

# Phylotranscriptomics Reveals the Genetic Basis of Climate Adaptation in *Acers*

Mengmeng Lu Jan 14, 2024

### **Principal Investigator**



### Nate Swenson

Gillen Director of the University of Notre Dame Environmental Research Center Professor of Biology

CV

Google Scholar



National Science Foundation



## Distribution of Acer (maples)

- ➤ ~132 extant species of trees or shrubs
- > Most are native to Asia



From https://en.wikipedia.org/wiki/Maple

# Acer (maples)



From https://www.thoughtco.com/the-national-flag-of-canada-508080



From https://www.cbc.ca/news/canada/montreal/maplesyrup-grade-producers-sugar-shack-1.6375742



From https://www.etsy.com/listing/1484227236/84-softmaple-hardwood-lumber-top-grade?gpla=1&gao=1&

## Acer (maples)









Mount Royal, Montreal, October, 2021

## Acer under climate change

Warming temperatures and changes in freeze and thaw cycles have affected maple trees and syrup production and quality



"... long periods of freezing in the middle of the season are really hard...and sap is not flowing..."

"...A few weeks later it's really warm and the trees just stop producing sap..."

## To study the genetic basis of local adaptation



#### **MOLECULAR ECOLOGY**

Molecular Ecology (2015) 24, 4348-4370

doi: 10.1111/mec.13322

INVITED REVIEWS AND SYNTHESES

A practical guide to environmental association analysis in landscape genomics

CHRISTIAN RELLSTAB,\* FELIX GUGERLI,\* ANDREW J. ECKERT,† ANGELA M. HANCOCK‡ and ROLF HOLDEREGGER\*§

\*WSL Swiss Federal Research Institute, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland, †Department of Biology, Virginia Commonwealth University, Richmond, VA 23284, USA, ‡Faculty of Molecular Biology, Max F. Perutz Laboratories and University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria, §ETH Zürich, Institute of Integrative Biology, Universitälstrasse 16, 8092 Zürich, Switzerland Phylotranscriptomics-based method to study the genetic basis of climate adaptation



## Studied species

## North America

Acer rubrum (red maple) Acer saccharum (sugar maple) Acer spicatum (mountain maple)



Acer amplum Asia Acer barbinerve (bearded maple) Acer buergerianum Acer caesium (Himalayan maple) Acer caudatum (candle-shape maple) Acer chingii Acer cordatum Acer franchetii Acer ginnala (Amur maple) Acer henryi Acer heptalobum (Japanese maple) Acer kuomeii Acer laevigatum (smooth maple or Nepal maple) Acer mandshuricum (Manchurian maple) Acer mono (painted maple or mono maple) Acer oblongum (Himalayan maple, evergreen maple or Kashmir maple) Acer paxii Acer pectinatum Acer robustum Acer sinense Acer stachyophyllum Acer tegmentosum (Manchurian stripedbark maple or Manchu striped maple) *Acer triflorum (*three-flowered maple) Acer truncatum (Shandong maple or purpleblow maple) Acer wilsonii (Three Gorges maple) Acer flabellatum

Dipteronia sinensis

Phylotranscriptomics-based method to study the genetic basis of climate adaptation



## Step 1

Performing ortholog test and phylogeny construction

- To find one-to-one orthogroups (OGs)
  - > *De novo* transcriptome assembly
  - Predicting the longest peptide sequence
  - Using OrthoFinder to predict the preliminary OGs
  - > Filtering OGs with sequences absent in one or more samples
  - > Aligning OGs and selecting conserved blocks with at least 100 amino acids
  - Removing the remaining paralogs and contamination

## Step 1

Performing ortholog test and phylogeny construction

**Coalescence analysis** 

## ≥211 one-to-one orthogroups (OGs)

Method 1

Sequence format

CDS (1<sup>st</sup> +2<sup>nd</sup>)

Amino acid

 $\succ$  CDS



Generating gene trees ---ightarrow inferring species tree using gene trees

Method 2 Maximum likelihood (ML) analysis using concatenated sequences

Method 3 Maximum likelihood (ML) analysis using partitioned sequences

Method 4 Bayesian inference (BI) using concatenated sequences

Method 5 Bayesian inference (BI) using partitioned sequences

> Finally, we yielded 15 phylogenetic trees.

