DNA in the service of selection and improvement of livestock. The development and application of genomics as a new tool for improvement

Mamuris Zissis, Professor of Genetics, Department of Biochemistry and Biotechnology, University of Thessaly

Introduction

Over the last 50 years, we have experienced an unprecedented increase in the human population. Based on current projections, world population will reach 9 billion in 2030. To meet the growing need for food using fewer resources is one of the greatest challenges that modern agriculture faces. Recent estimates by the Food and Agriculture Organization of the United Nations show that in order to meet the growing demand, food production must be doubled in the next 50 years. In other words, agriculture will have to produce more food than in the last 10,000 years altogether. When the year 2000 is used as a basis, predictions show an increase in global meat consumption by 68% and milk by 57% by 2030. The greatest demand for food based on animal proteins with the potential impact of climate change and water shortage, nutrients and energy will lead to big gaps in productivity. It is therefore vital that systematic application of technical and scientific methods to improve the food, nutrition, genetics, breeding, control of animal health, and the general improvement of livestock is performed in order to meet the upcoming productivity gaps. The greatest gains will come from innovations that accelerate agricultural productivity while reducing costs and reducing environmental impact.

The traditional breeding methods are successful, but limited

For centuries, farmers have very effectively manipulated the genomes of productive species, based on the fact that there are natural variations within a species, in a race, and in a population. Traditional breeding and improvement has happened due to the lack of knowledge of the molecular composition of genes controlling quantitative traits of animals. Breeders have improved productive traits in their herds, choosing the best individuals as ancestors for future generations. These improved "breeding values” were achieved by combining the phenotypic recording of individual performance with pedigree information. In Holstein dairy cattle, milk production continues to increase by 110 kg per animal per year. In pig production, the pounds of food required to produce one kilogram of pork, a parameter known as the feed conversion is estimated to have decreased by 50% between 1960 and 2005. Although these results are powerful examples of what can be achieved through traditional breeding methods, the effectiveness of these traditional methods is reduced when the traits are difficult to measure, have low heritability, or cannot be measured quickly, inexpensively and correctly on a large number of animals. Such difficulties in measuring and recording features are often extremely important because they include fertility, longevity, feed efficiency and disease resistance. The selection of these features must therefore be achieved through genomics.

The high expectations of Genomics

During the last 2 decades, the rapid development of genomics has opened new ways for the examination of the scientific basis of livestock biology and reproduction and has led to new production methods to achieve steady growth of feed yields and long-term improvements in efficiency of animal production. A new era, the “genomics era”, promises to enable an objective prediction of the effects that are based on direct access to the complete sequence of the DNA of many individuals, and therefore a renewed and more objective view of the genetic value of animals that is not limited to just a few features of production. One of the factors that triggered the development of this genomic era was the international program to decode the human genome. The goal of this project was to produce the first (de novo) complete DNA sequence of a human being. Along with this came the development and application of new genomic tools, especially improved technologies of sequencing of DNA and the increased availability of high-performance platforms for the analysis of genotypes. The value for the sequencing of a single nucleotide of DNA has been reduced by 100 million times since 1990. Technological innovations that have led to this cost reduction have also facilitated the sequencing of entire genomes of many species.

Genomics: The passage of the Science of Animal Production in a new dimension

From a scientific point of view, the acceleration of the analysis of genomes on a wide scale will have a significant impact. It will give new information to the understanding of the basic structure and function of the livestock’s genomes and it will further explain the control of complex traits. With the production of whole genome sequences for the major animal species at very low costs, a comparison of sequences from different individuals of different races to a reference sequence resulted in the discovery of an almost inexhaustible source of genetic markers, especially polymorphisms in the form of single nucleotide polymorphism ( SNP). Thus, the scientific community has gained access to millions of SNPs, but also free accessible databases (http://www.ncbi.nlm.nih.gov/projects/SNP). These databases contain information today for over 86 organisms and a total of over 50 million SNPs. Another important technological breakthrough was the development and continuous improvement of the technology of DNA arrays, allowing inexpensive analysis of SNPs in a given sample. The success of these DNA arrays is that they show a strong parallel processing ability and a remarkable ability to automate. Although first used for gene expression studies, these DNA arrays have proven very useful for the development of whole genome SNP panels for many species, including many productive species. With DNA-arrays, hundreds of thousands of SNP can be tested in parallel, allowing scientists to perform genome association studies, which would be impossible with other techniques or markers. In recent years, several studies have been published that demonstrate their effectiveness (e.g., BovineSNP50, OvineSNP50, EquineSNP50, PorcineSNP60).

