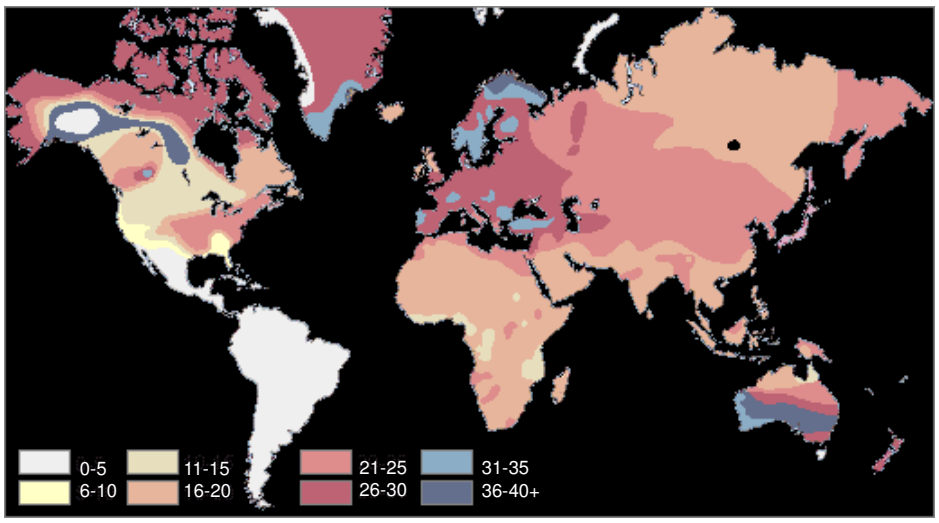
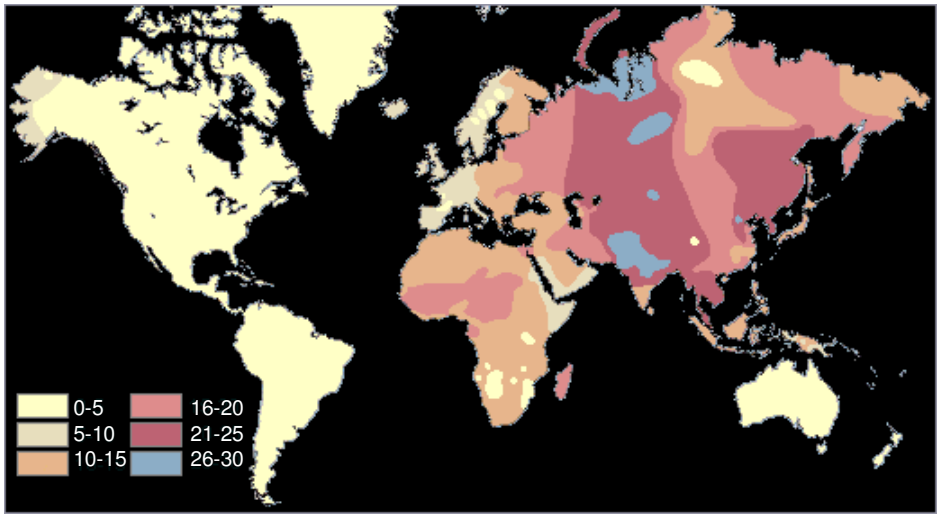
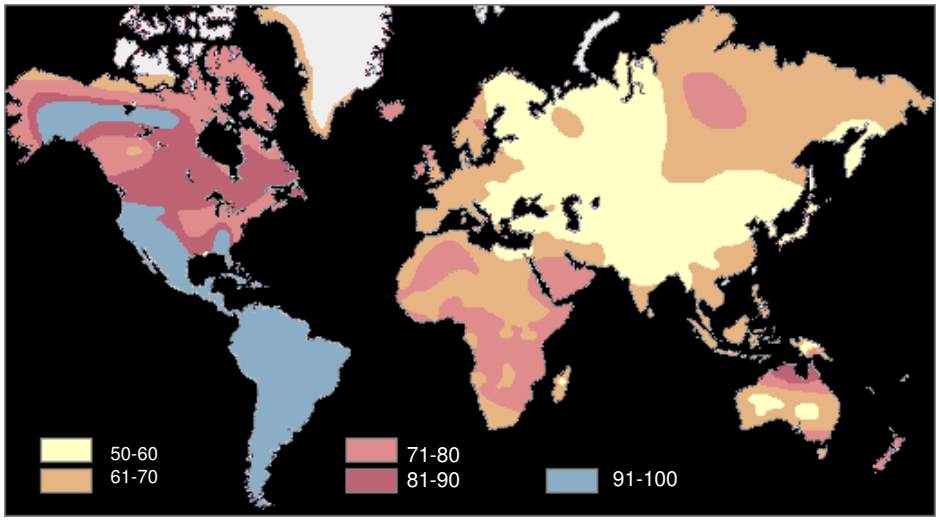
distribution of the A allele (I^A)



distribution of the B allele (I^B)



distribution of the O allele (*i*)



