



Genomic selection and assisted reproduction technologies to foster cattle breeding

J.F. Garcia^{1,2,5}, R.V. Alonso^{3,4}, Y.T. Utsunomiya², A.S. Carmo^{2,4}

¹Faculdade de Medicina Veterinária de Araçatuba, UNESP – Univ Estadual Paulista, Araçatuba, SP, Brazil.

²Faculdade de Ciências Agrárias e Veterinárias, UNESP – Univ Estadual Paulista, Jaboticabal, SP, Brazil.

³Faculdade de Medicina Veterinária e Zootecnia, USP, São Paulo, SP, Brazil.

⁴Deoxi Biotecnologia Ltda, Araçatuba, SP, Brazil.

Abstract

Genomics has been propagated as a “paradigm shifting” innovation in livestock during the last decade. The possibility of predicting breeding values using genomic information has revolutionized the dairy cattle industry and is now being implemented in beef cattle. In this paper we discuss how genomics is changing cattle breeding through genomic selection, and how this change is creating new ways to articulate assisted reproduction technologies with animal breeding. We also debate that the scientific community is still starting the long journey to reveal the functional aspects of the cattle genome, and that knowledge in this field is the frontier to a whole new venue for the development of novel applications in the livestock sector.

Keywords: breeding, cattle, genomic selection, reproductive biotechnologies.

How genomics is changing cattle breeding

Since the first initiatives to sequence the human genome made their data publically available back in 2001 (International Human Genome Sequencing Consortium *et al.*, 2001; Venter *et al.*, 2001), scientists have been looking for sustainable ways to apply genomic information in several areas of human activity. The community quickly realized that the early developments of this pursuit had great potential to be applied to the livestock sector due to its economic importance, which led to the prioritization of livestock as a major target for developing real world applications of genomics. In fact, this trend started to become reality in 2002, when the Bovine Genome Project was conceived and implemented. Seven years later, as one of the major achievements in livestock genomics, the conclusion of the *Bos primigenius taurus* genome in 2009 (Bovine Genome Sequencing and Analysis Consortium *et al.*, 2009) opened the possibility for the development of a range of analytical tools capable to help on the exploration of the genomic features of the *Bos* genus (Van Tassell *et al.*, 2008; Matukumalli *et al.*, 2009; Boichard *et al.*, 2012).

Among these genomic analytical tools, one of the most popular is commonly called “genomic chip”.

With these devices, it is possible to analyze up to few millions of genetic markers, known as single nucleotide polymorphism (SNP), from a single DNA sample. Thanks to the existence of the cow reference genome, any genetic marker can be easily linked to a chromosomal location, enabling to perform a multitude of association statistical tests to help in the integration of chromosomal coordinates, SNP markers and phenotypes, and enabling the development of applications in the livestock breeding sector.

In a very fundamental paper, written almost a decade before the completion of the cow genome and the launch of the first cattle “genomic chip”, Meuwissen *et al.* (2001) proposed that selection on estimated breeding values (EBVs) predicted from genetic markers could substantially increase the rate of genetic gain in animals and plants, especially if combined with reproductive techniques to shorten the generation interval. Although their analyses were performed using simulated DNA marker data, since no real data was available at that time, the possibility of predicting the genetic merit of an individual using genomic information associated with pedigree and phenotypic databases influenced the entire livestock genetics community and objective steps were taken in order to transform their idea in reality.

Later, with the concrete perspective of the cow genome conclusion and the imminent development of the first “genomic chip” specific for cattle, Schaeffer (2006) proposed that animals could be genotyped for thousands of SNPs located at roughly one centimorgan (1 cM) intervals throughout the genome, so that a “genomic” estimated breeding value (gEBV) could be determined if the effects of each analyzed SNP could be estimated. He also argued that, since the eventual gEBV could be calculated at birth, it should be possible to compare this new method with the traditional progeny testing strategy used in dairy cattle. His expectations at that time were that costs associated with proving bulls could be reduced by 92% and genetic change increased by a factor of 2. He proposed that genome-wide selection could become a popular tool for genetic improvement in livestock.

