

# Exploiting Homozygosity In The Era of Genomics: Inbreeding, ROH, and Genomic Mating Programs

Joint ADSA-Interbull Session: Breeding and Genetics: Ten Years of Genomic  
Selection

C. Maltecca\*, F. Tiezzi, C. Baes, J.Pryce, J. Cole  
\*cmaltec@ncsu.edu

ADSA Meeting, Cincinnati, Ohio – June 24, 2019

# Inbreeding and Genomic Information

Landscape Change



Genomics to predict breeding values is now standard in animal breeding

- ▶ Adopted in pretty much all species
- ▶ Mature and effective machinery for the prediction of GEBV

In dairy > 2 million individuals have genotype information.

- ▶ Their genotypes not fully integrated or exploited

Opportunity to incorporate population/herd management at the genomic level.

- ▶ Purebred:

- ▶ Management of Lethal and Sub-Lethal Mutations.
- ▶ Functional Inbreeding Depression.
- ▶ Genetic Diversity.
- ▶ Breed/Population divergence.
- ▶ Optimal Matching of Genomes.



The accumulation of inbreeding is unavoidable:

- ▶ Intense directional selection
- ▶ High variance of reproductive success
- ▶ Use of BLUP and truncation selection

Inbreeding "**can**" reduce the mean value of a trait



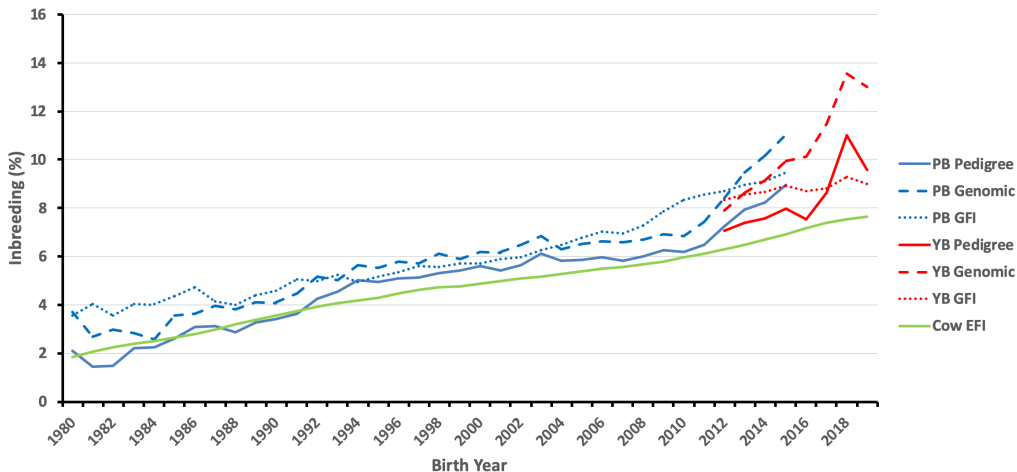
# How genomic selection influences inbreeding



- ▶ The rate of inbreeding per year increases due to shortening of generation interval
- ▶ The rate of inbreeding per generation decreases because MS is better assessed
- ▶ The rate of inbreeding per generation decreases because a larger pool of genotypes can potentially be sampled

All of these are true but the "**net**" effect is a faster accumulation of homozygosity in the population

# How genomic selection influences inbreeding



# What does Inbreeding Measures

And why do we care

# Inbreeding



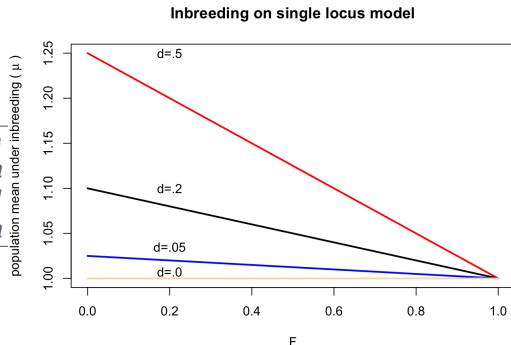
The probability of 2 random alleles at the same locus from 2 gametes which unite to be IBD from a common ancestor Crow and Kimura, 1970

genotype	Frequency	Genotypic Value
$A_1A_1$	$R_{11} = p^2(1 - F) + pF$	$a$
$A_1A_2$	$2R_{12} = 2pq(1 - F)$	$d$
$A_2A_2$	$R_{22} = q^2(1 - F) + qF$	$-a$

The mean of the inbred population will therefore be:

$$\mu F = a(p - q) + 2d(1 - F)pq$$

The reduction in the population mean due to inbreeding  $-2pqFd$ :



