HIV-1 genetic diversity a challenge for AIDS vaccine development: a retrospective bibliometric analysis

Abdur Rashid, Kang Li, Yi Feng, Tauseef Ahmad, Yimam Getaneh, Yueyang Yu, Xiaoyan Hu, Syed Hani Abidi, and Yiming Shao

School of Medicine, Nankai University, Tianjin, China; State Key Laboratory for Infectious Disease Prevention and Control, National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China; Key Laboratory of Molecular Microbiology and Technology, Ministry of Education, College of Life Sciences, Nankai University, Tianjin, China; Department of Epidemiology and Health Statistics, School of Public Health, Southeast University, Nanjing, China; Department of Biological and Biomedical Sciences, Aga Khan University, Karachi, Pakistan

ABSTRACT

Background: Despite recent advances in human immunodeficiency virus-1 (HIV-1) prevention, a fast, safe, and effective vaccine will probably be necessary to end the HIV/AIDS pandemic. This study was conducted to evaluate global research trends and map the key bibliometric indices in HIV-1 genetic diversity from 1998 to 2021.

Methods: A comprehensive online search was conducted in the Web of Science Core Collection database to retrieve published literature on HIV-1 genetic diversity. Key bibliometric indicators were calculated and evaluated using HistCiteTM, Bibliometric: An R-tool, and VOSviewer software for windows.

Results: A total of 2,060 documents written by 9,201 authors and published in 250 journals were included in the final analysis. Year 2012 was the most productive year with 121 (5.87%) publications. The most prolific author was Shao Yiming (n = 74, 3.59%) from Chinese Center for Disease Control and Prevention. The United States of America was the highly contributing and influential country (n = 681, 33.05%). AIDS Research and Human Retroviruses was the most productive journal (n = 562, 27.2%). Network visualization shows that HIV-1 was the most widely used author keyword.

Conclusion: This study provides global research trends and detailed information on HIV-1 genetic diversity. The amount of scientific literature on HIV-1 genetic diversity research has rapidly increased in the last two decades. The maximum number of articles on HIV-1 genetic diversity was published in developed countries; therefore, a scientific research collaboration among researchers and institutes in low-income countries should be promoted and supported.

Introduction

In the last few decades, human immunodeficiency virus-1 (HIV-1) seriously threatened global public health and disease surveillance systems, particularly in low-income countries. An estimated 37.6 million (30.2–45.0 million) people infected globally were living with HIV-1 in 2020, and 1.5 million (1.1–2.1 million) people were newly infected. Despite the availability of HIV-1 antiretroviral drugs, approximately 690,000 (480,000–1 million) people died from acquired immunodeficiency syndrome (AIDS) related illnesses in 2020. HIV-1, like other human RNA viruses, has a high genetic variability that emanates from its high mutation rates, retroviral recombination, and high viral turnover. HIV-1 has been divided into four genetically distinct groups: M, O, N, and P. The vast majority of infections worldwide are caused by group M lineages, which have further diversified into 10 distinct subtypes A, B, C, D, F, G, H, J, K, and L. Genetic distance within a subtype can be 15 to 20%, whereas genetic distances between subtypes are normally 25 to 35%. Recombination between group M subtypes has led to the generation of circulating recombinant forms (CRFs) and unique recombinant forms (URFs). Till now, 118 distinct CRFs have been identified globally, and these CRFs responsible for 16.7% of all global HIV infections.

The global HIV-1 genetic diversity is a major stumbling block to the successful design of an effective HIV/AIDS vaccine, as well as antiretroviral therapy, diagnostic and viral load assays. Information about country-level HIV-1 diversity help in the successful design of subtype-specific vaccines that can benefit specific countries and regions.

Bibliometric is the quantitative analysis of scholarly publications concerned about one special topic by statistical and mathematical methods. Bibliometric analysis is increasingly used when studying various aspects of science and how worldwide universities and institutions are ranked. In recent years, the quantity of bibliometric studies in various fields has significantly increased. Several bibliometric studies have been published on HIV/AIDS. However, to the best of our knowledge, no bibliometric analysis on HIV-1 genetic diversity research has been published to date. Bibliometric analysis of...
the high number of research publications related to HIV-1 genetic diversity is not only useful providing crucial information for epidemiological assessments and public health interventions, but also to determine future research direction, and new research frontiers. Therefore, in this study we evaluated the global research output and mapped key bibliometric indices in HIV-1 genetic diversity research from 1998 to 2021.

