Black Sea Journal of Agriculture
doi: 10.47115/bsagriculture.1103853

Open Access Journal
e-ISSN: 2618 – 6578

Research Article
Volume 5 - Issue 3: 234-239 / July 2022

BIBLIOMETRIC ANALYSIS FOR GENOME-WIDE ASSOCIATION STUDIES IN ANIMAL SCIENCE

Cem TIRINK[1]*

[1]*Iğdır University, Faculty of Agriculture, Department of Animal Science, 76000, Iğdır, Türkiye*

Abstract: The main idea of the study is to determine the trends in recent years in the field of animal science, by examining 379 studies with the term "genome-wide association studies" in the title of the article published within the scope of SCI-Expanded between 2007 and 2021, within the scope of bibliometric analysis. In this context, the term of "Genome-Wide Association Studies" was searched in the Web of Science database in the study titles and the bibliometric data of the studies were accessed in plaintext format. The bibliometric results show that GWAS within animal science is developing steadily as a field of scientific research and is currently a highly topical issue. GWAS has been one of the most popular research areas due to its application in many different fields such as cell biology, plant sciences, zoology, animal science, etc. In the light of this information, it can be listed as an important contribution that GWAS studies with bibliometric analysis are still up-to-date and that the studies to be done will increase their contribution to animal science.

Keywords: Genome-wide association studies, GWAS, Bibliometric analysis, Animal science

***Corresponding author:** Iğdır University, Faculty of Agriculture, Department of Animal Science, 76000, Iğdır, Türkiye
E mail: cem.tirink@gmail.com (C. TIRINK)
Cem TIRINK https://orcid.org/0000-0001-6902-5837

Received: April 15, 2022
Accepted: May 05, 2022
Published: July 01, 2022
Cite as: Tırınk C. 2022. Bibliometric analysis for genome-wide association studies in animal science. BSJ Agri, 5(3): 234-239.

1. Introduction

Breeding studies aiming to increase genetic capacity and improve environmental factors have come to the fore. Animal breeding is expressed as an effort to increase the relative proportion of economically used animals in the population in the next generations (Wellmann, 2019). Classical animal breeding studies, which have been continuing since the 18th century, continue today. In classical breeding methods, genotypic and phenotypic parameters that will help select individuals that will form future generations can be estimated by using the records of the yield characteristics of animals (Ertugrul et al., 2002). To achieve this goal, there are many methods for estimating the genotypic value. The advancing technology in biology and molecular genetics has allowed the identification of genes and polymorphisms responsible for developing functional traits that are important in breeding studies in farm animals (Smołucha et al., 2020). Modern breeding programs have been received by estimating the genetic values of selection candidates based on phenotypic and pedigree information and then making selection decisions on this information.

In this era, breeding value is achieved by utilizing the knowledge of all the markers in the genome on the contrary using the limited number of marker information of animals. This breeding value has been named genomic breeding value (Genomic Breeding Value: GEBV). The selection made according to the GEBV of animals is called "genomic selection" (Meuwissen et al., 2001). In genomic selection, parameter estimations obtained by using a training population with genetic marker and phenotypic values are also used to estimate the breeding values of individuals in the test population for which only the marker knowledge is available. It was possible to identify genes held by characteristics of economic value to microsatellite markers primarily by QTL mapping in the 90s (Lipkin et al., 1998). Today, however, with the onset of all genome sequencing technologies and the availability of affordable panels of all genomes single nucleotide polymorphisms (SNPs), SNPs are used for planning besides phenotype and pedigree information. To date, many QTLs have been described in several separate research papers (Sharma et al., 2015). However, it is impossible to determine gene effects in a genomic selection where genetic parameters are estimated. The Genome-Wide Association Studies (GWAS) method comes to the fore to overcome this situation.

