Modelling Fitness Trade-Offs of Rates of Horizontal Gene Transfer

Evolution 02-731 Project

Siddharth Reed

Carnegie Mellon University

May 6 2021

Background

Horizontal Gene Transfer (HGT)

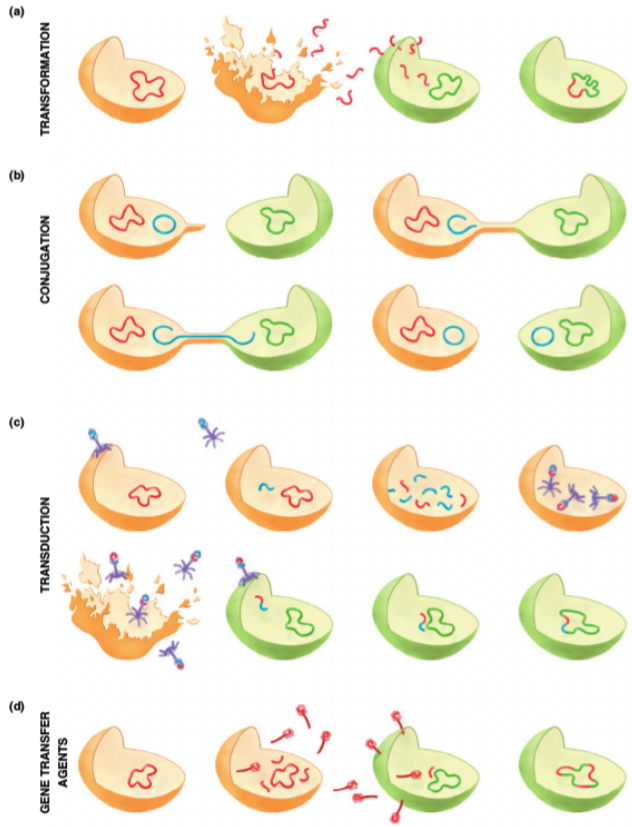


Figure 1: HGT Mechanisms

- **Transformation:** Incorporation of free-floating DNA into the genome

Horizontal Gene Transfer (HGT)

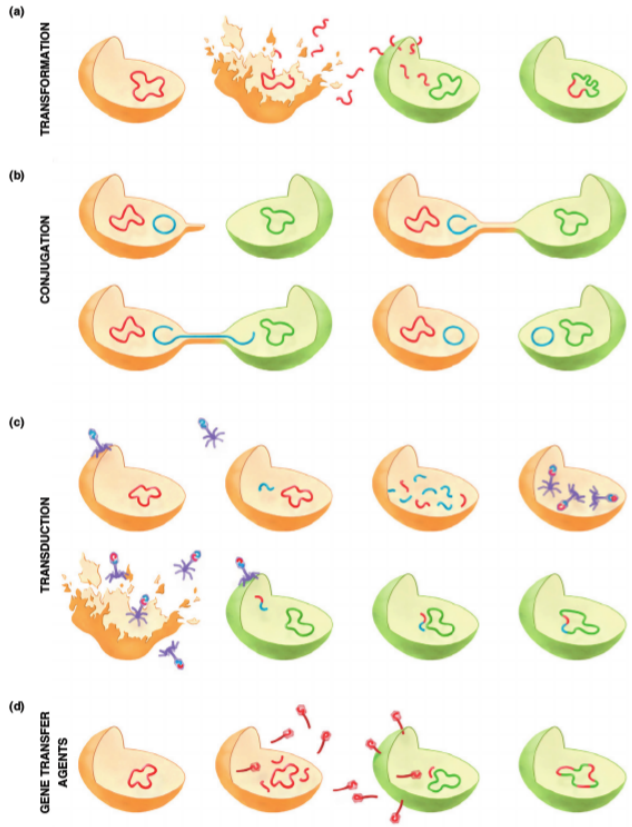


Figure 1: HGT Mechanisms

- **Transformation:** Incorporation of free-floating DNA into the genome
- **Conjugation:** Transfer of DNA through cell-cell connections

Horizontal Gene Transfer (HGT)

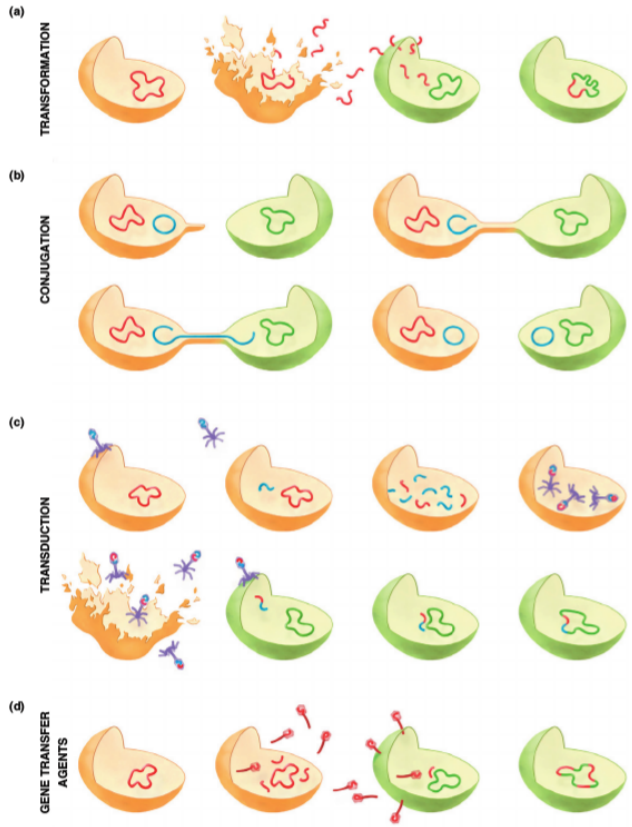


Figure 1: HGT Mechanisms

- **Transformation:** Incorporation of free-floating DNA into the genome
- **Conjugation:** Transfer of DNA through cell-cell connections
- **Transduction:** Transfer of DNA via phage

