

# Population Genetic Assessment of Comal Springs Riffle Beetle

---

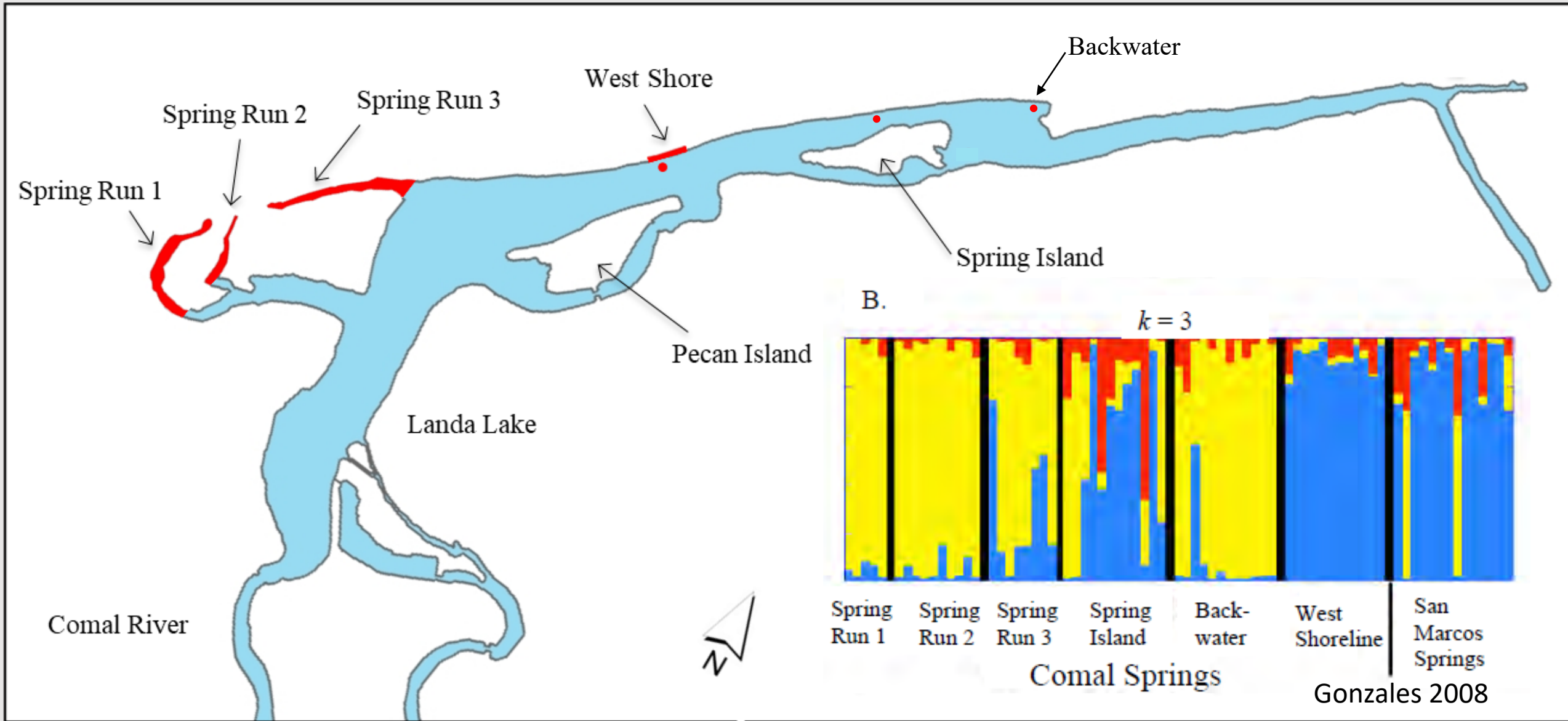
A 2022 proposal from the US Fish and Wildlife Service  
San Marcos Aquatic Resources Center

The background of the slide is a photograph of a pond with green trees and foliage reflected in the water. The text is overlaid on a white rectangular area.

# **To create a fully functioning refugia it is critical to determine the genetic diversity of the population**

- Inform Refugia collection needs to ensure the Refugia population reflects the wild population
- Assess current vs historical population status
- Assess population resiliency
- Inform restocking strategies in the event of a catastrophe

# Previous Genetic Studies Show Structure Among Sites



# Study Objectives

## Assess

Assess the fine-scale population-level genetic diversity of the Comal Springs riffle beetle across Landa Lake

## Estimate

Estimate effective population size (number of breeding adults)

## Investigate

Investigate changes in population size



# Metrics of Genetic Diversity

- Allelic Diversity
  - Number of alleles per locus ( $N_A$ )
  - Allelic richness ( $A_R$ )
    - Private alleles; show diversity specific to a site
- Heterozygosity
  - Expected Heterozygosity ( $H_E$ ) vs Observed heterozygosity ( $H_O$ )
    - Deviations from Hardy-Weinberg Equilibrium
- $F_{is}$  and  $F_{st}$  (AMOVA)
- STRUCTURE

# Metrics of Genetic Diversity: Heterozygosity

- Two Alleles at locus A
  - A and a

- Three genotypes at Locus A
  - AA, Aa, aa

- Heterozygosity measures deviations from Hardy-Weinberg Equilibrium (Null)
  - Equal allele frequency,  $A = 0.5$  and  $a = 0.5$ , heterozygosity (Aa) = 0.5
  - Unequal allele frequency,  $A = 0.33$  and  $a = 0.66$ , heterozygosity (Aa) = 0.55
  - Unequal allele frequency,  $A = 0.66$  and  $a = 0.33$ , heterozygosity (Aa) = 0.45

- Deviations from Null suggest one or more of the following assumptions are not true
  - Random mating
  - Infinite population size
  - No selection
  - No migration/gene flow
  - No mutation

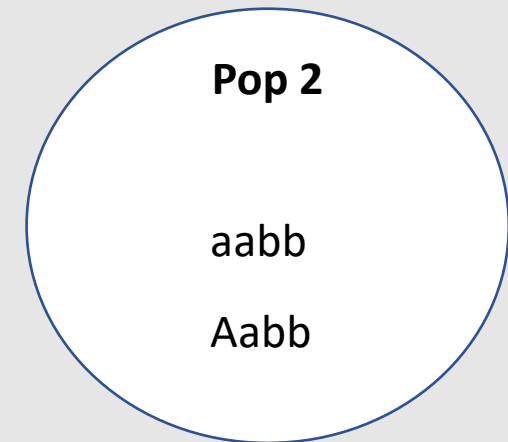
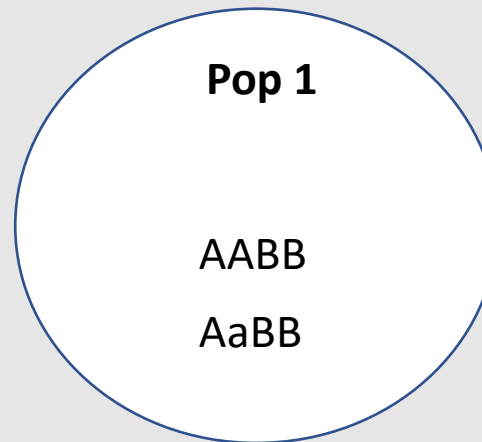
	A	a
A	AA	Aa
a	Aa	aa

