Use of Haplotypes to Predict Selection Limits and Mendelian Sampling

J.B. Cole P.M. VanRaden

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

john.cole@ars.usda.gov

August 2, 2010

Introduction

- Mendelian sampling is the difference between an individual's PTA and its PA
- Sustained genetic gain under selection depends on Mendelian sampling variance (Woolliams et al., 1999)
- Increased reliabilities from genomic selection are due to better estimates of Mendelian sampling (Hayes et al., 2009)

Genomic Selection

- You do not benefit substantially from genomic selection until you have a large enough pool of genotyped animals to provide good estimates of marker effects
- Good marker effects are essential for reliable prediction of Mendelian sampling
- Mendelian sampling can be calculated directly from marker effects without the need to wait for progeny test-based evaluations

Objectives

- Describe the predicted Mendelian sampling (MS) variation in Brown Swiss, Holstein, and Jersey cattle
- Calculate selection limits based on haplotypes present in the genotyped population
- Examine adjustments to breeding values for changes in heterozygosity

Materials

- 43,382 SNP from the Illumina BovineSNP50
  - SNP solutions from the June, 2010 evaluation
- Three breeds
  - 1,455 Brown Swiss males and females
  - 40,351 Holstein males and females
  - 4,064 Jersey males and females
- Three traits
  - Daughter pregnancy rate (DPR)
  - Milk yield (Milk)
  - Lifetime net merit (NM$)

Methods

- Haplotypes imputed with findhap.f90 (VanRaden et al., 2010)
- Calculations performed with SAS 9.2
- Plots produced with R 2.10.1 and ggplot2 0.8.7 (Wickham, 2009)
- Operating system is 64-bit RedHat Enterprise Linux 5

Predicted Mendelian Sampling

Mendelian sampling was computed assuming all loci were unlinked (MS_U) or that loci were in perfect linkage (MS_C):

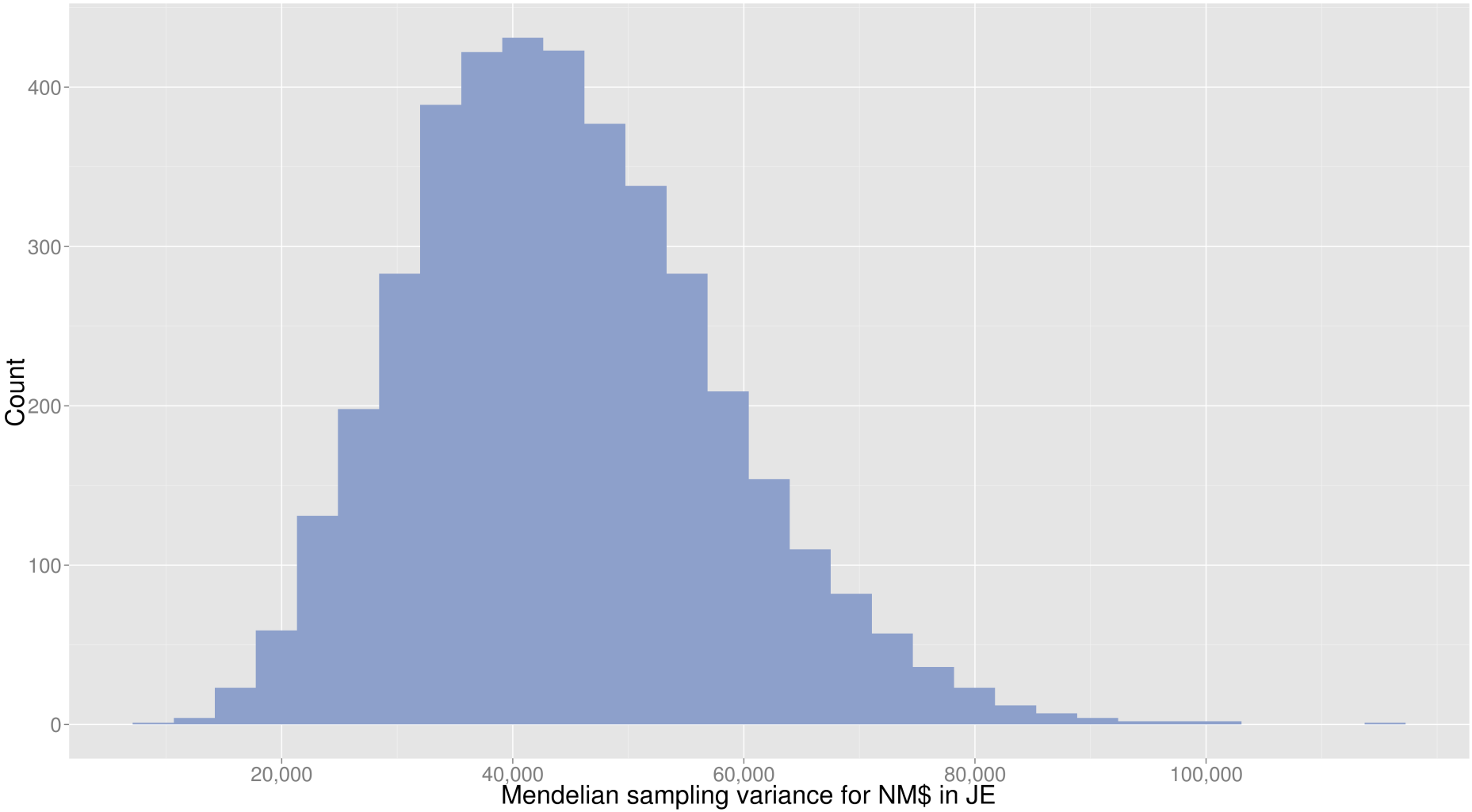
$$MS_U = \sum_{m=1}^{43{,}382} (s_m\alpha_m - d_m\alpha_m)^2$$

