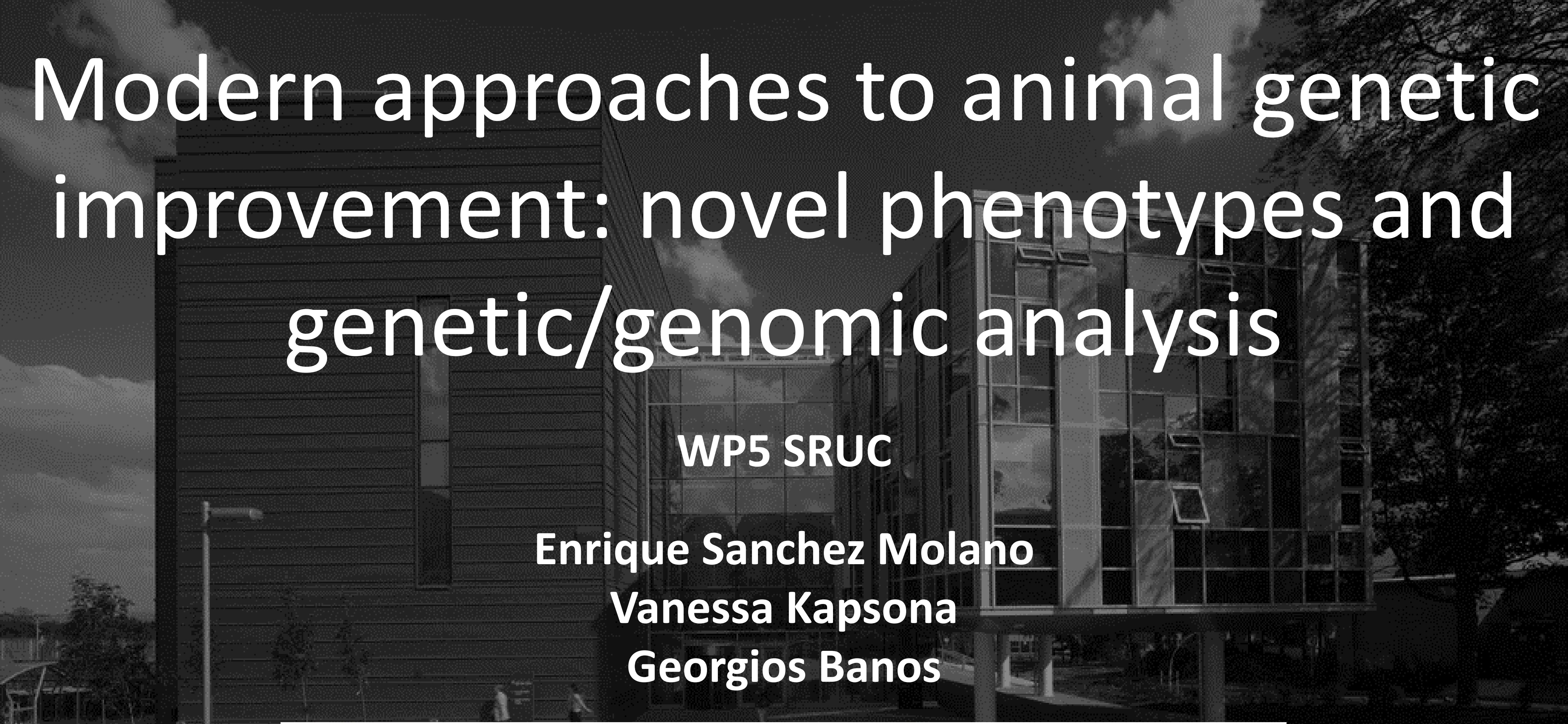


Modern approaches to animal genetic improvement: novel phenotypes and genetic/genomic analysis

WP5 SRUC

Enrique Sanchez Molano
Vanessa Kapsona
Georgios Banos

iSAGE
Innovation for Sustainable
Sheep and Goat
Production in Europe

KEY QUESTIONS

- **What is genetic improvement?**

KEY QUESTIONS

- **What is genetic improvement?**

  Improving animal performance with selective breeding

KEY QUESTIONS

- **What is genetic improvement?**
  Improving animal performance with selective breeding

- **What constitutes animal performance?**

KEY QUESTIONS

- **What is genetic improvement?**
  Improving animal performance with selective breeding

- **What constitutes animal performance?**
  Production, Health, Fertility, Longevity, …

KEY QUESTIONS

- **What is genetic improvement?**
  Improving animal performance with selective breeding

- **What constitutes animal performance?**
  Production, Health, Fertility, Longevity, …

- **5 steps to genetic improvement?**

KEY QUESTIONS

- **What is genetic improvement?**
  Improving animal performance with selective breeding

- **What constitutes animal performance?**
  Production, Health, Fertility, Longevity, …

- **5 steps to genetic improvement?**

Setting the breeding goal → **Animal recording (data)** → **Genetic evalution (data analysis)** → **Selection of the best animals** → **Appropriate mating strategies**

CHALLENGES AND OPPORTUNITIES

- **Changes in:**
  Environment, climate, markets, policies, …

- **Animal improvement must change accordingly and adopt new strategies/methods**

- **New genomic and sequencing tools, molecular advances**

- **Scientific inter-disciplinary collaboration is key (iSAGE)**

WHAT IS THIS ABOUT?

- **Tools for genetic improvement:**
    - Classic methods for genetics
    - New methods for genetic and environment interactions
    - Molecular tools to detect genes underlying performance

OUTLINE

- **Phenotypes and genotypes. DNA structure and variation. DNA markers**
- **Genetic and phenotypic analysis**
  - Classic model – pedigree
    - Variance components and heritability
    - Prediction of animal genetic values and accuracy
  - GxE – novel phenotypes with reaction norms
    - Phenotypic description
    - Heritability
  - Incorporating genotypes/genomics
    - GWAS
    - RHM

PHENOTYPES AND GENOTYPES.
DNA STRUCTURE.
DNA VARIATION.
DNA MARKERS.

Credit: irishtimes.com

PHENOTYPE

Phenotype (P) = Genotype (G) + Environment (E)

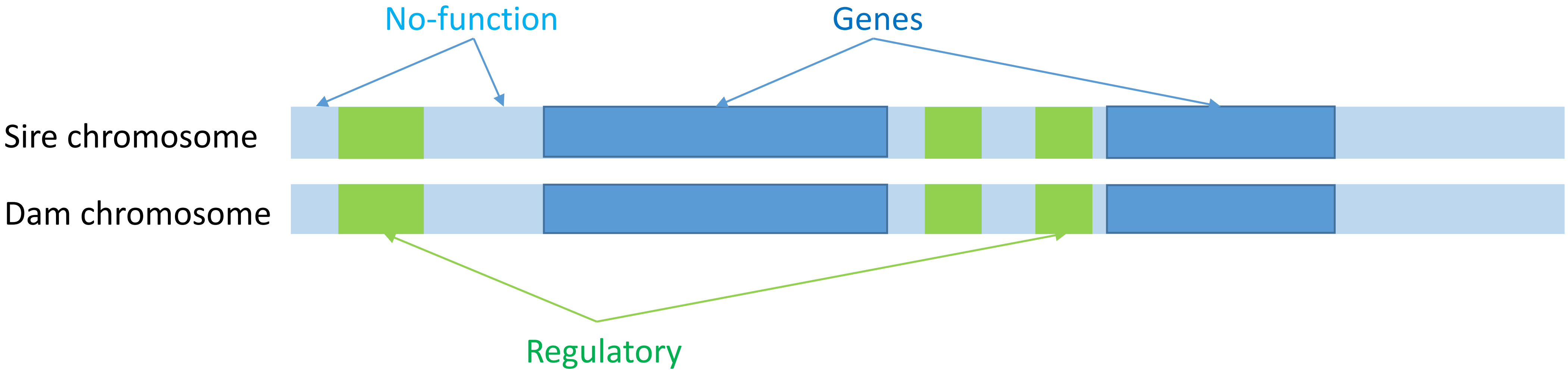
- **Phenotype** : Trait, observable characteristic of an animal
- **Genotype** : Genetic makeup of an individual. We can act on it. Inherited
- **Environment** : Events around the individual. We can partially act on it. Not inherited

