B4EST International Conference

Managing Forest Genetic Resources (FGR) for an uncertain future

Predict climate response and adaptation in Cork Oak using gradient forest

Octávio S. Paulo

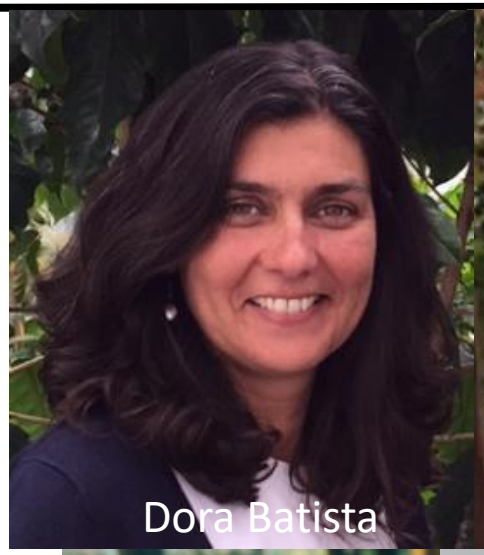
cE3c - Center for Ecology, Evolution and Environmental Changes & CHANGE - Global Change and Sustainability Institute,
Departamento de Biologia Animal. Faculdade de Ciências, Universidade de Lisboa

Ciências ULisboa Faculdade de Ciências da Universidade de Lisboa

CHANGE Global Change and Sustainability Institute

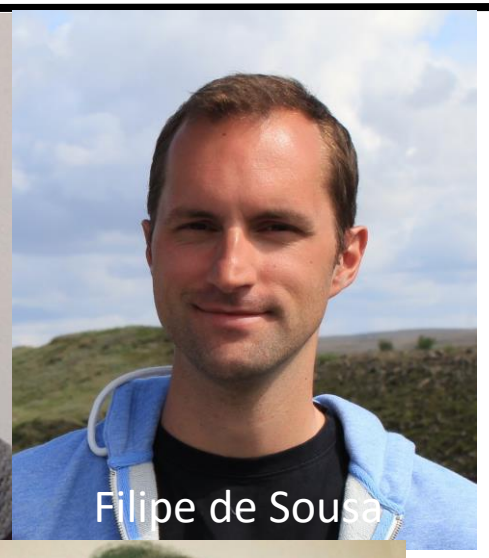
cE3c centre for ecology, evolution and environmental changes

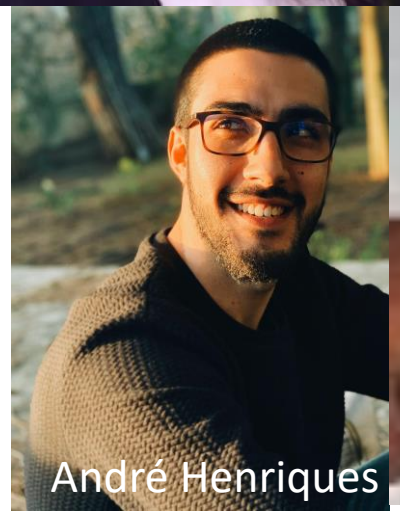
Acknowledgments

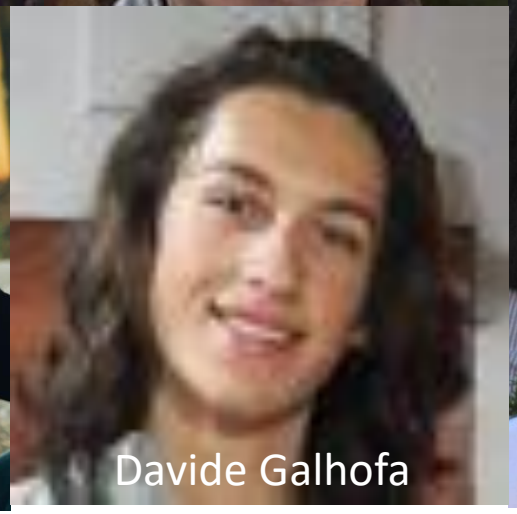

Dora Batista

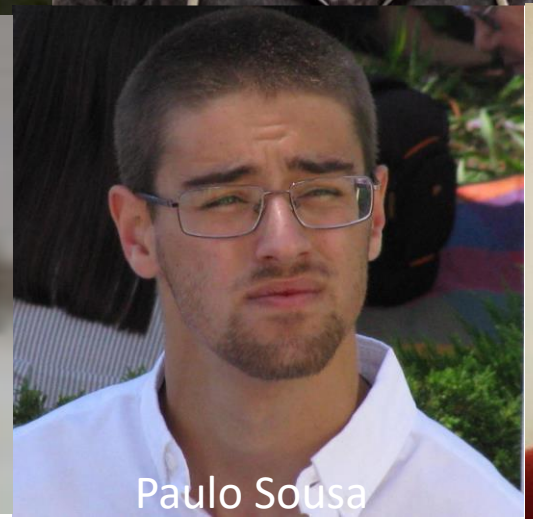
Francisco Pina.Martins

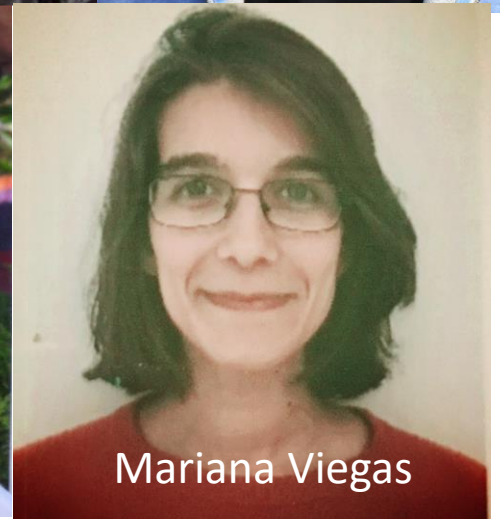
Mathieu Vanhove


Filipe de Sousa


André Henriques


Davide Galhofa


Paulo Sousa


Mariana Viegas

Ciências ULisboa
Faculdade de Ciências da Universidade de Lisboa

CoBiG²
Computational Biology & Population Genomics Group

cE3c
centre for ecology, evolution and environmental changes

Acknowledgments

Université Mohammed V
Faculté des Sciences
Rabat

UNIVERSITY OF MICHIGAN

iniav
Instituto Nacional de Investigação Agrária e Veterinária, I.P.

UAlg
UNIVERSIDADE DO ALGARVE

Ciências ULisboa
Faculdade de Ciências da Universidade de Lisboa

CHANGE
Global Change and Sustainability Institute

cE3c
centre for ecology, evolution and environmental changes

Cristina Branquinho

Lacey Knowles

Adriana Silva

Augusta Costa

Ana Cristina Coelho

Bouchra Belkadi

Acknowledgments

KeePace

Keep Pace: Selection of trees keeping pace with fast environmental changes, a science based approach for sustainable XXI century Oak forests – KeePace. LISBOA-01-0145-FEDER-029263 PTDC/ASP-SIL/29263/2017

Questions addressed by the Project KeePace

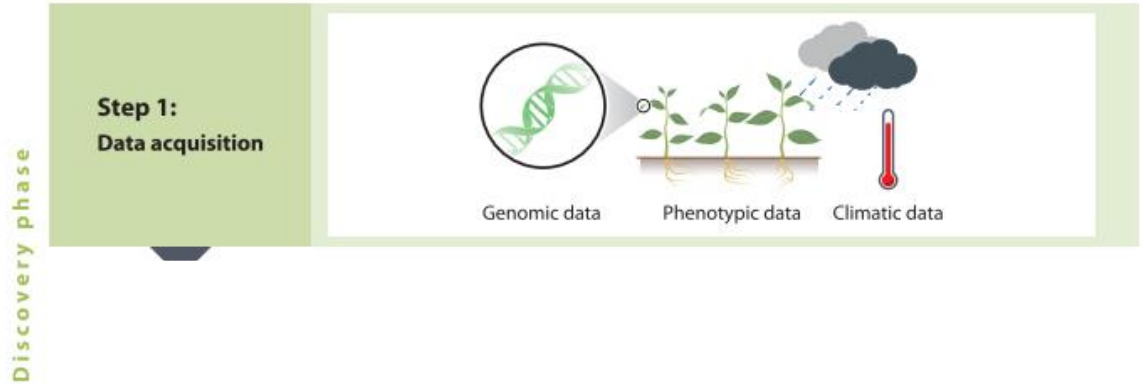
1 – Understand the population structure and demographic history

2 – Detection of the genomic signatures of natural selection in natural populations

3 – Which genes/genomic regions are associated with relevant environmental variables

4 – How populations will respond genetically to climate change

Discovery phase – Step 1: Data acquisition



Capblancq *et al.* 2020 AREES

Sampled Populations/Locations

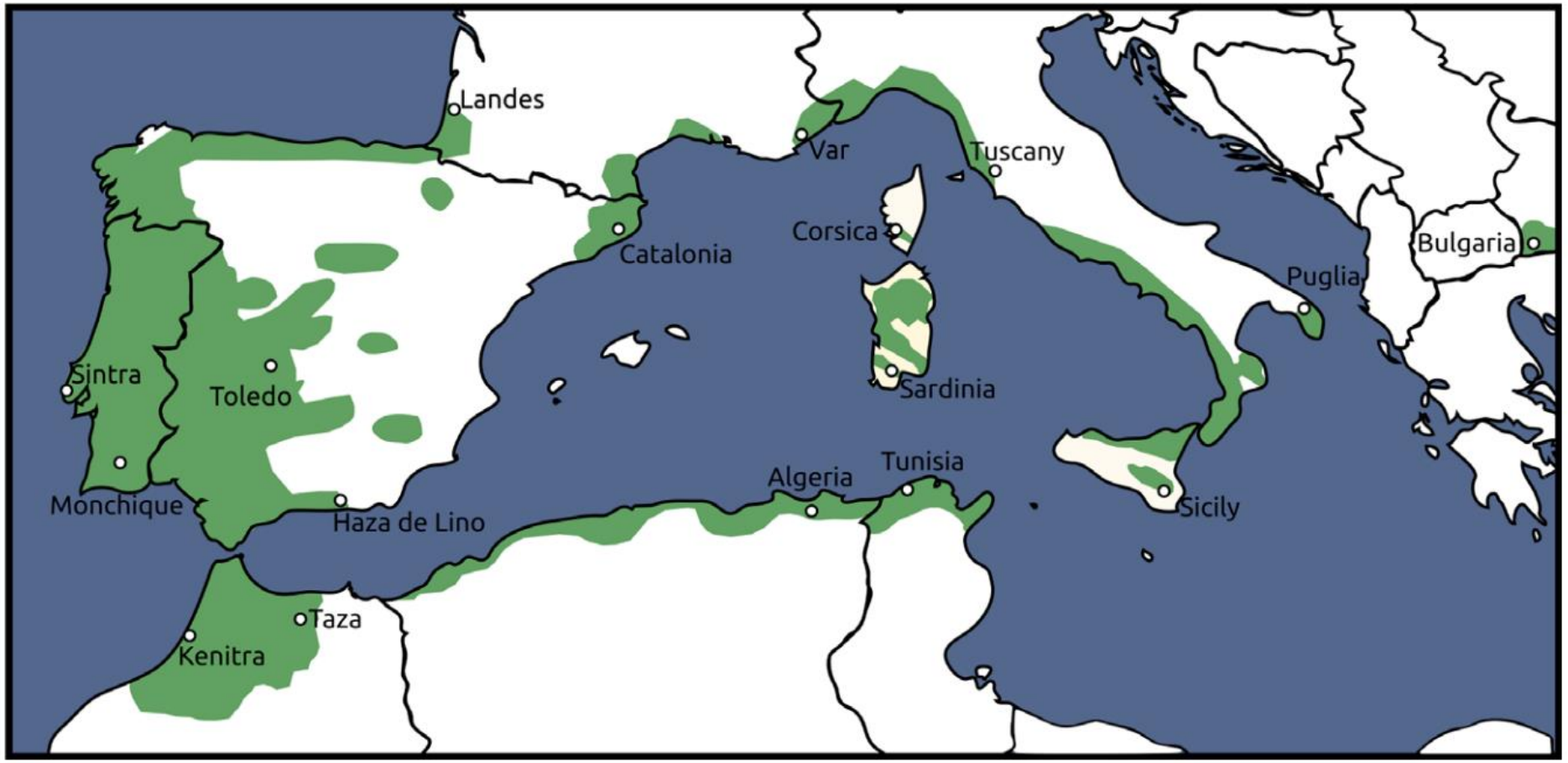
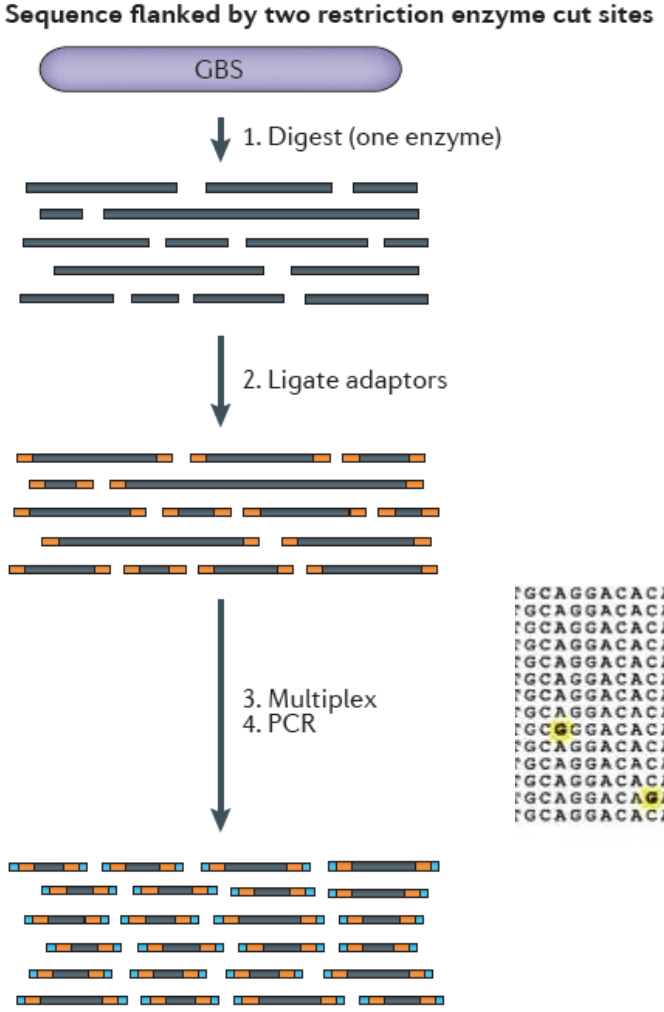
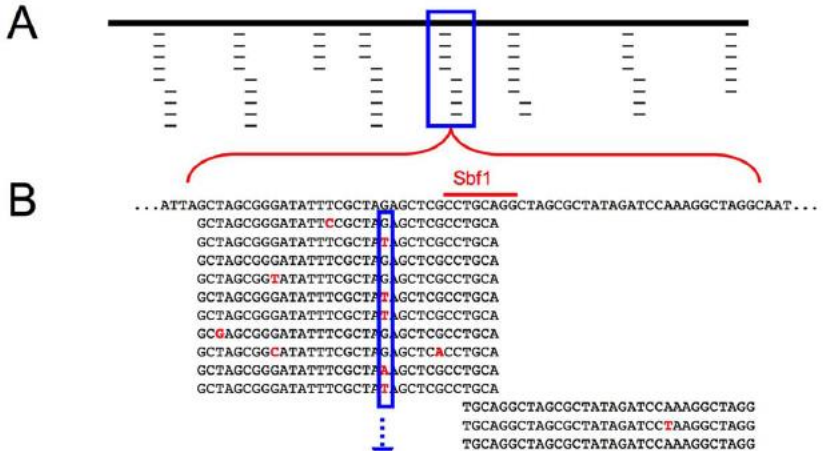
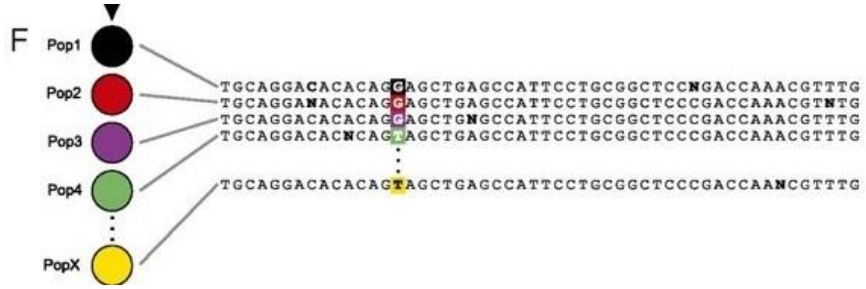


