

# THE RED QUEEN IN THE KINGDOM OF RECOMBINATION

ANR  
AGENCE NATIONALE DE LA RECHERCHE

THIBAULT LATRILLE, LAURENT DURET, NICOLAS LARTILLOT

LABORATOIRE DE BIOMÉTRIE ET BIOLOGIE ÉVOLUTIVE, UMR 5558

cnrs  
organisme de recherche

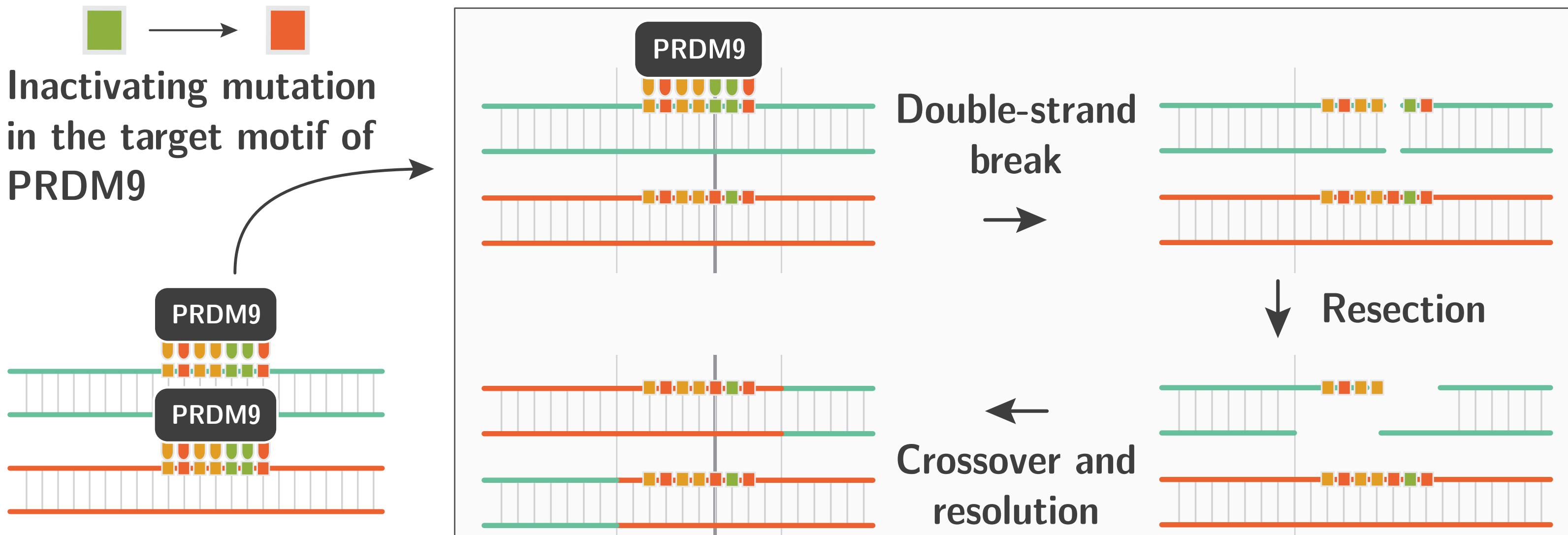


We propose a population-genetic model of the Red Queen dynamic of recombination. This model was implemented as a Wright-Fisher simulator, allowing exploration of the behaviour of the model (in terms of the mean recombination rate, diversity at the Prdm9 locus, or turnover rate) as a function of the parameters (effective population size, mutation and erosion rates). In a second step, analytical results, based on self-consistent mean-field approximations, were derived. These analytical results reproduce the scaling relations observed in the simulations, offering key insights about the detailed population-genetic mechanisms of the Red Queen model. Empirical fit of the model to current data from the mouse and humans suggests both a high mutation rate at PRDM9 and strong biased gene conversion on its targets.

