

# Základní poznatky o struktuře genomu

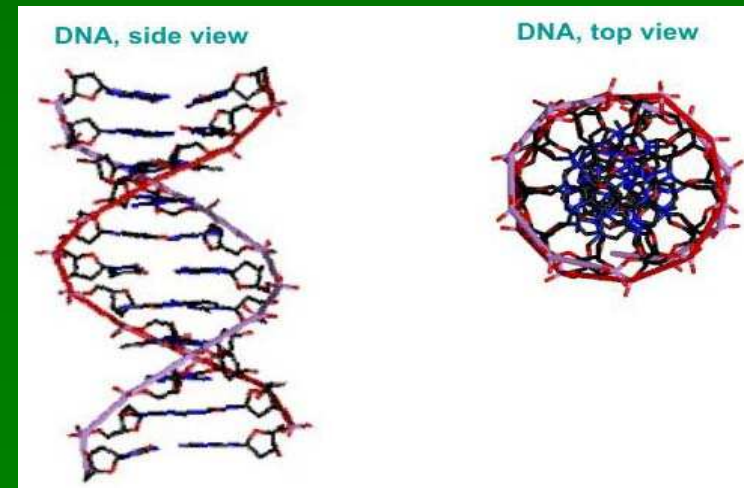
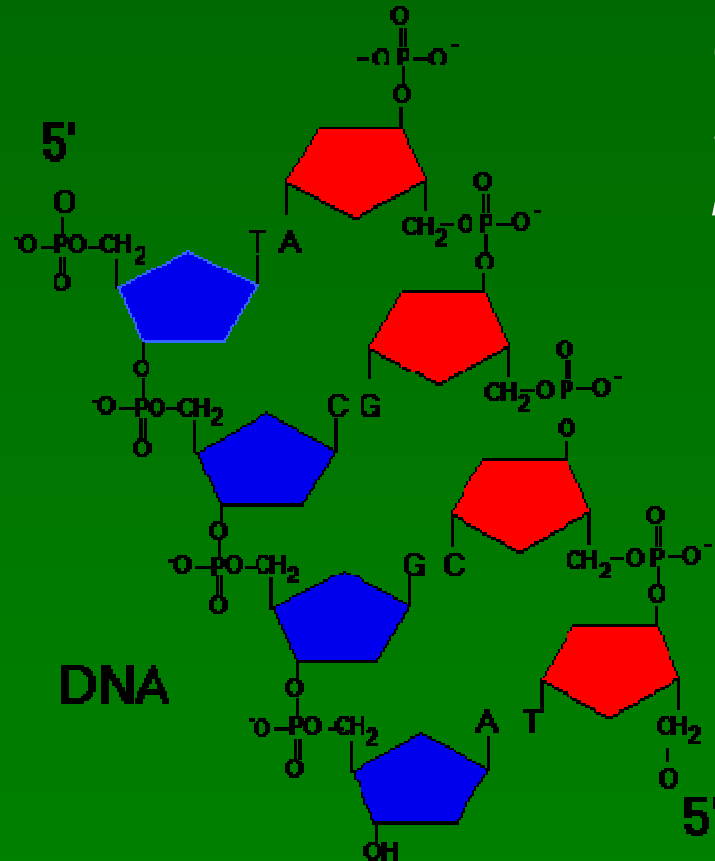
*Primární a sekundární struktura DNA*

*Nukleosom a chromatinové vlákno*

*Chromosomy a genom*

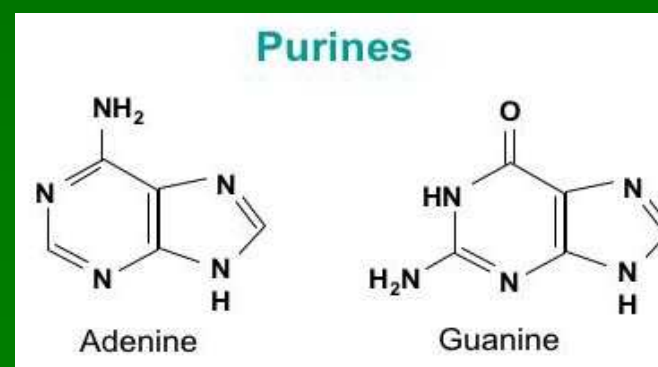
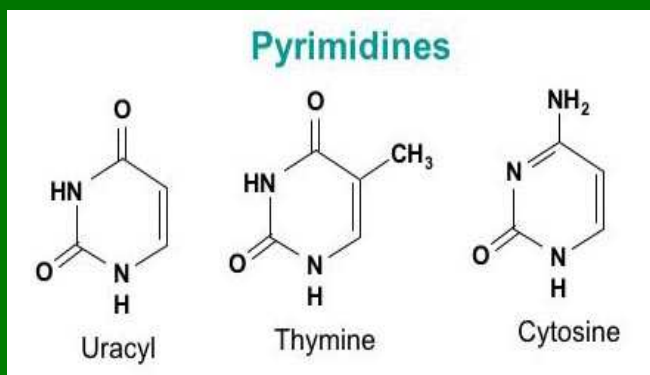
# Primární a sekundární struktura DNA

*DNA je tvořena nukleotidy (báze+cukr+fostátová skupina), které jsou uspořádány do řetězce. Dva řetězce pak tvoří dvoušroubovici.*



# Pyrimidiny a puriny v DNA

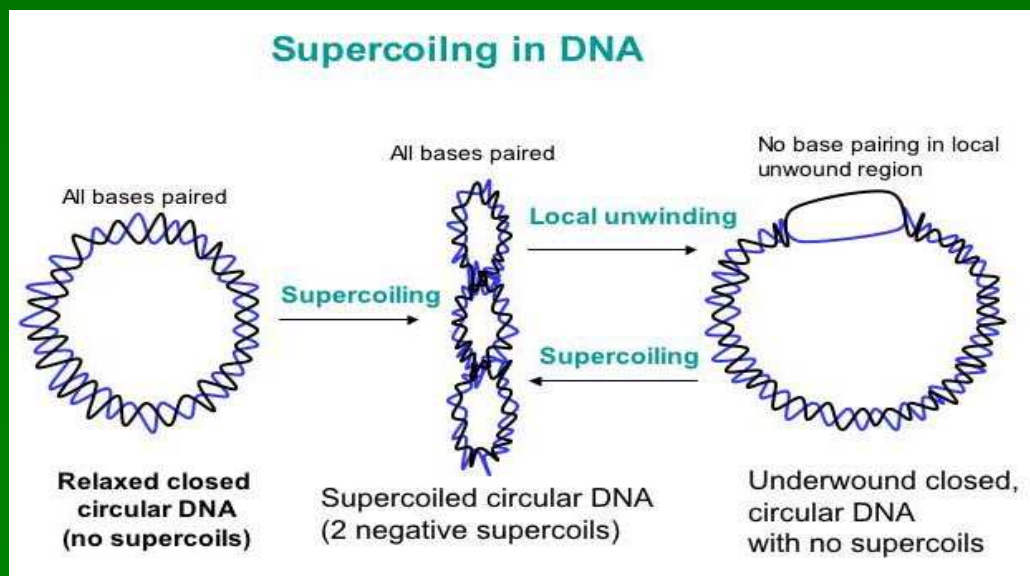
Báze nukleotidů jsou odvozeny od pyrimidinů nebo od purinů. Jsou to nenasycené sloučeniny s dvojitými vazbami, struktura je plošná, absorbují UV světlo. Hlavní pyrimidiny a puriny obsažené v nukleotidech jsou:



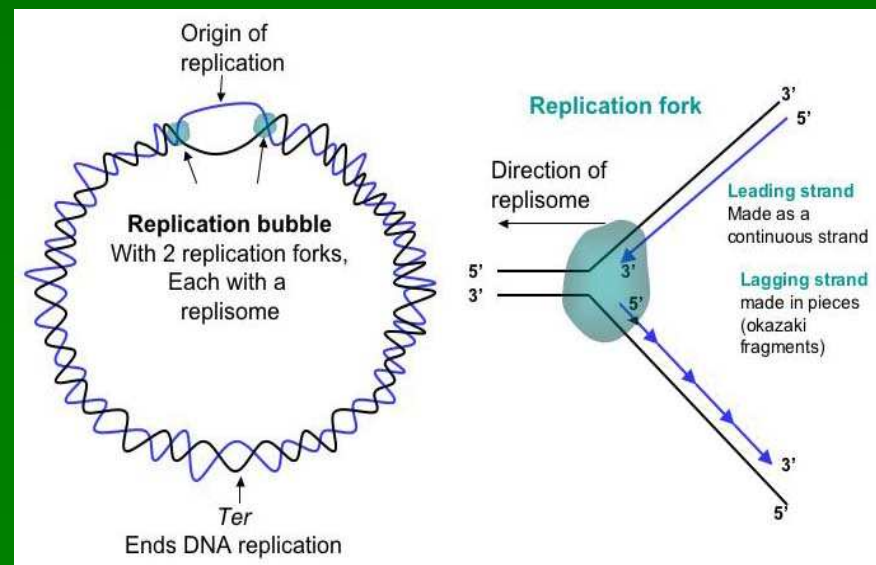
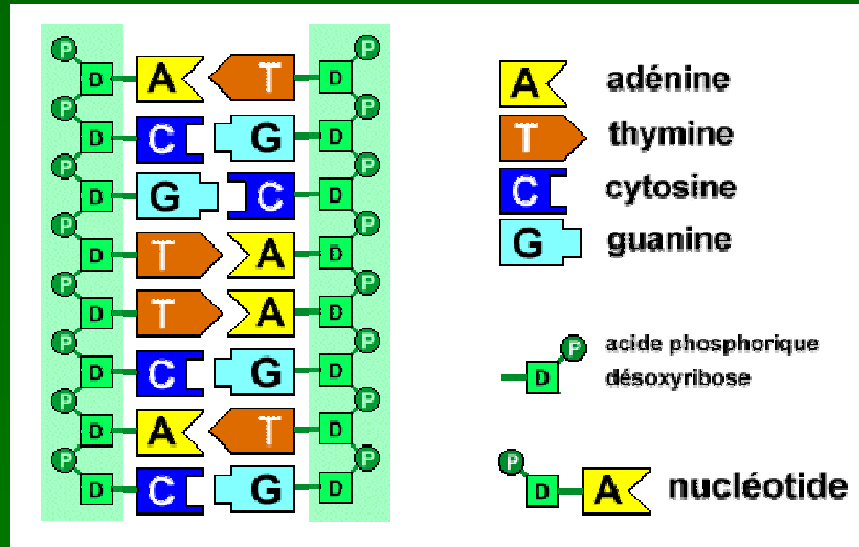
# Supercoiled DNA

Reálná DNA může mít větší nebo menší počet otáček než odpovídá její přirozené struktuře. To je kompenzováno vytvořením vyšší šroubovice (DNA supercoiling).

Kruhové genomy bakteriální DNA tvoří nadšroubovici (5 otáček na 1000 bp. Lineární chromosomy mohou mít také oblasti se supercoiled DNA.



# Komplementarita DNA umožňuje replikaci



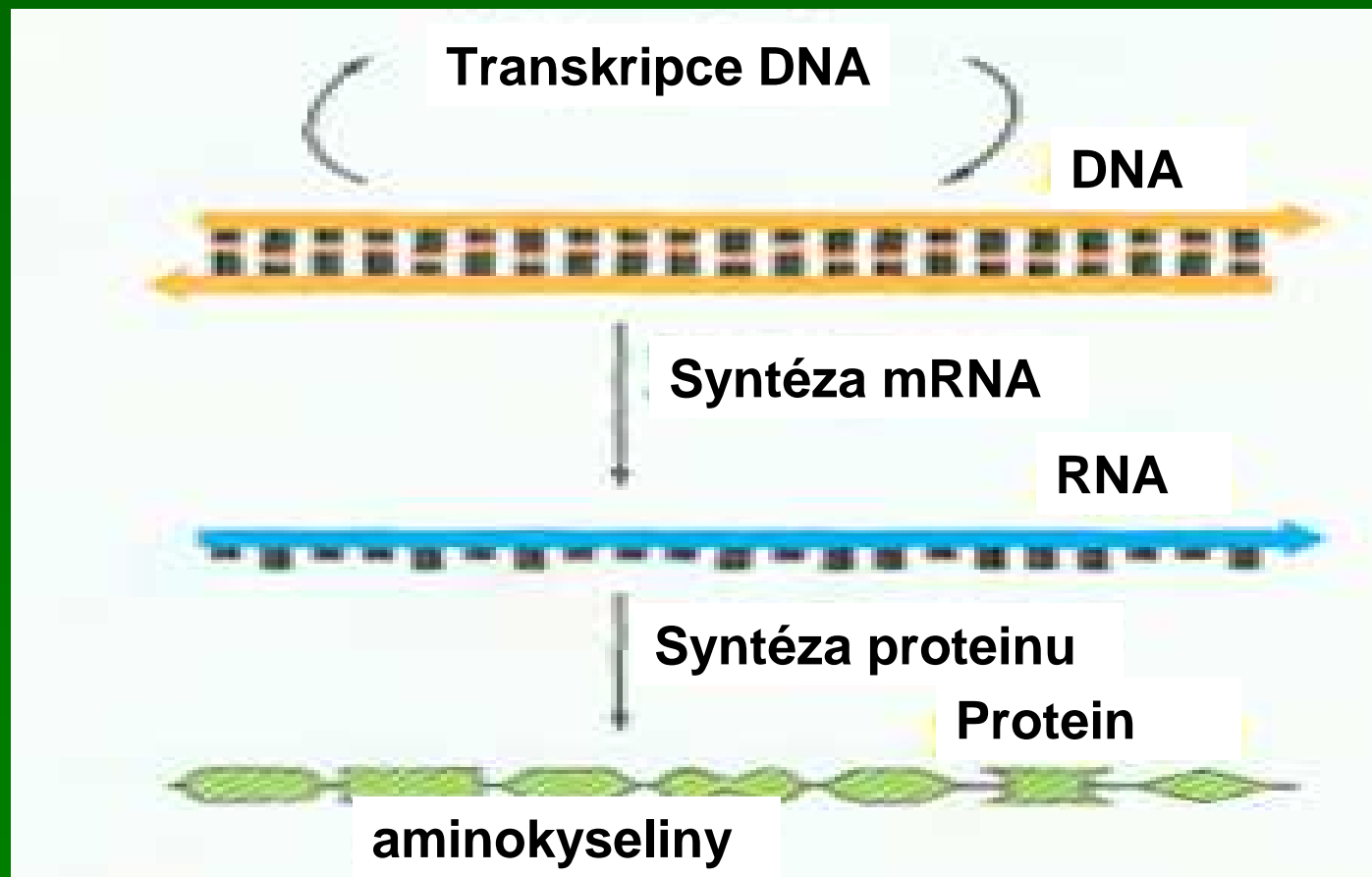
# DNA koduje aminokyseliny

Genetický kód je system pravidel, které přiřazují posloupnost nukleotidů v DNA aminokyselinám. Tento kód je využit v živých buňkách a je universální.

Vznik kódu je předmětem teorií:

- 1) afinita aminokyselin k jejich tripletům – možná se dříve syntéza děla přímo, tj. bez tRNA
- 2) vývoj prostřednictvím selekce tak, aby byl co nejmenší počet chyb (mutací)

# Od DNA k proteinu



# mRNA koduje aminokyseliny

1st position (5' end) ↓	2nd position				3rd position (3' end) ↓
	U	C	A	G	
U	Phe Phe Leu Leu	Ser Ser Ser Ser	Tyr Tyr STOP STOP	Cys Cys STOP Trp	U C A G
C	Leu Leu Leu Leu	Pro Pro Pro Pro	His His Gln Gln	Arg Arg Arg Arg	U C A G
A	Ile Ile Ile Met	Thr Thr Thr Thr	Asn Asn Lys Lys	Ser Ser Arg Arg	U C A G
G	Val Val Val Val	Ala Ala Ala Ala	Asp Asp Glu Glu	Gly Gly Gly Gly	U C A G

## Kód je

-degenerovaný  
(existuje  $4^3=64$  komb.)

-skryté mutace

-stop kodon (UAG,  
UGA, UAA)

-důležitá je fáze čtení

-variace:

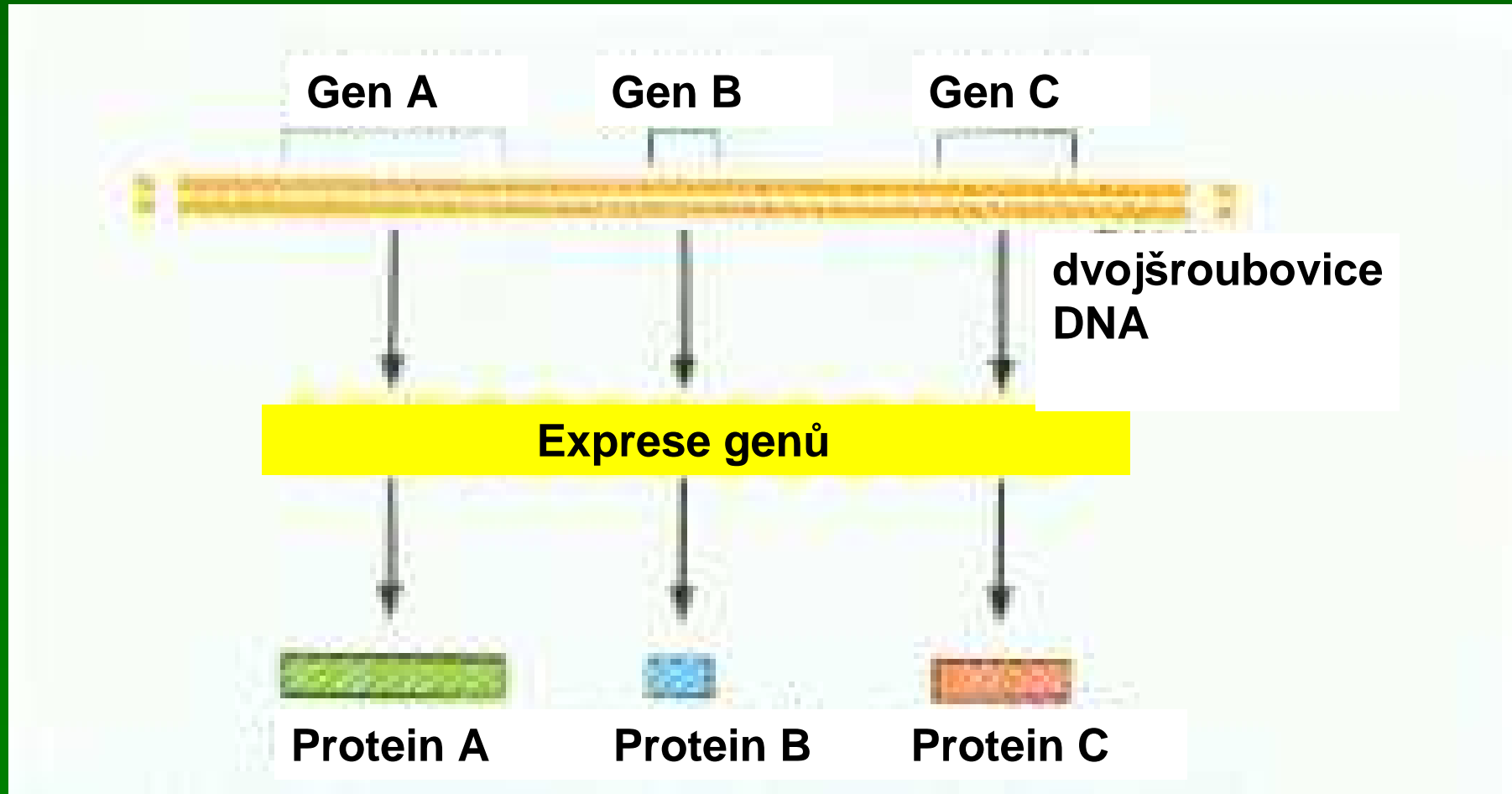
-mykoplasma (UGA  
je tryptofan)

-bičíkovci (UAG je  
glutamin)

-candida (CUG –  
Serin)



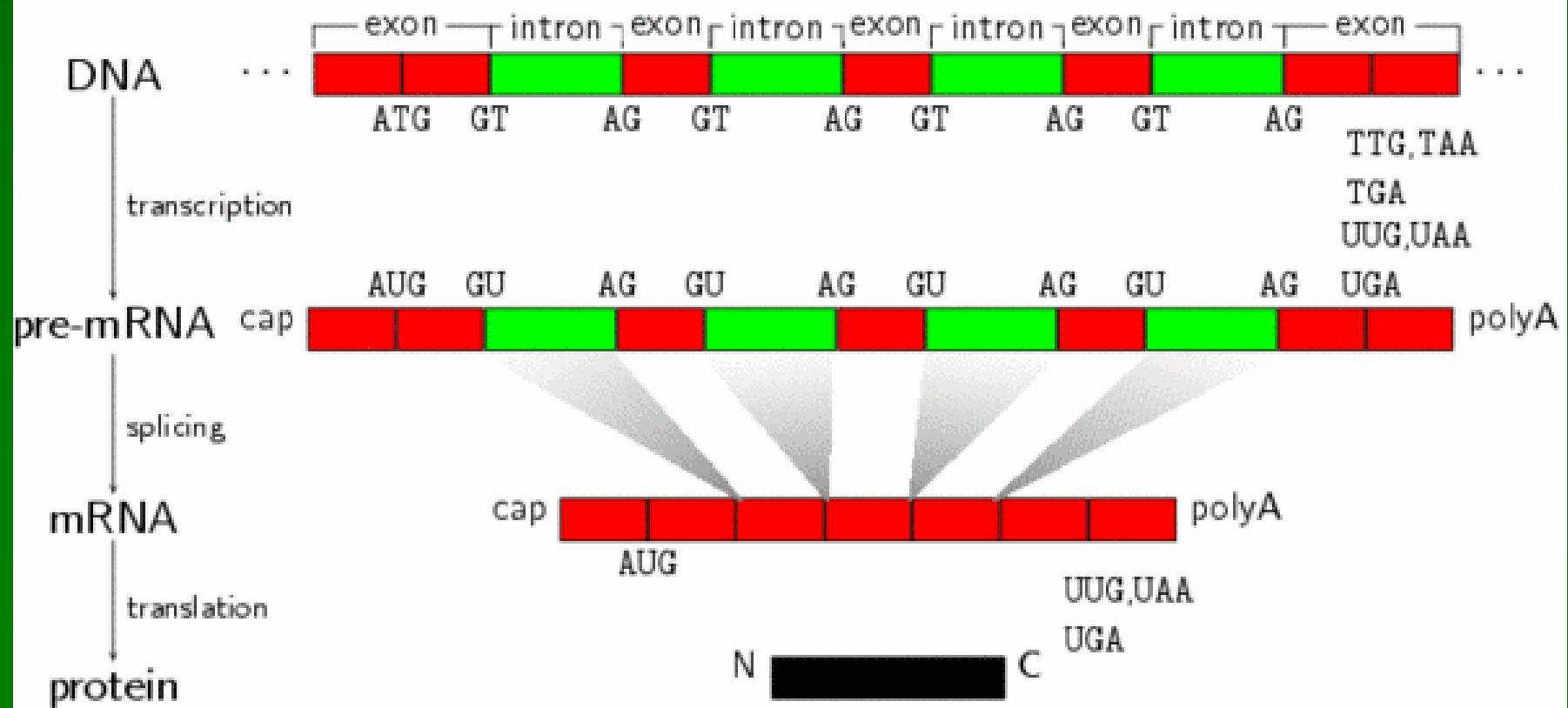
# Genetická informace v DNA



# Sestřih – u eukaryot

**Splice sites** are locations on DNA at boundaries of

- **exons** (which code for proteins)
- **introns** (which do not)

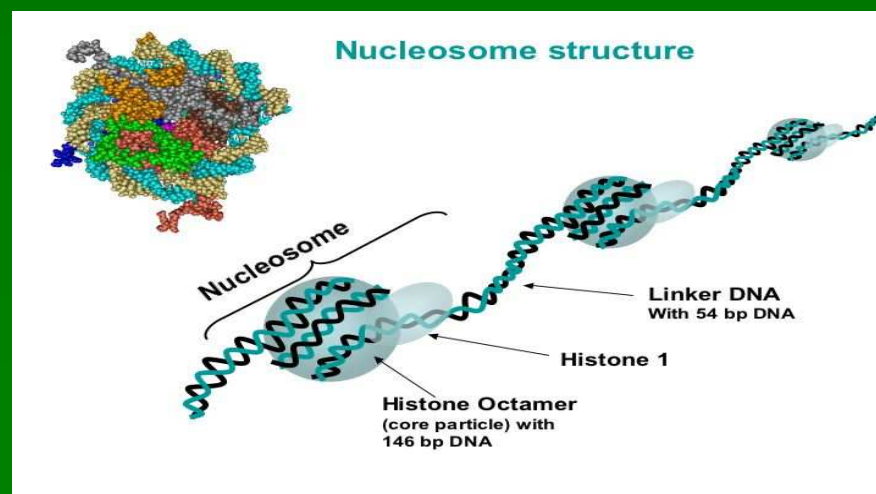


# DNA jako součást chromatinu

V eukaryotických jádrech – chromatin=DNA + histony + další proteiny. Histony (H2A a H2B) a (H3 a H4) tvoří dimery, ty tvoří tetramery a dohromady pak vzniká oktamer – kolem se vine DNA (146 bp) a vytváří tak nukleosom. Spojnice = 56 bp.