The first GWAS study published by Klein et al. (2005) discovered a genetic variation relating to a higher risk of age-related macular degeneration. After that, the National Human Genome Research Institute (NHGRI) announced an electronic record of GWAS in 2008 (Mansiaux and Carrat, 2012). GWAS in farm animals have made popularity in mapping QTL to economically essential features such as calving ease, fat and protein content, meat quality and quantity, milk yield, egg production, fertility characteristics, etc. (Sharma et al., 2015). GWAS examines genotyped SNPs in the genome and their relationship to phenotype (Zeng et al., 2015;

Khanzadeh et al., 2020). GWAS has been conducted for many species in livestock (Schook et al., 2005; Zimin et al., 2009; Gu et al., 2011). A GWAS yields by testing each SNP, in order, to associate with the yield trait under consideration. The critical hypothesis in GWAS is that important associations appear because the SNP is in LD with a causative mutation affecting its feature and thus close to it (Hayes and Goddard, 2010).

Since its discovery, GWAS has been one of the most popular research areas due to its application in many different fields such as cell biology, plant sciences, animal science, etc. According to the Web of Science data, there are 20208 publications related to GWAS, and also if we give some examples related to livestock, the first studies were in 2007. Abasht and Lamont (2007) studied the broiler population. Long et al. (2007) also studied GWAS for broilers. In addition, Charlier et al. (2008) reported that they examined fine-scale mapping of the disorder of five recessive genes in cattle using SNP associations. There are hundreds of studies on GWAS, such as these examples. However, as far as we know, no article in the livestock science has presented a general bibliometric perspective of the GWAS.

In this context, bibliometric analysis is a research area that attracts more and more attention in the scientific society and is determined by the rapid improvement of computers and internet (Bar-Ilan, 2008; Merigó and Yang, 2017; Celik, 2021). Bibliometric analysis is a basic approach used to analyze investigation and takes its foundation from the public library and information science.

This paper aims to afford a general indication of GWAS research using bibliometric methods from 2007 to 2021. The Web of Science (WoS) database was used to collect information for this aim. The objective is to identify the most beneficial and effective research in GWAS and make sure the current progression of the area by considering the most prominent papers and authors. Most of the results follow mutual advice. However, we get several specific cases that show in what manner the field of GWAS is increasing, with some topics being very widespread and highly cited. In contrast, some topics do not cite enough citations.

2. Material and Methods

We made a topical query in terms of "Genome-Wide Association Studies" to refine the animal science field using the WoS database. The bibliographic information of 379 animal science studies out of 20208 studies with GWAS was used as a material from 2007 to 2021.

Bibliometric analysis is a scientific computer-assisted analysis method that identifies research, authors, and their relations by covering all the publications published on a particular subject (Han et al., 2020). However, bibliometric analysis can provide comprehensive visualization and relational information on the chosen topic to understand the overall picture. The first bibliometric analysis examines the most influential publications, primarily according to author or citation information. In recent bibliometric analysis, sociometric analysis and network analysis based on keywords, titles and abstract data have been adopted.

Bibliometric analysis is a process from beginning the data collection to visualization of the analysis results. The data collection process begins the search of the issue with the essential keywords in Web of Science (WoS).

The term social network is a structure that shows the interaction, collaboration, and effects between people in a social context (Celik, 2021). Social network analysis examines the social structure that is the subject of research and its effects (Tindall and Wellman, 2001). The primary purpose of social network analysis is to define and visualize the social network structure, model it statistically, and generate information from the network (Celik, 2021).

This study performed a bibliometric analysis for GWAS, widely used in animal science studies in recent years. All statistical evaluations were made using R software with the package of “bibliometrix” (R Core Team, 2020; Aria and Cuccurullo, 2017). The data were obtained bibliographically from the WoS system in Plain text format. After that, the bibliographic data was converted to the data frame by using “convert2pdf” function with the package of “bibliometrix”. The bibliometric analysis was performed by the biblioAnalysis function. In this context, this article aims to perform bibliometric, collaboration, and co-citation analysis to determine the importance of GWAS in animal science over the years.

3. Results

The researchers published 379 studies from 38 sources such as journals, books, etc., about "GWAS" by 1741 authors. The annual percent growth rate for scientific production is approximately 7.598. 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 GWAS studies in animal science was 2 in 2007, how much this subject has been used over the years can be seen.

