Selection and pleiotropy constrain the human nonsynonymous mutation spectrum

David Castellano & Ryan Gutenkunst
University of Arizona

"Mutation is the ultimate source of genetic variation."

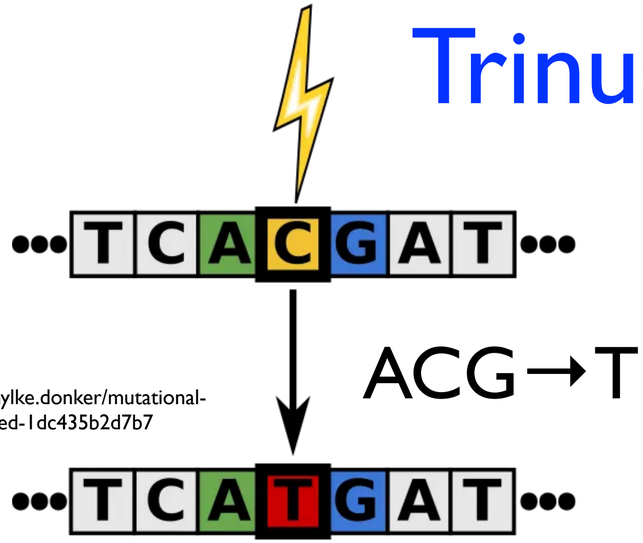
(too many papers and textbooks to count)