## 1. QUALITATIVE MODEL

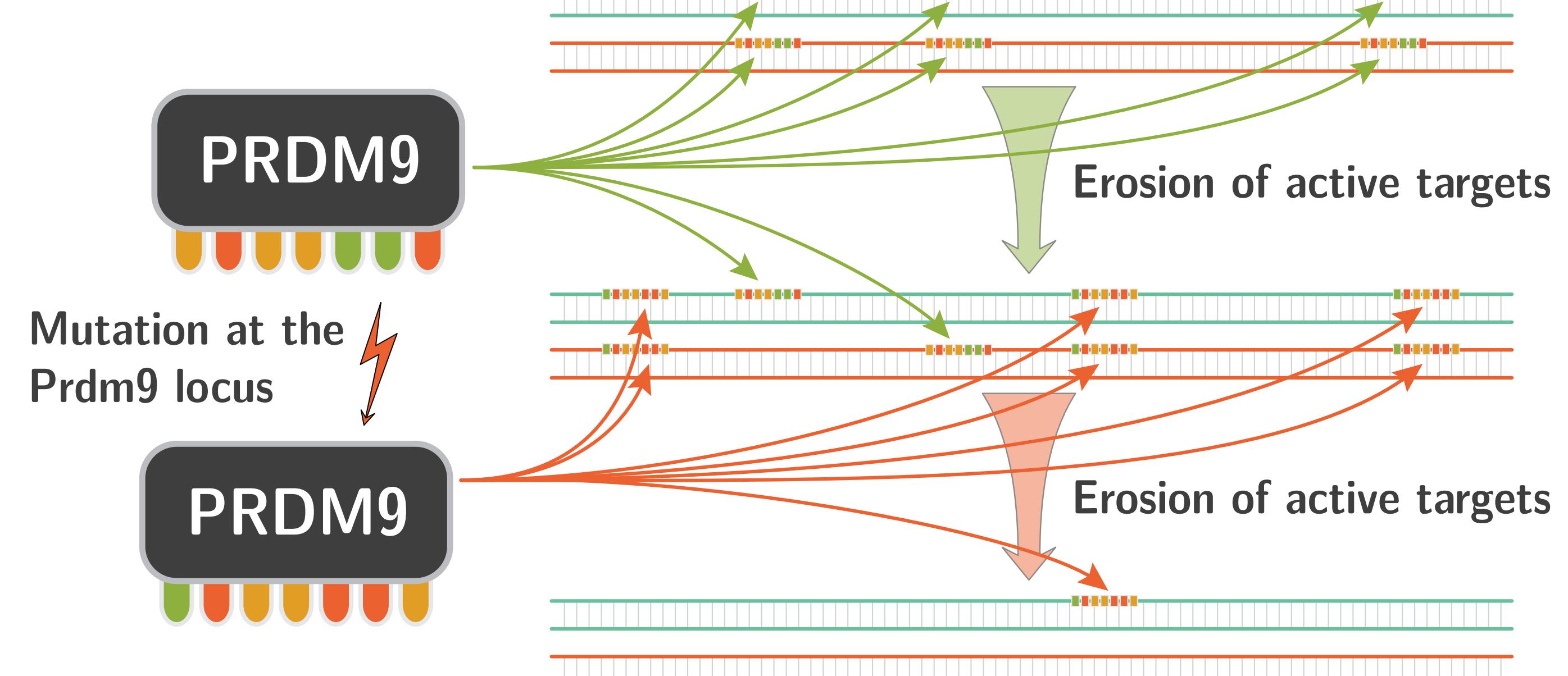
Hotspot location is determined by sequence motifs targeted by PRDM9 protein.  
PRDM9 binding triggers a double strand break event at the target location, initiating recombination.

Molecular mechanism of recombination results in biased gene conversion in favour of inactive mutant motifs. Leading to extinction of hotspots genome wide at the population level.