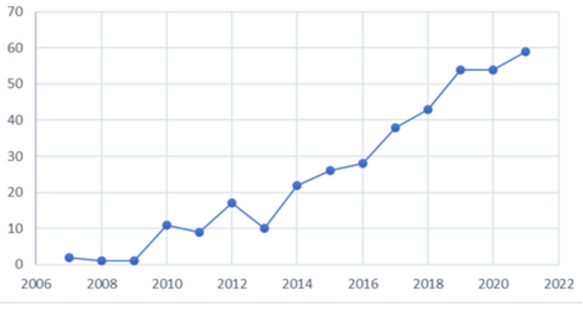

Figure 1. Number of articles per year.

In addition, the preliminary information about the bibliographic data is provided in Table 1. According to Table 1, a total of 379 studies were published in some sources such as journals, books, etc. Only 13 of these studies were done with a single-authored document.

Table 1. The primary information of the data

Information	Number
Documents	379
Sources (Journals, Books, etc)	38
Average years from publication	4.68
Average citations per document	15.09
Average citations per year per document	2.839
References	12702
Authors of single-authored documents	13
Documents per Author	0.221
Authors per Document	4.52
Co-Authors per Documents	6.91
Collaboration Index	4.65

A total of 20208 studies were utilized about the GWAS. However, 379 studies were used about the GWAS in the animal science field (Table 2). According to Table 2, most of these studies consist of articles. In addition, there are two book chapters, three early access studies, four papers, and 29 reviews about GWAS.

Table 2. Document Types for GWAS

Document Types	Number
Article	329
Book chapter	2
Early access	2
Proceeding paper	2
Meeting abstract	13
Review	28
Book chapter (Review)	1

Table 3 provides information on which journals have published the GWAS article. According to Table 3, the researchers published 76 articles in the Genetics Selection Evolution journal as the first chosen journal. The second journal was Animal Genetics journal with the number of 58 articles. The least selected journal was Animal Science Journal with eight articles.

Table 3. The most published articles in journals

Sources	Number
Genetics Selection Evolution	76
Animal Genetics	58
Journal of Dairy Science	52
Journal of Animal Science	34
Animals	29
Livestock Science	19
Animal	16
Journal of Animal Breeding and Genetics	14
Asian-Australasian J of Animal Sciences	12
Animal Science Journal	8

Figure 2 shows the most used keywords in publications. According to Figure 2, among the keywords used in studies about GWAS, expressions such as SNP, candidate genes, dairy cattle, genomic selection appear as the most frequently used keywords, apart from the name of GWAS. Figure 3 shows the most influential authors through the GWAS studies. According to Figure 3, the most effective author was Hayes BJ, with 12 documents about GWAS. Regarding the number of articles, the lowest number of publications so far, with 8, was made by Brito LF, Dekkers JCM, Lund MS and Pryce JE.

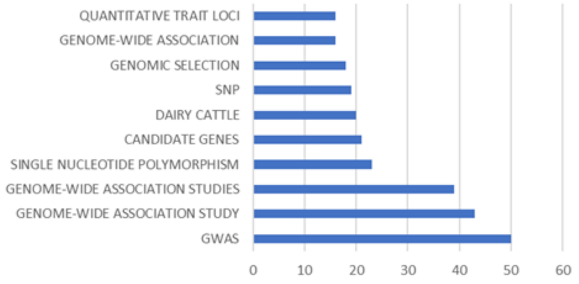

Figure 2. The most chosen keywords for GWAS.

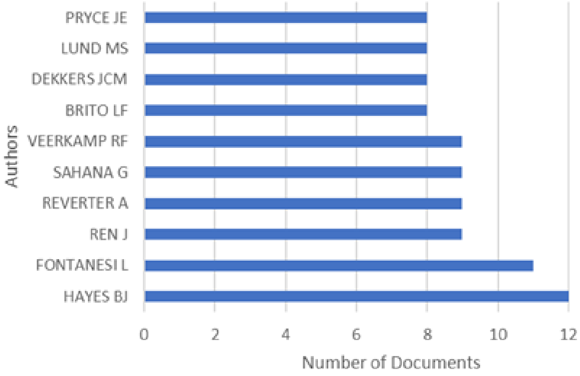

Figure 3. The most productive authors in GWAS generated by bibliometrix package.

