Assignment Tests

Question of interest: From what population did an individual originate?

Uses in fisheries and conservation

a. Allow managers to determine rates of movements among populations
c. Assigning individuals to breeding populations
d. Assigning individuals to hatcheries of origin

Assumptions:

a. Hardy-Weinberg equilibrium (the expected frequency of a genotype can be determine from population allele frequencies

b. loci are independent

c. accurate estimation of population allele frequencies (i.e., large population sample sizes)

Steps in classical assignment methods:

a. genotype unknown individual

b. calculate population allele frequencies for all possible populations of origin

c. calculate the expected frequency of the unknown individual’s genotype in each possible population of origin

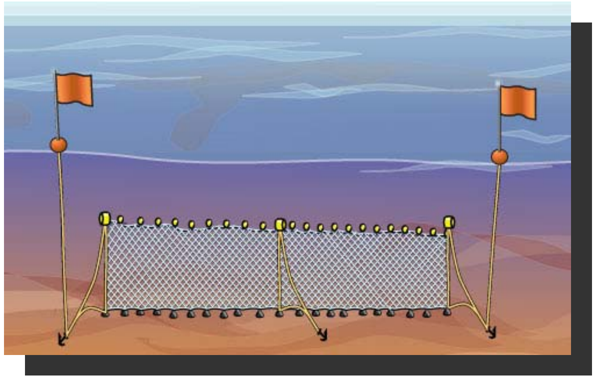
d. multiply these expected frequencies across loci for each population separately (see example)

e. assign unknown to population where its genotype is most likely to occur (highest expected frequency)

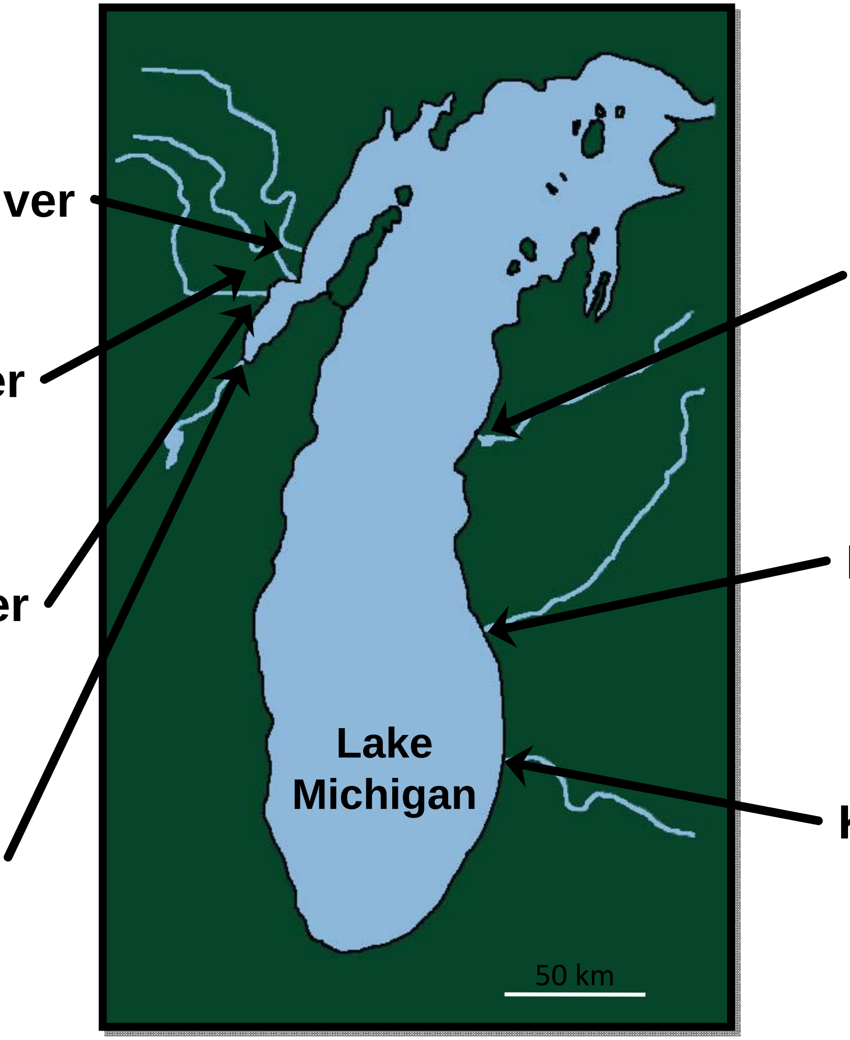
Here is an actual example from lake sturgeon sampled from locations in open waters of Lake Michigan