GENOTYPE

- **We can act on it**
- **Inherited**
- **Determined by genes (DNA segments coding for functions related to the trait)**
- **Different organisms/cells will have different number of DNA copies**
  - Haploid (one copy of DNA): Male wasps, animals gametes (sperm, egg cell)
  - Diploid (two copies of DNA): Most animals
  - Triploid (three copies of DNA): Bananas, ginger
  - Tetraploid (four copies of DNA): some species of fish
  - Etc

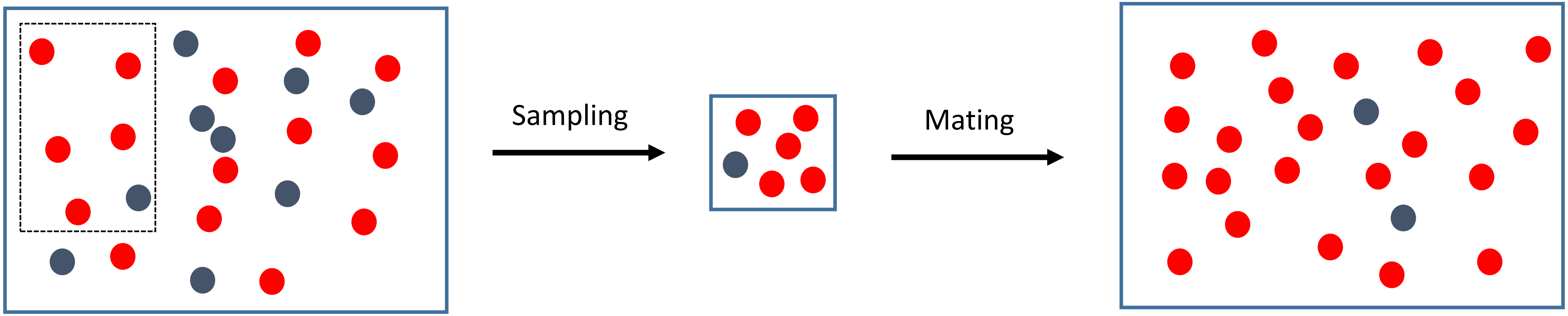
DNA STRUCTURE

- **Diploid animals: One copy inherited from the sire and other copy from the dam**
- **Each DNA copy is a sequence of nucleotides (biomolecules)**
- **In the DNA there are 4 types of nucleotides (A, G, C, T)**
- **DNA organises in chromosomes**



DNA VARIATION

- **Individuals are different**
- **Genetic (DNA) variation is important → Adaptation and improvement**
- **Mechanisms that decrease variation (eg: drift/randomness and selection)**



DNA VARIATION

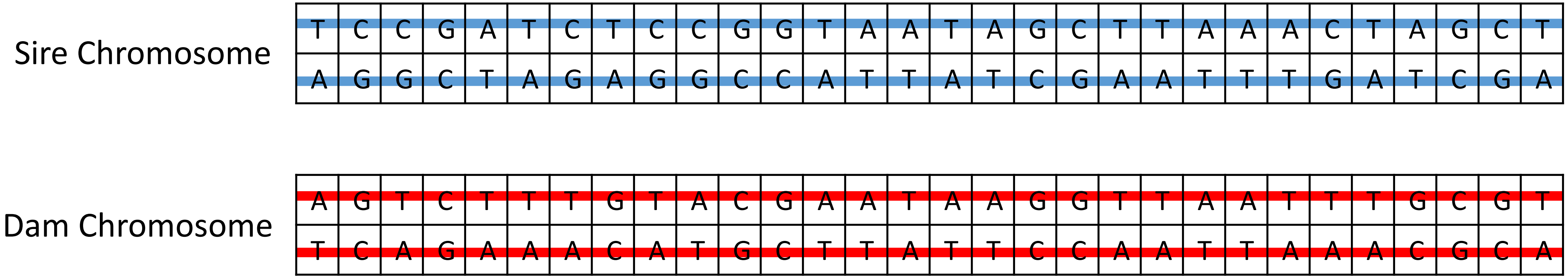
- **Mechanisms that increase variation such as recombination**

Sire Chromosome	T	C	C	G	A	T	C	T	C	C	G	G	T	A	A	T	A	G	C	T	T	A	A	A	C	T	A	G	C	T
	A	G	G	C	T	A	G	A	G	G	C	C	A	T	T	A	T	C	G	A	A	T	T	T	G	A	T	C	G	A

Dam Chromosome	A	G	T	C	T	T	T	G	T	A	C	G	A	A	T	A	A	G	G	T	T	A	A	T	T	T	G	C	G	T
	T	C	A	G	A	A	A	C	A	T	G	C	T	T	A	T	T	C	C	A	A	T	T	A	A	A	C	G	C	A

DNA VARIATION

- **Mechanisms that increase variation such as recombination**

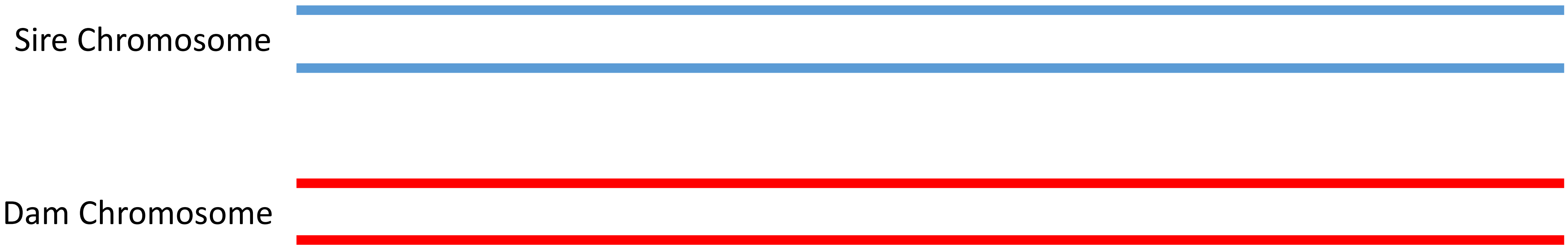


Sire Chromosome	T	C	C	G	A	T	C	T	C	C	G	G	T	A	A	T	A	G	C	T	T	A	A	A	C	T	A	G	C	T
	A	G	G	C	T	A	G	A	G	G	C	C	A	T	T	A	T	C	G	A	A	T	T	T	G	A	T	C	G	A

Dam Chromosome	A	G	T	C	T	T	T	G	T	A	C	G	A	A	T	A	A	G	G	T	T	A	A	T	T	T	G	C	G	T
	T	C	A	G	A	A	A	C	A	T	G	C	T	T	A	T	T	C	C	A	A	T	T	A	A	A	C	G	C	A