# Metrics of Genetic Diversity

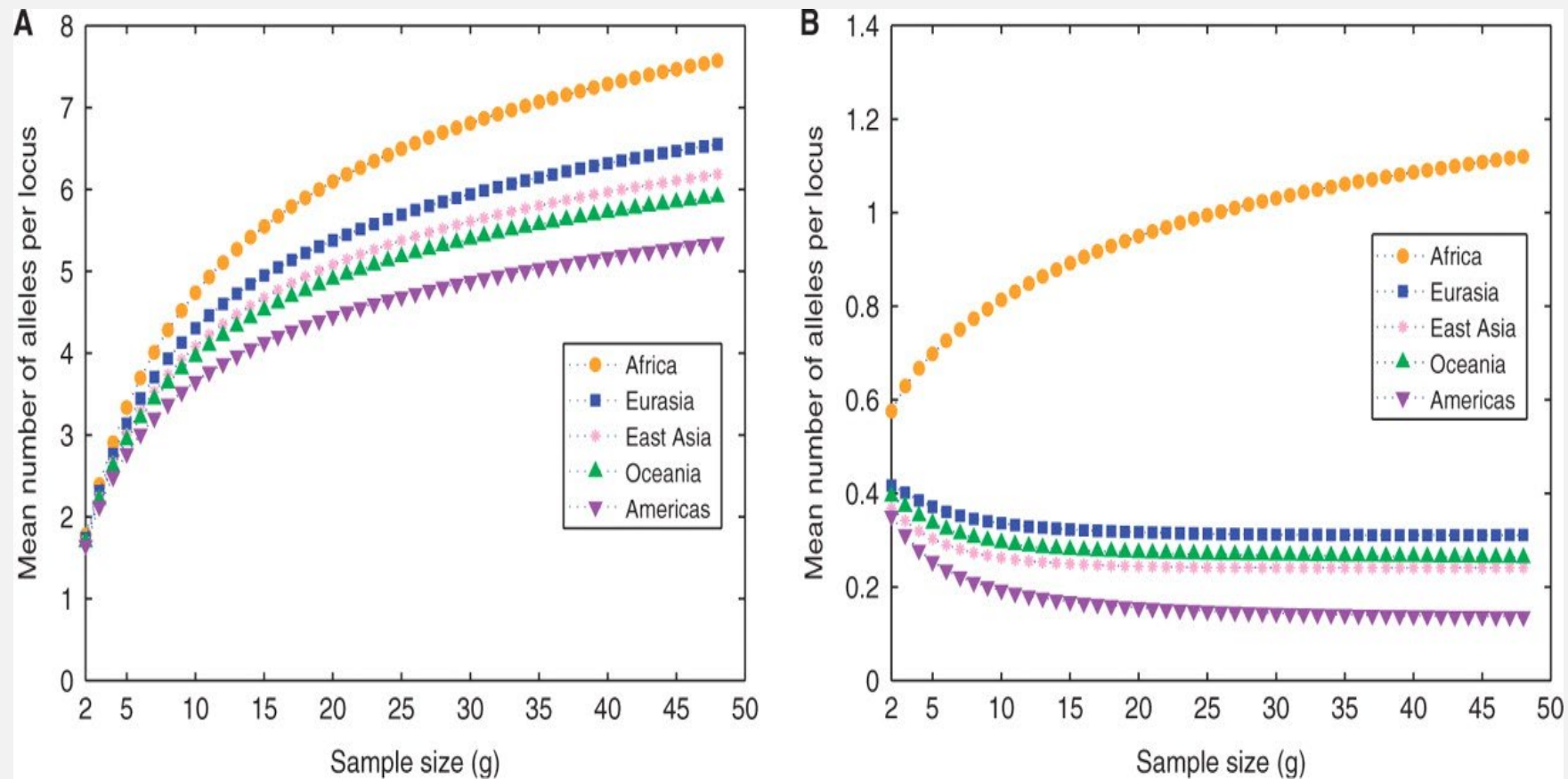
	AB	AB	ab	ab
AB	AABB	AABB	AaBb	AaBb
AB	AABB	AABB	AaBb	AaBb
ab	AaBb	AaBb	aabb	aabb
ab	AaBb	AaBb	aabb	aabb

## At this Locus

- Four alleles
  - A, a, B, b
- Three genotypes
  - AABB, aabb, AaBb
- Heterozygosity = 0.5
- Private alleles or **Allelic richness ( $A_R$ )** is increased







**Fig. 4.** The mean number of (A) distinct alleles per locus and (B) private alleles per locus, as functions of standardized sample size for five major geographic regions