CRISPR-Cas In Bacteria

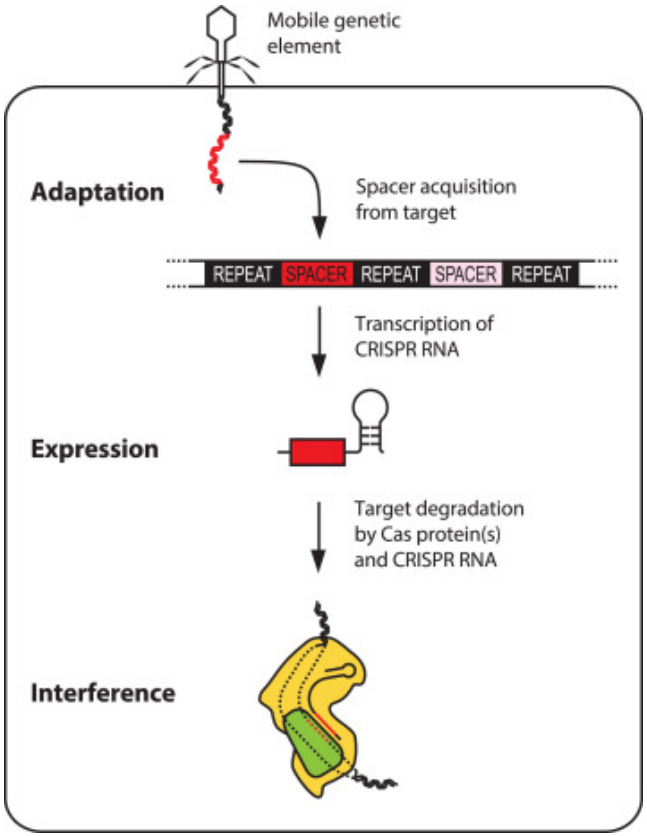


Figure 2: CRISPR-Cas Mechanism

- Adaptive Bacterial Immune System

CRISPR-Cas In Bacteria

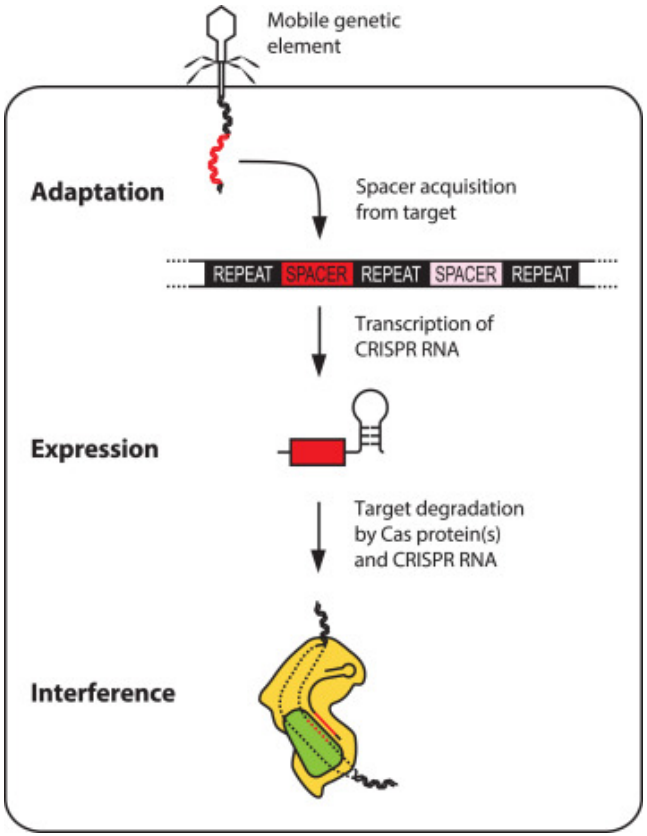


Figure 2: CRISPR-Cas Mechanism

- Adaptive Bacterial Immune System
- Requires CRISPR array and Cas proteins

CRISPR-Cas In Bacteria

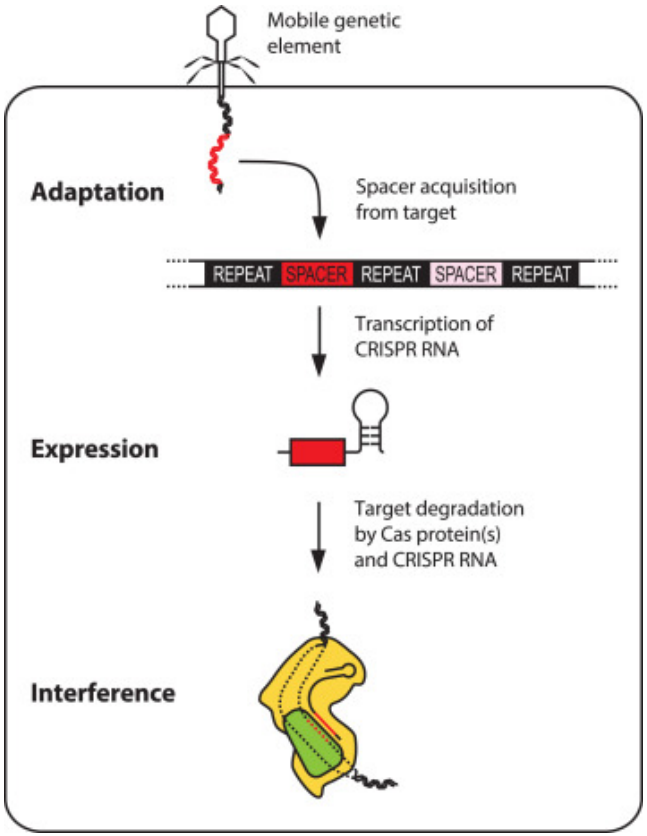


Figure 2: CRISPR-Cas Mechanism

- Adaptive Bacterial Immune System
- Requires CRISPR array and Cas proteins
- Steps

CRISPR-Cas In Bacteria

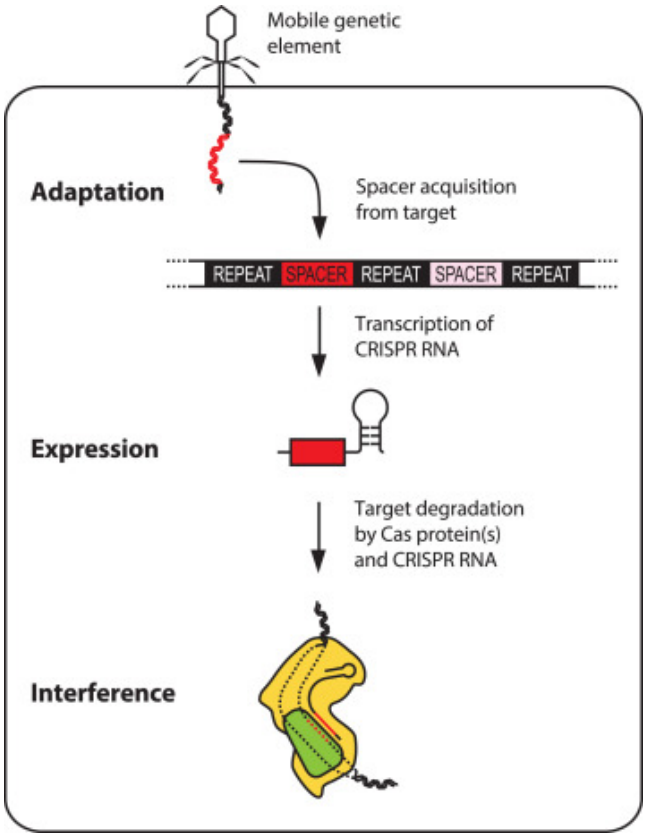


Figure 2: CRISPR-Cas Mechanism

- Adaptive Bacterial Immune System
- Requires CRISPR array and Cas proteins
- Steps
  1. Exposure