DNA VARIATION

- **Mechanisms that increase variation such as recombination**



Sire Chromosome

Dam Chromosome

DNA VARIATION

- **Mechanisms that increase variation such as recombination**

DNA VARIATION

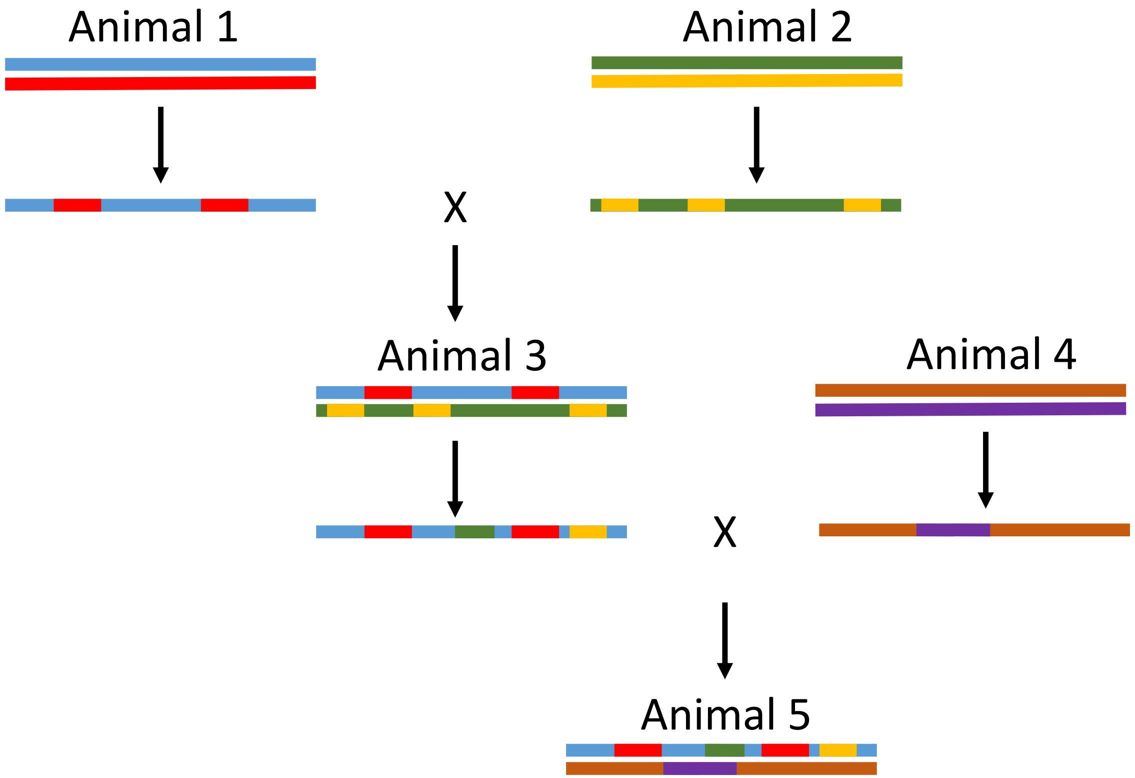
- **Mechanisms that increase variation such as recombination**

Sire Chromosome

Dam Chromosome

LINKAGE DISEQUILIBRIUM

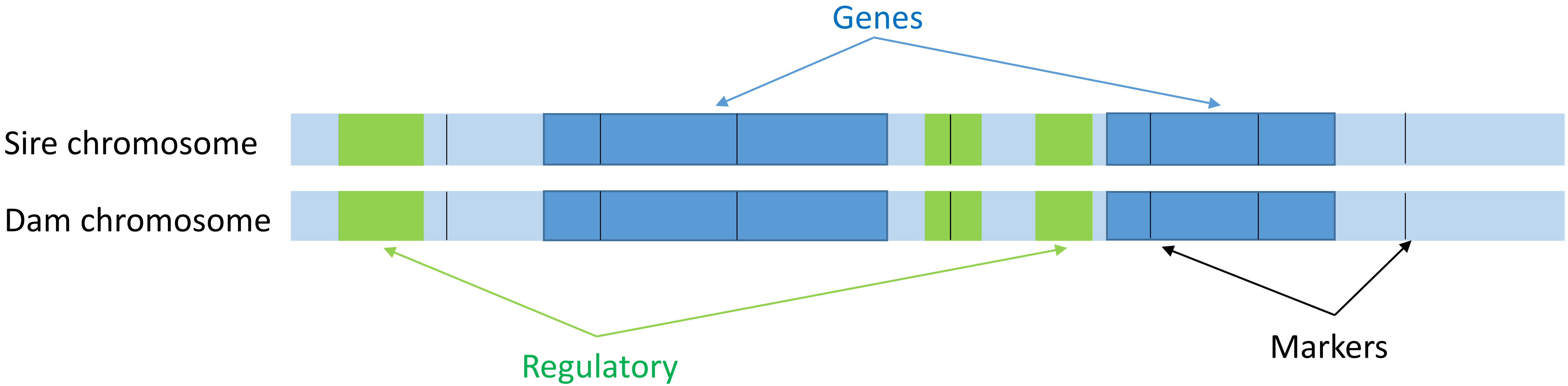
- **DNA is composed by “blocks” (haplotypes) of different length**
- **Haplotypes are inherited together (they are in LD)**



- **Recombination reduces LD (sorter haplotypes)**
- **Selection and drift increase LD (longer haplotypes)**

BASICS: GENOTYPE MARKERS

- **Markers: Variable sequences with known position**

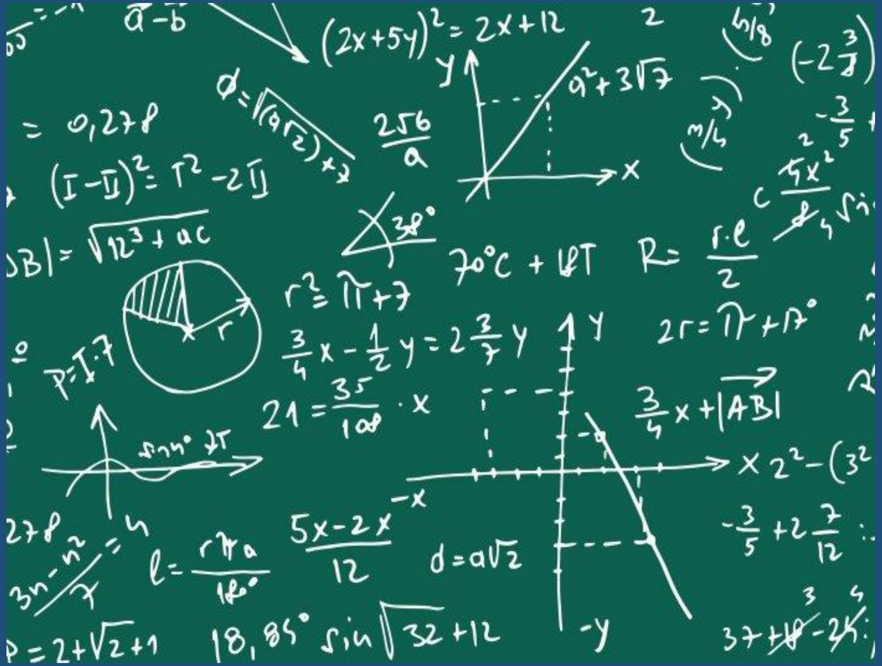


- **They can be in LD with genes related to a trait of interest**
- **If they are in LD with interesting genes, we can use them for selection**
- **We can also use them to see how animals are similar to each other (relationships)**

SUMMARY – TAKE HOME MESSAGE

- **A trait has a genetic component and a environment component**
- **We can act on the genetics in addition to the environment**
- **Genetics is dictated by the DNA**
- **DNA variation is important and it drives genetic improvement**
- **Associated DNA blocks called haplotypes are formed due to these forces**
- **Linkage disequilibrium allows us to select using markers and detect genes of interest**

GENETIC AND PHENOTYPIC ANALYSIS: CLASSIC GENETIC MODEL.



Credit: imperial.ac.uk

VARIANCE COMPONENTS

Phenotype (P) = Genotype (G) + Environment (E)

Phenotype (P) = additive (a) + dominance (d) + epistatic (i) + Environment (E)

- **Taking variances on the model, we have:**

$$V_P = V_a + V_d + V_i + V_E$$

- **Then, we define the heritability (h^2) as V_a/V_P , so the proportion of phenotypic variation that is additive genetic (heritable)**
- **Heritability ranges from 0 to 100%**

EXAMPLE OF VARIANCE COMPONENTS

Phenotype	V_P	V_A	h^2
Milk yield (kg. Yorkshire goats)	0.938	0.368	0.39
Weaning weight (kg. sheep)	9.004	3.441	0.27
Birth weight (kg. Jamunapari goats)[1]	0.324	0.046	0.14
Muscle depth (mm. Suffolk sheep)[2]	10.39	3.81	0.37
Fertility (service success. Dairy cattle)[3]	0.250	0.0179	0.071
Fertility (ewe lambed or not. Finnsheep)[4]	0.1317	0.0059	0.044

1 Rout et al. Genetic analysis of growth parameters and survival potential of Jamunapari goats in semiarid tropics. Small Rumin Res. 2018. 165: 124-130

2 Matheson et al. Genetic parameters for fitness and neonatal behavior traits in sheep. Behav Genet. 2012. 42: 899-911

3 Muuttoranta et al. Genetic parameters for female fertility in Nordic Holstein and Red Cattle dairy breeds. J. Dairy Sci. 2019. 102: 8184-8196

