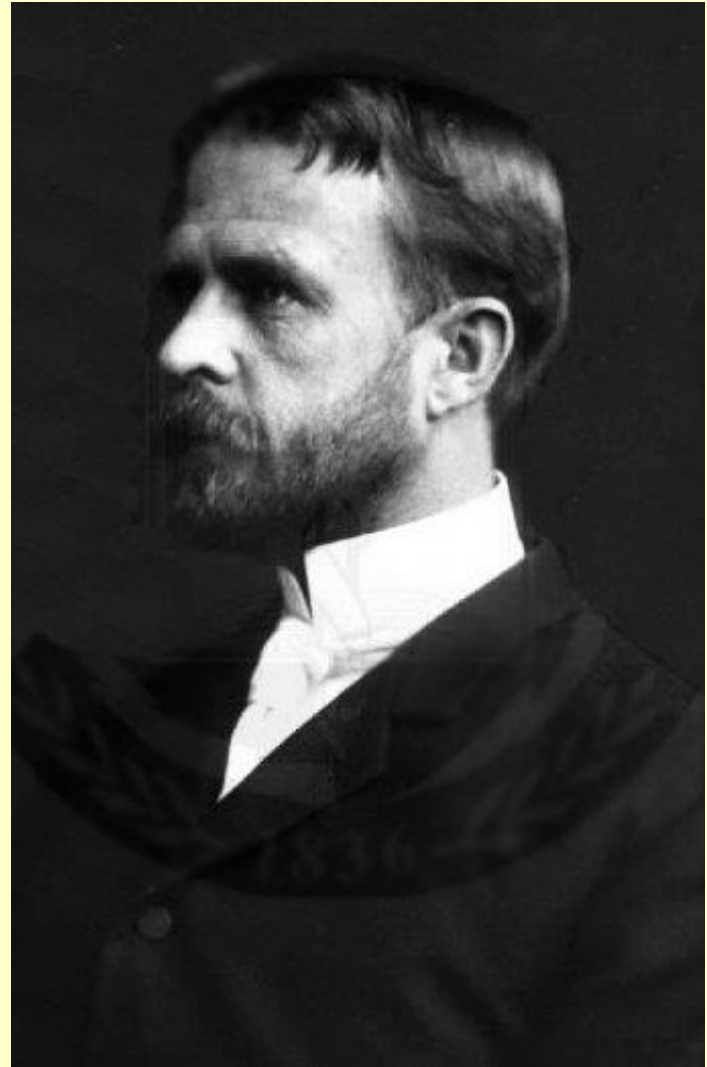


Vazba vloh

Thomas Hunt Morgan

(1866 – 1945)

- americký genetik a embryolog
- pokusy s octomilkou (*D. melanogaster*)



Morganova pravidla

1. geny jsou na chromozómu uspořádány lineárně za sebou
2. počet vazbových skupin odpovídá haploidnímu počtu chromozómů

Výjimky: - přeskupování genů nebo jejich částí
- překřížení a rekombinace během
meiózy

Geny na chromozomu

- a) vazbová skupina
- b) syntenní skupina

Syntenní skupina

- geny lokalizované na stejném chromozómu jsou vzdáleny natolik, že se chovají jako nezávislé
- pravděpodobnost rekombinací je 50% (vzdálenost mezi geny $>$ než 50cM)

Syntenii prokazujeme

- hybridologickou analýzou (genetické mapování)
- cytogenetickými a molekulárně-genetickými technikami (hybridizace *in-situ*, FISH)

Vazbová skupina

- blízké umístění
- podíl rekombinací $< 50\%$
- geny se nemusí přenášet společně

Pravděpodobnost, že se geny **nepřenesou** společně je pravděpodobností výskytu crossing-overů.

Druhy vazby

Úplná:

- geny zůstávají stále ve stejné vazbové skupině
- mezi chromozomy neprobíhá c. o.

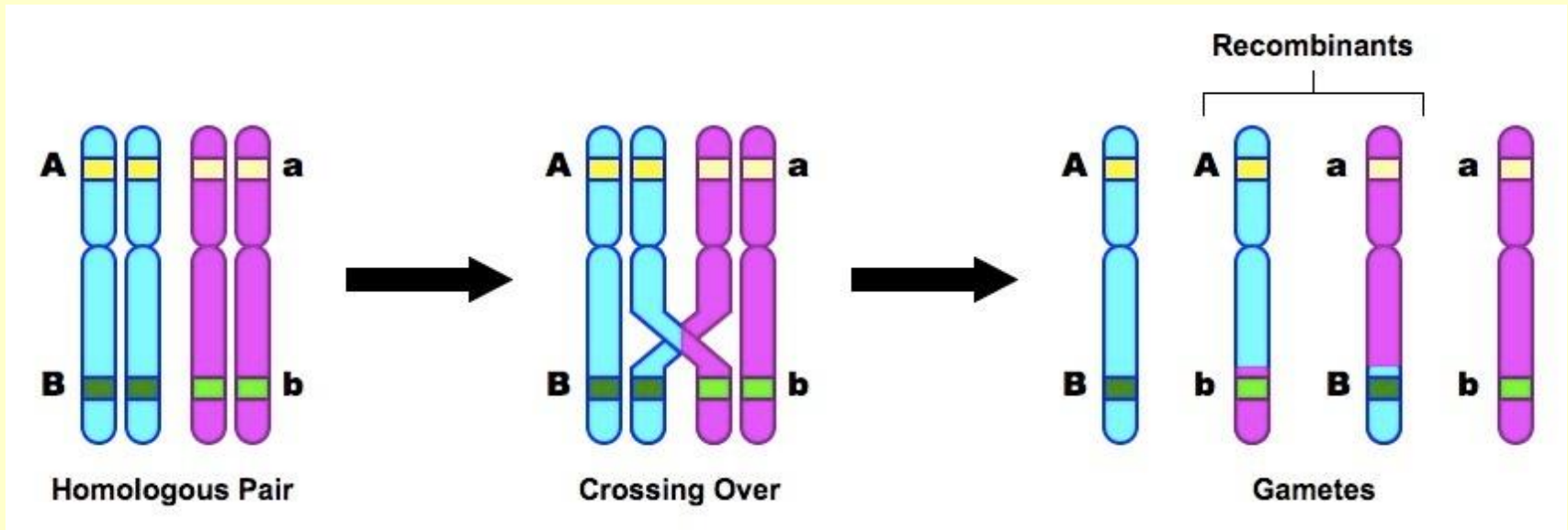
Neúplná:

- geny mohou přecházet z jednoho homolog. chromozomu na druhý
- mezi chromozomy probíhá c. o.
- vznik rekombinantních gamet, jejichž frekvence je vždy nižší než frekvence rodičovských, nerekombinovaných gamet

