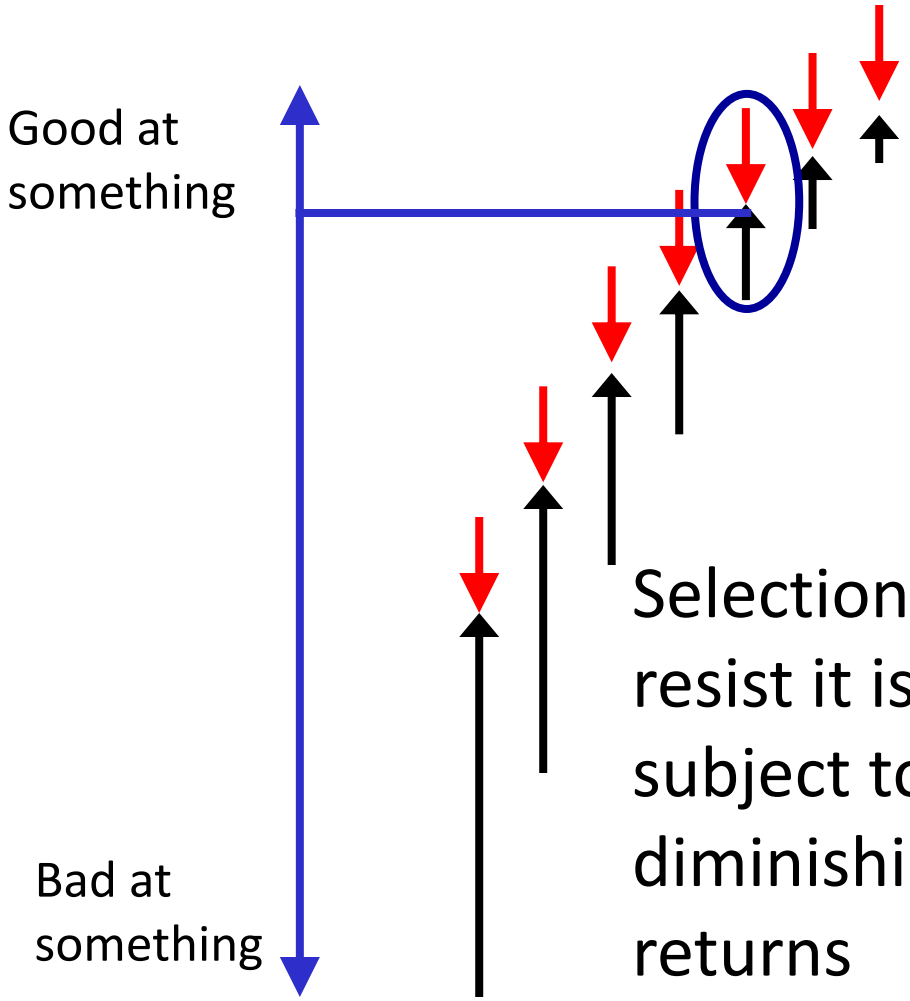
Drift barriers and evolvability

Joanna Masel

University of Arizona

Drift barrier is where selection becomes as weak as mutation bias



Mutation bias towards loss of function erodes a trait

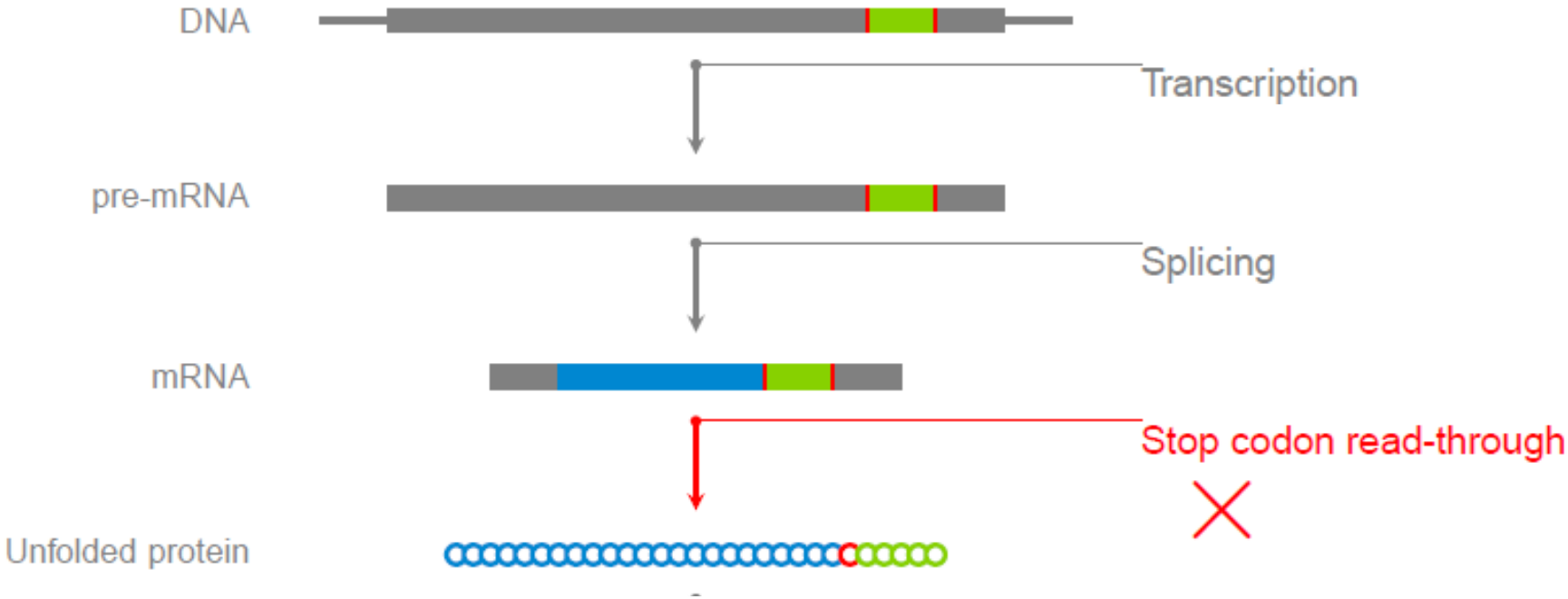
Selection to resist it is subject to diminishing returns

"Effective population size" N_e = location of drift barrier: large N_e allows selection to persist longer as returns diminish

Drift barrier theory

- Large N_e species are good at things, small N_e species are clumsy
- As problems accumulate in small N_e species, a second line of defense evolves: “mutational-hazard theory”
- In this second line of defense, it is the small N_e species that are the most exquisitely adapted, including many aspects of “complexity” at the level of genome architecture

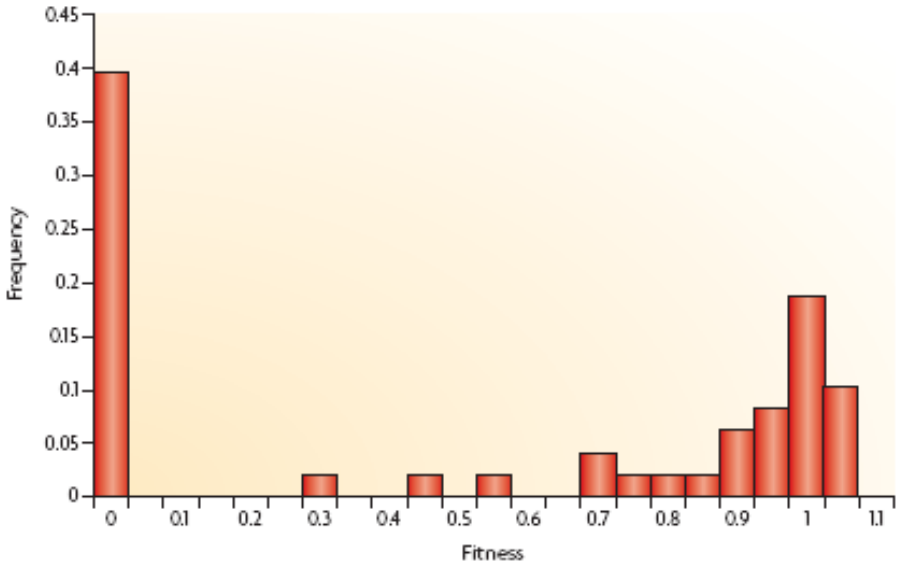
Application: adapting to the threat of failing to stop at a stop codon, expressing a cryptic sequence



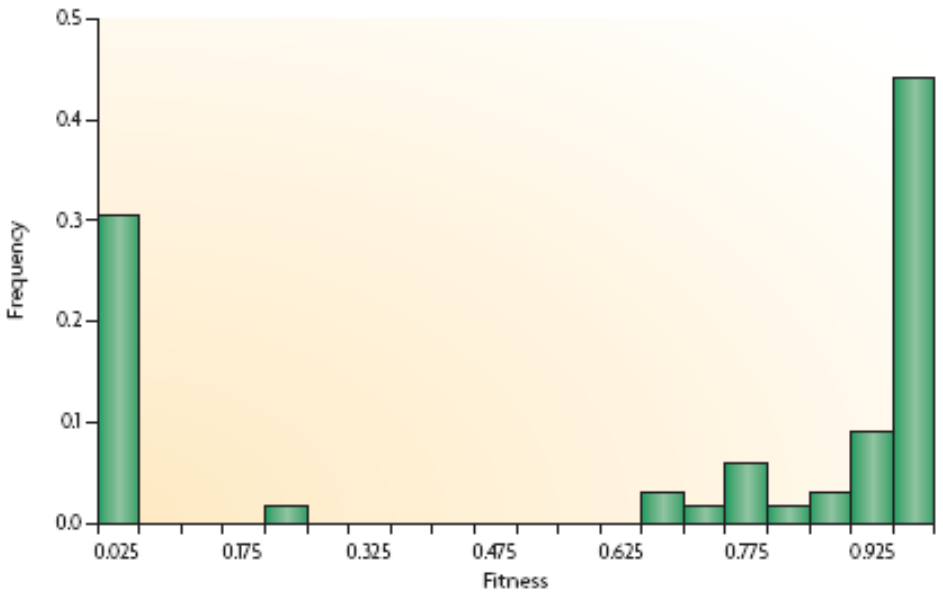
sequences of
cryptic
ther bad or
ess, rarely in

