

# Including genetic diversity as part of the international breeding goal for Holsteins

Effective Use of Subpopulations for Sustained Genetic Improvement.

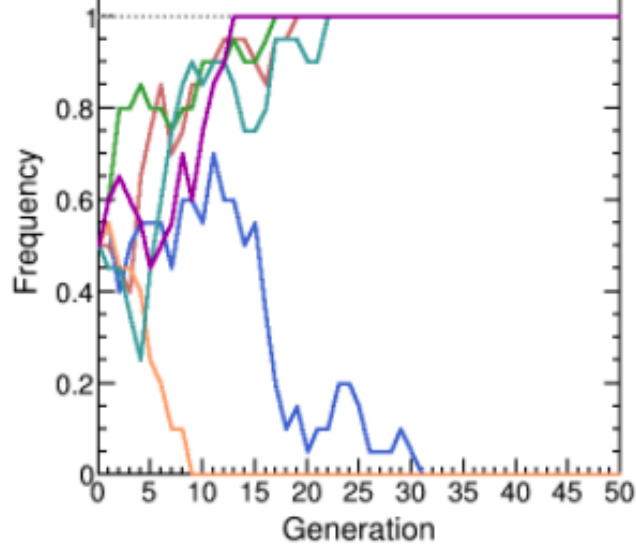
Thomas Lawlor, Holstein Association Inc., USA

Yvette Steyn, University of Georgia

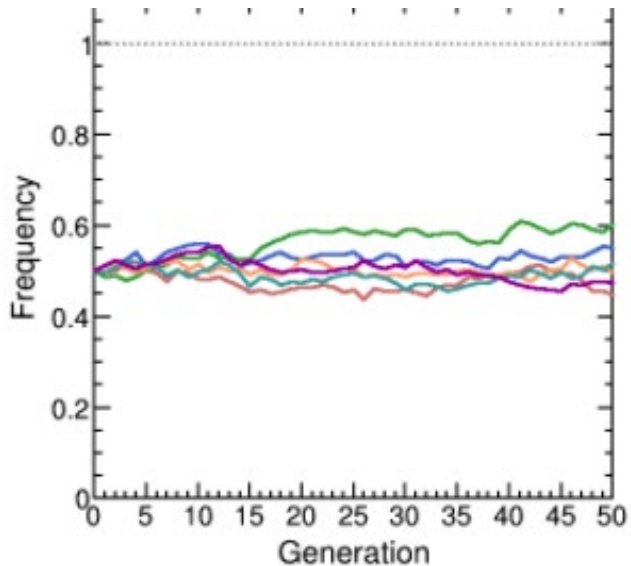


University of Georgia





Fisher described a single large population ... where alleles are described as being GOOD or BAD.



Sewall Wright describes subpopulations ... where each group is searching for the right COMBINATION of alleles

**Different subpopulations change the frequency of different alleles ..... fewer alleles will become fixed in the global population.**

# Evolve and Resequence

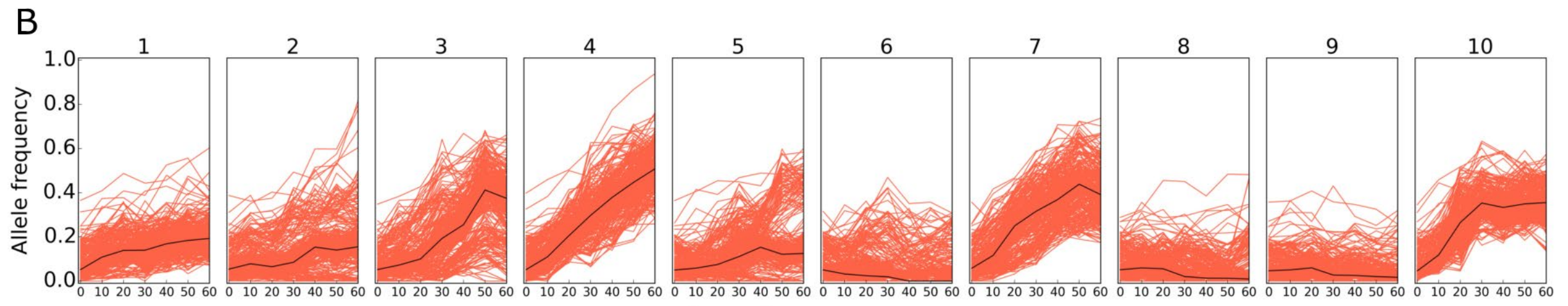
## Genetic redundancy fuels polygenic adaptation in *Drosophila*

Barghi et al., 2019

A single population was divided into ten sub-populations.



All subpopulations adapted to the higher temperature.