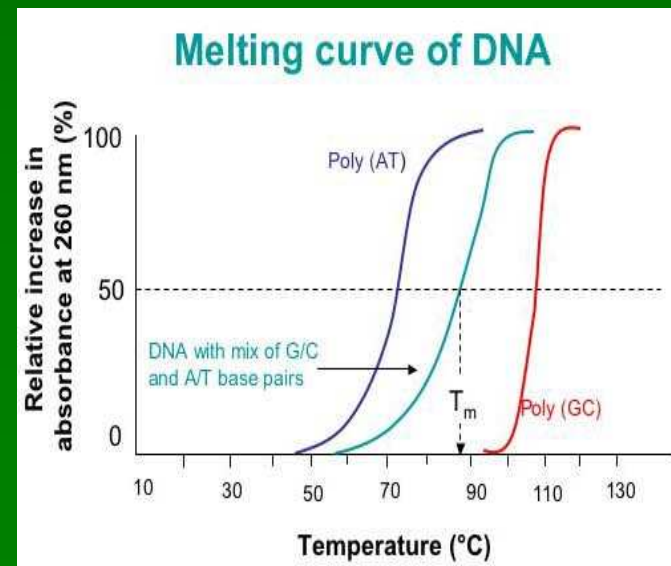
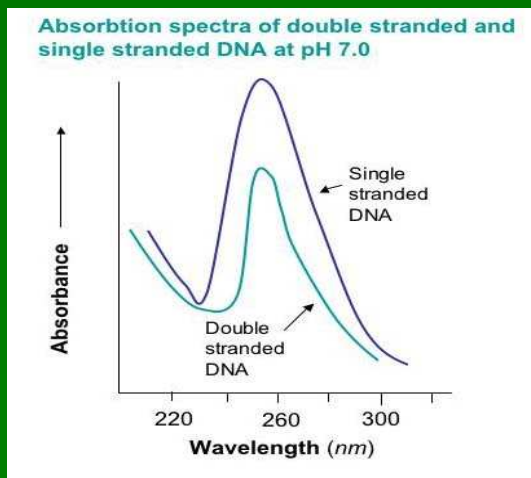
Histony jsou zásadité proteiny, bohaté na lysin a arginin (R-skupiny mají kladný náboj, který interaguje se záporným nábojem fosfátových skupin DNA).

Histony jsou vysoce konzervovány u eukaryot.

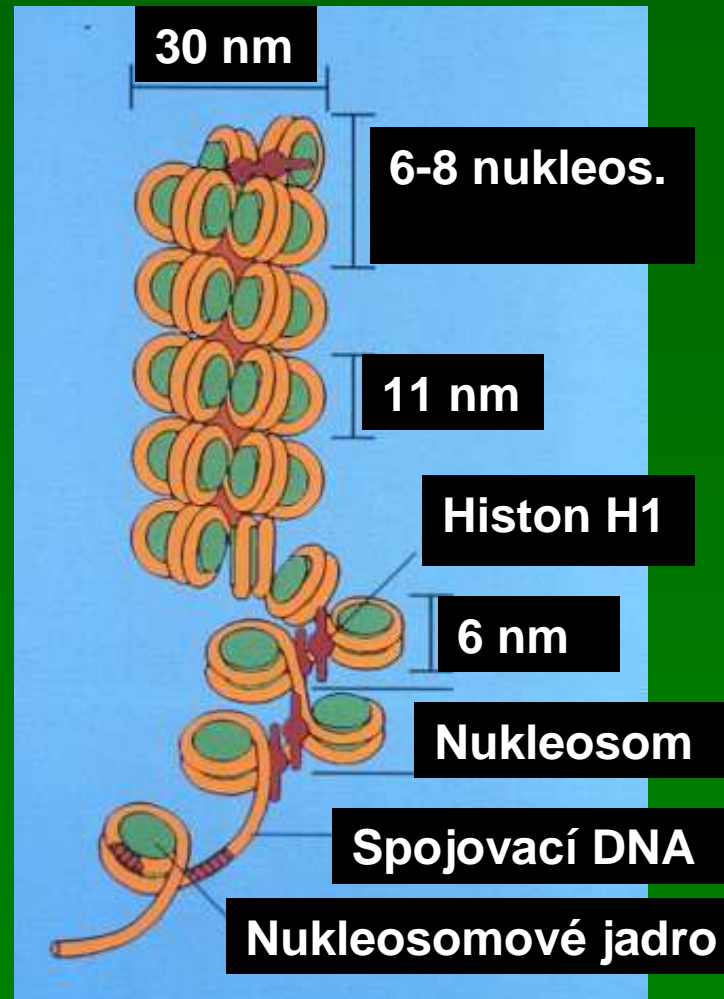


# Denaturace DNA

*Je to oddělení komplementárních řetězců. Může nastat teplotou (95°C – kompletní). Denaturační teplota ( $T_m$ ) – teplota, při které je polovina DNA oddělená. Denaturace je doprovázená zvýšením absorpce UV. Denaturační křivka – závislost absorpance na  $T$ .  $T_m$  závisí na obsahu GC párů.*

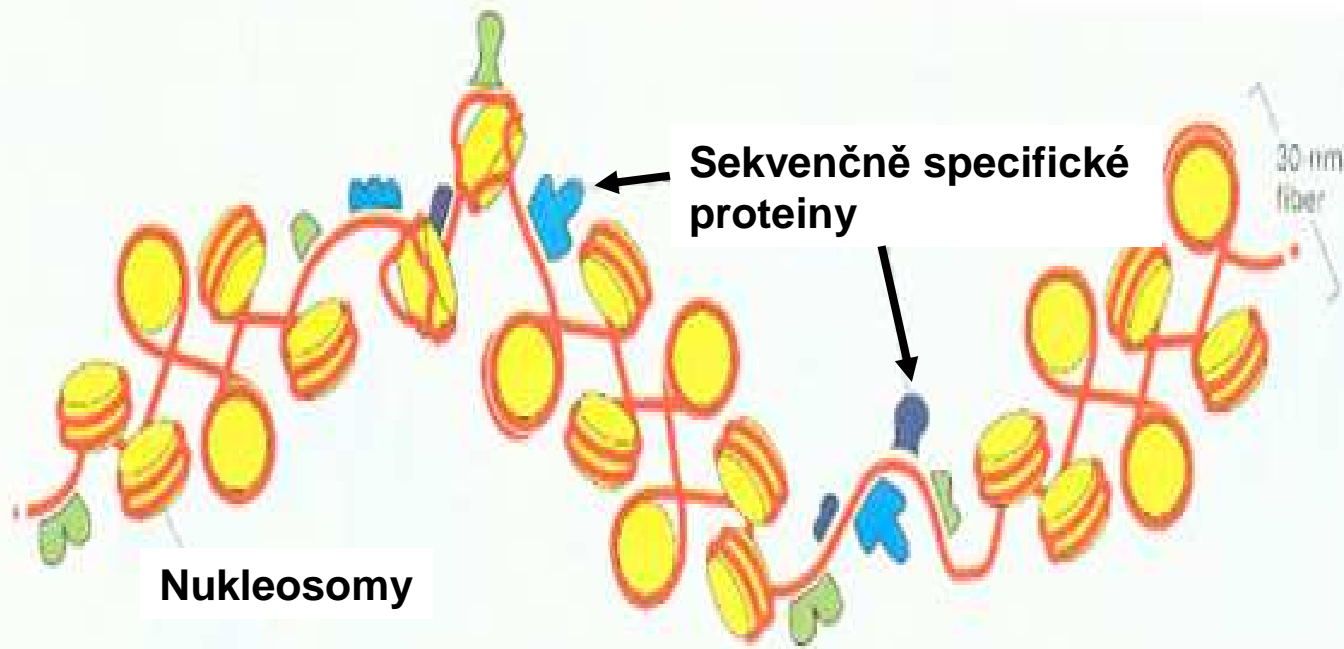


# Chromatinové vlákno

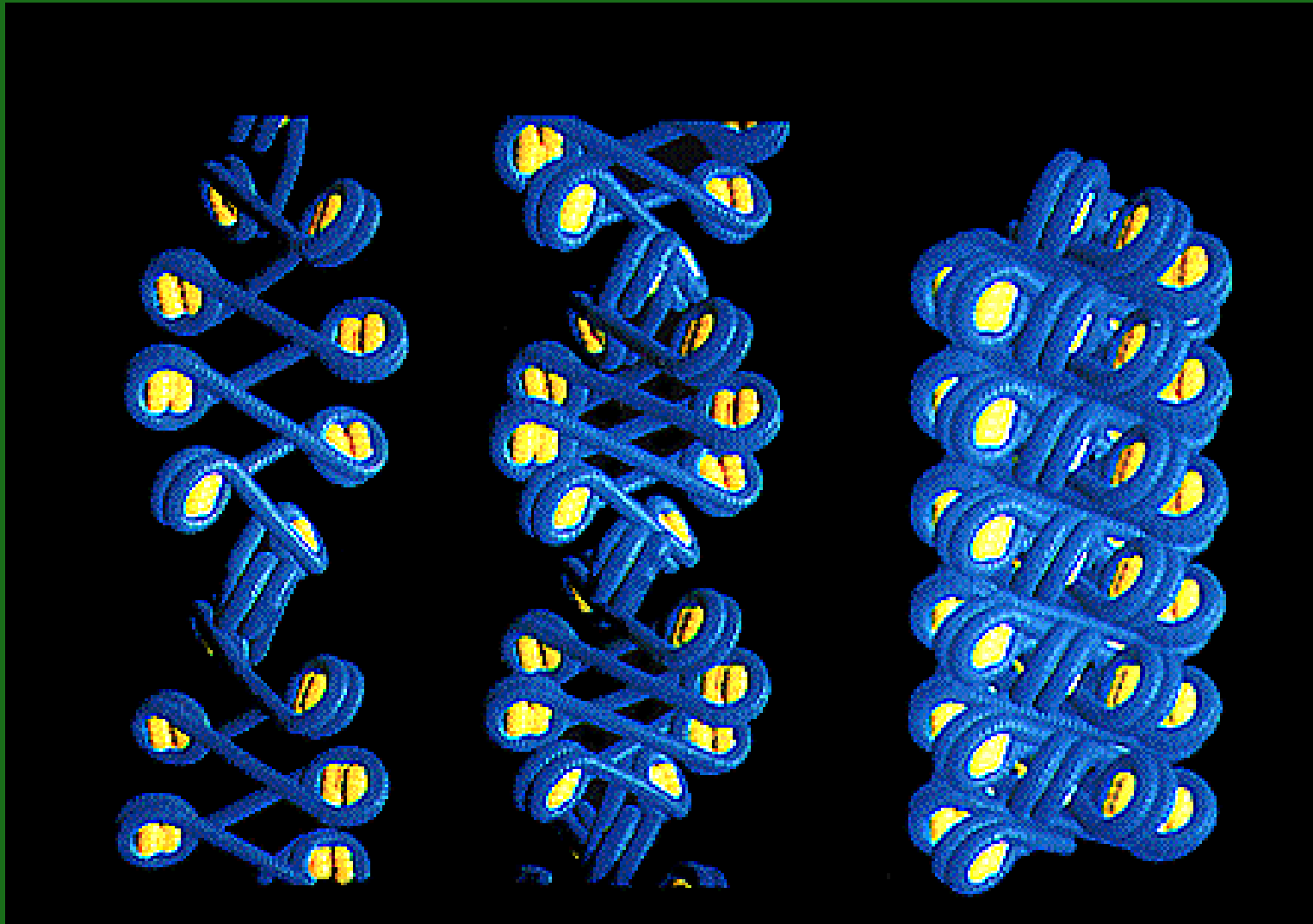


# Chromatinové vlákno

30-nm chromatinové vlákno s nukleosomy a navázanými proteiny



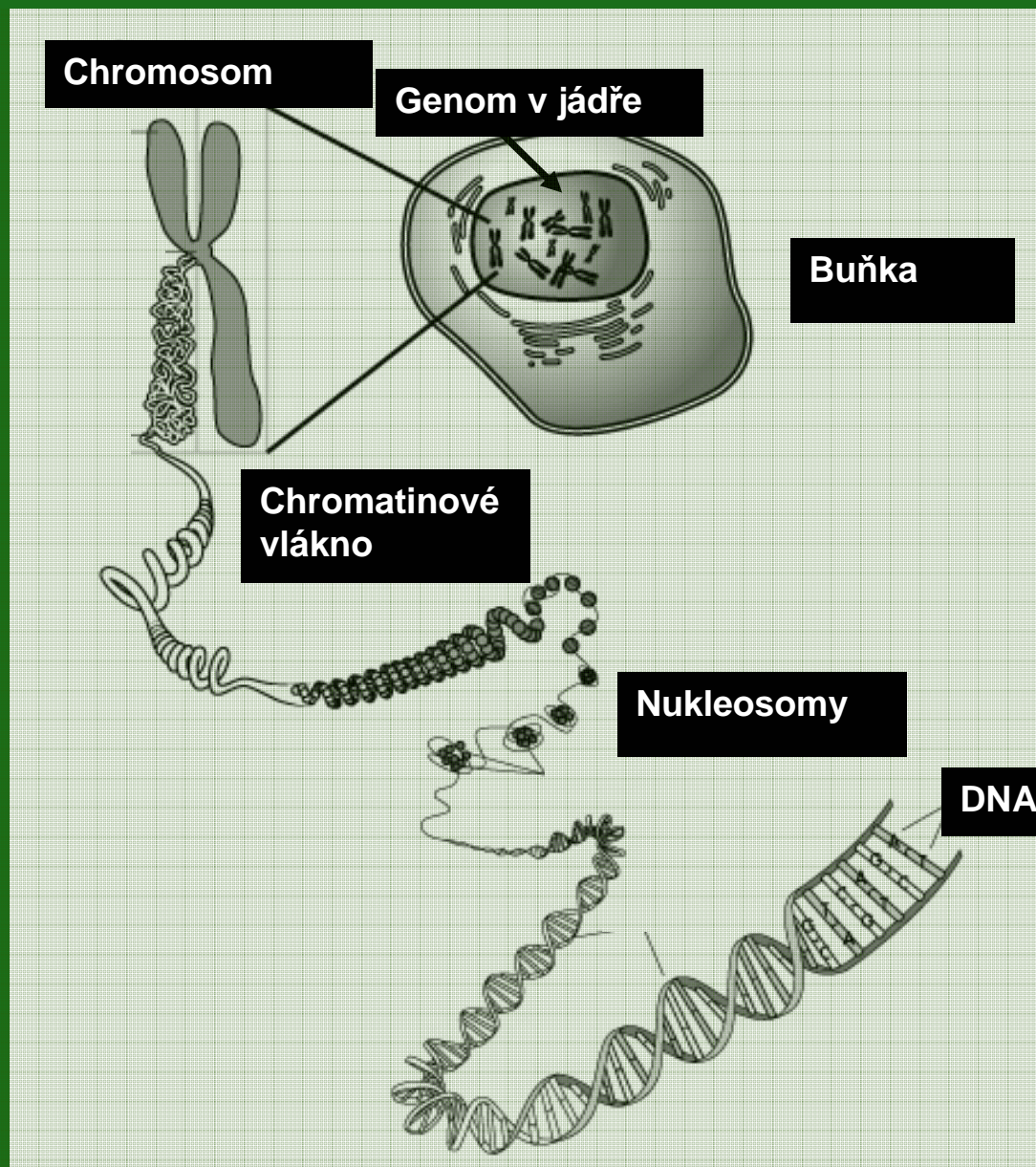
# Struktura chromatinu



Roztažená a částečně roztažená  
forma chromatinu

Kondensovaná forma  
forma chromatinu

# DNA, nukleosomy, chromatin, chromosom a genom





# RNA v buňce

**Ribosomální RNA (rRNA)** 80% celkové RNA,  
část ribosomů

**Transferová RNA (tRNA)** – nese aktivovanou  
aminokyselinu na ribosom

**Messengerová RNA (mRNA)** – nese info z  
DNA na ribosomy pro odpovídající protein

**Malá RNA** – regulace exprese

# Chromosomy

*Úvod - mitotické a interfázní chromosomy*

*Barvení chromosomů (klasika, FISH)*

*Významné elementy chromosomů*

*CT a jejich části jsou disjunktní*

*Subdomény CT*

*Chromosom jako náhodný polymer*

*Nenáhodná vnitřní struktura CT*

# Chromosomy-složení

*Chromosom obsahuje jedinou molekulu DNA spojenou s proteiny:*

- a) 5 druhů molekul histonů – proteiny bohaté na lysina a arginin – proto s pozitivním nábojem. Proto se vážou silně s negativně nabitými fosfátovými skupinami DNA*
- b) různé proteiny nehistonové povahy- většinou jsou to transkripční faktory*

# Chromosomy-struktura

*Interfáze vs mitóza – stupeň kondenzace, v mitóze dobře viditelné*

*V S-fázi dochází k syntéze DNA, a proto v mitóze pozorujeme chromosomy duplikované, složené ze dvou sesterských chromatid spojených v centromerické oblasti*

*V oblasti centromery jsou přítomny proteiny kinetochory, které se vážou na vlákna dělicího vřeténka*

*Ramena se označují p-raménko (kratší) a q-raménko*

*Barvení Giemsou zviditelní sérii pruhů (tzv. G-pruhy)*

# Chromosomy-počty

*Všechna zvířata mají charakteristický počet chromosomů ve svém těle, tzv. diploidní počet ( $2n$ )*

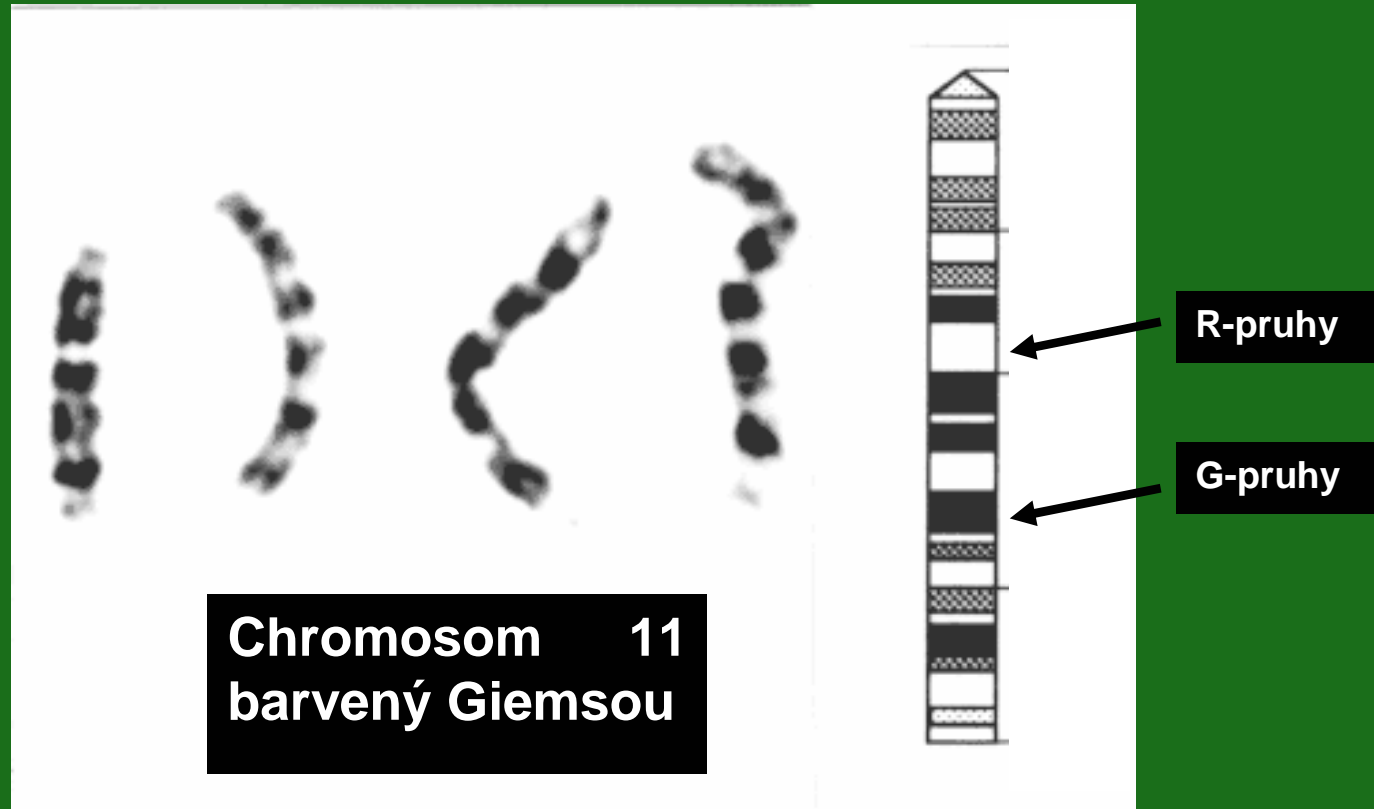
*Ty se vyskytují jako homologní páry, každý z páru byl získán od jednoho z rodičů*

