GENOMIC SELECTION - AN IMMINENT PROTOTYPE CHANGE THAT PROMISES PROGRESS OF LIVESTOCK BREEDING SECTOR IN INDIA

C.B. Bimal and G. Radhika

Department of Animal Breeding and Genetics, College of Veterinary and Animal Sciences,
Mannuthy, Thrissur, Kerala - 680 651

Received: 20-08-2017 Accepted: 30-08-2017

INTRODUCTION

Animal breeding, by definition involves the selective breeding of domestic animals, with the aim to improve beneficial traits that are heritable and hence propagated in the future generations. Conventional breeding and selection methodologies utilized information on phenotypes as well as on pedigree to predict the breeding value and helped to improve the production potential of livestock to a great extent. However, this process turned out to be time consuming and less accurate with traits of low heritability and sex limited traits like milk production, thereby decreasing the genetic gain.

With the advances in biotechnology and molecular biology, it has now become possible to identify genomic variations at a single nucleotide level and associate them with traits of economic importance. Many genetic markers were found in livestock that were closely linked to a Quantitative Trait Loci (QTL) with favourable effect on many traits and this in turn led to the implementation of Marker Assisted Selection (MAS) in livestock breeding programmes. But genetic variance was found to be overly estimated using this method, owing to the limited number of QTLs identified. Also, the efficiency of MAS was reduced due to the low association between markers and QTL (Boichard *et al*., 2016). Currently, the focus has moved from candidate gene and QTL level to genome wide scale, thereby shifting the selection methodology from MAS to Genomic Selection. High throughput genomic technologies like Next Generation Sequencing (NGS) and Single Nucleotide Polymorphism (SNP) genotyping together with the advancements in bioinformatics as well as statistics have made it possible to predict the outcome of selection more accurately. Genomic selection has currently become the most important tool for animal breeders. Genome wide SNP arrays are presently available for many of the livestock species that enable animal breeders to come up with a genomic estimated breeding value (gEBV). This gEBV helps the breeders to rank sires and dams as well as young animals based on genetic merit.

Genomic Selection

Genomic selection approach was first explained by Meuwissen *et al.* in 2001, where they suggested that the breeding value could be estimated from genetic markers which span the entire genome. It is defined as a form of MAS in which genomic markers covering the whole genome are utilized in such a way that all QTLs come in linkage disequilibrium with at least one marker (Goddard and Hayes, 2007). This technique demanded genotyping a large number of individuals as well as using higher number of markers covering the genome to make it effective. Eventually, this became possible with the development of large scale and a relatively cheaper genotyping method using DNA chips, by the Illumina® company in collaboration with an International Consortium. Initially, a chip that enabled genotyping over 54000 SNPs (Illumina BovineSNP50 chip) was developed (Matukumalli *et al.*, 2009). This assay became the international standard for genome wide studies and genomic selection in cattle. It became possible to scan all the chromosomes for all markers simultaneously in one single step, known as high throughput genotyping. Presently, a high density chip developed by Illumina® in 2010, that can genotype around 777000 SNPs simultaneously, is used widely for SNP genotyping in cattle.

Although these SNPs represented only a small proportion of all the discovered SNPs, they were evenly spaced in the genome and highly polymorphic in nature. The genomic breeding value estimated in progeny tested bulls using these chips, was highly accurate and hence it easily replaced the conventional progeny testing in the Western world. As a consequence, the selection process was simplified and its cost reduced substantially. Due to the marked decrease in generation interval, the annual genetic gain was doubled. Since genomic selection requires huge reference populations, large consortia were formed with international cooperation especially in Europe. Now, it has come to a point where efforts are made to fulfill a more balanced and sustainable objectives by including low heritable traits such as fertility, disease resistance, low methane emission etc in genomic selection (Hayes *et al.*, 2013 and Boichard *et al.*, 2016)

