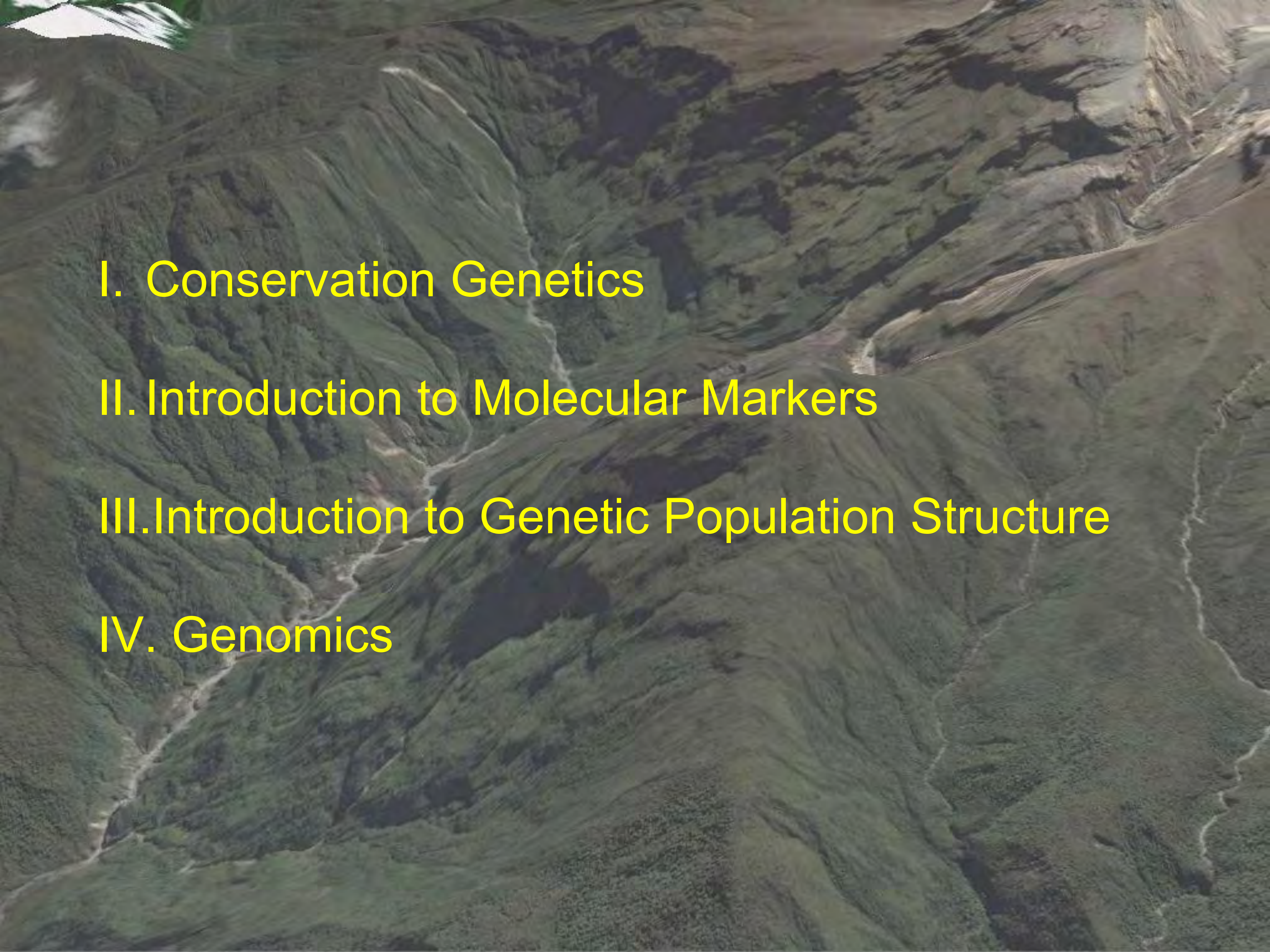
Conservation Genetics: Introduction to molecular markers, parameters, and sampling for monitoring wildlife

Dan Bingham

Michael Schwartz

Rocky Mountain Research Station



I. Conservation Genetics

II. Introduction to Molecular Markers

III. Introduction to Genetic Population Structure

IV. Genomics

What is Conservation Genetics?

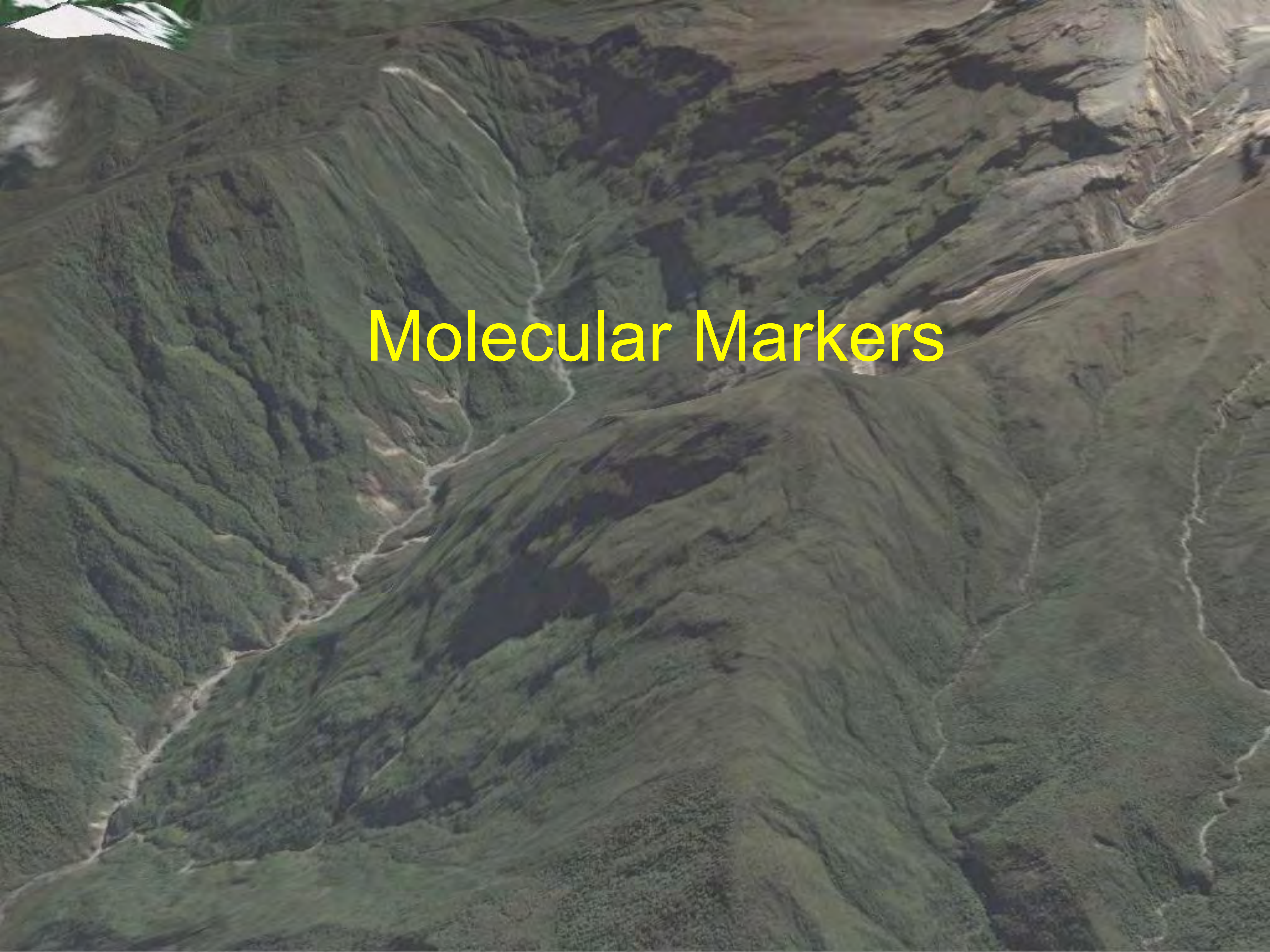
Conservation genetics is an applied science, involving the application of evolutionary and molecular genetics to biodiversity conservation

Frankham 2010

Questions We Address with Genetics

- What is the geographic distribution of a species?
- Does this species live in the forest I manage?
- Is the population increasing, decreasing, or stable?
- Is the population connected to others?
- Is there genetic structure (i.e., genetically distinct populations)?

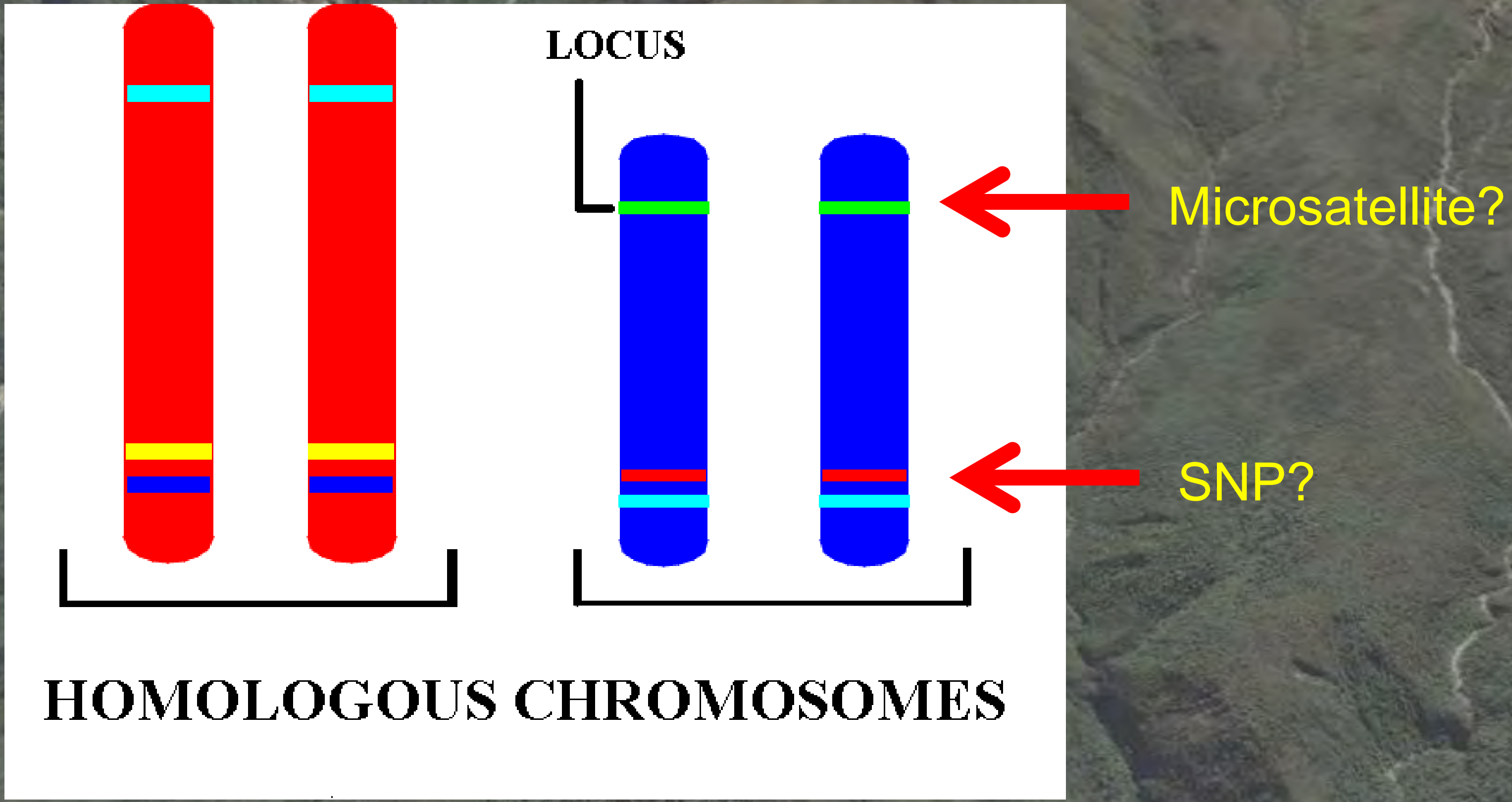
Molecular Markers

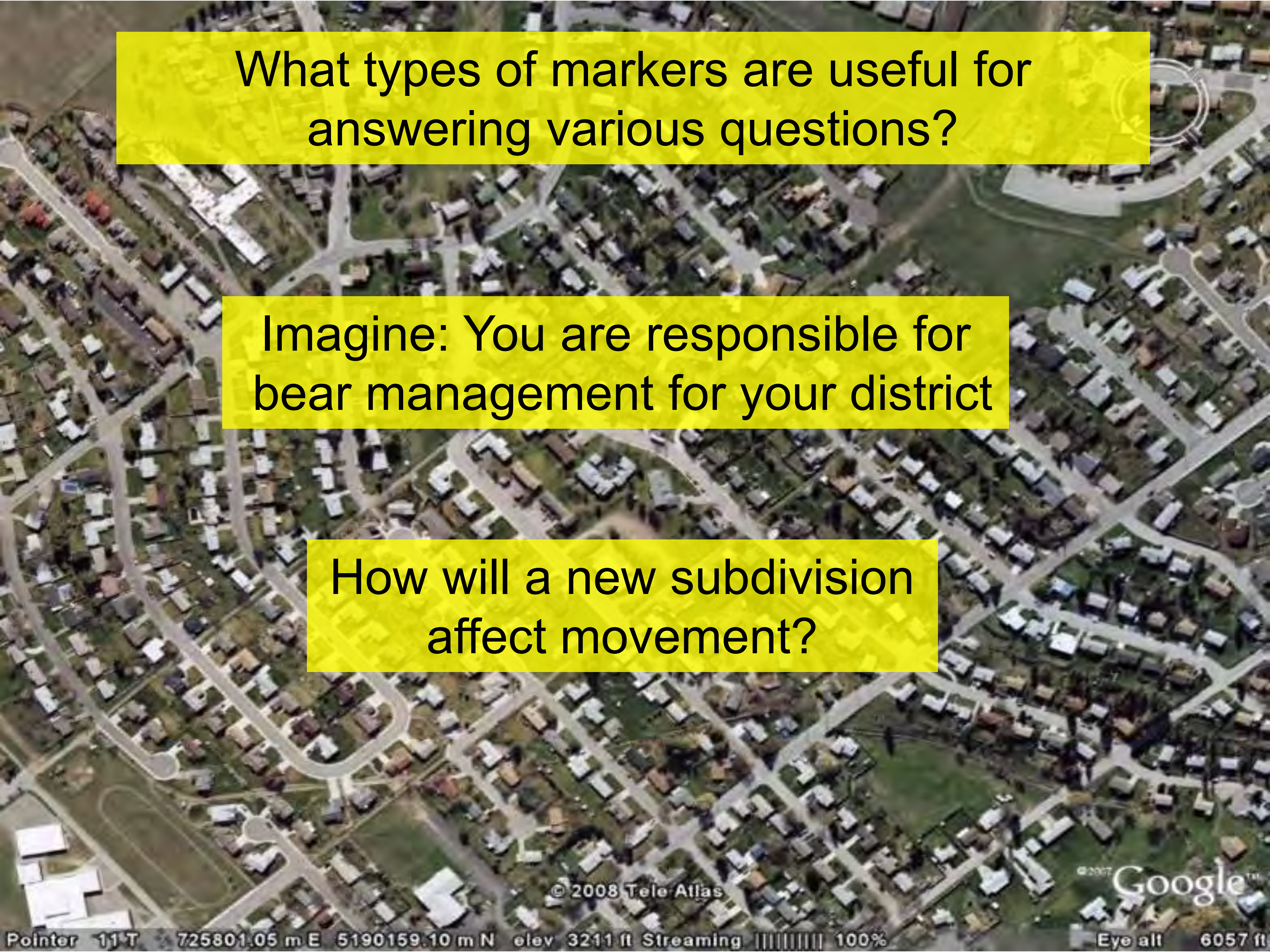


What is a Molecular Marker?

A Type of DNA Polymorphism

That is Associated with a Certain Locus



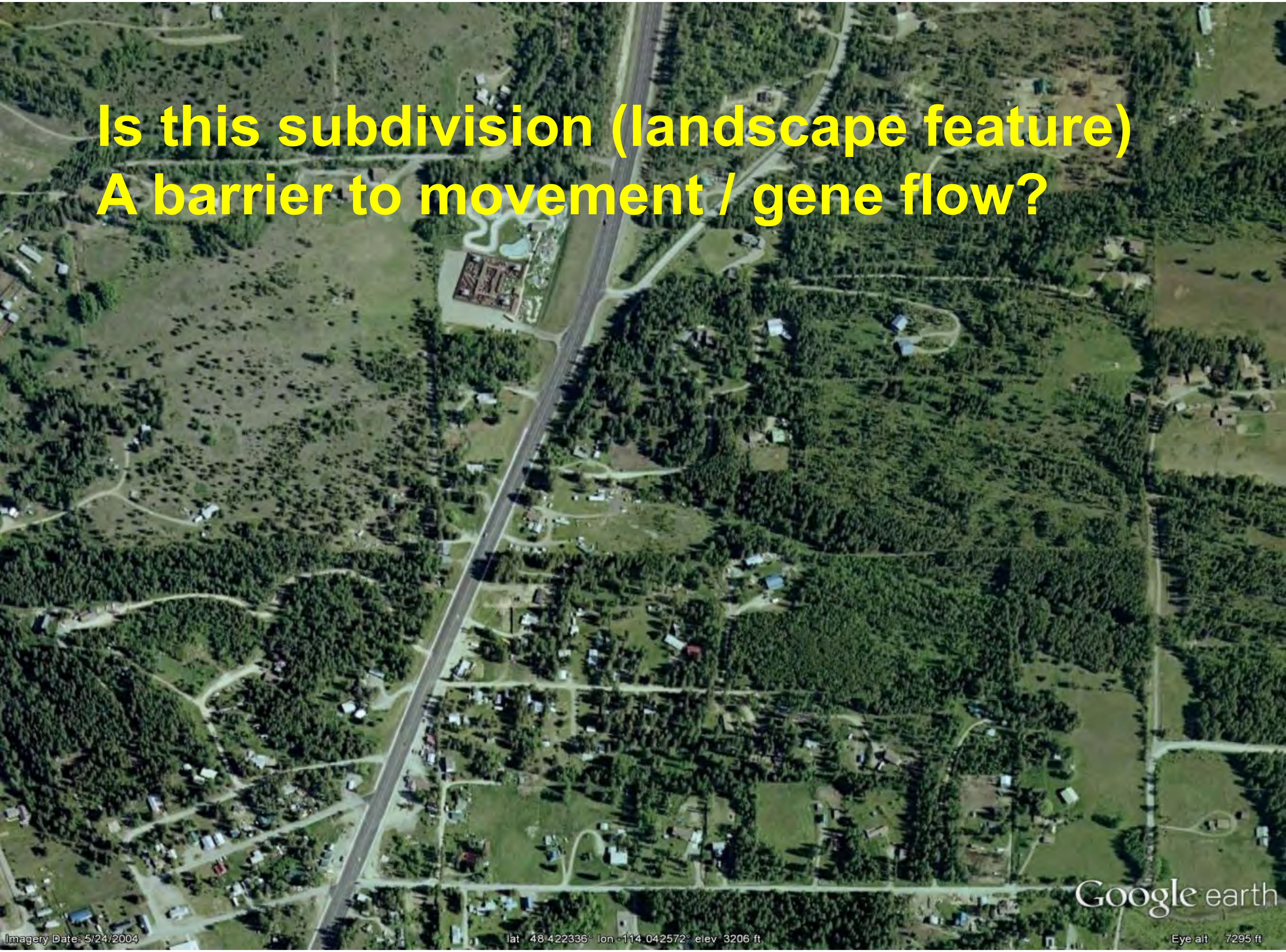


What types of markers are useful for answering various questions?