Genetic variation in a population is affected by two sets of processes:

1. Genetic -- mutation, recombination, independent assortment, transposition, meiotic drive

2. Ecological -- changes in population size, dispersal, mating system, environmental variation

How do these processes affect population genetic variation ?

Population Genetics

measuring population genetic variation

quantifying population genetic variation
- genotype frequencies
- allele frequencies

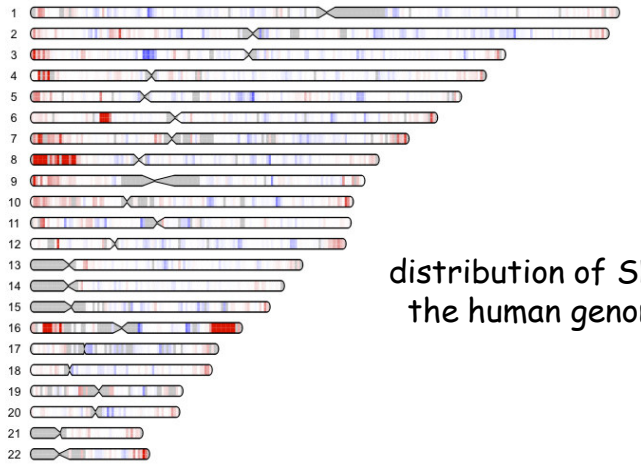
Hardy-Weinberg Equilibrium

testing for HWE

HWE and sex-linkage

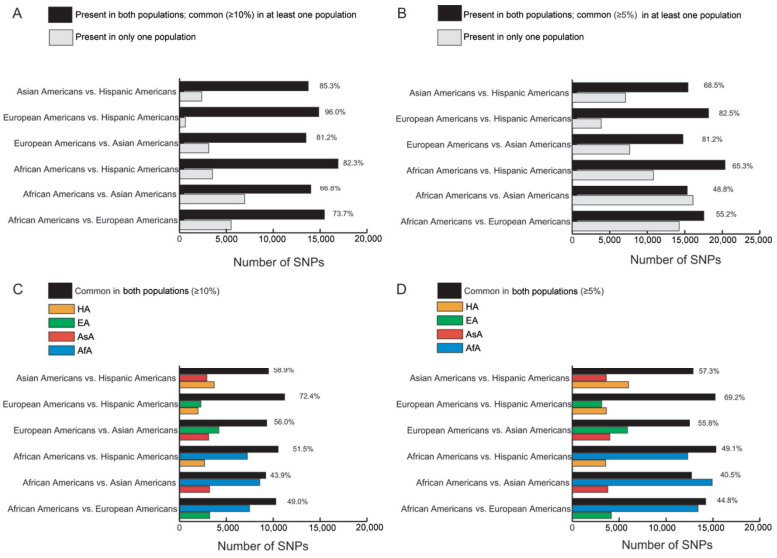
natural selection

Measuring population genetic variation -- SNPs



distribution of SNPs in the human genome

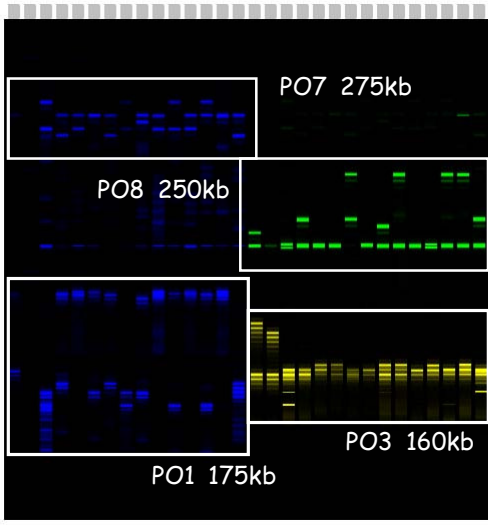
Wellcome Trust, Sanger Institute, 2010



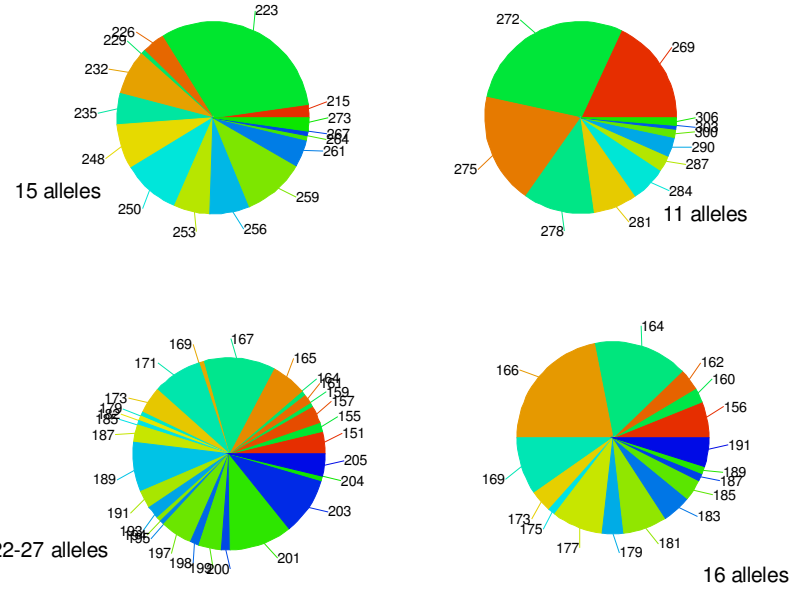
Distribution of common SNPs among Latino/Hispanic, African, Asian, and European Americans. *C* and *D*, The percentage of common SNPs common in both populations (*black bars*) compared with SNPs common in only each population compared: African Americans (AfA) (*blue bars*), Asian Americans (AsA) (*red bars*), European Americans (EA) (*green bars*), and Latino/Hispanic Americans (HA) (*orange bars*). Overall, only a modest percentage (44%–72%) of SNPs common in at least one population are common in both populations. A substantial proportion of common SNPs in African Americans are common only in African Americans.

