

Conservation of Combinatorial Structures in Evolution Scenarios

S everine B erard (INRA, Toulouse),

Anne Bergeron et Cedric Chauve (UQAM, Montr al)

Belle marquise, vos beaux yeux me font mourir d'amour.

Vos yeux beaux d'amour me font, belle marquise, mourir.

Me font vos beaux yeux mourir, belle marquise, d'amour.

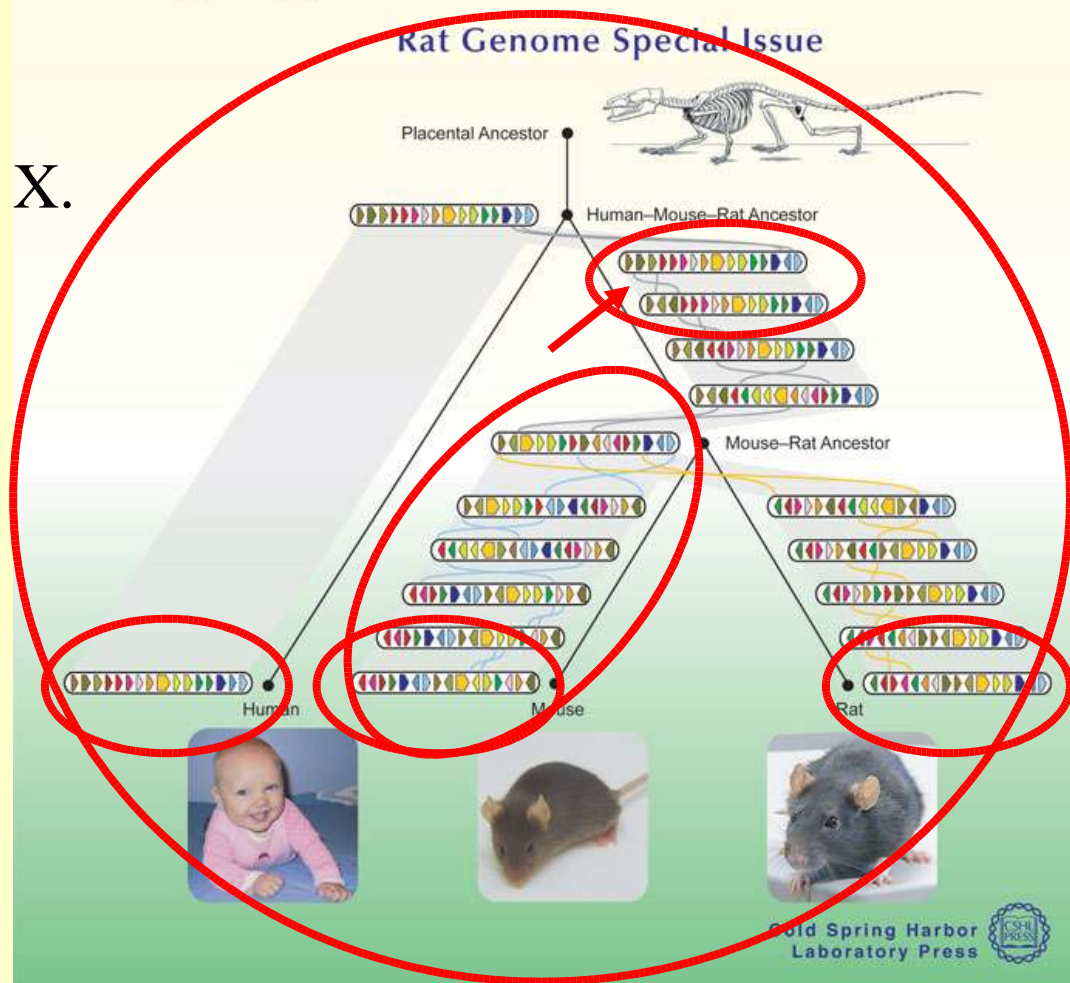
What is an “Evolution Scenario” ?

ISSN 1088-9051

GENOME RESEARCH

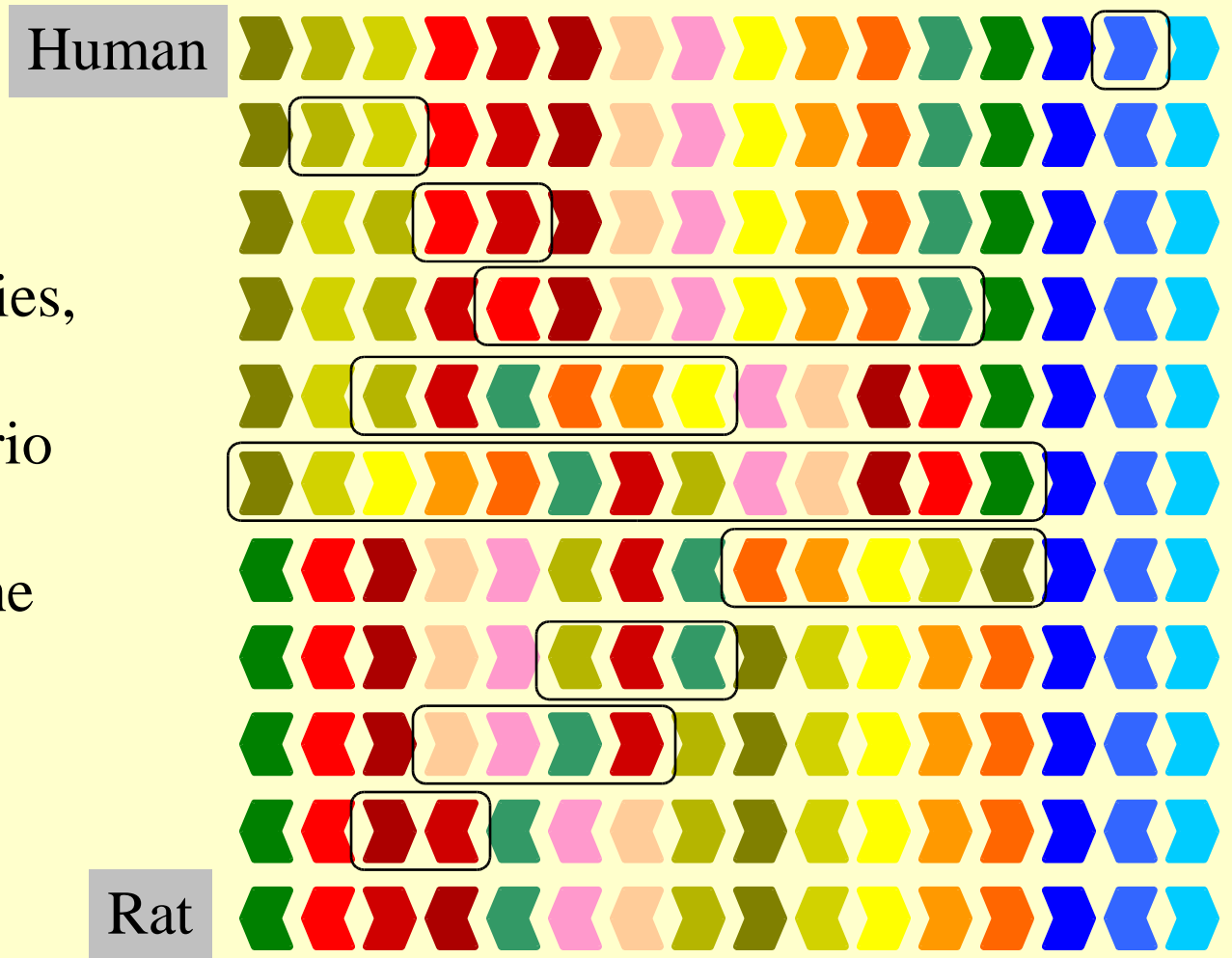
April 2004
Volume 14 Number 4

1. An evolution tree.
2. Genomes of existing species labeling the leaves of the tree, here blocks from the Mouse, Rat and Human chromosomes X.
3. Sequences of genomes on each branch of the tree.
4. Each successive genome differs from the next one by one inversion.



(Art work by Guillaume Bourque, scientific work by Guillaume Bourque, Pavel Pevzner and Glenn Tesler, 2004)

What is an “Evolution Scenario” ?



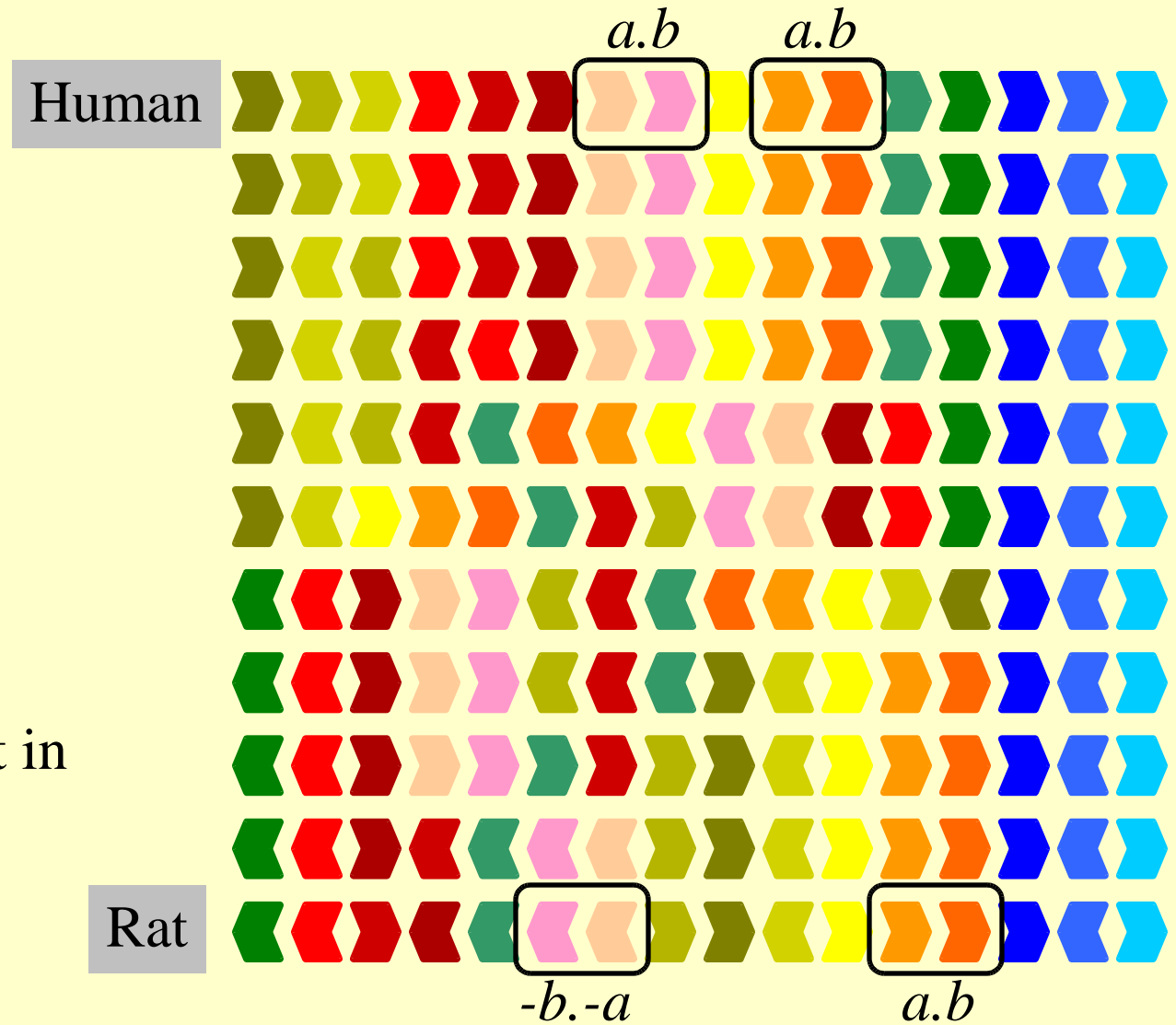
For each pair of species,
the tree induces a
rearrangement scenario
between the two
chromosomes X of the
two species.

What is a “Conserved Combinatorial Structure” ?

1. Adjacencies

Let a and b be two consecutive genes (or blocks) in one genome.

The adjacency $a.b$ is *conserved* if either $a.b$ or $-b.-a$ is present in the other genome.

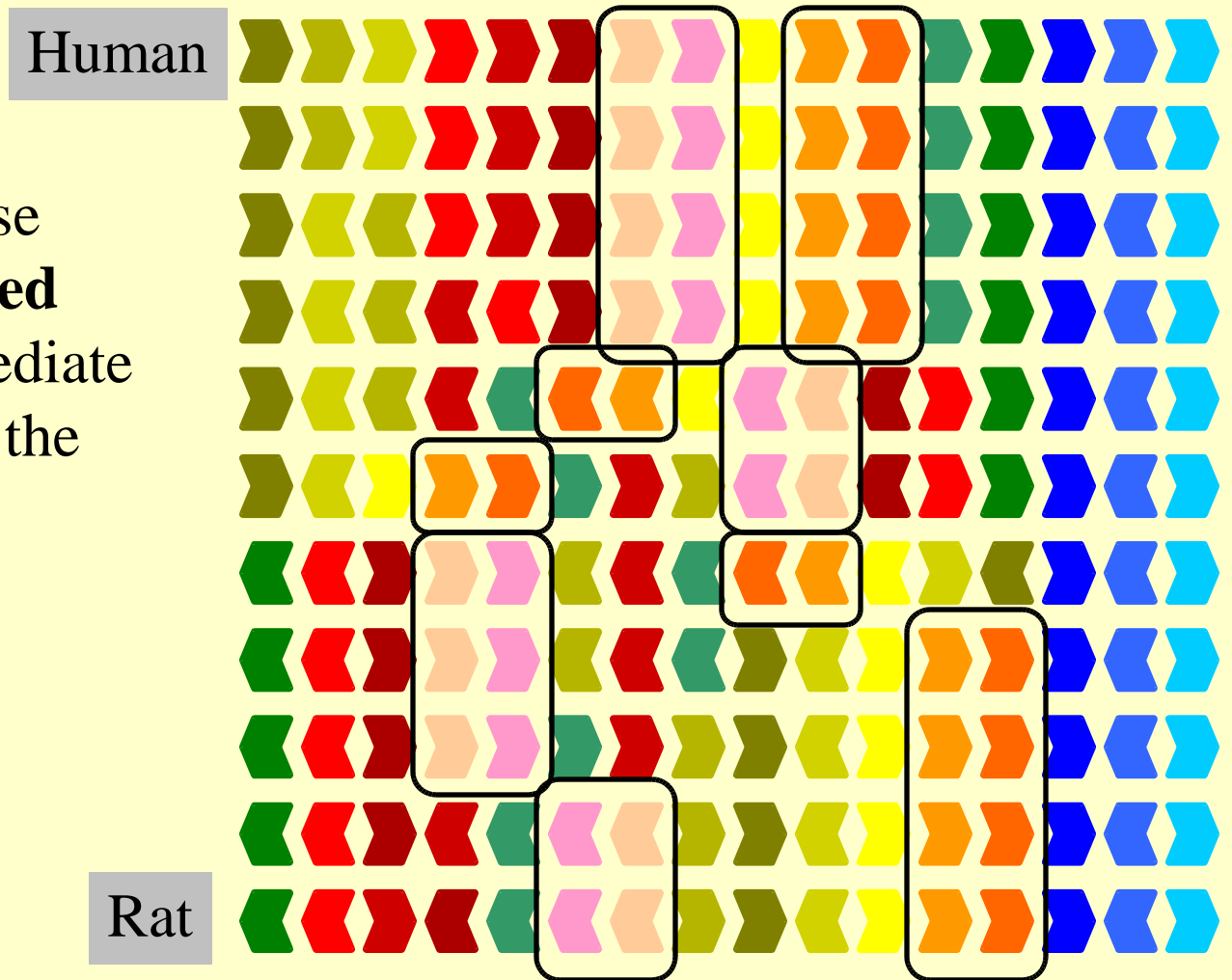


(Blanchette, Kunisawa, Sankoff, 1999)

What is a “Conserved Combinatorial Structure” ?

1. Adjacencies

Note that each of these adjacency is **conserved** in each of the intermediate genome proposed by the scenario.

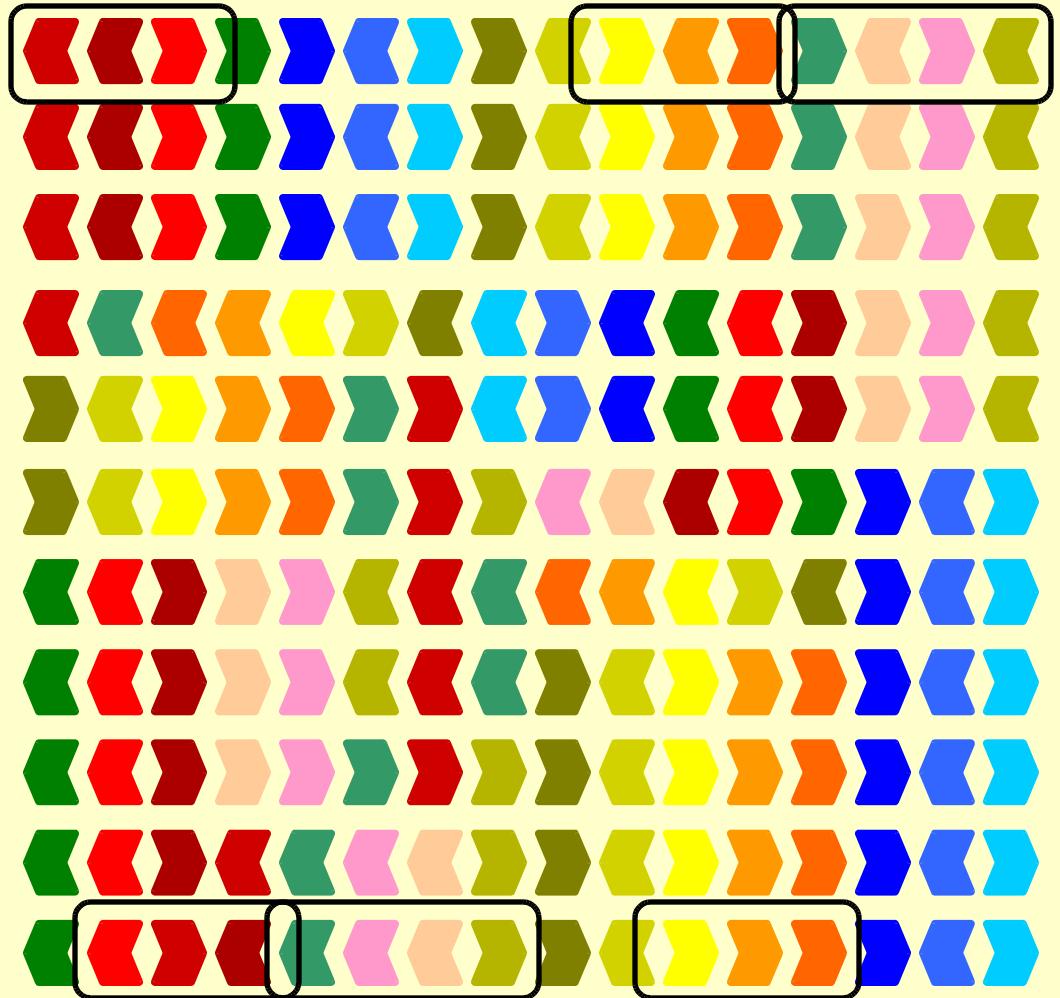


(Blanchette, Kunisawa, Sankoff, 1999)

What is a “Conserved Combinatorial Structure” ?