Baseline (river) spawning adult samples	Open water samples of unknown origin
Long handle dip nets	Commercial gill and trap nets
Captured in spring	Captured year round
1 cm^2 dorsal fin clip, stored dry	1 cm^2 dorsal fin clip, stored dry
Spawning in river	

River Mouth Samples	Open Water Samples
Near or within river mouths	Presumed independent of spawning population



Sampling locations



Menominee River

Peshto River

Oconto River

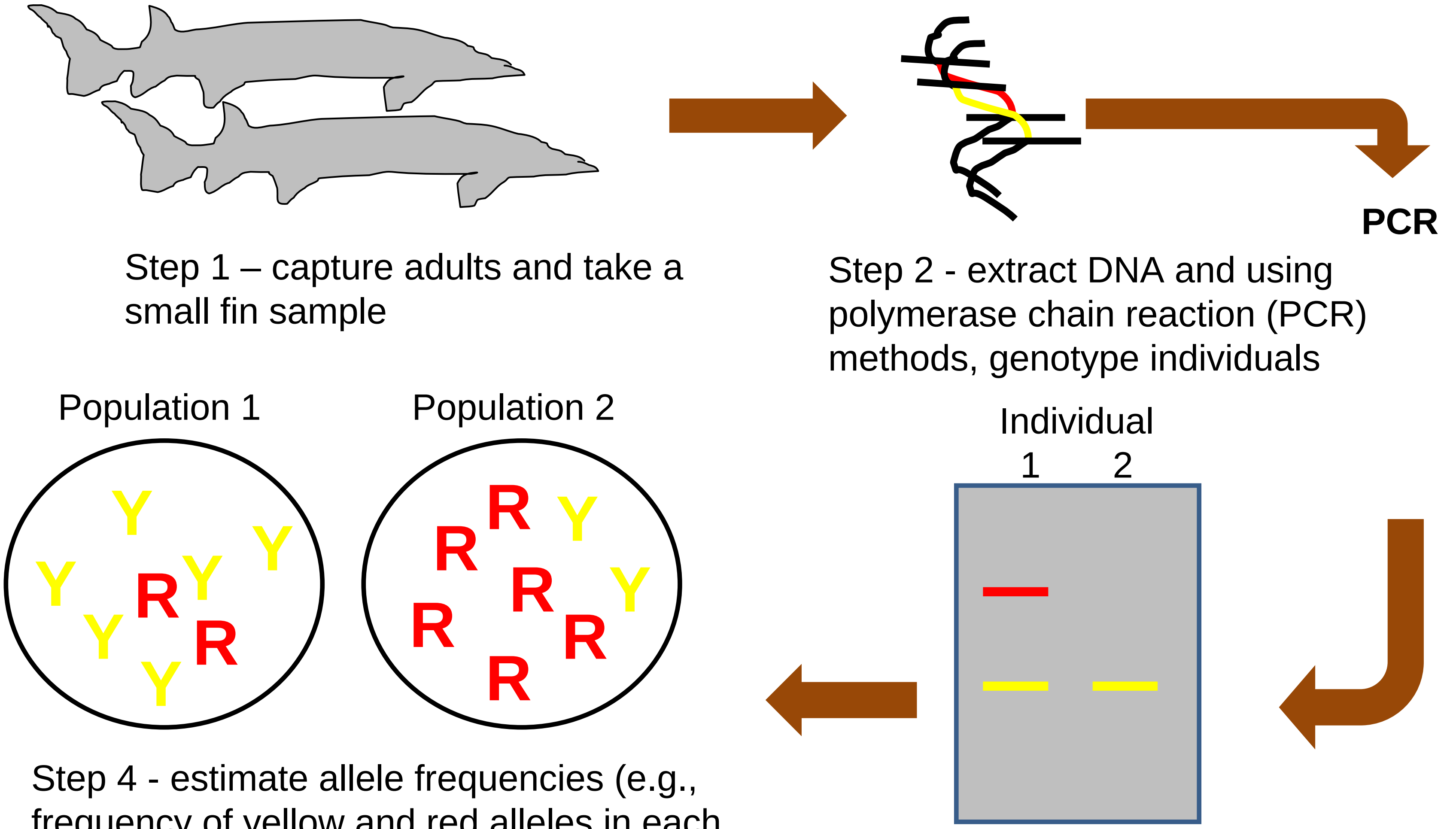
Fox River

Manistee River

Muskegon River

Kalamazoo River

Laboratory Methods

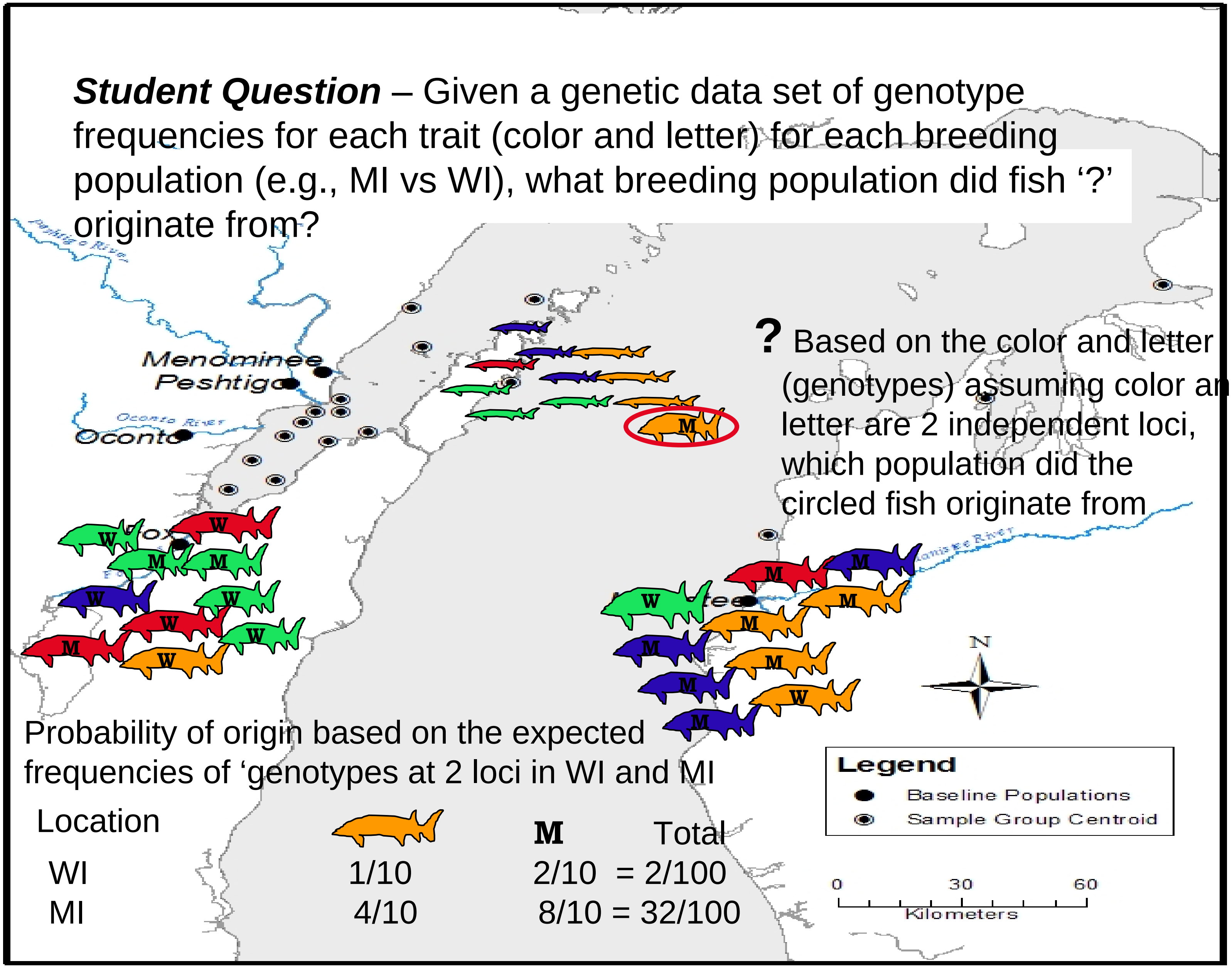


Step 1 – capture adults and take a small fin sample

Step 2 - extract DNA and using polymerase chain reaction (PCR) methods, genotype individuals

Step 3 – use electrophoresis to resolve genotypes

