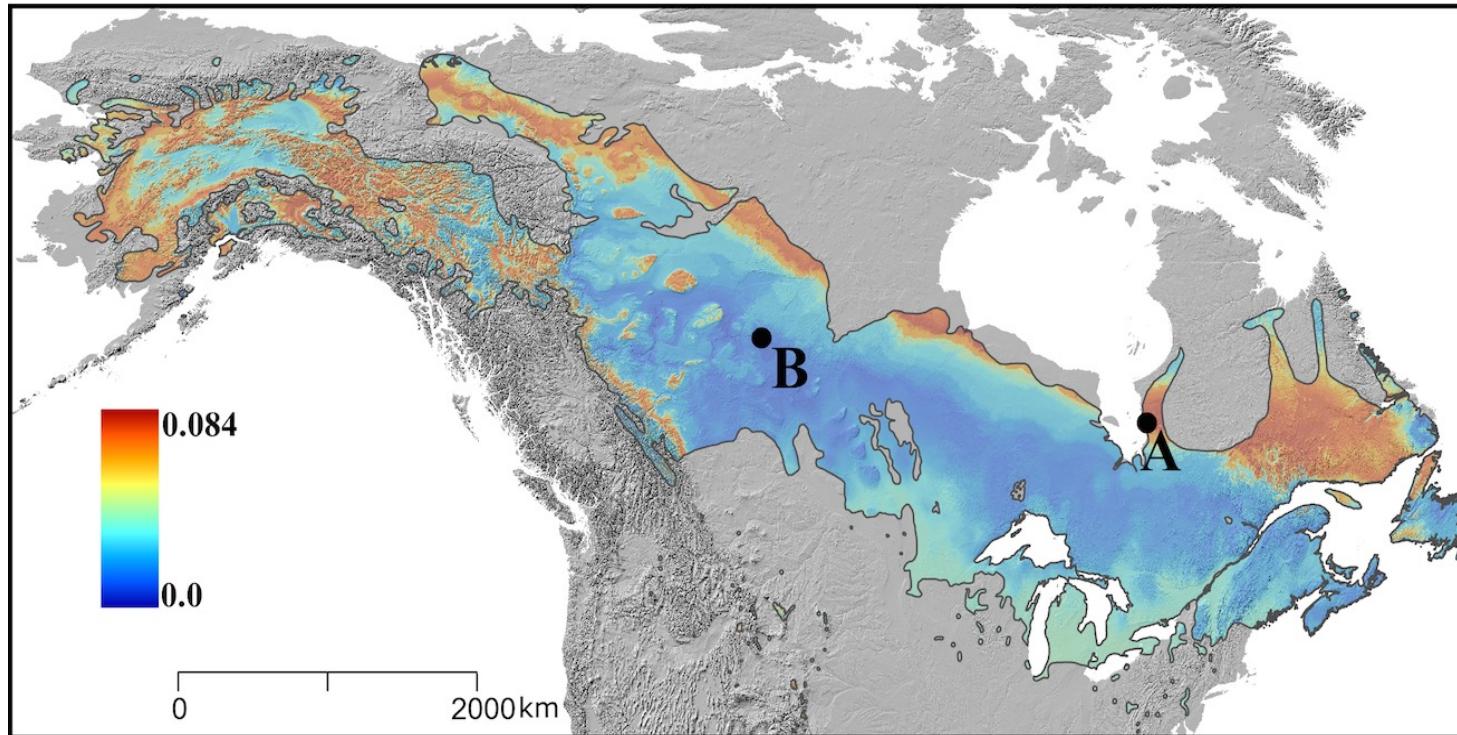


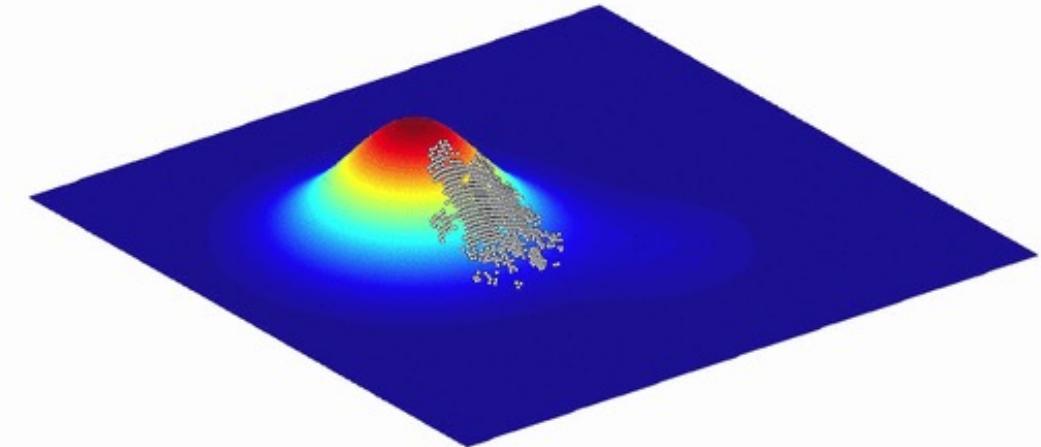
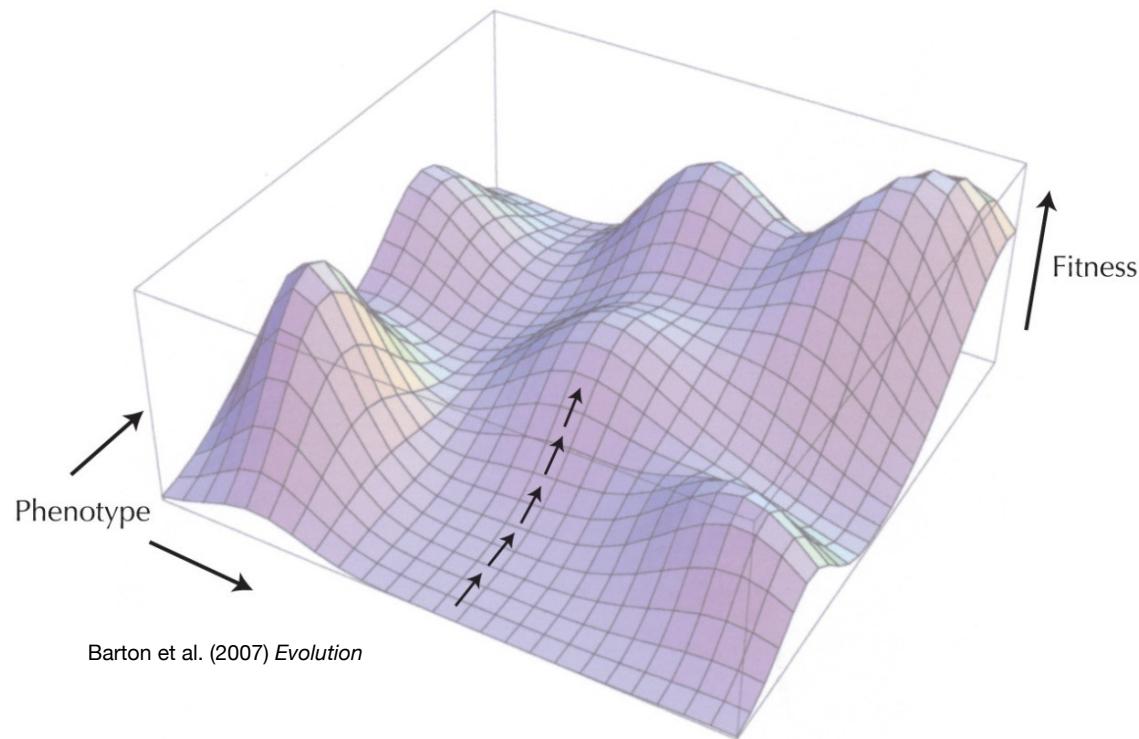
More than just pretty maps – ideas and strategies for validating genomic offset predictions



Stephen Keller
Department of Plant Biology
University of Vermont
Stephen.Keller@uvm.edu



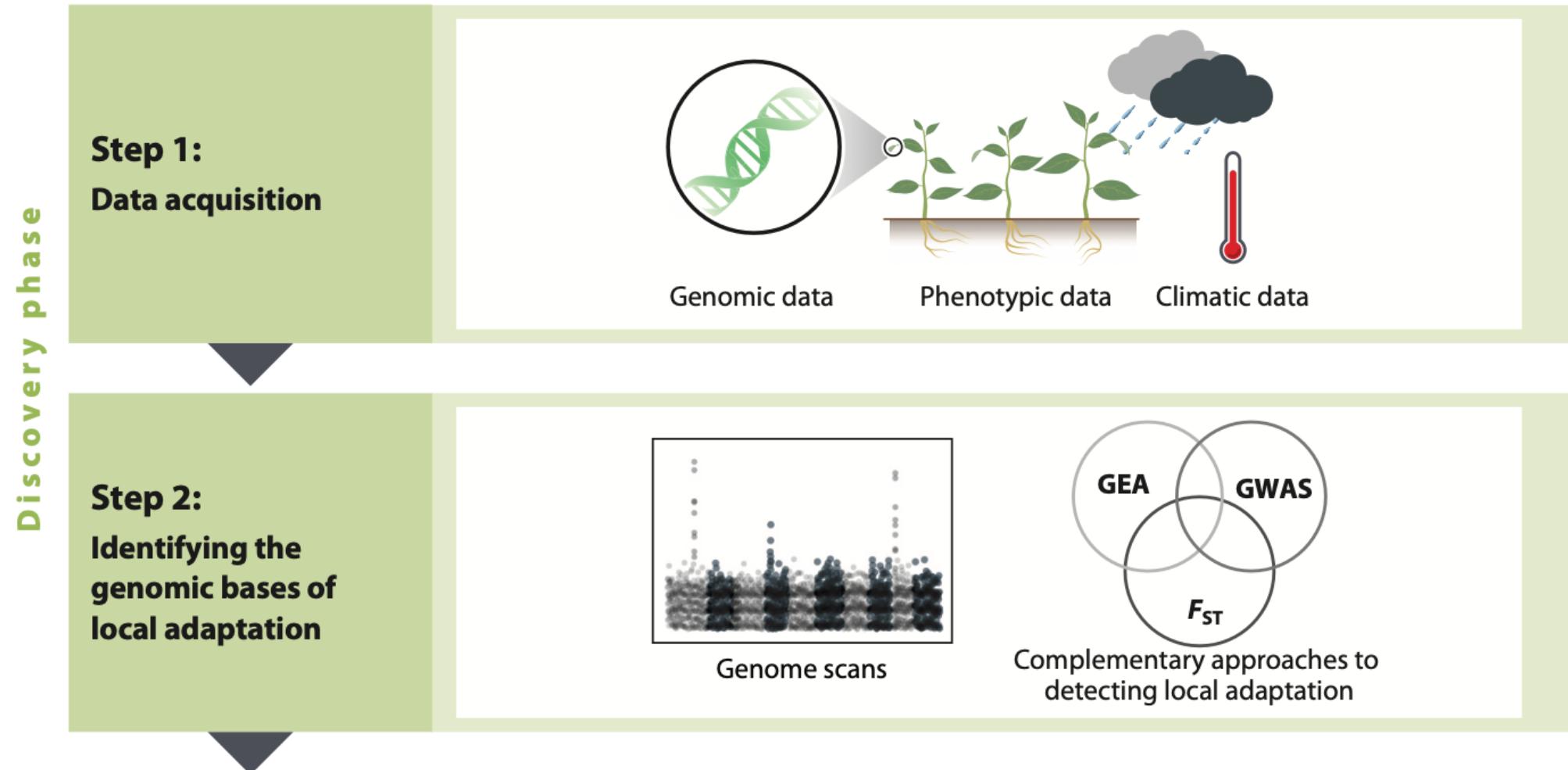
Today's adaptations may become tomorrow's maladaptations



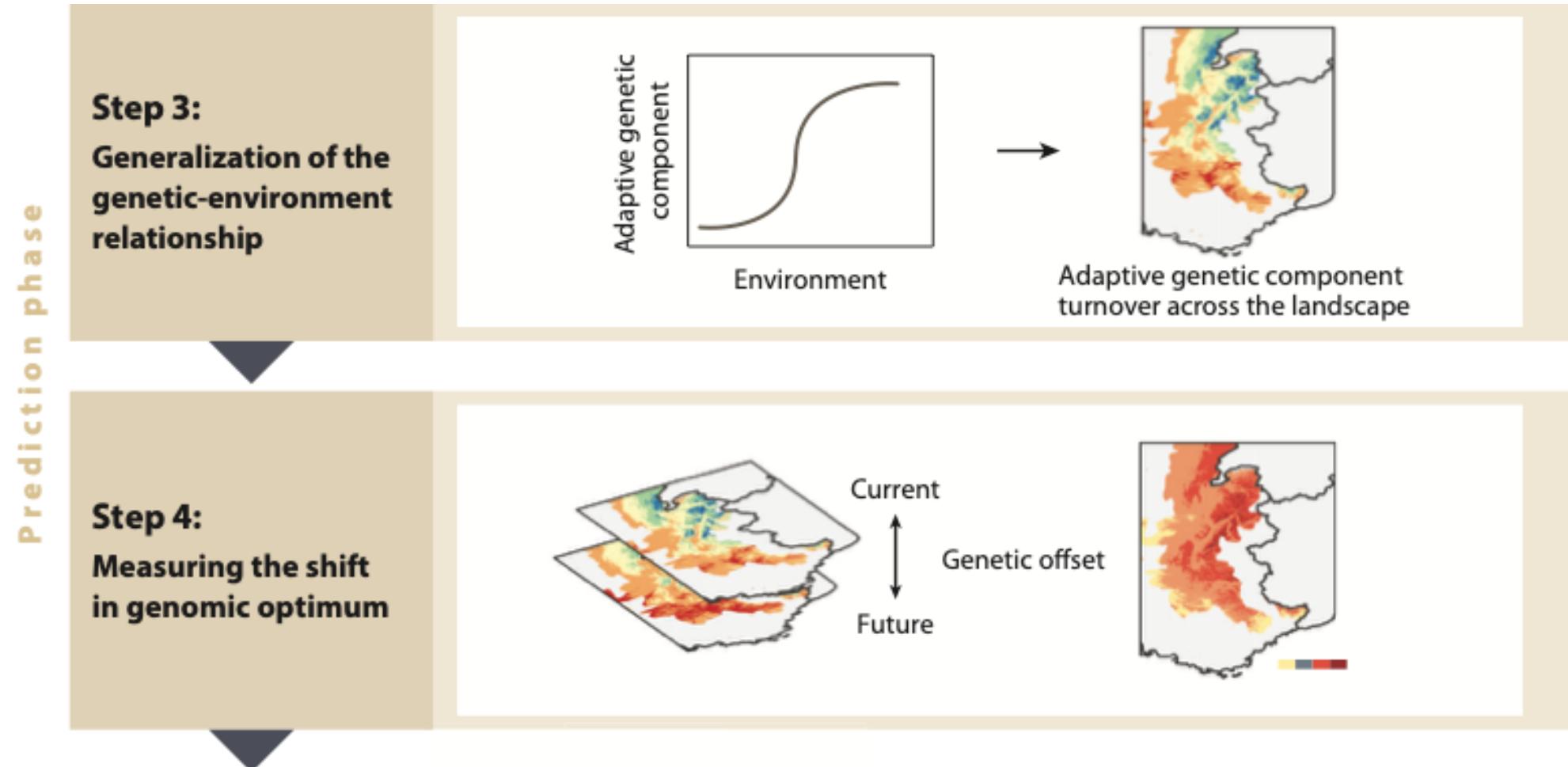
Randy Olson and Bjørn Østman, CC BY-SA 3.0
<https://creativecommons.org/licenses/by-sa/3.0>

Can we use genomic regions under **current local adaptation** to help ***predict*** future maladaptation under changing environments?

First, discover genomic regions with evidence of local adaptation



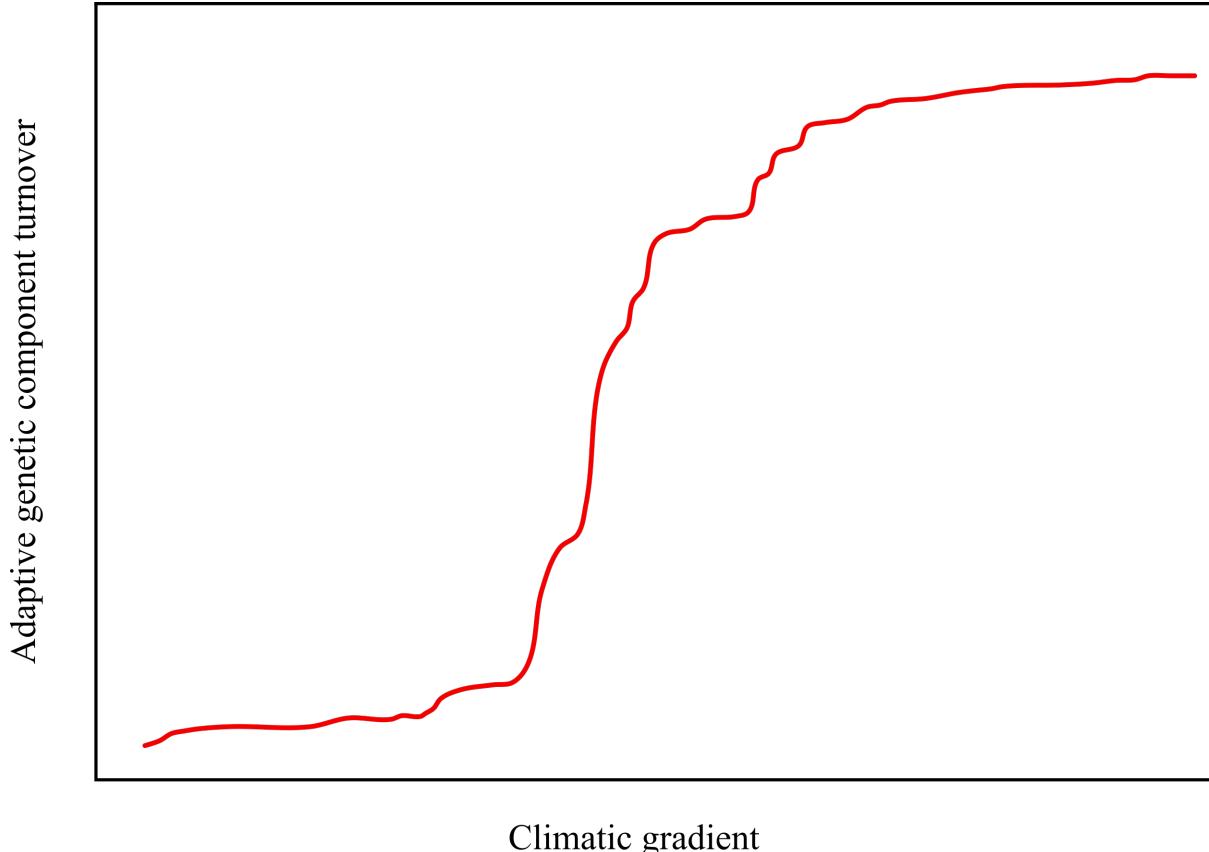
Then, determine which environmental gradients are most important to explaining adaptive allele frequencies



Can genomics *predict* maladaptation under changing environments?

