Genomic tools applied to dairy Buffaloes

Humberto Tonhati, Diercles Francisco Cardoso, Daniel Jordan de Abreu Santos

Depto Zootecnia, FCAV/Unesp, Jaboticabal

Abstract

Buffaloes (*Bubalus bubalis*) are domestic animals of great socioeconomic importance specially to developing countries where the herds are located. In Brazil, the dairy buffalo herds are especially important due to the great market acceptance and appreciation of the mozzarella cheese. New genomic technology applied to animal breeding has been the object of many studies with emphasis on the apparent advantages achieved by adopting this tool. This paper presents the results of studies that applied these genomic tools to dairy buffaloes. The results suggest that applying genomic selection to dairy buffaloes is advantageous; however, the available tools require improvements for the genomic evaluation of this species.

Introduction

The Brazilian buffalo herd is estimated at 1.2 million animals, according to FAO (2013), ranking it as the largest buffalo herd in the American continent and the eleventh in the world. Despite this substantial population, the records on the introduction of buffalo in Brazil describe the entry of about 300 animals, through independent imports of small batches coming mainly from Italy and the Asian continent, between the late XIX and mid-XX centuries (Bernardes, 2007; Santiago, 2013). One of the motivations for these imports was the interest of Italian immigrant families in the manufacture of dairy products appreciated in their country of origin, such as the "mozzarella" cheese.

The increasing demand and the number of industrial units specialized in the production of buffalo milk products led to the expansion of the dairy farming activity in the Southeast, Northeast and Midwest regions, which concentrate 37% of the current national herd. Despite the low number of animals from which the Brazilian buffalo population emerged, the large variability of the productive traits of the dairy herd observed in some studies (Tonhati et al. 2000; Pinto et al. 2007) indicates sufficient genetic variability for the implementation of breeding programs. The Brazilian Association of Buffalo Breeders (Associação Brasileira dos Criadores de Búfalos, ABCB) carried out a genetic assessment in herds of 13 dairy farms in 2007 and reported significant differences between individual breeding values (BV) for milk yield, and a coefficient of variation of 36.20% for the total milk yield at 305 days in 5434 lactations.

Genomics, the new tool in animal breeding, consists of applying statistical methodologies using the DNA molecule variations to predict the genomic breeding values (GBV) through genomic selection (GS) and genome-wide association studies (GWAS), identifying which regions of the genome influence certain traits. The GS has the potential to reduce the generation interval by anticipating the selection process for late measurement traits, and predict genetic values with greater accuracy for traits of low heritability and/or limited to sex. Thus, this technology has been successfully applied in the genetic improvement of dairy cattle.

Regarding buffaloes, the application of genomic tools is still somehow limited, mainly due to lack of a own reference sequence. The strategy used in buffalo genomic studies is based on using the SNP chips developed for cattle or buffalo-specific chips but with SNPs mapped in the bovine reference genome. This paper aims at presenting the results of genomic studies obtained using the chip BovineHD BeadChip 777K and Axiom Genotyping in Brazilian buffalo herds.

Material and Methods

The results were obtained from the genotyping of 384 animals using the BovineHD BeadChip 777K (Illumina) and 452 animals using the Axiom Buffalo Genotyping (Affymetrix). Both genotyping procedures were financed by the research project: "Application of genomic information in genetic improvement of dairy buffaloes, "FAPESP 2010 / 20887-1. The genotyped populations belong to two dairy farms in Rio Grande do Norte and São Paulo.

Results

Using the BovineHD BeadChip

The BovineHD BeadChip developed for cattle, contains 777,962 SNP markers. Using this chip for genotyping of buffalo populations idealizes the transferability of SNPs between species with similar genomic sequences. However, in the studied population, only 16,580 of the 688,593 successfully genotyped buffalo markers were polymorphic (Aspilcueta et al. 2014). This result indicates a high conservation level of the loci between the two species, but not for polymorphisms. It is noteworthy that the percentage of SNPs of the BovineHD BeadChip information in buffalo population (2%) agrees with those observed by Michelizzi et al. (2010) and Wu et al. (2013) in relation to low dense bovine chips (50K).

The average linkage disequilibrium (LD) between pairs of markers, estimated by r^2 was 0.28, suggesting that despite the low transferability of SNPs this chip provides sufficient dense markers to apply to buffalo selection genomics since the LD values greater than 0.2 allow estimating GBVs with an accuracy of 0.8 (Calus et al., 2009). Nevertheless, the genotyping cost with the HD panel to use only 2% of genomic information is likely to hamper the use of this panel in buffalo populations.

Using the Axiom Buffalo Genotyping

The Axiom Buffalo *Genotyping* 90K, developed especially for the buffalo species, contains 92,826 SNPs aligned according to the bovine reference sequence. In the studied population, 46,378 markers (50%) were polymorphic. The average LD between markers, estimated by r^2, was 0.29.

The effective population size estimated by the LD was 86 for 10 generations, close to the minimum levels recommended by FAO (1998) to maintain genetic diversity.

Viability of genomic selection for productive traits

One way of using the genomic information to estimate genomic values is the single step, which consists of combining the traditional relationship matrix (A) with the genomic relationship matrix (G) to obtain the so-called matrix H (Legarra et al., 2009). The matrix H, combining non-genotyped animals and those genotyped with the 90K chip, was used in a single-step model for genomic evaluations of the following traits, milk yield (MY), total fat (TF) and protein (TP) at 270 days; fat percentage (FP) and protein percentage (PP); mozarella production (MP), and somatic cell score (SCS).