*Pohlavní buňky obsahují haploidní počet chromosomů*

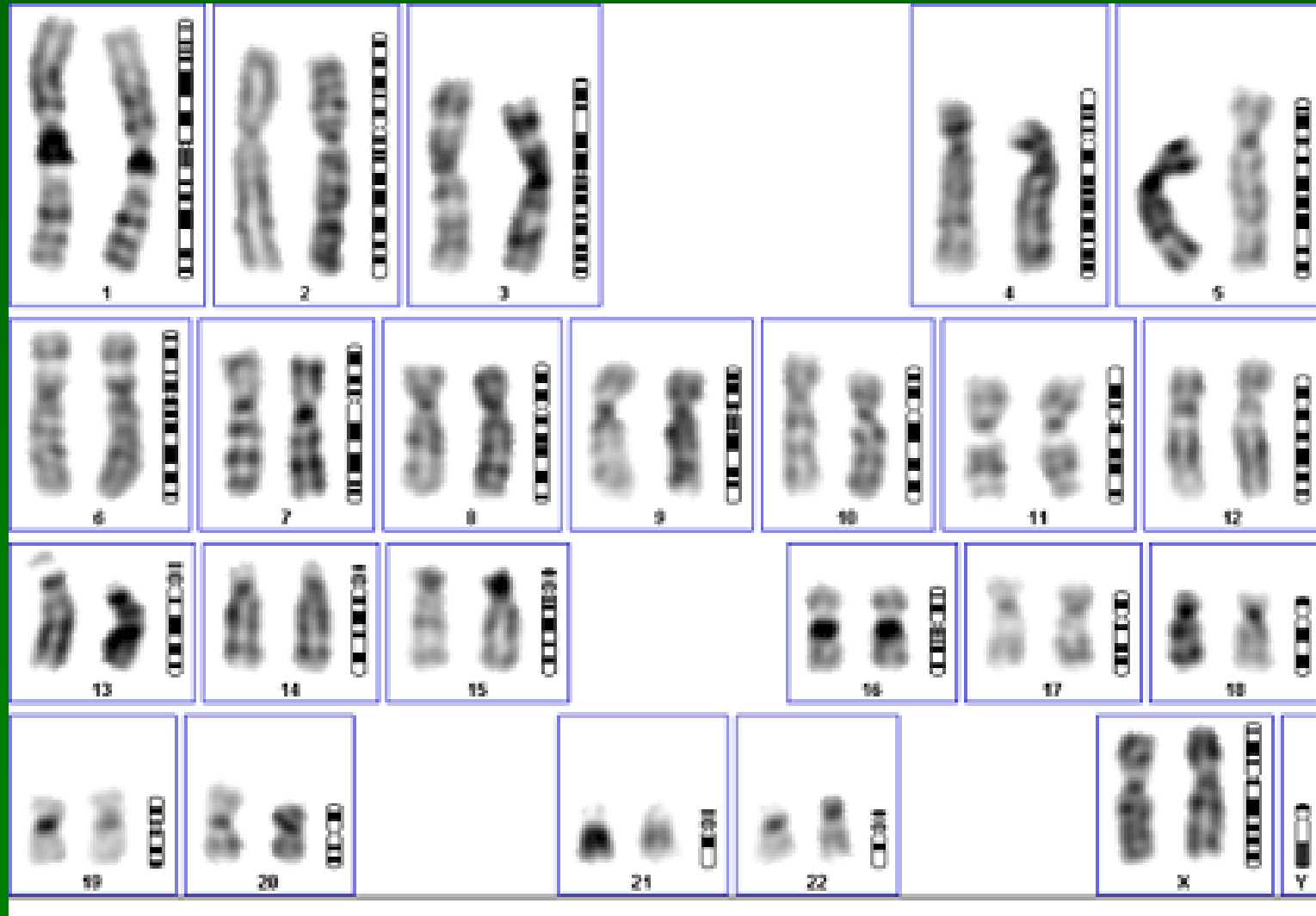
# Chromosomy-počty

Člověk	46
Myš	40
<i>Drosophila melanogaster</i>	8
Mikroskopický červ <i>Caenorhabditis elegans</i>	12
Kvasinky ( <i>Sacharomyces cerevisie</i> )	32
<i>Arabidopsis thaliana</i> (rostlina)	10
Žába <i>Xenopus laevis</i> (jižní Afrika)	36
Pes	78
Kuře	78
Pšenice	20
Mravenec <i>Myrmecia pilosula</i>	2
Červ <i>Parascaris equorum</i>	2
Rak <i>Cambarus clarkii</i>	200
Přeslička rolní <i>Equisetum arvense</i>	216

# Klasické barvení chromosomů

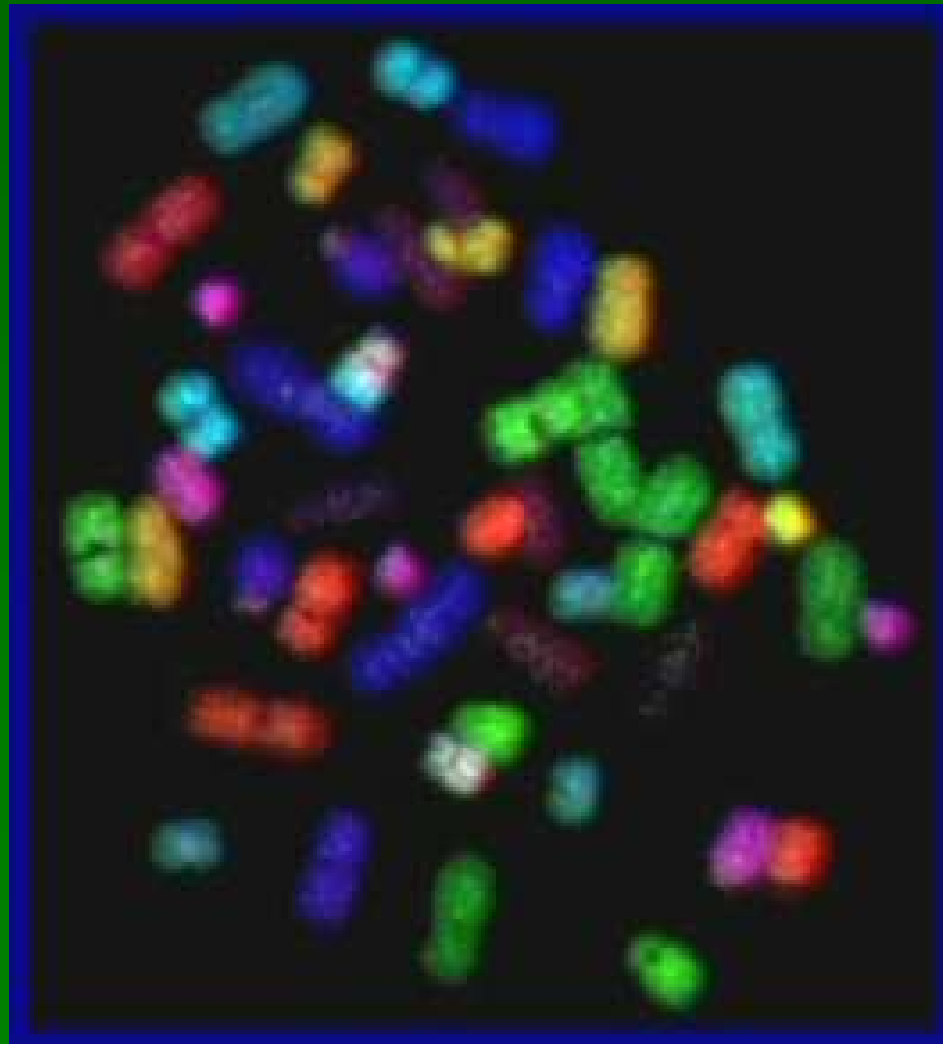


# Barvení chromosomů - G-banding

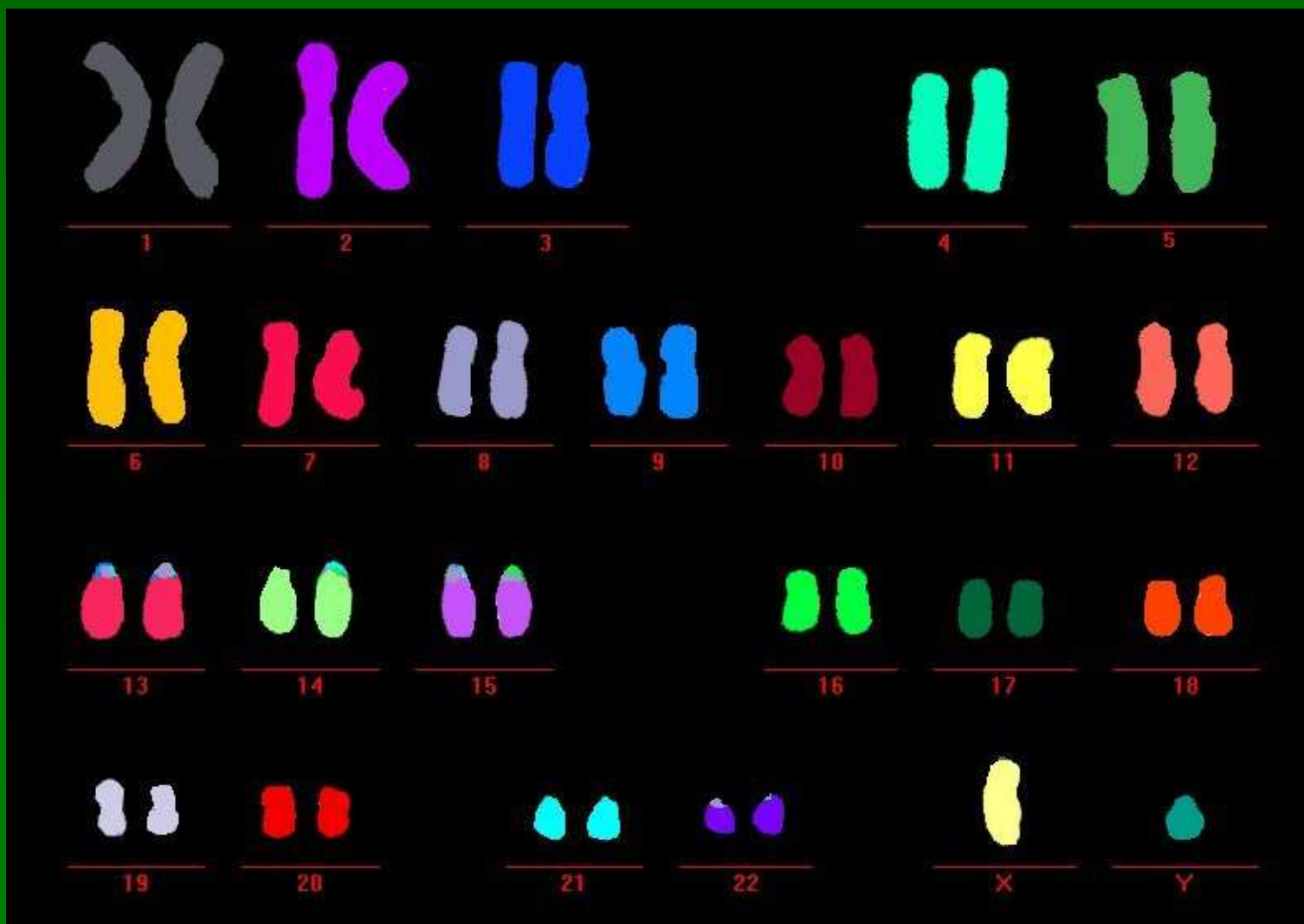




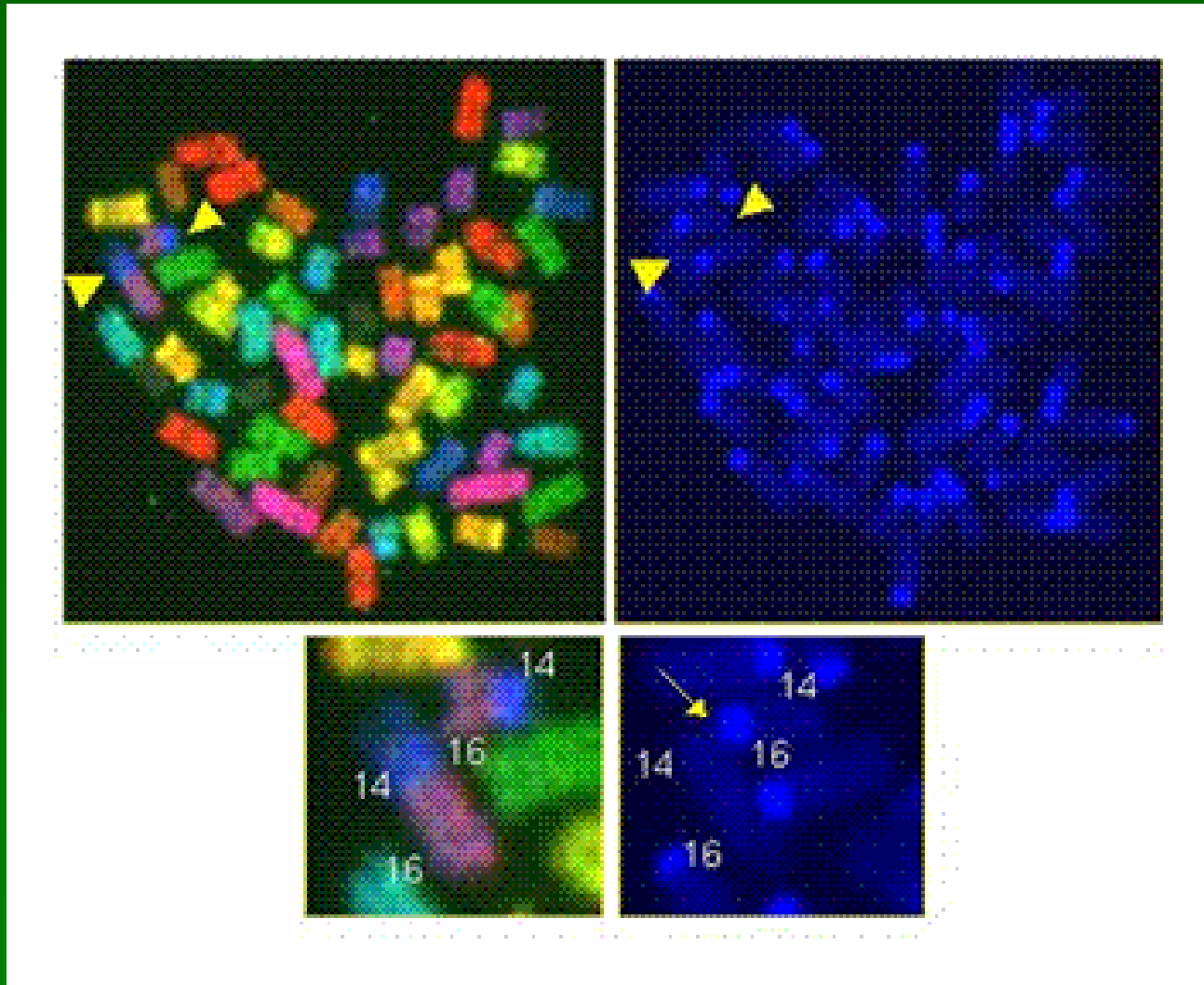
# Lidské chromosomy obarvené pomocí multicolor FISH



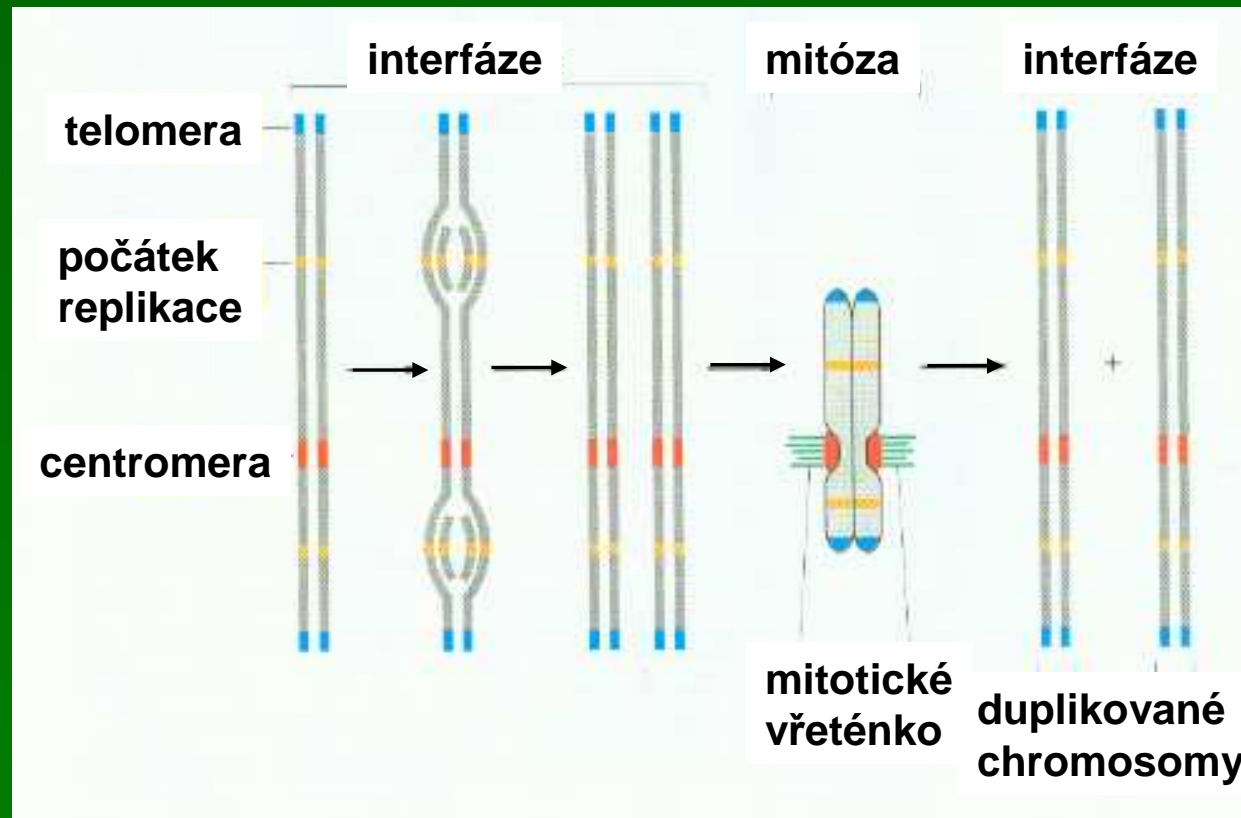
# Lidské chromosomy obarvené pomocí multicolor FISH



# Translokace detekovaná pomocí m-FISH



# Významné sekvence na chromosomech



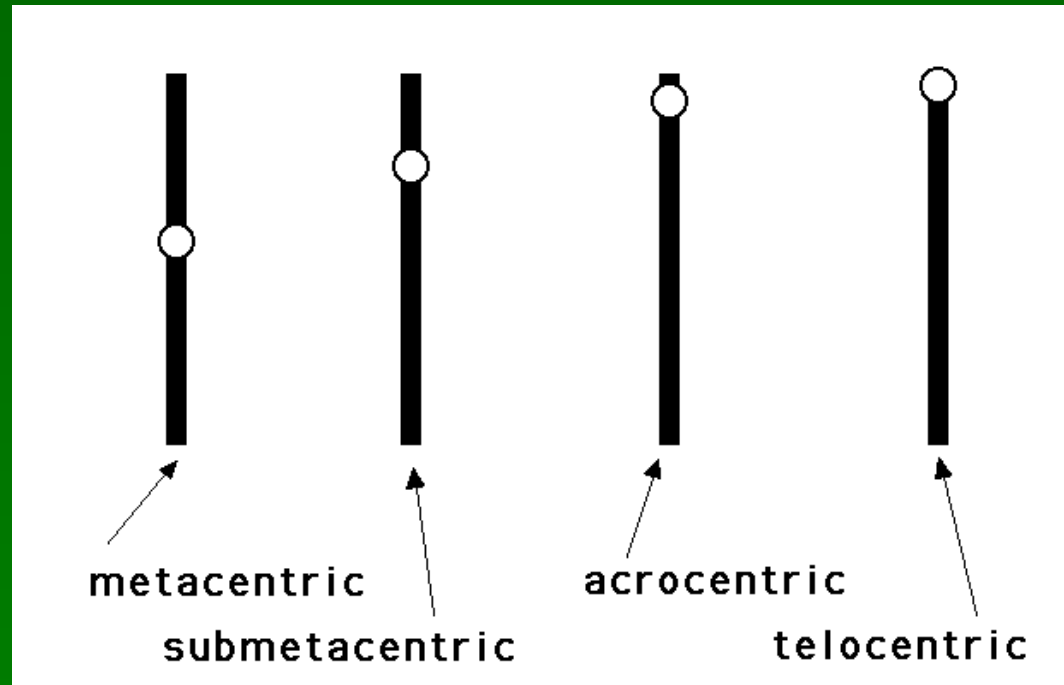
Na chromosomech rozlišujeme:

**telomery** (konce chromosomů obsahující sekvence zajišťující zachování délky chromosomu)

**centromery** (účastní se při segregaci chromosomů v mitóze)

**počátky replikace** (více míst, kde začíná replikace)

# Typy chromosomů podle umístění centromery



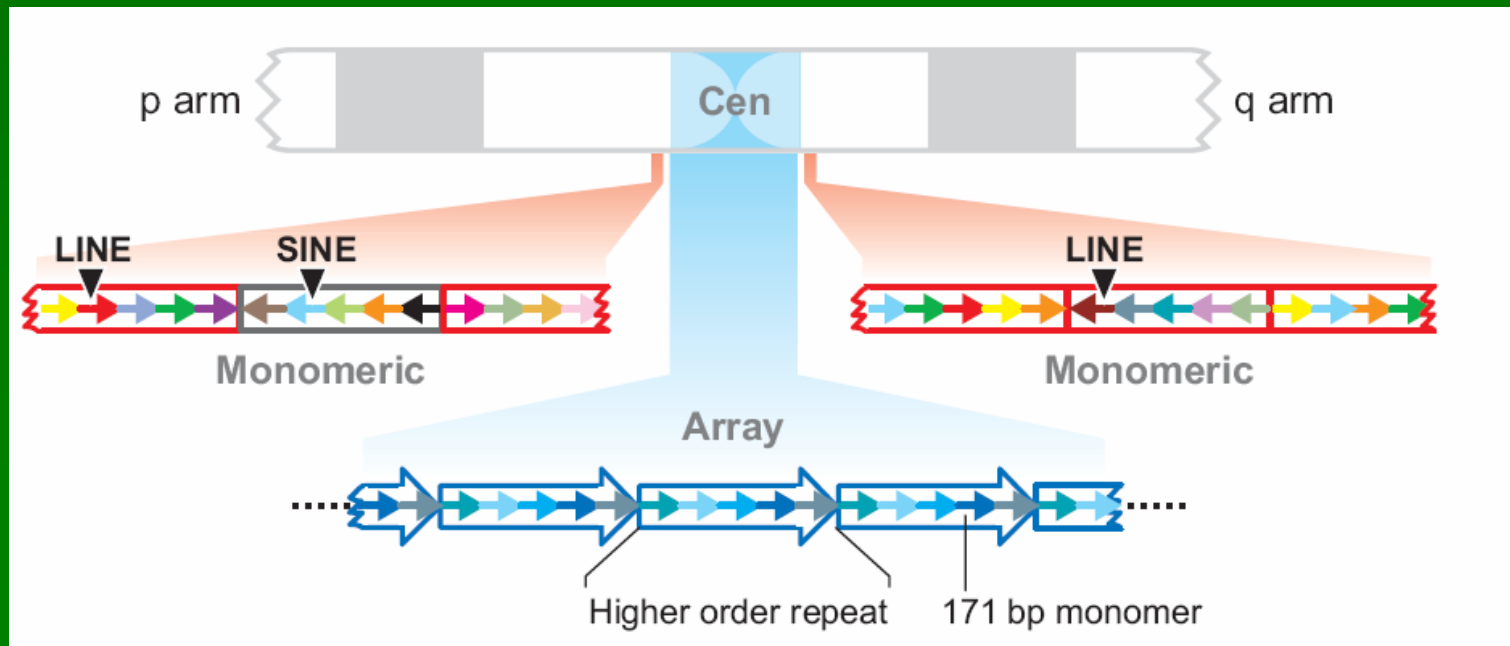
**Centromery** – identifikovatelné jako místa konstriktce, v nichž jsou chromatidy spojeny k sobě, odpovědné za segregaci CT, fixovány na určité místo v CT, mají určitou sekvenci. Tyto sekvence mohou být stejné nebo podobné pro více CT.

