

# Special course in Computer Science: Molecular Computing

## Lecture 11: Gene assembly. The Intermolecular and Intramolecular model

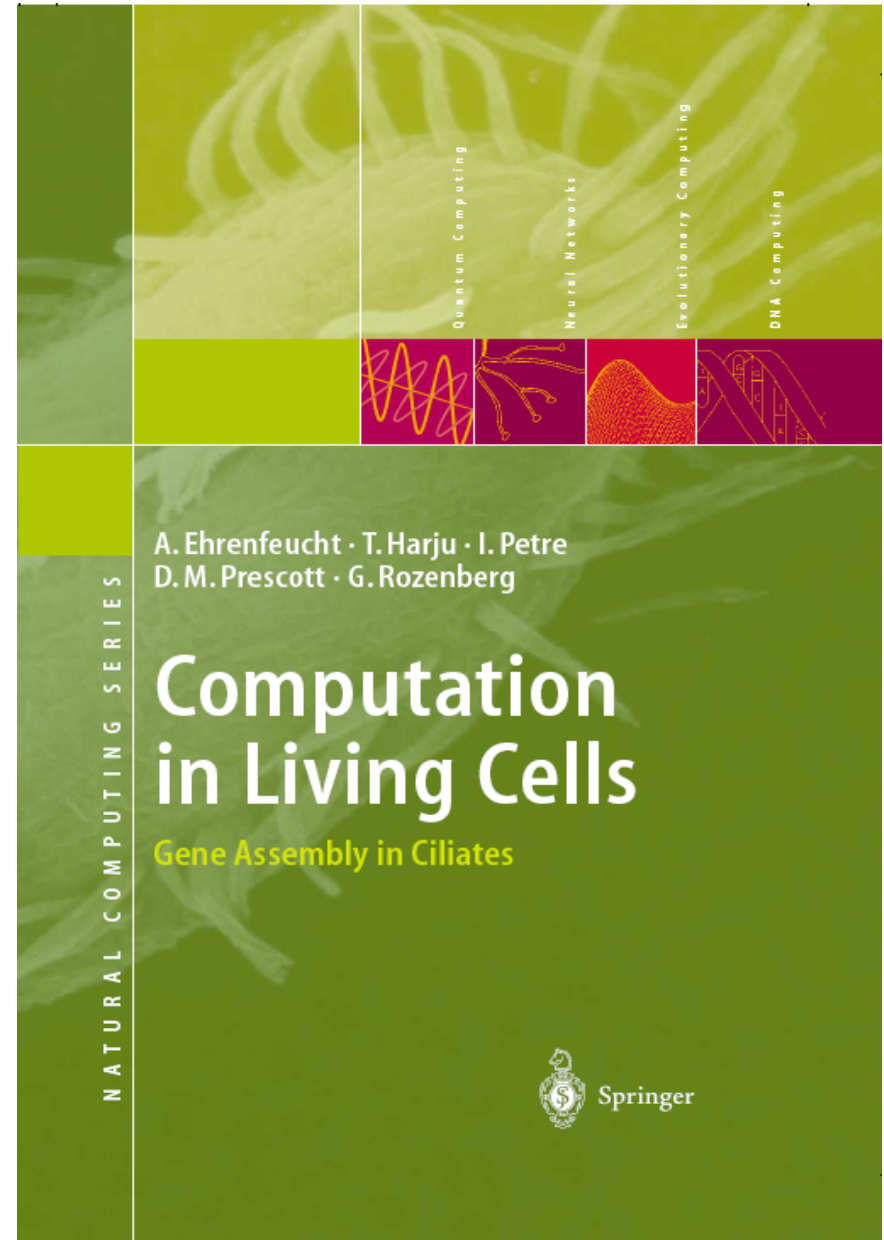
Vladimir Rogojin  
Department of IT, Åbo Akademi  
<http://combio.abo.fi/teaching/special>

Fall 2013

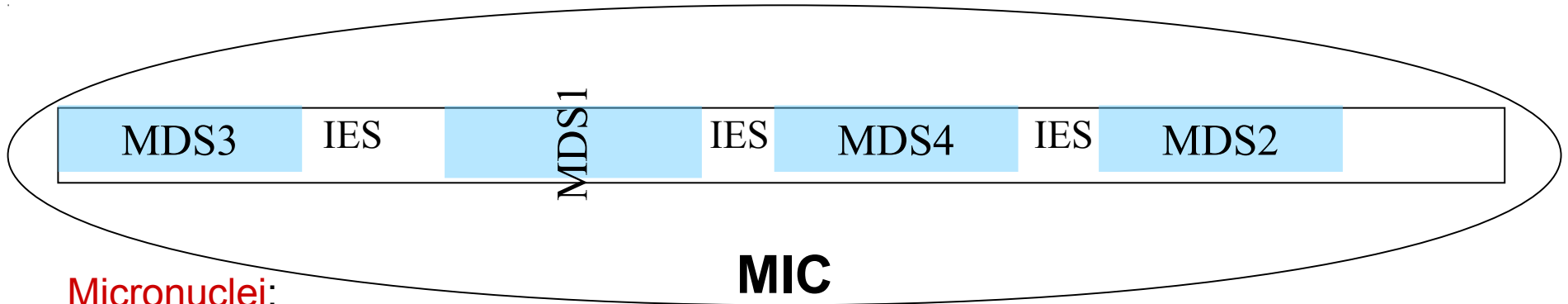
# Gene assembly in ciliates – course book

Lectures 10 – 13,  
recommended reading:

- A. Ehrenfeucht et al.,  
“Computation in Living  
Cells: Gene Assembly  
in Ciliates”, Springer,  
2003



# MIC and MAC

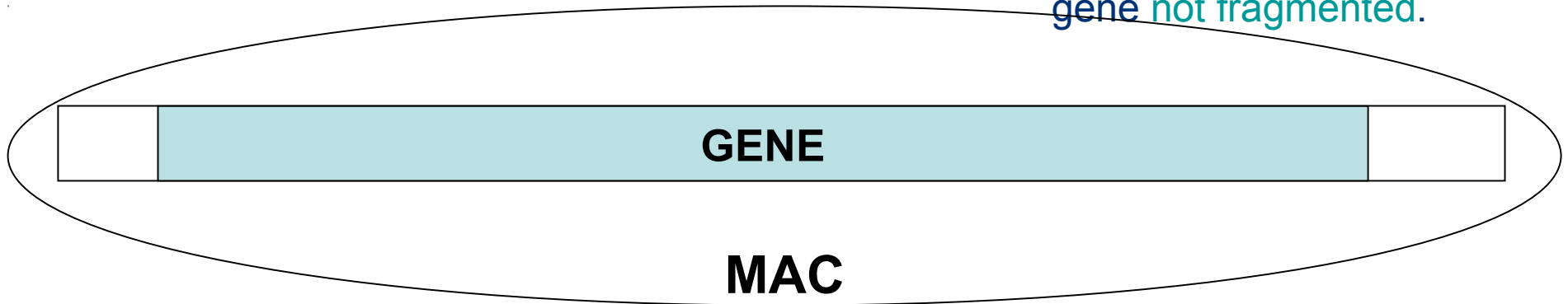


## Micronuclei:

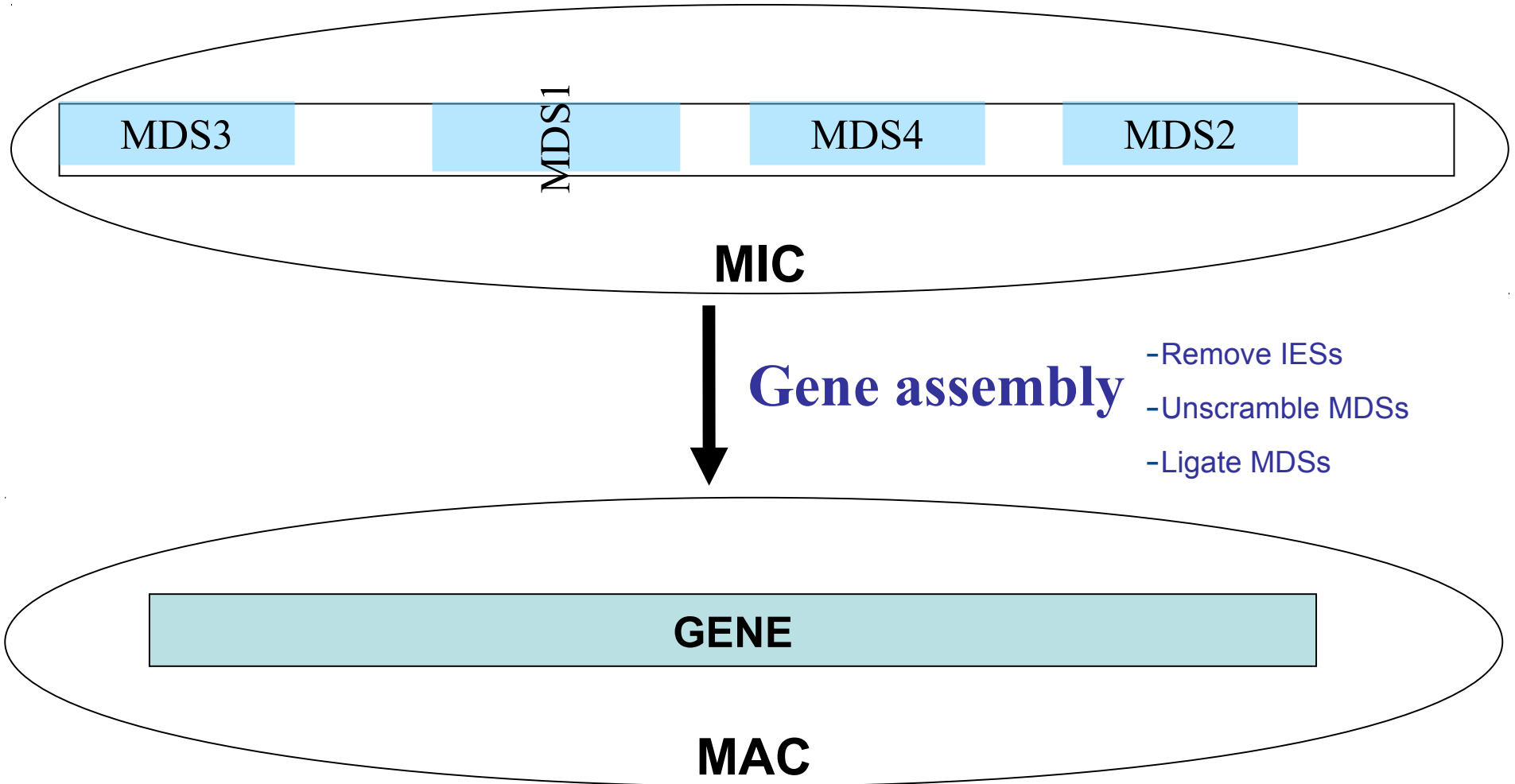
DNA organized on chromosomes,  
One DNA molecule - many genes,  
genes are fragmented,  
fragments (MDS's) are shuffled,  
some fragments are inverted,  
MDSs separated by IESs

## Macronuclei:

1 DNA - 1 gene,  
gene not fragmented.



# MIC to MAC

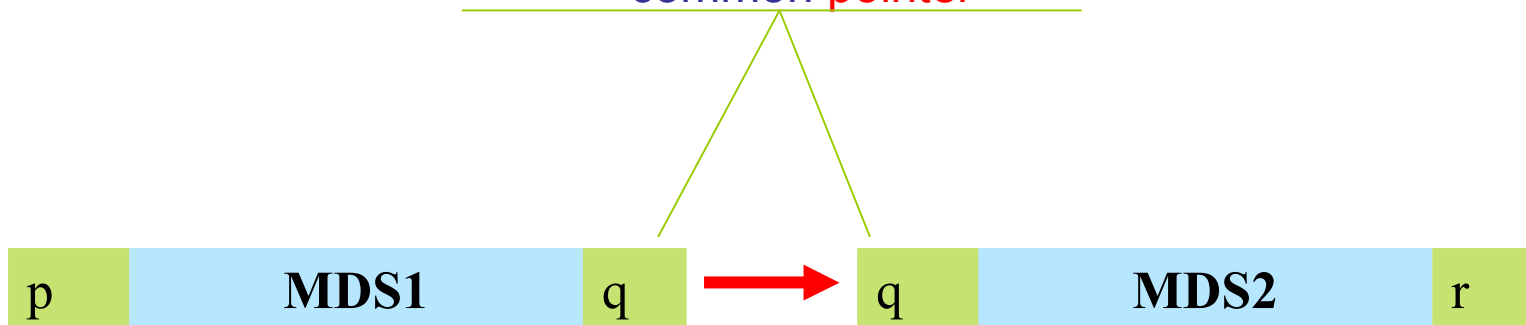


# Pointers

Pointers:

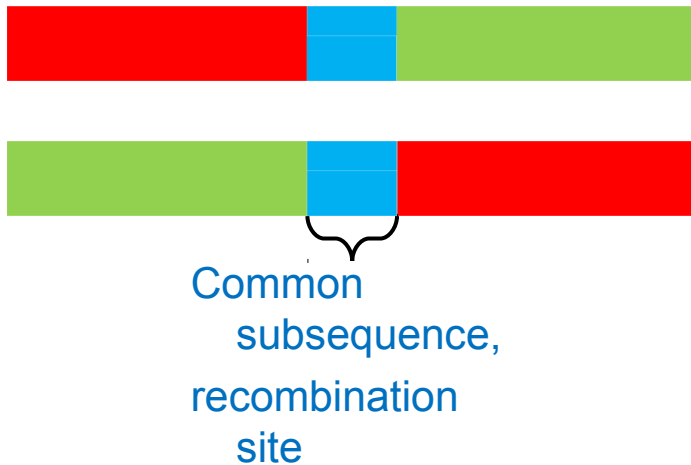
short nucleotide sequences  
on MDS edges

