

Bibliometric Analysis for Genomic Selection Studies in Animal Science

Hasan ÖNDER¹, Cem TIRINK^{2*}

ABSTRACT: The animal breeding studies rapidly increased over the last century. When the genomic selection tool introduced, scientists and animal breeders have a new area to work with increasing computer power and genomic tools. In this study, it was aimed to show the situation of last 25 years of the genomic selection studies. Results showed that the number of authors per article showed that the genomic selection is a collaborative work that its tasks should be shared by group of scientists. Only about 1/3 of the genomic selection studies related to animal science. Its reason may be hardness of working with animals and generation interval which is more easy for plant breeding. When the article issues examined it can be seen that most of the articles were related to dairy science because this method widely use for dairy industry especially to determine candidate sire. The keyword “genomic selection” is widely used even this is a prediction method. It is the proof that the genomic selection is generally accepted idiom. Citation values of the most cited articles also showed that this method mostly affect the dairy science.

Keywords: Genomic selection, genomic prediction, animal breeding, bibliometry, citation

¹ Hasan ÖNDER ([Orcid ID: 0000-0002-8404-8700](https://orcid.org/0000-0002-8404-8700)), Ondokuz Mayıs University, Faculty of Agriculture, Department of Animal Science, Biometry and Genetic Unit, TR55139, Samsun, Türkiye

² Cem TIRINK ([Orcid ID: 0000-0001-6902-5837](https://orcid.org/0000-0001-6902-5837)), Iğdır University, Faculty of Agriculture, Department of Animal Science, Biometry and Genetic Unit, TR76000, Iğdır, Türkiye

*Corresponding Author: Cem TIRINK, e-mail: cem.tirink@gmail.com

INTRODUCTION

In order to improve animal production and product quality, which are of great importance in human nutrition, it is necessary to improve environmental conditions and improve genetic structures for farm animals (Önder and Abacı, 2015; Abacı and Önder, 2020). In general, the aim of animal breeding is to genetically improve the population by changing the allele frequency of the relevant yield trait so that farm animals produce more efficiently under expected future production conditions. Genetic improvement for economic traits can be achieved by selecting the best individuals of the current generation and using them as parents of the next generation (Önder et al., 2015; Olfaz et al., 2019). Traditional genetic improvement in farm animals has been very successful in predicting breeding values using phenotypes and pedigree information. Unlike traditional animal breeding methods, it is able to estimate breeding values more accurately using information on variation in DNA sequence among animals (Goddard and Hayes, 2007; Wan et al., 2016). The frontier of the genomic studies began with marker-assisted selection (MAS) which appeared in 1980's for dairy cattle. It was assumed that the main benefit of marker-assisted selection would be extensive use of young bulls based on both pedigree and marker information. Several breeding programs based on this strategy were applied to dairy cattle beginning in the first decade of 21th century (Weller et al., 2017).

Animal breeding studies getting hard because decreasing the genetic variation by selection pressure and genomic selection methods can be used as alternative way to sustainability of animal breeding (Seyedsharifi et al., 2018). The revolution in animal breeding has been created by Genomic selection (GS). Unlike traditional selection methods, GS estimates the genetic value of individuals with all genotypes, even if the phenotype has not been determined, based on estimates of marker effects (Ding et al., 2013). With genomic selection (GS) theory, instead of using a limited number of marker belonging to individual animals, the breeding value obtained by using the information belonging to all the markers in the genome has been named as genomic estimated breeding value (GEBV) using the genomic prediction (GP) of additive genetic merits of selection candidates (Taylor et al., 2016; Abacı and Önder, 2020). Different statistical models for GP were first described by Meuwissen et al. (2001), who used simulation to evaluate the performance of linear mixed models and Bayesian mixture models for estimation of marker effects and GEBVs (Taylor et al., 2016). With the introduction of high-density (including >50,000 markers) SNP chips in 2008, genomic selection begun to take place. This technology was successfully implemented in the many countries. The adoption of breeding programs based on genomic selection in major milk producing countries has led to significant changes in the dairy industry worldwide (Weller et al., 2017).

In this study, the literature related to genomic selection since its started to be worked was examined to understand the evaluation and spreading of genomic selection.

