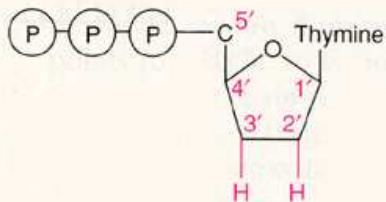


# Sekvence DNA

- ▶ Sekvence se dělá ve dvou fázích
  - ▶ Za použití DNA polymerázy a směsi normálních a modifikovaných deoxyribonukleotidů se syntetizují vlákna DNA (imitace replikace), templátem je sekvenovaná molekula DNA
  - ▶ Produkty polymerizace se v automatických strojích chromatograficky dělí podle velikosti a vypovídají o sekvenci
- ▶ Jedna reakce podá informaci o 500-1000 bazích



# Dideoxynukleotid neumožňuje navázání ďalších nukleotidů



Dideoxythymidine triphosphate (ddTTP)

# Polymeráza se zastaví při použití ddNTP

DNA Polymerase reads the template strand and synthesizes a new second strand to match:



IF 5% of the T nucleotides are actually dideoxy T, then each strand will terminate when it gets a ddT on its growing end:

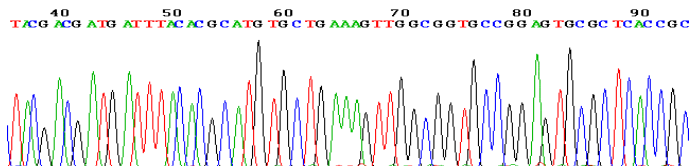




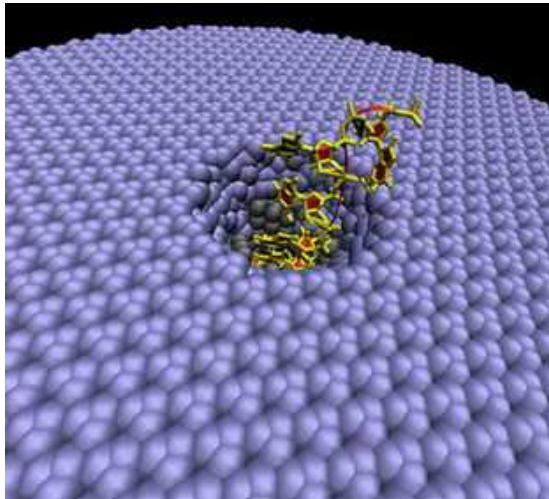
# Výsledek automatického sekvenování



# Výsledek automatického sekvenování



Přechod DNA pórem v membráně generuje elektrický signál, který umožňuje sekvenování

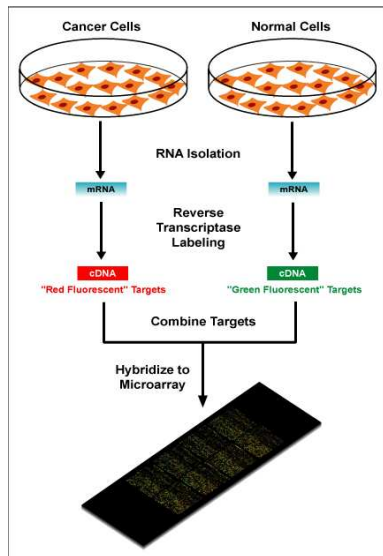


# Historie sekvenace genomů

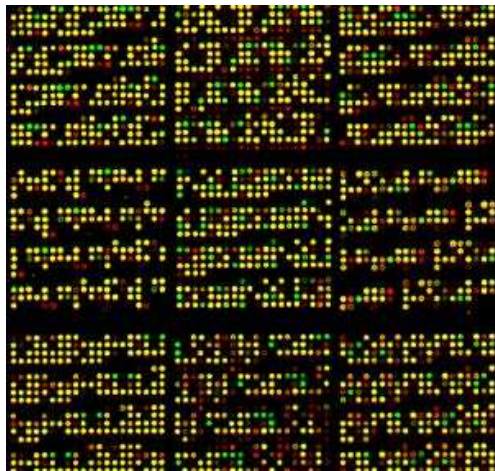
1975	bakteriofag MS2	
1977	bakteriofag PhiX174	5,375 bp
1982	bakteriofag Lambda	
1984	HIV-1	
1990	virus HCMV	230 Kbp
1995	H. influenzae	1,83 Mbp
1996	E. coli	4,60 Mbp
1996	S. cerevisiae	12,00 Mbp
1998	C. elegans	96 Mbp
2000	D. melanogaster	120 Mbp
2000	A. thaliana	130 Mbp
2001	H. sapiens	3 Gbp
2001	M. musculus	3 Gbp
2018	T. aestivum	14 Gbp