and

$$MS_C = \sum_{c=1}^{30} \left(\sum_{m=1}^{n_c} s_m\alpha_m - \sum_{m=1}^{n_c} d_m\alpha_m \right)^2$$

where m is a marker, s and d are sire and dam genotypes for the m^{th} marker, α_m is the allele substitution effect for the m^{th} marker, c is the chromosome, and n_c is the number of markers present on the c^{th} chromosome.

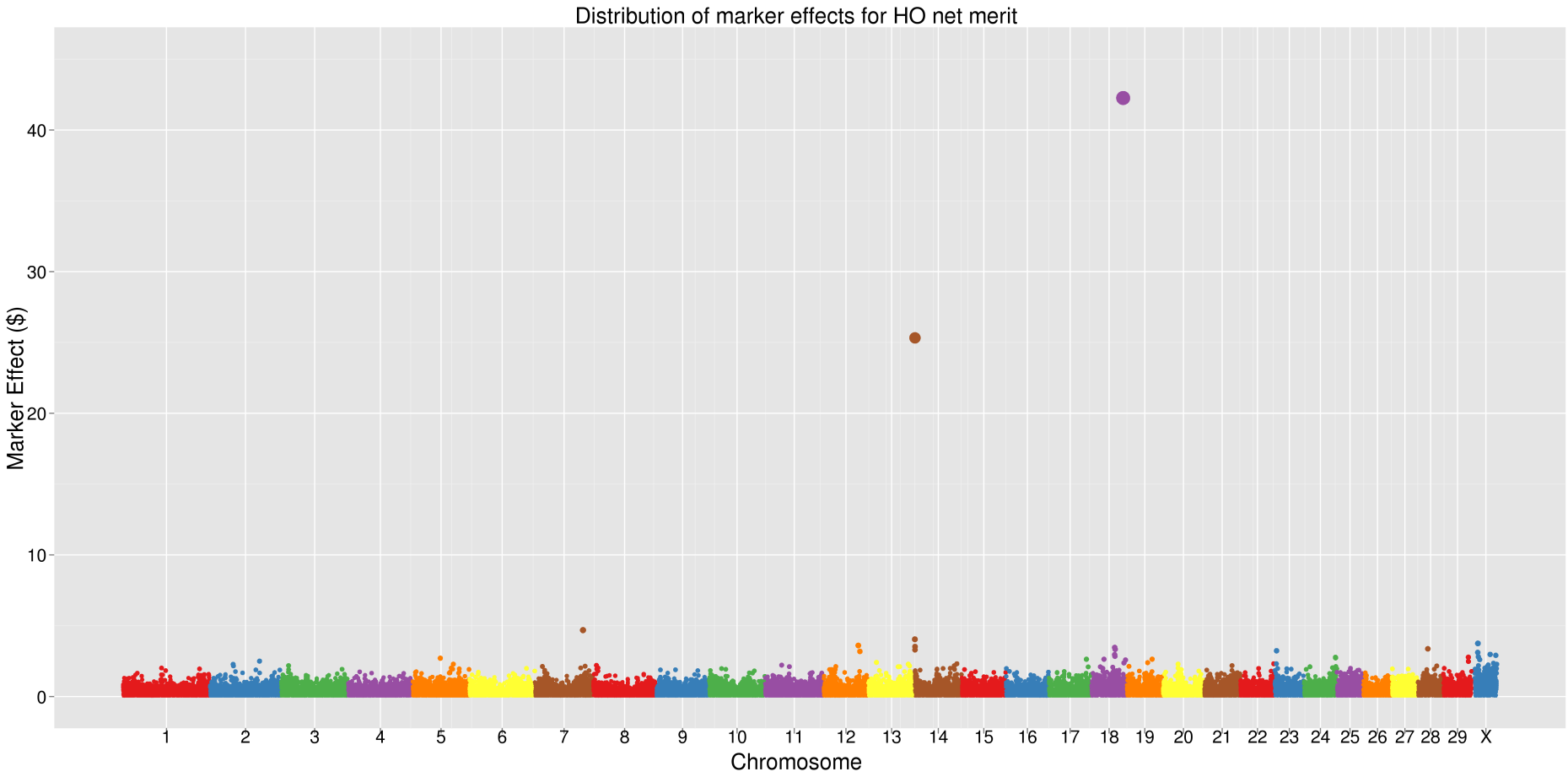
Distribution of Mendelian sampling variance