FIGURE 1 A map of cork oak (*Quercus suber*) distribution. Shaded land areas represent the species' range. White dots represent the sampling locations. Adapted from EUFORGEN 2009 (www.euforgen.org) [Colour figure can be viewed at wileyonlinelibrary.com]

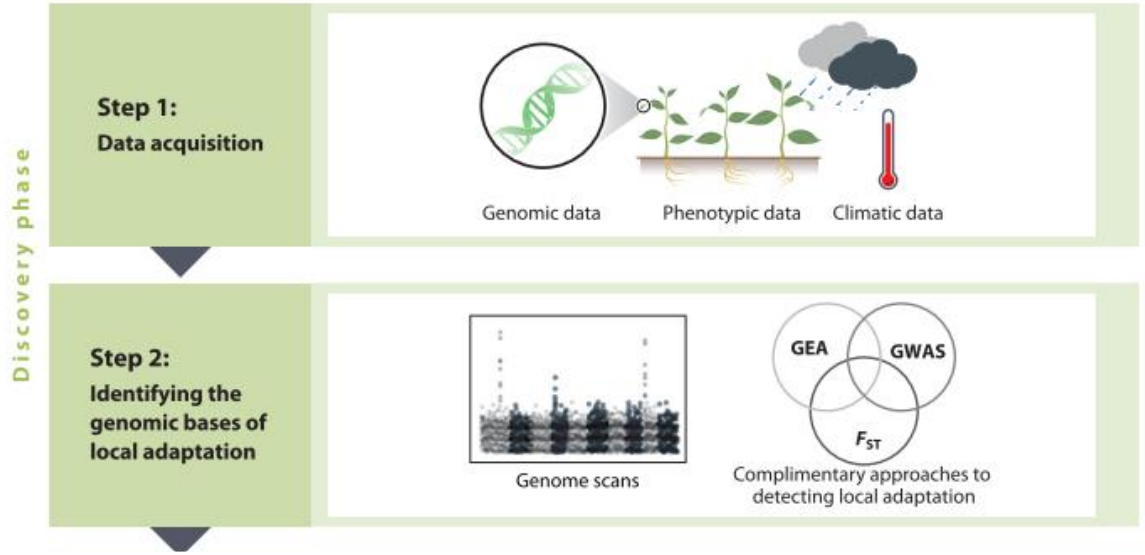
GBS – Genotyping-by-Sequencing





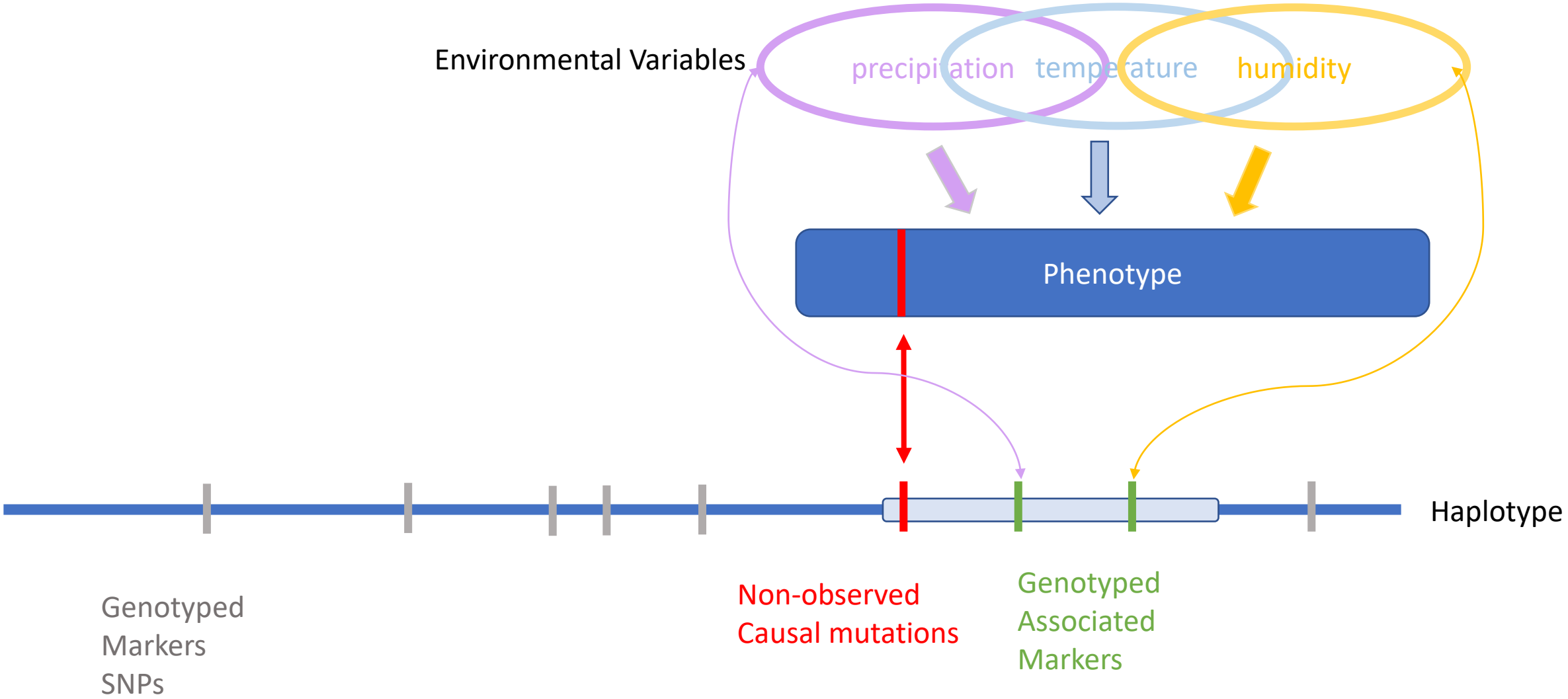


Discovery phase – Step 2: Identifying the genomic base of local adaptation



Capblancq *et al.* 2020 AREES

Genotype Environment Association Analysis



First results

Received: 10 August 2018 | Revised: 10 October 2018 | Accepted: 15 October 2018

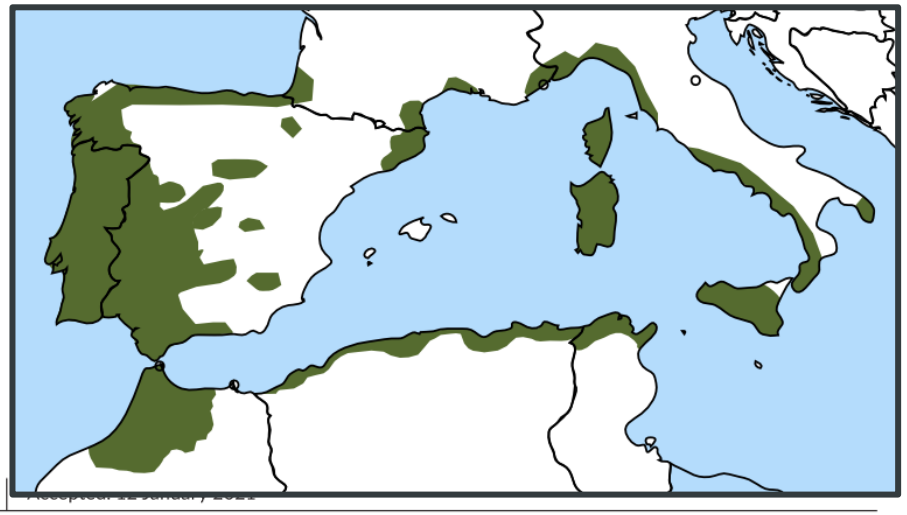
DOI: 10.1111/gcb.14497

PRIMARY RESEARCH ARTICLE

WILEY Global Change Biology

New insights into adaptation and population structure of cork oak using genotyping by sequencing

Francisco Pina-Martins[1] | João Baptista[2] | Georgios Pappas Jr[3] | Octávio S. Paulo[1]



Received: 13 May 2020 | Revised: 6 January 2021

DOI: 10.1111/jeb.13765

SPECIAL ISSUE

Journal of Evolutionary Biology esseb WILEY

Using gradient Forest to predict climate response and adaptation in Cork oak

Mathieu Vanhove[1] | Francisco Pina-Martins[1] | Ana Cristina Coelho[2] | Cristina Branquinho[1] | Augusta Costa[3,4] | Dora Batista[1,5] | Adriana Príncipe[1] | Paulo Sousa[1] | André Henriques[1] | Isabel Marques[1,6] | Bouchra Belkadi[7] | L. Lacey Knowles[8] | Octávio S. Paulo[1]

1996 SNPs, 95/81 Trees, 259 SNPs non-neutrals

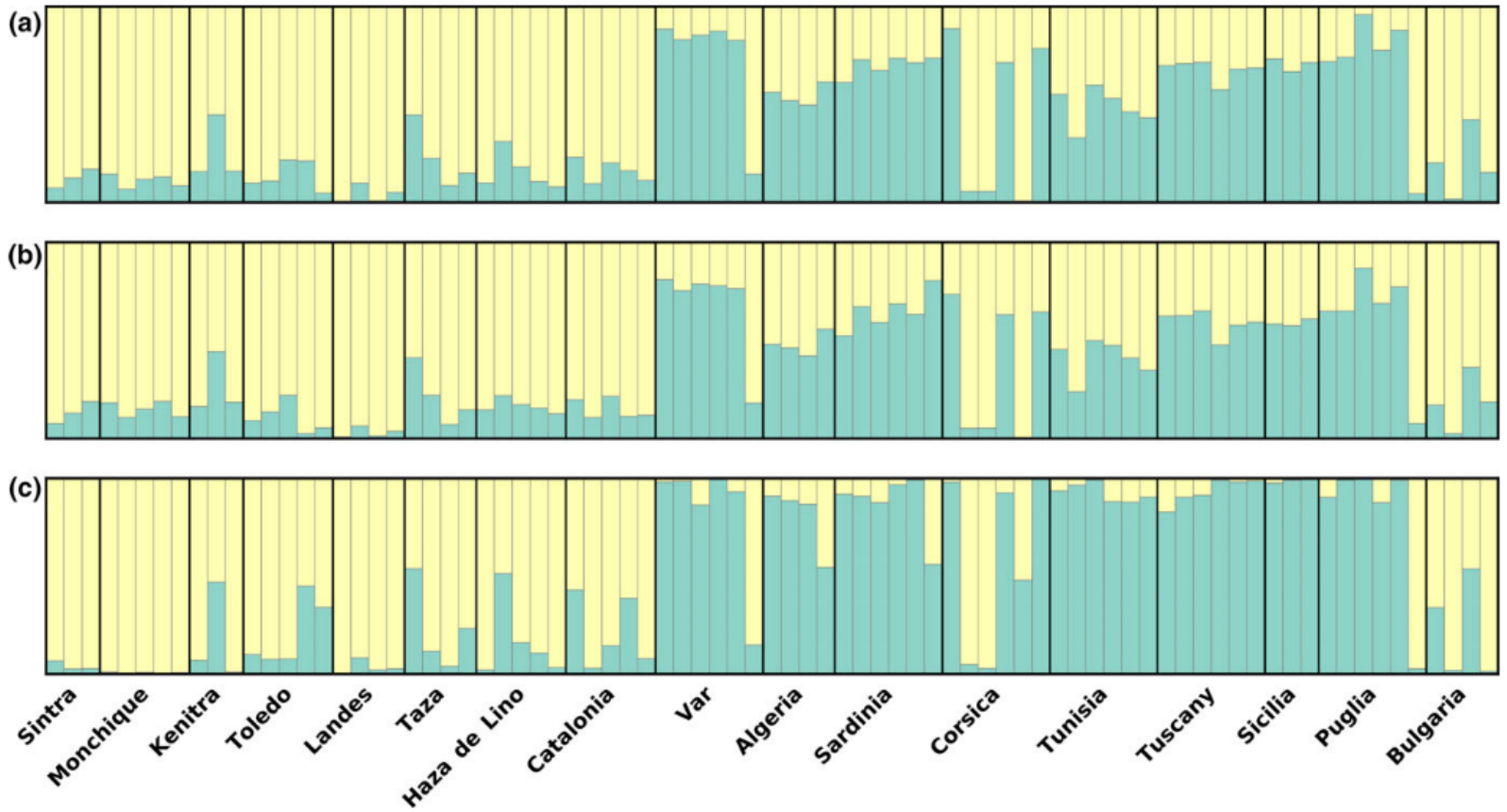


FIGURE 2 MAVERICK clustering plots for $K = 2$. Sampling sites are presented from West to East. "a" is the Q-value plot for the dataset with all loci, "b" is for the dataset with only "neutral" loci, and "c" if for the dataset with only "non-neutral" loci [Colour figure can be viewed at wile yonlinelibrary.com]

