

CONSERVATION GENOMICS



Emiliano Trucchi
Marche Polytechnic University

**WHAT IS THE DIFFERENCE BETWEEN
CONSERVATION GENETICS AND GENOMICS?**

**WHAT IS THE FIRST INGREDIENT OF
CONSERVATION GENOMICS?**

**HOW CAN GENOMICS INFORM
CONSERVATION BIOLOGY?**

**IN WHICH APPLICATION GENOMICS IS
A GAME CHANGER?**

with a close up on two topics:

Genetic load

ROHs

**WHAT AM I DOING TO MAKE USE OF
GENOMICS IN CONSERVATION?**

**WHAT IS THE DIFFERENCE BETWEEN
CONSERVATION GENETICS AND GENOMICS?**

WHAT IS THE FIRST INGREDIENT OF
CONSERVATION GENOMICS?

HOW CAN GENOMICS INFORM
CONSERVATION BIOLOGY?

IN WHICH APPLICATION GENOMICS IS
A GAME CHANGER?

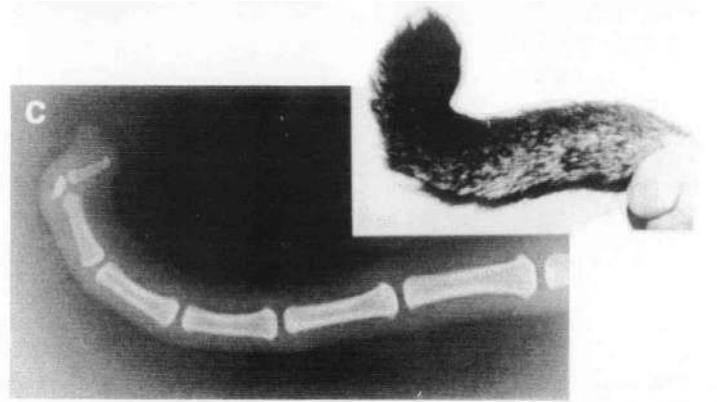
with a close up on two topics:

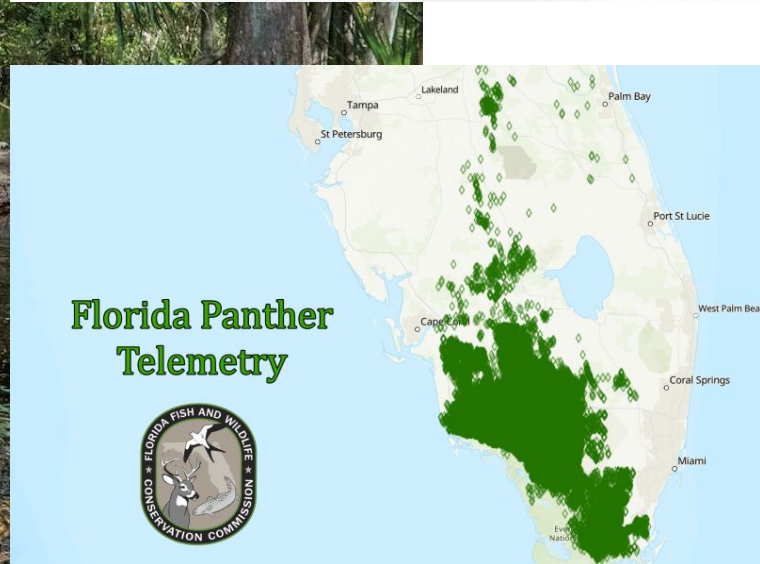
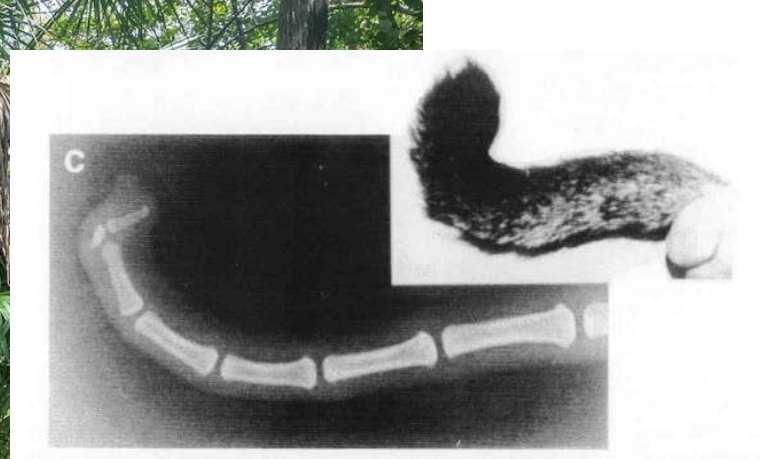
Genetic load

ROHs

WHAT AM I DOING TO MAKE USE OF
GENOMICS IN CONSERVATION?

WHOSE TAIL IS THIS?



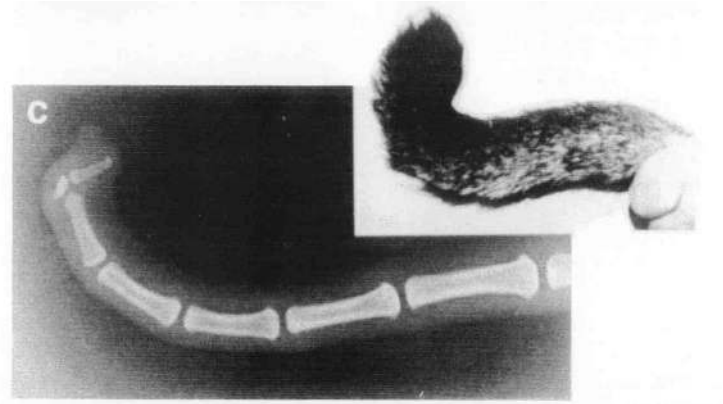


THE FLORIDA PANTHER

from genetics to genomics

Endangered since late '60 - early '70

Small relic population of **30 individuals** in
Big Cypress Swamp and Everglades NP



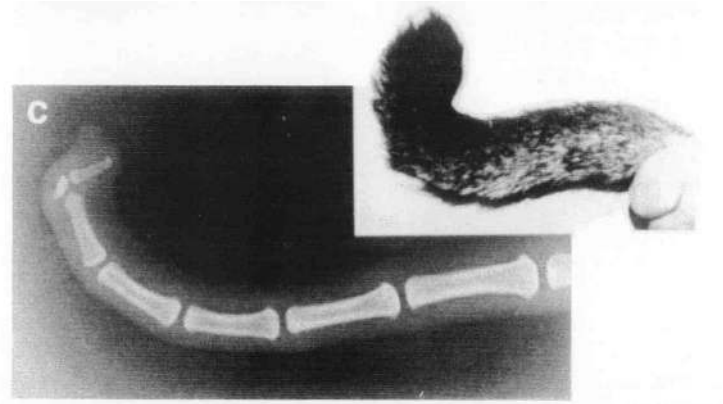
THE FLORIDA PANTHER

from genetics to genomics

Endangered since late '60 - early '70

Small relic population of 30 individuals in Big Cypress Swamp and Everglades NP

One mitochondrial haplotype, very low allozyme diversity, remarkably **low minisatellite DNA diversity** (as low as the Asian lions from the Gir Forest Sanctuary)



THE ASIATIC LION

Gir Forest Sanctuary

Year	Count	Male:Female:Cubs
1968	177	-
1974	180	-
1979	261	76:100:100
1984	252	88:100:64
1990	249	82:100:67
1995	265	94:100:71
2000	327	99:115:76
2005	359	-
2010	411	97:162:152
2015	523	109:201:213
2020	674	277:260:137



THE FLORIDA PANTHER

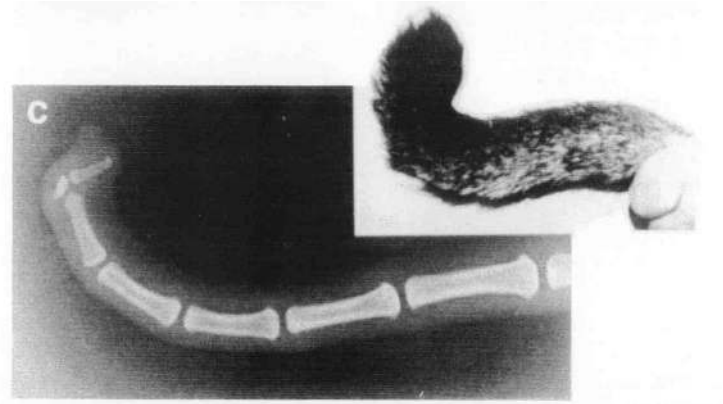
from genetics to genomics

Endangered since late '60 - early '70

Small relic population of 30 individuals in Big Cypress Swamp and Everglades NP

One mitochondrial haplotype, very low allozyme diversity, remarkably low minisatellite DNA diversity (as low as the Asian lions from the Gir Forest Sanctuary)

Clear signs of **inbreeding depression**: kinked tails, cowlick, poor seminal quality, cryptorchidism, cardiac defects



THE FLORIDA PANTHER

from genetics to genomics

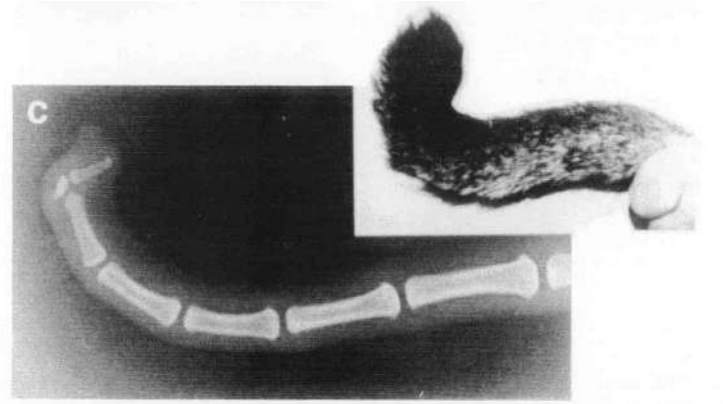
Endangered since late '60 - early '70

Small relic population of 30 individuals in Big Cypress Swamp and Everglades NP

One mitochondrial haplotype, very low allozyme diversity, remarkably low minisatellite DNA diversity (as low as the Asian lions from the Gir Forest Sanctuary)

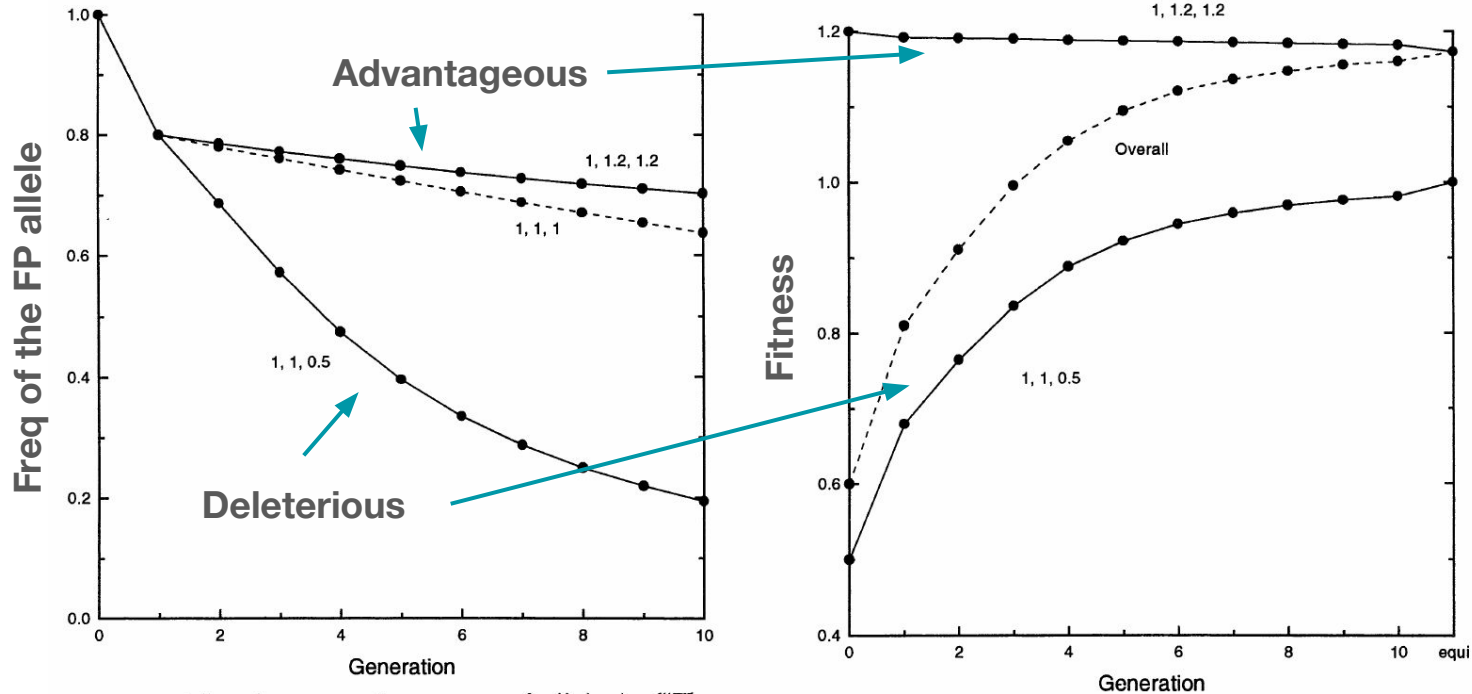
Clear signs of inbreeding depression: kinked tails, cowlick, poor seminal quality, cryptorchidism, cardiac defects

RECOMMENDATION: immediate augmentation of the population with Texas pumas! (Roelke et al 1993)



THE FLORIDA PANTHER

from genetics to genomics

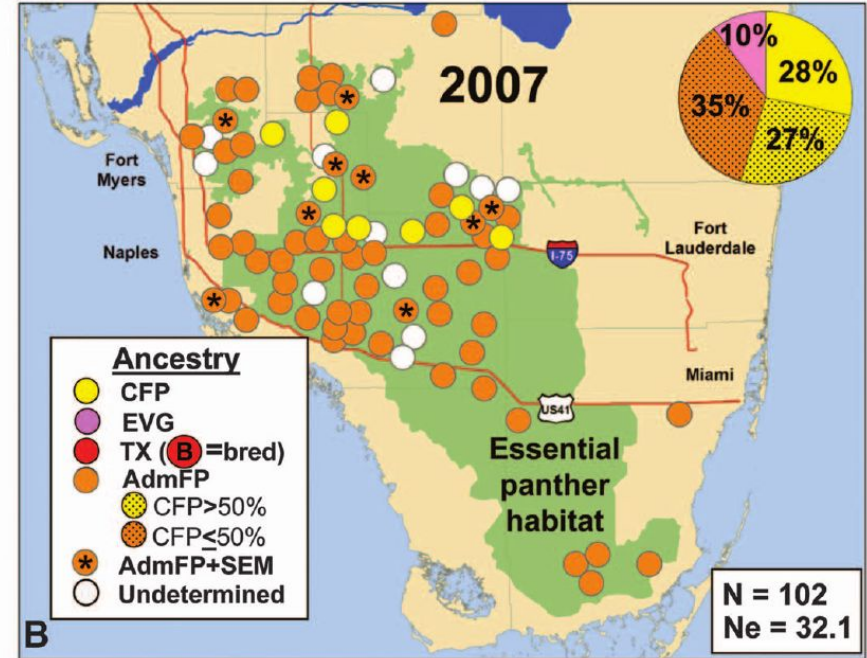
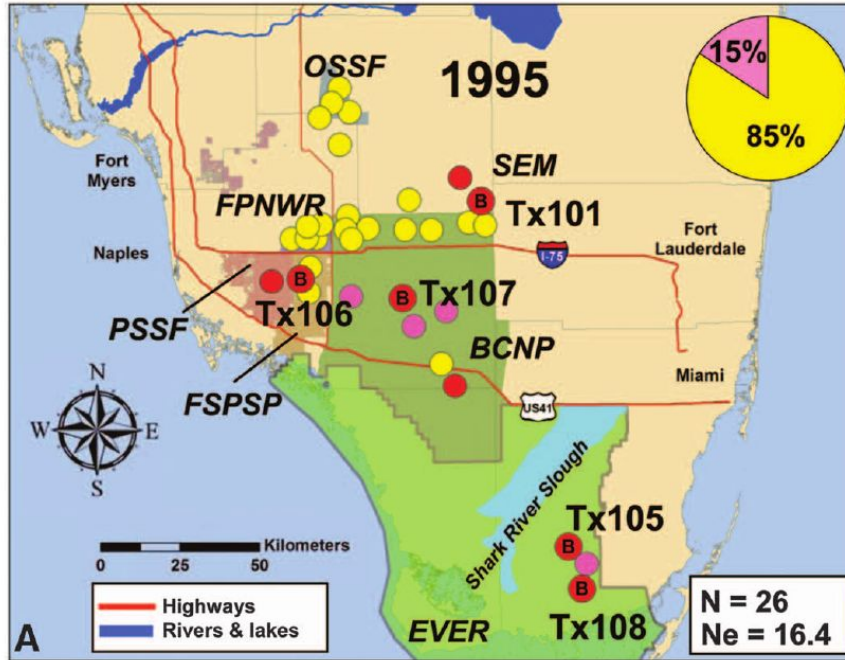


GENETIC RESCUE ($m=0.2$ females) will eliminate lowered fitness, restore genetic diversity, retain adaptive alleles (Hendrick 1995)

THE FLORIDA PANTHER

from genetics to genomics

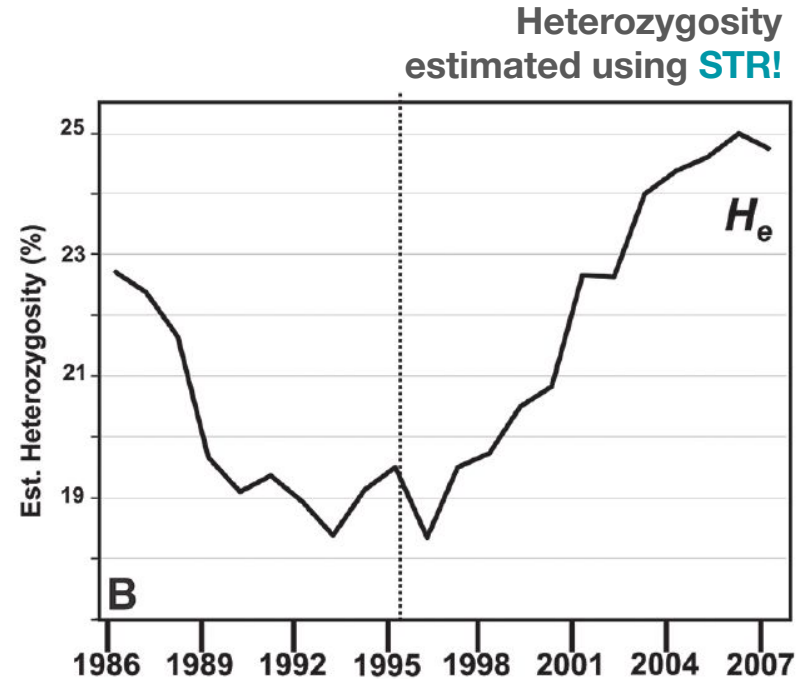
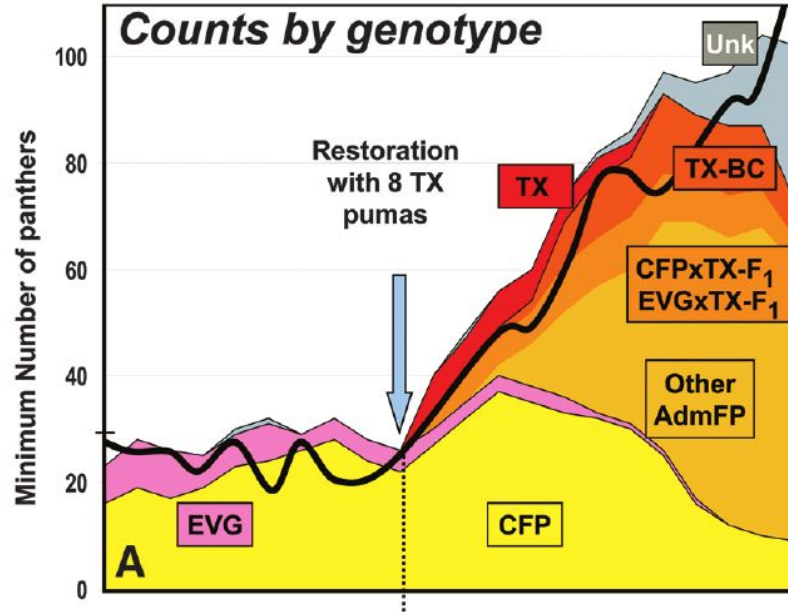
Hybrid females did better! 3X survival!



12 YEARS AFTER THE GENETIC RESCUE N, Ne and range increased due to hybrids individuals (Johnson et al 2010)

THE FLORIDA PANTHER

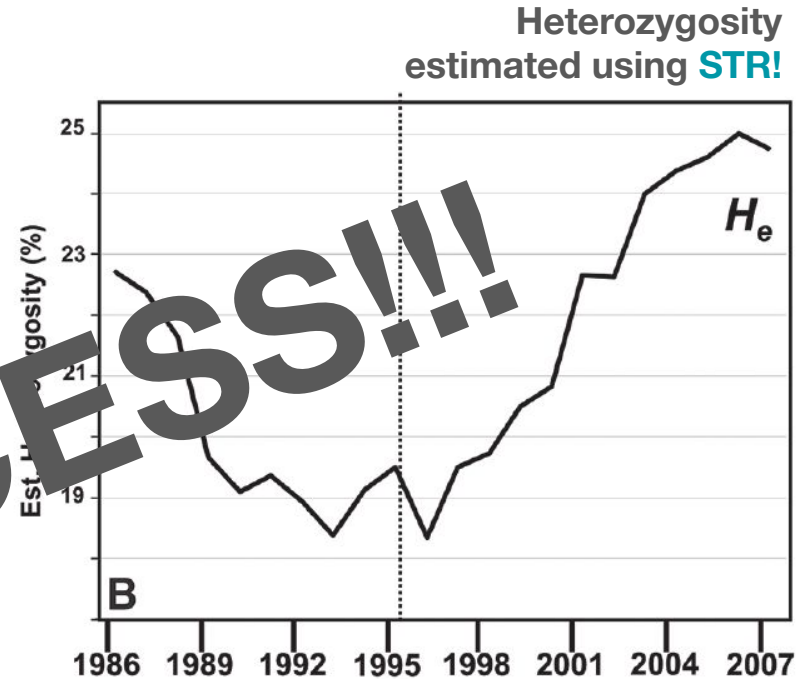
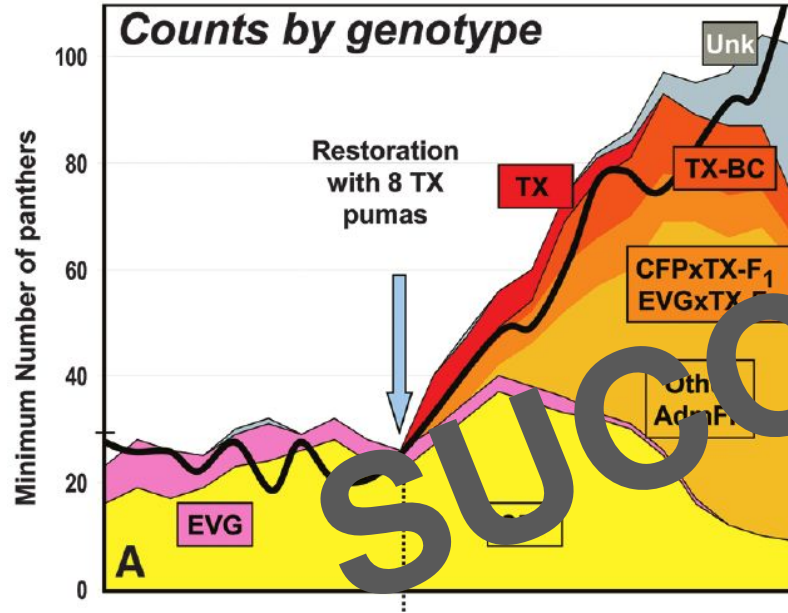
From genetics to genomics



12 YEARS AFTER THE GENETIC RESCUE N , N_e and range increased due to hybrids individuals, and heterozygosity, of course! (Johnson et al 2010)

THE FLORIDA PANTHER

From genetics to genomics



12 YEARS AFTER THE GENETIC RESCUE N , N_e and range increased due to hybrids individuals, and heterozygosity, of course! (Johnson et al 2010)

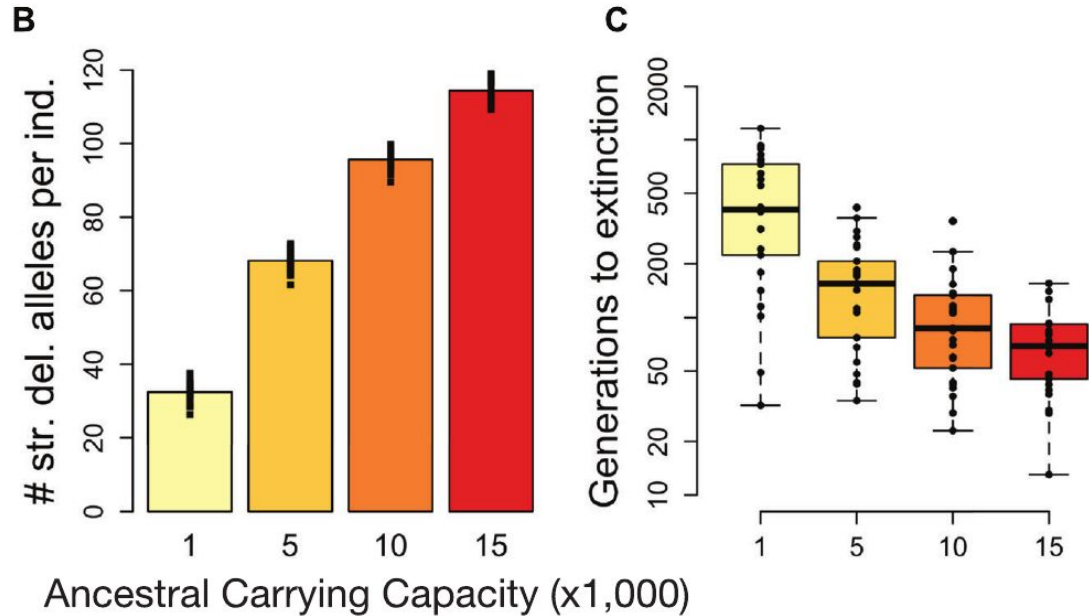
THE FLORIDA PANTHER

From genetics to genomics

BUT LET'S INVESTIGATE THIS SUCCESS STORY
THROUGH A **GENOMIC LENS**

THE FLORIDA PANTHER

From genetics to genomics



Time



Ancestral population

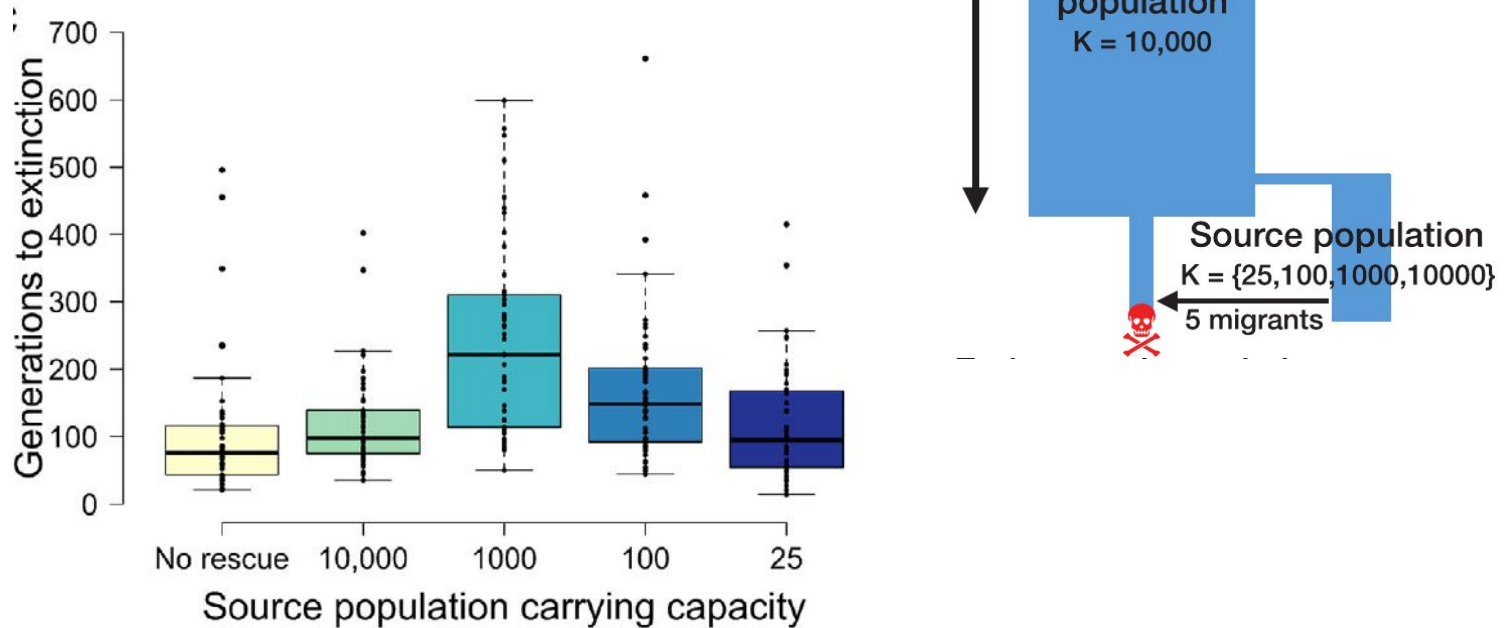
$K = \{1,000, 5,000, 10,000, 15,000\}$



GENOMICS SIMULATIONS indicate that larger populations have larger risk of extinction in case of a sudden bottleneck (Kyriazis et al 2021)

THE FLORIDA PANTHER

From genetics to genomics



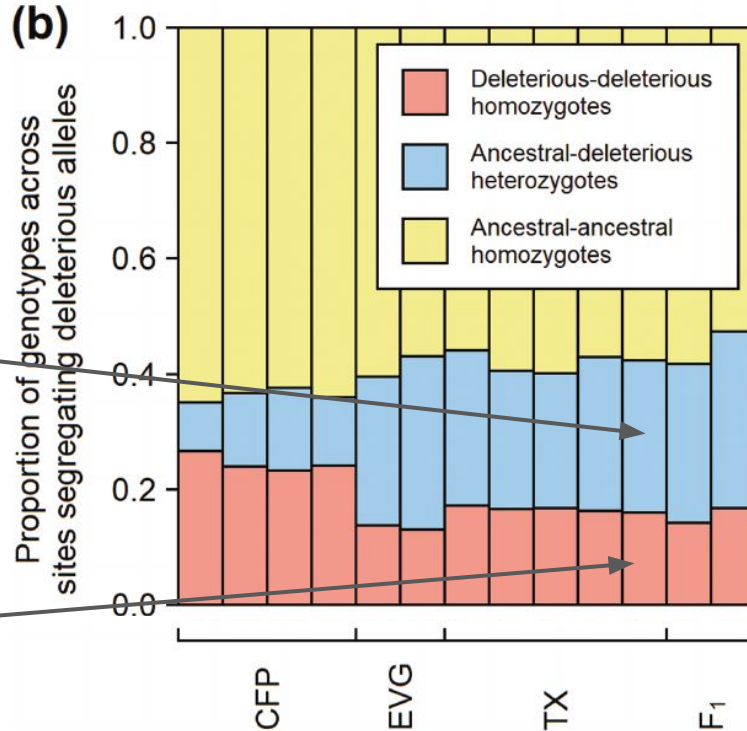
AND THAT RESCUES FROM LARGE POPS are expected to be less effective - if the receiving population stays small (Kyriazis et al 2021)

THE FLORIDA PANTHER

From genetics to genomics

Masked genetic load
has increased

Realized genetic load
has decreased



AS IT IS THE CASE OF THE FLORIDA PANTHER which is accumulating deleterious alleles from the large donor Texas pop! (Ochoa et al 2022)

WHAT IS THE DIFFERENCE BETWEEN CONSERVATION GENETICS AND GENOMICS?

All in all, it is still using **population genetics models** to inform conservation biology

WHAT IS THE DIFFERENCE BETWEEN CONSERVATION GENETICS AND GENOMICS?

All in all, it is still using population genetics models to inform conservation biology

But employing the full breadth of genomic data (**neutral and non-neutral genetic variation**) to get a more accurate picture of what is going on in the population

WHAT IS THE DIFFERENCE BETWEEN CONSERVATION GENETICS AND GENOMICS?

All in all, it is still using population genetics models to inform conservation biology

But employing the full breadth of genomic data (neutral and non-neutral genetic variation) to get a more accurate picture of what is going on in the population

The first simplest advantage: using lots of loci (whole-genome) **improves the accuracy** of our estimates

ESTIMATING EFFECTIVE POPULATION SIZE (N_e)

Are more individuals better than more loci?

How to estimate N_e (e.g., stationary size) from one diploid individual?

ESTIMATING EFFECTIVE POPULATION SIZE (N_e)

Are more individuals better than more loci?

How to estimate N_e (e.g., stationary size) from one diploid individual?

$$H = \theta = 4N_e\mu$$

$$N_e = \frac{H}{4\mu(1-H)} \approx \frac{H}{4\mu}$$

ESTIMATING EFFECTIVE POPULATION SIZE (N_e)

Are more individuals better than more loci?

$$N_e = 10000$$

$$\mu = 10^{-9}$$

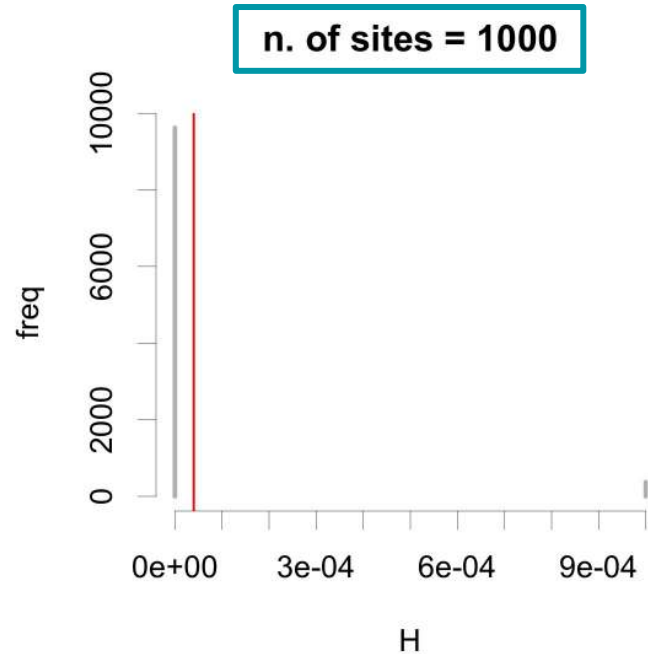
$$H = 4 \times 10^{-5}$$

Sample size: 1 diploid individual

How to estimate N_e (e.g., stationary size) from one diploid individual?

$$H = \theta = 4N_e\mu$$

$$N_e = \frac{H}{4\mu(1-H)} \approx \frac{H}{4\mu}$$



ESTIMATING EFFECTIVE POPULATION SIZE (N_e)

Are more individuals better than more loci?

$$N_e = 10000$$

$$\mu = 10^{-9}$$

$$H = 4 \times 10^{-5}$$

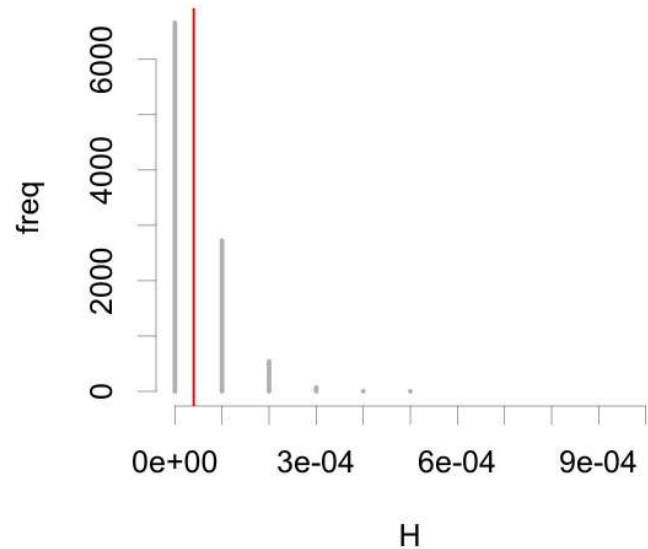
Sample size: 1 diploid individual

How to estimate N_e (e.g., stationary size) from one diploid individual?

n. of sites = 10000

$$H = \theta = 4N_e\mu$$

$$N_e = \frac{H}{4\mu(1-H)} \approx \frac{H}{4\mu}$$



ESTIMATING EFFECTIVE POPULATION SIZE (N_e)

Are more individuals better than more loci?

$$N_e = 10000$$

$$\mu = 10^{-9}$$

$$H = 4 \times 10^{-5}$$

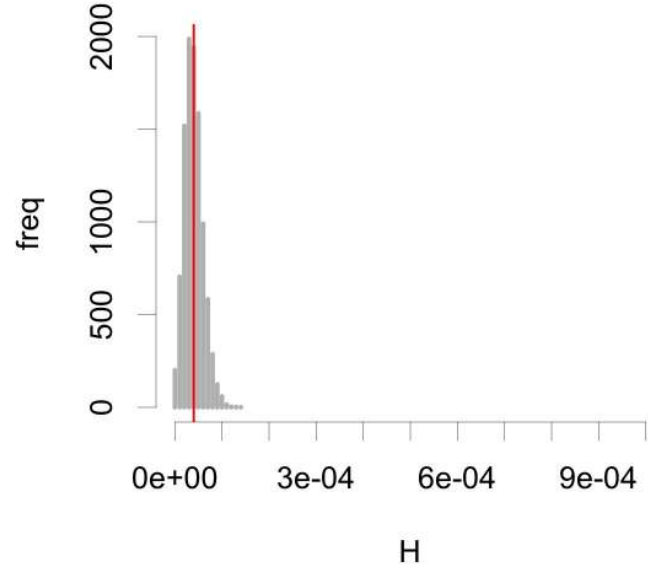
Sample size: 1 diploid individual

How to estimate N_e (e.g., stationary size) from one diploid individual?

$$H = \theta = 4N_e\mu$$

$$N_e = \frac{H}{4\mu(1-H)} \approx \frac{H}{4\mu}$$

n. of sites = 1e+05



ESTIMATING EFFECTIVE POPULATION SIZE (N_e)

Are more individuals better than more loci?

$$N_e = 10000$$

$$\mu = 10^{-9}$$

$$H = 4 \times 10^{-5}$$

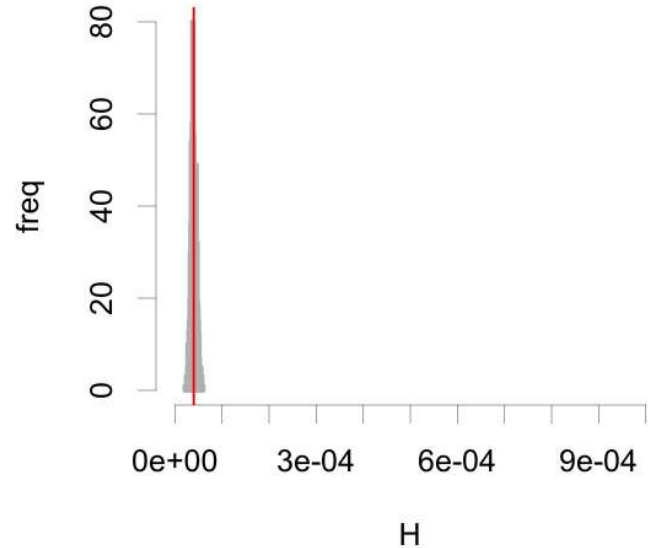
Sample size: 1 diploid individual

How to estimate N_e (e.g., stationary size) from one diploid individual?

$$H = \theta = 4N_e\mu$$

$$N_e = \frac{H}{4\mu(1-H)} \approx \frac{H}{4\mu}$$

n. of sites = 1e+06



ESTIMATING EFFECTIVE POPULATION SIZE (N_e)

Are more individuals better than more loci?

$$N_e = 10000$$

$$\mu = 10^{-9}$$

$$H = 4 \times 10^{-5}$$

How to estimate N_e (e.g., stationary size) from multiple individuals?

$$H = \theta = 4N_e\mu$$

ESTIMATING EFFECTIVE POPULATION SIZE (N_e)

Are more individuals better than more loci?

$$N_e = 10000$$

$$\mu = 10^{-9}$$

$$H = 4 \times 10^{-5}$$

How to estimate N_e (e.g., stationary size) from multiple individuals?

$$\hat{\theta} = \theta = 4N_e\mu$$

$$\hat{\theta} = \frac{S}{\sum_{i=1}^{n-1} 1/i}$$

ESTIMATING EFFECTIVE POPULATION SIZE (N_e)

Are more individuals better than more loci?

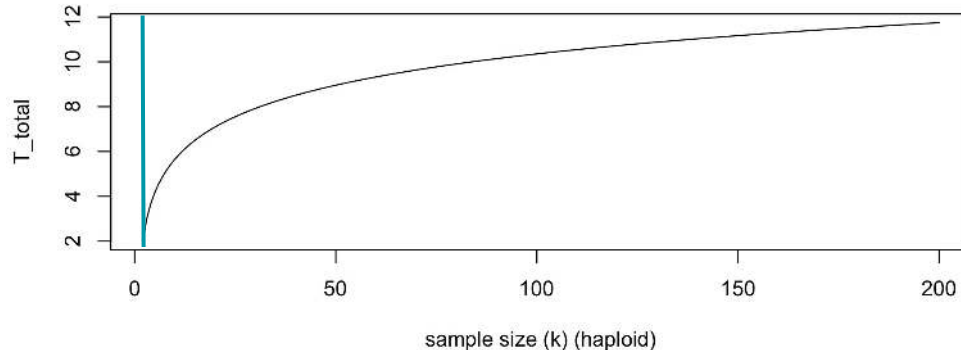
$$N_e = 10000$$

$$\mu = 10^{-9}$$

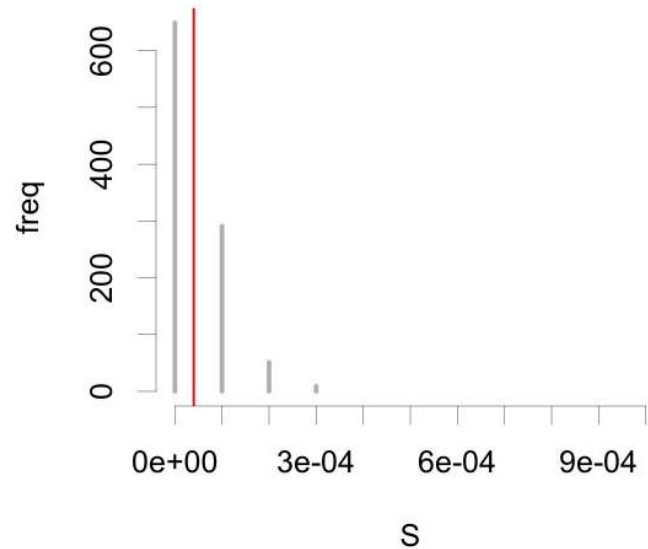
$$H = 4 \times 10^{-5}$$

How does precision of estimation of theta increase with sample size?

1 diploid individual



n. of sites = 10000 ; k = 2



ESTIMATING EFFECTIVE POPULATION SIZE (N_e)

Are more individuals better than more loci?

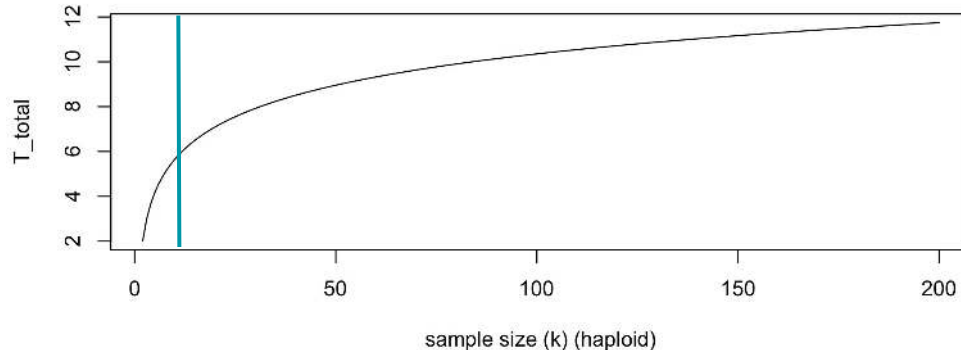
$$N_e = 10000$$

$$\mu = 10^{-9}$$

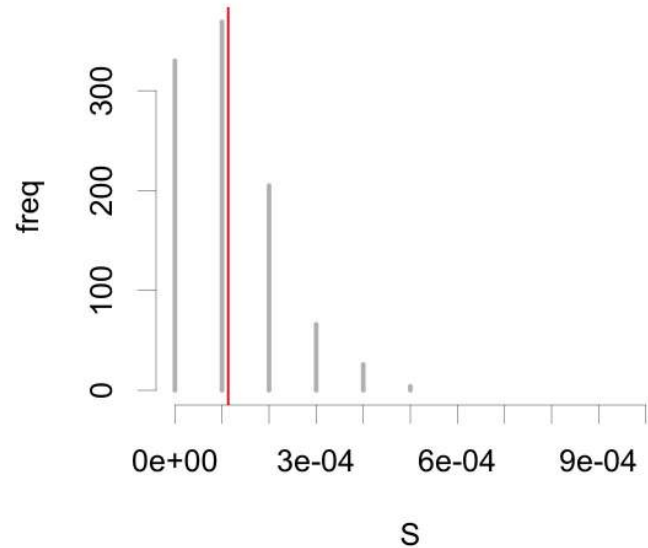
$$H = 4 \times 10^{-5}$$

How does precision of estimation of theta increase with sample size?

