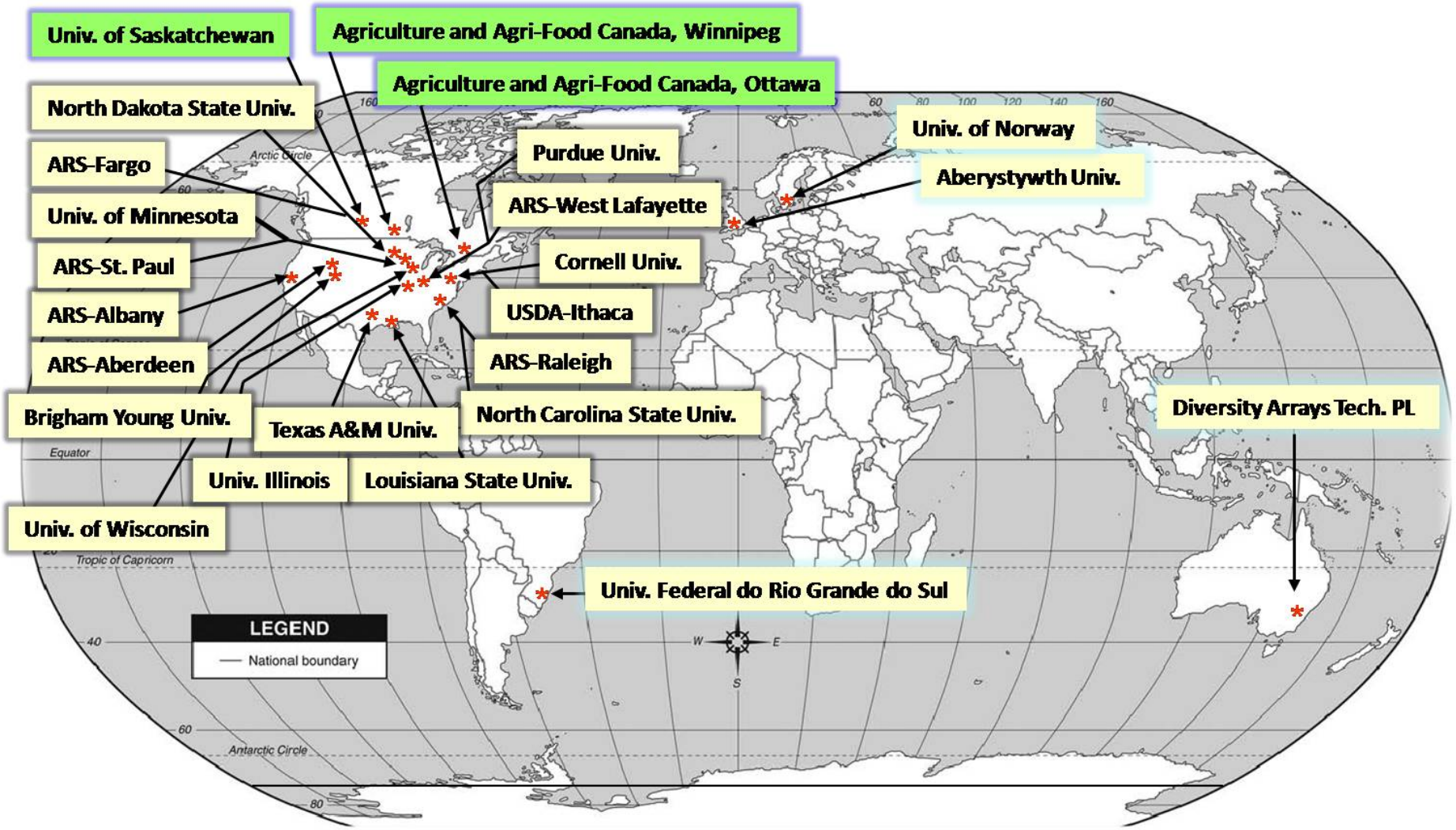
GENETIC DIVERSITY IN THE NORTH AMERICAN COLLABORATIVE OAT RESEARCH ENTERPRISE (CORE)