The concept:

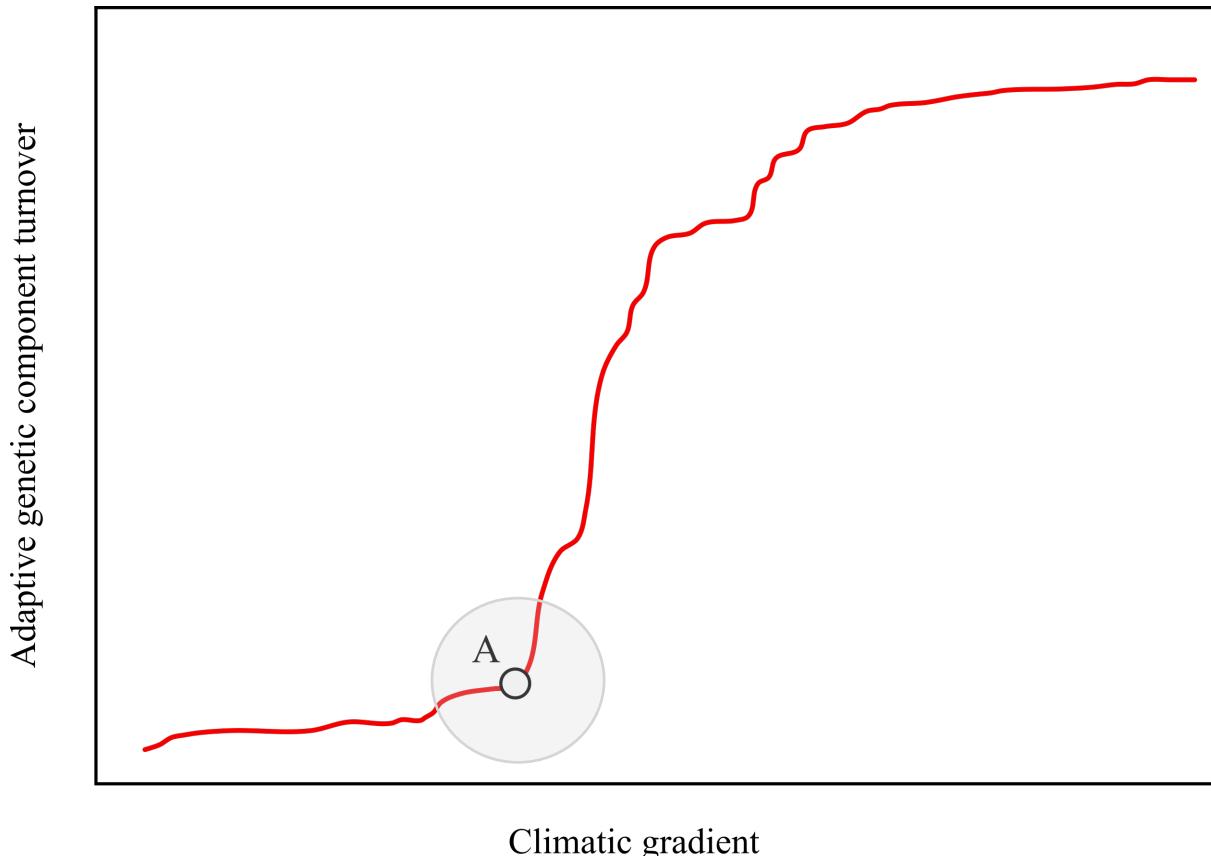
- Allele frequencies clines along climate gradients reflect local selection



Can genomics *predict* maladaptation under changing environments?

The concept:

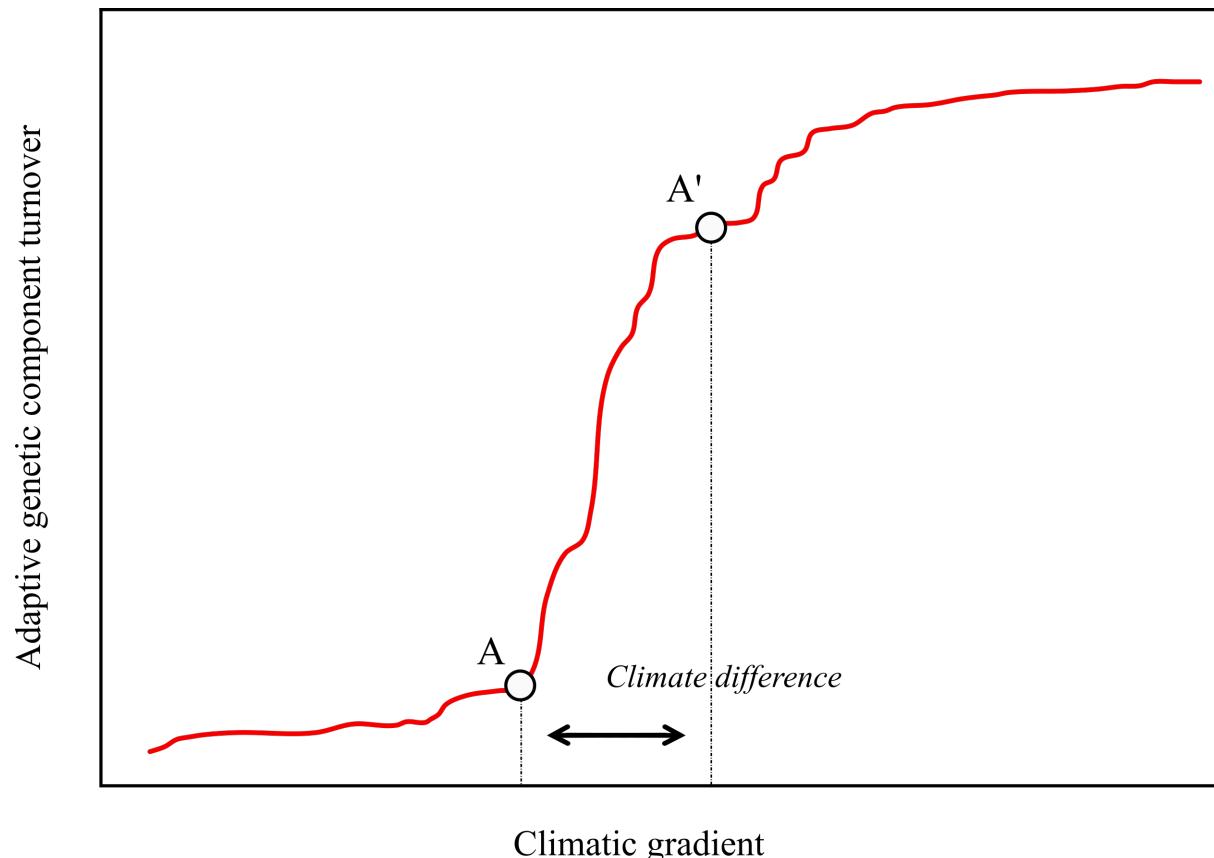
- Populations occupy different positions along the genetic-environment association



Can genomics *predict* maladaptation under changing environments?

The concept:

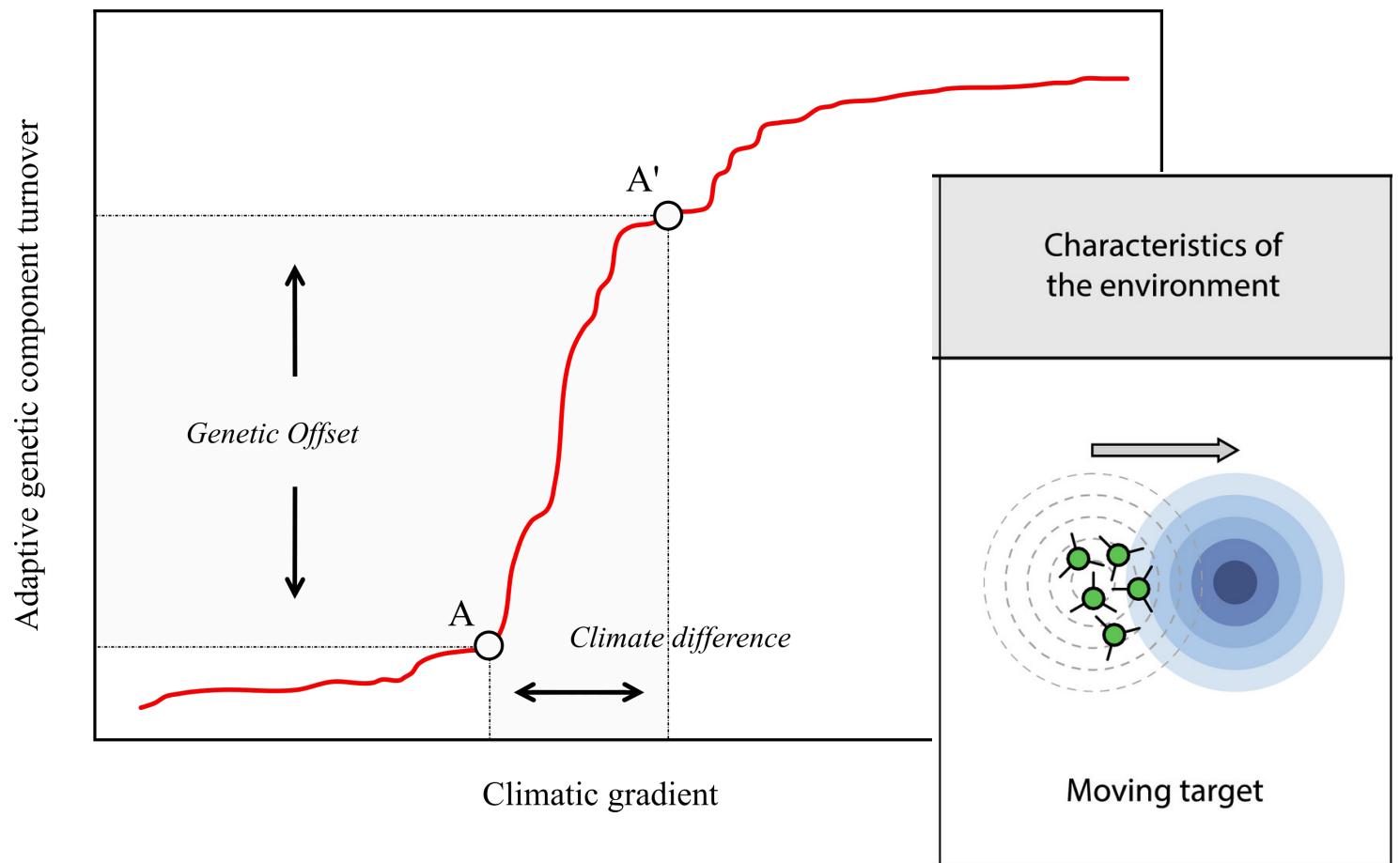
- Climate change can shift a population along the climatic gradient



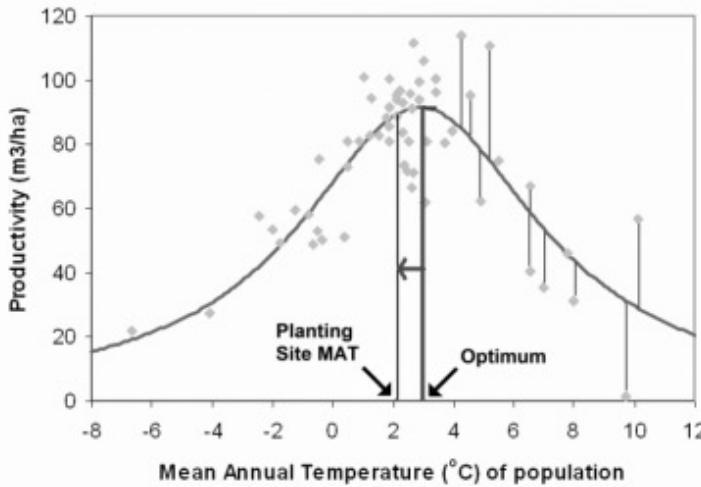
Can genomics *predict* maladaptation under changing environments?

The concept:

- Estimate the distance for a set of (putatively) adaptive loci between the current and future/changed environment
→ “*Genetic Offset*”



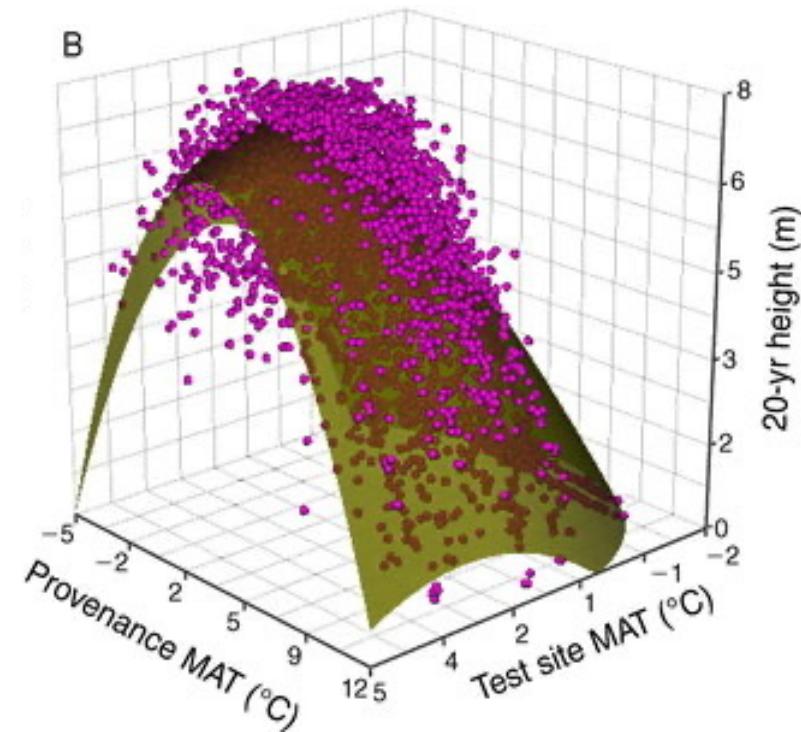
Genetic offset is conceptually related to climate transfer distance in forestry provenance trials



Ecological Applications, 20(1), 2010, pp. 153–163
© 2010 by the Ecological Society of America

Integrating environmental and genetic effects to predict responses
of tree populations to climate

TONGLI WANG,^{1,3} GREGORY A. O'NEILL,² AND SALLY N. AITKEN¹



Global Change Biology (2007) 13, 1441–1454, doi: 10.1111/j.1365-2486.2007.01385.x

Genetic maladaptation of coastal Douglas-fir seedlings to future climates

J. BRADLEY ST CLAIR* and GLENN T. HOWE†

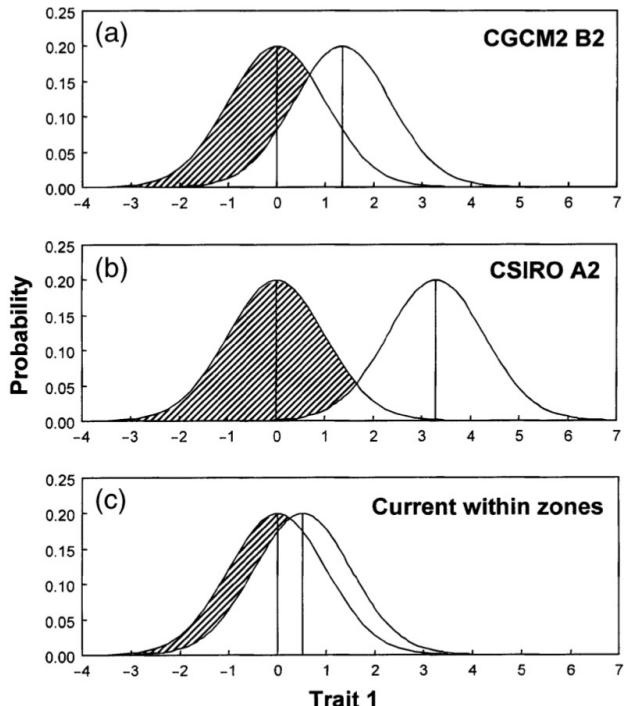
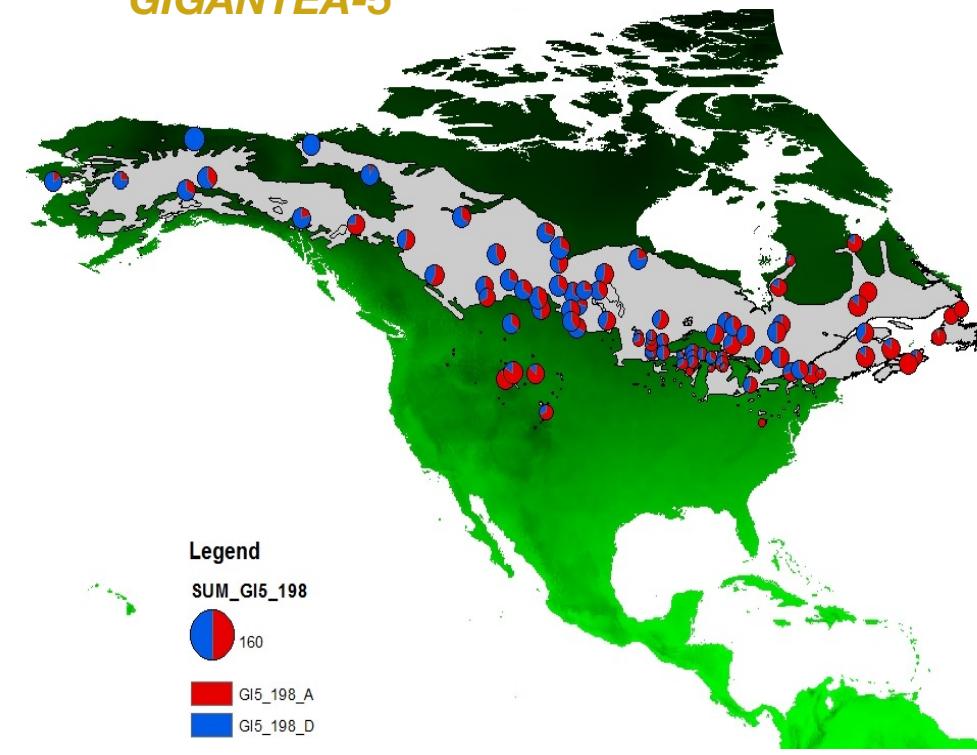


Fig. 1 Relative risk of maladaptation associated with current populations in future (2077–2099) climates for trait 1. In each

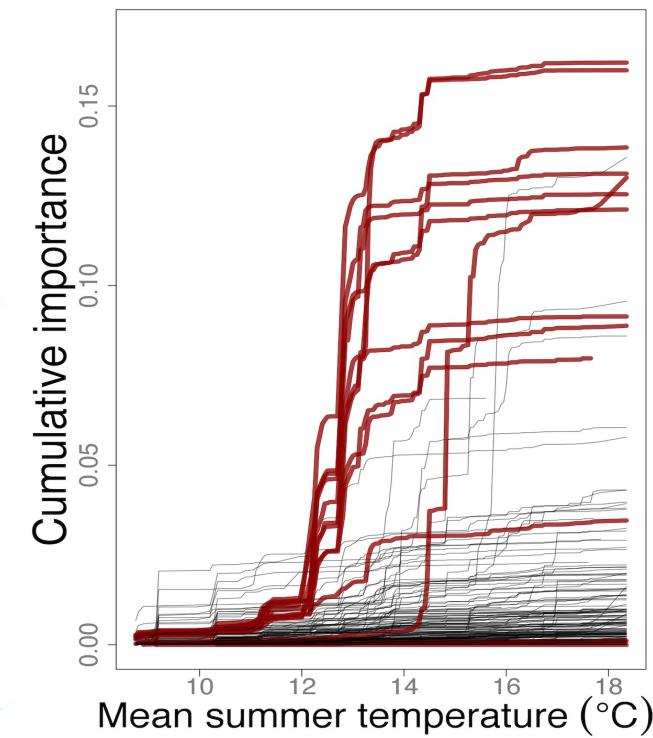
Gradient Forest Cumulative Importance (CI) curves

