

Genomic Models



Hossein Jorjani
Hossein.Jorjani@slu.se
018 - 67 19 64

1



SOME BACKGROUND

2



How did it start (in theory)?

Copyright © 2001 by the Genetics Society of America

Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps

T. H. E. Meuwissen,* B. J. Hayes[†] and M. E. Goddard^{†‡}

**Research Institute of Animal Science and Health, 8200 AB Leystad, The Netherlands, [†]Victorian Institute of Animal Science, Attwood 3049, Victoria, Australia and [‡]Institute of Land and Food Resources, University of Melbourne, Parkville 3052, Victoria, Australia*

Manuscript received August 17, 2000
Accepted for publication January 17, 2001

3



How did it start (in plans)?

J. Anim. Breed. Genet. ISSN 0931-2668

ORIGINAL ARTICLE

Strategy for applying genome-wide selection in dairy cattle

L.R. Schaeffer

Department of Animal and Poultry Science, Centre for Genetic Improvement of Livestock, University of Guelph, Guelph, ON, Canada

4



How did it start (in practice)?

Genomic Measures of Relationship and Inbreeding

P.M. VanRaden

*Animal Improvement Programs Laboratory, Agricultural Research Service,
United States Department of Agriculture, Beltsville, MD, USA 20705-2350*

5



How did it start (in practice)?

J. Dairy Sci. 91:4414–4423

doi:10.3168/jds.2007-0980

© American Dairy Science Association, 2008.

Efficient Methods to Compute Genomic Predictions

P. M. VanRaden¹

Animal Improvement Programs Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705-2350

6



How did it start (in practice)?

J. Dairy Sci. 92:16–24
doi:10.3168/jds.2008-1514

© American Dairy Science Association, 2009.

Invited review: Reliability of genomic predictions for North American Holstein bulls

P. M. VanRaden,^{*1} C. P. Van Tassell,^{*†} G. R. Wiggans,^{*} T. S. Sonstegard,[‡] R. D. Schnabel,[‡] J. F. Taylor,[‡]
and F. S. Schenkel[§]

^{*}Animal Improvement Programs Laboratory, and

[†]Bovine Functional Genomics Laboratory, Agricultural Research Service, USDA, Beltsville, MD 20705-2350

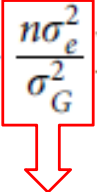
[‡]Division of Animal Sciences, University of Missouri, Columbia 65211

[§]Centre for Genetic Improvement of Livestock, Department of Animal and Poultry Science, University of Guelph, Guelph,
Ontario N1G 2W1, Canada

7



How many methods?

$$\begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} \\ \mathbf{Z}\mathbf{X}' & \mathbf{Z}'\mathbf{Z} + \frac{n\sigma_e^2}{\sigma_G^2} \mathbf{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \mathbf{G} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{Y} \\ \mathbf{Z}'\mathbf{Y} \end{bmatrix}$$


What kind of distribution assign to this?

How do you solve this equation system?

8



How many methods?

Whole-Genome Regression and Prediction Methods Applied to Plant and Animal Breeding

Gustavo de los Campos,^{*,1} John M. Hickey,[†] Ricardo Pong-Wong,[‡] Hans D. Daetwyler,[§] and Mario P. L. Calus^{**}

^{*}Department of Biostatistics, School of Public Health, University of Alabama, Birmingham, Alabama 35294, [†]School of Environmental and Rural Science, University of New England, Armidale 2351, New South Wales, Australia, [‡]The Roslin Institute, Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian, EH25 9RG, Scotland, [§]Biosciences Research Division, Department of Primary Industries, Bundoola 3083, Victoria, Australia, and ^{**}Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, 8200 AB Lelystad, The Netherlands

Genetics, Vol. 193, 327–345 February 2013

9



How do you interpret the methods?

