GLOBAL SCIENCE & TECHNOLOGY TRENDS REPORT

BIOLOGICS IN THERAPIES RESEARCH & DEVELOPMENT

CAS®
A DIVISION OF THE AMERICAN CHEMICAL SOCIETY

中国科学院文献情报中心
National Science Library, Chinese Academy of Sciences

PREFACE

The global landscape of scientific discovery and innovation has witnessed dramatic change during recent decades. Enhanced knowledge sharing, and economic incentives have fueled robust research and development in diverse areas of science and technology. While these changes offer exciting opportunities on many fronts, they also present unprecedented challenges. In order to stay abreast of this challenging landscape it is crucial to gather insights and current perspectives related to emerging scientific research and technological trends.

The dominant boom in global biotechnology over the past 20 years is exemplified by major breakthroughs in genetic and protein engineering, cell engineering and DNA sequencing. In particular, rapid advancements in medical biotechnology and the industrialization of related advanced biotechnologies have expedited the development of biopharmaceuticals (i.e. biologics) on an unprecedented scale. Commercialization of biologics is experiencing unprecedented returns on investment. Expanding success in the application of biologics in disease diagnosis and treatment has further stimulated innovation and investment in biotechnology.

Biologics have gained momentum and are on the forefront of biopharmaceutical innovation largely due to their robust pharmacological activity, comparatively low toxicity, negligible side effects and high target specificity. Biologics have been progressively applied to the treatment of a broad range of diseases including cancer, immune diseases, inflammation and endocrine diseases, indicative of a promising future. With positive advances and impressive breakthroughs in areas such as monoclonal antibodies, novel protein drugs, therapeutic vaccines, gene therapy, immune cell therapy and induced pluripotent stem cells produced by cell reprogramming, the biologics industry has entered a new phrase of rapid adaption and accelerated development. All major pharmaceutical companies and many biotech organizations worldwide have increased their investment in the R&D of biologics, which has enabled more drug candidates to flow into drug pipelines, enter clinical trials and gain approval for clinical use.

To understand fully the landscape and gather insights into biologics R&D, the National Science Library of the Chinese Academy of Sciences (NSLC), under the guidance of the Development Planning Bureau of the Chinese Academy of Sciences, and the CAS, a division of the American Chemical Society have conducted a collaborative and large scale data analysis in the area of biologics using advanced analysis and visualization tools. This analytic project assimilated biologics data contained in the CAS registered substance collection, which is regarded as being the industry "gold-standard" for literature-derived substance information.

This project's goal was to thoroughly investigate the overall landscape and trends in global biopharmaceutical R&D. It is our hope that the results of this study can enable and advance scientific progress and offer valuable insights to a broad class of readers including scientists, corporate decision makers, research management experts, institutional policy makers, industry practitioners, business and investment professionals as well as professors and students associated with academic and research institutions. This is the second collaboration between NSLC and CAS following the first effort on "The Global Science & Technology Trends Report: Graphene Research & Development" in 2017.

CAS headquartered in Columbus, Ohio, USA, has curated publicly disclosed publications and provided in-depth indexing of important technology discoveries in chemistry and related fields since 1907. The document database compiled by CAS, also referred as CAplus℠, covers scientific journals, books, dissertations, meeting proceedings and other related disclosures (collectively referred to as "papers") worldwide since early 19th century as well as patents from 63 patent offices, making it one of the most comprehensive document databases.

CAS's electronic workflow ensures the rapid transfer of original documents from publishers and patent offices to CAS experts, who manually review each document and create a record of index entries, using controlled terms to describe each studies' scientific emphasis and novelty.

The CAS substance database, also referred to as CAS REGISTRY℠, is the most comprehensive collection of chemicals and chemically-related substances in the world, each of which is assigned a unique CAS Registry Number®. Each registered substance record may include structural and/or descriptive information in addition to nomenclature.

Since 1963, over 200 million total substances have been added to CAS REGISTRY℠, including more than 143 million unique organic and inorganic compounds such as polymers, coordination compounds, natural products, chemical complexes, minerals, mixtures and salts. CAS REGISTRY℠ also contains more than 67 million biological sequences including naturally-occurring and synthetic nucleic acids, peptides and proteins with therapeutic implications. CAS has become the world's largest provider of integrated information for chemical and related disciplines.

CAS HAS BECOME THE WORLD'S LARGEST PROVIDER OF INTEGRATED INFORMATION FOR CHEMICAL AND RELATED DISCIPLINES.

The National Science Library, Chinese Academy of Sciences (NSLC) is the public library service system of the Chinese Academy of Sciences. NSLC, along with other information departments in the Chinese Academy of Sciences, is responsible for organizing, managing and coordinating literature information services.

It is also responsible for construction of the scientific and technological literature resource support system, and the setup and management of public document information services in natural sciences and high-tech fields for the Chinese Academy of Sciences.

NSLC aims to be an internationally recognized professional force, with 11 members that currently sit in various section standing committees of the International Federation of library Associations and Institutes (IFLA). It is also a member of Electronic Information for Libraries (EIFL), the Confederation of Open Access Repositories (COAR) and the International Council for Scientific and Technical Information (ICSTI).

NSLC collaborates with scientists, management experts, policy experts, think tank experts, and industry experts as well as other units of the document information system in the Chinese Academy of Sciences to create a competitive T-series intelligence product covering the innovation value chain for first-line decision making in scientific research and industry.

TABLE OF CONTENTS

FIGURES & TABLES

CHAPTER 1. OVERVIEW

1.1 RESEARCH BACKGROUND

Biologics (also known as biopharmaceuticals) refer to any products that are manufactured by recombinant DNA technology, or purified from organisms, tissues, living cells, or body fluids, or semi-synthesized from biological sources for disease prevention, treatment and/or diagnosis. These products primarily include macromolecules such as proteins and nucleic acids as well as vaccines and genetically engineered cells.

They specifically include vaccines, therapeutic antibodies, recombinant proteins (including fusion proteins), gene therapy, cell therapy, antimicrobial peptides, cytokines, proteinaceous hormones, enzymes, stem cells and so forth. Due to their robust pharmacological activities, low toxicity, minimal side effects, and high target specificity, biologics have been increasingly used in disease treatment. They have shown great potential for treatment and prevention of major diseases including cancer, infectious diseases, inflammation, autoimmune diseases, hereditary diseases, neurodegenerative diseases, diabetes, and cardiovascular and cerebrovascular diseases. Therefore, biologics have been recognized as the most promising category of new drugs for development in the 21st century.

Although biologics R&D has had a long history, early R&D was mostly limited to vaccines and blood products.

IT IS PREDICTED THAT BIOLOGICS WILL BECOME ONE OF THE MOST PROMISING TECHNOLOGY FIELDS WITH THE POTENTIAL FOR FAST RETURN ON INVESTMENT.

With advances in development and application of genetically engineered insulin in the 1970s, and rapid development and maturity of biotechnology in the past four decades, biologics R&D has gradually gained momentum and unlocked new directions for the diagnosis and treatment of diseases.

In the past decade or so, the biologics market and R&D have experienced broad expansion and rapid development. Large pharmaceutical companies have investigated significantly in biologics R&D. According to a forecast report from EvaluatePharma (Figure 1), a US market research institute, eight of the top 10 global best-selling drugs in 2018 were predicted to be monoclonal antibodies and fusion protein drugs, including the first anti-PD-1 antibody drug, Keytruda, approved by the US in 2015, and two additional small molecule amide drugs. Furthermore, it is predicted that the global biologics market will reach 326 billion US dollars in 2022, and that biologics will undoubtedly become one of the most promising technology fields with the potential for fast returns on investment in the pharmaceutical industry.

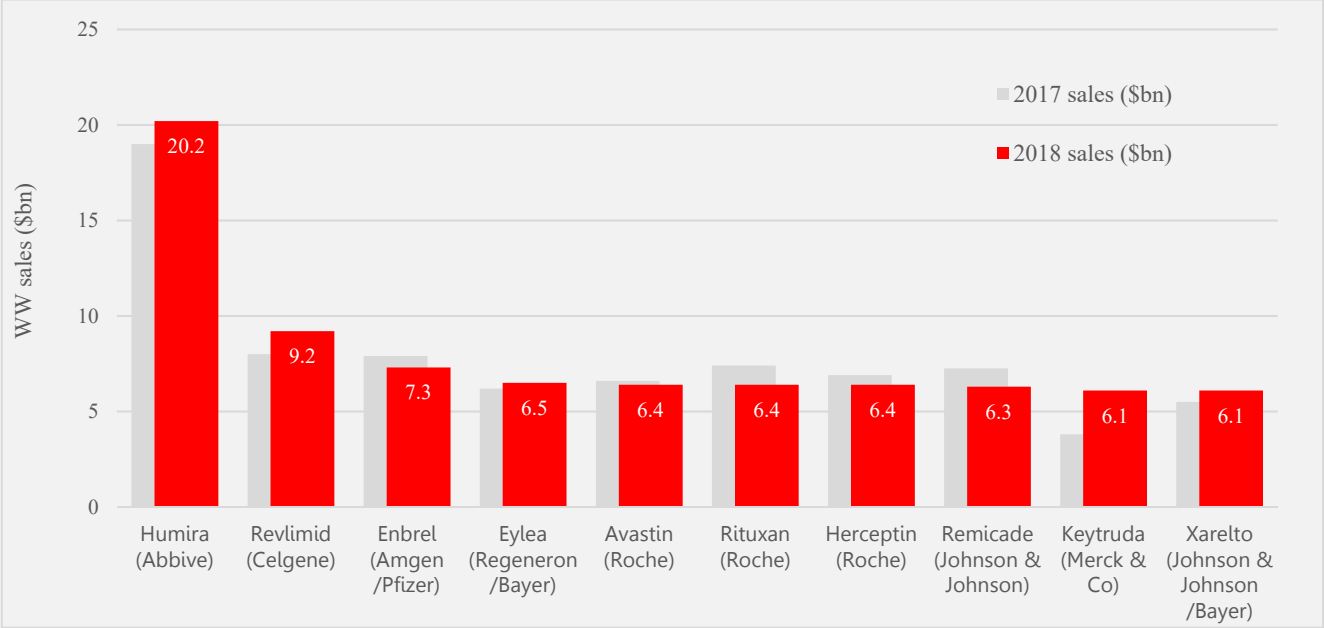


Figure 1. Predicted Top Ten Best Selling Drugs in the World in 2018

According to a report released by the Pharmaceutical Research and Manufacturers of America (PhRMA) in 2013, the United States had developed 907 biologics that could be used for the treatment of more than 100 diseases. These drugs were classified as monoclonal antibodies, vaccines, recombinant hormones/proteins, cell therapy or gene therapy, etc. (Figure 2).

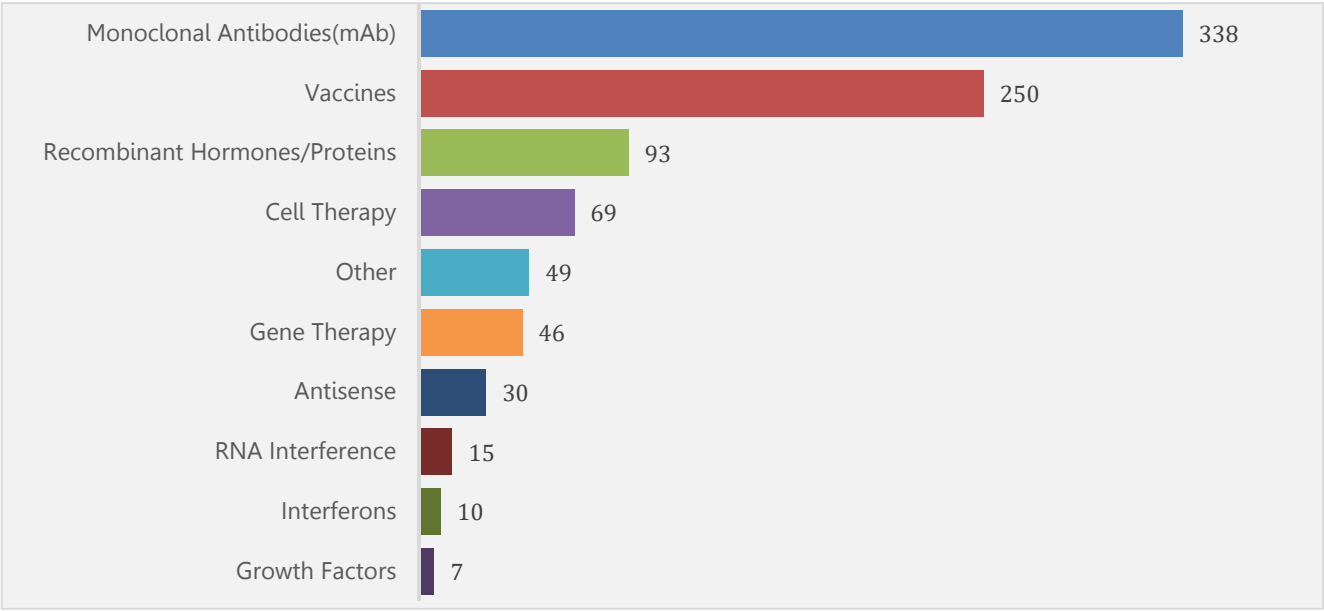


Figure 2. Biologics Classification and Quantity of America (2013)

Based on this classification scheme, this study selected four classes of biologics for further study and analysis using data extracted from the CAS databases. The four classes of biologics chosen for analysis were therapeutic antibodies, fusion proteins, gene and cell therapy and vaccines. An in-depth analysis of these four classes of biologics was conducted in order to gain insights into trends in their basic research and industrial development.

1.2 DATA SOURCES AND METHODS

In this report, scientific papers and patents on biologics acquired up through 2017 were provided by CAS. The analyses focused primarily on data published in recent decades, including conceptual and substance information as well as commercial or government agency and publishing source information.

Due to the long period of time covered in this report, the analysis data was grouped into five-year periods in order to provide a clearer visual presentation and trend analysis. Topical section/category information associated with biologic documents was provided by CAS and used to identify major specialty areas associated with biologics studies. Indexed concepts were used for research topic analysis, and indexed disease terms and a disease hierarchy were used to uncover the various diseases studied. The CAS registered biologics substance collection was also used to further identify patterns and trends in biologics R&D.

1.3 PATENT-RELATED TECHNICAL TERMS

A group of patent applications covering the same or similar technical content are called a patent family. Patents for the same technological invention may be filed in multiple countries or regions, and the CAS database maintains these related multiple applications as one record for indexing and this record is referred as one patent for most analyses in this report.

Individual patent applications from the same patent family may be applied for in different countries and/or patent offices. In order to properly represent the distribution of patent applications filed by assignees' countries (i.e., applicants from different countries or organizations), individual patent applications from the same family are counted separately for geographic distribution and patent flow analysis.

The country of origin of a patented technology is determined based on the country location of the patent assignee (or applicant, inventor). The country analysis in this report is based on the country location of the patent applicant for determination of the country of origin of the technology or actual country owning the technology.

CHAPTER 2. GLOBAL BIOLOGICS R&D TREND ANALYSIS

2.1 DISTRIBUTION OF PAPERS AND PATENTS BY CLASSES OF BIOLOGICS

Based on the CAS document database, since 1877 there have been 373,923 basic research papers in biologics, 137,175 patents (i.e., patent families), and 465,386 individual patent applications after patent families were expanded and redundant patent applications were de-duplicated according to the application number. The total numbers of papers and patents for all biologics as well as for the individual classes of biologics including antibody, fusion protein, gene and cell therapy and vaccine are shown in Table 1.

In terms of paper publications, the vaccine class had the largest number of published papers (150,897, or 40% of the total), whereas the fusion protein class had the least number of papers (12,198, or 3%). Regarding the distribution of patents among these classes, the antibody class had the largest number (54%), and the fusion proteins had the least number (8%).

Table 1. Total Biologics Papers and Patents and Distribution among Individual Classes

	Antibody	Fusion protein	Gene and cell therapy	Vaccine	Total
Number of papers	133,966	12,198	121,974	150,897	373,923
Number of patents (patent families)	73,400	11,337	37,728	45,271	137,175
Number of patents (applications)	278,769	50,164	116,064	148,452	465,386

Note: The overall number of biologics documents was the de-duplicated sum of the four individual classes.