Ex: Balsam Poplar flowering time genes

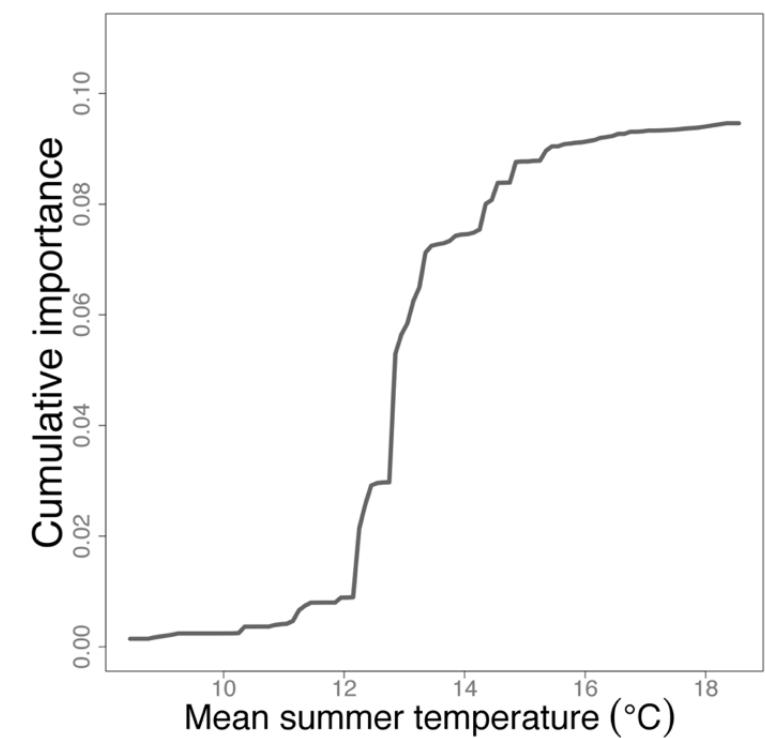
GIGANTEA-5



Locus-wise CI

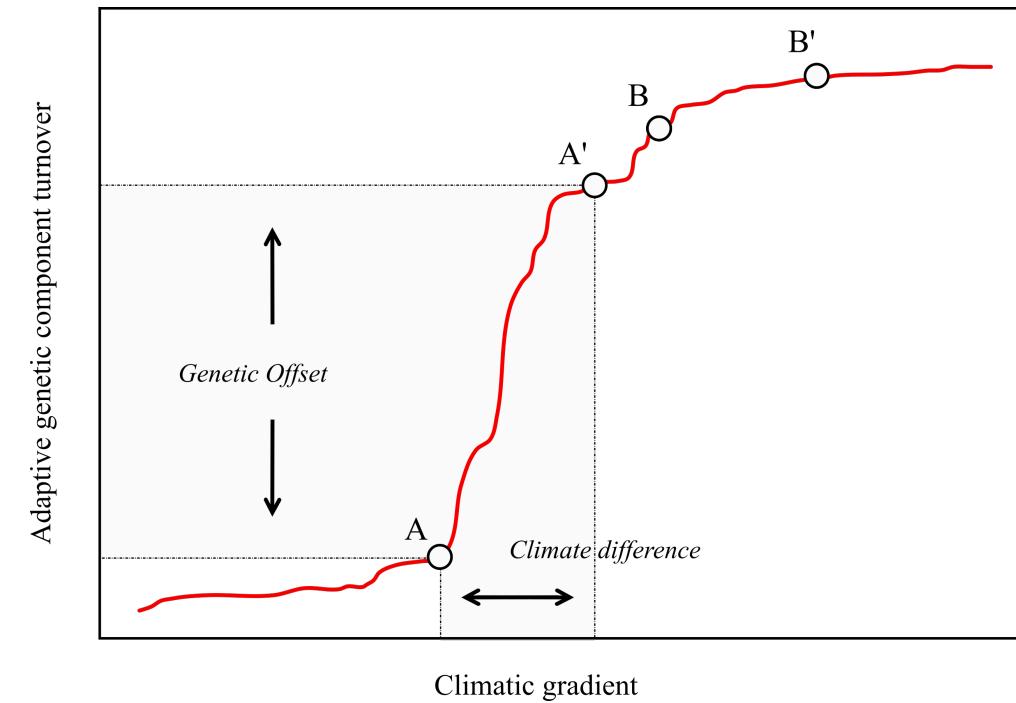
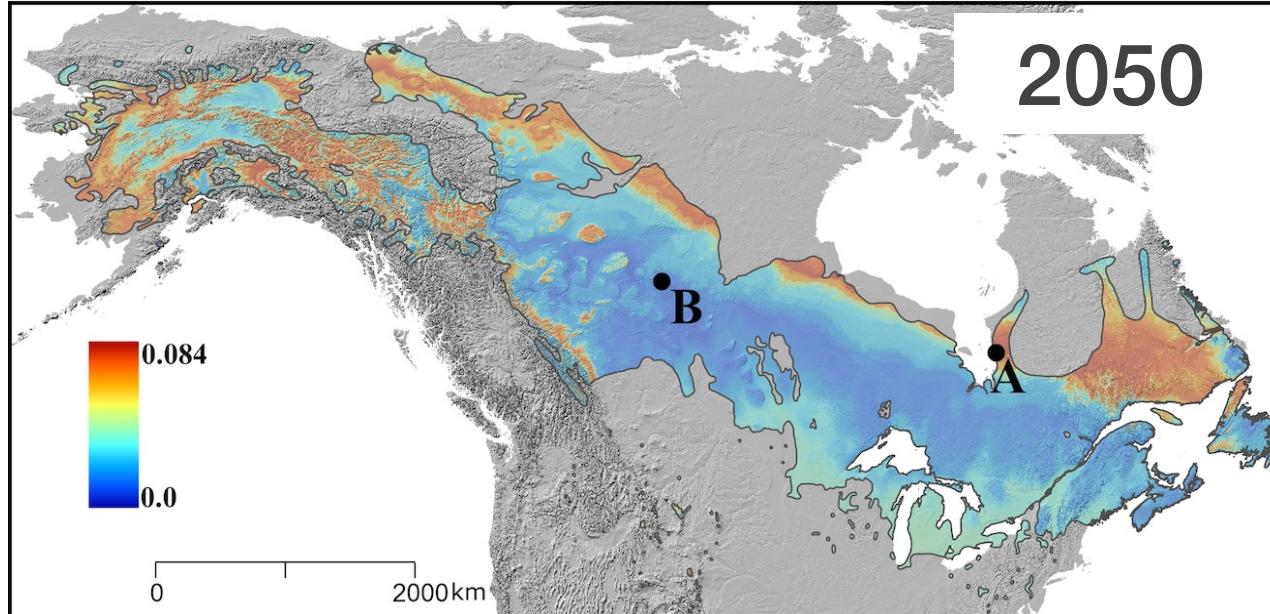


Multi-locus weighted CI



Genetic Offset from Gradient Forest (GF)

$$GFOffset_{A \rightarrow A'} = \sqrt{\sum_j (CI_{Aj} - CI_{A'j})^2}$$



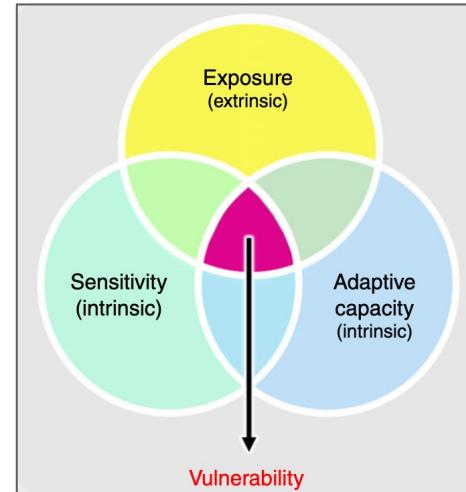
A Clarification on Terminology

“*Genomic Offset*” = *Genetic Offset* scaled up to genome-wide data

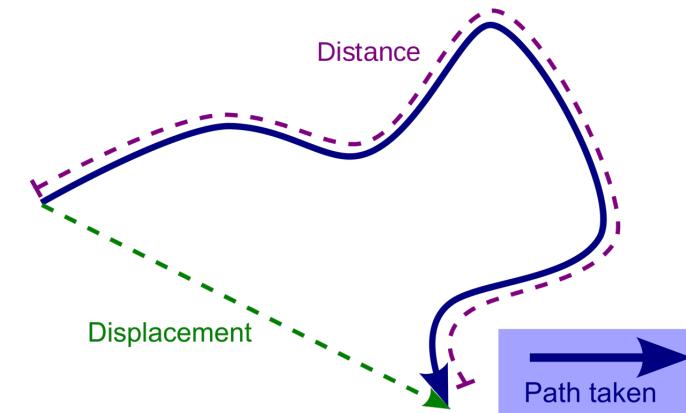
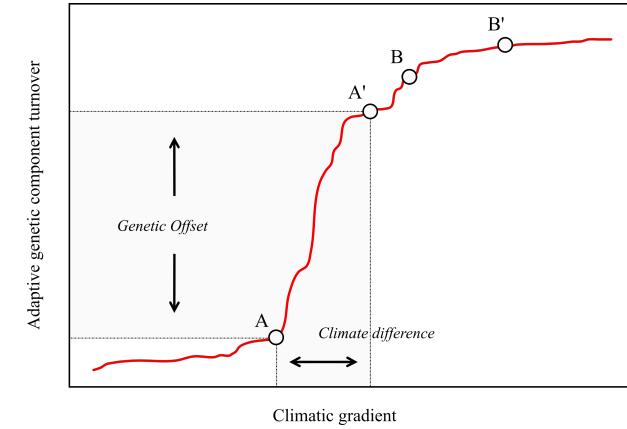
“*Genomic vulnerability*” is synonymous with Genomic Offset

...but, genomic offsets only describe 2 of the 3 components of climate change vulnerability

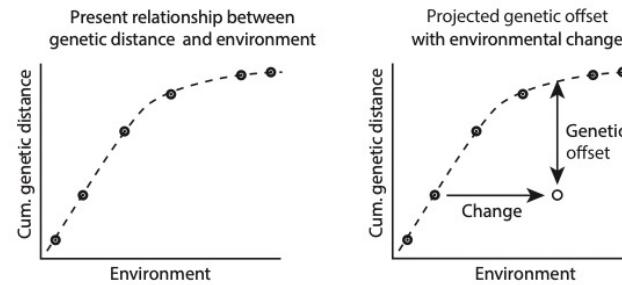
“*Offset*” (geometry) the position of a point or a particle in reference to an origin or to a previous position



Foden et al. (2019) *WIREs Clim. Ch.*

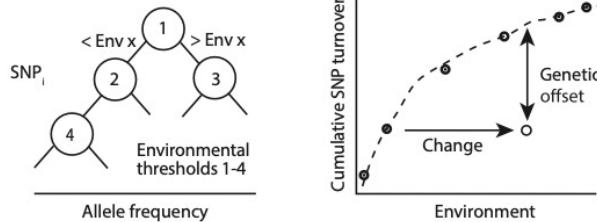


(b) Genetic offset
b1 Generalized dissimilarity modelling



b2 Gradient forest

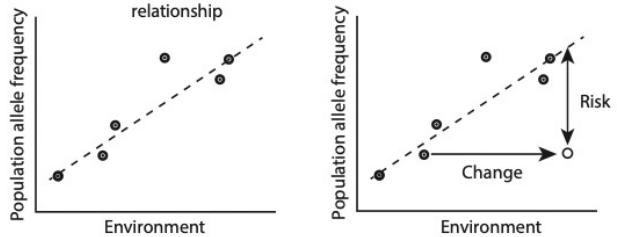
Model of present allele frequency



GDM

(c) Risk of non-adaptedness

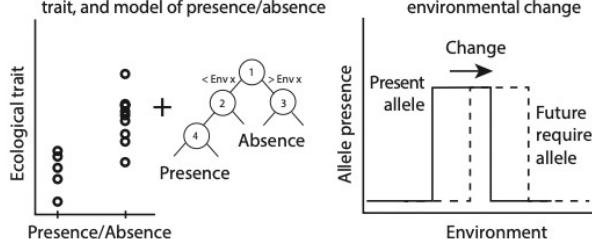
Present environment – allele frequency relationship



GF

(d) Random forest ecological niche model

Identify SNPs associated with ecological trait, and model of presence/absence



RF-ENM

A growing number of approaches estimate Genomic Offset

...and recently:

Latent Factor Mixed Models (LFMMs)

Gain and François et al. (2021) Mol. Ecol. Res.

Redundancy Analysis (RDA)

Capblancq and Forester (2021) Met. in Ecol. & Evol.

Spatial Areas of Genotype Probability (SPAG)

Rochat et al. (2021) Div. & Distr.

Studies estimating Genomic Offset are accumulating

Species	Type	Spatial scale	Data	Adaptive component identification	Predictive model	Validation procedure	Reference
<i>Populus balsamifera</i>		North America	Genomic (targeted genotyping of candidate genes)	F_{ST} (Arlequin, BayeScan), GEA (Bayenv), GPA	Gradient forest and GDM	None	Fitzpatrick & Keller 2015
<i>Arabidopsis thaliana</i>		Europe	Genomic (whole genome)	GWAS (EMMAX25, BSLMM28)	Ecological niche modeling on alleles	None	Exposito-Alonso et al. 2018
<i>Populus tremula</i>		Sweden	Genomic (whole genome)	GEA (LFMM)	GDM	None	Ingvarsson & Bernhardsson 2018
<i>Quercus rugosa</i>		Mexico	Genomic (GBS)	F_{ST}	Gradient forest	None	Martins et al. 2018
<i>Empidonax traillii</i>		North America	Genomic (RADseq)	GEA (LFMM)	Gradient forest	Correlation with population historical trends	Ruegg et al. 2018
<i>Setopagis petechia</i>		North America	Genomic (RADseq)	GEA (LFMM)	Gradient forest	Correlation with population historical trends	Bay et al. 2018
<i>Quercus suber</i>		Western Mediterranean	Genomic (GBS)	F_{ST} (BayeScan), SelEstim	RONA	None	Pina-Martins et al. 2019
<i>Myotis crypticus</i> and <i>Myotis escalerai</i>		Western Europe	Genomic (RADseq)	GEA (LFMM, RDA)	Ecological niche modeling separating two extreme adaptive groups	None	Razgour et al. 2019
<i>Eucalyptus melliodora</i>		Australia	Genomic (GBS)	None	GDM using the complete genomic data set	None	Supple et al. 2018
<i>Arabidopsis thaliana</i>		Europe	Genomic (whole genome) Phenotypic (common garden)	GWAS (LM-GEMMA)	GWES	Model evaluation procedure	Exposito-Alonso et al. 2019
<i>Zea mays</i>		Mexico	Genomic (published data)	GEA (Bayenv, Bayescenv)	Gradient forest	None	Aguirre-Liguori et al. 2019
<i>Fagus sylvatica</i>		French Alps	Genomic (RADseq)	GEA (RDA)	RDA	None	Capblancq et al. 2020
<i>Capra aegagrus hircus</i>		Morocco	Genomic (whole genome)	GEA (logistic regressions, SAM)	SPAG	Model evaluation procedure	Rochat & Joost 2019
<i>Mimosa acutistipula</i> and <i>Dioctria apurensis</i>		Carajás Mineral Province (Brazil)	Genomic (GBS) Phenotypic (sample traits)	GEA (RDA, LFMM), GPA (RDA, LFMM)	RDA	None	Carvalho et al. 2019
<i>Populus balsamifera</i>		North America	Genomic (targeted genotyping of candidate genes)	GEA (LFMM, Bayenv)	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; LFMM, latent factor mixed model; LM-GEMMA, linear model form of the genome-wide efficient mixed model; RADseq, restriction-site associated DNA sequencing; RDA, redundancy analysis; RONA, risk of nonadaptiveness; SAM, spatial analysis method; SPAG, spatial areas of genotype probability.

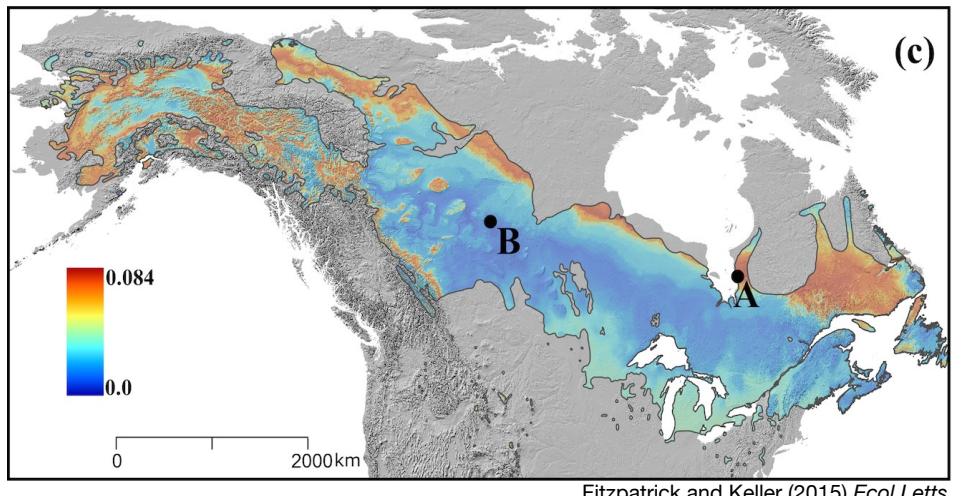
Organism type:  = Tree

 = Plant

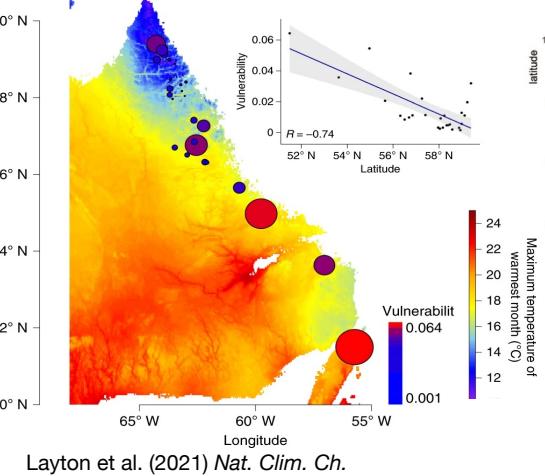
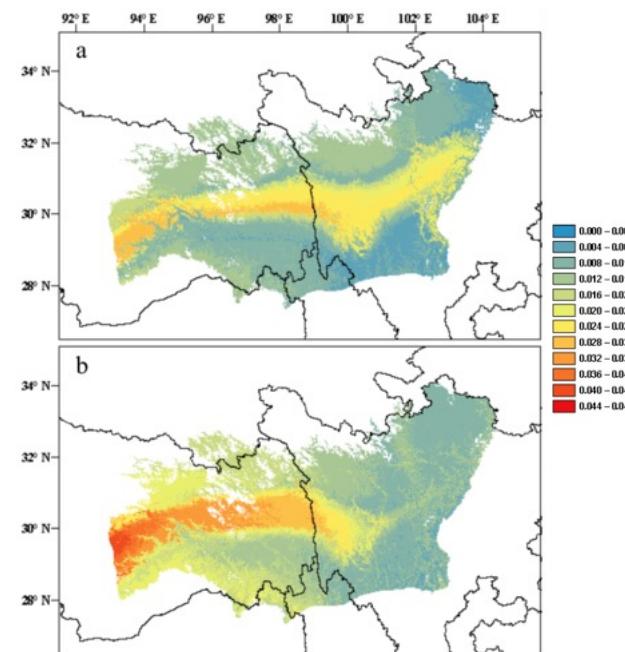
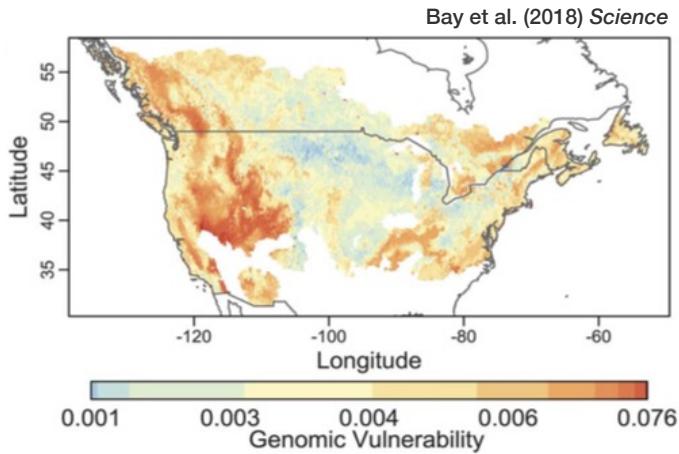
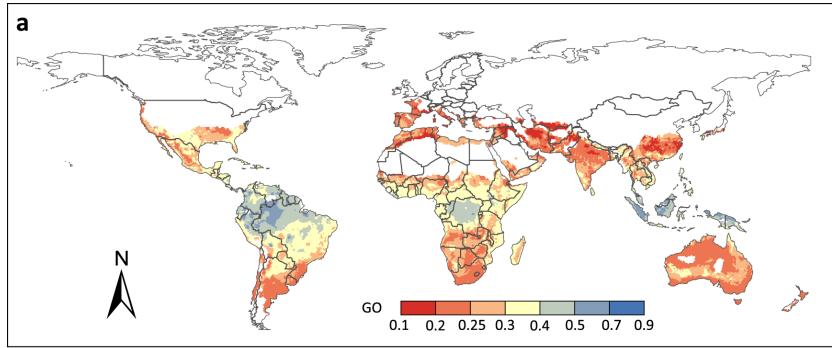
 = Domesticated mammal

