

After 15 years of genomic selection, we now know a lot more molecular biology.

**Heritable variation in gene expression  
is the key contributor of phenotypic change  
and genetic diversity in Holstein cattle**



Dr. Tom Lawlor  
Holstein Association USA

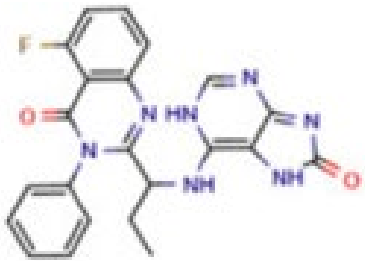


# Integrating molecular biology with Fisher's infinitesimal model.

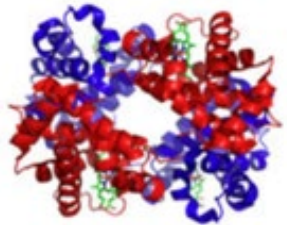
Phenotype



Metabolites



Protein



RNA



DNA



**SNP effects**



General population

94%

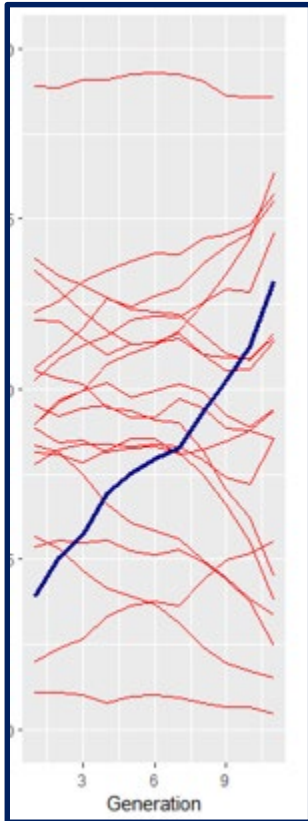


Single nucleotide polymorphism (SNP)

6%

# Genome changes in Holsteins is very dynamic

Rapid changes in allele frequency



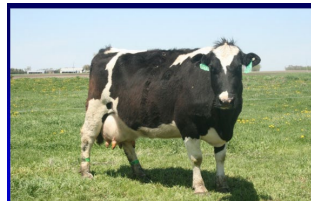
Subpopulations constantly forming, mixing and reforming.

Elite Cows – born 2015

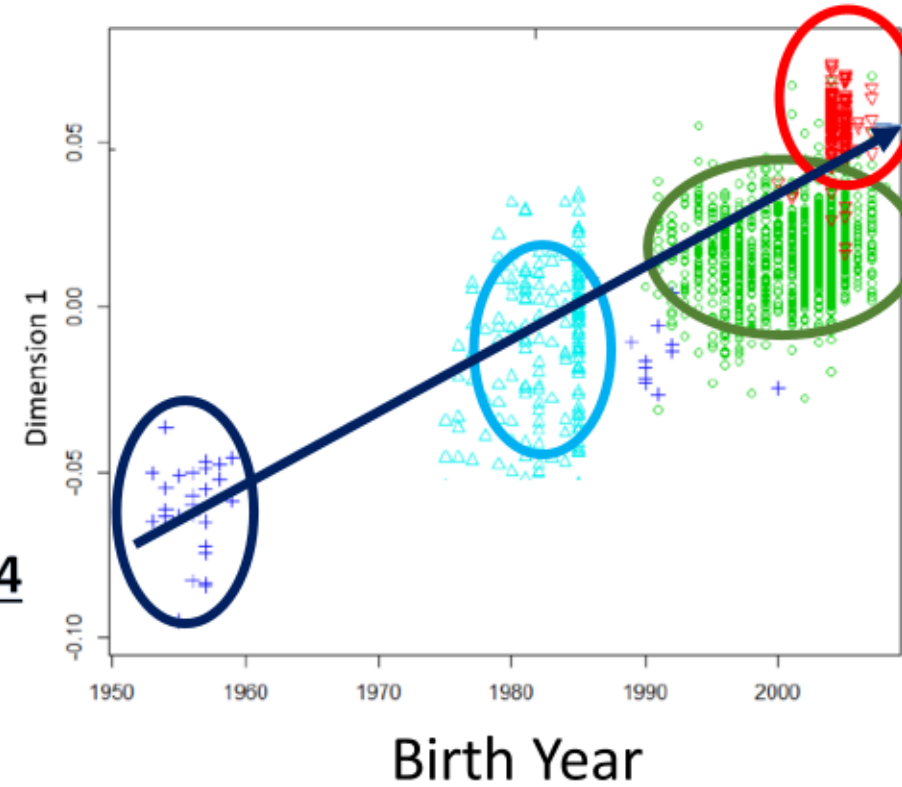
Contemporary Cows  
born around 2000

Bulls born around 1980

Cows with genetics from 1964



1964 Holstein Genetics

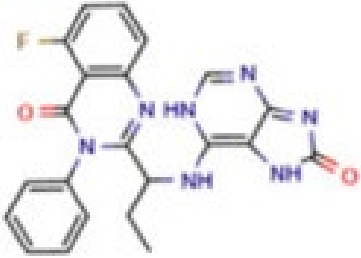


Phenotype



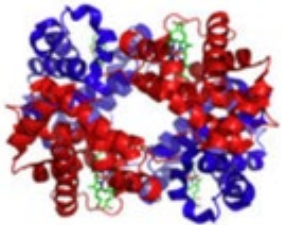
# What's new in molecular biology

Metabolites



**Gene Regulatory Networks**  
Coordinated response with redundancy

Protein



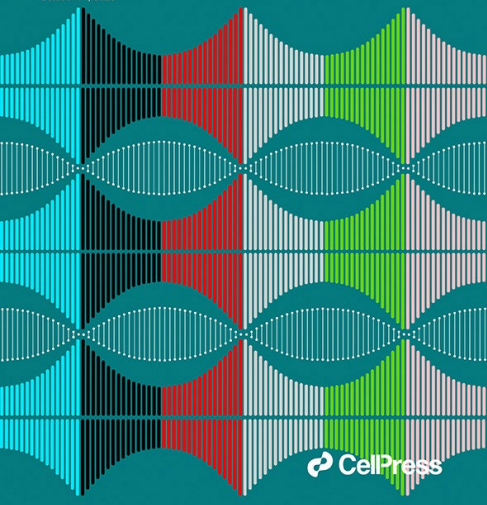
**Gene expression**  
Major driver of genetic change

RNA



DNA



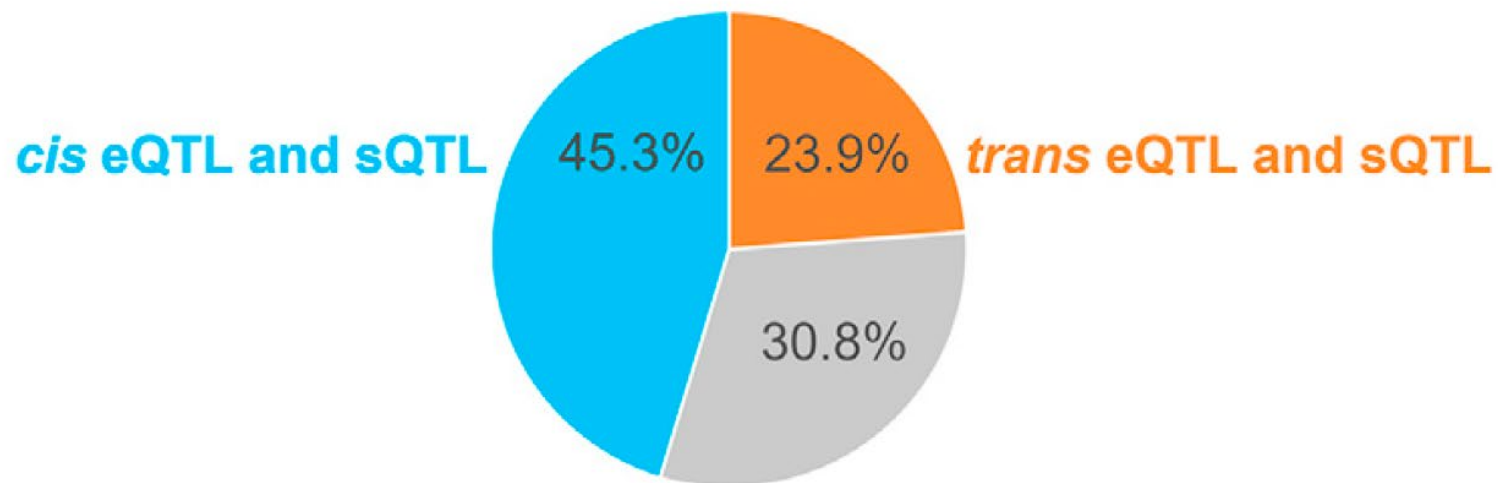


# Gene expression and RNA splicing explain large proportions of the heritability for complex traits in cattle

Published: October 11, 2023

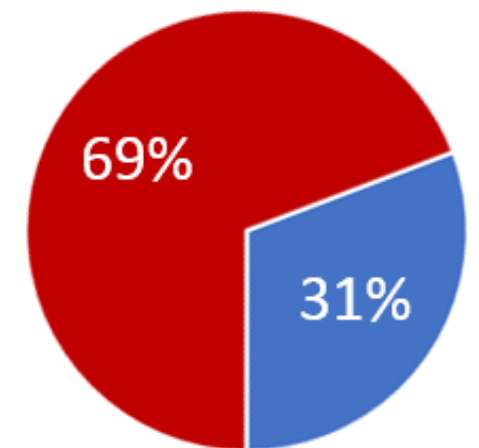
<https://doi.org/10.1016/j.xgen.2023.10038>

Averaged proportion of heritability explained by regulatory variants across 37 traits