2.2 GLOBAL TREND OF BIOLOGICS PAPERS AND PATENTS

2.2.1 GLOBAL TRENDS OF BIOLOGICS PAPERS

As shown in Figure 3, the international output of biologics research papers has been growing steadily. Based on CAS document data, a total of 373,923 research papers in biologics have been published worldwide up through 2017. Among these, 6,780 papers (1.8%) were published during 1988-1992, 19,747 (5.3%) during 1993-1997, 41,100 (11.0%) during 1998-2002, 70,318 (18.8%) during 2003-2007, 105,081 (28.1%) during 2008-2012 and 114,921 (30.7%) during 2013-2017.

The paper output associated with the individual antibody, fusion protein, gene and cell therapy, and vaccine classes is basically consistent with the overall trend. There is, however, large variability in the number of papers associated with each individual class (Table 1, Figure 3).

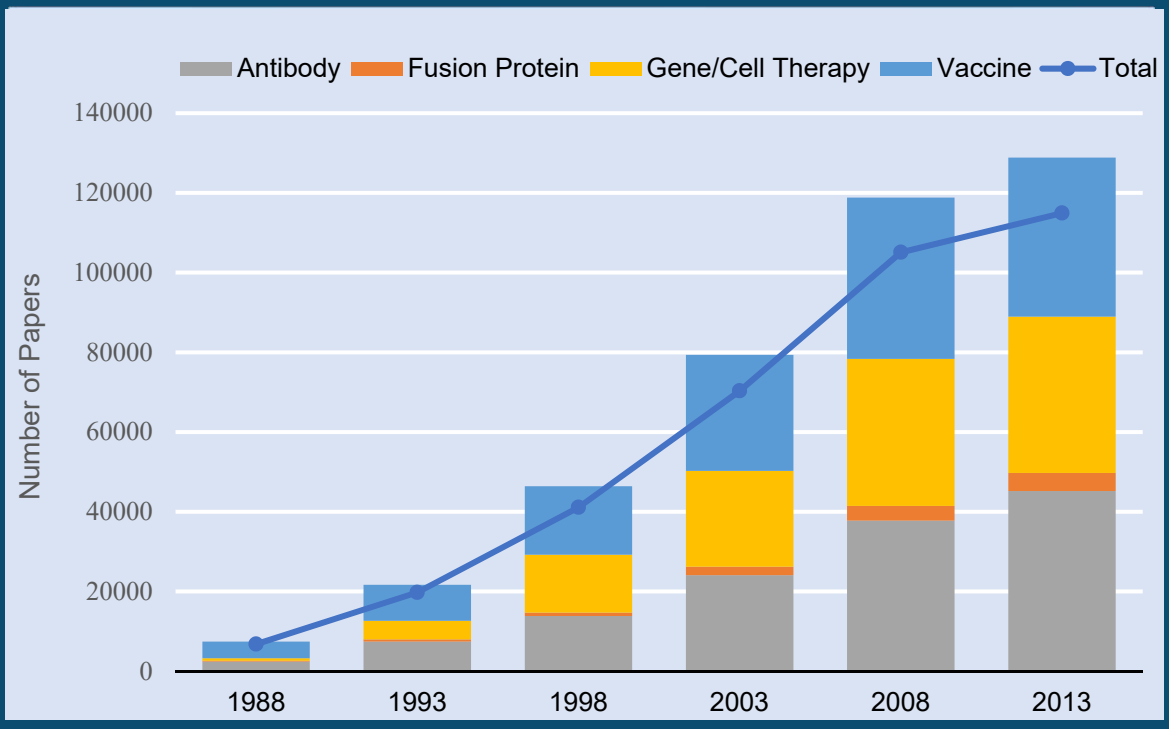


Figure 3. Numbers of Biologics Papers Published Over Time

2.2.2 GLOBAL TRENDS OF BIOLOGICS PATENTS

The global biologics patent output also showed an overall steady growth (Figure 4). In total, there were 137,175 patents worldwide covering 465,386 patent applications in the CAS document database up through 2017 (Table 1). Among them, 4,279 (0.9% of the total), 14,969 (3.2%), 56,670 (12.2%), 87,100 (18.7%), 116,232 (25.0%), and 174,277 (37.5%) patents were applied for in 1988-1992, 1993-1997, 1998-2002, 2003-2007, 2008-2012 and 2013-2017, respectively.

Patents in each class also increased continuously over time. The number of antibody-related patents was significantly higher than those relevant to gene and cell therapy, vaccines, and fusion proteins (Table 1, Figure 4). For example, the proportion of patents related to antibodies, vaccines, gene and cell therapies, and fusion proteins accounted for 60.88%, 27.94%, 26.42% and 12.93%, respectively, of the total number of patents during the period from 2013-2017.

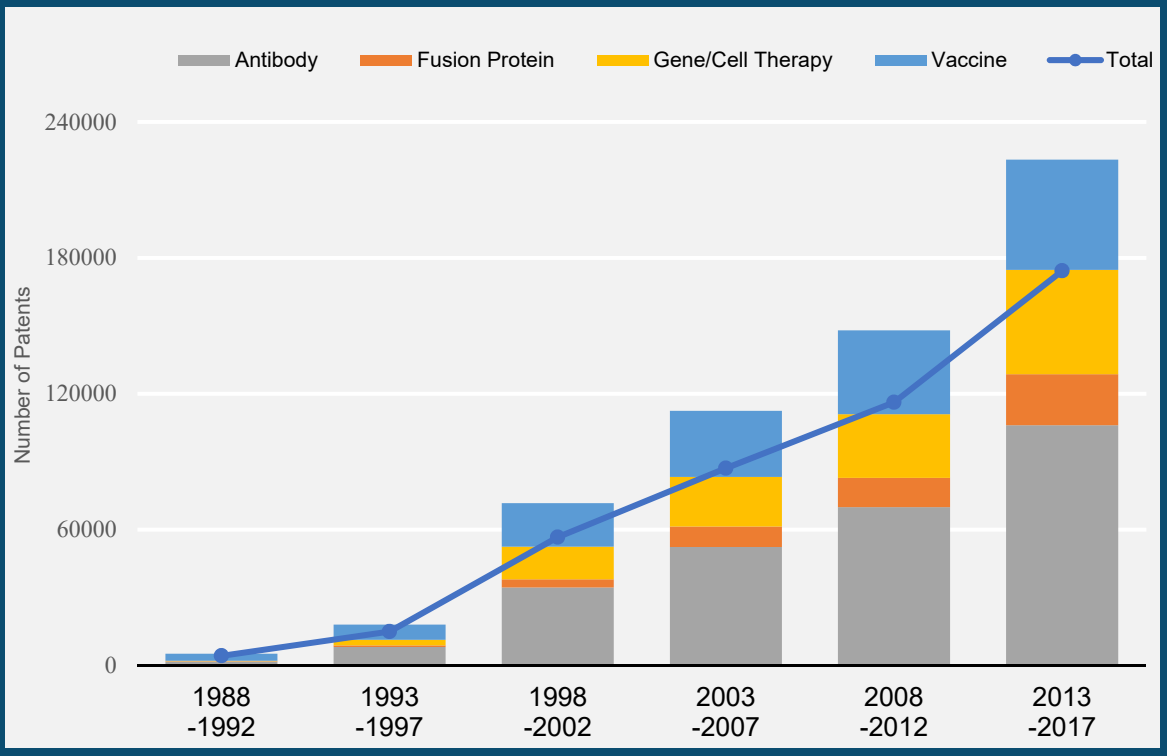


Figure 4. Numbers of Biologics Patents Over Time

2.3 SECTION/SPECIALTY DISTRIBUTION OF BIOLOGICS RESEARCH

Analysis of biologics studies based on area of specialty was conducted using section information provided by CAS. Table 2 lists the top 15 sections based on specialties for papers and patents, respectively. The top 5 specialty areas for both papers and patents were immunochemistry, pharmacology, mammalian pathological biochemistry, biochemical genetics and pharmaceuticals. In addition, the number of papers and patents in these five sections accounted for 111.78% and 186.80%, respectively, indicating overlaps of some documents with more than one specialty area.

This is conceivable because the CAS document database assigns a primary section and one or more secondary section specialties to each paper and patent.

Specialties that were ranked sixth to tenth for biologics papers were microbial/algal/fungal biochemistry, mammalian hormones, mammalian biochemistry, biochemical methods and radiation biochemistry. For patents, radiation biochemistry was ranked thirteenth as a specialty area, while the other five section specialties were among the top ten.

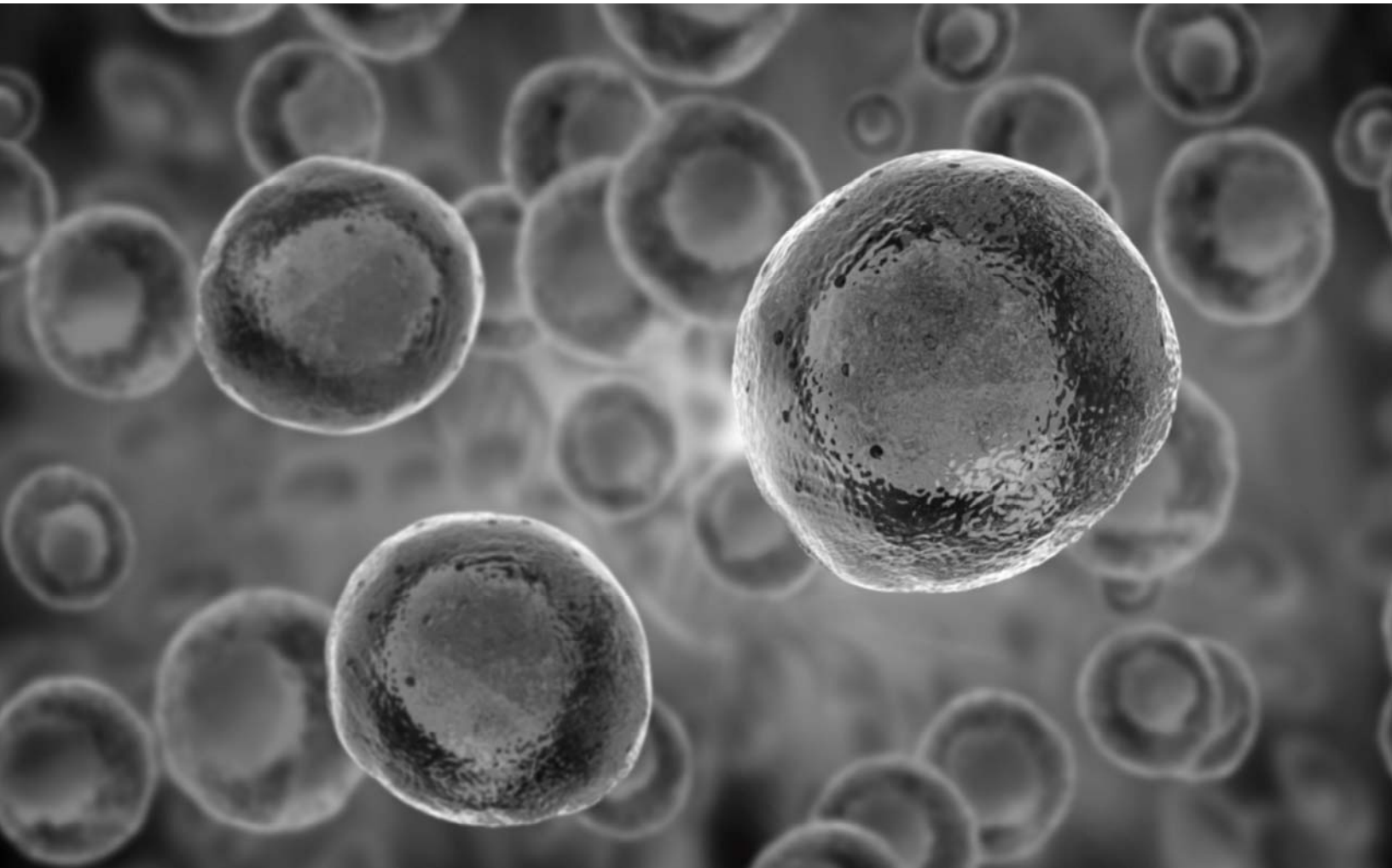


Table 2. Top Fifteen Specialty Sections Associated with Biologics Papers and Patents

Rank	Distribution of sections in papers			Distribution of sections in patents		
	Section	Number of papers	%	Section	Number of patents	%
1	Immunochemistry	185,363	49.57	Immunochemistry	63.305	46.15
2	Pharmacology	100,939	26.99	Pharmacology	61,647	44.94
3	Mammalian Pathological Biochemistry	48,858	13.07	Pharmaceuticals	57,852	42.17
4	Biochemical Genetics	45,346	12.13	Biochemical Genetics	48,151	35.10
5	Pharmaceuticals	37,463	10.02	Mammalian Pathological Biochemistry	25,294	18.44
6	Microbial, Algal, and Fungal Biochemistry	22,606	6.05	Biochemical Methods	20,591	15.01
7	Mammalian Hormones	13,762	3.68	Mammalian Biochemistry	13,879	10.12
8	Mammalian Biochemistry	10,160	2.72	Microbial, Algal, and Fungal Biochemistry	11,561	8.43
9	Biochemical Methods	9,427	2.52	Basic Biochemistry	10,413	7.59
10	Radiation Biochemistry	5,736	1.53	Mammalian Hormones	7,748	5.65
11	Basic Biochemistry	5,532	1.48	Enzyme	5,633	4.11
12	Nonmammalian Biochemistry	3,005	0.80	Fermentation and Bioindustrial Chemistry	3,427	2.50
13	Toxicology	2,974	0.80	Radiation Biochemistry	3,289	2.40
14	Fermentation and Bioindustrial Chemistry	2,632	0.70%	Food and Feed Chemistry	2,844	2.07
15	Enzyme	2,487	0.67%	Heterocyclic Compound	2,364	1.72

2.4 DISTRIBUTION AND EVOLUTION OF RESEARCH SUBJECTS

2.4.1 DISTRIBUTION OF RESEARCH SUBJECTS IN BIOLOGICS

Research papers and patents from 1988 to 2017 were clustered based on their indexed concepts to reveal trends in subject interest. Not surprisingly, in the period 1988-2017, primary topics of research were diseases and their treatments (Figure 5).

Popular biologics-related subjects included neoplasms and antitumor agents, inflammation, treatment strategies including antibodies (especially monoclonal antibodies) and immunoglobulins, gene therapy, protein sequences, vaccines, stem cells, interferon γ and proteins.