Copyright © 2009 by the Genetics Society of America
DOI: 10.1534/genetics.109.109592

Additive Genetic Variability and the Bayesian Alphabet

Daniel Gianola,^{*,†,‡,1} Gustavo de los Campos,^{*} William G. Hill,[§] Eduardo Manfredi,[‡] and Rohan Fernando^{**}

^{*}Department of Animal Sciences, University of Wisconsin, Madison, Wisconsin 53706, [†]Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, N-1432 Ås, Norway, [‡]Institut National de la Recherche Agronomique, UR631 Station d'Amélioration Génétique des Animaux, BP 52627, 32326 Castanet-Tolosan, France, [§]Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom and ^{**}Department of Animal Science, Iowa State University, Ames, Iowa 50011

Manuscript received April 14, 2009
Accepted for publication July 16, 2009

10



An example

INTERBULL BULLETIN NO. 46. Cork, Ireland, May 28 – 31, 2012

Status of Genomic Evaluation in the Brown Swiss Populations

Hossein Jorjani, Jette Jakobsen, Eva Hjerpe, Valentina Palucci and João Dürr

Interbull Centre, Department of Animal Breeding and Genetics,

Swedish University of Agricultural Sciences, Box 7023, S-75007, Uppsala Sweden

e-mail: Hossein.Jorjani@slu.se

11



Data

(All in flat files)



Genotype checks

- Handling SNPs for multiple reported genotypes;
- Handling of conflicting sire-son SNP conflicts;
- Handling of conflicts between A & G
-

13



SNP editing

- Is the SNP is bi-allelic;
- Call rate for each locus and animal;
- Minor allele/genotype frequency;
- Within and across country test of conformity to HW ratios;
- Across genome phase of all SNPs;

14



Pedigree

Do we want to use the pedigree?

$$\begin{bmatrix} X'X & X'Z \\ ZX' & Z'Z + \frac{n\sigma_e^2}{\sigma_G^2} \mathbf{I} \end{bmatrix} \begin{bmatrix} \beta \\ G \end{bmatrix} = \begin{bmatrix} X'Y \\ Z'Y \end{bmatrix}$$


Do we want to use A, G or a mix of them?

15



Phenotype

- Real phenotype;
- A proxy :
 - Predicted Genetic Merit (EBV, PTA);
 - De-regressed PGM;
 - Daughter Yield Deviation (or any "corrected phenotype")

16



Check of genomic results

- Correlations of PGM, DGV and GEBV for different groups of animals;
- Gain in reliability for three groups of animals;
- Correlation between “REL PGM” and the “difference between PGM & DGV/GEBV”;
- Etc.

17



Imputation of missing SNPs

- From low to high density;
- Speed;
- Accuracy;

18



Genomic evaluation method and implementation

- The genomic evaluation was the one proposed by VanRaden (2008; JDS: 91: 4414-4423).
- An iterative, nonlinear model with heavy-tailed prior for marker effects analogous to Bayes A is used (a curve parameter of 1.05 was used). Base population allele frequencies are subtracted from genotypes, and a polygenic effect (poly) with 10% of additive variance is fit in the model:
- $DPGM = \text{mean} + \sum \text{genotypes} * \text{effects} + \text{poly} + \text{error}.$

19



Results



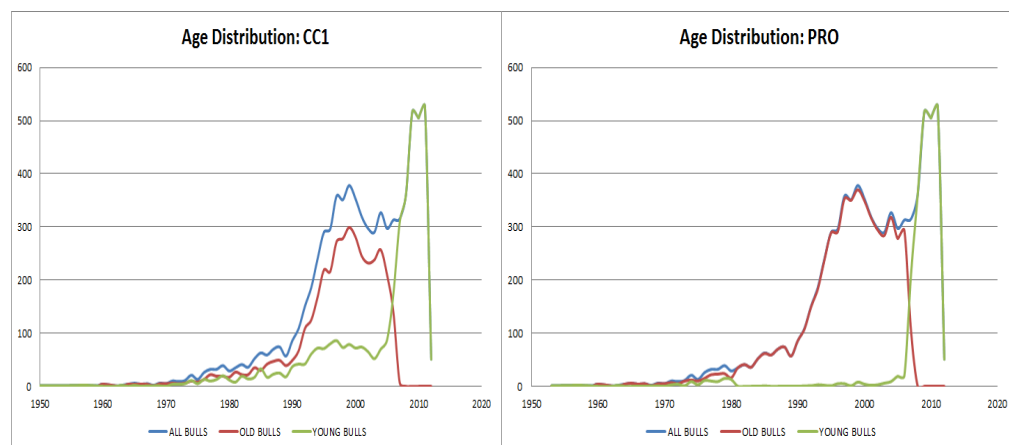
Genotypes (May 2012)