An MDS “points” to the  
next MDS by means of  
common pointer



# Homologous recombination

Two nucleotide sequences recombine at common short subsequence



# Molecular models

- We will overview 2 splicing-based molecular models explaining DNA folding, alignment and recombination during macronuclear gene assembly
  - Intermolecular model: considers DNA rearrangement operations involving two molecules
  - Intramolecular model: considers DNA rearrangement operations within a single molecule
- Also, we will present briefly a theory explaining homologous recombination at pointer sites
  - Template-based recombination: template DNAs help to align short pointer sites

# Intermolecular model

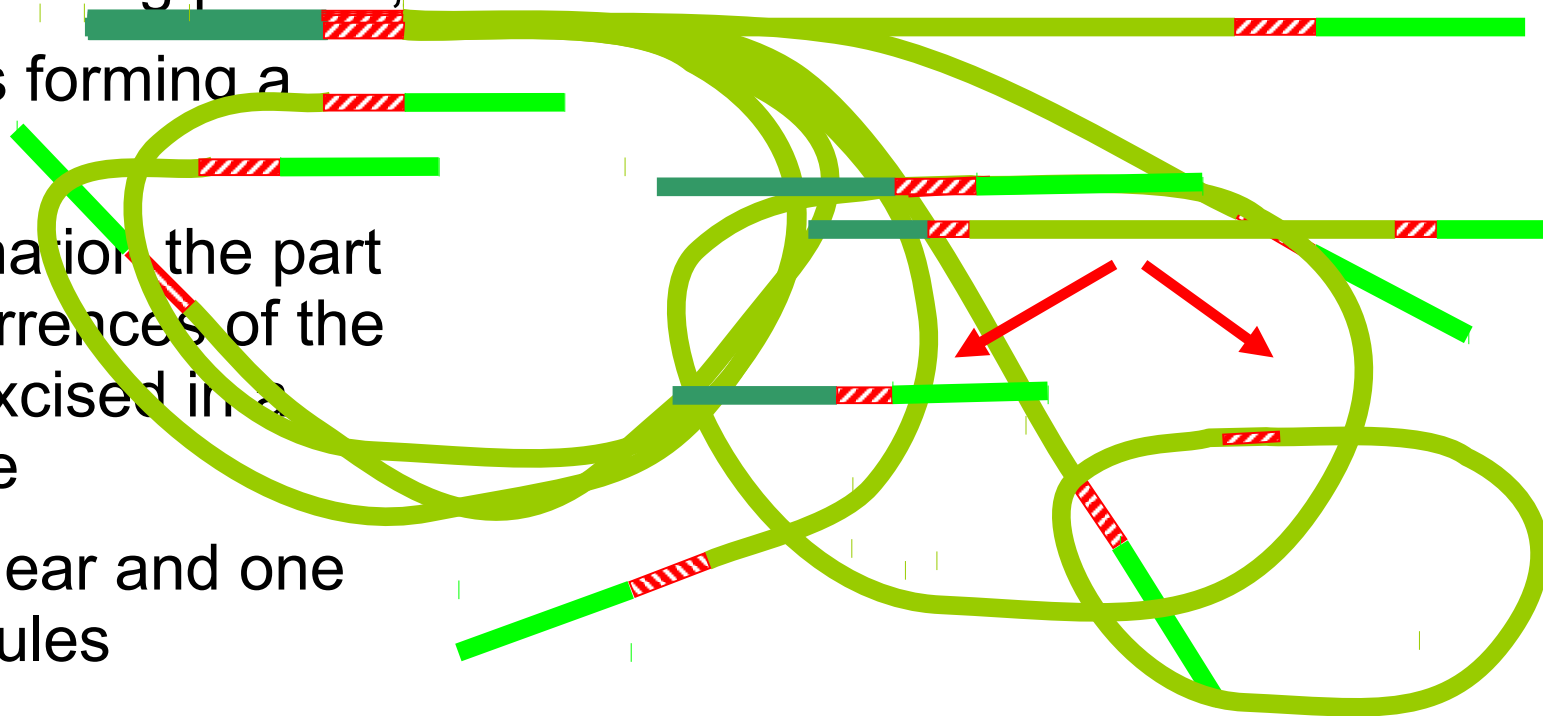
- Intramolecular recombination:

- applies on repeating pointer,

- molecule folds forming a loop,

- after recombination the part between occurrences of the pointer gets excised in the form of a circle

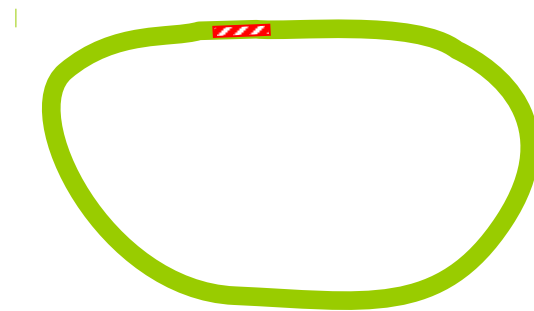
- Result: one linear and one circular molecules





# Intermolecular model

- The intramolecular recombination is reversible – intermolecular recombination:
  - Applied at one linear and at one circular molecules having same pointer
  - The linear and circular molecules get aligned next to each other on their common pointer
  - In the result of the recombination the circular molecule gets inserted into the linear molecule
  - Result: one linear molecule with the nucleotide sequence from the circular molecule



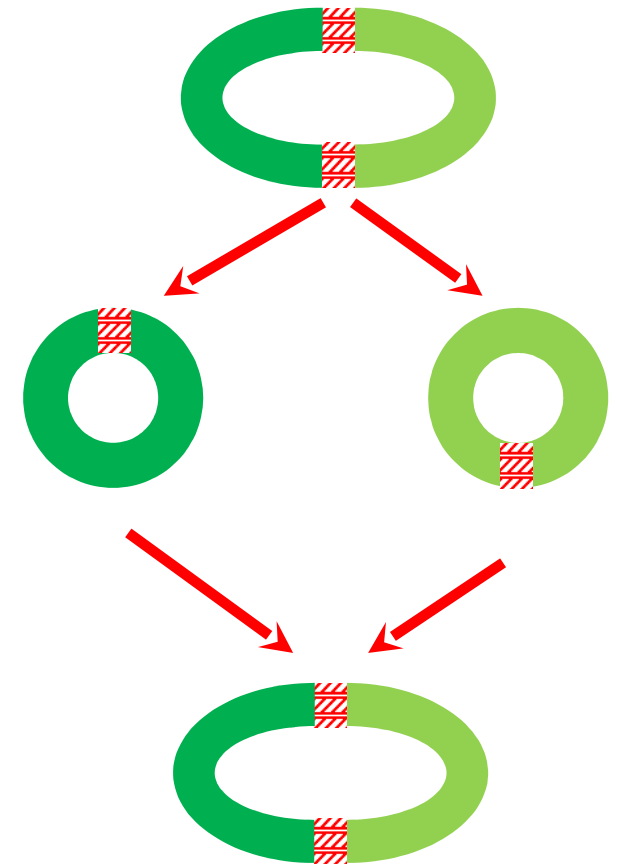
# Intermolecular model

- Circular intramolecular recombination:

- applies on repeating pointer,
- after recombination the part between occurrences of the pointer gets excised in a form of a circle
- Result: two circular molecules

- The inverse to the circular intramolecular recombination – the circular intermolecular recombination

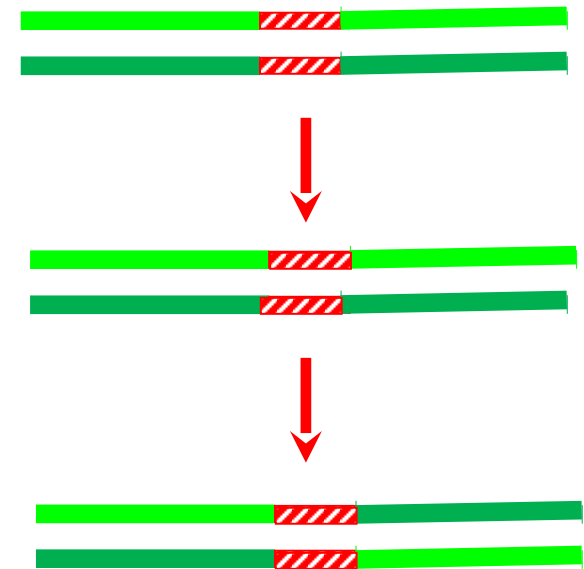
- Occurs between two circular molecules having same pointer
- As the result, a bigger circular molecule is formed with the contents of both molecules separated by the pointer



# Intermolecular model

- Intermolecular recombination: two linear molecules recombine on their common pointer

- Applied on two linear molecules with common pointer
- The result: two linear molecules with interchanged tails

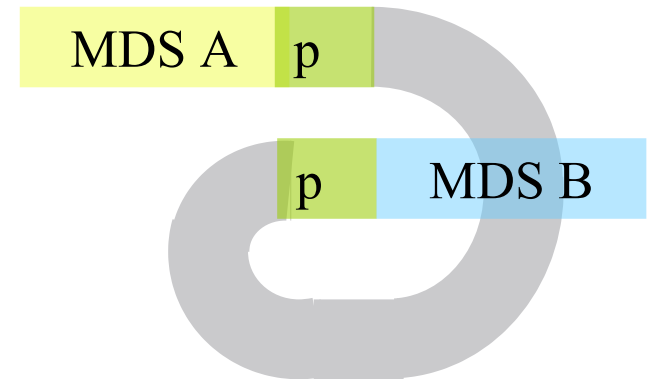


- Note: the operation is self reversible

# Assembling a gene with the intermolecular model

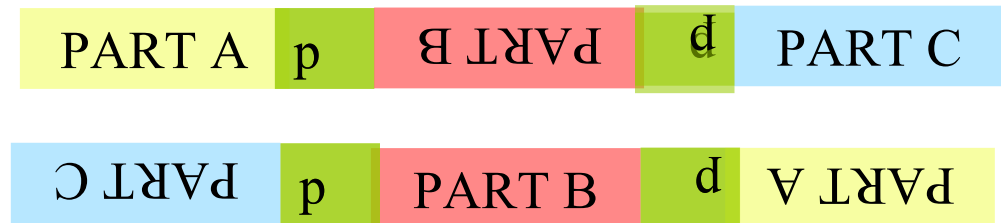
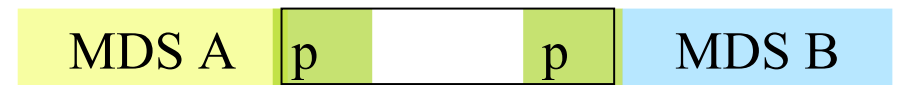
- Excising IES:

- intramolecular recombination, applied on a pointer flanking an IES



- Restoring the orientation of an inverted MDS:

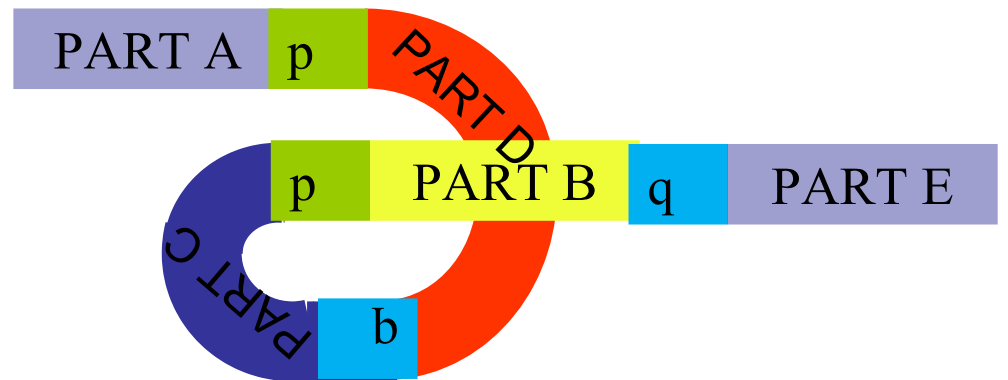
- intermolecular recombination between two copies of a linear molecule, the copies are aligned so that pointers have the same orientation



# Assembling a gene with the intermolecular model

- Sorting MDSs:
  - intramolecular recombination + intermolecular recombination, excision by one of the overlapping pointers and reinsertion by another pointer