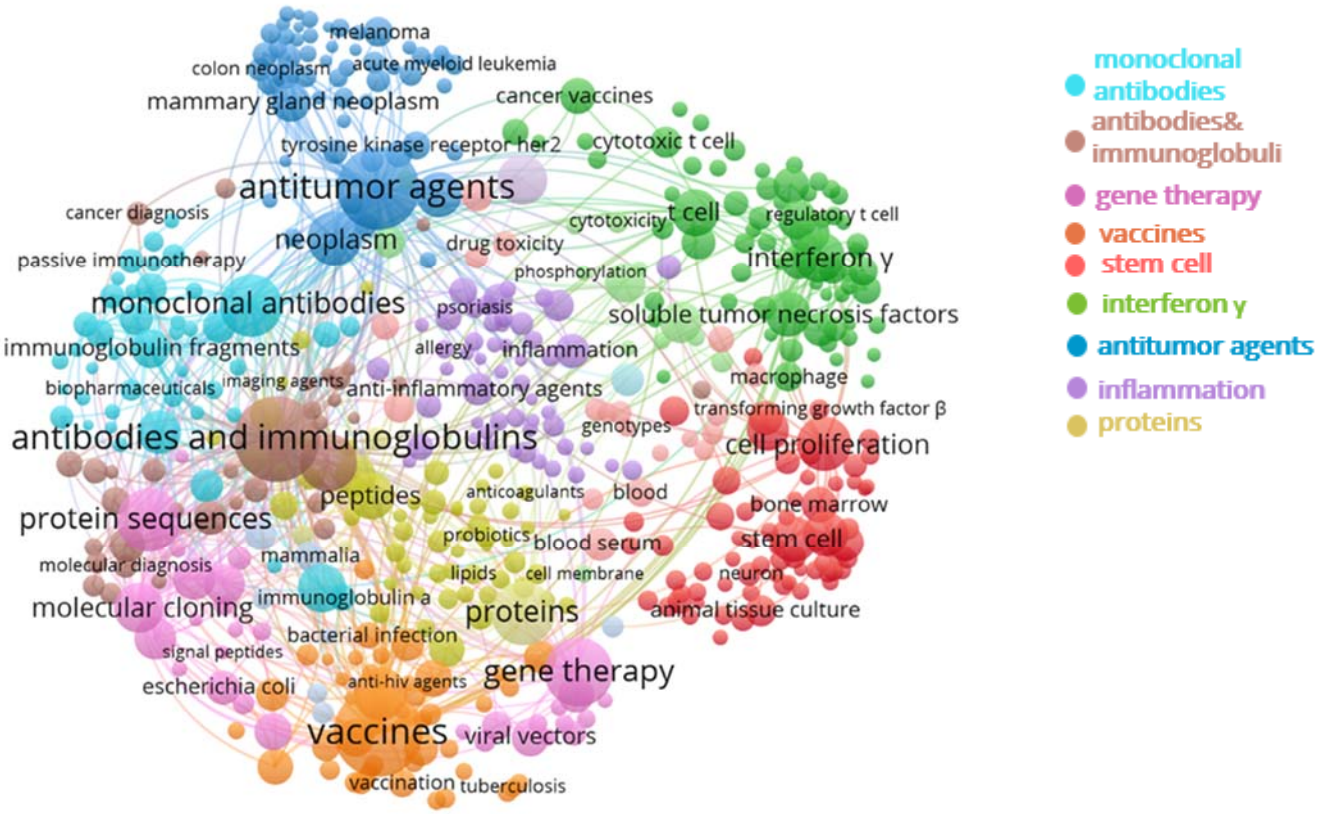


Figure 5. Distribution of Research Subjects of Biologics Papers and Patents

2.4.2 EVOLUTION OF RESEARCH SUBJECTS IN BIOLOGICS

Figure 6 shows the evolution of specific research subjects over time. As can be seen, although there were many research subjects during in the period 1988-1992, most studies focused on vaccines (green dots) and antibodies (red dots) and antigens (blue dots). Related subjects included antitumor agents, proteins, and T cells, etc.

From 1993 to 1997, in addition to the above subjects, new subjects such as gene therapy (yellow dots) and related topics such as DNA sequences, protein sequences, monoclonal antibodies, mutations and gene mapping emerged or gained in frequency. Furthermore, immunology-related studies on tumor necrosis factors, interleukins, various CD antigens, T cells and B cells increased, as can be seen in the green dot area for monoclonal antibodies.

From 1998 to 2002, there were growing numbers of studies on immunodiagnosis (purple dots), drug delivery systems (red dots), and proteins in general (red dots). Meanwhile, studies on molecular diagnosis and drug screening increased.

DURING 2013-2017, ANTIBODIES, VACCINES, ANTITUMOR AGENTS, STEM CELLS AND IMMUNOTHERAPY CONTINUED TO BE HIGHLY STUDIED TOPICS.

Studies on stem cells began to increase (red dots) during the 2003-2007 while antibodies and immunoglobulins, gene therapy, vaccines remained to be popular subjects. More importantly, the studies on immunotherapy increased (blue dots).

During 2008-2012, while those subjects remained to be popular, interests in γ-interferon (green dots) increased.

During 2013-2017, antibodies, vaccines, antitumor agents, stem cells, immunotherapy and γ-interferon and tumor necrosis factors continued to be highly studied topics.

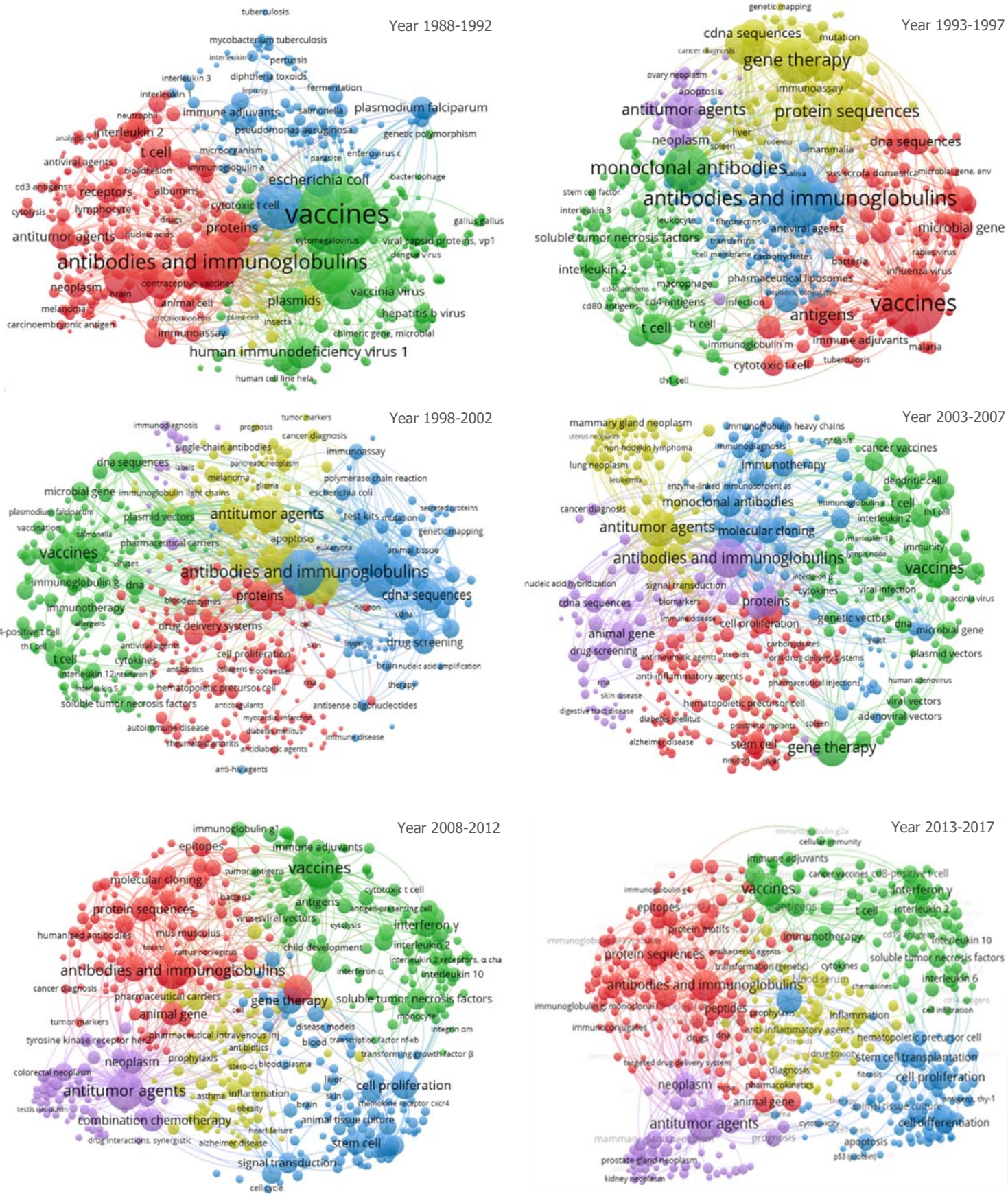


Figure 6. Evolution of Research Subjects in Biologics Papers and Patents

In summary, over the past 30 years, as biotechnology rapidly developed and matured, biologics-related research subjects have shifted focus from antibodies and preventive vaccines in earlier years to more diversified studies on therapeutic antibodies, recombinant protein drugs, tumor vaccines, gene therapy, cell therapy and stem cells in more recent years. While R&D activities in antibodies and immunoglobulins, vaccines, antitumor agents sustained throughout the entire period, gene therapy studies became more robust during 1993-1997 and 2008-2012. Interest in immunotherapeutics also gained momentum over time.

2.5 DISTRIBUTION OF GLOBAL R&D IN MAJOR COUNTRIES/REGIONS

2.5.1 DISTRIBUTION ANALYSIS BY COUNTRIES/REGIONS

Table 3 and Figure 7 reveal that the major countries contributing to biologics paper output up through 2017 were the United States, China, Japan, the United Kingdom, Germany, Italy, France, Canada, the Netherlands and South Korea, The total number of research papers from these ten nations accounted for 77.28% of the total. The combined five-nation output of papers from the United States, China, Japan, the United Kingdom and Germany accounted for 64.25% of the total.

In terms of numbers of biologics patents, the top ten nations consisted of the United States, China, Japan, Germany, South Korea, the United Kingdom, France, Switzerland, Canada and Russia, which produced 79.44% of the total patent applications (Table 3, Figure 7). The top five nations contributed 67.06% of the total.

The United States is a primary producer of published research in biologics, as demonstrated by its papers and patents accounting for 30.75% and 38.50%, respectively, of the global total. Although China and Japan came as second and third in terms of biologics document output, their contributions were much smaller than that of the United States. Additionally, the overall strength of the EU region cannot be underestimated. The total number of papers for the UK, Germany, Italy, France, the Netherlands, Spain, Sweden, Belgium and Denmark altogether has reached to 24% of the global total and the total number of patents from these countries accounted for 16%.

Table 3. Top 20 Countries/Regions in Biologics Papers and Patents

Rank	Distribution of papers			Distribution of patents		
	Countries/ Regions	Number of papers	Proportion of papers %	Countries/ Regions	Number of patents	Proportion of patents %
1	USA	114,997	30.75	USA	52,810	38.50
2	China	58,547	15.66	China	18,339	13.37
3	Japan	28,240	7.55	Japan	9,361	6.82
4	UK	19,345	5.17	Germany	6,069	4.42
5	Germany	19,113	5.11	South Korea	5,404	3.94
6	Italy	12,666	3.39	UK	4,646	3.39
7	France	12,211	3.27	France	4,525	3.30
8	Canada	8,941	2.39	Switzerland	3,244	2.36
9	Netherlands	7,706	2.06	Canada	2,624	1.91
10	South Korea	7,214	1.93	Russia	1,946	1.42
11	Australia	7,206	1.93	Netherlands	1,762	1.28
12	India	6,555	1.75	Australia	1,669	1.22
13	Spain	6,095	1.63	Belgium	1,667	1.22
14	Switzerland	4,807	1.29	Israel	1,499	1.09
15	Sweden	4,385	1.17	Italy	1,261	0.92
16	Brazil	3,670	0.98	Spain	1,114	0.81
17	Belgium	3,600	0.96	Denmark	1,099	0.80
18	Taiwan, China	2,976	0.80	India	1,010	0.74
19	Iran	2,900	0.78	Sweden	958	0.70
20	Israel	2,636	0.70	Taiwan, China	822	0.60

Note: Absolute numbers of publications were used in this analysis; other factors such as impact factor were not used

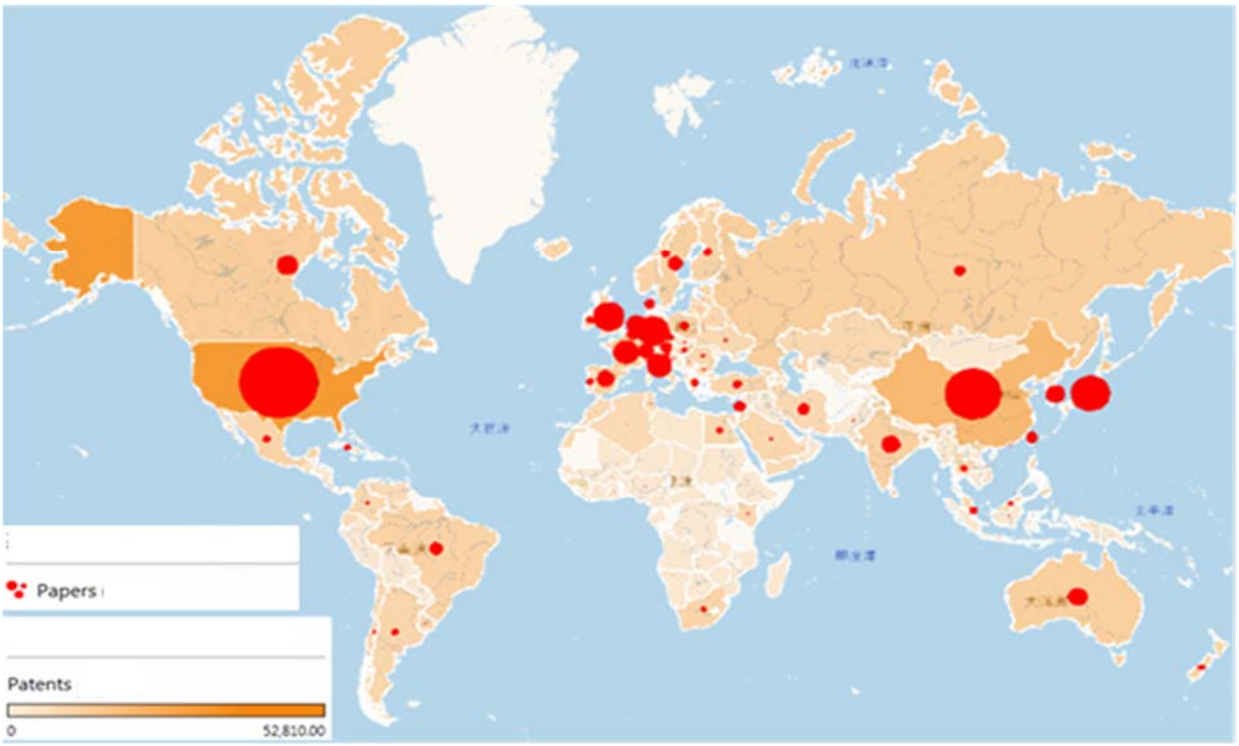


Figure 7. Countries/Regions (China excluding Hong Kong, Macau and Taiwan, the same below) in Biologics Papers and Patents

2.5.2 PUBLICATION TIMELINE ANALYSIS BY MAJOR COUNTRIES/REGIONS

Based on paper publication dates (Figure 8), five countries including the United States, China, Japan, the United Kingdom, and Germany conducted biologics research prior to the 1990s. The United States has led the world in biologics research publications in both the total paper publications and numbers of papers published in each 5-year period, and is continuing to maintain its lead. For example, the number of papers published by the United States exceeded 30,000 during 2013-2017. China is a rising star in biologics research demonstrating rapid progress in recent years. Since 2003-2007, China's publication numbers have surpassed those of Japan, the United Kingdom and Germany, and are second only to the United States. In addition, its number of papers was close to 21,000 during 2013-2017. Japan, Germany and the United Kingdom have shown a more steady development and moderate growth in biologics research. The number of papers published from each of the three countries during 2013-2017 period was less than 7,000.

Figure 9 shows the trend of patents by the top five countries: the United States, China, Japan, Germany and South Korea. This ranking is slightly different from that of paper publications because South Korea replaced the United Kingdom in the top five. The United States has also historically been the leader in biologics patent applications and has maintained its lead. Since 2003, the number of patents from China has grown rapidly, gradually surpassing those of Japan and Germany, and has narrowed its gap with the United States. The patents from Japan and Germany remain on a relatively slow pace.

The numbers of patents filed by the US and China during 2013-2017 were nearly 15,000, and 11,000, respectively, while those filed by Japan and Germany totaled less than 3,000, surpassed only slightly by South Korea.

