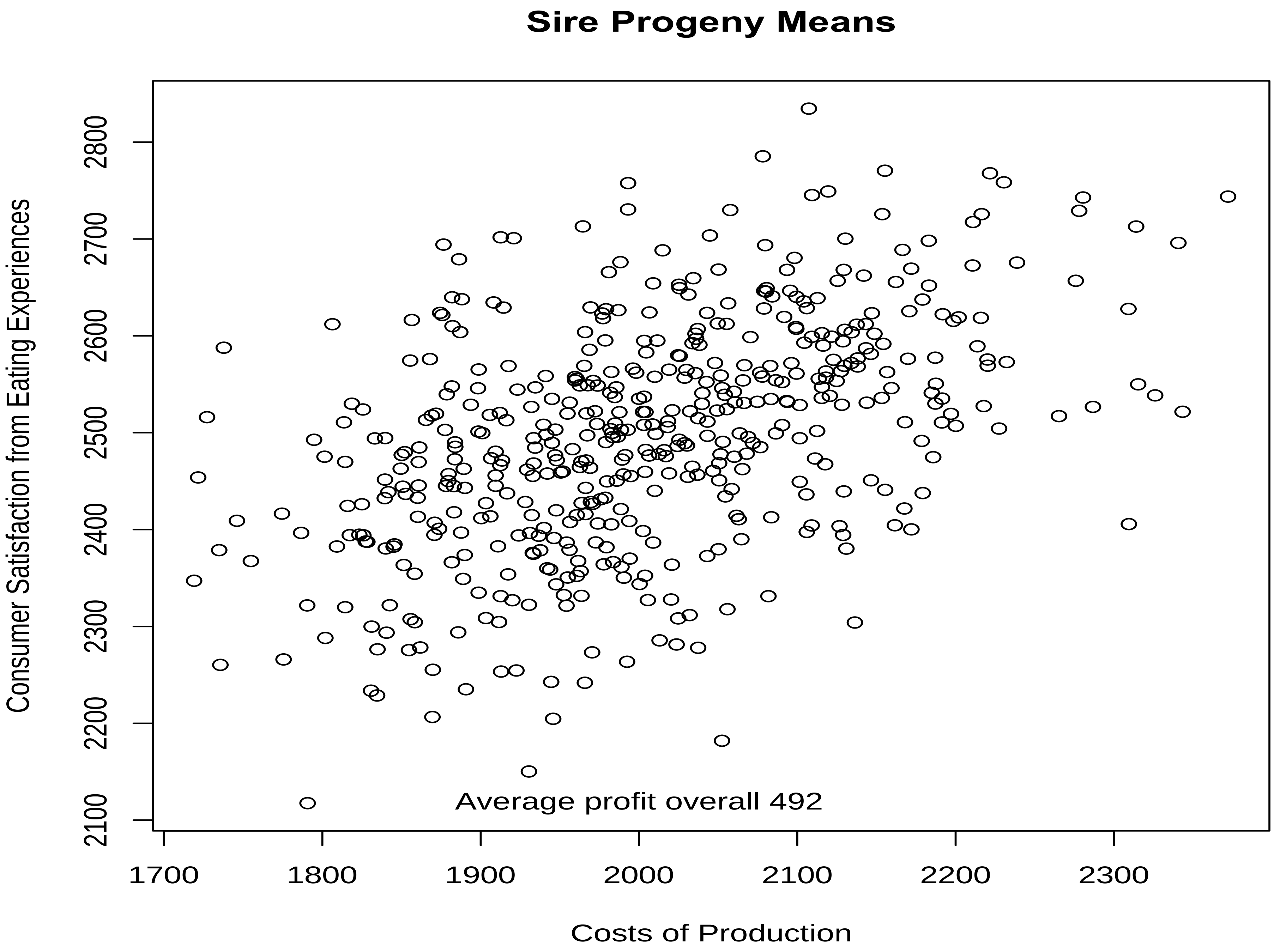
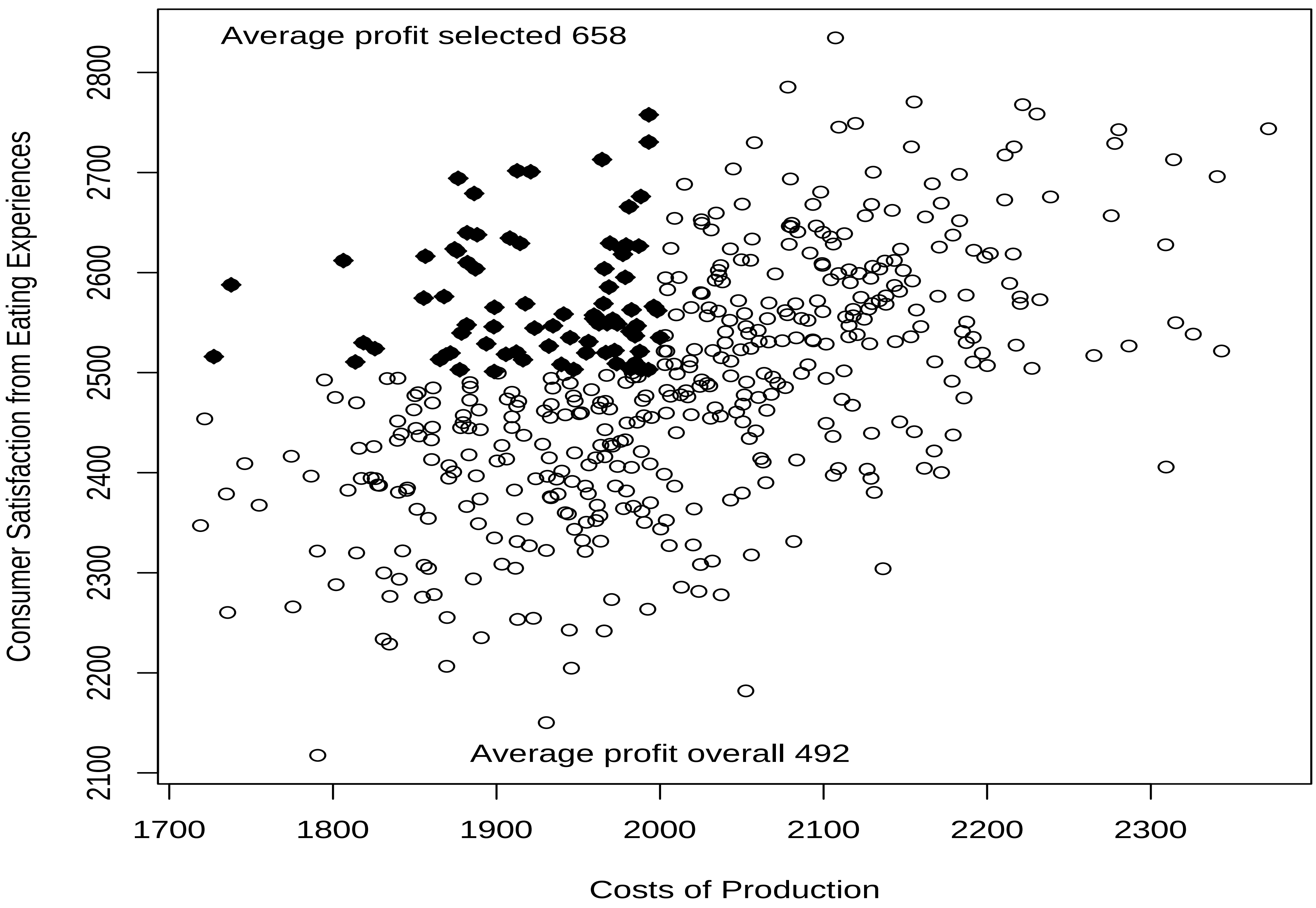
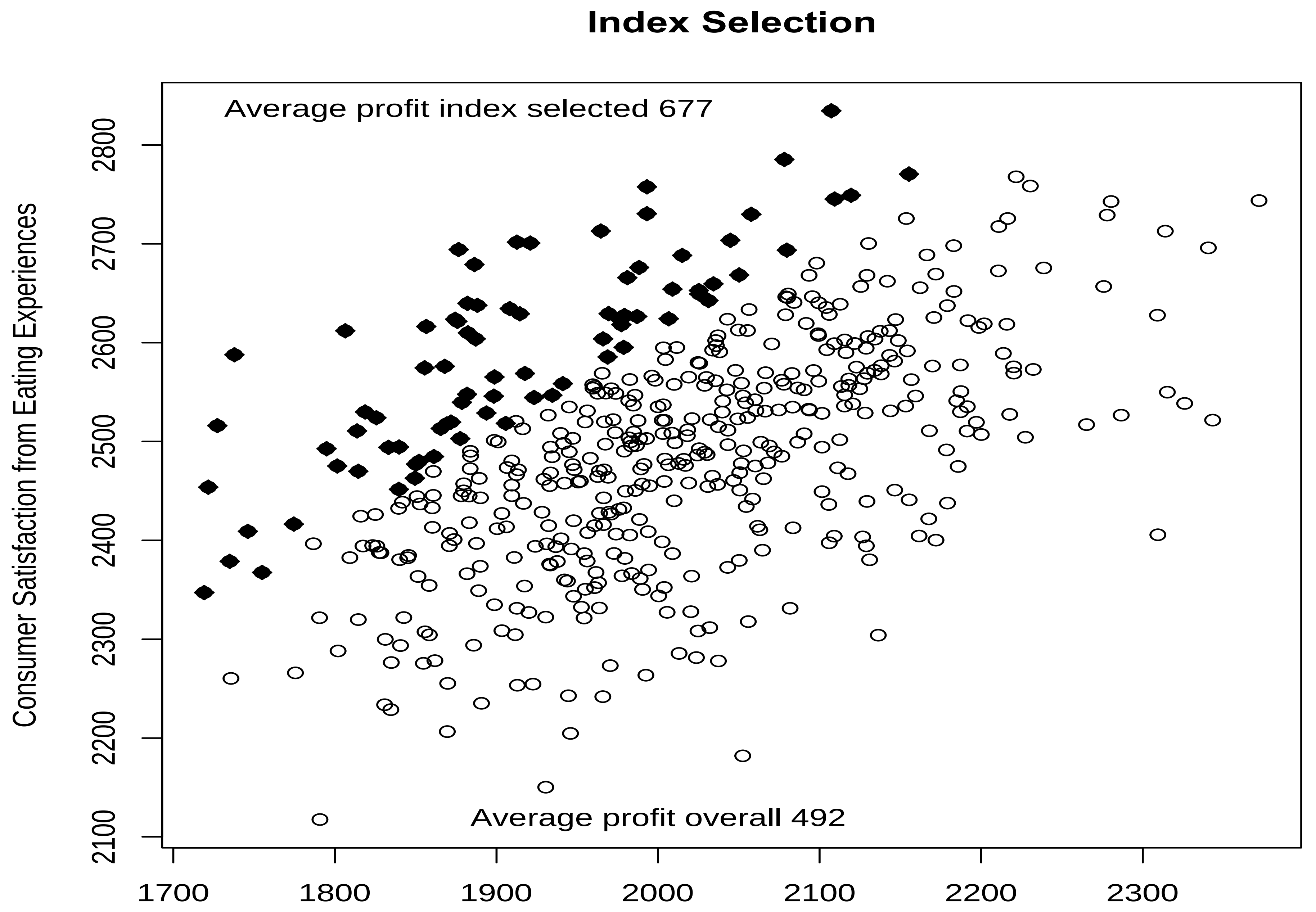
An International Perspective on Genomics

Dorian Garrick
dorian@iastate.edu

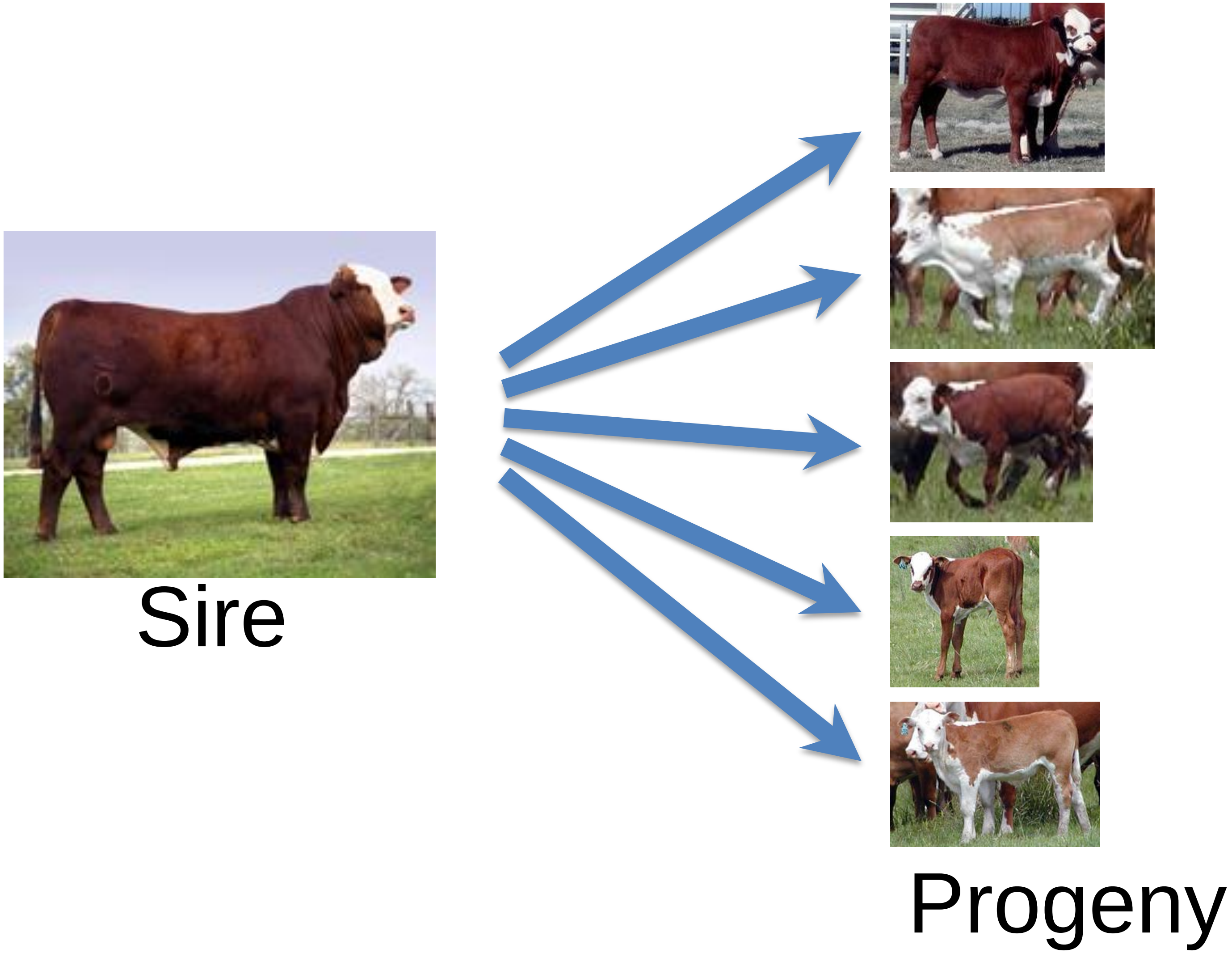

Sire Progeny Means
Consumer Satisfaction from Eating Experiences
2100
2200
2300
2400
2500
2600
2700
2800
1700
1800
1900
2000
2100
2200
2300
Costs of Production
Average profit overall 492

Independent Culling Levels

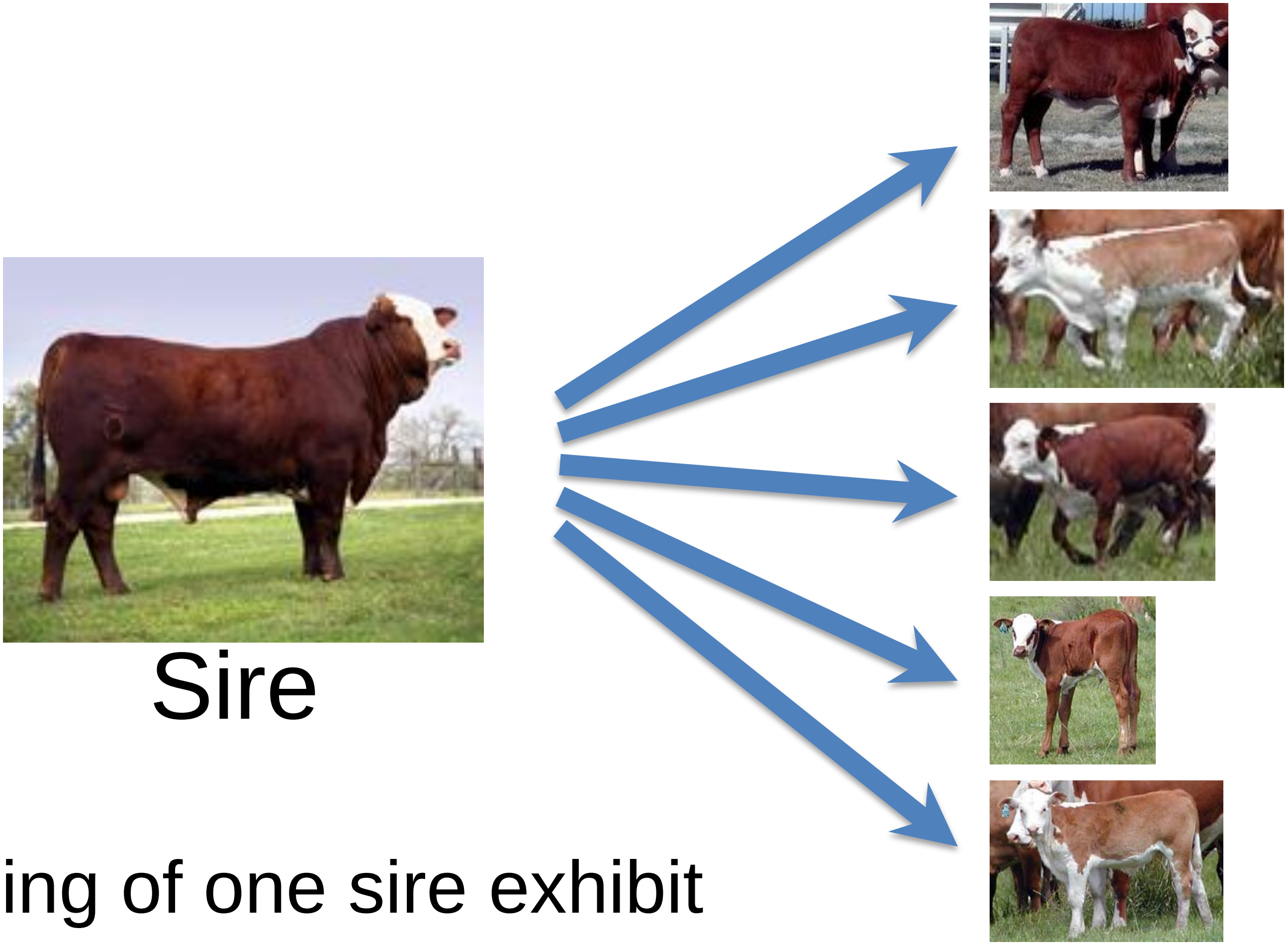



Index Selection
Average profit index selected 677
Average profit overall 492
Consumer Satisfaction from Eating Experiences
2100
2200
2300
2400
2500
2600
2700
2800
Costs of Production
1700
1800
1900
2000
2100
2200
2300

