The Red-Queen model of recombination hotspots evolution

Thibault Latrille, Laurent Duret, Nicolas Lartillot

Laboratoire de Biométrie et Biologie Évolutive, Université de Lyon

S-51 Causes and consequences of recombination rate evolution

What is the equilibrium recombination rate determined by the Red-Queen model of recombination hotspots?

Recombination is concentrated in hotspots

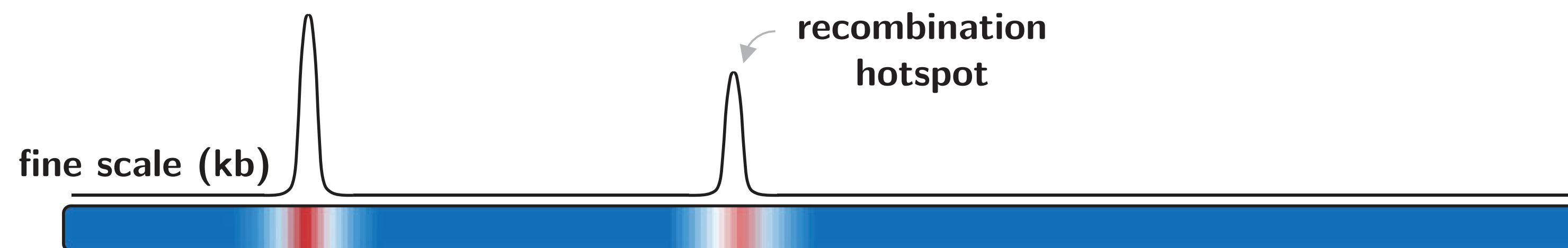

recombination hotspot

fine scale (kb)

Figure adapted from Dapper (2017)

- Hotspots are short lived, not shared between human and chimpanzee.

Kauppi *et al* **(2004)**, **Ptak** *et al* **(2005)**, **Arnheim** *et al* **(2007)**

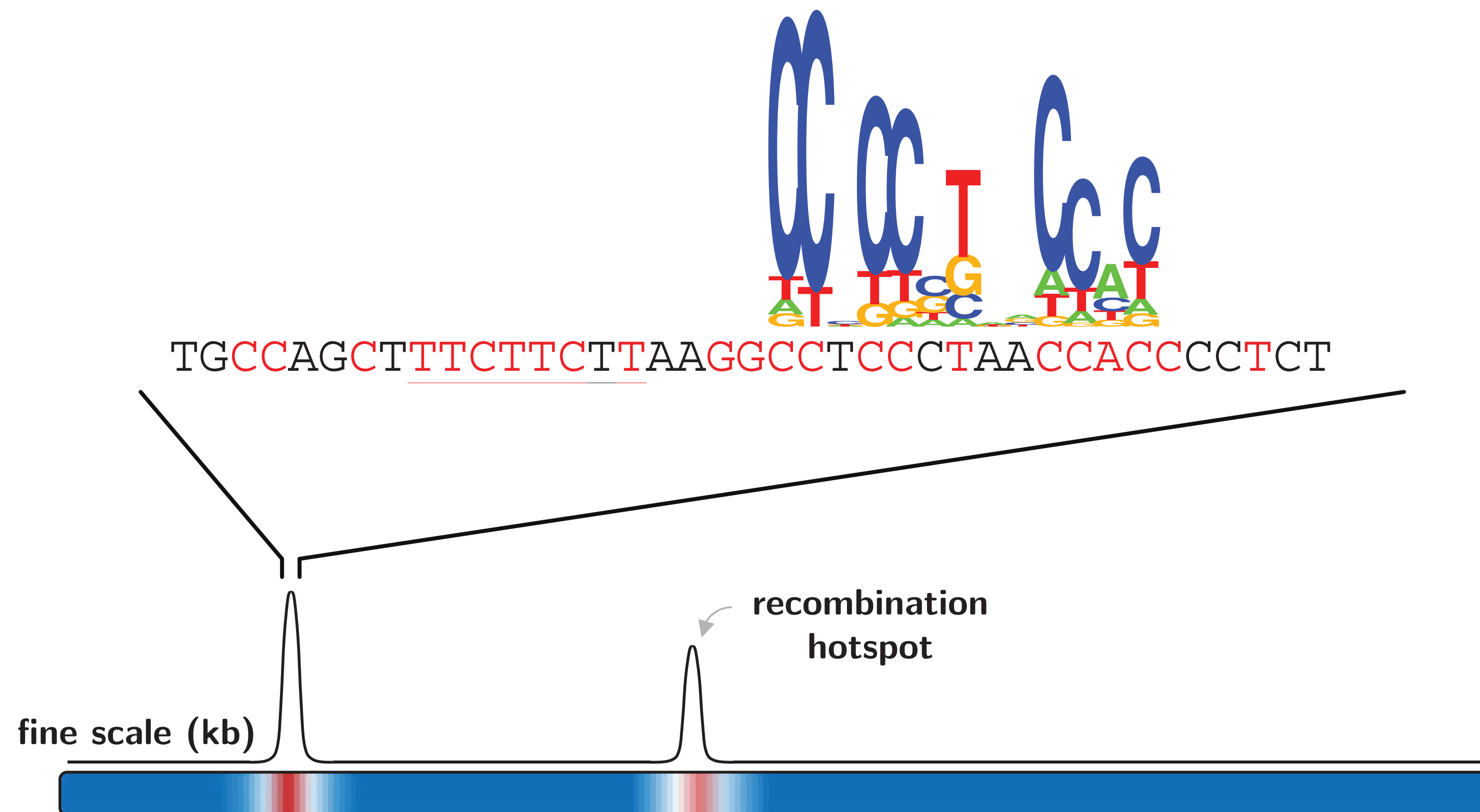# Recombination is concentrated in hotspots containing a specific motif

Figure adapted from Dapper (2017)

- Hotspots are short lived, not shared between human and chimpanzee.
- Human hotspots are enriched in a sequence motif.
- Human motif is not enriched in chimpanzee hotspots.
- Human motif has been depleted in the human lineage.

Wall *et al* **(2003)**, **Ptak** *et al* **(2004)**, **Winckler** *et al* **(2005)**, **Myers (2008)**