CRISPR-Cas In Bacteria

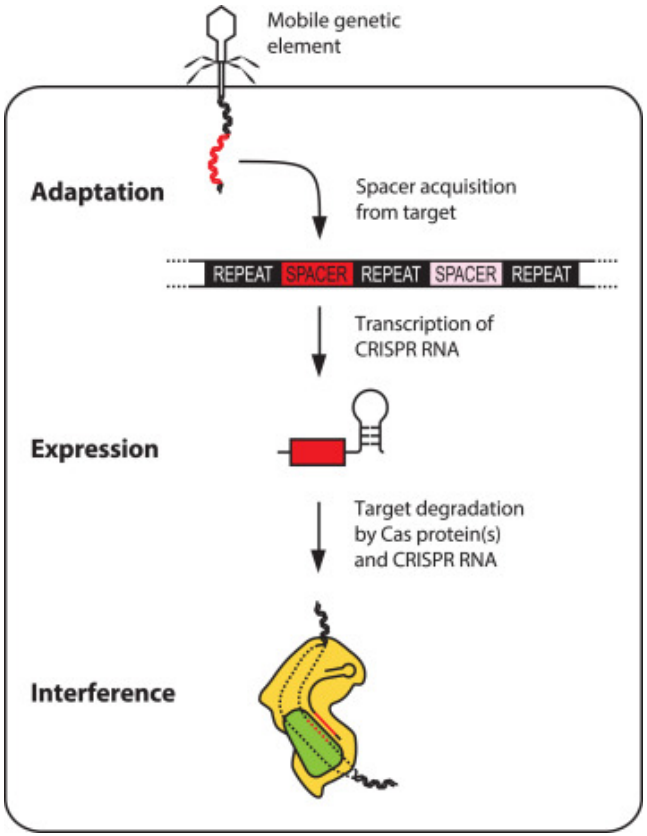


Figure 2: CRISPR-Cas Mechanism

- Adaptive Bacterial Immune System
- Requires CRISPR array and Cas proteins
- Steps
  1. Exposure
  2. Spacer Acquisition

CRISPR-Cas In Bacteria

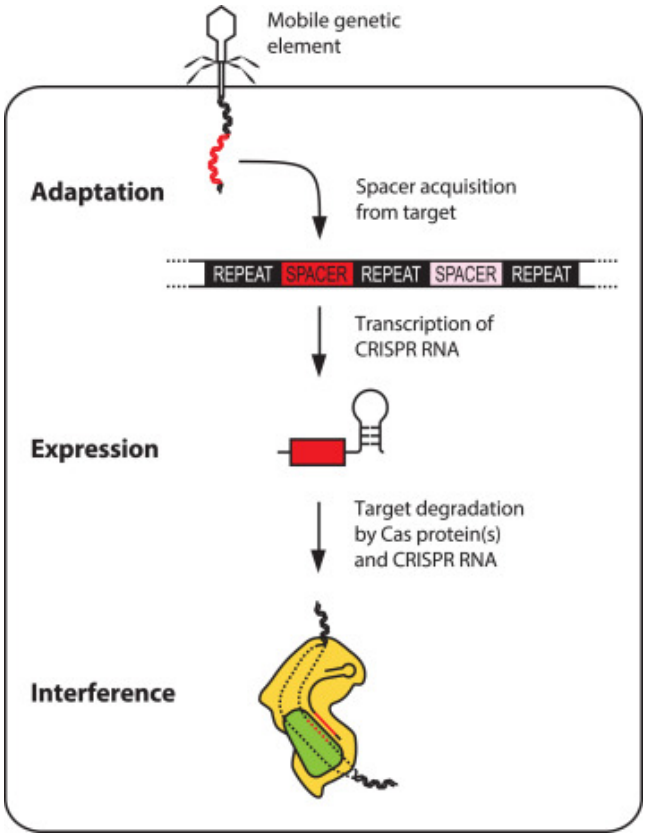


Figure 2: CRISPR-Cas Mechanism

- Adaptive Bacterial Immune System
- Requires CRISPR array and Cas proteins
- Steps
  1. Exposure
  2. Spacer Acquisition
  3. Targeted degradation in next exposure

CRISPR-Cas In Bacteria

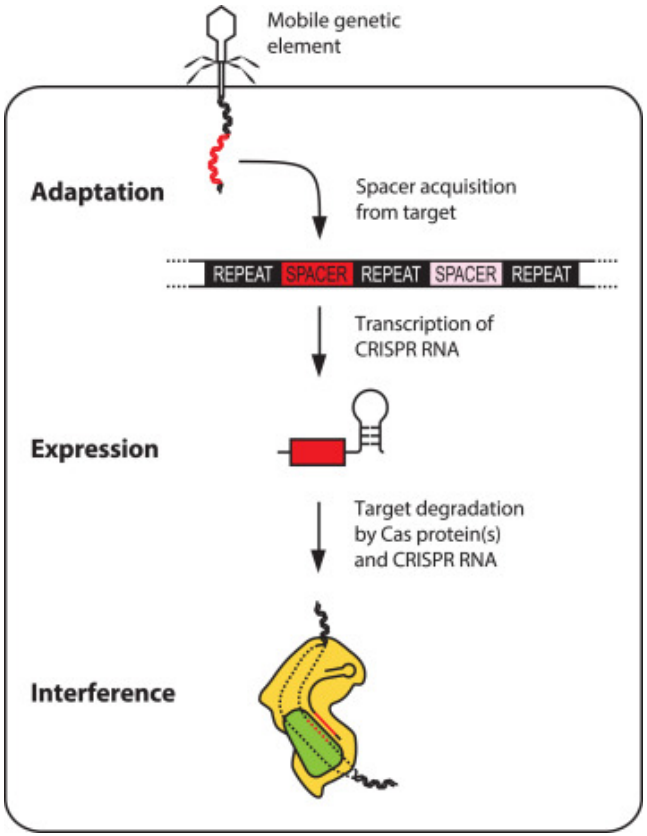


Figure 2: CRISPR-Cas Mechanism

- Adaptive Bacterial Immune System
- Requires CRISPR array and Cas proteins
- Steps
  1. Exposure
  2. Spacer Acquisition
  3. Targeted degradation in next exposure
- Protects against "foreign" DNA, but can acquire **any** DNA as a spacer

Fitness Trade-offs

- HGT can help gain genes increase fitness in a specific environment faster than mutation

Fitness Trade-offs

- HGT can help gain genes increase fitness in a specific environment faster than mutation
- CRISPR can block phage mediated HGT or uptake of environmental DNA

