

Persistence of Linkage Disequilibrium Phase among Five South African Beef Cattle Populations

S. O. Makina^{*†}, E. vanMarle-Köster[†], F.C. Muchadeyi[‡] & A. Maiwashe^{*}

^{*}Agricultural Research Council-Animal Production Institute, Private Bag X 2, Irene 0062, South Africa

[†]Department of Animal and Wildlife Sciences, University of Pretoria, Private Bag X 20, Hatfield 0028, South Africa

[‡]Agricultural Research Council-Biotechnology Platform, Private Bag X 5, Onderstepoort, 0110, South Africa

Abstract Text: Genomic selection relies on the assumption that phases of linkage disequilibrium between markers and quantitative trait loci are the same in selection candidate and the reference population. If marker phase are correlated across multiple breeds, it could be possible to pool several breeds into one common reference population. This study investigated persistence of linkage disequilibrium phase among five South African beef cattle including Afrikaner (44), Nguni (56), Drakensberger (47) Bonsmara (46) and Angus (31). Consistencies of SNP phase between breed pairs were correlated using the signed r values. The correlation of r values between populations for the same marker pairs did not reach high levels. The highest correlation was observed between Nguni-Bonsmara pair (0.6) for marker pairs separated by 10 kb. This suggested the necessity of breed-specific reference populations or the need to include adequate representation of each breed in the reference population if a multi-breed reference is to be used.

Keywords: Genomic selection

Introduction

Genomic selection which refers to selection decisions based on genomic estimated breeding values (GEBV) offer many advantages with regard to improving quantitative traits (Meuwissen et al., 2001). Genomic selection calculates GEBV based on large number of DNA markers such as single nucleotide polymorphism (SNP) utilizing linkage disequilibrium (LD) between markers and quantitative traits loci (QTL). The effects of these markers are estimated in a reference population where animals are both phenotyped and genotyped. In following generations animals are genotyped for markers effects and the effects of genotypes are summed over the whole genome to obtain GEBV (Goddard & Hayes, 2007). Recently the limitations of large number of markers have been overcome in cattle, through the availability of Bovine SNP50K and 770K BeadChip. SNP50K BeadChip was found to be useful for genome wide scans studies on South African cattle breeds (Qwabe et al., 2013).

One major limitation to the implementation of genomic selection in South African cattle populations would be the availability of large reference populations that have reliable phenotypic records. Majority of cattle breeds in South African have limited number of animals that have

reliable phenotypic records and this will limit the number of individuals to form reference populations. One possible solution could be to use a breed with large reference population to predict GEBV in other breeds, as genomic selection rely on the assumption that phases of linkage disequilibrium between markers and QTL are similar in selection candidate and reference population. Goddard et al. (2006) suggested that across breed genomic evaluations could increase the accuracy of genomic predictions especial for breeds with small population size. Therefore, the objective of this study was to determine the persistence of LD phase between five South African beef cattle populations including Afrikaner, Nguni, Drakensberger Bonsmara and Angus in order to establish if the prediction equation for one breed could be used to predict GEBV for another breed.

Material and Methods

Animal resources: Five South African beef cattle populations including Afrikaner (44), Nguni (56), Drakensberger (47), Bonsmara (46) and Angus (31) were used in this study. To maximize the genetic diversity within each sampled population, pedigree data were used to select against full and half sib animals.

Genotyping and quality control: Genomic DNA was extracted at the ARC-Biotechnology Platform using the Qiagen DNeasy extraction kit according to the manufacturer's protocol. Genotyping was conducted at the ARC-Biotechnology Platform with the Illumina BovineSNP50 BeadChip v2. Samples were processed according to the Illumina Infinium-II assay protocol (Illumina, Inc. San Diego, CA 92122 USA). Quality control criteria were used to remove from further analysis any SNPs with less than 95% call rate, SNPs with less than 0.05 MAF, SNPs which deviated significantly from HWE ($P < 0.001$) and to exclude loci assigned to unmapped contigs as well as SNP on the sex chromosomes (Purcell et al., 2007).