Figure 4 provides information about the most cited articles. In this context, it was determined that the article written by Yang ZW received the most citations. This article was published in the journal Anim Biotechnol in 2021. The first article Yang ZW has 311.50 citations per year. In this context, the second widely cited article is the article published by Pryce JE in 2010 in the journal J Dairy Sci. This article has 8.69 citations per year.

According to countries, the most collaborative information about GWAS studies is given in Figure 5. Figure 5 shows that the China and USA is the most productive country according to the number of single (SCP) and multiple (MCP) country publications. However, China was the most productive country in the single country publications.

Figure 6 shows the co-citation network for GWAS studies. Figure 7 shows the conceptual structure of the keywords in the research area for GWAS. There were 3 cluster for the authors' keywords. The cluster showing the greatest similarity of the three clusters consists of terms such as meat quality, milk production, fertility and reproduction, etc.

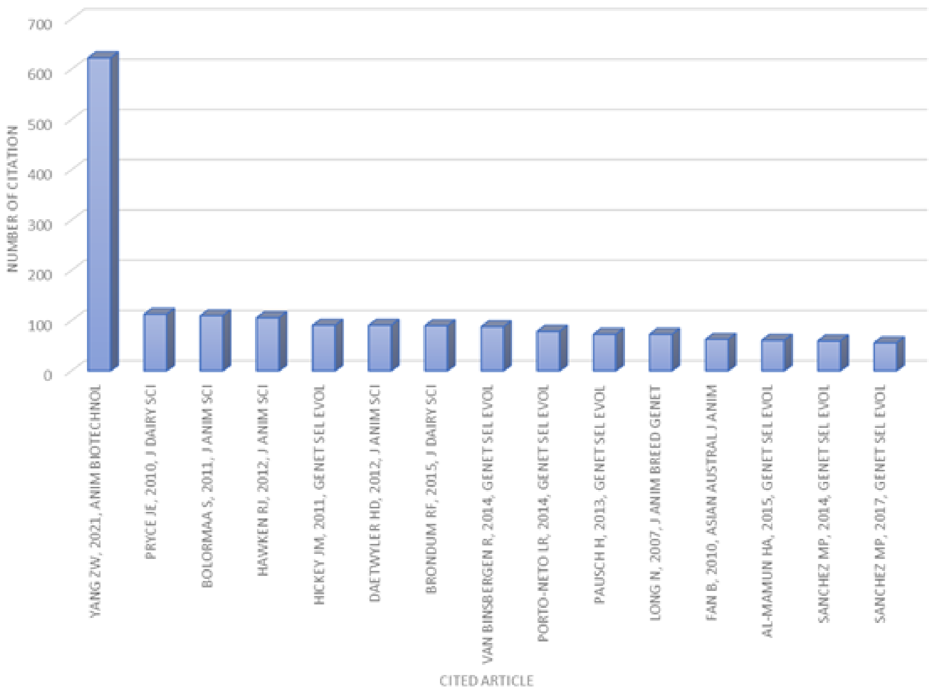

Figure 4. The most cited article.