The original predictions of Meuwissen *et al.* (2001) and Schaeffer (2006) were shown to be reasonably accurate just after the first cattle “genomic

⁵Corresponding author: jfgarcia@fmva.unesp.br

Phone: +55(18)3636-1365

Received: May 30, 2013

Accepted: June 26, 2013



chip” came to the market. The first commercially available bovine “genomic chip” was the so-called “BovineSNP50” (developed by the company Illumina Inc. in collaboration with the United States Department of Agriculture - USDA), being considered as a robust platform for mapping disease genes and quantitative trait loci (QTL) in cattle (Matukumalli *et al.*, 2009). In fact, this genomic tool was the basis for the development of genomic selection in dairy cattle, one of the most elegant and successful applications of genomics in livestock up to date.

In order to improve genomic analysis in cattle, new “genomic chips” with different marker densities and configurations were developed recently, enabling better research achievements and making it real the routine use of genomics for breeding purposes in cattle (Boichard *et al.*, 2012). Nowadays, there are “genomic chips” available for the major livestock species in different formats and marker densities, so it is expected that there will be an explosion of discoveries and applications in the livestock sector over the next years due to these tools.

Genomic selection as a tool to refine breeding strategies

It is safe to affirm that the only economically sustainable large application of genomics in livestock so far is the dairy cattle genomic selection initiative, more particularly in Holstein breeding programs. Trying to reduce the high costs associated with progeny tests, and in order to increase the velocity of genetic improvement in the Holstein breed from USA and Canada, artificial insemination and pharmaceutical companies joined forces in a consortium with research institutions from these countries in order to put in practice the ideas of Meuwissen *et al.* (2001). As the result of this initiative, initiated in 2004, the first bovine “genomic chip” was launched in 2007 and the first experimental genomic evaluation was performed in USA/Canada in 2008. Since 2009, genomic predictions are part of the Holstein Association routine services to dairy breeders (Scheffer and Weigel, 2012).

Genomic selection conceptually consists in the use of pre-existent pedigree and phenotypic records, associated with genomic information to generate gEBVs. In contrast with the traditional EBVs, the new methodology provides indexes with higher reliabilities. Table 1 illustrates the extent of the improvement in reliability rates by comparing traditional and genomic EBVs in dairy cattle phenotypes (in this case LPI, an index related to the profit offered by the cow during its entire life). A two-fold increase in the average reliability has doubled the genetic change rate in Holstein breeding programs as predicted by Schaeffer (2006). Additionally, the genomic identification of the best calves in early stages of life, who will be kept as young bull candidates for the progeny tests, has reduced the

maintenance costs and the generation interval as suggested by Meuwissen *et al.* (2001), representing the best justification for the massive application of genomic selection we witnessed in the recent years. Other traditional dairy breeds from North America, such as Jersey, Ayrshire and Brown Swiss are taking advantage of this technology by applying genomic selection in their breeding processes since 2009.

Diversely, the beef cattle sector has been struggling with the attempt to apply genomics in its genetic evaluation routine for years, with relatively less success. The first experiences were made using SNP panels containing low number of genetic markers (between 96 and 384 SNPs) and commercialized mainly by companies traditionally dedicated to livestock health and reproduction areas. Those pioneer initiatives were originated from the idea that the use of few SNP markers, associated with genes related to phenotypes of interest, would allow breeders to test a given animal and make strategic decisions.

In North America, this type of genomic test is still being applied to some beef breeds, in special to Angus (Gill *et al.*, 2009; Garrick, 2011), using a “blending” approach as the way to improve EBV’s reliability. Other breeds, such as the dual-purpose Simmental and Braunvieh Schweiz (Brown Swiss in Europe), are implementing the genomic selection protocols similarly as in dairy breeds, although still in a very preliminary way (Gredler *et al.*, 2009; Croiseau *et al.*, 2012)

In Brazil, since 2005 and specifically for Nelore (the major beef breed in the country), there were several frustrated attempts to implement genomics in breeding and selection. Several factors have concurred to the failure of implementing this type of initiative, but three of them were of capital importance and should be closely analyzed. The first one was the fact that in the beginning, the first SNP tests were not integrated to the traditional breeding programs, either because their costs were prohibitive or the breeders could not assess the benefits of investing in that new technology. Based on the market strategies and technical data presented at that time, the final user (the breeder) was not able to know how the investment would be recovered since the promised increase in reliability did not provide enough parameters for proper economic calculations. Another important factor was that the business model of those companies did not comprehend the sharing of real data information (SNP variants and its effects) to the breeders, offering results in a “score-type system” which breeders should use as the selection criteria. Finally, that approach was overcome by the genomic selection concept, widely applied in dairy breeds from 2008, and where instead of couple hundreds, several thousands of SNP effects are calculated and used to infer the genetic value of an individual.