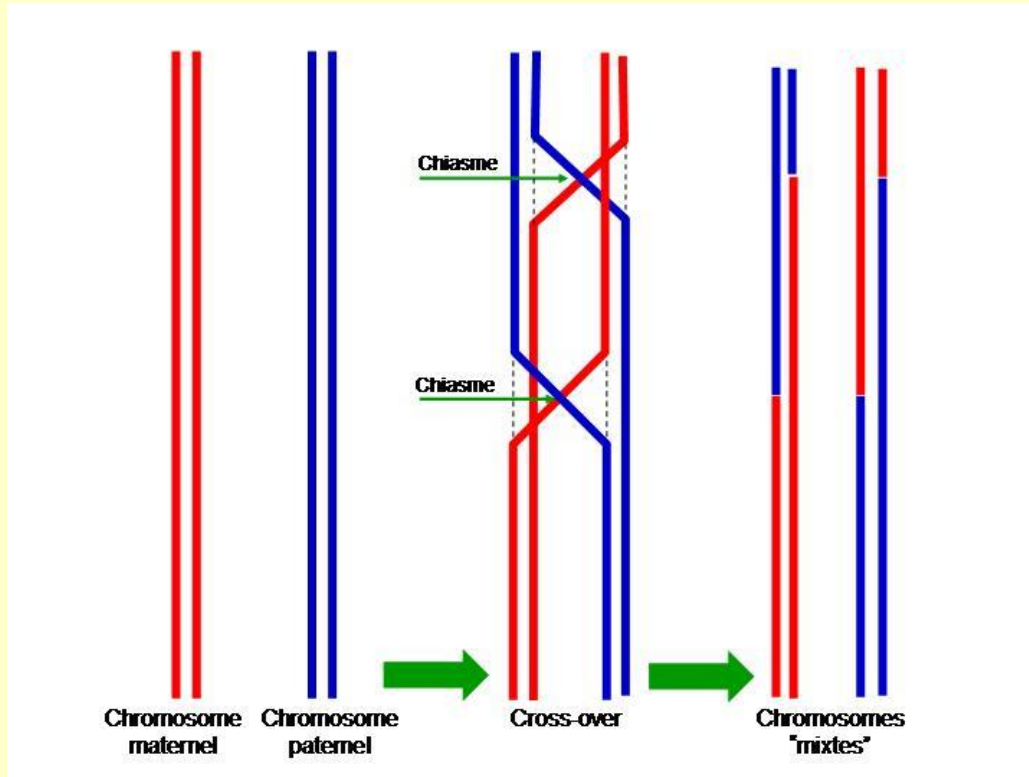
Příčina neúplné vazby

Crossing over v meiotické profázi I

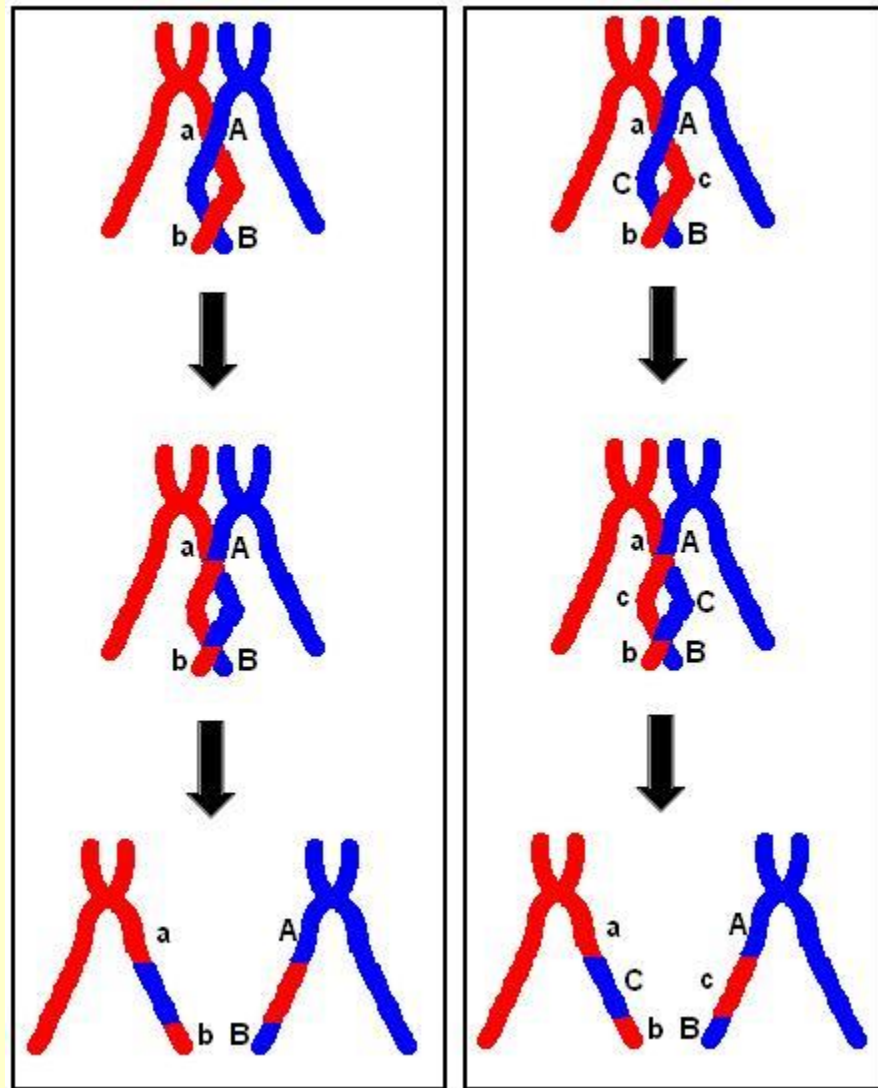
CROSSING-OVER



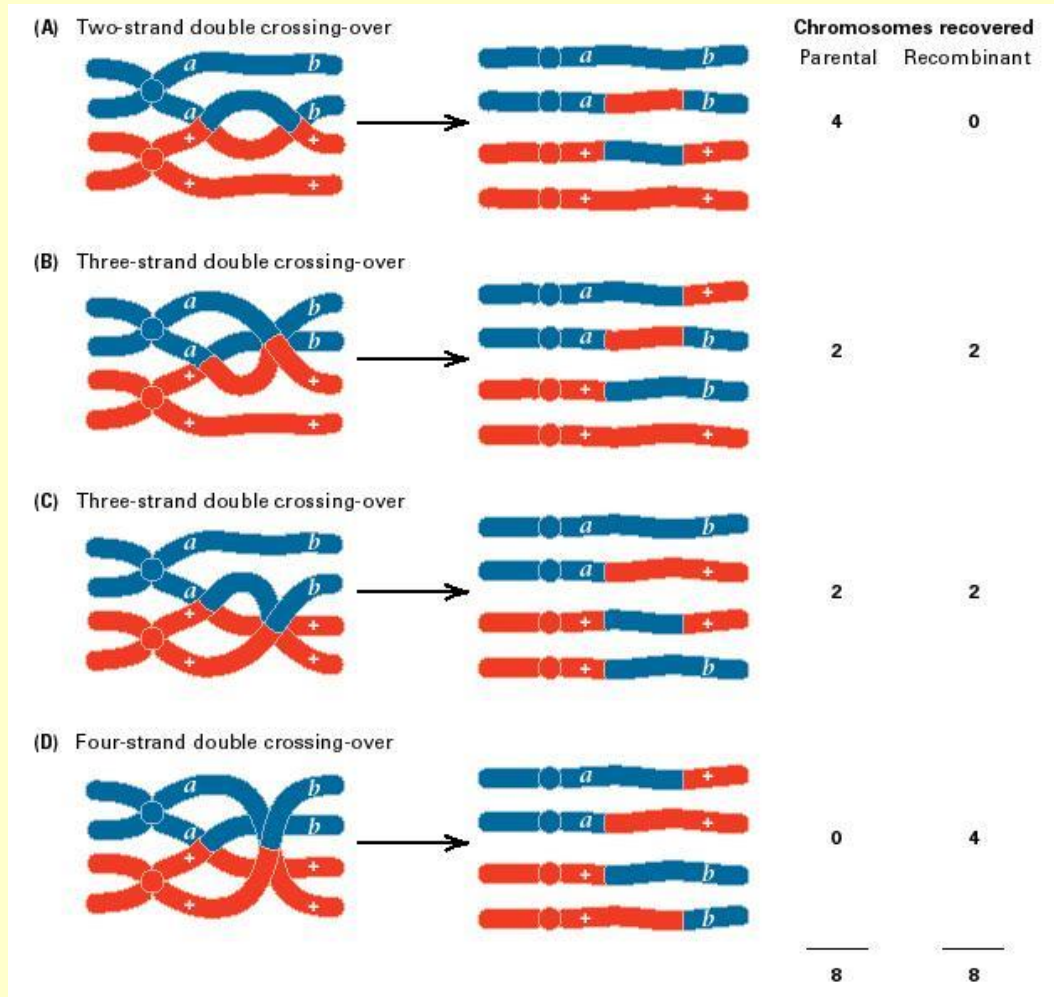
Mohou být rekombinované obě chromatidy, každá jinak



DVOJITÝ CROSSING-OVER



Také u dvojitého c.-o. mohou být rekombinované obě chromatidy,
každá jinak



Crossing-over není zcela náhodným procesem z hlediska toho, kde probíhá.

Běžně nenastává uvnitř genů nebo jejich exonů.

Frekvence (pravděpodobnost) crossing – overu stoupá se zvyšující se vzdáleností genů na chromozomu.

