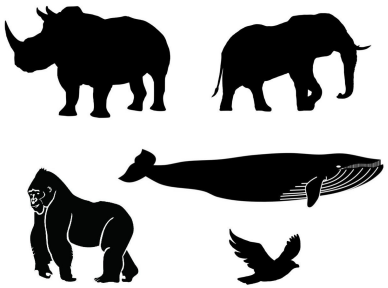
CPG Centre for Palaeogenetics

Stockholms universitet

Naturhistoriska riksmuseet

GenErode

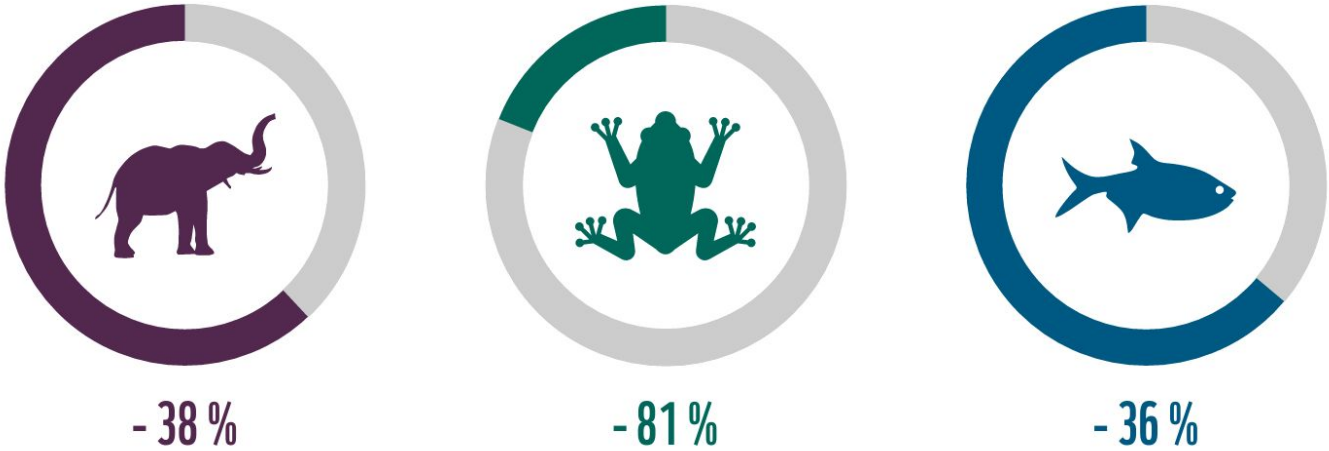
A bioinformatics pipeline to study genomic erosion



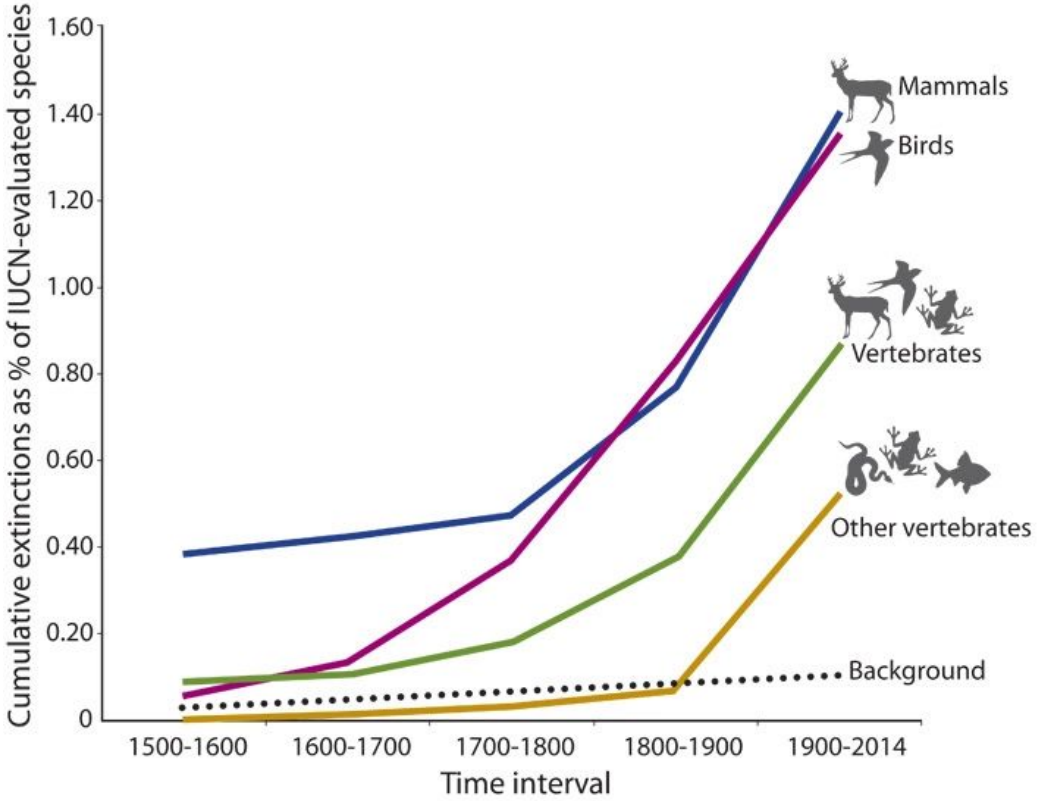
David Díez del Molino
Centre for Palaeogenetics

@indianadiez

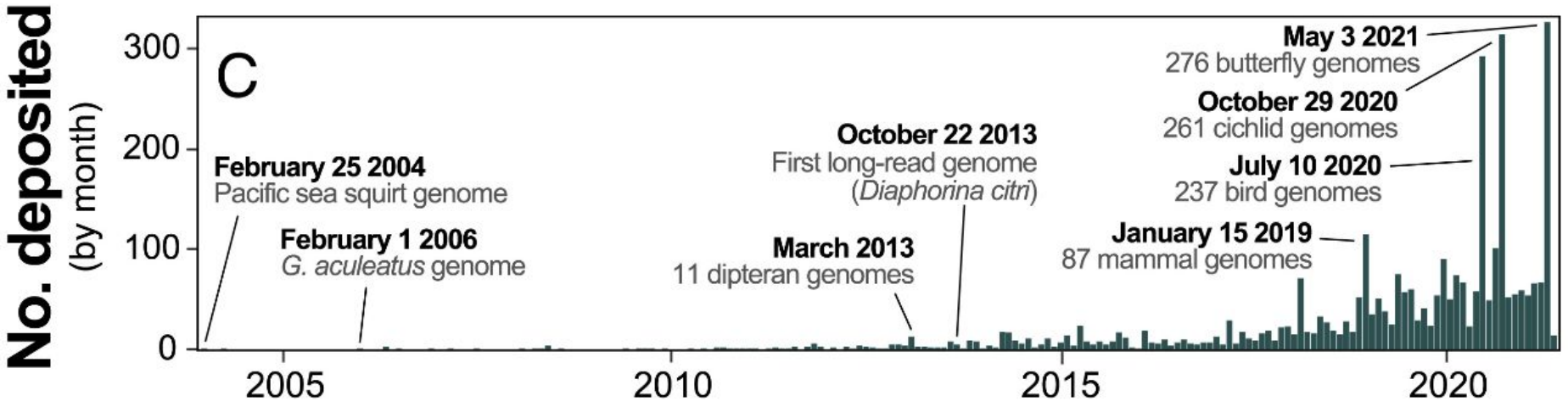
Dramatic population size reductions in the last 200 years