Nowadays, several research groups are depositing their efforts on the development of new low



density “genomic chips” (containing between three and twenty thousand SNPs) to be routinely used in beef cattle genomic selection programs. The idea behind this strategy is to create a DNA test that could reach reasonable price and high precision simultaneously, serving to the purposes and peculiarities of beef breeding systems. The use of inference methods to generate high density genotypes from low density ones, known as imputation, is the key for its success, allowing for

reduction in genotyping associated costs while maintaining the quality of genomic predictions. However, there are several unsolved expectations still to be clarified about how the beef sector will fully benefit from genomics, as the dairy sector did recently. The costs associated with DNA testing is one of the most striking factors, since in beef cattle (diversely from dairy), the unitary value of an individual makes prohibitive to test a large number of animals in a herd.

Table 1. Average gain in LPI (Lifetime Profit Index) reliability due to genomics in Canadian Holstein - August 2013.

Sub-group for Holstein breed	Average LPI Reliability (%)			
	Traditional	Genomics ^a	Gain	DGV Weight ^b
50k ^c Young bulls and heifers (born 2010-2012)	38	70	32	65%
LD ^d (3k or 6k) heifers (born 2011-2013)	34	67	33	66%
LD ^d Younger hows in 1 st or 2 nd lactation	50	69	19	58%
LD ^d Foreign hows with MACE ^e in Canada	42	69	27	63%
1st Crop proven sires in Canada	85	90	5	51%
Foreign sires with MACE ^e in Canada	69	83	14	55%

^aEvaluation based on the Genomic Parent Average (GPA), which combines the Parent Average (PA) with the Direct Genomic Value (DGV). ^bWeight attributed to the DGV in order to combine PA and DGV. ^cAnimals genotyped with Illumina® BovineSNP50 Genotyping BeadChip (a.k.a. 50k), which interrogates 54,609 SNPs. ^dAnimals genotyped either with Illumina® BovineLD Genotyping BeadChip (a.k.a. LD or 6k) or Illumina® GoldenGate® Bovine3k Genotyping BeadChip (a.k.a. 3k), which interrogates 6,909 and 2,900 SNPs, respectively. ^eMultiple-trait Across Country Evaluations. Adapted from Canadian Dairy Network <<http://www.cdn.ca/articles.php>>.

Articulating genomic selection and assisted reproduction technologies

For decades, genetic evaluations (GE) and assisted reproduction technologies (ART) have been tracking different scientific pathways. However, in the context of application, GE and ART have been working together to promote genetic changes in commercial herds. While breeding programs use mass selection approaches for GE to identify phenotypically superior animals, considering several traits simultaneously and weighting them in selection indexes based in EBV with high reliabilities, ART such as artificial insemination, embryo *in vitro* production, and embryo transfer, have been applied to the multiplication of animals considered to be superior. Genomic predictions now offer a new opportunity for using ART to shorten generation intervals and maximize genetic gain, in special the use of preimplantation genetic diagnosis (PGD).

The commercial use of PGD has been limited not only by technical restraints such as training technicians in embryo micromanipulation and embryo biopsy procedures, but also by the existence of a small number of genetic tests with justifiable benefit, low cost, and commercial potential available. Perhaps the largest present application of PGD is embryo sexing, which allows for the identification of male and female embryos prior to transfer to a recipient cow. However, semen sex

sorting is being rapidly adopted as a replacing technology to embryo sexing in the cattle industry.

As the genome of the embryo remains unchanged in the somatic cells of the adult, PGD can be used as a satellite technology to speed up genetic gain using genomic selection. In this context, an embryo biopsy can be performed, its DNA extracted, and a “genomic chip” used to estimate the gEBV of an animal that was never born. One may want to transfer only embryos exhibiting satisfactory gEBVs, and discard the others. Similarly, pregnancy can be marketed with fetuses with known and quantified genetic merit.