But see Sturtevant (1937) *Quart Rev Biol*

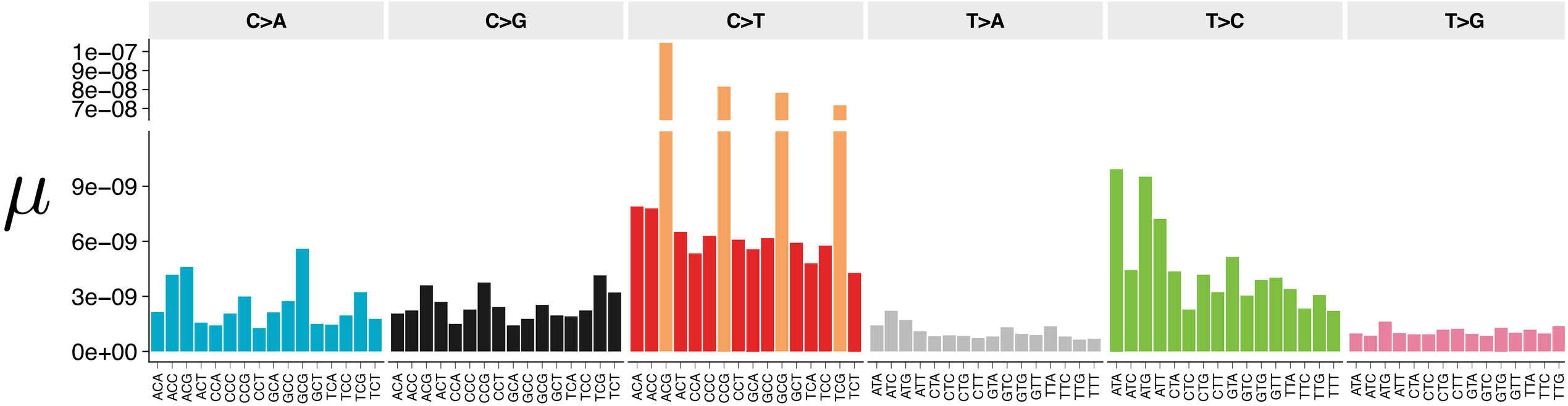
NIH NIGMS

ARIZONA

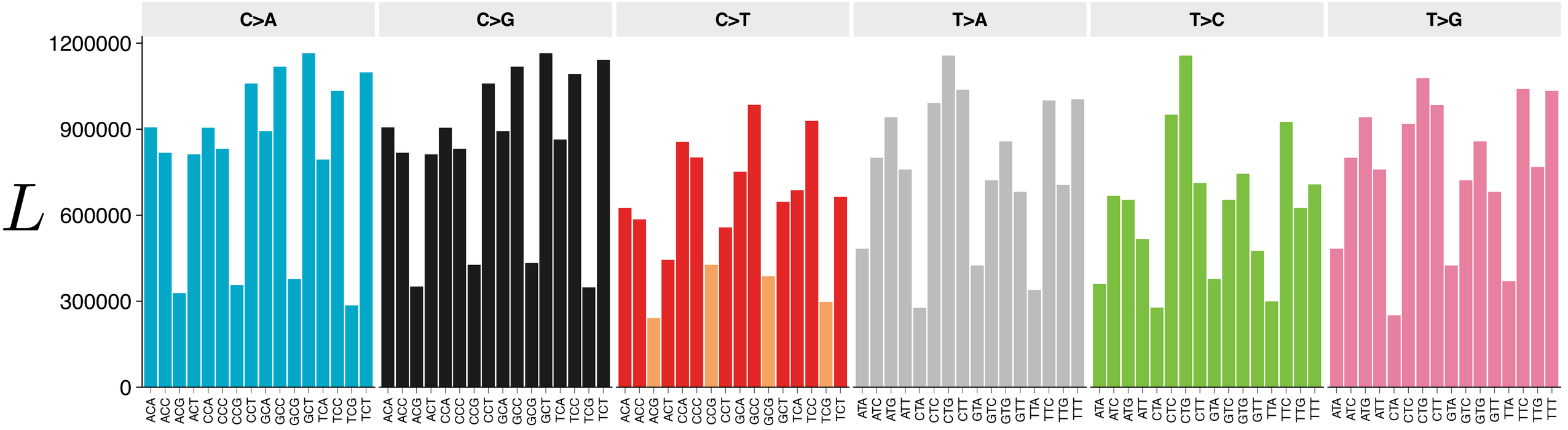
Trinucleotide mutation rates



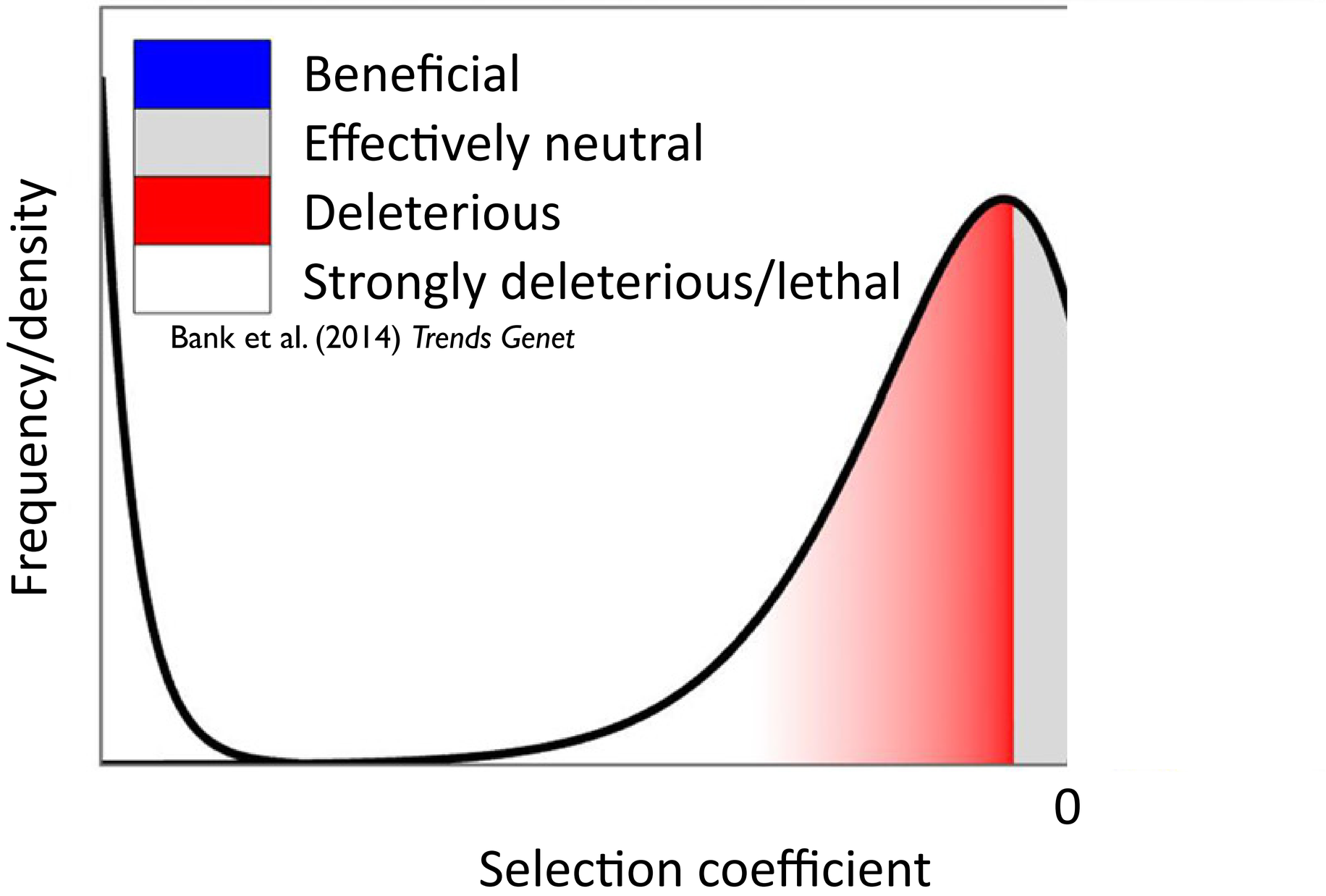
https://medium.com/@hylke.donker/mutational-signatures-explained-1dc435b2d7b7