## Step 1 Performing ortholog test and phylogeny construction

> These 15 phylogenetic trees are almost the same except for the placement of Acer ginnala & Acer spicatum.



A species tree inferred using 210 gene trees labeled with branch local posterior probability. An unpartitioned ML tree of *Acers* reconstructed from the concatenated alignment of amino acid sequences of 211 OGs labeled with SH-aLRT support (%) / ultrafast bootstrap support (%).



A consensus BI tree of *Acers* reconstructed from the partitioned alignment of CDS ( $1^{st} + 2^{nd}$ ) of 211 OGs labeled with support rate.



Removing A. spicatum & A. ginnala, the final phylogeny includes 27 species. ( ) contains the section names.





#### > Clustering in this phylogenetic tree is consistent with section classification of Acers.

### Step 1

### Performing ortholog test and phylogeny construction



> This is the final phylogenetic (species) tree.

> Derive phylogenetic distance from this tree.

Phylotranscriptomics-based method to study the genetic basis of climate adaptation



## Step 2

## Detecting correlation between phylogeny and climate distance

- > Download GBIF presence occurrence data for each of the 27 species in the phylogeny
- Plot occurrence distribution maps



- Download GBIF presence occurrence data for each species
- Plot occurrence distribution (red spots) maps





> Calculate the median values of each climate variable (WorldClim v2.1) within the natural range

Step 2 Detecting correlation between phylogeny and climate distance

1) Phylogenetic signal in climate variables

- > Blomberg's K and Pagel's lambda are two measures of phylogenetic signals
- > Calculate K and lambda in the median values of each climate variable (From WorldClim v2.1) in the natural range

0.98	0.00	1.00	0.00	bio1_Annual mean temperature	1
0.69	0.01	0.42	0.07	bio2_Mean diurnal range	0.8
0.45	0.46	0.00	1.00	bio3_lsothermality (bio2/bio7)	0.6
0.64	0.04	0.19	0.52	bio4_Temperature seasonality	0.4
0.58	0.08	0.10	0.78	bio5_Max temperature of warmest month	0.4
0.89	0.01	1.00	0.01	bio6_Min temperature of coldest month	0.2
0.74	0.01	1.00	0.06	bio7_Temperature annual range (bio5-bio6)	
0.72	0.01	0.75	0.11	bio8_Mean temperature of wettest quarter	
	0.00	1.00	0.01	bio9_Mean temperature of driest quarter	
0.68	0.03	0.47	0.24	bio10_Mean temperature of warmest quarter	
	0.00	1.00	0.00	bio11_Mean temperature of coldest quarter	
0.58	0.08	0.23	0.28	bio12_Annual precipitation	
0.61	0.06	0.17	0.54	bio13_Precipitation of wettest month	
0.67	0.05	1.00	0.07	bio14_Precipitation of driest month	
0.65	0.06	1.00	0.14	bio15_Precipitation seasonality	
0.61	0.06	0.22	0.38	bio16_Precipitation of wettest quarter	
0.64	0.07	1.00	0.17	bio17_Precipitation of driest quarter	
0.59	0.08	0.13	0.62	bio18_Precipitation of warmest quarter	
0.59	0.10	0.00	1.00	bio19_Precipitation of coldest quarter	
K_val	κ_p	lambda_val	lambda_p	> Dhylogopotic cign	



> Phylogenetic signals exists in climate variables related with minimum temperature of cold months.