MATERIALS AND METHODS

In this study, studies related to genomic selection between the years 1995-2021 were taken into account. In this context, the "genomic selection" expression was used for searching on the Web of Science (WoS) database. In addition, the WoS categories were refined for the "Agriculture dairy animal science". The bibliographic information under the heading "Agriculture dairy animal science" of 1293 of 4002 studies on genomic selection from 1995 to 2021 was used as material.

The bibliometric analysis method is a computer-assisted scientific analysis method that determines all published researches, authors and the relationships between them (Han et al., 2020). In addition, bibliometric analysis can afford an all-inclusive network of visualizations and associative information about the chosen topic to understand the overall picture. Initially, the bibliometric analysis method was

used when examining the most effective publications primarily by author or citation information. Recently, it has adopted using bibliometric analysis, sociometric analysis and network analysis methods based on data such as keywords, titles and abstracts. The main step to perform bibliometric analysis is the data collection process. In this context, the process started to search for the issue with the essential keywords in WoS.

In this study, the bibliometric analysis for genomic selection term in animal science studies was performed with R software (R Core Team, 2020). For this aim, the bibliometrix package were used (Aria and Cuccurullo, 2017). The bibliographic data were obtained from the WoS system in Plain text format. Further, the data was changed as the data frame by using “convert2pdf” function. The biblioAnalysis function was used for performing the bibliometric analysis.

RESULTS AND DISCUSSION

The researchers published a total of 1293 studies from 77 sources such as journals, books, etc., about "genomic selection in animal science" by 2973 authors. In the 1293 authors, only 61 studies has been single authored documents of afromentioned topic. The genomic selection in animal science examined within the scope of authors collaboration that there were 2.3 authors per document. When the collaboration index examined, there were 2.41. The annual percent growth rate for scientific production is approximately 7.134388. The graphic of the number of publications in terms of yearly scientific output is given in Figure 1. According to Figure 1, while the number of genomic selection studies in animal science was 1 in 1996, how much this subject has been used over the years can be seen.

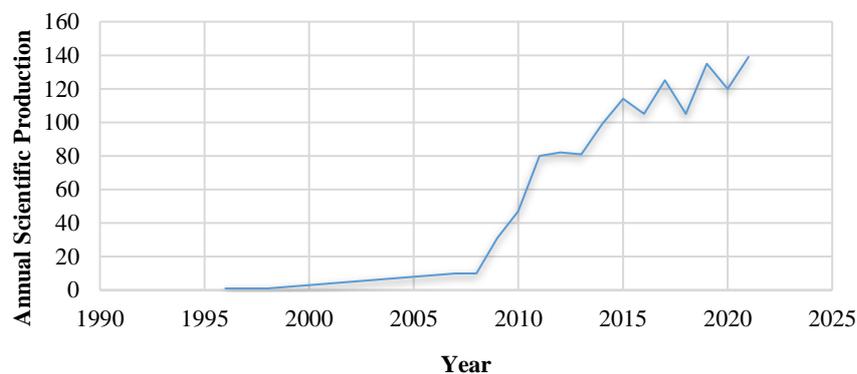


Figure 1. Annual scientific production about genomic selection in animal science

In addition, general information on bibliographic data on genomic selection in animal science is given in Table 1. According to Table 1, journals, books, etc. A total of 1293 studies have been published in some sources.

Table 1. The primary information of the data

Information	Number
Documents	1293
Sources (Journals, Books, etc)	77
Average years from publication	6.03
Average citations per document	23.75
Average citations per year per document	2.628
References	24661
Authors of single-authored documents	61
Documents per Author	0.435
Authors per Document	2.30
Co-Authors per Documents	5.30
Collaboration Index	2.41

A total of 4001 studies were utilized about the genomic selection. However, 1293 studies were used about the genomic selection in the animal science (Table 2). According to Table 2, the most common form of publication related to genomic selection is the article. In addition, there are twenty-four book chapters, two early access studies, forty-one proceeding papers, and 80 reviews about genomic selection in animal science.

