Romanian Biotechnological Letters *Vol. 22, No. 5, 2017*

REVIEW

The advent of genomics and its potential contribution to the development of quantitative genetics

Received for publication, December 13, 2015
Accepted, May 31, 2016

COȘIER VIORICA[1] and MONICA MARIAN[2*]
[1)]University of Agricultural Sciences and Veterinary Medicine of Cluj-Napoca, 3-5 Manastur Street, 400373 Cluj-Napoca, Romania, viorica.cosier@gmail.com
[2)]Technical University of Cluj-Napoca, North University of Baia Mare, 76 Victoriei Street, 430122, Baia Mare, Romania
***Address for correspondence to:** monicamarian17@gmail.com*

Abstract

As a result of technological advancement, in the past four decades, profound changes have occurred in the methodological approaches for the selection of animal populations. For this purpose, many revolutionary molecular tools have been developed. With the assistance of recent technologies (NGS-Next Generation Sequencing) and the detection of all SNP (SNP polymorphisms) by microarray technology, we are getting close to the time when all differences in genomic DNA can be found. Genomic selection, based on scanning very dense markers (SNPs) and incorporating them with phenotypic data in genomic breeding value prediction is revolutionising breeding programmes in agriculture. The present paper focuses on discussing, in a broad sense, the manner in which the advent of genomics has contributed to understanding the genetic variations of complex quantitative traits.

Keywords: MAS; GAS; QTL; SNPs; GEBV; BV; LE; LD; genomic selection

1. Introduction

From Gregor Mendel until today, geneticists have been trying to find genetic differences among individuals, the bases of high genetic merit in some animals and how these can be increased and multiplied in animal populations. The first criterion used in selection was phenotype, then, with progress of understanding in molecular field, the protein sequence, electrophoresis techniques, Restriction Fragment Length Polymorphisms (RFLPs) and microsatellite polymorphisms were applied.

The complex quantitative traits are determined by a high number of genes, each gene having a small additive effect. The study of such characters is based on a model in which those genes influence the traits and in which non-genetic factors may also be important. In classical quantitative genetics, the additive effect of all minor genes is looked upon as a black box, and for the association between traits and genetic variation, complex statistical methods were developed and applied in the attempt to decompose the effects of these genes. Continuous progress recorded in the last 30 years, both in molecular genetics and in statistical data interpretation methods, has radically changed this status.

First, the possibility to study the variation occurred in a gene and the role of this genetic variation in the determinism of quantitative traits was clarified, by means of statistical associations. In this view, identification and mapping on chromosome of a sufficient number of polymorphic loci, that capture the highest proportion of the genetic variance was the first challenge. Even though in the ’90 numerous techniques were available for polymorphism

detection, increasing knowledge on Quantitative Trait Loci (QTL) search became available in time, with progress in deciphering genome sequence, and after map saturation in markers, which led to Marker Assisted Selection (MAS). In the last decade, studies of association between genetic variations and economically relevant traits moved to the level of the genome, both in livestock and in crops, due to the development of massive parallel sequencing techniques of the entire genome. With the assistance of recent technologies (NGS-Next Generation Sequencing) and the detection of all SNP differences (SNP polymorphisms) by microarray technology, we are getting close to the time when all differences in genomic DNA can be found (WANG & al., 1998 [1]; HARTWELL & al., 2006 [2], RUSSELL, 2010 [3]). Knowing all the common and less common SNPs in the genome, through genome scanning followed by statistical association procedure (GWAS-Genome Wide Association Studies), the perspective opens up for high accuracy estimate of the genomic breeding value (GBV) (KWOK, 2001 [4]; KWOK, 2003[5]) for **genomic selection (GS)**. Recently, whole genome panels dense in SNP polymorphisms have become commercially available, making it possible to scan the entire genome for the detection of/ polymorphisms associated with economically important traits (microarray technology) in many species. At present, we have a much clearer description of the manner in which genes act to determine phenotypic variations (ZHU & ZHAO, 2007 [6.]) for GS. In animal populations, the first genomic selection revolution began with two developments: sequencing of the bovine genome to understand ruminant biology and evolution (ROSS &WORLEY 2009 [7]); the second development shows that highly accurate selection decisions are possible when breeding value (BV) is predicted from dense marker data alone (MEUWISSEN & al., 2001 [8], SCHEFERS, & WEIGEL. 2012 [9]). As in the last five years, genomic sequence is available for more and more species, and an increasing number of SNPs were identified as a result of genome sequencing and re-sequencing, microarray panels for the interrogation of whole genome SNP polymorphisms are emerging (GODDARD &. HAYES, 2009 [10]). Thus, *Illumina* Inc. has developed genome-wide SNP panels for cattle, dogs, sheep, pigs and horses, as well as for crops (e.g. maize), and other new ones are about to be developed. The present review will focus on discussing, in a broad sense, the manner in which the advent of genomics has contributed to the development of quantitative genetics. Unfortunately, progress in this field is only achieved in developed countries. In order to better understand the progress achieved in quantitative genetics in the last 25 years, which have allowed moving from the classical selection to MAS and then, to genomic selection (GS), the major molecular tools and techniques developed will be briefly outline.

2. Four tools that have revolutionized molecular genetics

Moving from classical selection to MAS is due to progress recorded in both technological advances and improving knowledge in the molecular field. For individual genotyping of economically important traits, in order to apply molecular selection, it was necessary to develop precise and cheap laboratory techniques, and to discover a large enough number of highly polymorphic loci which cover the whole genome. Four molecular tools – based on different principles – contributed to these achievements, which in time allowed the study of genomes and identification of QTL, whose variation was exploited in modern quantitative genetics through the introduction of MAS (Marker Assisted Selection).

