Historical habitat barriers prevent ring-like genetic connectivity throughout the distribution of the threatened Alameda striped racer

Jonathan Q. Richmond*, Dustin A. Wood, Karen E. Swaim, Robert N. Fisher, Amy Vandergast

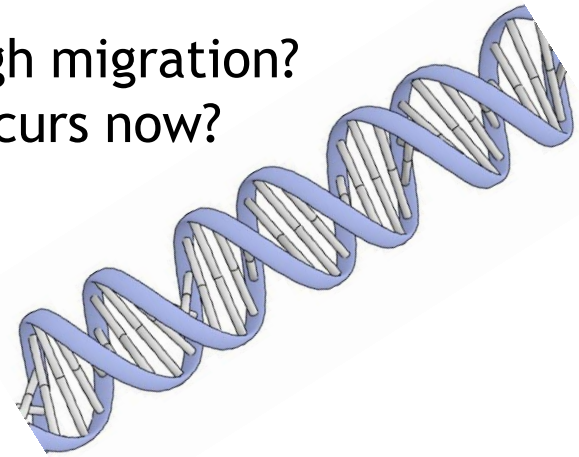
*U.S. Geological Survey, 4165 Spruance Rd. Suite 200,, San Diego CA 92101
jrichmond@usgs.gov, 619-225-6434

Photo by Chad M. Lane

Video by Nathan Smith

Why should we care about genetics when developing conservation management strategies?

- Genetic diversity
  - More diversity translates to higher potential for adaptation
  - Masks deleterious genetic variation that might exist at low frequency
- Gene flow
  - Maintains diversity across populations
  - Re-supplies diversity to localized areas that may have experienced losses
- Landscape genetic structure
  - How are gene pools structured across the landscape?
  - Where has there been mixing between gene pools, say through migration?
  - Where did gene exchange occur historically, but no longer occurs now?
- Relatedness
  - Within population measures (i.e. family groups)
  - Between or among sets of populations
  - Phylogenetic relatedness (across species)



Why should we be especially concerned about the genetics of threatened and endangered species?

- Most threatened and endangered species have that status because of restricted range size
  - As occupied land area gets smaller, so do population densities
  - Probability of mating with close relatives increases (i.e. inbreeding)
  - Genetic variation declines due to genetic drift, particularly if fragmentation prevents migration
    - In a closed population, genetic drift will always lead to a loss of genetic diversity over time
  - Reduced ability to adapt to a changing environment

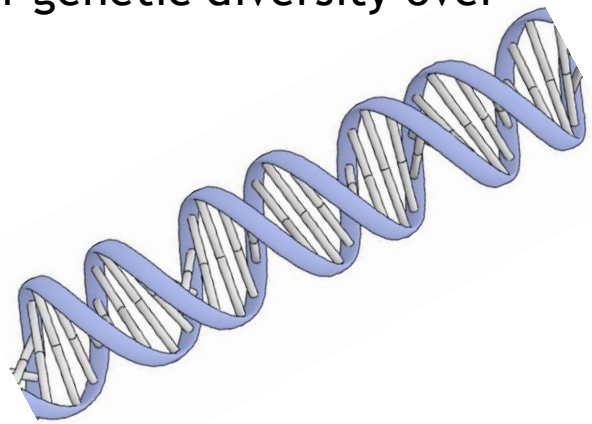


Photo by Chad M. Lane

Photo by Chad M. Lane

Photo by Jared Heald

Photo by Gary Nafis

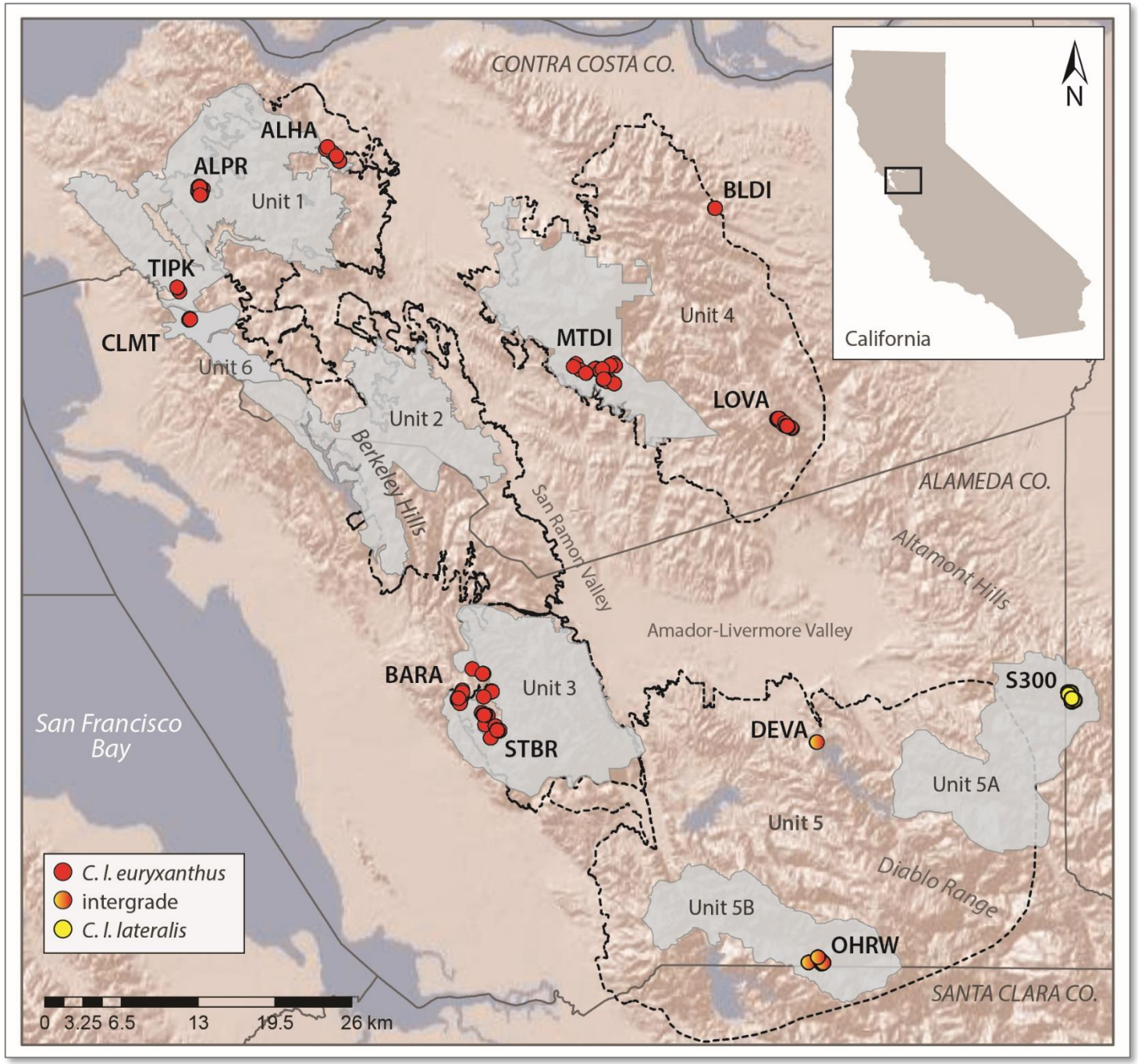
Genetics research to date:

- Management units for the Alameda whipsnake (*M. lateralis euryxanthus*)
- Rangewide phylogeography for the California striped racer (*M. lateralis*)

Photo by Jim Bennet

Sampling for Alameda whipsnake genetics

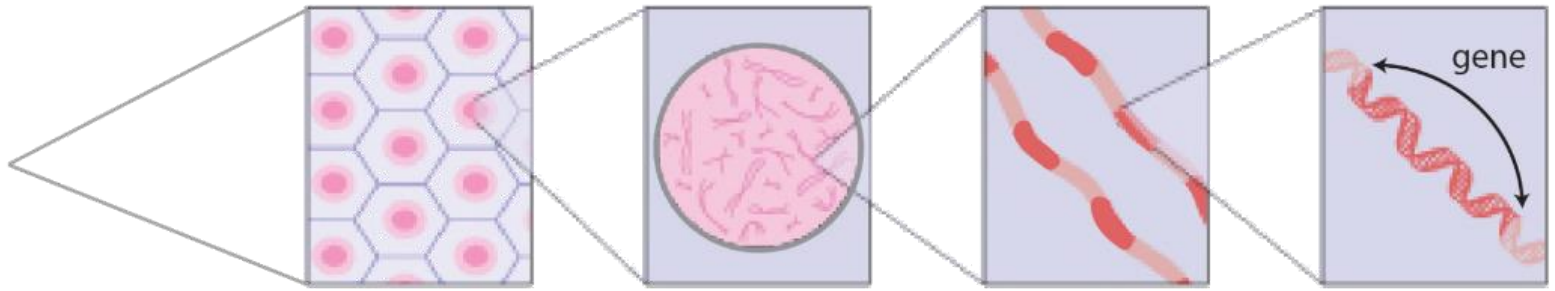
- Bullets are individual sampling points
- Four letter site codes; 12 sampling localities
- Colors denote different subspecies and intergrade forms



Main study questions:

- How well do the different management units capture the genetic diversity within the Alameda whipsnake?
- To what degree are populations within the units genetically distinctive or admixed?
- What are the patterns of gene flow across the landscape?

Photo by Demitri Penuelas



Constituent cells

Nucleus containing two sets of 23 chromosomes

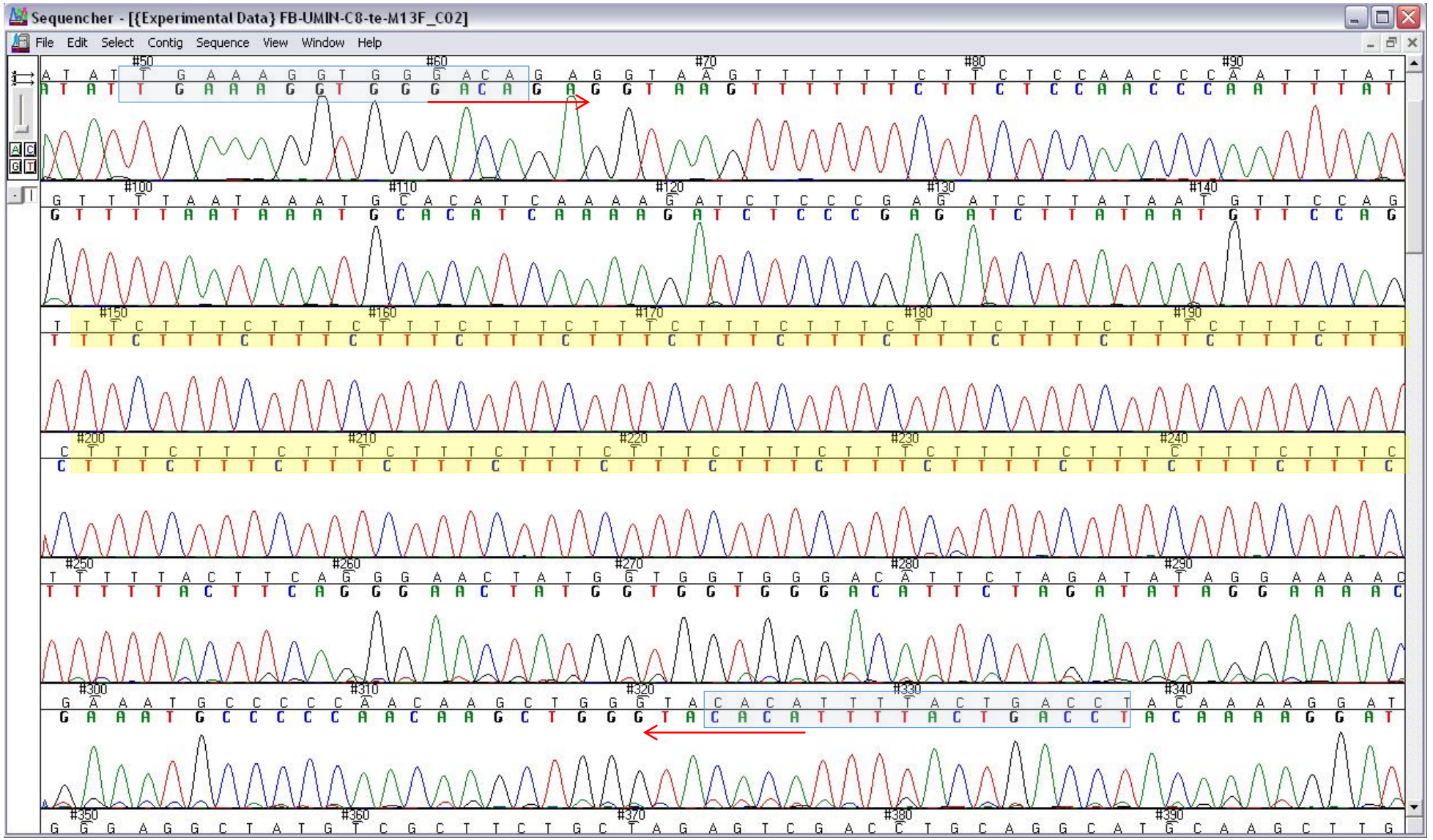
Enlargement of part of a pair of chromosomes