 = Wild mammal

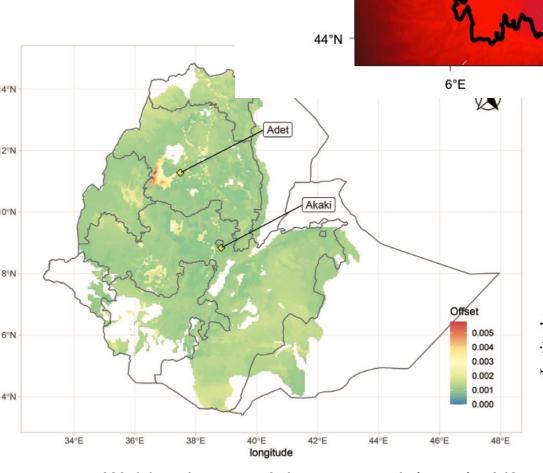
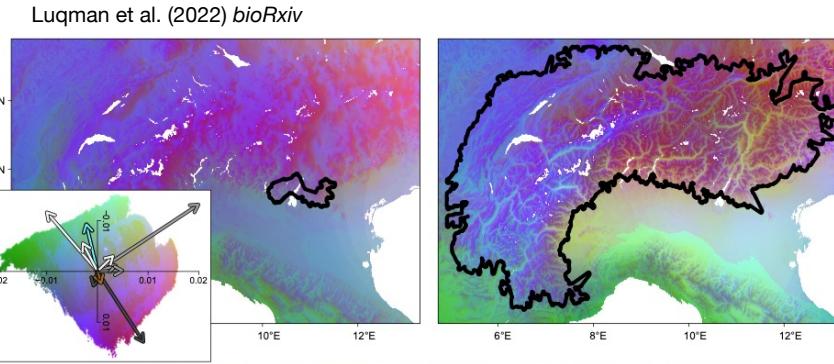
 = Bird



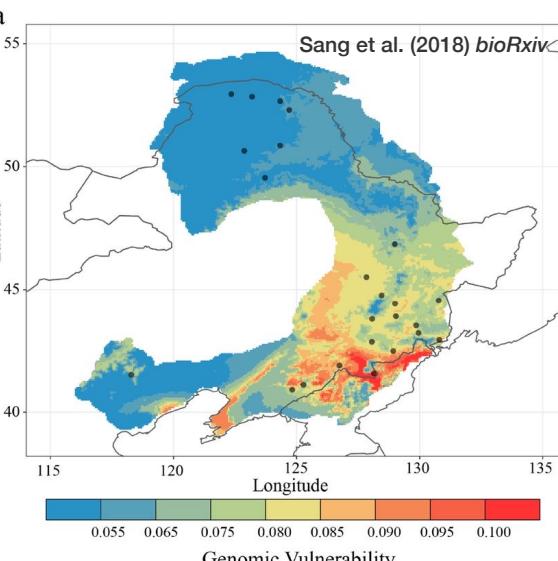
Chen et al. (2021) *Nat. Comm.*



Layton et al. (2021) *Nat. Clim. Ch.*



Woldeyohannes, Iohannes et al. (2022) *eLife*



Genomic Vulnerability

Genomic Offset shows promise, but also requires careful study, empirical support, and experimental validation

 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,*}

¹School of BioSciences, Bio21 Institute, The University of Melbourne, Parkville, VIC 3010, Australia

²School of Biological Sciences, Monash University, Clayton, VIC 3800, Australia

*Correspondence: carla.sgro@monash.edu
<https://doi.org/10.1016/j.cell.2021.02.006>

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

DOI: 10.1111/eva.13205

PERSPECTIVE

Evolutionary Applications
Open Access

WILEY

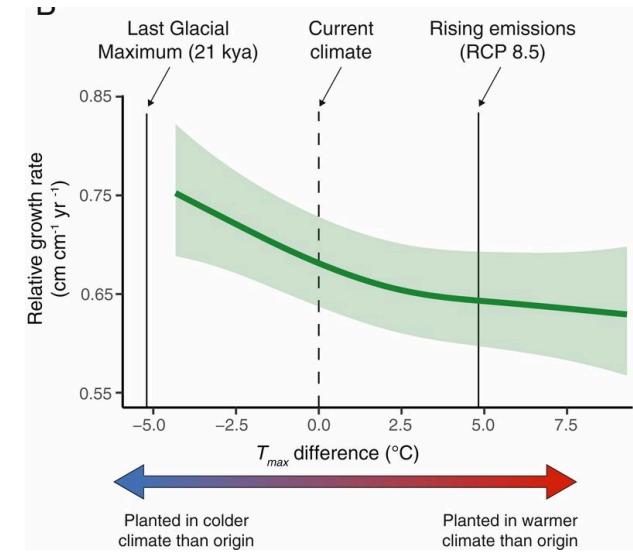
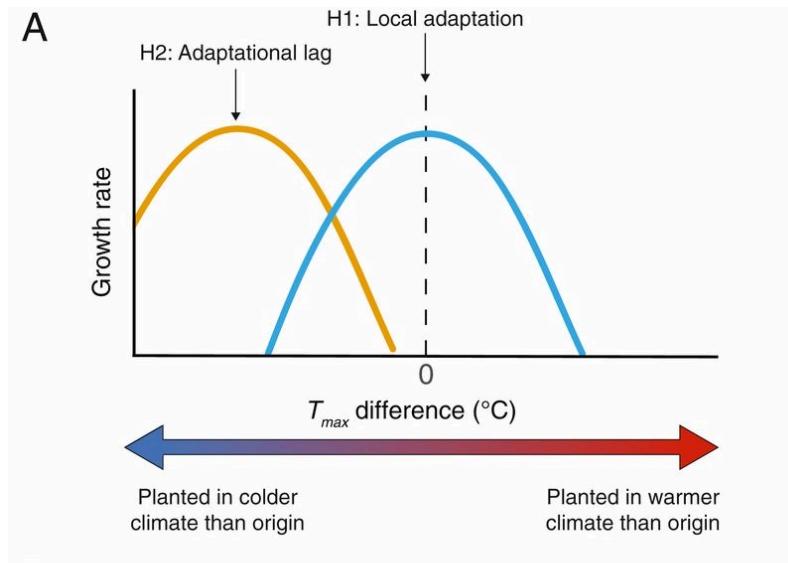
Prospects and limitations of genomic offset in conservation management

Rellstab, Dauphin, Exposito-Alonso (2021)

What assumptions being made should inspire caution?

Assumption: Populations are currently at their locally adaptive fitness peak

What about adaptation lag following range expansion?

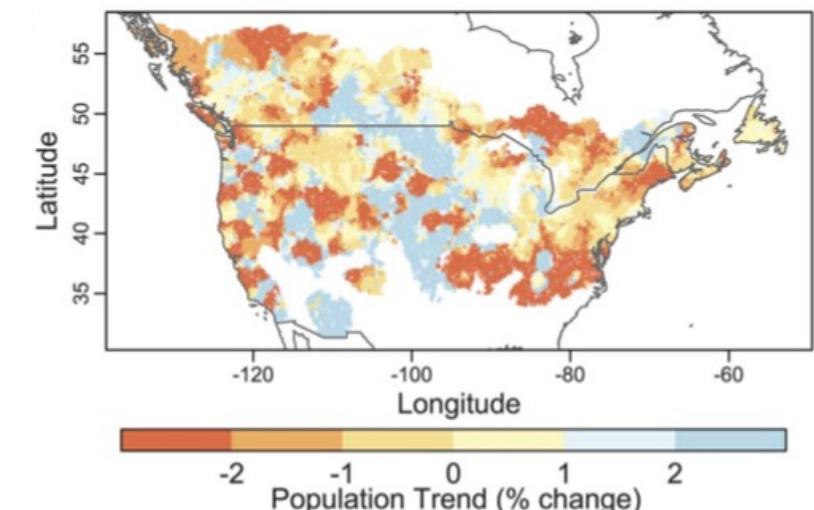
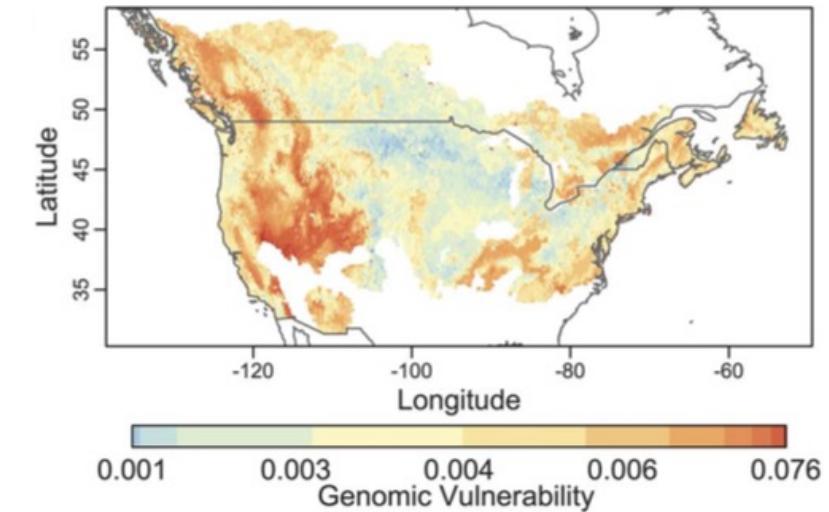


What assumptions being made should inspire caution?

Assumption: Allele frequency along environmental gradients is **due to selection**, and not drift or other neutral features of demography

Whenever possible, genome scans should employ procedures that correct for demography

Selected loci should be used to train the offset model



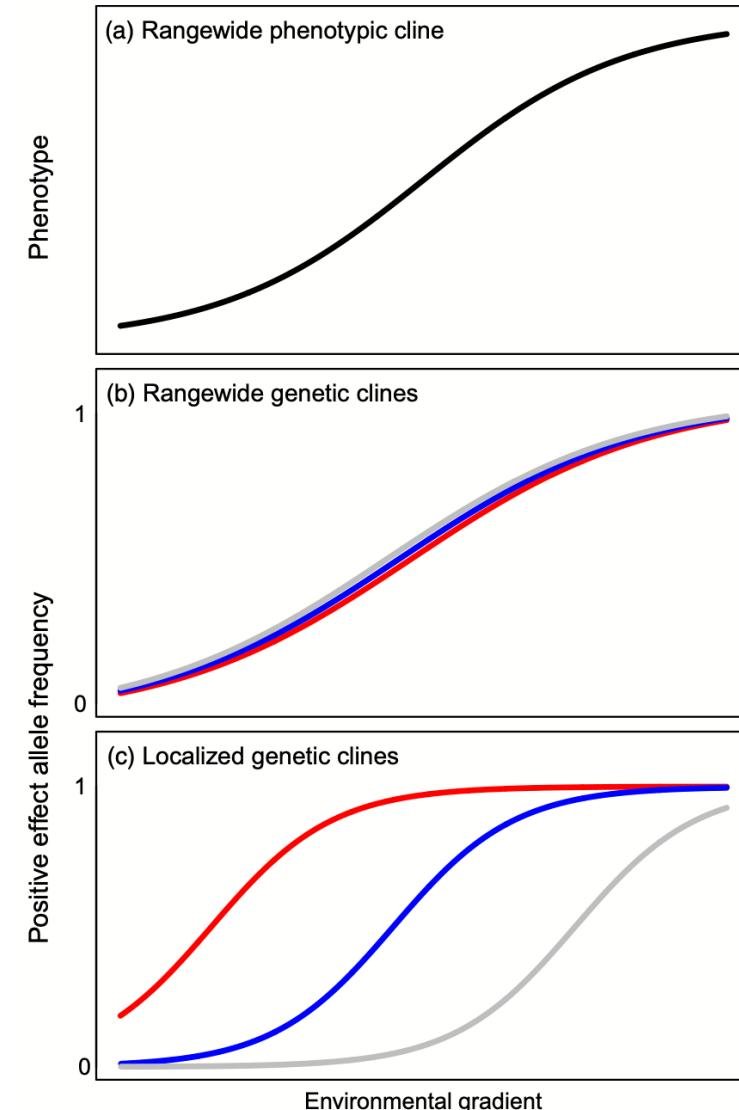
What assumptions being made should inspire caution?

Assumption: Adaptive allele frequency change is **monotonic** with respect to the environment

Do we have reason to think it might not be monotonic? Maybe...

- *(see KL's bioRxiv paper)*

Do current GEA genome scans for selection accommodate non-monotonicity? No...



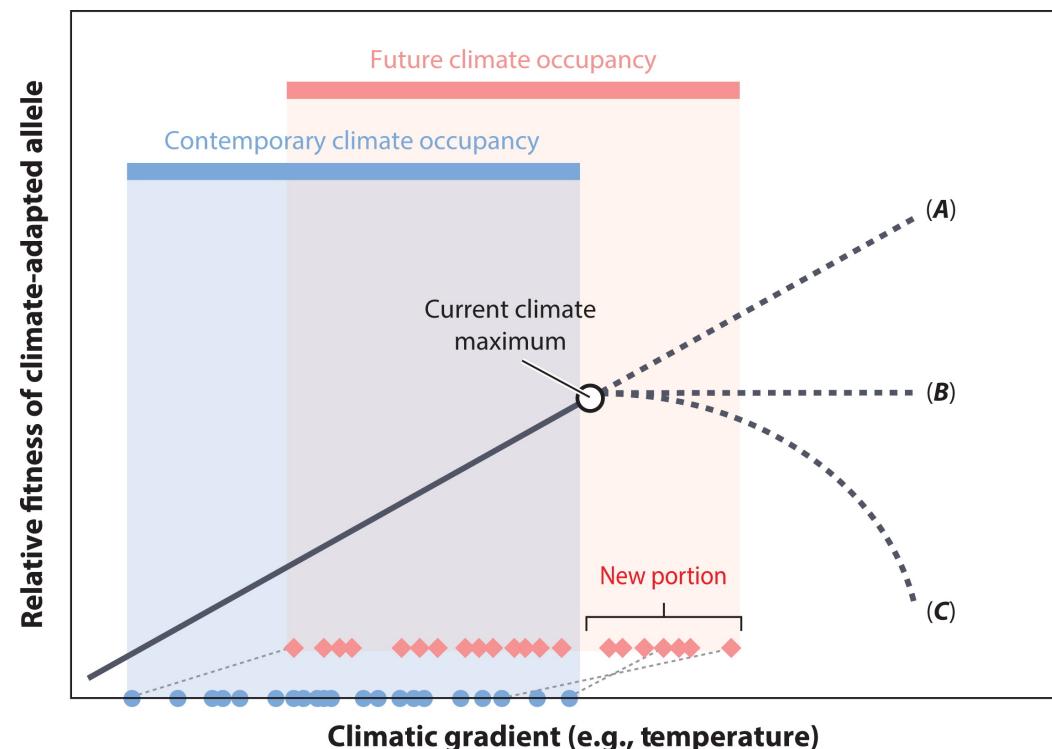
What assumptions being made should inspire caution?

Assumption: Extrapolation beyond the range of training environments is risky

(not really an assumption, but worth bringing up...)

**Extrapolation beyond the training data
is dangerous**

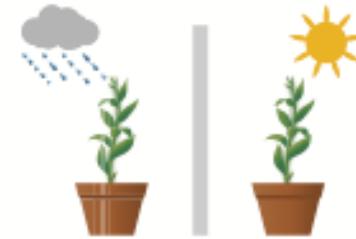
**Could address novel environments
experimentally with field
manipulations, growth chambers, etc**



How can we tell if the predictions are any good?