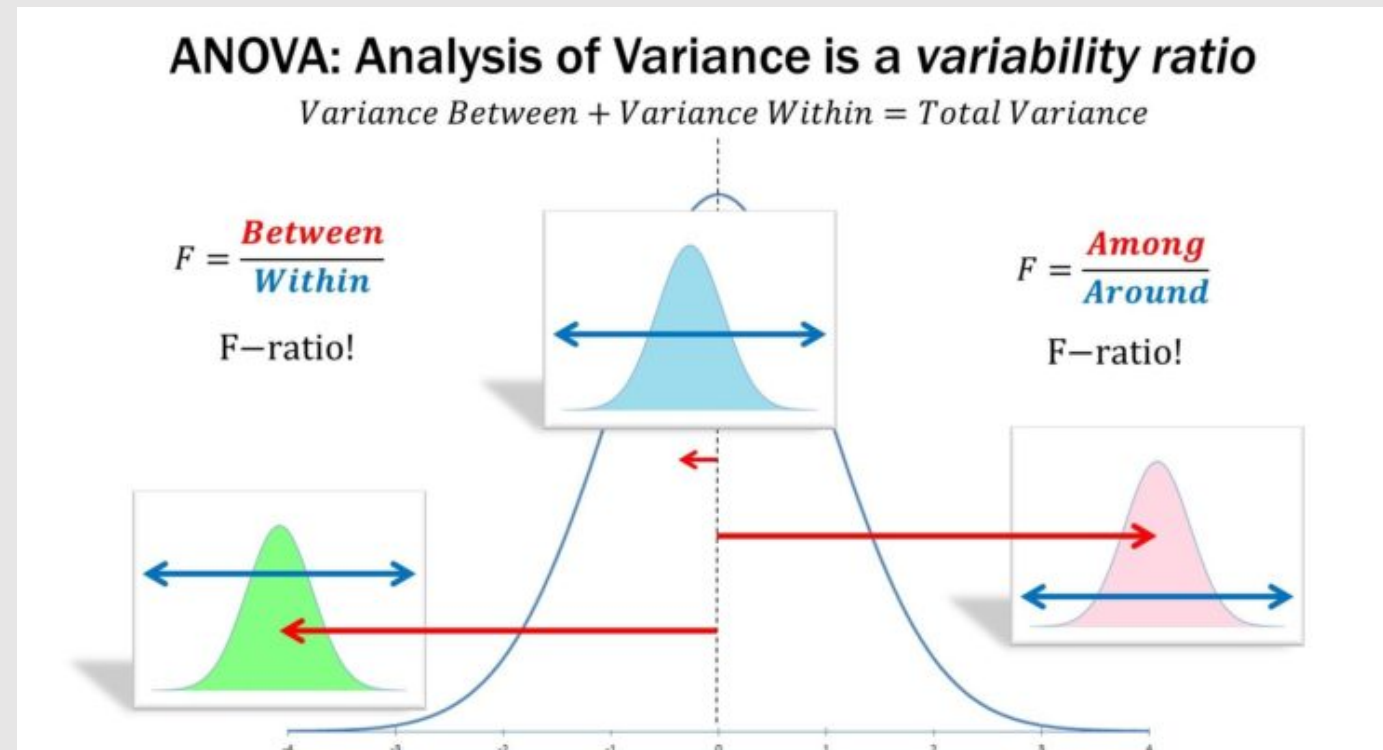
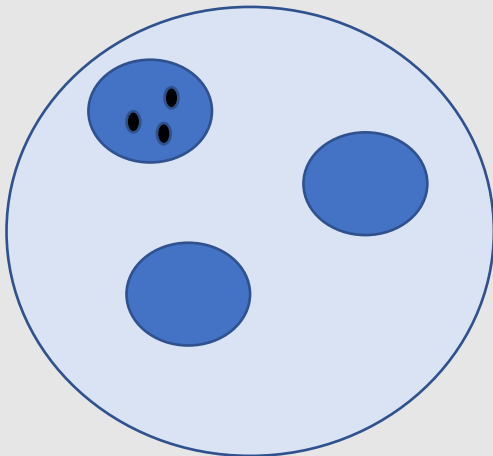
# Analysis of Molecular Variance (AMOVA)

- Basically, a one-way ANOVA
  - Determine if there are significant differences in the means of three or more independent groups

$F_{IS}$  = individual relative to subpopulation

$F_{ST}$  = subpopulation relative to total

$F_{IT}$  = individual relative to total



# STRUCTURE

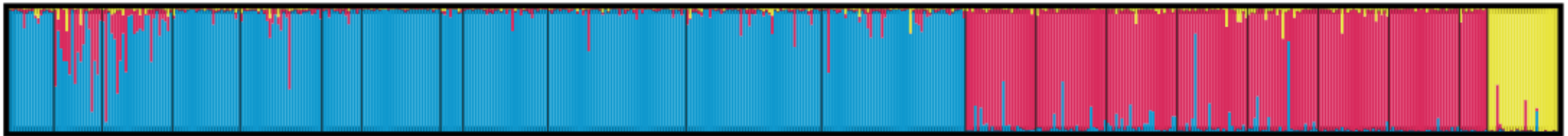
- Bayesian clustering approach using Markov Chain Monte Carlo (MCMC) estimation
- Randomly assign individuals to populations, allele frequencies are estimated and the probability of observing these estimated frequencies is calculated based on observed data
- Individuals are reassigned based on the estimated allele frequencies
- The process is repeated thousands of times until it converges on the highest probability of observing the estimated allele frequency estimates and probability of an individual's membership to a population
- STRUCTRE does this under an assumed number of populations or genetic groups (K)
- Assume No Admixture, Admixture, Gene Linkage

# STRUCTURE Plot

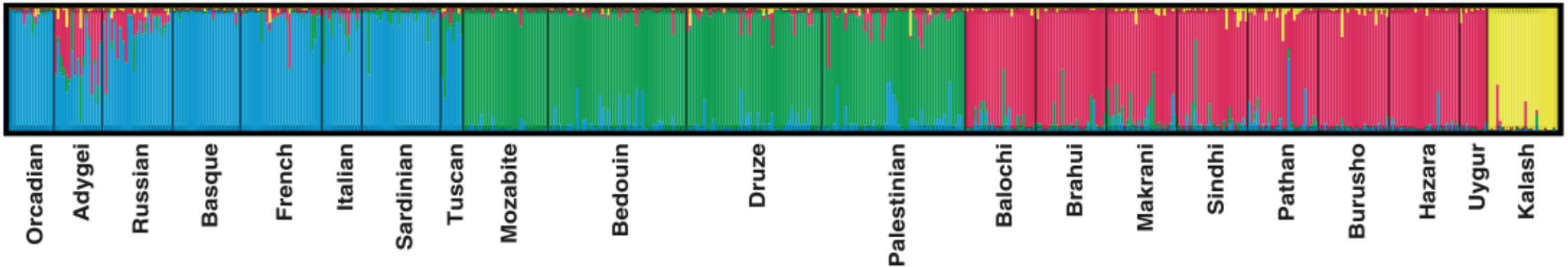
- K = the number of genetic lineages
- Different colors represent each K
- Vertical lines represent a single individual
- Individuals are grouped by population or sampling site