Rekombinace

- zvyšuje počet geneticky různých gamet
- rekombinace, náhodný rozchod chromozómů a chromatid a náhodné spojení gamet vytváří prakticky nevyčerpatelný počet kombinací

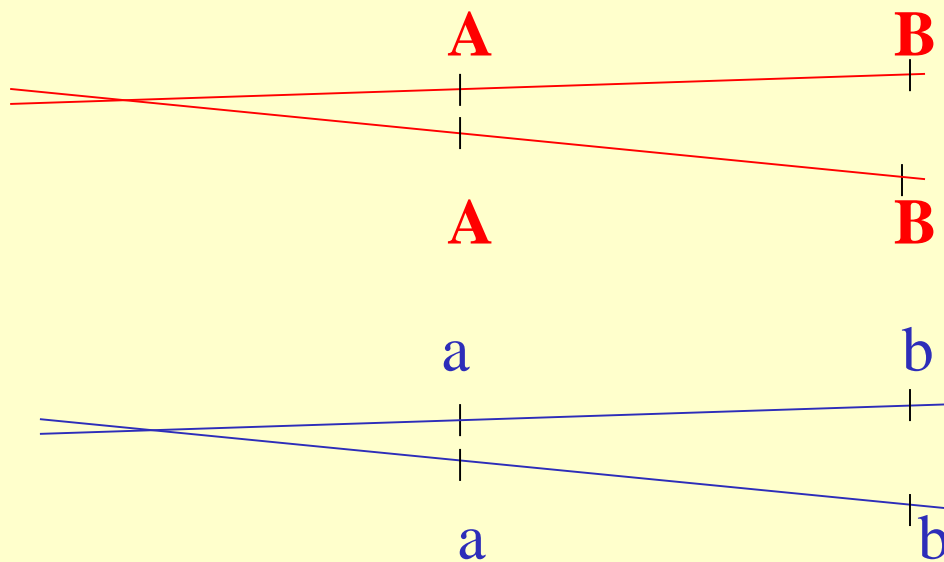
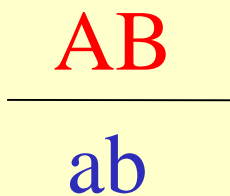
Vazbová fáze cis (coupling)

na jednom chromozómu alely dominantní,
na druhém recesivní

$$P: \frac{AB}{AB} \quad \times \quad \frac{ab}{ab}$$

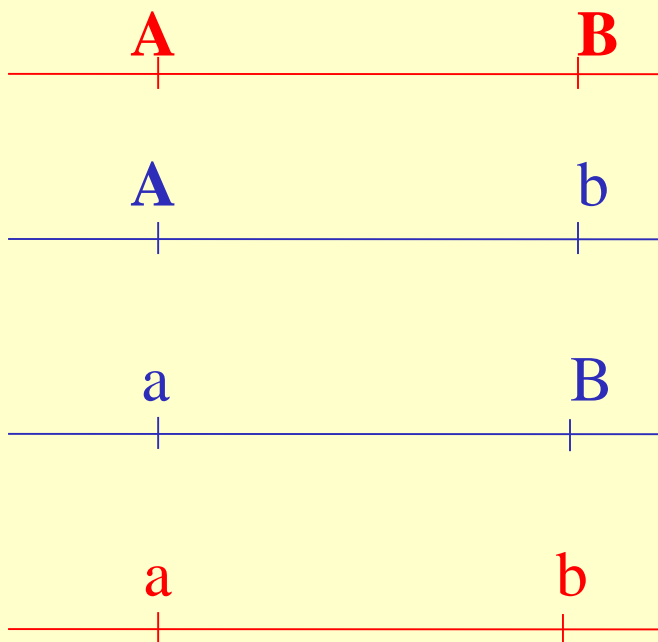
$$F_1: \frac{AB}{ab}$$

Vazbová fáze cis (coupling)



Vazbová fáze cis

Vznikají gamety rodičovské **AB, ab**
rekombinované **Ab, aB**



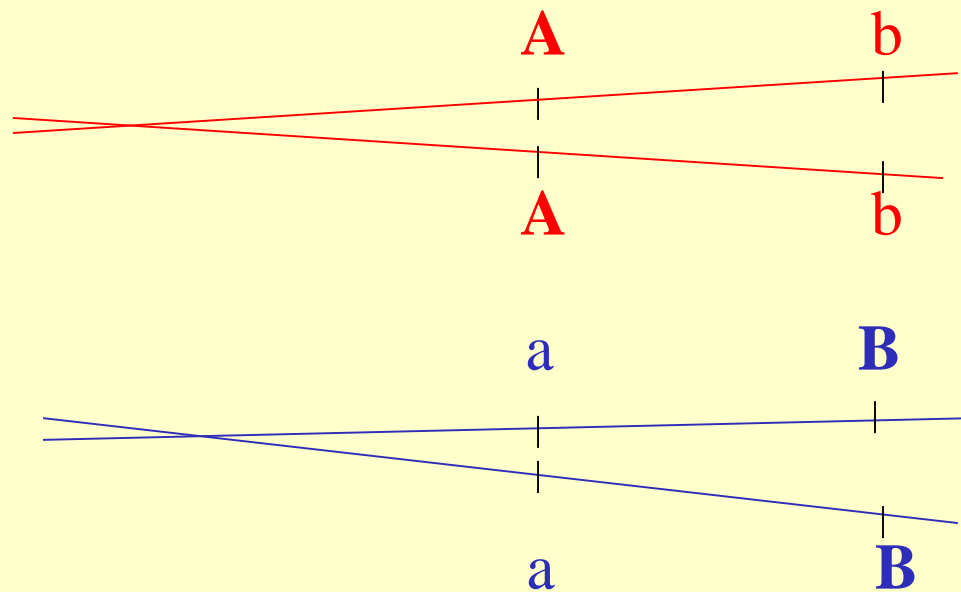
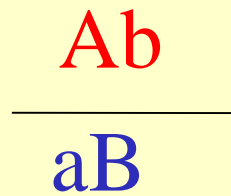
Vazbová fáze trans (repulsion)

na jednom chromozómu alela dominantní a recesivní, na druhém recesivní a dominantní

$$P: \frac{Ab}{Ab} \quad \times \quad \frac{aB}{aB}$$

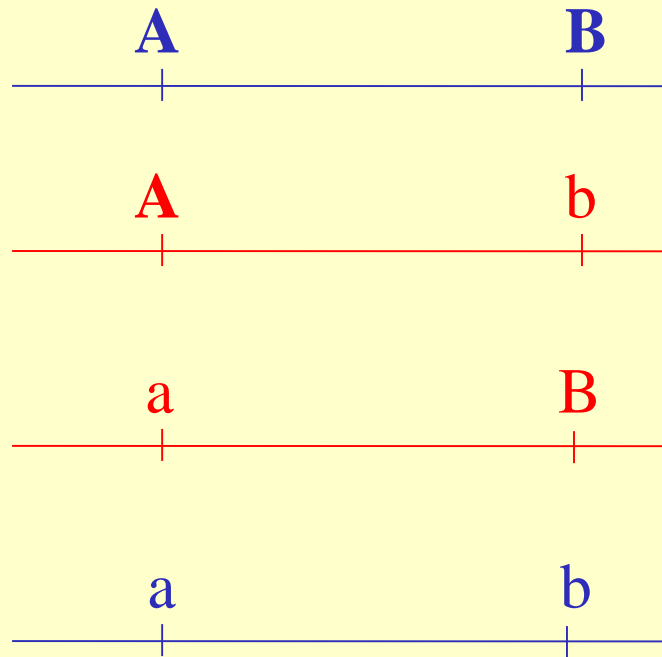
$$F_1: \frac{Ab}{aB}$$

Vazbová fáze trans (repulsion)



Vazbová fáze trans

Vznikají gamety rodičovské **Ab, aB**
rekombinované **AB, ab**



Vazbová fáze cis

Zpětné křížení **AB/ab** **x** **ab/ab**

gamety: **AB** **Ab** **aB** **ab** **ab**

genotypy:

AB/ab (a_1) **Ab/ab** (a_2) **aB/ab** (a_3) **ab/ab** (a_4)

pokud existuje **úplná vazba mezi lokusy „A“ a „B“**, je na
100 potomků

AB/ab (a_1) **Ab/ab** (a_2) **aB/ab** (a_3) **ab/ab** (a_4)

50

0

0

50

Vazbová fáze trans

Zpětné křížení **Ab/aB** **x** **aa/bb**
gamety: **AB** **Ab** **aB** **ab** **ab**