Predicted Mendelian sampling variance

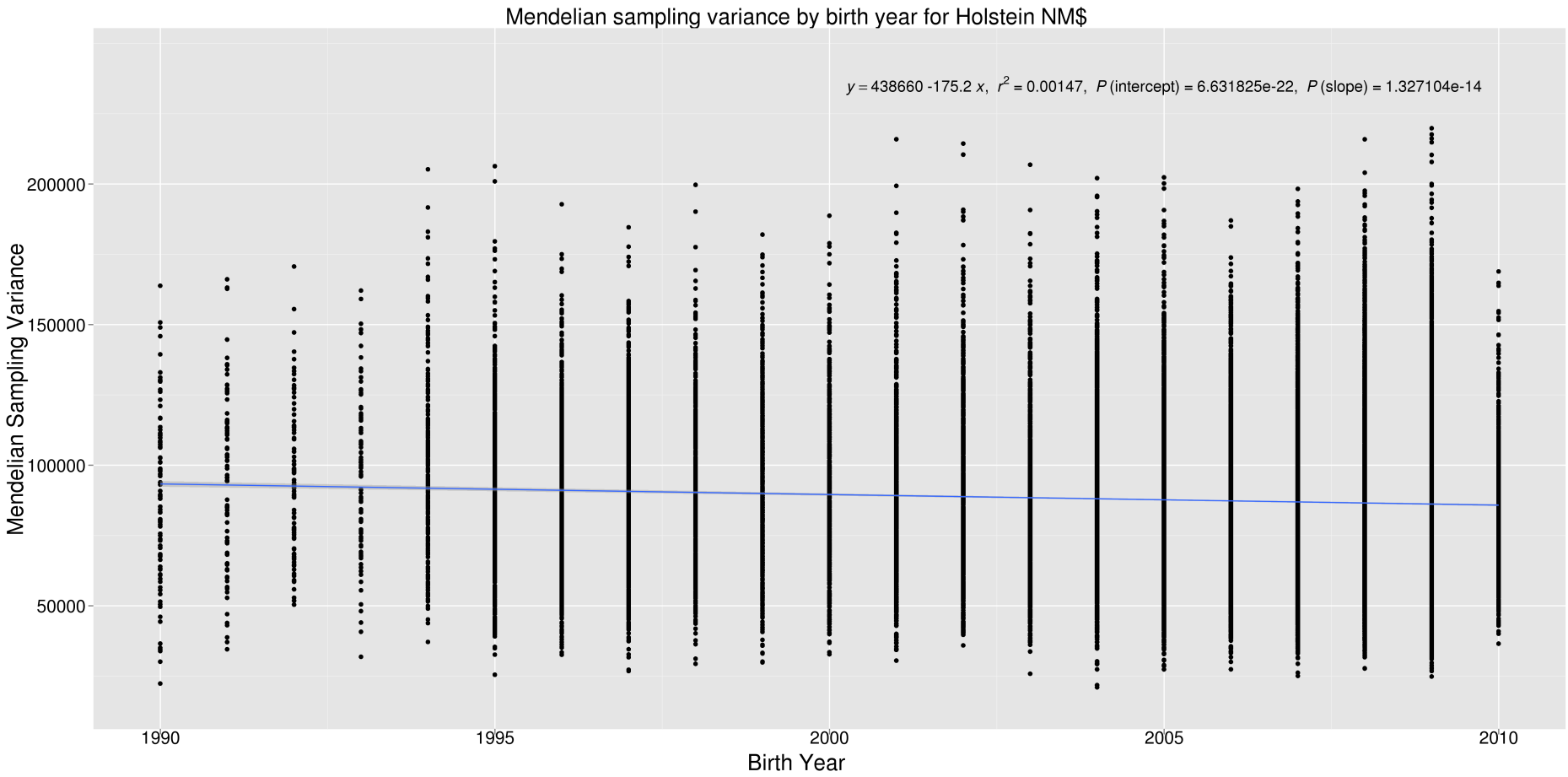
Trait	Breed	Lower	Expected	Upper
DPR	BS	0.09	1.45	1.57
	HO	0.57	1.45	4.02
	JE	0.09	0.98	1.27
Milk	BS	7,264	44,238	104,255
	HO	46,879	53,736	219,939
	JE	30,855	42,238	12,376
NM$	BS	2,539	19,602	40,458
	HO	16,601	19,602	87,449
	JE	3,978	19,602	44,552

Why are Holsteins so different?