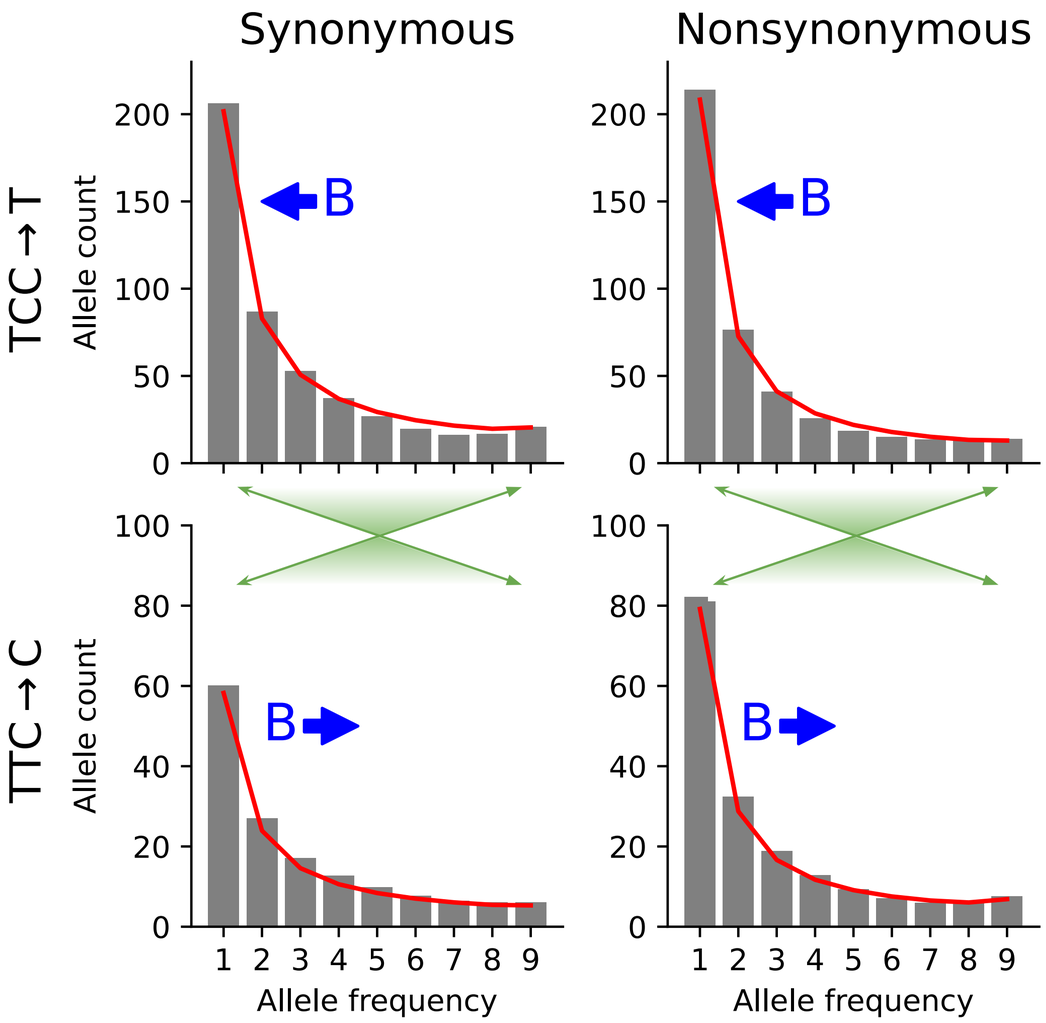
Nonsynonymous opportunities



Distributions of Fitness Effects (DFEs)



Our inference framework



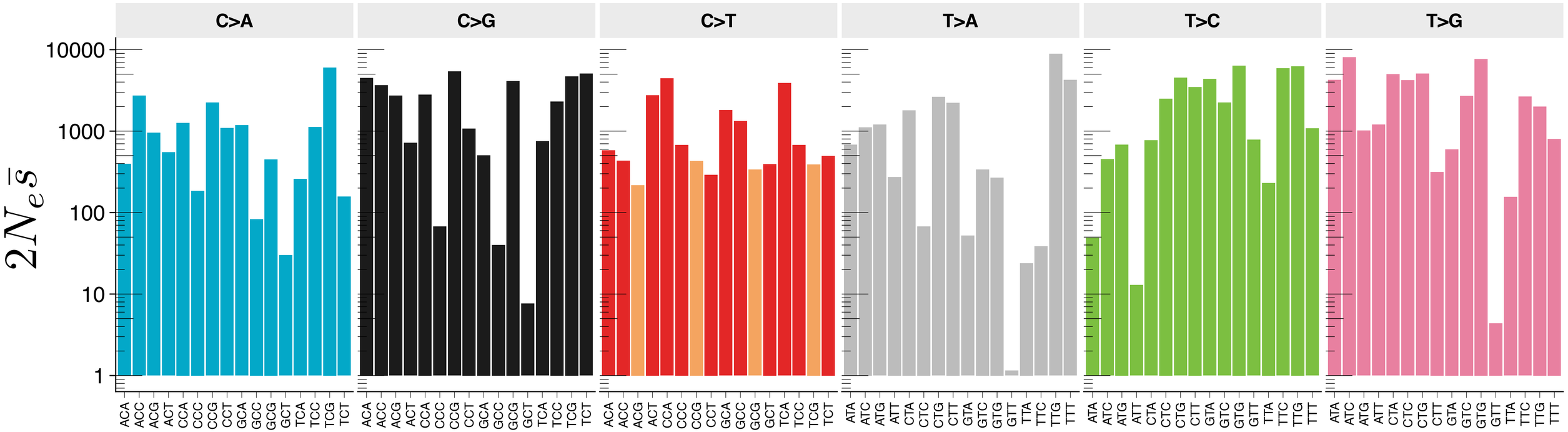
To account for ancestral state misID,
fit paired forward and backward SFS