Pina-Martins *et al.* 2019 GCB

2583 SNPs, 95/81 Tree, 265 SNPs non-neutrals

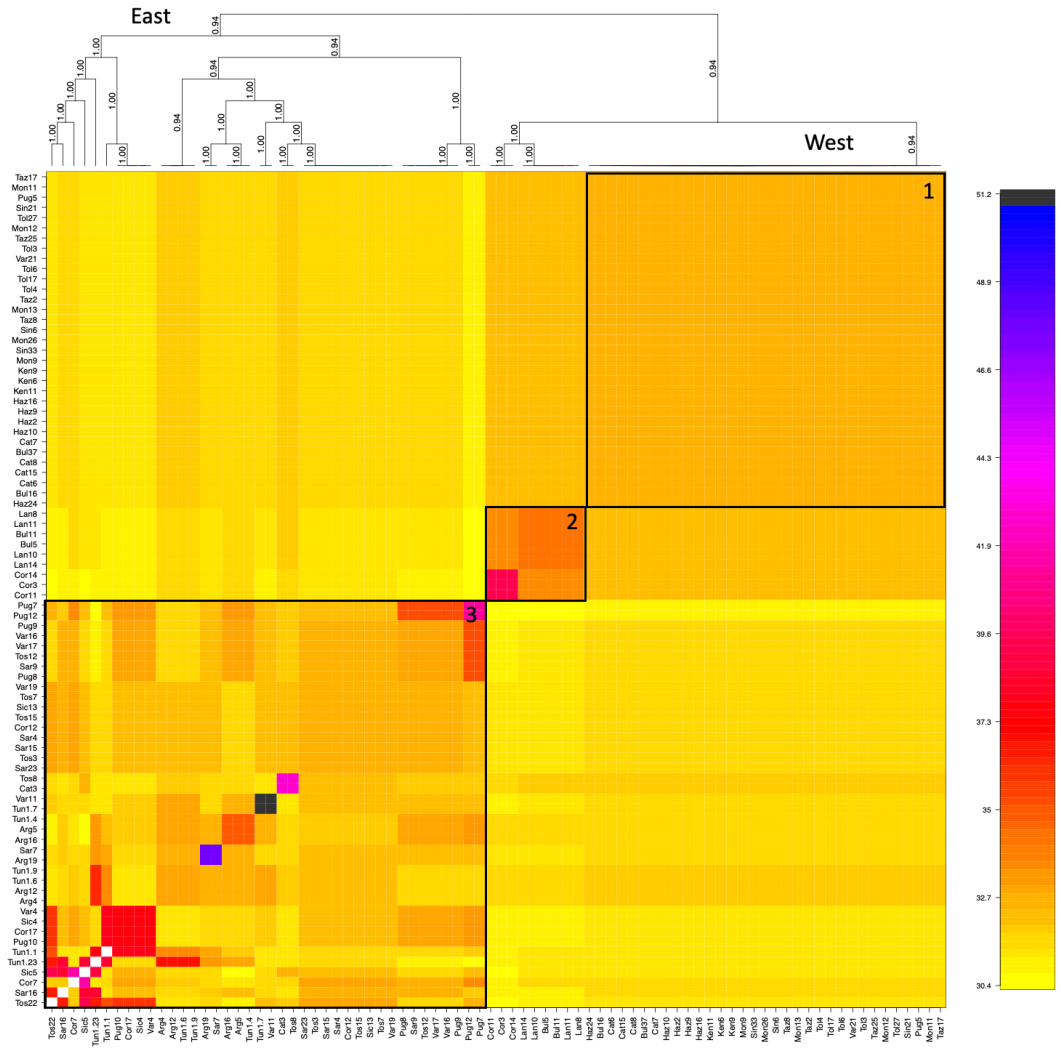


FIGURE 1 Ecological niche, dark green, of *Q. suber* with each sampling location coloured according to the associated and group of populations (blue, black and red)

Vanhove *et al.* 2021 JEB

Genotype Environment Association Results

274 associations between 249 SNPs and 12 of the 16 tested environmental variables

Of all environmental variables, the one with most markers associated is:
“Precipitation of Driest Month” with 71 associations
“ Isothermality” with 35 associations
“Mean Temperature of Driest Quarter”, with 29 associations.

Sequences containing 22 of the 249 markers associated with environmental variables were matched to entries in the *Q. lobata* genome, however, of these only 10 were annotated

TABLE 2 Summary of BLAST hits for loci with single nucleotide polymorphisms (SNPs) associated to one or more environmental variables. “MTDQ” and “MTWQ” stand for “Mean Temperature of Driest Quarter” and “Mean Temperature of Wettest Quarter,” respectively

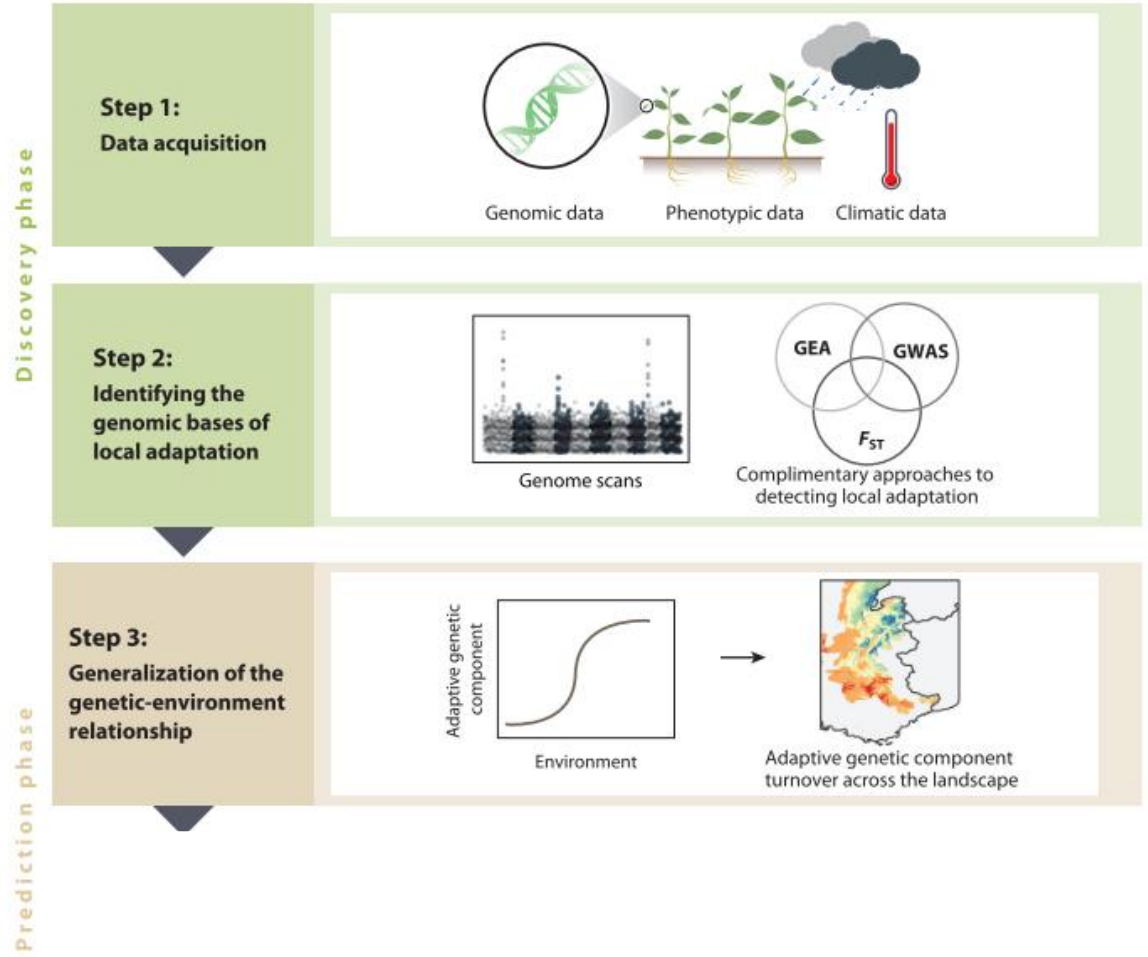
SNP name	Note (Similar to)	Associations
SNP 158	TRE1: Trehalase (*Arabidopsis thaliana*)	Mean Temperature of Driest Quarter
SNP 168	PER47: Peroxidase 47 (*Arabidopsis thaliana*)	Precipitation of Driest Month
SNP 233	CPSF160: Cleavage and polyadenylation specificity factor subunit 1 (*Arabidopsis thaliana*)	Annual Mean Temperature
SNP 333	Ascc1: Activating signal cointegrator 1 complex subunit 1 (*Mus musculus*)	Mean Temperature of Driest Quarter
SNP 455	GLCAT14A: Beta-glucurono syltransferase GlcAT14A (*Arabidopsis thaliana*)	Precipitation of Driest Month
SNP 619	GBP6: Guanylate-binding protein 6 (*Pongo abelii*)	Precipitation of Driest Month
SNP 834	NAC098: Protein CUP-SHAPED COTYLEDON 2 (*Arabidopsis thaliana*)	Longitude
SNP 880	TPP1: Thylakoidal processing peptidase 1%2C chloroplastic (*Arabidopsis thaliana*)	Mean Temperature of Warmest Quarter
SNP 1134	EMB2654: Pentatricopeptide repeat-containing protein At2g 41720 (*Arabidopsis thaliana*)	Mean Temperature of Driest Quarter
SNP 1589	At1g19525: Pentatricopeptide repeat-containing protein At1g19525 (*Arabidopsis thaliana*)	Temperature Seasonality

“Similar to TRE1: Trehalase,” which is known to play a role in drought stress (Houtte et al., 2013).

Located in a region matching the annotation of “Similar to PER47: Peroxidase 47,” which is known to play a role in drought response (Li et al., 2017).

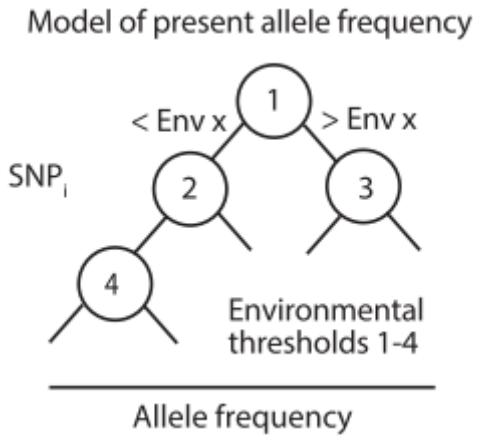
Pina-Martins *et al.* 2019 GCB

Prediction Phase – Step 3: Generalization of the genetic-environment relationship



