Genetics for Monitoring and Management Workshop

December 16, 2013

USGS San Diego Field Station

Organizers:

Amy Vandergast, US Geological Survey

Andrew Bohonak, San Diego State University

Genetics for Monitoring and Management Workshop

Genetics 101: State of the Science
(Andrew Bohonak)

Estimating Diversity and Effective Size within Populations
(Jonathan Richmond)

BREAK

Estimating Gene Flow and Movement Among Populations
(Amy Vandergast)

Informing Action: Mountain Yellow Legged Frogs
(Robert Fisher)

LUNCH & Small Group Discussions

Managing for Genetic Variation: When, Why and How?
(Ollie Ryder)

Special Consideration for Plants
(Norman Ellstrand)

BREAK

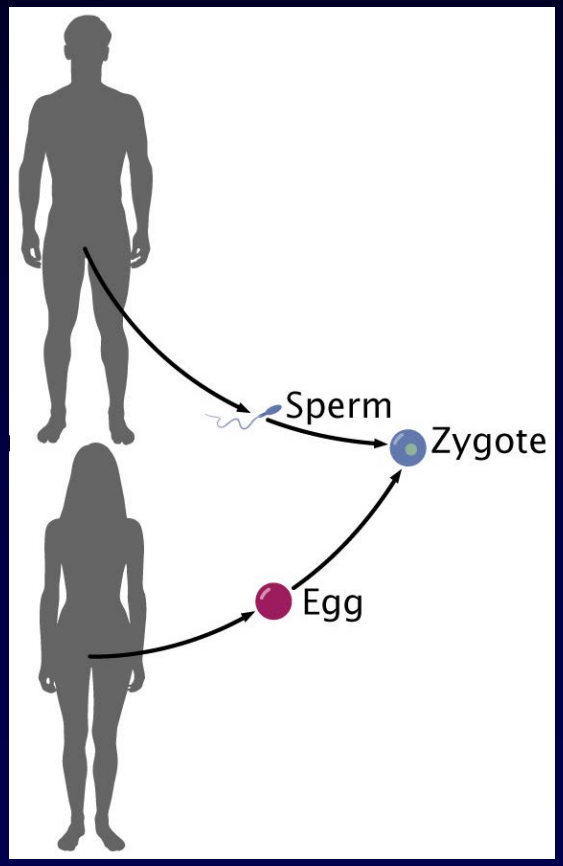
Genomics
(Bradley Shaffer)

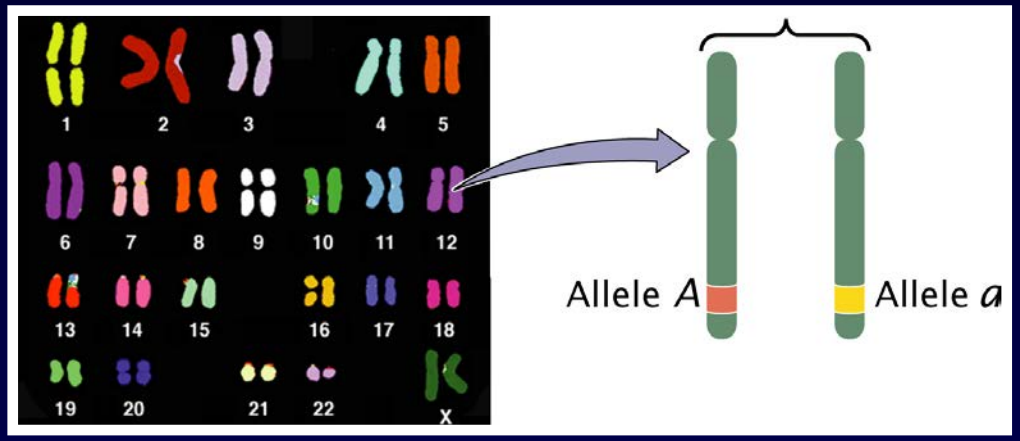
Informing Action: Mountain Lions
(Winston Vickers)

Informing Actions: Burrowing Owls
(Brenda Johnson)

Small Group Discussions

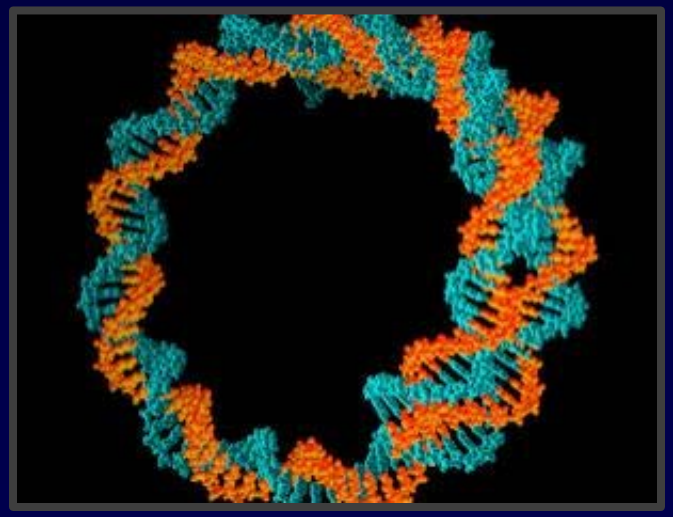
We most often study nuclear genes (inherited from both parents)



