Cornell University
College of Veterinary Medicine
Baker Institute for Animal Health

Genomic Analysis of the Nokota Horse

Summary

94 Nokota horses
Selected to be as unrelated as possible

Blood samples sent to Cornell
DNA extracted and stored in Nokota BioBank

DNA tested on Single Nucleotide Polymorphism Array containing 70,000 markers spread across the equine genome

First stage analysis using Principal Component Analysis (PCA)

Result: Nokota horses cluster as a group (breed), with highest similarity to the Standardbred Harness Horse breed, using 3 separate variations of PCA

Second stage analysis of Major Histocompatibility Complex (MHC) variation in the Nokota breed

AIM
to study the genetic makeup of the Nokota horse

METHODS
SNP assay & microsatellites

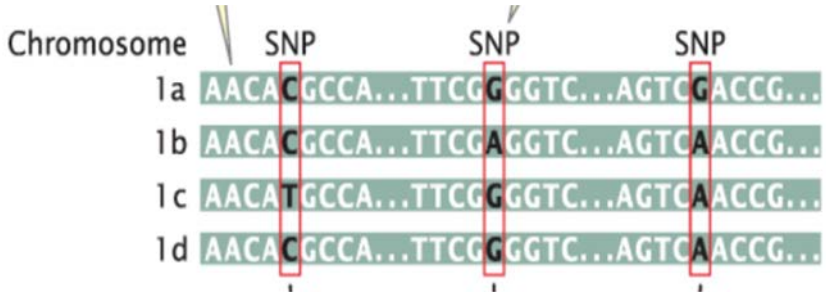
Definitions

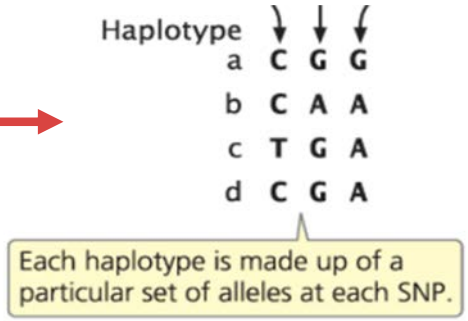
Single Nucleotide Polymorphism

- Loci with alleles that differ at a single basepair.
- Occur once every 1,000 basepairs on average.
- Occur in both exons and introns.

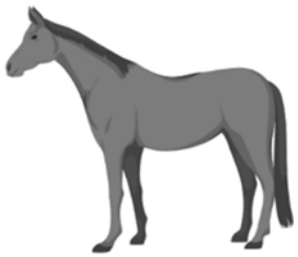
Haplotype

= constellation of alleles in a single region that is inherited.





SNP analysis of MHC



94 unrelated Nokota horses
(blood samples)