Kathy Esvelt Klos
Yung-Fen Huang
Nicholas Tinker
CORE Collaborators

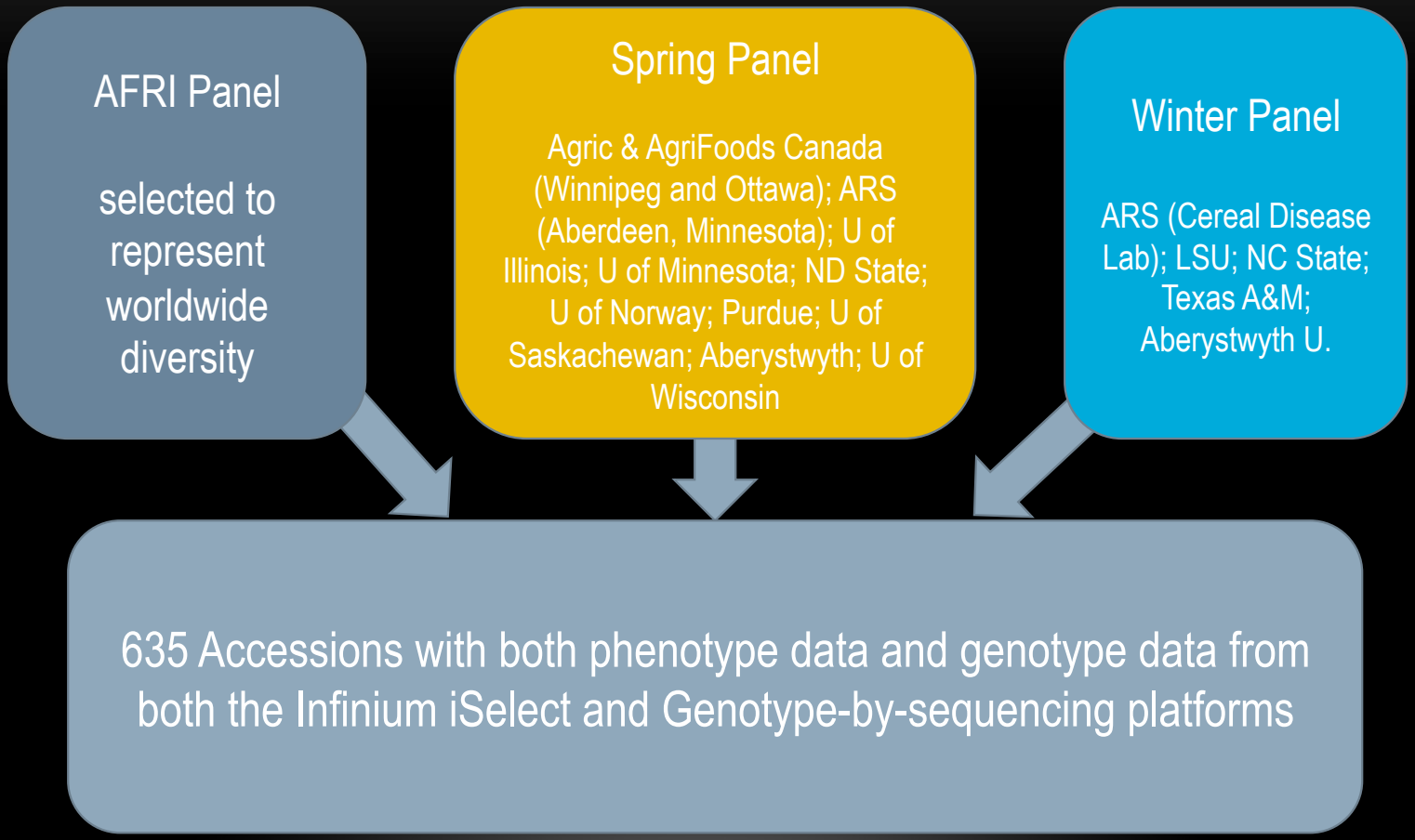

Univ. of Saskatchewan
Agriculture and Agri-Food Canada, Winnipeg
Agriculture and Agri-Food Canada, Ottawa
North Dakota State Univ.
ARS-Fargo
Univ. of Minnesota
ARS-St. Paul
ARS-Albany
ARS-Aberdeen
Brigham Young Univ.
Univ. of Wisconsin
Purdue Univ.
ARS-West Lafayette
Cornell Univ.
USDA-Ithaca
ARS-Raleigh
North Carolina State Univ.
Texas A&M Univ.
Univ. Illinois
Louisiana State Univ.
Univ. of Norway
Aberystywth Univ.
Diversity Arrays Tech. PL
Univ. Federal do Rio Grande do Sul
Arctic Circle
Equator
Tropic of Capricorn
Antarctic Circle
LEGEND
National boundary
W
E
S

CORE Team

SPOILERS

- In mixed linear model analyses of genome-wide association in the CORE sample, confounding between large-scale population structure and cryptic relatedness will tend to inflate the false negative rate.
- Empirical analyses of GWAS indicate high concordance between expected and observed location of loci.

DATA: ACCESSIONS



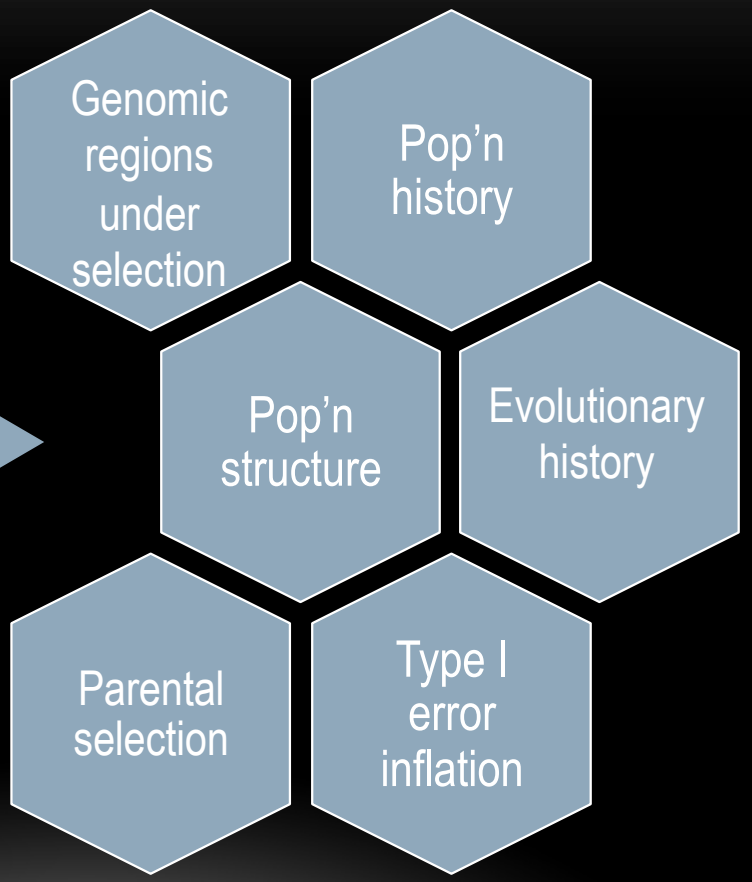
DATA: GENOTYPES

After filtering by ≤95% missing data, MAF ≥0.01 and heterozygocity ≤0.05

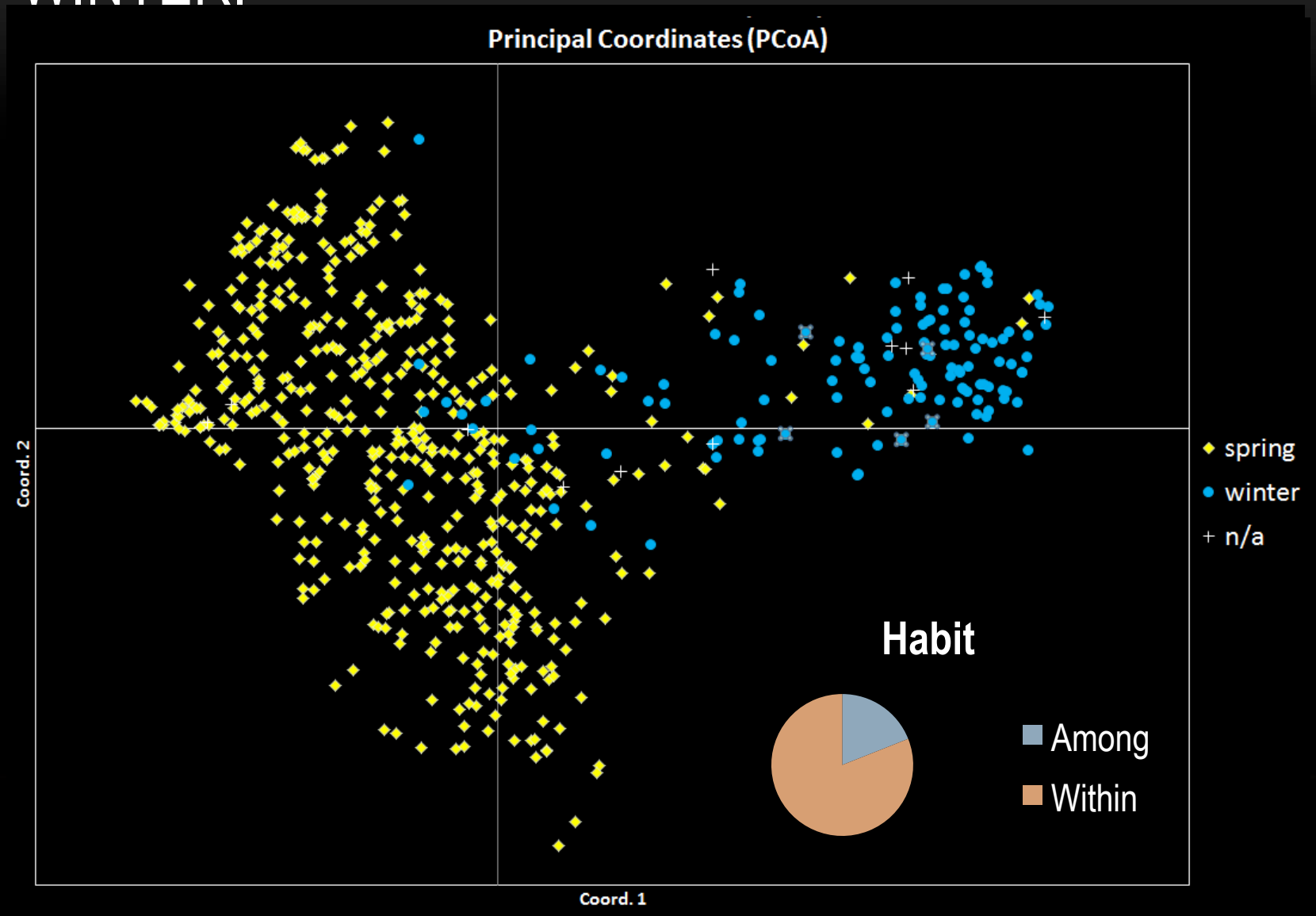
	Polymorphic Markers						
	All	**AFRI**	**Spring**	**Winter**	**Rare**	**MAF**	**Call Rate**
iSelect	1,926	1,925	1,926	1,876	116	0.25	0.999
GBS	9,995	9,411	9,986	9,765	8,160	0.05	0.971
	N=635	N=102	N=409	N=121			

