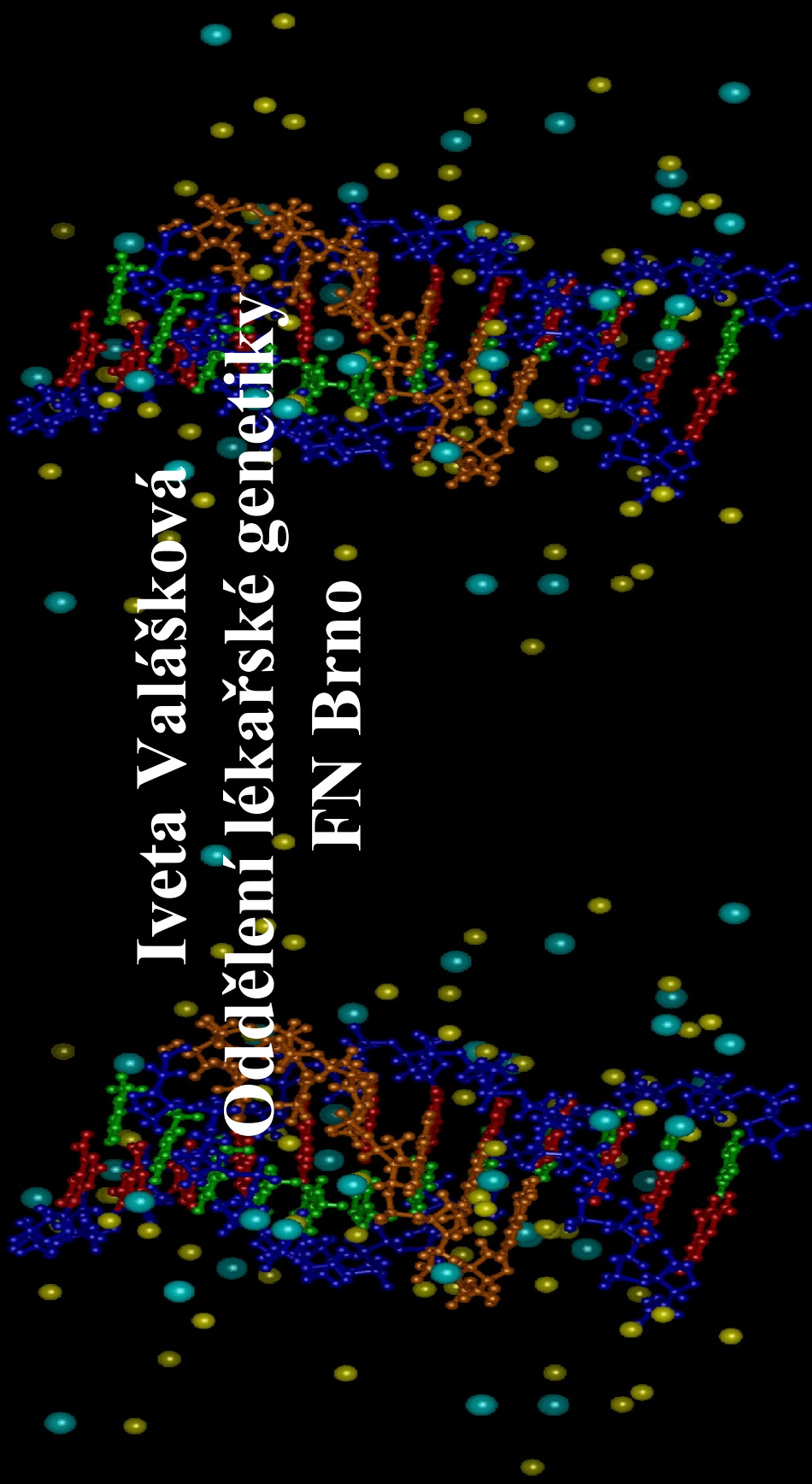


# Molekulárně genetická diagnostika

Iveta Valášková  
Oddělení lékařské genetiky  
FN Brno



Cell Nucleus Containing  
23 Pairs of Chromosomes

**Každá z bilionů buněk v lidském těle obsahuje kompletní sestavu chromozomů-genom (výjimky, např. erytrocyty)**

**3 miliardy bází od každého rodiče**

Genes

Chromosomes

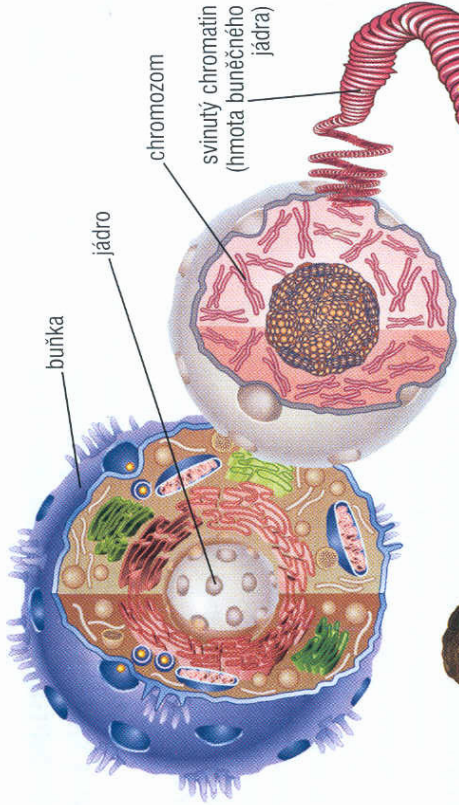
**50-263 milionů bází**

Bases

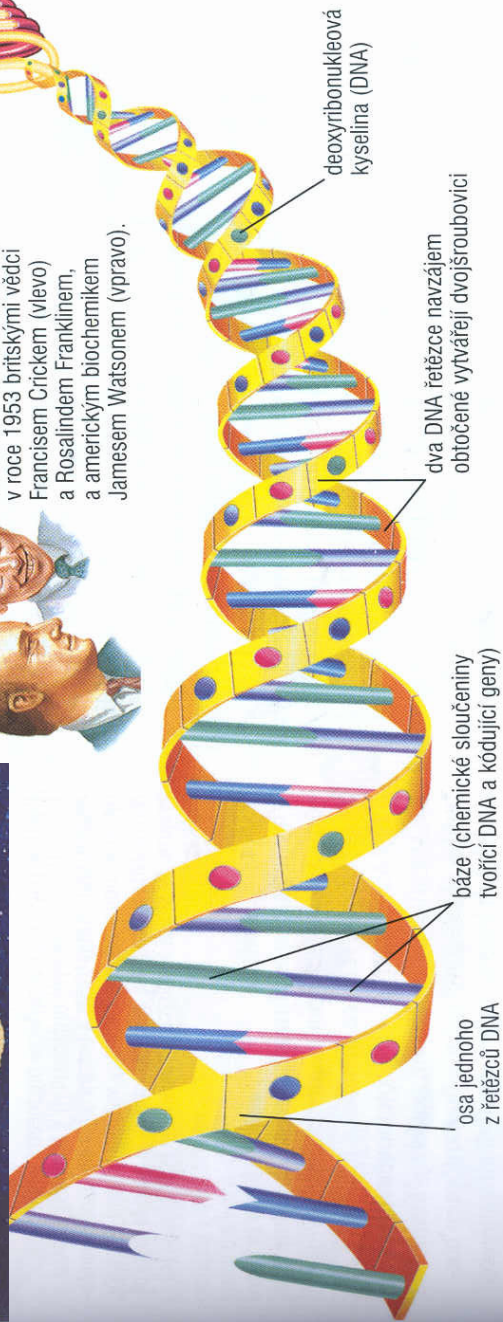
DNA Strand

**4 báze A,T,C,G se opakují mnohokrát v každém chromozomu**

Chromozomy se nacházejí v jádru téměř každé buňky. Každý z chromozomů obsahuje soubory informací, zvané geny.



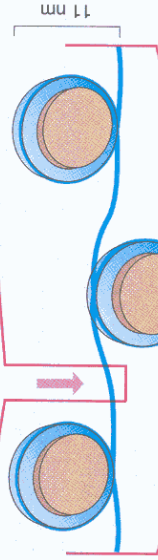
Struktura DNA byla objevena v roce 1953 britskými vědci Francisem Crickem (vlevo) a Rosalindem Franklinem, a americkým biochemikem Jamesem Watsonem (vpravo).



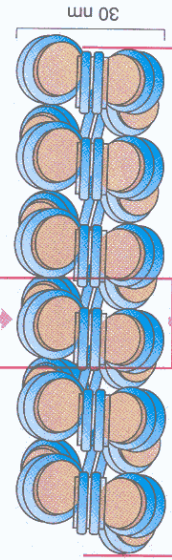
Short region of DNA double helix (five turns)



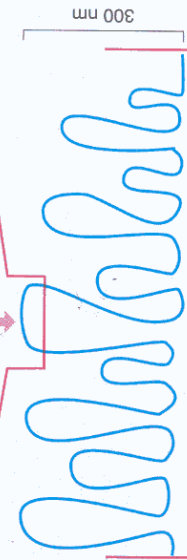
Chromatin section



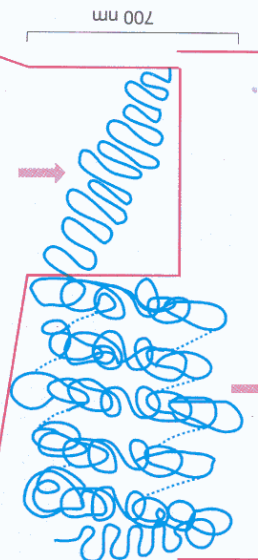
30 nm chromatin fiber with nucleosomes tightly packed



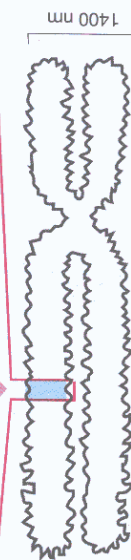
Part of a chromosome section



Condensed section of a metaphase chromosome



Metaphase chromosome



Chromatin  
osomes

- Každý z 46 lidských chromozomů je tvořen jednou molekulou dvoušrobovicové DNA
- DNA z jedné buňky je dlouhá přibližně 2 metry
- Proteiny zvané histony tvoří jádro, kolem kterého je DNA navinuta v dvojité smyčce tvořené přibližně 146 pb – NUKLEOZÓM
- Stovky sekvenčně specifických DNA-vazebných proteinů, které rozpoznávají krátké úseky DNA
- Jejich vazba má důležitý význam pro regulaci genové aktivity (genové exprese)

DNA Strand

©Chromosomes

# DNA

## *Deoxyribonukleová kyselina*

• je nositelkou genetické informace všech organismů

s výjimkou těch nebuněčných organismů, u nichž hraje tuto úlohu RNA (RNA-viry, virusoidy a viroidy).



chromosom



DNA

gen

• je pro život nezbytnou látkou, která ve své struktuře kóduje

a buňkám zadává jejich program a tím předurčuje vývoj a vlastnosti celého organismu

# DNA

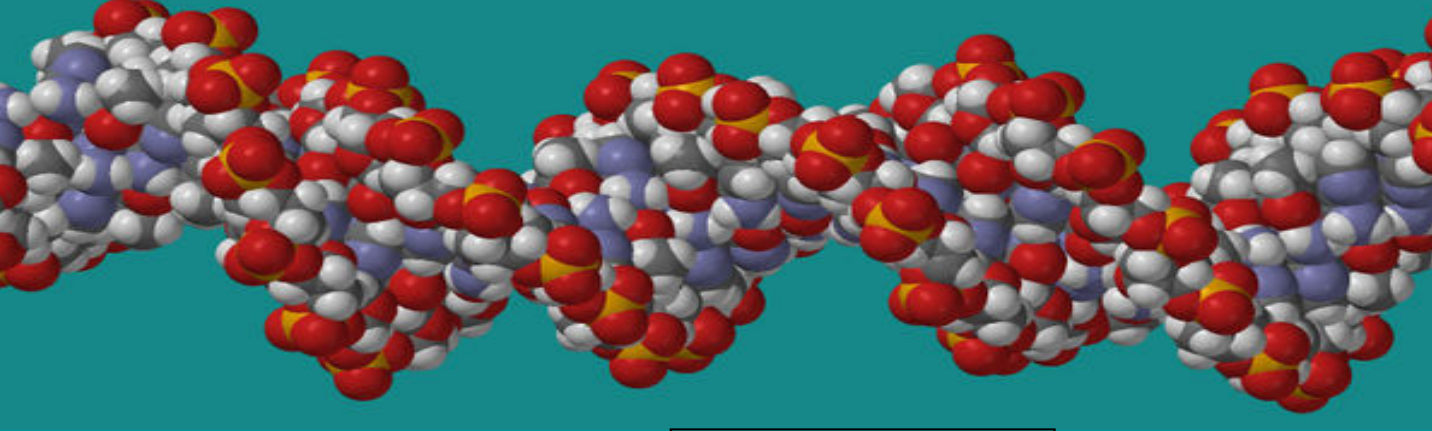
- **biologická makromolekula - polymer**

Biologická makromolekula čili biomakromolekula nebo **biopolymer**

je obecně látka vzniklá v organismu kondenzací z více stejných či různých nízkomolekulárních látek

Jejich molekulová hmotnost se pohybuje v řádech tisíců až milionů.

