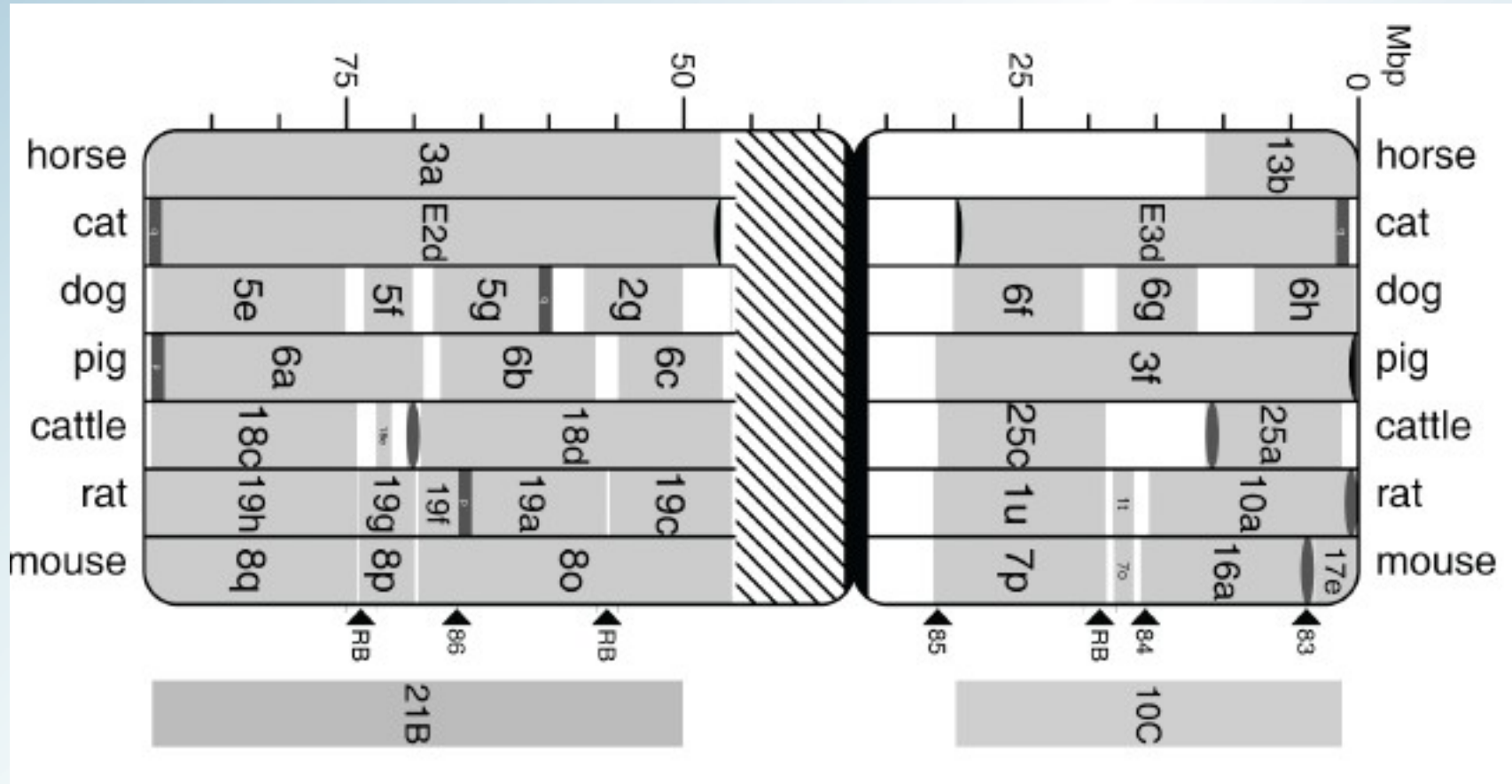




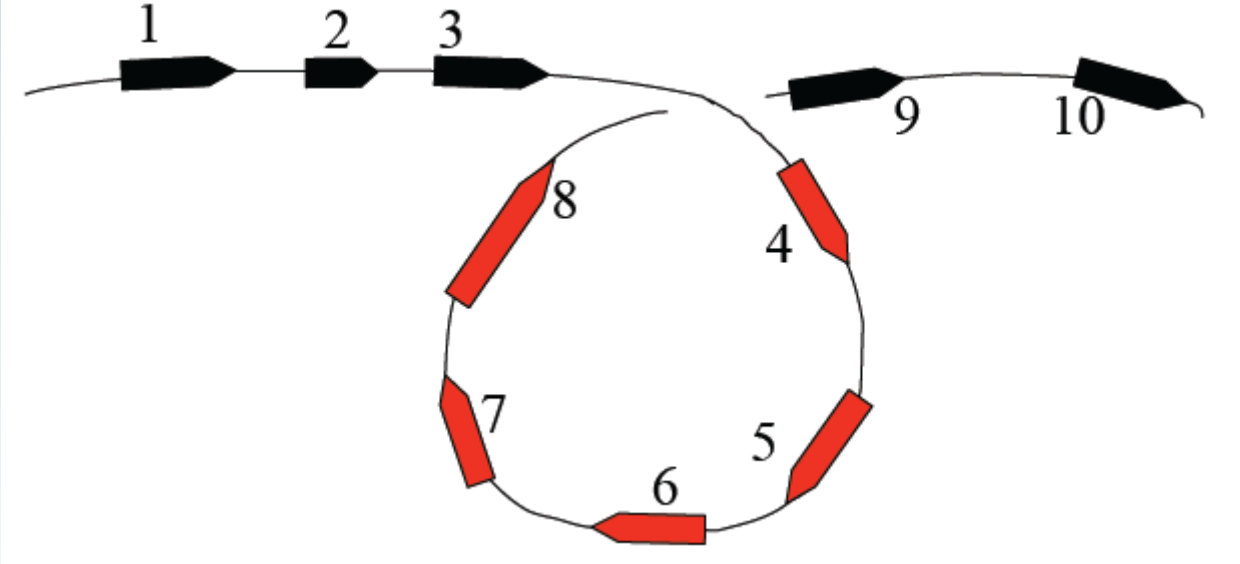
Evoluzione cromosomica

Antonio Miraglia
Raffaella Brasino

Intro

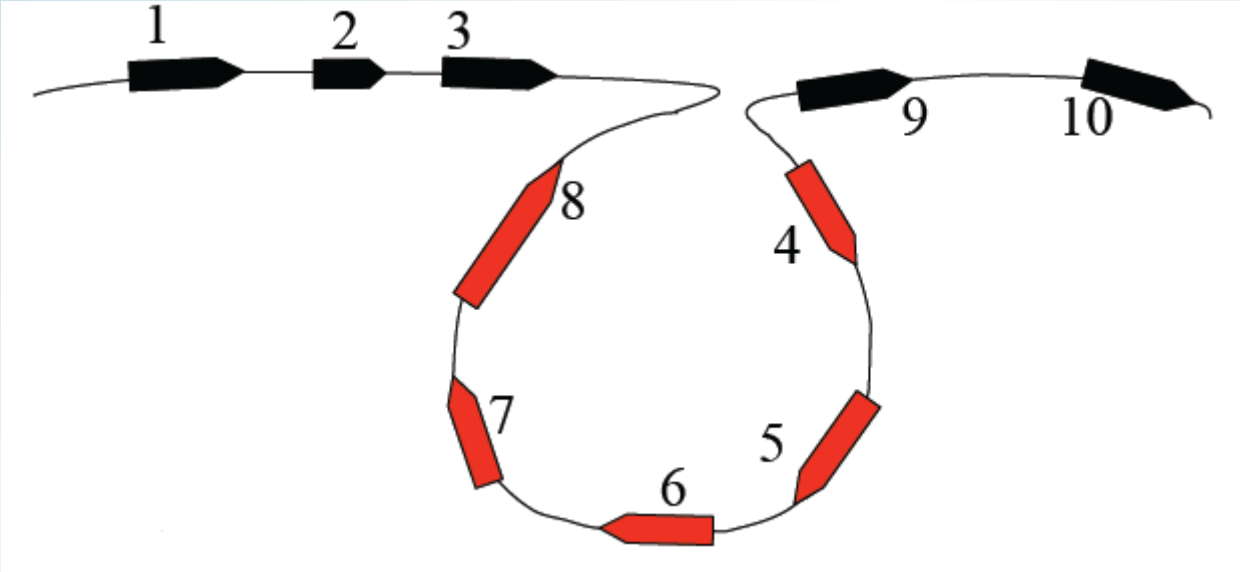


Riarrangiamento



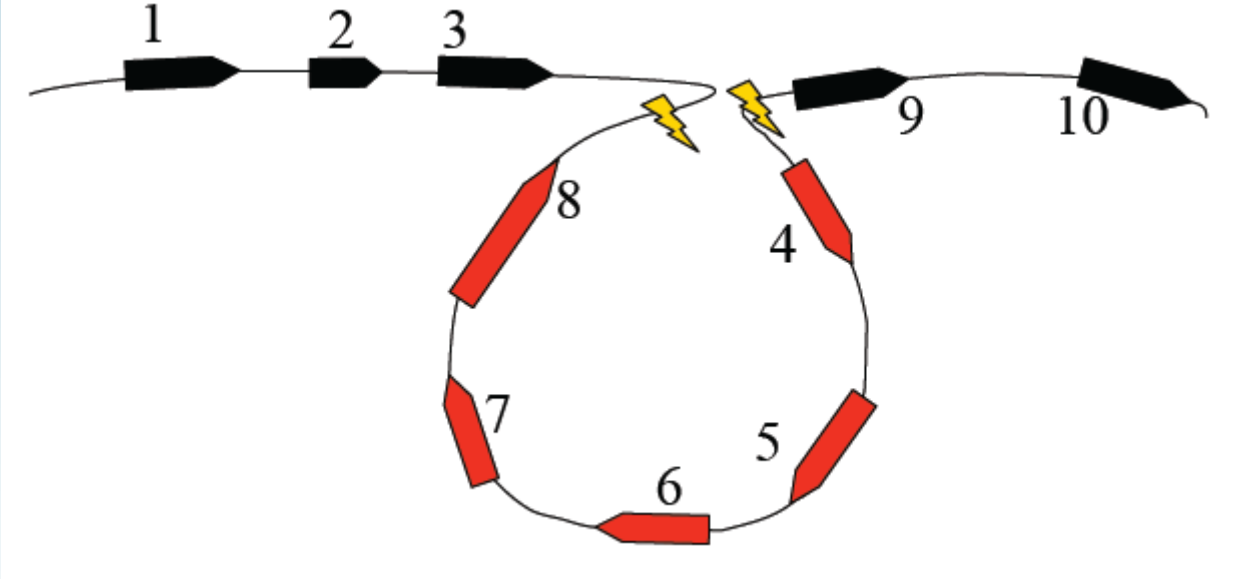
1 2 3 4 5 6 7 8 9 10

Riarrangiamento



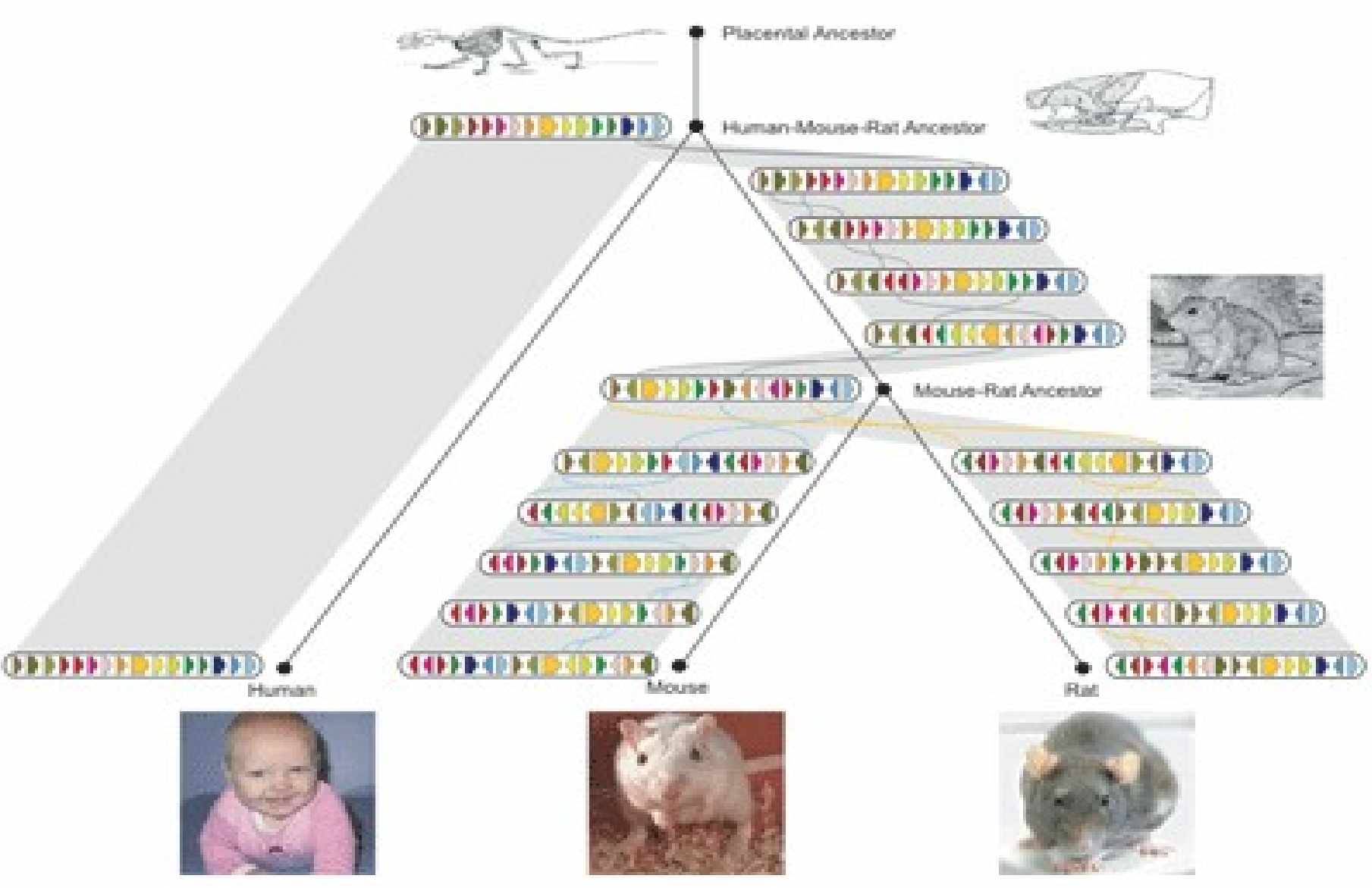
1 2 3 -8 -7 -6 -5 -4 9 10

Riarrangiamento



1 2 3 BP -8 -7 -6 -5 -4 BP 9 10

Gemoma uomo/topo



Synteny blocks uomo/topo

Blocchi di sintenia tra il genoma umano (A) e quello di topo (B)

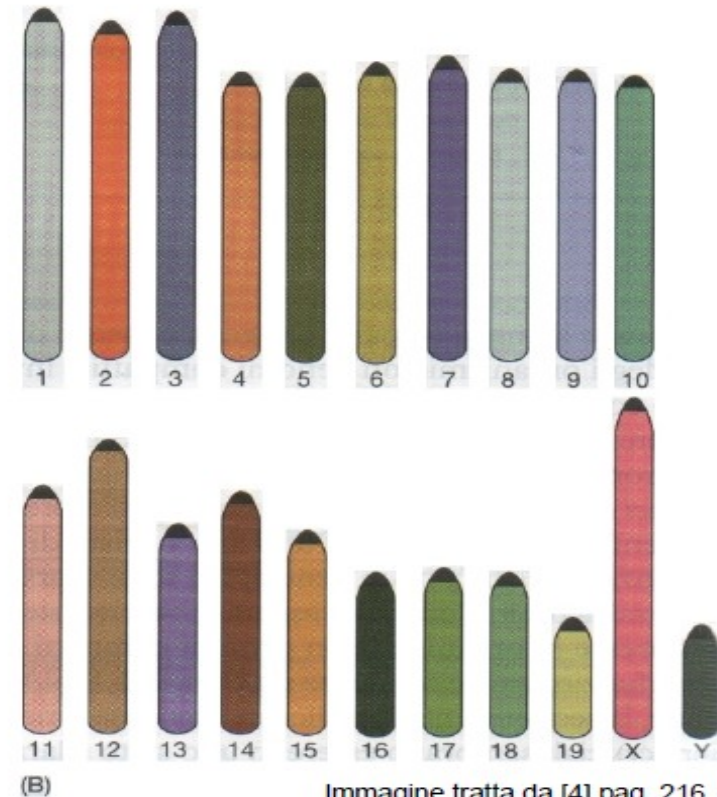
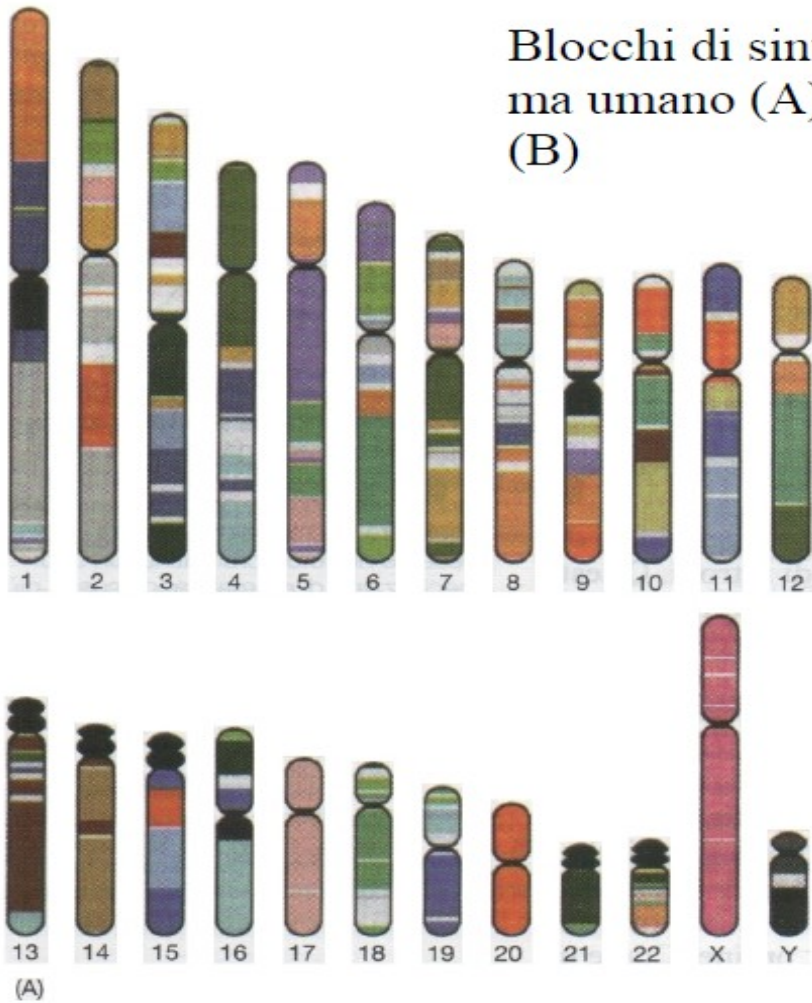


















Immagine tratta da [4] pag. 216

Distanza riarrangiamento

Step 0: π	2	<u>-4</u>	<u>-3</u>	5	-8	-7	-6	1
Step 1:	2	3	4	5	-8	-7	-6	1
Step 2:	-5	-4	-3	-2	<u>-8</u>	-7	-6	1
Step 3:	-5	-4	-3	-2	-1	6	7	8
Step 4: γ	1	2	3	4	5	6	7	8

$$\begin{aligned}
 \pi &= \pi_1 \dots \pi_{i-1} \pi_i \pi_{i+1} \dots \pi_{j-1} \pi_j \pi_{j+1} \dots \pi_n \Rightarrow \\
 &\pi_1 \dots \pi_{i-1} \pi_j \pi_{j-1} \dots \pi_{i+1} \pi_i \pi_{j+1} \dots \pi_n \Rightarrow \\