Fitness Trade-offs

- HGT can help gain genes increase fitness in a specific environment faster than mutation
- CRISPR can block phage mediated HGT or uptake of environmental DNA
- HGT can result in genome disruption or gaining toxic gene products

Fitness Trade-offs

- HGT can help gain genes increase fitness in a specific environment faster than mutation
- CRISPR can block phage mediated HGT or uptake of environmental DNA
- HGT can result in genome disruption or gaining toxic gene products
- metabolic cost to maintain CRISPR or HGT machinery expression

Fitness Trade-offs

- HGT can help gain genes increase fitness in a specific environment faster than mutation
- CRISPR can block phage mediated HGT or uptake of environmental DNA
- HGT can result in genome disruption or gaining toxic gene products
- metabolic cost to maintain CRISPR or HGT machinery expression
- CRISPR systems can also be transferred between bacteria via plasmids

Model

Genotypes

Allele		Description
Major	Minor	
R	r	has/does not have resistance gene
H	h	HGT machinery is expressed/not expressed
C	c	CRISPR-Cas is expressed/not expressed

Table 1: Allele definitions

Basics

- Haploid population

Basics

- Haploid population
- infinite population

Basics

- Haploid population
- infinite population
- Each generation we have 1) gene transfer 2) mutation and 3) selection

Basics

- Haploid population
- infinite population
- Each generation we have 1) gene transfer 2) mutation and 3) selection
  - no sexual reproduction, consider gene transfer step

Basics

- Haploid population
- infinite population
- Each generation we have 1) gene transfer 2) mutation and 3) selection
  - no sexual reproduction, consider gene transfer step
  - Gene transfer is analogous to oblique learning from Fogarty L. 2018

Basics

- Haploid population
- infinite population
- Each generation we have 1) gene transfer 2) mutation and 3) selection
  - no sexual reproduction, consider gene transfer step
  - Gene transfer is analogous to oblique learning from Fogarty L. 2018
  - mutation is $r \rightarrow R$ or $R \rightarrow r$

Environment Switching

- We consider 3 different environments

Environment Switching

- We consider 3 different environments
  - E_n: Neutral, no threats

Environment Switching

- We consider 3 different environments
  - E_n: Neutral, no threats
  - E_b: Bacteriophage, increased risk of phage contact

Environment Switching

- We consider 3 different environments
  - E_n: Neutral, no threats
  - E_b: Bacteriophage, increased risk of phage contact
  - E_a: Antibiotic, increased risk of antibiotic contact

Environment Switching

- We consider 3 different environments
  - E_n: Neutral, no threats
  - E_b: Bacteriophage, increased risk of phage contact
  - E_a: Antibiotic, increased risk of antibiotic contact
- **Singular Threat:** Singular threat event (antibiotic dosage or phage outbreak)

Environment Switching

- We consider 3 different environments
  - E_n: Neutral, no threats
  - E_b: Bacteriophage, increased risk of phage contact
  - E_a: Antibiotic, increased risk of antibiotic contact
- **Singular Threat:** Singular threat event (antibiotic dosage or phage outbreak)
- **Cyclical Threat:** Regular threat events every $2l$ generations, lasting l generations.

Environment Switching

- We consider 3 different environments
  - E_n: Neutral, no threats
  - E_b: Bacteriophage, increased risk of phage contact
  - E_a: Antibiotic, increased risk of antibiotic contact

- **Singular Threat:** Singular threat event (antibiotic dosage or phage outbreak)
- **Cyclical Threat:** Regular threat events every $2l$ generations, lasting l generations.
- **Alternating Threat:** Same as Cyclical Threat model but switching between threats threats each event.

Fitness

Genotype	*Environment*		
	E_n	E_b	E_a
RCH	$1 - 2s_m$	$(1 + s_p)(1 - 2s_m)$	$(1 + s_p)(1 - 2s_m)$
RCh	$1 - s_m$	$(1 + s_p)(1 - s_m)$	$(1 + s_p)(1 - s_m)$
RcH	$1 - s_m$	$1 - s_m$	$(1 + s_p)(1 - s_m)$
Rch	1	1	$1 + s_p$
rCH	$1 - 2s_m$	$(1 + s_p)(1 - 2s_m)$	$1 - 2s_m$
rCh	$1 - s_m$	$(1 + s_p)(1 - s_m)$	$1 - s_m$
rcH	$1 - s_m$	$1 - s_m$	$1 - s_m$
rch	1	1	1

Table 2: Relative fitness values for each genotype in each environment

- s_m reflects cost of maintaining HGT/CRISPR, $s_m << s_p$

Genotype Frequencies

- g represents each genotype

Genotype Frequencies

- g represents each genotype

1. **Gene Transfer:** $x_g^t = x_g + \sum_{x_R} x_{\neg g} x_R h(x_{\neg g}, x_R)$

Genotype Frequencies

- g represents each genotype

1. **Gene Transfer:** $x_g^t = x_g + \sum_{x_R} x_{\neg g} x_R h(x_{\neg g}, x_R)$
   - if $g = RCH$ then $\neg g = rCH$, same for CH, cH, Ch, ch

Genotype Frequencies

- g represents each genotype

1. **Gene Transfer:** $x_g^t = x_g + \sum_{x_R} x_{\neg g} x_R h(x_{\neg g}, x_R)$
   - if $g = RCH$ then $\neg g = rCH$, same for CH, cH, Ch, ch
   - defined for R genotypes (x_R), for r genotypes subtract the sum

Genotype Frequencies

- g represents each genotype

1. **Gene Transfer:** $x_g^t = x_g + \sum_{x_R} x_{\neg g} x_R h(x_{\neg g}, x_R)$
   - if $g = RCH$ then $\neg g = rCH$, same for CH, cH, Ch, ch
   - defined for R genotypes (x_R), for r genotypes subtract the sum
   - $h()$ probability of transfer, increases for each H allele $(g_h, g_H, 2g_H)$

Genotype Frequencies

- g represents each genotype

1. **Gene Transfer:** $x_g^t = x_g + \sum_{x_R} x_{\neg g} x_R h(x_{\neg g}, x_R)$
   - if $g = RCH$ then $\neg g = rCH$, same for CH, cH, Ch, ch
   - defined for R genotypes (x_R), for r genotypes subtract the sum
   - $h()$ probability of transfer, increases for each H allele ($g_h, g_H, 2g_H$)
2. **Mutation:** $x_g^s = (1 - \mu(g))x_g^t + \mu(g)x_{\neg g}^t$

Genotype Frequencies

- g represents each genotype

1. **Gene Transfer:** $x_g^t = x_g + \sum_{x_R} x_{\neg g} x_R h(x_{\neg g}, x_R)$
   - if $g = RCH$ then $\neg g = rCH$, same for CH, cH, Ch, ch
   - defined for R genotypes (x_R), for r genotypes subtract the sum
   - $h()$ probability of transfer, increases for each H allele ($g_h, g_H, 2g_H$)
2. **Mutation:** $x_g^s = (1 - \mu(g))x_g^t + \mu(g)x_{\neg g}^t$
   - $\mu(g)$ is $\mu_{r \to R}$ for r genotypes and$\mu_{R \to r}$ for R genotypes