# Inbreeding Depression



Model	Parent genotypes	F <sub>1</sub> hybrid genotypes and their fitness (or quality) relative to the parent genotypes
<b>Recessive deleterious mutations</b>		
Dominance hypothesis		
Single locus	$A/A \times a/a$	$A/a$ Intermediate fitness but above the parental average (homozygote shows inbreeding depression)
Multiple loci (effects of different mutant alleles marked in hybrids)	$A/A \times b/b$ $a/a \times B/B$	$A/a \quad B/b$ High fitness (heterosis and inbreeding depression)
<b>Recessive deleterious mutations at closely linked loci</b>		
Pseudo-overdominance		
	$A \quad b$ $\times$ $a \quad B$	$A \quad b$ $a \quad B$ Higher fitness than the parent genotypes
<b>Single loci with heterozygous advantage</b>		
True overdominance		
	$A_1/A_1 \times A_2/A_2$	$A_1/A_2$ High fitness (heterosis and inbreeding depression in homozygote)

# Genetic Variance Under Inbreeding



The total genetic variance in the population following inbreeding is (Weir and Cockerham, 1977)

$$V_{GF} = (1 + F)V_A + (1 - F)V_D + 4FC_{AD} + FV_H - F^2D_H$$

where  $V_A$  and  $V_D$  are the additive genetic and dominance variances in the base population ( $F = 0$ ).

$$C_{AD} = \sum_i P_i \alpha_i \delta_{ii}, V_H = \sum_i P_i \delta_{ii}^2, D_H = (\sum_i P_i \delta_{ii})^2$$

When dominance variance is zero,

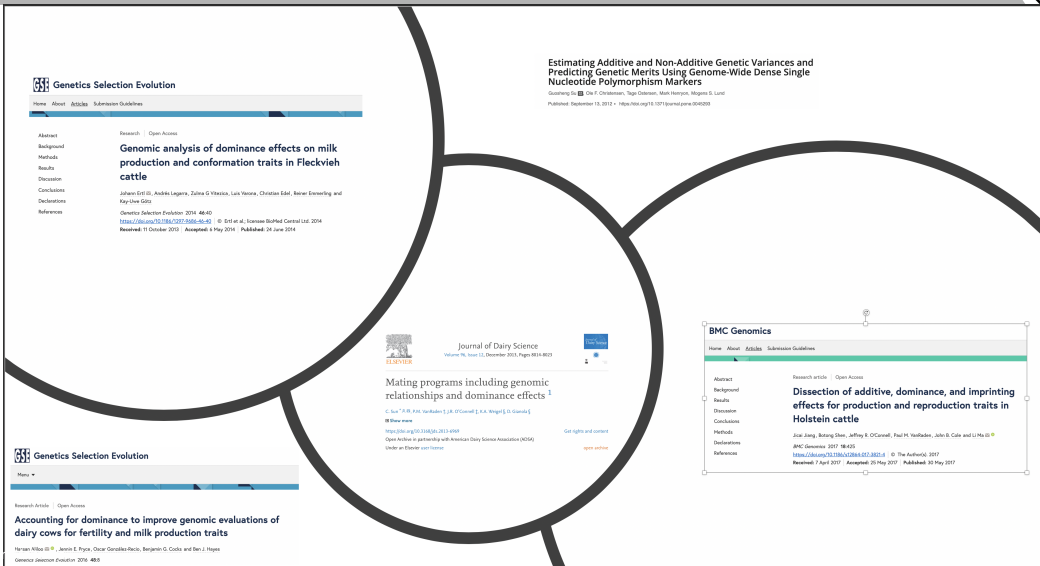
$$V_{AF} = (1 + F)V_A$$

Under correct conditions inbreeding may have advantages for selection because the variance of breeding values is increased.

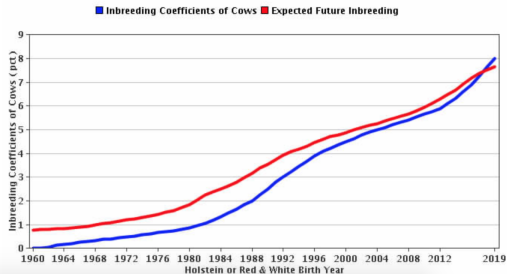
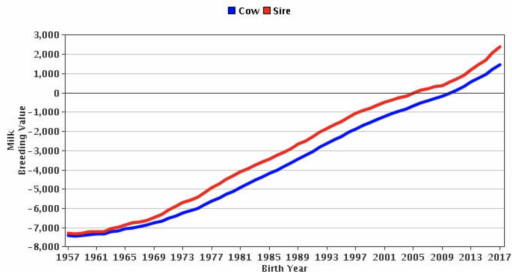
# OK then. How Much Dominance is there then?