2.1 Restriction enzymes

Molecular genetics underwent radical changes in the mid-1970's, when restriction enzymes were first isolated and characterised (SMITH & WILCOX, 1970[11], SMITH &

NATHANS, 1973 [12], ARBER, S. LINN, 1969.[13], GERLT, 1993 [14], ALBERTS & al.,, 2002 [15], LODISH & al., 2003 [16]) and different molecular procedures were developed, allowing researchers to isolate genes and to construct recombinant DNA molecules (COHEN & al., 1973 [17], BROWN & al., 2008 [18], COSIER, 2008 [19]). These breakthroughs facilitated the emergence of modern biotechnologies which put genes to produce recombinant bio molecules, needed in research, medicine (ITAKURA & al., 1977 [20], GOEDDEL & al., 1979 [21], ARENTZEN & al., 1979 [22]), bioremediation (CHAKRABARTY & FRIELLO, 1974 [23]) or in transgenic plant and animal production (BATTEY & al., 1979 [24], GASSER &. FRALEY, 1989, [25], HIATT & al., 1989 [26], CLARK & al., 1989 [27]) and also in agricultural biotechnologies for increasing genetic progress in crops and livestock. Not lastly, molecular techniques which use restriction enzymes were developed and utilised in the construction of the first restriction molecular maps and in genomic QTL identification. Of these, the PCR-RFLP and RFLP techniques based on hybridization were extensively utilised for the above-named purposes, and they were at the base of wide plant and animal population genotyping. Also, restriction enzymes produced raw material for genomic library construction, the first step towards genome's deciphering. In parallel with genome deciphering, detailed physical maps were produced containing the position of genes and other important landmarks in the genome. In time, maps were saturated in markers and constituted the base for developing sophisticated applications for individual genotyping in animal and plant populations, allowing for MAS and, more recently, for GS.

2.2 PCR technique

Ten years later, a new technique developed by Karry Mullis – the PCR (*Polymerase Chain Reaction*) – revolutionised research in the molecular field (CLARK & al., 1988 [28}, MULLIS, 1990 [29], MULLIS, 1990 [30]). The high sensitivity of the technique, by which a DNA fragment can be multiplied in billions of copies, led to numerous variants (PCR-RFLP, iPCR, AFLP, Multiplex PCR, Real Time PCR, etc.) which differ from each other in terms of a priori knowledge about the studied locus/loci. PCR allows the interrogation of polymorphism in the genome in a couple of hours. Applications of these techniques soon appeared in various field of life sciences: identification of DNA variations in alternative allelic forms, study of mutations and of mutational-gene effects, molecular diagnosis, DNA typing for identity testing (JEFFREYS & al., 1985 [31], JEFFREYS & al., 1985 [32], MUELLER & WOLFENBARGER. [33]), comparative analysis of genes with similar functions in different species (EELES & STAMP, 1993 [34]), etc.

The possibility of applying DNA analysis through PCR has led to the development of specific applications in plant and animal genotyping, at loci of economic interest. Consequently, the mutations and variations at the DNA level (genotype), which determine a distinct phenotype, are highlighted by PCR techniques. Once the favourable allele from one or multiple loci is established by association procedure, the molecular information can be integrated in selection. The process targets indirectly the increase in animal production by modifying the genetic structure of the populations.

2.3 Nucleic acid hybridization

Nucleic acid hybridization is a procedure based on the ability of a single-stranded nucleic acid to form double-stranded molecules when the complementarity between nucleotides exists in a relatively large proportion. Depending on the type of nucleic acid, hybridization can take place between DNA-DNA, cDNA-mRNA chains. Depending on the application developed on hybridization assay, it involves a labelled nucleic acid probe to identify related DNA or RNA molecules (ZHU & ZHAO, 2007 [6.], COSIER, 2008 [19], BROWN, 2002 [35]) in a complex mixture of unlabelled molecules. Although much more

expensive and more time-consuming, it is an alternative or a complementary tool for the identification of QTL. As part of the gene and genomic homology analysis, it can be utilised to identify sequences and regions in the undeciphered genomes in different species, even if they are distantly related species; to detect and a priori characterize nucleotide sequences; to screen QTL loci in the genome and genomic libraries; to annotate genomes, etc. (BROWN, 2002 [35]). Microarray provides a platform where, in a single experiment, up to hundreds of thousands of SNP polymorphisms can be identified, or where the expression of hundreds of genes can be measured. Fluorescent labelled nucleic acid probes are directed to a multiple collection of specific short DNA or cDNA sequences immobilized on miniature solid support (the DNA chip) for hybridization (LOCKHART & al., 1996 [36], SCHIMENTI & BUCAN, 1998 [37]. When hybridization is done with cDNA, it is called gene expression profiling, a set of expressed genes (transcriptome) that determine the phenotype of a particular cell (BROWN, 2002 [35]), SCHENA & al., [38]). In selection, the SNP panels which cover the entire genome are targeted to identify such polymorphisms for the high accuracy estimation of GEBV (Genomic Estimated Breeding Value). Although the technology for producing microarray chips was only developed in the last few years, already there have been numerous important applications and their impact on future biomedical research, diagnostic approaches and genomic selection in crops and livestock is expected to be high (RUSSELL, 2010 [3]). Recently in animal field, for the interrogation of genetic variation in cattle genome, the Bovine SNP50v2 and BovineHD Beadchips were developed (by Illumina and USDA ARS, University of Missouri and Alberta University). In one experiment this application makes possible the interrogation of 54,609 or 777,000 informative SNP, or more than 2900 strategically selected SNPs of the cattle genome, very useful in evaluation of genetic merit of individuals and genome-wide selection.

2.4 Sequencing based on capillary electrophoresis and Next Generation Sequencing (NGS)

Genome's sequence is a powerful tool in genome database analyses for comparing sequences from different genomes, identifying gene sequences and the regulatory regions (i.e. promoter, enhancers) (RUSSELL, 2010 [3]), fine mapping and detecting mutations related to diseases or variations associated to QTL. DNA sequencing dates back to 1977, when the first genomes were deciphered (SV40 virus and phage ϕX174, respectively) (MAXAM & GILBERT, 1977 [39], SANGER & al., [40], HUTCHINSON, 2007 [41)] using the Maxam Gilbert chemical method. The chain terminator method, developed almost concomitantly by Sanger & Coulson (1977) (SANGER & al., 1992 [42]), which is based on modified nucleotide incorporation (dideoxy-nuclotide), rapidly attracted researchers' attention, and in a short span of time, it replaced the Maxam Gilbert method. Another improvement was the automation of the procedure to separate reaction products based on capillary electrophoresis (in 1996 by Applied Biosystems, USA), which replaced polyacrylamide gel electrophoresis, due to significant gains in workflow, throughput, and ease of use. For large genome sequencing, different approaches were proposed but the prevailing technique in use was shotgun sequencing (ANDERSON, 1981 [43], PEIFENG & al., [44]). Regardless of the approach, the brunt of the genomes sequencing work for the last 30 years was handled by dideoxy DNA sequencing (Sanger) based on capillary electrophoresis (SMITH & al., 1986 [45]). Recently developed, the massive parallel sequencing techniques, which rely on Next-Generation Sequencing (NGS) systems, began to replace the automated Sanger method. The power of NGS is vastly greater than obtained by capillary electrophoresis (MARDIS, 2008 [46], LISTER & al., 2009 [47]) and lies in the ability to process in parallel hundreds of thousands to millions of DNA templates, resulting in lower running costs per base of