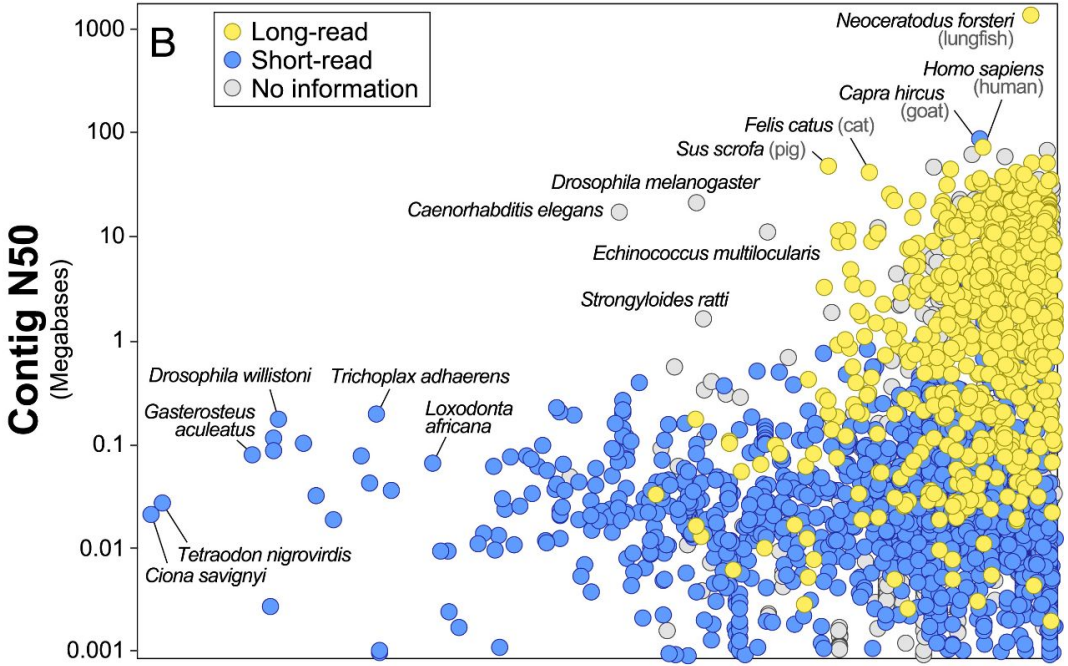
WWF
LIVING PLANET REPORT 2016

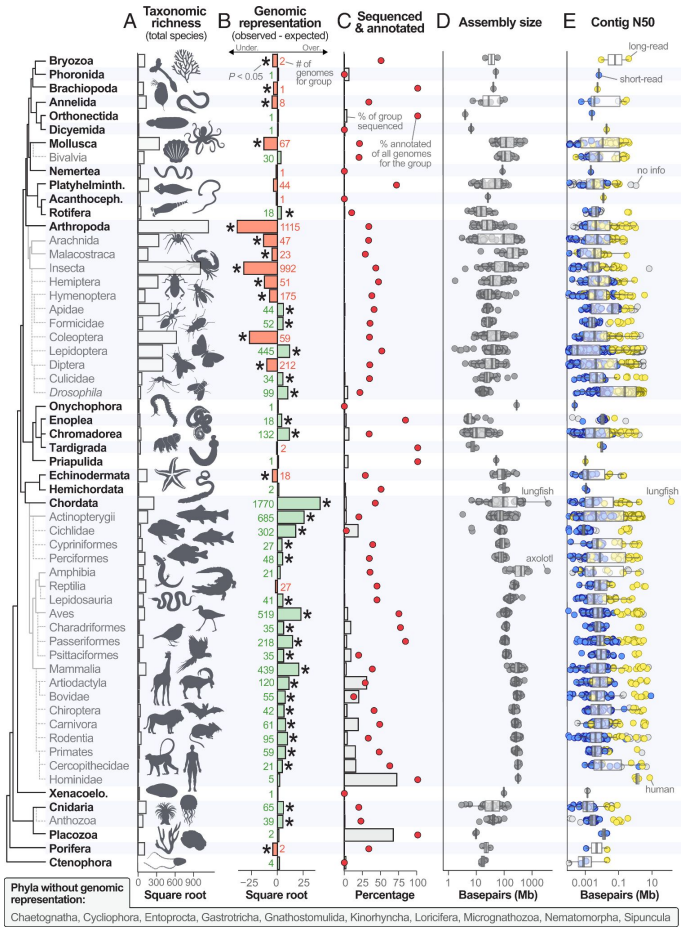


Ceballos et al. 2015



Hotaling et al. 2021





Hotaling et al. 2021

EARTH BIOGENOME PROJECT

VGP VERTEBRATE GENOMES PROJECT
A PROJECT OF THE G10K CONSORTIUM

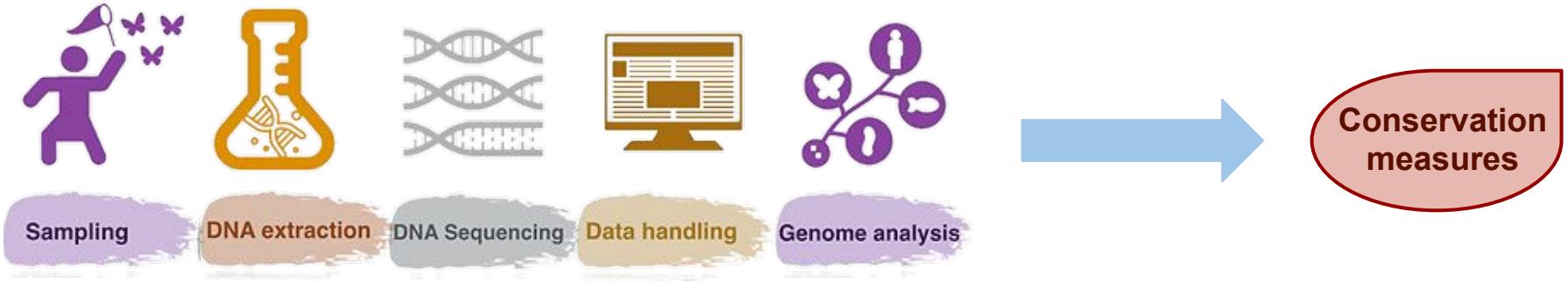
B10K

ERGA
EUROPEAN REFERENCE GENOME ATLAS

BAT 1K

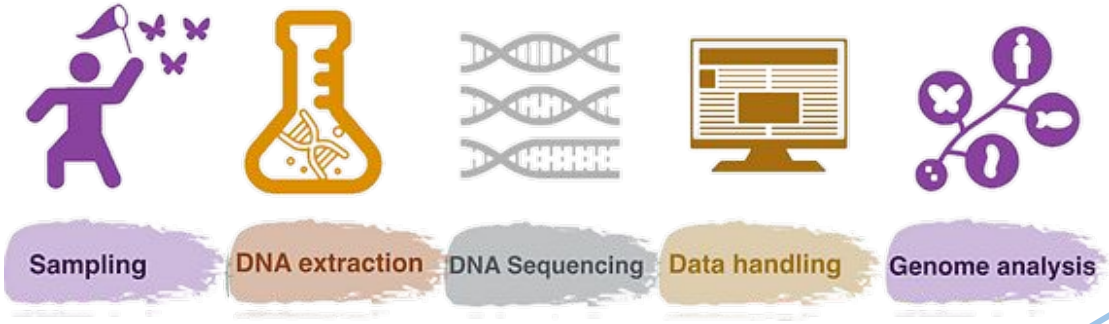
Darwin TREE of LIFE

Lack of reference genomes problem solved!

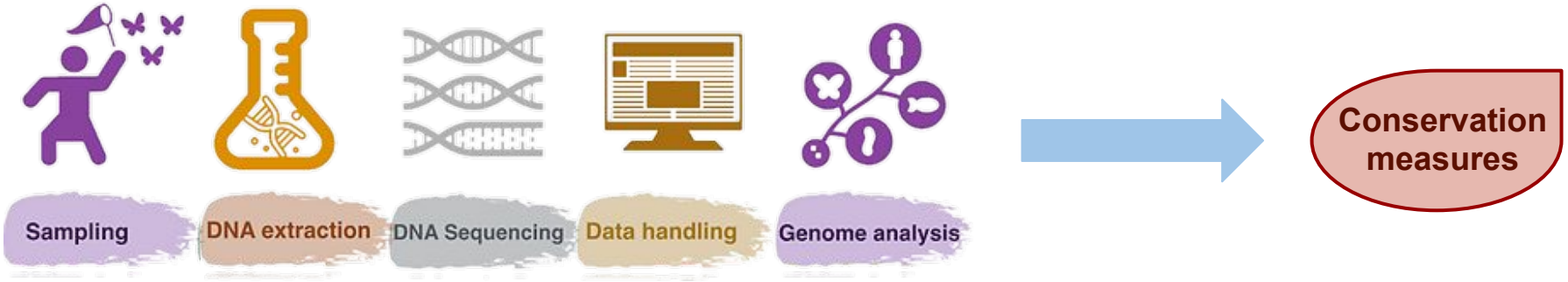


Lack of reference genomes problem solved!