They observed different allele frequency changes in the different subpopulations

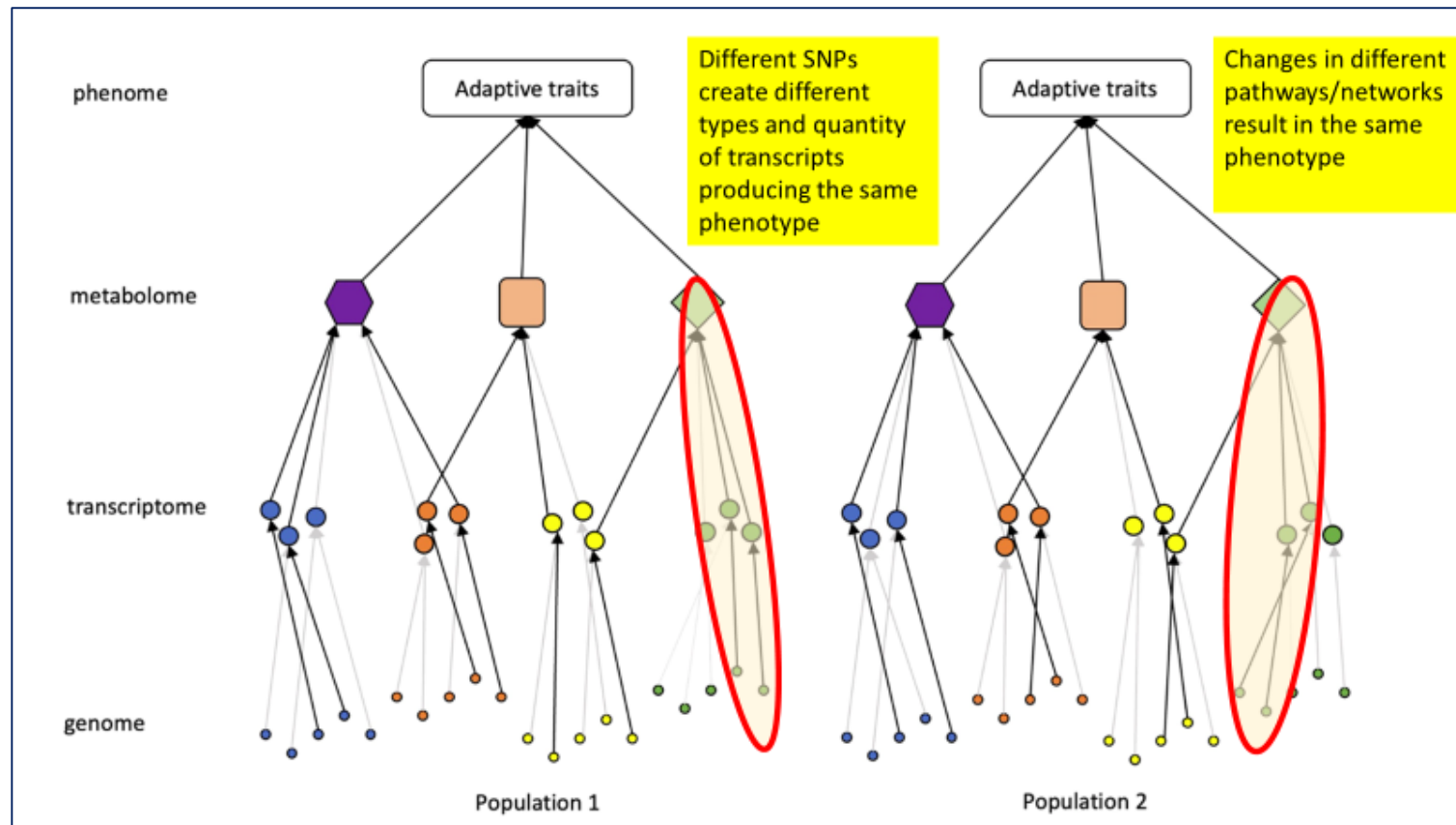
•The **Different combinations of genes and genetic networks were used to adapt.**

**Gene regulation** was the major driver of the adaptation to higher temperatures.

**Different SNPs** lead to different types and quantity of transcripts being created.

**Different pathways/networks** were utilized to produce the same phenotype.

Lai et al, 2023



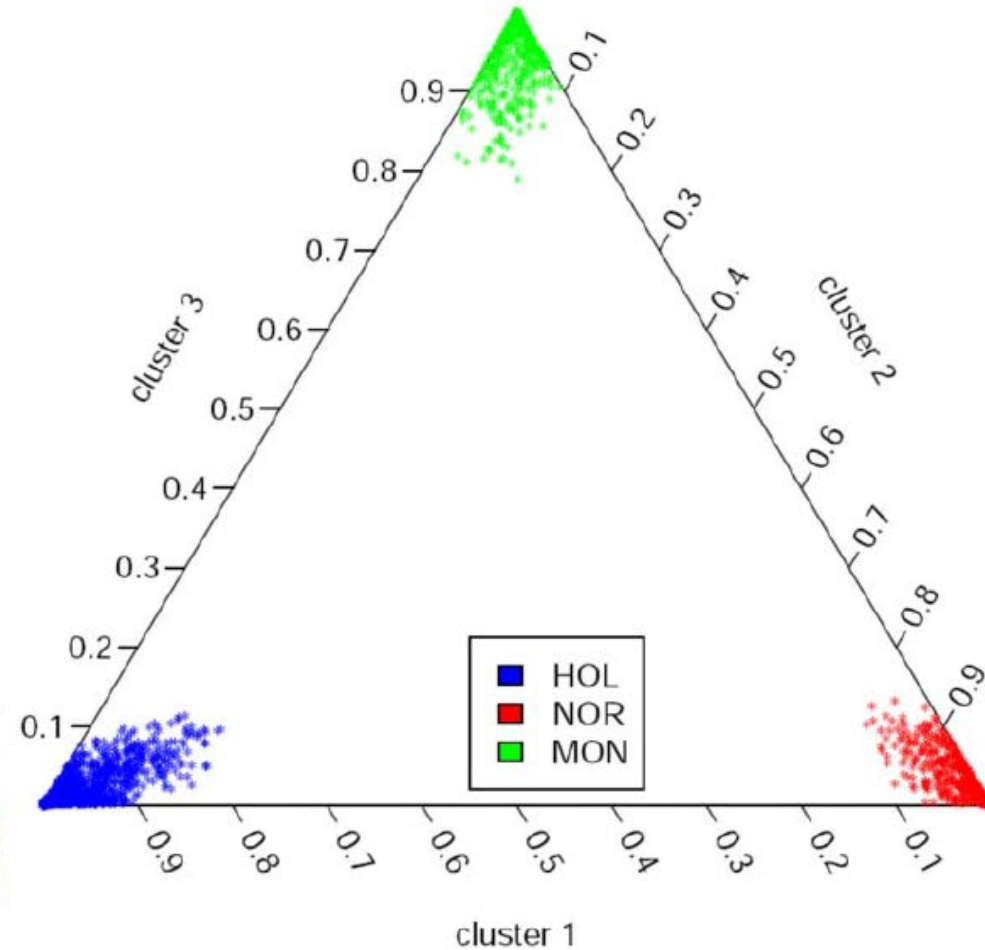
# The Genome Response to Artificial Selection: A Case Study in Dairy Cattle (Flori et al., 2009)



