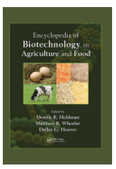

Encyclopedia of Biotechnology in Agriculture and Food
Publication details, including instructions for authors and subscription information:
http://tandfonline.com/doi/book/10.1081/E-EBAF

Livestock Marker-Assisted Selection

Kwame A. Darfour-Oduro [a] , Lawrence B. Schook [a]
[a] Department of Animal Sciences, University of Illinois, Urbana, Illinois, U.S.A.

Published online: 25 Jun 2012

To cite this entry: Kwame A. Darfour-Oduro, Lawrence B. Schook. Livestock Marker-Assisted Selection. *In* Encyclopedia of Biotechnology in Agriculture and Food. Taylor and Francis: New York, Published online: 25 Jun 2012; 1-3.

To link to this chapter: http://dx.doi.org/10.1081/E-EBAF-120043003

Livestock Marker-Assisted Selection

Kwame A. Darfour-Oduro
Lawrence B. Schook
Department of Animal Sciences, University of Illinois, Urbana, Illinois, U.S.A.

Abstract
Livestock marker-assisted selection (MAS) exploits linkage disequilibrium between DNA markers and quantitative trait loci of economic importance. Next-generation sequencing technologies have led to the discovery of genomic markers rapidly and in a cost-effective manner, and criteria for identifying single nucleotide polymorphism (SNP) with appropriate minimum allelic frequencies (MAF >0.05) to ensure informative SNPs are incorporated into genotyping arrays. Genome-wide association studies rely on genotyped animals to identify associations between SNPs and the presence of a trait of interest in existing populations. In livestock production, the future for MAS will be the prediction of genetic merit of animals by the simultaneous use of highly dense, uniformly distributed genome-wide markers through genomic selection approaches.

INTRODUCTION

Genetic markers are heritable specific DNA sequence differences located in the genome. Marker-assisted selection (MAS) is the selection of traits of interest indirectly by selecting on genetic markers in linkage disequilibrium (LD) with quantitative trait loci (QTLs). In this entry, we first describe the concepts and technologies underlying MAS as it pertains to livestock. We further discuss the drawbacks of MAS and finally conclude by noting future trends of MAS in livestock production.

CONCEPTS AND TECHNOLOGIES OF LIVESTOCK MAS

For livestock species, the majority of traits of economic importance are under the influence of numerous genes. Genomic regions containing one or more of the contributing genes that influence a given trait are referred to as a QTL. MAS seeks to increase the number of animals exhibiting favorable marker alleles linked to a trait. Over the course of time, single nucleotide polymorphisms (SNPs) have become the markers of choice. Compared with short tandem repeats (STRs), SNP markers are more powerful markers because of their abundance (1/100 to 1/1000 bp vs. 1/10 to 1/100 kbp) in the genome and the fact that they possess mutation rates (10^{-8} vs. 10^{-3}) that are approximately 100,000 times lower than STRs.[1]

NEXT-GENERATION SEQUENCING (NGS)

In NGS, whole genome or specific regions within the genome are randomly digested with restriction enzymes into small fragments (short reads) that are sequenced and aligned to a reference genome. Reads generated by NGS are mapped to the reference genome, and sequence variation between reference genome and reads are detected as SNPs. Developments of NGS methods have led to a fast decline in the cost of DNA sequencing, with a doubling of the number of base pairs that can be sequenced for one dollar ($1) every 2 yr and a quantum leap in efficiency.[2]

SNP ARRAYS

High-density SNP arrays are made up of tens of thousands of SNPs distributed throughout the genome and support the interrogation of hundreds of loci at a low cost. With improvements in genotyping technology, there has been a 10,000-fold reduction in the cost of genotyping overtime (Table 1). Candidate SNPs to be included in the design of SNP arrays should be validated to ensure the inclusion of SNPs that are informative. In the design of PorcineSNP60 array,[3] SNPs were assigned to waves based on an Illumina design score, the number of beads required to interrogate an SNP, and SNP minimum allelic frequency (MAF). Rounds of SNP selection were then conducted to select SNPs from waves with high stringency. The usefulness of this assay to MAS is evidenced by an average MAF of >0.25 for all SNPs.

GENOME-WIDE ASSOCIATION STUDIES

With the advent of NGS methods and SNP arrays for livestock species, genome-wide association studies (GWAS) are now used to test each SNP marker for a significant association with a trait. Such significant markers are then

Encyclopedia of Biotechnology in Agriculture and Food DOI: 10.1081/E-EBAF-120043003

Table 1 Reduction in the cost of genetic marker genotyping over time.

