A bibliometric analysis of research updates and tendencies on steroid biotransformation

Zhaoyu Song
Key Laboratory of Industrial Fermentation Microbiology, Ministry of Education, College of Biotechnology, Tianjin University of Science & Technology, Tianjin 300457, People's Republic of China.
E-mail:13502076152@163.com

Abstract. Steroid biotransformation, as a powerful tool for generation of steroid active pharmaceutical ingredients and key intermediates, has received widespread attention with increasing market demand for steroid-based drugs. In our study, a bibliometric analysis of steroid biotransformation was performed to trace the research updates and tendencies from 1993 to 2016, based on the Science Citation Index Expanded (SCIE) database. Results showed a notable growth trend in publication outputs. Although the USA was the most productive country between 1993 and 2016, developing nations, including China and India, contributed the prominent growth in recent years (2005–2016). Steroids was the leading journal in this field, and the research outputs had notably increased in the field of ‘Chemistry’, ‘Pharmacology and Pharmacy’ and ‘Biotechnology and Applied Microbiology’. Finally, research focused mainly on the efficient production of novel steroid active pharmaceutical ingredients and key intermediates through steroid biotransformation. Furthermore, cytochrome P450 involved in the side-chain oxidation of sterols has gradually become a hotspot issue in recent years.

1. Introduction
Rapid development of medical technology in the recent years, steroid drugs are ranked among the most marketed medical products and represent the second large category next to antibiotics [1]. Compared with traditional chemical synthesis, the mild conditions required for biotransformation are important for molecules unstable at high temperatures or extremal pH values, and biotransformation are more ecologically friendly. These advantages determine a broad expansion of biocatalysis technologies into the steroid field. For nearly eight decades, numerous studies have been carried out on the various aspects of steroid biotransformation, such as raw materials for steroid production, different types of catalytic reactions and corresponding microorganisms and methods for enhancement of steroid biotransformation [2]. Meanwhile, these studies were published in diverse journals of many subject categories and were written by researchers from a number of countries all over the world. However, no systematic analysis of the scientific research on steroid biotransformation has been carried out to date. Based on this situation, a comprehensive bibliometric analysis was used to trace global trends and hotspot issues in steroid biotransformation research.

2. Data sources and methods
The data on the documents used in the present study were based on the online SCI-Expanded database, where we extracted information about scientific output on 24 October, 2017. “Stero* and (biotransform* or bio-transform* or microbial transform* or bioconvers* or bio-convers* or microbial
bioconvers* or biocataly* or bio-cataly* or microbial cataly* or catabol*)” were used as keywords to search titles, abstracts, and author keywords from 1993 to 2016. Data analysis was performed by the method described by Mao et al. [3], and Microsoft Excel 2007 software was used to deal with the relevant data.

3. Results and discussion

3.1. Document type and language of publication
From 1993 to 2016, document types and languages of publication 6,376 publications related to steroid biotransformation were identified in SCI-EXPANDED, which were categorized by 9 document types. The article was the dominant document type, accounting for 87% (5,516) of total publications. The remaining publications were reviews (669), proceedings (79), meeting abstracts (44), notes (32), editorial materials (26), letters (7), corrections (2), reprints (1).

As the dominant type of document, the languages used in articles were then analyzed. 5,411 articles (98%) were published in English, followed by German (24), French (23), Chinese (13), Spanish (9), Portuguese (8), Russian (8), Polish (5), Japanese (5), Hungarian (3), Korean (2), Czech (2), and only one article each in Italian, Turkish and Rumanian.

3.2. Characteristics of publication outputs
As shown in figure 1, although steroid biotransformation has been studied for nearly 80 years so far, but no more than 40 publications were published each year before 1990. Since 1991, the research of steroid biotransformation began to attract widespread interests and experienced a tremendous growth. In the past five years (2012-2016), the number of publications each year exceeded 300 and was about 10 times that of 1991. The possible reason for the dramatic increase is that the efficacy of steroidal drugs is further explored as medical advances, leading to increasing market demand year after year, and, compared with traditional chemical synthesis, steroid bioconversion has the mild reaction conditions, high yield and more ecological friendliness. In general, the SCI-Expanded publications were still showing a rapid growth in recent years. As a result, the number of scientific publications on steroid biotransformation is expected to continue to grow rapidly. To obtain an overview of steroids biotransformation research, the characteristics of annual articles during 1993–2016 were further analyzed (data not shown). we found that the average number of authors per article increased by 65.8% from 3.81 in 1993 to 6.32 in 2016, with a similar 18.3% increase in the number of references cited per article over the same period. The average article length increased slightly, with an maximum length of 10.92 pages per article at 2015. The progressive increase of authors, references and article length indicated a growth trend and an increasing communication in the field of steroid biotransformation during the past 24 years. Conversely, the average number of times cited per article reached a maximum of 52.78 in 1996 and then showed roughly declining trend year after year, the value dropped to 1.95 by 2016. The possible reason was that the just emerging of steroid biotransformation was relatively simple and basic research, which provided the fundamental theoretical guidance for later research, yet the latest research involved a wider range of fields and richer content.
3.3. Characteristics of journals and subject categories