Genomics: a paradigm shift in animal breeding

The greatest evolution accomplished in the application of genomics in the design and implementation of breeding programs and promises profits throughout the value chain. For farmers and other members of the livestock industry, genomics is expected to increase the efficiency and productivity of livestock and for consumers and the manufacturing sector, is expected to enhance the safety and quality of animal products. The new knowledge acquired on the development, nutrition, health, and animal welfare should allow a better understanding of the molecular mechanisms of the traits of commercial interest. Therefore, genomics, tapping into new sources of genetic polymorphisms, creates further opportunities to improve the selection accuracy, while reducing the costs and time between generations.

Selection principle via Genomics

Selection via genomics was first described by Meuwissen et al. (2001) and is based on the fundamental principle that the information of a large number of markers could be used for the estimation of breeding value without precise knowledge of the location on the genome of these specific genes that control those characteristics. With tens of thousands of SNPs, well selected to be representative of the entire genome, it is expected that there will always be a SNP in the immediate vicinity of a specific gene or part of the DNA of interest and that the existing linkage disequilibrium between one (or more) SNPs and a gene allele will be significant and may then be used to explain a significant fraction of the variation of the desired characteristic. The first step in the selection process through genomics is, therefore, access to a large group of animals, or in a reference population with specific phenotypes for the targeted characteristic (s). This population should also be genotyped using either a series of SNPs across the genome, either already known SNPs in cases where these have been designated. Then, the data obtained will serve as a reference for the development of a statistical interaction estimation model of each SNP with the targeted characteristic (s). The result is a predictive equation for calculating a genomic estimated breeding value (GEBV). After a validation step, the genomic breeding value of new animals can be estimated using the prediction equation, based on the genotypes of the SNP array and the absence of any precise information about the phenotype of these animals (Figure 1 and Table 1). The accuracy of the GEBV depends on the size of the population and the heritability of the targeted characteristic.

Table 1. Example of a simplified calculation of the genomic breeding value, with 4 SNPs and estimated effects





Figure 1. A reference animal population is evaluated and graded for key production traits and genotypes using SNPs. The genotypes are represented by the variable X, with values ​​0, 1, 2 (homozygous, heterozygous or homozygous alternative). A predictive equation is generated by combining all genotypic markers with their results to calculate a genomic estimated breeding value for each animal. This prediction equation can be applied even in a group of animals which have not been phenotyped, breeding values ​​can be evaluated, and the best animals can be selected for playback. Adapted from Goddard and Hayes (2009)

The Application of Genomic Selection

The Genomic selection is based on existing breeding programs in which collecting genealogical information combined with phenotypic data is already a routine. It provides a new level of information that can be incorporated into the decision-making process for identifying and selecting the most productive animals. The main advantages of genomic selection is that it can be implemented very early in the life of the animal, is not limited to a particular sex, and may be extended to any characteristics which are recorded in a reference population. It provides, especially for hard to improve traits, a better selection accuracy while reducing the time between generations, thus increasing the intensity of selection. Moreover, it is not limited to particular groups. Since 2006, Schaeffer (2006) showed that using genomic selection, genetic gain per year could double in dairy cattle, with the potential to reduce costs for the supply of bulls by more than 90%. Instead of cattle passing through a long and costly progeny testing by recording the phenotypic information for a large number of daughters, the exact GEBV data could be calculated through a cost-effective genomic study of the most suitable genotypes. Obvious benefits observed in dairy cattle can also be transferred to other species. The wider use of the genomics approach, or even the genomic selection, promises to be one of the next major advances in breeding programs for all animal species. As efficient genomics tools for animals, plants and fish continue to be developed, the respective breeding industries will be able to take decisions on selection earlier, to improve the characteristics that are difficult to treat with traditional methods, and provide customers with high-quality, safer food, while reducing the impact of farming on the environment (environmental footprint) and ensuring long-term sustainability products and livestock. The applications of genomics approach on the field extend far beyond reproduction, because genomics tools can also provide accurate information on animal identification, validation of paternity (parentage) and traceability of food and animal breeds. Genomics can also be applied to the flock management process to optimize the mating and to reduce inbreeding.