22 SNPs were observed only in the Spring panel,
8 SNPs were observed only in the Winter Pane.

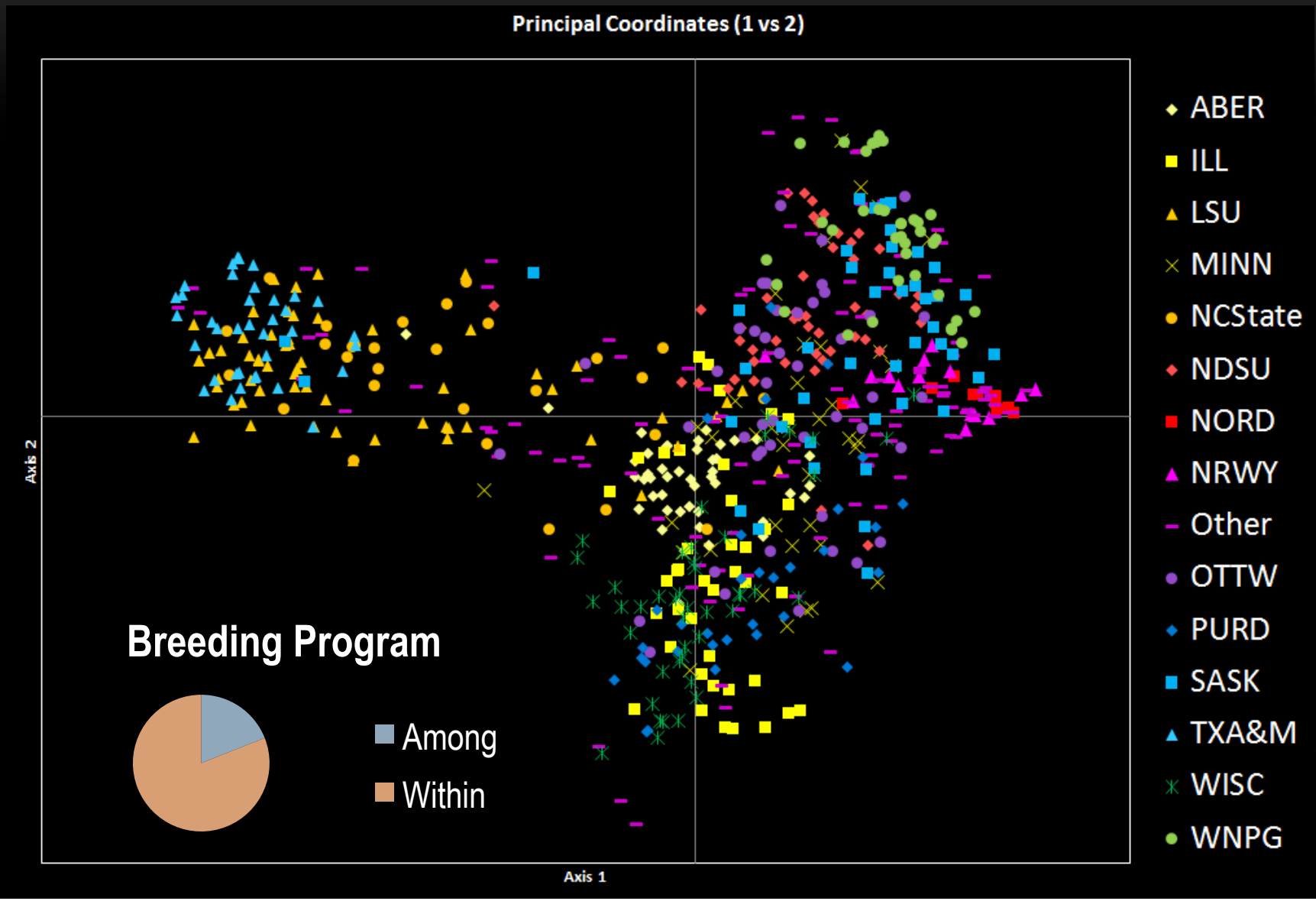
Genetic Diversity



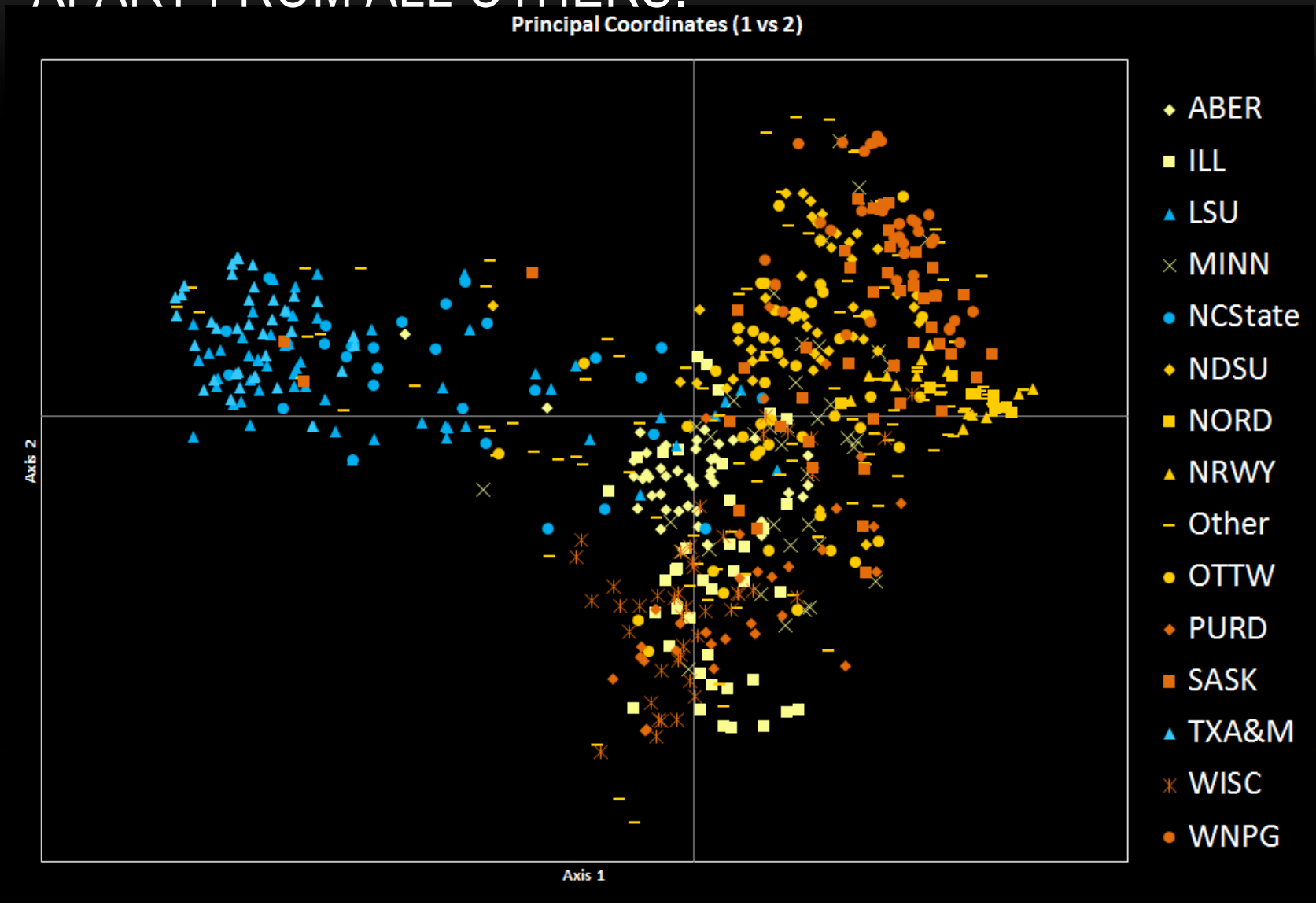
PCA SEPARATES SPRING ACCESSIONS FROM WINTER.