	SNP panel				
	LD	50KV1	50KV2	HD	Sum
CHE	-	1799	948	-	2733
DEA	-	1821	639	-	2445
FRA	-	91	135	-	226
ITA	-	1547	185	-	1644
SVN	-	191	-	-	191
USA	4	938	44	15	994
Used					7670

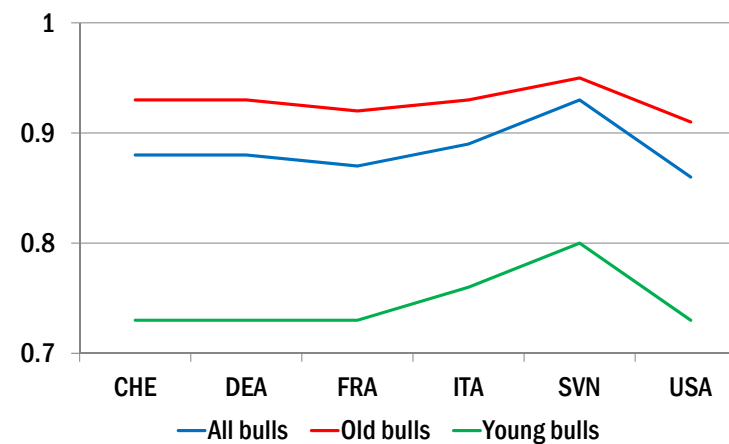


Phenotype (May 2012)

Trait	Number of		Trait	Number of	
	Populations	Bulls		Populations	Bulls
ANG	3	2162	MCE	3	3554
BDE	4	4660	MIL	6	5320
CC1	4	3882	MSP	3	4653
CC2	4	3997	OCS	5	3584
CRC	4	4009	OFL	4	4517
CWI	5	4439	OUS	4	4579
DCE	3	4460	PRO	6	5320
DLO	5	4964	RAN	5	4872
FAN	5	4868	RLS	5	4872
FAT	6	5320	RTP	4	3560
FTL	5	4871	RUH	5	4868
FTP	4	3723	RUW	5	4868
FUA	5	3864	RWI	4	3959
HCO	3	2697	SCS	5	5067
HDE	4	4597	STA	5	4872
INT	2	2782	UDE	5	4872
MAS	5	5067	USU	5	4872
			Total	150	