Z chemického hlediska jde o **nukleové kyseliny, proteiny a polysacharidy**

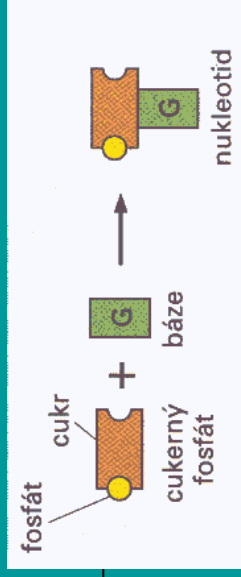
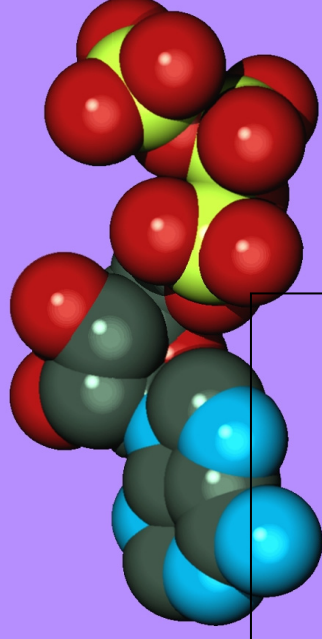


# DNA

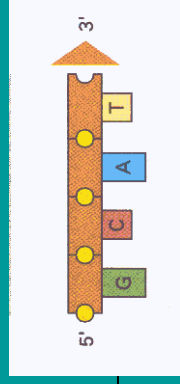
- je tvořená čtyřmi typy **nukleotidů**

skládají se ze tří složek:

- **fosfátu** (zbytek kyseliny fosforečné)
- **deoxyrybózy** (pětiuhlíkový cukr-pentóza)
- **nukleové báze** (konkrétní dusíkaté heterocyklické sloučeniny).



jsou kovalentně spojeny v polynukleotidovém řetězci cukerfosfátovou páteří, ze které vyčnívají jednotlivé báze



# Nukleové báze

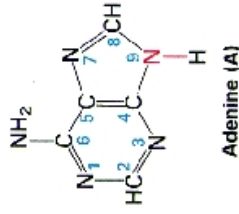
V DNA se v různých kombinacích vyskytují čtyři nukleové báze:

purinové báze - adenin(A)  
guanin (G)

pyrimidinové báze - thymin - (T)  
cytosin(C).

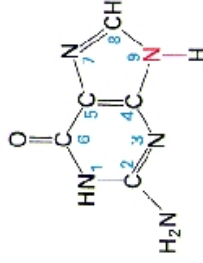
## Purinové báze

☞ adenin



Adenine (A)

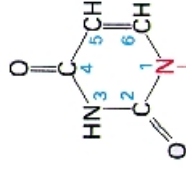
☞ guanin



Guanine (G)

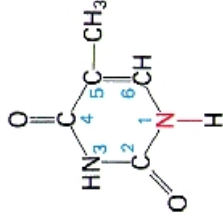
## Pyrimidinové báze

☞ uracil ... RNA



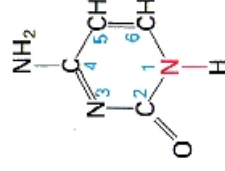
Uracil (U)

☞ thymin ... DNA



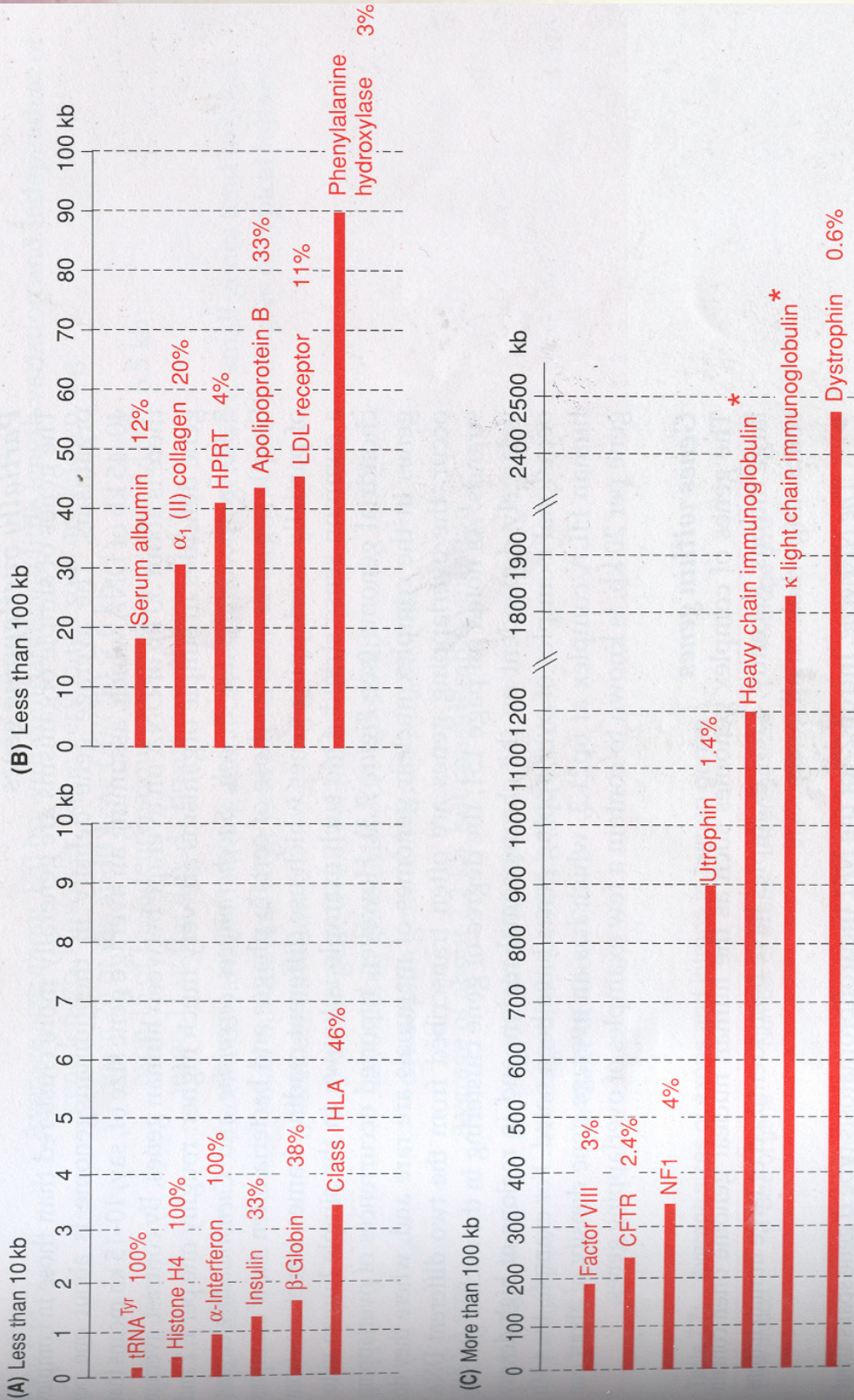
Thymine (T)

☞ cytosin



Cytosine (C)



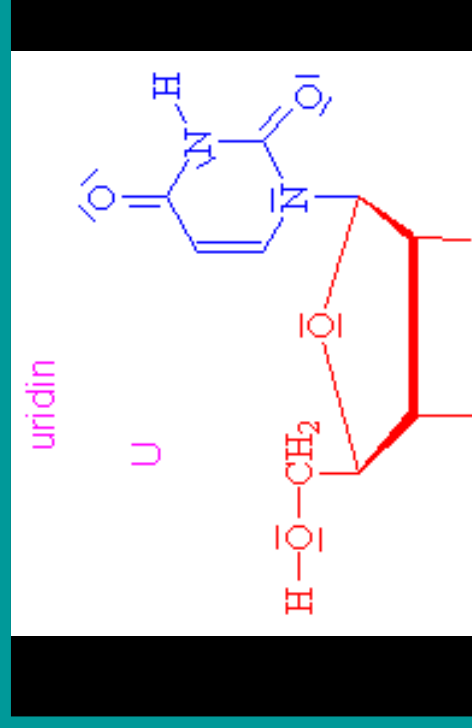
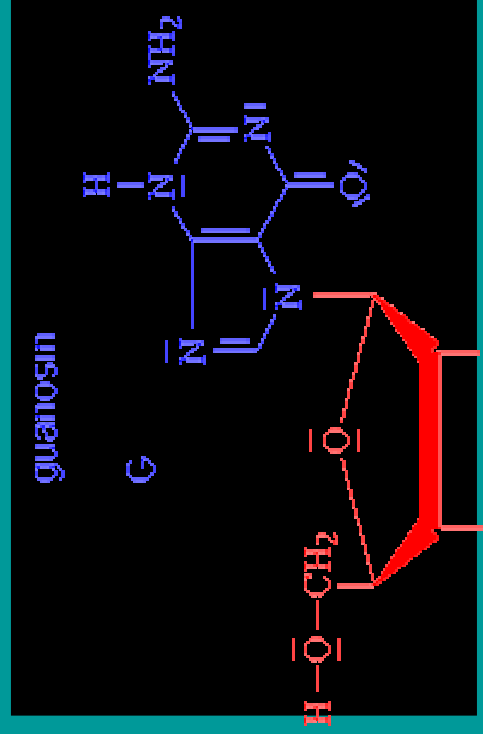
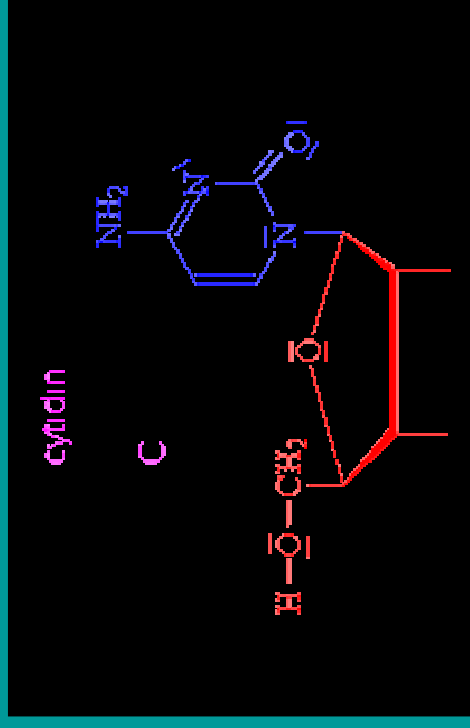
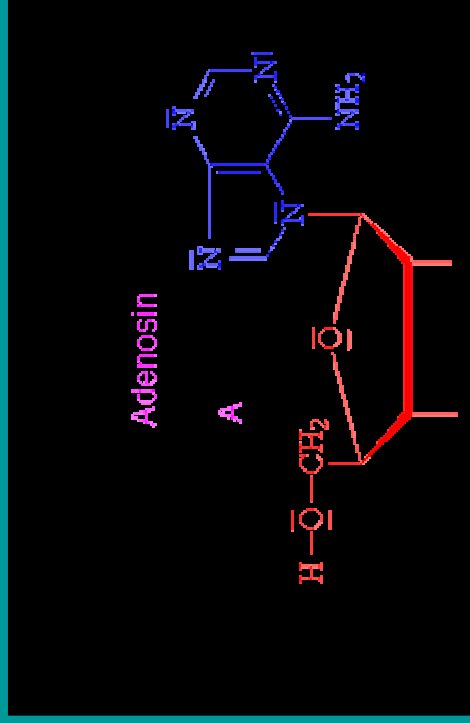


**Figure 7.6:** Human genes vary enormously in size and exon content.

Exon content is shown as a percentage of the lengths of indicated genes. Note the generally inverse relationship between gene length and percentage of exon content. Asterisks emphasize that the lengths given for the indicated Ig heavy chain and light chain loci correspond to the germline organizations. Immunoglobulin and T-cell receptor genes have unique organizations, requiring cell-specific somatic rearrangements in order to be expressed in B or T lymphocytes respectively (see page 177). CFTR, cystic fibrosis transmembrane regulator; HPRT, hypoxanthine phosphoribosyl transferase; NF1, neurofibromatous type 1.

