



Evolution of Population Genetics Research: Insights into Hardy-Weinberg Equilibrium and Mathematical Modeling

Priyambodo Priyambodo*, Agung Putra Wijaya, Nindy Permatasari, Hidayatulloh Hidayatulloh, Siti Hamidatul Aliyah, and Elly Lestari Rustiati

Received : March 6, 2026

Revised : May 9, 2026

Accepted : June 2, 2026

Online : July 30, 2026

Abstract

Population genetics has long relied on mathematical frameworks to interpret patterns of genetic variation and evolutionary processes within populations. Among these frameworks, the Hardy-Weinberg equilibrium is one of the most fundamental models for evaluating allele and genotype frequency distributions under theoretical population conditions. Despite its central role in population genetics, the development and research trends related to the Hardy-Weinberg equilibrium and mathematical modeling have not been systematically examined at the bibliometric level. This study aimed to analyze the evolution of scientific research on Hardy-Weinberg equilibrium and mathematical modeling within population genetics using bibliometric approaches. Bibliographic data were retrieved from the Scopus database, resulting in a final dataset of 1,899 research articles. Publication trends, leading journals, productive authors, citation impact, international collaboration networks, thematic structures, and historiographic evolution were analyzed using bibliometric techniques, VOSviewer, and Bibliometrix in R Studio. The results indicate a sustained growth in publications, particularly since the early 2000s, reflecting the increasing role of quantitative and computational approaches in population genetics. Citation and thematic analyses reveal a substantial intellectual transition from classical equilibrium-based population genetics toward molecular, computational, and genomics-oriented research. Historiographic periodization further demonstrates that Hardy-Weinberg-based mathematical frameworks have continuously evolved alongside advances in high-throughput sequencing, statistical genetics, and large-scale genomic analysis. An important insight from this study is that modern genomics research has not replaced the classical Hardy-Weinberg principles, but has instead expanded their application through the integration of computational biology, bioinformatics, and genomics technologies. These findings highlight the continuing relevance of mathematical population genetic frameworks in contemporary genetic science and provide a clearer understanding of the state of the art and future direction of population genetics research.

Keywords: computational genetics, genomics, Hardy-Weinberg equilibrium, historiographic evolution, mathematical modeling, population genetics

1. INTRODUCTION

Mathematical approaches have become increasingly important in modern biology because many biological phenomena involve complex interactions that require quantitative interpretation [1]. Over time, mathematical reasoning has transformed biology from a predominantly descriptive discipline into one capable of explaining biological patterns and processes through formal models. In particular, mathematical frameworks enable researchers to describe biological systems using measurable parameters, thereby improving the precision of predictions and interpretations in

scientific research [2].

One of the biological disciplines most strongly shaped by mathematical reasoning is population genetics [3]. This field seeks to understand how genetic variation is distributed within and among populations and how evolutionary processes influence genetic composition over time [4]. Because changes in populations can be represented through allele and genotype frequencies, population genetics has long relied on mathematical formulations to analyze evolutionary dynamics and population structure [5]. As a result, quantitative approaches have become central to the development of modern population genetic research.

Among the most fundamental mathematical frameworks in population genetics is the Hardy-Weinberg equilibrium, which describes the expected distribution of allele and genotype frequencies under theoretical population conditions in the absence of evolutionary forces such as mutation, migration, natural selection, and genetic drift [6]. Although relatively simple, this model provides an essential baseline for evaluating genetic variation within populations. Deviations from Hardy-Weinberg expectations frequently indicate

Publisher's Note:

Pandawa Institute stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright:

© 2026 by the author(s).

Licensee Pandawa Institute, Metro, Indonesia. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

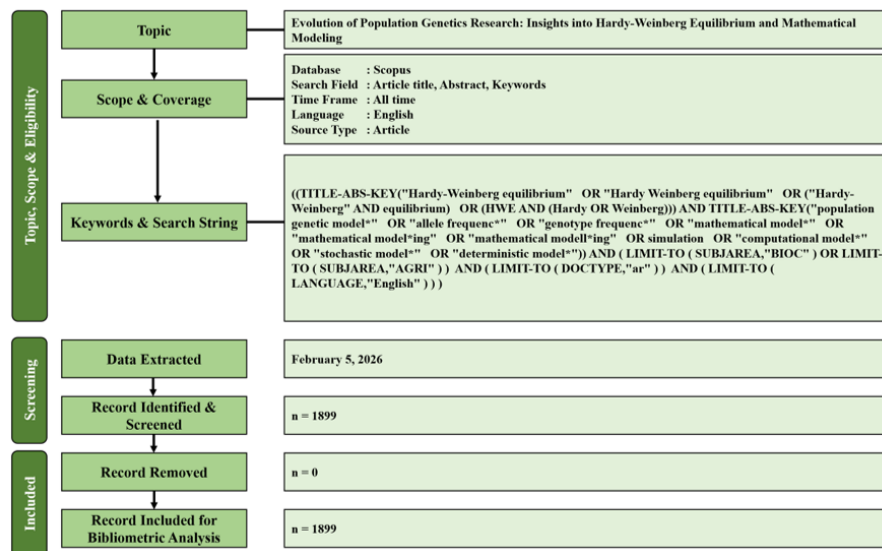


Figure 1. Literature identification and screening process.

the influence of biological and demographic processes affecting population structure [7][8]. Consequently, the Hardy-Weinberg framework has become one of the most widely used analytical foundations in population genetics.

Beyond its theoretical significance, Hardy-Weinberg equilibrium is extensively applied across diverse areas of biological research, including evolutionary biology, conservation genetics, breeding programs, forensic genetics, and population monitoring [9][10]. Previous studies have primarily focused on the application of Hardy-Weinberg principles in empirical genetic analyses, molecular marker studies, forensic identification, and population structure assessment [11][12]. In recent decades, advances in molecular biology, computational methods, and high-throughput sequencing technologies have further expanded the analytical capacity of population genetics, leading to broader applications involving microsatellites, SNP markers, genomic datasets, and computational simulations [13][14]. More recently, the rapid expansion of population genomics has transformed population genetics into a more data-intensive and computationally driven discipline [15]. Large-scale genomic datasets now require increasingly sophisticated mathematical and statistical frameworks to interpret genetic variation, demographic history, and evolutionary relationships across populations [16][17]. As a consequence, Hardy-Weinberg-based mathematical models continue to serve as important conceptual and

analytical foundations in contemporary genetic and genomic research.

Despite the extensive use of Hardy-Weinberg equilibrium across multiple biological disciplines, the overall evolution of scientific research related to this mathematical framework remains insufficiently explored from a bibliometric perspective [18][19]. Existing studies have largely concentrated on theoretical population genetic models, empirical genetic applications, or specific molecular and genomic investigations [20]-[22]. However, limited attention has been given to systematically examining how Hardy-Weinberg-related research has evolved over time, how thematic research directions have shifted across different periods, and how mathematical modeling has influenced the broader development of population genetics research [23]. In particular, the transition from classical equilibrium-based population genetics toward genomics-oriented and computational approaches has not been comprehensively mapped within the scientific literature.

Bibliometric analysis provides an effective approach for addressing this gap because it enables the identification of publication growth patterns, influential publication sources, international collaboration structures, and thematic developments within a research field [24][25]. In the context of population genetics, bibliometric analysis can provide a broader understanding of how Hardy-Weinberg-based mathematical frameworks have shaped scientific development across theoretical,

molecular, forensic, and genomics-related studies [26]. Such analysis is also important for establishing the current state of the art and for identifying major scientific transitions that continue to influence contemporary population genetics research.

Therefore, this article aims to analyze the evolution of scientific research related to Hardy-Weinberg equilibrium and mathematical modeling within population genetics using bibliometric approaches. Specifically, this study addresses several key questions, including how publication trends have evolved over time, which journals, authors, and countries have contributed most prominently to this research field, how citation impact reflects the intellectual influence of scientific publications, how thematic research clusters are structured, and how research themes have shifted from classical population genetic models toward genomics-oriented and computational approaches. This work also examines the historiographic development and periodization of research trends to establish the broader state of the art within this field. Through this approach, the study seeks to provide a comprehensive overview of how Hardy-Weinberg-based mathematical frameworks have contributed to the development of population genetics from classical theoretical models to modern genomics-driven research.

2. MATERIALS AND METHODS

2.1. Data Source and Search Strategy

The bibliographic data used in this study were obtained from the Scopus database, which is widely recognized as one of the most comprehensive sources of peer-reviewed scientific literature [27]. Scopus was selected because of its extensive coverage of journals across biological and interdisciplinary research fields, its standardized bibliographic metadata, and its compatibility with bibliometric and scientometric mapping tools such as VOSviewer and Bibliometrix [28]. Although other databases such as Web of Science and PubMed also contain relevant literature, Scopus was selected as the primary data source because of its broad interdisciplinary coverage and consistent export format suitable for large-scale bibliometric analysis.

To identify relevant publications, a structured search strategy was developed to capture studies related to population genetics, Hardy-Weinberg equilibrium, and mathematical modeling approaches. The search was conducted using the TITLE-ABS-KEY field to ensure that the retrieved records contained relevant terms within their titles, abstracts, or keywords. The search query used in the Scopus database was as follows: ((TITLE-ABS-KEY("Hardy-Weinberg equilibrium" OR "Hardy Weinberg equilibrium" OR ("Hardy-Weinberg" AND equilibrium) OR (HWE AND (Hardy OR Weinberg))) AND TITLE-ABS-KEY("population genetic model*" OR "allele frequenc*" OR "genotype frequenc*" OR "mathematical model*" OR "mathematical model*ing" OR "mathematical modell*ing" OR simulation OR "computational model*" OR "stochastic model*" OR "deterministic model*")) AND ((LIMIT-TO (SUBJAREA,"BIOC")) OR LIMIT-TO (SUBJAREA,"AGRI")) AND (LIMIT-TO (DOCTYPE,"ar")) AND (LIMIT-TO (LANGUAGE,"English"))). No restriction on publication year was applied in order to capture the full historical development of research related to Hardy-Weinberg equilibrium and mathematical modeling in population genetics. The search was conducted on February 5, 2026, and the retrieved records were exported for further analysis. Based on the records indexed in the Scopus database, the retrieved publications span the period from 1970 to 2026.