## Eurasia

K=3



K=4

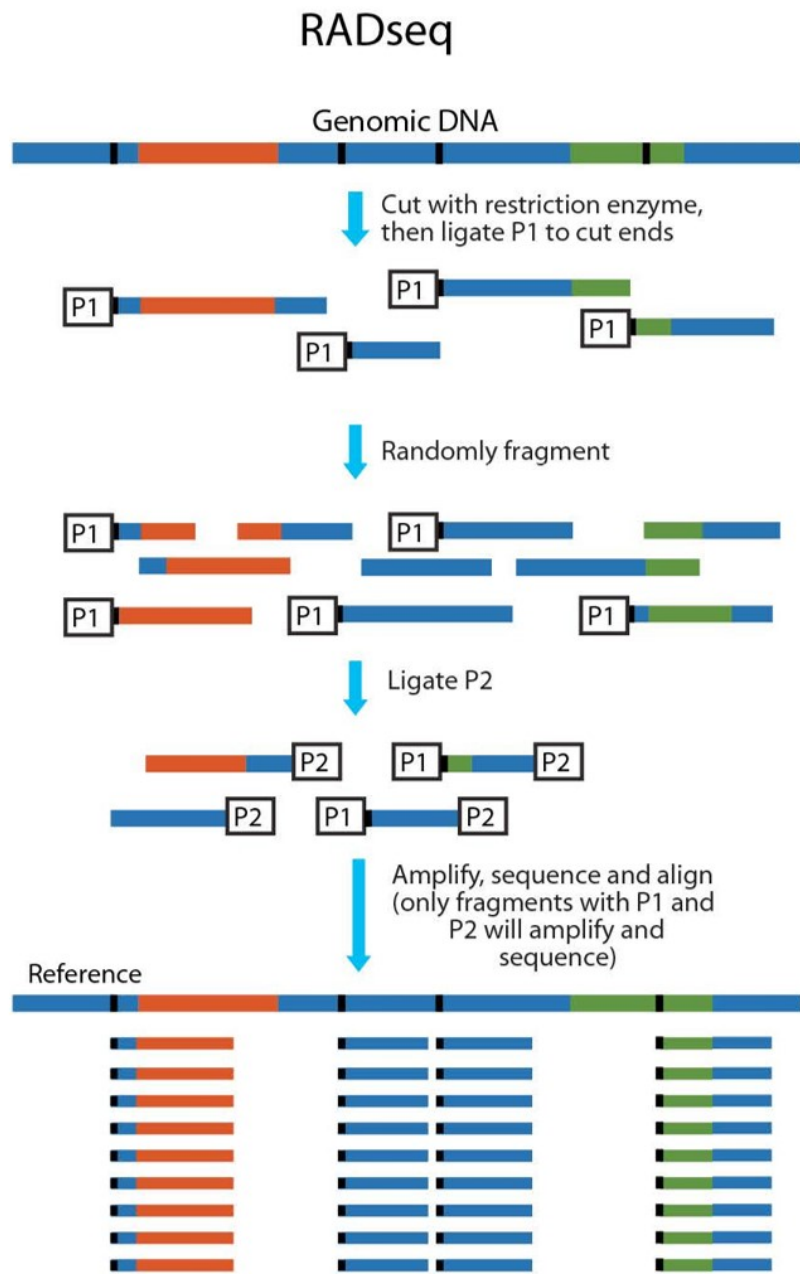


# Acquiring Samples Across Landa Lake

- Two-year occupancy study sampling 70+ sites across Landa Lake
  - Year 1 – acquiring samples
  - Year 2 – genetic analysis
- Need approximately 10 individuals per site
- Total need is approximately 700 individuals

# Assess Genetic Diversity and Structure

- Use 3RAD protocol and high through-put sequencing to generate thousands of SNPs (data points)
- Use STACKS to process the data to identify the number of loci and genetic variation
- Use the output from STACKS to generate a STRUCTURE Plot



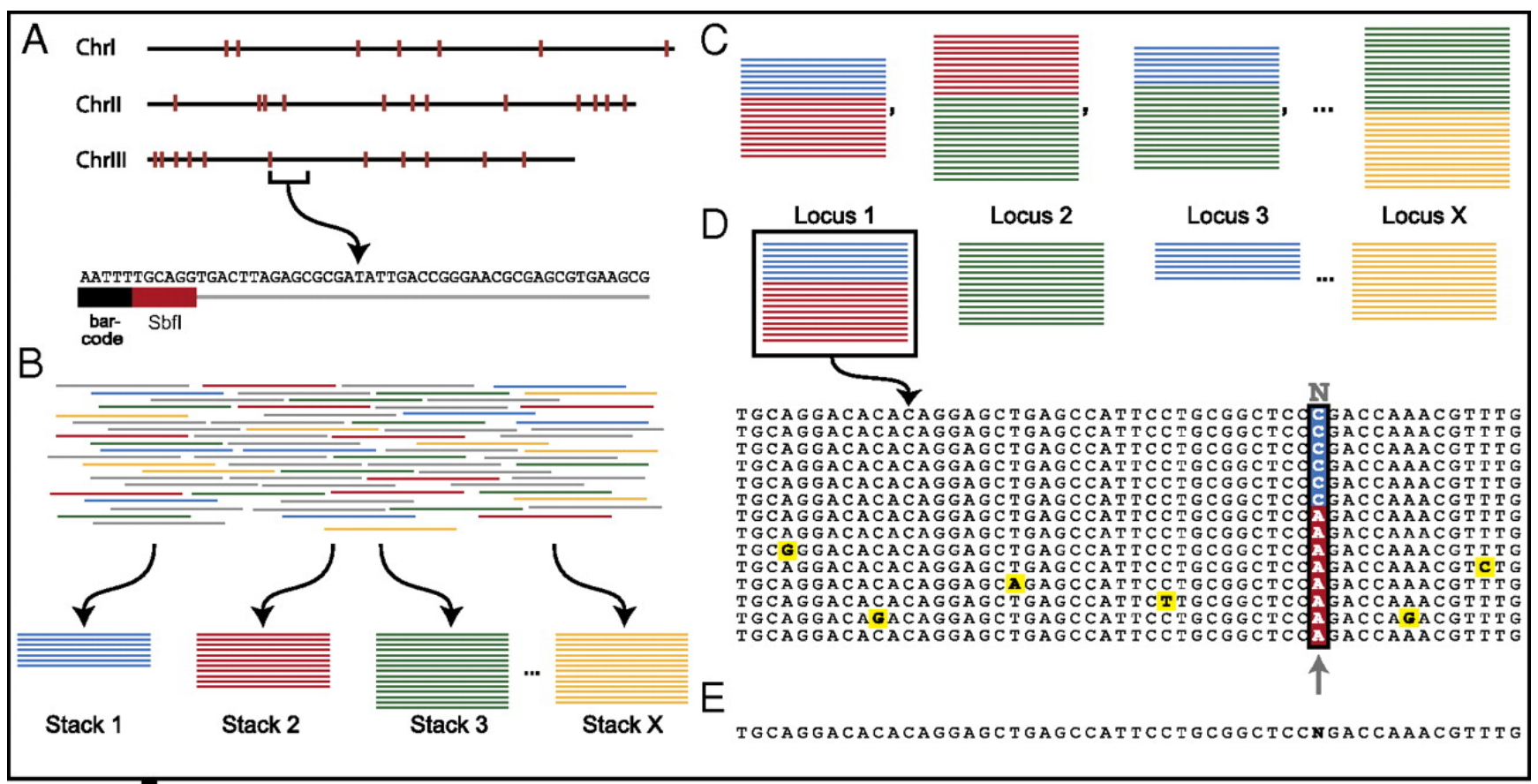
Perry et al. 2017

## Restriction-site Associated DNA Sequencing (RADseq)

- Enzymes cut genomic DNA into chunks
- Adapters (P1 and P2) are attached to DNA chunks at the enzyme cut sites
- DNA chunks are filtered and only those with both P1 and P2 adaptors are sequenced
- Thousands of DNA chunks are sequenced generating thousands of datapoints
- Thousands of copies of each datapoint is sequenced, providing high confidence in the data