2. Common intervals

Mouse



Rat

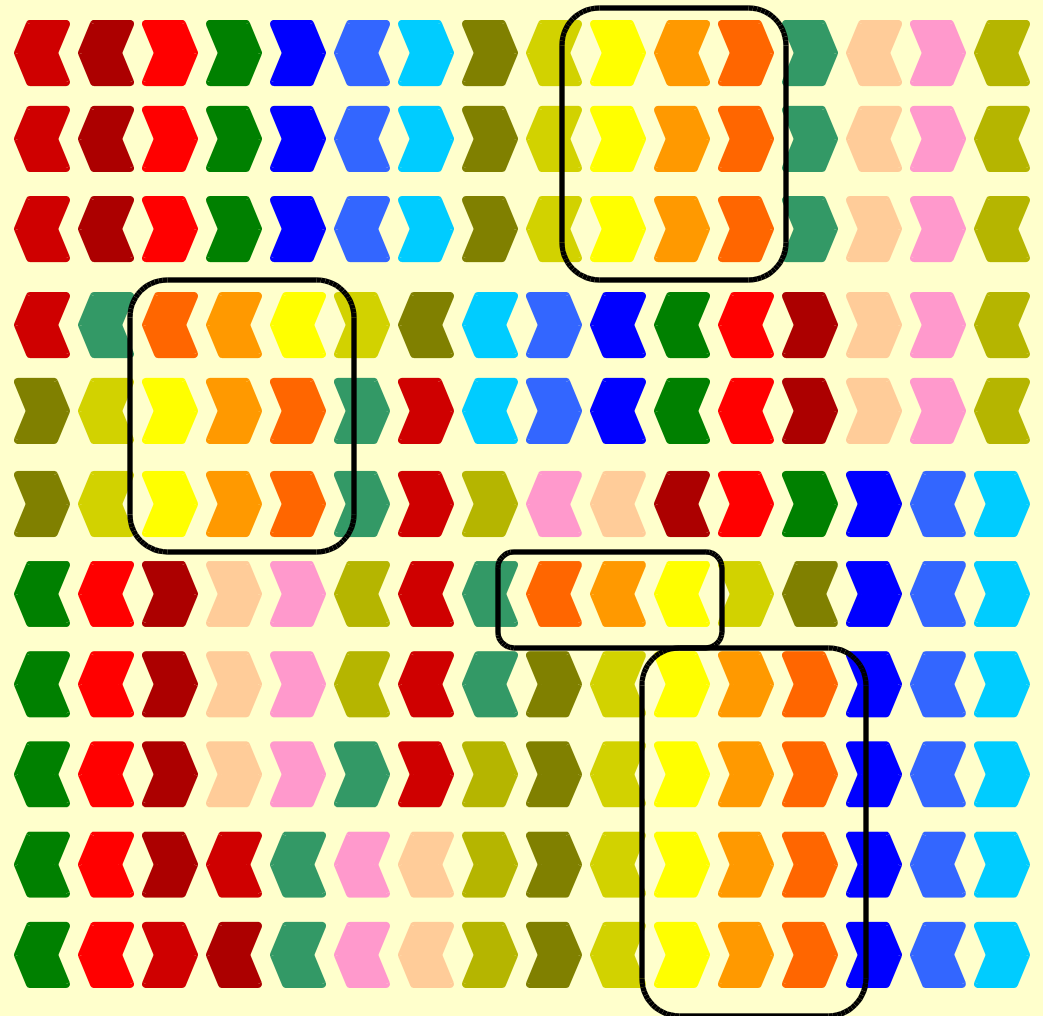
A common interval is a set of genes that appear consecutively in each genome, but not necessarily in the same order, or orientation.

What is a “Conserved Combinatorial Structure” ?

2. Common intervals

Mouse

Some common intervals are **conserved** in the intermediate genomes of proposed scenario.



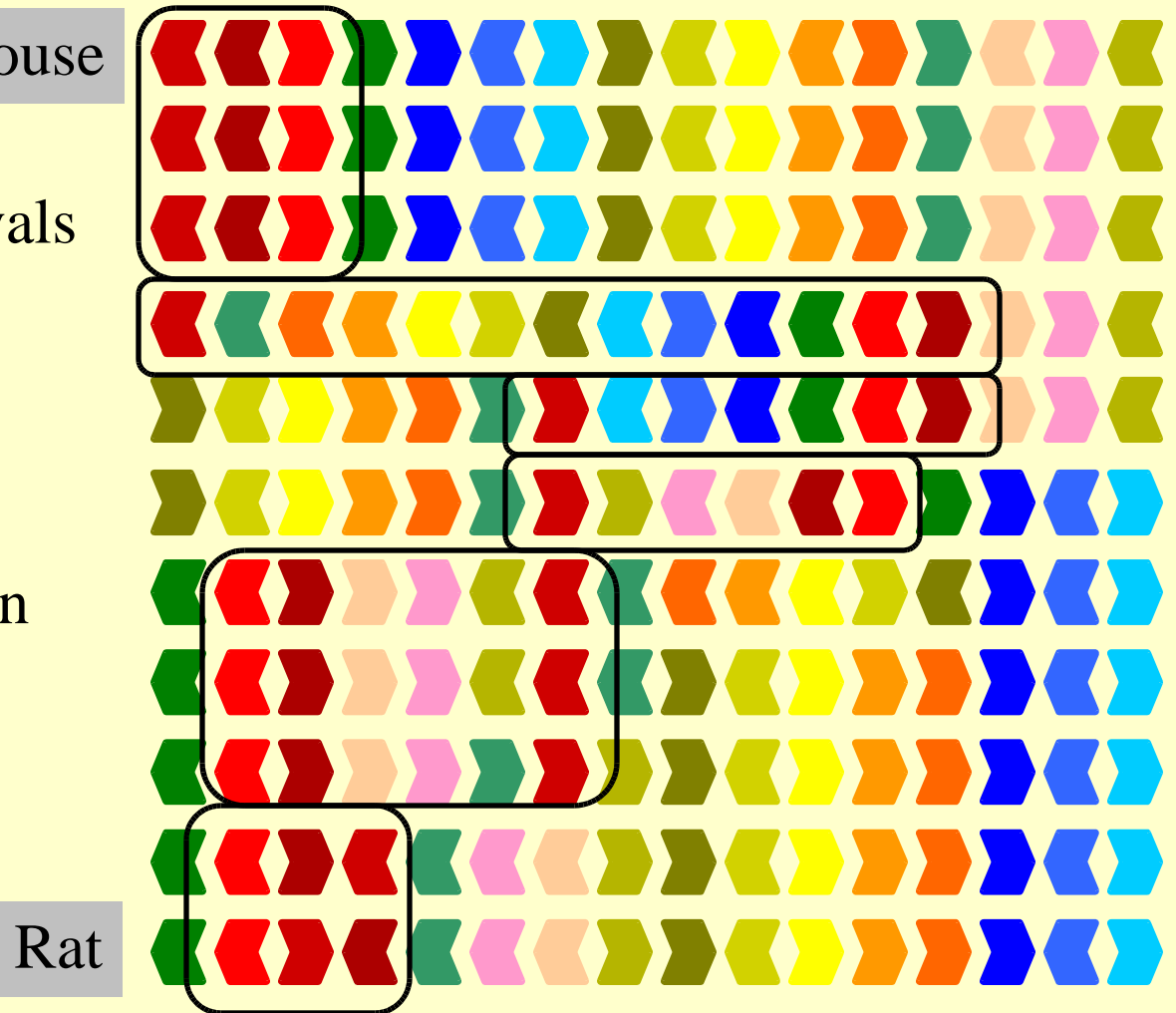
Rat

What is a “Conserved Combinatorial Structure” ?

2. Common intervals Mouse

But some common intervals are **not conserved** in the intermediate genomes of proposed scenario.

They are broken, and then reassembled.



Rat

(Uno and Yagiura, 2000;
Heber and Stoye, 2001)

Perfect Scenarios

Definition:

A *perfect* scenario is a rearrangement scenario in which no inversion breaks a common interval.

Questions:

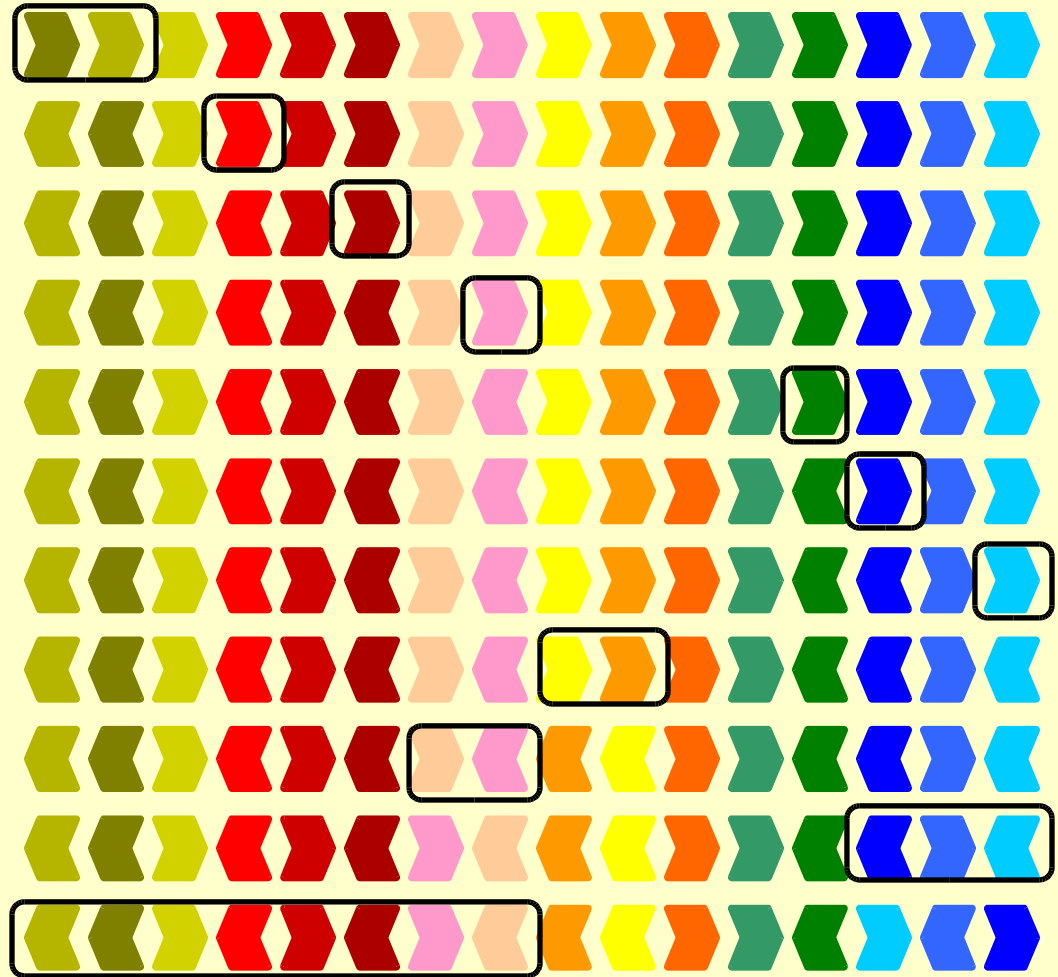
1. Do perfect scenarios exist in nature? **Yes.**
2. Should we question scenarios that are not perfect? **Yes.**
3. Is it easy to construct perfect scenarios? **Yes and no.**

Do perfect scenarios exist in nature?

This scenario transforms 16 blocks of a region of Human chromosome 17 into a region of Mouse chromosome 11.

It conserves **all common intervals** between Human and Mouse.

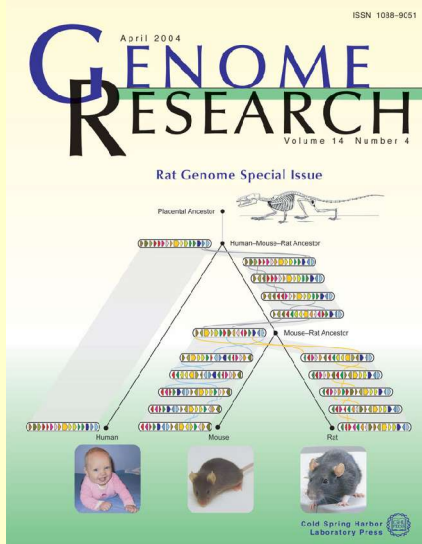
Human



Mouse

Should we question scenarios that are not perfect?

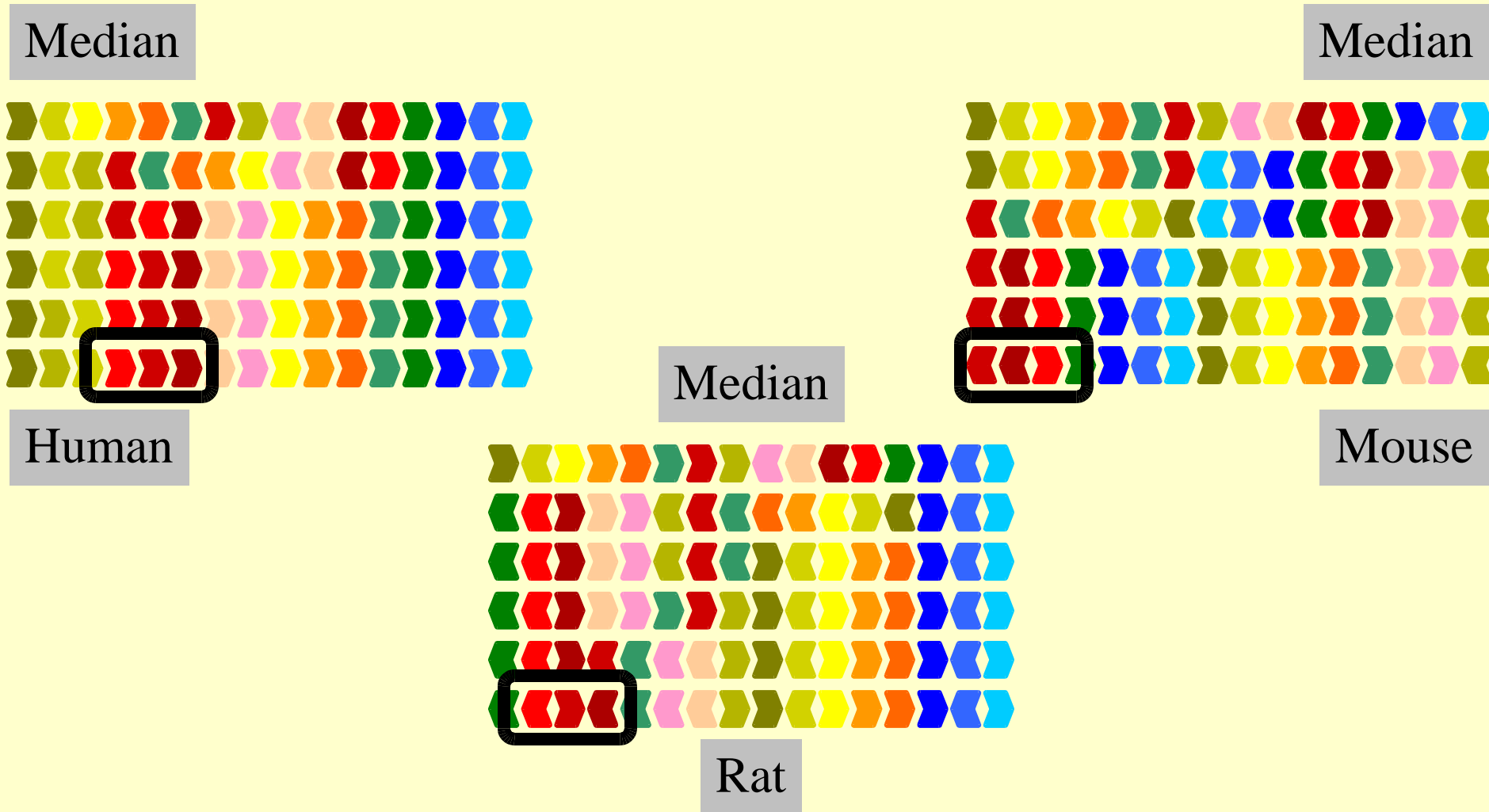
Let's go back to the scenarios induced by the evolution scenario comparing the Mouse, Rat and Human chromosomes X.



The three induced scenarios, from Human to Mouse, from Human to Rat, and from Mouse to Rat, have unusual features.

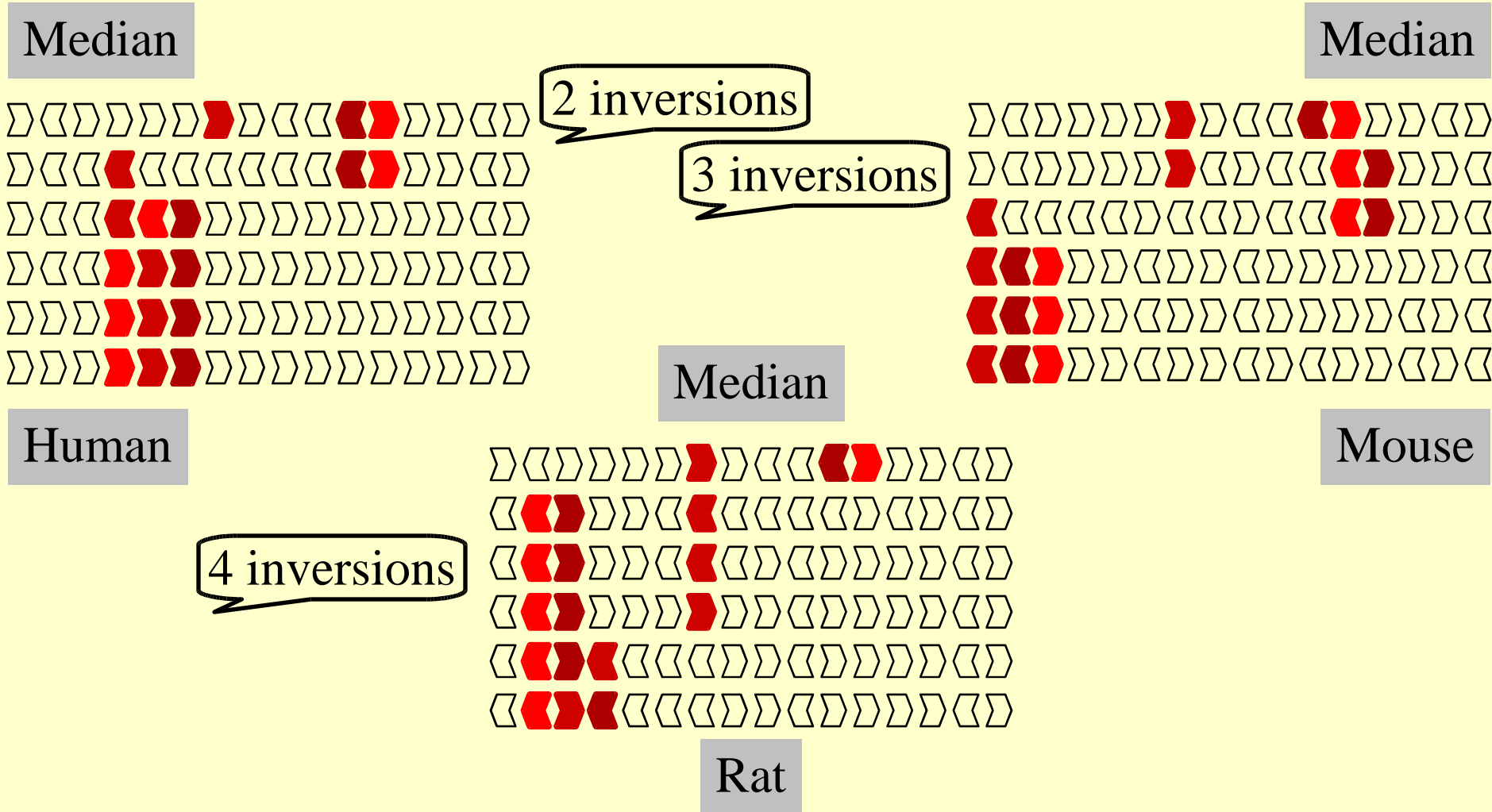
Should we question scenarios that are not perfect?