4 Matos et al. Genetic Analysis of Discrete Reproductive Traits in Sheep Using Linear and Nonlinear Models: I. Estimation of Genetic Parameters. J. Anim. Sci. 1997. 75: 76-87

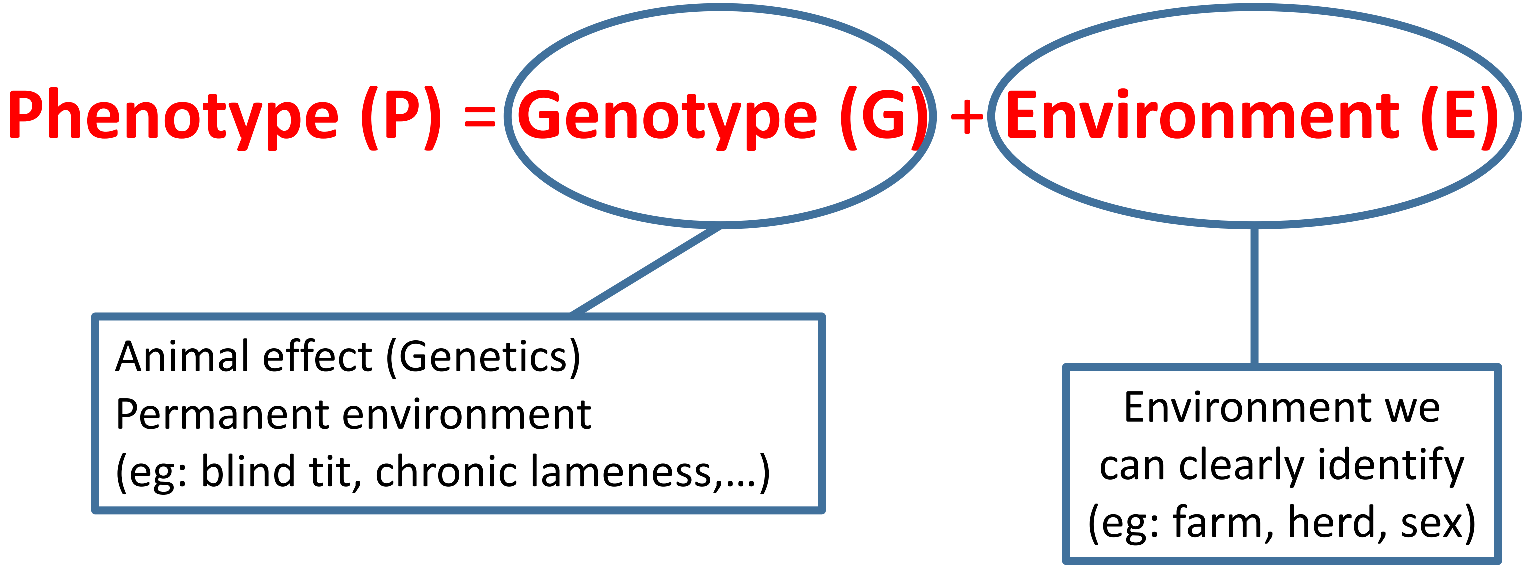
PREDICTION OF ANIMAL GENETIC EFFECT

- **How good are animals for a given phenotype?**
- **Statistic models to estimate how genetically good an animal is**

Phenotype (P) = Genotype (G) + Environment (E)

PREDICTION OF ANIMAL GENETIC EFFECT

- **How good are animals for a given phenotype?**
- **Statistic models to estimate how genetically good an animal is**



Phenotype (P) = Genotype (G) + Environment (E)

Animal effect (Genetics)
Permanent environment
(eg: blind tit, chronic lameness,…)

Environment we
can clearly identify
(eg: farm, herd, sex)

PREDICTION OF ANIMAL GENETIC EFFECT

- **How good are animals for a given phenotype?**
- **Statistic models to estimate how genetically good an animal is**

Phenotype (P) = Genotype (G) + Environment (E)

$$Y_i = X + u_i + \varepsilon$$

- **Phenotype (Y)** : Trait, observable characteristic of an animal
- **Fixed (X)** : Part of the environment. Eg: Farm, sex, herd, …
- **Random (u)** : Genetic value (a) + permanent environment (pe)
- **Error (ε):** Residual due to imperfection of the model

PREDICTION OF ANIMAL GENETIC EFFECT

- **We need pedigree data to disentangle genetics from environment**
  - We know genetic relations
  - We can see how the phenotype varies in animals that are related
- **Estimates of the true animal genetic value (additive breeding values, EBVs)**
- **The accuracy (r) of an EBV indicates how well it represents the true breeding value**
- **The accuracy depends on the heritability and how informative the pedigree is**
- **We can use EBVs in selection to avoid environmental bias**

EXAMPLES OF EBV ESTIMATION

- **EBVs estimated on progeny testing (n progeny) and different heritabilities (h^2)**

n (h^2=0.3)	accuracy
10	0.67
20	0.79
50	0.90
100	0.94
500	0.99
1000	0.99

h^2 (n=100)	accuracy
0.05	0.75
0.10	0.85
0.20	0.92
0.30	0.94
0.40	0.96
0.50	0.97

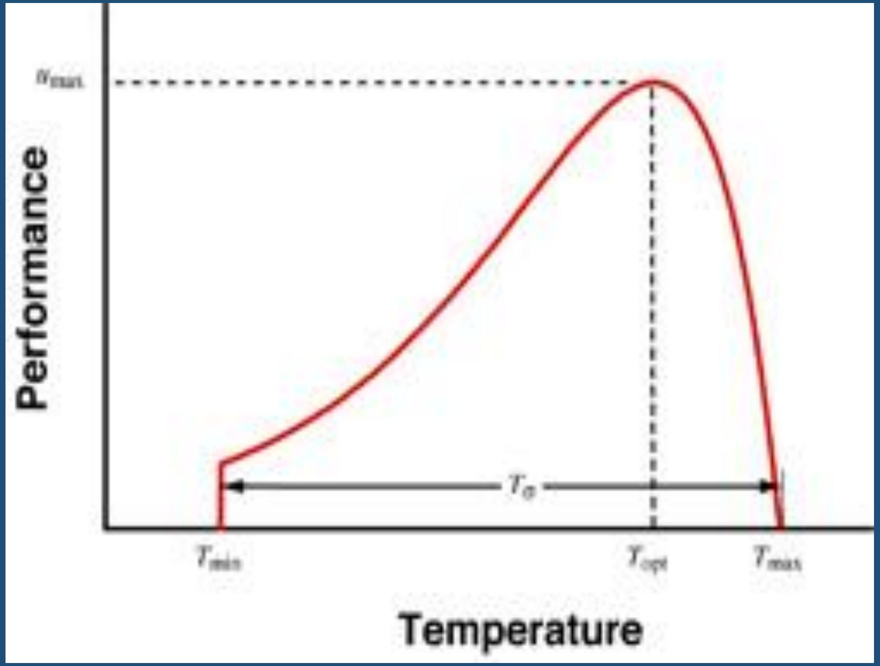
SUMMARY – TAKE HOME MESSAGE

- **Estimated breeding values (EBVs) for a trait provide estimates of the animal genetic value for that trait independently of the environment**
- **We use statistic models to estimate them**
- **We correct the phenotype for the known environment**
- **Use pedigree to obtain estimates of the genetic values (EBVs)**
- **EBVs reflect the true genetic values with a given accuracy**
- **Use EBVs for selection rather than phenotypes to avoid bias**

BREAK

BUT YOU CANNOT RUN AWAY FROM GENETICS

GENETIC AND PHENOTYPIC ANALYSIS:
GxE
REACTION NORMS.