**Methods**

**Study design**
The study used a retrospective bibliometric design.

**Searching database**
On May 15, 2021, the online search was conducted in the Web of Science Core Collection (WoSCC) database hosted by Clarivate Analytics to retrieve scientific publications on HIV-1 genetic diversity from 1998 to May 15, 2021. Web of Science is a comprehensive digital repository of published literature used globally.19

**Search strategy and data collection**
Two researchers (Abdur Rashid and Kang Li) independently conducted the online search, any disagreements related to inclusion and exclusion were resolved with the help of another researcher (Tauseef Ahmad). The search criteria for the included publications were the following: Topic: (HIV-1 genetic diversity), or Topic: (HIV-1 subtype), or Topic: (HIV-1 CRF), or Topic: (HIV-1 recombinant form), or Topic: (HIV-1 unique recombinant form), or Topic: (HIV-1 second-generation recombinant form), or Topic: (Novel HIV Type 1 recombinant form), or Topic: (HIV-1 genetic diversity and vaccine), or Topic: (HIV-1 genetic diversity impact on vaccine). The search was limited to English language and time span (from 1998 to May 15, 2021). To ensure data accuracy, the titles and abstracts of the published documents were screened carefully. The retrieved data were downloaded in plaintext and CSV format. The following data were extracted from the retrieved documents; title, author name, year of publication, journal, institution, country, and citations count. All the data were retrieved from WoSCC on May 15, 2021, to avoid bias in data collection because the WoSCC database is updated daily.

**Data analysis**
The obtained data were exported into Microsoft Excel 2019 and the values were presented in frequencies and percentages. The required graphs were generated using GraphPad Prism 8.0 software.

**Bibliometric indicators analyses**
Various bibliometric indicators, including the number of publications per year, most prolific authors, top journals, most active institutions, countries, frequently used author keywords, and highly cited articles were calculated and evaluated to assess the impact of the published articles using HistCite™ (http://www.histcite.com/).20 Furthermore, Bibliometrix: An R-tool was used to construct corresponding authors’ countries and wordcloud analysis, and calculate mean total citations per year.21 In addition, the obtained data were plotted for coauthorship countries, author keywords, and cited sources network visualization using VOSviewer software version 1.6.16 for windows (http://www.vosviewer.com).22

**Ethical consideration**
No ethical approval was required.

**Results**

**The annual quantity of publications**
In this study, a total of 2,060 documents published in English language on HIV-1 genetic diversity were analyzed. Majority of the documents were published in the form of original articles (n = 1831, 88.89%) and reviews (n = 98, 4.76%) as shown in Table 1. The changing trend of the annual quantity of publications between 1998 and 2021 is presented in Figure 1. After 1999, growing research productivity was observed, where the highest number of documents (n = 121, 5.87%) were published in 2012, followed by year 2019 (n = 120, 5.82%), while only (n = 22, 1.06%) documents were published in 1998. In the first few months of 2021, (n = 27, 1.31%) documents have been published, which are predicted to rise more in the coming months of the year 2021.

**Most active authors**
A total of 9,201 authors have been involved in HIV-1 genetic diversity publications. In total, only 184 authors had published at least 10 articles each. Based on number of publications, Shao YM was the most prolific author (n = 74, 3.59%), followed by

| Table 1. Main information about the retrieved data on HIV-1 genetic diversity from WoSCC database. |
|--------------------------------------------------|---------------------------|
| **Description**                               | **Results**               |
| Time span                                      | 1998–2021                |
| Countries                                      | 125                      |
| Journals                                       | 263                      |
| Documents                                      | 2060                     |
| References                                     | 27556                    |
| Document type                                  |                          |
| Articles                                       | 1831                     |
| Reviews                                        | 98                       |
| Proceedings paper                              | 43                       |
| Meetings abstract                              | 36                       |
| Letters                                        | 26                       |
| Editorial materials                            | 11                      |
| Early access                                   | 10                      |
| Corrections                                    | 3                       |
| Data papers                                    | 1                       |
| Book chapters                                  | 1                       |
| Authors collaboration                          |                          |
| Single-authored documents                      | 31                       |
| Documents per Author                           | 0.224                    |
| Authors per Document                           | 4.47                     |
| Coauthors per Documents                        | 8.91                     |
| Collaboration Index                            | 4.52                     |
Peeters M (n = 60, 2.91%), and Feng Y (n = 59, 2.86%) as presented in Table 2. Furthermore, the top 10 authors’ productions over the time as shown in Figure 2.

