

Genomic evaluations and breed composition for crossbred U.S. dairy cattle

P. M. VanRaden and T. A. Cooper*

*Animal Genomics and Improvement Laboratory,
Agricultural Research Service, USDA, Beltsville, MD, USA*

paul.vanraden@ars.usda.gov

Introduction

- **Traditional evaluations are on all-breed base**
- **Pedigree breed composition (PBC) has been computed, distributed, and used to calculate heterosis since we went to the all-breed model in 2007**
- **Theoretically genomic heterosis could be computed using Genomic breed composition (GBC) instead of PBC, but only if the parents are genotyped**

Introduction

- **Genomic evaluations are separate by breed**
 - **HOL, JER, BSW, RDC, possibly GUE**
 - **Crossbreds have mixture of SNP effects**
 - **Parent averages are incomplete within breed**
- **Growing demand to genotype crossbreds**

Background

- **Harris et al. (2009) used multibreed relationship matrix to obtain GEBVs in NZL**
- **Strandén and Mäntysaari (2012) used random regressions on pedigree breed composition**
- **Olson et al (2010) estimated genomic breed composition using all markers**
- **Many researchers tested benefit of combining breeds, but few had genotyped crossbreds**

Background at AGIL

- **Breed determining SNP introduced in 2010**
 - monomorphic in 1 breed and have fewer than 30% of animals homozygous for that allele in another breed
- **Initially used to identify misidentified samples**
- **Later, used to exclude crossbreds from genomic evaluation**
- **Recently, breed SNP have been used to aid in determining breed composition**

Breed check edits

- **Exclusion of genotypes from being evaluated**
 - ▶ **Genotype was identified as being from a different breed completely**
 - ▶ **Animal had a pedigree sire or dam of another breed**
 - ▶ **Based on breed SNP, animal was >40% non breed of evaluation for high density genotypes or >~20% non breed of evaluation for low density genotypes**

Goals

- **Compare imputation strategies for crossbreds**
- **Estimate each animal's breed composition**
- **Examine categories of animals genotyped**
- **Develop genomic evaluations for crossbreds**
- **Reduce breeder's need to guess before genotyping if an animal will pass breed check edit and be evaluated**

Crossbred genotypes

- **6,296** animals that had failed breed check test
- Expected to be crossbred (some wrong breed)
 - ▶ Only 33 animals had $\geq 50K$ genotypes
 - ▶ The other 99.5% had lower density chips containing **3K** to **13K** usable markers
 - ▶ Imputed to **60K** for evaluation
- Genotypes worth about **\$300,000**

Imputation

- **Part:** Crossbred animals and their genotyped ancestors
 - ▶ The pedigree file for the 6,296 crossbreds included 72,939 ancestors, but only 3,119 were genotyped
 - ▶ Those were included to improve imputation.
- **Full:** All 828,754 purebred and crossbred animals from March 2015 imputed together

Imputation time

- **Part**

- ▶ 9,425 animals
- ▶ 6 processors
- ▶ ~30 min

- **Full**

- ▶ 828,754 animals
- ▶ 6 processors
- ▶ ~22 days

Imputation comparison

- Correlations between **part** and **full** GBC were:
 - ▶ 0.997 for HO, 0.998 for JE, 0.998 for BS, and 0.990 for AY breed fractions
- Similar estimates of GBC from part and full data imply that genotypes were imputed consistently even when fewer purebred animals were included
- GBC can be computed much more quickly from part than full data

Adjusted Breed Composition (ABC)

- Adjust the GBC mean by subtracting from each GBC value the sum of GBC / number of breeds (Nbrd)
- Obtain the range of the adjusted GBC from the maximum and minimum adjusted breed GBC
- Adjust the SD if any adjusted GBC are > 100 or < 0 , using $\max(\text{largest adjusted GBC} - 100 / \text{Nbrd}) / [100 * (1 - 1 / \text{Nbrd})]$, or $(100 / \text{Nbrd} - \text{smallest adjusted GBC}) / (100 / \text{Nbrd})$
- Obtain $ABC = 100 / \text{Nbrd} + (\text{adjusted GBC} - 100 * \text{Nbrd}) / \text{SD}$.

Adjusted Breed Composition (ABC)

- **Genomic evaluations for crossbreds can be computed by weighting the marker effects for separate breeds by ABC instead of PBC as in some previous reports**
- **Marker effects must be computed on the all-breed base rather than within-breed bases**
- **An advantage of ABC over PBC is that pedigrees are often incomplete or inaccurate for crossbred animals**

Comparison of genomic, adjusted and pedigree breed composition

Statistic	HO	JE	BS	AY
Average GBC	85.7	11.5	2.3	0.5
Average ABC	84.3	11.7	2.7	0.9
Average PBC	85.7	11.5	2.2	0.5
Minimum GBC	-8	-3	-4	-3
Minimum ABC	0	0	0	0
Maximum GBC	106	108	104	110
Maximum ABC	100	100	100	100
Corr(GBC,ABC)	.999	.999	.999	.997
Corr(GBC,PBC)	.996	.996	.998	.990

Categories removed by breed check

Number	Description	Edit
733	JER x HOL F1	> 40% of both
55	BSW x HOL F1	> 40% of both
2300	HOL backcrosses	> 67% and < 90% HOL
2026	JER backcrosses	> 67% and < 90% JER
27	BSW backcrosses	> 67% and < 90% BSW
502	Other crosses	Not in groups above
1024	Purebreds	> 90% of ID breed
133	Wrong breed	< 20% of ID breed

Crossbred example

- **F1 JE X HO (50/50) - Dam**
 - ▶ **61% JE 38% HO**
- **JE X HO X JE (75/25) – 16 progeny**
 - ▶ **81% JE 18% HO**
 - ▶ **JE% ranged between 72% - 91%**
- **JE X HO X JE X JE (87.5/12.5) – 291 grand progeny (limited to 50 animals/progeny)**
 - ▶ **89% JE 10% HO**
 - ▶ **JE% ranged between 76% - 97%**

gPTA methods for crossbreds

- **Convert traditional evaluations of all purebred genotypes to the all breed base**
- **Calculate individual breed SNP effects**
- **Apply individual SNP effects to all crossbred animals**
- **Combine individual breed gptas weighted by breed composition**

gPTA example for milk

- JE X HO Dam (61/38)
- JE milk gPTA – (-1,354 lbs.)
- HO milk gPTA – (-3,611 lbs.)
- Combined all breed gPTA – (-2,555.5 lbs.)
- JE scale gPTA – (261.7 lbs.)
- Traditional JE evaluation – (1,182 lbs. 89 rel)

Purebred and Crossbred gPTA comp.

Breed	Corr(PB_{gpta}, CB_{gpta})	Corr(PBC, GBC)
Holstein	0.85	0.86
Jersey	0.62	0.87
Brown Swiss	0.93	0.78
Ayrshire	0.97	0.72

Conclusions

- **Imputing with only ancestors or with all animals gave similar GBC**
- **Adjusted GBC sums to 100, min 0, max 100**
- **Most crossbred animals were HOL or JER backcrosses**
- **Genomic evaluations were computed on all-breed scale, weighting SNP effects by GBC**
- **Purebreds not yet affected by other breeds**

Thank You!