

# Computing Genomic Offset Statistics with the R Package LEA

Olivier François

*Université Grenoble-Alpes*

SSMPG 2025, Aussois, Savoie

# Genetic offset (GO)

- **Genomic offset statistics** evaluate maladaptation of populations to rapid habitat alteration based on **association of genotypes with environmental factors**.
- Gos predict whether **sudden changes** in a species ecological niche could be sustained by non-migrating natural populations or not.

Capblancq et al [Genomic prediction of \(mal\)adaptation across current and future climatic landscapes](#) 2020

# Ecological predictors, $\mathbf{x} = (x_1, \dots, x_d)$

- **Bioclimatic predictors**: temperature, precipitation or solar radiation
- **Biotic predictors**: local abundance of species sharing ecological interactions with the studied organisms
- **Altered environment**,  $\mathbf{x}^\star$ , could result from **translocation** in geographic space or represent **future conditions** at unchanged geographic location
- Predictors are **unitless**, and expressed as deviations from the sample mean (e.g., centered)

# GO principles (2 stages)

1. Adjust a genotype-environment association (GEA) model predicting allelic frequencies from environmental variables at many genetic loci.
2. Evaluate dissimilarity between allelic frequencies predicted under current conditions and altered conditions.

Capblancq et al [Genomic prediction of \(mal\)adaptation across current and future climatic landscapes](#) 2020

# Background (main approaches)

GO	Authors	Predicted allelic frequencies	distance
Gradient Forest	<a href="#">Fitzpatrick &amp; Keller 2015</a>	Random forests	Quadratic distance between importance curves
Rona	<a href="#">Rellstab et al. 2016</a>	Linear regression models	Absolute difference
RDA GO	<a href="#">Capblancq et al. 2021</a>	Multivariate linear analysis	Euclidean distance in projection space
Geometric GO	<a href="#">Gain et al. 2023</a>	Latent factor regression models	Quadratic distance

# Goals of this tutorial

Genomic offset in the R package LEA

1. Modeling assumption
2. Basic usage
3. Advanced usage

# Modeling assumption: relative fitness

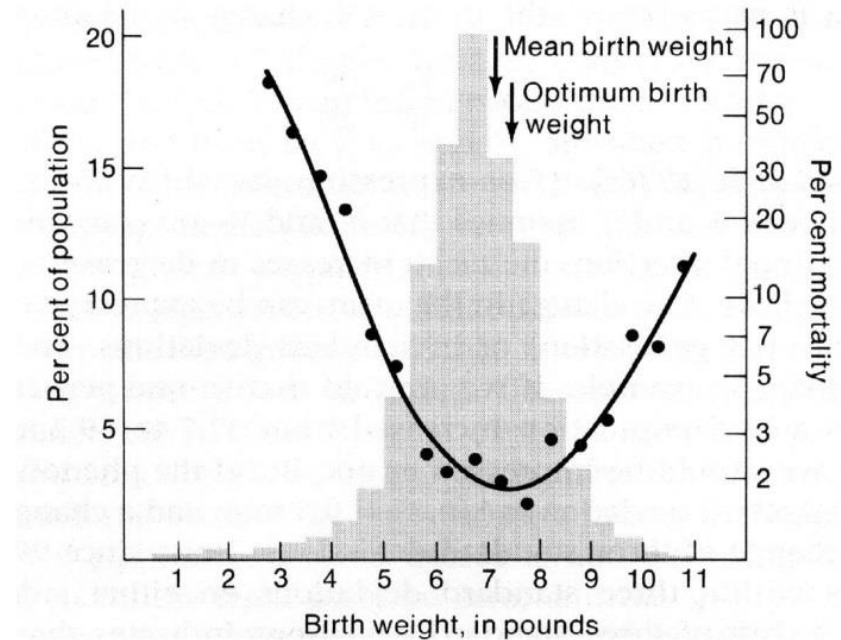
- Gaussian stabilizing selection = A type of natural selection that favors the average local phenotype

$z$  = phenotype,  $z_{\text{opt}}$  = average local value

$$\omega(z) = \exp\left(-\frac{(z - z_{\text{opt}})^2}{2V_s}\right)$$

- Fitness decreases exponentially fast as the phenotype shifts away from the optimum