DNA extraction
with Qiagen DNEasy kit

PCA analysis

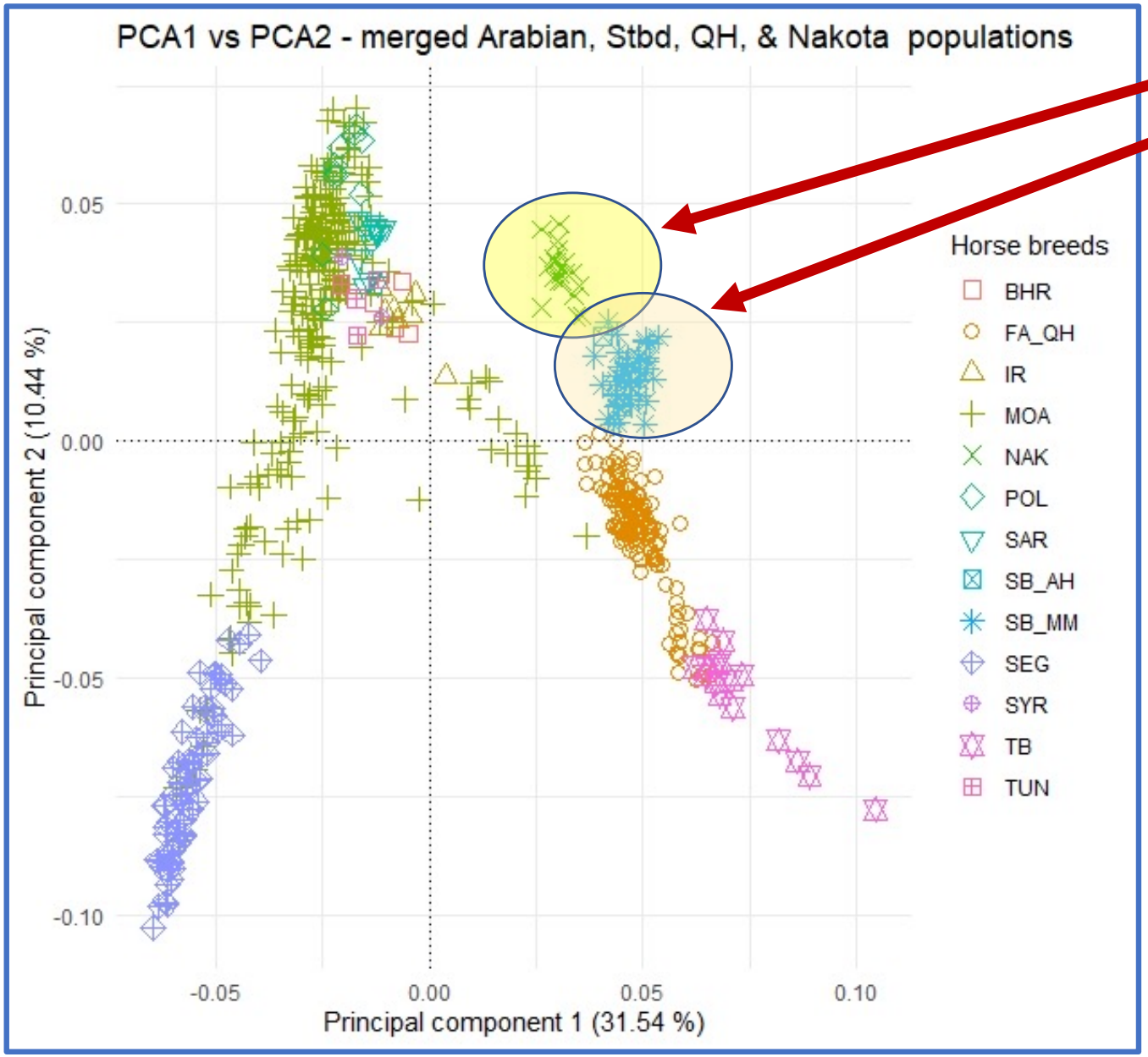
QC and filtering steps on McCue, QH, and Nokota horses

1. Filtered each population for the SNPs in HWE
2. Additional filtering – geno >90%, maf >0.01, mind >95%, only GGP70 SNPs
3. Combined AHS, McCue, & QH samples
4. Ran the PCA with --maf 0.02 --geno 0 --indep-pairwise 50 5 0.5 (bin, step, r2)

5. Filtered the Nakota horses for HWE SNPs, then kept only 20 random Nokotas with inbreeding coefficients of 0.
6. Combined the AHS, McCue, QH, and Nokota samples
7. Ran the PCA with --maf 0.02 --geno 0 --indep-pairwise 50 5 0.5 (bin, step, r2)

Plot 3. 587 Horses Merged Arabian, Stbd, QH, Nokota
PCA1 vs PCA2 - GGP70 SNPs only