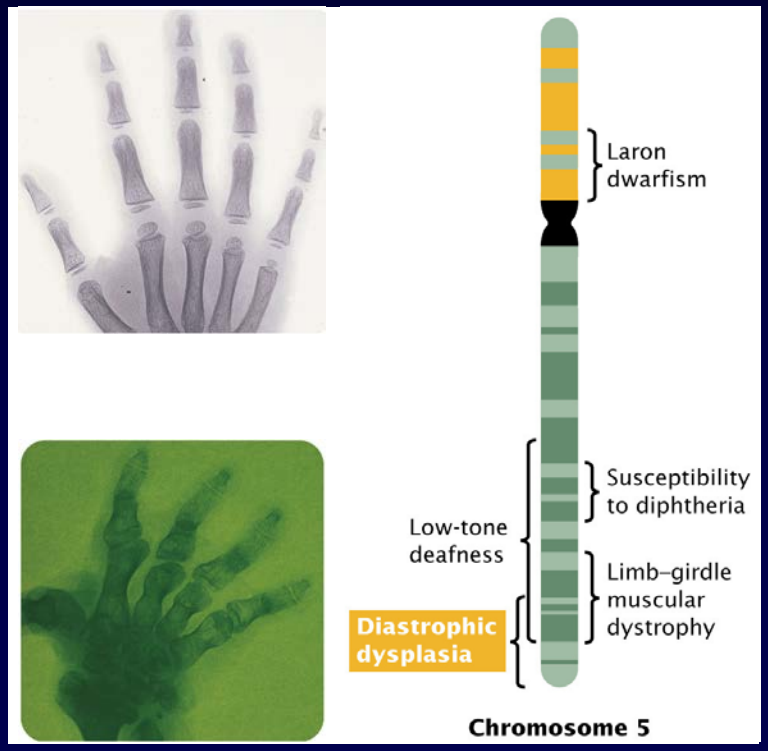


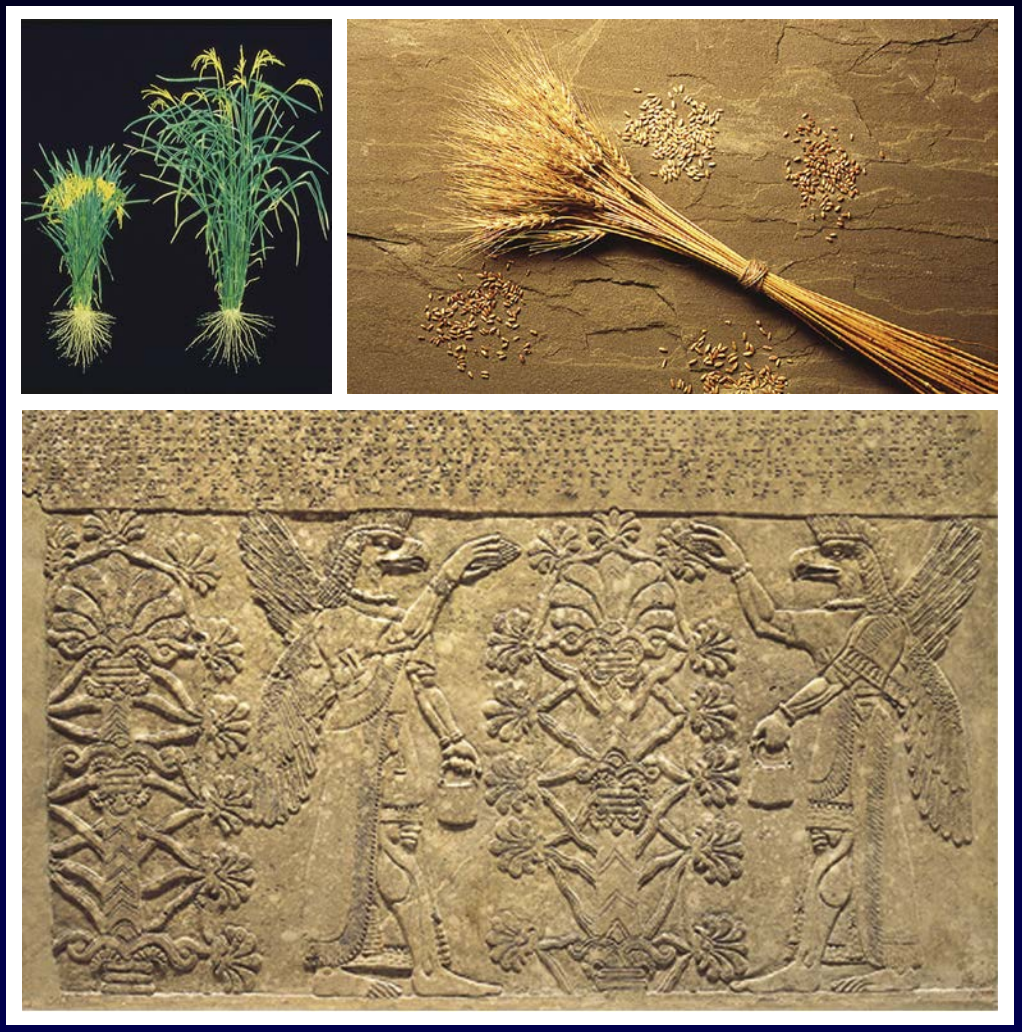
… but also …

mitochondrial genes, chloroplast genes, organisms with clonal reproduction or mixed reproductive modes, viruses, polyploid species (which have > 2 gene copies) …



Genetic variation is the norm, not the exception





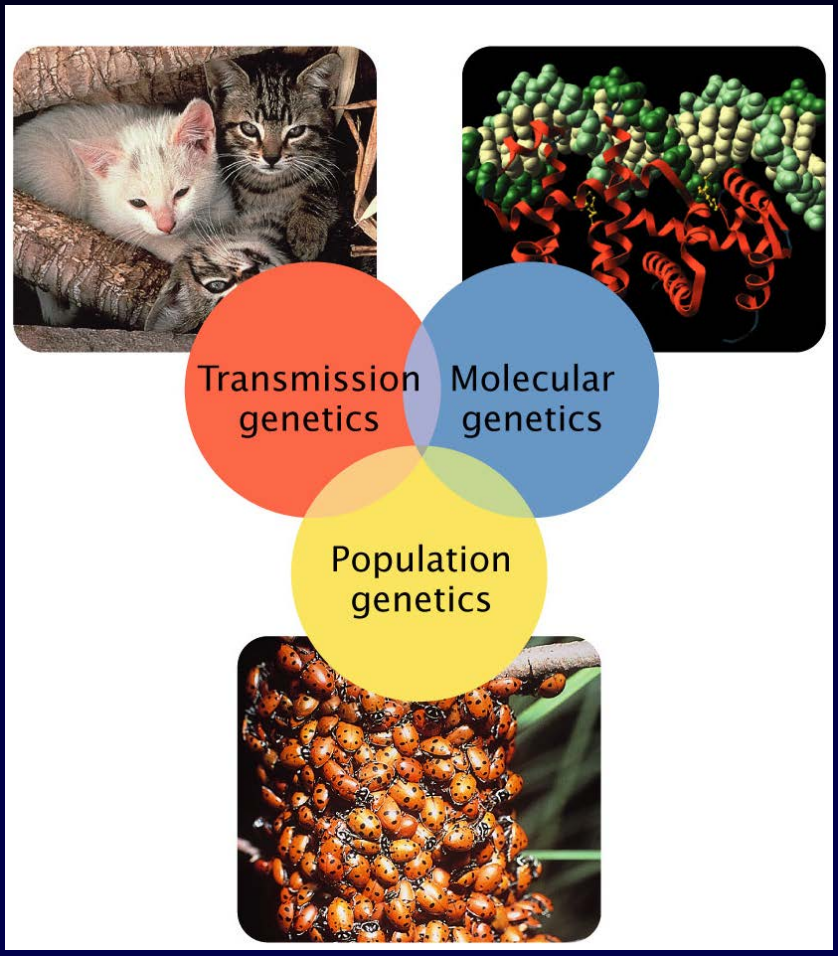
Areas of genetic research

Subdisciplines of genetics:

Molecular genetics
- primarily concerned with biochemical mechanisms

Transmission genetics
- patterns of inheritance from one generation to the next

Population genetics
- why the genetic composition of populations changes in time and space



Why “population” genetics?