Suppose we generate 100 progeny on 1 bull



Performance of the Progeny



Sire

+30 kg

+15 kg

-10 kg

+ 5 kg

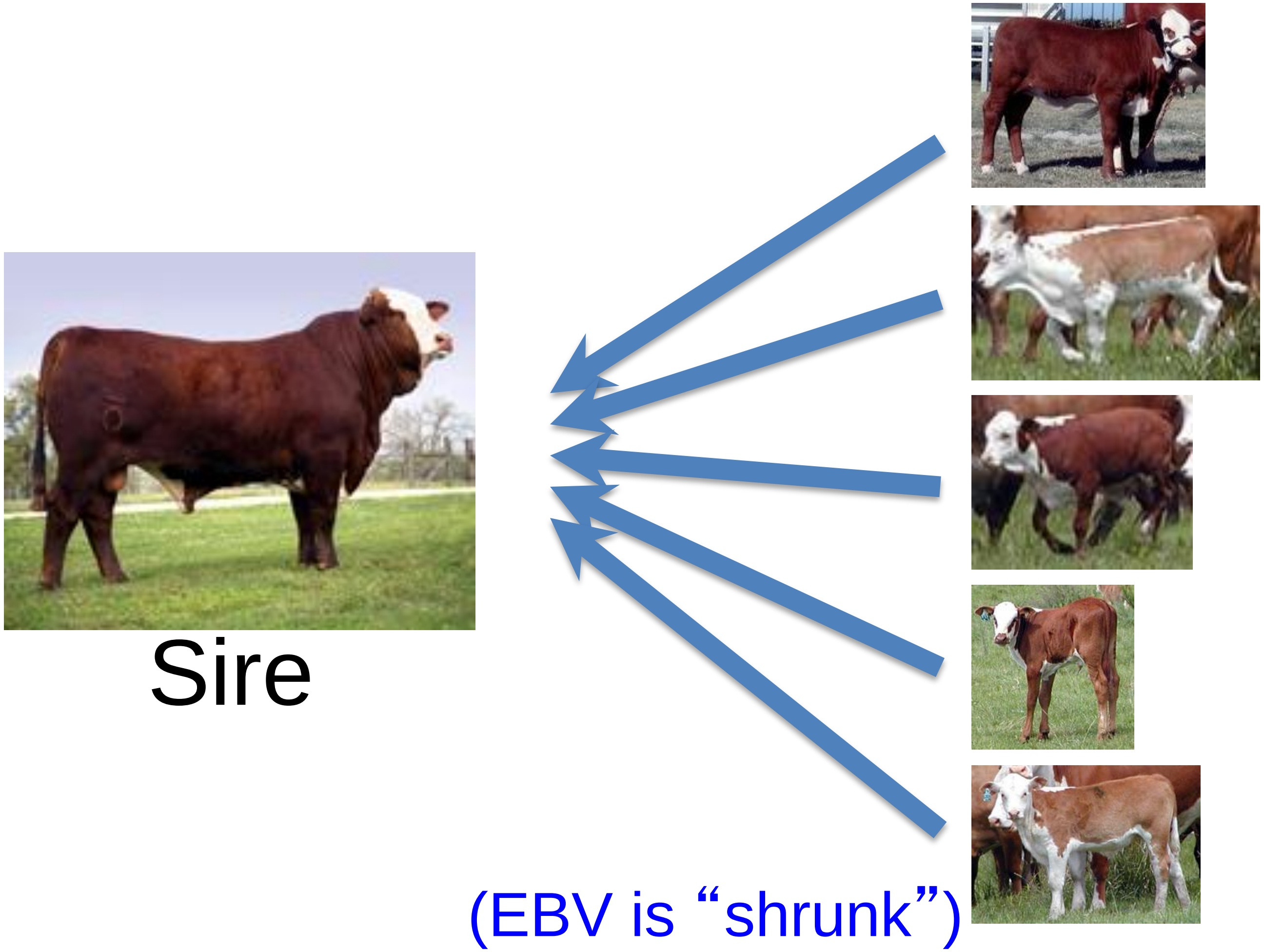
+10 kg

+10 kg

Progeny

Offspring of one sire exhibit more than ¾ diversity of the entire population

We learn about parents from progeny



Sire

+30 kg

+15 kg

-10 kg

+ 5 kg

+10 kg

(EBV is “shrunk”)

Sire EBV +16-18 kg

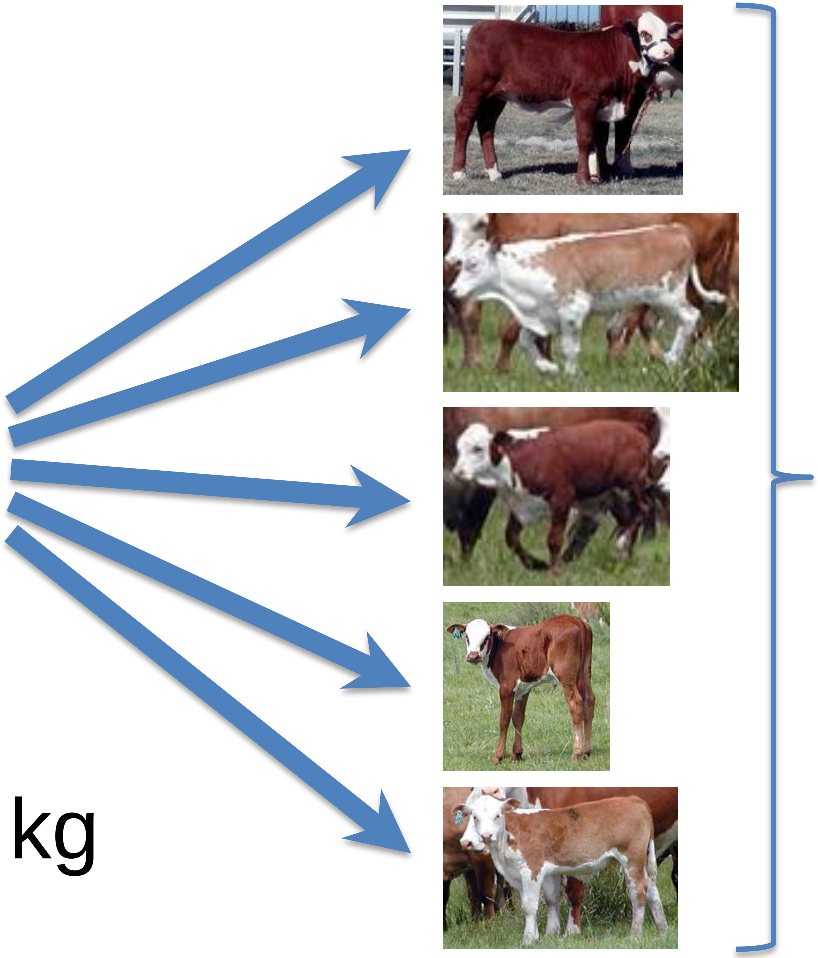
Progeny +10 kg

<2x progeny difference

Suppose we generate new progeny

Sire

Sire EBV +16-18 kg



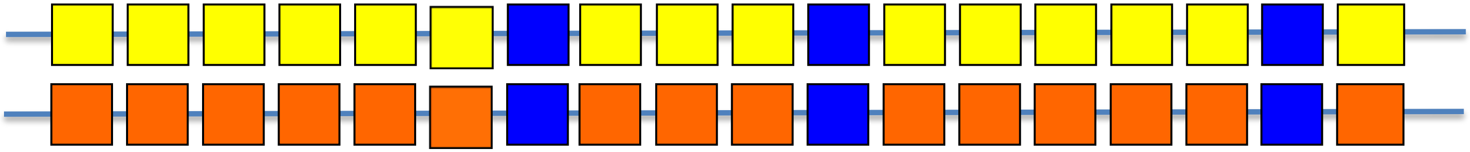
Expect them to be 8-9 kg heavier than those from an average sire

Some will be more others will be less but we cant tell which are better without "buying" more information

Progeny

Chromosomes are a sequence of base pairs

Part of 1 pair
of chromosomes



Paternal

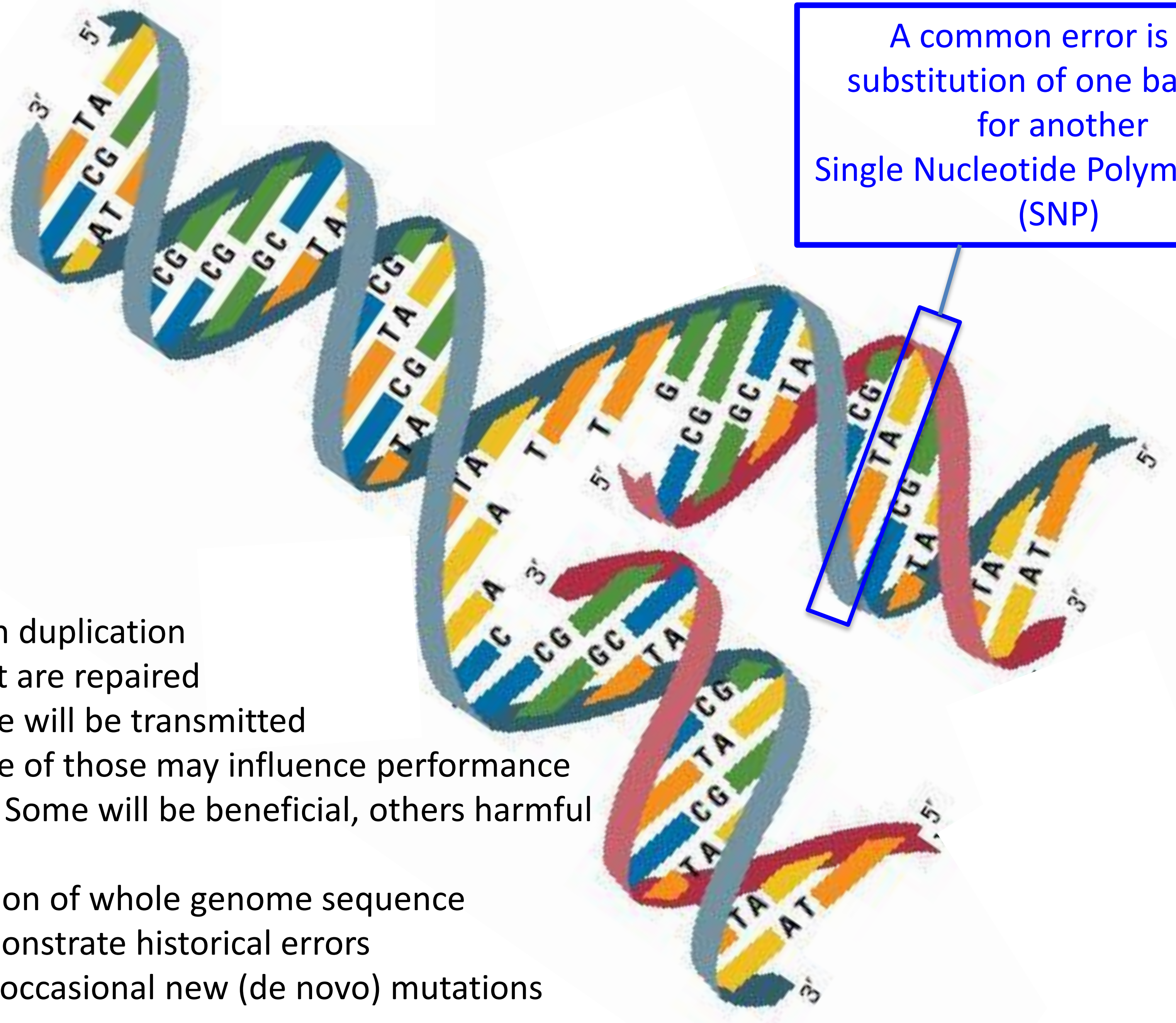
Maternal

Cattle usually have 30 pairs of chromosomes
One member of each pair was inherited from the sire, one from the dam
Each chromosome has about 100 million base pairs (A, G, T or C)
About 3 billion describe the animal

🟦 Blue base pairs represent genes/exons

🟨 Yellow represents the strand inherited from the sire

🟧 Orange represents the strand inherited from the dam



A common error is the substitution of one base pair for another
Single Nucleotide Polymorphism (SNP)

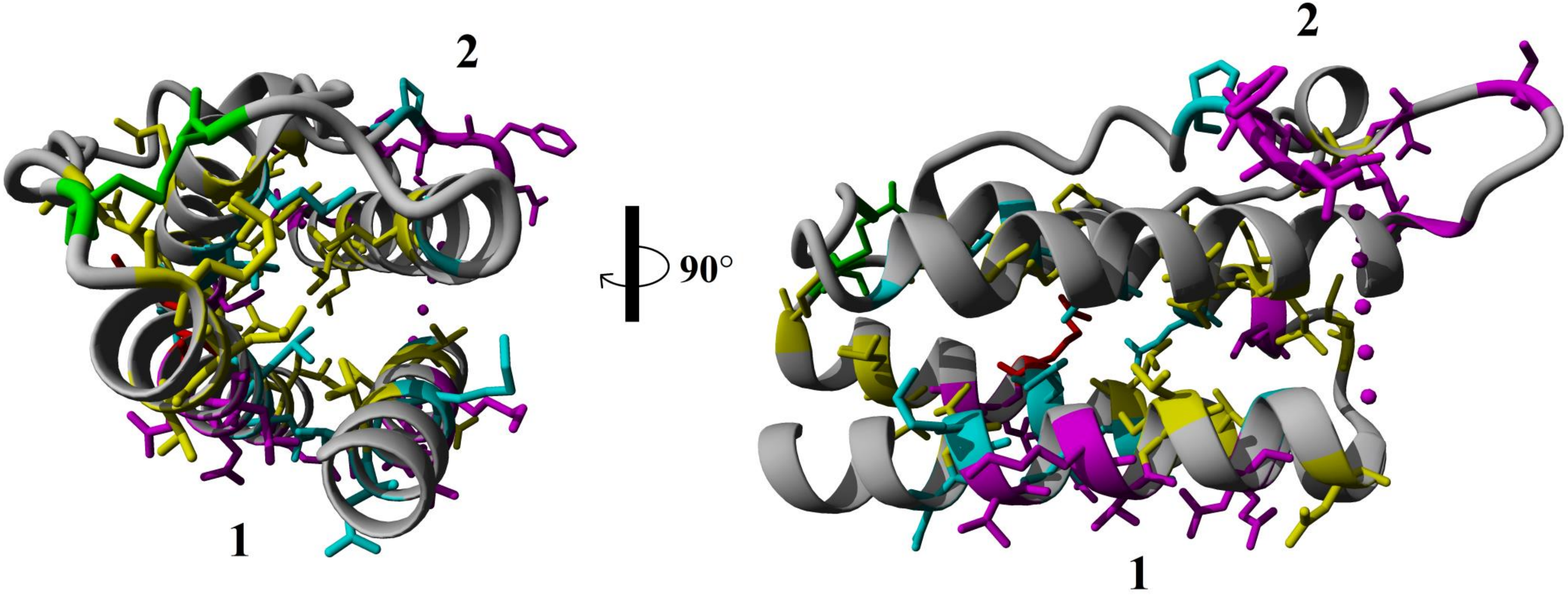
Errors in duplication

- Most are repaired
- Some will be transmitted
- Some of those may influence performance
  - Some will be beneficial, others harmful

Inspection of whole genome sequence

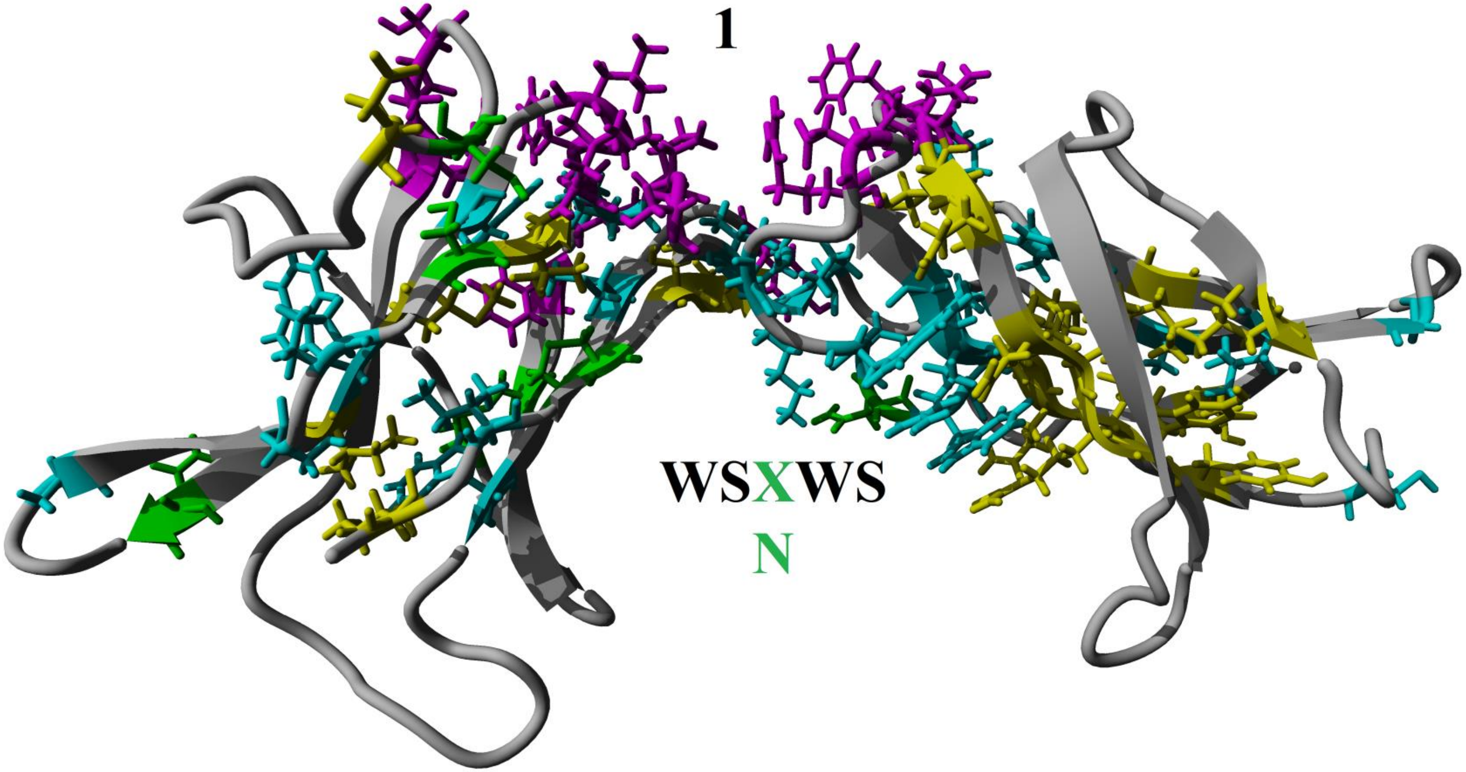
- Demonstrate historical errors
- And occasional new (de novo) mutations