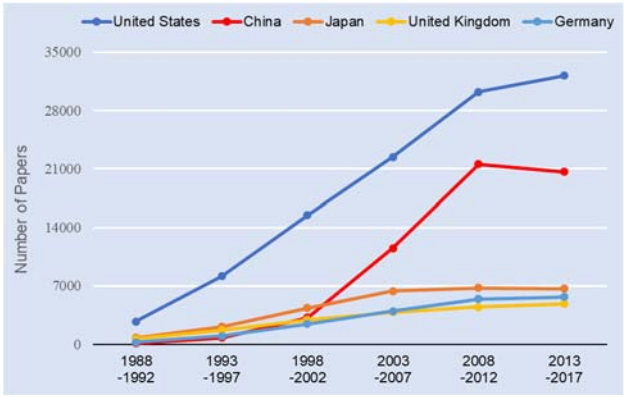


Figure 8. Biologics Papers in Major Countries Over Time

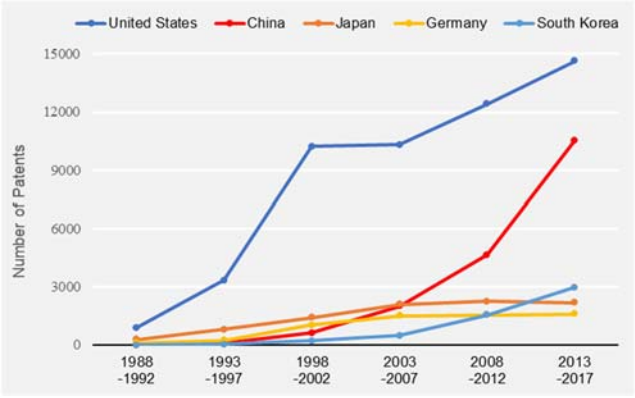


Figure 9. Biologics Patents in Major Countries Over Time

2.5.3 FLOW OF PATENT APPLICATIONS BETWEEN MAJOR COUNTRIES/REGIONS

The flow of patents between countries refers to the number of patents with inventors in one country that were applied for in another country. Figure 10 shows the flow of biologics patent applications between the top five patent countries including the United States, China, Japan, Germany, and South Korea during the past three decades. The United States, Japan, and Germany are in an obvious position of technological export, while China and South Korea are in a state of technological import. The United States protects its intellectual properties heavily in the Japan market, followed by the Chinese and Korean markets. For example, the United States exported 11,742, 6,355 and 4,660 patents to Japan, China and South Korea, respectively. These are relatively high numbers in comparison to the 72 patents it exported to Germany.

Both Japan and Germany focused more on patenting their technologies in the US market. They applied for 3,874 and 3,272 patents, respectively, in the US. In addition, Japan exported 1122, 853 and 231 patents to China, South Korea and Germany, respectively. Germany exported 1650, 913 and 681 patents to Japan, China and South Korea, respectively. Although China had a large number of patent applications, it only exported 630, 217 and 104 patents to the United States, Japan and South Korea, respectively and only one to Germany. In general, the United States, Japan and Germany strive to protect patented technology in overseas markets. China's export of patented technology to other countries is relatively low, and its patents are still limited to protection within China.

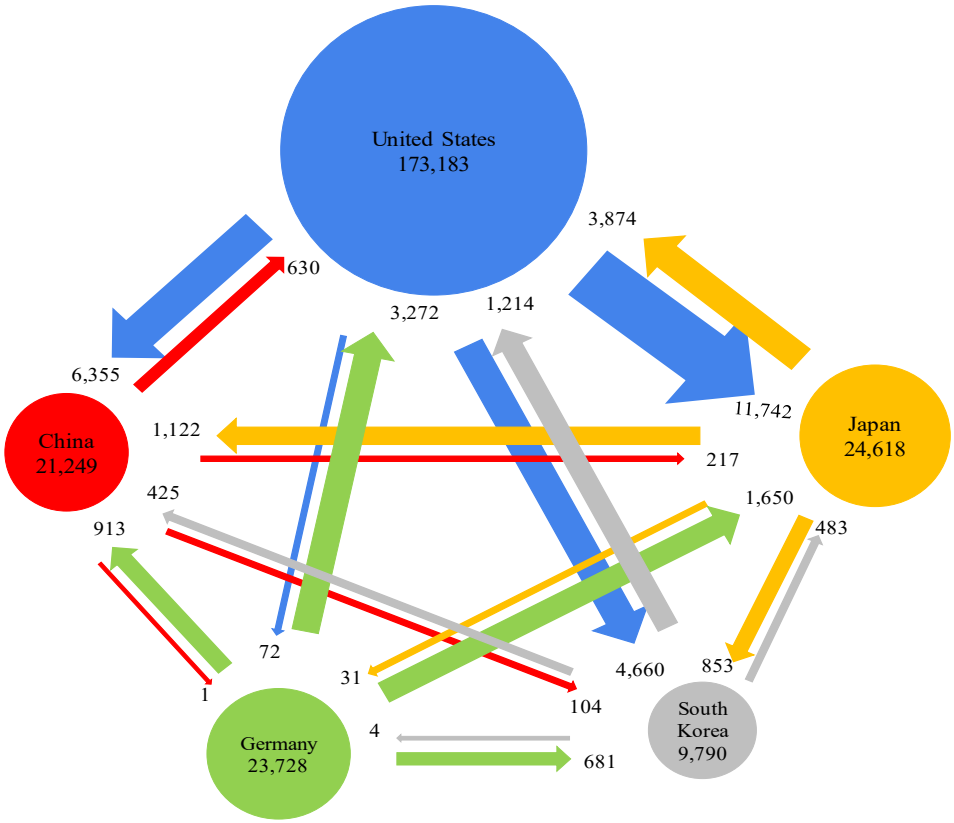


Figure 10. Flow of Patented Technology in Major Countries (the thickness of the arrow is positively correlated with the proportion of patent applications)

2.6 DISTRIBUTION OF BIOLOGICS R&D ORGANIZATIONS WORLDWIDE

2.6.1 MAJOR ORGANIZATIONS IN PAPER PUBLICATION

As shown in Table 4 and Figure 11, basic research in biologics, published as papers, has been dominated by universities and research institutes. All top 20 organizations were universities and research institutes (or clinical research institutes). Among them, 13 were from the United States and the top eight were all American institutes, indicating that the United States plays a strong leading role.

There are four from China, two from Japan, and one from France in the top 20 organizations. The top three organizations were the University of California, National Institutes of Health, and the University of Texas, with 5,366, 4,500, and 4,343 papers, respectively.

Table 4. Top 20 Organizations in Paper Publication

Rank	Agencies with basic research publications	Number of papers	Country	Type of Organization
1	University of California	5,366	USA	University
2	National Institutes of Health	4,500	USA	Research institute
3	University of Texas System	4,343	USA	University
4	Harvard University	2,291	USA	University
5	University of Pennsylvania	2,103	USA	University
6	Johns Hopkins University	2,003	USA	University
7	University of Pittsburgh	1,902	USA	University
8	Mayo Clinic	1,744	USA	Clinical Research institute
9	Osaka University	1,603	Japan	University
10	Academy of Military Medical Sciences	1,602	China	Research institute
11	Chinese Academy of Sciences	1,576	China	Research institute
12	Stanford University	1,553	USA	University
13	University of Washington	1,544	USA	University
14	Duke University	1,447	USA	University
15	Institut Pasteur	1,442	France	Research institute
16	Fourth Military Medical University	1,431	China	University
17	Chinese Academy of Medical Sciences	1,427	China	Research institute
18	University of Tokyo	1,426	Japan	University
19	University of Michigan	1,389	USA	University
20	University of Alabama	1,356	USA	University

2.6.2 DISTRIBUTION OF PATENT APPLICANTS

As shown in Table 5 and Figure 11, the United States shows strong leadership in biologics R&D, possessing 10 of the top 20 biologics patent producers, the majority of which are corporations/enterprises. Additional patent-producing organizations among the top 20 include three from France, two each from Switzerland and the United Kingdom and one each from China, Japan and Germany. Except for the fourth-ranked University of California system, all four of the top five patent producers were corporations.

GlaxoSmithKline from the UK, and F. Hoffmann-La Roche & Co. AG and Novartis AG from Switzerland took the top three ranking positions with patent outputs of 3,509, 2,416 and 1,965, respectively. The top 20 list of patent producers includes a number of multinational pharmaceutical giants such as Merck and Co., Inc., Pfizer Inc., Bristol-Myers Squibb, Amgen Inc. and Incyte Corp from America, Bayer AG from Germany, Takeda Pharmaceutical Co., Ltd. From Japan, Sanofi-Aventis from France and AstraZeneca from the United Kingdom. The Chinese Academy of Sciences was the only Chinese organization in the top 20, ranking ninth in the world with 1,017 patent publications.

Table 5. Top 20 Biologics Patent Applicants

Rank	Patenting Organization	Number of Patents	Country	Agency category
1	GlaxoSmithKline	3,509	UK	Corporation
2	F. Hoffmann-La Roche & Co. AG	2,416	Switzerland	Corporation
3	Novartis AG	1,965	Switzerland	Corporation
4	University of California	1,548	USA	University
5	Merck and Co., Inc.	1,395	USA	Corporation
6	U.S. Department of Health and Human Services	1,385	USA	Research institute
7	Pfizer Inc.	1,181	USA	Corporation
8	Institut National de la Sante et de la Recherche Medicale	1,067	France	Research institute
9	Chinese Academy of Sciences	1,017	China	Research institute
10	Bayer AG	1,006	Germany	Corporation
11	Takeda Pharmaceutical Co., Ltd.	950	Japan	Corporation
12	Sanofi-Aventis	937	France	Corporation
13	University of Texas System	883	USA	University
14	Amgen Inc.	835	USA	Corporation
15	Bristol-Myers Squibb	805	USA	Corporation
16	Centre National de la Recherche Scientifique	795	France	Research institute
17	AstraZeneca	736	UK	Corporation
18	Incyte Corp.	686	USA	Corporation
19	Johns Hopkins University	678	USA	University
20	University of Pennsylvania	669	USA	University

Figure 11 summarizes the distribution of research papers and patented technologies among universities, research institutes and enterprises in major countries contributing to biologics R&D activities.

It further demonstrates the overwhelmingly leading position of the United States in the research and development of biologics.

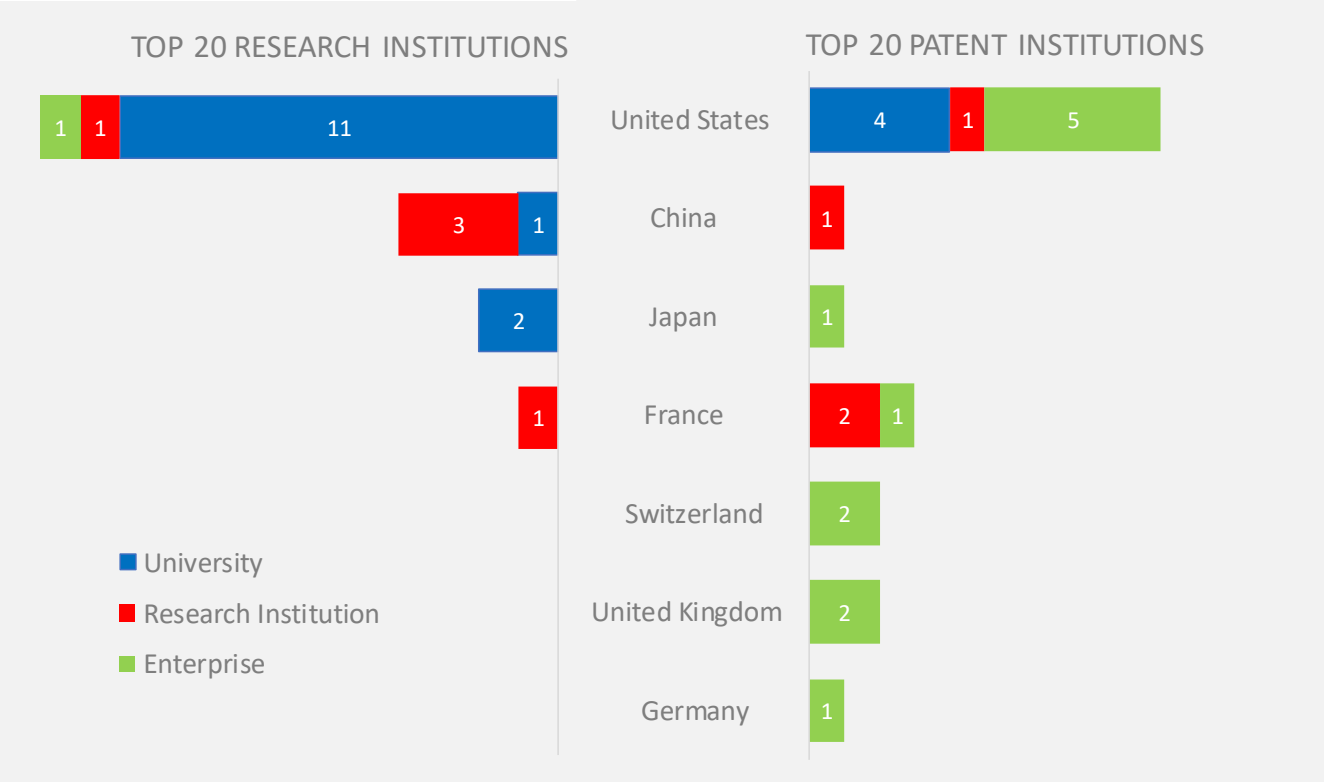


Figure 11. Summary of Various Types of Organizations in Major Countries in Their Contributions to Biologics R&D

CHAPTER 3. DISEASES STUDIED IN BIOLOGICS RESEARCH

3.1 OVERVIEW OF DISEASES STUDIED

The top ten disease types associated with research reported in the biologics papers and patents analyzed in this report were neoplastic diseases/cancer, digestive system diseases, infection, lymphatic system diseases, immune diseases, respiratory system diseases, inflammation, musculoskeletal diseases, hematopoietic disorders and cardiovascular diseases.

In general, disease studies were more prevalent in papers than patents. However, mental and behavioral disorders, endocrine system diseases and degenerative diseases appeared more frequently in patent applications.

Neoplastic diseases (cancer) were the most widely studied disease, ranking first in published research related to antibodies, gene and cell therapy and vaccines. Digestive diseases ranked second with equal distribution between patents and papers. Infectious diseases, primarily associated with vaccine studies, ranked third, with greater occurrence in papers than patents. Studies on neoplastic and lymphatic diseases were significantly more numerous among paper publications than patents in general.

STUDIES ON NEOPLASTIC AND LYMPHATIC DISEASES WERE SIGNIFICANTLY MORE NUMEROUS AMONG PAPER PUBLICATIONS THAN PATENTS IN GENERAL.

In contrast, a more balanced distribution between paper and patent coverage was observed for immune diseases and inflammation as related to antibodies, fusion proteins, genes and cell therapy and vaccines (Figure 12).

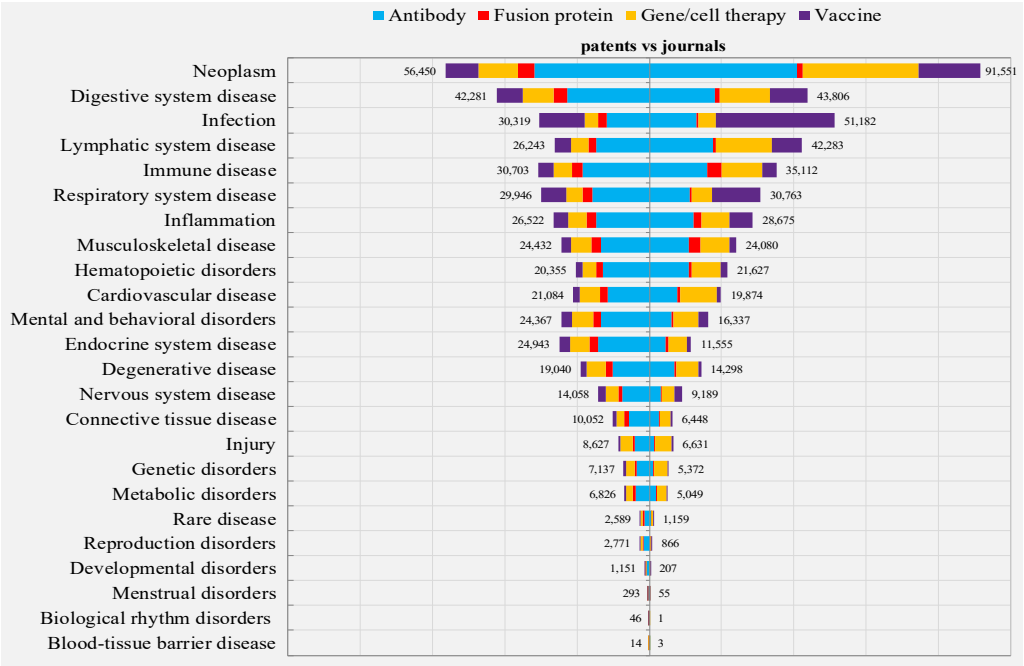


Figure 12. Distribution of Diseases (no de-duplication between sub-areas)

Figure 13 illustrates the distribution of the top 15 diseases studied in each of the four biologics classes. As shown in the figure, the majority of disease studies were associated with the therapeutic antibody class, followed by the gene and cell therapy, vaccine and fusion protein classes. Infectious diseases were mostly studied in the vaccine class.