Table 2. Document types for genomic selection

Document Types	Number
Article	1045
Book chapter	24
Early access	2
Proceeding paper (Article)	36
Correction	2
Editorial material	15
Meeting abstract	84
Review	80
Proceeding paper	5

Table 3 provides information about which journals the published articles are published in about the genomic selection in animal science. In this context, the top 10 journals of the list were shared. According to Table 3, the researchers published 259 articles in the Journal of Dairy Science journal as the first chosen journal. The second journal was Genetics Selection Evolution with the number of 201 articles. The least selected journal was INRA Productions Animales with thirty-one articles.

Table 3. The most published articles in journals

Sources	Number of Articles
Journal of Dairy Science	259
Genetics Selection Evolution	201
Journal of Animal Science	189
Journal of Animal Breeding and Genetics	96
Animal	68
Livestock Science	56
Animal Genetics	48
Animals	42
Animal Production Science	33
INRA Productions Animales	31

Figure 2 shows that the most preferred keywords in publications by the authors. According to Figure 2, genomic selection expression was used as the 10 most preferred keywords. In addition, expressions such as dairy cattle, genomic prediction, accuracy, beef cattle, SNP have been used extensively.

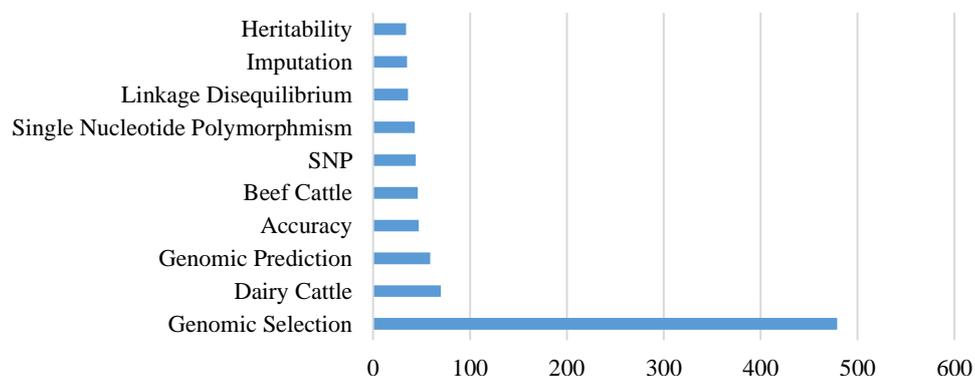


Figure 2. The most chosen keywords for genomic selection

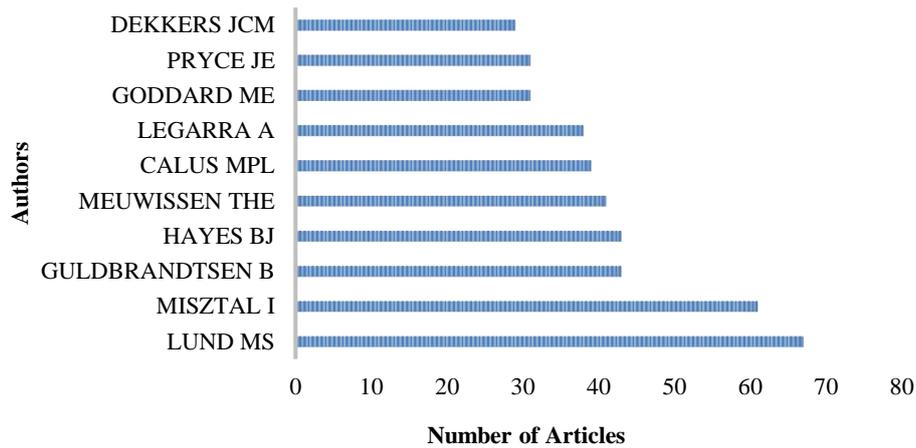


Figure 3. Productive authors for genomic selection

Figure 3 shows the most leading authors for the genomic selection researches. According to Figure 3, the most leading author was Lund MS and Misztal I, with 67 and 61 articles about genomic selection, respectively. In addition, Dekkers JCM is the author with the least number of publications in the top 10 today, with 29 articles.