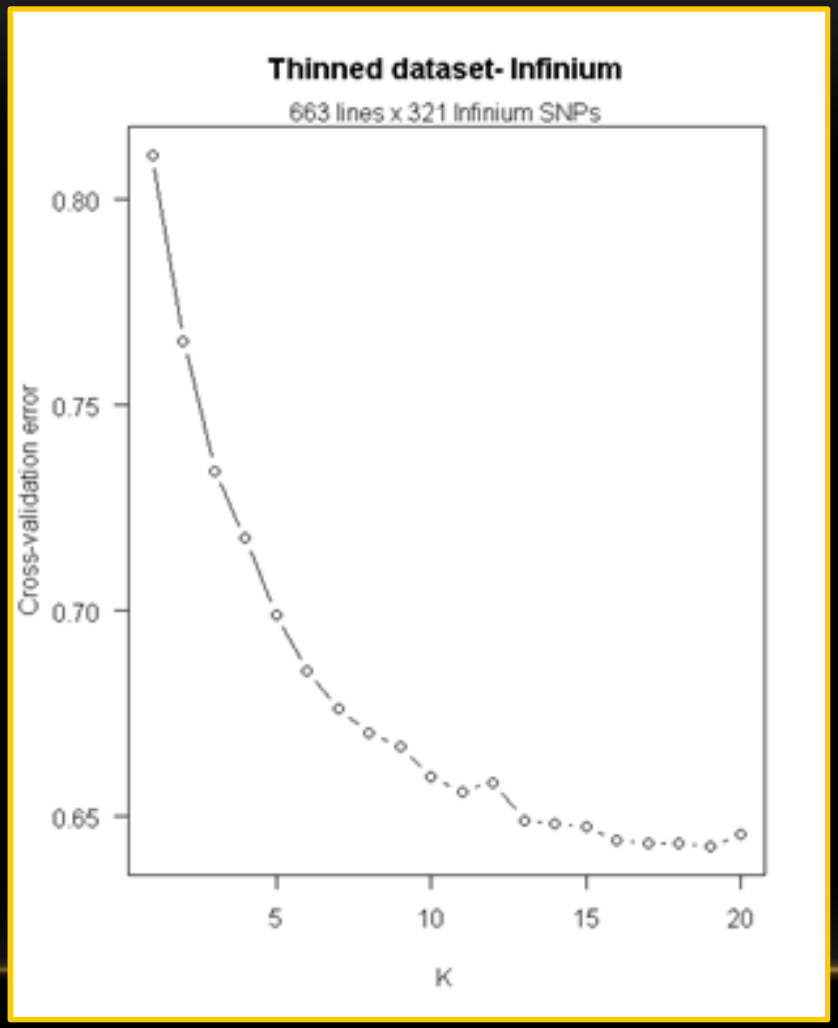
PCA ALSO CLUSTERS MANY ACCESSIONS BY BREEDING PROGRAM.



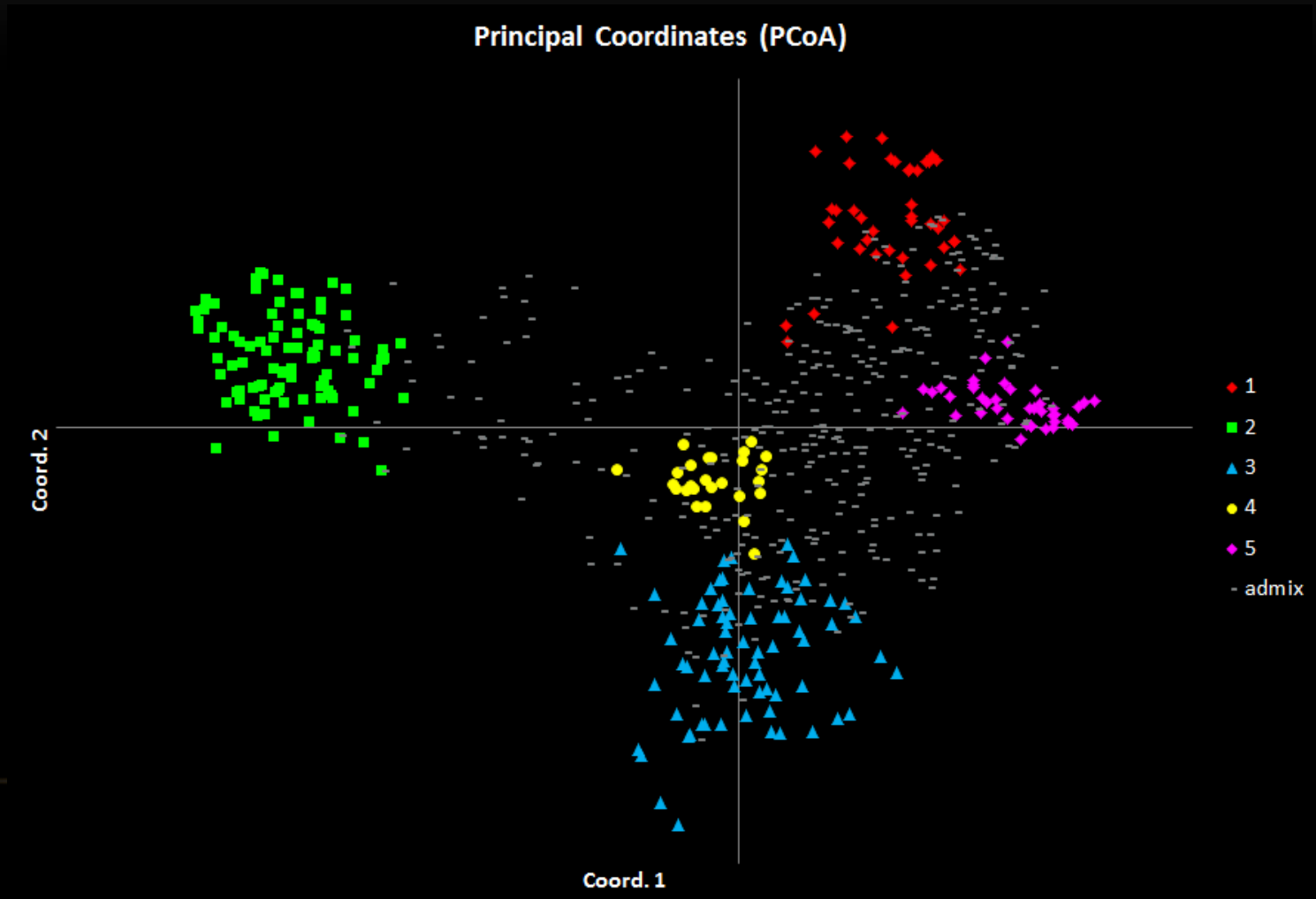
ACCESSIONS NOMINATED BY THE LSU, NC STATE, AND TX A&M PROGRAMS GENERALLY CLUSTER APART FROM ALL OTHERS.