Between 1993 and 2016, a total of 5,516 articles were published in 1632 journals in 104 subject categories in SCI EXPANDED. Table 1 summarized the top 6 most productive journals (TP > 50) and their IF in 2016 and h-index. In these journals, Steroids, with an IF of 2.280 in 2016 and an h-index of 33, published the most articles (189) in this field, taking 3.4% of all, followed by Journal of Steroid Biochemistry and Molecular Biology (162, 2.9%) and Journal of Biological Chemistry (126, 2.3%), the last of which has the highest h-index of all journals in this field. Subsequent 3 of journals published articles accounted for the percentage of the total articles were no more than 1.5%, as the most productive journals, the proportion of published articles was relatively low, which indicated the broad interest in biotransformation research from various fields. Furthermore, due to the assignment of journals to multiple subject categories, the changed trends of the top 6 ISI categories were shown in Figure 2. ‘Biochemistry and Molecular Biology’ was almost always in the leading position, mainly because of the ISI category of top 3 productive journals, including Steroids, Journal of Steroid Biochemistry and Molecular Biology and Journal of Biological Chemistry, belonged to ‘Biochemistry and Molecular Biology’. Moreover, ‘Chemistry’, ‘Pharmacology and Pharmacy’ and ‘Biotechnology and Applied Microbiology’ had similar trends which kept elevated momentum during 1993-2016. Conversely, the number of published articles belonging to ‘Endocrinology and Metabolism’ and ‘Cell Biology’ subject categories decreased year by year.

**Figure 1.** Trends of SCI-Expanded publications on steroid biotransformation since 1900.

**Table 1.** Top 6 productive journals (1993–2016) with the number of articles, IF and h-index.

| Journal                                                      | TP (%) | IF     | h-index |
|--------------------------------------------------------------|--------|--------|---------|
| Steroids                                                     | 189 (3.4) | 2.280  | 33      |
| Journal of Steroid Biochemistry and Molecular Biology        | 162 (2.9) | 4.561  | 40      |
| Journal of Biological Chemistry                              | 126 (2.3) | 4.125  | 65      |
| Endocrinology                                                | 83 (1.5)  | 4.286  | 50      |
| Molecular and Cellular Endocrinology                         | 59 (1.1)  | 3.754  | 29      |
| Journal of Clinical Endocrinology and Metabolism             | 59 (1.1)  | 5.455  | 37      |

TP (%): Number of articles (percentage of total articles).
IF: Impact factor.

h-index: The h of Np papers have at least h citations each and the other (Np-h) papers have \( \leq h \) citations each, and Np is the number of papers published over n years.

![Figure 2. The changed trends of the top 6 subject categories.](image)

3.4. Contribution of country or territory

The distribution of different countries and territories were analyzed according to the address of at least one author. A total of 5,516 articles with author address information published between 1993 and 2016 were analyzed, covering 105 countries or territories. 4067 (73.7 \%) were independent articles and 1449 (26.3 \%) were internationally collaborative articles. The top 10 productive countries, including two North American countries (USA, Canada), three Asian countries (China, India, Japan) and five European countries (UK, Germany, France, Italy, Spain), accounted for about 75.7 \% of the total articles. Development trends of the top 10 productive countries during 1993–2016 were displayed in figure 3, the size of each bubble indicated the number of published articles of each country in each period. In the meantime, the USA showed the most counts of article publications, followed by UK, Germany and Japan. Although the USA has been in the lead of the world in this field, China has maintained rapid growth in recent years (2005-2016). During the period 2011-2016, China has become the second largest productive country after the USA, accounting for 16.1 \% of the total published articles in this period, and, it is estimated that the number of articles published on steroid biotransformation will equal or surpass the USA in the next six years. Meanwhile, India also achieved faster growth over the past decade. The reasons for this phenomenon were that developed countries started much earlier than developing countries, and established relatively scientific research systems in the early days, but with China and India as the representative of the developing countries have enjoyed rapid development of science and technology and medicine levels in recent years, coupled with a larger population size, steroid drugs market demand is greater.