Genotype Frequencies

- g represents each genotype

1. **Gene Transfer:** $x_g^t = x_g + \sum_{x_R} x_{\neg g} x_R h(x_{\neg g}, x_R)$
   - if $g = RCH$ then $\neg g = rCH$, same for CH, cH, Ch, ch
   - defined for R genotypes (x_R), for r genotypes subtract the sum
   - $h()$ probability of transfer, increases for each H allele ($g_h, g_H, 2g_H$)
2. **Mutation:** $x_g^s = (1 - \mu(g))x_g^t + \mu(g)x_{\neg g}^t$
   - $\mu(g)$ is $\mu_{r \to R}$ for r genotypes and$\mu_{R \to r}$ for R genotypes
3. **Selection:** $x_g' = \frac{x_g^s f(g)}{\bar{w}}$

Genotype Frequencies

- g represents each genotype

1. **Gene Transfer:** $x_g^t = x_g + \sum_{x_R} x_{\neg g} x_R h(x_{\neg g}, x_R)$
   - if $g = RCH$ then $\neg g = rCH$, same for CH, cH, Ch, ch
   - defined for R genotypes (x_R), for r genotypes subtract the sum
   - $h()$ probability of transfer, increases for each H allele ($g_h, g_H, 2g_H$)
2. **Mutation:** $x_g^s = (1 - \mu(g))x_g^t + \mu(g)x_{\neg g}^t$
   - $\mu(g)$ is $\mu_{r \to R}$ for r genotypes and$\mu_{R \to r}$ for R genotypes
3. **Selection:** $x_g' = \frac{x_g^s f(g)}{\bar{w}}$
   - $f(g)$ picks the correct fitness modifier from Table 2