Leptin



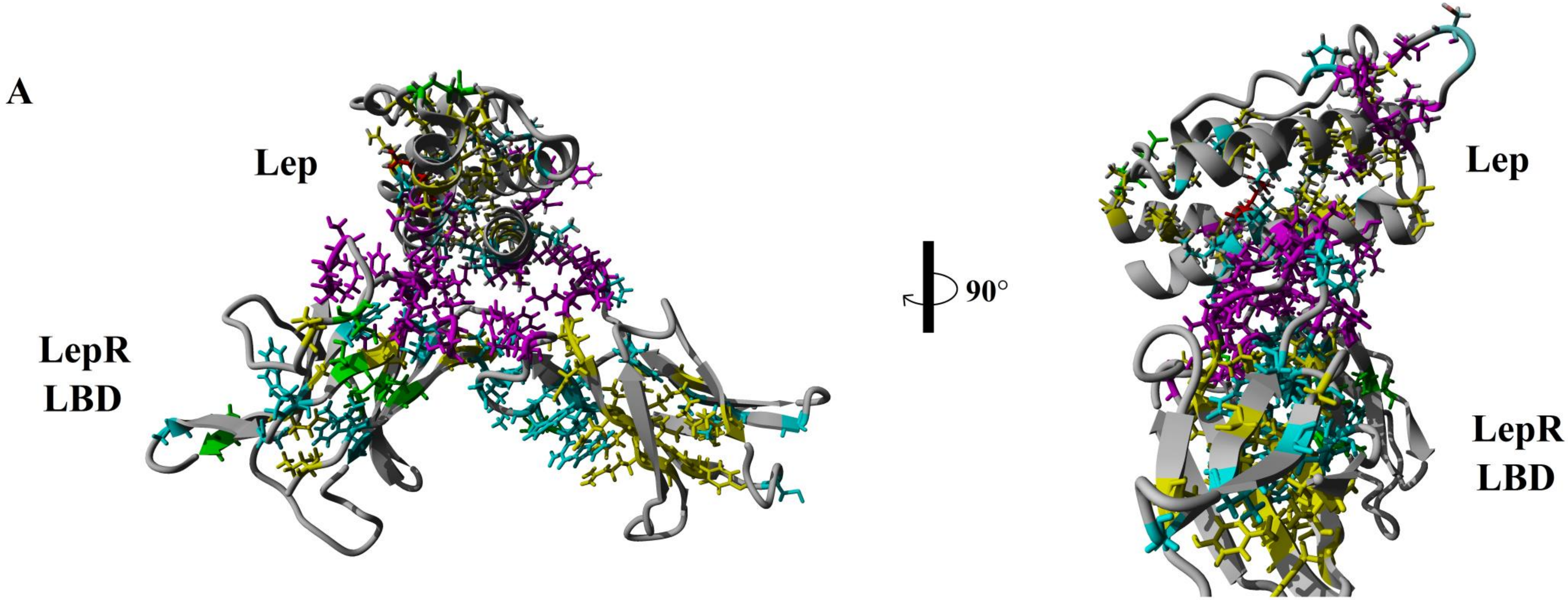
Prokop et al, Peptides, 2012

Leptin Receptor



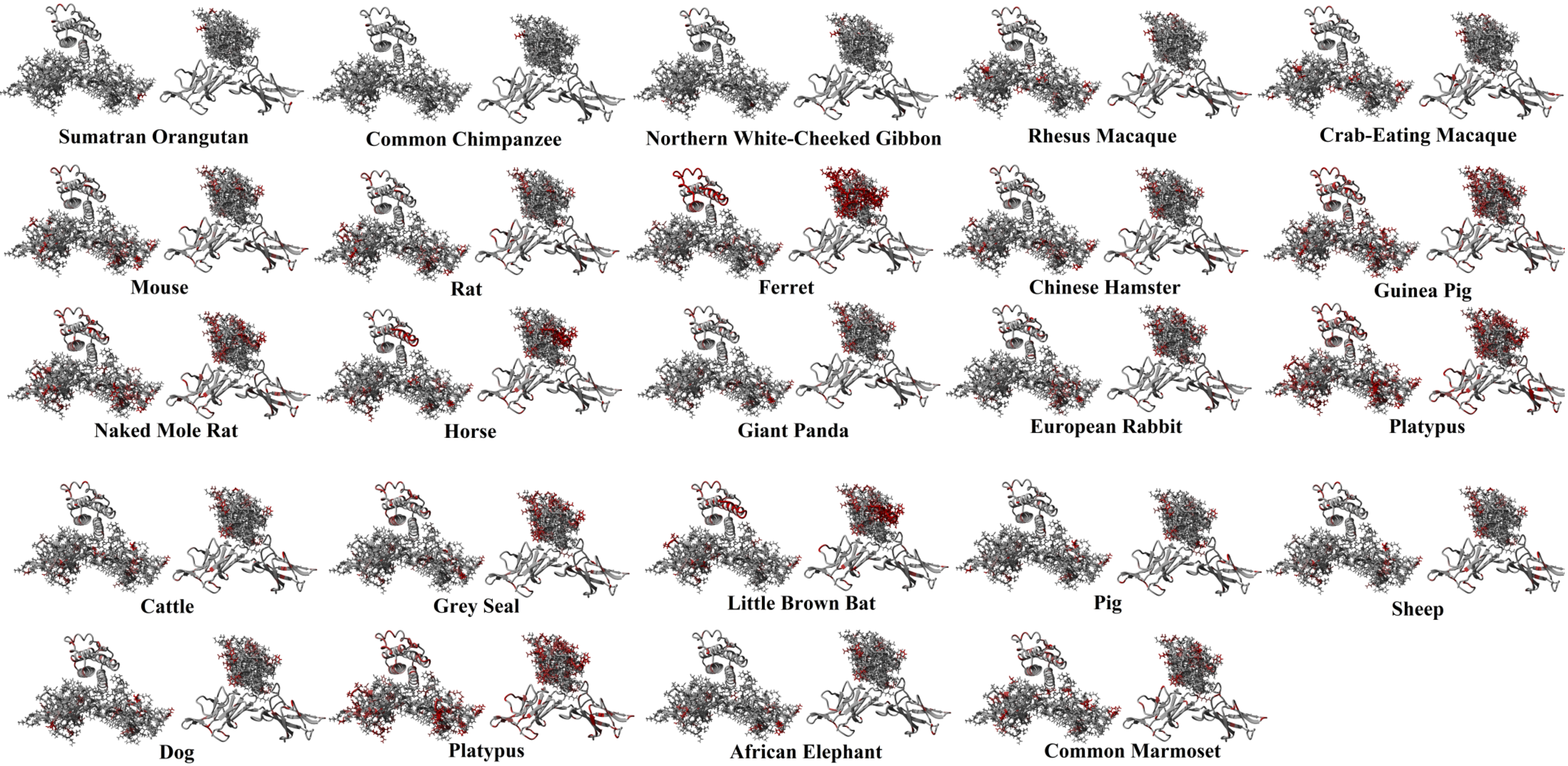
Prokop et al, Peptides, 2012

Joining the two



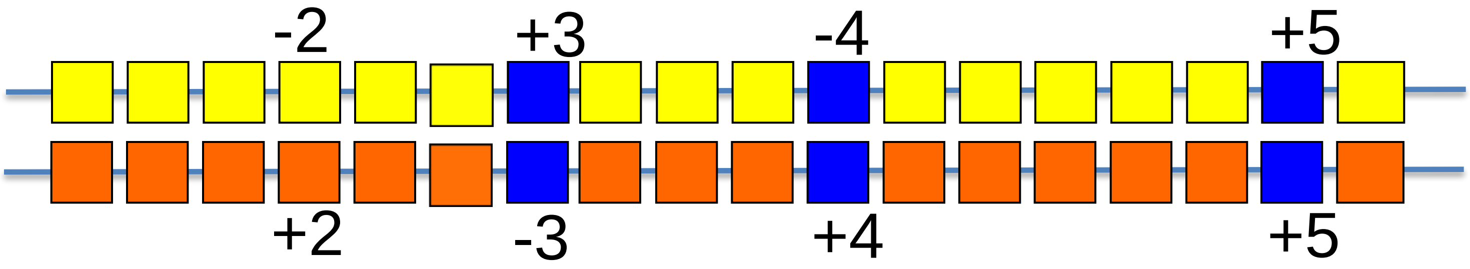
Prokop et al, Peptides, 2012

Leptin and its Receptor Across Species



Prokop et al, Peptides, 2012

Breeding Merit is sum of average gene effects



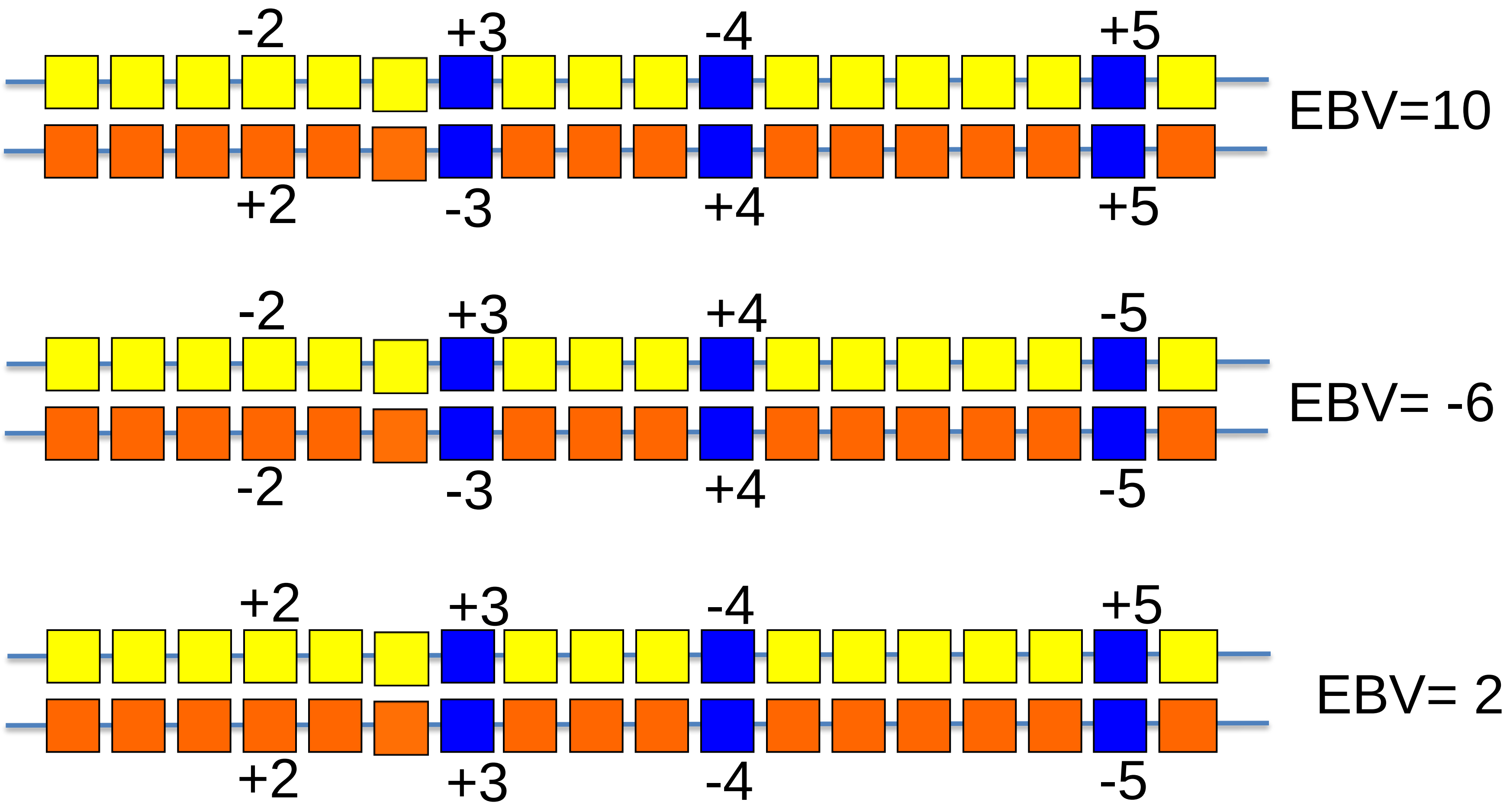
Sum=+2

Sum=+8

EBV=10

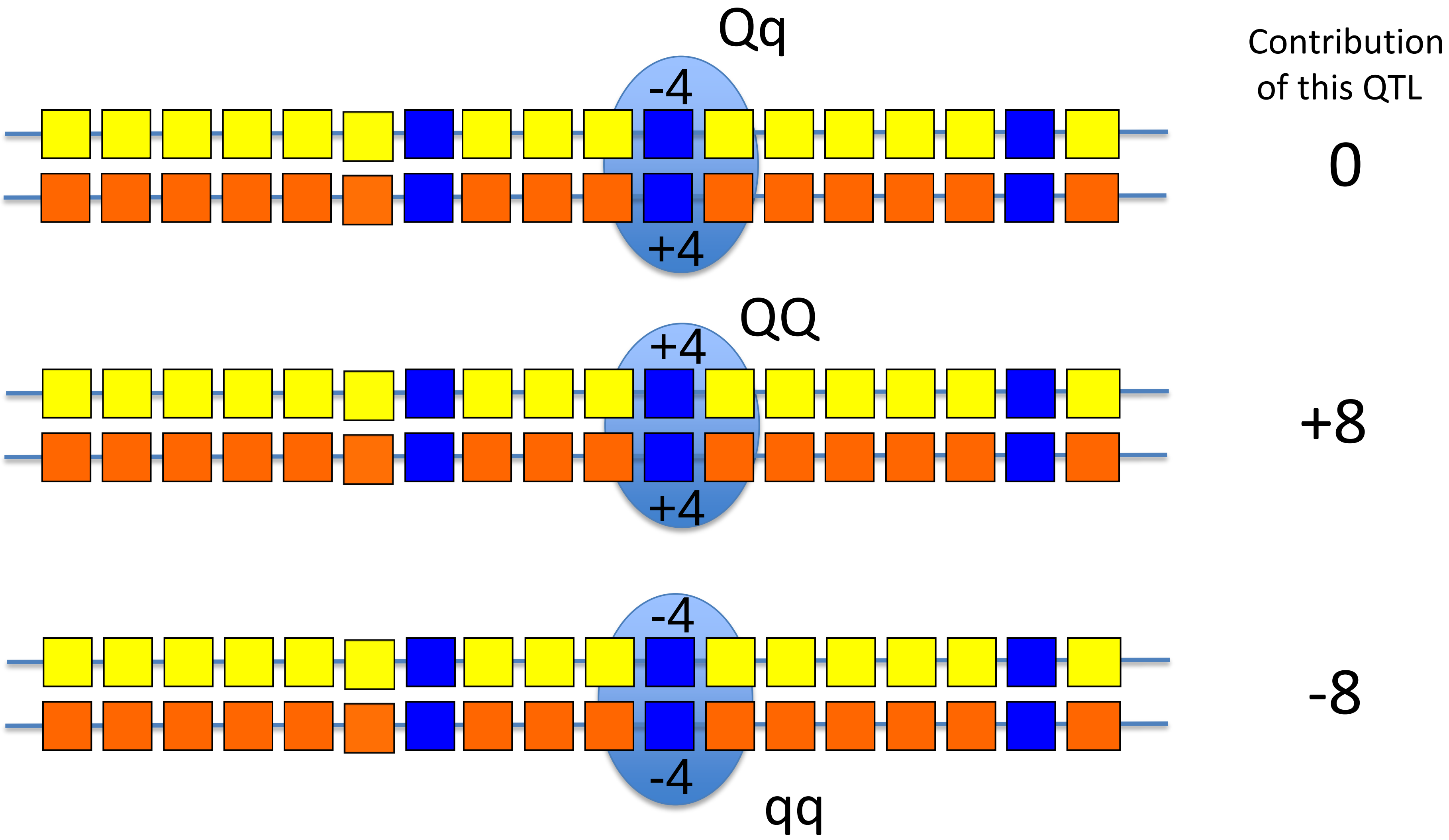
🟦 Blue base pairs represent genes/exons

Consider 3 Bulls

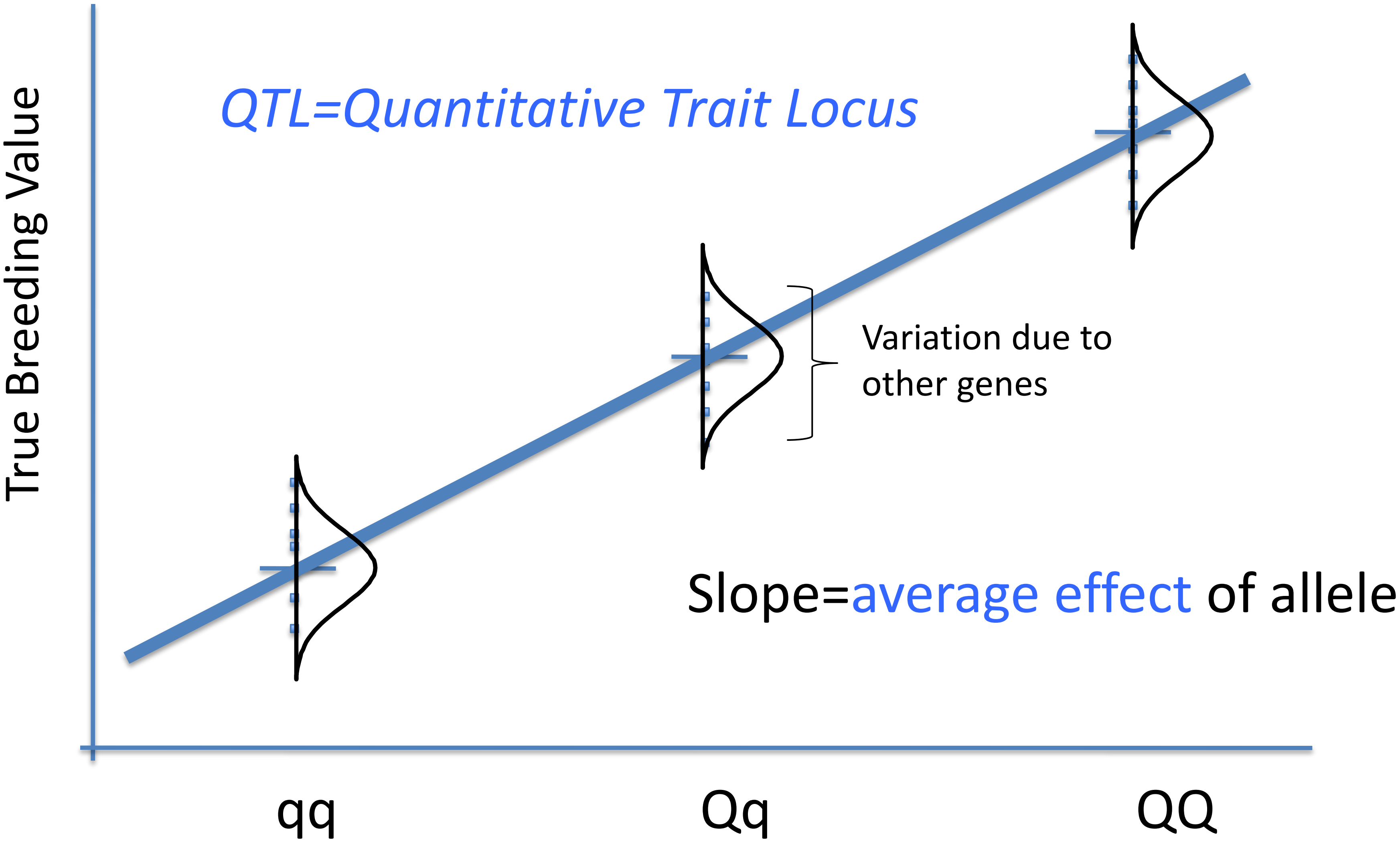


Below-average bulls will have some above-average alleles and vice versa!