Latrille Thibault Red-Queen of recombination hotspots

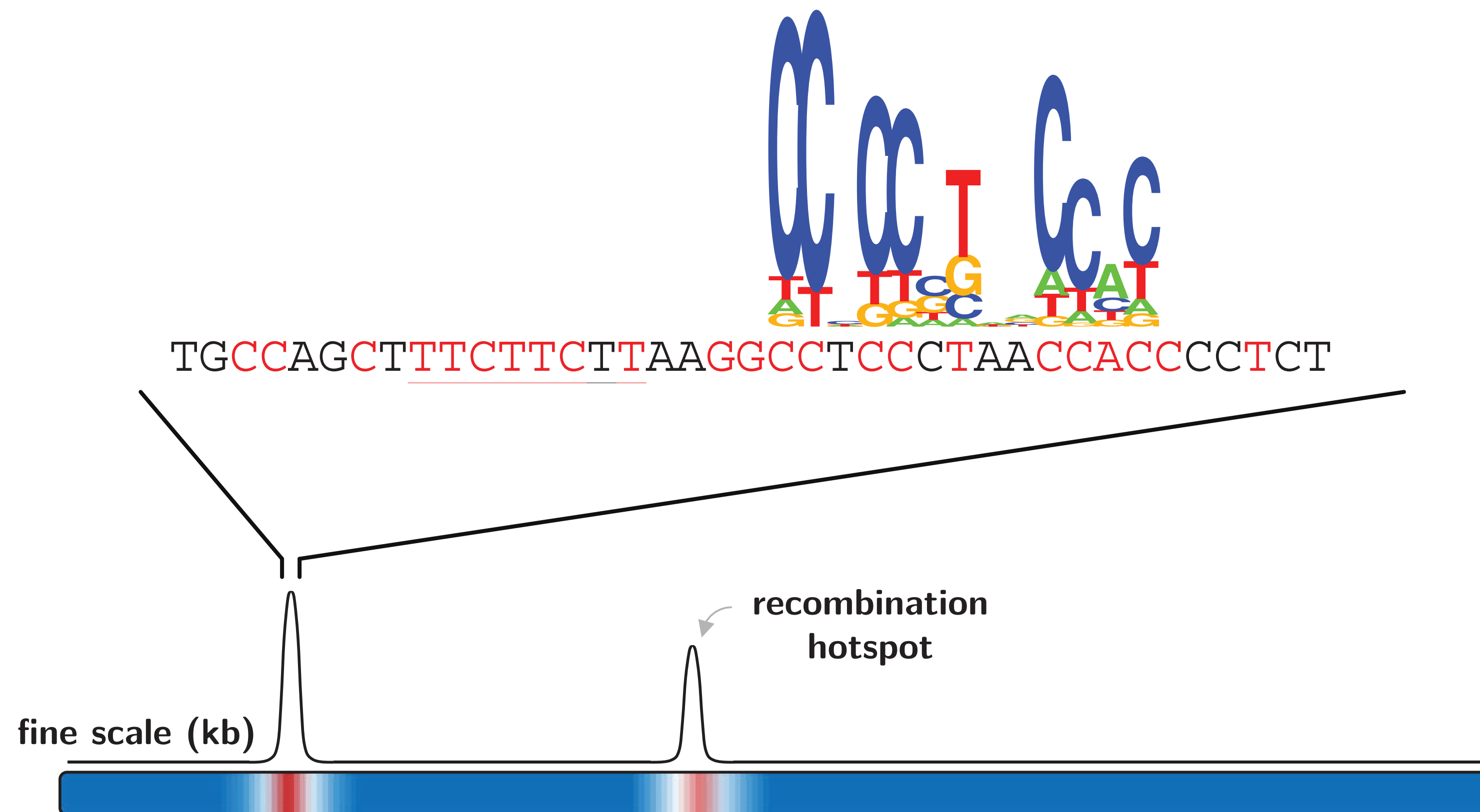# Recombination is concentrated in hotspots containing a specific motif

Figure adapted from Dapper (2017)

- Hotspots are short lived, not shared between human and chimpanzee.
- Human hotspots are enriched in a sequence motif.
- Human motif is not enriched in chimpanzee hotspots.
- Human motif has been depleted in the human lineage.

- Why are chimpanzee & human hotspots not using the same motif?
- Why are human motifs depleted in the human lineage?

**Wall** *et al* (**2003**), **Ptak** *et al* (**2004**), **Winckler** *et al* (**2005**), **Myers** (**2008**)

PRDM9

- PRDM9 determines the location of hotspots, by binding to the motif and then recruiting the recombination machinery.

Protein-protein interaction domain   **Nucleosomes marking domain**   **Zn-fingers DNA-binding domain**

KRAB   PR/SET

GAACACCAGACGACGAAGAAGAACCGCCGTAACCACCGAT

Hotspots motif

Figure adapted from Hochwagen (2010)

- PRDM9 is fast evolving protein, different in Chimpanzee, with both a high dN/dS and a high mutation rate.

Baudat *et al* (**2010**), **Myers** *et al* (**2010**), **Parvanov** *et al* (**2010**)

Latrille Thibault   Red-Queen of recombination hotspots

PRDM9

- PRDM9 determines the location of hotspots, by binding to the motif and then recruiting the recombination machinery.

Figure adapted from Hochwagen (2010)
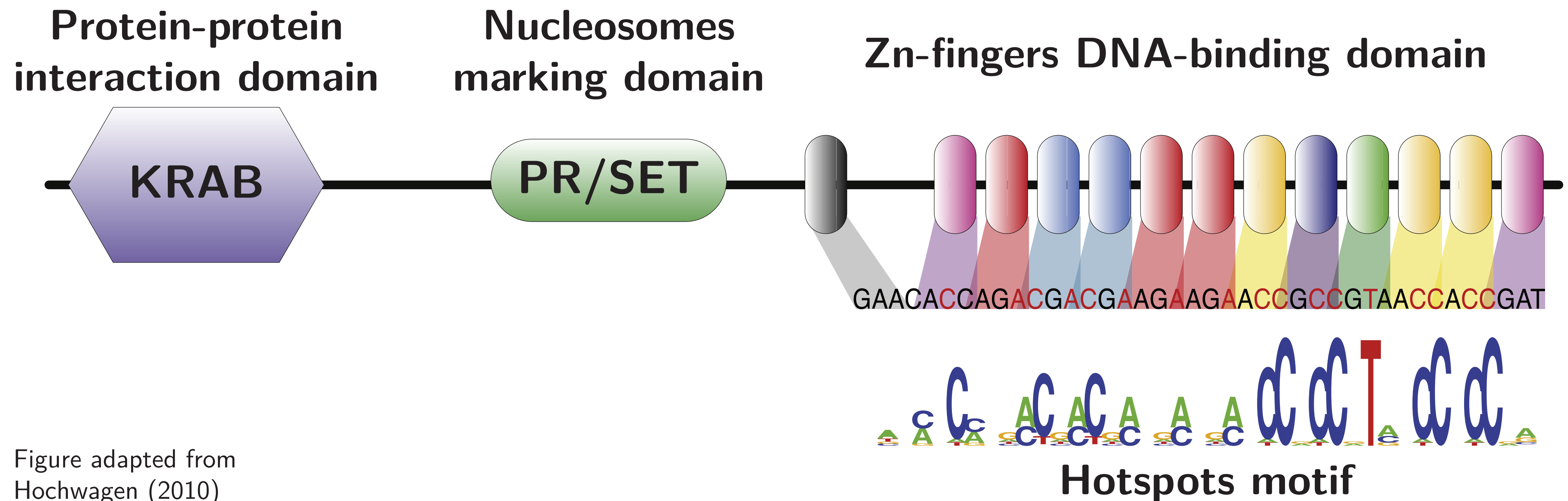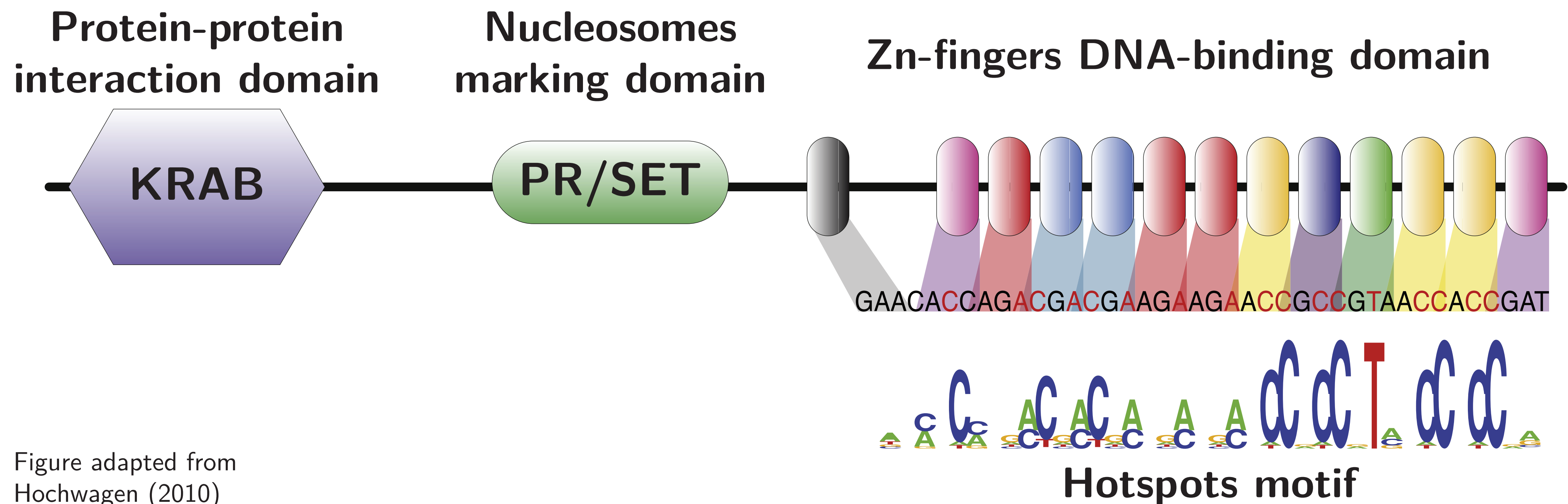
- PRDM9 is fast evolving protein, different in Chimpanzee, with both a high dN/dS and a high mutation rate.