# Nukleoizidy

vznikají vazbou bází na 1' uhlík pentózy

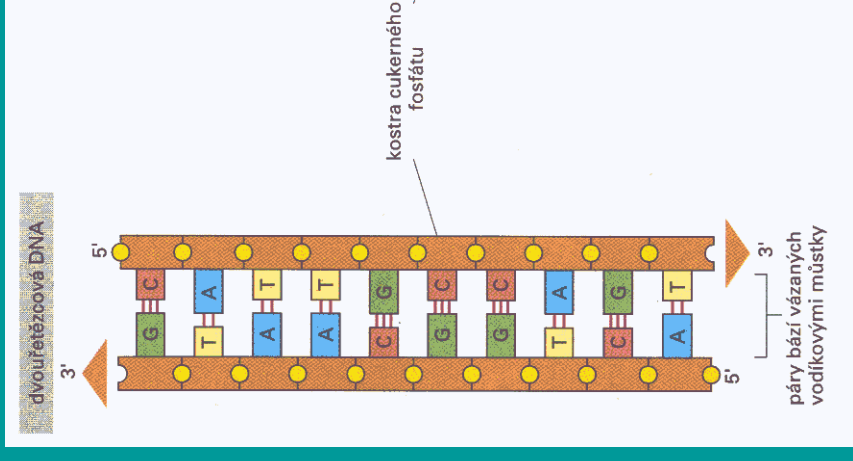
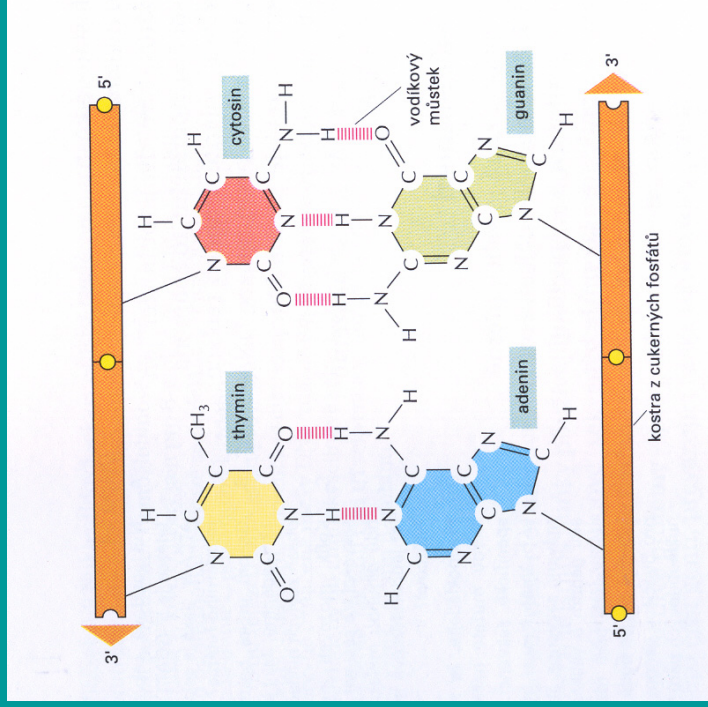


# DNA

Molekulu DNA tvoří dva polynukleotidové řetězce



jsou navzájem spojeny vodíkovými můstky mezi páry bází

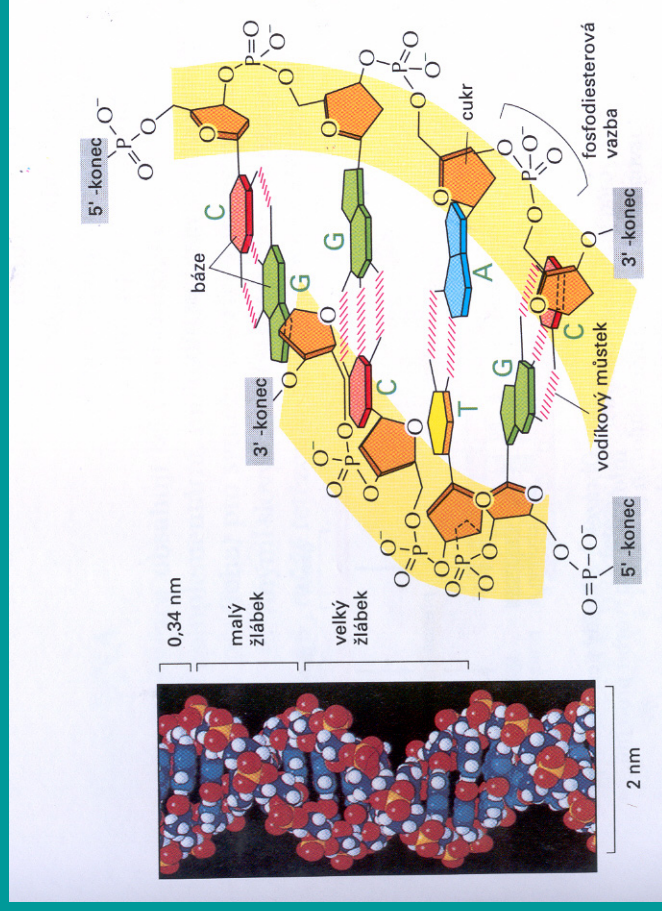


Šipky na řetězcích DNA označují jejich polaritu, obě vlákna jsou vůči sobě antiparalelní

**Komplementární páry bází**  
A-T dva vodíkové můstky  
G-C tři vodíkové můstky

Báze se mohou párovat jen v případě,  
že jsou oba řetězce antiparalelní

# Dvoušroubovice DNA



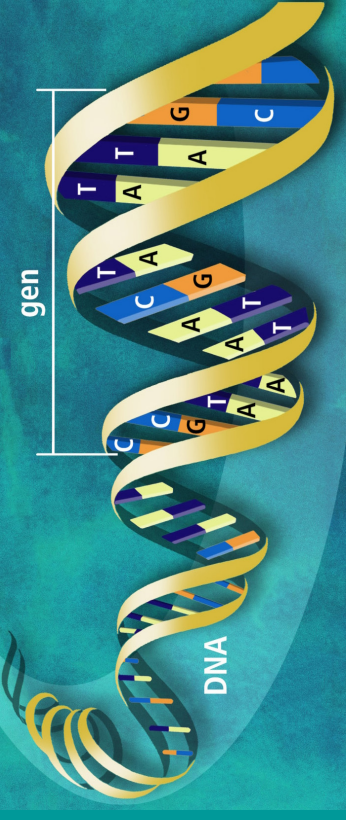
**Nukleotidy jsou spojeny navzájem fosfodiesterovými vazbami**

mezi 3'hydroxylovou skupinou (-OH) jednoho sacharidu a 5'fosfátovou skupinou (-P) druhého sacharidu

Oba konce vláknů jsou chemicky odlišné → každý polynukleotidový řetězec má jinou polaritu

# Gen

základní jednotka genetické informace



- Každý gen podmiňuje vytvoření určitého konkrétního znaku

gen obsahuje informaci k tvorbě určitého proteinu

↓  
má určitou funkci

(katalyckou, regulační, stavební)

↓  
*T.H.Morgan*

vytváří konkrétní znak

Každý gen lze ztotožnit s určitým úsekem chromozomu

Morganovy zákony

- A) Geny jsou na chromozomech uspořádány lineárně za sebou, Jejich umístění na daném chromozomu je neměnné  
Každý gen má na chromozomu svoje místo - **lokus**
- B) Všechny geny, které se nacházejí na stejném chromozomu, jsou ve **vazbě**, neboli vytvářejí **vazebnou skupinu**

# Geny

## □ Strukturální geny

obsahují informace o primární struktuře proteinů, popř. polypeptidů

➤ geny kódující proteiny se stavební funkcí

↪ základní složky cytoskeletu a mezibuněčné hmoty  
(tubulin, kolagen, elastin, aktin, myosin)

➤ geny kódující proteiny s biochemickou nebo fyziologickou funkcí

↪ enzymy, buněčné receptory, regulační proteiny, protilátky, hormony

## □ Geny pro funkční RNA

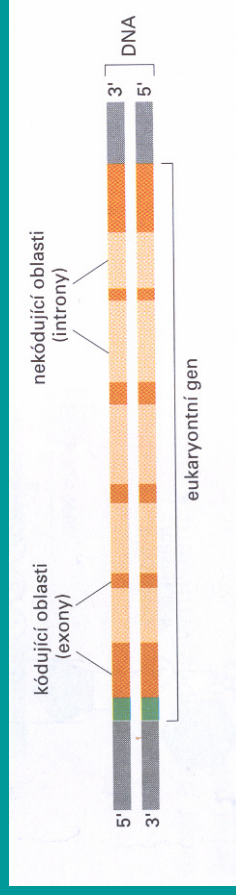
jejich transkripční produkty nepodléhají translaci

↪ geny kódující tRNA a rRNA

**Table 7.7:** Average sizes of exons and introns in human genes

Gene product	Size of gene	Number of exons	Average size of exon (bp)	Average size of intron (kb)
tRNA <sup>yr</sup>	0.1	2	50	0.02
Insulin	1.4	3	155	0.48
$\beta$ -Globin	1.6	3	150	0.49
Class I HLA	3.5	8	187	0.26
Serum albumin	18	14	137	1.1
Type VII collagen	31	118	77	0.19
Complement C3	41	29	122	0.9
Phenylalanine hydroxylase	90	26	96	3.5
Factor VIII	186	26	375	7.1
CFTR (cystic fibrosis)	250	27	227	9.1
Dystrophin	2400	79	180	30.0

# Struktura genu

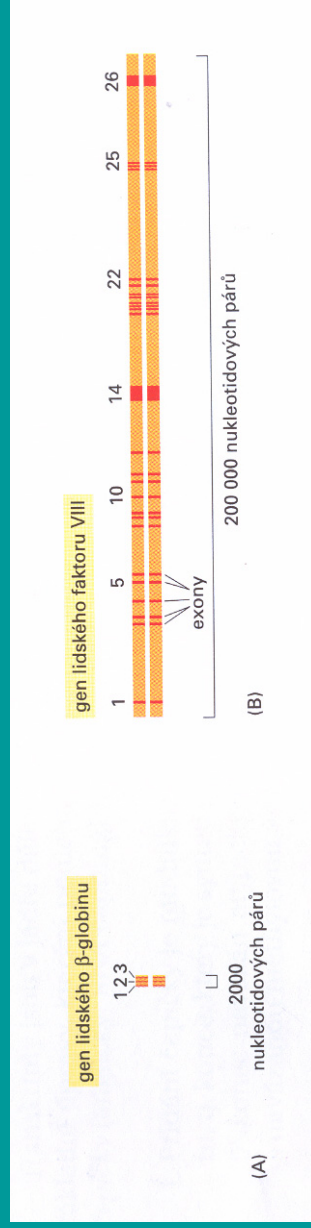


## Exony

úseky kódujících sekvencí  
jsou jen malým zlomkem délky celého genu

## Introny

nekódující sekvence  
nejsou překládány do proteínu  
velikost 80-10 000 nukleotidů





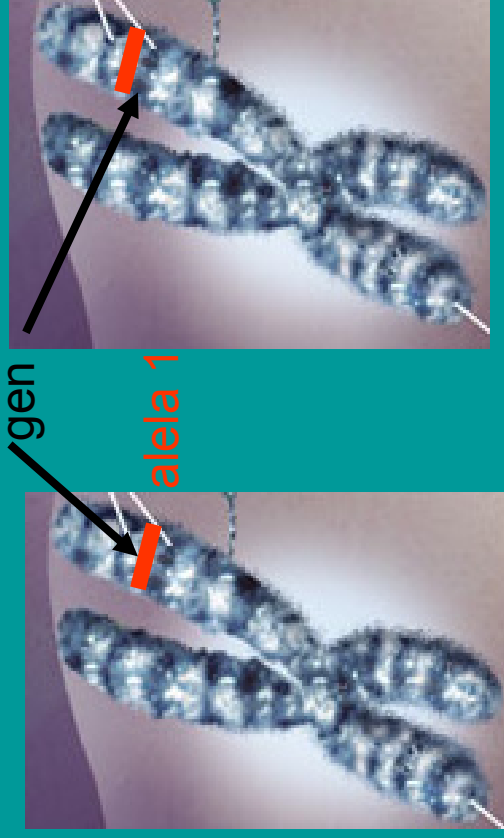
# Alela

Forma genu

Každý gen se vyskytuje  
v různých konkrétních formách  
alelách

**Gen**  
podmiňuje určitý znak

**Alela**  
obsahuje „návod“ k vytvoření  
určité konkrétní  
podoby znaku



## Alela

- **dominantní** - převažující
- **recesivní** - ustupující
- **kodominance** – alely se společně podílejí na konečné podobě znaku