Step 4 - estimate allele frequencies (e.g., frequency of yellow and red alleles in each spawning (baseline) population



Student Question – Given a genetic data set of genotype frequencies for each trait (color and letter) for each breeding population (e.g., MI vs WI), what breeding population did fish '?' originate from?

? Based on the color and letter (genotypes) assuming color and letter are 2 independent loci, which population did the circled fish originate from

Probability of origin based on the expected frequencies of 'genotypes at 2 loci in WI and MI

Location		**M**	Total
WI	1/10	2/10	= 2/100
MI	4/10	8/10	= 32/100

Example: two loci where the allele frequencies for alleles ‘R’ and ‘r’ for locus 1 and ‘T’ and ‘t’ for locus 2 are estimated for 2 populations (A and B). Unlike to example with color and letters above, for most loci, genotypes are composed of 2 alleles (one inherited from the mother and father). Thus, we must use the Hardy Weinberg principle to estimate the ‘expected’ frequency of genotypes for individuals from each population based on the population allele frequencies that can be estimated by sampling and genotyping fish from each spawning population

	Locus 1			Locus 2		
pop A	freq(R)	freq(r)		freq(T)	freq(t)	
	0.5	0.5		0.4	0.6	
	Pr(RR)	Pr(Rr)	Pr(rr)	Pr(TT)	Pr(Tt)	Pr(tt)
	p^2	$2p(1-p)$	$(1-p)^2$	p^2	$2p(1-p)$	$(1-p)^2$
Hardy Weinberg Expected frequency	**0.25**	**0.5**	**0.25**	**0.16**	**0.48**	**0.36**
pop B	0.9	0.1		0.6	0.4	
	0.81	**0.18**	**0.01**	**0.36**	**0.48**	**0.16**