Glémin et al. (2015) *Genome Res*

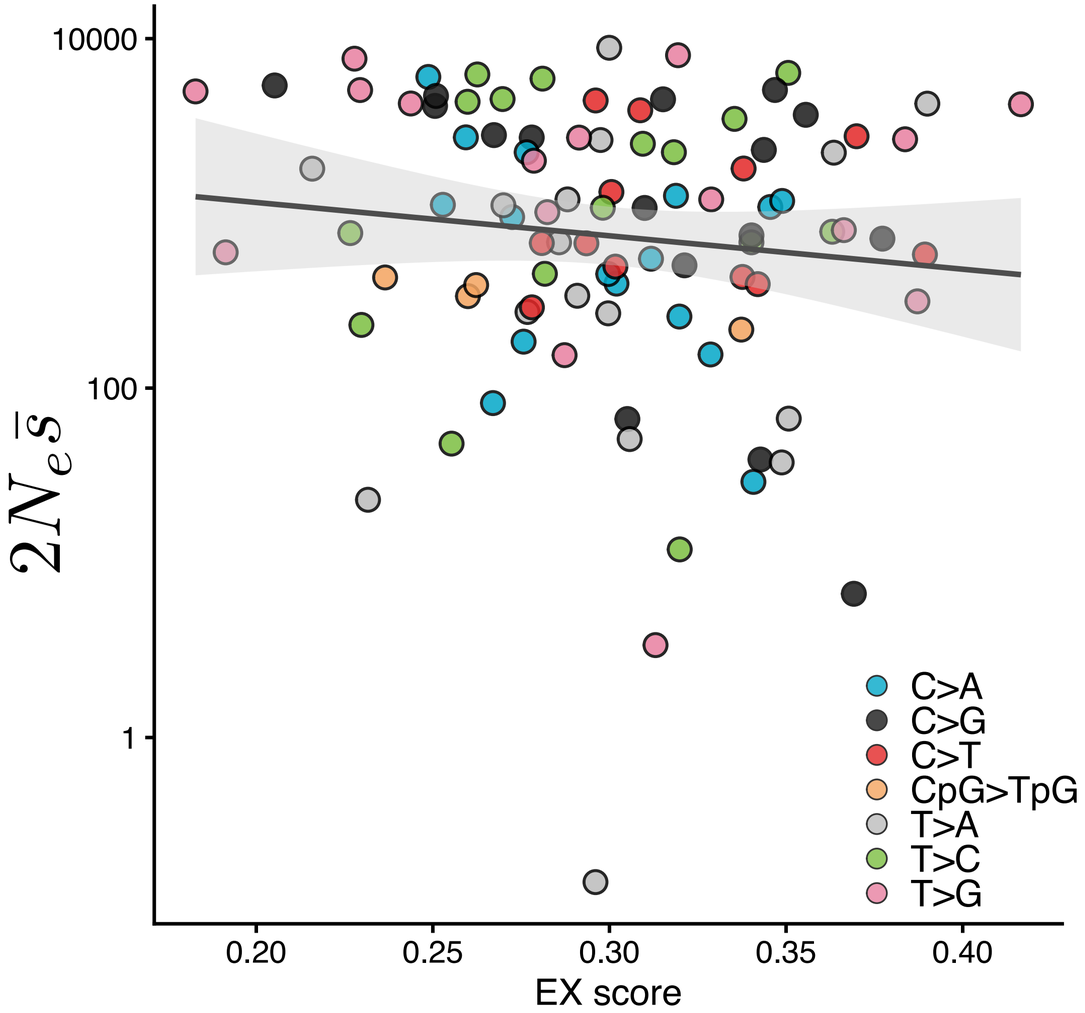
1. Using gcBGC-neutral types, fit demographic history to synonymous sites
2. For other types, fit gcBGC to synonymous sites (including demography)
3. For all types, fit gamma DFEs to nonsynonymous sites (including demography and gcBGC)

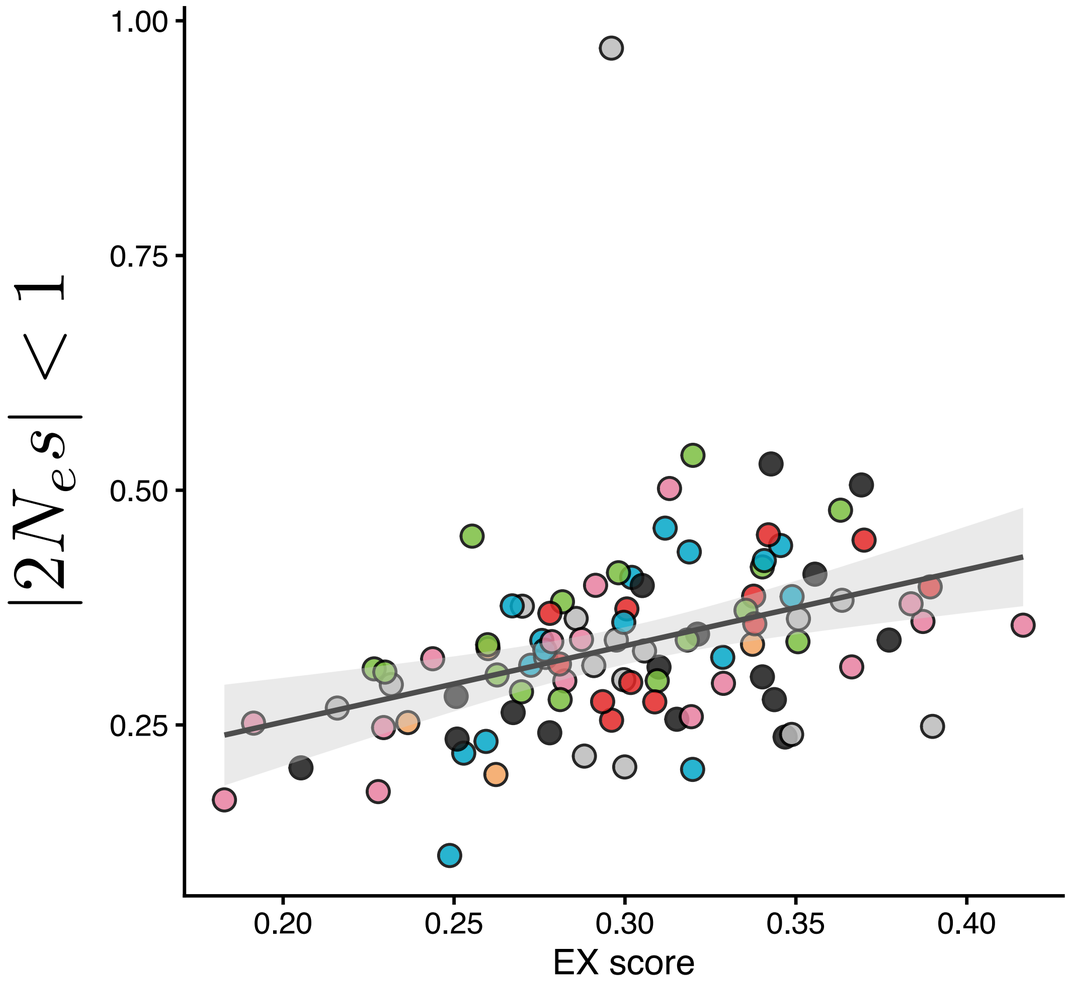
Analyze Yoruba data from 1000 Genomes project, with dadi