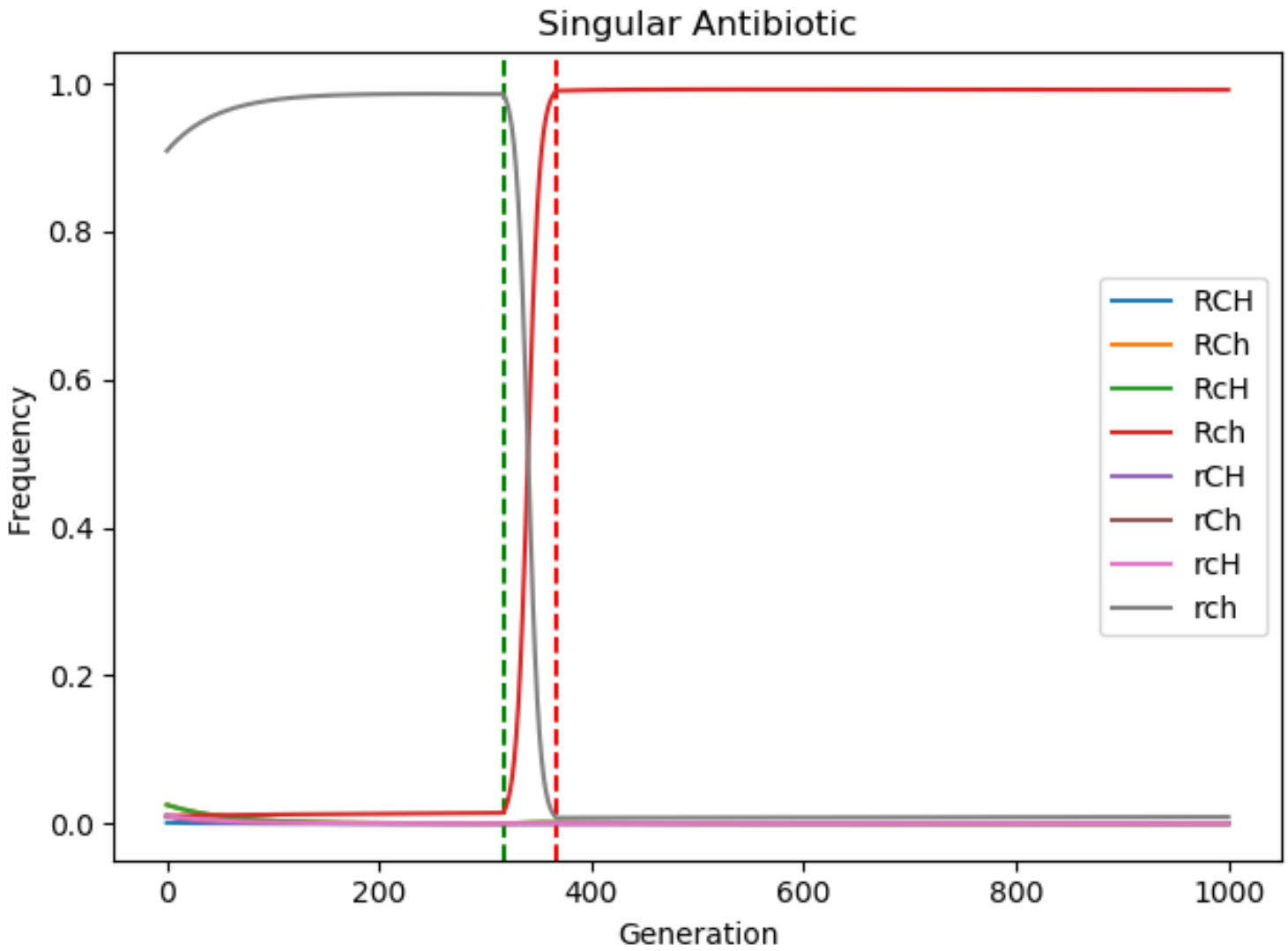
Genotype Frequencies

- g represents each genotype

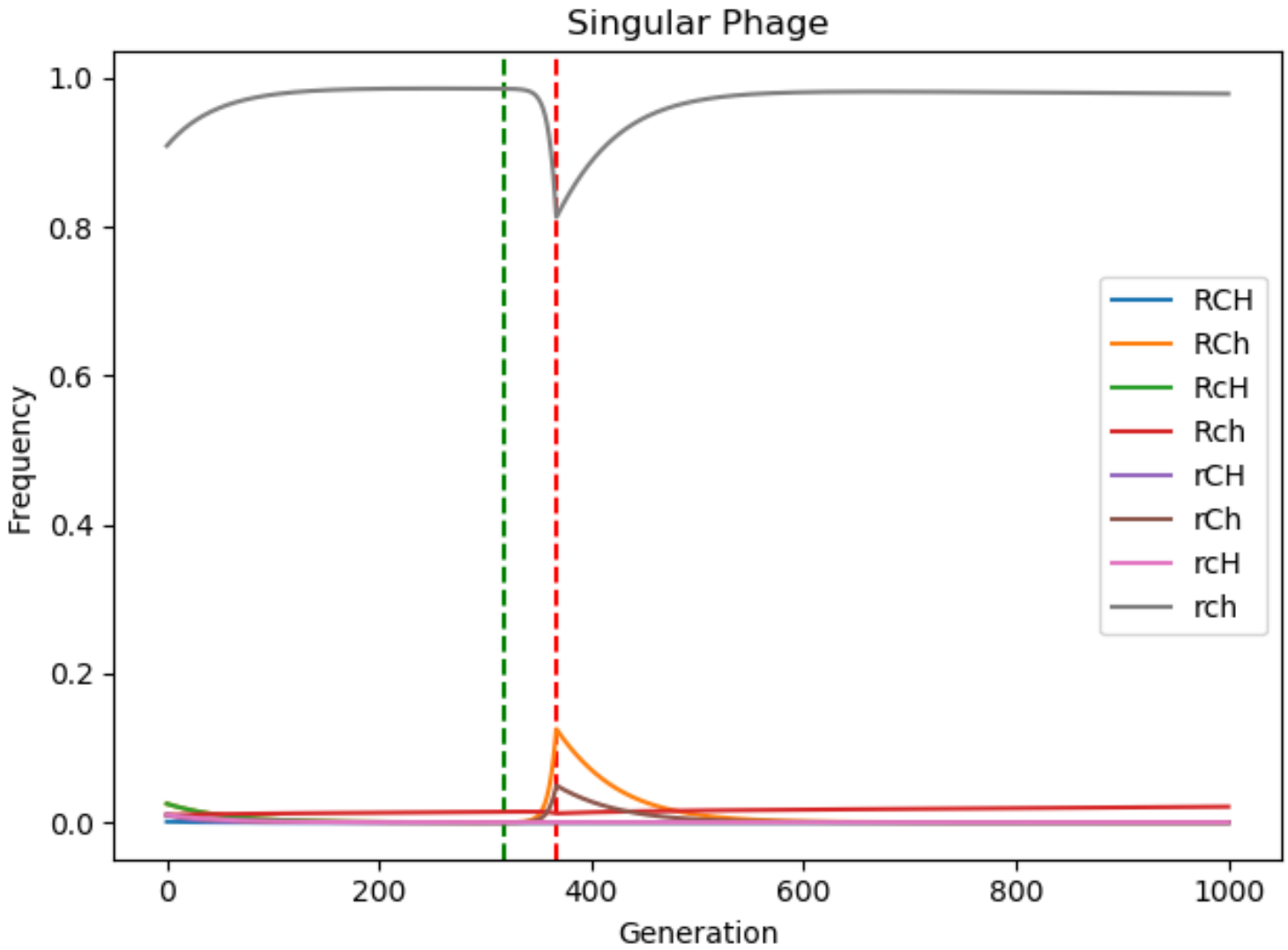
1. **Gene Transfer:** $x_g^t = x_g + \sum_{x_R} x_{\neg g} x_R h(x_{\neg g}, x_R)$
   - if $g = RCH$ then $\neg g = rCH$, same for CH, cH, Ch, ch
   - defined for R genotypes (x_R), for r genotypes subtract the sum
   - $h()$ probability of transfer, increases for each H allele ($g_h, g_H, 2g_H$)
2. **Mutation:** $x_g^s = (1 - \mu(g))x_g^t + \mu(g)x_{\neg g}^t$
   - $\mu(g)$ is $\mu_{r \to R}$ for r genotypes and$\mu_{R \to r}$ for R genotypes
3. **Selection:** $x_g' = \frac{x_g^s f(g)}{\bar{w}}$
   - $f(g)$ picks the correct fitness modifier from Table 2
   - average fitness $\bar{w} = \sum_g x_g^s f(g)$