1. **Evolution is “genetic change in a population over time”**

-> the population is the unit of evolution.

2. **Despite regulatory focus on species, populations are usually the unit of management and conservation efforts.**

“Population” genetic variation

Population: an intraspecific group of individuals

- in a single geographic area
- with the capability of interbreeding

For data analysis, we almost always assume that individuals within a population **mate**

- completely randomly
- or according to a set of rules not dependent on spatial position

“Population” genetic variation

Population: an intraspecific group of individuals

- in a single geographic area
- with the capability of interbreeding

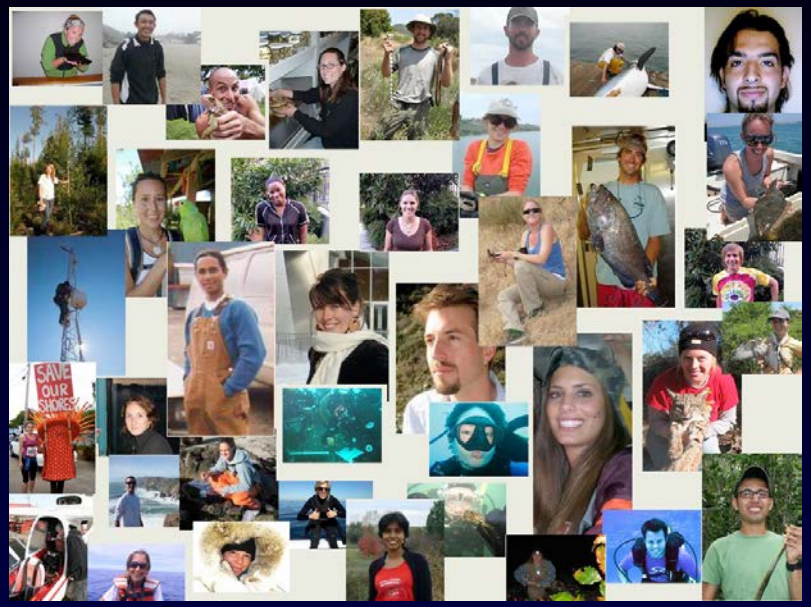
For data analysis, we almost always assume that individuals within a population mate

- completely randomly
- or according to a set of rules not dependent on spatial position

Gene pool: all gene copies in the population

For a genetic study in which “populations” are defined, keep in mind the biology of the organism, the scale of the study, and the way in which samples were obtained.

Gene pools



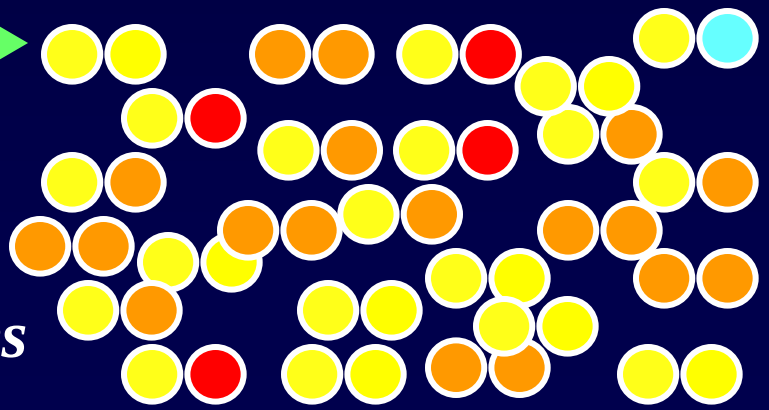
POPULATION

25 randomly mating individuals

GENE POOL

50 gene copies

for one particular gene



4 alleles

The Hardy-Weinberg Model

The simplest model of a gene pool, with many unrealistic assumptions.

1. **Provides a starting point for more realistic models**
2. **Provides a null model for statistical tests. When the Hardy-Weinberg model is rejected, there may be many reasons. Examples:**

   - nonrandom mating
   - predefined population does not fit assumptions
   - natural selection on the gene
   - methodological artifacts

Population genetic analyses

Diversity within and divergence among populations can be understood in terms of only 5 factors:

1. **drift**
2. **mutation**
3. **nonrandom mating**
4. **gene flow**
5. **natural selection.**

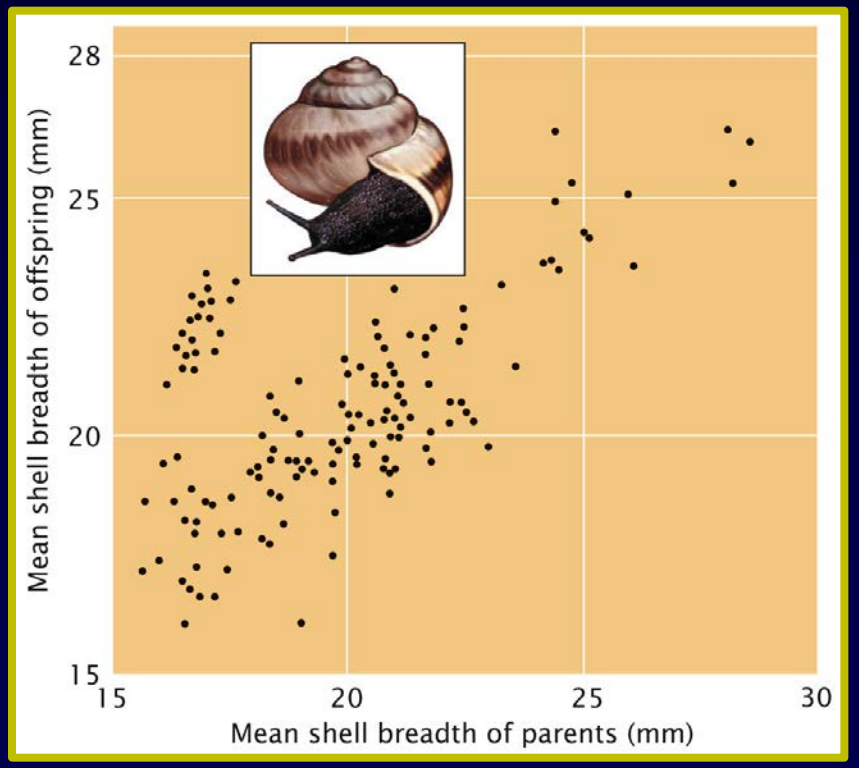
When population parameters change,

6. **nonequilibrium conditions**

must also be considered.

