



HAL
open science

Predicting species invasiveness with genomic data: is Genomic Offset related to establishment probability?

Louise Camus, Simon Boitard, Mathieu Gautier

► To cite this version:

Louise Camus, Simon Boitard, Mathieu Gautier. Predicting species invasiveness with genomic data: is Genomic Offset related to establishment probability?. Conservation Genomics Paris, Jun 2024, Paris, France. hal-04659080

HAL Id: hal-04659080

<https://hal.inrae.fr/hal-04659080v1>

Submitted on 22 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Predicting species invasiveness with genomic data: is Genomic Offset related to establishment probability ?

Louise Camus, Simon Boitard, Mathieu Gautier

Conservation Genomics, Paris

June 5th 2024

Invasive species pose a major challenge

Negative impacts on :

- Biodiversity (Mollot et al., 2017)
- Economy (Haubrock et al., 2021)
- Human health (Mazza et al., 2014)
- Food security (Paini et al., 2016)

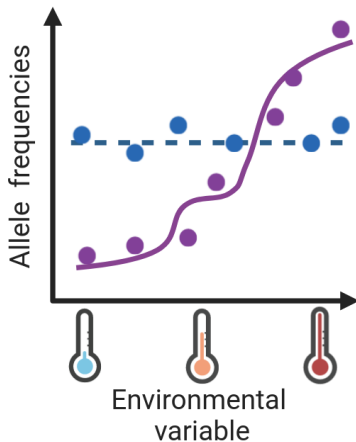


Accelerating effect of globalization and climate change

⇒ Need to improve the understanding and prediction of invasions to enhance biological control efficiency (Hulme, 2017, Reaser et al., 2020)

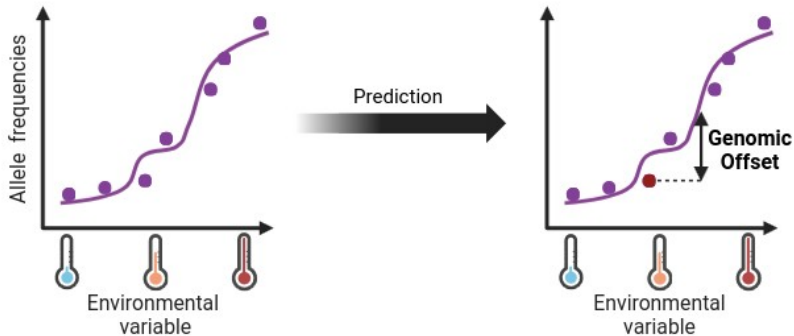
How can we use genomics in this context ?

Genotype Environment Association



- Identification of markers/environmental variables with specific roles in adaptation

Population-level genomic prediction ?



Genomic Offset = a measure of the level of maladaptation

- Temporal (future environmental conditions)
- Spatial (new geographic areas)

Relationship between GO and fitness already studied

Common garden : Rhoné *et al.* (2020), Fitzpatrick *et al.* (2021)

Simulations : Laruson *et al.* (2021), Gain *et al.* (2023)

But not applied to biological invasions

Simulations

- Relationship between GO and the establishment probability of invasive populations?

Methods

Geometric GO

- Linear (regression coefficients)
- Correction for neutral structure :
 - Latent factors (*LFMM*)
 - or**
 - Allelic frequencies covariance (*Baypass*)

Gradient Forest (optimized)

- Non linear (Random Forest)
- Allelic frequencies corrected with LFMM

Simulations design

2 steps for simulations (SLiM v4)