The average  $F_{st}$  value for markers across three different dairy breeds

**Holstein, Montbéliarde, and Normande was 0.07**

$F_{st}$  is a measure of the allele frequency differences between populations

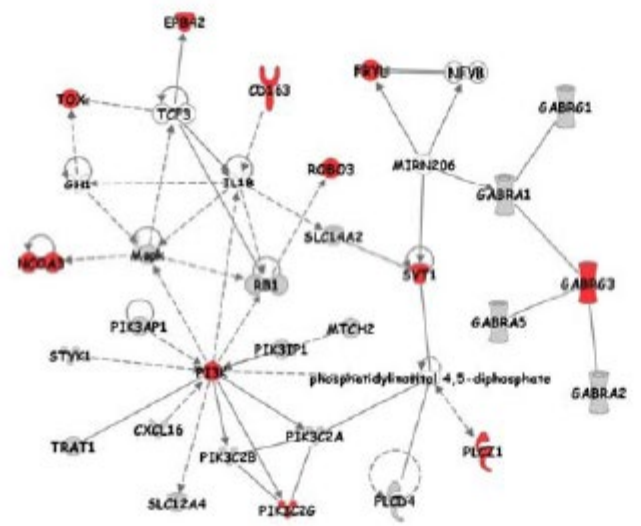
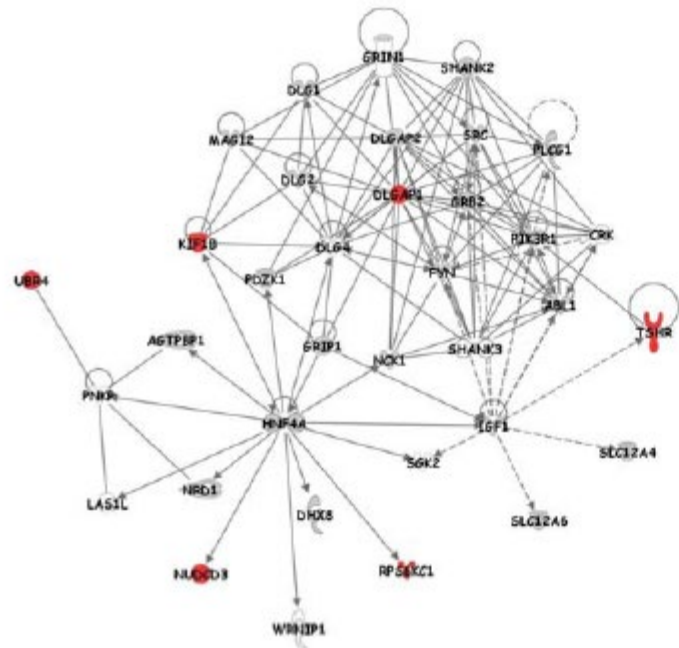


**Figure 1. Population structure.** The triangle plot represents the estimated membership of each 2803 bulls in each of the 3 assumed clusters. Each bull is represented by a point colored according to its breed of origin.

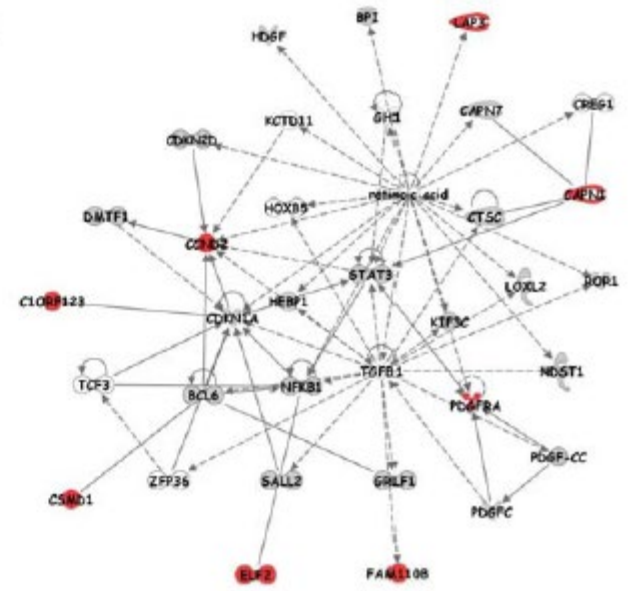
Each breed used a different gene network to improve its performance.

“Although centered on the same physiological pathways, set of differentiated genes were almost not overlapping among the breeds. This suggests a kind of plasticity in the genome allowing different solutions to respond to a

A



B

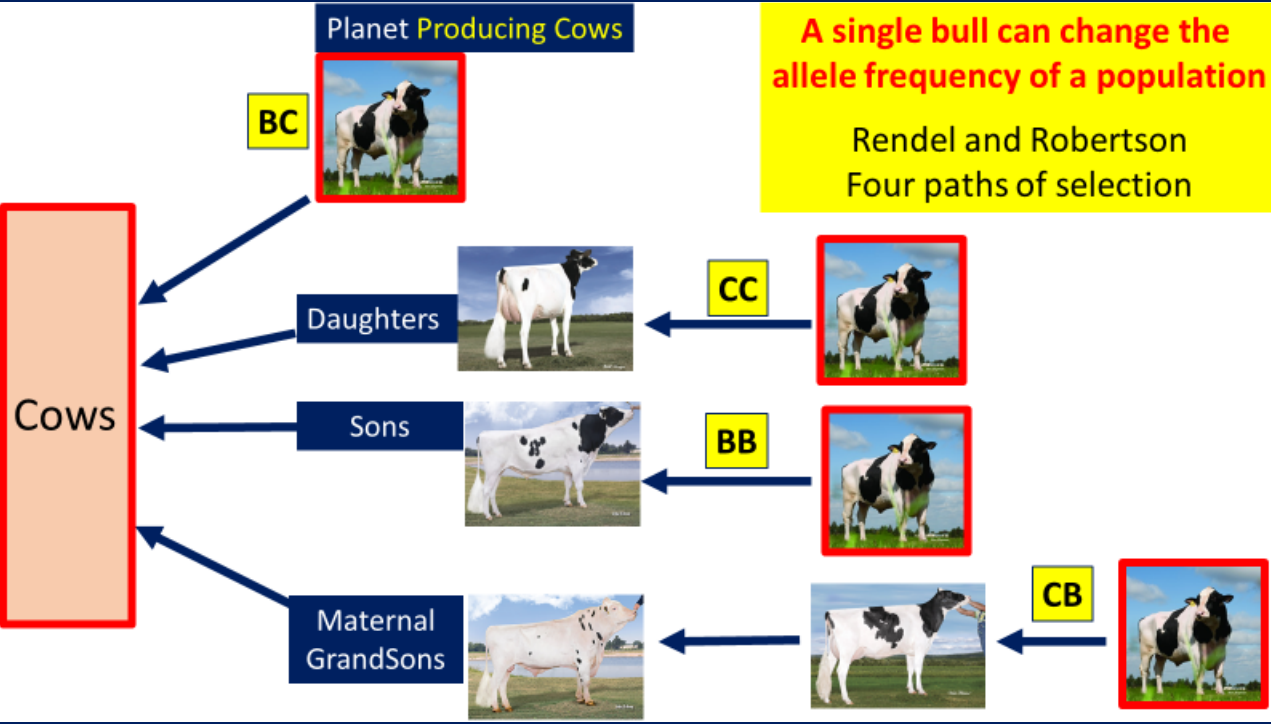


- Complex
- ▽ Cytokine/Growth Factor
- ◻ Chemical/Toxicant
- ⚡ Enzyme
- ⌘ G-protein Coupled Receptor
- ⊖ Group/Complex/Other
- ⬆ Growth factor
- ⌘ Ion Channel
- ⚡ Kinase
- ⌘ Peptidase
- ⌘ Transcription Regulator
- ⌘ Transmembrane Receptor
- ⌘ Transporter
- Unknown
- Direct Relation
- - - Indirect Relation