generated sequences and a throughput on the gigabase (Gb) scale. Three of the genome sequencing platforms belonging to the NGS are: Illumina from Solexa Genome Analyser, Applied Biosystem SOLiD ™ System, and the 454 FLX Pyrosequencer (Life Sciences, Roche) (LISTER & al., 2009 [47]). More recently, another two NGS systems were developed: the Helicos Helyscope™ and Pacific Biosciences SMRT (MEUWISSEN & GODDARD, 1996 [54]). The remarkable applications for genome analysis are: genome sequencing and polymorphisms discovery; mutation mapping by deep sequencing, analysis of DNA-protein interactions through ChIP-seq, genome-wide detection of sites of DNA methylation. The sequencing of RNA has also transitioned and now includes full-length cDNA analyses, serial analysis of gene expression (SAGE)-based methods, and non-coding RNA discovery (MARDIS, 2008 [46]). Applications for transcriptome sequencing are: deep sequencing of small RNA populations and mRNA sequencing for transcript discovery and profiling microRNA and siRNA (MARDIS , 2008 [46], LISTER & al., 2009 [47], WANG & al., 1998 [48]). The genomic variations related to phenotype changes, which could be targeted in the selection, belong to a few types of DNA polymorphisms: Restriction Fragment Length Polymorphisms (RFLPs), Single Nucleotide Polymorphism (SNPs), Variable Number in Tandem Repeats (VNTRs) and Short Tandem Repeats (STRs).

3. Classical selection, Marker Assisted Selection and Genomic selection

3.1 Classical selection

Classical selection for economically important traits in animal populations is based on phenotypic records of individuals and their relatives. Even though, to date, important advances have been achieved by classical selection methods for certain economically important traits, several limitations of these methods of improvement are becoming evident in time (CLARK & al., 1989 [27]). In traditional quantitative genetics, genetic progress was recorded without knowledge of genes involved in the expression of quantitative traits. However, the genetic variation of most quantitative traits is determined by polygenes which act additively, and also by major genes. Together, the major and minor genes determine the total genetic variation (COSIER & VLAIC, 2011 [49]). Because these complex determinisms were incompletely understood, the breeding methods utilised were complicated and slow. The problem was mainly due to some simplified models applied for the recording of polygene effect, which proved to be of little efficiency in the case of traits with low heritability and which are difficult to measure. These aspects complicated the breeders' work, who based their selection on estimated breeding values (EBV) calculated from phenotypic records and genealogical information, and on knowledge of the heritability of each trait. BLUP (Best Liniar Unbiases Prediction) (Henderson, 1984) was the prevalently used method for estimating breeding values. This has been successful, but the process was very slow for traits that can only be measured in one sex (for example, milk yield), after death (for example, meat quality), late in life (i.e. longevity), or if measuring the trait was expensive (for example, methane production, feed requirement or disease resistance (GODDARD &. HAYES, 2009 [10]). In the following years, with the development of molecular techniques, selection methods based on molecular information started to be applied confidently, in the belief that some of the current limitations of the methods could be overcome. With the development of genotyping techniques and improved knowledge in the molecular field (i.e. deciphering sequences of more and more genes, polymorphism identification and techniques used to highlight them), inclusion of molecular information in selection started to be studied. One justification for application of molecular genetics on livestock and crops was that the

expectation of information at the DNA level would lead to faster genetic gain than that achieved based on phenotypic data only.

3.2 Marker assisted selection (MAS)

Starting with the 1980's, these methods began to be applied at a smaller scale, using one or a few markers, while in the mid-1990's, due to considerable cuts in genotyping costs, they started to be applied in larger animal populations. In this context, molecular researchers and breeders worked together to find the most appropriate methods to include molecular information in selection, starting with the identification of polymorphic loci in the genome which affect important traits in livestock. This has enabled opportunities to enhance genetic improvement programs in livestock by direct selection on genes or genomic regions that affect economic traits through Marker-Assisted Selection (MAS) and Gene Introgression (GI) (DEKKERS, 2004 [50]). GI means that favourable alleles can be introduced in a recipient line by special mating procedures. In subsequent generations the genetic structure of the population is modified positively. Because of the longer generation intervals and lower reproduction rates, GI was feasible in livestock only for major genes, such as the successful halothane GI in Pietrain line or Booroola gene (FecB) in dairy sheep breeds (reviewed by Dekkers (DEKKERS, 2004 [50]). Because of the limited number of major genes, in the next years, research was reoriented to finding other genome-wide reference variations (markers). For 20 years, markers were mapped and tested in selection, leading to more or less important progress. Although significant financial and human resources were mobilised in MAS in developed countries, the methods did not yield the expected benefits in livestock, crops, forest trees and farmed fish (JUANE & SONNINO, 2007 [51]). However, the use of markers remains important. They can be utilised in QTL mapping; in the estimation of the population's genetic structure, in genetic diversity studies, etc. Molecular markers are mutation or variation in DNA fragments, usually non-informational (hyper-variable DNA regions), sometimes even informational (functional mutation) involved in the detection of chromosomal regions that affect single-gene traits and QTL (COSIER & VLAIC, 2003 [52]). MAS applications started by identifying the genes and the genomic regions involved in coding complex quantitative traits in both animals and crops. To this end, two different directions were pursued. One involved basic understanding of physical maps or search for Quantitative Trait Loci (QTL), regions of the genome that control variations in quantitative traits. The second started from the biology of the trait, directly concerning one gene or several genes involved in the expression of the characteristics, and which are thus candidate genes in selection (ZHU & ZHAO, 2007 [6.], MEUWISSEN & al., 2001 [8]). This approach has proved useful in selection for traits that are difficult to improve by conventional methods. For the purpose of genetic improvement, markers can be used to enhance within-breed selection based on GAS (Gene Assisted Selection), LD-MAS (Linkage Disequilibrium - Marker Assisted Selection), or LE-MAS (Linkage Equilibrium - Marker Assisted Selection), or to enhance programs to capitalize on between-breed variation by selection within a cross. Dekkers, (DEKKERS, 2004 [50]). reports three types of polymorphic loci that can be used in selection: 1) direct markers, loci that code for the functional mutation; 2) LD markers (Linkage Disequilibrium), loci that are in population-wide linkage disequilibrium with the functional mutation; 3) LE markers (Linkage Equilibrium): loci that are in population-wide linkage equilibrium with the functional mutation in outbreed populations. The three types of loci are characterised by their own methods of polymorphism detection, and different selection methodologies. While direct markers require identification of the causative mutation in the genome, LD markers can be random markers and LE markers require close linkage with the causative mutation (DEKKERS, 2004 [50]). For quantitative traits, selection based