10



# OK then. How Much Dominance is there then?



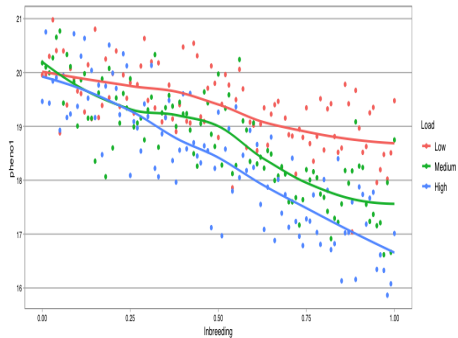
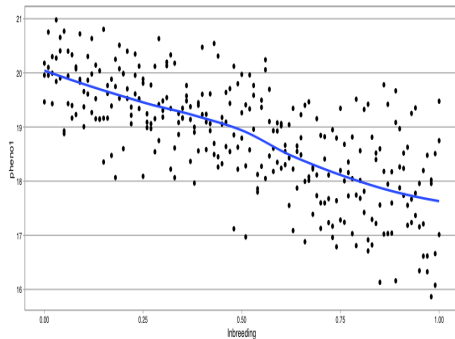
3

$$V_{GF} = (1 + F)V_A + (1 - F)V_D$$



- ▶ How can two individuals with similar breeding value be further discriminated on the basis of genomic diversity and fitness
- ▶ How should we do breeding balancing short and long term gains as well as fitness and overall variability?

# Primary Question



Johnsson et al. *Genet Sel Evol* (2019) 51:14  
<https://doi.org/10.1186/s12711-019-0456-8>

**GSE** Genetics  
Selection  
Evolution

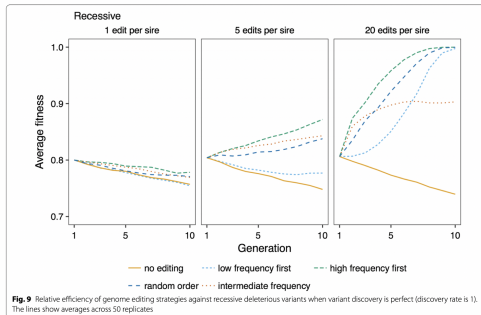
## RESEARCH ARTICLE

Open Access



## Removal of alleles by genome editing (RAGE) against deleterious load

Martin Johnsson<sup>1,2</sup>, R. Chris Gaynor<sup>1</sup>, Janez Jenko<sup>1</sup>, Gregor Gorjanc<sup>1</sup>, Dirk-Jan de Koning<sup>2</sup> and John M. Hickey<sup>1\*</sup>



# Determining Inbreeding Depression Based on Dominance Estimates



Finding Dominance is not easy:

- ▶ The proportion of the genetic variance at a causal variant that is captured by markers is  $\rho^2$  ( correlation between SNP and causal variant) for additive variance and  $\rho^4$  for dominance variance <sup>4</sup>
- ▶ Independence between additive and dominance effects does not hold in reality <sup>5</sup>
- ▶ Given the directionality of dominance, the a priori dominance effect of a gene should be positive so we account for it with inbreeding <sup>6</sup>

---

<sup>4</sup>Zhu et al 2015

<sup>5</sup>Huang and Mackay 2016

<sup>6</sup>Xiang et al. 2016

## ORIGINAL RESEARCH ARTICLE

Front. Genet., 14 May 2019 | <https://doi.org/10.3389/fgene.2019.00412>



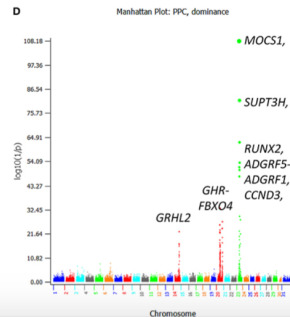
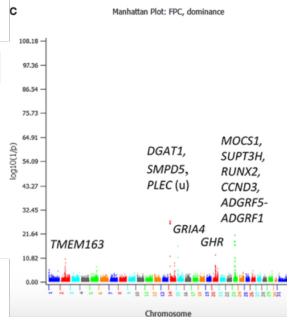
## A Large-Scale Genome-Wide Association Study in U.S. Holstein Cattle

Jicai Jiang<sup>1†</sup>, Li Ma<sup>1†</sup>, Dzianis Prakapenka<sup>2†</sup>, Paul M. VanRaden<sup>3†</sup>, John B. Cole<sup>3†</sup> and Yang Da<sup>2\*</sup>

<sup>1</sup>Department of Animal and Avian Sciences, University of Maryland, College Park, MD, United States

<sup>2</sup>Department of Animal Science, University of Minnesota, Saint Paul, MN, United States

<sup>3</sup>Animal Genomics and Improvement Laboratory, USDA-ARS, Beltsville, MD, United States



# Is Inbreeding still a useful Measure?

Yes it is

## ▶ **Pedigree Based Inbreeding**

- ▶ Expected probability of IBD
- ▶ Underestimate true inbreeding coefficient

## ▶ **Genomic Based Inbreeding**

- ▶ Realized proportion of genome IBS
- ▶ Approximate IBD probability

## ▶ **ROH Based Inbreeding**

- ▶ Estimate local (recent) autozygosity
- ▶ Approximate IBD probability

## ▶ **HBD Based Inbreeding**

- ▶ Estimate local autozygosity
- ▶ Models IBD probability

## Pedigree vs. Genomic Inbreeding



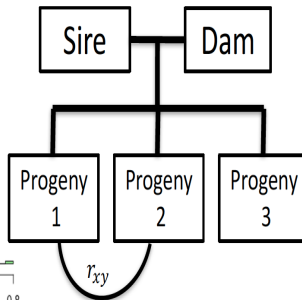
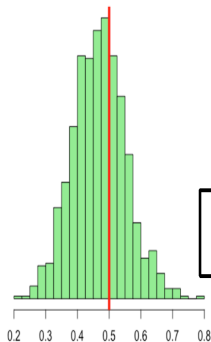
# Pedigree and Genomic Relatedness