# Genotyp

Soubor alel daného jedince

Genotypy jedinců stejného druhu mohou být rozdílné

# Fenotyp

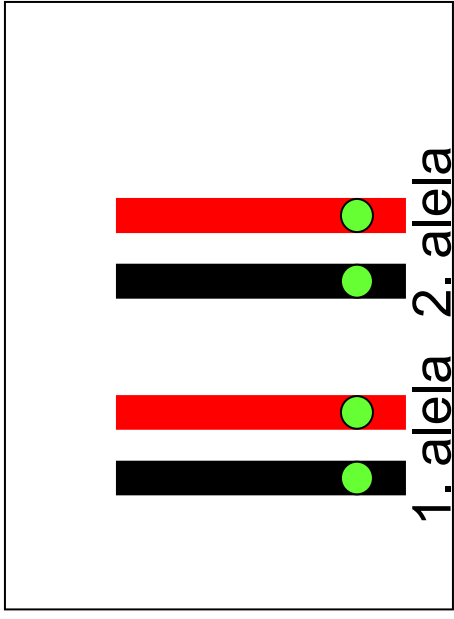
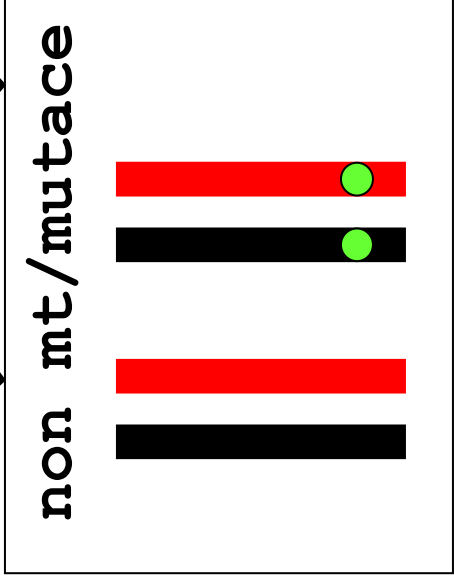
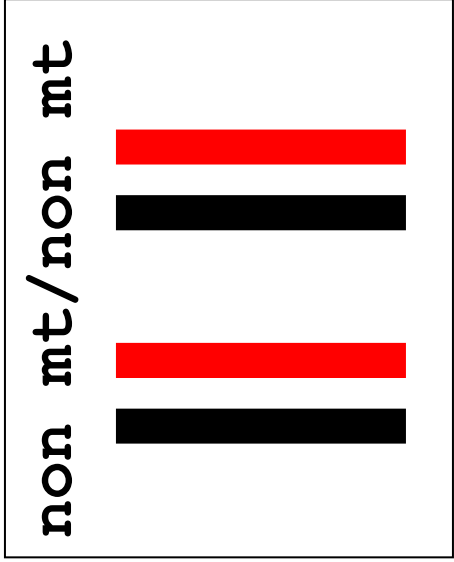
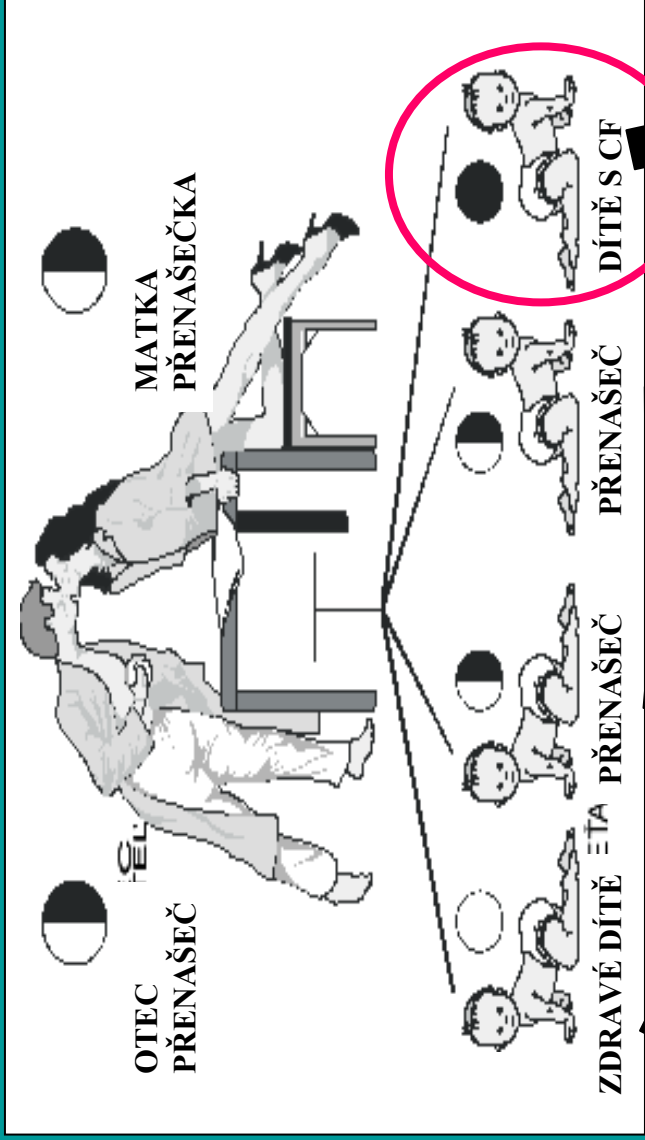
Vnější projev genotypů  
Soubor všech znaků organismu.

# Homozygot

Jedinci se shodným párem alel příslušného genu

# Heterozygot

Jedinci s odlišnými alami příslušného genu

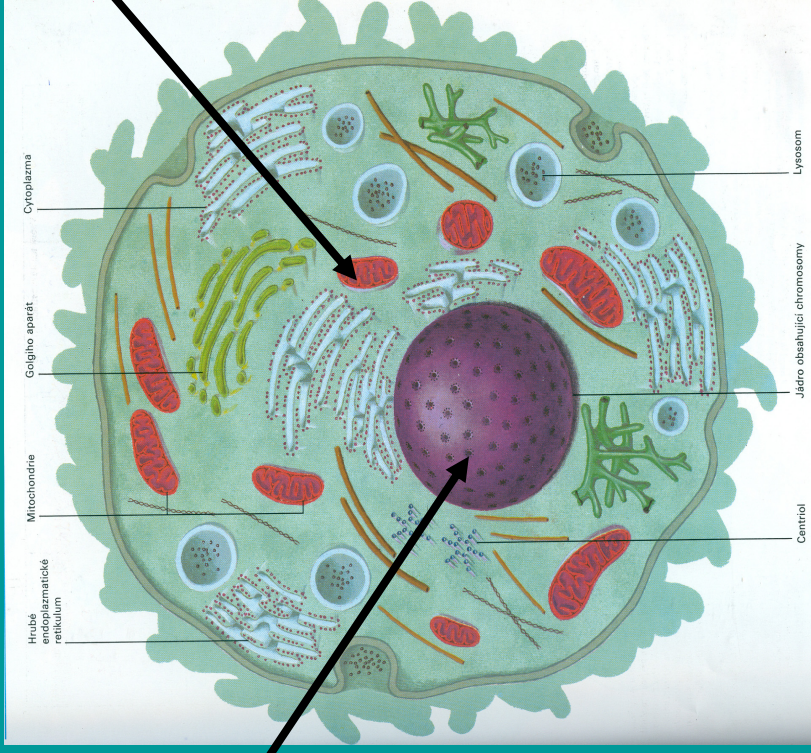
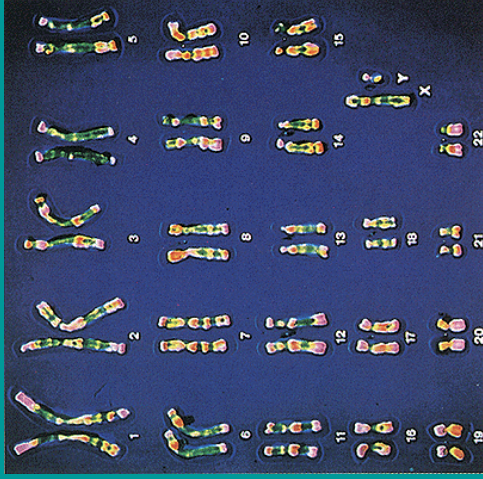


# Genom

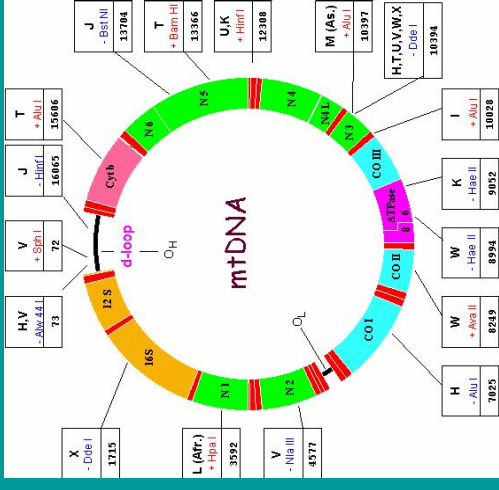
Soubor všech genů daného organismu

**Veškerá genetická informace obsažená v DNA v lidských buňkách**

jaderný genom

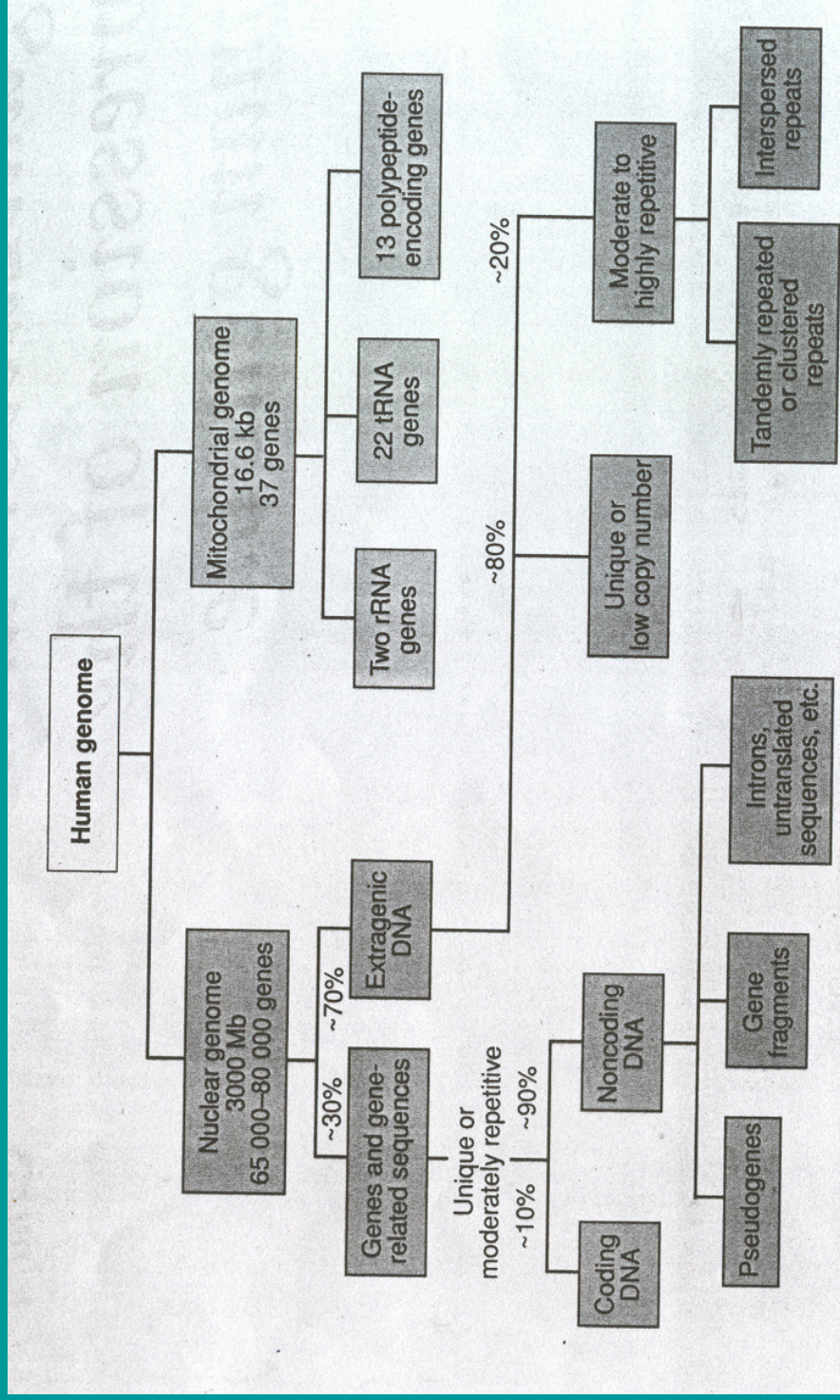


mimojaderný genom  
– mitochondriální



Genom u jedinců stejného druhu je stejný

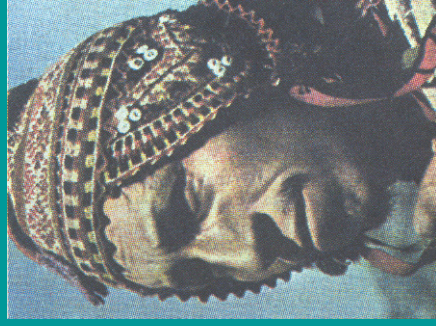
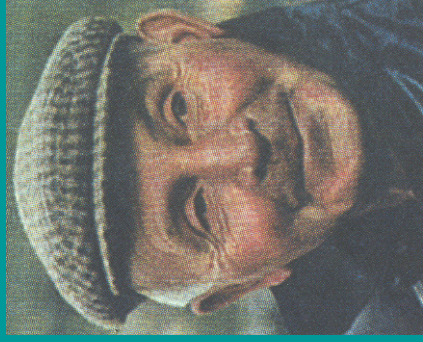
# Lidský genom