 &\dots \Rightarrow \dots \Rightarrow \dots \Rightarrow \gamma
 \end{aligned}$$

Tasso di riuso

Step 0: π		2		-4	-3		5		-8	-7	-6		1	
Step 1:		2	3	4	5		-8	-7	-6		1			
Step 2:		-5	-4	-3	-2		-8	-7	-6		1			
Step 3:		-5	-4	-3	-2	-1		6	7	8				
Step 4: γ		1	2	3	4	5	6	7	8					

$$\text{BRR} = 2*d / \# b$$

BRR => breackpoint re-use rate

d => distanza del riarrangiamento

b => breackpoint

Random Breackage vs Fragile Breackage

- Simulazione di Sankoff e Trinh
- ST-Synteny
- Difetti di ST-Synteny
- GRIMM-Synteny
- Analisi differenze ST-GRIMM-Synteny
- Simulazione migliorata

Simulazione(n,m,k,w)

1 2 n => Sequenza genomica

Input

- n => numero di elementi
- m => inversioni casuali
- k => micro-inversioni su w elementi

Output

- π => permutazione di n elementi

ST-Synteny(π, w, Δ)

Input

- $\pi \Rightarrow$ permutazione di n elementi
- $w \Rightarrow$ distanza massima del microriarrangiamento
- $\Delta \Rightarrow$ numero minimo di elementi nel blocco di sintenia

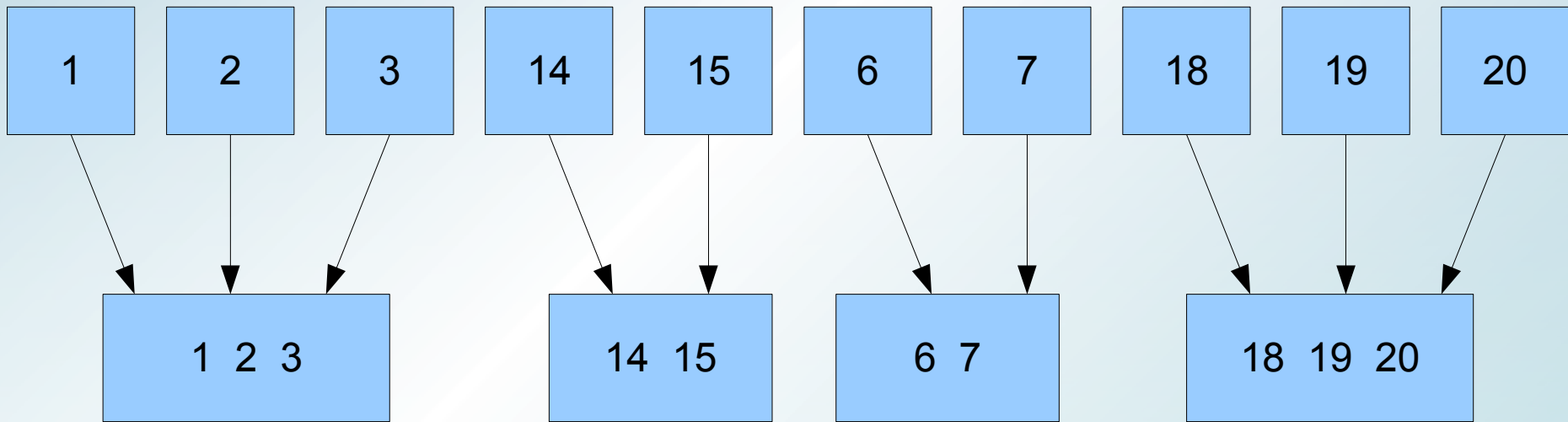
Output

- Synteny blocks

ST-Synteny($\pi, 1, 3$)