One gene, a functional region of chromosomal DNA

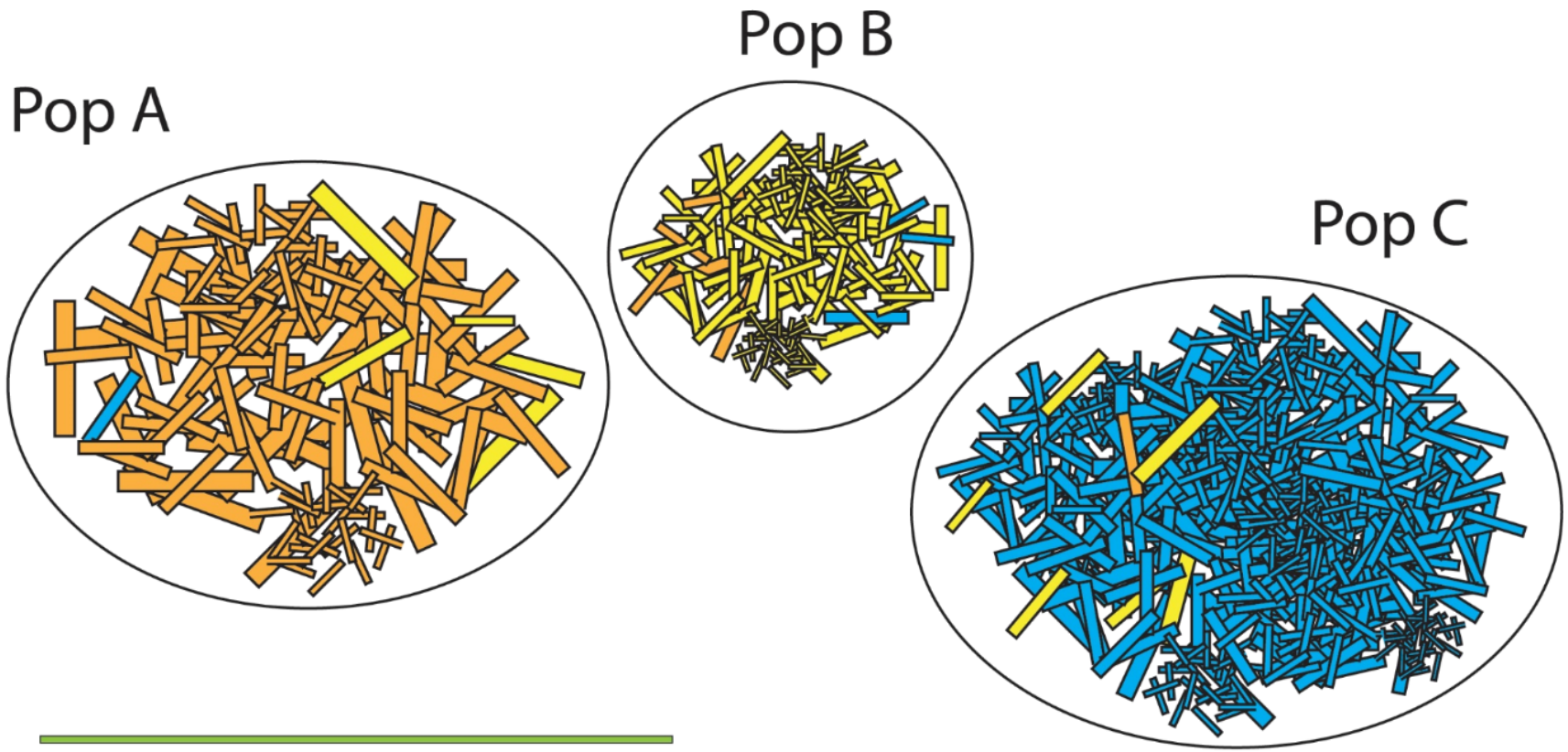
The types of DNA data we collect:

- Single nucleotide polymorphisms (SNPs)
- Microsatellites
- DNA sequence data

- Non-coding nuclear gene regions
- Bi-parently inherited
- Rapid rates of mutation
- Used to study ‘shallow history’

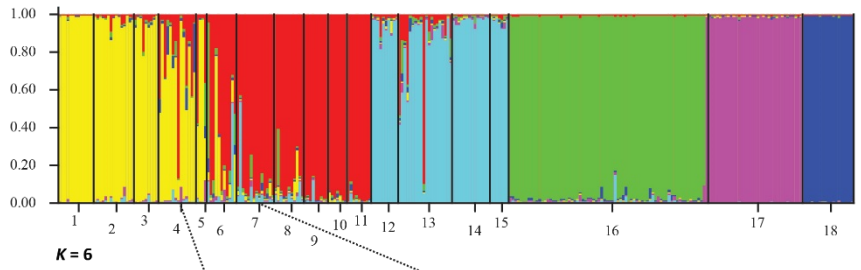


$TTTC_{25}$ tetramer motif

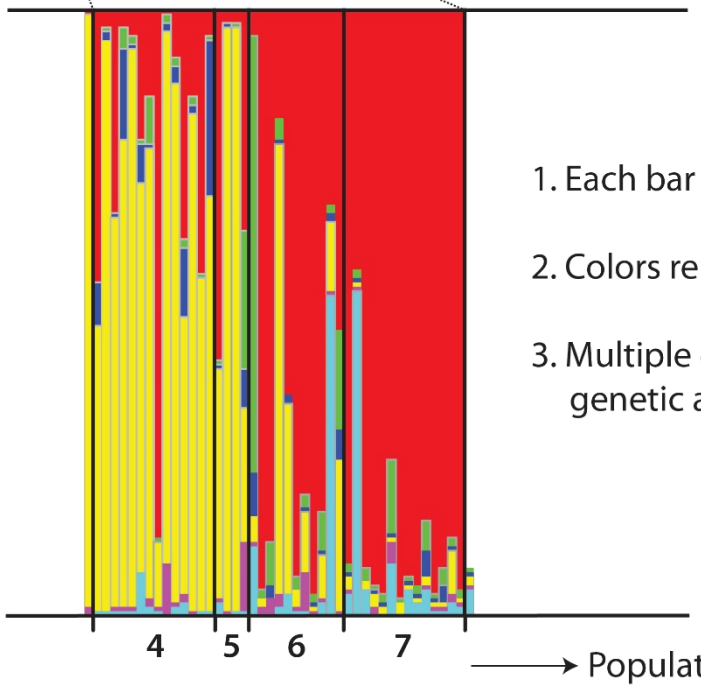

Pop A
Pop B
Pop C
= CTACTACTACTACTA
= CTACTACTACTA
= CTACTACTA

Data analysis: cluster assignments

- Probabilistic assignment of individuals to different clusters/groups using Bayesian statistical methods
- Estimate the degree to which clusters are genetically distinctive or admixed