7,905 SNPs

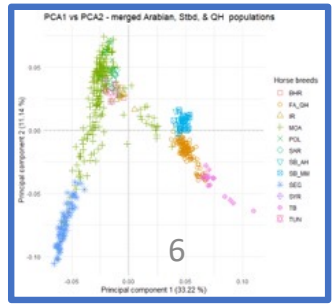


Nokotas

Standardbreds

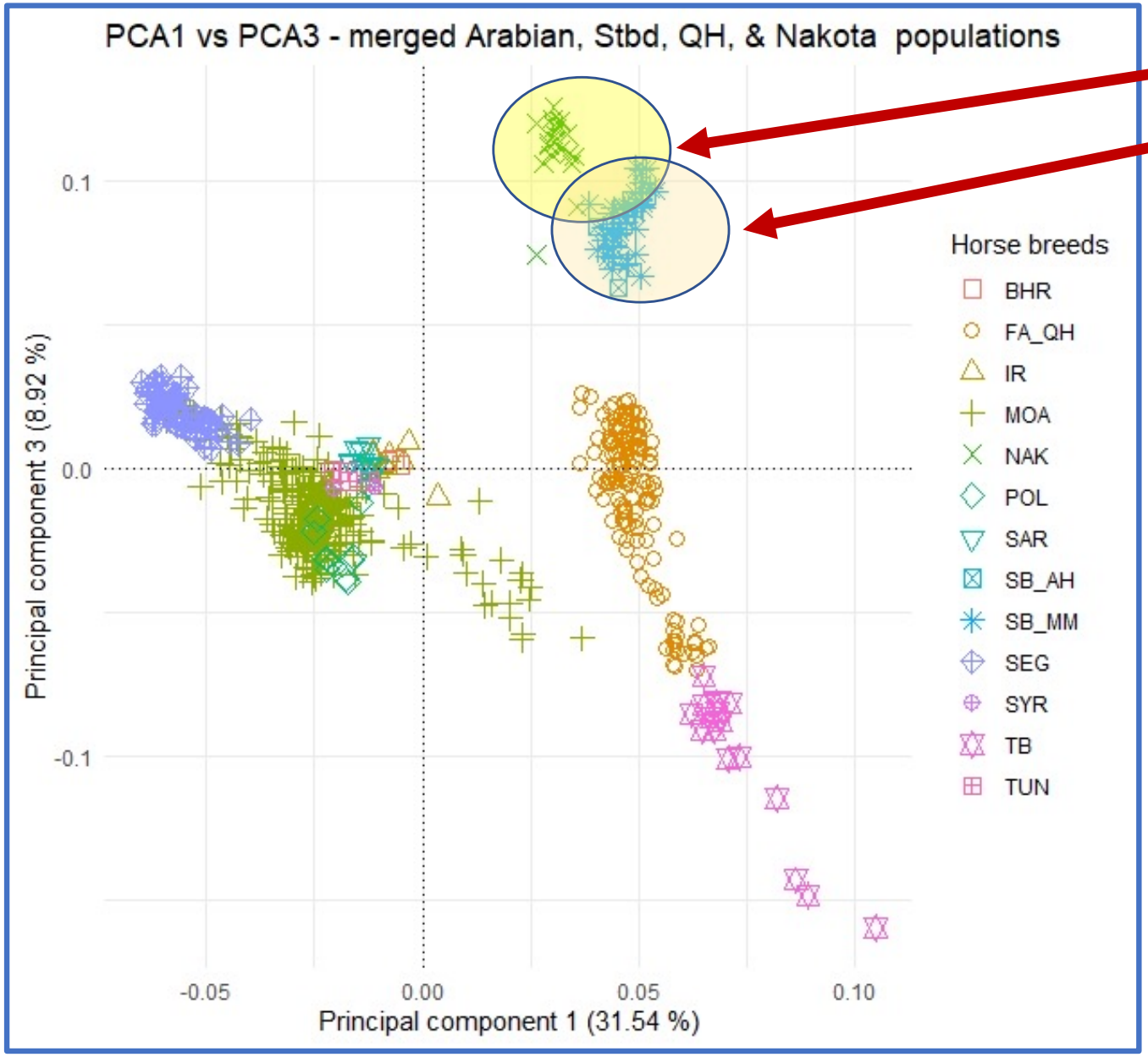
Bahrainian
Avila - QH
Iranian
Multi-origin ancestry
Nokota
Polish
Saudi Arabian
Standardbred – this study
Standardbred - McCue
Straight Egyptian
Syrian
Thoroughbred
Tunisian

9,925 SNPs



Plot 3. 587 Horses Merged Arabian, Stbd, QH, Nokota
PCA1 vs PCA3 - GGP70 SNPs only