$\pi = 1 \ 2 \ 3 \ 14 \ 15 \ 6 \ 7 \ 18 \ 19 \ 20$

Step 1



ST-Synteny($\pi, 1, 3$)

$\pi = 1 \ 2 \ 3 \ 14 \ 15 \ 6 \ 7 \ 18 \ 19 \ 20$

Step 2

1 2 3

~~14 15~~

~~6 7~~

18 19 20

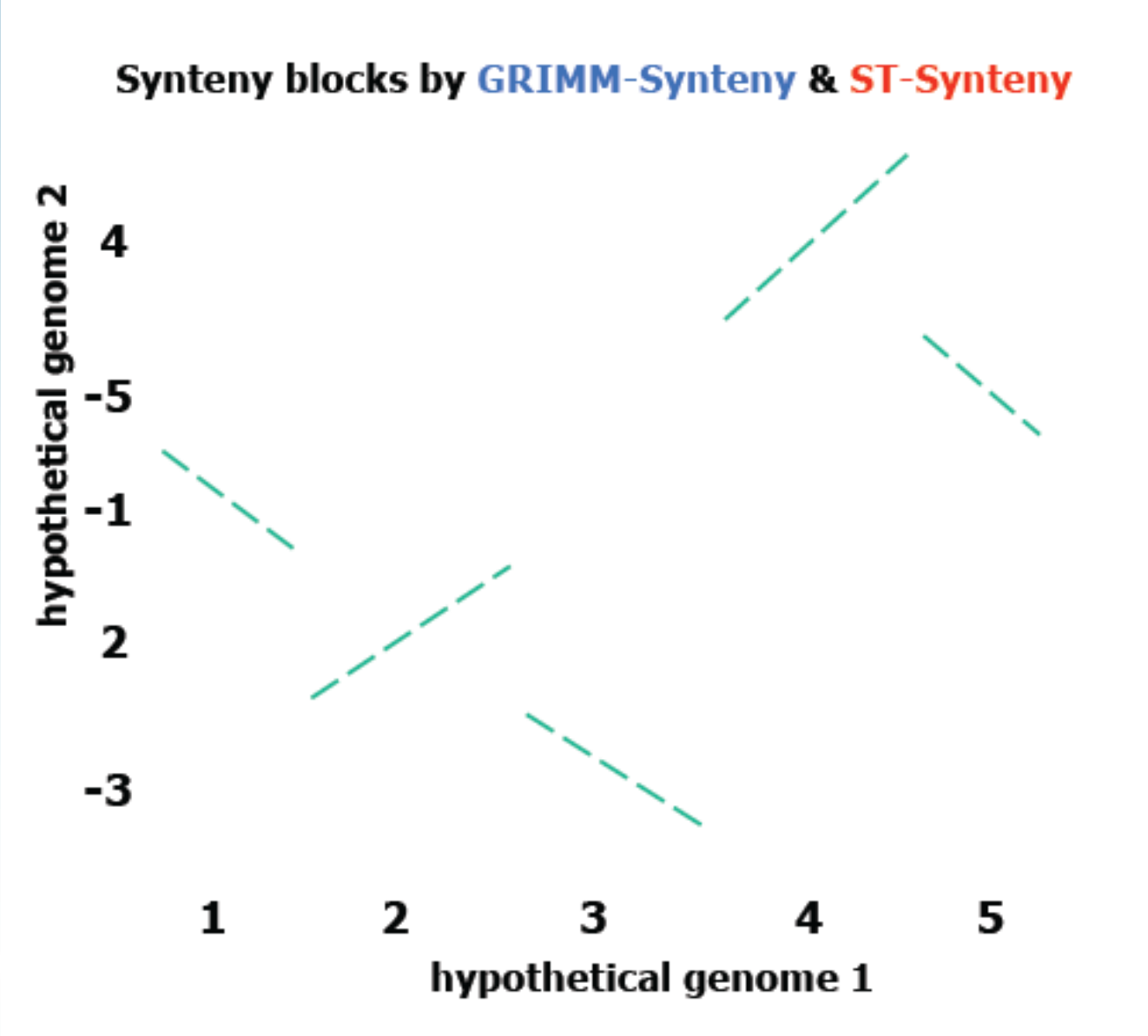
Difetti ST-Synteny

$\pi = 1 \ 2 \ 3 \ 4 \ 5 \Rightarrow$ identity permutation

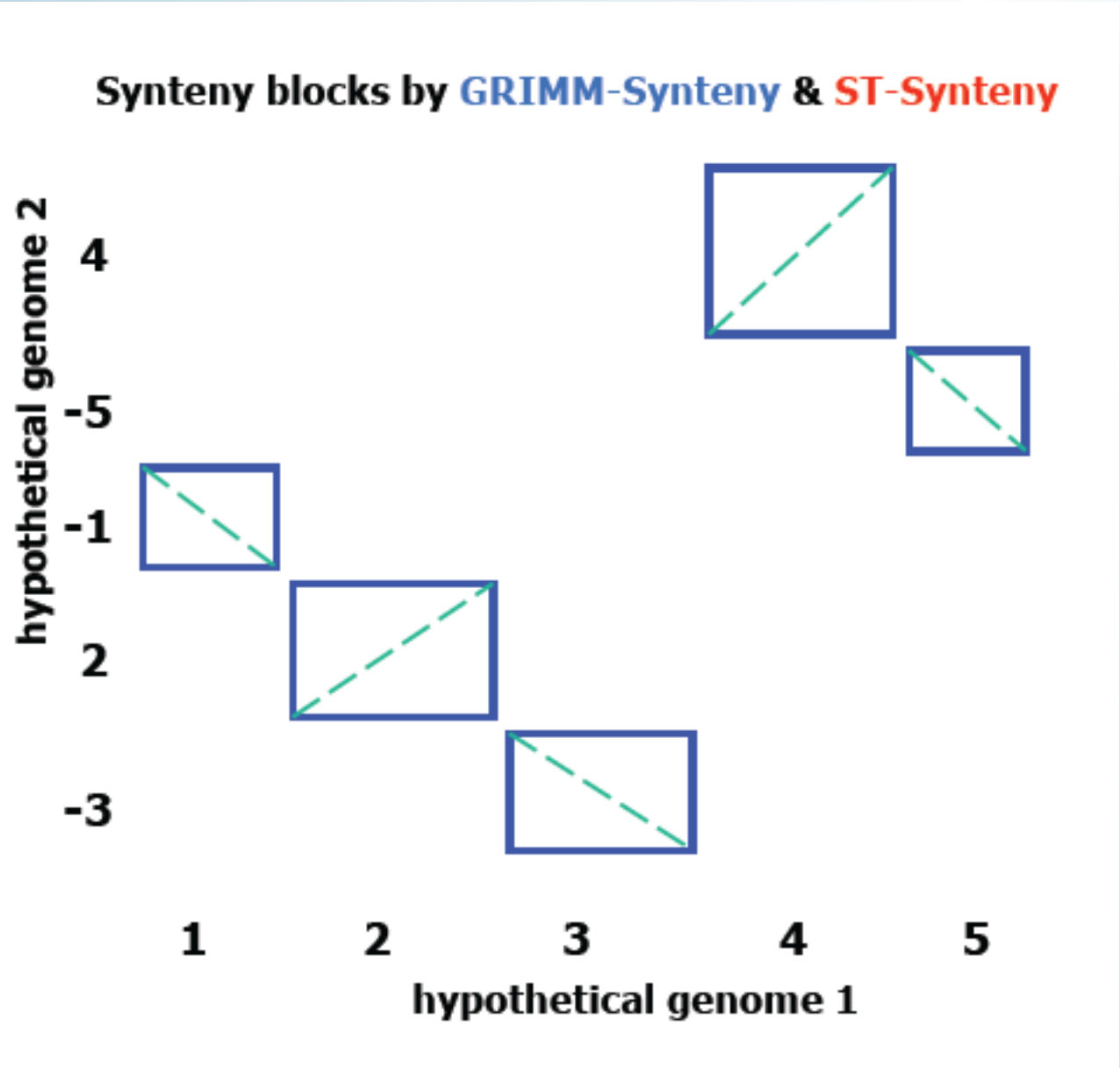


$\pi = -3 \ 2 \ -1 \ -5 \ 4$

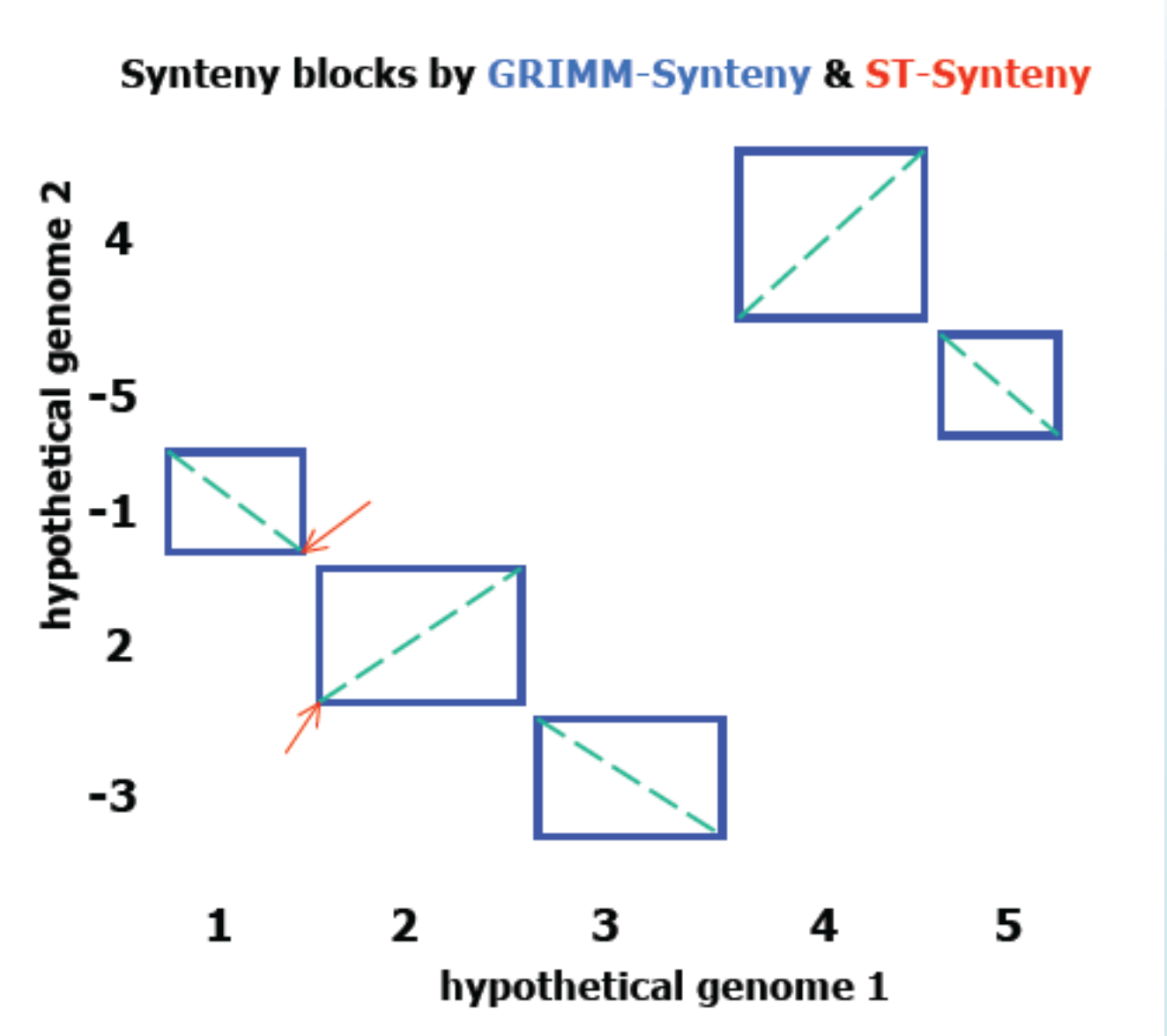
Difetti ST-Synteny



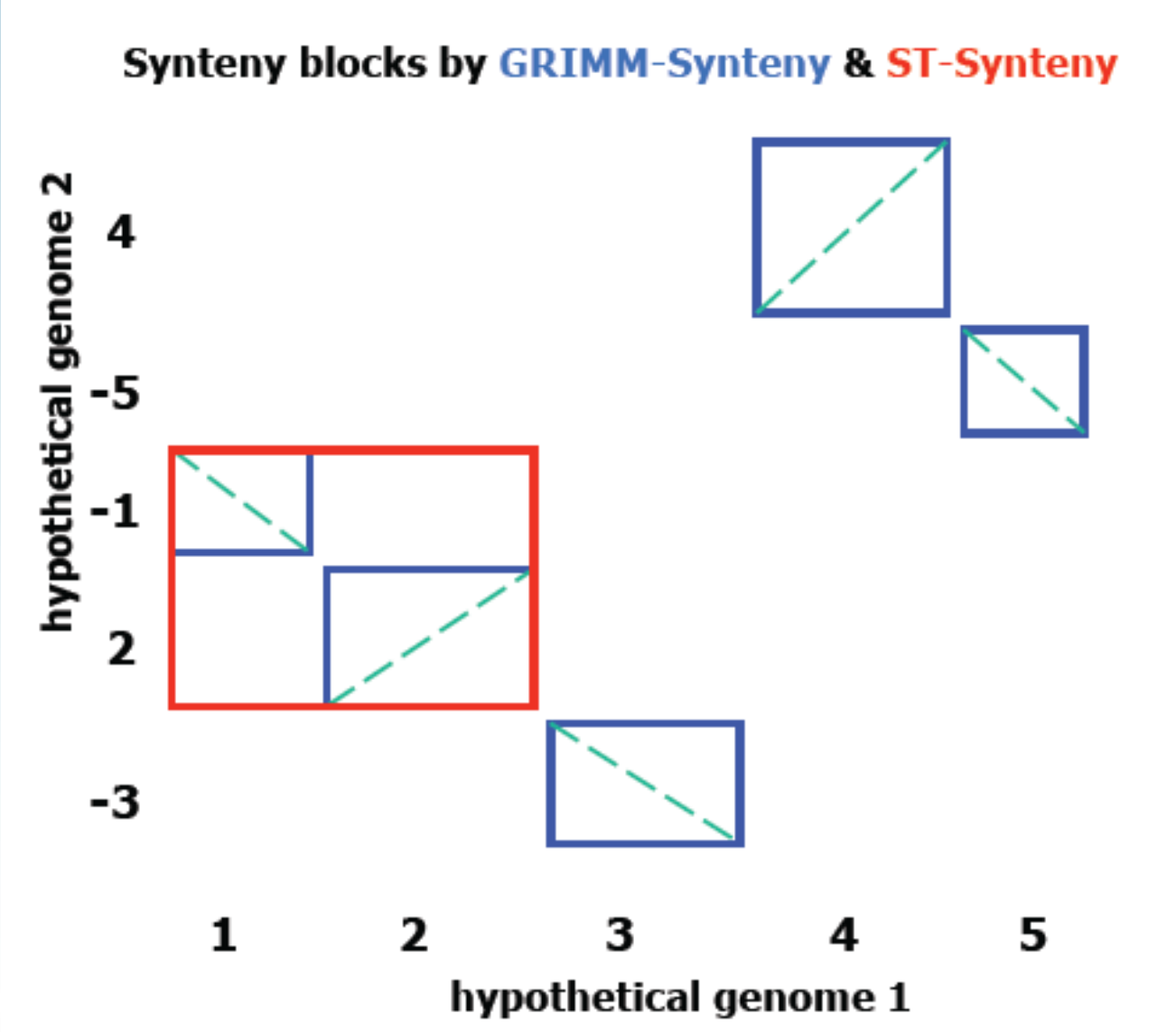