- Funding agencies
- General public
- Conservation practitioners
- Policy makers

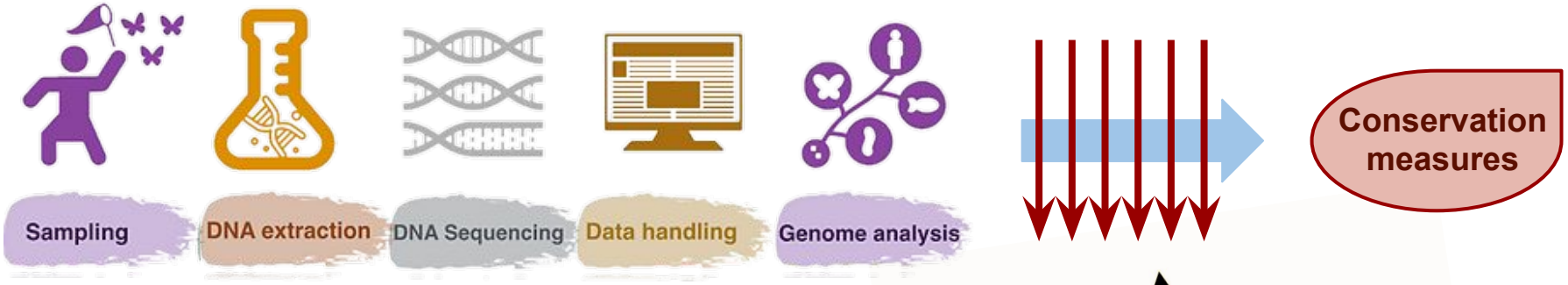


Conservation measures



Reproducibility

Standardization

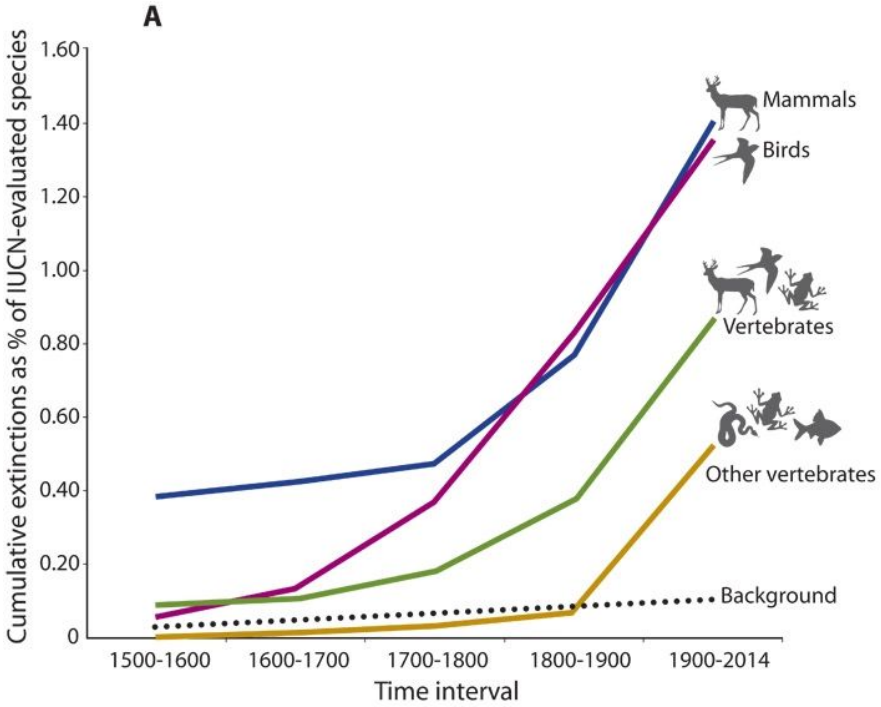


Reproducibility

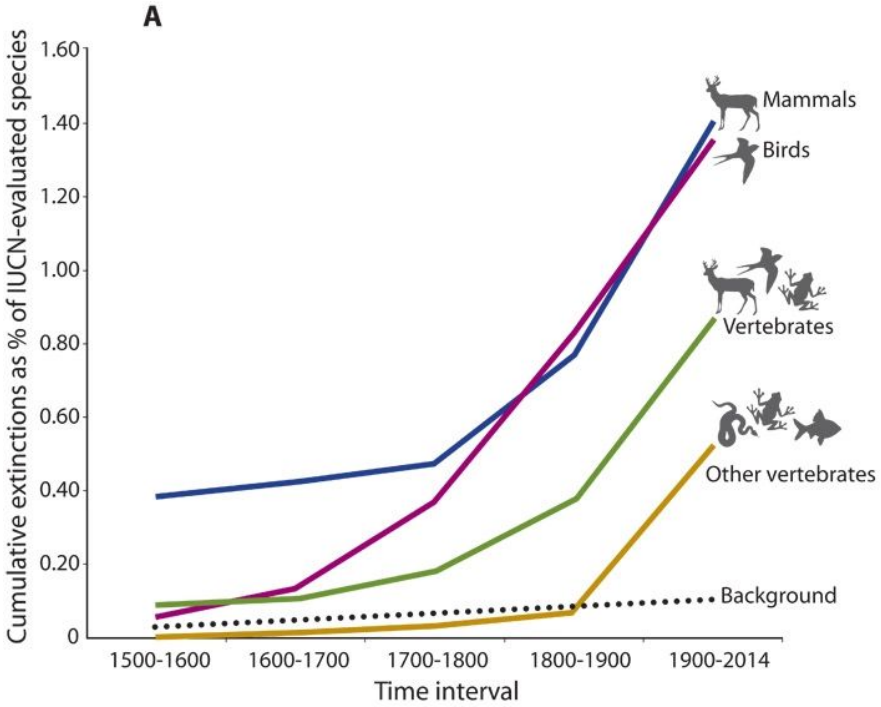
Standardization

GenErode pipeline objectives

Species get closer to extinction because of the small population size

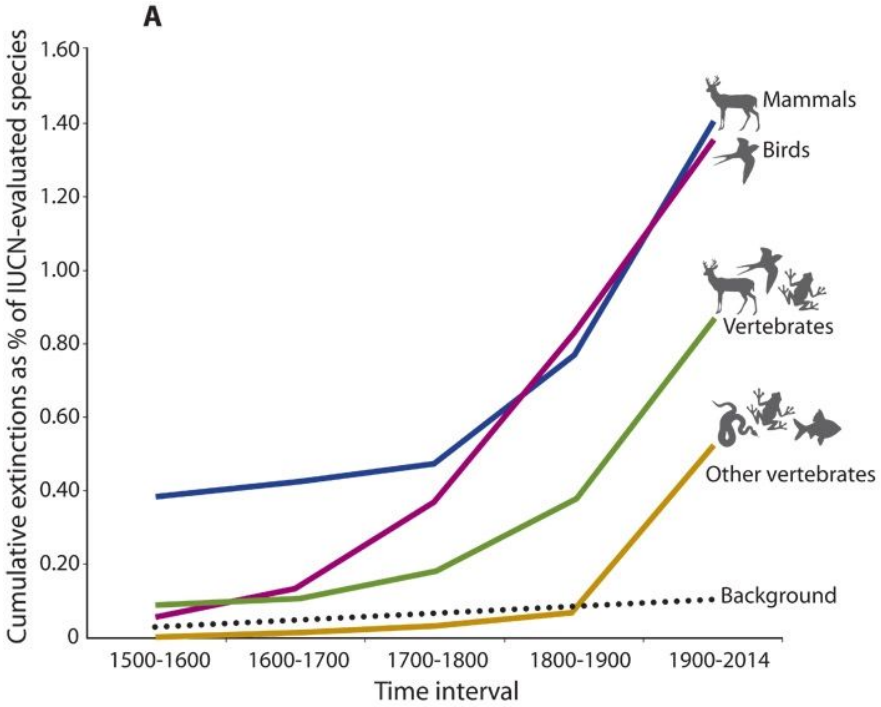


Species get closer to extinction because of the small population size



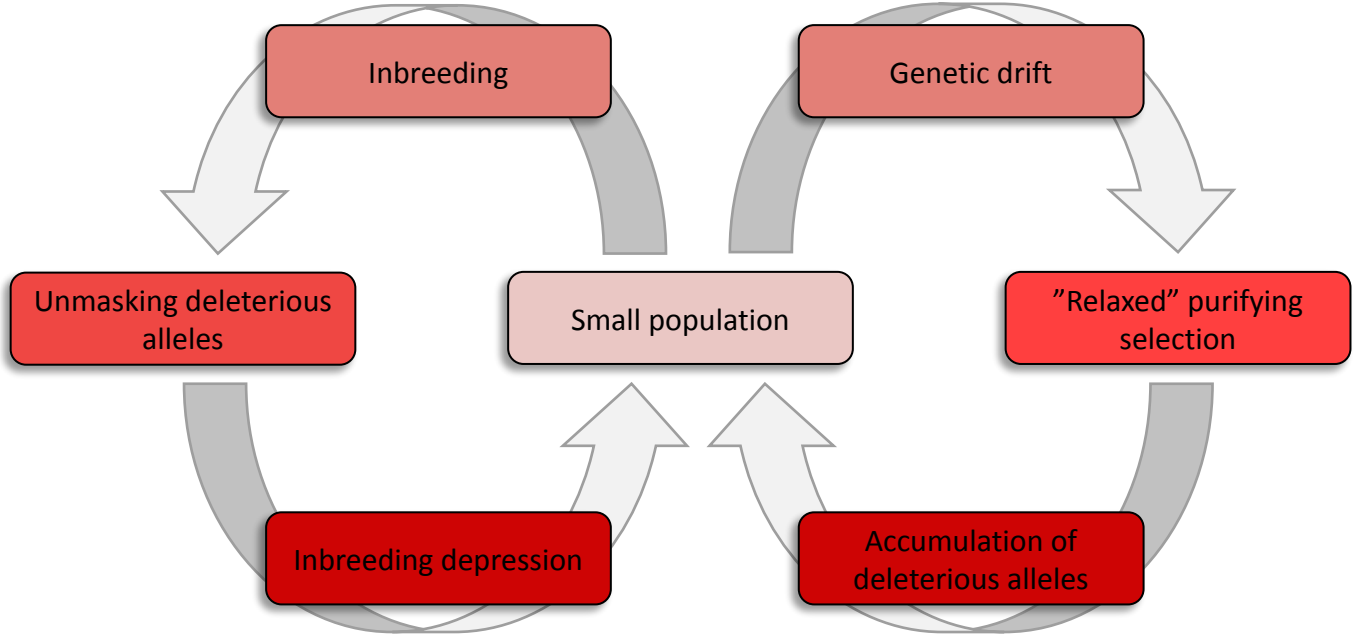
- Stochastic demographic and environmental events