## STABILIZING SELECTION ON HUMAN BIRTH WEIGHT



FROM: Cavalli-Sforza & Bodmer 1971

# Model assumptions

- **Environmental effects:** At any genetic locus, the **allelic frequency**,  $y$ , is linked to the ecological predictors,  $\mathbf{x}$ , and some unobserved covariates,  $\mathbf{u}$ , through a GEA model

$$y = \mathbf{x}\mathbf{b}^T + \mathbf{u}\mathbf{v}^T + \epsilon .$$

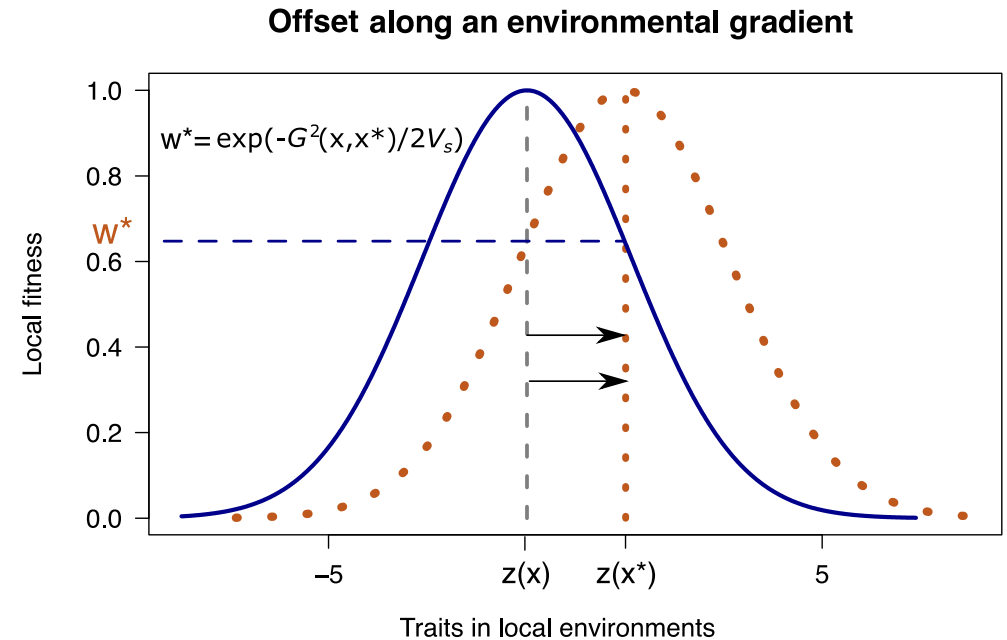
- **Fisher's infinitesimal model** : Traits are determined by a polygenic model in which allelic effects are « infinitesimally » small



# Evolutionary mismatch ( $\omega^\star$ )

- **Evolutionary mismatch** occurs when a trait at equilibrium in environment  $\mathbf{x}$  is placed in an altered environment  $\mathbf{x}^\star$

$$\omega^\star = \omega(\mathbf{x}, \mathbf{x}^\star)$$



# Theory

- The genomic offset in LEA is defined as

$$G^2(\mathbf{x}, \mathbf{x}^\star) = (\mathbf{x} - \mathbf{x}^\star) \mathbf{C}_b (\mathbf{x} - \mathbf{x}^\star)^T,$$

Covariance of environmental effect sizes

- Relationship between altered fitness and GO

$$-\log \omega(\mathbf{x}, \mathbf{x}^\star) \approx \alpha_s G^2(\mathbf{x}, \mathbf{x}^\star) + \alpha_0$$

# Basic usage: Starting a session

1. Start an R session with the LEA package

```
library(LEA)
```

2. Load the example dataset

```
# loading the simulated data in R  
data("offset_example")  
  
# Y contains 100 genotypes for 200 individuals  
Y <- offset_example$geno  
  
# X contains 4 environmental variables for 200 individuals  
X <- offset_example$env
```