between

Bimodality is like that of the distribution of fitness effects of new mutations



vesicular stomatic virus



yeast

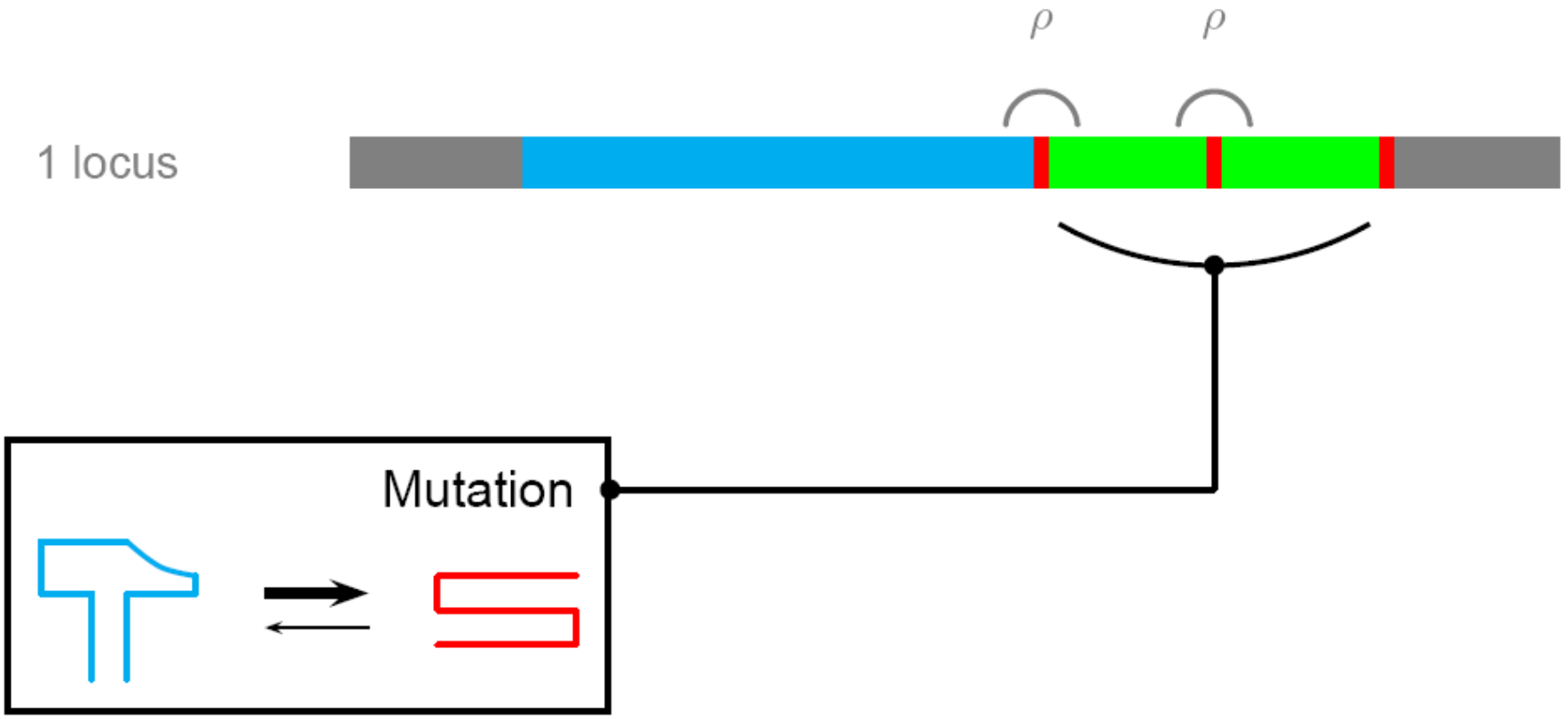
Eyre-Walker & Keightley 2007

Outline

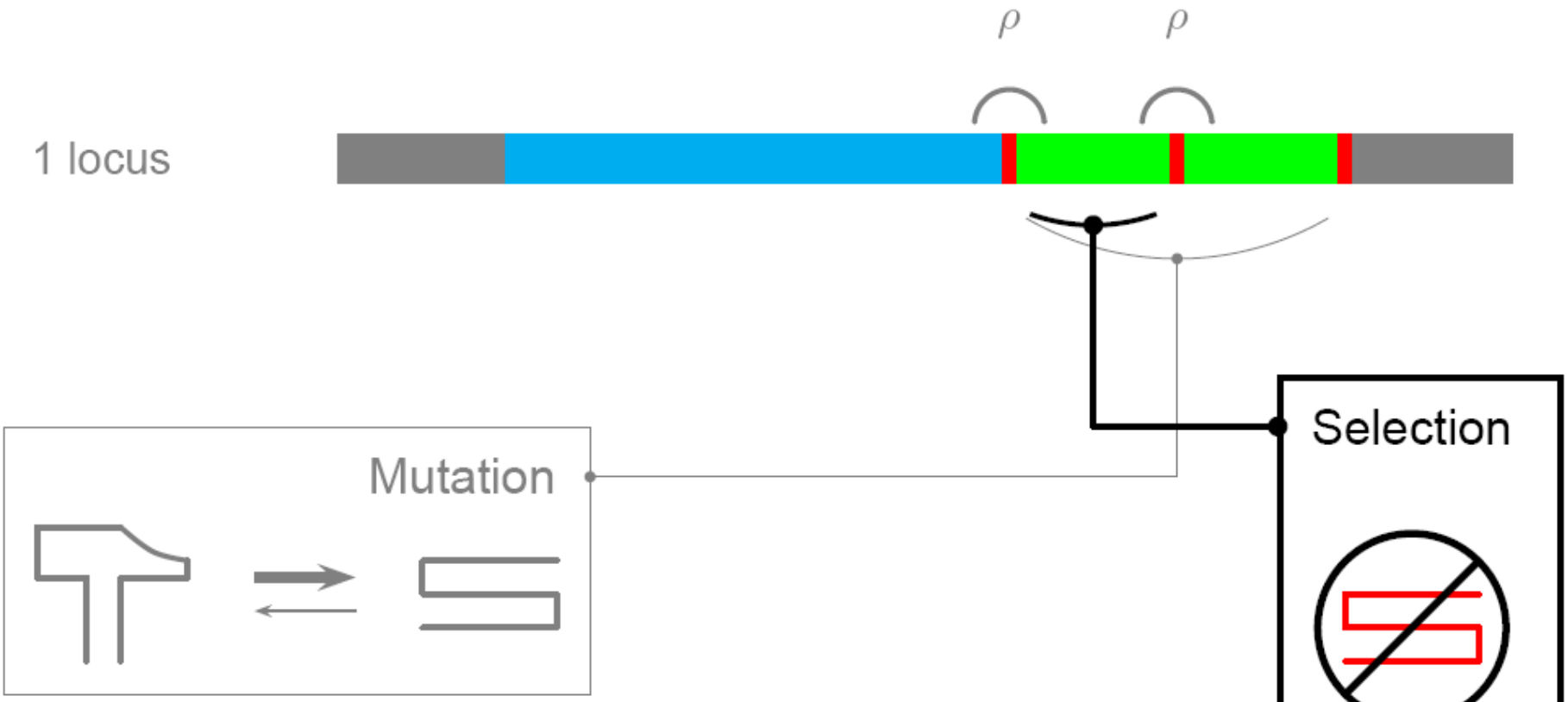
1. 1st line of defense is benign cryptic sequences
   2nd line of defense is low error rate
2. 1st promotes evolvability
3. Evolution of evolvability
4. Data: high errors are associated with high evolvability
5. De novo gene birth

2nd line of defense:
reduce the readthrough error rate ρ

Mutation bias favors misfolding of cryptic sequence



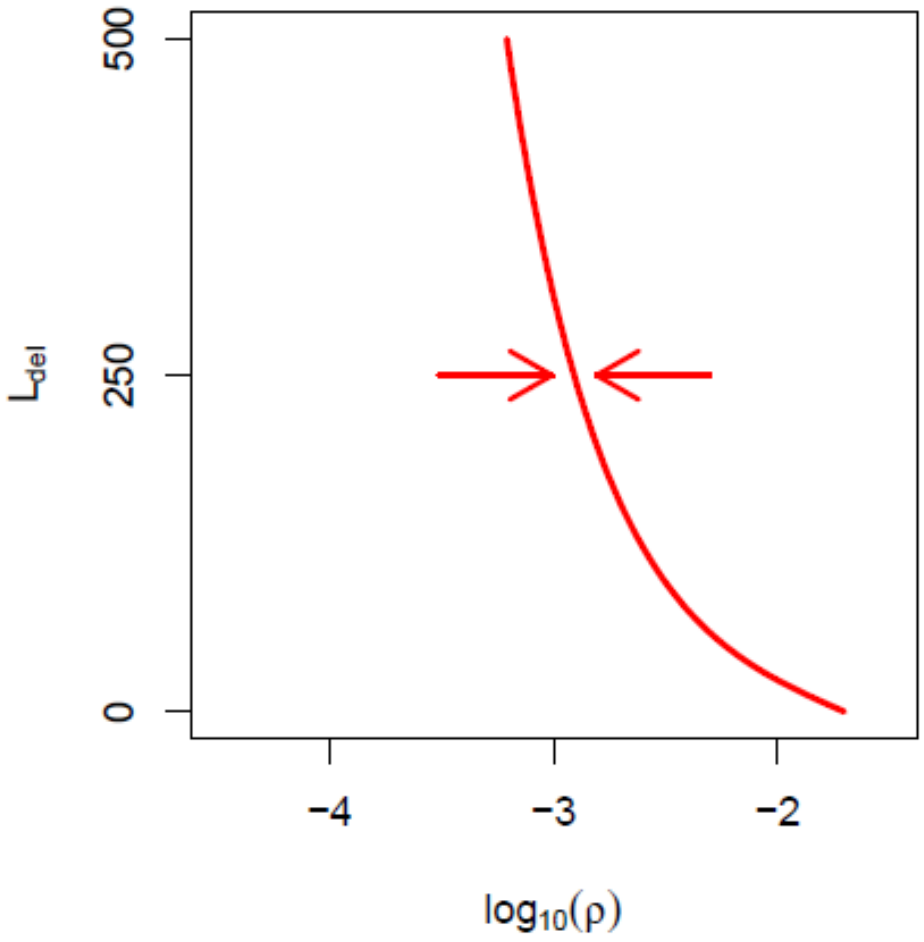
1st line defense: selection for a stable fold even after a readthrough error



1st line of defense, subject to tough drift barrier, is to reduce the number of deleterious cryptic sequences L_{del}

2nd line of defense is costly proofreading to lower ρ

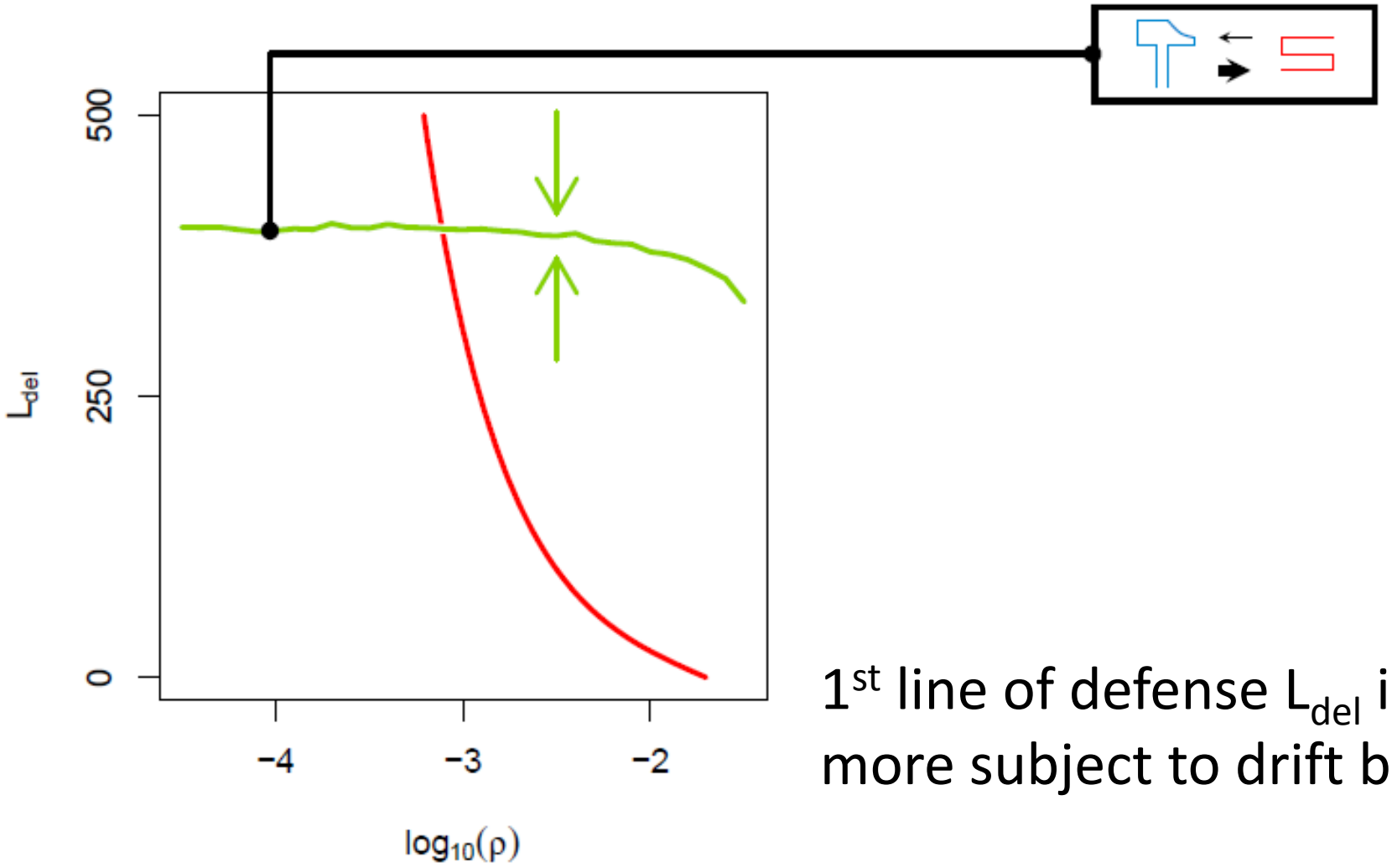
Coevolution of ρ and L_{del}



Strong stabilizing selection on 2nd line of defense ρ

Rajon & Masel (2011)