Types of population genetic analysis

1. Quantitative genetics: study of continuous traits for which the specific genetic basis is not known
   - breeding experiments
   - common garden experiments
   - reciprocal transplant experiments
   - estimating heritability
   - estimating the strength of natural selection



Types of population genetic analysis

1. **Quantitative genetics**
2. **Population genomics: Study of numerous genes to better understand the five microevolutionary processes**

   (mating, mutation, drift, gene flow, natural selection)

   More narrow definition: separate genome-wide effects (drift, gene flow, phylogenetic history) from gene-specific effects (mutation, recombination, natural selection).

Types of population genetic analysis

1. Quantitative genetics
2. Genomics
3. Summary statistics for genetic diversity and divergence
4. Analyze specific models of a particular process or scenario
5. Analysis with gene genealogies (trees)

BIOLOGICAL CONSERVATION 141 (2008) 1600–1611

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/biocon

ELSEVIER

Genetic isolation and morphological divergence of Black Sea bottlenose dolphins

Karine A. Viaud-Martinez[a,*], *Robert L. Brownell Jr.*[b], *Anastasia Komnenou*[c], *Andrew J. Bohonak*[a]

[a]Department of Biology, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182-4614, USA
[b]Southwest Fisheries Science Center, 1352 Lighthouse Avenue, Pacific Groove, CA 93950, USA
[c]Veterinary Faculty, Aristotle University of Thessaloniki, St. Voutyra 11, 546 27 Thessaloniki, Greece

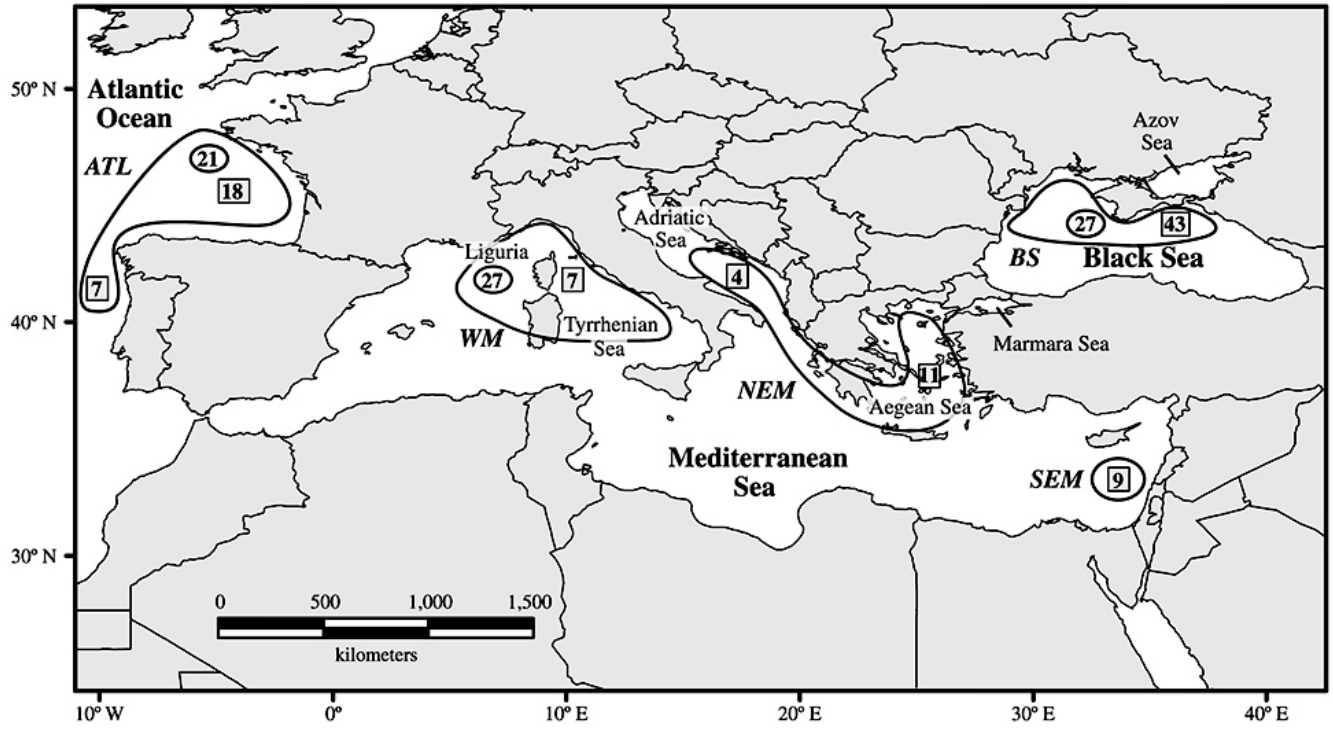


Fig. 1 – Sample sizes across the primary study area. Circled numbers represent the number of skulls measured from each basin. Shaded numbers in squares indicate the sample size at each site for genetic analyses. The large circles represent five designated populations: Atlantic (ATL), western Mediterranean (WM), northeastern Mediterranean (NEM), southeastern Mediterranean (SEM) and Black Sea (BS).

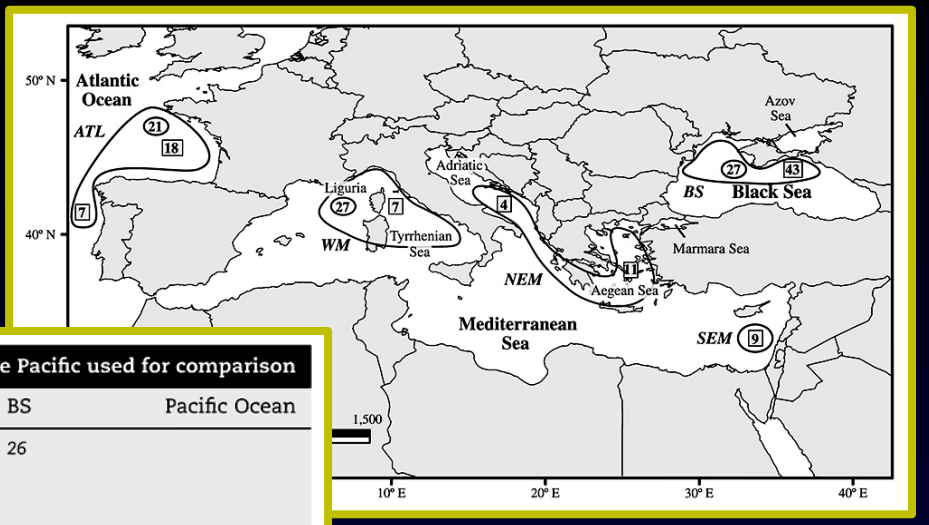


Table 2 – Haplotype affiliations across the five putative populations and 10 samples from the Pacific used for comparison

Haplotype	ATL	WM	NEM	SEM	BS	Pacific Ocean
I	1	2	8		26	
II	1					
III	2					
IV	1					
V	1					
VI	1					
VII	1					
VIII	2					
IX	2					
X	1					
XI	2					
XII	1					
XIII	1					
XIV	1		1	1		
XV	7		3			
XVI		1				
XVII		1				
XVIII		2				
XIX			1			
XX				7		
XXI				1		
XXII			2		1	
XXIII		1			11	
XXIV					2	
XXV					1	
XXVI					2	
XXVII						2
XXVIII						1
XXIX						1
XXX						1
XXXI						1
XXXII						1
XXXIII						1
XXXIV						2
N	25	7	15	9	43	10

Population abbreviations are as in Table 1. Italicized rows represent shared haplotypes among locations.