Assignment plot

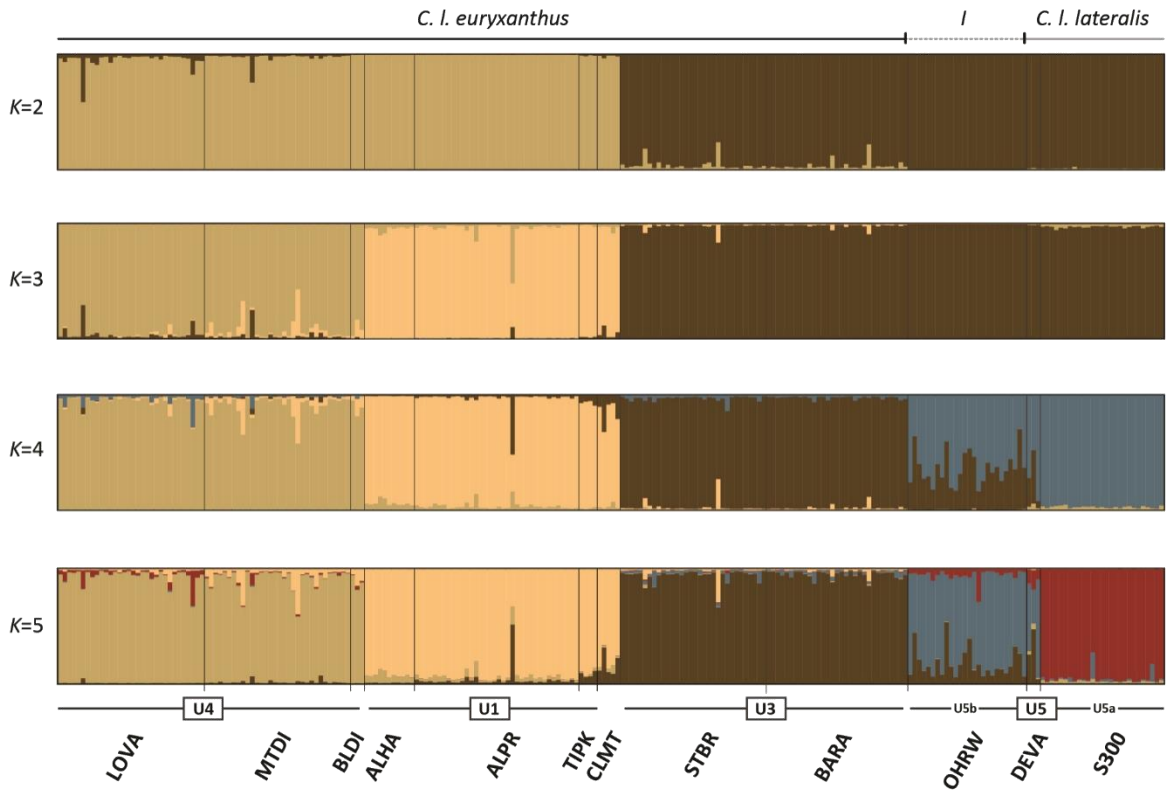


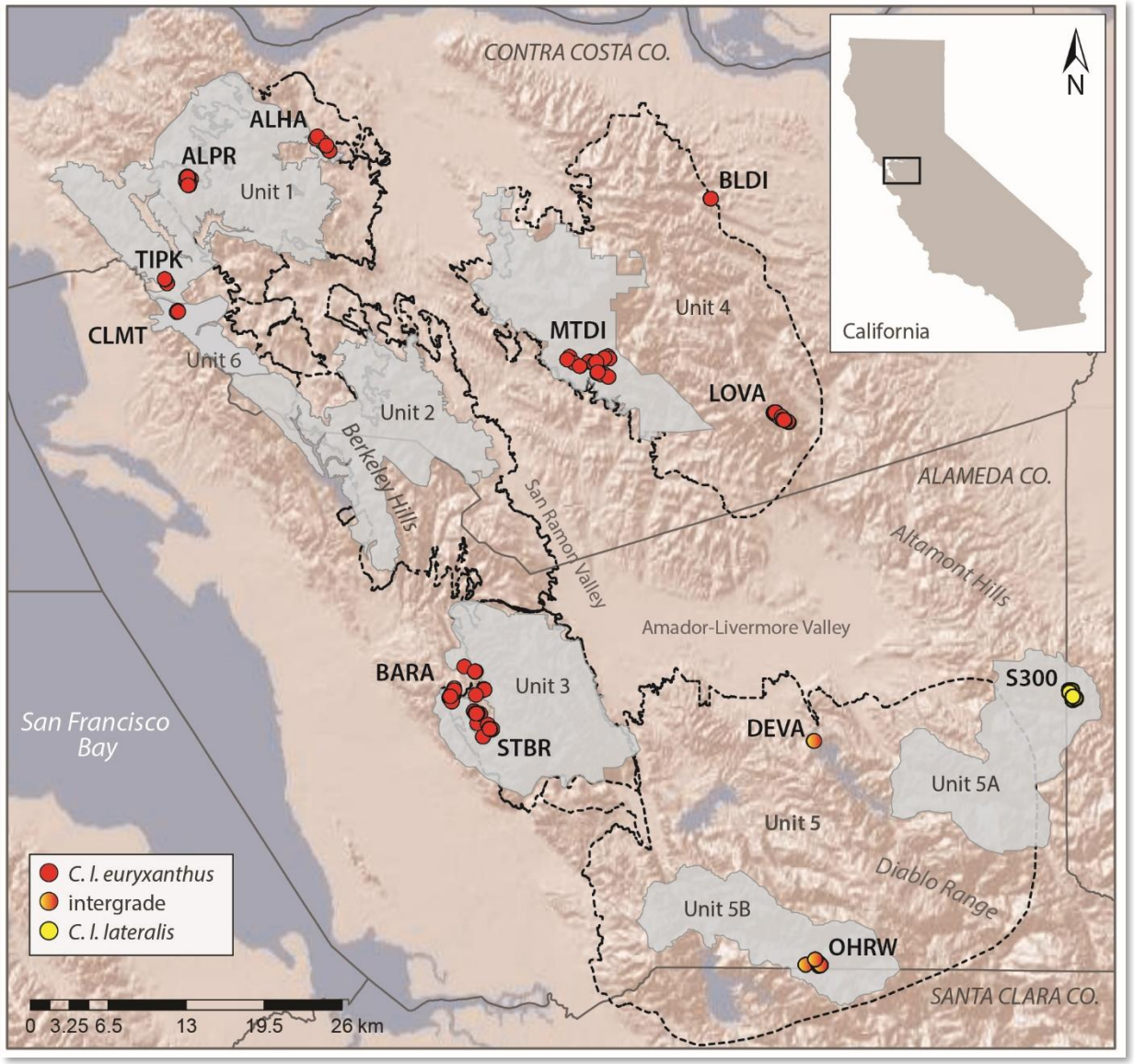
1. Each bar in the histogram is an individual.

2. Colors represent genetic groups.

3. Multiple colors within a bar represent genetic admixture.

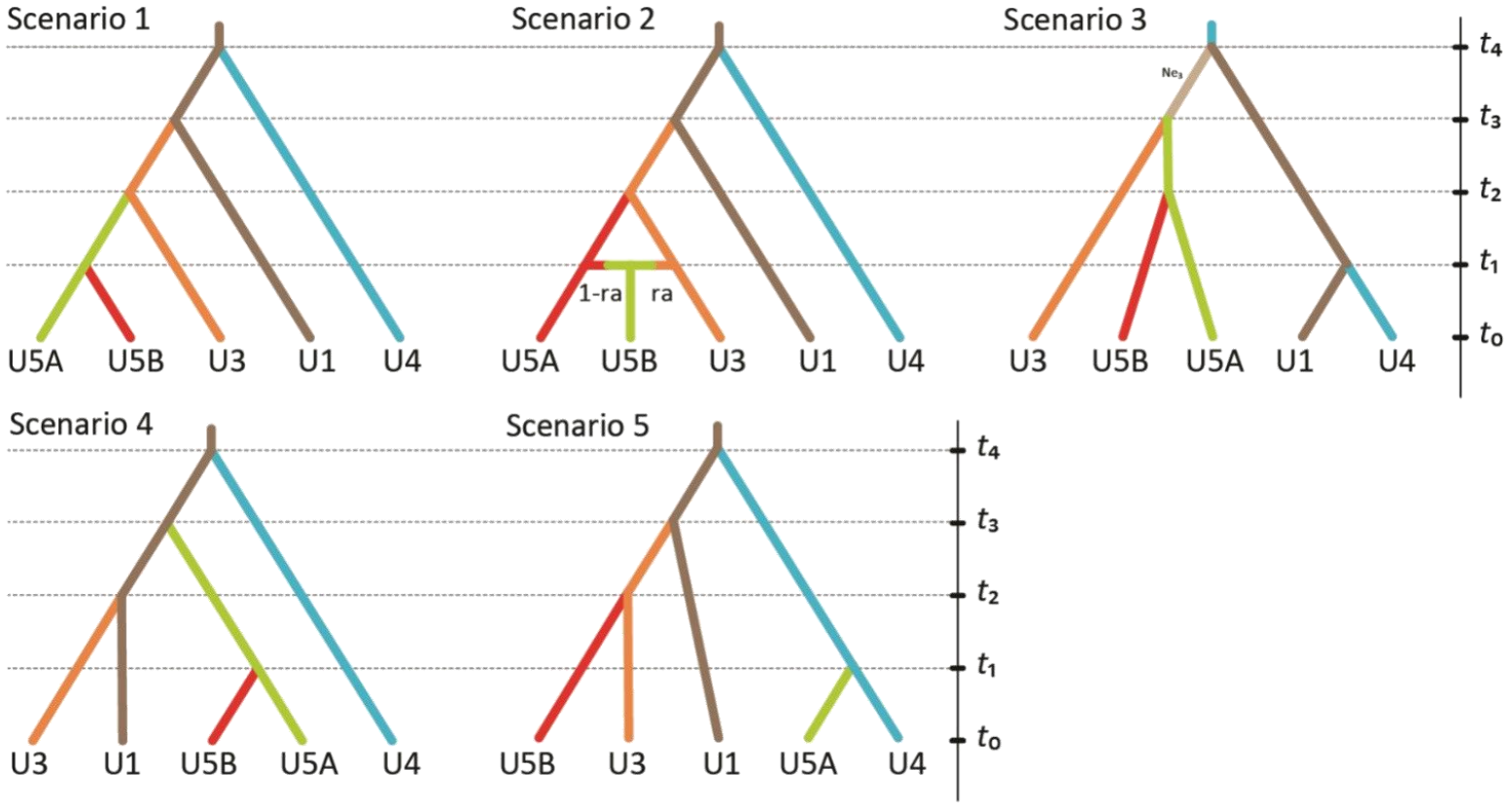
Results: cluster assignments



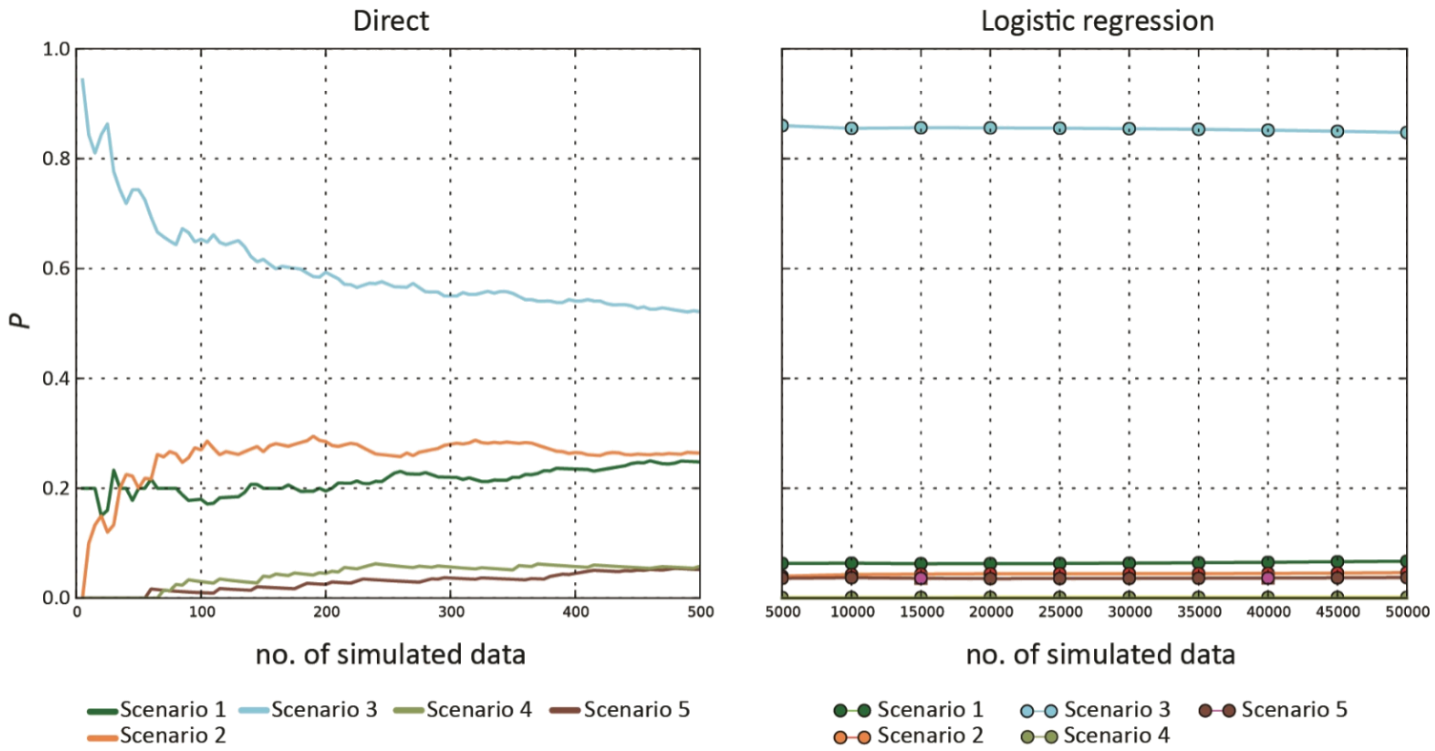


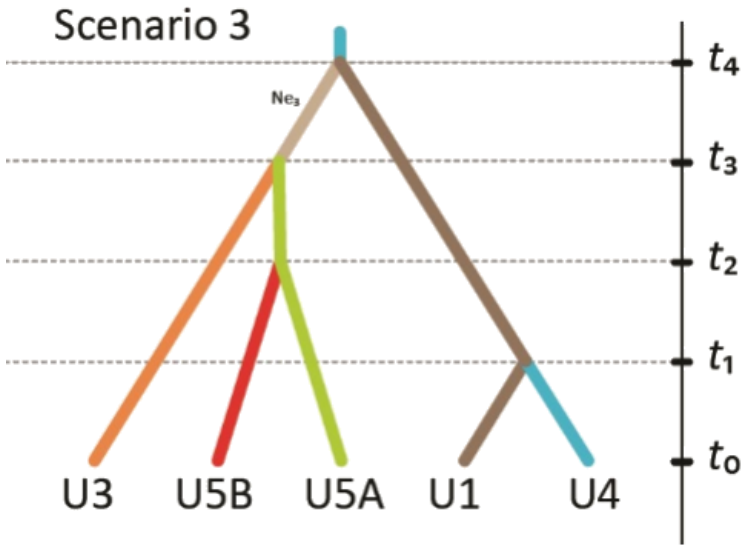
Data analysis: historical demographic modeling

