

NAME _____

POPULATION GENETICS AND MICROEVOLUTIONARY THEORY
MIDTERM EXAMINATION

WRITE YOUR NAME ON EVERY PAGE!

The following table may be useful in answering one or more of the questions. Use it as needed:

Probability	Chi-Square Test Statistic Value			
	df = 1	df = 2	df = 3	df=4
0.10	2.706	4.605	6.251	7.779
0.05	3.841	5.991	7.815	9.488
0.01	6.635	9.210	11.345	13.277

1. Briefly define (5 points each):

a) Evolution

A change through time of the frequencies of alleles or allele combinations in the gene pool.

b) Infinite sites model

Every mutation is at a different nucleotide site from all previous mutations.

c) AMOVA

A partitioning of genetic variation within and among populations that is similar to f_{st} but that uses a molecule genetic distance to compare paired homologous sequences rather than heterozygosity.

2. A captive population of an endangered species is surveyed at an autosomal locus with two alleles, yielding the following results:

Genotype	AA	Aa	aa	Total
Number	30	60	10	100

- a. (5 points) Test the null hypothesis that this population has Hardy-Weinberg genotype frequencies.
b.

Genotype	AA	Aa	aa	Sum:	
Numbers	30	60	10	100	
Genotype Freq.	0.3000	0.6000	0.1000	1	
Alleles	A	a			
Allele Freq.	0.600	0.400		1	1 pt
H-W. Freq.	0.3600	0.4800	0.1600	1	
Exp.	36.000	48.000	16.000	100	1 pt
(o-e) ² /e	1	3	2.25	6.25	0.012419332
			Df=1, 1 pt.	chi-square, 1 pt	p-value<.05, 1 pt

NAME _____

b. (4 points) An analysis of the pedigrees of the individuals in the genetic survey reveals that their average inbreeding coefficient is 0.271. Is this result consistent with your answer to part a), and why or why not?

It is consistent. Even though part a) shows a significant excess of heterozygotes, and hence $f < 0$, since this inbreeding coefficient is calculated from pedigree data, it must be the average F , and reflects more of the role of genetic drift and not system of mating (just like in the example of Speke's gazelles).

c. (6 points) The captive population was established 3 generations ago. Using this founder generation as the reference generation, what is the inbreeding effective size of the population surveyed in part a)?

Use the equation $N_{ef} = \frac{1}{2\{1! [1! \bar{F}(t)]^{1/t}\}}$ (2 pts) with $t=3$ and $F=0.271$ (2 pts) to get $N_{ef} = 5$ (2 pts).

3. Two populations, say I and II, are polymorphic for two autosomal loci, each with two alleles (A and a at locus 1; B and b at locus 2) such the two loci are in two locus Hardy-Weinberg equilibrium within each population. The two loci are on different autosomes. The frequency of A is 0.2 and the frequency of B is 0.5 in population I; whereas the frequency of A is 0.6 and the frequency of B is 0.1 in population II. The two populations are then mixed together in equal frequencies.

a. (6 points) What is the linkage disequilibrium between these two loci in the mixed population?

Take the product of the allele frequencies to calculate the gamete frequencies within each pop:

Population I	1 pt			
gametes	AB	Ab	aB	ab
HW. Freq.	0.1	0.1	0.4	0.4
Population II	1 pt			
gametes	AB	Ab	aB	ab
HW. Freq.	0.06	0.54	0.04	0.36

Then, mix the gene pools together 50:50 to obtain:

gametes	AB	Ab	aB	ab
50:50 mix	0.08	0.32	0.22	0.38

2 pts.

Use the equation $D = (g_{AB}g_{ab} - g_{Ab}g_{aB})$ to obtain $D = -0.04$ (2 pts)

b. (6 points) What is the theoretical maximum magnitude of linkage disequilibrium in the mixed population?

$$\frac{D}{\min(p_A p_B, p_a p_b)}, D < 0$$

Use the equation $D' =$

$$\frac{D}{\min(p_A p_b, p_a p_B)}, D > 0$$

Since $D < 0$, $\min(p_A p_B, p_a p_b)$ is the theoretical maximum magnitude. (2 pts). p_A in the mixed populations is $\frac{1}{2}(0.2+0.6)=0.4$, and similarly p_B is 0.3 (2 pts), so $p_A p_B = 0.12$ and $p_a p_b = 0.42$. Hence, the theoretical maximum magnitude of D is 0.12 (2 pts).

NAME _____

c). (3 points) Assume now that the individuals in the mixed population mate at random? What is the linkage disequilibrium in the next generation produced by this random mating?

Use the equation $D_t = D_0(1-r)^t$ with $t=1$ and $r= \frac{1}{2}$ (on different autosomes) 2 pts. Hence:

$D_1 = -0.04(\frac{1}{2}) = -0.02$ (1 pt)

d). (5 points) Now assume that all members of population I are red and all members of population II are yellow, with the color difference being entirely due to the fact that the two populations had different diets? Moreover, the diet is culturally inherited such that all offspring have the diet of their mother, so these diet differences persist in the mixed population. Suppose now that there is 100% assortative mating by color phenotype in the initial mixed population, but random mating within a color phenotype. What is the linkage disequilibrium in the next generation produced by this assortative mating.