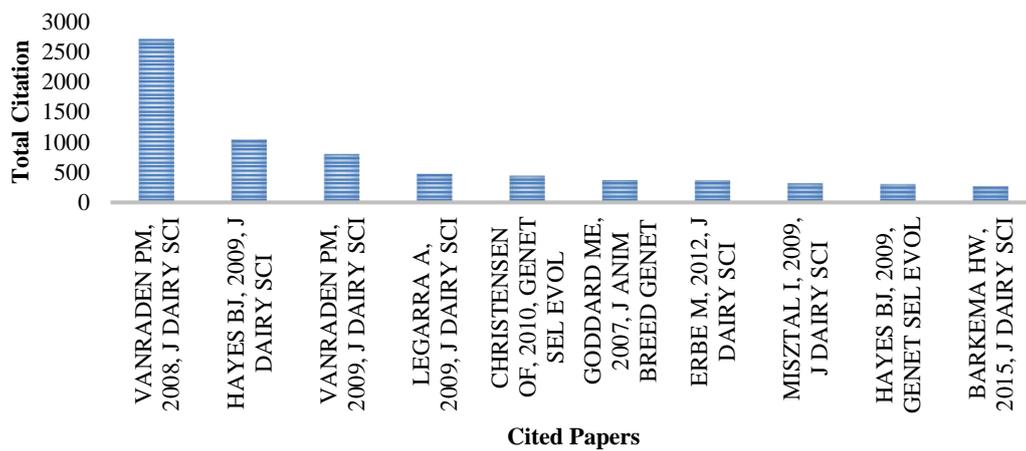


Figure 4. The most cited article

Figure 4 provides information on the most cited articles on genomic selection. As a result of the bibliometric analysis made in this context, it was determined that the most citations were made to the article written by Vanraden PM. Published in the journal J Dairy Sci in 2008, this article receives 180.3 citations per year. In this context, the second most cited article is the article published by Hayes BJ in 2009 in the journal J Dairy Sci. This article receives an average of 74.8 citations per year. When the articles cited in the top 10 were examined, it was determined that the most cited journal was J Dairy Sci with 7 articles.

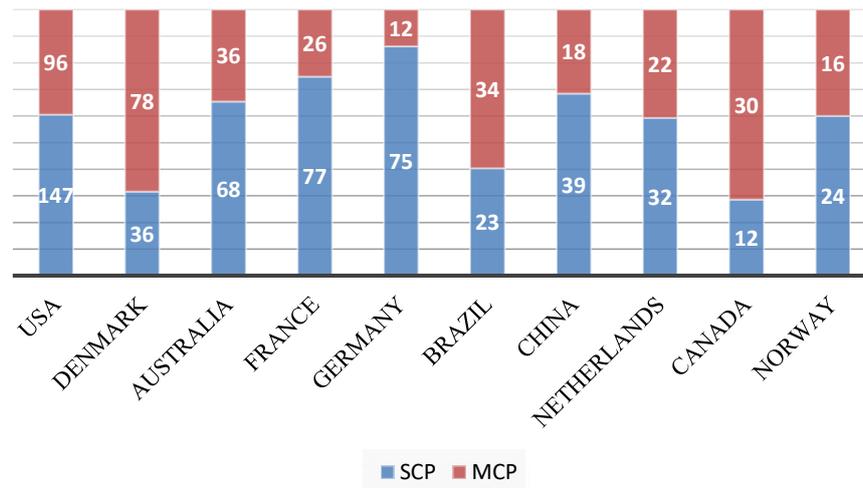


Figure 5. Collaborative information about genomic selection studies

The collaborative evaluation of genomic selection by country is given in Figure 5. Figure 5 shows that the USA is the most productive country by number of single country publication (SCP) and multiple country publication (MCP). Germany is the country that follows the USA in terms of SCP, while Denmark follows the USA in terms of MCP.