Variants contributing to heritability

**regulatory control**  
**NO regulatory control**



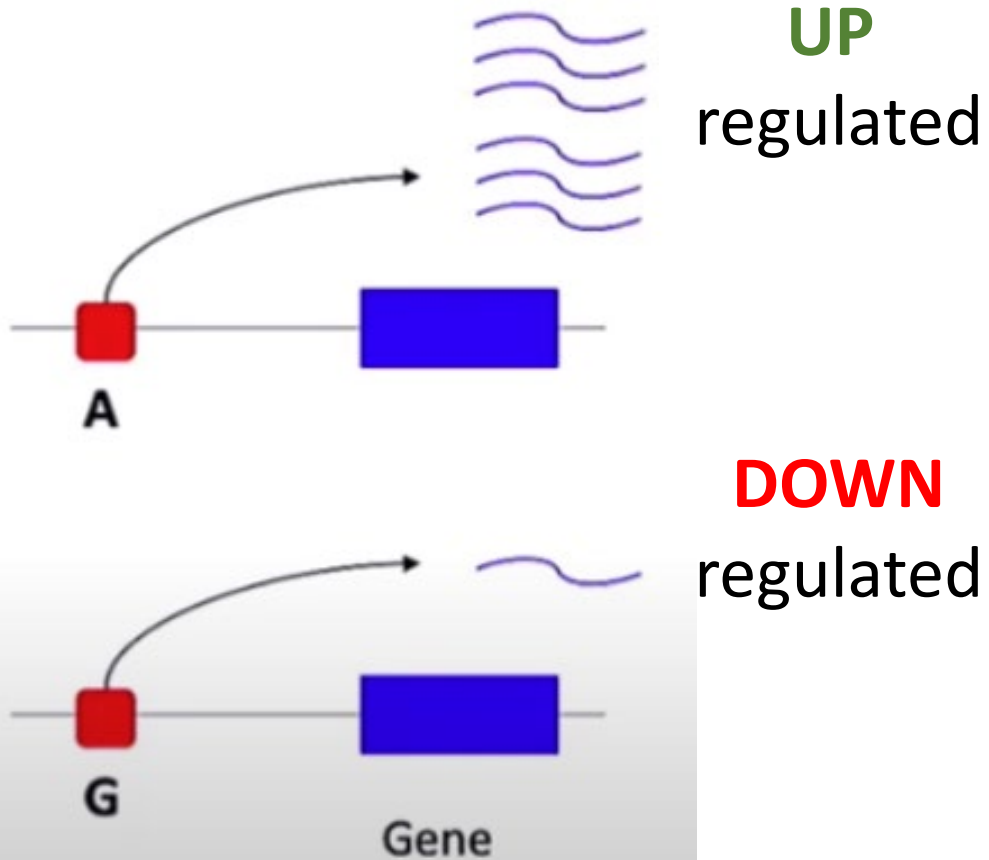
Cell Genomics, October 2023

# Gene expression

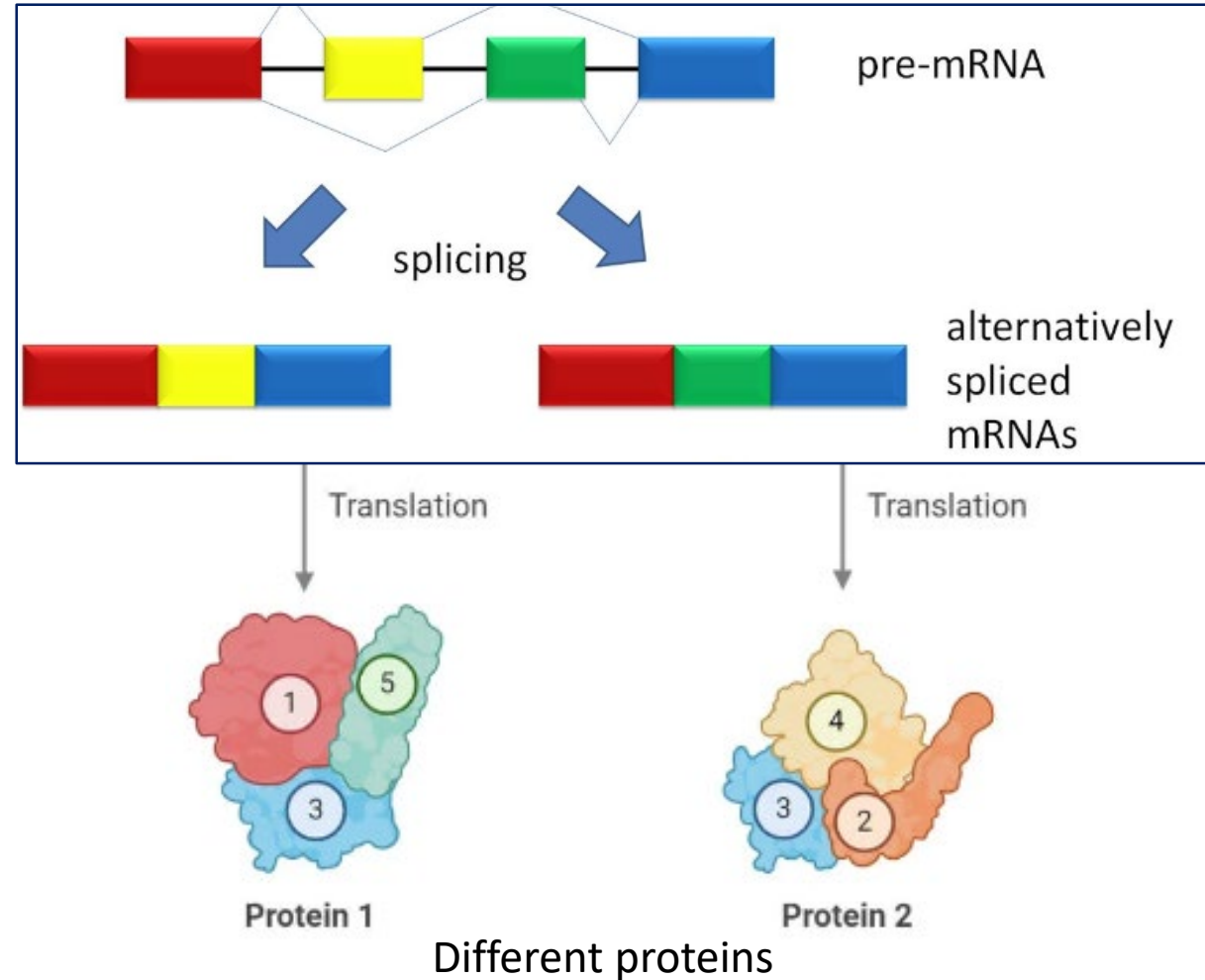
*Regulation* – when, where and how much of a protein