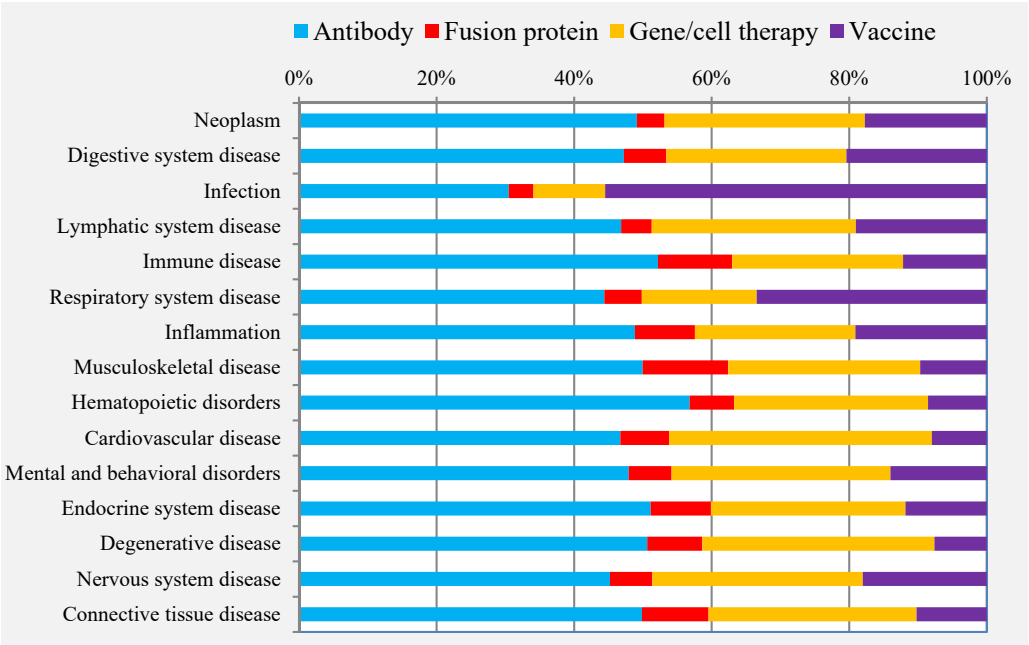


Figure 13. Distribution of Top 15 Diseases Based on Biologics Class

3.2 RECENT RESEARCH ON DISEASES (2013-2017)

The bulk of the disease-related research reported in both biologics papers and patents has been published within recent years. With the exception of immune disease, publications associated with the other top 14 diseases during the 2013-2017 time period consistently accounted for 35-40% of the total number of biologics studies reported in the historical literature (Table 6). Immune disease studies in the past five years accounted for 31% of the total in the entire period.

Table 6. Comparison of the Amount of Literature on Diseases from 2013-2017 with the Historical Total

Disease	Total Historical Publications (de-duplicated)	Publications from 2013-2017	Percentage of Publications from 2013-2017
Neoplasm	78,215	28,882	36.9
Digestive system disease	45,063	17,378	38.6
Infection	33,616	12,200	36.3
Lymphatic system disease	34,810	12,735	36.6
Immune disease	36,691	11,348	30.9
Respiratory system disease	31,553	12,593	39.9
Inflammation	29,430	10,697	36.3
Musculoskeletal disease	25,611	9,560	37.3
Hematopoietic disorders	24,641	8,742	35.5
Cardiovascular disease	20,135	7,186	35.7
Mental and behavioral disorders	21,058	8,272	39.3
Endocrine system disease	20,167	7,590	37.6
Degenerative disease	17,782	6,613	37.2
Nervous system disease	11,682	3,998	34.2
Connective tissue disease	8,728	3,315	38.0

3.3 ANALYSIS OF RESEARCH INTENSITY ON MAJOR DISEASES

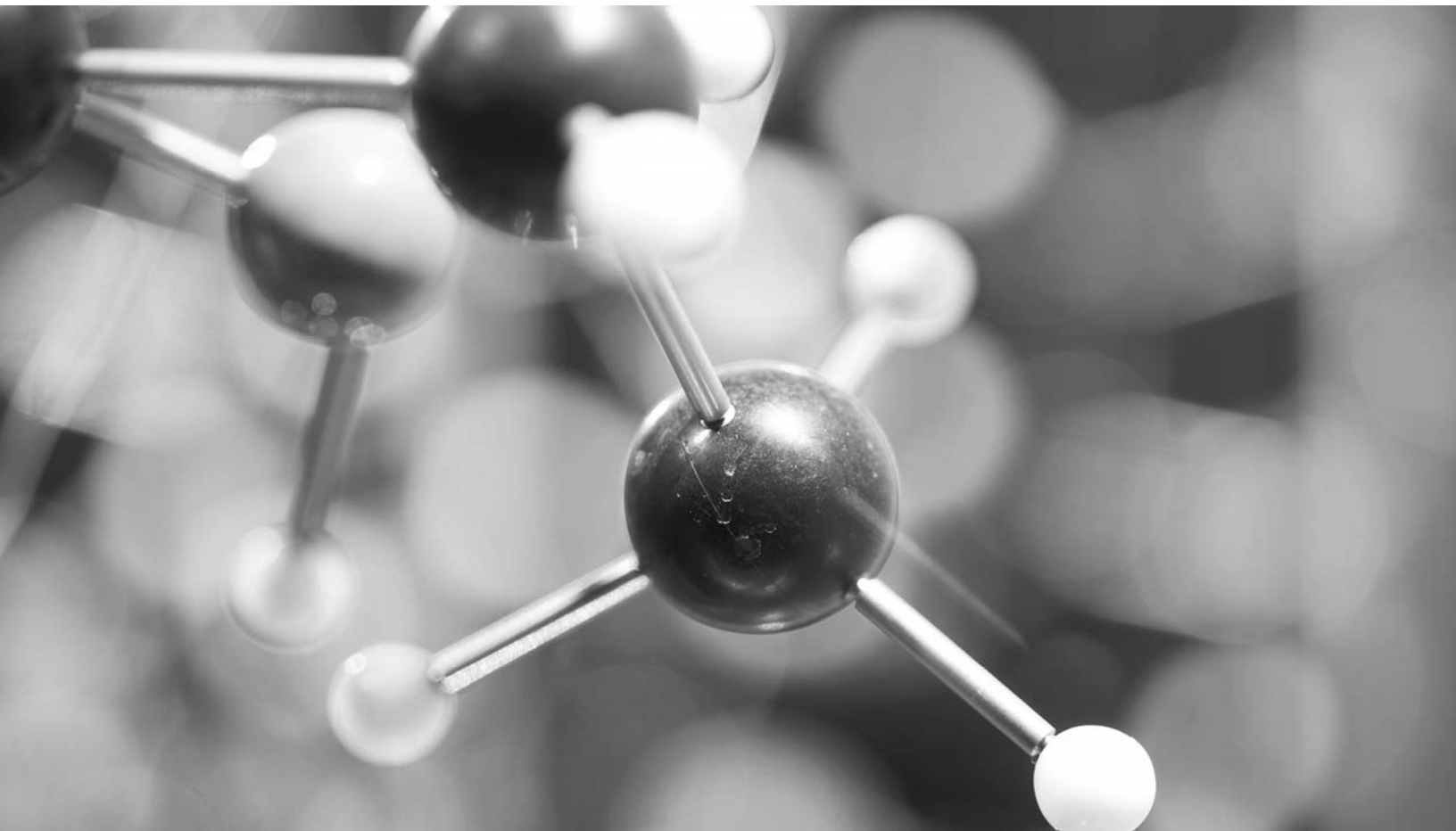
An in-depth analysis was performed on three different disease classes: neoplasms, immune diseases and inflammation. Based on a hierarchical classification of these diseases, tree diagrams were constructed to show relationships between research intensity related to individual diseases (or disease classes) (Figures 14-16).

3.3.1 Research focused on neoplasms

Neoplasms (cancerous tumors) constitute the most widely studied disease class in the biologics papers and patents. Figure 14 shows that more than 20 different types of neoplastic diseases were reported in over 50 documents at the second hierarchy level and over 100 types of
specific neoplastic diseases were reported in over 50 documents at the third hierarchy level.

Breast cancer (mammary gland neoplasm) has been the most studied disease among the various neoplastic diseases with a total of 17,540 associated publications in the document database. Specific neoplastic forms of the disease, including carcinoma, sarcoma, and endocrine system neoplasm, were indexed relatively more frequently.

At the third hierarchical level of disease classification, lung neoplasms were reported most frequently with a total of 9,545 associated publications, indicating that this is also one of the most intensely studied neoplastic disease types. Additionally, ovary neoplasms, pancreatic neoplasms, liver neoplasms, lymphoma and leukemia are popular areas of research, with their published references exceeding 5,000.



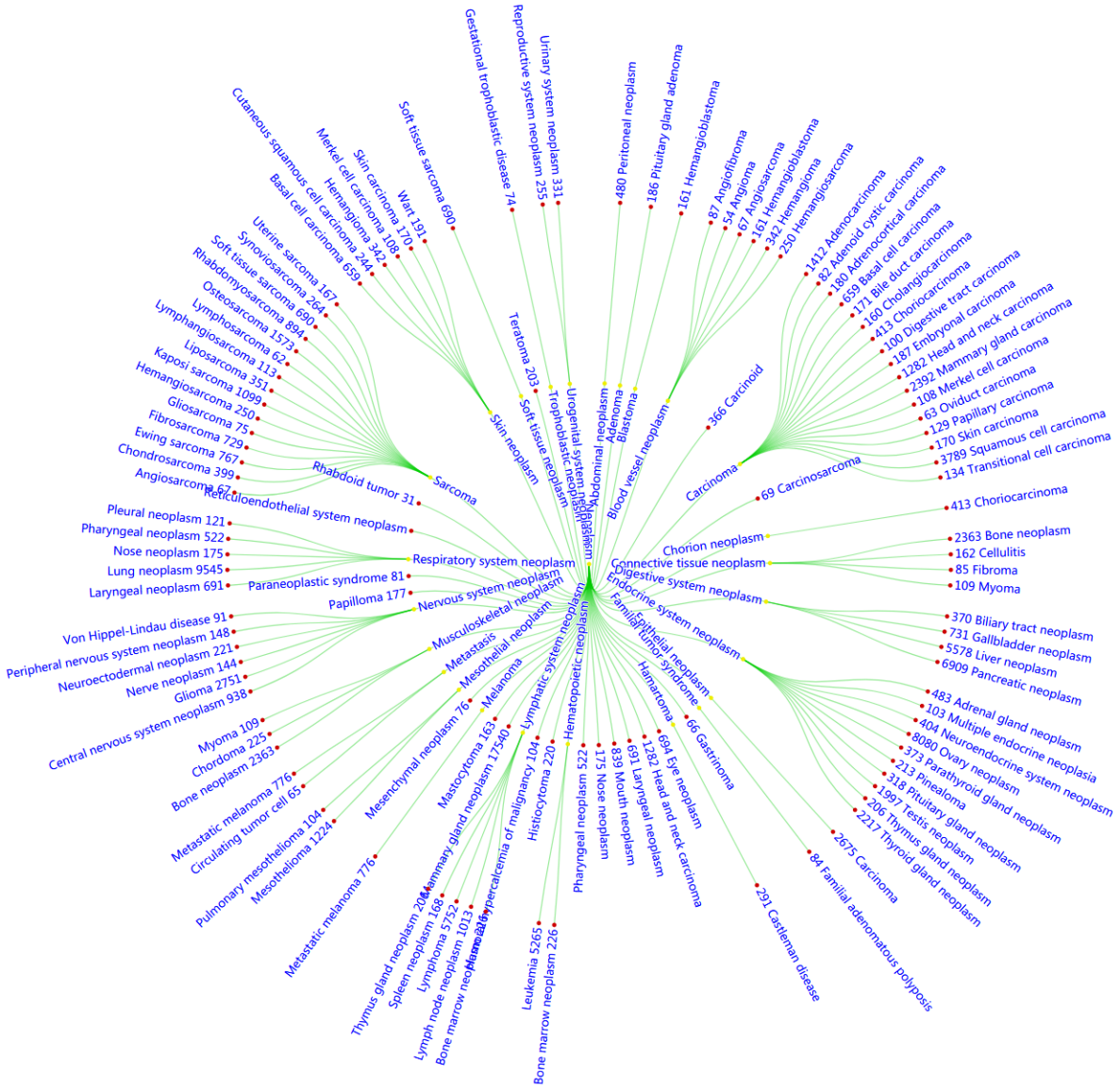


Figure 14. Neoplastic Disease Frequency

Note: Types of neoplasm terms indexed in more than 50 documents in the entire period studied

3.3.2 RESEARCH FOCUSED ON IMMUNE DISEASES

Immune diseases rank fifth in publication frequency with a total of over 30,000 associated documents. Based on the hierarchical tree classification system, there were about 21 types of second-level immune disease classes and 51 types of third-level diseases reported in at least 50 or more document database records
(Figure 15).

Autoimmune disease was the most frequently reported immune disease type, based on the second level hierarchy. Rheumatoid arthritis was the most extensively reported disease, occurring in 10,089 publications. It is also a popular topic of research in relation to inflammation studies. Multiple sclerosis (5,999 documents), multiple myeloma (5,092 documents) and psoriasis (5,009 documents) were also highly published areas of disease research.

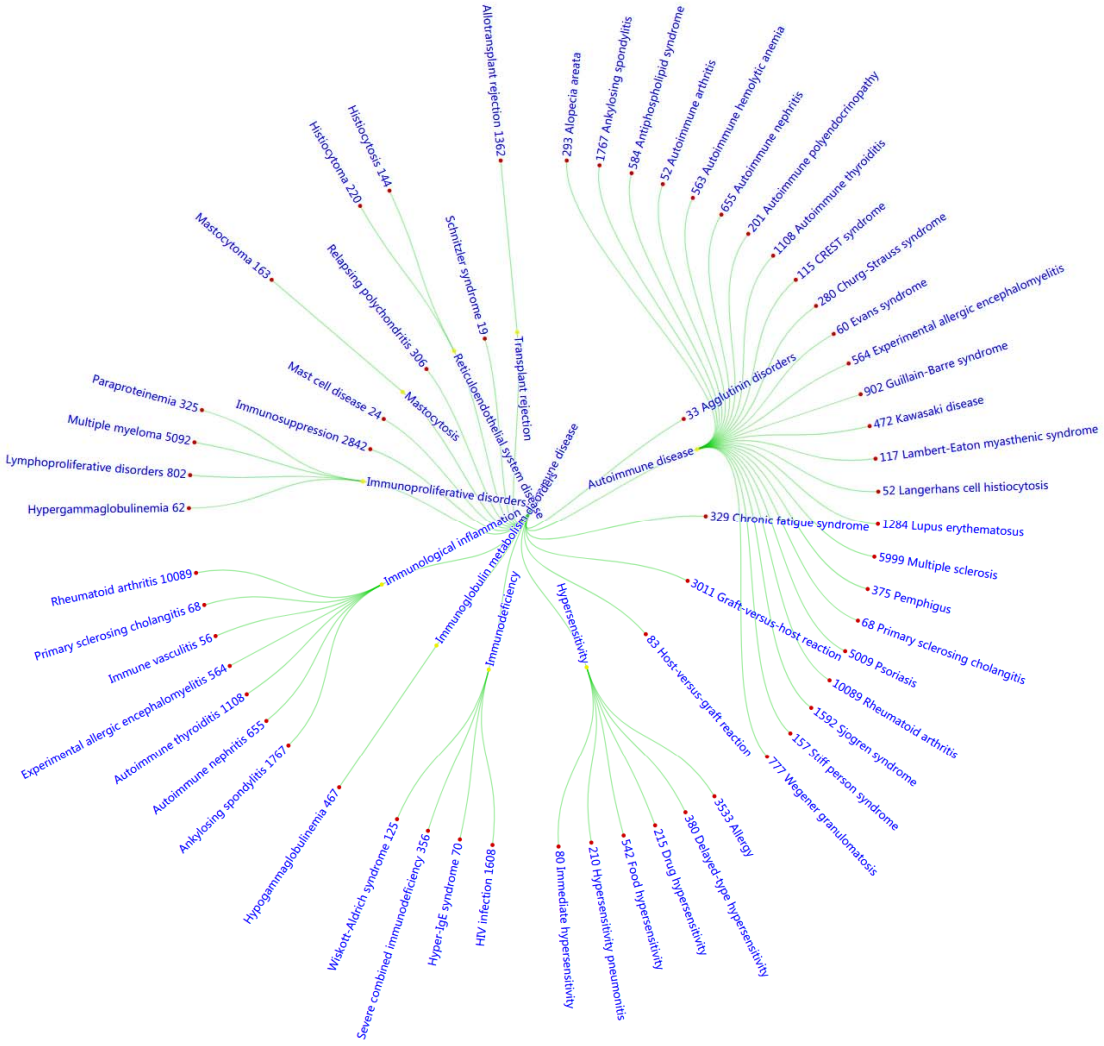


Figure 15. Immune Disease Frequency

Note: Types of immune disease terms indexed in more than 50 documents in the entire period studied

3.3.3 RESEARCH FOCUSED ON INFLAMMATORY DISEASES

Inflammatory diseases ranked seventh in biologics publication frequency, being associated with close to 30,000 documents. Based on its hierarchical tree classification (Figure 16), there were about 62 second-level inflammatory diseases and more than 100 types of specific third-level inflammatory diseases reported in more than 50 documents with a relatively balanced distribution of publication frequency.