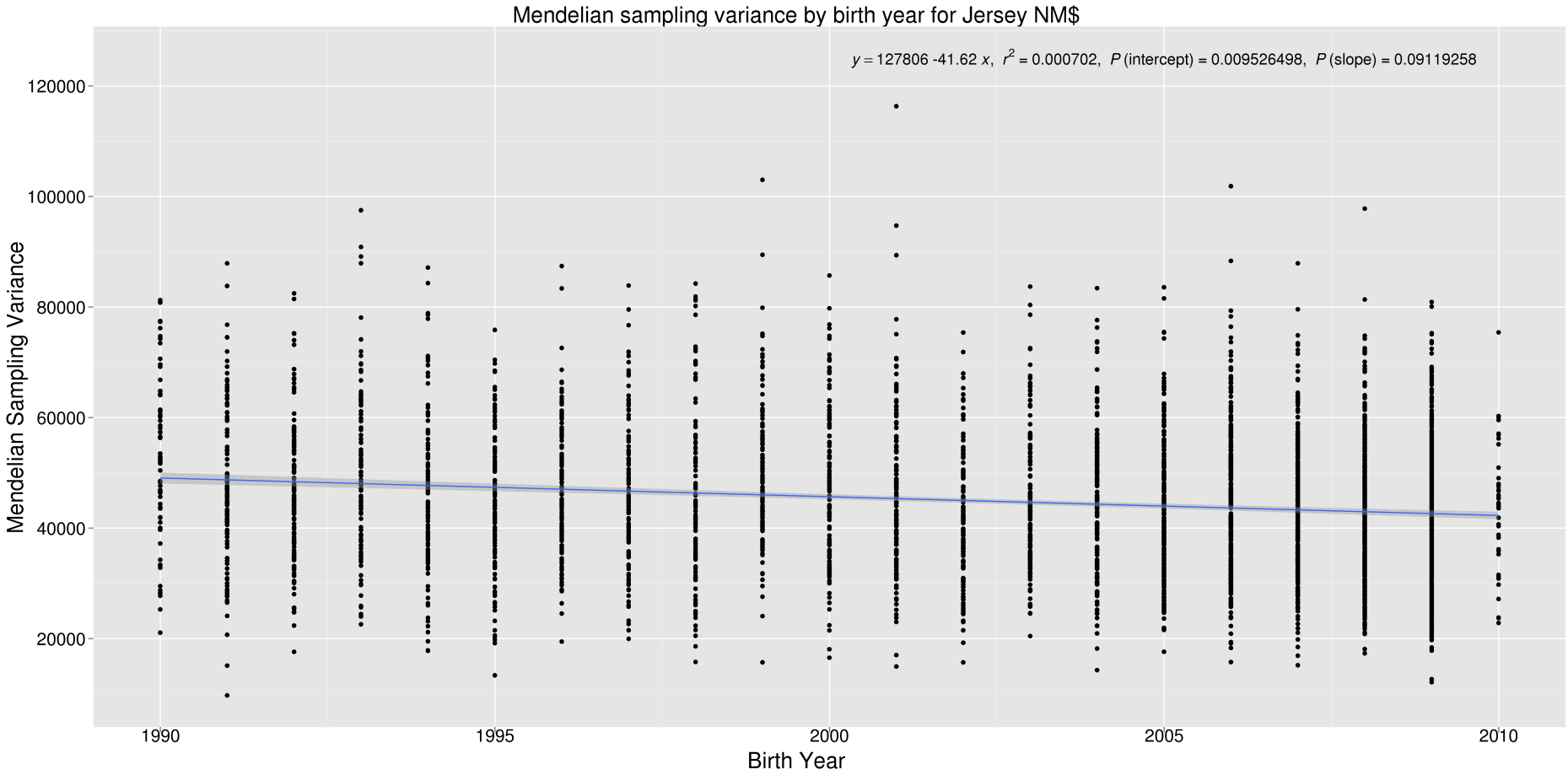


- There are many more HO haplotypes represented
- There are multiple QTL affecting NM$ in HO