Enzyme cuts DNA and all sequences include cut site. Acts as an anchor to assemble similar reads

Mixed sequences that are grouped by similarity into "Stacks"



Unique Stacks are cataloged as Loci with variable Single Nucleotide Sequences or "data points"

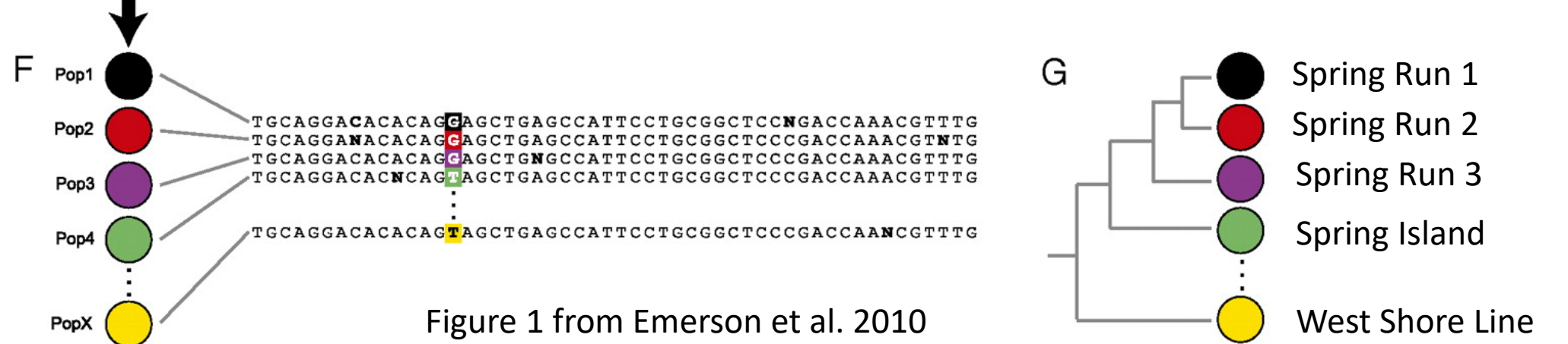


Figure 1 from Emerson et al. 2010



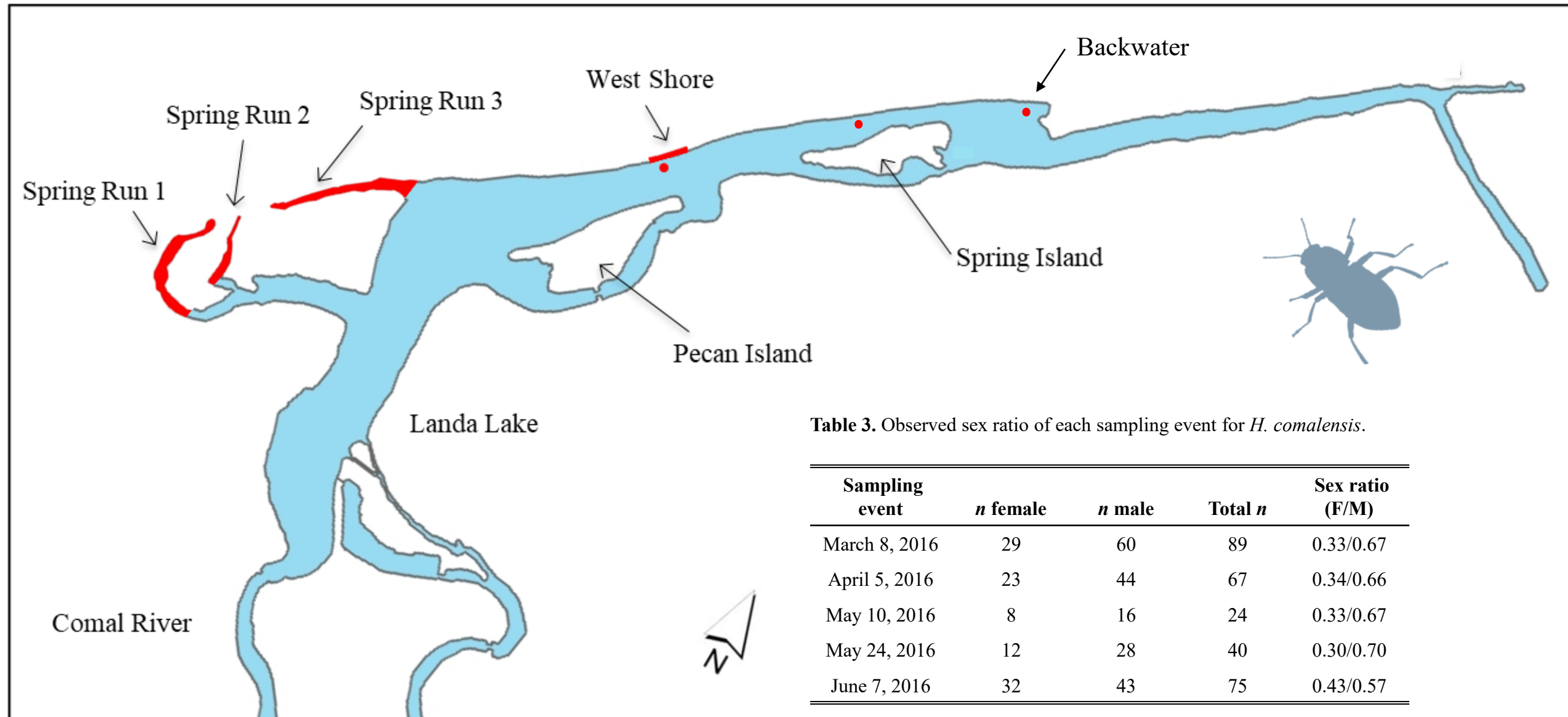
# Estimate Effective Population Size

## Why is effective population size ( $N_e$ ) important?

- Estimate of the total number of individuals contributing to reproduction
- Estimate how many males ( $N_m$ ) and females ( $N_f$ ) are breeding
  - Physical surveys estimate actual populations size
  - Not all individuals reproduce
  - Not all who reproduce produce the same number of offspring
- $N_e$  can be used to inform historical changes in population size
- Show how vulnerable the contemporary population may be to change.