on marker effects alone has dramatically changed standard procedure used in animal breeding between 1980-1990 (GEORGES & al., 1995 [53], MEUWISSEN & GODDARD, 1996 [54], DEKKERS, and J.H.J VAN DER WERF, 2007 [55]) concluded that for successful application of MAS in breeding programmes is needed: gene and marker genotyping, QTL detection and association with economically important traits, integration of phenotypic and genotypic data in statistical methods and development of breeding strategies and programmes for the use of molecular genetic information in selection and mating programmes.

3.2.1. Direct markers

The direct markers are functional polymorphisms in genes that control variation in the quantitative traits (WILLIAMS, 2005 [56]). When direct markers exist for a QTL, the marker genotype gives a direct indication of the QTL genotype, and thus the marker genotype can be used in selection. As an application for rapid marker genotyping, commercial tests soon became available. Dekkers (DEKKERS, 2004 [50]) reviewed direct markers, LE and LD markers, available in commercial breeding programmes. Some examples of these direct markers are: a) markers in dairy cattle: k-Casein, ß-lactoglobulin, FM03; b) milk yield and composition markers: DGAT1 (reviewed by Williams) (WILLIAMS, 2005 [56]), GRH; c) growth and composition markers: in beef cattle – Myostatin; in pork – PRKAG3, RYR; d) disease markers: PrP in sheep and goats; e) reproduction markers: Booroola and Inverdale in sheep. Examples of LD markers: a) growth, composition and meat quality: CAST in pork; CAST, THYR, leptine in beef cattle; b) reproduction: ESR and PRLR in pig.

3.2.2. LE markers

Linked markers are only near QTL on the genome, and not the causative mutation in the gene concerned. When the marker is linked to QTL, the information on which marker genotype is linked to the positive QTL allele is family specific. LE markers can be detected in the genome using breed crosses, or through the analysis of large populations of half-sib families within the breed. In this case, scanning of the genome requires saturated maps of about 150 – 300 markers, depending on the marker's informativity and genotyping costs. For LE markers, it is necessary to gather trait and pedigree information because for each family the linkage phase between marker and QTL needs to be established. This was a limitation of this approach. Because the QTLs were mapped very imprecisely (often at confidence intervals of 50 cM) and the marker and QTL were in linkage equilibrium, the linkage phase varied between families. Consequently, the linkage phase had to be determined in each family before the marker could be used in selection (GODDARD & al., 2010 [57]). Nonetheless, this was approach was thought to be markedly enhanced by the identification of numerous highly variable microsatellite markers, and for this purpose, linkage analysis was often performed with large half-sibling families (GODDARD &. HAYES, 2009 [10]). Estimation of BV using markers in linkage with QTL provided few successful examples (DEKKERS, 2004 [50], BOICHARD & al., 2002 [58], so this method was rarely used in selection. Therefore, a new type of selection was necessary to emerge, which did not require the linkage phase to be determined for each family (GODDARD & al., 2010 [57]).

3.2.3 LD markers

In *LD - Marker assisted selection (MAS)* molecular markers are in linkage disequilibrium (LD) with quantitative trait loci (QTL). LD markers must be close to the functional mutation so that there is enough linkage disequilibrium between marker and QTL in the entire population (as a rule, within 1 cM, depending on the size of LD in the population, which in turn depends on the population structure and history). Linkage disequilibrium refers to alternative haplotype frequencies in a population and can be revealed by crossbreeding which differ in allele and haplotype frequencies. Although these loci are in parental

population in LE with QTL, after crossbreeding, they will be in partial LD with QTL, but only if the QTL frequencies are different in parental populations. This LD in the population can be used for QTL scanning and mapping. Identification of LD markers can be done by using the candidate gene approach (ROTHSCHILD & RUVINSKY, 2011 [59]) or the fine-mapping approach (ROTHSCHILD & SOOLER, 1997 [60]). Using LD between markers and QTLs, more precise mapping has become possible because LD decays quickly as the distance between marker and QTL increases (GODDARD & al., 2010 [57]). In conclusion, the LD markers will be by necessity close to the functional mutation for sufficient population-wide LD between the marker and QTL to exist (DEKKERS, 2004 [50]). The major factors that contribute to LD occurrence are mutation, selection, migration and drift. For the integration of phenotypic and genotypic data in statistical methods, several approaches have been suggested. These models can be distinguished by attribution of QTL effect as random or fixed, and by the type of genetic information from LE, LD or direct markers (GODDARD & al., 2010 [57]). Integration of molecular data in BLUP was demonstrated by Fernando and Grossman (Fernando & Grossman, 1989 [61]), who predicted increase in genetic gain by 8-38% in conventional MAS schemes. Nevertheless, because of the continuous decrease of variation of QTL, newer and newer QTLs had to be detected (MEUWISSEN , 2003 [62]). As a result, expectation of genetic progress enhancement failed in many quantitative traits, except for those with low heritability, sex-limited, and available after slaughter (DEKKERS, 2004 [50], MEUWISSEN & GODDARD, 1996 [54], (GODDARD & al., 2010 [57]). They found that, in low heritability characteristics, MAS could enhance the rate of genetic improvement up to 64 percent. Some QTLs of larger effect were however discovered after genome mapping of genes, such as the DGAT1 polymorphism which explains about 40% of the genetic variation in fat content in the milk of Holstein cattle (GRISART & al., 2002 [63]); while other mutations, such as some functional mutations in the myostatin gene, seem to be breed specific in beef cattle (GROBET & al., 1998 [64]) which limits their use in selection. In order to detect a new QTL for MAS, it was necessary to perform very stringent tests for statistical significance because many false positives can be found. In GS, the idea of significance test is to be omitted and the effects of all genes are simply estimated simultaneously (MEUWISSEN & al., 2001 [8]), which will eventually lead to high accuracy Estimate Genomic Breeding Value (EGBV). With the dense panels of markers covering the whole genome (SNPs) and in LD with the QTL, large increases in genetic response to selection were estimated by Meuwissen *et al*. (MEUWISSEN & al., 2001 [8]) to be achieved in the new form of genomic selection.