Imagine: You are responsible for bear management for your district

How will a new subdivision affect movement?

Is this subdivision (landscape feature) A barrier to movement / gene flow?

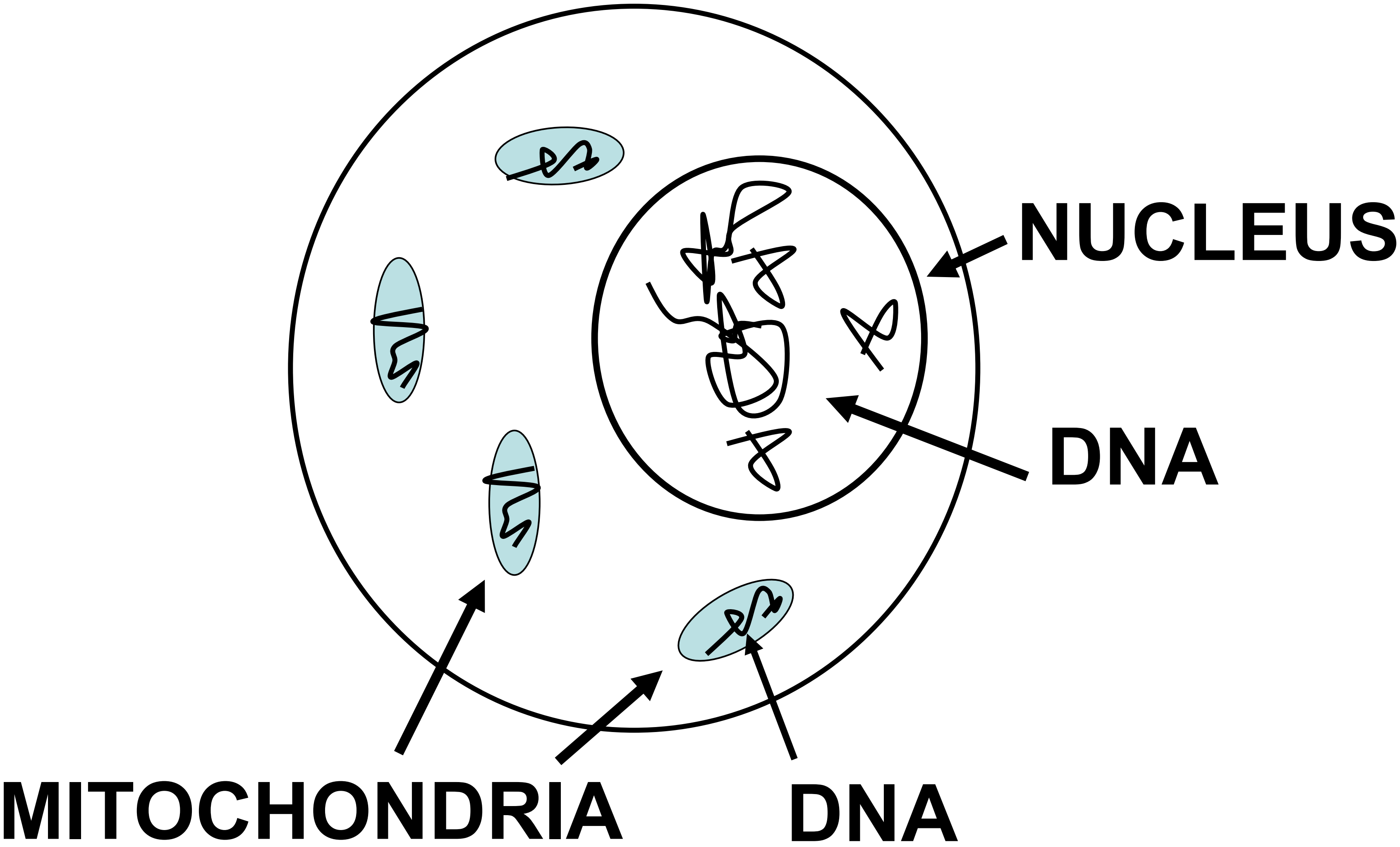


DNA Obtained From:

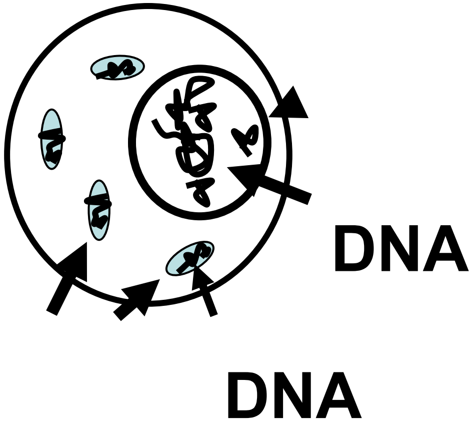
- "Optimal Samples"
  - Tissue
  - Blood
- "Sub-Optimal Samples"
  - Hair
  - Scat
  - Urine
  - Skins/Museum Specimens
  - Feathers
  - Guano


A. Macleod

Quick Genetics Primer



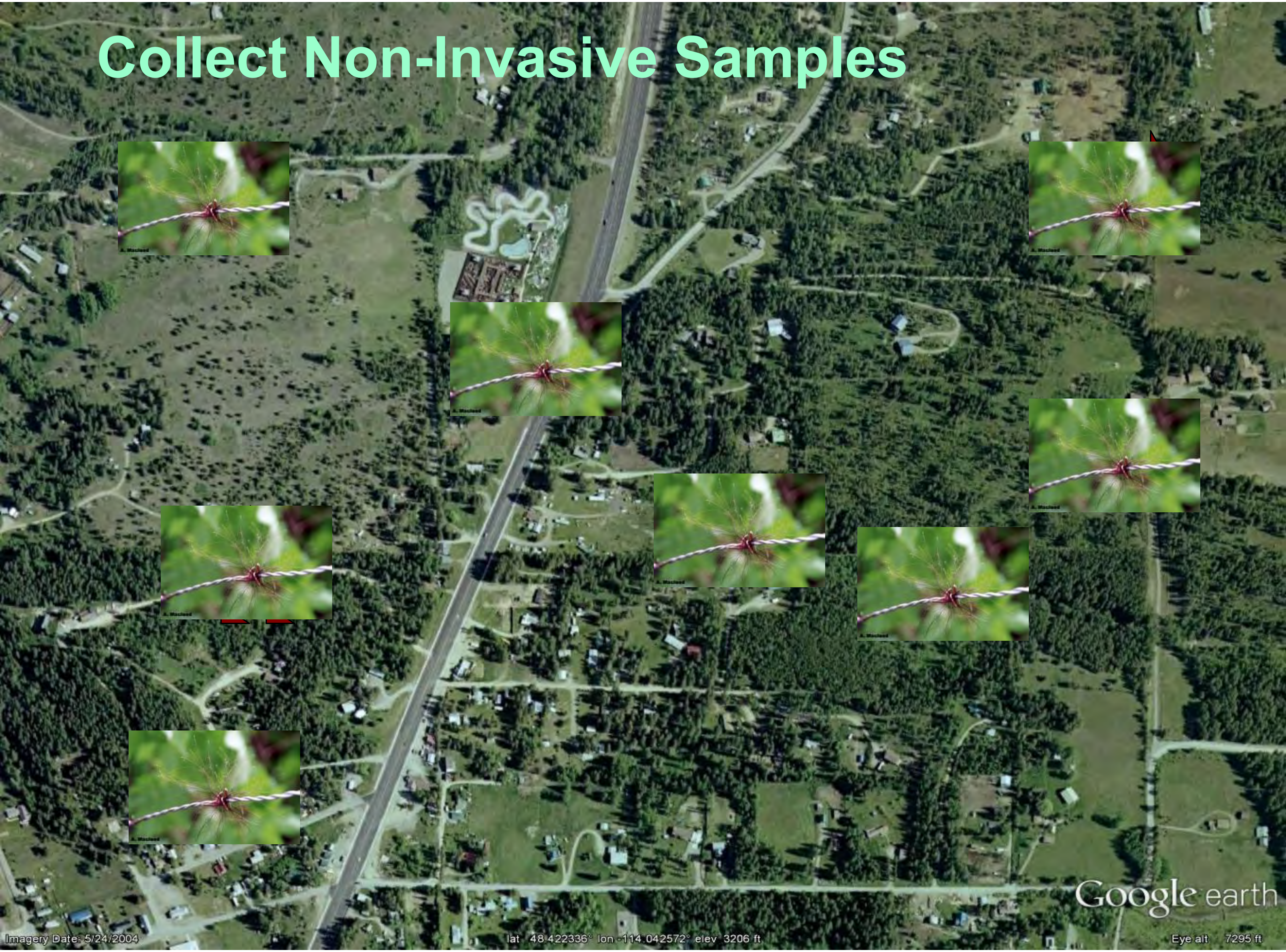
Two Types of DNA



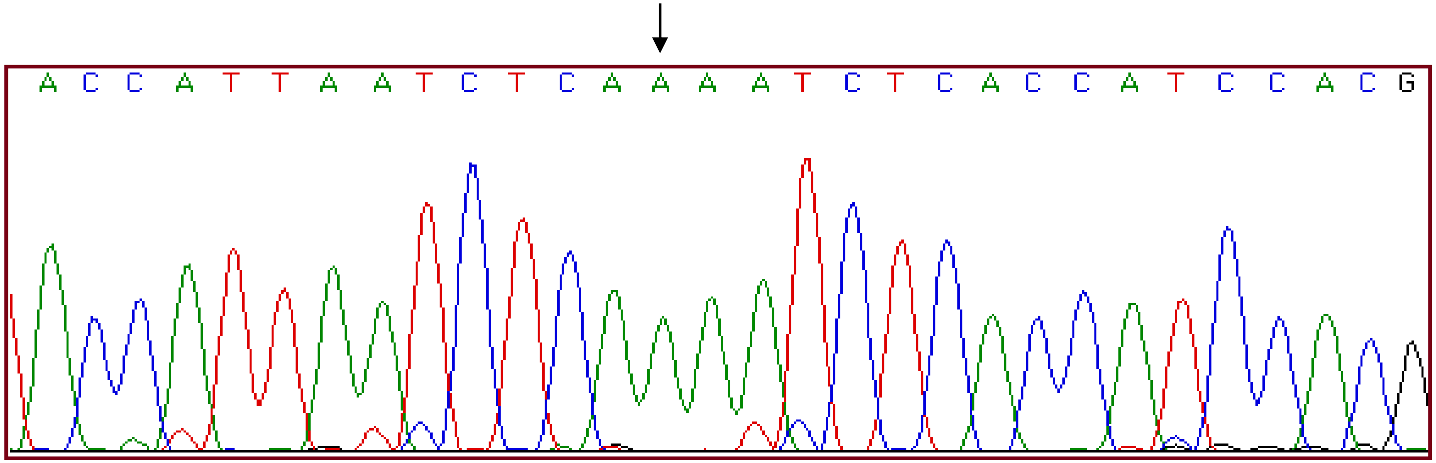
- **Mitochondrial DNA (mtDNA)**
  - 1000's of copy per cell (20x more)
  - Maternally inherited
  - Highly conserved (Species ID)

- **Nuclear DNA**
  - Two copies per cell
  - Inherited from both parents
  - Highly variable regions (microsatellite DNA) for distinguishing individuals

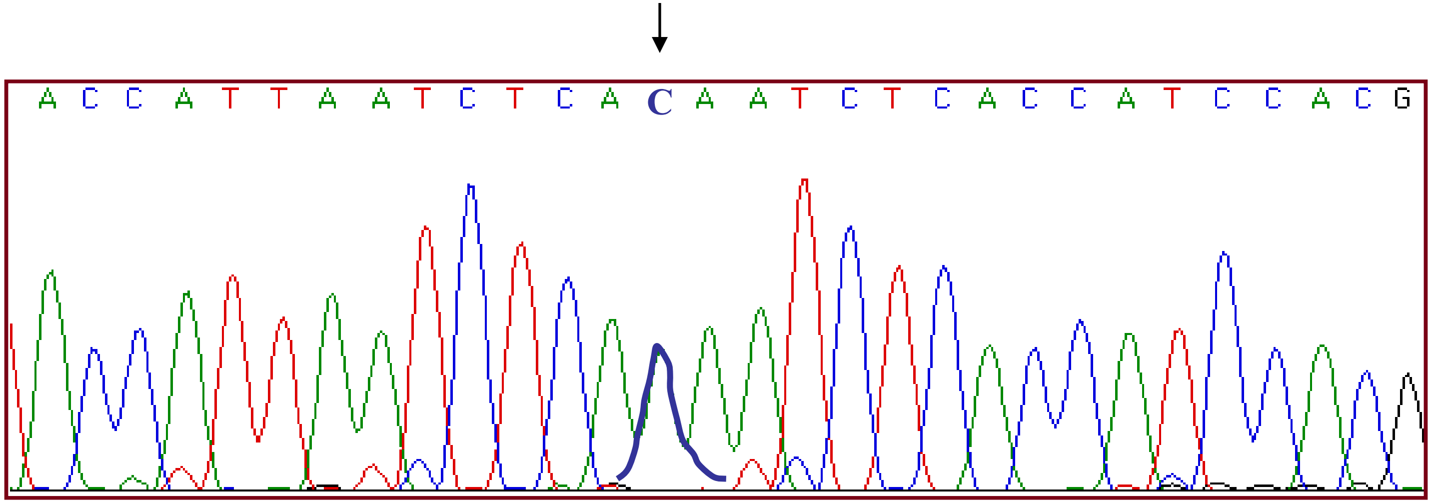
Collect Non-Invasive Samples



Species Identification



Black Bear

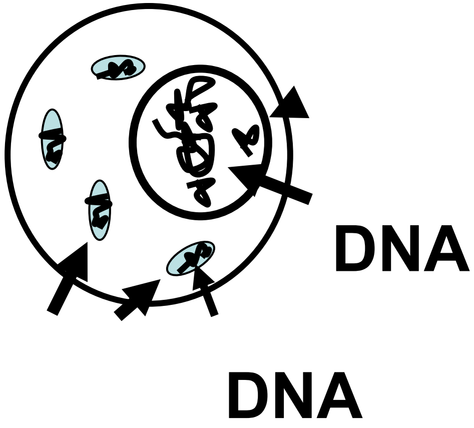


Brown Bear

Compare the mtDNA sequences to online databases

How many individuals?