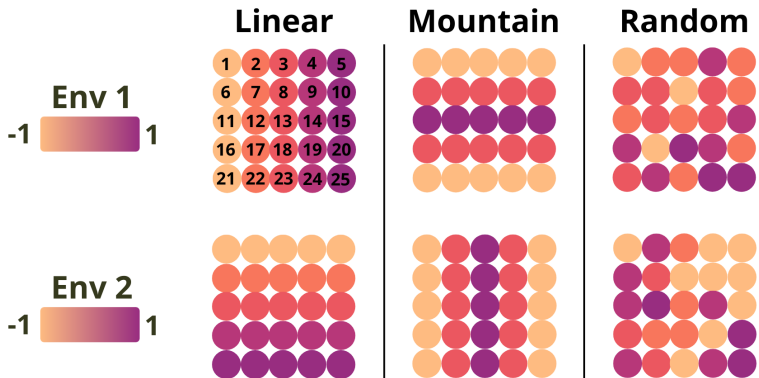
1. Native area

- Population grid (5x5) , stepping stone
- 1000 individuals/population
- High (0.05) or low (0.005) migration rate
- 3000 simulated generations



Simulations design : native area

- 2 environmental variables, related to adaptation *via* QTLs (Quantitative Trait Loci)
 - Σ effect sizes = phenotype \leftrightarrow environment = fitness
- 3 different environment type (replicated 10 times each) :



Simulations design : invaded area

2. Invaded area

- Panmictic population characterized by 2 the env. variables
↳ **9 possible environments to invade** : $(-1,0,1) \times (-1,0,1)$

Simulations design : invaded area

2. Invaded area

- Panmictic population characterized by 2 the env. variables
↳ **9 possible environments to invade** : $(-1,0,1) \times (-1,0,1)$
- 10 or 100 invading individuals
↳ **3 possible source populations** : $-1/-1, 0/0, 1/1$

Simulations design : invaded area

2. Invaded area

- Panmictic population characterized by 2 the env. variables
↳ **9 possible environments to invade** : $(-1,0,1) \times (-1,0,1)$
- 10 or 100 invading individuals
↳ **3 possible source populations** : $-1/-1, 0/0, 1/1$
- Non constrained population size ("Non-WF") :
 - Overlapping generations
 - Death \neq Reproduction

Simulations design : invaded area

2. Invaded area

- Panmictic population characterized by 2 the env. variables
↳ **9 possible environments to invade** : $(-1,0,1) \times (-1,0,1)$
- 10 or 100 invading individuals
↳ **3 possible source populations** : $-1/-1, 0/0, 1/1$
- Non constrained population size ("Non-WF") :
 - Overlapping generations
 - Death \neq Reproduction
- Simulation stops when the population :
 - goes extinct
 - **goes established**
 - ↳ reaches 50 000 individuals
 - ↳ survives for 100 generations

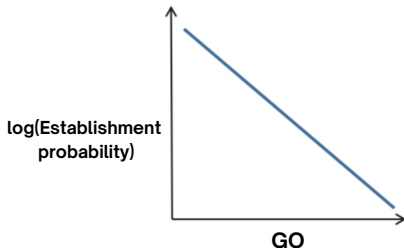
Analyses

GO can be computed :

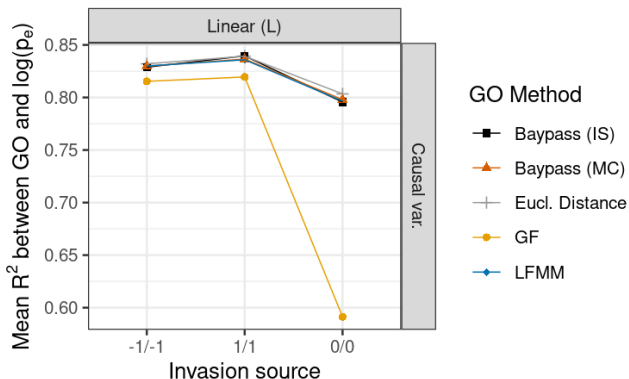
Variables



Each invasion is repeated 250 times to compute establishment probability

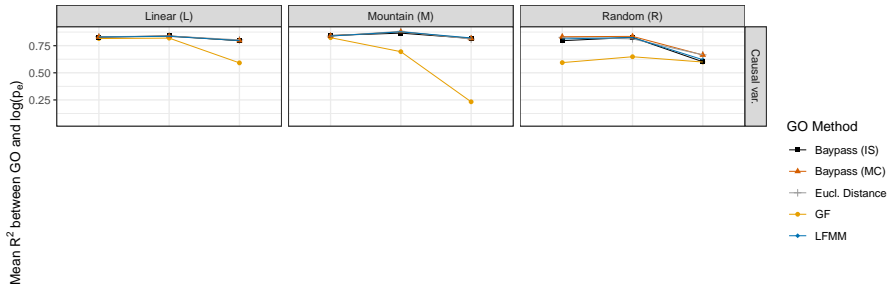


Results in an "ideal" case

Mean R^2 between GO and establishment prob. (10 ind., mig 0.005)