5 diploid individual



n. of sites = 10000 ; k = 10



ESTIMATING EFFECTIVE POPULATION SIZE (N_e)

Are more individuals better than more loci?

$$N_e = 10000$$

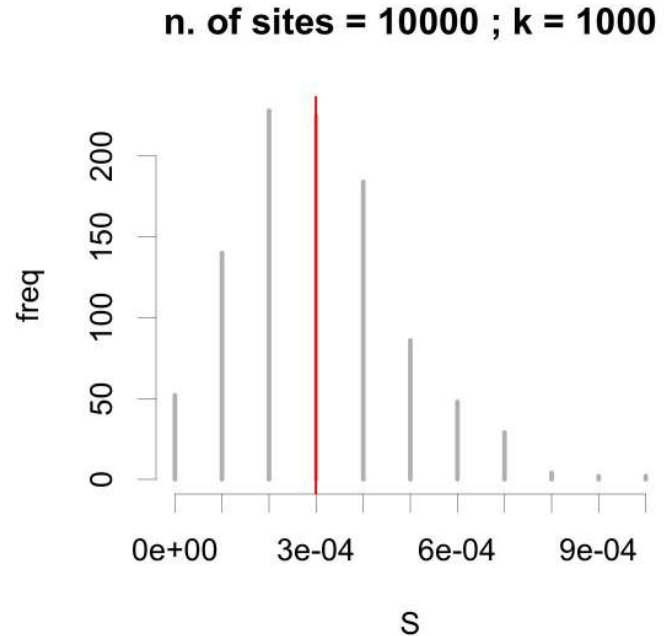
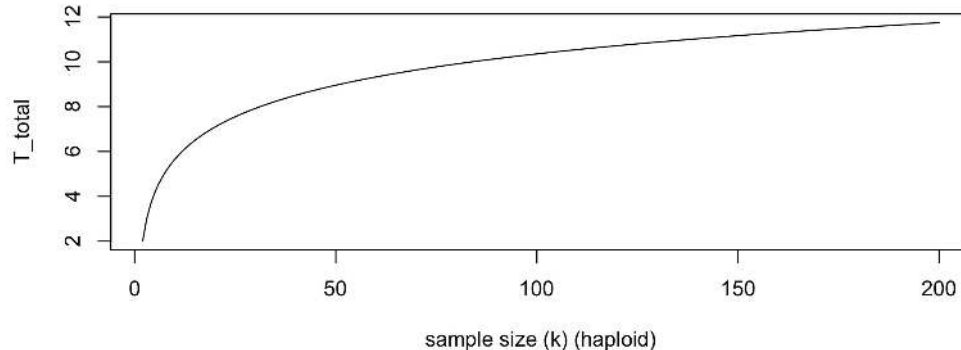
$$\mu = 10^{-9}$$

$$H = 4 \times 10^{-5}$$

How does precision of estimation of theta increase with sample size?

500 diploid individual

Not much really!!



ESTIMATING EFFECTIVE POPULATION SIZE (N_e)

Are more individuals better than more loci?

$$N_e = 10000$$

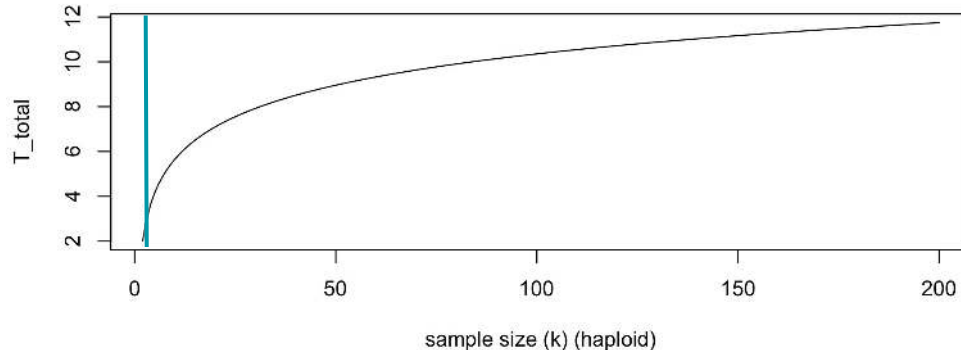
$$\mu = 10^{-9}$$

$$H = 4 \times 10^{-5}$$

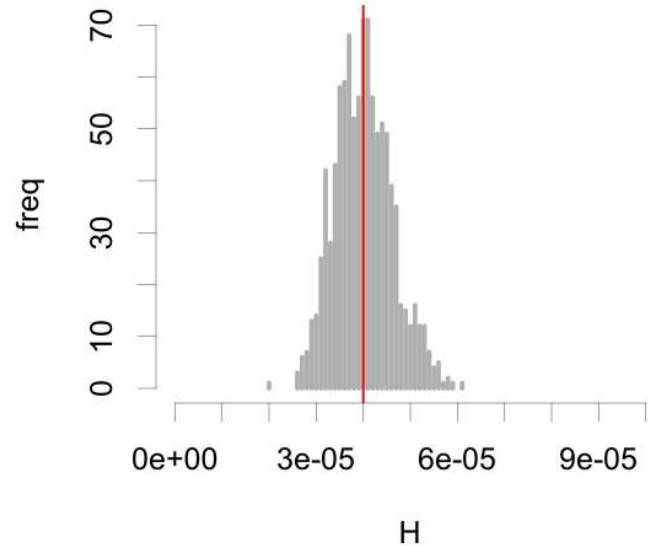
How does precision of estimation of theta increase with sample size?

1 diploid individual

Better to increase the number of sites
(or both)



n. of sites = $1e+06$



WHAT IS THE DIFFERENCE BETWEEN
CONSERVATION GENETICS AND GENOMICS?

**WHAT IS THE FIRST INGREDIENT OF
CONSERVATION GENOMICS?**

HOW CAN GENOMICS INFORM
CONSERVATION BIOLOGY?

IN WHICH APPLICATION GENOMICS IS
A GAME CHANGER?

with a close up on two topics:

Genetic load

ROHs

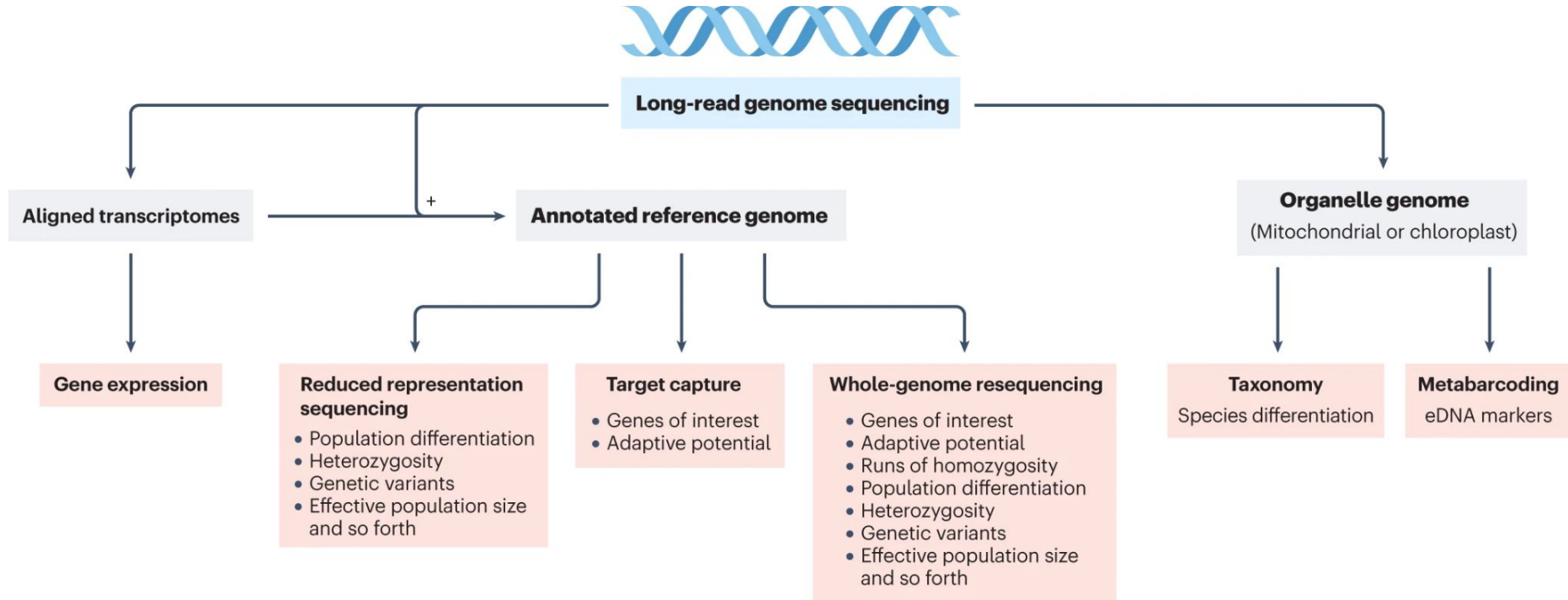
WHAT AM I DOING TO MAKE USE OF
GENOMICS IN CONSERVATION?

FIRST, A WELL ANNOTATED REFERENCE GENOME

A blossoming of genome assembly initiatives

FIRST, A WELL ANNOTATED REFERENCE GENOME

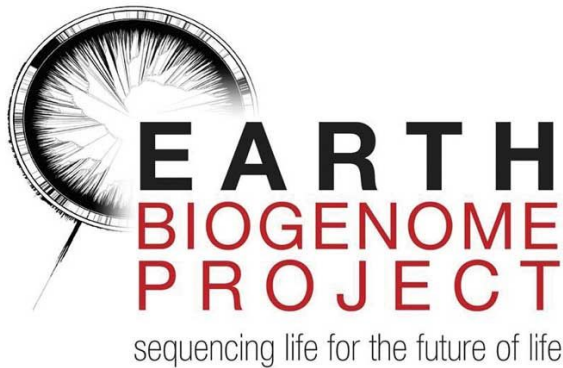
A blossoming of genome assembly initiatives



A reference genome can lead to many downstream applications (Hogg 2023)

FIRST, A WELL ANNOTATED REFERENCE GENOME

A blossoming of genome assembly initiatives



Earth BioGenome Project: Sequencing life for the future of life

Harris A. Lewin , Gene E. Robinson, W. John Kress, +20, and Guojie Zhang [Authors Info & Affiliations](#)

Edited by John C. Avise, University of California, Irvine, CA, and approved March 15, 2018 (received for review January 6, 2018)

April 23, 2018 | 115 (17) 4325-4333 | <https://doi.org/10.1073/pnas.1720115115>



Article | [Open Access](#) | Published: 28 April 2021

Towards complete and error-free genome assemblies of all vertebrate species

Arang Rhie, [Shane A. McCarthy](#), [Olivier Fedrigo](#), [Joana Damas](#), [Giulio Formenti](#), [Sergey Koren](#), [Marcela](#)

FIRST, A WELL ANNOTATED REFERENCE GENOME

A community effort to set the golden standard in EU

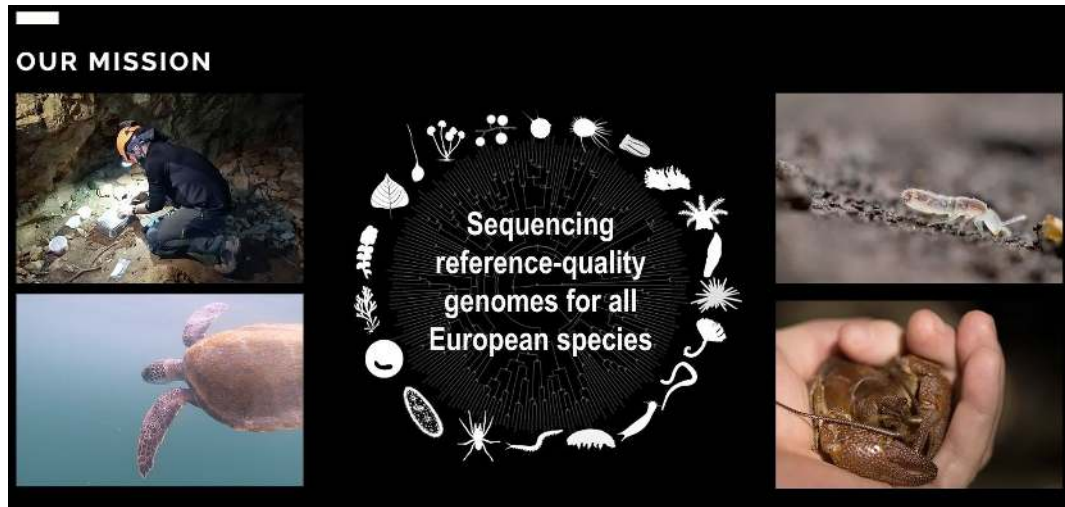


The era of reference genomes in conservation genomics

Giulio Formenti ²⁹ • Kathrin Theissinger ²⁹ • Carlos Fernandes ²⁹ • ... Erich D. Jarvis • Miklós Bálint  •

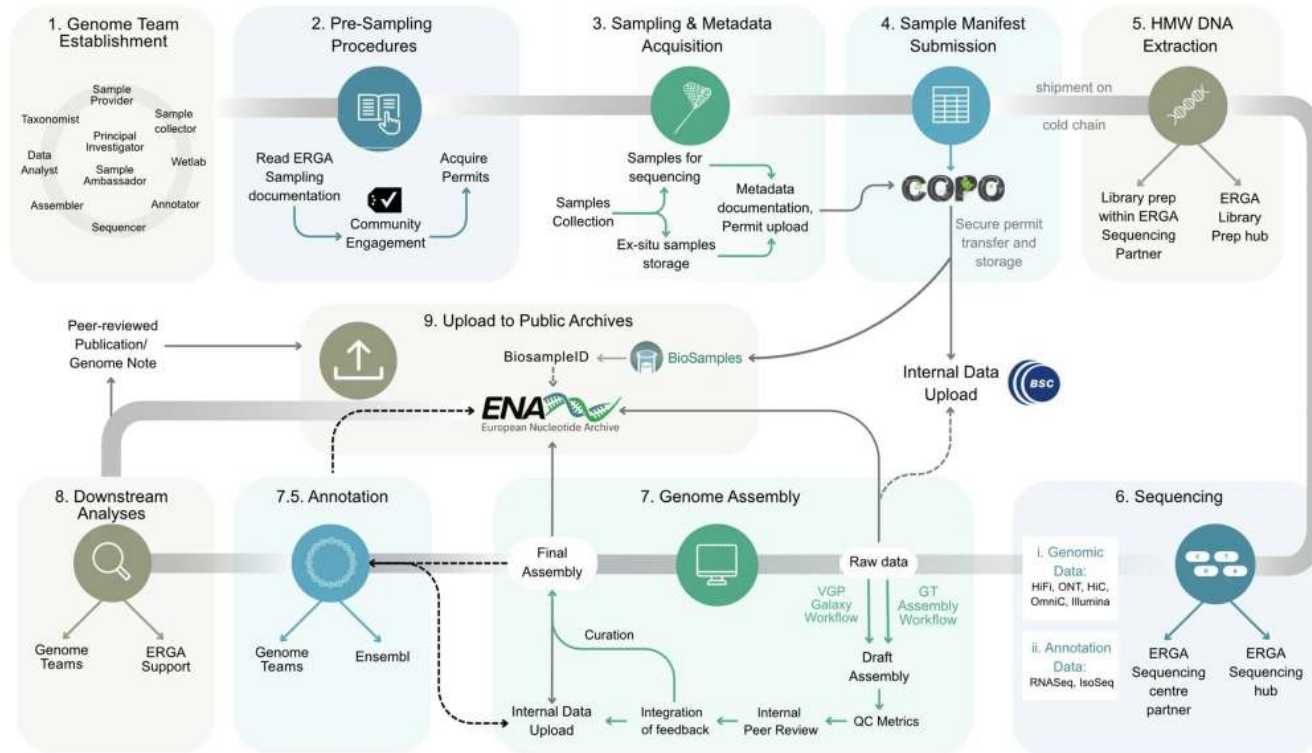
European Reference Genome Atlas (ERGA) Consortium ^{30, 31} • [Show all authors](#) • [Show footnotes](#)

[Open Access](#) • Published: January 24, 2022 • DOI: <https://doi.org/10.1016/j.tree.2021.11.008> •



FIRST, A WELL ANNOTATED REFERENCE GENOME

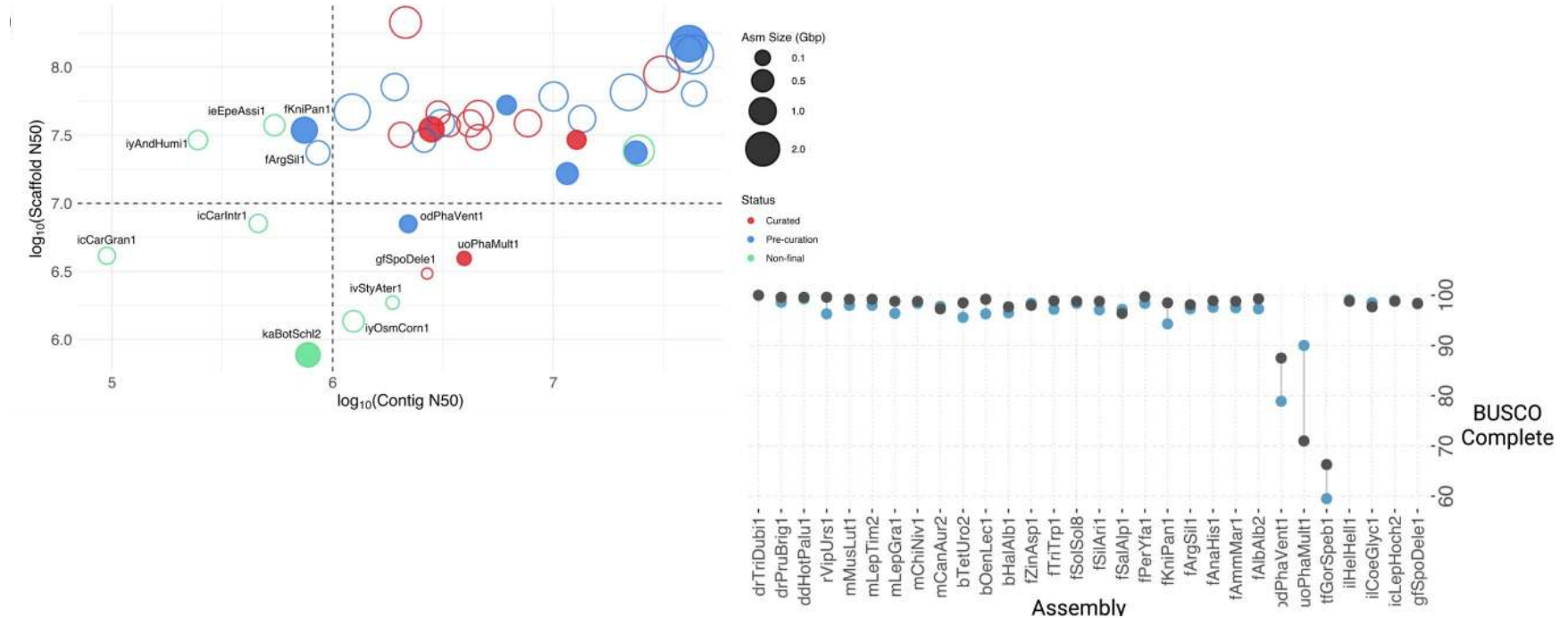
From sample to high quality genome assemblies



BEST PRACTICE in the ERGA pilot study (Mc Cartney et al 2024)

FIRST, A WELL ANNOTATED REFERENCE GENOME

Scaffold length (up to chrs!), QVs, and BUSCO requirements



QUALITY THRESHOLDS in the ERGA pilot study (Mc Cartney et al 2024)

WHAT IS THE FIRST INGREDIENT OF CONSERVATION GENOMICS?

The gold standard is a high-quality, highly-contiguous (e.g., **chromosome level**) reference genome assembly

WHAT IS THE FIRST INGREDIENT OF CONSERVATION GENOMICS?

The gold standard is a high-quality, highly-contiguous (e.g., **chromosome level**) reference genome assembly

But **pangenomes** are going to be the next level soon!

WHAT IS THE FIRST INGREDIENT OF CONSERVATION GENOMICS?

The gold standard is a high-quality, highly-contiguous (e.g., chromosome level) reference genome assembly

But pangenomes are going to be the next level soon!

Well **annotated genome features** (e.g., genes, transposable elements, long-non coding RNAs, promoters, etc)

WHAT IS THE FIRST INGREDIENT OF CONSERVATION GENOMICS?

The gold standard is a high-quality, highly-contiguous (e.g., chromosome level) reference genome assembly

But pangenomes are going to be the next level soon

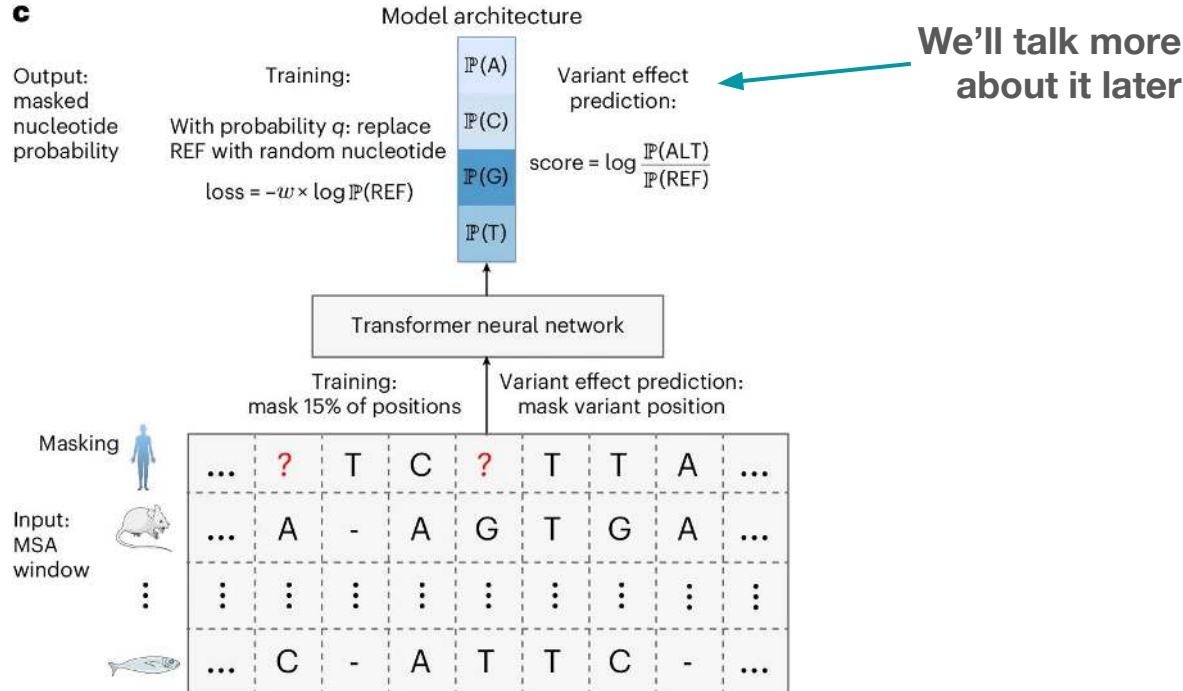
Well **annotated genome features** (e.g., genes, transposable elements, long-non coding RNAs, promoters, etc)

There's room for improvements: technological (e.g., PacBio Kinnex), methodological (e.g., **machine learning approaches**)

GENOMIC FEATURES ANNOTATION

New AI powered approaches

c

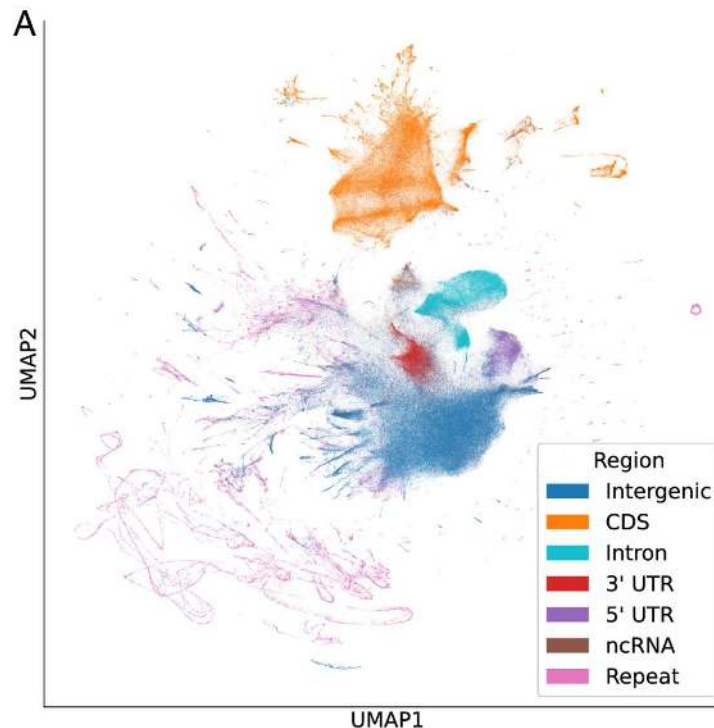


Genomic Pretrained Network /with multiple-sequence alignment
GPN/GPN-MSA (Benegas et al 2023 and 2025).

GENOMIC FEATURES ANNOTATION

New AI powered approaches

It can learn **gene structure** ...



Genomic Pretrained Network /with multiple-sequence alignment
GPN/GPN-MSA (Benegas et al 2023 and 2025).

GENOMIC FEATURES ANNOTATION

New AI powered approaches

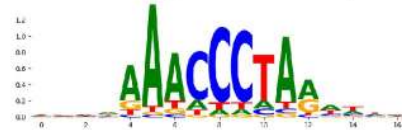
... and **DNA motifs**
without supervision

B GPN motif extracted by TF-MoDISco

Motif 18 (780 occurrences)



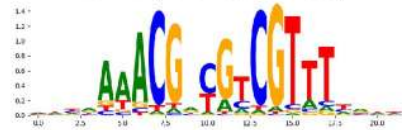
Motif 2 (3386 occurrences)



Motif 9 (1385 occurrences)

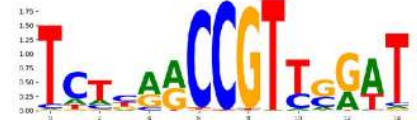


Motif 10 (1057 occurrences)



Reported match in PlantTFDB

AT4G24470



AT1G72740



(no match)

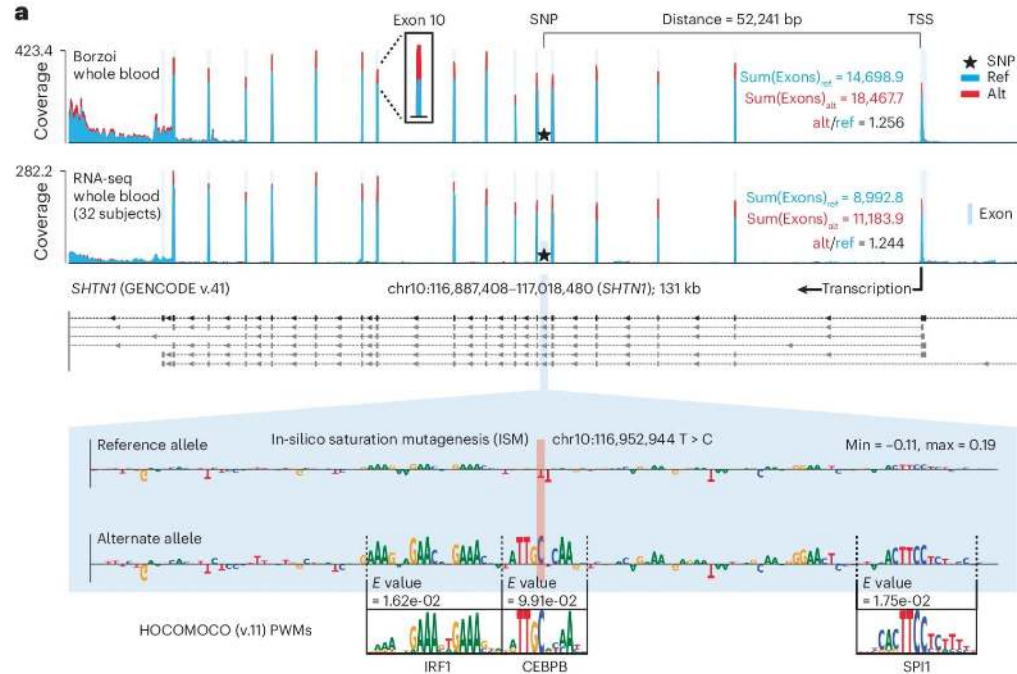
(no match)

Genomic Pretrained Network /with multiple-sequence alignment
GPN/GPN-MSA (Benegas et al 2023 and 2025).

GENOMIC FEATURES ANNOTATION

New AI powered approaches

It can predicts far away
TF variants effect on
gene expression



Borzoi derives TF motifs and a genome-wide map of nucleotide influence on gene structure and expression (Linder et al **2025**).

GENOMIC FEATURES ANNOTATION

Technological improvements

Both ONT and PacBio can
now provide accurate
methylation data

SMRTbell® library

PacBio® long-read systems

5-base HiFi sequencing with A, C, G, T, +5mC



Nucleotide incorporation kinetics are measured in real time

5-base HiFi sequencing with A, C, G, T, +5mC

ACTGACGGACTGATCGACTG

5mC

5mC

5mC encoded with standard BAM tags³

MM:Z:C+m,4,12,16,4,16,19,44,10

ML:B:C,249,4,247,177,210,228,245,244

The PacBio long-read systems directly output long, highly accurate HiFi reads with annotation of 5mC methylation at all CpG sites.
No special library preparation like bisulfite treatment is required.

EPIGENETIC PROFILING from long-range sequencing data is much more affordable with falling costs of LRS

WHAT IS THE DIFFERENCE BETWEEN
CONSERVATION GENETICS AND GENOMICS?

WHAT IS THE FIRST INGREDIENT OF
CONSERVATION GENOMICS?

**HOW CAN GENOMICS INFORM
CONSERVATION BIOLOGY?**

IN WHICH APPLICATION GENOMICS IS
A GAME CHANGER?

with a close up on two topics:

Genetic load

ROHs

WHAT AM I DOING TO MAKE USE OF
GENOMICS IN CONSERVATION?

PROGRESS IN ADOPTION OF GENOMICS IN CONSERVATION

Broadly used  Getting better  To be improved  (Hogg 2023)

Increase accuracy estimating parameters that require neutral markers

PROGRESS IN ADOPTION OF GENOMICS IN CONSERVATION

Broadly used  Getting better  To be improved  (Hogg 2023)

Increase accuracy estimating parameters that require neutral markers

-  Effective population size
-  Migration rates (gene flow)
-  Inferring haplotypes from linked loci to determine directionality of migration
-  Accuracy of kin estimates
-  Pedigree reconstruction of wild populations
-  Proportion of admixture to assess population delineations
-  Landscape genetics

PROGRESS IN ADOPTION OF GENOMICS IN CONSERVATION

Broadly used  Getting better  To be improved  (Hogg 2023)

Increase accuracy estimating parameters that require neutral markers

-  Effective population size
-  Migration rates (gene flow)
-  Inferring haplotypes from linked loci to determine directionality of migration
-  Accuracy of kin estimates
-  Pedigree reconstruction of wild populations
-  Proportion of admixture to assess population delineations
-  Landscape genetics

Estimate inbreeding depression (ID)

PROGRESS IN ADOPTION OF GENOMICS IN CONSERVATION

Broadly used  Getting better  To be improved  (Hogg 2023)

Increase accuracy estimating parameters that require neutral markers

-  Effective population size
-  Migration rates (gene flow)
-  Inferring haplotypes from linked loci to determine directionality of migration
-  Accuracy of kin estimates
-  Pedigree reconstruction of wild populations
-  Proportion of admixture to assess population delineations
-  Landscape genetics

Estimate inbreeding depression (ID)

-  Selection coefficients (interactions between drift, selection and migration)
-  Molecular basis and genetic architecture of ID
-  Identifying loci contributing to ID by sequencing parents and offspring

PROGRESS IN ADOPTION OF GENOMICS IN CONSERVATION

Broadly used  Getting better  To be improved  (Hogg 2023)

Detect local adaptation

PROGRESS IN ADOPTION OF GENOMICS IN CONSERVATION

Broadly used  Getting better  To be improved  (Hogg 2023)

Detect local adaptation

-  Genome-wide variation vs. specific genomic regions responding to selection
-  Signatures of selection to assess local adaptations
-  Mapping associations btw. adaptive genome regions - environmental gradients

PROGRESS IN ADOPTION OF GENOMICS IN CONSERVATION

Broadly used  Getting better  To be improved  (Hogg 2023)

Detect local adaptation

-  Genome-wide variation vs. specific genomic regions responding to selection
-  Signatures of selection to assess local adaptations
-  Mapping associations btw. adaptive genome regions - environmental gradients

Units of conservation and hybridization

PROGRESS IN ADOPTION OF GENOMICS IN CONSERVATION

Broadly used  Getting better  To be improved  (Hogg 2023)

Detect local adaptation

-  Genome-wide variation vs. specific genomic regions responding to selection
-  Signatures of selection to assess local adaptations
-  Mapping associations btw. adaptive genome regions - environmental gradients

Units of conservation and hybridization

-  Rates of introgression across different genomic regions following hybridization
-  Detection of hybridization
-  Predicting probability of outbreeding depression if populations mixed

PROGRESS IN ADOPTION OF GENOMICS IN CONSERVATION

Broadly used  Getting better  To be improved  (Hogg 2023)

Detect local adaptation

-  Genome-wide variation vs. specific genomic regions responding to selection
-  Signatures of selection to assess local adaptations
-  Mapping associations btw. adaptive genome regions - environmental gradients

Units of conservation and hybridization

-  Rates of introgression across different genomic regions following hybridization
-  Detection of hybridization
-  Predicting probability of outbreeding depression if populations mixed

Captive breeding and assisted migration

PROGRESS IN ADOPTION OF GENOMICS IN CONSERVATION

Broadly used  Getting better  To be improved  (Hogg 2023)

Detect local adaptation

-  Genome-wide variation vs. specific genomic regions responding to selection
-  Signatures of selection to assess local adaptations
-  Mapping associations btw. adaptive genome regions - environmental gradients

Units of conservation and hybridization

-  Rates of introgression across different genomic regions following hybridization
-  Detection of hybridization
-  Predicting probability of outbreeding depression if populations mixed

Captive breeding and assisted migration

-  Founder relationships in captive breeding programmes
-  Genome-wide heterozygosity to manage inbreeding depression
-  Genetic rescue
-  Minimizing adaptation to captivity

WHAT IS THE DIFFERENCE BETWEEN
CONSERVATION GENETICS AND GENOMICS?

WHAT IS THE FIRST INGREDIENT OF
CONSERVATION GENOMICS?

HOW CAN GENETICS INFORM
CONSERVATION BIOLOGY?

**IN WHICH APPLICATION GENOMICS IS
A GAME CHANGER?**

with a close up on two topics:

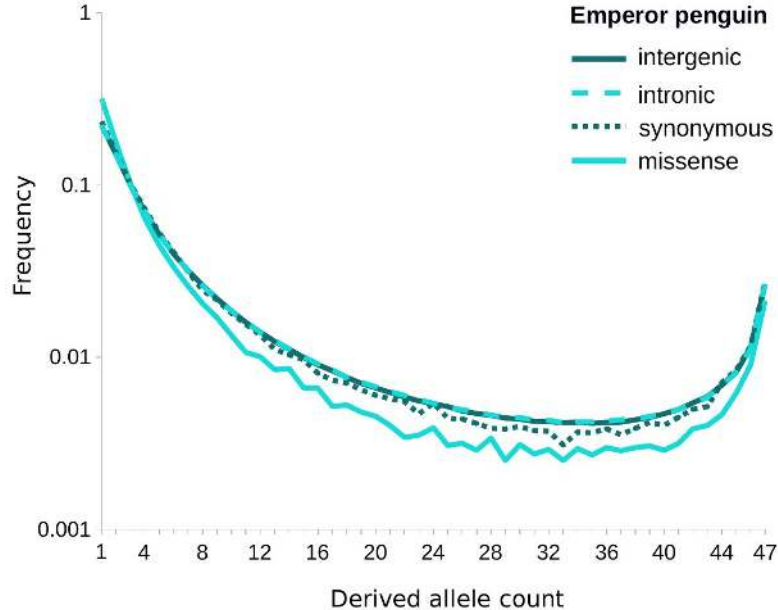
Genetic load

ROHs

WHAT AM I DOING TO MAKE USE OF
GENOMICS IN CONSERVATION?

GETTING BEYOND NEUTRAL VARIATION

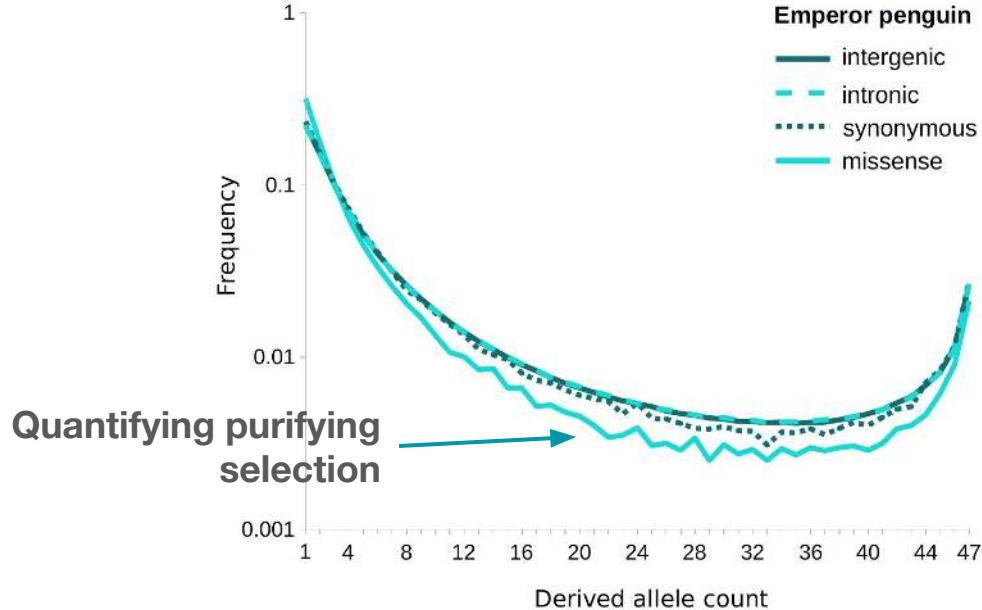
Investigating adaptive and maladaptive processes



ESTIMATE FUNCTIONAL DIVERSITY: nonsynonymous, exonic, in transcription factor binding sites, etc. (Trucchi et al 2024)

GETTING BEYOND NEUTRAL VARIATION

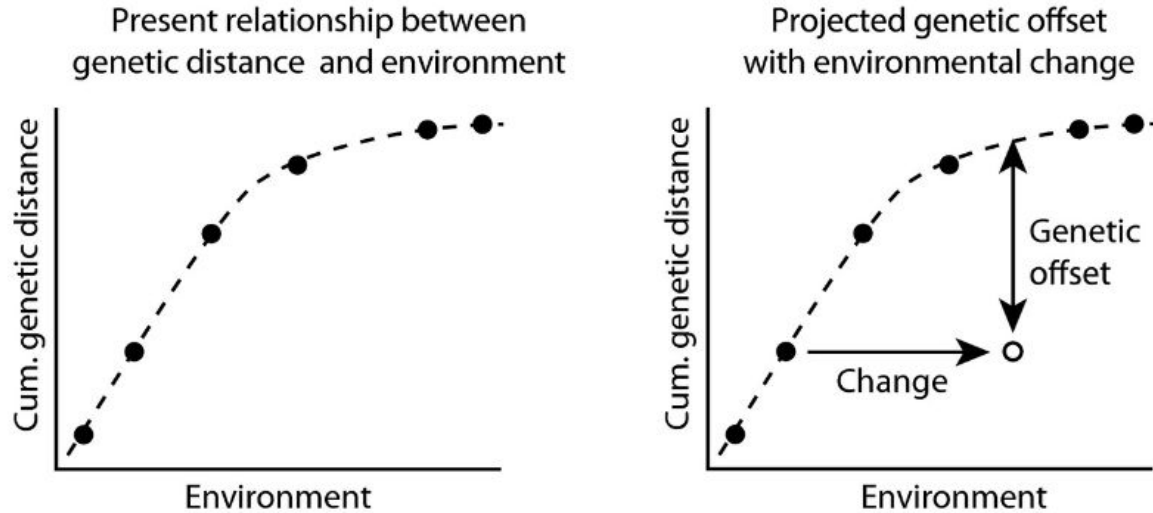
Investigating adaptive and maladaptive processes



ESTIMATE FUNCTIONAL DIVERSITY: nonsynonymous, exonic, in transcription factor binding sites, etc. (Trucchi et al 2024)

GETTING BEYOND NEUTRAL VARIATION

Investigating adaptive and maladaptive processes



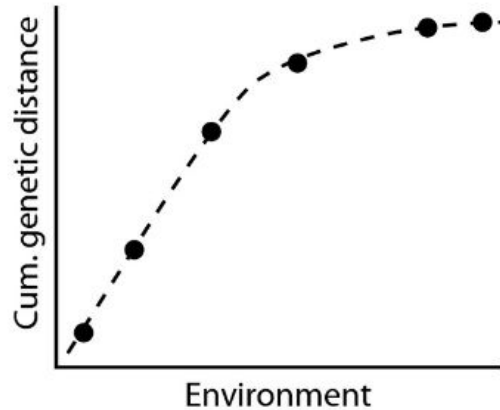
Reillstab et al 2012

DETECT LOCAL ADAPTATION: from selective differentiation and sweeps to genotype to environment association (GEA), to genomic offset or vulnerability.

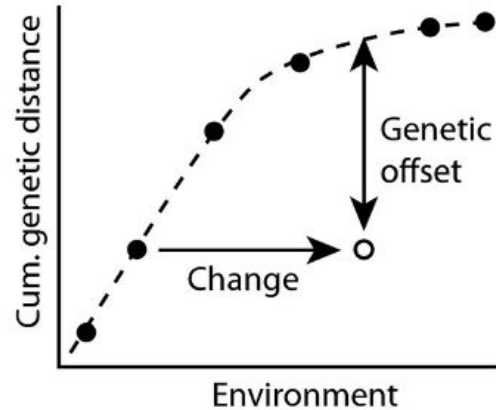
GETTING BEYOND NEUTRAL VARIATION

Investigating adaptive and maladaptive processes

Present relationship between genetic distance and environment



Projected genetic offset with environmental change

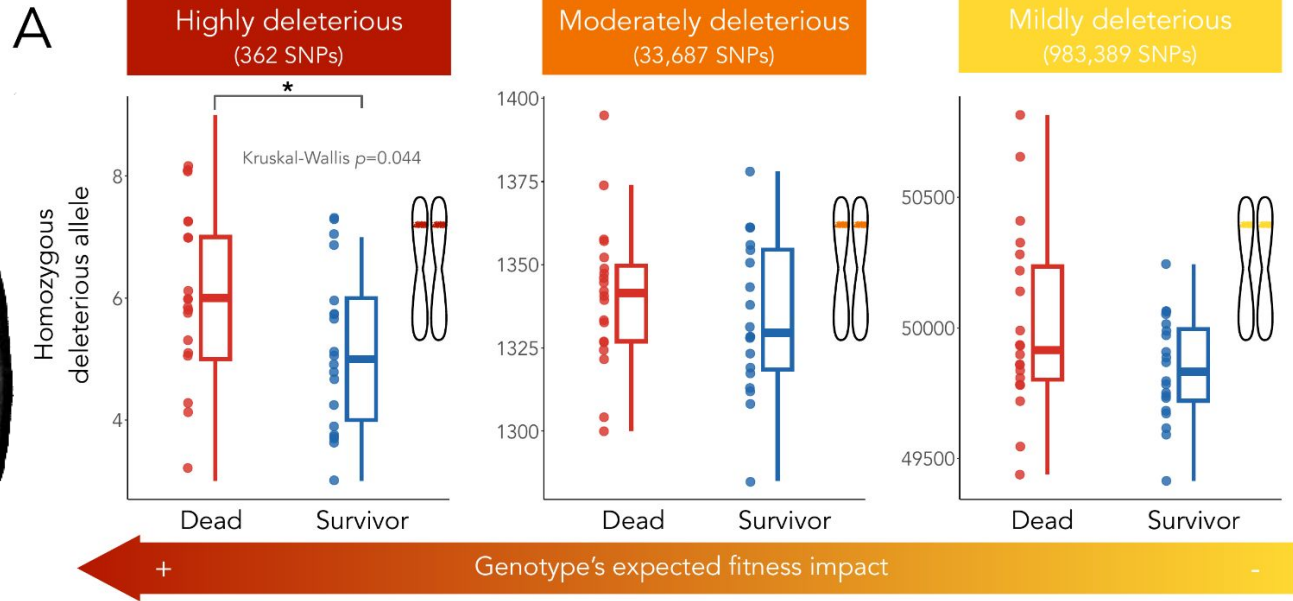


for [more details and more complex](#) approaches see Hoffman et al 2021 or Lachmuth et al 2023

DETECT LOCAL ADAPTATION: from selective differentiation and sweeps to genotype to environment association (GEA), to genomic offset or vulnerability.