Figure 5. Representation of the gene networks N\_MON (A), N\_NOR (B) and N\_HOL (C). Symbols corresponding to candidate genes are colored in red. Genes colored in grey were represented in our study but did not display any evidence of selection.

Prominent bulls with high reproductive value can create subpopulations.

A single bull can change the allele frequency of a population  
 Rendel and Robertson  
 Four paths of selection



Animals with genetic evaluations

Signal Strength

Gene Source	Average Year of birth	Number of Descendants	Relationship to Planet	Genome copies
CB	2015.8	622,905	0.125	77,863
CC	2014.8	51,683	0.25	12,921
BB	2013.7	363,164	0.25	90,791
BC	2012.2	58,288	0.5	29,144

K-means clustering allows us to cluster or group animals with the same genetic relationship together

## Prominent bulls for different clusters

Top 4 sires for Cluster A	Number of genotyped offspring
Planet	843
Observer – Planet son	594
Shamrock – Planet son	593
Bookem – Planet son	424

Top 4 sires for Cluster D	Number of genotyped offspring
AltaIOTA – O Man son	436
Man-O-Man O Man son	434
Freddie – O Man son	382
Gerard – O Man son	199



**Genetic contribution (%)**  
of prominent bulls in each of the five clusters

Cluster	Planet	Goldwyn	Shottle	O Man	Several
1	<b>28.1</b>	1.0	5.1	4.2	0.3
2	0.6	<b>18.8</b>	3.7	1.5	0.6
3	0.6	2.7	<b>19.8</b>	1.2	1.6
4	1.0	2.0	2.7	<b>21.6</b>	0.6
5	0.2	0.5	2.4	2.0	4.3

**28.1%** of all genes in **Cluster 1** trace back to **Planet**  
**18.8%** of all genes in **Cluster 2** trace back to **Goldwyn**  
 etc.