- Why are human motifs depleted in the human lineage?

Baudat *et al* (**2010**), **Myers** *et al* (**2010**), **Parvanov** *et al* (**2010**)

Latrille Thibault Red-Queen of recombination hotspots

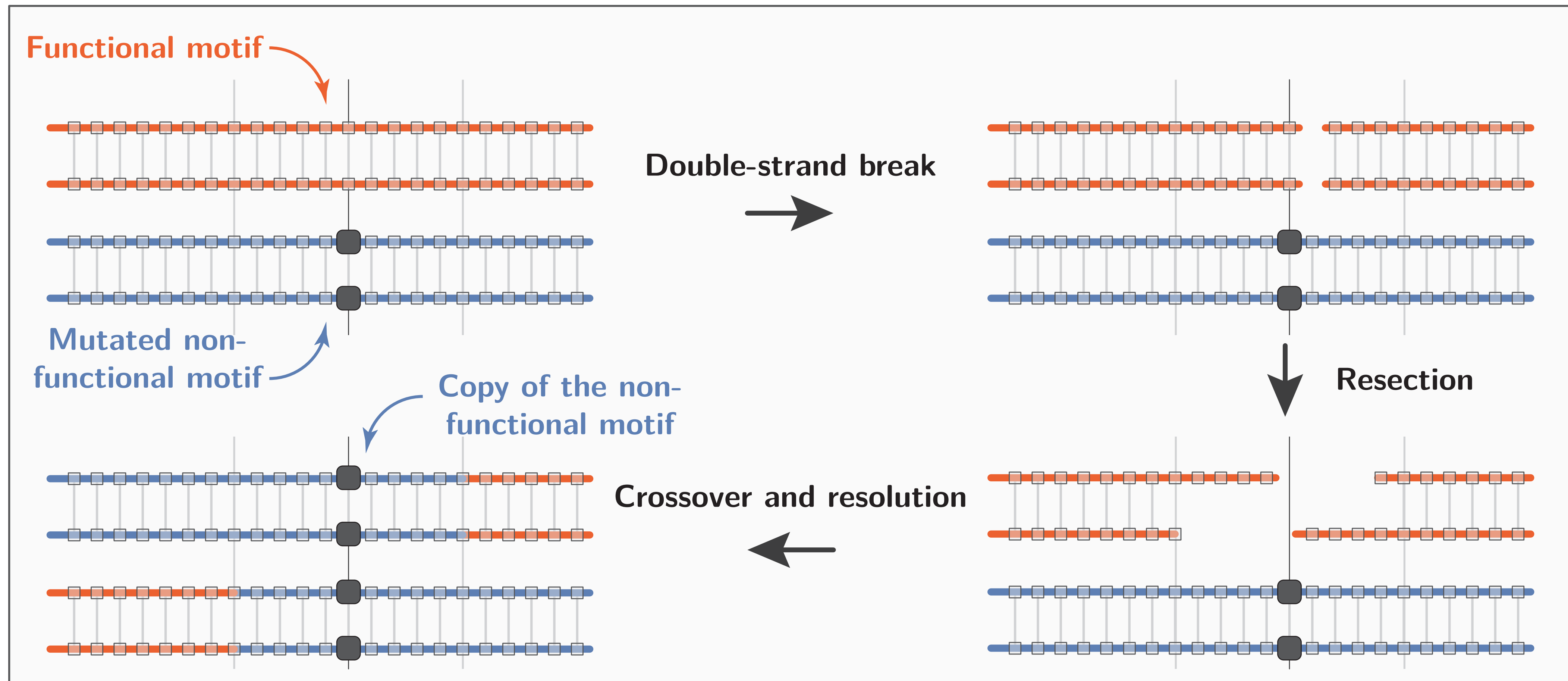# Motifs at the hotspots are architects of their own destruction
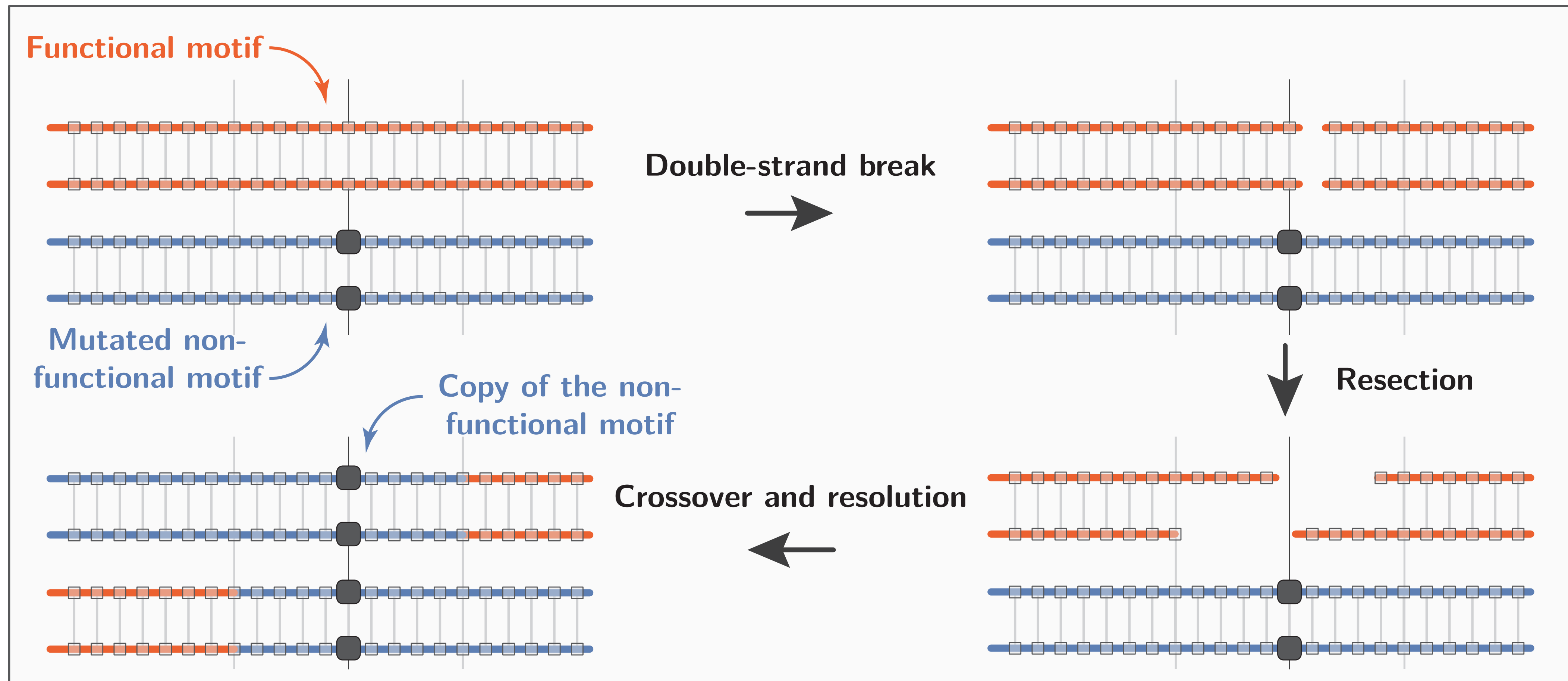
- Biased Gene Conversion favors a mutated motif over a functional motif.
- Mutations disrupting the motif are "selected for".
- Motifs (and thus hotspots) are not conserved.