Two Types of DNA



- **Mitochondrial DNA (mtDNA)**
  - 1000’s of copy per cell (20x more)
  - Maternally inherited
  - Highly conserved (good for species ID)

- **Nuclear DNA**
  - Two copies per cell
  - Inherited from both parents
  - Highly variable regions (microsatellite DNA for distinguishing individuals)

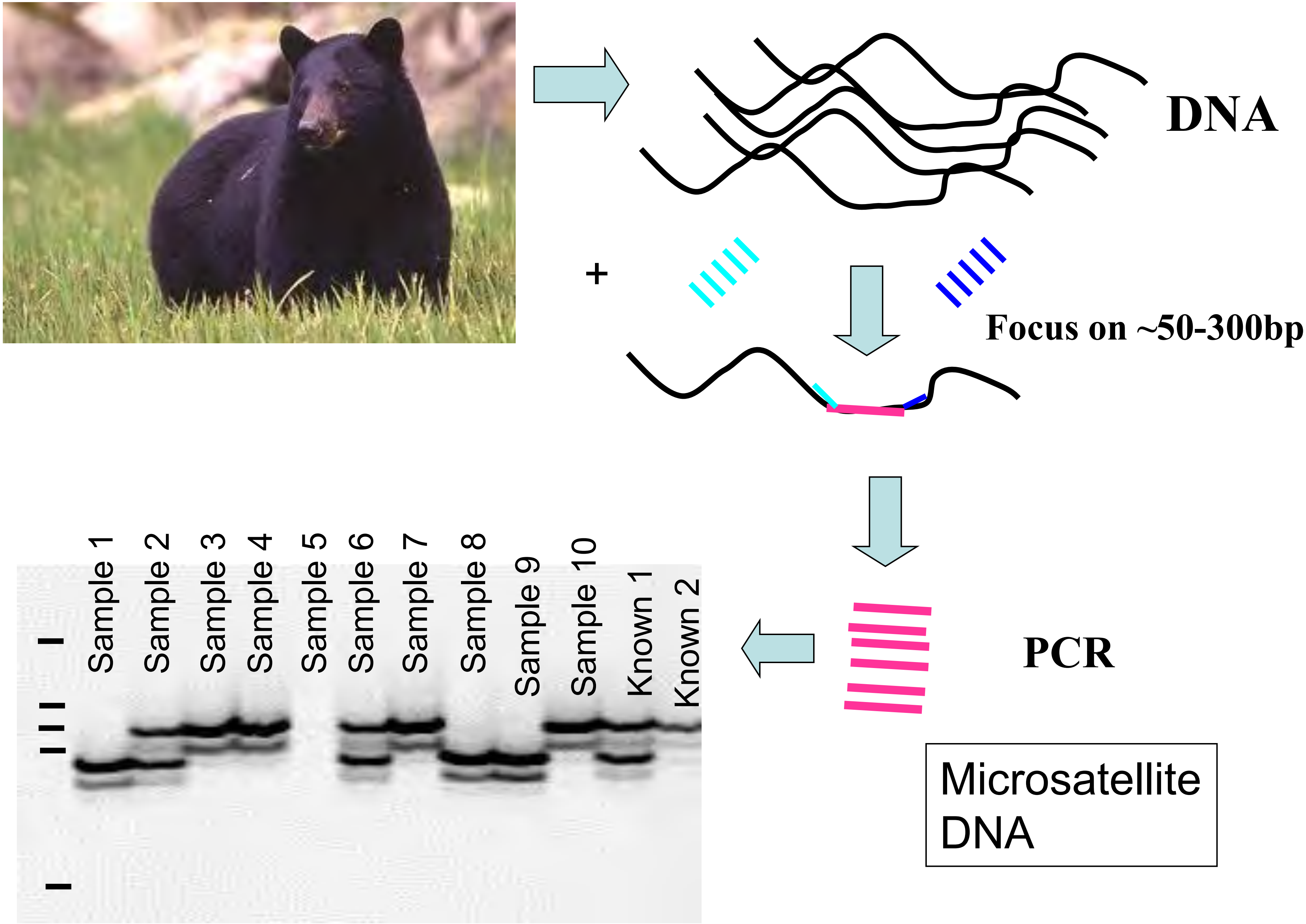
Microsatellite DNA

- Highly Variable DNA Region

A stretch of DNA with mono-, di-, tri- or tetranucleotide units repeated

Examples:

AAAAAAAAAAAAAAAAAA….
GTGTGTGTGTGTGTGTGT…..
CATCATCATCATCATCAT…..
ACGGACGGACGGACGGA….



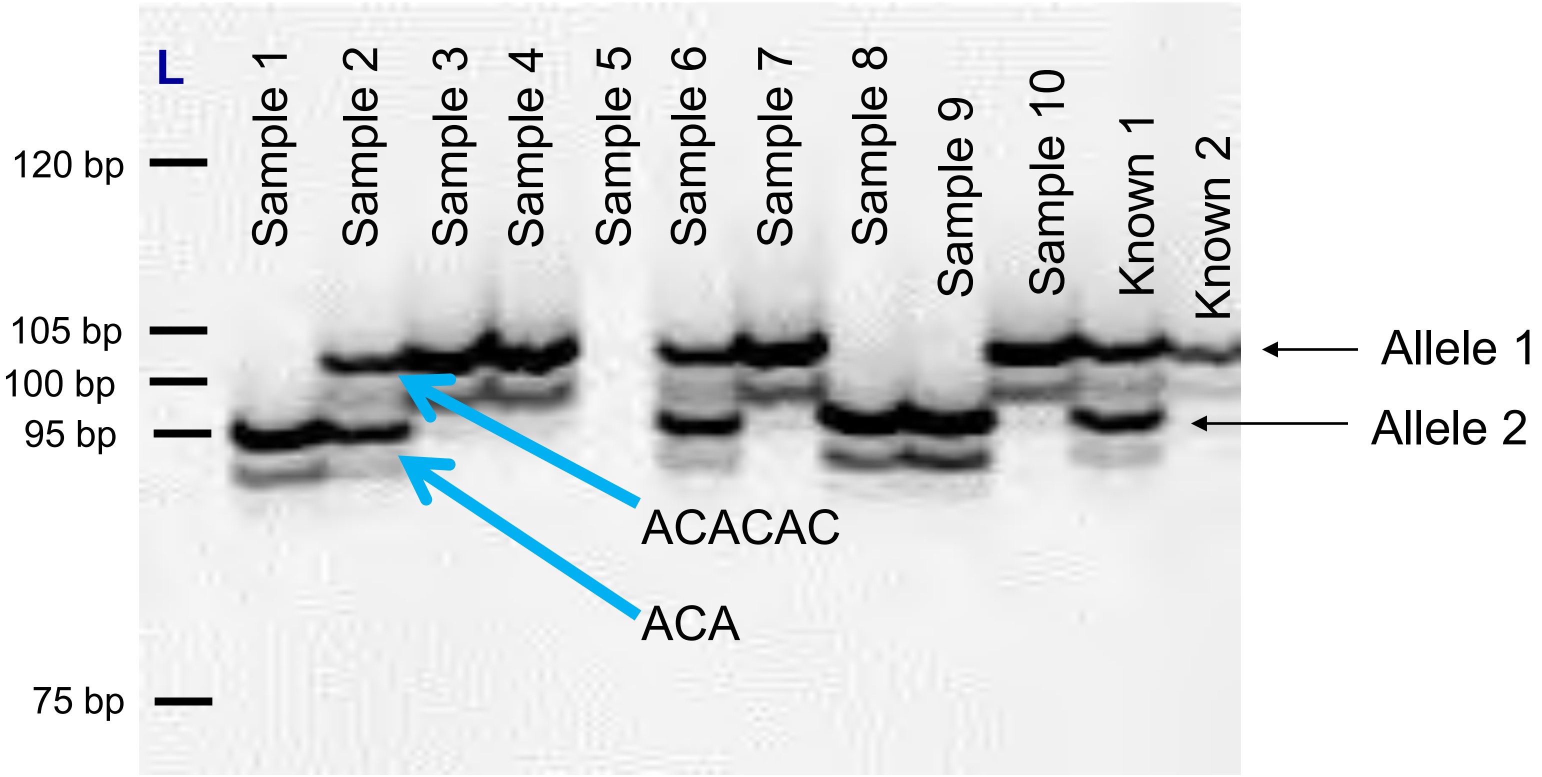
DNA

+

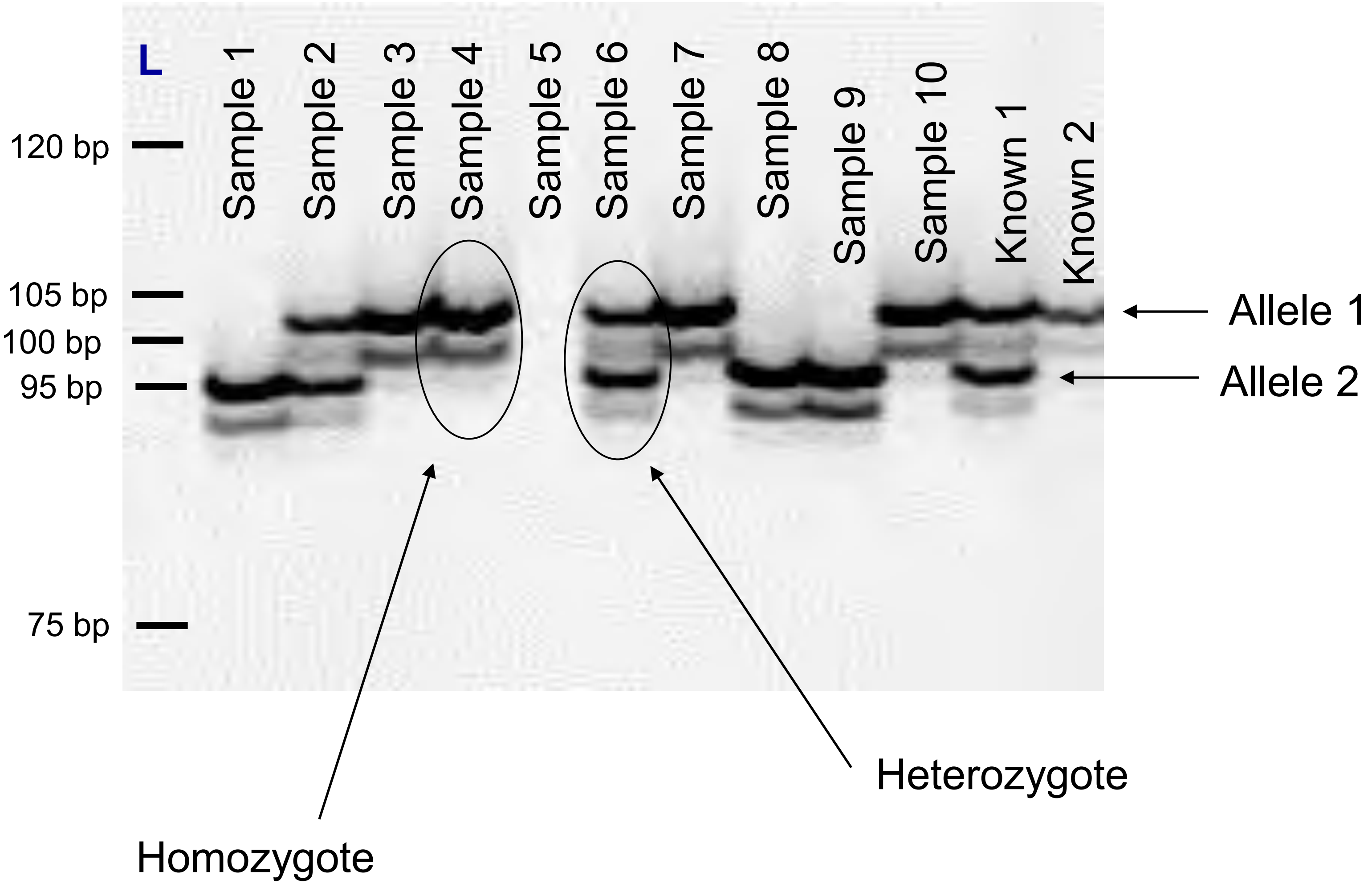
Focus on ~50-300bp

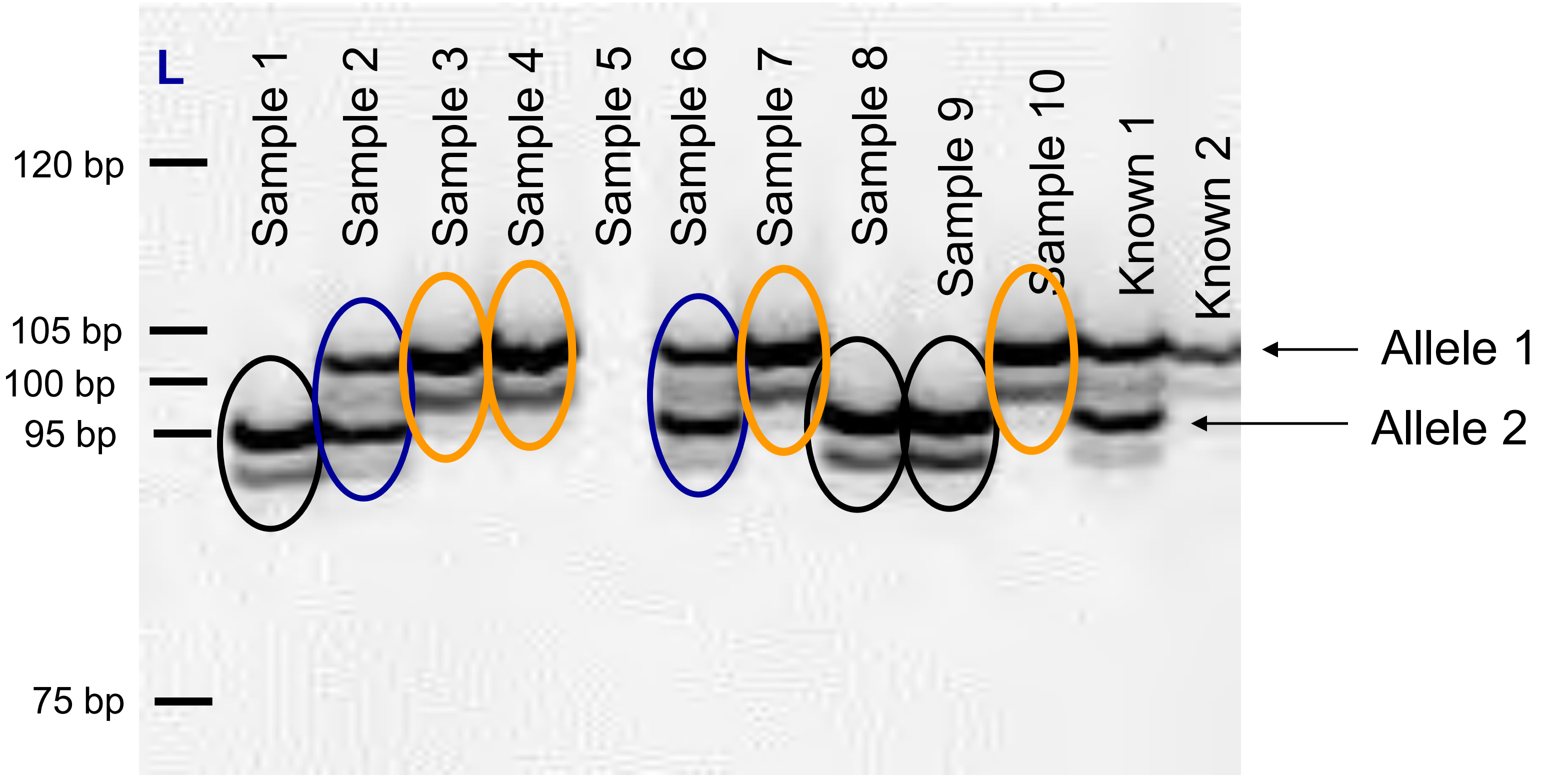
PCR

Microsatellite DNA



Allele = Polymorphisms that exist at a particular locus


L
Sample 1
Sample 2
Sample 3
Sample 4
Sample 5
Sample 6
Sample 7
Sample 8
Sample 9
Sample 10
Known 1
Known 2
120 bp
105 bp
100 bp
95 bp
75 bp
Allele 1
Allele 2
Homozygote
Heterozygote

