

## 224. Estimation of old and new inbreeding and their effects on growth traits using pedigree and genomic information

P. Sumreddee<sup>1</sup>, A. Ling<sup>2</sup>, S. Toghiani<sup>3</sup>, E.H. Hay<sup>3</sup>, E. Hartono, S.E. Aggrey<sup>4</sup> and R. Rekaya<sup>2</sup>

<sup>1</sup>The Bureau of Biotechnology in Livestock Production, Department of Livestock Development, Pratuntani, 12000, Thailand; <sup>2</sup>Department of Animal and Dairy Science, The University of Georgia, Athens, GA 30602, USA; <sup>3</sup>USDA Agricultural Research Service, Miles City, MT 59301, USA; <sup>4</sup>Department of Poultry Science, The University of Georgia, Athens, GA 30602, USA; rrekaya@uga.edu

### Abstract

Inbreeding caused by the mating of animals related through a recent ancestor is expected to have more pronounced effects on traits than old inbreeding. Estimating these effects requires a clear definition of new (recent) and old (ancient) inbreeding. Current methods to classify inbreeding are largely based on heuristic criteria. The subjectivity of the thresholds used to distinguish between ancient and recent inbreeding has led to inconsistency in results. To remedy this situation, we developed a new method to classify inbreeding based on a search algorithm driven by the hypothesis that new inbreeding tends to have a more pronounced effect on traits. Four growth traits were used to quantify the effects of old and new inbreeding in a Hereford herd. Thresholds to classify inbreeding varied across traits and sources of information. Using pedigree information, inbreeding generated in the last 10 to 11 generations was considered recent. When genomic information was used, thresholds ranged between 4 to 7 generations. In general, the proposed method seemed to have similar performance compared to existing approaches using ROH segments and to be slightly better using pedigree information.

### Introduction

Although genealogy information has been traditionally used to estimate individual level of inbreeding ( $F$ ), the pedigree-based measure of inbreeding ( $F_{ped}$ ) is highly influenced by the quality of the pedigree and its completeness (Cassell *et al.*, 2003). The availability of high-density single nucleotide polymorphisms (SNPs) provides an alternative to quantify the genome-wide homozygosity through runs of homozygosity (ROH) segments. Several studies showed that inbreeding coefficients calculated based on ROH ( $F_{ROH}$ ) are more accurate, and they are more powerful in assessing the effects of inbreeding (Keller *et al.* 2011). The distribution of ROH segment length could be a valuable resource to distinguish between new and old inbreeding. Long ROH segments are likely produced by recent common ancestors (recent inbreeding), whereas shorter segments are generated many generations in the past (older inbreeding). Knowing the distribution of ROH segments will provide a useful tool to understand the purging of deleterious alleles and to assess inbreeding depression. Several approaches based on tracing the pedigree back to a pre-specified number of generations (threshold) to identify recent inbreeding were proposed (Doekes *et al.* 2019). Unfortunately, the threshold is arbitrarily set, and it is likely to be population or even breed dependent. Additionally, multiple studies have recently attempted to discriminate between recent and ancient inbreeding using genomic autozygous segments. Several approaches have been proposed to categorize ROH segment length into different age-classes based on arbitrary thresholds. These approaches to discriminate between recent and old inbreeding have led to inconsistent results. It was hypothesized that recent inbreeding has a greater impact on traits compared to inbreeding due to distant common ancestors. Ancient inbreeding allows selection to have enough time to purge deleterious recessive mutations. The ability to identify inbreeding classes associated with a greater impact on phenotypes could be useful not only for quantifying its impact on phenotypes, but also for better herd management. The objectives of this study are to: (1) develop a new method to distinguish between recent and old inbreeding using pedigree and ROH information; and (2) compare the performance of the proposed approach with existing methods.