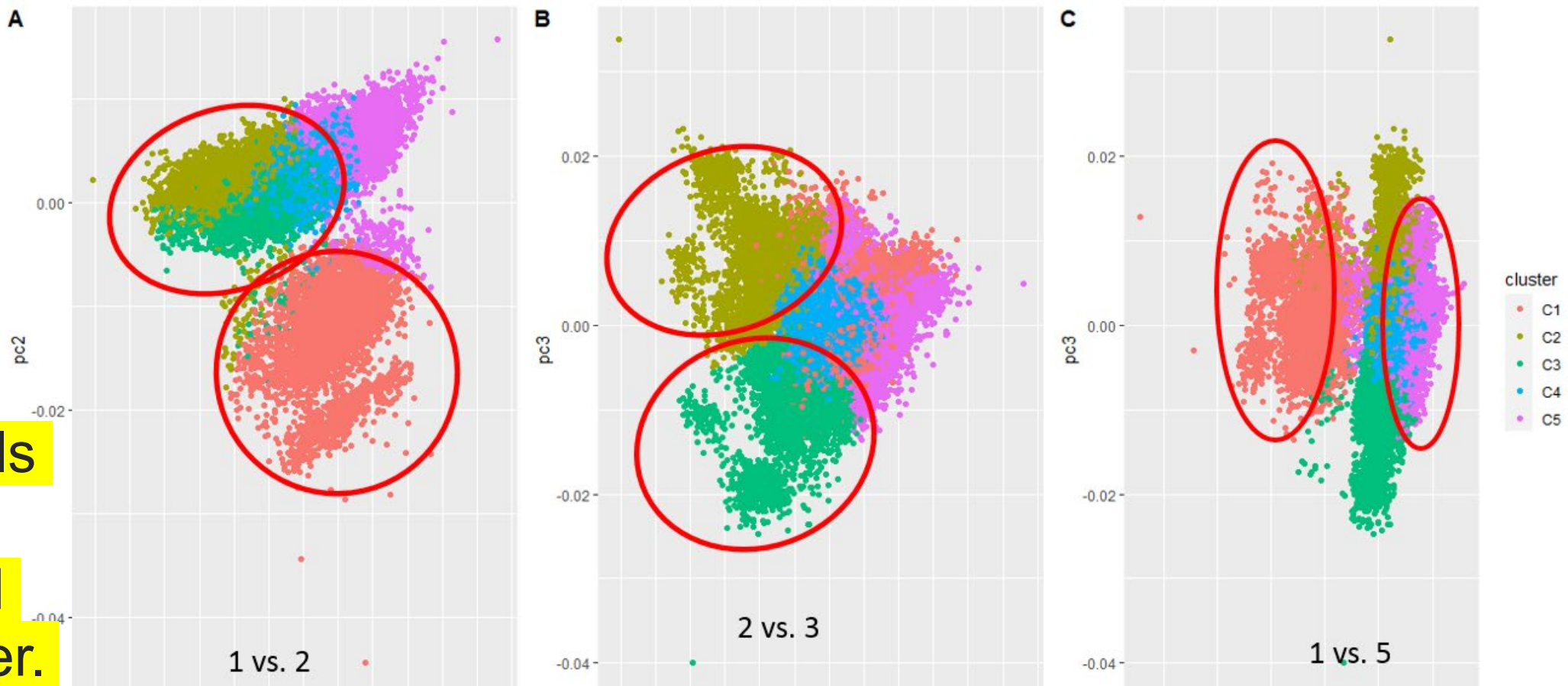


**Average  $F_{st}$  value across Holstein clusters was 0.03**

**Indicating average allele frequency differs between the clusters.**

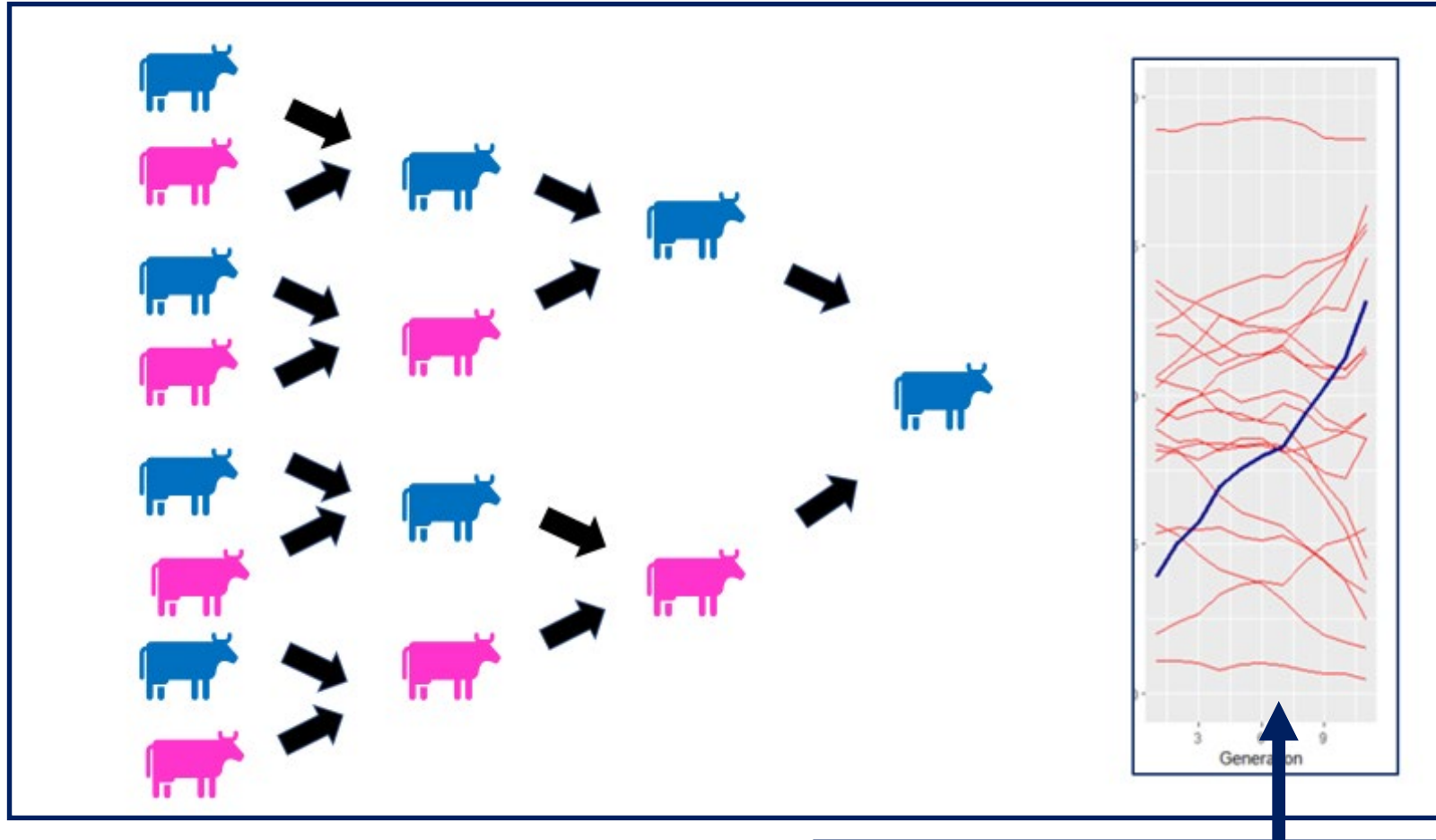


Animals with similar allele frequencies cluster together. These animals are more highly related to one another.



# Pedigrees of animals within a cluster were traced back 10 generations

Allele frequency was calculated for each generation.