Definition and advantages



## Visualizing Genomic Relationships

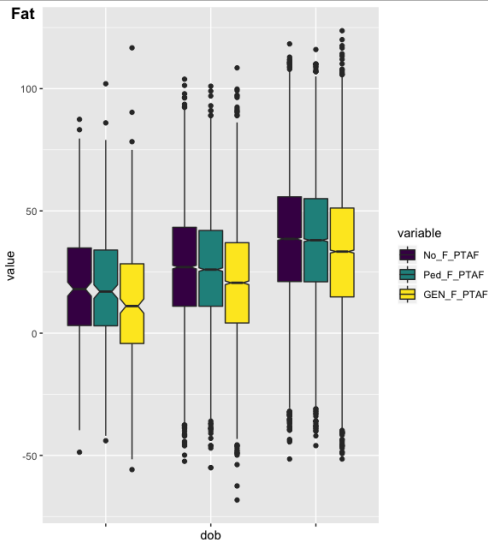
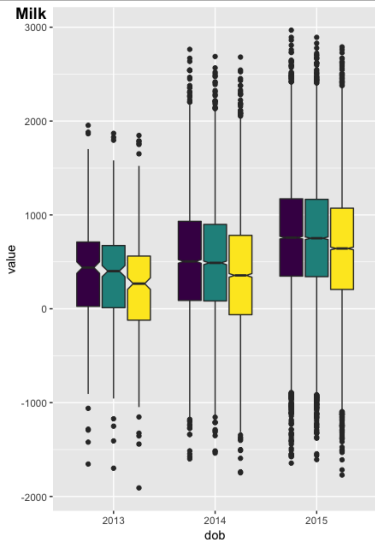
Full-Sib Relationships



Relationships

# Not all Inbreeding is Created Equal

## Inbreeding Load



# Not all Inbreeding is Created Equal

Inbreeding Load

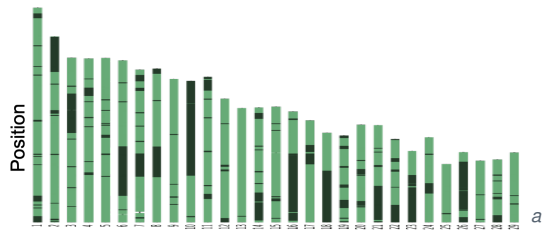
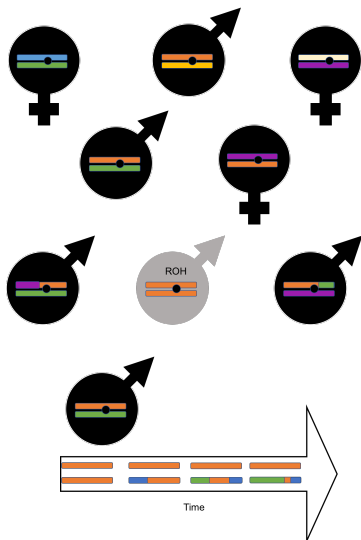


Trait	-10logp Pedigree	-10logp Genomic
Milk	4.95	18.06
Fat	4.67	9.96
Pro	2.18	13.47
PL	0.33	1.5
DPR	0.57	0.08
SCS	0.11	0.14

## Characterizing Local Inbreeding

# ROH

## Definition and advantages



<sup>a</sup>Baes, 2018

Identifying and managing inbreeding at the global and local level with the use of genomic information

**BMC Genomics**

Home About Articles Submission Guidelines

Abstract  
Background  
Results  
Discussion  
Conclusion  
Methods  
Declarations  
References

Research article | Open Access

**Investigation of regions impacting inbreeding depression and their association with the additive genetic effect for United States and Australia Jersey dairy cattle**

Jeremy T. Howard, Mekonnen Haile-Mariam, Jennie E. Pryce and Christian Maltecca

**GSE Genetics Selection Evolution**

Home About Articles Submission Guidelines

Abstract  
Background  
Methods  
Results and discussion  
Conclusions

Research Article | Open Access

**A simple strategy for managing many recessive disorders in a dairy cattle breeding program**

John B. Cole

**GSE Genetics Selection Evolution**

Home About Articles Submission Guidelines

Abstract  
Background  
Methods  
Results  
Discussion  
Conclusions

Research | Open Access

**Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle**

Jennie E Pryce, Mekonnen Haile-Mariam, Michael E Goddard and Ben J Hayes

**BMC Genomics**

Home About Articles Submission Guidelines

Abstract  
Background  
Results  
Discussion  
Conclusions  
Methods  
Declarations