Species get closer to extinction because of the small population size

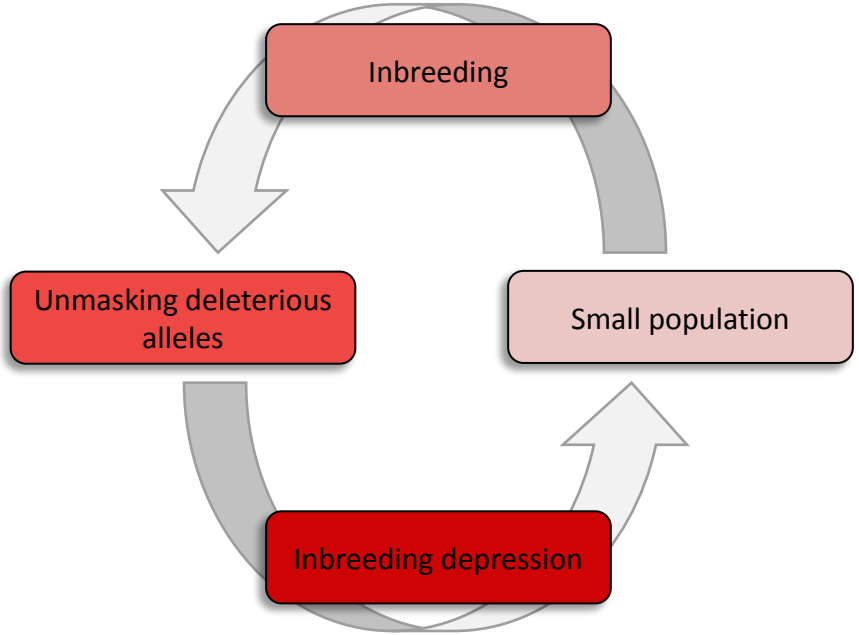


- Stochastic demographic and environmental events
- Genetic processes

Extinction vortex



Inbreeding depression



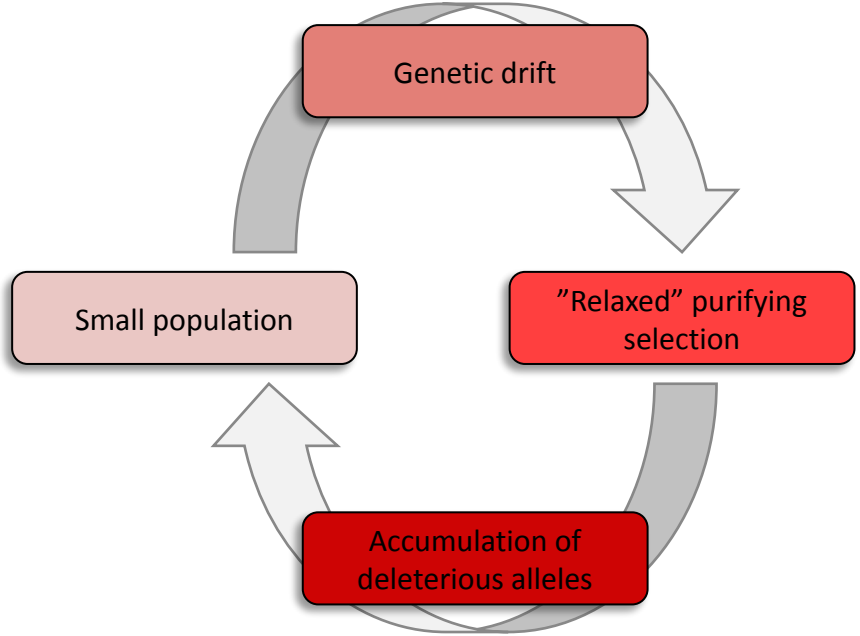
Genomic evidences:

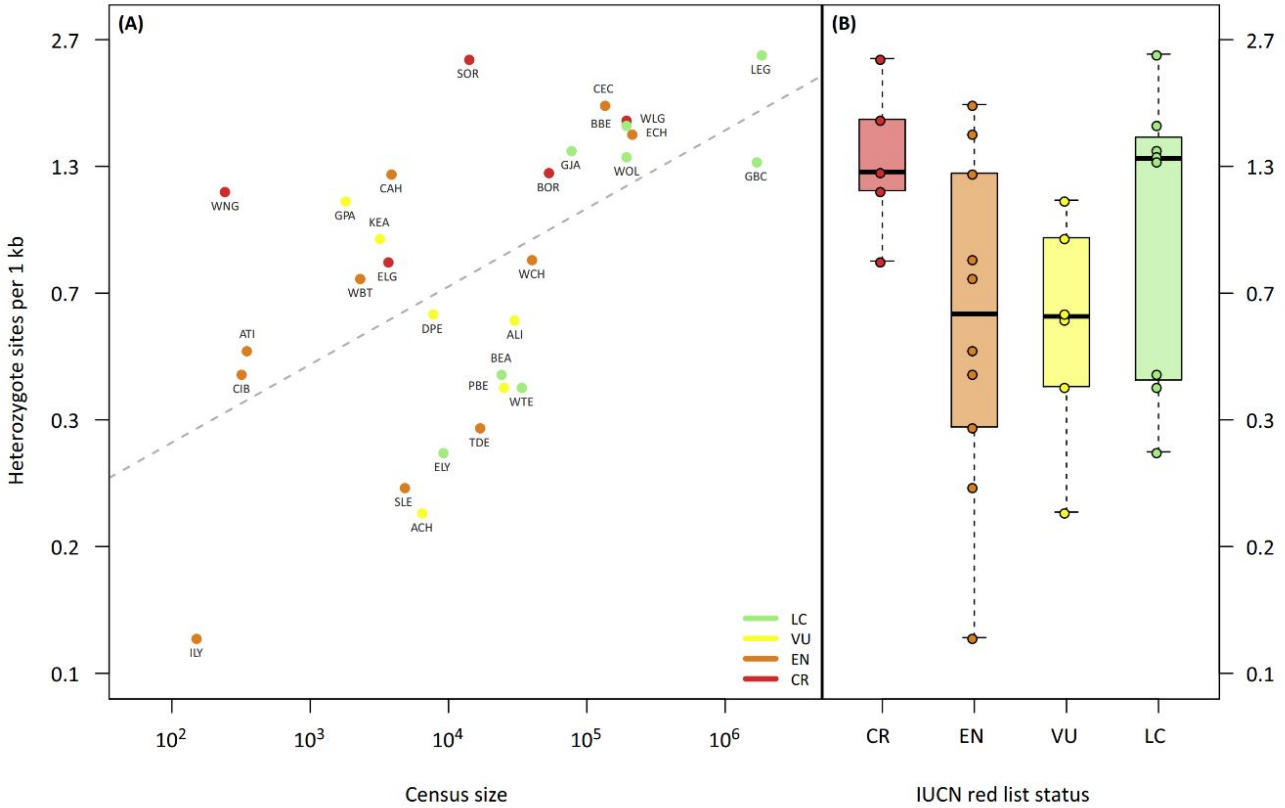
- Reduced heterozygosity
- Increased inbreeding

Mutational meltdown

Genomic evidences:

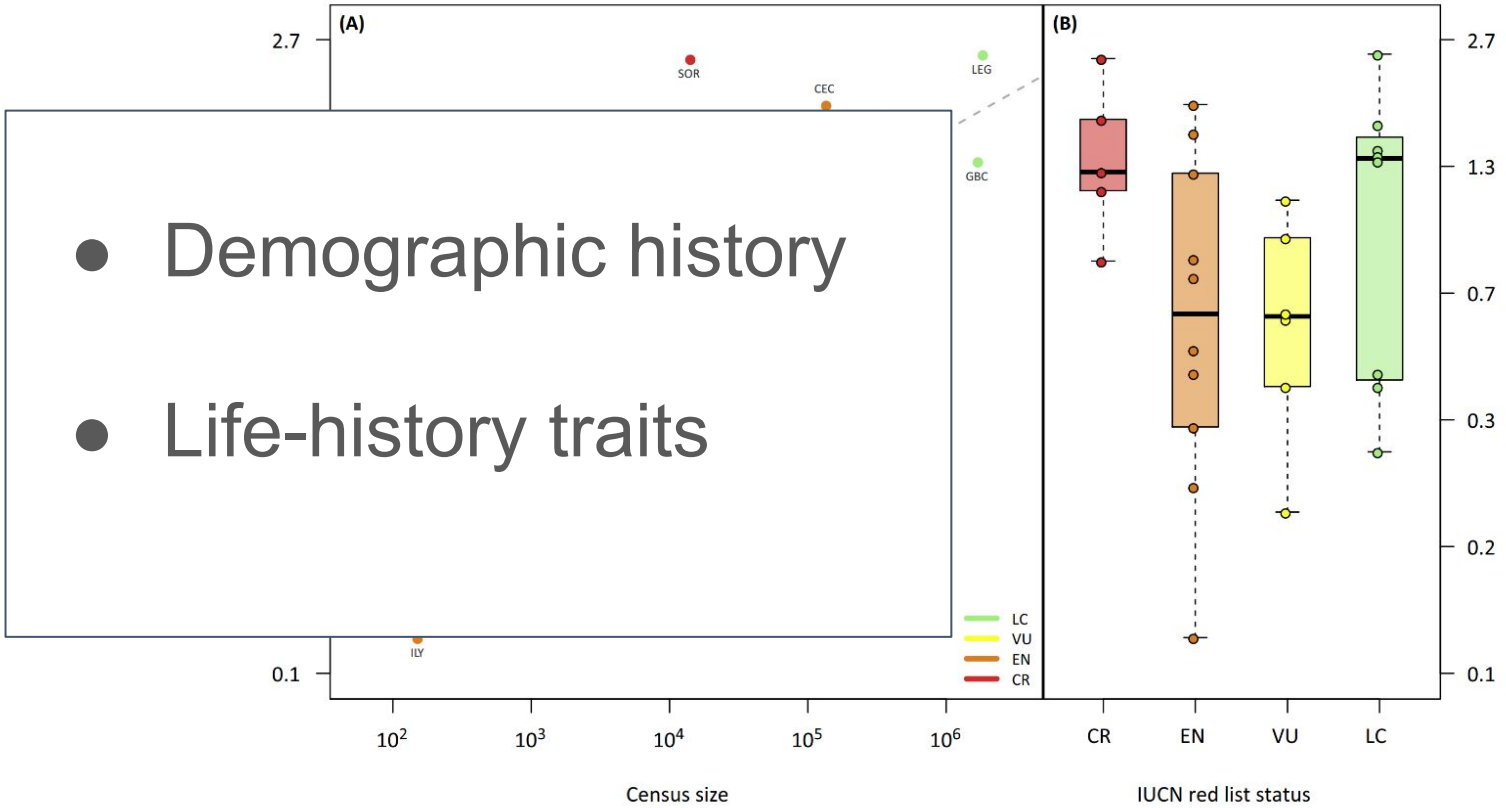
- Increased number of deleterious alleles





Diez-del-Molino et al. 2018

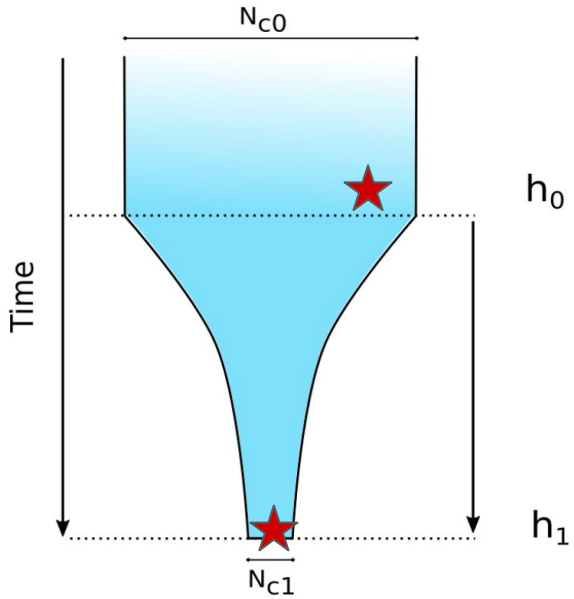
- Demographic history
- Life-history traits



Diez-del-Molino et al. 2018

Interspecific comparisons of contemporary genomic erosion indices are poor proxies for conservation status of wild organisms

Temporal data

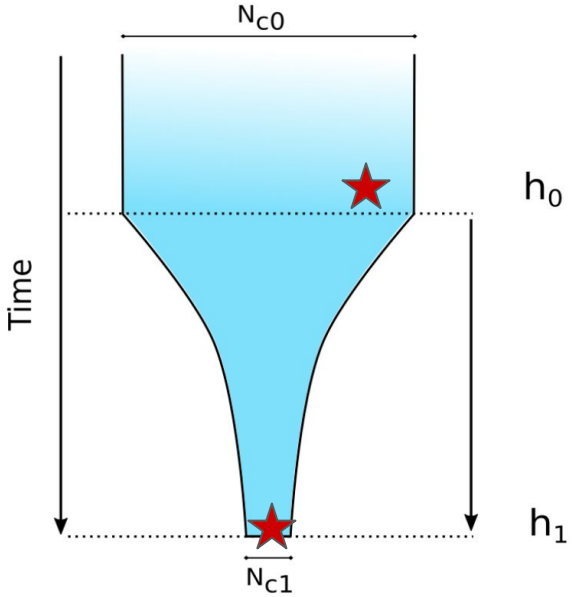


Dicerorhinus sumatrensis

CR (30 individuals)