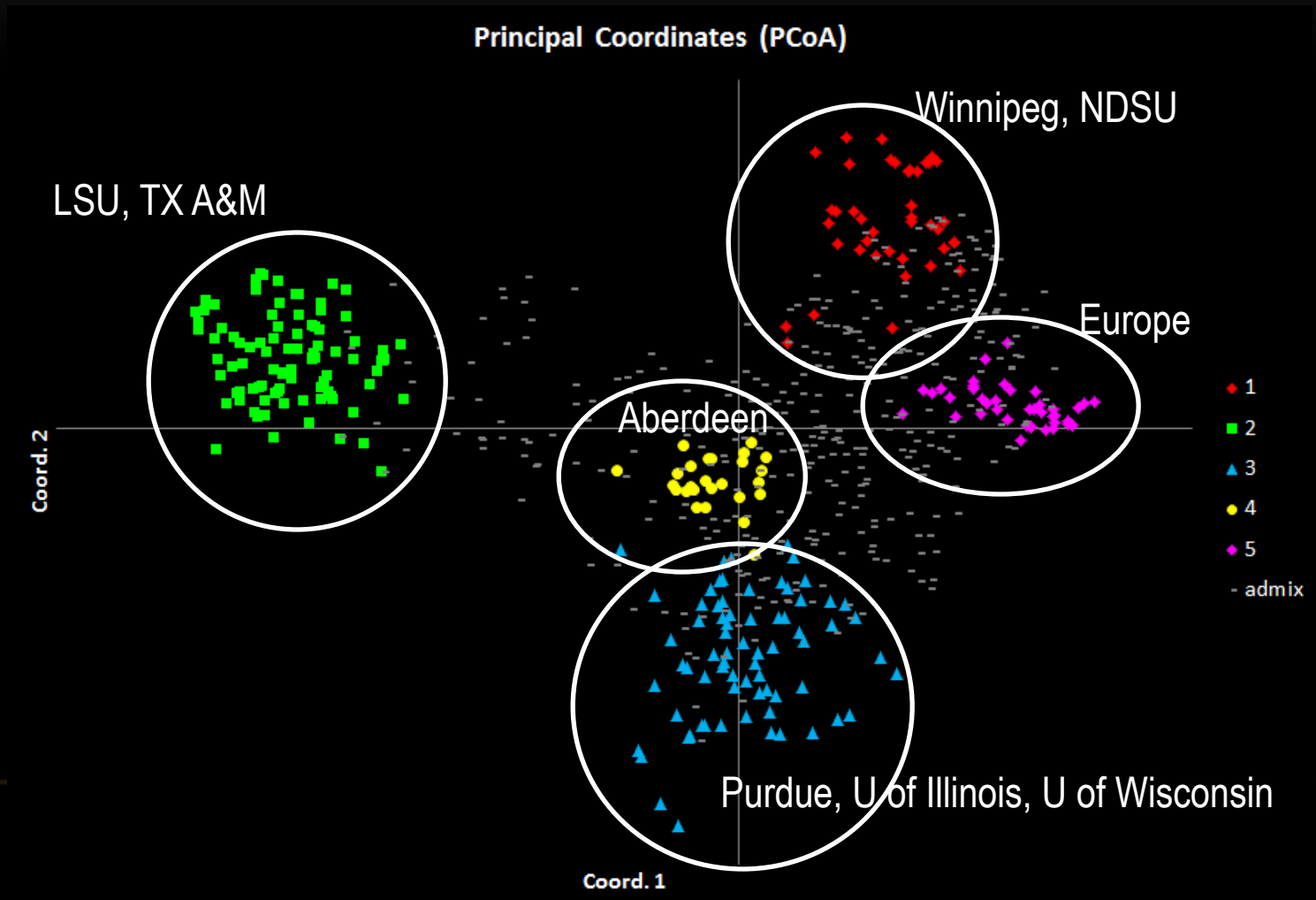
SUBPOPULATION NUMBERS INFERRED BY MODEL-BASED ANALYSES



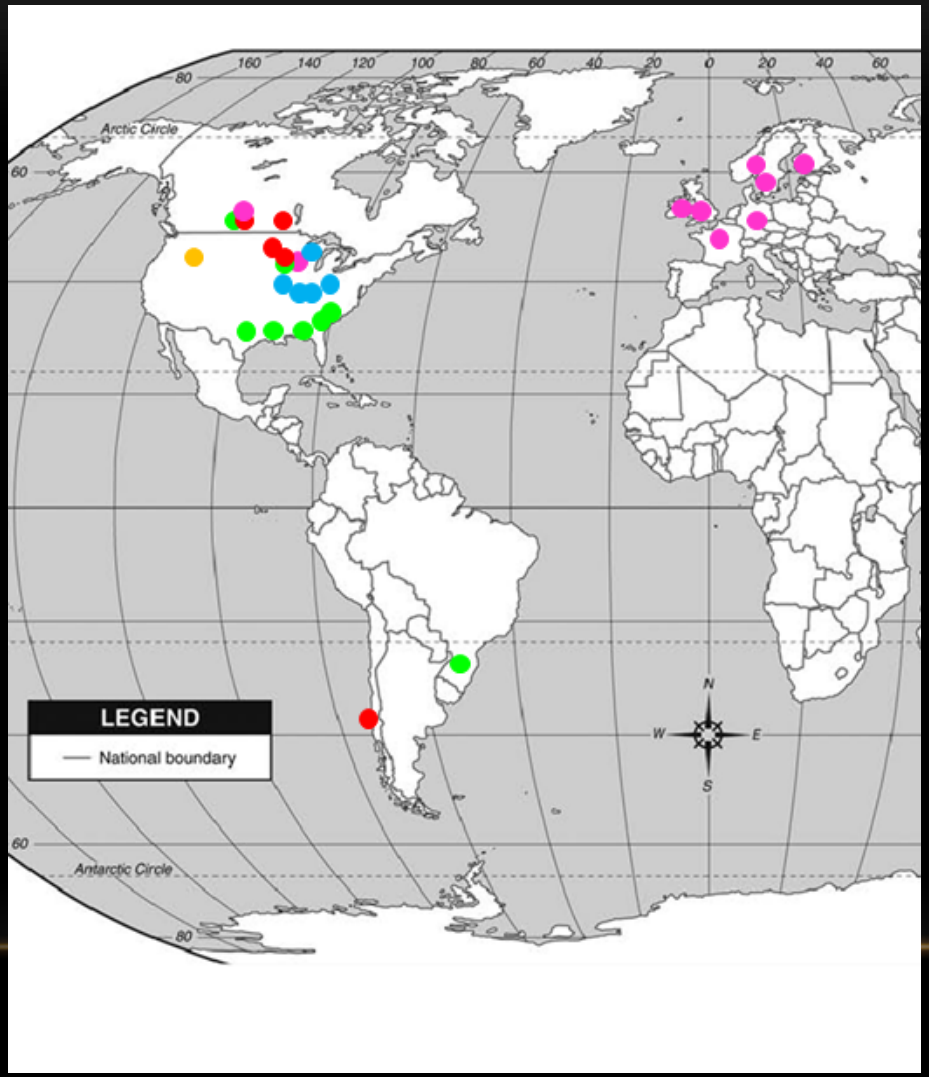
CLUSTERING OF LINES UNAMBIGUOUSLY ASSIGNED TO ONE OF 5 SUBPOPULATIONS