Year	Genetic marker	Genotyping technology	Cost of genotyping ($)/genotype
1986	RFPL	Southern blot	10
1996	STR	PCR assays	1
2006	SNP	SNP arrays	0.001

used in MAS. GWAS relies on LD arising as a result of physical linkage where the SNP marker is found at a chromosomal location near the gene affecting a trait.[4]

WHOLE GENOME APPROACHES FOR GENE DETECTION

Strategies to localize and characterize genes affecting complex traits of livestock broadly fall under two categories: genome scan and candidate gene approach.[5] The genome scan approach identifies chromosomal regions associated with a trait by studying the relationship between the trait and DNA markers selected across the genome. The candidate gene approach exploits the relationship between the trait of interest and genes that have possible roles in the physiology of the trait. Genome scans are labor intensive and locate the chromosomal region of a QTL with a large number of candidate genes. Candidate gene approach relies on information on biological function of genes, which may sometimes be scanty or unavailable. Comparative genomic analysis has also been employed in gene detection. This approach relies on the availability of whole genome sequences for species to discover genes affecting traits in livestock. The strategy is efficient if functionally conserved or structurally homologous genes influencing phenotypic variation in a trait are already confirmed in other species apart from the species of interest.[6] Digital candidate gene approach has recently emerged for the identification of candidate genes. According to Zhu and Zhao,[6] the approach makes computational identification of potential candidate genes of interest by extracting, filtering, or reanalyzing all possible resources derived from the public web databases mainly in accordance with the principles of biological ontology and complex statistical methods.

APPLICATION OF MAS IN LIVESTOCK POPULATIONS

MAS enhances within-breed selection to bring about genetic improvements in livestock populations through the use of causative mutation within a gene that controls variation in a trait and markers that are in LD with QTLs affecting a trait. Breeding values (BVs) are then estimated for selection candidates based on three types of information: marker, pedigree, and phenotypic information. DNA marker testing determines the marker allele an animal has for qualitative and quantitative traits in livestock. A list of some traits in livestock for which marker tests are available is shown in Table 2.

STRENGTHS AND LIMITATIONS OF MAS APPLICATION FOR LIVESTOCK

In a French MAS program in dairy cattle, estimated breeding values (EBVs) for MAS for all traits considered were more reliable than EBVs estimated from classical selection methods (mean gains of reliability ranged from 0.015 to 0.094 in 2004 and from 0.038 to 0.114 in 2006),[7] demonstrating that MAS may lead to increases in genetic improvement as compared with traditional breeding methods. Marker tests have also been used in identifying both recessive alleles associated with diseases. A major drawback in the implementation of livestock MAS is that population-based GWAS is unable to detect SNP loci showing association with a trait if the rarer allele at that loci has a

Table 2 Some traits of livestock for which marker tests are available.

Species	Traits	Reference
Pig	Porcine stress syndrome Porcine reproductive and respiratory syndrome Pork quality	http://www.biogeneticservices.com
Sheep	Scrapie resistance Spider lamb syndrome Johne's disease	http://www.biogeneticservices.com
Cattle (dairy)	Coat color Deficiency of blood coagulation factor XI Complex vertebral malformation Bovine leukocyte adhesion deficiency	www.genomnz.co.nz
Cattle (beef)	Carcass composition components –Tenderness –Fat thickness –Marbling	http://us.igenity.com

frequency below 5% or 1%.[4] In addition, MAS relies on a limited number of markers and is therefore unable to account for a large proportion of the genetic variation in a trait.[8]

GENOMIC SELECTION

As MAS utilizes SNPs that are significant in GWAS,[8] not all QTLs affecting a trait are identified. With thousands of SNPs across the genomes of livestock species, genomic selection (GS) uses all these markers simultaneously in predicting the genomic estimated breeding value (GEBV) for traits of animals. Animals are then selected based on the rank of their GEBV. Methods of estimating GEBV generally fall into two categories: best linear unbiased prediction (BLUP) and Bayesian methods. BLUP assumes a constant genetic variance for all genotyped SNPs, and effects of SNP are sampled from a normal distribution.[9] For the Bayesian method, prior knowledge about the distribution of SNP effects is assumed and that is many SNPs are likely to have small individual effects and only a few will have large effects.[10] Two measures of the effectiveness of GS methods in estimating the BV of animals are accuracy, which is the correlation between GEBV and true BV, and reliability, which is the square of the accuracy.

