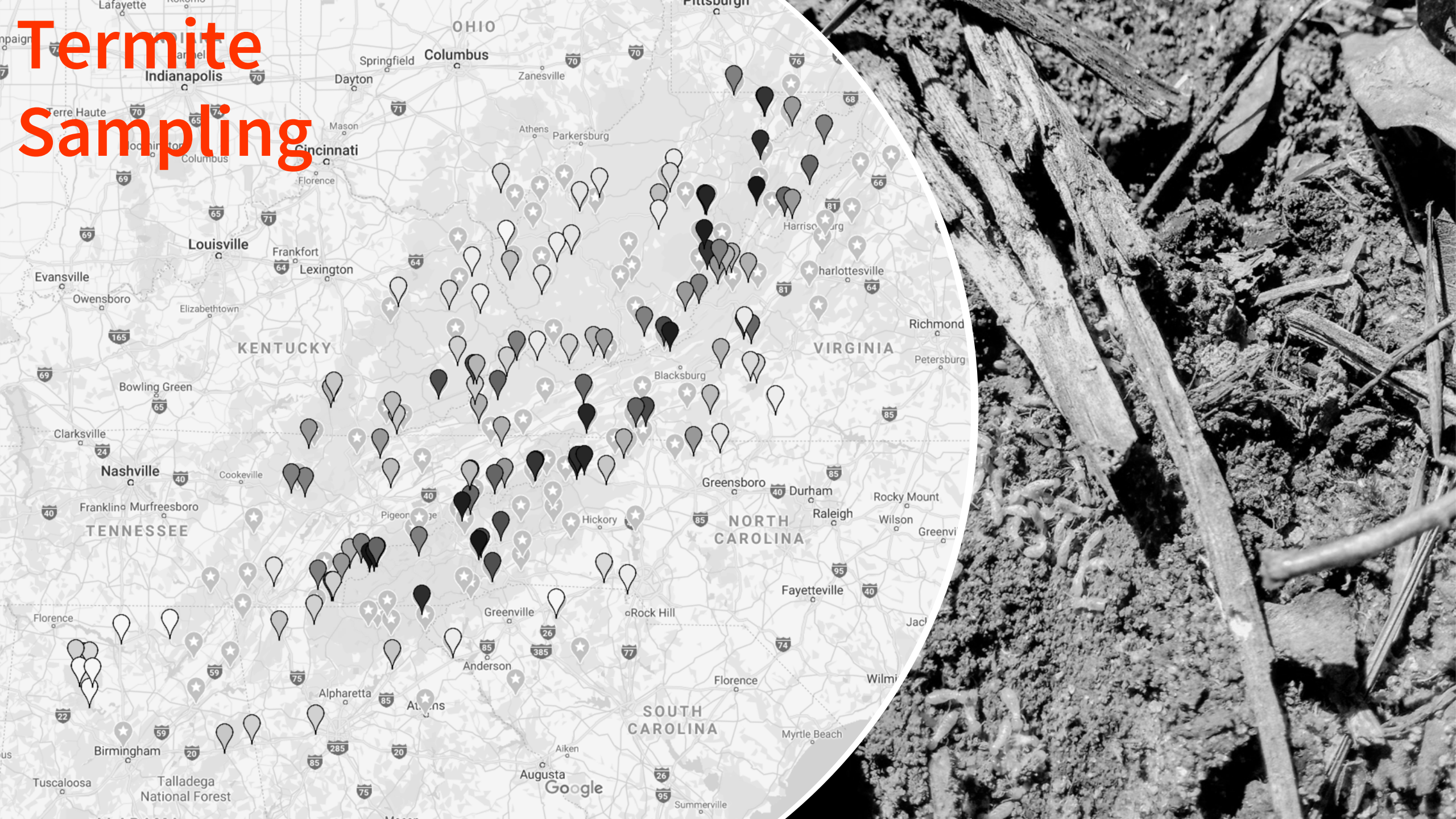
Epigenetic Change in the Eastern Subterranean Termite

Modulated by Ecological Disturbance in Appalachian Forest Ecosystems

Chaz Hyseni

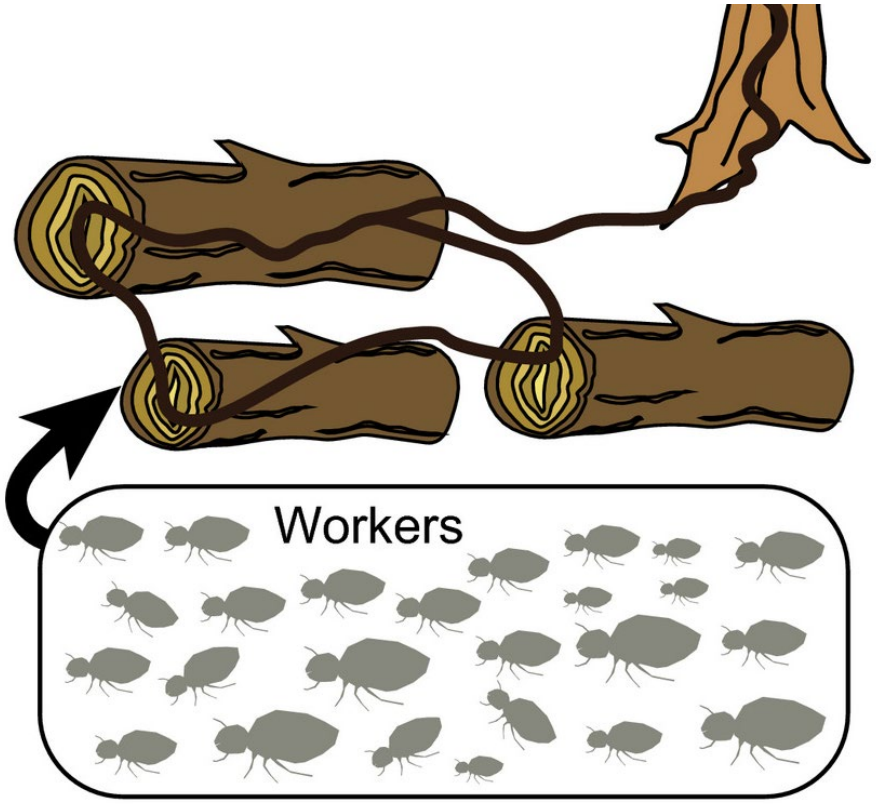
SpEcEvGen http://specevgen.net

Termite Sampling



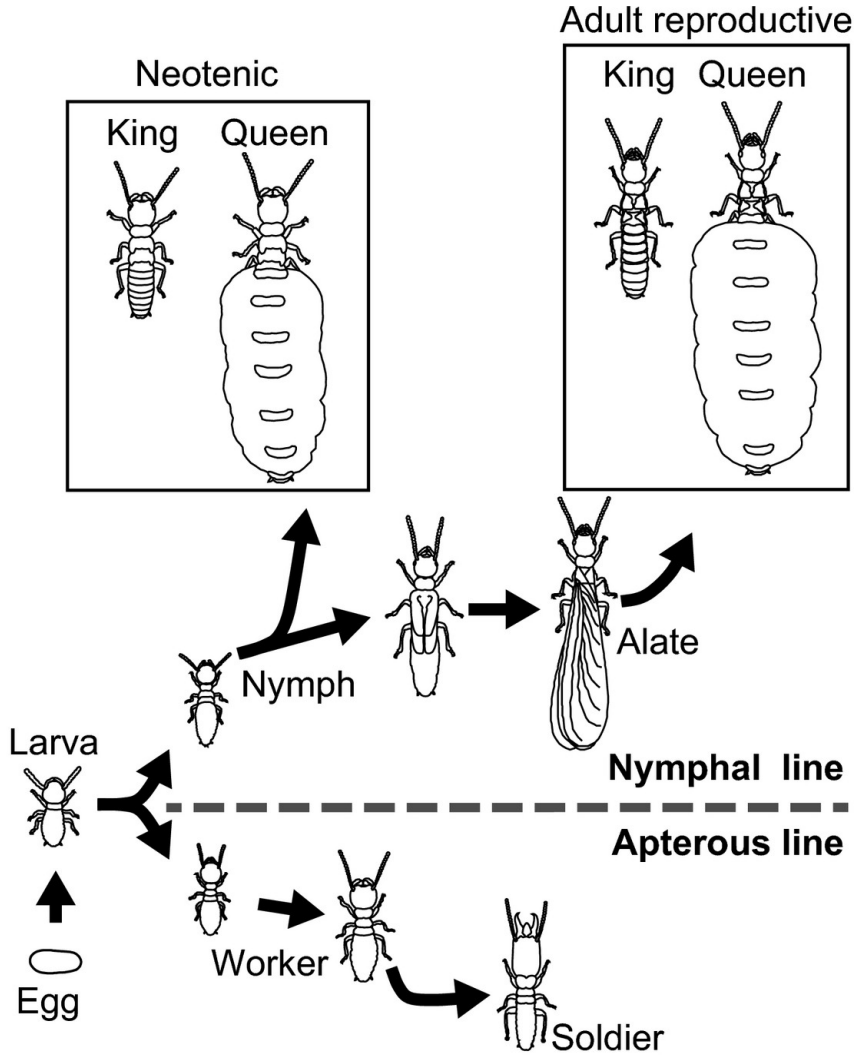
Subterranean Termites (*Reticulitermes*)

Ecosystem Engineering through a Foraging Lifestyle



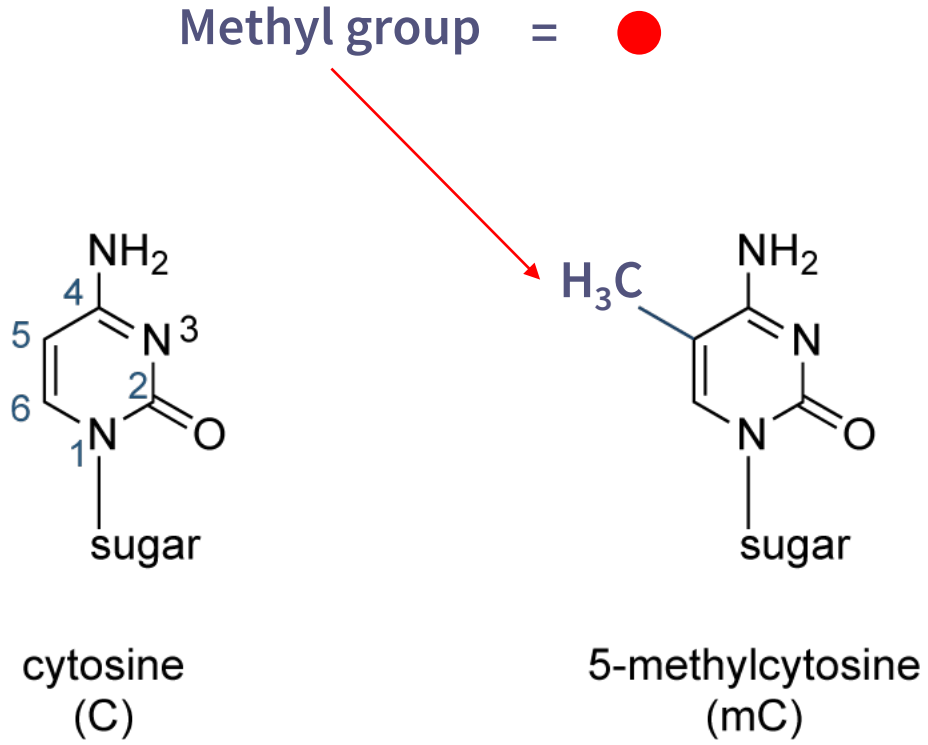
Nozaki & Matsuura. *Ecol. Evol.* 2019

Subterranean Termites: Phenotypic Plasticity



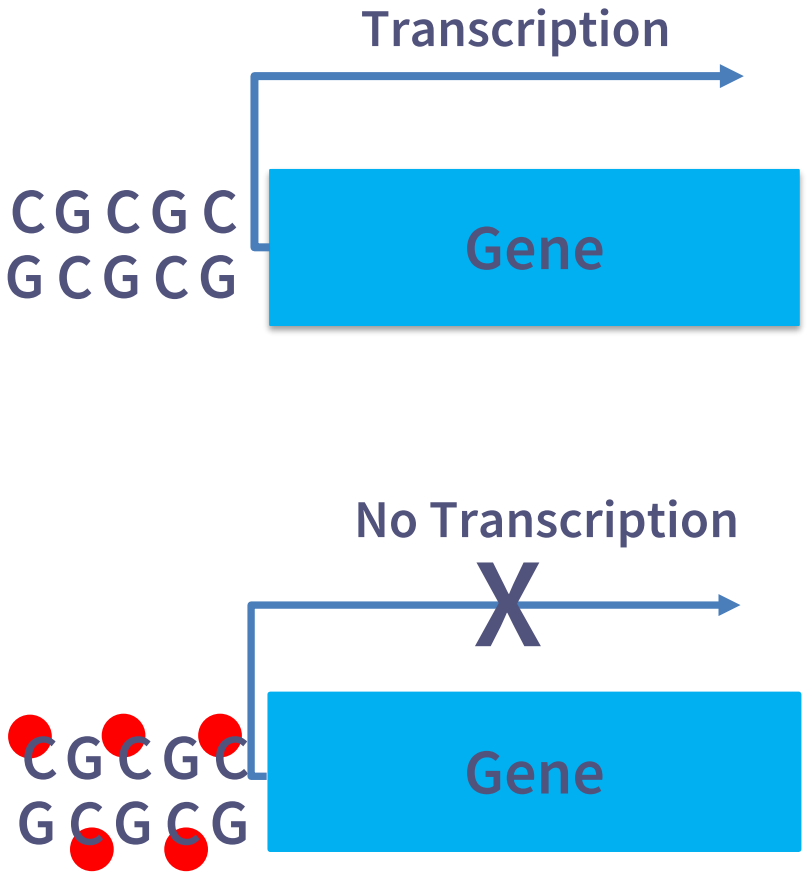
Nozaki & Matsuura. *Ecol. Evol.* 2019

Epigenetics: DNA methylation

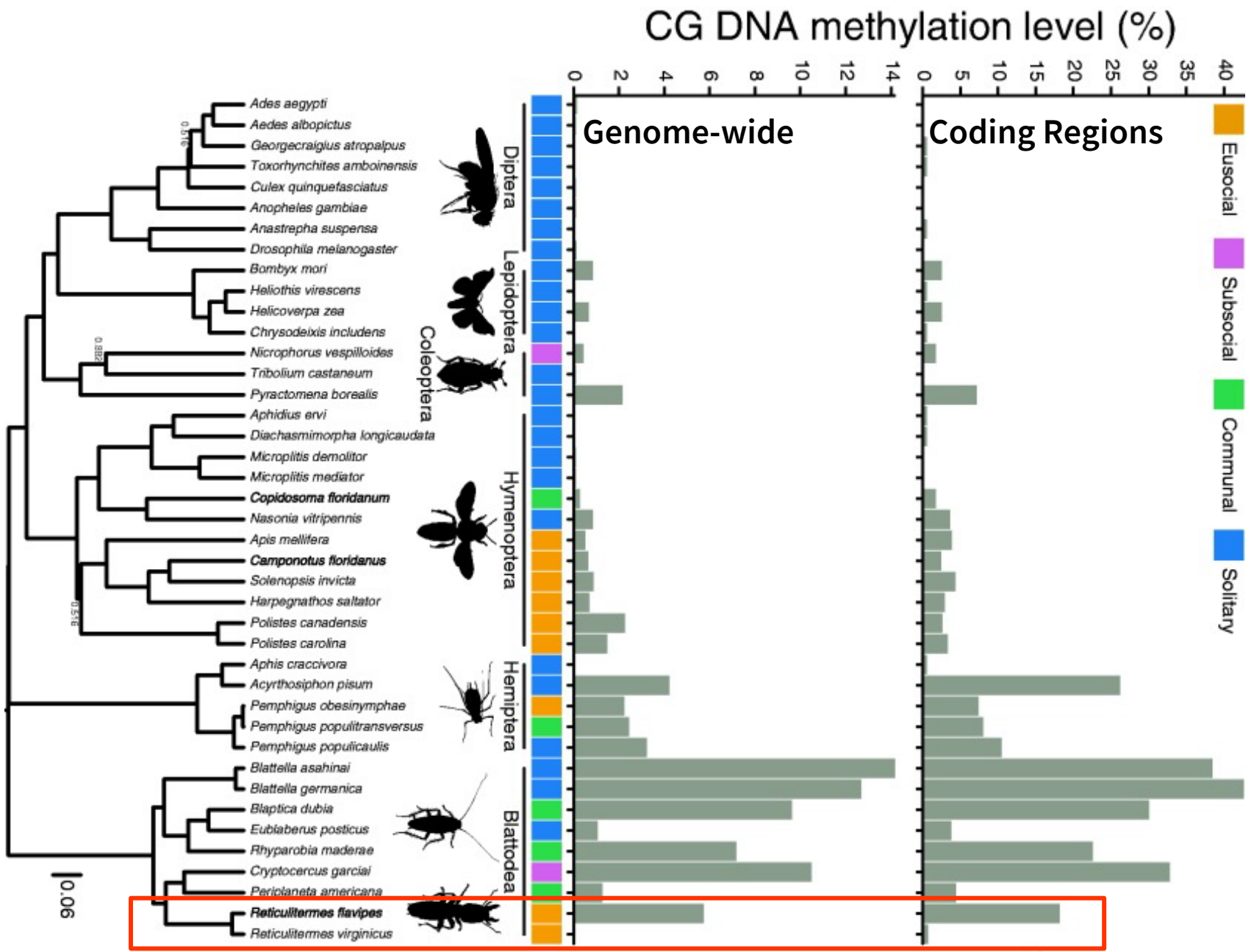


cytosine
(C)

5-methylcytosine
(mC)



Epigenetic Machinery of *Reticulitermes*

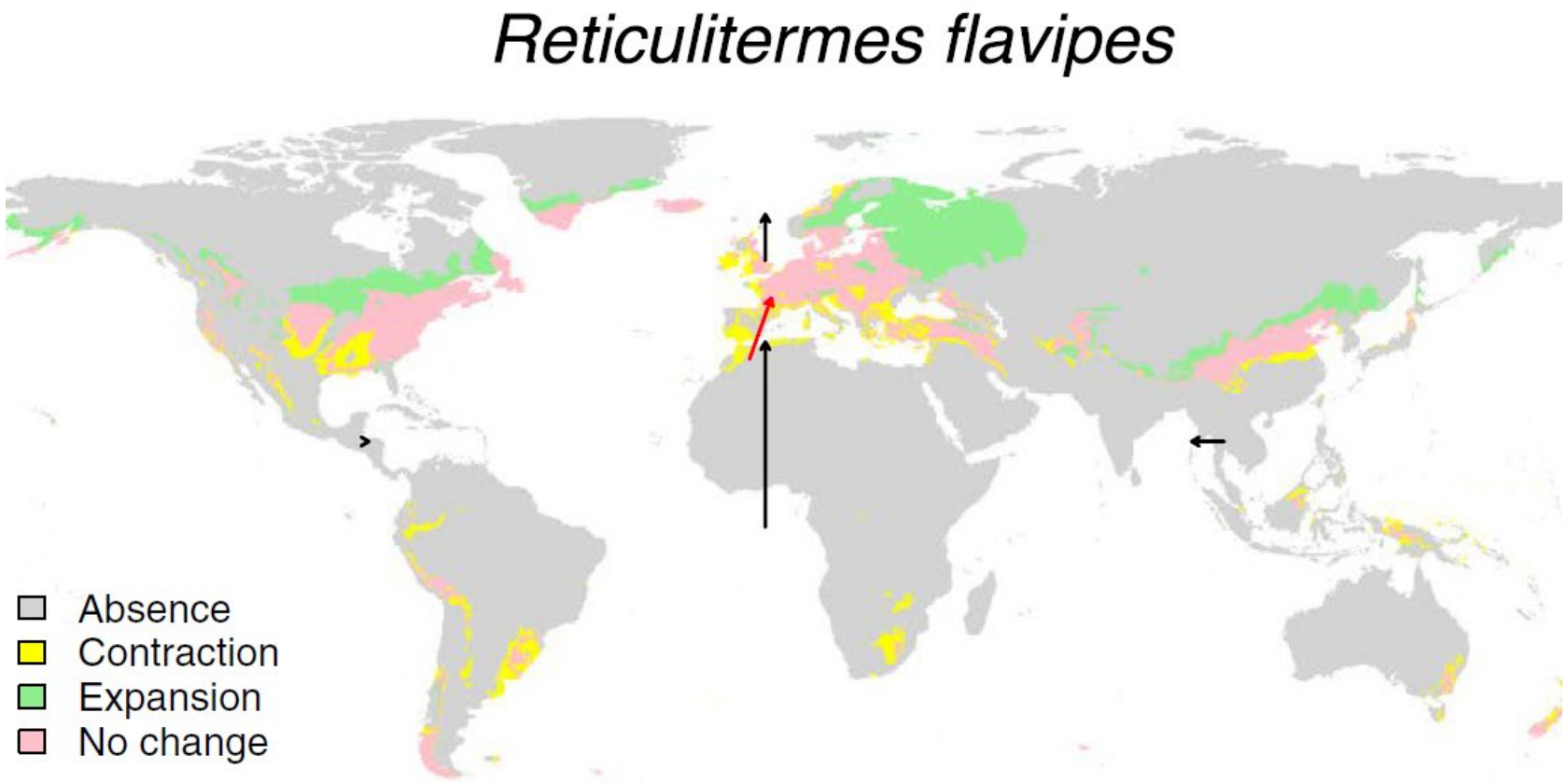


Bewick et al. *Mol. Biol. Evol.* 2017

R. flavipes is a pest in N. America and invasive elsewhere in the world

Could DNA methylation have something to do with this?