CLUSTERING OF LINES UNAMBIGUOUSLY ASSIGNED TO ONE OF 5 SUBPOPULATIONS

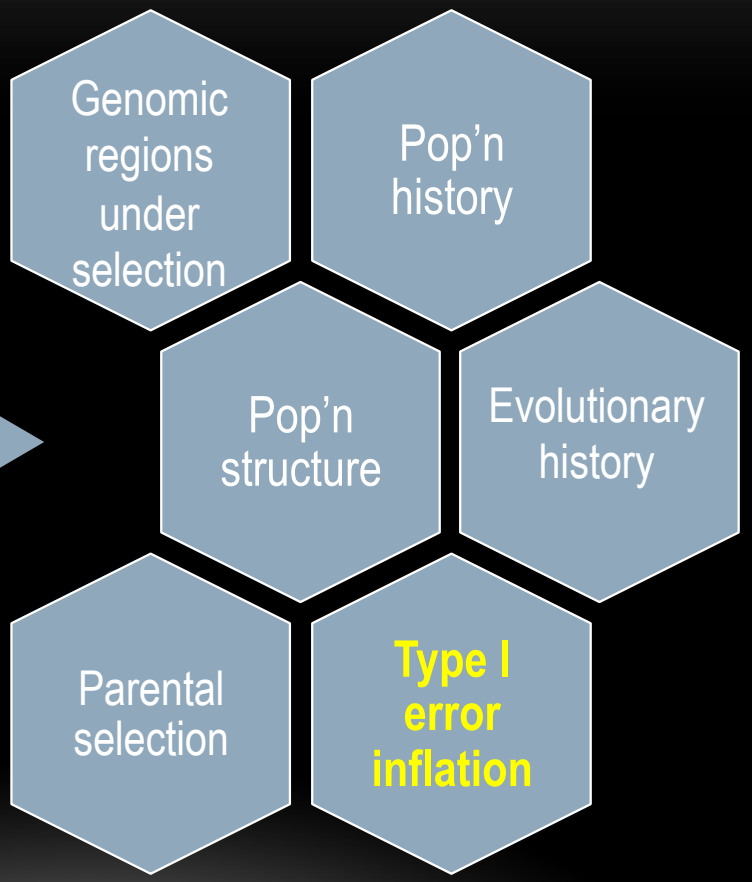


DISTRIBUTION OF ACCESSIONS ASSIGNED TO SUBPOPULATIONS BY BAYESIAN MODELLING

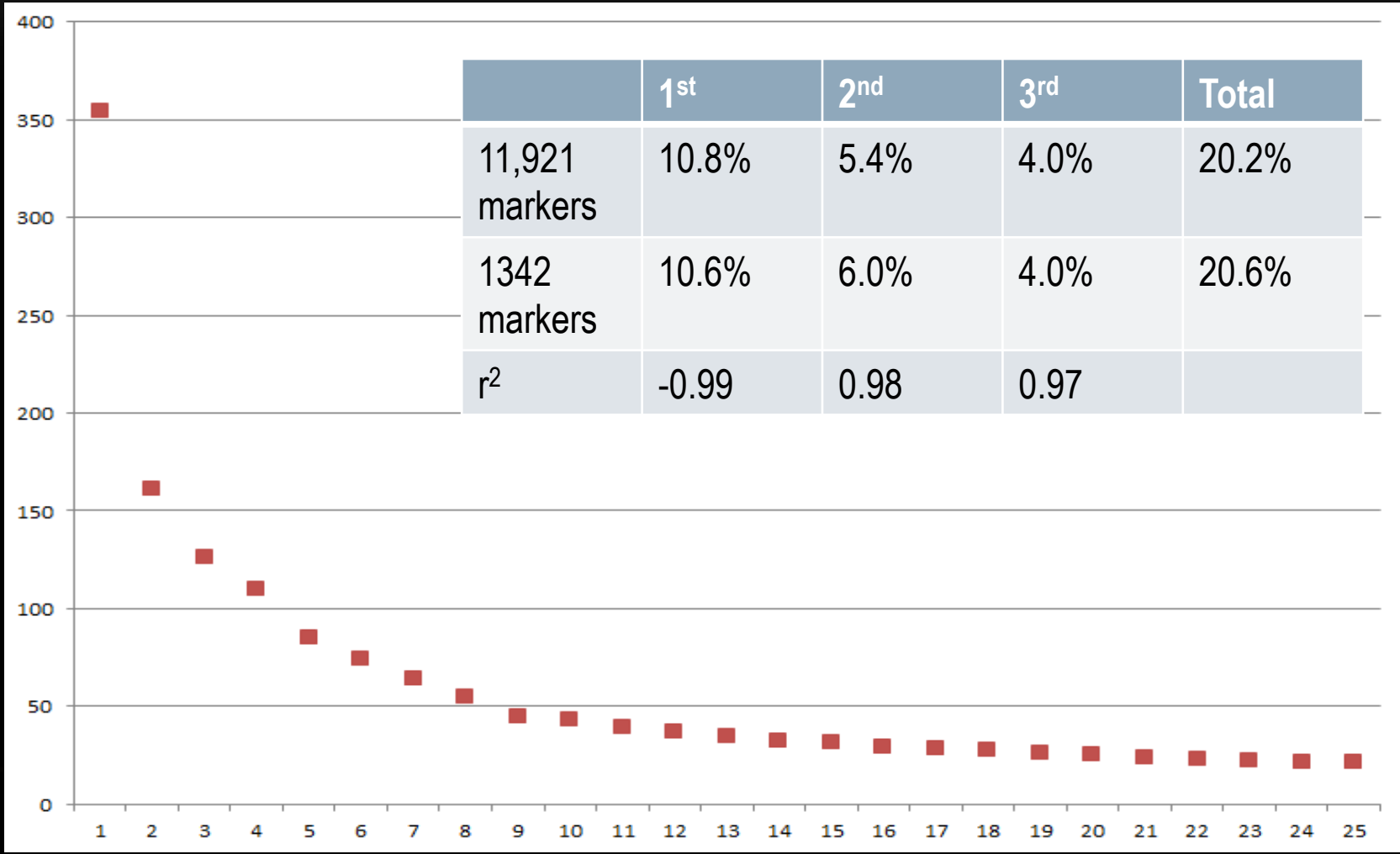


WHAT'S THE QUESTION?