3.3 Genomic selection (GS)

Genomic selection (GS) is an approach proposed by Meuwissen *et al.* [8]. In fact, GS is a form of marker-assisted selection in which genetic markers (SNPs), covering the whole genome are used so that all quantitative trait loci (QTL) are in linkage disequilibrium with at least one marker (MEUWISSEN & al., 2001 [8], GODDARD &. HAYES, 2009 [10], GODDARD & HAYES, 2007 [65]. GS also refers to selection decisions based on genomic breeding values (GBV) calculated as the sum of the effects of dense genetic markers, or haplotypes of these markers (surrounding every 1-cM region) across the entire genome, thereby potentially capturing all the quantitative trait loci (QTL) that contribute to variation in a trait (HAYES & al., 2001 [66], HAYES & GODDARD , 2001[67], HAYES & al., 2009 [68]). In practice, GEBV are calculated by estimating SNP effects from prediction equations, which are derived from a subset of animals in the population (i.e., a reference population) that have SNP genotypes and phenotypes for traits of interest. Accuracy of the prediction of genome wide breeding value (GEBV) (HAYES & al., 2009 [68]) depends on the size of the

reference population, the heritability of the trait, and the extent of relationships between selection candidates and the reference population (HAYES & al., 2009 [68]), MULDER & al., 2010 [70], CHAMBERLAIN & al., 2012 [71]). In 2001, Meuwissen demonstrated by simulation that high accuracy in breeding value prediction (0.85) is possible based on data markers alone. Accuracy is the correlation between the true breeding value and the estimated breeding value, and reliability is the squared value of this result (HAYES & al., 2001 [66]). Recently, Lillehammer (LILLEHAMMER & al., 2013 [72]) demonstrated (by stochastic simulation) that within-family genomic breeding values increased genetic gain by 15% and reduced rate of inbreeding by 15% in aquaculture species in which genotyping of a large number of individuals would make the breeding program expensive. Notwithstanding, Jannick (JANNINK & al., 2010 [73]) considers that these studies in which predictions are made refer only to the first or second generations of GS and that long-term reliability of genetic gain is not justified, and also low frequency alleles that are favourable in the population can be lost. In stochastic simulation studies over several generations, on barley with 983 polymorphic markers, Jannick (JANNINK & al., 2010 [73]) designed different scenarios of GS application (before and after phenotyping; applying additional weight on low frequency favourable alleles) and concluded that (1) applying GS before phenotyping strongly increased early selection gain, but also caused the loss of many QTL favourable alleles, leading to loss of genetic variance and loss of GS accuracy, while (2) placing additional weight on low frequencies favourable alleles allowed GS to increase their frequency early on and lead to higher long term gain rather than short-term one. Moreover, the prediction equations derived for one breed may not predict accurate genome-wide breeding values when applied to other breeds, since family structures differ among breeds (LUAN & al., 2009 [69]).

4. Conclusions

Based on the latest data from animal production and crops, Steinfeld and Gerbe (STEINFELD & GERBE, 2010 [74]) state that progress obtained by the broad application of science and advanced technology in feeding and nutrition, genetics and reproduction, and animal health control as well as general improvements in animal husbandry was not nevertheless insignificant. The relationship between technological advancement and population growth has been concretised in the increase by 90% in crop production due to increased productivity rather than expansion of cultivated surface. Although the surface of arable land increased only by 30%, and the surface area of grassland by under 10%, the global meat and milk production increased by 245% and 75%, respectively. It is foreseen that soon the livestock sector will exceed the recently published threshold of sustainability in the area of climate change. The forecast global increase (for 2000-2030) of meat consumption by 67% and of milk by 57% cannot be sustained by ruminants because of the large quantities of methane gas produced; as an alternative, monogastrics (pigs and chicken) will play a substantial role in supporting consumption. With the availability of genome-wide dense molecular markers, GS has now become practical and its effectiveness in dairy cattle breeding has been demonstrated in many countries. (HAYES & al., 2009 [68], LUAN & al., 2009 [69], HARRIS & al., 2008 [75], BERRY & al., 2009 [76], VANRADEN & al., 2009 [77], GONZALEZ-RECIO & al., 2009 [78]). This method offers many advantages with regard to improving the rate of genetic gain in meat and dairy cattle breeding programs and is expected to double the rate of genetic improvement per year in many livestock systems (SCHEFERS, & WEIGEL. 2012 [9], HAYES & al., 2009 [68]). With the advent of cheap and high-density array-based DNA marker technologies, genome wide dense molecular markers are becoming available for livestock and crop species. Studies of the accuracy of genomic predictions have

emerged in some animal species, including chickens (GONZALEZ-RECIO & al., 2009 [78]), pork sheep, dogs, and also in crop species: barley (ZHONG & al., 2009 [79]), sorghum, wheat (ZHAO & al., 2012 [80]), *Arabidopsis thaliana* (L.), (HESLOT & al., 2012 [81]), maize (GUO & al., 2011 [82], ZHAO & al., 2011 [83]; and sugar beet (WÜRSCHUM & al., 2013 [84]). This is indicative of the major role that advanced technology can play in limiting environmental impact, particularly in developing country agriculture and livestock production where there are large productivity gaps (STEINFELD & GERBE, 2010 [74]).