7,905 SNPs

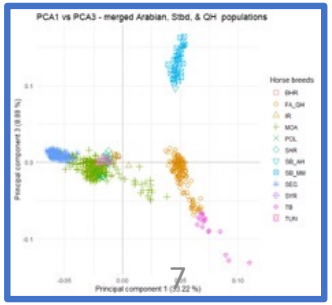


Nokotas

Standardbreds

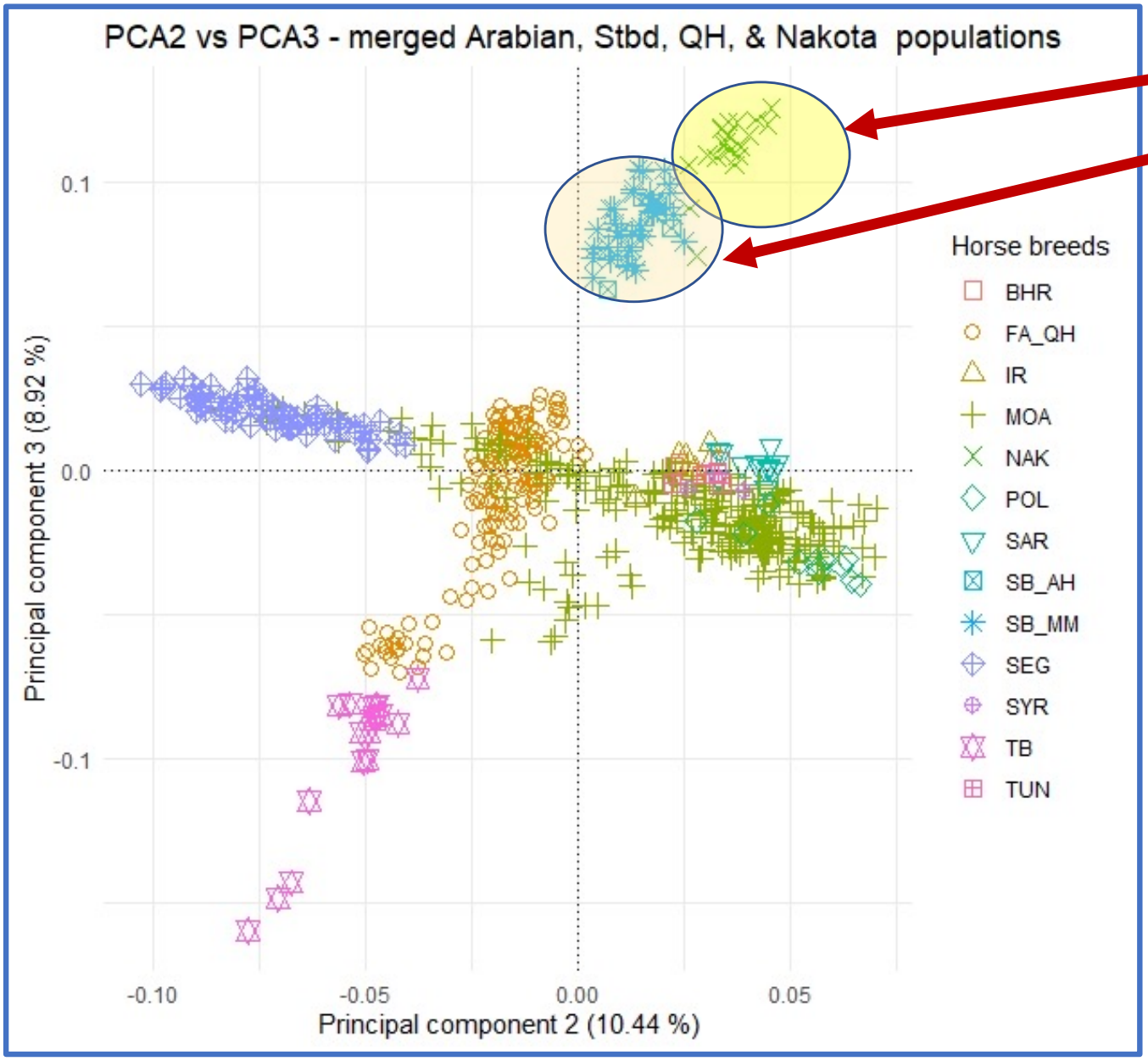
Bahrainian
Avila - QH
Iranian
Multi-origin ancestry
Nokota
Polish
Saudi Arabian
Standardbred – this study
Standardbred - McCue
Straight Egyptian
Syrian
Thoroughbred
Tunisian