# Assembling a gene with the intermolecular model



## •Sorting MDSs:

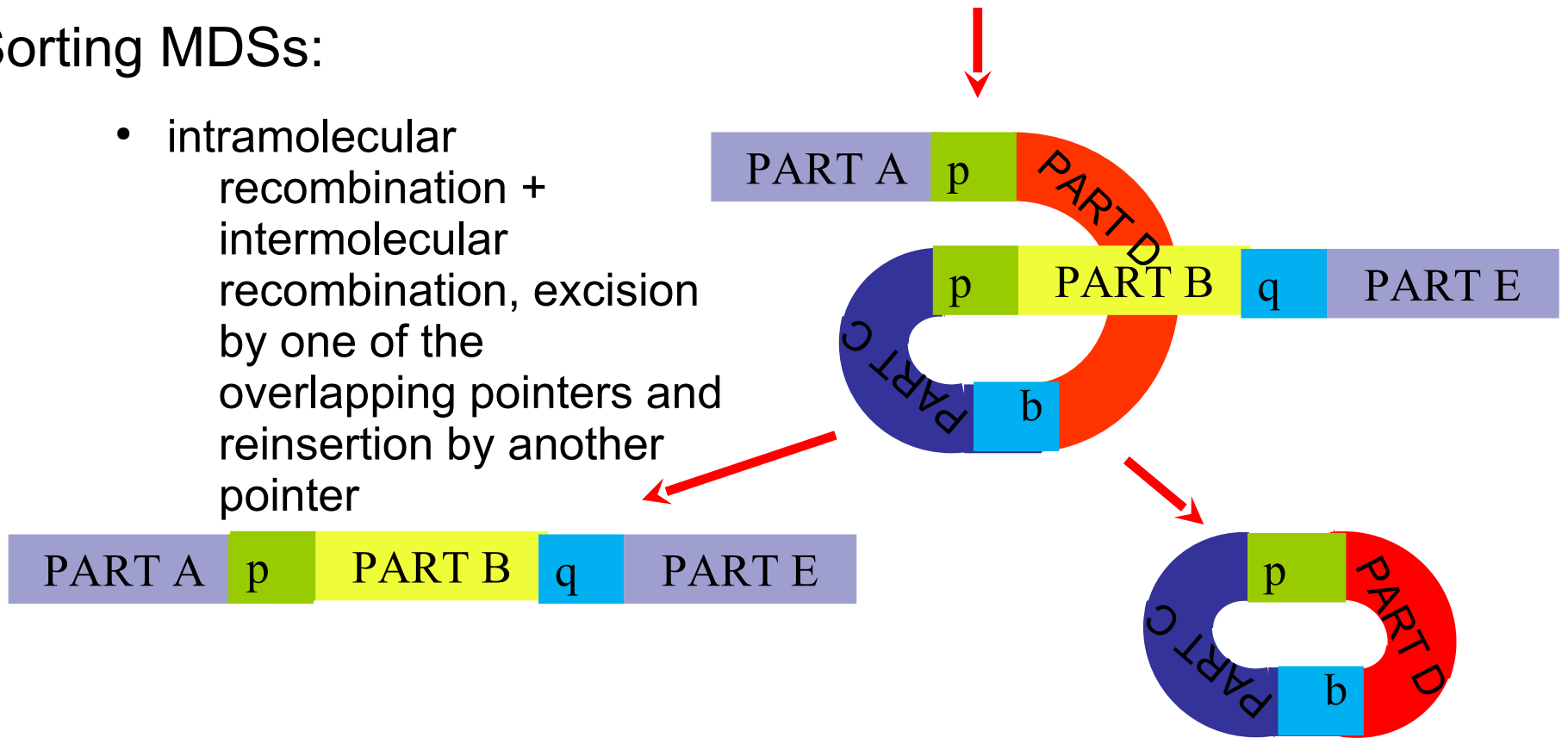
- intramolecular recombination + intermolecular recombination, excision by one of the overlapping pointers and reinsertion by another pointer

# Assembling a gene with the intermolecular model



- Sorting MDSs:

- intramolecular recombination + intermolecular recombination, excision by one of the overlapping pointers and reinsertion by another pointer

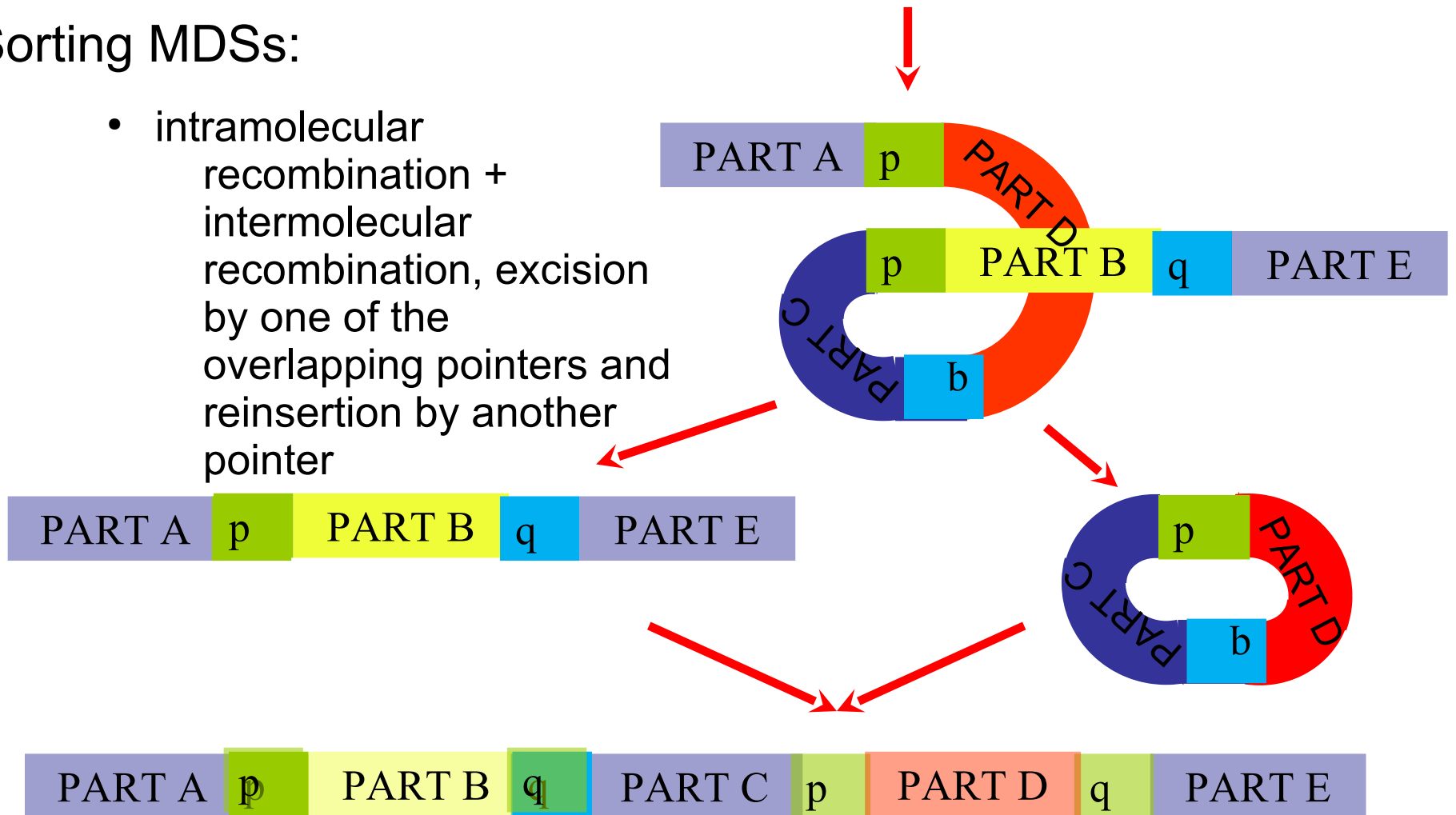


# Assembling a gene with the intermolecular model



- Sorting MDSs:

- intramolecular recombination + intermolecular recombination, excision by one of the overlapping pointers and reinsertion by another pointer





# Assembling a gene with the intermolecular model

- Excising IES:

- intramolecular recombination, applied on a pointer flanking an IES

- Restoring the orientation of an inverted MDS:

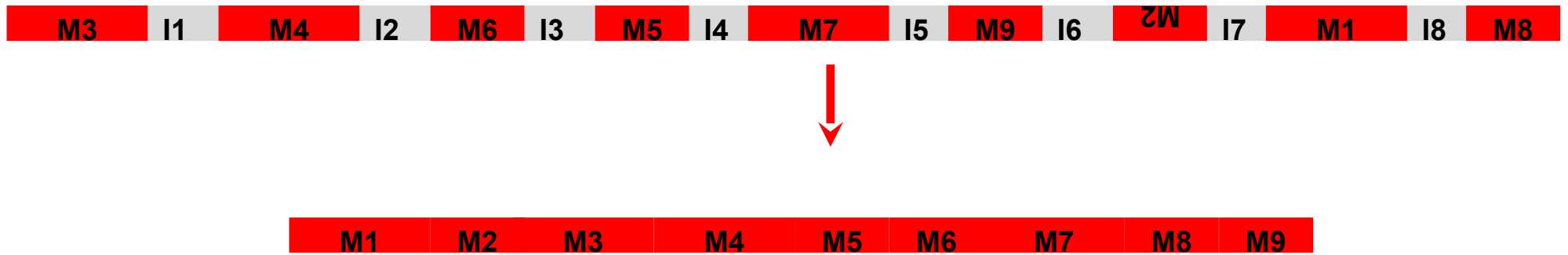
- intermolecular recombination between two copies of a linear molecule, the copies are aligned so that pointers have the same orientation

- Sorting MDSs:

- intramolecular recombination + intermolecular recombination, excision by one of the overlapping pointers and reinsertion by another pointer

# Example

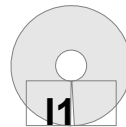
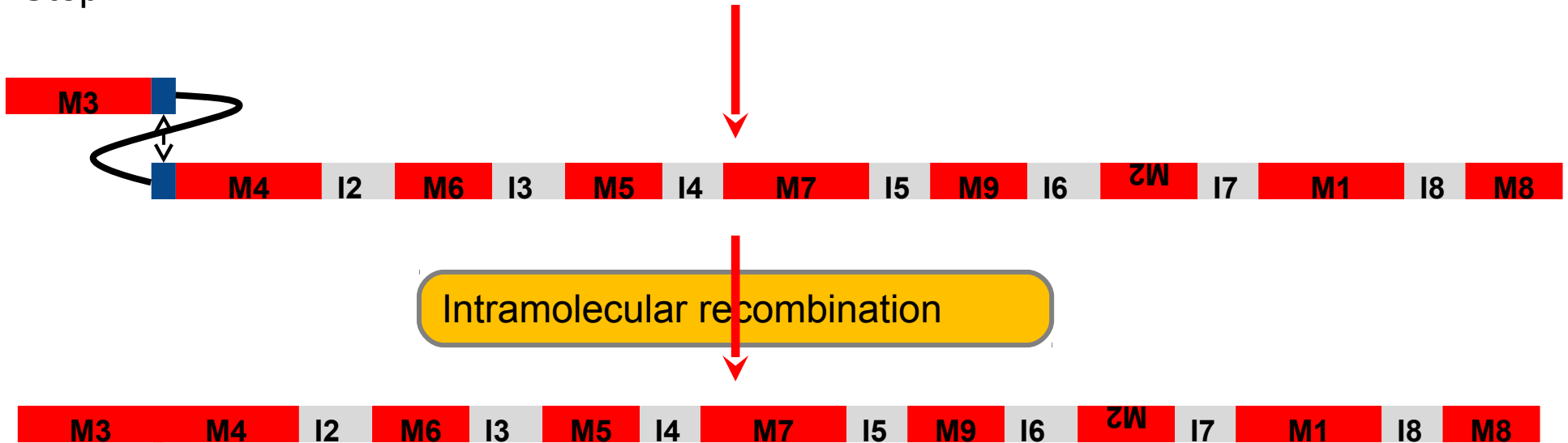
**Example:** assembling gene *actin I* in *S.Nova*



# Example



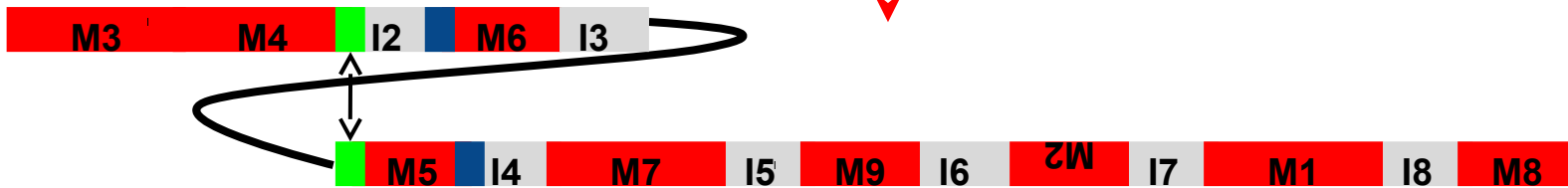
Step 1:



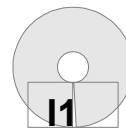
# Example



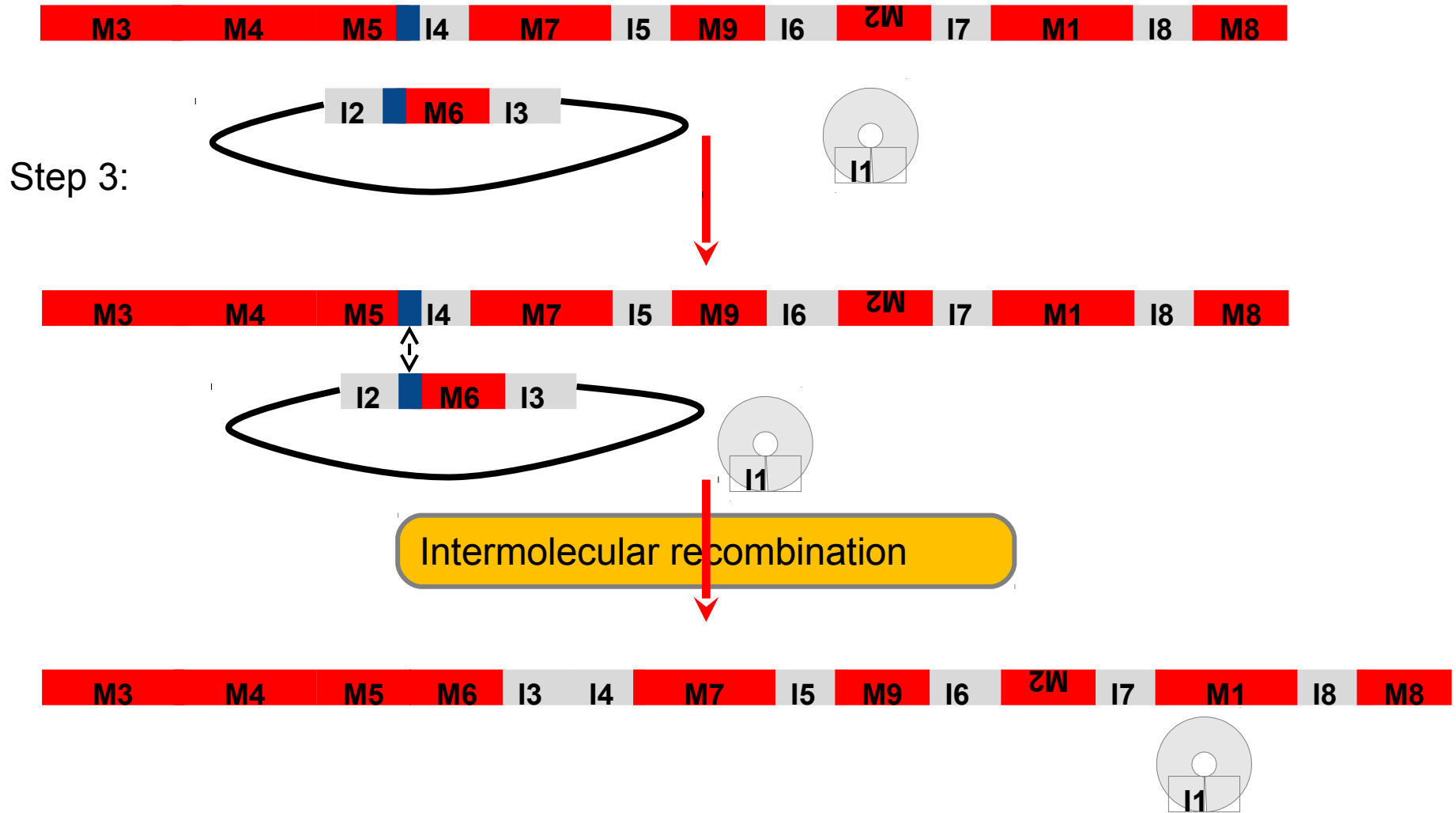
Step 2:



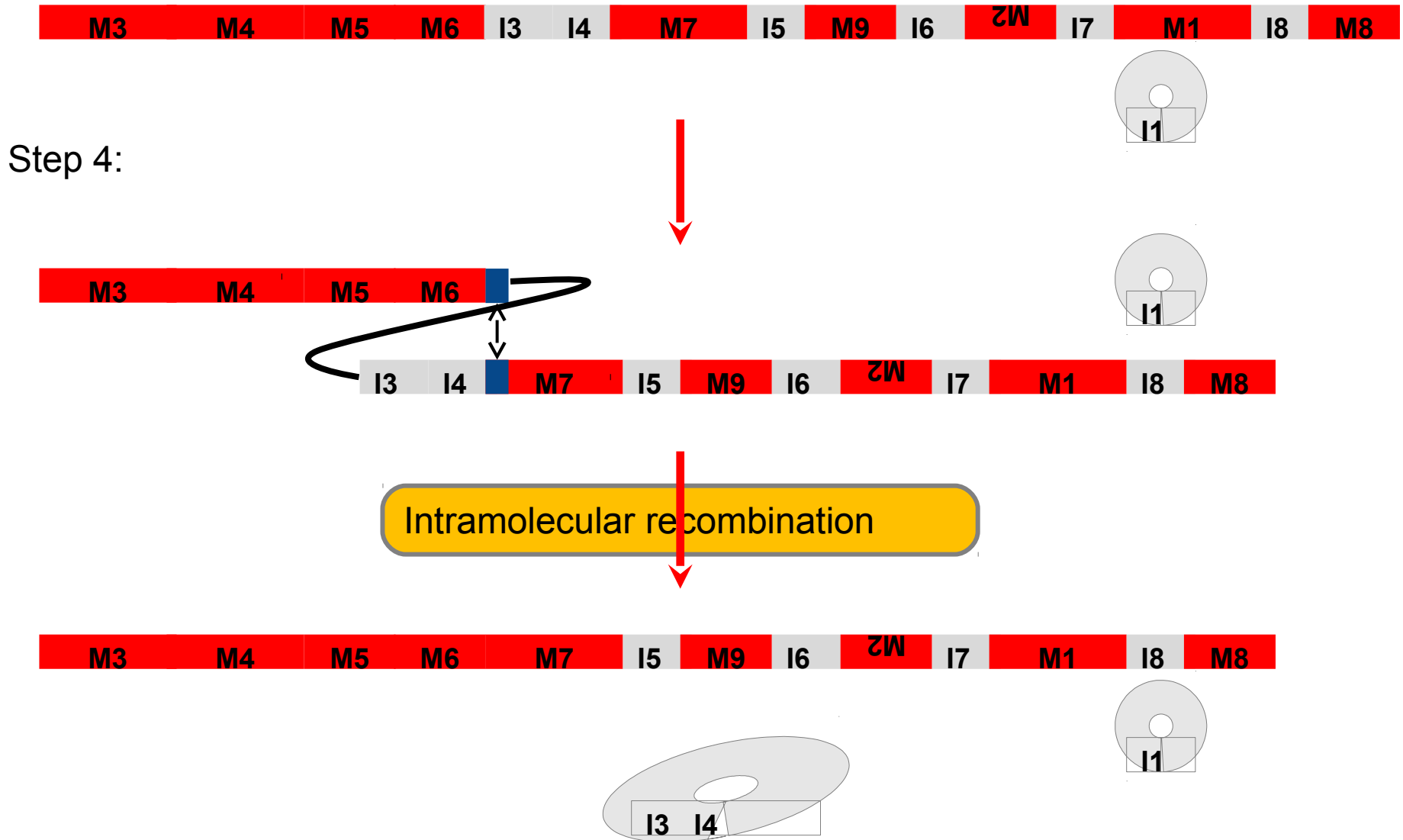
Intramolecular recombination



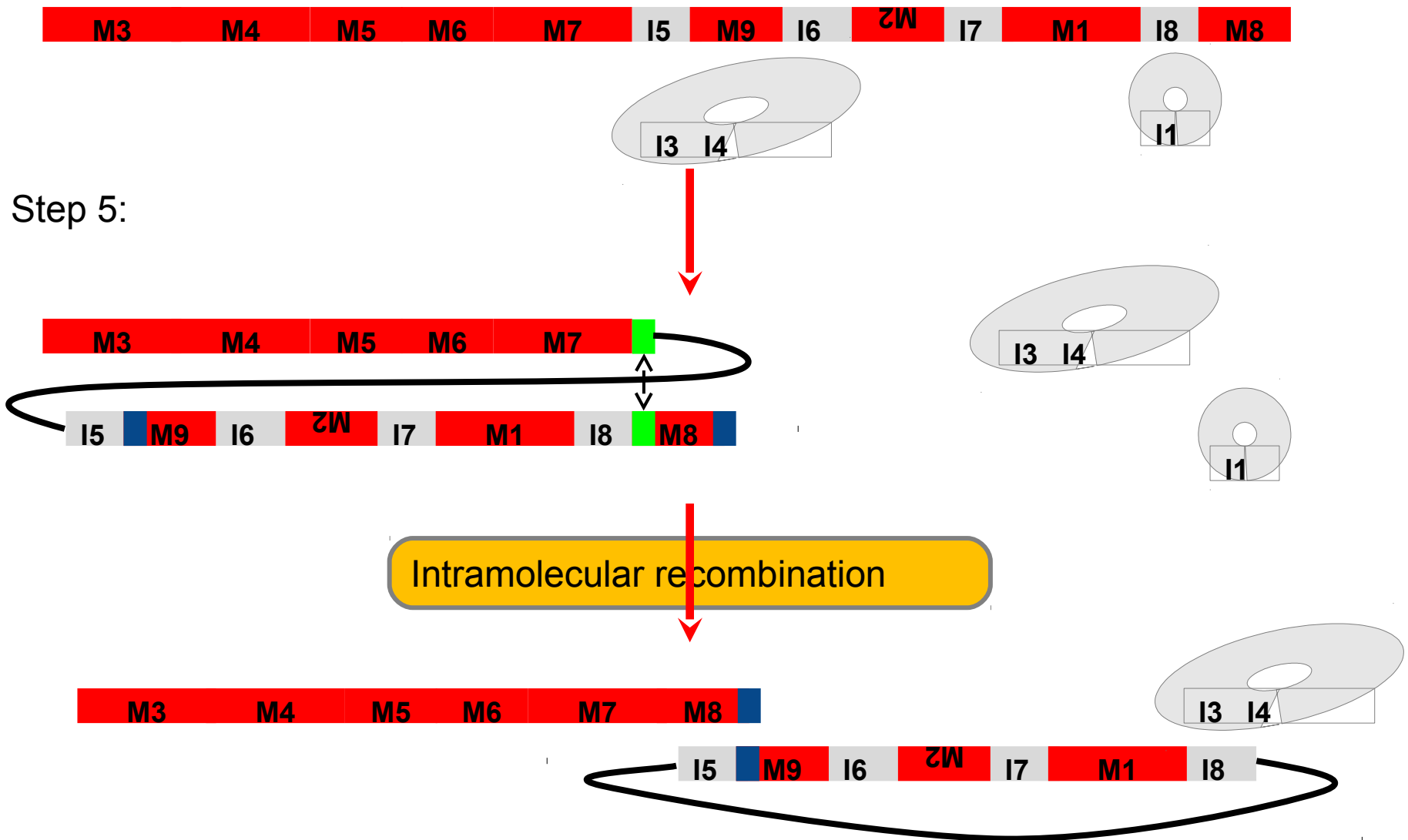
# Example



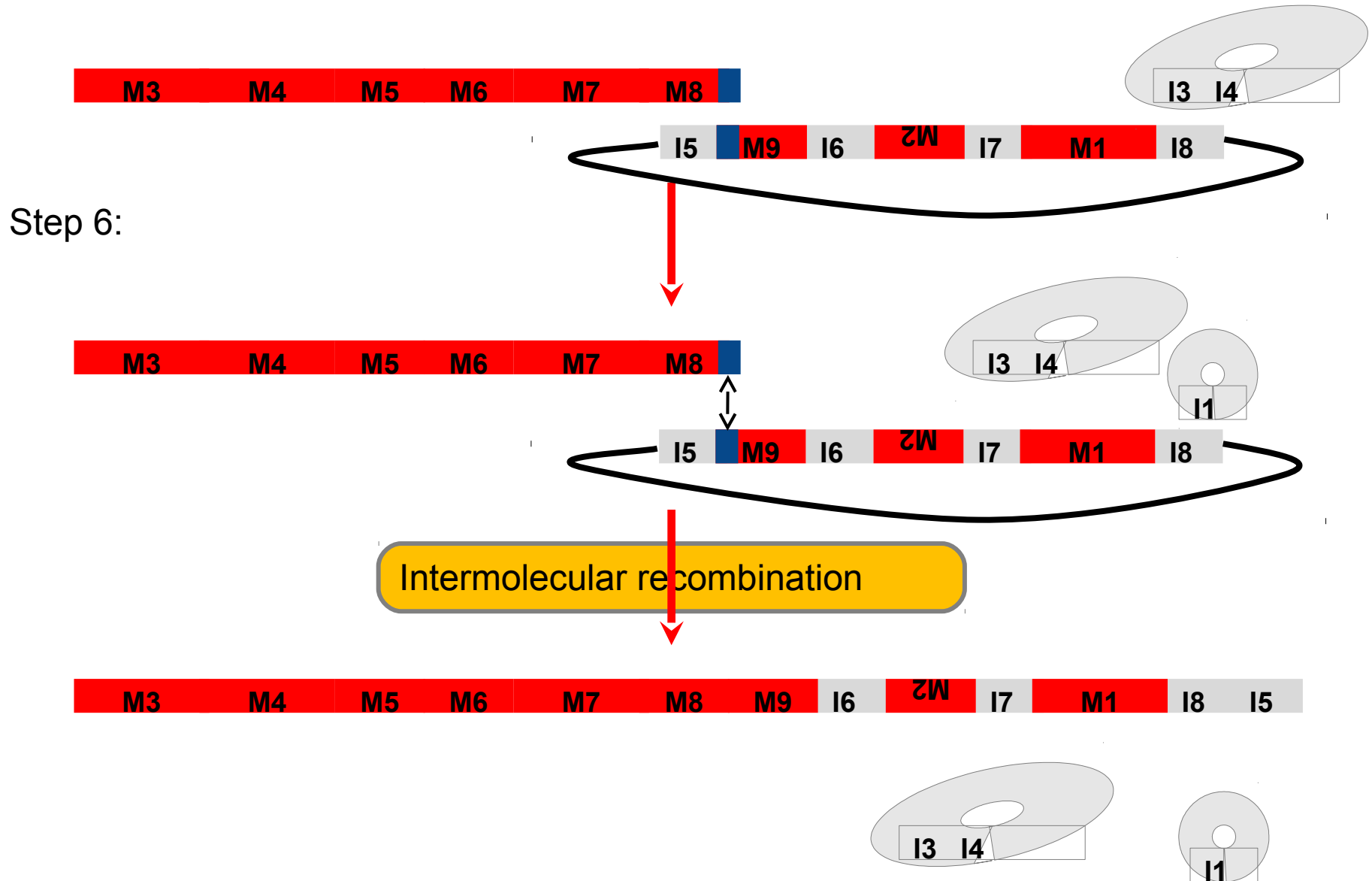
# Example



# Example



# Example

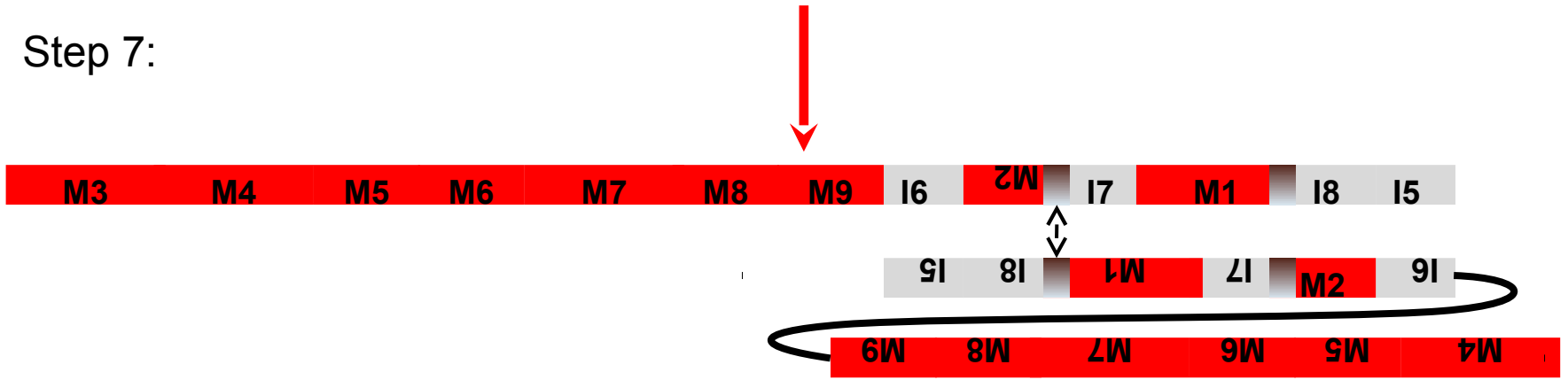




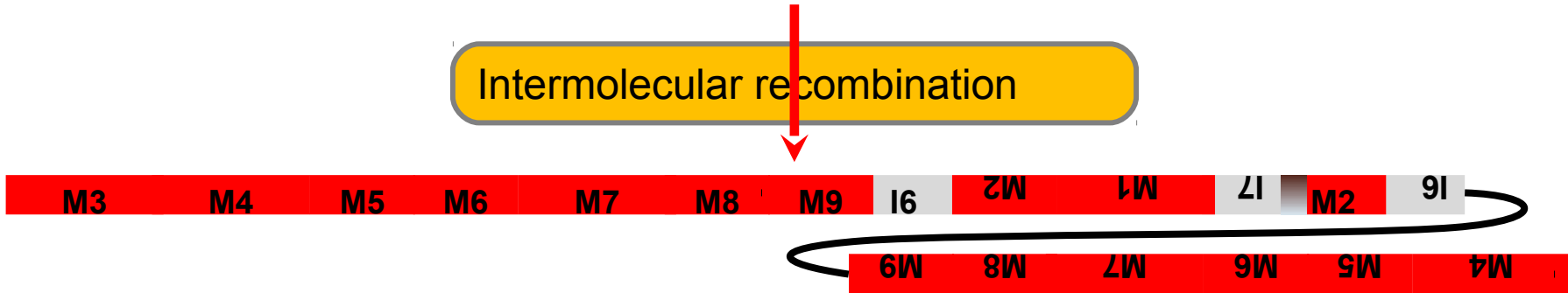
# Example



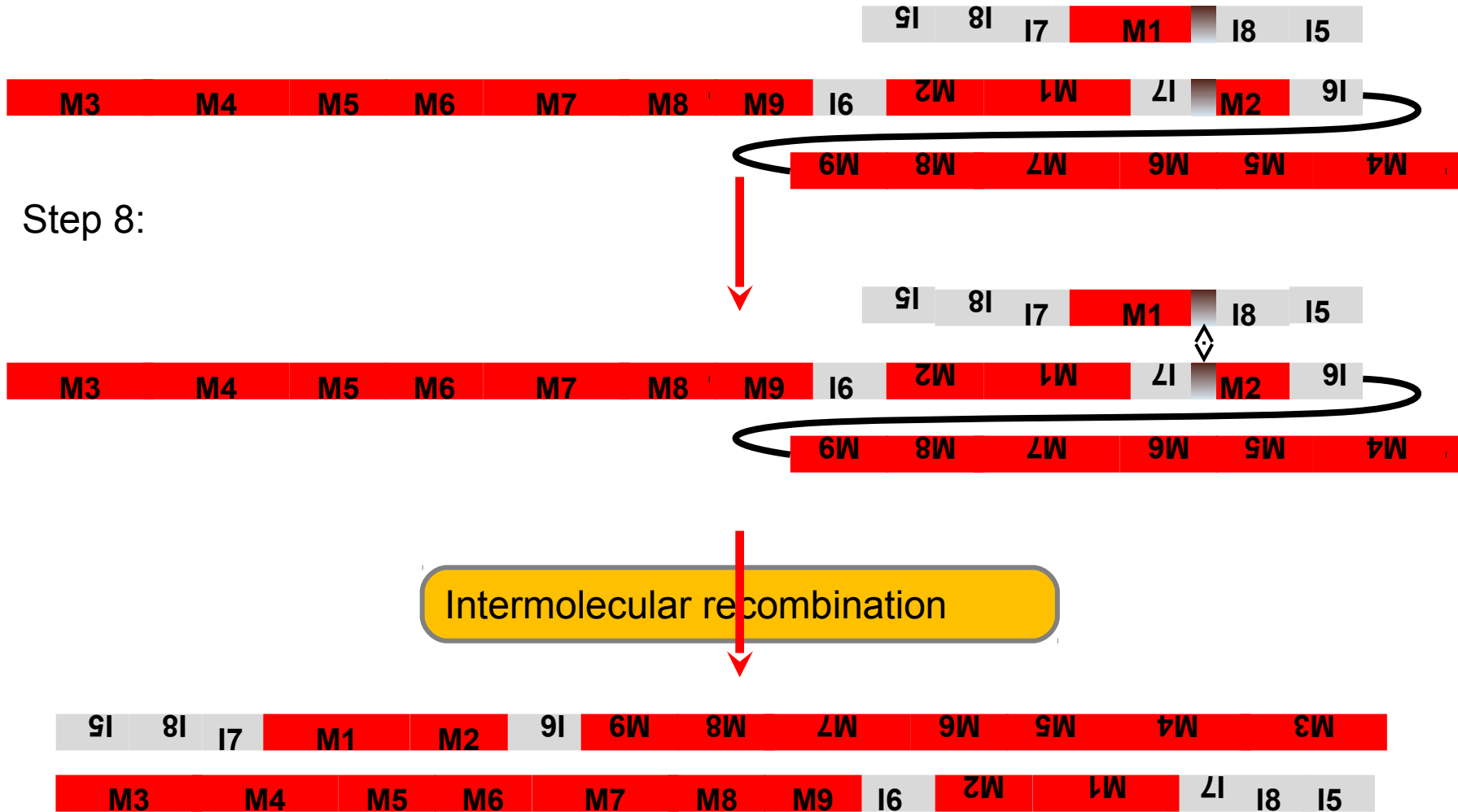
Step 7:



Intermolecular recombination



# Example



# Example



Step 9:



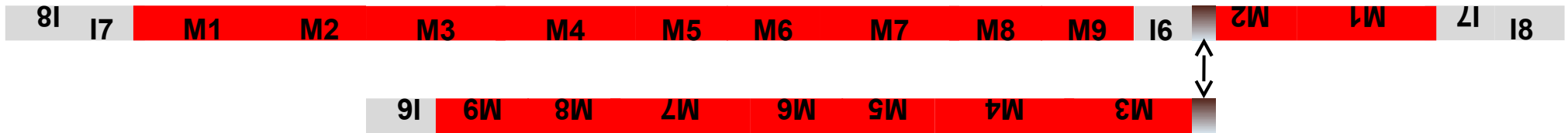
Intermolecular recombination



# Example



Step 10:



Intermolecular recombination



# Gene assembly, recall

- Goals:

- Remove the IESs
- Unscramble (place in the orthodox order) the MDSs
- Ligate the MDSs
- Excise the gene from the chromosome
- Add telomeres
- Multiply the DNA molecule to reach to right number of copies, depending on the species

- The complexity of the problem

- For one single gene, there can be more than 50 MDSs that have to be ordered, spliced, and then excised with precision
- *O.trifallax*: in its whole genome, there are more than 100 000 IESs that have to be excised and more than 100 000 MDSs that have to be ordered
  - Amazing accuracy: more than 98% survival rate after mating

# Intramolecular Model

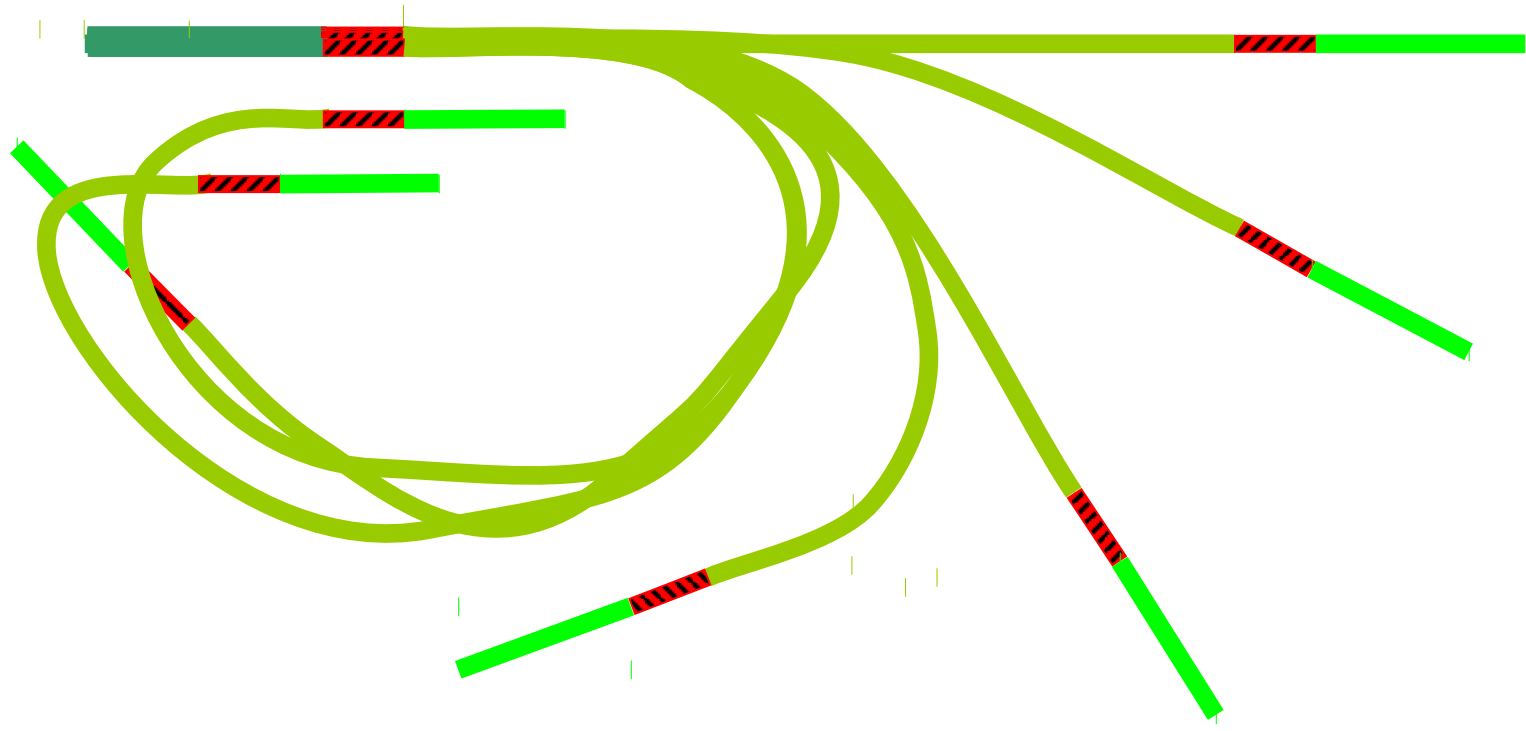
## Molecular models:

- Intermolecular model: considers DNA rearrangement operations involving two molecules
- Intramolecular model: considers DNA rearrangement operations within a single molecule

## Intramolecular model, three operations:

- **Ld-excision:** loop, direct repeat – excision
- **Hi-excision/reinsertion:** hairpin, inverted repeat – excision/reinsertion
- **Dlad-excision/reinsertion:** double loop, alternating direct repeat – excision/reinsertion