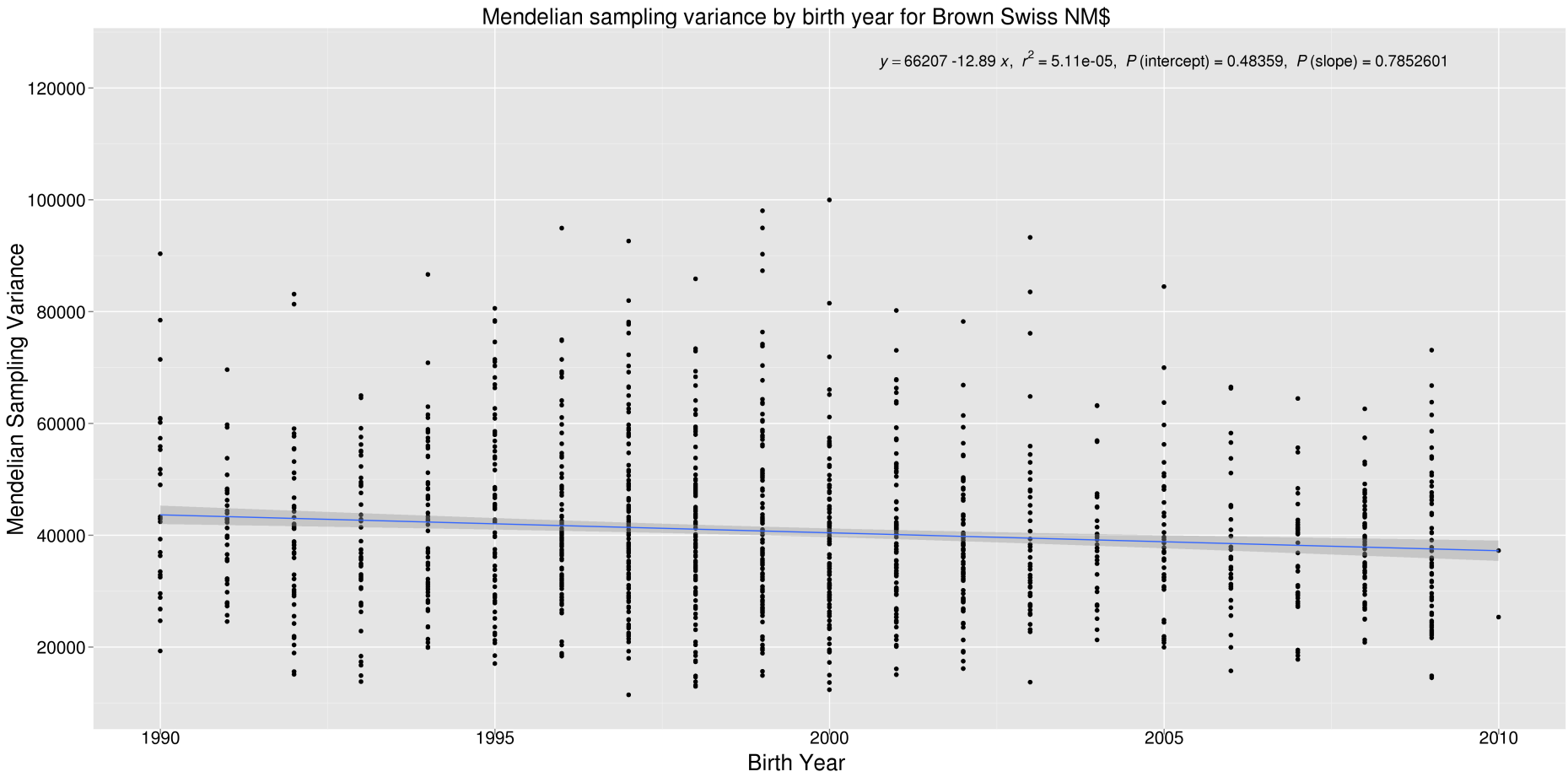
Holstein sampling variance over time (NM$)