Further, we observed that international collaboration have become the dominate in the steroid biotransformation research, which is a trend shared with other fields [4]. To obtain important information about international cooperation, 19 countries with no less than 70 articles were chosen as study samples. As shown in figure 4, the size of the circle is proportional to the number of articles published by the country, and, the line thickness is proportional to the extent of cooperation from the two countries. Undoubtedly, the largest and most complicated groups was the first productive countries of USA with 14 countries, followed by the UK centered group (3) and the France centered group (3). Among them, USA-UK collaborations ranked first, with 111 cooperative articles, followed by Japan-UK (104) and China-USA (56). International collaboration is vitally important to boost the
progress of biomedical research and is the major trend in the development of science and technology in the future.

Figure 3. Development trends of top 10 productive countries.

Figure 4. Network diagram showing cooperation between countries with a minimum of 70 articles.

3.5. Research hotspots

Analyses of research hotspots can be made through finding author keywords [5]. From 1993 to 2016, 1,871 articles without author keywords information from Web of Science were excluded, the other 3,645 articles with author keywords information were analyzed, a total of 10,825 author keywords further were listed. Among them, 8,565 (79.1 %) keywords were used once and 1,165 (10.8 %) were used twice, and a lot of single-used author keywords presumably indicated more extensive differences in research aims and a lack of continuous research [6, 7]. However, only 159 (1.5 %) author keywords were used more than ten times, indicating that the mainstream research in steroid biotransformation mainly focused on a very small field [8]. Except for the searching words studied in the present work, the top 10 frequently used author keywords and the proportion of relevant articles were shown in figure 5. Most of the top 10 author keywords were related to the steroid active pharmaceutical ingredients and key intermediates, including estrogen (114, 13.3 %), progesterone (108, 12.6 %), cholesterol (104, 12.1 %), estradiol (89, 10.4 %), testosterone (87, 10.4 %), dehydroepiandrosterone (82, 9.6 %), androgen (73, 8.5 %), indicating that researchers dedicated to explore how to efficiently produce novel steroid active pharmaceutical ingredients and key intermediates through steroid biotransformation over the past years. Besides, cytochrome P450, which is widely present in Mycobacterium and Rhodococcus, catalyzes the hydroxylation of the C-26 or C-27, the first step in the side-chain oxidation of sterols [9, 10]. Cytochrome P450 has gradually become a hotspot issue in recent years (2005-2016).
Figure 5. The top 10 frequently used author keywords and the proportion of relevant articles.

References
[1] Gao J M, Shen J W, Wang J Y, Yang Z and Zhang A L. 2011 Steroids 76:43–47.
[2] Marina V D and Olga V E. 2012 Appl Microbiol Biotechnol 94:1423–1447.
[3] Mao N, Wang M H and Ho Y S. 2010 Hum Ecol Risk Assess 16:801–824.
[4] Niu B B, Hong S, Yuan J F, Peng S, Wang Z and Zhang X. 2014 Scientometrics 98: 511–529.
[5] Rosloniec K Z, Wilbrink M H, Capyk J K, Mohn W W, Ostendorf M, Dijkhuizen L, Eltis L D. 2009 Mol Microbiol 74:1031–1043.
[6] Wang L H, Wang Q H, Zhang X, Cai W W and Sun X H. 2013 J Mater Cycles Waste Manag 15: 1-8.
[7] Zheng T L, Wang J, Wang Q H, Meng H M and Wang L H. 2017 Appl Water Sci 7:13-30.
[8] Chuang K, Huang Y and Ho Y. 2007 Scientometrics 2:201–212.
[9] Capyk J K, Kalscheuer R, Stewart G R, Liu J, Kwon H, Zhao R, Okamoto S, Eltis L D and Mohn W W. 2009 J Biol Chem 284:35534–35542.
[10] Rosloniec K Z, Wilbrink M H, Capyk J K, Mohn W W, Ostendorf M, Geize R, Dijkhuizen L and Eltis L D. 2009 Mol Microbiol 74:1031–1043.