# Microarray



# Microarray

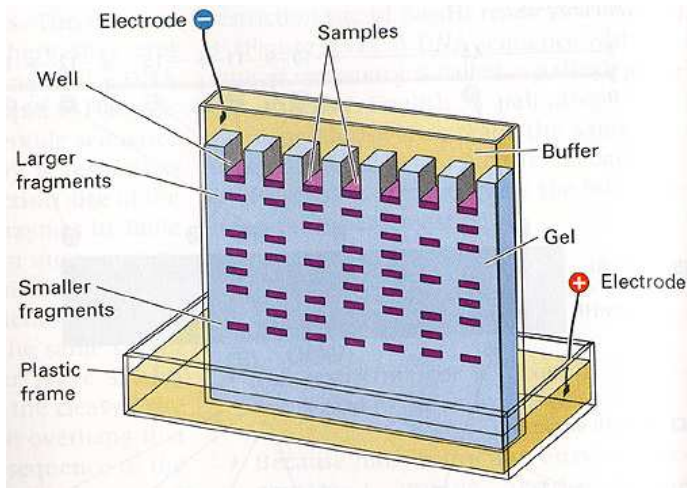


# Techniky manipulující proteiny

- ▶ izolace z buněk
- ▶ elektroforéza (dělení podle velikosti)
- ▶ zjišťování aktivity
- ▶ štěpení peptidázami
- ▶ určování sekvence
- ▶ generování protilátek pro daný protein
- ▶ ELISA a podobné testy
- ▶ produkce rekombinantních proteinů (např lacZ, GFP)
- ▶ krystalizace a určování struktury
- ▶ hmotnostní spektrometrie



# SDS-PAGE (elektroforéza v polyakrylamidovém gelu)



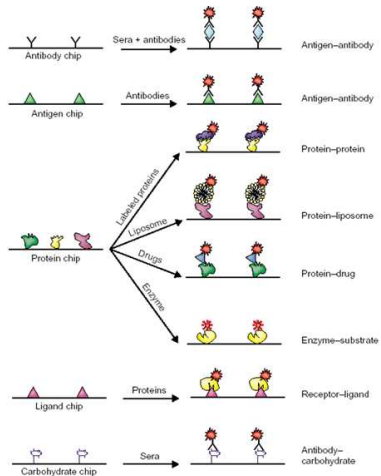
Sodium

dodecyl sulfát denaturuje proteiny a dává vzniklým komplexům záporný náboj



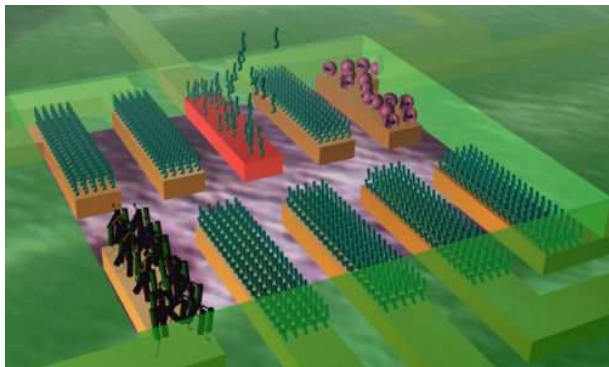


# Pole na bázi proteinů

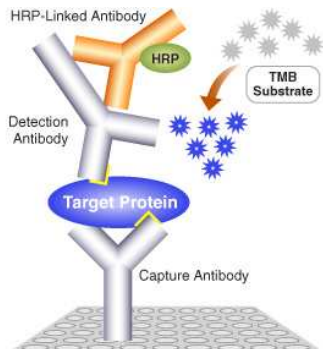
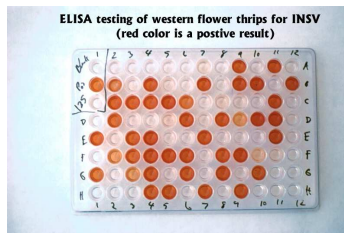


Current Opinion in Chemical Biology

# Biosenzory v proteinových čípech



# ELISA (enzyme-linked immunosorbent assay)



# Slovní pojmů molekulární biologie

central dogma	gene	RNA
genome	ORF	transcription
DNA	gene structure	mRNA
nucleotide	promoter	gene expression
3'/5'	intron	microarray
hybridization	exon	probe
replication	prokaryote	EST
DNA polymerase	ATG	translation
vector	GC content	codon
plasmid	eukaryote	protein
sequence	alignment	TATA
PCR	enhancer	mass spectrometry
DNA sequence	silencer	signal transduction
proteome	antibody	western blot
restriction enzyme	endonuclease	phylogenetics



## Analýza sekvencí



# Outline



# For Further Reading

X