At any 1 locus there are 3 genotypes



Regress BV on QTL genotype

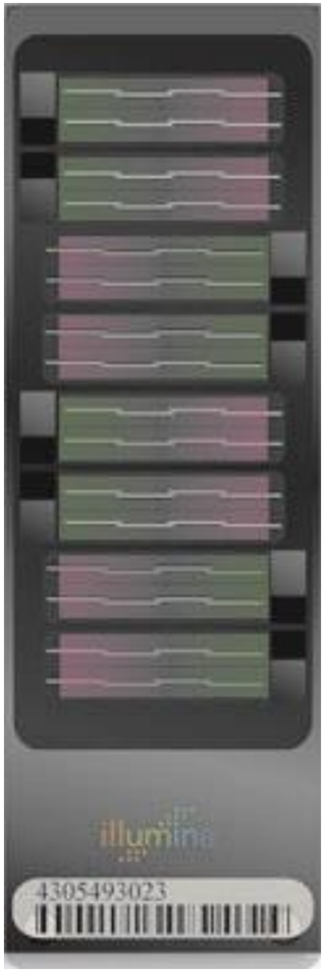


QTL=Quantitative Trait Locus

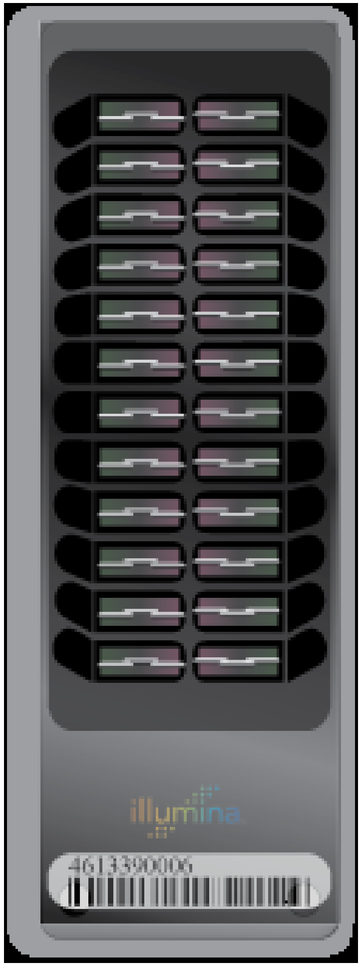
Variation due to other genes

Slope=average effect of allele

Illumina Bovine 770k, 50k (v2), 3k



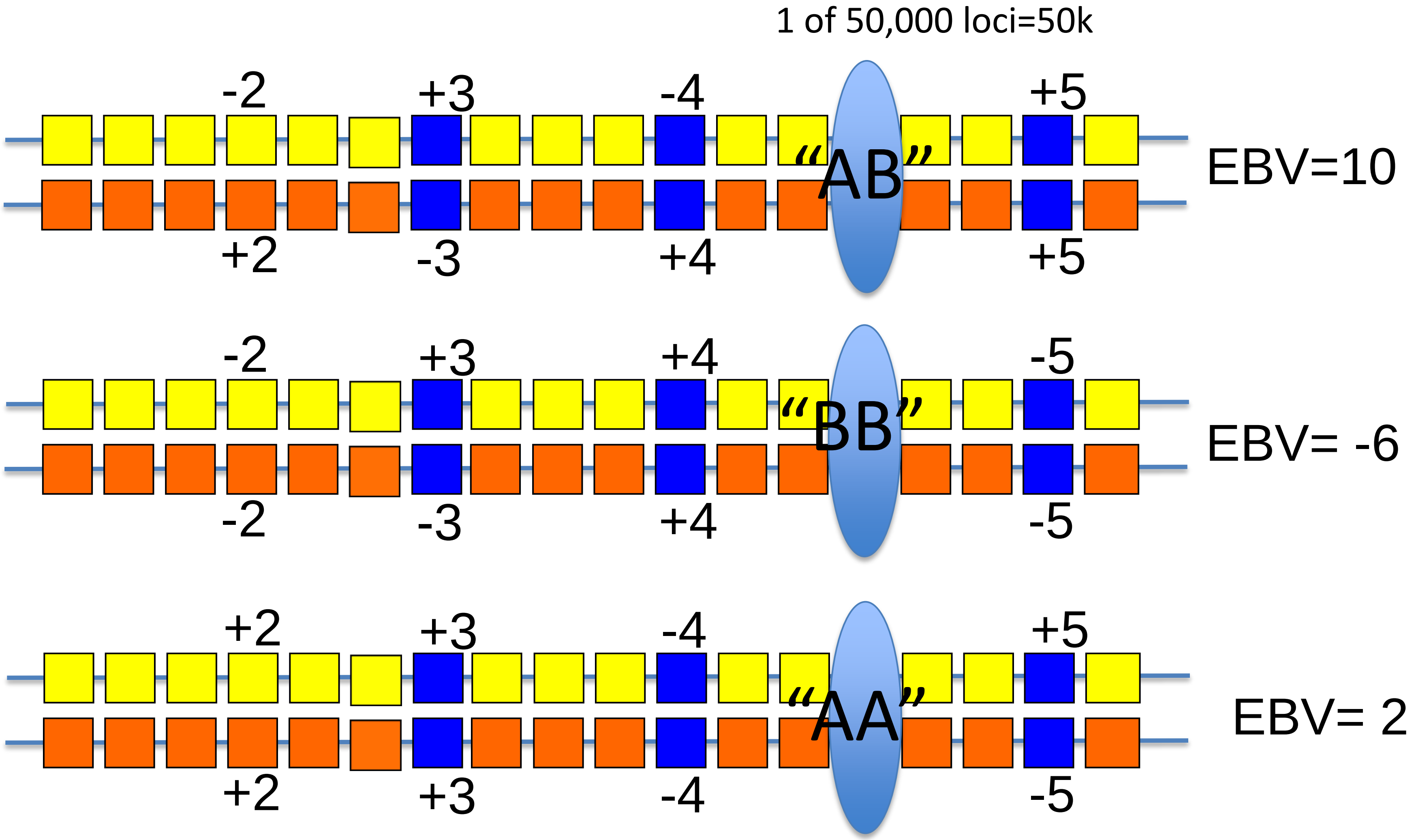
700k (HD)



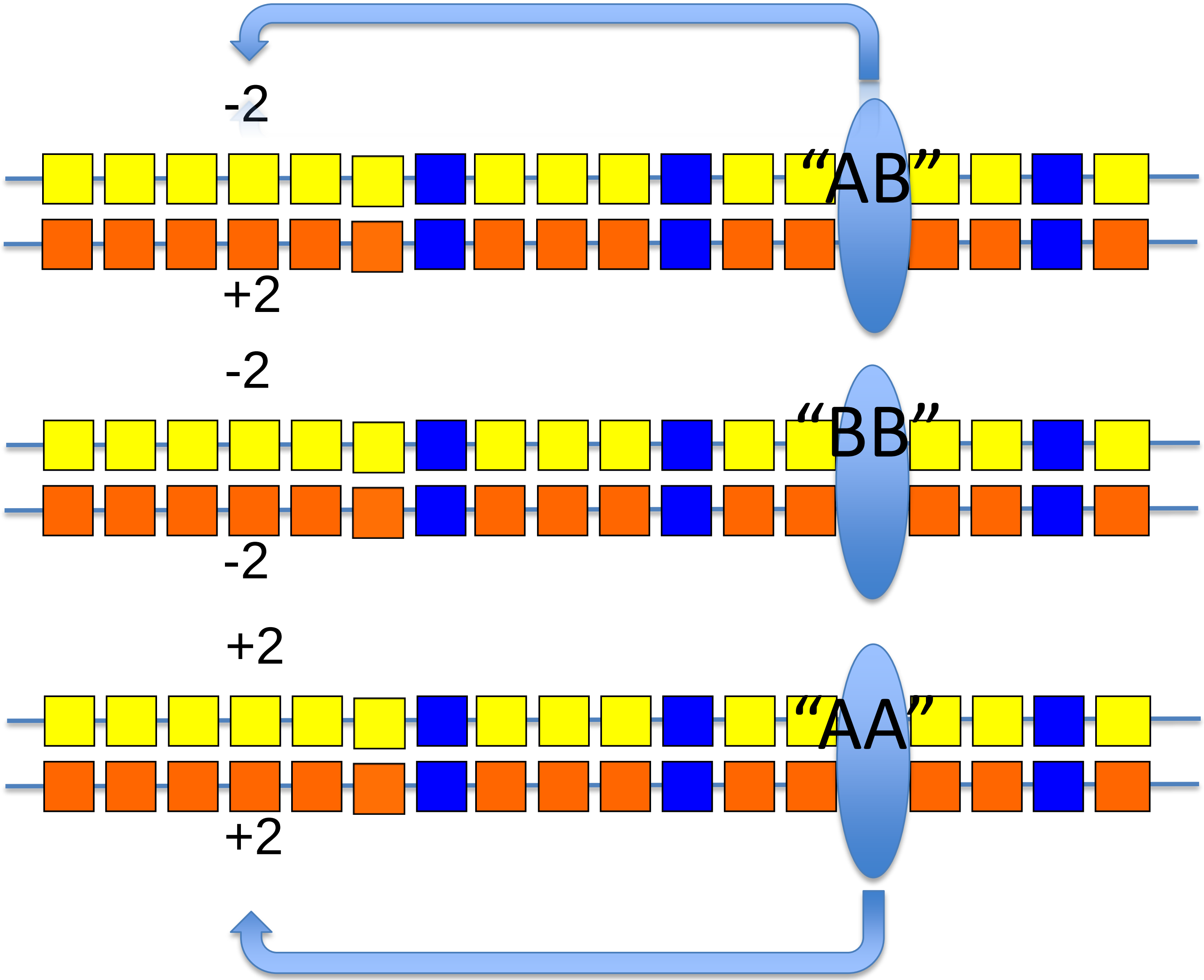
50k (Several versions)

3k (LD)

SNP Genotyping the Bulls

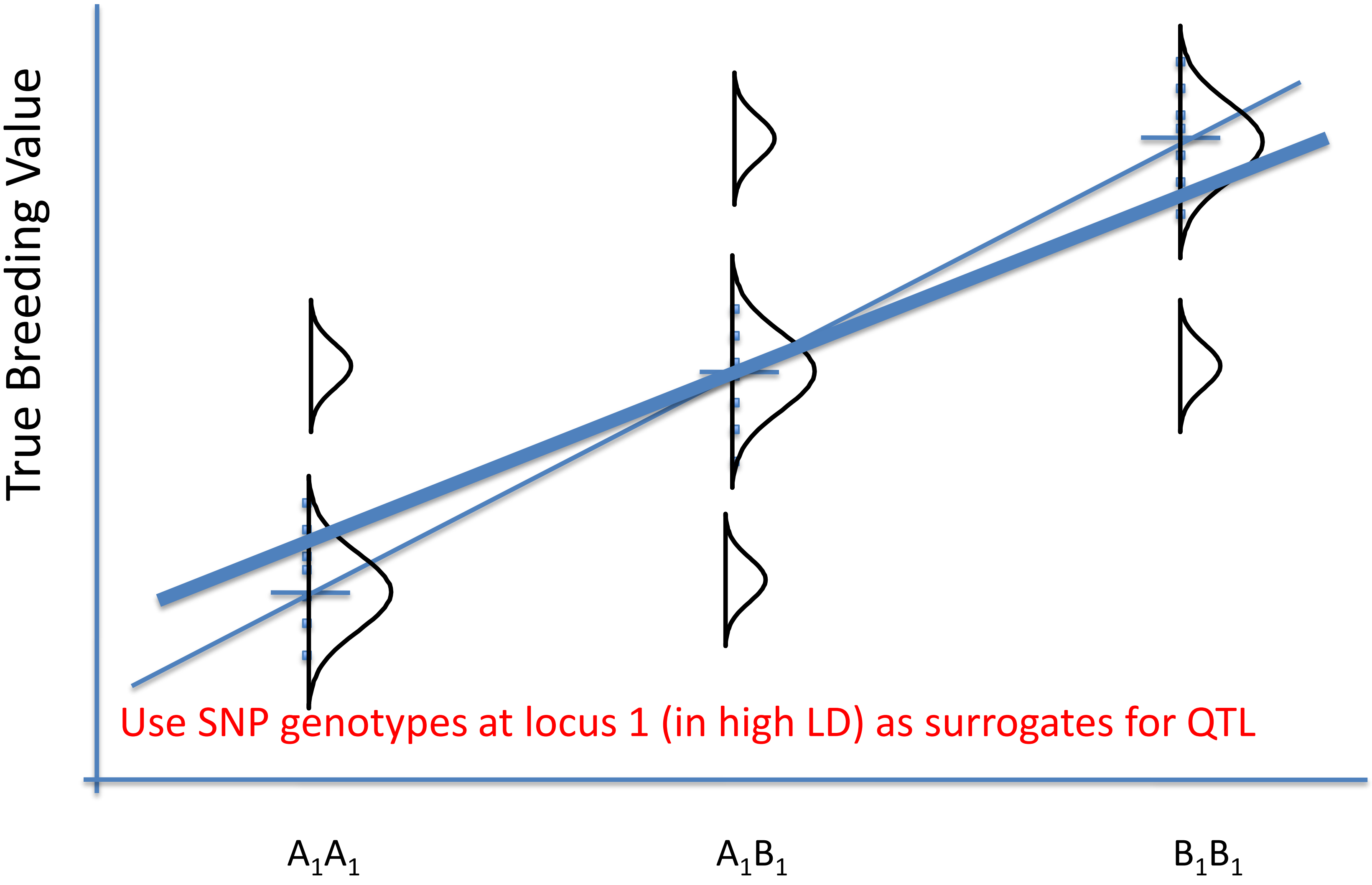


Linkage Disequilibrium (LD)