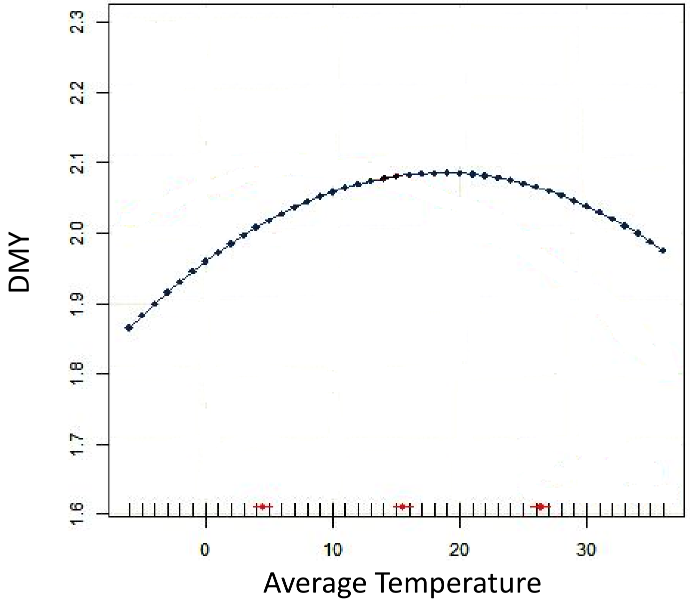
Credit: gwgilc.people.wm.edu

REACTION NORMS

- **What if our genotype behaves differently depending on the environment?**

Phenotype (P) = Genotype (G) + Environment (E) + Gen_by_Env (GxE)

- **Reaction norm: how the phenotypic expression changes in a range of environments**



REACTION NORMS

- **Interested in average animal performance AND how it changes with environment**
- **We use “Random regression models” (RRM)**
- **RRM model population/animal changes in performance along different values of an environmental variable (eg: daily milk yield changes with temperature)**
- **In the model we will have two terms of interest**
  - The changes in population performance across the environmental effect of interest
  - The changes in individual performance across the environmental effect of interest

REACTION NORMS: GENERAL MODEL

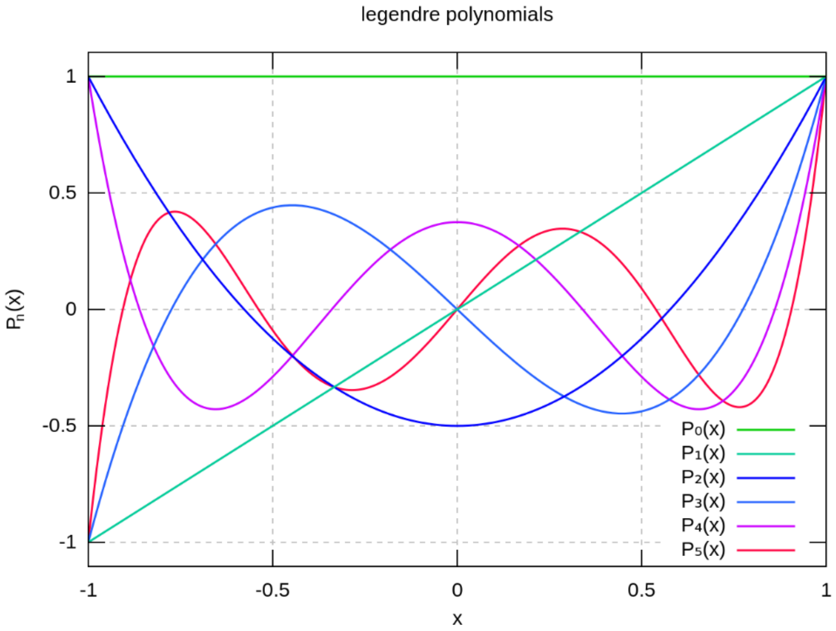
$Y_{ij} = X + f(\beta , X_j) + a_i + e_{ij}$ ***Population level***

$Y_{ij} = X + f(\beta , X_j) + f_i(a_i , X_j) + f_i(p_i , X_j) + e_{ij}$ ***Individual level***

Y_{ij}	Phenotype	Performance of an animal *i*
X	Environment	Set of fixed effects (excluding the environment of interest)
$f(\beta , X_j)$	Environment	General relation between performance and environment *j* (eg: temperature)
a_i	Genetic	Animal effect (as described before) of animal *i*
$f_i(a_i , X_j)$	GxE	Describe the relation between individual animal genetic performances and environment *j* deviated from the general relation f(β , Xj)
$f_i(p_i , X_j)$	Other	Describe the relations between individual animal permanent environment and environment *j* deviated from the general relation f(β , Xj)
e_{ij}	Other	Error due to model imperfection

REACTION NORMS: GENERAL MODEL

- **We need functions to describe the terms $f(\beta , X_j)$ and $f_i(a_i+p_i , X_j)$**
- **We used Legendre polynomials**
- **Legendre polynomials allow to describe linear and non-linear relations between the population or the individuals and the environmental variable**



REACTION NORMS: PHENOTYPIC MODEL

- **As we want individual phenotypes (and not genetic effects), we do not fit pedigree**
- **Thus, the general model becomes**

$$Y_{ij} = X + f(\beta\ , X_j) + f_i(a_i\ , X_j) + f_i(p_i\ , X_j) + e_{ij} \quad \rightarrow \quad Y_{ij} = X + f(\beta\ , X_j) + \boxed{f_i(a_i + p_i\ , X_j)} + e_{ij}$$

PxE : Describes the relation between the individual phenotypic performance and environment

REACTION NORMS: POPULATION PHENOTYPE

- **Step 1:** We use Legendre polynomials of degree 2 to explore non-linearity, so we calculate the corresponding *b* coefficients for temperature *T*

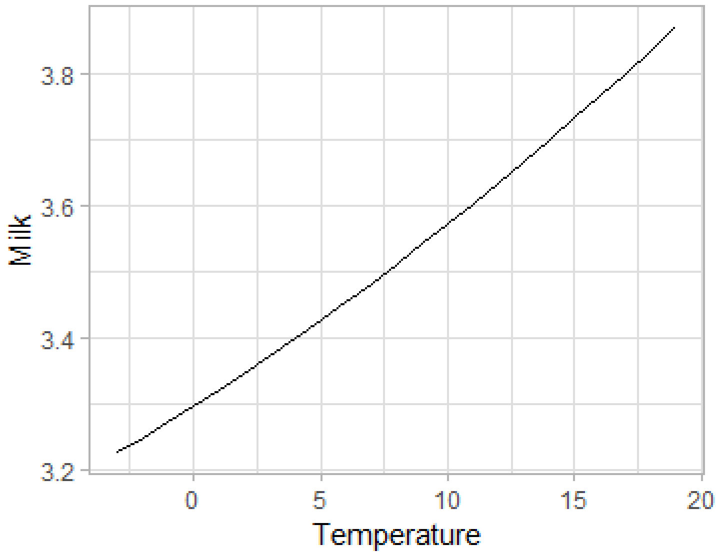
$$\boldsymbol{f(\beta , X_j)} = b_0 + b_1 T + b_2 T^2$$

- **Step 2:** Fit the population model. Animal effect is fitted as random.