- Develop a range of plausible historical demographic scenarios
- Simulate datasets based on those scenarios
- Statistically measure the fit of the observed data to the simulated data generated from the different scenarios



Results: historical demographic modeling





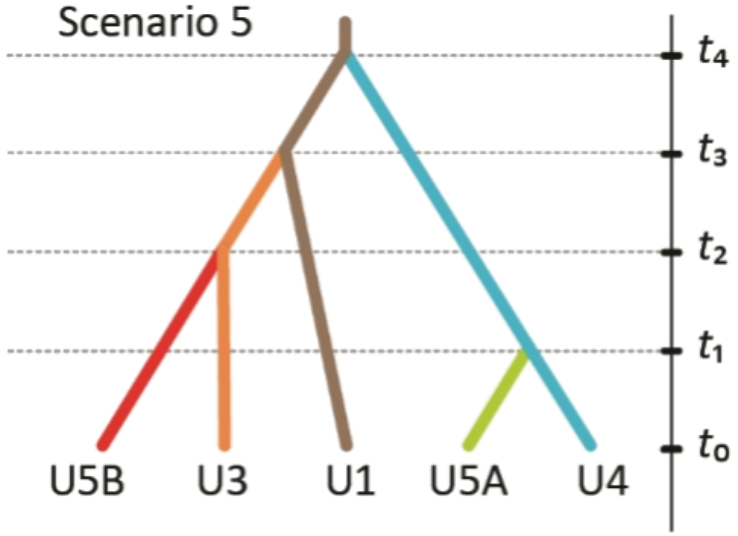
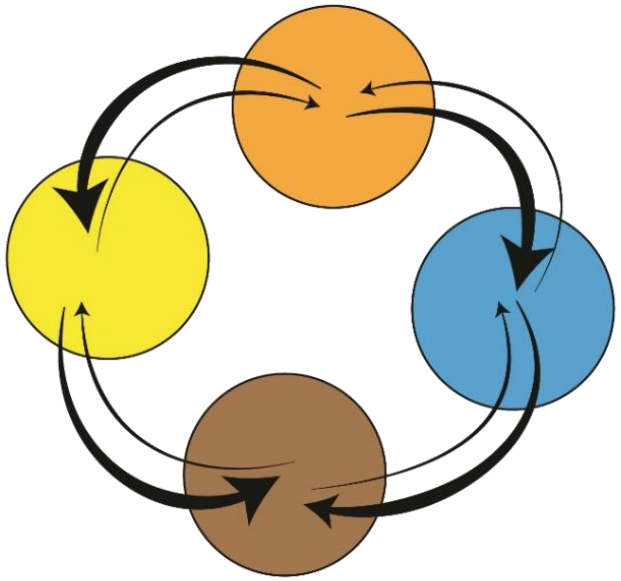


Photo by MC Rider

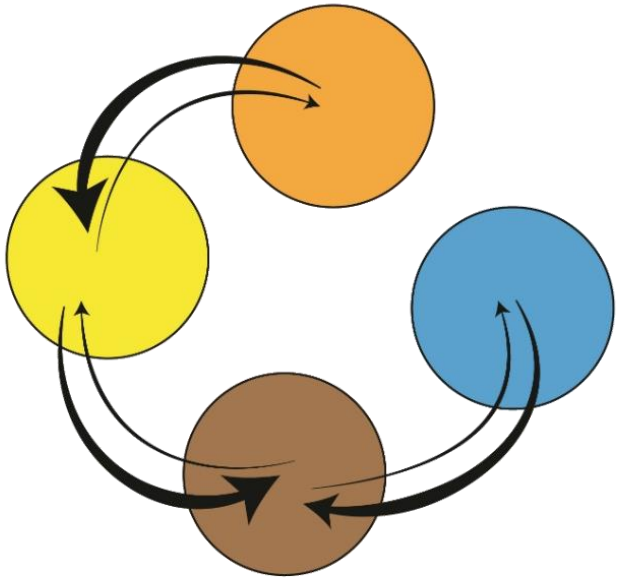
Data analysis: Inferring patterns of historic gene flow

Migrate N (Beerli & Palczewski. 2010)

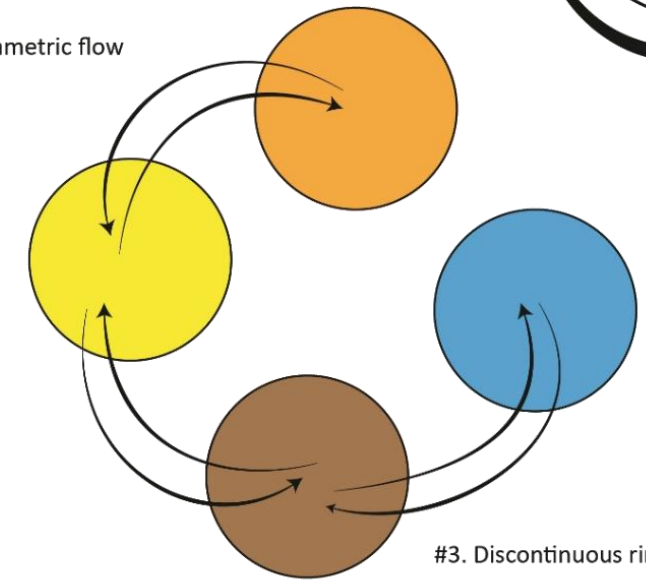
- Compare the fit of models that describe different patterns of population connectivity and differences in rates of gene flow among populations
- Estimates effective population sizes and migration rates based on coalescent theory