Step 2 Detecting correlation between phylogeny and climate distance

2) Correlations between pairwise phylogenetic distance and climate variables

For each pair of the 27 species in the phylogenetic tree,

- Calculate the pairwise phylogenetic distances using its branch lengths
- Calculate the pairwise Euclidean distances in the 11 temperature BIOs and the 12 monthly minimum temperature
- Run Mantel test

	Temperature BIOs vs. phylogenetic distance	Monthly minimum temperature vs. phylogenetic distance
Mantel coefficient (r)	0.098	0.154
One-tailed p-value (null hypothesis: r<=0)	0.067	0.018

> Phylogenetic distance is correlated with monthly minimum temperature distance.



Step 2 Detecting correlation between phylogeny and climate distance

3) Pairwise temperature ecological niche (ENM) similarities

#### @ ENMTools

#### R-CMD-check failing

This package implements various tests, visualizations, and metrics for use with environmental niche models (ENMs) and species distribution models (SDMs).

#### Citation

Warren, D.L., Matzke, N.J., Cardillo, M., Baumgartner, J.B., Beaumont, L.J., Turelli, M., Glor, R.E., Huron, N.A., Simões, M., Iglesias, T.L. Piquet, J.C., and Dinnage, R. 2021. ENMTools 1.0: an R package for comparative ecological biogeography. Ecography, 44(4), pp.504-511.



- Generate PCs using the 11 temperature BIOs of the 27 species according to their occurrence records
- Construct ENM using the first 4 PCs (98.9% variance)
- Calculate the Spearman rank correlation coefficient between each pair of rasters
- Plot heatmap of the correlation with dendrogram cluster function
- The temperature ENM of "cold maples" (A. saccharum, A. mandshuricum, A. tegmentosum) tend to cluster together.



Phylotranscriptomics-based method to study the genetic basis of climate adaptation



1) Selecting ortholog groups (OGs) highly correlated with climate variables

- Ortholog assignment including at least 10 species
- Align OGs and select the conserved alignment
- Remove paralogs
- > Construct gene trees for each OG, and calculate the phylogenetic distance between each pair of species
- Calculate the pairwise Euclidean distances in the 12 monthly minimum temperature
- Run Mantel test between phylogenetic distance and climate distances
- Find out the OGs significantly correlated with climate distances

141 OGs (r > 0.5 & p-value <= 0.05) are highly correlated with pairwise distance in minimum temperature

1) Selecting ortholog groups highly correlated with climate variables

141 OGs (r > 0.5 & p-value <= 0.05) are highly correlated with pairwise distance in minimum temperature

Check the GO term enrichment in these highly correlated OGs.



2) Detecting genes for positive selection and selection intensity



Set "A. saccharum, A. mandshuricum, A. tegmentosum" as foreground and the other species as background,

Test selection intensity using RELAX

k>1 indicates that selection strength has been intensified;

k<1 indicates that selection strength has been relaxed

Test positive selection using PAML

Set "A. saccharum, A. mandshuricum, A. tegmentosum" as foreground and the other species as background,

Test selection intensity using RELAX

> Calculate genome-wide median of K (stringency of selection) value for each species

### $\succ$ Run PGLS between K and BIOs

bio_var	PGLS_pva
bio1_Annual mean temperature	0.05
bio2_Mean diurnal range	0.27
bio3_lsothermality (bio2/bio7)	0.43
bio4_Temperature seasonality	0.06
bio5_Max temperature of warmest month	0.8
bio6_Min temperature of coldest month	<mark>0.02</mark>
bio7_Temperature annual range (bio5-bio6)	0.05
bio8_Mean temperature of wettest quarter	0.69
bio9_Mean temperature of driest quarter	<mark>0.01</mark>
bio10_Mean temperature of warmest quarter	0.51
bio11_Mean temperature of coldest quarter	<mark>0.02</mark>
bio12_Annual precipitation	0.71
bio13_Precipitation of wettest month	0.95
bio14_Precipitation of driest month	0.9
bio15_Precipitation seasonality	0.56
bio16_Precipitation of wettest quarter	0.86
bio17_Precipitation of driest quarter	0.87
bio18_Precipitation of warmest quarter	0.7
bio19_Precipitation of coldest quarter	0.73

tmin_var	PGLS_pval
Minimum temperature_01	<mark>0.02</mark>
Minimum temperature_02	<mark>0.02</mark>
Minimum temperature_03	<mark>0.04</mark>
Minimum temperature_04	0.08
Minimum temperature_05	0.13
Minimum temperature_06	0.18
Minimum temperature_07	0.57
Minimum temperature_08	0.6
Minimum temperature_09	0.09
Minimum temperature_10	<mark>0.03</mark>
Minimum temperature_11	<mark>0.03</mark>
Minimum temperature_12	<mark>0.02</mark>