With this system of mating, no individuals in the mixed population will mate with an individual from the other ancestral population. Hence, the two historical populations simply coexist together, but in a genetically isolated state, and the original mixed gene pool persists. Hence, $D = -0.04$.

4. (5 points) Two formerly isolated populations begin to exchange 25% of their gametes every generation. How many generations will it take to reduce any initial allele frequency differences between the former isolated populations by 50%?

Use the equation $d_t = d_0(1 - 2m)^t$ with $m=0.25$ (2 pts).

The reduction per generation is therefore $1 - 2m = 0.5$. (1 pt).

Hence, it will take 1 generation to reduce the initial allele frequency differences by 50% (2 pts).

5. Two closely related species live on adjacent islands, with species A having 10,000 individuals and species B having 10,000,000 individuals. The islands used to be one island, but rising sea levels split them into two, and individuals and gametes from both species A and B are incapable of moving from one island to the other. Two autosomal (A1 and A2) and two X-linked (X1 and X2) regions of the genome characterized by no recombination in both species are sequenced in a single female individual from each species. All regions are characterized by random mating with one generation per year. Using an outgroup and the molecular clock, the following times to the most recent common ancestral DNA molecule for each genomic region is estimated as follows (measured in millions of years):

Species	A1	A2	X1	X2
A	0.9	1.1	0.6	0.9
B	0.7	0.3	0.3	0.45

a. (6 points) What are the long-term inbreeding effective sizes for species A and B averaged across all genomic regions?

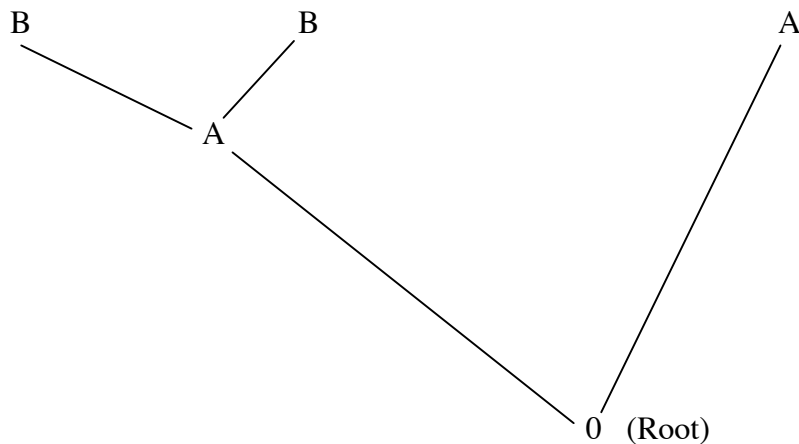
Because a single female individual is sequenced within each species, and there is random mating, we are dealing with the coalescence time of a randomly drawn pair of genes within each species. Hence,

use the equation $\text{Expected}(\text{Time to Coalesce}) = \sum_{t=1}^{\infty} t \left(\frac{1}{xN_{ef}} \right)^t \left(\frac{x-1}{xN_{ef}} \right)^{t-1} = xN_{ef}$ with time in generations=years (2 pts). Hence,

x	2	2	1.5	1.5	
Species	A1	A2	X1	X2	Average
A	0.9	1.1	0.6	0.9	
N _{ef}	450000	550000	400000	600000	500000
B	0.7	0.3	0.3	0.45	
N _{ef}	350000	150000	200000	300000	250000

(2 pts for each N_{ef}).

b. Coalescence occurs within a species for each genomic region but A2. For that region, the haplotype tree is as follows, where A indicates an existing haplotype in species A, and B indicates an existing haplotype in species B and 0 indicating the inferred root of the tree with no corresponding haplotype in the current sample:



(2 points) Is this haplotype tree evolutionarily inconsistent with the other three genomic regions and why?

No, haplotype trees are not always species trees and can have different topologies from one another.

(2 points) What evolutionary force(s) could explain the result at A2, assuming neutrality?

Lineage sorting due to genetic drift of ancestral polymorphisms.

c. (5 points) Assuming neutrality and the infinite sites model and a mutation rate of 10^{-6} per genomic region, what are the expected heterozygosities for a randomly chosen autosomal region in species A and in species B?

Use the equations $\bar{F}_{eq} = \frac{1}{4N\mu+1}$ and $1! \bar{F}_{eq} = H_{eq} = 1! \frac{1}{+1} = \frac{1}{+1}$ (2 pts).

Thus use the N_{ef} estimated in part a) (1 pt) to get that the expected heterozygosity in A is 0.667 and the expected heterozygosity in B is 0.5 (2 pts).

d. (5 points) Which species, A or B, should have the higher rate of molecular evolution under the assumptions described in part c)?

Neither, the rate of molecular evolution is the same in both species.

NAME _____

6. (25 points) Discuss and contrast the role of recombination in the following measured genotype approaches in quantitative genetics: linkage analysis with interval mapping, linkage disequilibrium mapping (GWAS), admixture mapping, and nested clade analysis (use the other side of this page if necessary).