**Most productive journals**

The retrieved documents (n = 2,060) were published in 250 journals. In total, only 30 journals published at least 10 articles on HIV-1 genetic diversity. Majority of documents were published in *AIDS Research and Human Retroviruses* (n = 562, 27.2%), followed by *PLoS ONE* (n = 156, 7.57%), and *AIDS* (n = 108, 5.2%). The top 10 journals published 61% of the analyzed documents as shown in Table 3. Among the top 10 most productive journals, the *Journal of Clinical Microbiology* holds the highest IF 5.345, a Q1 category journal published from the USA. An annual number of publications published by the top 10 listed journals as shown in Figure 3.

**Most active institutions**

A total of 2,821 institutions in 125 countries contributed to the articles, and amongst them, only 131 institutions published at least 10 articles on HIV-1 genetic diversity. Institutes that published 30 or more documents have been considered as the most active institutions. The top 10 active institutions, in terms of publishing original articles on HIV-1 genetic diversity, were the Chinese Center for Disease Control and Prevention (n = 78, 3.78%), followed by the Center for Disease Control and Prevention, USA (n = 76, 3.68%), and Walter Reed Army Institute of Research, USA (n = 72, 3.5%) as shown in Table 4.

**Most active corresponding authors’ countries**

A total of 125 countries were found to be involved in HIV-1 genetic diversity research; among them the most productive countries were USA (NP = 681, 33.0%), followed by China (NP = 329, 15.9%), Brazil (NP = 213, 10.3%), and UK (NP = 177, 8.5%). In addition, the USA was the most collaborative country with publications having authors from multiple countries publications (NP = 284), followed by China (NP = 65), and Brazil (NP = 44), as shown in Table 5. The annual distribution of HIV-1 genetic diversity research and the contributions of different countries are presented in Figure 4. The mutual collaboration between the most active corresponding authors’ countries contributed to HIV-1 genetic diversity research is presented in Figure 4.

**Top 20 highly cited publications**

The top 20 highly cited publications on HIV-1 genetic diversity are presented in Table 6. Among the top 20 publications, 17 were research articles, and 3 were review articles.

---

Table 2. Top 10 most prolific authors ranked by total number of publications.

| Rank | Author    | NP  | Country        | Affiliation                          | PY_start | h_index | TC   | %    |
|------|-----------|-----|----------------|--------------------------------------|----------|---------|------|------|
| 1    | Shao YM   | 74  | China          | CDC, China                           | 2000     | 19      | 1270 | 3.59 |
| 2    | Peeters M | 60  | France         | University of Montpellier            | 1998     | 30      | 3053 | 2.91 |
| 3    | Feng Y    | 59  | China          | CDC, China                           | 2004     | 13      | 570  | 2.86 |
| 4    | McCutchan FE | 52 | USA            | Henry M. Jackson Foundation         | 1998     | 31      | 3040 | 2.52 |
| 5    | Xing H    | 49  | China          | CDC, China                           | 2004     | 16      | 722  | 2.37 |
| 6    | Carr JK   | 48  | USA            | School of Medicine University of Maryland | 1998 | 26      | 2266 | 2.33 |
| 7    | Li JY     | 47  | China          | Beijing Institute of Microbiology and Epidemiology | 2008 | 13      | 473  | 2.28 |
| 8    | Bello G   | 45  | Brazil         | Oswaldo Cruz Institute               | 2006     | 16      | 608  | 2.18 |
| 9    | Morgado MG | 45 | Brazil         | Oswaldo Cruz Institute               | 2000     | 18      | 962  | 2.18 |
| 10   | Delaporte E | 43 | France         | University of Montpellier            | 1998     | 24      | 1959 | 2.08 |

NP: Number of Publications, TC: Total Citations, PY: Publication Year.
articles. The most highly cited document was a research article “Human immunodeficiency virus type 1 env clones from acute and early subtype B infections for standardized assessments of vaccine-elicited neutralizing antibodies” published in *Journal of Virology* by Li M et al., 2005 cited 870 times, followed by “Timing the ancestor of the HIV-1 pandemic strains” published in *Science* by Korber B et al., 2000 cited 622 times. The top 20 highly cited publications on HIV-1 genetic diversity had citations ranging from 195 to 870.