# Basic usage: GO statistics

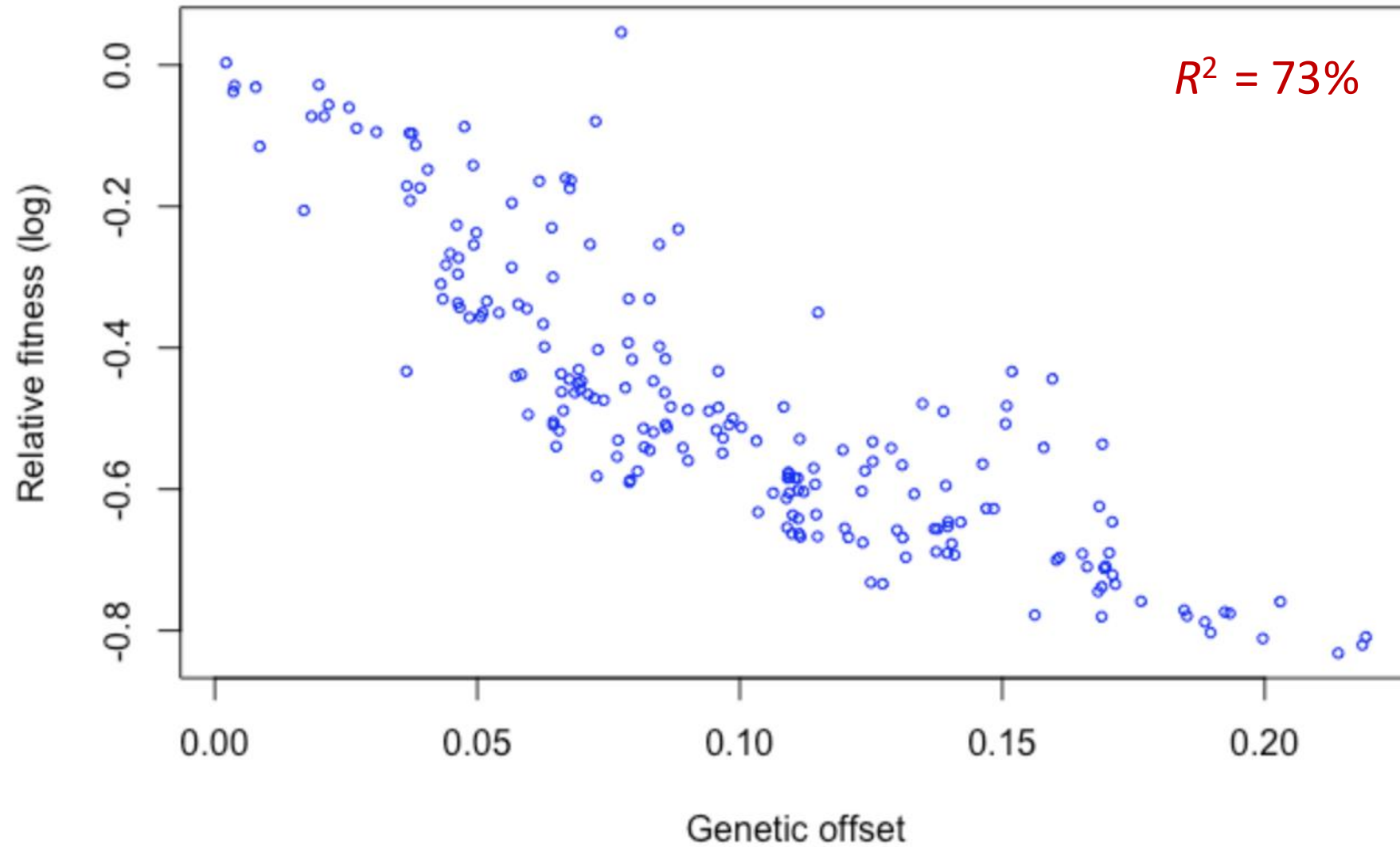
1. Load future conditions (altered environmental variables)

```
## modified environment  
X_pred <- offset_example$env.pred
```

2. Compute LEA's genomic offset

```
## computes genetic offset using all loci  
g_offset <- genetic.offset(input = Y,  
                           env = X,  
                           pred.env = X_pred,  
                           K = 3)
```

# Results



# Advanced usage: GEA study

1. Fit an LFMM with  $K = 3$  latent factors

```
# use lfmm2 to estimate the latent factors from the data  
mod_lfmm2 <- lfmm2(input = Y,  
                   env = X,  
                   K = 3)
```

2. Compute  $p$ -values (one for each locus)

```
# GEA analysis computing locus-specific p-values, P  
pv_lfmm2 = lfmm2.test(object = mod_lfmm2,  
                      input = Y,  
                      env = X,  
                      full = TRUE)$pvalues
```