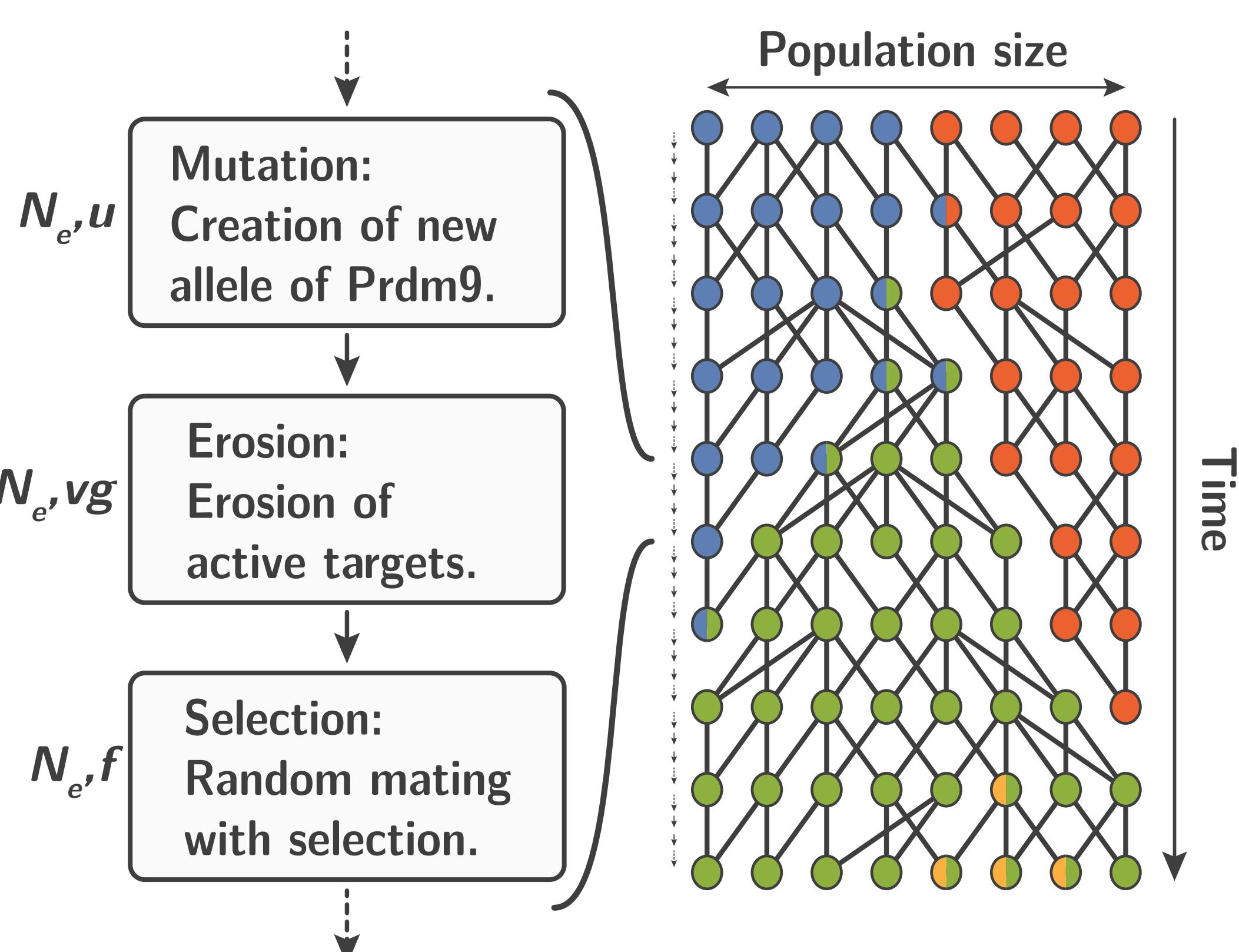
Extinction of hotspots by biased gene conversion

Depleted recombination induces positive selection on new Prdm9 alleles recognizing new sequence motifs. Thus restoring genome wide recombination.