9,925 SNPs



Plot 3. 587 Horses Merged Arabian, Stbd, QH, Nokota
PCA2 vs PCA3 - GGP70 SNPs only

7,905 SNPs

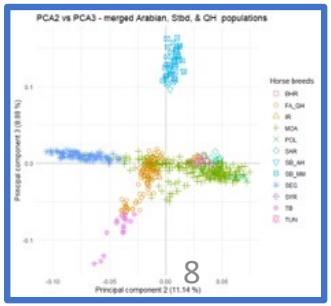


Nokotas

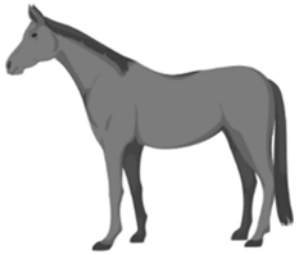
Standardbreds

Bahrainian
Avila - QH
Iranian
Multi-origin ancestry
Nokota
Polish
Saudi Arabian
Standardbred – this study
Standardbred - McCue
Straight Egyptian
Syrian
Thoroughbred
Tunisian

9,925 SNPs

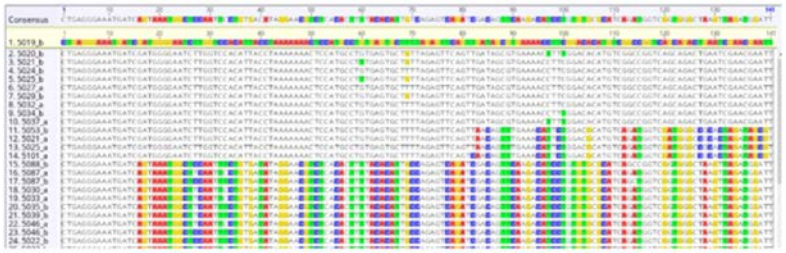


SNP analysis of Major Histocompatibility Complex (MHC) of Nokota horses



94 unrelated Nokota horses
(blood samples)

DNA extraction
with Qiagen DNEasy kit



SNP analysis of MHC class I, II & III

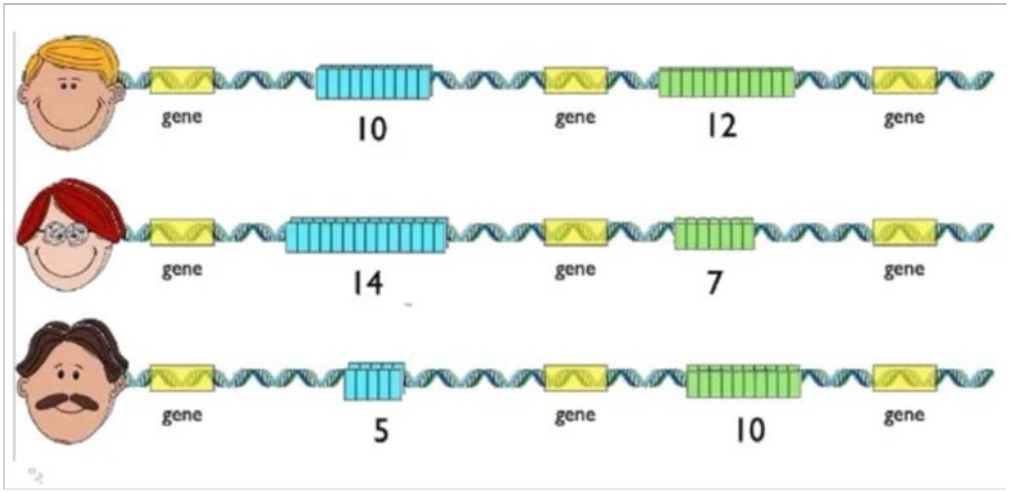
1. Neogen GGP70 SNP array
2. Phasing with SHAPEIT