GETTING BEYOND NEUTRAL VARIATION

Investigating adaptive and maladaptive processes



ESTIMATE GENETIC LOAD in relation to individual fitness (Fernandes et al *in prep*)
or to compare different populations (Gabielli et al *in review*)

WHAT IS THE DIFFERENCE BETWEEN
CONSERVATION GENETICS AND GENOMICS?

WHAT IS THE FIRST INGREDIENT OF
CONSERVATION GENOMICS?

HOW CAN GENETICS INFORM
CONSERVATION BIOLOGY?

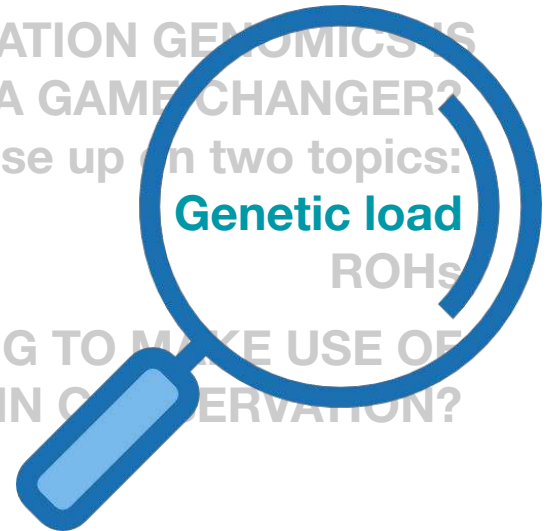
IN WHICH APPLICATION GENOMICS IS
A GAME CHANGER?

with a close up on two topics:

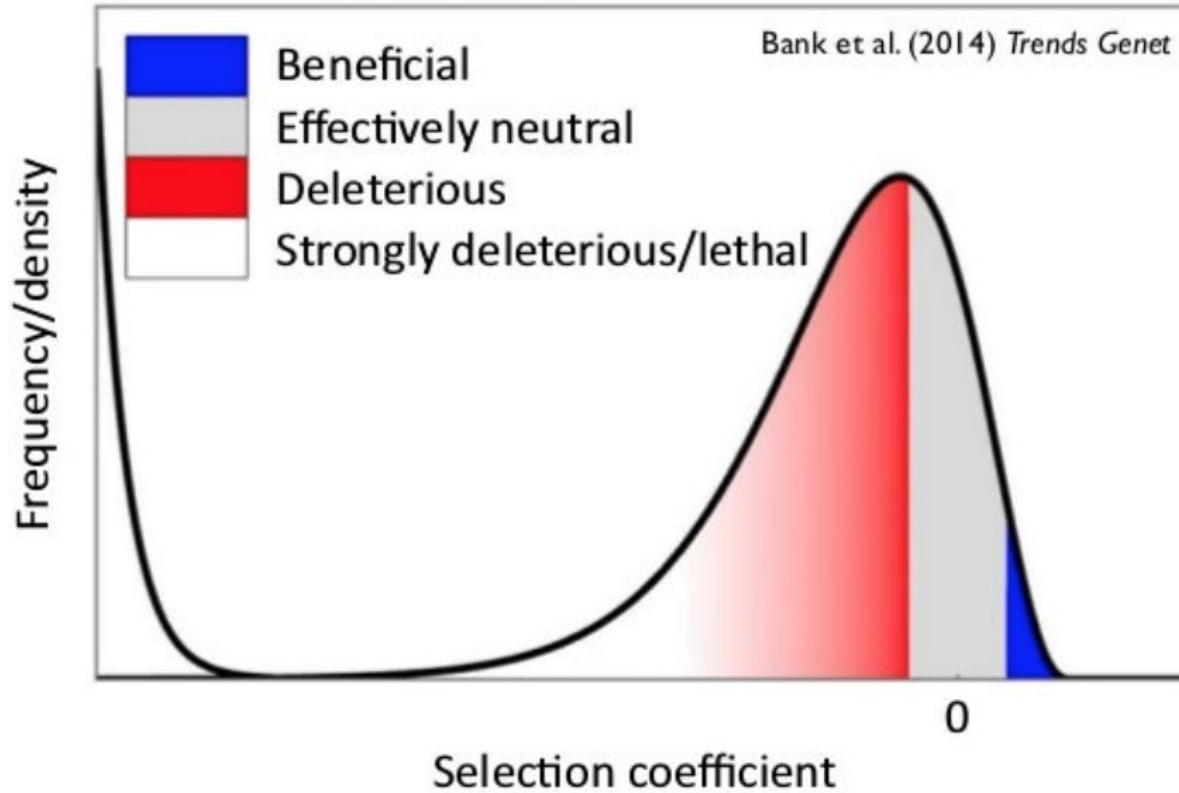
Genetic load

ROHs

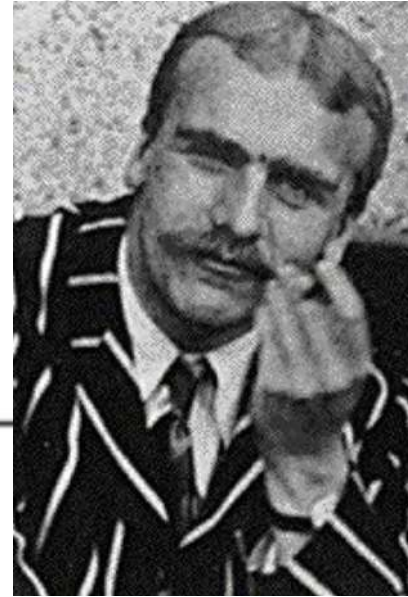
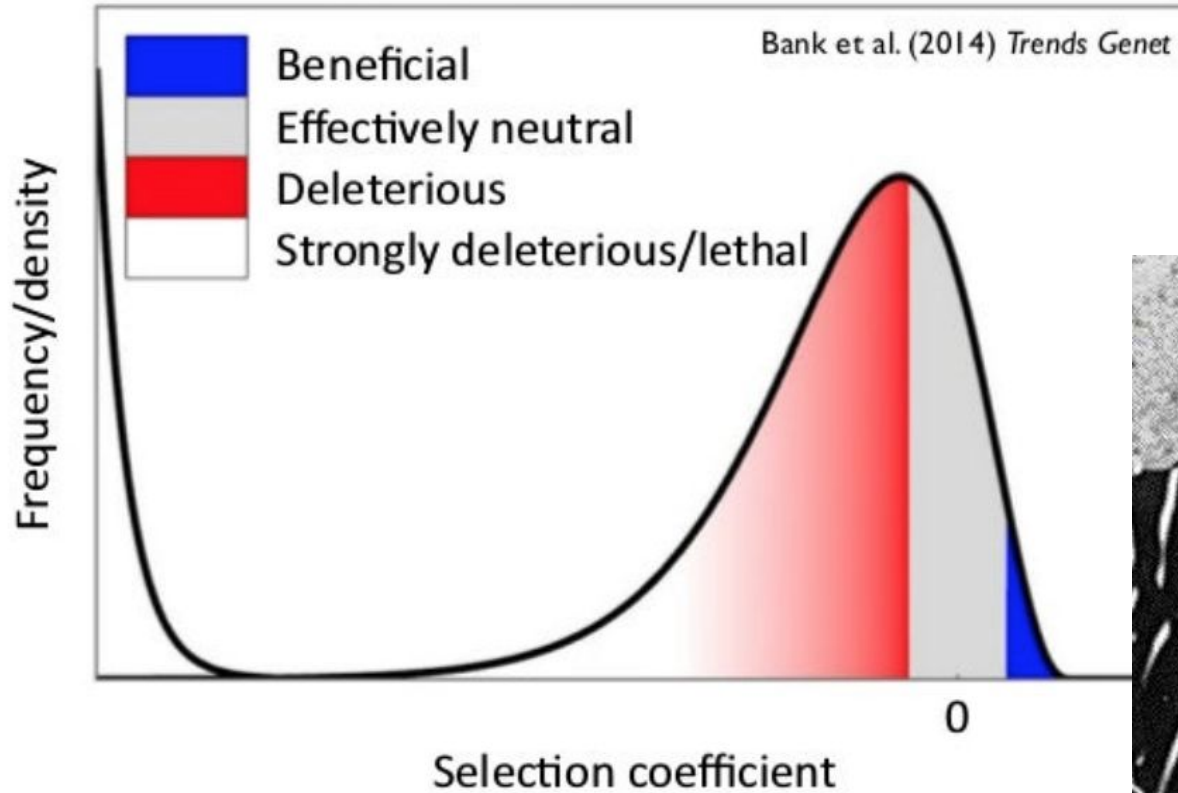
WHAT AM I DOING TO MAKE USE OF
GENOMICS IN CONSERVATION?



WHAT IS GENETIC LOAD?



WHAT IS GENETIC LOAD?



COST OF NATURAL SELECTION: the price paid by a population for its capacity for further evolution

FITNESS AND GENETIC LOAD

$$\text{GENETIC LOAD} = \frac{W_o - \bar{W}}{W_o}$$

W_o Optimal fitness

$\bar{W}$ Average reduced
fitness caused by
mutation(s)

FITNESS AND GENETIC LOAD

REDUCTION OF FITNESS DUE TO:

New deleterious mutations appearing -> **MUTATION LOAD**

Increase in frequency of deleterious mutations -> **DRIFT LOAD**

A more beneficial allele appearing -> **EVOLUTION LOAD**

Deleterious recessive mutations unmasked by inbreeding -> **INBREEDING LOAD**
(**masked** load or potential load)

Deleterious mutations introgressed from a different population -> **MIGRATION LOAD**
(hybrid load)

The break up of favourable combinations of alleles at different loci due to recombination -> **RECOMBINATION LOAD**

See Bertorelle et al 2022

FITNESS AND SELECTION COEFFICIENTS

	AA	Aa	aa
Frequency	$(1-q)^2$	$2q(1-q)$	q^2
Fitness (w)	1	$1-hs$	$1-s$

FITNESS AND SELECTION COEFFICIENTS

	AA	Aa	aa
Frequency	$(1-q)^2$	$2q(1-q)$	q^2
Fitness (w)	1	$1-hs$	$1-s$

$$\bar{w} = 1 - sq^2 - 2q(1 - q)hs$$

GENETIC LOAD AND SELECTION COEFFICIENTS

	AA	Aa	aa
Frequency	$(1-q)^2$	$2q(1-q)$	q^2
Fitness (w)	1	$1-hs$	$1-s$

$$\bar{w} = 1 - sq^2 - 2q(1 - q)hs$$

$$\text{REALIZED GENETIC LOAD} = \frac{1 - \bar{w}}{1} = sq^2 + 2q(1 - q)hs$$

GENETIC LOAD AND SELECTION COEFFICIENTS

$$\text{REALIZED GENETIC LOAD} = \sum_{i=1}^n s_i q_i^2 + 2 \sum_{i=1}^n q_i [1 - q_i] h_i s_i$$

(n loci)

See Bertorelle et al 2022

INDIVIDUAL GENETIC LOAD PARTITION

$$\text{TOTAL LOAD} = \text{REALIZED LOAD} + \text{MASKED LOAD}$$

See Bertorelle et al 2022

INDIVIDUAL GENETIC LOAD PARTITION

TOTAL LOAD = **REALIZED LOAD** + **MASKED LOAD**

$$\sum_{i=1}^{L(hom)} s_i + \sum_{i=1}^{L(het)} 0.5 s_i =$$

See Bertorelle et al 2022

INDIVIDUAL GENETIC LOAD PARTITION

TOTAL LOAD = **REALIZED LOAD** + **MASKED LOAD**

$$\sum_{i=1}^{L(hom)} s_i + \sum_{i=1}^{L(het)} 0.5 s_i = \sum_{i=1}^{L(hom)} s_i + \sum_{i=1}^{L(het)} h_i s_i$$

See Bertorelle et al 2022

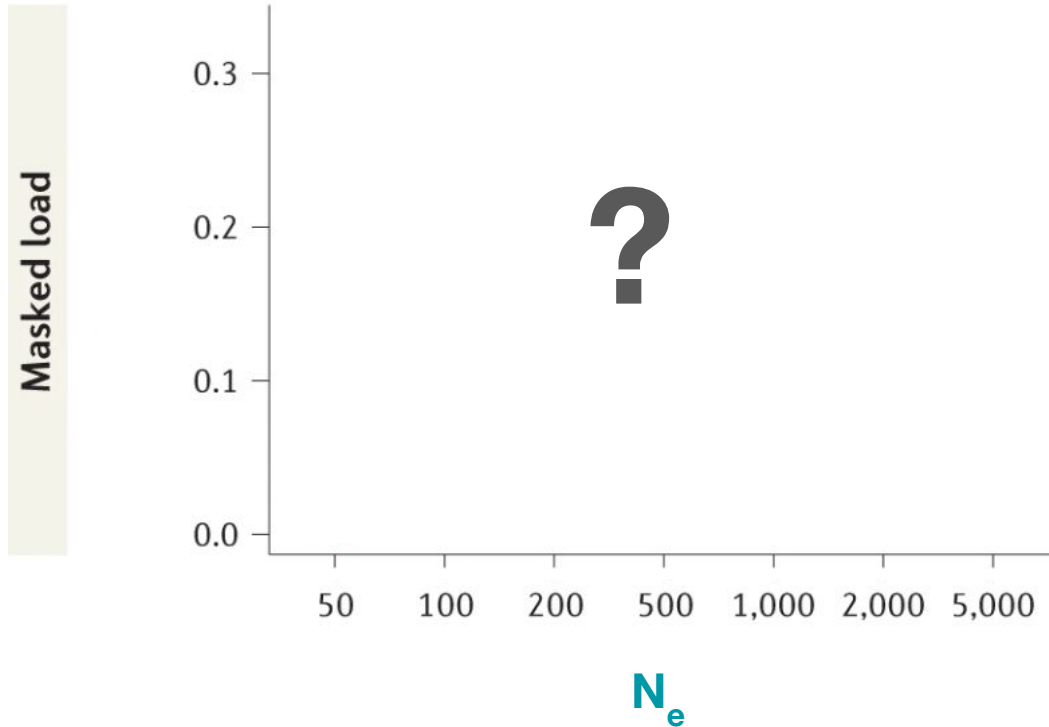
INDIVIDUAL GENETIC LOAD PARTITION

TOTAL LOAD **=** **REALIZED LOAD** **+** **MASKED LOAD**

$$\sum_{i=1}^{L(hom)} s_i + \sum_{i=1}^{L(het)} 0.5 s_i = \sum_{i=1}^{L(hom)} s_i + \sum_{i=1}^{L(het)} h_i s_i + \sum_{j=1}^{L(het)} (0.5 - h_j) s_j$$

See Bertorelle et al 2022

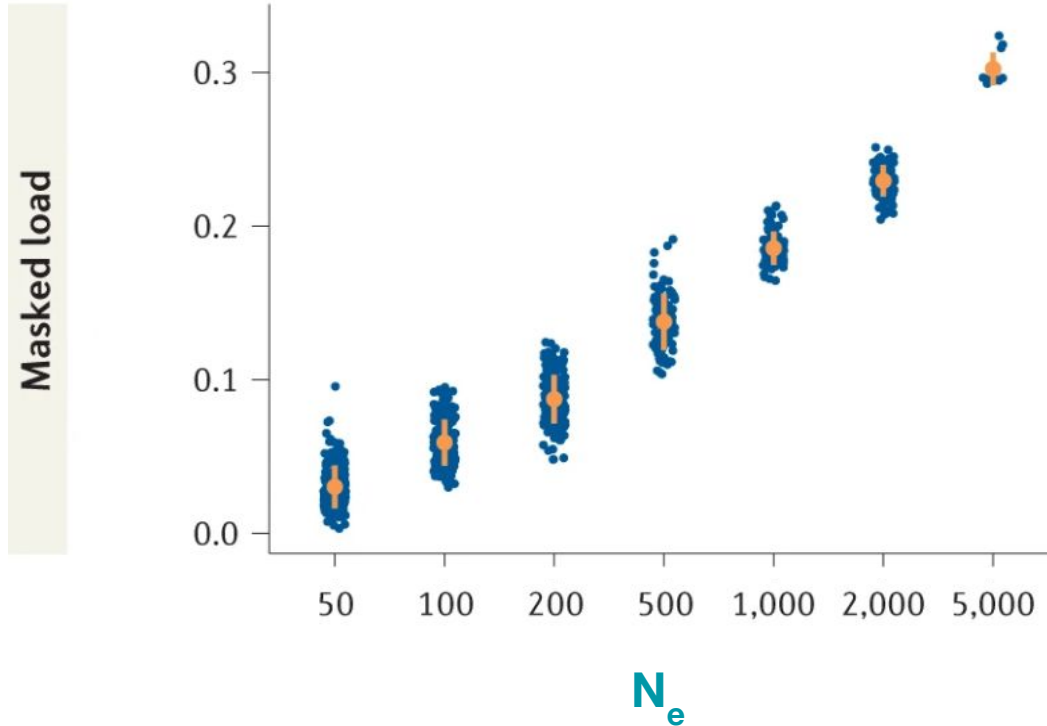
INDIVIDUAL GENETIC LOAD AND POPULATION SIZE



$$\sum_{j=1}^{L(het)} (0.5 - h_j) s_j$$

See Bertorelle et al 2022

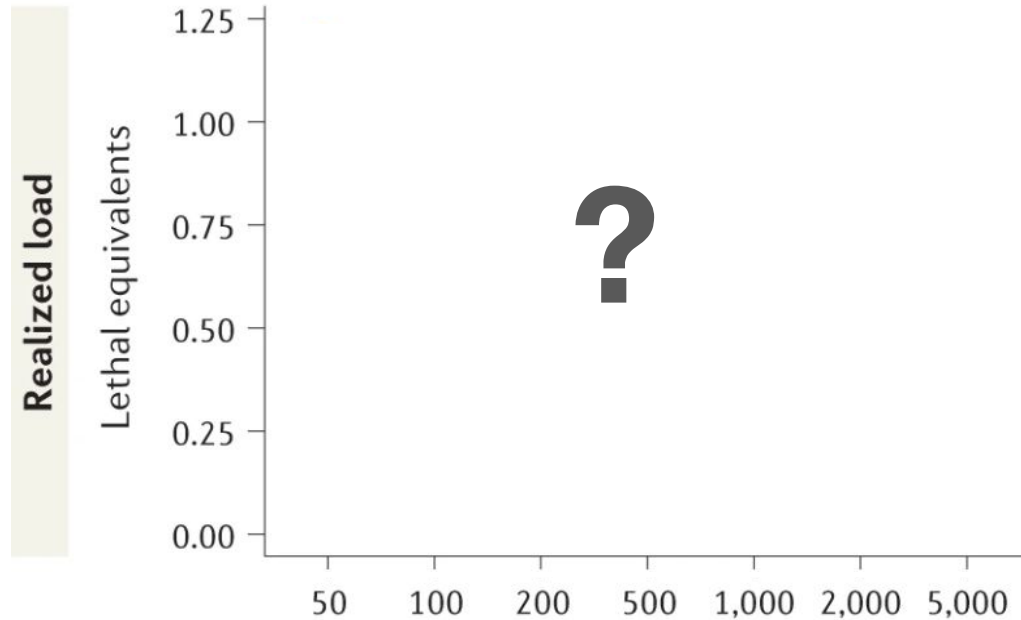
INDIVIDUAL GENETIC LOAD AND POPULATION SIZE



$$\sum_{j=1}^{L(het)} (0.5 - h_j) s_j$$

See Bertorelle et al 2022

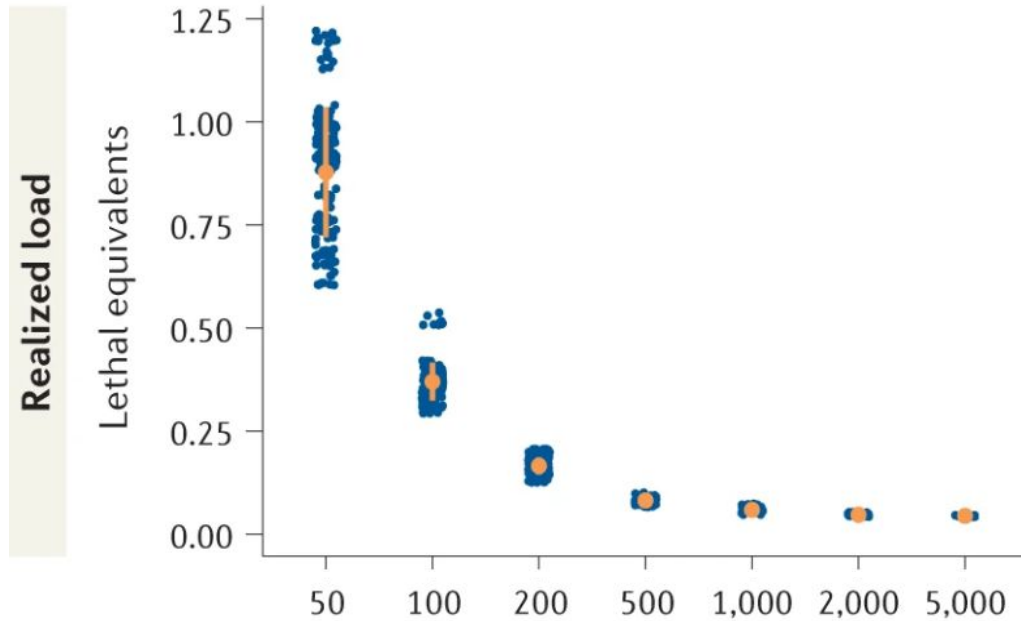
INDIVIDUAL GENETIC LOAD AND POPULATION SIZE



$$\sum_{i=1}^{L(hom)} s_i + \sum_{i=1}^{L(het)} h_i s_i$$

See Bertorelle et al 2022

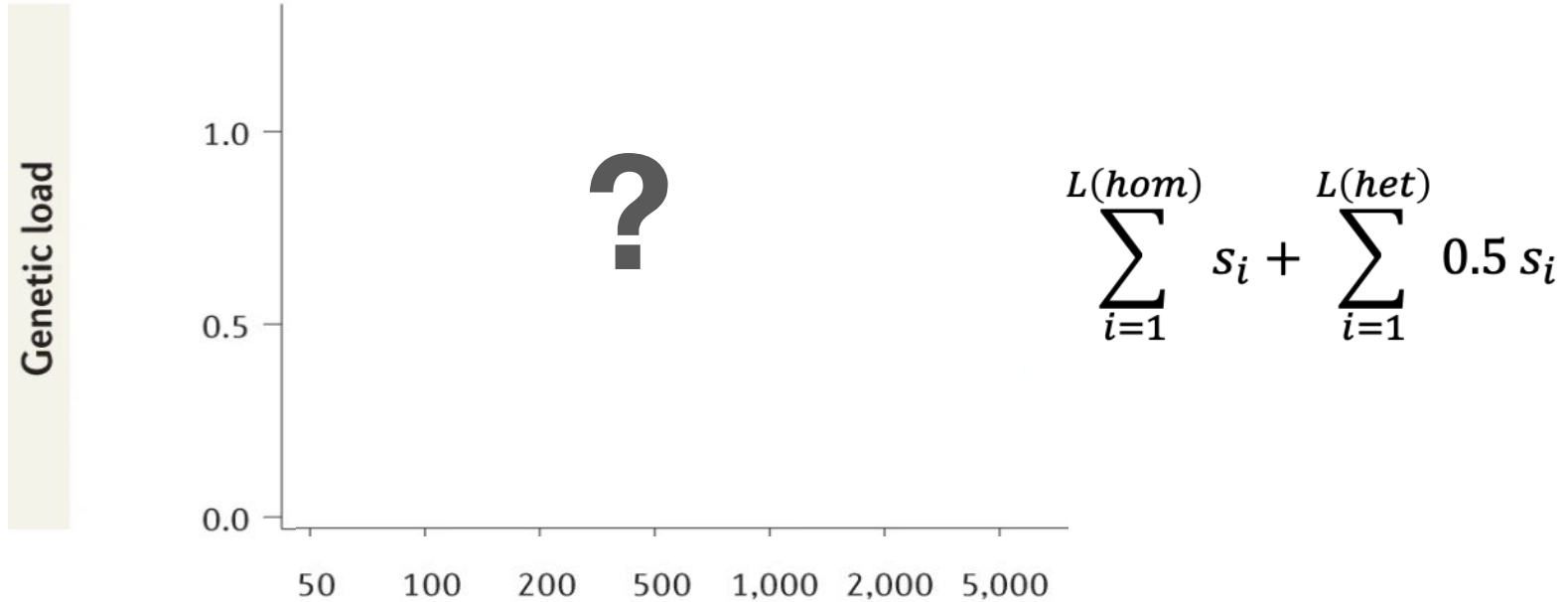
INDIVIDUAL GENETIC LOAD AND POPULATION SIZE



$$\sum_{i=1}^{L(hom)} s_i + \sum_{i=1}^{L(het)} h_i s_i$$

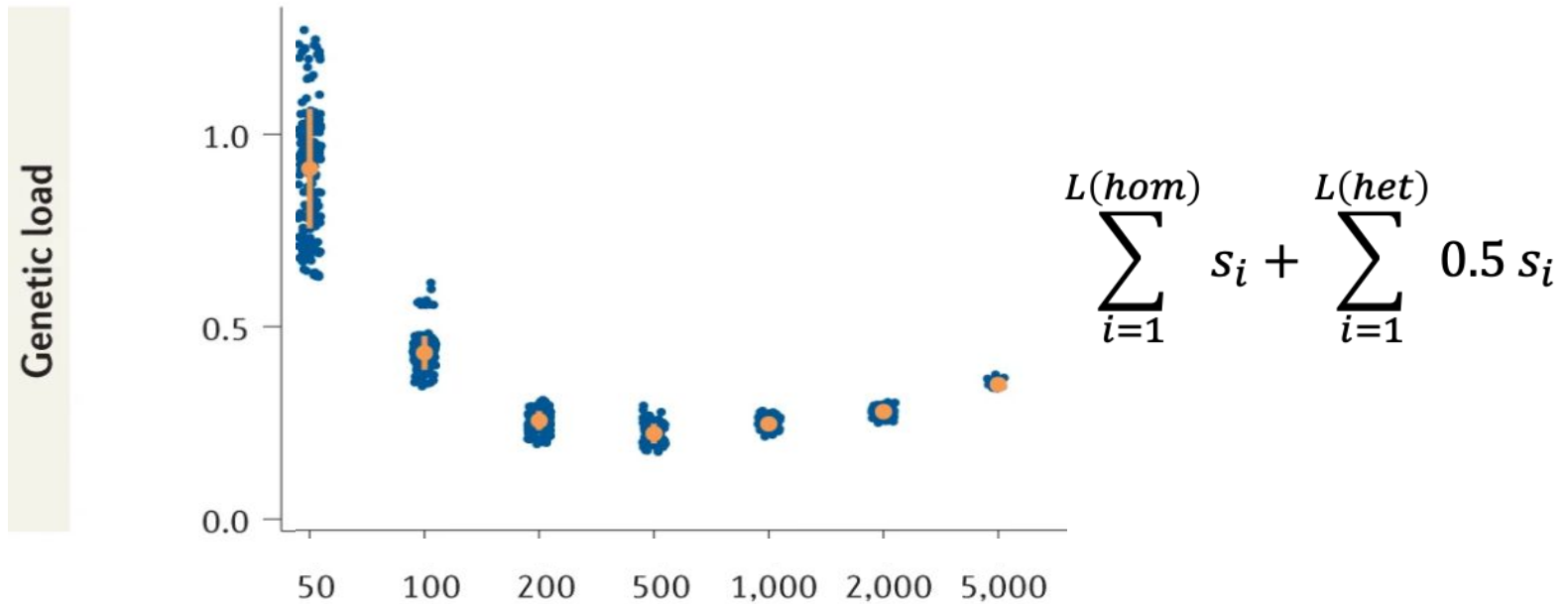
See Bertorelle et al 2022

INDIVIDUAL GENETIC LOAD AND POPULATION SIZE



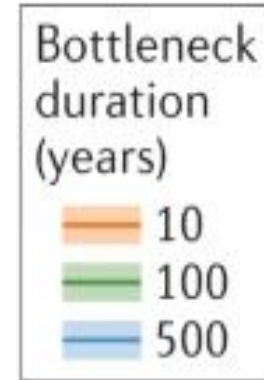
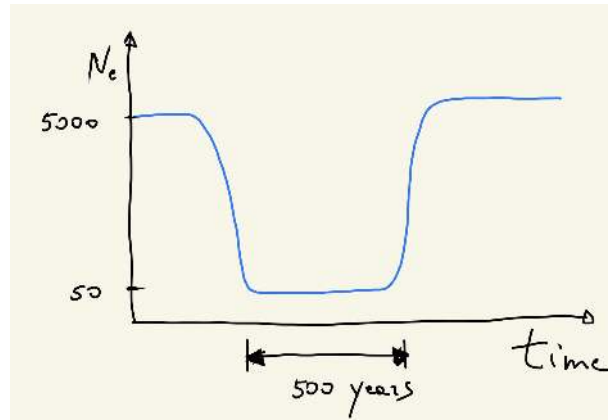
See Bertorelle et al 2022

INDIVIDUAL GENETIC LOAD AND POPULATION SIZE



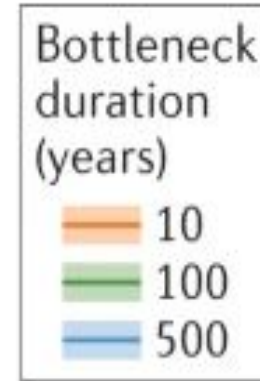
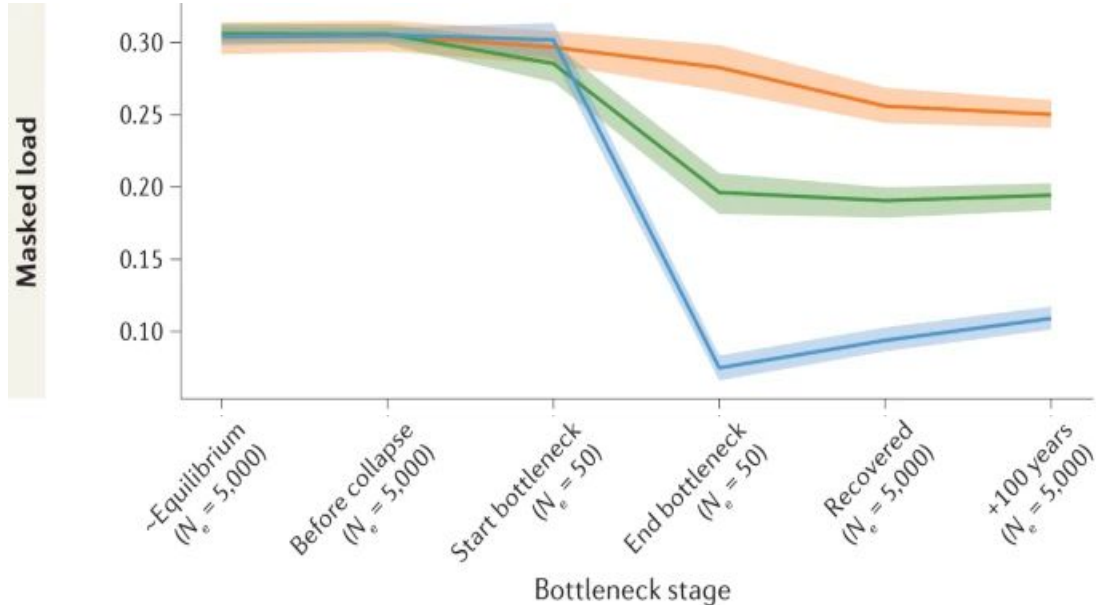
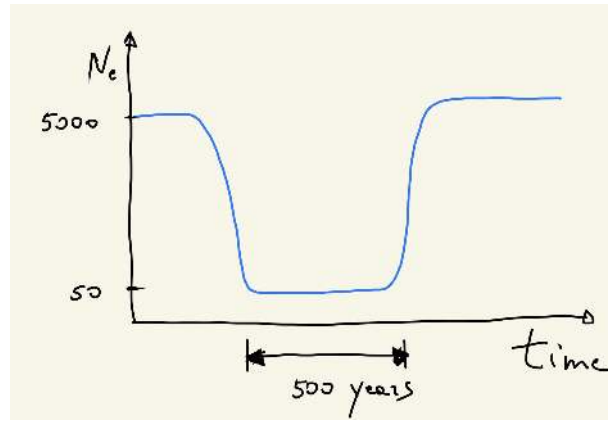
See Bertorelle et al 2022

INDIVIDUAL GENETIC LOAD AND POPULATION SIZE CHANGES



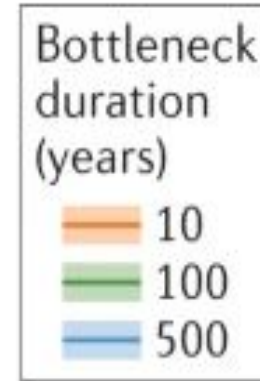
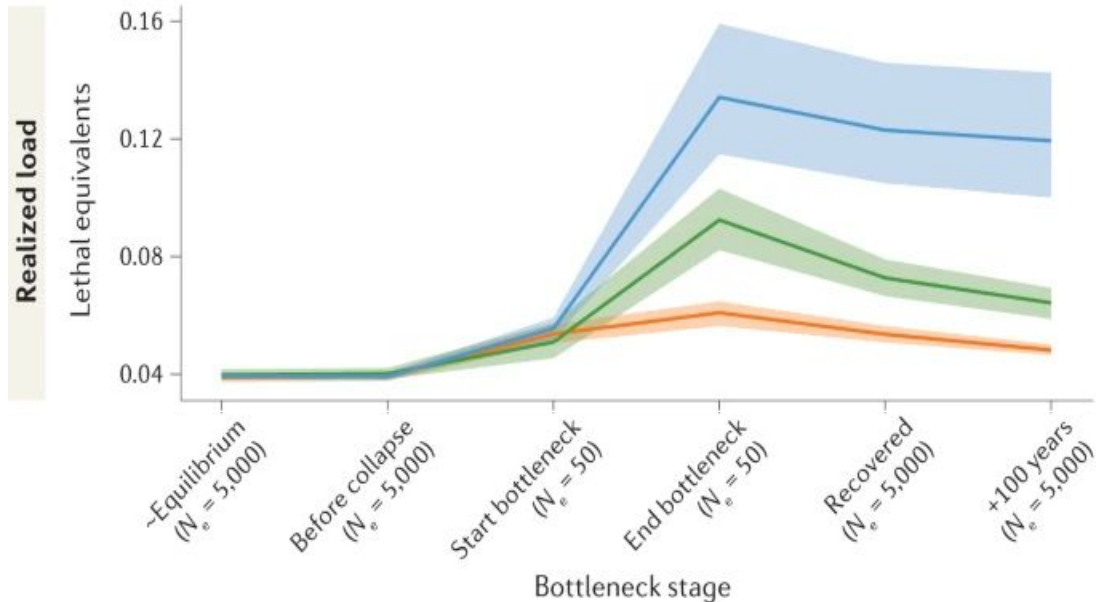
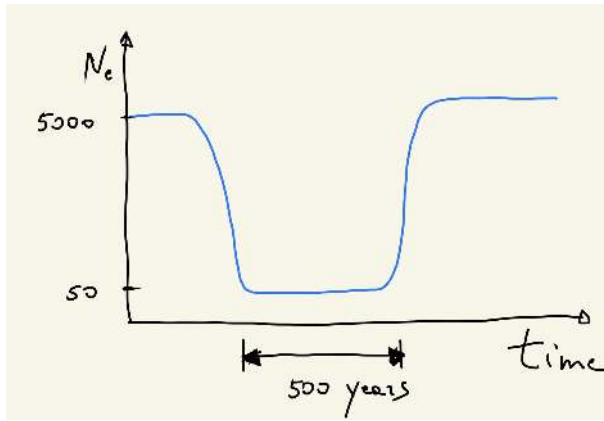
See Bertorelle et al 2022

INDIVIDUAL GENETIC LOAD AND POPULATION SIZE CHANGES



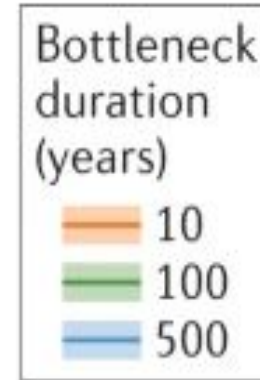
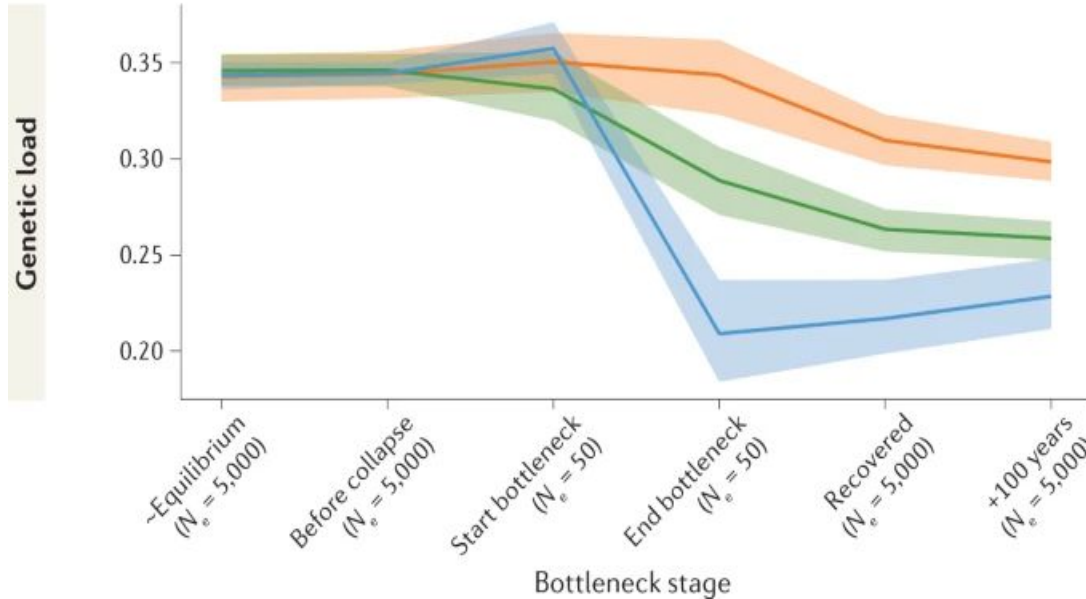
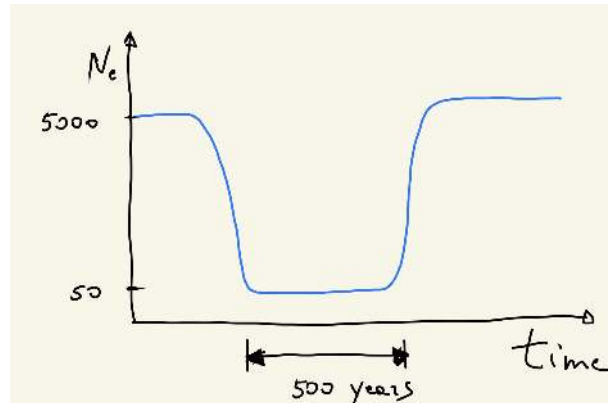
See Bertorelle et al 2022

INDIVIDUAL GENETIC LOAD AND POPULATION SIZE CHANGES



See Bertorelle et al 2022

INDIVIDUAL GENETIC LOAD AND POPULATION SIZE CHANGES



See Bertorelle et al 2022

ONE “LITTLE” PROBLEM:
HOW TO GET THE SELECTION COEFFICIENTS?

ONE “LITTLE” PROBLEM: HOW TO GET THE SELECTION COEFFICIENTS?

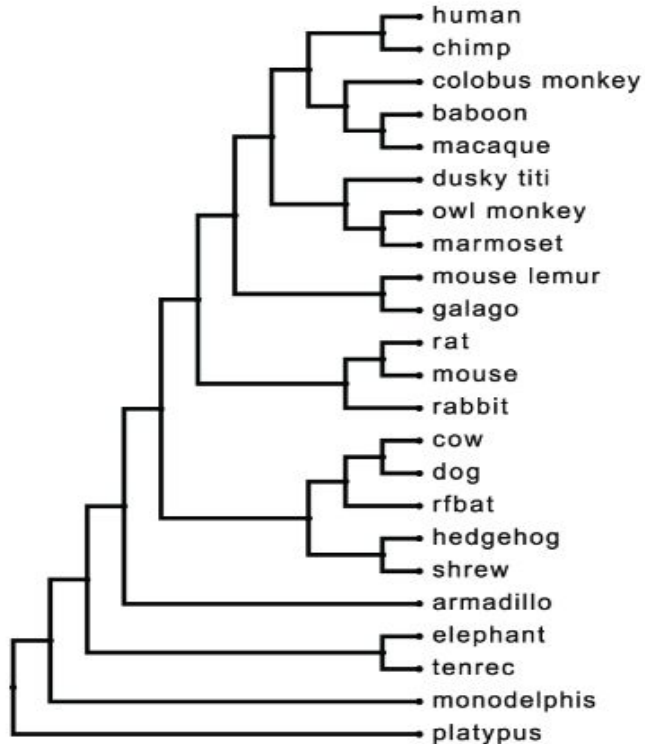
PREDICTIONS BASED ON:

- 1) **CONSERVATION SCORES** through long evolutionary times
(GERP, PhyloP, GPN-MSA*)

Relies on **alignments of genomes** (e.g., Cactus)

CONSERVATION SCORES

Phylogenetic Tree



Multiple Sequence Alignment

human	AATACGG	A	ACTTCATTTCATT
chimp	AATATGG	A	ACTTCATTTCATT
colobus monkey	AGTATGG	A	ACTTCATTTCATT
baboon	AGTATGG	A	ACTTCATTTCATT
macaque	AGTATGG	A	ACTTCATTTCATT
dusky titi	AGTATGG	A	ACTTCATTTCATT
owl monkey	AGTATGG	A	ACTTCATTTCATT
marmoset	AGTATGG	A	ACTTCATTTCATT
mouse lemur	AGTACGG	A	ACTTCATTTCATT
galago	AGTACGG	A	ACTTCATTTCATT
rat	AGTATGG	A	ACATCGTTCATT
mouse	AGTATGG	A	ACATCTTTCATT
rabbit	AGTATGG	A	ACATCATTTCATT
cow	AGTATGG	A	ACATCATTTCATT
dog	AGTACGG	A	ACATCATTTCATT
rfbat	AGTATGG	A	ACATCGTTCATT
hedgehog	AGTATGG	A	ACATCATTTCATT
shrew	AGTATGG	G	ACATCCTTCATT
armadillo	-----	-	-----
elephant	AGTATGG	A	ACATCGTTCATT
tenrec	AGTATGG	A	ACATCGTTCATT
monodelphis	AGTATGG	G	ACATCTTTCATT
platypus	AGTATGG	A	ACGTCATTTCATT

ONE “LITTLE” PROBLEM: HOW TO GET THE SELECTION COEFFICIENTS?

PREDICTIONS BASED ON:

- 1) CONSERVATION SCORES through long evolutionary times
(GERP, PhyloP, GPN-MSA*)

Relies on alignments of genomes (e.g., Cactus,...)

- 2) **Impact of changes** on coding sequence

Relies on reference genome **annotation** (SNPEff)

SnpEff

Genetic variant annotation and effect prediction toolbox.

CODING SEQUENCE CHANGES

Cingolani et al 2012

CODING SEQUENCE CHANGES

SnpEff

Genetic variant annotation and effect prediction toolbox.

putative impact

description of consequence
exon_loss_variant, stop_lost,
frameshift_variant

impact

estimation of level of impact
HIGH, LOW, MODERATE

CODING SEQUENCE CHANGES

SnpEff

Genetic variant annotation and effect prediction toolbox.

putative impact

description of consequence
exon_loss_variant, stop_lost,
frameshift_variant

impact

estimation of level of impact
HIGH, LOW, MODERATE

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO
chr1	123456	.	C	A	.	.	ANN=A ...
chr1	234567	.	A	G,T	.	.	ANN=G ... , T ...

t
C

ANN = Annotation aka effect or consequence

CODING SEQUENCE CHANGES

SnpEff

Genetic variant annotation and effect prediction toolbox.

putative impact

description of consequence
exon_loss_variant, stop_lost,
frameshift_variant

impact

estimation of level of impact
HIGH, LOW, MODERATE

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO
chr1	123456	.	C	A	.	.	ANN=A ...
chr1	234567	.	A	G,T	.	.	ANN=G ... , T ...

t
C

ANN = Annotation aka effect or consequence

Type (alphabetical order)	Count	Percent
HIGH	39	0.04%
LOW	395	0.407%
MODERATE	239	0.246%
MODIFIER	96,412	99.307%

WHAT IS THE DIFFERENCE BETWEEN
CONSERVATION GENETICS AND GENOMICS?

WHAT IS THE FIRST INGREDIENT OF
CONSERVATION GENOMICS?

HOW CAN GENETICS INFORM
CONSERVATION BIOLOGY?