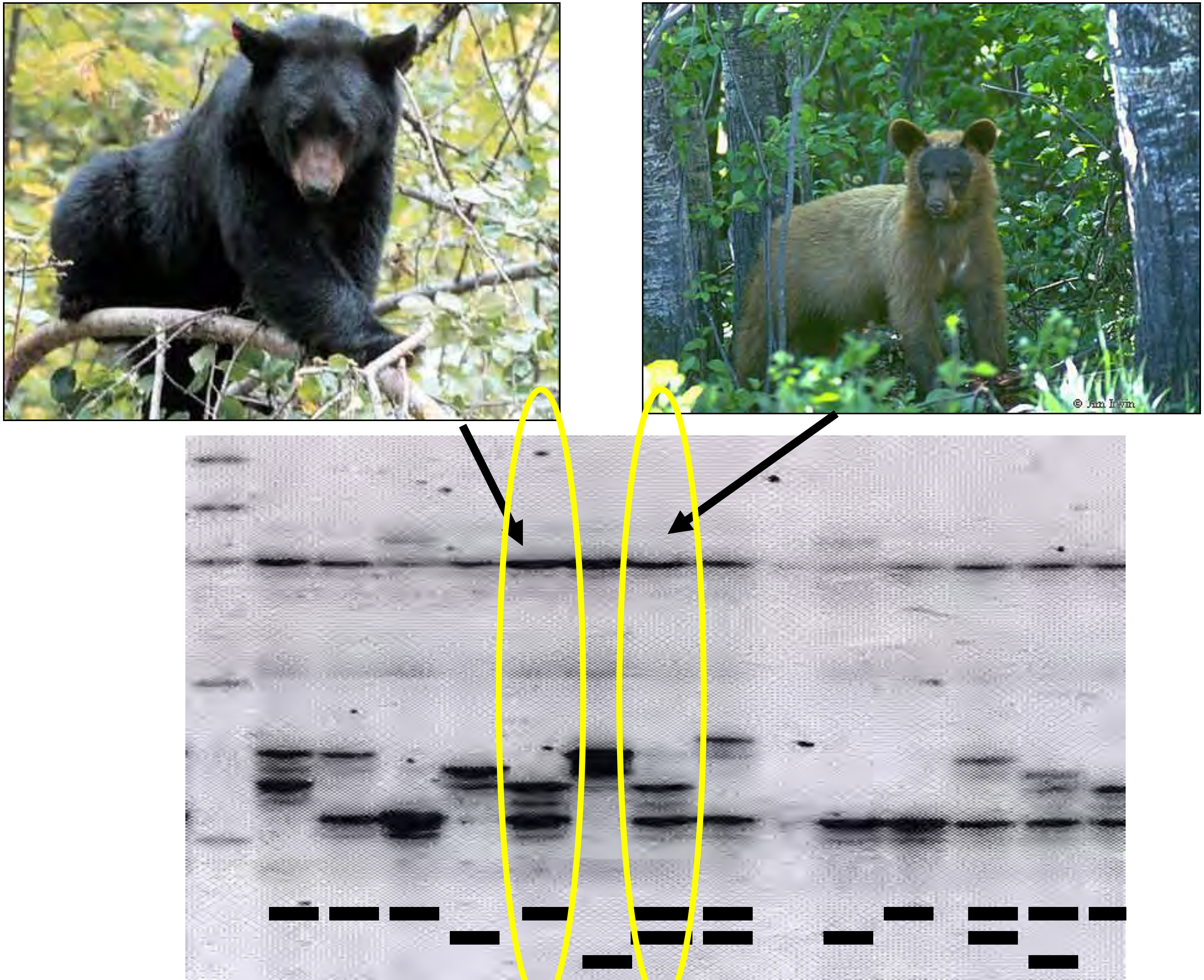


Individual 1

Individual 2

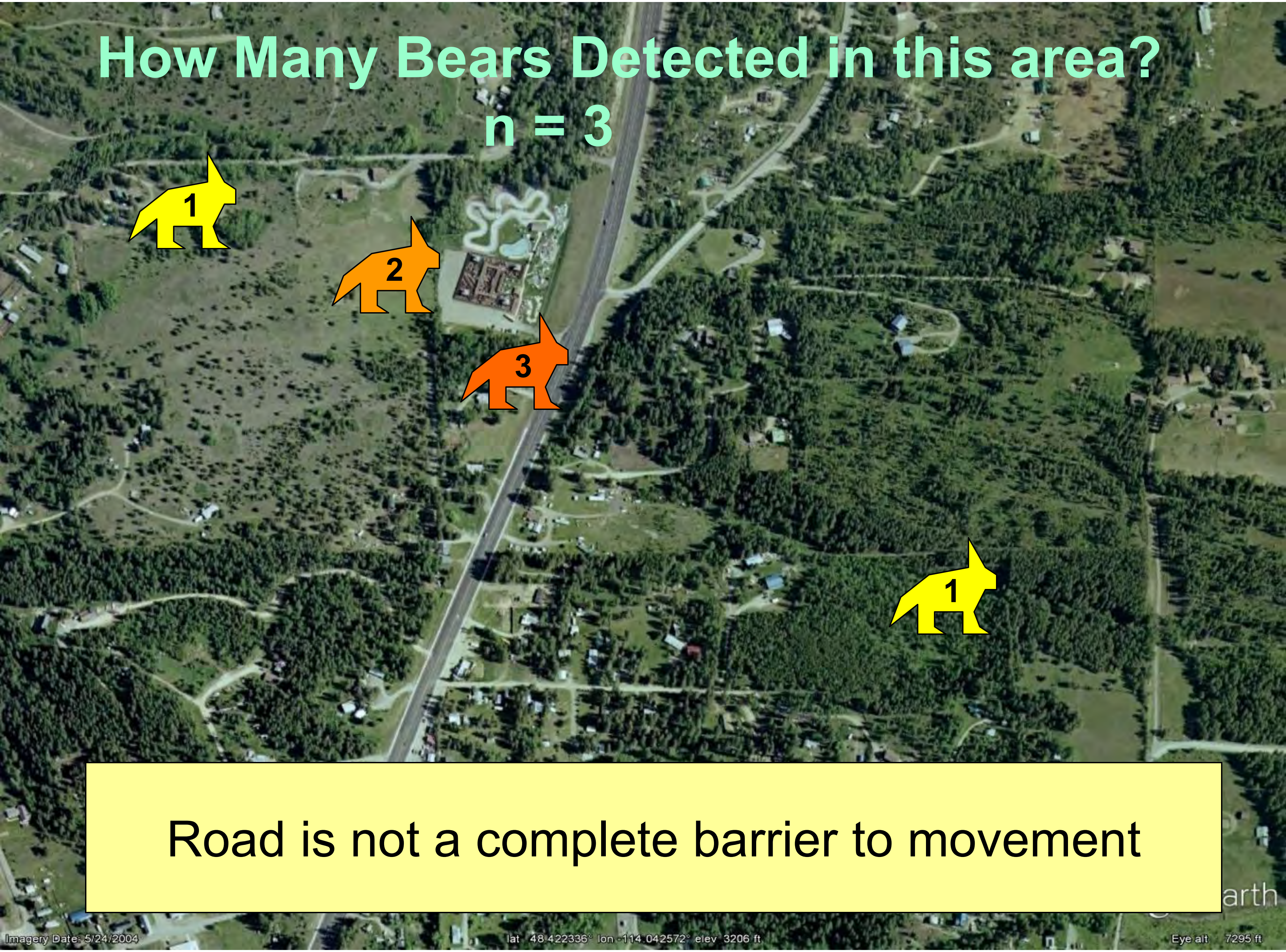
Individual 3

Important to use many loci



How Many Bears Detected in this area?

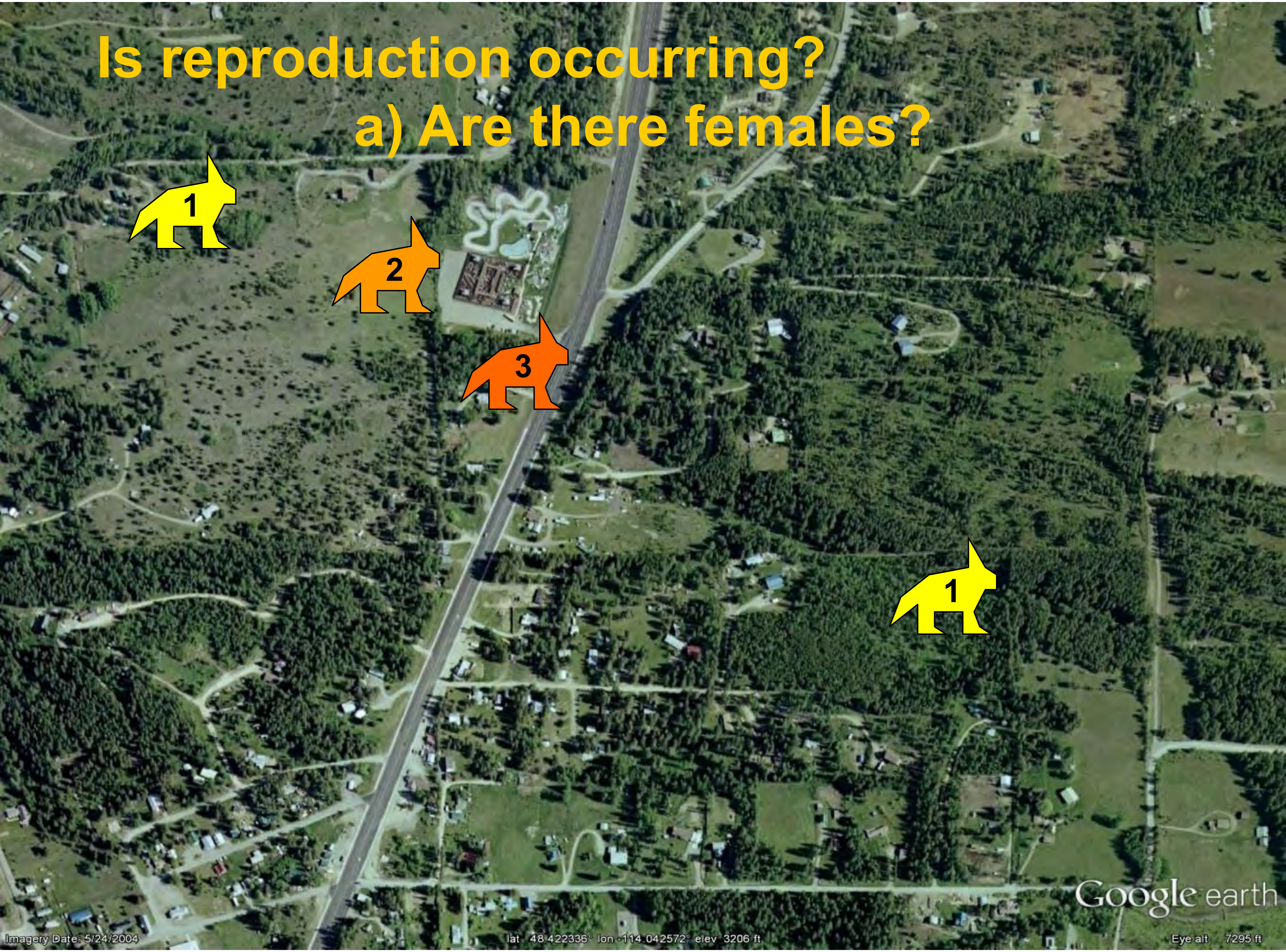
n = 3



Road is not a complete barrier to movement

Is reproduction occurring?

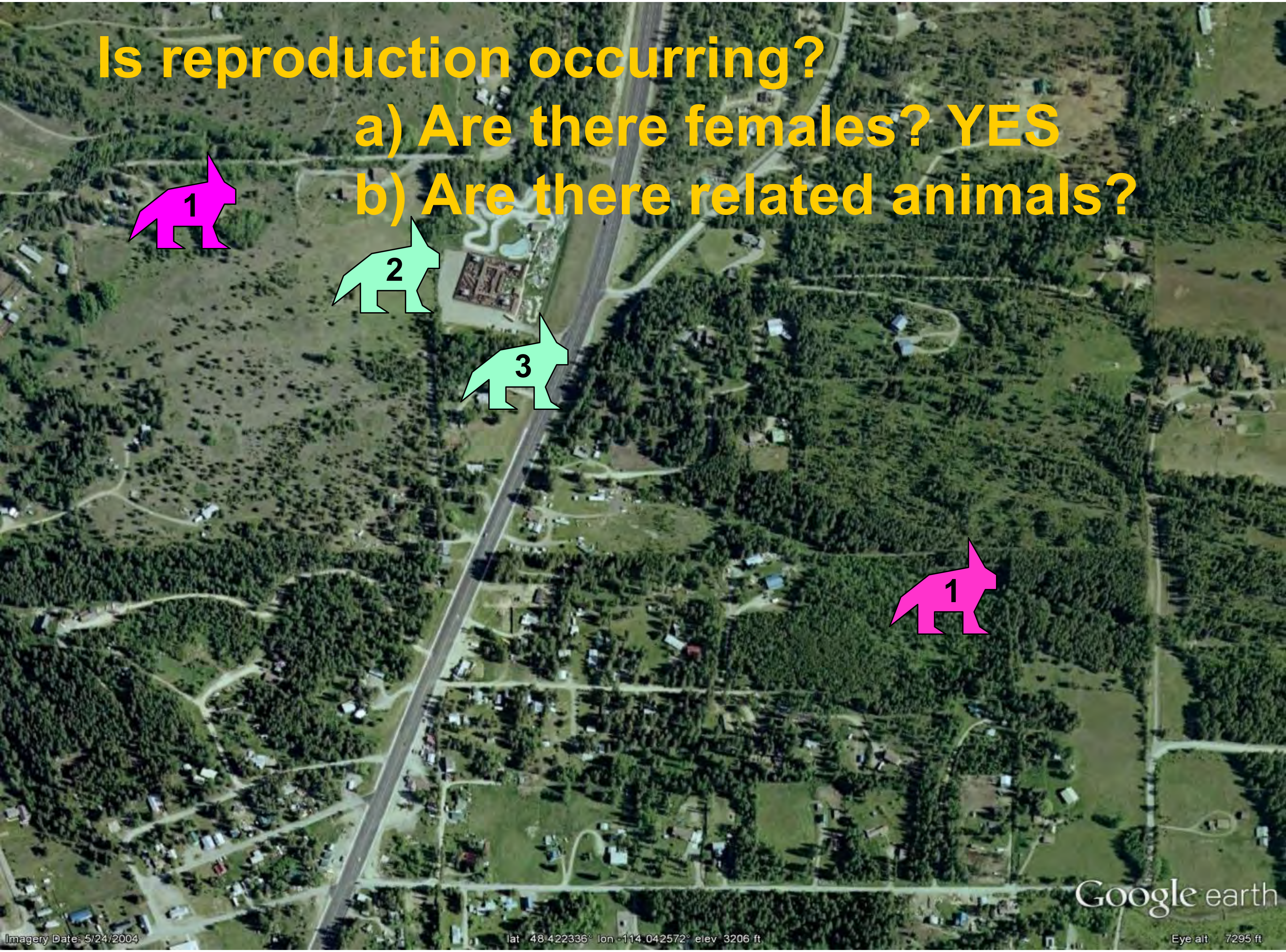
a) Are there females?



Sex Determination

- Females=XX
- Males=XY

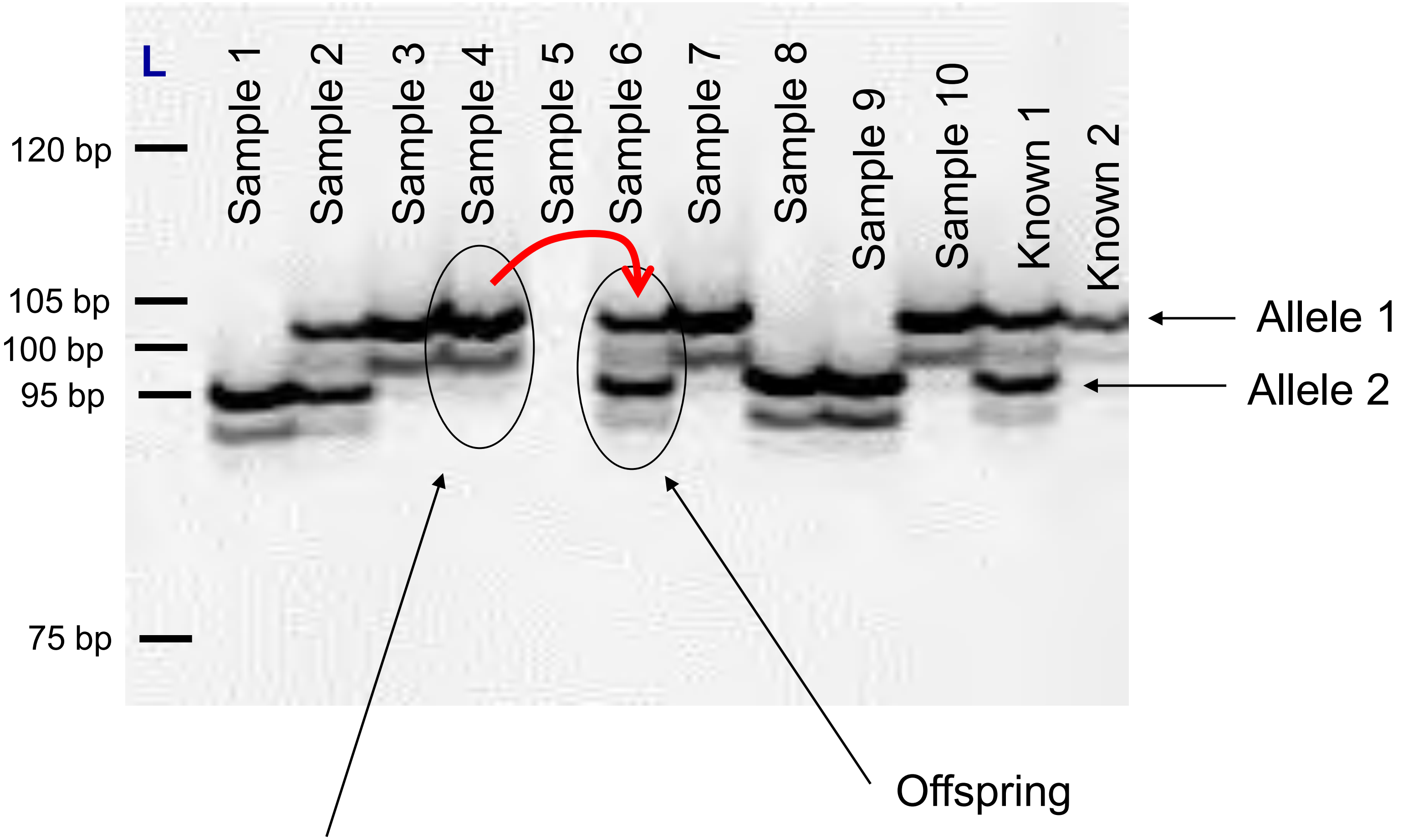
Zf, Amelogenin



Is reproduction occurring?

a) Are there females? YES

b) Are there related animals?