2.2. Study Selection and Screening Procedure

The selection of publications followed a systematic screening procedure adapted from the PRISMA framework to ensure transparency in the identification and selection of relevant studies [29]. The PRISMA flow diagram describing the article selection process is presented in Figure 1. Initially, all records retrieved from the Scopus database were collected and exported. Duplicate entries were removed to ensure that each publication was represented only once in the dataset. The remaining records were then screened based on their titles and abstracts to assess their relevance to population genetics research involving Hardy-Weinberg equilibrium and mathematical modeling approaches. Articles that were not directly related to

these topics were excluded during the screening stage. The remaining publications were subsequently evaluated to ensure that they met the predefined inclusion criteria. After completing the screening and eligibility assessment, the final set of publications was retained for bibliometric analysis.

In addition to the search query, several filters were applied within the Scopus database to refine the dataset. The search results were restricted to the subject areas of Biochemistry, Genetics and Molecular Biology and Agricultural and Biological Sciences, as these categories encompass most research related to population genetics. Only documents classified as research articles and written in English were included to ensure consistency in document type and language. Following database filtering, the retrieved records underwent a screening process based on titles and abstracts to remove publications not directly related to Hardy-Weinberg equilibrium or population genetic modeling. Duplicate records were also identified and removed prior to the final dataset construction.

2.3. Inclusion Criteria

Several inclusion criteria were applied to ensure the relevance and quality of the dataset. First, only documents classified as research articles were included in the analysis, as these represent original scientific contributions. Second, only publications written in English were selected to ensure consistency in text processing and interpretation.

Finally, all retrieved articles were required to contain relevant concepts associated with Hardy-Weinberg equilibrium and mathematical modeling within population genetics. After applying these criteria and completing the screening process, a total of 1,899 articles were included in the final dataset used for analysis.

2.4. Data Extraction

Bibliographic information from the selected articles was exported from Scopus in CSV format. The exported metadata included authors, article titles, publication years, source titles, author keywords, abstracts, affiliations, and citation information. These data were used to examine publication trends, collaboration networks, and thematic relationships within the research field. Before performing the bibliometric analysis, the dataset was reviewed to identify potential inconsistencies in keyword usage and to reduce redundancy caused by variations in terminology. Minor adjustments were conducted to standardize similar keywords where appropriate, ensuring more accurate network analysis and visualization. In addition, the dataset was examined to identify duplicate entries and inconsistencies in author keywords prior to network analysis. This data cleaning step helped ensure that the bibliometric mapping more accurately reflected relationships among research themes. To improve the consistency of the keyword co-occurrence analysis,

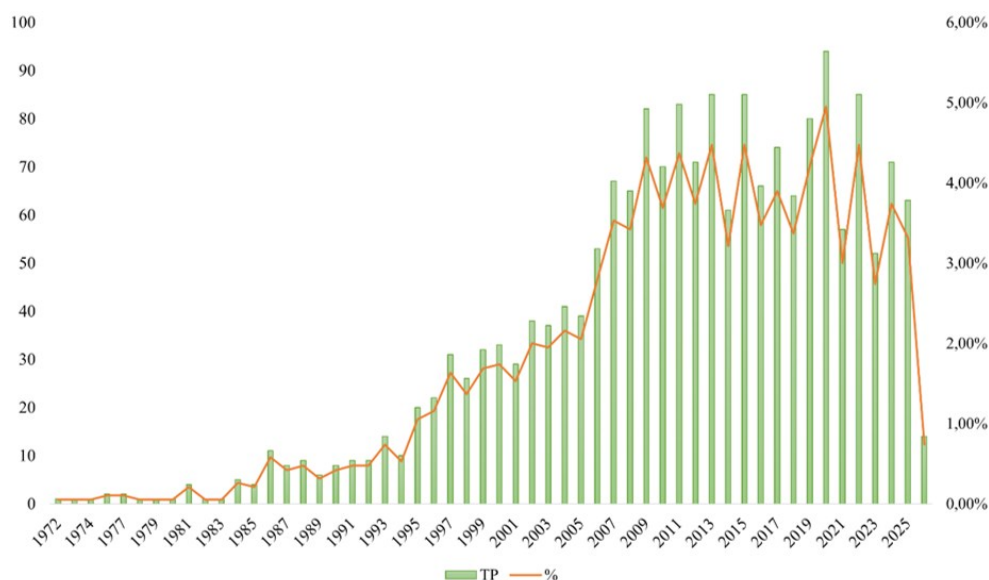


Figure 2. Annual publication trends related to Hardy-Weinberg equilibrium.

a manual standardization procedure was applied to merge synonymous or closely related keywords prior to network construction. This process involved identifying variations in terminology, differences in singular and plural forms, and alternative expressions referring to the same concept. For example, terms such as “Hardy Weinberg equilibrium” and “Hardy-Weinberg equilibrium” were unified under a single standardized keyword. Similarly, variations such as “single nucleotide polymorphism” and “SNP” were treated as equivalent concepts where appropriate. Minor spelling variations and formatting differences were also harmonized to avoid fragmentation of keywords within the bibliometric network. This normalization process ensured that conceptually identical terms were represented consistently in the keyword co-occurrence analysis.

2.5. Bibliometric Analysis

Bibliometric analysis was conducted to explore the development and structure of research related to Hardy-Weinberg equilibrium and mathematical modeling in population genetics. Several analytical approaches were employed to identify patterns and relationships within the scientific literature. First, annual publication trends were analyzed to examine the temporal growth of scientific output related to Hardy-Weinberg equilibrium and mathematical modeling in population genetics. The number of publications per year was extracted from the Scopus metadata and summarized using Microsoft Excel to visualize changes in research activity over time. Second, leading journals and the most productive authors were identified by calculating the number of publications associated with each source title and author within the dataset. These descriptive bibliometric indicators were compiled and analyzed using Microsoft Excel, allowing the identification of major publication outlets and researchers contributing most actively to this research field. Moreover, citation analysis was also conducted to evaluate the scientific impact of publications within the dataset. Citation-based indicators were used to identify highly cited articles, influential authors, and major publication sources that contributed substantially to the intellectual development of research related to Hardy-Weinberg equilibrium and mathematical modeling in population genetics.

Citation data extracted from the Scopus database were summarized and analyzed descriptively to assess patterns of scientific influence within the research field.

Third, leading countries and international collaboration networks were examined based on author affiliation data. The co-authorship relationships among countries were analyzed to explore patterns of international collaboration in population genetics research. The collaboration network visualization was generated using VOSviewer, which enables the construction of bibliometric maps based on co-authorship relationships. Fourth, keyword co-occurrence analysis was performed to identify the main thematic clusters within the literature. Author keywords extracted from the dataset were analyzed using VOSviewer, which constructs network maps based on the frequency with which keywords appear together in scientific publications. In this study, author keywords were used as the unit of analysis, and a minimum occurrence threshold of 10 was applied to identify the most relevant terms.

The keyword co-occurrence network was constructed using the association strength normalization method implemented in VOSviewer. Clusters were automatically generated using the software’s built-in clustering algorithm, which groups keywords based on the strength of their co-occurrence relationships. Default clustering resolution parameters were applied to identify coherent thematic structures within the network. The clustering procedure groups keywords based on the strength of their co-occurrence relationships, allowing conceptually related research topics to emerge as distinct thematic clusters. Finally, the thematic evolution of research topics was examined through overlay visualization of the keyword co-occurrence network. This analysis illustrates the temporal distribution of research topics by assigning colors to keywords based on their average publication year, thereby revealing the progression of research themes over time. Through these analytical procedures, the study aims to reveal the growth patterns, global research contributions, collaboration networks, thematic structures, and temporal evolution of scientific research related to Hardy-Weinberg equilibrium and mathematical modeling in population genetics.

Table 1. Leading publication sources.

Source Title	Publication	Percentage from Overall Publication (%)
Conservation Genetics Resources	69	3.63
Genetics and Molecular Research	51	2.69
Forensic Science International Genetics	47	2.47
Molecular Biology Reports	46	2.42
Gene	39	2.05
Forensic Science International Genetics Supplement Series	36	1.90
Journal of Forensic Sciences	34	1.79
Annals of Human Biology	25	1.32
Human Heredity	25	1.32
Frontiers in Genetics	23	1.21

In addition to overlay visualization, thematic evolution and historiographic analyses were conducted using the Bibliometrix package through the Biblioshiny interface in R Studio. These analyses were used to examine the temporal development of research themes and to identify major shifts in the intellectual structure of the field across different periods. To facilitate the analysis of research development trends and periodization, the publication timeline was divided into several chronological phases based on publication growth patterns and thematic transitions observed within the dataset. This approach enabled the identification of major historical phases in the evolution of Hardy-Weinberg-related research, ranging from classical theoretical population genetics to contemporary genomics-oriented and computational approaches.

3. RESULTS AND DISCUSSIONS

3.1. Annual Publication Trends

The temporal development of research related to Hardy-Weinberg equilibrium and mathematical modeling in population genetics can be observed through the annual distribution of scientific publications. Examining publication trends provides valuable insight into how scholarly attention toward a particular topic evolves over time while also reflecting broader developments in theoretical frameworks, analytical methods, and biological applications. [Figure 2](#) illustrates the annual publication output together with the percentage contribution of each year to the total number of

publications included in the dataset.

The publication trend indicates that research on Hardy-Weinberg equilibrium and mathematical approaches in population genetics has experienced a long-term increase. During the early decades, particularly from the 1970s through the late 1980s, the number of publications remained relatively low and sporadic. This period likely reflects the early stage of integrating formal mathematical modeling into population genetics research, when analytical tools and computational resources were still limited [30]. Beginning in the mid-1990s, the number of publications started to increase more steadily, suggesting a growing interest in quantitative approaches to population genetics. This growth coincides with the rapid development of molecular genetics techniques and advances in computational methods, which enabled researchers to analyze genetic variation and population structure with greater precision. Mathematical models based on Hardy-Weinberg principles became increasingly useful for interpreting genetic data and exploring theoretical scenarios related to evolutionary processes.

A more pronounced expansion in publication output can be observed from the late 2000s onward, when annual publications consistently reached higher levels and fluctuated within a relatively elevated range. This pattern suggests that Hardy-Weinberg-based analyses and mathematical modeling have become well-established components of population genetics research. The sustained research activity in recent years also

indicates that quantitative frameworks continue to play an important role in understanding genetic variation, population dynamics, and evolutionary mechanisms across diverse biological contexts.

Beyond the numerical increase in publications, the observed trend reflects the broader transformation of population genetics into a more data-intensive and quantitatively driven discipline. Advances in molecular genetics, particularly the development of high-throughput sequencing technologies and large-scale genetic datasets, have significantly expanded the scope of population genetic analyses [31]. In this context, Hardy-Weinberg equilibrium continues to function as a foundational mathematical framework for evaluating genetic variation and testing theoretical assumptions about population structure [32]. The consistent growth in publication output therefore suggests that mathematical modeling remains central to contemporary population genetics research, supporting both theoretical investigations and empirical analyses across diverse biological systems [33]. This growth is also closely linked to the rapid expansion of genomic technologies and large-scale sequencing projects over the past two decades. The availability of genome-wide datasets has created new opportunities for testing population genetic models at unprecedented resolution, thereby increasing the relevance of Hardy-Weinberg-based mathematical frameworks in modern genetic research [34].

3.2. Leading Journals and Most Productive Author

The scientific literature on Hardy-Weinberg equilibrium and mathematical modeling in population genetics is distributed across numerous journals. Identifying the leading publication outlets and the most productive author helps reveal the primary channels through which knowledge in this field has developed. The distribution of publications across journals provides an overview of the primary outlets contributing to research on Hardy-Weinberg equilibrium and mathematical modeling in population genetics. Identifying the journals that publish the largest number of articles helps reveal the main platforms through which knowledge in this research area is disseminated. Table 1 presents the ten journals with the highest number of publications in the dataset.

As shown in Table 1, Conservation Genetics Resources ranks as the leading journal in terms of publication output, contributing 69 articles, which represents approximately 3.63% of the total publications in the dataset. This indicates the strong relevance of Hardy-Weinberg-based analyses and mathematical approaches in studies related to conservation genetics, where understanding genetic diversity and population structure is essential for species management and preservation. Hardy-Weinberg-based analyses are frequently used in conservation genetics to evaluate genetic diversity, detect inbreeding, and assess whether small or fragmented populations maintain expected genotype frequencies [35]. These analyses provide essential baseline information for managing endangered species and conserving genetic resources. The prominence of this journal suggests that population genetic models remain an important tool in conservation-oriented biological research.

Several other journals also demonstrate substantial contributions to the field. Genetics and Molecular Research, Forensic Science International: Genetics, and Molecular Biology Reports appear among the most productive sources, each publishing a considerable number of studies related to Hardy-Weinberg equilibrium and quantitative genetic analysis. The presence of these journals highlights the interdisciplinary nature of population genetics research, where mathematical frameworks are applied not only in theoretical genetics but also in molecular biology and forensic science contexts. In addition, Gene, Journal of Forensic Sciences, Annals of Human Biology, Human Heredity, and Frontiers in Genetics further illustrate the broad disciplinary scope of this research domain. The inclusion of journals from fields such as human genetics, forensic genetics, and molecular biology indicates that Hardy-Weinberg equilibrium and mathematical modeling approaches are widely applied across multiple branches of biological science [36]. This diversity of publication outlets reflects the continued importance of quantitative frameworks in advancing the understanding of genetic variation and population dynamics [37].

In addition to identifying the main publication outlets, examining the most productive authors provides insight into the individuals who have

contributed most actively to research on Hardy-Weinberg equilibrium and mathematical modeling in population genetics. Productive authors often play an important role in shaping the development of a research field by consistently producing studies that advance theoretical understanding and methodological approaches [38]. Table 2 presents the ten authors with the highest number of publications within the dataset.

Budowle, B. emerges as the most productive author in this research domain, contributing 16 publications, which accounts for approximately 0.84% of the total articles in the dataset. This level of productivity highlights a sustained contribution to studies related to population genetics and the application of Hardy-Weinberg principles in genetic analysis. The prominence of this author also reflects the continued relevance of quantitative genetic approaches in fields such as forensic genetics and population structure analysis [39]. Other authors also demonstrate notable research productivity. Carracedo, A. and Pichèth, G. each contributed nine publications, while Chen, Y., Graffelman, J., Gusmão, L., and Zan, L.S. each produced eight articles within the dataset. The presence of these authors suggests the existence of multiple active research groups working on topics related to Hardy-Weinberg equilibrium, population genetic modeling, and applied genetic analyses. Their consistent publication output indicates a strong engagement with methodological and empirical investigations within this research area [40]. Additional contributors such as Chen, H., Lai, J.,

and Liu, Y. further illustrate the distributed nature of scholarly productivity in this field. Although each of these authors contributed slightly fewer publications, their work still represents a meaningful share of the research output within the dataset. Overall, the distribution of productive authors indicates that research on Hardy-Weinberg equilibrium and mathematical modeling is driven by a diverse group of scholars across different research contexts, reflecting the interdisciplinary and collaborative nature of population genetics research.

Taken together, the distribution of leading journals and productive authors reflects the multidisciplinary character of research on Hardy-Weinberg equilibrium and mathematical modeling in population genetics. The prominence of journals in conservation genetics, molecular biology, and forensic science suggests that mathematical population genetic frameworks are widely applied across both theoretical and applied biological disciplines. At the same time, the relatively dispersed distribution of productive authors indicates that research activity in this field is not concentrated within a single dominant research group, but rather distributed across multiple institutions and scientific communities. Such a pattern highlights the broad applicability of Hardy-Weinberg-based analyses and quantitative genetic models in addressing diverse biological questions, ranging from conservation and evolutionary studies to human genetics and forensic investigations.

Table 2. Most productive author.

Author Name	Publication	Percentage from Overall of Publication (%)
Budowle, B.	16	0.84
Carracedo, A.	9	0.47
Picheth, G.	9	0.47
Chen, Y.	8	0.42
Graffelman, J.	8	0.42
Gusmão, L.	8	0.42
Zan, L.S.	8	0.42
Chen, H.	7	0.37
Lai, J.	7	0.37
Liu, Y.	7	0.37

3.3. Citation Analysis

Citation analysis was conducted to evaluate the scientific impact and intellectual influence of publications related to Hardy-Weinberg equilibrium and mathematical modeling within population genetics. By identifying highly cited publications, this analysis provides insight into the studies that have contributed most substantially to the conceptual and methodological development of the field [41]. Citation-based evaluation is particularly important in bibliometric studies because highly cited publications often reflect influential scientific advances, methodological innovations, and major shifts in research direction within a discipline.

Table 3 presents the ten most highly cited publications identified within the dataset. The publication with the highest citation count was “Bayesian inference of recent migration rates using multilocus genotypes” by Wilson and Rannala, published in *Genetics* in 2003, with 1,708 citations. Several other highly cited studies were associated with molecular polymorphism, Bayesian inference, haplotype estimation, and computational genetic modeling. The dataset also shows a strong representation of journals such as *American Journal of Human Genetics*, *Genetics*, and *Genomics*, indicating the important role of human genetics and molecular genetics in shaping the development of population genetic research. Chronologically, most highly cited publications were concentrated between the early 1990s and mid-2000s, reflecting a period of rapid methodological and computational expansion in the field.

The citation patterns observed in this study reveal a substantial intellectual transition in population genetics research. Earlier influential studies were predominantly focused on genetic variation, polymorphism detection, inheritance-related disease analysis, and population-level allele distribution. These studies established important empirical and theoretical foundations for understanding genetic structure and variation within populations. However, subsequent highly cited publications increasingly emphasized statistical inference, Bayesian approaches, computational simulations, and haplotype estimation methods, demonstrating the growing importance of mathematical and computational frameworks in population genetics research.

This transition also reflects the broader evolution of population genetics from classical equilibrium-based analysis toward genomics-oriented and data-intensive approaches [42]. The emergence of computational tools such as GENELAND, alongside Bayesian and expectation-maximization algorithms, indicates that modern population genetics increasingly relies on sophisticated mathematical modeling to interpret large and complex genetic datasets [43][44]. In this context, Hardy-Weinberg equilibrium remains an important conceptual foundation, but its application has expanded considerably through the integration of computational biology, statistical genetics, and genomic technologies. These findings support the view that mathematical modeling has played a central role in shaping the intellectual development and methodological transformation of contemporary population genetics research.

3.4. Leading Countries and International Collaboration Network

The geographic distribution of publications reflects the countries that contribute most actively to research on Hardy-Weinberg equilibrium and mathematical modeling in population genetics. Analyzing leading countries together with international collaboration patterns provides insight into the global structure of research activity in this field. The geographical distribution of scientific publications provides an overview of the countries contributing to research on Hardy-Weinberg equilibrium and mathematical modeling in population genetics. Mapping publication output by country helps illustrate the global spread of research activity and highlights regions that play prominent roles in advancing this field. Figure 4 presents the spatial distribution of publications across different countries based on the total number of articles included in the dataset. As illustrated in Figure 4, research on Hardy-Weinberg equilibrium and mathematical modeling in population genetics is distributed across many regions of the world, although the intensity of research activity varies substantially between countries. The United States and China emerge as the two most prominent contributors, each producing a large number of publications compared with other countries. Their leading positions likely reflect strong research

Table 3. Most cited documents.

No.	Document Title	Authors	Source	Year of Publication	Total of Citations
1	Bayesian inference of recent migration rates using multilocus genotypes	Wilson & Rannala	Genetics	2003	1,708
2	Genetic variation at five trimeric and tetrameric tandem repeat loci in ...	Edwards et al.	Genomics	1992	1,254
3	High risk of thrombosis in patients homozygous for factor V Leiden ...	Rosendaal et al.	Blood	1995	1,066
4	GENELAND: A computer package for landscape genetics	Guillot et al.	Molecular Ecology Notes	2005	962
5	From genotypes to genes: Doubling the sample size	Sasieni	Biometrics	1997	773
6	Bayesian haplotype inference for multiple linked single-...	Niu et al.	American Journal of Human Genetics	2002	538
7	Human microsomal epoxide hydrolase: Genetic polymorphism and ...	Hassett et al.	Human Molecular Genetics	1994	482
8	Rational inferences about departures from Hardy-Weinberg equilibrium	Wittke-Thompson et al.	American Journal of Human Genetics	2005	362
9	A log-linear approach to case-parent-triad data: Assessing effects of ...	Weinberg et al.	American Journal of Human Genetics	1998	355
10	Accuracy of haplotype frequency estimation for biallelic loci, via ...	Fallin & Schork	American Journal of Human Genetics	2000	353

infrastructures, extensive funding for genetic and molecular biology studies, and the presence of numerous institutions actively engaged in population genetics research [45]. The availability of shared genomic datasets and international genetic repositories has further strengthened the capacity of these countries to conduct population genetic research [46][47]. Access to large-scale datasets enables researchers to test mathematical population genetic models across diverse populations and geographic regions [48].

The strong research presence of these countries may also be associated with their participation in large-scale genomic initiatives and population-based genetic studies [49]. International projects such as global genome sequencing programs and population genetic surveys generate extensive datasets that require population genetic frameworks, including Hardy-Weinberg-based analyses, to interpret patterns of genetic variation and population structure [50]. Access to shared genomic datasets and advanced computational infrastructure therefore plays an important role in supporting the development of population genetics research in these regions [51].

Beyond these two major contributors, several countries in Europe and Asia also demonstrate substantial research activity. Countries such as India, Germany, Spain, and Turkey show notable publication outputs, indicating that population genetics research is well established across multiple scientific communities. The presence of these countries among the leading contributors suggests that mathematical modeling and Hardy-Weinberg-based analyses are widely applied in diverse research contexts, including evolutionary biology, human genetics, and forensic genetics.

Research contributions are also visible in other regions, including South America, Africa, and Oceania, although the number of publications from these areas is generally lower. Countries such as Brazil, Australia, and South Africa contribute a moderate share of publications, reflecting growing engagement in population genetics research. Overall, the global distribution of publications indicates that studies involving Hardy-Weinberg equilibrium and mathematical modeling have developed into a widely distributed research field, with contributions emerging from a broad range of

countries and scientific environments. This global distribution also reflects the increasing importance of international population sampling and collaborative data sharing, which are essential for analyzing genetic variation across diverse populations [52][53]. These patterns also reflect the concentration of large-scale genomic initiatives and population-based sequencing projects in technologically advanced research environments. Such initiatives generate extensive datasets that require population genetic models, including Hardy-Weinberg-based analyses, to interpret patterns of genetic variation across populations.

Beyond examining national research productivity, analyzing international collaboration patterns provides additional insight into how scientific knowledge in population genetics is generated and shared across countries [54]. Collaborative networks reflect the extent to which researchers from different regions work together to address common scientific questions [55]. [Figure 5](#) illustrates the international collaboration network among countries based on co-authorship relationships within the dataset.

As shown in [Figure 5](#), the international collaboration network reveals a highly interconnected structure involving numerous countries across different regions. The United States appears as the most central node in the network, forming extensive collaborative links with many other countries. This prominent position indicates that the United States plays a key role in facilitating international research partnerships in studies related to Hardy-Weinberg equilibrium and mathematical modeling in population genetics. China also emerges as an important hub within the collaboration network, connecting with several countries across Asia, Europe, and North America. The strong presence of both the United States and China suggests that these countries not only contribute a large number of publications but also actively participate in international research collaborations. Such interactions often enhance the exchange of knowledge, methodological approaches, and research resources, which are essential for advancing complex scientific fields such as population genetics.

In addition to these major hubs, several European and regional clusters can be observed

within the network. Countries such as Spain, Germany, France, and Brazil form visible collaboration groups, reflecting active research partnerships within and across regions. The presence of these interconnected clusters indicates that international collaboration plays an important role in the development of population genetics research, enabling scientists from different countries to jointly investigate genetic variation, population structure, and evolutionary processes using shared analytical frameworks. The observed patterns of publication distribution and collaboration networks underscore the increasingly global character of research on Hardy-Weinberg equilibrium and mathematical modeling in population genetics. The strong presence of countries such as the United States and China reflects not only their substantial research capacity but also their central role in facilitating cross-border scientific cooperation. At the same time, the participation of countries from Europe, Asia, and other regions indicates that this field is supported by a broad international research community. Such collaborative networks are particularly important in population genetics, where comparative analyses across different populations and geographic regions often require shared datasets, methodological expertise, and coordinated research efforts [56]. Consequently, international collaboration continues to play a crucial role in advancing the understanding of genetic variation, population structure, and evolutionary dynamics at a global scale [57].

3.5. Keyword Co-occurrence Analysis and Thematic Clusters

The thematic labels assigned to each cluster were determined by examining the most frequent and centrally positioned keywords within each group of the network [58]. These keywords represent the dominant research topics associated with each cluster and therefore provide the basis for interpreting the thematic structure of the field [59]. Analyzing keyword co-occurrence provides insight into the conceptual structure of a research field by identifying terms that frequently appear together in scientific publications [60]. Through this approach, it is possible to reveal thematic clusters that represent the principal research directions and

methodological approaches within the literature [61]. Figure 6 presents the network visualization of author keywords related to Hardy-Weinberg equilibrium and mathematical modeling in population genetics, highlighting the relationships among frequently occurring terms.

As illustrated in Figure 6, the keyword co-occurrence analysis reveals several interconnected thematic clusters that characterize the research landscape of population genetics studies involving Hardy-Weinberg equilibrium and mathematical modeling. These clusters represent different yet complementary research directions, ranging from molecular marker analysis and genetic polymorphism studies to population structure analysis and applications in forensic genetics. The network structure indicates that while these themes form distinct conceptual groups, they remain strongly interconnected through shared methodological approaches and analytical frameworks.

3.5.1. Cluster 1: Genetic Polymorphism and Molecular Markers

The red cluster is primarily centered on keywords such as polymorphism, single nucleotide polymorphism, genetic polymorphism, genotype, and association. This cluster reflects a major research direction focusing on genetic variation at the molecular level and its association with phenotypic traits or disease conditions. The prominence of terms such as breast cancer, obesity,

and hypertension suggests that many studies apply population genetics frameworks to investigate genetic associations in biomedical contexts. In addition, methodological terms such as *PCR-RFLP* and references to livestock species including cattle indicate that this research theme also extends to agricultural and animal genetics. These studies often aim to identify genetic polymorphisms associated with economically important traits or biological characteristics. Overall, this cluster highlights the central role of molecular genetic variation in advancing research on population genetics and applied genomics. The red cluster is primarily centered on keywords such as polymorphism, single nucleotide polymorphism, genetic polymorphism, genotype, and association, indicating a strong focus on molecular genetic variation studies.

3.5.2. Cluster 2: Population Genetics and Forensic Applications

The blue cluster centers around population genetics, forensic science, DNA typing, short tandem repeat, and polymerase chain reaction. This thematic group represents research focusing on the use of genetic markers to analyze population structure and genetic identity. The presence of forensic-related keywords indicates that Hardy-Weinberg equilibrium plays a crucial role in validating allele frequency distributions used in forensic DNA analysis. Studies within this cluster often involve the use of short tandem repeat (STR)

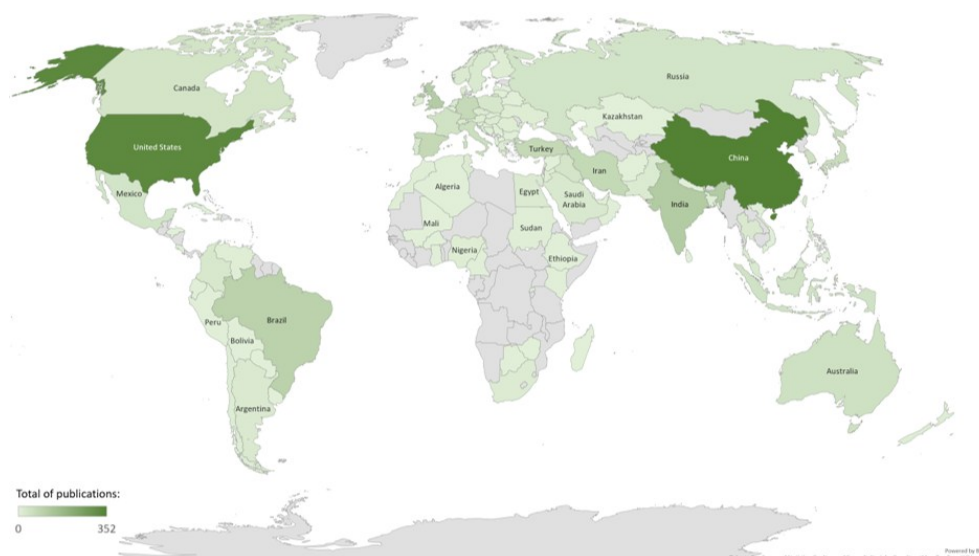


Figure 4. Global distribution of publications by country.

markers and PCR-based techniques to evaluate genetic variation within and between populations. These approaches are widely applied in forensic investigations, identity testing, and population structure studies. Consequently, this cluster reflects the intersection between theoretical population genetics and practical applications in forensic and legal contexts. The blue cluster is characterized by keywords including forensic science, DNA typing, short tandem repeat, and polymerase chain reaction, which collectively represent research directions related to forensic genetic identification.

3.5.3. Cluster 3: Genetic Diversity and Population Structure

The green cluster focuses on themes related to genetic diversity, allele frequency, genetic variation, microsatellites, and genetic structure. This group of keywords reflects studies aimed at understanding the distribution of genetic variation within and among populations. Research in this cluster often applies Hardy-Weinberg equilibrium as a theoretical framework to assess whether populations exhibit expected genotype frequencies under conditions of random mating. Furthermore, the presence of keywords such as linkage disequilibrium, heterozygosity, and conservation suggests that many studies in this cluster are concerned with evolutionary processes and conservation genetics. These investigations frequently use microsatellite markers or other molecular tools to analyze population differentiation, genetic diversity levels, and the evolutionary dynamics of species. In conservation genetics, these analytical approaches are widely used to monitor genetic diversity in endangered populations and to evaluate whether conservation interventions maintain sufficient genetic variation for long-term population viability [62]. The green cluster includes dominant keywords such as genetic diversity, allele frequency, genetic structure, and microsatellites, reflecting studies that focus on population structure and evolutionary genetic processes.

3.5.4. Cluster 4: Genomic Markers and High-throughput Sequencing

The yellow cluster represents a more recent research direction centered on genomic

technologies and advanced molecular markers. Keywords such as SNP marker, RAD-seq, and transcriptome indicate the increasing use of high-throughput sequencing techniques in population genetics research. These approaches allow researchers to analyze large numbers of genetic loci simultaneously, thereby providing more detailed insights into population structure and evolutionary patterns. The integration of genomic-scale data has significantly expanded the analytical capacity of population genetics studies. By combining SNP-based analyses with mathematical models and statistical frameworks, researchers are now able to investigate complex evolutionary processes with greater resolution. This cluster therefore reflects the ongoing transition of population genetics toward genomics-driven research methodologies. The yellow cluster is dominated by keywords such as SNP marker, RAD-seq, and transcriptome, indicating a research direction associated with genomic technologies and high-throughput sequencing approaches.

The overall structure of the keyword network highlights the interdisciplinary nature of research on Hardy-Weinberg equilibrium and mathematical modeling in population genetics. While some themes emphasize classical population genetics concepts such as allele frequency, genetic diversity, and population structure, others reflect methodological developments in molecular genetics, forensic applications, and genomic technologies. The strong interconnections among these thematic clusters indicate that advances in population genetics increasingly rely on the integration of theoretical models, molecular markers, and large-scale genomic data [63]. Understanding how these themes interact provides a foundation for examining how research topics in this field have evolved over time, which is explored further in the following section [64]. These thematic patterns also illustrate how research on Hardy-Weinberg equilibrium has evolved beyond its traditional role as a theoretical population genetics model. The presence of clusters related to molecular polymorphisms, forensic genetics, conservation genetics, and genomic technologies indicates that the Hardy-Weinberg framework has become an essential analytical tool across multiple areas of biological research [65]. Rather than remaining

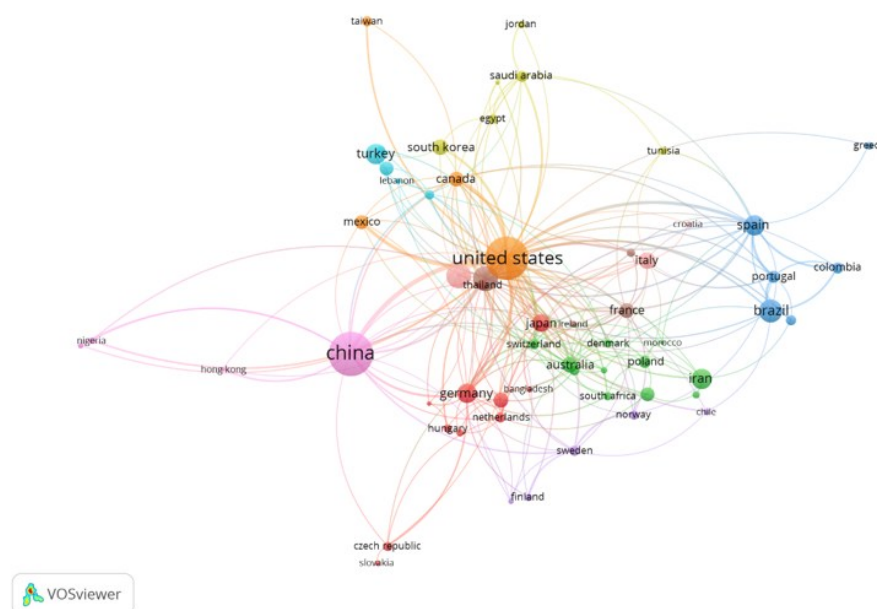


Figure 5. International collaboration network among countries.

confined to theoretical population studies, the model is now widely integrated into empirical investigations involving molecular markers, genomic datasets, and applied genetic analyses [66]. This broad applicability highlights the continuing importance of mathematical frameworks in interpreting genetic data and demonstrates how classical population genetics concepts continue to support modern research in molecular biology, evolutionary studies, and applied genomics [67].

The continued prominence of genetic diversity, allele frequency, and population structure within the keyword clusters also highlights the strong relevance of Hardy-Weinberg-based analyses for conservation genetics. Monitoring deviations from equilibrium conditions provides important insights into inbreeding, genetic drift, and population fragmentation, which are key concerns in biodiversity conservation and genetic resource management. The emergence of clusters related to genomic markers and population structure also reflects the growing demand for high-resolution genetic analyses in conservation genetics and biodiversity monitoring. As genomic datasets become increasingly available for both model and non-model organisms, mathematical population genetic frameworks remain essential tools for interpreting genetic diversity and guiding conservation strategies.

3.6. Historiographic Evolution and Research Periodization

Examining the temporal evolution of research themes provides important insight into how scientific attention and methodological priorities within a field change over time [68]. In bibliometric studies, thematic evolution analysis enables the identification of major transitions in scientific development by tracing how dominant research themes emerge, expand, and shift across different chronological periods [69]. In the present study, thematic evolution analysis was performed using the Bibliometrix package through the Biblioshiny interface in R Studio to examine the historiographic development of research related to Hardy-Weinberg equilibrium and mathematical modeling in population genetics. Figure 7 illustrates the thematic evolution of research topics across four chronological periods spanning from 1972 to 2026.

As illustrated in Figure 7, the thematic evolution analysis reveals a clear transition in the intellectual development of population genetics research over time. During the earliest period (1972–2006), the dominant research themes were primarily associated with classical population genetics concepts, including allele, alleles, gene frequency, genotype, and genetic variability. These themes indicate that early studies largely focused on theoretical population genetic frameworks, allele

distribution patterns, and equilibrium-based analyses used to understand genetic variation within populations. The prominence of these foundational terms reflects the strong influence of classical Hardy-Weinberg principles during the formative stage of population genetics research.

In the subsequent period (2007–2013), the thematic structure began to shift toward more molecularly oriented approaches. Keywords such as haplotype, genotype, and gene frequency became increasingly prominent, indicating the growing integration of molecular marker analysis and inheritance-based genetic studies. This transition suggests that advances in molecular biology and statistical genetics substantially influenced the analytical direction of population genetics research during this period. The emergence of haplotype-related themes also reflects the increasing use of multilocus genetic data and computational inference methods for evaluating genetic relationships and population structure.

A more substantial transition became evident during the 2014–2019 period, when themes related to genomic DNA emerged prominently within the thematic network. The appearance of genomics-related terminology indicates the expansion of population genetics into data-intensive and genomics-oriented research approaches. During this phase, population genetic analyses increasingly relied on high-throughput sequencing technologies,

large-scale molecular datasets, and computational analytical frameworks capable of evaluating genome-wide variation. At the same time, classical themes such as gene frequency and genotype remained strongly connected within the network, suggesting that traditional Hardy–Weinberg-based mathematical frameworks continued to provide an important conceptual foundation despite major technological advancements.

In the most recent period (2020–2026), the thematic evolution analysis demonstrates a stronger emphasis on genetic variability, gene frequency, and allele-related studies. This pattern indicates that contemporary population genetics research increasingly focuses on understanding genetic diversity and population-level genomic variation using more advanced molecular and computational approaches. The persistence of gene frequency and allele-related themes across all chronological periods further demonstrates the continuing relevance of Hardy-Weinberg-based population genetic principles within modern genomics research. Rather than being replaced by newer technologies, classical mathematical population genetic frameworks appear to have evolved alongside computational and genomics-driven approaches.

Overall, the thematic evolution analysis reveals that the development of population genetics research has progressed through several major

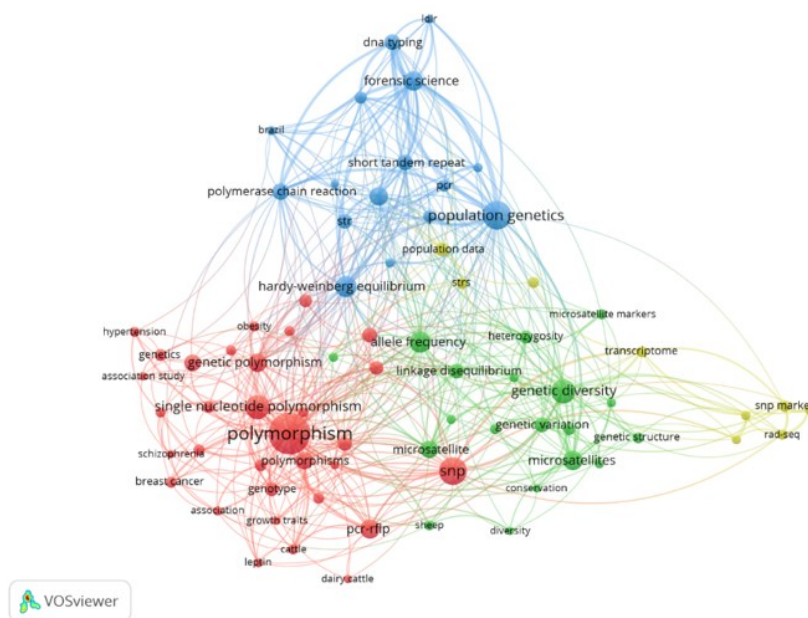


Figure 6. Keyword co-occurrence network and thematic clusters.

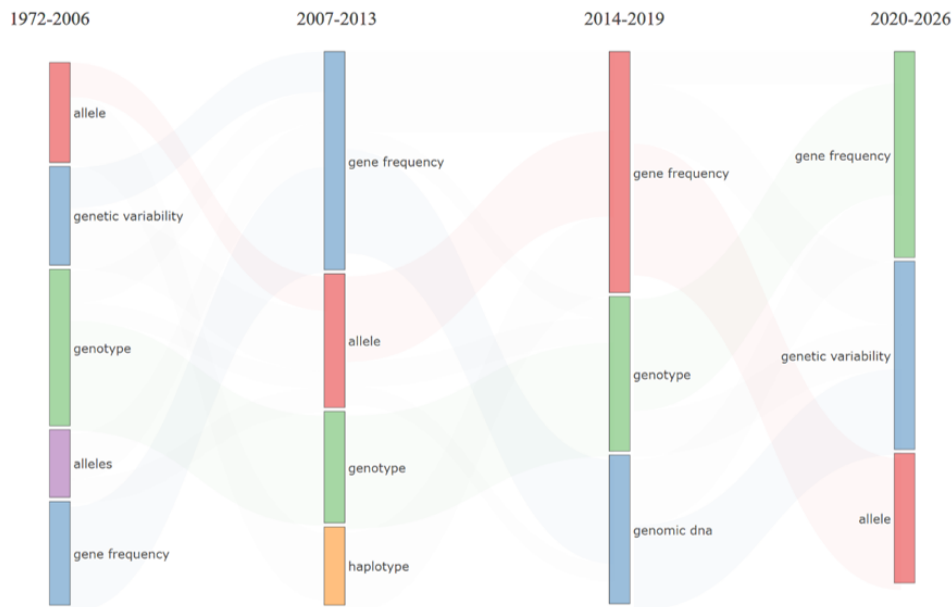


Figure 7. Historiographic evolution and thematic periodization of research topics related to Hardy-Weinberg equilibrium and mathematical modeling in population genetics from 1972 to 2026.

historiographic phases, beginning with classical equilibrium-based population genetics, followed by the expansion of molecular and haplotype-based analyses, and ultimately transitioning toward genomics-oriented and computationally intensive research. These findings demonstrate that mathematical models derived from Hardy-Weinberg principles continue to play a central role in interpreting increasingly complex genetic datasets generated through modern molecular and genomic technologies. The observed transition toward genomics-driven population genetics also reflects broader technological shifts in biological sciences, particularly the widespread adoption of high-throughput sequencing, computational biology, and large-scale genomic analysis frameworks.

3.7. Practical Implications for Applied Natural Science

The historiographic evolution identified in this study demonstrates that Hardy-Weinberg equilibrium has progressed from a classical theoretical framework into an essential analytical foundation within modern genomics-oriented population genetics. The transition from allele frequency-based analyses toward computational, molecular, and genomics-driven approaches highlights the increasingly important role of

mathematical and statistical modeling in interpreting large-scale genetic datasets. These findings indicate that Hardy-Weinberg-based mathematical frameworks remain highly relevant for applied natural science fields such as biodiversity conservation, forensic genetics, breeding programs, and population monitoring, while also emphasizing the growing integration of computational biology, bioinformatics, and genomics methodologies in contemporary population genetics research.

3.8. Study Limitations

This study has several limitations that should be acknowledged when interpreting the results. First, the bibliometric dataset was derived exclusively from the Scopus database. Although Scopus is widely recognized as one of the most comprehensive sources of peer-reviewed literature, it does not include all relevant publications indexed in other databases such as Web of Science, PubMed, or Google Scholar. Consequently, some relevant studies on Hardy-Weinberg equilibrium and mathematical modeling in population genetics may not have been captured in the present analysis. Second, the study included only research articles written in English and relied primarily on bibliographic metadata such as titles, abstracts, and author keywords. As a result, research published in

other languages or in different document types (e.g., reviews, conference papers, or book chapters) was not considered. In addition, the bibliometric analysis focused on keyword co-occurrence and collaboration networks derived from metadata rather than full-text content. Future studies could expand the scope of analysis by integrating multiple databases and incorporating additional text-mining or citation-based approaches to provide a more comprehensive representation of the research landscape in population genetics.

In addition, several methodological factors inherent to bibliometric studies should be considered when interpreting the results. The construction of the search query and the selection of keywords may introduce bias by potentially excluding relevant studies that use alternative terminology. Differences in database indexing practices may also influence the visibility of certain publications or research fields within the dataset. Furthermore, bibliometric mapping techniques such as keyword co-occurrence analysis depend on the structure of the available metadata and the parameters used in visualization software. As a result, the identified thematic clusters represent general patterns within the literature rather than exhaustive representations of all research directions in population genetics.

4. CONCLUSIONS

This bibliometric study demonstrates that research related to Hardy-Weinberg equilibrium and mathematical modeling in population genetics has evolved substantially from classical equilibrium-based population genetics toward molecular, computational, and genomics-oriented approaches. Publication trends, citation analysis, collaboration networks, and thematic evolution collectively indicate that mathematical modeling has remained a central foundation throughout the intellectual development of the field. We found that classical Hardy-Weinberg principles has not been replaced by modern genomics research. In contrast, modern genomics research instead expanded their application through the integration of computational biology, statistical genetics, and large-scale genomic technologies. These findings highlight the continuing relevance of Hardy-Weinberg-based

mathematical frameworks in contemporary population genetics and provide a clearer understanding of the historiographic development, state of the art, and future research direction within this field.

AUTHOR INFORMATION

Corresponding Author

Priyambodo Priyambodo — Department of Biology, Universitas Lampung, Bandar Lampung-35145 (Indonesia);

 orcid.org/0000-0003-1856-2637

Email: priyambodo@fmipa.unila.ac.id

Authors

Agung Putra Wijaya — Department of Mathematics Education, Universitas Lampung, Bandar Lampung-35145 (Indonesia);

 orcid.org/0000-0001-6993-757X

Nindy Permatasari — Department of Plantation Crop Cultivation, Politeknik Negeri Lampung, Bandar Lampung-35141 (Indonesia);

 orcid.org/0009-0002-2659-0531

Hidayatulloh Hidayatulloh — Department of Mathematics Education, Universitas Muhammadiyah Pringsewu Lampung, Pringsewu-35373 (Indonesia);

 orcid.org/0000-0002-8958-6276

Siti Hamidatul Aliyah — Center for Biomedical Research, National Research and Innovation Agency (BRIN), Cibinong-16911 (Indonesia);

 orcid.org/0000-0003-0017-3123

Elly Lestari Rustiati — Department of Biology, Universitas Lampung, Bandar Lampung-35145 (Indonesia);

 orcid.org/0000-0003-4257-0004

Author Contributions

P. P. was responsible for the data analysis and the overall preparation of the manuscript. N. P. and E. L. R. contributed by providing critical review and academic input in the fields of biology and genetics. S. H. A. assisted in retrieving bibliographic data from the Scopus database and contributed to the review of the manuscript from a genetics perspective. A. P. W. and H. H. contributed to drafting sections related to

mathematical aspects and their connections to biological applications.

Conflicts of Interest

The authors declare no conflict of interest.

ACKNOWLEDGEMENT

This research supported by the Institute of Research and Community Service (Lembaga Penelitian dan Pengabdian kepada Masyarakat) Universitas Lampung, with collaborate with other researcher from Politeknik Negeri Lampung, Universitas Muhammadiyah Pringsewu Lampung, and National Research and Innovation Agency.

DECLARATION OF GENERATIVE AI

While preparing this work, authors utilize ChatGPT to refine English grammar and edit sentences more effectively. The results generated by the AI tools are reviewed by the author members, and all authors are solely responsible for the publication content.

REFERENCES

- [1] S. Forestiero. (2022). "The Historical Nature of Biological Complexity and the Ineffectiveness of the Mathematical Approach to It". *Theory in Biosciences*. **141** (2): 213-231. [10.1007/s12064-022-00369-7](https://doi.org/10.1007/s12064-022-00369-7).
- [2] J. R. Banga and A. F. Villaverde. (2025). "Mechanistic Dynamic Modelling of Biological Systems: The Road Ahead". *Current Opinion in Systems Biology*. **42** : 100553. [10.1016/j.coisb.2025.100553](https://doi.org/10.1016/j.coisb.2025.100553).
- [3] M. Montévil. (2022). "Historicity at the Heart of Biology". *Theory in Biosciences*. **141** (2): 165-173. [10.1007/s12064-020-00320-8](https://doi.org/10.1007/s12064-020-00320-8).
- [4] D. Nichol, M. Robertson-Tessi, A. R. A. Anderson, and P. Jeavons. (2019). "Model Genotype-Phenotype Mappings and the Algorithmic Structure of Evolution". *Journal of the Royal Society Interface*. **16** (160). [10.1098/rsif.2019.0332](https://doi.org/10.1098/rsif.2019.0332).
- [5] G. S. Bradburd and P. L. Ralph. (2019). "Spatial Population Genetics: It's About Time". *Annual Review of Ecology, Evolution, and Systematics*. **50** (1): 427-449. [10.1146/annurev-ecolsys-110316-022659](https://doi.org/10.1146/annurev-ecolsys-110316-022659).
- [6] W. Hao and J. D. Storey. (2019). "Extending Tests of Hardy-Weinberg Equilibrium to Structured Populations". *Genetics*. **213** (3): 759-770. [10.1534/genetics.119.302370](https://doi.org/10.1534/genetics.119.302370).
- [7] N. Abramovs, A. Brass, and M. Tassabehji. (2020). "Hardy-Weinberg Equilibrium in the Large-Scale Genomic Sequencing Era". *Frontiers in Genetics*. **11** : 210. [10.3389/fgene.2020.00210](https://doi.org/10.3389/fgene.2020.00210).
- [8] M. W. Klymkowsky. (2023). "Rethinking (Again) Hardy-Weinberg and Genetic Drift in Undergraduate Biology". *Frontiers in Genetics*. **14** : 1199739. [10.3389/fgene.2023.1199739](https://doi.org/10.3389/fgene.2023.1199739).
- [9] W. S. Pearman, L. Urban, and A. Alexander. (2022). "Commonly Used Hardy-Weinberg Equilibrium Filtering Schemes Impact Population Structure Inferences Using RADseq Data". *Molecular Ecology Resources*. **22** (7): 2599-2613. [10.1111/1755-0998.13646](https://doi.org/10.1111/1755-0998.13646).
- [10] J. G. Ramírez-García, S. P. Maciel-Torres, M. Hernández-Rodríguez, P. Arenas-Báez, J. F. Orzuna-Orzuna, and L. D. Granados-Rivera. (2025). "Non-Invasive Sampling for Population Genetics of Wild Terrestrial Mammals (2015-2025): A Systematic Review". *Diversity*. **17** (11): 760. [10.3390/d17110760](https://doi.org/10.3390/d17110760).
- [11] S. Li, D. Li, L. Zhang, W. Shang, B. Qin, and Y. Jiang. (2021). "High Levels of Genetic Diversity and an Absence of Genetic Structure Among Breeding Populations of the Endangered Rufous-Backed Bunting in Inner Mongolia, China: Implications for Conservation". *Avian Research*. **12** (1): 7. [10.1186/s40657-020-00236-3](https://doi.org/10.1186/s40657-020-00236-3).
- [12] R. S. Waples. (2025). "The Idiot's Guide to Effective Population Size". *Molecular Ecology*. **34** (23): e17670. [10.1111/mec.17670](https://doi.org/10.1111/mec.17670).
- [13] Y. X. C. Bourgeois and B. H. Warren. (2021). "An Overview of Current Population Genomics Methods for the Analysis of Whole-Genome Resequencing Data in Eukaryotes". *Molecular Ecology*. **30** : 6036-6071. [10.1111/mec.15989](https://doi.org/10.1111/mec.15989).

- [14] P. A. Hohenlohe, W. C. Funk, and O. P. Rajora. (2021). "Population Genomics for Wildlife Conservation and Management". *Molecular Ecology*. **30** : 62-82. [10.1111/mec.15720](https://doi.org/10.1111/mec.15720).
- [15] P. A. Sofi, S. M. Zargar, A. Hamadani, S. Shafi, A. Zaffar, D. Bijarniya, and P. V. V. Prasad. (2024). In: "A Biologist's Guide to Artificial Intelligence". Academic Press. 47-66. [10.1016/B978-0-443-24001-0.00004-X](https://doi.org/10.1016/B978-0-443-24001-0.00004-X).
- [16] J. Griesemer. (2020). In: "S. Leonelli and N. Tempini (Eds.) Data Journeys in the Sciences." Cham: Springer. [10.1007/978-3-030-37177-7_8](https://doi.org/10.1007/978-3-030-37177-7_8).
- [17] R. M. Schweizer, N. Saarman, K. M. Ramstad, B. R. Forester, J. L. Kelley, B. K. Hand, R. L. Malison, A. S. Ackiss, M. Watsa, T. C. Nelson, A. Beja-Pereira, R. S. Waples, W. C. Funk, and G. Luikart. (2021). "Big Data in Conservation Genomics: Boosting Skills, Hedging Bets, and Staying Current in the Field". *Journal of Heredity*. **112** (4): 313-327. [10.1093/jhered/esab019](https://doi.org/10.1093/jhered/esab019).
- [18] Y. H. Hsu. (2026). "Hardy-Weinberg Equilibrium Filtering in Population Genomics: Empirical Review and Decision Framework for Improved Practice". *Ecology and Evolution*. **16** (1): e72688. [10.1002/ece3.72688](https://doi.org/10.1002/ece3.72688).
- [19] M. J. Simpson and M. J. Plank. (2025). "Inference and Prediction for Stochastic Models of Biological Populations Undergoing Migration and Proliferation". *Journal of the Royal Society Interface*. **22** (231). [10.1098/rsif.2025.0536](https://doi.org/10.1098/rsif.2025.0536).
- [20] S. Nagar. (2024). "Exploring Mathematical Models in Genetics: Insights, Applications and Future Directions". *Arya Bhatta Journal of Mathematics and Informatics*. **16** (1): 31-36. [10.5958/2394-9309.2024.00005.0](https://doi.org/10.5958/2394-9309.2024.00005.0).
- [21] L. Sun, J. Gan, L. Jiang, and R. Wu. (2021). "Recursive Test of Hardy-Weinberg Equilibrium in Tetraploids". *Trends in Genetics*. **37** (6): 504-513. [10.1016/j.tig.2020.11.006](https://doi.org/10.1016/j.tig.2020.11.006).
- [22] J. N. Hellwege, J. M. Keaton, A. Giri, X. Gao, D. R. Velez Edwards, and T. L. Edwards. (2017). "Population Stratification in Genetic Association Studies". *Current Protocols in Human Genetics*. **95** (1): 1-22. [10.1002/cphg.48](https://doi.org/10.1002/cphg.48).
- [23] D. Nettle and W. E. Frankenhuis. (2019). "The Evolution of Life-History Theory: A Bibliometric Analysis of an Interdisciplinary Research Area". *Proceedings of the Royal Society B: Biological Sciences*. **286** (1899): 20190040. [10.1098/rspb.2019.0040](https://doi.org/10.1098/rspb.2019.0040).
- [24] H. Neamatzadeh, S. A. Dastgheib, M. Mazaheri, A. Masoudi, A. Shiri, A. Omid, A. Rahmani, A. Golshan-Tafti, M. Aghasipour, M. Yeganegi, and M. Bahrami. (2024). "Hardy-Weinberg Equilibrium in Meta-Analysis Studies and Large-Scale Genomic Sequencing Era". *Asian Pacific Journal of Cancer Prevention*. **25** (7): 2229-2236. [10.31557/APJCP.2024.25.7.2229](https://doi.org/10.31557/APJCP.2024.25.7.2229).
- [25] K. Raza and S. Ahmad. (2019). "Recent Advancement in Next-Generation Sequencing Techniques and Its Computational Analysis". *International Journal of Bioinformatics Research and Applications*. **15** (3): 191-220. [10.1504/IJBRA.2019.10022508](https://doi.org/10.1504/IJBRA.2019.10022508).
- [26] Y. C. Fu, M. Marques, Y. H. Tseng, J. J. Powell, and D. P. Baker. (2022). "An Evolving International Research Collaboration Network: Spatial and Thematic Developments in Co-Authored Higher Education Research, 1998-2018". *Scientometrics*. **127** (3): 1403-1429. [10.1007/s11192-021-04200-w](https://doi.org/10.1007/s11192-021-04200-w).
- [27] A. Parlina, K. Ramli, and H. Murfi. (2020). "Theme Mapping and Bibliometrics Analysis of One Decade of Big Data Research in the Scopus Database". *Information*. **11** (2): 69. [10.3390/info11020069](https://doi.org/10.3390/info11020069).
- [28] J. Baas, M. Schotten, A. Plume, G. Côté, and R. Karimi. (2020). "Scopus as a Curated, High-Quality Bibliometric Data Source for Academic Research in Quantitative Science Studies". *Quantitative Science Studies*. **1** (1): 377-386. [10.1162/qss_a_00019](https://doi.org/10.1162/qss_a_00019).
- [29] R. Nezameslami, A. Nezameslami, B. Mehdikhani, A. Mosavi-Jarrahi, A. Shahbazi, A. Rahmani, A. Masoudi, M. Yeganegi, R. Akhondzardaini, M. Bahrami, and K. Aghili. (2025). "Adapting PRISMA Guidelines to Enhance Reporting Quality in Genetic

- Association Studies: A Framework Proposal". *Asian Pacific Journal of Cancer Prevention*. **26** (5): 1641-1648. [10.31557/APJCP.2025.26.5.1641](https://doi.org/10.31557/APJCP.2025.26.5.1641).
- [30] M. Campbell-Kelly, W. F. Aspray, J. R. Yost, H. Tinn, and G. C. Díaz. (2023). "Computer: A History of the Information Machine". Routledge. [10.4324/9781003263272](https://doi.org/10.4324/9781003263272).
- [31] Y. Lu, M. Li, Z. Gao, H. Ma, Y. Chong, J. Hong, J. Wu, D. Wu, D. Xi, and W. Deng. (2025). "Advances in Whole Genome Sequencing: Methods, Tools, and Applications in Population Genomics". *International Journal of Molecular Sciences*. **26** (1): 372. [10.3390/ijms26010372](https://doi.org/10.3390/ijms26010372).
- [32] C. Liu, D. L. Dreger, S. Liu, A. Samarapungavan, S. M. Gardner, K. L. Clase, and N. Pelaez. (2024). "The Conceptual Analysis of Disciplinary Evidence (CADE) Framework as a Guide for Evidentiary Reasoning: A Practical Implementation in a Hardy-Weinberg Equilibrium (HWE) Laboratory Investigation". *Journal of Biological Education*. **58** (3): 626-650. [10.1080/00219266.2022.2100450](https://doi.org/10.1080/00219266.2022.2100450).
- [33] J. R. Adrion, C. B. Cole, N. Dukler, J. G. Galloway, A. L. Gladstein, G. Gower, C. C. Kyriazis, A. P. Ragsdale, G. Tsambos, F. Baumdicker, and J. Carlson. (2020). "A Community-Maintained Standard Library of Population Genetic Models". *eLife*. **9** : e54967. [10.7554/eLife.54967](https://doi.org/10.7554/eLife.54967).
- [34] J. Kelleher, Y. Wong, A. W. Wohns, C. Fadil, P. K. Albers, and G. McVean. (2019). "Inferring Whole-Genome Histories in Large Population Datasets". *Nature Genetics*. **51** (9): 1330-1338. [10.1038/s41588-019-0483-y](https://doi.org/10.1038/s41588-019-0483-y).
- [35] E. A. Snegin, V. M. Makeeva, A. P. Kaledin, A. M. Ostapchuk, I. D. Alazneli, and A. V. Smurov. (2021). "Genetic Diversity of the Central European Wild Boar (*Sus scrofa scrofa*) Population and Domestic Pig (*Sus scrofa domestica*) Breeds Based on a Microsatellite DNA Locus". *Vavilov Journal of Genetics and Breeding*. **25** (8): 822-829. [10.18699/VJ21.095](https://doi.org/10.18699/VJ21.095).
- [36] A. M. González-Ferreras, S. Leal, J. Barquín, and A. Almodóvar. (2022). "Patterns of Genetic Diversity of Brown Trout in a Northern Spanish Catchment Linked to Structural Connectivity". *Aquatic Sciences*. **84** (4): 48. [10.1007/s00027-022-00877-0](https://doi.org/10.1007/s00027-022-00877-0).
- [37] J. Benn Torres. (2020). "Anthropological Perspectives on Genomic Data, Genetic Ancestry, and Race". *American Journal of Physical Anthropology*. **171** : 74-86. [10.1002/ajpa.23979](https://doi.org/10.1002/ajpa.23979).
- [38] C. Post, R. Sarala, C. Gatrell, and J. E. Prescott. (2020). "Advancing Theory with Review Articles". *Journal of Management Studies*. **57** (2): 351-376. [10.1111/joms.12549](https://doi.org/10.1111/joms.12549).
- [39] J. Nicolaisen. (2025). "Research Persistence: A Comparative Study of Core and Peripheral Journal Authors in Five Medical Research Topics". *Journal of Scholarly Publishing*. **56** (4): 797-814. [10.3138/jsp-2024-0134](https://doi.org/10.3138/jsp-2024-0134).
- [40] M. H. Arsalan, O. Mubin, A. A. Mahmud, and S. Perveen. (2025). "What Factors Influence Research Impact? An Empirical Study on the Interplay of Research, Publications, Researchers, Institutions, and National Conditions". *Journal of Data and Information Science*. **10** (1): 188-227. [10.2478/jdis-2025-0001](https://doi.org/10.2478/jdis-2025-0001).
- [41] A. Klarin. (2024). "How to Conduct a Bibliometric Content Analysis: Guidelines and Contributions of Content Co-Occurrence or Co-Word Literature Reviews". *International Journal of Consumer Studies*. **48** (2): e13031. [10.1111/ijcs.13031](https://doi.org/10.1111/ijcs.13031).
- [42] T. Leroy and M. Heuertz. (2026). "Adaptive Introgression in the Context of Climate Adaptation". *Molecular Ecology*. **35** (8): e70333. [10.1111/mec.70333](https://doi.org/10.1111/mec.70333).
- [43] K. Korfmann, O. E. Gaggiotti, and M. Fumagalli. (2023). "Deep Learning in Population Genetics". *Genome Biology and Evolution*. **15** (2): evad008. [10.1093/gbe/evad008](https://doi.org/10.1093/gbe/evad008).
- [44] M. W. Khan. (2024). "A Comprehensive Survey of Genetic Programming Applications in Modern Biological Research". *MEDIN*. [10.47852/bonviewMEDIN42023692](https://doi.org/10.47852/bonviewMEDIN42023692).
- [45] F. Kershaw, M. W. Bruford, W. C. Funk, C.

- E. Grueber, S. Hoban, M. E. Hunter, L. Laikre, A. J. MacDonald, M. H. Meek, C. Mittan, and D. O'Brien. (2022). "The Coalition for Conservation Genetics: Working Across Organizations to Build Capacity and Achieve Change in Policy and Practice". *Conservation Science and Practice*. **4** (4): e12635. [10.1111/csp2.12635](https://doi.org/10.1111/csp2.12635).
- [46] J. B. Byrd, A. C. Greene, D. V. Prasad, X. Jiang, and C. S. Greene. (2020). "Responsible, Practical Genomic Data Sharing That Accelerates Research". *Nature Reviews Genetics*. **21** (10): 615-629. [10.1038/s41576-020-0257-5](https://doi.org/10.1038/s41576-020-0257-5).
- [47] D. M. Leigh, A. G. Vandergast, M. E. Hunter, E. D. Crandall, W. C. Funk, C. J. Garroway, S. Hoban, S. J. Oyler-McCance, C. Rellstab, G. Segelbacher, and C. Schmidt. (2024). "Best Practices for Genetic and Genomic Data Archiving". *Nature Ecology & Evolution*. **8** (7): 1224-1232. [10.1038/s41559-024-02423-7](https://doi.org/10.1038/s41559-024-02423-7).
- [48] X. Huang, A. Rymbekova, O. Dolgova, O. Lao, and M. Kuhlwilm. (2024). "Harnessing Deep Learning for Population Genetic Inference". *Nature Reviews Genetics*. **25** (1): 61-78. [10.1038/s41576-023-00636-3](https://doi.org/10.1038/s41576-023-00636-3).
- [49] R. González-José, E. Alfaro, V. Arencibia, C. F. Argüelles, S. Avena, G. Bailliet, M. Berenstein, C. M. Bravi, M. Cuello, J. E. Dipierri, and H. Dopazo. (2025). "Program of Genomic Reference and Biobank of the Argentinian Population: A National Initiative for Genomic Equity and Population-Based Research in Argentina". *Biopreservation and Biobanking*. **23** (6): 512-522. [10.1177/19475535251390754](https://doi.org/10.1177/19475535251390754).
- [50] D. S. Zarlenga, E. P. Hoberg, and J. T. Detwiler. (2022). In: "Helminth Infections and Their Impact on Global Public Health". Cham: Springer International Publishing. 35-72. [10.1007/978-3-031-00303-5_2](https://doi.org/10.1007/978-3-031-00303-5_2).
- [51] Z. Stark, D. Glazer, O. Hofmann, A. Rendon, C. R. Marshall, G. S. Ginsburg, C. Lunt, N. Allen, M. Effingham, J. Hastings Ward, and S. L. Hill. (2025). "A Call to Action to Scale Up Research and Clinical Genomic Data Sharing". *Nature Reviews Genetics*. **26** (2): 141-147. [10.1038/s41576-024-00776-0](https://doi.org/10.1038/s41576-024-00776-0).
- [52] W. D. Chawinga and S. Zinn. (2019). "Global Perspectives of Research Data Sharing: A Systematic Literature Review". *Library & Information Science Research*. **41** (2): 109-122. [10.1016/j.lisr.2019.04.004](https://doi.org/10.1016/j.lisr.2019.04.004).
- [53] M. Kwiek. (2021). "What Large-Scale Publication and Citation Data Tell Us About International Research Collaboration in Europe: Changing National Patterns in Global Contexts". *Studies in Higher Education*. **46** (12): 2629-2649. [10.1080/03075079.2020.1749254](https://doi.org/10.1080/03075079.2020.1749254).
- [54] A. Middleton, R. Milne, M. A. Almarri, S. Anwer, J. Atutornu, E. E. Baranova, P. Bevan, M. Cerezo, Y. Cong, C. Critchley, and J. Fernow. (2020). "Global Public Perceptions of Genomic Data Sharing: What Shapes the Willingness to Donate DNA and Health Data?". *The American Journal of Human Genetics*. **107** (4): 743-752. [10.1016/j.ajhg.2020.08.023](https://doi.org/10.1016/j.ajhg.2020.08.023).
- [55] S. Beck, C. Bergenholtz, M. Bogers, T. M. Brasseur, M. L. Conradsen, D. Di Marco, A. P. Distel, L. Dobusch, D. Dörler, A. Effert, and B. Fecher. (2022). "The Open Innovation in Science Research Field: A Collaborative Conceptualisation Approach". *Industry and Innovation*. **29** (2): 136-185. [10.1080/13662716.2020.1792274](https://doi.org/10.1080/13662716.2020.1792274).
- [56] S. V. Edwards, V. V. Robin, N. Ferrand, and C. Moritz. (2022). "The Evolution of Comparative Phylogeography: Putting the Geography (and More) into Comparative Population Genomics". *Genome Biology and Evolution*. **14** (1): e176. [10.1093/gbe/evab176](https://doi.org/10.1093/gbe/evab176).
- [57] S. Hoban, F. I. Archer, L. D. Bertola, J. G. Bragg, M. F. Breed, M. W. Bruford, M. A. Coleman, R. Ekblom, W. C. Funk, C. E. Grueber, and B. K. Hand. (2022). "Global Genetic Diversity Status and Trends: Towards a Suite of Essential Biodiversity Variables (EBVs) for Genetic Composition". *Biological Reviews*. **97** (4): 1511-1538. [10.1111/brv.12852](https://doi.org/10.1111/brv.12852).
- [58] C. Mejia, M. Wu, Y. Zhang, and Y. Kajikawa. (2021). "Exploring Topics in Bibliometric Research Through Citation Networks and Semantic Analysis". *Frontiers*

- in *Research Metrics and Analytics*. **6** : 742311. [10.3389/frma.2021.742311](https://doi.org/10.3389/frma.2021.742311).
- [59] P. Casadei, M. Bloom, R. Camerani, M. Masucci, J. Siepel, and J. V. Ospina. (2023). "Mapping the State of the Art of Creative Cluster Research: A Bibliometric and Thematic Analysis". *European Planning Studies*. **31** (12): 2531-2551. [10.1080/09654313.2022.2158722](https://doi.org/10.1080/09654313.2022.2158722).
- [60] S. Radhakrishnan, S. Erbis, J. A. Isaacs, and S. Kamarthi. (2017). "Novel keyword co-occurrence network-based methods to foster systematic reviews of scientific literature". *PLoS ONE*. **12** (3): e0172778. [10.1371/journal.pone.0172778](https://doi.org/10.1371/journal.pone.0172778).
- [61] A. Tontodimamma, E. Nissi, A. Sarra, and L. Fontanella. (2021). "Thirty Years of Research into Hate Speech: Topics of Interest and Their Evolution". *Scientometrics*. **126** (1): 157-179. [10.1007/s11192-020-03737-6](https://doi.org/10.1007/s11192-020-03737-6).
- [62] M. E. Hunter, S. M. Hoban, M. W. Bruford, G. Segelbacher, and L. Bernatchez. (2018). "Next-Generation Conservation Genetics and Biodiversity Monitoring". *Evolutionary Applications*. **11** (7): 1029-1034. [10.1111/eva.12661](https://doi.org/10.1111/eva.12661).
- [63] Y. Zhou and Z. Zhou. (2025). "Unlocking Rice's Genetic Potential: Big Data-Driven Insights from Population Genomics". *Genomics Communications*. **2** (1). [10.48130/gcomm-0025-0012](https://doi.org/10.48130/gcomm-0025-0012).
- [64] S. Sanokulova, F. Matyakubova, O. Saydayev, F. Azizova, A. Sidikov, and R. Sultanov. (2026). "Teaching Gene Discovery and Population Genomics Through Authentic Analysis of Large-Scale Sequencing Datasets". *Genetics and Molecular Research*. **25** (1): 1-3. [10.4238/fvxcrh73](https://doi.org/10.4238/fvxcrh73).
- [65] M. Saadat. (2024). "Evaluation of Hardy-Weinberg Equilibrium in Genetic Association Studies". *International Journal of Immunogenetics*. **51** (3): 183-186. [10.1111/iji.12668](https://doi.org/10.1111/iji.12668).
- [66] J. Nishino, F. Miya, and M. Kato. (2024). "Gene-Based Hardy-Weinberg Equilibrium Test Using Genotype Count Data Identifies Novel Cancer-Related Genes". *medRxiv*. [10.1101/2024.03.20.24304529](https://doi.org/10.1101/2024.03.20.24304529).
- [67] F. Prosdocimi, M. Garbin, and F. Dondero. (2026). "From Natural Theology to the Extended Synthesis: Historical Milestones and Conceptual Expansions in Evolutionary Biology". *Genetics and Molecular Biology*. **49** (Suppl 2): e20250179. [10.1590/1678-4685-GMB-2025-0179](https://doi.org/10.1590/1678-4685-GMB-2025-0179).
- [68] S. Wang, Y. Chen, X. Lv, and J. Xu. (2023). "Hot Topics and Frontier Evolution of Science Education Research: A Bibliometric Mapping from 2001 to 2020". *Science & Education*. **32** (3): 845-869. [10.1007/s11191-022-00337-z](https://doi.org/10.1007/s11191-022-00337-z).
- [69] A. Ayaz. (2026). "The 500 Most-Cited Articles in Management Information Systems: A Bibliometric and Thematic Approach". *Information Systems Frontiers*. **28** : 317-340. [10.1007/s10796-025-10687-y](https://doi.org/10.1007/s10796-025-10687-y).