**IN WHICH APPLICATION GENOMICS IS
A GAME CHANGER?**

with a close up on two topics:

Genetic load

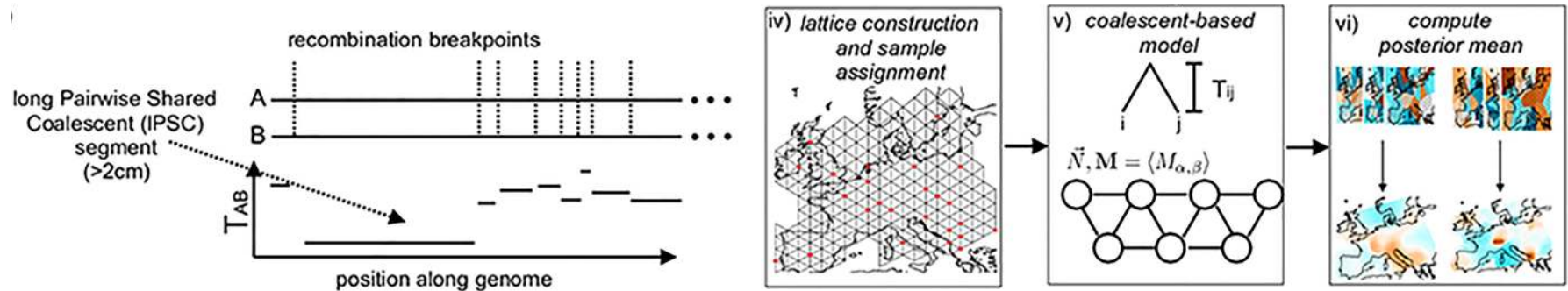
ROHs

WHAT AM I DOING TO MAKE USE OF
GENOMICS IN CONSERVATION?

USING EXTENDED HAPLOTYPE INFORMATION

Which is getting more affordable with long-range sequencing

Sharing **Identity-By-Descent (IBD)** blocks to infer recent migration (Al-Asadi 2019).

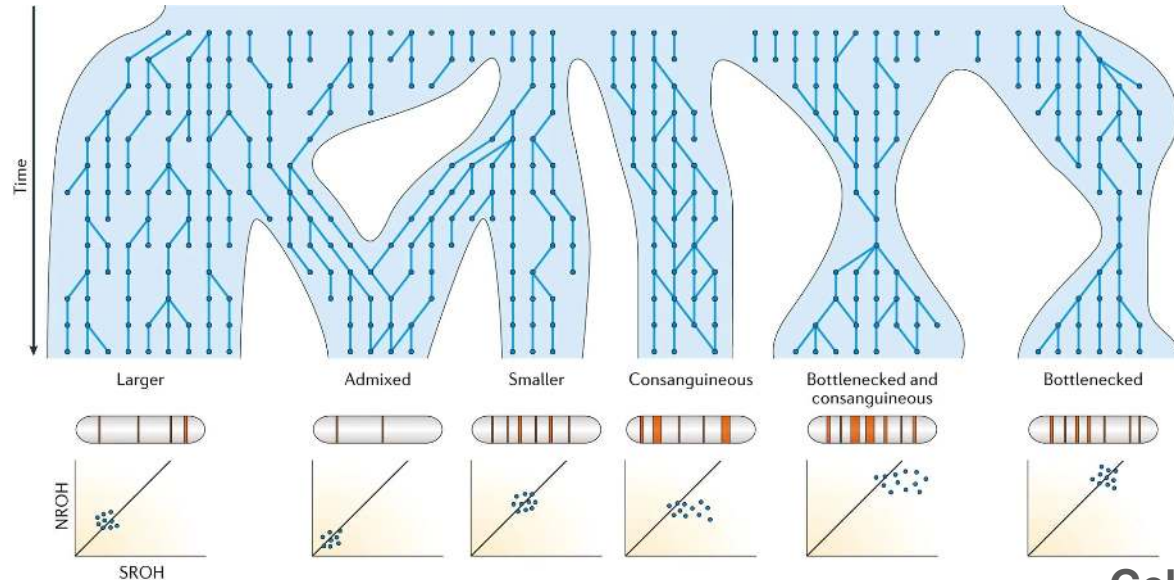


(but see AI non IBD block based alternative by Andy Kern (Smith et al 2024))

USING EXTENDED HAPLOTYPE INFORMATION

Which is getting more affordable with long-range sequencing

Using **Runs-Of-Homozygosity (ROH)** distribution length to infer recent inbreeding and population size changes



Ceballos et al 2018

WHAT IS THE DIFFERENCE BETWEEN
CONSERVATION GENETICS AND GENOMICS?

WHAT IS THE FIRST INGREDIENT OF
CONSERVATION GENOMICS?

HOW CAN GENETICS INFORM
CONSERVATION BIOLOGY?

IN WHICH APPLICATION GENOMICS IS
A GAME CHANGER?

with a close up on two topics:

Genetic load

ROHs

WHAT AM I DOING TO MAKE USE OF
GENOMICS IN CONSERVATION?





**WHAT IS A
RUN OF HOMOZYGOSITY?**

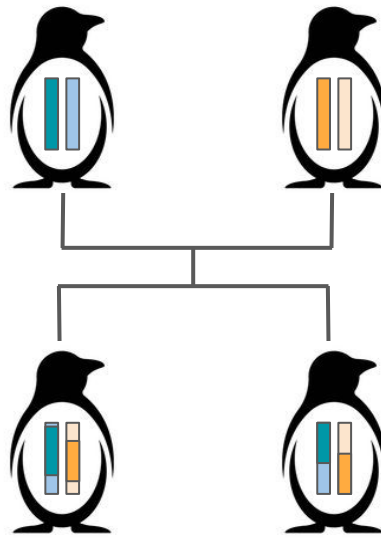


**WHAT IS A
RUN OF HOMOZYGOSITY?**

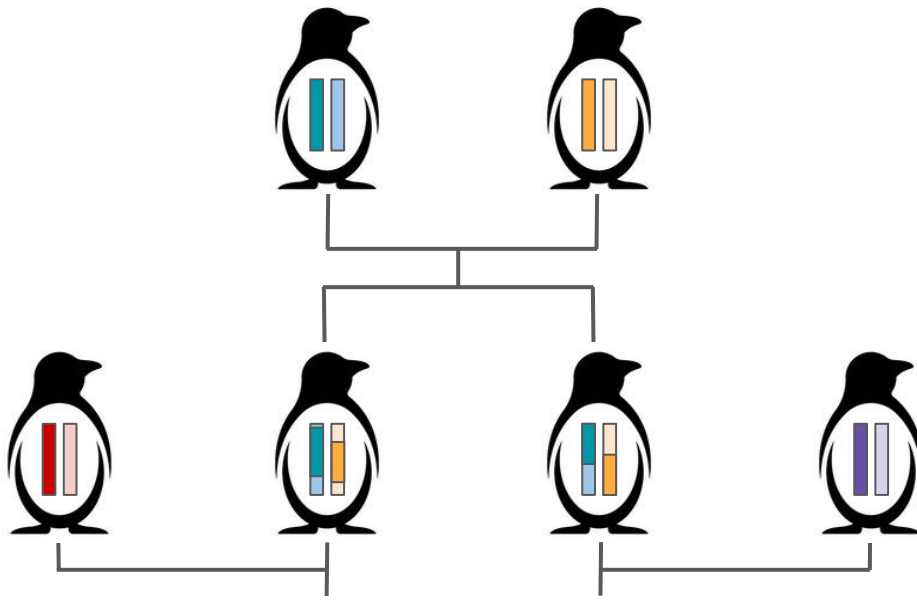


**WHAT IS A
RUN OF HOMOZYGOSITY?**

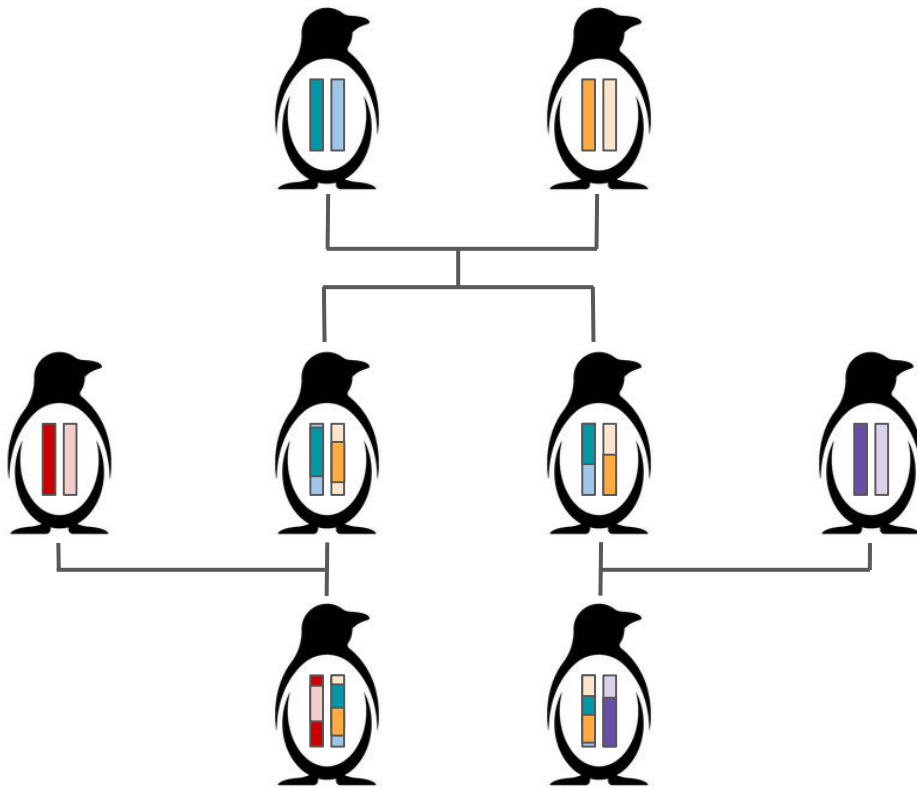
**WHAT IS A
RUN OF HOMOZYGOSITY?**



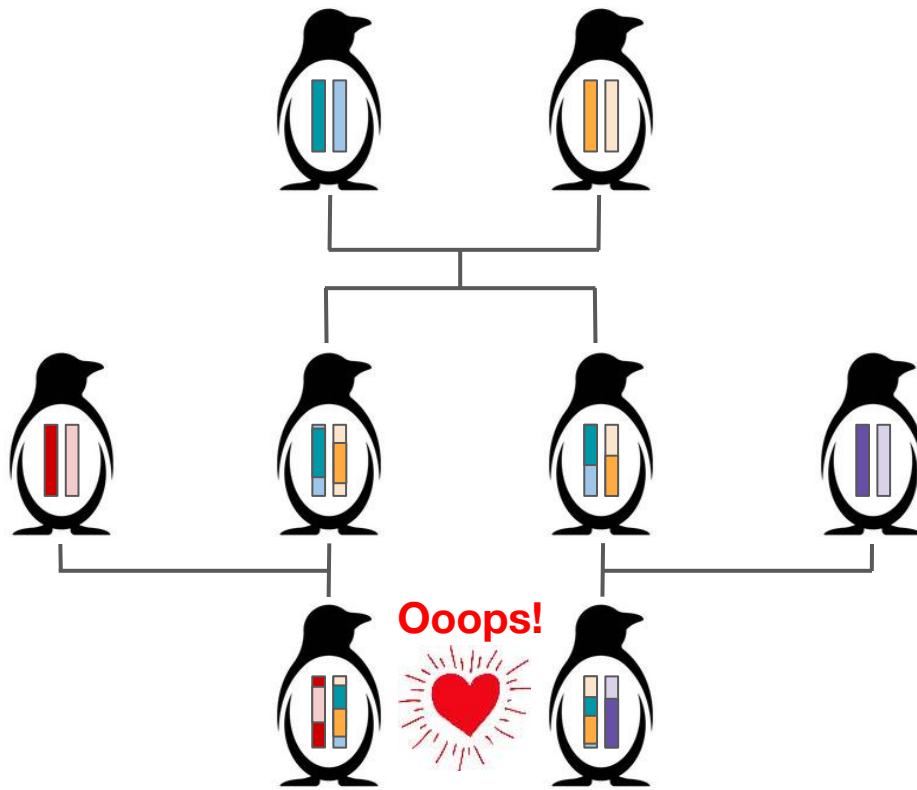
WHAT IS A RUN OF HOMOZYGOSITY?



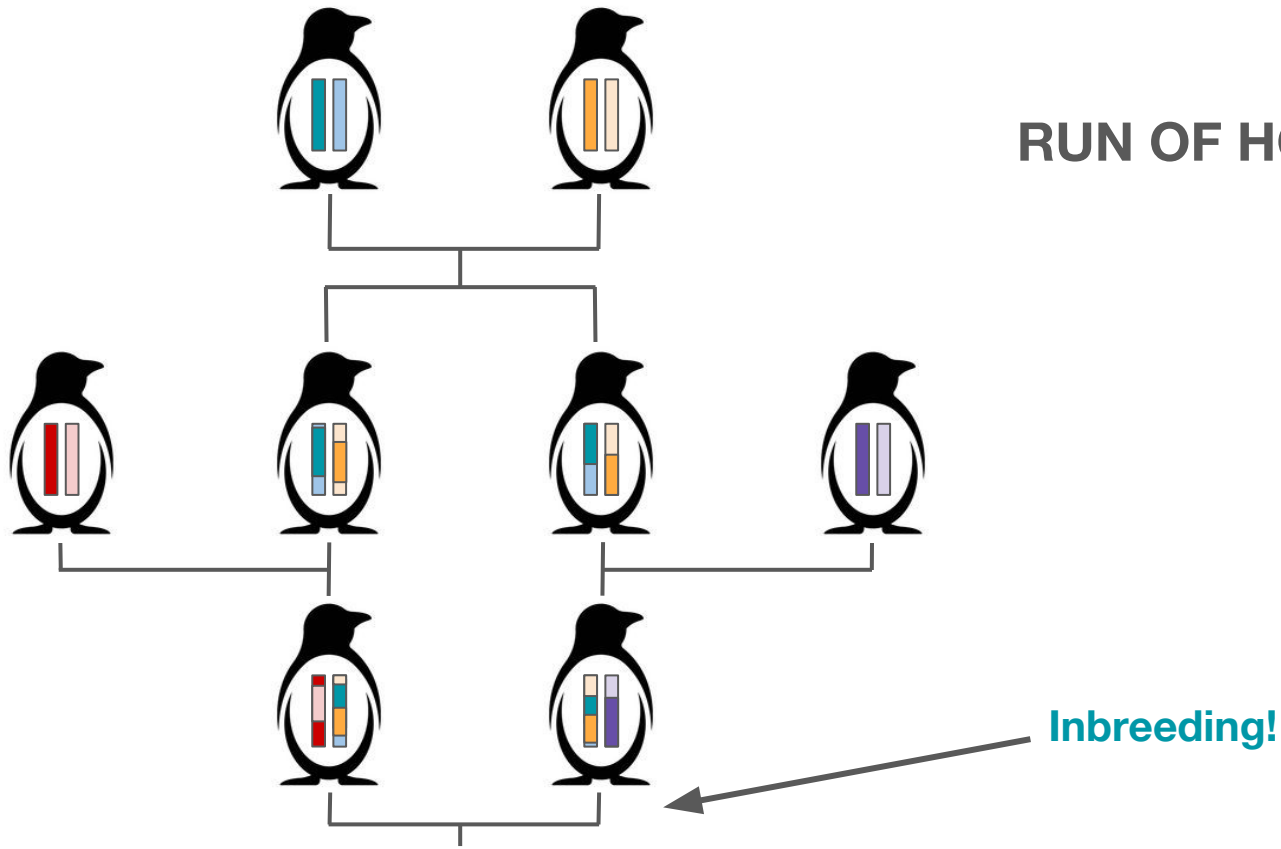
WHAT IS A RUN OF HOMOZYGOSITY?



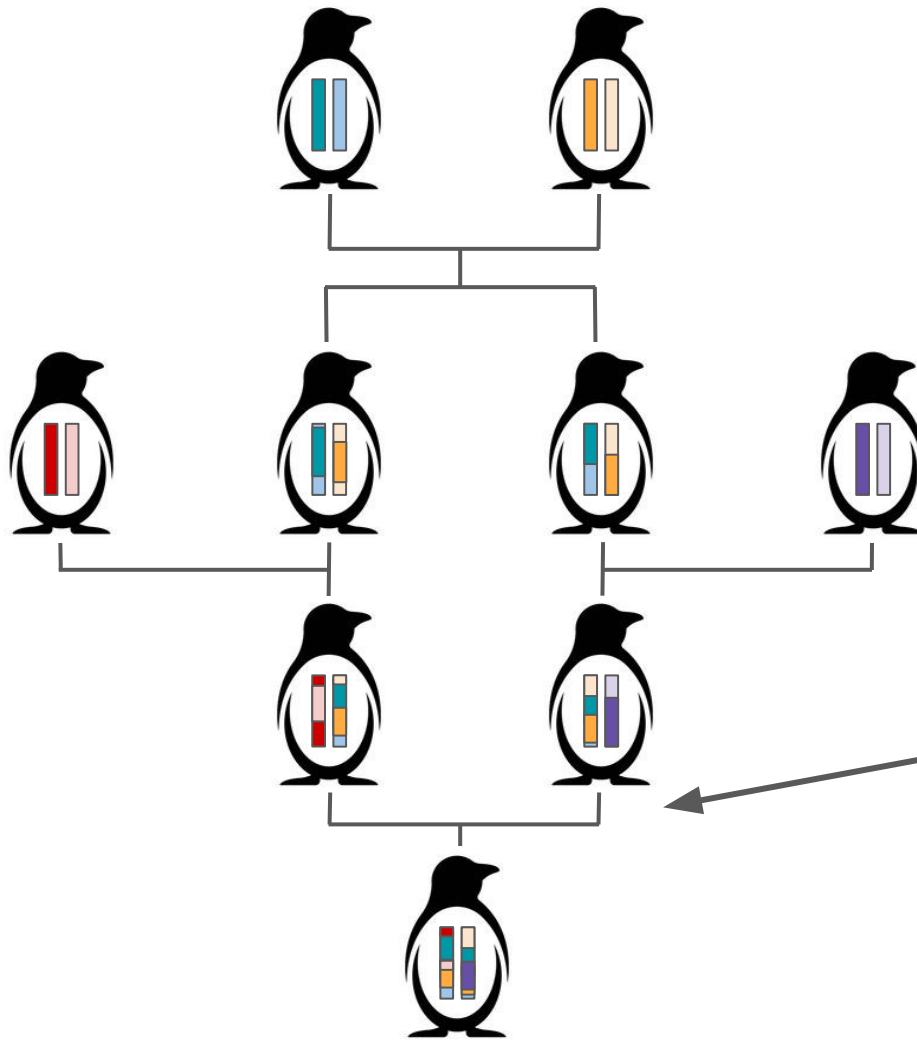
WHAT IS A RUN OF HOMOZYGOSITY?



WHAT IS A RUN OF HOMOZYGOSITY?

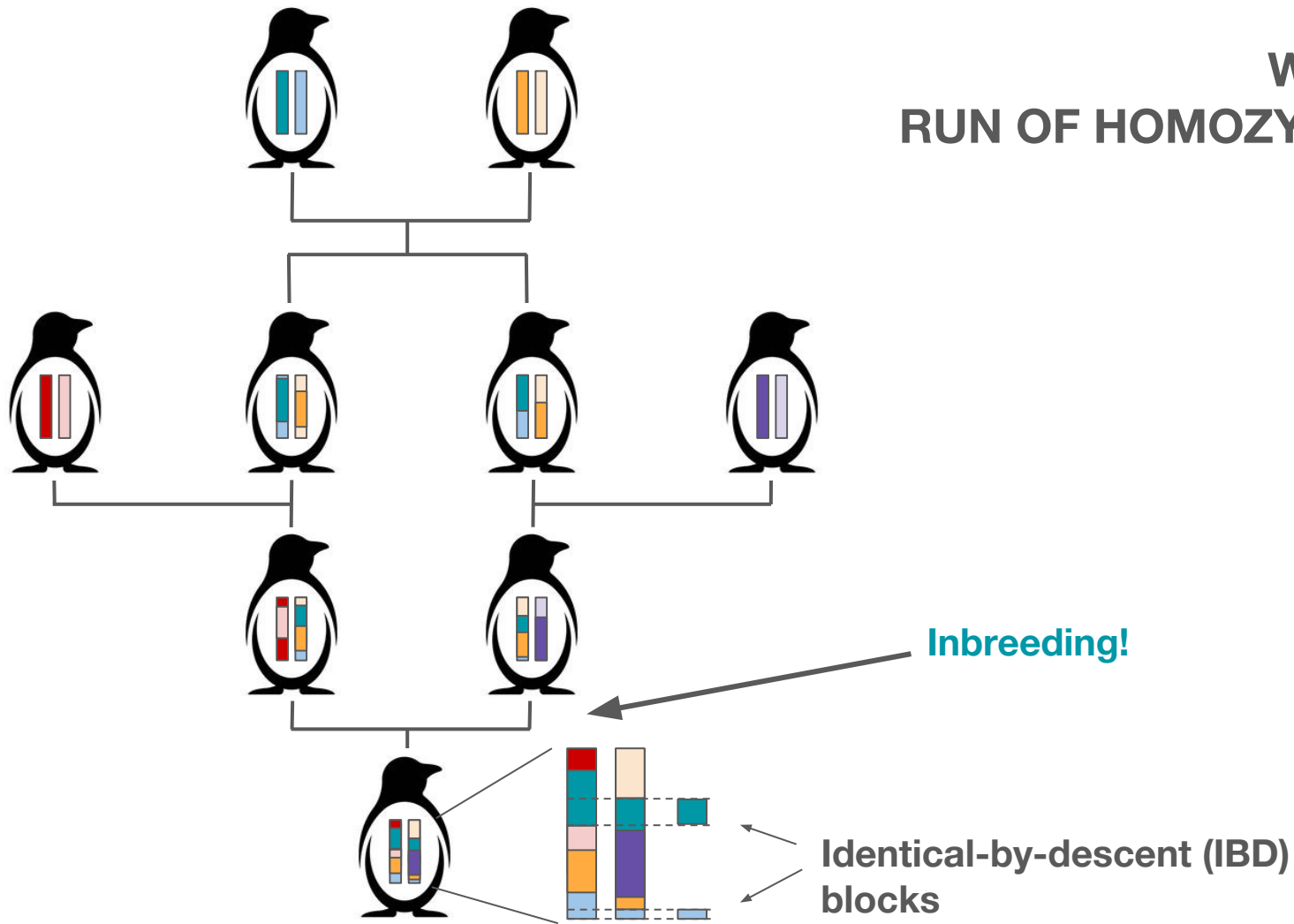


WHAT IS A RUN OF HOMOZYGOSITY?



Inbreeding!

WHAT IS A RUN OF HOMOZYGOSITY?





WHAT'S THE FATE OF A RUN OF HOMOZYGOSITY?



$t_{\text{inbreeding}}$



Mating with someone
not closely related



$t_{\text{inbreeding}}$

t_1

WHAT'S THE FATE OF A RUN OF HOMOZYGOSITY?



Mating with someone
not closely related



$t_{\text{inbreeding}}$

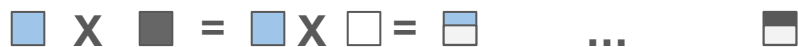
t_1

t_2

WHAT'S THE FATE OF A RUN OF HOMOZYGOSITY?



WHAT'S THE FATE OF A RUN OF HOMOZYGOSITY?



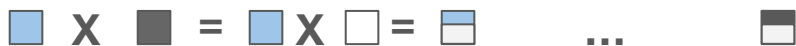
Getting shorter and shorter

tinbreeding

 t_1 t_2 t_x



Mating with someone
not closely related



$t_{\text{inbreeding}}$

t_1

t_2

t_x

WHAT'S THE FATE OF A RUN OF HOMOZYGOSITY?

Recombination rate (ρ)

Chromosome length (L)

Time (t)

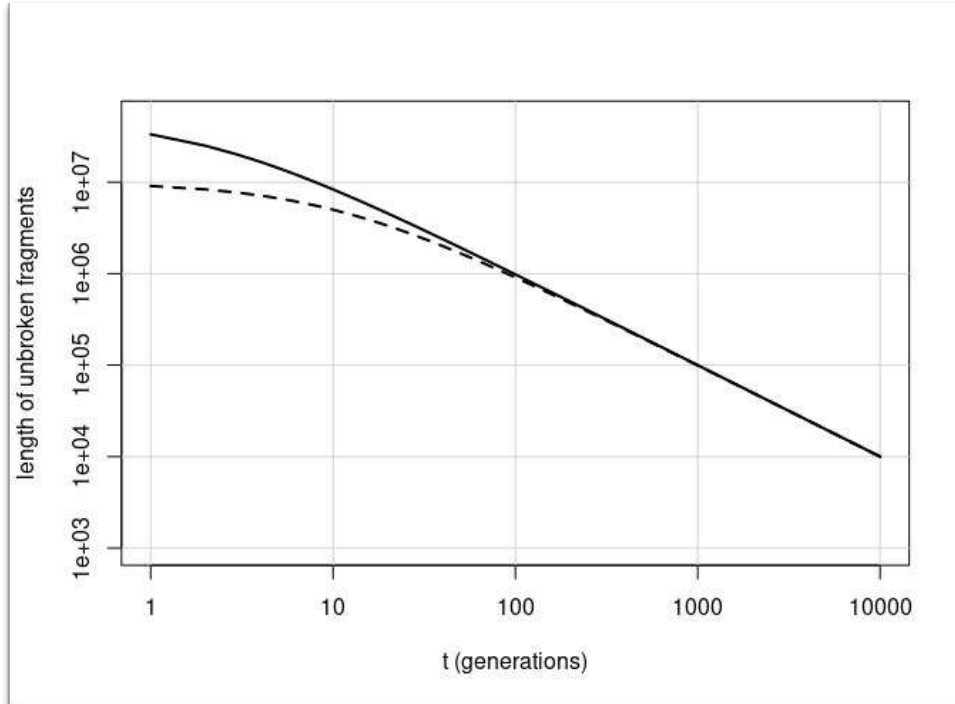
Expected number of breaks

ρLt

Expected unbroken fragment length

$L/(1+\rho Lt)$

WHAT'S THE FATE OF A RUN OF HOMOZYGOSITY?



Recombination rate (ρ): $1 \times 10^{-8} \text{ gen}^{-1} \text{ bp}^{-1}$

Chromosome length (L_0):

5×10^7 bp (solid) or 1×10^7 bp (dashed)

Time (t): 10000 generations

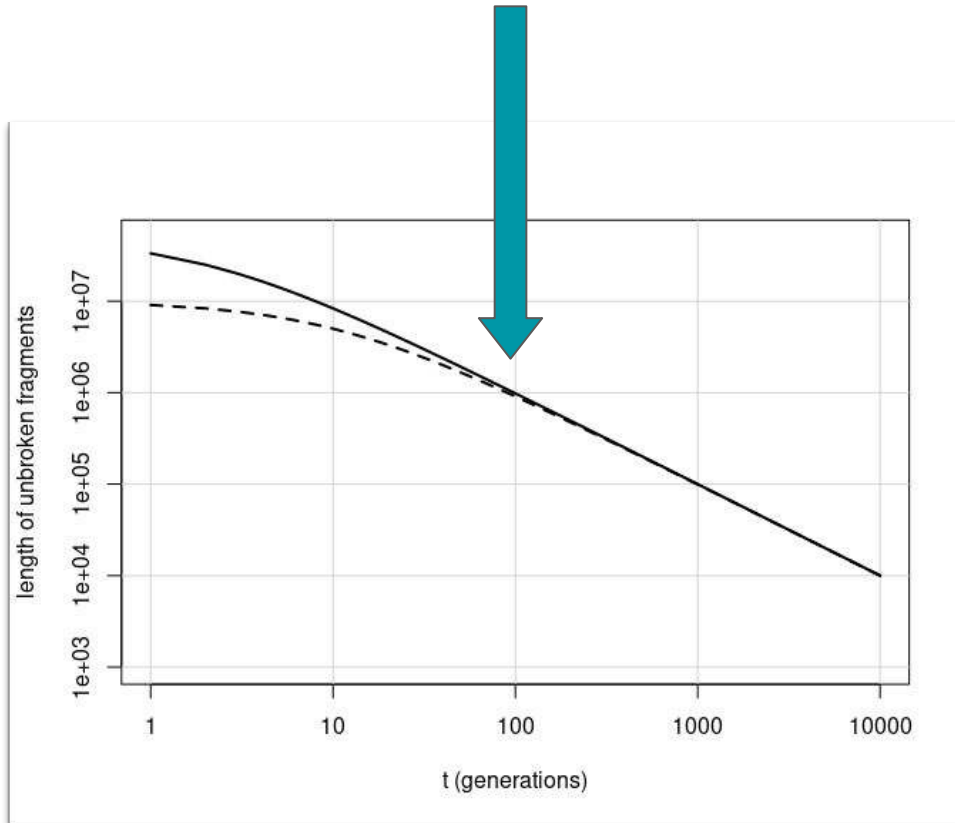
Expected number of breaks:

$$\rho L_0 t$$

Expected unbroken fragment length

$$L_t = L_0 / (1 + \rho L_0 t)$$

WHAT'S THE FATE OF A RUN OF HOMOZYGOSITY?



Recombination rate (ρ): $1 \times 10^{-8} \text{ gen}^{-1} \text{ bp}^{-1}$

Chromosome length (L_0):

5×10^7 bp (solid) or 1×10^7 bp (dashed)

Time (t): 10000 generations

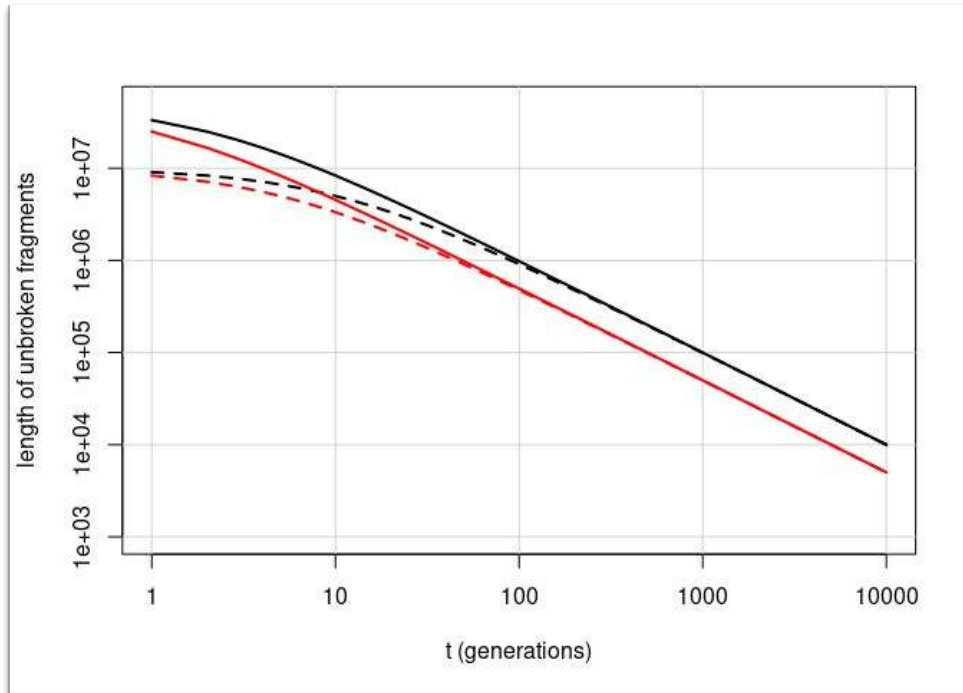
Expected number of breaks:

$$\rho L_0 t$$

Expected unbroken fragment length

$$L_t = L_0 / (1 + \rho L_0 t)$$

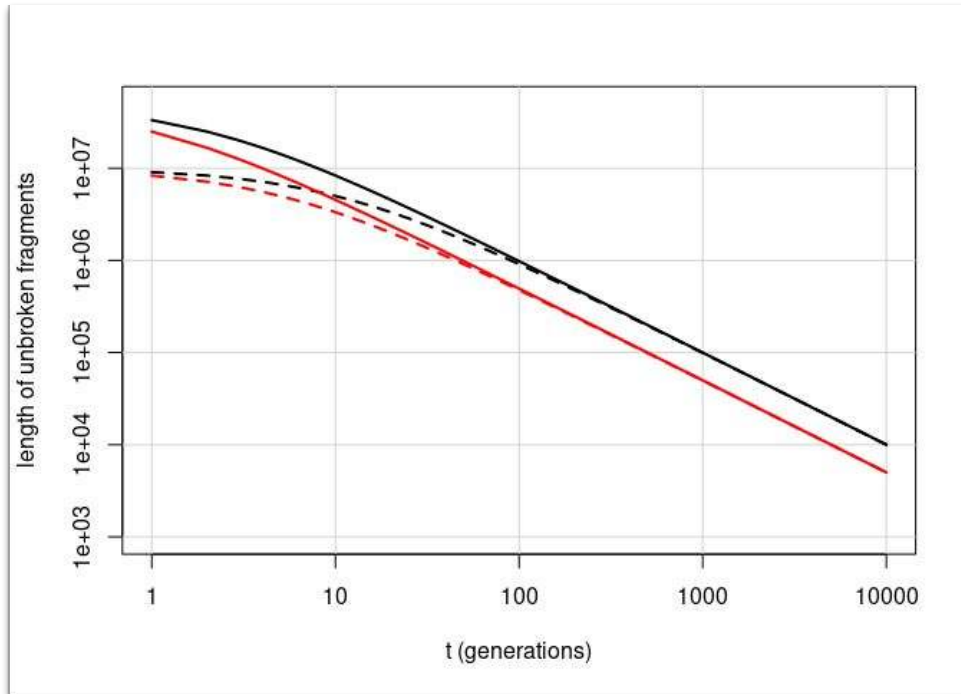
WHAT'S THE FATE OF A RUN OF HOMOZYGOSITY?



For **two lineages (ROH)**, the expected length of an Identity by descent (IBD) block is simply

$$L_t = L_0 / (1 + 2\rho L_0 t)$$

WHAT'S THE FATE OF A RUN OF HOMOZYGOSITY?



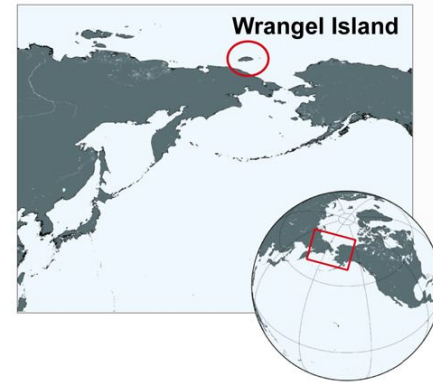
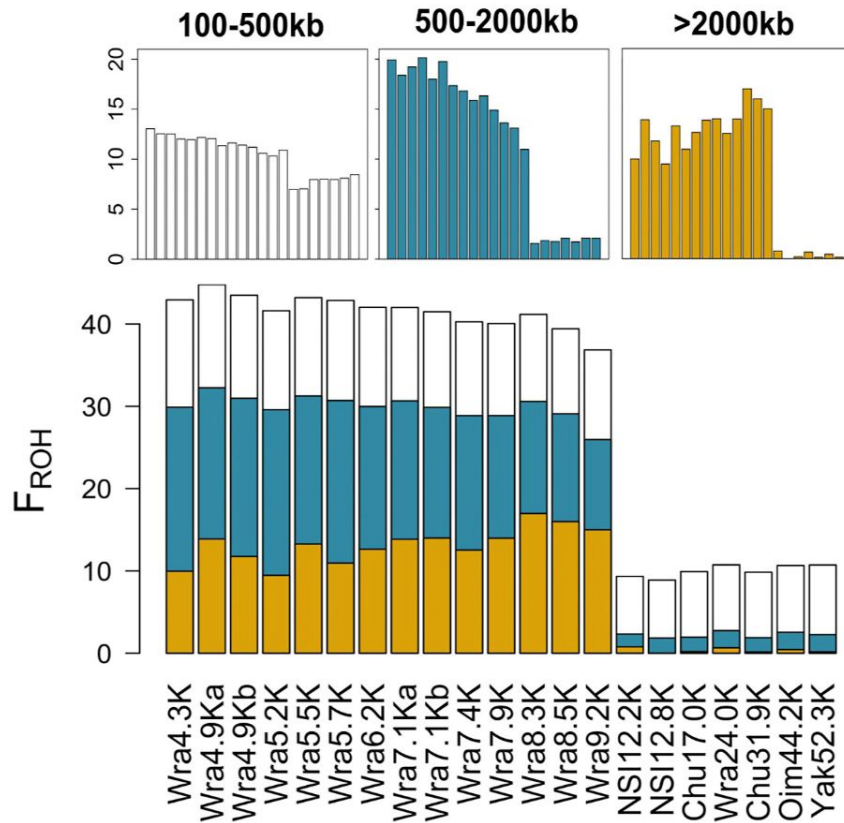
For **two lineages (ROH)**, the expected length of an Identity by descent (IBD) block is simply

$$L_t = L_0 / (1 + 2\rho L_0 t)$$

From this we can get an expectation on the time of the inbreeding loop (and on the population size at that time*)

$$t = 1 / (2\rho L_t)$$

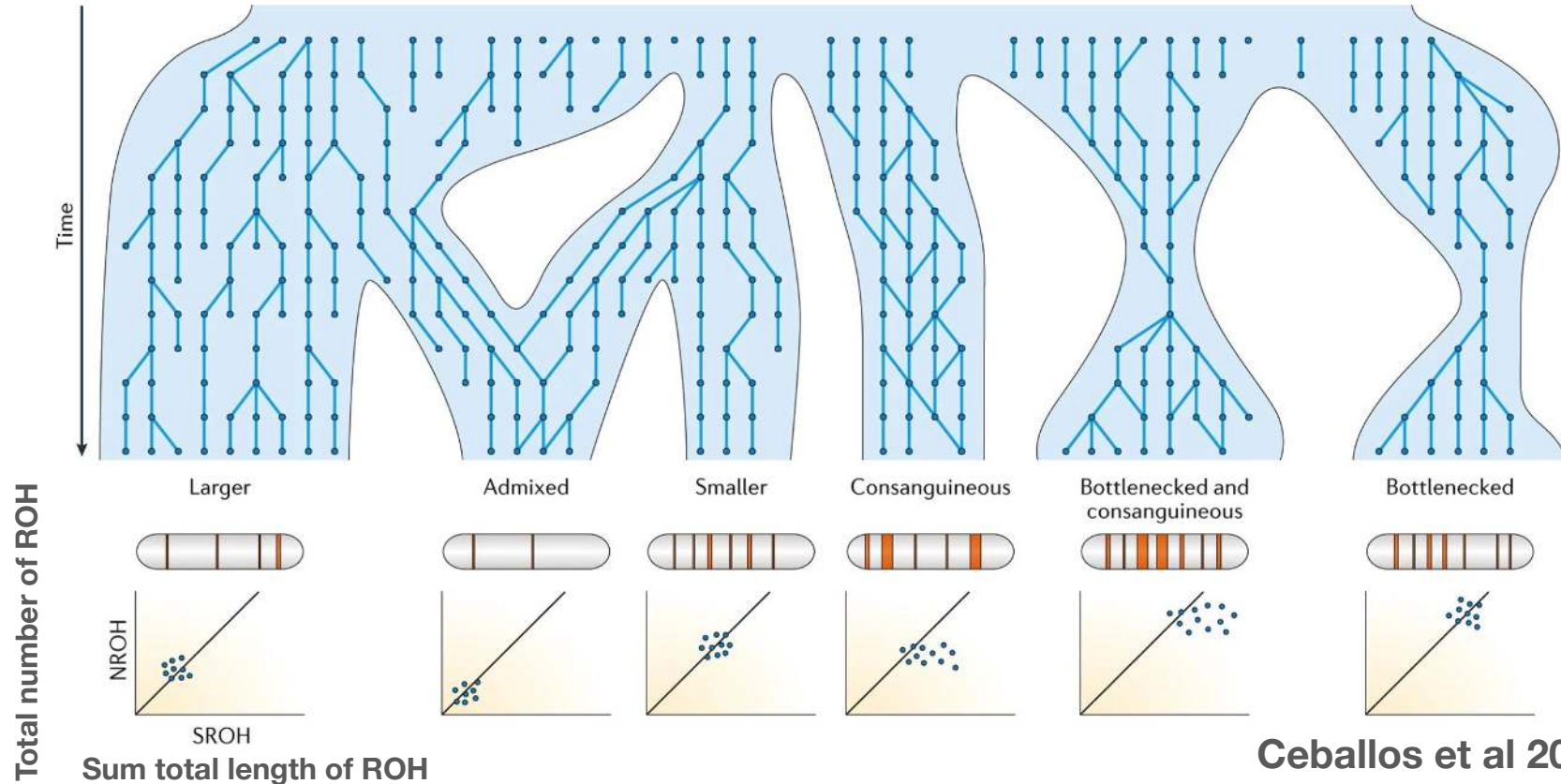
ROHS CAN INFORM ABOUT PAST **INBREEDING EVENTS**



Infer when most recent
inbreeding event happened by
ROHs of maximum length*

Dehasque et al 2024

ROHS CAN INFORM ABOUT PAST **DEMOGRAPHY**



WHAT IS THE DIFFERENCE BETWEEN
CONSERVATION GENETICS AND GENOMICS?

WHAT IS THE FIRST INGREDIENT OF
CONSERVATION GENOMICS?

HOW CAN GENETICS INFORM
CONSERVATION BIOLOGY?

IN WHICH APPLICATION GENOMICS IS
A GAME CHANGER?

with a close up on two topics:

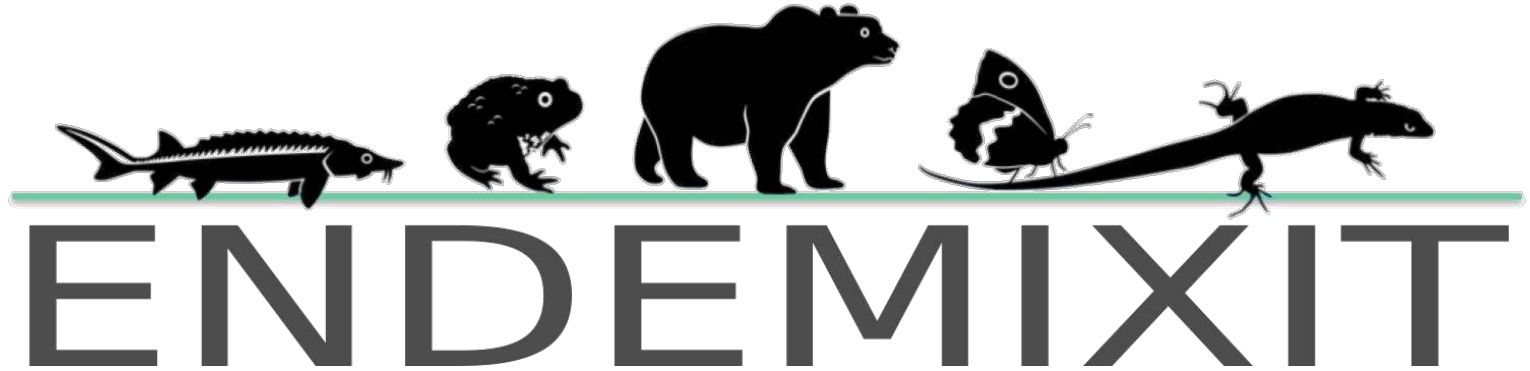
Genetic load

ROHs

WHAT AM I DOING TO MAKE USE OF
GENOMICS IN CONSERVATION?

GENOMIC SUSCEPTIBILITY TO EXTINCTION

A whole-genome approach to study and protect endangered Italian endemics

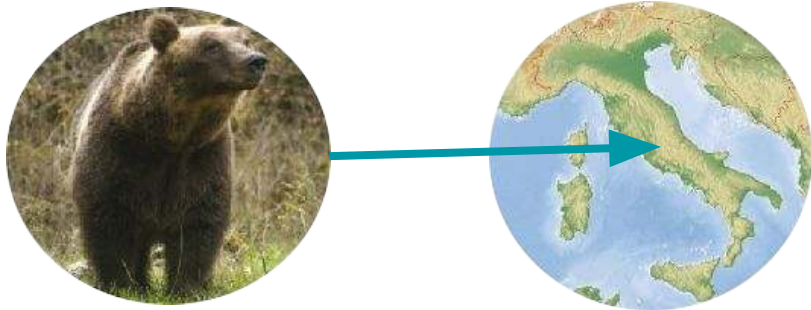


PI: **Giorgio Bertorelle**, University of Ferrara

<https://endemixit.com/>
https://youtu.be/mL_JzgOqk7c

GENOMICS OF FIVE ENDEMIC SPECIES

Very to very small population size



Marsican bear
(2.2Gb)

GENOMICS OF FIVE ENDEMIC SPECIES

Very to very small population size



Marsican bear
(2.2Gb)



Ponza grayling
(0.4Mb)

GENOMICS OF FIVE ENDEMIC SPECIES

Very to very small population size



Marsican bear
(2.2Gb)



Ponza grayling
(0.4Mb)



Aeolian lizard
(1.4Gb)

GENOMICS OF FIVE ENDEMIC SPECIES

Very to very small population size



Marsican bear
(2.2Gb)



Ponza grayling
(0.4Mb)



Aeolian lizard
(1.4Gb)



Apennine yellow-bellied toad (10Gb)

GENOMICS OF FIVE ENDEMIC SPECIES

Very to very small population size



Marsican bear
(2.2Gb)



Adriatic sturgeon
(1.4Gb, tetraploid)



Ponza grayling
(0.4Mb)



Aeolian lizard
(1.4Gb)



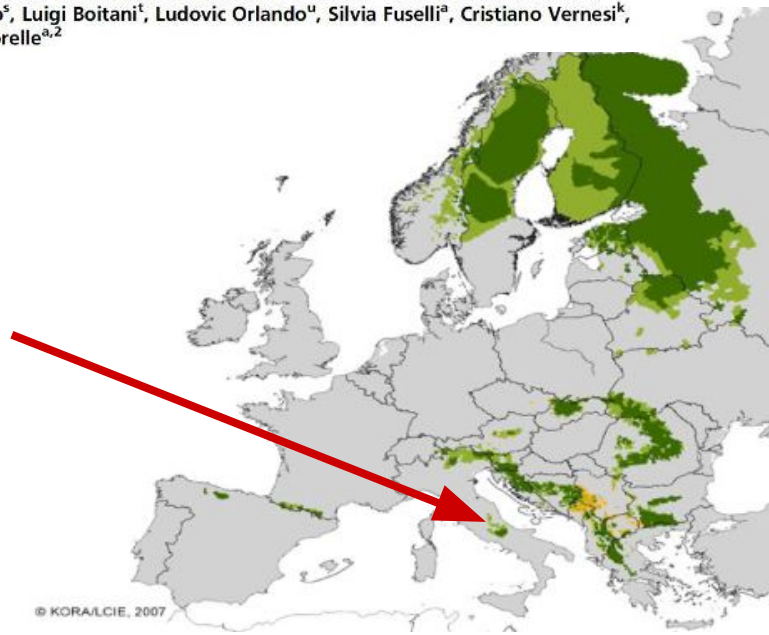
Apennine yellow-bellied toad (10Gb)

BEFORE ENDEMIXIT THERE WAS JUST THE APENNINE BEAR!