Unknown genotype: RRTT

Likelihood of observing RRTT from pop A: (0.25)(0.16) = **0.04**

Likelihood of observiing RRTT from pop B: (0.81)(0.36) = **0.29**

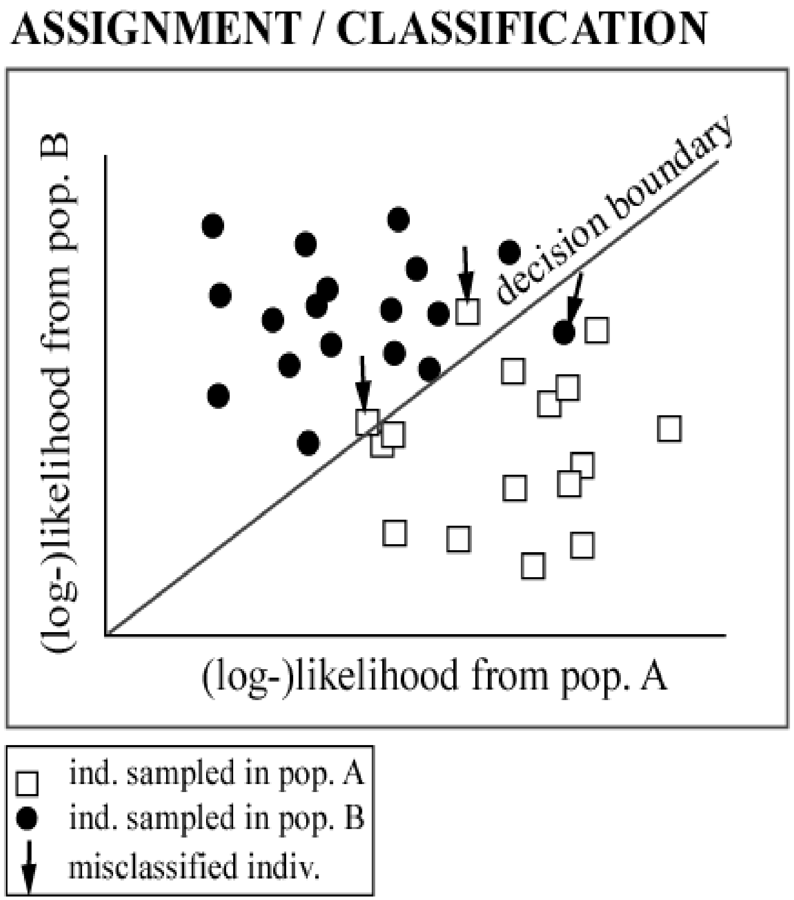
Likelihood estimators for individual assignment

$$\boxed{\frac{L_A}{L_B} = \frac{\text{Prob}(\text{genotype} \mid \hat{\theta}_A)}{\text{Prob}(\text{genotype} \mid \hat{\theta}_B)}}$$

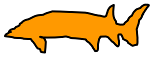
where $\hat{\theta}_A$ and $\hat{\theta}_B$ are estimates of population allele frequencies for populations A and B, respectively

Assign individual to population A of ratio >1 and to population B if ratio <1

Likelihood estimators for individual assignment



$$\frac{L_A}{L_B} = \frac{\text{Prob}(\text{genotype}\,|\,\hat{\theta}_A)}{\text{Prob}(\text{genotype}\,|\,\hat{\theta}_B)}$$

where $\hat{\theta}_A$ and $\hat{\theta}_B$
are estimates of population allele
frequencies for populations
A and B (e.g., WI and MI), respectively
for traits '**M**' and 

<u>Decision Rule</u>

Assign individual to population A
if ratio >1 and to population B
if ratio <1

<u>Incorporating Uncertainty Into Assignment Decisions</u>

Figure describing locations associated with Lesson 10 (population analysis of allele frequencies and assignment testing. The rivers circled are the Wolf and Fox Rivers in Wisconsin and the Cheboygan River in Michigan.

Scenario – In the 1980's some young lake sturgeon from the Cheboygan River and from the Wolf River were stocked into Burt Lake in Michigan. Some years later, biologists were conducting surveys in Burt Lake and found that a number of fish captured were of a size that were consistent with the fish stocked. The fish were untagged and so visually, biologists were not able to determine whether the fish were from the stocked Wisconsin fish or Michigan fish. Biologists requested that a genetic analysis be conducted.

There are several steps to the genetic analysis. First, population allele frequencies have to be estimated from the two rivers (Wolf/Fox and Cheboygan). If the rivers differ in allele frequencies, then an assignment test can be used to genotype the fish sampled in Burt Lake to determine origin.