Při **úplné vazbě** je na 100 potomků

AB /ab (a_1)	Ab /ab (a_2)	aB /ab (a_3)	ab /ab (a_4)
0	50	50	0

Bez vazby

$AaBb (a_1) = 25$ $Aabb (a_2) = 25$ $aaBb (a_3) = 25$ $aabb (a_4) = 25$

A – nafialovělá barva
aleuronu kukuřičného zrna

a - žlutá barva

B – kulaté zrno

b - hranaté zrno

Testovací křížení $AaBb \times aabb$, **úplná vazba:**

cis: $AB/ab \times aabb$

50% zrn nafialovělých kulatých

(a_1) AB/ab

50% zrn žlutých hranatých

(a_4) ab/ab

**žádné fialové hranaté ani
žluté kulaté**

trans: Ab/aB

50% zrn nafialovělých hranatých

(a_2) Ab/ab

50% zrn žlutých kulatých

(a_3) aB/ab

**žádné fialové kulaté ani žluté
hrnaté**

Neúplná vazba

- výskyt rekombinantů, jejichž frekvence je vždy nižší než frekvence rodičovských nerekombinovaných genotypů;
- % rekombinant zjistíme z dvojnásobných testovacích křížení:

AaBb x aabb

Vzdálenost mezi geny – síla vazby

- cM centimorgan
- 1 cM je **pravděpodobnost rekombinací** mezi lokusy 1%

Morganovo číslo p

% podíl rekombinant ze všech informativních potomků

$$\text{cis : cM}(p) = \frac{a_2 + a_3}{a_1 + a_2 + a_3 + a_4} \times 100$$

$$\text{trans : cM}(p) = \frac{a_1 + a_4}{a_1 + a_2 + a_3 + a_4} \times 100$$

Musíme znát vazbovou fázi!

Batesonovo číslo c

Kolikrát častěji vznikají gamety rodičovské, než rekombinované

$$cis : c = \frac{a_1 + a_4}{a_2 + a_3} \qquad trans : c = \frac{a_2 + a_3}{a_1 + a_4}$$

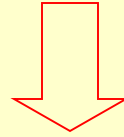
$$c = \frac{100 - p}{p}$$

Musíme znát vazbovou fázi!

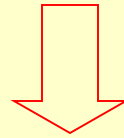
Určení vazbové fáze

z genotypu potomků zpětného křížení,
zpětná dedukce genotypů rodičů a
prarodičů

$B_1: AaBb \times aabb$



$AaBb (a_1)$	$Aabb (a_2)$	$aaBb (a_3)$	$aabb (a_4)$
40	10	10	40

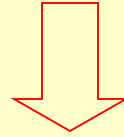


převaha **nerekombinant** a_1, a_4 , tzn. fáze

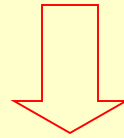
cis

$$p = \frac{10(a_2) + 10(a_3)}{40 + 10 + 10 + 40} \times 100 = 20\% = 20\text{cM}$$

$B_1: AaBb \times aabb$



$AaBb (a_1)$	$Aabb (a_2)$	$aaBb (a_3)$	$aabb (a_4)$
10	40	40	10



převaha **nerekombinant** a_2, a_3 , tj. fáze

trans

$$p = \frac{10(a_1) + 10(a_4)}{10 + 40 + 40 + 10} \times 100 = 20\% = 20\text{cM}$$

Určení vzdálenosti v cM

Vzdálenost mezi dvěma body, např. geny A, B na genetické mapě chromozomu se rovná průměrnému počtu crossing – overů mezi nimi během meiozy.

- Celkem 100 gamet – u 70 bez c.-o.
- u 20 jednoduchý c.-o.
 - u 8 dvojitý c.-o.
 - u 2 trojnásobný c.-o.

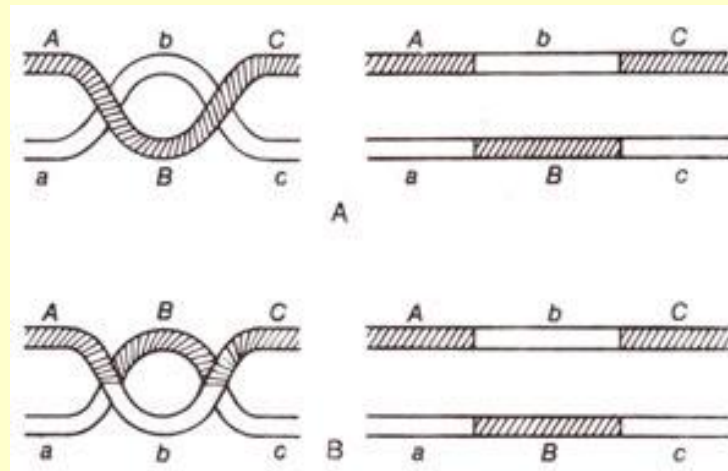
Průměrný počet crossing - overů mezi A a B

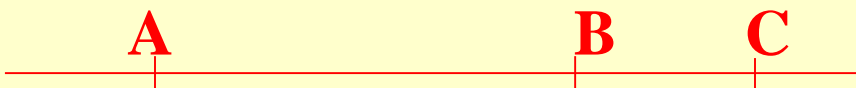
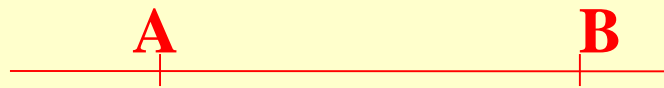
$$0 \times (70/100) + 1 \times (20/100) + 2 \times (8/100) + 3 \times (2/100) = 0,42 \text{ M}$$

V průměru došlo u 42 chromozomů ze sledovaných 100 chromozomů k výměně mezi oběma geny A - B.

Pořadí genů na chromozomu lze určit analýzou dvojitého crossing – overu:

tříbodový test





Lidský chromozóm č. 21 = 0,5 M (50 cM)

Lidský chromozóm č. 1 = 2 M (200 cM)

Celková mapová délka lidského genomu je cca 3000 cM u mužů,
4200 cM u žen.

Fyzická délka haploidního lidského genomu je asi 3×10^9 bp.

1cM odpovídá přibližně 1×10^6 bp u muže a 7×10^5 u žen.

Pro stanovení vzdálenosti genů je Morgan příliš velká jednotka.

1cM pojme asi 0,7 – 2 miliony párů bází.

Do 1cM se vejde několik stovek až tisíc genů.

Lineární vztah mezi % rekombinant a vzdáleností v cM je do cca **20-25 cM**. Kvůli **dvojitým** a **vícenásobným** crossing overům se vzdálenosti v genomu jakoby prodlužují.

Interference

Jsou rekombinace na jednom chromozómu na sobě nezávislé?

$$\text{Frekvence dvojitého c.-o.} = \frac{a \cdot b}{100} [\%]$$

a - frekvence jednoduchého c.-o. v jedné oblasti

b - frekvence jednoduchého c.-o. v druhé oblasti

$$\text{Př.: } 23,84 \cdot 6,16 / 100 = 1,47 \%$$

ale **skutečná** frekvence dvojit. c.-o. = 0,72%

Tj. **potlačení** vzniku dalšího c.-o. na chromozómu, kde již jeden c.-o. proběhl.

Příčina – mechanické schopnosti chromozómu překřížit se na více místech. Další příčiny, protože interference se vyskytuje i mezi geny vzdálenými více než 30 cM.

Interference

se vyjadřuje jako

$$\text{Koef. koincidence} = \frac{\text{skutečný podíl dvoj. c.-o.}}{\text{teor. podíl dvoj. c.-o.}}$$

$$0,72 / 1,47 = 0,49$$

Počet dvojnásobných a vícenásobných crossing-overů do 20 cM je zanedbatelný.

Nad 20-25 cM začíná být mapování nepřesné, stanovení vzdálenosti genů na základě **dílčích rekombinačních frekvencí není aditivní!**

Nad 25 cM je nutno genetickou vzdálenost korigovat:

Kosambiho mapovací funkce

Dále, v mnoha částech genomu savců existuje cca 2 x vyšší frekvence rekombinací u **homogametního** pohlaví (samic) než u samců (viz výše).

Např. u sameček drosofily neprobíhá rekombinace vůbec!

Délkové a rekombinační hodnoty jsou však zkresleny také nerovnoměrnou frekvencí rekombinací v různých částech genomu, v jednotlivých chromozomech a jejich segmentech.