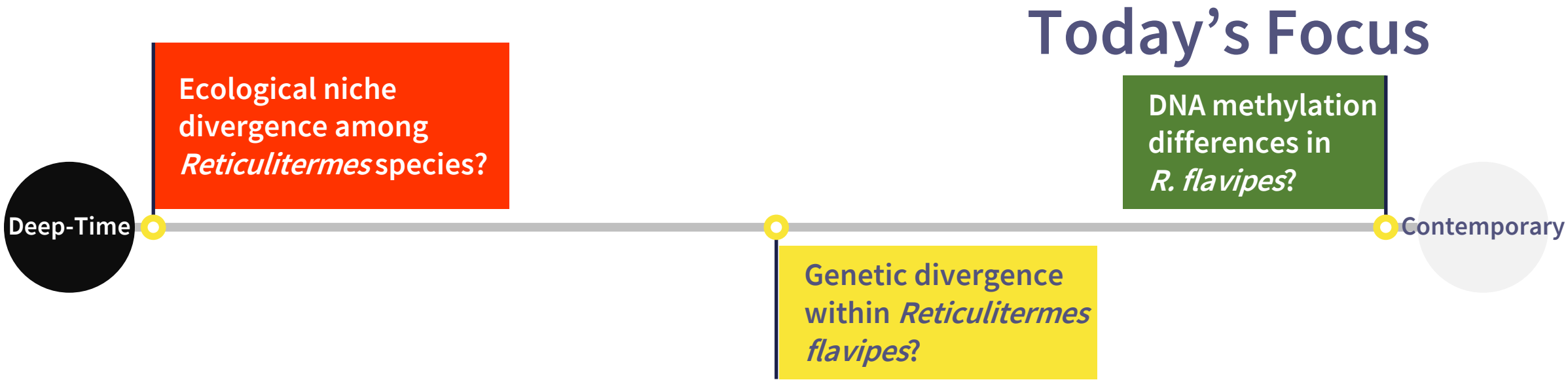
2050
Predicted
Geographic
Distribution



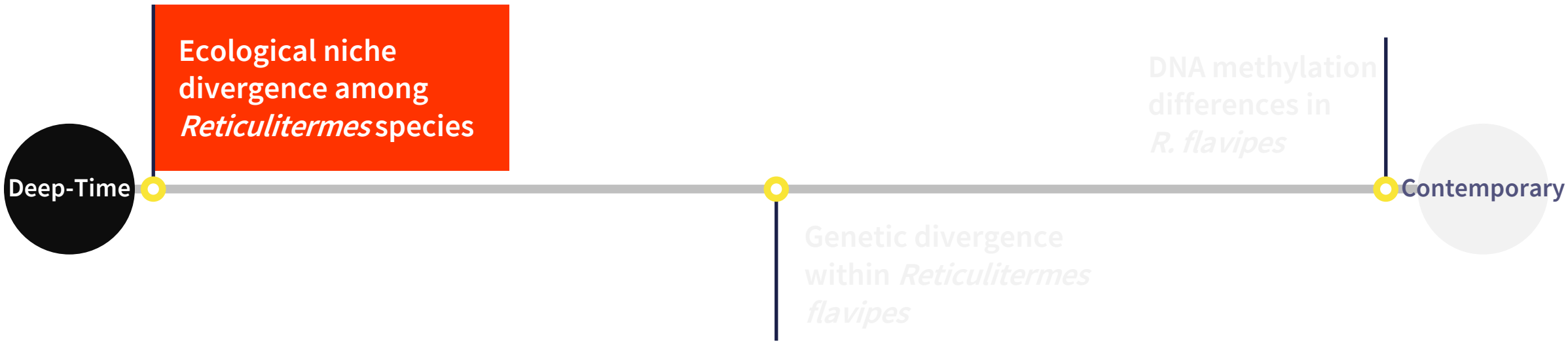
Buczkowski & Bertelsmeier. *Ecol. Evol.* 2017

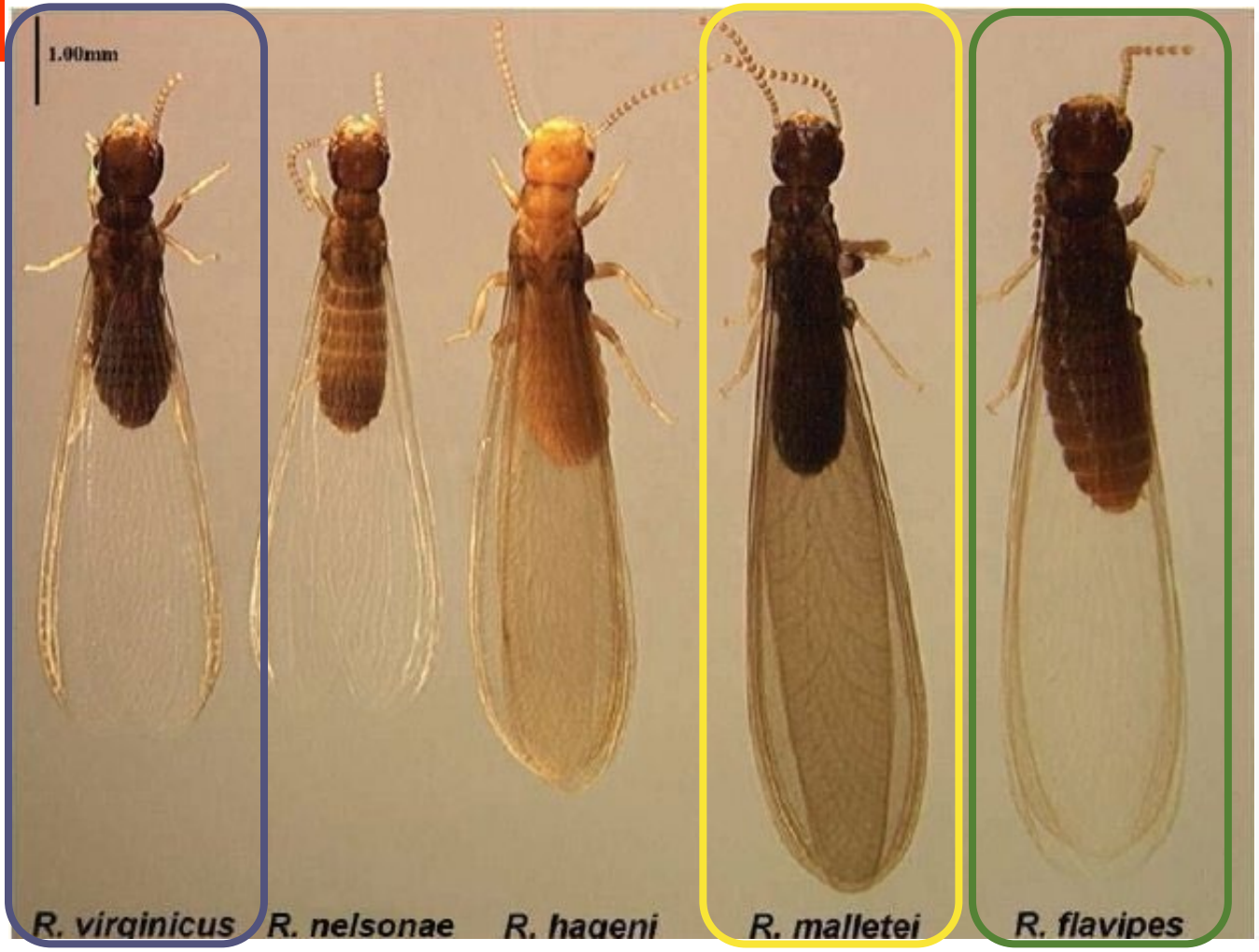
N/S/E/W Shift Overall Shift

Questions at Different Evolutionary Timescales

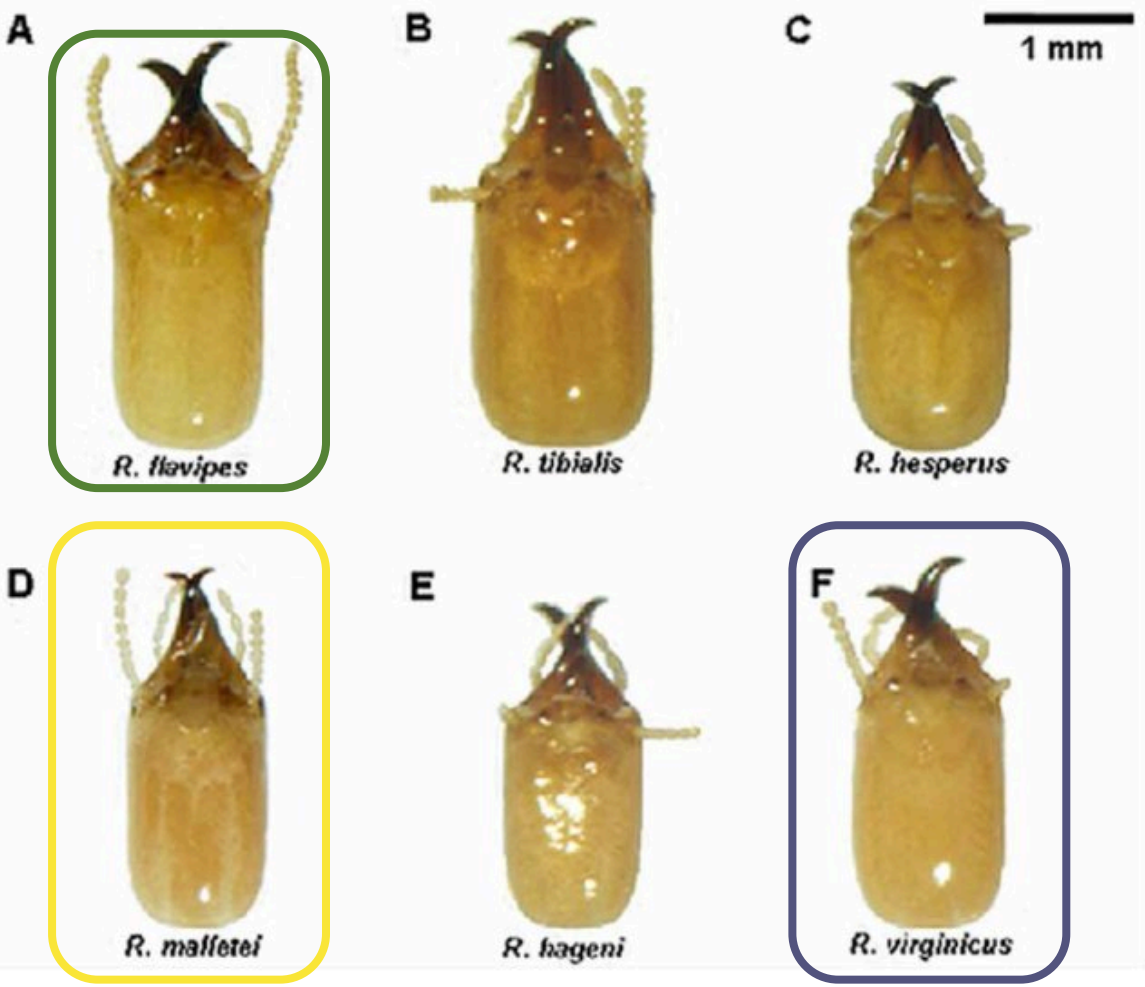


Timescale of Evolution



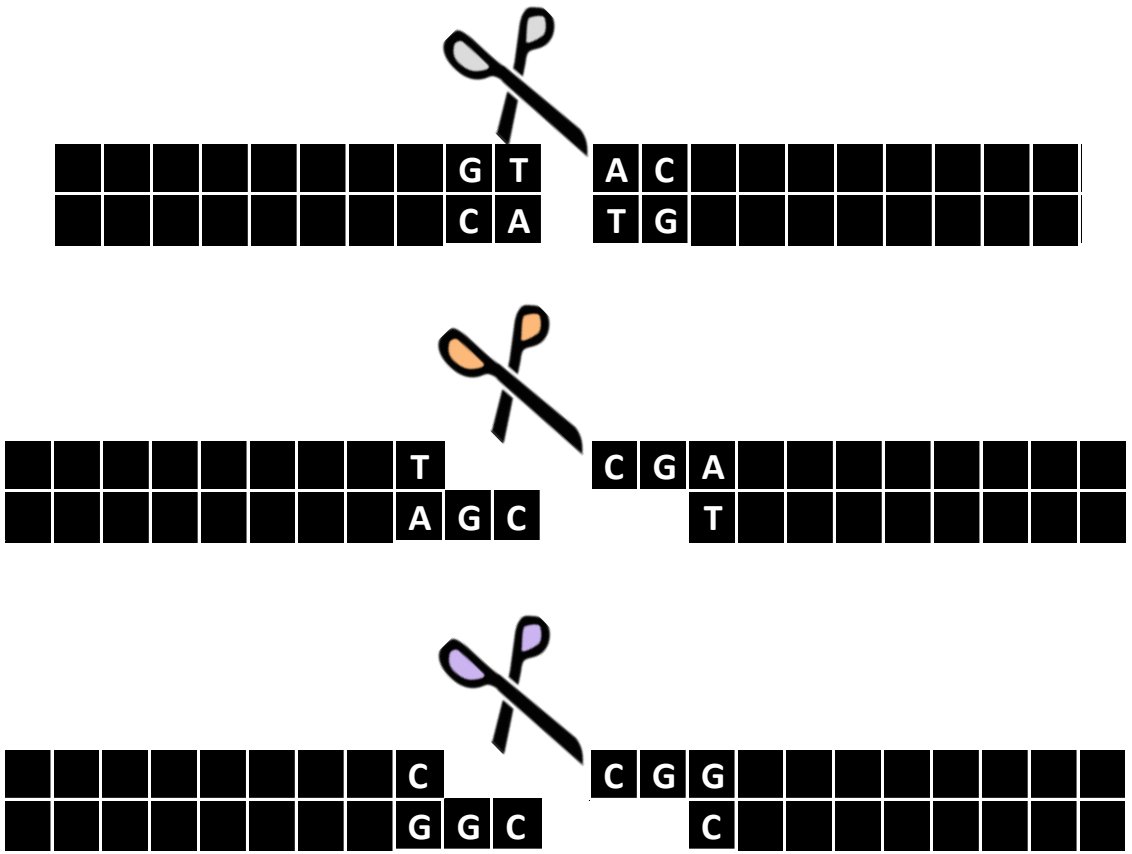


Lim & Forschler. *Insects* 2012



Austin et al. *Zootaxa* 2007

Restriction Digestion



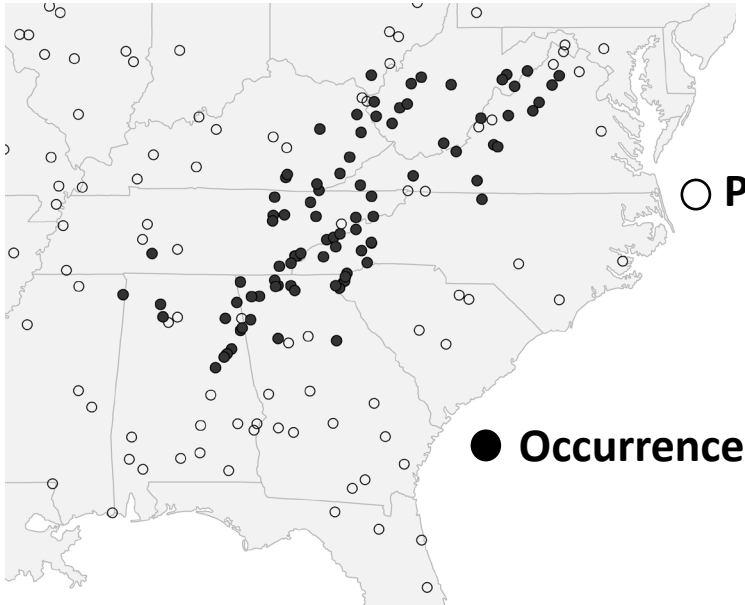
Restriction Enzymes: RsaI TaqI MspI

Digestion of 376 bp fragment of mitochondrial DNA

Restriction Enzyme	Fragment Sizes (bp)	Species				
		R. flavipes	*R. hageni*	*R. malletei*	*R. nelsonae*	*R. virginicus*
Rsa I	175, 201	✓	×	×	✓	✓
	48, 127, 201	×	×	✓	✓	×
	86, 115, 175	✓	✓	×	×	×
Taq I	376	×	✓	✓	✓	✓
	153, 223	×	×	×	✓	×
	183, 193	✓	×	×	×	×
	67, 126, 183	✓	×	×	×	×
	30, 67, 126, 153	✓	×	×	×	×
Msp I	376	✓	×	×	✓	✓
	37, 339 *	✓	×	✓	×	×
	77, 299	×	✓	×	✓	×
	37, 40, 299	×	×	✓	×	×

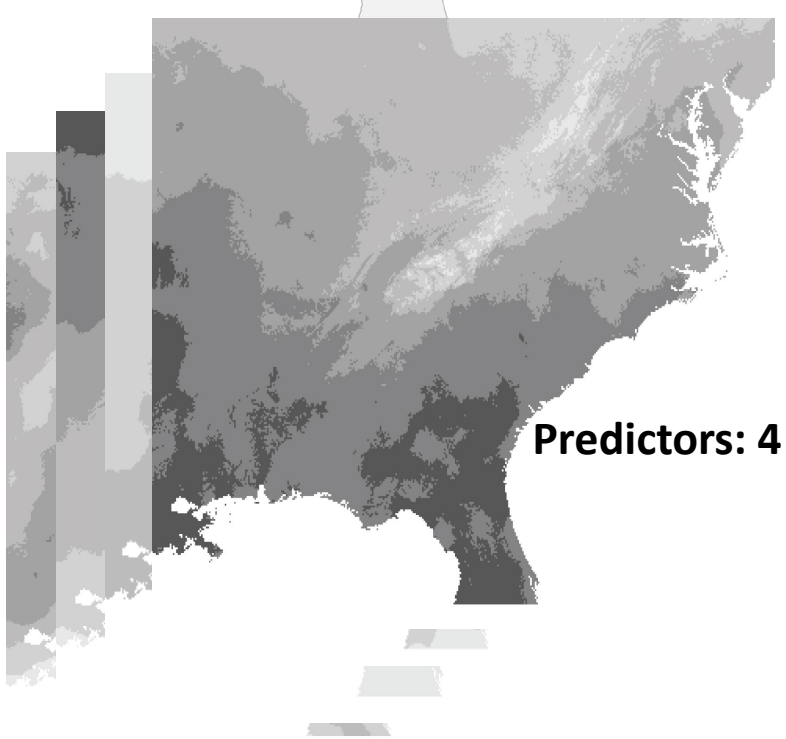
Ecological Niche Modeling:
occurrence data (132) and environmental predictors (4)

R. flavipes = Rf
R. malletei = Rm
R. virginicus = Rv



○ Pseudoabsence: 100

● Occurrence: Rf = 91, Rm = 17, Rv = 30



Predictors: 4

Model Training and Testing

Machine Learning Algorithms:

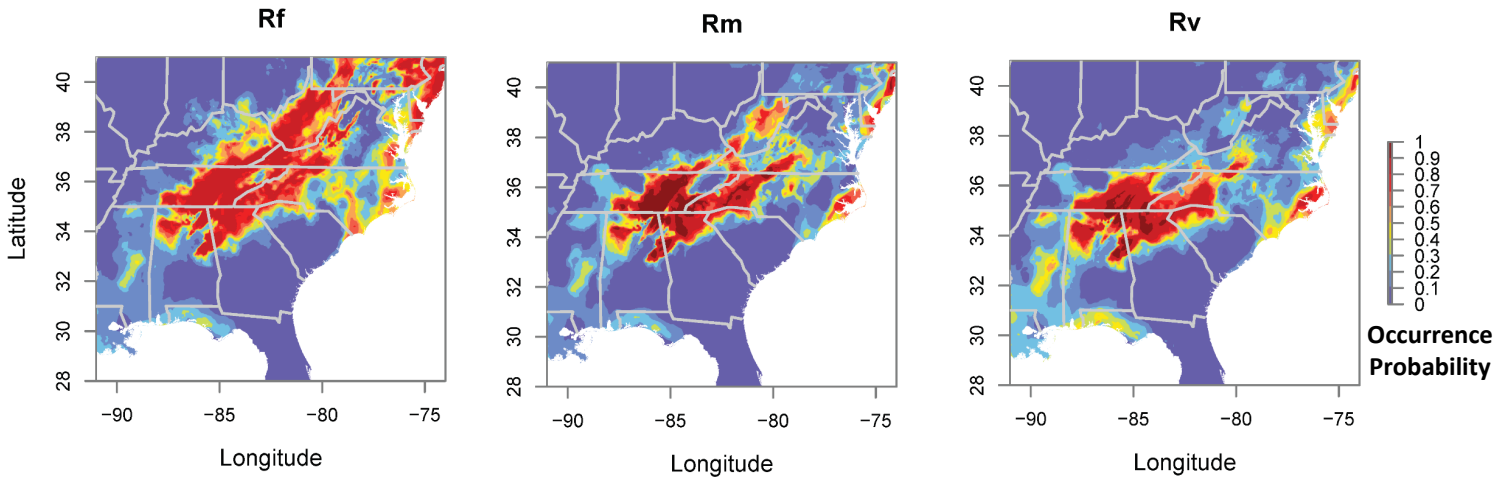
- Random Forest
- Artifical Neural Network
- Generalized Boosted Regression Trees
- Maximum Entropy

x20
20 Sets of
100 Pseudoabsence Points

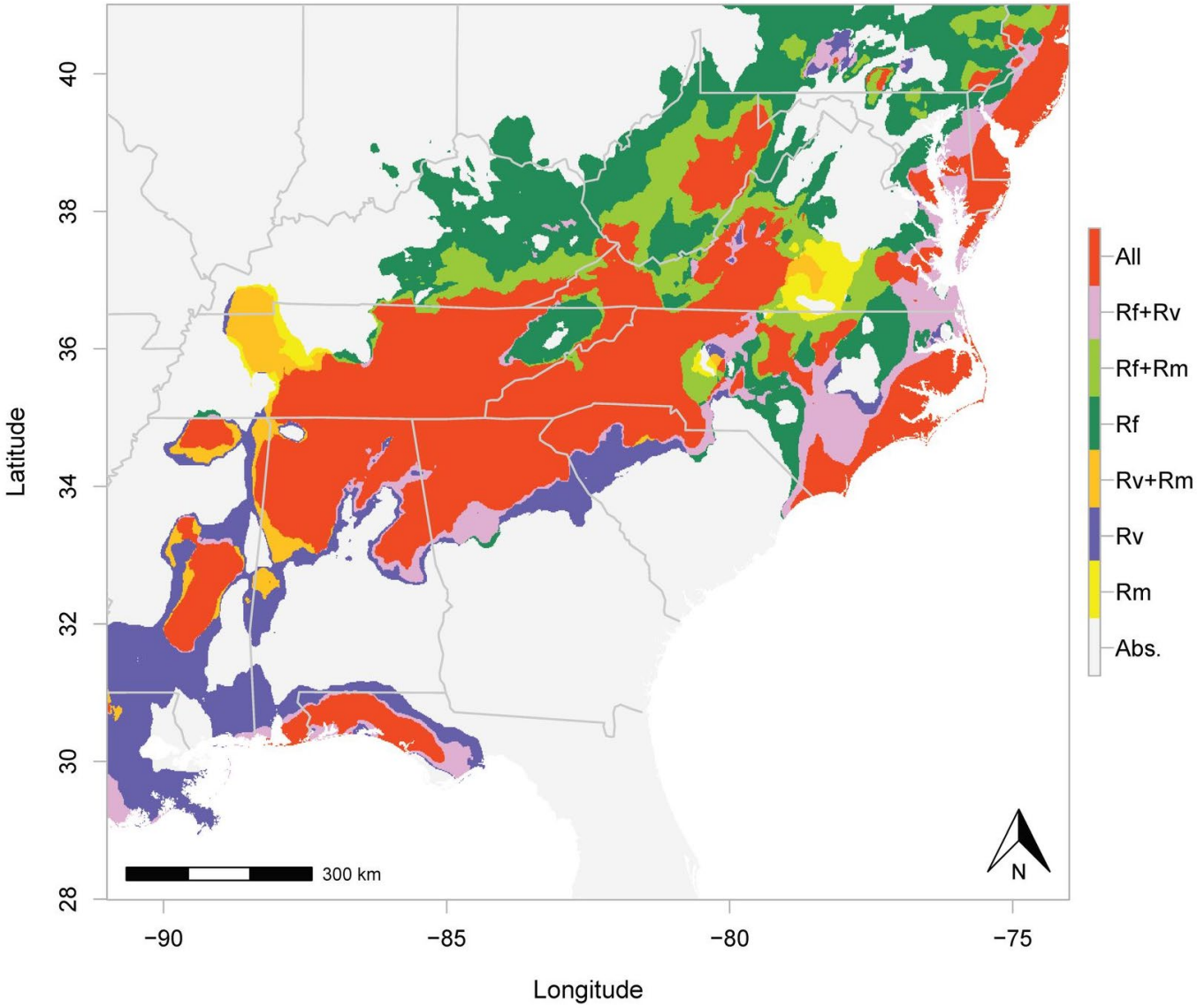
x5

100
Runs

Consensus Projections:

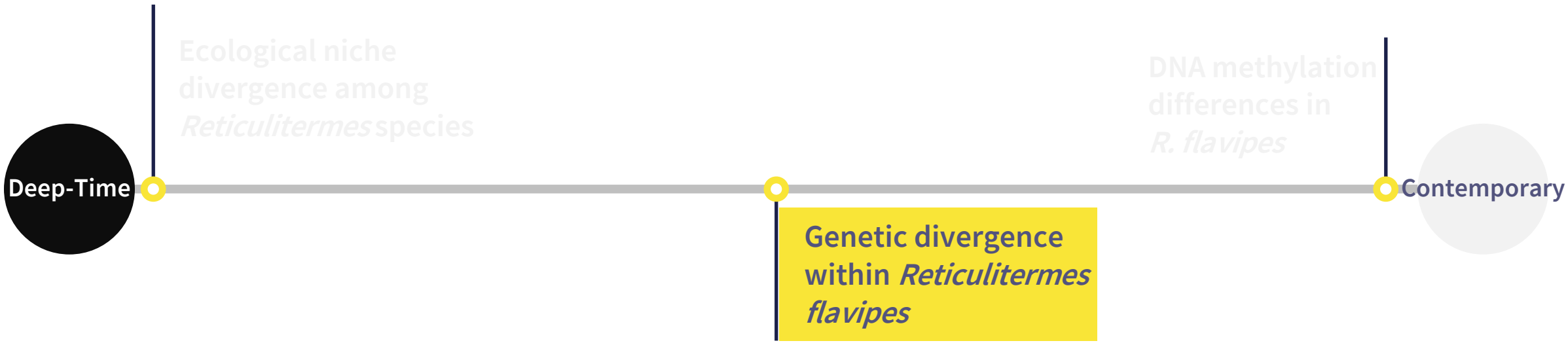


Niche Divergence of *Reticulitermes* Species

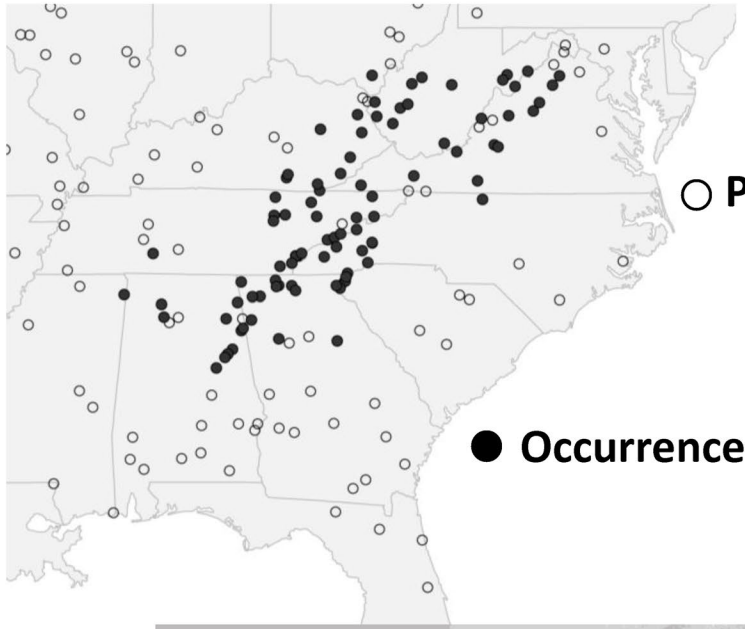


Significant Rf-Rv niche divergence:
Schoener's D statistic = 0.582
***p*-value < 0.01**

Timescale of Evolution

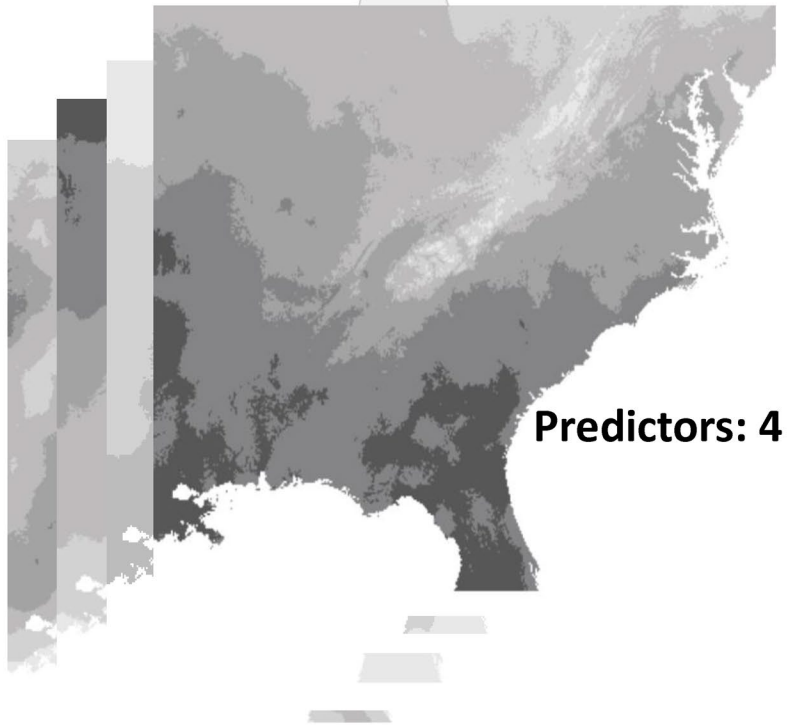


Distribution Modeling



○ Pseudoabsence: 100

● Occurrence: 91



Predictors: 4

Model Training and Testing

x5

Machine Learning Algorithms:

Random Forest

Artifical Neural Network

Generalized Boosted Regression Trees

Maximum Entropy

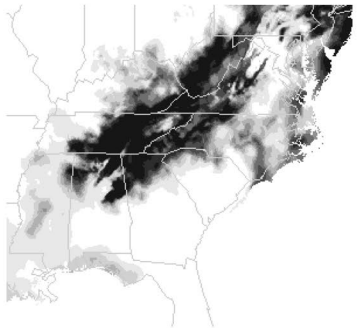
x20

20 Sets of
100 Pseudoabsence Points

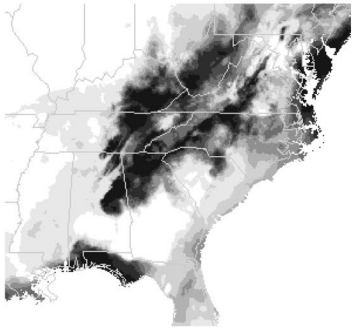
100
Runs

Consensus Projections:

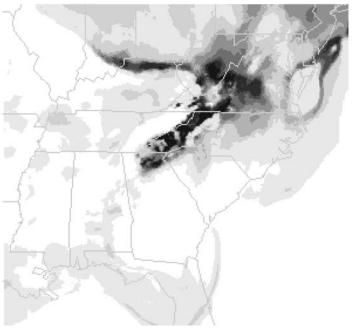
Present-Day



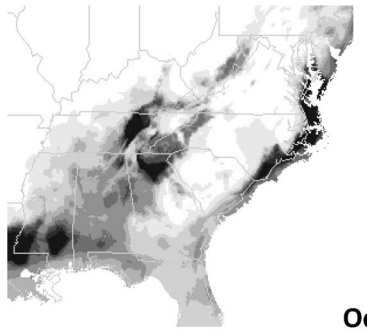
6,000 ya

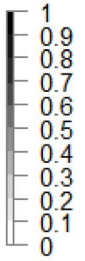


22,000 ya



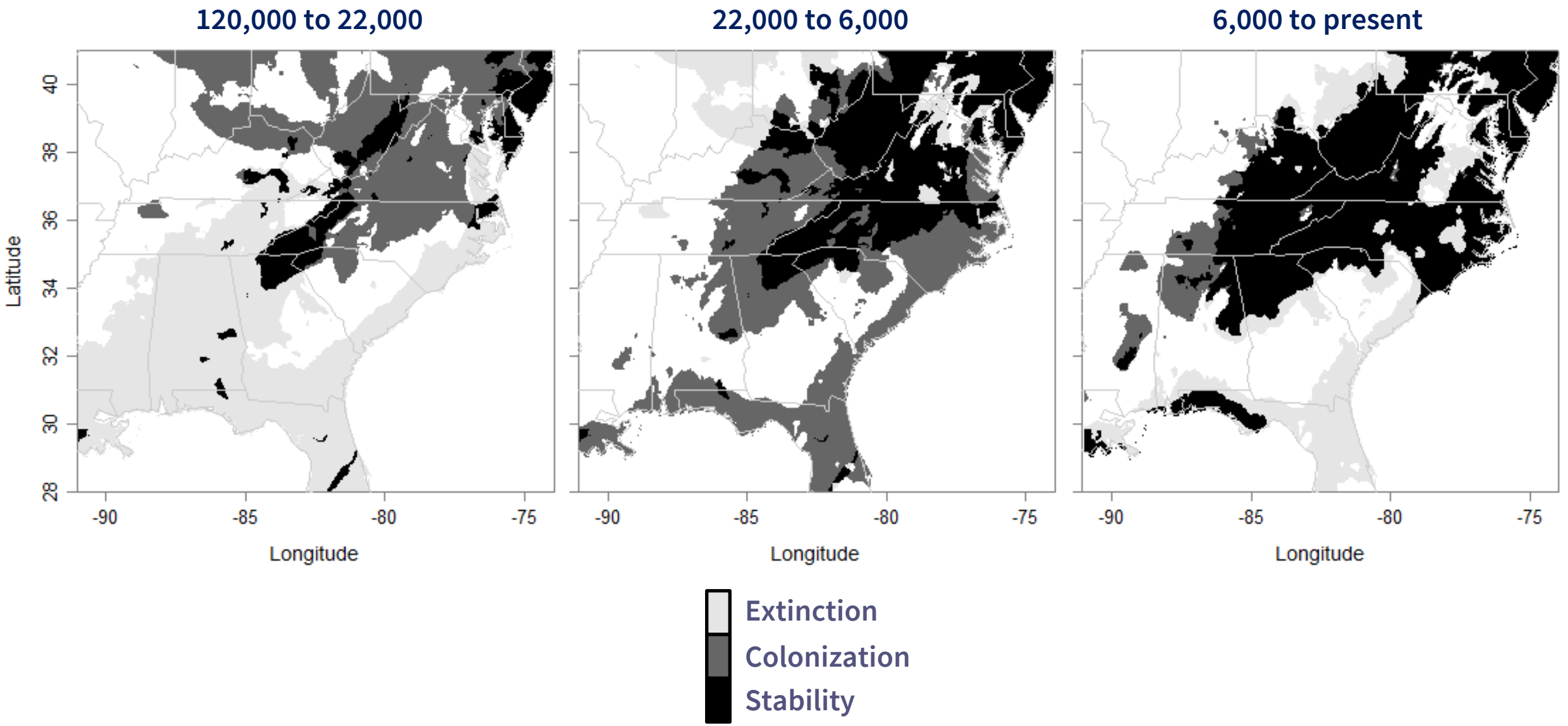
120,000 ya



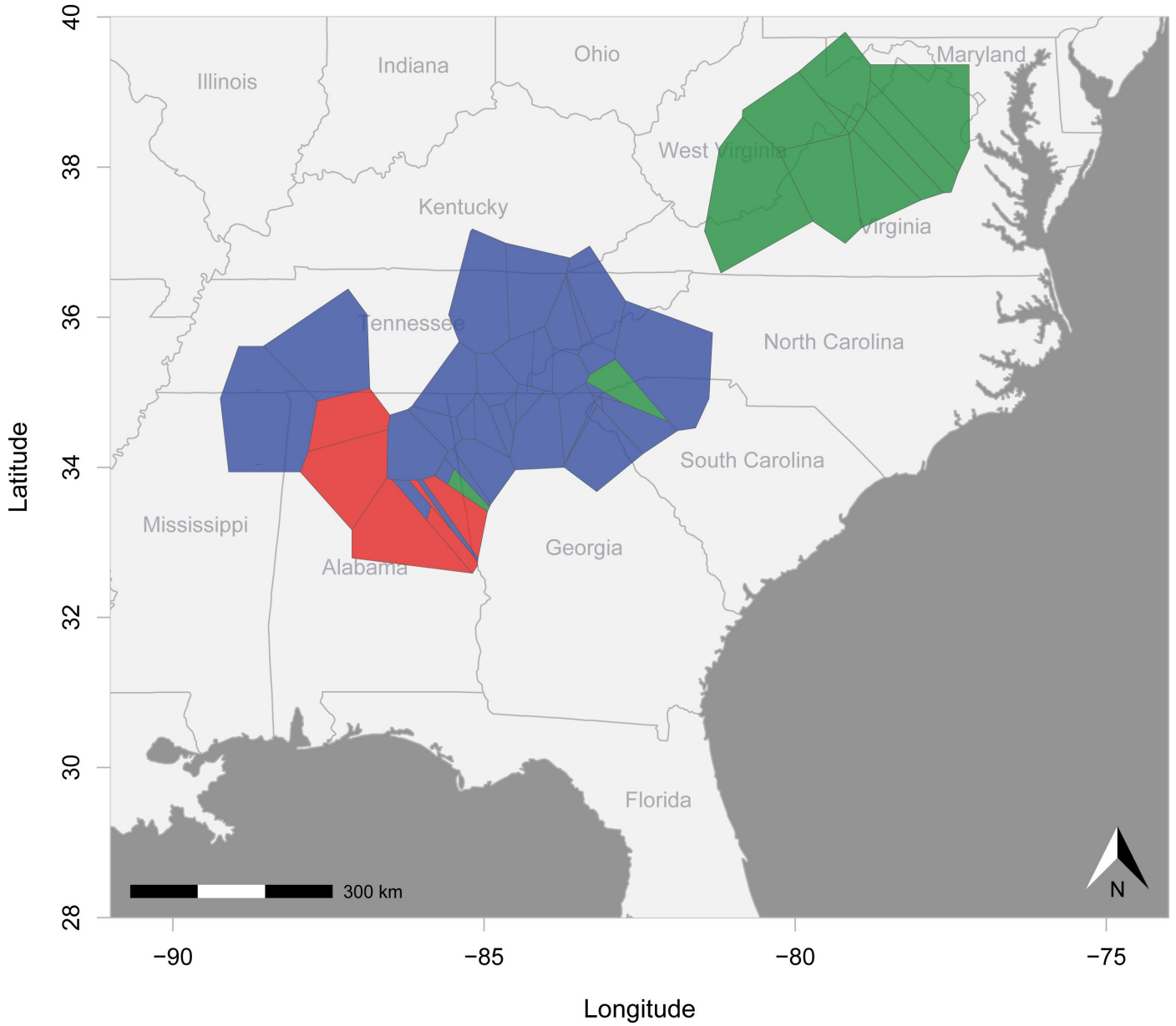


Occurrence
Probability

Paleoclimatic cycles resulted in distributional shifts



Genetic Divergence: Three Lineages

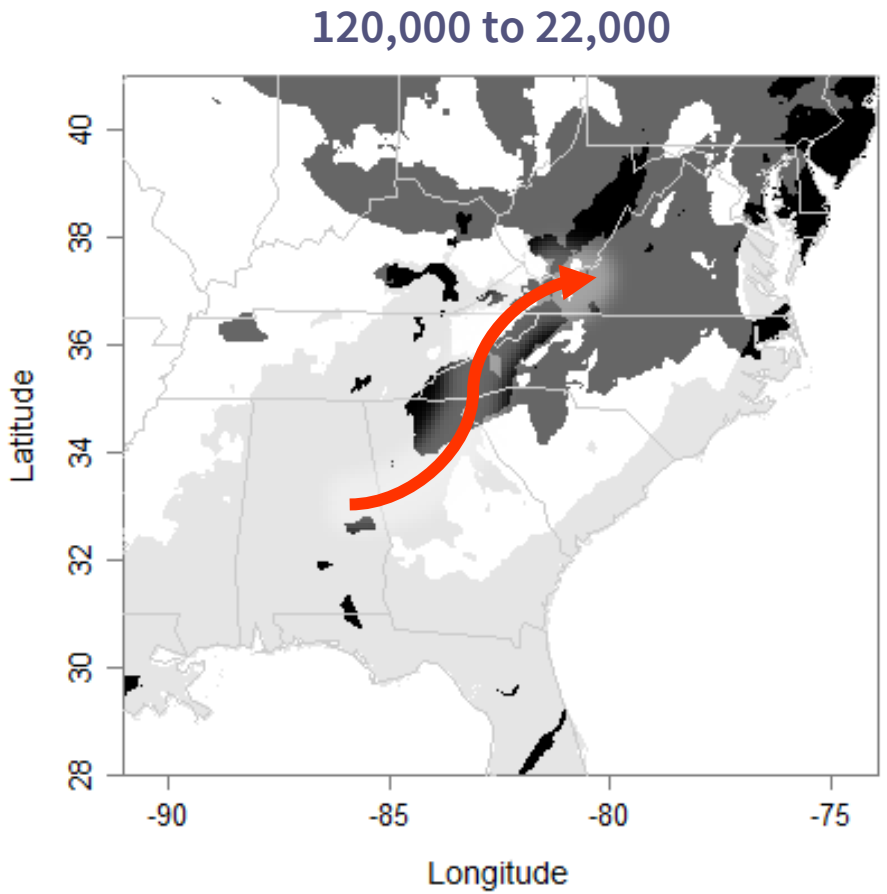


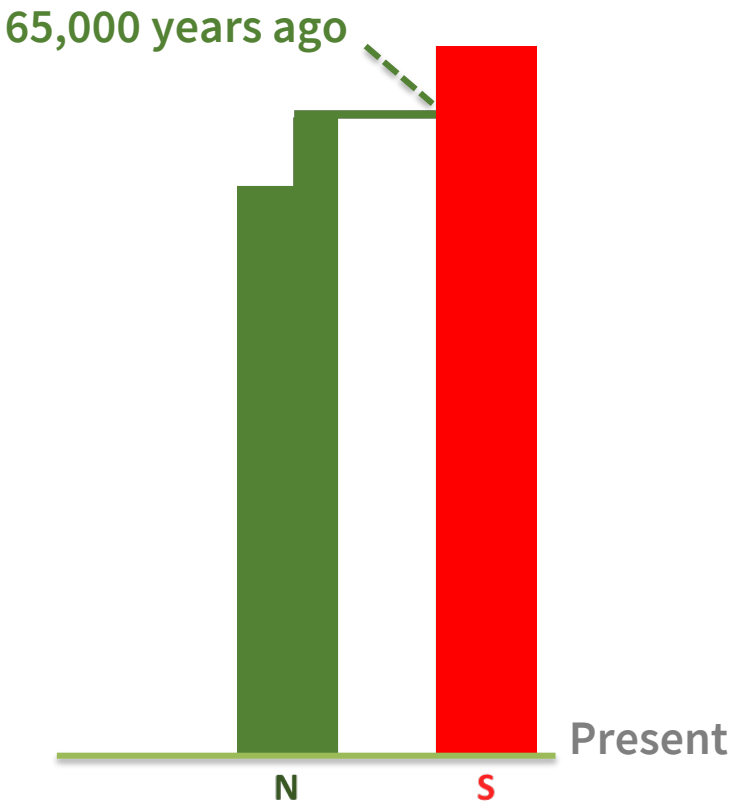
N = Northern

C = Central

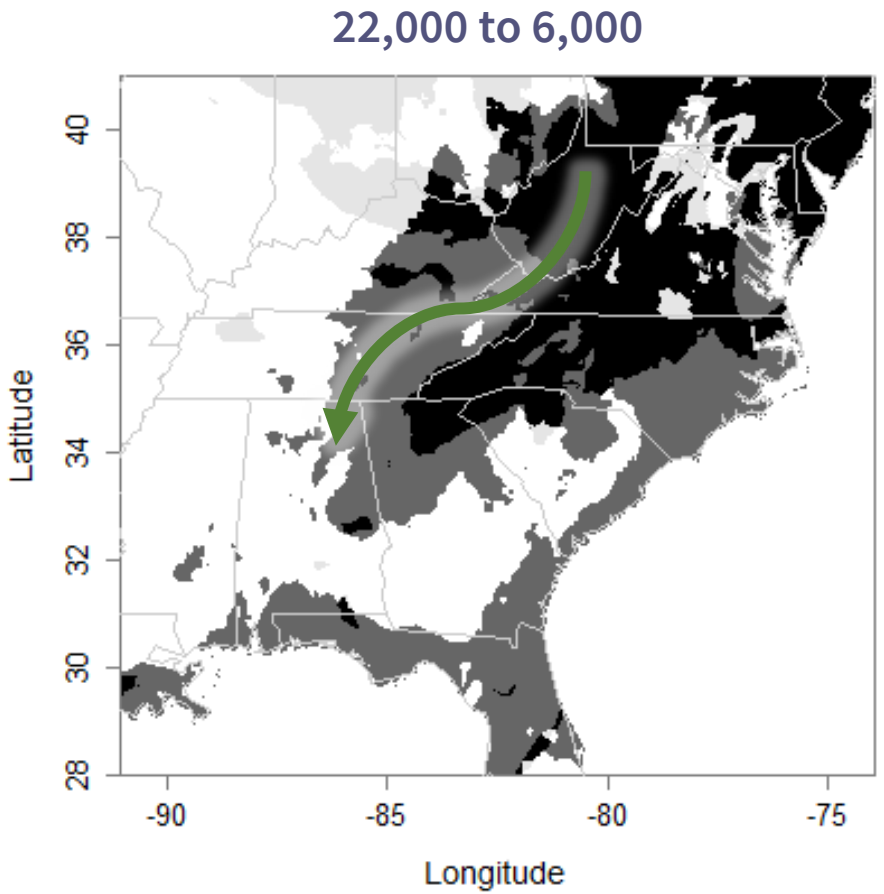
S = Southern

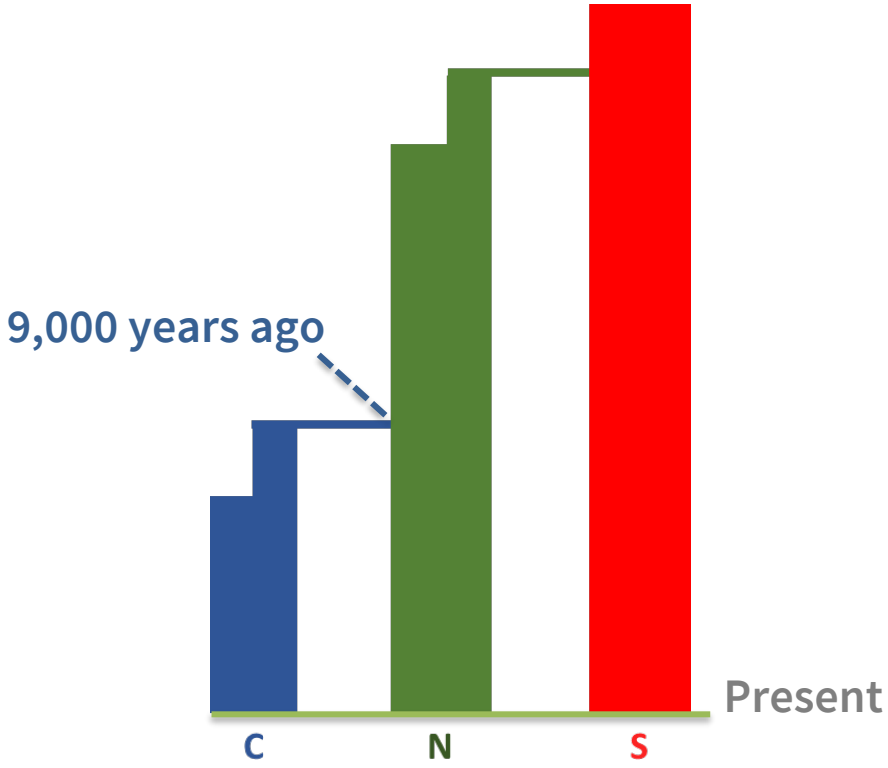
Distributional Shift: South-to-North



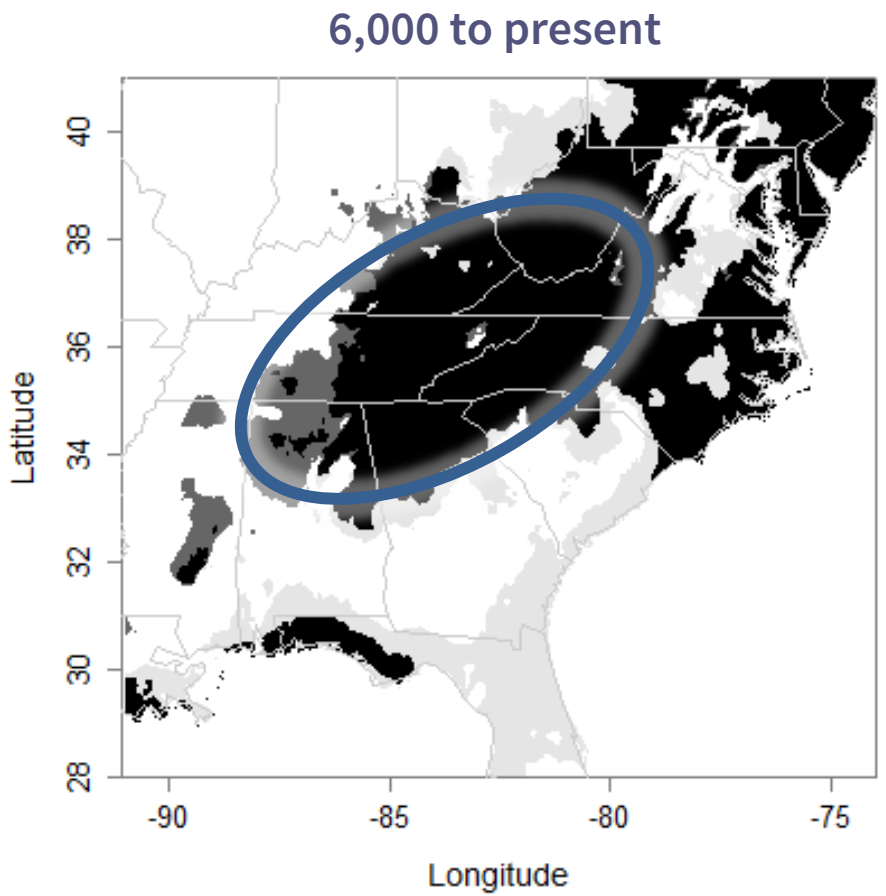


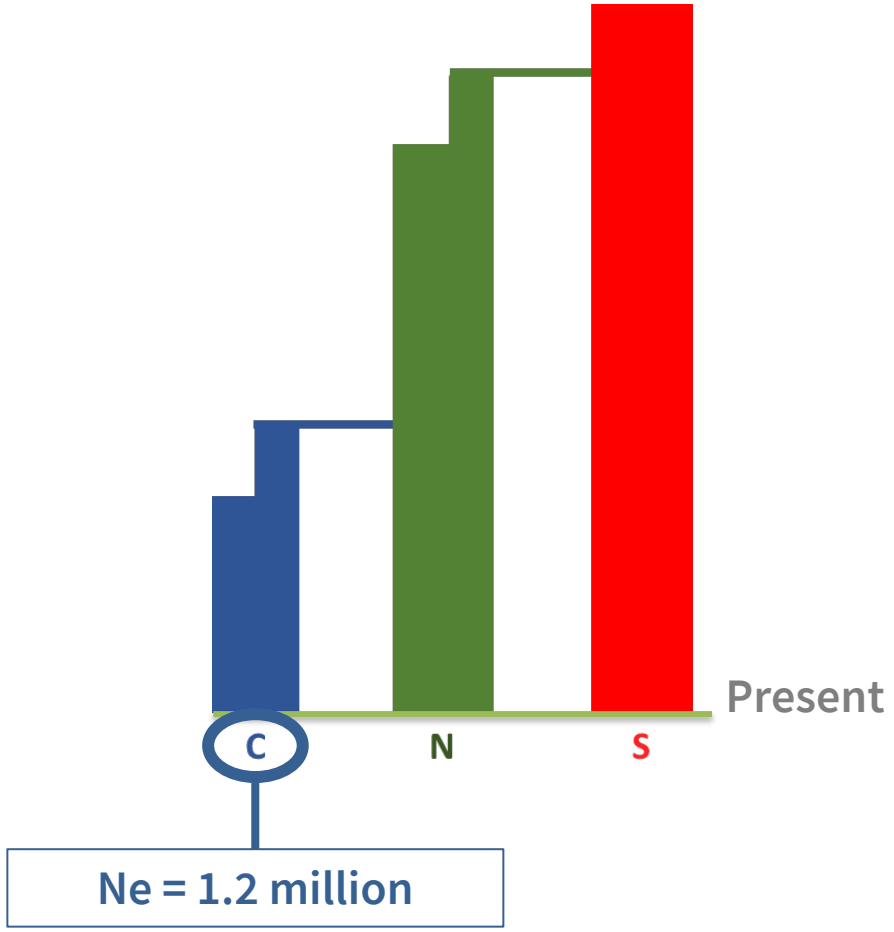
Distributional Shift: North-to-Center



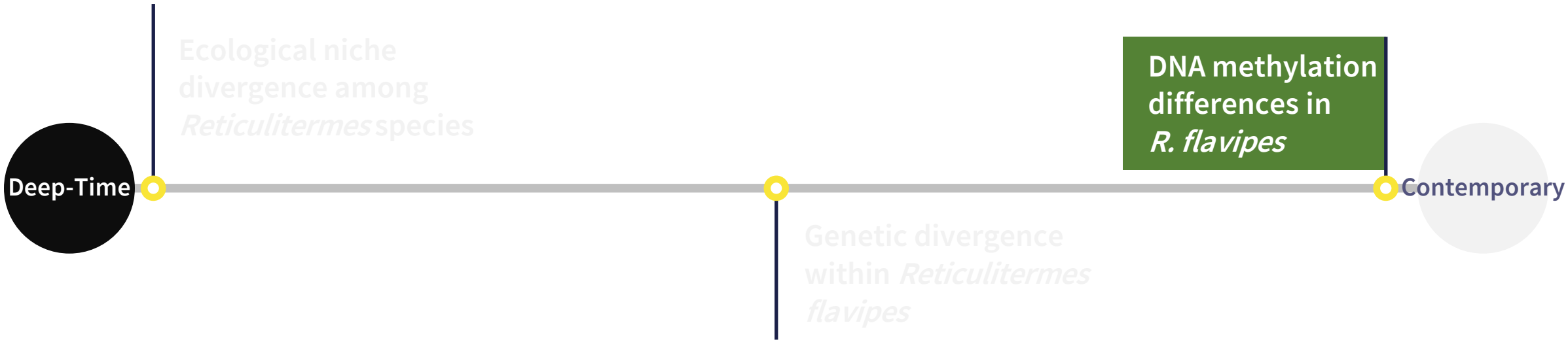


Distributional Shift: Central Expansion





Timescale of Evolution



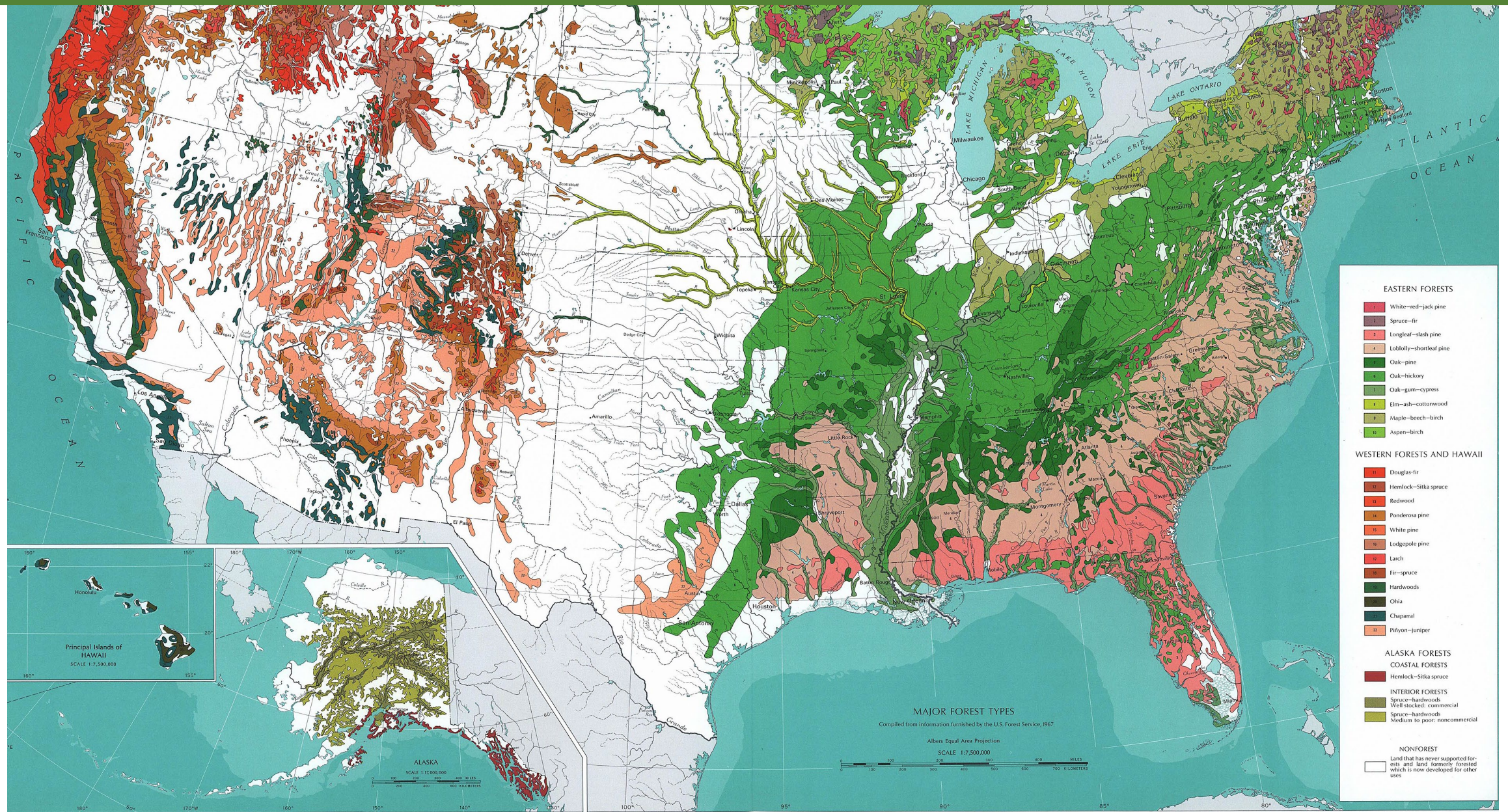
Questions

- **Disturbance of forest ecosystems**

1. Are there long-term methylation differences in *R. flavipes*?
2. Did ecological disturbance lead to (short-term) changes in DNA methylation?

Appalachian Forest Ecosystems

Pine? Oak?