# Centromery mají odlišné histony

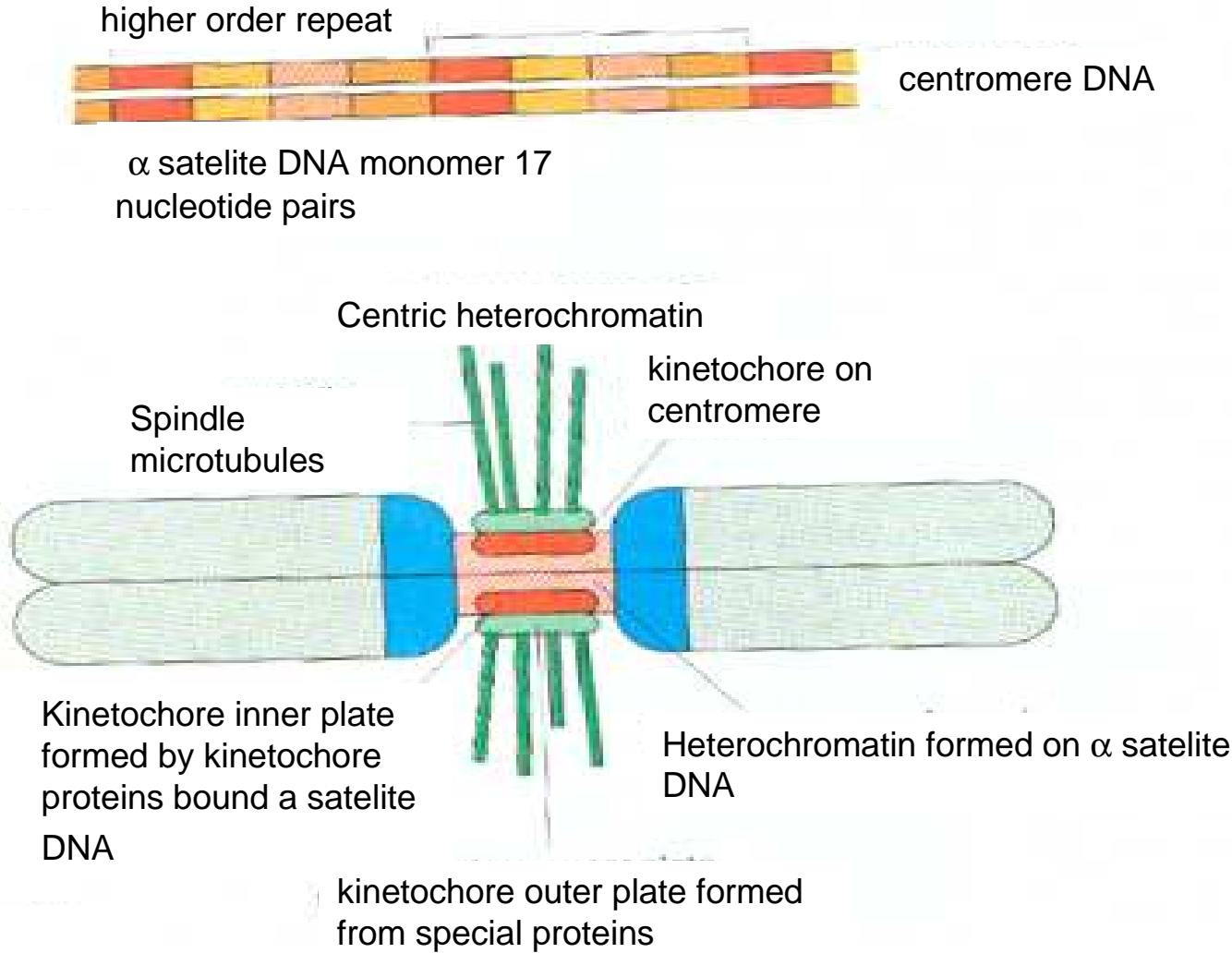
Centromery – jsou tvořeny alfa-satelitní DNA, obsahují kinetochorové proteiny, zajišťují připojení chromosomu k vřeténku a dělení chromosomu do dceřiných buněk.

Centromery obsahují specifický histon CENPA, mají schopnost vytvářet chromocentra a účastní se regulace exprese.

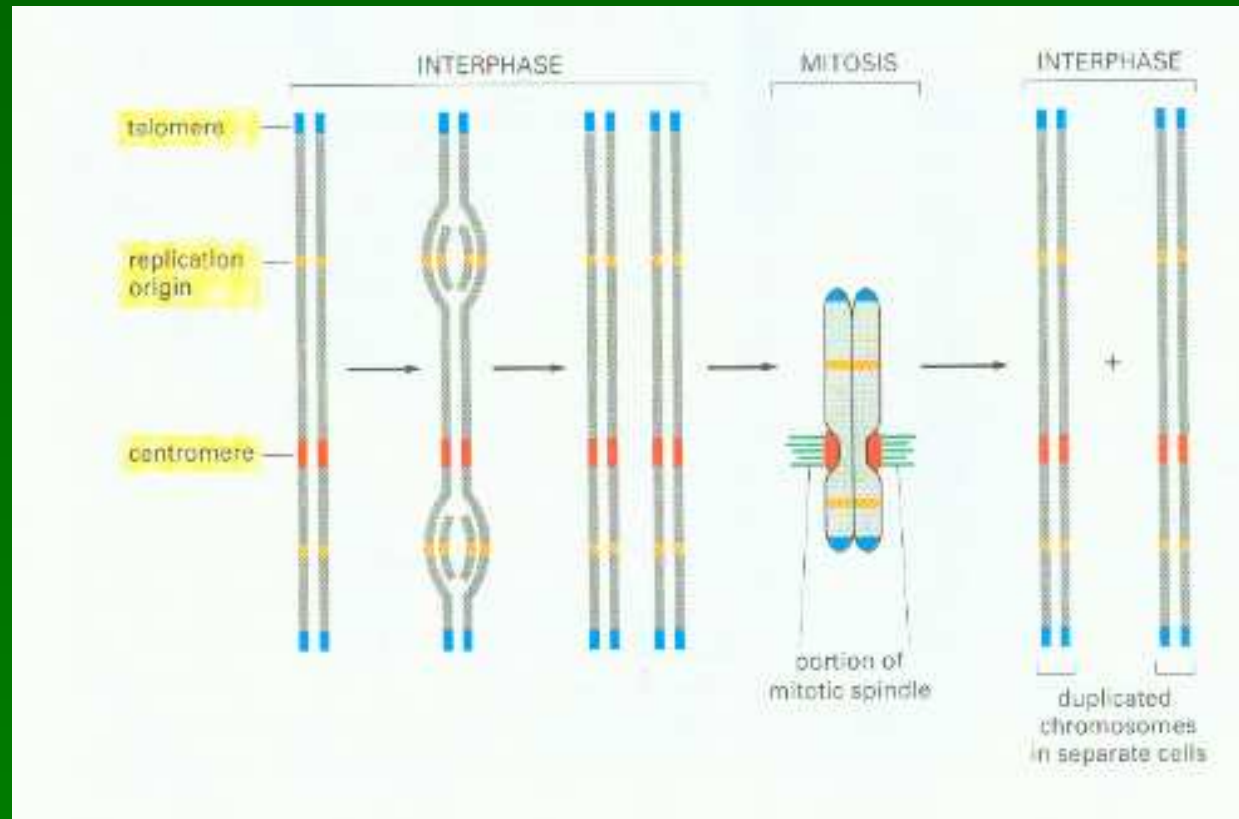
Na obr. pericentromerické satelity (šipky). Bloky repetice obsahují občas LINEs a SINEs (interspersed elements). Orientace mohou být stejné nebo opačné. Centromerické oblasti obsahují alfa satelity (obsahují 5 monomerů) – modře.



# Struktura centromery



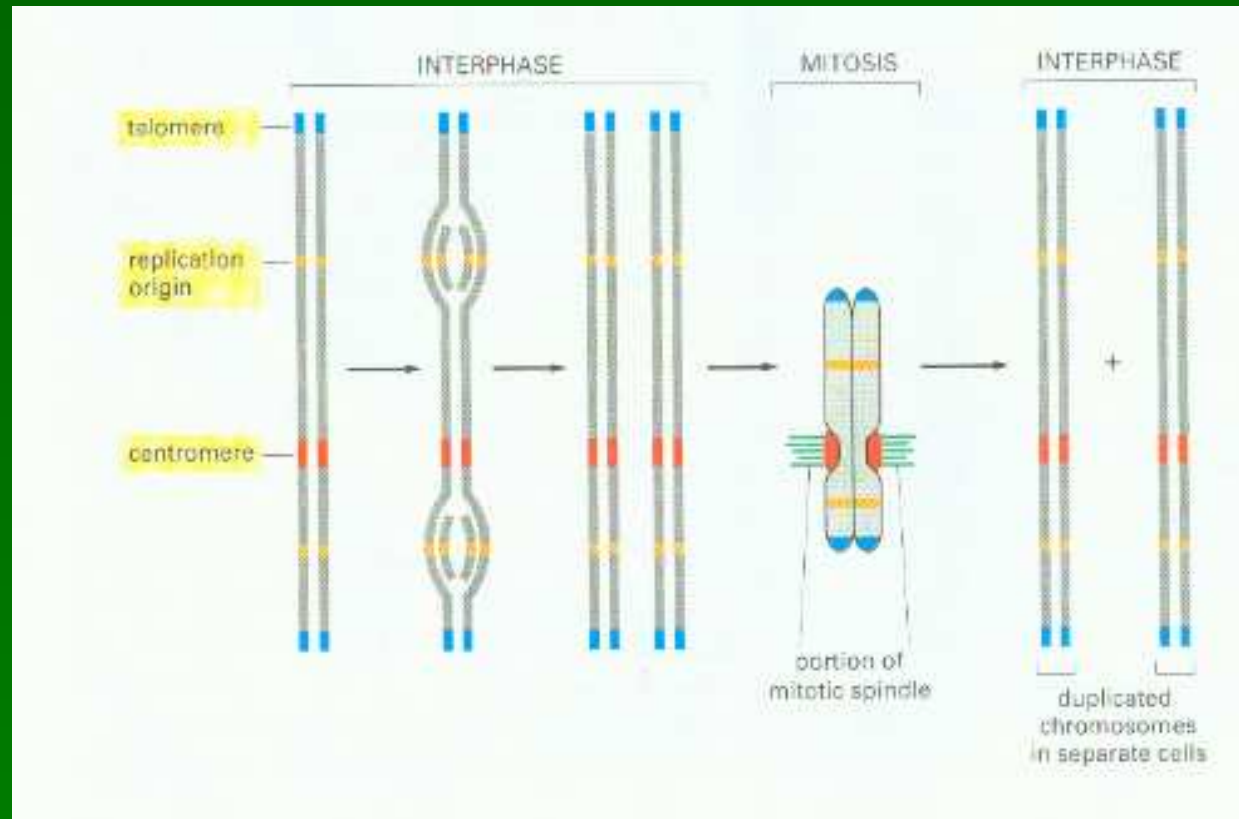
# Významné sekvence na chromosomech



**Telomery** – problém při replikaci na konci lineárních molekul DNA se řeší speciální sekvencí (tandem GGGTTA – až 10 kbp), která váže enzym telomerázu, tj. enzym, který funguje v komplexu s krátkým úsekem RNA, podle kterého syntetizuje (jako reverzní transkriptáza) malé úseky DNA na koncích chromosomů.

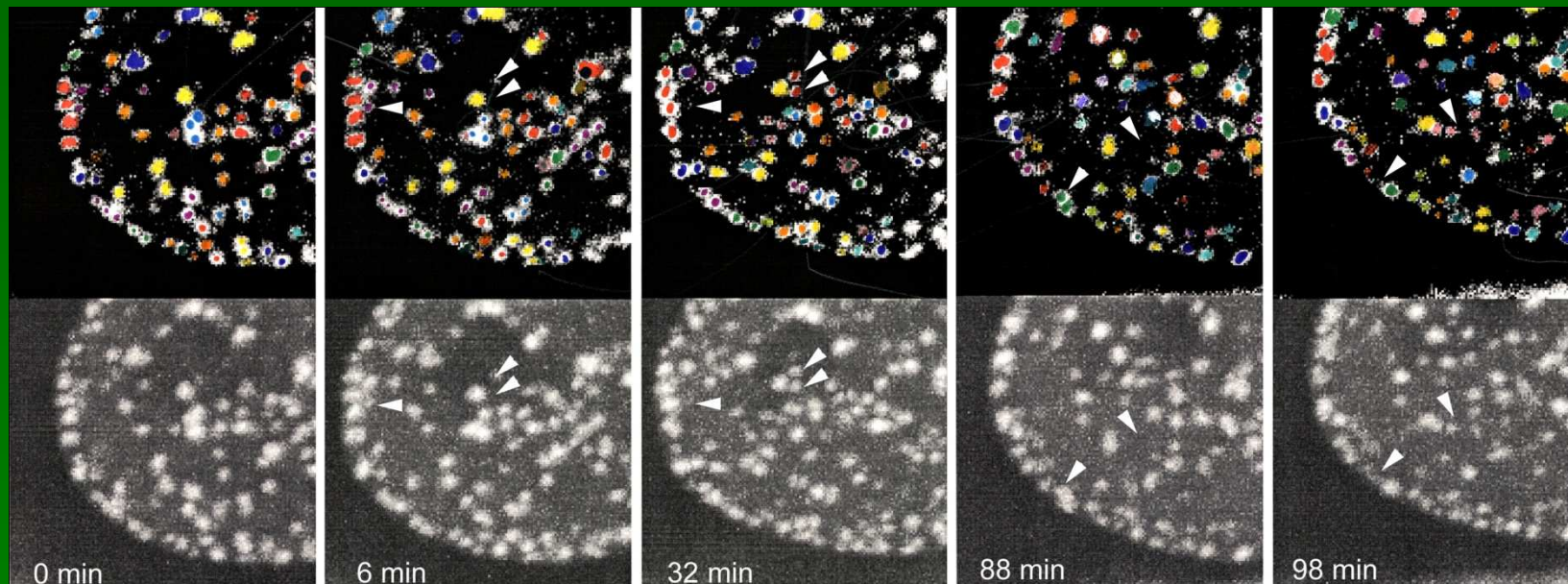


# Významné sekvence na chromosomech



**Počátky replikace** – replikace běží v obou směrech od společného počátku (50 bp/s), který je charakterizován určitým typem sekvence; v lidském genomu je mnoho počátků replikace ( $10^3$ ). Replikace je aktivována v klastrech a probíhá nejdříve pro méně kondenzovaný chromatin.

# Vizualizace více DNA sekvencí *in vivo* pomocí GFP

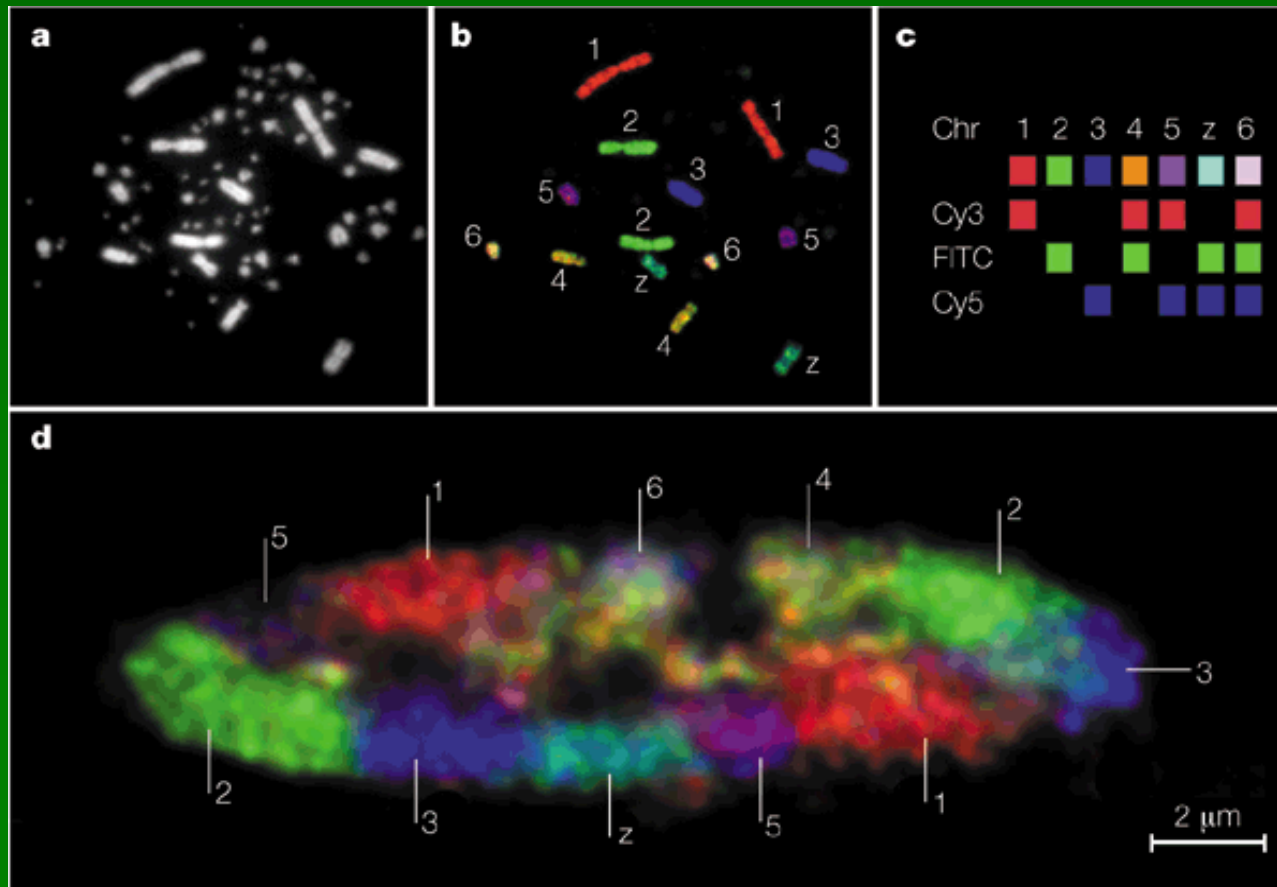


Sadoni et al., JCB 2000



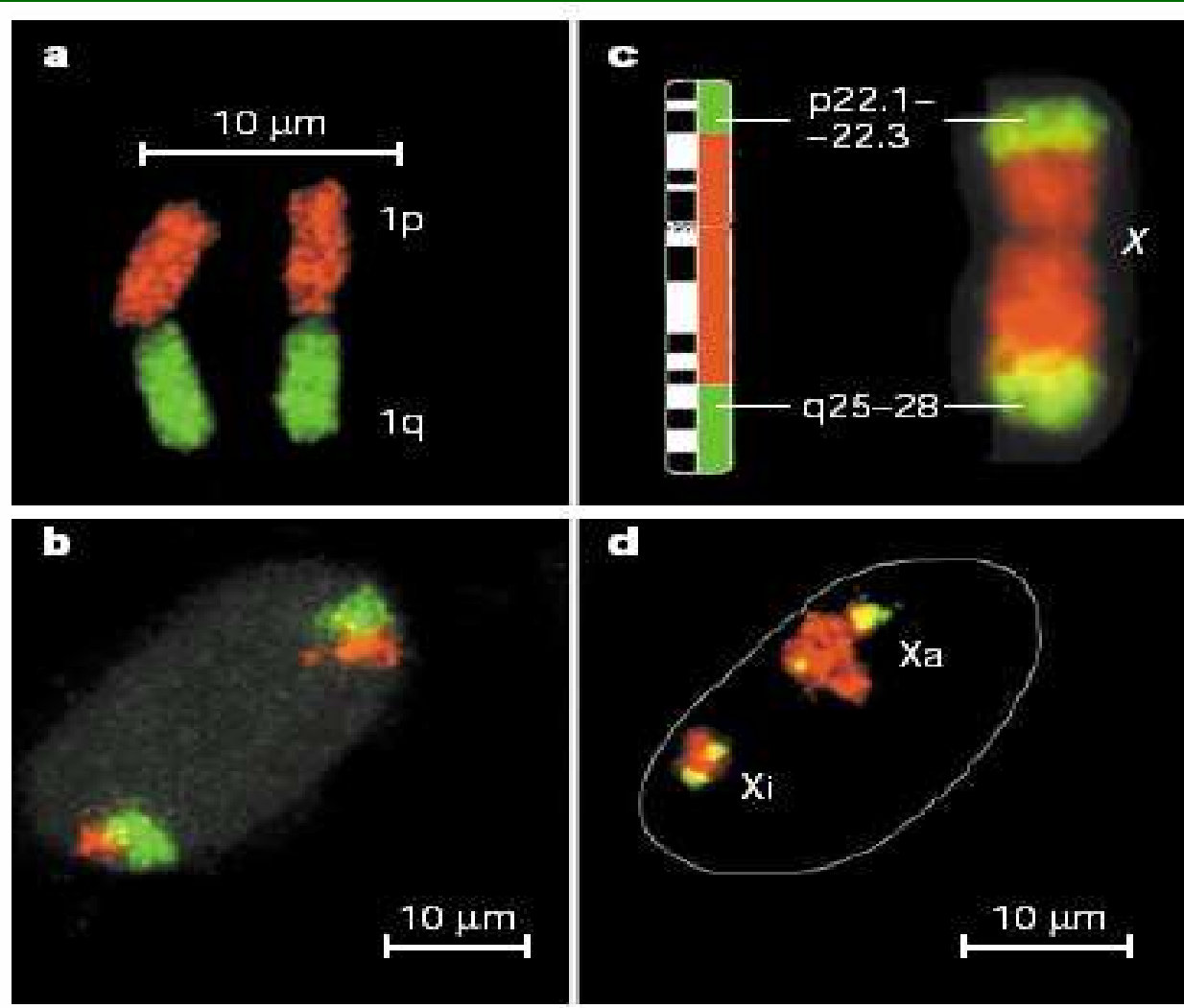
# Chromosomová teritoria (CT)

První experimenty, které vedly k závěru, že chromosomy se nacházejí v jádře v podobě ohraničených domén, byly pokusy T. Cremera v letech 1982-1984. Zavedení FISH podstatně urychlilo poznání chromosomů jak v mitóze, tak v interfázi.