Research article | Open Access

**Characterizing homozygosity across United States, New Zealand and Australian Jersey cow and bull populations**

Jeremy T Howard, Christian Maltecca, Mekonnen Haile-Mariam, Ben J Hayes and Jennie E Pryce



ORIGINAL ARTICLE

**Geno-Diver: A combined coalescence and forward-in-time simulator for populations undergoing selection for complex traits**

**A heuristic method to identify runs of homozygosity associated with reduced performance in livestock**

J. T. Howard, F. Tiezzi, Y. Huang, K. A. Gray, C. Maltecca

*Journal of Animal Science*, Volume 95, Issue 10, October 2017, Pages 4318–4332,  
<https://doi.org/10.2527/jas2017.1664>

Published: 01 October 2017 Article history ▼

# GenoDiver

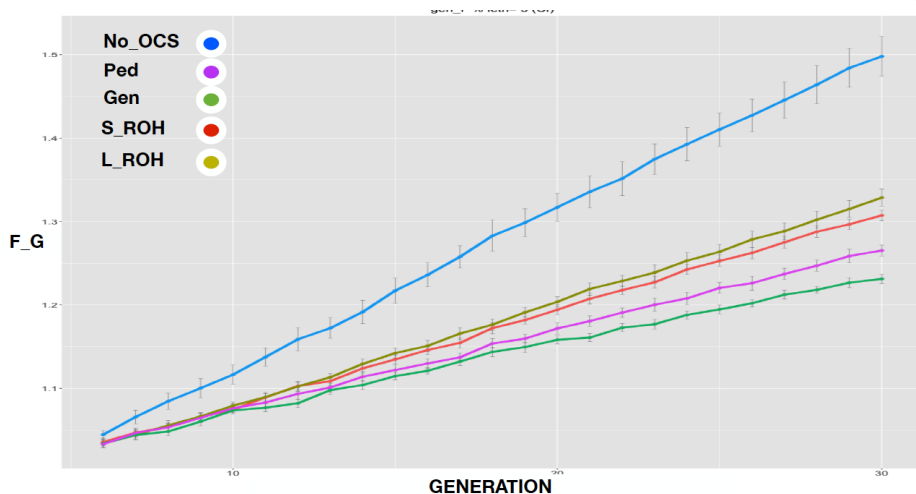
A Coalescent Forward in Time Simulation Toolkit

[Home](#) [About](#) [Parameters](#) [Output Files](#) [Examples](#)