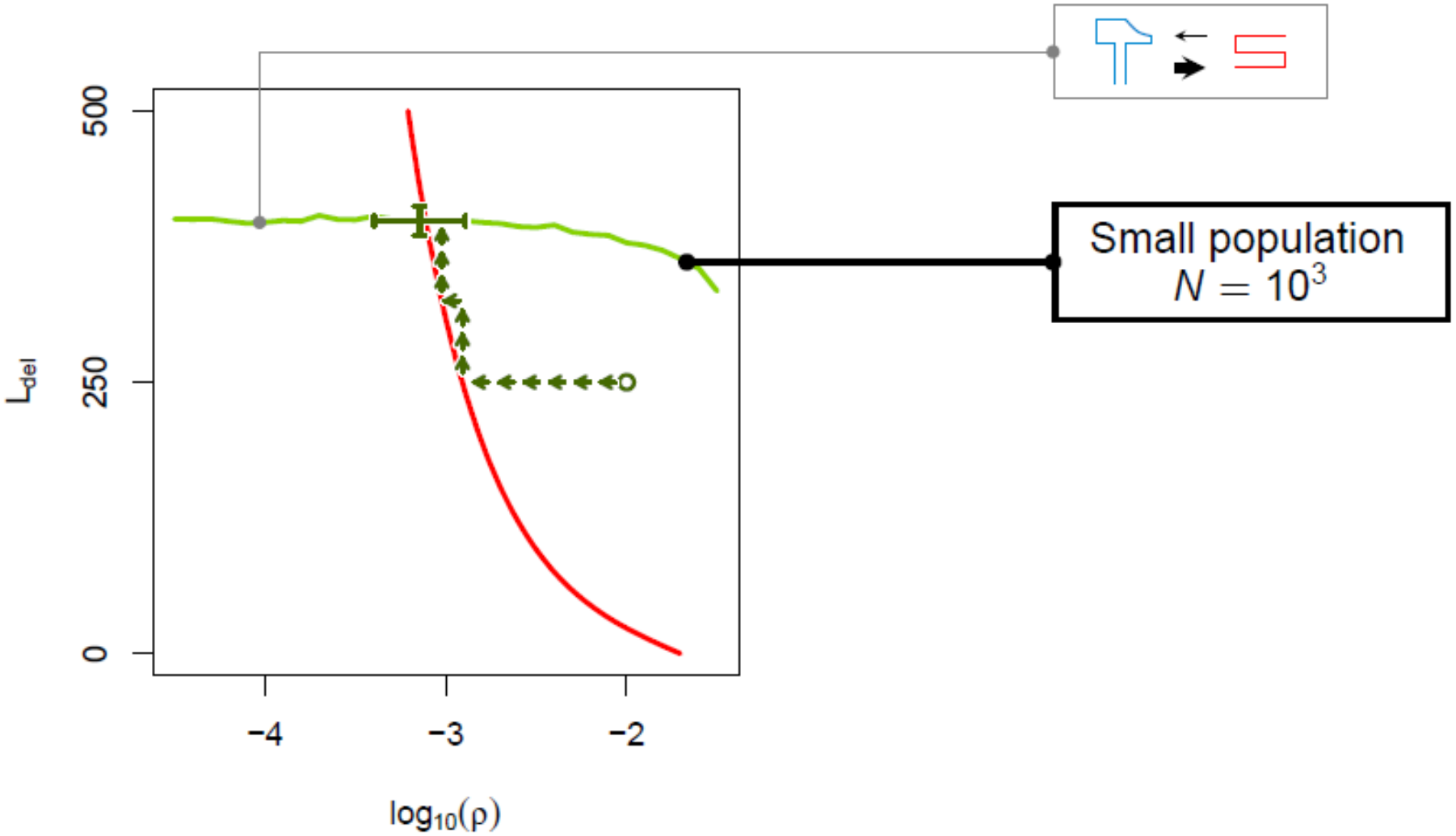
Coevolution of ρ and L_{del}



1st line of defense L_{del} is more subject to drift barrier

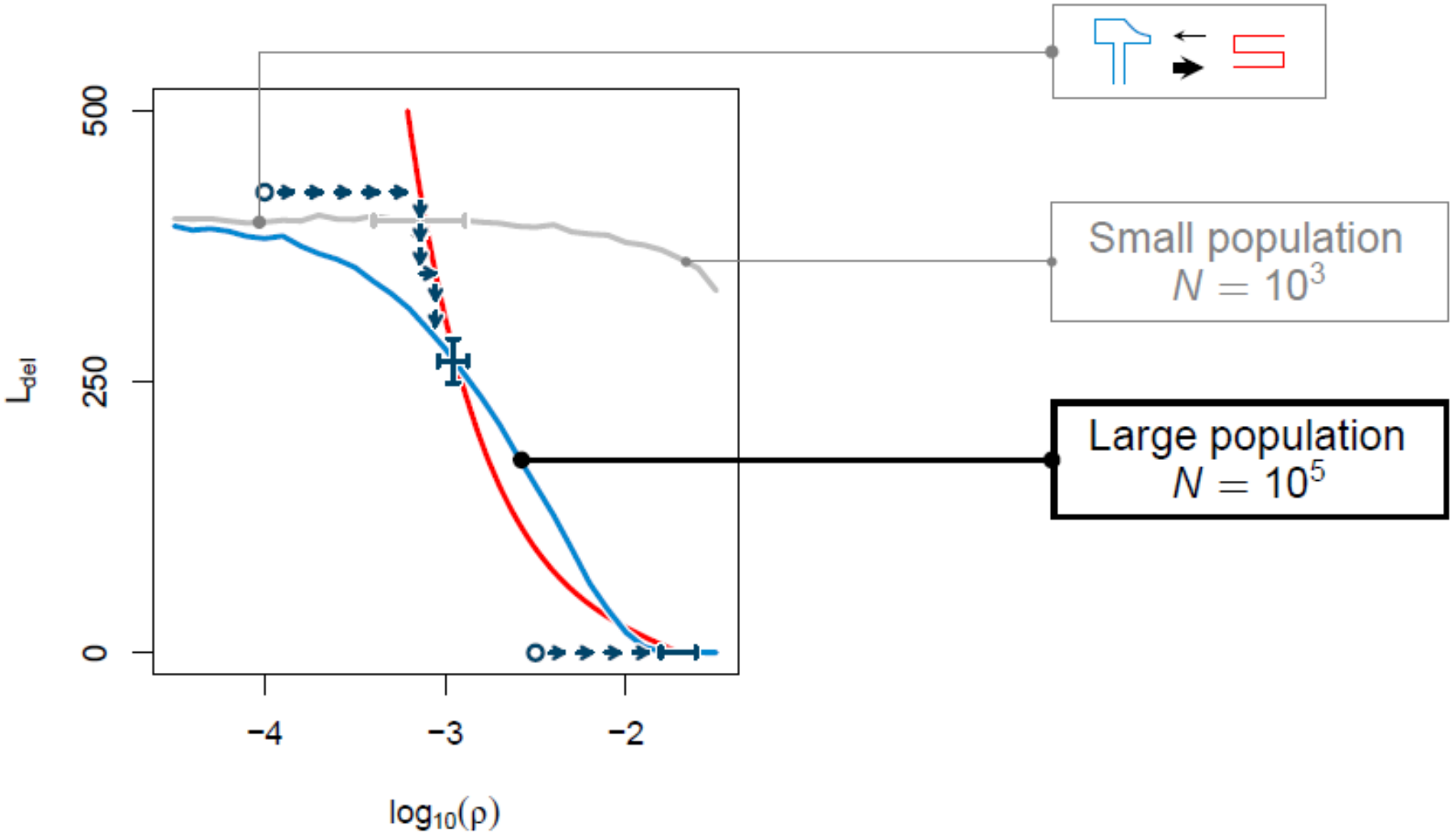
Rajon & Masel (2011)

Coevolution of ρ and L_{del}



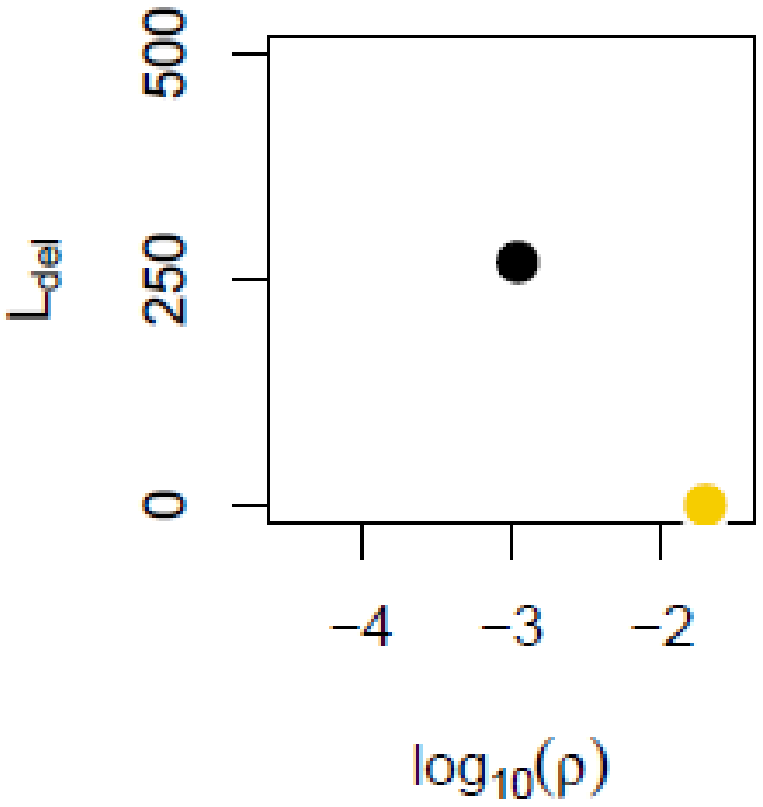
Rajon & Masel (2011)

Two attractors in large populations



Rajon & Masel (2011)

Two strategies are quite different



2 strategies:

●: allowing deleterious sequences, but hiding them

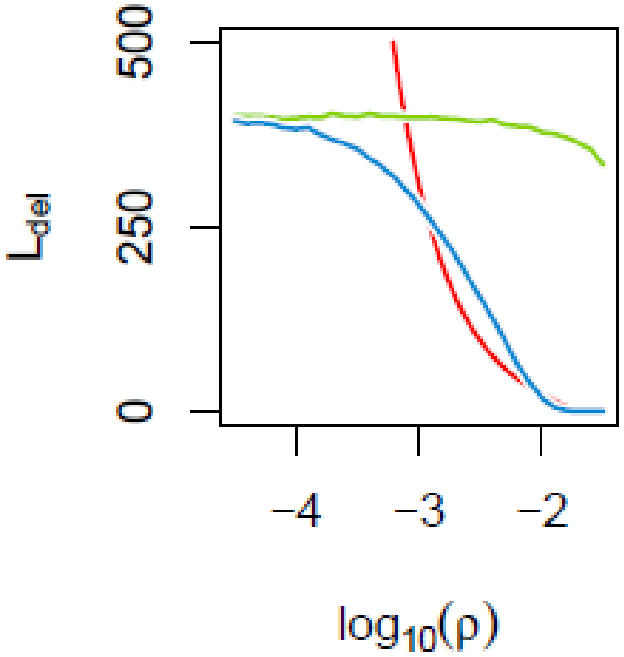
relies on 2nd line defense

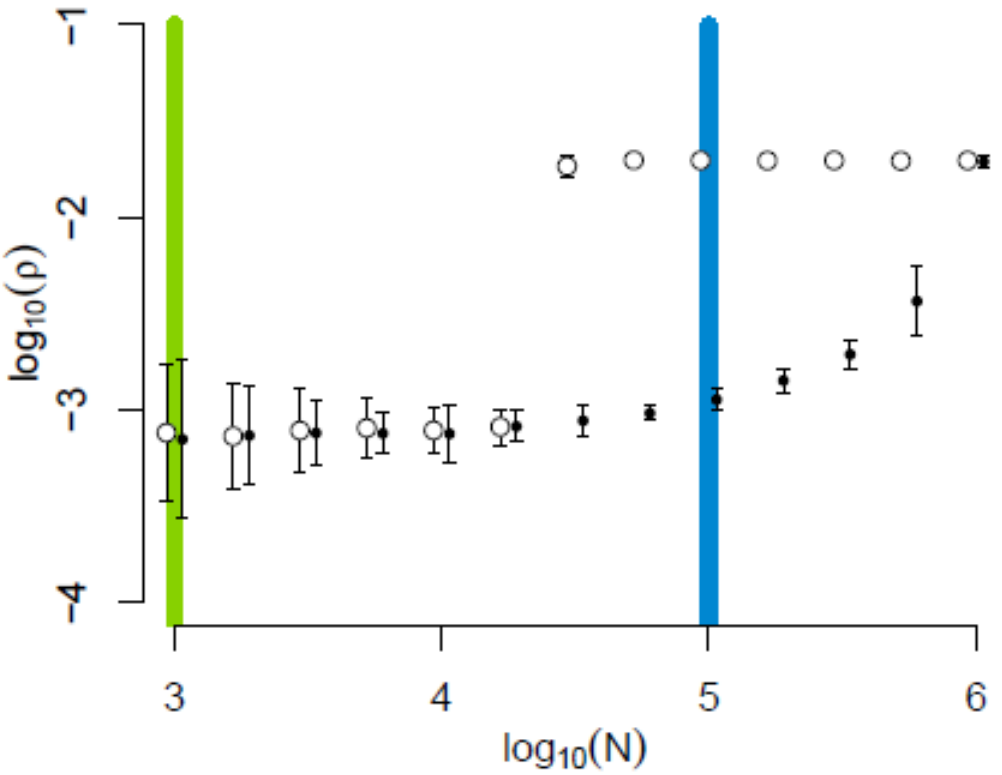
●: eliminating deleterious sequence by expressing them

emphasizes 1st line of defense, superior but subject to tough drift barrier

Rajon & Masel (2011)

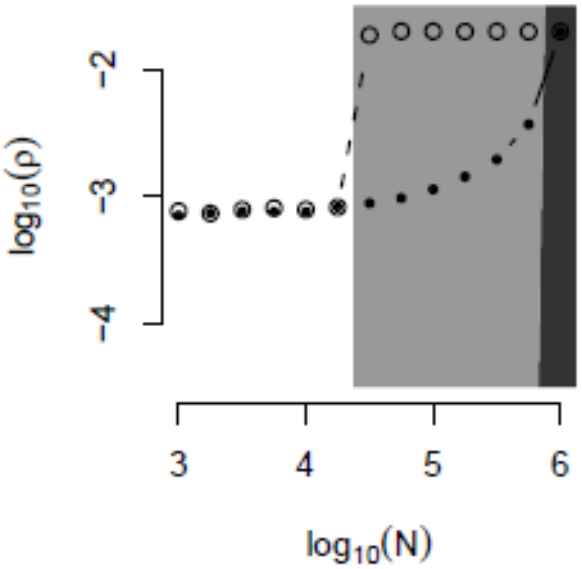
Two attractors for a range of population sizes (i.e. drift barrier locations)

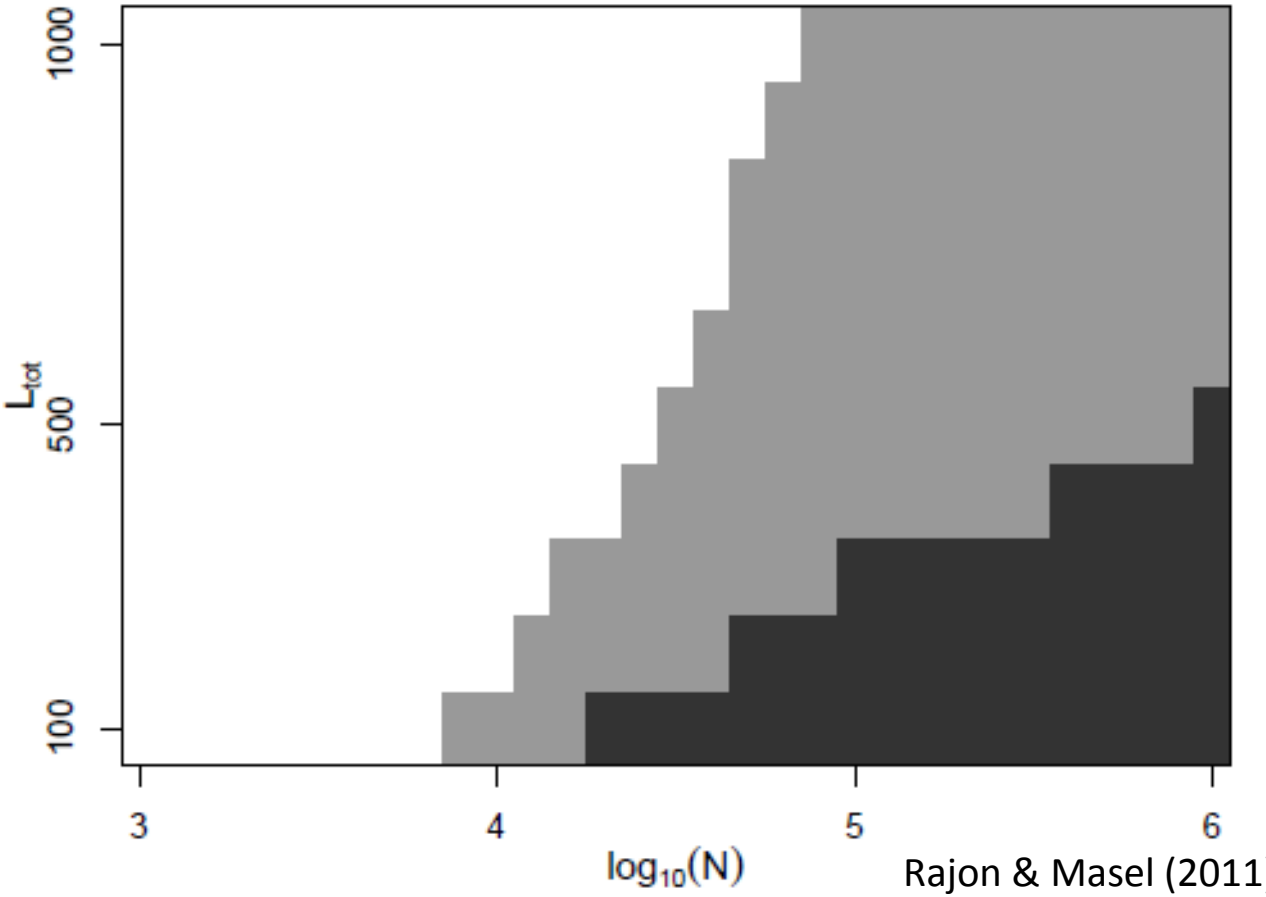




Rajon & Masel (2011)