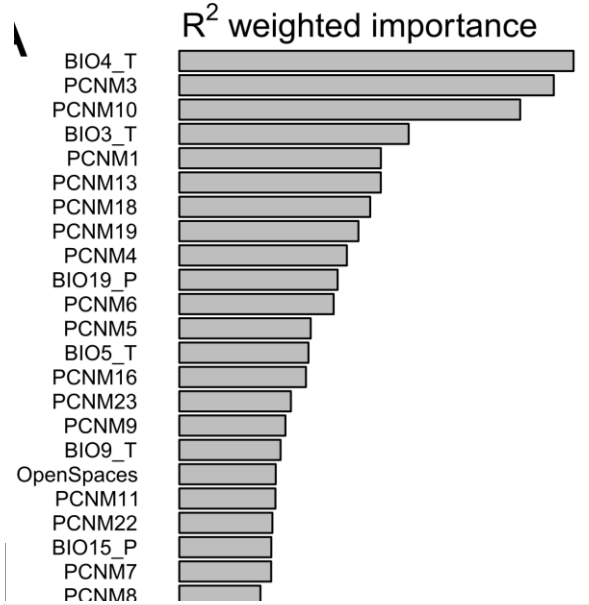
Capblancq *et al.* 2020 AREES

Gradient Forest GF

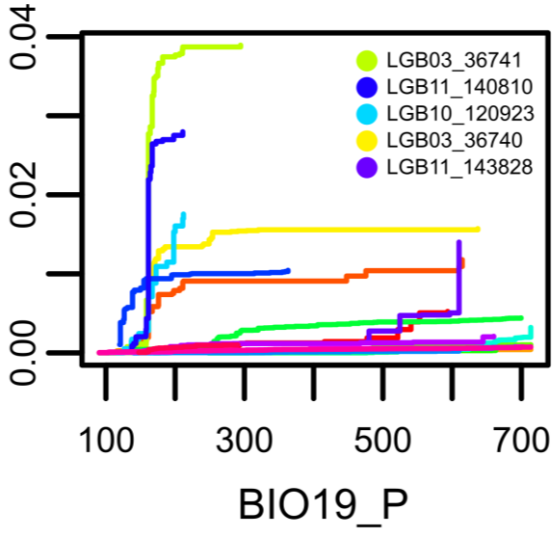


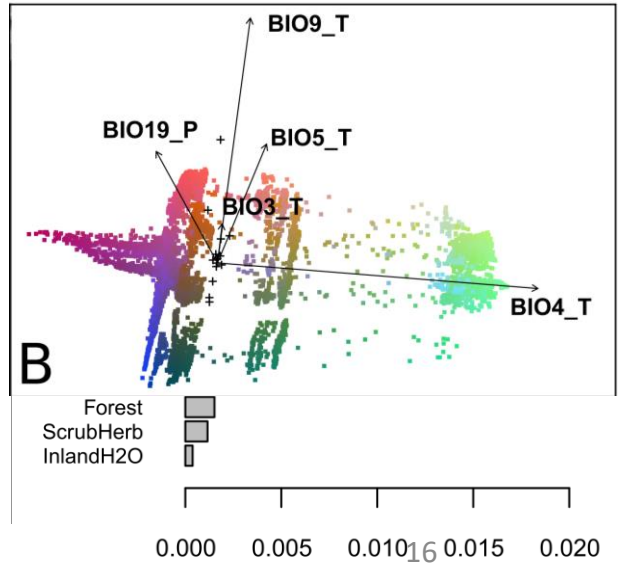
GF method construct a monotonic turnover function for each SNP and an aggregate (for all SPNs) turnover function across all SNPS using a strategy of "predict, then assemble"

Fitzpatrick & Keller 2015 Ecol Let

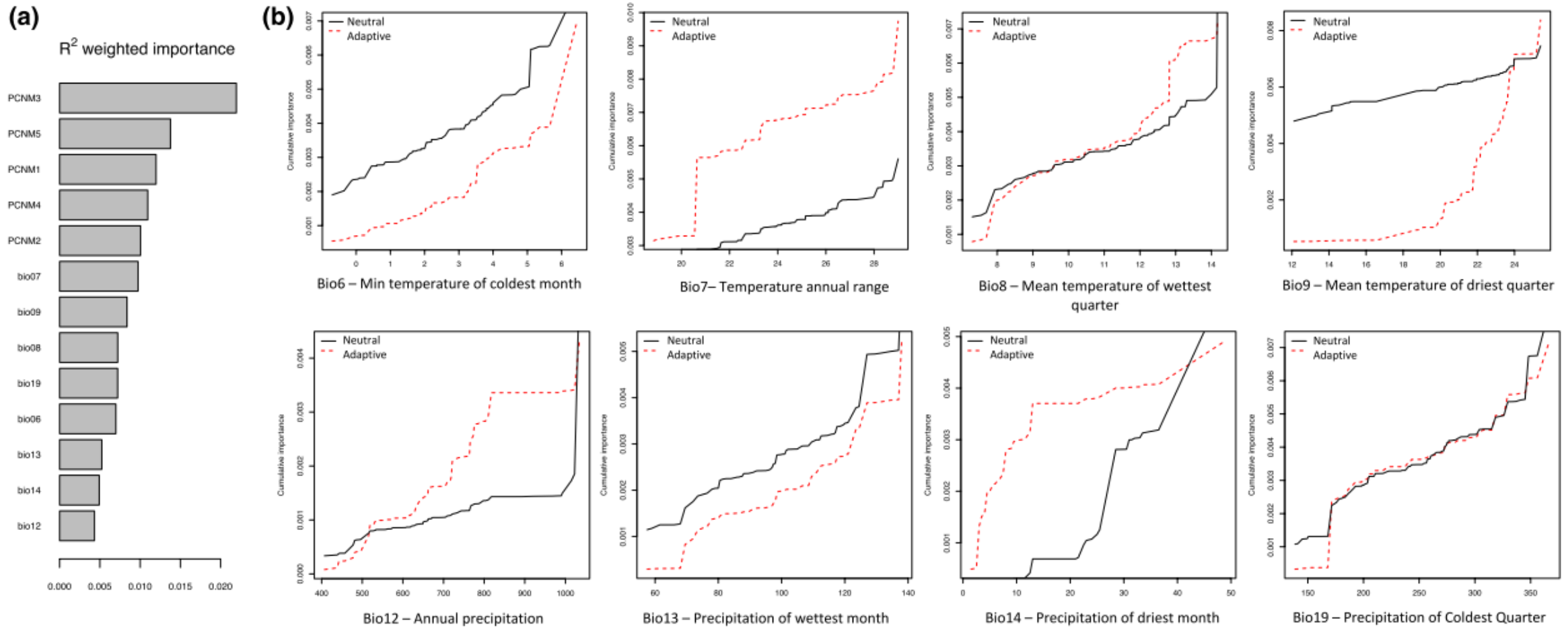


Gradient forest is non-parametric method that use "machine-learning regression tree approach" known as Random Forest





Gradient Forest - SNPs Neutral vs 265 SNPs Adaptive

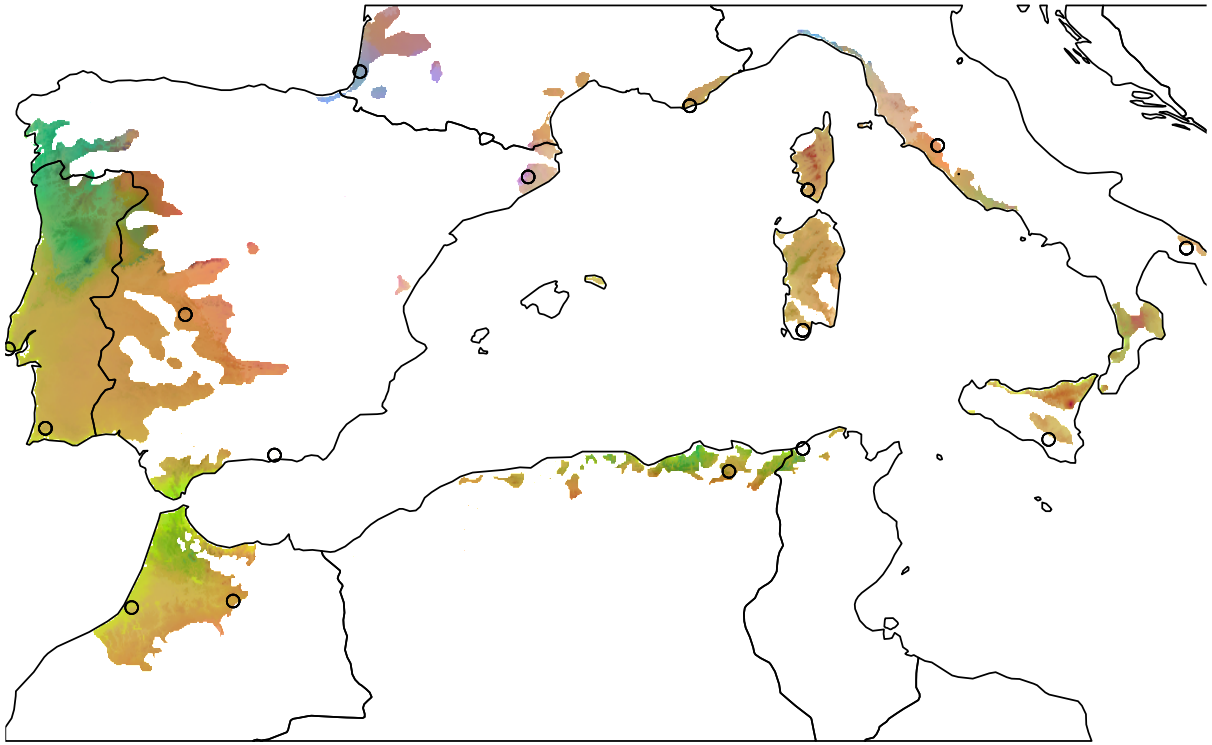


(a) R2- weighted importance of environmental and spatial variables for the adaptive SNP model.
(b) Cumulative importance of allelic change along environmental gradients for the adaptive and neutral SNP model

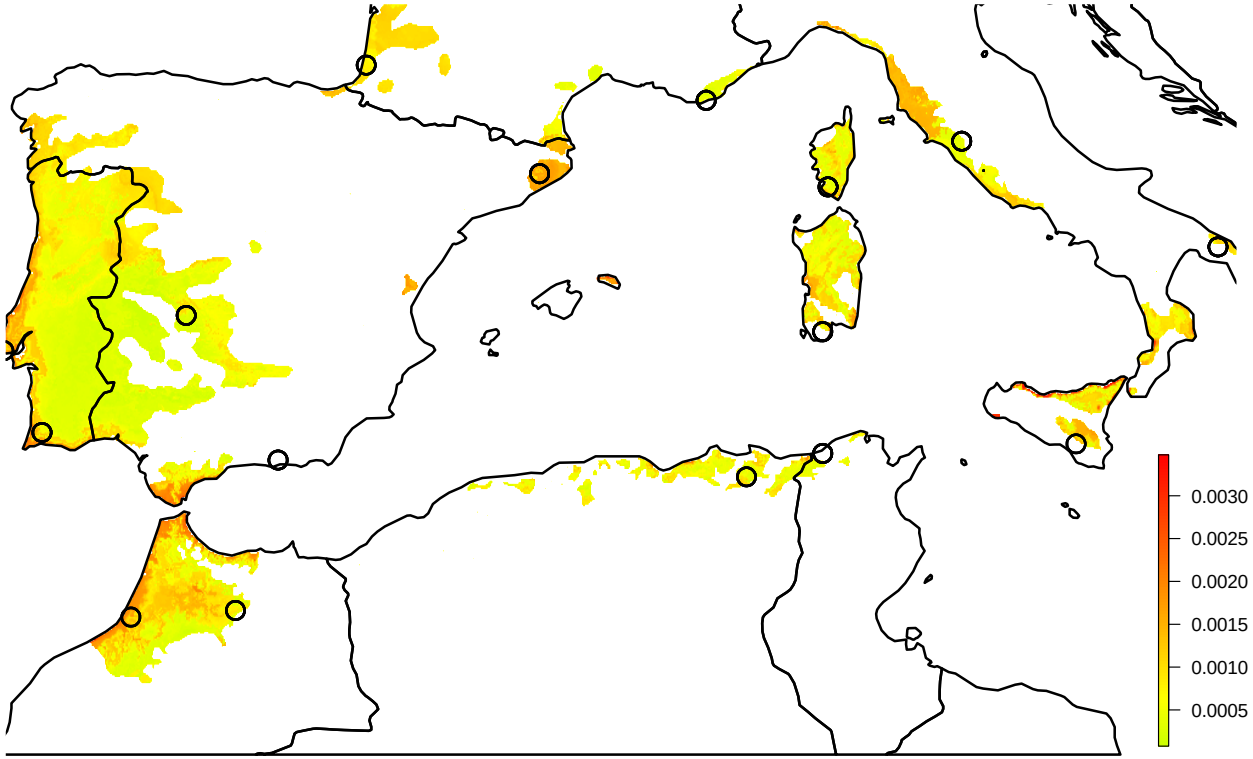
Vanhove *et al.* 2021 JEB

Maps of predicted turnover in allele frequencies - Gradient Forest for adaptive SNPs

Current landscape patterns of allelic composition predicted under the gradient forest for the adaptive SNPs dataset.



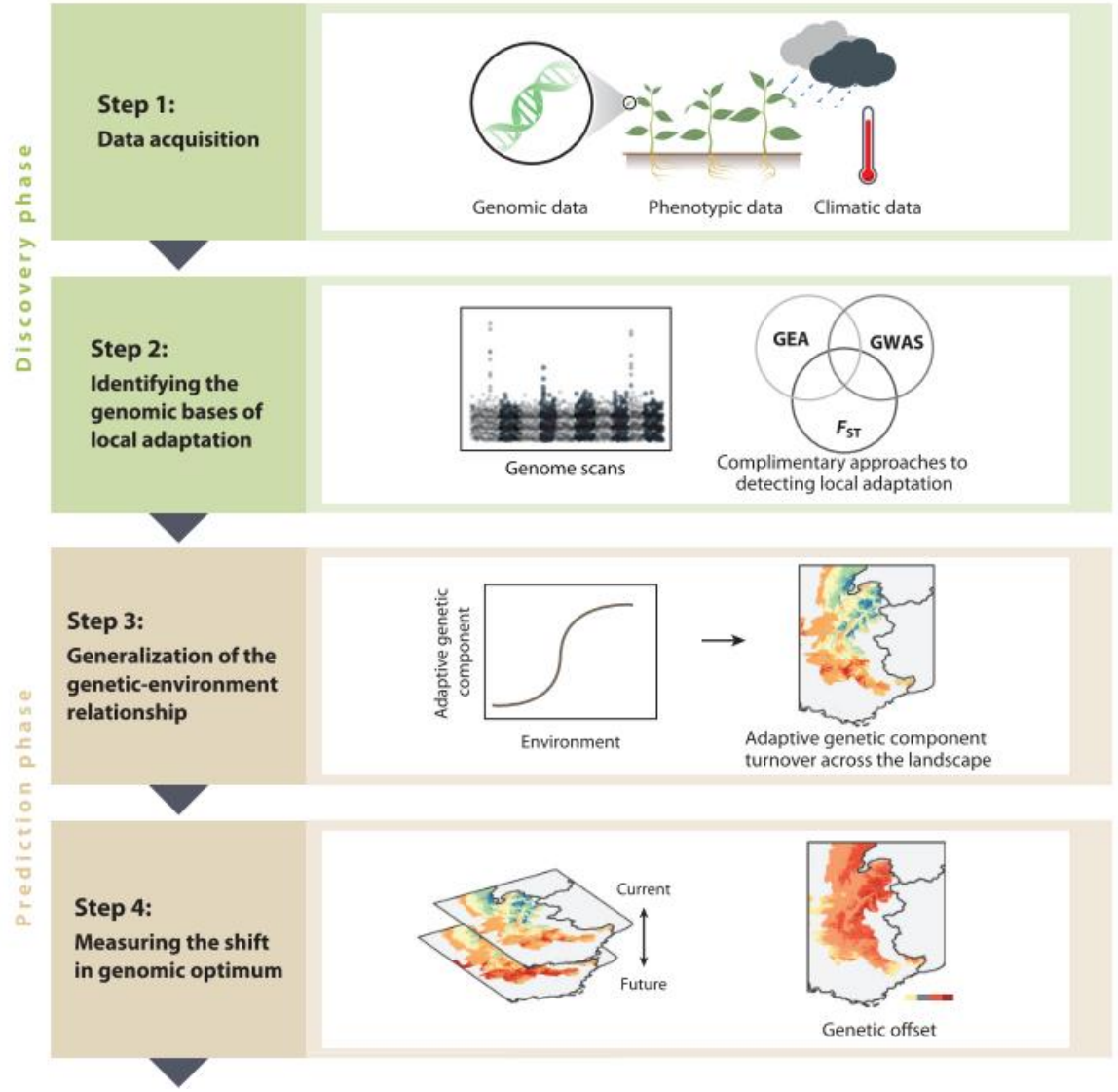
Difference between full SNP and adaptive SNPs datasets based on Procrustes residuals.