# Advanced usage: candidate loci

1. Decide which loci to include in GO

```
# FDR control: computing qvalues  
qv_lfmm2 <- qvalue::qvalue(pv_lfmm2, fdr.level = 0.2)  
  
# the most interesting targets  
candidates <- which(qv_lfmm2$significant)
```

2. Compute LEA's genomic offset

```
g_offset_candidate <- genetic.offset(input = Y,  
                                     env = X,  
                                     pred.env = X_pred,  
                                     candidate.loci = candidates,  
                                     K = 3)
```

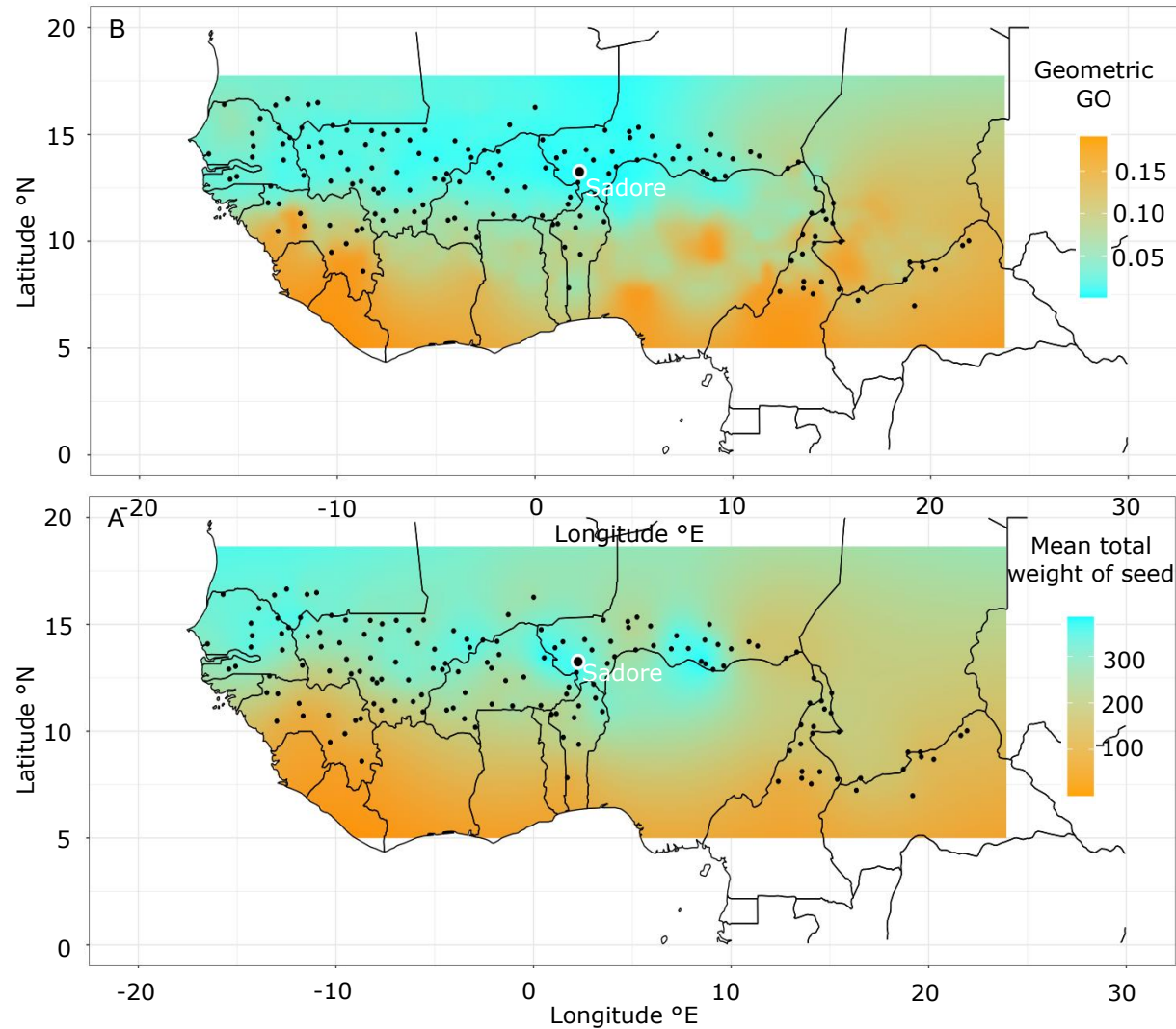
# Advanced usage: maps

Required packages

```
# Required packages  
# Loading worldclim/cimp6 bioclimatic data  
library(terra)  
library(geodata)  
  
# displaying images and maps  
library(fields)  
library(maps)
```