Boulton et al (1997), Pineda-Krch and Redfield (2005), Coop & Myers (2007)

Latrille Thibault   Red-Queen of recombination hotspots

# Motifs at the hotspots are architects of their own destruction

- Biased Gene Conversion favors a mutated motif over a functional motif.
- Mutations disrupting the motif are "selected for".
- Motifs (and thus hotspots) are not conserved.

- Whom is running after whom?

Boulton et al (1997), Pineda-Krch and Redfield (2005), Coop & Myers (2007)

Latrille Thibault Red-Queen of recombination hotspots

Erosion-invasion cycle

- Biased Gene Conversion (BGC) erodes the motifs targeted by PRDM9.
- Once motifs are depleted, invasion of a new mutant of PRDM9 restores hotspots by targeting 'virgin' motifs.
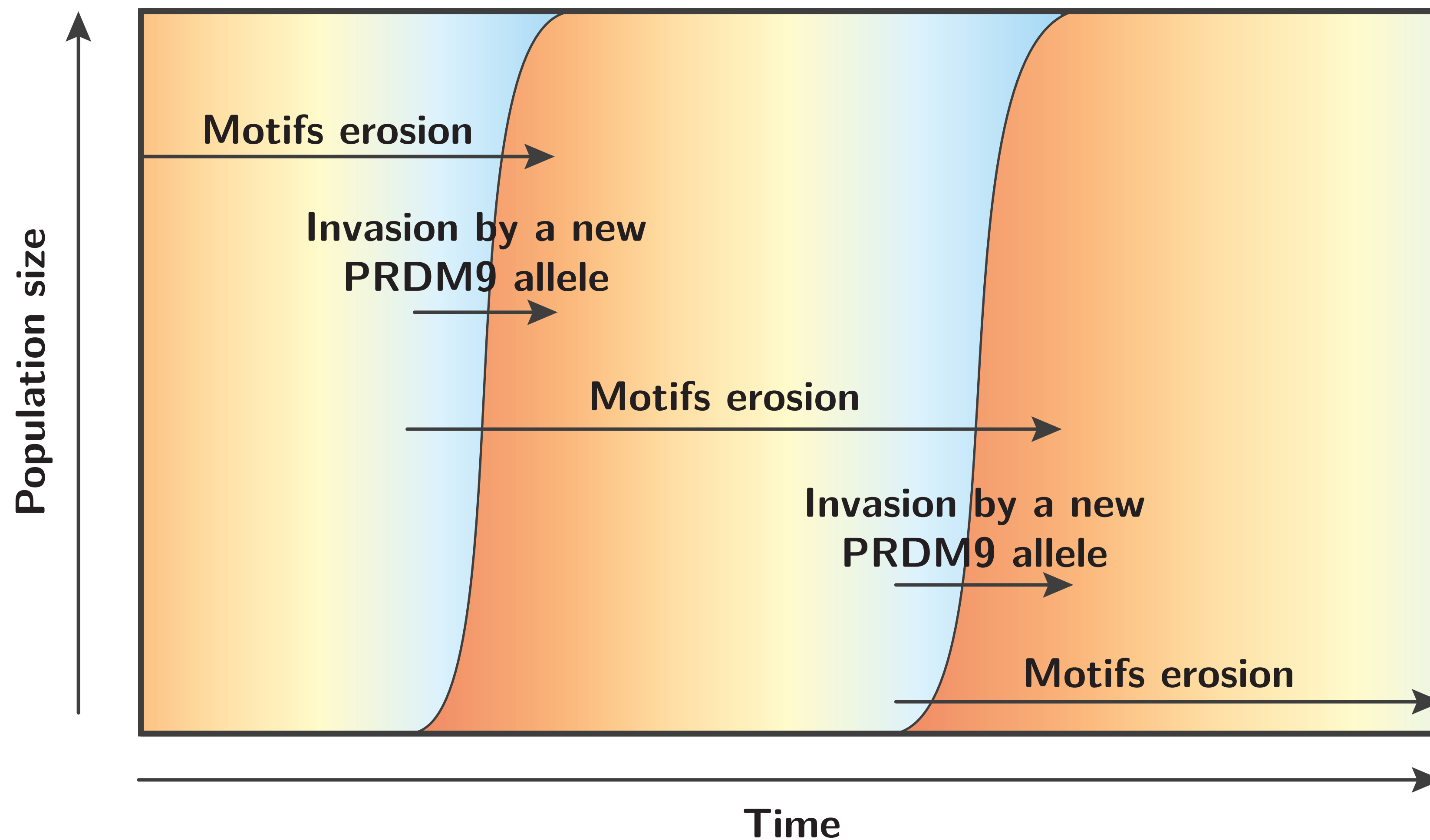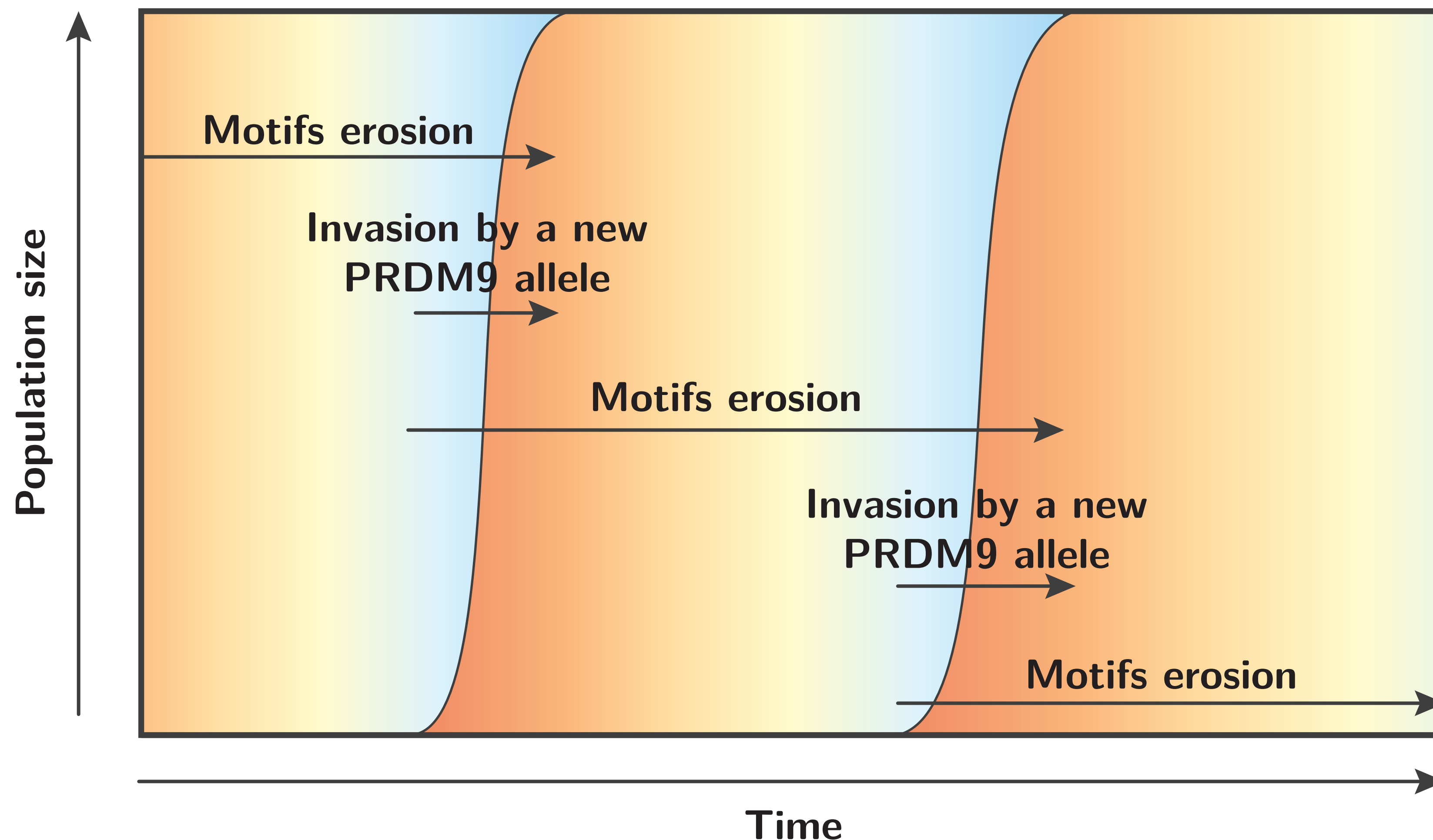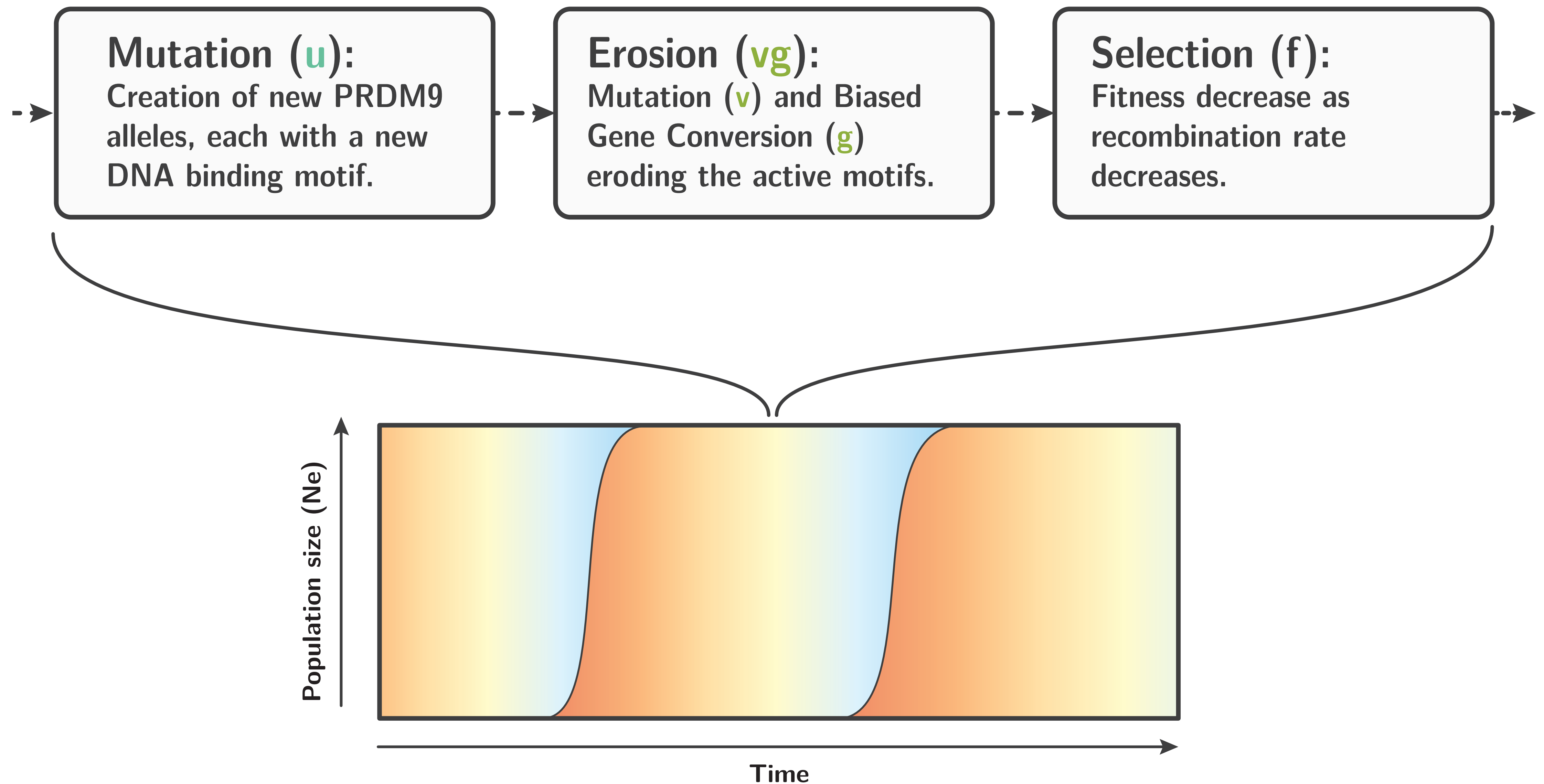
Erosion-invasion cycle

- Biased Gene Conversion (BGC) erodes the motifs targeted by PRDM9.
- Once motifs are depleted, invasion of a new mutant of PRDM9 restores hotspots by targeting 'virgin' motifs.

- On which condition is the PRDM9 locus polymorphic?
- What is the equilibrium recombination rate?

Population genetic modeling

Wright-Fisher simulation with non-overlapping generations

Mutation (u):
Creation of new PRDM9 alleles, each with a new DNA binding motif.

Erosion (vg):
Mutation (v) and Biased Gene Conversion (g) eroding the active motifs.

Selection (f):
Fitness decrease as recombination rate decreases.