STRATEGIES TO INTEGRATE WHOLE GENOME SELECTION IN FUTURE LIVESTOCK BREEDING PROGRAMS

The implementation of GS involves the derivation of an equation that predicts BV based on phenotypic and genotypic information on individuals in a reference population. This predictive equation is then used to calculate the GEBV of individuals with genotypic information in a candidate population. Dairy cattle are amenable to GS as they have a long generation interval (average age when parents give birth to offspring), and a good data recording scheme is in place.[11] In traditional dairy breeding programs, the selection of young bulls is based on EBV with an approximate accuracy of 75% estimated from progeny test results. GS can be integrated into such a breeding program by first selecting the young bulls based on their GEBVs before the progeny test or completely abdicating the progeny test with the advantage of reduced generation interval[12] and reduced cost for the breeding program. Bull dams may also be initially selected based on phenotype and pedigree, and a final selection of the number to produce young bulls is made based on the ranking of their GEBV. The limited data recording scheme in small ruminants such as sheep and goat, the conservative nature of the poultry industry, and the shorter generation interval of pigs make these species lag behind dairy cattle in GS implementation.[11]

CONCLUSION

The success of MAS in livestock populations is limited by the reliance of MAS on a limited number of markers and the inability of GWAS to detect very rare marker alleles. GS uses genome-wide markers that simultaneously predict genomic breeding values of appreciable accuracy for use in animal breeding programs. Dairy cattle have a long generation interval, and dairy breeding programs have a long history of good record keeping and are therefore favorably disposed to the implementation of GS. The success of GS will depend on the willingness of livestock keepers to embrace this new technology, the accuracy of GEBV, and efforts to remove barriers that make the implementation of GS in other livestock species apart from dairy cattle less feasible.

REFERENCES

1. Butler, J.M.; Coble, M.D.; Vallone, P.M. STRs vs SNPs: Thoughts on the future of forensic DNA testing. Forensic Sci. Med. Pathol. **2007**, *3*, 200–205.
2. Hauser, L.; Seeb, J.E. Advances in molecular technology and their impact on fisheries genetics. Fish Fisheries **2008**, *9,* 473–486.
3. Ramos, A.M.; Crooijmans, R.P.M.A.; Affara, N.A.; Amaral, A.J.; Archibald, A.L.; Beever, J.E.; Bendixen, C.; Churcher, C.; Clark, R.; Dehais, P.; Hansen, M.S.; Hedegaard, J.; Hu, Z-H.; Kerstens, H.H.; Law, A.S.; Megens, H.-J.; Milan, D.; Nonneman, D.J.; Rohrer, G.A.; Rothschild, M.F.; Smith, P.L.; Schnabel, R.D.; Van Tassel, C.P.; Taylor, J.F.; Wiedmann, R.T.; Schook, L.B.; Groenen, M.A.M. Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. PLoS ONE **2009**, *4* (8), e6524.
4. Brookfield, F.Y. Q&A: Promise and pitfalls of genome-wide association studies. BMC Biol. **2010**, *8*, 41.
5. Andersson, L. Genetic dissection of phenotypic diversity in farm animals. Nat. Genet. **2001**, *2*, 130–138.
6. Zhu, M.; Zhao, S. Candidate gene identification approach: Progress and challenges. Int. J. Biol. Sci. **2007**, *3* (7), 420–427.
7. Guillaume, F.; Fritz, S.; Boichard, D.; Druet, T. Estimation of the efficiency of French marker-assisted selection program in dairy cattle. Genet. Sel. Evol. **2008**, *40*, 91–102.
8. Hayes, B.; Goddard, M. Genome-wide association and genomic selection in animal breeding. Genome **2010**, *53* (11), 876–883.
9. Meuwissen, T.H.E.; Hayes, B.J.; Goddard M.E. Prediction of total genetic values using genome-wide dense marker maps. Genetics **2001**, *157*, 1819–1829.
10. Hayes, B.J.; Goddard, M.E. The distribution of the effects of genes affecting qualitative traits in livestock. Genet. Sel. Evol. **2001**, *23*, 209–229.
11. Ibañez-Escriche, N.; Gonzalez-Recio, O. Review. Promises, pitfalls and challenges of genomic selection in breeding programs. Span. J. Agric. Res. **2011**, *9* (2), 404–413.
12. Hayes, B.J.; Bowman, P.J.; Chamberlain, A.J.; Goddard, M.E. Invited review: Genomic selection in dairy cattle: Progress and challenges. J. Dairy Sci. **2006**, *92*, 433–443.