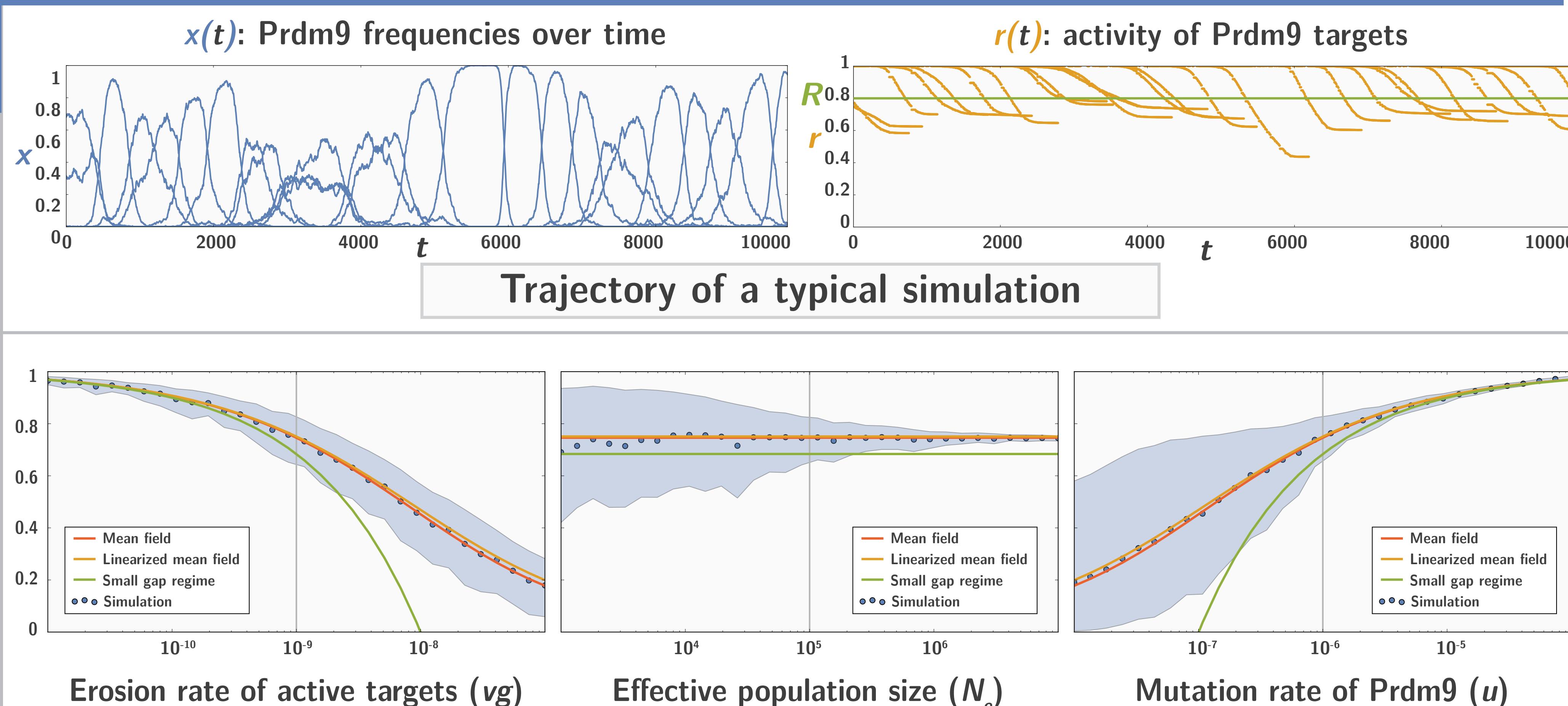


Positive selection on Prdm9

## 2. SIMULATION MODEL



Wright-fisher