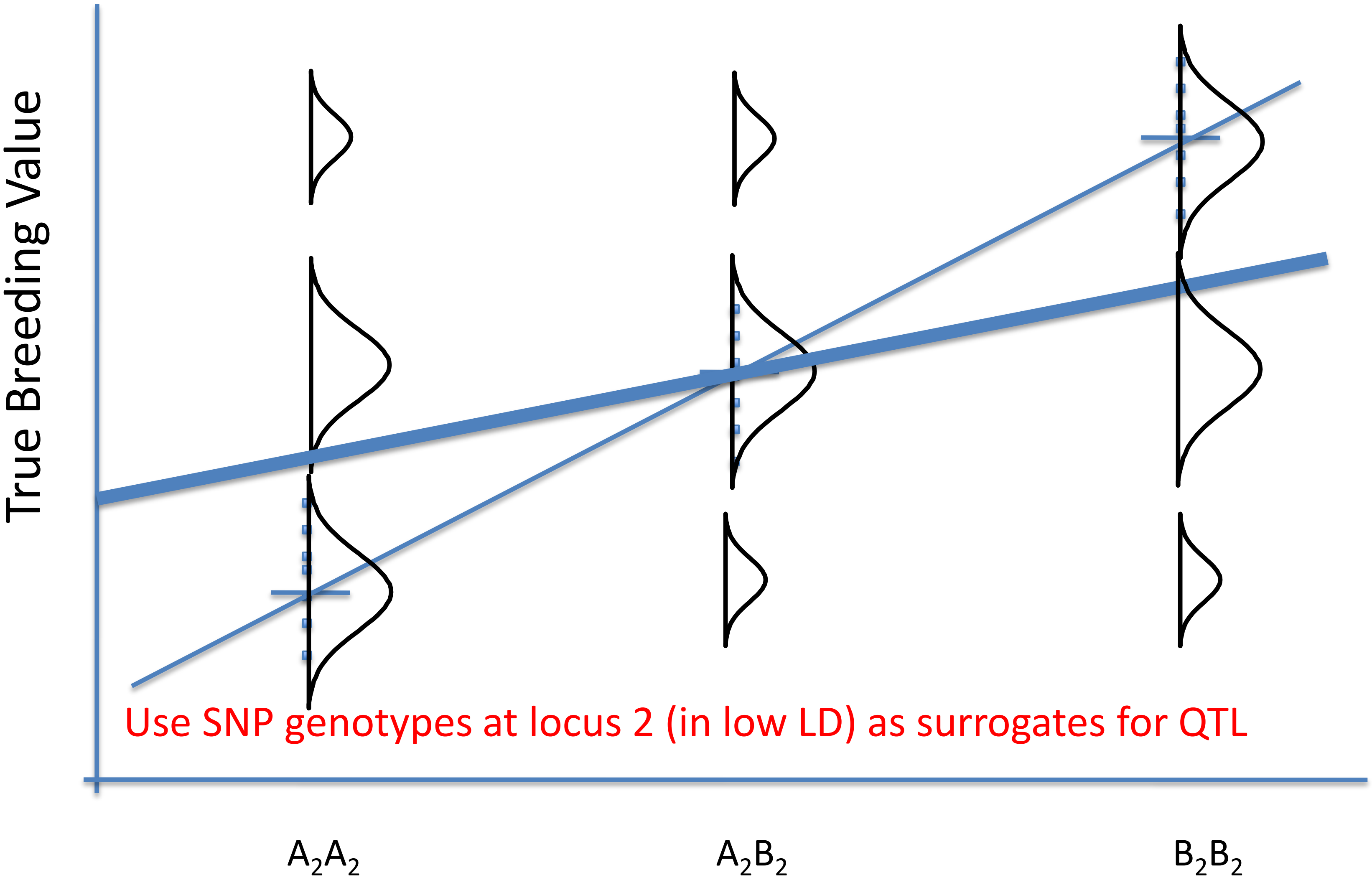


LD occurs when genotypes at one locus are predictive of genotypes at another

Practice – EBV on SNP



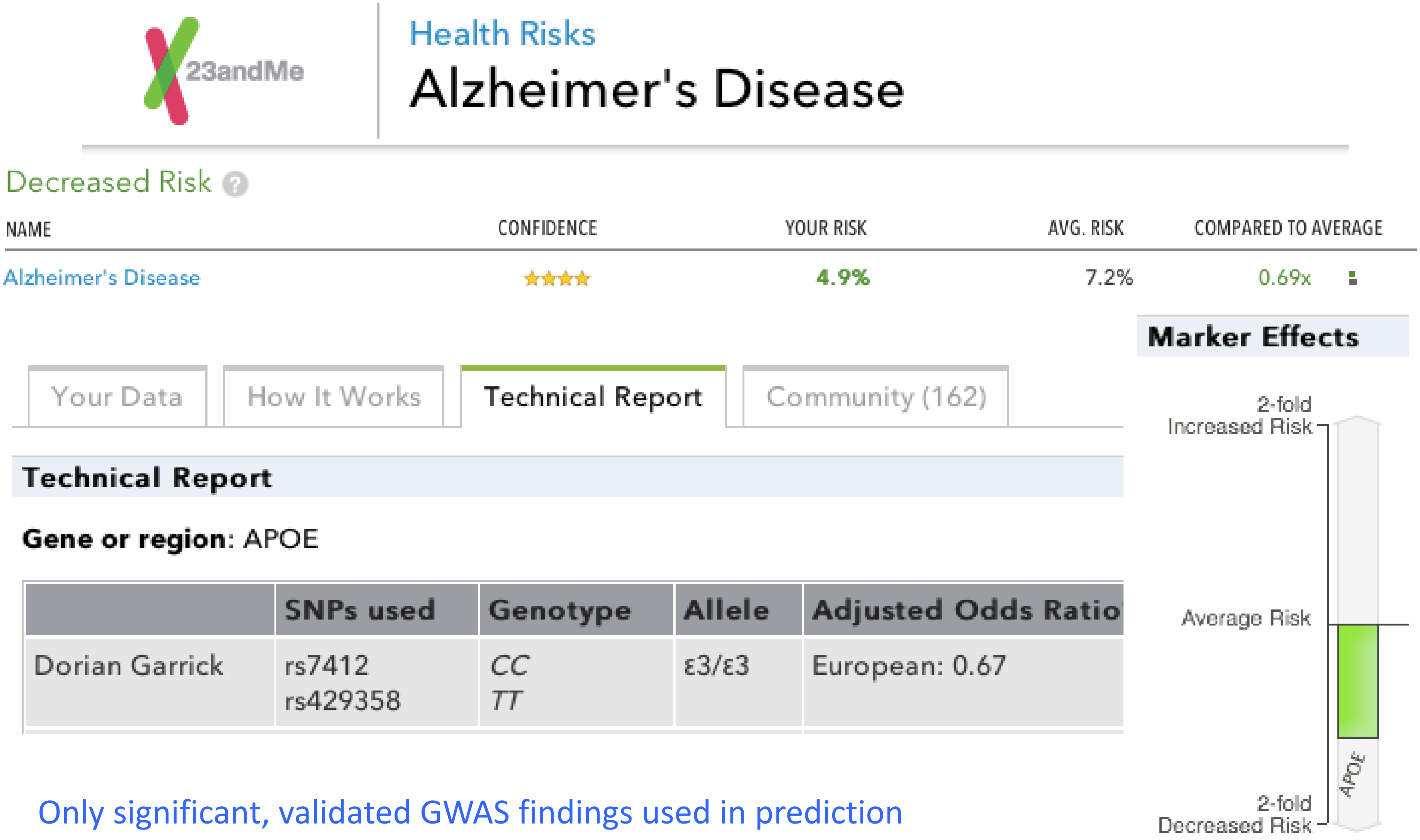
Practice – EBV on SNP



Use SNP genotypes at locus 2 (in low LD) as surrogates for QTL

In practice fitting all SNP simultaneously
Meuwissen, Hayes and Goddard (2001)

www.23andme.com



Health Risks

Alzheimer's Disease

Decreased Risk

NAME	CONFIDENCE	YOUR RISK	AVG. RISK	COMPARED TO AVERAGE
Alzheimer's Disease	★★★★	4.9%	7.2%	0.69x

Technical Report

Gene or region: APOE

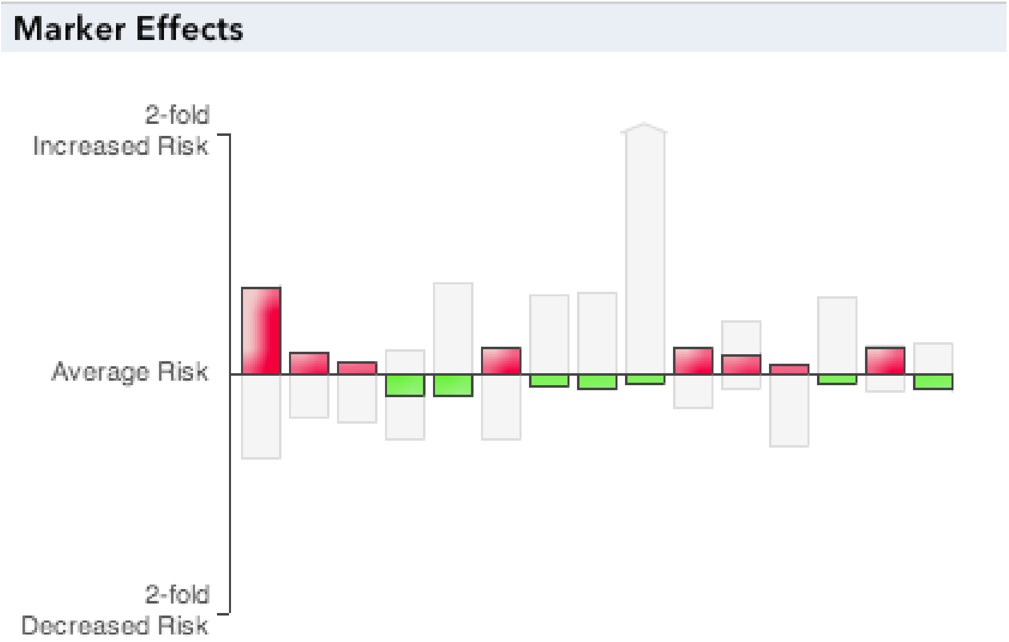
	SNPs used	Genotype	Allele	Adjusted Odds Ratio
Dorian Garrick	rs7412 rs429358	*CC* *TT*	ε3/ε3	European: 0.67

Marker Effects

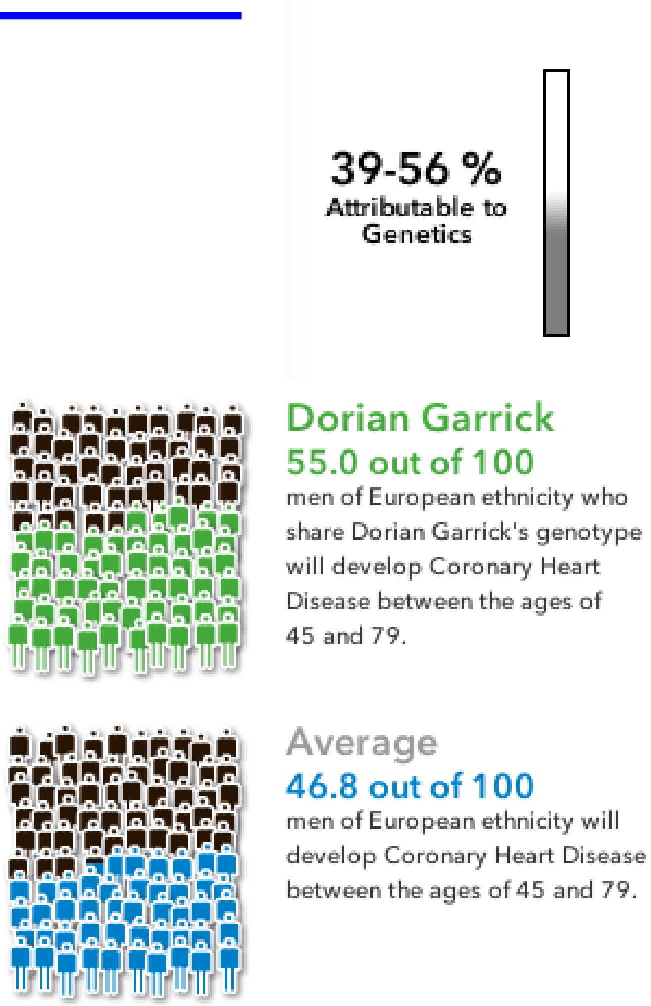
Only significant, validated GWAS findings used in prediction

www.23andme.com

- Coronary Heart Disease



Each bar represents a different risk QTL allele
(mouseover shows the allele and links to the research publications)
QTL=Quantitative Trait Locus



39-56 %
Attributable to Genetics

Dorian Garrick
55.0 out of 100
men of European ethnicity who share Dorian Garrick's genotype will develop Coronary Heart Disease between the ages of 45 and 79.

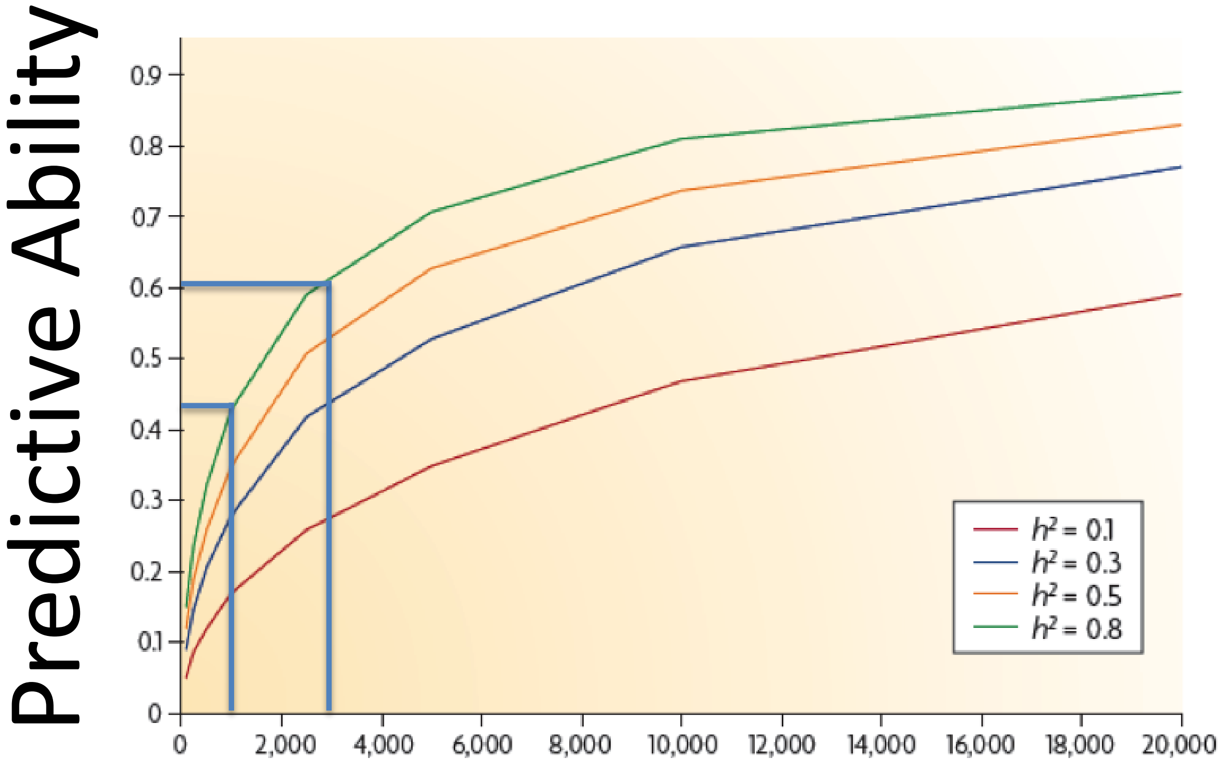
Average
46.8 out of 100
men of European ethnicity will develop Coronary Heart Disease between the ages of 45 and 79.

Only significant, validated GWAS findings used in prediction

Plant & Animal Perspective

- Typically more SNP loci than subjects
- Landmark concepts were suggested by Meuwissen, Hayes & Goddard (2001)
  - Could simply fit all the SNP together (regardless of “significance”) by treating as random effects
    - They referred to these methods as “BLUP” or “BayesA”
  - Or use a variable selection model to fit as random effects some subset of the most informative SNP
    - They proposed a method called “BayesB”

Theoretical Basis for Accuracy



Size of Training Population

Goddard & Hayes (Nature Reviews Genetics, 2009)

Heritability=0.8

N_e=100
like Holsteins & Jerseys

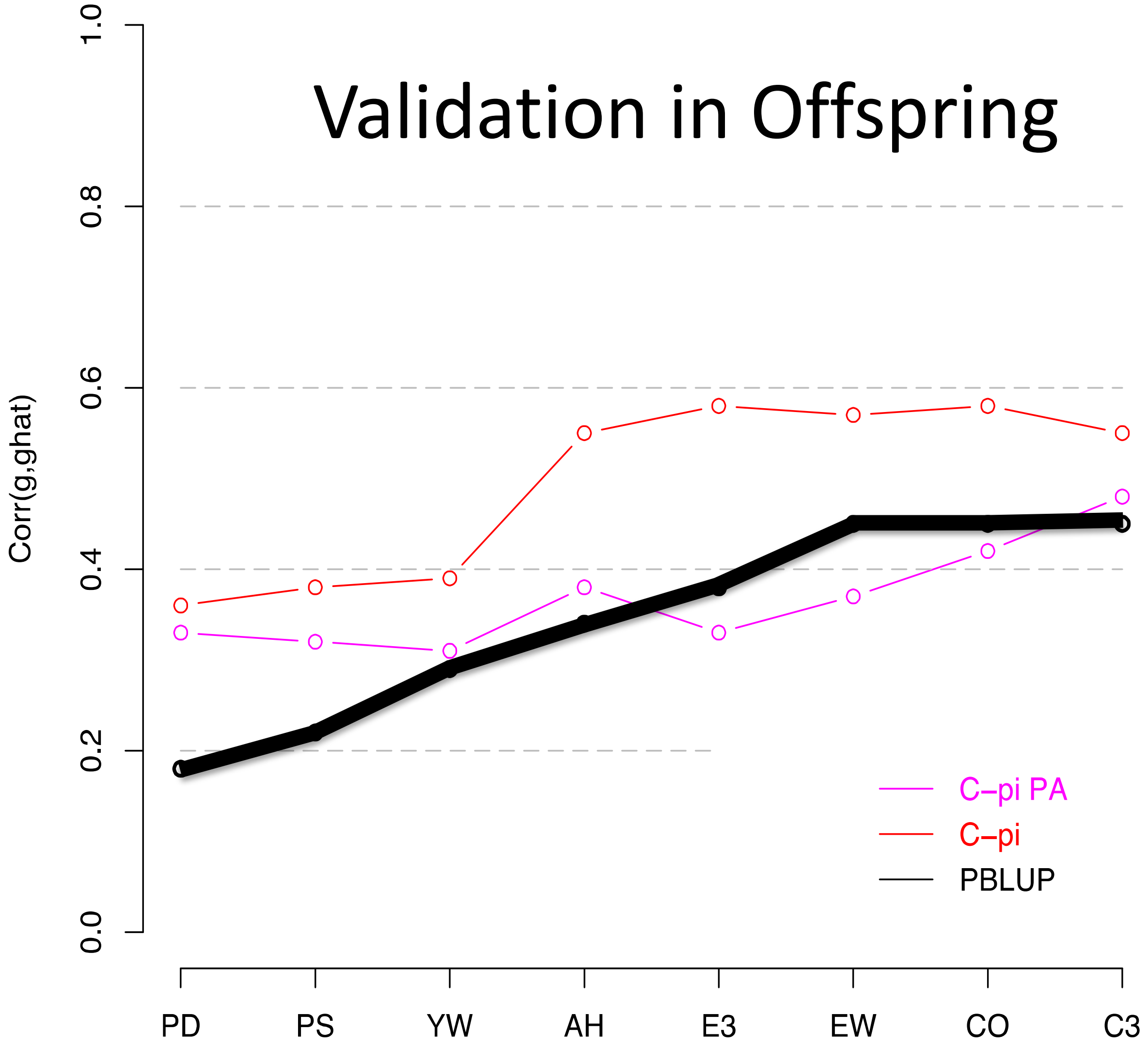
1,000 training animals
r=0.43 20% genetic variance

3,000 training animals
r=0.6 36% genetic variance

Reliable prediction requires large training populations of genotyped and phenotyped individuals