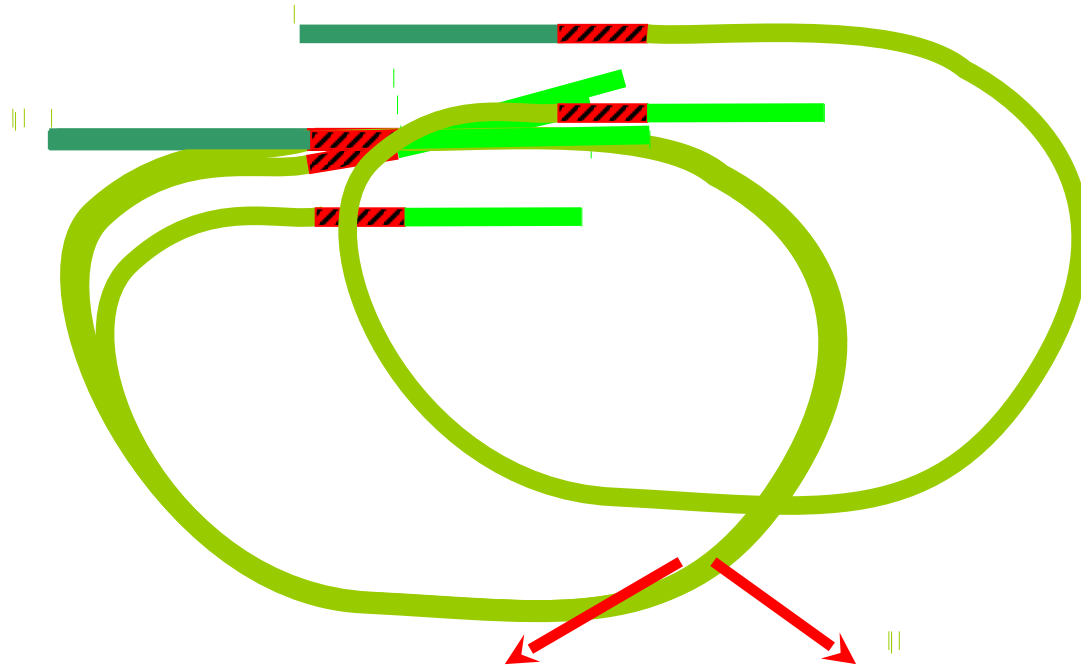
LD



**LD**



**The repeat**

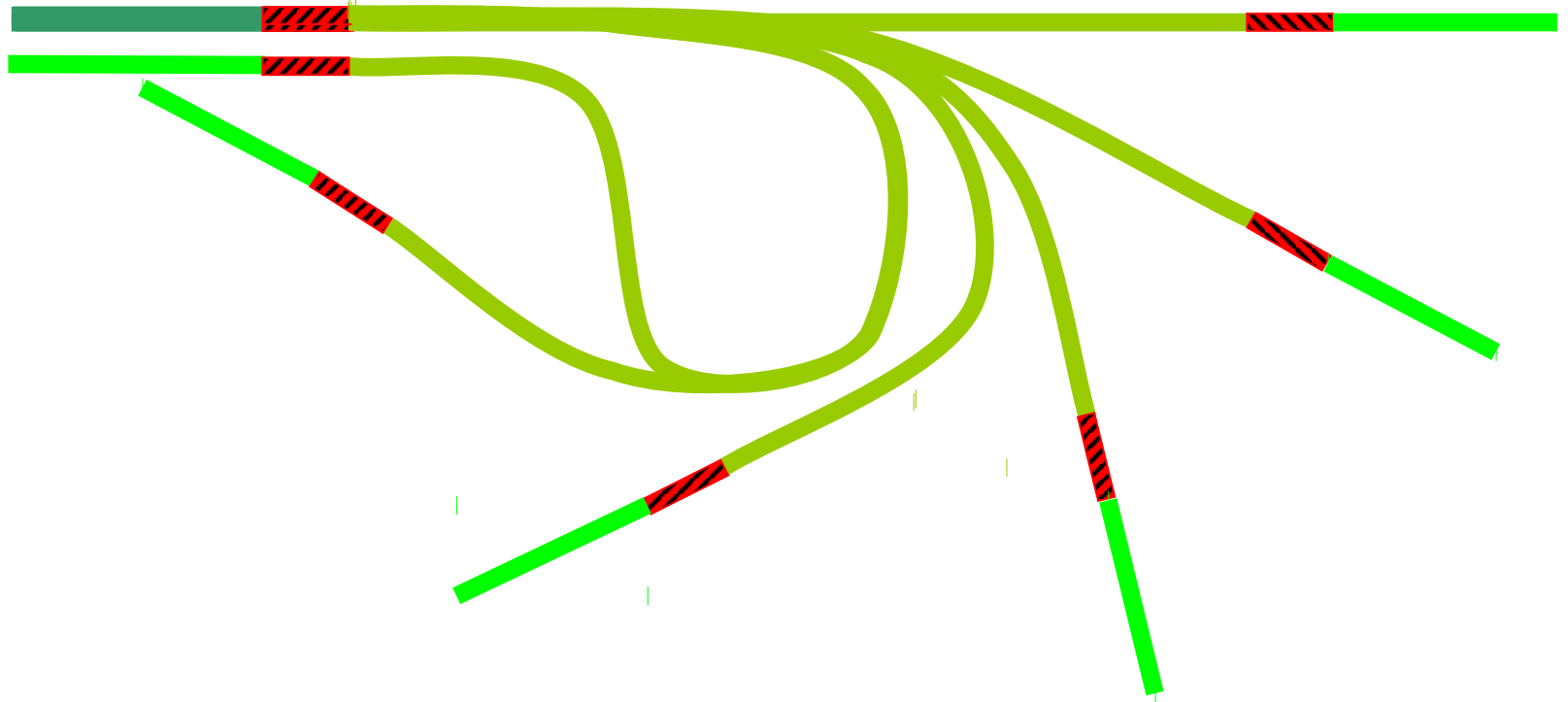


**The fold**

**The result**

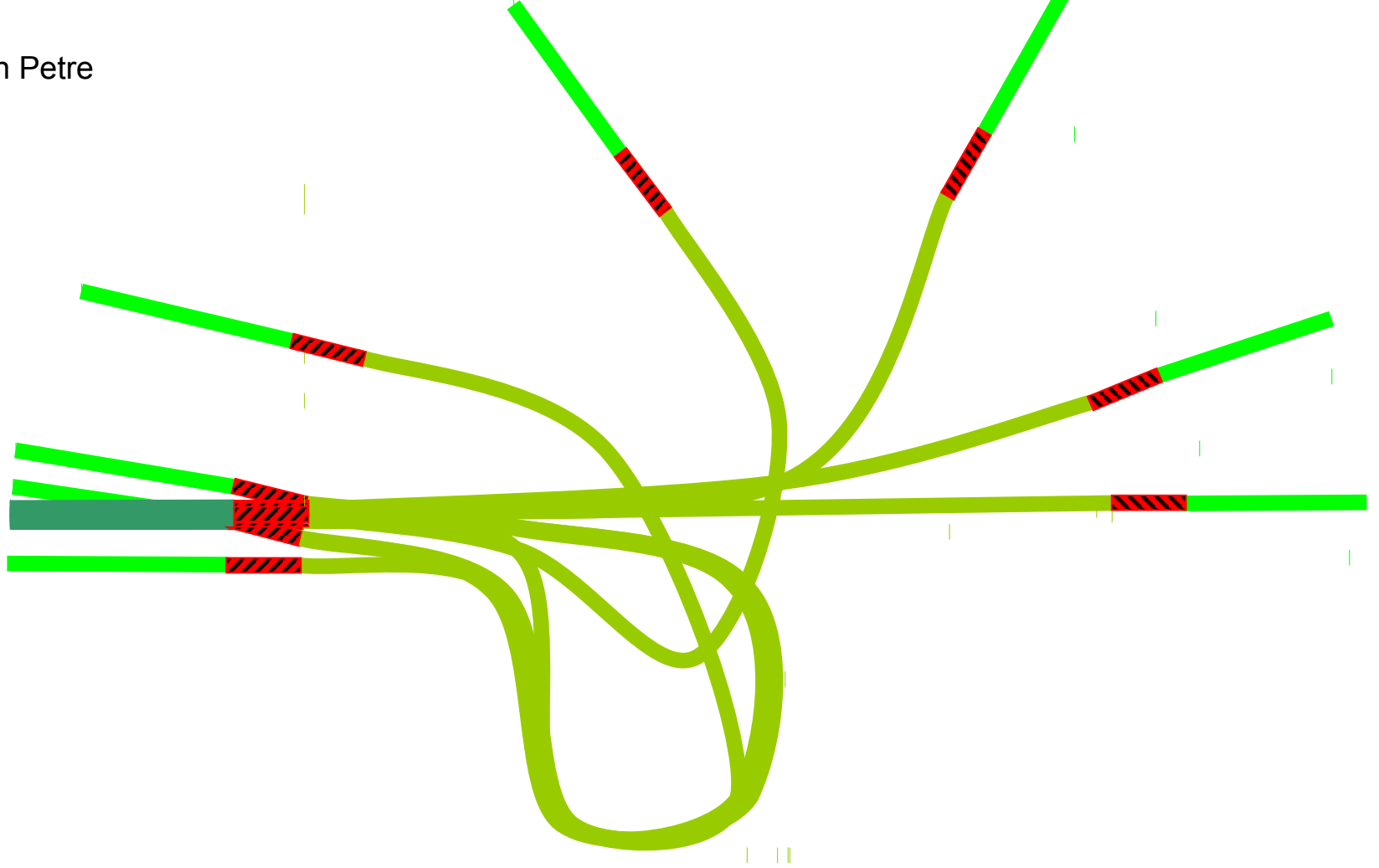


# HI



Courtesy of Ion Petre

**HI**



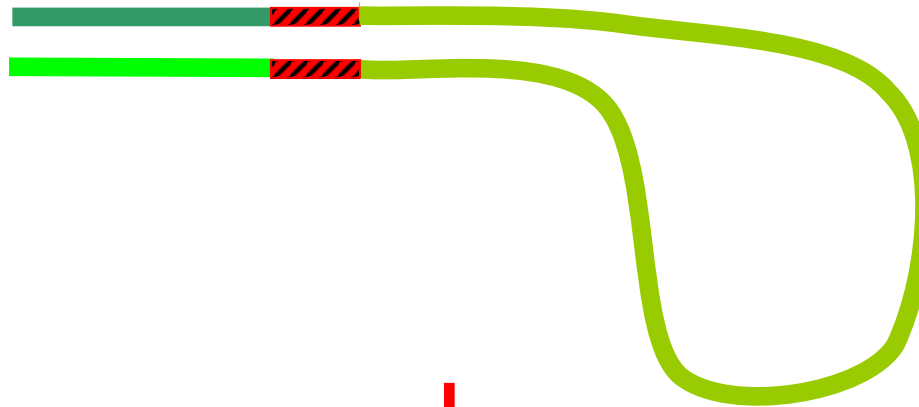
Courtesy of Ion Petre



**The repeat**



**HI**



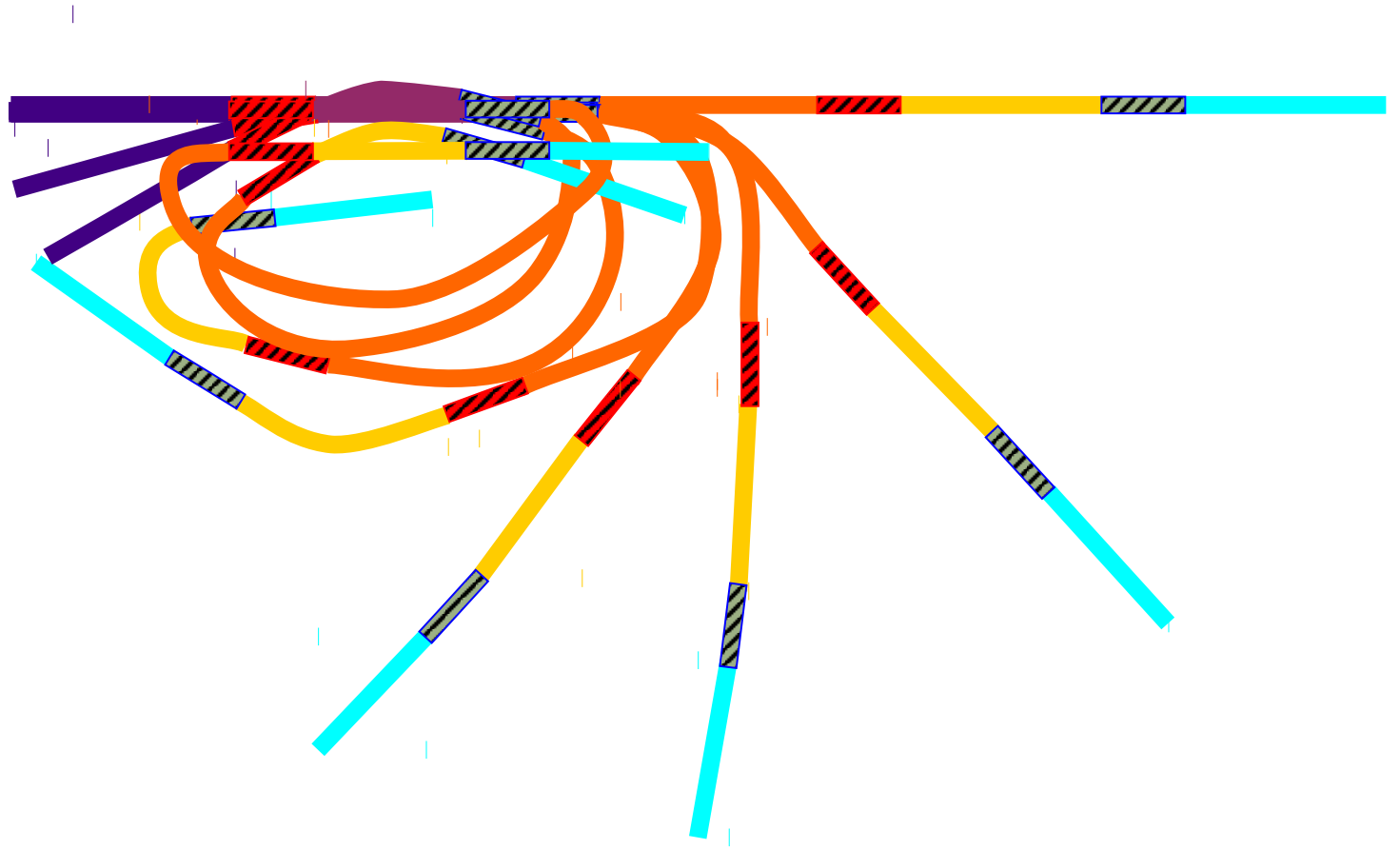
**The fold**



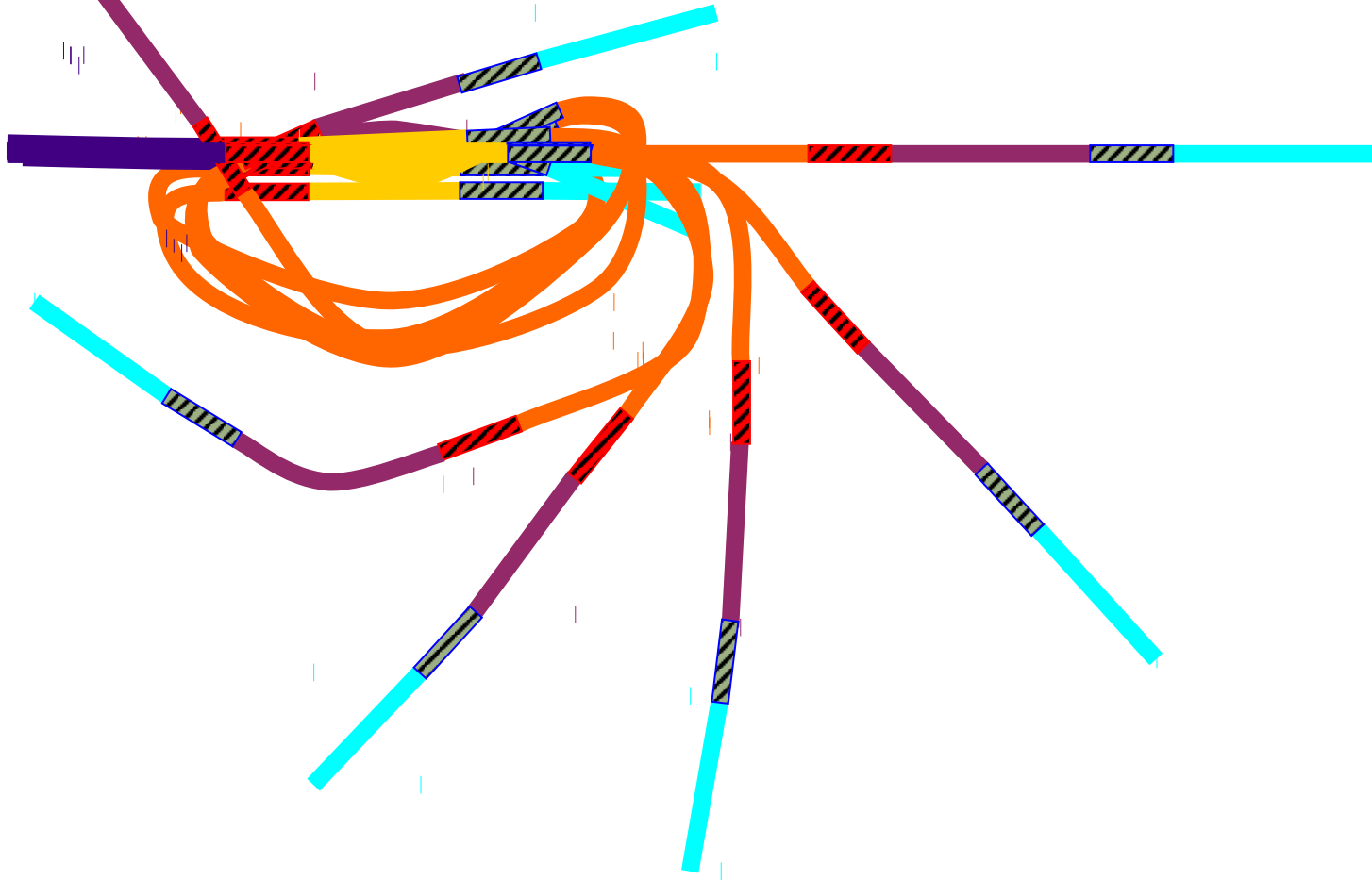
**The result**

Courtesy of Ion Petre

# DLAD



# DLAD



Courtesy of Ion Petre



**The repeat**

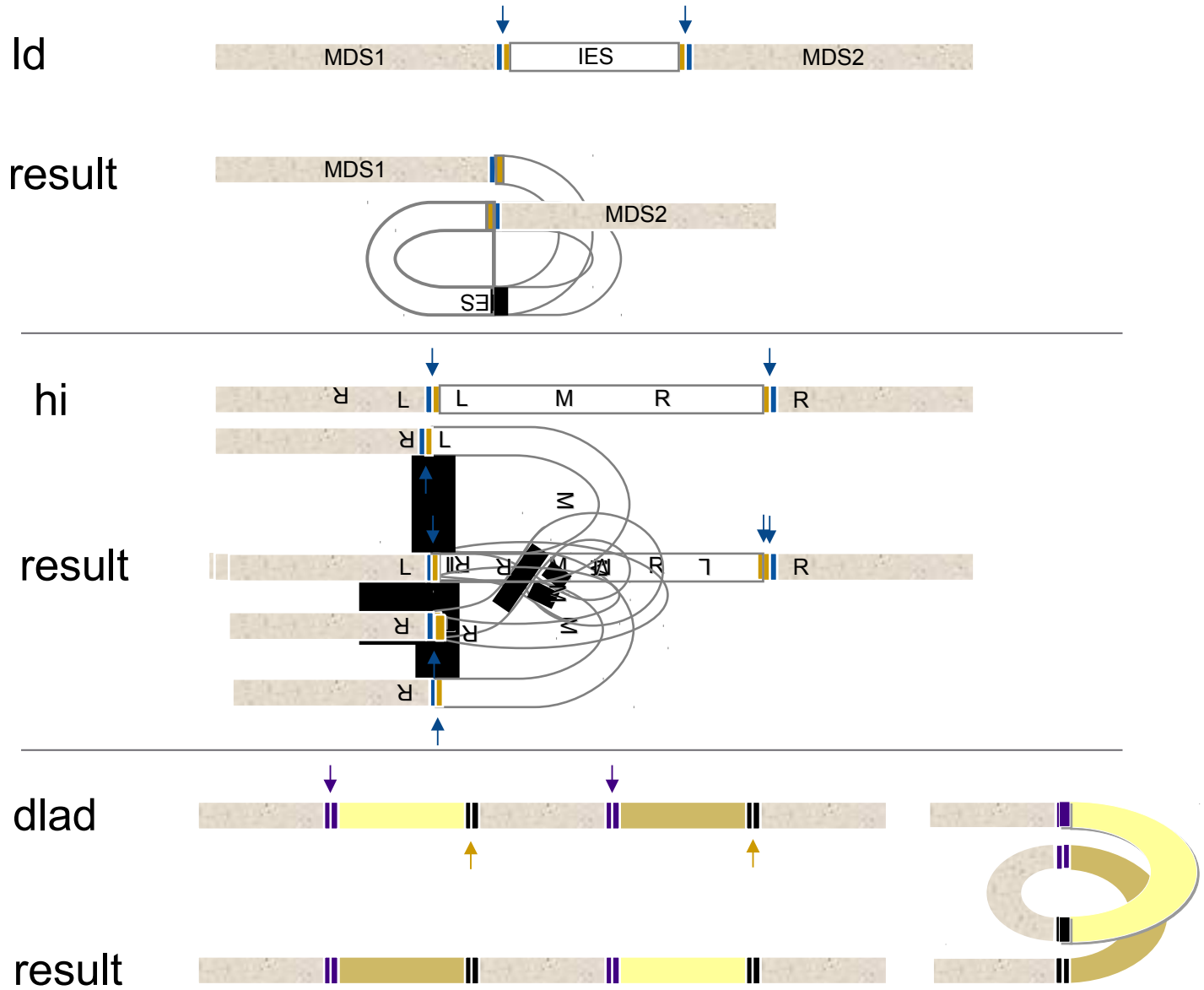
**DLAD**



**The fold**



**The result**

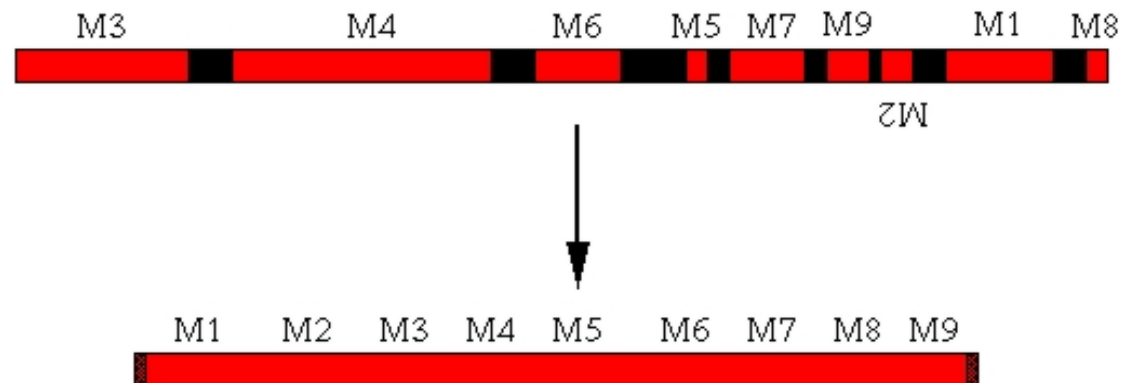