# Estimate Effective Population Size

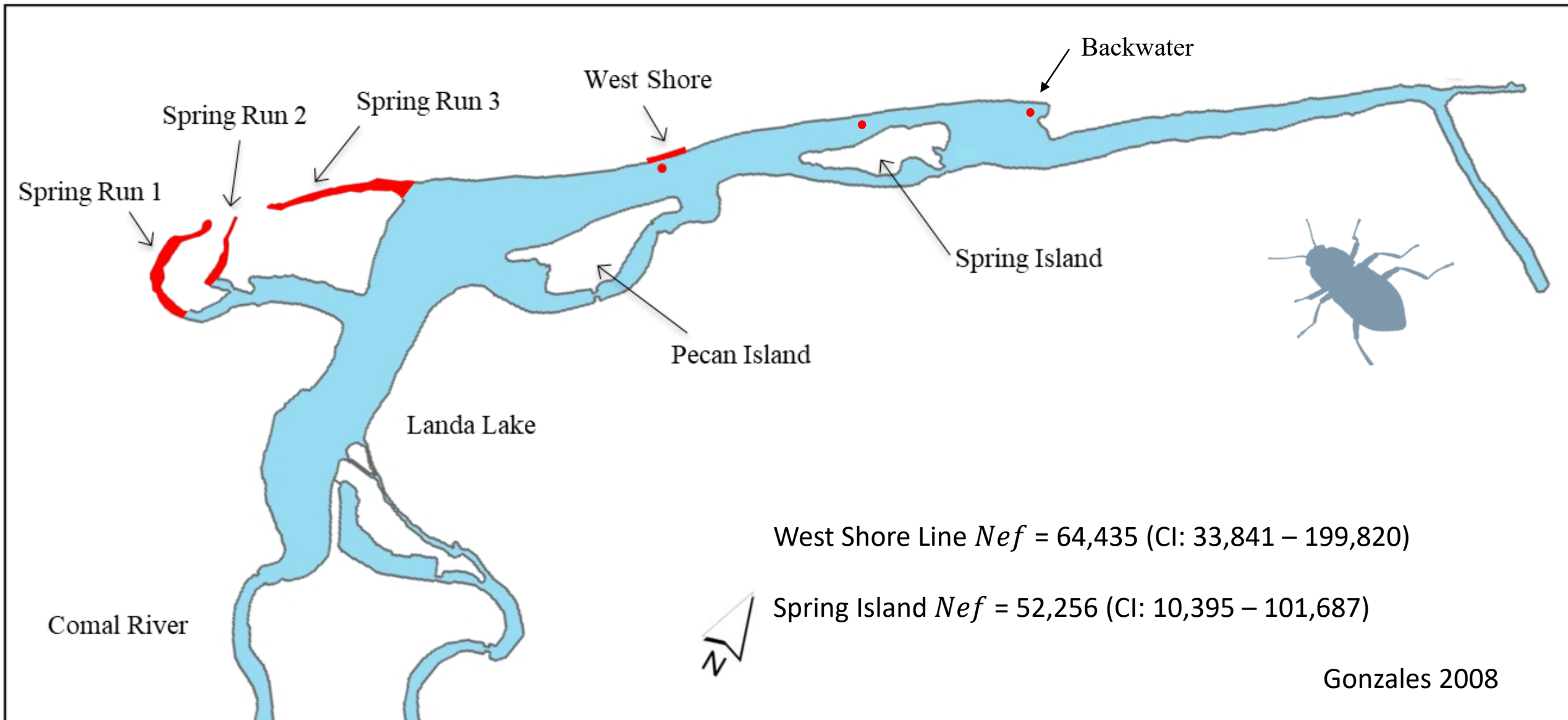


Assess

Estimate

Investigate

# Estimate Effective Population Size



Assess

Estimate

Investigate

# Estimate Effective Population Size

Effective population ( $N_e$ ) size is not equivalent to actual population size

$$N_e = \frac{4NmNf}{Nm + Nf}$$

$m$  = # of Males  
 $f$  = # of Females

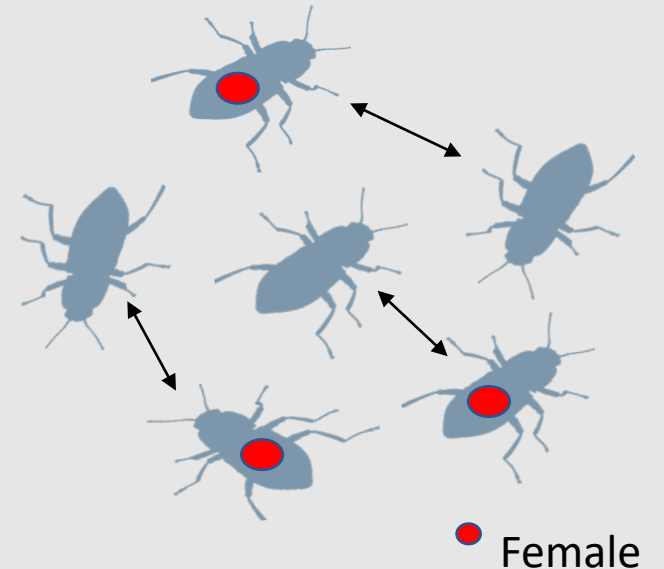
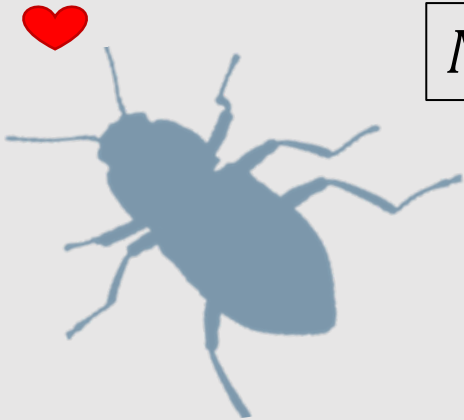
Null Hypothesis: Adults contribute equally to the next generation