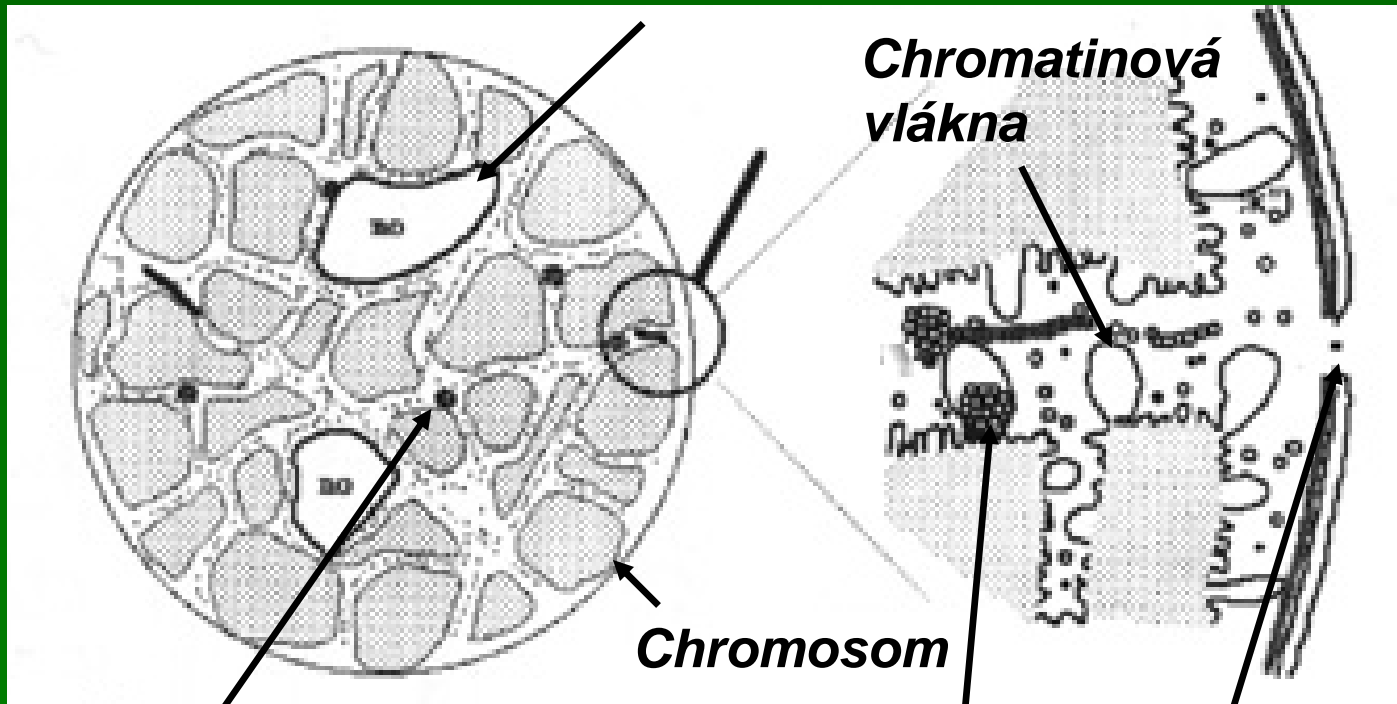


# Chromosomová teritoria se nepřekrývají; ani jejich části



# Funkční uspořádání genomu v jádře

*Jadérko*



*Chromatinová vlákna*

*Chromosom*

*Jaderné pory*

*Proteinový komplex*

*Transkripce*

# Chromosomy

*Úvod - mitotické a interfázní chromosomy*

*Barvení chromosomů (klasika, FISH)*

*Významné elementy chromosomů*

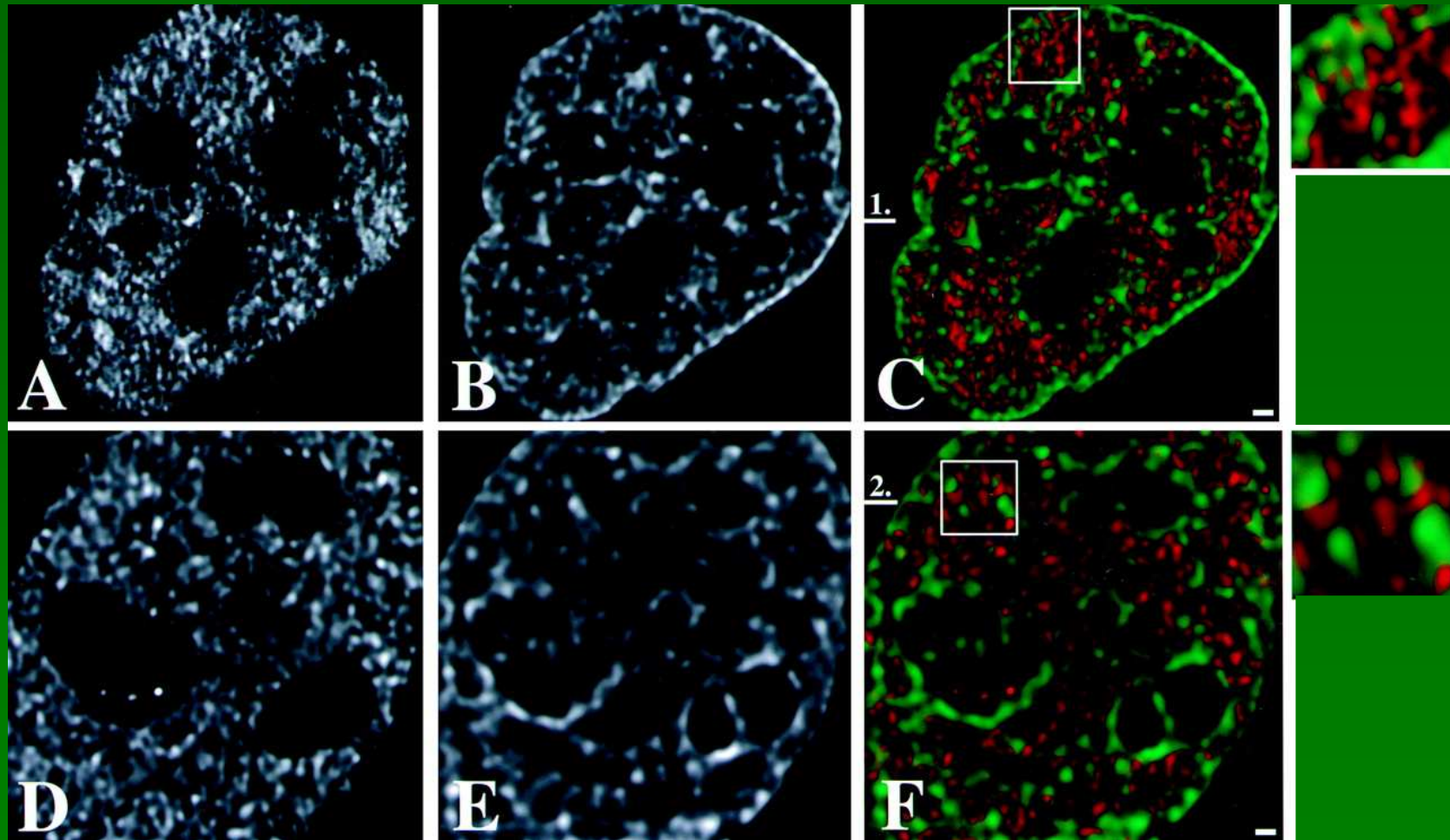
*CT a jejich části jsou disjunktní*

*Subdomény CT*

*Chromosom jako náhodný polymer*

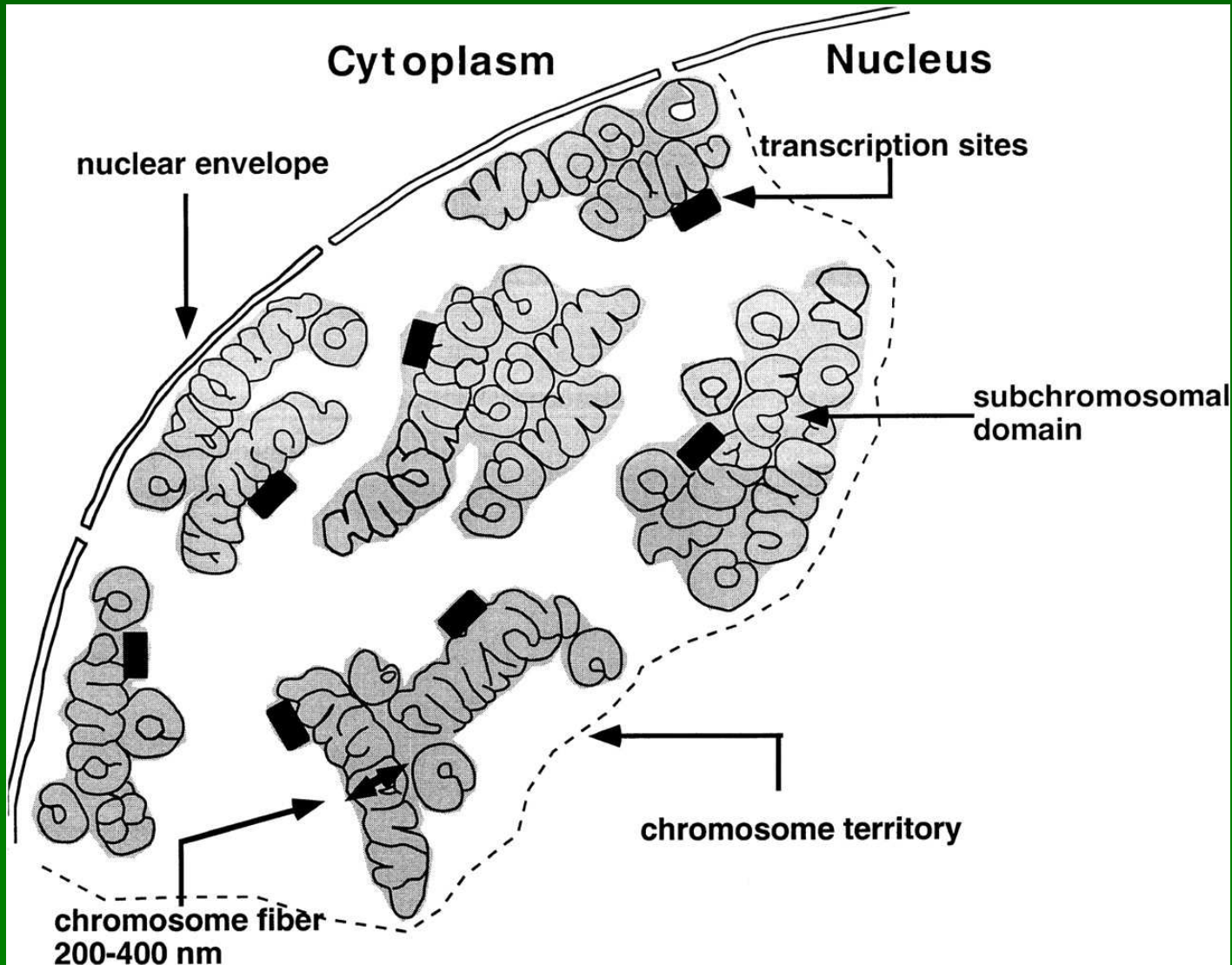
*Nenáhodná vnitřní struktura CT*

# Struktura CT – CT je tvořeno subdoménami



Verschure et al., 1999 – H2B+GFP a nascentní mRNA (BrUTP) + Cy3 ukazuje, že mRNA se nachází mezi chromatinem vizualizovaným pomocí GFP.

# Struktura CT - schéma





# Chromosomy

*Úvod - mitotické a interfázní chromosomy*

*Barvení chromosomů (klasika, FISH)*

*Významné elementy chromosomů*

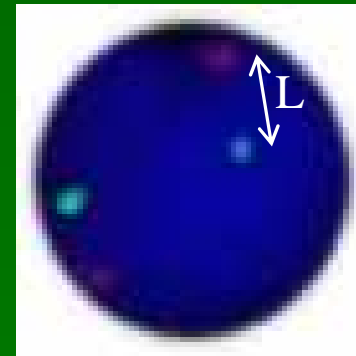
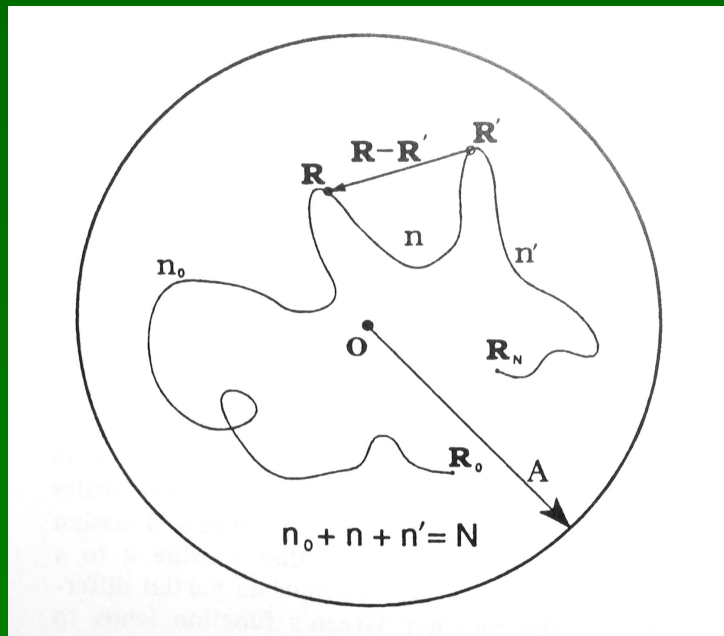
*CT a jejich části jsou disjunktní*

*Subdomény CT*

*Chromosom jako náhodný polymer*

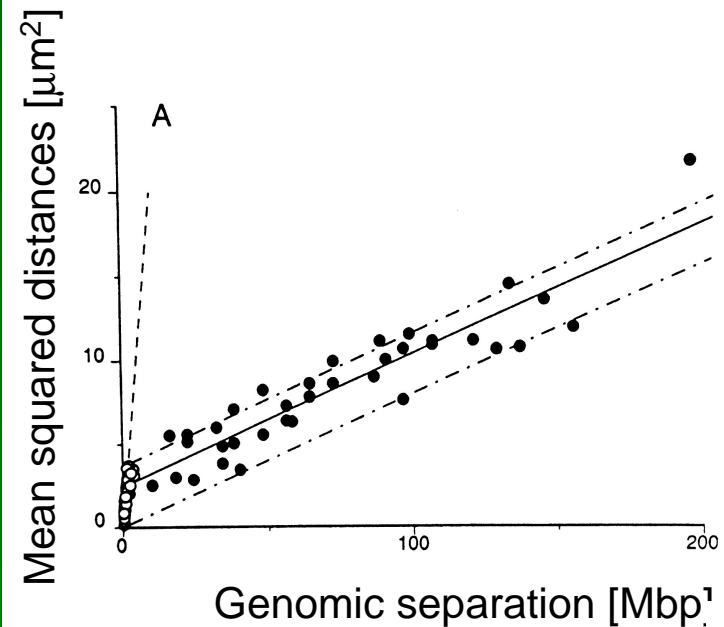
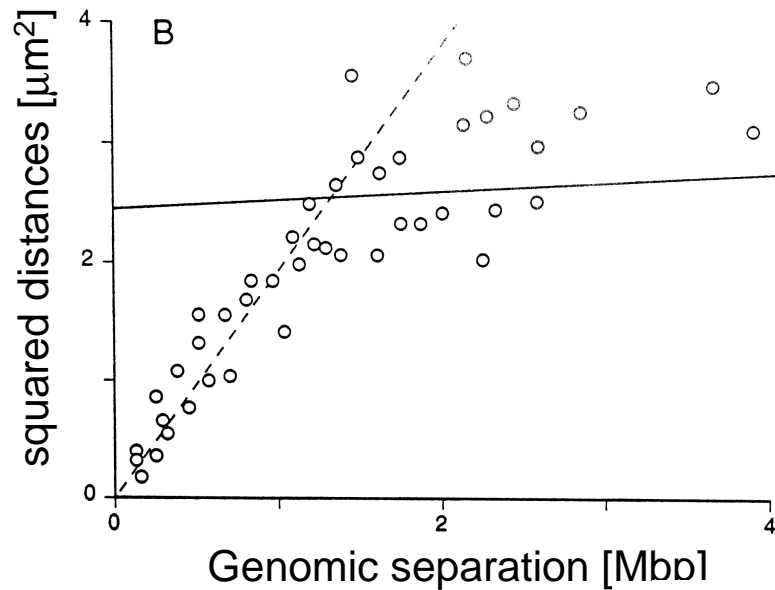
*Nenáhodná vnitřní struktura CT*

# Náhodný (Gausovský) polymer v jádře buňky (jsou zobrazeny 2 body polymeru $R$ a $R'$ )

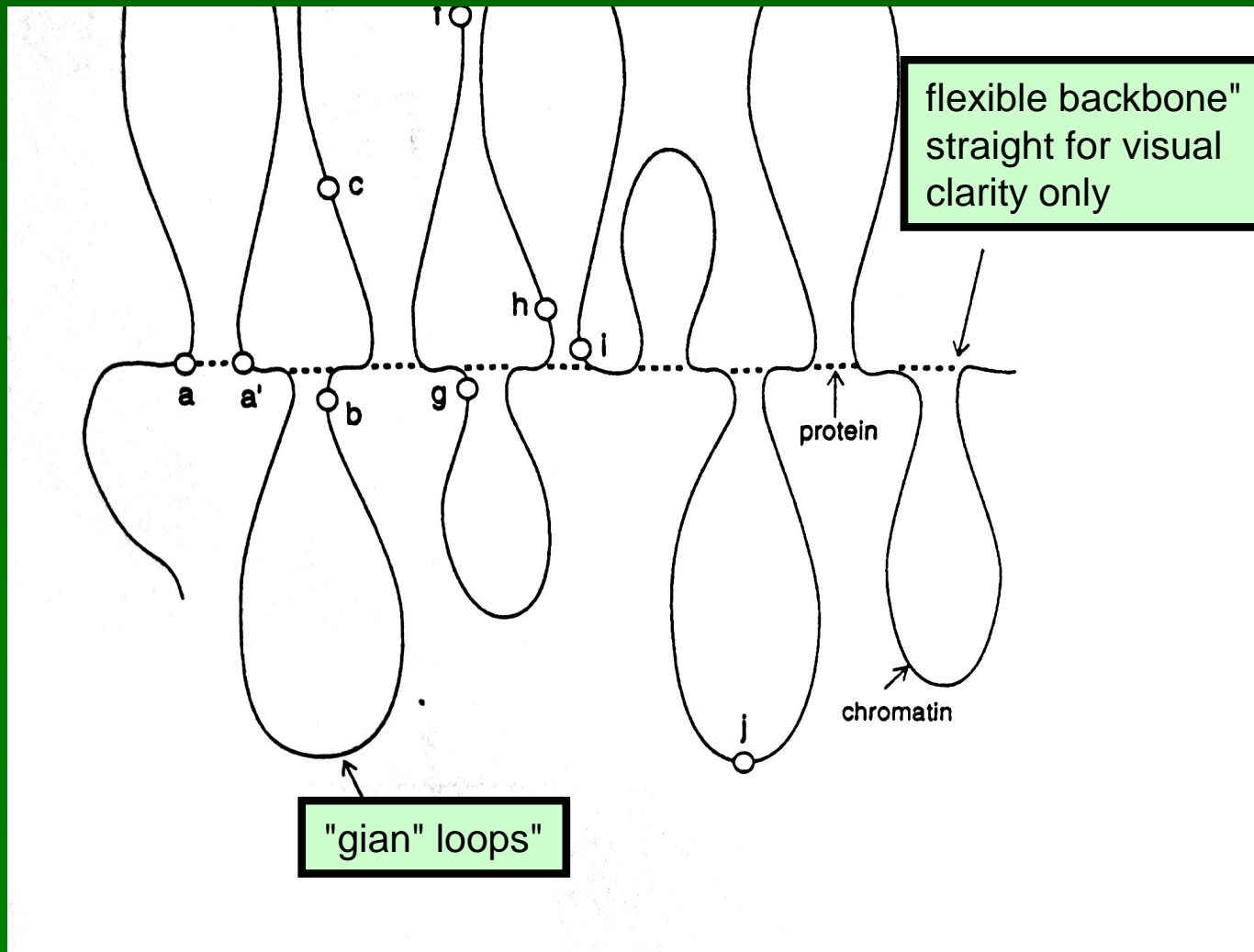


$$\langle L^2 \rangle \approx n$$

# Střední vzdálenosti mezi genetickými elementy v závislosti na jejich molekulárních vzdálenostech (H. Yokota et al., 1995)

