

$\pi$  per site is greater when there are more bases differing among individuals i.e. if you have a very rare variant, it's not actually going to raise  $\pi$  all that much within a population

Loss of variation acting on advantageous mutation. - would lower  $\pi$ , especially in a region of high recombination

nearby alleles go to fixation with the advantageous mutation

Muller's ratchet: steady deterioration of fitness in a population.

Spread of other nearby alleles along with the advantageous one because of linkage (lack of recombination)

Spread of an advantageous allele throughout population by natural selection & associated loss of variation near it

Average out by the number of sites being investigated.

Average number of pairwise mismatches (2pg)

measure of nucleotide variation/diversity

if genes get more mutations, than likely to accumulate more differences between species

a lot of ways to change a.a. to nonsynonymously

$\pi$

selective sweeps

adaptive alleles can sweep

hitch-hiking

if you have high recomb, you're able to sustain variation, so we expect this to be a positive association.

recomb. rates not constant across genome.

recombination

get some offspring without bad mutations

allows population to "unload" itself from bad mutations (stopping the "ratchet")

may be particularly helpful in variable environments

recombination makes combinations of alleles across two or more loci that may be advantageous.

assuming no recombination, adaptive mutation will spread & erase all variation.

Size of exposed window varies with recombination rate.

"bad" mutations are very common; when a bad mutation happens, it dooms the alleles near it to eventual death

background selection prediction: - depends on high rates of bad mutation - regions of high recombination (close hotspots) will preserve more variation (high  $\pi$ )

same prediction as hitchhiking.

Contrast "present" (within) with "historical" (between)

as the bad mutations are eliminated by negative selection, they drag away neutral variation at nearby genes and non-coding regions.

Natural Selection

accelerates adaptation

background selection

Selection's impact on genomes of most species is HUGE!

Purifying/neg. Selection acting on some genes, eliminating bad mutations

McDonald-Kreitman

	fixed	lost	within
non-syn	A	B	
syn	C	D	

• If  $A/C = B/D \rightarrow$  can't reject neutrality  
• If  $A/C > B/D \rightarrow$  positive selection occurs if some non-syn. differences are selected.  
• If  $A/C < B/D \rightarrow$  negative selection occurs if maladaptive non-syn differences persist w/in species (could be overdominance).

• Purifying/negative/constraint selection: removing negative mutations & acting to keep the alleles unchanging  
• Positive selection: increasing a good mutation  $\rightarrow$  acting on new (not required) variant to increase its frequency  
• Balancing selection: selection is maintaining polymorphism  $\rightarrow$  could be negative frequency dominance selection or overdominance.  
• Neutral selection: null hypothesis; assume this unless find contrary evidence to convince us otherwise.

Positive association between recombination rate &  $\pi$  in many species

synonymous

differences presumed to be neutral (no effect on protein phenotype)

accumulate differences faster

actually more potential nonsynonymous changes than there are synonymous changes between species, if we assumed all base changes were neutral

dN/ds

ratio of nonsyn. to synon. differences estimates non-neutral changes relative to neutral changes.

Estimates how much "non-neutral" (nonsyn.) evolution has happened relative to "neutral" (syn.) evolution.

dN/ds can be too conservative for finding adaptive a.a. changes  $\rightarrow$  many false negatives

ME & dN/ds tests apply to specific genes of interest & identify some evolutionary forces acting on them in particular.

nonsynonymous

differences affect the protein; cause a different amino acid to be placed into the protein, & may affect phenotype.

differences may be: - detrimental (if original protein was better) - neutral (if no effect on protein function) or advantageous (if new protein is better!)

may spread quickly w/in one species if positively selected

dN = # of non-synonymous changes per non-synonymous site

dS = # of synonymous changes per synonymous site.

• dN/ds = 1  $\rightarrow$  can not reject neutrality (unlikely realistically)  
• dN/ds < 1  $\rightarrow$  lots of "constraint"; changes favored  
• dN/ds > 1  $\rightarrow$  selection driving rapid change, multiple a.a. changes favored

dN/ds < 1 much more common than dN/ds > 1