$$N_e = \frac{4NmNf}{Nm + Nf}$$

20 males  
20 females

$$N_e = \frac{4(20 * 20)}{20 + 20}$$

$N_e = 40$  reflects actual population size  $N = 40$



Assess

Estimate

Investigate

# Estimate Effective Population Size

Effective population ( $N_e$ ) size is not equivalent to actual population size

$$N_e = \frac{4NmNf}{Nm + Nf}$$

$m$  = # of Males  
 $f$  = # of Females

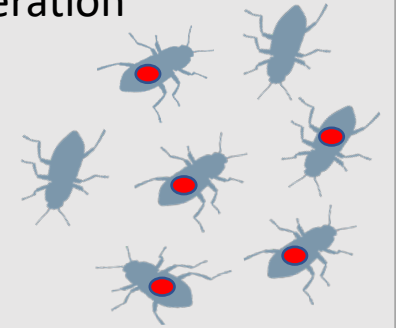


Alternate Hypothesis 2: Unequal sex ratio contributing to the next generation

$$N_e = \frac{4NmNf}{Nm + Nf}$$

$$N_e = \frac{4(10 * 30)}{10 + 30}$$

$N_e = 30$   
vs  
 $N = 40$

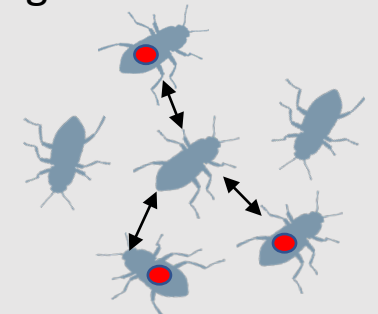


Alternate Hypothesis 1: Adults contribute disproportionally to the next generation