## Materials & methods

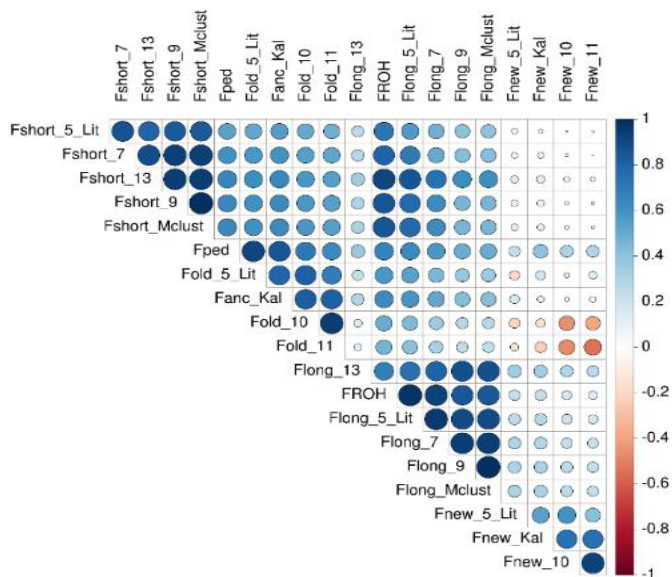
Data used in this study was collected from purebred Line 1 Hereford cattle herd maintained at USDA-ARS, Fort Keogh Livestock and Range Research Laboratory, Miles City, MT. Although the population was historically selected for increased yearling weight, since 2011, the selection was primarily focused on improving calving ease. A complete pedigree comprising 10,478 animals was used. Pedigree analyses of the Line 1 population showed relatively complete, deep genealogical records with an average of 17.15 equivalent complete generations for all animals. A subset of 797 animals born between 1953 and 2016 was genotyped with different density panels (3k to 50k markers). Imputed genotypes for 30,220 SNPs on 29 autosomal chromosomes for 785 animals, generated as described in Sumreddee *et al.* (2019), were used.

**Grid search approach to discriminate between old and recent inbreeding.** We developed a grid search algorithm to determine the appropriate threshold for discriminating between new and old inbreeding based on the basic assumption that recent autozygosity has more harmful effects than old inbreeding (Sumreddee *et al.*, 2019). Thus, the threshold separating old and recent inbreeding can be identified by comparing their effects on traits. By varying the threshold to cluster ROH into short and long segment classes, the optimal cut-off-point (threshold) can be inferred. Standardized coefficients for new and old inbreeding based on varying thresholds (ROH length or number of generations) were included as covariates in the model for analyzing inbreeding depression. The magnitudes of the estimated inbreeding effects associated with partial inbreeding were then compared, and a threshold for recent inbreeding was determined. The threshold was defined as the minimum number of ancestral generations ( $t$ ) or the longest length  $m$  (in Mb) used to classify ROH segments that resulted in a larger negative effect of recent inbreeding compared to its old counterpart. The resulting new and old inbreeding are denoted hereafter as  $F_{\text{new } t}$  and  $F_{\text{old } t}$  using pedigree information and  $F_{\text{long } m}$  and  $F_{\text{short } m}$  using ROH segments. The proposed method was compared to several existing approaches used to partition inbreeding into new and old classes.

## Results

Figure 1 presents the correlations between different estimates of new and old inbreeding. In general,  $F_{\text{ped}}$  has moderate to high correlation ( $r=0.63$  to  $0.91$ ) with old age-class inbreeding (e.g.  $F_{\text{old } t}$ ,  $F_{\text{anc Kal}}$ ). However, it was weakly correlated ( $r=0.24$  to  $0.41$ ) with recent inbreeding (e.g.  $F_{\text{new } t}$ ,  $F_{\text{new Kal}}$ ). Different patterns were observed between  $F_{\text{ROH}}$  and partial inbreeding coefficients derived based on different definitions of ROH segment length classifications.  $F_{\text{ROH}}$  was in general highly correlated with both short and long ROH-based inbreeding coefficients ( $r=0.69$  to  $0.98$ ). However, high correlations were observed only between  $F_{\text{ROH}}$  and old inbreeding derived based on pedigree ( $r=0.47$  to  $0.62$ ). Correlation between  $F_{\text{ROH}}$  and pedigree based new inbreeding ranged between 0.15 to 0.24. The higher correlations of both  $F_{\text{ped}}$  and  $F_{\text{ROH}}$  with pedigree based old inbreeding seem to suggest that the majority of autozygosity in the used pedigree is ancient. Moderate negative correlations ( $r=-0.45$  to  $-0.55$ ) were observed between pairs of new and old inbreeding based on pedigree using our proposed thresholds (10 and 11 generations), while almost no linear relationships were observed between pairs of new and old inbreeding using a 5 generations literature threshold ( $r=-0.18$ ) and between Kalinowski's new and ancestral inbreeding ( $r=0.09$ ).