Table 1 – Genetic differentiation estimated with Φ_{st} (above the diagonal) and exact tests (p values below the diagonal)

Locality	N	ATL	WM	NEM	SEM	BS	Pacific Ocean
ATL	25		0.254**	0.059	0.509**	0.521**	0.363**
WM	7	0.015 ± 0.006		0.084	0.809**	0.191*	0.468**
NEM	15	0.006 ± 0.01	0.029 ± 0.005		0.535**	0.324**	0.372**
SEM	9	<0.001	0.001 ± 0.001	<0.0001		0.895**	0.771**
BS	43	<0.0001	0.003 ± 0.002	<0.0001	<0.0001		0.745**
Pacific Ocean	10	<0.001	0.007 ± 0.002	<0.0001	0.001 ± 0.0007	<0.0001	

All exact test contrasts are significant at $p < 0.05$, and those $\leqslant 0.003$ are significant after Bonferroni correction for 15 tests. *$\Phi_{st} > 0$ ($p < 0.05$), **$\Phi st > 0$ ($p \leqslant 0.003$). Populations are: Atlantic (ATL), western Mediterranean (WM), northeastern Mediterranean (NEM), southeastern Mediterranean (SEM) and Black Sea (BS). Sample sizes for each population are reported in the second column (N).

Table 2 – Haplotype affiliations a

Haplotype	ATL					
I	1	2	8		26	
II	1					
III	2					
IV	1					
V	1					
VI	1					
VII	1					
VIII	2					
IX	2					
X	1					
XI	2					
XII	1					
XIII	1					
XIV	1		1	1		
XV	7		3			
XVI		1				
XVII		1				
XVIII		2				
XIX			1			
XX				7		
XXI				1		
XXII			2		1	
XXIII		1			11	
XXIV					2	
XXV					1	
XXVI					2	
XXVII						2
XXVIII						1
XXIX						1
XXX						1
XXXI						1
XXXII						1
XXXIII						1
XXXIV						2
N	25	7	15	9	43	10

Population abbreviations are as in Table 1. Italicized rows represent shared haplotypes among locations.

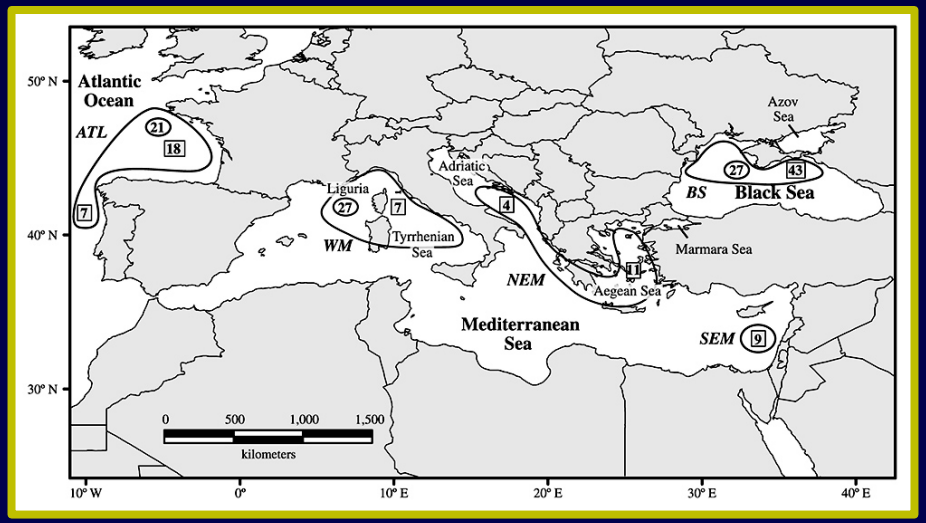
3. Summary statistics for genetic diversity and divergence

a) Estimate summary statistics
b) Test statistical hypotheses
c) Interpret in terms of the 5 microevolutionary processes (mating, mutation, drift, gene flow, natural selection)

Table 1 – Genetic differentiation estimated with Φ_{st} (above the diagonal) and exact tests (p values below the diagonal)

Locality	N	ATL	WM	NEM	SEM	BS	Pacific Ocean
ATL	25		0.254**	0.059	0.509**	0.521**	0.363**
WM	7	0.015 ± 0.006		0.084	0.809**	0.191*	0.468**
NEM	15	0.006 ± 0.01	0.029 ± 0.005		0.535**	0.324**	0.372**
SEM	9	<0.001	0.001 ± 0.001	<0.0001		0.895**	0.771**
BS	43	<0.0001	0.003 ± 0.002	<0.0001	<0.0001		0.745**
Pacific Ocean	10	<0.001	0.007 ± 0.002	<0.0001	0.001 ± 0.0007	<0.0001	

All exact test contrasts are significant at $p < 0.05$, and those $\leqslant 0.003$ are significant after Bonferroni correction for 15 tests. *$\Phi_{st} > 0$ ($p < 0.05$), **$\Phi st > 0$ ($p \leqslant 0.003$). Populations are: Atlantic (ATL), western Mediterranean (WM), northeastern Mediterranean (NEM), southeastern Mediterranean (SEM) and Black Sea (BS). Sample sizes for each population are reported in the second column (N).



3. Summary statistics for genetic diversity and divergence

a) Estimate summary statistics
b) Test statistical hypotheses
c) Interpret in terms of the 5 microevolutionary processes (mating, mutation, drift, gene flow, natural selection)