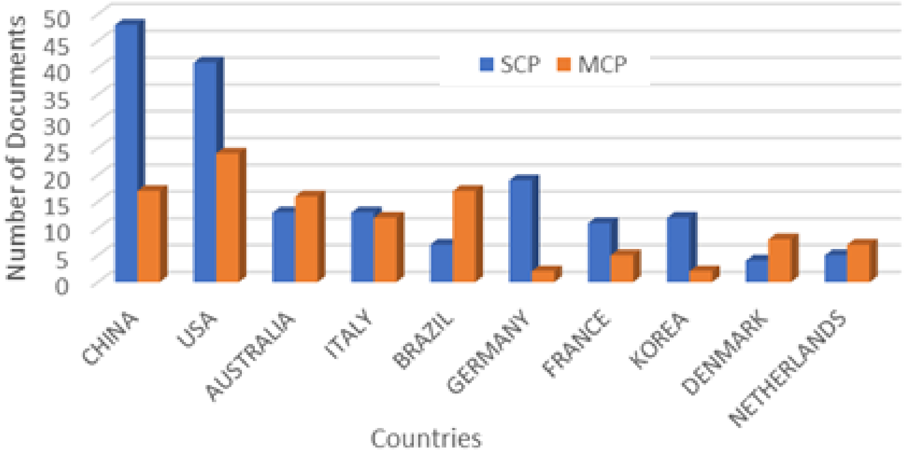

Figure 5. Collaborative information about GWAS studies.

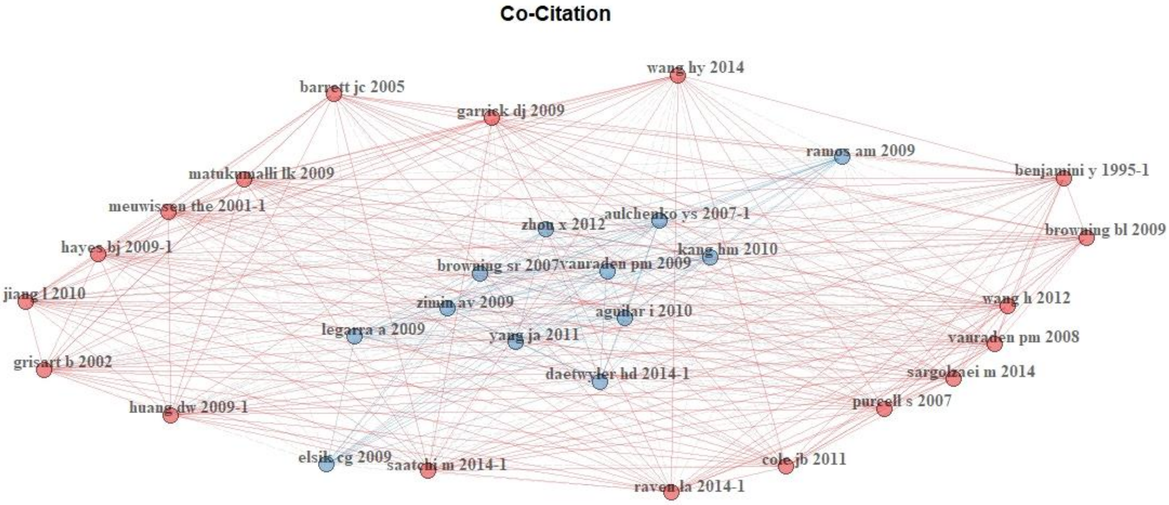

Figure 6. Co-citation report about GWAS.