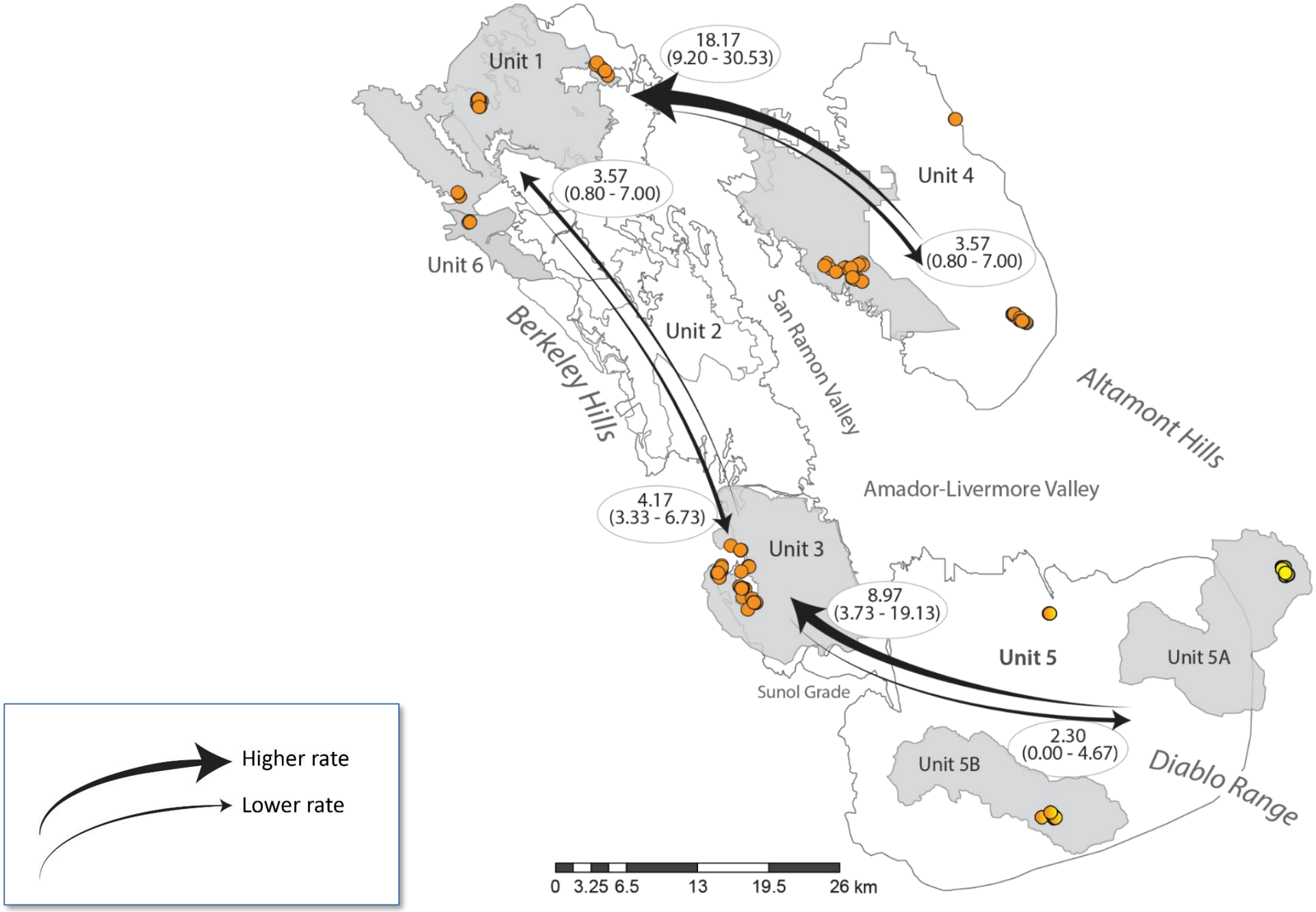
#1. Continuous ring, asymmetric flow

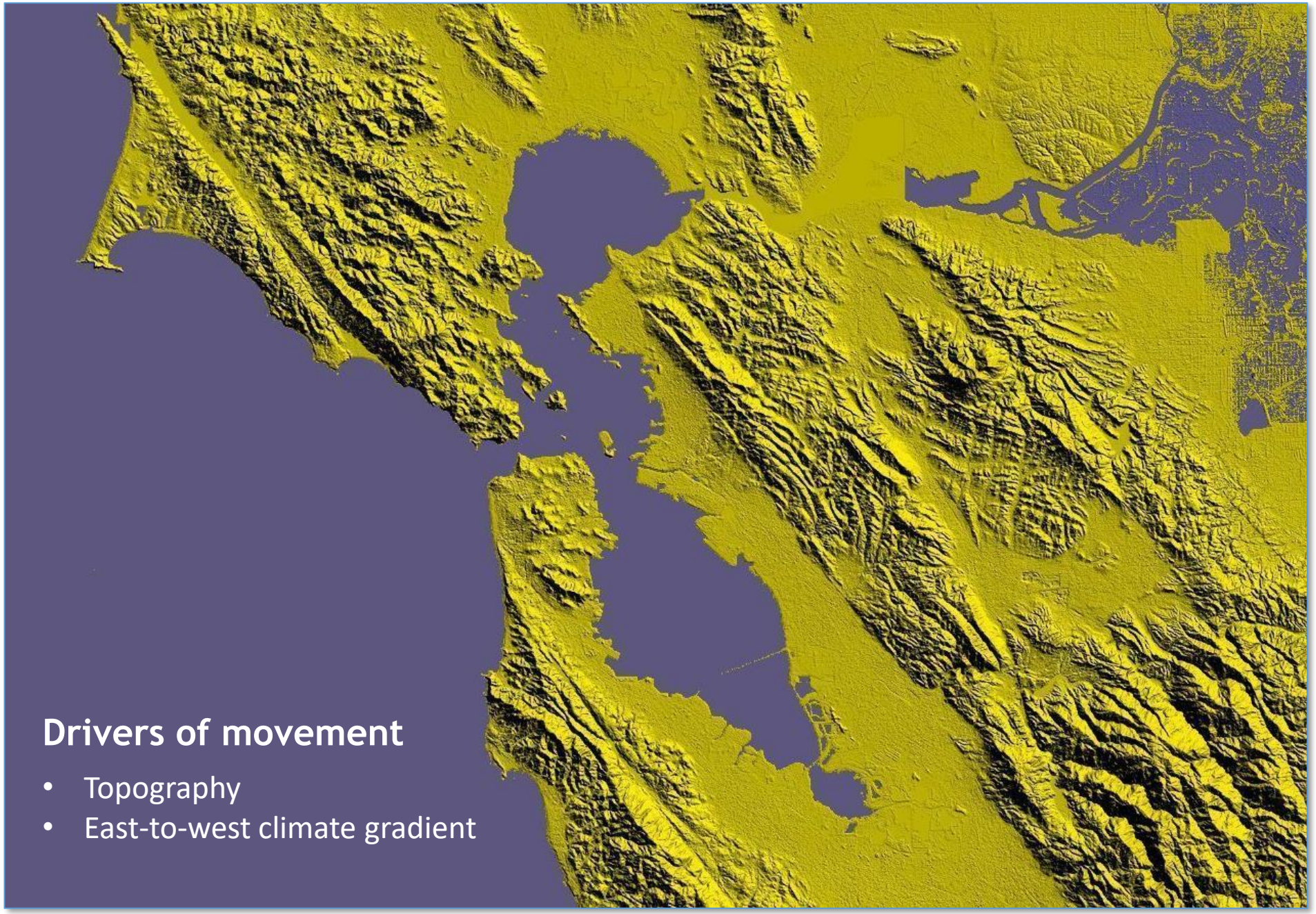


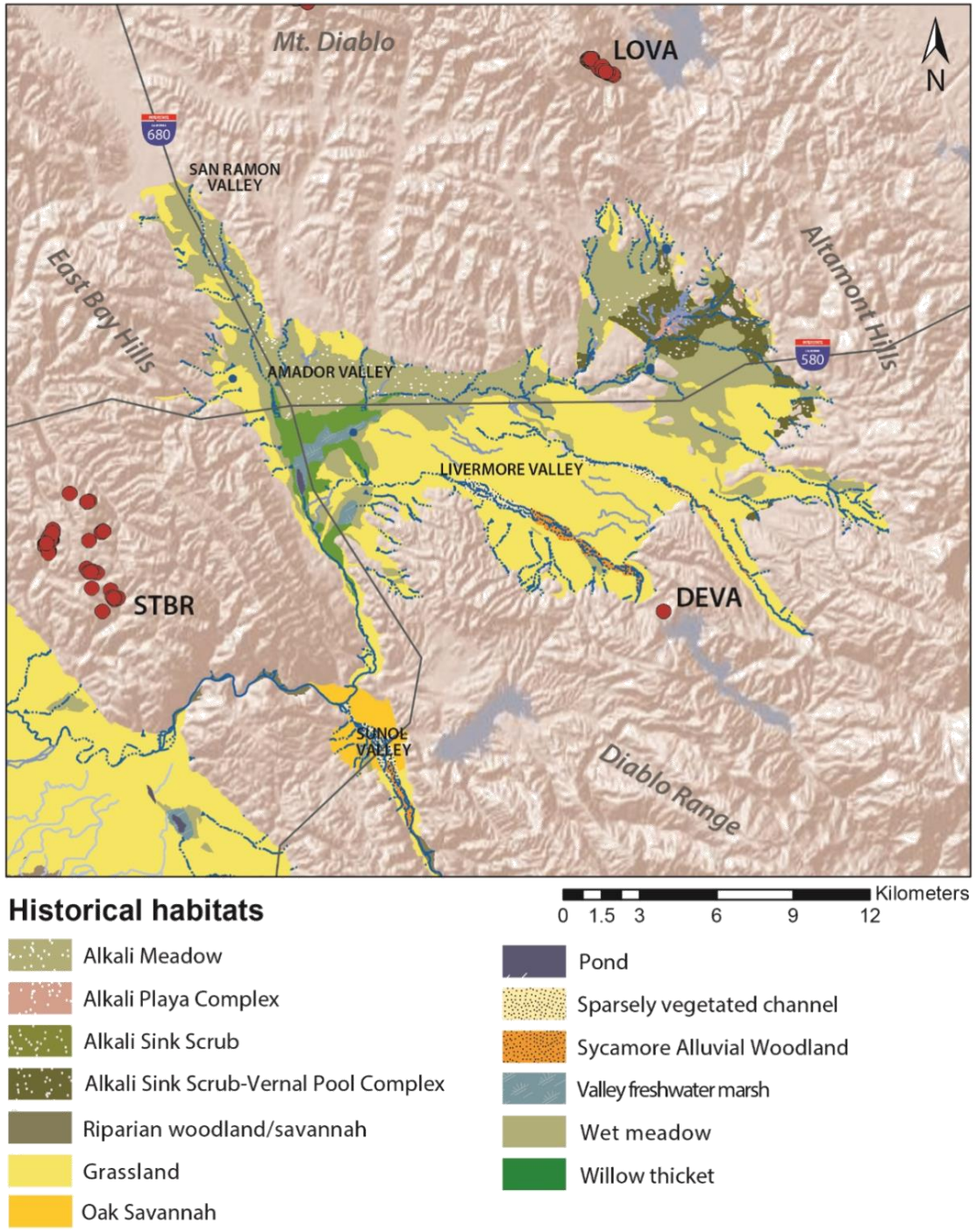
#2. Discontinuous ring, asymmetric flow



#3. Discontinuous ring, symmetric flow


Unit 1
18.17
(9.20 - 30.53)
Unit 4
3.57
(0.80 - 7.00)
3.57
(0.80 - 7.00)
Unit 6
Unit 2
Berkeley Hills
San Ramon Valley
Altamont Hills
Amador-Livermore Valley
4.17
(3.33 - 6.73)
Unit 3
8.97
(3.73 - 19.13)
Unit 5
Unit 5A
Sunol Grade
2.30
(0.00 - 4.67)
Unit 5B
Diablo Range
Higher rate
Lower rate
0 3.25 6.5 13 19.5 26 km


Drivers of movement
• Topography
• East-to-west climate gradient



Stanford et al. 2013. Alameda Creek Watershed Historical Ecology Study, SFEI Publication 679. San Francisco Estuary Institute, USA.

Photo by Richard Porter

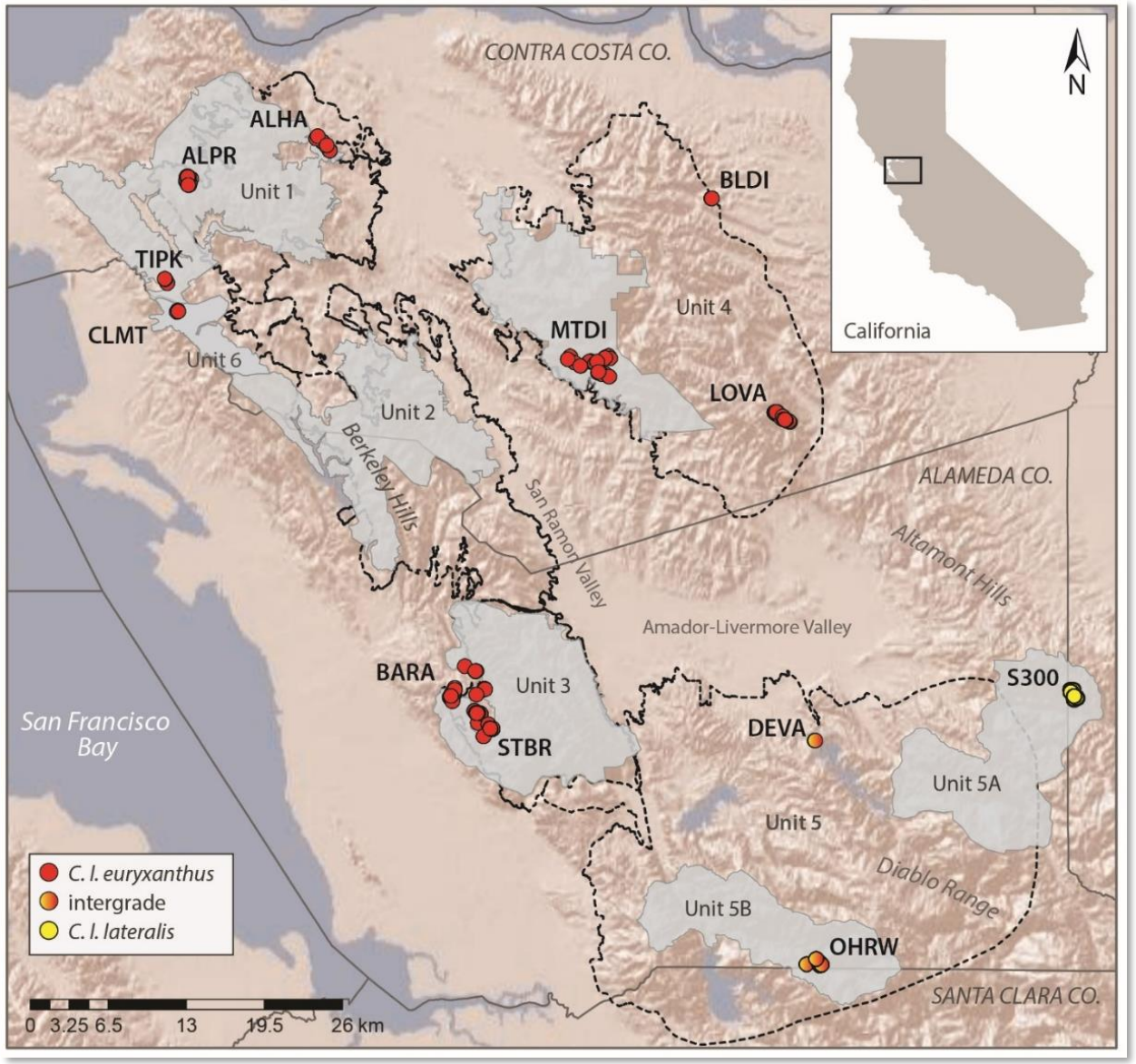
What did we learn?