Because rheumatoid arthritis is an autoimmune disease with inflammation, it appears in both of these hierarchies. It was associated with the largest number of publications (10,089 documents) and was the only disease with over 10,000 documents. Research studies on psoriasis (5,009 documents) and inflammatory bowel disease (3,791 documents) were also numerous.

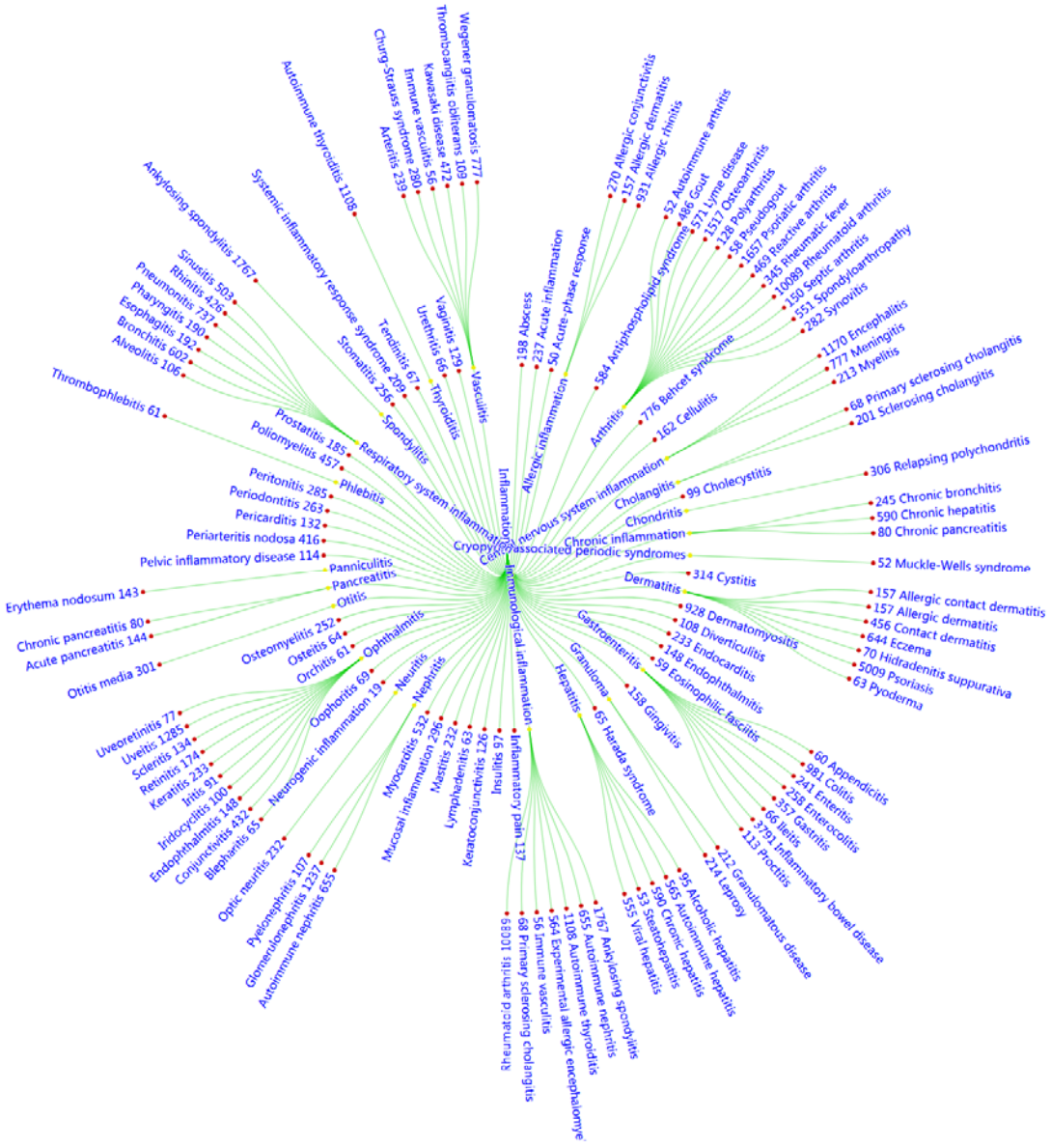


Figure 16. Inflammatory Disease Frequency

Note: Types of Inflammatory Disease terms indexed in more than 50 Documents

CHAPTER 4. BIOPHARMACEUTICAL SUBSTANCES

4.1 TOTAL NUMBERS OF BIOPHARMACEUTICAL SUBSTANCES

Biopharmaceutical substances are chemical or chemically-related entities assigned unique CAS Registry Numbers. Some may have associated commercial trade-mark names and/or known sequence information as well as chemical modification information. They include peptides, proteins, nucleic acids and vaccines, which may be in the early stages of R&D, in clinical trials or in clinical use. Some biologics are registered based on functional descriptions in cases where the chemical structures/sequences are not available (e.g. some vaccines).

This report utilized data from CAS Registry® for 207,880 biopharmaceutical substances, including 164,839 antibodies (79%), 37,235 fusion proteins (18%), 5,050 gene and cell therapy-related substances (2%) and 2,142 vaccine substances (1%) (Figure 17). A small number of these biologics (1,383) were categorized into two or more classes.

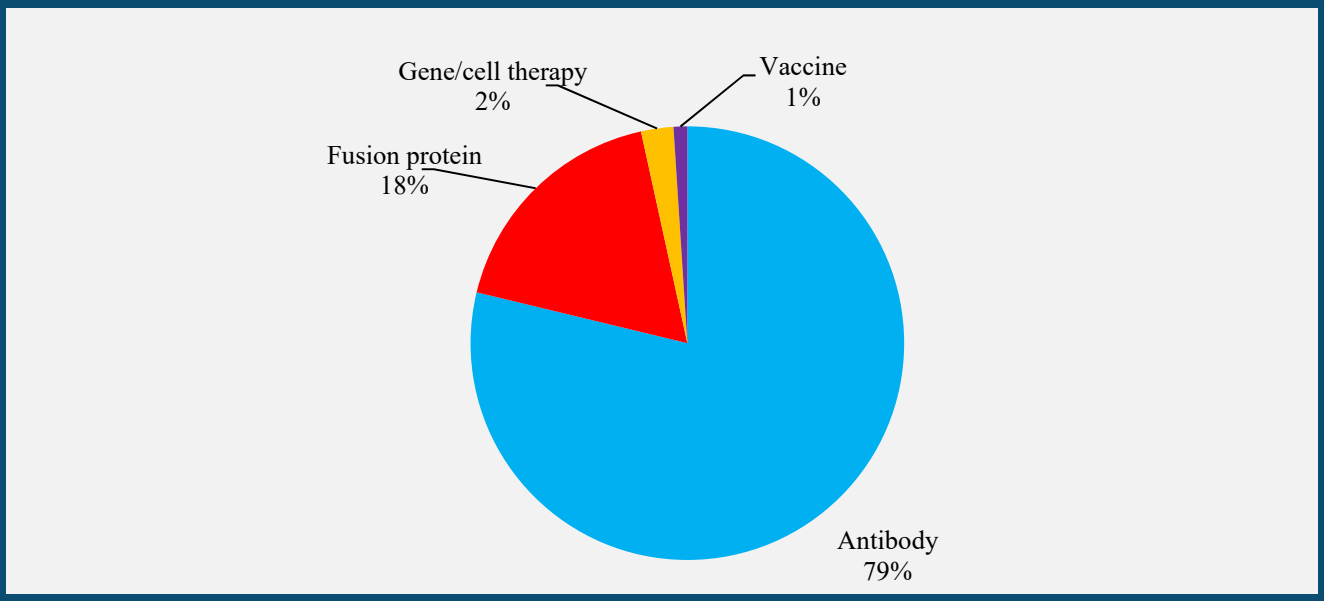


Figure 17. Distribution of Biologics Substance Classes

4.2 SUBSTANCE REGISTRATION OVER TIME

The number of registered biopharmaceutical substances has shown a historically steady increase, which has tapered off slightly since 2013 (Figure 18). The most pronounced increase in registered biologics occurred during the 1998-2012 timeframe. The growth rate in antibody registrations paralleled that of the overall growth rate in substance registrations. The registration of fusion proteins was highest from 2008-2012.

The number of registered substances related to gene and cell therapy increased rapidly during 2013-2017, with nearly four times as many such substances (2,867) registered as compared with the 2008-2012 time period. Although vaccine substance registrations occurred earlier historically than other biologic classes, their registration rate has remained low in the past two decades.

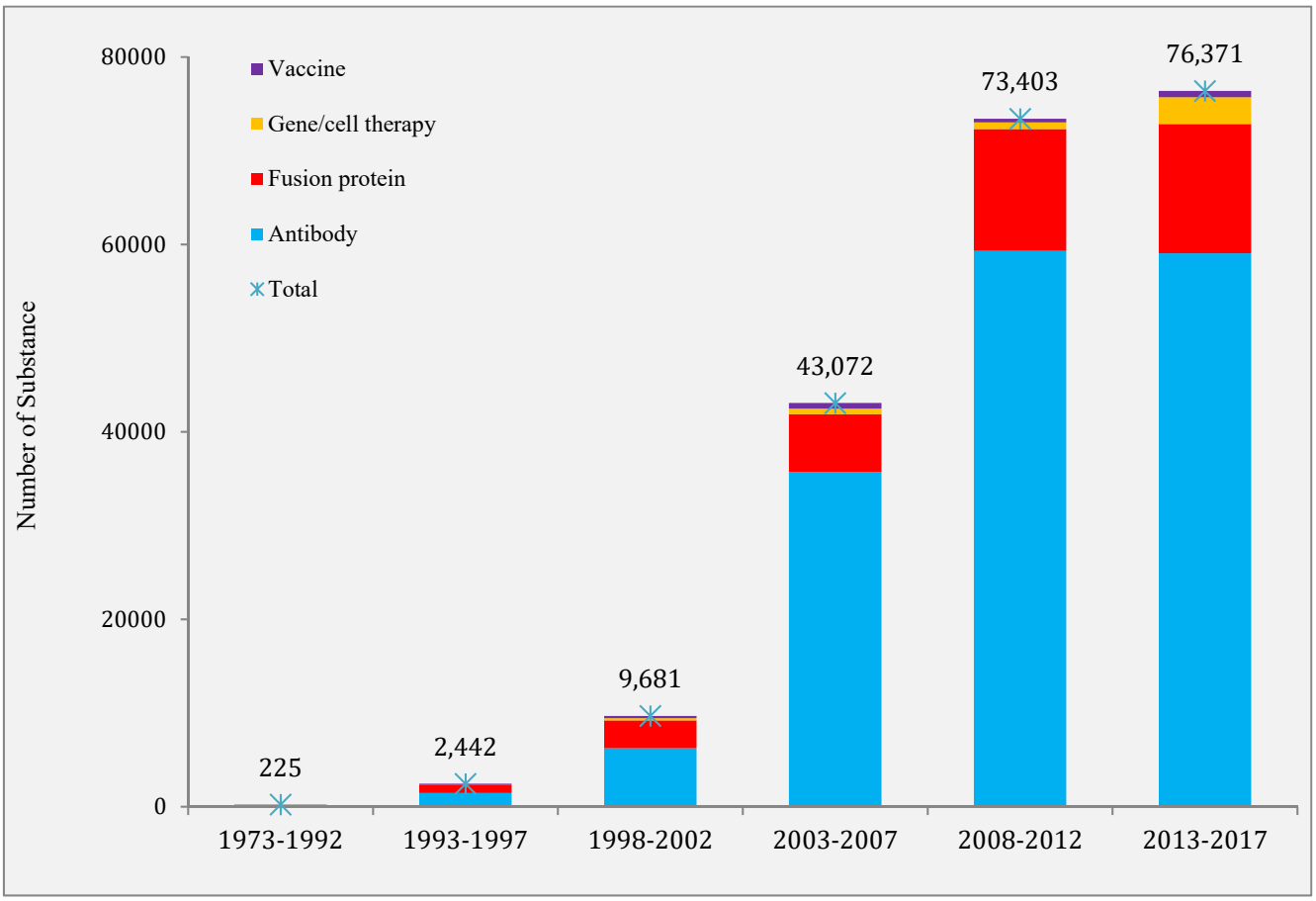


Figure 18. Timeline of Biopharmaceutical Substances Published in CAS REGISTRY℠

CAS not only generates registrations for substances appearing in CAS-indexed journals and patents, but also registers important substances from other sources. For example, CAS provides registrations for pharmaceutical substances during the application process for international non-proprietary names (INNs) by the World Health Organization (WHO) or US Adopted Drug Names (USANs) by the United States Medical Association special council. Prior to submission of a name proposal to either organization, the pharmaceutical/biotech company often needs to request a CAS Registration Number and Index Name for their chemical entity via CAS Client Services. As shown in Figure 19, although the total number of substances registered via CAS Client Services has been relatively low, these substances often are drugs that have either undergone, or are undergoing clinical trials, therefore, reliably reflect developing trends of biologics substances. Notably, the number of these substances has grown rapidly in recent years. In particular, 425 substances of these types have been registered during the last five years, which is almost three times of that during the previous five year period.