All three species share a common interval, that is not a common interval of their median.



Should we question scenarios that are not perfect?

This implies three independent reconstructions of the common interval!



Should we question scenarios that are not perfect?

What might be wrong with this reconstruction?

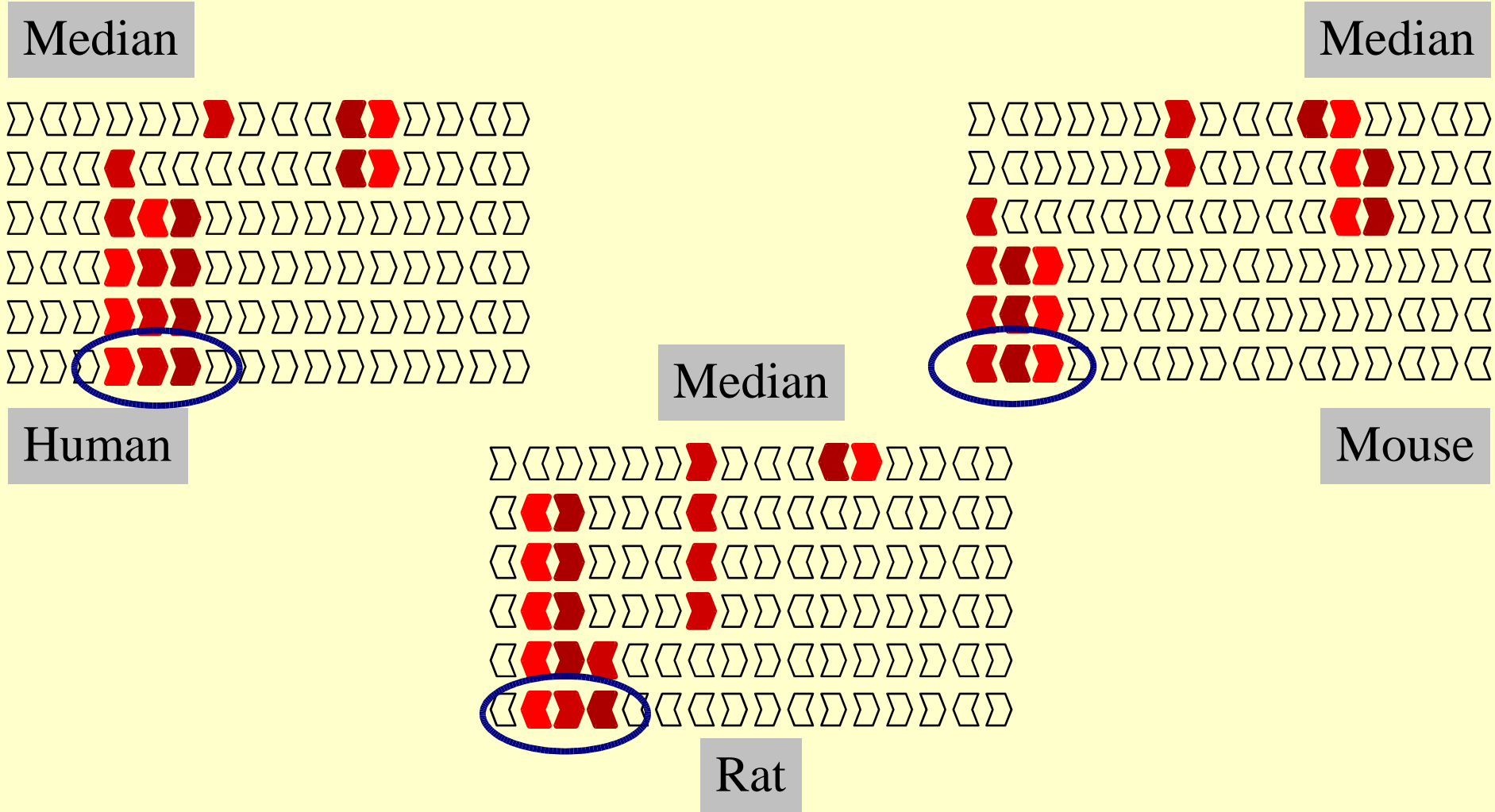
1. Nothing. It happened that way.
2. The “inversion only” model does not apply.
3. The parsimony assumption does not apply.

4. The data is wrong.

Easiest to check!

Should we question scenarios that are not perfect?

We investigated these three blocks in the various assemblies of the Mouse, Human and Rat X chromosomes.



Should we question scenarios that are not perfect?

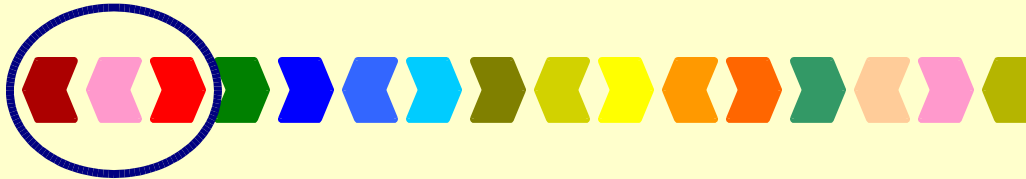
The data *was* wrong.

Mouse 30

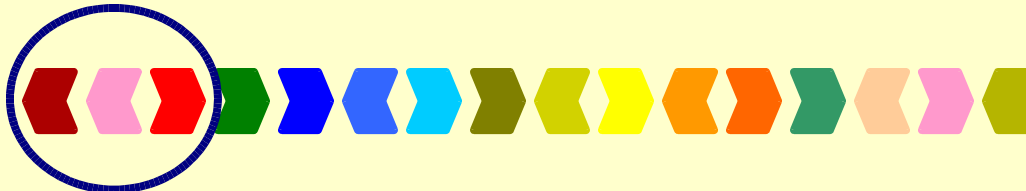


Assembly used in Bourque,
Pevzner and Tesler, 2004.

Mouse 32



Mouse 33



Should we question scenarios that are not perfect?

Mouse 30

Mouse 32

Human to Mouse

4/10

8/10

Human to Rat

2/10

6/10

Rat to Mouse

4/10

4/10

Number of inversions that do not break a common interval.
Comparison of scenarios using two different assemblies of the mouse.

Is it easy to construct perfect scenarios?

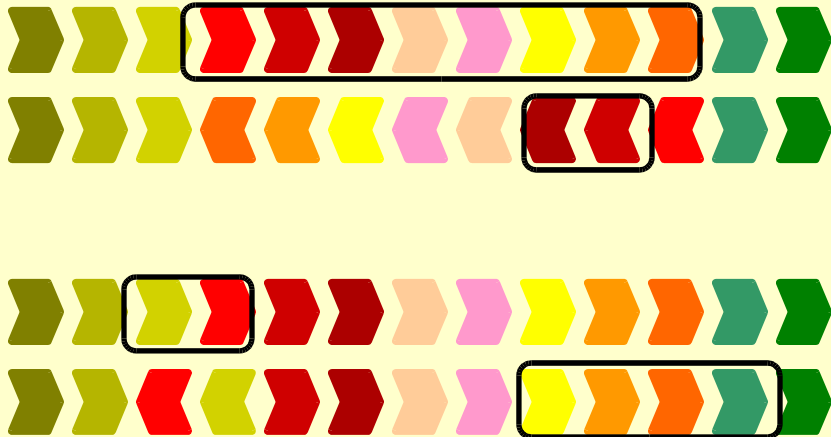
1. The construction of perfect scenarios is computationally difficult.
(Figeac and Varré, 2004)
2. The construction of optimal *commuting* scenarios is computationally easy.
(Bérard, Bergeron and Chauve, 2004)

Is it easy to construct perfect scenarios?

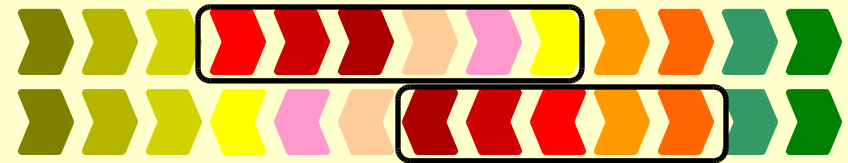
Definition:

A *commuting* scenario is a rearrangement scenario in which all pairs of inversions trivially overlap.

Trivial overlaps



Non-trivial overlap



Note that the second inversion is not an interval of the first genome.

Is it easy to construct perfect scenarios?

Basic Property:

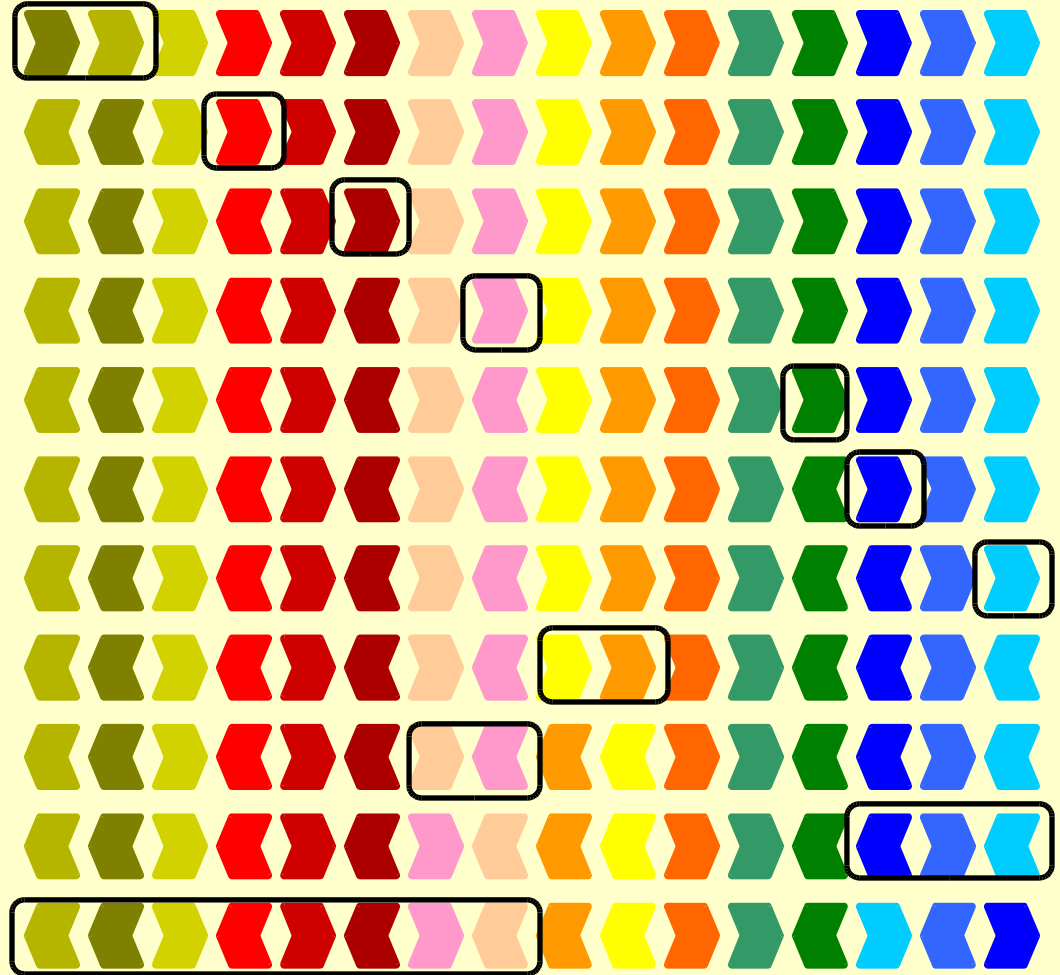
Commuting scenarios are perfect!

A Commuting Scenario

This scenario transforms 16 blocks of a region of Human chromosome 17 into a region of Mouse chromosome 11.

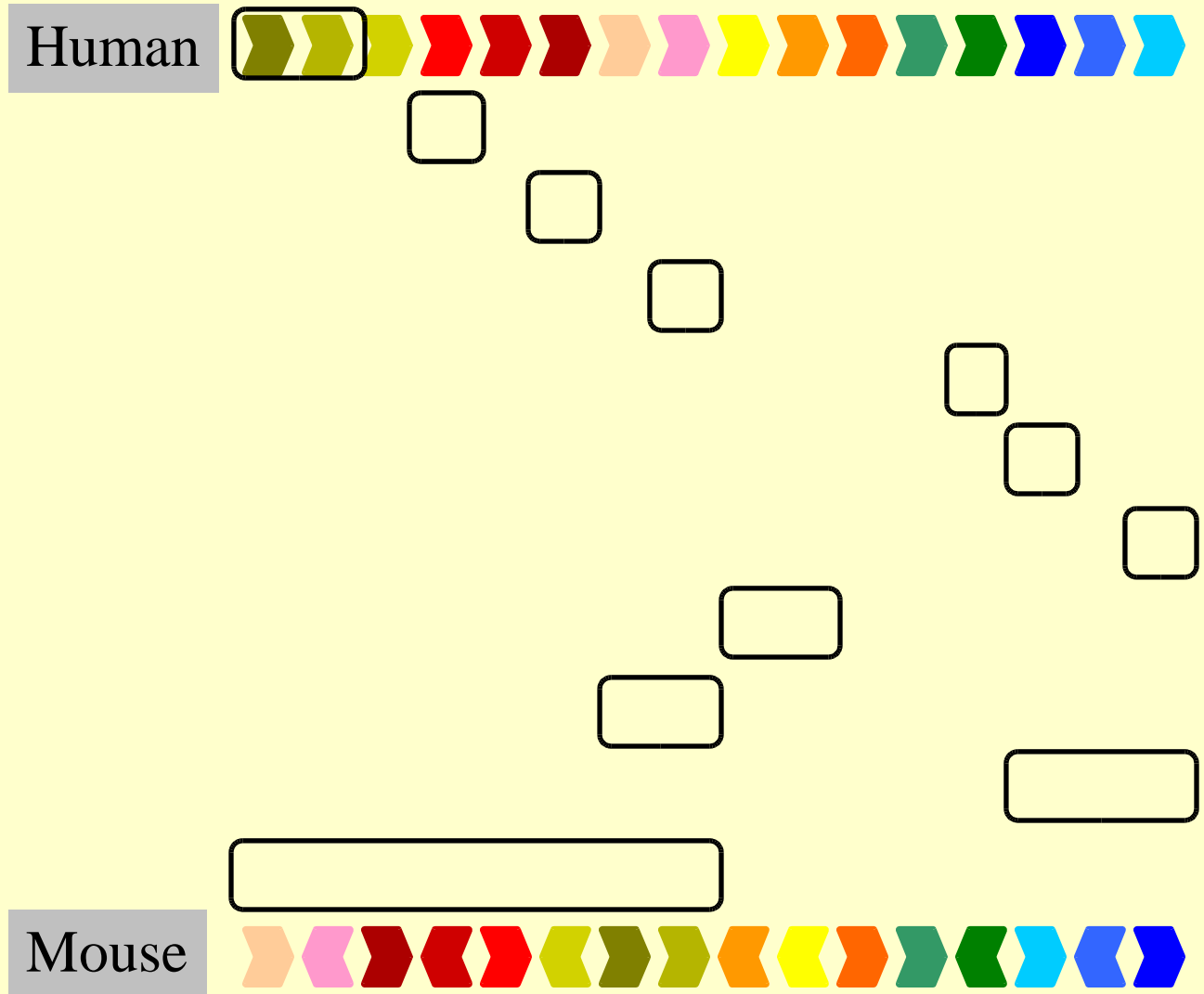
It conserves **all common intervals** between Human and Mouse.

Human

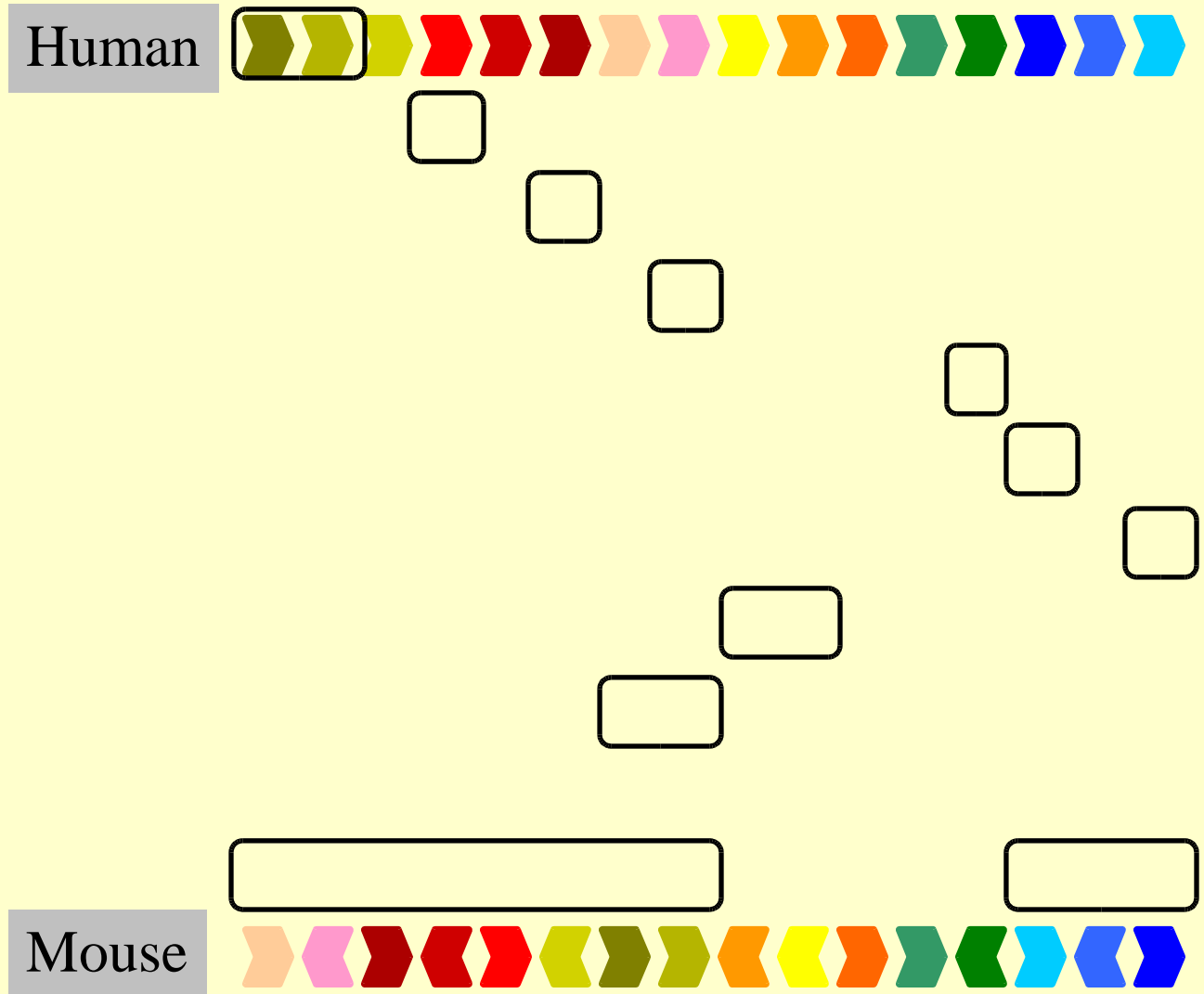


Mouse

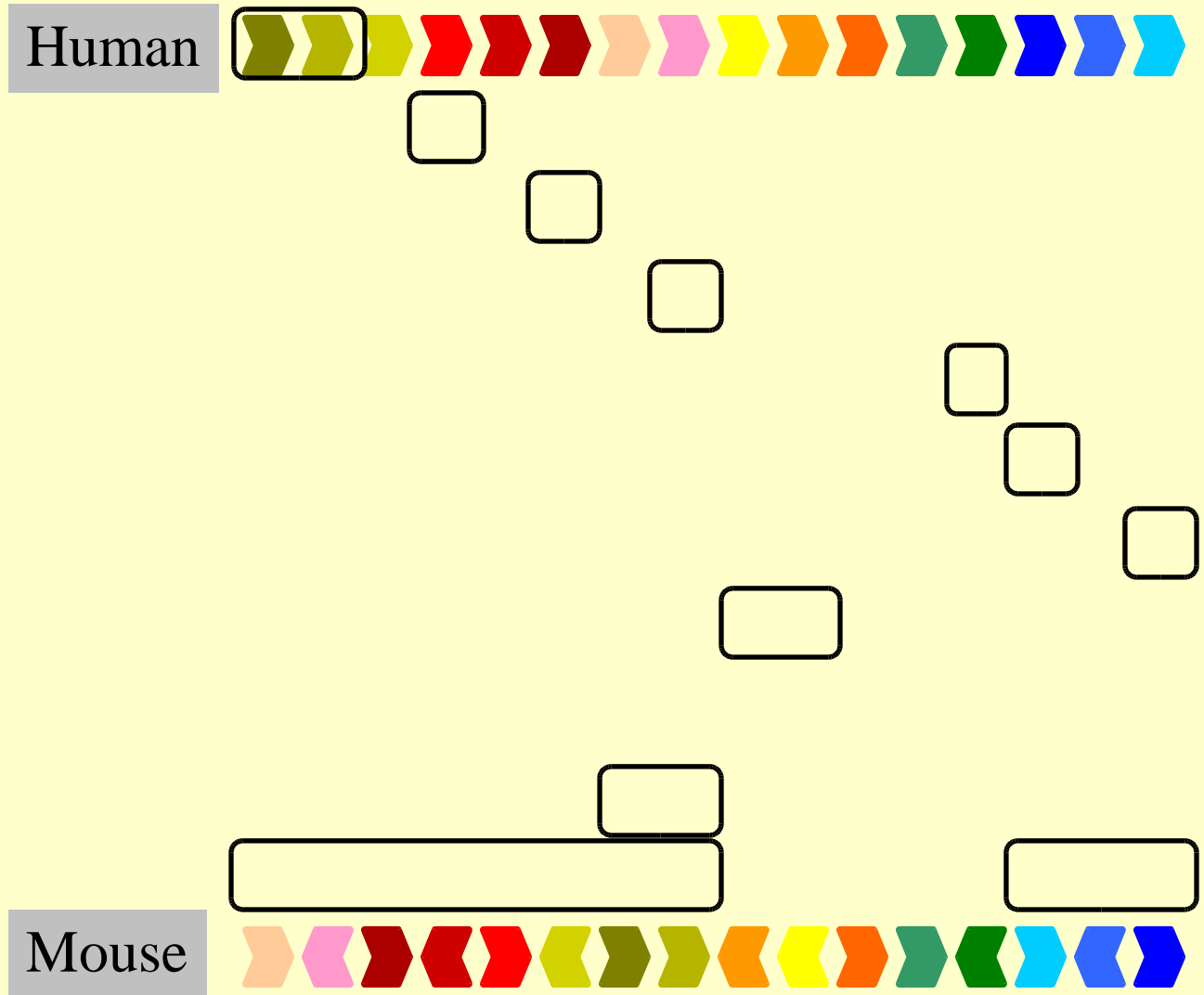
A Commuting Scenario



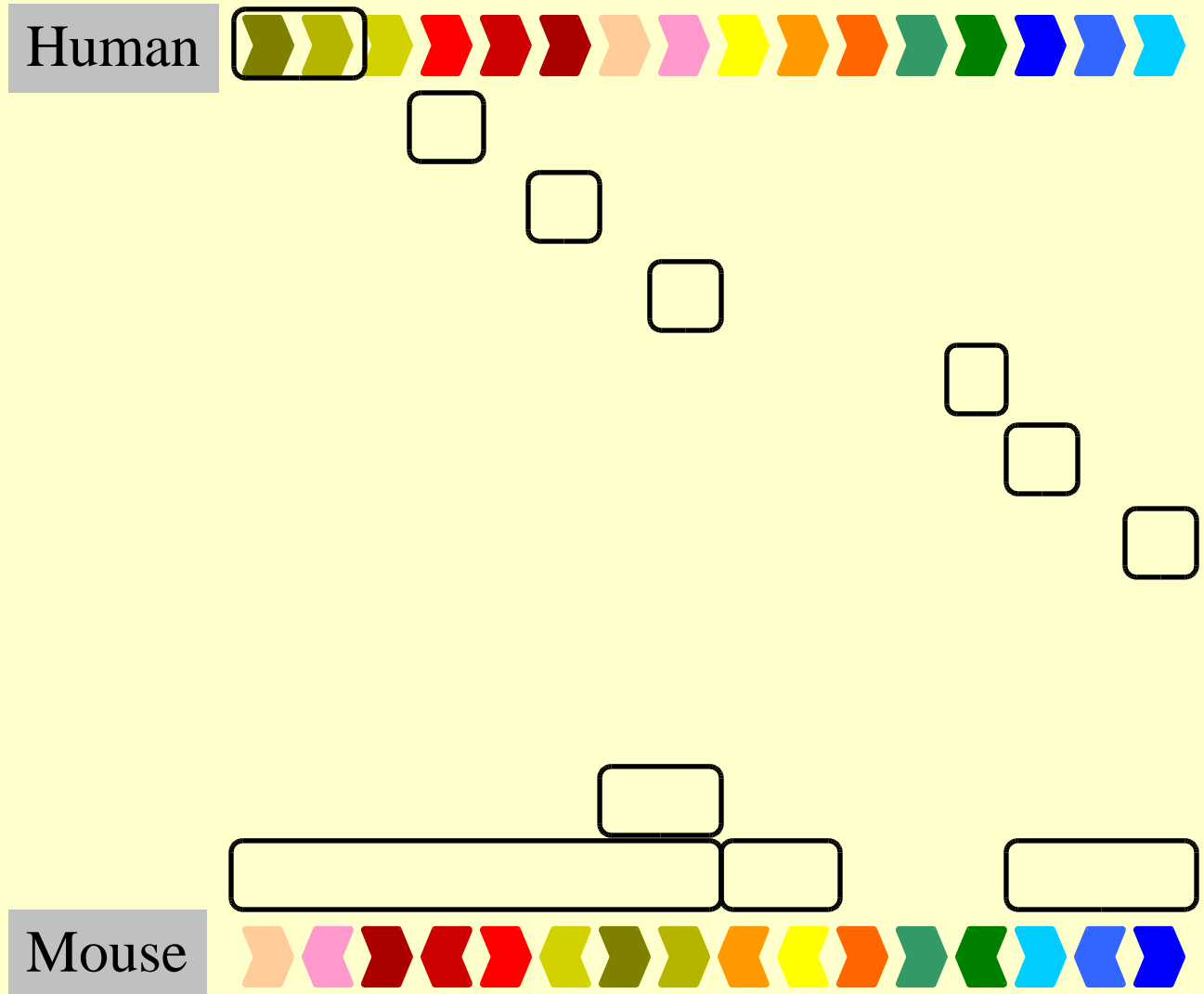
A Commuting Scenario



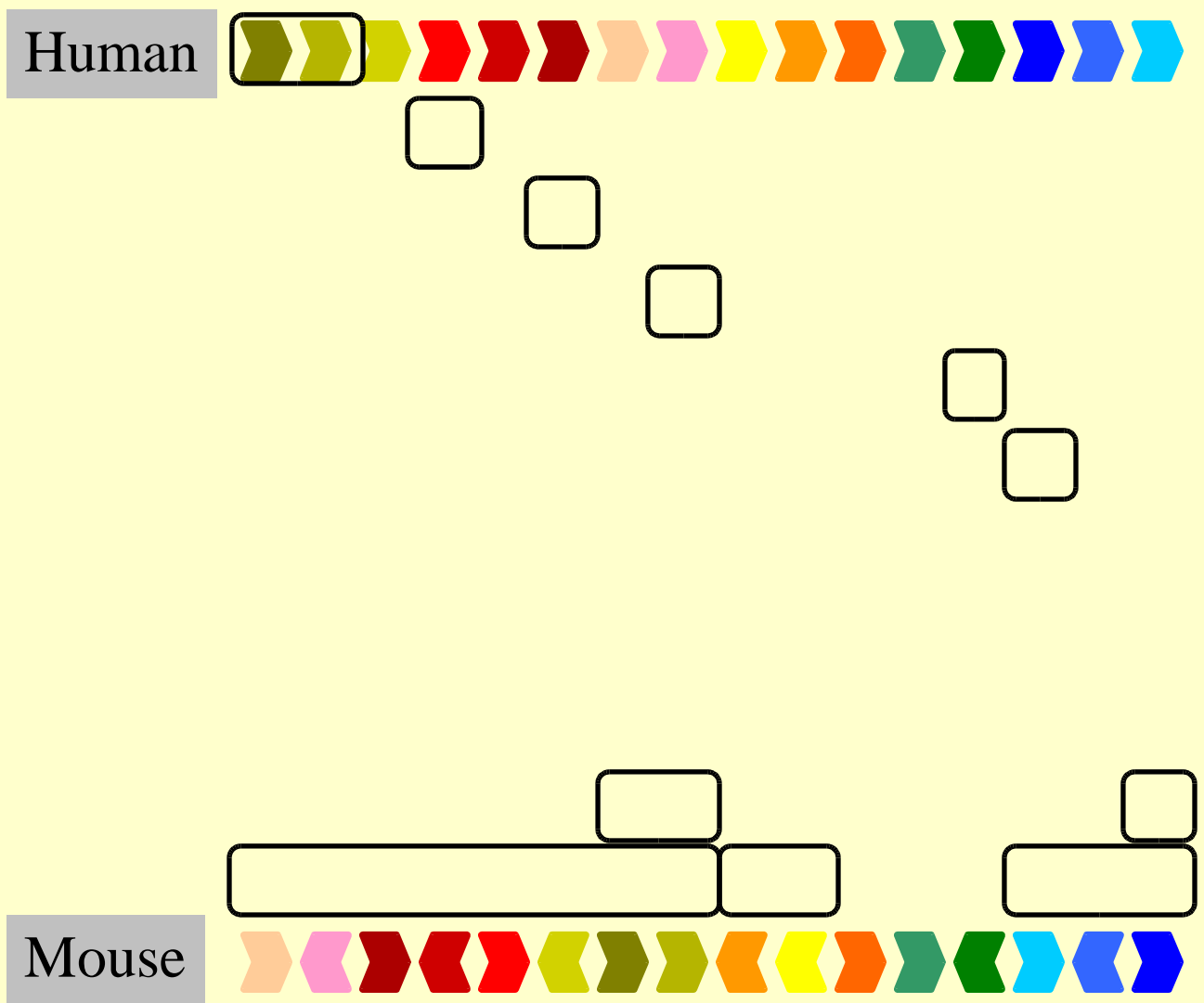
A Commuting Scenario



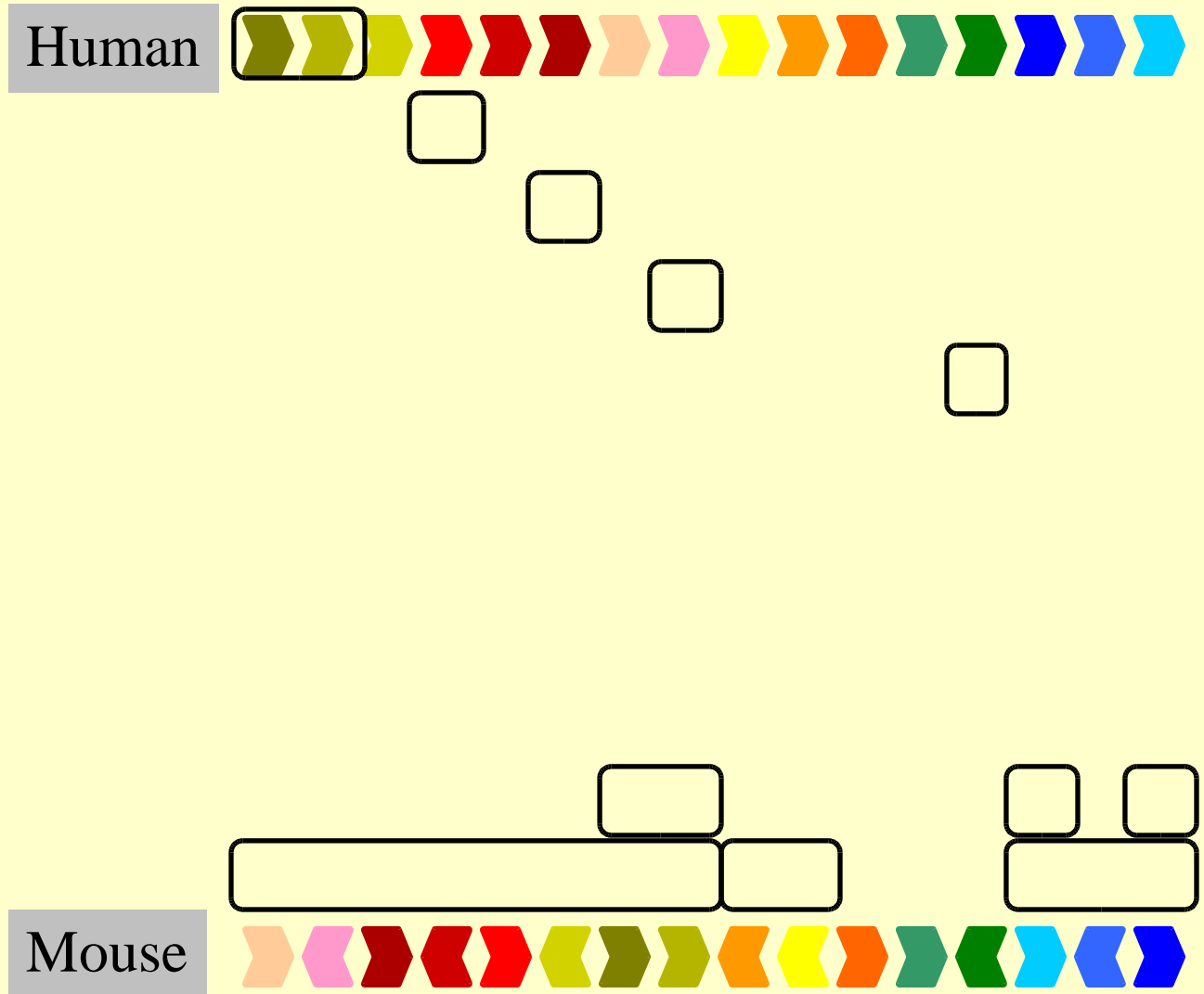
A Commuting Scenario



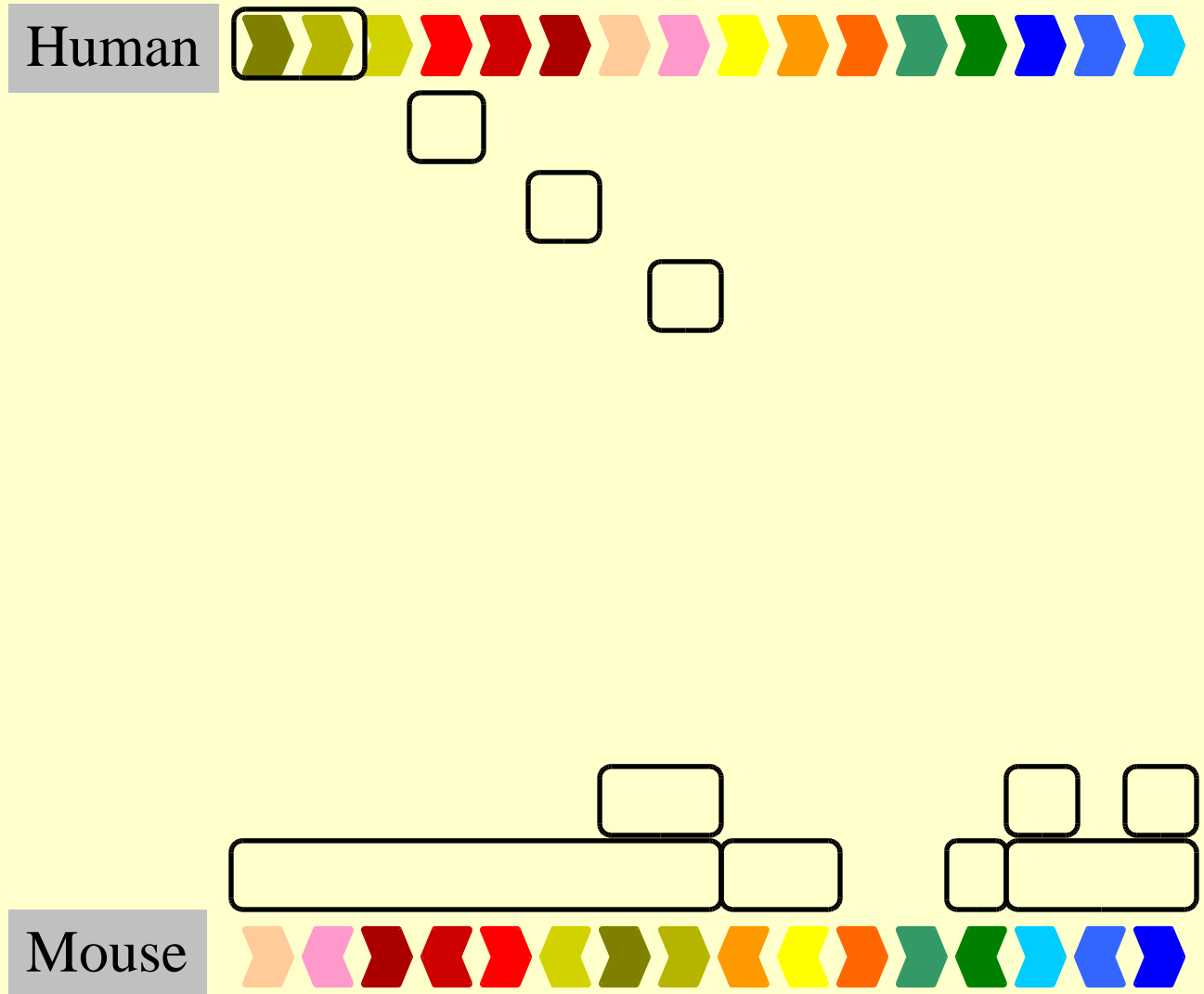
A Commuting Scenario



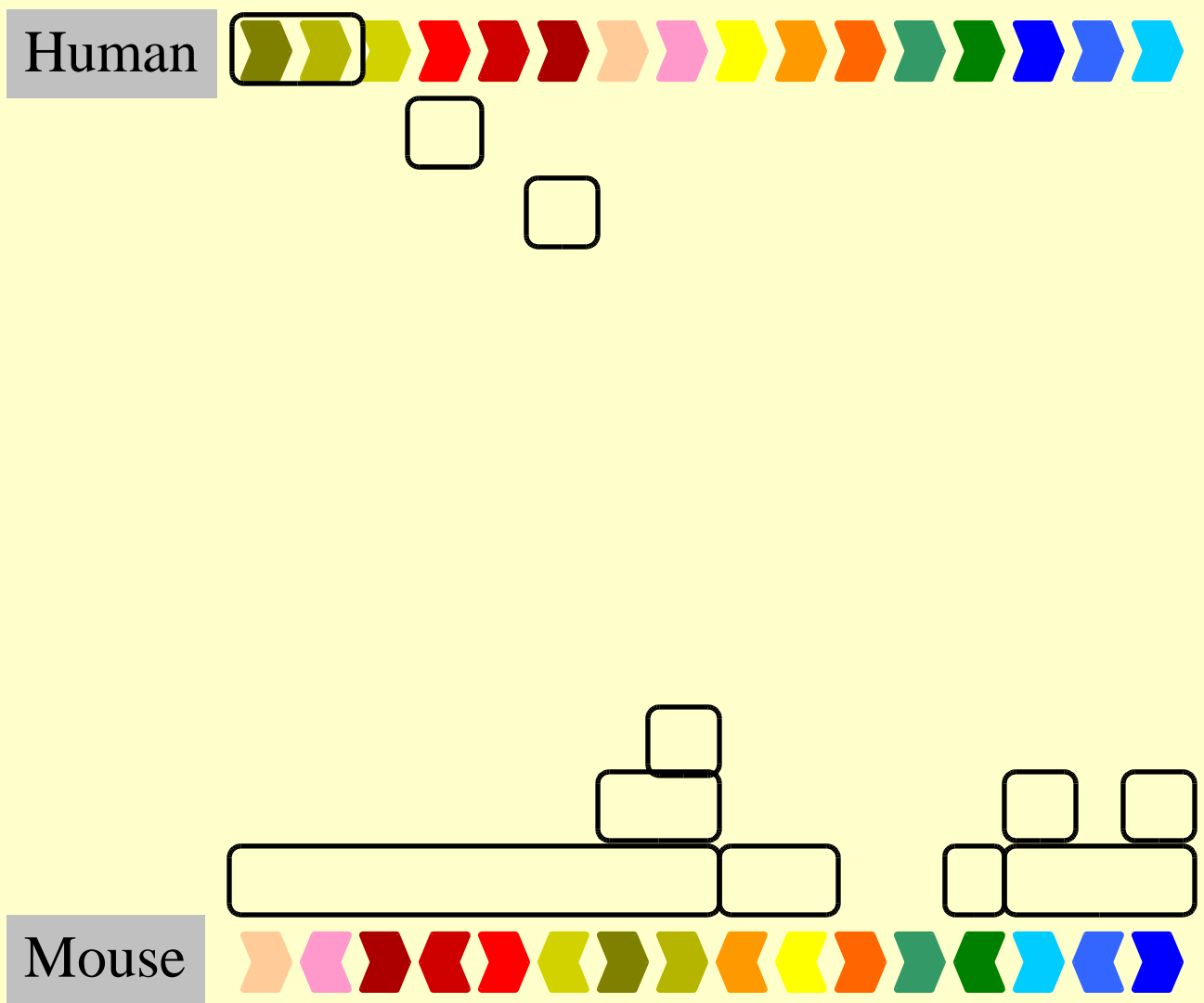
A Commuting Scenario



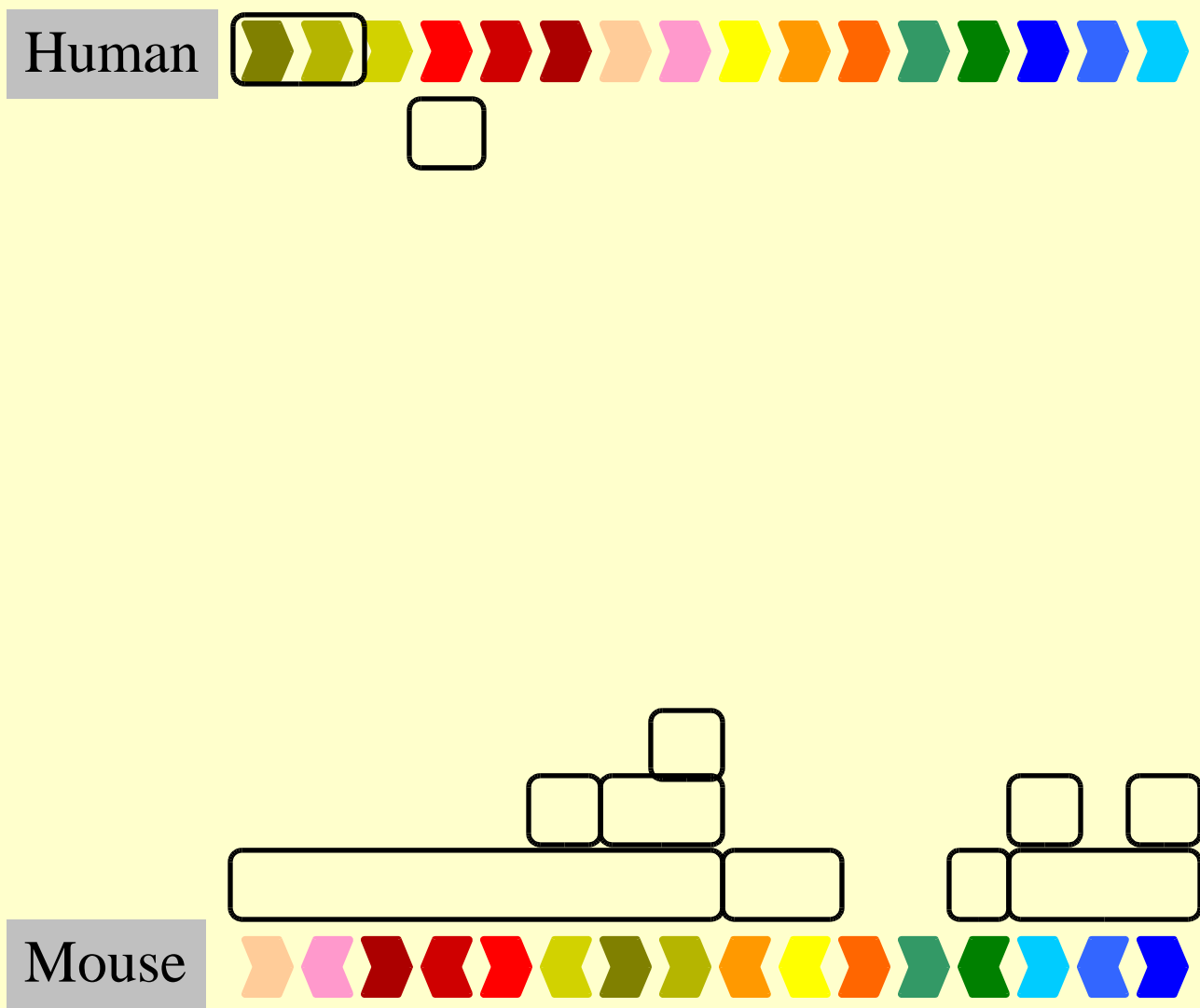
A Commuting Scenario



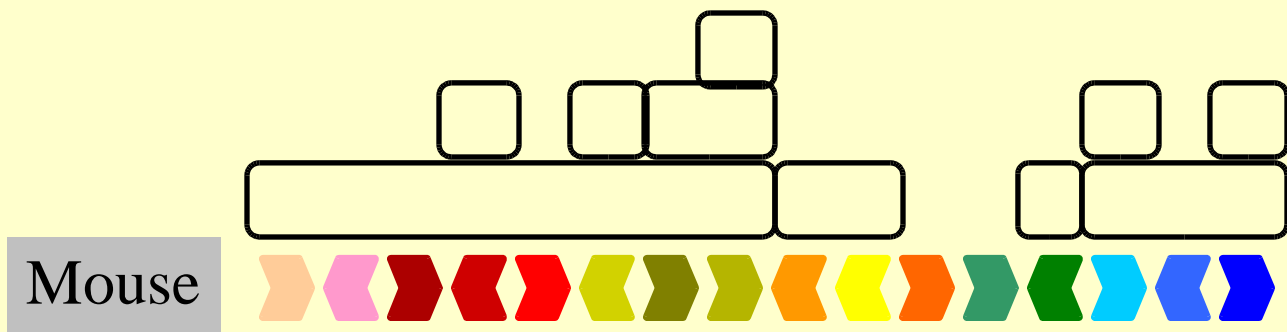
A Commuting Scenario



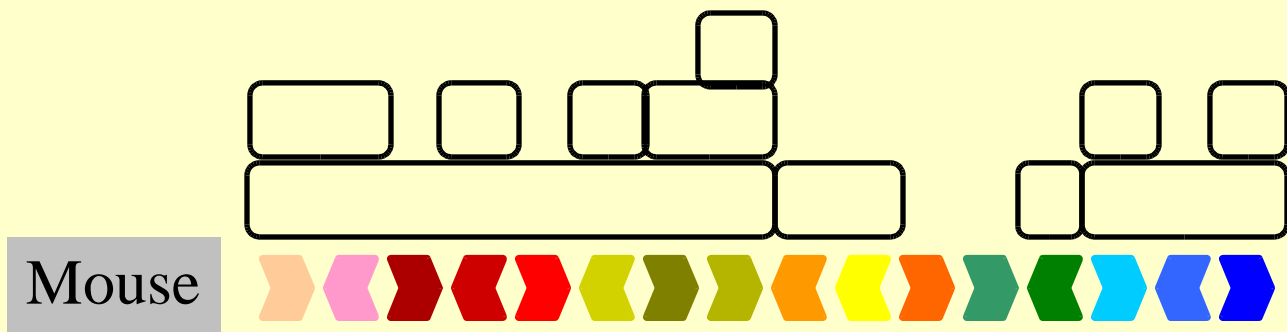
A Commuting Scenario



A Commuting Scenario

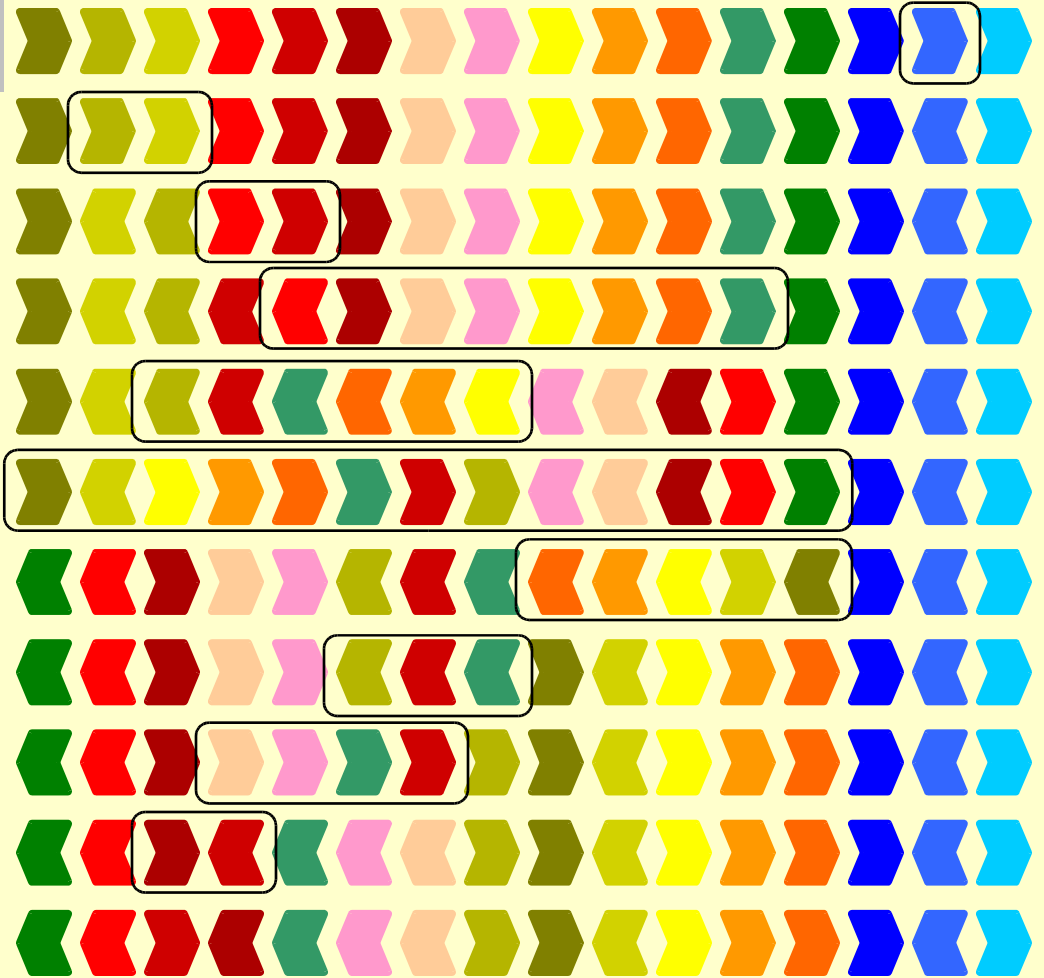


A Commuting Scenario



A Non-Commuting Scenario

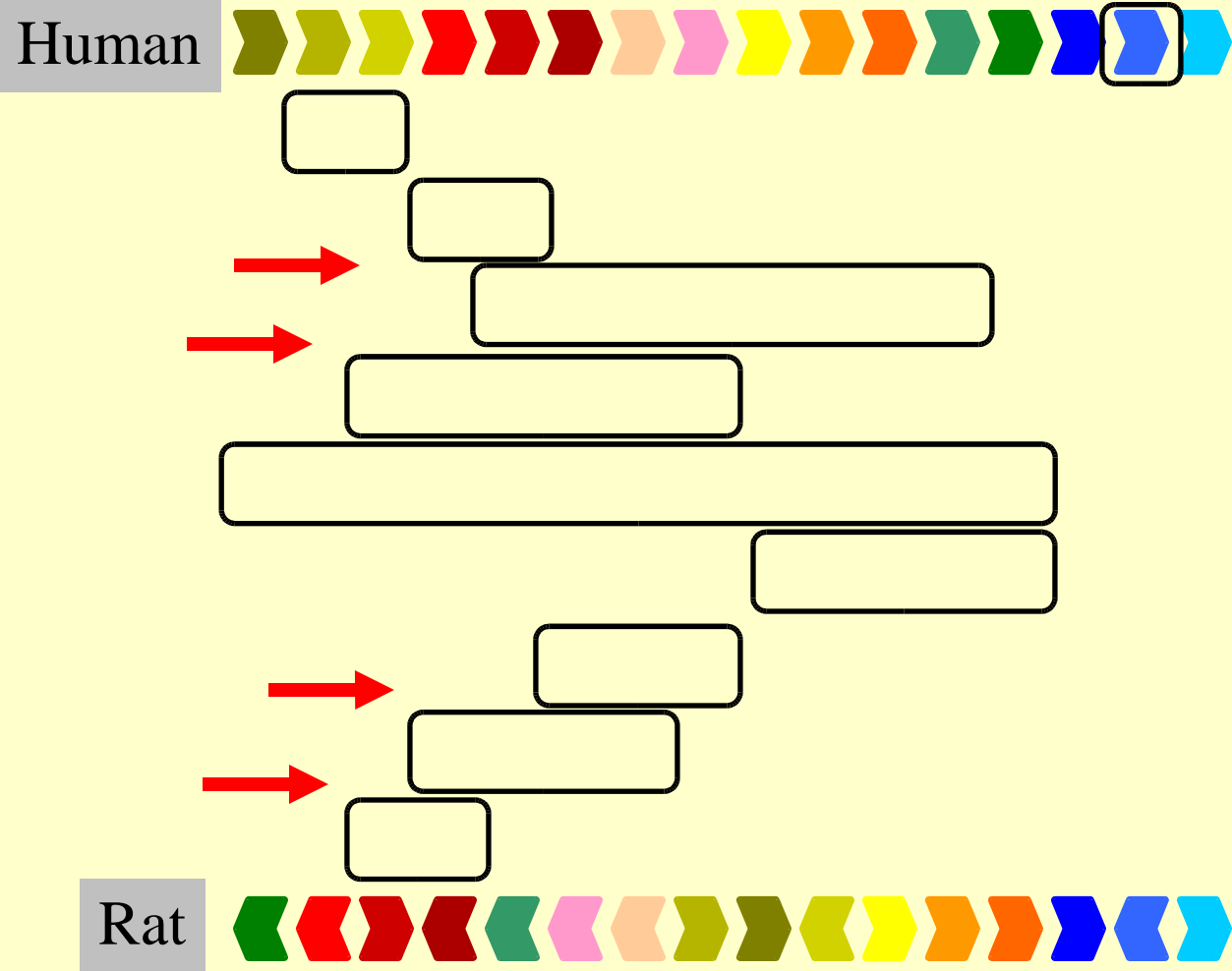
Human



Rat

Transforming the
Human chr. X into the
Rat chr. X.

A Non-Commuting Scenario



Theoretical results

Observations:

3. Perfect scenarios always exist between genomes.

Example:

$(1\ 2\ 3\ 4)$ and $(3\ 1\ 4\ 2)$ have only one common interval, therefore, any rearrangement scenario is perfect.

Theoretical results

Observations:

3. Two genomes can have an optimal commuting scenarios, **and** optimal non-commuting ones.

Example: [from the Human chr. 17 and Mouse chr. 11]

Human = (1 2 3 4 5 6 7 8)

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

Theoretical results

Theorem:

The following statements are equivalent:

5. There exists an optimal commuting scenario between two genomes.
7. There exists an optimal scenario in which each inversion is a common interval.
3. The connected components of the overlap graph of the two corresponding permutations are trees, or cycles of odd length.

Theoretical results

Consequences:

Deciding whether a genome can be optimally transformed in another with a commuting scenario can be done in $O(n)$ time.

Identifying a sequence of commuting inversions that transforms a genome in another can be done in $O(n)$ time.

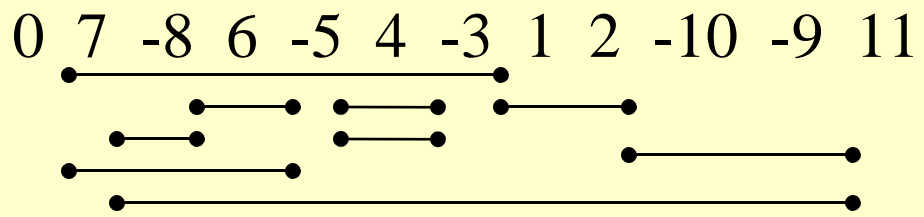
Conclusions and Questions

1. Does nature commute?
2. Can we measure "perfection"?
3. How many broken common intervals can one tolerate in an evolution scenario? How many broken *conserved* intervals? How many broken adjacencies?
4. Perfect scenarios are not necessarily commuting. Deciding efficiently if an optimal perfect scenario exists is an open problem.

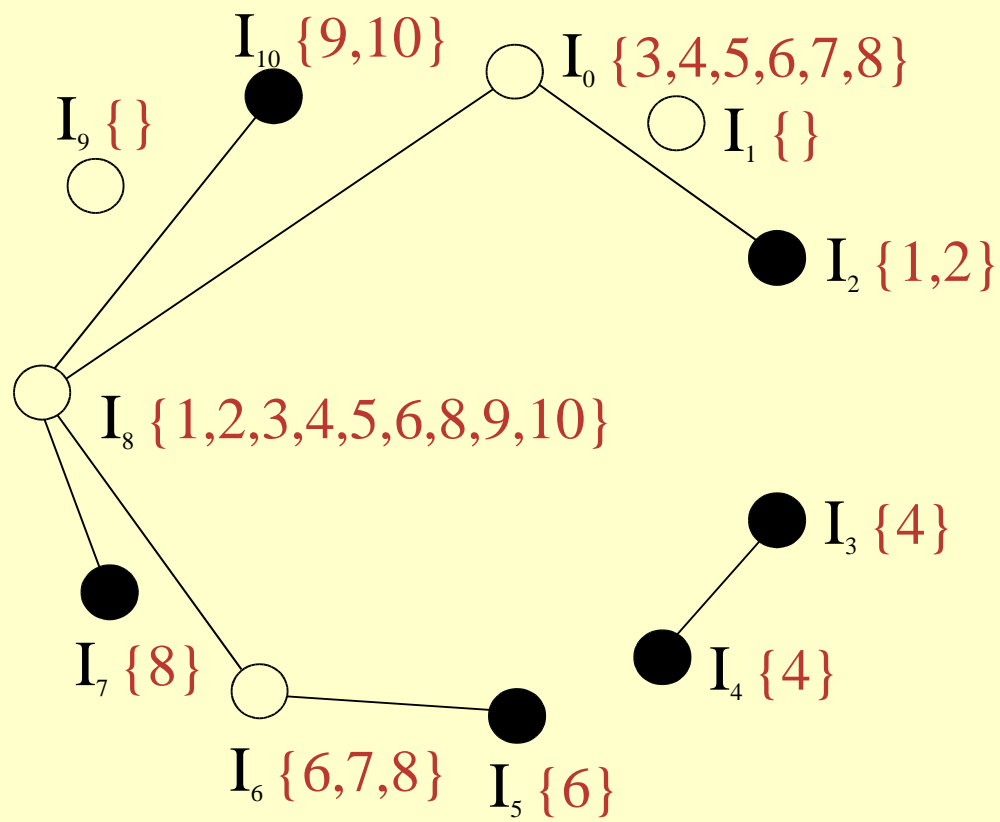
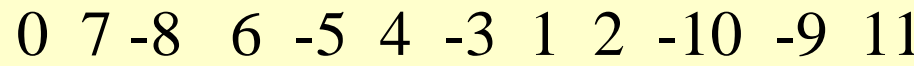
Sorting Chromosome 17

Sorting Chromosome 17

Elementary intervals

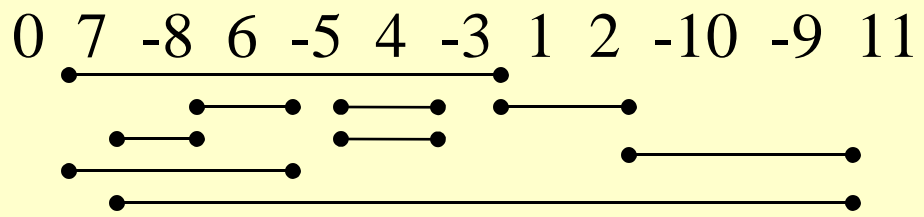


Inversions

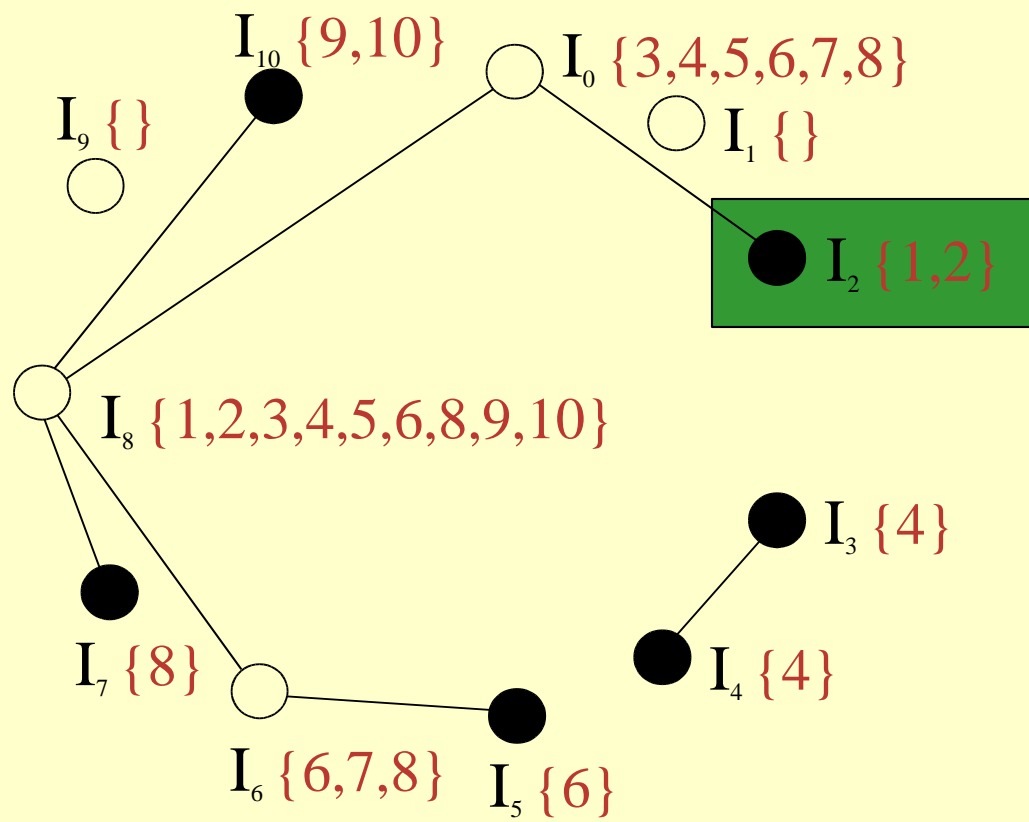
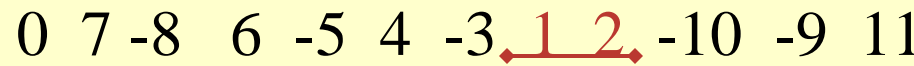


Sorting Chromosome 17

Elementary intervals

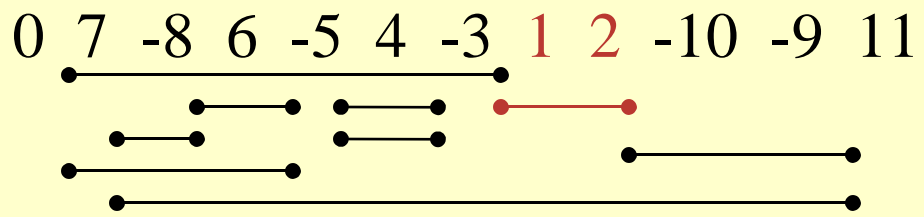


Inversions

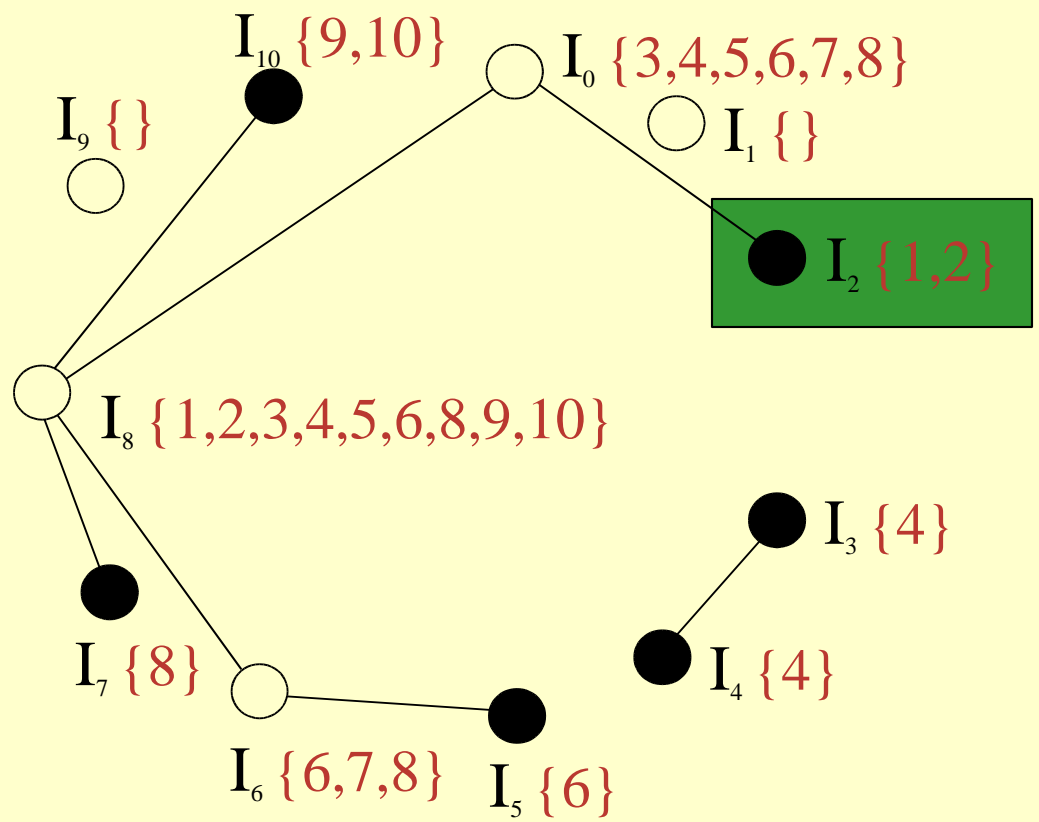
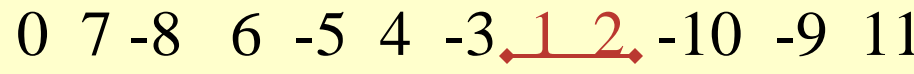


Sorting Chromosome 17

Elementary intervals

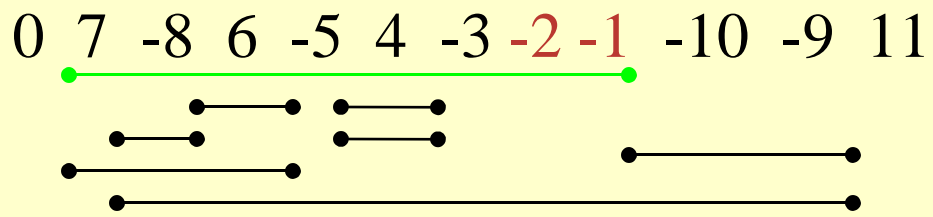


Inversions

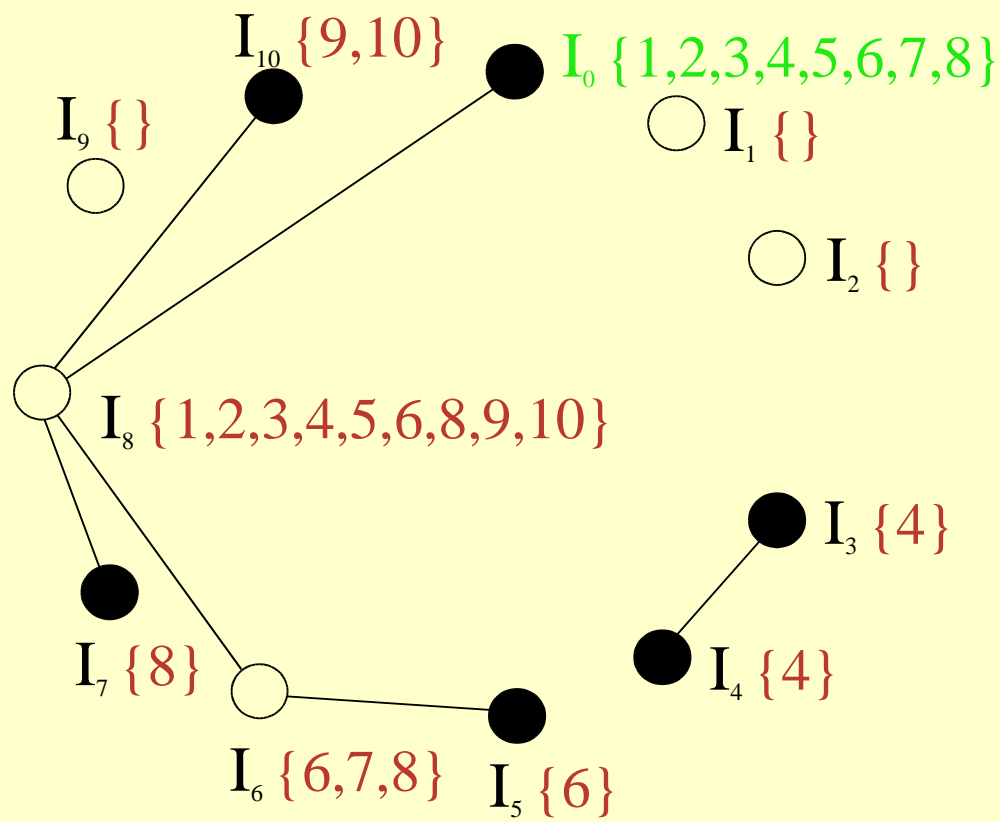
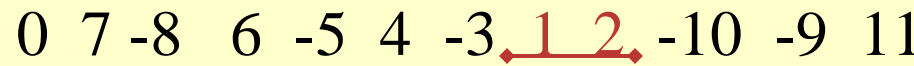


Sorting Chromosome 17

Elementary intervals

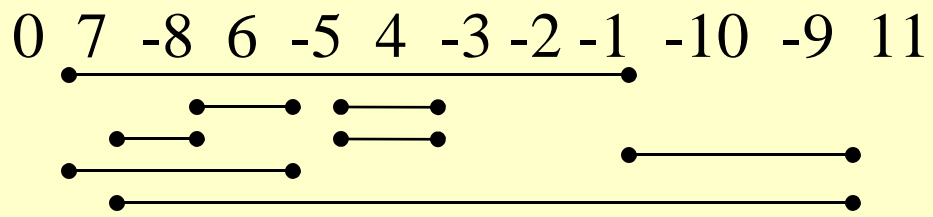


Inversions

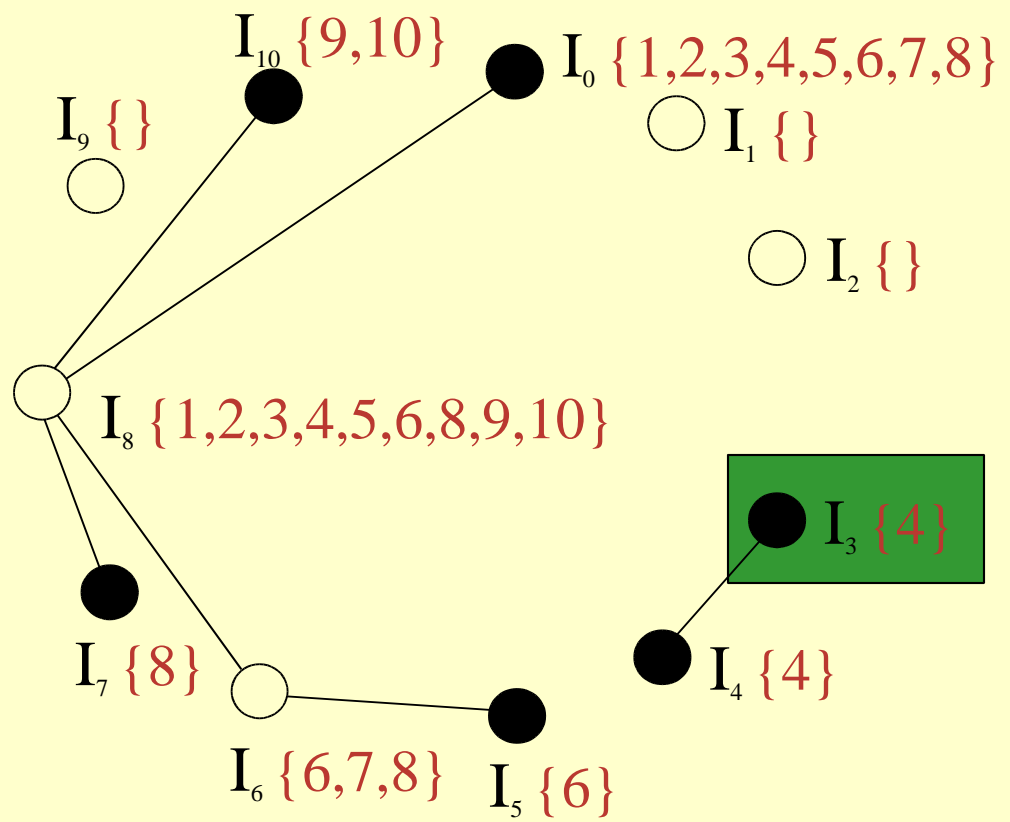


Sorting Chromosome 17

Elementary intervals

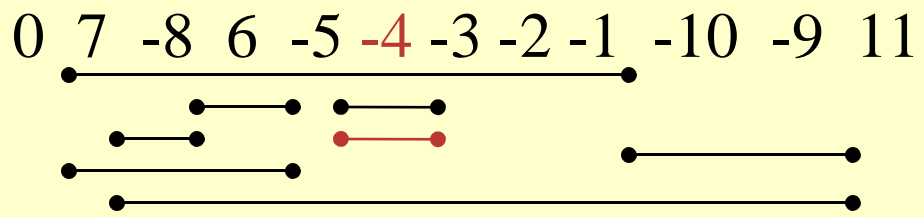


Inversions

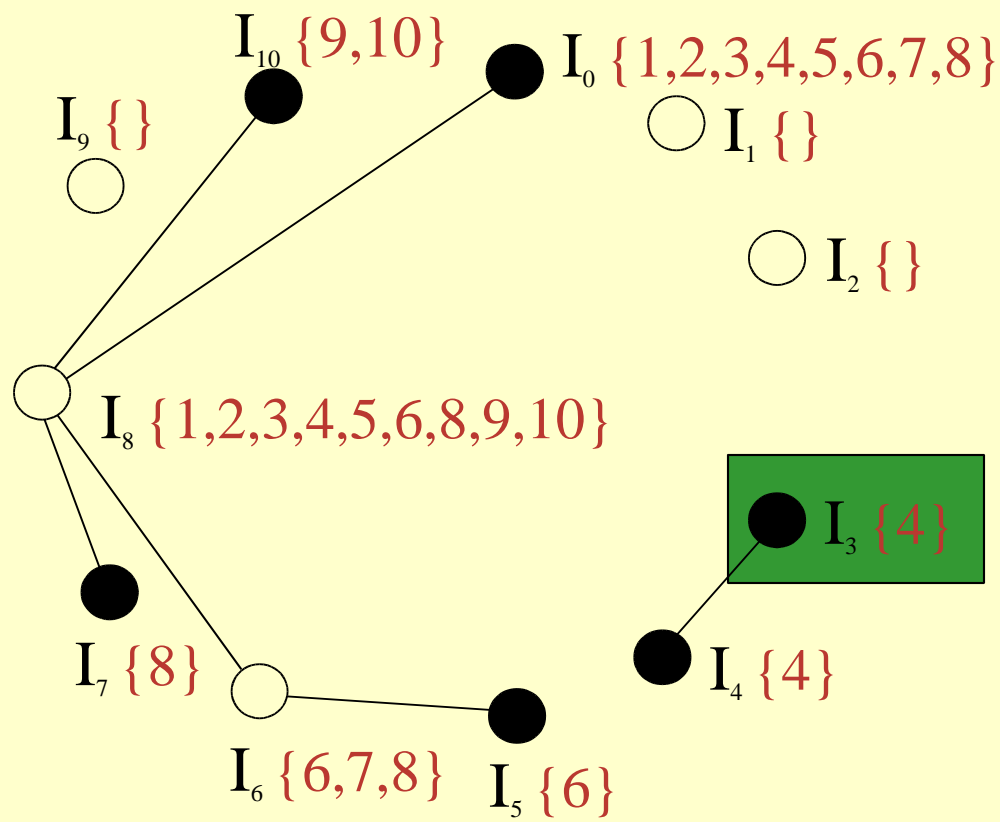


Sorting Chromosome 17

Elementary intervals

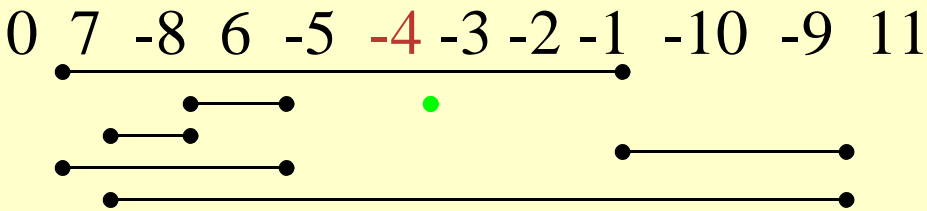


Inversions

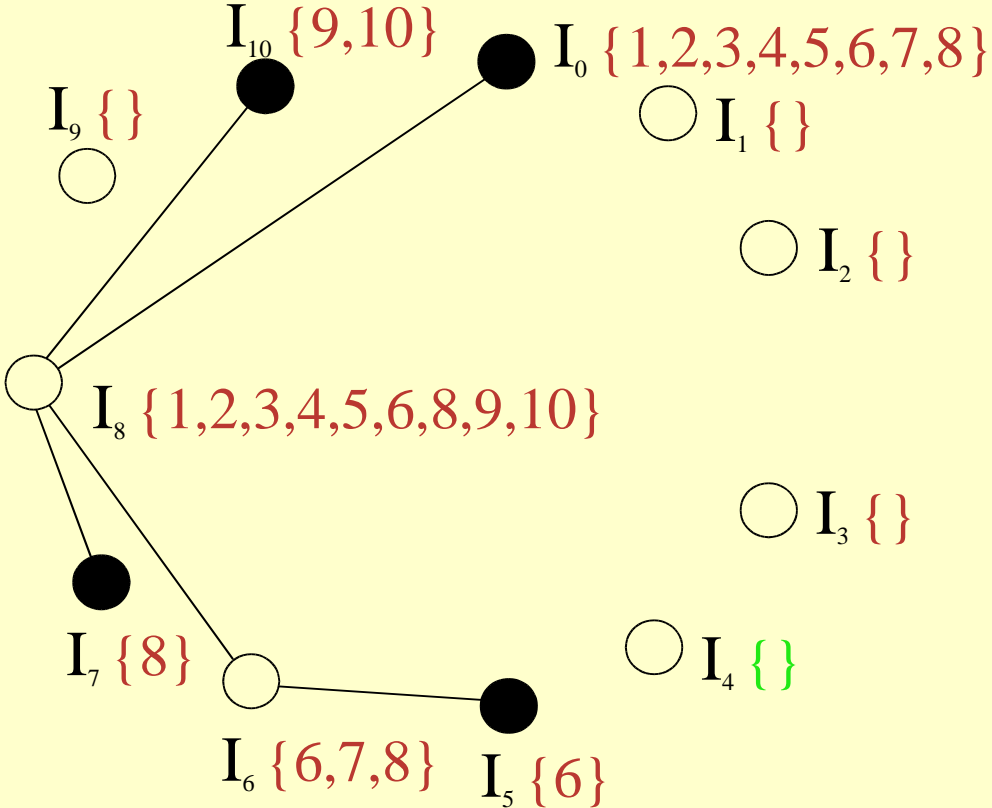
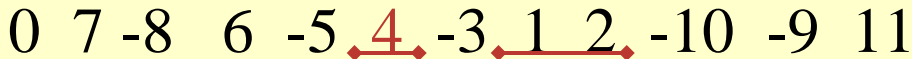


Sorting Chromosome 17

Elementary intervals

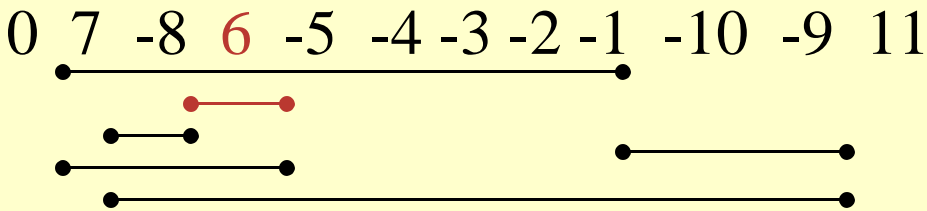


Inversions

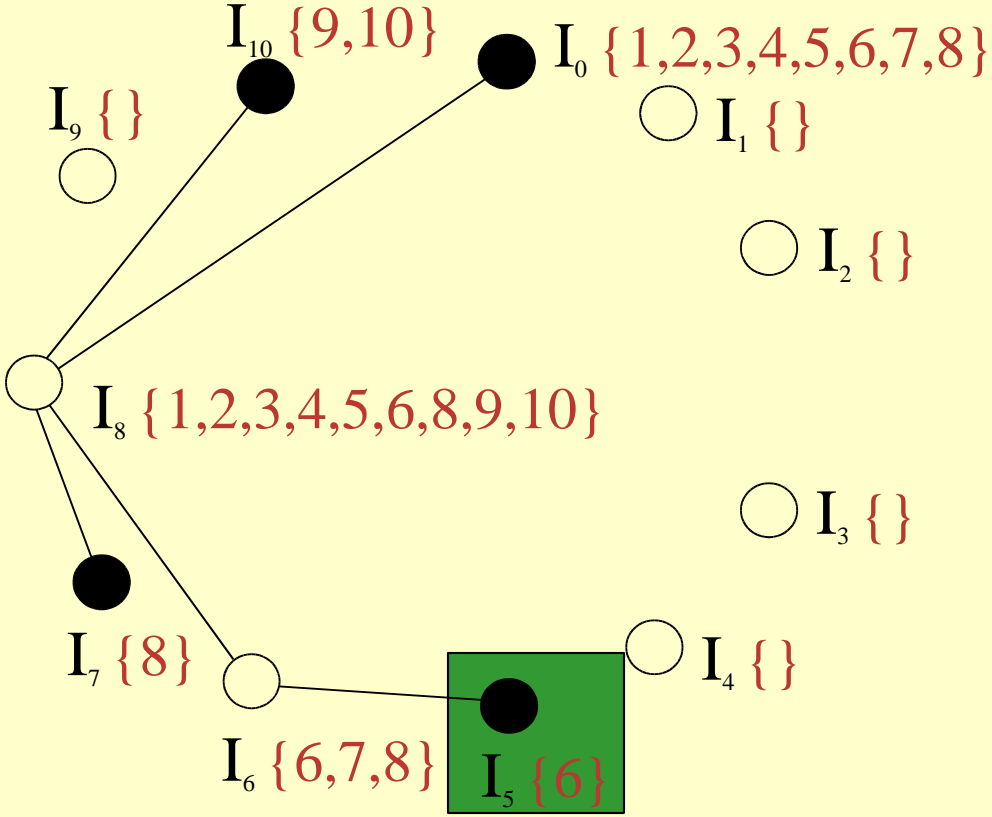
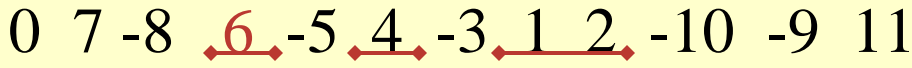


Sorting Chromosome 17

Elementary intervals

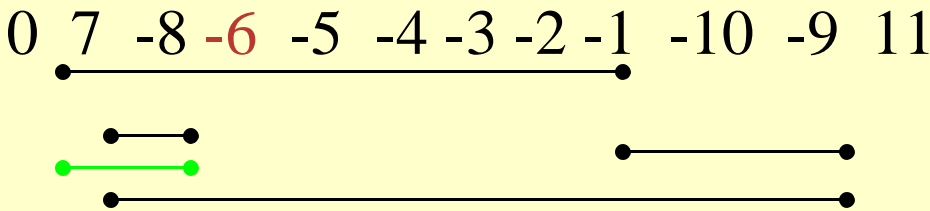


Inversions

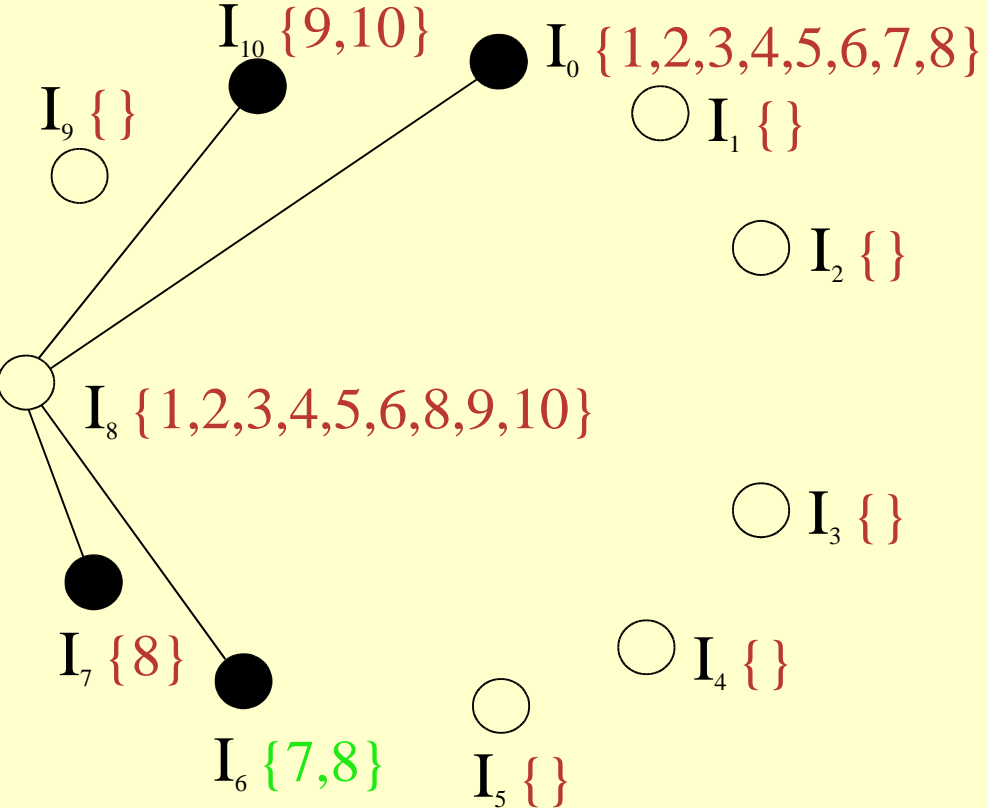
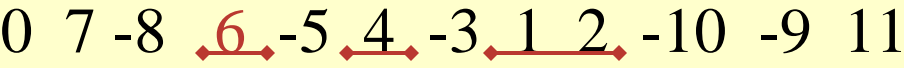


Sorting Chromosome 17

Elementary intervals

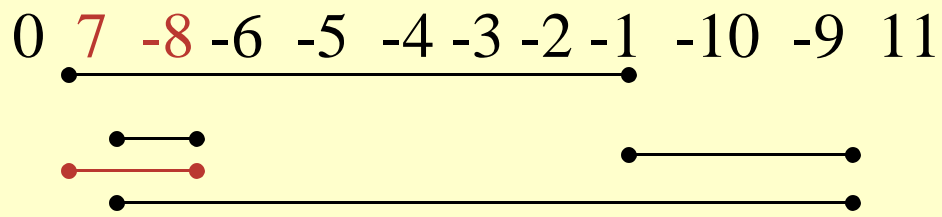


Inversions

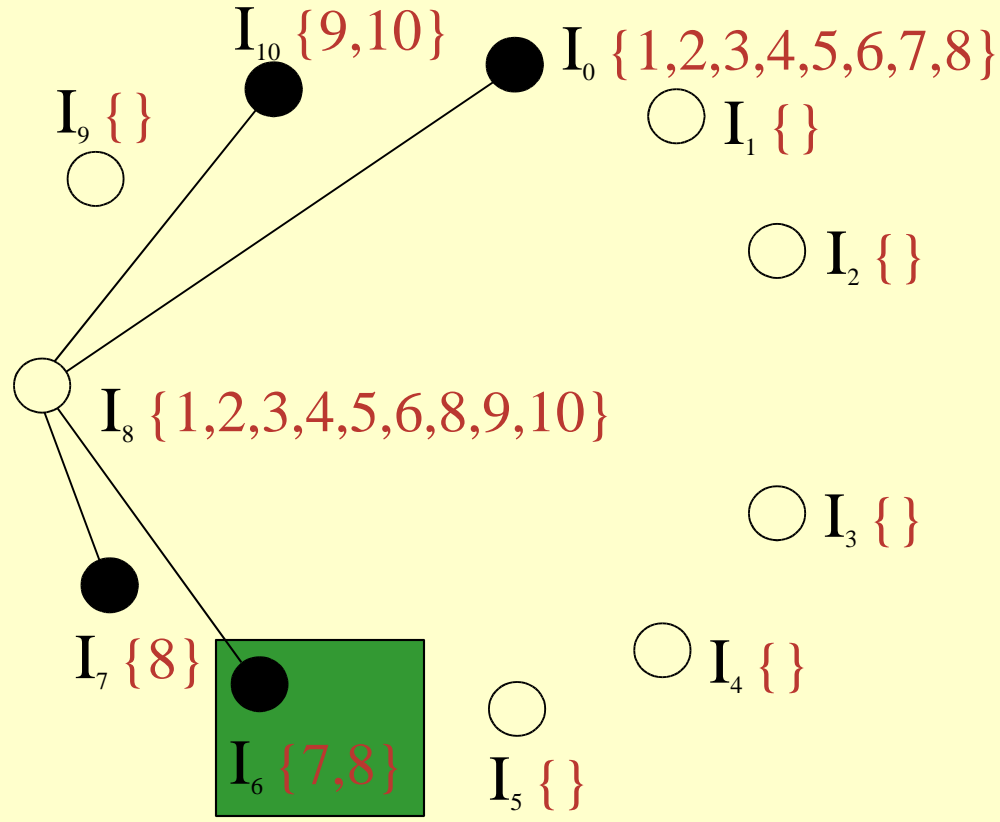


Sorting Chromosome 17

Elementary intervals

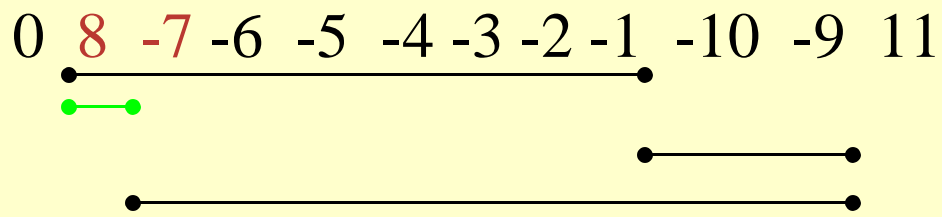


Inversions

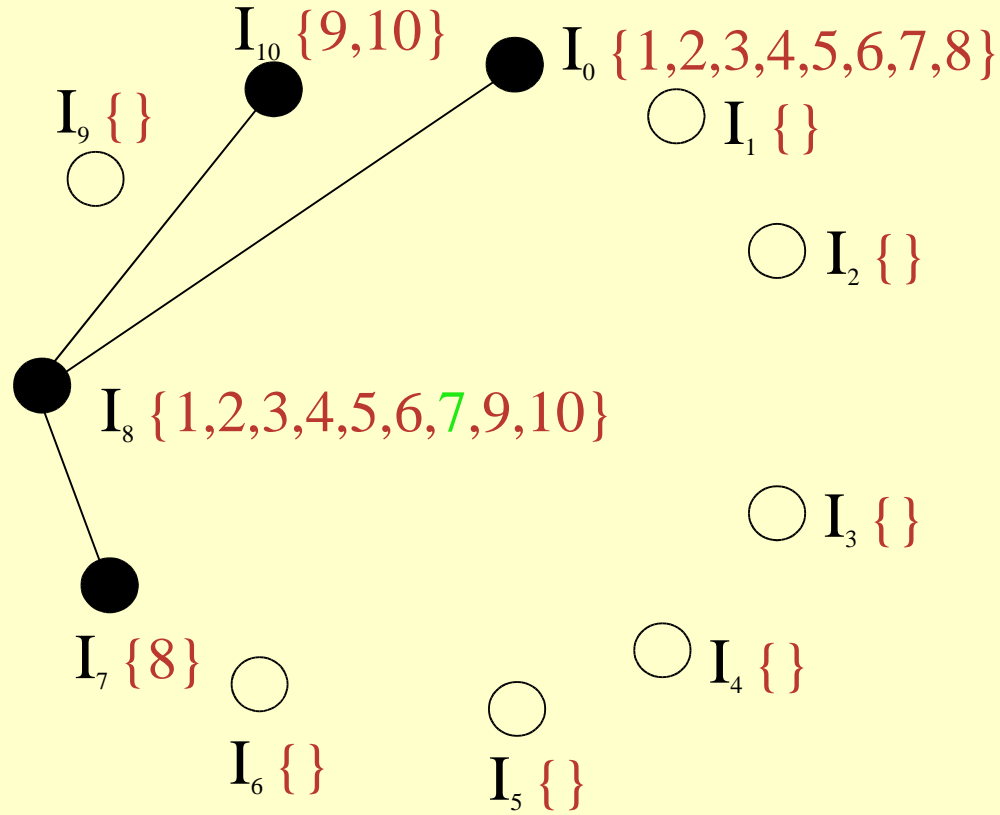


Sorting Chromosome 17

Elementary intervals

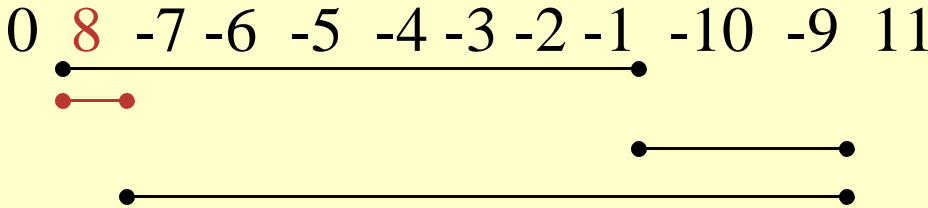


Inversions

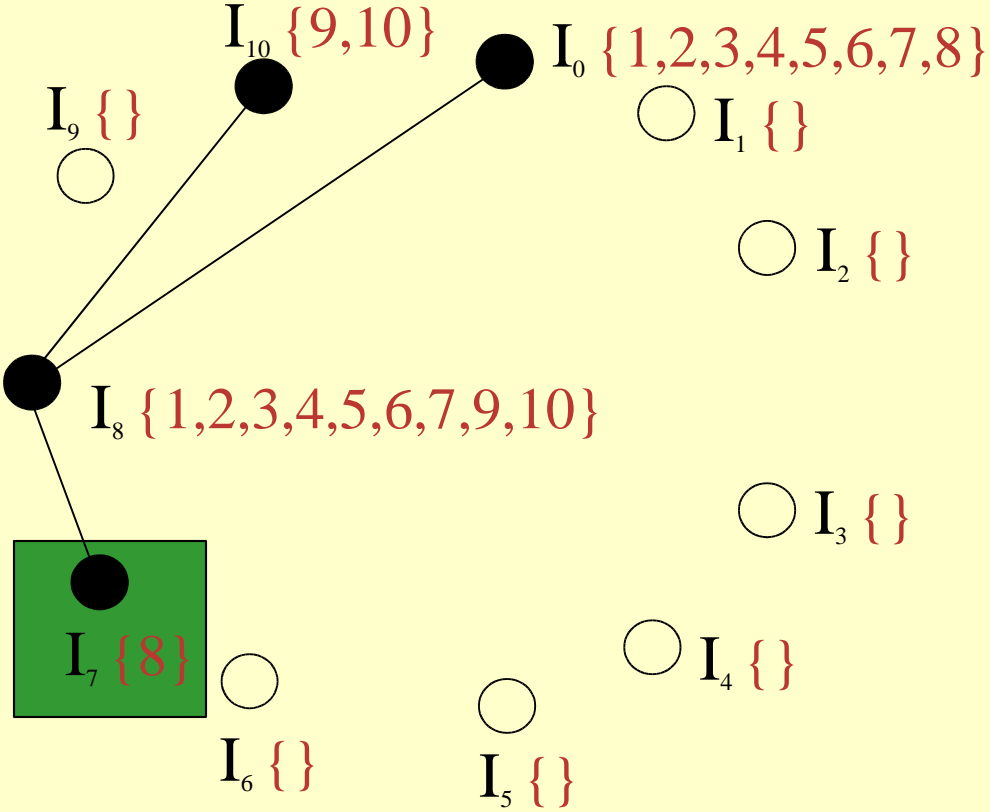
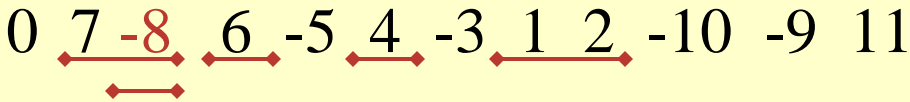


Sorting Chromosome 17

Elementary intervals

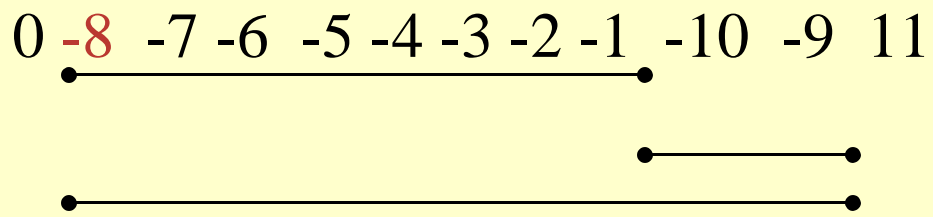


Inversions

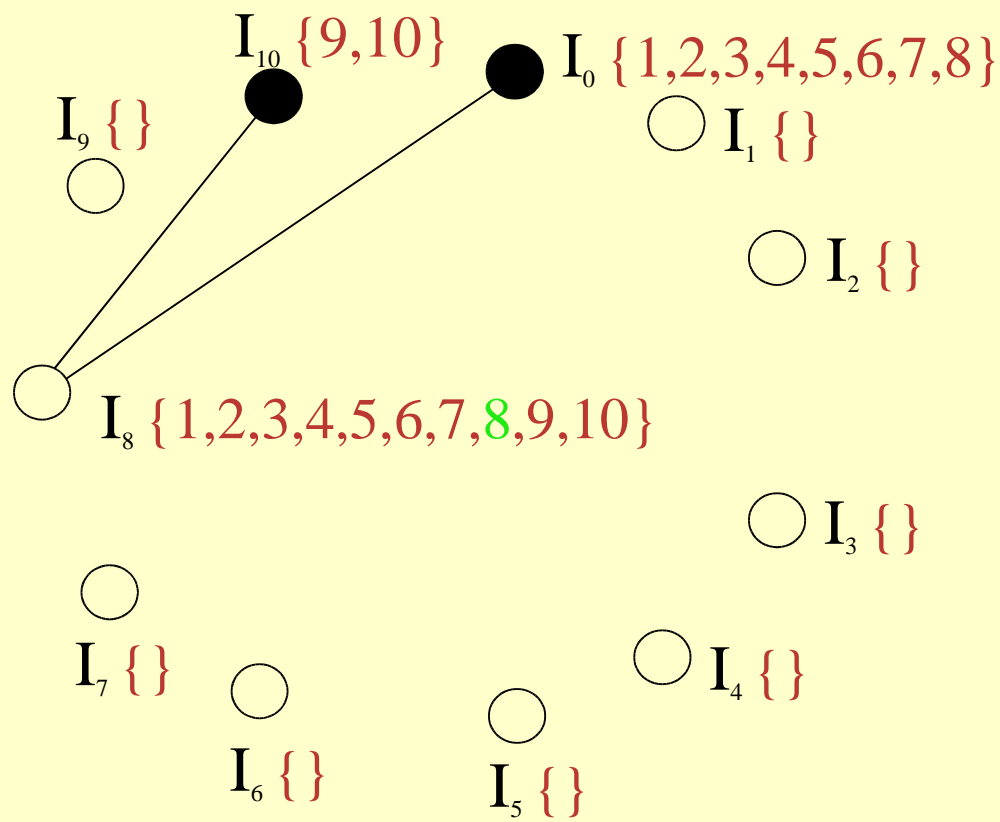
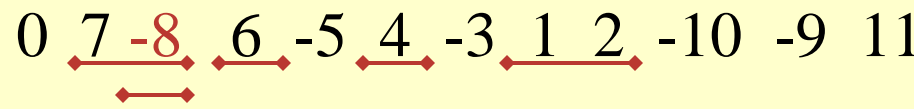


Sorting Chromosome 17

Elementary intervals

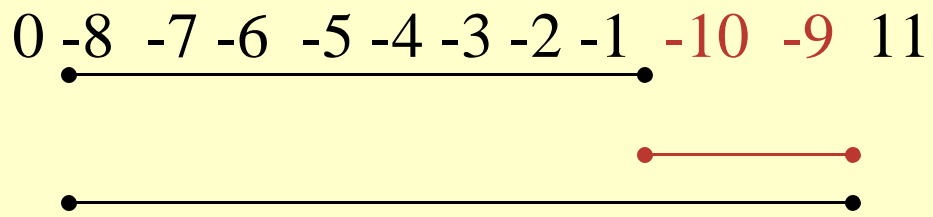


Inversions

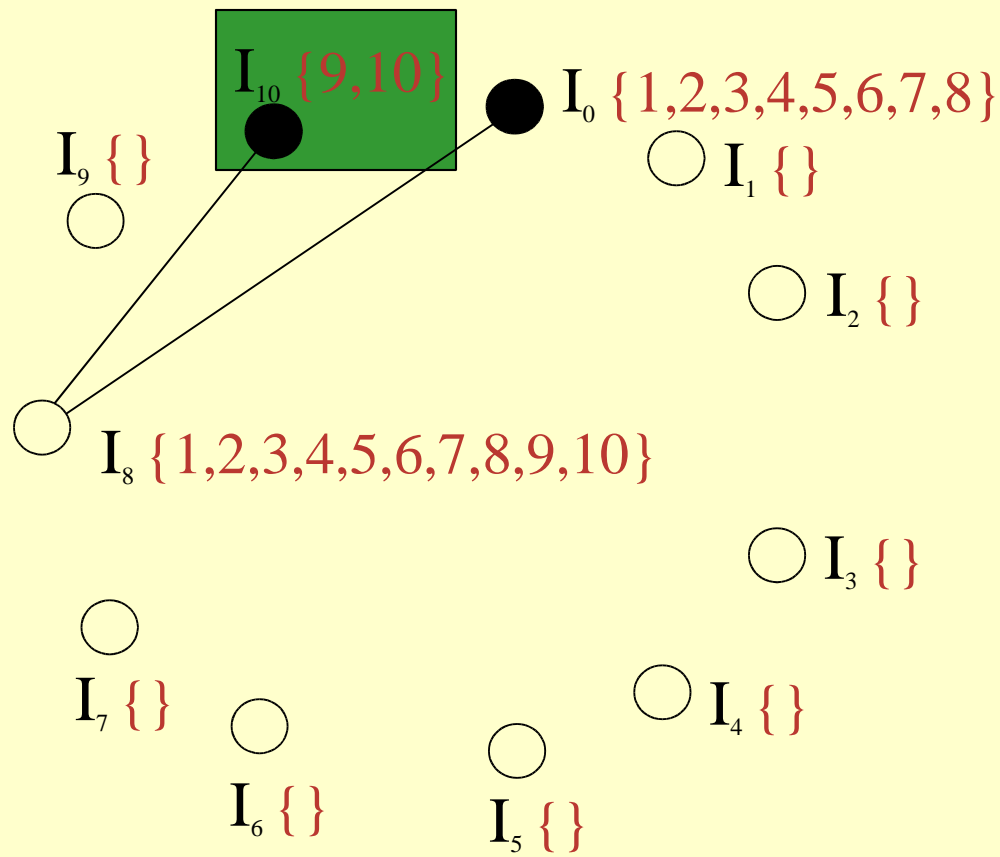
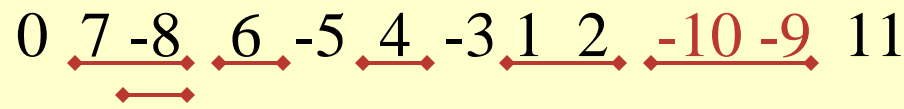


Sorting Chromosome 17

Elementary intervals

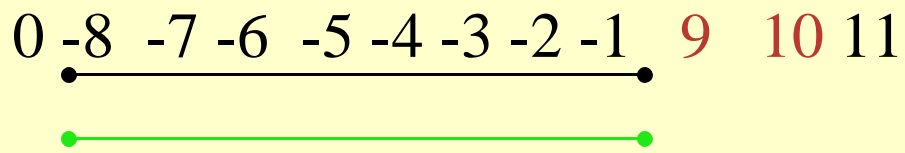


Inversions

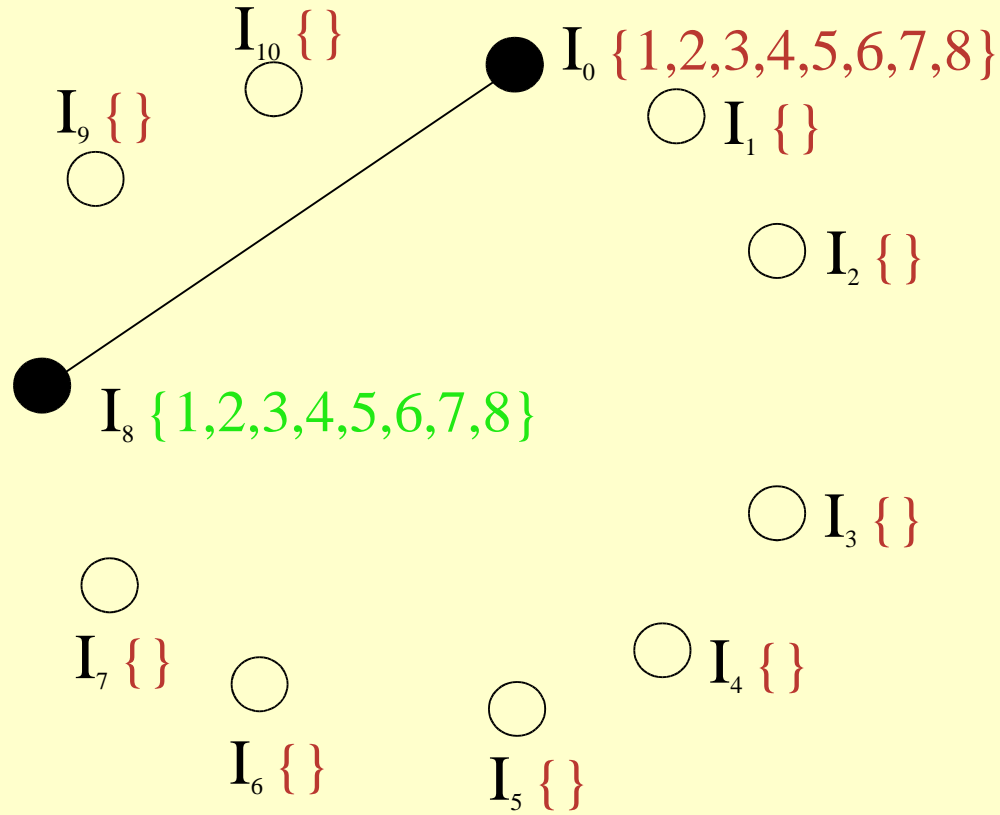
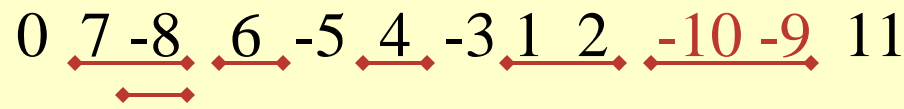


Sorting Chromosome 17

Elementary intervals

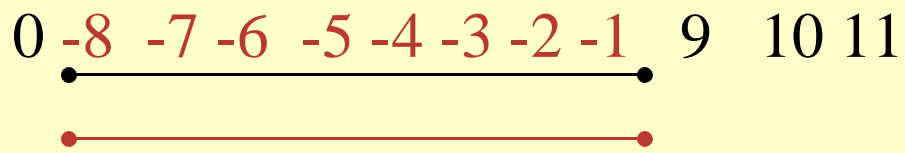


Inversions

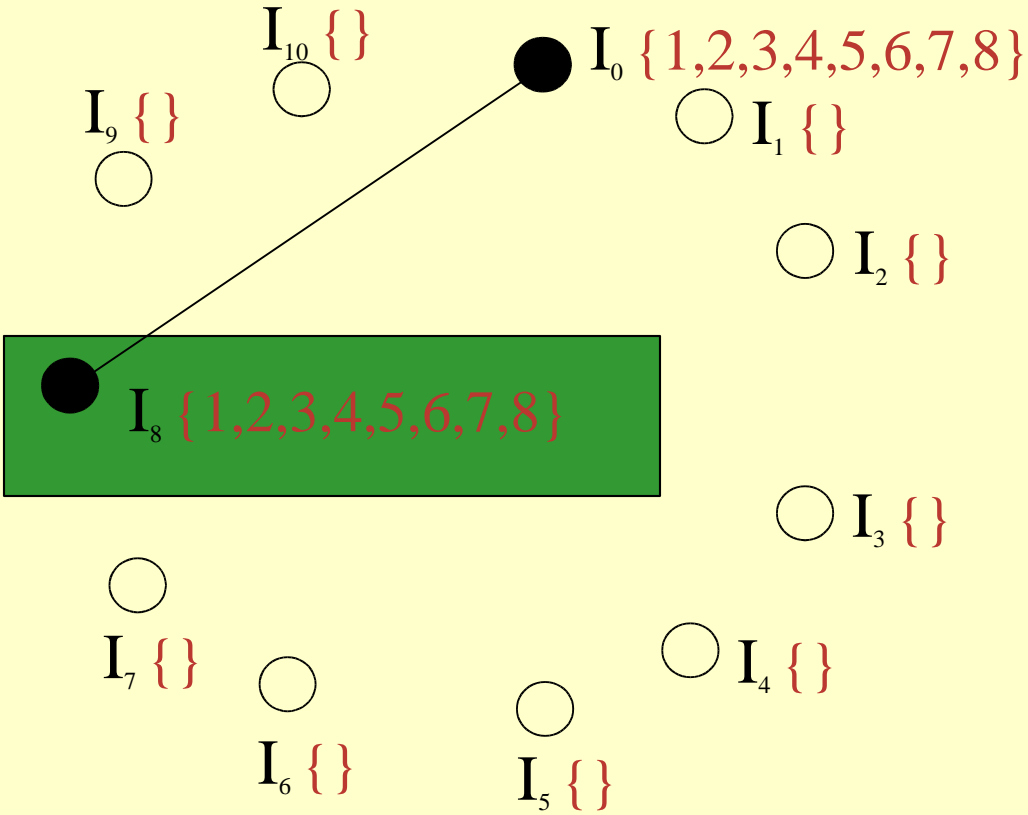
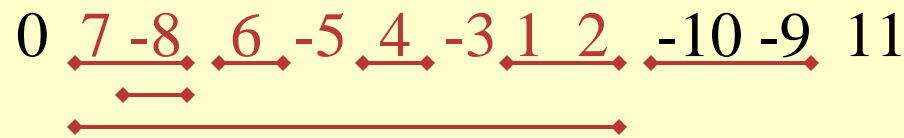


Sorting Chromosome 17

Elementary intervals



Inversions



Sorting Chromosome 17

Elementary intervals

0 1 2 3 4 5 6 7 8 9 10 11

Inversions