Validation

Step 5: **Validation of the predictions**



Testing the decrease in fitness associated with an increasing genetic offset in controlled environments



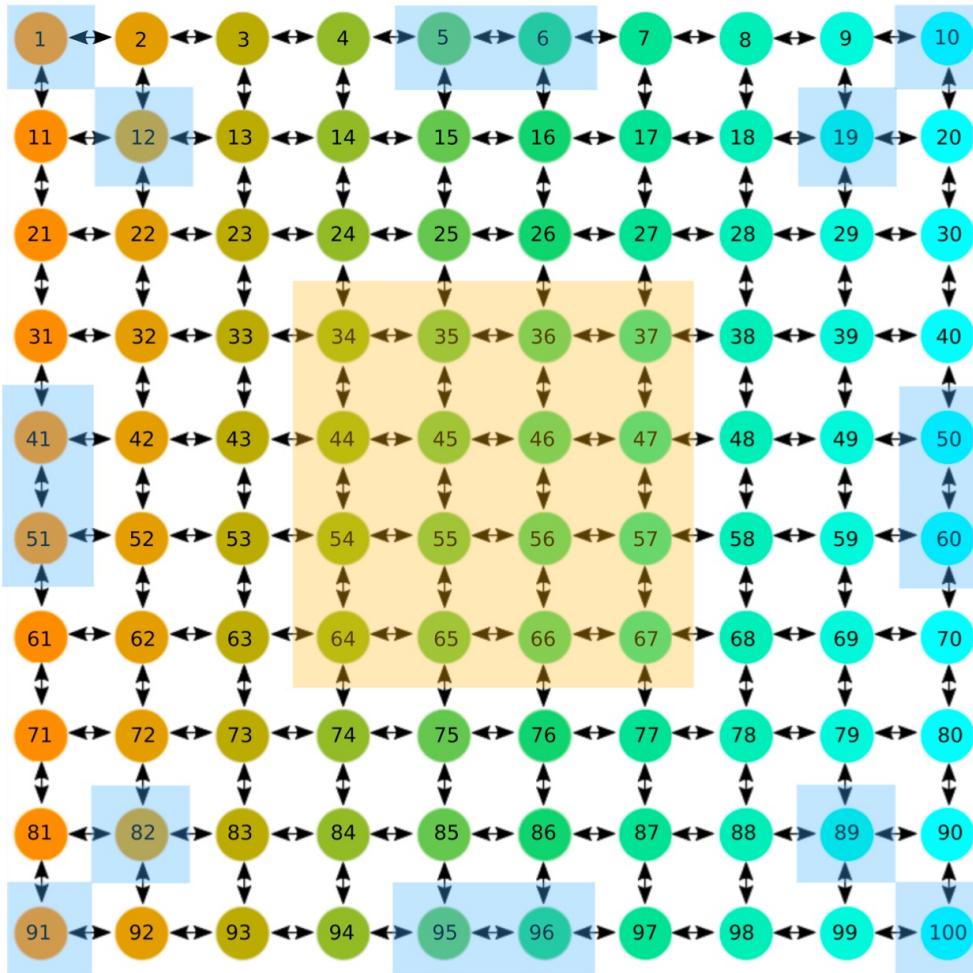
Linking population trend surveys and predicted genetic offset



Genotyping biological specimens to look for genetic changes between historical and contemporary climates

Only 8 studies to date have associated offsets with a metric of fitness, selection, or population performance, and most don't provide a true validation (*sensu* Lotterhos et al. 2019 *PLoS Biol.*)

Analysis validation using “truth known” simulations



- 10 x 10 array of demes with 2D stepping-stone migration
- 10 LGs (50K sites each) with mutation and recombination (LD)
- Stabilizing selection within demes; clinal selection across demes
- Single locus and polygenic genetic architectures
- Causal, “fake”(correlated) and random environmental predictors

Áki Láruson



Matt Fitzpatrick



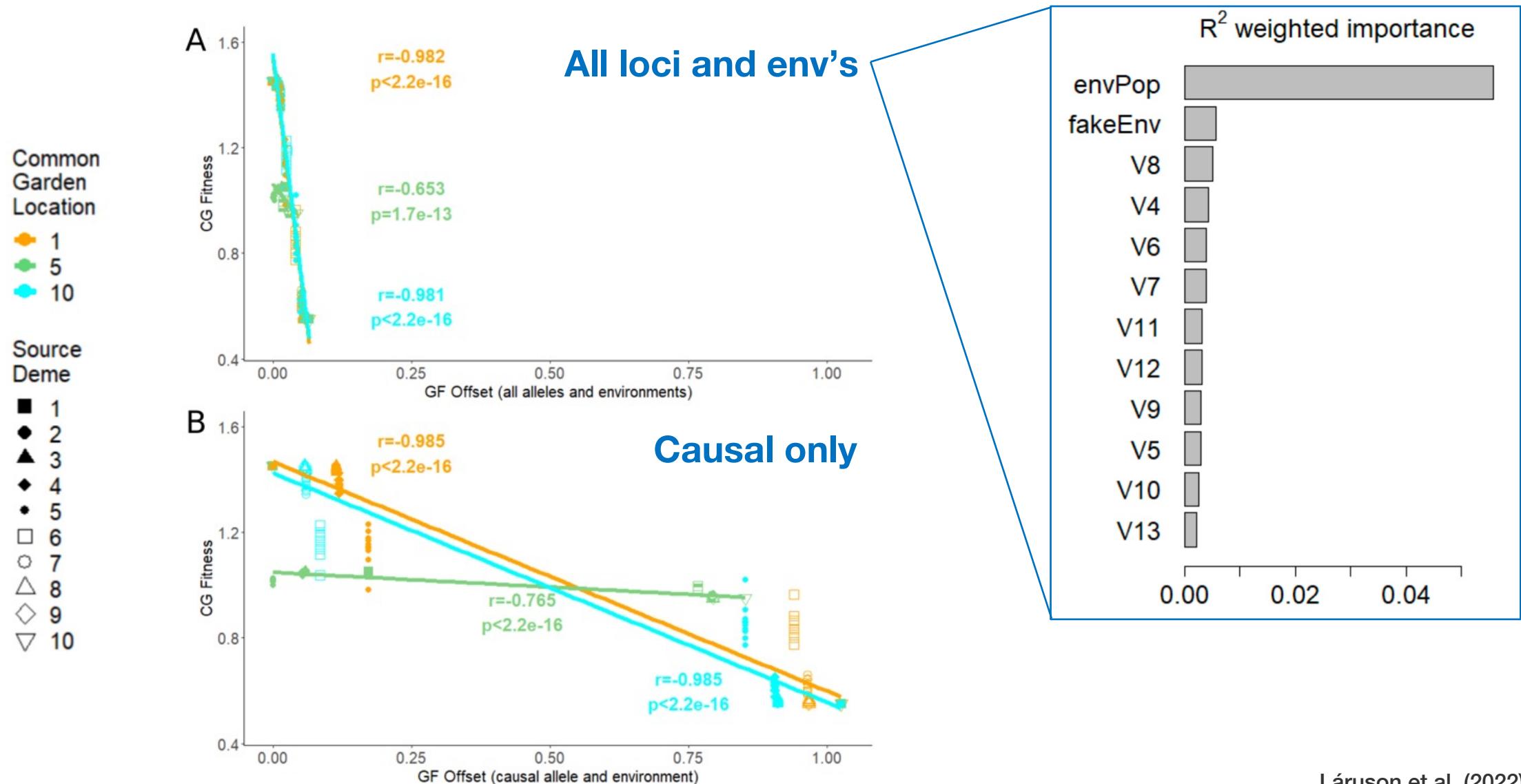
Ben Haller



Katie Lotterhos



Single-locus model: *Genomic Offset* predicts maladaptation



Environments vary linearly
with geography

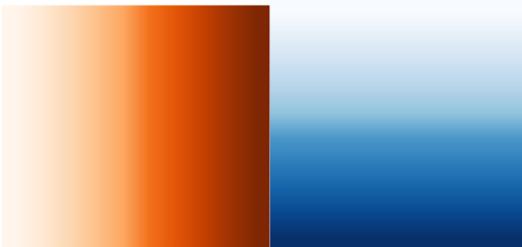
Environments vary non-
linearly with geography

Environments vary linearly;
narrower gradient for Env 2

Environments vary linearly;
weaker stabilizing selection
within demes for Env 2

Phenotypic Optima Gradient **Phenotype correlation
with environment** **Mean evolved
phenotype**

Case 1



Case 2



Case 3



Case 4



Env 1

Env 2

Time

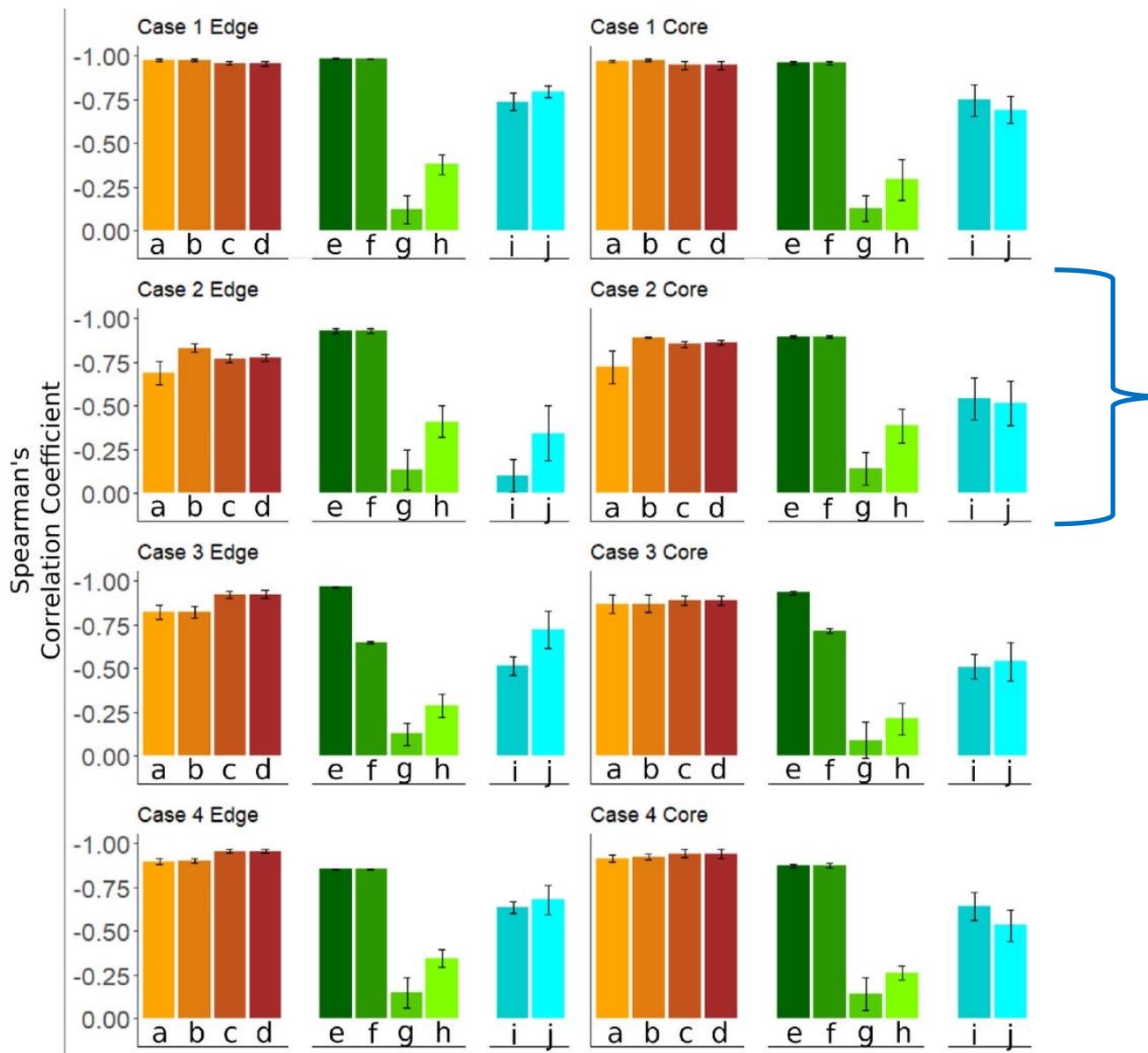
Env

Polygenic model: *Genomic Offset* predicts maladaptation

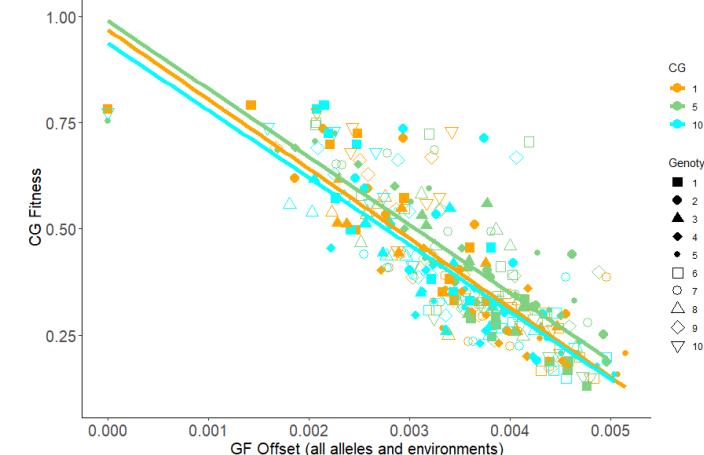
a GF off. genome, all env.
b GF off. genome, causal env.
c GF off. causal, all env.
d GF off. causal, causal env.

e Causal Env. dist. (ED)
f Causal Env. dist. (MD)
g All Env. dist. (ED)
h All Env. dist. (MD)

i FST genome
j FST causal



Case 2:



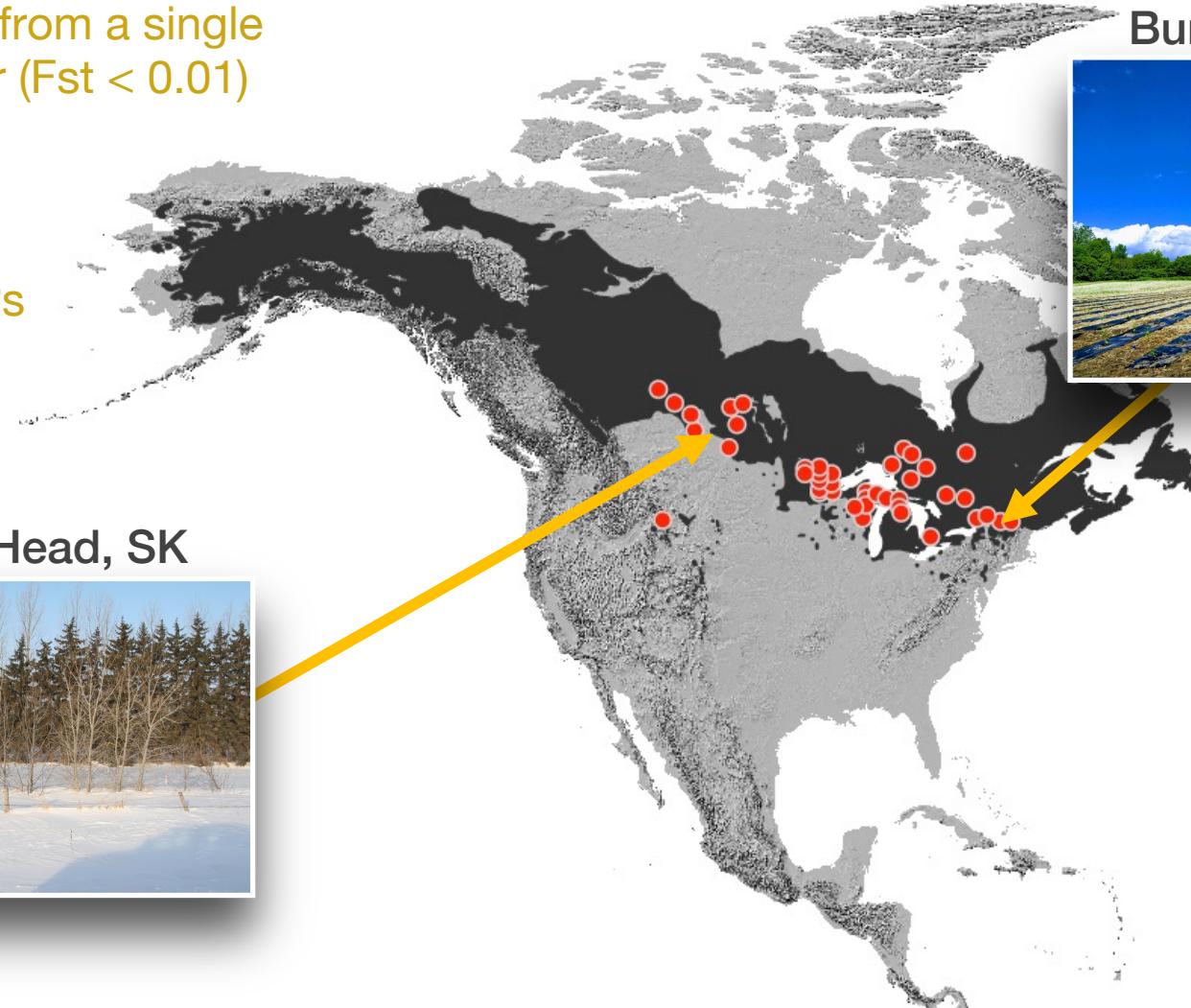
Case Study: Ground truthing with common gardens

Balsam poplar (*Populus balsamifera* L.)

- 42 populations from a single ancestry cluster ($F_{ST} < 0.01$)
- 336 individuals
- 106K GBS SNPs



Indian Head, SK



Burlington, VT



Fitzpatrick



Chhatre

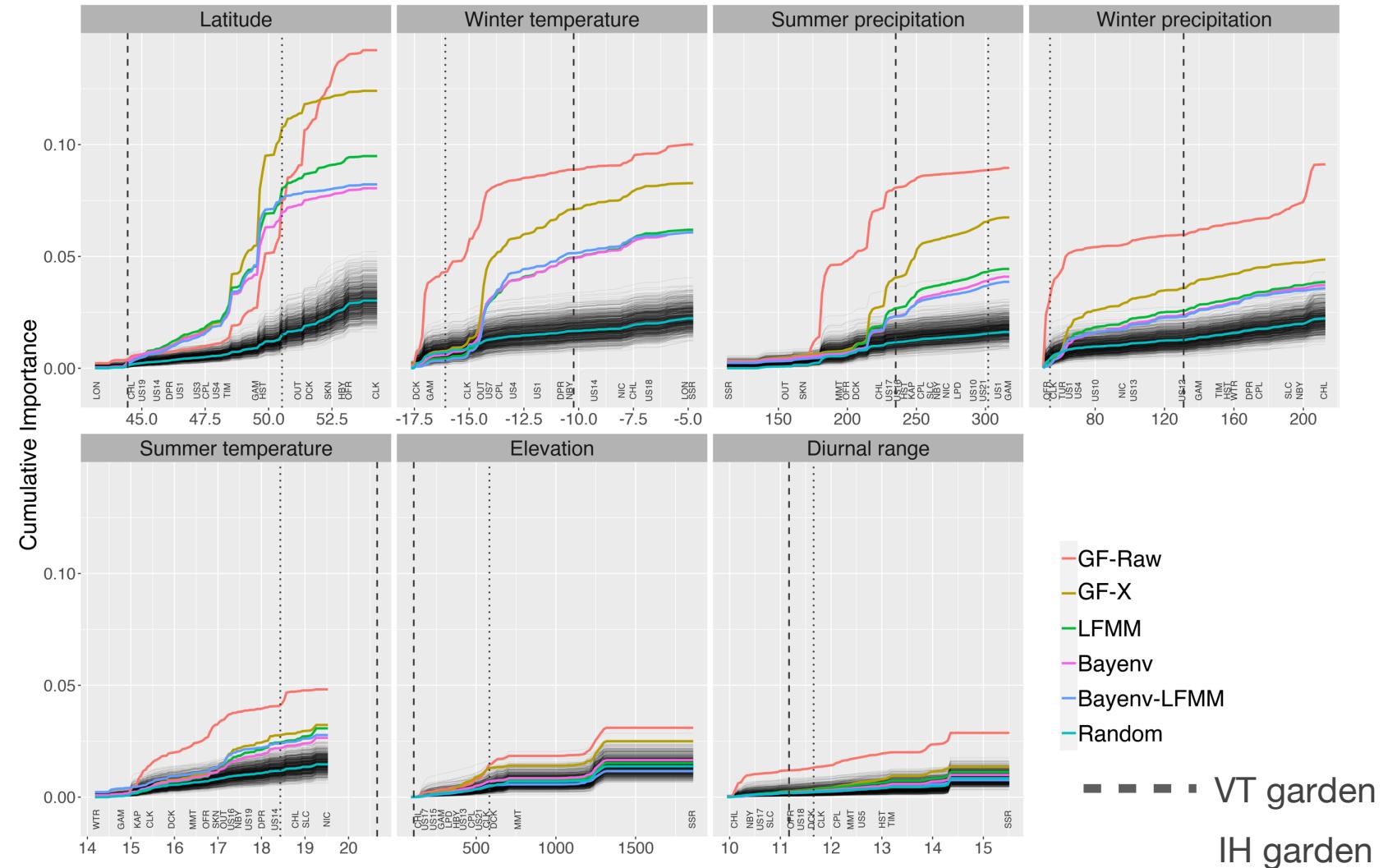
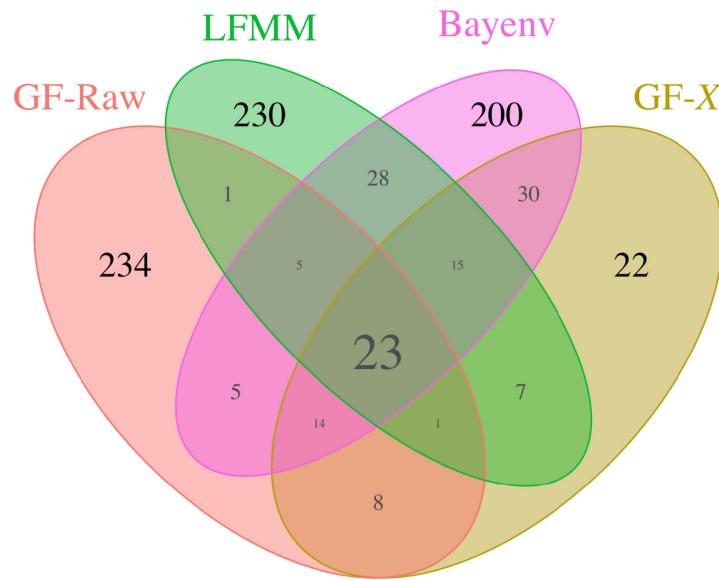