# Pearl millet common garden experiment

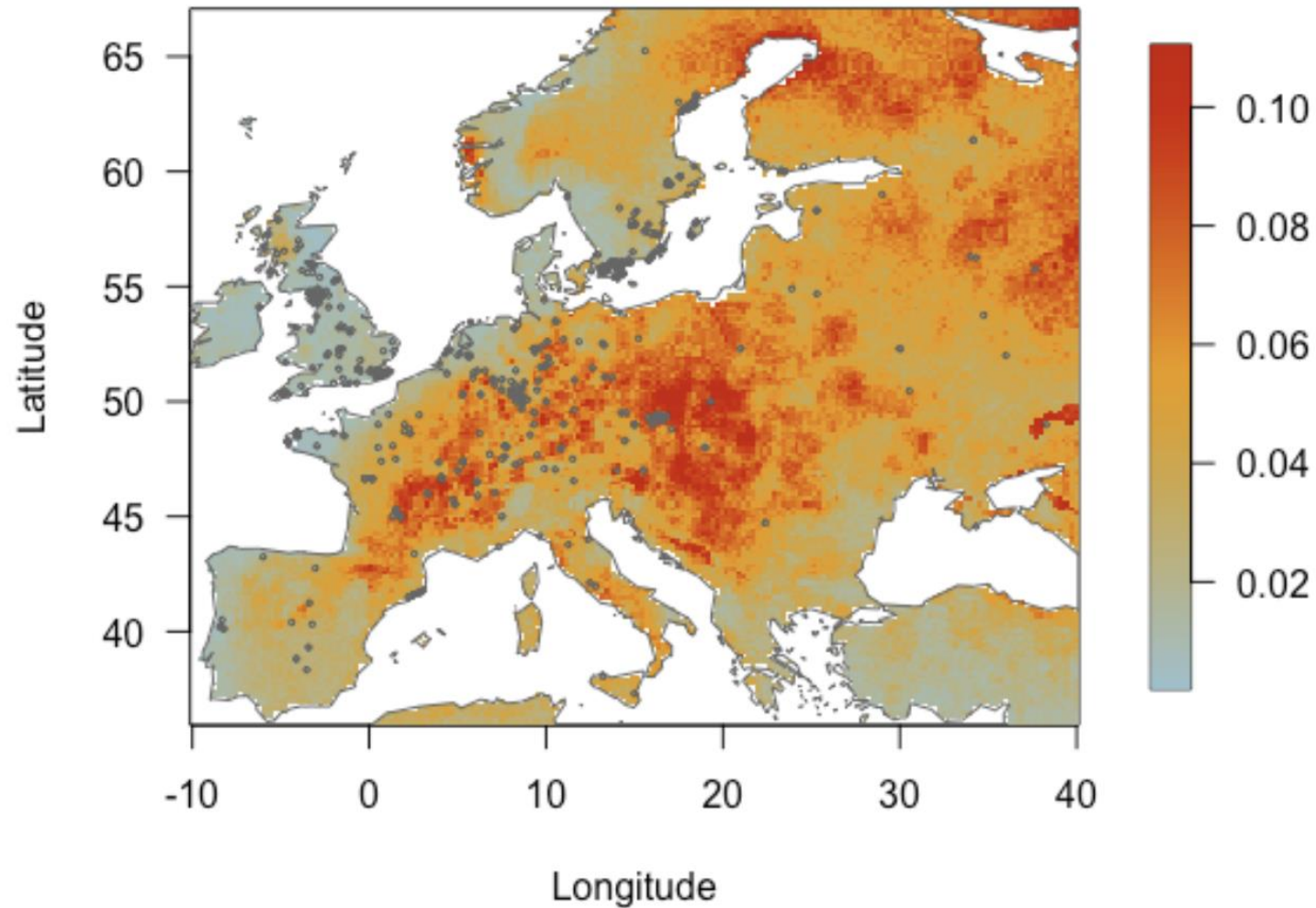


154 landraces of pearl millet grown in Niger

Whole Genome Pool-seq data (100 individuals)