*Exon splicing* – what type of protein is produced

## Regulation



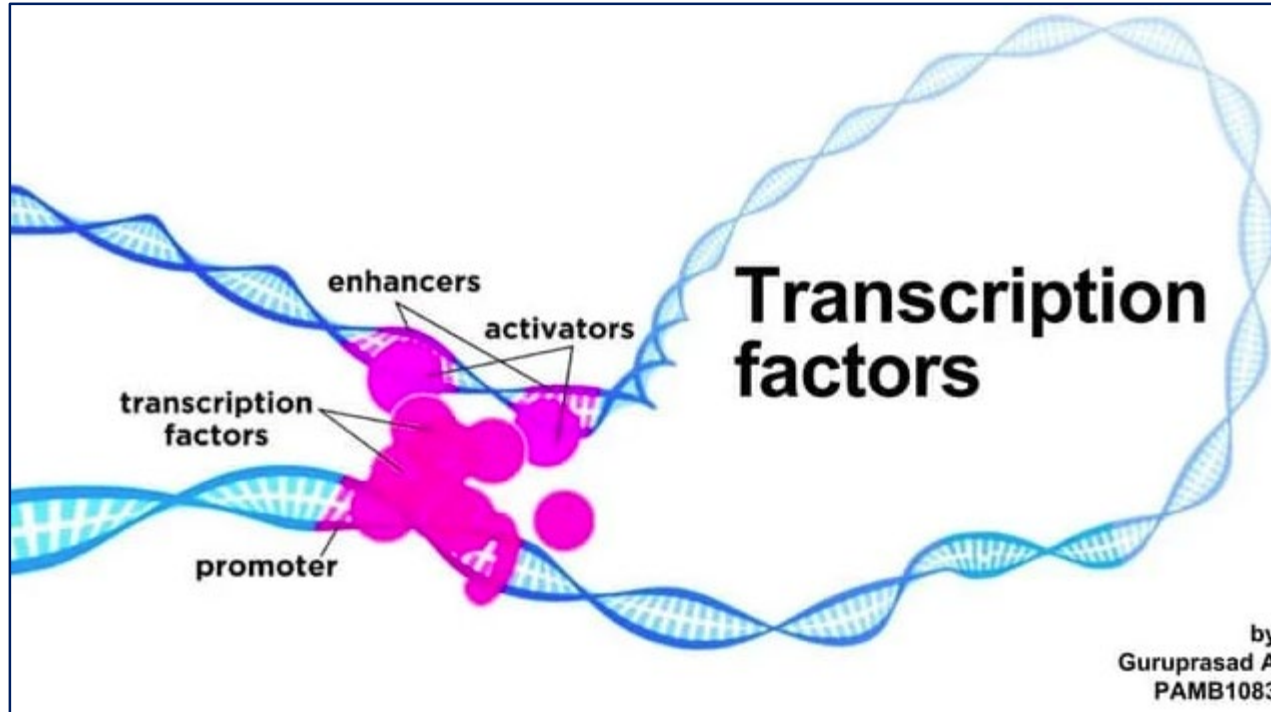
## Exon Splicing



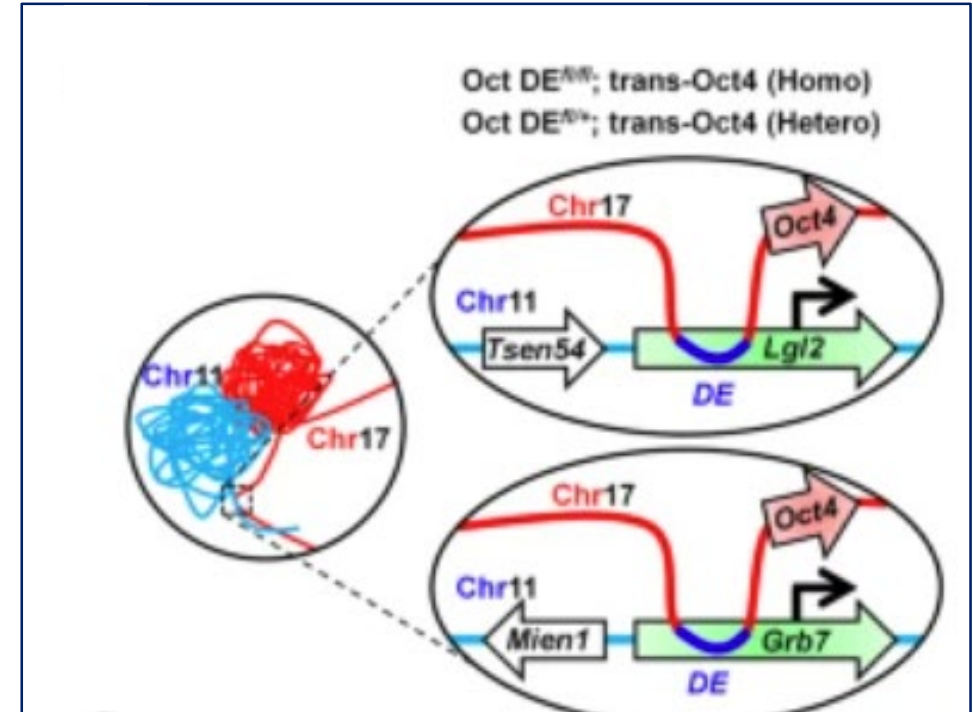
# Large number of possible interactions

*Cis* – nearby

*Trans* – long distance



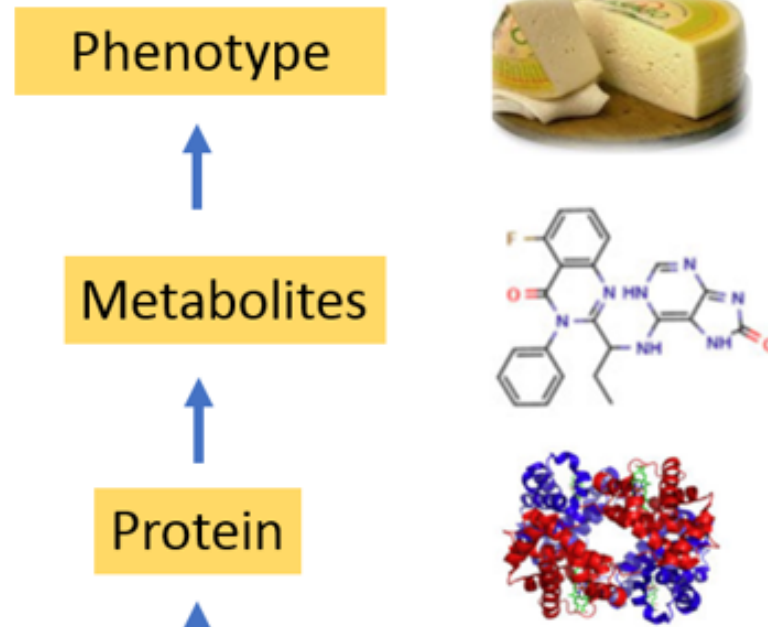
These loops allow interactions between different regions of DNA



Long range inter-chromosomal interactions

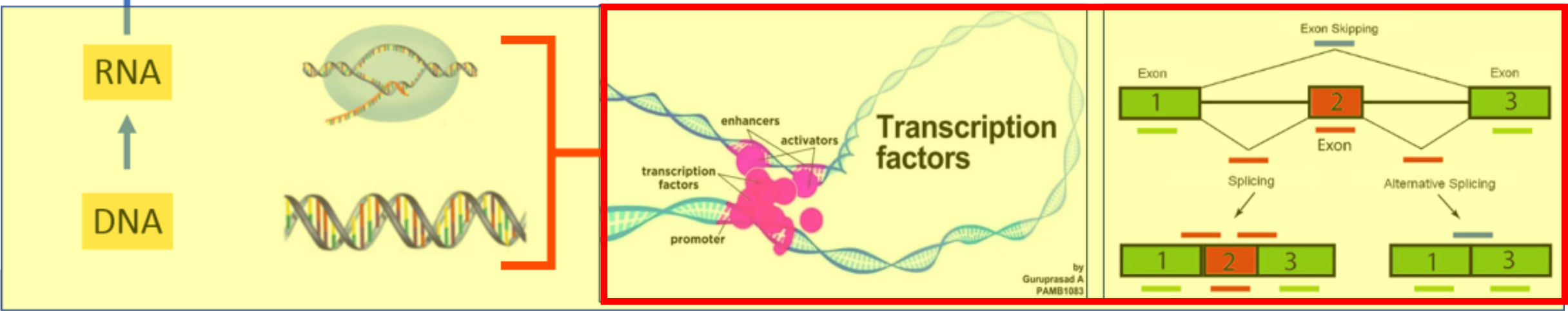
# Genetic Redundancy:

only a subset of all these interactions is needed



Hundreds of millions of combinations

Exon splicing provides more variation



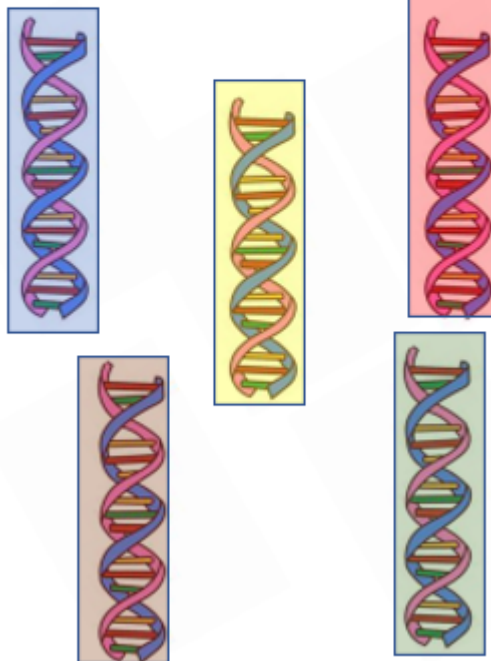
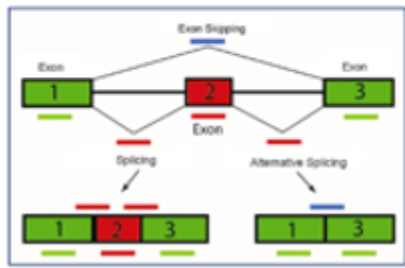
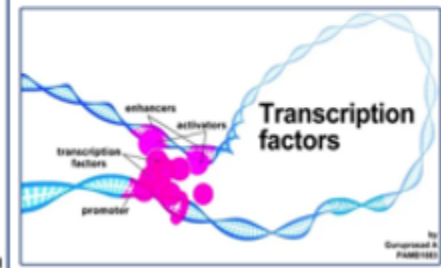


Hundreds of millions of combinations

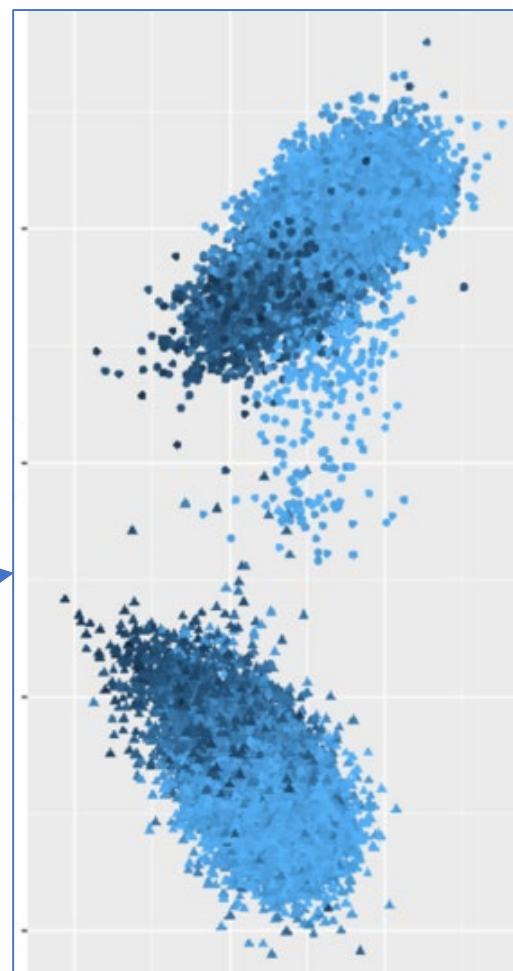
Exon splicing provides more variation

Wide variety of genetic material to select from

Population differentiation



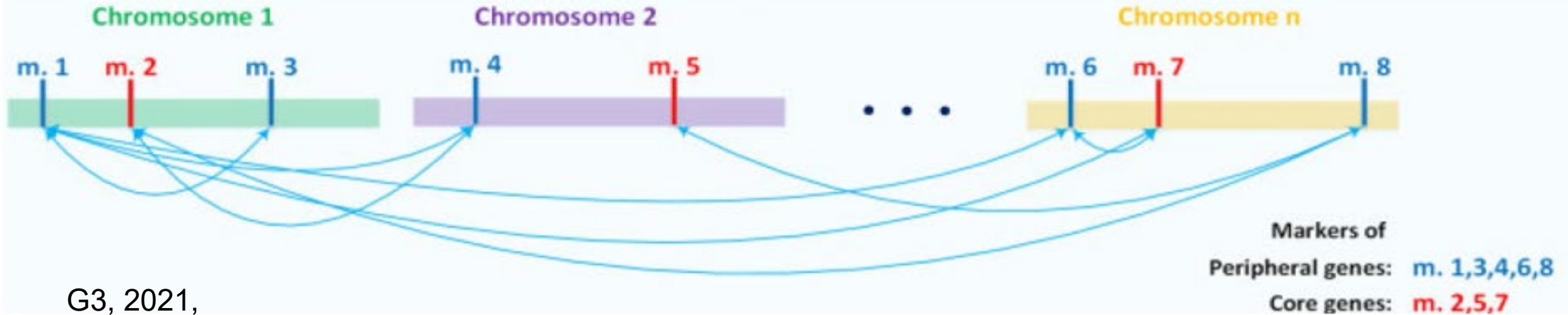
Parents of the next generation will have a unique genotype