Predictive Ability = Accuracy (r) = correlation true & predicted merit

Accuracy of Genomic Prediction



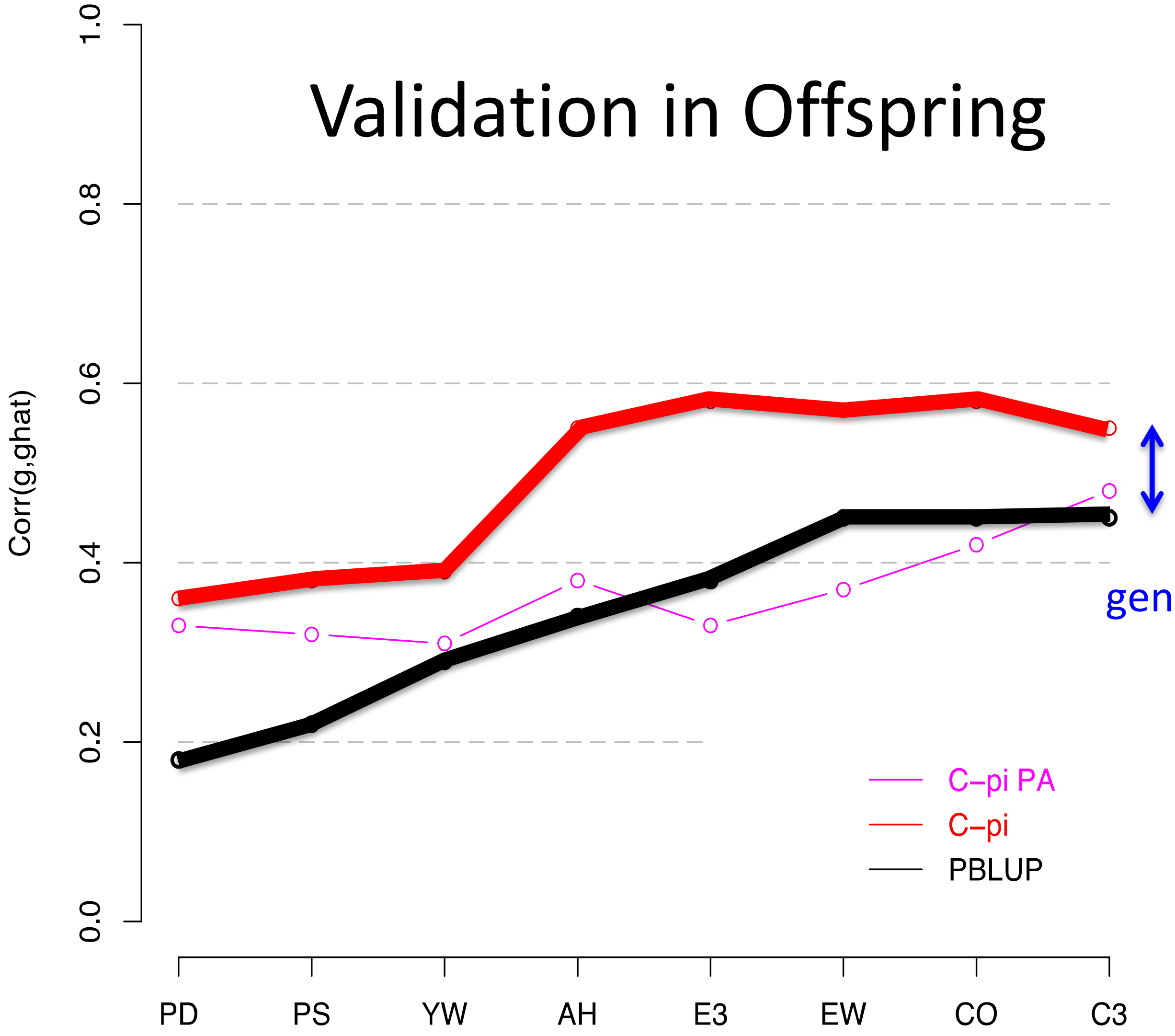
Validation in Offspring

Correlation(g,g-hat)
Early Selection
Layers

Conventional
pedigree
relationships

Wolc et al 2010 9WCGALP

Accuracy of Genomic Prediction



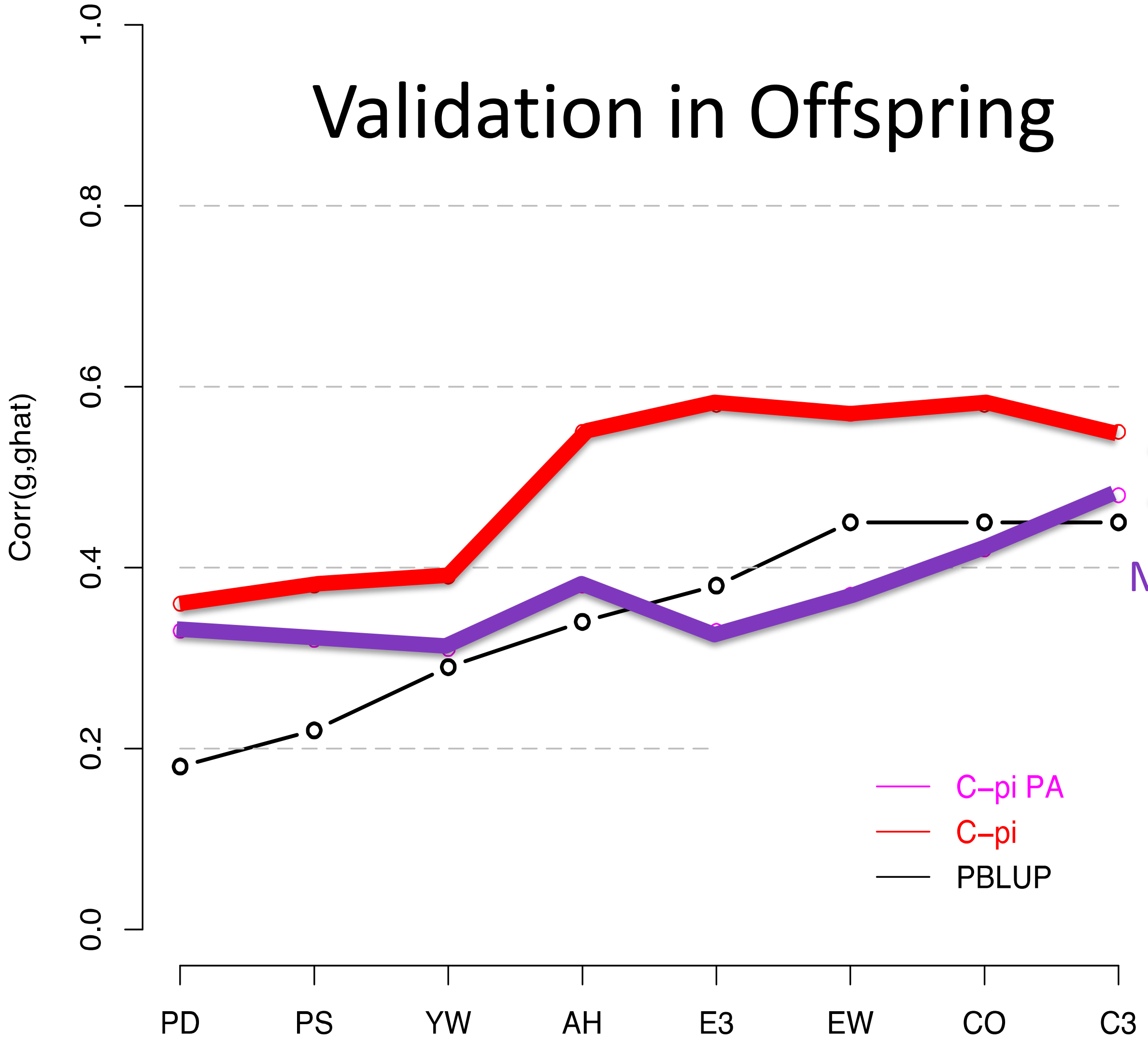
Validation in Offspring

Correlation(g,g-hat)
Early Selection
Layers

Superiority
of prediction
using
genomic relationships

Wolc et al 9WCGALP

Accuracy of Genomic Prediction



Validation in Offspring

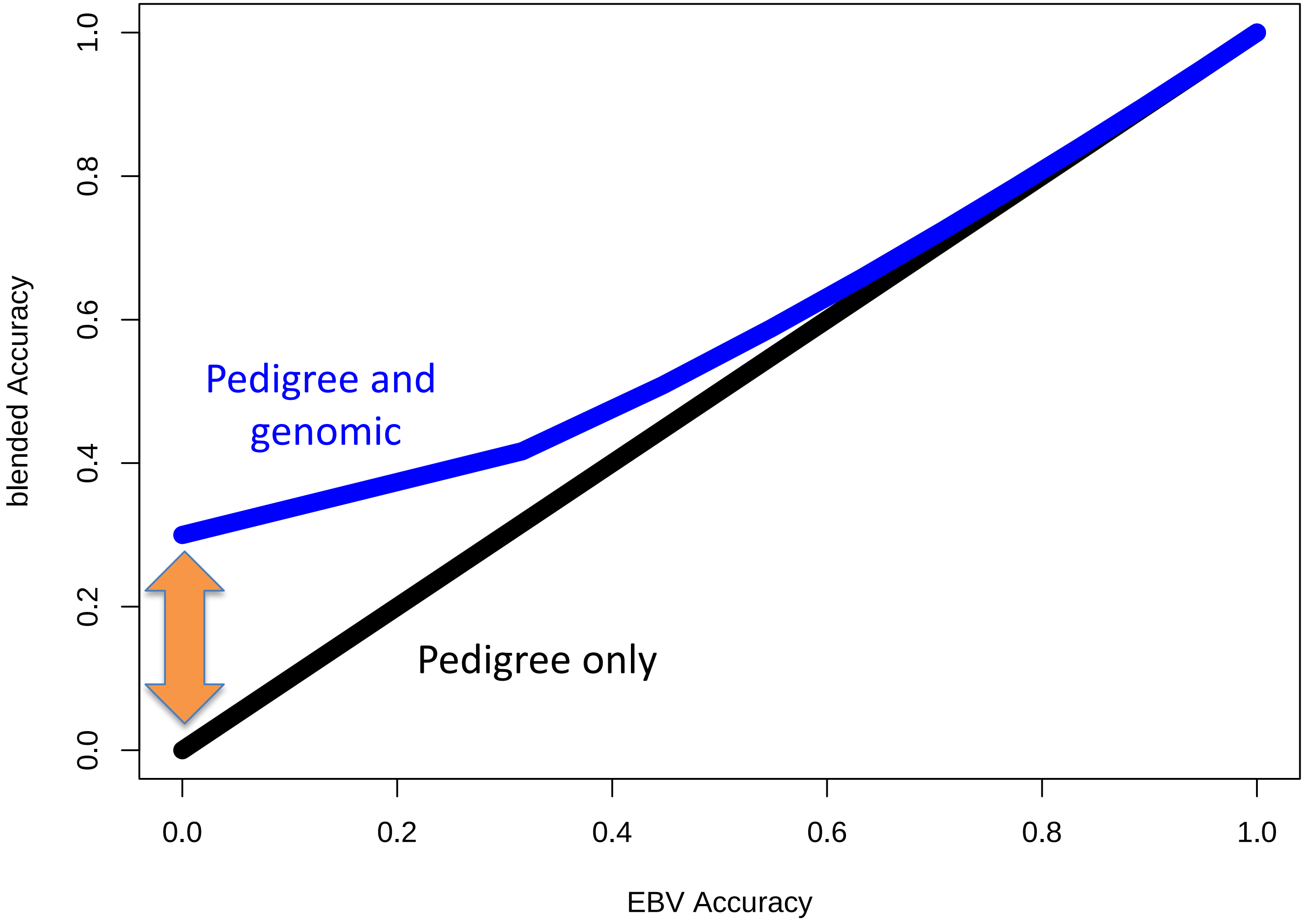
Correlation(g,g-hat)
Early Selection
Layers

Extent genomic
prediction
captures
Mendelian Sampling

Wolc et al 9WCGALP

Impact on Accuracy--%GV=10%

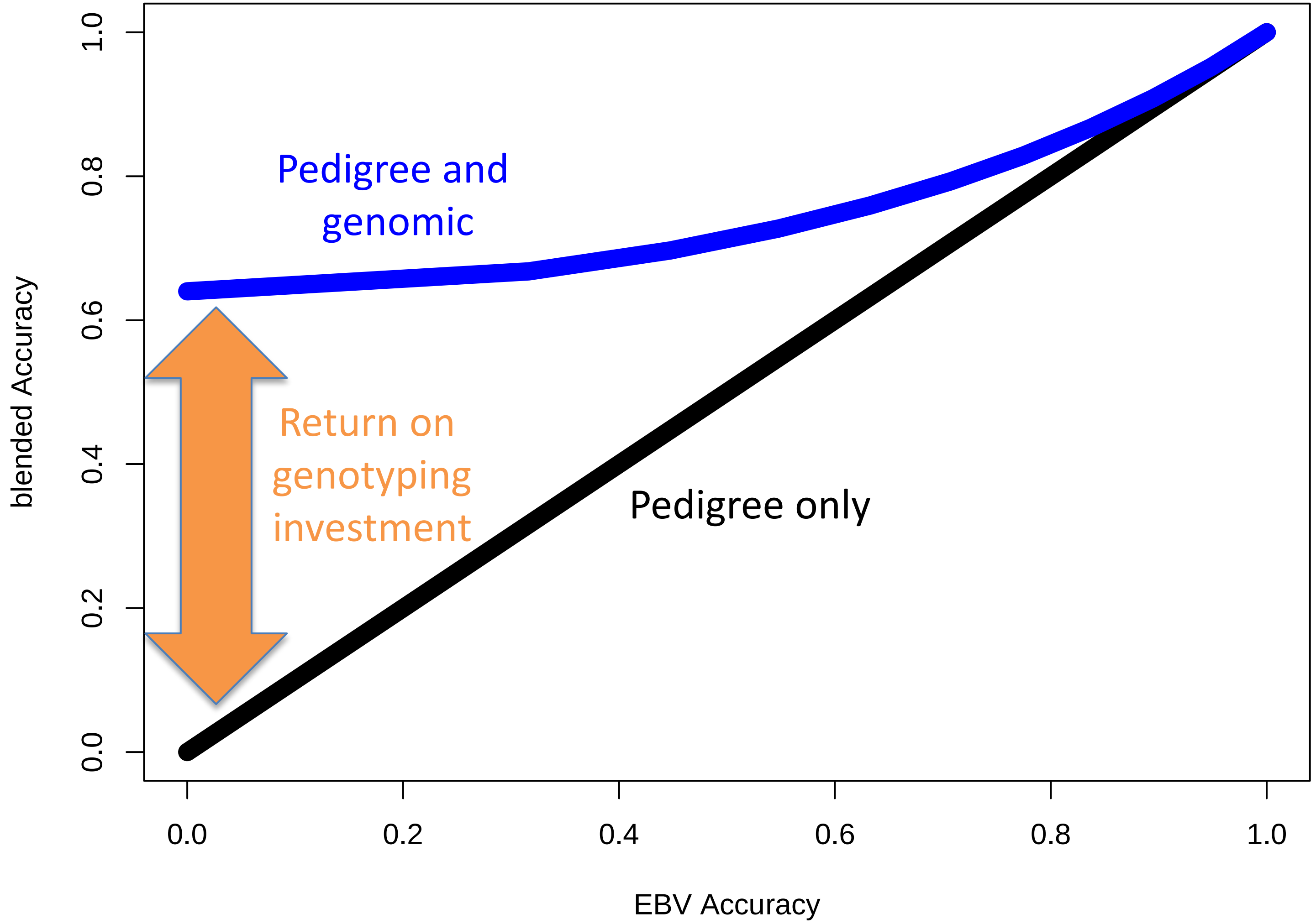
Genetic correlation=0.3



Blending will not improve the accuracy of a bull that already has a reliable EBV

Impact on Accuracy--%GV=40%

Genetic correlation=0.64



Blended EBVs are equally likely to be better or worse than the preblended EBVs

Layer Hens – Dekkers scheme

Strategy	Traditional	
	Male	Female
#candidates with phenotype	1000	3000
# selected	60	360
Generation interval (months)	13	
Information	Own Phenotype	

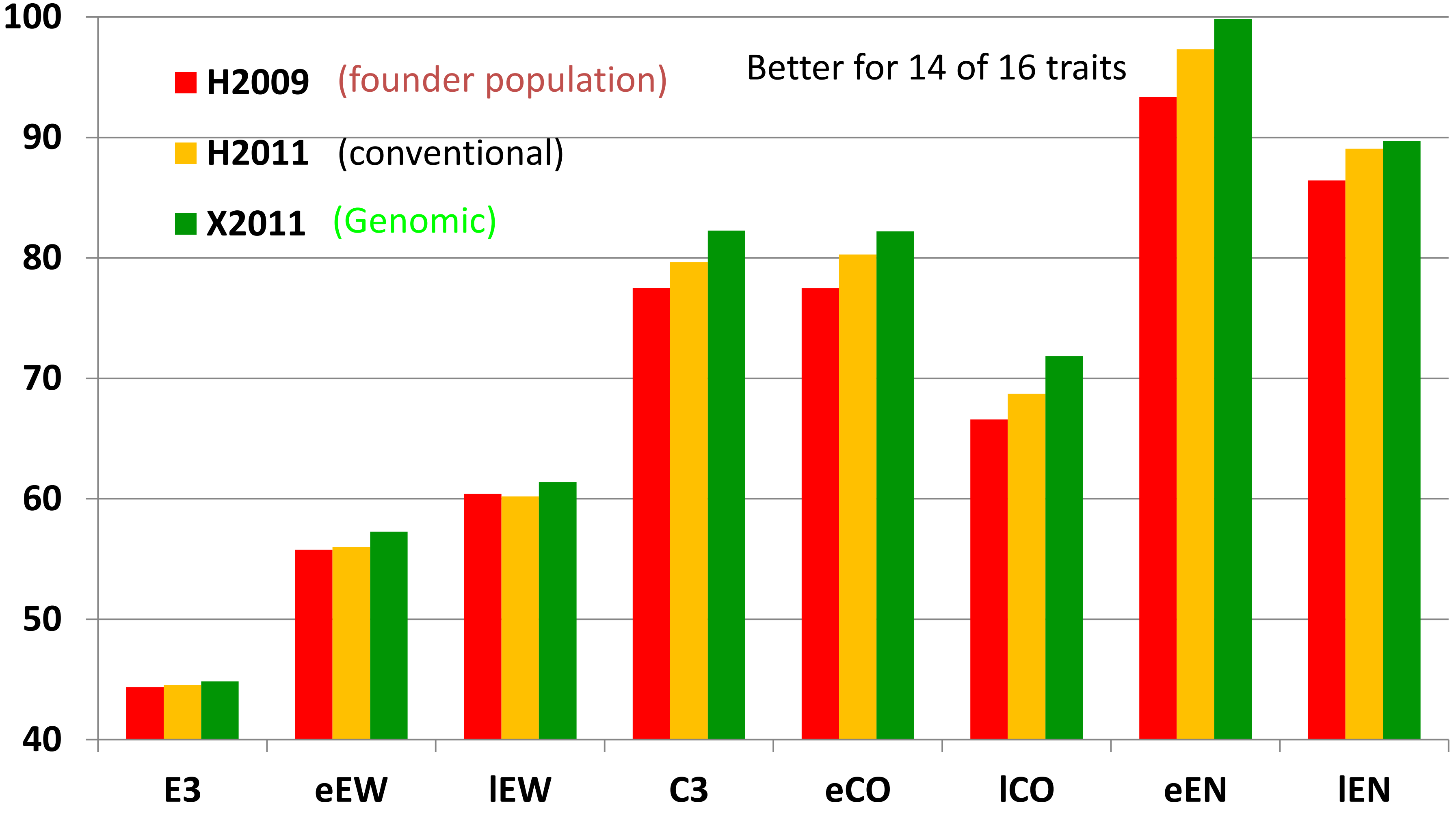
Layer Hens – Dekkers scheme

Strategy	Traditional		GS	
	Male	**Female**	**Male**	**Female**
#candidates with phenotype	1000	3000	300	300
# selected	60	360	50	50
Generation interval (months)	13		6-7	
Information	Own Phenotype		Genotype+Phenotype	