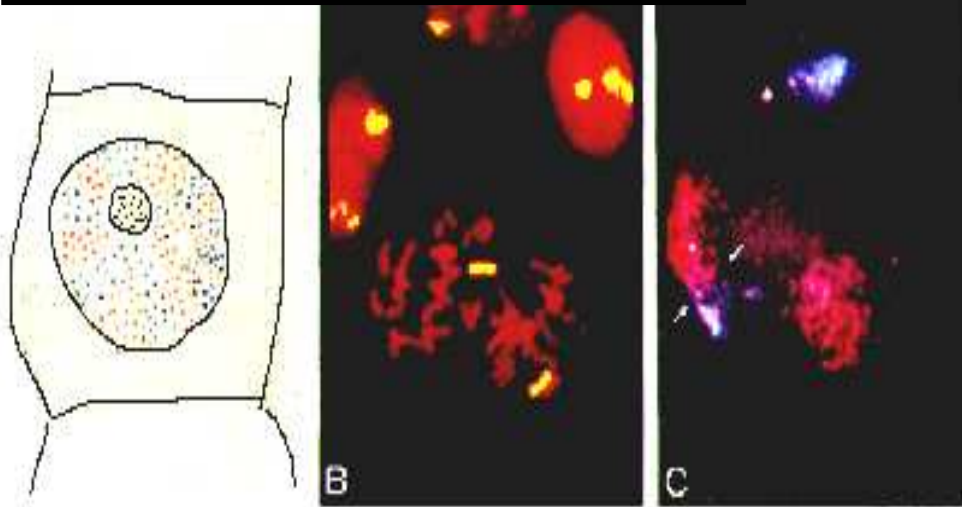


# A random-walk/giant-loop model Interfázniho jádra (Sachs et al., 1995)

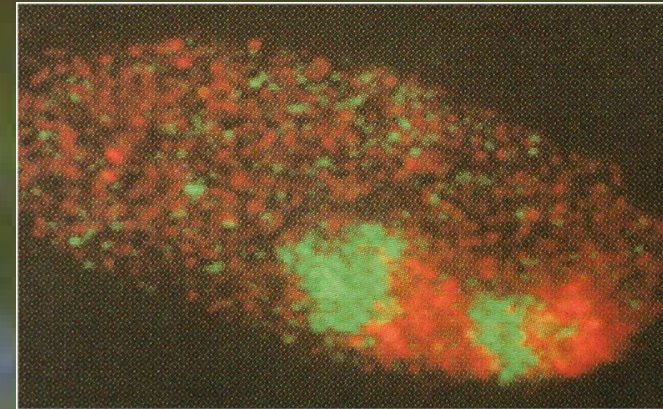


# Multi-loop subcompartment model (Münkel et al., 1999)

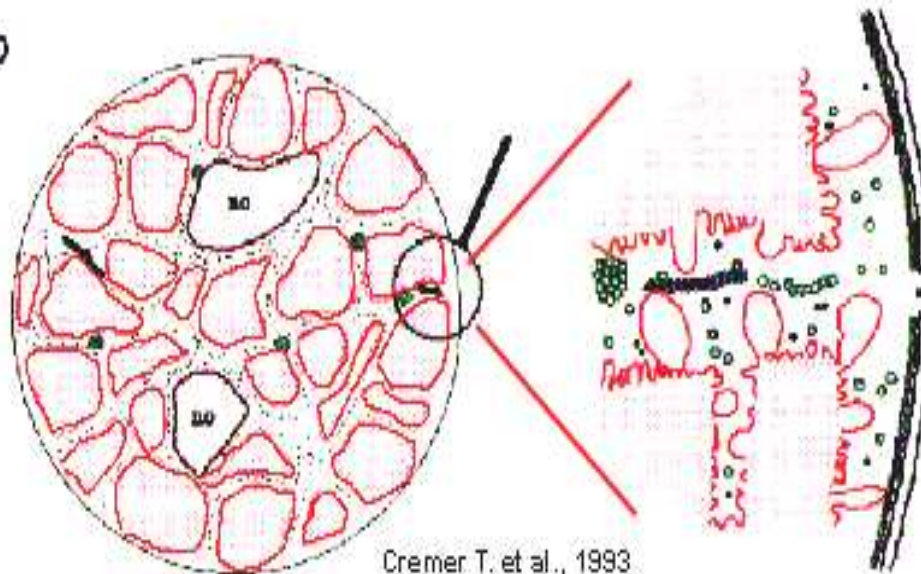
Existence disjunktních CT



Existence subdomén uvnitř CT

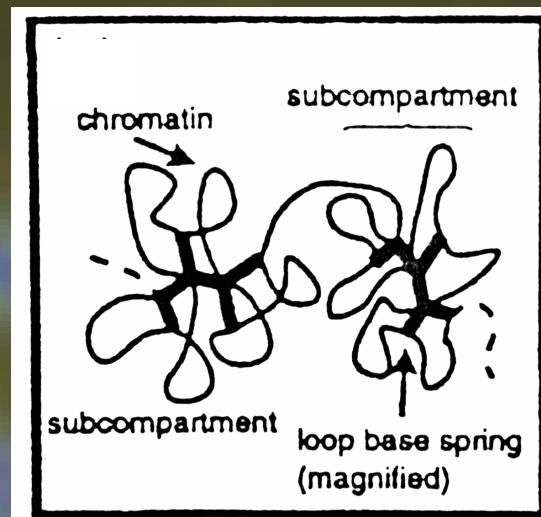


D

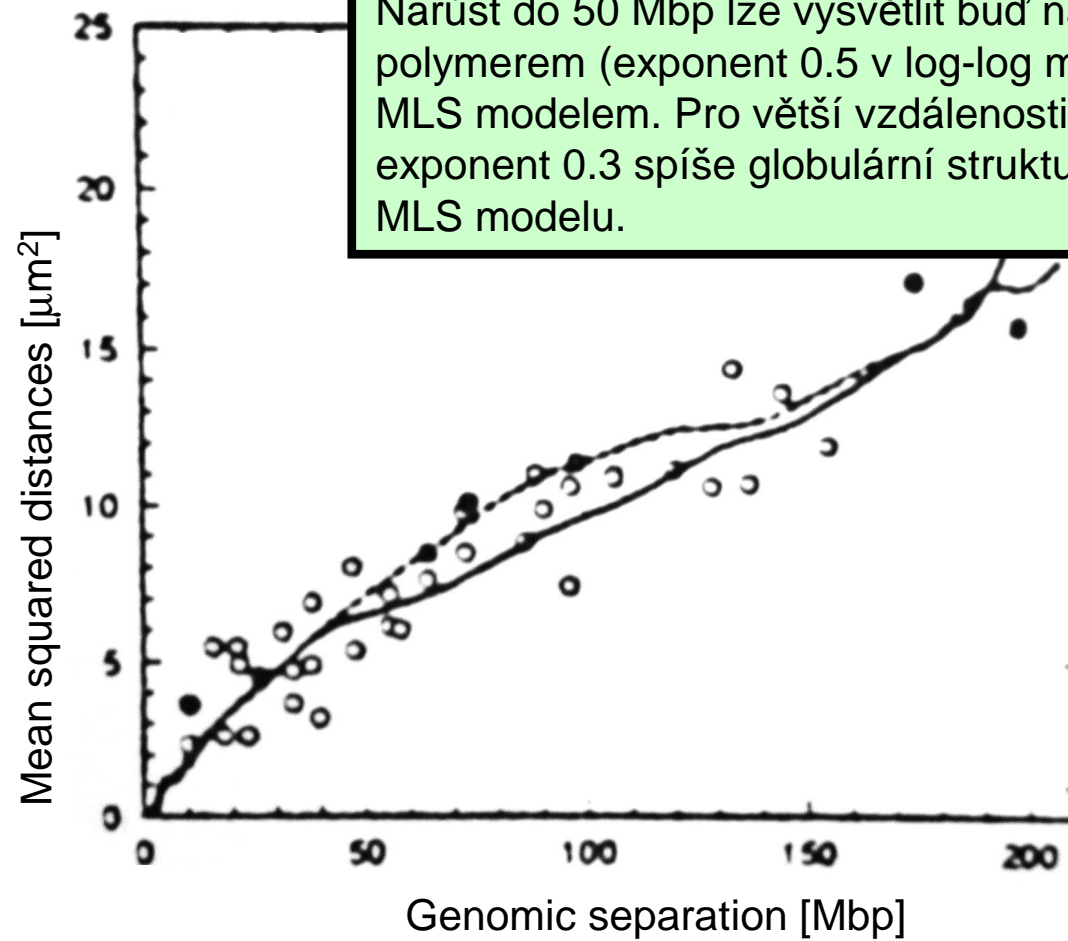


Cremer T. et al., 1993

Schéma modelu



## $L^2(n)$ závislost pro velké molekulární vzdálenosti podle MLS modelu



Nárůst do 50 Mbp lze vysvětlit buď náhodným polymerem (exponent 0.5 v log-log měřítku) nebo MLS modelem. Pro větší vzdálenosti odpovídá exponent 0.3 spíše globulární struktuře popsané v MLS modelu.

# Chromosomy

*Úvod - mitotické a interfázní chromosomy*

*Barvení chromosomů (klasika, FISH)*

*Významné elementy chromosomů*

*CT a jejich části jsou disjunktní*

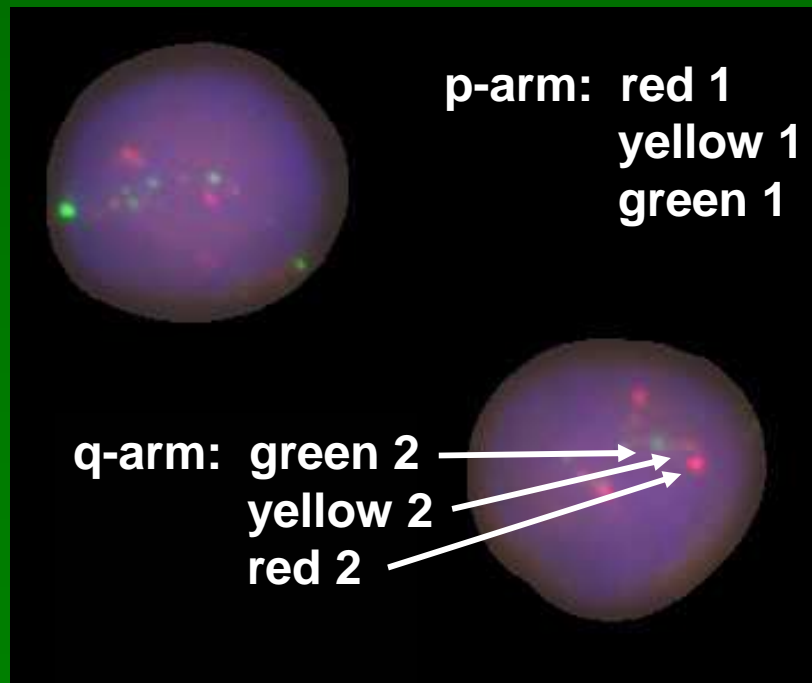
*Subdomény CT*

*Chromosom jako náhodný polymer*

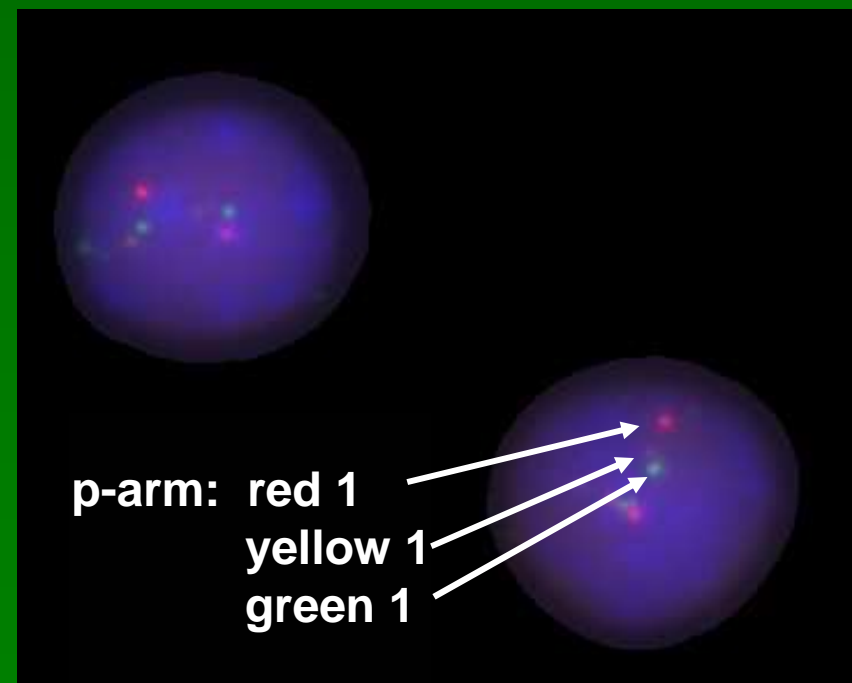
*Nenáhodná vnitřní struktura CT*

# Použití více DNA prob pro tentýž chromosom

*Signály z první hybridizace lze rozlišit v druhé*



1<sup>st</sup> hybridization – 6 probes

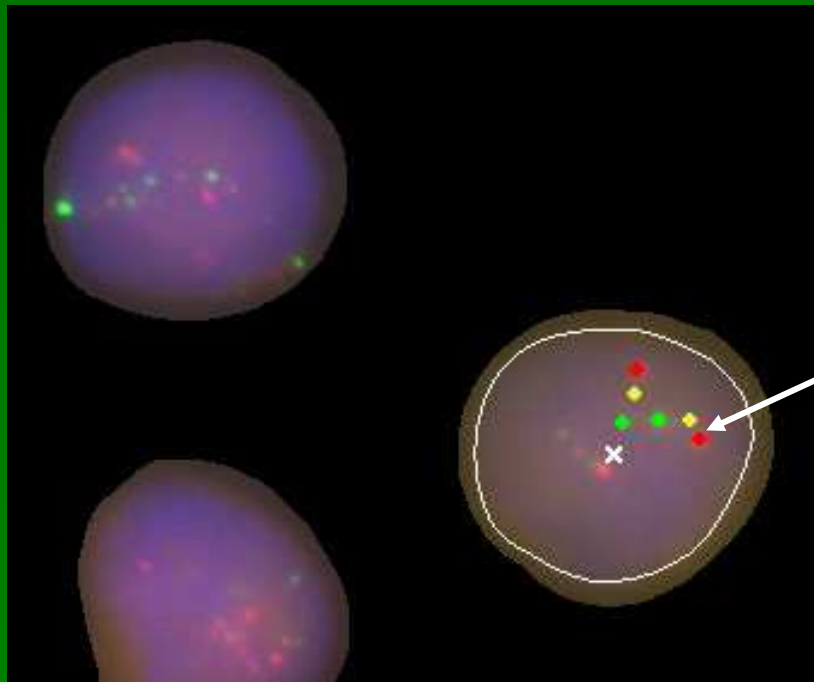


2<sup>nd</sup> hybridization – 3 probes



# Cytometrie s vysokým rozlišením – nalezení signálů

Topografické parametry jsou vypočteny pro každý signál – souřadnice x, y, z v těžišťovém systému jádra, intenzita, výška, velikost apod.



**Signal attributes:**

Hybridization #1	Probe #1	Signal type: Gene	
X: 621	Y: 189	Z: 35	
dX: 36	dY: 6	dZ: 2	
%R: 75	Intensity: 239	Height: 126	Size: 24
A1(FWHMx): 0	A2(FWHMy): 0	A3(FWHMz): 0	

Intensity criterion: 1 of 4  
Height criterion: 1 of 4  
Size criterion: 2 of 4

Do you want to delete this signal?

## Zpracování dat

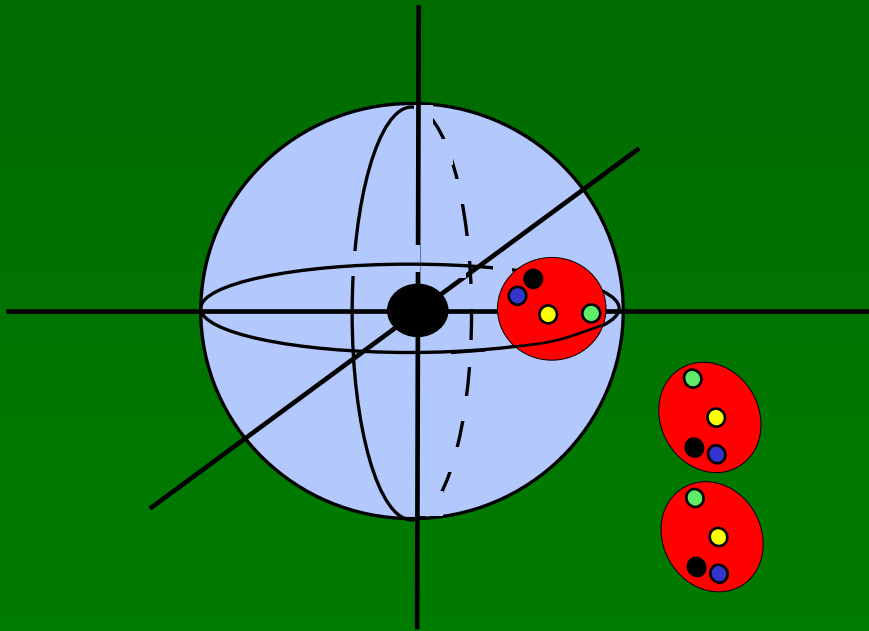
*Souřadnice a další parametry lze uložit do textového souboru a exportovat z programu FISH 2.0*

*V průběhu další analýzy **CT lze vzájemně porovnat a zkoumat míru podobnosti:***

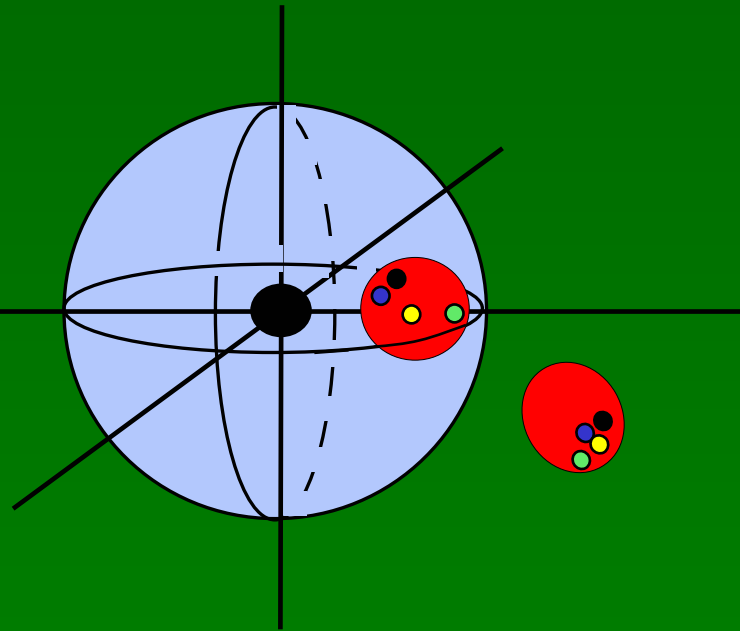
- Podobnostní funkce – definice a výpočet pro sérii CT*
- Jsou vypočtena těžiště CT a posunuta do jednoho bodu*
- CT jsou rotována v 3D prostoru tak, abychom dostali co nejlepší hodnotu podobnostní funkce*
- Definice klastrů genetických lokusů a výpočet jejich disperze*

# CT jsou vzájemně porovnány

*CT jsou stejná*

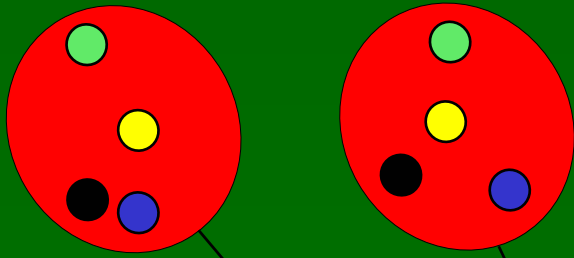


*CT se velmi liší*



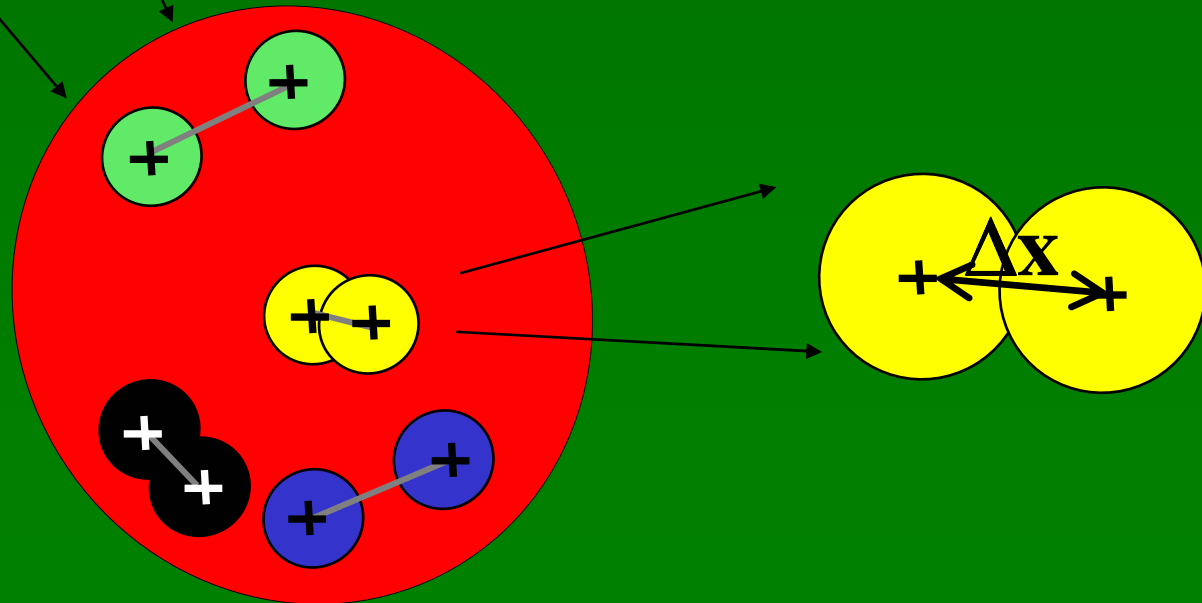
# Definice podobnostní funkce pro 2 teritoria