**Figure 2.** Top 10 most prolific authors’ in HIV-1 genetic diversity.

**Table 3.** Top 10 journals in HIV-1 genetic diversity research.

| Rank | Journal Name                                      | NP | h_index | TC  | 2020 IF  | 5 Year IF | JQ   | Publisher                  | PC |
|------|---------------------------------------------------|----|---------|-----|----------|-----------|------|---------------------------|----|
| 1    | AIDS Research and Human Retroviruses              | 562| 40      | 8477| 27.28    | 2.205     | Q4   | Mary Ann Liebert            | USA|
| 2    | PLoS ONE                                          | 156| 23      | 2115| 7.57     | 3.24      | Q2   | Open Access Scholarly Publisher | USA|
| 3    | AIDS                                              | 108| 41      | 5352| 5.24     | 4.177     | Q3   | Wolters Kluwer              | USA|
| 4    | Journal of Virology                               | 91 | 46      | 6946| 4.41     | 5.103     | Q1   | American Society for Microbiology | USA|
| 5    | JAIDS-Journal of Acquired Immune Deficiency Syndromes | 76 | 30      | 2160| 3.68     | 3.731     | Q3   | Lippincott                 | USA|
| 6    | Journal of Medical Virology                       | 69 | 19      | 1115| 3.34     | 2.327     | Q4   | Wiley-Blackwell             | USA|
| 7    | Infection Genetics and Evolution                  | 60 | 15      | 932 | 2.91     | 3.342     | Q3   | Elsevier                   | Netherlands|
| 8    | Retrovirology                                     | 48 | 18      | 1080| 2.33     | 4.602     | Q2   | BioMed Central              | England|
| 9    | Journal of Clinical Microbiology                  | 41 | 25      | 1654| 1.99     | 5.948     | Q1   | American Society for Microbiology | USA|
| 10   | Virology                                          | 38 | 19      | 1170| 1.84     | 3.616     | Q3   | Elsevier                   | USA|

NP: Number of Publications. TC: Total Citations. IF: Impact Factor. JQ: Journal Quartile. PC: Publisher Country.

**Coauthorship countries network visualization**

Growing evidence shows that a scientific collaboration network between countries helps admire the countries with potential in the relevant research area. The coauthorship countries analysis reflects the degree of collaboration among researchers from different countries based on total link strength. The minimum number of country publications was set at 10, and the number of coauthorship countries per/publication was set at 25. Among the analyzed countries, only 50 countries met the
thresholds and formed 9 clusters. The size of each node represents the number of published documents by a country. The USA was the leading country with (TLS = 524), followed by China (TLS = 93), Brazil (TLS = 75), France (TLS = 119), and England (TLS = 141). The USA and China are the most active countries, and both have a high frequency of international collaboration. The thickness of each link represents the strength of collaboration between two countries. The distance between two nodes represents the relatedness of the links they each have. The network visualization map of coauthorship countries is shown in Figure 5.

**Author keywords network visualization**

Before submitting a manuscript to a journal, the author is required to write keywords that reflect the main ideas presented in the article. The keywords analysis provides a medium for discovering key research areas. In this context, occurrence shows the number of articles where a keyword occurs, while co-occurrence means how often a keyword occurs with other

### Table 4. Top 10 most productive institutions ranked by number of publications.

| Rank | Institution | Country | NP | TC | % |
|------|-------------|---------|----|----|---|
| 1    | CDC, China  | China   | 78 | 988| 3.78|
| 2    | CDC, USA    | USA     | 76 | 2466| 3.68|
| 3    | Walter Reed Army Institute of Research | USA | 72 | 2799| 3.49|
| 4    | University of Oxford | UK | 68 | 3457| 3.30|
| 5    | Instituto de Salud Carlos III | Spain | 60 | 1724| 2.91|
| 6    | Henry M. Jackson Foundation | USA | 54 | 4073| 2.61|
| 7    | National Institute of Infectious Diseases | USA | 52 | 1299| 2.52|
| 8    | Oswaldo Cruz Institute | Brazil | 50 | 864| 2.42|
| 9    | Harvard University | USA | 50 | 1505| 2.42|
| 10   | John Hopkins University | USA | 49 | 1874| 2.37|

NP: Number of Publications, CDC: Center for Disease Control and Prevention, TC: Total Citations.