Inferred DFEs



Correlation with AA exchangeability

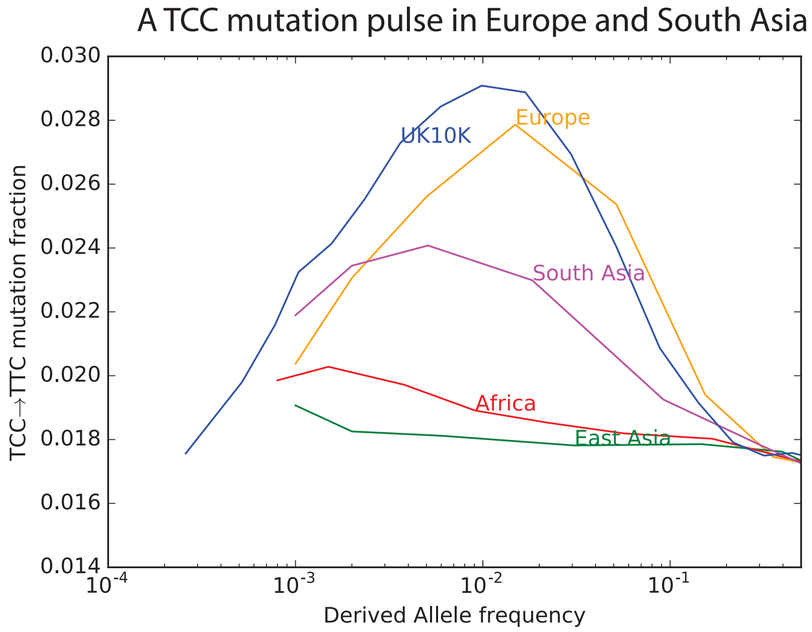


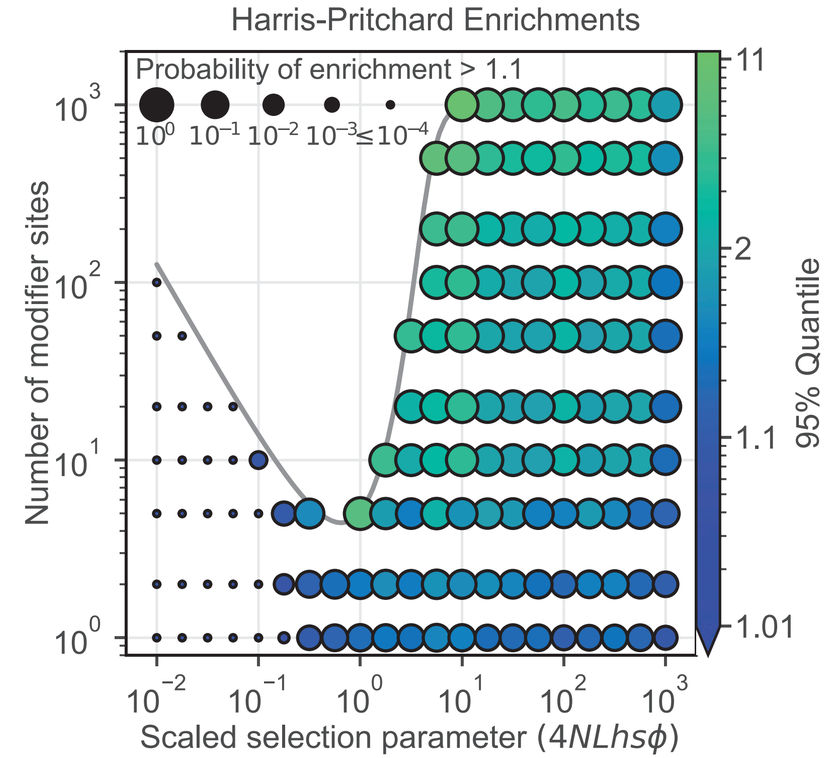


Yampolsky & Stoltzfus (2005) *Genetics*

Transient mutators

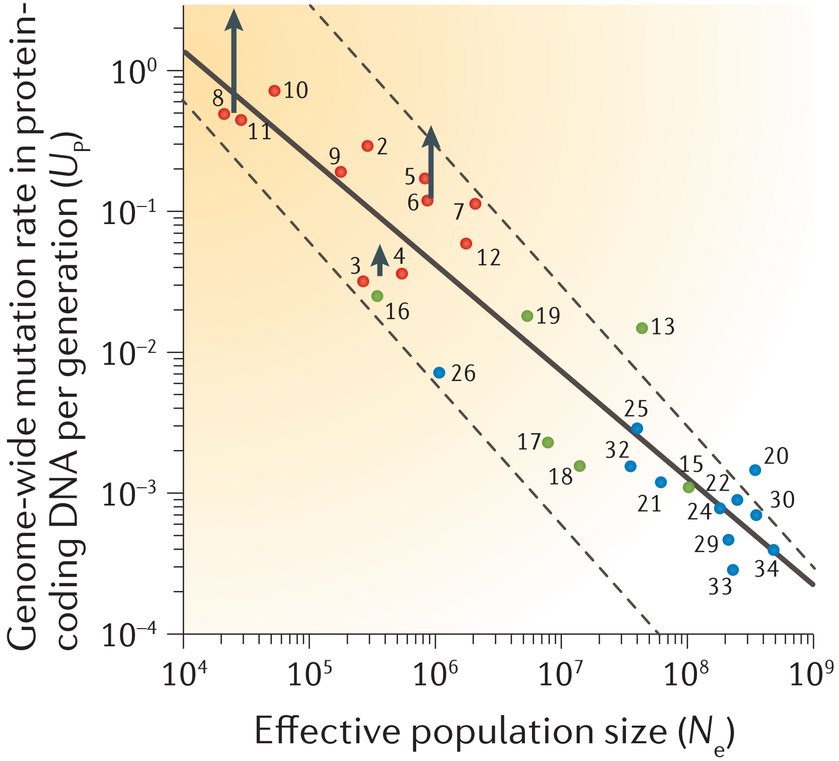
- Harris and Pritchard (2015, 2017) discovered a population-specific +50% pulse of TCC→T mutations
- The indirect fitness effect of a mutator that raises rate by a factor ϕ is $s_\phi = 2\phi\mu L\bar{s}$
- Milligan (2022) theoretically identified parameter regime for observable pulses
- For the TCC→T pulse, mutator at 10-25% allele frequency would have $\phi \sim 2-5$
- For TCC→T, our results thus imply $2N_e s_\phi = 4N_e L\phi\bar{s} \sim 29-72$, which is outside Milligan (2022)'s regime





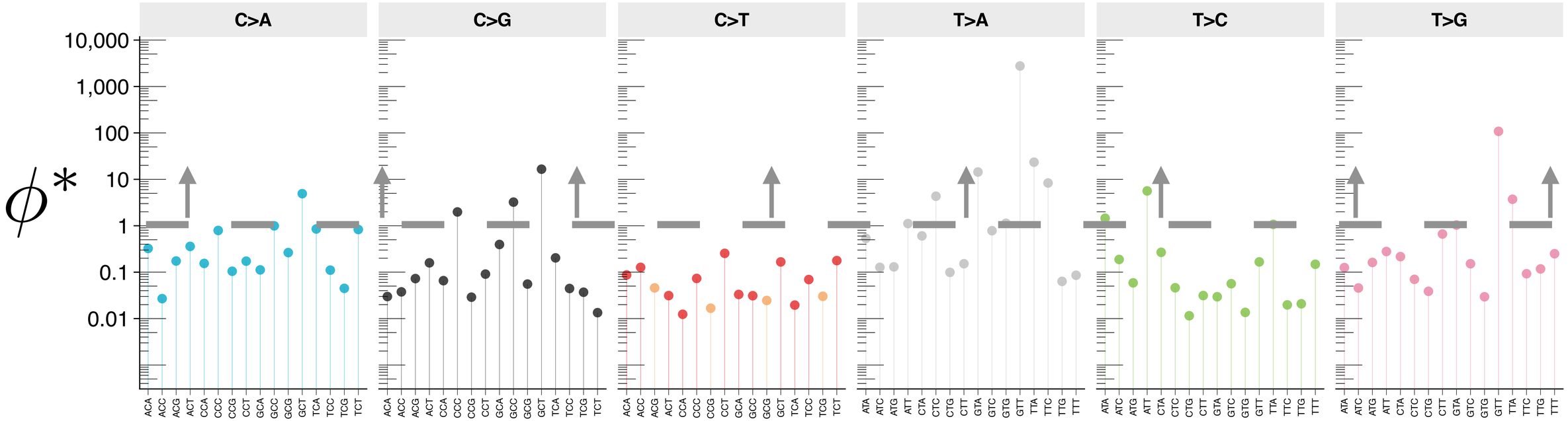
Mutator mutation pressure

- Mutator alleles are much more common than antimutator alleles.