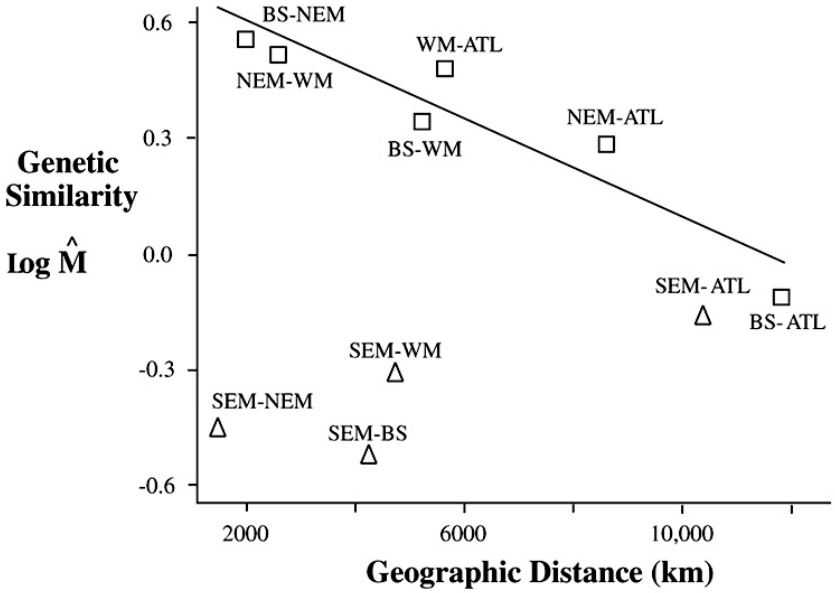
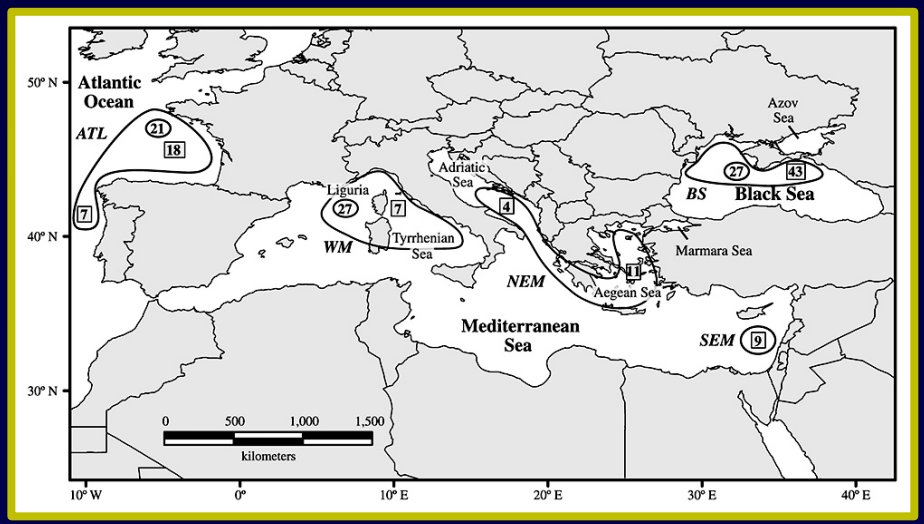


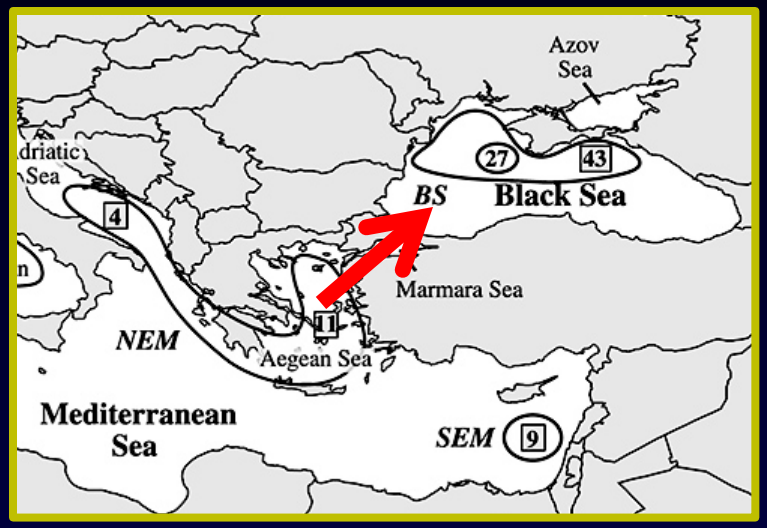
Fig. 3 – "Isolation by distance" plot of genetic similarity vs. geographic distance for all pairs of populations. Triangles represent comparisons with the southeastern Mediterranean population. The association is statistically significant with the southeastern Mediterranean population excluded ($p < 0.001$).