### Table 5. Top 20 prolific corresponding authors countries and their collaborative publications.

| Country            | Total Articles | Total citations | Average Article Citations | Single Country Publications | Multiple Country Publications | Multiple Country Publications Ratio |
|--------------------|----------------|-----------------|---------------------------|-------------------------------|-----------------------------|-----------------------------------|
| USA                | 486            | 19,187          | 39.48                     | 202                          | 284                         | 0.584                             |
| CHINA              | 323            | 3,161           | 9.79                      | 258                          | 65                          | 0.201                             |
| BRAZIL             | 181            | 2,894           | 15.99                     | 137                          | 44                          | 0.243                             |
| FRANCE             | 115            | 2,991           | 26.01                     | 52                           | 63                          | 0.548                             |
| SPAIN              | 94             | 2,168           | 23.06                     | 60                           | 34                          | 0.362                             |
| ITALY              | 66             | 964             | 14.61                     | 47                           | 19                          | 0.288                             |
| JAPAN              | 62             | 1,332           | 21.48                     | 18                           | 44                          | 0.71                              |
| UNITED             | 60             | 2,597           | 43.28                     | 29                           | 31                          | 0.517                             |
| KINGDOM            | 54             | 892             | 16.52                     | 18                           | 36                          | 0.667                             |
| SOUTH AFRICA       | 48             | 1,153           | 24.02                     | 28                           | 20                          | 0.417                             |
| INDIA              | 43             | 529             | 12.3                      | 35                           | 8                           | 0.186                             |
| BELGIUM            | 36             | 1,286           | 35.72                     | 8                            | 28                          | 0.778                             |
| AUSTRALIA          | 28             | 465             | 16.61                     | 17                           | 11                          | 0.393                             |
| GERMANY            | 26             | 454             | 17.46                     | 4                            | 22                          | 0.846                             |
| THAILAND           | 26             | 707             | 27.19                     | 10                           | 16                          | 0.615                             |
| ARGENTINA          | 25             | 325             | 13                        | 11                           | 14                          | 0.56                              |
| SWEDEN             | 25             | 659             | 26.36                     | 11                           | 14                          | 0.56                              |
| PORTUGAL           | 24             | 542             | 22.58                     | 10                           | 14                          | 0.583                             |
| RUSSIA             | 23             | 296             | 12.87                     | 16                           | 7                           | 0.304                             |
| NETHERLANDS        | 19             | 401             | 21.11                     | 9                            | 10                          | 0.526                             |
Keywords. Keywords with a strong co-occurrence relationship can divulge research hotspots more accurately than a single keyword. A minimum occurrence of author keywords was set at 10. Among the analyzed keywords, only 62 met the threshold and formed 8 clusters. HIV-1 was the most frequently used author keyword (TLS = 362), followed by HIV (TLS = 160), molecular epidemiology (TLS = 97), subtypes (TLS = 80), and vaccine (TLS = 52). Keywords with the greatest total link strength as shown in Figure 6. The link strength between two keywords is premised on the number of documents both keywords mutually occur, exploring the specific research spot’s relationship. The link’s thickness is proportional to the strength of the link in the network. In our visualization analysis, closely connected keywords were colored the same, indicating that they might share something in common. The orange color keywords nodes are related to vaccine research (envelope, evolution, gp120, and diversity). The keywords in red and blue nodes are mainly related to epidemiology. The nodes in purple are related to virology (genotype, protease, mutation, drug resistance, antiretroviral therapy, reverse transcriptase, and AIDS). Wordcloud visualization of frequently used authors’ keywords in the WoSCC database is presented in Figure 6.
Table 6. Top 20 highly cited publications on HIV-1 genetic diversity.