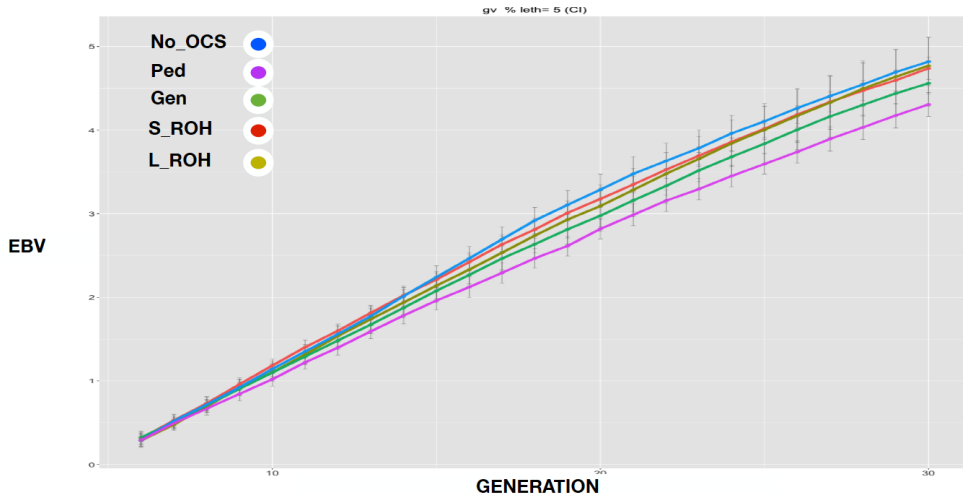
# Optimal Contribution Selection

ROH



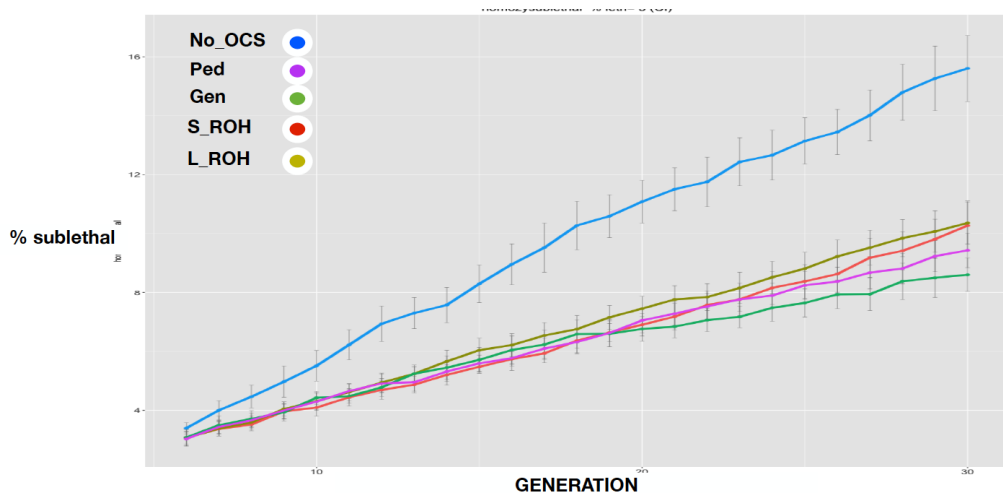
# Optimal Contribution Selection

ROH



# Optimal Contribution Selection

ROH



## There are many theoretically satisfying ideas...

- ...but nobody uses them
- Geneticists don't breed cows, farmers do
- Many cows now mated at random to a portfolio of bulls
- Everyone thinks their neighbor should use different bulls

### Some methods for avoiding inbreeding

Optimal contribution theory

Minimization of progeny inbreeding

Linear programming

Look-ahead mate selection

Selection against lethal alleles

Index selection including Mendelian sampling variance

Genomic selection including dominance

## A simple strategy for managing many recessive disorders in a dairy cattle breeding program

John B. Cole 

Haplotype	Functional/ <i>Gene</i> name	Chromosome	Location (Mbp)	Haplotype frequency (%)	Timing <sup>1</sup>
HBR	Black/red coat color/ <i>MC1R</i> ( <i>MSHR</i> )	18	14.71	0.75	—
HCD	Cholesterol deficiency/ <i>APOB</i>	11	77.87	2.28	W
HDR	Dominant red color/ <i>MC1R</i> ( <i>MSHR</i> )	3	9.36	0.03	—
HH0	Brachyspina/ <i>FANCI</i>	21	20.77	1.65	E,B
HH1	<i>APAF1</i>	5	62.81	1.28	E
HH2	—	1	93.50-95.58	1.21	E
HH3	<i>SMC2</i>	8	93.75	2.64	E
HH4	<i>GART</i>	1	1.99	0.23	E
HH5	<i>TFB1M</i>	9	91.85-91.94	2.39	E
HH6	<i>SDE2</i>	16	29.01-29.05	0.44	Old age
HHB	<i>BLAD/ITGB2</i>	1	144.77	0.21	W
HHC	<i>CVM/SLC35A3</i>	3	43.26	1.10	E,B
HHD	<i>DUMPS/UMPS</i>	1	69.15	0.01	E
HHM	Mulefoot/ <i>LRP4</i>	15	76.81	0.05	B
HHP	Polledness (dominant)/ <i>POLLED</i>	1	2.57	0.88	—
HHR	Red coat color/ <i>MC1R</i> ( <i>MSHR</i> )	18	14.71	3.29	—

<sup>1</sup>Timing of embryonic loss/calf death for homozygous animals: B = calf death at/shortly after birth, E = embryonic loss/abortion, W = calf death weeks/months after birth (Cole et al., 2016; Cole et al., 2018).



<sup>a</sup>Cole, 2019

- Large amount of research has been conducted on finding lethal haplotypes (VanRaden et al., 2011; Sahana et al., 2013; Hoff et al. 2017).
- Within livestock populations undergoing selection lethals mutations are effectively removed from the population, while sub-lethal mutations are difficult to remove.
- What about sub-lethal mutations (i.e. affected animal **doesn't die**, but instead has **reduced performance**)?

## Mate Two Carriers

	D	d
D	DD	Dd
d	Dd	<b>dd</b>

Appears in  
genotype data  
as an **ROH**





## Use heuristics to "tag" unfavorable ROH genotype

[illegible]



## Inbreeding Load Matrix (ILM)

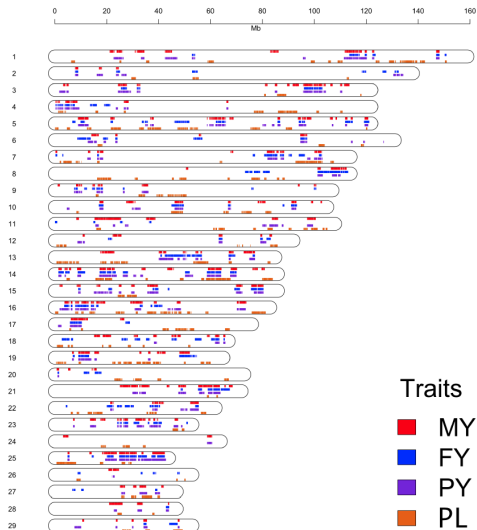
- Summarize effects of unfavorable haplotypes identified.
- Across two individuals:

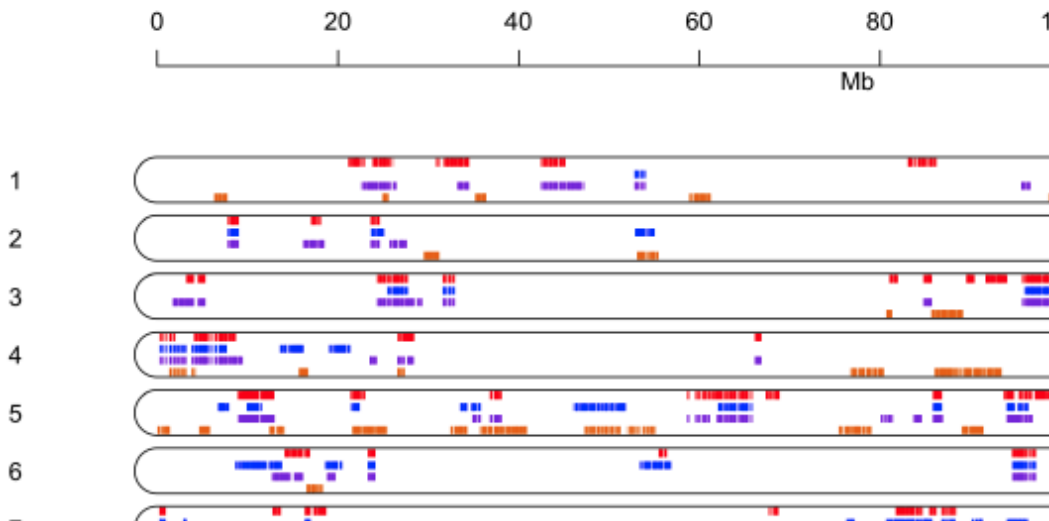


- Within an individual:



$$\begin{pmatrix} \sum_{i=0}^{i=\text{Haplotypes}} \left( \frac{1}{4} * (X_1X_1 + X_1X_2 + X_2X_1 + X_2X_2) * \beta_i \right) & \sum_{i=0}^{i=\text{Haplotypes}} \left( \frac{1}{4} * (X_1Y_1 + X_1Y_2 + X_2Y_1 + X_2Y_2) * \beta_i \right) & \vdots \\ \vdots & \vdots & \vdots \end{pmatrix}$$





Trait	-10logp Pedigree	-10logp Genomic	-10logp Diag. ILM
Milk	4.95	18.06	27.45
Fat	4.67	9.96	16.56
Pro	2.18	13.47	21.43
PL	0.33	1.5	4.75
DPR	0.57	0.08	1.21
SCS	0.11	0.14	0.58

- A mate allocation matrix ( $B$ ; Pryce et al. 2012) constructed:

$$B_{ij} = \lambda_1 * EBV_{ij} - \lambda_2 * F_{ij} - \lambda_3 * ILM_{ij},$$

- Can be constructed based on lethal and unfavorable haplotypes/mutations.
- Penalize breeding value based on inbreeding metric and functional inbreeding.

	Dam1	Dam2	Dam3	Dam4	Dam5
Sire 1	$B_{\text{Sire 1, Dam 1}}$	$B_{\text{Sire 1, Dam 2}}$	$B_{\text{Sire 1, Dam 3}}$	$B_{\text{Sire 1, Dam 4}}$	$B_{\text{Sire 1, Dam 5}}$
Sire 2	$B_{\text{Sire 2, Dam 1}}$	$B_{\text{Sire 2, Dam 2}}$	$B_{\text{Sire 2, Dam 3}}$	$B_{\text{Sire 2, Dam 4}}$	$B_{\text{Sire 2, Dam 5}}$
Sire 3	$B_{\text{Sire 3, Dam 1}}$	$B_{\text{Sire 3, Dam 2}}$	$B_{\text{Sire 3, Dam 3}}$	$B_{\text{Sire 3, Dam 4}}$	$B_{\text{Sire 3, Dam 5}}$
Sire 4	$B_{\text{Sire 4, Dam 1}}$	$B_{\text{Sire 4, Dam 2}}$	$B_{\text{Sire 4, Dam 3}}$	$B_{\text{Sire 4, Dam 4}}$	$B_{\text{Sire 4, Dam 5}}$
Sire 5	$B_{\text{Sire 5, Dam 1}}$	$B_{\text{Sire 5, Dam 2}}$	$B_{\text{Sire 5, Dam 3}}$	$B_{\text{Sire 5, Dam 4}}$	$B_{\text{Sire 5, Dam 5}}$