Genetic Diversity

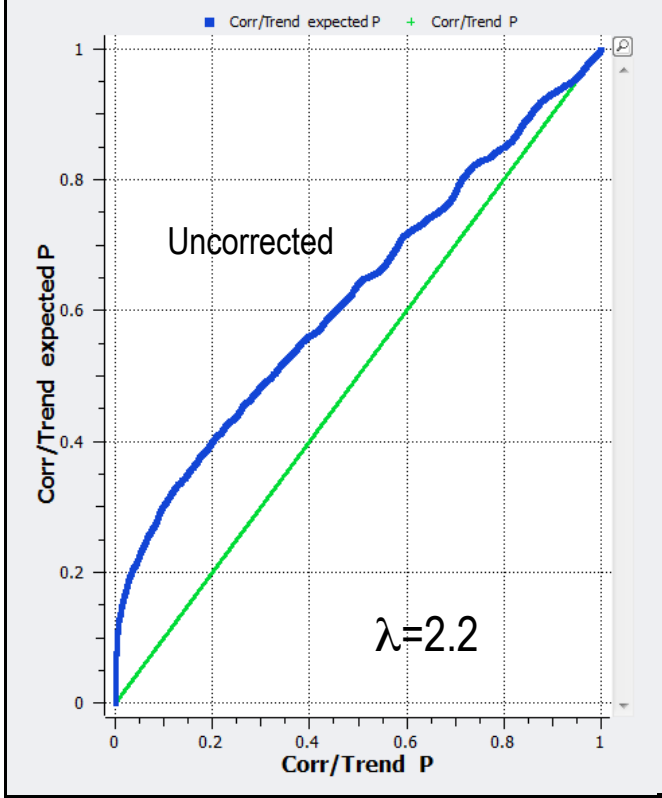


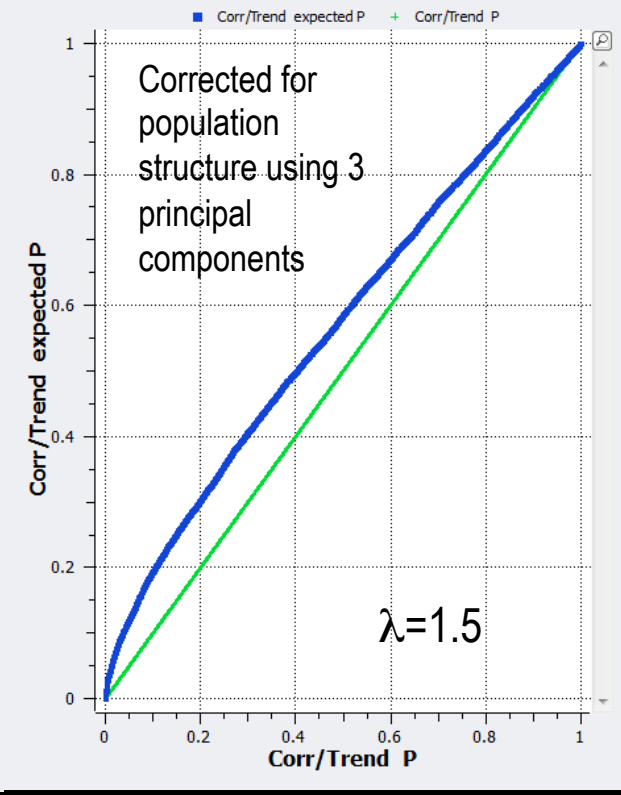
EIGENVALUES OF THE FIRST 25 PRINCIPAL COMPONENTS

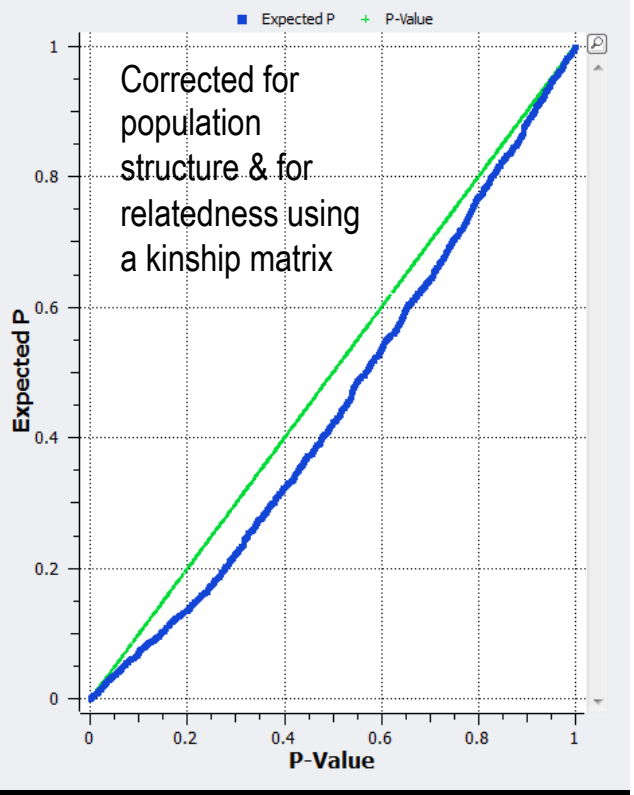


	1st	2nd	3rd	Total
11,921 markers	10.8%	5.4%	4.0%	20.2%
1342 markers	10.6%	6.0%	4.0%	20.6%
r^2	-0.99	0.98	0.97	

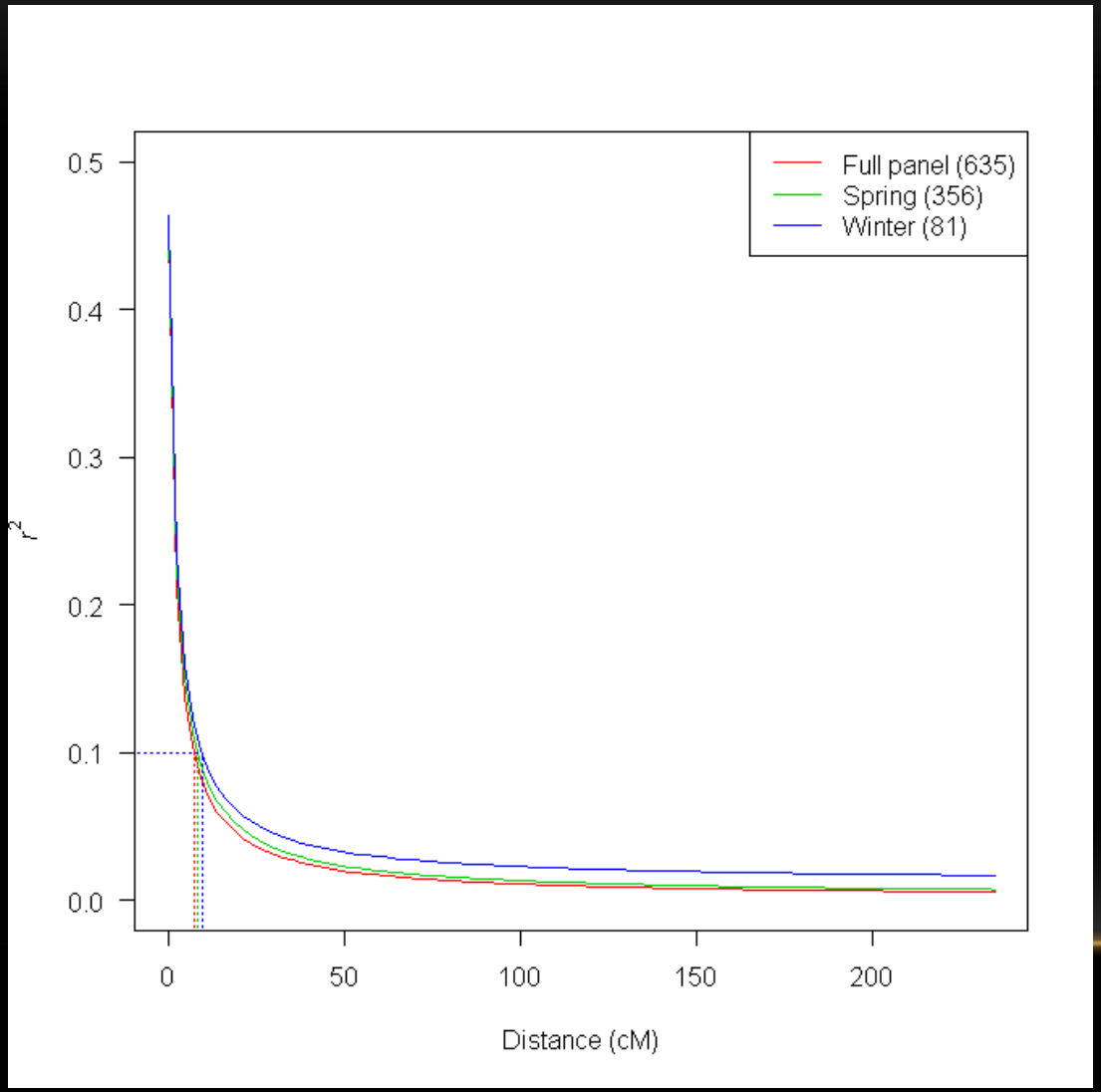
3 PRINCIPAL COMPONENTS SHOULD EFFECTIVELY REDUCE TYPE I ERROR INFLATION DUE TO POPULATION STRUCTURE