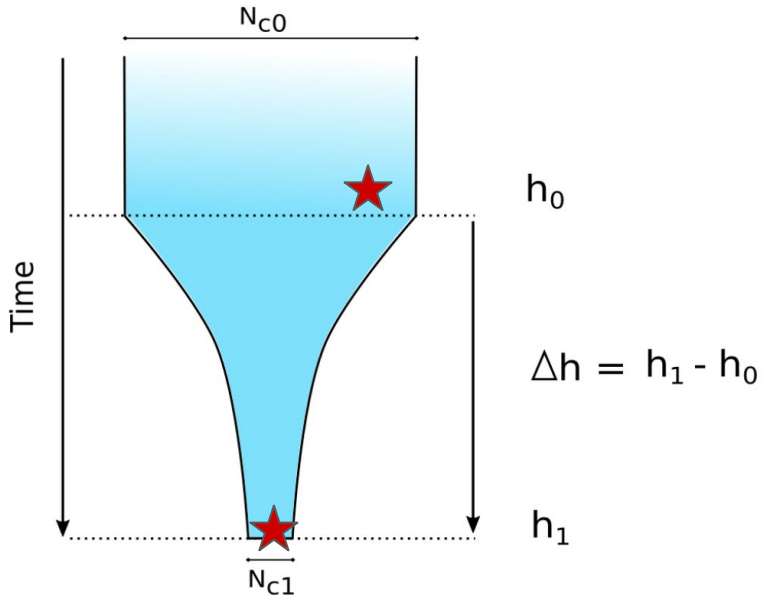
Diez-del-Molino et al. 2018

Temporal data



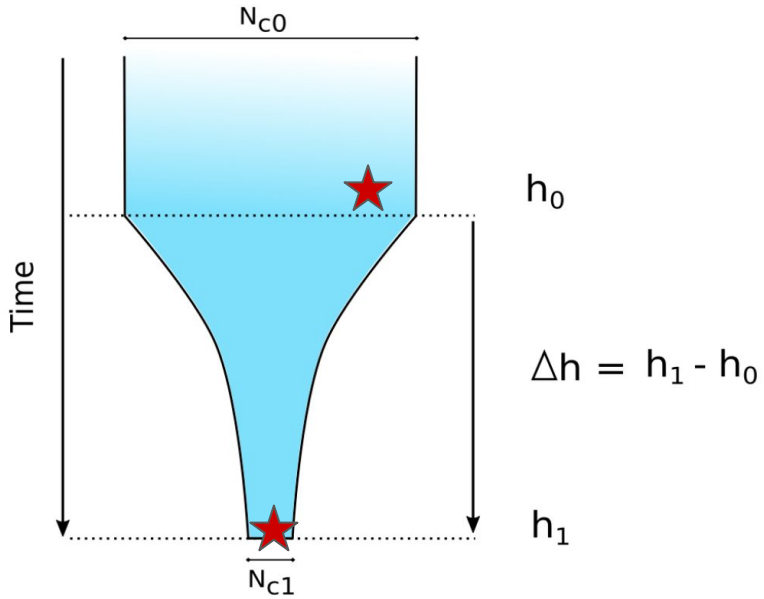
Diez-del-Molino et al. 2018

Temporal data



Diez-del-Molino et al. 2018

Temporal data

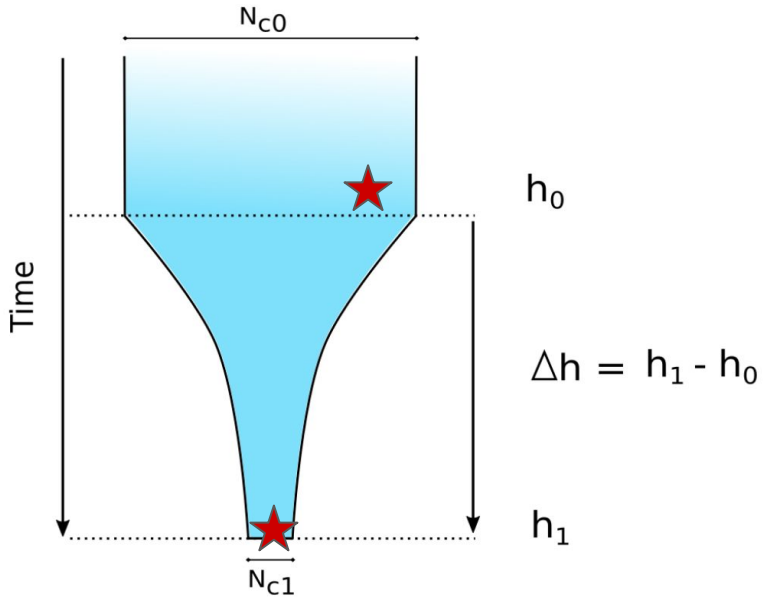


Temporally sampled data useful for:

- Inform of the rate of change in genomic erosion indices.
- Provide comparable trends for modern-day endangered populations.

Diez-del-Molino et al. 2018

Temporal data



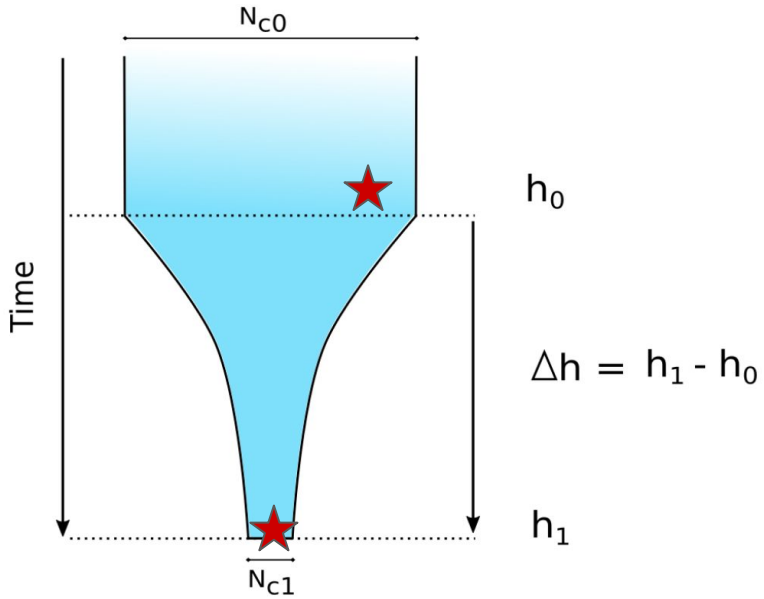
Temporally sampled data useful for:

- Inform of the rate of change in genomic erosion indices.
- Provide comparable trends for modern-day endangered populations.

IUCN RED LIST

Diez-del-Molino et al. 2018

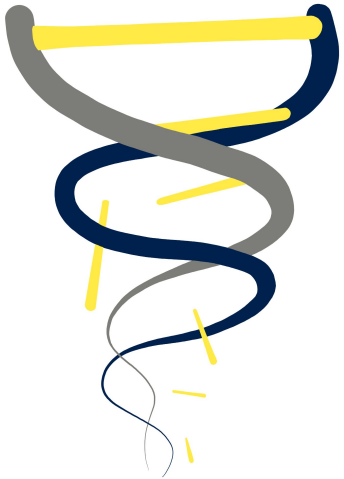
Temporal data



GenErode pipeline objectives

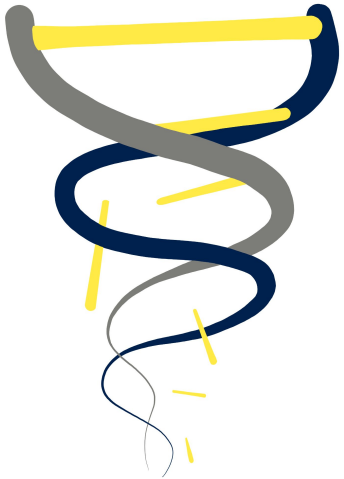
Diez-del-Molino et al. 2018

GenErode



A bioinformatics pipeline to simultaneously analyze genome re-sequencing data from **ancient/historical and modern samples**, with the aim to look for patterns of **genome erosion.**

GenErode

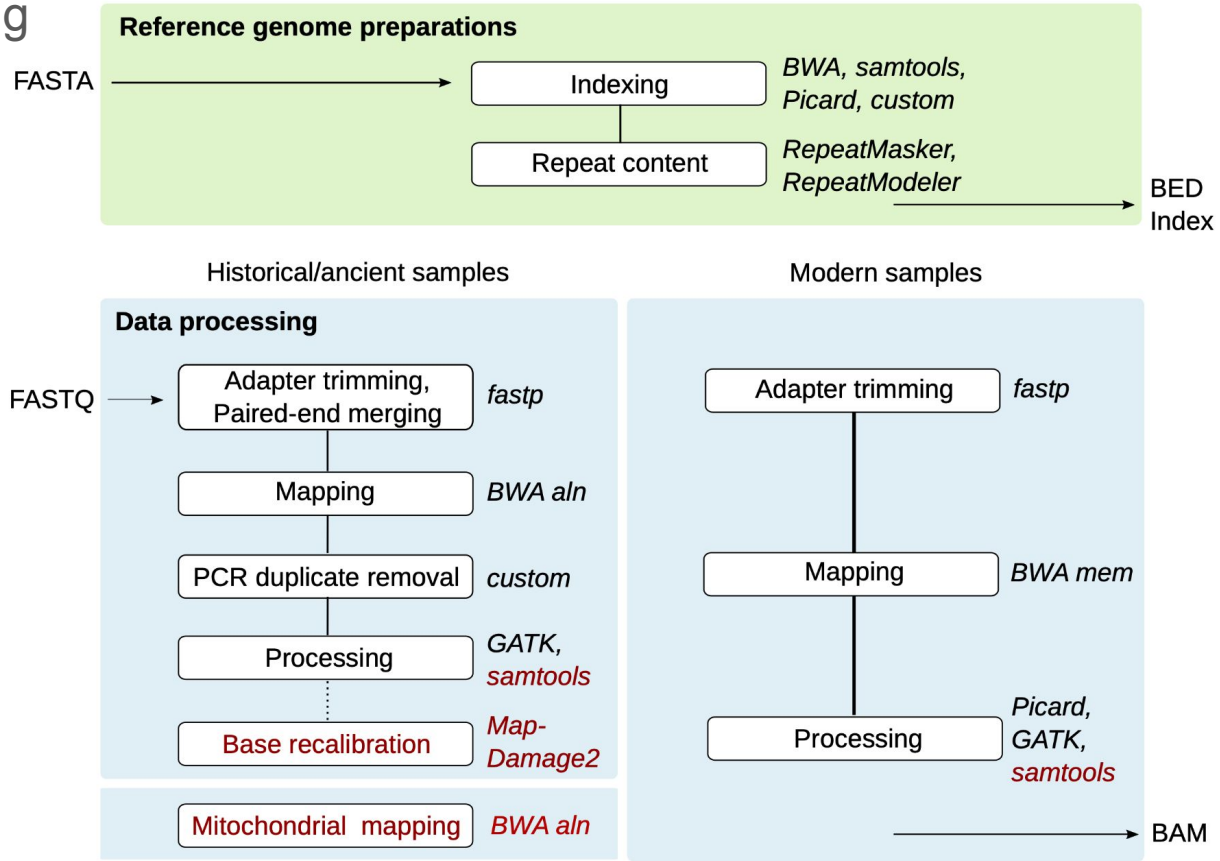


- Snakemake
- No programming knowledge required
- Conda and Singularity containers
- Designed for HPC clusters
- Modular and highly flexible
- Well documented (github wiki + article)

GenErode

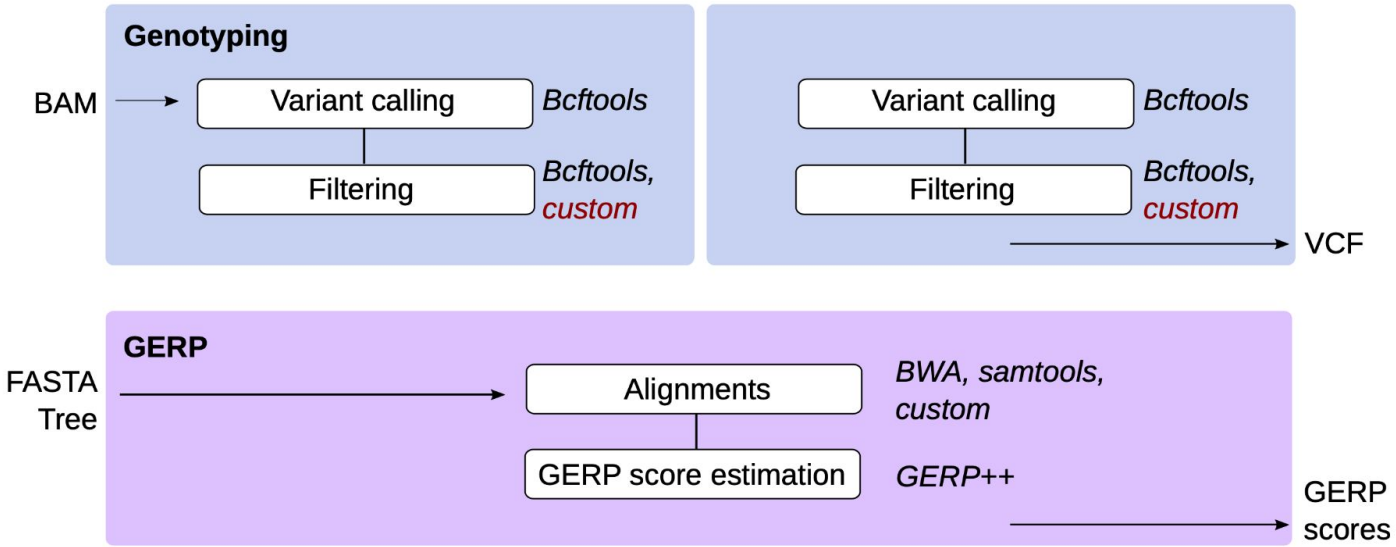
- Reproducible analyses
- Standardization of indices
- Comparable indices historical/modern samples
- Accessible to non-experts

Data processing



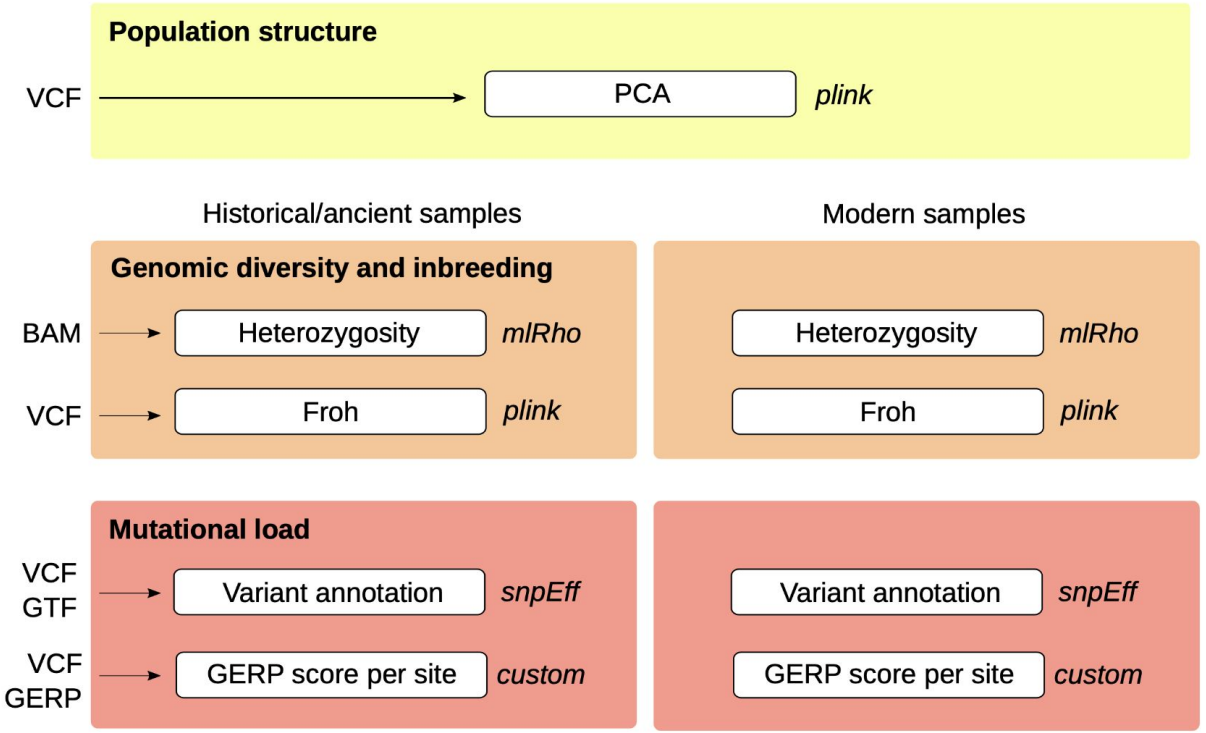
Kutschera et al. 2022

Data processing



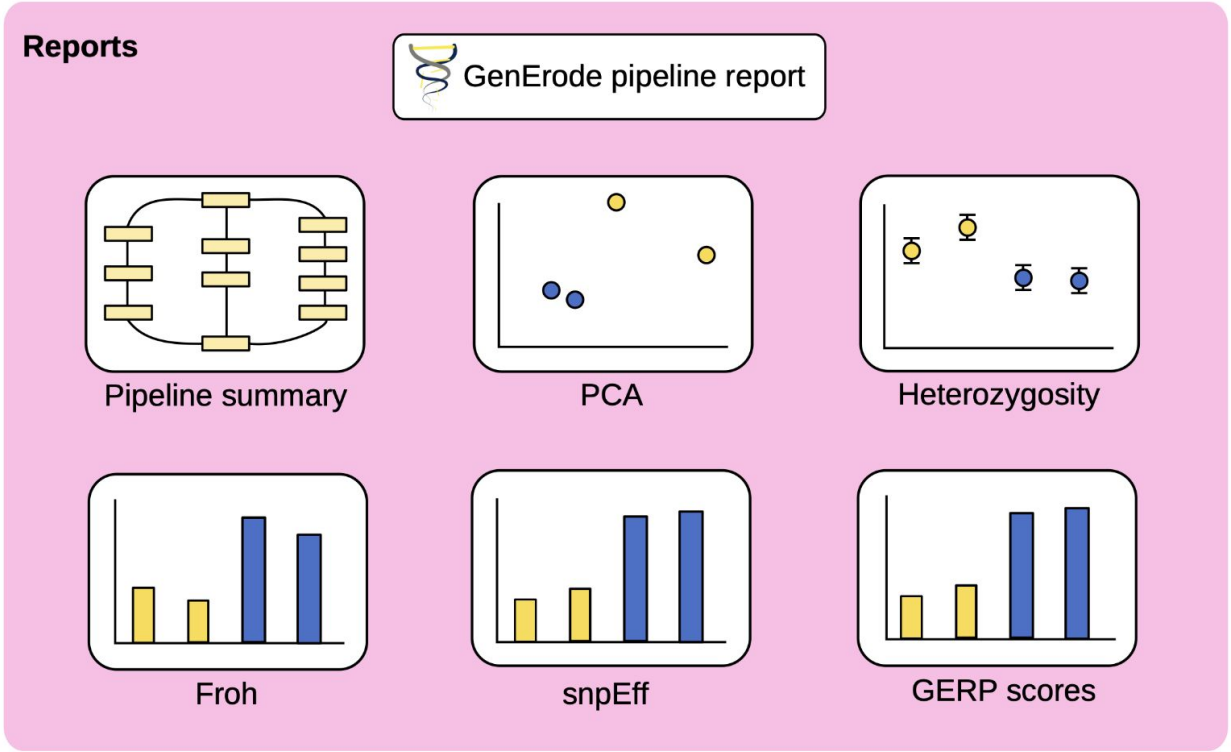
Kutschera et al. 2022

Data analysis



Kutschera et al. 2022

Pipeline outputs



Kutschera et al. 2022

nature COMMUNICATIONS

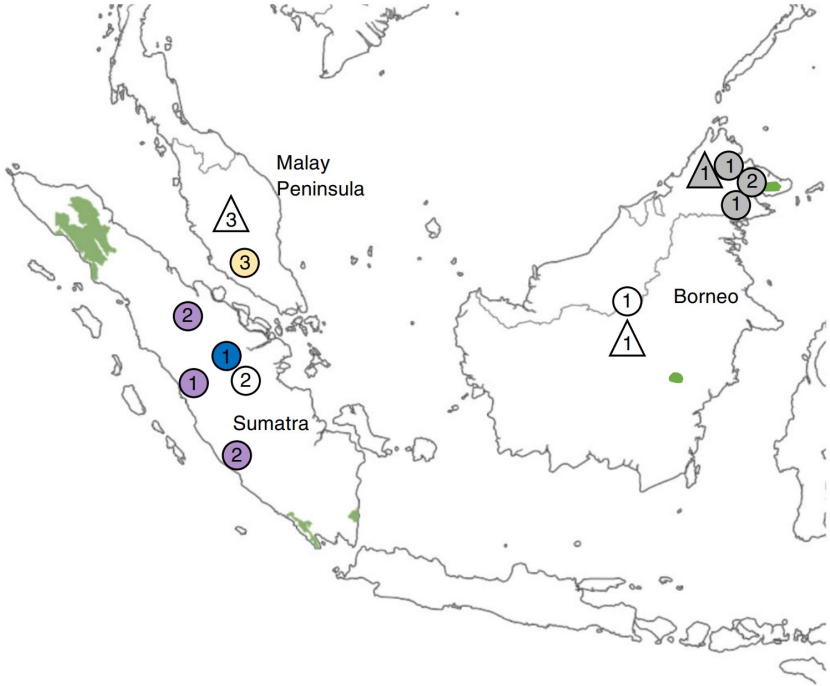
ARTICLE

https://doi.org/10.1038/s41467-021-22386-8 OPEN

Check for updates

Genomic insights into the conservation status of the world's last remaining Sumatran rhinoceros populations

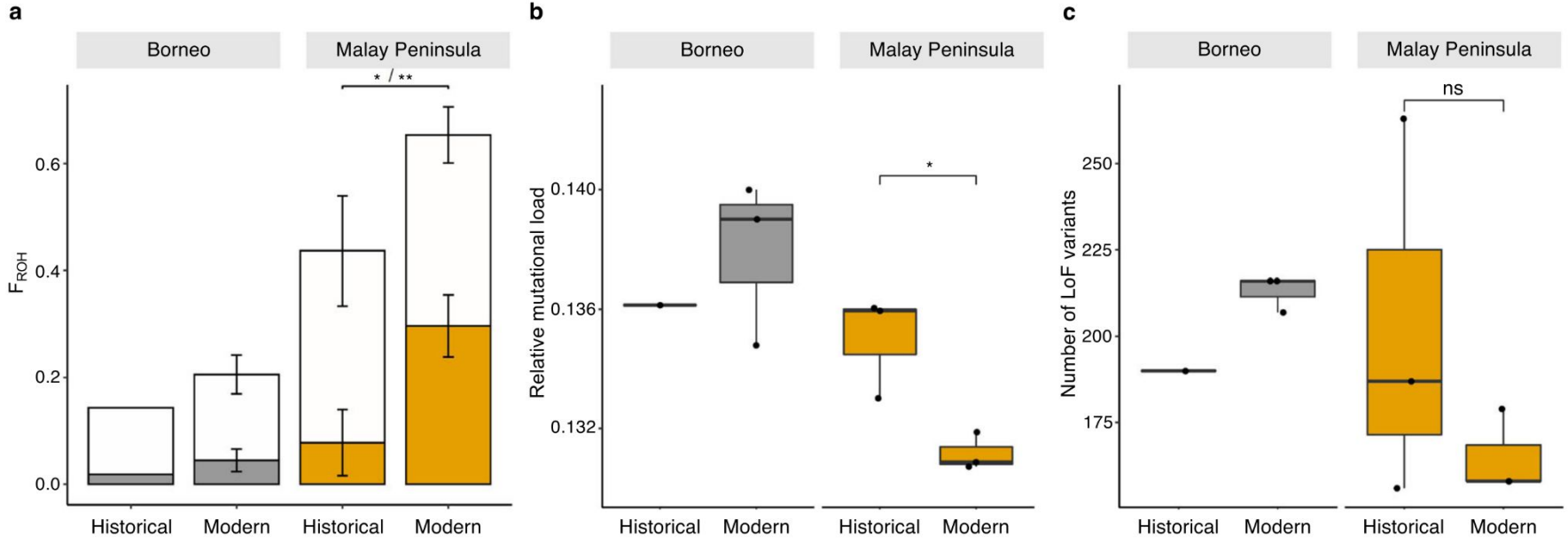
Johanna von Seth[1,2,3,21✉], Nicolas Dussex[1,2,3,21✉], David Díez-del-Molino[1,2,3], Tom van der Valk[1,2,4], Verena E. Kutschera[5], Marcin Kierczak[6], Cynthia C. Steiner[7], Shanlin Liu[8], M. Thomas P. Gilbert[8,9], Mikkel-Holger S. Sinding[8,10], Stefan Prost[11,12], Katerina Guschanski[4,13], Senthilvel K. S. S. Nathan[14], Selina Brace[15], Yvonne L. Chan[1,2], Christopher W. Wheat[3], Pontus Skoglund[16], Oliver A. Ryder[7], Benoit Goossens[14,17,18,19], Anders Götherström[1,20] & Love Dalén[1,2,3✉]

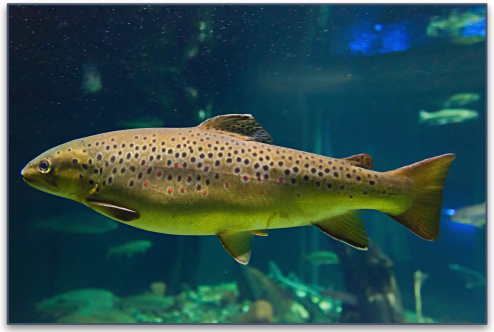


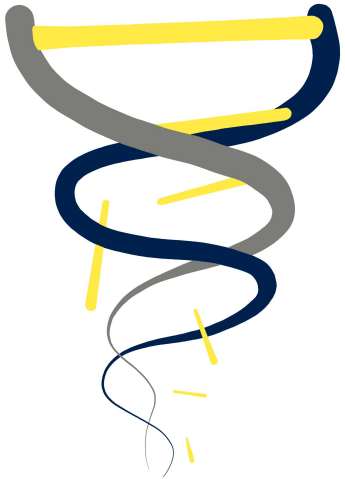
Dicerorhinus sumatrensis

CR (30 individuals)

Von Seth et al. 2021


a
Borneo
Malay Peninsula
* / **
F_{ROH}
0.0
0.2
0.4
0.6
Historical
Modern
Historical
Modern
b
Borneo
Malay Peninsula
*
Relative mutational load
0.132
0.136
0.140
Historical
Modern
Historical
Modern
c
Borneo
Malay Peninsula
ns
Number of LoF variants
175
200
225
250
Historical
Modern
Historical
Modern





github.com/NBISweden/GenErode

Verena Kutschera
Marcin Kierczak
Payam Khoonsari
Björn Nystedt

Johanna von Seth
Tom van der Valk
Nicolas Dussex
Edana Lord
Marianne Dehasque
Dave Stanton
Love Dalén

& more!

NBIS

Stockholms universitet

Naturhistoriska riksmuseet

CPG Centre for Palaeogenetics

Cold Spring Harbor Laboratory

bioRxiv

THE PREPRINT SERVER FOR BIOLOGY

GenErode: a bioinformatics pipeline to investigate genome erosion in endangered and extinct species

Verena E. Kutschera, Marcin Kierczak, Tom van der Valk, Johanna von Seth, Nicolas Dussex, Edana Lord, Marianne Dehasque, David W. G. Stanton, Payam Emami Khoonsari, Björn Nystedt, Love Dalén, David Díez-del-Molino