- Snakes in the different management units are largely genetically distinctive, with the exception of the western portion of Unit 5
- Greater differentiation in the south/southeastern part of the study area
- Historic demography is consistent with a ring-like pattern of expansion, but without ring closure across the Altamont Hills
- Gene flow is directionally biased, with higher movement rates towards the more mesic, western parts of the study area

Photo by Mike Pingleton

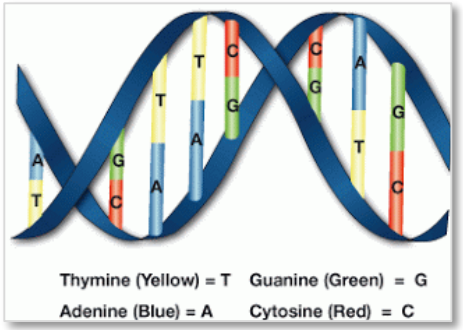
What should we do next?

- Fill in sampling gaps, particularly in Unit 2
- Increase sample sizes for certain areas
- Sample through and beyond the subspecies boundary in Santa Clara County
- Collect genomic data (i.e. restriction site-associated DNA sequencing [RADseq] to identify single nucleotide polymorphisms, or SNPs)
- Identify ‘outlier’ genetic markers that potentially distinguish *M. l. euryxanthus* from *M. l. lateralis*



Acknowledgments

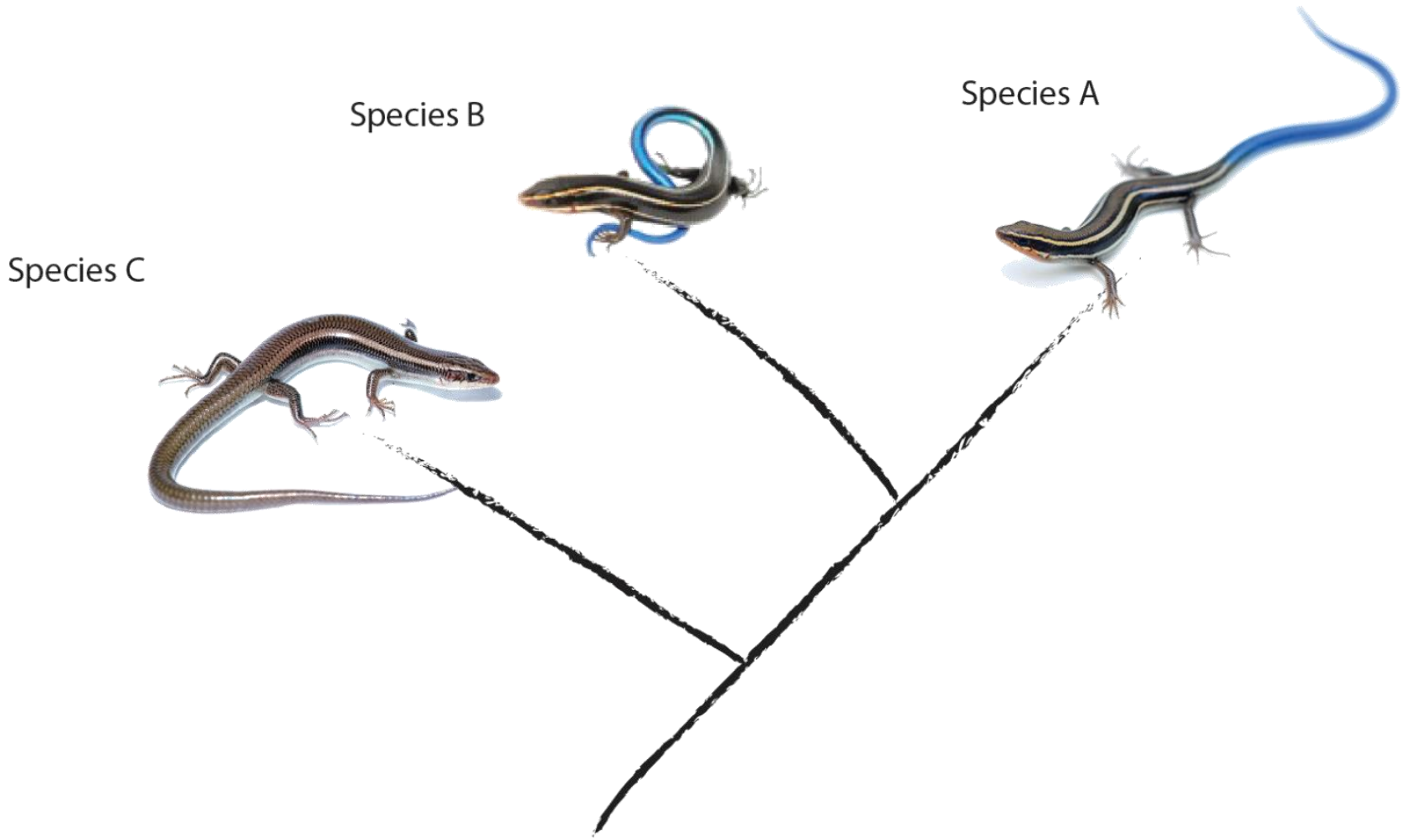
- **Funding**: U.S. Fish and Wildlife Service (USFWS) Science Support Program, the Central Valley Habitat Restoration Program (USFWS), and the U.S. Geological Survey Ecosystems Mission Area.
- We thank the East Bay Regional Parks District for land access, permission to conduct survey work, and general support of this research.
- The project benefited greatly from discussions, comments and insight by B. Solvesky (USFWS) and R. Stoelting (Swaim Biological Inc.).
- R. Stoelting, B. Solvesky, E. Britt, A. Murphy, and four anonymous reviewers provided helpful comments on earlier versions of the manuscript.