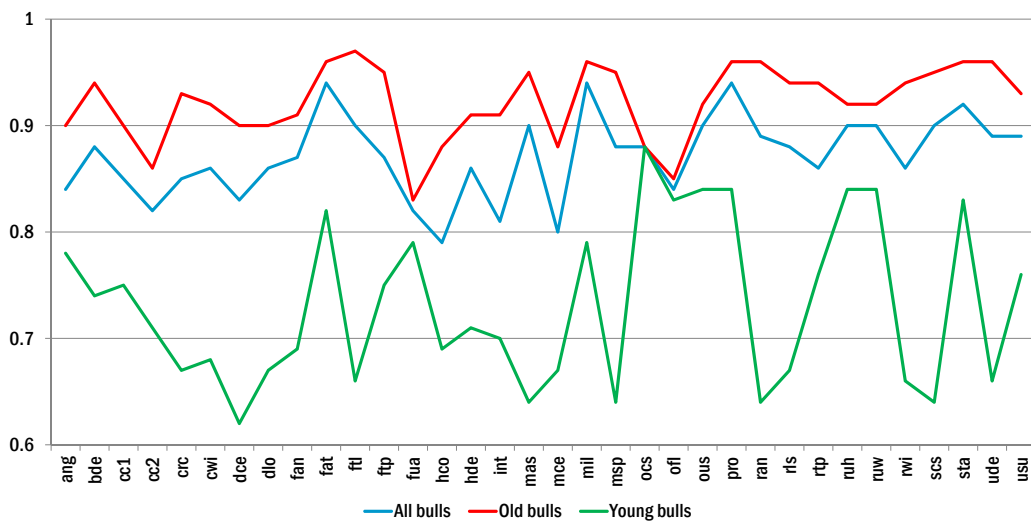


$$r_{(EBV, GEBV)}$$

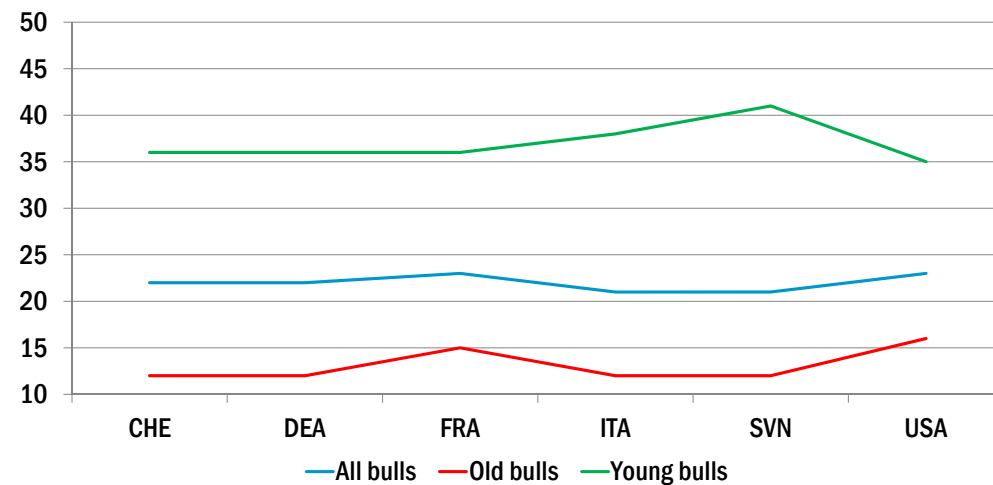




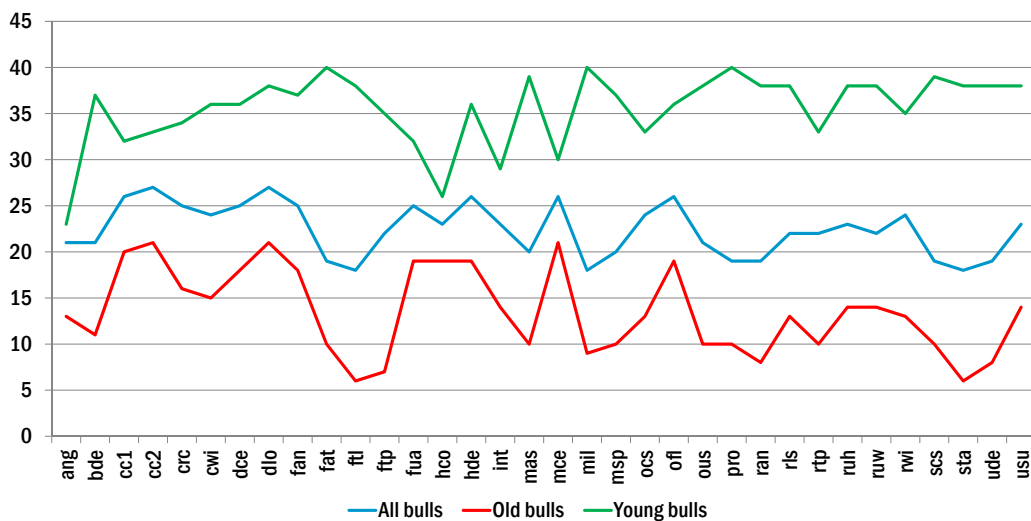
$r_{(EBV, GEBV)}$



Gain in reliability



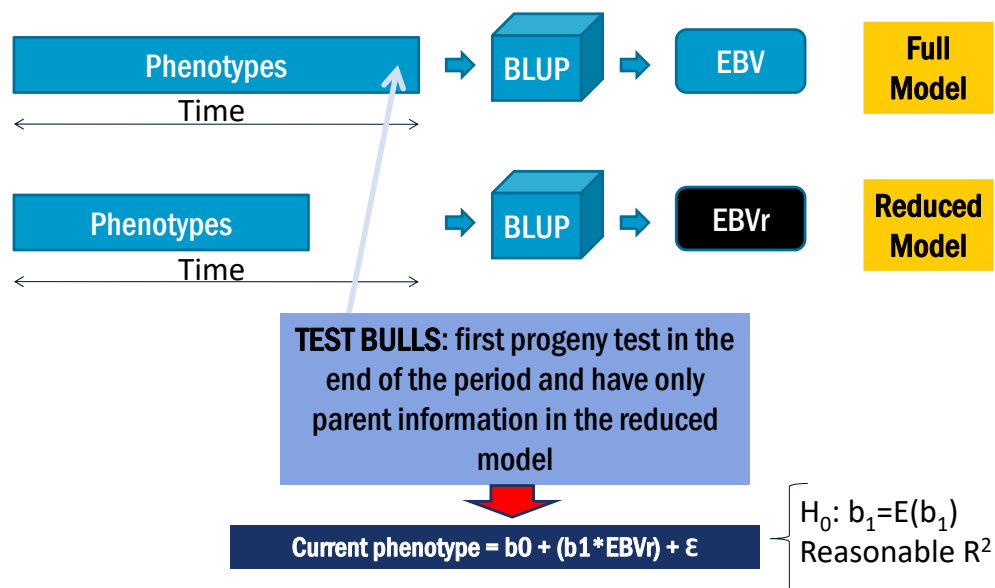
Gain in reliability



Validation



Validation principle



Validation: 73 “country-trait” combinations

GROUP	TRAIT	POPULATIONS	REFERENCE POPULATION	VALIDATION POPULATION
CONF	ang	3	1394	450
FERT	cc2	3	2383	922
LONG	dlo	5	3756	656
CONF	fan	5	3125	1107
PROD	fat	6	3505	1270
CONF	ftl	5	3128	1107
CONF	ftp	4	2288	927
FERT	int	2	1637	605
PROD	mil	6	3505	1270
WORK	msp	3	2980	1058
PROD	pro	6	3505	1270
CONF	ran	5	3129	1107
CONF	rls	5	3129	1107
CONF	ruh	5	3125	1107
UDER	scs	5	3319	1218
CONF	sta	5	3129	1107



Validation

Across country validation

All bulls used in validation

Within country validation

Only domestic bulls used in validation

2 bull selection criteria * (16 traits) * (2-6 countries / trait) = 146 “country-traits” to validate