Crossclassified mating

Halve the generation interval and reduce costs by (less phenotyping) to get same gain & same inbreeding

Selection Response - Difference between the lines



Better for 14 of 16 traits

After 3 generations of conventional or 6 gens of genomic selection

Genomic selection was as good, if not better in terms of realized response

Predictions in Beef Cattle Breeds

Trait	RedAngus (6,412)	Angus (3,500)	Hereford (2,980)	Simmental (2,800)	Limousin (2,400)	Gelbvieh (1,321)+
BirthWt	0.75	0.64	0.68	0.65	0.58	0.62
WeanWt	0.67	0.67	0.52	0.52	0.58	0.52
YlgWt	0.69	0.75	0.60	0.45	0.76	0.53
Milk	0.51	0.51	0.37	0.34	0.46	0.39
Fat	0.90	0.70	0.48	0.29		0.75
REA	0.75	0.75	0.49	0.59	0.63	0.61
Marbling	0.85	0.80	0.43	0.63	0.65	0.87
CED	0.60	0.69	0.68	0.45	0.52	0.47
CEM	0.32	0.73	0.51	0.32	0.51	0.62
SC		0.71	0.43		0.45	
Average	**0.67**	**0.69**	**0.52**	**0.47**	**0.57**	**0.56**

Genetic correlations from k-fold validation Saatchi et al (GSE, 2011; 2012; J Anim Sc, 2013)

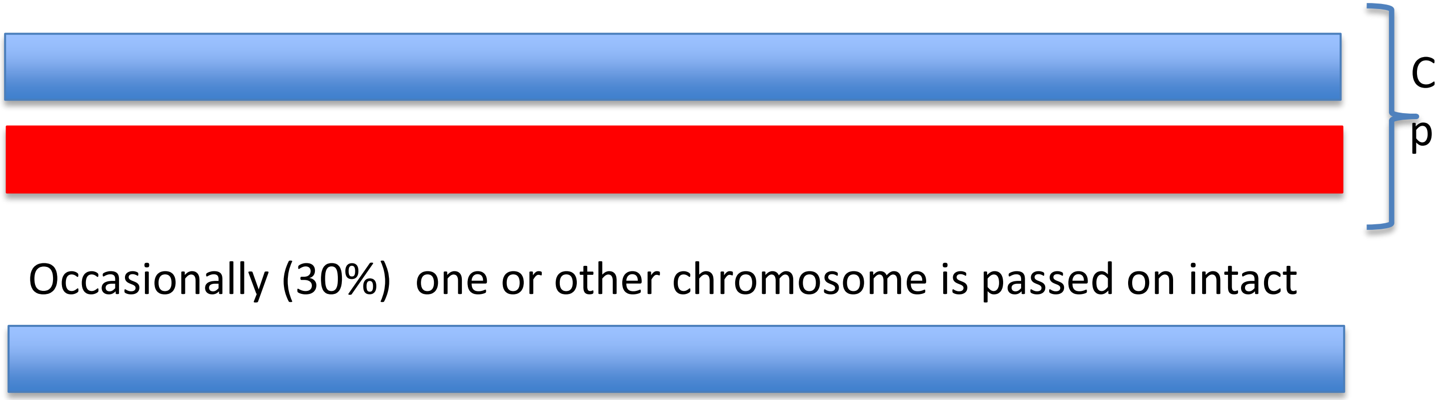
SNP Alleles are inherited in blocks

paternal

maternal

Chromosome pair

SNP Alleles are inherited in blocks



paternal

maternal

Chromosome pair

Occasionally (30%) one or other chromosome is passed on intact

e.g

SNP Alleles are inherited in blocks

paternal

maternal

Chromosome pair

Typically (40%) one crossover produces a new recombinant gamete

Recombination can occur anywhere but there are “hot” spots and “cold” spots

SNP Alleles are inherited in blocks

paternal

maternal

Chromosome pair

Sometimes there may be two (20%) or more (10%) crossovers

Never close together

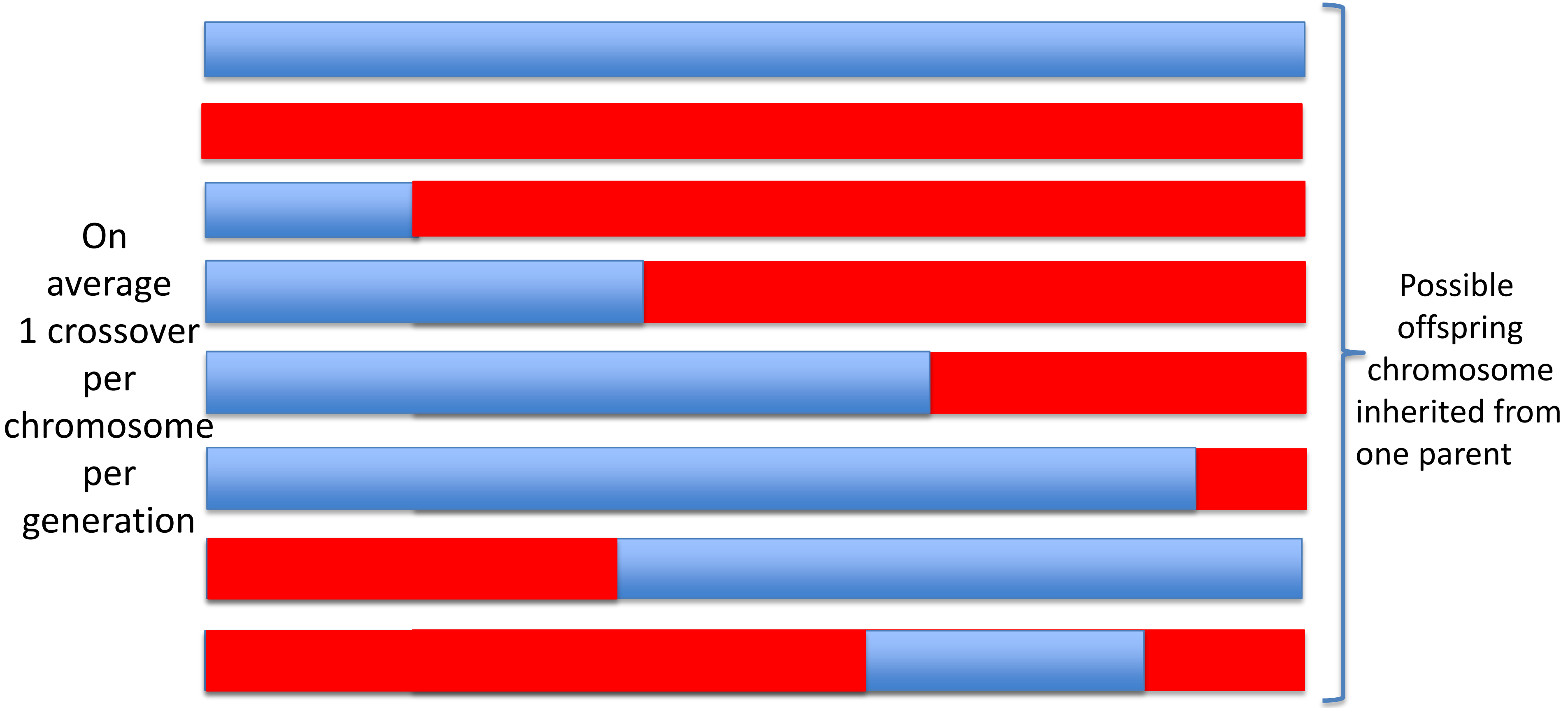
SNP Alleles are inherited in blocks

paternal

maternal

Chromosome pair

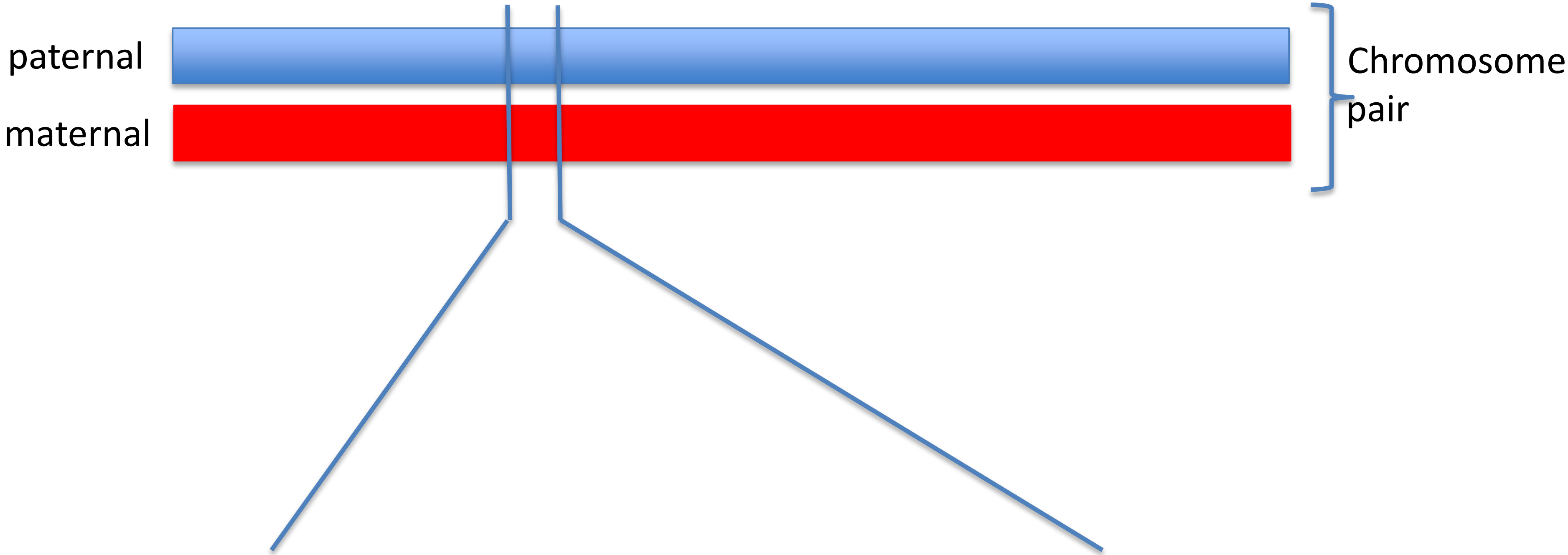
Interestingly the number of crossovers varies between sires and is heritable



On average 1 crossover per chromosome per generation

Possible offspring chromosome inherited from one parent

SNP Alleles are inherited in blocks



paternal

maternal

Chromosome pair

Consider a small window of say 1% chromosome (1 Mb)

SNP Alleles are inherited in blocks

paternal

maternal

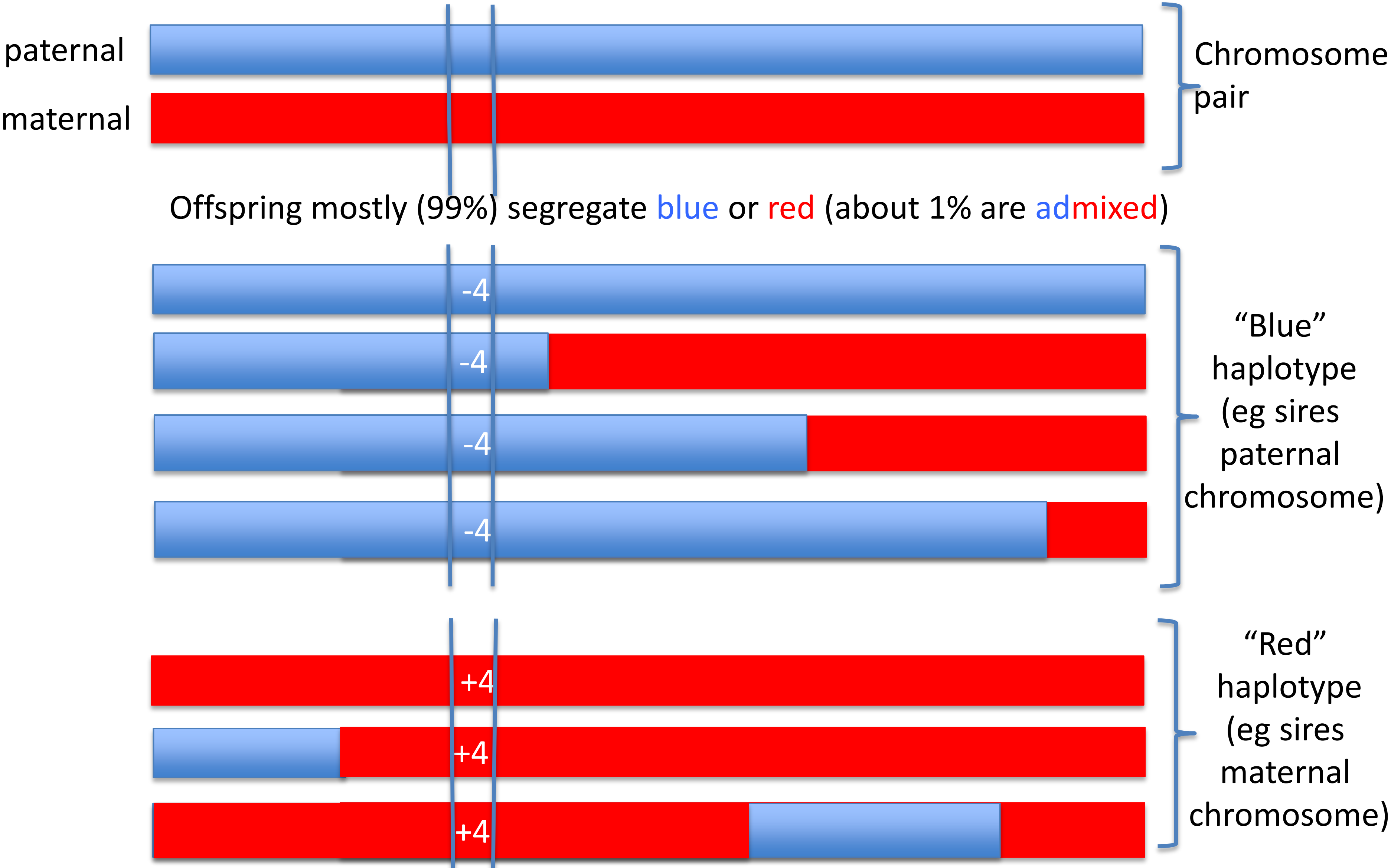
Chromosome pair

Offspring mostly (99%) segregate blue or red (about 1% are admixed)

“Blue” haplotype (eg sires paternal chromosome)

“Red” haplotype (eg sires maternal chromosome)

SNP Alleles are inherited in blocks



paternal

maternal

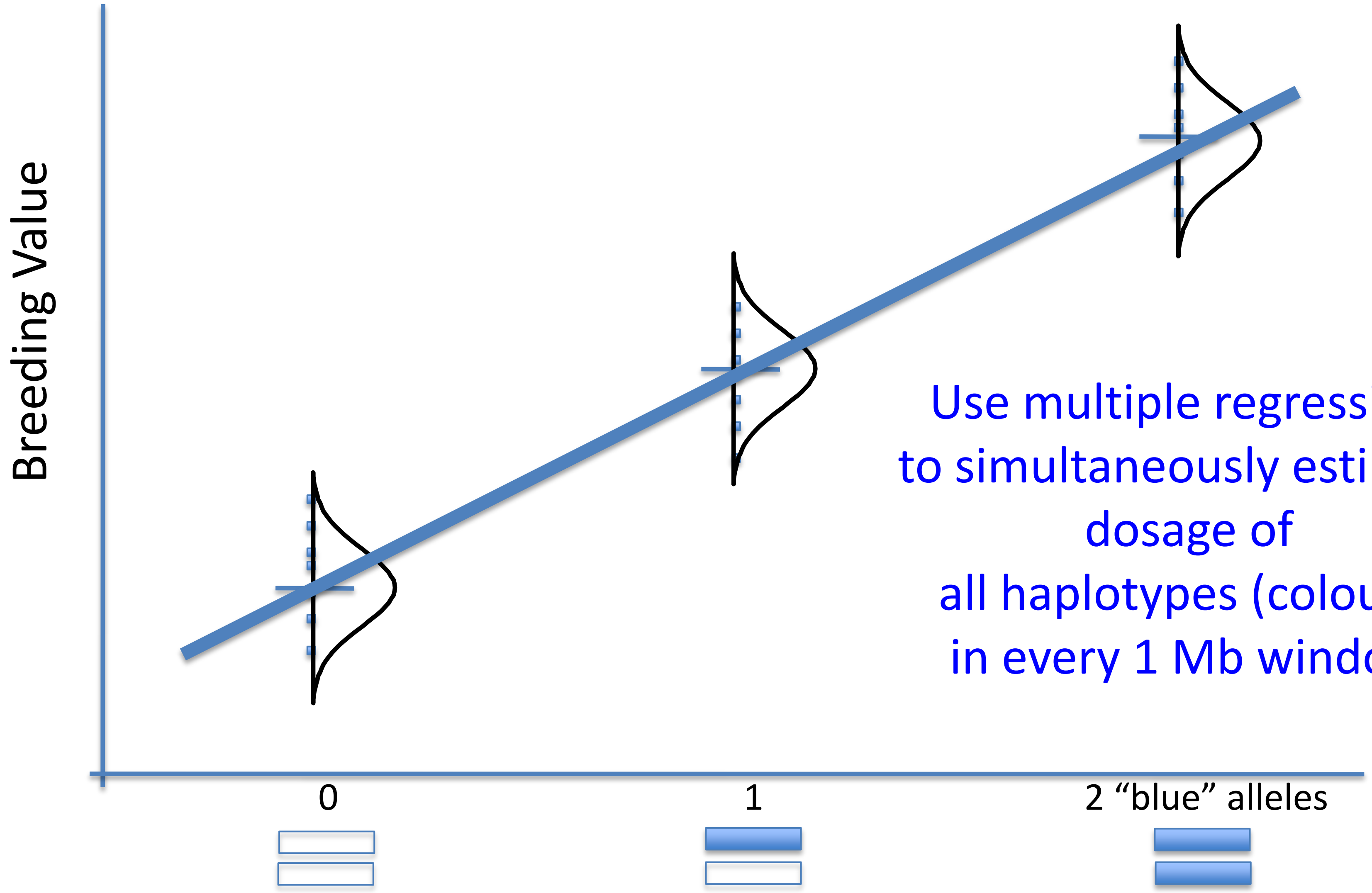
Chromosome pair

Offspring mostly (99%) segregate blue or red (about 1% are admixed)