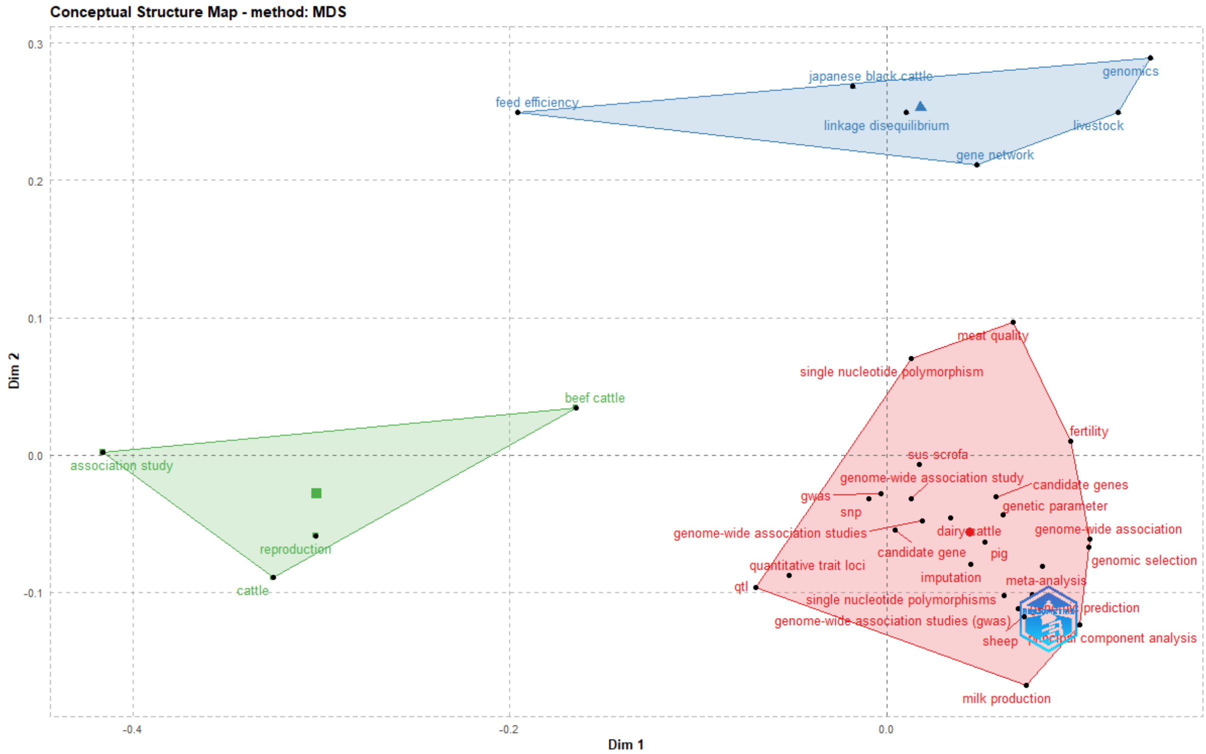

Figure 7. The conceptual structure for GWAS.

Considering the distribution of 379 studies published between 2007 and 2021 on GWAS applications in animal science, it is seen that the most studies were done in 2021. In this context, the issue has not lost its importance and is a current issue. USA and China stand out as the countries with the highest broadcasting rate.

GWAS has been one of the most popular research areas due to its application in many different fields such as cell biology, plant sciences, zoology, animal science, etc. Especially in animal science, it is seen that there are many studies on subjects such as meat quality, milk yield, gene expressions and reproduction, etc. Considering the distribution of 379 studies published between 2007 and 2021 on GWAS applications in animal science, it is seen that the most studies were done in 2021. In this context, the issue still has not lost its importance and stands out as a current issue.

In this context, as a result of the bibliometric analysis of GWAS in animal science, Genetics Selection Evolution has the status of the journal with the most publications on this subject. In addition, when the number of citations was examined, it was determined that the most effective author was Yang ZW. It was determined that the most cited countries were China, Australia and the USA.

In the light of this information, it can be listed as an important contribution that GWAS studies with bibliometric analysis are still up-to-date and that the studies to be done will increase their contribution to animal science.

Author Contributions

All task made by C.T. (100%) data acquisition and analysis, writing up, submission and revision. The author reviewed and approved final version of the manuscript.

Conflict of Interest

The author declared that there is no conflict of interest.

Ethical Consideration

Ethics committee approval was not required for this study because of there was no study on animals or humans.

References

Abasht B, Lamont SJ. 2007. Genome-wide association analysis reveals cryptic alleles as an important factor in heterosis for fatness in chicken F2 population. Anim Genet, 38(5): 491-498.

Aria M, Cuccurullo C. 2017. bibliometrix: An R-tool for comprehensive science mapping analysis. J Informet, 11(4): 959-975.

Bar-Ilan J. 2008. Informetrics at the beginning of the 21st century-A review. J Informetr, 2(1): 1-52.

Celik S. 2021. Social network analysis and a model practice. J Orig Stu, 2(1): 29-41.

Charlier C, Coppieters W, Rollin F, Desmecht D, Agerholm JS, Cambisano N, Carta E, Dardano S, Dive M, Fasquelle, C, Frennet JC. 2008. Highly effective SNP-based association mapping and management of recessive defects in livestock. Nat Genet, 40(4): 449-54.