PNAS

Survival and divergence in a small group: The extraordinary genomic history of the endangered Apennine brown bear stragglers

Andrea Benazzo^{a,1}, Emiliano Trucchi^{a,b,1}, James A. Cahill^c, Pierpaolo Maisano Delser^{d,e,f}, Stefano Mona^{d,e}, Matteo Fumagalli^g, Lynsey Bunnefeld^{h,i}, Luca Cornetti^j, Silvia Ghirotto^a, Matteo Girardi^k, Lino Ometto^{l,m}, Alex Panziera^a, Omar Rota-Stabelli^l, Enrico Zanetti^a, Alexandros Karamanlidisⁿ, Claudio Groff^o, Ladislav Paule^p, Leonardo Gentile^q, Carles Vilà^r, Saverio Vicario^s, Luigi Boitani^t, Ludovic Orlando^u, Silvia Fuselli^a, Cristiano Vernesi^k, Beth Shapiro^c, Paolo Ciucci^t, and Giorgio Bertorelle^{a,2}



MOVE TO THE OLD PRESENTATION ON THE BEAR!



A brown bear is walking through a rocky, grassy landscape. The bear is positioned in the center-right of the frame, facing left. The terrain is covered with green grass and numerous grey, jagged rocks of various sizes. The background shows a continuation of this landscape, with more rocks and grass under a bright sky.

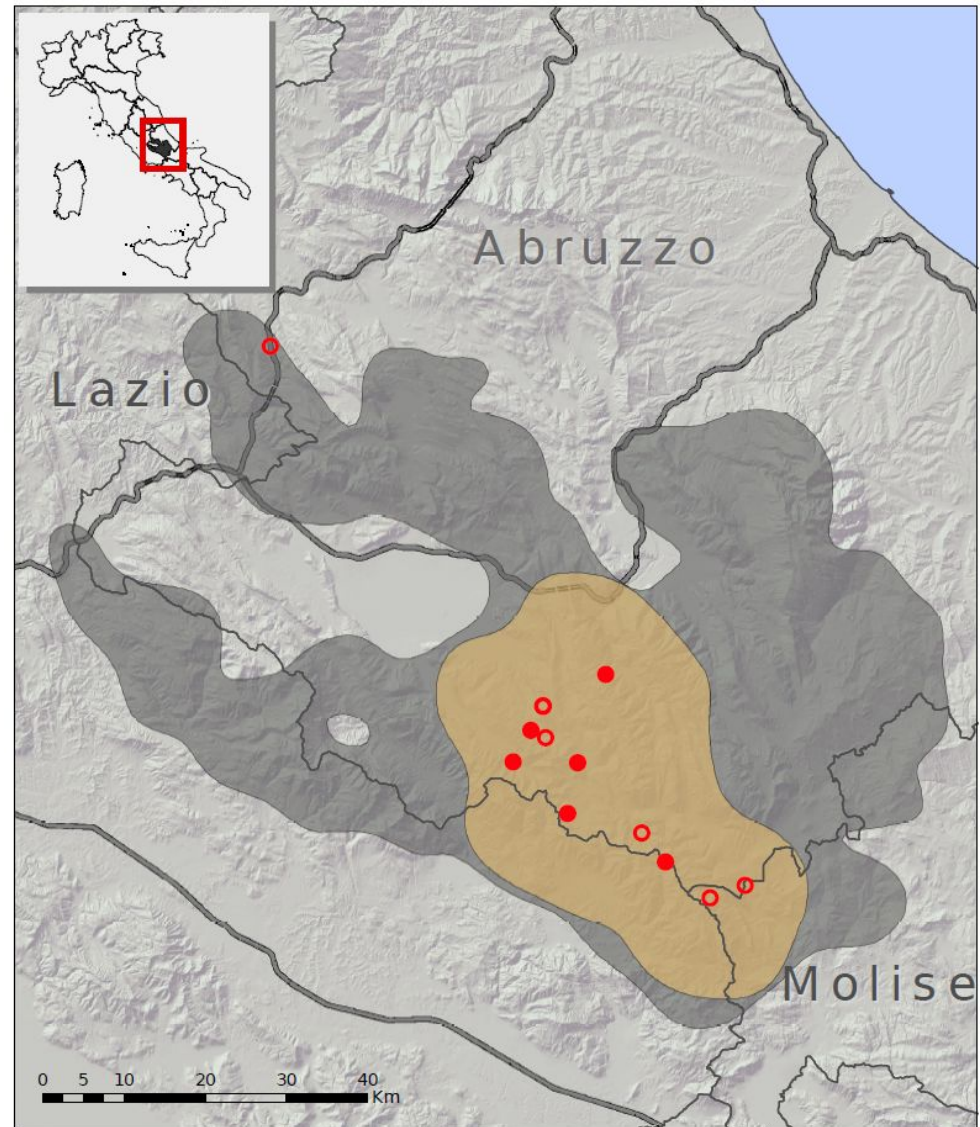
The extraordinary genomic history of the endangered Apennine Brown Bear

Brown bear distribution

A small and isolated population in Central Italy

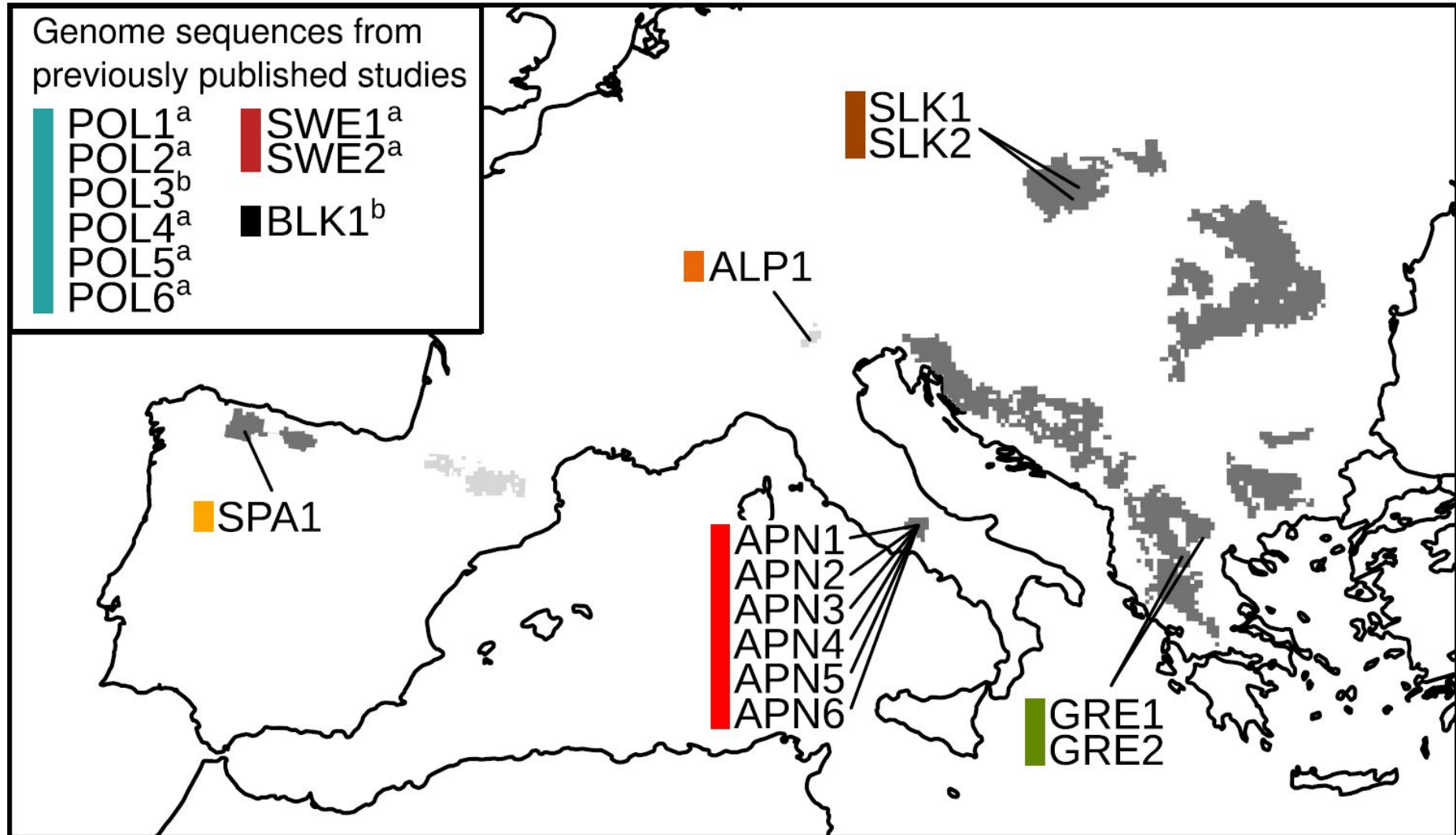


Less than **100** individuals



Whole-genome data

A few other samples from other European populations

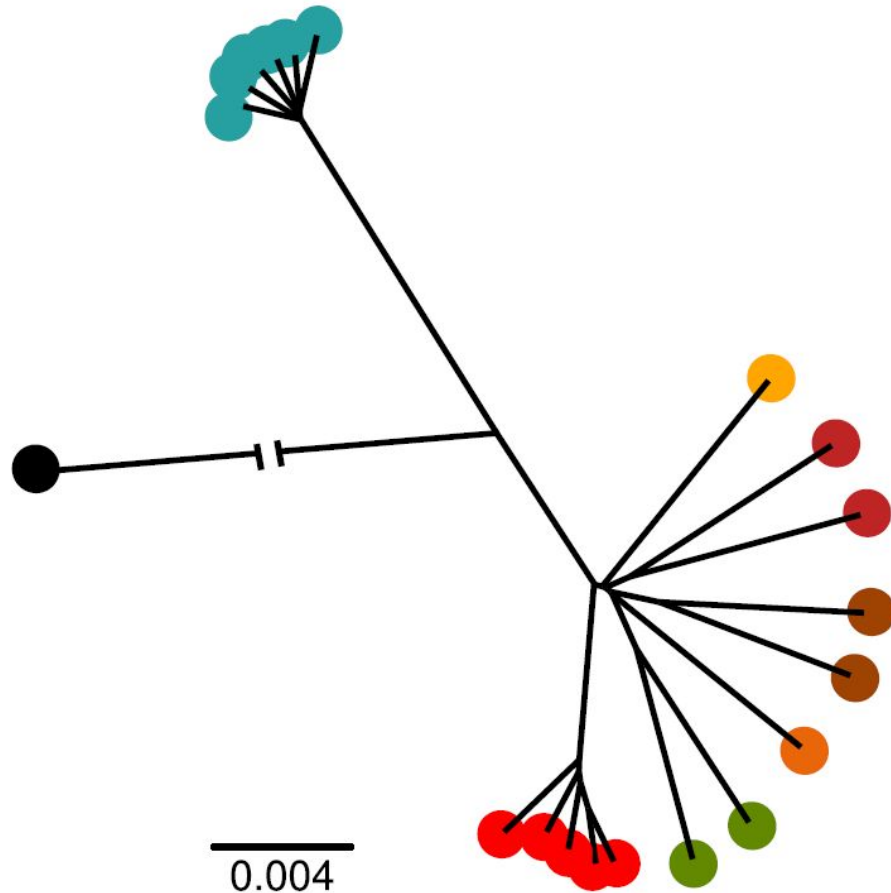


Structure of genomic diversity

Whole-genome and mitochondrion contrasting histories

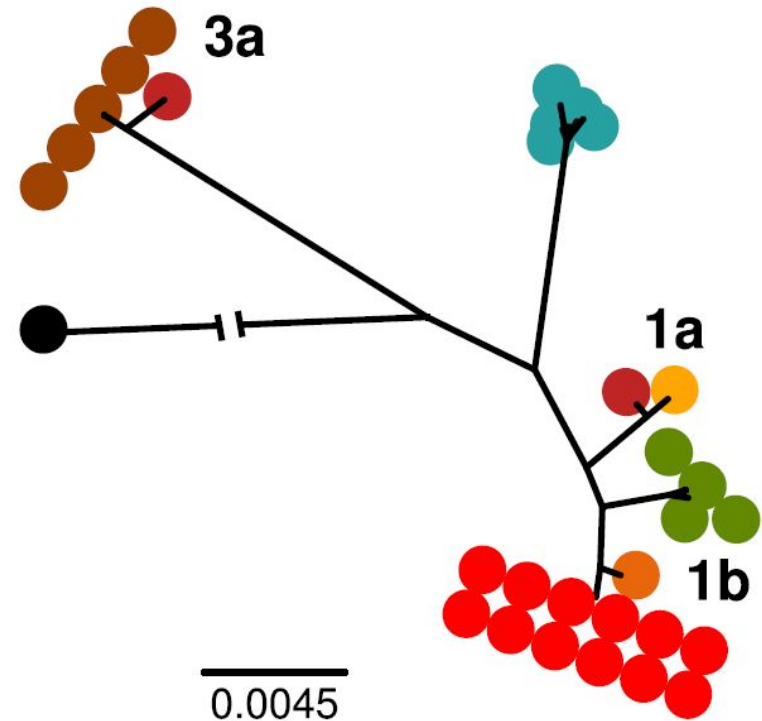


1,842,042,551 nuclear bp



ANGSD, ngsDist, nj from ape R package

16,485 mitochondrial bp



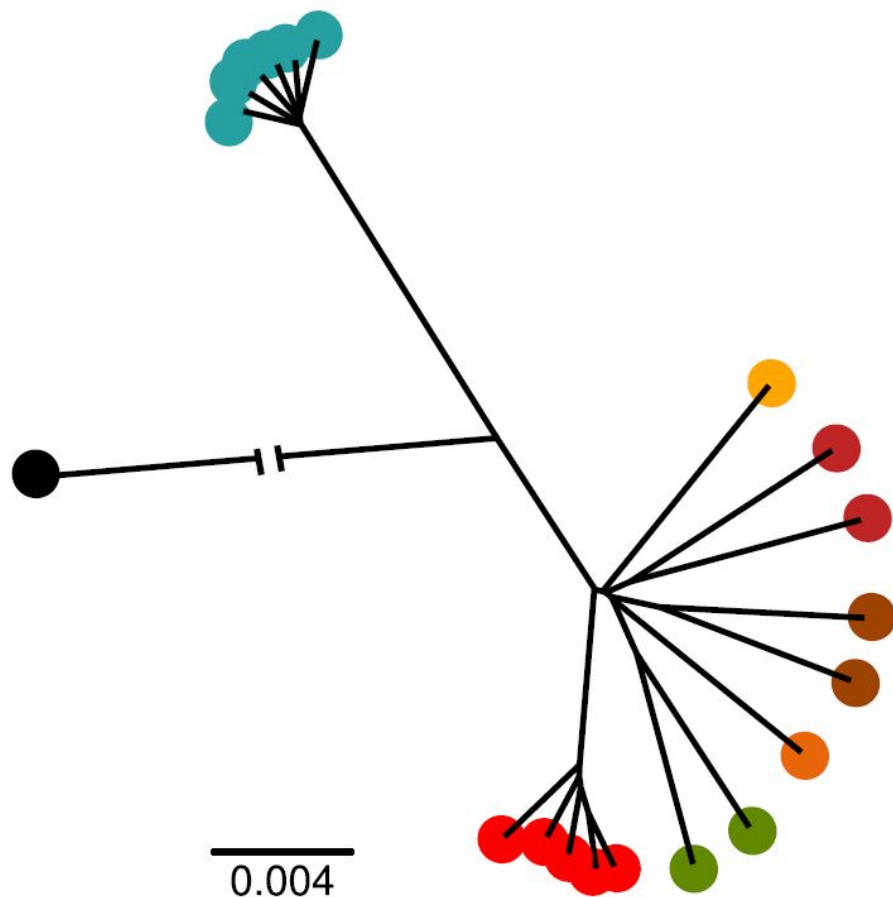
Geneious, MAFFT, genetic distance (HKY), NJ

Structure of genomic diversity

Y-chr agrees with the genome: sex-biased introgression?

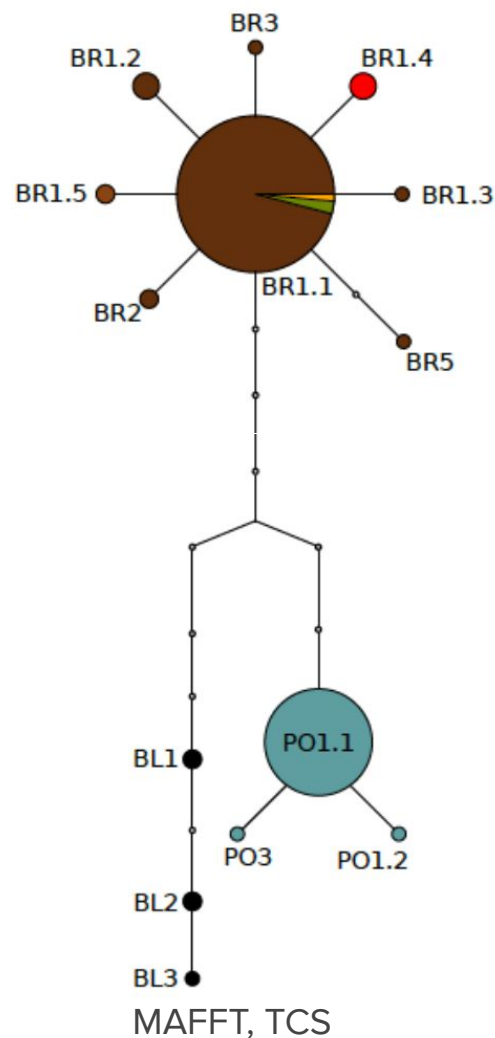


1,842,042,551 nuclear bp



ANGSD, ngsDist, nj from ape R package

5.3 Kb Y-chromosome



Genomic diversity - average

Low but the polar bear is worse

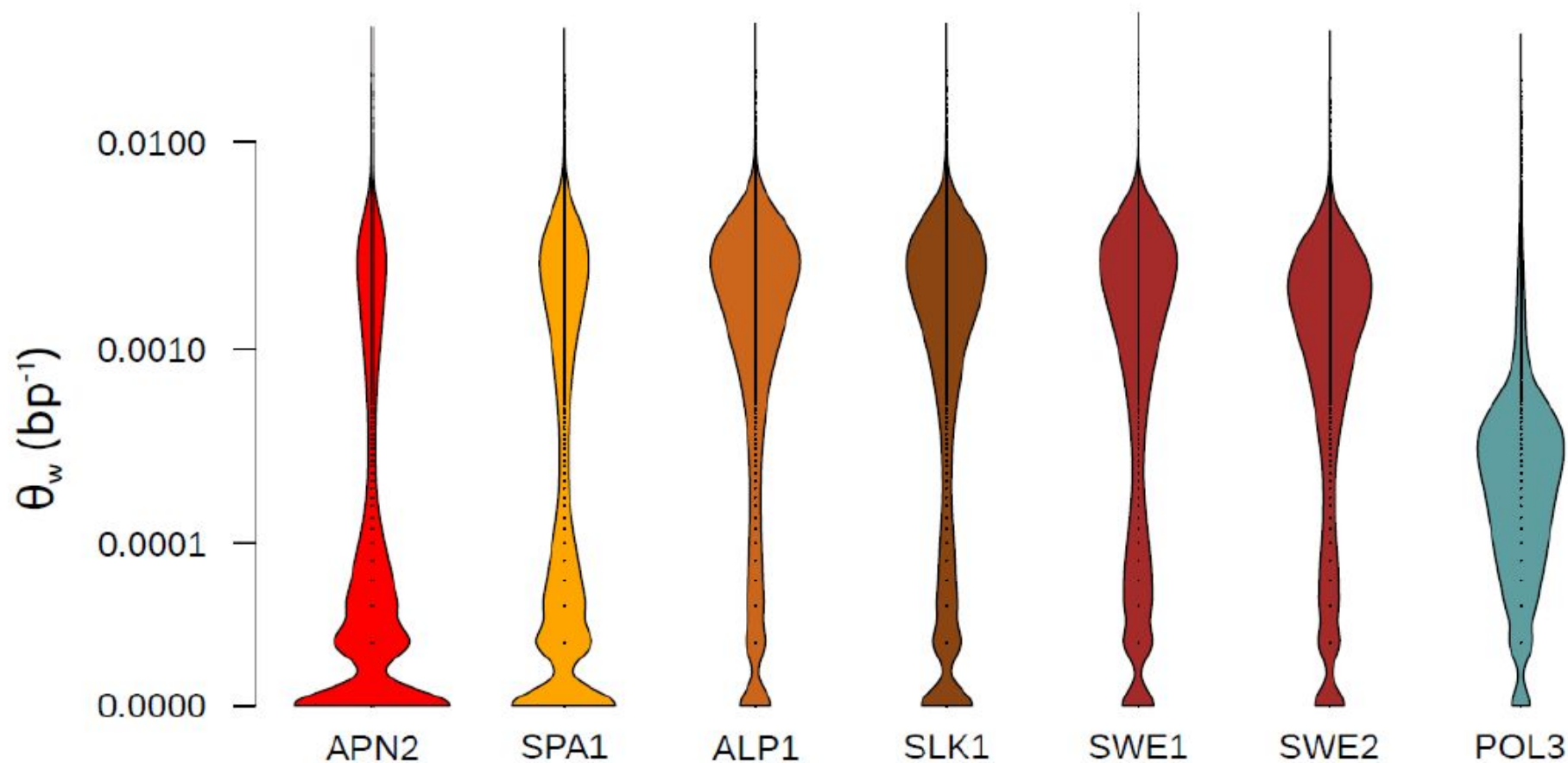


High coverage:
vcftools, SNP density, 50kb
windows
(then downsampled and analyzed
as low coverage)

Low coverage:
ANGSD, realSFS, do_theta,
50kb windows

Genomic diversity - distribution

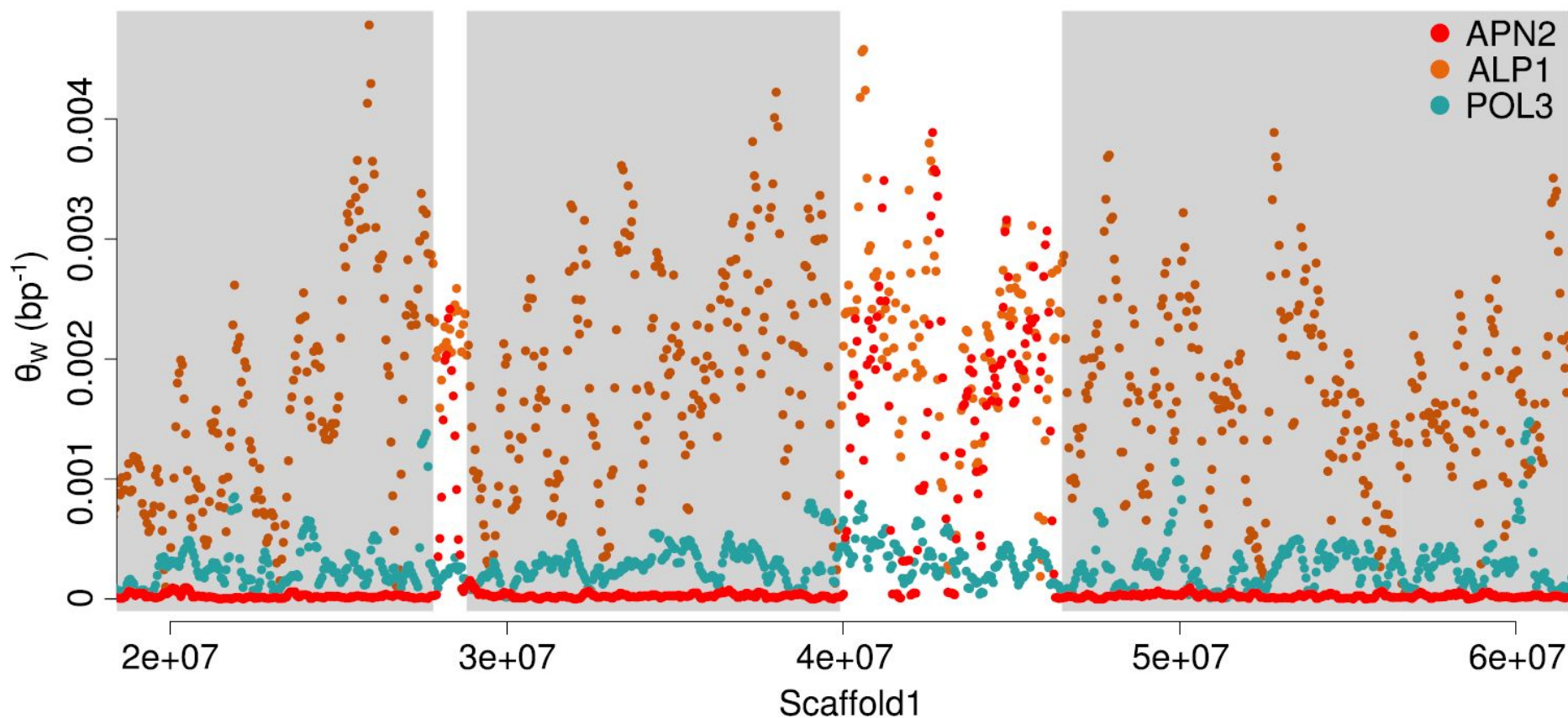
Low but the polar bear is worse...maybe not!



vcftools, SNP density in 50kb windows

Genomic diversity - distribution

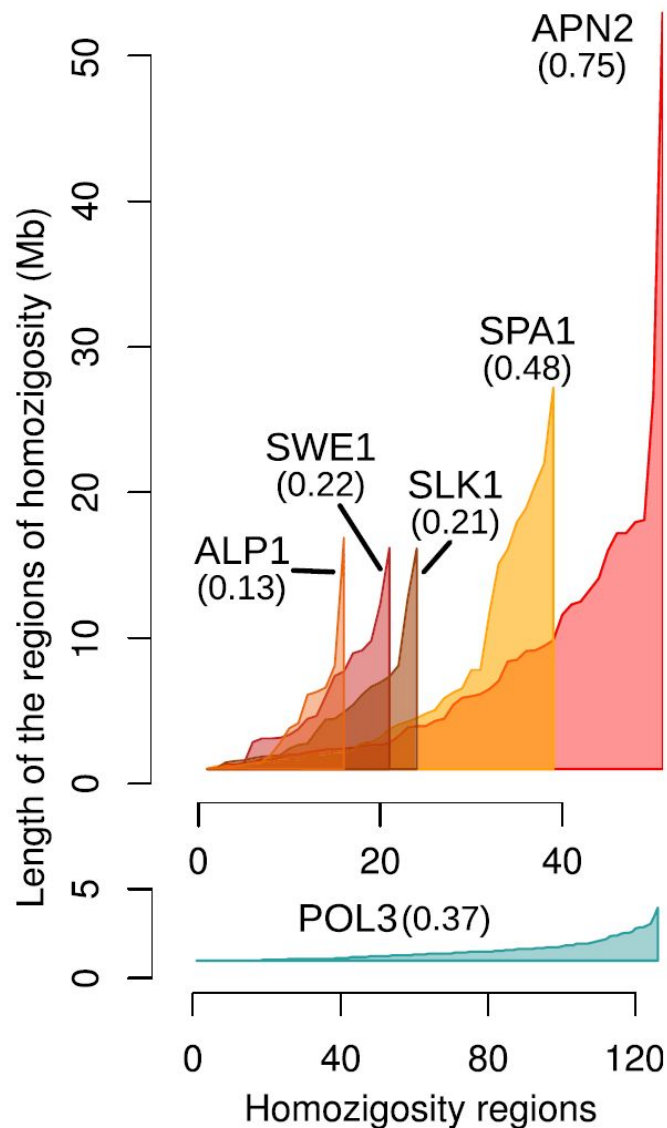
Long stretches with no diversity in the Apennine bear



vcftools, SNP density in 50kb windows

Inbreeding estimates

Much higher than any other European bear

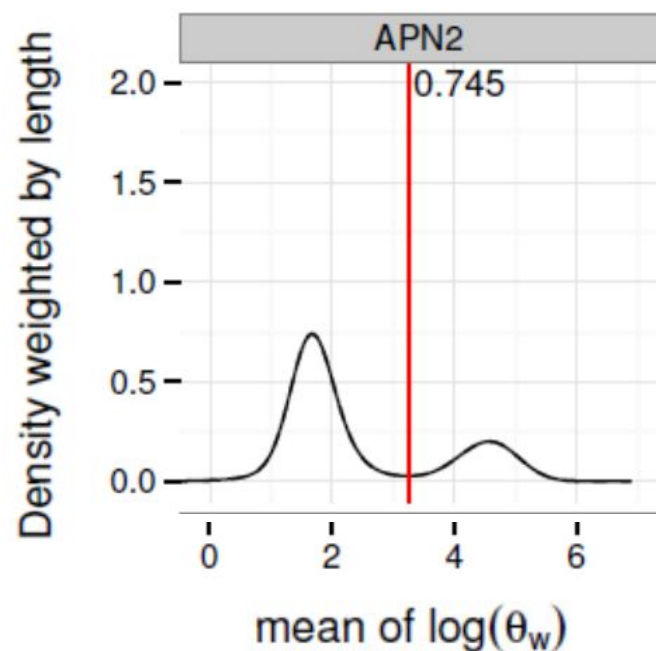
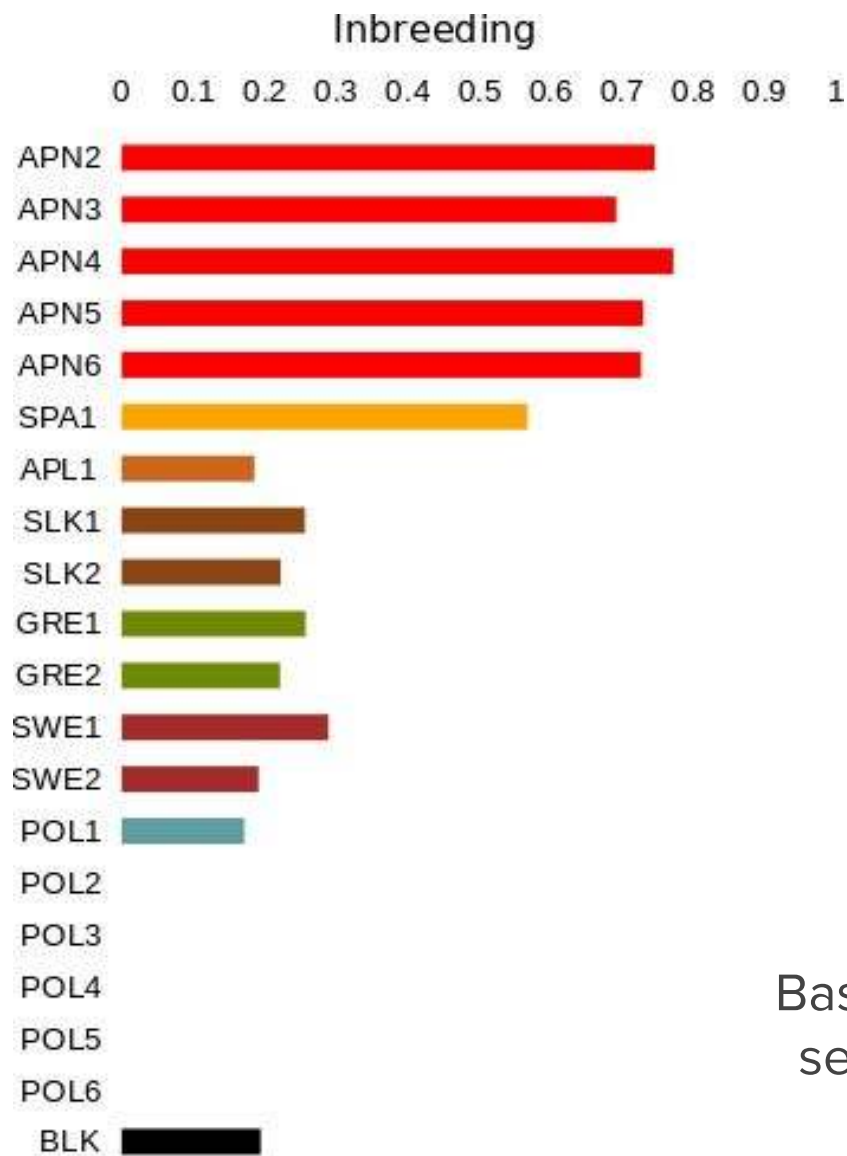


ROH: Regions longer than 1Mb with less than 25 segregating sites in the 13 longest scaffolds, ca. 0.5 Gb - sorted

(percent of the analyzed scaffolds)

Inbreeding estimates

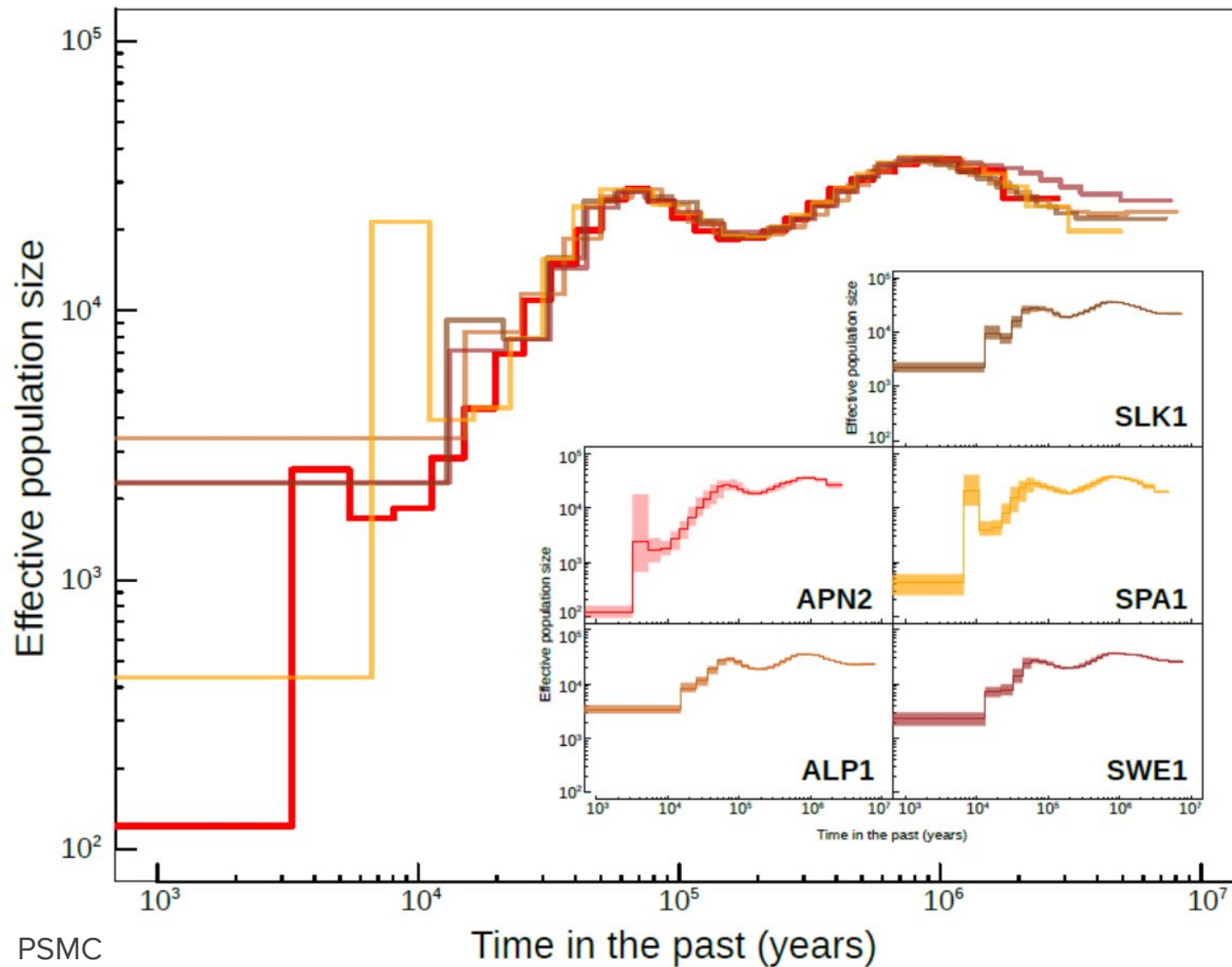
Much higher than any other European bear



Based on the proportion of the genome segments that are mostly homozygous (Prüfer et al 2014)

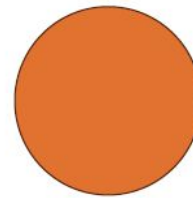
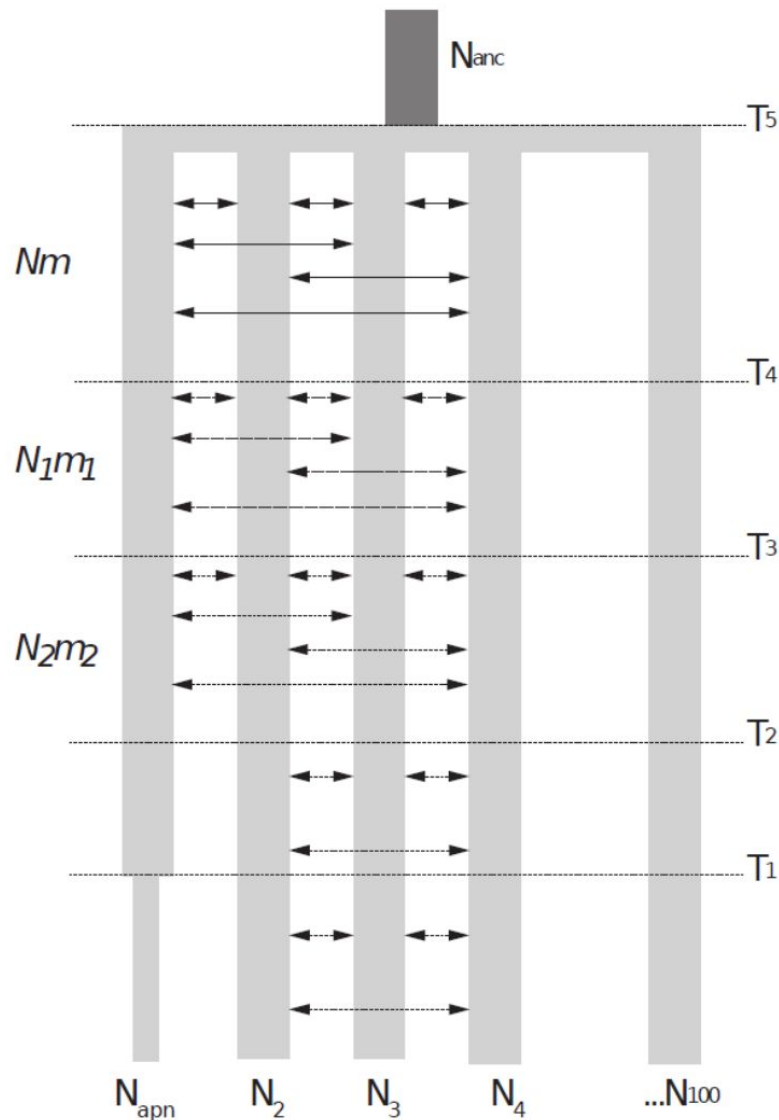
Past demography

Apennine population declined more than other Eu pops



Past demography

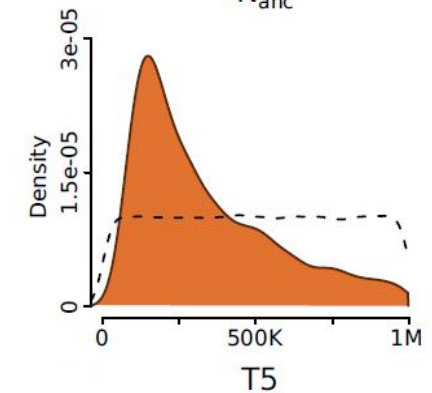
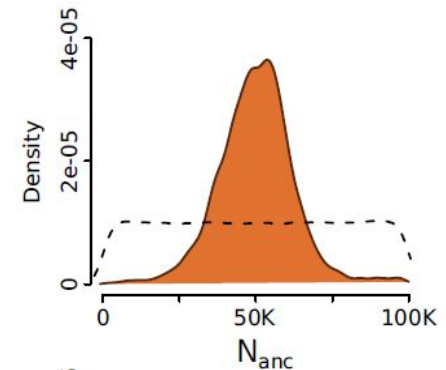
Fragmentation of a large European population