The following table shows the heritability and average breeding values estimated by traditional evaluation and genomics. It is noteworthy that prediction accuracy for some traits increased when using the genomic information.

Table – Heritability and average of estimated accuracies for PTAs of all animals and those genotyped using traditional evaluation and genomics.

Trait	Heritability		Total		Genotyped	
	Traditional	Genomic	Traditional	Genomic	Traditional	Genomic
MY	0.21	0.22	0.518±0.15	**0.522±0.15**	0.693±0.07	**0.703±0.07**
SCS	0.13	0.13	0.349±0.16	**0.350±0.16**	0.610±0.08	**0.611±0.08**
TP	0.21	0.21	0.513±0.15	**0.516±0.15**	0.696±0.08	**0.702±0.07**
TF	0.19	0.19	0.487±0.15	**0.490±0.15**	0.679±0.07	**0.682±0.07**
MP	0.21	0.21	0.509±0.15	**0.512±0.15**	0.692±0.07	**0.697±0.07**
FP	0.18	0.18	0.381±0.16	0.381±0.16	0.653±0.08	0.653±0.08
PP	0.23	0.26	**0.403±0.17**	0.400±0.17	**0.696±0.08**	0.690±0.08

The use of the matrix G for estimating genomic PTAs regardless of pedigree information allowed estimating the PTAs for milk yield with an accuracy of 0.40. In this case, accuracy was measured by the correlation between the genomic PTAs predicted for the genotyped subgroups (validation groups) and phenotypic values observed in the dataset. These results suggest that the use of genomic information may facilitate the implementation of selection programs even when there is no pedigree information available.

Genome-Wide Association Study (GWAS)

The deregresed PTAs of the six productive traits (MY, TF, TP, FP, PP, SCS) and four reproductive traits, age at first calving (AFC), calving interval (CI), number of artificial inseminations per conception (NSC) and days open (OD) were used for association studies in the genome, considering the population genotyped with the 90k chip (Camargo et al., 2015). The identification of the genes closest to the SNPs associated with each trait was performed using the "Map Viewer" of the bovine genome available on the NCBI, considering the 1Mb distance to the right and left of the SNPs significantly associated with each trait.

SNPs that explained higher proportions of phenotypic variance of each trait indicated the regions that influence the studied traits while some of them mapped genes related to the physiology of each trait (e.g. ZNF385D and TAC1 genes are knowingly related to lipid metabolism, LOC100336939 is related to the implementation of bovine embryos, etc.). The identified genes are candidates for fine mapping to find causal mutations and incorporate this information to a low-density chip. The discovery of new SNPs, particularly those in and near the buffalo gene sequences, raises the possibility that markers in linkage disequilibrium with the QTL may be inserted in SNPs chips.

Final considerations

The application of genomic tools can assist the implementation of selection programs for dairy buffaloes, promoting benefits over traditional selection. However, buffaloes lack a complete genome available for public consultation. Therefore, the advance of new approaches and development of technologies for the genetic evaluation of the species become more difficult. Additionally, some QTL previously identified in buffaloes as the region of the *DGAT1* gene that is highly related to fat production

(Cardoso et al., 2015) was not detected in this study, suggesting that there are no markers in the chip in LD with this QTL.

REFERENCES

ASPILCUETA-BORQUIS RR (2014) Water buffalo genome characterization by the Illumina BovineHD BeadChip. Genetics and Molecular Research, 13, 4202-4215.

BERNARDES O (2007) Bubalinocultura no Brasil: situação e importância econômica **Revista Brasileira de Reprodução Animal**. v.31, p.293-298.

CALUS MP et al. (2009) Estimating genomic breeding values from the QTL-MAS Workshop data using a single SNP and haplotype/IBD approach. BMC Proc. 3.

CAMARGO GMF et al. (2015) Prospecting major genes in dairy buffaloes. BMC genomics, v. 16, p. 872.

CARDOSO DF et al (2015) Variable number of tandem repeats polymorphism in DGAT1 gene of buffaloes (*Bubalus bubalis*) is associated with milk constituents. Journal of Dairy Science, v.98, p.3492-4395.

FAO. **Food Agriculture Organization**. The Global Livestock Sector – a Growth Engine, 2012. Disponível em: http://faostat3.fao.org/faostat-gateway/go/to/download/Q/QA/S

MALHADO CH et al. (2007) Parâmetros e tendências da produção de leite em bubalinos da raça Murrah no Brasil. Revista Brasileira de Zootecnia, v. 36, n. 2, p. 376–379.

MICHELIZZI VN et al. (2010) A global view of 54,001 single nucleotide polymorphisms (SNPs) on the Illumina BovineSNP50 BeadChip and their transferability to water buffalo. International journal of biological sciences, v. 7, n. 1, p. 18–27.

TONHATI H et al. (2000) Parâmetros Genéticos para a Produção de Leite, Gordura e Proteína em Bubalinos. Revista Brasileira de zootecnia, v. 29, n. 6, p. 2051–2056.

SANTIAGO, A. A. **Introdução dos búfalos no Brasil.** Disponível em: http://www.bufalo.com.br/info_criador/historico_bufalos.pdf.

Wu JJ et al (2013) Investigation of transferability of BovineSNP50 BeadChip from cattle to water buffalo for genome wide association study. Molecular Biology Reports, 40,743–750.