Guthery et al 2007 AJHG81:1221

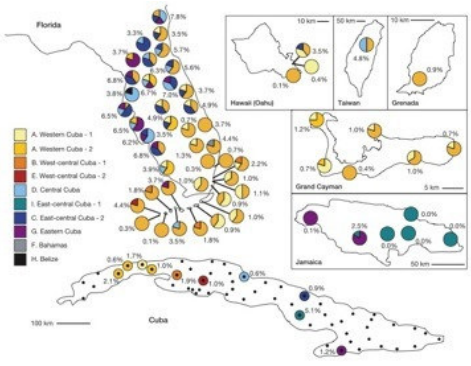
Measuring population genetic variation -- microsatellites



harvester ants,
Pogonomyrmex occidentalis



Measuring population genetic variation -- haplotypes



Anolis sagrei

Quantifying population genetic variation:

$$\text{genotype frequency} = \frac{\#\text{ particular genotype}}{\text{total number of individuals}}$$

$$\text{allele frequency} = \frac{\#\text{ particular allele}}{\text{total number of alleles}}$$

$$\text{allele frequency} = f(\text{homozygotes}) + \tfrac{1}{2} f(\text{heterozygotes})$$

$$f(A_1) = f(A_1A_1) + \tfrac{1}{2} f(A_1A_2)$$

$$f(A_2) = f(A_2A_2) + \tfrac{1}{2} f(A_1A_2)$$

by the law of proportions, both genotype and allele frequencies <u>always</u> sum to one

$$f(A_1A_1) + f(A_1A_2) + f(A_2A_2) = 1$$

and $f(A_1) + f(A_2) = 1$

genotype	A_1A_1	A_1A_2	A_2A_2	total
number	670	200	130	1000
genotype frequency	$\frac{670}{1000}$	$\frac{200}{1000}$	$\frac{130}{1000}$	
genotype frequency	0.67	0.20	0.13	

frequency of A_1 = $0.67 + \frac{1}{2}(0.20) = 0.77$

frequency of A_2 = $0.13 + \frac{1}{2}(0.20) = 0.23$

Hardy-Weinberg Equilibrium

In the presence of certain conditions, the genotype frequencies of a population will be stable over time, and will be directly predictable from the allele frequencies.

If the population is not at equilibrium, under these conditions, it will achieve it after one generation of random mating.