Co-Citation

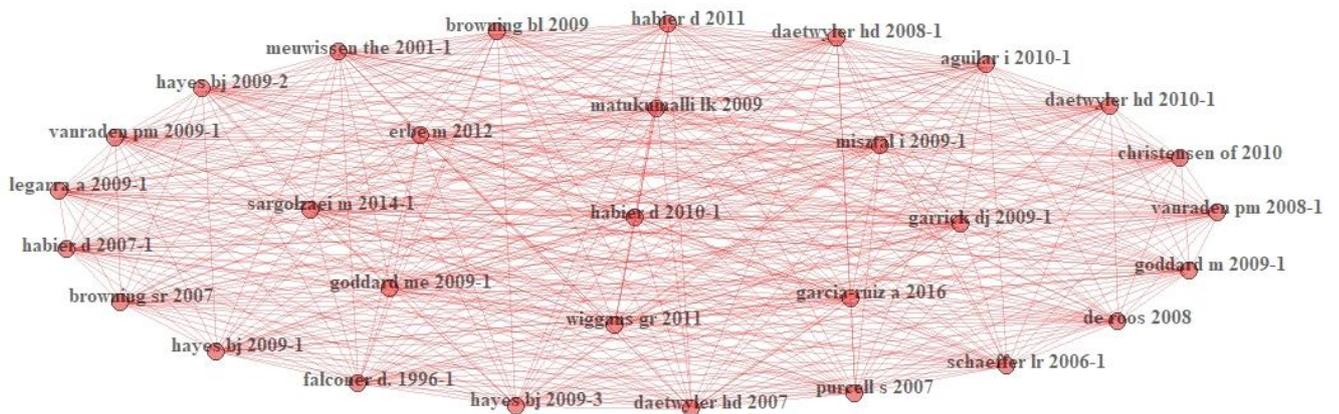


Figure 6. Co-citation report about genomic selection

Figure 6 shows the common citation network of studies on genomic selection. Figure 7 shows the conceptual structure of the keywords used in studies in the field of genomic selection. In the Figure, 3 clusters were formed according to the keywords used by the authors.

The bibliographic information under the heading "Science of Agricultural Dairy Livestock" of 1293 of 4002 studies on genomic selection conducted between 1995-2021 was used as material. In this context, the subject is a subject that has not lost its originality. In genomic selection studies, the USA stands out as the country with the highest publication speed.

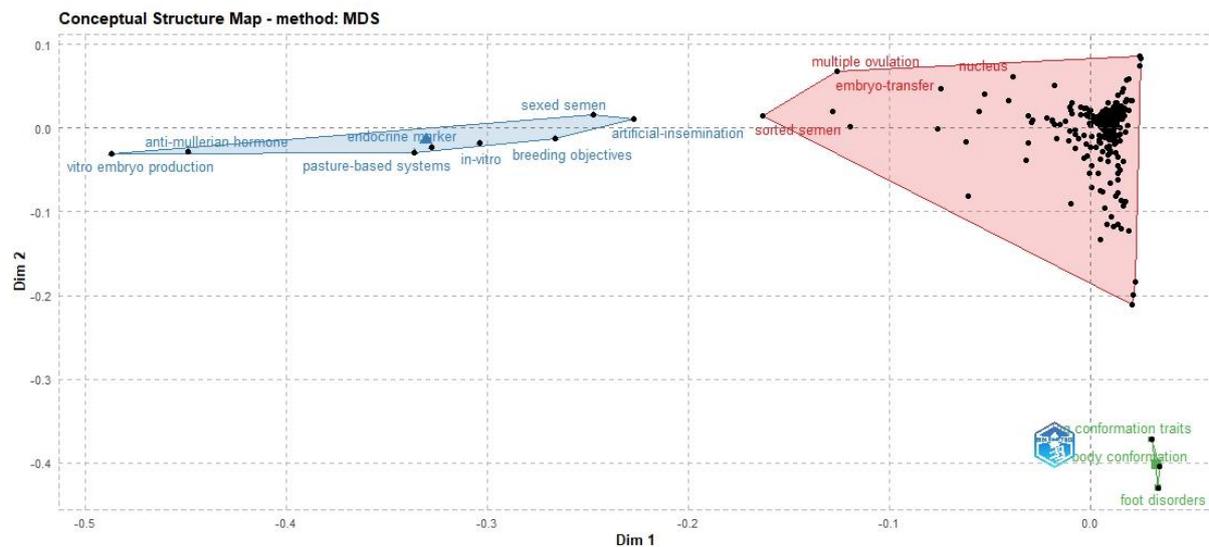


Figure 7. The conceptual structure for genomic selection

CONCLUSION

The number of authors per article showed that the genomic selection is a collaborative work that its tasks should be shared by group of scientists. Only about 1/3 of the genomic selection studies related to animal science. Its reason may be hardness of working with animals and generation interval which is more easy for plant breeding. When the article issues examined it can be seen that most of the articles were related to dairy science because this method widely use for dairy industry especially to determine candidate sire. The keyword “genomic selection” is widely used even this is a prediction method. It is the proof that the genomic selection is generally accepted idiom. Citation values of the most cited articles also showed that this method mostly affect the dairy science. Results showed that the Denmark and Canada have higher multiple country publications where the Germany leader to produce single country publication. This situation may be the result of educational relationships.