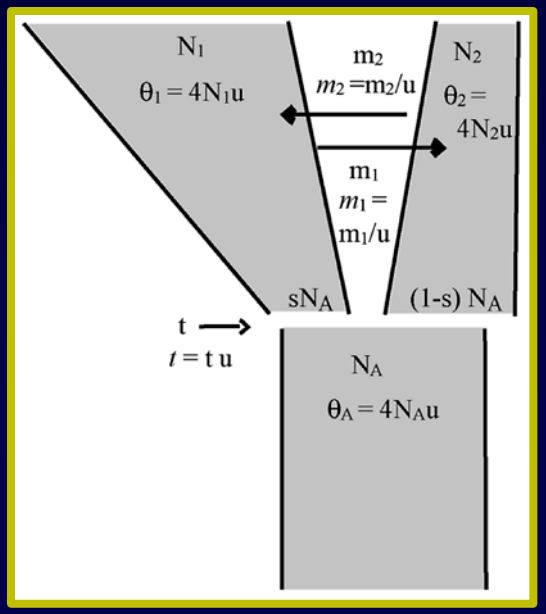


4. *A priori* models of a particular process or scenario

a) Estimate model parameters

b) Compare alternative models





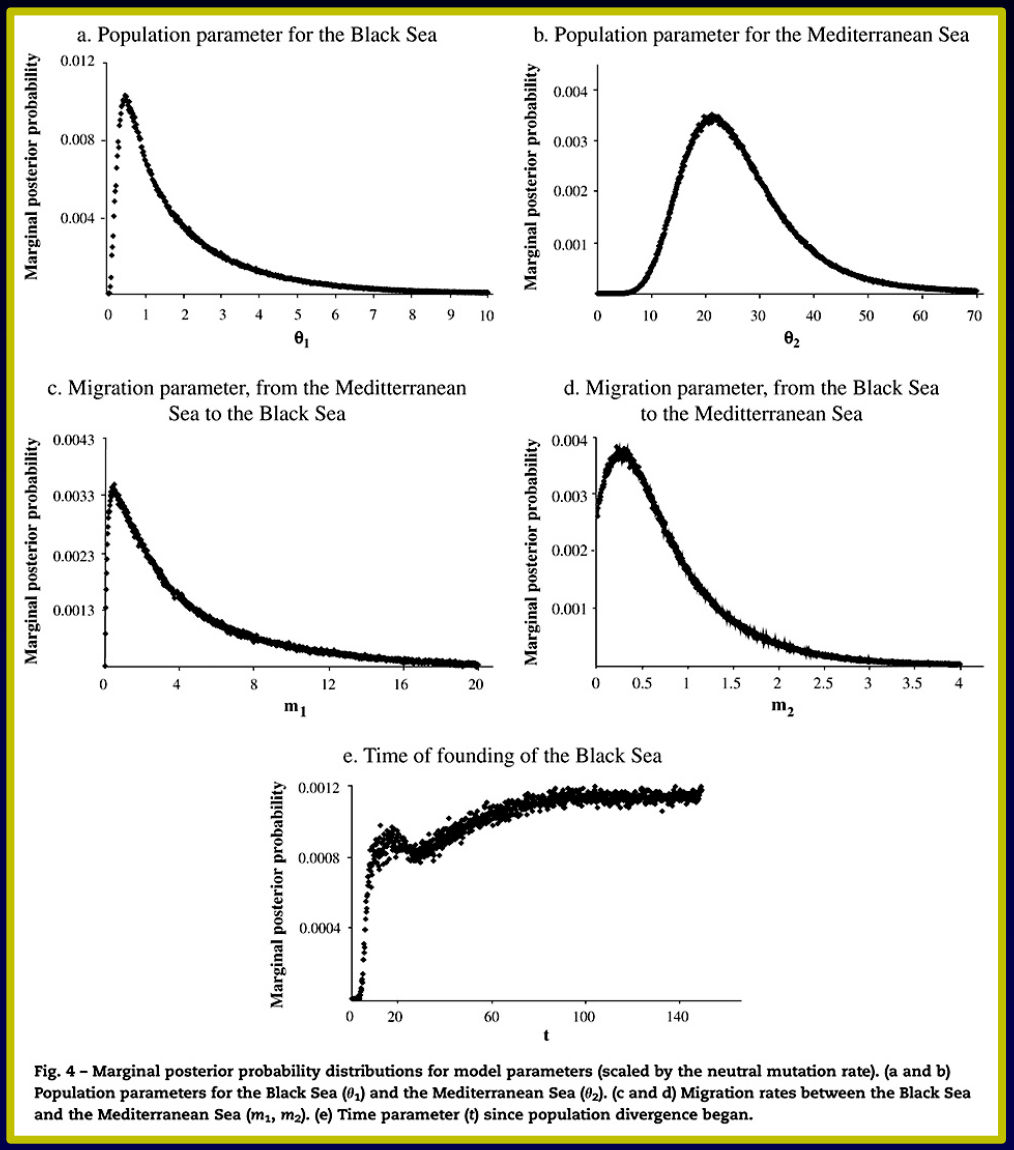


Fig. 4 – Marginal posterior probability distributions for model parameters (scaled by the neutral mutation rate). (a and b) Population parameters for the Black Sea (θ_1) and the Mediterranean Sea (θ_2). (c and d) Migration rates between the Black Sea and the Mediterranean Sea (m_1, m_2). (e) Time parameter (t) since population divergence began.

5. Analysis with gene genealogies (trees)

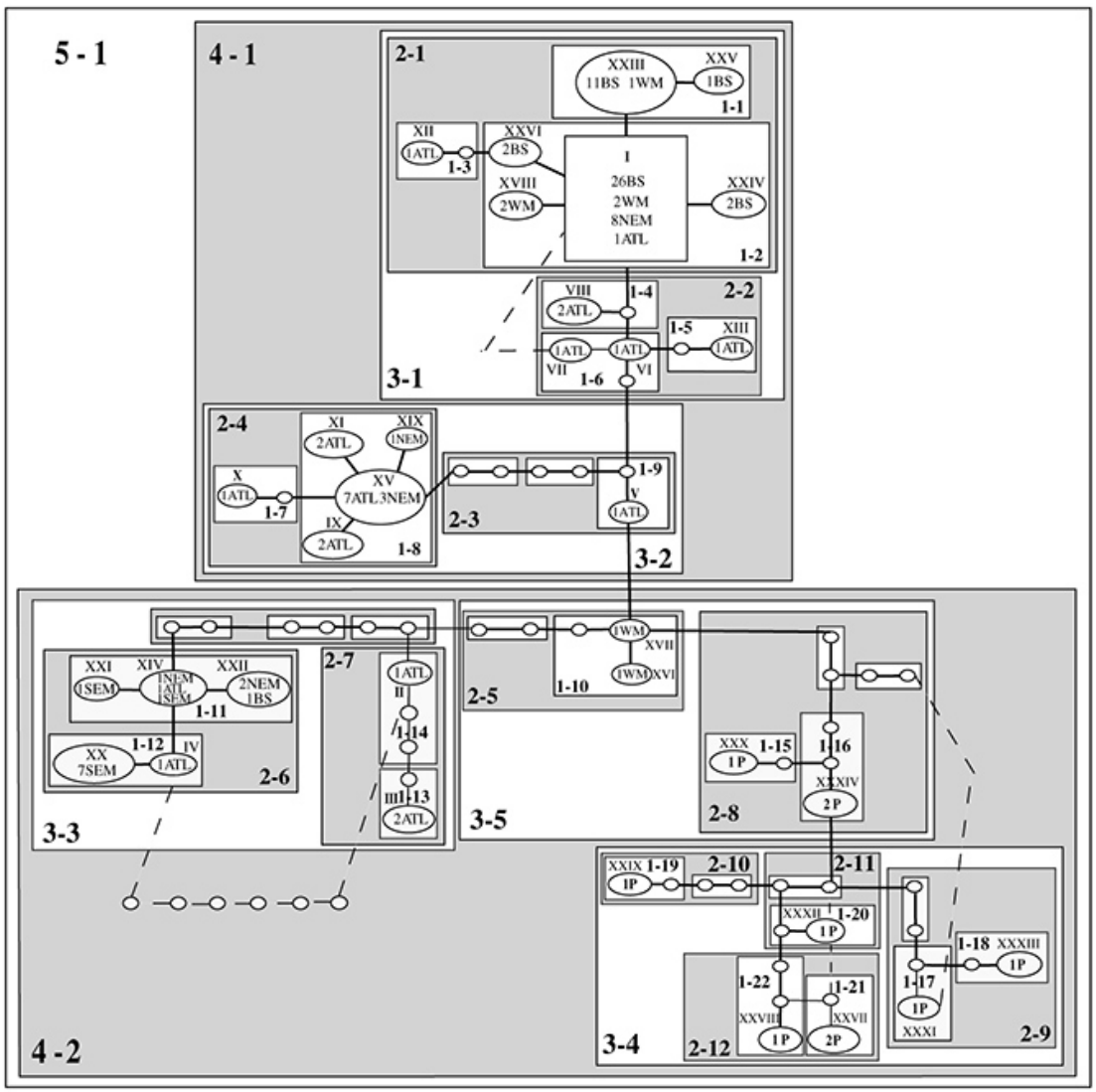


Fig. 2 – Haplotype network for mitochondrial DNA. Population codes are as in Fig. 1, with the addition of samples from the eastern Pacific (P). Oval size is proportional to the number of individuals, which precedes the population code. Each line represents one mutational step, and empty ovals are unsampled haplotypes. Four levels of nesting are enclosed in successively larger boxes. Three ambiguous loops (unresolved portions of the network) were resolved according to the criteria of Crandall et al. (1994) and Templeton et al. (1995).