Larger bistable range with more loci





Rajon & Masel (2011)

Model applies to many kinds of molecular errors

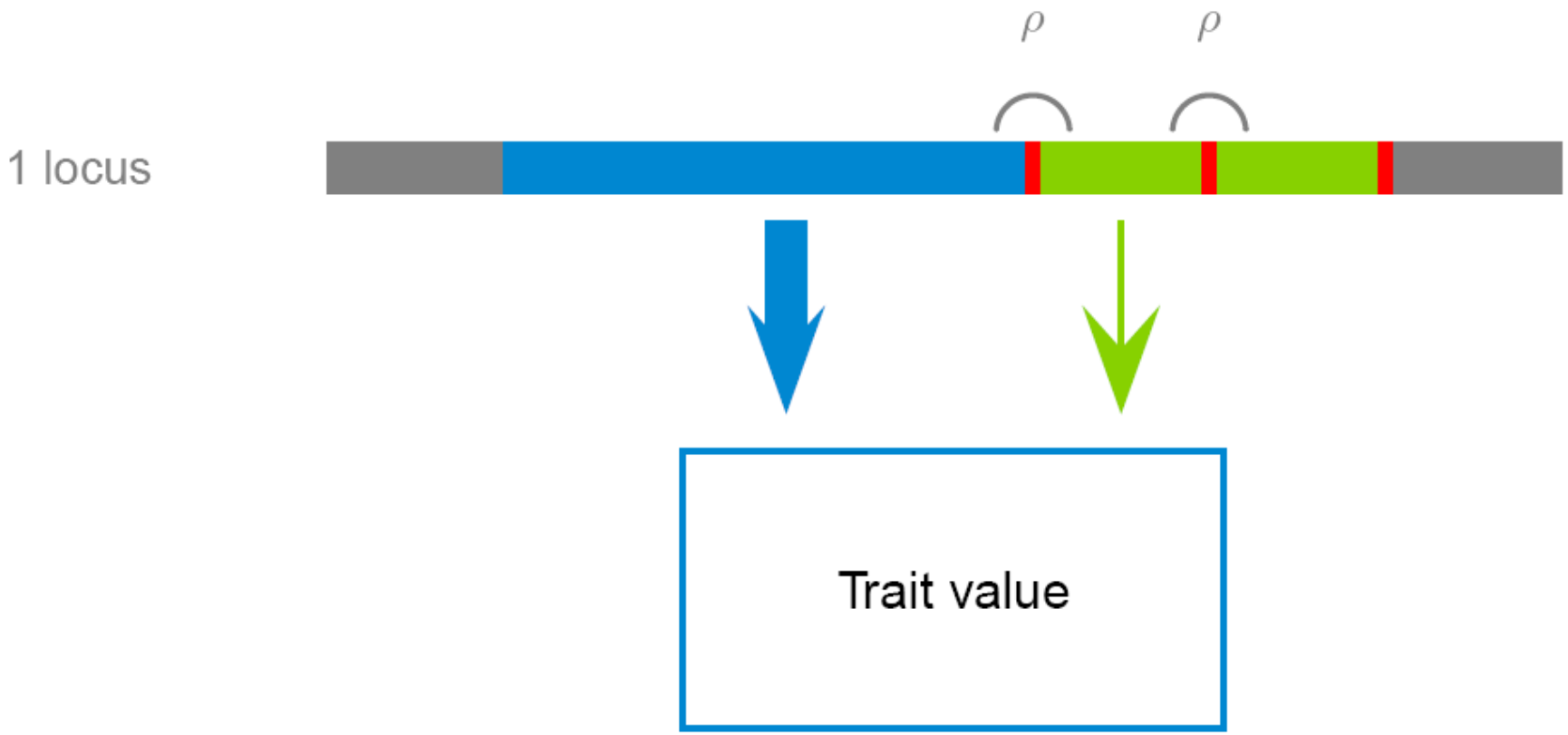
	2nd line defense	1st line defense
Error	Global solution	Local solution
Stop codon readthrough	Accurate ribosome & release factors	Benign 3'UTR

Rajon & Masel (2011)

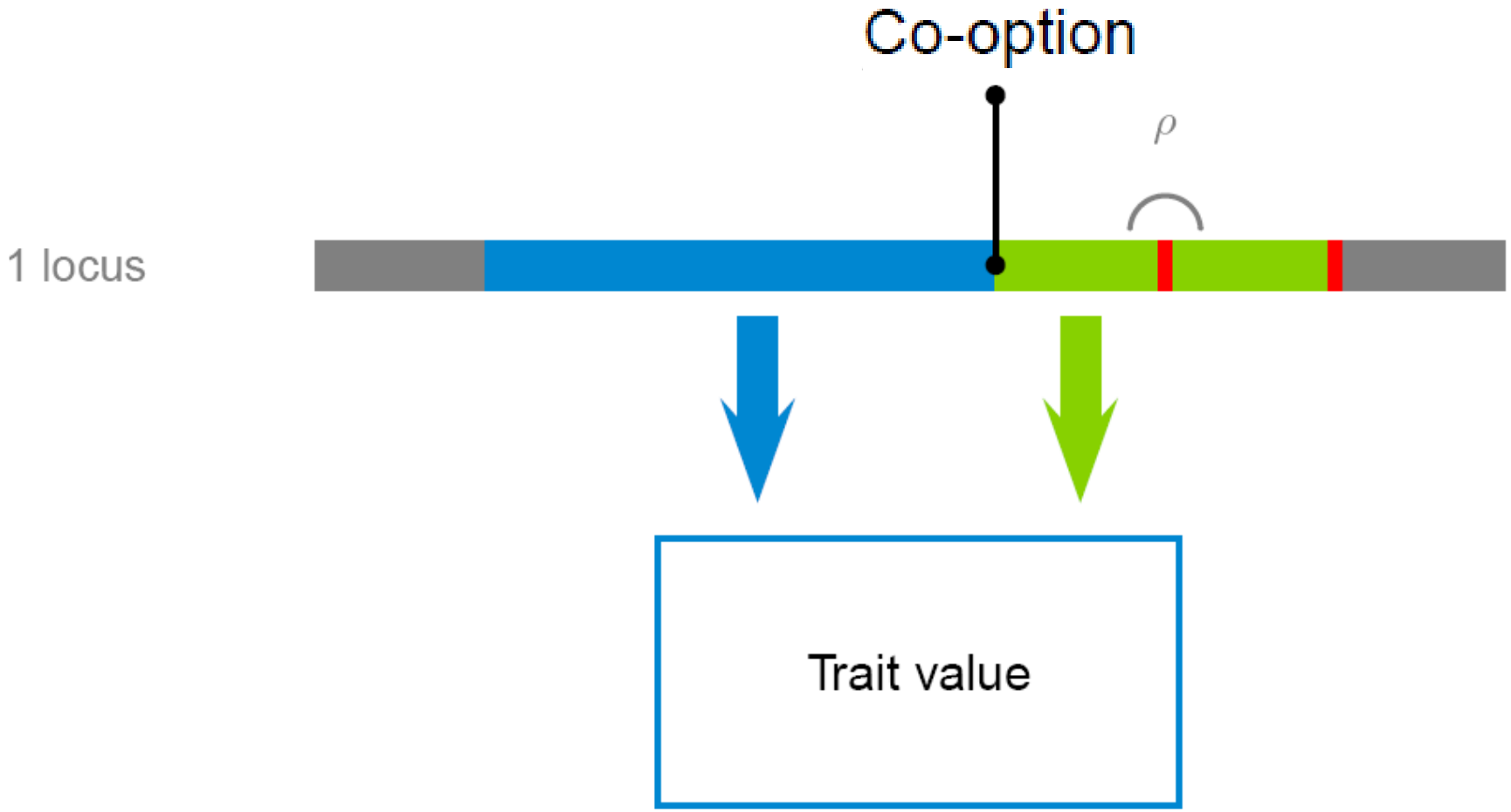
Outline

1. 1st line of defense is benign cryptic sequences (local)
   2nd line of defense is low error rate (global)
2. 1st line of defense promotes evolvability
3. Evolution of evolvability
4. Data: high errors are associated with high evolvability
5. De novo gene birth

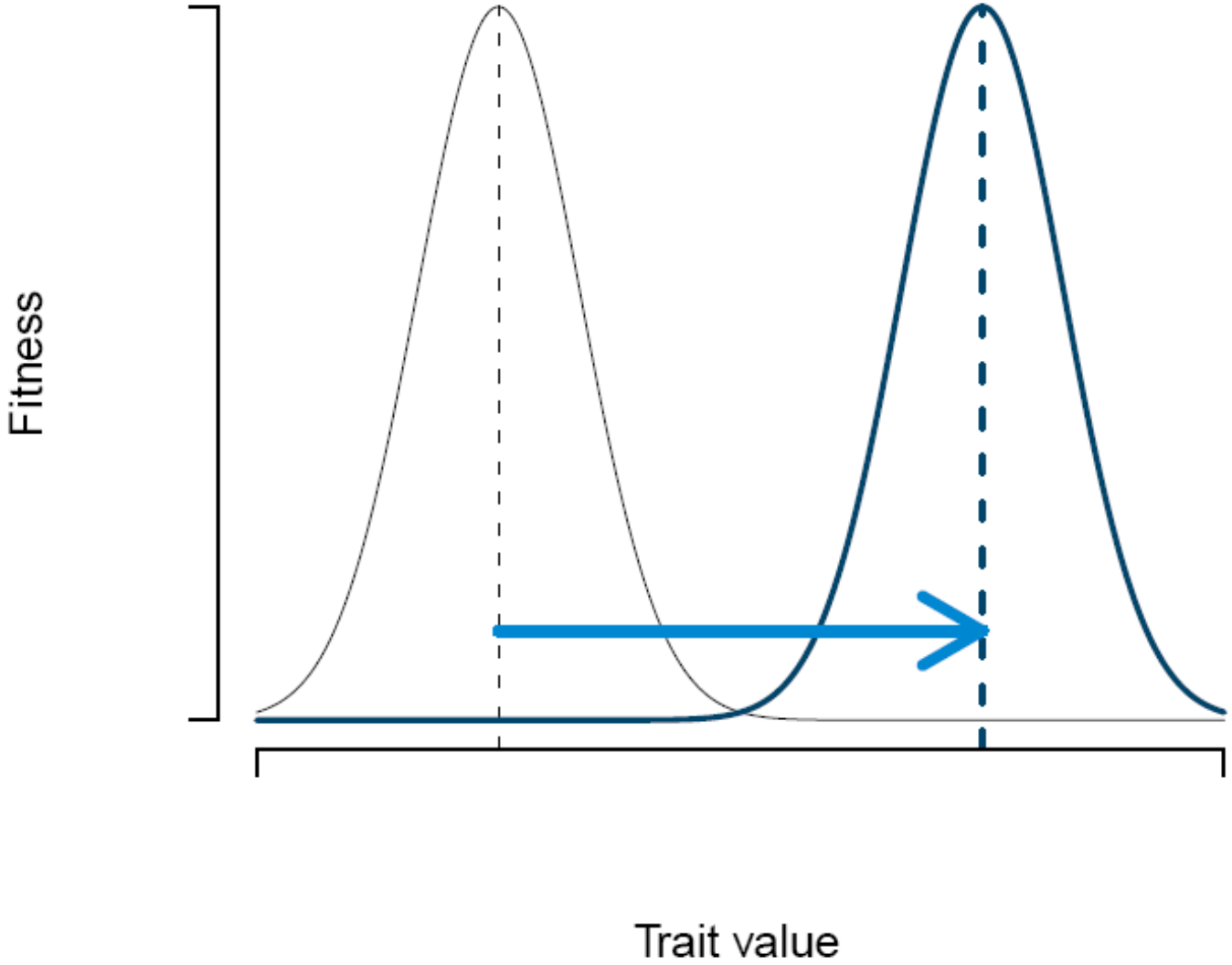
Effect on quantitative trait proportional to expression



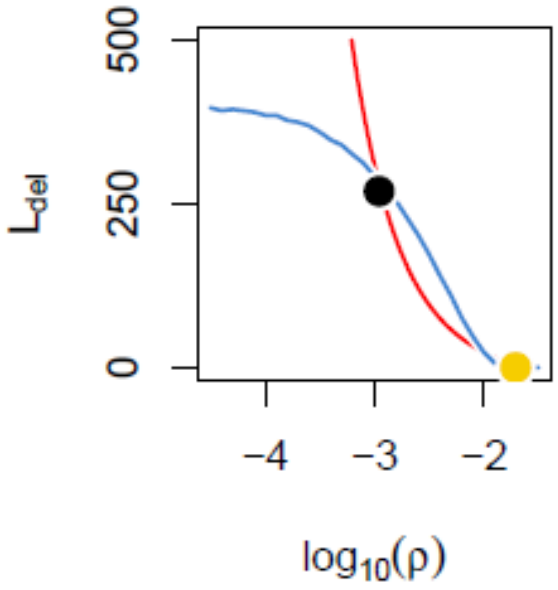
Point mutation in stop codon → full expression of previously cryptic sequence (that won't misfold if error rate was high)