Selection of 45 Nokota horses

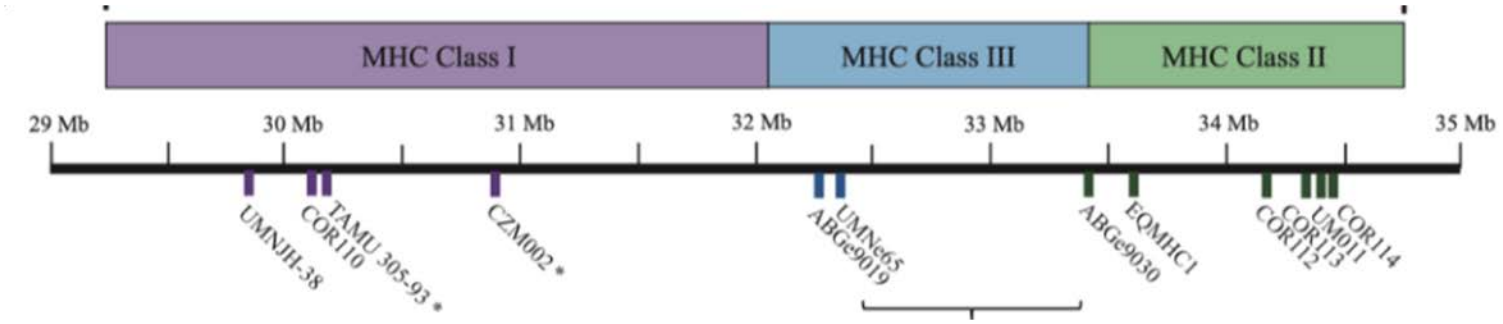
- MHC homozygotes
- MHC heterozygotes for the two most common haplotypes (A & B)

Goal of second stage analysisl:
confirm SNP MHC-haplotype = microsatellite MHC-haplotype assignments

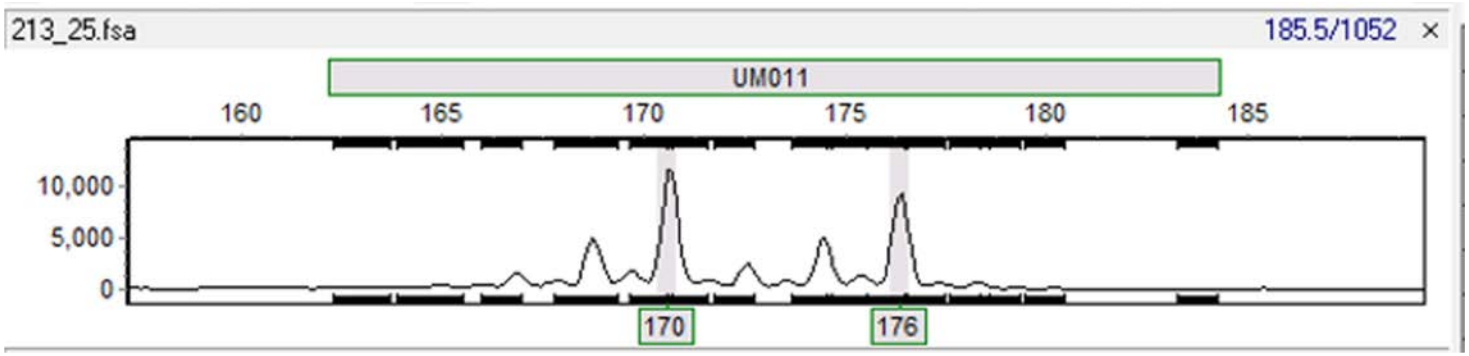
Microsatellites: how do they work?