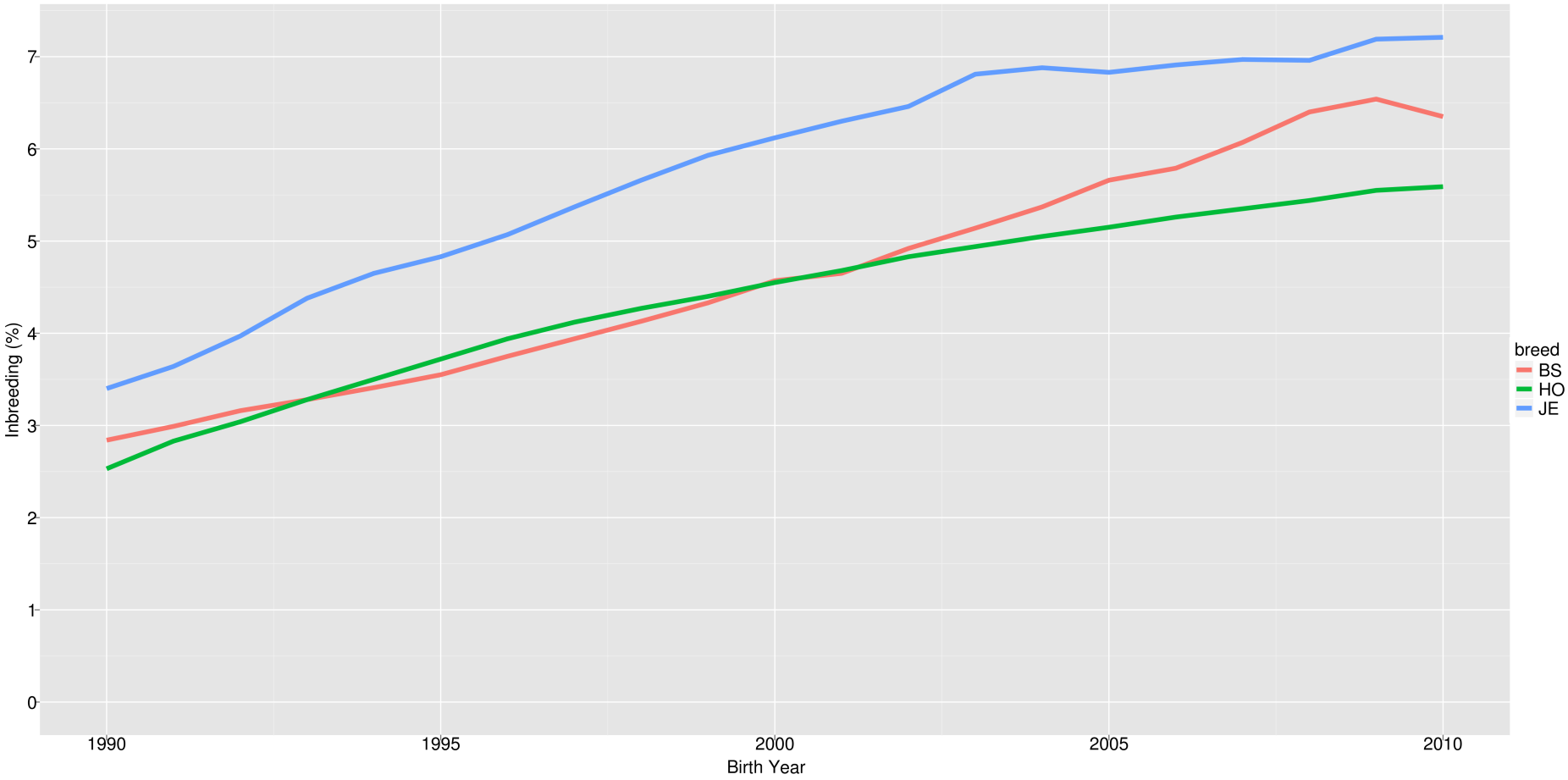
Jersey sampling variance over time (NM$)



Brown Swiss sampling variance over time (NM$)



Inbreeding over time



Correlations with inbreeding

		Lower		Upper	
Trait	Breed	Genomic	Pedigree	Genomic	Pedigree
DPR	BS	-0.73	-0.38	-0.02	0.09
	HO	-0.77	-0.40	-0.11	-0.03
	JE	-0.83	-0.53	-0.01	0.06
Milk	BS	-0.86	-0.55	-0.05	0.03
	HO	-0.12	-0.05	-0.10	-0.03
	JE	-0.01	0.03	-0.04	0.04
NM$	BS	-0.85	-0.49	0.03	0.13
	HO	-0.21	-0.12	-0.11	-0.03
	JE	-0.86	-0.53	-0.11	-0.02

Selection limits

- Lower bound found by summing the best chromosomes
- Upper bound found by summing the best alleles at each locus
- NM$ was adjusted to account for changes in heterozygosity