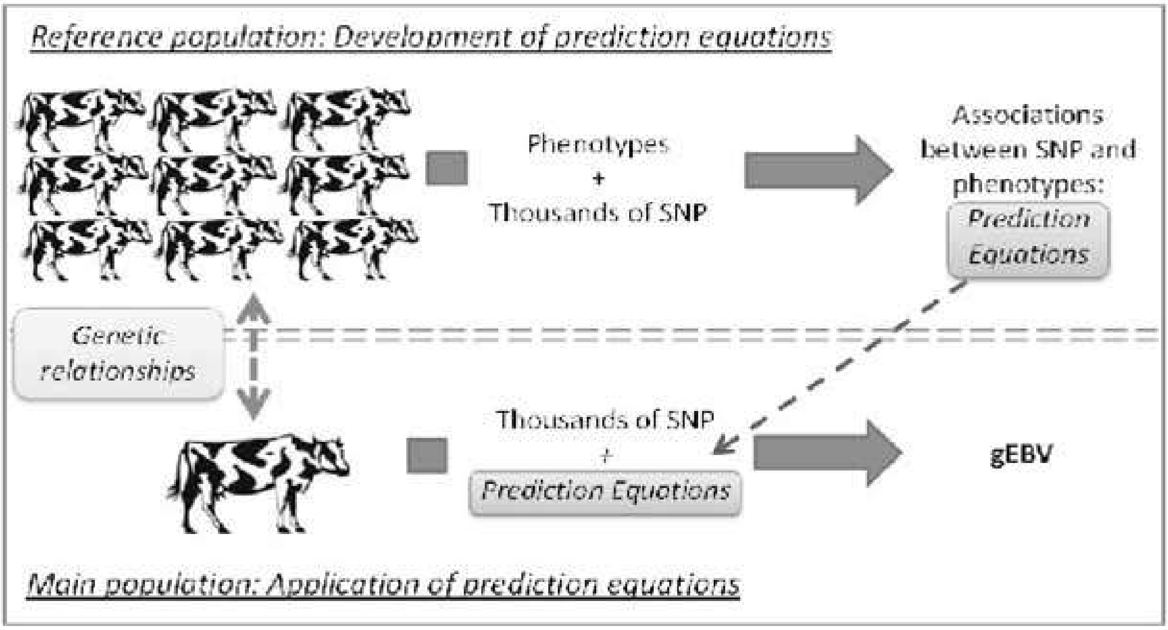
Principles and Methodologies

Genomic selection involves making use of the estimated association between many SNPs and phenotypes in a reference population, to estimate the breeding value of animals whose phenotype is not known but are typed for the SNPs. This is particularly useful in cases where i) phenotypes are very difficult or expensive to measure, especially the low heritable traits, ii) to estimate the breeding value of very young animals before they can produce the phenotypes and iii) in case of sex limited traits as well as those traits expressed later in life or after slaughter. The reference population provides the information to estimate associations between phenotype and SNP genotypes. These associations are then translated into prediction equations that are used to estimate genomic breeding values for animals outside the reference population with the help of SNP genotypes and not with phenotypic expressions.

With genomic selection it is possible to estimate an animal's breeding value quite accurately without its own performance or performance of a large number of offspring. Genomic selection is based on estimation of comprehensive associations between a very dense set of genetic markers (SNP) and phenotypes on a selected group of animals. These associations can then be used to predict the so-called genomic estimated breeding values (gEBV) for related animals that have been genotyped for a large set of SNP, but do not have information on their own performance or a large number of offspring, which are needed for accurate breeding value estimation. Simulation results and limited experimental results suggest that breeding values can be predicted with high accuracy using genetic markers alone but validation is required especially in samples of the population different from that in which the effect of the markers was estimated (Goddard and Hayes, 2007).

The genomic selection is based on the analysis of 10000 to 800000 SNPs. This high number of genetic markers is used as input in a genomic prediction formula that predicts the breeding value of an animal. In genomic selection, the information on DNA of the animal alone will be enough for estimating the breeding value, without collecting its own or its close relatives' phenotypic values. The reference population of animals are scored for key production traits and genotyped using a commercial or customized SNP array. The genotypes are represented by the variable (x) with values 0, 1, 2 (homozygous, heterozygous, or alternate homozygous). A linear prediction equation is generated combining all the marker genotypes with their effects to generate a gEBV for each animal. This prediction equation can be applied to a group of animals that have not been phenotyped to estimate their breeding values and the best animals can be selected for breeding with the help of this gEBV (Fig. 1). Advanced statistical methods like GBLUP, RRBLUP or Bayesian statistical methods are used in estimating the effect of markers on breeding goal traits (Gaur and Sahoo, 2017).