Across country validation (Sep 2011)

TRAIT	Total P	Total F
ang	2	1
cc2	2	1
dlo	2	3
fan	2	3
fat	6	0
ftl	5	0
ftp	2	2
int	1	1
mil	6	0
msp	3	0
pro	6	0
ran	3	2
rls	4	1
ruh	2	3
scs	4	1
sta	5	0
Total P	52	
Total F		21



Within country validation (Sep 2011)

TRAIT	Total P	Total F
ang	2	1
cc2	3	0
dlo	3	2
fan	2	3
fat	6	0
ftl	5	0
ftp	3	1
int	2	0
mil	5	1
msp	3	0
pro	5	1
ran	4	1
rls	5	0
ruh	3	2
scs	4	1
sta	5	0
Total P	60	
Total F		13



Gain and Gain (Sep 2011)

		Genomic validation	Genomic evaluation
		R ² - Gain	REL Gain
CHE	fat	22	38
DEA	fat	22	39
FRA	fat	21	39
ITA	fat	21	39
SVN	fat	20	40
USA	fat	20	38
MEAN	All traits	16.3	36.3
STD	All traits	8.0	3.9



Across country validation (Jan 2012)

	Passed		Failed	
	Mean	SD	Mean	SD
b₁	0.94	0.07	0.75	0.10
SE_{b1}	0.10	0.11	0.04	0.01
R² Gain	17.21	7.93	12.98	5.87



Conclusions



Conclusions

Pooling of genotype data from several countries is an effective way of building a large enough reference population.