# Depending on animal's ancestry



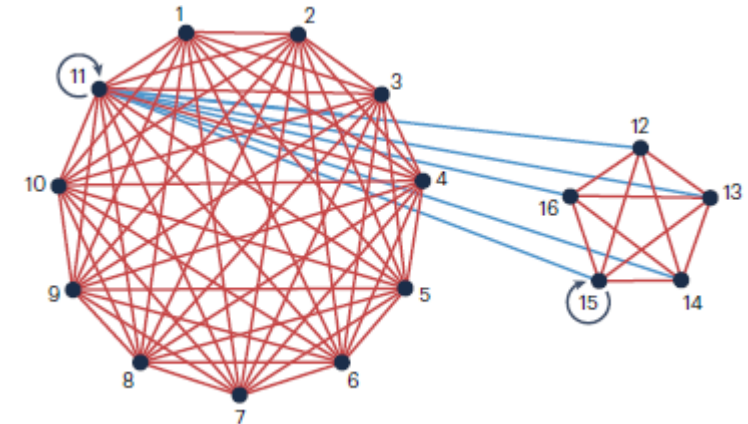
Different subpopulations will have different genetic architecture



G3, 2021,  
11(7),

DOI: 10.1093/g3journal/jkab133

Different transcripts leads to different  
**Gene Regulatory Networks**



# The Genome Response to Artificial Selection: A Case Study in Dairy Cattle



Holstein



Montbéliarde



Normande

These three breeds are genetically different from each other.

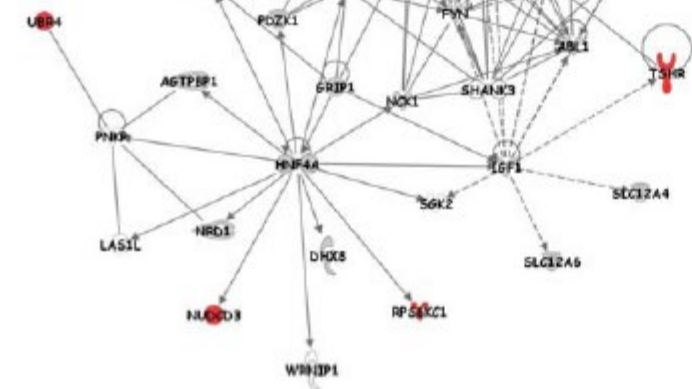
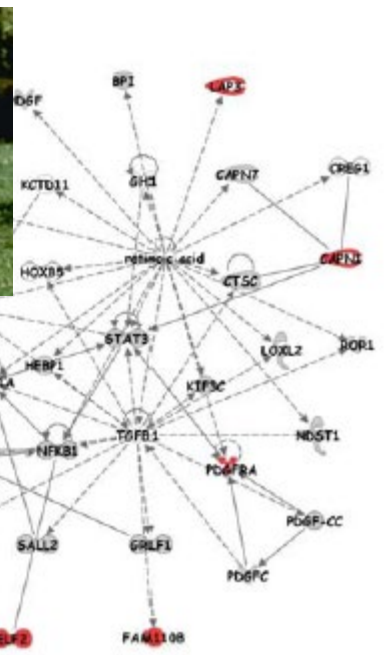
$F_{st}$  measures allele frequency differences between populations.

$F_{st} = 0$  same breed

$F_{st} = 0.07$  different dairy breeds

$F_{st} = 0.15$  dairy breed compared to a beef breed

Each of the breeds has found its own pathway to a genetic solution



Wired differently  
Connections between genes are different in the different breeds

“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

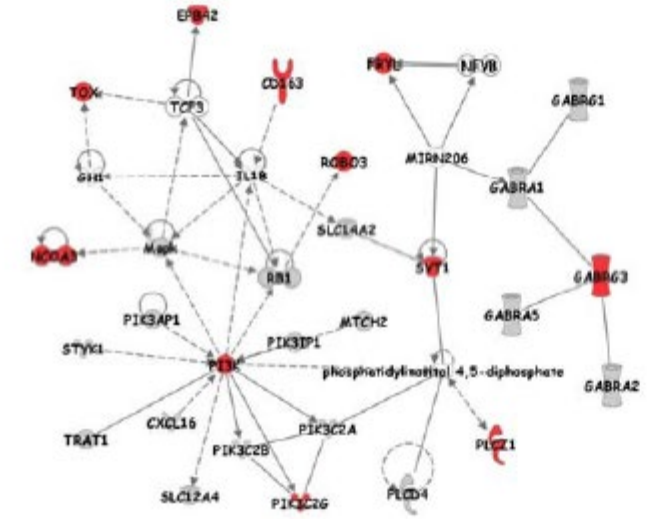
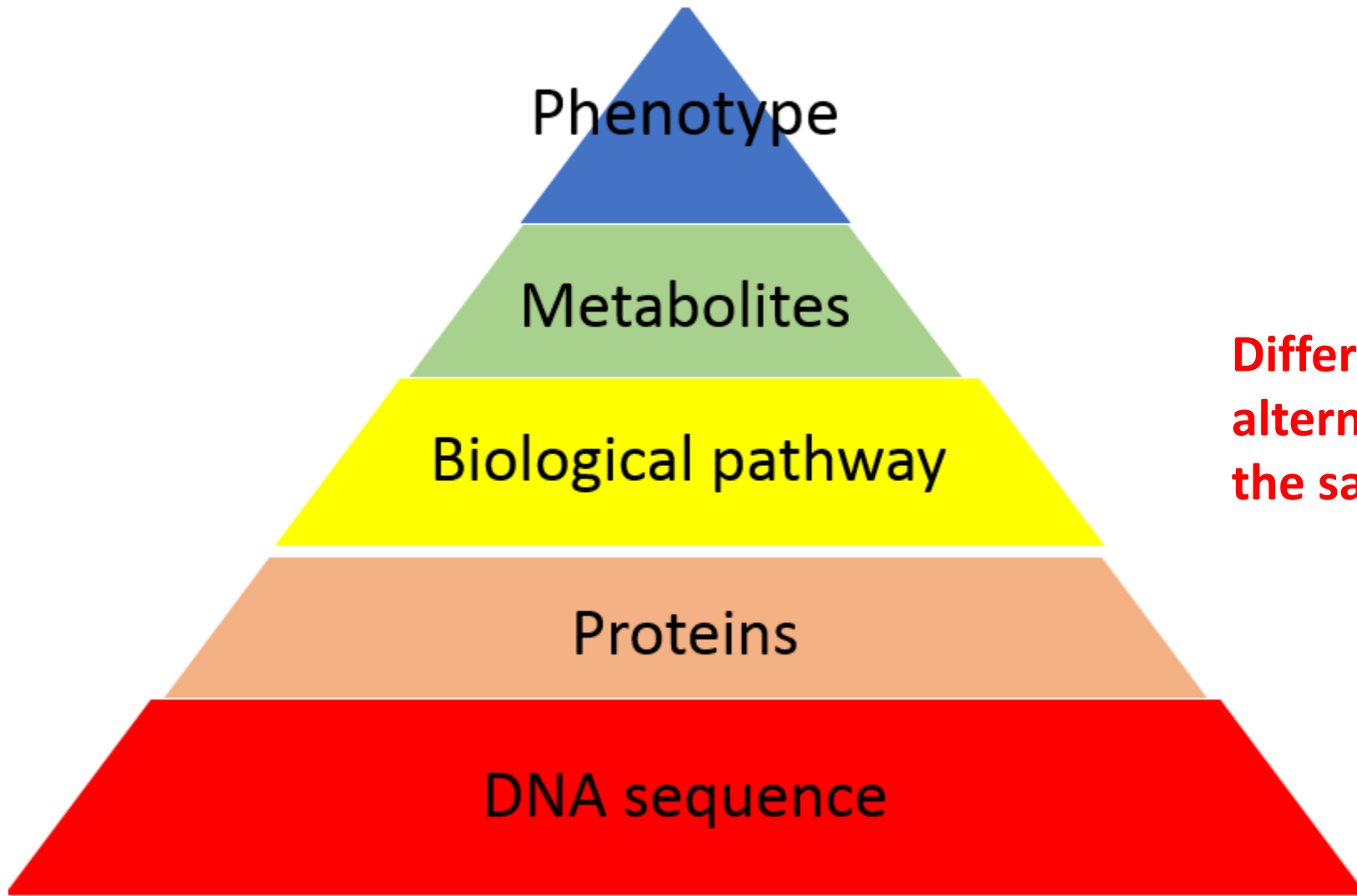


Figure 5. Representation of the gene networks N\_MON (A), N\_NOR (B) and N\_HOL (C). Symbols corresponding to candidate genes are colored in red. Genes colored in grey were represented in our study but did not display any evidence of selection.