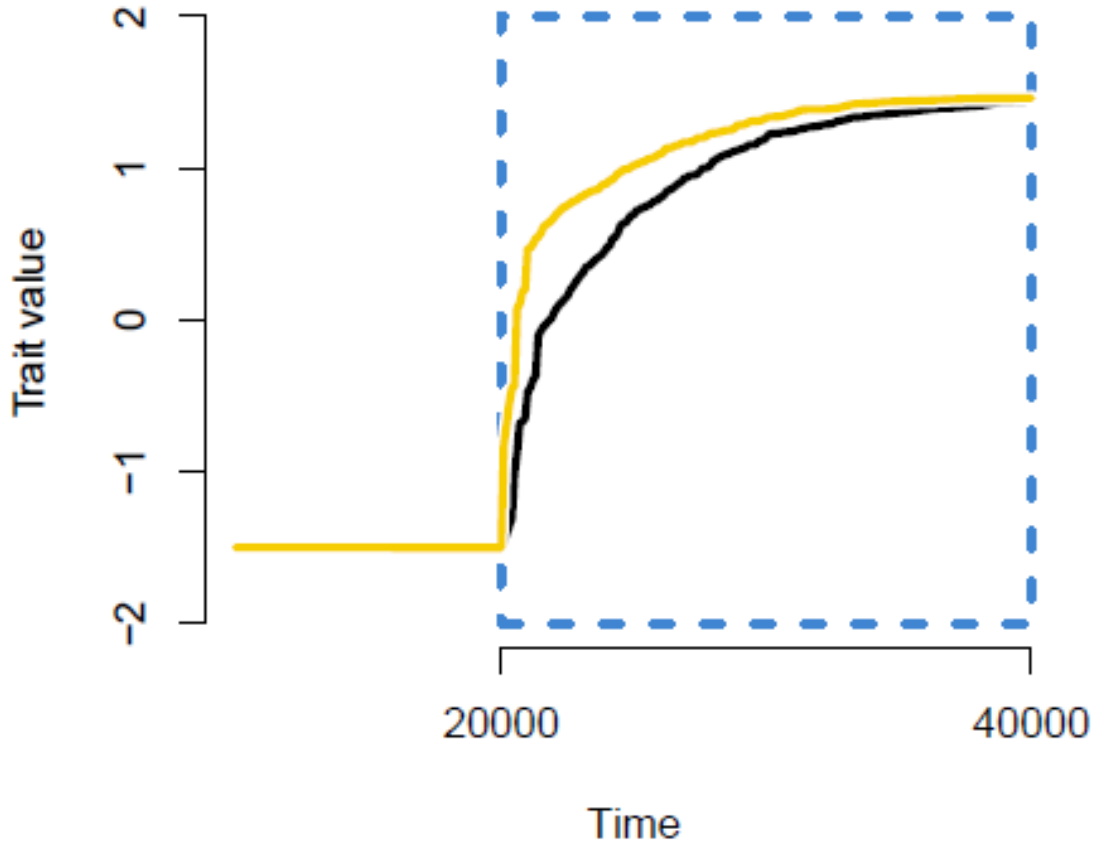


Environmental change in optimal trait value



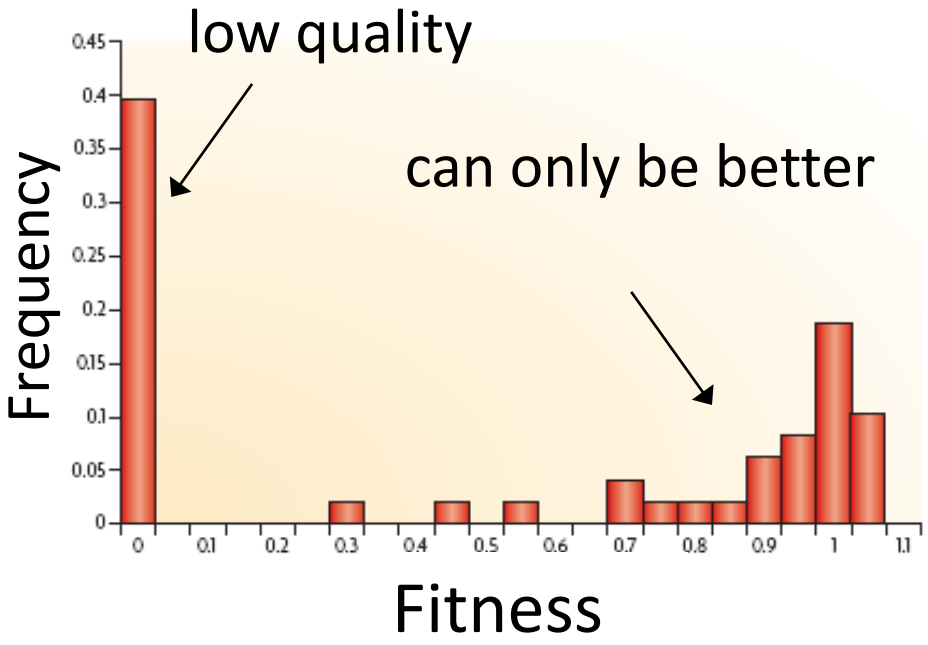
Populations with high error rates evolve faster



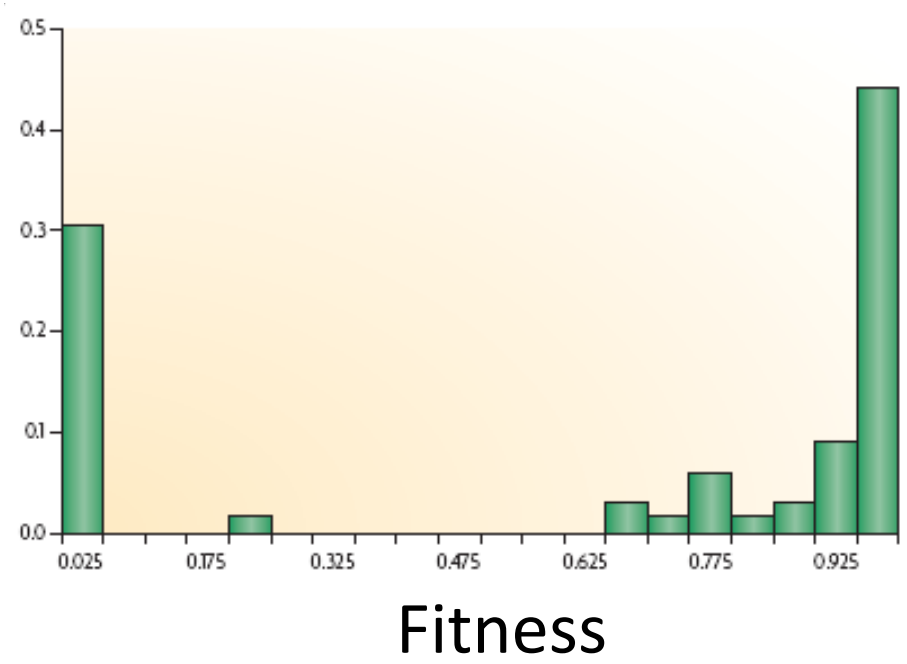


Rajon & Masel (2011)

New mutations



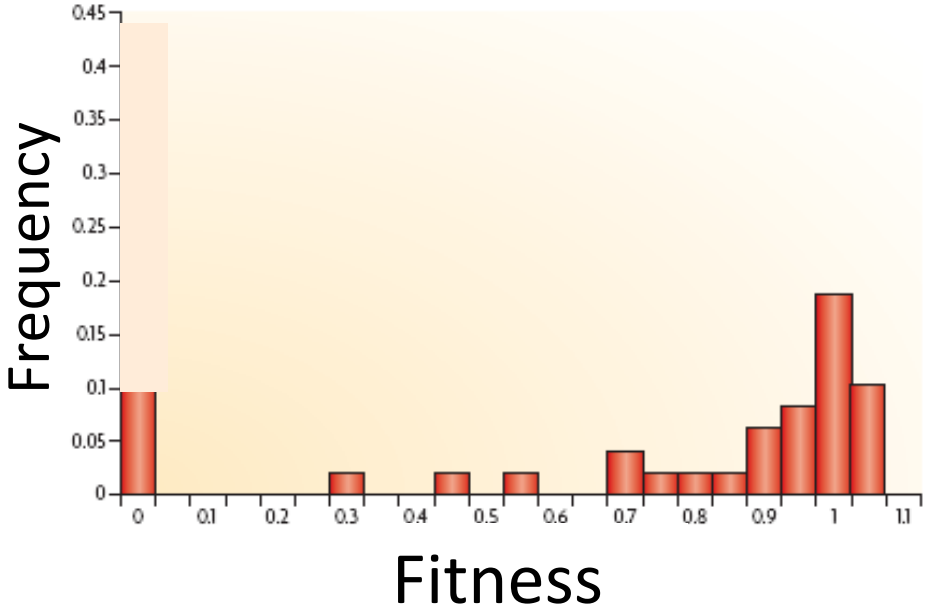
vesicular stomatic virus



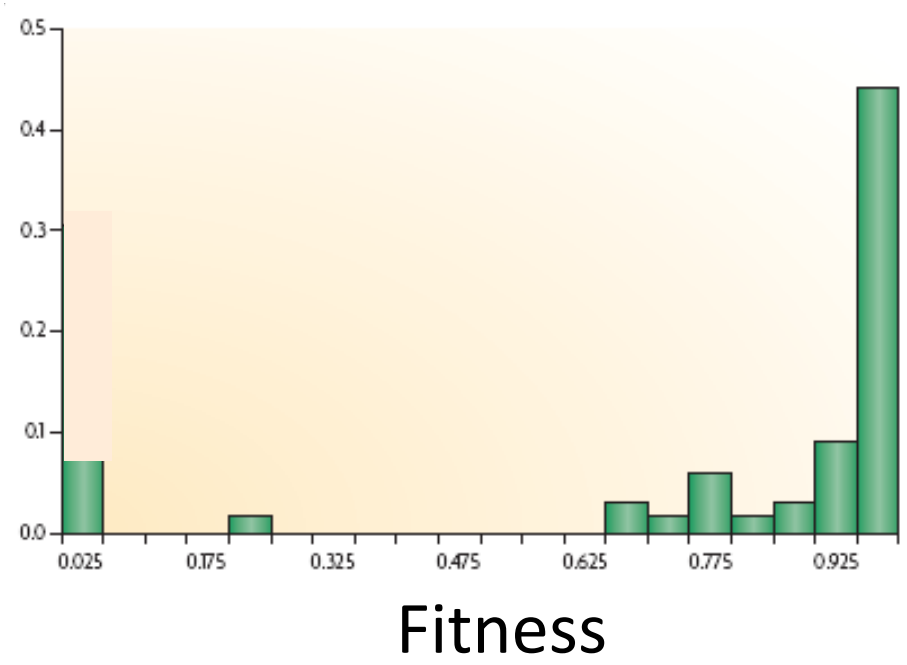
yeast

Eyre-Walker & Keightley 2007

Cryptic variants



vesicular stomatic virus



yeast

Pre-adapting selection

Masel 2006, Rajon & Masel 2011

Evolvability comes from tapping into cryptic variants

- Molecular errors in the present
  mimic mutations in the future
- Strongly deleterious sequences are
  pre-purged in favor of benign ones
- Benign sequences are co-optable for adaptation

Benefits go to any "high error" locally benign cryptic sequences

More examples

- Promiscuous enzyme activities
- Rare protein-protein interactions (PPIs) that lose crypticity when proteins see each other more often

Aside: "cryptic" PPIs (deliberately bad yeast-2-hybrid data) are biologically meaningful

They predict gene noise and plasticity better than "real" PPIs (best practice affinity capture mass spec)

"Stickiness" trumps "hubness"

Brettner & Masel (2012)

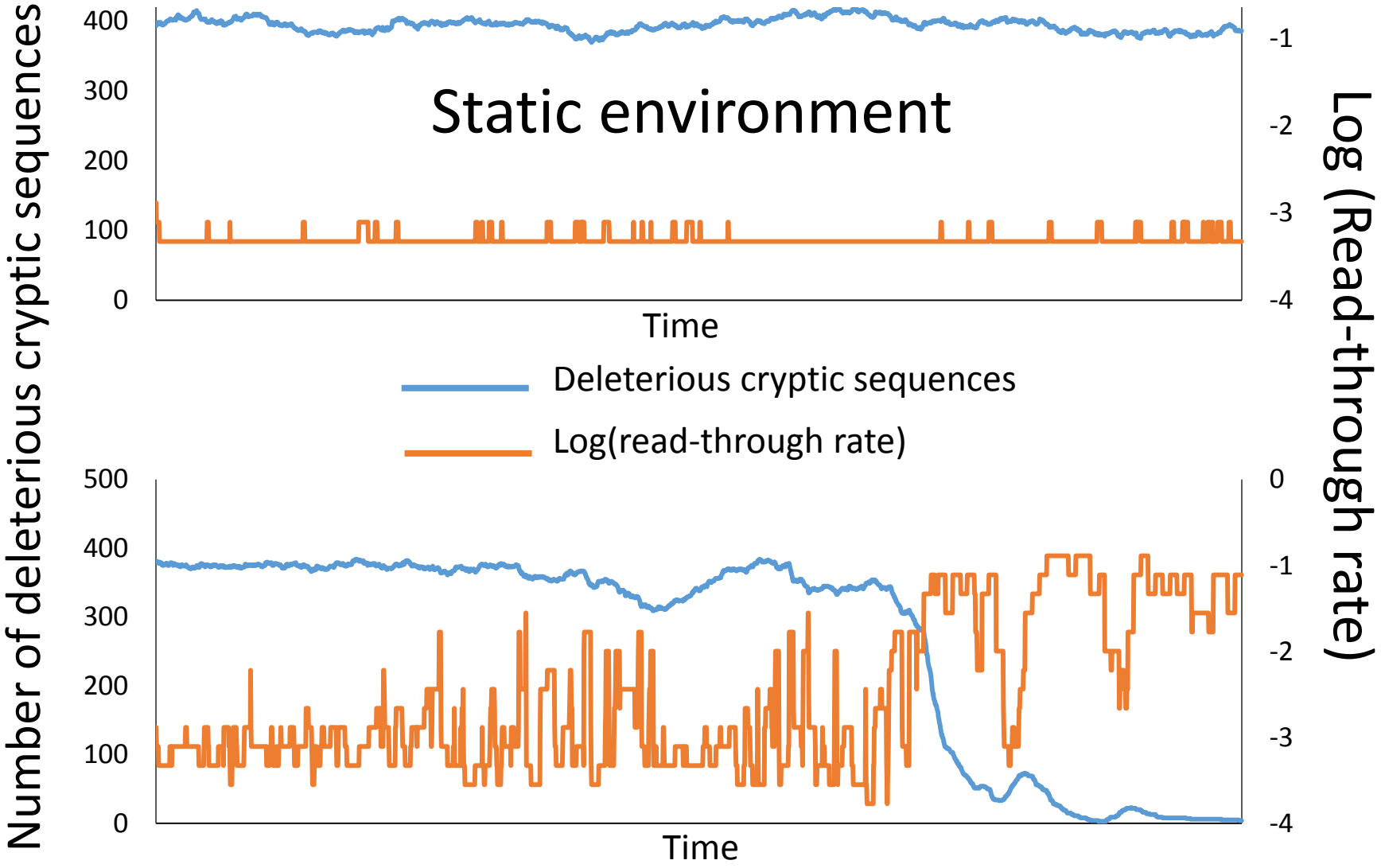
Outline

1. 1st line of defense is benign cryptic sequences (local)
   2nd line of defense is low error rate (global)
2. 1st line of defense promotes evolvability
3. Evolution of evolvability
4. Data: high errors are associated with high evolvability
5. De novo gene birth