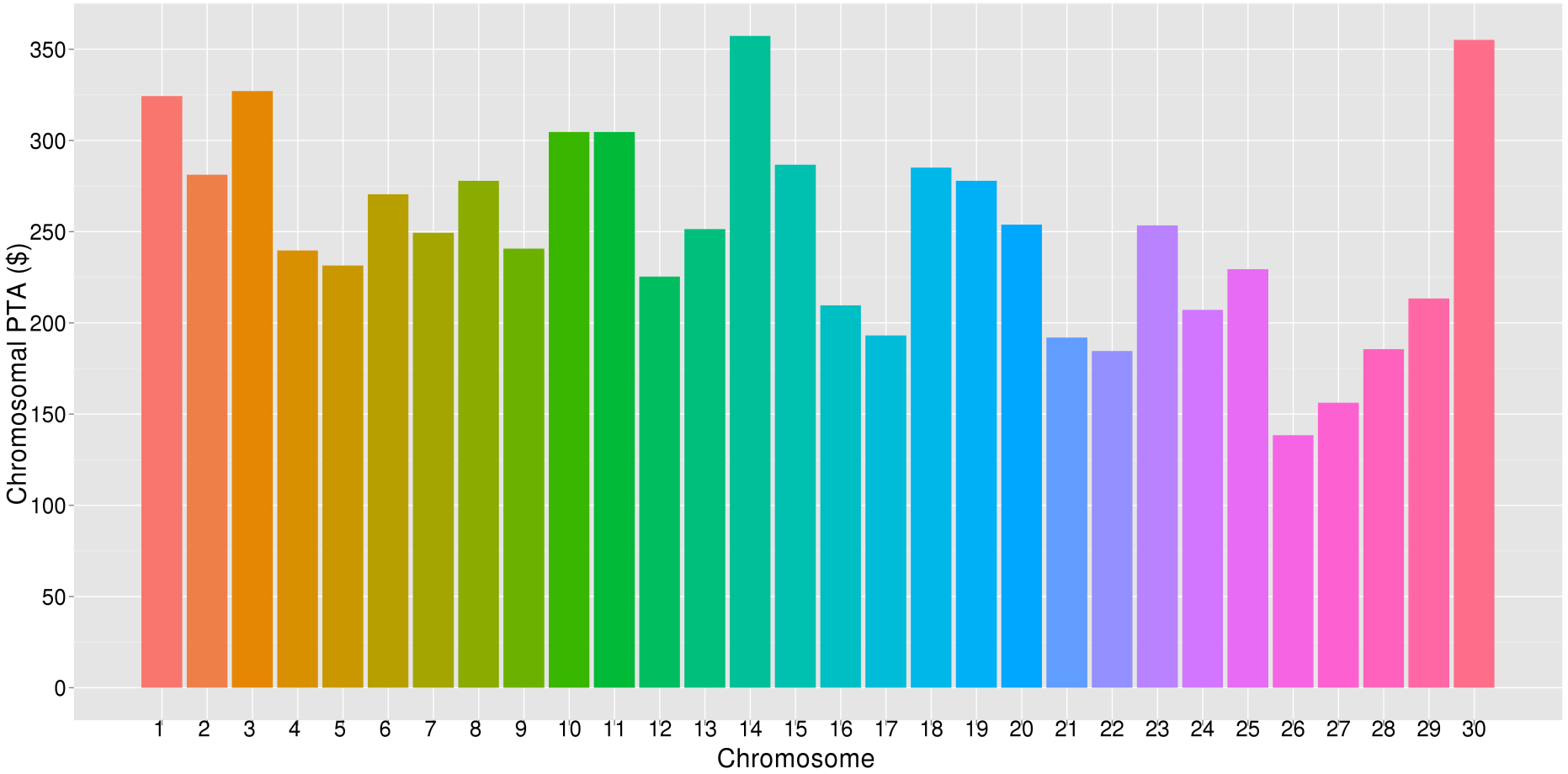
Calculated selection limits

Trait	Breed	Lower	Upper	Largest DGV
DPR	BS	20	53	8
	HO	40	139	8
	JE	19	53	5
Milk	BS	6,461	15,465	2,065
	HO	11,310	35,419	3,634
	JE	7,333	18,295	2,554
NM$	BS	3,857	9,140	1,102
	HO	7,515	23,588	2,528
	JE	4,678	11,517	1,556

Accounting for changes in heterozygosity

- Chromosomal PTA for NM$ were adjusted by adding or subtracting 6% of an additive genetic standard deviation ($11.88) per 1% change in heterozygosity (Smith et al., 1998)
- Only 4 chromosomes (BTA 6, 10, 14, and 16) differed between the adjusted and unadjusted groups

What's the best cow we can make?



A "Supercow" constructed from the best haplotypes in the Holstein population would have a PTA(NM$) of $7,515.

Conclusions

- Selection limits and sampling variances differ among breeds
- Sampling variance has decreased slightly over time
- The top animals in each breed are well below predicted selection limits
- Adjustment for heterozygosity had little effect on PTA for NM$