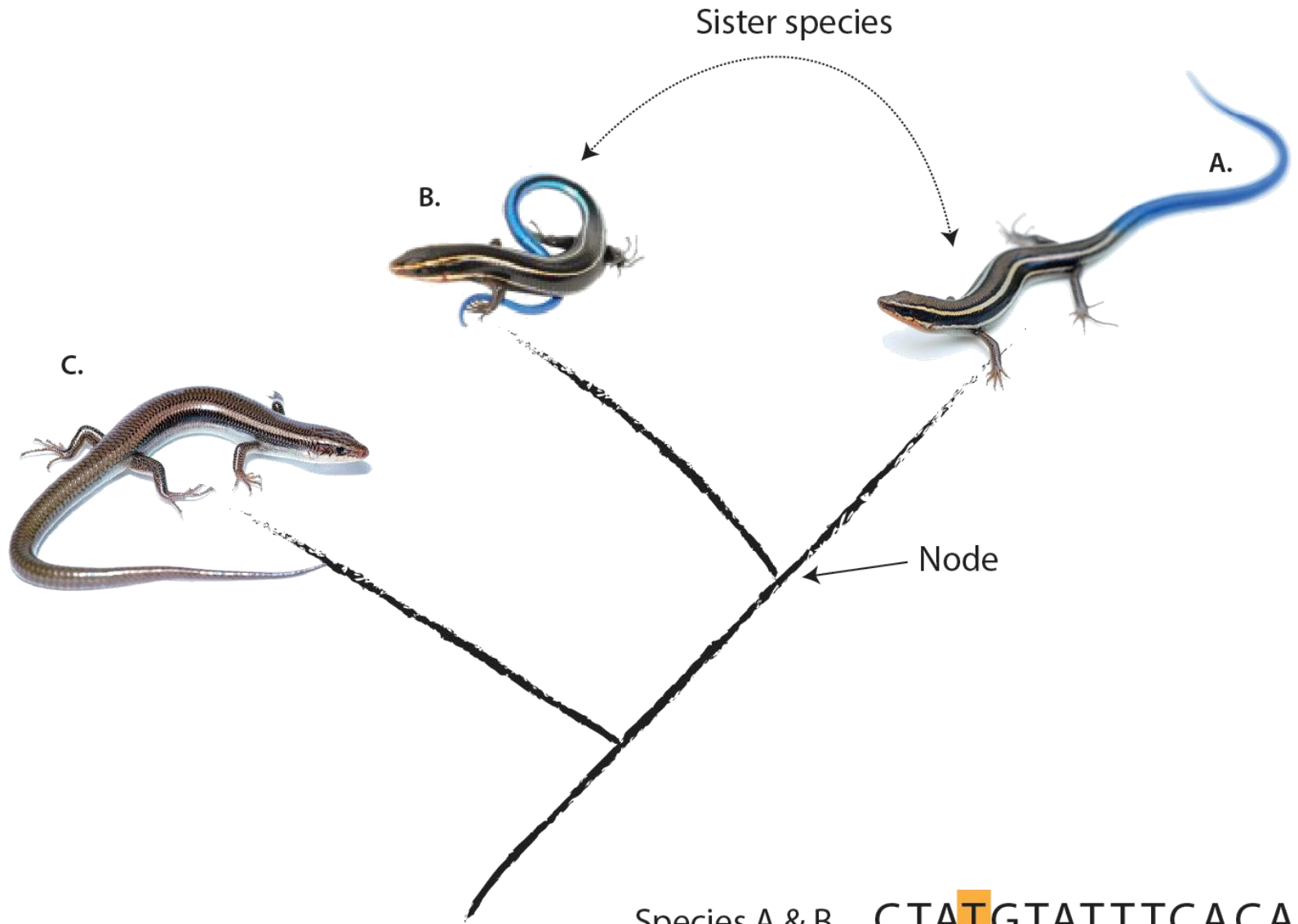


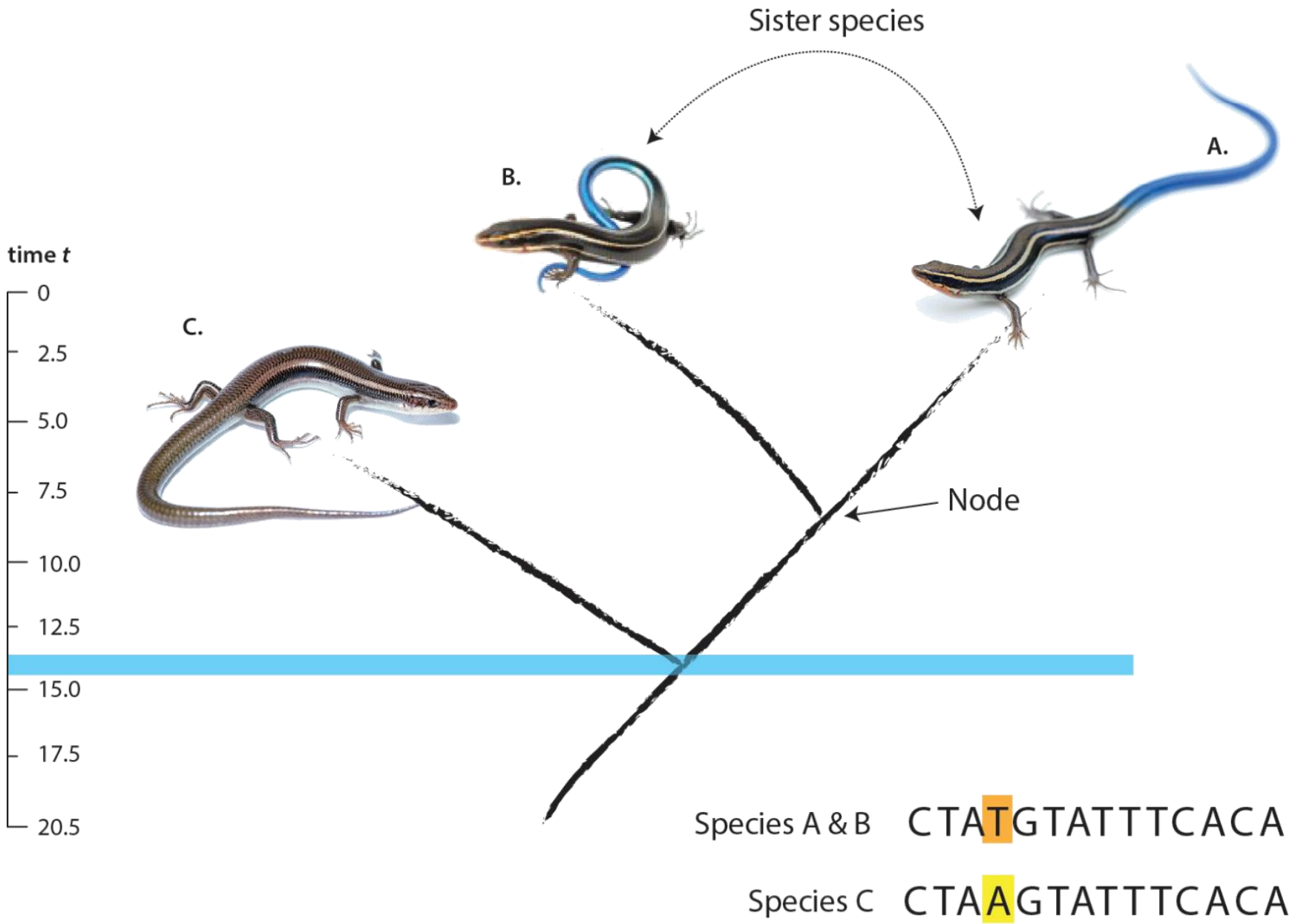
Example of a DNA sequence alignment

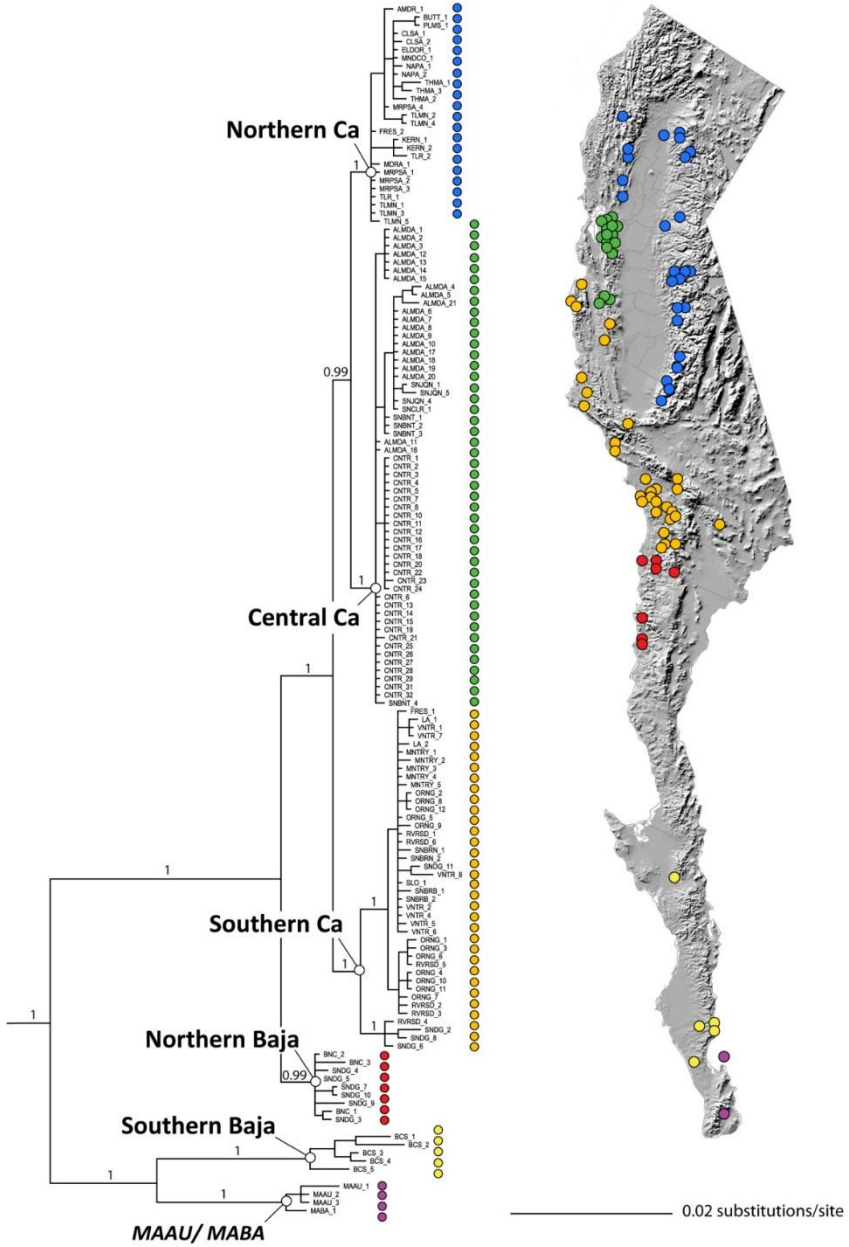
*Emoia*_1 AACCTATTATCTCATTCCCACATCTGCCACACTAAGTATAATAACACCTCA
*Emoia*_2 AACCTATTATCTCATTCCCACATCTGCCACACTAAGTATAATAACACCTCA
*Emoia*_3 AACCTATTATCTCATTCCTACAACTGCCACACTAAGTATAATAACACCTCA
*Emoia*_4 AACCTATTATCTCATTCCTACAACTGCCACACTAAGTATAATAACACCTTA
*Emoia*_5 AACCTATTATTATATTCCCACAACTGCCATACTGAGTATAATAACACTTCA
*Emoia*_6 AACCTATTATTATATTCCCACAACTGCCATACTGAGTATAATAACACTTCA
*Emoia*_7 AACCTACTATTTCATTCCCATAACCGCCACACTGAGTATAATAACACCTCA
*Emoia*_8 AACCTACTATTTCATTCCCACAACCGCCACACTGAGTATAATAACACCTCA
*Emoia*_9 AACCGATTATTATATCCCCATAACTGCCACACTGAGTATAATAACACTTCA
*Emoia*_10 AACCGATTATTATATCCCCATAACTGCCACACTGAGTATAATAACACTTCA

The DNA sequences of closely related individuals are more similar than distantly related individuals.


Species A
Species B
Species C


Sister species
A.
B.
C.
Node
Species A & B CTATGTATTTCACA
Species C CTAAGTATTTCACA


Sister species
A.
B.
C.
time t
0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.5
Node
Species A & B CTATGTATTTCACA
Species C CTAAGTATTTCACA