Použijeme translace a rotace pro superponování CT



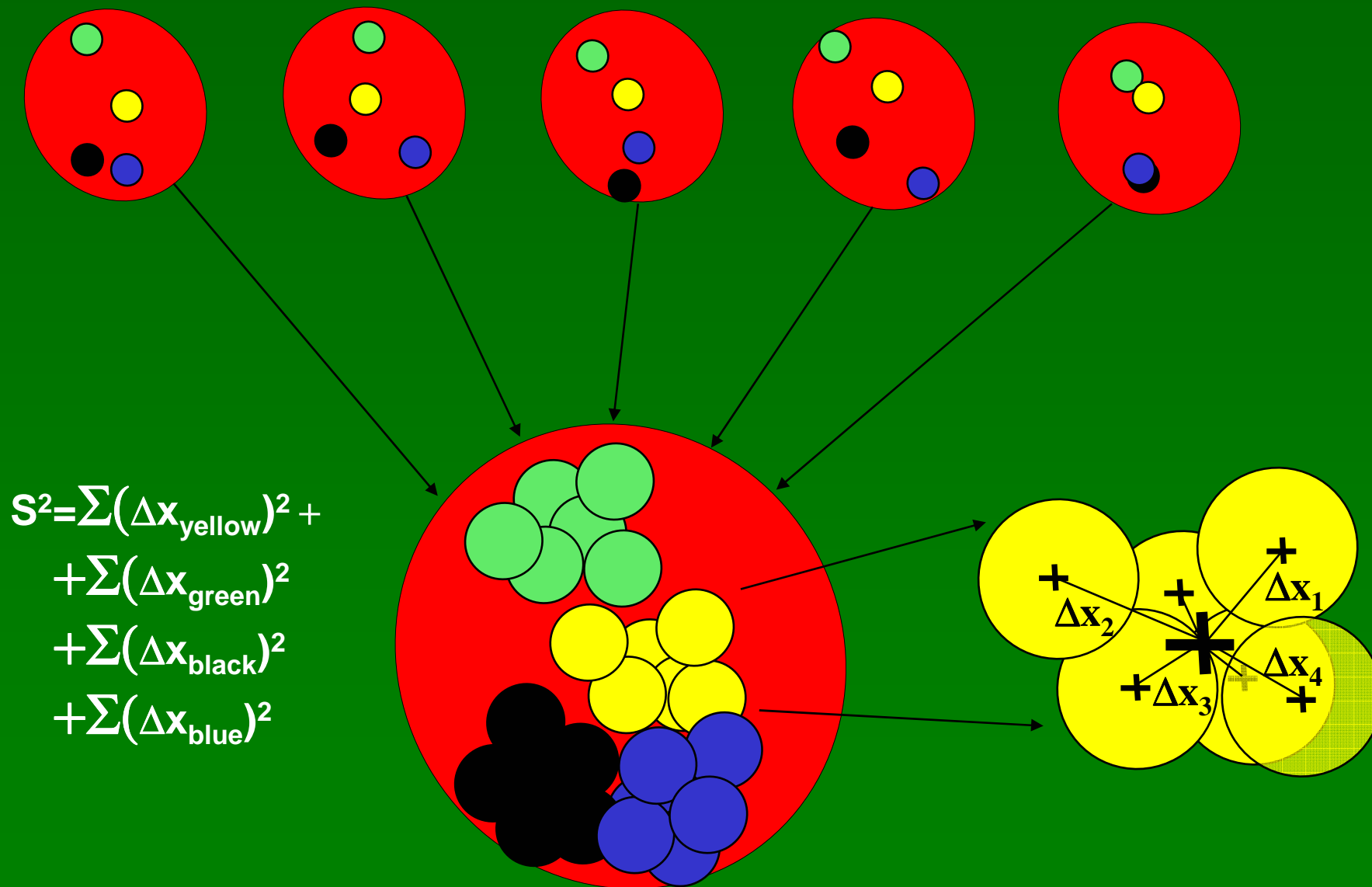
$$S^2 = \sum(\Delta x_{\text{yellow}})^2 + \sum(\Delta x_{\text{green}})^2 + \sum(\Delta x_{\text{black}})^2 + \sum(\Delta x_{\text{blue}})^2$$

Podobnost je definována jako minimální možná hodnota funkce  $S^2$

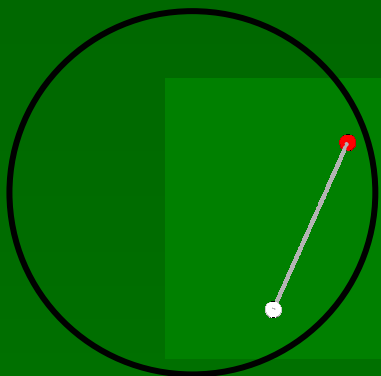


# Definice míry podobnosti pro více CT

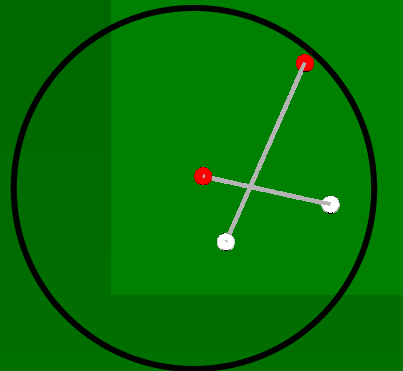
Superposice teritorií jedné na druhou metodou nejmenších čtverců



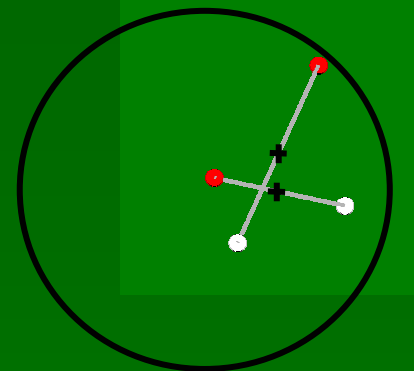
## Ukážeme metodu na příkladu 2 lokusů



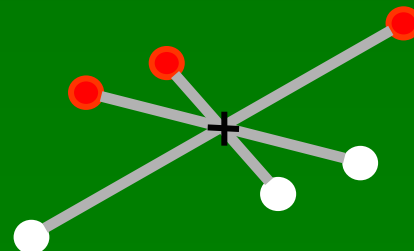
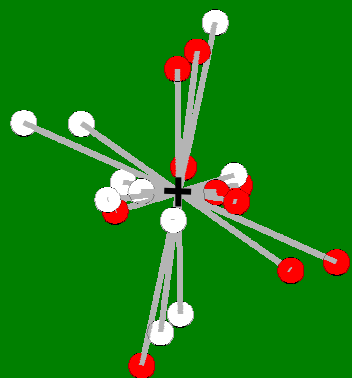
CT se 2 lokusy



2 CT se 2 lokusy  
každá

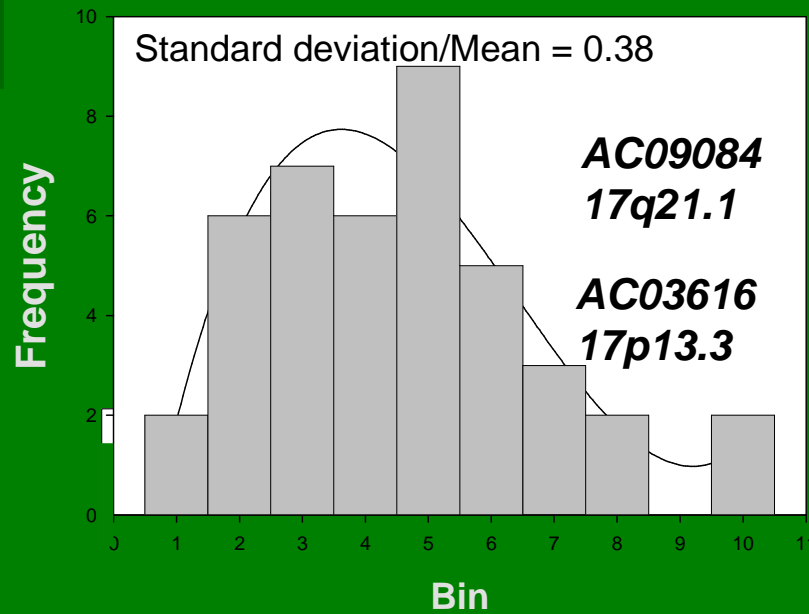
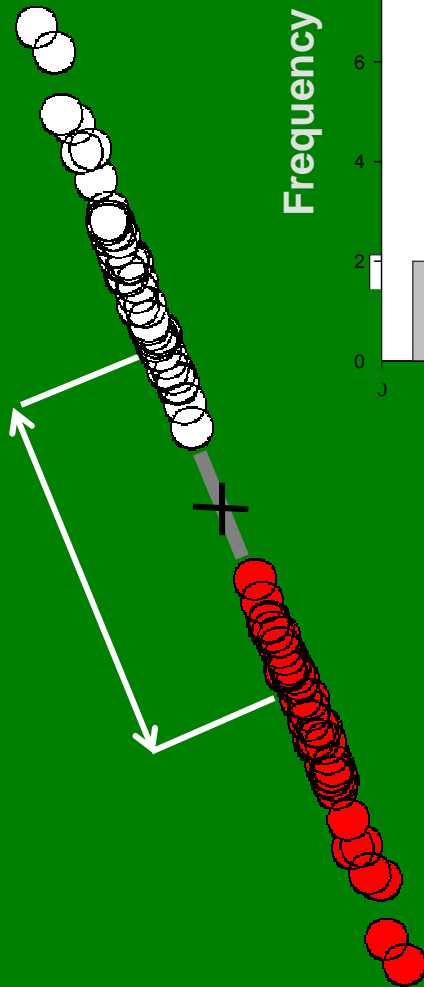


Křížkem jsou označena  
těžiště



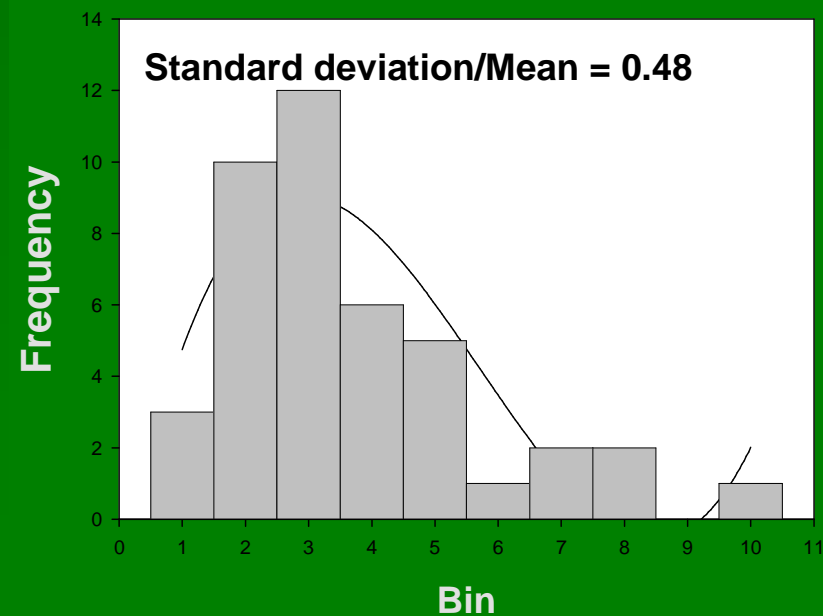
Těžiště lze posunout  
do stejného bodu a CT  
lze rotovat tak, abychom  
dostali lokusy na stejnou  
osu, červené k sobě a  
bílé k sobě.

# Výsledky rotace vidíme pro reálné lokusy HSA 17



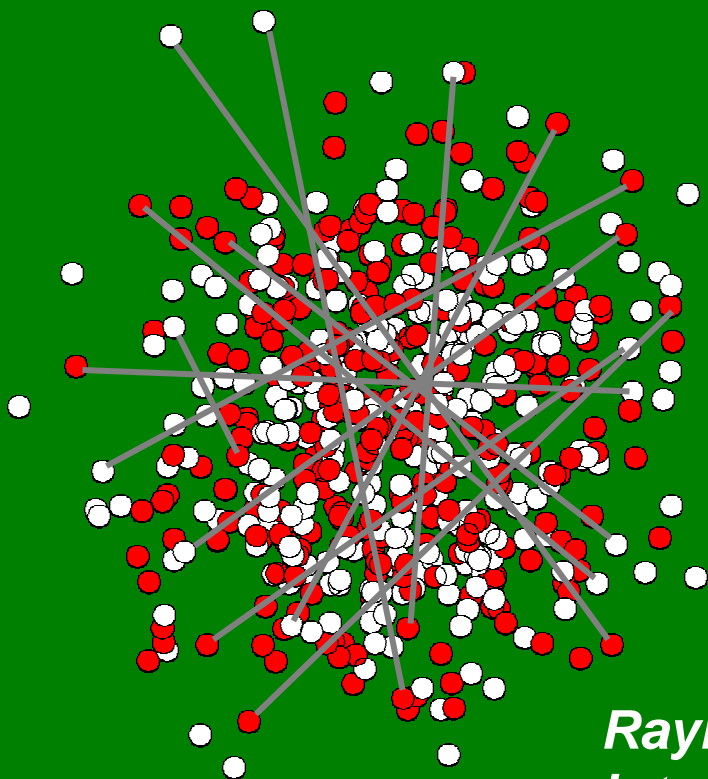
Lze vypočítat  
rozdělení vzdáleností  
a poměr

$$R = S.D./Mean$$

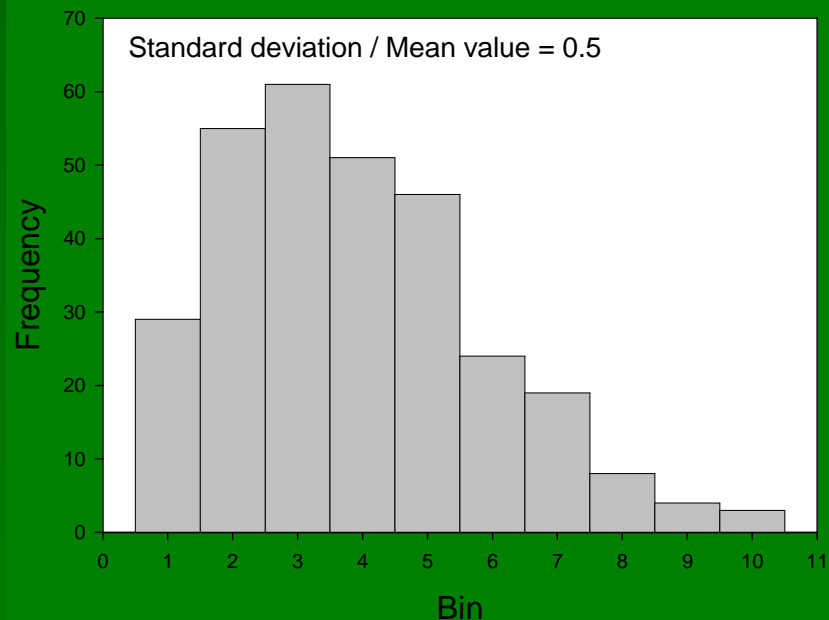


## Porovnání s náhodným modelem:

300 dvojic bodů byly náhodně generovány v prostoru podle Gaussovy distribuce



### Rayleigh distribution, $R=0.5$

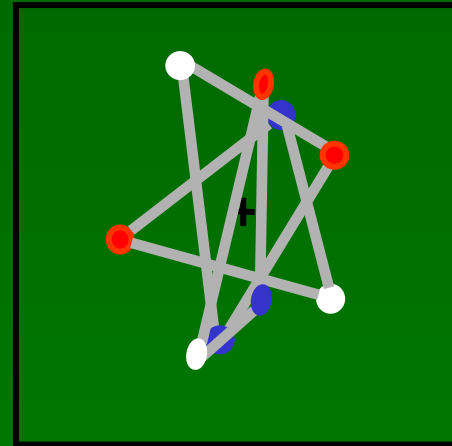
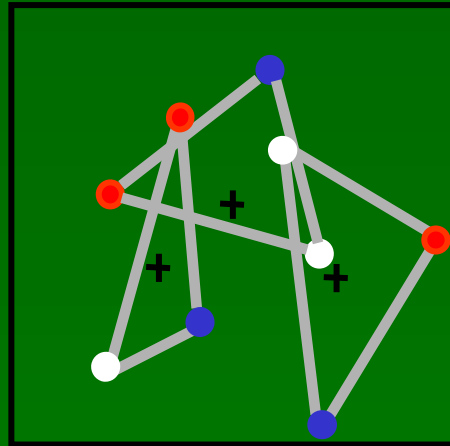
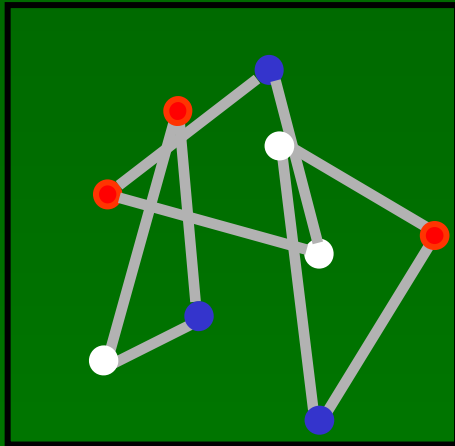


*Rayleighovo rozdělení vede k závěru, že interfázní chromosom se chová jako volně plovoucí polymer*



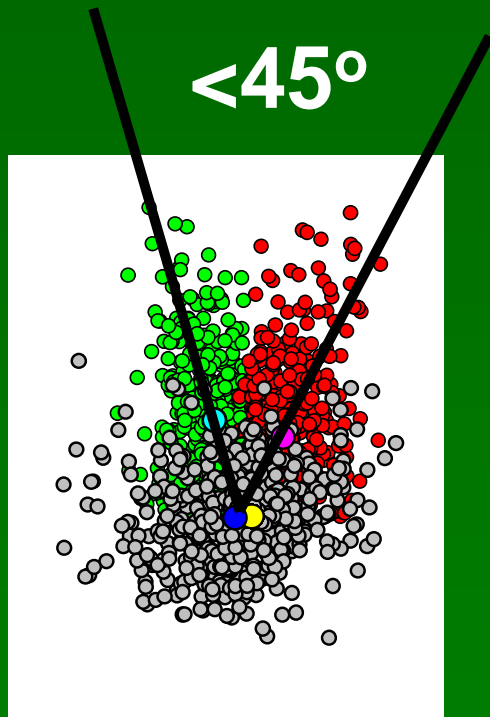


## Uvažujme nyní 3 genetické elementy



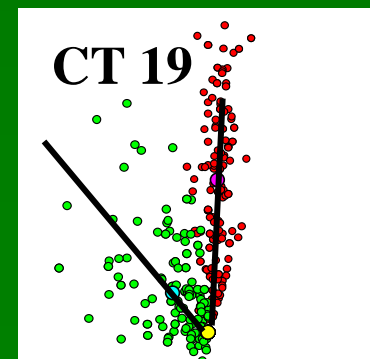
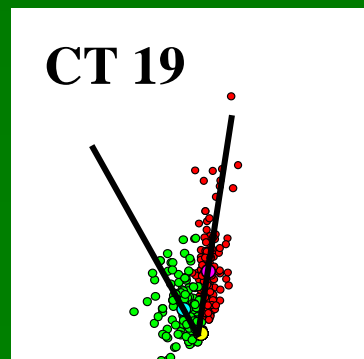
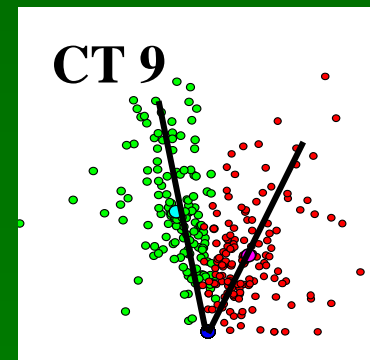
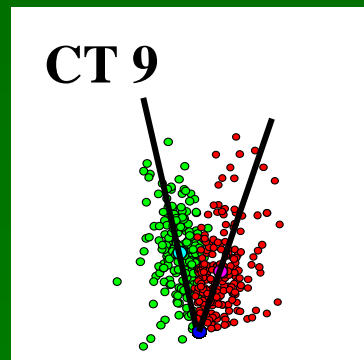
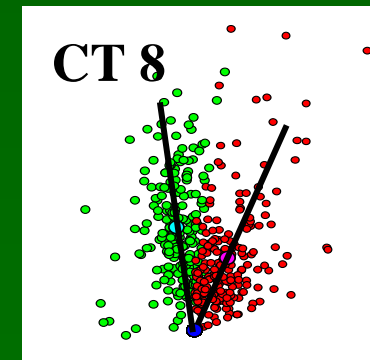
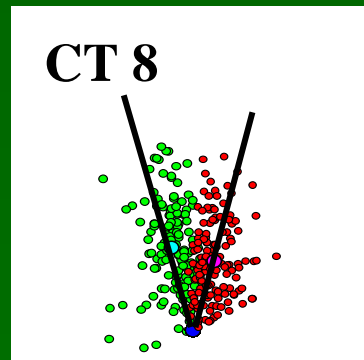
- Vypočteme těžiště trojúhelníků,
- Položíme je do jednoho bodu
- Rotujeme trojúhelníky tak, abychom dostali signály stejné barvy co nejbliže k sobě
- Genetické elementy stejného typu vytvoří klastry

# Polární struktura chromosomových teritorií

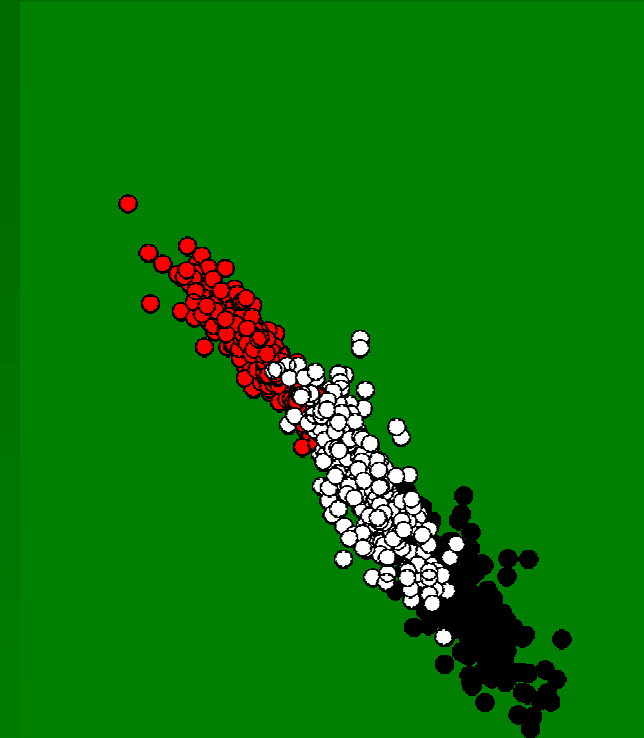
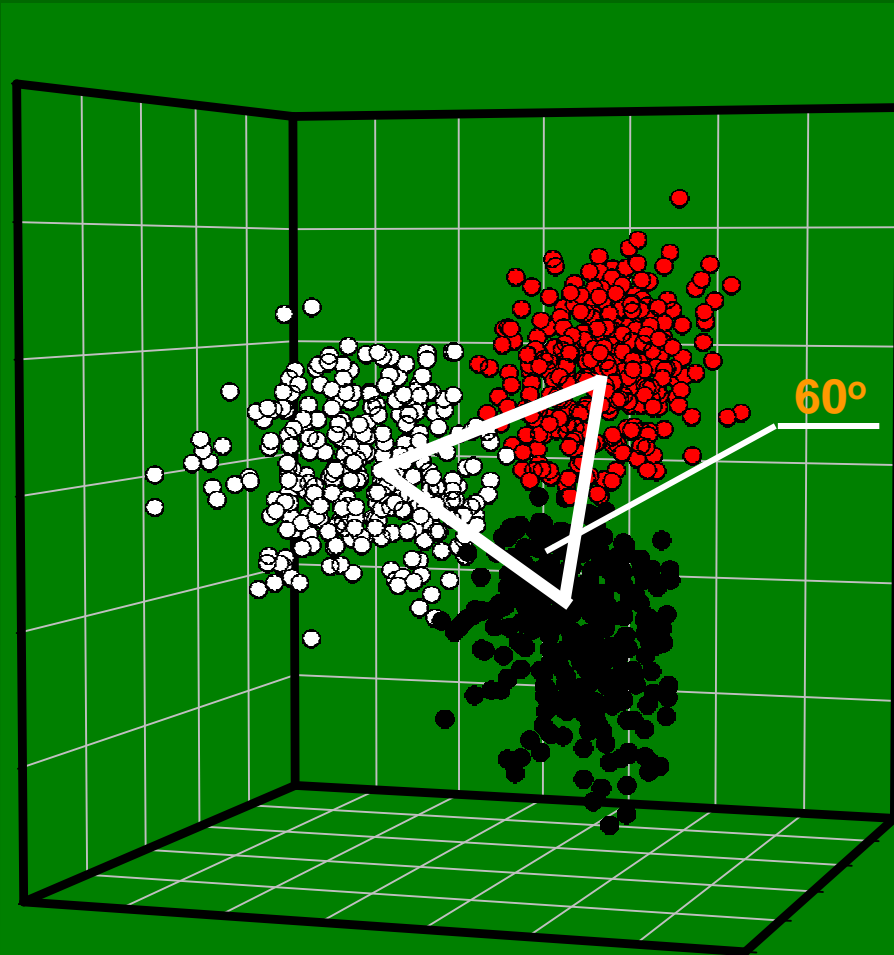