References:

1. D. G. WANG, J.-B. FAN, C.-J. SIAO, A. BERNO, P. YOUNG, R. SAPOLSKY. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science*, 280: 1077-1082 (1998**).**
2. H. HARTWELL, G. REYNOLDS, S. VERES. Genetics - From gene to genome. McGraw-Hill College, (2006).
3. P. RUSSELL. Genetics. Pearson Education, Inc., San Francisco, USA, (2010), pp 254.
4. P.Y KWOK. Methods for genotyping single nucleotide polymorphisms. *Annu Rev Genomics Hum Genet.* (b), 2: 235-58(2001).
5. KWOK, P.Y. Single Nucleotide Polymorphism. Methods in Molecular Biology, Humana Press Inc., New-Jersey, USA, (2003).
6. M. ZHU, S. ZHAO. Candidate Gene Identification Approach: Progress and Challenges. *Int J Biol Sci*, 3(7):420-427 (2007).
7. L.T. ROSS, K.C. WORLEY. The genome sequence of taurine cattle: A window to Ruminant Biology and Evolution. *Science*, 348: 522-528. (2009).
8. T.H.E. MEUWISSEN, B.J. HAYES, M.E. GODDARD. Prediction of total genetic values using genome-wide dense marker maps. *Genetics*, 157, 1819-1829 (2001).
9. J.M. SCHEFERS, K.A. WEIGEL. Genomic selection in dairy cattle: Integration of DNA testing into breeding programs. *Anim Front*, 2(1), 4-9.11(2012).
10. M.E. GODDARD, B.J. HAYES. Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nat Rew Gen.* 10: 381-391(2009).
11. H. SMITH, K.W. WILCOX. A Restriction enzyme from *Hemophilus influenzae 1I.* Purification and general properties. *J of Mol Biol*, 51 (2): 379–391 (1970).
12. H.O. SMITH, D. NATHANS. Letter: A suggested nomenclature for bacterial host modification and restriction systems and their enzymes. *J. Mol. Biol.* 81(3): 419–23 (1973).
13. W. ARBER, S. LINN. DNA modification and restriction. *Annu Rev Biochem*, 38, 467–500 (1969).
14. J.A. GERLT. Mechanistic Principles of Enzyme-Catalyzed Cleavage of Phosphodiester Bonds. In Nucleases, 2nd Ed., Cold Spring Harbor Archive, (1993).
15. B. ALBERTS, A. JOHNSON, J. LEWIS. Molecular Biology of the Cell, 4th ed., Garland Science, New York, NY, USA. (2002).
16. H. LODISH., A. BERK, P. MATSUDAIRA, C.A KAISER, M. KRIEGER, M.P. SCOTT, L. ZIPURSKY, J. DARNELL. *Mol Cell Biol*, 5th ed, Freeman W.H., (2003).
17. S.N. COHEN, A.C.Y. CHANG, H.W. BOYER, R.B. HELLING. Construction of biologically functional bacterial plasmids *in vitro*, *Proc Natl Acad Sci USA.*, 70 (11), 3240-3244 (1973).
18. T.A. BROWN. Gene cloning and DNA analysis, 5th Ed., Blackwell Publishing, Oxford, UK.(2008).
19. V. COSIER. Inginerie genetică, Ed. Risoprint, Cluj-Napoca, Ro (2008), pp 113-126.
20. K. ITAKURA, T. HIROSE, R. CREA, A. RIGGS; H.L. HEYNEKER, F. BOLIVAR, H. BOYER. Expression in *E.coli* of a chemically synthesized gene for the hormone somatostatin. *Science*, 198, 1056-1063(1977).
21. D.V. GOEDDEL, H.L. HEYNEKER, T. HOZUMI, D.V. GOEDDEL, H.L. KLEID, D.G. BOLIVAR, F. HEYNEKER, H.L. YANSURA, D.G. CREA, R. HIROSE, R., T. KRASEWSKI, K. ITAKURA, A.D. RIGSS. Expression of chemically synthesized genes for human insulin. *Proc. Natl. Acad. Sci. USA*, 76, 106-110(1979).
22. R. ARENTZEN, K. ITAKURA, D.G. YANSURA, M.J. ROSS, G. MIOZZARI, G., R. CREA, P. SEEBURG. Direct expression in *Escherichia coli* of a DNA sequence coding for human growth hormone. *Nature*, 281, 544-548(1979).