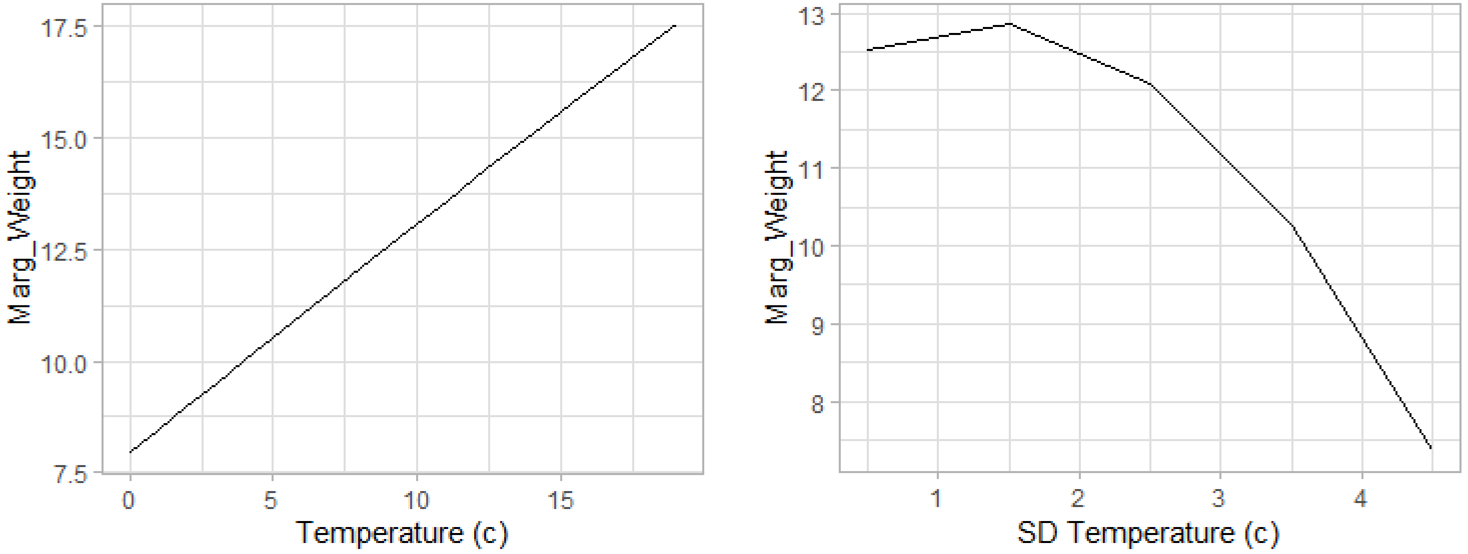
$$\boldsymbol{Y_{ij}} = \boldsymbol{X} + \boldsymbol{f(\beta , X_j)} + \boldsymbol{a_i} + \boldsymbol{e_{ij}}$$

- **Step 3:** Use the output of the model (model coefficients) to predict based on given temperature values and plot results

REACTION NORMS: POPULATION PHENOTYPE



Reaction norm for milk yield (kg) depending on average temperature



Reaction norms for weight gain (kg) depending on average temperature (left) and temperature variation (right)

REACTION NORMS: INDIVIDUAL PHENOTYPE

- **Step 1:** We use Legendre polynomials of degree 1 (as population RN were mostly linear), so we calculate the corresponding *b* coefficients for temperature *T*

$$f(\beta, X_j) = b_0 + b_1 T$$

- **Step 2:** Fit the individual model. Animal effect is fitted as random and multiplies the Legendre polynomials

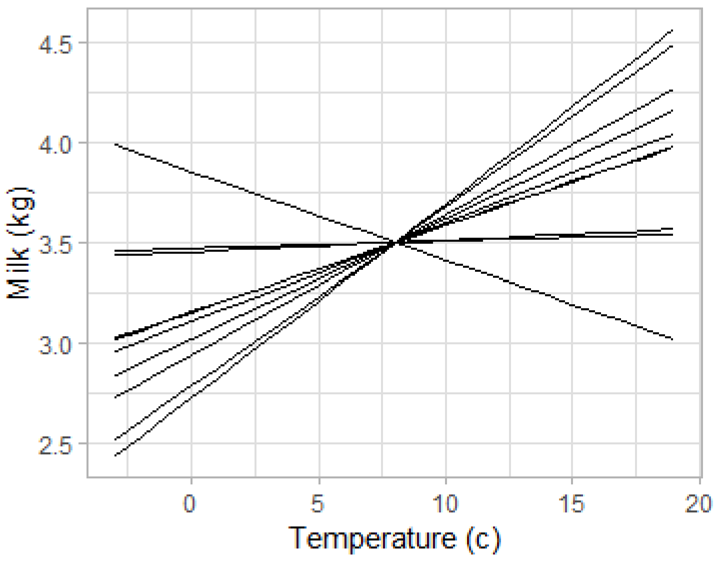
$$Y_{ij} = X + f(\beta, X_j) + f_i(a_i + p_i, X_j) + e_{ij}$$

$$DMY \sim \text{Fixed effects} + f(\beta, X_j) + f(\beta, X_j) * animal + e_{ij}$$

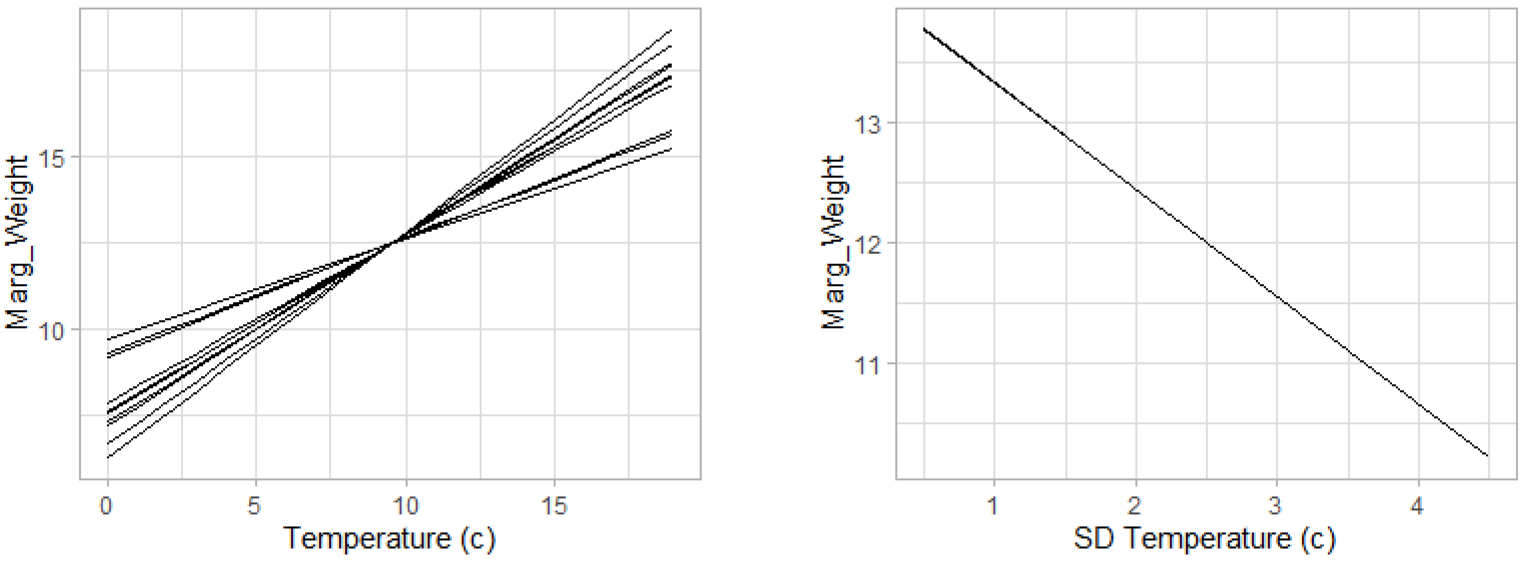
The term $f(\beta, X_j)*animal$ is called "nested effect". It is a random effect where each member of a group is contained in a single unit of another group

REACTION NORMS: INDIVIDUAL PHENOTYPE

- **Step 3:** Use the output of the model (model coefficients) to predict based on given temperature values and plot results (eg: 10 animals)



Reaction norms for milk yield (kg) depending on average temperature



Reaction norms for weight gain (kg) depending on average temperature (left) and temperature variation (right)

REACTION NORMS: INDIVIDUAL PHENOTYPE

- **Step 4:** We can estimate the slope of the phenotypic reaction norm. This is a measure of how the performance of each animal will change with weather

Stability phenotype: Slopes of the phenotypic reaction norms

Slopes ~ 0 → Animal production is not affected by weather changes
Slopes ≠ 0 → Animal production is affected by weather changes

GENETIC ANALYSIS FOR INDIVIDUAL PHENOTYPES

- **We have animal phenotypes for stability (slopes)**
- **We can use an animal model and pedigree information to estimate breeding values**
- **The model estimates the variance components and heritability of these phenotypes**
- **If there is genetic variance and heritability, we can use these phenotypes to improve stability of performance**

Phenotype	V_P	V_A	h^2
Milk yield to aver. temperature (goats)	2.06E-3	2.18E-4	0.110
Gain weight to aver. temperature (sheep)	0.9513E-2	0.1388E-2	0.146
Gain weight to var. temperature (sheep)	0.3095E-7	0.8137E-8	0.263