  So under neutrality, mutation pressure should push mutation rates upward.
- Rates are constrained by selection against indirect mutation effects: $s_\phi = 2\phi\mu L\bar{s}$
- Long-term, an allele will be effectively deleterious if $2N_e s_\phi > 1$
- We define ϕ^* as the threshold mutator effect size above which alleles will be effectively deleterious



Lynch (2016) *Nat Rev Genet*

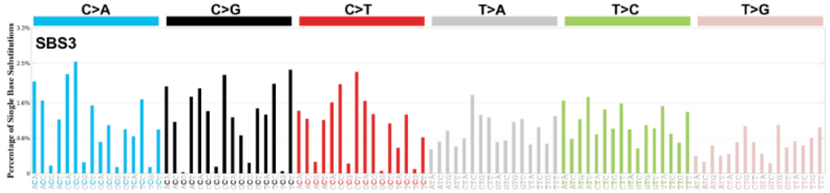
Mutator mutation pressure



Rates of many mutation types could move upward dramatically

Genetic architecture of mutation

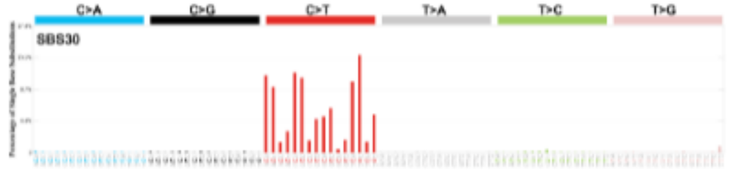
- Mutator alleles may necessarily be pleiotropic, affecting more than one trinucleotide mutation type
- Mutation signatures from human tumors are cataloged by COSMIC
- We use 15 signatures with proposed biological etiologies