Soolanayakanahally



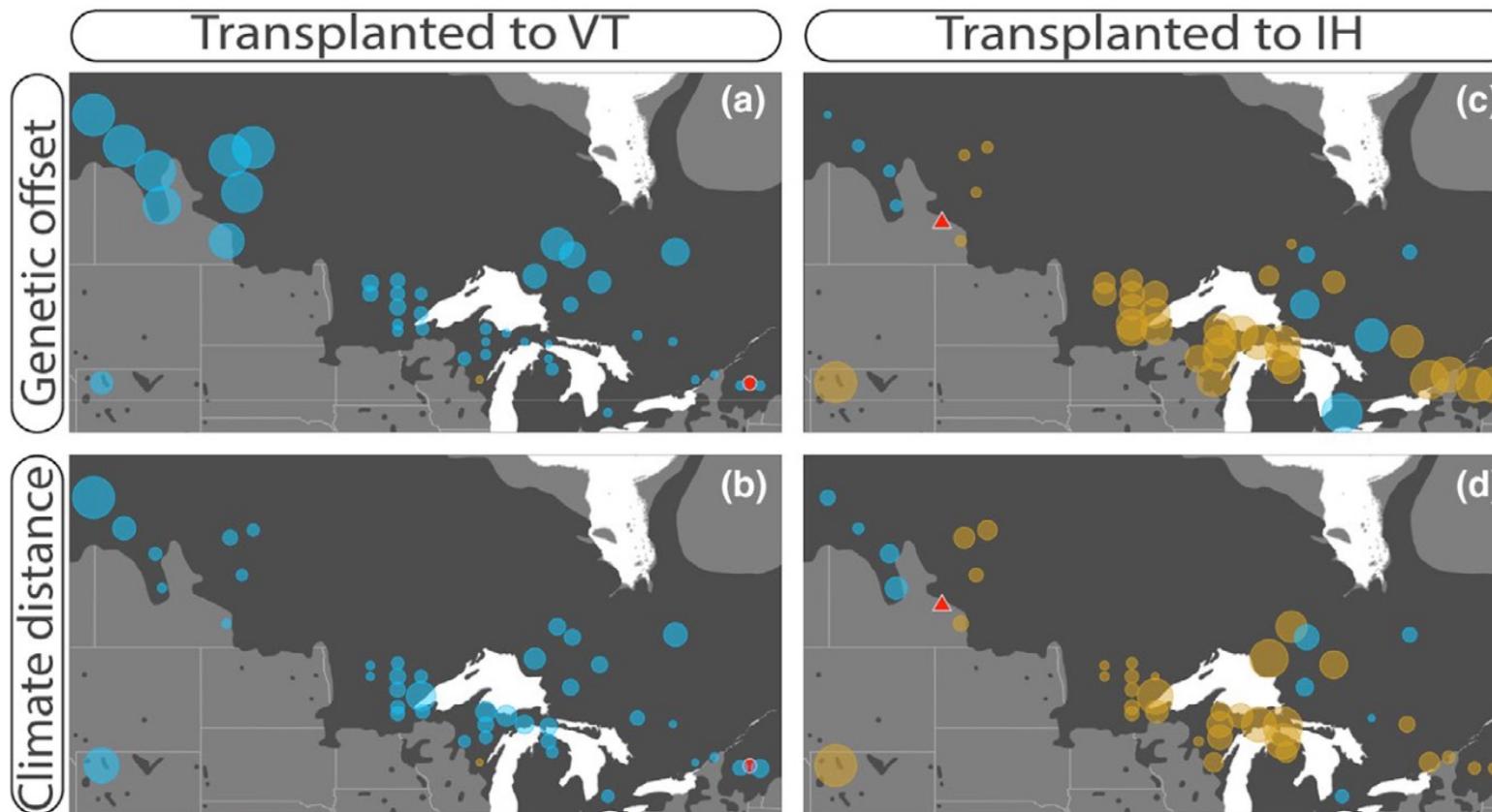
GF Cumulative Importance functions based on different outlier sets

Genome scans for local adaptation

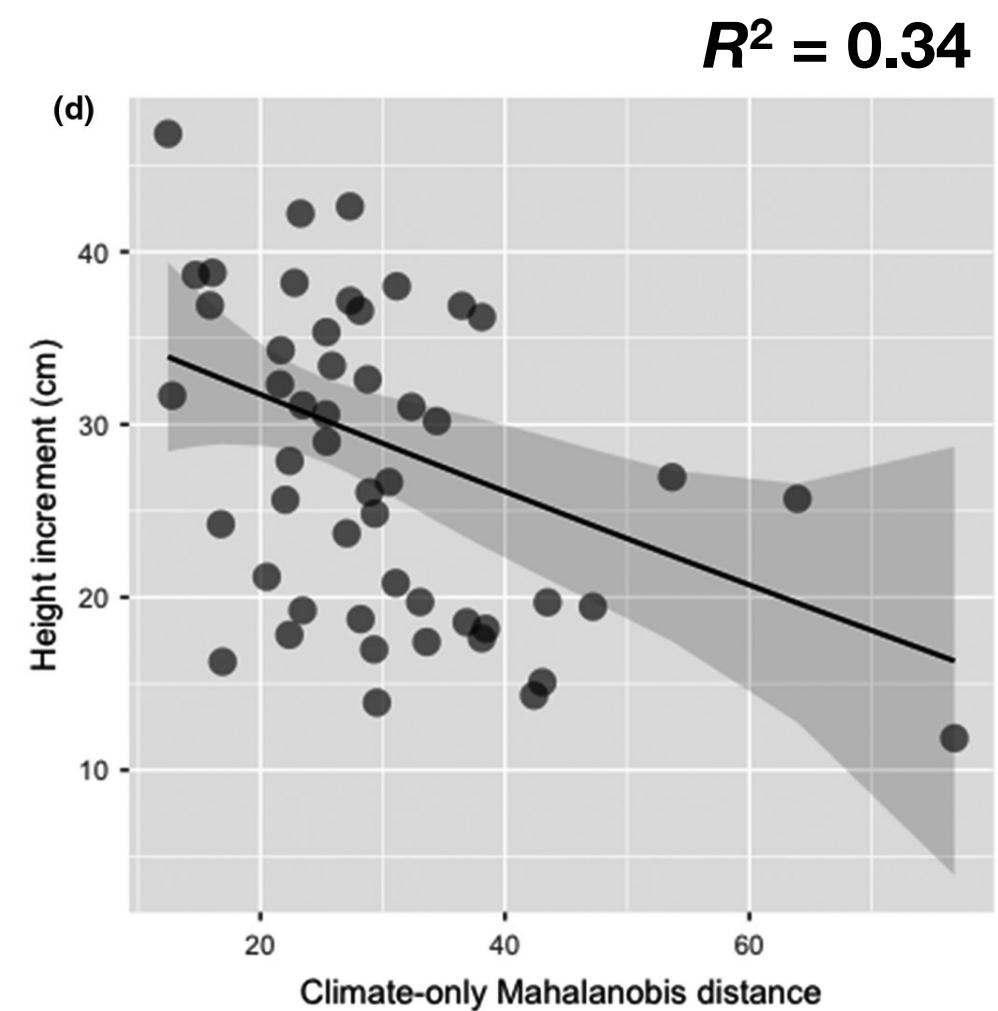
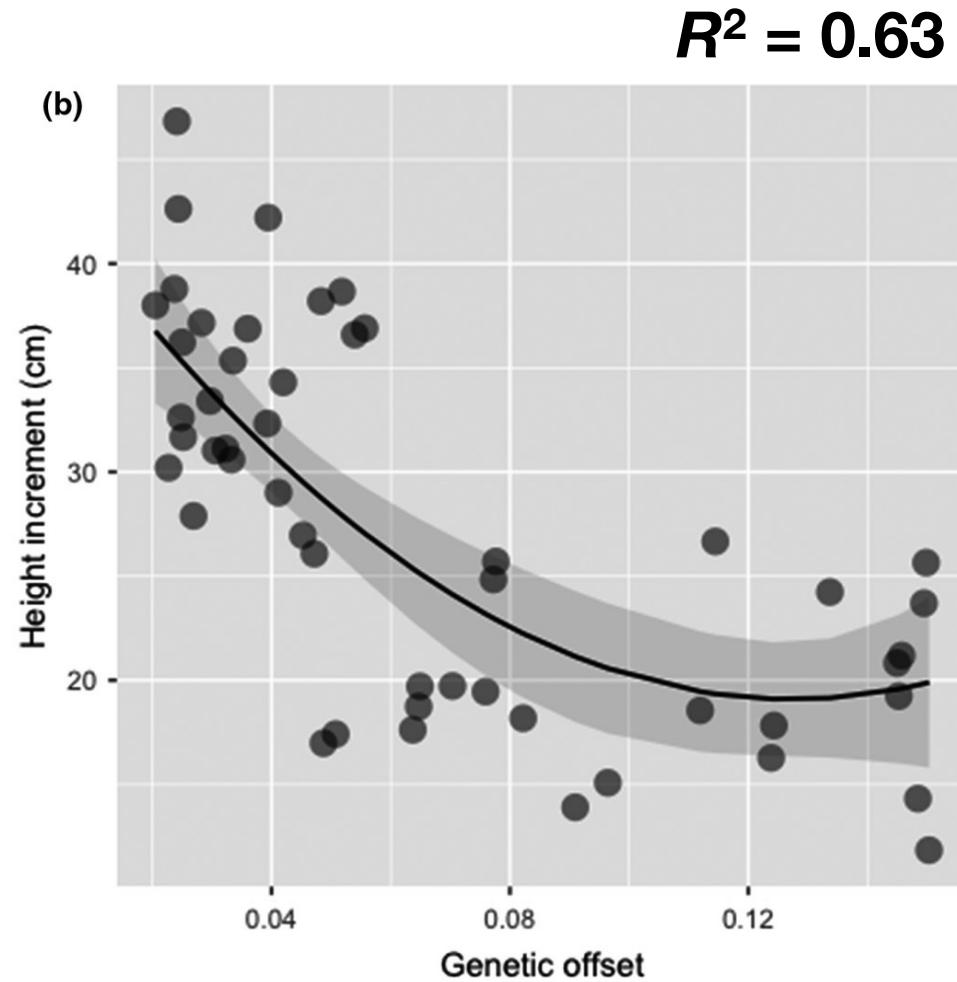


Space for time substitution: source climate → garden climate

- Clonal replicates transplanted to common gardens
- Height increment growth phenotyped after 3 years
- Same genotypes as used for offset estimations



Genomic Offset negatively associated with height growth





Ground truthing with common gardens

Red spruce (*Picea rubens* Sarg.)

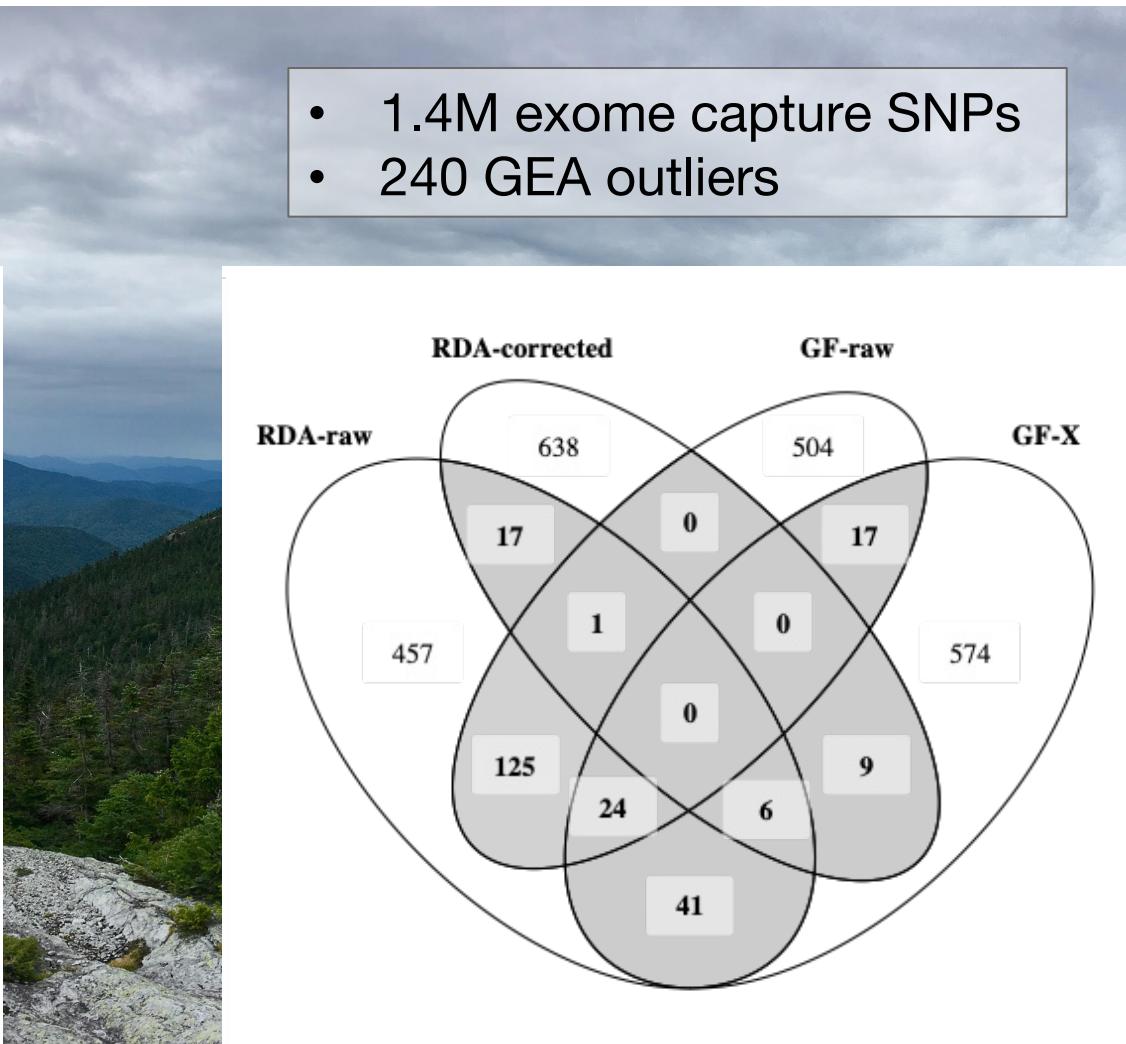
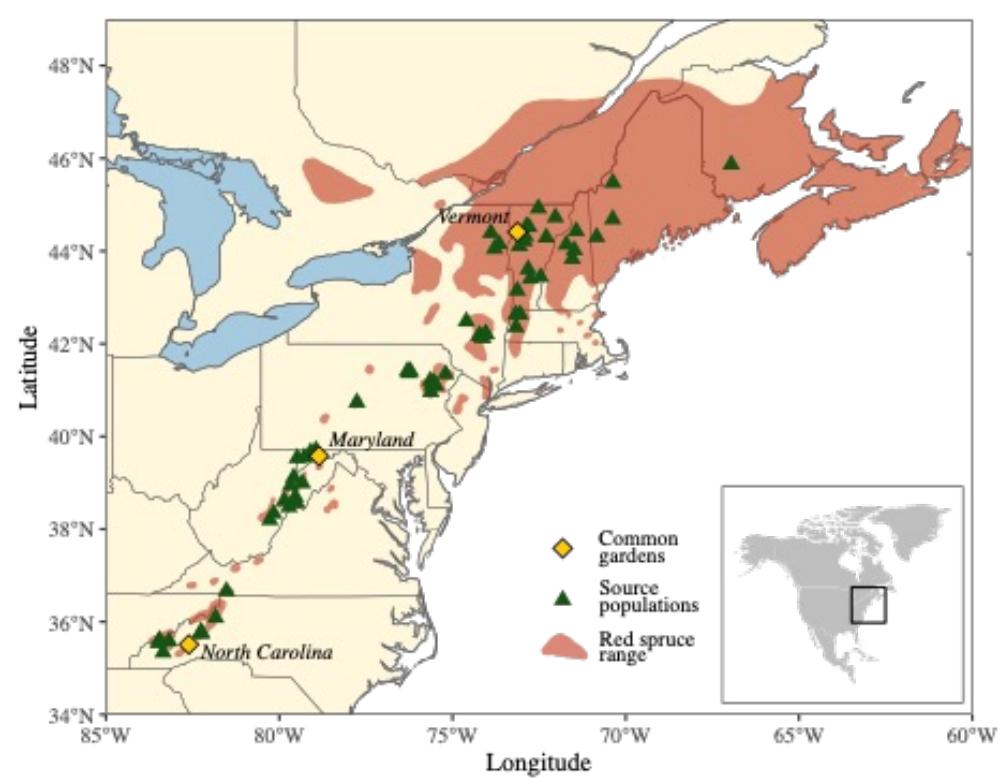