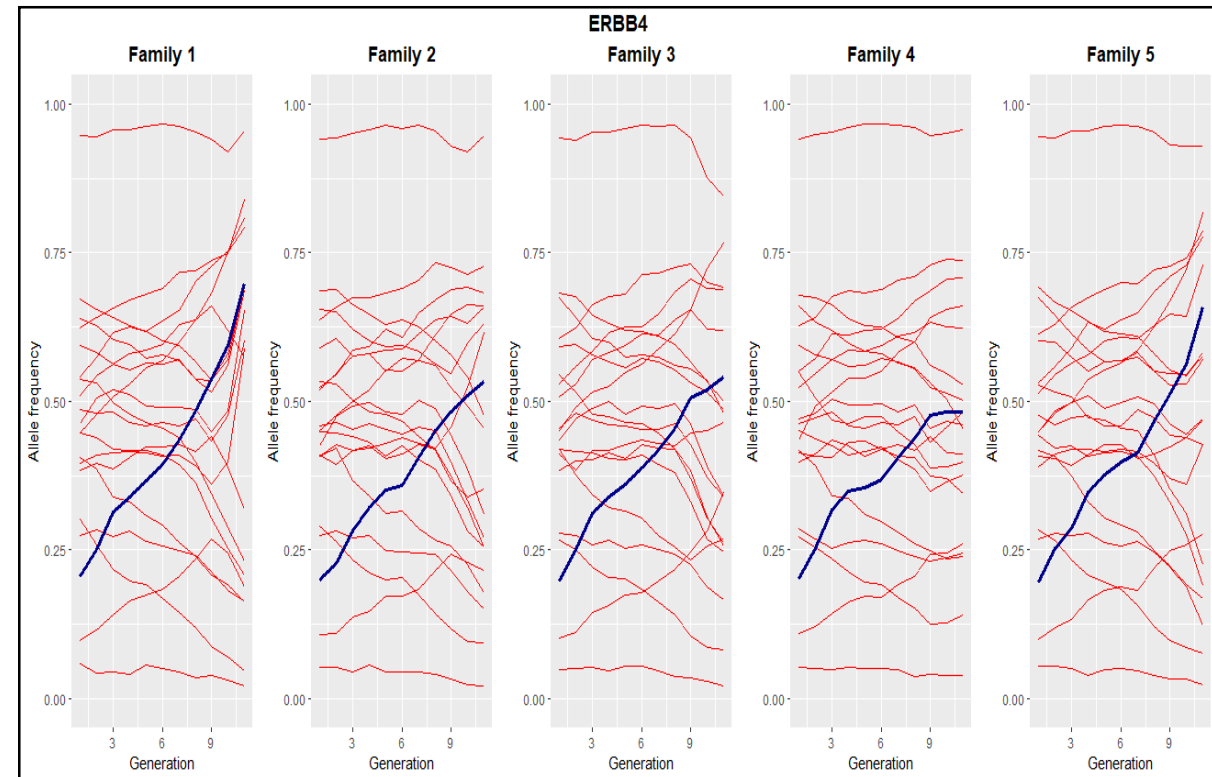
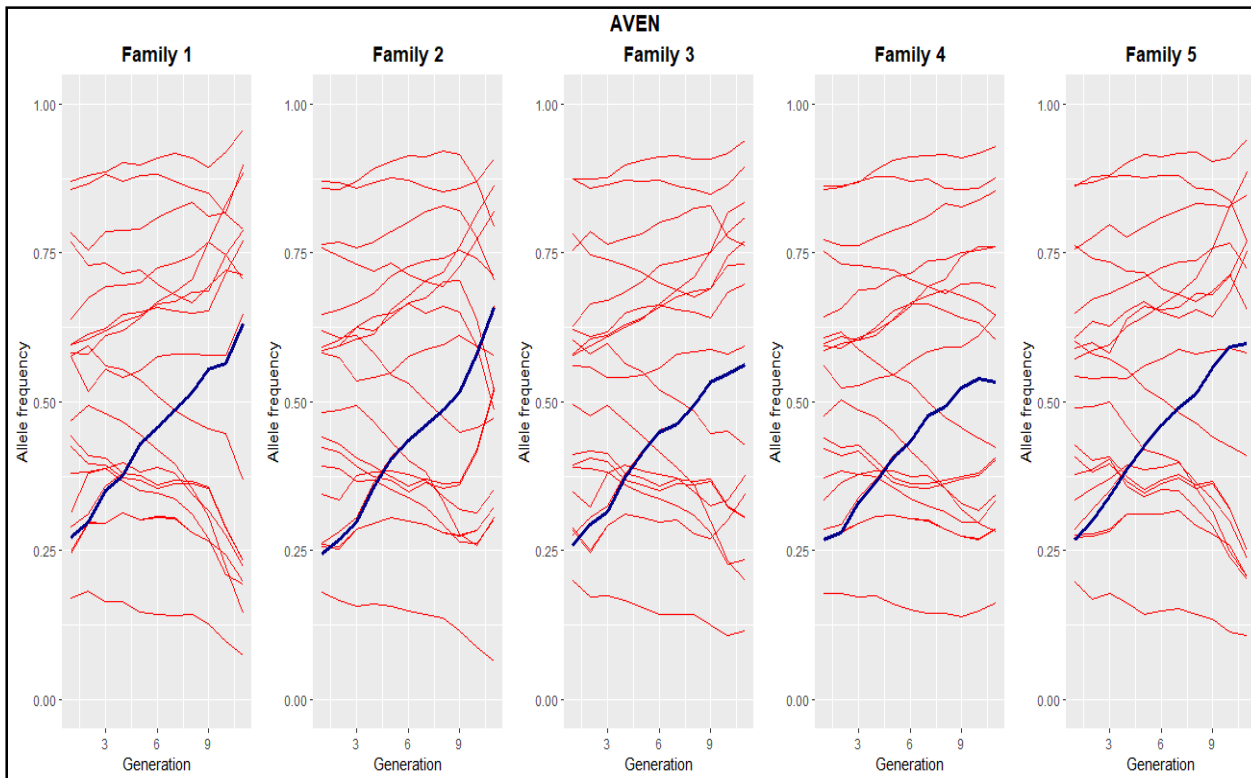


Graph of allele frequency change over 10 generations

# Some directional SNP changes are common to all families

**Blues lines** are the SNPs closest to the Fertility genes: AVEN and ERBB4

**Red lines** are the 20 closest SNPs surrounding the marker of interest



# Many allele frequency changes differ across families

## Replicate Frequency Spectrum (RFS)

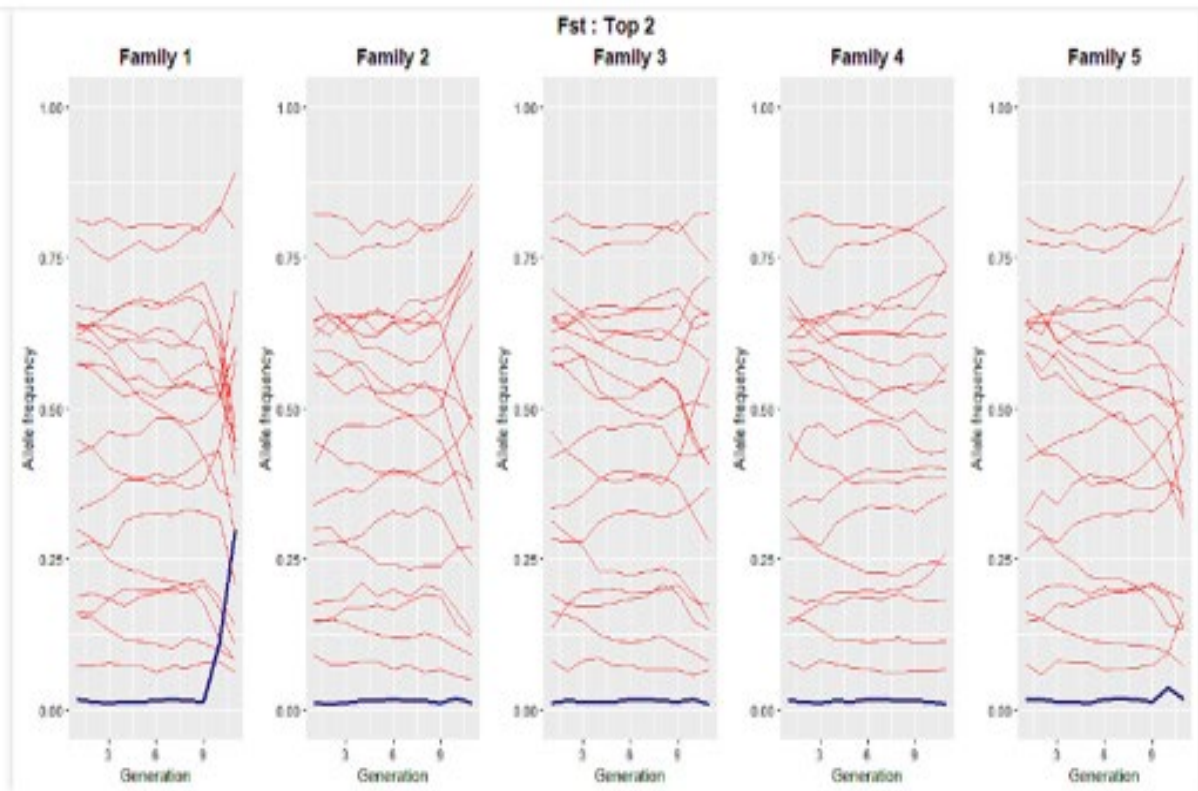
	Number of Alleles changing the same			
	Planet	Goldwyn	Shottle	O Man
Planet	100	69	82	73
Goldwyn	61	100	78	57
Shottle	64	71	100	52
O Man	65	78	82	98