"Blue" haplotype (eg sires paternal chromosome)

"Red" haplotype (eg sires maternal chromosome)

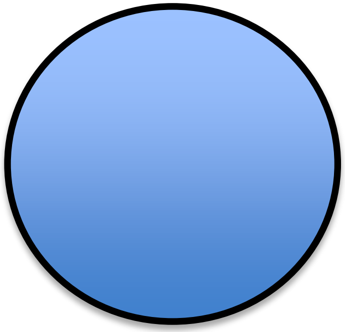
Regress BV on haplotype dosage



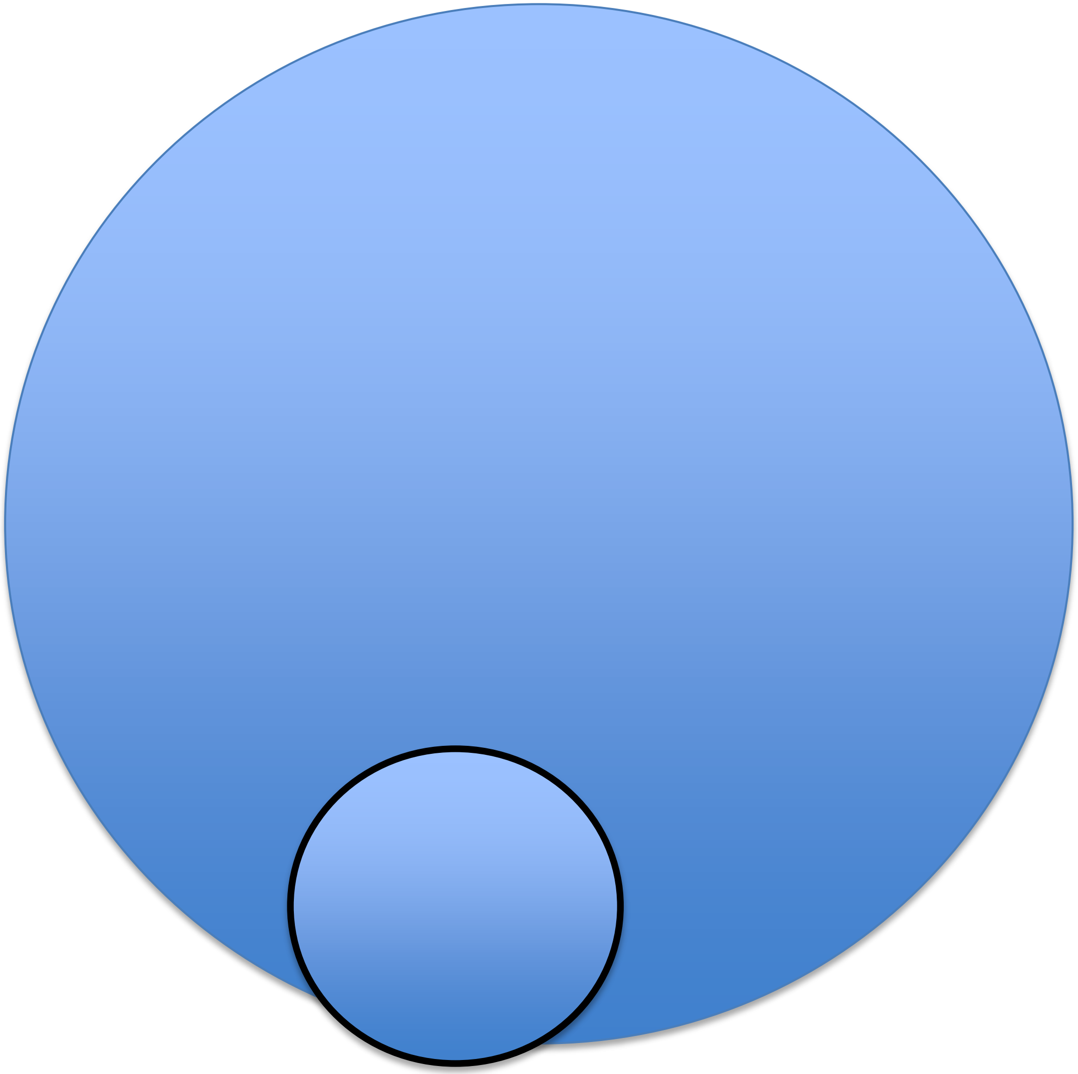
Use multiple regression
to simultaneously estimate
dosage of
all haplotypes (colours)
in every 1 Mb window

Panel Comparison

Black = Illumina 50K



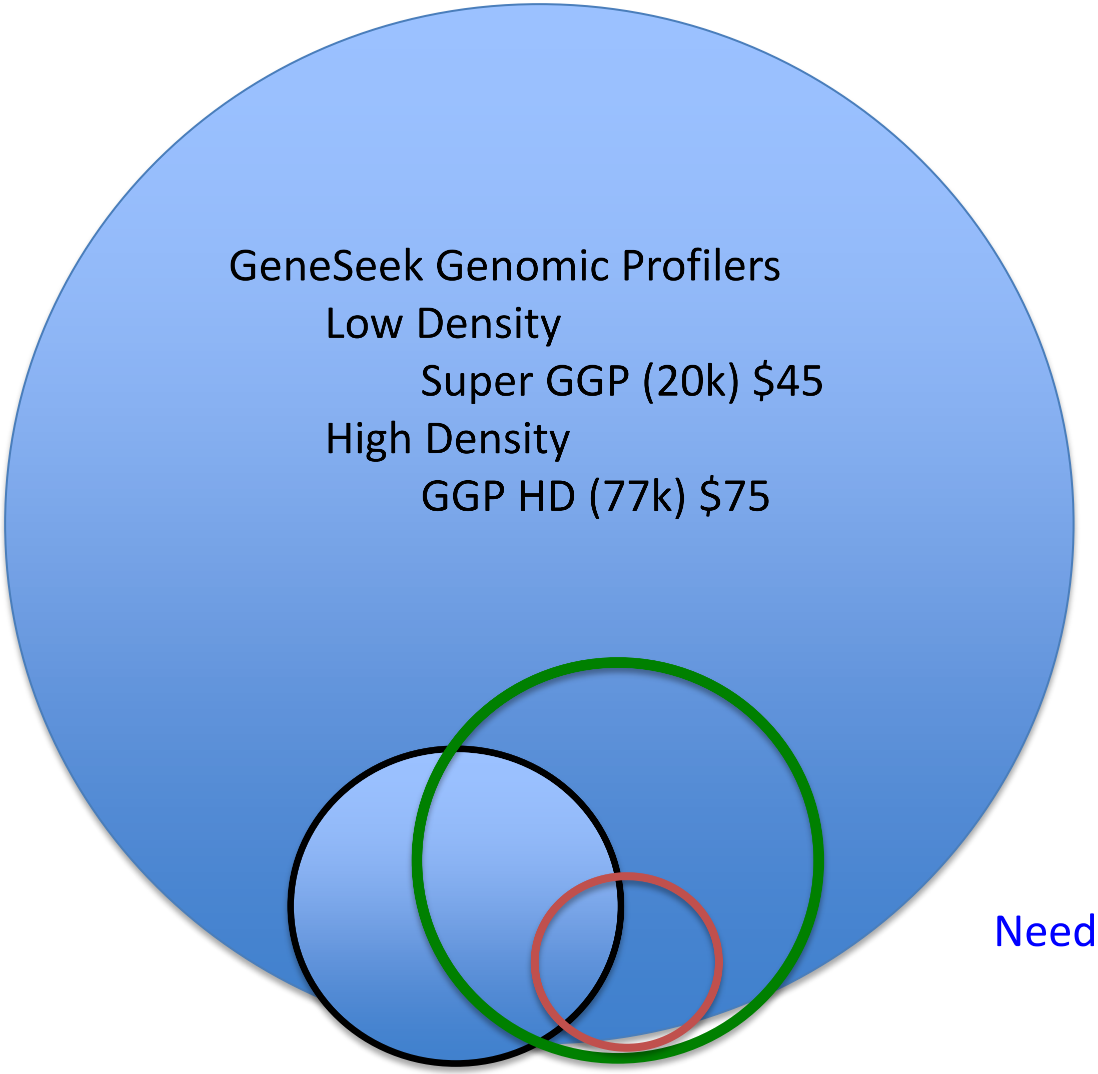
Panel Comparison



Black = Illumina 50K
Blue = Illumina HD (700K)

Panel Comparison

No longer using Illumina 50k



GeneSeek Genomic Profilers
- Low Density
  - Super GGP (20k) $45
- High Density
  - GGP HD (77k) $75

Orange = GGP-Super LD 19k
Green = GGP-HD (taurus) 70k
Black = Illumina 50K

GGP also include custom SNP

50k and GGP-HD share 28K
50k and GGP-Super LD share 8k

Need to genotype more individuals/yr
Need cheaper genotyping

There are multiple minor variants of all these panels!

Also a separate GGP-HD-I (Indicus)

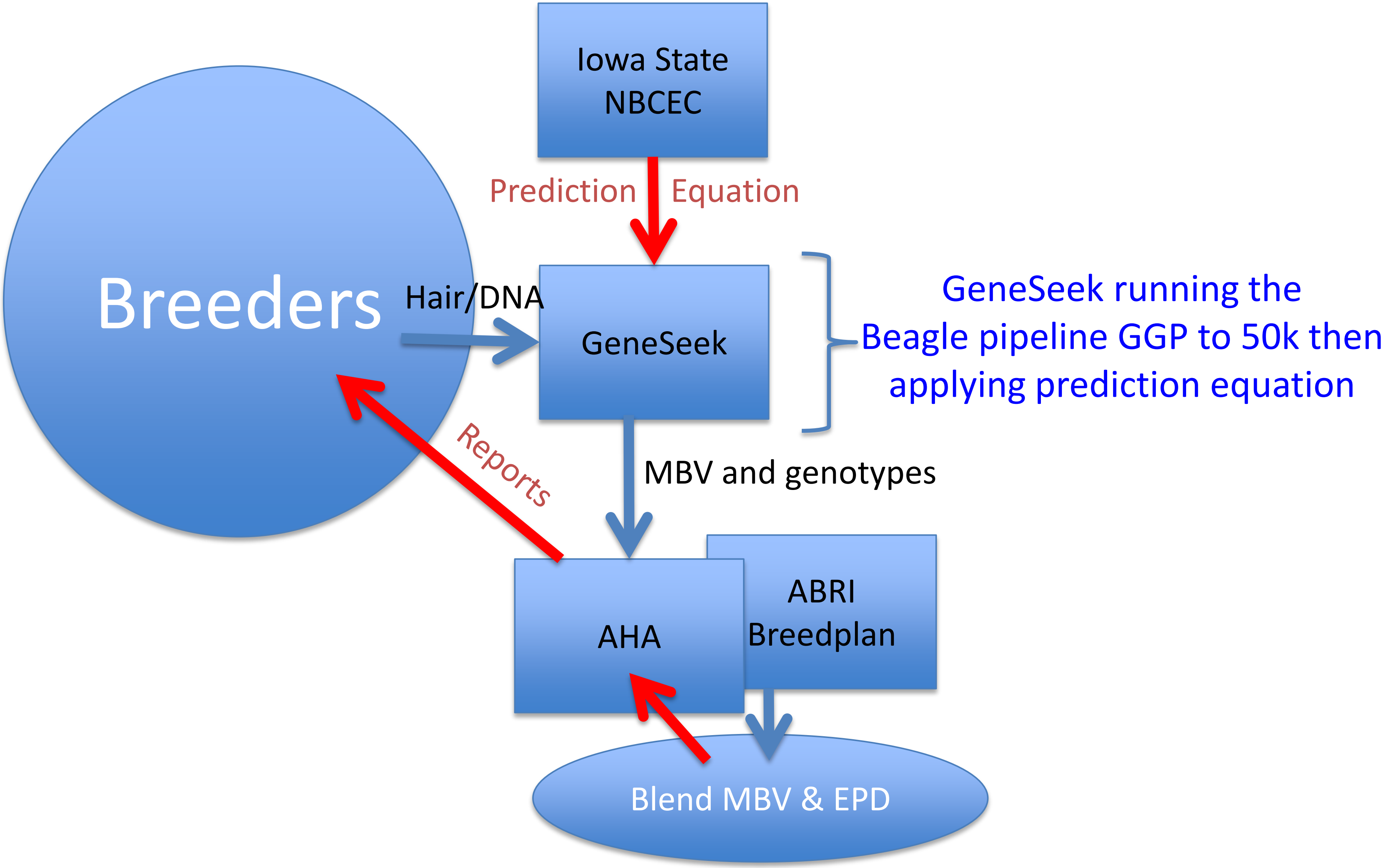
Lower Density Panels

AHA Predictive Accuracy 2,980 6-fold

Trait	Actual	Imputed
Birth Weight	0.67	0.65
Calving Ease Direct	0.68	0.67
Calving Ease Maternal	0.51	0.50
Fat Thickness	0.47	0.46
Marbling	0.42	0.42
Mature cow weight	0.64	0.62
Rib Eye Muscle Area	0.49	0.46
Scrotal Circumference	0.43	0.42
Weaning Weight Direct	0.53	0.50
Weaning Weight Maternal	0.37	0.35
Yearling Weight	0.61	0.59
Mean	0.53	0.51

Actual = 50k
Imputed = 10k
(from GGP-LD)

Genomic Prediction Pipeline



Iowa State NBCEC

Prediction Equation

Breeders

Hair/DNA

GeneSeek

GeneSeek running the Beagle pipeline GGP to 50k then applying prediction equation

MBV and genotypes

Reports

AHA

ABRI Breedplan

Blend MBV & EPD

Early 2014 Genotype Counts

Breed	9k	GGP-LD	50k	GGP-HD	BOS-1	700k HD	TOTAL
AAN		911	13,409	787		947	16,054
BRG			1,128	173		243	1,544
BSH			325			136	461
CHA			1,617			525	2,142
GVH	186	209	1,643	371	414	430	3,253
HER			7,064	1,887	471	850	10,272
LIM		429	3,420	8	461	675	4,993
NEL						2,571	2,571
RAN			1,931	1,183	226		3,340
RDP			1,394				1,394
SIM	5,223	7,026	6,501	1,347	1,601	674	22,372
TOTALS	5,409	8,575	38,432	5,756	3,173	7,051	68,396

Major Regions for Birth Weight

Genetic Variance %

Chr_mb	Angus	Hereford	Shorthorn	Limousin	Simmental	Gelbvieh
7_93	7.10	5.85	0.01	0.02	0.18	0.02
6_38-39	0.47	8.48	11.63	5.90	16.3	4.75
20_4	3.70	7.99	1.19	0.07	1.53	0.03
14_24-26	0.42	0.01	0.01	0.71	3.05	8.14

Adding Haplotypes
3.20%
5.90%

Imputed 700k
Collective 3 QTL
30% GV

Some of these same regions have big effects on one or more of weaning weight, yearling weight, marbling, ribeye area, calving ease

PLAG1 on Chromosome 14 @25 Mb

Effect of 1 copy	**Growth**
Birthweight	5 lb (10 lb for QQ – qq)
Weaning weight	10 lb
Feedlot on weight	16 lb
Feedlot off weight	24 lb
Carcass weight	14 lb

PLAG1 on Chromosome 14 @25 Mb

Effect of 1 copy	Growth
Birthweight	5 lb (10 lb for QQ – qq)
Weaning weight	10 lb
Feedlot on weight	16 lb
Feedlot off weight	24 lb
Carcass weight	14 lb

Effect of 1 copy	Reproduction
Age CL (1st Corpus Luteum)	38 days (76 days QQ – qq)
PPAI (post partum anoestrus)	15 days
Presence CL before weaning	-5%
Weight at CL	36 lb
Age at 26 cm Scrotal Circumf	19 days

Sequence

- Now sequencing individual sires
  - Identify loss-of-function alleles to compare to underrepresented haplotype alleles
  - Identify mutations that are perfectly concordant with haplotype allelic effect
    - More powerful across breed

Genomic Prediction

- Exploits advances in quantitative genetics, statistical genetics, computing, molecular biology, and bioinformatics
- Is the basis for some aspects of personalized medicine
- Will revolutionize plant and animal improvement programmes, but to different extents in different industries

Genomic Prediction

- Its application in humans, plants and animals is still an immature but maturing technology
- Its development will greatly benefit from collaborative activities with other researchers across the entire range of disciplines with interests in genomics

Acknowledgments

- Dr. Rohan Fernando
- Dr. Jack Dekkers
- Dr. Max Rothschild

- Dr. Ania Wolc
- Dr. Bruce Golden

- Dr. Mahdi Saatchi
- Dr. Kadir Kizilkaya
- Dr. David Habier
- Dr. Hailin Su
- Dr. Jungjae Lee
- Dr. Jingjing Yan
- Ziging Weng

- GeneSeek
- Beef Breed Associations
  - American Angus Assoc
  - American Hereford Assoc
  - American Simmental Assoc
  - American Gelbvieh Assoc
  - Red Angus Association
- Aviagen (Broilers)
- HyLine (Layers)
- Livestock Improvement Corp