Scale - a red-green-blue color scale was generated using the first three axes of the principal components of the gradient forest prediction. Regions with similar colors are expected to have similar genetic composition.

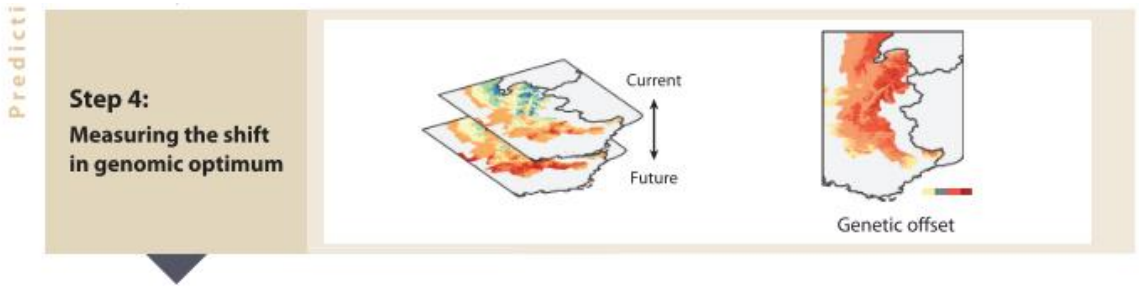
Vanhove *et al.* 2021 JEB

Prediction Phase – Step 4 – Measuring the shifts in genomic optimum

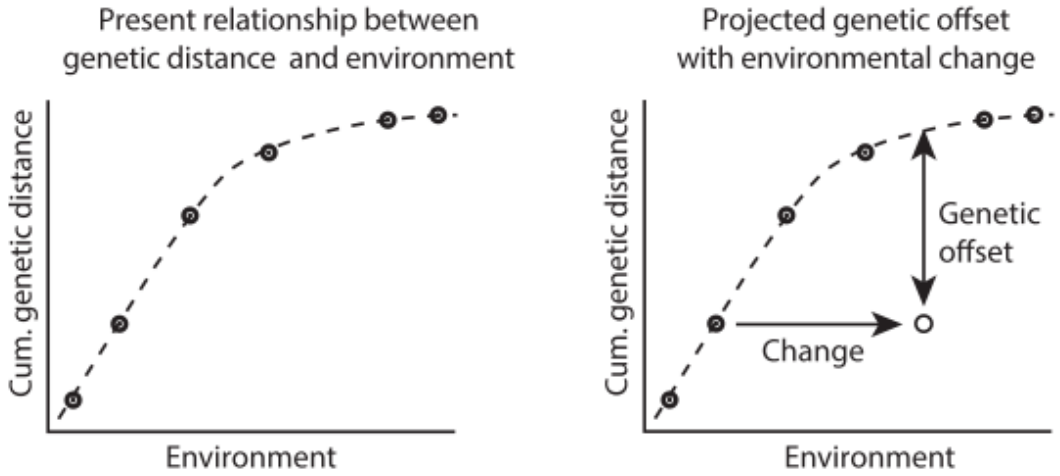


Capblancq *et al.* 2020 AREES

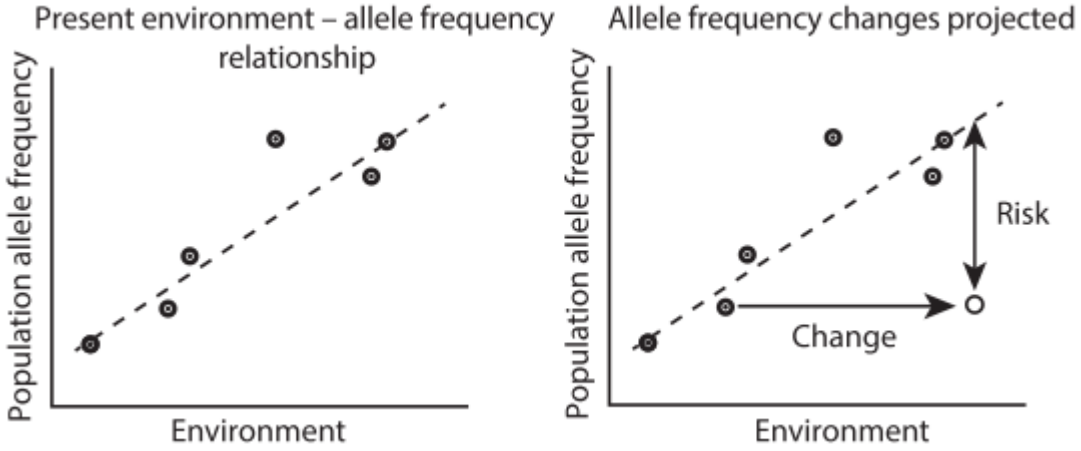
Prediction Phase - Genetic offset



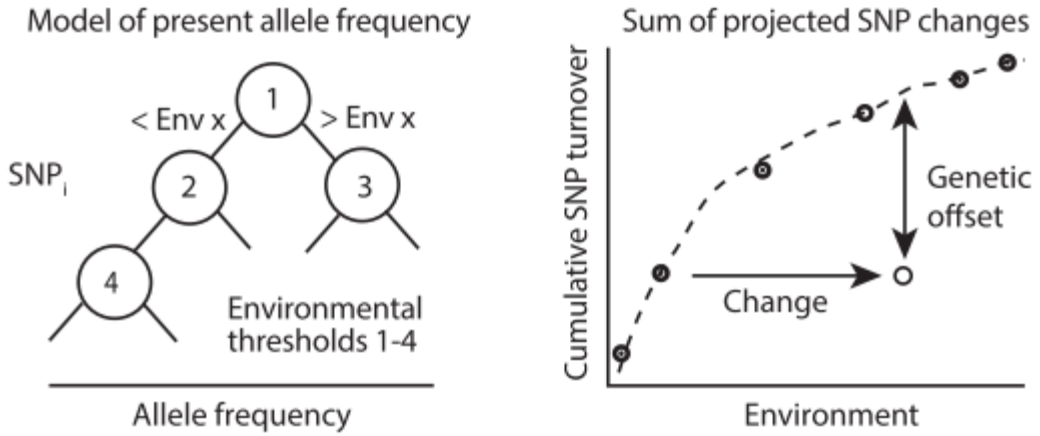
Generalized dissimalirity modelling



Risk of non-adaptedness



Gradient forest

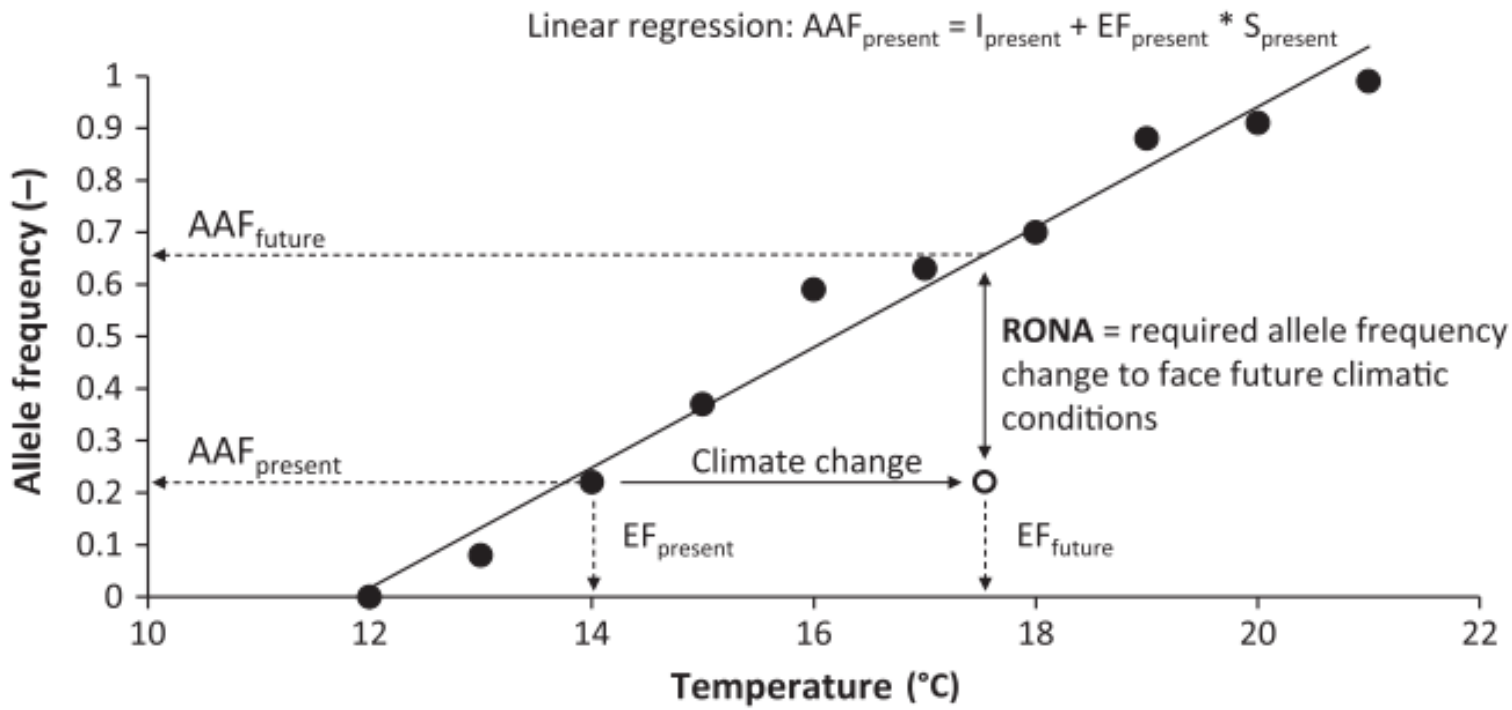


Rellstab *et al.* 2021 Evol Appl

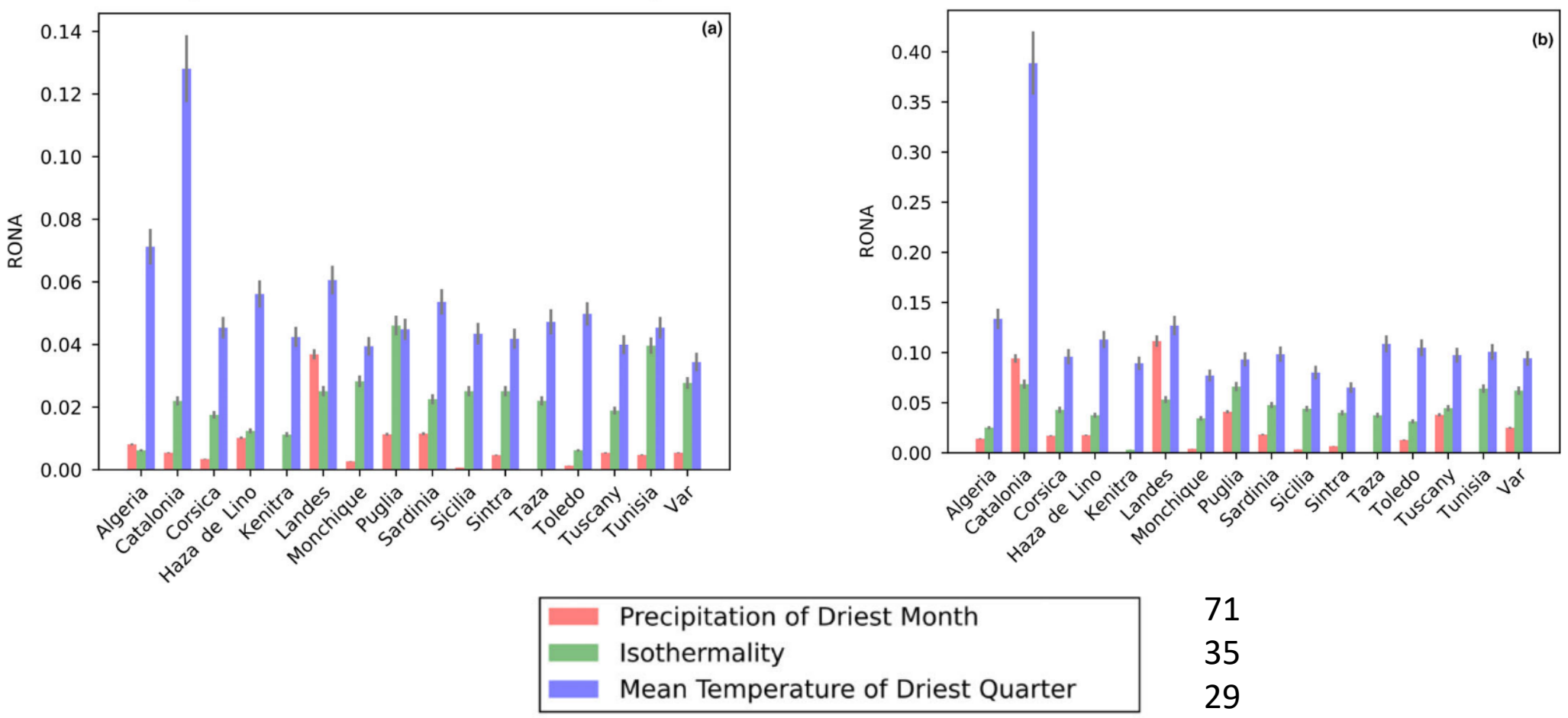
Risk of non-adaptedness RONA

The average absolute difference between $AAF_{present}$ and AAF_{future} represents the RONA of a population to future conditions, or, in other words, the average change in allele frequency at the *n* relevant loci that is needed to be adapted to altered conditions. This results in the following formula:

$$\mathrm{RONA} = \frac{1}{n}\sum_{i=1}^{n} \left|\left(S_{\mathrm{present}_i} * \mathrm{EF}_{\mathrm{future}} + I_{\mathrm{present}_i}\right) - \mathrm{AAF}_{\mathrm{present}_i}\right|$$



