



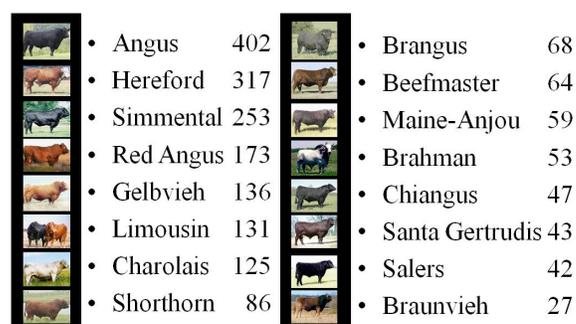
Larry Kuehn, PhD  
 Research Geneticist  
 U.S. Meat Animal Research Center  
 P.O. Box 166, Clay Center, NE 68933  
 Ph: (402) 762-4352 Fax: (402) 762-4173  
 Email: [Larry.Kuehn@ars.usda.gov](mailto:Larry.Kuehn@ars.usda.gov)

## Relationships of Beef Breeds Using the 50K Chip

Written by Larry Kuehn

Breeds of cattle were established primarily through selection based on physical appearance (color patterns, stature), easily observed performance (milk production, muscling), or due to geographic constraints. These various selection pressures produced the uniform populations of cattle we define as breeds. Selection occurred over several generations; hence, breeds of cattle also diverged genomically due to selection pressure, inbreeding, and drift. Based on genomic differences, we should be able to **determine the breed composition of animals of unknown origin**. Knowledge of breed composition of cattle would be useful for predicting heterosis, evaluating adaptability to production environments (e.g., Brahman influence), and sorting animals into management groups. Breed composition is also useful for tracing an animal's history from birth for the purposes of tracking disease transmission and sources of contamination in meat. **Therefore, our objective was to determine whether we could accurately predict breed percentage in crossbred animals using the 2,000 bull project bulls as a genomic breed reference pool.**

**2,000 Bull Project:** In order to test whether breed composition of crossbred animals could be accurately predicted, we needed a reference population of bulls from several pertinent beef cattle breeds. The 2,000 bull project sample of bulls was ideal for this purpose. The U.S. Meat Animal Research Center genotyped over 2,000 bulls on the Illumina BovineSNP50 in collaboration with the 16 largest U.S. breed associations that have a national cattle evaluation (Figure 1). The purpose of this project was to evaluate the effectiveness of producing marker tests from a discovery population (USMARC) and exporting these genomic predictions to an industry representative set of animals. However, because the bulls were a **highly representative sample of the major beef breeds in the United States**, they were also an excellent resource to estimate breed allele frequencies for the Illumina BovineSNP50 markers in the major U.S. beef breeds. These allele frequencies were used in a statistical model to determine the percentage breed makeup of a sample of crossbred steers and heifers. Our sample of crossbreds were part of the USMARC Germplasm Evaluation Program. Specifically, they were 2-, 3-, and 4-breed crosses of Angus, Hereford, Red Angus, Charolais, Gelbvieh, and Shorthorn. We tested whether their pedigree based breed frequency was accurately determined using the reference allele frequencies from the 2,000 bull project. These crossbred animals were already genotyped using the Illumina BovineSNP50 for genomic discovery work.



2026  
 Figure 1. Number of bulls sampled in the USMARC 2,000 Bull Project.

**Breed Relationships/Distances:** Based on the frequencies of each marker for each breed in the 2,000 Bull Project, breeds can be shown as genetically distant from one other (Figure 2).

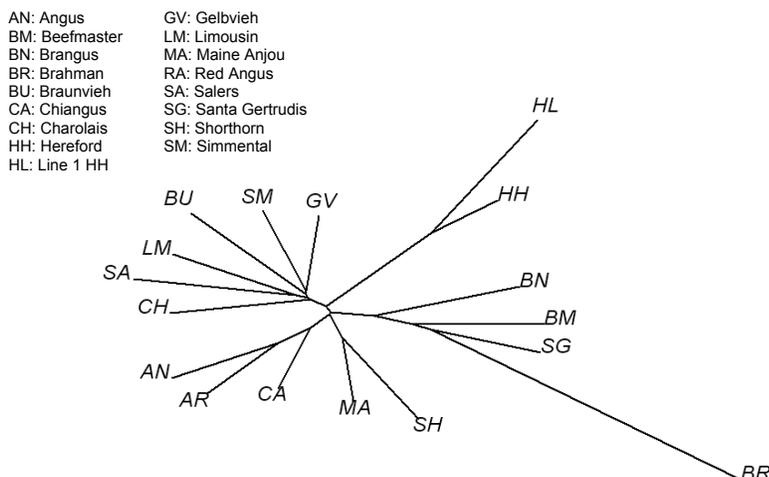


Figure 2. Approximate genetic distance between breeds using data from the 2,000 Bull Project.

Although some of the distances are inflated due to biases from the discovery source of the DNA markers on the 50K chip (e.g., Hereford origin vs. Angus origin), these distances can be exploited to determine breed composition of animals with unknown pedigree. In general, Brahman-influenced breeds segregated from European breeds. Hereford was the most distant European breed. Line 1 Herefords (from the Ft. Keogh Livestock and Range Research Laboratory; Miles City, MT) were even more distantly

removed from the other breeds, as would be expected. Most continental breeds grouped separately. Angus and Red Angus were the most closely related breeds.

**Results:** Breed composition could be predicted accurately relative to the pedigree based estimate for each breed. For instance, the percentage of Hereford relative to the amount of Hereford in the animal was estimated with an accuracy of approximately 95% ( $R^2 = .92$ ). Other breeds were predicted with slightly lower (1-2%) accuracy but still reasonably well with the exception of Red Angus and Angus. Because of the close genetic distance between these two breeds (Figure 2), representing the relatively recent divergence between the breeds, our statistical model tended to predict animals with either Red Angus or Angus in their pedigree as having high proportions of both breeds. As a solution, we combined the two breeds (Figure 3), and pedigree percentages were predicted with much greater accuracy. It is important to note that the pedigree percentage of each breed is not completely accurate due to genomic segregation; an animal whose pedigree would indicate it being  $\frac{1}{4}$  Angus,  $\frac{1}{4}$  Hereford,  $\frac{1}{4}$  Simmental, and  $\frac{1}{4}$  Charolais will actual vary in their genomic proportion of each of these breeds. Based on these genomic predictions, we estimate some of these real variations relative to pedigree based breed composition.

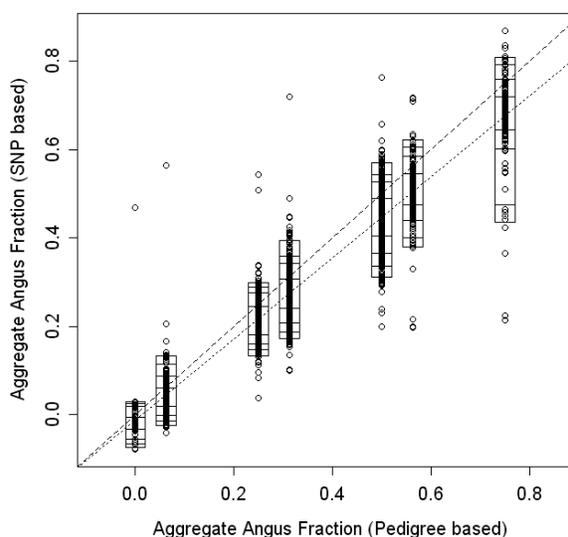


Figure 3. Prediction of percentage of Angus or Red Angus (Aggregate Angus) in crossbred animals relative to the pedigree based percentage of either breed.