Allele frequency changed by at least 30%

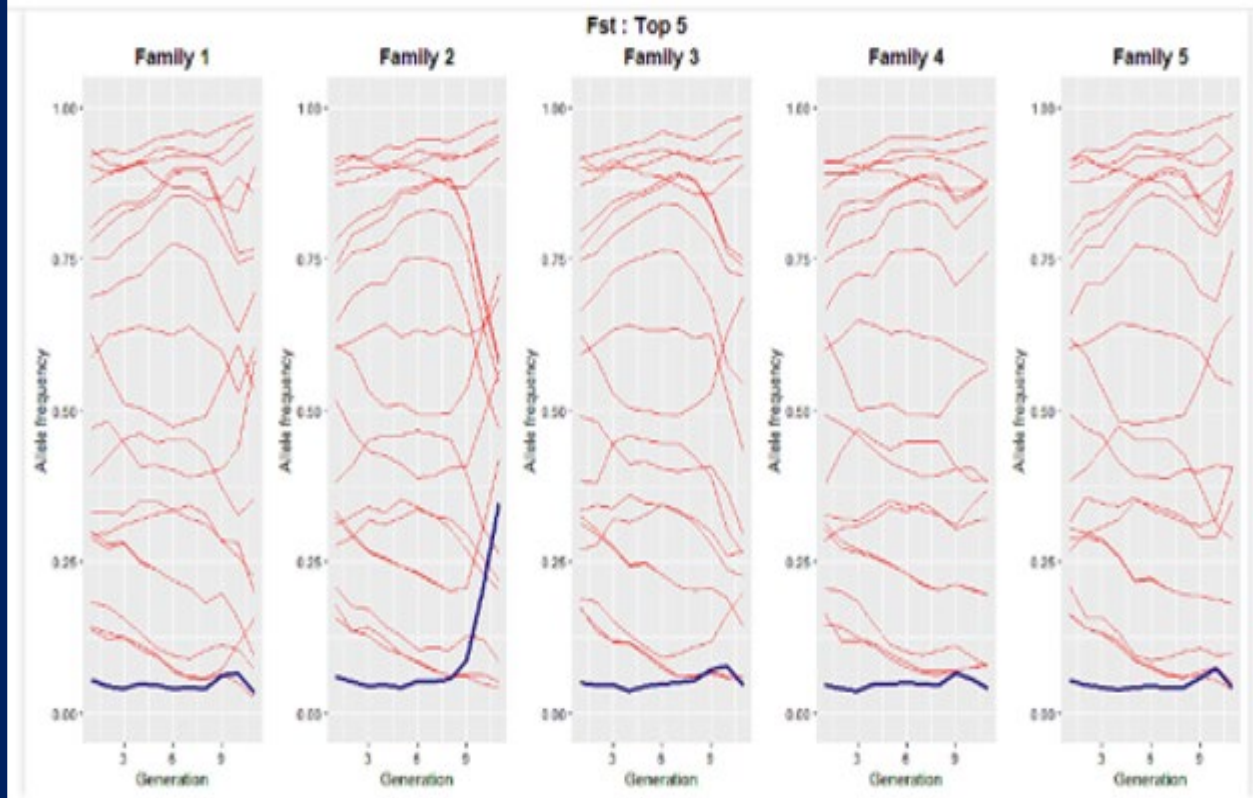
Between 18 to 48 of the Top 100 within-family SNPs were not shared across-families.