23. A.M. CHAKRABARTY, D.A. FRIELLO. Dissociation and interaction of individual components of a degradative plasmid aggregate in *Pseudomonas. Proc Nat Acad Sci USA*, 71 (9): 3410–4 (1974).
24. J.F.K. BATTEY, M. SCHMIDT, J.B. OHLRROGE. Genetic engineering for plant oils: potential and limitation. *Trends Biotech*, 7, 122-125 (1989).
25. C.S. GASSER, R.T. FRALEY. Genetically engineering plants for crop improvement. *Science*, 244, 1293-1299 (1989).
26. A. HIATT, R. CAFFERKEY, K. BOSDISH. Production of antibodies in transgenic plants, *Nature*, 342, 76-78 (1989).
27. A.J. CLARK, H. BESSOS, O. BISHOP, S. HARRIS, R. LATHE, M. MCCLEANAGHAN, C. PROWSE, J.P. SOMONS, C.B.A. WHITELAW, I. WILMUT. Expression of human antihemophilic Factor IX in the milk of transgenic sheep. *BioTech* 7, 487 – 492 (1989).
28. R.K. SAIKI, D.H. GELFAND, S. STOFFEL, S.J. SCHARF, R. HIGUCHI, G.T. HORN, K.B. MULLIS, H.A. ERLICH. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase, *Science*, 239 (4839): 487–91 (1988).
29. K.B. MULLIS. The unusual origin of the polymerase chain reaction. *Sci Amer*, 262(4), 64-5 (1990).
30. K.B. MULLIS. Target amplification for DNA analysis by the polymerase chain reaction. *Ann Biol Clin* (Paris), 48(8):579-82 (1990).
31. A.J. JEFFREYS, V. WILSON, S.L. THEI. Hypervariable "minisatellite" regions in human DNA, *Nature*, 314, 67–73(1985).
32. A.J. JEFFREYS, V. WILSON, S.L. THEI. Individual-specific "fingerprints" of human DNA. *Nature*, 316, 76–79 (1985).
33. U.G. MUELLER, L.L. WOLFENBARGER. AFLP genotyping and fingerprinting. *Trends Ecol Evol*. 14(10):389-394 (1999).
34. R.A. EELES, C.A. STAMP. Polymerase Chain Reaction (PCR): The Technique and Its Application. R. G. Landes, Austin TX USA, (1993).
35. T.A. BROWN. *Genomes*. 2nd Ed., Oxford: Wiley-Liss.UK (2002).
36. .D.J. LOCKHART, H. DONG, M.C. BYRNE, M.T. FOLLETTIE, M.V. GALLO, M.S. CHEE, M. MITTMANN, C. WANG, M. KOBAYASHI, H. HORTON, E.L. BROWN. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol.* 14(13), 1675-80 (1996).
37. J. SCHIMENTI, M. BUCAN. Functional genomics in the mouse: phenotype-based mutagenesis screens. *Genome Res*. 8 (7):698-710 (1998).
38. M. SCHENA, R.A. HELLER, T.P. THERIAULT, K. KONRAD, E. LACHENMEIER, R.W. DAVIS. Microarrays: biotechnology's discovery platform for functional genomics. *Trends Biotechnol*, 16, 301-306 (1998).
39. A.M. MAXAM, W. GILBERT. A new method for sequencing DNA. *Proc. Natl Acad. Sci.* USA, 74, 560–564 (1977).
40. F. SANGER, A.R. COULSON, T. FRIEDMANN, G.M. AIR, B.G. BARRELL, N.L. BROWN, J.C. FIDDES., C.A. HUTCHISON, 3RD, P.M. SLOCOMBE, M. SMITH. The nucleotide sequence of bacteriophage phiX174. *J Mol Biol.*, 125(2):225-46 (1978).
41. C.A. HUTCHINSON. DNA sequencing: bench to bedside and beyond, *Nucleic Acids Res*, 35(18), 6227–6237(2007) .
42. F. SANGER, S. NICKLEN, A.R. COULSON. DNA sequencing with chain-terminating inhibitors. *Biotechnol*, 24, 104-8(1992).
43. S. ANDERSON. Shotgun DNA sequencing using cloned DNase I-generated fragments. *Nucleic Acids Res*. 9, 3015–3027 (1981).
44. J. PEIFENG, Z. YAN, L. CHAO, Z. ZIXIA, W. JIAN, L. JIONGTANG, X. PENG, S. XIAOWEN. High Throughput Mining and Characterization of Microsatellites from Common Carp Genome. *Int. J. Mol. Sci.* 13, 9798-9807 doi:10.3390/ijms13089798 (2012).
45. L.M. SMITH, J.Z. SANDERS, R.J. KAIZER, P. HUGHES, C. DODD, C.R. CONNELL, C. HEINER, S.B. KENT, L.E. HOOD. Fluorescence detection in automated DNA Sequence analysis, *Nature*. 321, 674-679 (1986).
46. R.E. MARDIS. Next Generation DNA Sequencing Methods. *Annu. Rev. Genomics Hum. Genetics*. 9, 387-402(2008).
47. R. LISTER, B.D. GREGORY, J.R. ECKER. Next is now: New technologies for sequencing of genomes, transcriptomes and beyond. *Curr Oppin Plant Biol.* 12 (2), 107-118(2009).
48. D.G. WANG, J.-B. FAN, C.-J. SIAO, A. BERNO, P. YOUNG, R. SAPOLSKY. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science*, 280, 1077-1082 (1998).