| Title                                                                 | Journal                        | Document type | Date of Publication | Corresponding author’s Country | Total Citations (TC) | Average TC per Year | DOI                                      |
|----------------------------------------------------------------------|--------------------------------|----------------|---------------------|-------------------------------|----------------------|---------------------|------------------------------------------|
| Human immunodeficiency virus type 1 envelope clones from acute and early subtype B infections for standardized assessments of vaccine-elicited neutralizing antibodies | *Journal of Virology*        | Article         | AUG 2005            | USA                           | 870                  | 51.1765             | 10.1128/JVI.79.16.10108–10125.2005       |
| Timing the ancestor of the HIV-1 pandemic strains                     | *Science*                     | Article         | JUN 2000            | USA                           | 622                  | 28.2727             | 10.1126/science.288.5472.1789            |
| Global trends in molecular epidemiology of HIV-1 during 2000–2007    | *AIDS*                         | Article         | MAR 2011            | England                       | 520                  | 47.2727             | 10.1097/QAD.Ob013e328342ff9            |
| Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing | *Journal of Virology*        | Article         | APR 2008            | USA                           | 448                  | 32                  | 10.1128/JVI.02660–07                     |
| Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004 | *AIDS*                         | Article         | OCT 2006            | Switzerland                   | 415                  | 25.9375             | 10.1097/01                               |
| Human Immunodeficiency Virus Type 1 Elite Neutralizers: Individuals with Broad and Potent Neutralizing Activity Identified by Using a High-Throughput Neutralization Assay together with an Analytical Selection Algorithm | *Journal of Virology*        | Article         | JUL 2009            | USA                           | 397                  | 30.5385             | 10.1128/JVI.00110–09                     |
| Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960   | *Nature*                      | Article         | OCT 2008            | USA                           | 327                  | 23.3571             | 10.1038/nature07390                    |
| The early spread and epidemic ignition of HIV-1 in human populations   | *Science*                     | Article         | OCT 2014            | England                       | 288                  | 36                  | 10.1126/science.1256739                  |
| Understanding the genetic diversity of HIV-1 and implications for the origin of the epidemic | *AIDS*                         | Review          | 2000                | USA                           | 285                  | 12.9545             | NA                                        |
| An African HIV-1 sequence from 1959 and implications for the origin of the epidemic | *Nature*                      | Article         | FEB 1998            | USA                           | 268                  | 11.1667             | 10.1038/35400                           |
| Estimated global distribution and regional spread of HIV-1 genetic subtypes in the year 2000 | *JAIDS*                        | Article         | FEB 2002            | Switzerland                   | 259                  | 12.95               | NA                                        |
| A recent outbreak of human immunodeficiency virus type 1 infection in southern China was initiated by two highly homogeneous, geographically separated strains, circulating recombinant form AE and a novel BC recombinant | *Journal of Virology*        | Article         | DEC 2000            | USA                           | 250                  | 11.3636             | 10.1128/JVI.74.23.11286–11295.2000       |
| Accurate sampling and deep sequencing of the HIV-1 protease gene using a Primer ID | *Proceedings of the national academy of sciences of the United States of America* | Article | DEC 2011            | USA                           | 246                  | 22.3363             | 10.1073/pnas.1110064108                  |
| Antigenic conservation and immunogenicity                              | *Journal of Experimental medicine* | Article       | May 2005            | USA                           | 241                  | 14.1765             | 10.1084/jem.20042510                    |
| Inflammatory Genital Infections Mitigate a Severe Genetic Bottleneck in Heterosexual Transmission of Subtypes A and C HIV-1 | *PLOS PATHOGENS*              | Article         | JAN 2009            | USA                           | 237                  | 18.2308             | 10.1371/journal.ppat.10000274           |
| A comprehensive panel of near-full-length clones and reference sequences for non-subtype B isolates of human immunodeficiency virus type 1 | *Journal of Virology*        | Article         | JUL 1998            | USA                           | 237                  | 9.875               | 10.1128/JVI.72.7.5680–5698.1998         |
| Unprecedented degree of human immunodeficiency virus type 1 (HIV-1) group M genetic diversity in the Democratic Republic of Congo suggests that the HIV-1 pandemic originated in Central Africa | *Journal of Virology*        | Article         | NOV 2000            | France                        | 230                  | 10.4545             | 10.1128/JVI.74.22.10498–10507.2000       |
| Minireview- Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: Pathogenic and therapeutic implications | *Journal of Virology*        | Review          | OCT 2007            | Italy                         | 217                  | 14.4667             | 10.1128/JVI.00872–07                     |
| Selection for human immunodeficiency virus type 1 envelope glycosylation variants with shorter V1-V2 loop sequences occurs during transmission of certain genetic subtypes and impact viral RNA levels | *Journal of Virology*        | Article         | May 2005            | USA                           | 206                  | 12.1176             | 10.1128/JVI.79.10.6528–6531.2005         |
| Genetic diversity of HIV-1: the moving target                           | *AIDS*                         | Review          | 2000                | France                        | 195                  | 8.8636              | NA                                        |
Republic Australia, Greece, India, Israel, Italy, Luxembourg, Mexico, Netherlands, Romania, South Korea, Sweden; Cluster 2 (Countries: 8: Burkina Faso, Cameroon, Cote Ivoire, Democratic Republic of Congo, France, Gabon, Republic of Congo, Senegal); Cluster 3 (Countries: 6: England, Pakistan, Scotland, Uganda, USA, Zambia); Cluster 4 (Countries: 6: Australia, Japan, Malaysia, China, Taiwan, Thailand); Cluster 5 (Countries: 5: Angola, Argentina, Brazil, Peru, Portugal); Cluster 6 (Countries: 4: Cuba, Russia, Spain, Switzerland); Cluster 7 (Countries: 3: Canada, Iran, Kenya); Cluster 8 (Countries: 3: Germany, South Africa, Tanzania); Cluster 9 (Countries: 1: Nigeria). Correct the label number... it is Figure 5 not 3