SBS3

Proposed Aetiology

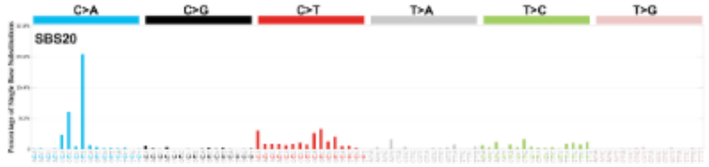
Defective homologous recombination DNA damage repair



SBS30

Proposed Aetiology

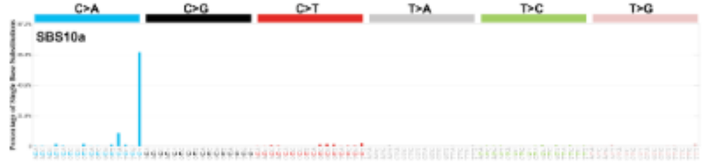
Defective DNA base excision repair due to NTHL1 mutations



SBS20

Proposed Aetiology

Concurrent POLD1 mutations and defective DNA mismatch repair



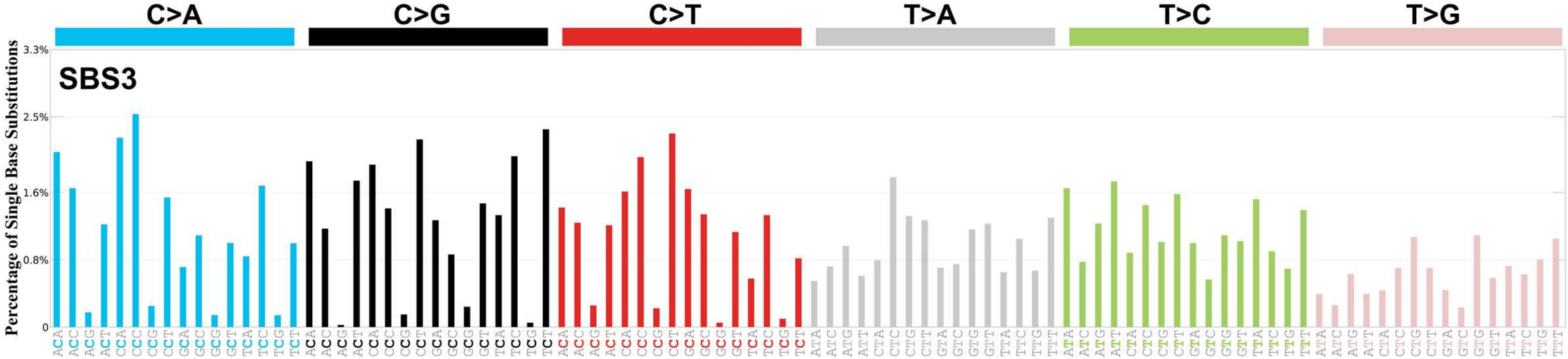
SBS10a

Proposed Aetiology

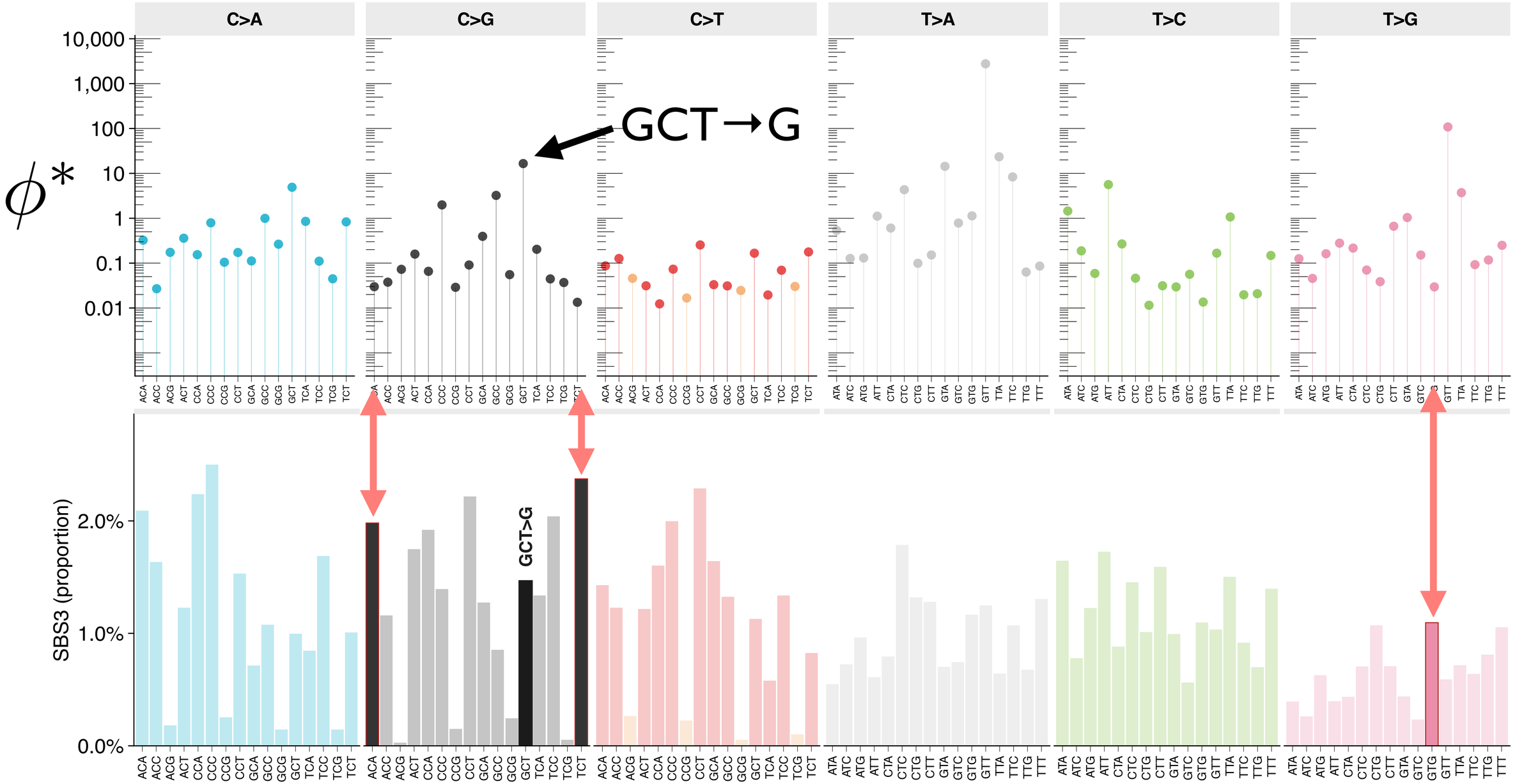
Polymerase epsilon exonuclease domain mutations