Lauri *et al.* (2013) have proved the concept that faithful genotypes can be retrieved from amplified DNA of embryo biopsies by using samples of cloned embryos. However, the quality of the genotypes obtained is sensitive to the number of embryonic cells biopsied, i.e., to the initial amount of DNA available for amplification (Alonso *et al.*, 2013; FMVZ/USP, São Paulo, SP; unpublished data). Satisfactory results can be obtained from a typical embryo biopsy, and the quality of the genotypes can be substantially improved if genotypes of the sire and dam are available to be compared with the genotypes of the embryo. Embryo genotyping has already been used in reproductive programs of dairy breeds in Europe (Le Bourhis *et al.*, 2011) and North America (Sargolzaei *et al.*, 2013), and will soon be available for commercial application with



in vitro produced bovine embryos in Brazil.

Next frontier in cattle genomics

Although the current development of genomic selection is a promising field allowing for broader applications in dairy and beef cattle, the generation of large amount of genomic data (either SNP or DNA sequence information) has the potential to create completely new lines of research, and take us beyond using whole genome information to predict breeding values. It is gradually becoming clear that exploiting the functional meaning of the genes underlying economically important traits can generate invaluable knowledge to develop new technologies.

Using SNP data, it is now possible to scan whole cattle genomes, and look for signals of associations between chromosome coordinates and traits of interest (Bush and Moore, 2012), or even seek genomic regions where past selection has taken place (Oleksyk *et al.*, 2010). By extracting functional information about the genes located nearby these signals, it is possible to give rise to a handful of interesting hypotheses on the biological pathways that underlie economically important traits. This is the revolution of the “hypothesis generating research”, which allows for setting experiments to elucidate physiological and pathological roles in phenotype differences based on findings from large scale genomics studies.

Studies taping the genomes of Brazilian cattle are only now emerging. For instance, a genome-wide scan for birth weight in Nellore cattle was recently performed and led to the discovery of a strong signal pointing out to a region in chromosome 14, which harbors a series of genes previously found to affect human and cattle stature (Utsunomiya *et al.*, 2013a). Another study looking for evidence of breed-specific recent selection detected genetic variations that may have been shaped by human-driven selection in four different cattle breeds (Utsunomiya *et al.*, 2013b). Among these, the dairy Gyr breed presented a signal mapping to the *ST6GALNAC5* gene, which is known to participate in the deposition of gangliosides in milk fat globules. These studies represent important resources for characterizing genome regions that affect traits of economic interest in Brazilian cattle.

Conclusions

Analyzing the large amount of scientific literature related to cattle genomics produced in the past five years, it is possible to observe the extremely fast pace on the adoption of scientific knowledge in the practical daily applications. With this in perspective, it is feasible to postulate that in the near future the ART and GE approaches combined will be the driving forces to lead cattle breeding to a finer process than it is nowadays.

From one side, GE improved methods will make possible to know which gene alleles are the exact ones desired for a given type of animal. On the other hand, ART will allow to check the presence of these favorable alleles in early stage *in vitro* produced embryos, making the whole selection and breeding process extremely more accurate.

We also foresee the development of specific “genomic-audited” lineages, carrying specific and interconnected alleles selected inside the traditional breeds, which would offer better chances to the livestock industry to produce the animal required for each type of application.

This vision can also be beneficial to tropical animal production systems where traits related to environment adaptation (such as heat tolerance, low quality forage grazing ability, disease challenge resistance) play fundamental roles for its development, although still having their physiological basis to be uncovered.

Acknowledgments

Scientific Projects: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (Grants: 560922/2010-8 and 483590/2010-0), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (Grants: 2011/16643-2 and 2010/52030-2).

Technological Innovation Projects: CNPq RHAÉ (Grants: 551809/2011-6 and 350291/2012-8), FAPESP – PIPE (Grants: 2006/56683-5 and 2009/51784-6).