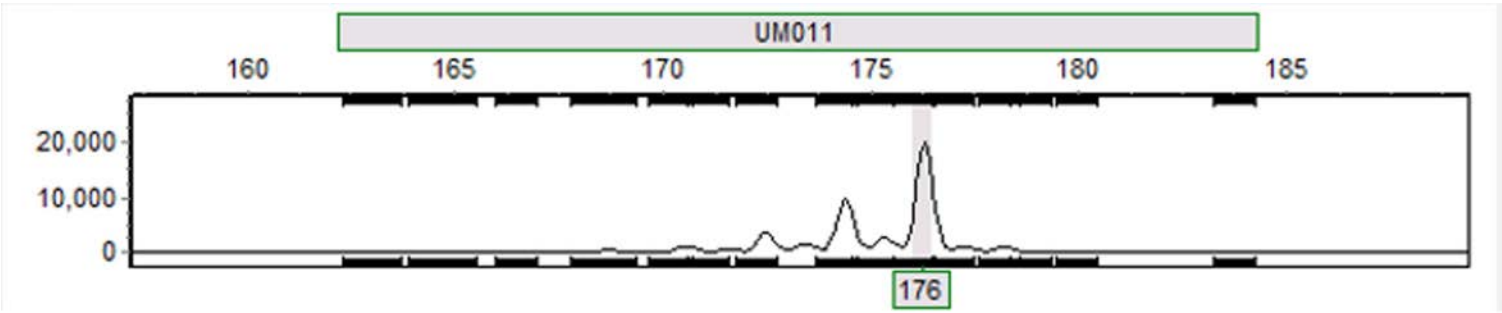
- Simple, repeating sequences of 2 – 6 basepairs. (**CT CT CT CT**)
- Can be repeated 3 – 100 times.
- Amount of repeats can differ between individuals.



Calling microsatellites alleles using Genemarker Software



Heterozygote: 2 peaks



Homozygote: 1 peak

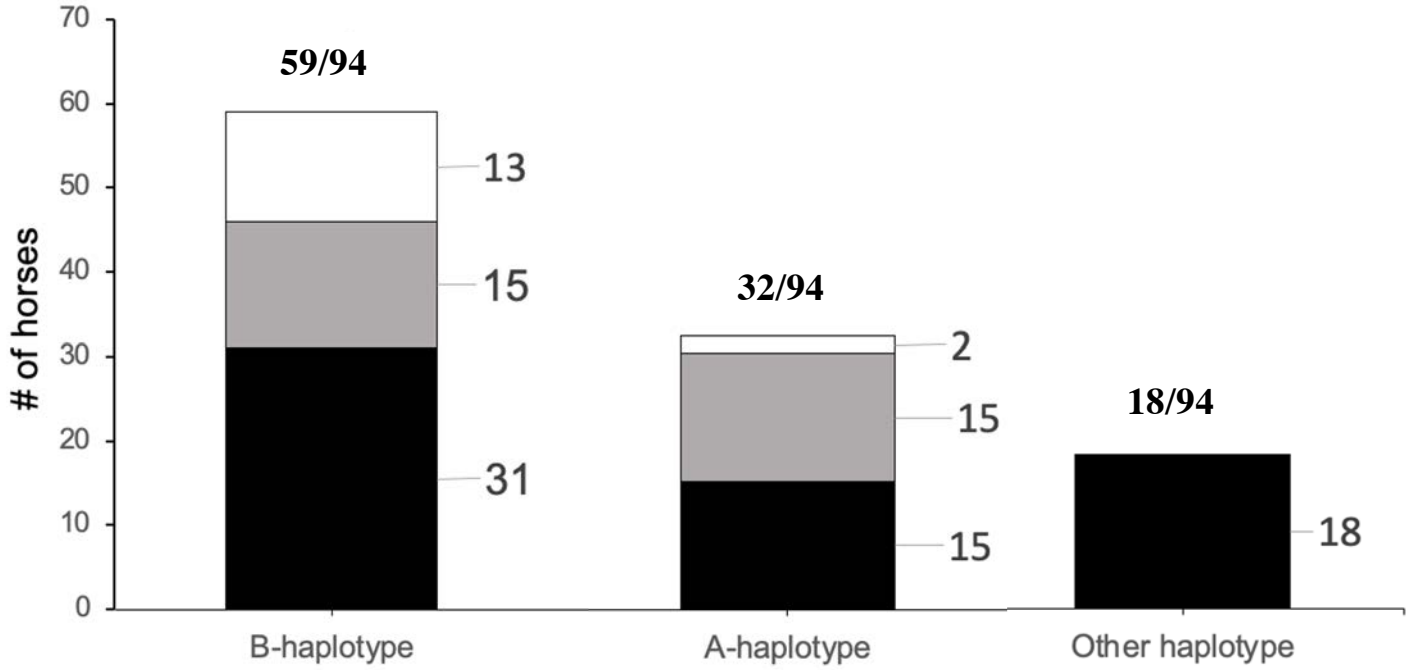
MHC haplotype calling with microsatellites