## Characterizing Age of Inbreeding

# Identify Homozygous by descent

Characterizing HBD segments and age related autozygosity



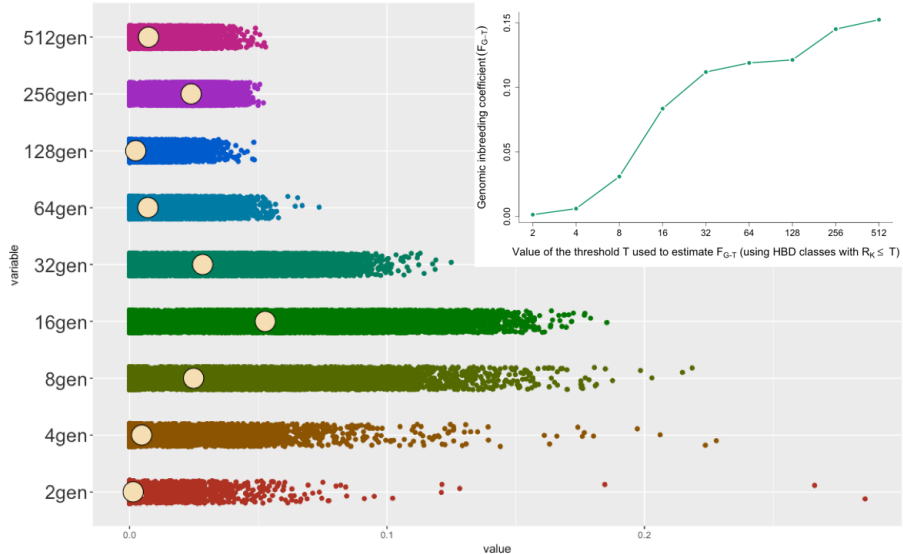
- ▶ HBD happens when individuals inherit copies of ancestral chromosome
- ▶ Copies are inherited through different patterns
- ▶ The length of HBD depend on the number of generations and the population structure
- ▶ Inbreeding can then be divided into different age classes and these related to the overall depression load based on their age<sup>8</sup>

---

<sup>8</sup>Druet, et al. 2017

# HBD

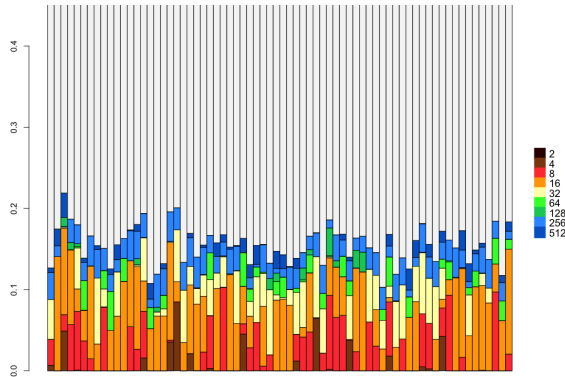
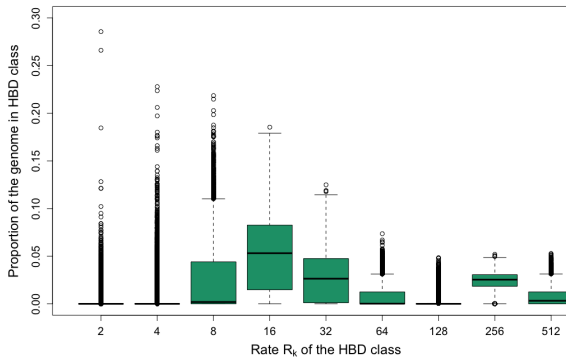
## Characterizing the population through HBD





# HBD

Characterizing the population through HBD



# HBD

Characterizing the population through HBD



Trait	Pedigree Regression	Genomic Regression
Milk	-91.1	-108.2
Fat	-3.63	-3.58
Pro	-1.81	-2.86
PL	-0.56	-0.85
DPR	-0.12	-0.02
SCS	-0.08	~0

HBD	Trait	Regression
2-8 Gen	Milk	-138.1
18-128 Gen		-88.3
>128 Gen		-185.3
2-8 Gen	Fat	-4.44
18-128 Gen		-3.82
>128 Gen		-5.12
2-8 Gen	Pro	-3.38
18-128 Gen		-2.60
>128 Gen		-4.43
2-8 Gen	PL	-0.83
18-128 Gen		-0.32
>128 Gen		0.08
2-8 Gen	DPR	-0.14
18-128 Gen		-0.04
>128 Gen		0.02

## Near future

- ▶ Characterization annotation of haplotypes
- ▶ Refining HBD structure

## Mid future

- ▶ Validation of haplotypes in larger populations (across breeds)
- ▶ Implement ILM for validated haplotypes






$$EBV_{c0} = EBV - \lambda_0 F_P$$

$$EBV_{c1} = EBV - \lambda_1 F_G$$

$$EBV_{c2} = EBV - \lambda_1 F_{G_{hbd_1}} \dots - \lambda_n F_{G_{hbd_n}}$$

$$EBV_{c3} = EBV - \lambda_1 F_{G_{hbd_1}} \dots - \lambda_n F_{G_{hbd_n}} - \lambda_{n+1} F_{ILM}$$

Future Direction

- Optimal contribution 
- Minimization of progeny inbreeding 
- Selection against lethal alleles 
- Index selection including Mendelian sampling variance 
- Genomic selection including dominance 

# Acknowledgments

- ▶ Christine Baes
- ▶ John Cole
- ▶ Jennie Pryce
- ▶ Francesco Tiezzi
- ▶ Matteo Bergamaschi
- ▶ Jeremy Howard
- ▶ Emmanuel Lozada