Susanne
Lachmuth

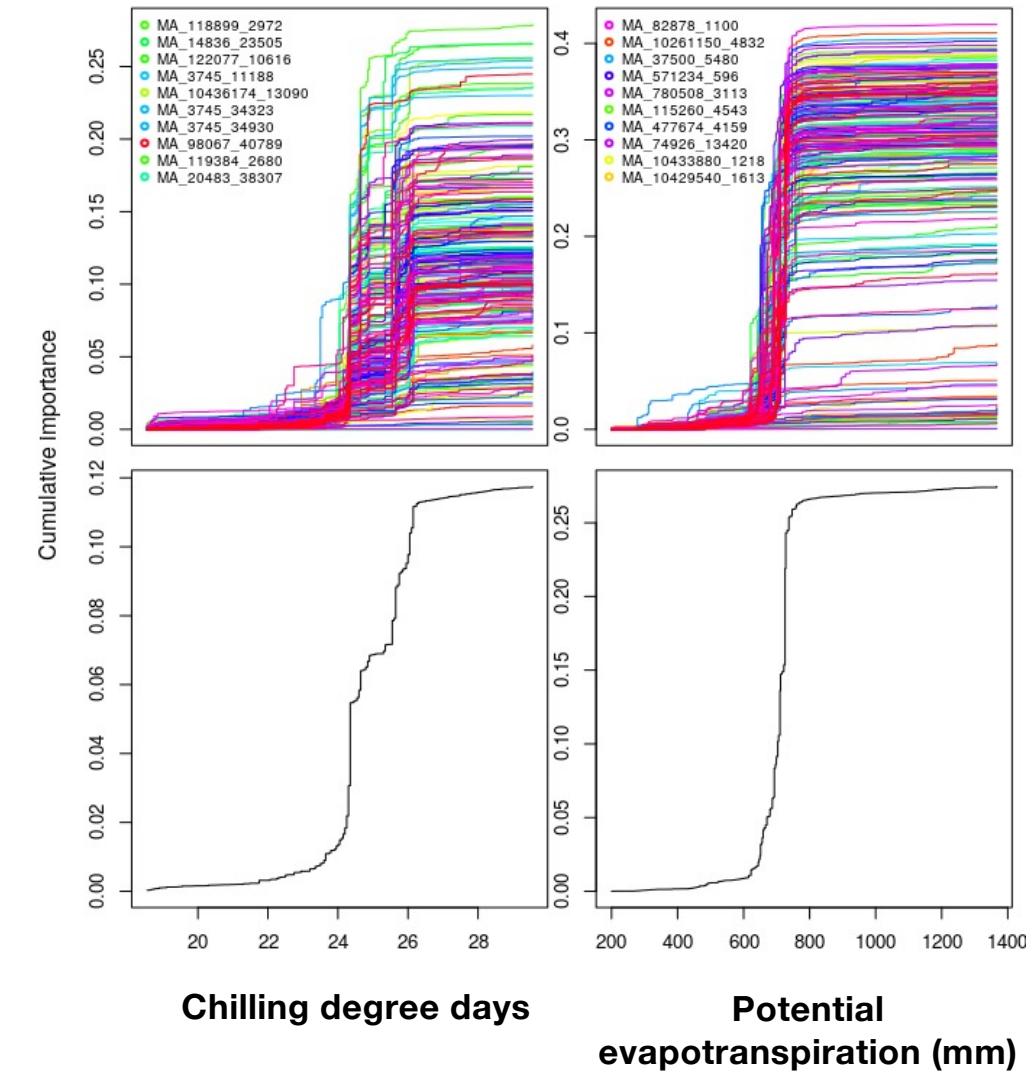
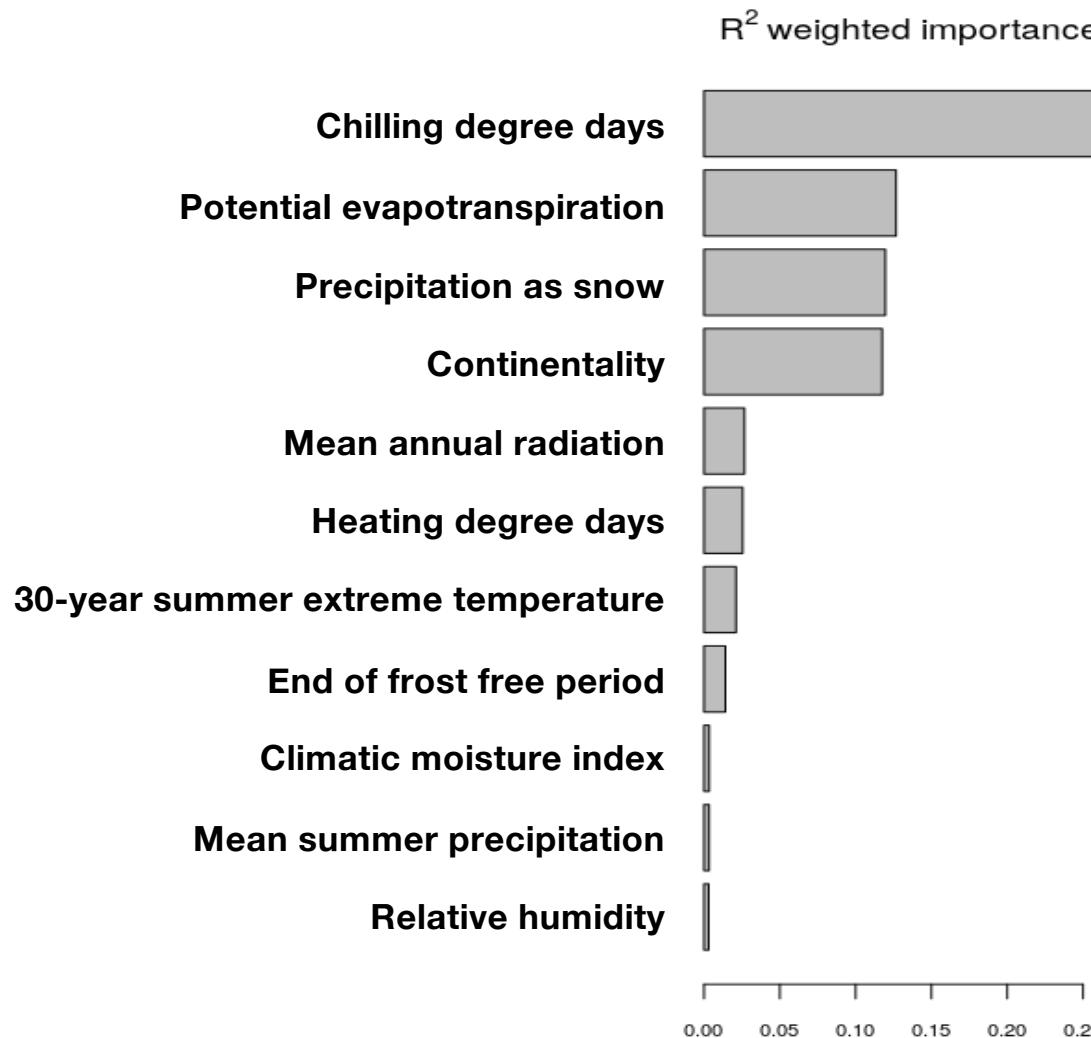
Thibaut
Capblancq

- 65 populations
- 340 individuals

- 1.4M exome capture SNPs
- 240 GEA outliers

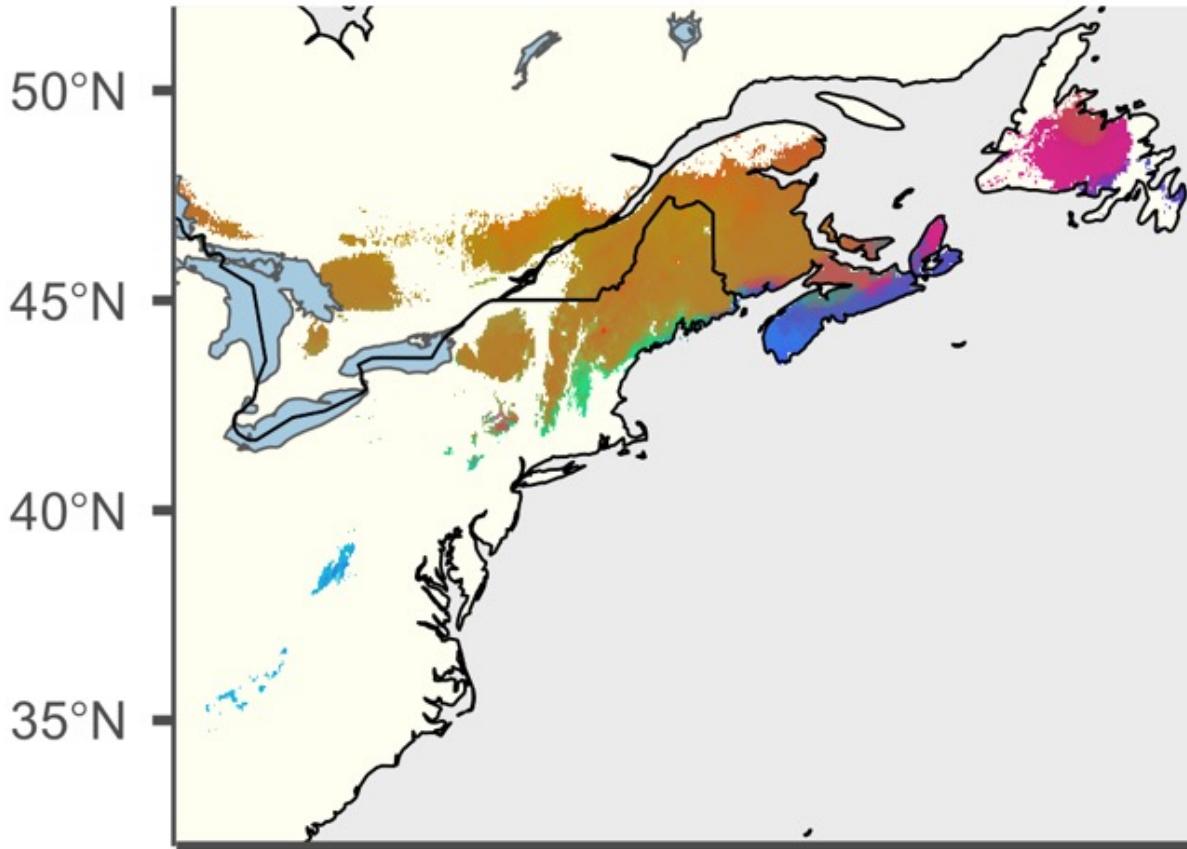


Gradient Forest turnover along climatic gradients

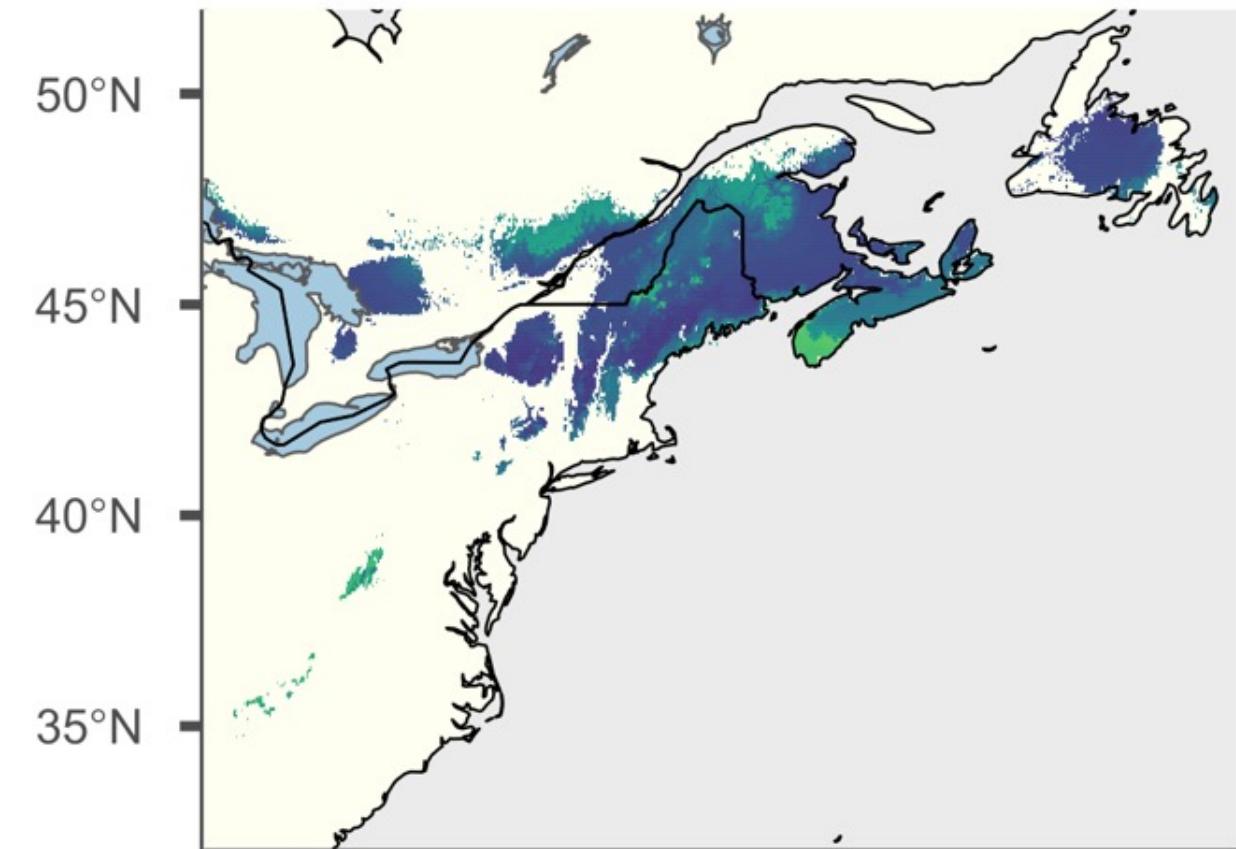


Adaptive turnover and *Genomic Offset* under future climates

Adaptive genomic turnover



Local offset under SSP2-45

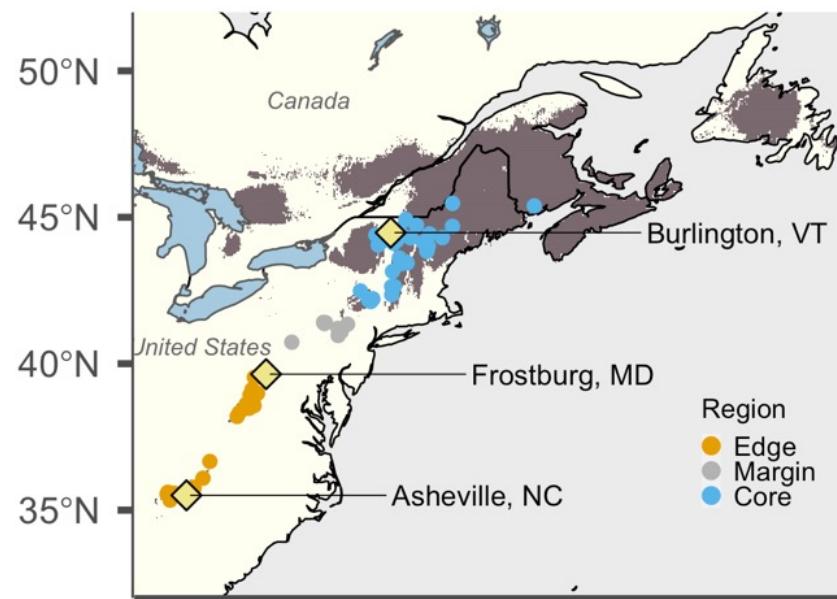


Local offset (z-score)



Common gardens to evaluate Genomic Offset predictions

Field sampling in 65 red spruce populations across the range



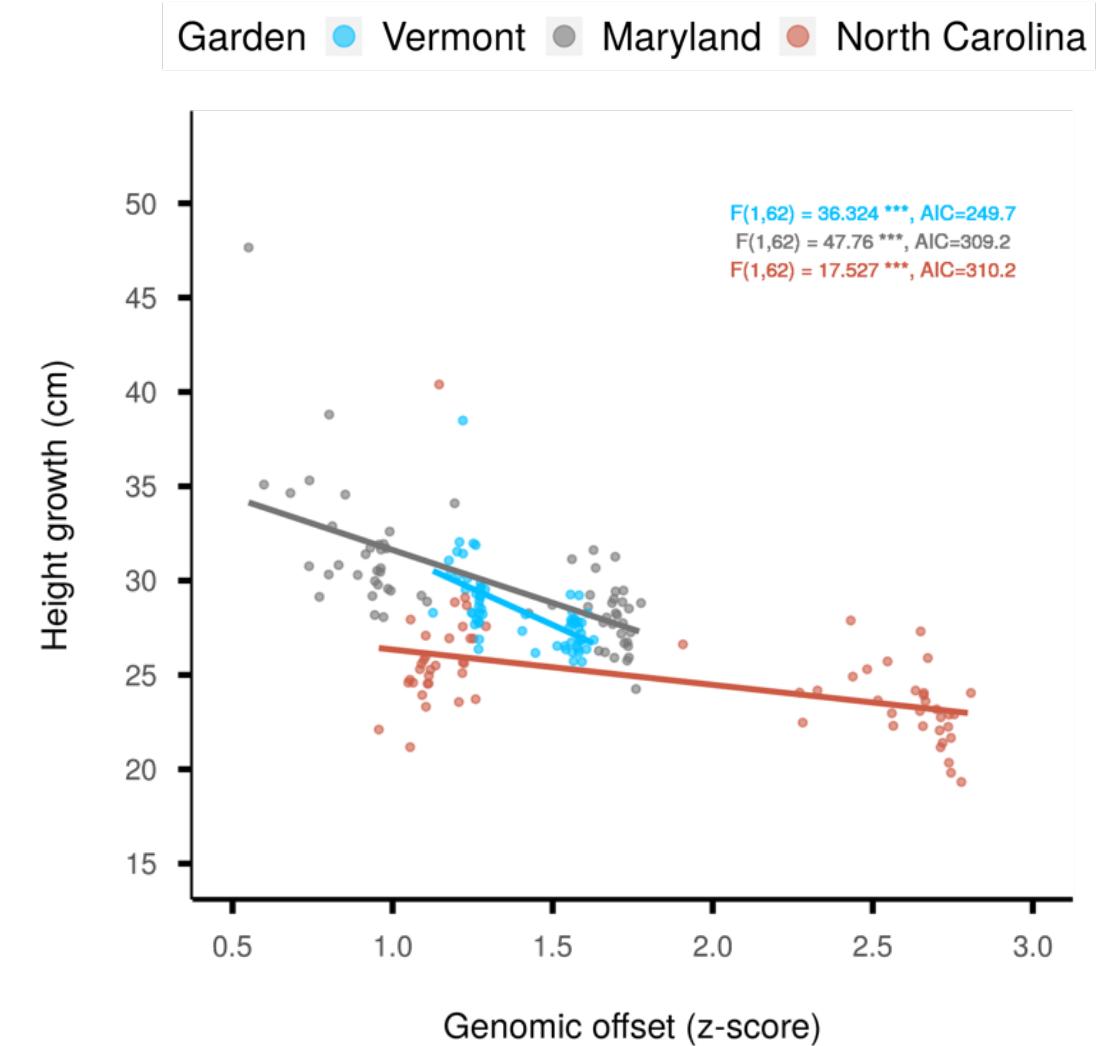
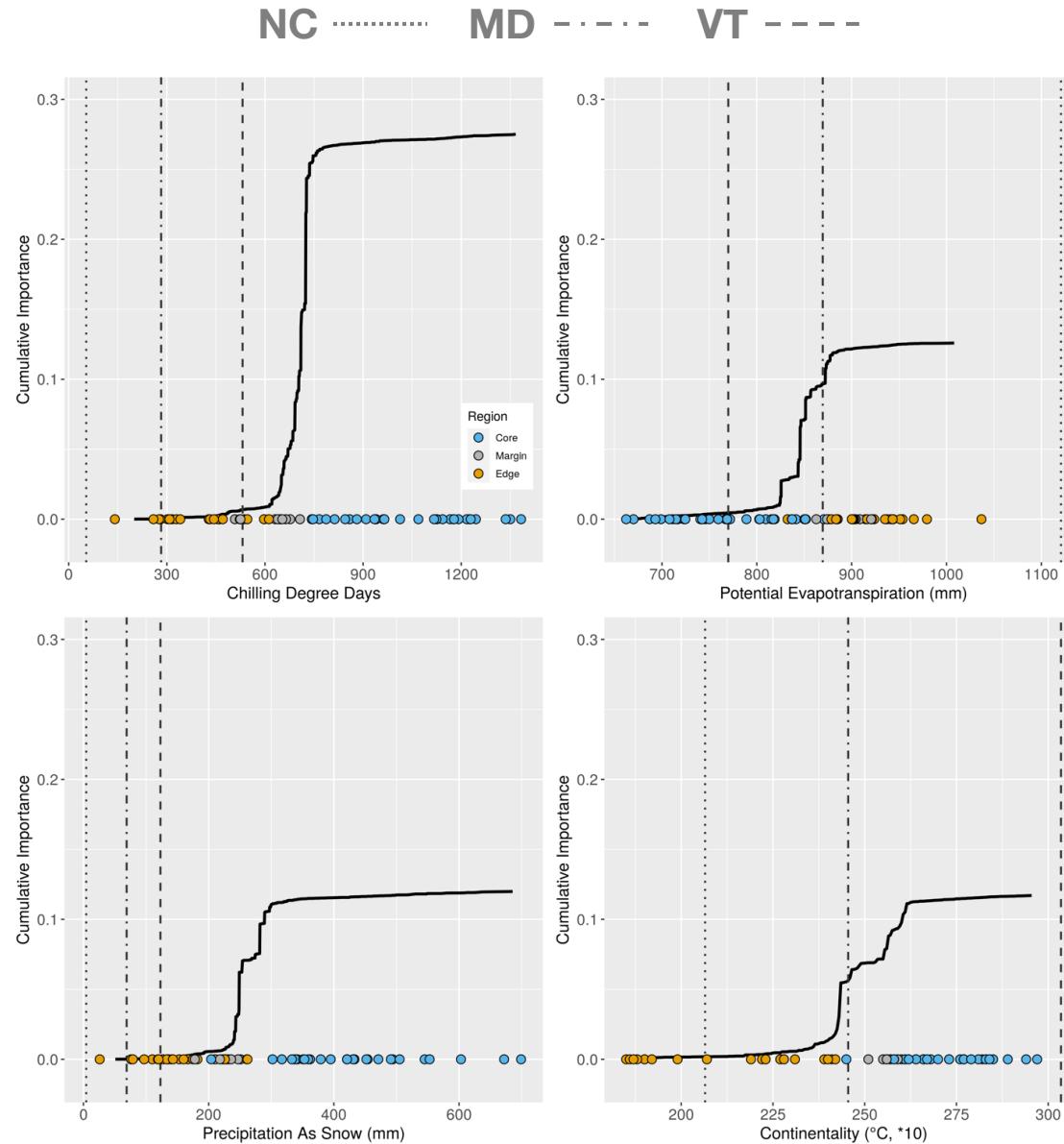
Three common gardens each with 1700 juvenile red spruce (340 families x 5 reps)