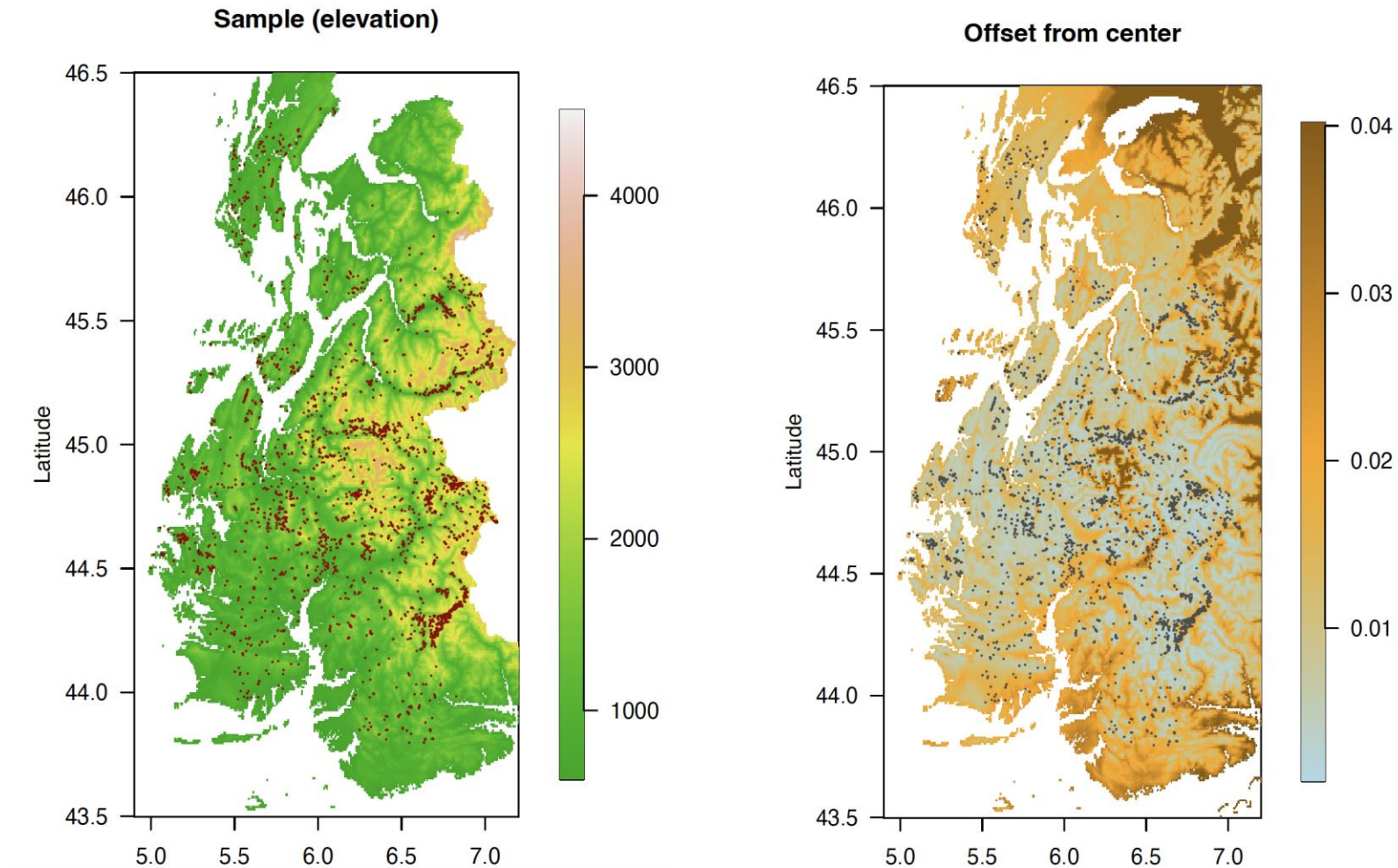
27 bioclimatic variables

## A. thaliana example: 1,096 European accessions



Climate change scenario: SSP2-4.5 from IPCC AR6. Climate model ACCESS-ESM1-5 (2041-2060)

# Alpine plant communities (presence/absence data)



Compositional offset for translocation from the center of the climatic niche (ecological mismatch)

# Resources

- For basic GOs and maps, there are tutorials in the SSMPG 2025 GitHub repository and in LEA's vignette

<https://github.com/bcm-uga/SSMPG2025/blob/main/Tutorials/>

- For more advanced questions on GEAs, GOs and maps, ask us during the practical sessions