The Next Wave of discoveries and new approaches

Arrays with higher density SNPs with several hundred thousand SNP have been developed for different species. In cattle, the success of genomic selection is extended by combining various breeds in order to increase the size of the reference population and generate ratings between tribes. This approach will prove very beneficial for breeds with a limited number of individuals or phenotypic records, or for species for which the reproduction between the races is an effective tool in the improvement process. As the cost of sequencing continues to be reduced and the access to the sequencing of the entire genome for specific individual is accessible, one of the next steps would be to include the sequencing data of the whole genome in routine genetic evaluations. According to a simulation presented by Meuwissen and Goddard (2010), a 40% gain in the accuracy of prediction of genetic values ​​could be achieved by using the full sequencing instead of only the data of data from 30,000 SNPs arrays. Furthermore, using the sequencing data of the whole genome, the prediction of the genetic value remained accurate even when the reference and evaluation data were 10 generations apart, since the accuracy observed was similar to that for data from the same generation.

Opportunities for developing countries

In developed countries, phenotypes and lineages have been recorded for certain species, such as cattle, for more than 100 years. Progeny control is applied for 50 years. The implementation of improvement programs in developing countries is often limited by the absence of phenotype recording programs for various animal breeds and the lack of evaluation and national testing programs for assessing the genetic value of sires. The genomics approach should help identify critical populations for the maintenance of certain local breeds that are well adapted and could be used further to reproduce valuable animals through a combination of selection and breeding. Of course, as with genomics, we can manage only that we can measure and the collection of a minimum number of phenotypes will remain one of the crucial and difficult steps for the further development of genomic selection in developing countries.

Conclusions and Prospects for the Future

The ability to explore the genome, the trascriptome the epigenome and metagenome of any species with high efficiency sequencing methods opens up a new world of possibilities. Further reduction of sequencing costs will continue to lead to wider acceptance of new approaches and their application for the benefit of research on livestock and consumers. All cost-effective species, subspecies, and their pathogens will undoubtedly be sequenced in the near future. Thousands of related genomes will also be sequenced to test the genetic diversity within and between genetic groups of stocks, providing important information for applying selection genomics programs in developed countries. The Genomic selection will surpass the conventional methods and specific genotypes that are detected by high performance genotyping systems, can be directly related to economic values. Breeding programs will be guided mainly by genomics data because of the higher cost and much higher performance. New expertise in the field of animal pharmacogenomics will help increase vaccines and the specificity of the drug while nutri-genomics will help create diets designed based on the genomic profile. As genomics will continue to provide extremely valuable biological information, the key to the further success of genomic selection and genomic approach will be to collect the most suitable phenotypes to identify the alleles that determine them and the precise mechanisms by which these are produced, and to combine these alleles in breeding lines in as few generations as possible. We are entering a truly exciting time powered by genomics.

Review

* There is a significant difference between demand, based on population growth, and the current state of the livestock performance. This is the entry point of genomics to livestock improvement.
* Although the traditional breeding methods have been proven effective in the selection of animals with easy measurement of productive characteristics, such methods have substantially reached their limits and characteristics that are more difficult to measure (and often the most important) are not selected efficiently by using traditional methods
* The race of competition for sequencing the first human genome, and then the struggle to make sequencing affordable for dozens, if not hundreds of thousands, additional human genomes has led to an unprecedented (100 million times) reduction the price of sequencing DNA from 1990. Sequencing of animal genomes has benefited from this.
* Sequencing the genome of economically important species of livestock has resulted in the discovery of millions of single nucleotide polymorphisms (SNP). These single nucleotide polymorphisms have been developed parallel to DNA microarrays, allowing massive genome association studies to identify the genotype-phenotype correlations for both simple and, more importantly, for complex traits.
* Guided by a growing reduction of the cost of measuring genetic diversity, we are entering a new era in which the information from these studies on genome association level should be effectively used in routine tests using genomic selection. Selection through genomics analysis promises more promising results because it does not have the requirement of a priori knowledge for the position of alleles or molecular markers and the requirement that the selection by molecular marker should be implemented within families.