Cited sources network visualization

Many studies give attention to the importance of a journal in any scientific field. Such analysis helps an author publish their work in a well-reputed journal, and the reader can easily find the best source of information relevant to their work. The obtained data were plotted for cited sources network visualization is shown in Figure 7. The minimum number of citations of a source was set at 30. Among the analyzed sources, only 178 sources met the threshold and formed 4 clusters. AIDS Research and Human Retroviruses journal was the leading journal with (TLS = 7523.8) followed by Journal of Virology (TLS = 7100.08), AIDS (TLS = 6166.73), and Science (TLS = 2302.88).

Discussion

HIV-1 unusual degree of genetic variability makes the virus highly heterogeneous that has hampered the HIV/AIDS vaccine development. This foundation of high genetic variability is the consequence of high replication rates, error-prone reverse transcriptase enzyme, and retroviral recombination that shuttle the mutations between recombinant genomes. In the second half of the 20th century, the worldwide transmission of HIV-1 preceded the distinctive distribution of HIV-1 subtypes and recombinants. This bibliometric analysis was a solely descriptive analysis of global publications in HIV-1 genetic diversity. Significant research progress has been observed on HIV-1 genetic diversity. The quantity of publications has been steadily increasing over the 24 years study period, particularly since 2000. According to the retrieved documents from WoSCC, most of the research was regarding the epidemiology of HIV-1. Research concerned about the impact of HIV-1 genetic diversity on vaccine development or the consideration of genetic diversity in HIV-1 vaccine development is comparatively less in number since virologists paid less assiduity to this research area. According to the retrieved data, the highly cited document was a research article published by the researcher from Duke University, USA. They focus on the priority of broadly neutralizing antibodies for HIV vaccine development. A holy grail of the HIV-1 vaccine field is to elicit broadly neutralizing antibodies. All vaccines try to induce antibodies against the native envelope protein to prevent the virus from infecting host cells, but the HIV-1 envelope is much more devious. Most immunogens were used; none of them give rise to broadly neutralizing antibodies Abs due to lack of consistency in target strains used by scientists. For wide distribution, an extreme need to set up HIV reference strains drawn from a mosaic of HIV-1 variants in order to protect against divergent HIV-1 subtypes and recombinants. The second most cited document was an article published by Los Alamos National Laboratory of USA concern about the time estimation of the ancestor sequence of the HIV-1 M group responsible for the HIV pandemic. They estimated HIV-1’s last common ancestor date to be 1931 (1915–1941).

A total of 125 countries were involved in research on HIV-1 genetic diversity, but most research contributions were from developed countries such as the USA, China, Brazil, and the UK. The USA maintained its leading role in single and multiple countries’ research contributions, followed by China, Brazil, and France, as is also evident from many bibliometric studies in other field. In the most productive country list, South Africa on the eighth position, which shows their researchers are fostering international collaboration links in HIV/AIDS research, particularly with the USA and UK. The disparities between developed and undeveloped countries can be attributed to treatment and prevention coverage, economic stability,
adequate research funding and high international collaboration. *AIDS Research and Human Retroviruses* was the most prominent and productive journal contributing to HIV-1 genetic diversity research, followed by the *Journal of Virology* with total citations of 8,477 and 6,946. These top journals are very responsive to new emerging subtypes or recombinants of HIV-1 and with coverage ranging from the molecular basis of the virus to clinical trials on vaccines. All the top 10 journals’ contributions to HIV-1 genetic diversity are stunning and promoted the effective dissemination of pedantic evidence of contend the variability of HIV-1 in developing an HIV-1 vaccine. Shao YM was the most active researcher in the field of HIV-1 genetic diversity research. His research laboratory identified different HIV-1 circulating recombinant forms CRFs among HIV-1 infected Chinese patients and submitted to Los Alamos National Laboratory CRF62_BC,