Evolution of evolvability

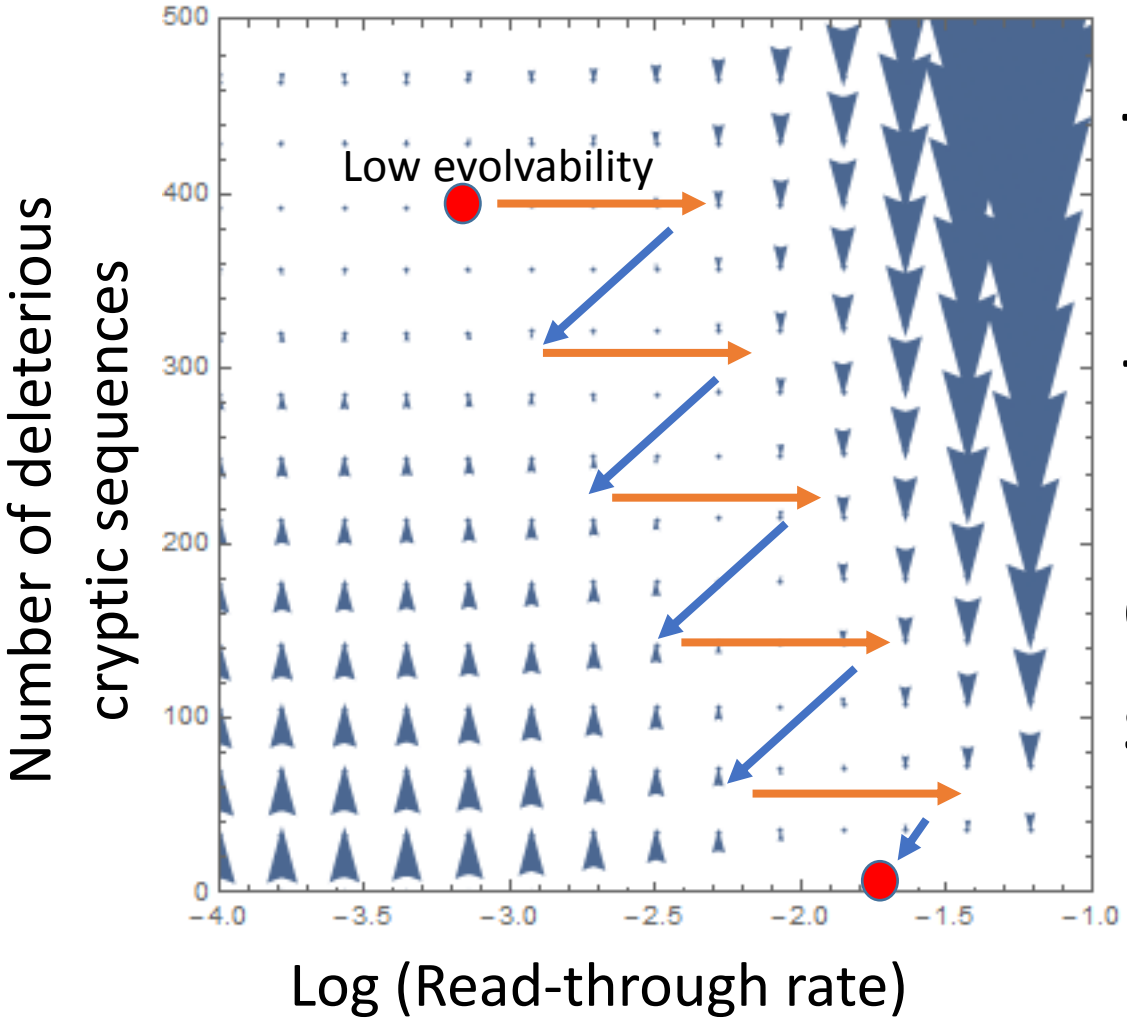
- Evolvability = byproduct of purging deleterious cryptic sequences at high N_e
- Adaptive "capacitors" switch on benign sequences during environmental change
  - E.g. yeast prion [PSI^+] is a heritable but reversible way to increase stop codon readthrough
  - Only works when sequences are benign
- Are sequences more likely to be benign when "needed" often?

Recurrent environmental change tips a bistable system towards the high evolvability attractor



Nelson & Masel in prep

Environmental change briefly favors high errors, acting as an evolutionary capacitor



Temporary pulses in the read-through rate result in a loss of deleterious cryptic sequences.

Nelson & Masel in prep

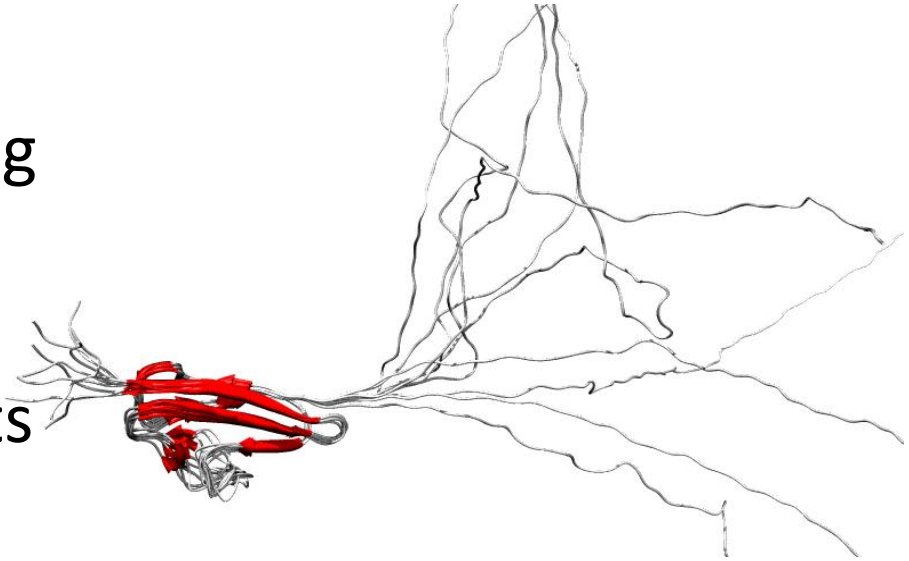
Outline

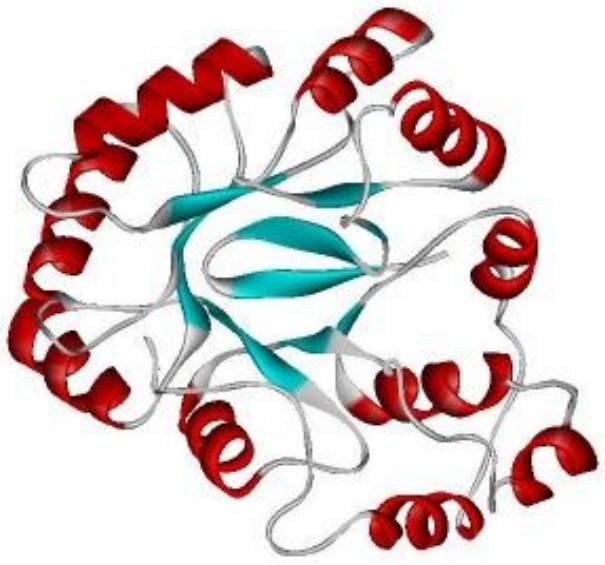
1. 1st line of defense is benign cryptic sequences (local)
   2nd line of defense is low error rate (global)
2. 1st line of defense promotes evolvability
3. Evolution of evolvability
4. Data: high errors are associated with high evolvability
5. De novo gene birth

Can we find a predictor of evolvability?

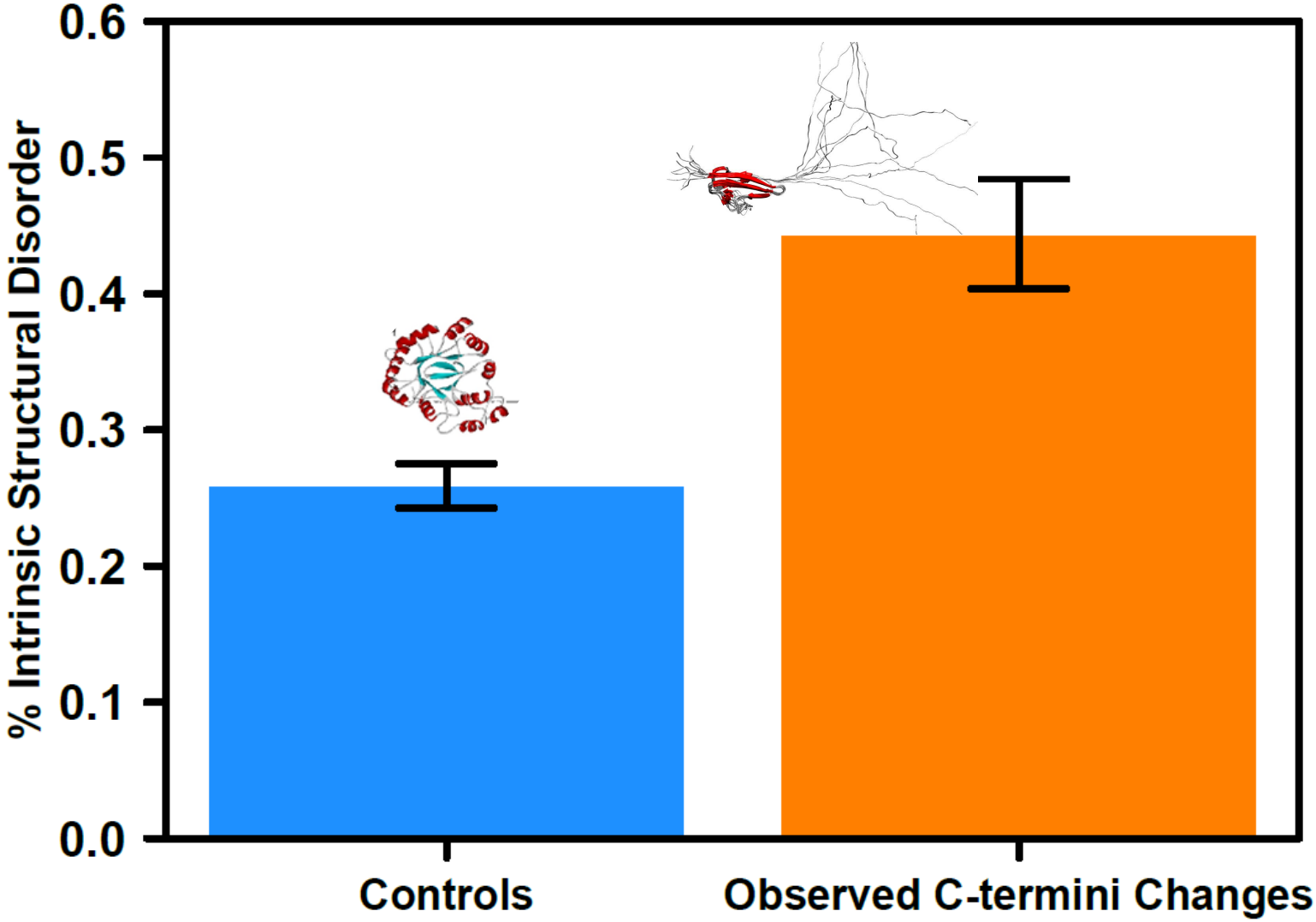
- Hypothesis: a hydrophilic, “floppy” addition from losing a stop codon should be tolerated better than a hydrophobic one that inserts into an existing protein fold
- Test by looking at 54 recent (polymorphic) yeast cases of stop codon loss → de novo C-termini. So recent that they are proxies for ancestor.