Allele 1

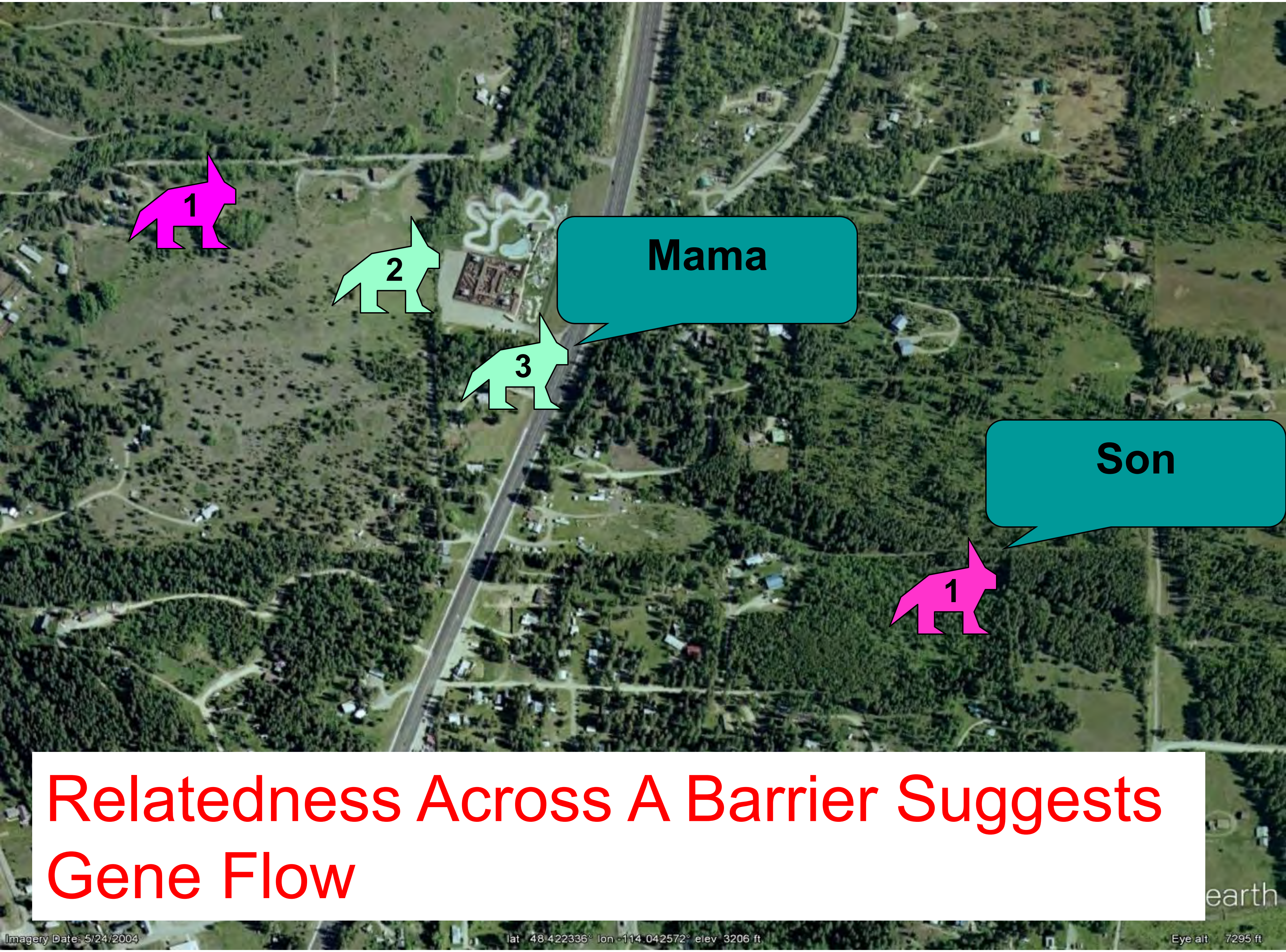
Allele 2

Offspring

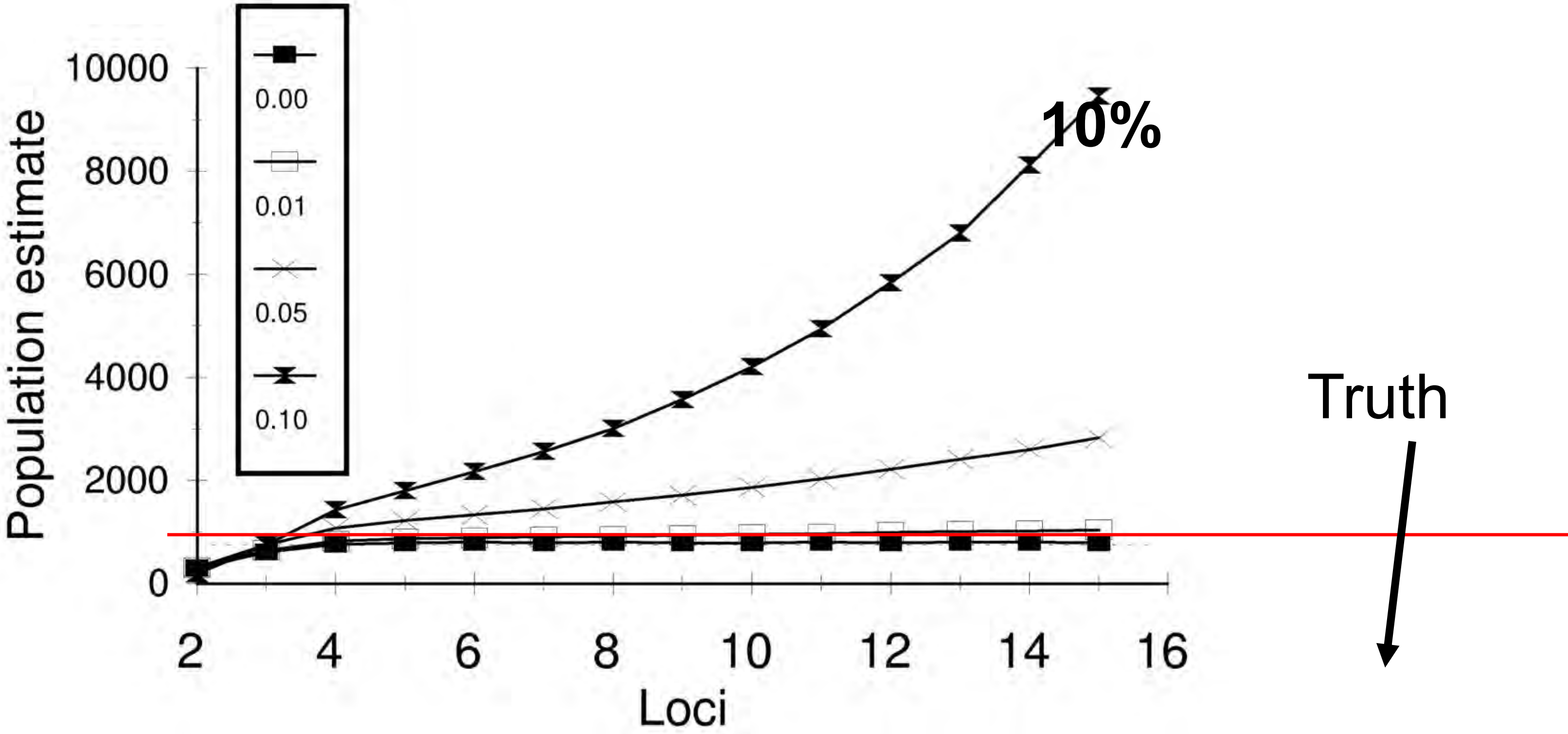
Mother

Who could be Dad?

Relatedness Across A Barrier Suggests Gene Flow



Errors Cause Overestimation of Abundance



True N = 800

Molecular Ecology (2003)

SHORT COMMUNICATION

Population size estimation in Yellowstone wolves with error-prone noninvasive microsatellite genotypes

Scott Creel, Goran Spong, Jennifer L. Sands, Jay Rotella, Janet Zeigle, Lawrence Joe, Kerry M. Murphy and Douglas Smith

Abstract

...Consequently, a substantial proportion of multilocus genotypes held one or more errors, despite multiple PCR. These genotyping errors created several genotypes per individual and caused overestimation (up to 5.5-fold) of population size....

5.5X Wolves

Evidence for over estimation of black bears due to genotyping errors

McKelvey and Schwartz 2004, 2004b J. Wild. Manage.
McKelvey and Schwartz 2005 Molec. Ecol. Notes.
Schwartz et al. 2006 Ursus
Cushman, Hayden, Schwartz, McKelvey 2006 Am Naturalist

28% Overestimation of Bears

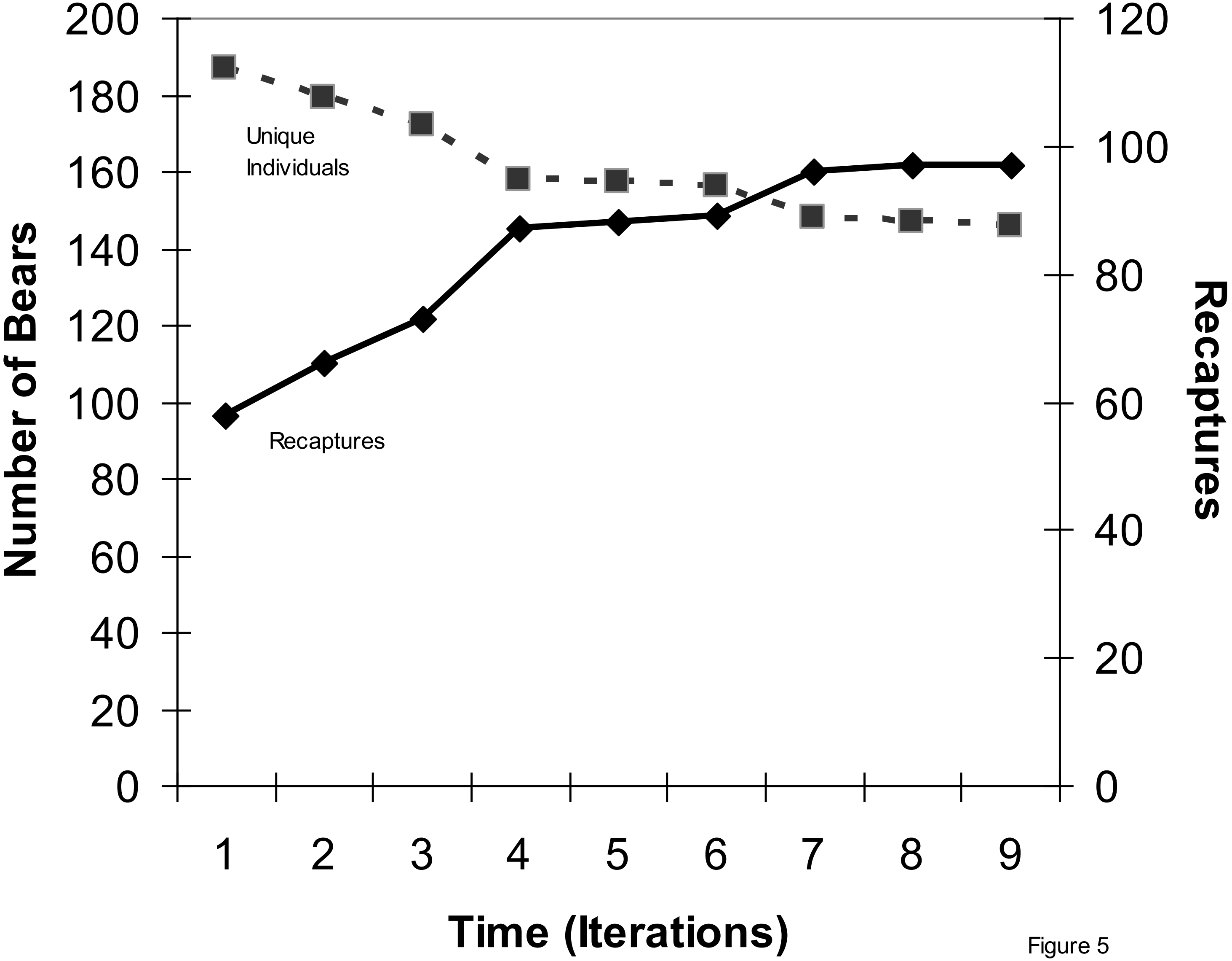
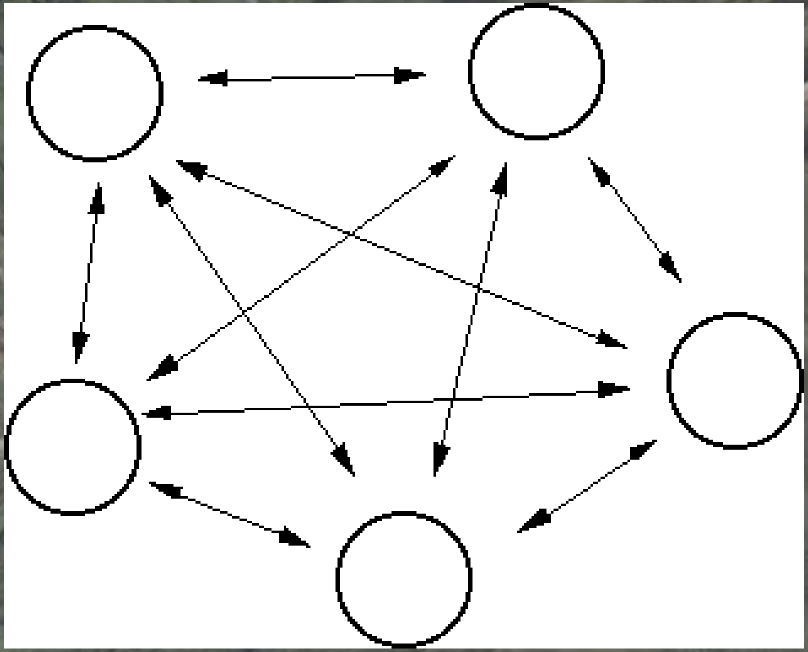


Figure 5

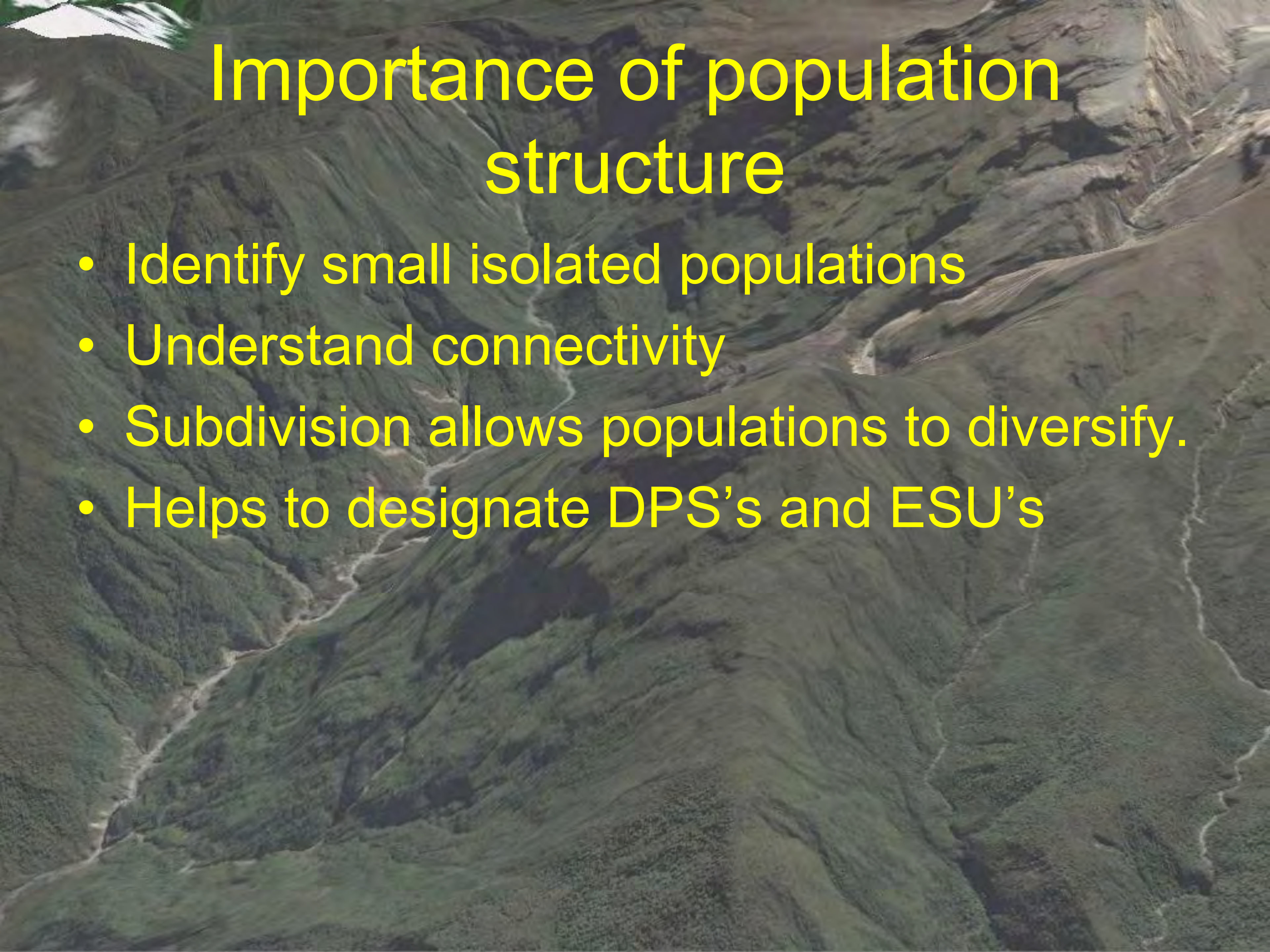
Describing Genetic Population Structure
Genetic Parameters



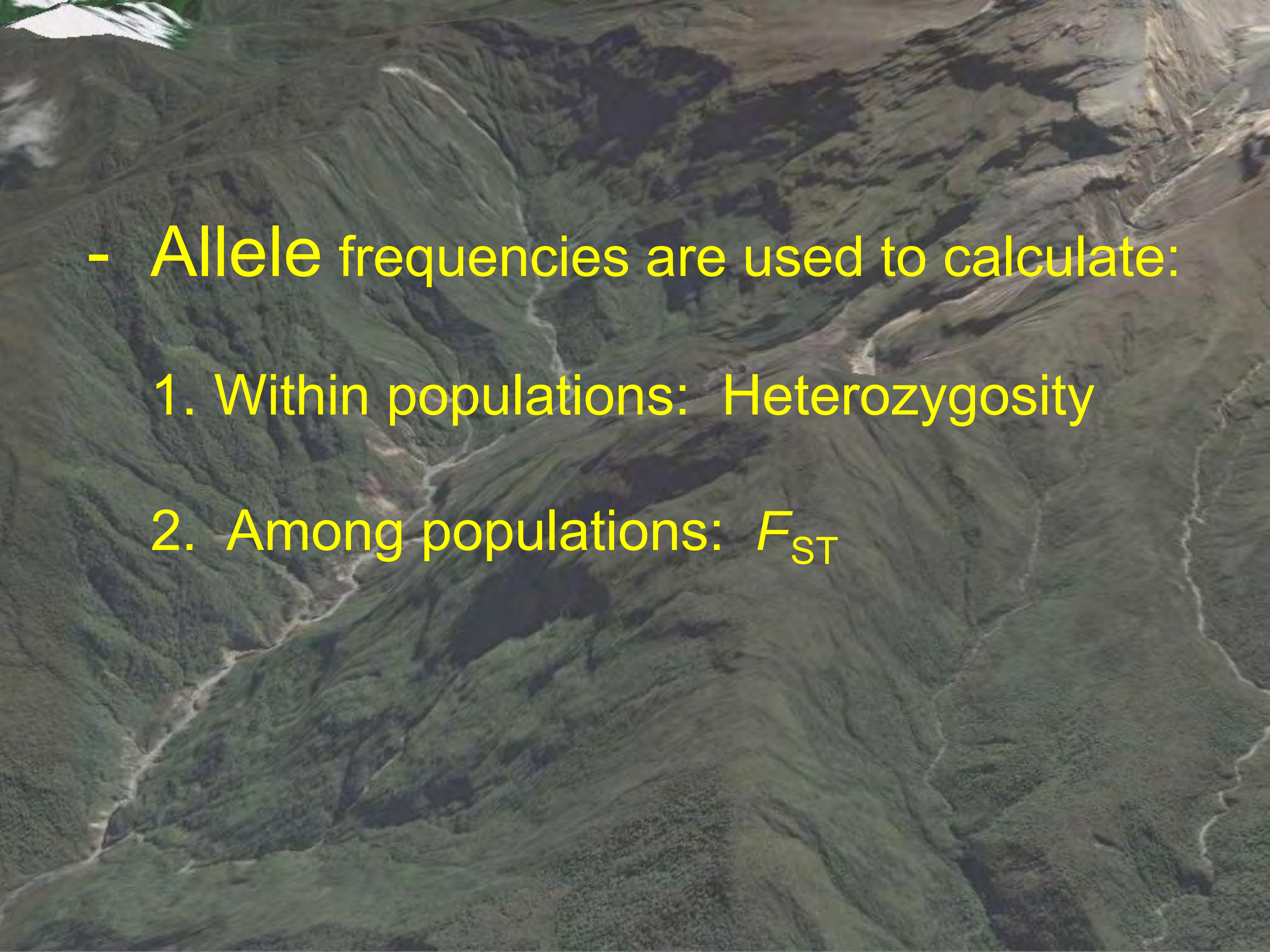
Populations exchanging more migrants will be more similar.

Small populations exchanging fewer migrants will diverge and lose variation through genetic drift: random sampling of alleles.

Importance of population structure



- Identify small isolated populations
- Understand connectivity
- Subdivision allows populations to diversify.
- Helps to designate DPS's and ESU's



- Allele frequencies are used to calculate:

1. Within populations: Heterozygosity
2. Among populations: F_{ST}

Allele Frequencies

12 total alleles

A =0.66

a =0.33

Heterozygosity

- Observed H_o and Expected H_e

H_O=0.33 H_e=0.44

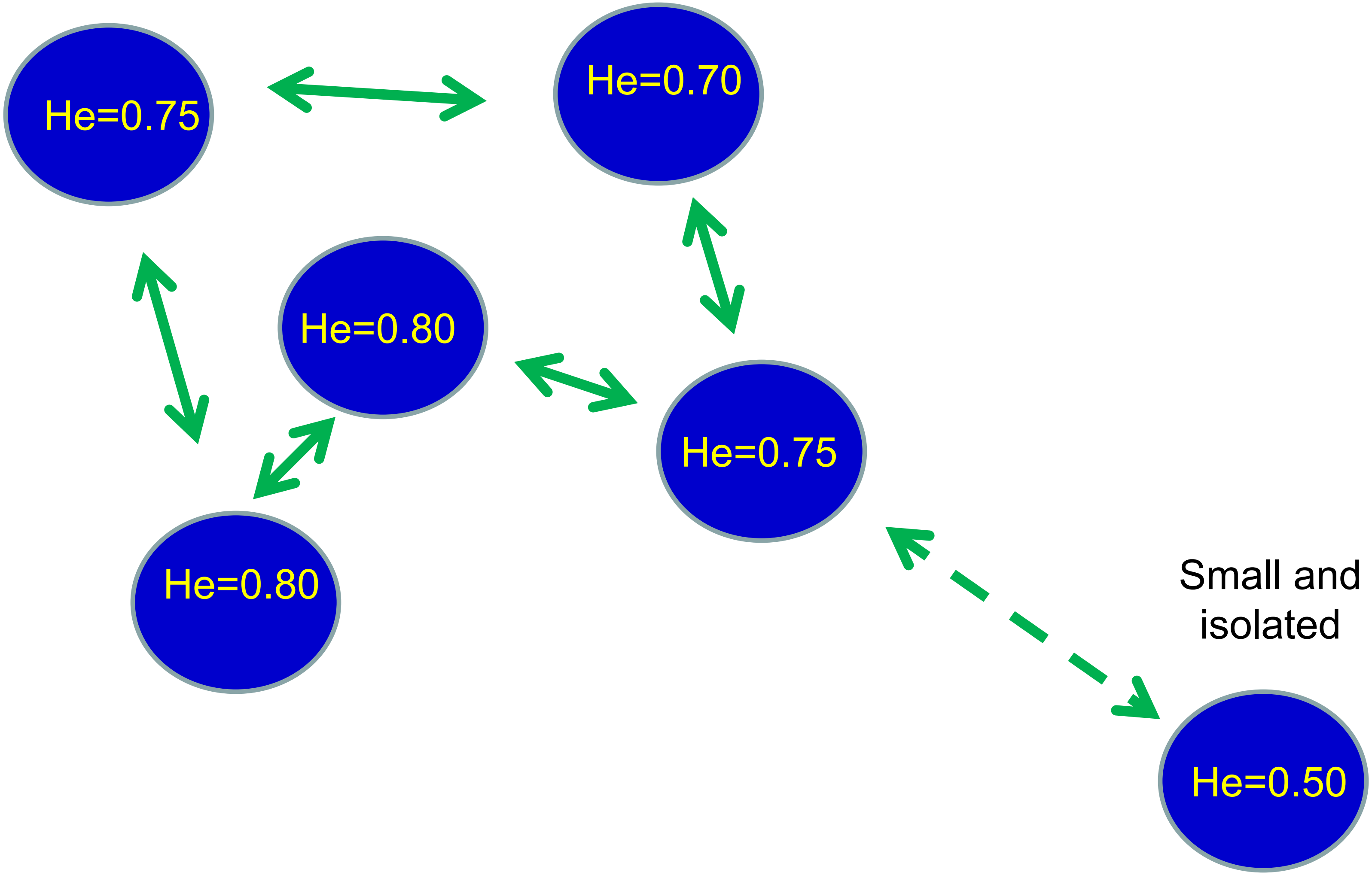
"Deficit of heterozygotes" might indicate that more than one population was sampled: Wahlund Effect.

H_e provides a comparable estimate of diversity that is directly dependent on allele frequencies

H_O=1.00

H_e=0.50

Amount of genetic diversity is relative



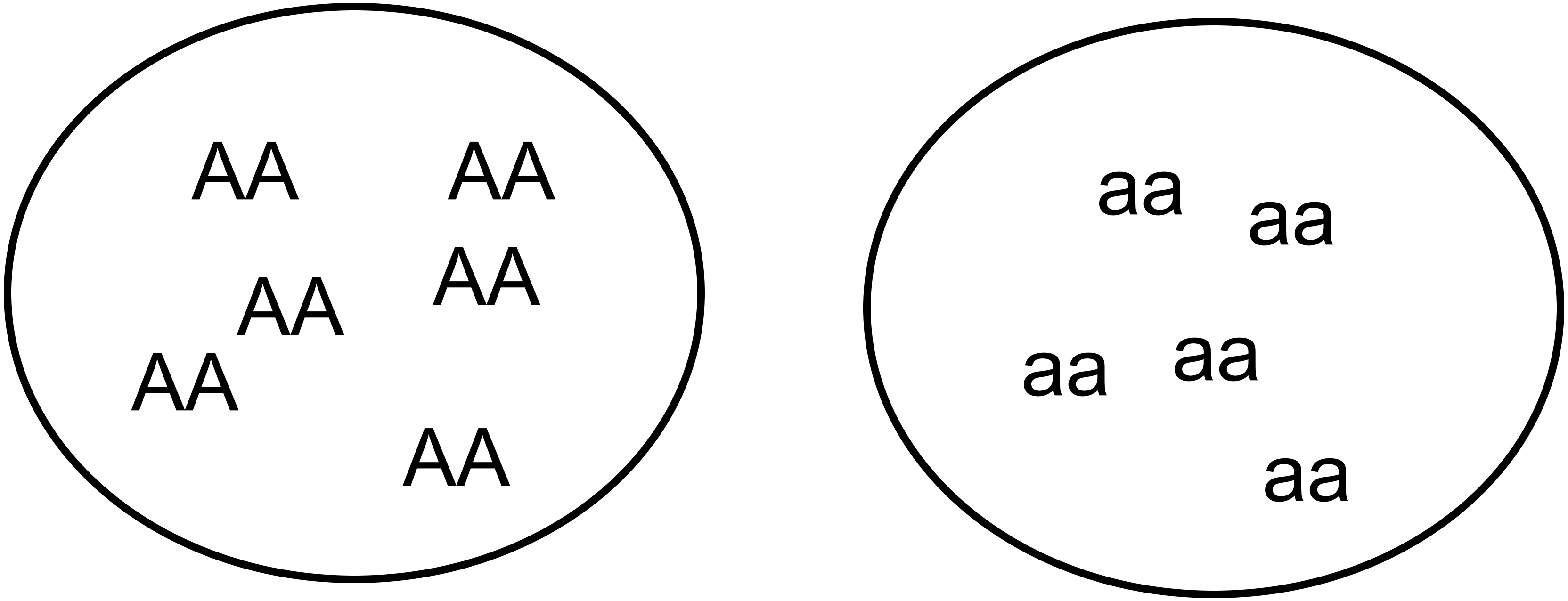
Small and
isolated

Population Structure

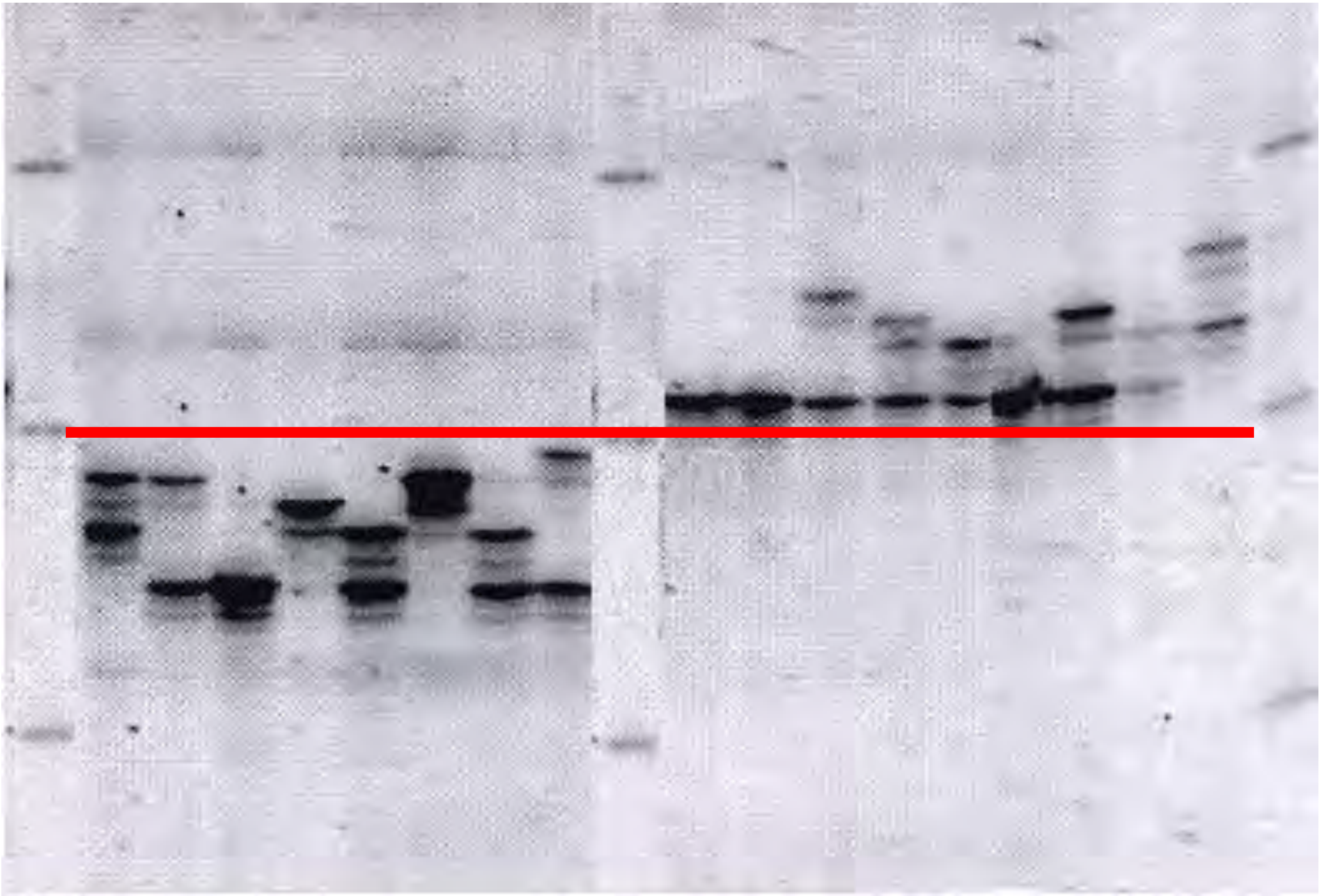
- F_{ST} is the proportional reduction in heterozygosity due to population subdivision: ranges from 0.0 to 1.0.

- High gene flow drives F_{ST} to 0.0.

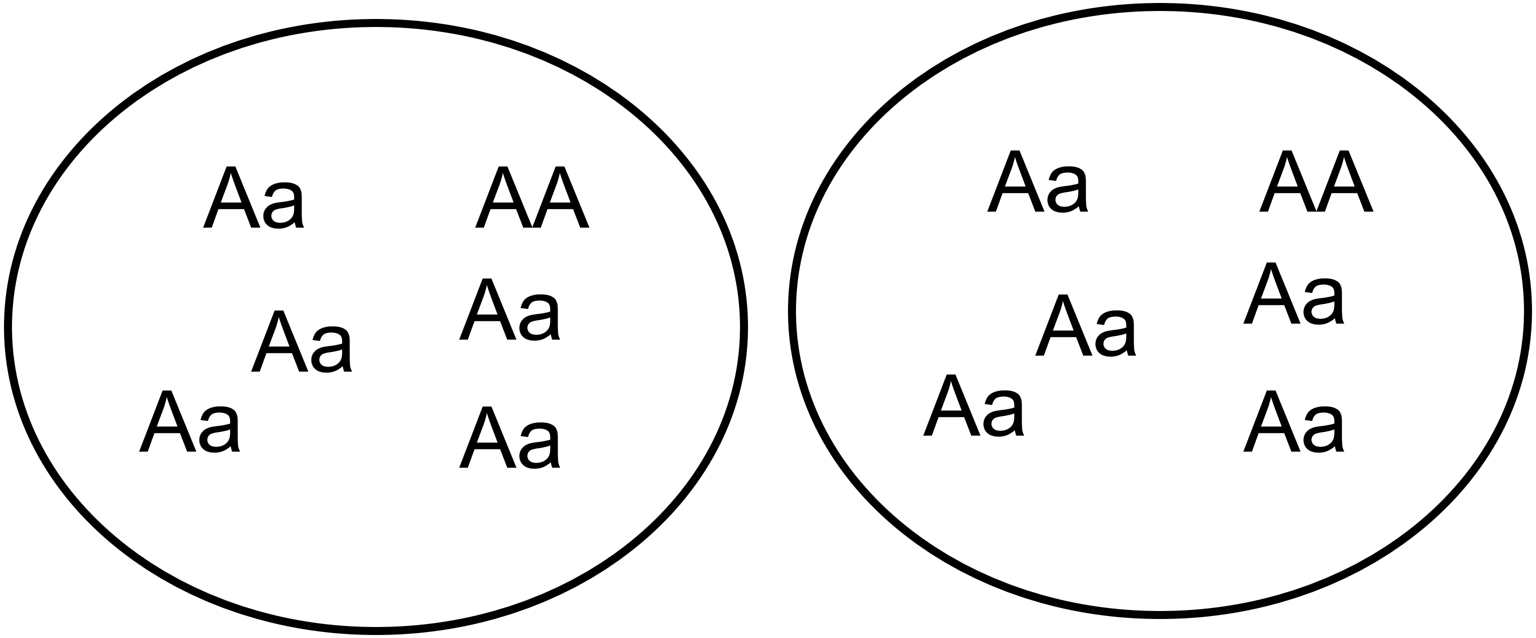
F_{ST} is the proportional reduction in heterozygosity due to population substructure



F_{ST}=1.00



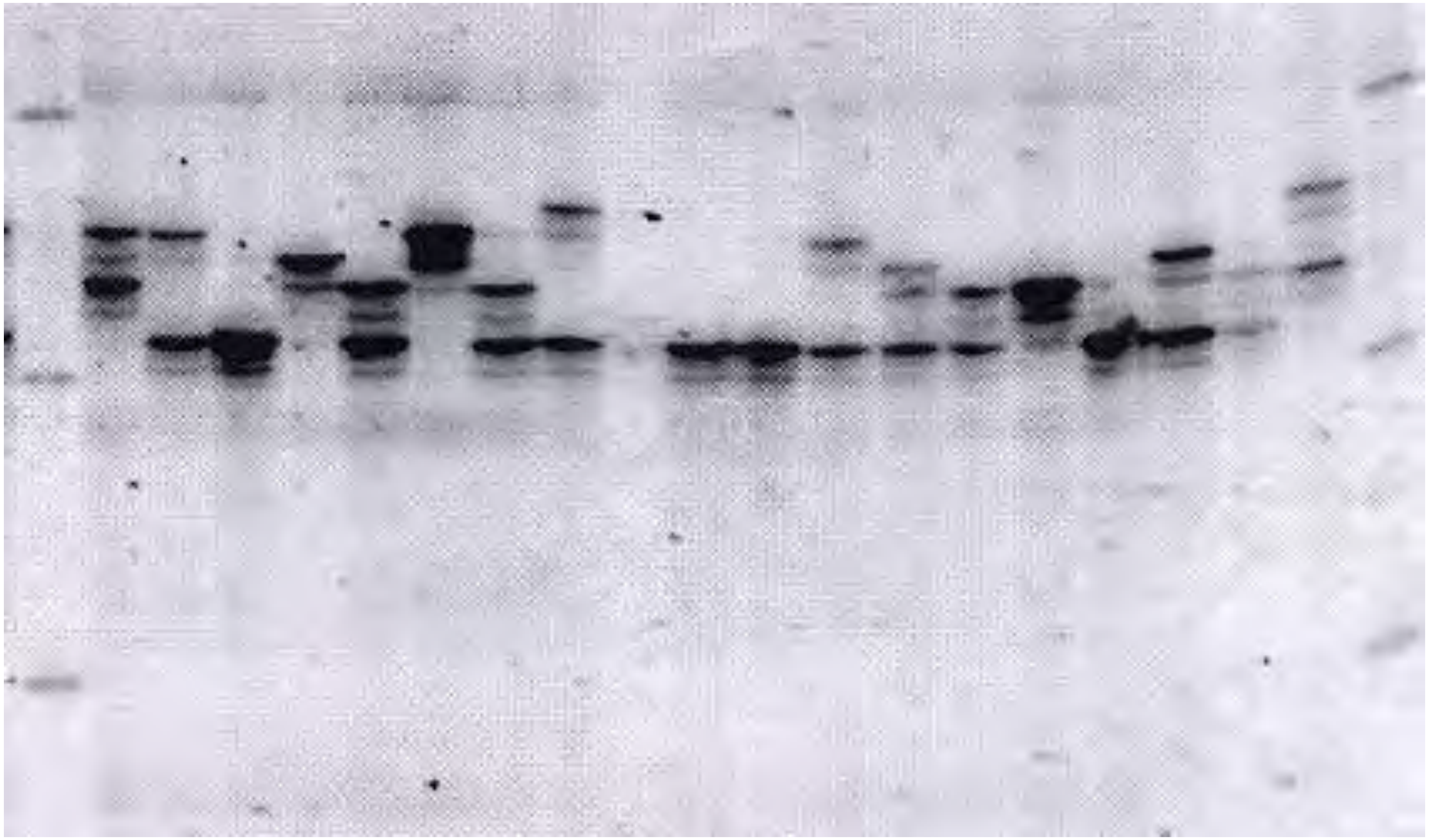
F_{ST} is the proportional reduction in heterozygosity due to population substructure



Aa AA
Aa Aa
Aa Aa

Aa AA
Aa Aa
Aa Aa

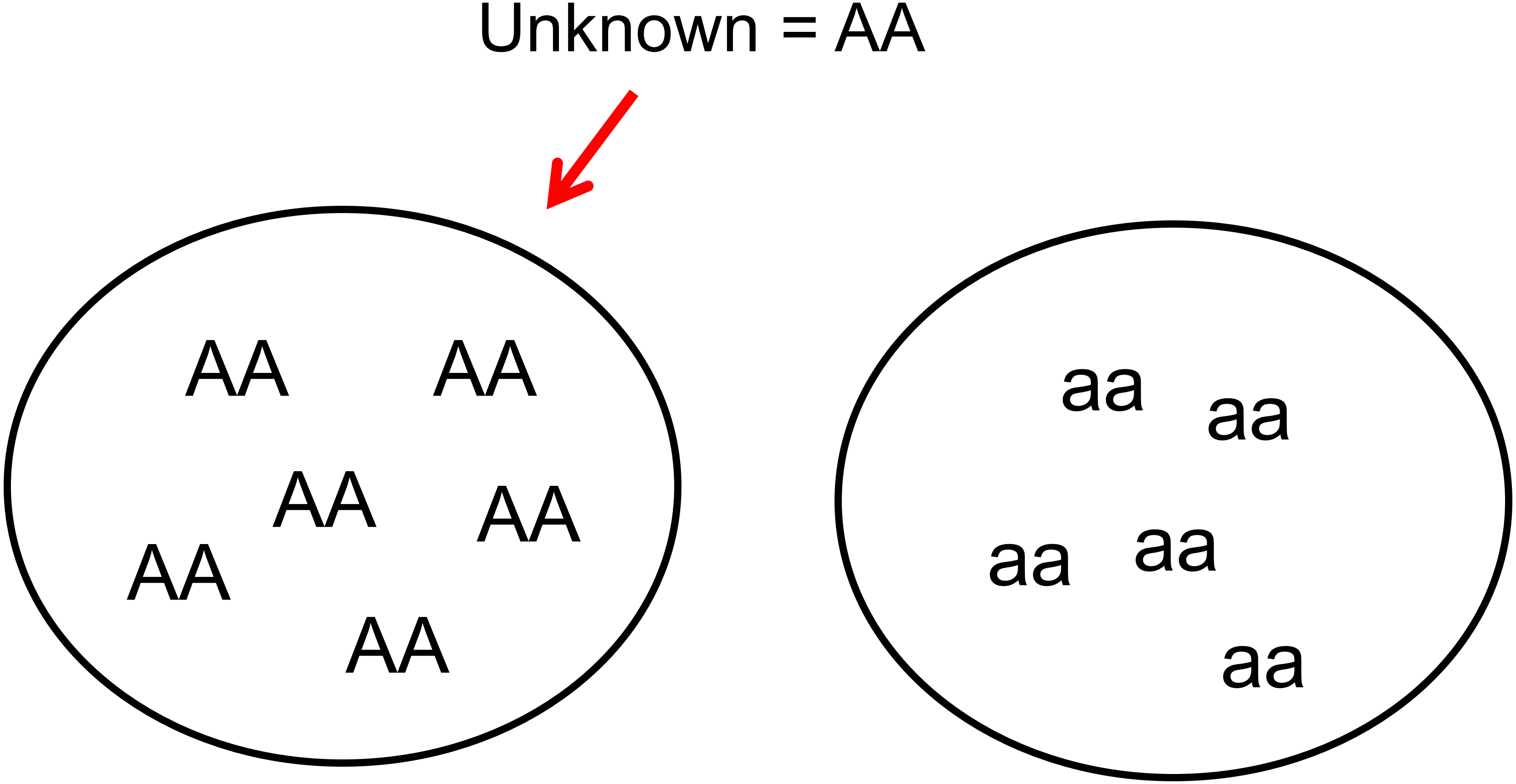
F_{ST}=0.00



Assignment Test

- Assigns individuals to a population based on likelihood of genotypic frequency

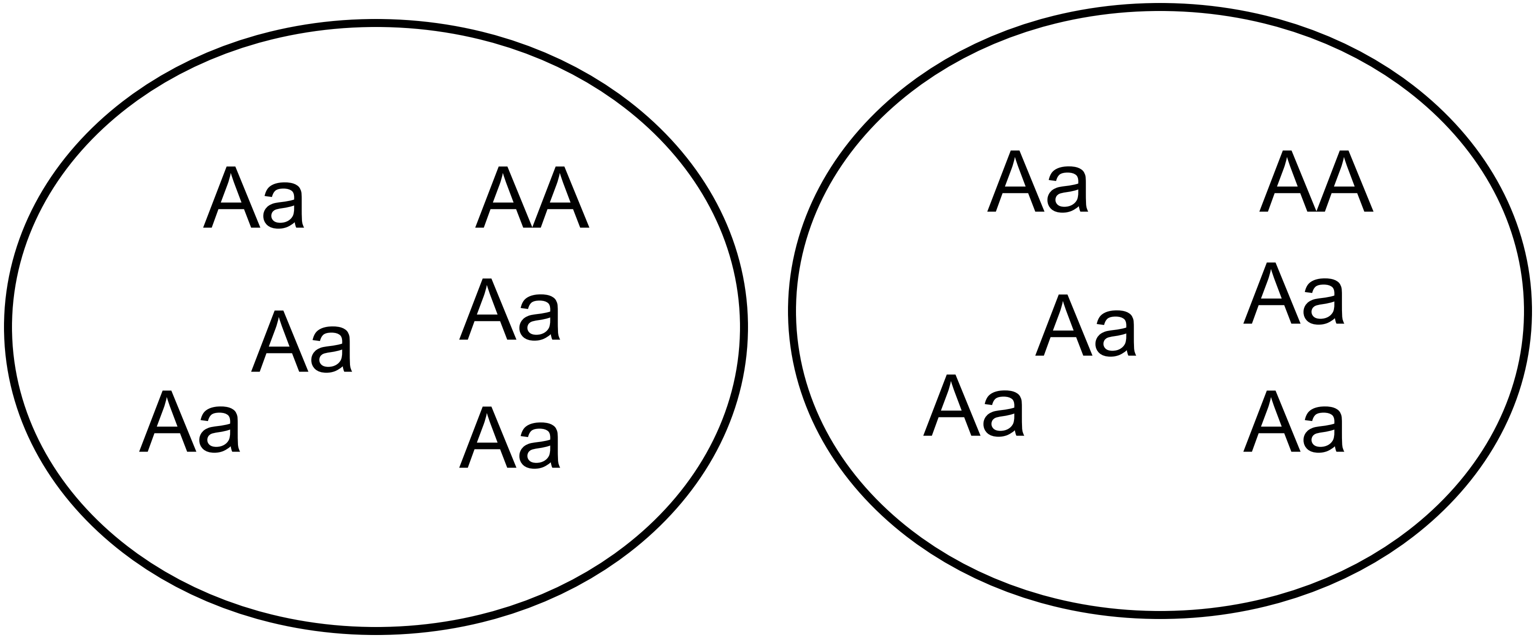
Assignment Test



F_{ST}=1.00

Assignment Test

Unknown = AA?



Aa AA

Aa Aa

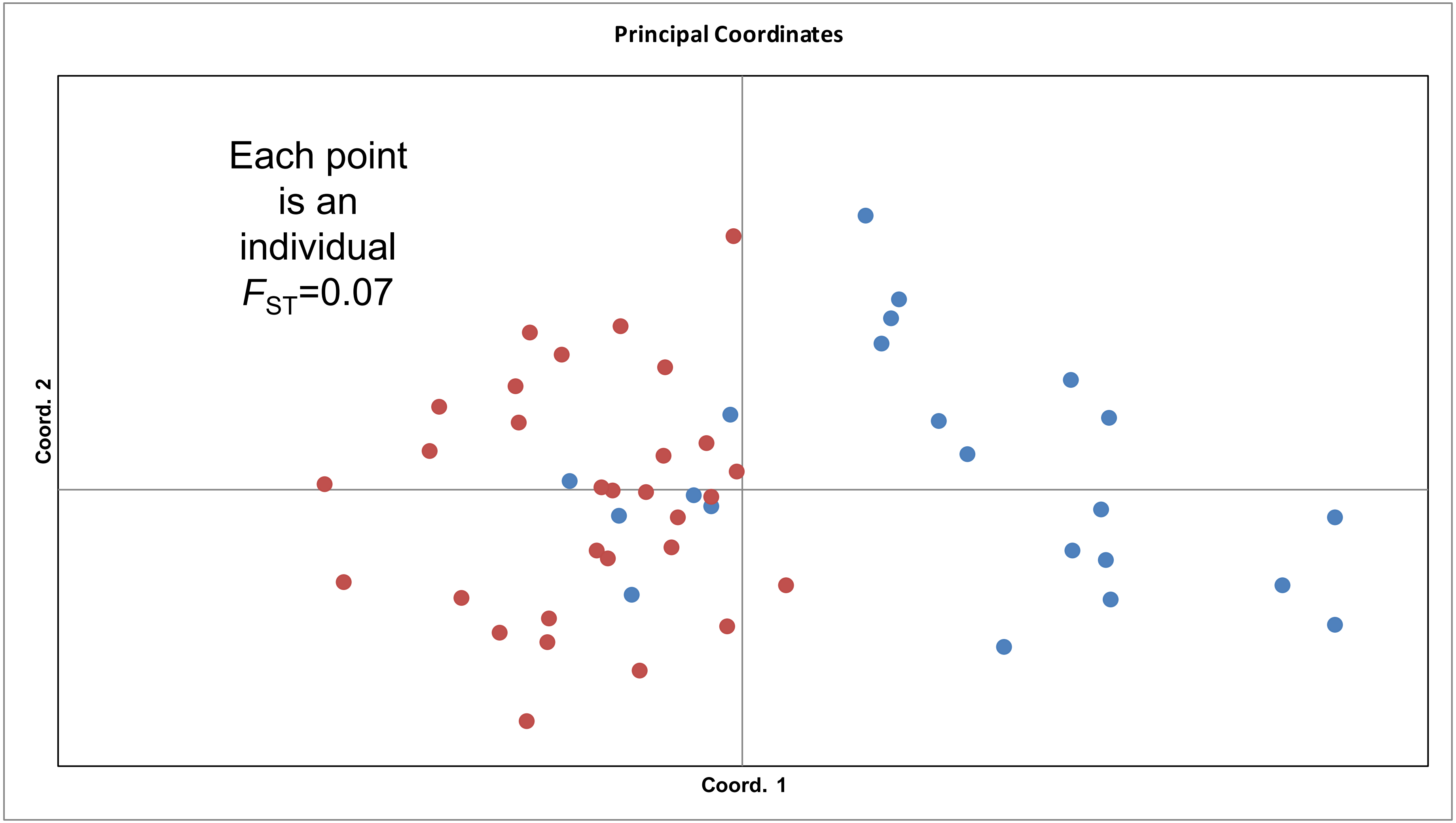
Aa Aa

Aa AA

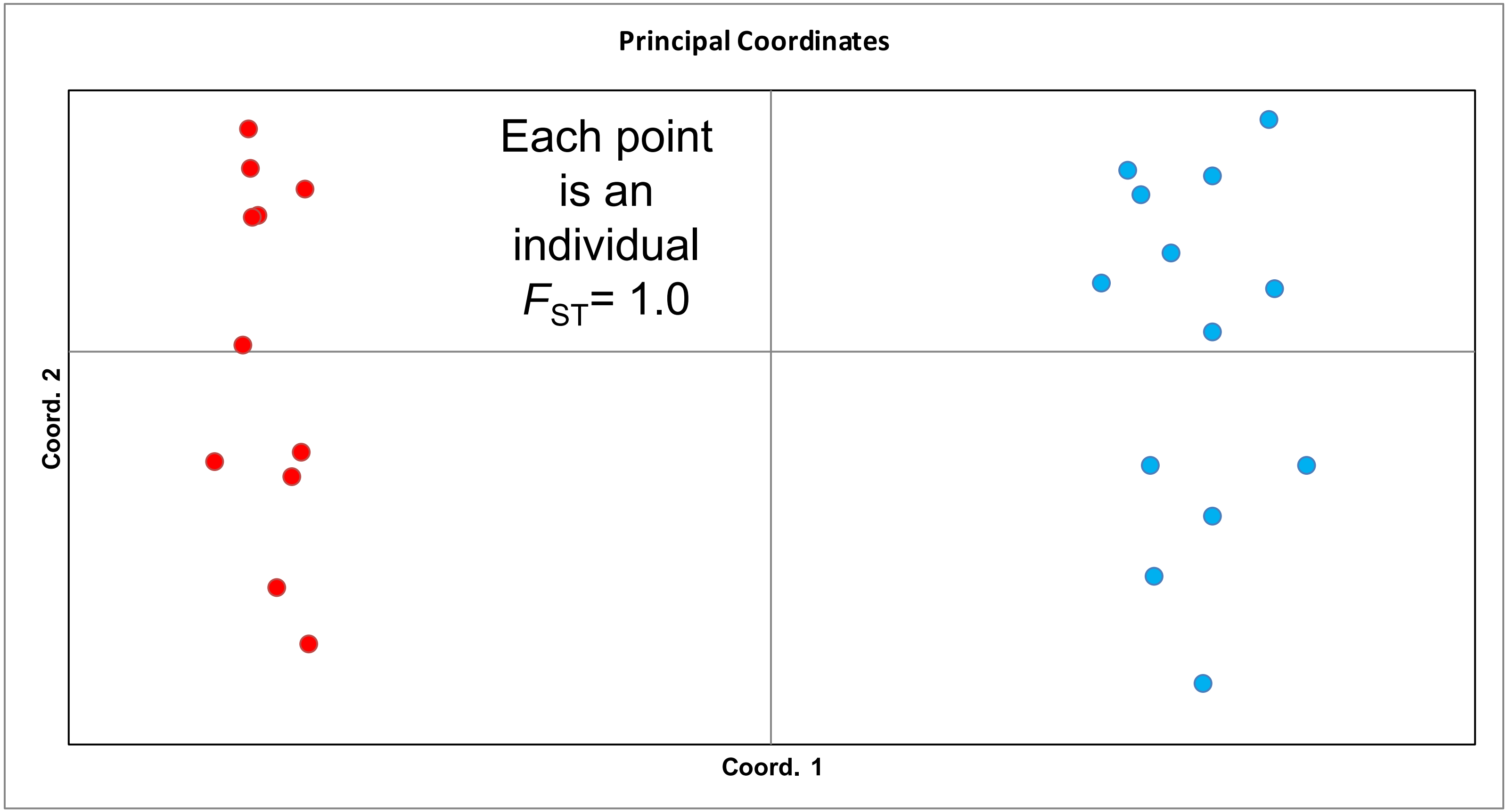
Aa Aa

Aa Aa

F_{ST}=0.00



Each point
is an
individual
F_{ST}=0.07



Each point is an individual

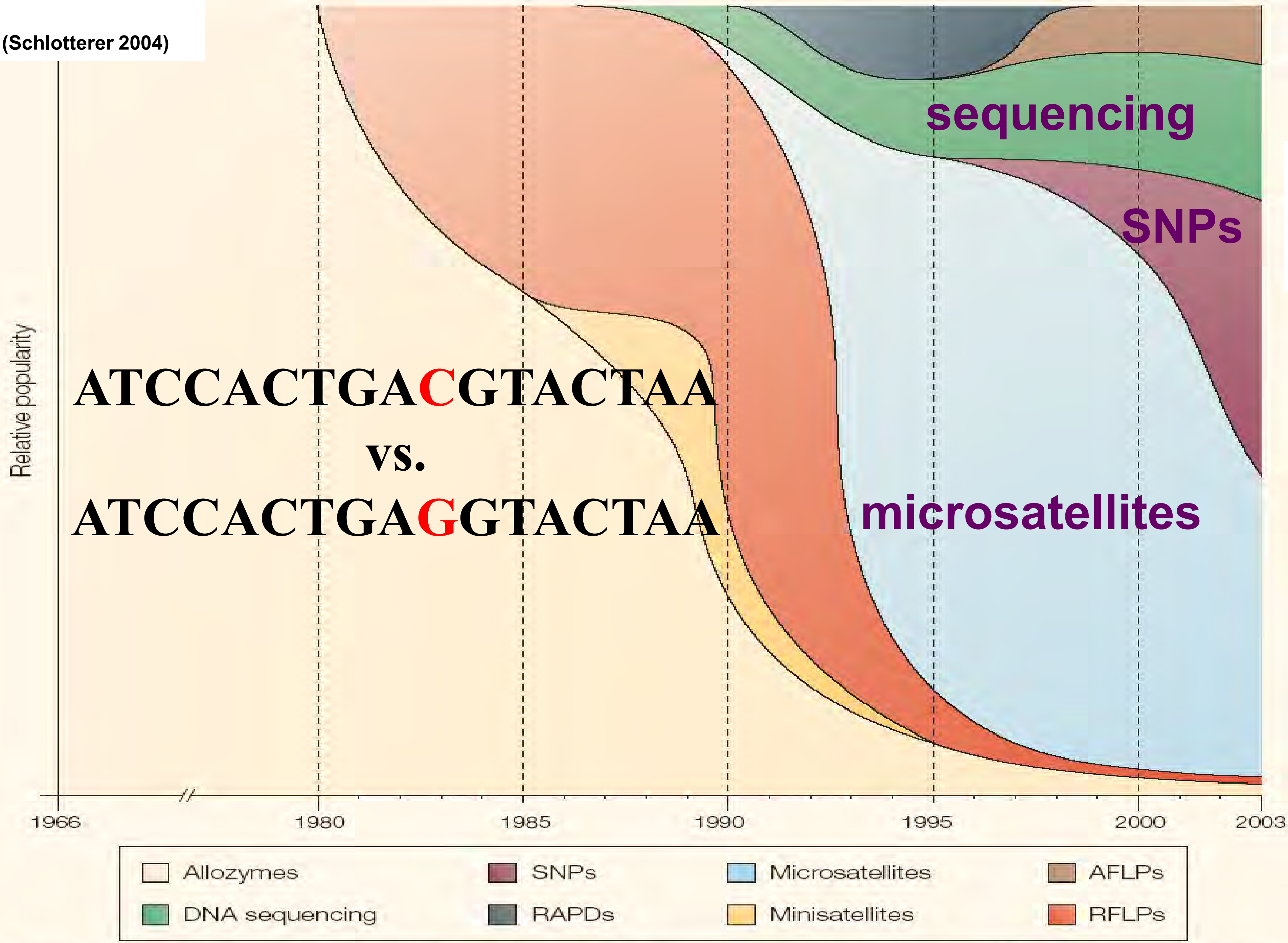
$F_{ST} = 1.0$

Future of Molecular Markers…….

The Era of Bioinformatics and Genomics

“In less than half a century, molecular markers have totally changed our view of nature….”

- Schlotterer (2004)



(Schlotterer 2004)

sequencing

SNPs

microsatellites

ATCCACTGACGTACTAA

vs.

ATCCACTGAGGTACTAA

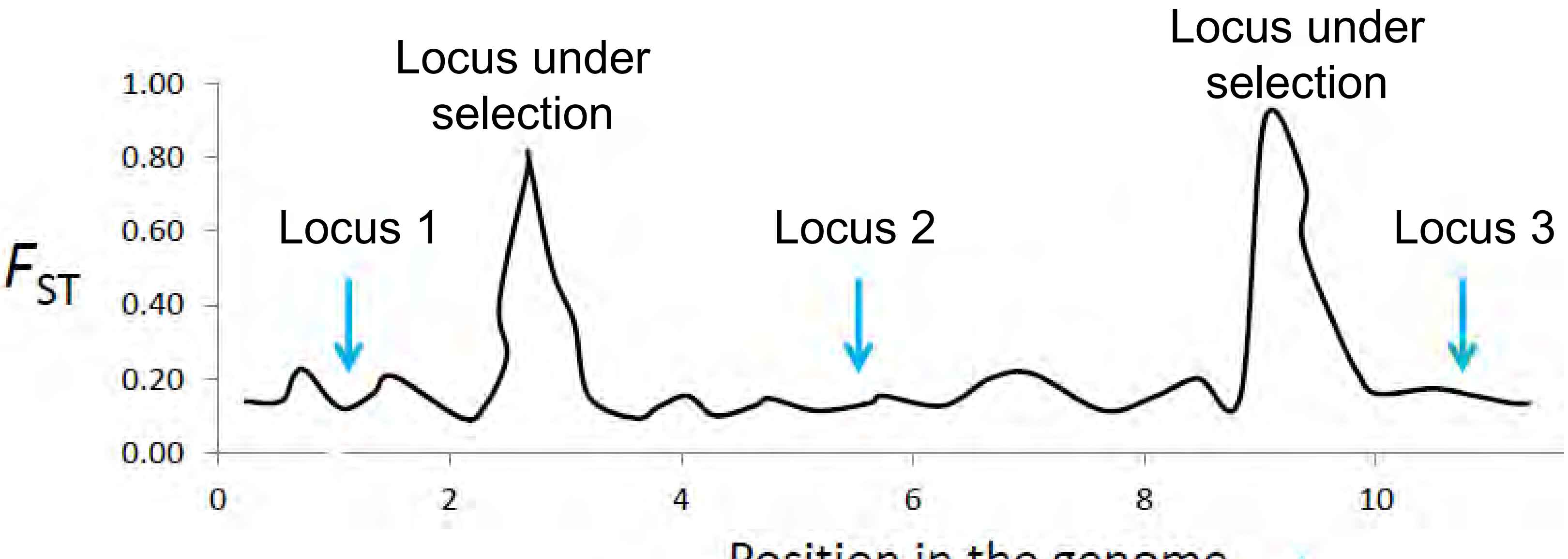
Figure 1 | **Subjective view of the the changing relative importance of different molecular markers.**

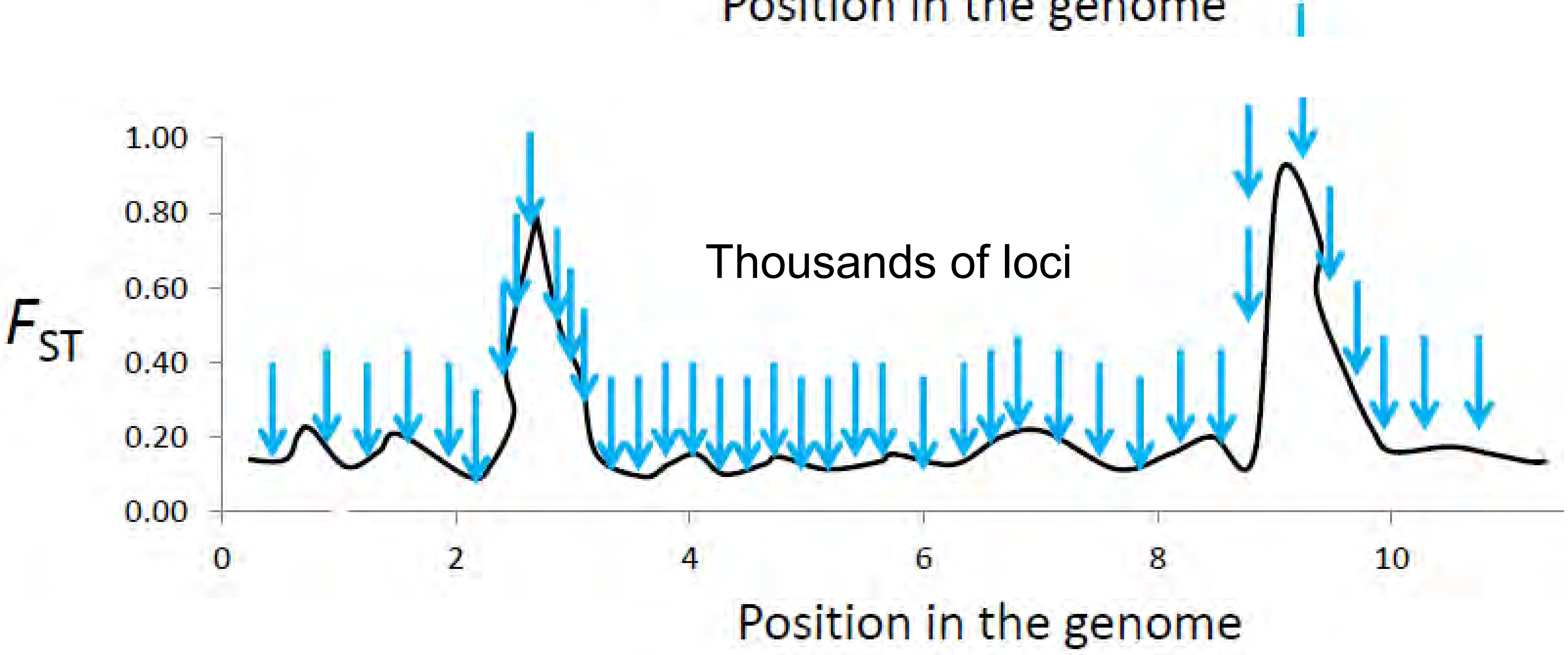
Conservation Genetics and the "Information Explosion"

- Historically: 10's of neutral markers to survey neutral variation
  - Provides a basic understanding of genetic variation influenced by drift and gene flow: not selection

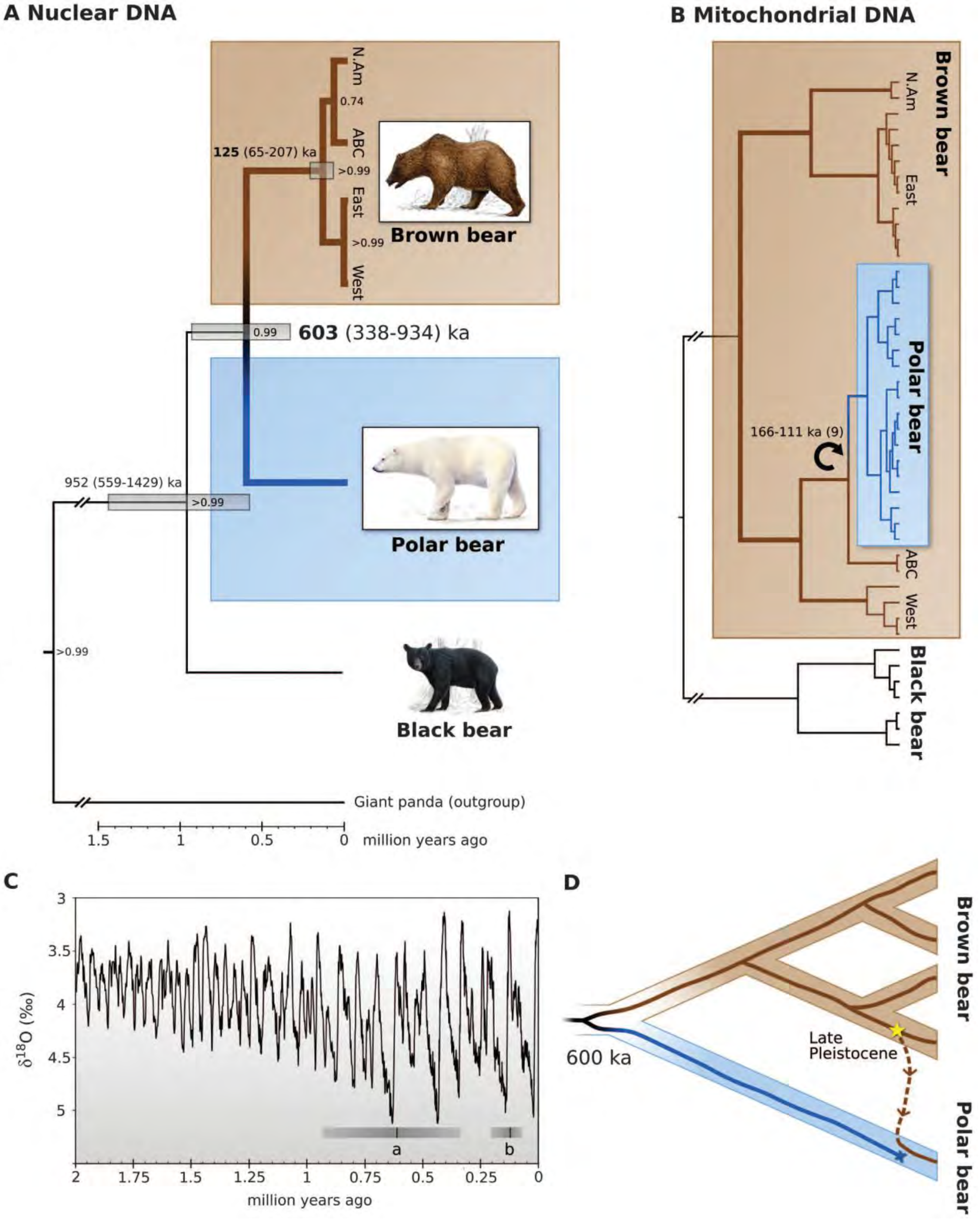
- Future: 1,000's of markers to survey functional variation
  - Identify populations that warrant protection due to adaptive differentiation





Locus under selection
Locus under selection
Locus 1
Locus 2
Locus 3
F_{ST}
1.00
0.80
0.60
0.40
0.20
0.00
0
2
4
6
8
10
Position in the genome


Thousands of loci
F_{ST}
1.00
0.80
0.60
0.40
0.20
0.00
0
2
4
6
8
10
Position in the genome

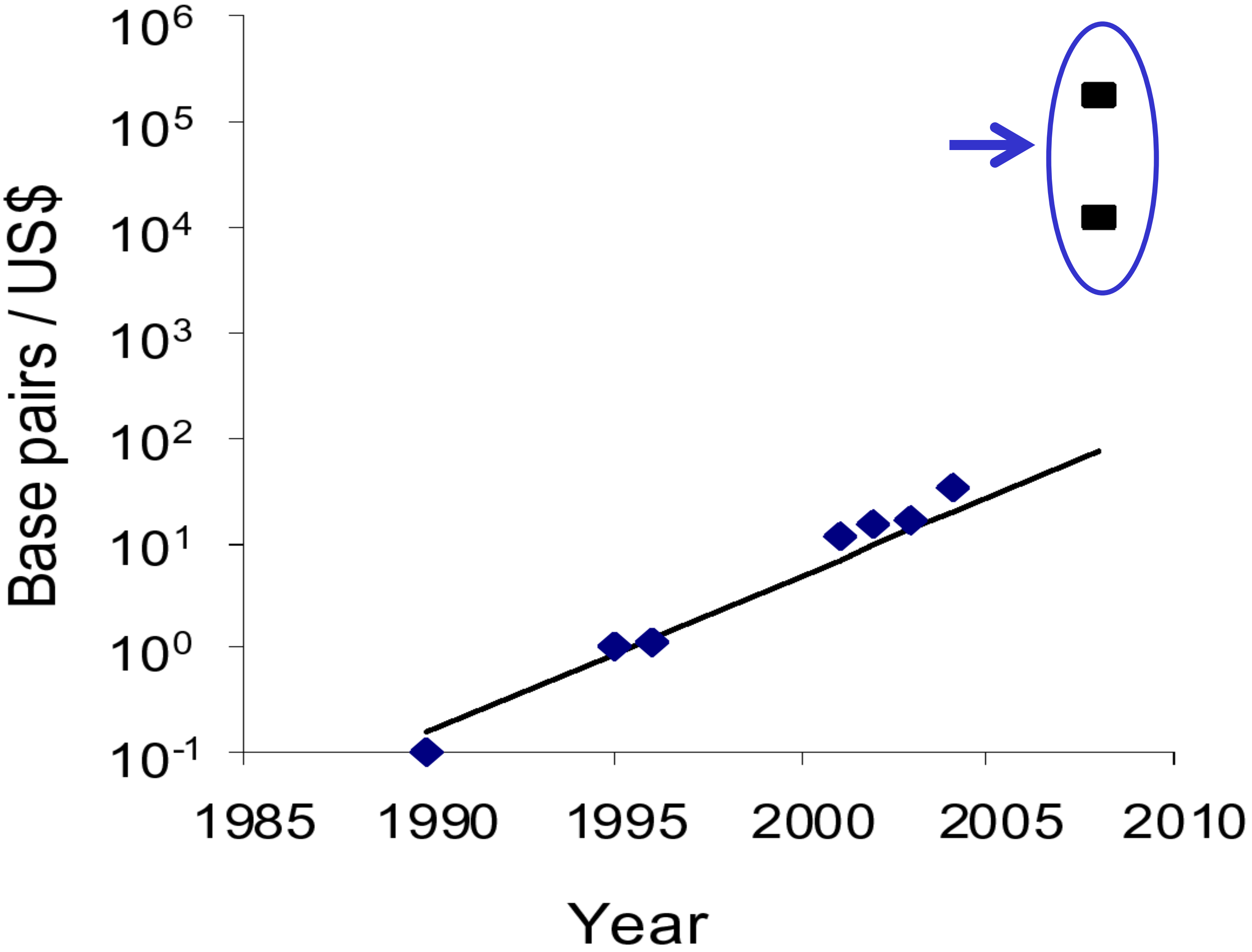
More accurate patterns of neutral divergence



Nuclear Genomic Sequences Reveal that Polar Bears Are an Old and Distinct Bear Lineage

Frank Hailer,[1*] Verena E. Kutschera,[1] Björn M. Hallström,[1] Denise Klassert,[1] Steven R. Fain,[2] Jennifer A. Leonard,[3] Ulfur Arnason,[4] Axel Janke[1,5*]

Increase in base pairs sequenced for $1 US



Hauser and Seeb, 2009; Mardis 2008

Questions?

- Thanks to IDFG!