Ancestral population



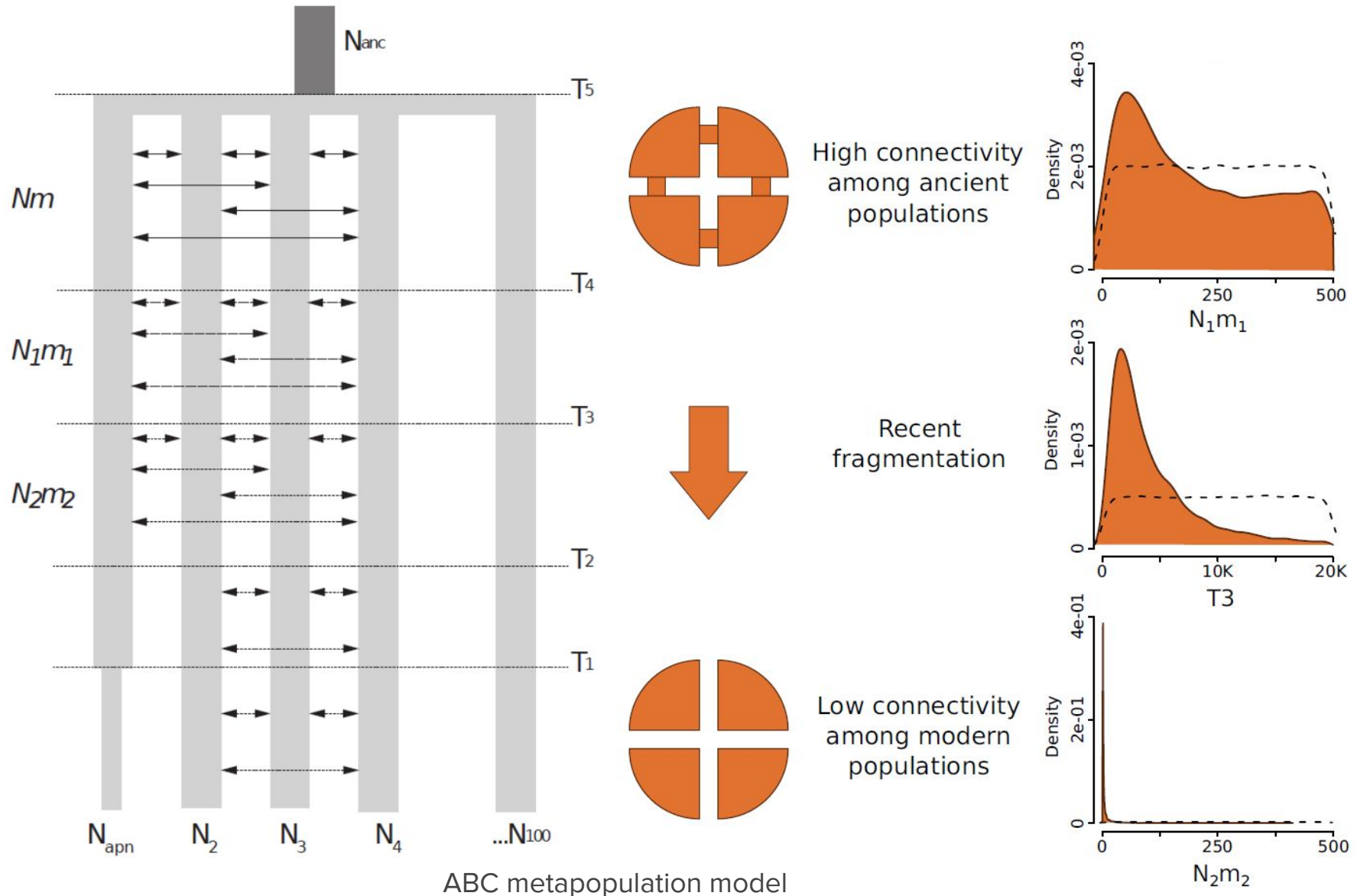
Ancient European spread



ABC metapopulation model

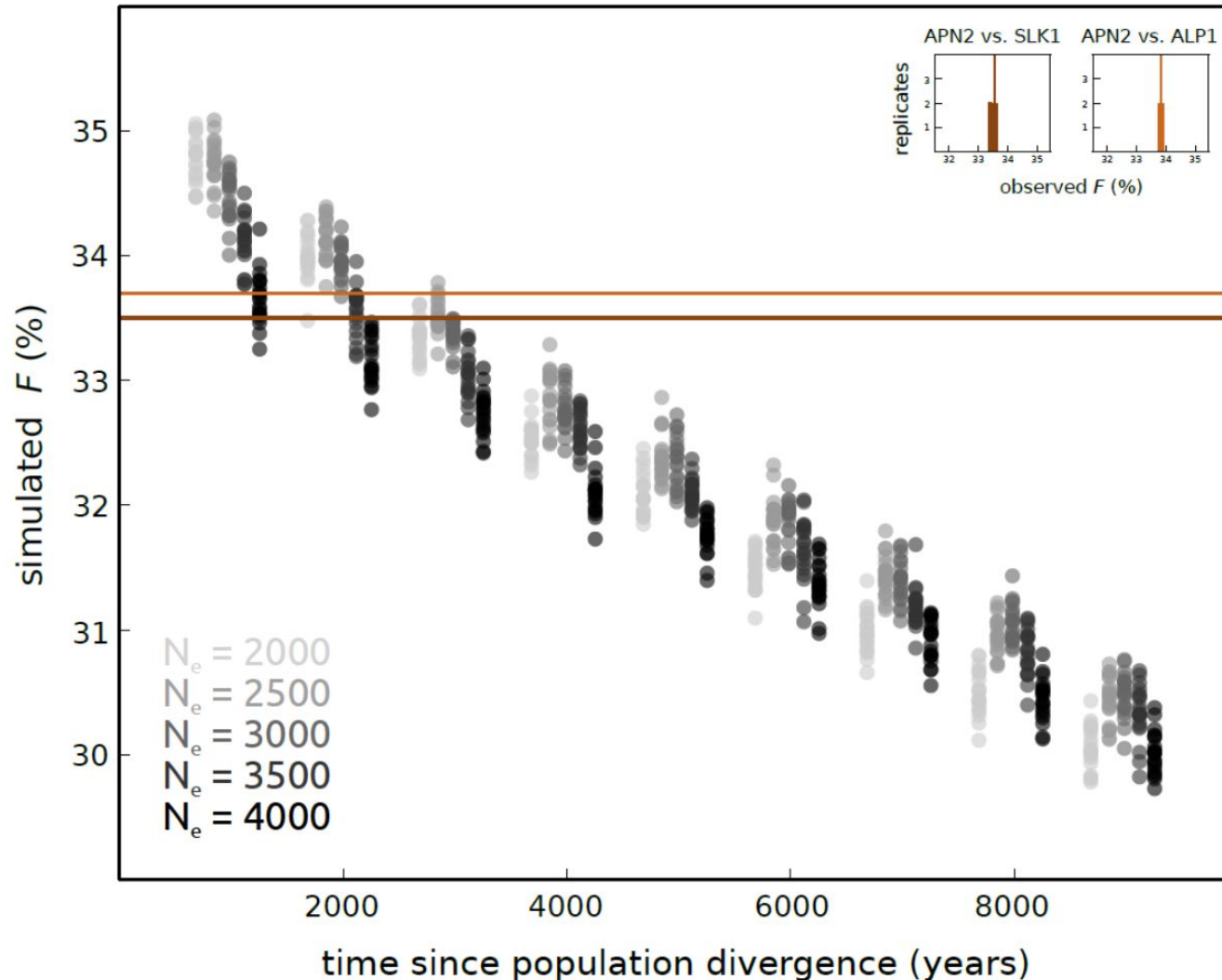
Past demography

Fragmentation of a large European population



Past demography

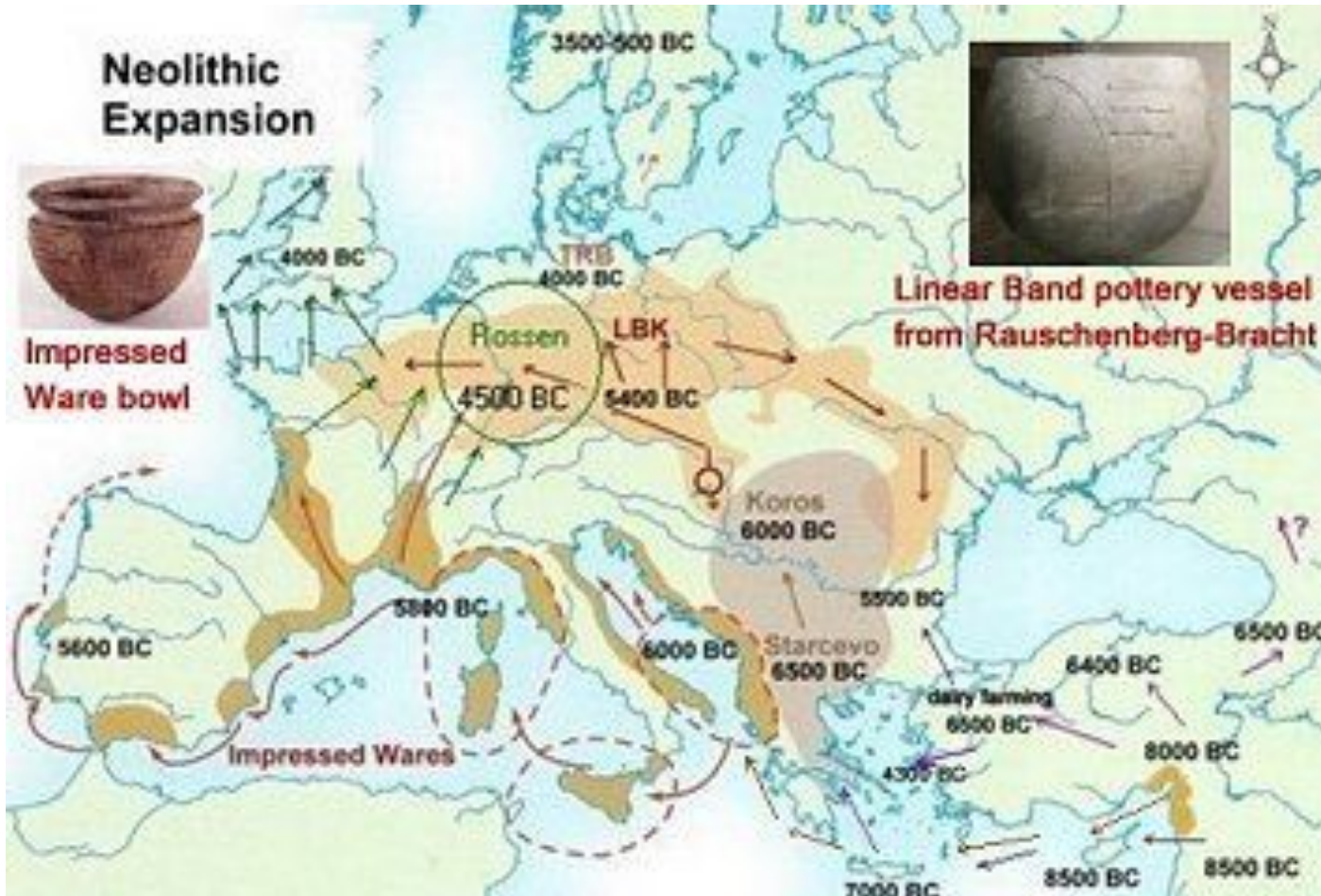
Apennine bears have been isolated since then



F statistic (Green et al 2010), simulation with ms

Past demography

Expansion of Neolithic farming in Europe burning forests



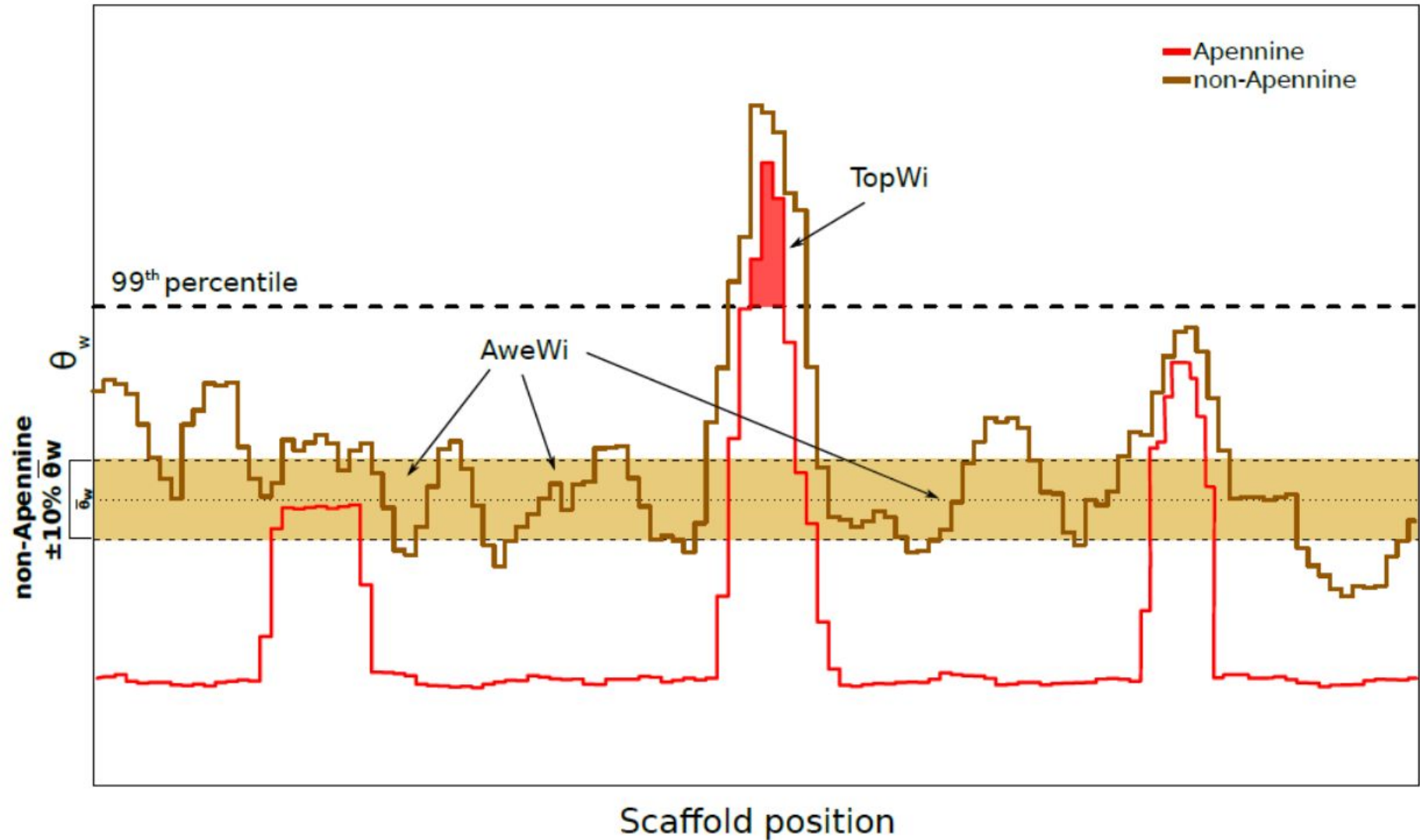
Why is this population still there?

Given its likely high extinction probability



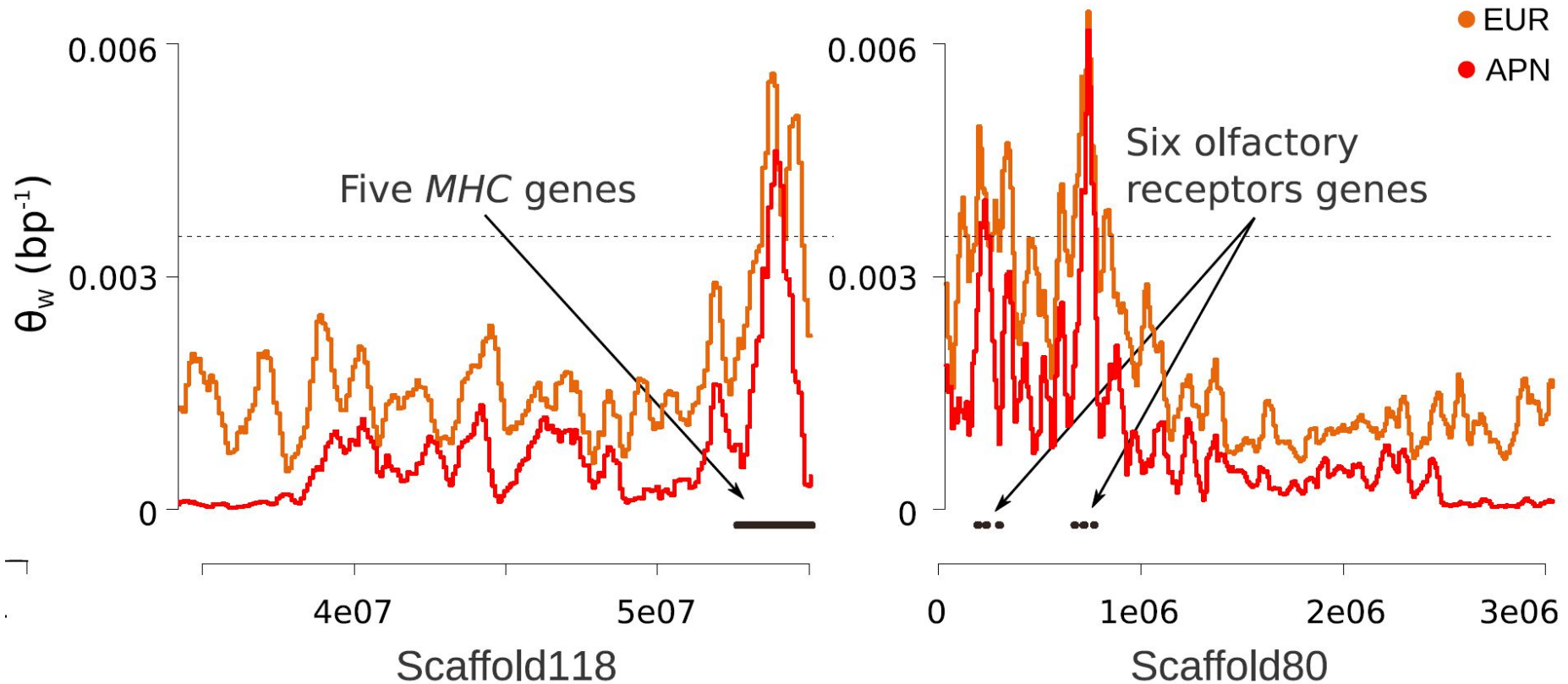
Selective processes: balancing

Retention of high diversity regions (HDR)



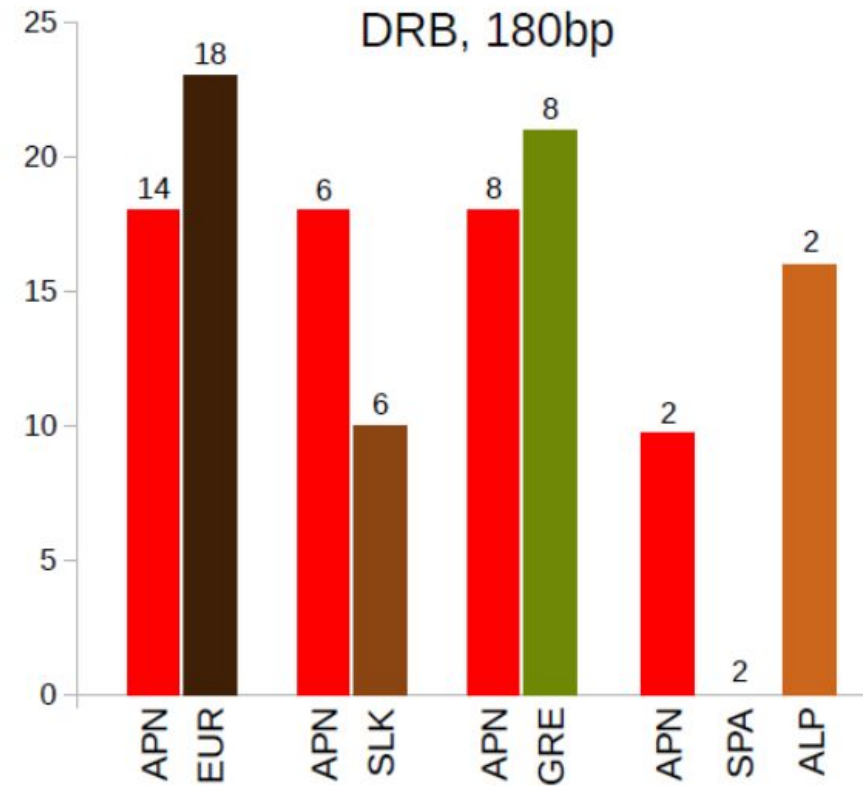
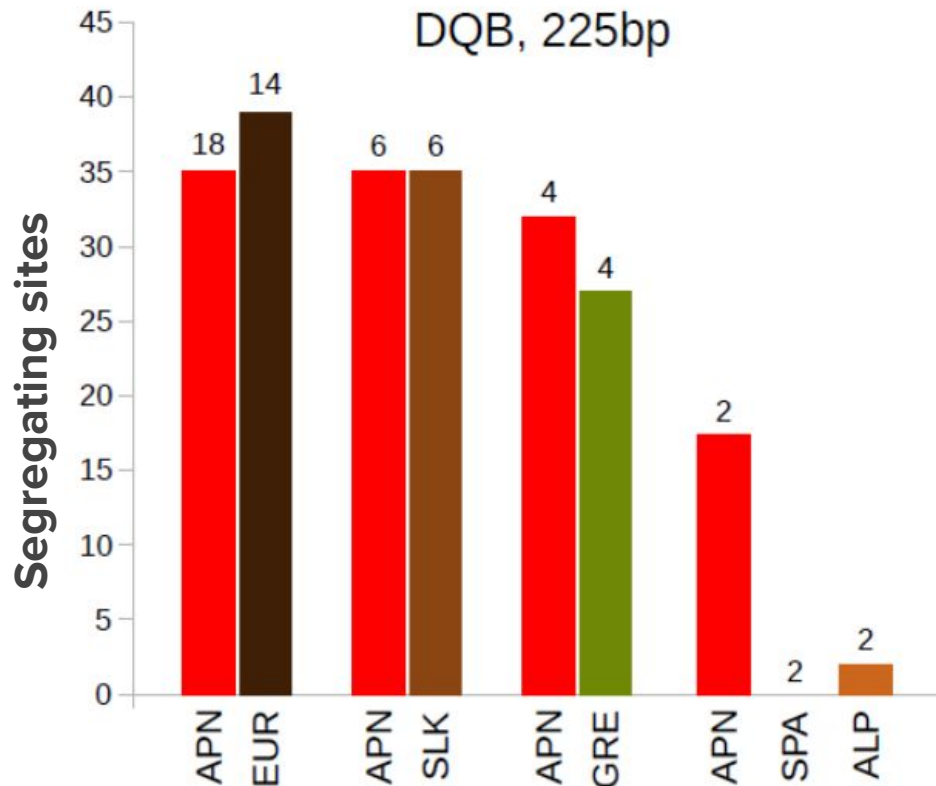
Selective processes: balancing

HDR are enriched in immune and olfactory genes



Selective processes: balancing

MHC genes are as diverse as in the rest of Europe



Sanger-sequencing of two *MHC* class II loci in additional individuals from Apennine and the rest of Europe

Genetic load

Accumulation of deleterious effects



Predicted deleterious substitution

fixed in the Apennine bear genomes

(by Panther + Polyphen):

40 + 4 stop codon

None in the other European bears!

Genetic load

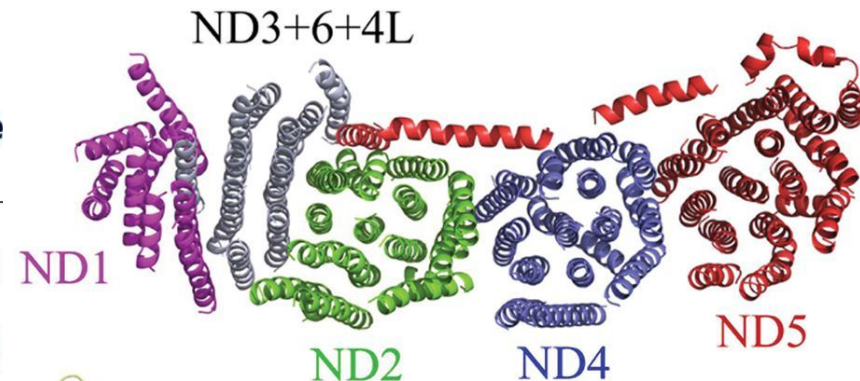
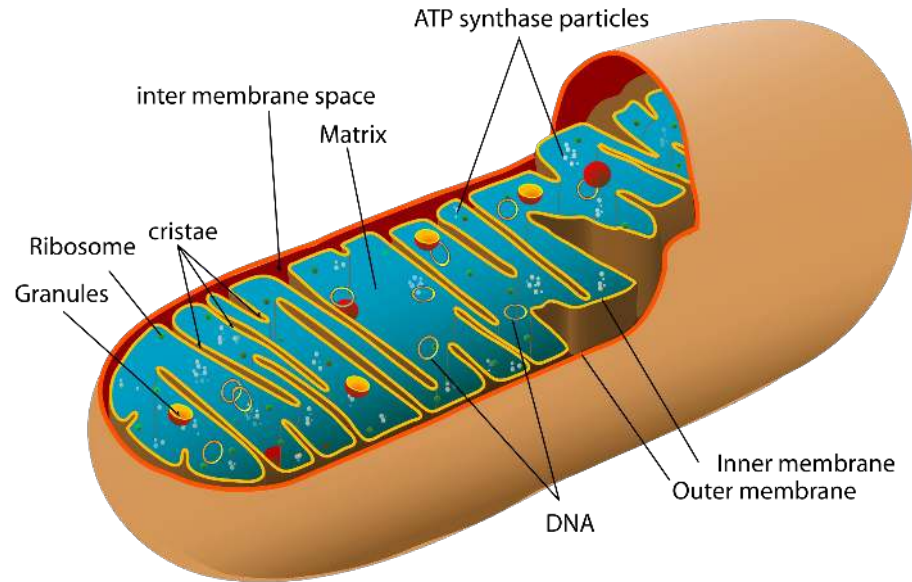
Accumulation of deleterious effects



Predicted deleterious substitution fixed in the Apennine bear genomes (by Panther + Polyphen):
40 + 4 stop codon

None in the other European bears!

Five subs in the **mt ND5** of which the most deleterious one is not found elsewhere



Gene	Position	#APNs (out of 6)	#NonAPNs (out of 45)	Score
ND5	G526E	6	0	-4.28
ND5	P447S	6	5	-3.59
ND5	T555A	6	3	-2.52

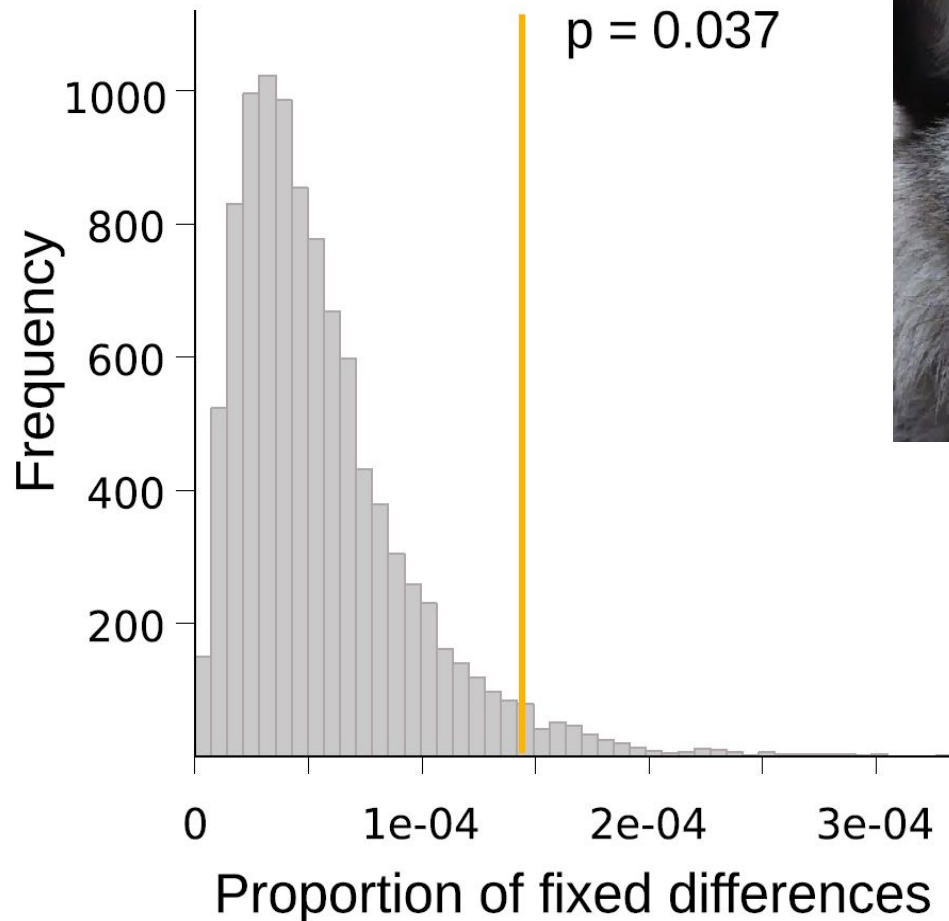
Genetic load in specific group of genes

Are Apennine bears less aggressive?



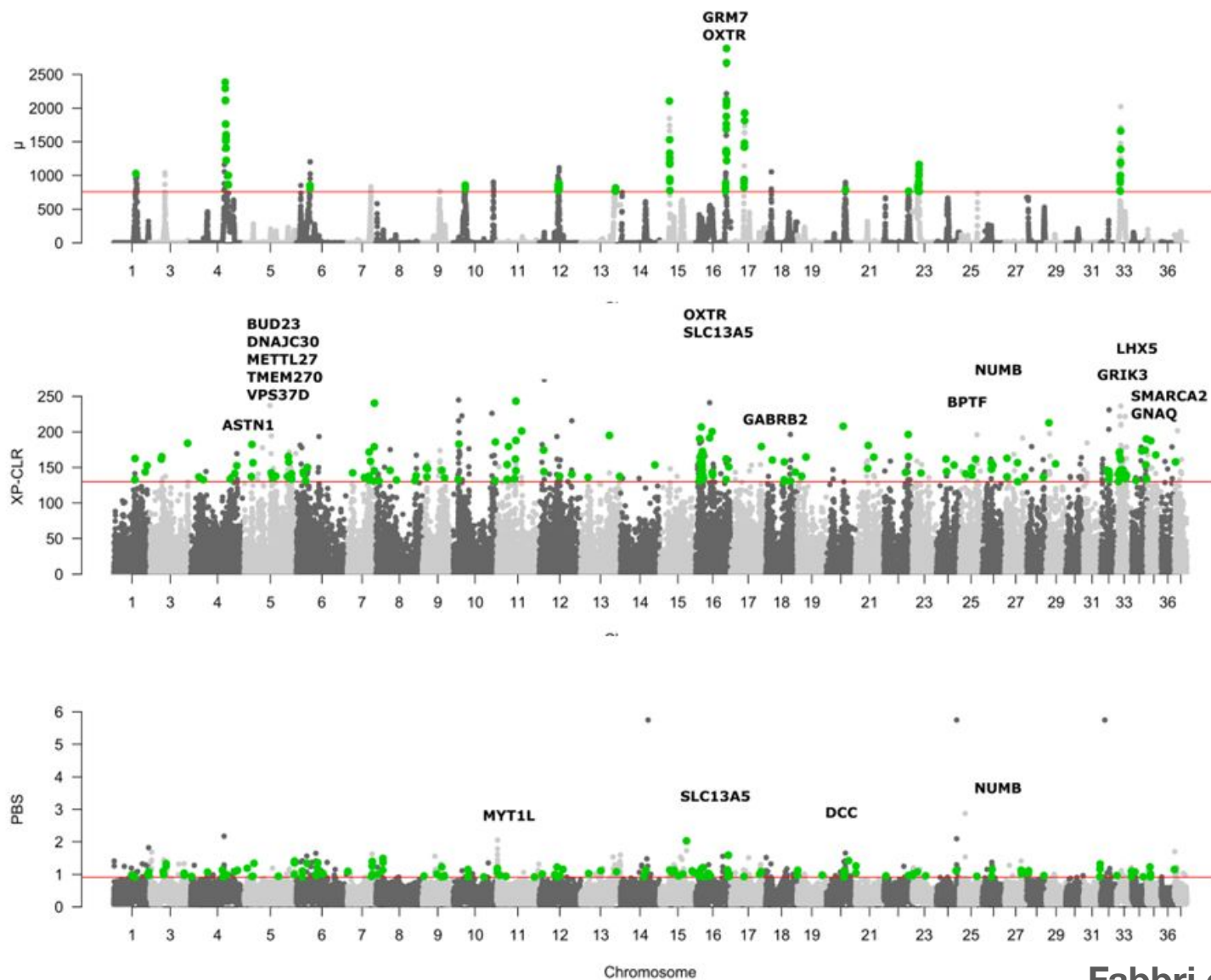
Fixed differences in 22 “tameness” genes

Not all deleterious fixations come to harm



Pattern confirmed with **additional individuals** in three of these genes:
PLAXNB1, DCC, DLL3

FURTHER EVIDENCE OF DIVERGENCE AT BEHAVIORAL GENES WITH MORE RECENT DATA AND ANALYSES



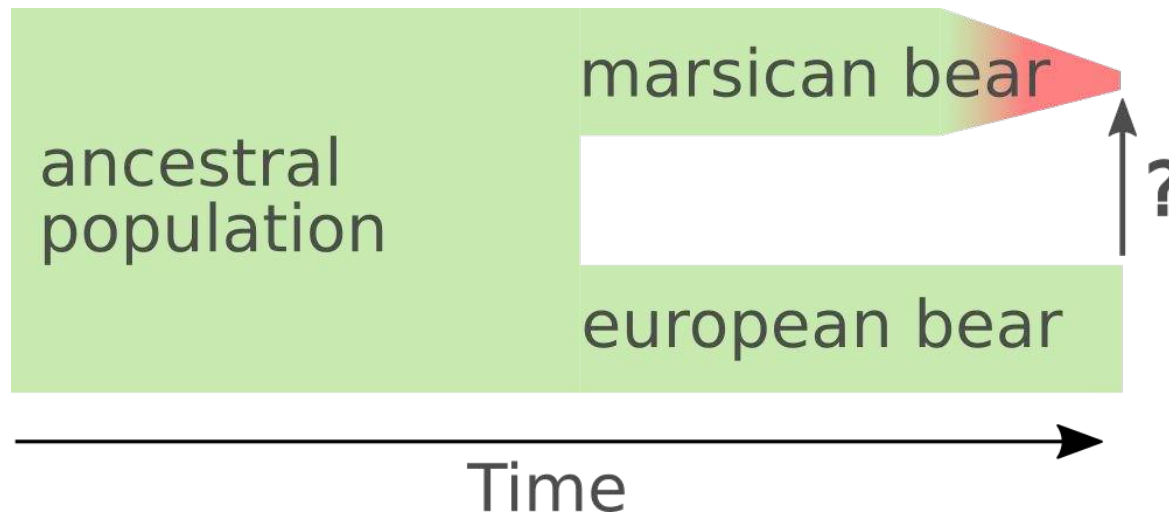
Future conservation strategies

Trade-off between increasing fitness without losing ancestry



Genetic rescue vs. doubling population size

Realistic forward simulation using Slim



How many individuals? How to choose them? -> 5 with the lowest **genetic load** as proportion of deleterious alleles

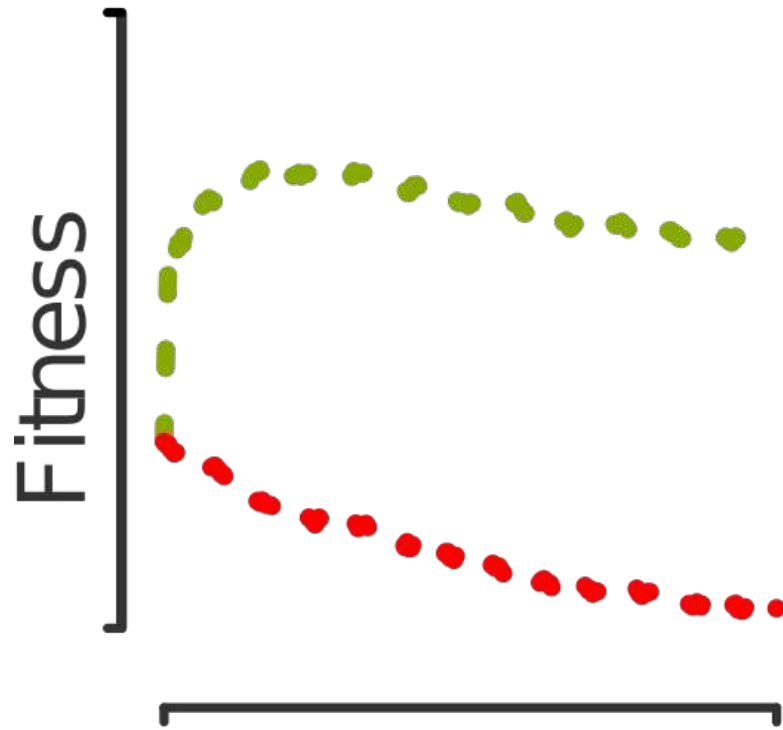
Future conservation strategies

Trade-off between increasing fitness without losing ancestry

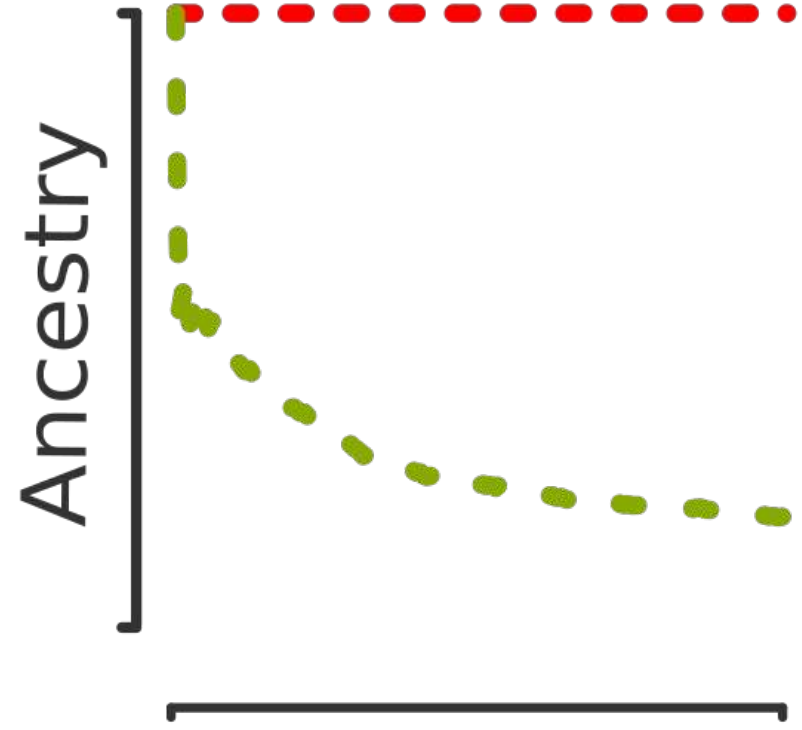


GR

1 x PopSize



NO GR



Years after GR and/or 2 x PopSize

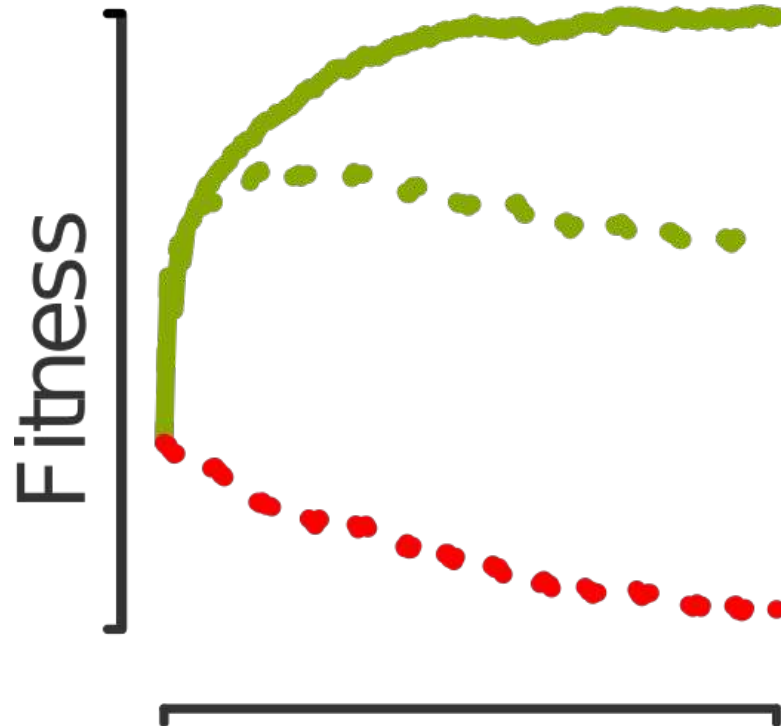
Future conservation strategies

Trade-off between increasing fitness without losing ancestry



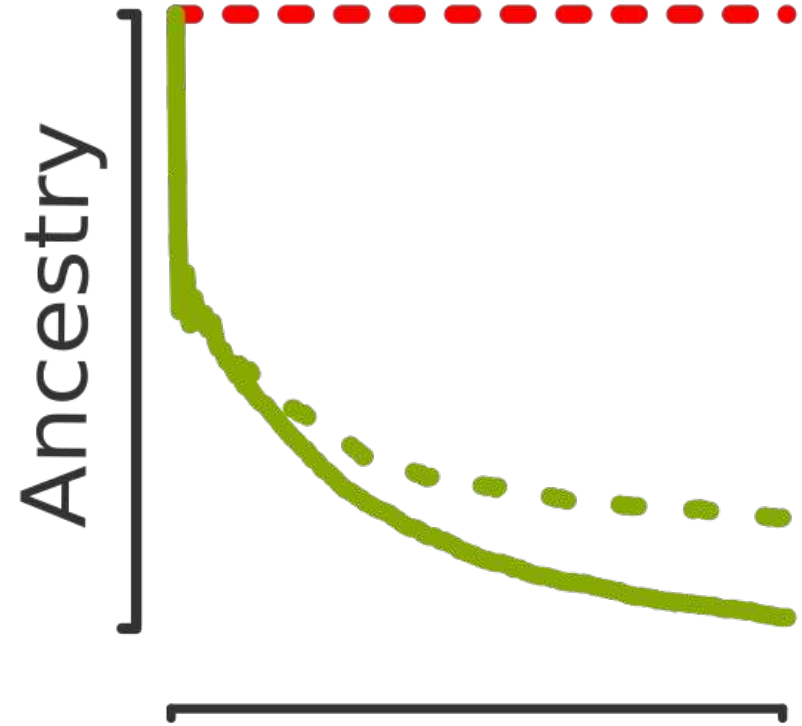
GR

1 x PopSize



NO GR

2 x PopSize



Years after GR and/or 2 x PopSize

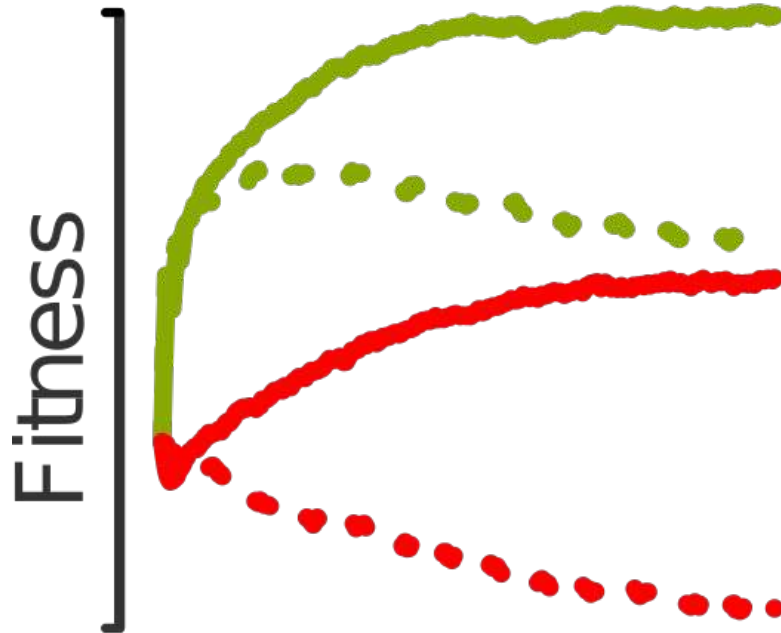
Future conservation strategies

Trade-off between increasing fitness without losing ancestry



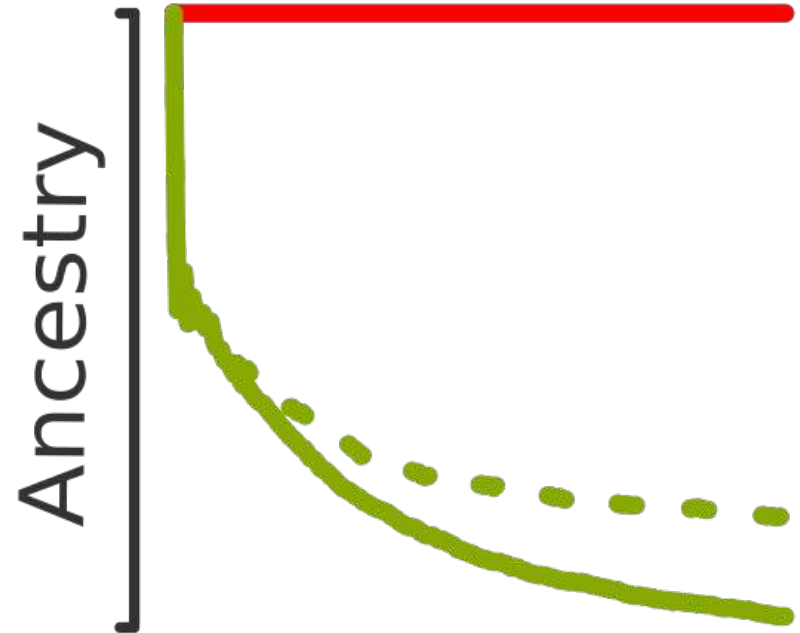
GR

1 x PopSize



NO GR

2 x PopSize



Years after GR and/or 2 x PopSize

GENOMICS OF FIVE ENDEMIC SPECIES

Very to very small population size



Marsican bear
(2.2Gb)



Adriatic sturgeon
(1.4Gb, tetraploid)



Ponza grayling
(0.4Mb)



Aeolian lizard
(1.4Gb)



Apennine yellow-bellied toad (10Gb)