Information gathered from this bibliometric work showed that genomic selection studies will be increased rapidly in the future as a long time. Also SNP densities will grow up (still the scientific world reach about 270k of 3000k SNP of cattle) and with increasing computer power and help of new algorithms more detailed and reliable works will be conducted for genomic selection.

Conflict of Interest

The article authors declare that there is no conflict of interest between them.

Author's Contributions

CT and HO collected and analyzed data and wrote the manuscript. HO and CT revised the manuscript. All authors read and approved the final manuscript.

REFERENCES

- Abacı SH, Önder H, 2020. Comparison of BLUP and Bayesian methods for different sizes of training population in genomic selection. *Turkish Journal of Veterinary & Animal Sciences*, 44(5): 994-1002. doi:10.3906/vet-2001-52.
- Aria M, Cuccurullo C, 2017. bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of informetrics*, 11(4): 959-975.

- Ding X, Zhang Z, Li X, Wang S, Wu X, Sun D, Yu Y, Liu J, Wang Y, Zhang Y, Zhang S, Zhang Y, Zhang Q, 2013. Accuracy of genomic prediction for milk production traits in the Chinese Holstein population using a reference population consisting of cows. *Journal of Dairy Science*, 96(8): 5315-5323. doi:10.3168/jds.2012-6194.
- Goddard ME, Hayes BJ, 2007. Genomic selection. *Journal of Animal breeding and Genetics*, 124(6): 323-330.
- Han J, Kang HJ, Kim M, Kwon GH, 2020. Mapping the intellectual structure of research on surgery with mixed reality: Bibliometric network analysis (2000–2019). *Journal of Biomedical Informatics*, 109: 103516.
- Meuwissen THE, Hayes BJ, Goddard ME, 2001 Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157(4): 1819-1829.
- Olfaz M, Tırınk C, Önder, H, 2019. Use of CART and CHAID algorithms in Karayaka sheep breeding. *Kafkas Üniversitesi Veteriner Fakültesi Dergisi*, 25 (1): 105-110. doi: 10.9775/kvfd.2018.20388.
- Önder H, Abacı SH, 2015. Path Analysis for Body Measurements on Body Weight of Saanen Kids. *Kafkas Üniversitesi Veteriner Fakültesi Dergisi*, 21(3): 351-354. DOI: 10.9775/kvfd.2014.12500.
- Önder H, Şen U, Takma Ç, Ocak S, Abacı SH, 2015. Genetic Parameter Estimates for Growth Traits in Saanen Kids. *Kafkas Üniversitesi Veteriner Fakültesi Dergisi*, 21(6): 799-804. DOI: 10.9775/kvfd.2015.13407.
- R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Seyedsharifi R, Azizyan S, Boustan A, Seifdavati J, Mojtahedin A, 2018. Investigating the Bias Resulted from Ignoring Bulmer Effect on the Genetic and Economic Output in Progeny Test and Genomic Selection Program. *Kafkas Üniversitesi Veteriner Fakültesi Dergisi*, 24(1): 137-144. DOI: 10.9775/kvfd.2017.18618.
- Taylor JF, Taylor KH, Decker JE, 2016. Holsteins are the genomic selection poster cows. *PNAS*, 113: 28. doi: 10.1073/pnas.1608144113.
- Wan C, Huang Y, Shi S, Fu G, Cheng L, Chen H, Fu Q, Chen C, Hu K, 2016. Complete Genome Sequence of Goose Parvovirus Isolated from *Anser cygnoides* in China. *Kafkas Üniversitesi Veteriner Fakültesi Dergisi*, 22(1): 155-158, DOI: 10.9775/kvfd.2015.14242.
- Weller JI, Ezra E, Ron M, 2017. Invited review: A perspective on the future of genomic selection in dairy cattle. *Journal of dairy science*, 100(11): 8633-8644. doi: 10.3168/jds.2017-12879.