Cold adaption is correlated with relaxation of selection (all median K<1, 0.21~0.29).</p> "Relaxation of selective strength can drive evolutionary innovation ... " Selection can be relaxed from the removal of an existing selective constraint ...."

---- Wertheim et al., 2014

### **RELAX: Detecting Relaxed Selection in a Phylogenetic** Framework

Joel O. Wertheim,<sup>\*1</sup> Ben Murrell,<sup>1</sup> Martin D. Smith,<sup>2</sup> Sergei L. Kosakovsky Pond,<sup>1</sup> and Konrad Scheffler<sup>\*,1,3</sup> <sup>1</sup>Department of Medicine, University of California, San Diego <sup>2</sup>Bioinformatics and Systems Biology Graduate Program, University of California, San Diego

<sup>3</sup>Department of Mathematical Sciences, Stellenbosch University, Stellenbosch, South Africa

\*Corresponding author: E-mail: jwertheim@ucsd.edu; kscheffler@ucsd.edu.

Associate editor: Beth Shapiro

#### Abstract

# Relaxed selection is a precursor to the evolu <sup>Selife</sup> phenotypic plasticity

Brendan G. Hunt<sup>a,1</sup>, Lino Ometto<sup>b,1,2</sup>, Yannick Wurm<sup>b,c,3</sup>, DeWayne Shoemaker<sup>d</sup>, Soojin V. Yi<sup>a</sup>, Lau and Michael A. D. Goodisman<sup>a,4,5</sup>

<sup>a</sup>School of Biology, Georgia Institute of Technology, Atlanta, GA 30332; <sup>b</sup>Department of Ecology and Evolution, University of L Lausanne, Switzerland; <sup>c</sup> Vital-IT Group, Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland; and <sup>d</sup>Center for Medical, Entomology, United States Department of Agriculture- Agricultural Research Service, Gainesville, FL 32608

Edited by Michael Lynch, Indiana University, Bloomington, IN, and approved August 16, 2011 (received for review March 30, 20

Phenotypic plasticity allows organisms to produce alternative phenotypes under different conditions and represents one of the

with differences in expression associate (10, 16–19). However, few studies have relation and differentially expression 2 October 31, 2023 (this version)

Trends in Genetics



Opinion

# Relaxed Selection and the Rapid Evolution of Reproductive Genes

Amy L. Dapper<sup>1,2,3,\*</sup> and Michael J. Wade<sup>2</sup>

Evolutionary genomic studies find that reproductive protein genes, those directly involved in reproductive processes, diversify more rapidly than most other gene categories. Strong postcopulatory sexual selection acting within species is the predominant hypothesis proposed to account for the observed pattern.

#### Highlights

Reproductive protein genes evolve more rapidly than almost any other gene category.

#### Evolutionary Biology Plant Biology

### Positive selection and relaxed purifying selection contribute to rapid evolution of male-biased genes in a dioecious flowering plant

#### Lei Zhao, Wei Zhou, Jun He, De-Zhu Li 🎴 , Hong-Tao Li 🗳

Germplasm Bank of Wild Species, Yunnan Key Laboratory of Crop Wild Belatives Omics, Kunming Institute of Go to page 29 Botany, Chinese Academy of Sciences, Kunming, Yunnan Occur, Crima - Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan 650201, China

d https://en.wikipedia.org/wiki/Open\_access

Reviewed preprint version 1 September 1, 2023

**Reviewed Preprint** 

About eLife's process

review.