# The genome-to-phenotype map has a hierarchical organization



**Different subpopulations can use alternative pathways to achieve the same phenotypic response**

**Hundreds of million potential combinations**

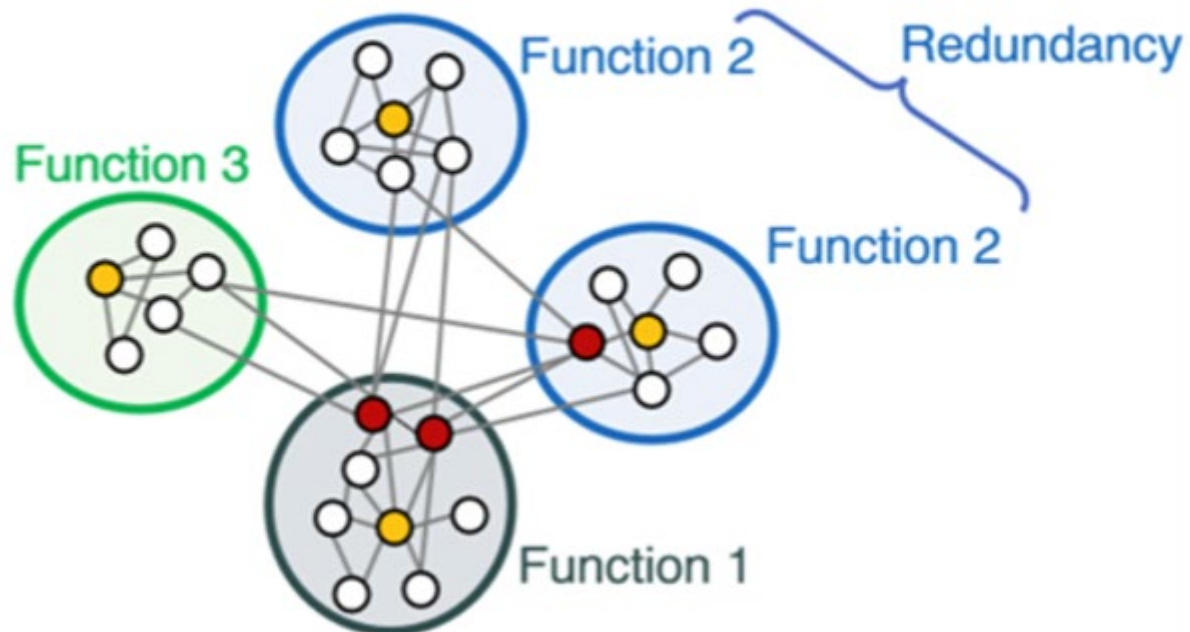
# Evolutionary advantages to having redundancy and hierarchical organization

Improved or novel functions can evolve gradually, rather than in a single step.

**Modularity:** components may be separated and recombined, often with the benefit of flexibility and variety in use.



Gene Regulatory Network model



## Coordinated response with redundancy

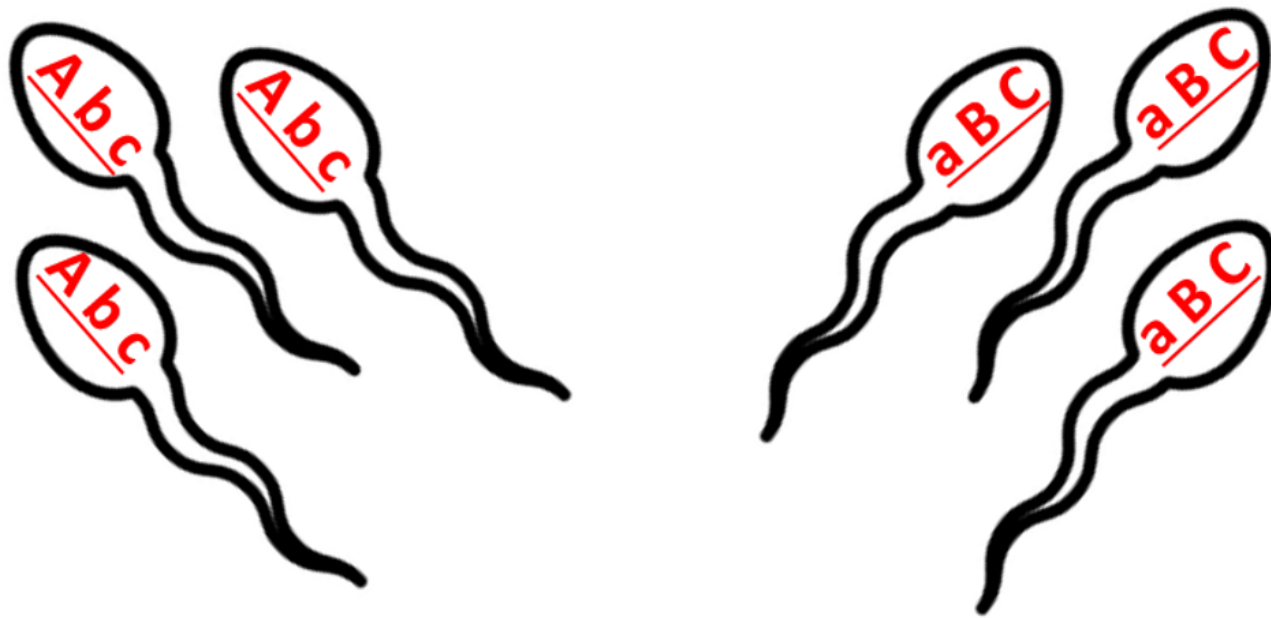
**similar** - compensates for mistakes in other pathways

**improved** – better use of a similar pathway

**new** - modified pathway leading to new phenotype (evolution)

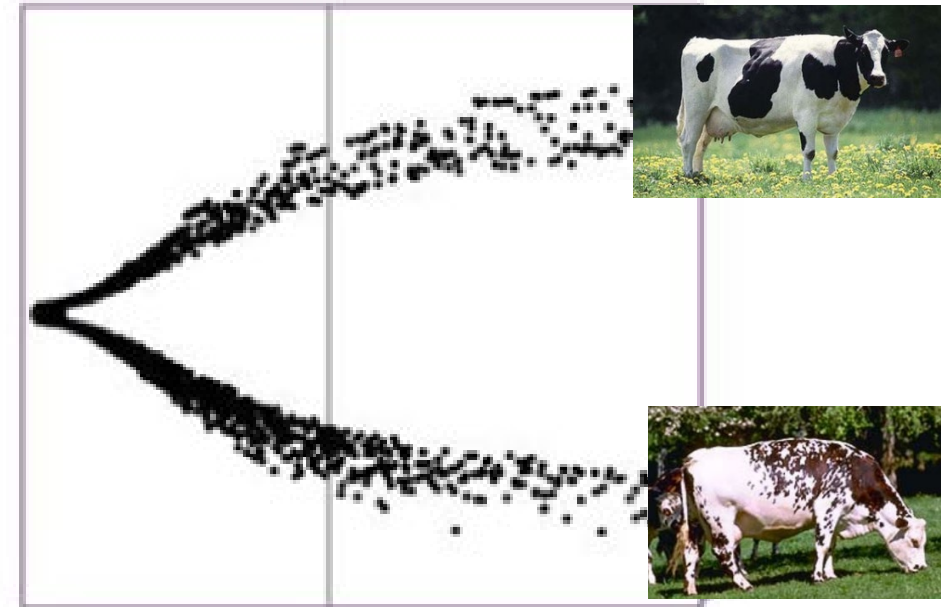
# Gametic disequilibrium

Different gametes with different combinations of alleles leads to different genetic solutions (pathways) being favored in different subpopulations.



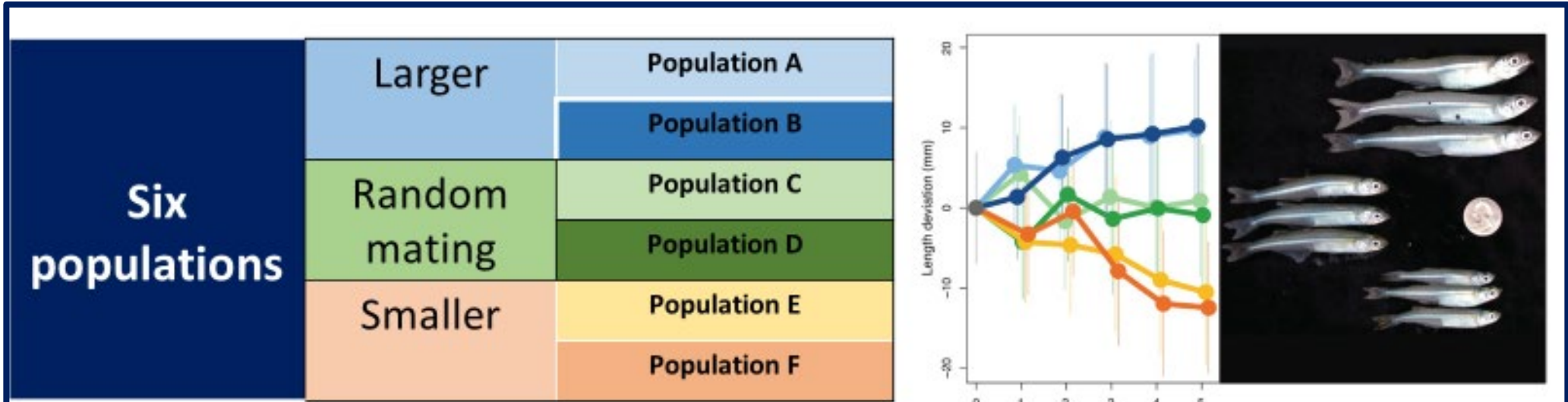
Gamete **A\_b\_c** is desirable in Family 1  
Gamete **a\_B\_C** is desirable in Family 2