Results

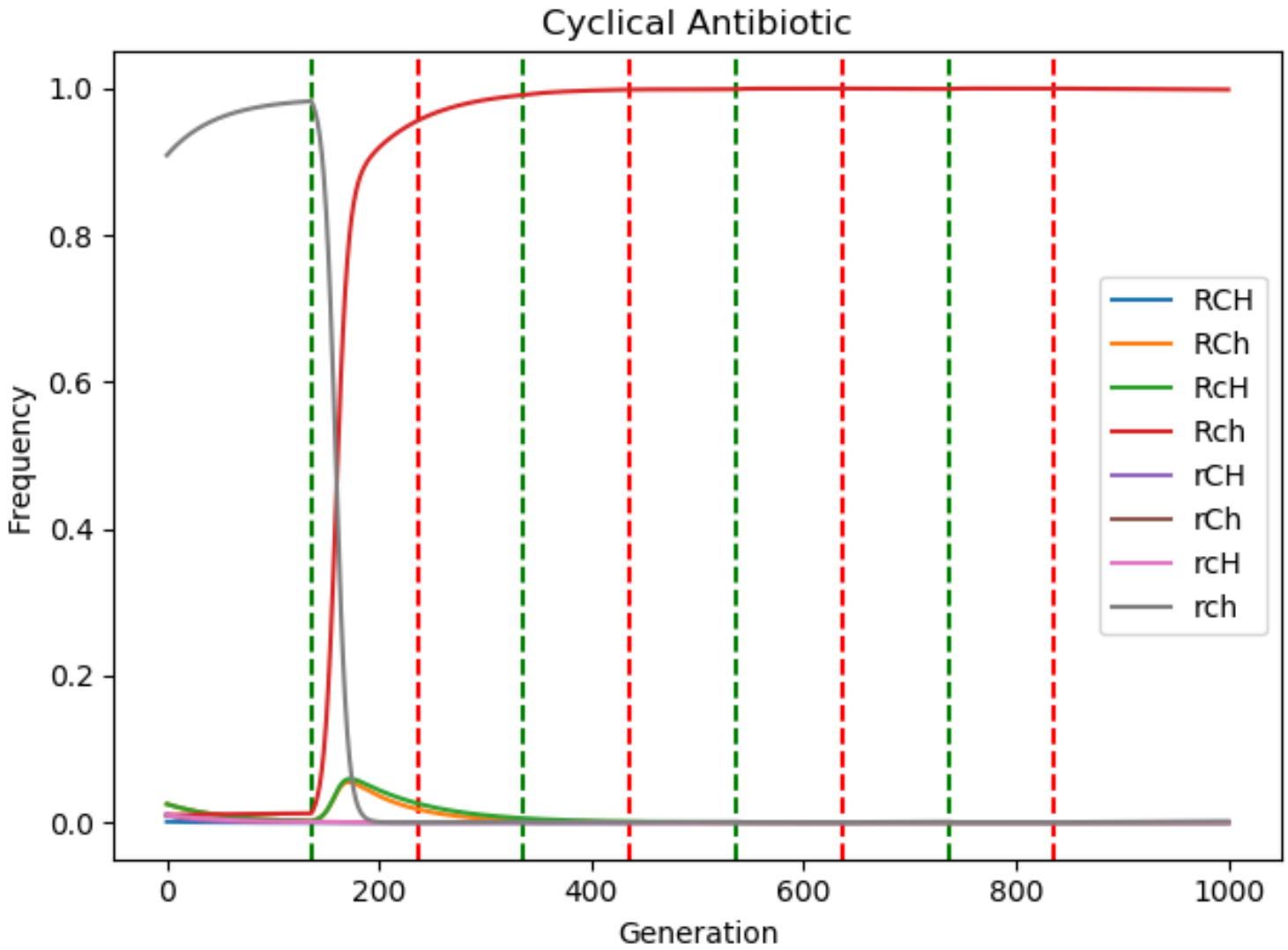
Single Antibiotic Event



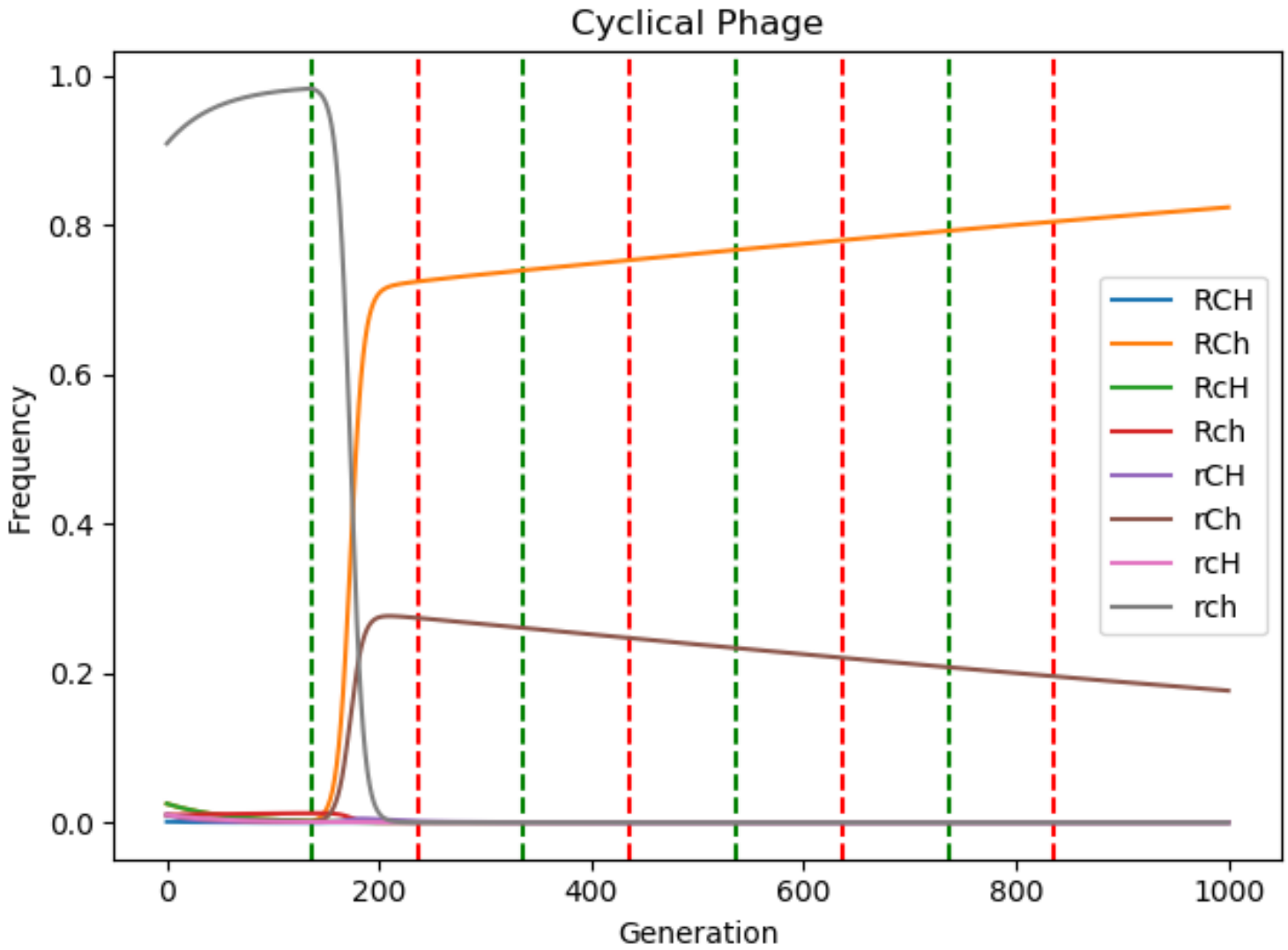
Single Phage Event



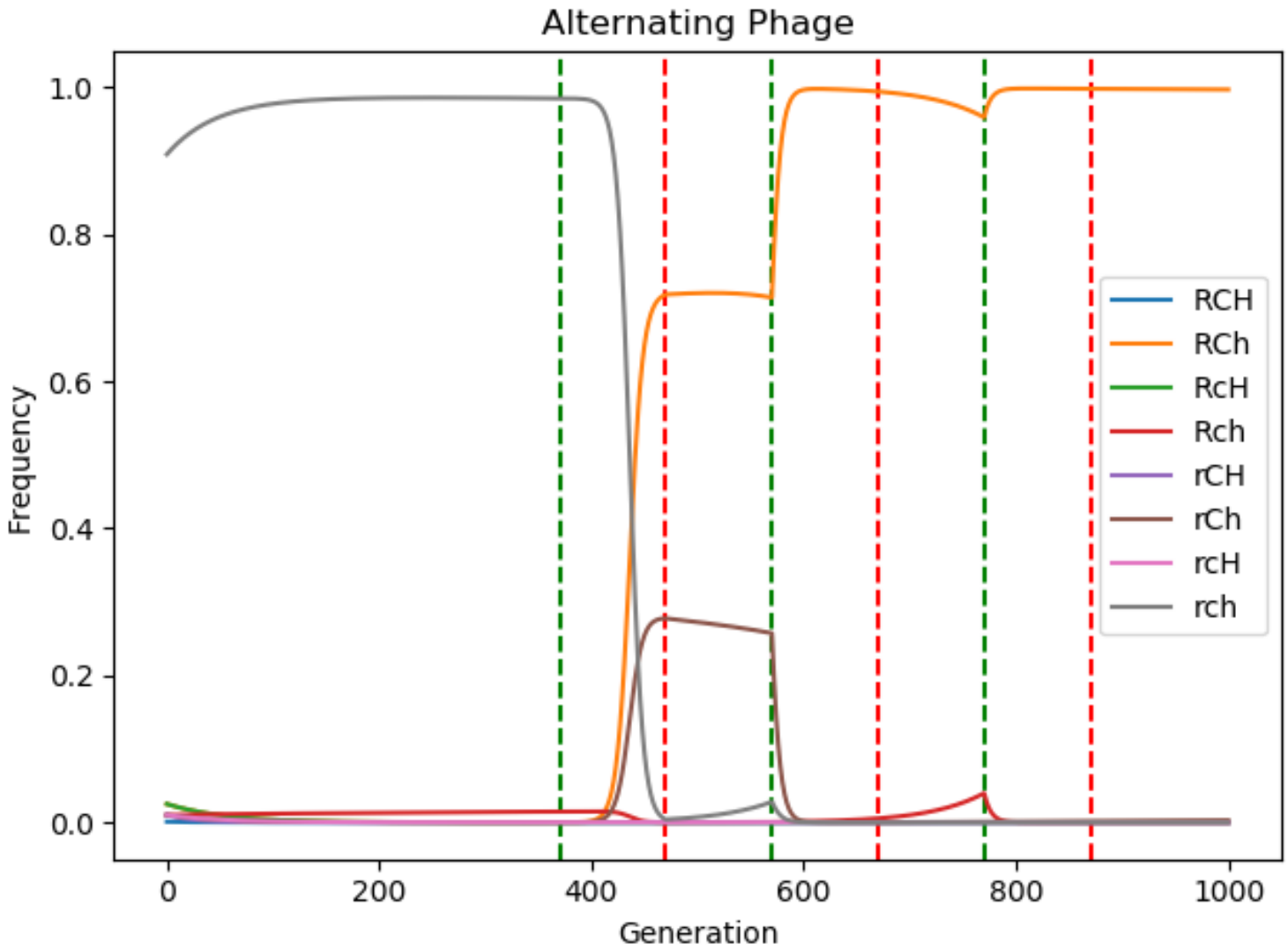
Cyclical Antibiotic Events



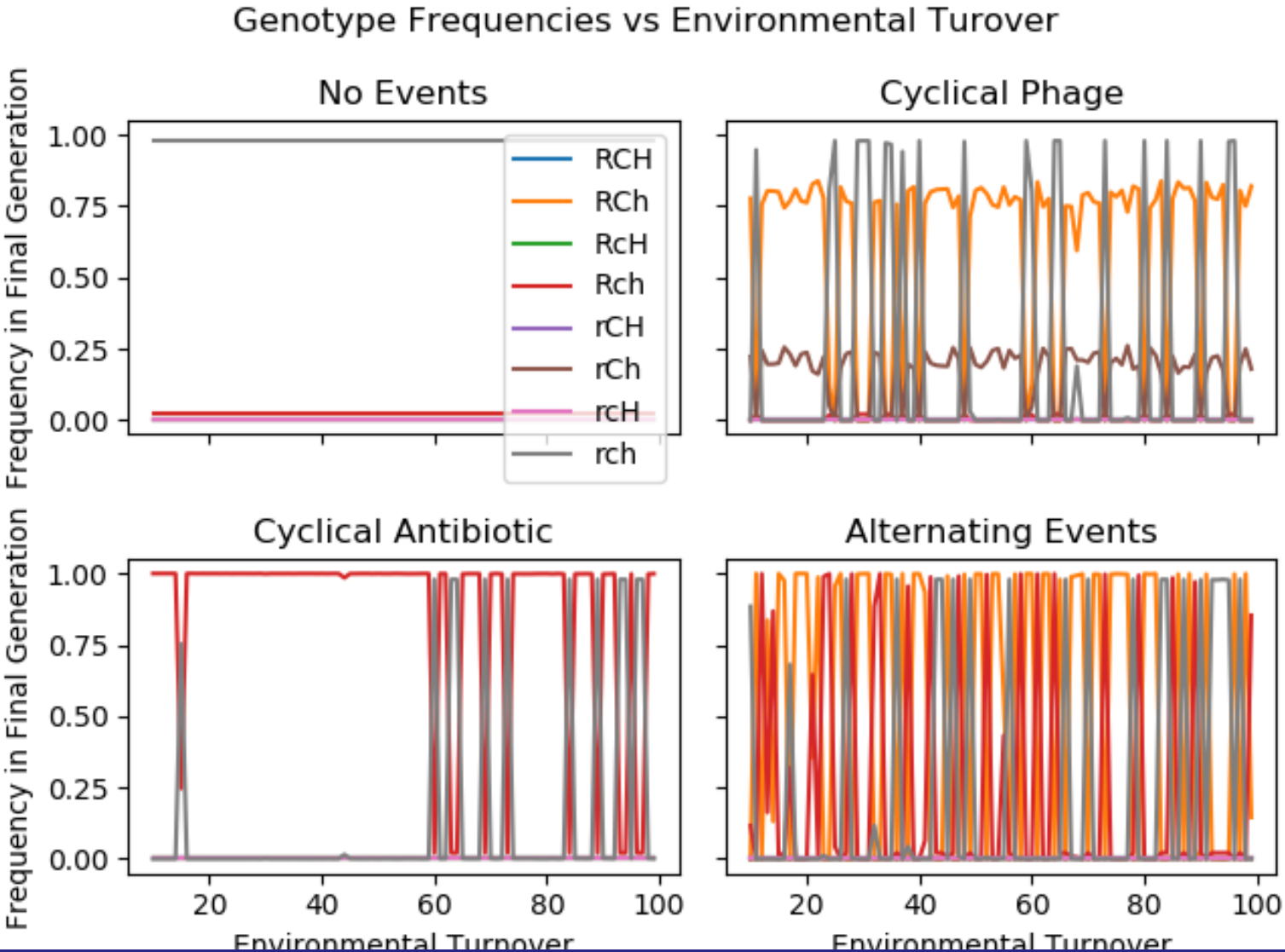
Cyclical Phage Events



Alternating Events



Environmental Stability



Conclusion

Conclusion

- resistance allele dominates even outside of antibiotic pressure

Future Directions

Conclusion

Conclusion

- resistance allele dominates even outside of antibiotic pressure
- environmental turnover rate significantly affects genotype frequencies

Future Directions

Conclusion

Conclusion

- resistance allele dominates even outside of antibiotic pressure
- environmental turnover rate significantly affects genotype frequencies

Future Directions

- explore parameter space and look for empirical justifications

Conclusion

Conclusion

- resistance allele dominates even outside of antibiotic pressure
- environmental turnover rate significantly affects genotype frequencies

Future Directions

- explore parameter space and look for empirical justifications
- model phage population dynamics directly

Conclusion

Conclusion

- resistance allele dominates even outside of antibiotic pressure
- environmental turnover rate significantly affects genotype frequencies

Future Directions

- explore parameter space and look for empirical justifications
- model phage population dynamics directly
- incorporate terms that reflect biological trade-off of HGT/CRISPR