# New mutations or rare alleles in different families

New Mutation in Family 1



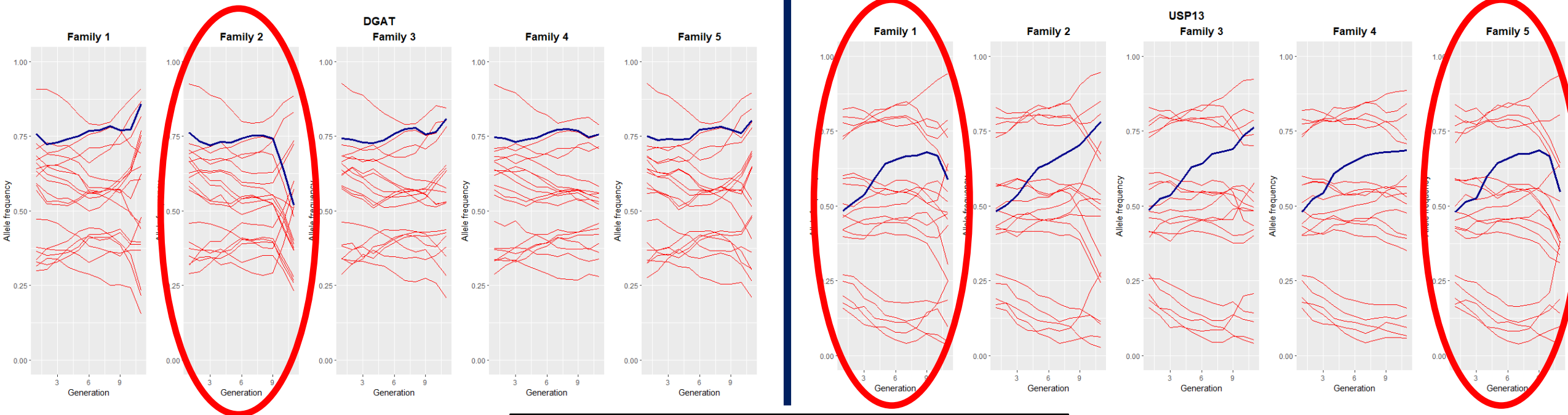
New Mutation in Family 2



# Different changes in different families for SNPs located within key genes

Fat and Protein: DGAT

Immunity: USP13

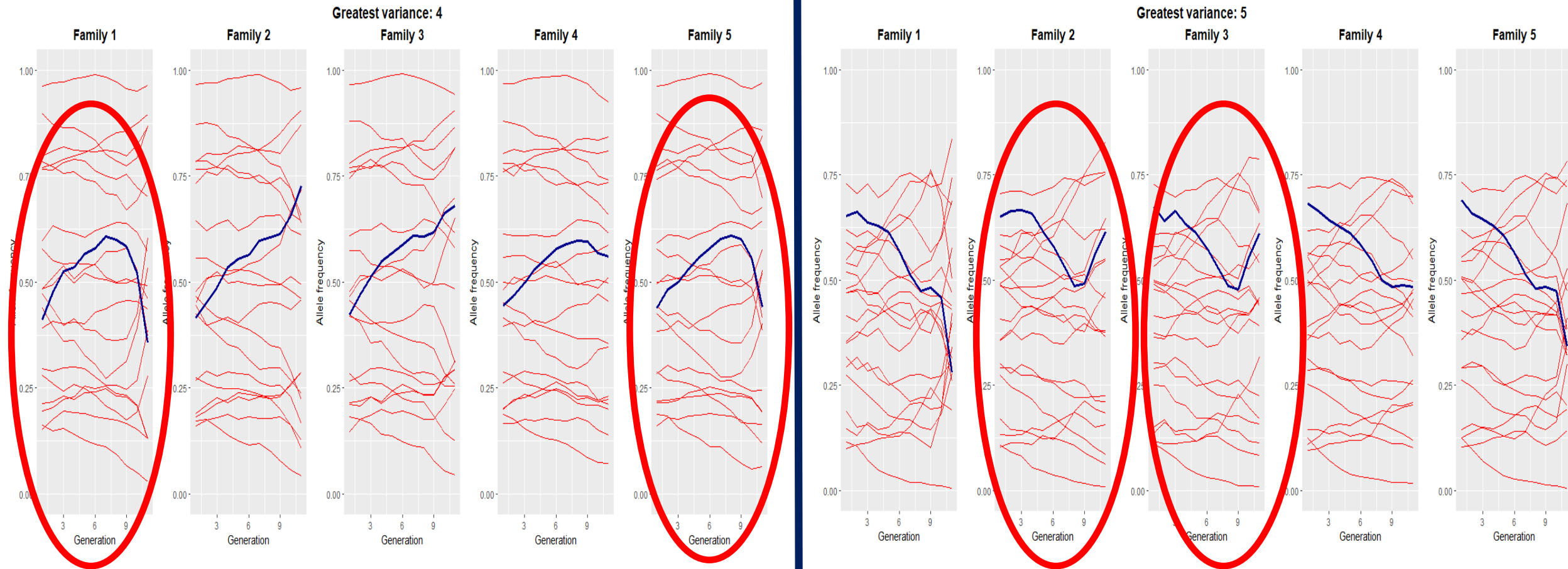


Reasons

1. Change in selection goals.
2. Breakage of linkage.
3. Epistasis.

# Reversal of direction is a genomic signature of epistasis

With epistasis...value of allele is contingent on the genetic background





Having different subpopulations changing the frequency of different alleles results in fewer alleles becoming fixed in the global population.

The number of SNPs that became fixed **within each family**  
**38, 22, 22, 40, and 59**

Number of SNPs across the **whole population** that went to a frequency of 100%

**3 SNP markers** out of **58,990**

The number of SNPs that **reversed direction**

	Number	Percentage %
Family 1	<b>6765</b>	11
Family 2	<b>5986</b>	10
Family 3	<b>6238</b>	11
Family 4	<b>6285</b>	11
Family 5	<b>2172</b>	10% change.

# Consequences of having a different genetic architecture in different subpopulations

- Genetic correlations between subpopulations are less than 1.0.
- Individual bulls rank differently in the different subpopulations.



# Using the procedure of Duenk et al. 2020

This is illustrated for **FAMILY 2** for the trait **STATURE**

**Genetic correlation** of additive genetic values using SNP effects calculated from females in the five different subpopulations

**Family 2 - GOLDWYN**

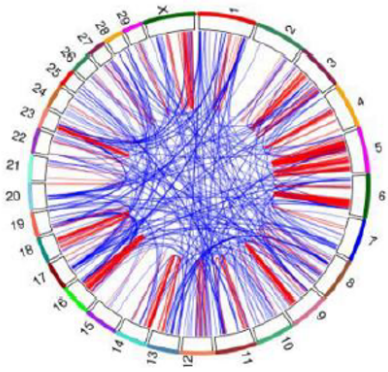
Own Family	Other Families			
1.00	0.89	0.56	0.59	0.87

Example bulls showing **rank** in different subpopulations

Goldwyn's		Combined Population	Own Family	Other Families			
grandson	Airlift	1	3	101	64	276	44
son	G.W. Atwood	22	8	81	150	446	212

## Cis and Trans epistatic interactions

Protein yield in U.S. Holsteins.



Each number on the outer circle is the chromosome number, red color represents intra-chromosome epistasis effects, and blue color represents inter-chromosome epistasis effects.

GWAS using 294,079 first-lactation U.S. Holstein cows.  
Source: [Prakapenka et al. 2021, Genes 12071089](#)

# Going forward

- Do not combine all Holstein SNP effects together.
- Genetic models – including additive-by-additive effects lowers allele fixation (Wientjes et al., 2023).
- **Breeding for niche markets creates subpopulations.**
- **AI breeding programs are becoming more isolated with limited exchange.**
- Complex genetic architectures favors establishment of separate genetic lines (Technow, 2021).

Bulls	$F_{ST}$
Select	<b>0.020</b>
SexTech	<b>0.032</b>
Peak	<b>0.042</b>
ABS	<b>0.063</b>
Germany	<b>0.018</b>

# Any Questions



Use of genotypes from CDCB is appreciated.