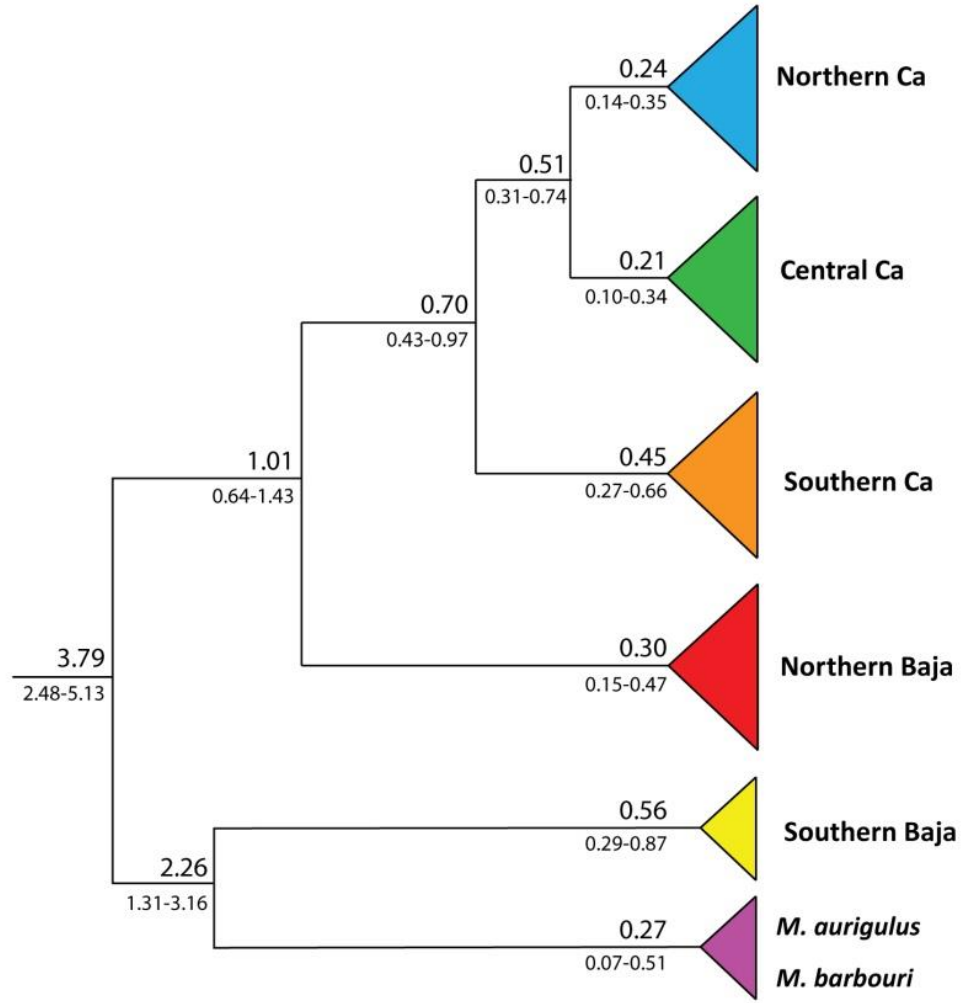
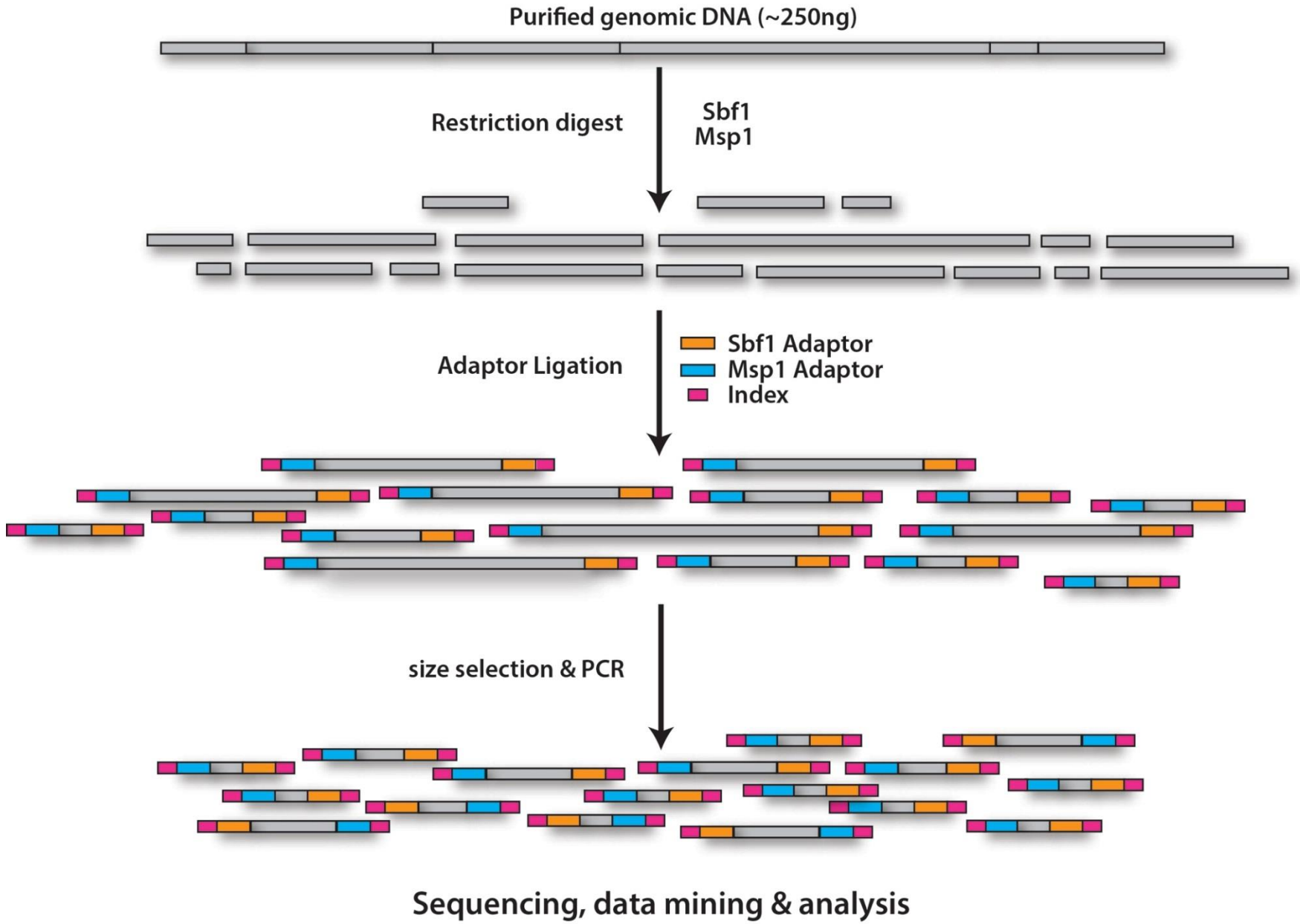
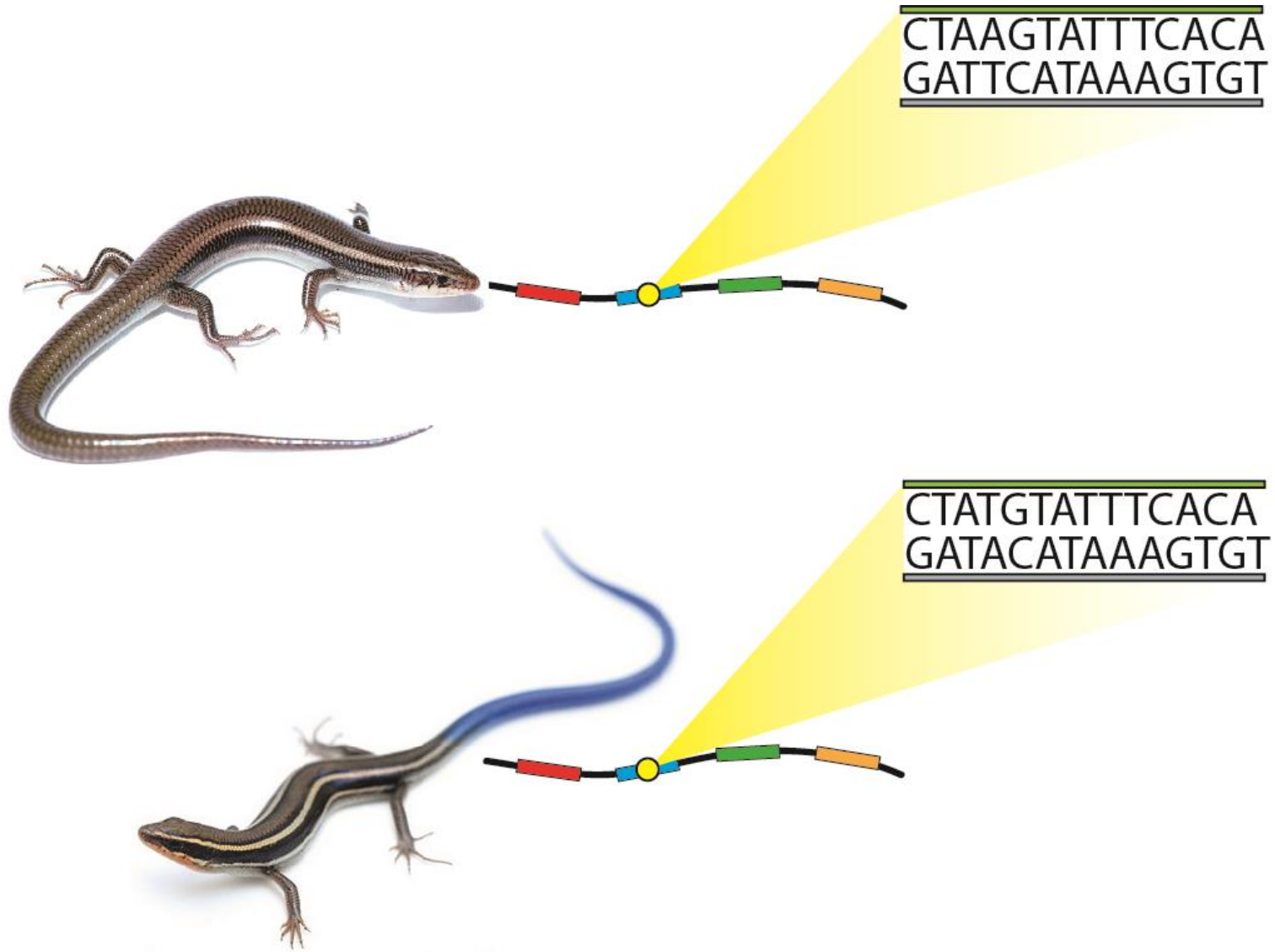


Figure 4. Bayesian relaxed-clock estimates of divergence times for the major mtDNA lineages of *M. lateralis*. Numbers above the branches represent time (millions of years before present); numbers below the branches are the 95% highest posterior densities, and represent the uncertainty of the age estimates. Color codes for the major lineages match those in Figure 3.

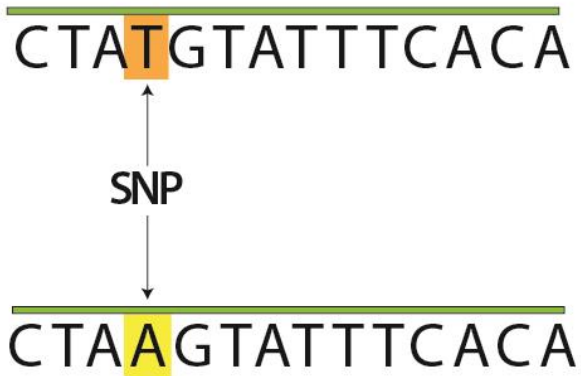
ddRADseq


Purified genomic DNA (~250ng)
Restriction digest
Sbf1
Msp1
Adaptor Ligation
Sbf1 Adaptor
Msp1 Adaptor
Index
size selection & PCR
Sequencing, data mining & analysis

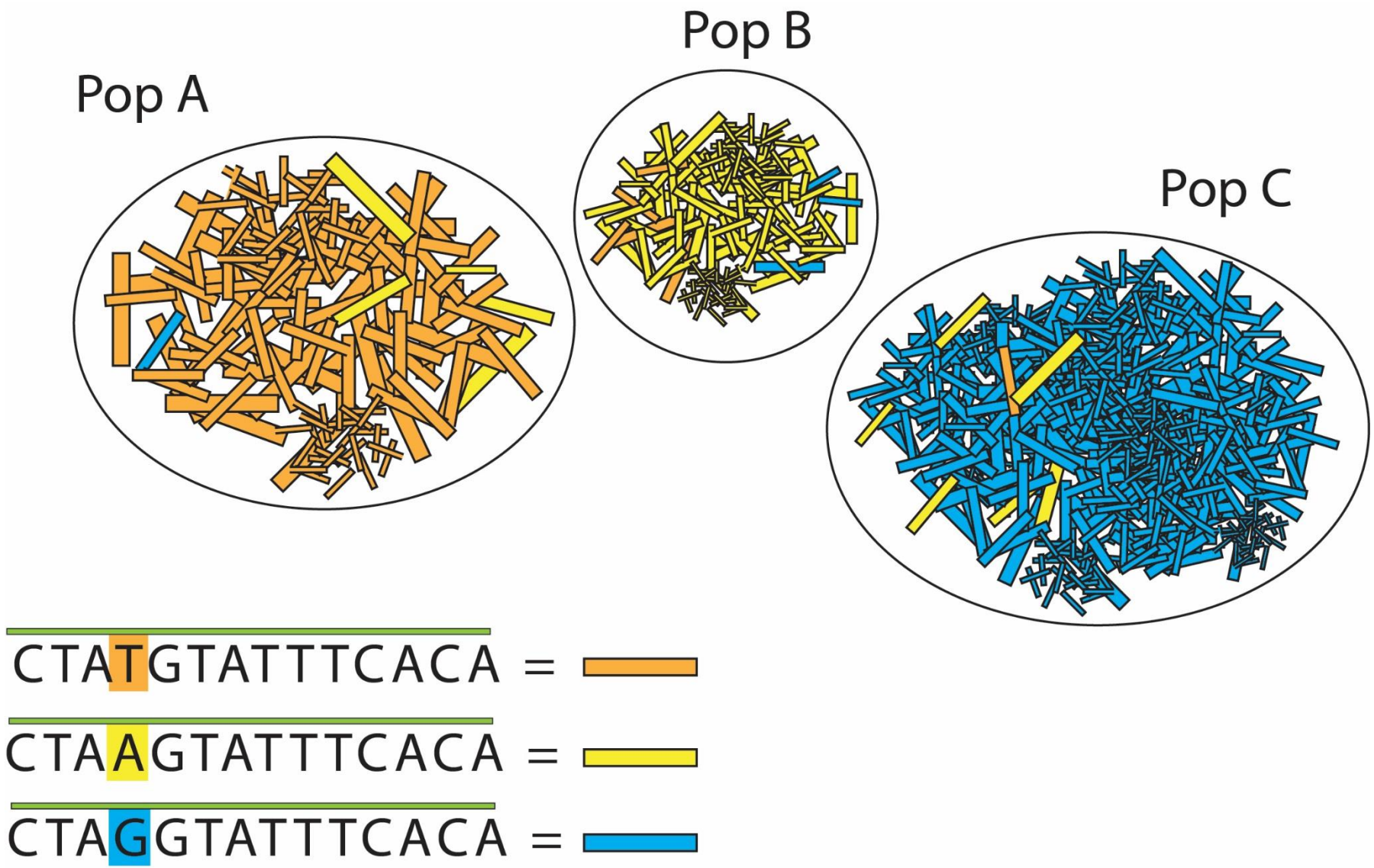


Here we have two short, double stranded DNA sequences taken from the same region of the genome in two different skinks.

Keeping it simple, we'll just show one strand....



The sequences are almost identical except for one nucleotide position.


Pop A
Pop B
Pop C
CTATGTATTTCACA =
CTAAGTATTTCACA =
CTAGGTATTTCACA =