Population size (Ne)

Time

Latrille Thibault · Red-Queen of recombination hotspots
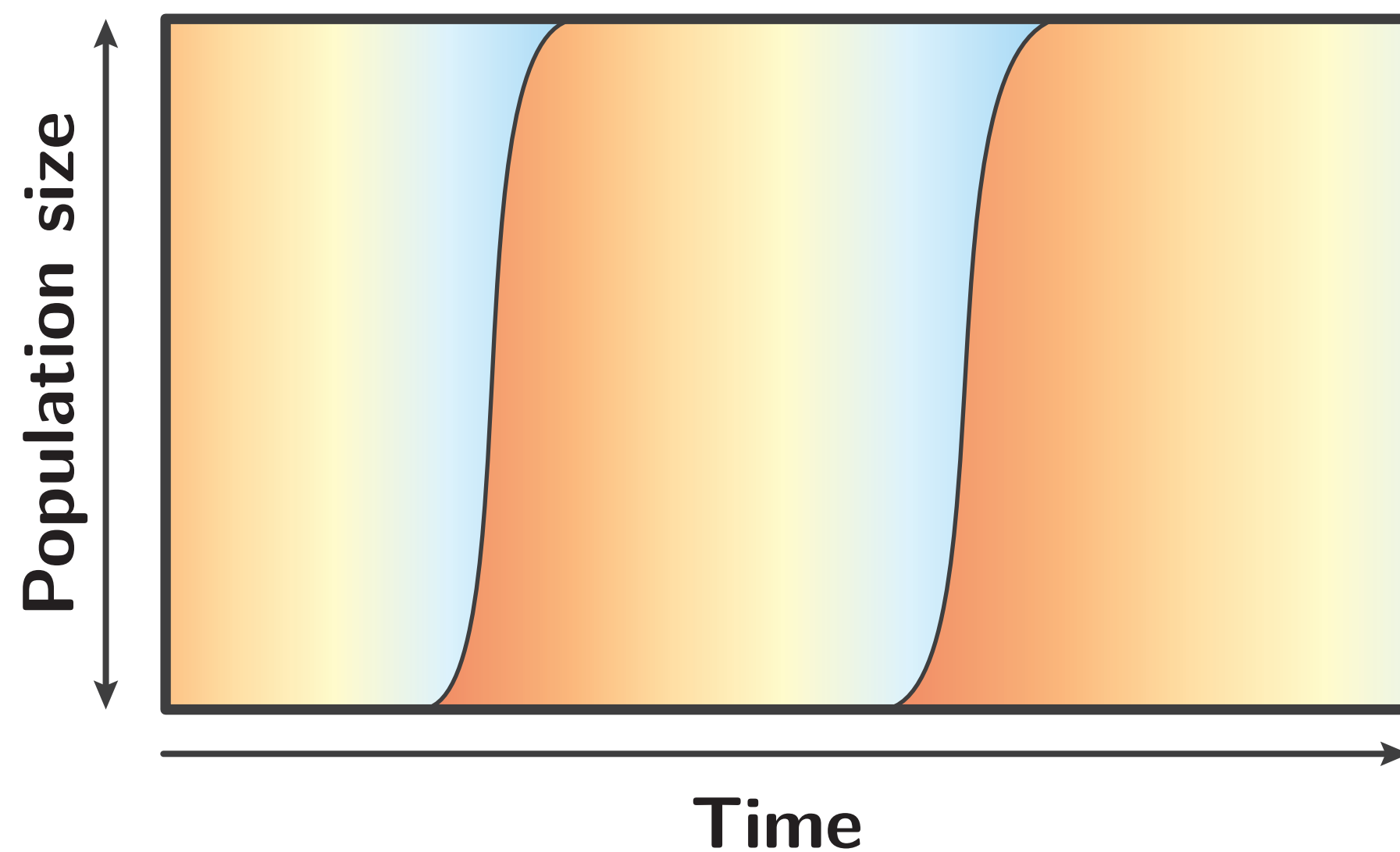
Is the PRDM9 locus polymorphic?

$$N_e, u, vg, f$$

$$N_e u$$

The expected number of new PRDM9
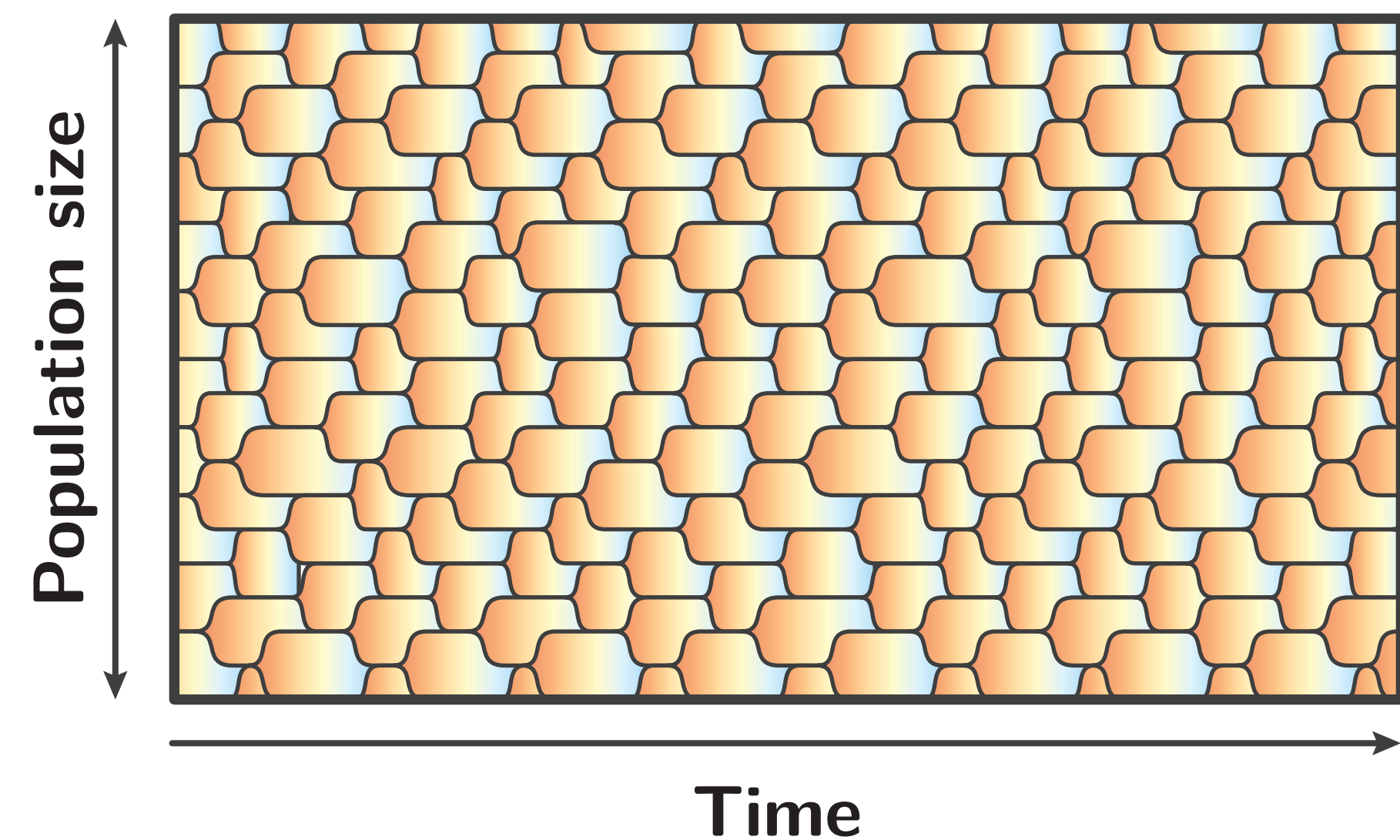alleles in the population at each generation.