simulation at the Prdm9 locus



Mean activity of targets ( $R$ ) as a function of parameters of the model

## 3. ANALYTICAL APPROXIMATION

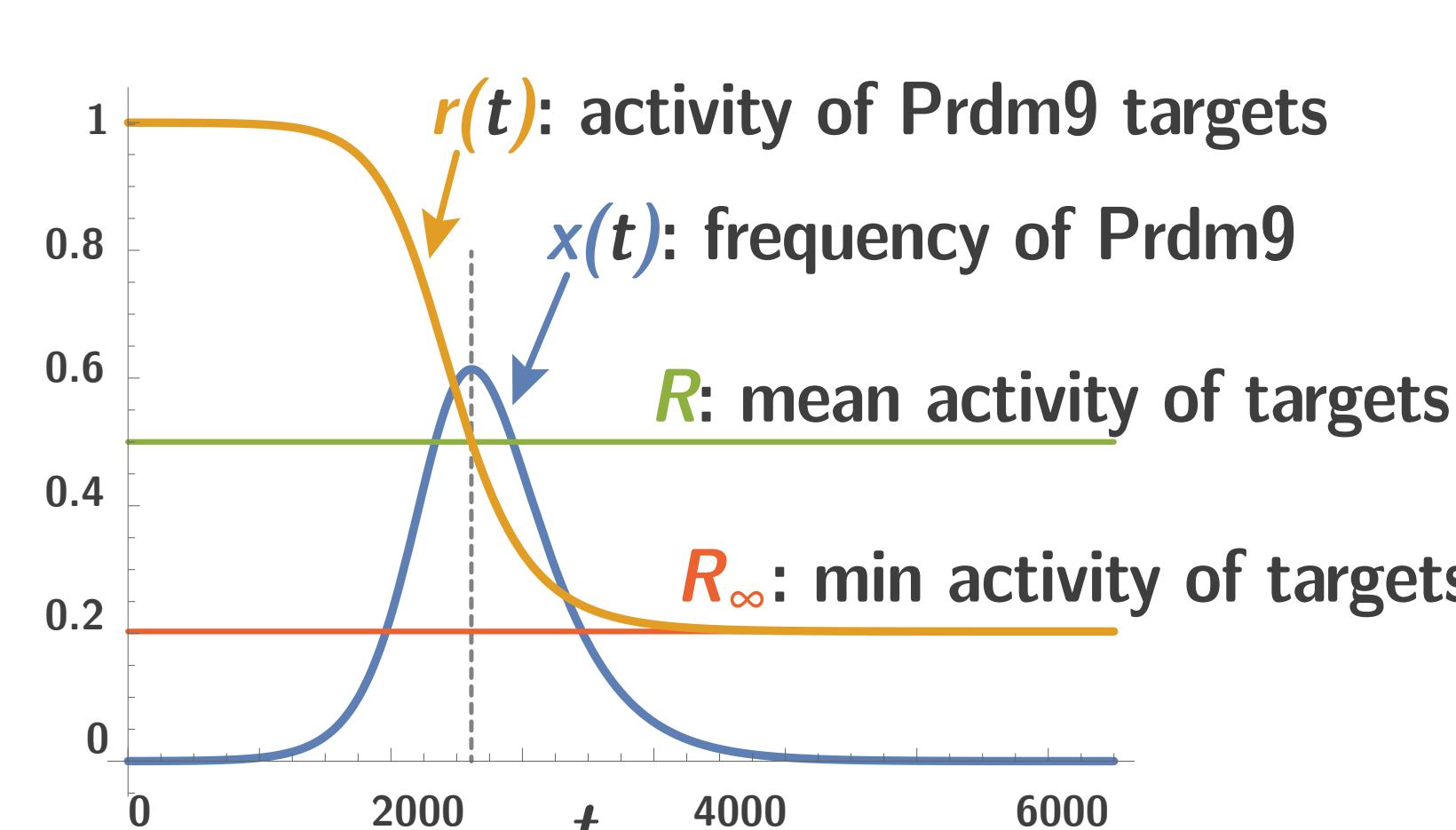
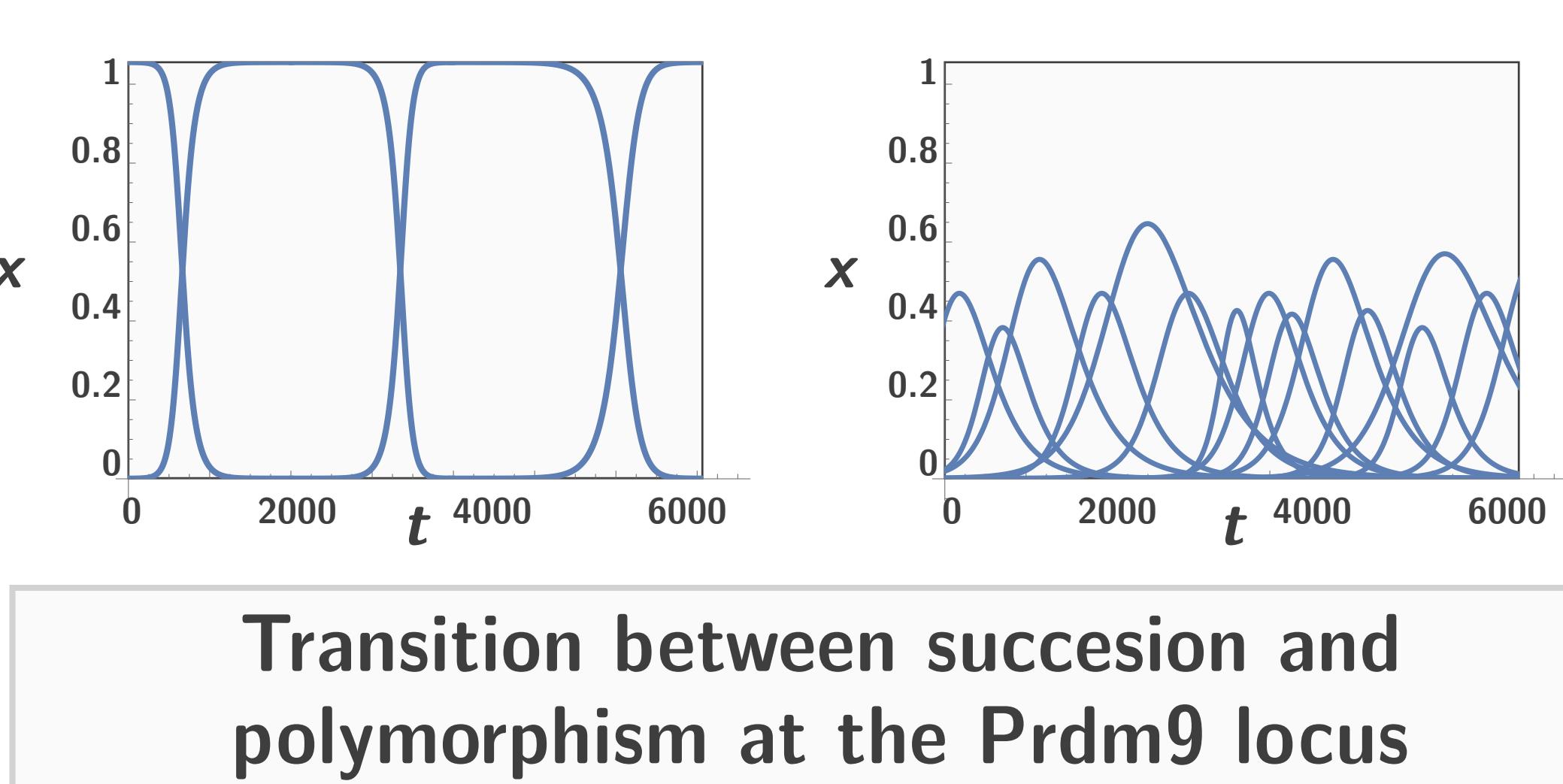
$N_e, u, vg, f$