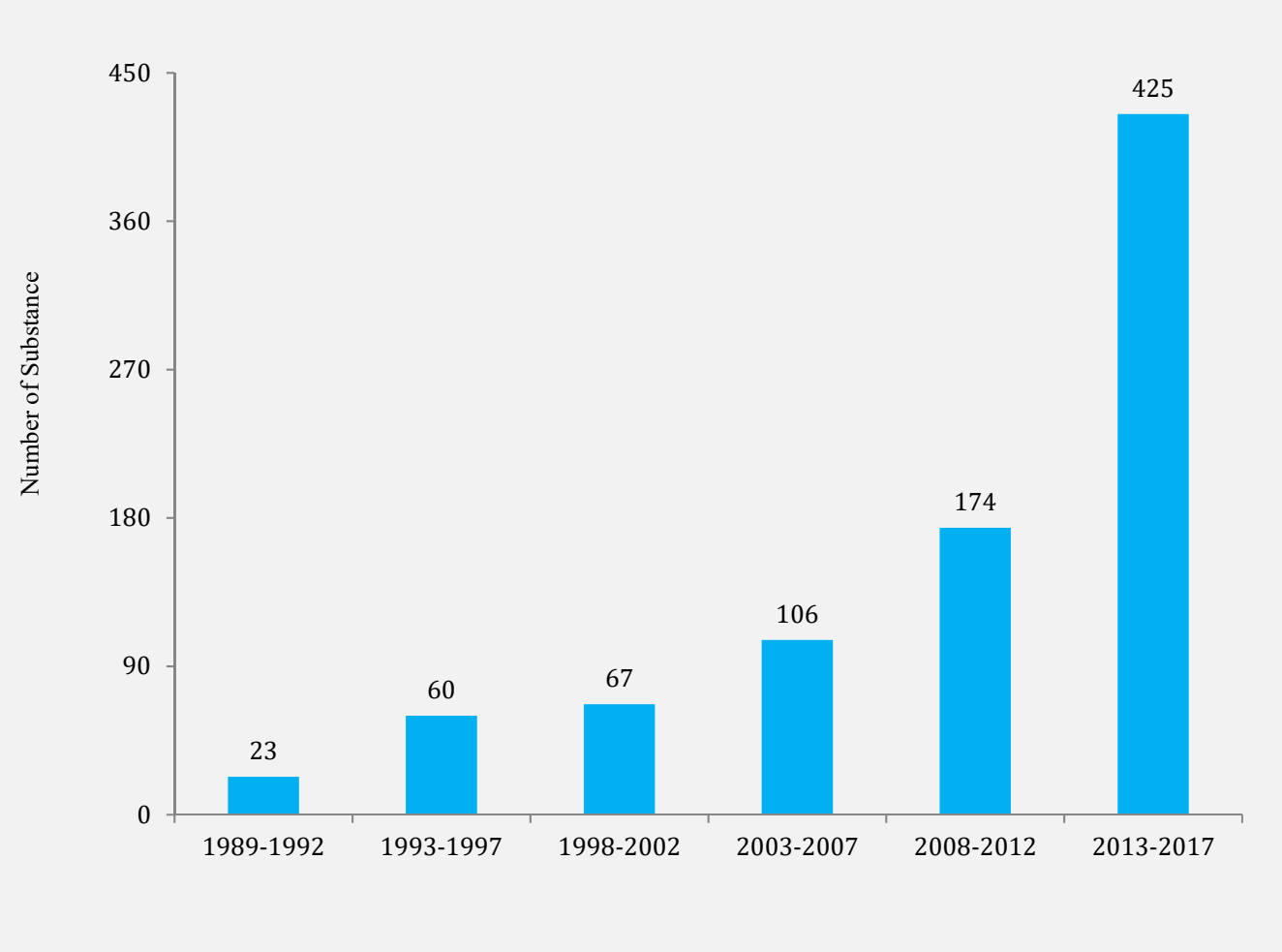


Figure 19. Number of biopharmaceutical Substances Registered via CAS's Client Services

4.3 DISTRIBUTION OF PROTEIN AND NUCLEIC ACID SEQUENCES

Proteins comprise a large subset of the registered biopharmaceutical substances. Because many such substances are produced via DNA recombinant technology and the development of protein drugs may be derived from their corresponding DNA sequences, the CAS's biopharmaceutical substances also include DNA sequences. The vast majority of registered biopharmaceutical substances in CAS's substance collection include bio-sequence information in addition to names. Protein sequences associated with antibodies and fusion proteins outnumbered the associated DNA sequences about 4:1 (Fig. 20).

Biologics associated with gene and cell therapy included 50% DNA sequences and 50% protein sequences. The vaccine class contains less associated sequence information from publication sources. Approximately 31.5% of the vaccine substances do not have sequence structural information as part of their registration records. These include inactivated/killed bacteria/viruses, attenuated viruses and a small number of toxoids.

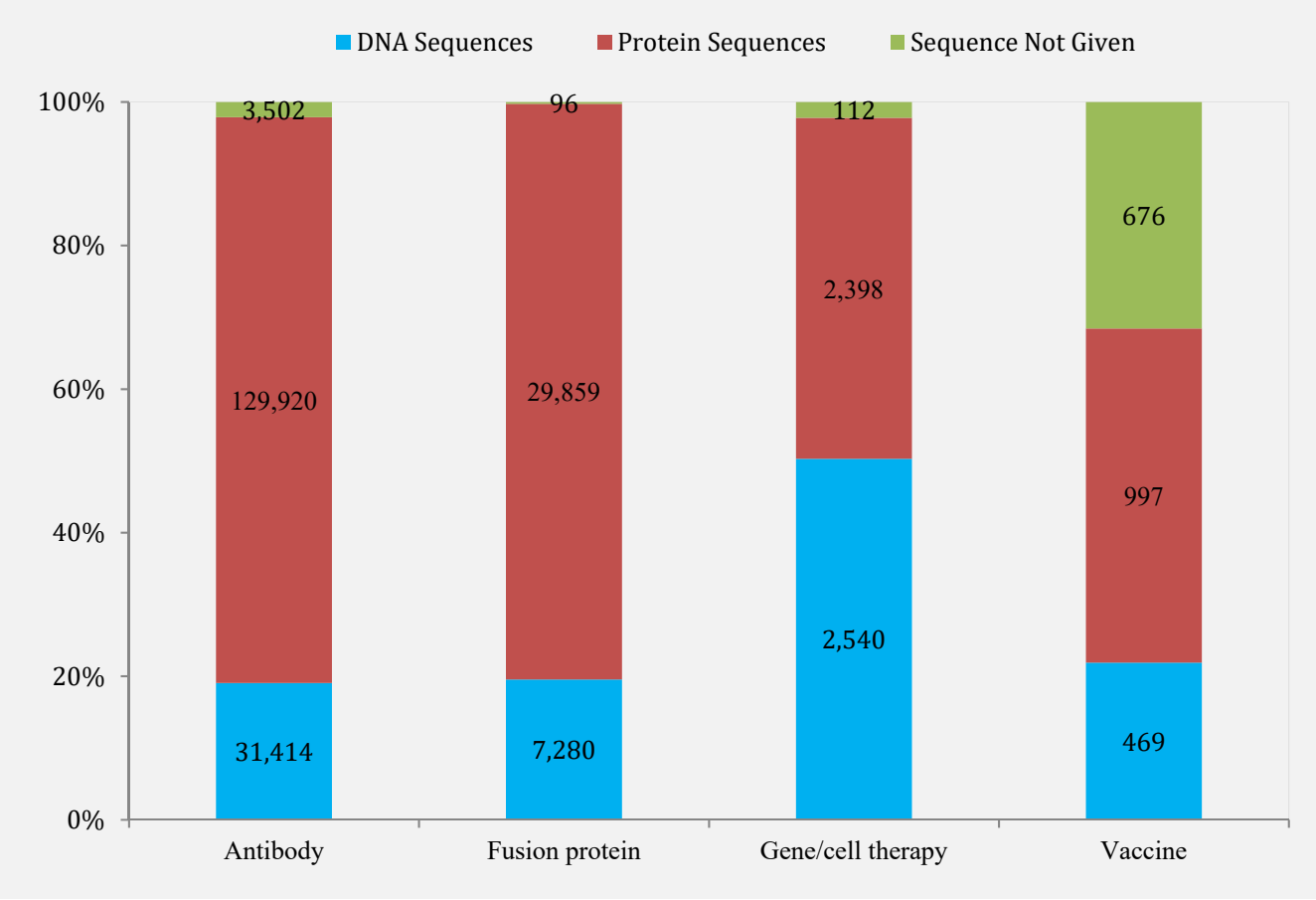


Figure 20. Sequence Information for Biopharmaceutical Substances

4.4 MAJOR BIOPHARMACEUTICAL SUBSTANCES IN THE FOUR CLASSES OF BIOLOGICS

Among the four classes of biologics, the antibody class has the largest number of substances, and the top ten most-studied substances are all commercially available antibody preparations (Figure 21). Among them, the most studied is rituximab, appearing in 12,439 papers and patents, followed by bevacizumab, trastuzumab, infliximab and cetuximab.

Fusion proteins rank second in number of biologic substance registrations. The top ten most studied substances are commercial fusion proteins. Among them, etanercept has been mostly studied in literature (5,125 documents), followed by abatacept, aflibercept, alefacept and sipuleucel-T.

The number of registered substances belonging to the gene and cell therapy class is relatively small, totaling only 5,050. Among them, picibanil has been mostly studied (1,022 documents), followed by sipuleucel-T, talimogene laherparepvec, and GVAx, and a 24 amino acid peptide (CAS RN 1387555-93-2) used in preparation of chimeric antigen receptors.

There were a total of 2, 142 registered vaccine substances, including both preventive and therapeutic vaccines. The top five most studied vaccines were preventive vaccines for HIV and influenza infections, a mycobacterium bovis BCG vaccine, tetanus toxoids and malaria vaccines. Sipuleucel-T was the most studied recombinant vaccine for immunotherapy of certain cancer with 327 published occurrences, followed closely by Gardasil, engerix-B, cervarix and rotarix.

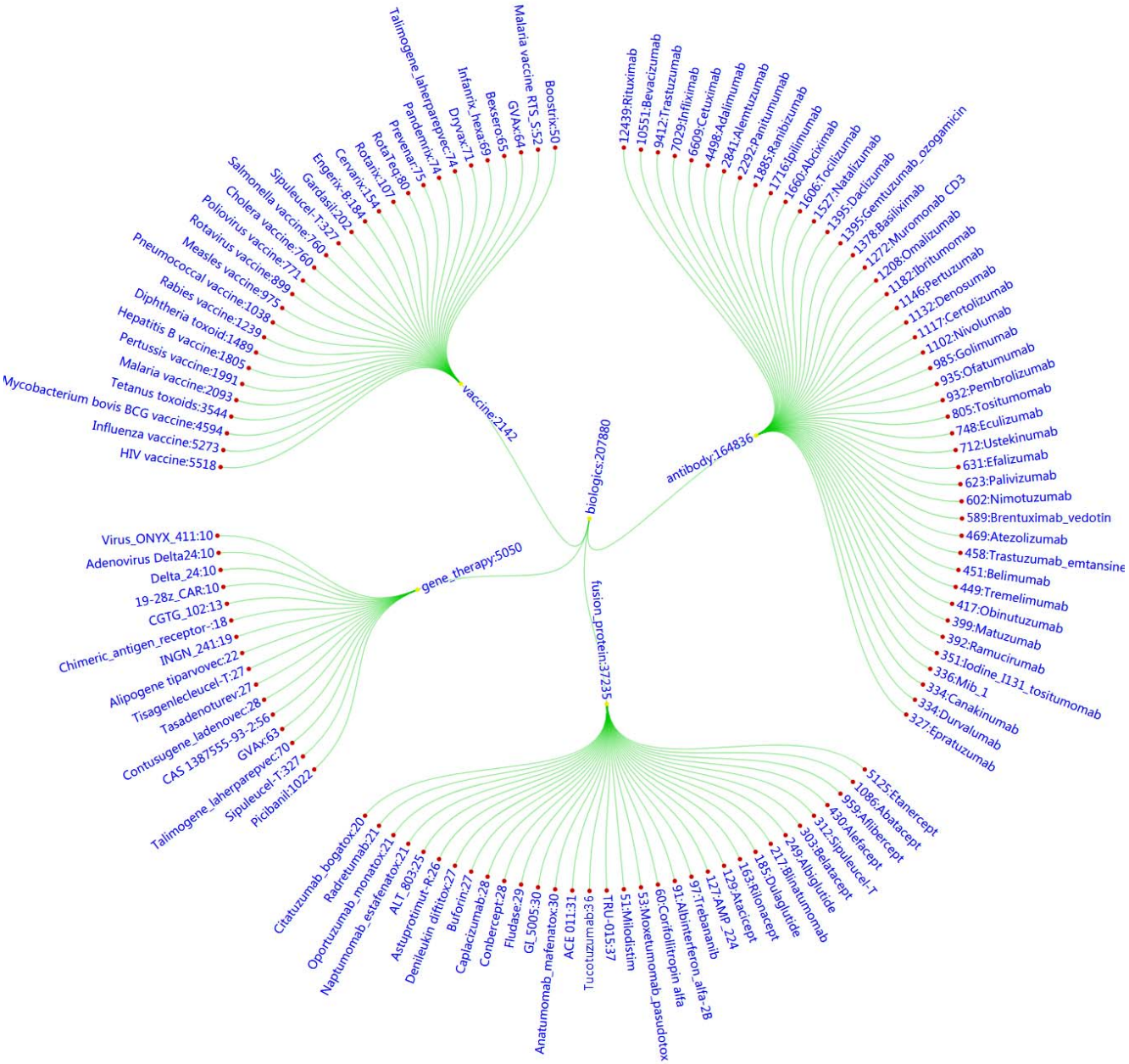


Figure 21. Substances in Four Biologics Classes

Notes for Figure 21: 1) The outer layer shows the substance name and its indexing frequency, and the inner layer shows the name of the biologics class and the number of substances in that class。 2) The values in the outer layer of the figure are the indexing frequency of each substance, which is used as an indicator for the research popularity of that substance, 3) CAS registry numbers are used for a few substances instead of their much longer names. 4) Because there was a large gap between the total number for each substance and the corresponding publication totals, different thresholds were used in the ranking.

4.5 SUBSTANCE-DISEASE RELATIONSHIPS

An analysis of the relationships between substances and diseases is valuable to revealing the potential drug applications. Figure 22 shows the substance-disease relationship heat map for biologics, with the numbers in the boxes indicating the number of times the substance and disease were reported together in the same publication. The color of the box indicates the significance or strength of the correlation. Overall, the number of antibodies reported in association with diseases was significantly higher than that of fusion proteins, gene and cell therapy, indicating that antibodies have been the most widely studied class of biologics in relation to diseases.

Antibodies occurred in publications associated with research on neoplasms, lymphatic diseases, immune diseases, inflammatory diseases and blood diseases. In contrast, fusion proteins were primarily associated with publications dealing with inflammatory diseases and immune diseases. The gene and cell therapy class was heavily concentrated in the area of oncology research, but the overall number of studies for this class of biologics was significantly lower.

Subfield	Substance \ Disease	Neoplasm	Digestive system disease	Infection	Lymphatic system disease	Immune disease	Respiratory system disease	Inflammation	Musculoskeletal disease	Hematopoietic disorders	Cardiovascular disease	Mental and behavioral disorders	Endocrine system disease	Degenerative disease	Nervous system disease	Connective tissue disease
Antibody	Rituximab	7,445	2,215	1,359	6,968	4,818	1,866	3,146	4,171	6,073	1,652	1,820	1,455	890	1,279	1,899
	Bevacizumab	7,745	4,271	547	1,384	1,038	2,729	1,089	1,538	1,526	1,541	2,274	2,174	1,045	1,208	1,159
	Trastuzumab	7,887	2,774	599	1,722	1,245	2,091	794	1,705	1,593	901	1,587	1,821	451	1,213	1,253
	Infliximab	858	2,926	849	736	3,751	789	5,702	3,228	640	919	664	528	2,435	428	575
	Cetuximab	5,702	3,664	414	1,117	887	2,191	601	1,066	1,079	506	1,196	1,258	274	771	811
	Adalimumab	588	1,501	607	454	2,821	587	3,621	2,416	427	582	489	381	1,253	306	417
Fusion protein	Etanercept	569	931	613	463	2,632	588	2,875	2,327	478	635	538	425	692	341	446
	Abatacept	126	204	135	111	602	131	582	565	113	124	127	128	142	84	145
	Aflibercept	417	343	66	127	110	189	140	141	130	168	153	261	253	109	106
Gene/Cell therapy	Piclabanil	118	65	24	53	46	53	21	46	51	17	46	43	14	40	31
	Sipuleucel-T	164	52	24	44	45	56	24	50	38	20	33	37	14	20	33
	Talimogene laherparepvec	39	20	5	18	12	21	4	18	16	2	20	15	3	20	16
Vaccine	Sipuleucel-T	164	52	24	44	45	56	24	50	38	20	33	37	14	20	33
	Cardasil	54	32	54	18	21	21	16	10	11	8	15	11	4	12	6
	Engerix-B	6	77	82	7	13	3	76	3	3	3	4	4	4	5	1