Assumes: no mutation, no selection, infinite population size, no gene flow, random mating

Null model for describing population genetic variation

single locus with two alleles: A_1, A_2

$f(A_1) = p \quad f(A_2) = q$

genotype frequencies:

$f(A_1A_1) = P \quad f(A_1A_2) = H \quad f(A_2A_2) = Q$

allele frequencies:

$p = P + \frac{1}{2} H \qquad q = Q + \frac{1}{2} H$

Type of Mating male x female	Freq.	Offspring genotype frequencies A_1A_1	A_1A_2	A_2A_2
$A_1A_1 \times A_1A_1$	P^2	$1\,P^2$		
$A_1A_1 \times A_1A_2$	PH	$\frac{1}{2}PH$	$\frac{1}{2}PH$	
$A_1A_2 \times A_1A_1$	PH	$\frac{1}{2}PH$	$\frac{1}{2}PH$	
$A_1A_1 \times A_2A_2$	PQ		$1\,PQ$	
$A_2A_2 \times A_1A_1$	PQ		$1\,PQ$	
$A_1A_2 \times A_1A_2$	H^2	$\frac{1}{4}H^2$	$\frac{1}{2}H^2$	$\frac{1}{4}H^2$
$A_1A_2 \times A_2A_2$	HQ		$\frac{1}{2}HQ$	$\frac{1}{2}HQ$
$A_2A_2 \times A_1A_2$	HQ		$\frac{1}{2}HQ$	$\frac{1}{2}HQ$
$A_2A_2 \times A_2A_2$	Q^2			$1\,Q^2$
Total	$(P + H + Q)^2$	$(P + \frac{1}{2}H)^2$	$2(P + \frac{1}{2}H)(Q + \frac{1}{2}H)$	$(Q + \frac{1}{2}H)^2$
	1	p^2	$2pq$	q^2

Determing whether a population is in HWE:

1) calculate observed genotype frequencies
2) calculate allele frequencies
3) calculate expected genotype frequencies
4) compare observed and expected genotype frequencies
   - test for goodness of fit using a chi-square

Note: if there are only two alleles with complete dominance, it is not possible to test for HWE

testing whether a population is in HWE:

genotype	MM	MN	NN	total
number	270	180	550	1000
genotype frequency	$\frac{270}{1000}$	$\frac{180}{1000}$	$\frac{550}{1000}$	
geno. freq.	0.27	0.18	0.55	

allele frequencies:

$f(M) = 0.27 + \frac{1}{2}(0.18) = 0.36$

$f(N) = 0.55 + \frac{1}{2}(0.18) = 0.64$

testing whether a population is in HWE:

genotype	MM	MN	NN
observed frequencies	0.27	0.18	0.55
expected frequencies	$f(M)^2$	$2f(M)f(N)$	$f(N)^2$
$f(M) = 0.36$ $f(N) = 0.64$	$(0.36)^2$	$2(0.36)(0.64)$	$(0.64)^2$
	~0.13	~0.46	~0.41
expected number	130	460	410

Chi-square test for goodness-of-fit:

$$\chi^2_{dof} = \sum \frac{(obs - exp)^2}{exp}$$

$$= \frac{(270 - 130)^2}{130} + \frac{(180 - 460)^2}{460} + \frac{(550 - 410)^2}{410}$$

$$= 369.0 \quad p \ll 0.001$$

degrees of freedom = 3 - 1 - 1 = 1

Type of Mating male x female	Freq.	Offspring genotype frequencies MM	MN	NN
MM x MM	(.27)(.27)	$1(.27)^2$		
MM x MN	(.27)(.18)	$\frac{1}{2}(.27)(.18)$	$\frac{1}{2}(.27)(.18)$	
MN x MM	(.18)(.27)	$\frac{1}{2}(.18)(.27)$	$\frac{1}{2}(.18)(.27)$	
MM x NN	(.27)(.55)		$1(.27)(.55)$	
NN x MM	(.55)(.27)		$1(.55)(.27)$	
MN x MN	(.18)(.18)	$\frac{1}{4}(.18)^2$	$\frac{1}{2}(.18)^2$	$\frac{1}{4}(.18)^2$
MN x NN	(.18)(.55)		$\frac{1}{2}(.18)(.55)$	$\frac{1}{2}(.18)(.55)$
NN x MN	(.55)(.18)		$\frac{1}{2}(.55)(.18)$	$\frac{1}{2}(.55)(.18)$
NN x NN	(.55)(.55)			$1(.55)^2$
Total	$(.27 + .18 + .55)^2$	$(.27 + \frac{1}{2}(.18))^2$	$2(.27 + \frac{1}{2}(.18))(.55 + \frac{1}{2}(.18))$	$(.55 + \frac{1}{2}(.18))^2$
	1	$(.36)^2$	$2(.36)(.64)$	$(.64)^2$