Environmental Predictors:

DP = Dry-season Precipitation

ST = Summer Temperature

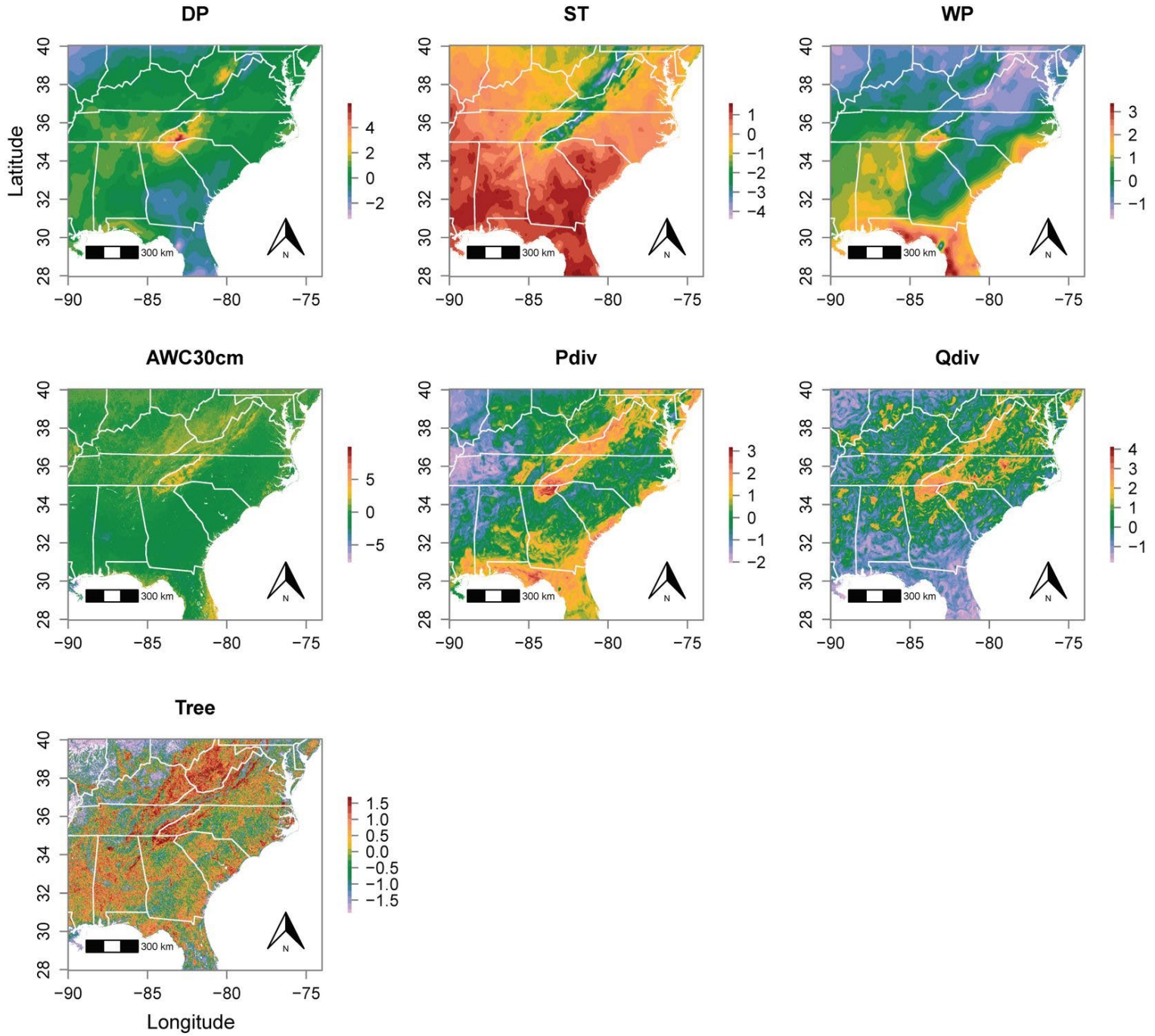
WP = Wet-season Precipitation

AWC30cm = Available Water Capacity at a Soil Depth of 30 cm

Pdiv = *Pinus* (pine) species richness

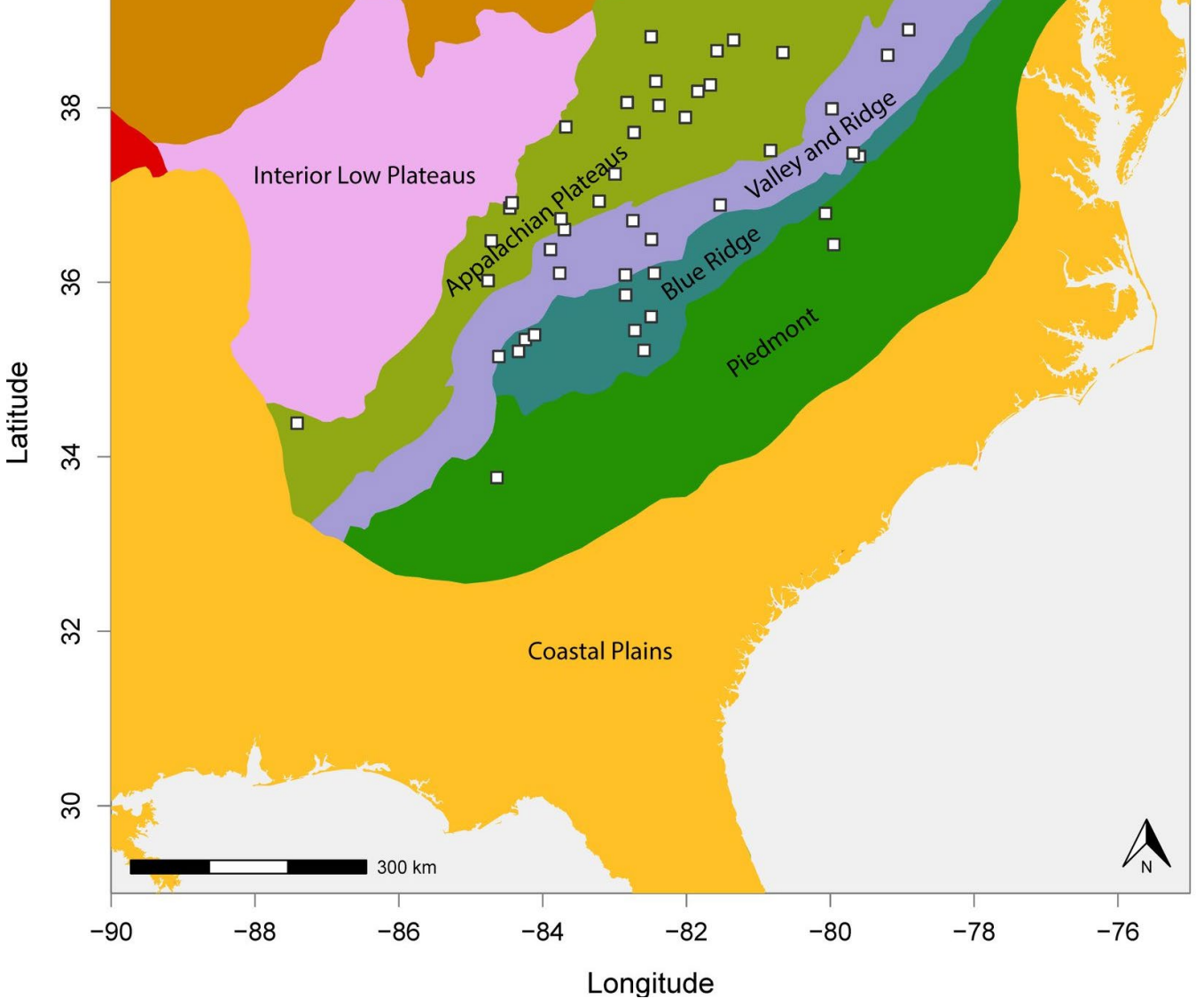
Qdiv = *Quercus* (oak) species richness

Tree = tree canopy cover



Epigenetic Data Collection:
Four Ecoregions: Appalachian Plateaus, Valley and Ridge, Blue Ridge, Piedmont

45 sites,
167 individuals
(workers AND soldiers)



DNA methylation assessed using

methylation-sensitive Amplified Fragment Length Polymorphism (msAFLP)

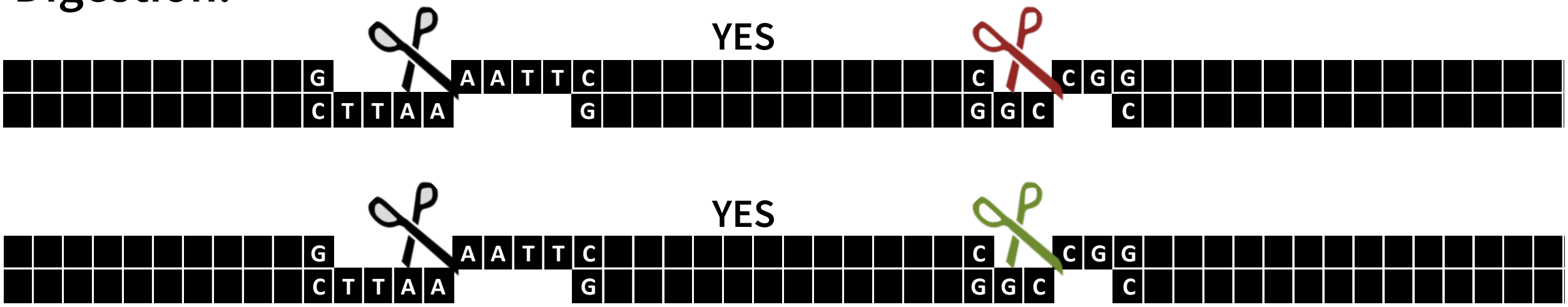
470 epigenetic loci

msAFLP

CCGG

Restriction Digestion:

Fragment Present:



Restriction Enzymes:

EcoRI

HpaII

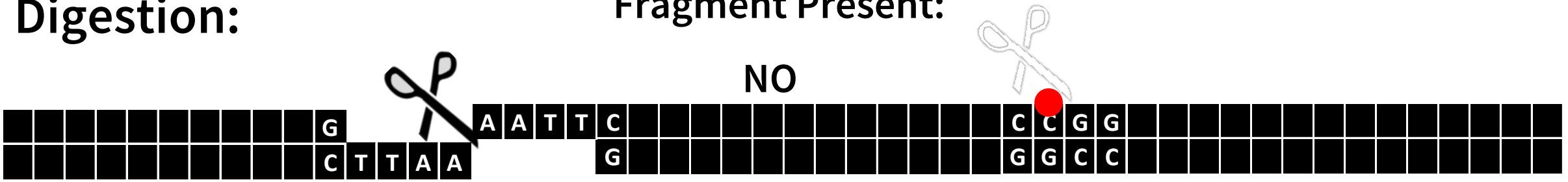
MspI

CmCGG

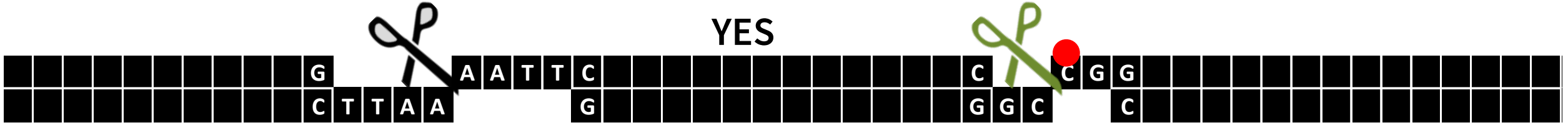
Restriction Digestion:

Fragment Present:

NO



YES



Restriction Enzymes:

EcoRI

HpaII

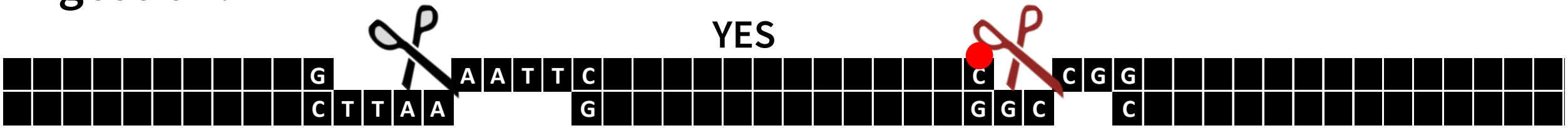
MspI

mCCGG

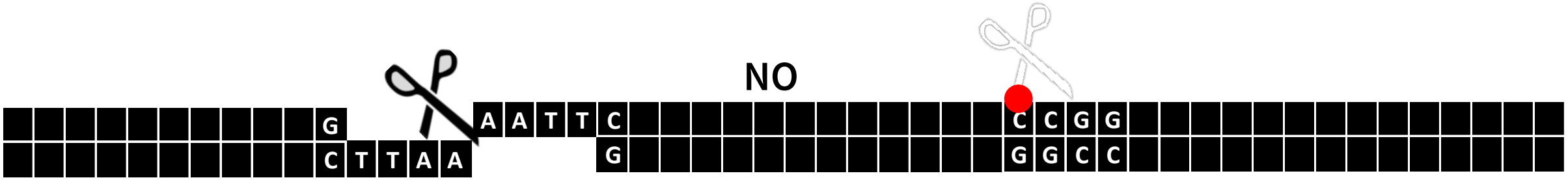
Restriction Digestion:

Fragment Present:

YES



NO



Restriction Enzymes:

EcoRI

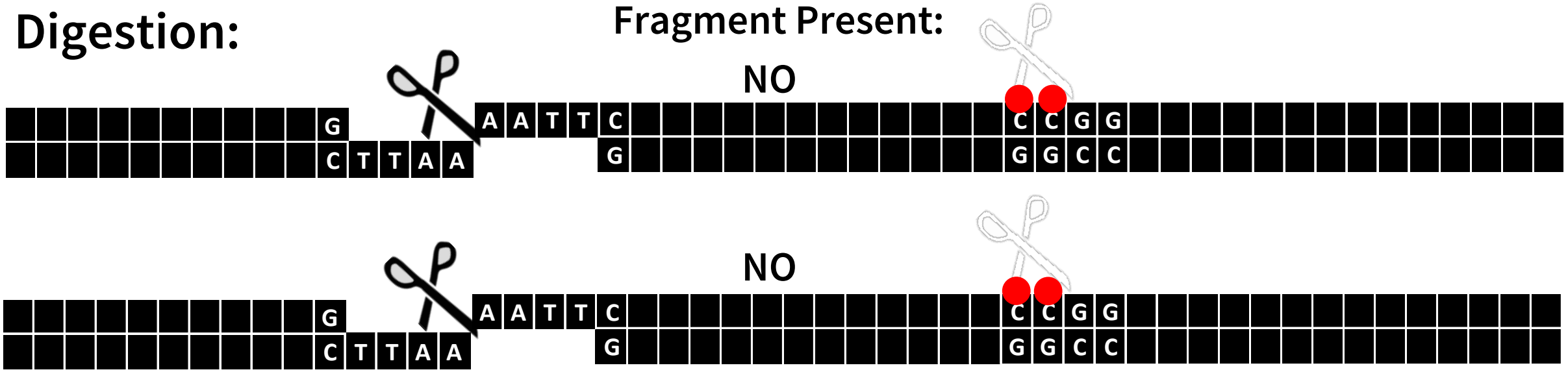
HpaII

MspI

Both cytosines methylated

Restriction Digestion:

Fragment Present:



Restriction Enzymes:

EcoRI

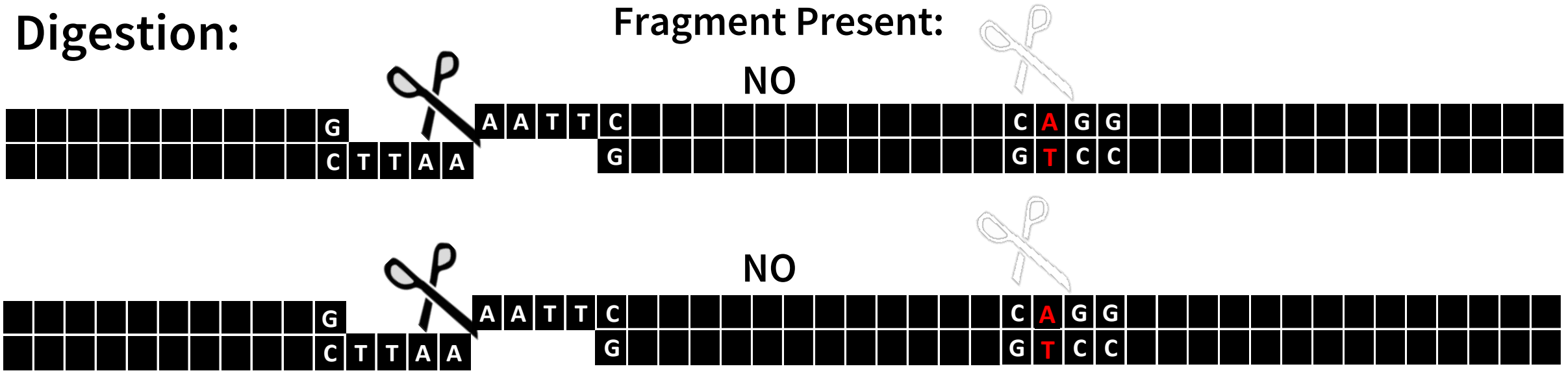
HpaII

MspI

Mutation at cut site

Restriction Digestion:

Fragment Present:



Restriction Enzymes:

EcoRI

HpaII

MspI

msAFLP data (each locus has 3 states)

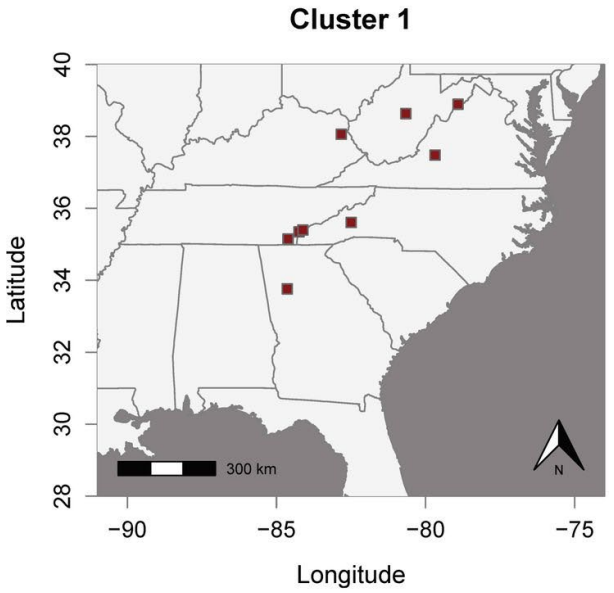
Ind.	Enz.	Loc.1
Ind.1	MspI	1
Ind.1	HpaII	1
Ind.2	MspI	1
Ind.2	HpaII	0
Ind.3	MspI	0
Ind.3	HpaII	1
Ind.4	MspI	0
Ind.4	HpaII	0

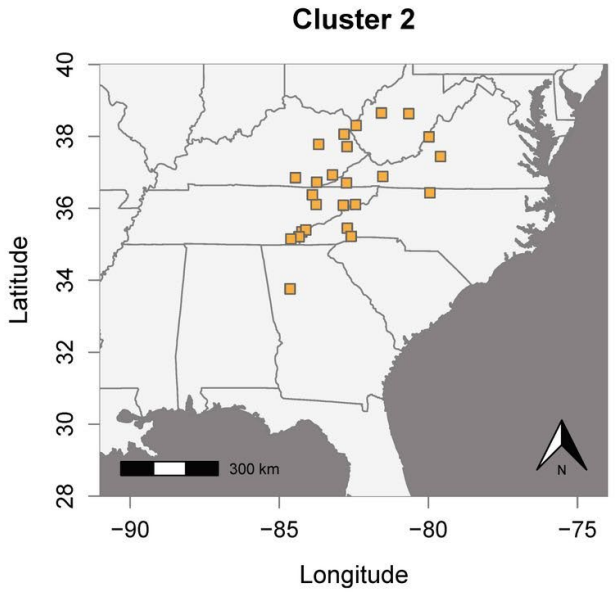
Ind.	Loc.1 (CCGG)	Loc.1 (CmCGG)	Loc.1 (mCCGG)
Ind.1	1	0	0
Ind.2	0	1	0
Ind.3	0	0	1
Ind.4	0	0	0

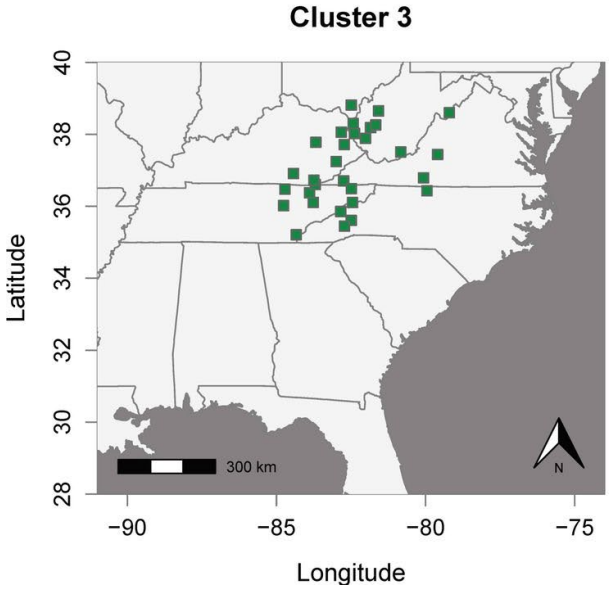
Epigenetic Clustering:

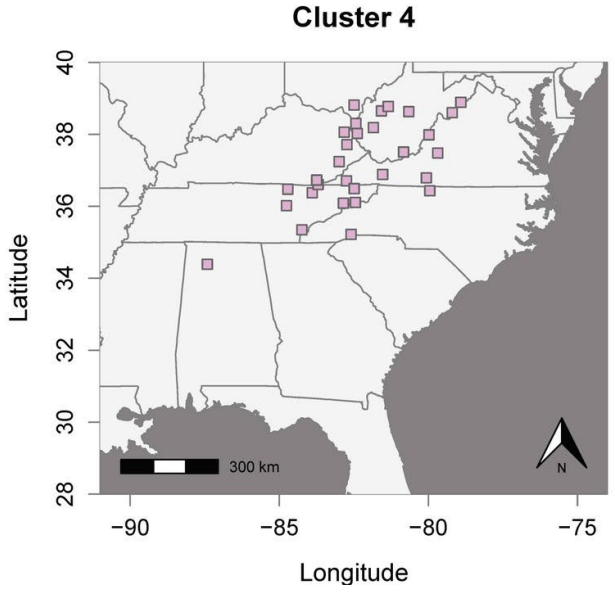
Genetic vs. Epigenetic?