International (MACE) breeding values work well as the input/phenotype.

The reliability gain for the old bulls is 5%-10% for high h^2 traits, and 10%-20% for low h^2 traits (interpret these numbers cautiously).

The reliability gain for the young bulls is 20%-40% irrespective of h^2 (interpret these numbers cautiously).

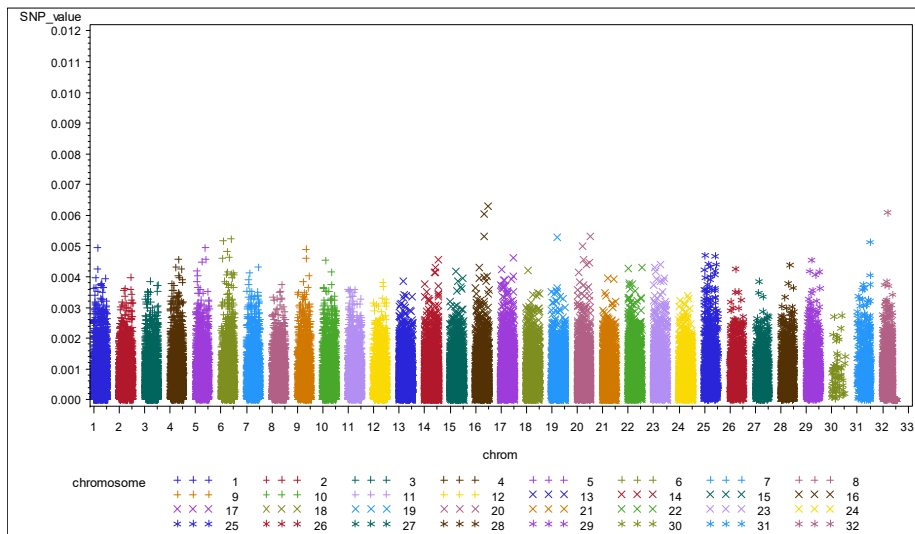


A QUESTION

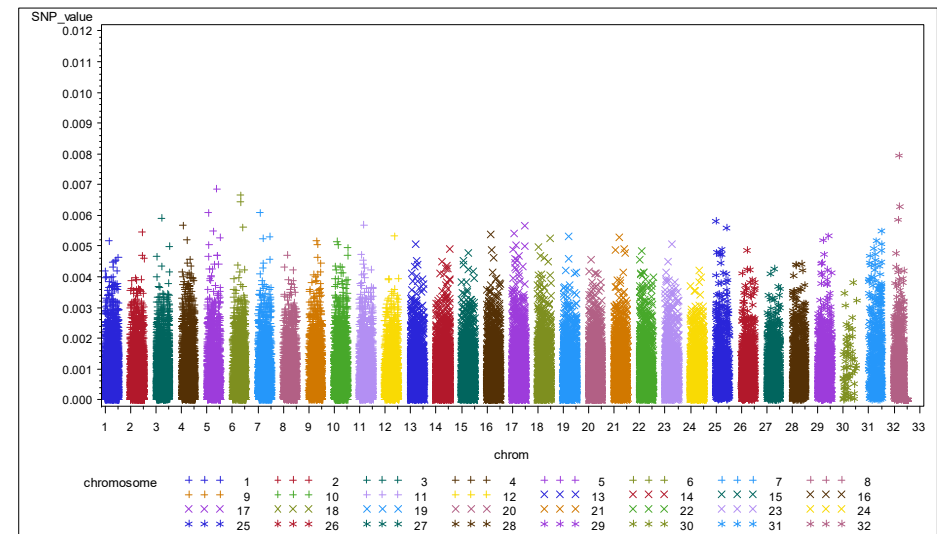
38

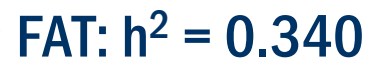