$$N_e = \frac{4NmNf}{Nm + Nf}$$

$$N_e = \frac{4(10 * 20)}{10 + 20}$$

$N_e = 27$   
vs  
 $N = 40$



● Female

Effective population size ( $N_e$ ) does not reflect actual population size

Assess

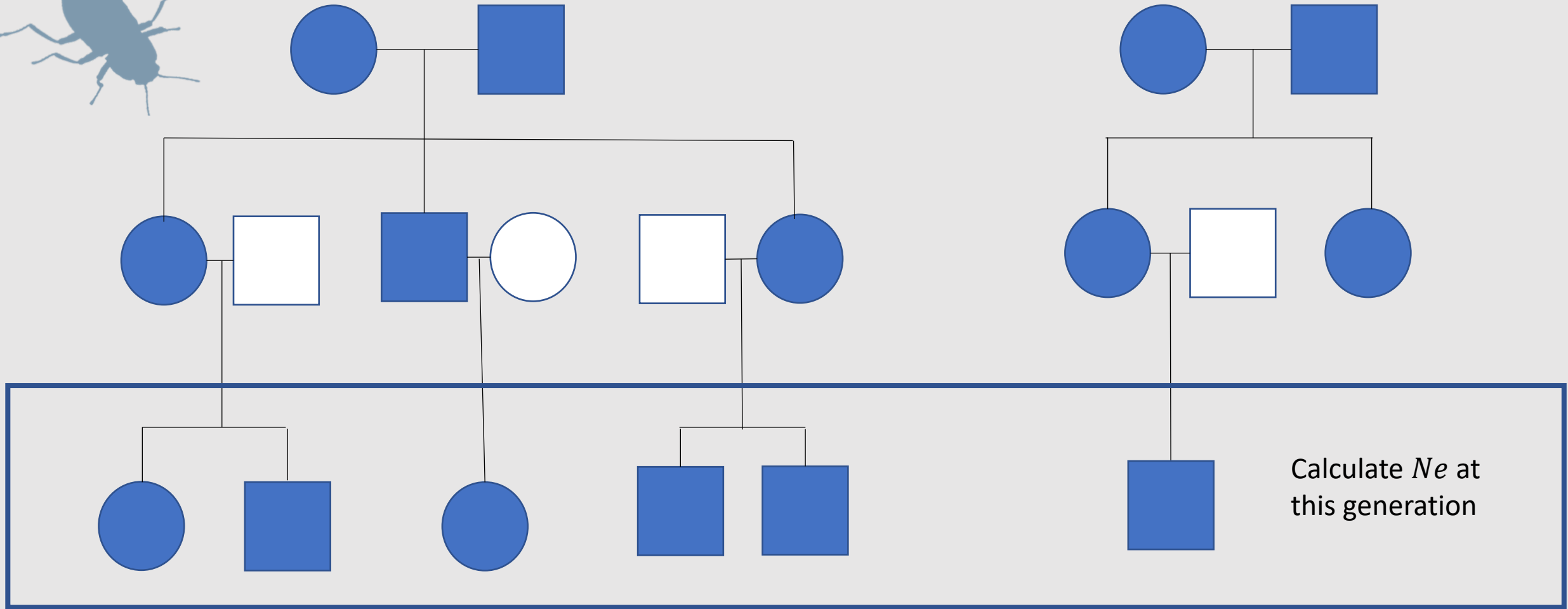
Estimate

Investigate

# Estimate Effective Population Size



Use sibship relationships to back calculate to  $N_m$  and  $N_f$



Assess

Estimate

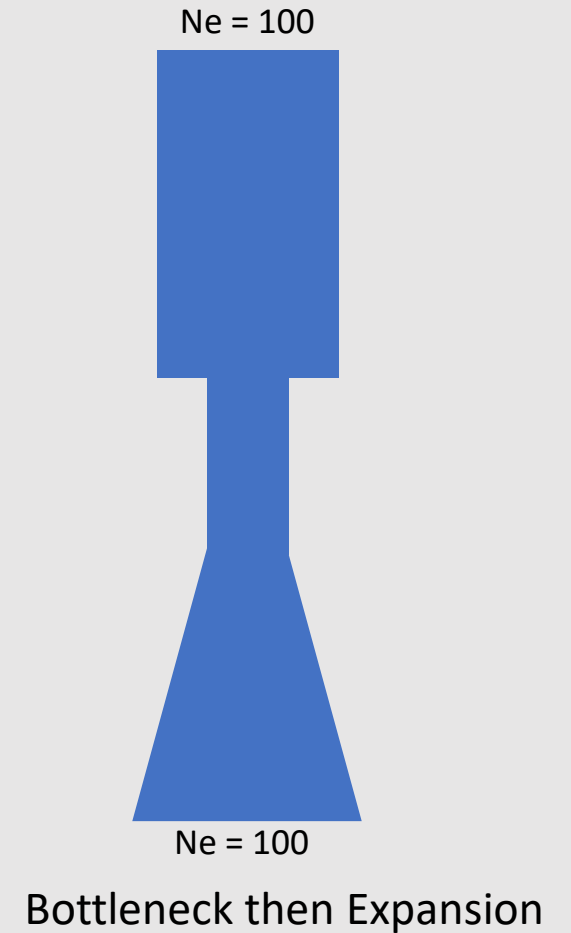
Investigate

# Investigate Changes in Population Size



Population size fluctuates over time and  $N_e$  can tell us how

Then  
Time  
Now



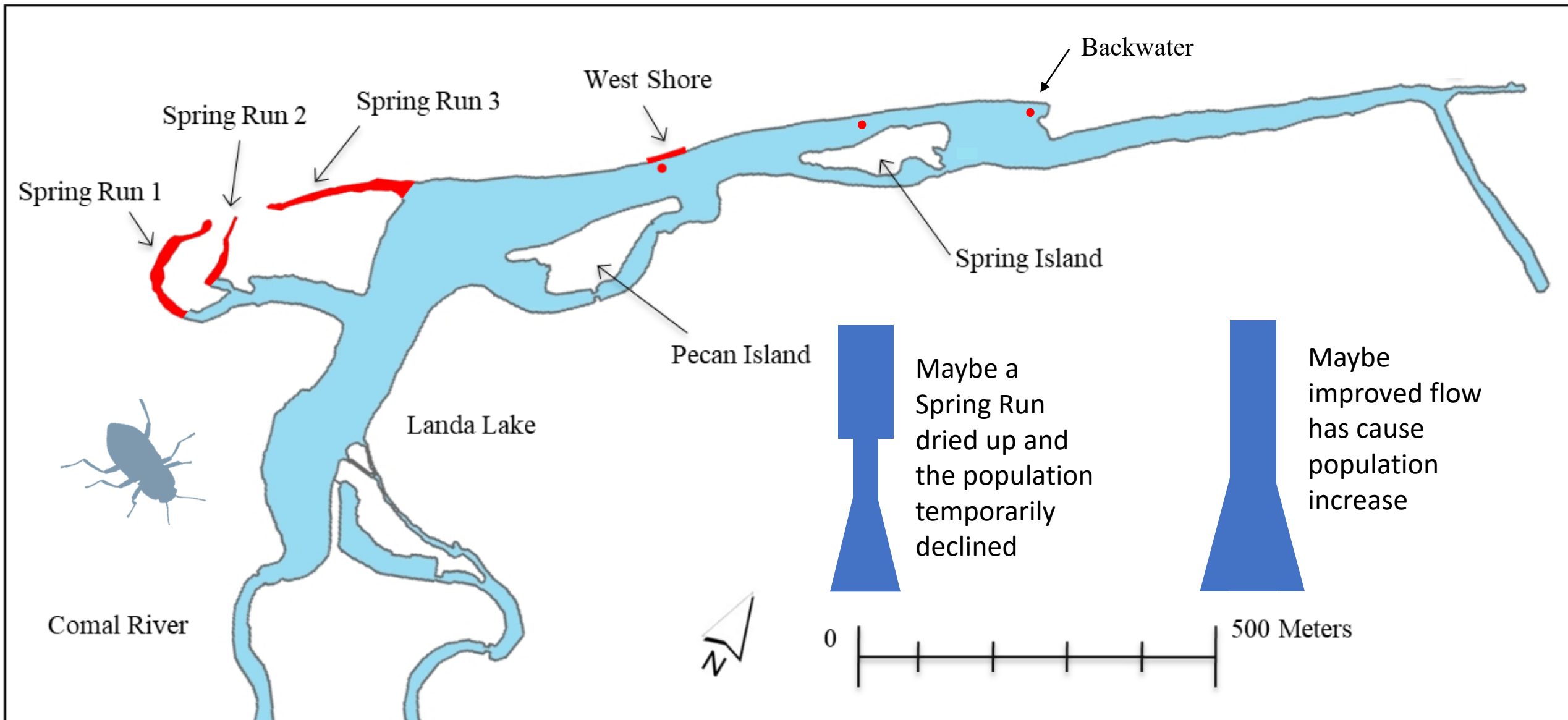
Assess

Estimate

Investigate



# Investigate Changes in Population Size



Assess

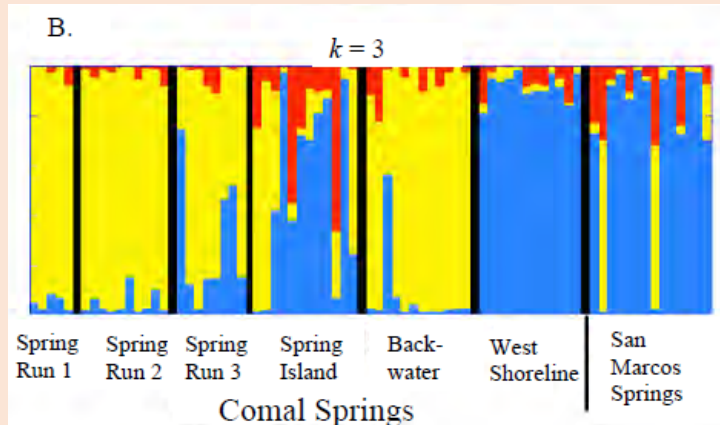
Estimate

Investigate

# Genetic Analysis of Comal Springs Riffle Beetle to:

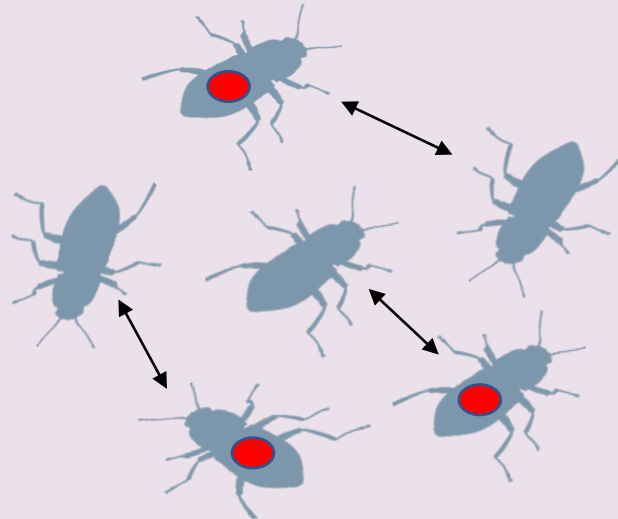
## Assess

Assess fine-scale diversity of the Comal Springs riffle beetle across Landa Lake



## Estimate

Estimate the effective population size ( $N_e$ ) (number of breeding adults)



## Investigate

Investigate historical changes in population size



# Benefits to the Refugia Include

Inform the minimum number of individuals needed in refugia to reflect wild population

Inform refugia collections

Better assess threats to the wild population

Inform long-term impact of reintroducing an assurance population

# Thank You For Your Attention And Support

---

## Acknowledgments

- Edwards Aquifer Authority
- Dr. Chad Furl
- Kristy Kollaus
- Will Colman
- Randy Gibson
- Dr. Ely Kosniki and BIO-WEST



# Questions?