Mutation signature selection

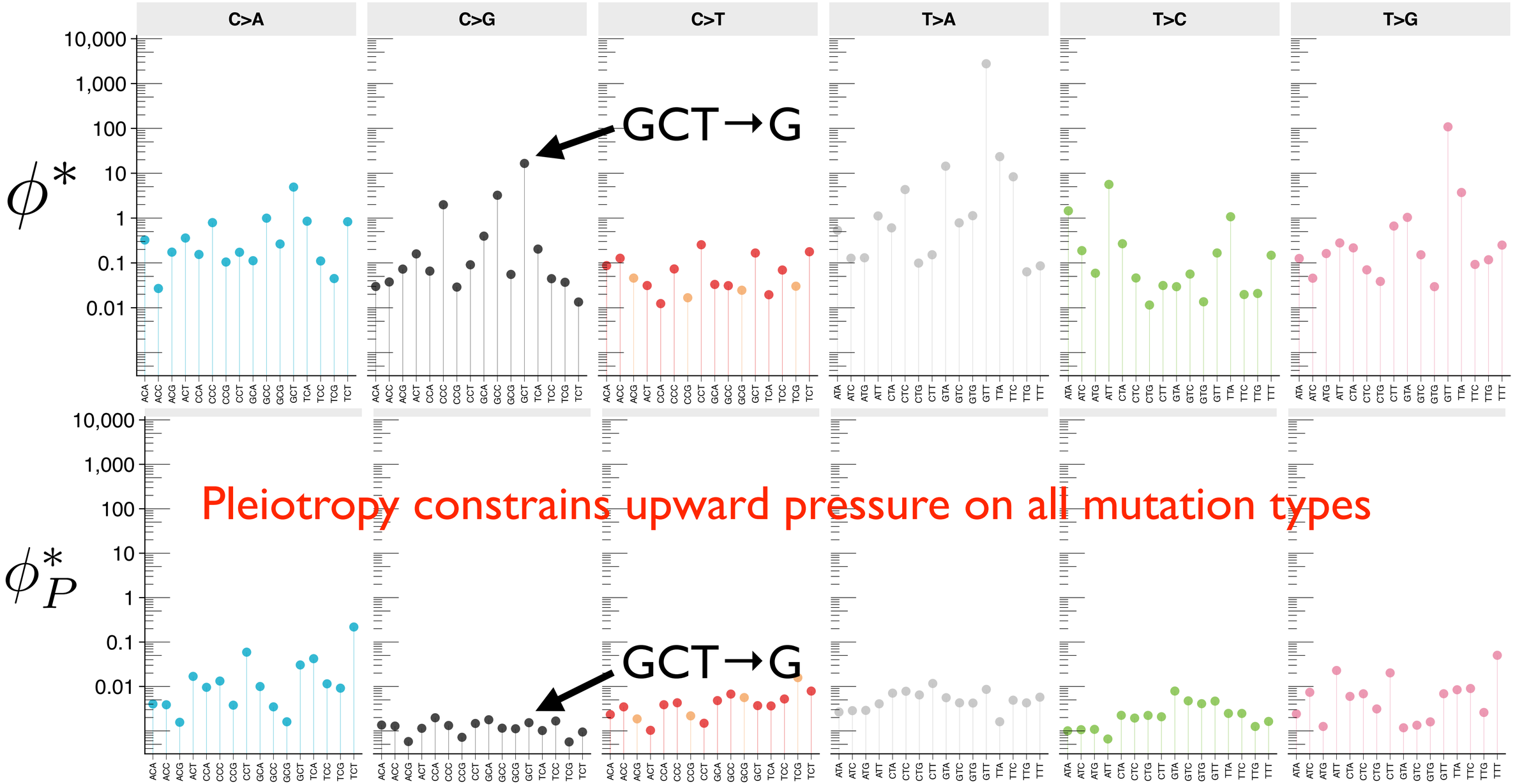
- The indirect fitness effect of a mutator allele that increases the activity of a given signature by an amount A is $s_A = 2A \sum_{\text{types } t} w_t L_t \bar{s}_t$
- Each signature thus has a maximum activation level A^* above which a mutator would be effectively deleterious
- We define ϕ_P^* as the pleiotropic threshold effect size for each type, which is its largest relative change among max-activated signatures



Mutation signature selection



Mutation signature selection



Pleiotropy constrains upward pressure on all mutation types

Summary

- DFEs of nonsynonymous trinucleotide mutation types vary dramatically and are correlated with amino acid effects
- TCC→T pulse is unlikely to have been a mutator allele
- Mutation pressure could drive some trinucleotide mutation rates dramatically upward
- Pleiotropy among known mutator alleles constrains that pressure