Existují tzv. horká místa (**hot spots**) s častou tvorbou chiasmat, na druhé straně dlouhé segmenty 15-20 cM, v nichž je suprese rekombinací a pokud nastanou, mají často patologické následky, např. T/t genový komplex u myši, projevující se řadou malformací a infertilitou.

Nevýhody výpočtu Morganova čísla odstraňuje
metoda

LOD skóre

LOD skóre:

- lze určit existenci vazby a souč. její sílu
- výsledky segregace potomků z různých křížení lze jednoduše akumulovat
- k vyhodnocení stačí dvougenerační rodokmen (P a F1), v P nemusí být známa vazbová fáze
- pro důkaz vazby stačí minimální počet potomků
- potomci nemusí pocházet jen z dvojnásobných testovacích křížení, tj. $AaBb \times aabb$, ale i z jiných typů, jednoduché nebo dvojnás. intercrossy $AaBb \times Aabb$; $AaBb \times AaBb$ bez dominance či s dominancí v jedné nebo obou alelách.

LOD skóre

- statistická metoda pro detekci vazby z poměru dvou pravděpodobností
- **L**og logaritmus
- **OD**s šance pro převahu pravděpodobnosti
- skóre je zde poměr skutečných nebo hypotetických rekombinant k nerekombinantám

Dekadický logaritmus poměru pravděpodobností, šance, že pozorovaná část potomků vznikla jako výsledek genetické vazby k pravděpodobnosti, že vznikla náhodou.

Porovnávají se dvě alternativní hypotézy:

H₁ vazba existuje, rekombinační zlomek v intervalu

$0 < 1/2$, přitom 0 znamená úplnou vazbu

H₀ vazba neexistuje, volná kombinovatelnost vloh,

rekombin. zlomek $1/2$

V čitateli je hodnoceno konkrétní skóre ve prospěch existence vazby a ve jmenovateli je zvažováno skóre, jakoby vazba neexistovala.

LOD skóre

$$\sum Z = \log \frac{\text{pravd. sekvence s danou hodnotou v azby}}{\text{pravd. vzniku sekvence bez vazby}}$$

Výsledek Lod skóre z jednoho typu křížení (páření, rodiny) se sčítá s dalšími $Z_1, Z_2, Z_3 \dots \Sigma Z$.

$$LOD = Z = \log_{10} \frac{\text{probability of birth sequence with a given linkage value}}{\text{probability of birth sequence with no linkage}} = \log_{10} \frac{(1 - \theta)^{NR} \times \theta^R}{0.5^{(NR+R)}}$$

NR je počet nerekombinant v potomstvu

R počet rekombinant

0.5 ve jmenovateli znamená, že jakékoliv alely, které jsou zcela volně kombinovatelné, tzn. že jsou na různých chromozómech, mají 50% šanci rekombinace v důsledku nezávislého třídění

θ (theta) je rekombinační zlomek, tj. podíl porodů v nichž proběhla rekombinace mezi sledovanými genetickými markery, tj. je roven $R / (NR + R)$

Pravděpodobnost θ

(theta)

podíl rekombinovaných gamet ze všech sledovaných

- 0.00 při úplné vazbě bez rekombinace
- 0.50 vzdálenost genů je velká - volná kombinovatelnost

Je – li LOD skóre > 3 , vazba na autozomu je přítomna (pro vazbu na chrom. X stačí > 2), protože pravděpodobnost existence vazby je 1000x větší než pravděpodobnost nulové hypotézy.

Je- li LOD skóre < -2 , vazba není přítomna, je volná kombinovatelnost, vazba je 100x méně pravděpodobná než žádná vazba.

Je-li vypočtená hodnota mezi -2 a 3 , nelze učinit rozhodnutí, je nutné pokračovat v analýzách.

Je velmi nepravděpodobné, že LOD skóre nad 3 bude dosaženo u jednoho rodokmenu.

Matematické vlastnosti testu dovolují kombinaci dat z různých rodokmenů součtem jejich LOD skóre.

Příklad

$N = 632$ počet potomků ze zpětného testovacího křížení ($AaBb \times aabb$)

$10 + 19 = 29$ počet rekombinantů = r

$318 + 285 = 603$ počet nerekombinantů = $N - r$

$d = \text{rekombinantní zlomek} = r/N = 29/632 = 0,046$

$$\text{LOD} = \log_{10} (0,046^{29} * (1-0,046)^{603} / 0,5^{632}) = 242$$

vazba je přítomna

Vazbová nerovnováha (linkage disequilibrium)

- nerovnoměrná frekvence 4 možných kombinací (AB, Ab, aB, ab) v populaci
- častější výskyt kombinací 2 alel

příčiny:

- evoluční výhoda (pouze pro cis)
- rovnováha dosud nebyla navozena pro krátký evoluční interval

Biologický význam vazby

- uchování stálé sestavy funkčně spřízněných genů
- evoluční konzervace genových rodin (funkční geny + pseudogeny)
- funkční záloha – změnou podmínek prostředí mohou být pseudogeny „zapnuty“

Biologický význam rekombinace

- zvyšování proměnlivosti při uchování vazbové lokalizace genů
- vznik nových genotypů

Biologický význam vazby a rekombinace

Pro schopnost populace přizpůsobit svůj genofond změněným podmínkám má význam i uspořádání genů na chromozomech.

Na určitou vlastnost působí geny XYZ. Je-li optimální fenotyp intermediární, je optimální genotyp heterozygotní $XxYyZz$. Potom je výhodná velmi silná vazba, protože nejvýhodnějším genotypem je XyZ/xYz (trans), zaručuje heterozygotnost, suboptimální genotypy vznikají s mnohem menší pravděpodobností než při volné kombinovatelnosti vloh, zachovává se maximální genetická variabilita.

Dojde-li k takové změně prostředí, že optimální bude homozygotní genotyp XYZ/XYZ nebo xyz/xyz , je vazba méně výhodná, neboť k přechodu na vazbovou fázi cis XYZ/xyz je nutný crossing-over.

Mutací vzniklé alely tedy musí být rekombinacemi zapojeny do genofondu populace.

Mapování genomu

Chceme znát účinek genů podílejících se na určité vlastnosti. Protože podíl strukturních genů z genomu je malý (2-3%), obsahují mapy kromě kódujících lokusů

rovněž nekódující

lokusy (mikrosatelity, SNP).

Mapování genomu

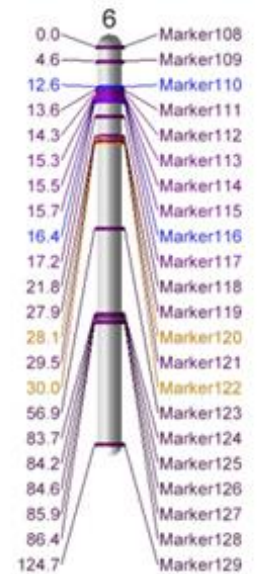
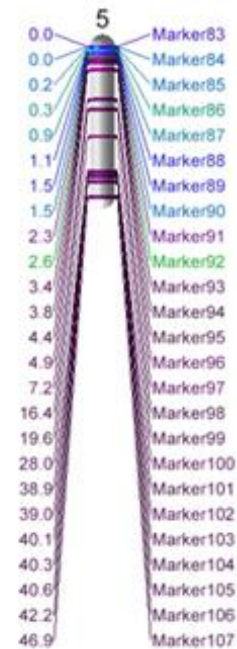
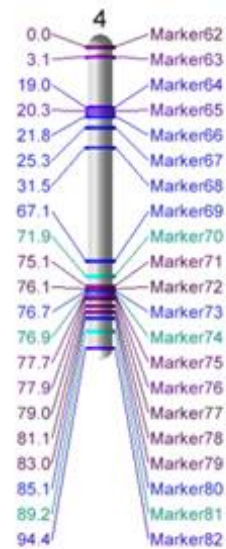
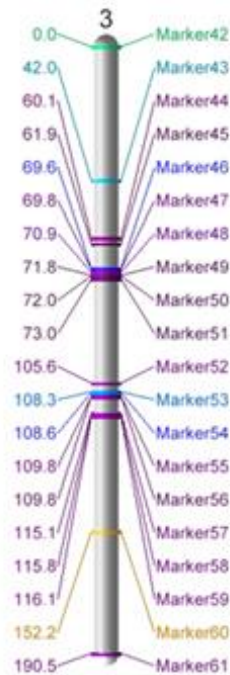
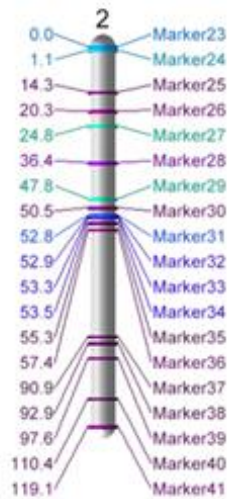
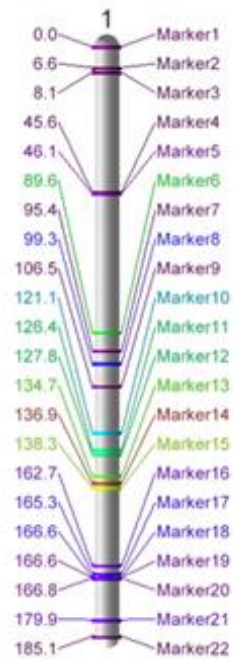
- určení pořadí genů
- zjištění umístění genů na chromozómech
- stanovení nukleotidové posloupnosti genů nebo celého genomu

Typy map

- genetické (rekombinační, vazbové)
- fyzické
- cytogenetické
- kombinované

Genetické (rekombinační, vazbové) mapy

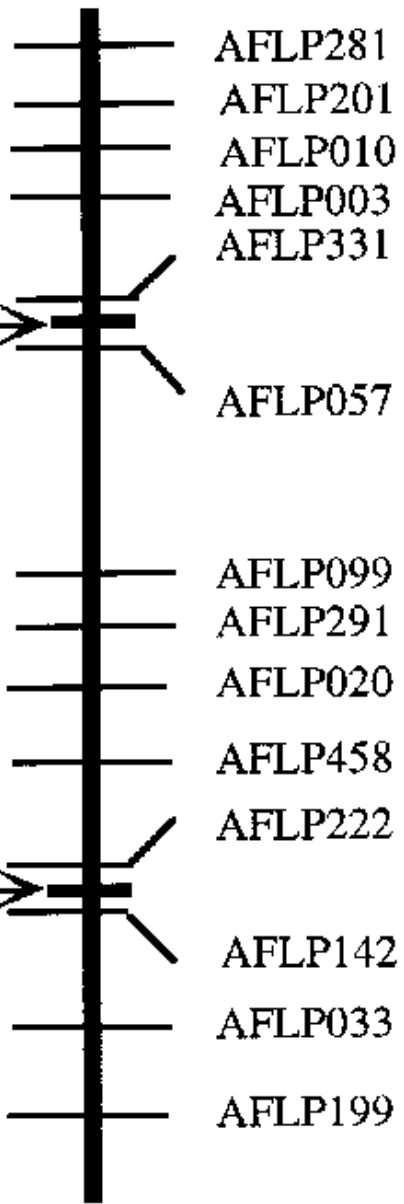
- neúplné mapy, udávají pořadí a vzdálenosti genů nebo nekódujících markerů
- vzdálenosti vyjádřené v **cM**
- založené na analýze frekvencí rekombinací pomocí dvoubodového nebo třibodového testu.
- Nyní detekce vazby a vzdáleností pomocí LOD skóre





← trait A

← trait B



AFLP281
AFLP201
AFLP010
AFLP003
AFLP331

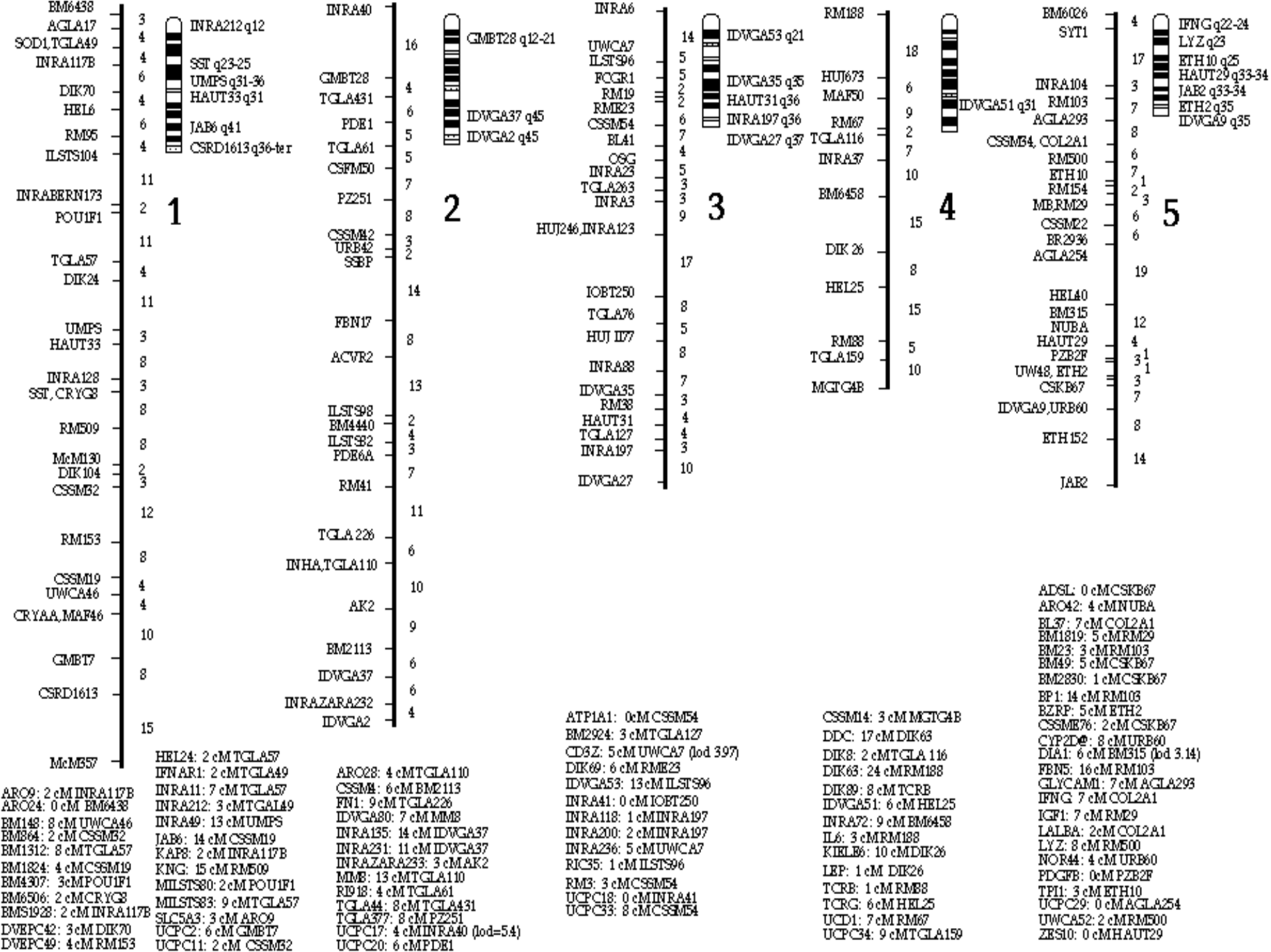
AFLP057

AFLP099
AFLP291
AFLP020
AFLP458
AFLP222

AFLP142
AFLP033
AFLP199

↙ DNA Markers
↘ Linked to A

↙ DNA Markers
↘ Linked to B

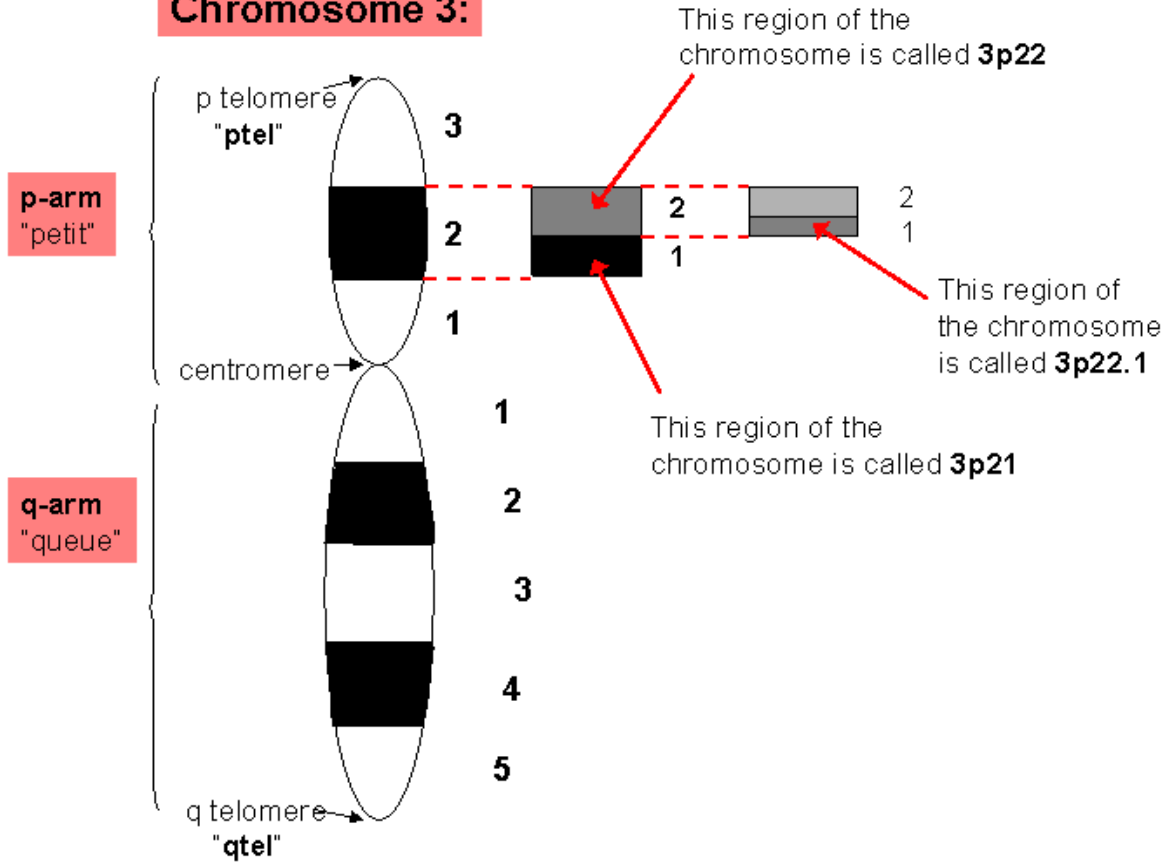


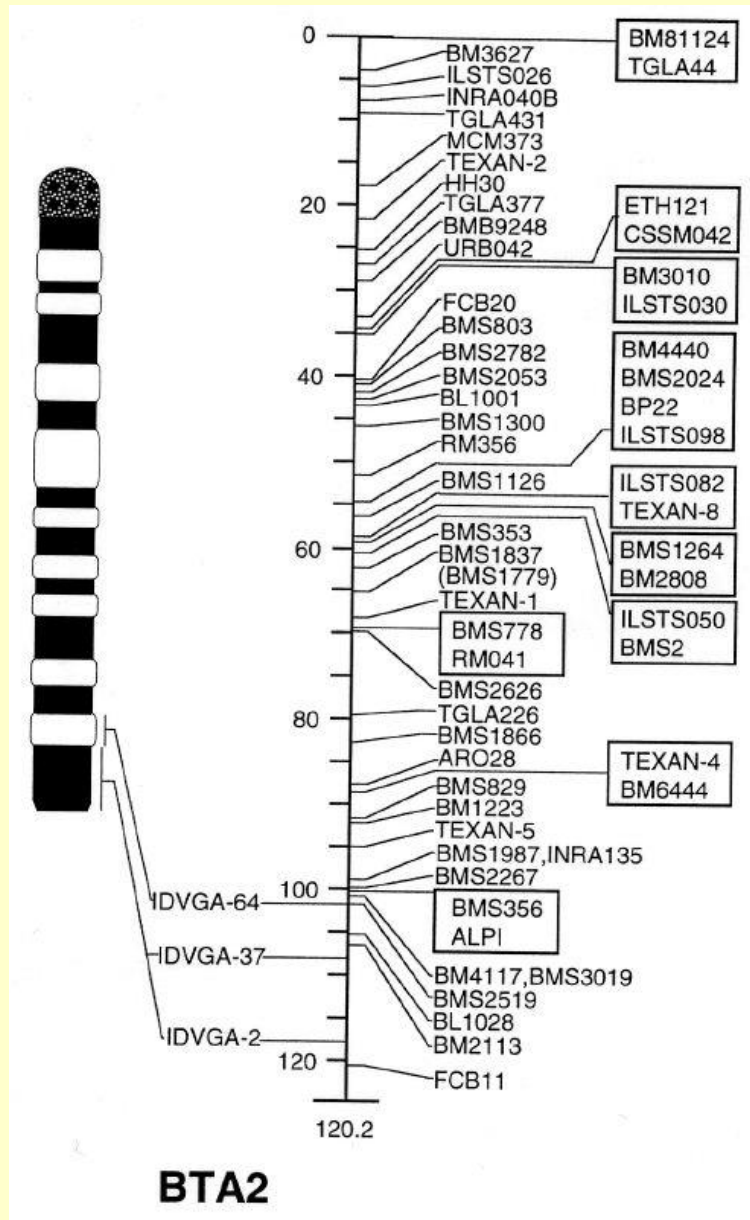
Chromozómové (cytogenetické) mapy

- založené na testování karyotypu buď fixovaného na skle, nebo pomocí průtokové cytometrie.
- poloha genu dána číslem chromozómu a pruhem (pruhovací, banding techniky).
- ev. hybridizace *in situ* se značenou sondou. Značení radioaktivně nebo fluorescenčně (FISH).
- ev. využití panelu hybridních somatických buněk.

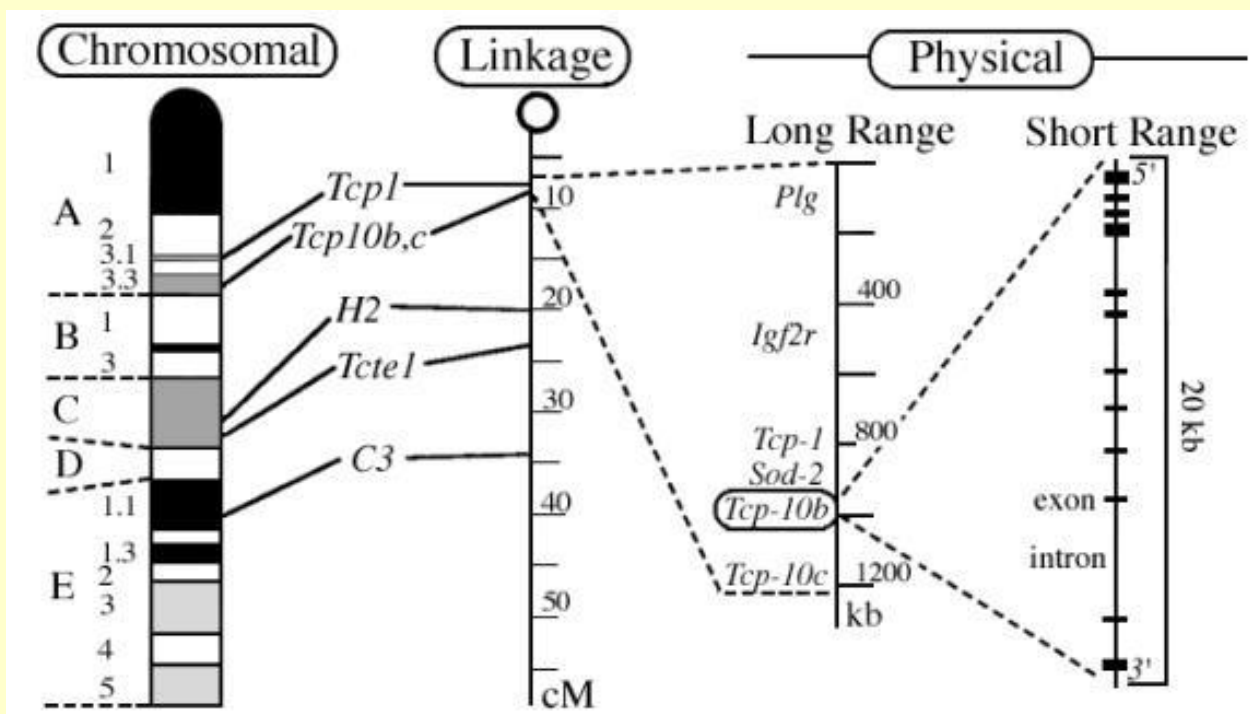
Cytogenetic Banding Nomenclature

Chromosome 3:





Integrované mapy



Integrovaná chromozómová (cytogenetická) a
vazbová mapa Bta 1 (bovinní chromozóm 1).

U10

APP
C9
CBS
CD18
COL6A2
CP
D1S3 (GMBT59)
D1S5 (TGLA52)
D1S7 (TGLA415)
D1S15 (INRA049)
D1S16 (INRA054)
D1S17 (INRA128)
D1S20 (INRA117)
D1S21 (HEL6)
D1S37 (CSSM004)
D1S38 (CSSM011)
D1S39 (CSSM024)
ETS2
FIM3
GAP43
GART
IFNBR
PFKL
S100B
SI
SMPP
UPK1B



1

COL6A1

SST

UMPS

D1S11 (DSMS46)

D1S1 (CSRD1613)

D1S12 (cIOBL26)

Barendse et al., 1994

IFNAR
SOD1
D1S14 (TGLA49)

PIT1
D1S13 (RM095)

D1S8 (TGLA57)

CRYG8

D1S9 (CSSM032)

D1S10 (CSSM019)

D1S4 (MAF46)

CRYA1

D1S2 (GMBT7)

Bishop et al., 1994

D1S24 (BM6438)
6.4
D1S25 (AGLA17)

23.5

10.2

D1S26 (ILST004)

5.3

D1S27 (BM4307)

17.4

D1S6 (INRA011)

D1S28 (BM1312)

13.5

D1S29 (BM6506)

18

D1S30 (BR2724)

D1S31 (BM864)

9.8

D1S32 (TGLA130)

6.9

D1S33 (BL28)

8.6

D1S34 (BM1824)

6.5

TF
D1S36 (BM148)

20.3

D1S35 (BM3205)

Fyzické mapy

Úplné - sekvenční

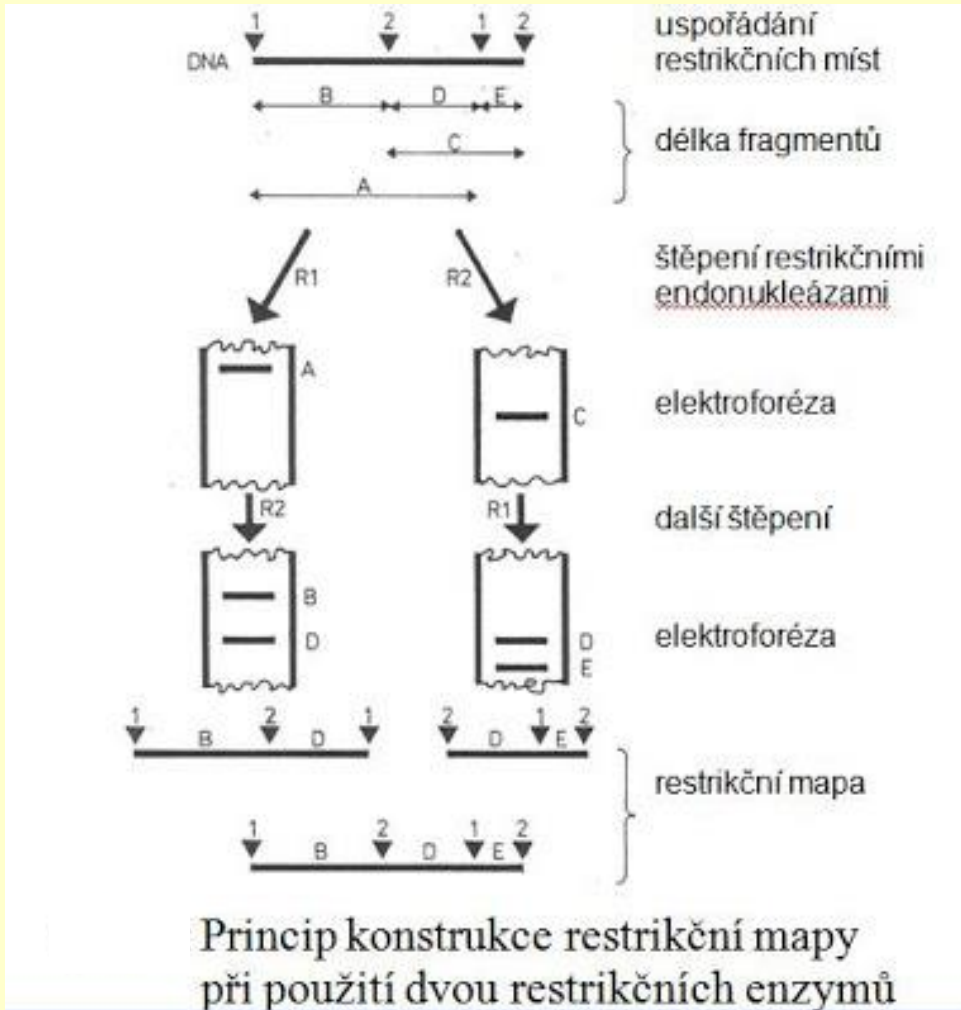
- založeny na sekvenování, tj. zjištění posloupnosti nukleotidů
- jednotkou je nukleotid

Neúplné

Neúplné mapy, např.:

Restrikční

schématické znázornění poloh a vzdáleností restrikčních míst na molekule DNA, vzdálenost mezi místy – počty nukleotidů (kilobáze)



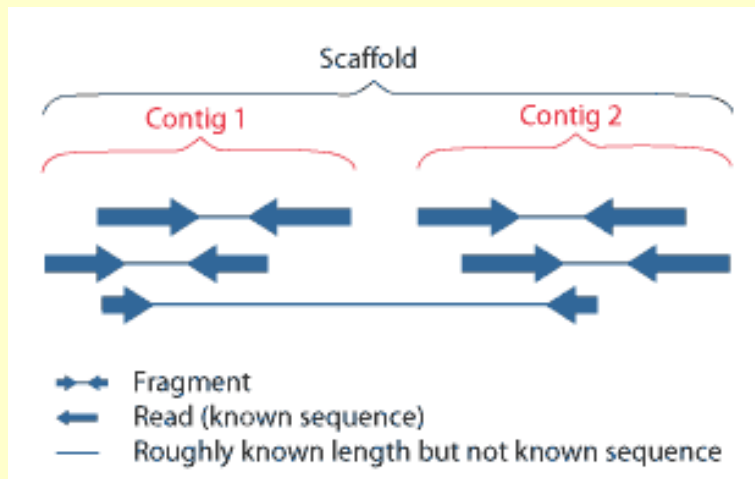
Výhodou je mj. že umožňuje porovnat fyzickou mapu chromosomu s mapou genetickou.

Přítomnost či absence restrikčního místa na zkoumaném úseku DNA může být rovněž použita k rekonstrukci fylogeneze, umožňují srovnávat DNA příbuzných i nepříbuzných jedinců

Contigové mapy

Každý chromozom je jednou contigovou mapou

- Contig - překrývající se sady klonů
- a) nejprve se připraví restriční mapy jednotlivých klonů (klony jsou náhodně klonované fragmenty 200 – 500 kb)
- b) počítač vyhledá případné překryvy (klony vždy vykazují několik společných restričních fragmentů).
- c) překrývající se klony se sestaví do kontigových map



Scaffold je tvořen contigy a mezerami

Kombinované mapy

- kombinace rekombinačních s fyzickými
- kombinace s cytogenetickými mapami

Příklad: STS, EST

kombinace sekvenčního a rekombinačního postupu

- STS sequence tagged site, „poloha definovaná pomocí specifické sekvence“. Krátké sekvence DNA 200 až 500 bp, které se vyskytují jednou v genomu a jejichž lokalizace a sekvence bází jsou známy. Mohou být snadno detekovány v PCR za použití specifických primerů. Proto jsou užitečné pro konstrukci genetických a fyzických map ze sekvenčních dat poskytnutých mnoha různými laboratořemi.
- EST expressed sequence tag, „místo na cDNA definované specifickou sekvencí“. K amplifikaci použita jen cDNA.

Cíl kombinovaných map:

sjednocení genetických map s fyzickými
i cytogenetickými