$N_e u \ll 1$
Succession of
PRDM9 alleles

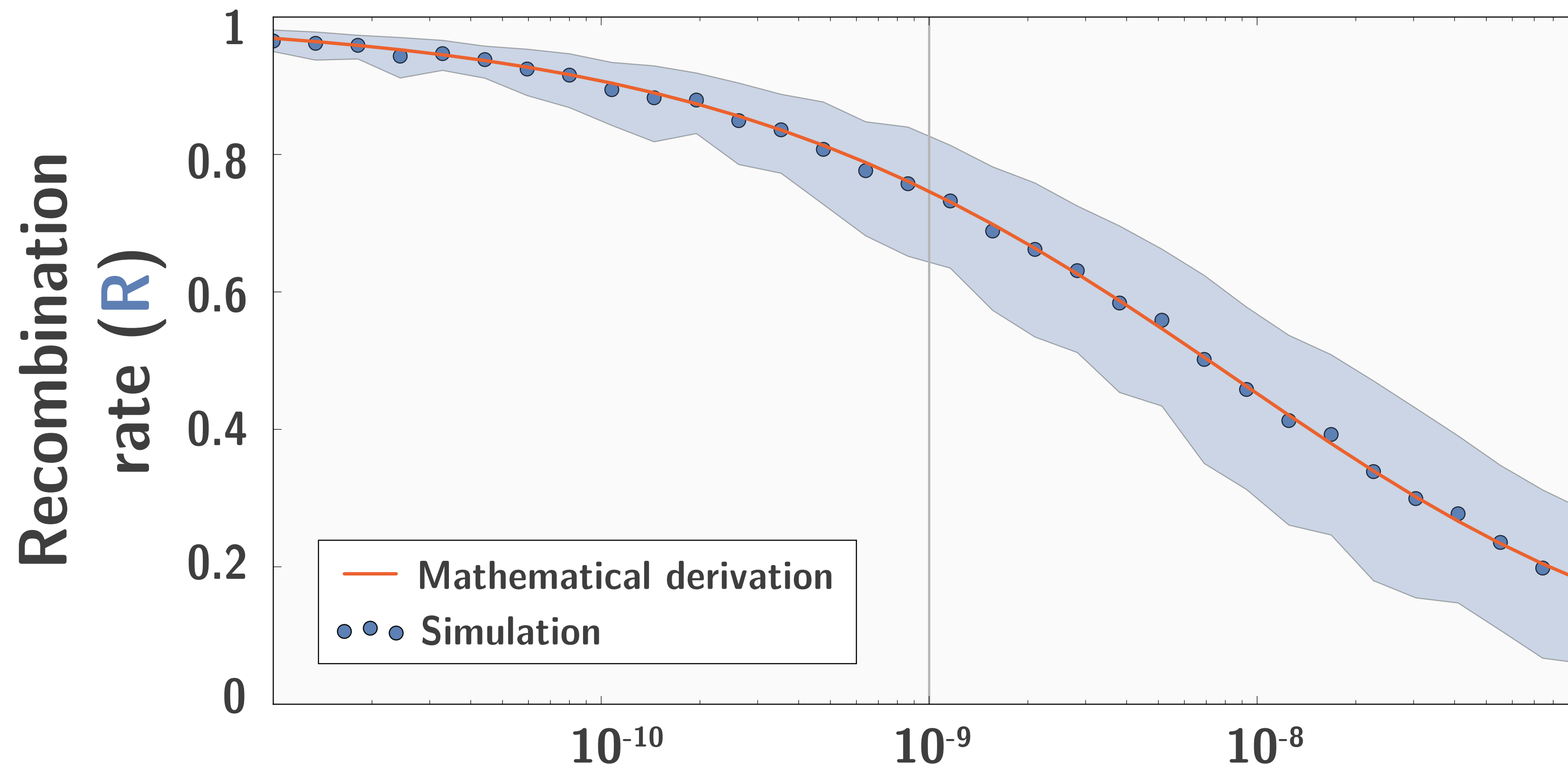

$N_e u \gg 1$
Polymorphism of
PRDM9 alleles

Latrille *et al* **(2017)**

What is the equilibrium recombination rate?

Increase erosion rate

⇓
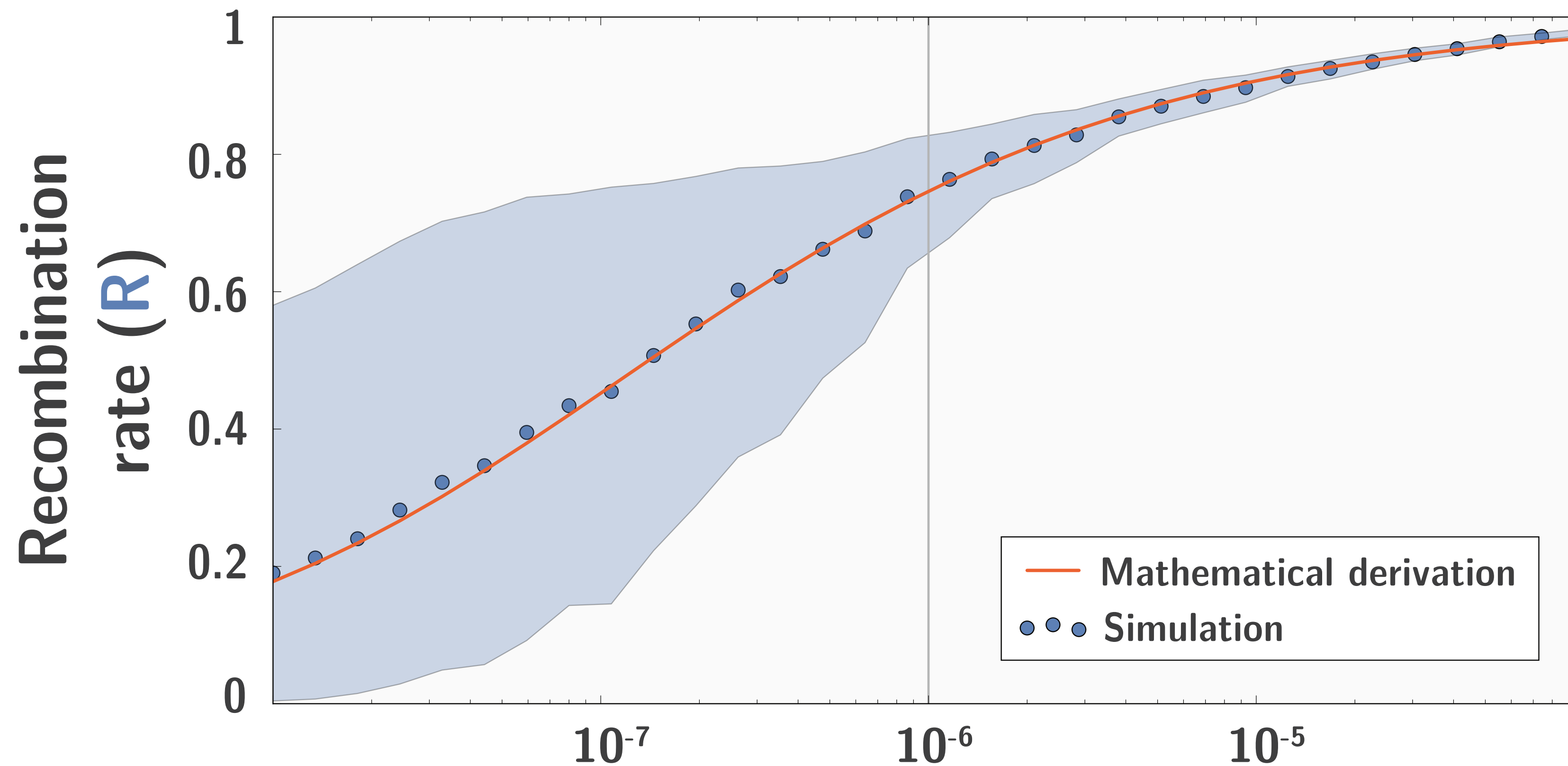
Increase erosion of motifs

⇓

Decrease recombination rate

What is the equilibrium recombination rate?