Hardy-Weinberg Equilibrium

IS NOT

$p^2 + 2pq + q^2 = 1$

genotype frequencies must always sum to 1
allele frequencies must always sum to 1

$(p + q)^2 = p^2 + 2pq + q^2 = 1$

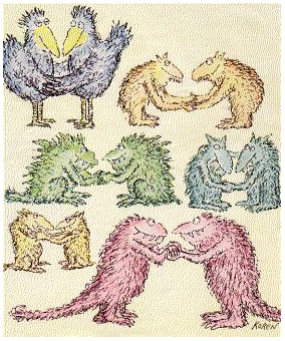
HWE for multiple alleles: $(p + q + r + ..)^2$

Processes that maintain or increase genetic variation

mutation
recombination
disassortative mating
gene flow
natural selection
- heterozygote advantage
- frequency-dependent
- disruptive

Processes that reduce genetic variation

asexual reproduction
small population size
inbreeding
assortative mating
natural selection
- directional



HWE and sex-linked genes

autosomes: half of the alleles in each sex
sex chromosomes: 2/3 alleles in the homogametic (XX) sex
1/3 alleles in the heterogametic (XY) sex

if females are heterogametic,

A_1A_1 A_1A_2 A_2A_2 A_1^ A_2^

males females

allele frequencies are sex-specific: p_m, q_m and p_f, q_f

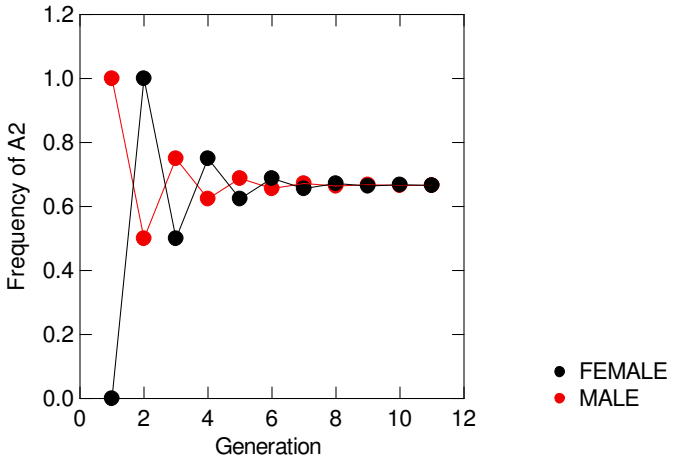
Under random mating:

$q_m' = \frac{1}{2}(q_m + q_f)$ males get an X-chromosome from each parent

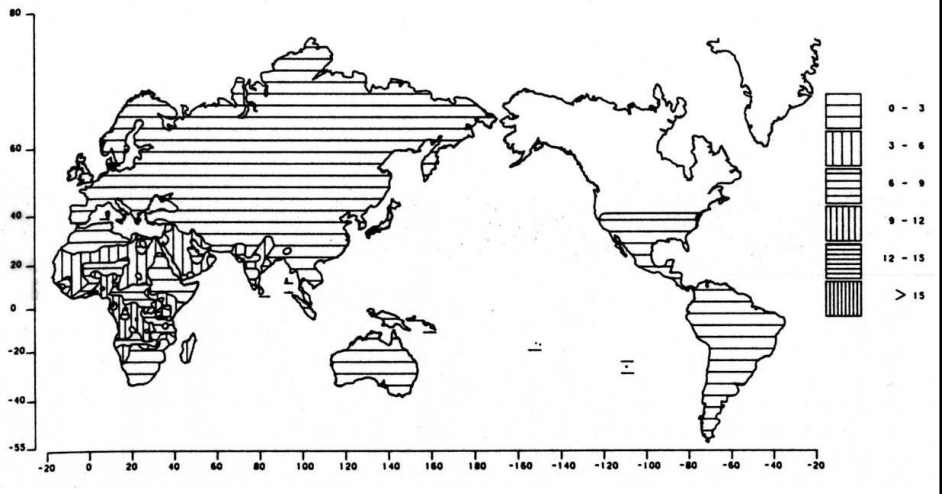
$q_f' = q_m$ females get their only X-chromosome from their father

if $q_m \neq q_f$, oscillatory approach to equilibrium

$\hat{p} = \frac{2}{3}p_m + \frac{1}{3}p_f$ $\hat{q} = \frac{2}{3}q_m + \frac{1}{3}q_f$