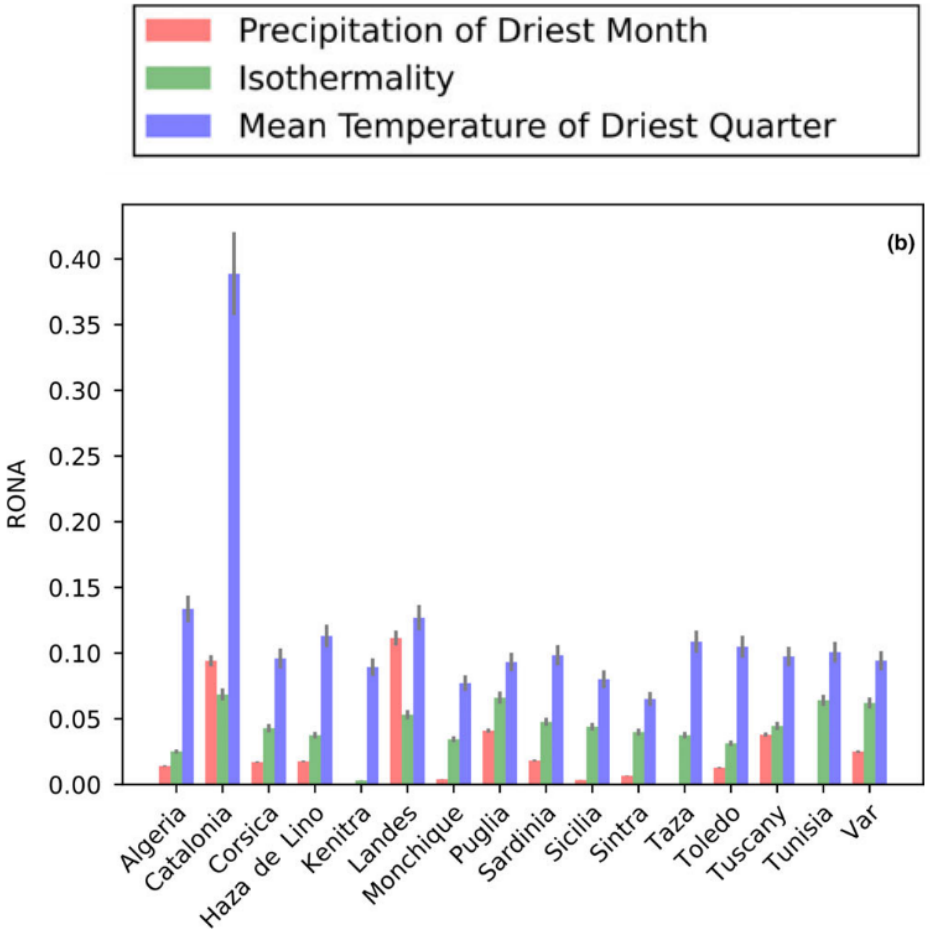
RONA RCP26 (a) e RCP85 (b)



71
35
29

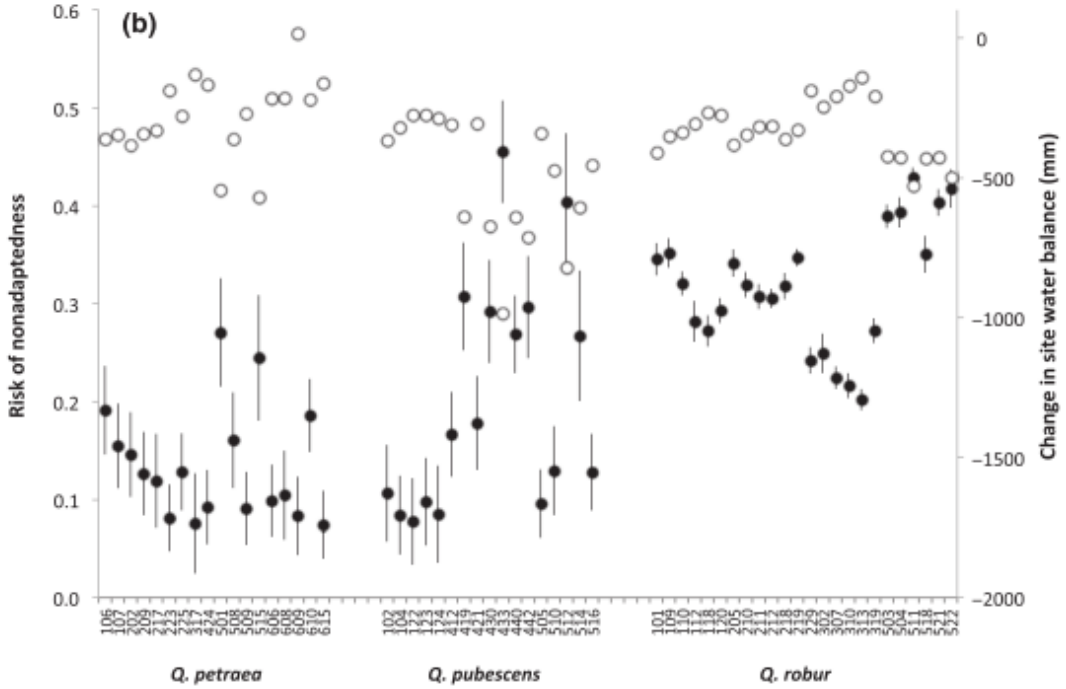
Pina-Martins *et al.* 2019 GCB

Risk of non-adaptedness RONA



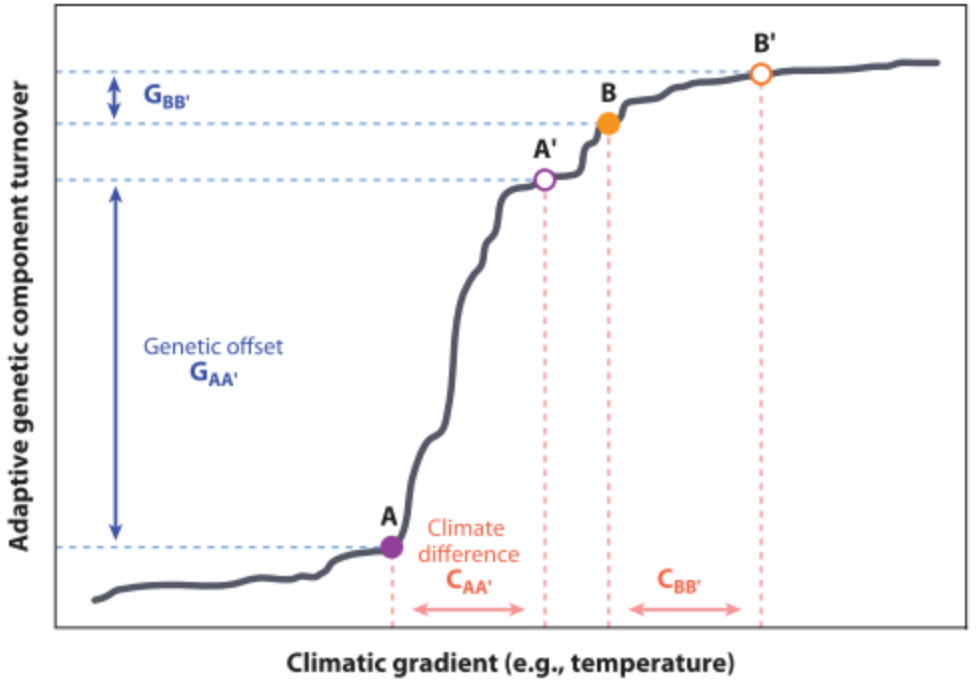
Pina-Martins *et al.* 2019 GCB

Climate change predict between (1931 and 1960) e (2071–2100) in the habaitats of the 71 population and RONA



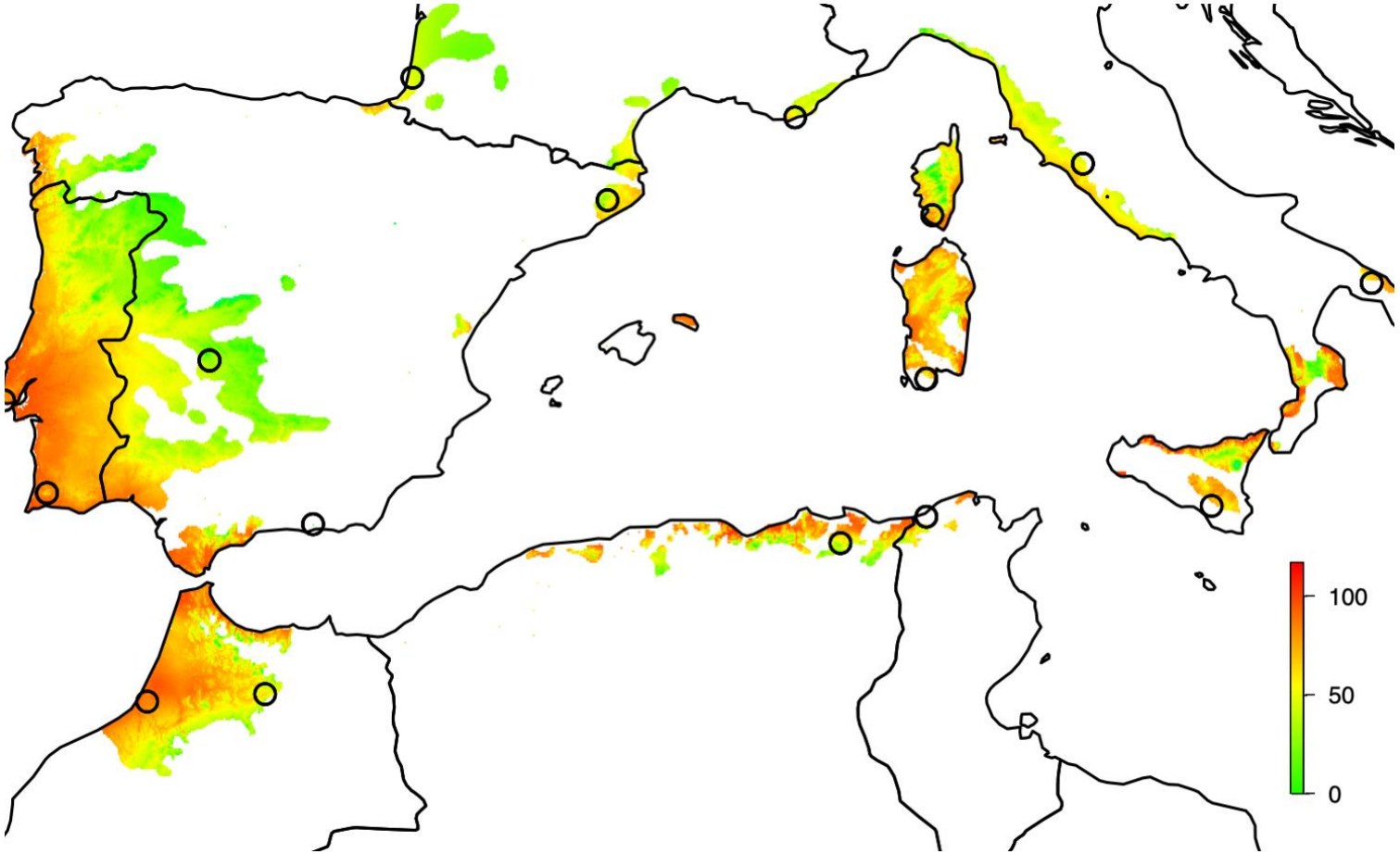
Rellstab et al. 2016 MEC

Genetic offset



Capblancq *et al.* 2020 AREES

KeePace – Genetic offset



Predictive genetic offset (full SNPs dataset) under climate change for 2070. Euclidian distances between current and future climate were calculated for each model and regions with larger euclidian distances are expected to have a larger genetic offset.