Andreatta et al. 2015



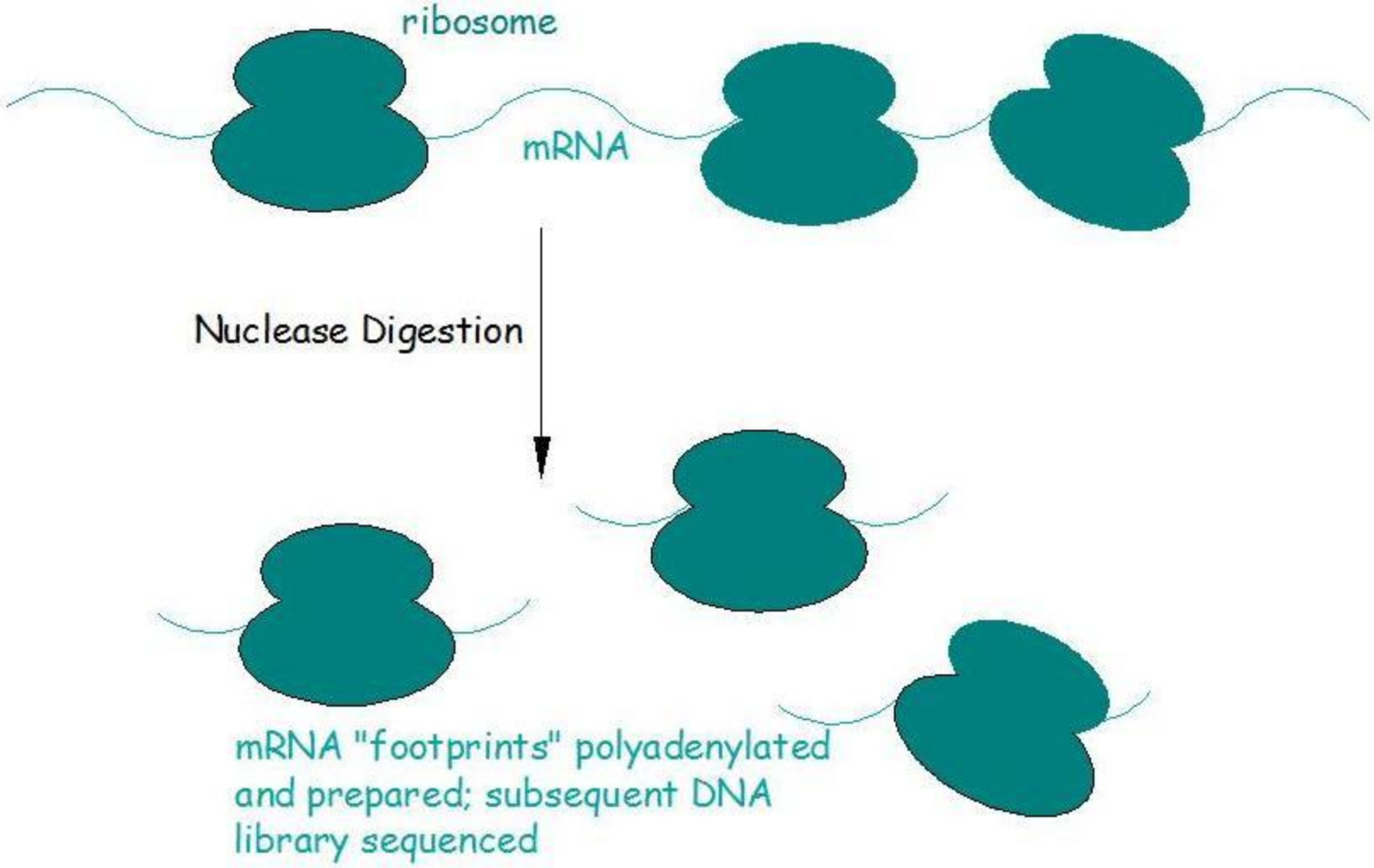


High intrinsic structural disorder is a preadaptation for joining a protein

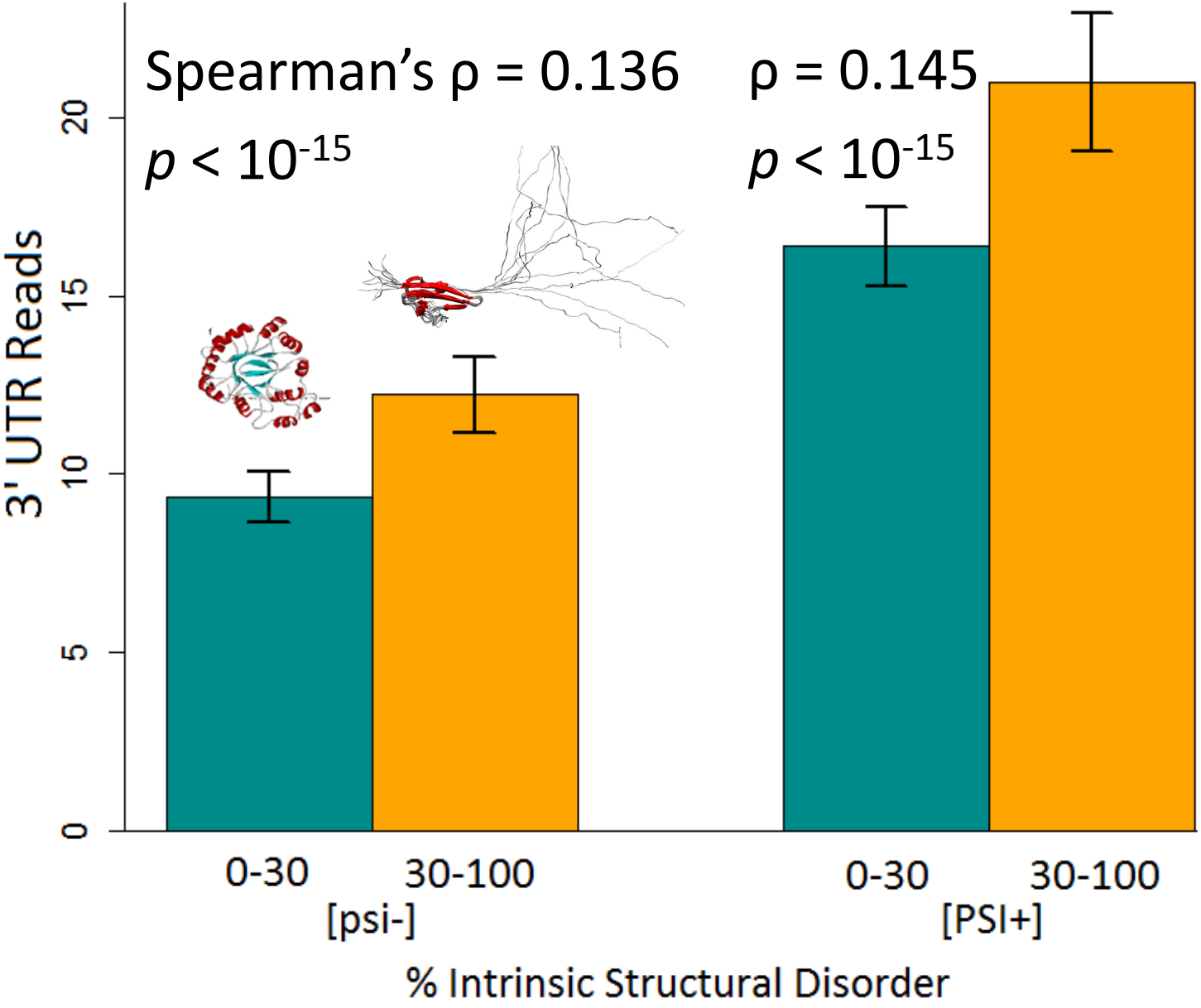


Kosinski et al. in prep

Now link preadaptation (ISD beyond stop codon) to error rate (ribosomal profiling hits past stop codon)



Across all yeast genes, the high ISD preadaptation (high evolvability) is associated with high error rates



Kosinski et al. in prep

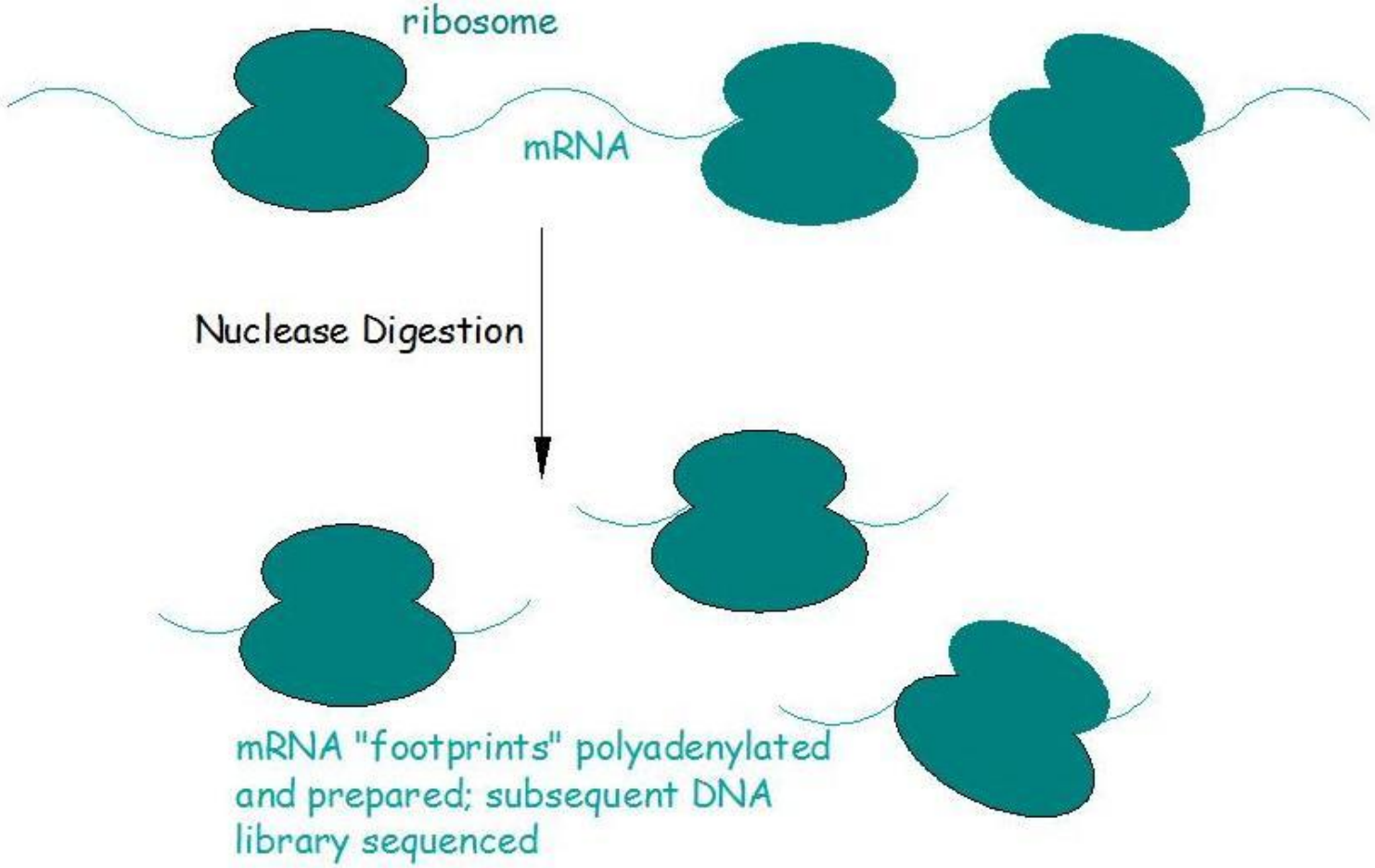
Outline

1. 1st line of defense is benign cryptic sequences (local)
   2nd line of defense is low error rate (global)
2. 1st line of defense promotes evolvability
3. Evolution of evolvability
4. Data: high errors are associated with high evolvability
5. De novo gene birth

De novo gene birth

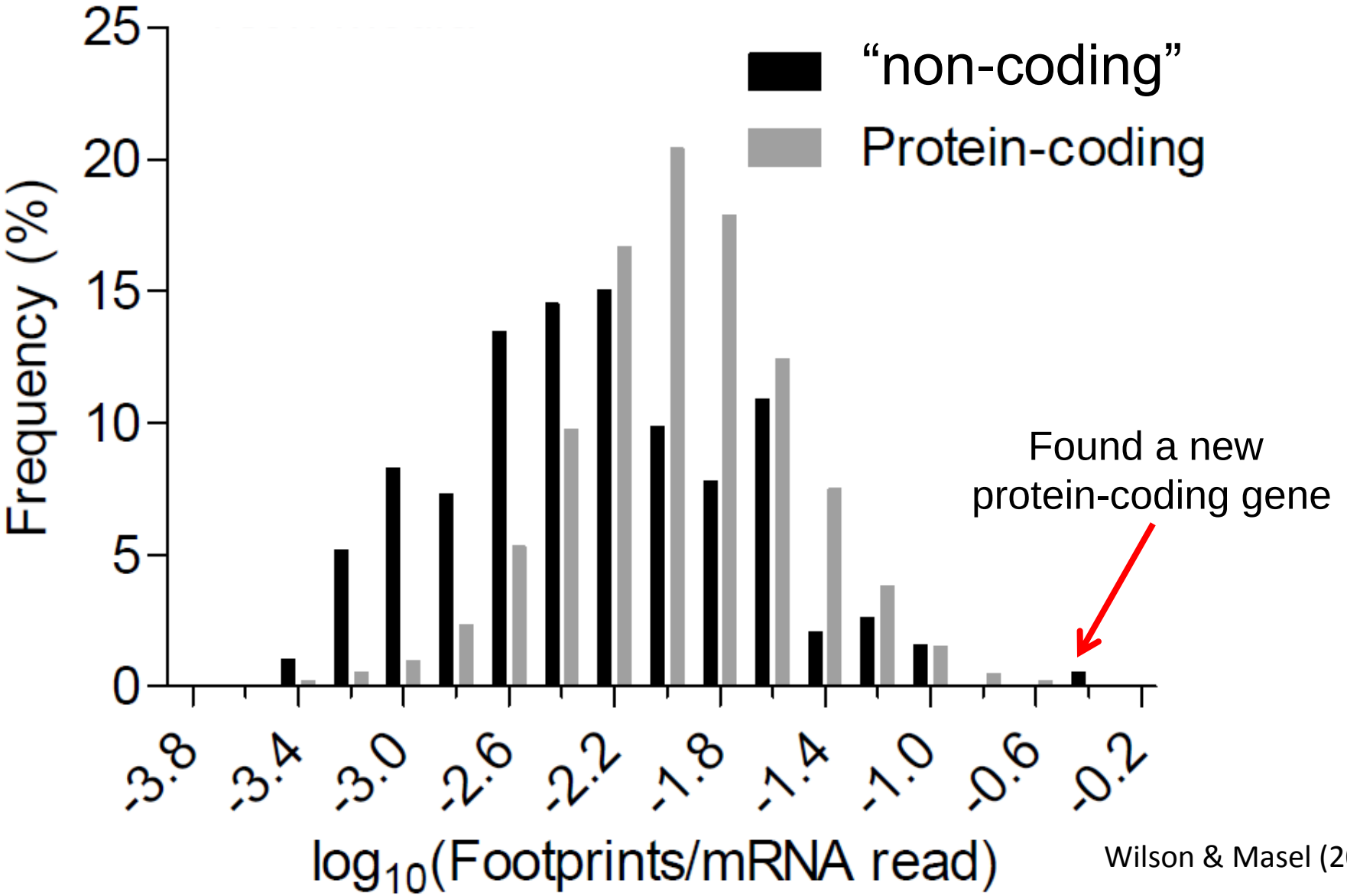
- Why aren't random polypeptides toxic?
- Explained if they are already under preadapting selection.
- Are there "proto-genes", i.e. non-coding transcripts that end up translated just a little bit, by accident, enough to purge out the deleterious options?