Revised by authors after peer

© Copyright information

Set "A. saccharum, A. mandshuricum, A. tegmentosum" as foreground and the other species as background,

Test relaxation of selection using RELAX

k>1 indicates that selection strength has been intensified;

k<1 indicates that selection strength has been relaxed

Cold adaption is correlated with relaxation of selection (all median K<1, 0.21~0.29).



### My interpretation:

- A. saccharum, A. mandshuricum, A. tegmentosum were originated in other regions, possibility warm/hot regions.
- At some historical point, these species migrated from their ancient niche to their current habitat, so the original constraint were released.
- > Thus, relaxation of selection was applied to help cold adaptation.

2) Detecting genes for positive selection and selection intensity

➢ 155 out of 2828 OGs have significant relaxed selection (K<1) on test lineages</p>

Check the GO term enrichment in these identified OGs.



Set "A. saccharum, A. mandshuricum, A. tegmentosum" as foreground and the other species as background,

> Test relaxation of selection using RELAX

k>1 indicates that selection strength has been intensified;

k<1 indicates that selection strength has been relaxed

Cold adaption is correlated with relaxation of selection (all median K<1, 0.21~0.29).



### My interpretation:

- A. saccharum, A. mandshuricum, A. tegmentosum were originated in other regions, possibility warm/hot regions.
- At some historical point, these species migrated from their ancient niche to their current habitat, so the original constraint were released.
- > Thus, relaxation of selection was applied to help cold adaptation.

### Relaxation of selection involves genes controlling epigenetic regulation:

DNA and histone modifications and higher-order chromatin structures through alternative polyadenylation (mRNA 3'-End Processing) in flower development genes.

2) Detecting genes for positive selection and selection intensity

Set "A. saccharum, A. mandshuricum, A. tegmentosum" as foreground and the other species as background,

Test relaxation of selection using RELAX k>1 indicates that selection strength has been intensified; k<1 indicates that selection strength has been relaxed</p>

Test positive selection using PAML

> 155 out of 2828 OGs have significant relaxed selection (K<1) on test lineages

> 137 out of 2828 OGs have significant intensified selection (K>1) on test lineages

 $\succ$  109 out of 3279 OGs were tested positive selection on test lineages.

Check the GO term enrichment in the identified intensively selected and positively selected OGs.

2) Detecting genes for positive selection and selection intensity

## GO term enrichment

	Intensified selection	Positive selection
Biological process	<ul> <li>transcription by RNA polymerase I</li> <li>glucose metabolic process</li> <li>purine-containing compound metabolic process</li> <li>ATP transport</li> <li>intracellular protein transport</li> <li>nitrogen compound transport</li> </ul>	<ul> <li>endoplasmic reticulum unfolded protein response</li> <li>cellular response to unfolded protein</li> <li>immune response</li> <li>defense response</li> <li>glucose metabolic process</li> <li>intracellular chemical homeostasis</li> <li>phosphate ion transport</li> </ul>
Molecular function	<ul> <li>ATP:ADP antiporter activity</li> <li>RNA polymerase I activity</li> <li>electron transfer activity</li> </ul>	monosaccharide binding
Cellular component	<ul> <li>RNA polymerase III complex</li> <li>RNA polymerase I complex</li> </ul>	<ul> <li>Mitochondrial inner membrane complex</li> <li>heterotrimeric G-protein complex</li> <li>extrinsic component of membrane</li> <li>nitrite reductase complex</li> </ul>





Phylotranscriptomics-based method to study the genetic basis of climate adaptation in Acers





Phylotranscriptomics-based method to study the genetic basis of climate adaptation in Acers





> Phylogeny-based method is efficient to identify orthologs related to climate adaptation.

Relaxation of selection in genes involved in epigenetic regulation is the main mechanism for cold adaption in *Acers*.

Caution: Genetic variation that arise from gene duplication (paralogs) were not considered in the present study.

# Thank you!

