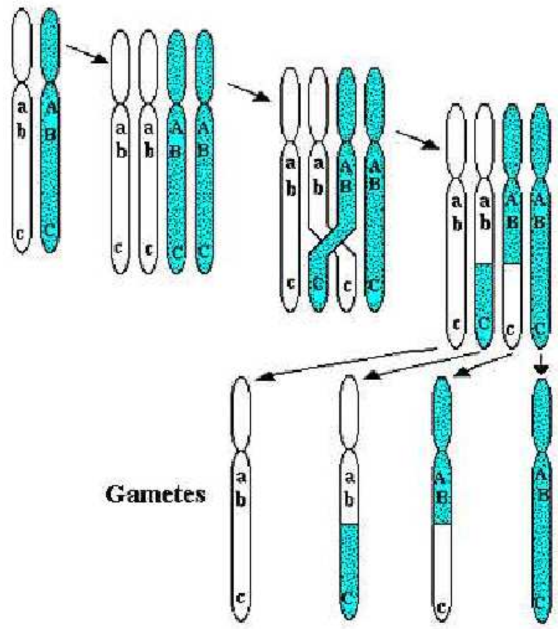
Vergleich von Genomen



Crossing-over and recombination during meiosis

Auf Basis der Anordnung der Gene

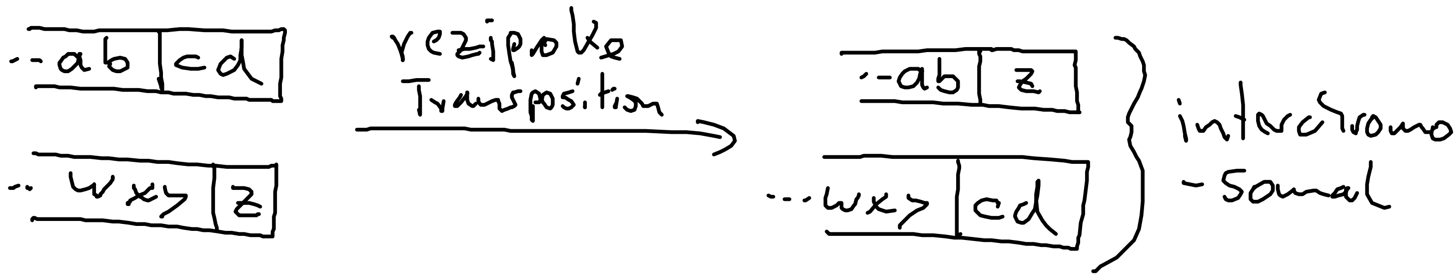
Beispiele für Änderungen der Anordnung durch Evolution

(*)

x | a b c | y $\xrightarrow{\text{Reversal}}$ x | c b a | y

w | z ... x | a b c | y ... $\xrightarrow{\text{Transposition}}$ w | a b c | z ... x | y ...

(x) intrachromosomale Transformationen



Modell: Permutation der Gene (Namen) ohne doppelte Elemente

Def: $\pi=(\pi_1,\dots,\pi_n)$ Permutation der Ordnung n.
$1 \le i < j \le n$ dann (i,j)-Reversal eine Permutation $\rho(i,j)$ mit

$$\pi\cdot\rho(i,j) = (\pi_1,\dots,\pi_{i-1},\pi_j,\pi_{j-1},\dots,\pi_i,\pi_{j+1},\dots,\pi_n)$$

Problem: Sortiere beliebige Permutation mit minimaler Anzahl Reversals.
MINSR

Bsp.

2 1 3 7 5 4 8 6

1 2 3 7 5 4 8 6

1 2 3 4 5 7 8 6

1 2 3 4 5 $\underline{7\ 6}$ 8

<u>Satz:</u> Das Problem MINSR ist NP-schwer.

<u>Def:</u> Für Permutation π der Ordnung $\underline{n}$ def. die erweiterte Darstellung $ext(\pi)$ über $\pi_0 := 0$
$\pi_{n+1} := n+1$, d.h. $ext(\pi) = (\pi_0, \pi_1, \dots, \pi_n, \pi_{n+1})$

<u>Def:</u> π Permutation der Ordnung n. Ein Breakpoint von π ist ein Paar $(i, i+1) \in \{0, \dots, n\} \times \{1, \dots, n+1\}$ mit $|\pi_i - \pi_{i+1}| \neq 1$.
Mit $brp(\pi)$ bezeichnen wir Anzahl breakp. von π.

<u>Bsp:</u> $\pi = (4, 3, 2, 7, 1, 5, 6, 8)$

$ext(\pi) = (0, 4, 3, \dots, 6, 8, 9)$

$0 \mid 4\ 3\ 2 \mid 7 \mid 1 \mid 5\ 6 \mid 8\ 9$

<u>Lemma:</u> Sei π Permutation der Ordnung n. Dann sind mind. $\left\lceil \frac{brp(\pi)}{2} \right\rceil$ Reversals nötig, um π zu sortieren.

<u>Beweis:</u> Identische Perm. ist die einzige ohne Breakpoints.

Ein (i,j)-Reversal kann nur die Breakpoints $(i-1,i)$ und $(j,j+1)$ beeinflussen. $\Rightarrow$ Jeder Reversal eliminiert höchstens 2 Breakpoints. $\boxtimes$

<u>Def.:</u> π Permutation der Ordnung n. Sei $k = bp(\pi)$ mit $(i_1, i_1+1), (i_2, i_2+1), \ldots, (i_k, i_k+1)$ die BPs von π. Dann nennen wir die $k+1$ Folgen

$S_0 = (\pi_0, \ldots, \pi_{i_1}),\ S_1 = (\pi_{i_1+1}, \ldots, \pi_{i_2}), \ldots, S_k = (\pi_{i_k+1}, \ldots, \pi_{n+1})$

die Strips von π.

Strip aufsteigend, falls $\pi_{j_i+1} < \ldots < \pi_{j_{i+1}}$ sonst absteigend.

Strip der Länge 1 per Def. absteigend außer im Fall, dass S_0 oder S_k nur 1 Element diese dann aufsteigend.

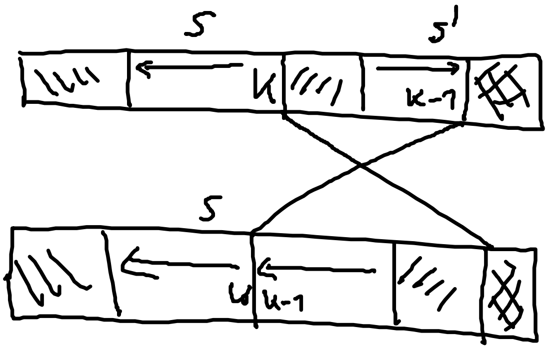
<u>Lemma:</u> π Permutation der Ord. n, $k \in \{0, \ldots, n+1\}$ ein Element aus $ext(\pi)$.

(a) Falls k in einem absteigenden Strip von $ext(\pi)$ liegt und $k-1$ in einem aufsteigenden, dann existiert Reversal ρ, so dass $bp(\pi\rho) < bp(\pi)$

(b) Falls l in einem absteigenden Strip und $l+1$ in einem aufsteigenden, dann $\exists\, \sigma$ mit $bip(\pi\sigma) < bip(\pi)$

Beweis:

(a) Sowohl k als auch $k-1$ sind die Letzten Elemente in ihrem Strip



analog für umgekehrte Anordnung von s & s'

(b) Hier l bzw. $l+1$ erstes Element ihres Strips und Reversal in Analogie zu (a) bringt l und $l+1$ an benachbarte Positionen bei angepa. Orientierung beider Strips $\square$

<u>Lemma:</u> Sei π Permutation mit einem absteigenden Strip. Dann existiert Reversal ρ mit $bip(\pi \cdot \rho) < bip(\pi)$.