Increase mutation rate of PRDM9

⇓

Increase invasion rate of PRDM9 alleles and recruitment of 'virgin' motifs

⇓

Increase recombination rate

Latrille Thibault Red-Queen of recombination hotspots

Equilibrium recombination rate in equations

In succession regime, assuming the invasion time of a PRDM9 allele is short compared to the lifespan of this allele:

$$R = g\left(\frac{vg}{u}\right) \simeq 1 - \sqrt{\frac{vg}{u}}$$

Latrille *et al* (2017)

Equilibrium recombination rate in equations

In succession regime, assuming the invasion time of a PRDM9 allele is short compared to the lifespan of this allele:

$$R = g\left(\frac{vg}{u}\right) \simeq 1 - \sqrt{\frac{vg}{u}}$$

In polymorphic regime, assuming many PRDM9 alleles to apply mean-field approximation:

Latrille _et al_ (2017)

Equilibrium recombination rate in equations

In succession regime, assuming the invasion time of a PRDM9 allele is short compared to the lifespan of this allele:

$$R = g\left(\frac{vg}{u}\right) \simeq 1 - \sqrt{\frac{vg}{u}}$$

In polymorphic regime, assuming many PRDM9 alleles to apply mean-field approximation:

$$R = g\left(\frac{vg}{u}\right) \simeq 1 - \sqrt{\frac{vg}{u}}$$

Latrille *et al* (2017)

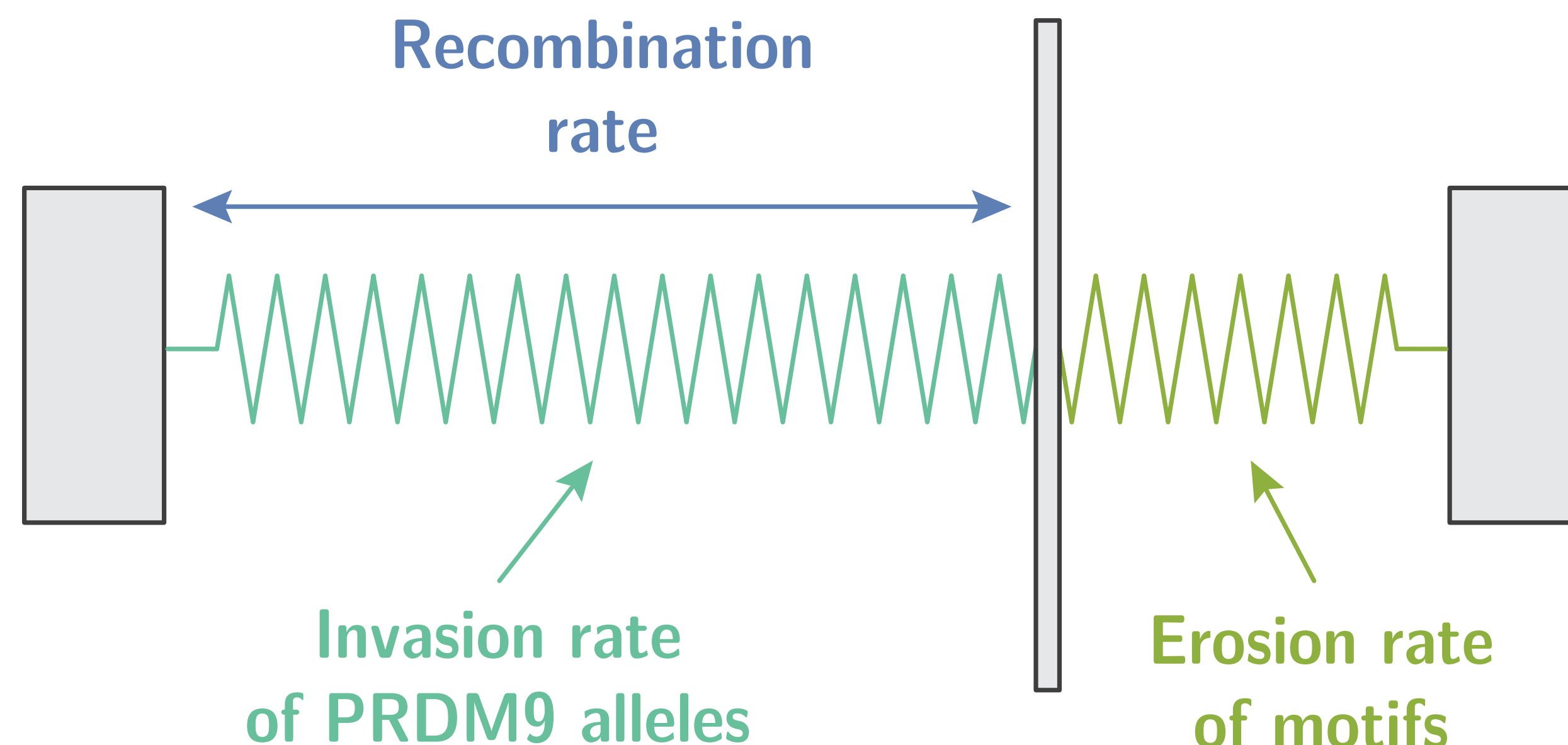Latrille Thibault   Red-Queen of recombination hotspots

Equilibrium recombination rate in equations

In succession regime, assuming the invasion time of a PRDM9 allele is short compared to the lifespan of this allele:

$$R = g\left(\frac{vg}{u}\right) \simeq 1 - \sqrt{\frac{vg}{u}}$$

In polymorphic regime, assuming many PRDM9 alleles to apply mean-field approximation:

$$R = g\left(\frac{vg}{u}\right) \simeq 1 - \sqrt{\frac{vg}{u}}$$

Recombination rate

Invasion rate of PRDM9 alleles

Erosion rate of motifs

Latrille *et al* (2017)

Latrille Thibault Red-Queen of recombination hotspots

Conclusions

- **Mutation, not selection explains the PRDM9 diversity**

- **Balance between invasion rate of PRDM9 alleles and erosion rate of motifs determines the equilibrium recombination rate.**

- **Prediction of the model can be tested in multiple species.**

- **Calibration in mouse suggests a high mutation rate of PRDM9 ($\sim$3e^{-6}) and a strong selection coefficient of BGC ($\sim$3e^{-3}).**

Latrille Thibault   Red-Queen of recombination hotspots

Conclusions & Perspectives

- **Mutation, not selection explains the PRDM9 diversity**

- **Balance between invasion rate of PRDM9 alleles and erosion rate of motifs determines the equilibrium recombination rate.**

- **Prediction of the model can be tested in multiple species.**

- **Calibration in mouse suggests a high mutation rate of PRDM9 ($\sim$3e^{-6}) and a strong selection coefficient of BGC ($\sim$3e^{-3}).**

- **Why is PRDM9 mutation rate so high in the first place?**

- **Is our model compatible with hybrid sterility due to hotspots asymmetry in a meta-population context?**

Latrille Thibault Red-Queen of recombination hotspots

Acknowledgments

- **Organizers of the symposium, of the conference.**

- **The LBBE lab (in Lyon).**

- **Funding agencies (ANR)**

- **You for your attention**

My PhD advisor

Me