2 Year height growth

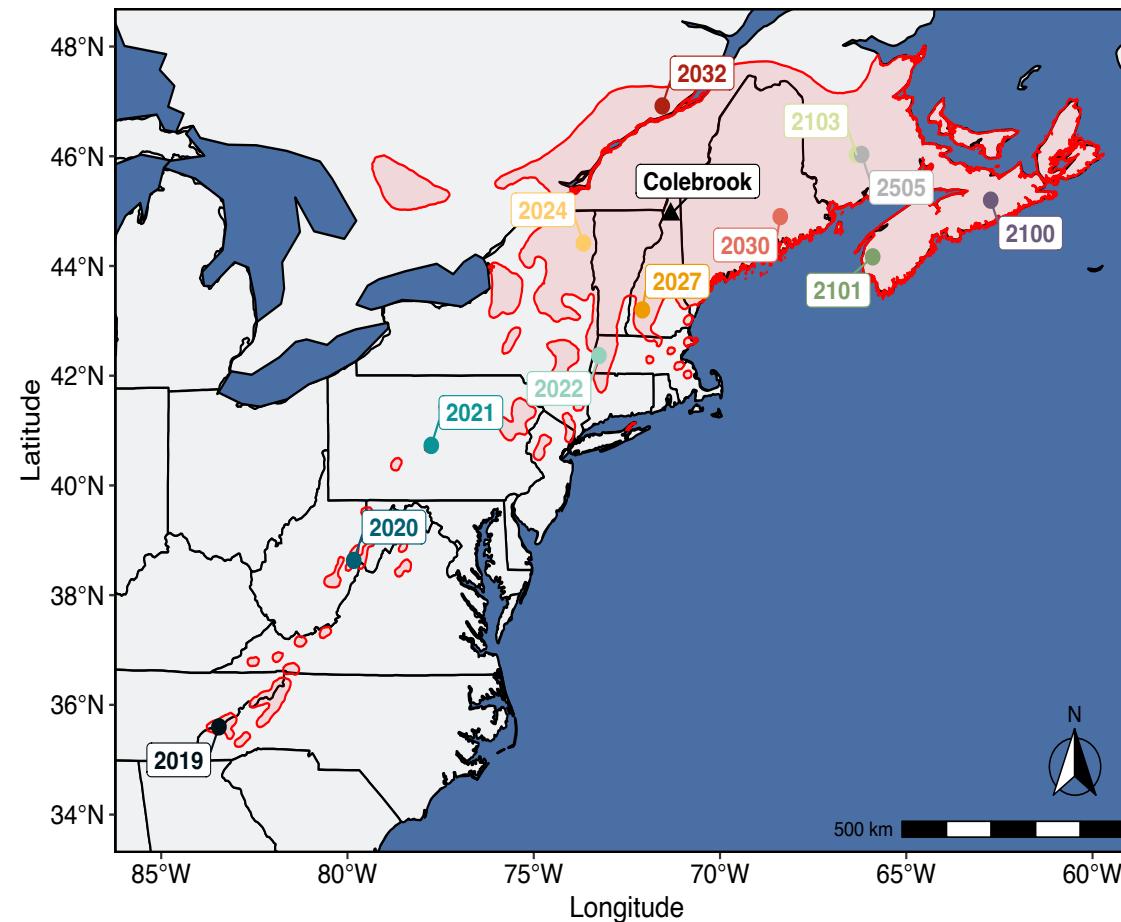


Genomic Offset: garden climate v. source climate



Empirical validation using an independent provenance trial

- Red spruce GF model from Lachmuth et al. training set (65 pops)
- *Genomic offset* projected for 12 independent pops planted in Colebrook, NH
- Tested against growth and survival across 3 time periods



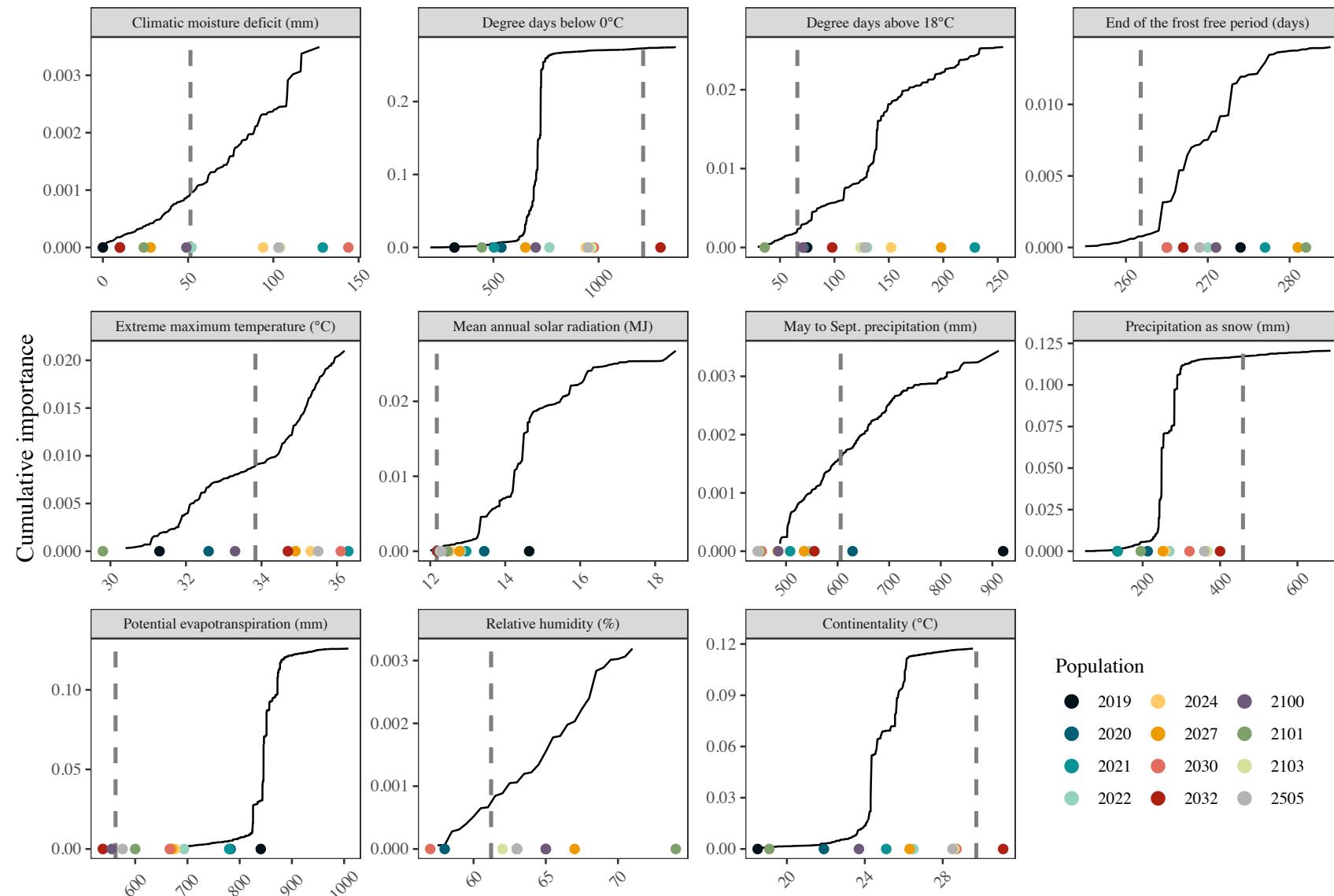
Brittany Verrico



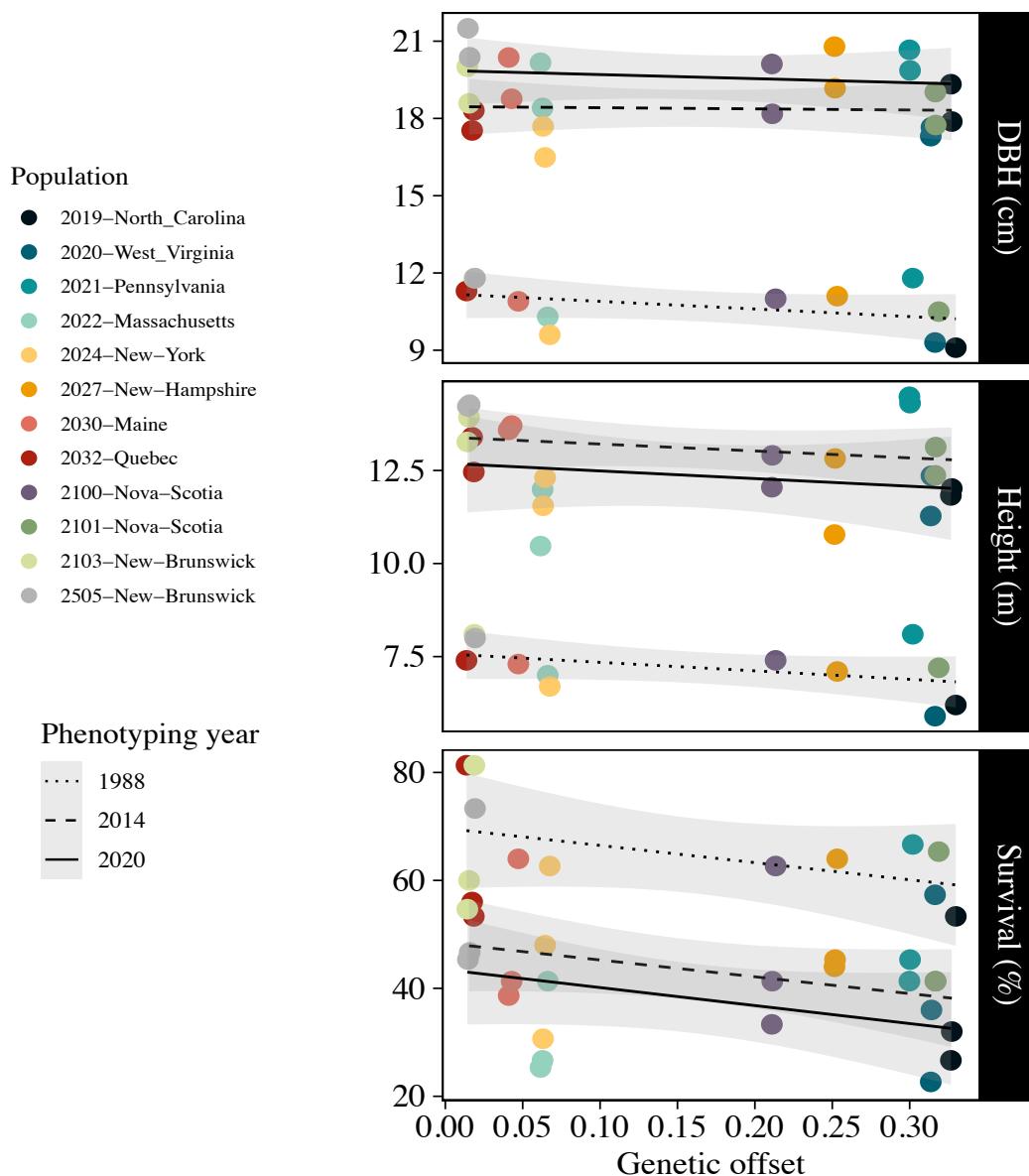
Thibaut Capblancq



Testing pops overlaid on turnover functions from training set



Weak association between offset and performance



- Weak to moderate decline in traits with increasing *Genomic Offset*
 - All P -values > 0.05
- Trend is consistent over years 1988, 2014, 2020
- Compare to seedling relationship (winner's curse? Life-stage differences? Directionality of climate shift?)

Limitations:

- Small sample size in testing data ($N=12$ pops)
- Traits measured on >30 y.o trees → asymptotic relationship between size and age

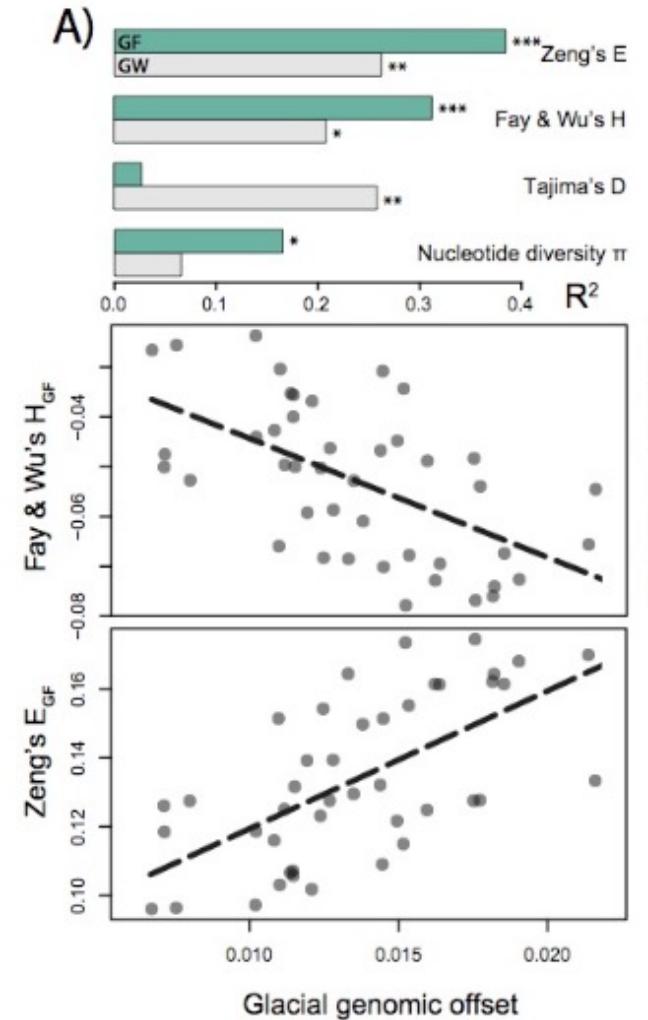
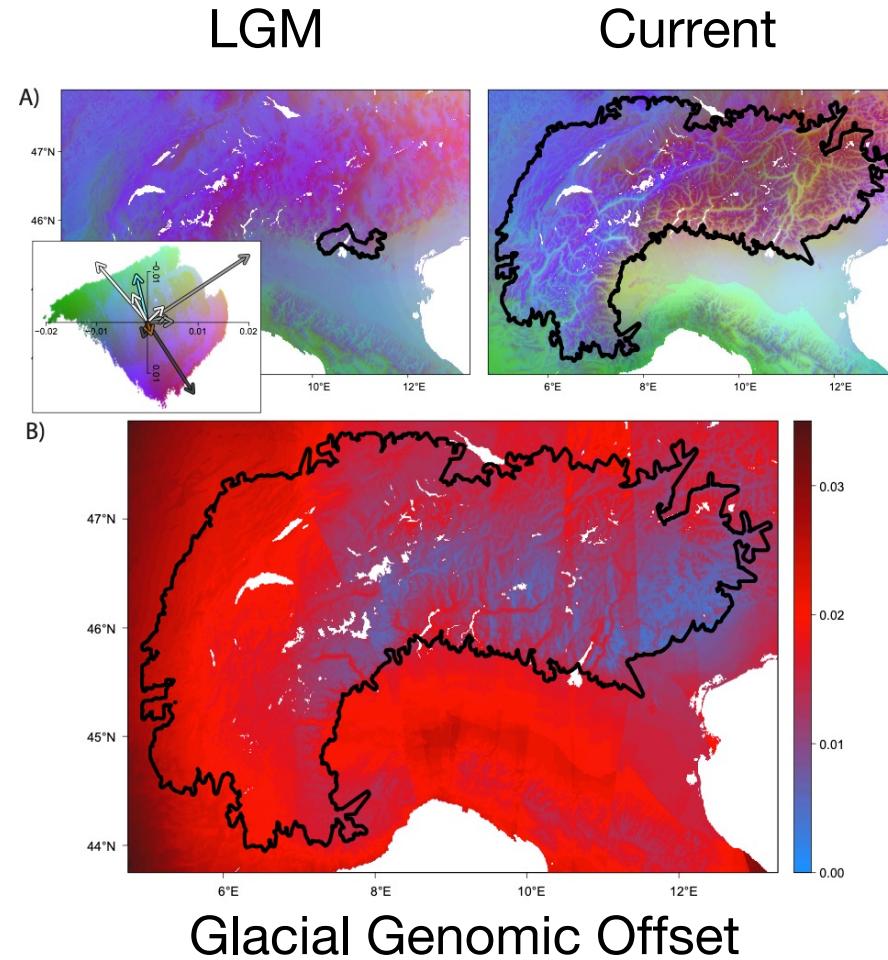
Ground truthing offset with selection signatures

Dianthus sylvestris



Glacial Genomic Offset:

GF turnover calculated from current GEA outliers and then hindcast to the LGM refugia



Where is *Genomic Offset* headed in the future?

- Rigourous comparisons of different *Genomic Offset* estimators to refine existing methods and/or develop new ones
- More simulation studies to better understand the effects of sampling, experimental design, genetic architecture, and population structure
- Carefully designed validation studies using different approaches (e.g., experimental evolution, resurrection studies, dendrochronology, common gardens, growth chambers, ...)
- Explore applications beyond climate change (Biological invasions, Restoration, Seed Zones, Agriculture, ...)

Acknowledgements

Thibaut and Olivier for organizing and hosting the SSMPG!

Methodology/Conceptual Development:

Matt Fitzpatrick (UMCES)
Thibaut Capblancq (LECA)
Andy Gougherty (Purdue)
Susanne Lachmuth (UMCES)

Simulation Testing:

Aki Laruson (Cornell)
Katie Lotterhos (NEU)
Matt Fitzpatrick (UMCES)
Ben Haller (Cornell)
Vikram Chhatre (UWYO)

Common Garden Ground Truthing:

Anoob Prakash (UVM)
Raju Soolanayakanahally (AAFC)
Matt Fitzpatrick (UMCES)
Susanne Lachmuth (UMCES)
Thibaut Capblancq (LECA)
John Butnor (USDA)
Brit Verrico (UVM)
Vikram Chhatre (UWYO)

Funding:

