Bibliometric analysis of scientific research on cocoa (*Theobroma cacao* L.) in Mexico

Castillo-Méndez, R. C.[1]; Aguirre-Andrade, A.[1]; Ulloa-Pimienta, A. R.[2]; Leue-Luna, M. C.[1]; Sánchez-Trinidad, R. C.[1*]

[1] Universidad Popular de la Chontalpa, Carretera Cárdenas-Huimanguillo, km 2.0, R/a Pazo y Playa, Cárdenas, Tabasco.
[2] Tecnológico Nacional de México / Instituto Tecnológico Superior de Villa la Venta.
* Correspondence: rosa.sanchez@upch.mx

ABSTRACT

Objective: to analyze the scientific contributions made in Mexico on Cocoa tree cultivation during the period 1999-2022 and to quantify those contributions by gender.

Design/Methodology/Approach: scientific articles were collected from major publishers (Elsevier, MDPI, Wiley and Springer), database of articles from Open Access Journals (CONRICYT, Scielo, Redalyc, Latindex, Claryvate Analytics, Periodica and DOAJ), and the open access web search engine Google Scholar. The Gephi software was used to build networks of the institutions, the researchers' affiliation and co-authorship networks.

Results: scientific production showed an exponential growth of scientific texts at the national scale where Cocoa tree was the subject of research. In the last 22 years, scientific productivity was concentrated in three main topics: biotechnology (18.51%), economy (15.26%), and diseases, (13.31%). The research was focused on the Mexican southeast, and was led by Colegio de Postgraduados (COLPOS), the National Institute of Forestry, Agricultural and Livestock Research (INIFAP), and the Juarez Autonomous University of Tabasco. The male gender presented a higher percentage of articles published as first author and as author for correspondence.

Limitations/Implications of the study: the documents analyzed were exclusively scientific articles.

Findings/Conclusions: There is a research gap in species propagation techniques, ethnobotany, irrigation, plant physiology, and the influence of indigenous cultures and groups on the transfer of knowledge.

Keywords: *Theobroma cacao*, bibliometric analysis, co-authorship networks, gender.

Citation: Castillo-Méndez, R.C., Aguirre-Andrade, A., Ulloa-Pimienta, A.R., Leue-Luna, M.C., & Sánchez-Trinidad, R.C. (2024). Bibliometric analysis of scientific research on cocoa (*Theobroma cacao* L.) in Mexico. *Agro Productividad*. https://doi.org/10.32854/agrop.v17i10.2879

Academic Editor: Jorge Cadena Iñiguez
Associate Editor: Dra. Lucero del Mar Ruiz Posadas
Guest Editor: Daniel Alejandro Cadena Zamudio

Received: April 18, 2024.
Accepted: September 05, 2024.
Published on-line: November XX, 2024.

Agro Productividad, 17(10). October. 2024. pp: 83-90.

INTRODUCTION

Cocoa tree (*Theobroma cacao* L.) originated in the Neotropics of South America, in the Amazonian region. Mexican cocoa tree has developed across cultural groups, in different geographical environments where it has thrived and evolved with intense gene flows among populations (Lanaud *et al.*, 2024).

In Mexico, cocoa tree is economically important because it sustains the income of many families. On a social scale, it represents a source of employment, and in a cultural aspect, the annual consumption of chocolate is very important (0.5 kg of processed cocoa nut per capita). Mexico's domestic production satisfies 18% (28 thousand tons, Megagrams –Mg) of the domestic demand for cocoa, so about 120 thousand Mg are needed to satisfy it (CEDRSSA, 2020).

Cocoa yields in the South-Southeast region, where more than 97% cocoa is produced, are very low (0.49 Mg ha^{-1}) (SIAP, 2024). Cocoa is subject to many threats in Mexico, including its susceptibility to disease and climate change (Lahive *et al.*, 2019).

Bibliometric studies allow retrospective analysis of how the scientific advances generated specifically in this crop have been achieved and made known. The results of this type of study can be used for the development of scientific policy and its management, to the extent that scientific knowledge is perceived as a strategic value for the generation of useful results (Sanz-Valero *et al.*, 2014). For this reason, the objective of this research was to analyze the scientific contributions accomplished in Mexico on cocoa tree cultivation during the period 1999-2022, and to quantify those contributions by gender.

MATERIALS AND METHODS

Origin of Data

In this study, we considered the set of scientific articles where cocoa was the object of research and carried out by Mexican researchers or with an area of study in Mexico. Keywords used in the search for the scientific articles were Cacao and its English equivalent, cocoa. The scientific articles were collected from the main publishers (Elsevier, MDPI, Wiley and Springer); databases of articles from Open Access Journals (CONRICYT, Scielo, Redalyc, Latindex, Claryvate Analytics, Periodica and DOAJ), and the open access web search engine Google Scholar.

The scientific articles were collected from January to June 2023, and the number of texts available until December 2022 was considered. Through content analysis, the studies that were not developed in Mexico or by Mexican researchers were discarded. In addition, the "snowball" technique was used to expand the search for reports, based on the list of references of articles found in the initial search (Leipold, 2014).

Bibliometric indicators

Variables analyzed for each of the scientific articles collected were authors, year of publication, title, journal, and bibliographic citations. Then, through content analysis, the authors' affiliation institutions, the country of origin of the first author, and the topic of each research were identified: 1) Biotechnology (included studies associated with food and beverage production processes); 2) Economics (studies where added value is given to direct and indirect cocoa products); 3) Diseases (pathogens associated with cocoa production); 4) Pests (parasites associated with cocoa production); 5) Fertilization (compounds added to the cocoa tree to improve yields in the field); 6) Agroforestry (association of cocoa with other crops); 7) Plant breeding (studies related to propagation techniques of the cocoa tree); 8) Ecology (research on the influence of cocoa on the environment where it grows); and 9) Botany (studies where the cocoa species is described).

Data was captured in a spreadsheet and the original language of each of the texts was kept. During the capture of data, records that presented variants, but with the same significance, were standardized. In addition, special characters were changed such as: š (for s) ç (for c), ã (for a), or eliminated, accents, superscripts, subscripts, ® or © among others, in order to facilitate the analysis.

Data Analysis

The methodology described by Santillán-Fernández *et al.* (2023) was followed. A graph of temporality of scientific production was created with the variables year of publication and number of citations. For the variable frequency of scientific articles per year, an ordinary least squares regression model was estimated to determine the trend in the frequency of publications. A graph was also generated with the main research topics, with the intention of determining potential areas for the development of new research about cocoa.

Bibliometric indicators were generated for journals that published scientific articles in which cocoa was studied; more frequently, and of greater relevance; which is measured by the number of times they appear in bibliographic citations. Through the Gephi program (Bastian *et al.*, 2009), networks of the institutions of affiliation of the researchers were identified. This in order to determine those institutions that have generated most of the knowledge about cocoa, and the co-authorship networks which identify the main researchers. Once the articles were selected, the number of researchers was reclassified by gender. In those articles where the full name of the author and co-authors does not appear, an email was sent to the author for correspondence, asking for the full names of the authors.

RESULTS AND DISCUSSION

During the period analyzed (1999-2022), 199 scientific articles have been published on cocoa cultivation. A linear increase in the number of articles published was observed, which denotes the importance of cocoa as a local crop to contribute to food sovereignty and to the generation of self-employment of families living in rural areas where cocoa is grown. During the period 2015-2020, more articles were published per year; the increase was due to the greater number of research carried out on the control of cocoa moniliasis, a disease that causes necrotic damage to the internal tissue of the fruit, causes its fall and therefore decreases the yield of cocoa almonds (Hipólito-Romero *et al.*, 2020).

Most of the research in Mexico has focused on topics such as how to generate cocoa-based food and beverages (biotechnology 18.51%); adding value (economics 15.26%); and disease control (13.31%) (Figure 2). But, there are research gaps in agronomic management such as fertilization, irrigation, shade levels, physiology, uses of GRAS substances (substances generally recognized as safe), that could improve cocoa almond yields, as well as the understanding of the potential effects of climate change on this crop.

Mexican researchers are used to publish their findings in Mexican journals (7 out of 10), which limits the scope of research at the international scale. These journals are mostly published in Spanish without impact factor- JIF (Table 1).

The journal Agro Productividad ranks first with 28 (14.07%) articles and REMEXCA the second, with 16 (8.04%). The preference for these magazines may be due to a few reasons. Agro Productividad in its beginnings was classified as a scientific communication journal by CONACYT (today CONAHCYT). But now it is a multidisciplinary scientific journal, which allows the publication of a large number of diverse topics. In addition, it has color illustrations and publishes images of field and lab experiments. This features make this journal more attractive.

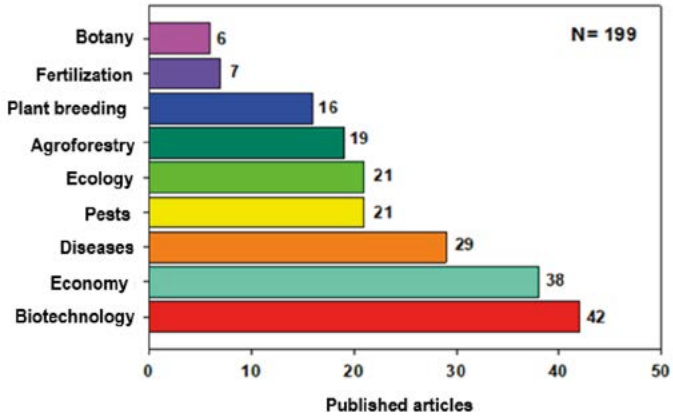

Figure 1. Temporal evolution of scientific production and bibliographic citations on the species *Theobroma cacao* in Mexico.

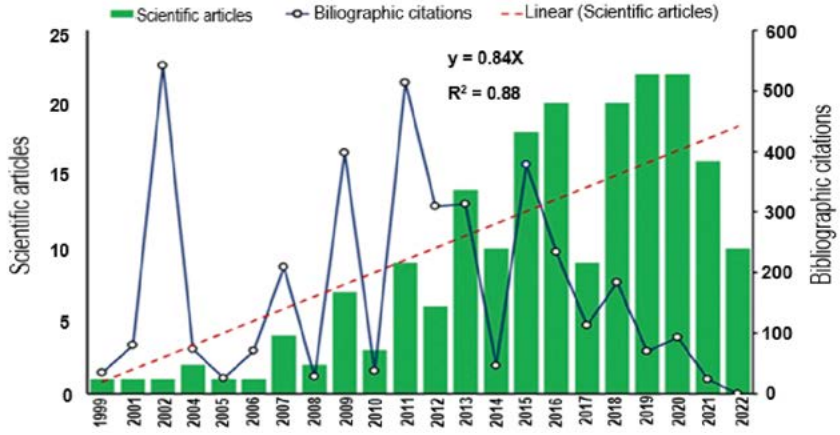

Figure 2. Main research topics where cocoa (*Theobroma cacao*) was the object of research in Mexico from 1999 to 2022.

Table 1. Bibliometric indicators of the main journals that published scientific articles where cocoa (*Theobroma cacao*) was the object of research in Mexico from 1999 to 2022, ranked according to the number of published articles.

Journal	Country	Publisher	Language	Topic	Articles		Citation	
					Number	%	Number	%
AgroProductividad	México	ColPos	Spanish	Agriculture	28	14.07	126	3.34
REMEXCA	México	INIFAP	Spanish	Agriculture	16	8.04	204	5.40
RFM	México	SM_Fitogenetica A. C.	Spanish/English	Agriculture	6	3.02	74	1.96
TSA	México	UADY	Spanish	Agroforestry	6	3.02	38	1.01
Universidad y Ciencia	México	UJAT	Spanish	Biotechnology	6	3.02	296	7.84
Botanical Sciences	México	SBM A.C.	English	Botany	4	2.01	5	0.13
Agroforest Syst	USA	Springer	English	Agroforestry	3	1.51	20	0.53
Interciencia	Venezuela	Asociación Interciencia	Spanish/English	Biotechnology	3	1.51	165	4.37
RMB	México	UNAM	Spanish/English	Ecology	3	1.51	91	2.41
Tecnología en Marcha	Costa Rica	ETCR	Spanish	Biotechnology	3	1.51	30	0.79
Otras (106)					121	60.80	2728	72.23
Total (116)					199	100.00	3777	100.00

In contrast, REMEXCA is published by the Institute of Forestry, Agricultural and Livestock Research-INIFAP, which is the institution where much of the cocoa research on various topics is carried out in Mexico. In fact, INIFAP is a national leader in the genetic improvement of *Theobroma cacao*.

Tabasco is the state where the most studies on cocoa are developed. This is because it is the state with the highest production and planted area of Cocoa tree. It should be noted that the most cited studies were published mostly in high-impact journals, under world-class publishers which publish only in English. Interestingly, although there is little research on the influence of pre-Hispanic cultures and current ethnicities on cacao (Ethnobotany), they turn out to be the reports with the highest number of citations. In addition, those journals show the tendency of indigenous groups of people to consume local resources and their influence on food sovereignty, which are interesting topics to be developed.

The analysis of networks of research institutions revealed that in the 199 scientific articles published, there was co-participation of different research institutions on 114 instances (nodes) and 204 links are apparent (Figure 3).

Table 2. Bibliometric indicators of the main scientific articles where cocoa (*Theobroma cacao*) was the object of research in Mexico from 1999 to 2022, ranked according to the number of bibliographic citations obtained.

First author			Scientific article						
Name	Institución	Country	Área of stady	Journal	lLanguage	Publisher	WoS (2021)	Subject	Citations
Motamayor *et al.* (2002)	CIRAD	France	Tabasco	Heredity	English	Springer	3.8	Ethnobotany	542
Rodriguez-Campos *et al.* (2012)	IPN	Mexico	Tabasco	Food Chemistry	English	Elsevier	8.8	Biotechnology	262
Rodriguez-Campos *et al.* (2011)	IPN	Mexico	Tabasco	Food_RI	English	Elsevier	8.1	Biotechnology	259
Crown y Hurst (2009)	UN_Méx	USA	Sureste Mexico	PNAS	English	PNAS	11.1	Ethnobotany	195
Powis *et al.* (2011)	KSU	USA	Sureste Mexico	PNAS	English	PNAS	11.1	Ethnobotany	123
Salgado-Mora *et al.* (2007)	ECOSUR_ Chiapas	Mexico	Chiapas	Interciencia	Spanish/ English	Interciencia	0.4	Ecology	112
Cordova_Avalos (2001)	ColPos_ Puebla	Mexico	Tabasco	Universidad y Ciencia	Spanish	UJAT	Sin factor	Agroforestry	80
Muñoz *et al.* (2006)	ECOSUR_ Chiapas	Mexico	Tabasco	AJ_ Primatology	English	Wiley	2.4	Ecology	71
Romero-Cortes *et al.* (2013)	TecNM_ Veracruz	Mexico	Tabasco	J Sci Food Agric	English	Wiley	4.1	Biotechnology	59
Priego-Castillo *et al.* (2009)	ColPos_ Tabasco	Mexico	Tabasco	Universidad y Ciencia	Spanish	UJAT	Sin factor	Agroforestry	56

CIRAD: Centre de Coopération Internationale en Recherche Agronomique pour le Développement; IPN: Instituto Politécnico Nacional; UN_ Méx: Univerity if new Mexico; KSU: Kennesaw State University; ECOSUR: El Colegio de la Frontera Sur; ColPos: Colegio de Postgraduados; TecNM: Tecnológico Nacional de México; PNAS: Proceedings of the National Academy of Sciences; Food_RI: Food Research International; AJ_Primatology: American Journal of Primatology; Interciencia: Asociación Interciencia; UJAT: Universidad Juárez Autónoma de Tabasco.

Figure 3 shows that six institutions lead cocoa research in Mexico; all of them located in the producing areas within the Mexican southeast. Institutions located in the state of Tabasco tend to a greater degree of association with each other. On the other hand, the institutions located in the state of Chiapas have broader collaboration networks with other research centers in the country.

The institutions that develop research in cocoa cultivation are located in the cocoa production areas. This helps to improve technology transfer, because the information was generated in the cocoa area, in field conditions with cooperating producers. In other words, the spatial validation of knowledge is not an impediment to the transfer.

In the 199 scientific articles analyzed, 485 different authors (nodes) and 661 links were found (Figure 4). A network is composed of nodes interconnected by edges, where the

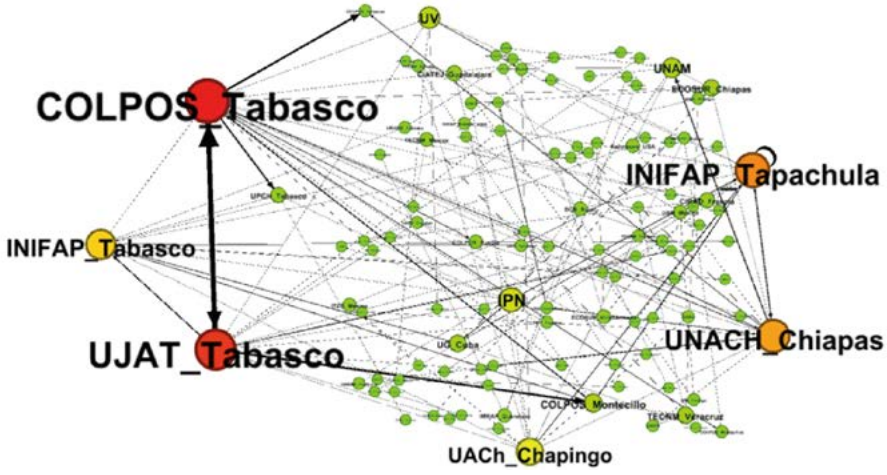

Figure 3. Network of research institutions that published scientific articles on cacao (*Theobroma cacao*) in Mexico from 1999 to 2022. The size of the node corresponds to their productivity.

Figure 4. Network of authors who published scientific articles on cacao (*Theobroma cacao*) in Mexico from 1999 to 2022. Node size corresponds to their productivity.

nodes represent the individuals and the edges represent the links that unite them (Aguilar-Gallegos *et al.*, 2016). Links in an analysis of co-authorship networks are important because, through them, an author can access ideas, knowledge, and sources of information that are socially distant (Granovetter, 1973).

Main authors who developed research where cocoa was the object of study from 1999 to 2022 were: Avendano_Azarrate_CH of INIFAP-Tapachula at the Rosario Izapa experimental station; Torres de la Cruz_M de la UJAT_Tabasco, and Cordova_Avalos_V of CP_Tabasco. However, unlike CP_Tabasco, UJAT_Tabasco, and INIFAP_Tapachula, where one researcher is apparently full specialist in cocoa issues, in UNACH_Chiapas a network of researchers was presented, composed mainly of Lopez_Baez_O, Suarez_Venero_GM, Vázquez_Ovando_A, and López_Hernández_JG who are linked to other specialists in different institutions in the country.

Silva *et al.* (2014) found that authors from the same institution tend to associate with each other. This restricts constructive criticism and reduces feedback on the relevance of research. In addition, institutional research groups tend to replicate the same methodologies in different study sites, which limits innovation in research and allows circularity (redundancy in the object of study) of publications (Santillán-Fernández *et al.*, 2023). Therefore, synergies with authors from other institutions can be a good strategy to improve the quantity and quality of research.

The total number of researchers who conducted research in the last 22 years is 485 people. Of this total, 332 (68.45%) are males (men) and 153 (31.55%) are females (women) (Figure 5A). Figure 5B shows that the males published 147 (73.86%) articles as first author, and females 52 (26.13%). The same trend is observed when analyzing the number of articles published as author for correspondence (Figure 5C). Men published 151 (75.87%) and women, 48 (24.12%).

Our findings showed a gender inequality in the results of research conducted on cocoa in Mexico. Thus, this study provides empirical evidence on that imbalance in scientific

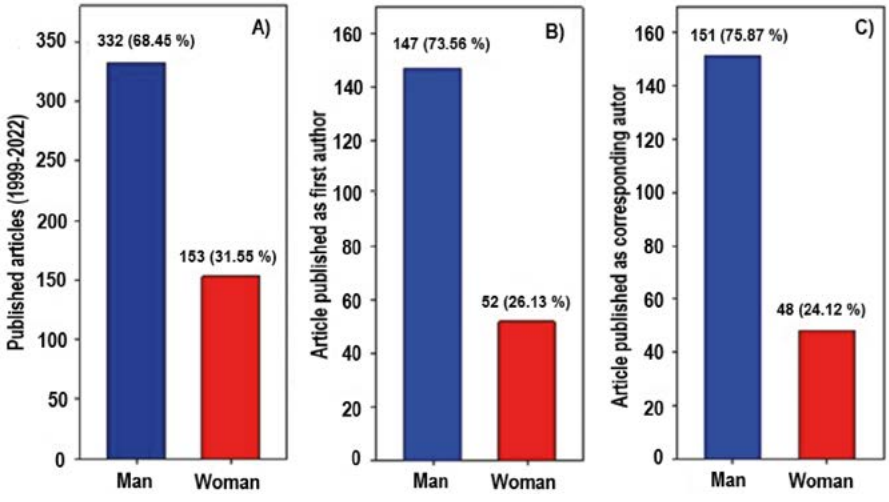

Figure 5. Number of articles on cocoa (*Theobroma cacao*) published by A) gender (only female or male); B) as first author and C) as author for correspondence.

production. This imbalance is in line with global results in which women represent less than 30% of shared authorships, while men represent over 70% (Larivière *et al.*, 2013).

CONCLUSIONS

Scientific production showed an exponential growth of papers at the national scale about the cultivation of cocoa (*Theobroma cacao*) in the recent 22 years past. The greatest scientific productivity was concentrated in topics such as biotechnology, economics and disease control. Research is focused on the Mexican southeast, and was led by Colegio de Postgraduados, the National Institute of Agricultural and Livestock Forestry Research-INIFAP, and the Juárez Autonomous University of Tabasco-UJAT. The male gender is more represented as publisher of the most articles. The highest percentage of publications as first author and corresponding author are also represented by males.

REFERENCES

Aguilar-Gallegos, N., Martínez-González, E. G., Aguilar-Ávila, J., Santoyo-Cortés, H., Muñoz-Rodríguez, M., & García-Sánchez, E. I. (2016). Análisis de redes sociales para catalizar la innovación agrícola: de los vínculos directos a la integración y radialidad. *Estudios Gerenciales*, 32:197-207. http://dx.doi.org/10.1016/j.estger.2016.06.0060.

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. *Icwsm*, 8:361-362.

CEDRSSA. (Centro de Estudio para el Desarrollo Rural Sustentable y la Soberanía Alimentaría). 2020. Industria del cacao en México. 7 pag. http://www.cedrssa.gob.mx/files/b/9/71Industria%20del%20cacao%20en%20M%C3%A9xico.pdf.

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology, 78*(6):1360-1380. http://www.jstor.org/stable/2776392.

Hipólito-Romero, E., Cocoletzi-Vásquez, E., Ramos-Prado, J. M., Espinoza, C., Torres-de la Cruz, M., & Ricaño-Rodríguez, J. (2020). Breve aproximación a la naturaleza genómica de *Moniliophthora roreri* CPMRT01 aislado de cacao en Tabasco, México. *Biotecnia, 22*(2):39-49. https://doi.org/10.18633/biotecnia.v22i2.1244.

Lanaud, C., Vignes, H., Utge, J., Valette, G., Rhoné, B., Garcia Caputi, M., & Argout, X. (2024). A revisited history of cacao domestication in pre-Columbian times revealed by archaeogenomic approaches. *Scientific Reports, 14*(1):2972. https://doi.org/10.1038/s41598-024-53010-6.

Lahive, F., Hadley, P., & Daymond, A.J. (2019). The physiological responses of cacao to the environment and the implications for climate change resilience. A review. *Agronomy for Sustainable Development*, 39:1-22. https://doi.org/10.1007/s13593-018-0552-0.

Larivière, V., Ni, C., Gingras, Y., Cronin, B., & Sugimoto, C.R. (2013). Bibliometrics: Global gender disparities in science. *Nature, 504*(7479):211-213. http://dx.doi.org/10.1038/504211a.

Leipold, S. (2014). Creating forests with words-A review of forest-related discourse studies. *Forest Policy and Economics*, 40:12-20. https://doi.org/10.1016/j.forpol.2013.12.005.

PAN (Planeación Agrícola Nacional 2017-2030). (2017). Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA). Primera edición. https://www.gob.mx/cms/uploads/attachment/file/256425/B_sico-Cacao.pdf.

Santillán-Fernández, A., Vásquez-Bautista, N., Pelcastre-Ruiz, L., Ortigoza-García, C., Padilla-Herrera, E., Tadeo-Noble, A., Carrillo-Ávila, E., Juárez-López, J., Vera-López, J., & Bautista-Ortega, J. (2023). Bibliometric Analysis of Forestry Research in Mexico Published by Mexican Journals. *Forests*, 14:648. https://doi.org/10.3390/f14030648.

Sanz-Valero, J., Casterá, V.T., & Wanden-Berghe, C. (2014). Estudio bibliométrico de la producción científica publicada por la Revista Panamericana de Salud Pública/Pan American Journal of Public Health en el período de 1997 a 2012. *Rev Panam Salud Publica, 35*(2): 81-88. https://www.scielosp.org/pdf/rpsp/v35n2/a01v35n2.pdf.

SIAP (2024). Atlas Agroalimentario 2012-2023: Sistema de Información Agroalimentaria y Pesquera. Disponible en: https://nube.siap.gob.mx/index.php/s/oI2ukvy4dVX6DOl?utm_source=NUBE (acceso 15 febrero de 2024).

Silva, T., Moro, M., Silva, A., Meira, W., & Laender, A. (2014). Community-based endogamy as an influence indicator. IEEE/ACM Joint Conference on Digital Libraries, London, UK, pp. 67-76.