Figure 1. Schematic overview of steps in Genomic selection (Adapted from Textbook of Animal breeding and genetics for BSc students, Kor Oldenbroek and Liesbeth van der Waaij, 2015)



All animals in the reference population are genotyped for a very large number of SNP that are spread across the entire genome. Genotyping for more SNPs though expensive, will also result in a more accurately estimated association. A possible solution to reduce genotyping expense could be the use of haplotype markers instead of SNPs. A haplotype marker makes it possible to type a group of alleles spread across the genome, but are inherited together, all in a single step. Another question under debate is about the optimum number of animals in the reference population. A larger population is going to be more expensive as phenotyping and genotyping of these animals are costly. But a larger population also allows for more accurate estimation of the SNP effects. Like many aspects in animal breeding, the choice of size of the reference population and number of SNP will be a matter of cost-benefit analysis (Hayes *et al.*, 2009).

International experiences in Genomic Selection

Genomic selection in Europe

In Europe, Institut National de la Recherché Agronomique (INRA), France played a major role in implementing genomic selection in cattle. In 2002, Apis-Gene company started genotyping the bulls, later Euro genomics consortium was set up for Holstein breed in 2009 and Intergenomics was set up for Brown Swiss breed in 2014. At present, Illumina Euro G10K chip is being used for genotyping of French cows.

Genomic selection implemented in French dairy cattle breeds (Holstein, Normande and Montbeliarde) denoted the use of haplotypes instead of SNPs to maximize linkage disequilibrium between markers and QTLs (Biochard *et al.*, 2012). They also reported that QTL-BLUP model (best linear unbiased prediction model including QTL random effects) included 300 to 700 trait-dependent chromosomal regions selected by linkage disequilibrium and linkage analysis. Boichard *et al.* (2012) concluded that the QTL-BLUP model was the most efficient one, though it required more effort to phase genotypes, detect QTLs and select SNPs. In Europe, by 2009 more than 65000 animals were genotyped for selection by Labogena (genotyping laboratory) with the 50k chip.

Patry and Ducrocq (2013) reported that INRA, based on their experience in genomic selection in dairy cattle, has formed collaboration with research partners like Brazil and South Africa for development of accurate genomic evaluations in diversified populations and environments. They opined that local genomic selection, which properly accounts genotype by environment interaction, should guarantee increased genetic gains serving local farmer interests.

Genomic Selection in United States (US) Cattle Breeding Scenario

Dairy cattle genomic evaluation reports from US is released since 2008. At present, the number of males and females estimated during 2017 exceed 40,000 and 400,000, respectively. Genotyping laboratories, breed associations and artificial insemination organizations work hand in hand for analyzing DNA and phenotypic data appropriate for evaluating the animal in a genomic way. As far as US genomic evaluation of dairy cattle is concerned, almost all animals evaluated had their sires genotyped and some had their dams evaluated too. More than 60000 SNPs are used for evaluation and tests for genetic abnormalities are also performed. A stage has reached where AI organizations are now relying on genomic evaluation to generate next generation of bulls. Generation interval has dropped which led to an increase in annual genetic improvement. Research is now focusing on including additional traits for selection and also for developing a system to evaluate the already existing, good producing crossbreds. Though more emphasis is given towards the purebred, it would be beneficial if best performing crossbreds are also included in this program.

In US, the first commercial SNP genotyping chip was released with a set of 54001 SNPs in the year 2007. USDA released the genomic evaluation of US dairy cattle from 2009 to 2016, for Holstein, Jerseys, Brown Swiss, Ayrshires and Guernseys. Wiggans *et al.* (2017) opined that the integration of DNA markers and the implementation of genomic technology into the traditional genetic evaluation system has doubled the rate of genetic progress, decreased generation interval, increased selection accuracy, reduced previous costs of progeny testing and has also allowed identification of recessive lethals.