**Inbreeding depression based on new and old pedigree-based inbreeding.** The effects of new and old inbreeding based on pedigree information were present in Table 1. Using the proposed method, the threshold for discriminating between new and old inbreeding was 10 (for BW, WW and YW) and 11 (for ADG) generations. An additional 1% of new inbreeding resulted in a loss of 1.78 kg and 8 gram/day in YW and ADG, respectively. Old inbreeding also showed a significant negative effect of 5.5 gram/day on ADG but was not significant for YW. All growth traits consistently showed a larger negative effect of new inbreeding compared to its old counterpart. The proposed method yielded consistent results with those using the ancestral inbreeding approach ( $F_{\text{new Kal}}$  and  $F_{\text{anc Kal}}$ ) based on Kalinowski *et al.* (2000) in terms



**Figure 1.** Correlations between different estimates of inbreeding coefficients with hierarchical clustering order. Blue and red circles represent positive and negative correlations, respectively. Color intensity and the size of the circle are proportional to the correlation coefficients.  $F_{ped}$  = pedigree-based inbreeding from the full pedigree;  $F_{old\_t}$ ,  $F_{new\_t}$  = old and new inbreeding based on tracing the pedigree back to  $t$  ancestral generations;  $F_{old\_5\_Lit}$  and  $F_{new\_5\_Lit}$  = old and new inbreeding based on 5 generations for recent generations according to the literature;  $F_{anc\_Kal}$ ,  $F_{new\_Kal}$  = ancestral and new inbreeding based on Kalinoski *et al.* (2000);  $F_{ROH}$  = genome-wide ROH-based inbreeding;  $F_{short\_Mclust}$  and  $F_{long\_Mclust}$  = inbreeding due to short and long ROH based on model-based clustering method;  $F_{short\_5\_Lit}$  and  $F_{long\_5\_Lit}$  = inbreeding due to short and long ROH based on 5 Mb as defined from the literature;  $F_{short\_m}$  and  $F_{long\_m}$  = inbreeding due to short and long ROH determined using proposed method with  $m$  Mb threshold.

**Table 1.** Estimates of the regression coefficients (SE)<sup>1</sup> for genome-wide and partial inbreeding coefficients based on pedigree using the proposed and existing methods.

Measure of inbreeding <sup>2</sup>		BW <sup>3</sup>	WW	YW	ADG
Genome-wide	$F_{ped}$	-0.018 (0.065)	-0.137 (0.312)	-1.170 (0.518)	-0.0065 (0.0016)
	$F_{ROH}$	-0.043 (0.040)	-0.168 (0.211)	-0.890 (0.364)	-0.0050 (0.0013)
Existing	$F_{old\_5\_Lit}$	-0.017 (0.075)	-0.233 (0.361)	-1.436 (0.592)	-0.0070 (0.0017)
	$F_{new\_5\_Lit}$	-0.020 (0.105)	0.106 (0.553)	-0.405 (0.972)	-0.0043 (0.0038)
	$F_{anc\_Kal}$	-0.038 (0.112)	0.468 (0.550)	-0.265 (0.973)	-0.0059 (0.0032)
	$F_{new\_Kal}$	-0.152 (0.401)	-3.701 (2.059)	-7.282 (3.632)	-0.0276 (0.0129)
Proposed	$F_{old\_10}$	0.0007 (0.079)	0.257 (0.383)	-0.693 (0.640)	-
	$F_{new\_10}$	-0.038 (0.081)	-0.637 (0.421)	-1.779 (0.707)	-
	$F_{old\_11}$	-	-	-	-0.0055 (0.0018)
	$F_{new\_11}$	-	-	-	-0.0080 (0.0020)

<sup>1</sup> Significant effects are indicated by a bold font. The estimates of inbreeding effects from the proposed method are shown only for specific traits due to a trait-specific threshold to define new inbreeding.