Four Clusters

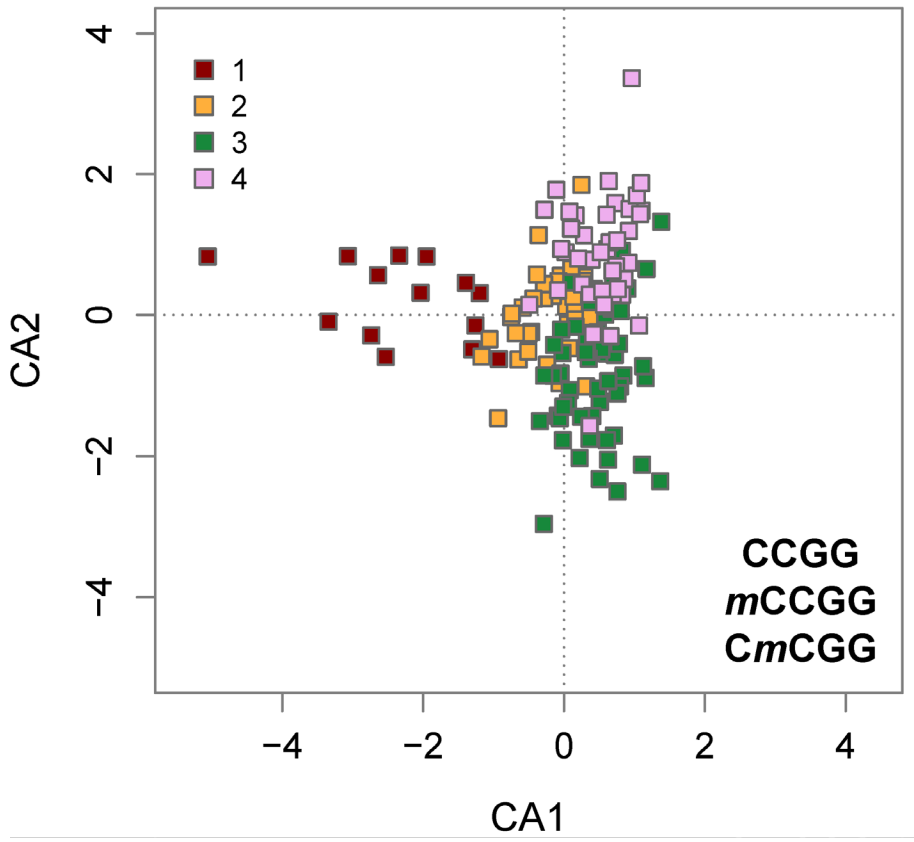






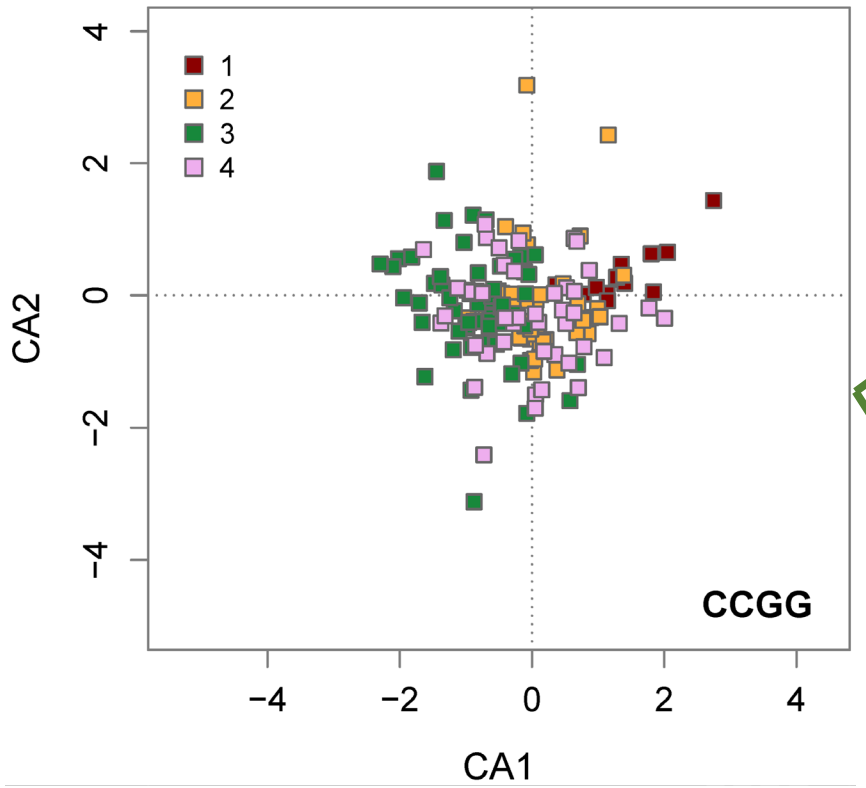


Four Clusters



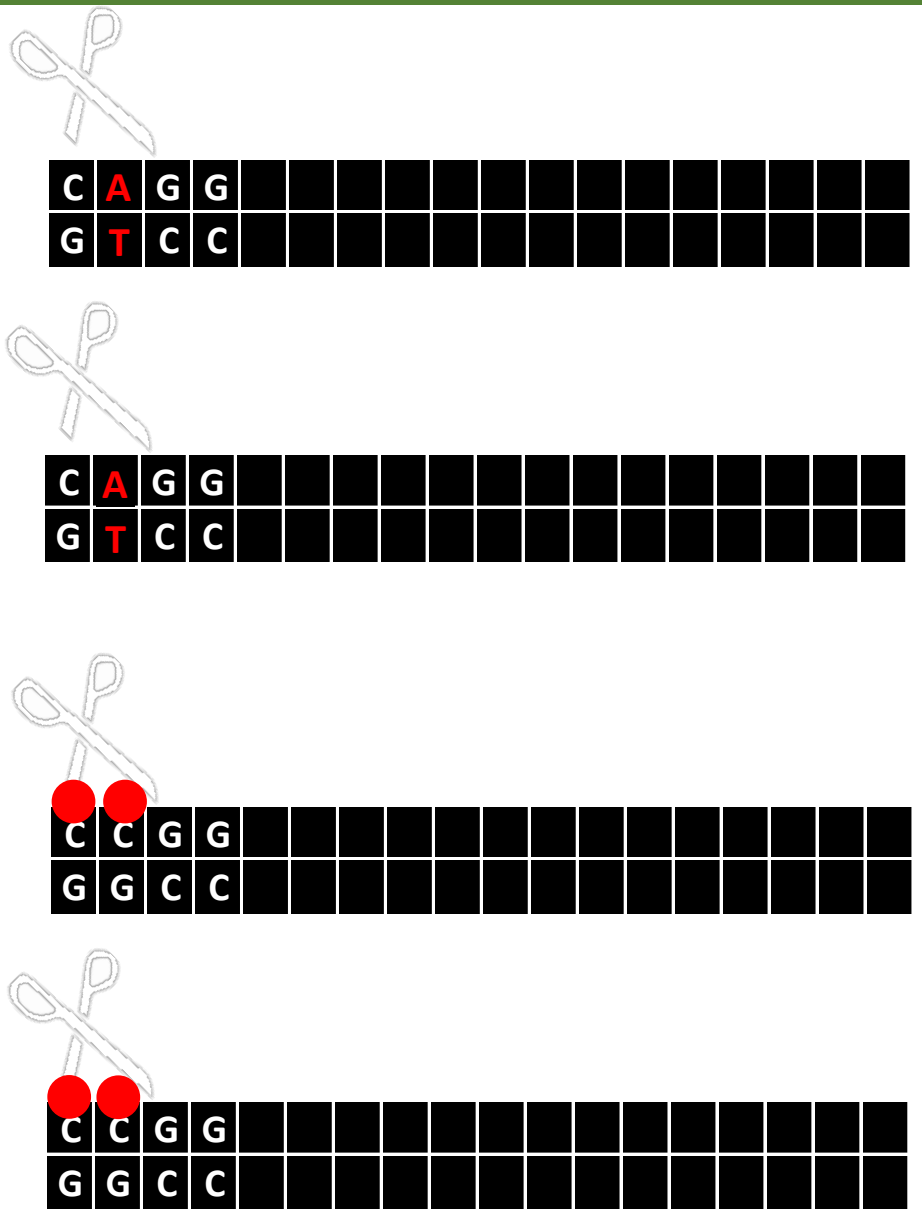
Genetic or Epigenetic Variation?

This genetic(?) variation
ASSUMES

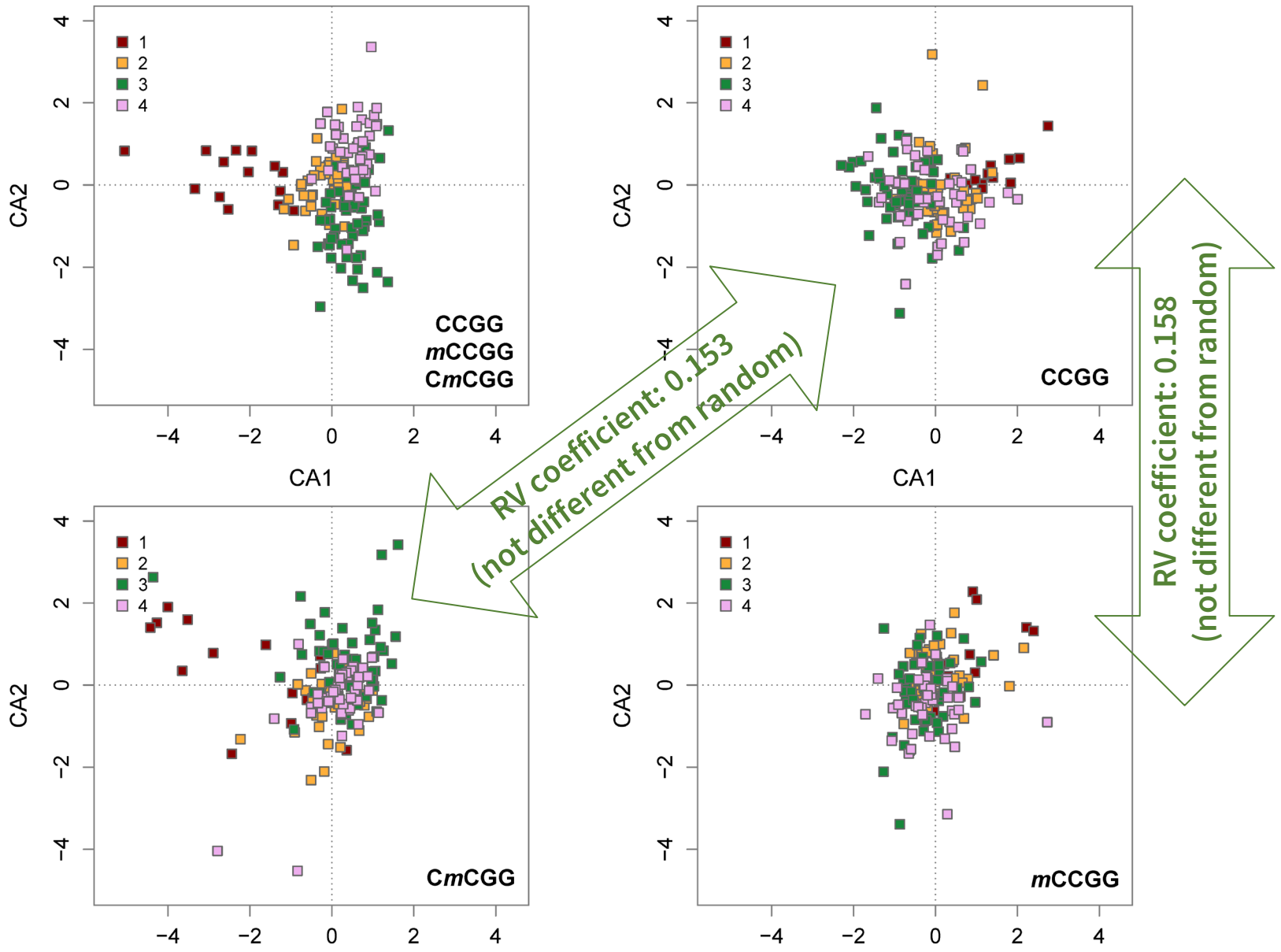


THIS!

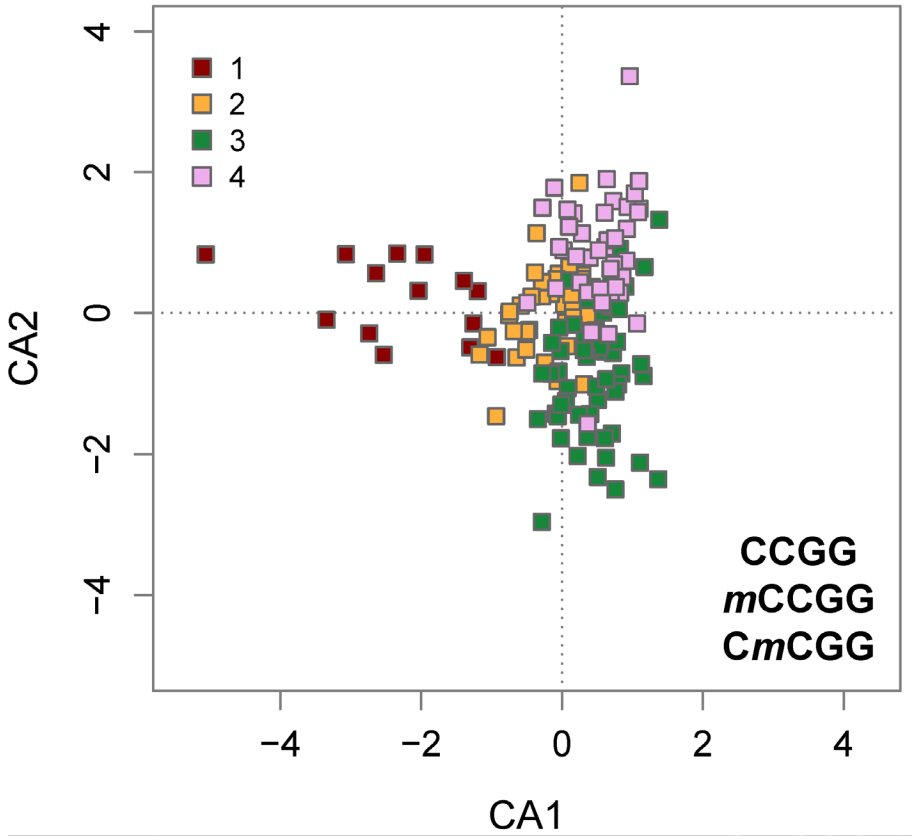
NOT THIS!



Genetic(?) vs. Epigenetic Variation



Genetically-dependent Epigenetic Variation

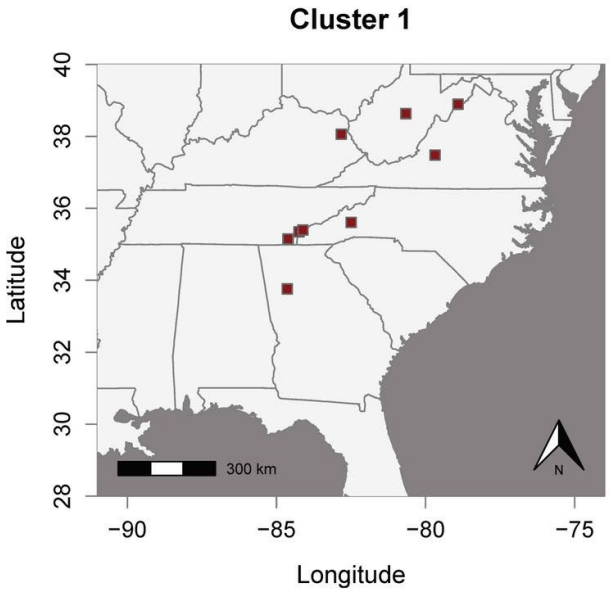


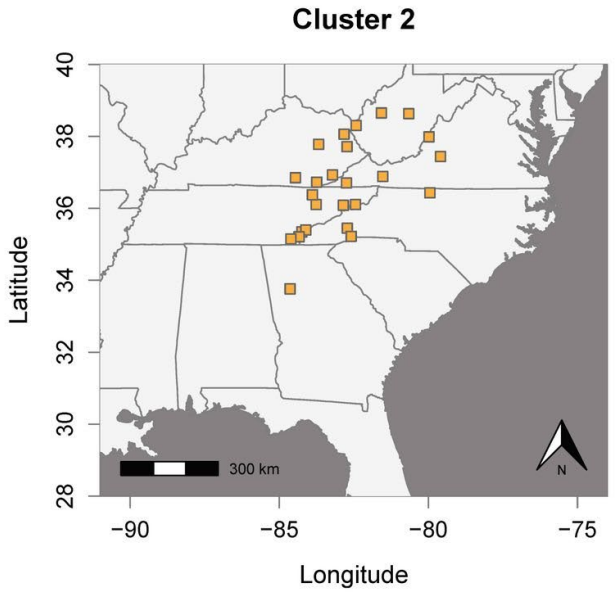
“Tricky” Terms:

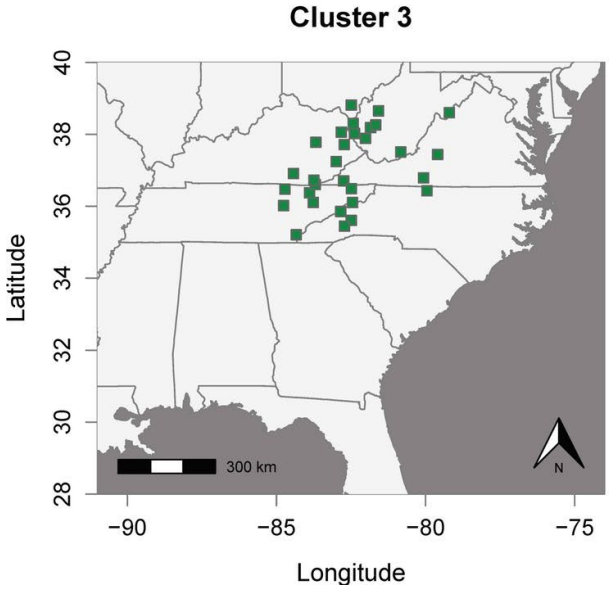
Long-term ≈ genetically-dependent ≈ ~~trans-generational(?)~~
Short-term ≈ environmentally-induced ≈ ~~within-generation(?)~~