SUMMARY – TAKE HOME MESSAGE

- **Random regression models (RRM) are used to study evolution of traits across the variation of the environment**
- **We are interested in how each animal reacts to the changes in weather for a particular trait**
- **We use linear and non-linear functions to relate performance and weather variables**
- **We can estimate the slopes of the reaction norm (even at different points for non linear functions) as a measure of stability of performance at that value of the weather variable**
- **These slopes measure the stability of performance. If they are heritable, they can be used for selection to improve stability of performance**

GENETIC AND PHENOTYPIC ANALYSIS: INCORPORATING GENOTYPES AND GENOMICS

Credit: treehugger.com

GENERAL OVERVIEW

- **By including information (pedigree), we can estimate the genetic value of an animal**
  - Pedigree relationships are averages
- **Genetic markers also provide information**
  - Markers across all the genome
  - Estimate actual genetic relationship
  - Markers in LD with genes of interest
- **We can use genetic markers (genotypes) to provide information instead of/together with pedigree in the animal models**
- **We can also use genotypes to detect genes or genomic regions of interest**

GWAS

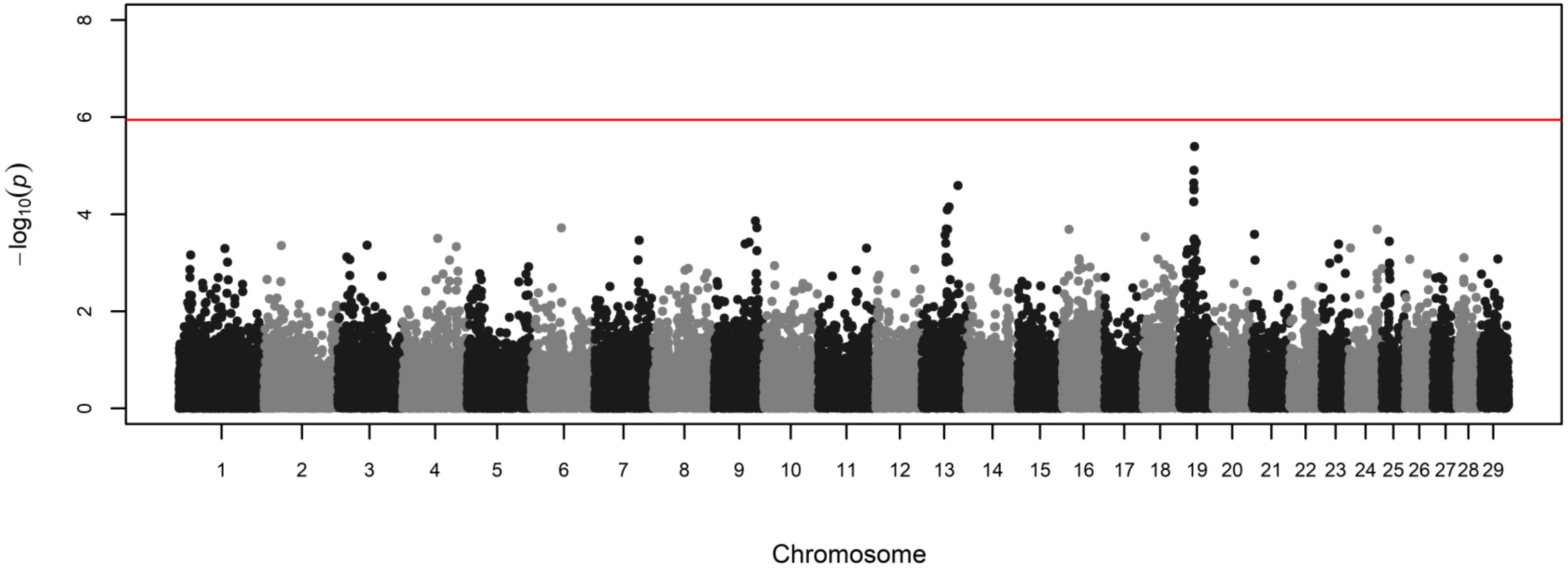
- **We can use genotypes to detect genes or genomic regions of interest**
- **Tens of thousands of markers distributed across the genome**
- **Some of these markers might be linked (LD) to genes related with the trait of interest**
- **If we have animals with genotypes and phenotypes, we can:**
  - Statistical analyses to test the association of each marker with the phenotype

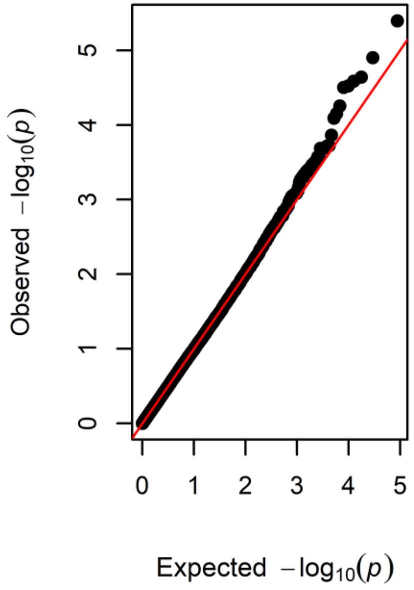
Classic animal model ➔ $\boldsymbol{Y} = \boldsymbol{X} + \boldsymbol{u} + \boldsymbol{\varepsilon}$

GWAS model ➔ $\boldsymbol{Y} = \boldsymbol{X} + \textbf{\textit{Markers}} + \boldsymbol{u} + \boldsymbol{\varepsilon}$

  - Look at the markers that are significantly associated and search for genes around

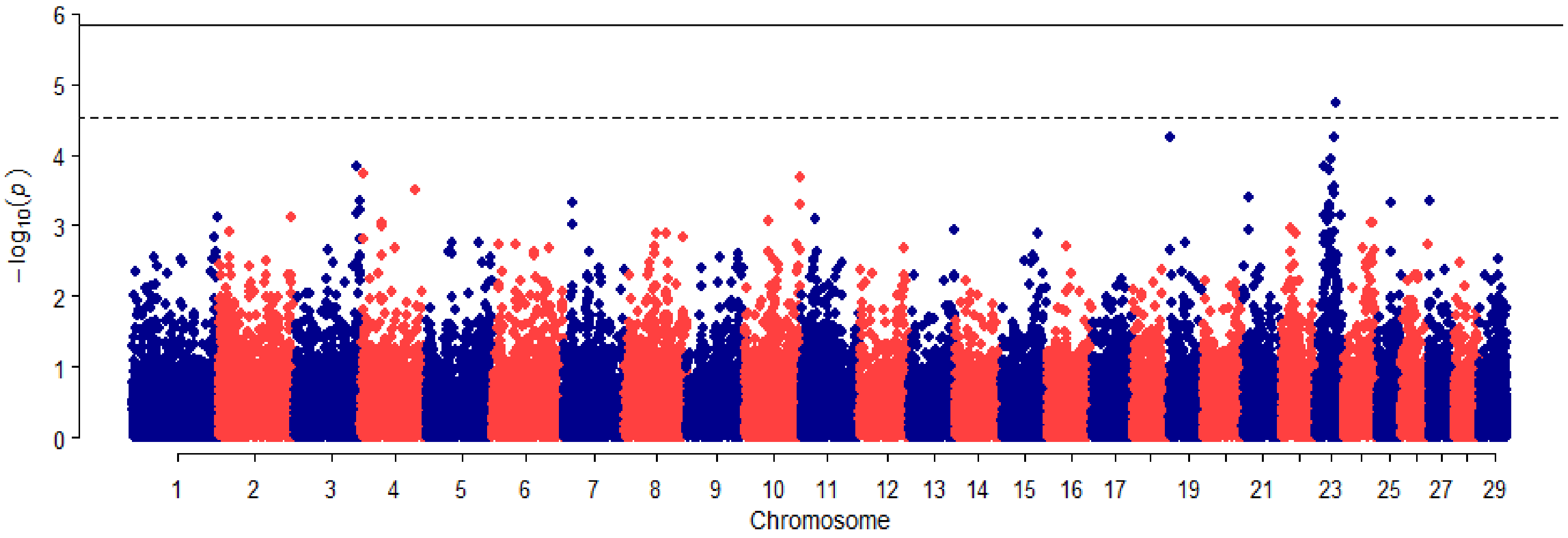
GWAS





Potential candidate genes were detected in relation to resilience of milk production to changes in average temperature in goats (Sánchez-Molano *et al*. 2019. BMC Genetics)