CRF64_BC, CRF65_cpx, CRF79_0107, CRF106_cpx. Among the most active authors who published more articles on HIV-1 genetic diversity came from the Chinese CDC, French National Research Institute for Sustainable Development (IRD), or Henry M. Jackson Foundation. Most of these authors have collaborated in publishing articles on HIV-1 genetic diversity and belong to the same institutions. The above statement highlights that research collaboration among researchers from different countries needs to be promoted. It is interesting to note that 39.4% of the publications on HIV-1 genetic diversity were collaborative publications. Like other research areas, the USA was the leading country in HIV-1 genetic research collaboration; the collaborative publications between the USA and China were 3.4%. Compared to developed countries, the collaboration between developing countries was negligible. Chinese institutions research is largely conducted in the area of HIV-1 epidemiology. The majority of highly cited articles research was about HIV-1 vaccine published by the institutions in USA. The country-wise publication pattern indicates that research on HIV-1 diversity in a developing country is very low highlights that research must be supported in developing countries to identify the most common HIV-1 strains circulating that can help in the development of a therapeutic vaccine against...
prevalent strains in each country. The current study might be helpful for researchers to understand the research achievements and trends in HIV-1 genetic diversity. The promotion and sharing of scientific knowledge and medical technologies with developing countries are extremely necessary, which can help to curb the HIV/AIDS in the coming future.

**Limitations**

The potential limitation of our study lies that the publications related to HIV-1 genetic diversity are only extracted from the WoSCC database; databases like Scopus, PubMed, Google Scholar and Chinese databases were not searched; therefore, the results may be different using other databases. We also evaluated the top cited publications in HIV-1 genetic diversity based on the total citation score. Although, authors have self-citations that can influence the overall number of h-index and citations count.

**Conclusion**

This study provides global research trends and detailed information on HIV-1 genetic diversity for the first time. The amount of scientific literature on HIV-1 genetic diversity research has rapidly increased in the last two decades. The USA was the leading country in term of publications. Therefore, scientific research in low-income countries should be promoted and supported. The development of an effective HIV-1 vaccine remains a major goal of AIDS research.

**Acknowledgments**

The authors acknowledge Nankai University and Chinese Center for Disease Control and Prevention CDC for providing online access to the Clarivate Analytics Web of Science Core Collection (WoSCC) database.

**Disclosure statement**

No potential conflict of interest was reported by the author(s).

**Funding**

This work was supported by the National Science and Technology Major Project of China [grant number: 2018ZX10731-101], Science Priority [Grant 2019SKLID602] from the State Key Laboratory of Infections Disease Prevention and Control, National Natural Science Foundation International/Inter-Organization Cooperation and Exchange Study-NSFC-VR Project (China and Switzerland) [grant number 81861138011].

**ORCID**

Tauseef Ahmad http://orcid.org/0000-0001-8793-273X

**Authors’ contributions**

All the authors potentially contributed and met the authorship criteria.

**References**

1. Global HIV & AIDS Statistics – 2021 fact sheet. 2021 [accessed 2021 Oct 5]. https://www.unaids.org/en/resources/fact-sheet .
2. Dorrell L. Therapeutic immunization strategies for the control of HIV-1. Expert Rev Vaccines. 2005 Aug;4(4):513–20. doi:10.1586/ 14760584.4.4.513.
3. De Cock KM, Jaffe HW, Curran JW. Reflections on 40 years of AIDS, Emerg Infect Dis. 2021 June;27(6):1553–60. doi:10.3201/ eid2706.210284.
4. Bbosa N, Kaleebu P, Ssemwangwa D. HIV subtype diversity worldwide. Curr Opin HIV AIDS. 2019 May;14(3):153–60. doi:10.1097/COH.0000000000000534.
5. Thomson MM, Pérez-Alvarez L, Nájera R. Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy. Lancet Infect Dis. 2002 Aug;2(8):461–71. doi:10.1016/S1473-3099(02)00343-2.