Beweis: Wir wählen k als das kleinste Element das in π in einem absteigenden Strip vorkommt. Dann ist $k-1$ Element eines aufsteigenden Strips.

Fall $k=1$ gem. Def. entsprechend.

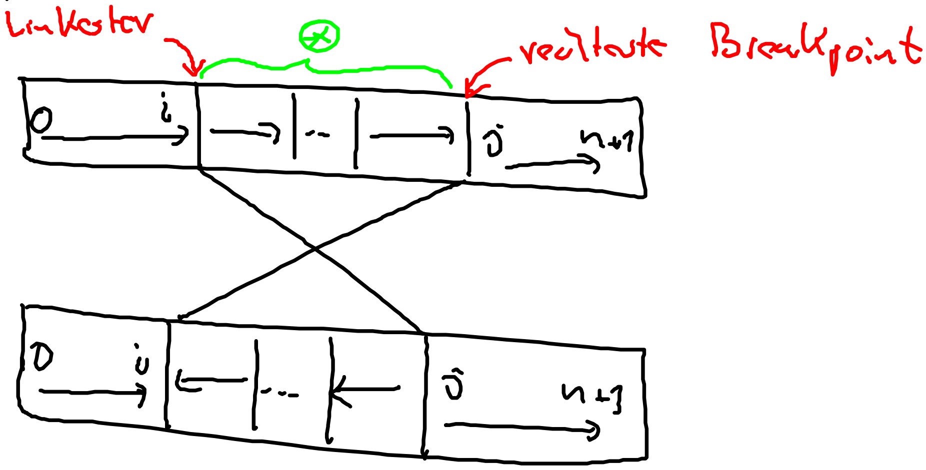
$\curvearrowright$ Wende Lemma zuvor (a) an

□

Lemma: Sei π eine Permutation ohne absteigenden Strip. Dann ist π die identische Permutation oder es existiert Reversal ρ mit $\pi \cdot \rho$ enthält einen absteigenden Strip und $brp(\pi\cdot\rho) \leq brp(\pi)$.

Beweis: π nicht identische Permutation

$\curvearrowright$ π mind. 2 Breakpoints.



⊛ nicht leer, da linkester $\neq$ rechtester BP.

□

Lemma: Sei π Permutation mit einem absteigenden Strip. Sei k das kleinste Element in einem absteigenden Strip und l das größte Element in einem abst.

Strip von π.

Sei ρ das Reversal das $k-1$ neben k plaziert und σ Reversal das $l+1$ neben l plaziert. Wenn sowohl $\pi\rho$ als auch $\pi\sigma$ keinen absteigenden Strip enthalten, dann gilt $\sigma = \rho$ und $bp(\pi\rho) = bp(\pi) - 2$.

Beweis: k kleinste in abst. Strip

$\rightarrow k-1$ liegt in aufst. Strip.

$\Longrightarrow$ Lemma von oben

l größte Element in abst. Strip

$\Longrightarrow l+1$ in einem aufsteigenden strip.

$\Longrightarrow$ Lemma von oben.

K muß in dem von σ ungedrehten Intervall liegen, da sonst adst. Strip. Analog muß L in dem von ρ ungedrehten Intervall liegen.

Annahme $\rho \neq \sigma$, dann gibt es einen Strip, der zu dem Intervall von nur einem der beiden Reversals gehört. Falls dieser Strip in π aufsteigend, dann ist er in entweder $\pi \cdot \rho$ absteigend, (oder $\pi \cdot \sigma$ absteigend) Falls ein in π absteigend ist, bleibt er in $\pi \cdot \sigma$ $(\pi \cdot \rho)$ erhalten oder wird höchstens verlängert. Beides nach Voraussetzung unmöglich.

$\Longrightarrow \sigma = \rho$

Damit eliminiert $\rho = \sigma$ 2 BP, da sowohl K und K-1 ab auch L und L+1 in $\pi \cdot \rho$ aufeinander folgen.

$\square$