Difetti ST-Synteny



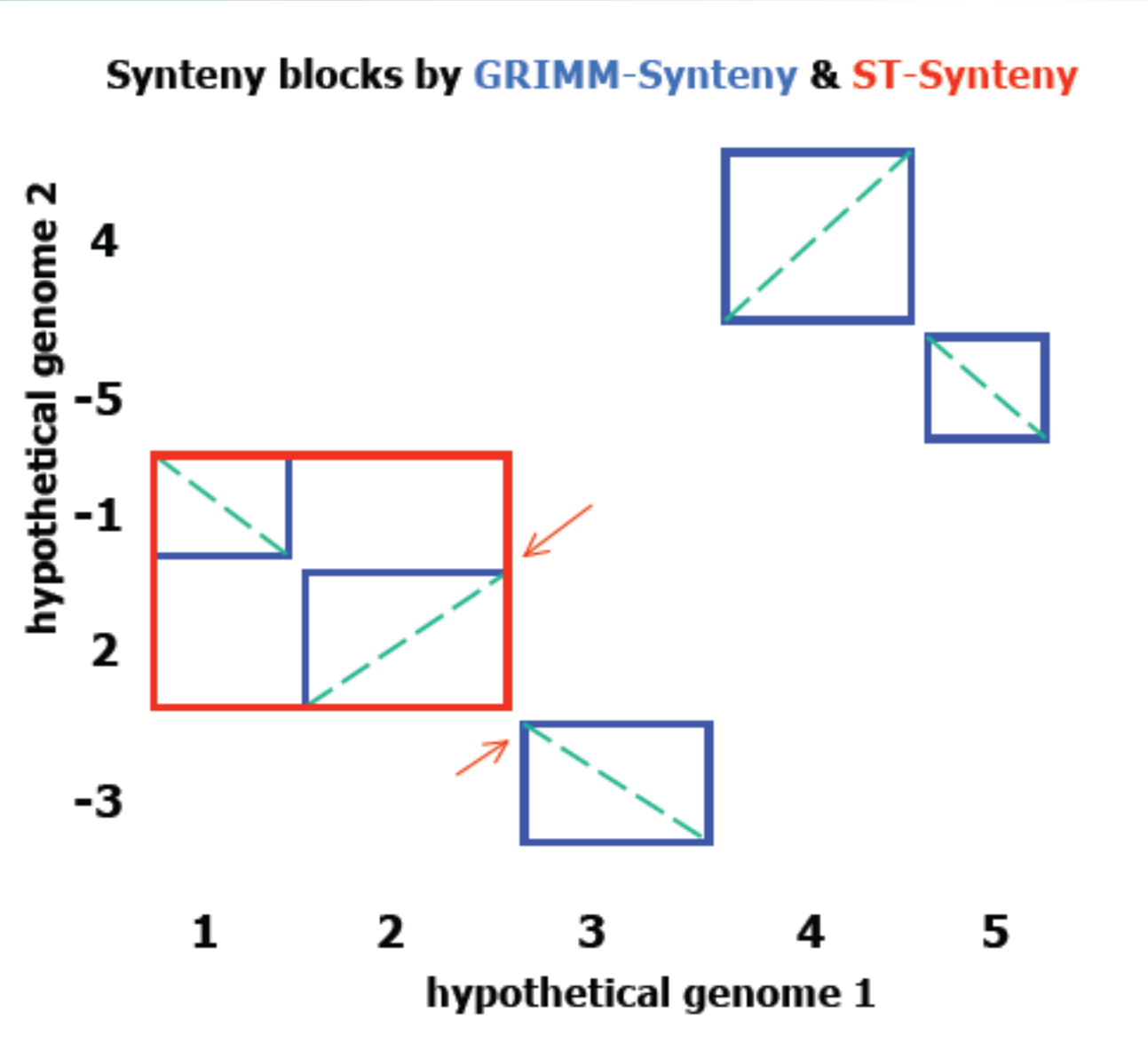
Difetti ST-Synteny



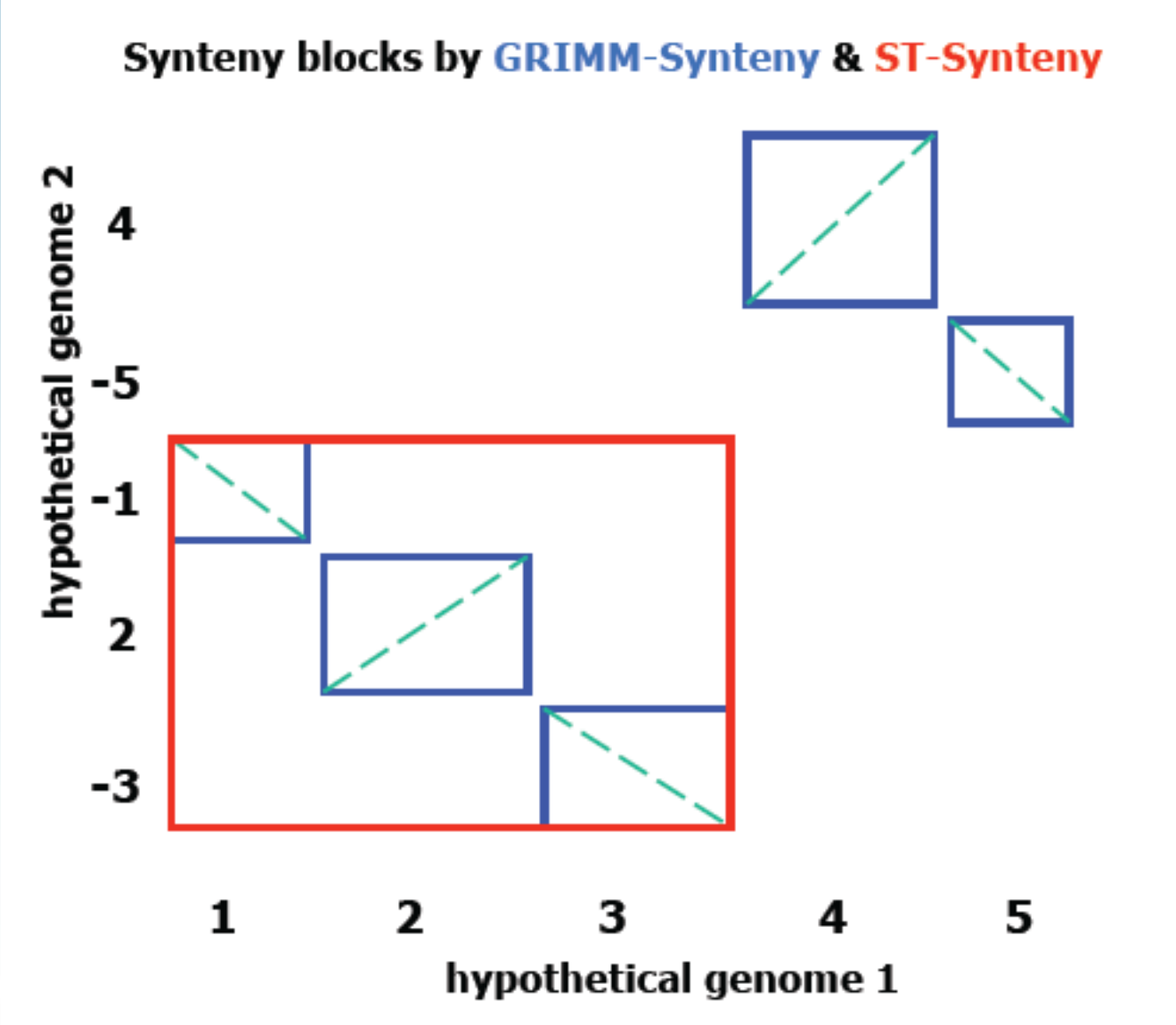
Difetti ST-Synteny



Difetti ST-Synteny

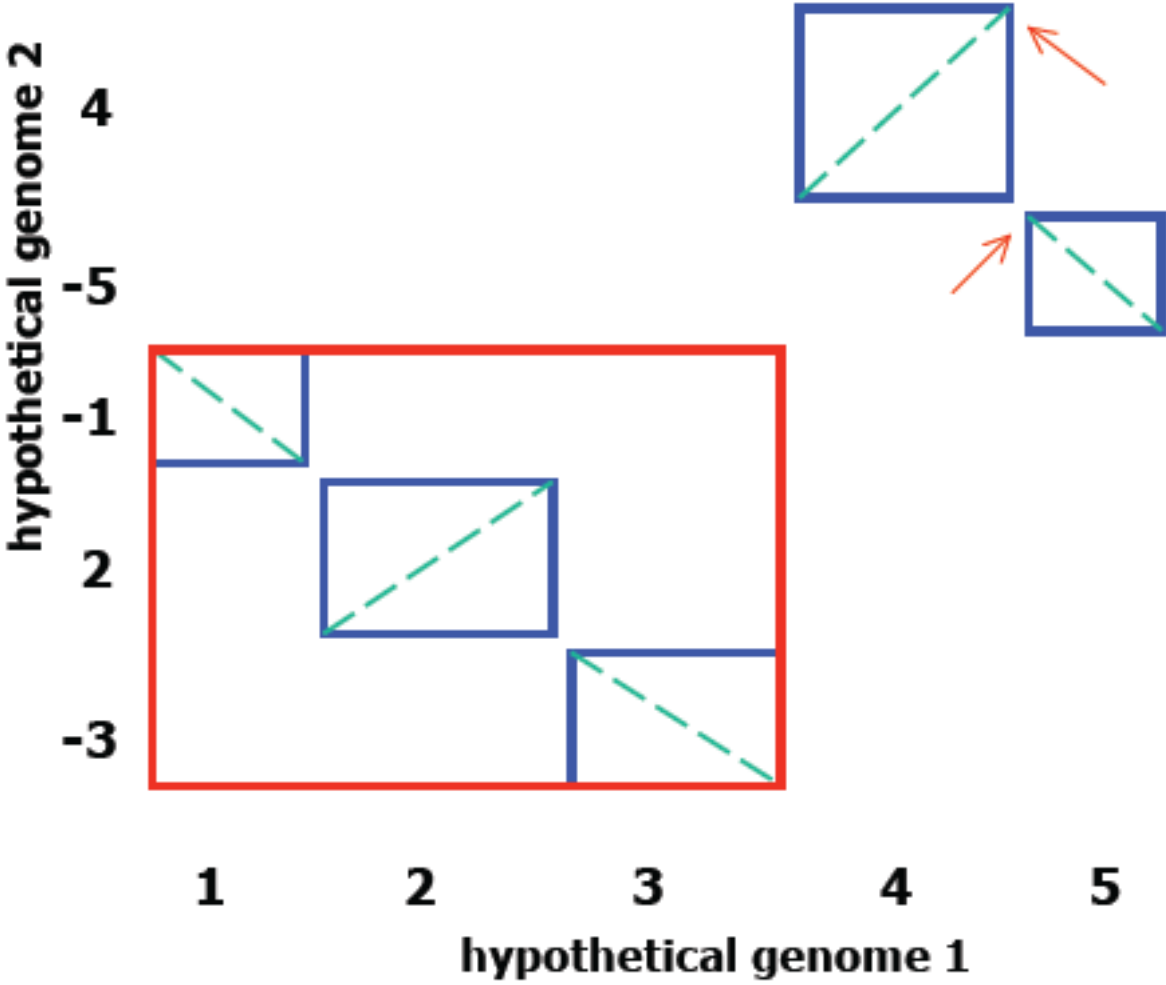


Difetti ST-Synteny

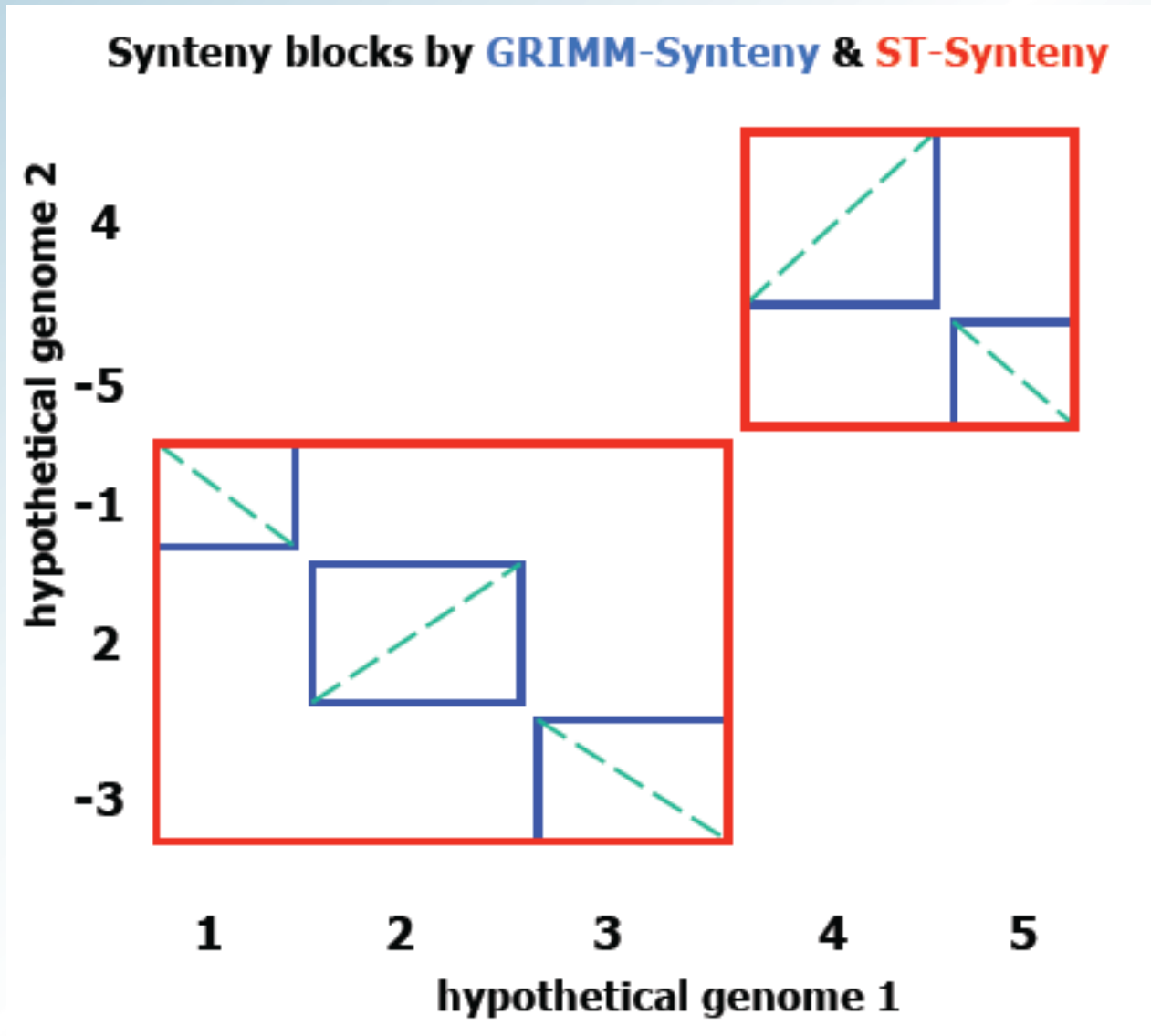


Difetti ST-Synteny

Synteny blocks by GRIMM-Synteny & ST-Synteny

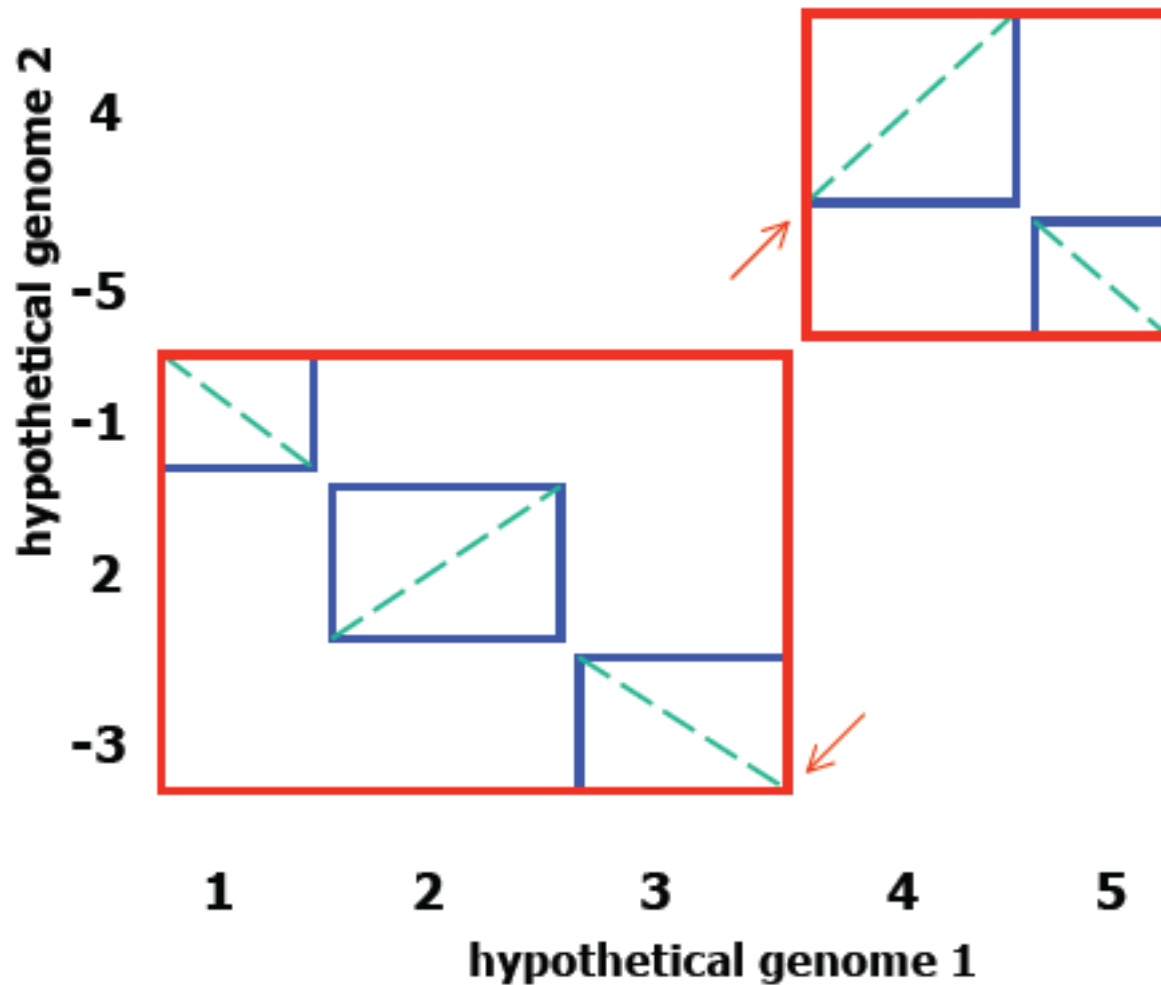


Difetti ST-Synteny



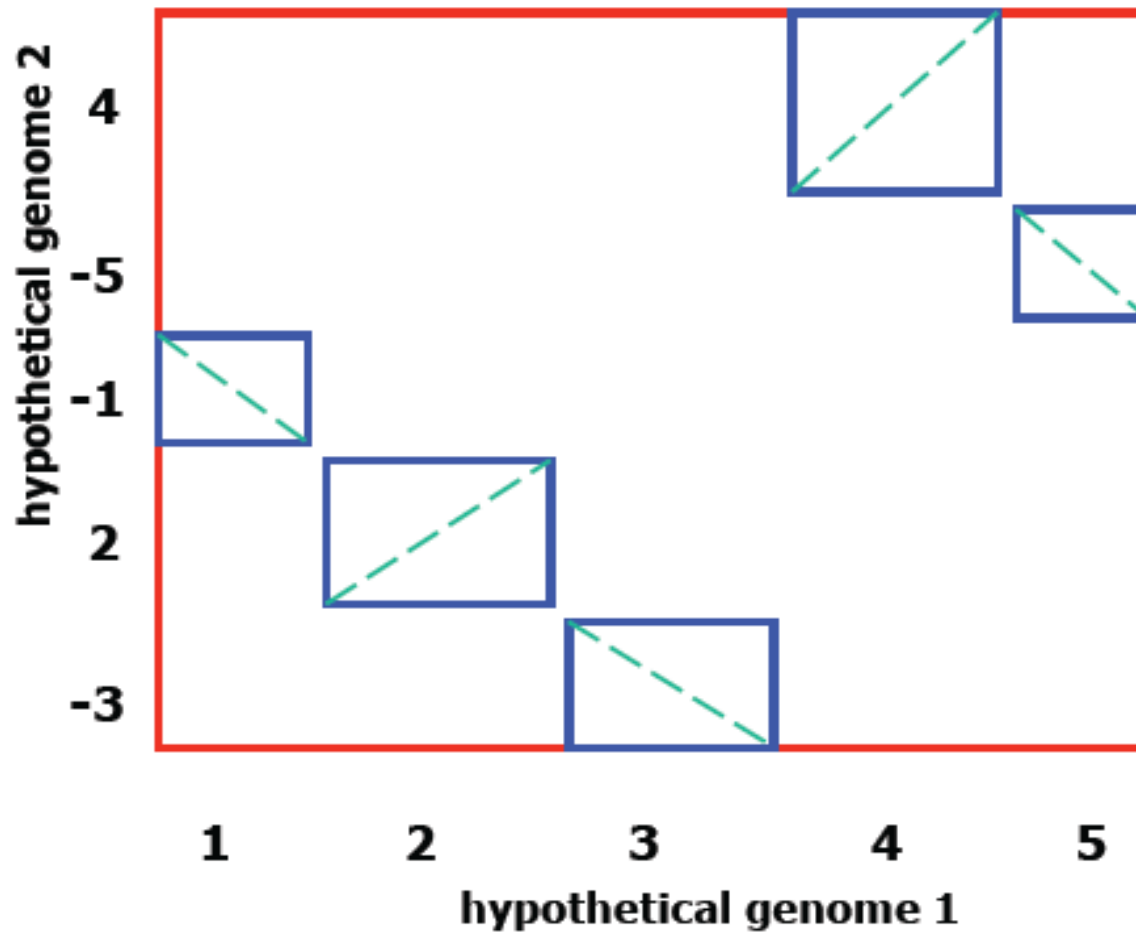
Difetti ST-Synteny

Synteny blocks by GRIMM-Synteny & ST-Synteny



Difetti ST-Synteny

Synteny blocks by GRIMM-Synteny & ST-Synteny



Difetti ST-Synteny

$\pi = \dots 100 \ 101 \ 200 \ 102 \ 103 \ 104 \ 300 \ 105 \ 106 \ 107 \dots$

$w \Rightarrow$ relarivamente piccolo



(100 101) (200) (102 103 104) (300) (105 106 107)

$\Delta = 3$



(102 103 104) (105 106 107)

Difetti ST-Syntenen

1) $\pi = 1\ 3\ 6\ 9\ 2\ 4\ 7\ 10\ 5\ 8$

2) $\pi^{-1} = 1\ 5\ 2\ 6\ 9\ 3\ 7\ 10\ 4\ 8$

$w = 2$

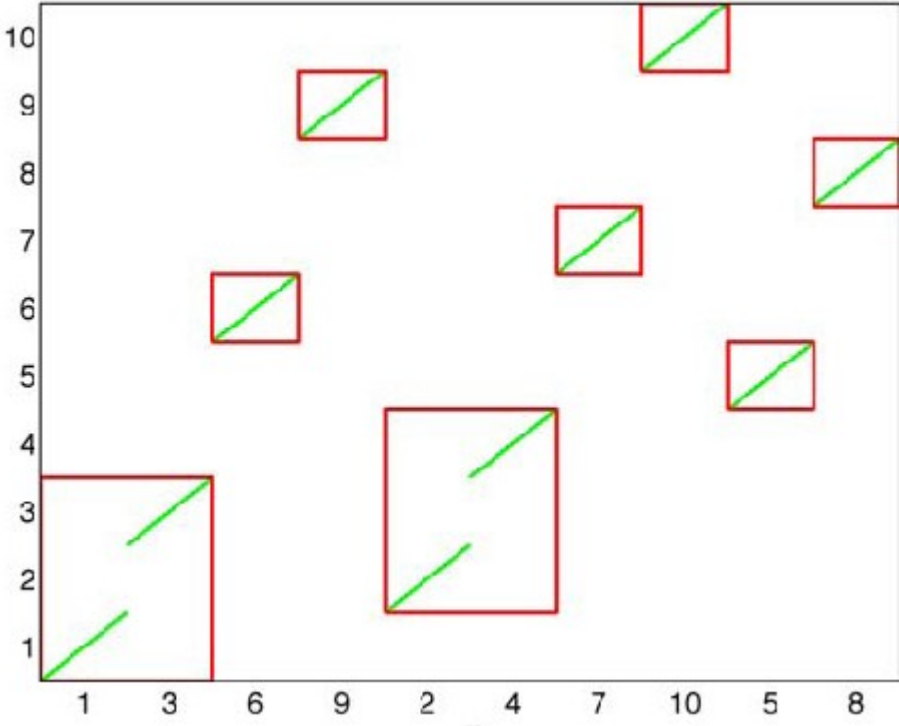


1) $(1\ 3)\ (6)\ (9)\ (2\ 4)\ (7)\ (10)\ (5)\ (8)$

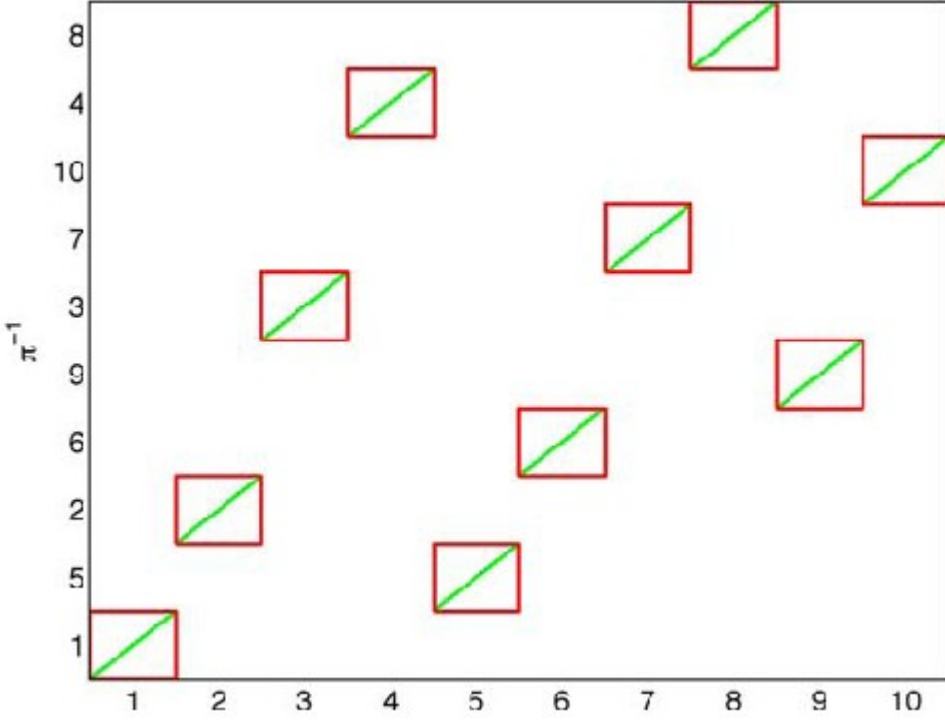
2) $(1)\ (5)\ (2)\ (6)\ (9)\ (3)\ (7)\ (10)\ (4)\ (8)$

Difetti ST-Synteny

A



B



GRIMM-Synteny(π, G, C)

Input

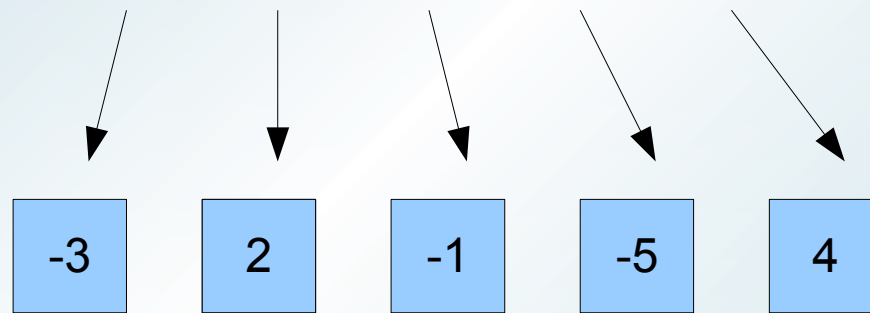
- $\pi \Rightarrow$ permutazione di n elementi
- $G \Rightarrow$ gap massimo
- $C \Rightarrow$ cluster minimo

Output

- Synteny blocks

GRIMM-Synteny(π, G, C)

$\pi = -3 \quad 2 \quad -1 \quad -5 \quad 4$



GRIMM-Synteny(π, G, C)

$\pi = \dots 100 101 200 102 103 104 300 105 106 107 \dots$

$w \Rightarrow$ relativamente piccolo



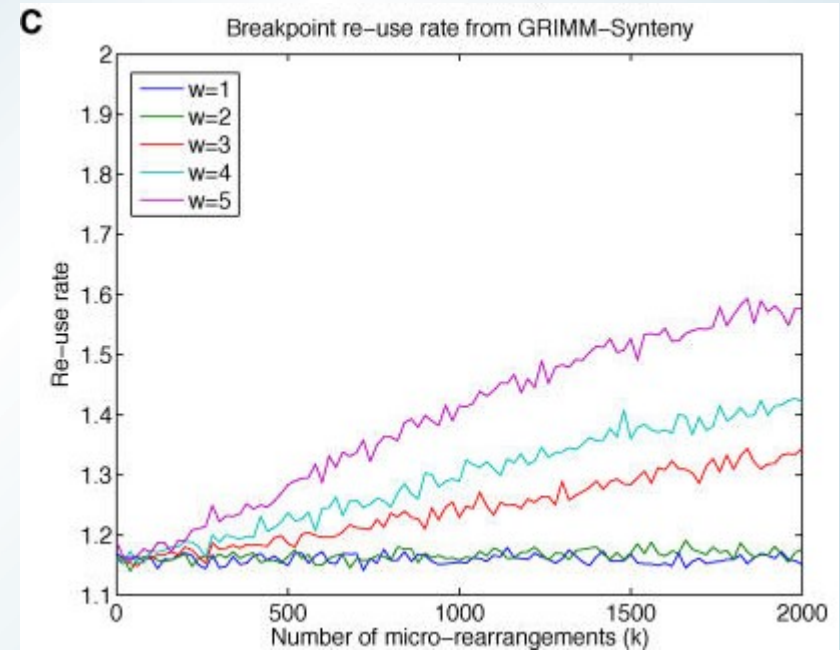
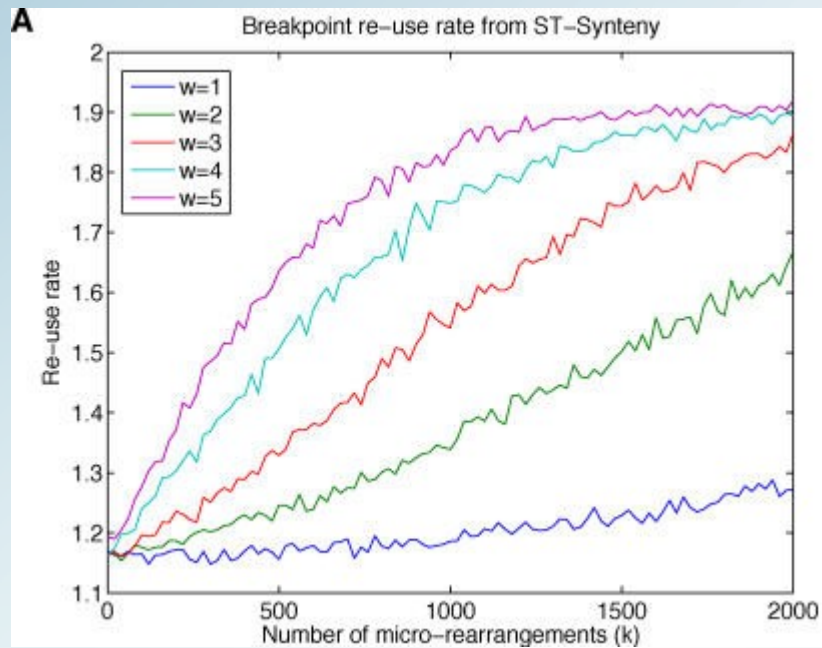
(100 101 102 103 104 105 106 107) (200) (300)

$\Delta = 3$



(102 103 104 105 106 107)

Analisi delle differenze ST-GRIMM-Syntyeny



Analisi delle differenze ST-GRIMM-Synteny

Simulation10 – Simulation15


Lmr = 6Mb - 9Mb >> Lmmr = 196kb


Genoma uomo/topo

Analisi delle differenze ST-GRIMM-Synteny

$\pi = \text{Simulation}(5000, 15, 500, w)$

$w = 1, 2, 3, 4, 5$



$\text{ST-Synteny}(\pi, w, 3) - \text{GRIMM-Synteny}(\pi, w+3, 3)$



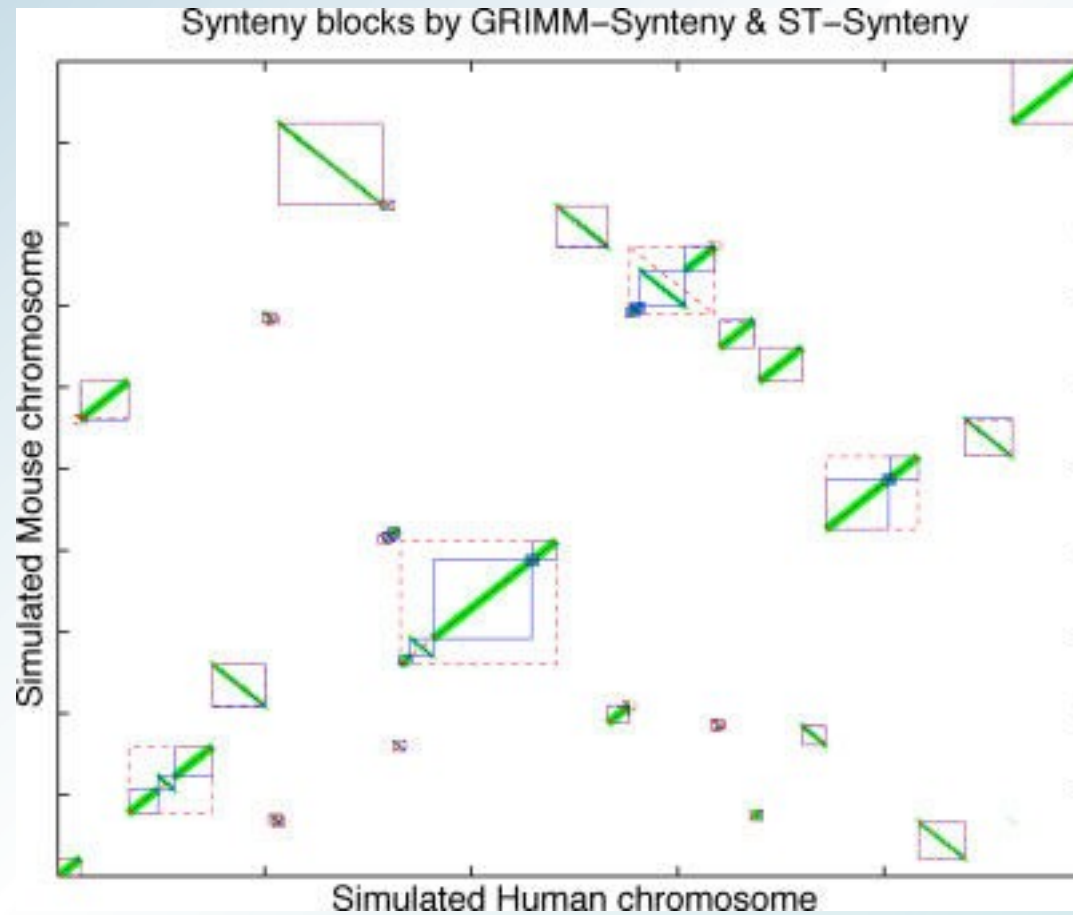
$w = 1, 2, 3, 4, 5$

$\text{Media-BRR}(\text{ST}) = 1,07 \Rightarrow 1,30$



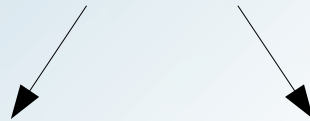
$\text{Media-BRR}(\text{GRIMM}) = 1,03 \Rightarrow 1,09$

Analisi delle differenze ST-GRIMM-Synteny



Analisi delle differenze ST-GRIMM-Synteny

Cromosoma X



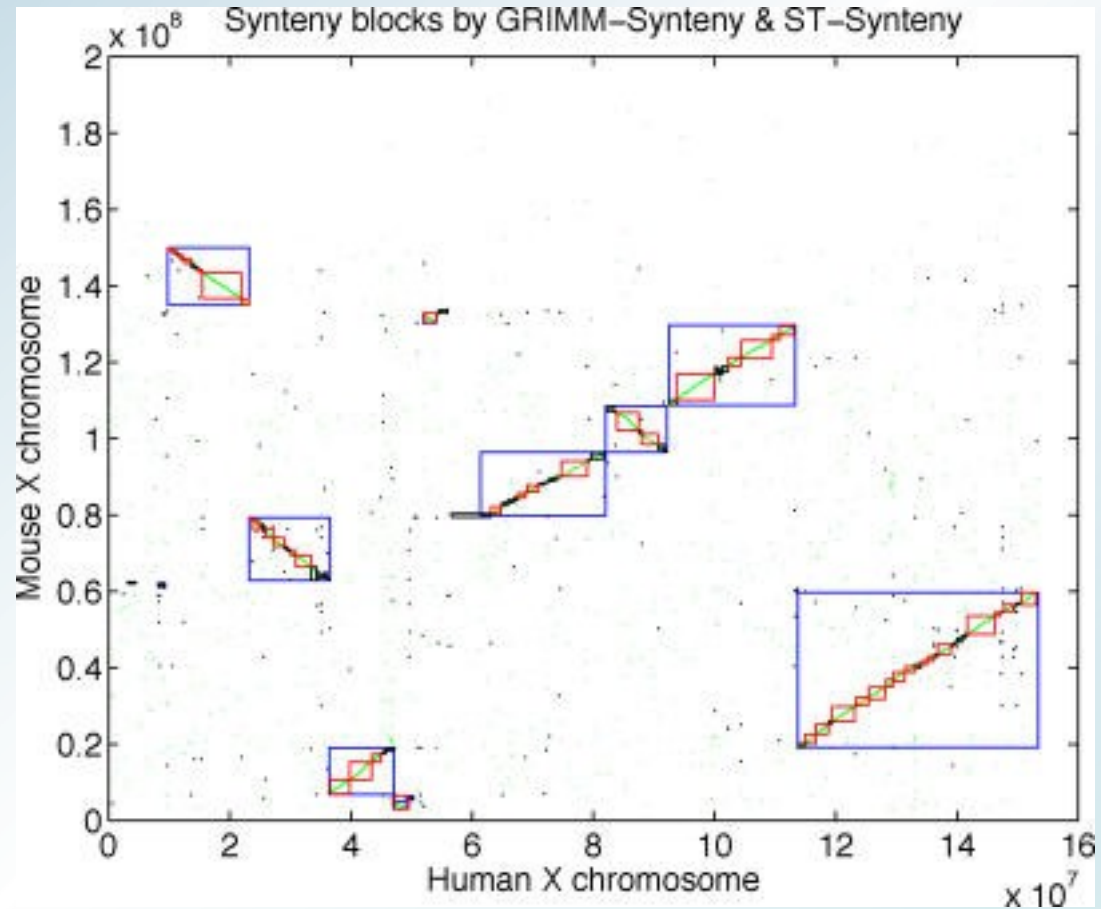
Ancore = 58930

Lunghezza = 153 Mbp

ST-Synteny(π ,378,379)

GRIMM-Synteny(π ,1Mb ,1Mb)

Analisi delle differenze ST-GRIMM-Synteny



Analisi delle differenze ST-GRIMM-Synteny

Characteristic	GRIMM-Synteny	ST-Synteny
Number of blocks	10	44
Total block length (bp)	139,781,782	95,317,543
Breakpoint region (% of chromosome length)	9.05%	37.98%

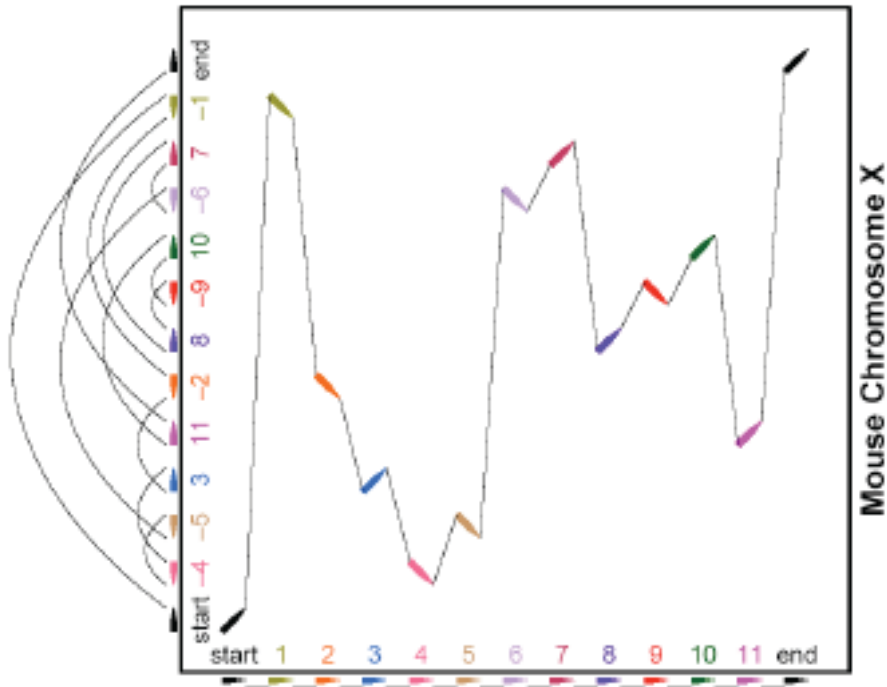
DOI: 10.1371/journal.pcbi.0020014.t001

Simulazione migliorata

Step 1

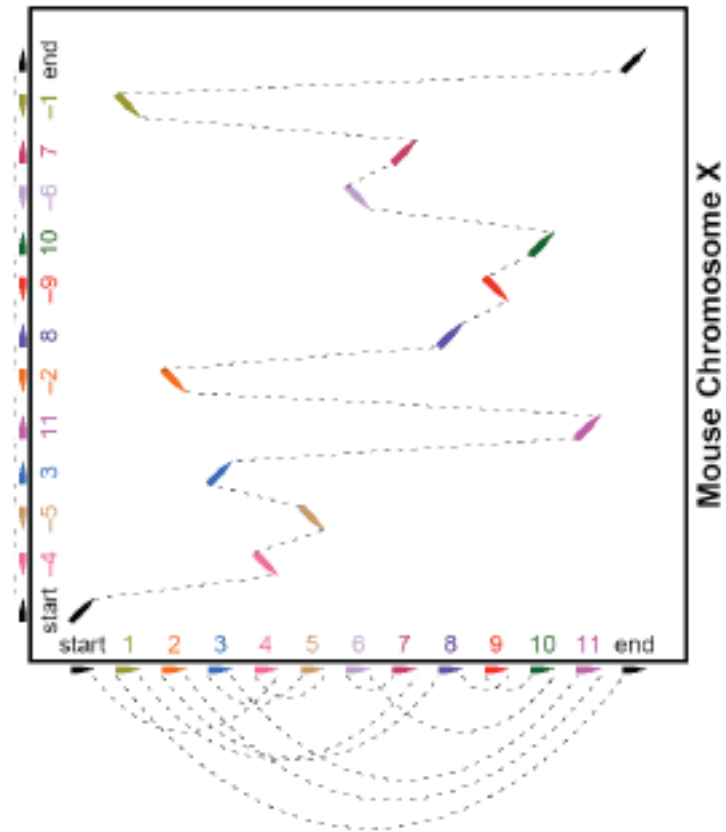
(f) Human path

Human Chromosome X



(g) Mouse path

Human Chromosome X



Simulazione migliorata

Step 2

(1 2 3 4 5 6 7 8)

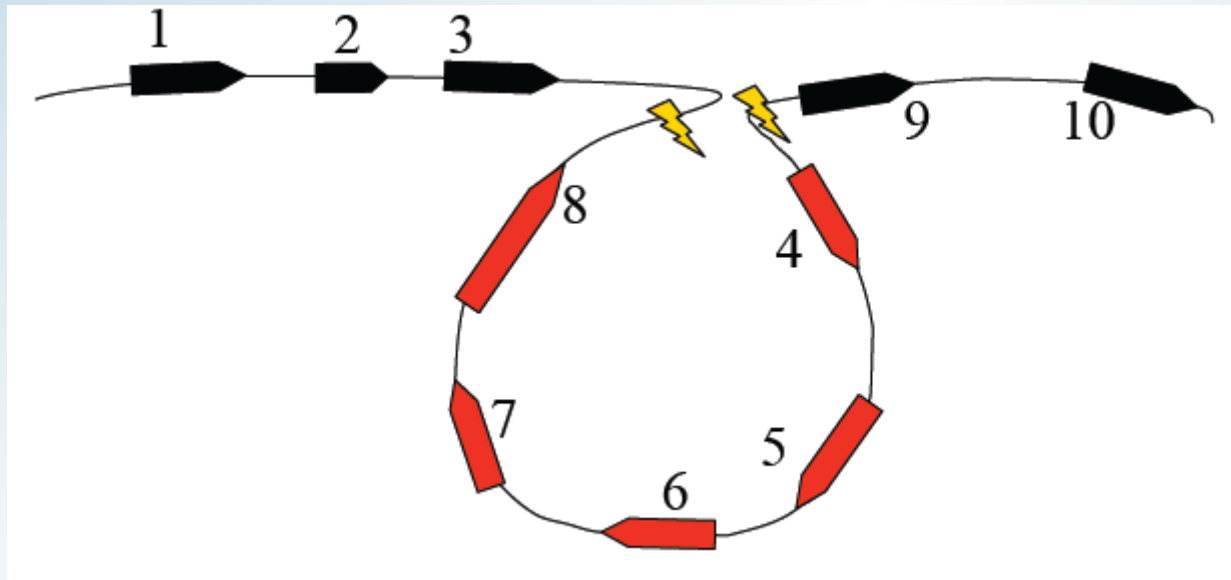
Reversal



(1 2 -5 -4 -3 6 7 8)

Simulazione migliorata

Step 2



Simulazione migliorata

Step 2

(5 9 4 10)
(-6 -1 11 7 -2)

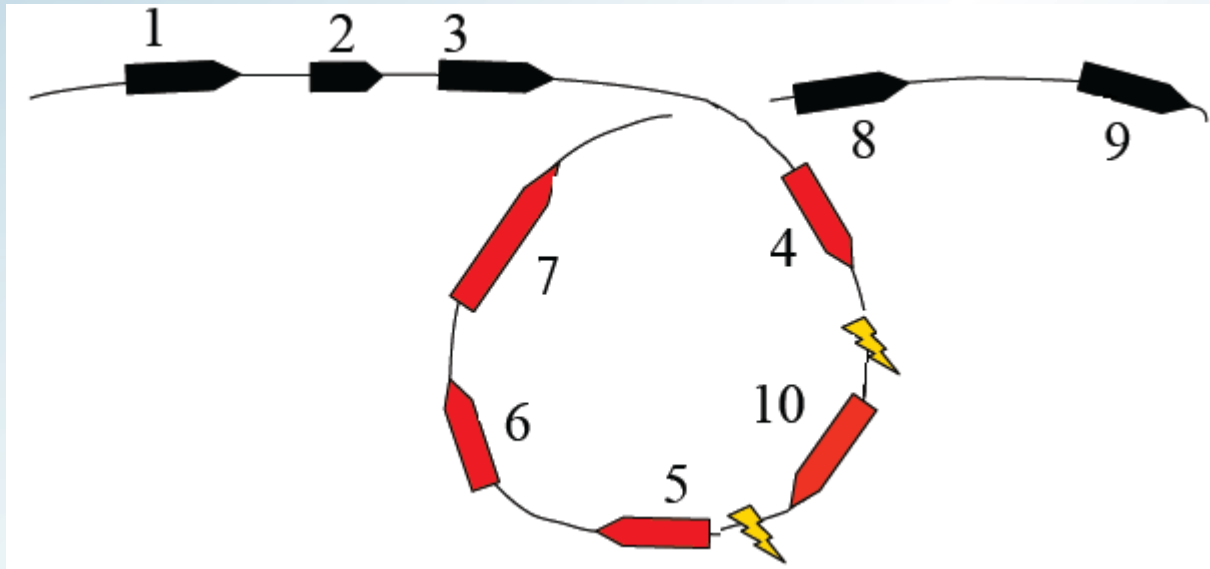
Translocation



(5 9 11 7 -2)
(-6 -1 4 10)

Simulazione migliorata

Step 2



Ancora = 4 10 5 6 7

10 4 5 6 7 4 5 6 7 10

Simulazione migliorata

Step 3

- $k \Rightarrow$ micro-inversioni casuali
- Distanza d'inversione \Rightarrow $\text{random}(1,w)$

Step 4

- 0,2% rumore ancore
- $\text{moveRandom}(\text{ancore})$

Step 5

- $\text{GRIMM-Synteny}(\pi, 1\text{Mb}, 1\text{Mb})$

Simulazione migliorata

k	W = 0.5 Mb				W = 1 Mb			
	Number of Blocks	bk (%)	r _e (%)	Reuse Rate	Number of Blocks	bk (%)	r _e (%)	Reuse Rate
0	9	3.5	0.0	1.25	12	2.9	0.0	1.09
100	9	3.5	55.6	1.25	12	3.5	33.3	1.09
200	9	3.5	55.6	1.25	12	3.6	83.3	1.09
300	9	3.5	66.7	1.25	12	3.5	100.0	1.09
400	9	3.5	88.9	1.25	12	3.3	91.7	1.09
500	9	3.5	88.9	1.25	16	4.3	100.0	1.20
600	9	3.5	100.0	1.25	13	4.1	100.0	1.17
700	9	3.5	100.0	1.25	13	3.6	100.0	1.17
800	9	3.5	100.0	1.25	13	3.3	92.3	1.17
900	9	3.5	100.0	1.25	13	3.8	92.3	1.17
1,000	9	3.5	100.0	1.25	13	4.1	92.3	1.17

Simulazione migliorata

k	W = 0.5 Mb				W = 1 Mb			
	Number of Blocks	bk (%)	r _e (%)	Reuse Rate	Number of Blocks	bk (%)	r _e (%)	Reuse Rate
0	282	6.7	0.0	1.12	286	6.9	0.0	1.12
1,000	282	6.7	18.1	1.12	286	7.0	28.0	1.12
2,000	282	6.7	30.1	1.12	284	7.1	44.4	1.12
3,000	282	6.7	37.2	1.12	283	7.4	60.8	1.12
4,000	282	6.7	44.7	1.12	284	7.6	70.1	1.13
5,000	282	6.7	52.8	1.12	285	7.7	75.8	1.13
6,000	282	6.7	58.9	1.12	283	7.9	78.8	1.14
7,000	282	6.7	66.0	1.12	283	7.6	82.0	1.14
8,000	282	6.7	72.7	1.12	284	7.7	88.7	1.14
9,000	282	6.7	75.2	1.12	290	7.6	91.4	1.14
10,000	282	6.7	76.6	1.12	293	7.6	93.2	1.14
11,000	282	6.7	79.8	1.12	292	7.6	93.2	1.14
12,000	282	6.7	81.2	1.12	293	7.5	95.2	1.14
13,000	282	6.8	80.9	1.12	297	7.5	95.6	1.14
14,000	282	6.8	83.3	1.12	297	7.5	95.6	1.16
15,000	282	6.8	85.8	1.12	298	7.4	96.6	1.17

Conclusioni

Genoma completo uomo/topo

- 294 => synteny blocks
- 10900 => micro-riarrangiamenti
- 196 kb => distanza media d'inversione
- 39,1% => r_e
- 9,06% => bk
- 1,67 => BRR