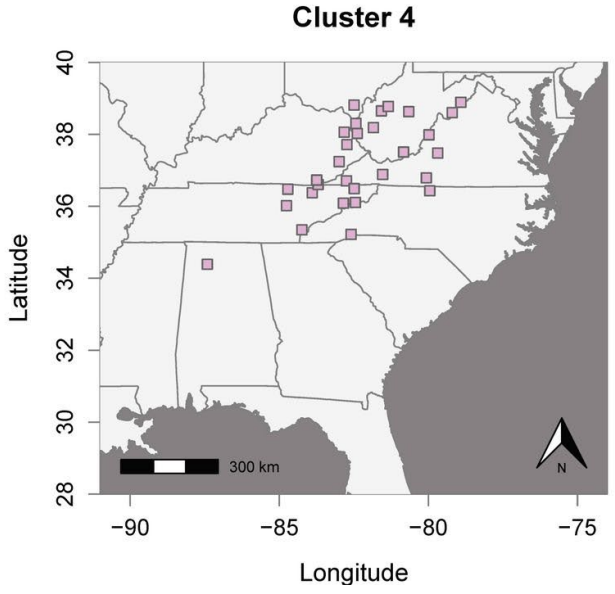
Long-term Influences on Epigenetic Variation

Epigenetic Clusters

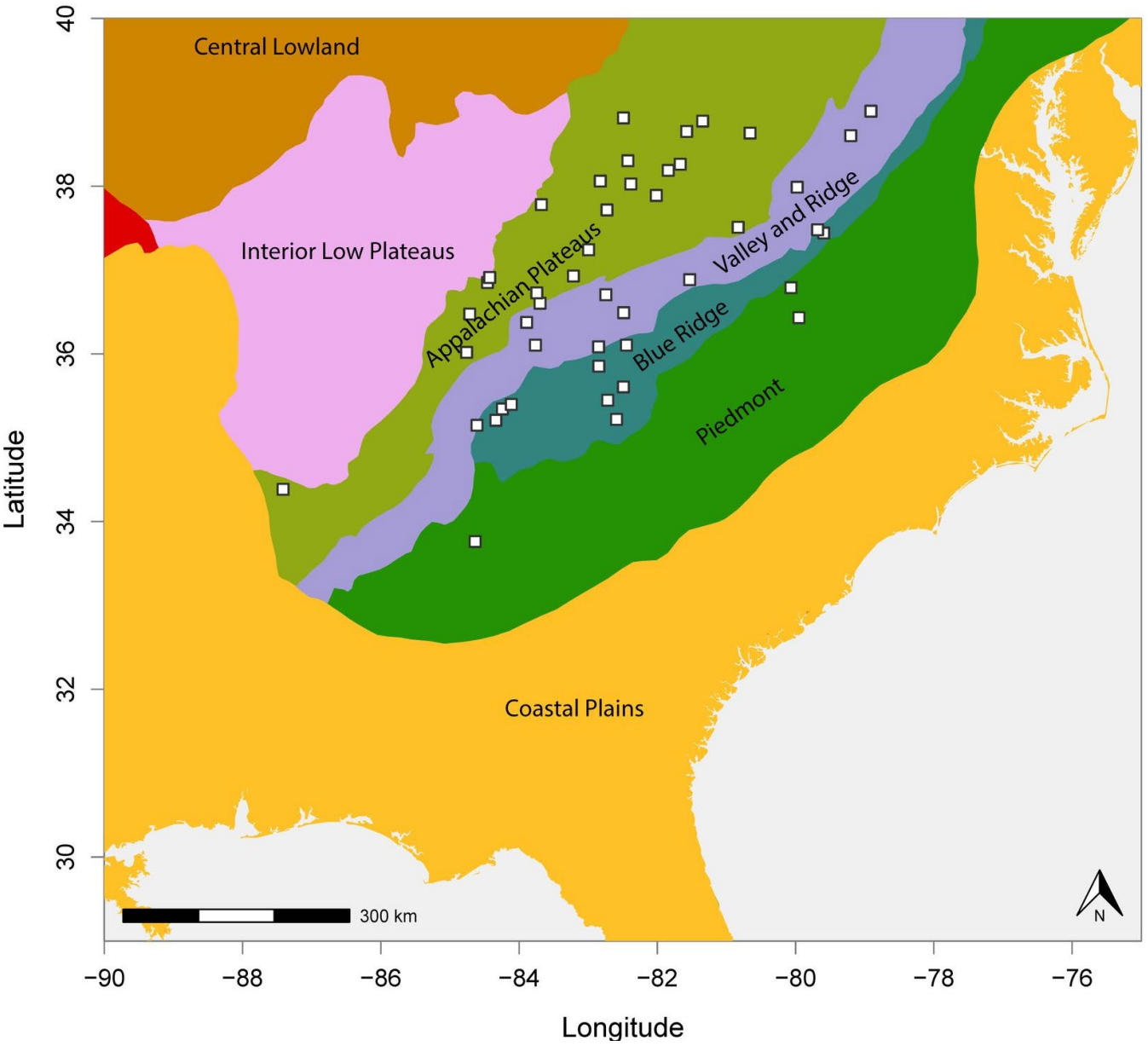








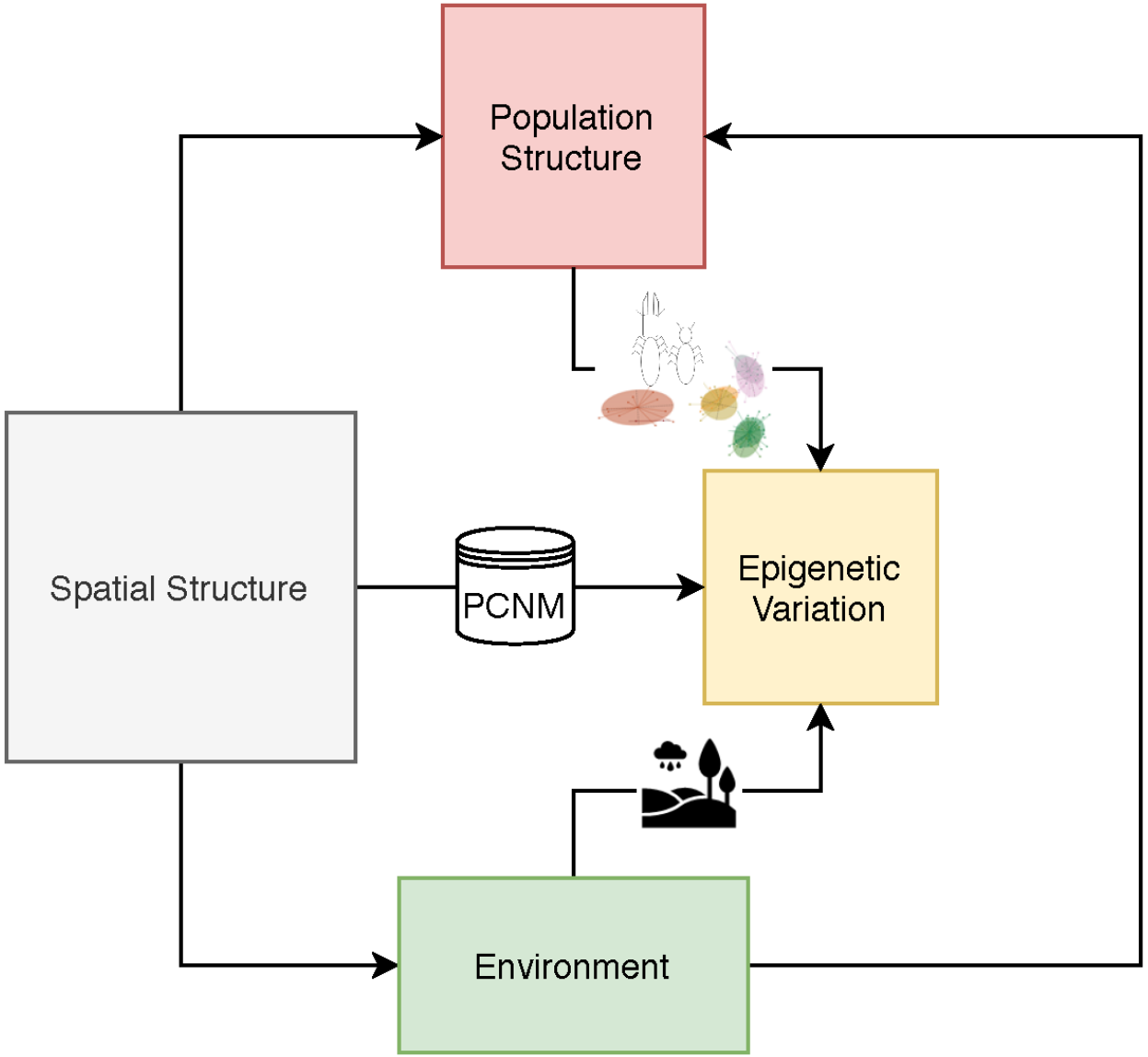
Epigenetic Clusters



	Proportion of Cluster in Ecoregion			
Appalachian Plateaus	0.21	0.37	0.57	0.48
Blue Ridge	0.64	0.30	0.19	0.14
Piedmont	0.07	0.09	0.07	0.07
Valley and Ridge	0.07	0.23	0.17	0.32

Cluster 1:	64% of cluster 1 is in Blue Ridge
Cluster 2:	37% in Plateaus, 30% in Blue Ridge
Cluster 3:	57% of cluster 3 is in Plateaus
Cluster 4:	48% of cluster 4 is in Plateaus

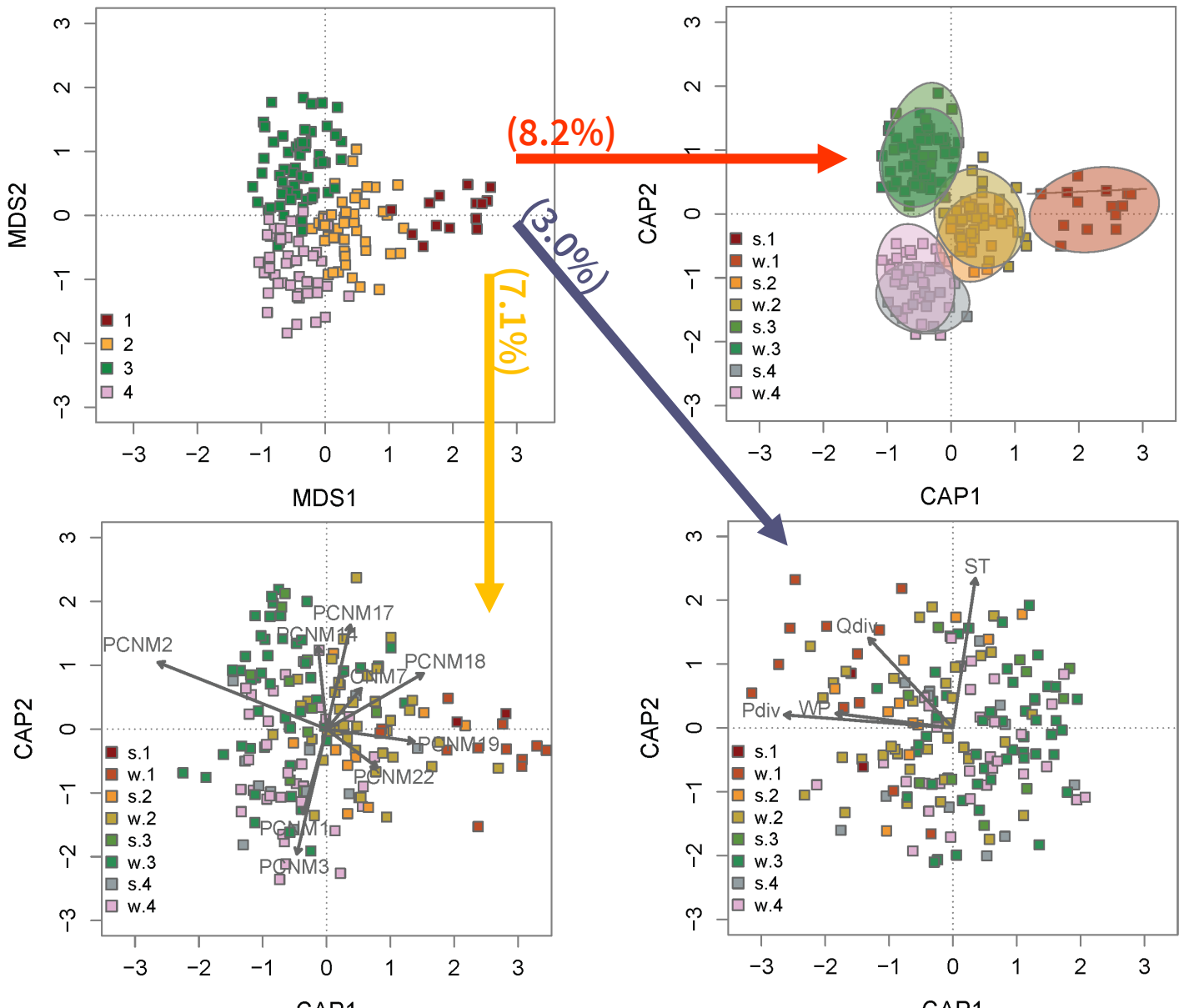
Partitioning Epigenetic Variation



Partitioning Epigenetic Variation

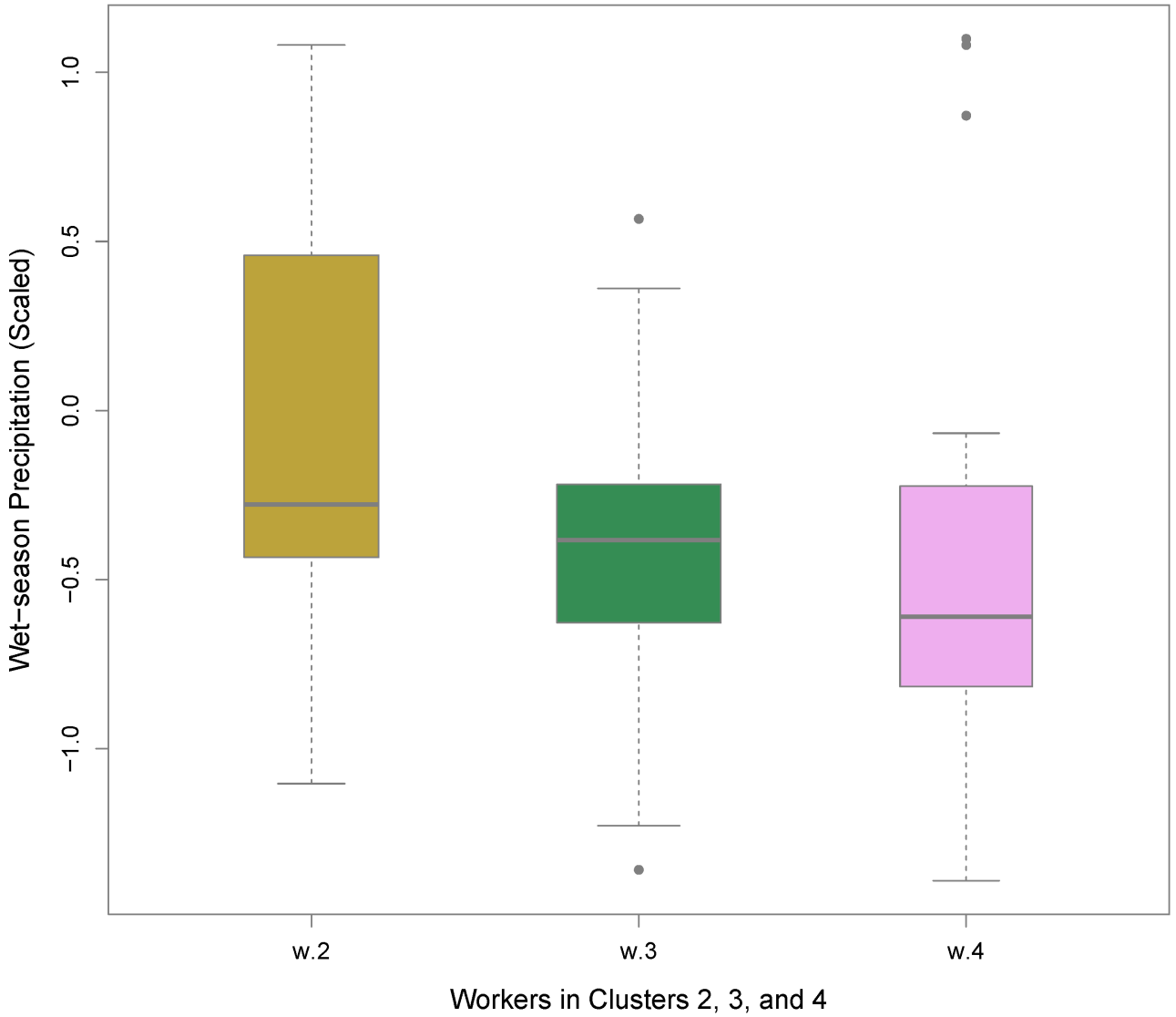
Constraints (Variation Explained):

- **Geography**
- **Environment**
- **Pop. Strata:**
  - Epigenetic Clusters
  - Caste Identity



Partitioning Epigenetic Variation: Environment

Univariate:

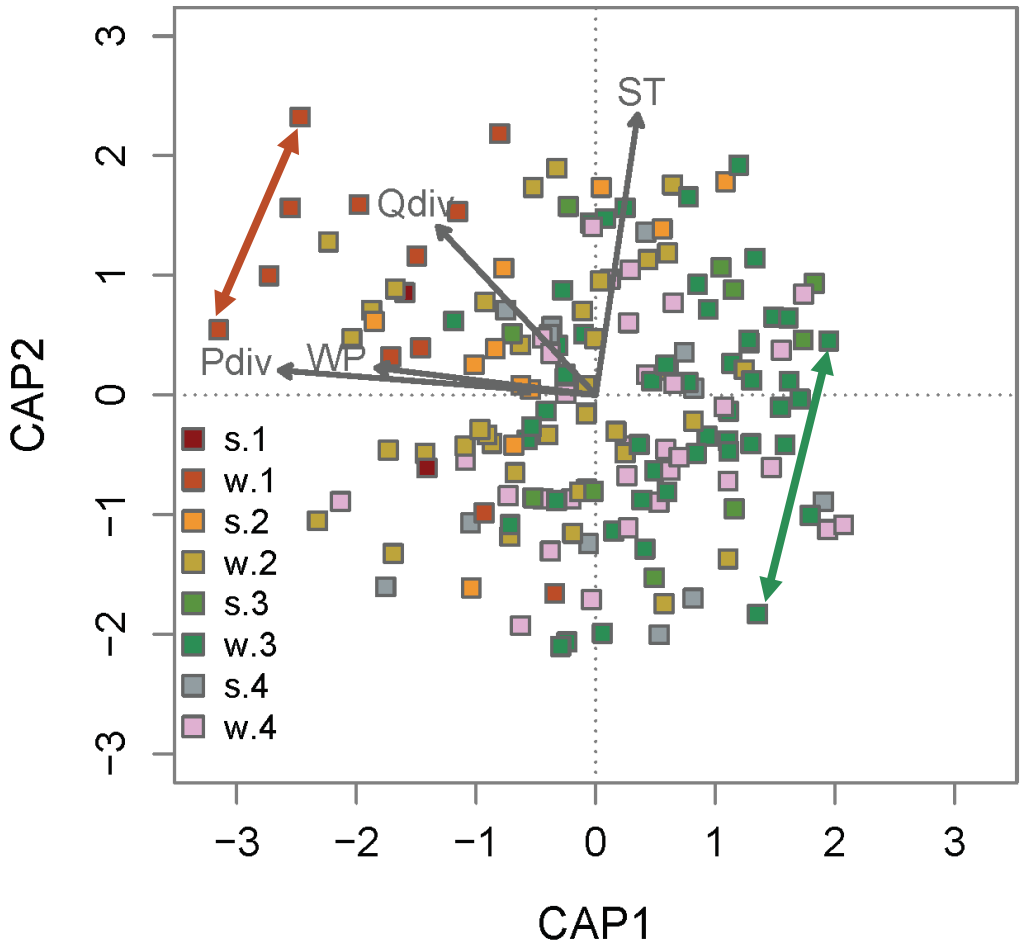


Multivariate:

Environment | Caste*Cluster + Geography				
	d.f.	Sum of Squares	F	p
DP	1	0.406	1.084	0.178
ST	1	0.449	1.199	0.025
WP	1	0.496	1.324	0.001
AWC30cm	1	0.447	1.191	0.031
Pdiv	1	0.491	1.309	0.003
Qdiv	1	0.429	1.143	0.060
Tree	1	0.408	1.088	0.170
Residual	143	53.599		

Short-term Influences on Epigenetic Variation

Short-term Influences: Differences at the Individual Level



Pine species richness: negative association (uniformly)