# Lidský jaderný a mitochondriální genom

	Nuclear genome	Mitochondrial genome
Size	3000 Mb	16.6 kb
No. of different DNA molecules	23 (in XX) or 24 (in XY) cells, all linear	One circular DNA molecule
Total no. of DNA molecules per cell	23 in haploid cells; 46 in diploid cells	Several $\times 10^3$
Associated protein	Several classes of histone and nonhistone protein	Largely free of protein
No. of genes	~65 000–80 000	37
Gene density	~1/40 kb	1/0.45 kb
Repetitive DNA	Large fraction, see <i>Figure 7.1</i> .	Very little
Transcription	The great bulk of genes are transcribed individually	Continuous transcription of multiple genes
Introns	Found in most genes	Absent
% of coding DNA	~3%	~93%
Codon usage	See <i>Figure 1.22</i>	See <i>Figure 1.22</i>
Recombination	At least once for each pair of homologs at meiosis	None
Inheritance	Mendelian for sequences on X and autosomes; paternal for sequences on Y	Exclusively maternal

# Genom x Genotyp

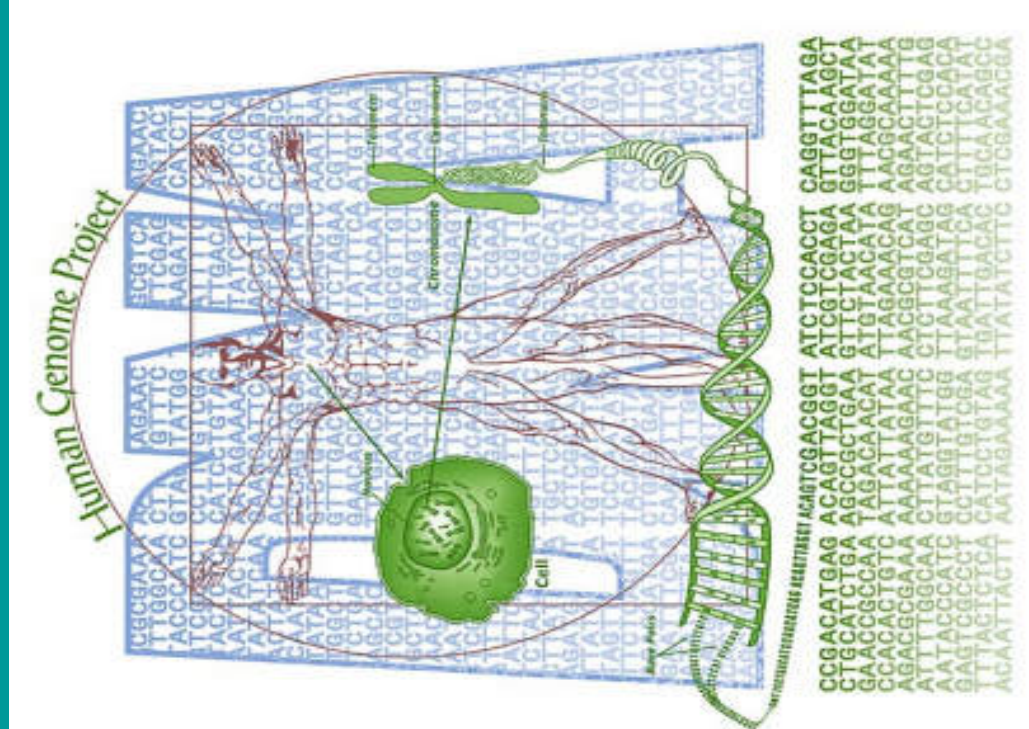


Genom u jedinců stejného druhu je stejný  
Genotypy jedinců stejného druhu mohou být rozdílné



# Human Genome Project

## Projekt lidský genom



└1986: Santa Fe - DoE Intl. Meeting

└James Watson:

„vstupitna cestu od dvojišroubovice k 3 miliardám schodů lidského genomu“

└1988: Kongres USA schválil 15 letý projekt a dotaci 3 mld USD

└1990: začátek projektu

└2005: předpokladané ukončení

Koordinace:

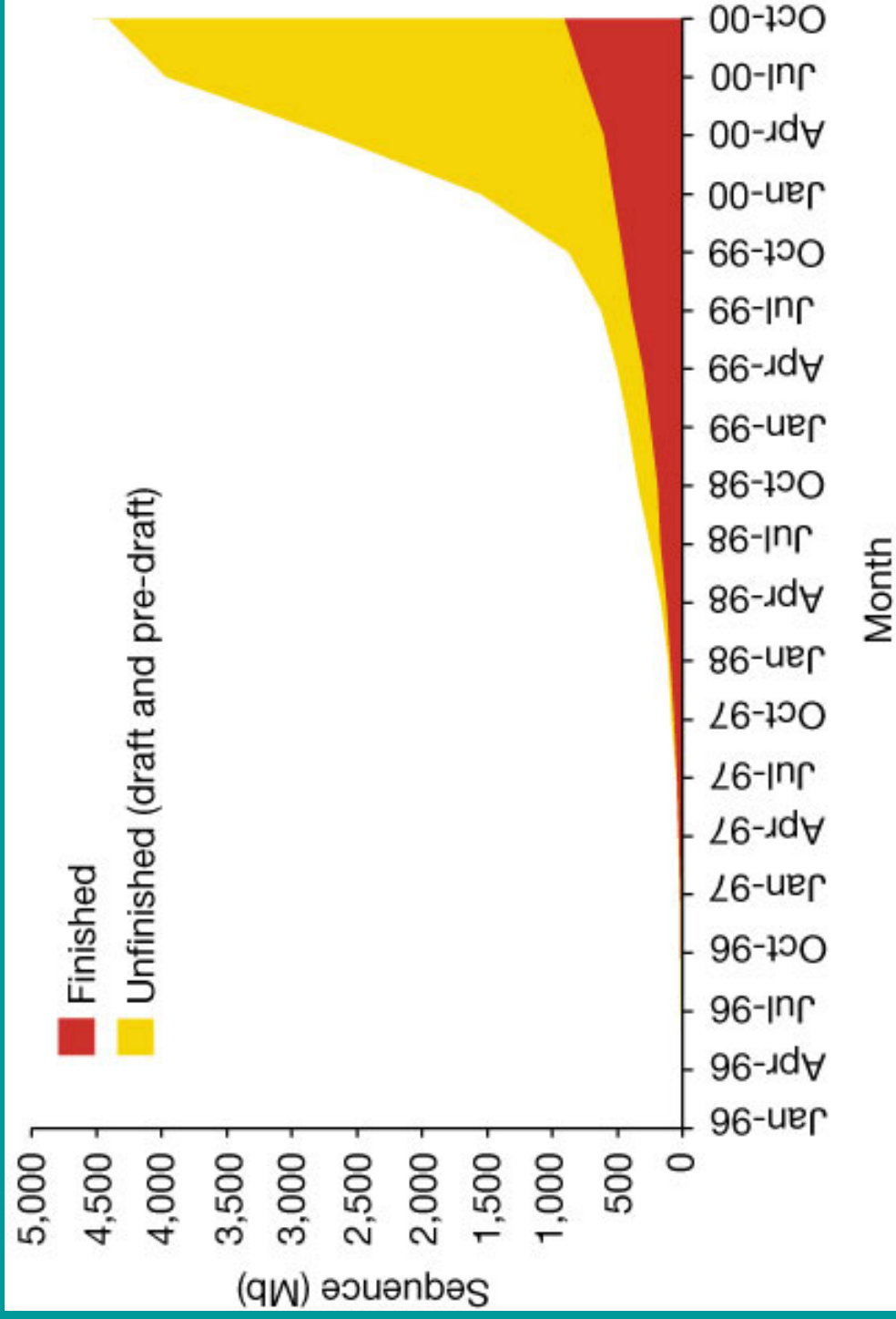
1988: HUGO (Human Genome Organization)

HUGO Europe – Americas - Pacific

# HUGO - cíle

- └ Triangulovat genom pomocí DNA markerů
- └ Identifikovat a mapovat geny, určit jejich strukturu a funkci v zdraví i v patologii
- └ Identifikovat důležité mimogenové sekvence
- └ Určit úplnou sekvenci genomu (3,2 Gb)
- └ Všechny data uložit v veřejně přístupných databázích
- └ Sekvenovat genomy modelových organismů (drozofila, myš, *C. elegans*, *A. thaliana* a i.)
- └ Zkoumat etické, právní a sociální aspekty

# Nárúst údajů o sekvenci DNA



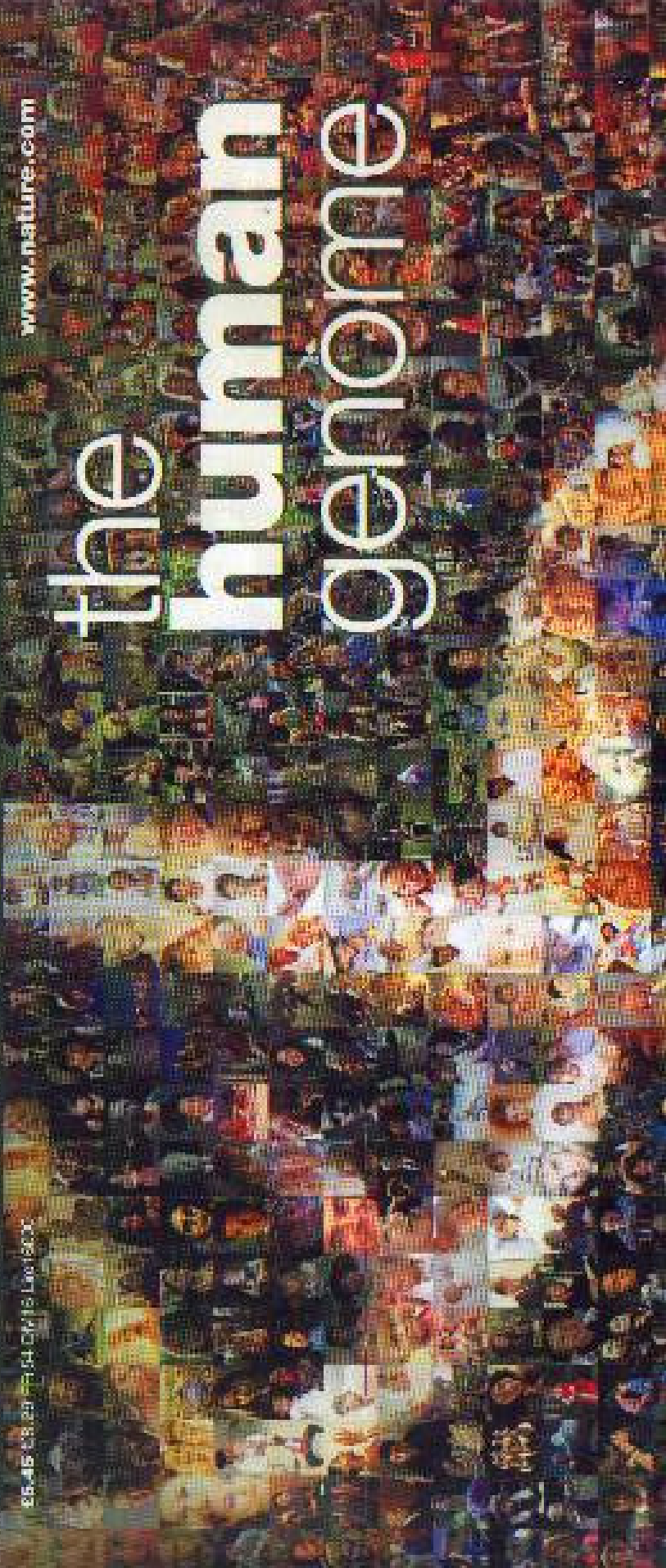
15 February 2001

# Nature

ISSN 0950-0804 PUBLISHED WEEKLY

www.nature.com

the  
**human**  
genome



# Human Genome Project

**Table 11.** Genome overview.

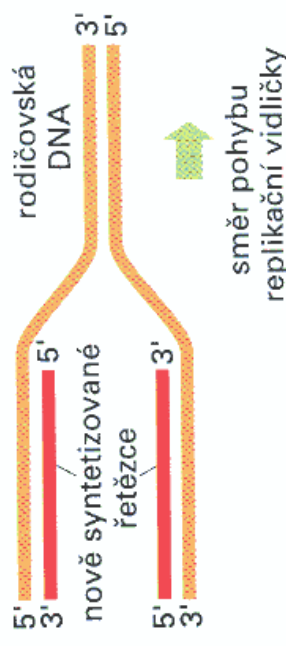
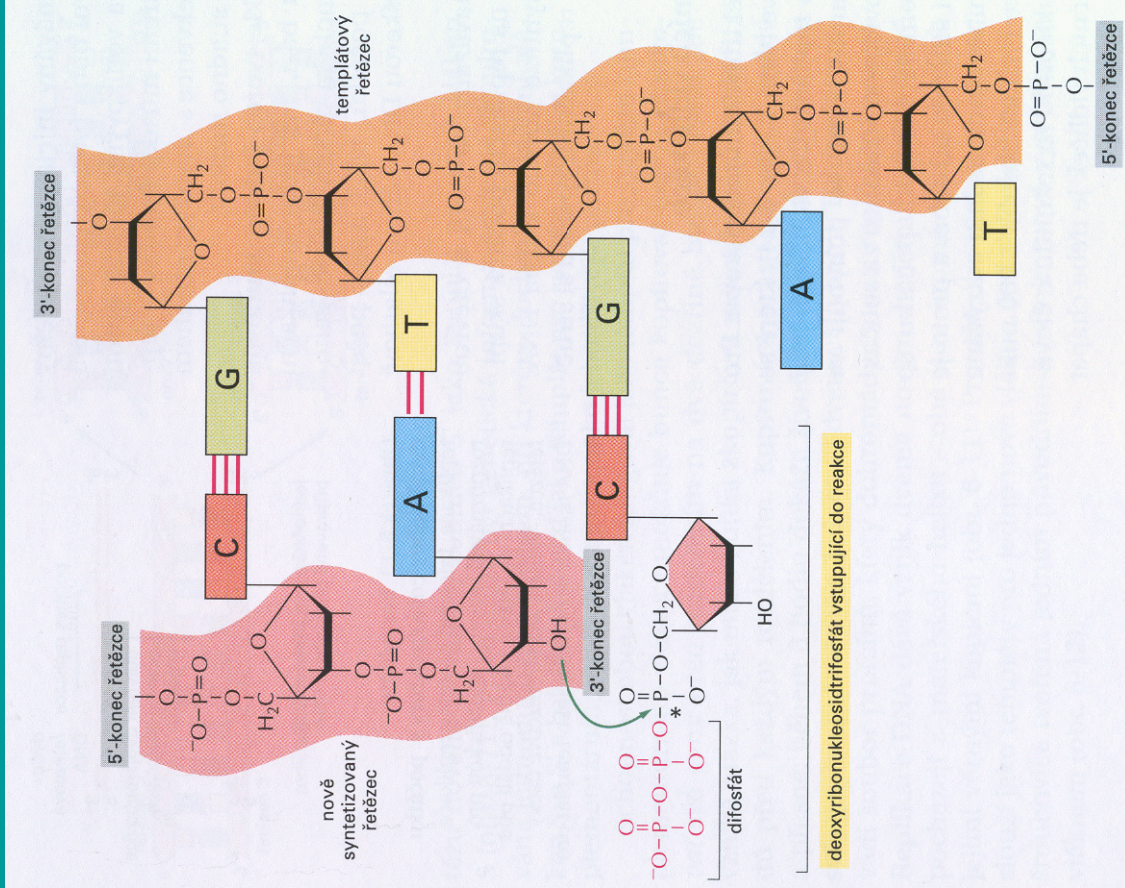
Size of the genome (including gaps)	2.91 Gbp
Size of the genome (excluding gaps)	2.66 Gbp
Longest contig	1.99 Mbp
Longest scaffold	14.4 Mbp
Percent of A+T in the genome	54
Percent of G+C in the genome	38
Percent of undetermined bases in the genome	9
Most GC-rich 50 kb	Chr. 2 (66%)
Least GC-rich 50 kb	Chr. X (25%)
Percent of genome classified as repeats	35
Number of annotated genes	26,383
Percent of annotated genes with unknown function	42
Number of genes (hypothetical and annotated)	39,114
Percent of hypothetical and annotated genes with unknown function	59
Gene with the most exons	Titin (234 exons)
Average gene size	27 kbp
Most gene-rich chromosome	Chr. 19 (23 genes/Mb)
Least gene-rich chromosomes	Chr. 13 (5 genes/Mb), Chr. Y (5 genes/Mb)
Total size of gene deserts (>500 kb with no annotated genes)	605 Mbp
Percent of base pairs spanned by genes	25.5 to 37.8*
Percent of base pairs spanned by exons	1.1 to 1.4*
Percent of base pairs spanned by introns	24.4 to 36.4*
Percent of base pairs in intergenic DNA	74.5 to 63.6*
Chromosome with highest proportion of DNA in annotated exons	Chr. 19 (9.33)
Chromosome with lowest proportion of DNA in annotated exons	Chr. Y (0.36)
Longest intergenic region (between annotated + hypothetical genes)	Chr. 13 (3,038,416 bp)
Rate of SNP variation	1/1250 bp

\*In these ranges, the percentages correspond to the annotated gene set (26,383 genes) and the hypothetical + annotated gene set (39,114 genes), respectively.

# Přenos genetické informace

Genetická informace je přenášena z generace na generaci v nezměněné formě

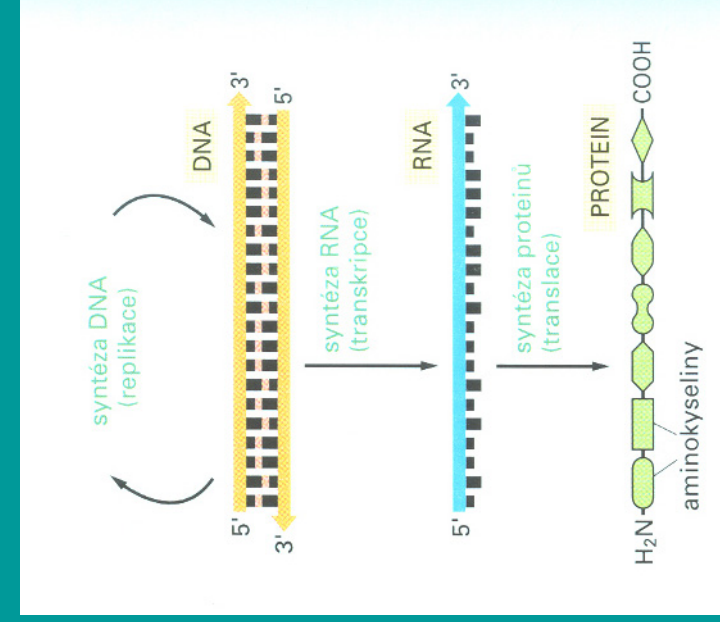
## replikace DNA



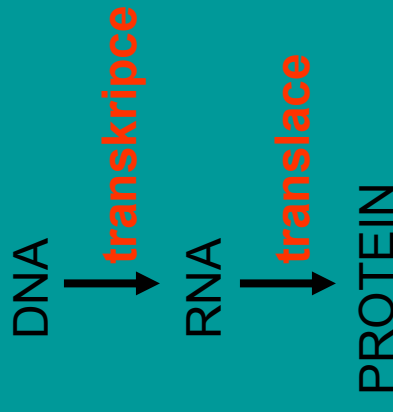
Obrázek 6-16 Polarita DNA-řetězců v replikační vidličce.

# Ústřední dogma molekulární biologie

Všechny typy buněk dekódují genetickou informaci tímto způsobem



Tok genetické informace



# RNA

RNA se liší od DNA v těchto aspektech:

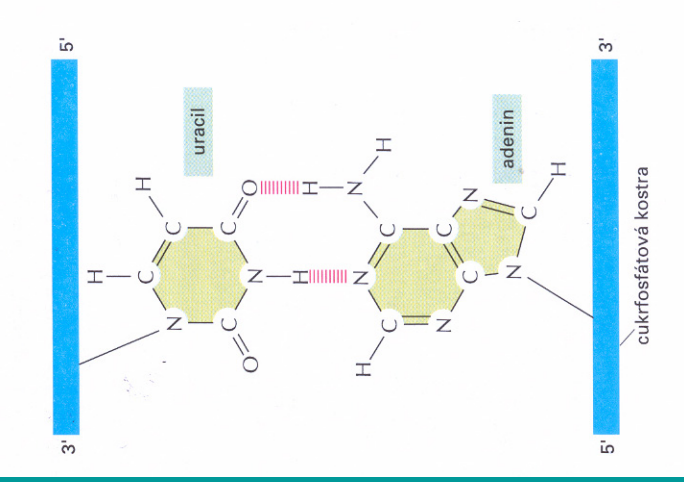
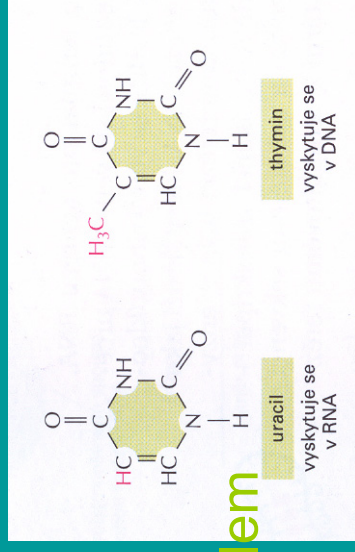
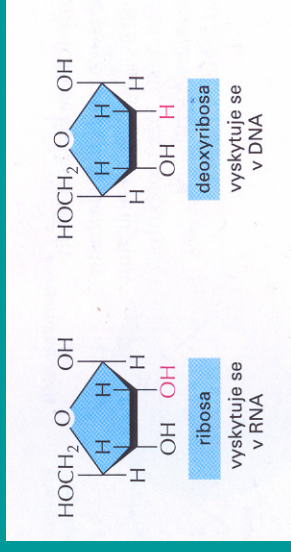
□ nukleotidy v RNA se nazývají **rybonukleotidy**, protože cukernou složkou je **riboza**

□ v RNA se vyskytují báze **adenin**

**cytosin**

**guanin**

**thymin nahrazen uracilem**

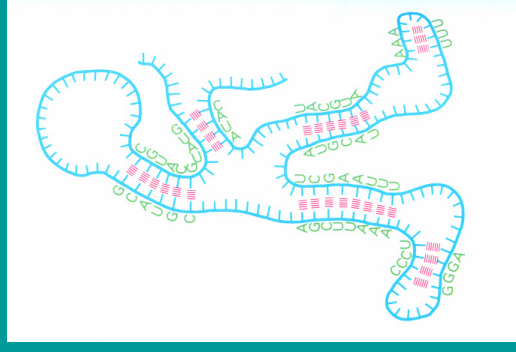




# RNA

jednořetězcová molekula

na základě párování bazí se může sbalit do různých tvarů



## Typ RNA

mRNA

rRNA

tRNA

malé RNA

## Funkce

kódování proteinů

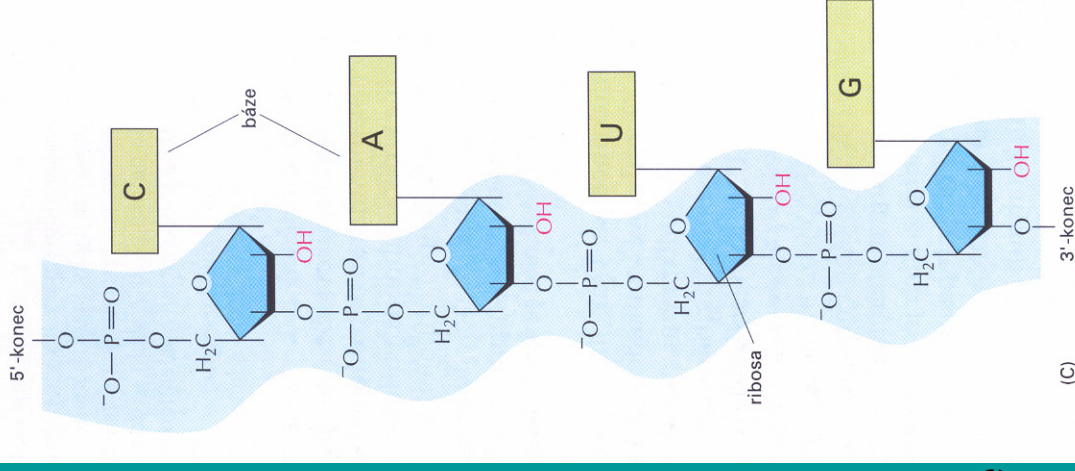
součást ribozomů, účast na proteosyntéze

adaptor mezi mRNA a aminokyselinou

při proteosyntéze

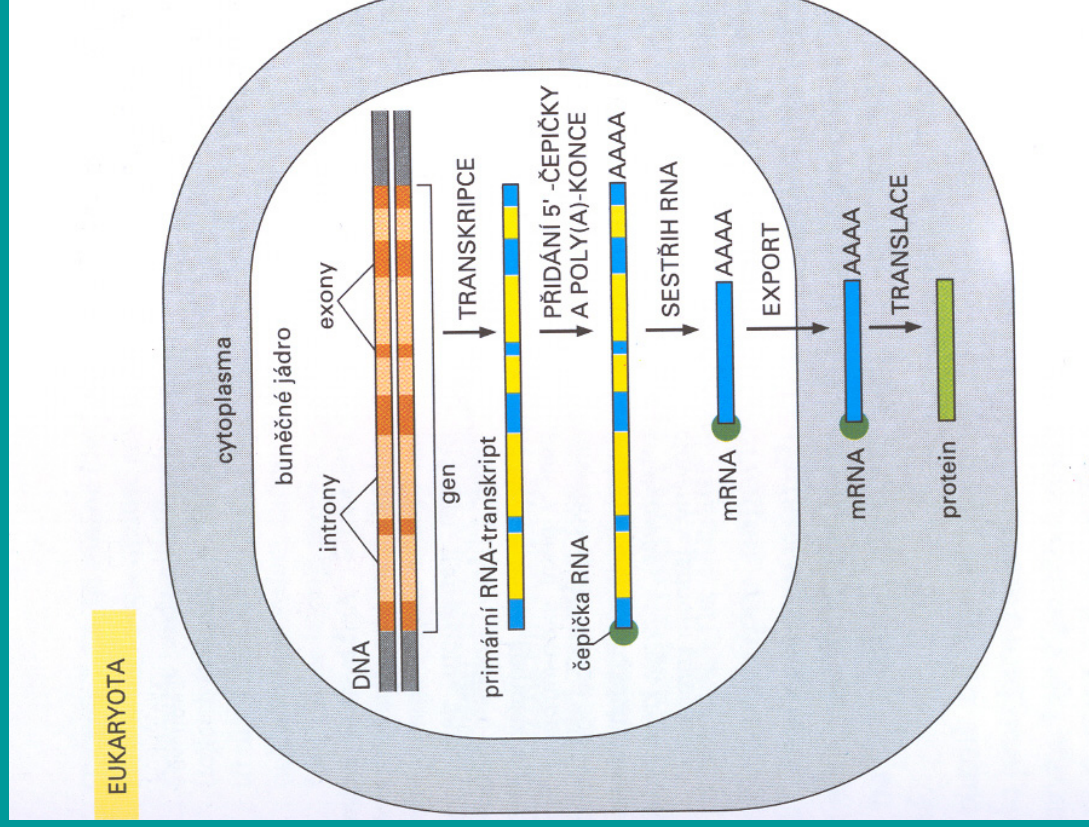
účast na sestřihu pre-RNA,

účast na transportu proteinů do ER



(C)

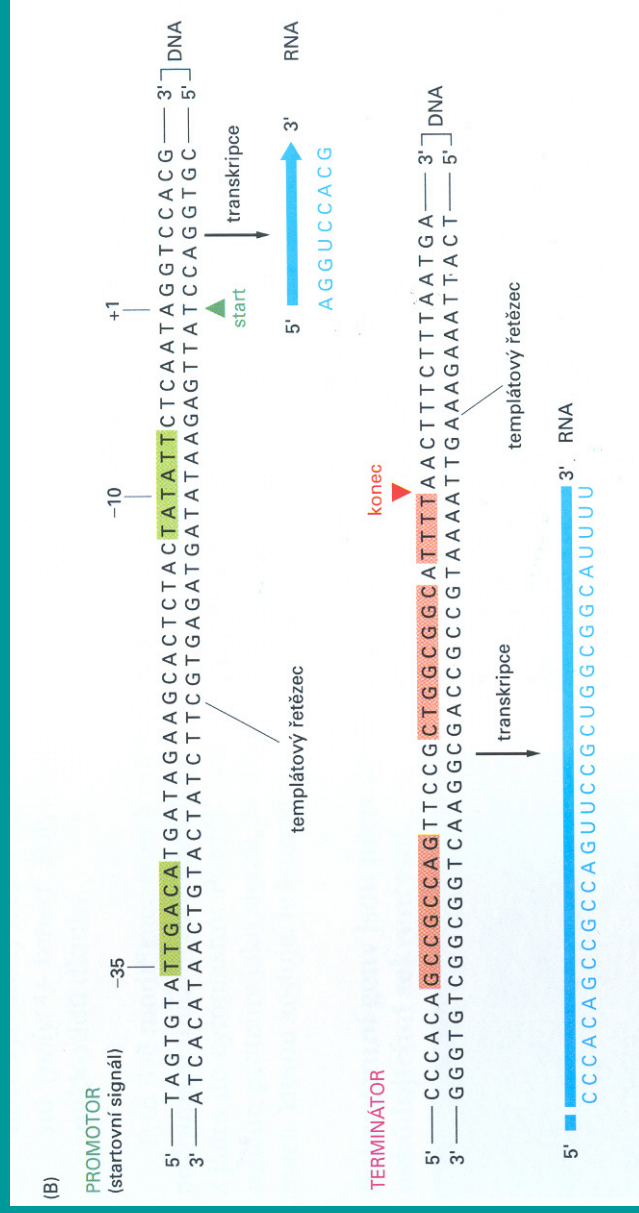
# Tok genetické informace



# Transkripce

Nukleotidové sekvence signalizují **RNA-polymeráze** začátek a konec transkripce

↓  
**promotor**  
**terminátor**



# Posttranskripční úpravy

## RNA processing

Transkripční vzniká primární transkript upravován v jádře

### 1. Přidání čepičky

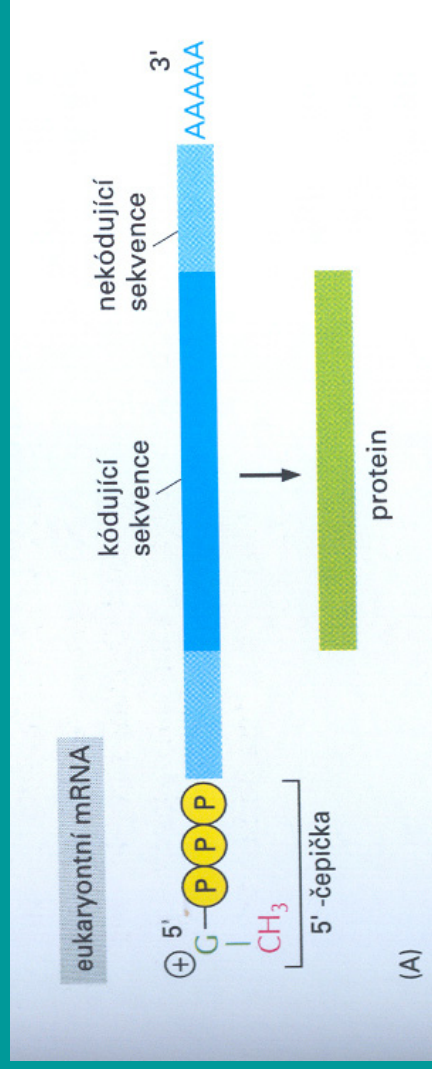
modifikace 5' konce primárního transkriptu

mnoho eukaryotních mRNA má modifikaci na 2'OH skupině skupině ribosy

### 2. Polyadenylace

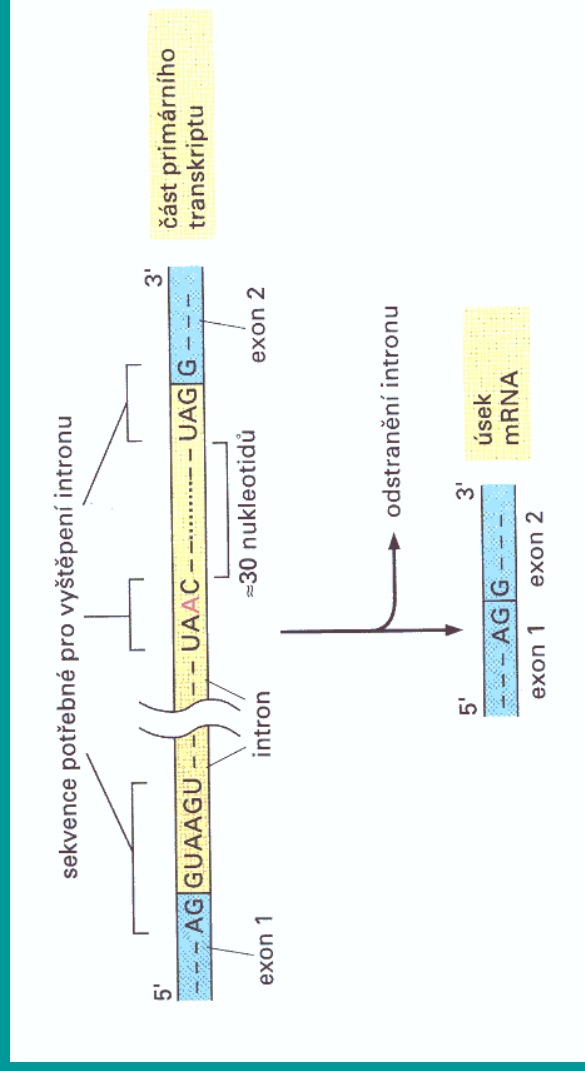
odštěpení 3' konce primárního transkriptu,

je nahrazen přidáním několika adeninů (poly(A)konec)



# Sestřích RNA RNA splicing

Z primárního transkriptu jsou vystřiženy všechny introny a exony jsou spojeny  
Výsledkem je molekula mRNA, která obsahuje nepřerušenu kódující sekvenci



Tři nukleotidové sekvence jsou nutné k rozpoznávání a vyštěpení intronu  
rozpoznávány snRNP částicemi, které rozštěpí exon-intronové hranice  
a spojí sousední exony

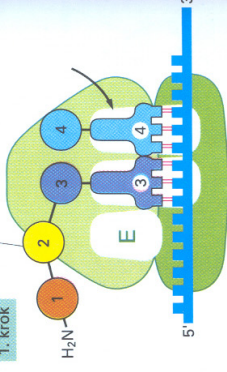
# Translace mRNA

Tříkrokový cyklus:

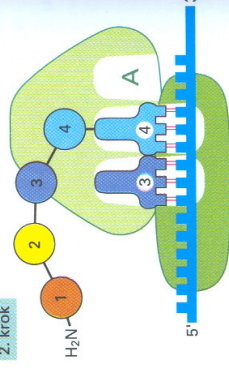
1. krok  
aminoacyl-tRNA navázána do rybosomu
2. krok  
vzniká nová peptidická vazba
3. Krok  
uvolnění tRNA bez navázané aminokyseliny

rostoucí polypeptidový řetězec

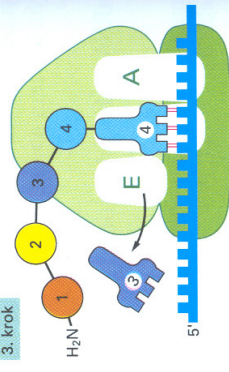
1. krok



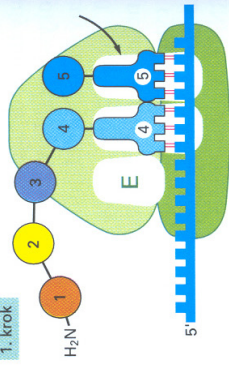
2. krok



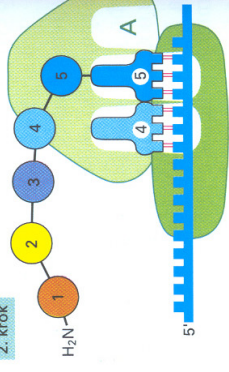
3. krok



1. krok



2. krok



# Čtecí rámce

Při translaci je nukleotidová sekvence mRNA čtena z 5' konce po trojicích nukleotidů



— Leu — Ser — Val — Thr —



— Ser — Ala — Leu — Pro —



— Gln — Arg — Tyr — His —





# Mutace



- jakákoliv dědičná změna v genetické materiálu
- změny v DNA, jejichž frekvence je nižší než 1%

*Epimutace – nejsou to změny ve fenotypu, které jsou důsledkem změny sekvence DNA (poruchy metylace aj.)*

**Mutace – pozitivní (zdroj variability, selekce)**

- negativní ( 4500 monogenně dědičných chorob, stárnutí)
- neutrální

*Každý člověk 5 – 10 patologických mutací*

# Mutace

spontální x indukované

gametická x somatická

dominantní x recesivní

přímé x zpětné

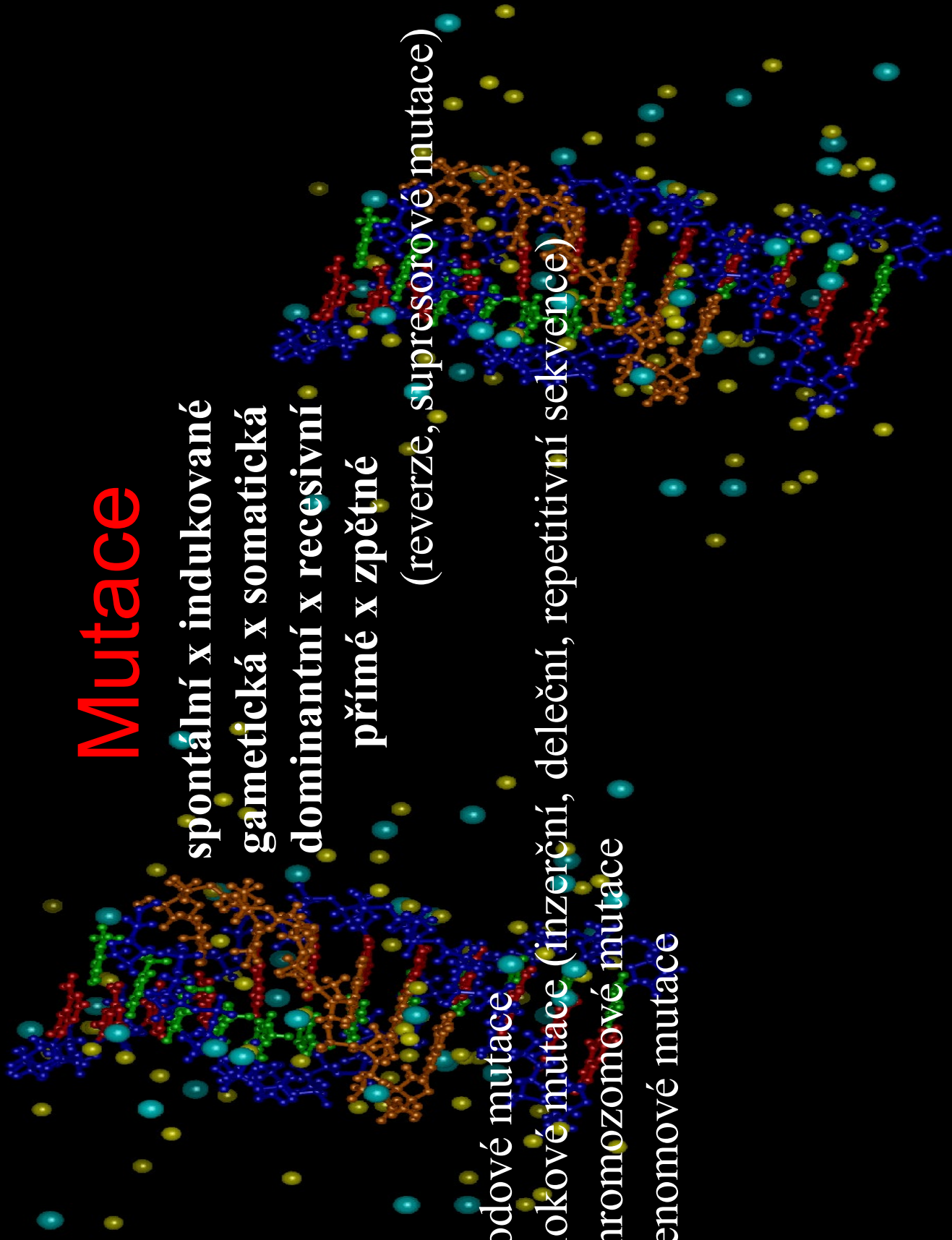
(reverze, supresorové mutace)

Bodové mutace

Blokové mutace (inzerční, deleční, repetitivní sekvence)

Chromozomové mutace

Genomové mutace

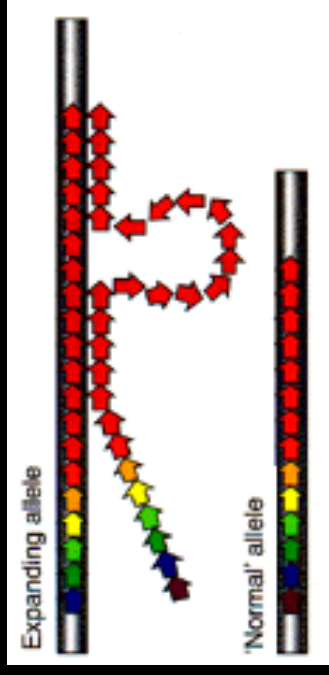


# Mutace

## Typy mutací

- **delece**
  - od 1 bp po megabáze
- **inzerce**
  - včetně duplikací
- **jednobázové substituce**
  - missense transverze
  - nonsense tranzice
  - splice site
  - silent
- **posunové (frameshift**
  - v důsledku delecí, inzercí, poruch splicingu
- **„dynamické“ (expansia rep. sekv.)**
  - tandemové repetice, které často expandují během meiózy

Podle sekvence:  
- v kódující sekvenci  
- v nekódující sekvenci



# Mutace

## Klasifikace mutací z hlediska efektu na genový produkt

### 1. Produkt se sníženou až nulovou funkcí (*loss-of-function*)

- nejčastěji je produktem genu enzym
- častým typem mutace je delece genu

### 2. Produkt s abnormální funkcí (*gain-of-function*)

- nejčastěji je produktem neenzymatický protein
- časté u nádorů (somat. mutace), zřídka u monogen. chorob
- delece genu se neuplatňuje (nevede k nové funkci)

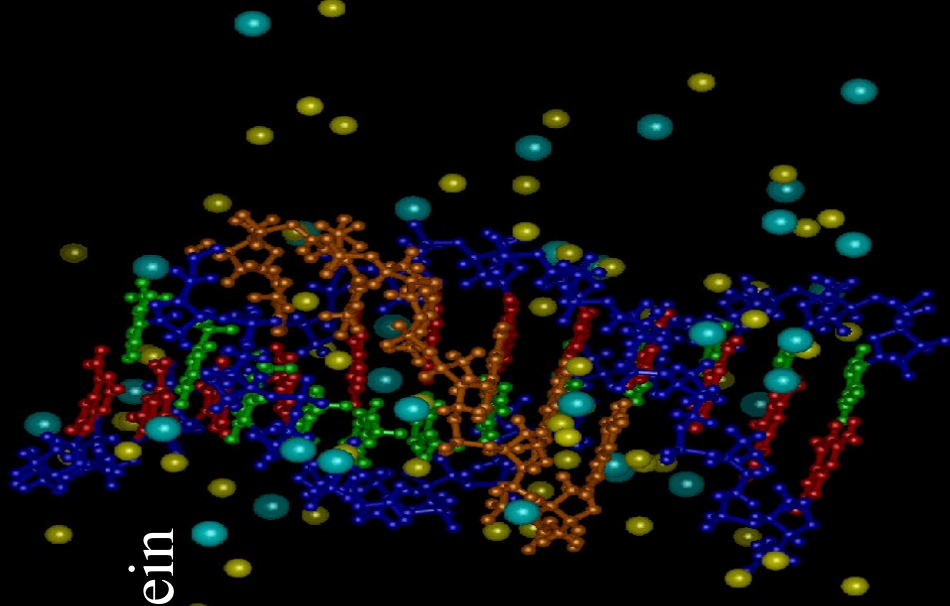
všeobecně mutace typu 1 jsou nejčastěji recesivní,  
typu 2 dominantní

- u některých genů jsou známy oba typy (výsledkem je *alelová heterogenita*)

# Typy mutací dle funkce

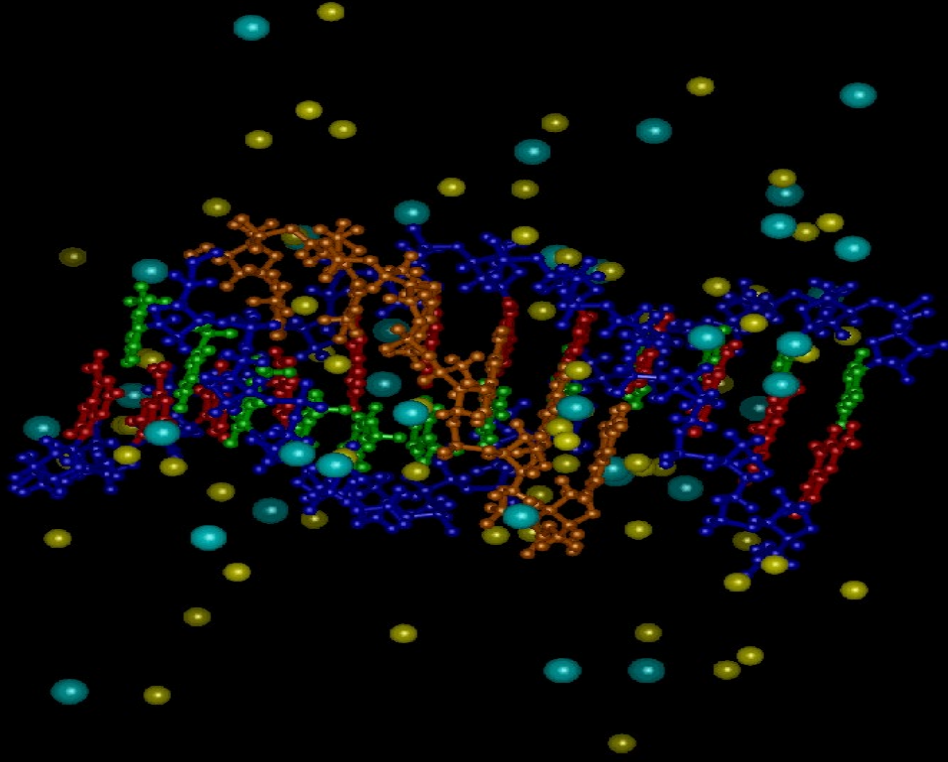
Nuliformní - žádný funkční protein  
- většinou recesivní  
- enzymopatie

Hypomorfní  
Hypermorfní  
Zisk nových vlastností



# Nomenklatura mutací

Human Gene Mutation Database:  
[www.hgmd.org](http://www.hgmd.org)



- **Substituce aminokyselin**  
(jedno- alebo trojprísmenové symboly)
  - R117H nebo Arg117His: substituce arg za his na AA 117
  - G542X: glycín 542 nahrazený stop-kodónem

## Substituce nukleotidů

(A z inic. kodónu ATG je vždy +1)

- 1162G>A: výmena G na pozicii 1162 za A
- IVS4+2G>T: záměna G za T v druhém bázi intronu 4

## Delece a inzerce

(symboly del resp. ins)

- F508del: delece fenylalaninu 508
- 622-626del: delece 5 nt začínající 622
- 409-410insC: vložení C mezi 409 a 410