$N_e$

The expected number of new Prdm9 alleles in the population at each generation.

$N_e u \ll 1$   
Succession of Prdm9 alleles

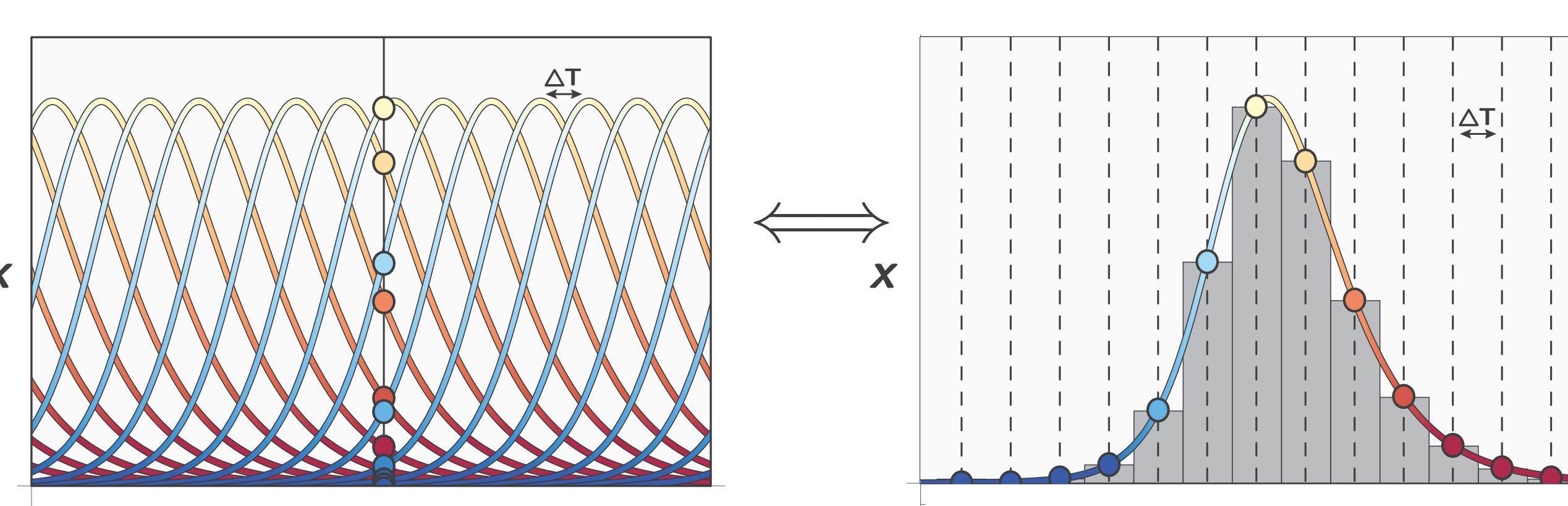
$N_e u \gg 1$   
Polymorphism of Prdm9 alleles



$$\begin{cases} \frac{dx(t)}{dt} = \frac{f'(R)}{2f(R)} (r(t) - R) x(t) \\ \frac{dr(t)}{dt} = -4N_e vg x(t) r(t) \end{cases}$$

$$\Rightarrow \begin{cases} x(r) = \frac{f'(R)}{2f(R)} \frac{1 - r + R \ln(r)}{4N_e vg} + x_{initial} \\ 0 = 1 - R_\infty + R \ln(R_\infty) \end{cases}$$

Trajectory of a single allele for a given  $R$



$$\begin{cases} \frac{1}{\tau} = 4N_e u \frac{f'(R)}{2f(R)} (1 - R) \\ \tau = \int_0^\infty x(t) dt = \frac{1 - R_\infty}{4N_e vg R} \end{cases}$$

$$\Rightarrow R = g(R, \frac{vg}{u})$$

$$\Rightarrow 1 - R \propto \sqrt{\frac{vg}{u}}$$

Mean-field approximation in polymorphic regime

## CONCLUSION

- For small  $N_e u$ , one single PRDM9 allele dominates the population at any time.
  - For large  $N_e u$ , polymorphism is maintained at the PRDM9 locus.
  - Mean activity of targets ( $R$ ) is independent of population size, in polymorphic and successional regimes.
  - The activity gap ( $1 - R$ ) scales as  $\sqrt{vg/u}$ .
- Empirical fit of the model to current data from the mouse and humans suggests both a high mutation rate at Prdm9 locus ( $u$ ) and strong biased gene conversion on its targets ( $g$ ).
- More complex models (e.g. including population structure) can be explored.