Determining consistency of phase across the breeds: Linkage disequilibrium (LD) phase, r , for marker pairs that were common amongst five breeds was estimated using plink software version 1.9 (Purcell et al., 2007), using the following parameters: `--r, --ld-window-kb 1000 and --ld-window 9999`. Data was sorted into groups based on their inter-marker distance to determine the break down in

Table 1 Pearson correlation of signed r value between ten breed pairs at given genomic distance

Distance, Mb	AFR-NGU	AFR-DRA	AFR-BON	AFR-ANG	NGU-DRA	NGU-BON	NGU-ANG	DRA-BON	DRA-ANG	BON-ANG
0.001-0.01	0.37	0.29	0.41	0.20	0.46	0.60	0.22	0.46	0.32	0.23
0.01-0.02	0.27	0.23	0.29	0.04	0.39	0.31	0.18	0.36	0.26	0.17
0.02-0.03	0.34	0.29	0.31	0.17	0.36	0.36	0.21	0.37	0.25	0.25
0.03-0.04	0.33	0.23	0.30	0.11	0.33	0.34	0.16	0.31	0.15	0.19
0.04-0.05	0.29	0.21	0.28	0.14	0.31	0.32	0.14	0.33	0.20	0.19
0.05-0.06	0.28	0.20	0.25	0.10	0.31	0.33	0.12	0.27	0.16	0.16
0.06-0.07	0.26	0.18	0.24	0.10	0.24	0.29	0.12	0.24	0.16	0.15
0.07-0.08	0.25	0.17	0.20	0.08	0.26	0.26	0.09	0.24	0.16	0.14
0.08-0.09	0.27	0.15	0.19	0.09	0.22	0.27	0.08	0.23	0.14	0.10
0.09-0.1	0.22	0.15	0.23	0.05	0.24	0.27	0.05	0.24	0.12	0.11
0.1-0.2	0.20	0.11	0.15	0.03	0.19	0.22	0.05	0.17	0.07	0.07
0.2-0.3	0.13	0.07	0.10	0.01	0.12	0.14	0.02	0.13	0.03	0.04
0.3-0.4	0.11	0.06	0.10	0.00	0.08	0.12	0.01	0.09	0.02	0.02
0.4-0.5	0.10	0.04	0.08	0.01	0.07	0.09	0.01	0.07	0.03	0.03

Afrikaner (AFR), Drakensburg (DRA), Nguni (NGU), Bonsmara (BON), Angus (ANG)

correlation across distance. The r values were correlated between breeds using the PROC CORR procedure in SAS (SAS Institute Inc., Cary, USA). Correlations were performed between all pairs of breeds for same pairs of SNP markers amongst the five breeds. This resulted in ten independent breed pairs.

Results

Consistency of phase was investigated between same pairs of SNP, for pairs of breeds. Results are presented in Table 1 as well as Figure 1. The correlations were moderate over short distances and decreased as the distance increases. The correlation of r decreased more rapidly between all Angus-African breed pairs compared to African-African breed pairs. The greatest correlation amongst African breeds was observed between the Nguni and Bonsmara pair followed by Nguni and Drakensberger and Drakensburg and Bonsmara, whilst the lowest correlation was observed between Afrikaner and Drakensburger pair.

Correlations of linkage phase over short distance are useful to determine ancestral genetic relationship among breeds, and to estimate the time since breeds' divergence as LD over short distance reflects historical LD (Hayes et al., 2003). To illustrate historical genetic relationship between studied breed pairs, the correlation of r for breed pairs was plotted over short genomic distances (reflecting past generations ago) (Figure 1). A higher correlation of r suggests a close genetic relationship between breed pairs thousand generations ago (Lu et al., 2012). In this study, the Nguni-Bonsmara pair had the closest historical genetic relationship followed by Nguni-Drakensburger and Drakensberger-Bonsmara pairs compared to other breeds pairs.

Discussion

Linkage phase represents the genetic relationship between the populations, based on how much they have diverged over time (de Roos et al., 2008). Maudet et al. (2002) showed that breeds that are known to be more related based on descent are more likely to be more in consistence of phase. In the current study breed pairs of Nguni and Bonsmara (0.60), followed by Nguni and Drakensberger (0.46) and Drakensberger and Bonsmara (0.46) were mostly correlated compared to other breed pairs for marker pairs that were less than 10 kb apart. This means that if two markers at this distance were in LD in one population they will show similar level of LD in the other populations and that if a QTL is found in a chromosome region in Nguni, SNPs linked to the QTL have 60% chance to carry the same effect in the Bonsmara population given that region has an LD of at least 0.2. The correlation of r among Angus and African breed pairs were lower than the African-African breed pairs and mostly decreased over increasing genomic distance. This was agreement with great divergence between African and British breeds (Gautier et al., 2007). The correlation of LD phase between Afrikaner-Nguni and Nguni-Bonsmara pairs were higher at larger distance (0.4-0.5 Mb) than other breeds pairs. This may suggest that these breeds' pairs share longer segment of haplotypes compared to other breed pairs. It also implies that there has been a more admixture among these two breed pairs in recent generations compared to longer generations. This agrees with the history of Bonsmara population as this breed was developed from the Afrikaner cattle in 1960s through crossbred trials with Herford and Milking Shorthorn at Mara Research Station (Bonsma, 1980).