MIL: $h^2 = 0.360$



PRO: $h^2 = 0.350$





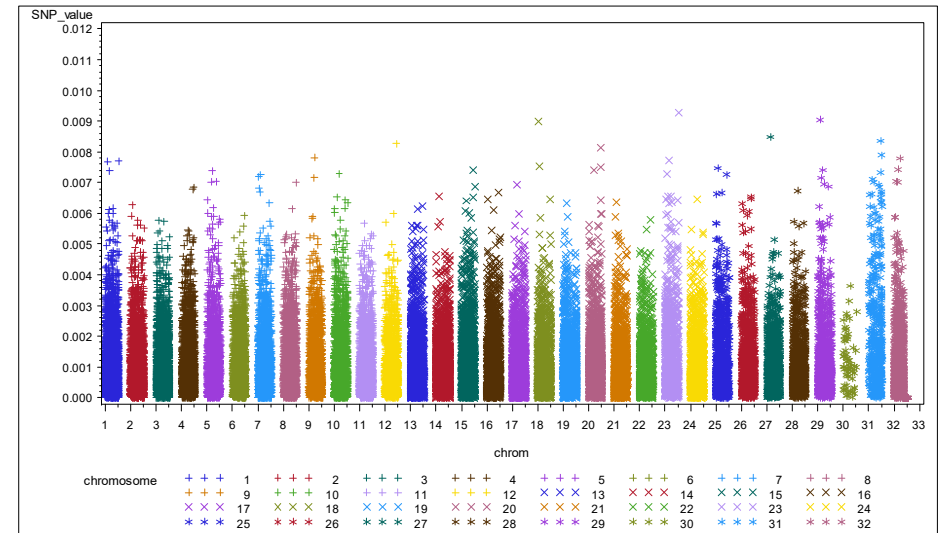
SCS: $h^2 = 0.270$

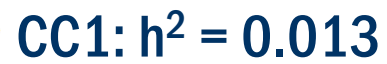


DLO: $h^2 = 0.163$



CRC: $h^2 = 0.039$





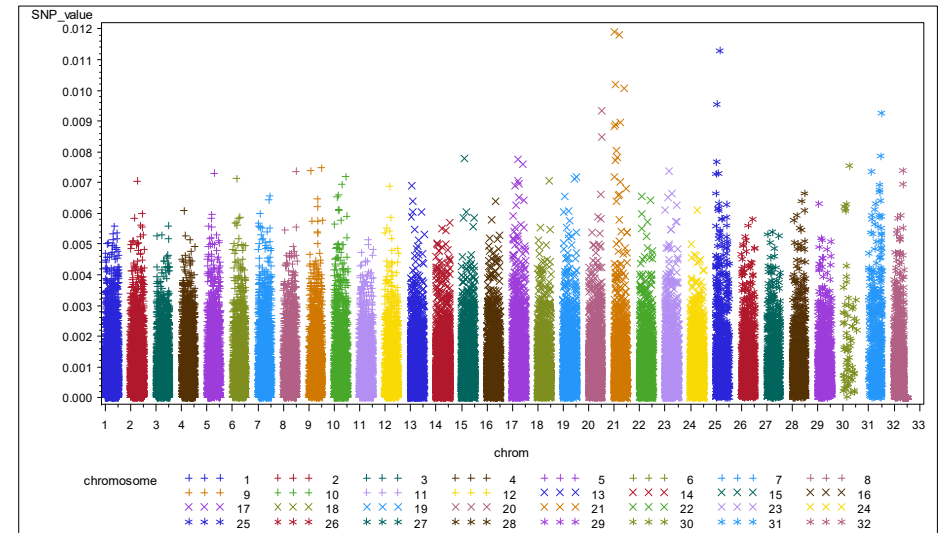
CC2: $h^2 = 0.013$



MCE: $h^2 = 0.011$



DCE: $h^2 = 0.009$





Hossein Jorjani

Interbull Centre

Department of Animal Breeding & Genetics

Swedish University of Agricultural Sciences

Box 7023

S – 750 07 Uppsala, Sweden