GWAS

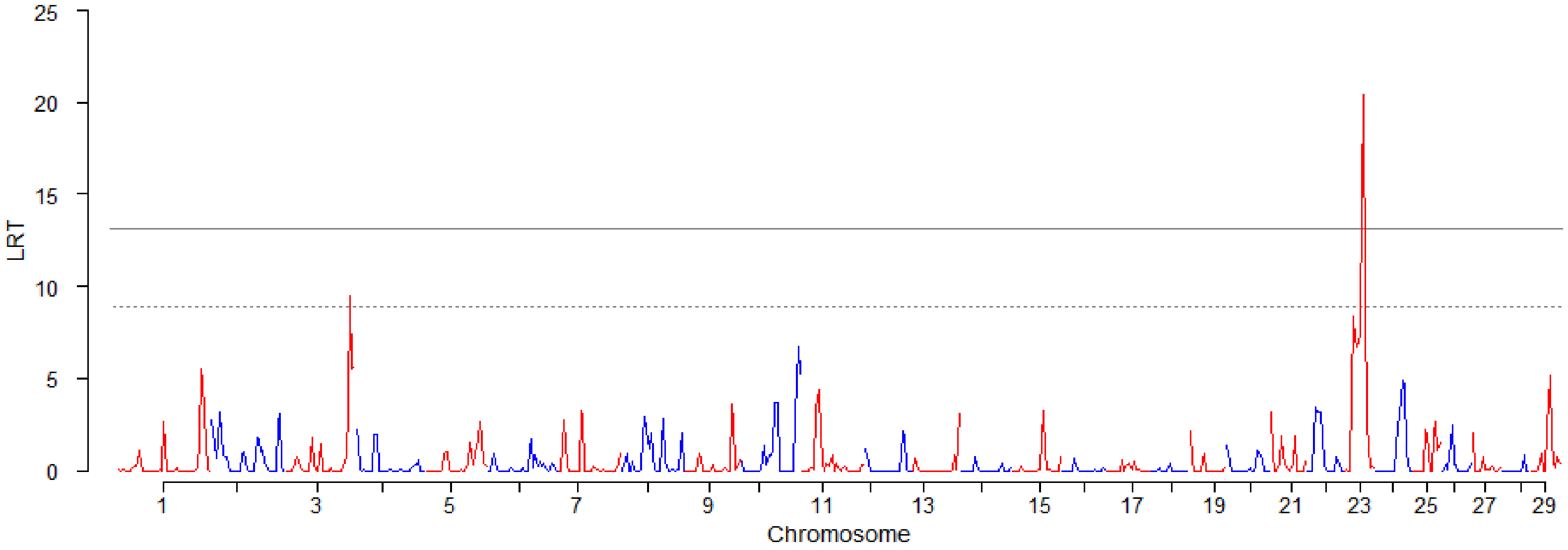


Genome-wide association analyses of three bovine tuberculosis susceptibility traits in Holstein-Friesian cattle (Raphaka et al. 2017. BMC Genetics)

REGIONAL HERITABILITY MAPPING

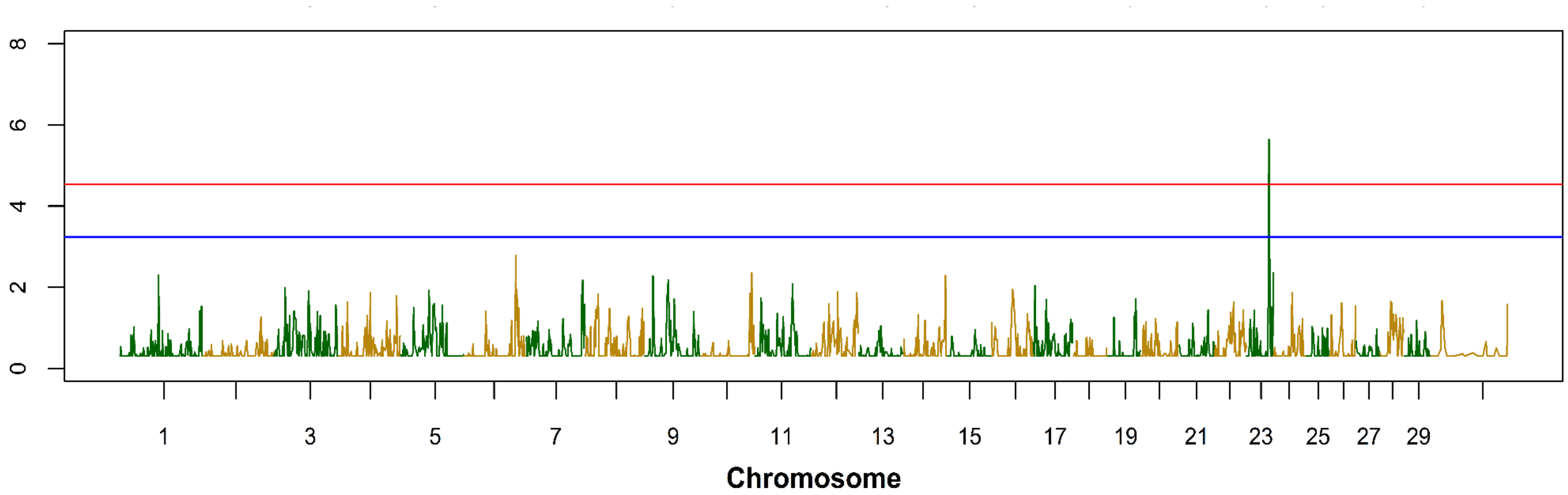
- **Instead of using markers, we can use groups of markers**
- **Some of these groups/regions will be associated with traits of interest**
- **Similar approach and model to GWAS, but some extra properties**
  - Less number of regions, so less statistical problems when the number of animals is small
  - GWAS detects marker effects that are very big. Regional heritability allows to estimate a collective effect for each region.

REGIONAL HERITABILITY MAPPING



Genome-wide association analyses of three bovine tuberculosis susceptibility traits in Holstein-Friesian cattle (Raphaka et al. 2017. BMC Genetics)

REGIONAL HERITABILITY MAPPING



Quantitative Trait Loci Mapping for Lameness Associated Phenotypes in Holstein–Friesian Dairy Cattle

(Sánchez-Molano *et al*. 2019. Frontiers in Genetics)

SUMMARY – TAKE HOME MESSAGE

- **While the classical method uses pedigree to inform about relationships in order to estimate the genetic effect of an animal, we can also use genetic markers**
- **Markers can be more powerful than pedigree if we have enough information**
- **Both pedigree and genotypes can be used to complement each other**
- **Markers can be also used to detect regions of interest through association analyses (GWAS and RHM)**

OVERALL SUMMARY

(YES, WE REACHED THE FINAL AND YOU ARE STILL ALIVE…HOPEFULLY)

OVERALL SUMMARY

- **Performance of animals is and will continue changing with weather**
- **Increased weather variation (climate change) will impact animal performance and we need to account for it in selection programmes**
- **Reaction norms can be used to develop phenotypes of resilience to weather variation**
- **These resilience phenotypes are heritable and can be used for selection**

FURTHER QUESTIONS AND ISSUES

- **Use of resilience phenotypes in selection. Which is the goal?**
  - Depends on what is more important: increase in variation or in average
  - If increase in variation: We probably want slopes around 0 (eg: animals with stable production regardless of the temperature)
  - If increase in average: We probably want positive/negative slopes (eg: animals with higher production under higher temperatures)

- **Depends entirely on the future changes expected in the place as well as on the species**

THANKS!

E-mail: georgios.banos@sruc.ac.uk
enrique.sanchez-molano@roslin.ed.ac.uk

Telephone: +44 (0)131 651 9342

Twitter: @GgabRoslinSRUC

Address: SRUC, Roslin Institute Building, Easter Bush, Midlothian EH25 9RG. Scotland