CT 8 in Go-lymphocytes

Go-lymphocytes      Stimulated lym

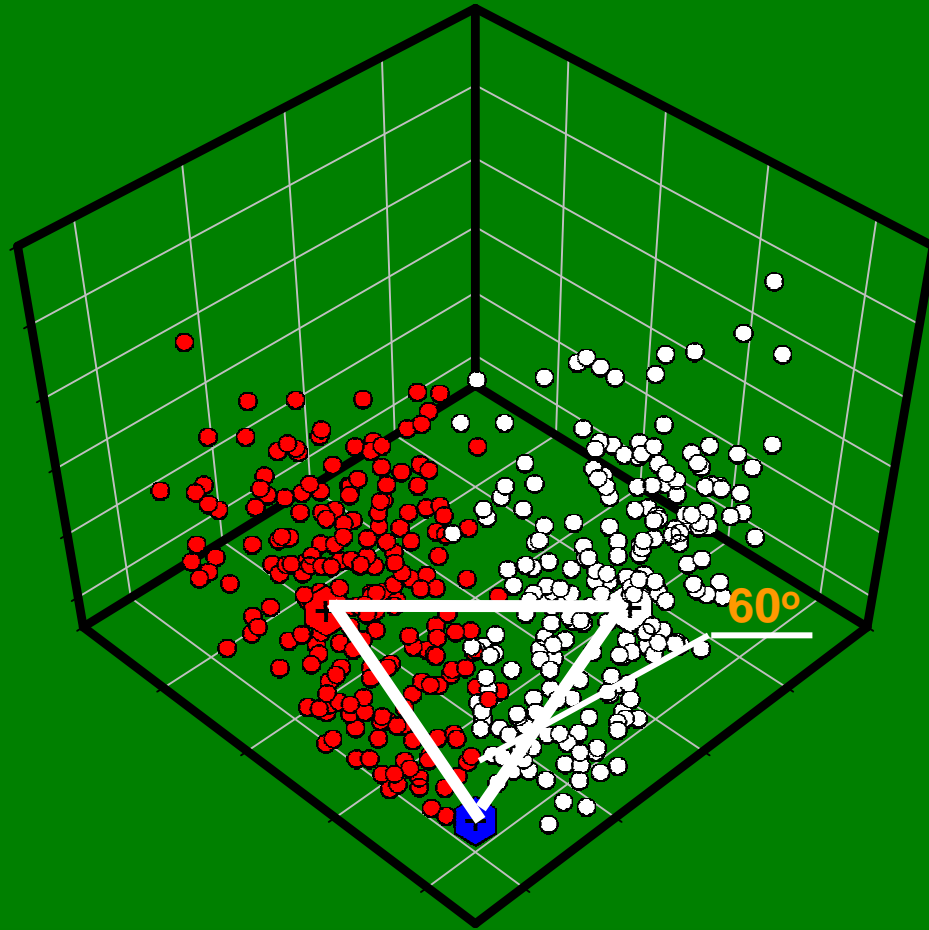


# Porovnání s náhodným modelem: 3 elementy náhodně generovány po rotaci



Side view – elements are  
located in one plane

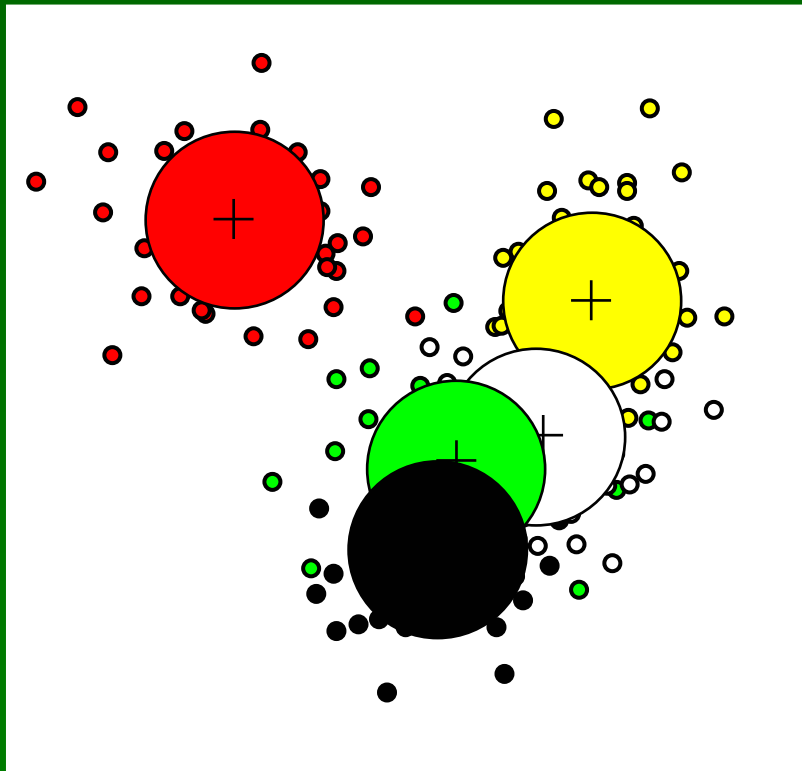
# Porovnání s náhodnými modely: 3 elementy náhodně generovány v prostoru po rotaci kolem jednoho z nich



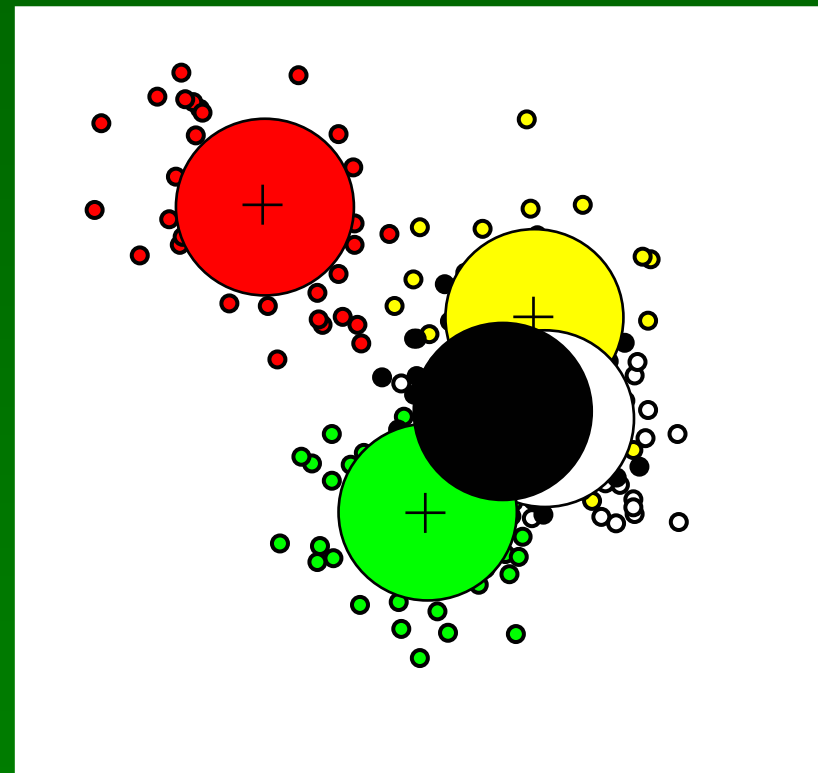
Podle náhodných modelů by  
úhly rovnostranného  
Trojúhelníku měly být **60°**

**Menší hodnoty získané  
v experimentu představují  
důkaz nenáhodného  
uspořádání CT**

# Více genetických elementů: 5 lokusů CT 17 u Go- a stimulovaných lymfocytů

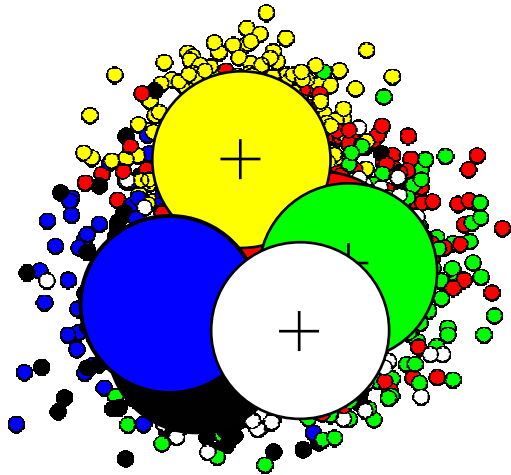


Go-lymphocytes

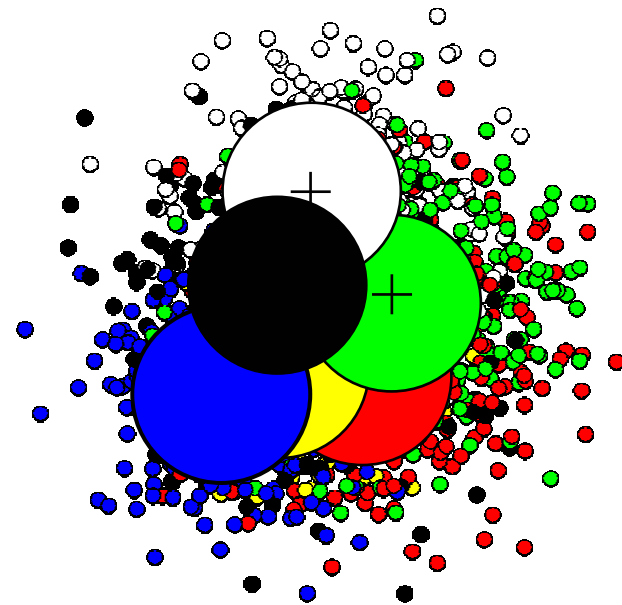


Stimulated lymphocytes

# Porovnání s náhodnými modely: pro více elementů se zmenšuje jejich vzájemná vzdálenost po rotaci



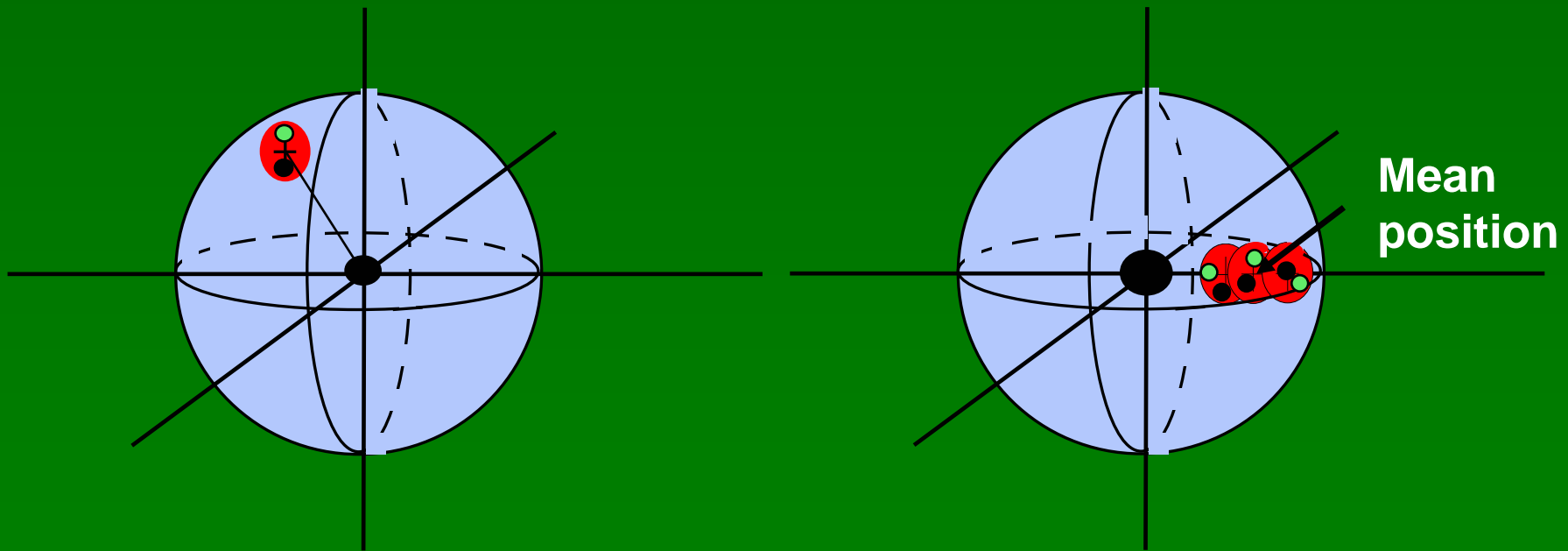
Points restricted in a spatial volume



Gaussian distribution

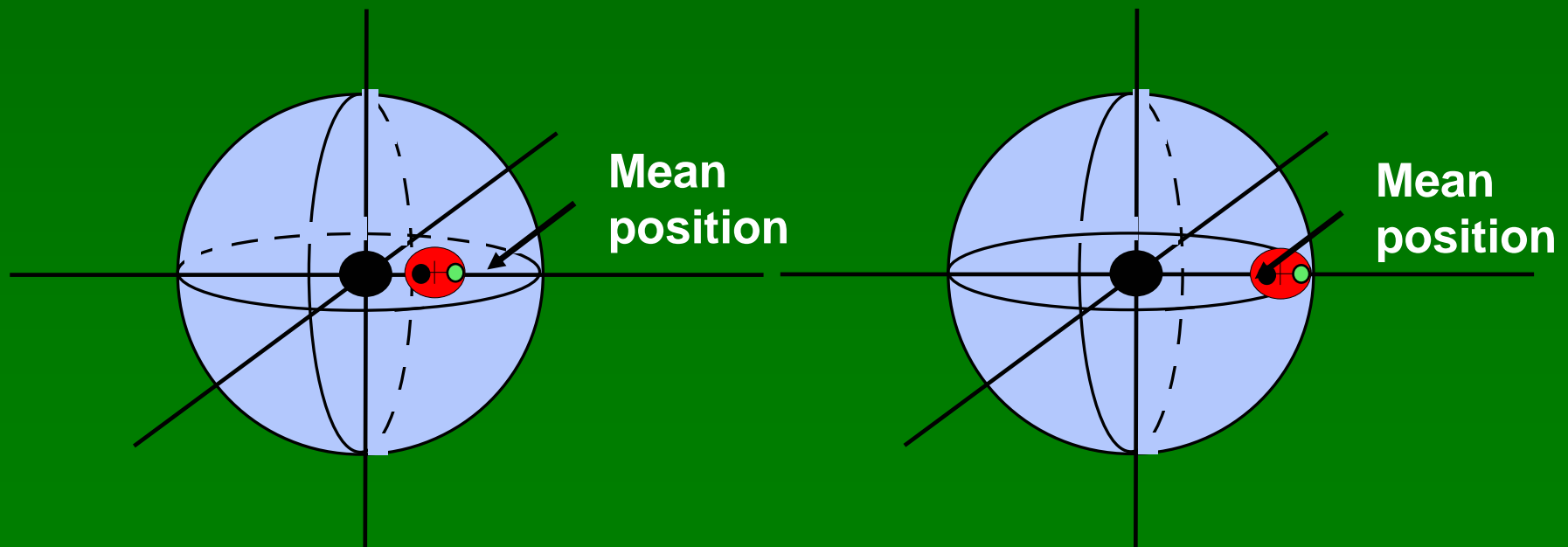
## Orientace CT v jádrech buněk

- 1) Buněčná jádra rotujeme tak, abychom dostali CT na jednu osu
- 2) Vypočte se střední poloha CT na této ose



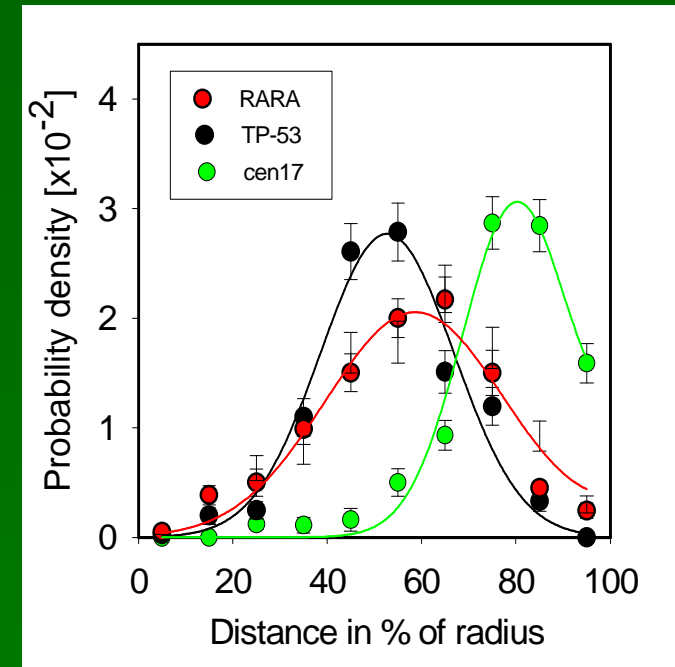
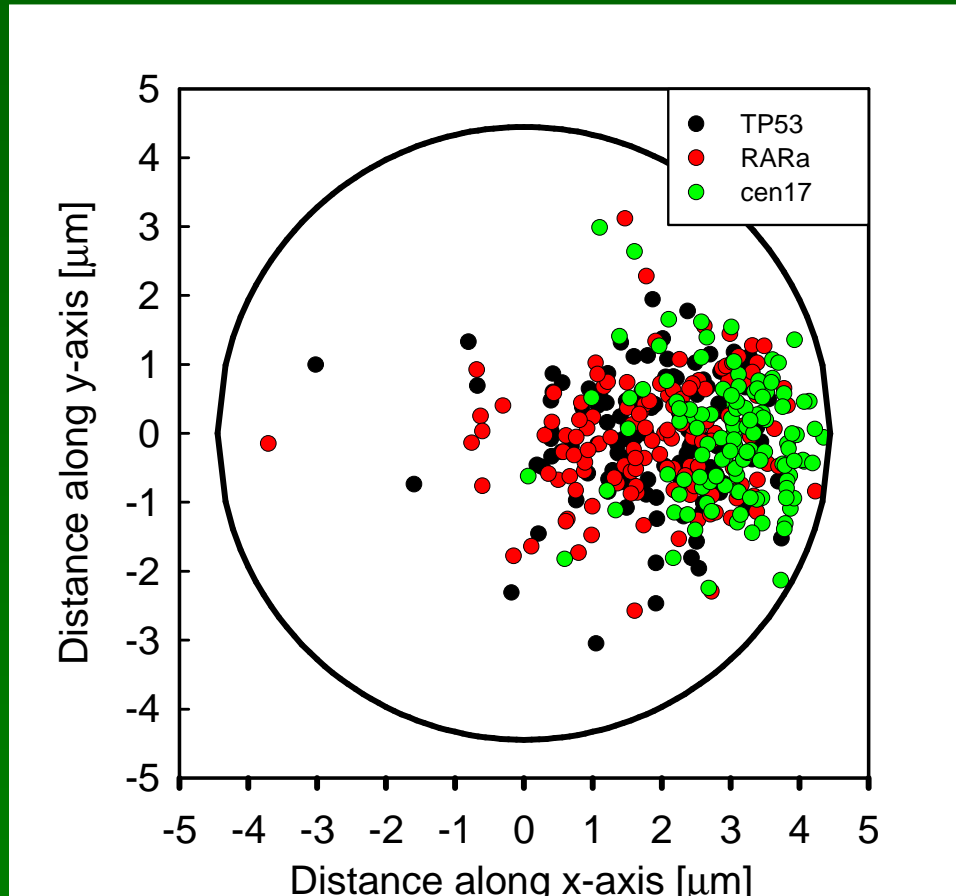
## Transformace CT v jádře buňky

- 3) Těžiště CT posuneme podél osy do jednoho bodu – střední polohy
- 4) Vezmeme tenký řez jádrem kolem roviny vedoucí přes střed jádra a střed CT





# Orientace CT v jádře buňky



Výše popsaným způsobem dostaneme reálné polohy genetických lokusů vzhledem k CT a také do značné míry vzhledem k jádru. Transformace odstraní fluktuace polohy genetických elementů jež jsou způsobené fluktuacemi CT uvnitř jádra

# Struktura chromosomových teritorií

## Závěry o nenáhodném uspořádání chromosomových teritorií:

- 1) Poměr mezi standardní odchylkou a střední hodnotou vzdálenosti mezi dvěma genetickými lokusy je v mnoha případech menší než 0.5, což je hodnota typická pro Rayleighovu distribuci.
- 2) Úhly trojúhelníku, jež vznikne ze tří genetických elementů stejného CT v jádře, jsou v mnoha případech menší než  $60^\circ$ , tj. hodnota, kterou lze očekávat pro náhodnou strukturu.
- 3) Pro větší počet genetických oblastí, polohy elementů ve 3D prostoru mají menší rozptyl než klastry vzniklé pro náhodně generovanou strukturu.
- 4) Orientace jader se zviditelněnými CT a jejich genetickými elementy poskytuje důkaz o orientaci a polaritě CT.