Fixed Effect	Random Effect	Locus	Methylation	Estimate	S.E.	z	p-value
Tree	Caste & Cluster	AG113	*CmCGG*	-0.642	0.333	-1.930	*0.054*
		AG113	*mCCGG*	0.904	0.423	2.136	0.033
		AG145	*CmCGG*	-0.902	0.405	-2.228	0.026
		AG174	*CmCGG*	2.465	1.114	2.214	0.027
		AG262	*CCGG*	-1.457	0.543	-2.684	0.007
		AT104	*CCGG*	0.765	0.307	2.495	0.013
		AT118	*CCGG*	-0.781	0.383	-2.038	0.042
		AT126	*mCCGG*	1.101	0.441	2.496	0.013
		AT206	*mCCGG*	-1.222	0.409	-2.987	0.003
		AT240	*CmCGG*	1.409	0.622	2.265	0.023
DP	Caste & Cluster	AT112	*CCGG*	0.506	0.212	2.389	0.017
		AG174	*CmCGG*	0.715	0.264	2.705	0.007
WP	Caste & Cluster	AT188	*CmCGG*	0.908	0.422	2.155	0.031
		AT206	*mCCGG*	-1.418	0.598	-2.372	0.018
ST	Caste & Cluster	AG134	*mCCGG*	0.758	0.346	2.194	0.028
		AG148	*CCGG*	0.790	0.308	2.563	0.010
		AG212	*mCCGG*	1.095	0.564	1.942	*0.052*
Pine	Caste & Cluster	AG113	*CCGG*	-0.797	0.329	-2.421	0.015
		AG134	*mCCGG*	-0.473	0.245	-1.928	*0.054*
		AT104	*CCGG*	-0.619	0.255	-2.425	0.015
		AT166	*CmCGG*	-0.853	0.341	-2.503	0.012
		AT216	*CCGG*	-1.301	0.562	-2.316	0.021
Oak	Caste & Cluster	AG113	*mCCGG*	0.530	0.246	2.154	0.031
		AG174	*CmCGG*	0.755	0.388	1.947	*0.052*
		AT166	*CmCGG*	-0.775	0.311	-2.491	0.013
		AT188	*CmCGG*	0.753	0.348	2.163	0.031
		AT240	*CmCGG*	0.672	0.307	2.192	0.028

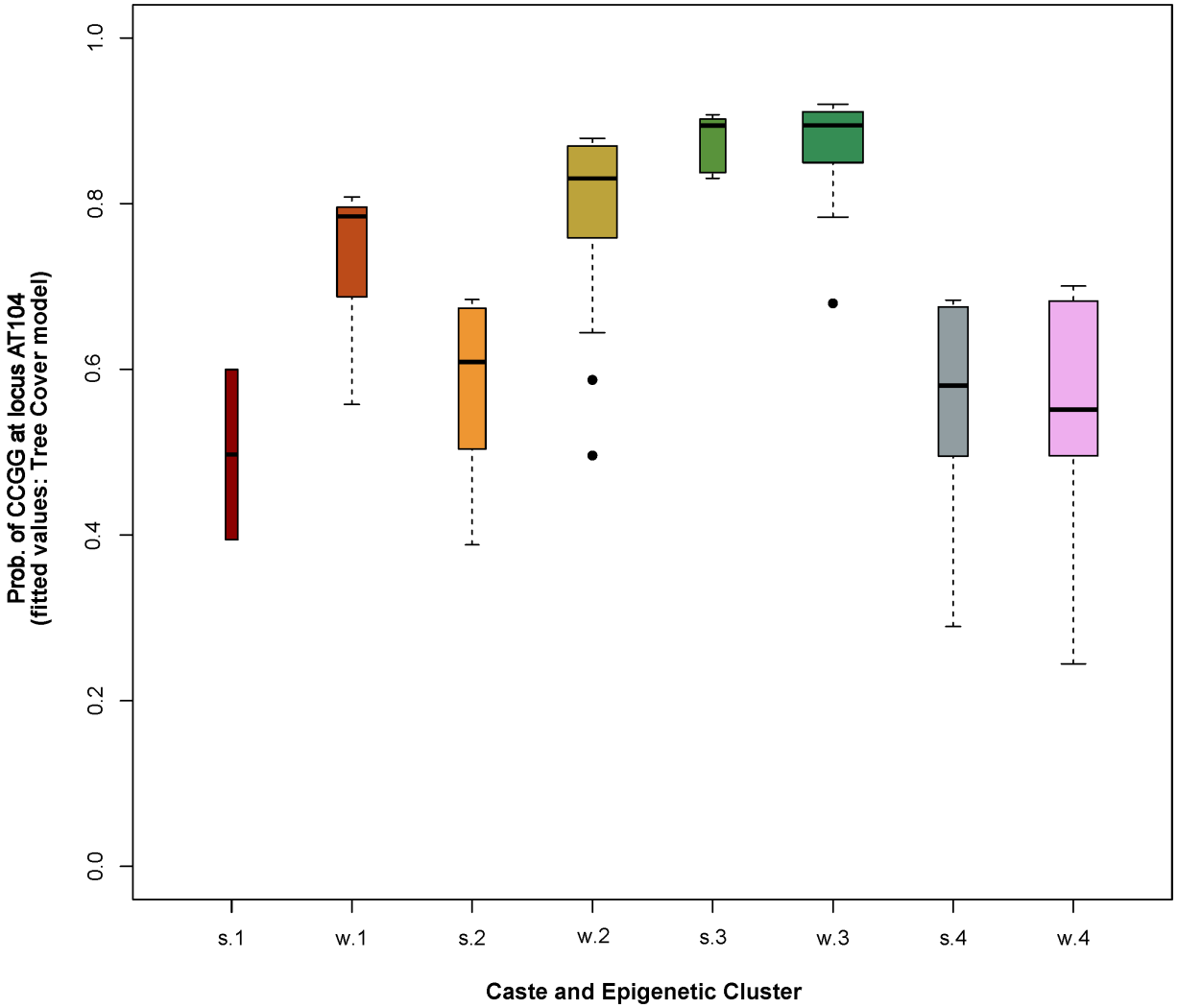
Oak species richness: positive association (mostly)

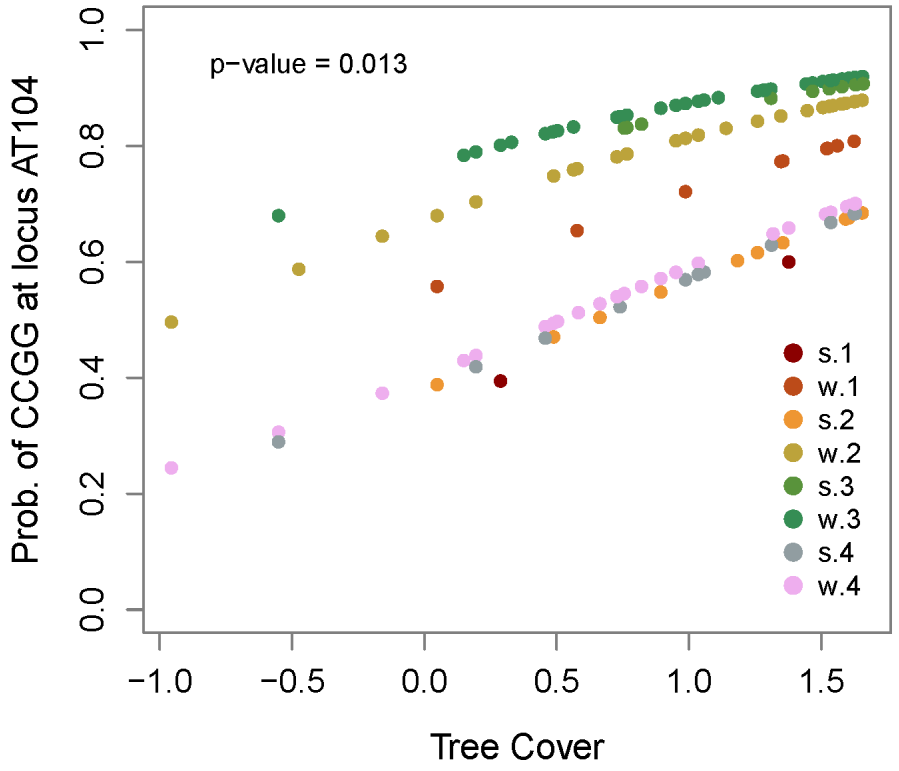
Fixed Effect	Random Effect	Locus	Methylation	Estimate	S.E.	z	p-value
Tree	Caste & Cluster	AG113	*CmCGG*	-0.642	0.333	-1.930	*0.054*
		AG113	*mCCGG*	0.904	0.423	2.136	0.033
		AG145	*CmCGG*	-0.902	0.405	-2.228	0.026
		AG174	*CmCGG*	2.465	1.114	2.214	0.027
		AG262	*CCGG*	-1.457	0.543	-2.684	0.007
		AT104	*CCGG*	0.765	0.307	2.495	0.013
		AT118	*CCGG*	-0.781	0.383	-2.038	0.042
		AT126	*mCCGG*	1.101	0.441	2.496	0.013
		AT206	*mCCGG*	-1.222	0.409	-2.987	0.003
		AT240	*CmCGG*	1.409	0.622	2.265	0.023
DP	Caste & Cluster	AT112	*CCGG*	0.506	0.212	2.389	0.017
		AG174	*CmCGG*	0.715	0.264	2.705	0.007
WP	Caste & Cluster	AT188	*CmCGG*	0.908	0.422	2.155	0.031
		AT206	*mCCGG*	-1.418	0.598	-2.372	0.018
ST	Caste & Cluster	AG134	*mCCGG*	0.758	0.346	2.194	0.028
		AG148	*CCGG*	0.790	0.308	2.563	0.010
		AG212	*mCCGG*	1.095	0.564	1.942	*0.052*
Pine	Caste & Cluster	AG113	*CCGG*	-0.797	0.329	-2.421	0.015
		AG134	*mCCGG*	-0.473	0.245	-1.928	*0.054*
		AT104	*CCGG*	-0.619	0.255	-2.425	0.015
		AT166	*CmCGG*	-0.853	0.341	-2.503	0.012
		AT216	*CCGG*	-1.301	0.562	-2.316	0.021
Oak	Caste & Cluster	AG113	*mCCGG*	0.530	0.246	2.154	0.031
		AG174	*CmCGG*	0.755	0.388	1.947	*0.052*
		AT166	*CmCGG*	-0.775	0.311	-2.491	0.013
		AT188	*CmCGG*	0.753	0.348	2.163	0.031
		AT240	*CmCGG*	0.672	0.307	2.192	0.028

Tree canopy cover

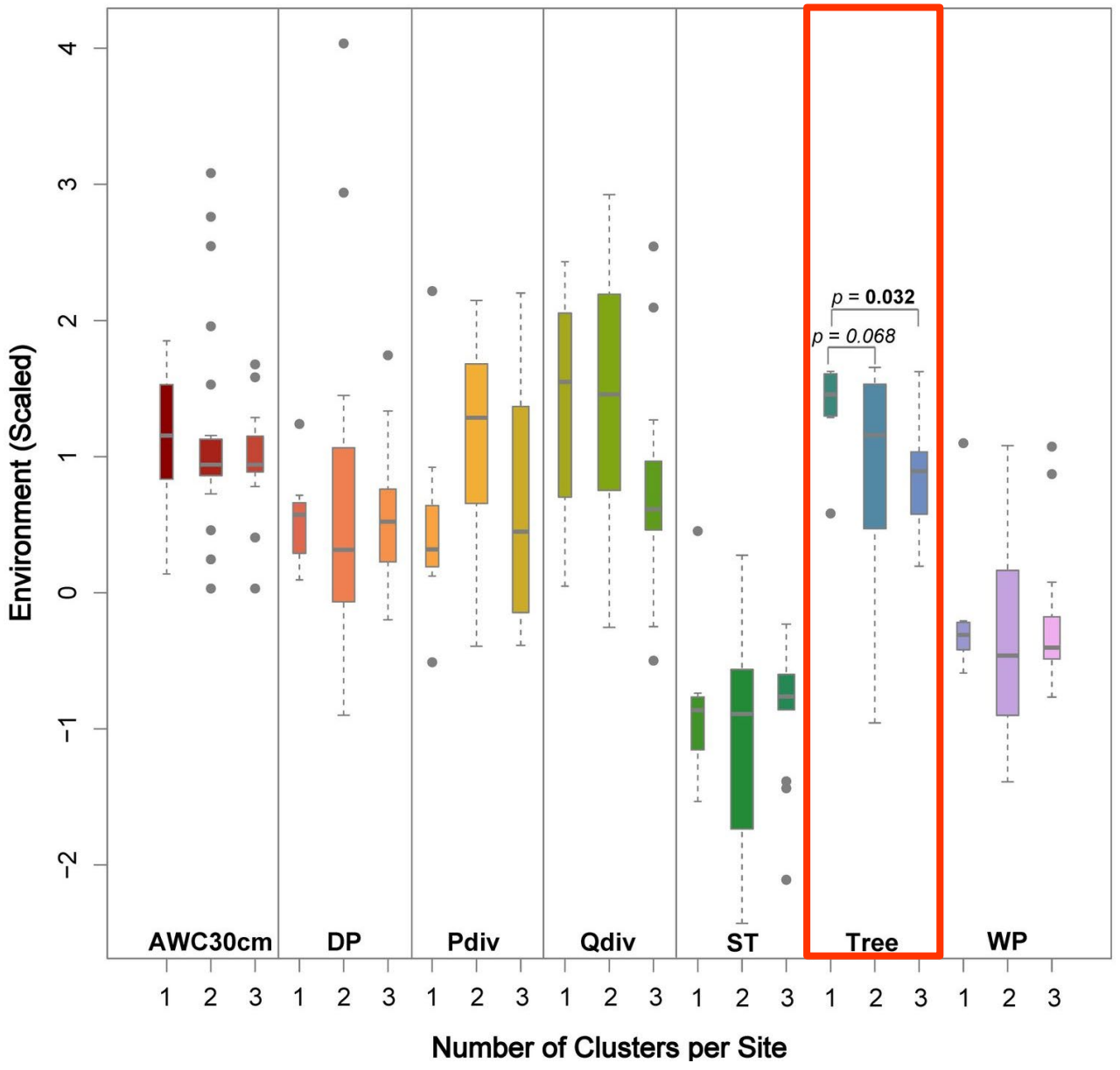
Fixed Effect	Random Effect	Locus	Methylation	Estimate	S.E.	z	p-value
Tree	Caste & Cluster	AG113	*CmCGG*	-0.642	0.333	-1.930	*0.054*
		AG113	*mCCGG*	0.904	0.423	2.136	0.033
		AG145	*CmCGG*	-0.902	0.405	-2.228	0.026
		AG174	*CmCGG*	2.465	1.114	2.214	0.027
		AG262	*CCGG*	-1.457	0.543	-2.684	0.007
		AT104	*CCGG*	0.765	0.307	2.495	0.013
		AT118	*CCGG*	-0.781	0.383	-2.038	0.042
		AT126	*mCCGG*	1.101	0.441	2.496	0.013
		AT206	*mCCGG*	-1.222	0.409	-2.987	0.003
		AT240	*CmCGG*	1.409	0.622	2.265	0.023
DP	Caste & Cluster	AT112	*CCGG*	0.506	0.212	2.389	0.017
		AG174	*CmCGG*	0.715	0.264	2.705	0.007
WP	Caste & Cluster	AT188	*CmCGG*	0.908	0.422	2.155	0.031
		AT206	*mCCGG*	-1.418	0.598	-2.372	0.018
ST	Caste & Cluster	AG134	*mCCGG*	0.758	0.346	2.194	0.028
		AG148	*CCGG*	0.790	0.308	2.563	0.010
		AG212	*mCCGG*	1.095	0.564	1.942	*0.052*
Pine	Caste & Cluster	AG113	*CCGG*	-0.797	0.329	-2.421	0.015
		AG134	*mCCGG*	-0.473	0.245	-1.928	*0.054*
		AT104	*CCGG*	-0.619	0.255	-2.425	0.015
		AT166	*CmCGG*	-0.853	0.341	-2.503	0.012
		AT216	*CCGG*	-1.301	0.562	-2.316	0.021
Oak	Caste & Cluster	AG113	*mCCGG*	0.530	0.246	2.154	0.031
		AG174	*CmCGG*	0.755	0.388	1.947	*0.052*
		AT166	*CmCGG*	-0.775	0.311	-2.491	0.013
		AT188	*CmCGG*	0.753	0.348	2.163	0.031
		AT240	*CmCGG*	0.672	0.307	2.192	0.028

Tree Canopy Cover:
Higher canopy cover = less methylation at locus AT104

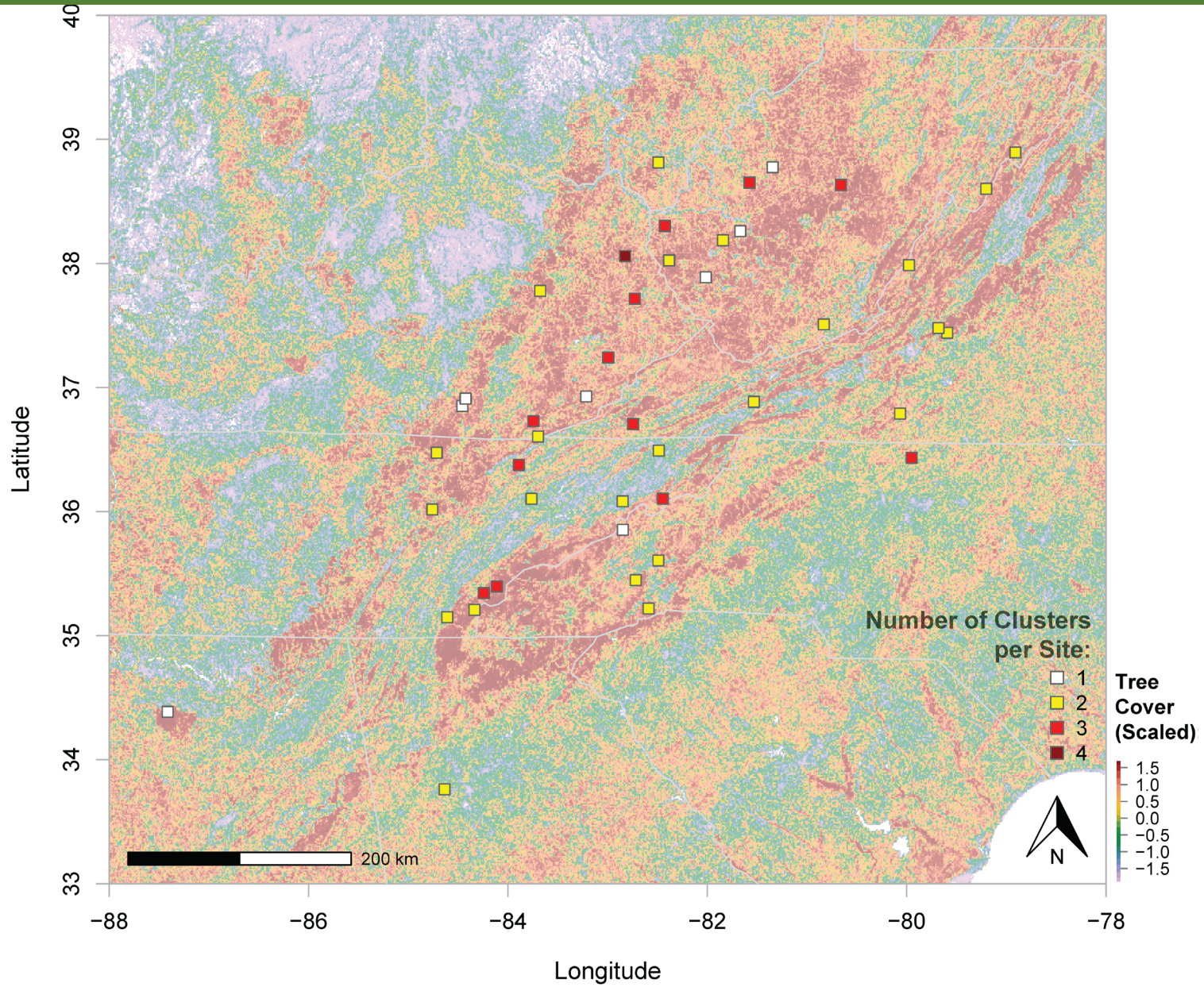




Tree Canopy Cover:
Lower canopy cover = more “mixed” colonies



Tree Canopy Cover:
"Mixed" colonies



Conclusions

1. Long-term methylation differences were detected
2. Canopy cover affects short-term methylation differences
3. Canopy cover affects epigenetic colony mixing

Details (analysis scripts and results) available at:

http://github.com/chazhyseni/msaflp

Questions?

SpECEvGen http://specevgen.net