Courtesy of Tseren-Onolt Ishdorj

# Example: assembling actin I gene in S.nova

- Example: assembling gene actin I in S.Nova
  - Red rectangles represent (coding) MDSs and black rectangles represent (noncoding) IESs
  - To simplify, in the following slides, IESs will also be shown as black thick lines

The MIC/MAC form  
of gene *actin I* in *S.Nova*



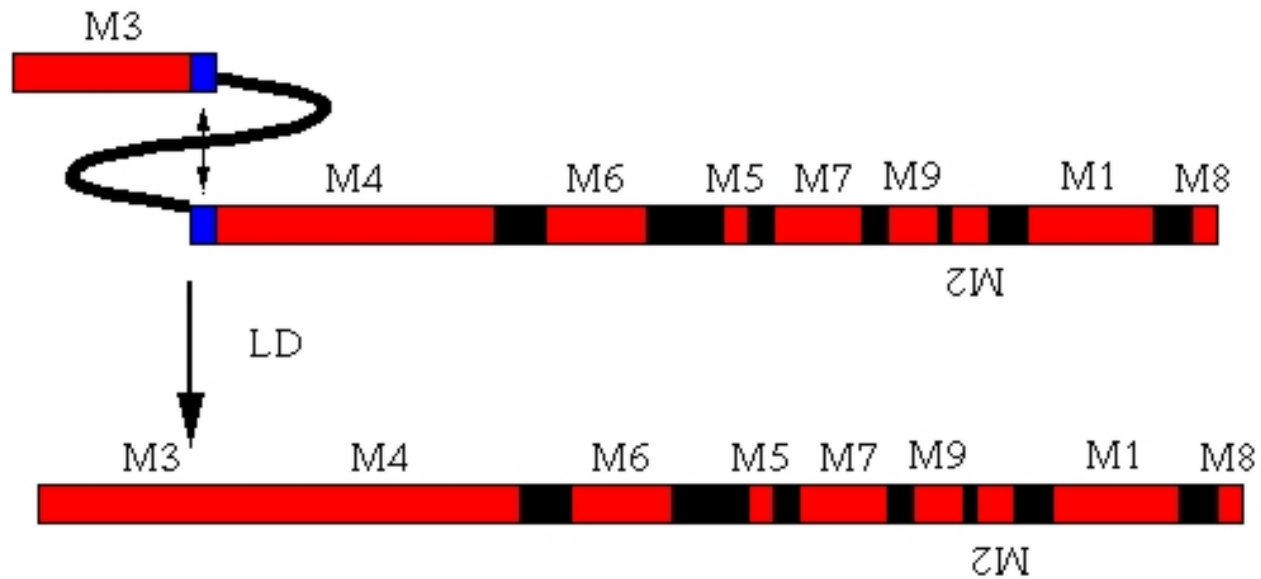


# Example: assembling actin I gene in *S.nova*



Step 1:

LD

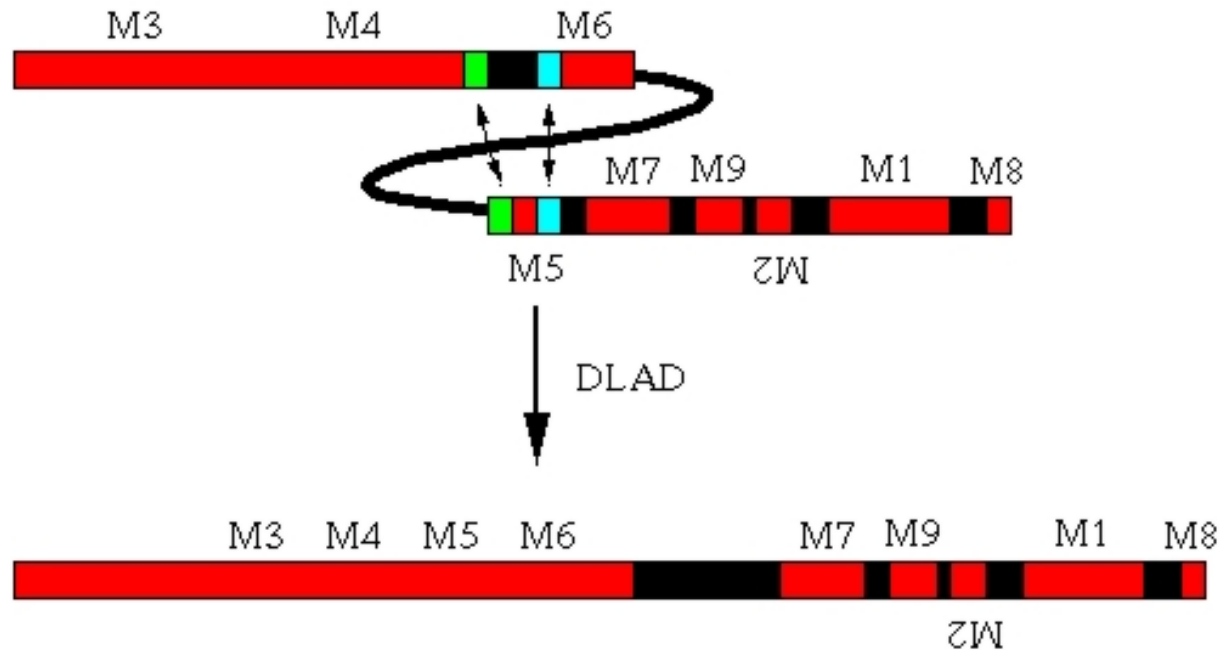


# Example: assembling actin I gene in *S.nova*



Step 2:

**DLAD**

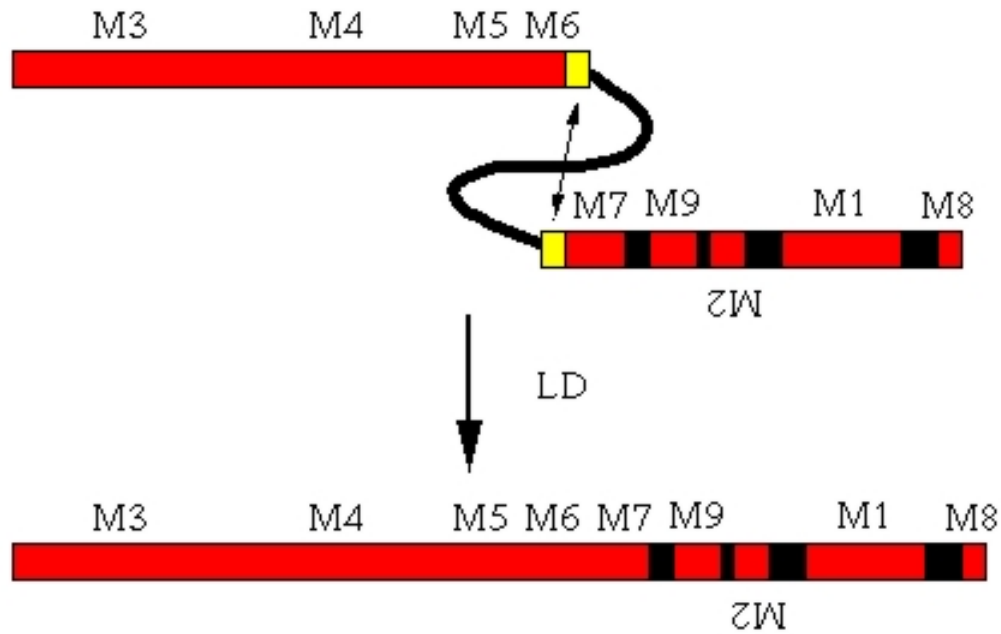


# Example: assembling actin I gene in *S.nova*



Step 3:

LD

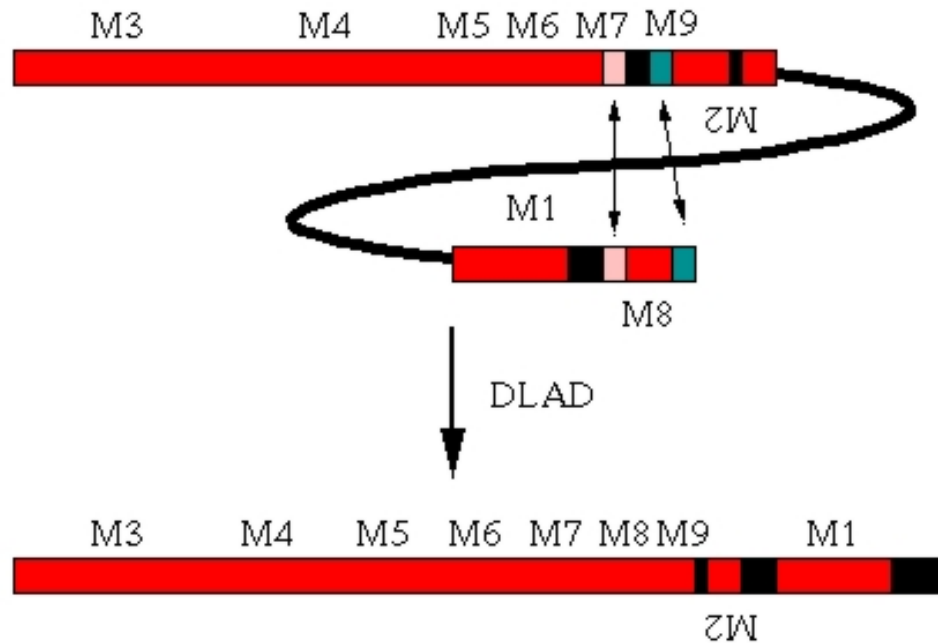


# Example: assembling actin I gene in *S.nova*



Step 4:

**DLAD**

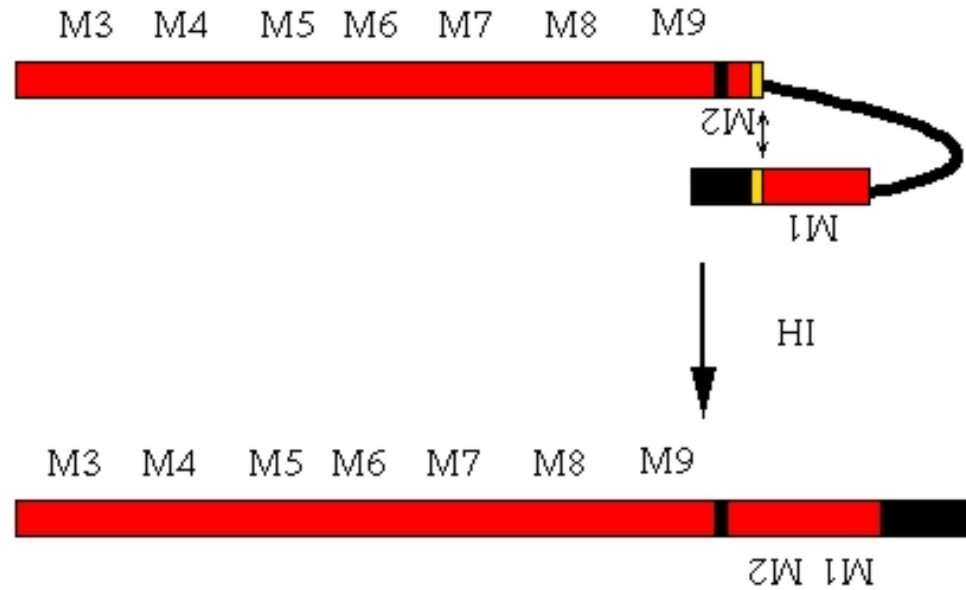


# Example: assembling actin I gene in *S.nova*

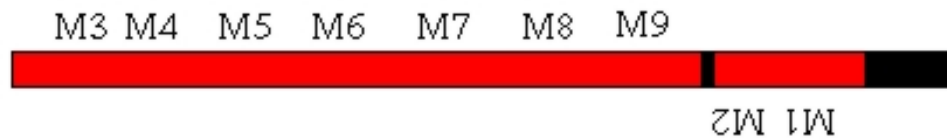


Step 5:

HI

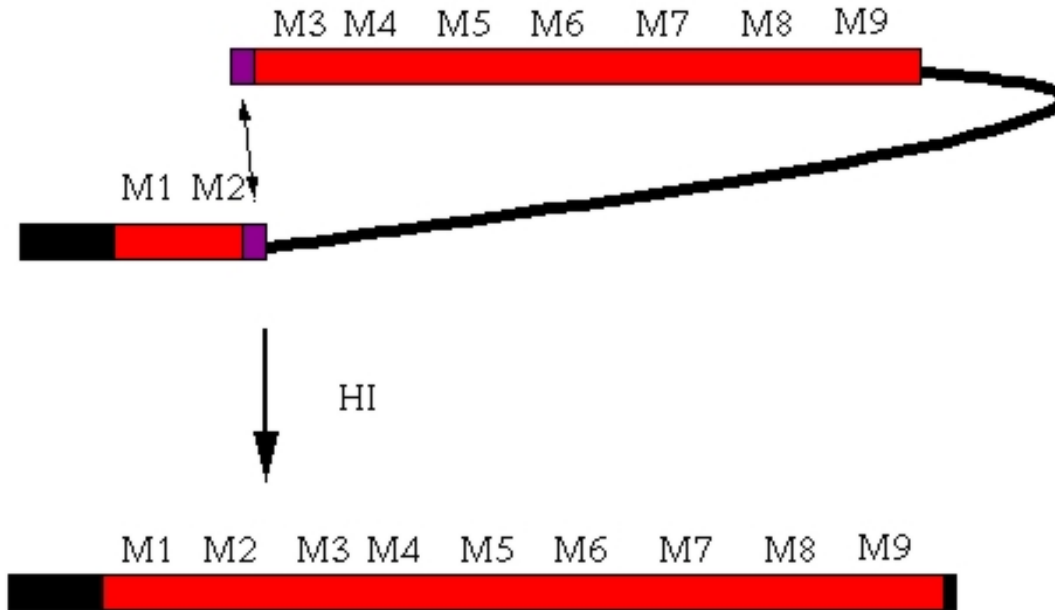


# Example: assembling actin I gene in *S.nova*



Step 6:

LD



# Assembling a gene with the intermolecular model

- Excising IES (LD): intramolecular recombination, applied on a pointer flanking an IES
- Restoring the orientation of an inverted MDS (HI): intermolecular recombination between two copies of a linear molecule, the copies are aligned so that pointers have the same orientation
- Sorting MDSs (DLAD): intramolecular recombination + intermolecular recombination, excision by one of the overlapping pointers and reinsertion by another pointer

- One can assemble any gene pattern in the intramolecular model (this will be shown later during the course);
- the intermolecular model can predict the assembly of any gene pattern as well!!!

# Intermolecular vs Intramolecular models

## Intramolecular model:

- only one molecule folds on itself and is involved in recombinations
  - only noncoding material (or an entire gene) may be excised
  - irreversible operations
- Splicing based molecular models

## Intermolecular model:

- two molecules may interact with each other
- big pieces of DNA (several MDSs and IESs) may be excised
- reversible operations



# Problems with pointer alignment

- Pointers are short (2-20 bp)
- They could occur also in the middle of MDSs and IESs
- How does one identify correctly pointers, MDSs and IESs?

# Problems with pointer alignment

- Molecular mechanisms that execute DNA processing in hypotrichs must perform two major functions:
  - Recognition with great accuracy of the cutting sites
  - Enzymes to cut and splice the DNA at the many different recognized sites
    - Issue: how are the relevant enzymes directed so precisely to the processing sites ?
    - IESs: 72%-100% AT-rich; MDSs: 45-65% AT-rich
    - Pointers: 6-19 bp in length, average of 10 bp
    - Non-scrambled genes: pointers of 2-4 bp
- Recognition of pointer sites: two explanations
  - Detecting differences in the TA content of the two neighboring regions
  - The old macronucleus acts as a template and the pointers are used just for the splicing – there are experimental evidences

# Template guided recombination

- DNA-template guided recombination model proposed by G. Rozenberg et al. in 2003 - a DNA molecule from the old macronucleus enters the newly developing macronucleus and acts as a template to guide pointer alignment
- RNA-template guided recombination model proposed by N. Jonoska et al. in 2007 - the RNA molecule from the old macronucleus serves as a catalyst that remains unchanged after DNA recombination (supported by experimental results, by L. Landweber et al. in 2007)



# RNA template-guided recombination (By Jonoska and others, 2007)

RNA-template guided recombination model proposed by N. Jonoska et al. in 2007 - the RNA molecule from the old macronucleus serves as a catalyst that remains unchanged after DNA recombination (supported by experimental results, by L. Landweber et al. in 2007)

This model proposes that a double- or single stranded RNA from the maternal MAC facilitates the MIC DNA and allows recombination without modifying any portion of the template

# Template-guided recombination - Summary

The mechanism implementing TGR:

- recognizes the alignment configuration
- makes the specific cuts in the phosphodiester backbones
- facilitates appropriate rotations of the resulting molecular segments
- relegates the phosphodiester backbones in their new configurations

this mechanism is independent of the sequences that form the *X-T-Y* alignment structure.

# Template-guided recombination - Summary

- The template-guided recombination model provides solutions to several troubling problems:
  - identifies correct copies of pointer sequence
  - defines precisely the boundary between MDSs and adjacent IESs
  - provides for irreversible recombination (the intermolecular model with the template guided recombination becomes irreversible)

# Template-guided recombination - Experimental results

From *RNA-mediated epigenetic programming of a genome-rearrangement pathway* by M. Nowacki, V. Vijayan, Y. Zhou, K. Schotanus, T. G. Doak, L. F. Landweber (2007):

- A complete RNA cache of the maternal somatic genome may be available during gene assembly
- disruption of specific RNA molecules disables rearrangement of the corresponding gene
- Injection of artificial templates reprogrammes the DNA rearrangement pathway

- The intermolecular model. During gene assembly several DNA may recombine with each other. Reversible operations.
  - The intramolecular model. During gene assembly a single DNA recombines with itself. Irreversible operations. Pointer DNA sequences cease act as pointers after participating in a intramolecular operation
- Template-guided recombination. DNA or RNA from maternal MAC helps to align correctly pointers, to excise precisely IESs and to unscramble shuffled MDSs. Two models: DNA based (2003 Rozenberg, template is reconstructed, problem with trash, non-compatible with intra.model) and RNA based (2007 Jonoska, Landweber, experimentally supported, one can suppress or reprogram gene assembly by altering RNA-templates pool, compatible with both inter. and intra.models)