Ertugrul O, Orman MN, Guneren G. 2002. Some genetic parameters of milk production in the Holstein breed. Turk J Vet Anim Sci, 26: 463-469.

Gu X, Feng C, Ma L, Song C, Wang Y, Da Y, Li H, Chen K, Ye S, Ge

C, Hu X, Li N. 2011. Genome-wide association study of body weight in chicken F2 resource population. PLoS ONE, 6(7): e21872.

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). J Biomed Inform, 109: 103516.

Hayes B, Goddard M. 2010. Genome-wide association and genomic selection in animal breeding. Genome, 53(11): 876-883.

Inci H, Celik S, Sogut B, Sengul T, Karakaya E. 2015. Examining the Effects of Different Feather Color on the Characteristics of Interior and Exterior Egg Quality of Japanese quail by Using Kruskal-Wallis Tests. Turk J Agric Nat Sci, 2 (1): 112–118.

Khanzadeh H, Ghavi Hossein-Zadeh N, Ghovvati S. 2020. Genome wide association studies, next generation sequencing and their application in animal breeding and genetics: a review. Iran J Appl Anim Sci, 10(3): 395-404.

Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, Henning AK, San Giovanni JP, Mane SM, Mayne ST, Bracken MB, Ferris FL, Ott J, Barnstable C, Hoh J. 2005. Complement factor H polymorphism in age-related macular degeneration. Science, 308(5720): 385-389.

Lipkin E, Mosig MO, Darvasi A, Ezra E, Shalom A, Friedmann A, Soller M. 1988. Quantitative trait locus mapping in dairy cattle by means of selective milk DNA pooling using dinucleotide microsatellite markers: Analysis of milk protein percentage. Genetics, 149: 1557-1567.

Long N, Gianola D, Rosa G, Weigel K. 2007. Machine learning classification procedure for selecting SNPs in genomic selection: application to early mortality in broilers. J Anim Breed Genet, 124(6): 377-389.

Mansiaux Y, Carrat F. 2012. Contribution of Genome-Wide Association Studies to Scientific Research: A Bibliometric Survey of the Citation Impacts of GWAS and Candidate Gene Studies Published during the Same Period and in the Same Journals. PLoS ONE, 7(12): e51408.

Merigó JM, Yang JB. 2017. A bibliometric analysis of operations research and management science. Omega, 73: 37-48.

Meuwissen THE, Hayes BJ, Goddard ME. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics, 157(4): 1819–1829.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Schook LB, Beever JE, Rogers J, Humphray S, Archibald A, Chardon P, Milan D, Rohrer G, Eversole K. 2005. Swine Genome Sequencing Consortium (SGSC): A strategic roadmap for sequencing the pig genome. Comp Funct Genomics, 6: 251-255.

Sharma A, Lee JS, Dang CG, Sudrajad P, Kim HC, Yeon SH, Kang HS, Lee SH. 2015. Stories and challenges of genome wide association studies in livestock-a review. Asian-Australas J Anim Sci, 28(10): 1371.

Smołucha G, Gurgul A, Jasielczuk I, Kawęcka A, Miksza-Cybulska A. 2021. A genome-wide association study for prolificacy in three Polish sheep breeds. J Appl Genet, 62(2): 323-326.

Tindall DB, Wellman B. 2001. Canada as social structure: Social network analysis and Canadian sociology. Can J Sociol, 26(3): 265-308.

Wellmann R. 2019. Optimum contribution selection for animal breeding and conservation: the R package optiSel. BMC Bioinformat, 20: 25.

Zeng P, Zhao Y, Qian C, Zhang L, Zhang R, Gou J, Liu J, Liu L, Chen F. 2015. Statistical analysis for genome-wide association study. J Biomed Res, 29(4): 285-292.

Zimin AV, Delcher AL, Florea L, Kelley DR, Schatz MC, Puiu D, Hanrahan F, Pertea G, Van Tassell CP, Sonstegard TS, Marçais G, Roberts M, Subramanian P, Yorke JA, Salzberg SL. 2009. A whole-genome assembly of the domestic cow, Bos taurus. Genom Biol, 10: R42.