Experiences from Brazil and Kenya on genomic selection

Brazil has 80 per cent of its herd being composed of Holstein x Gyr cross called Giralando. Since DNA information of bulls were not stored, the large reference population needed for Gyr was reported to be formed with addition of cow data to the limited bull data set. The Brazilian Kankrej called the Guzera breed was also studied using 50k illumina panel. GBLUP was considered as the best method for estimating genomic breeding value. Another pertinent point to be noted from the experiences of Brazil is that, genomic selection is practiced among crossbred cattle population also (Vinicius, 2017). The crossbred dairy cattle of Kenya were genotyped using 50k chips (Kim and Rothschild, 2014) and analysis of data using Principal Component Analysis (PCA) revealed admixture of local cattle with several imported breeds. They concluded that the analysis of local ancestry was a useful approach to understand the detailed genomic structure and to reveal regions of genome required for specialized adaptation by looking for changes of haplotype frequency in an admixed population.

Overseas performance information for increasing accuracy in genomic evaluation

Mariam *et al.* (2015) reported that reliability of direct genomic values (DGV) can be increased by increasing the size of reference set used for prediction. For this the bulls should be genotyped with daughter information in countries other than the evaluating country. But this increase in DGV will depend on the genotype by environment interaction between the evaluating country and countries contributing information.

Ducrocq (2017a) reported that genomic selection has radically transformed dairy cattle breeding in developed countries. They emphasized the importance of expanding the reference population by exploiting the data from INTERBULL and Multiple Across Country Evaluations (MACE). Ducrocq (2017b) explained his experiences in India with BAIF, in developing a female reference population of crossbred *Bos taurus* x *Bos indicus* animals.

Indian Initiatives in Genomic Selection

India still follows the conventional progeny testing method in the selection of bulls used for breeding programmes in cattle and buffaloes. But even this has been limited to a few organized herds under the supervision of research institutes of the country like Central Institute for Research on Cattle supported by ICAR having research projects in institutes of different parts of the country including Kerala Veterinary and Animal Sciences University, State Government backed institutes such as Kerala Livestock Development Board, as well as Non Governmental Organizations like BAIF. Owing to the importance of genomic selection, steps have been initiated to implement it at a country wide level. It is proposed to create a "National Bovine Genome Centre" at ICAR-NBAGR, Karnal, including six breeds of cattle namely Sahiwal, Tharparkar, Red Sindhi, Rathi, Gir and Kankrej as well as four breeds of buffalo namely Murrah, Nili-Ravi, Mehsana and Jaffarabadi (Vijh, 2017).

BAIF has initiated a process of genotype assay with the Bovine HD Genotyping Bead Chip, in 1100 animals sampled from ten Indian breeds and also in their crosses to exotic breeds. The objectives include determination of the information content in various breeds and their crosses and to design a medium density SNP array suitable for regular genome analysis and genomic selection in Indian cattle populations (Swaminathan *et al.*, 2017).

SUMMARY

The SNP based marker assay is a pre-requisite for genome wide studies. But the majority of commercial SNP arrays have been developed using data mostly from the exotic cattle (*Bos taurus*) and a limited number from Indian breeds. This might not help in estimating the genotype environment interaction. Hence it is necessary to validate these chips in Indian scenario before genomic selection can be fully adapted. Alternatively, a commercial chip could be custom made incorporating SNPs exclusive to our cattle and must be tested under the management and climatic conditions of our country. Also, lack of an integrated recording system can hamper the success of genomic selection in India. Hence, emphasis must be given to implement a comprehensive data recording system that gives due weightage to productivity, fertility, longevity as well as disease resistance and thermo tolerance. However, a shift from traditional system to a genome based breeding is imminent and promises a new revolution in the livestock sector of our country.

REFERENCES

Boichard, D., Ducrocq, V., Croiseau, P and Fritz, S. 2016. Genomic selection in domestic animals: Principles, applications and perspectives. *C. R. Biologies*. **339**: 274-277.