Types of population genetic analysis

1. Quantitative genetics
2. Genomics
3. Summary statistics for genetic diversity and divergence
4. Analyze specific models of a particular process or scenario
5. Analysis with gene genealogies (trees)

Methodological and logistical issues

1. Tissue collection and preservation

- **Project- and species-specific**
- **General goal: to inactivate DNA-detroying enzymes as quickly as possible by freezing, dessicating or otherwise preserving the sample.**
- **Ultracold freezer for long-term storage**
- **DNA and RNA preservation and extraction methods differ.**

Methodological and logistical issues

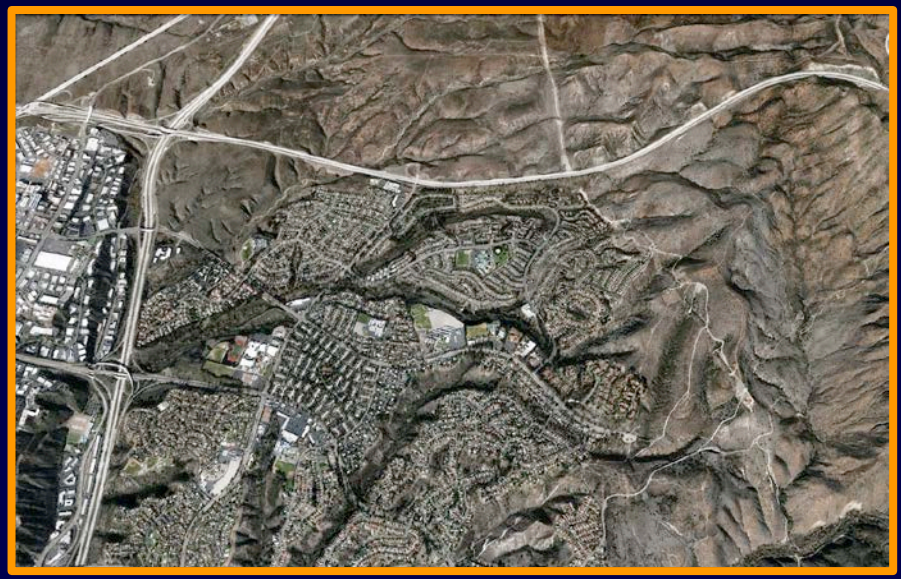
2. Identify goals for data collection

- Delineate population boundaries
- Quantify intrapopulation diversity
- Quantify interpopulation divergence
- Estimate parentage
- Estimate average interpopulation movement rates
- Estimate individual movement rates or ranges

Example of project goal for fragmented habitat

Primary goal: Quantify genetic diversity in habitat fragments of varying sizes

Secondary goal: Estimate average interfragment movement rates



Methodological and logistical issues

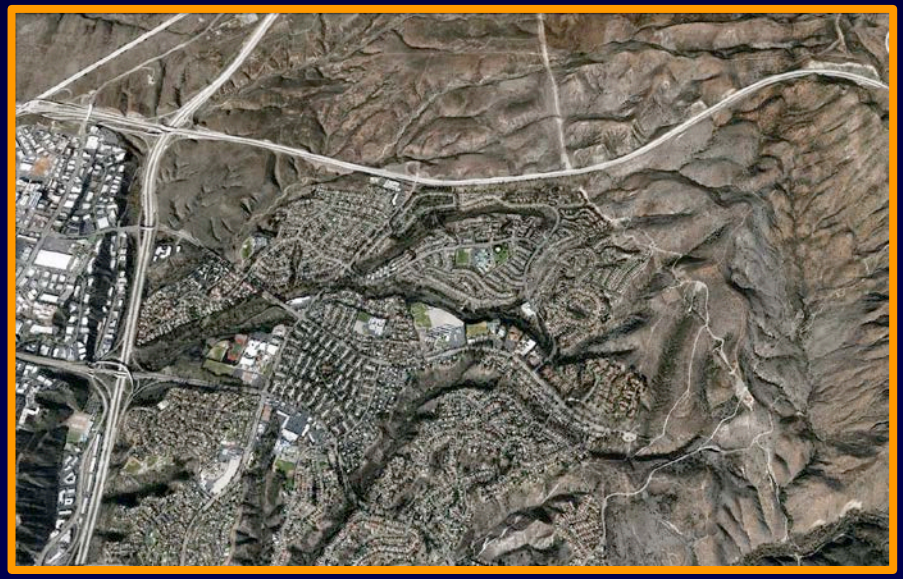
2. **Identify goals for data collection**
   - **Quantifying relationships or estimating parameters is preferable to testing hypotheses.**

Example of project goal for fragmented habitat

Estimate how much lower genetic diversity (H_e) is in small fragments (< 10 hectares) than in large fragments. Report the estimated decrease with a 95% CI.

or

Test null: H_e [fragments < 10 ha] = H_e [fragments ≥ 10 ha]



Methodological and logistical issues

3. Study design and analysis are goal-specific

<u>Examples</u>

a. Compare genetic diversity among sites?

-> sample as many sites as possible, modest to high number of genes (10+), 20+ individuals per site

b. Estimate gene pool boundaries?

-> modest to high number of moderately or highly polymorphic genes (12+), 20+ individuals per putative gene pool, even spatial coverage

c. Estimate the effects of a particular road on connectivity?

-> similar genetic requirements. Samples focused on multiple paired sites on same and opposite sides of road.

Methodological and logistical issues

4. Clarify short and long-term goals

<u>Examples</u>

a. Single study with specific goals and no long-term plans

-> choose most effective set of molecular markers to estimate parameters with a high degree of accuracy

b. Plans for genetic monitoring / future studies

-> consider investing in additional genetic markers/technologies: cost-effectiveness may be a high priority

-> plans for continuing tissue collection and storage

c. Model species / additional questions about physiology, adaptation, disease susceptibilty, etc.

-> potentially costly investment in genomics reources

Genetics for Monitoring and Management Workshop

Genetics 101: State of the Science
(Andrew Bohonak)

Estimating Diversity and Effective Size within Populations
(Jonathan Richmond)

BREAK

Estimating Gene Flow and Movement Among Populations
(Amy Vandergast)

Informing Action: Mountain Yellow Legged Frogs
(Robert Fisher)

LUNCH & Small Group Discussions

Managing for Genetic Variation: When, Why and How?
(Ollie Ryder)

Special Consideration for Plants
(Norman Ellstrand)

BREAK

Genomics
(Bradley Shaffer)

Informing Action: Mountain Lions
(Winston Vickers)

Informing Actions: Burrowing Owls
(Brenda Johnson)

Small Group Discussions