217/404 “non-coding” transcripts showed ribosomal association



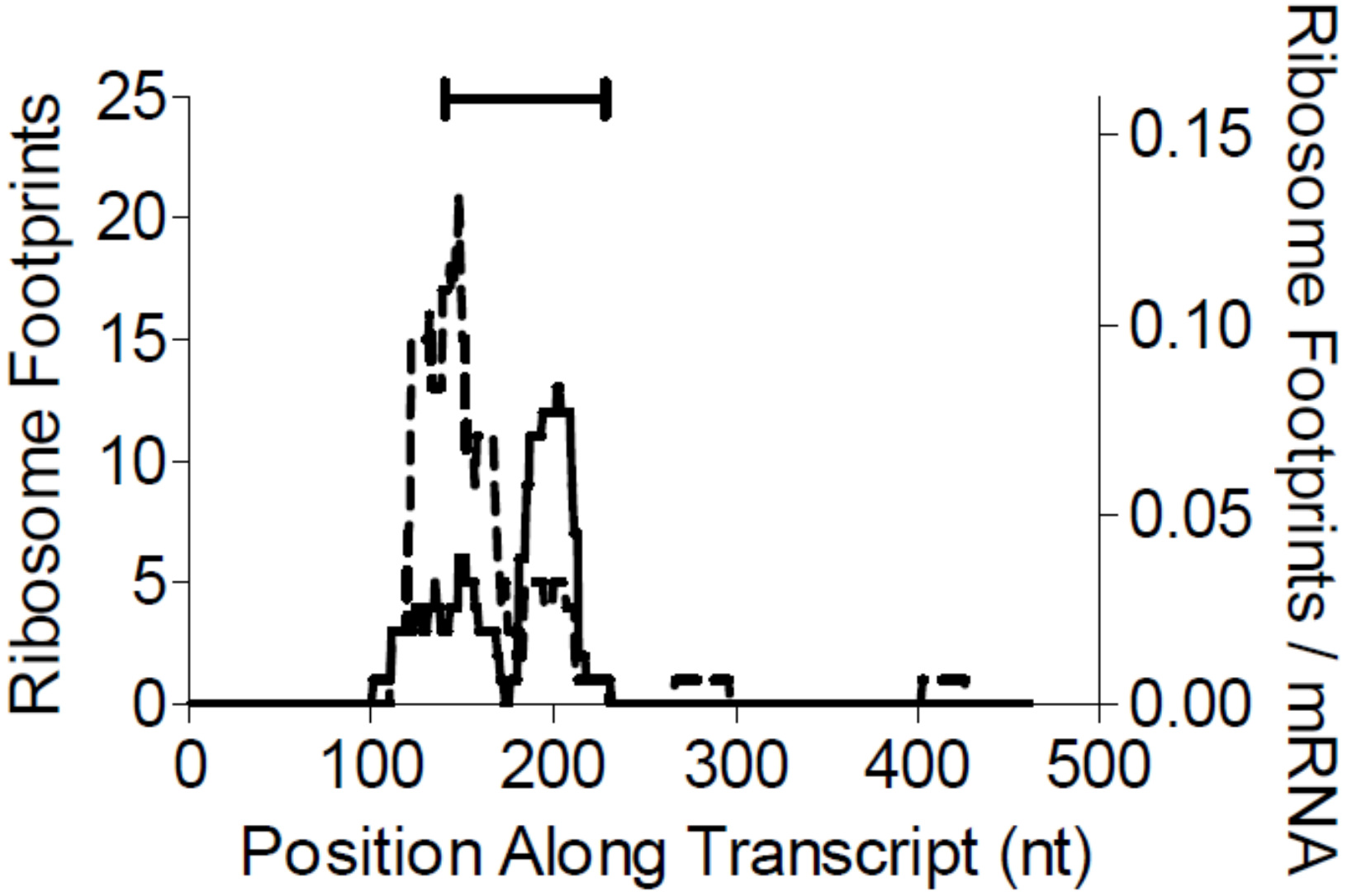
Wilson & Masel (2011)

Many individual “non-coding” transcripts have ORF-like ribosome densities



Wilson & Masel (2011)

Ribosomal footprint locations match a 28aa ORF



Wilson & Masel (2011)

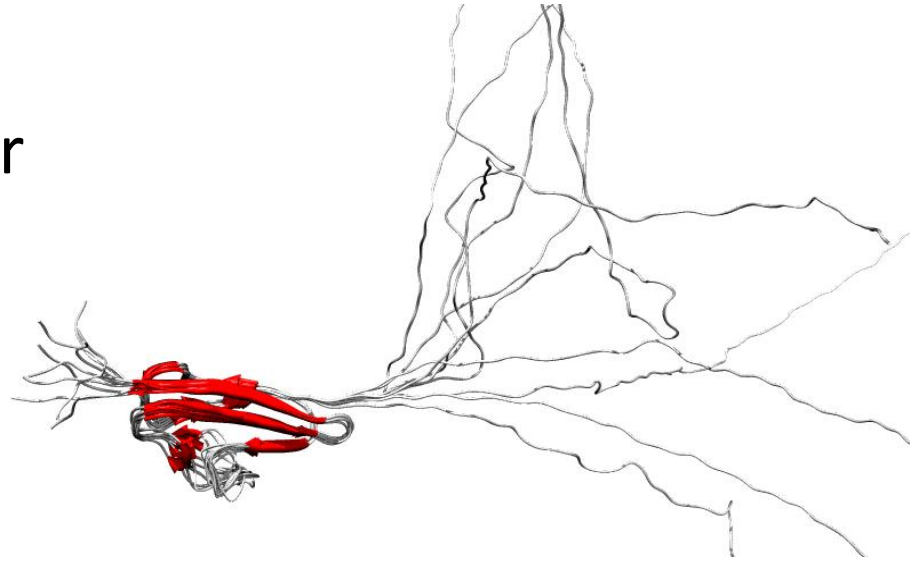
Summary of ribosome profiling results

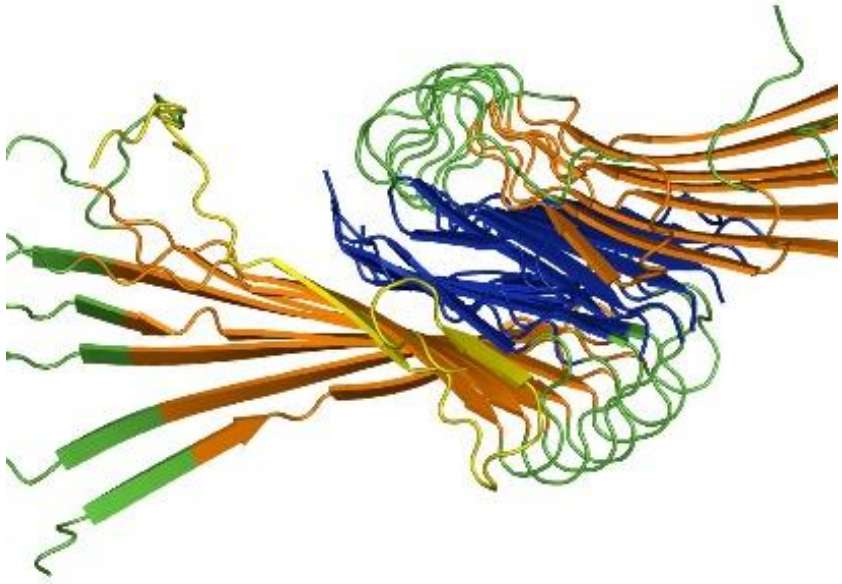
- Looks like a new coding sequence, but we don't know if polypeptide is functional
- Looks like de novo evolution
- Proof of principle of powerful method to annotate short de novo proteins
- Penultimate stage of gene birth is widespread

Wilson & Masel (2011)

Most eukaryotic ORFans may have arisen de novo: what is special about them while young?

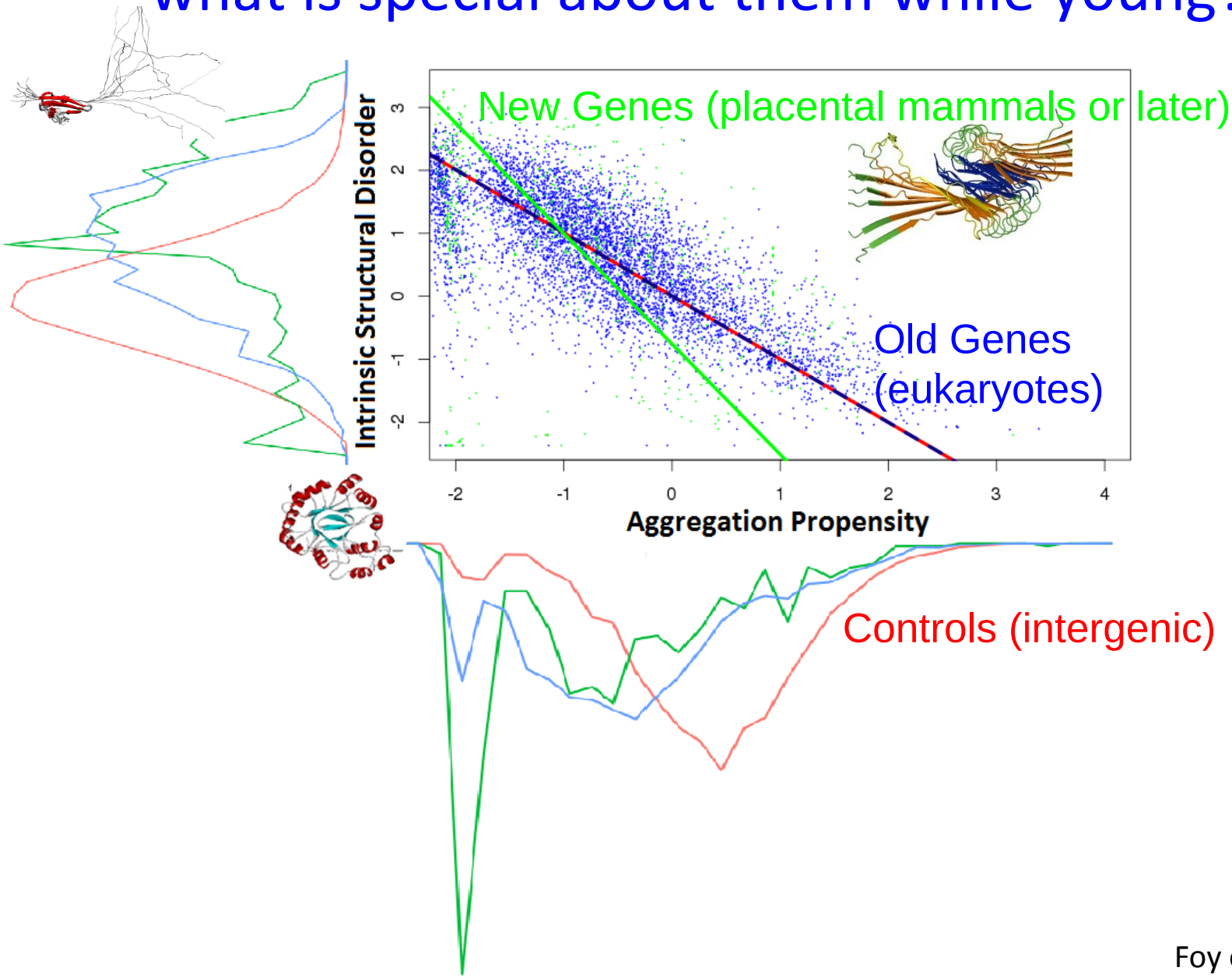
- Previous reports of high intrinsic structural disorder
- We hypothesize a need to avoid protein aggregation, although evidence on this has been scant





- The two are confounded: hydrophobic proteins have low disorder and high aggregation propensity

Most eukaryotic ORFans may have arisen de novo: what is special about them while young?



Foy et al. (in prep)

Conclusions

- 2 solutions to many molecular errors
  - high error rate, but robustness to each separate error
    (local solution, 1st line of defense)
  - low error rate via a proofreading mechanism for all sites
    (global solution, 2nd line of defense)
- High error rates pre-screen future variants,
  and so promote evolvability
- Biochemical correlates in the role of intrinsic disorder and aggregation propensity during de novo gene birth

Broader picture

- Waste and mess and errors are not just a typical biological nuisance
- Without waste and mess, creative evolutionary innovations may not be possible
- Looking for a clean molecular machine can miss the essence of biology

Thanks!

Funding

NIH
Pew Charitable Trusts
John Templeton Foundation

Theory

Etienne Rajon
Paul Nelson

C-terminal extensions

Luke Kosinski
Matt Andreatta
Josh Levine
Eden Eaton
Lynette Guzman
Matt Cordes
Premal Shah (U Penn)
Dan Jarosz (Stanford)

De novo genes

Ben Wilson
Scott Foy
Rafik Neme

PPIs

Leandra Brettner