References

- Boichard D, Chung H, Dasonneville R, David X, Eggen A, Fritz S, Gietzen KJ, Hayes BJ, Lawley CT, Sonstegard TS, Van Tassell CP, VanRaden PM, Viaud-Martinez KA, Wiggans GR, Bovine LD Consortium. 2012. Design of a bovine low-density SNP array optimized for imputation. *PLoS One*, 7:e34130.
- Bush WS, Moore JH. 2012. Genome-wide association studies. *PLoS Comput Biol*, 8:e1002822.
- Croiseau P, Guillaume F, Fritz S. 2012. Comparison of genomic selection approaches in Brown Swiss within intergenomics. *Interbull Bull*, 46:28-31.
- Elsik CG, Tellam RL, Worley KC, *et al*, Bovine Genome Sequencing and Analysis Consortium. 2009. The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science*, 324:522-528.
- Garrick DJ. 2011. The nature, scope and impact of genomic prediction in beef cattle in the United States. *Genet Sel Evol*, 43:17.
- Gill JL, Bishop SC, McCorquodale C, Williams JL, Wiener P. 2009. Association of selected SNP with carcass and taste panel assessed meat quality traits in a commercial population of Aberdeen Angus-sired beef cattle. *Genet Sel Evol*, 41:36.
- Gredler B, Nirea KG, Solberg TR, Egger-Danner C,



- Meuwissen THE, Sölkner J.** 2009. Genomic selection in Fleckvieh/Simmental: first results. *Interbull Bull*, 40:209-213.
- Lander ES, Linton LM, Birren B, et al, International Human Genome Sequencing Consortium.** 2001. Initial sequencing and analysis of the human genome. *Nature*, 409:860-921.
- Lauri A, Lazzari G, Galli C, Lagutina I, Genzini E, Braga F, Mariani P, Williams JL.** 2013. Assessment of MDA efficiency for genotyping using cloned embryo biopsies. *Genomics*, 101:24-29.
- Le Bourhis AD, Mullaart E, Schrooten C, Fritz S, Coppieters W, Ponsart C.** 2011. Breeding values concordance between embryos and corresponding calves. *Reprod Fertil Dev*, 24:180. (abstract).
- Matukumalli LK, Lawley CT, Schnabel RD, Taylor JF, Allan MF, Heaton MP, O'Connell J, Moore SS, Smith TP, Sonstegard TS, Van Tassell CP.** 2009. Development and characterization of a high density SNP genotyping assay for cattle. *PLoS One*, 4:e5350.
- Meuwissen TH, Hayes BJ, Goddard ME.** 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157:1819-1829.
- Oleksyk TK, Smith MW, O'Brien SJ.** 2010. Genome-wide scans for footprints of natural selection. *Phil Trans R Soc B*, 365:185-205.
- Sargolzaei M, Vigneault C, Blondin P, Schenkel F, Chesnais J.** 2013. Genotyping of bovine embryo biopsies with a high density panel for use in genome selection. *In: International Plant & Animal Genome Conference*, 21, 2013, San Diego, CA. pp. 541. Available on: <<https://pag.confex.com/pag/xxi/webprogram/Paper7976.html>>. Accessed on: 15 march. 2013.
- Schaeffer LR.** 2006. Strategy for applying genome-wide selection in dairy cattle. *J Anim Breed Genet*, 123:218-223.
- Scheffers JM, Weigel KA.** 2012. Genomic selection in dairy cattle: Integration of DNA testing into breeding programs. *Anim Front*, 2:4-9.
- Utsunomiya YT, Carmo AS, Carneiro R, Neves HHR, Matos MC, Zavarez LB, Pérez O'Brien AM, Sölkner J, McEwan J, Cole JB, Van Tassell CP, Schenkel FS, Silva MVGB, Porto Neto LR, Sonstegard TS, Garcia JF.** 2013a. Genome-wide association study for birth weight in Nelore cattle points to previously described orthologous genes affecting human and bovine height. *BMC Genetics*, 14:52.
- Utsunomiya YT, Pérez O'Brien AM, Sonstegard TS, Van Tassell CP, Carmo AS, Mészáros G, Sölkner J, Garcia JF.** 2013b. Detecting loci under recent positive selection in dairy and beef cattle by combining different genome-wide scan methods. *PLoS One*, 8:e64280.
- Van Tassell CP, Smith TP, Matukumalli LK, Taylor JF, Schnabel RD, Lawley CT, Haudenschild CD, Moore SS, Warren WC, Sonstegard TS.** 2008. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. *Nat Methods*, 5:247-252.
- Venter JC, Adams MD, Myers EW, et al.** 2001. The sequence of the human genome. *Science*, 291:1304-1351.
-