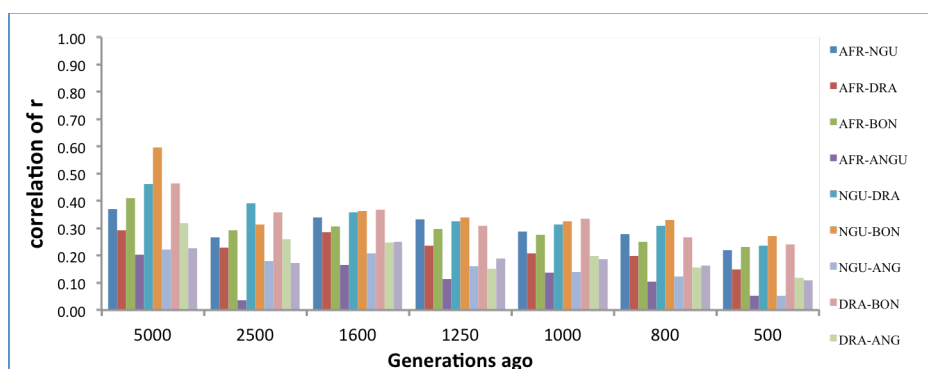


Figure 1. Pearson correlation of signed r between Afrikaner (AFR), Drakensberger (DRA), Nguni (NGU), Bonsmara (BON) and Angus (ANG) populations at given generations calculated using linkage disequilibrium

Persistence of LD phase between populations is also most useful to determine the marker density to conduct multi-breed genomic selection. In the current study LD markers that were found amongst African breeds will have a moderate effect across these populations because correlation of r remains just between 0.4 and 0.6 for marker pairs that were within 10 kb apart. In general this study revealed that there is still great degree of diversity amongst South African cattle populations. Therefore in order to find markers that works consistently, i.e. r correlation > 0.8 (de Roose et al., 2008), across South African beef cattle, the marker to QTL interval should be less than 5 kb, which correspond to approximate 300 000 or more markers distributed across the genome. The results of this study compares favorably with previous studies of Goddard et al. (2006) and de Roos et al. (2008) who indicated the value of having a large number of SNPs to cover the genome for genomic selection when looking across populations.

Conclusions

The persistence of linkage disequilibrium was moderate between any pair of African breeds and lower for any pair of Angus-African breeds for SNP pairs separated by less than 10 kb in the current study. This suggested the necessity of breed-specific reference populations or the need to include adequate representation of each breed in the reference population if a multi-breed reference is to be used.

References

- Bonsma J.C. (1980). J. C. Bonsma (Ed.), Livestock production. A global approach, pp. 126–36
- Dekkers, J. C. (2010). *Vet. Sci. Nutr. and Nat. Resour.* 5:1-13
- de Roos A. P., Hayes B. J., Spelman R. J. et al. (2008). *Genet.* 179:1503–1512
- Gautier M., Faraut T., Moazami-Goudarzi K. et al. (2007). *Genet. Soc. of Am* 177:1059–70
- Goddard M. E., Hayes B. J., McPartlan H. et al. (2006). Proc 8th WCGALP, CD-ROM Com. No 22-16
- Goddard, M. E., & Hayes. B. J. (2007). *J. Anim. Breed. Genet.*, 124:323–330
- Hayes, B. J., Bowman, P. J., Chamberlain, A. J. et al. (2009). *J. Dairy Sci.*, 92:433-443
- Lu D., Sargolzaei M., Kelly M. et al. (2012). *FGene.* 3:152-60
- Matukumalli L.K., Lawley C.T., Schnabel R.D. et al. (2009) *PLoS ONE*, 4: e5350-63
- Maudet, C., Luikart, G. & Taberlet, P. 2002. *J Anim Sci* 80 (4): 942-50
- Mckay S.D., Schnabel D., Murdoch M. et al. (2007). *BMC Genet.* 8:74, doi: 10.1186/1471/2156-8-74
- Meuwissen T. H., Hayes B. J., Goddard M. E. (2001). *Genet.* 157:1819–1829
- Purcell S., Neale B., Todd-Brown K. et al. (2007). *Am. J of Hum. Genet.*, 81:559–75
- Qwabe, S.O., vanMarle-Köster, E., Maiwashe, A. & Muchadeyi, F.C. (2013). *S. Afr. J. Anim. Sci.*, 1(43):64-7
- Snelling W. M., Allan M. F., Keele J. et al. (2010). *J. Anim. Sci.* 88:837–848
- Van Eenennaam, A. L. (2006). <http://animalscience.ucdavis.edu/animalbiotech>
- Van Marle-Köster, E., Visser, C. & Berry, D.P. (2013). *S. Afr. J. Anim. Sci.*, 43:1-17