49. V. COSIER, A. VLAIC. Abordarea practică a problemelor de genetică animală, Ed.Rispoprint, Cluj-Napoca, CN, Ro, (2011), pp 172.
50. J.C.M. DEKKERS. Commercial application of marker- and gene-assisted selection in livestock: Strategies and lessons. *J Anim Sci*, 82:E313-E328 (2004**).**
51. J. JUANE, A. SONNINO. Marker assisted selection as a tool for genetic improvement of crops, livestock, forestry and fish in developing countries: An overview of the issues. Marker Assisted Selection. Current status and further perspectives in crops, livestock, forestry and fish. FAO of the United Nations, Rome, (2007).
52. V. COSIER, A. VLAIC. DNA Polymorphism and techniques used in generating molecular markers. *Bull USAMV – CN ASB*, 59, 215-220 (2003).
53. M.D. GEORGES, M. NIELSEN, A. MACKINNON, R. MISHRA, I. OKIMOTO. Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing. *Genetics*, 139, 907–920 (1995).
54. T.H.E. MEUWISSEN, M.E. GODDARD. The use of markers haplotypes in animal breeding schemes. *Gen Sel Evol.* 28, 161-176 (1996).
55. J.C.M. DEKKERS, and J.H.J VAN DER WERF. Marker-assisted selection – Current status and future perspectives in crops, livestock, forestry and fish. Chapter 10. Guimaraies EP (ISBN 978-92-5-105717-9) (2007).
56. J.L. WILLIAMS. The use of marker assisted selection in animal breeding and biotechnology. *Rev Sci Techn Off Int Epiz*. 24(1):379-388 (2005).
57. M.E. GODDARD, B.J. HAYES, T.H. MEUWISSEN. Genomic selection in livestock populations. *Genet Res* (Camb)**,** 92(5-6): 413-21 (2010**).**
58. D. BOICHARD, S.FRITZ, M.N. ROSSIGNOL, M.Y. BOSCHER, A. MALAFOSSE, J.J. COLLEAU. Implementation of Marker Assisted Selection in French dairy cattle. El. Communication 22-03 in Proc. 7th World Cong *Genet. Appl. Livest*. *Prod*., Montpellier, France (2002).
59. M.F. ROTHSCHILD, A. RUVINSKY. The genetics of the pig, Second Ed., Hardback, UK, 407-410, (2011).
60. M.F. ROTHSCHILD, M. SOLLER. Candidate gene analysis to detect genes controlling traits of economic importance in domestic livestock. *Probe* 8:13-29 (1997).
61. R.L. FERNANDO, M. GROSSMAN. Marker assisted selection using best linear unbiased prediction. *Genet. Sel. Evol*, 21: 467-477 (1989**)**.
62. T.H.E. MEUWISSEN. Genomic selection: The future of Marker Assisted Selection and Animal Breeding. Workshop "Marker assisted selection: A fast track to increase genetic gain in plant and animal breeding?" Turin, Italy, 17-18 October (2003).
63. B. GRISART, F. COPPIETERS, F. FARNIR, L. KARIM, C. FORD, N. CAMBISANO, M. MNI, S. REID, R. SPELMAN, M. GEORGES, R. SNELL. Positional candidate cloning of QTL in dairy cattle: identification of missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition. *Genome Res* 12: 222-231(2002).
64. L. GROBET, D. PONCELET, L.J. ROYO, B. BROUWERS, D. PIROTTIN, C.MICHAUX, F. MENISSIER, M. ZANOTTI, S. DUNNER, M. GEORGES. Molecular definition of an allelic series of mutation disrupting the myostatin function and causing double-muscling in cattle. *Mamm. Genome* 9:210-213 (1998).
65. M.E. GODDARD, B.J. HAYES. Genomic selection, *J Anim Breed Genet*, 124(6), 323-30 (2007**)**.
66. B.J. HAYES, P.J. BOWMAN, M.E. GODDARD. Linkage disequilibrium and accuracy of predicting breeding values from marker haplotypes: 269-272 in *Procc Assoc Adv mt Anim Breed Genet.*, Queenstown, New Zeeland, (2001).
67. B.J. HAYES, M.E. GODDARD. The distribution of effects of genes affecting quantitative traits in livestock. *Genet. Sel. Evol,* 33 (3): 209-229 (2001).
68. B.J. HAYES, P.J BOWMAN, A.J. CHAMBERLAIN, M.E. GODDARD. Genomic selection in dairy cattle: Progress and challenges. *J. Dairy Sci*., 92: 433–443(2009).
69. T. LUAN, J.A. WOOLLIAMS, S. LIEN, M. KENT, M. SVENDSEN, T.H.E. MEUWISSEN. The Accuracy of Genomic Selection in Norwegian Red Cattle Assessed by Cross-Validation. *Genetics* 183 (3): 1119-1126. (2009).
70. H.A. MULDER, M.P.L. CALUS, R.F. VEERKAMP. Prediction of haplotypes for ungenotyped animals and its effect on marker-assisted breeding value estimation, *Gen Sel Evol.*42, 10 (2010**).**
71. A.J. CHAMBERLAIN, B.J. HAYES, K. SAVIN, S. BOLORMAA, H.C. MCPARTLAN, P. BOWMAN, C. VAN DER JAGT, S. MACEACHERN, M.E. GODDARD. Validation of single nucleotide polymorphisms associated with milk production traits in dairy cattle. *J Dairy Sci.* 95(2): 864-75 (2012).

72. M. LILLEHAMMER, T.H. MEUWISSEN, A.K. SONESSON. A low-marker density implementation of genomic selection in aquaculture using within-family genomic breeding values. *Genet Sel Evol.*, 45(1): 39 (2013).
73. J.L. JANNINK, A.J. LORENZ, H. IWATA. Genomic selection in plant breeding: from theory to practice. *Func Genom and Proteom*: 9 (2): 166 – 177 (2010).
74. H. STEINFELD, P. GERBE. Livestock production and the global environment: Consume less or produce better? *PNAS* 107 (43) :18237–18238 (2010).
75. B.L. HARRIS, D.L. JOHNSON, R.J. SPELMAN. Genomic selection in New Zealand and the implications for national genetic evaluation. Proceedings of the Interbull Meeting, Niagara Falls, NY, USA, 16-18 June, (2008).
76. D. BERRY, F. KEARNEY, B. HARRIS. Genomic selection in Ireland. In Proceedings of the Interbull International Workshop: Genomic Information in Genetic Evaluations, Uppsala, Sweden. *Interbull Bull*., 39: 29–34 (2009).
77. P.M. VANRADEN, C.P. VAN TASSELL, G.R. WIGGANS, T.S. SONSTEGARD, R.D. SCHNABEL, J.F. TAYLOR, F. SCHENKEL. Reliability of genomic predictions for North American Holstein bulls. *J. Dairy Sci*. 92: 16–24 (2009).
78. O. GONZALEZ-RECIO, D. GIANOLA, G.J.M. ROSSA, K.A.WEIGEL, A. KRANIS. Genome-assisted prediction of a quantitative trait measured in parents and progeny: application to food conversion rate in chickens. *Genet .Sel. Evol.* 41(1): 1–10 (2009).
79. S.J. ZHONG, C.M. DEKKERS, R.L. FERNANDO, J.L. JANNINK. Factors affecting accuracy from genomic selection in populations derived from multiple inbred lines: a barley case study. *Genetics*, 182: 355–364 (2009**)**.
80. Y. ZHAO, M. GOWDA, W. LIU, T. WÜRSCHUM, H.P. MAURER, C.F.H. LONGIN, N. RANC, J.C. REIF. Accuracy of genomic selection in European maize elite breeding populations. *Theor Appl Genet.* 124: 769-776 (2012).
81. N. HESLOT, H.-P. YANG, M.E. .SORRELLS, J.L. JANNINK, Genomic Selection in Plant Breeding: A Comparison of Models. *Crop Sci.* 52(1): 146-160 (2012).
82. .Z. GUO, D.M. TUCKER, J. LU, V. KISHORE, G. GAY. Evaluation of genomewide selection efficiency in maize nested association mapping populations. *Theor Appl Genet*. 124, 261-275 (2011**)**.
83. Y. ZHAO, J. ZENG, R.L. FERNANDO, J.C. REIF. Genomic prediction of hybrid wheat performance. Crop Sci, 53, 802-810 (2013).
84. T. WÜRSCHUM, J.C. REIF, T. KRAFT, G. JANSSEN, Y. ZHAO. Genomic selection in sugar beet breeding populations. *BMC Genetics*, 14, 85 (2013).