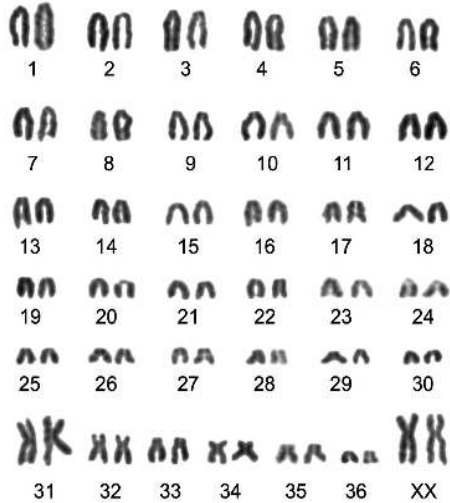
SAMPLING FOR GENOME ASSEMBLY AND RESEQUENCING

Small and large populations to compare

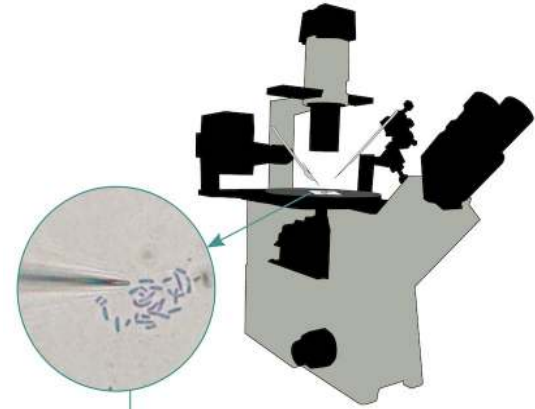


KARYOTYPE AND CHROMOSOME SORTING

The toad has few but huge chromosomes



MICRODISSECTION



Inverted microscope

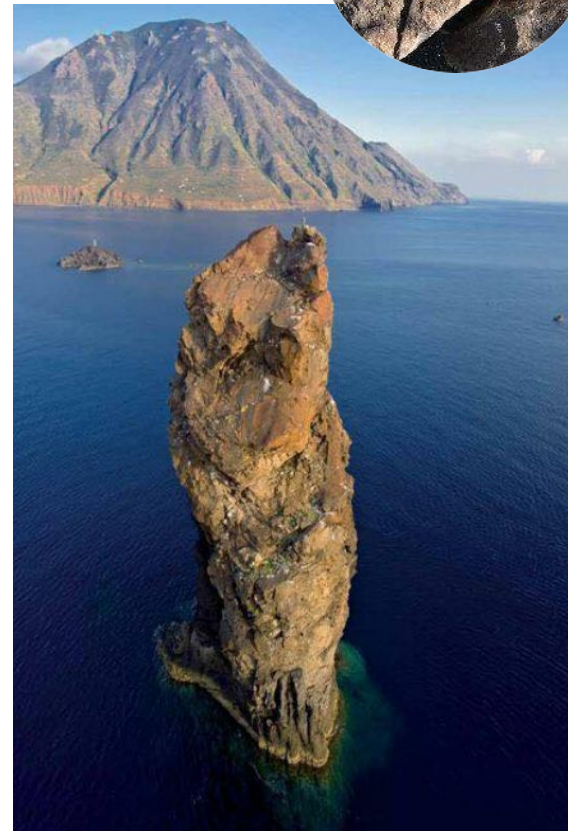
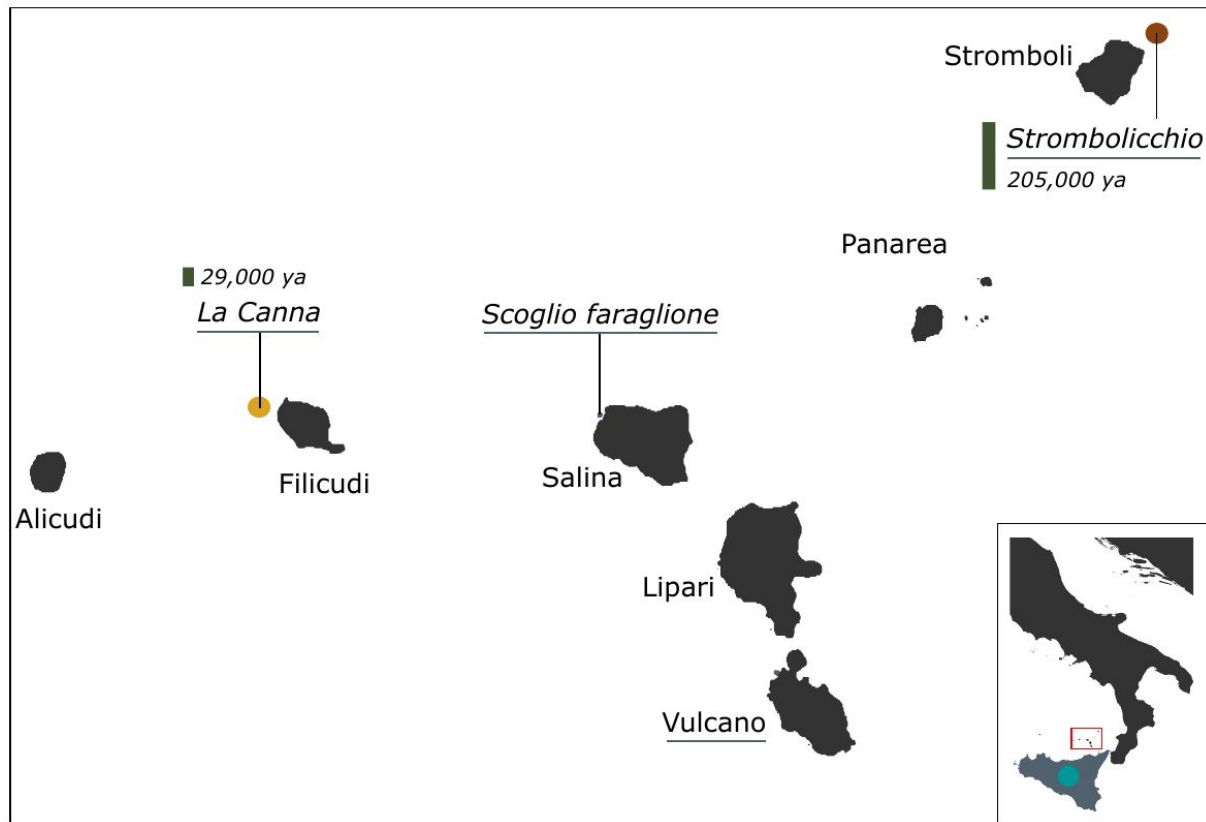


Chromosome 2

SEQUENCING

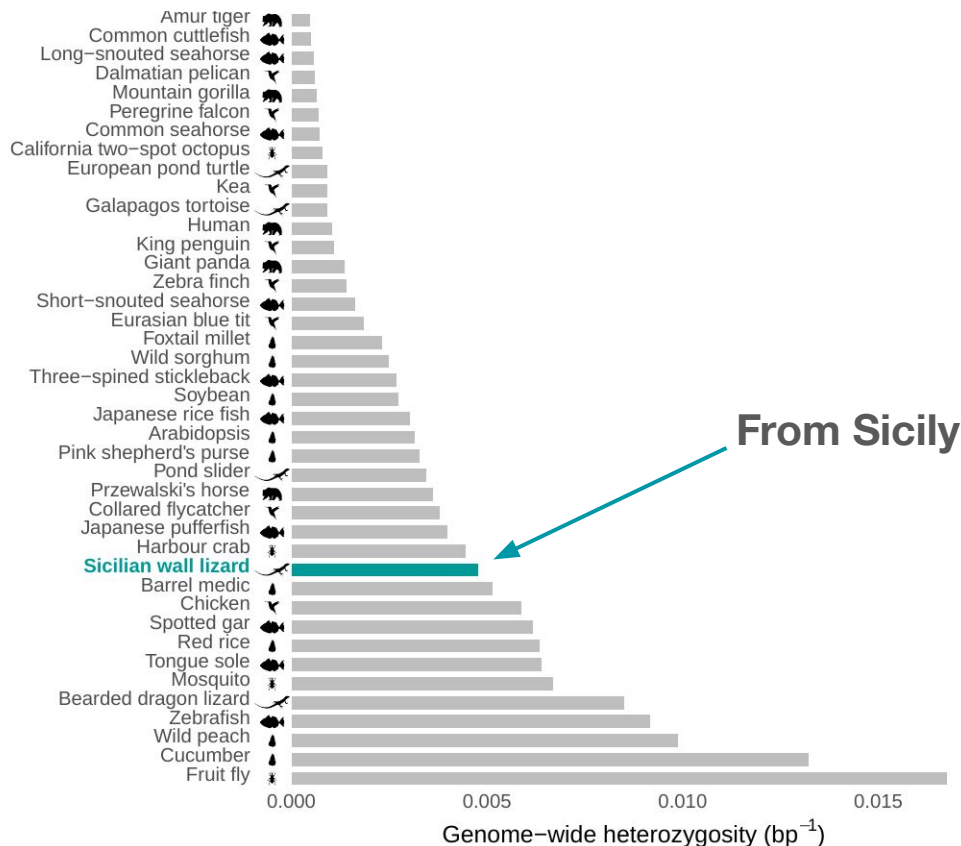
LIVING ON A ROCK AS A WILD INBRED STRAIN

Setting the limit for genomic variation and genetic load



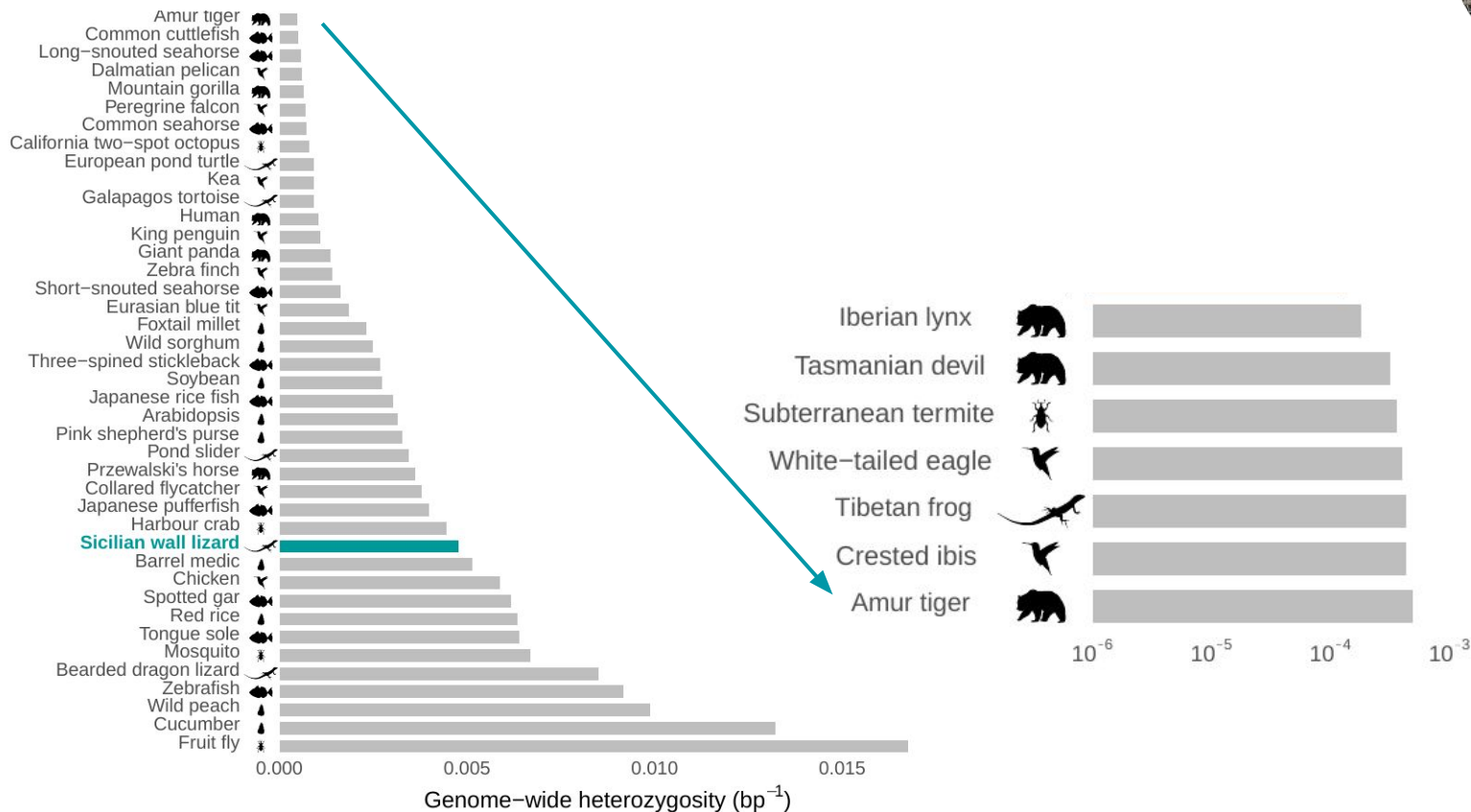
LIVING ON A ROCK AS A WILD INBRED STRAIN

Setting the limit for genomic variation and genetic load



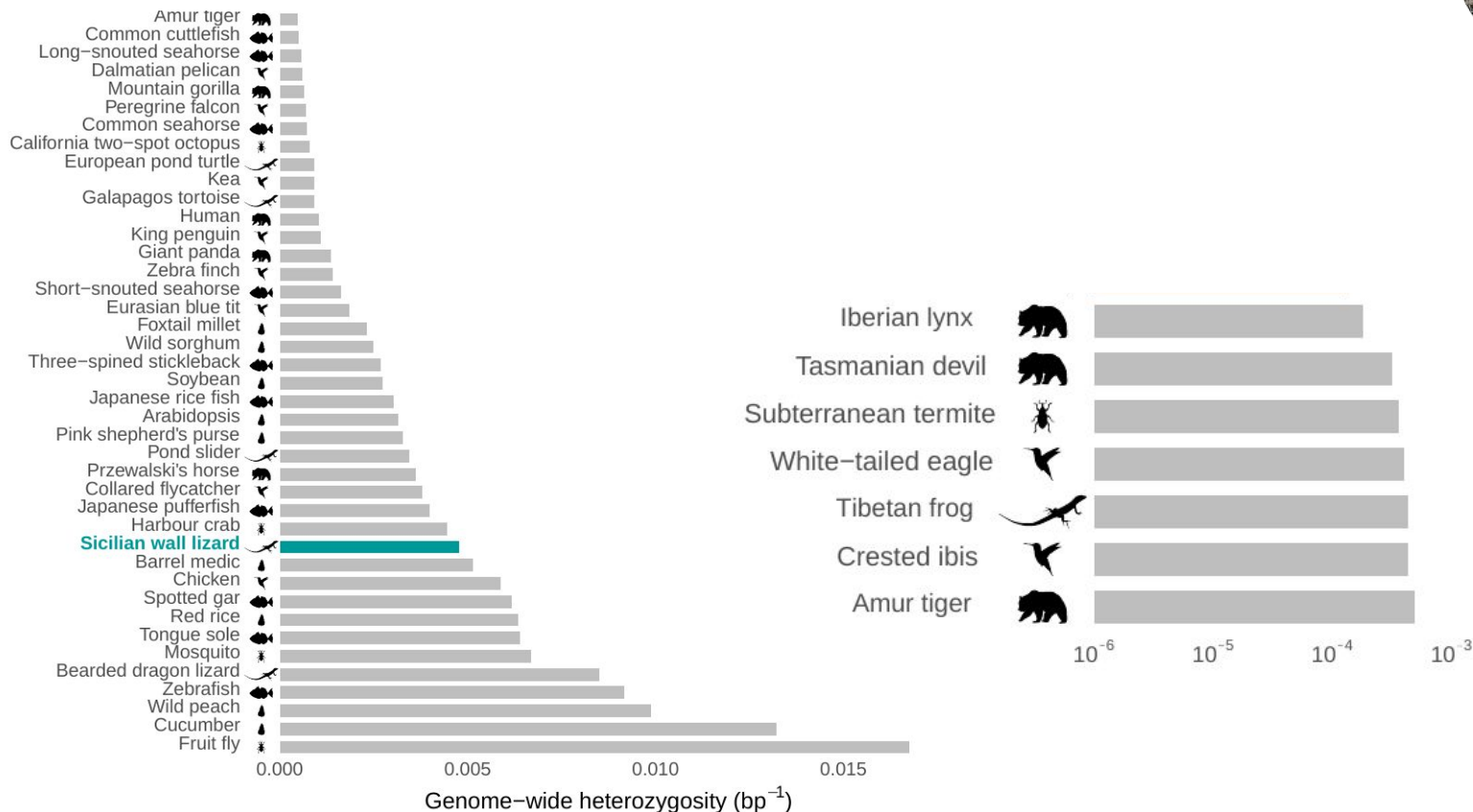
LIVING ON A ROCK AS A WILD INBRED STRAIN

Setting the limit for genomic variation and genetic load



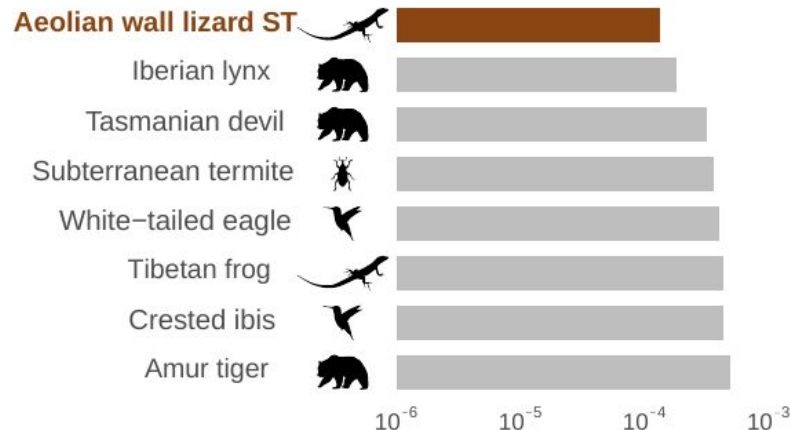
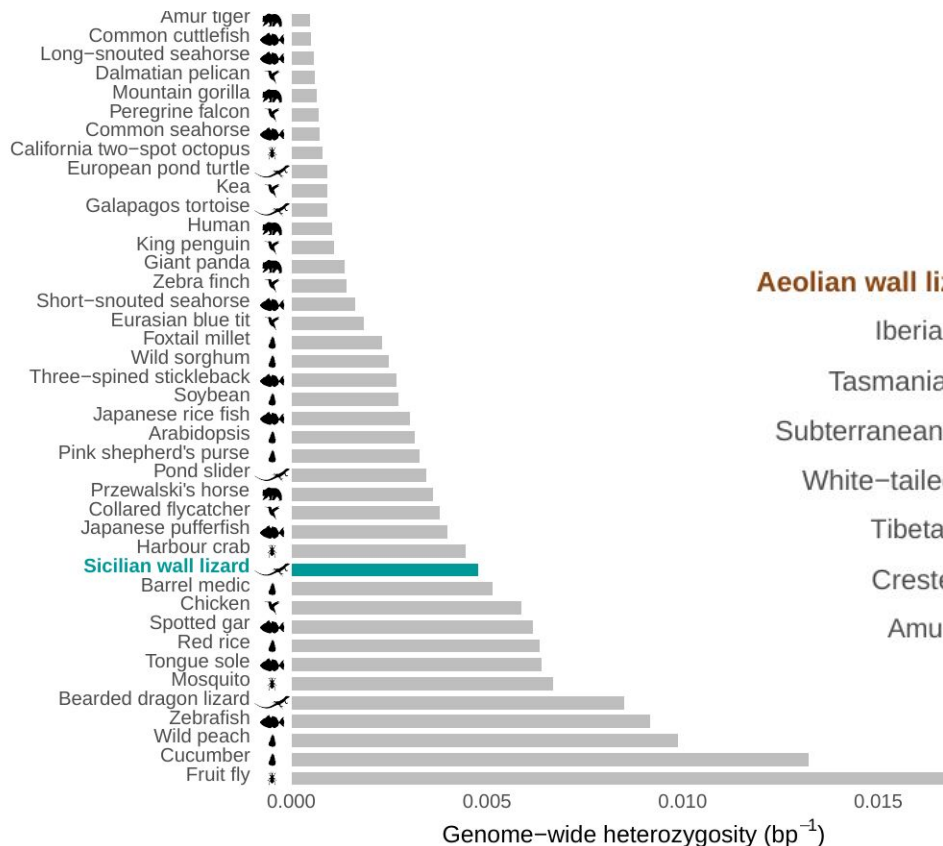
LIVING ON A ROCK AS A WILD INBRED STRAIN

Setting the limit for genomic variation and genetic load



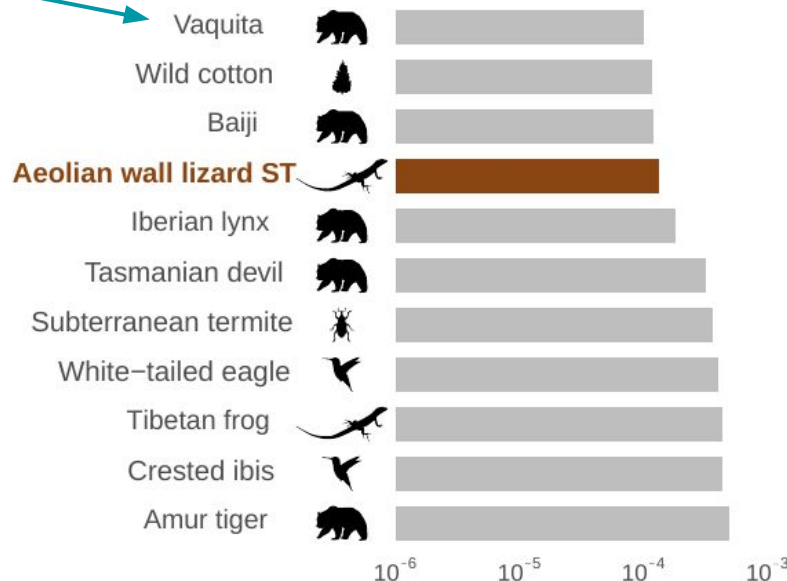
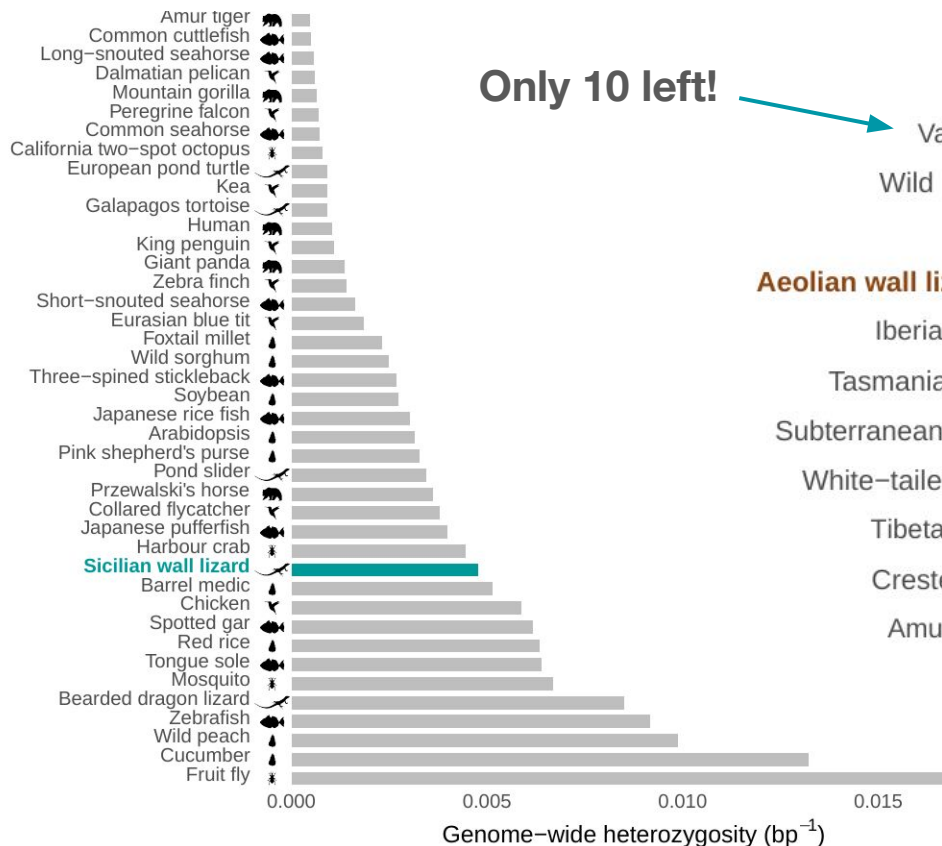
LIVING ON A ROCK AS A WILD INBRED STRAIN

Setting the limit for genomic variation and genetic load



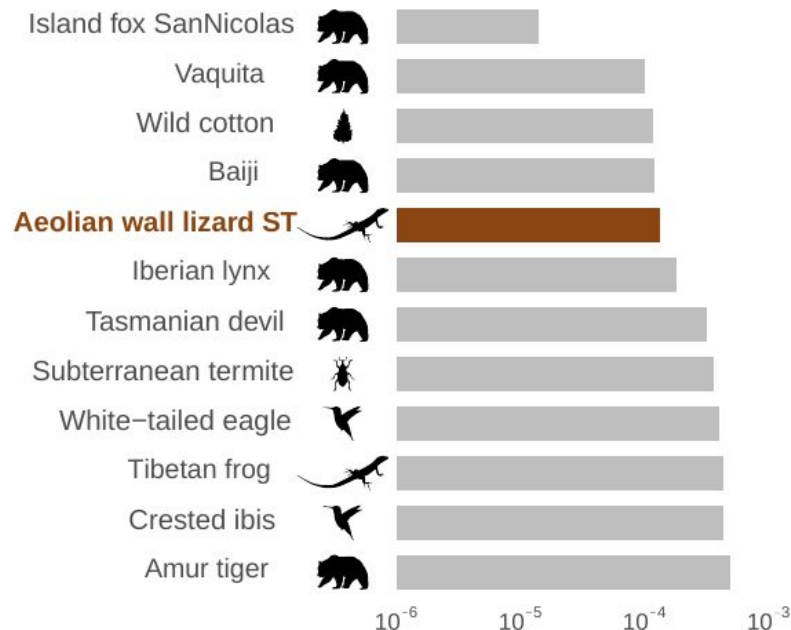
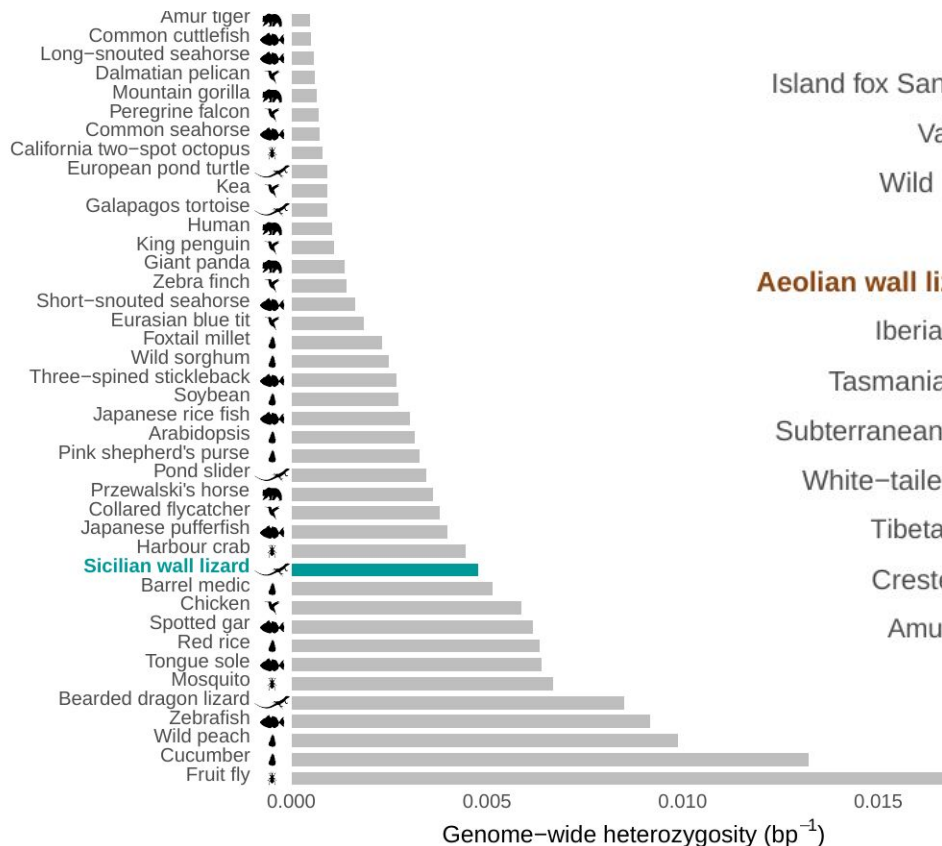
LIVING ON A ROCK AS A WILD INBRED STRAIN

Setting the limit for genomic variation and genetic load



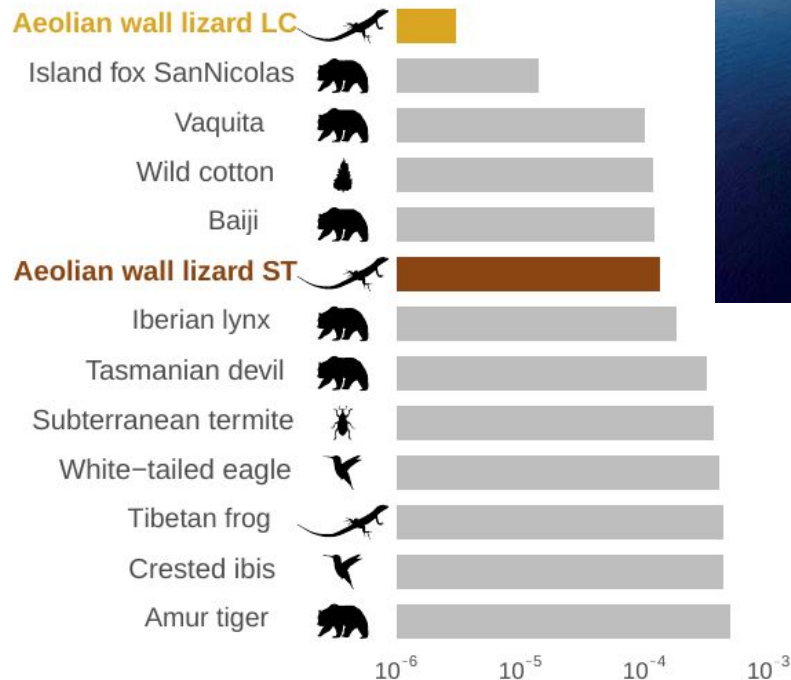
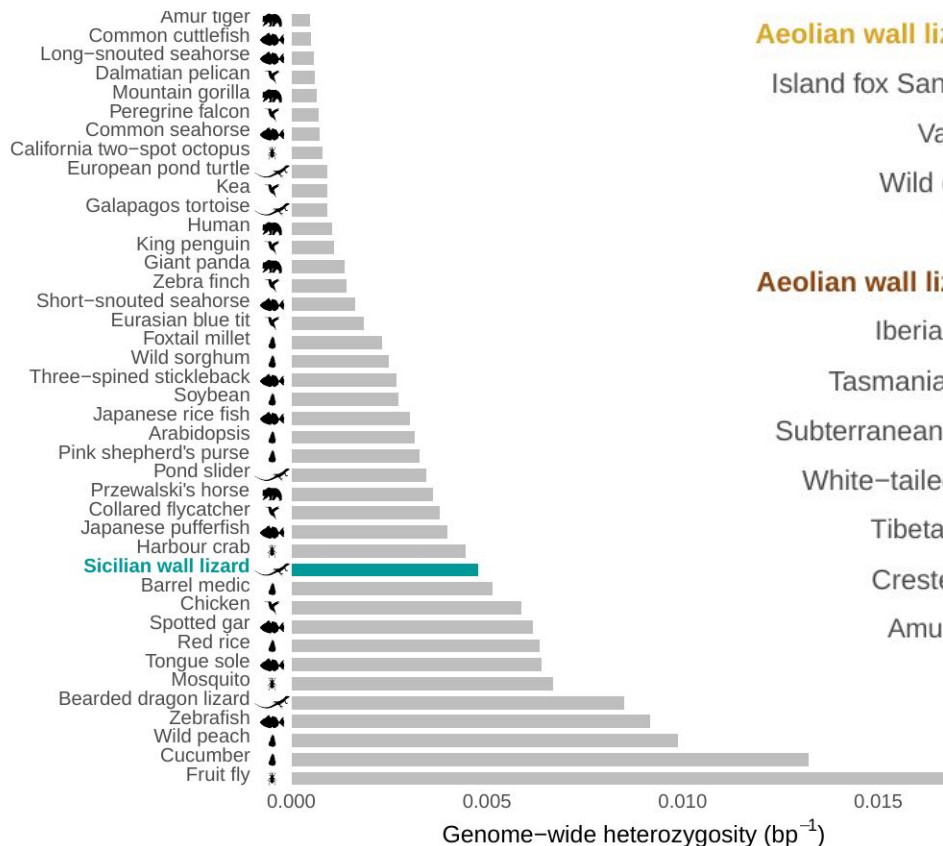
LIVING ON A ROCK AS A WILD INBRED STRAIN

Setting the limit for genomic variation and genetic load



LIVING ON A ROCK AS A WILD INBRED STRAIN

Setting the limit for genomic variation and genetic load



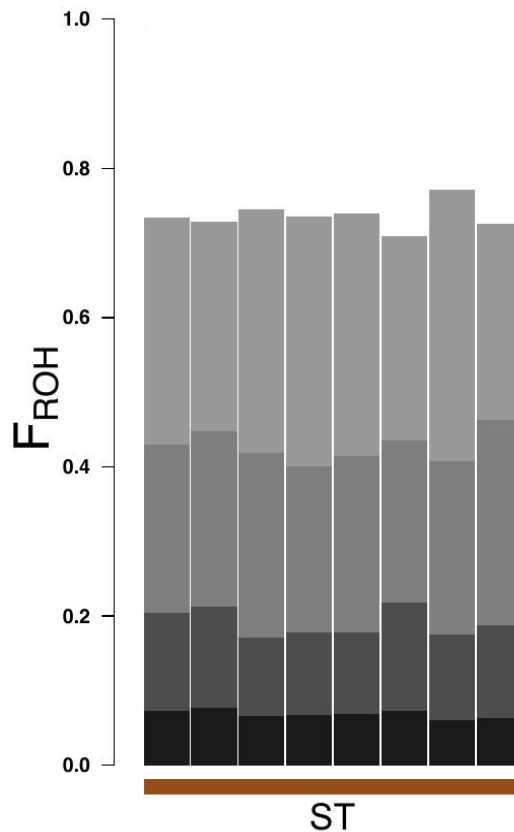
LIVING ON A ROCK AS A WILD INBRED STRAIN

Setting the limit for genomic variation and genetic load



LIVING ON A ROCK AS A WILD INBRED STRAIN

Setting the limit for genomic variation and genetic load

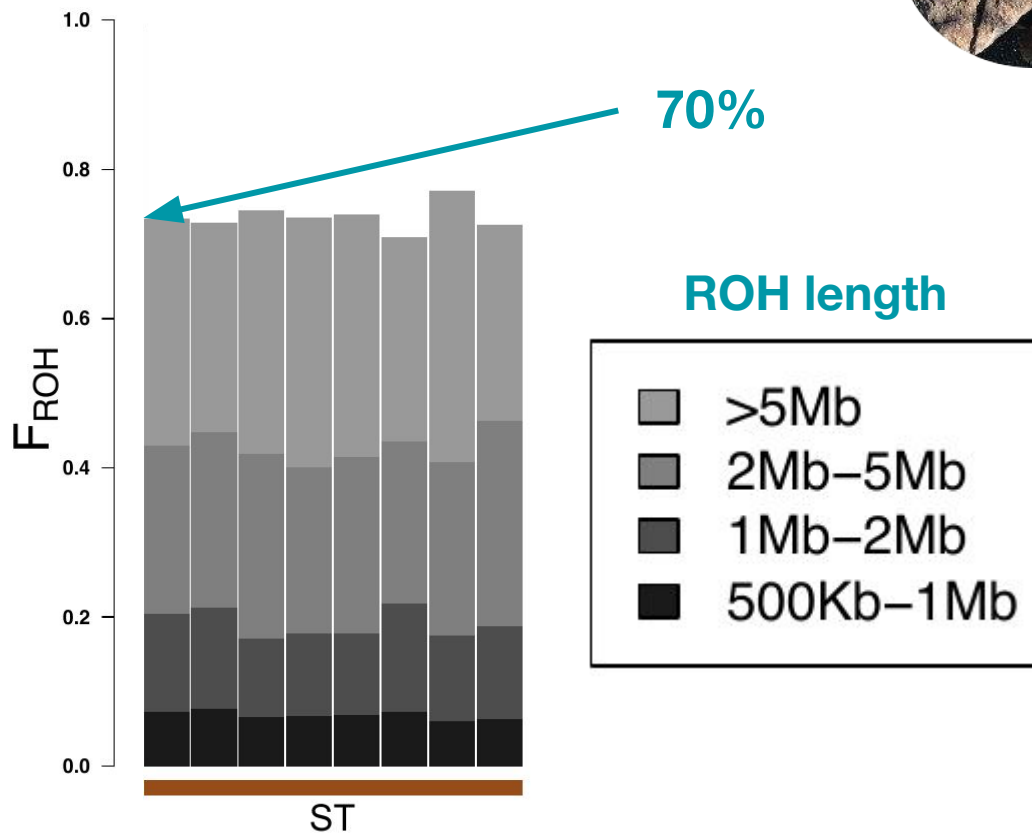


ROH length

- >5Mb
- 2Mb–5Mb
- 1Mb–2Mb
- 500Kb–1Mb

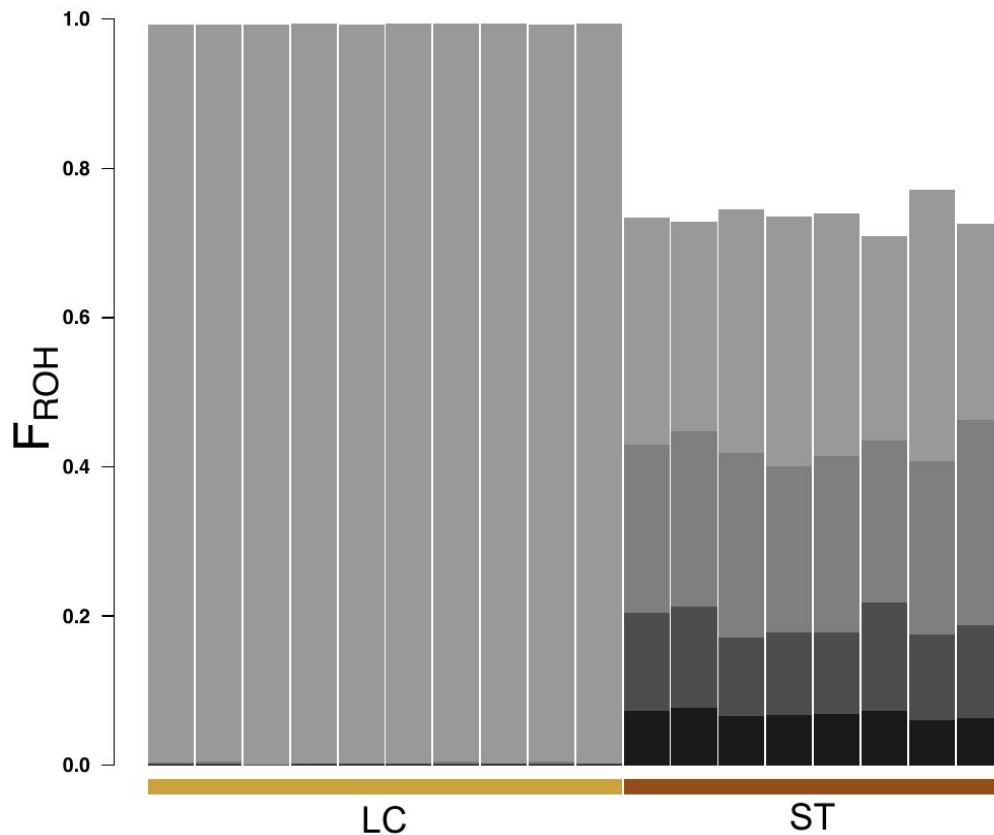
LIVING ON A ROCK AS A WILD INBRED STRAIN

Setting the limit for genomic variation and genetic load



LIVING ON A ROCK AS A WILD INBRED STRAIN

Setting the limit for genomic variation and genetic load

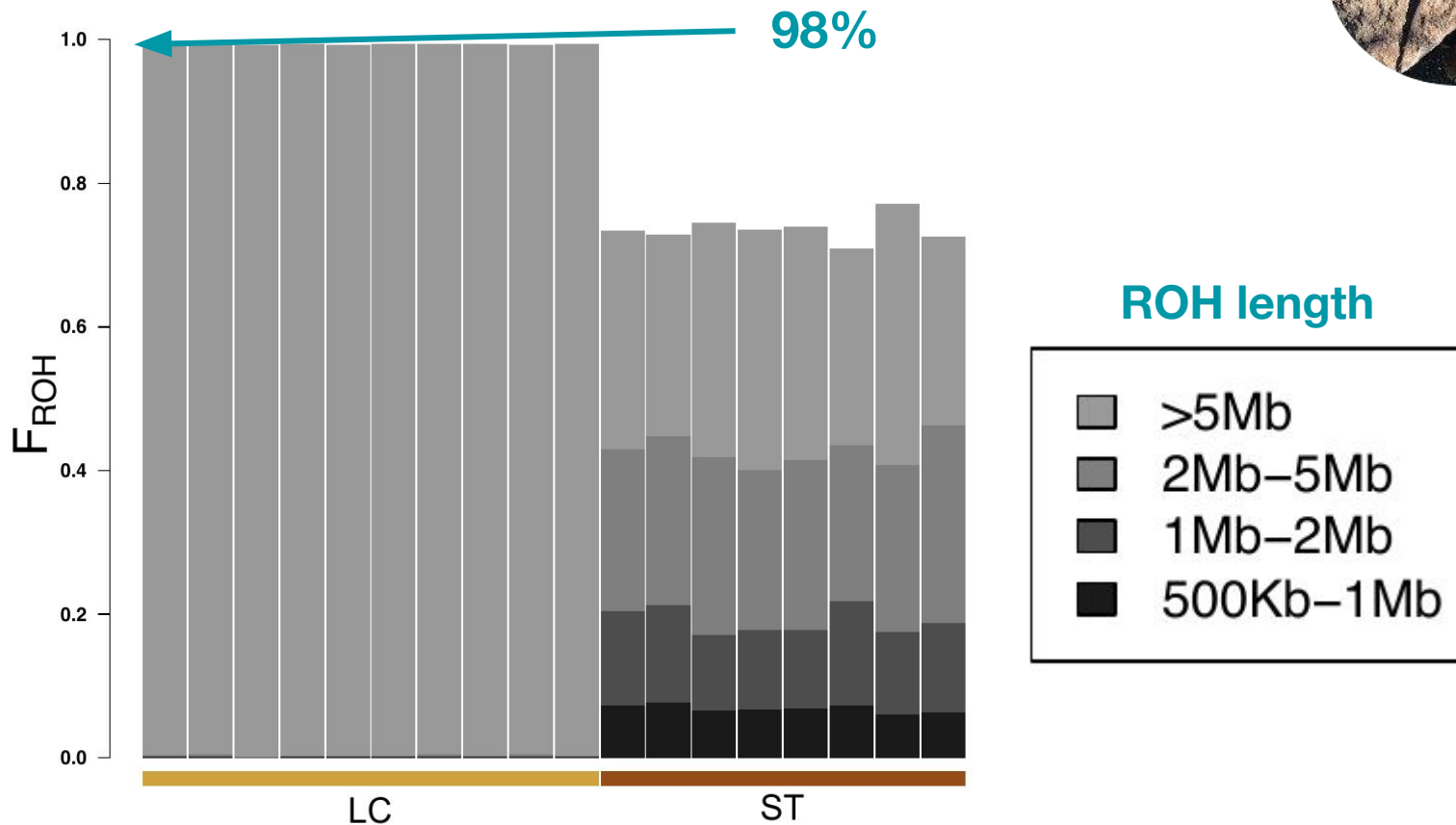


ROH length

- >5Mb
- 2Mb–5Mb
- 1Mb–2Mb
- 500Kb–1Mb

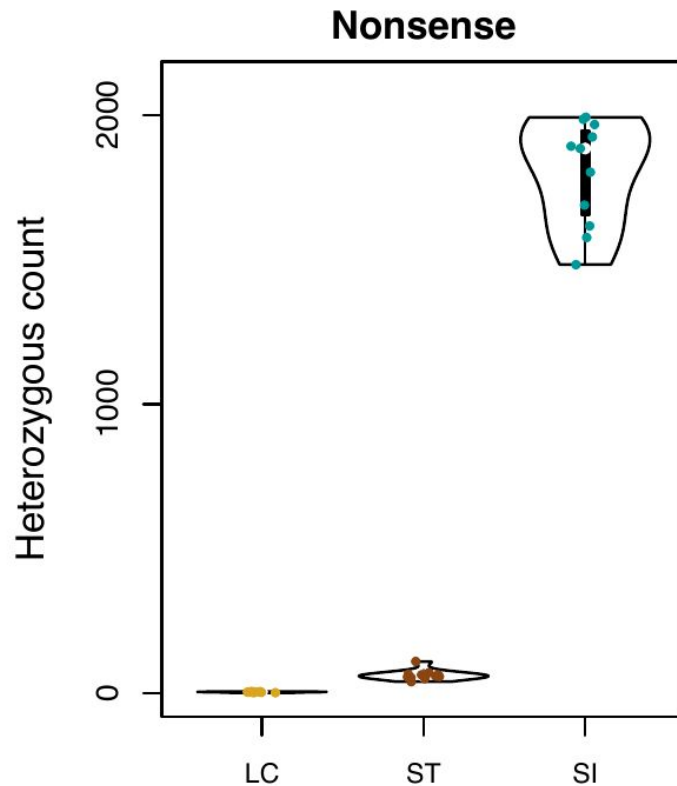
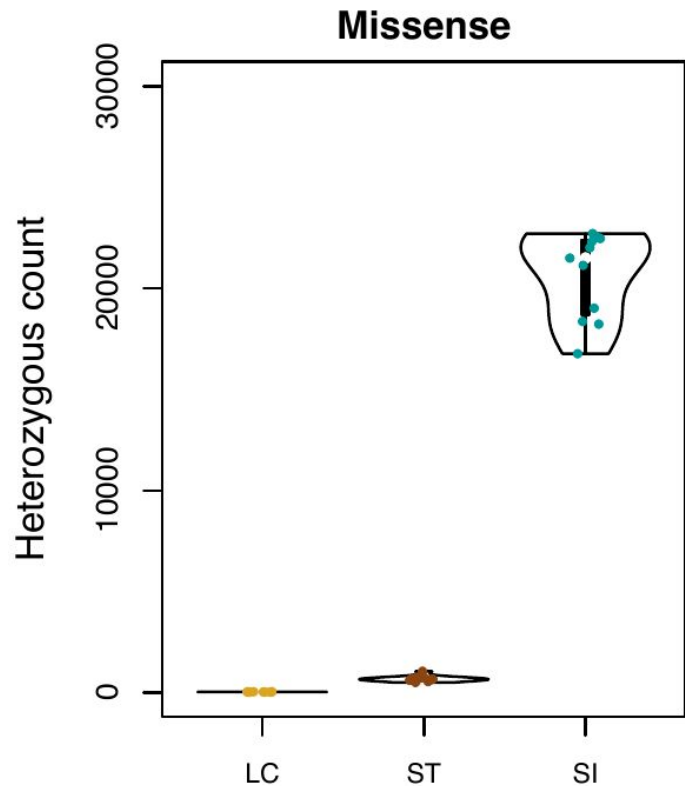
LIVING ON A ROCK AS A WILD INBRED STRAIN

Setting the limit for genomic variation and genetic load



LIVING ON A ROCK AS A WILD INBRED STRAIN

Setting the limit for genomic variation and genetic load



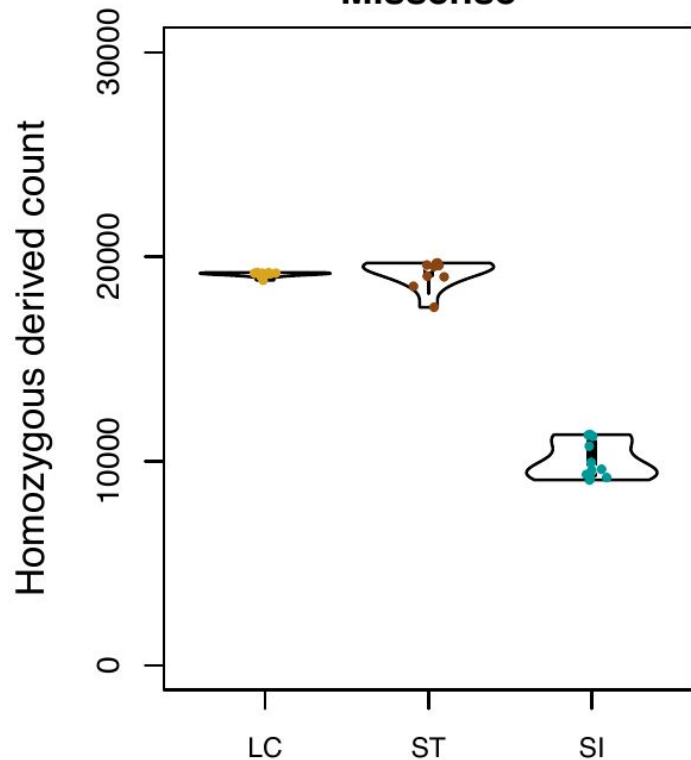
MASKED
GENETIC LOAD

LIVING ON A ROCK AS A WILD INBRED STRAIN

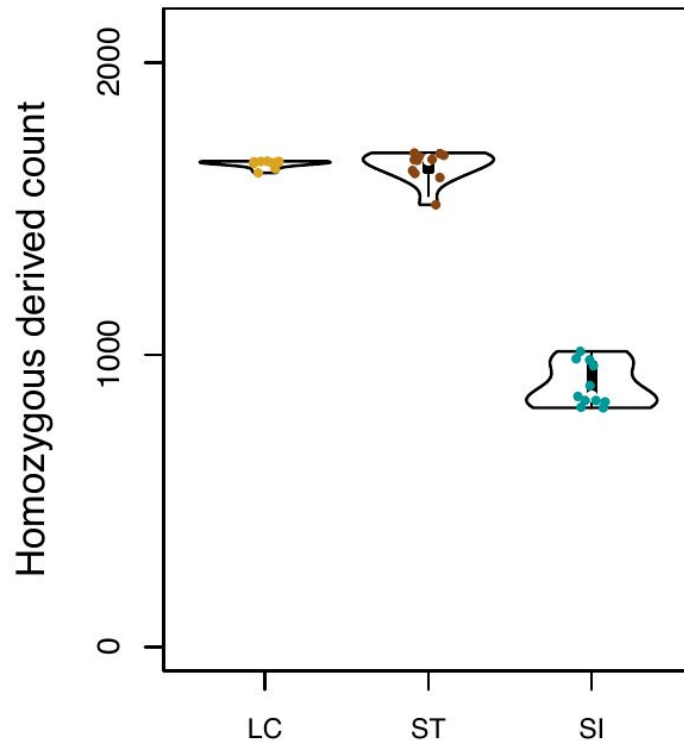
Setting the limit for genomic variation and genetic load



Missense



Nonsense



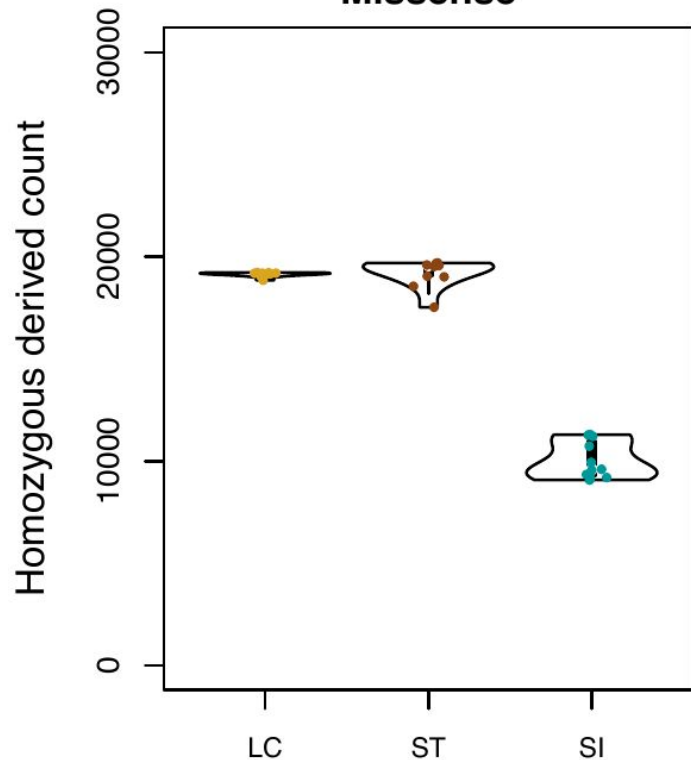
REALIZED
GENETIC LOAD

LIVING ON A ROCK AS A WILD INBRED STRAIN

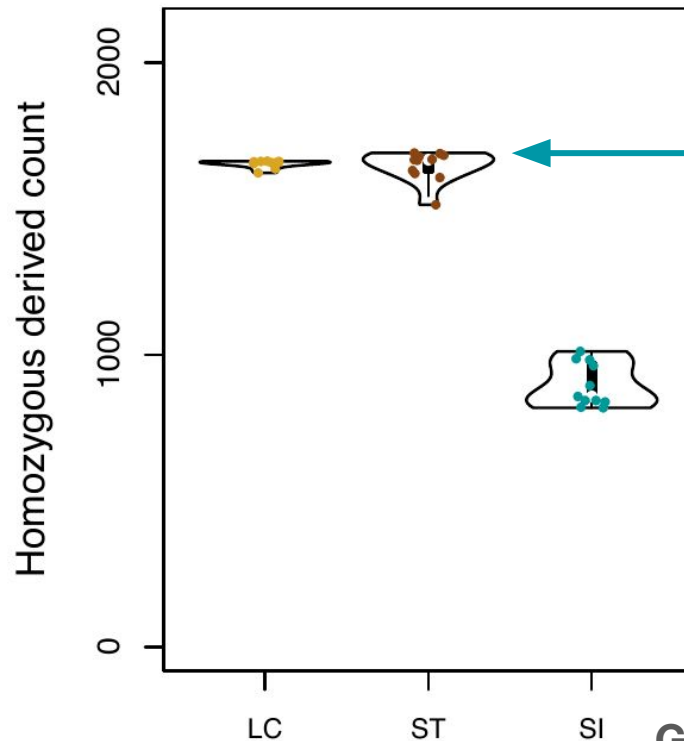
Setting the limit for genomic variation and genetic load



Missense



Nonsense



MAX
THRESHOLD?

REALIZED
GENETIC LOAD

TESTING PREDICTIONS OF DELETERIOUSNESS

Impaired bioenergetics in the Apennine brown bear



Slovakian

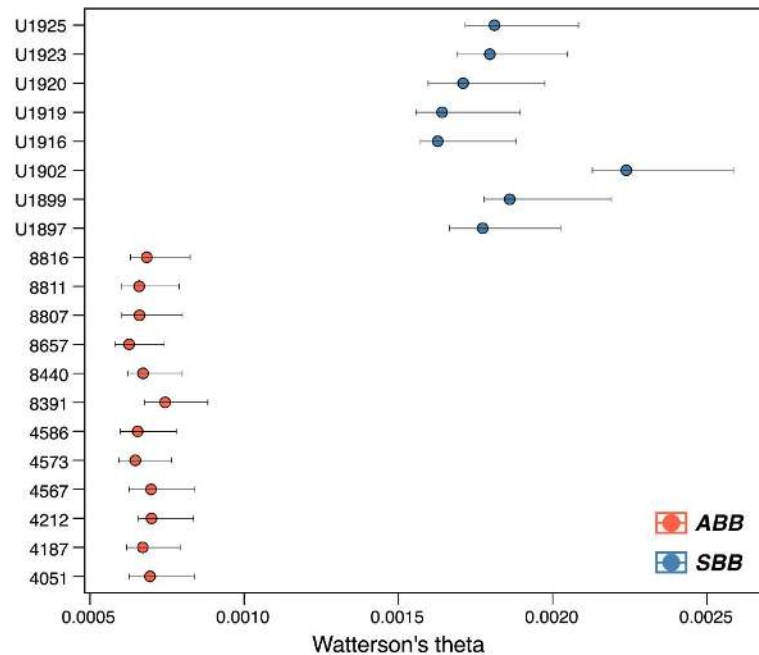
Apennine

TESTING PREDICTIONS OF DELETERIOUSNESS

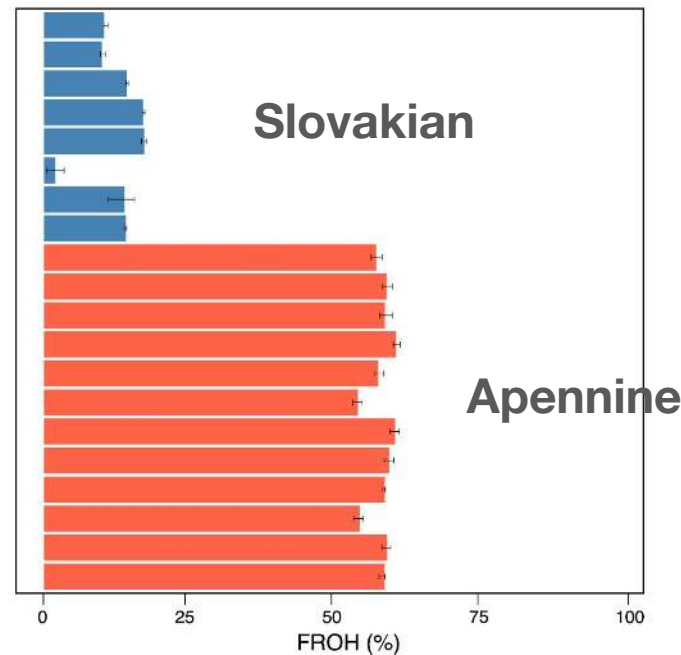
Impaired bioenergetics in the Apennine brown bear



a)

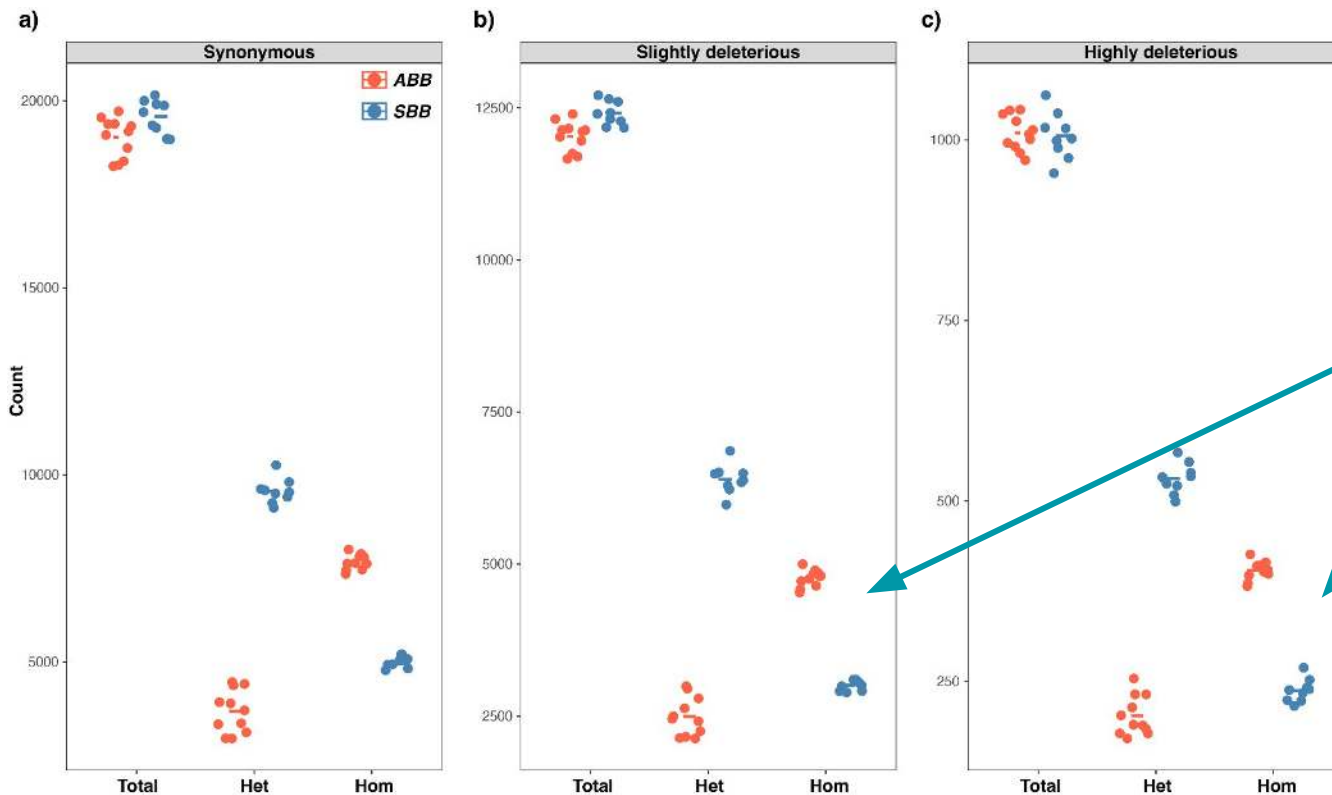


b)



TESTING PREDICTIONS OF DELETERIOUSNESS

Impaired bioenergetics in the Apennine brown bear



**REALIZED
GENETIC LOAD**

Fabbri et al *in prep*

TESTING PREDICTIONS OF DELETERIOUSNESS

Impaired bioenergetics in the Apennine brown bear



Ca. 40 **predicted deleterious** substitution in the Apennine bears

TESTING PREDICTIONS OF DELETERIOUSNESS

Impaired bioenergetics in the Apennine brown bear



3 in the mitochondrial ND5 -> **RESPIRATORY COMPLEX I**

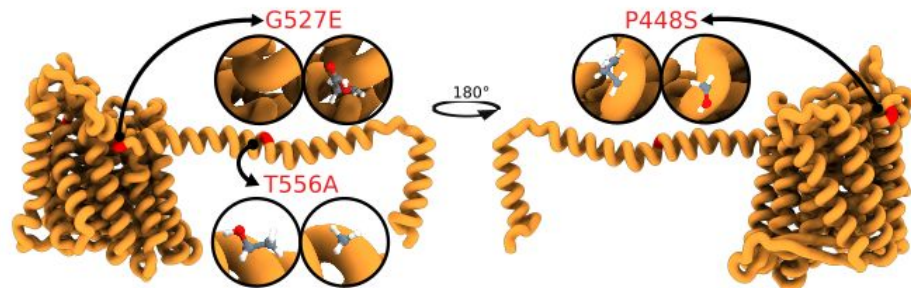
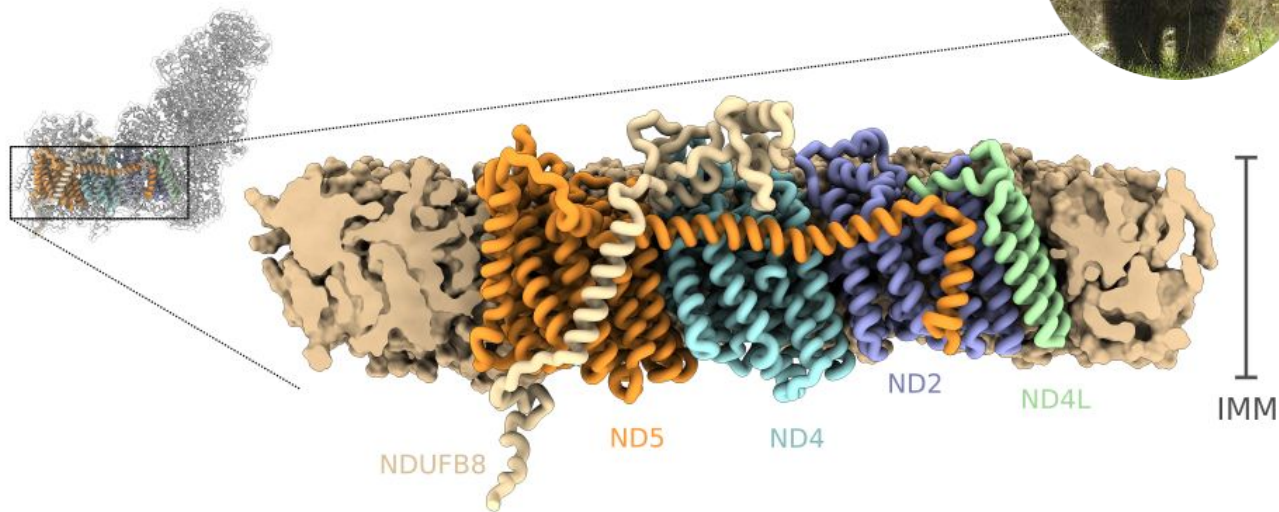
	P448S	G527E	T556A
APOGEE			
CADD			
CAROL			
Condel			
COVEC WMV			
EFIN HD			
EFIN SP			
FatHmm			
Meta-SNP			
Mitoclass.1			
MtoolBox			
MutAssessor			
MutTaster			
PANTHER			
PhD-SNP			
PolyPhen2			
PROVEAN			
SIFT			
SNAP			

TESTING PREDICTIONS OF DELETERIOUSNESS

Impaired bioenergetics in the Apennine brown bear



	P448S	G527E	T556A
APOGEE			
CADD			
CAROL			
Condel			
COVEC WMV			
EFIN HD			
EFIN SP			
FatHmm			
Meta-SNP			
Mitoclass.1			
MtoolBox			
MutAssessor			
MutTaster			
PANTHER			
PhD-SNP			
PolyPhen2			
PROVEAN			
SIFT			
SNAP			



TESTING PREDICTIONS OF DELETERIOUSNESS

Impaired bioenergetics in the Apennine brown bear



		ND5 positions			Experiments		
		448	527	556	MDCK cells	Fibroblasts	MD simulations
Apennine		S	E	A	✓	✓	✓
S448		S	G	T	✓		
E527		P	E	T	✓		✓
A556		P	G	A	✓		
CTRL		P	G	T	✓		✓
Alpine		S	G	A		✓	

TESTING PREDICTIONS OF DELETERIOUSNESS

Impaired bioenergetics in the Apennine brown bear



		ND5 positions			Experiments		
		448	527	556	MDCK cells	Fibroblasts	MD simulations
Apennine		S	E	A	✓	✓	✓
S448		S	G	T	✓		
E527		P	E	T	✓		✓
A556		P	G	A	✓		
CTRL		P	G	T	✓		✓
Alpine		S	G	A		✓	

TESTING PREDICTIONS OF DELETERIOUSNESS

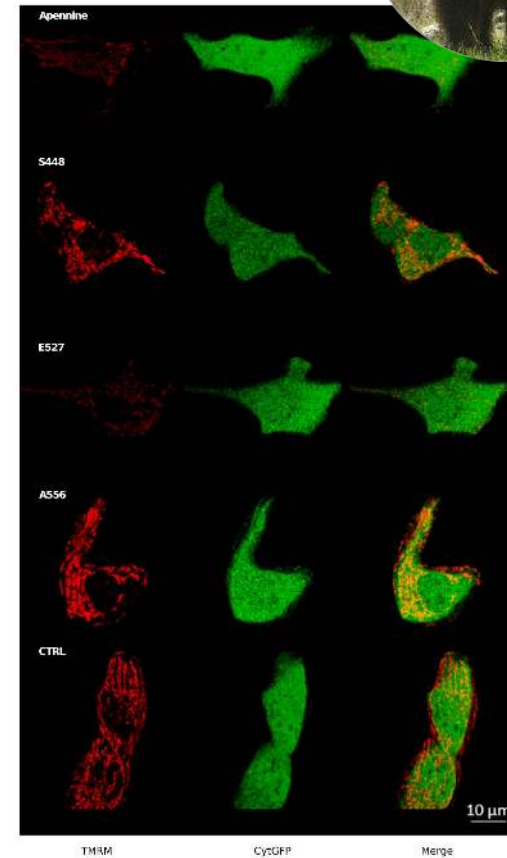
Impaired bioenergetics in the Apennine brown bear



		ND5 positions			Experiments		
		448	527	556	MDCK cells	Fibroblasts	MD simulations
Apennine		← S	E	A	✓	✓	✓
S448		S	G	T	✓		
E527		← P	E	T	✓		✓
A556		P	G	A	✓		
CTRL		P	G	T	✓		✓
Alpine		S	G	A		✓	

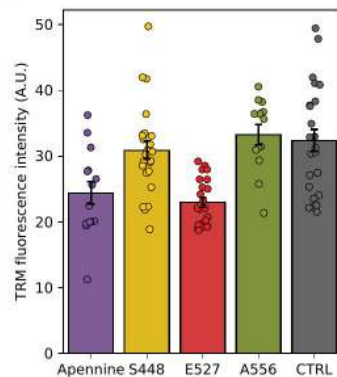
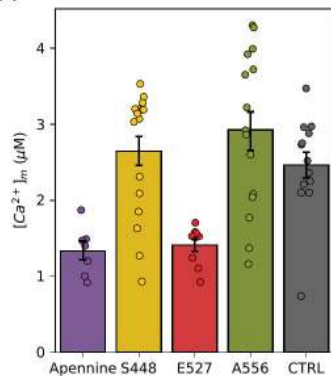
EXPERIMENT IN VITRO - MDCK CELLS

Impaired bioenergetics in the Apennine brown bear

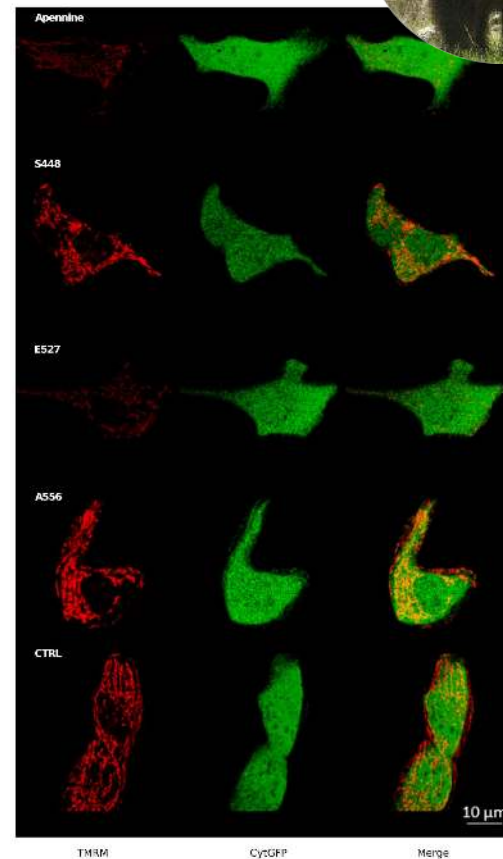


EXPERIMENT IN VITRO - MDCK CELLS

Impaired bioenergetics in the Apennine brown bear

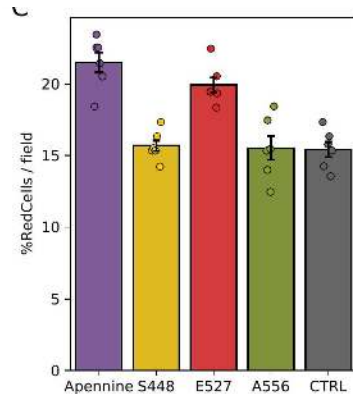
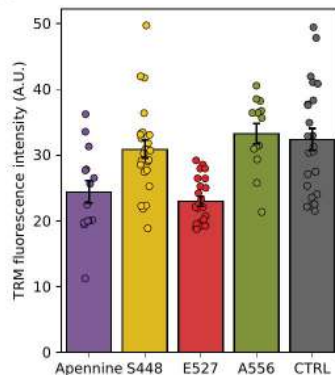
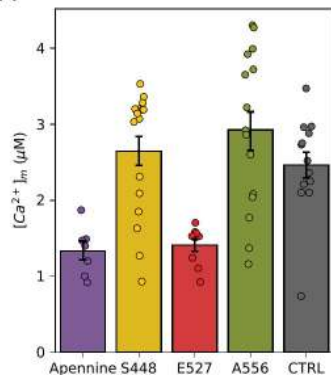


Transmembrane potential is lower



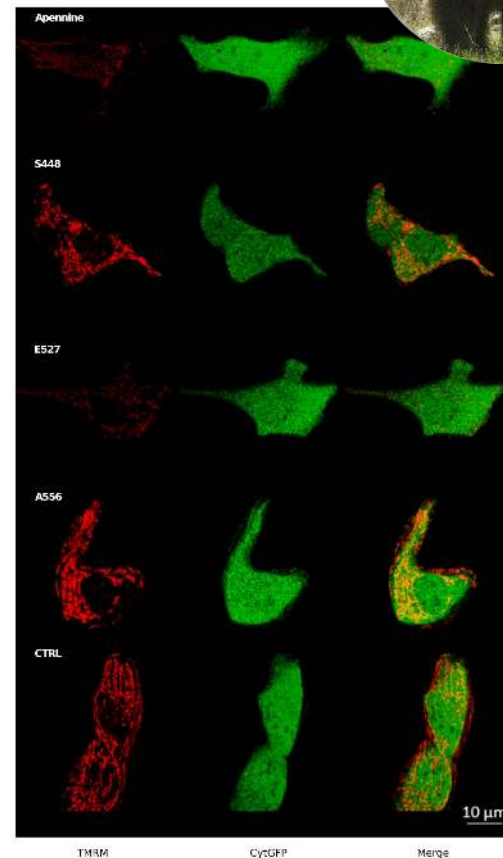
EXPERIMENT IN VITRO - MDCK CELLS

Impaired bioenergetics in the Apennine brown bear



Transmembrane potential is lower

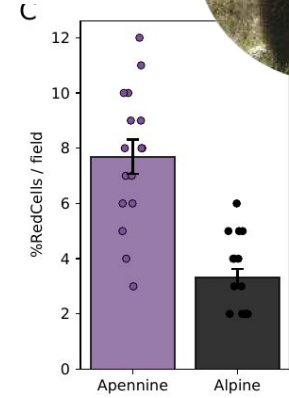
ROS production is higher



EXPERIMENT EX VIVO - FIBROBLASTS

Impaired bioenergetics in the Apennine brown bear

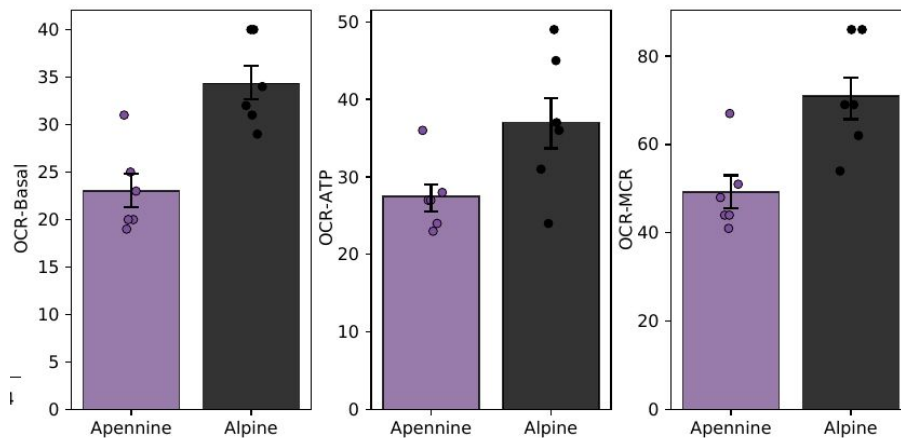
ROS
are higher



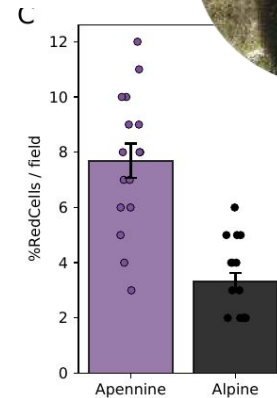
EXPERIMENT EX VIVO - FIBROBLASTS

Impaired bioenergetics in the Apennine brown bear

ATP production
is lower



ROS
are higher

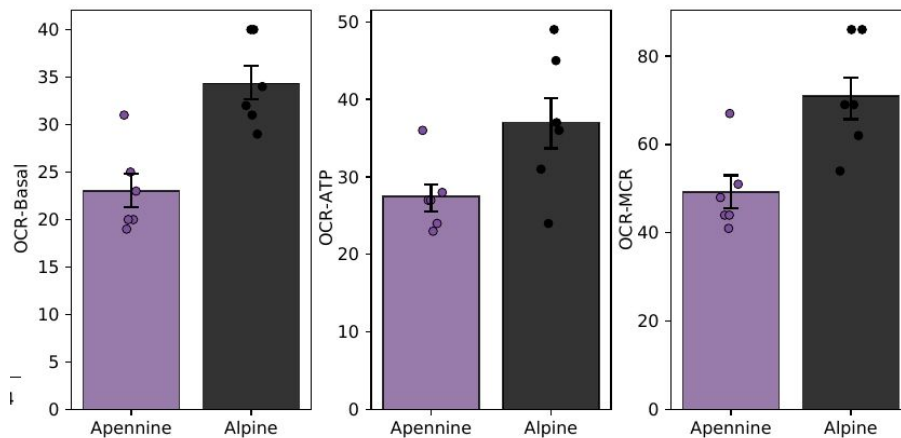


EXPERIMENT EX VIVO - FIBROBLASTS

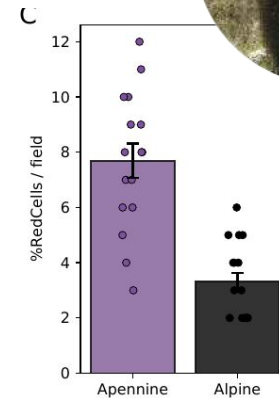
Impaired bioenergetics in the Apennine brown bear



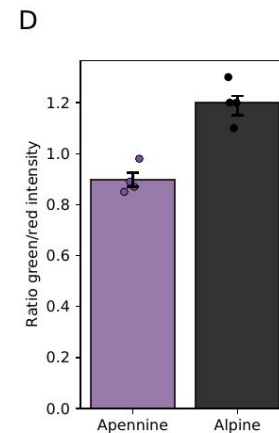
ATP production
is lower



ROS
are higher



MT
are older



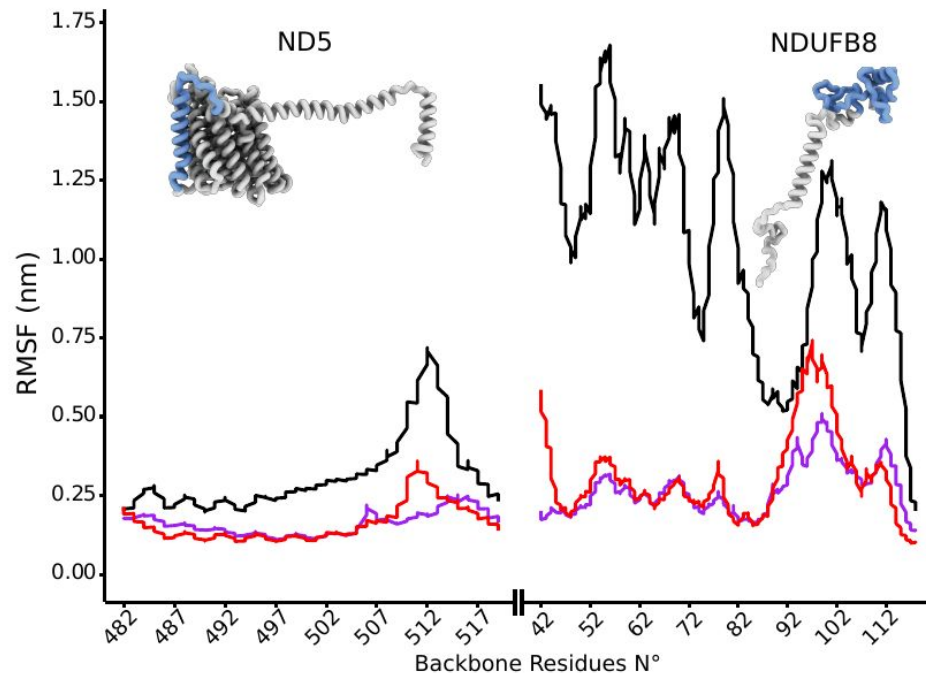
EXPERIMENT IN SILICO - MOLECULAR DYNAMICS

Impaired bioenergetics in the Apennine brown bear



EXPERIMENT IN SILICO - MOLECULAR DYNAMICS

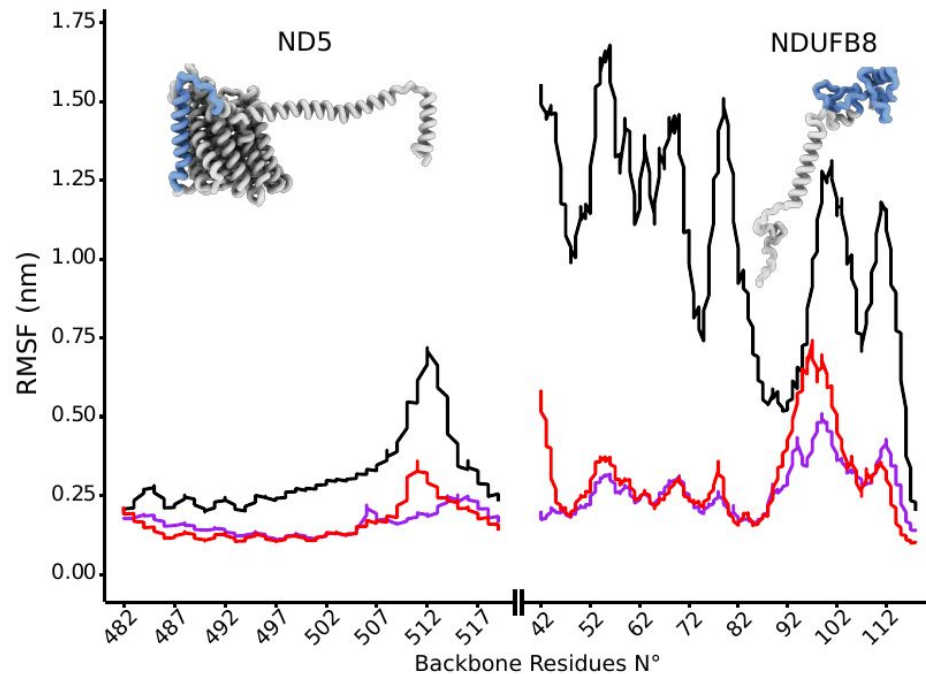
Impaired bioenergetics in the Apennine brown bear



Rigidity
increases

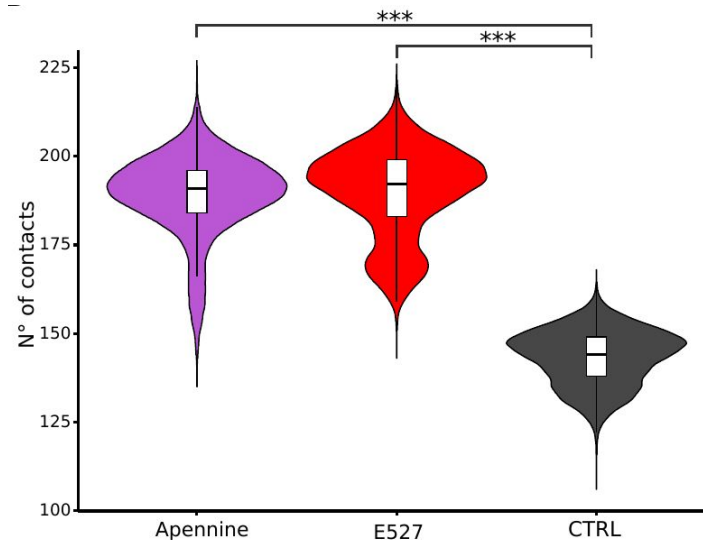
EXPERIMENT IN SILICO - MOLECULAR DYNAMICS

Impaired bioenergetics in the Apennine brown bear



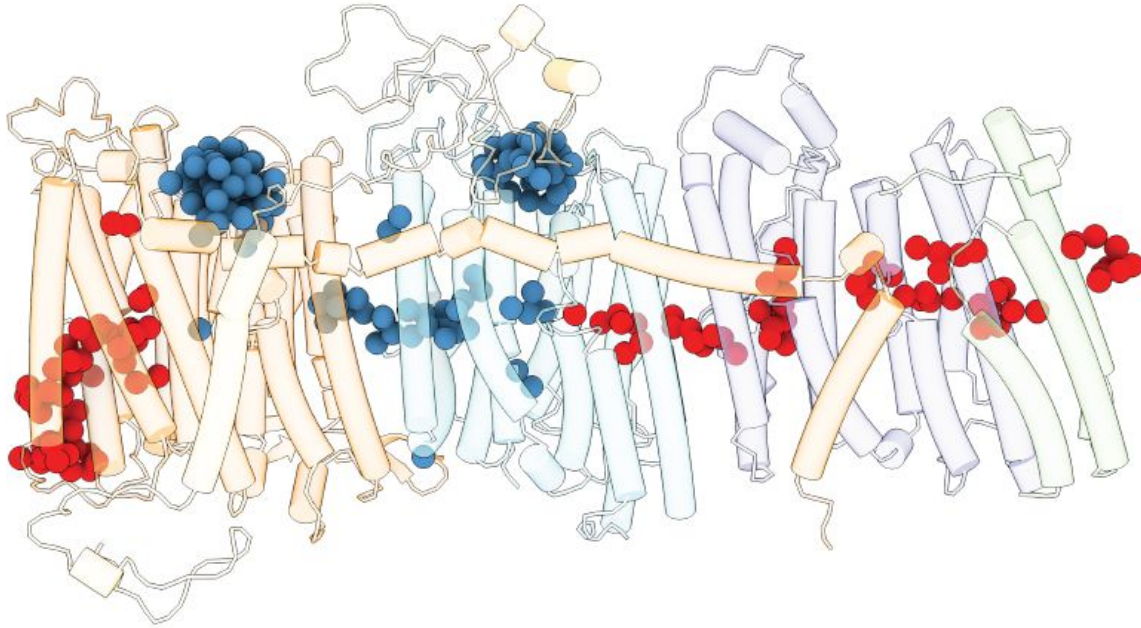
Rigidity
increases

Number of contacts
Is higher



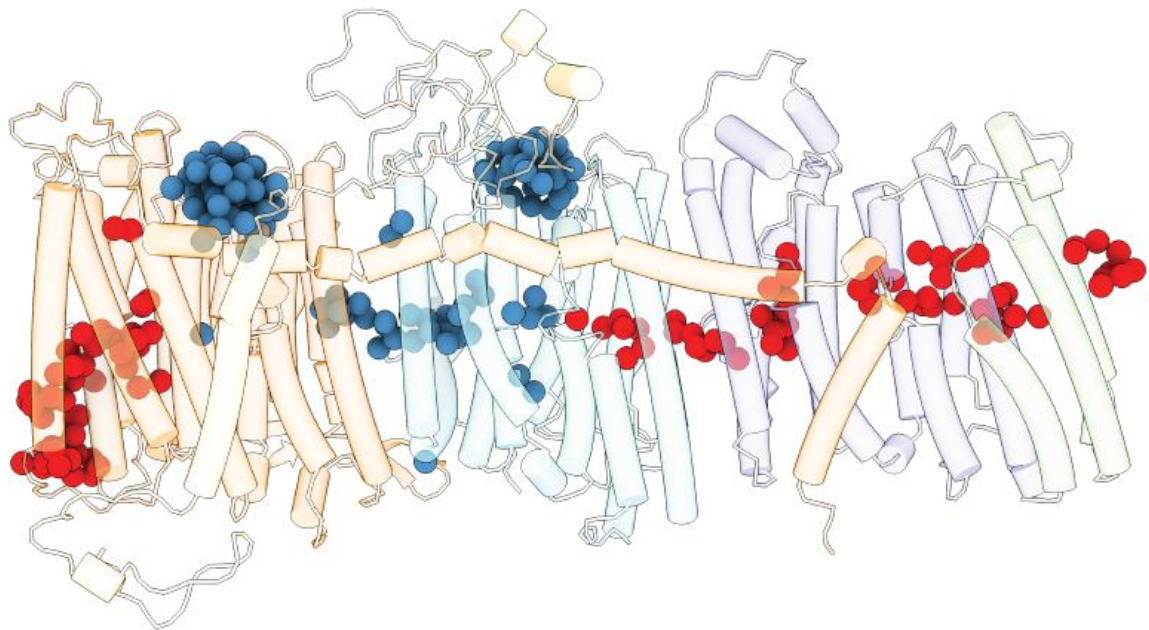
EXPERIMENT IN SILICO - MOLECULAR DYNAMICS

Impaired bioenergetics in the Apennine brown bear

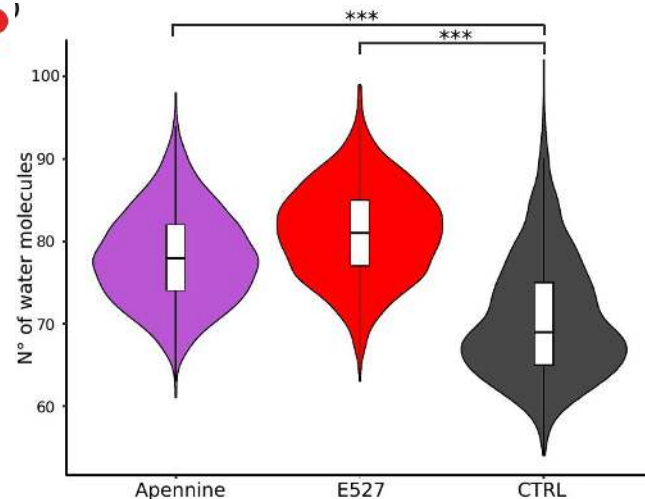


EXPERIMENT IN SILICO - MOLECULAR DYNAMICS

Impaired bioenergetics in the Apennine brown bear



Hydration
increases



Trucchi et al *submitted*

THE ENDEMIXIT FOLKS

Project coordinator: Giorgio Bertorelle, University of Ferrara



Univ. Ferrara

Andrea Benazzo
Silvia Fuselli
Maeva Gabrielli
Francesca Raffini
Victor Munoz Mora
Francesco Maroso
Roberto Biello
Paolo Pinton
Simone Patergnani
Alberto Danese
Giada Padovani
Giulia Fabbri

Marche Politec. Univ.

Emiliano Trucchi
Lorena Ancona
Francesco Giannelli
Sebastiano Fava
Nicola Rauli
Sara Laghetti

Univ. Tor Vergata

Valerio Sbordon
Donatella Cesaroni
Paolo Gratton

Univ. Firenze

Claudio Ciofi
Alessio Iannucci

Univ. Trieste

Marco Gerdol
Alberto Pallavicini
Samuele Greco
Marco Sollitto
M Barsid

Univ. Padova

Leonardo Congiu
Stefano Dalle Palle

FUTURE DIRECTIONS

FUTURE DIRECTIONS

INTEGRATE ADVANCED MOLECULAR BIOLOGY METHODS UP TO
GENETIC EDITING TO CORRECT FIXED DELETERIOUS MUTATIONS
IN ENDANGERED SPECIES AND SYNTHETIC BIOLOGY UP TO
DE-EXTINCTION

FUTURE DIRECTIONS



FUTURE DIRECTIONS

INTEGRATE ADVANCED MOLECULAR BIOLOGY METHODS UP TO
GENETIC EDITING TO CORRECT FIXED DELETERIOUS MUTATIONS
IN ENDANGERED SPECIES AND SYNTHETIC BIOLOGY UP TO
DE-EXTINCTION

IMPROVE ADAPTIVE AND MALADAPTIVE (GENETIC LOAD)
ESTIMATES **MAKING THEM COMPARABLE** ACROSS POPULATIONS
AND SPECIES

FUTURE DIRECTIONS

INTEGRATE ADVANCED MOLECULAR BIOLOGY METHODS UP TO GENETIC EDITING TO CORRECT FIXED DELETERIOUS MUTATIONS IN ENDANGERED SPECIES AND SYNTHETIC BIOLOGY UP TO DE-EXTINCTION

IMPROVE ADAPTIVE AND MALADAPTIVE (GENETIC LOAD) ESTIMATES MAKING THEM COMPARABLE ACROSS POPULATIONS AND SPECIES

LEVERAGE NEW COST-EFFECTIVE LONG-RANGE SEQUENCING TO ANALYZE ALSO SVS, METHYLATION DATA, ETC

FUTURE DIRECTIONS

INTEGRATE ADVANCED MOLECULAR BIOLOGY METHODS UP TO GENETIC EDITING TO CORRECT FIXED DELETERIOUS MUTATIONS IN ENDANGERED SPECIES AND SYNTHETIC BIOLOGY UP TO DE-EXTINCTION

IMPROVE ADAPTIVE AND MALADAPTIVE (GENETIC LOAD) ESTIMATES MAKING THEM COMPARABLE ACROSS POPULATIONS AND SPECIES

LEVERAGE NEW COST-EFFECTIVE LONG-RANGE SEQUENCING TO ANALYZE ALSO SVS, METHYLATION DATA, ETC

IMPROVE PREDICTION OF **GENOMIC FEATURES WITH AI**

CONSERVATION GENOMICS



QUESTIONS?