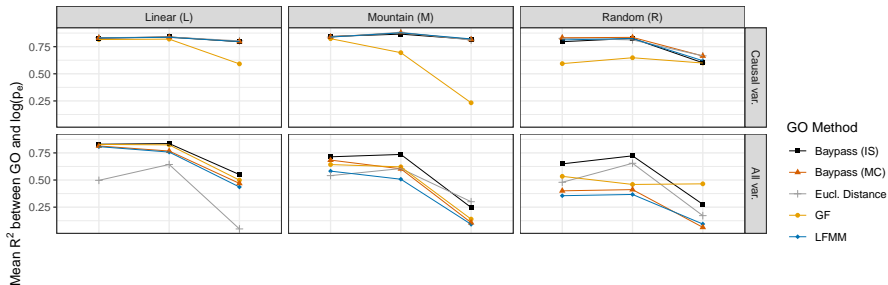
Good performances with only causal variables

Mean R^2 between GO and establishment prob. (10 ind., mig 0.005)



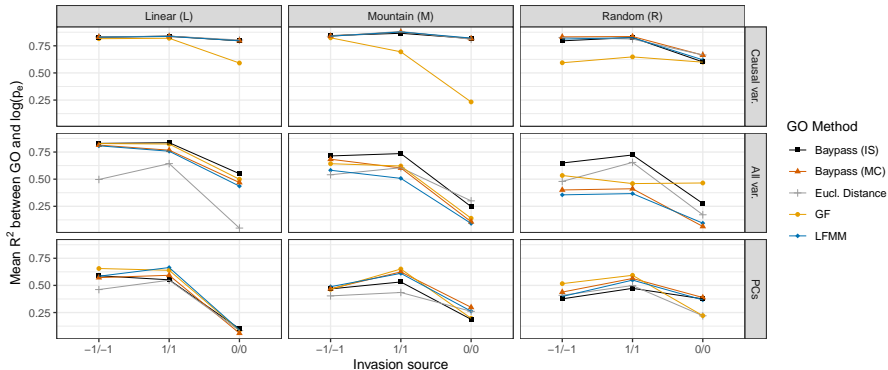
Performance reduction linked to confounding variables

Mean R^2 between GO and establishment prob. (10 ind., mig 0.005)



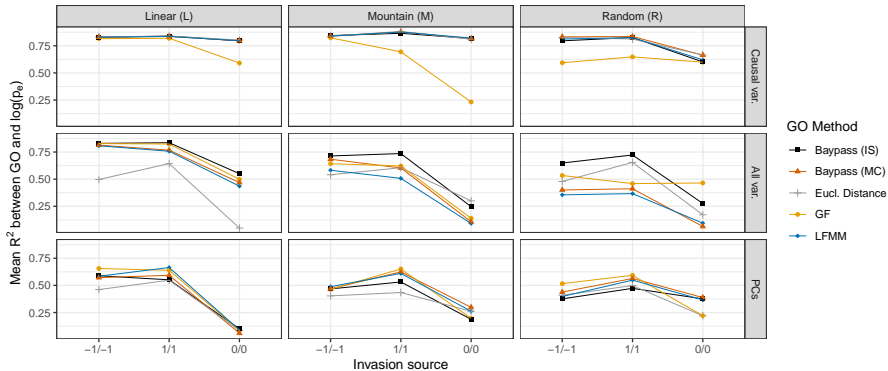
PCs bridge the performance gap between methods

Mean R^2 between GO and establishment prob. (10 ind., mig 0.005)