Boichard, D., Guillaume, F., Baur, A. and Fritz, S. 2012. Genomic selection in French dairy cattle. *Anim. Prod. Sci.* **52**: 115-120.

Ducrocq, V. 2017a. Development of genomic selection in dairy cattle in two emerging countries: South Africa and India. In: *68th Annual Meeting of the European Association for Animal Production (EAAP)*; August 2017, Estonia.

Ducrocq, V. 2017b. European experience in implementing genomic selection in major and minor dairy breeds. In: *Compendium on International Workshop for "Genomic Selection for Genetic Improvement in Dairy Animals*"; 28th to 29th November, 2017, BAIF, Pune.

Gaur, G.K. and Sahoo, N.R. 2017. Phenotpyic data analysis for genomic prediction. In: *Training Manual of ICAR sponsored winter school on "Omictechnologies and modern breeding approaches for conservation and productivity enhancement of indigenous cattle resources*"; 1st to 21st November, 2017. pp. 9-16.

Goddard, M.E. and Hayes, B.J. 2007. Genomic selection. *J. Anim. Breed. Genet.* **124**: 323-330. doi:10.1111/j.1439-0388.2007.00702.x. [04 Aug. 2017].

Hayes, B.J., Bowman, P.J., Chamberlain, A.J. and Goddard, M.E. 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. *J. Dairy Sci.* **92**(2): 433-443.

Hayes, B.J., Lewin, H.A. and Goddard, M.E. 2013. The future of livestock breeding: genomic selection for efficiency, reduced emissions intensity, and adaptation. *Trends Genet.* **29**(4): 206-214.

Kim, E.S. and Rothschild, M.F. 2014. Genomic adaptation of admixed dairy cattle in East Africa. *Front. Genet.* **5**: 443.

Kor Oldenbroek and Liesbeth van der Waaij. 2015. *Textbook Animal Breeding and Genetics for BSc students*. Centre for Genetic Resources, The Netherlands and Animal Breeding and Genomics Centre, Groen Kennisnet. https://wiki.groenkennisnet.nl/display/TAB/. [07 Aug. 2017].

Mariam, H., Pryce, J.E., Schrooten, C. and Hayes, B.J. 2015. Including overseas performance information in genomic evaluations of Australian dairy cattle. *J. Dairy Sci.* **98**: 3443-3459.

Matukumalli, L.K., Lawley, C.T., Schnabel, R.D., Taylor, J.F., Allan, M.F., Heaton, M. and Van Tassell, C.P. 2009. Development and Characterization of a High Density SNP Genotyping Assay for Cattle. *PLoS One*, **4**(4), e5350. [14 Aug. 2017].

Meuwissen, T.H., Hayes, B.J and Goddard, M.E. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics,* **157**: 1819-1829.

Patry, C. and Ducrocq, V. 2013. Genomic selection strategy in France: experiences in dairy cattle - how to make it beneficial for other breeds and countries? http://www.dairyknowledge.in/sites/default/files/patry_and_ducroq.pdf. [25 Jul. 2017].

Swaminathan, M., Pande, A.B., Gibson, J.P. and Ducrocq, V. 2017. BAIF initiatives in application of genomics and genomic selection in Indian dairy animals. In: *Compendium on International Workshop for "Genomic Selection for Genetic Improvement in*

Dairy Animals"; 28th to 29th November, 2017, BAIF, Pune.

Vijh, R.K. 2017. Country level efforts in application of genomic selection in dairy animals including buffaloes. In: *Compendium on International Workshop for "Genomic Selection for Genetic Improvement in Dairy Animals"*; 28th to 29th November, 2017, BAIF, Pune.

Vinicius, M. 2017. Brazilian experiences of genomic selection in *Bos indicus* breeds and its crosses. In: *Compendium on International Workshop for "Genomic Selection for Genetic Improvement in Dairy Animals"*; 28th to 29th November, 2017, BAIF, Pune.

Wiggans, G.R., Cole, J.B., Hubbard, S.M. and Sonstegard, T.S. 2017. Genomic selection in Dairy Cattle: The USDA Experience. *Annu. Rev. Biosci*. **5**: 309-327.

■