Vanhove *et al.* 2021 JEB

Predict the response to climate change

Table 1 Recent studies predicting a future change in local genomic optima due to climate change

Species	Type	Spatial scale	Data	Adaptive component identification	Predictive model	Validation procedure	Reference
Populus balsamifera	Tree	North America	Genomic (targeted genotyping of candidate genes)	F_{ST} (Arlequin, BayeScan), GEA (Bayenv), GPA	Gradient forest and GDM	None	Fitzpatrick & Keller 2015
Arabidopsis thaliana	Plant	Europe	Genomic (whole genome)	GWAS (EMMAX25, BSLMM28)	Ecological niche modeling on alleles	None	Exposito-Alonso et al. 2018
Populus tremula	Tree	Sweden	Genomic (whole genome)	GEA (LFMM)	GDM	None	Ingvarsson & Bernhardsson 2018
Quercus rugosa	Tree	Mexico	Genomic (GBS)	F_{ST}	Gradient forest	None	Martins et al. 2018
Empidonax traillii	Bird	North America	Genomic (RADseq)	GEA (LFMM)	Gradient forest	Correlation with population historical trends	Ruegg et al. 2018
Setophaga petechia	Bird	North America	Genomic (RADseq)	GEA (LFMM)	Gradient forest	Correlation with population historical trends	Bay et al. 2018
Quercus suber	Tree	Western Mediterranean	Genomic (GBS)	F_{ST} (BayeScan), SelEstim	RONA	None	Pina-Martins et al. 2019
Myotis crypticus and *Myotis escalerai*	Wild mammal	Western Europe	Genomic (RADseq)	GEA (LFMM, RDA)	Ecological niche modeling separating two extreme adaptive groups	None	Razgour et al. 2019
Eucalyptus melliodora	Tree	Australia	Genomic (GBS)	None	GDM using the complete genomic data set	None	Supple et al. 2018
Arabidopsis thaliana	Plant	Europe	Genomic (whole genome) Phenotypic (common gardens)	GWAS (LM-GEMMA)	GWES	Model evaluation procedure	Exposito-Alonso et al. 2019
Zea mays	Plant	Mexico	Genomic (published data)	GEA (Bayenv, Bayescenv)	Gradient forest	None	Aguirre-Liguori et al. 2019
Fagus sylvatica	Tree	French Alps	Genomic (RADseq)	GEA (RDA)	RDA	None	Capblancq et al. 2019
Capra aegagrus hircus	Domesticated mammal	Morocco	Genomic (whole genome)	GEA (logistic regressions, SAM)	SPAG	Model evaluation procedure	Rochat & Joost 2019
Mimosa acutistipula and *Dioclea apurensis*	Plant	Carajás Mineral Province (Brazil)	Genomic (GBS) Phenotypic (sample traits)	GEA (RDA, LFMM), GPA (RDA, LFMM)	RDA	None	Carvalho et al. 2019
Populus balsamifera	Tree	North America	Genomic (targeted genotyping of candidate genes)	None	GDM	None	Gougherty et al. 2020

Abbreviations: GBS, genotype-by-sequencing; GDM, generalized dissimilarity modeling; GEA, genotype-environment association; GPA, genotype-phenotype association; GWAS, genome-wide association studies; GWES, genome-wide environment selection; RADseq, restriction-site associated DNA sequencing; RDA, redundancy analysis; RONA, risk of nonadaptiveness; SPAG, spatial areas of genotype probability.

Organism type: = Tree = Plant = Domesticated mammal = Wild mammal = Bird

Annual Review of Ecology, Evolution, and Systematics

Genomic Prediction of (Mal)Adaptation Across Current and Future Climatic Landscapes

Thibaut Capblancq,[1] Matthew C. Fitzpatrick,[2] Rachael A. Bay,[3] Moises Exposito-Alonso,[4,5] and Stephen R. Keller[1]

[1]Department of Plant Biology, University of Vermont, Burlington, Vermont 05405, USA; email: srkeller@uvm.edu

[2]Appalachian Laboratory, University of Maryland Center for Environmental Science, Frostburg, Maryland 21532, USA

[3]Department of Evolution and Ecology, University of California, Davis, California 95616, USA

[4]Department of Plant Biology, Carnegie Institution for Science, Stanford, California 94305, USA

[5]Department of Biology, Stanford University, Stanford, California 94305, USA

Capblancq *et al.* 2020 AREES

Limitations

Received: 4 December 2020 | Revised: 6 February 2021 | Accepted: 9 February 2021

DOI: 10.1111/eva.13205

PERSPECTIVE

Evolutionary Applications Open Access WILEY

Prospects and limitations of genomic offset in conservation management

CellPress

Cell
Leading Edge

Commentary

Opportunities and challenges in assessing climate change vulnerability through genomics

Ary A. Hoffmann,[1] Andrew R. Weeks,[1] and Carla M. Sgrò[2,*]
[1]School of BioSciences, Bio21 Institute, The University of Melbourne, Parkville, VIC 3010, Australia
[2]School of Biological Sciences, Monash University, Clayton, VIC 3800, Australia
*Correspondence: carla.sgro@monash.edu

021.02.006

past selection has affected allele frequencies across space, genomic tools are into adaptive evolutionary processes. Now researchers are considering how this an be used to predict the future vulnerability of species under climate change. ssessments show promise, but challenges remain.

Received: 22 September 2021 | Revised: 22 January 2022 | Accepted: 30 January 2022

DOI: 10.1111/eva.13354

ORIGINAL ARTICLE

Evolutionary Applications Open Access WILEY

Seeing the forest for the trees: Assessing genetic offset predictions from gradient forest

Áki Jarl Láruson[1] | Matthew C. Fitzpatrick[2] | Stephen R. Keller[3] | Benjamin C. Haller[4] | Katie E. Lotterhos[5]

Predict the response to climate change - ENM

Mitig Adapt Strateg Glob Change
DOI 10.1007/s11027-017-9738-z

ORIGINAL ARTICLE

Contribution of spatially explicit models to climate change adaptation and mitigation plans for a priority forest habitat

Ricardo A. Correia [1,2,3,4] · **Miguel N. Bugalho** [5] · **Aldina M. A. Franco** [2] · **Jorge M. Palmeirim** [1]

Received: 15 July 2016 / Accepted: 31 January 2017
© Springer Science+Business Media Dordrecht 2017

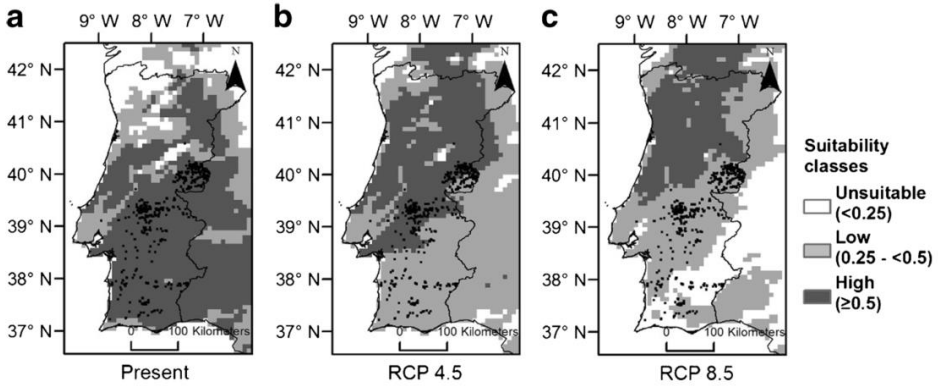


Fig. 4 Location of recently established cork oak stands in Portugal in relation to cork oak environmental suitability as predicted by MaxEnt models for present (**a**) and future (year 2070) environmental conditions based on RCP 4.5 (**b**) and RCP 8.5 (**c**) climate change scenarios. Cork oak stands are represented by *black dots*

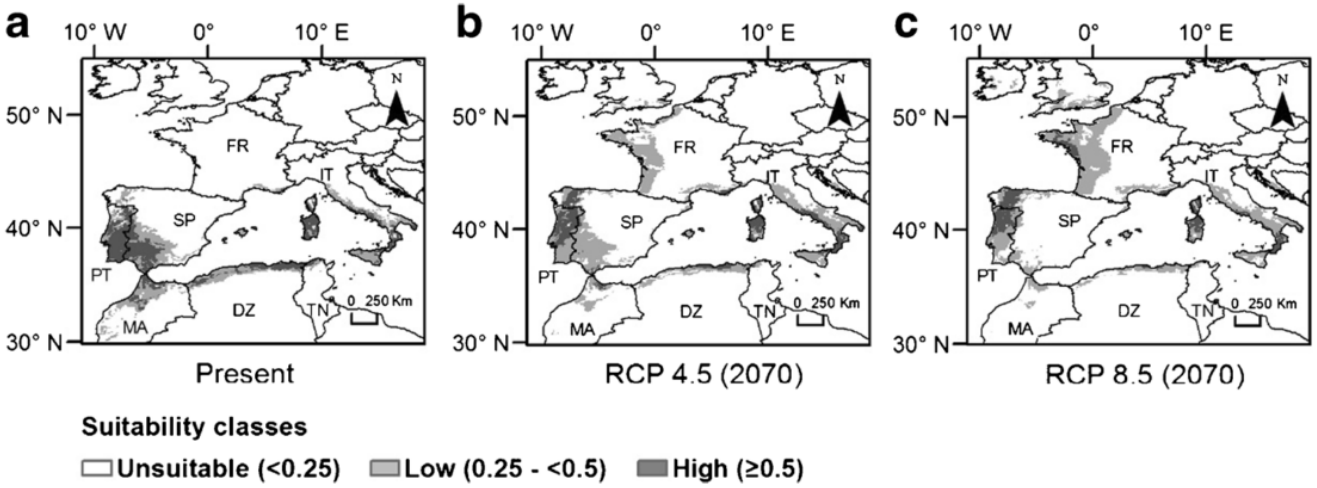


Fig. 2 Cork oak environmentally suitable areas as predicted by MaxEnt models for present (**a**) and future scenarios (year 2070) based on RCP 4.5 (**b**) and RCP 8.5 (**c**). Country codes identify the nations where the species currently occurs: Portugal (*PT*), Spain (*SP*), France (*FR*), Italy (*IT*), Morocco (*MA*), Algeria (*DZ*) and Tunisia (*TN*)

Predict the response to climate change - ENM

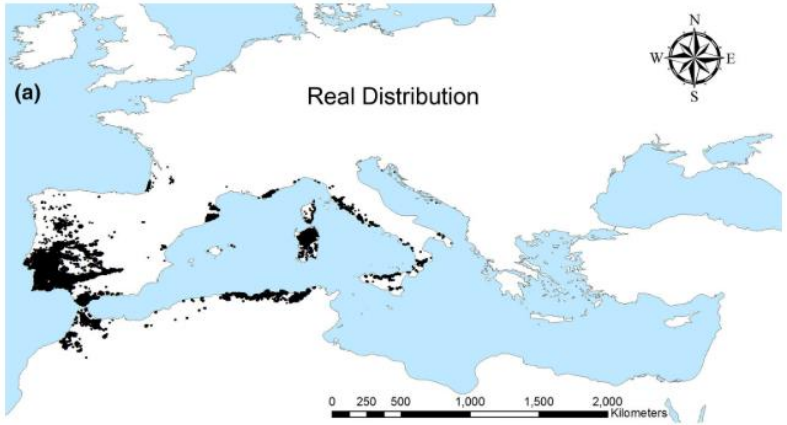
Eur J Forest Res (2017) 136:555–569
DOI 10.1007/s10342-017-1055-2

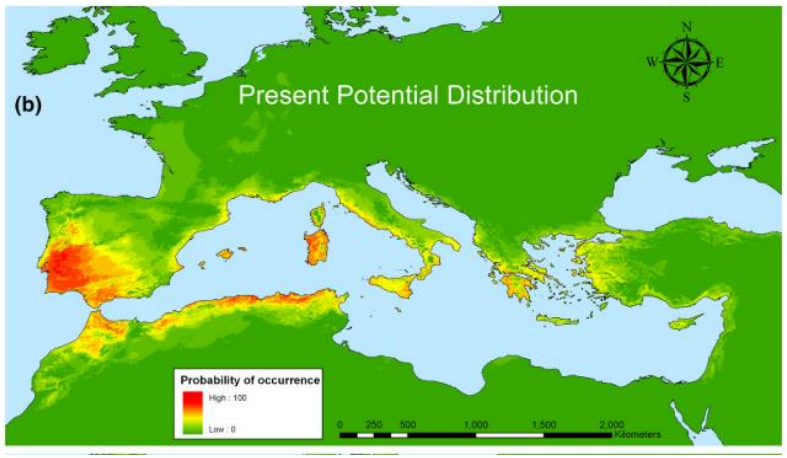
CrossMark

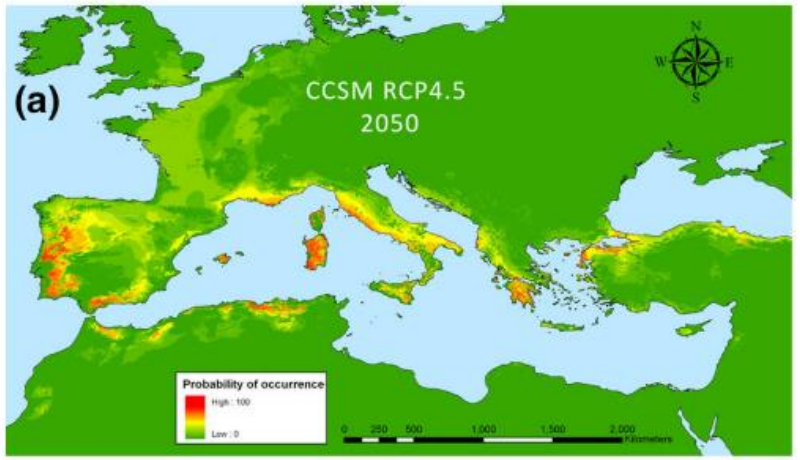
ORIGINAL PAPER

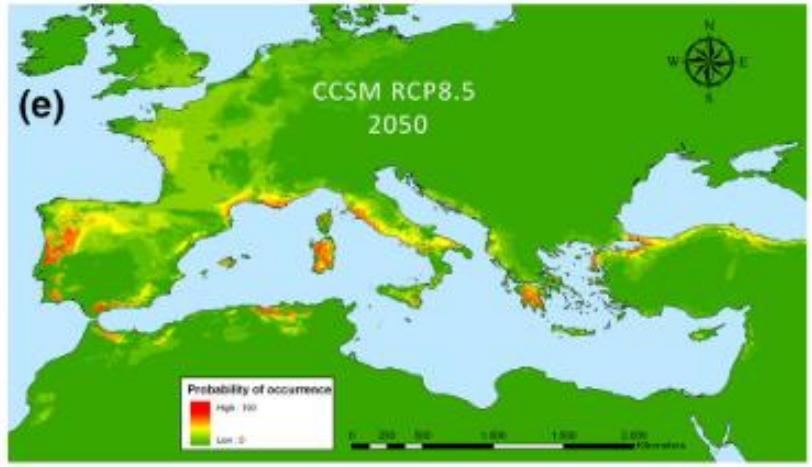
A tree species range in the face of climate change: cork oak as a study case for the Mediterranean biome