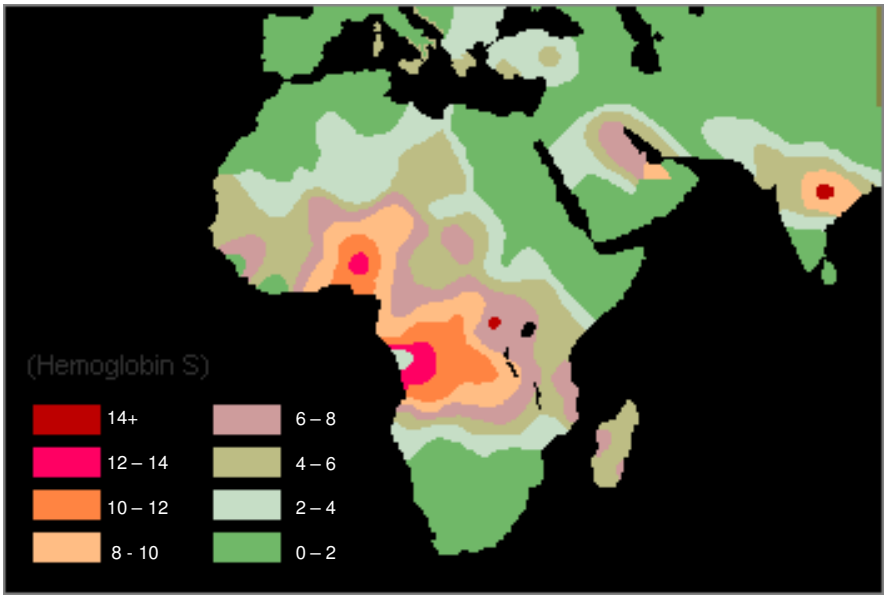
distribution of the sickle cell allele

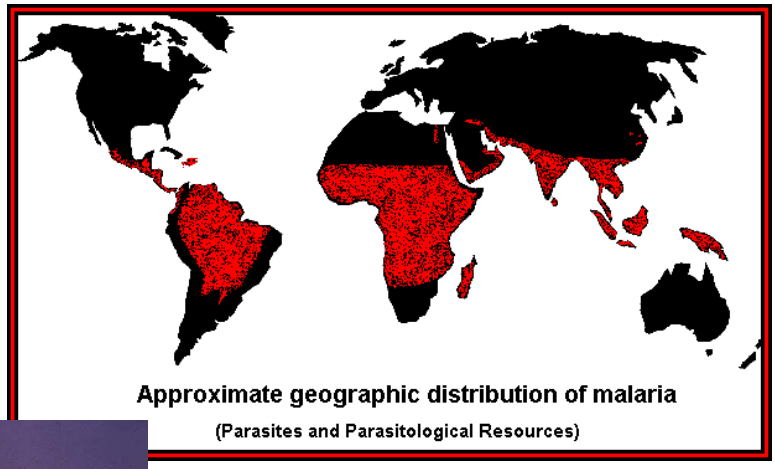


Polymorphism at the sickle cell locus

	hb^Ahb^A	hb^Ahb^S	hb^Shb^S
phenotype	normal	normal	severe anemia
pr(survival)	100%	100%	1%

distribution of the Sickle Cell allele





Fitness = individual's genetic contribution to the next generation due to differential survival and/or reproduction

absolute fitness, W_i = #offspring, lifespan, etc
relative fitness, w_i = contribution relative to other genotypes

$w_i = W_i/W_{max}$

		A_1A_1	A_1A_2	A_2A_2
Pr(escape)	W_i	80%	60%	20%
	w_i	1.0	0.75	0.25

Polymorphism at the sickle cell locus

	hb^Ahb^A	hb^Ahb^S	hb^Shb^S
phenotype	normal	normal	severe anemia
relative fitness			
temperate	1.0	1.0	0.01
malaria	0.88	1.0	0.14

Take-home points

- population genetic variation is described by genotype and allele frequencies
- genotype and allele frequencies can be calculated from any sample with codominant alleles
- the Hardy-Weinberg Equilibrium is a null model that describes genotype frequency variation in the absence of evolution
- deviations from Hardy-Weinberg Equilibrium indicate that a population is subject to mutation, selection, gene flow, genetic drift or non-random mating