		I	I	I	I	III	III	II	II	II	II	II	II	
		UMNJH-38	COR110	305-93	CZM002	ABGe9019	UMNe65	ABGe9030	EQMHC1	COR112	COR113	UM011	COR114	
	Acc. #	29.814.528	30.140.839	30.197.833	30.925.265	32.249.535	32.339.323	33.454.044	33.599.372	34.192.010	34.390.587	34.419.882	34.426.065	
1.	5021_A/C	163	207	343	251	312	261	211	192	262	268	176	247	ELA-A3b
		163	207	343	251	312	261	**	196	264	266	170	245	C
2.	5025_A/C	163	207	343	251	312	261	211	192	262	268	176	247	ELA-A3b
		163	207	343	251	312	261	**	196	264	266	170	245	C
3.	5038_D/E	156	207	345	247	307	255	**	196	264	266	170	245	D
		156	209	343	261	316	249	211	184	252	280	172	253	E
4.	5045_A/A	163	207	343	251	312	261	211	192	262	268	176	247	ELA-A3b
		163	207	343	251	312	261	211	192	262	268	176	247	ELA-A3b
5.	5046_B/B	156	221	342	259	305	259	205	194	260	274	172	247	B
		156	221	342	259	305	259	205	194	260	274	172	247	B

Microsatellites confirm SNP analysis is correct:

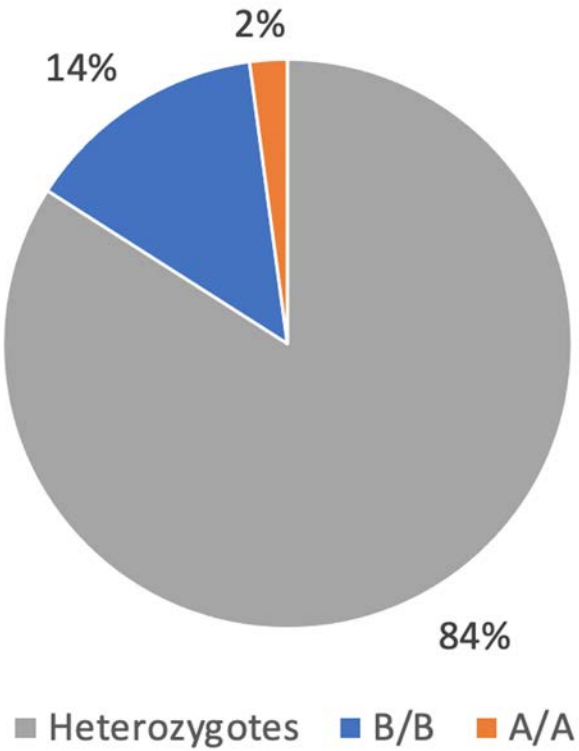
12 different haplotypes, 9 of which are new

2. SNP analysis detects 2 MHC haplotypes at high frequency



***White** = MHC homozygotes, **grey** = A/B MHC heterozygotes, **black** = B/x, A/x, and x/y heterozygotes.*

3. We found a high level of MHC homozygosity in the Nokota horse: 16%



Persian Arabians: 3 out of 124 horses → **2.4%**
(Sadeghi et al., 2018)

Icelandic Horses: 1 out of 156 horses → **< 1%**
(Holmes et al., 2019)

Take-away

➢There was a high correlation between MHC haplotype assignments made, using either SNPs or microsatellites.

➢Nokota horses appear to have exceptionally **high frequency of two haplotypes** compared to what was observed in other horse breeds. As a consequence there was **a high level of homozygosity.**

Future steps

Sample more Nokota's to find out some about the extent of the variation in haplotypes.

Look for evidence of recombination within the MHC → driver for genetic diversity

Acknowledgements

The Antczak Lab

Dr. Doug Antczak

Don Miller

Josephine Marchand

Maya Kulikowski

Scott Hoffay

Brandon Garcia

Leadership Program

Dr. John Parker

Dr. Gerlinde van de Walle

Dr. David Fraser

Elaine Lu

Cornell University
College of Veterinary Medicine
Baker Institute for Animal Health