$F_{st}$

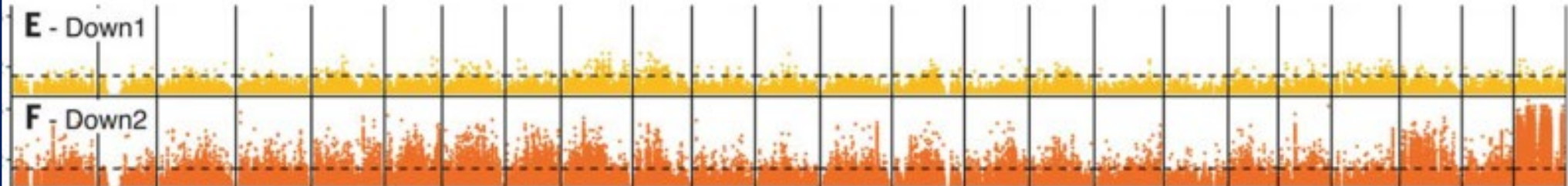


The two subpopulations become genetically different

# Rapid changes in genetic architecture are seen in nature: with selection and population division



*Science*, August 2019, 365: 487–490



The genetics of different subpopulations change differently  
even when they're selected for the same objective



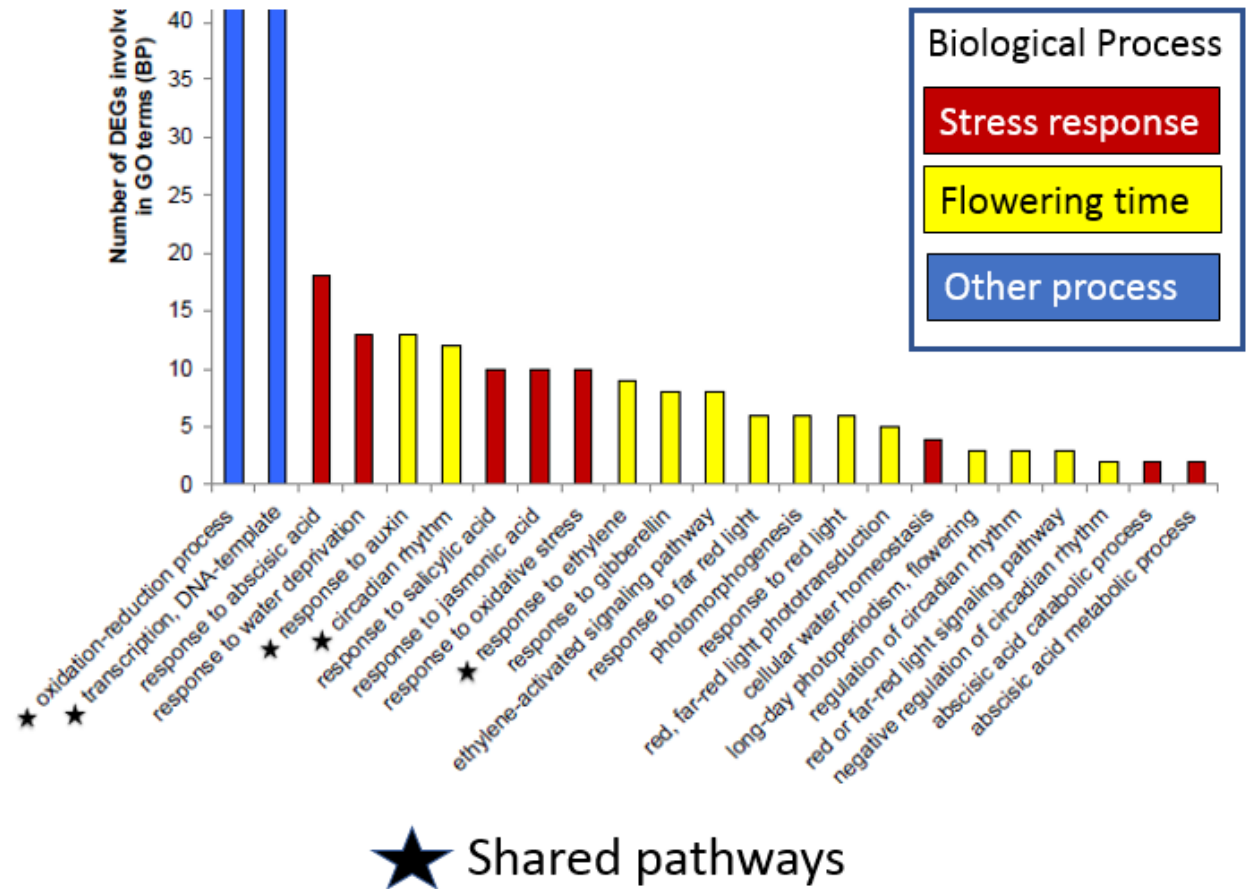
# Rapid evolutionary changes in gene expression in response to climate fluctuations

[Mol Ecol.](#) 2021 Jan; 30(1): 193–206



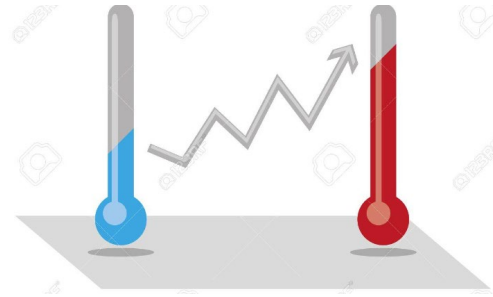
**“Biggest heritable change in response to drought was changes in gene expression”**

Different subpopulations took different pathways in response to the same drought conditions.



# Evolve and Resequence

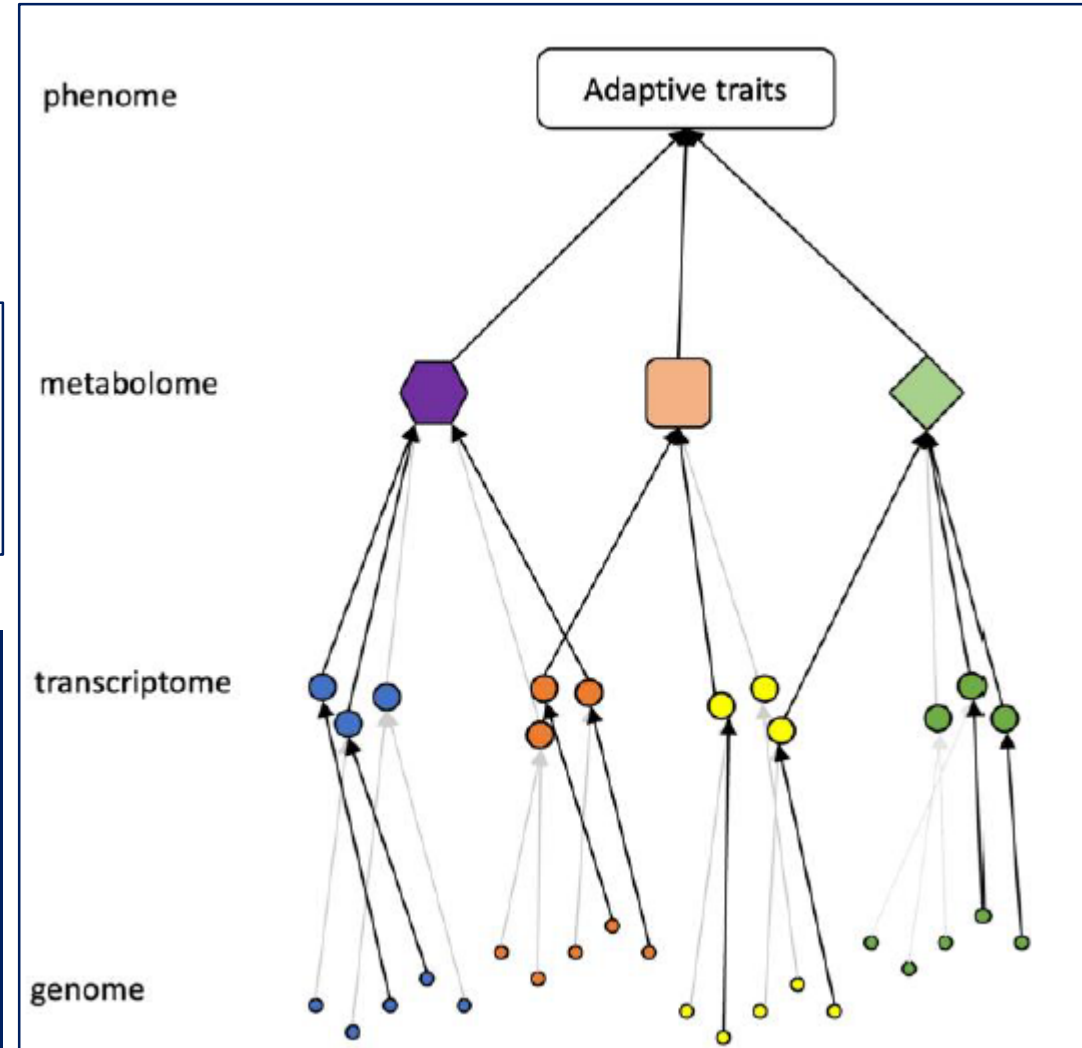
Barghi et al., 2019



A single population was divided into ten sub-populations. Then exposed to **HIGHER TEMPERATURES**.