PCs bridge the performance gap between methods

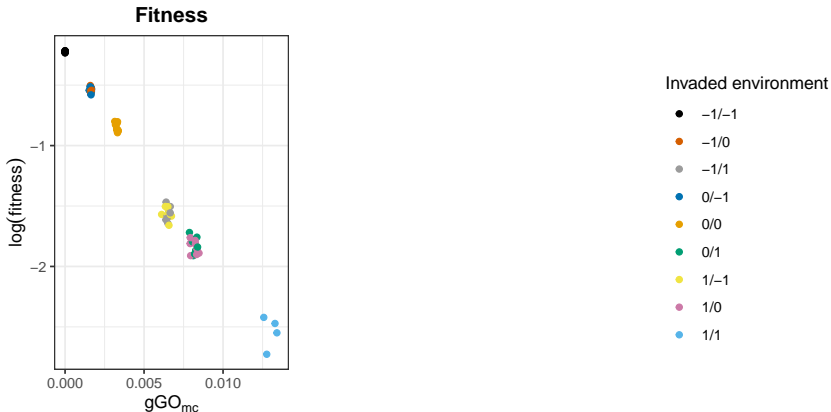
Mean R^2 between GO and establishment prob. (10 ind., mig 0.005)



Similar results with high migration rate

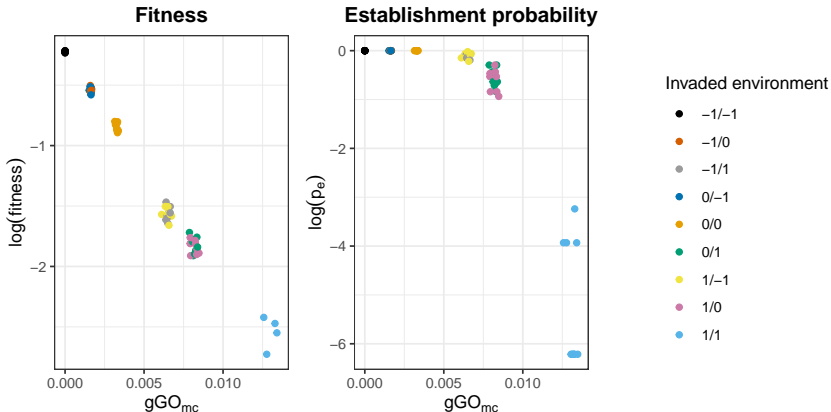
The issue of interpreting GO values

With 100 individuals → strong relationship with fitness



The issue of interpreting GO values

With 100 individuals → strong relationship with fitness
But adaptive challenge buffered



Conclusions

Relationship between GO and establishment probability?

- Strong correlation between GO and establishment probability
- Robust to the strenght of adaptation in the native area
- Disturbed by confounding variables and environmental complexity

Conclusions

Relationship between GO and establishment probability?

- Strong correlation between GO and establishment probability
- Robust to the strenght of adaptation in the native area
- Disturbed by confounding variables and environmental complexity

Methods performances?

- Use of univariate methods or PCs recommended
- Geometrics GO perform better

Conclusions

Relationship between GO and establishment probability?

- Strong correlation between GO and establishment probability
- Robust to the strenght of adaptation in the native area
- Disturbed by confounding variables and environmental complexity

Methods performances?

- Use of univariate methods or PCs recommended
- Geometrics GO perform better

Some questions remains ...

- Influence of admixture? Number of successive introduction?
Genetic Load?
- Interpretability of absolute GO values?

Preprint

File available

Predicting species invasiveness with genomic data: is Genomic Offset related to establishment probability?

February 2024

DOI: [10.1101/2024.02.20.581132](https://doi.org/10.1101/2024.02.20.581132)

License · [CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)

 Louise Camus · Mathieu Gautier · Simon Boitard

Preprints and early-stage research may not have been peer reviewed yet.

Thank you for your attention !

Simulations design : native area



- Mutation rate : 10^{-7} (QTLs and neutral markers)
- Recombination rate : 10^{-5} (50 cM)
- Randomly drawn effect sizes for each QTLs
- Σ effect sizes = phenotype \leftrightarrow environment = fitness
- Neutral mutations added after the simulation