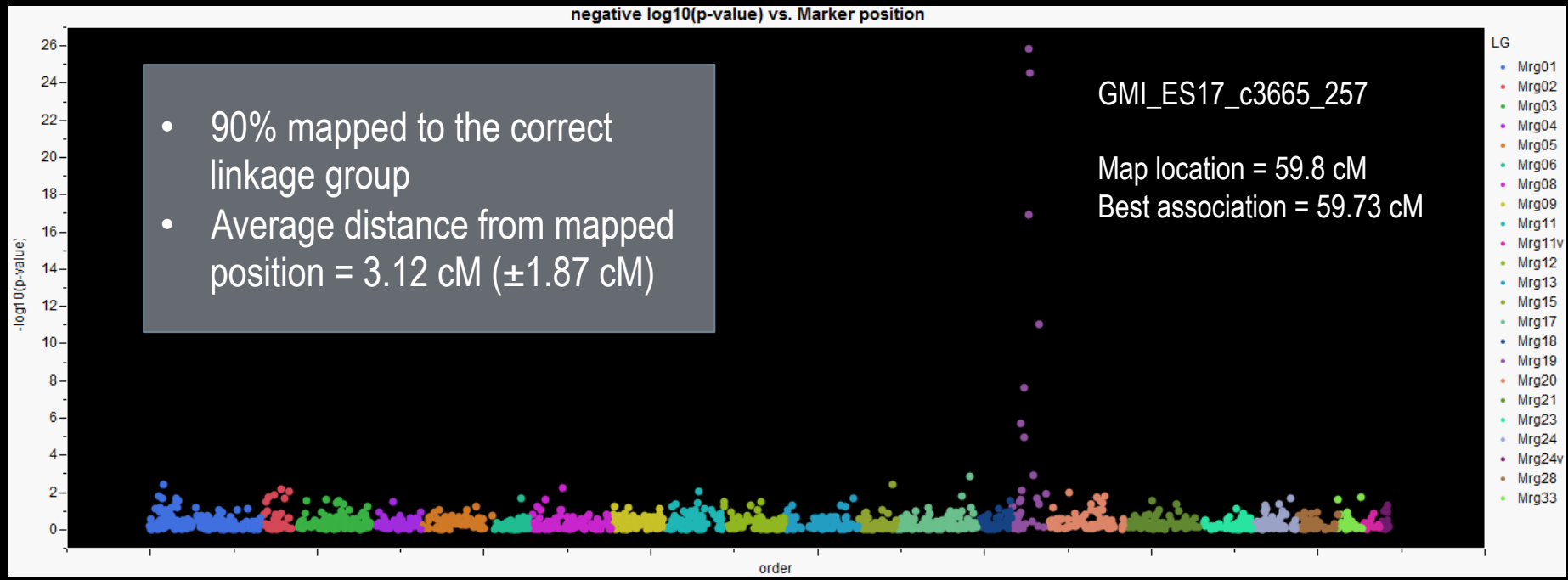
PAIRWISE LINKAGE DISEQUILIBRIUM BETWEEN MARKERS AS IT RELATES TO MAP DISTANCE



LD is the relative ability to predict genotype at an unknown locus using genotype information at a nearby locus.

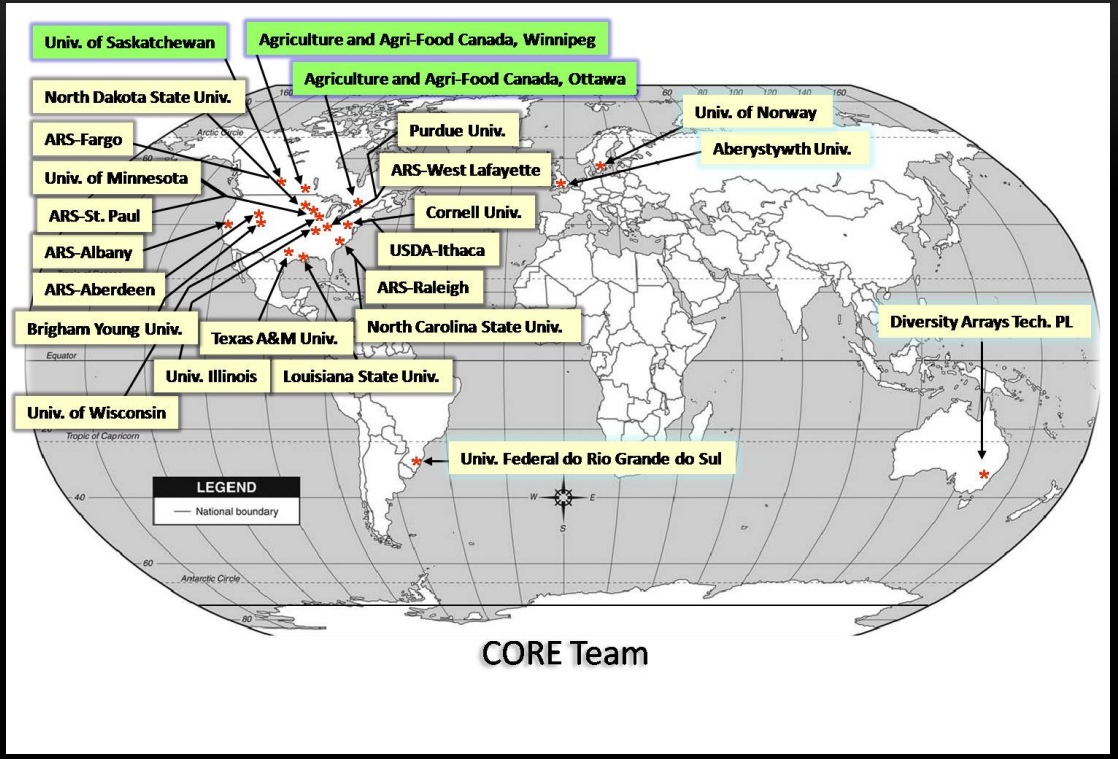
GWAS PERFORMANCE:

- 100 mapped markers randomly selected and converted to binary dependent variables
- GWAS performed for each marker with 1st 3 PC and a kinship matrix used to control for population structure
- Location of the best evidence of association was compared with the actual map location



THE ANALYSIS TEAM

- USDA-ARS
  - Michael Bonman
  - Ebraheim Babiker
  - Shiaoman Chao
  - Emir Islamovic
- Agriculture and Agri-Foods Canada
  - Nicholas Tinker
  - Yung-Fen Huang
  - Gnanesh Nanjappa



CORE Team

www.poga.ca
THE SOURCE FOR INFORMATION ABOUT THE OAT INDUSTRY ON THE CANADIAN PRAIRIES
Prairie Oat Growers Association

North American Millers' Association
NAMA

QUAKER

Agriculture and Agri-Food Canada | Agriculture et Agroalimentaire Canada

USDA Agricultural Research Service

GENERAL MILLS