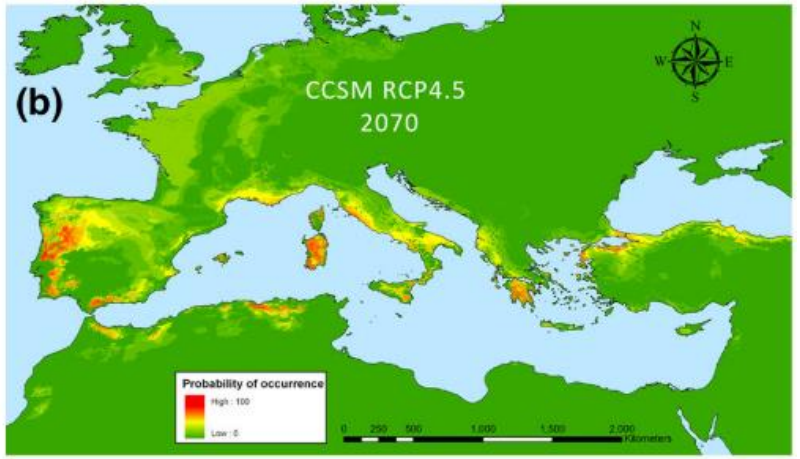
Federico Vessella[1] · Javier López-Tirado[2] · Marco Cosimo Simeone[1] · Bartolomeo Schirone[1] · Pablo J. Hidalgo[2]

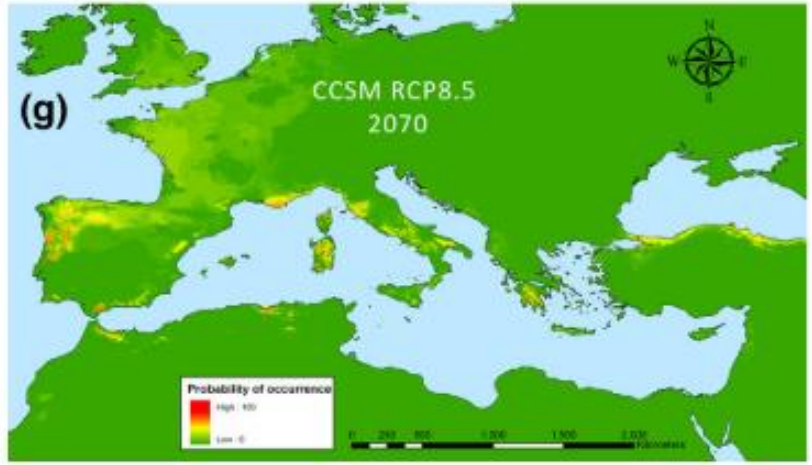












Assisted Migration or Assisted Gene Flow

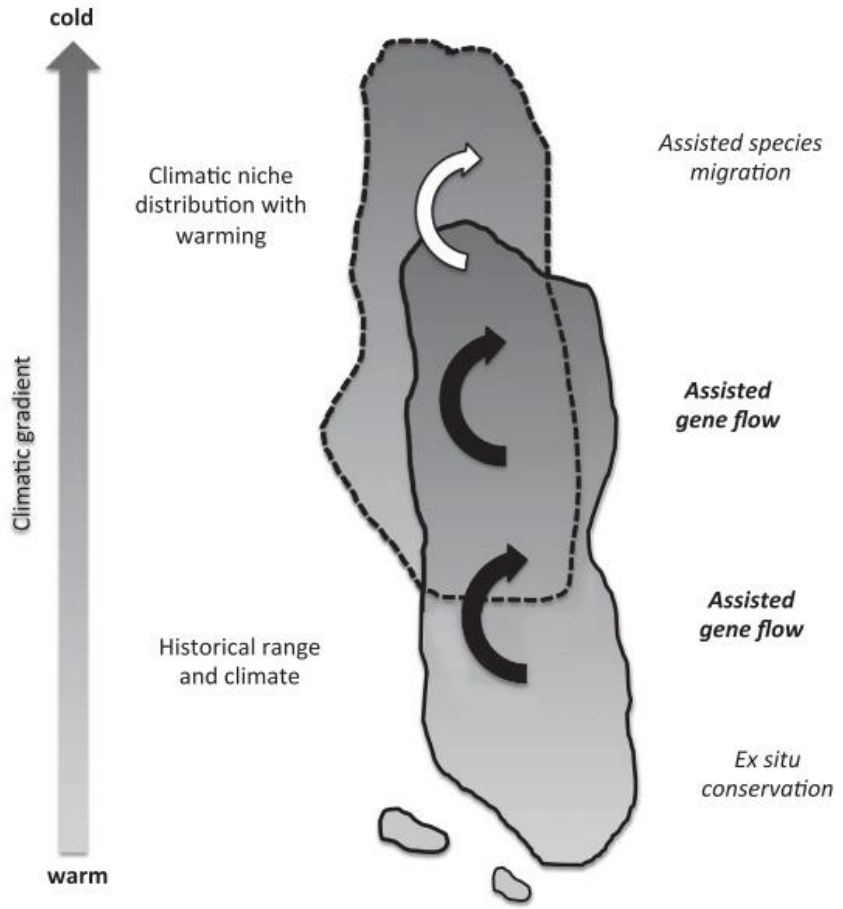


Figure 1 Schematic diagram of management options for reforestation and restoration in a changing climate. While this illustrates the northward movement of individuals, assisted gene flow may also occur along elevational or longitudinal climatic gradients.

Evolutionary Applications
Open Access
Evolutionary Applications ISSN 1752-4571

REVIEW AND SYNTHESES

Time to get moving: assisted gene flow of forest trees

Sally N. Aitken[1] and Jordan B. Bemmels[2]

1 Department of Forest and Conservation Sciences, University of British Columbia, Vancouver, BC, Canada
2 Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA

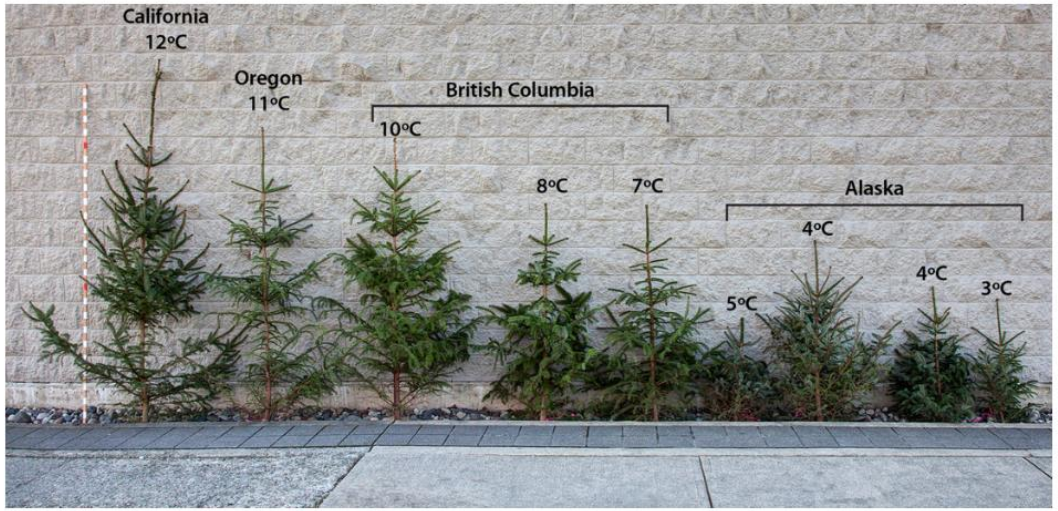


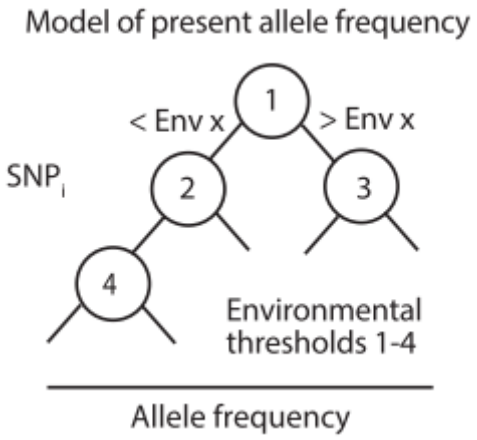
Figure 5 Provenance variation in 8-year-old *Picea sitchensis* from across the species range grown in a common garden in Vancouver, BC, Canada. For illustrative purposes, trees were selected from each population that were closest to the provenance mean. The state or province of origin and the provenance mean annual temperature are indicated. This species shows the steepest cline of all species for height at age 2 in Figure 4(A). The population from Kodiak Island, AK (Fig. 3) has poorer growth than expected for an MAT of 5°C, likely due to inbreeding (Mimura and Aitken 2007b).

Questions?

Population structure in *Quercus suber* L. revealed by nuclear microsatellite markers

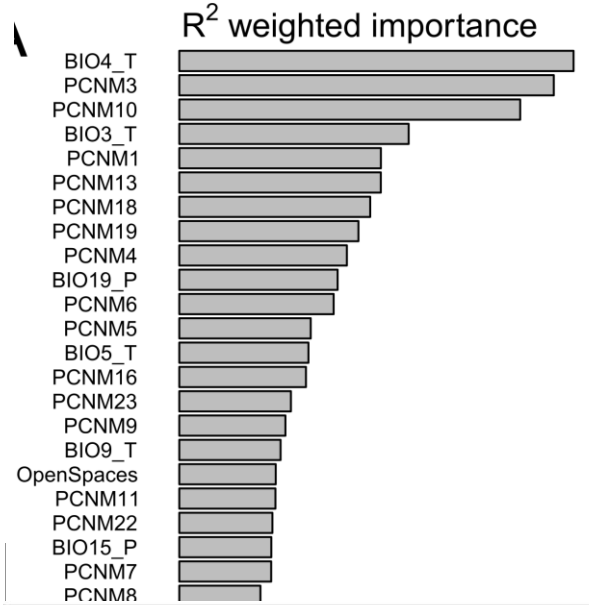
Filipe Sousa[1], Joana Costa[1,2], Carla Ribeiro[1], Marta Varandas[3], Francisco Pina-Martins[1,4], Fernanda Simões[3], José Matos[3], Maria Glushkova[5], Célia Miguel[6,7], Maria Manuela Veloso[3], Margarida Oliveira[8], Cândido Pinto Ricardo[8], Dora Batista[1,9] and Octávio S. Paulo[1]

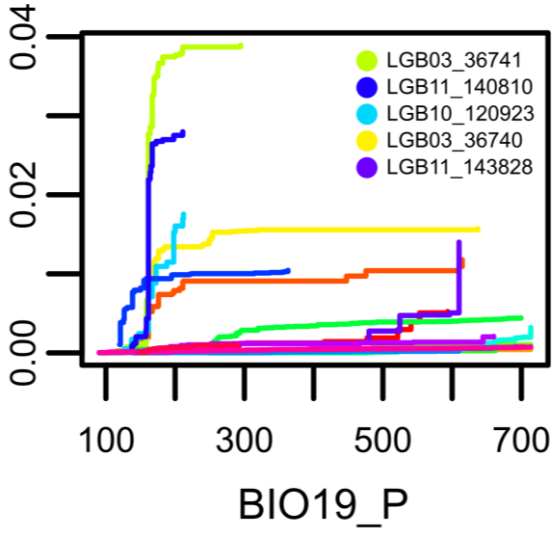
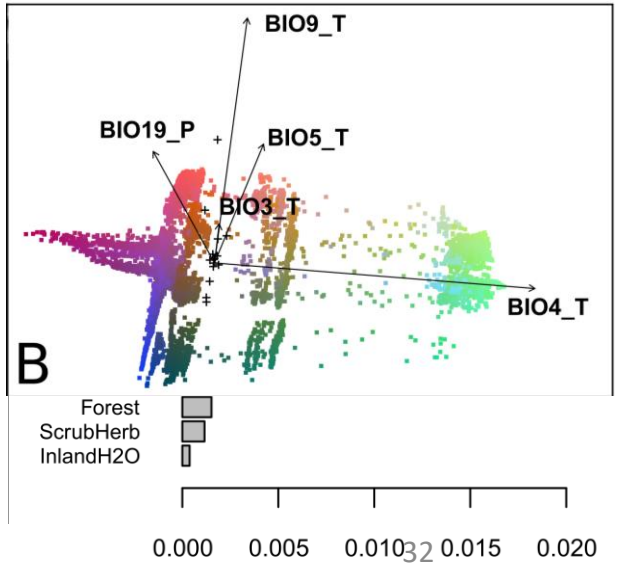
Gradient Forest GF



GF uses three primary outputs from random forests to model compositional turnover directly, including:

(1) the overall goodness of fit (R2) of the forest for each response (in this case, responses are individual SNPs),
(2) the accuracy importance for each environmental predictor within a forest and
(3) the importance of that predictor at a given split value "split importance" in determining changes in SNP frequency in a particular tree.



Gradient forest is non-parametric method that use "machine-learning regression tree approach" known as Random Forest

GF method construct a monotonic turnover function for each SNP and an aggregate (for all SPNs) turnover function across all SNPS using a strategy of "predict, then assemble"