Differences in the genetics of the original founders steers the subpopulation towards using a different set of SNPs to achieve the same phenotypic goal.

## Genetic redundancy fuels adaptation in *Drosophila*

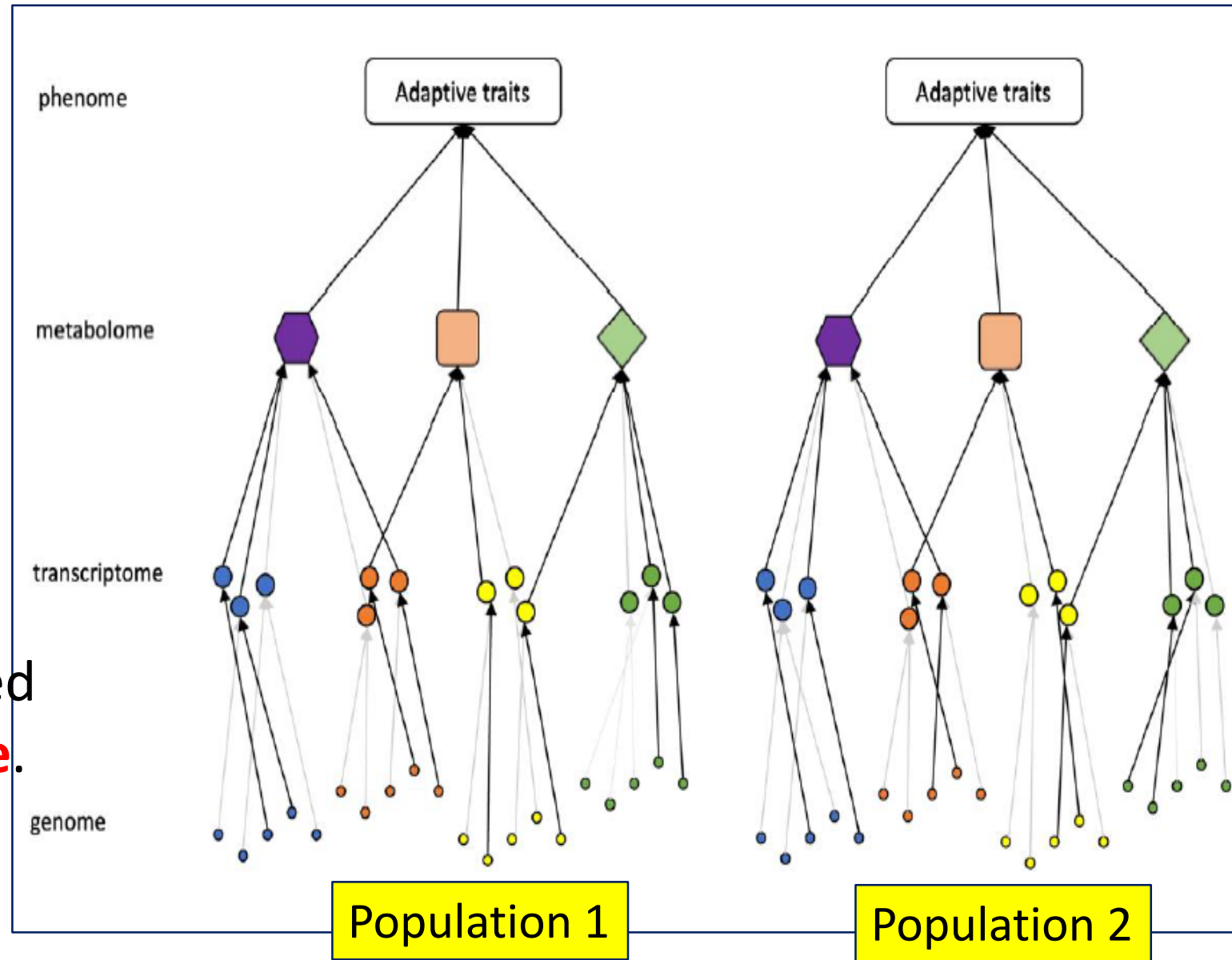


*Genome Biol. Evol.*

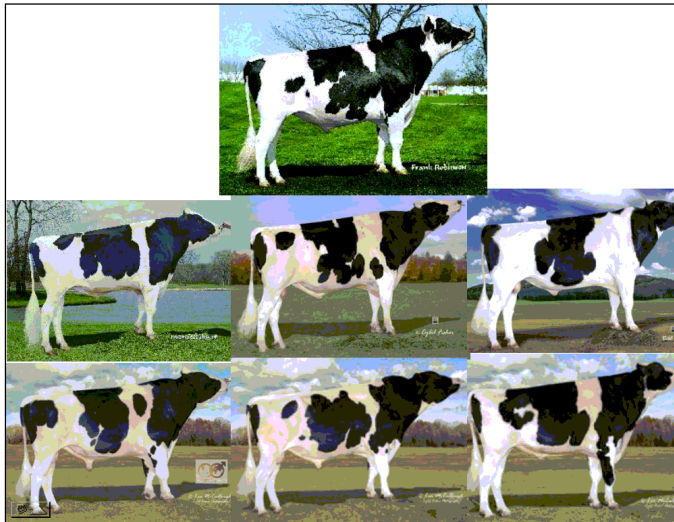
• The **Different combinations of SNPs** allows for different genetic networks to be used to adapt to the same phenotype.

**Different genotypes** leads to **different transcripts** leading to **different proteins** leading to **different interactions** leading to **different networks** leading to **different pathways** leading to **similar metabolites** being utilized to produce the **same phenotype**.

Lai et al, 2023



By understanding how genetic variation is created and organized, along with the proper **management of our population structure**, we can achieve rapid genetic progress and maintain genetic diversity.



# “SNP effects” or substitution effects

The average effect of **SNP i** in a population:

$$\alpha_i = a_i + (1 - 2p_i)d_i + \sum \alpha_{ij}^i$$

Main effects

Depends upon **frequency** of the SNP

$p_i$

**Gene action** (additive and dominance)

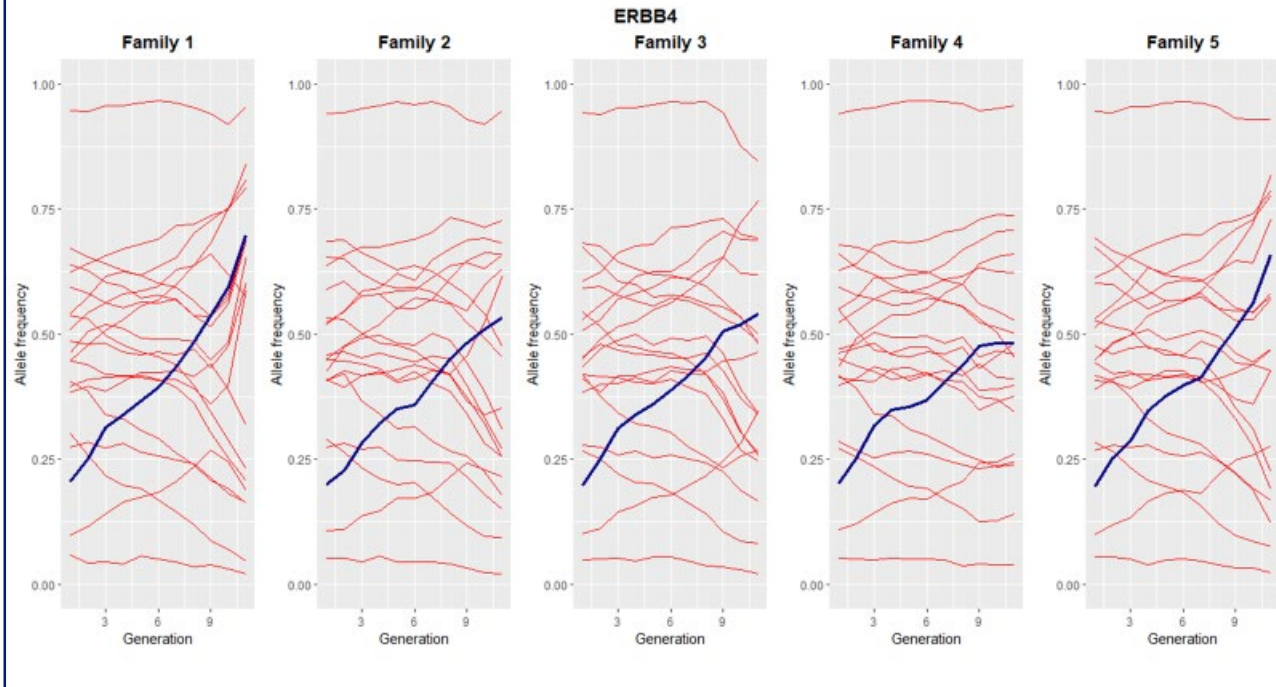
$a_i$  and  $d_i$

Interactions

**Interactions** with other SNPs

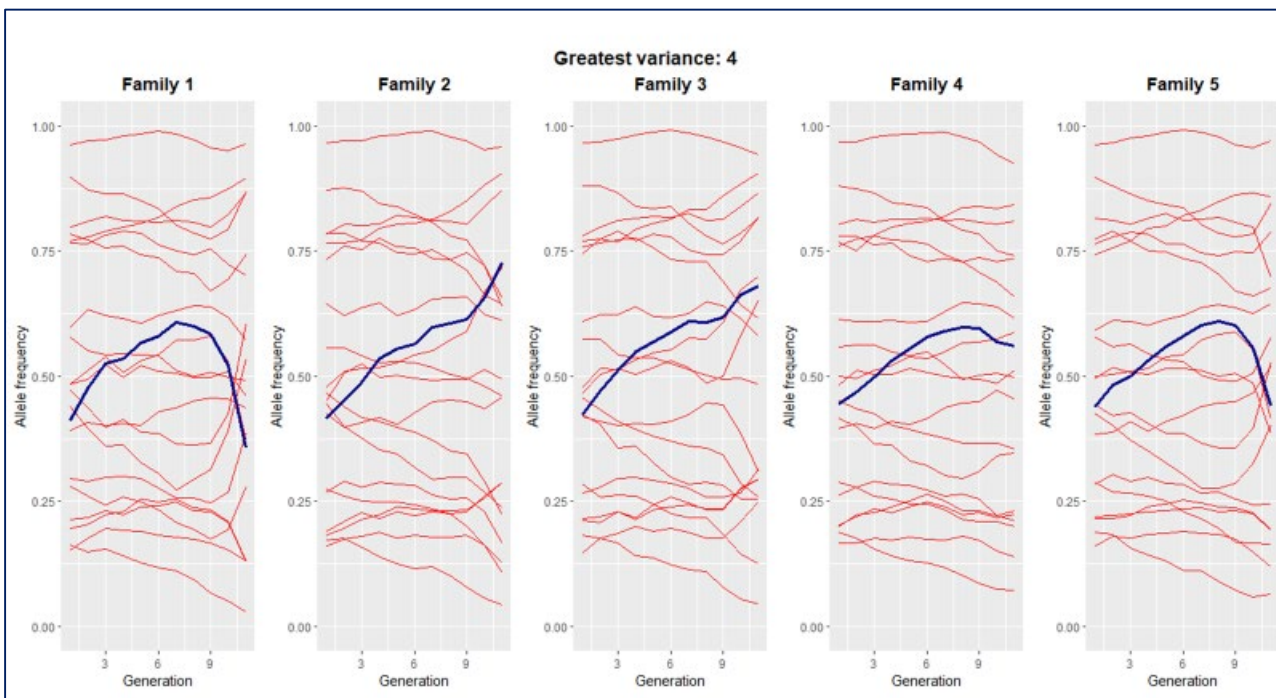
$\alpha_{ij}^i$

We don't ignore interactions, we averaged them out.



Assuming phenotypic value is related to change in SNP frequency

Without epistasis, consistency across families gives us a **high** SNP effect

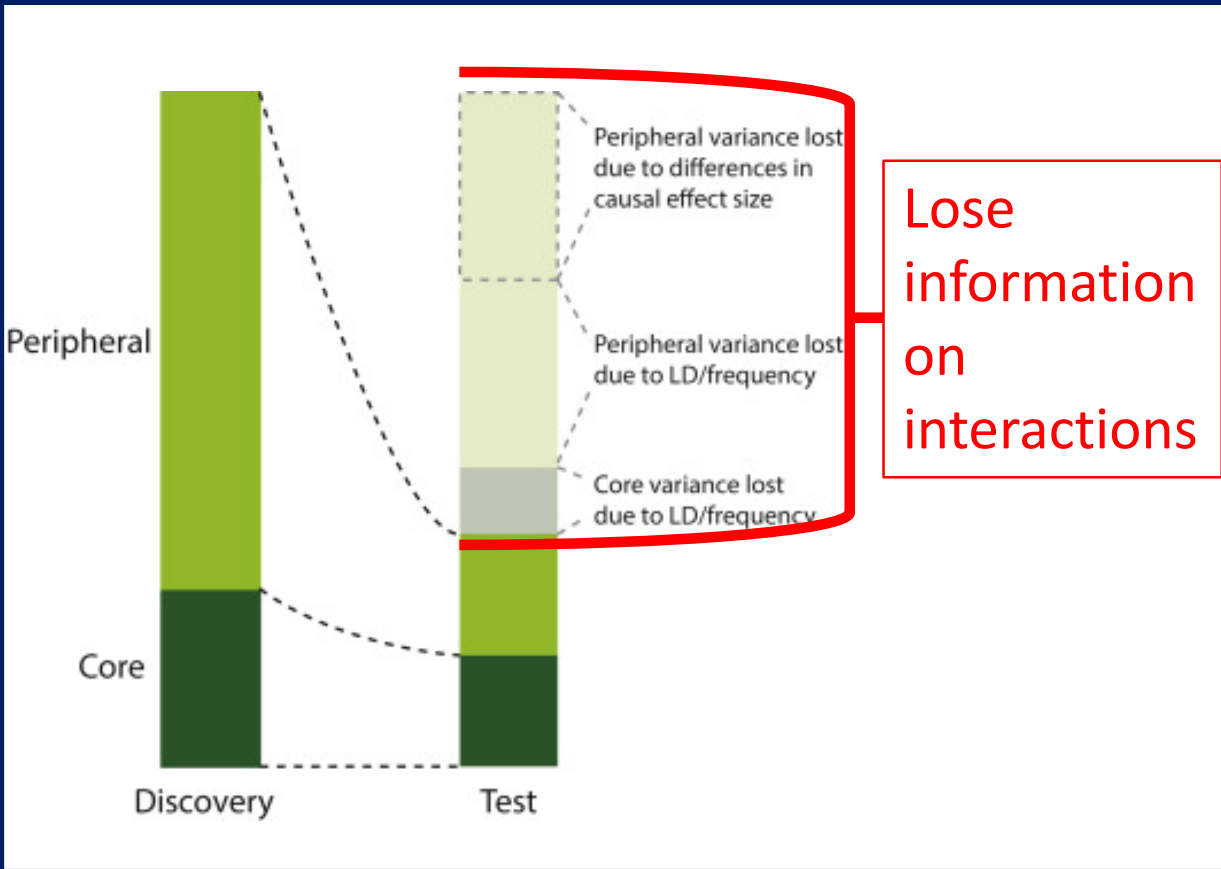


With epistasis, the inconsistency across families gives us a much **lower** SNP effect

Epistasis appears to be small because we're looking at the "average effect" across all subpopulations

# Analyzing Data

## Problem with pooling data from multiple subpopulations together



In a pooled data set.

We limit our selection to the SNPs associated with core genes or cis-regulation

### The conundrum:

More observations leads to a more accurate prediction.

By pooling, I get a more accurate prediction on a LIMITED set of SNPs.

Separate analysis, I get a less accurate prediction on a LARGER set of SNPs.

# Better use of the available genetic variation

Multiple breeding populations, each seeking its own path to obtain a similar goal.

## Wright's *Shifting Balance Theory*

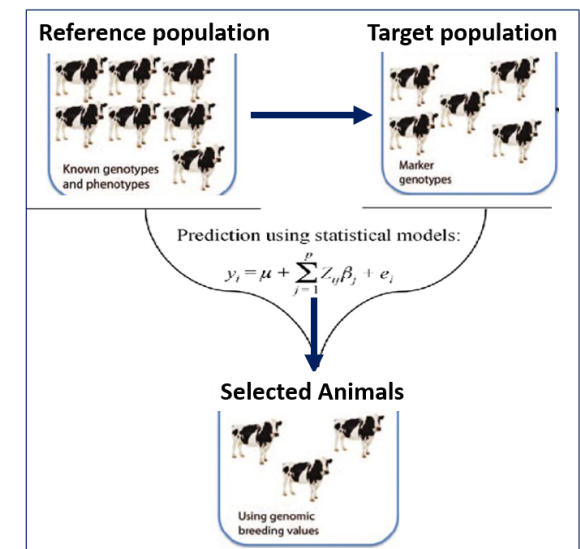
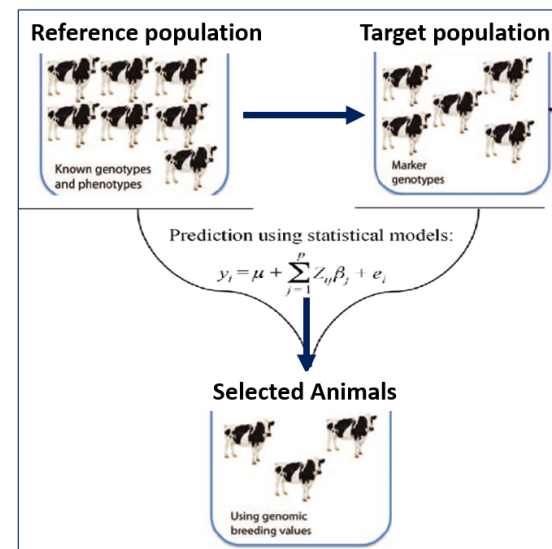
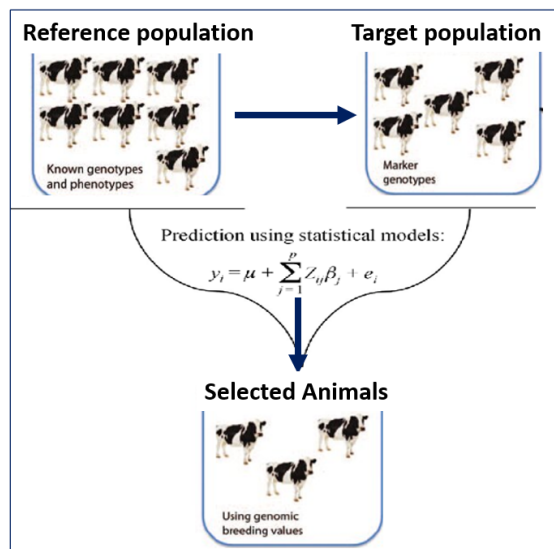
Create large family units.

Each with its own reference population.

Select best within-family animals.

Evaluate animals from all breeding organizations together on a national scale.

Evaluate all animals from all countries for the global population.





2022

Within-stud selection

U.S.  
Average

concentrating on their own families



$F_{st} = 0.04$

$F_{st} = 0.00$



$F_{st} = 0.02$



German  
bred

$F_{st} = 0.02$

$F_{st} = 0.03$



$F_{st} = 0.06$



$F_{st}$  measures genetic distance from overall population

Any Questions