<sup>2</sup>  $F_{old\_5\_Lit}$  and  $F_{new\_5\_Lit}$  = old and new inbreeding based on 5 generations for recent generations according to the literature;  $F_{anc\_Kal}$  and  $F_{new\_Kal}$  = ancestral and new inbreeding based on Kalinoski *et al.* (2000);  $F_{old\_t}$  and  $F_{new\_t}$  = old and new inbreeding based on tracing the pedigree back to  $t$  ancestral generations determined using the proposed method.

<sup>3</sup> BW = birth weight (kg); WW = weaning weight (kg); YW = yearling weight (kg); ADG = average daily gain (kg/d)

of inbreeding depression, especially for YW and ADG. Kalinowski's new inbreeding ( $F_{anc\_new}$ ) was more harmful than ancestral inbreeding ( $F_{anc\_Kal}$ ) for most of the growth traits; however, the significant effects were found only for YW and ADG. Using the literature-defined threshold of 5 generations to separate new and old inbreeding only the latter ( $F_{old\ 5\ Lit}$ ) showed a significant unfavorable effect on both YW and ADG. Although a greater harmful impact would be expected for a very recent inbreeding (5 generations), this small threshold value favored older inbreeding to capture the majority of the genome-wide inbreeding. In fact, about 93% of the total inbreeding ( $F_{ped}$ ) was captured by old inbreeding when the threshold was fixed at 5 generations. Consequently, magnitudes of effects of  $F_{old\ 5\ Lit}$  and  $F_{ped}$  on YW and ADG were similar. In this respect, although the old inbreeding is likely to have a small deleterious effect, the accumulation of a relatively larger number of old inbreeding variants could dominate the overall inbreeding depression. Likewise, a smaller effect of new inbreeding obtained using 5 recent generations ( $F_{new\ 5\ Lit}$ ) could be due to its minimal contribution (about 7%) to the total inbreeding. These results are concerning regarding the potential bias in the estimation of inbreeding depression using an arbitrary threshold.

**Inbreeding depression based on recent and ancient ROH-based inbreeding.** Partitioning of ROH segments into long and short classes offers the opportunity to assess the unfavorable effects of recent ( $F_{long}$ ) and ancient ( $F_{short}$ ) inbreeding. Using the proposed method, both inbreeding coefficients had negative effects on growth performances; however, the effects are statistically significant only for YW and ADG. Inbreeding based on ROH segments longer than 7 Mb ( $F_{long\_7}$ ) estimated using the proposed method resulted in a significant decrease in ADG (4.4 gram/day per 1% increase in inbreeding). Overall, the differences in the effects of new and old inbreeding based on the length of ROH segments were less evident compared to those observed using pedigree information. The proposed method seemed to have similar performance compared to existing approaches using ROH segments and to be slightly better using pedigree information.

## References

- Cassell B.G., Adamec V. and Pearson R.E. (2003) *J. Dairy Sci.* 86(9): 2977-2983. [https://doi.org/10.3168/jds.S0022-0302\(03\)73895-8](https://doi.org/10.3168/jds.S0022-0302(03)73895-8)
- Doekes H.P., Veerkamp R.F., Bijma P., De Jong G., Hiemstra S.J. *et al.* (2019) *Genetics Selection Evolution* 51(1): 54. <https://doi.org/10.1186/s12711-019-0497-z>
- Kalinowski S.T., Hedrick P.W. and Miller P.S. (2000) *Conservation Biology* 14(5): 1375-1384. <https://doi.org/10.1046/j.1523-1739.2000.98209.x>
- Keller M.C., Visscher P.M. and Goddard M.E. (2011) *Genetics* 189(1): 237-249. <https://doi.org/10.1534/genetics.111.130922>
- Sumreddee P., Toghiani S., Hay E.H., Roberts A., Agrrey S.E. *et al.* (2018) *J. Anim. Sci.* 97(1): 1-18. <https://doi.org/10.1093/jas/sky385>