Figure 22. Substance-Disease Relationships

4.6 BIOPHARMACEUTICAL SUBSTANCE DESCRIPTIONS

Table 7. Descriptions of Major Biopharmaceutical Substances

Drug Name	Earliest Approval Date by FDA	Approved Indications	Description	Invention Owner
Rituximab (174722-31-7)	1997	Initially applied to chemotherapy-resistant B-cell non-Hodgkin's lymphoma	Antibody: Chimeric monoclonal antibody, targeting B cell surface antigen CD20	IDEC Pharmaceuticals
Bevacizumab (216974-75-3)	2004	Initial metastatic colorectal cancer	Antibody: Recombinant humanized antibody, targeting VEGF	Genentech
Trastuzumab (180288-69-1)	1998	Breast cancer	Antibody: Targeting HER2 receptor	Genentech & UCLA
Infliximab (170277-31-3)	1998	Autoimmune diseases such as psoriasis and arthritis	Antibody: Chimeric monoclonal antibody targeting TNF-α)	Centocor (Janssen Biotech)
Cetuximab (205923-56-4)[1]	2004	Initial colorectal cancer spread	Antibody: Chimeric antibody, targeting EGFR	ImClone/BMS
Adalimumab (331731-18-1)	2002	Rheumatoid arthritis, psoriasis, etc.	Antibody: Fully humanized monoclonal antibody, targeting TNFα	BASF Pharma (Abbott)
Etanercept (185243-69-0)	1998	Arthritis, psoriasis, etc.	Fusion protein: targeting TNFα	Immunex (Amgen)
Abatacept (332348-12-6)	2005	Rheumatoid arthritis	Fusion protein: immune response interfering with T cell	Bristol-Myers Squibb
Aflibercept (862111-32-8)	2011	Wet macular degeneration	Fusion protein: targeting VEGF	Sanofi and Regeneron
Picibanil (39325-01-4)	1986	Cystic phlegm (lymphangioma)	Gene and Cell Therapy: Mixture of group A streptococcus and anti-tumor properties	——
Sipuleucel-T (917381-47-6)	2010	Prostate cancer	Gene and Cell Therapy: Cell vaccine	Dendreon
Talimogene laherparepvec[1] (1187560-31-1)	2015	Melanoma	Gene and Cell Therapy: Herpes simplex virus type 1	BioVex (Amgen)
Gardasil (910046-32-1)	2006	Treatment of HPV-induced cancer	Vaccine: Human papillomavirus (HPV) vaccine	University of Queensland
Engerix-B (351186-51-1)	1986	Hepatitis B	Vaccine: Hepatitis B vaccine	——

A. Antibodies

- Rituximab: Human-mouse chimeric monoclonal antibody mainly used to treat neoplasm-related diseases. Single most highly indexed biopharmaceutical substance. Also highly associated with lymphatic diseases and hematopoietic disorders, indicating that its therapeutic effect in hematological tumors has been highly studied.
- Bevacizumab: Recombinant humanized monoclonal antibody also known as Avastin. Indexed with neoplasm, digestive diseases and respiratory diseases. Bevacizumab binds to human vascular endothelial growth factor (VEGF) to block its biological activity.
- Trastuzumab: Humanized monoclonal antibody targeting ErbB2 Primarily indexed with neoplasm in the literature but also with digestive diseases and respiratory diseases.
- Infliximab: Human-mouse chimeric monoclonal antibody- blocking tumor necrosis factor (TNF) α, which is highly associated with inflammation, immune diseases and locomotor system diseases.
- Cetuximab: Chimeric antibody targeting the EGF receptor (EGFR). Mainly indexed in is mainly found in neoplasm-related documents and less so indigestive disease- and respiratory disease-related documents.
- Adalimumab: Fully humanized monoclonal antibody targeting TNF. Mainly indexed with inflammation, immune diseases and locomotor system diseases.

B. Fusion Proteins

- Etanercept: Fusion protein of extracellular ligand binding site of human TNFR2/p75 linked to an Fc fragment of human IgG1. Most highly indexed fusion protein. Indexed with inflammation, immune diseases (autoimmune diseases), and locomotor system diseases. Originally used to treat diseases such as rheumatoid arthritis (RA) and ankylosing spondylitis (AS).
- Abatacept: Used to treat inflammation and immune diseases (autoimmune diseases). It inhibits T cell activation by binding to CD80 and CD86 on antigen-presenting cells.
- Aflibercept: Index primarily with digestive system neoplasms. It treats tumors by binding to a soluble receptor associated with human VEGF, resulting in reduced angiogenesis and vascular permeability.

C. Gene and Cell Therapy

- Picibanil: Most highly indexed gene and cell therapy-related substance. Mainly indexed withneoplasmand, (bladder cancer) and digestive diseases.
- Sipuleucel-T: Autologous cellular immunotherapy drug mainly indexed in neoplasm-related disease documents. Also categorized as a therapeutic vaccine and fusion protein. Used for the treatment of metastatic castration-resistant prostate cancer without symptoms or with mild symptoms.
- Talimogene laherparepvec: Also known as T-Vec or Imlygic. First approved by the US Food and Drug Administration in 2015 as a genetically modified herpes simplex virus type 1 for topical treatment of melanoma. Mainly indexed with neoplasm, digestive and respiratory diseases as well as mental and behavioral disorders.

D. Vaccines

- Sipuleucel-T: Mainly indexed with neoplasm, as described above
- Gardasil: Cervical cancer vaccine, prepared using the outer envelope of human papillomavirus (HPV) as antigen. Mainly indexed with neoplasm and infectious diseases.
- Engerix-B: Mainly indexed with infectious diseases, digestive diseases, and inflammatory diseases.

CHAPTER 5. CONCLUSIONS

This report contains the results of our scientific analyses of a wide diversity of biologics-related data provided by CAS and compiled over many years as part of their document and substance databases. It illustrates historical and recent trends in biologics research and development worldwide.

The past 30 years have witnessed rapid developments in global biologics research and development, accompanied by a steady upward trend in the number of published papers and patent applications. Of the four biologics classes selected for inclusion in this report, antibodies are associated with the largest number of publications (papers and patents), followed by vaccines, gene and cell therapy, and fusion proteins, respectively.

The scientific disciplines associated with biologics publications are primarily in the areas of immunochemistry, pharmacology, mammalian pathobiochemistry, biochemical genetics and pharmaceutics. Popular research subjects are biologics for cancer treatment, inflammation, antibodies (especially monoclonal antibodies), gene therapy, vaccines, stem cells, and various protein biologics.

The United States, China, Japan, Britain and Germany produced the most biologics research papers, while the US, China, Japan, Germany and South Korea had the most patents related to biologics. Most of the global patent applications flowed to the US, China, Japan and Germany.

MOST OF THE MAJOR RESEARCH INSTITUTIONS PUBLISHING BIOLOGIC RESEARCH PAPERS AND RELATED PATENT APPLICATIONS ARE FROM THE UNITED STATES.

The US, Germany, and Japan have applied for a large number of overseas' patents, in contrast to China who has shied away from exporting its patent technology. The United States has held the primary lead in biologics R&D. Most of the major research institutions publishing biologic research papers and related patent applications are from the United States. The vast majority of these research papers originated from academic research institutions and universities while most patent applications originated from commercial enterprises.

Also included in the analysis were biopharmaceutical substances in the CAS Registry database. Many of these substances contain sequence-related information. Antibodies comprised the highest number of registered biologics substances, accounting for 79% of the total. The most frequent mentioned biopharmaceutical substances in the scientific literature include rituximab, bevacizumab, trastuzumab, infliximab and cetuximab, all of which are antibodies.

The overall number of biopharmaceutical substances has been steadily increasing over time.

In spite of the small number of substances in the gene and cell therapy class, there was a significant growth of such substances over the past five years.

The diseases most studied in biologics documents were determined to be neoplastic diseases, digestive diseases, infectious diseases, lymphatic diseases, immune diseases, respiratory diseases and inflammatory diseases.

CHAPTER 6. OUTLOOK

Persistent innovation and rapid progress in the area of biologics-related biotechnology, have heightened awareness of its market prospects and potential for transformation of the health-care and medical fields. Rapid developments in biologics research have accelerated the integration of biotechnology with other disciplines such as chemistry and clinical medicine and have fueled growth in biologics R&D. Although this report has not specifically discussed the broader range of biologics including antibody-drug conjugates, antisense nucleic acids, RNA interference drugs, protein hormones, cytokines and enzymes, there is reason to believe that R&D in these areas will also increase in the future.

Medical biotherapy, involving the use of biologics for treatment of various diseases including cancer, autoimmune diseases, inflammation, infectious diseases, endocrine diseases and cardiovascular diseases has become an important emerging field. Biotherapy has become an important fourth approach in cancer therapy regimens in addition to surgery, radiotherapy and chemotherapy.

Biologics have helped clinicians to consolidate and accelerate the application of immunotherapy in cancer treatment protocols, enhancing the cure rates for certain cancers.

CANCER, AUTOIMMUNE DISEASES, INFLAMMATION, INFECTIOUS DISEASES, ENDOCRINE DISEASES AND CARDIOVASCULAR DISEASES HAVE BECOME AN IMPORTANT EMERGING FIELD.

Biologics have been instrumental in helping to promote the development of precision medicine and personalized medicine as a result of their targeting specificity and capability for patient customization. Much of the data analyzed in this report was derived from basic research papers and patent publications rather than from clinical trials or commercial data sources. Although a large majority of emerging biologics are still in the research stage, it is expected that a significant number will make it through clinical trials in the near future to be applied to the diagnosis and treatment of various diseases.

Despite their potential benefits, biologics present challenges due to their fragile biology and demand for complex manufacturing conditions and processes. Current production life cycles for biologics are longer and more complex than for small molecule drugs, making them more costly to produce.

Appendix 1: Top 100 Organizations Producing the Most Papers in Biological Medicine and Therapies

No.	Organization Names	Papers
1	University of California	5,366
2	National Institutes of Health (USA)	4,500
3	University of Texas System	4,343
4	Harvard University	2,291
5	University of Pennsylvania	2,103
6	Johns Hopkins University	2,003
7	University of Pittsburgh	1,902
8	Mayo Clinic	1,744
9	Osaka University	1,603
10	Academy of Military Medical Sciences	1,602
11	Chinese Academy of Sciences	1,576
12	Stanford University	1,553
13	University of Washington	1,544
14	Duke University	1,447
15	Institut Pasteur	1,442
16	Fourth Military Medical University of PLA	1,431
17	Chinese Academy of Medical Sciences	1,427
18	University of Tokyo	1,426
19	University of Michigan	1,389
20	University of Alabama	1,356
21	Sun Yat-sen University	1,330
22	Sloan-Kettering Institute for Cancer Research	1,294
23	Shanghai Jiao Tong University	1,274
24	F. Hoffmann-La Roche & Co. AG	1,273
25	Baylor College of Medicine	1,242
26	Huazhong University of Science and Technology	1,218
27	Zhejiang University	1,211
28	Karolinska Institutet	1,206
29	Novartis AG	1,202
30	Third Military Medical University	1,200
31	U.S. Department of Defense	1,196
32	Emory University	1,193
33	The University of Maryland System	1,182
34	Peking University	1,179
35	Sichuan University	1,163
36	Chinese Academy of Agricultural Sciences	1,159
37	Fudan University	1,150
38	Kyoto University	1,136
39	University of Oxford	1,129
40	Imperial College London	1,101
41	The Scripps Research Institute	1,090

No.	Organization Names	Papers
42	The Ohio State University	1,071
43	Jilin University	1,057
44	University College London	1,051
45	University of Wisconsin	1,043
46	Vanderbilt University	1,030
47	University of Toronto	1,021
48	Seoul National University	1,017
49	Regents of the University of Minnesota	1,005
50	Massachusetts General Hospital	993
51	University of North Carolina	987
52	Universiteit Leiden	960
53	Merck and Co., Inc.	955
54	Second Military Medical University	933
55	Chongqing Medical University	932
56	University of Florida	919
57	Washington University in St. Louis	915
58	Cornell University	887
59	Kyushu University	858
60	University of Melbourne	847
61	Southern Medical University	834
62	University of Colorado	818
63	GlaxoSmithKline	817
64	Pfizer Inc.	802
65	Yale University	801
66	U. S. Food and Drug Administration	762
67	State University of New York	735
68	Icahn School of Medicine at Mount Sinai	719
69	Northwestern University	713
70	Fred Hutchinson Cancer Research Center	708
71	University of Illinois	707
72	Universite Catholique de Louvain	701
73	University of Chicago	700
74	Hokkaido University	699
75	Central South University	694
76	Dana-Farber Cancer Institute	693
77	Radboud Universiteit Nijmegen	687
78	New York University	665
79	University of Iowa	651
80	Columbia University	637
80	University of Queensland	637
82	University of Southern California	632
83	Keio University	629
84	Universitaet Zuerich	624
85	Medizinische Universitaet Wien	619
86	Nagoya University	614

No.	Organization Names	Papers
87	Lunds Universitet	606
88	Amgen Inc.	591
89	Brigham and Women's Hospital	588
90	Yeshiva University	581
91	University of Alberta	578
92	National University of Singapore	570
92	University of Cambridge	570
94	Thomas Jefferson University	566
95	Shandong University	559
96	Natl. Inst. Health Japan	549
96	University of Massachusetts	549
98	Universiteit Van Amsterdam	546
99	Nanjing Medical University	529
100	Universitaet Heidelberg	528

Appendix 2: Top 100 Organizations with Most Patents in Biological Medicine and Therapies

No.	Organization Names	Papers
1	GlaxoSmithKline	3,509
2	F. Hoffmann-La Roche & Co. AG	2,416
3	Novartis AG	1,965
4	University of California	1,548
5	Merck and Co., Inc.	1,395
6	U.S. Department of Health and Human Services	1,385
7	Pfizer Inc.	1,181
8	Institut National de la Sante et de la Recherche Medicale	1,067
9	Chinese Academy of Sciences	1,017
10	Bayer AG	1,006
11	Takeda Pharmaceutical Co., Ltd.	950
12	Sanofi	937
13	University of Texas System	883
14	Amgen Inc.	835
15	Bristol-Myers Squibb	805
16	Centre National de la Recherche Scientifique	795
17	AstraZeneca	736
18	Incyte Corp.	686
19	Johns Hopkins University	678
20	University of Pennsylvania	669
21	Johnson & Johnson	645
22	Academy of Military Medical Sciences	590
23	Institut Pasteur	566
24	Massachusetts General Hospital	557

No.	Organization Names	Papers
25	Chinese Academy of Agricultural Sciences	527
26	Harvard University	522
27	Merck KGaA	480
28	Leland Stanford Junior University	471
29	Baxter International Inc.	447
30	Seoul National University	443
31	Columbia University	435
32	Biogen, Inc.	426
33	Chinese Academy of Medical Sciences	419
34	Massachusetts Institute of Technology	418
35	Abbott Laboratories	413
36	Brigham and Women's Hospital	407
37	Dana-Farber Cancer Institute	402
38	Boehringer Ingelheim GmbH	393
39	The Scripps Research Institute	390
40	Duke University	389
41	Eli Lilly & Company	380
42	Chugai Pharmaceutical Co., Ltd.	367
43	Sloan-Kettering Institute for Cancer Research	335
44	Fudan University	330
45	Zhejiang University	325
46	Yeda Research and Development Co. Ltd.	319
47	CSL Ltd.	316
48	Regeneron Pharmaceuticals, Inc.	308
48	University of Michigan	308
50	The University of Maryland System	304
51	Yale University	300
52	University of Florida	291
53	Korea Research Institute of Bioscience and Biotechnology	284
54	Osaka University	276
55	Shanghai Jiao Tong University	268
56	Cornell University	267
57	Ludwig Institute for Cancer Research	264
58	Societe des Produits Nestle S.A.	263
59	Novo Nordisk A/S	259
60	U.S. Department of Defense	252
61	University of Pittsburgh	247
62	New York University	246
63	Sun Yat-sen University	245
64	Emory University	244
65	Agency for Science, Technology & Research	243
65	Children's Medical Center Corp.	243
65	University System of Georgia	243
68	Kyoto University	241
68	Mayo Clinic	241

No.	Organization Names	Papers
70	UAB Research Foundation	239
71	Baylor College of Medicine	235
72	Max-Planck-Gesellschaft	233
73	University of Tokyo	232
74	Third Military Medical University	230
75	University of North Carolina	228
76	University of Oxford	224
77	Beth Israel Deaconess Medical Center	221
78	Japan Science and Technology Agency	220
79	Wisconsin Alumni Research Foundation (WARF)	218
80	University of Washington	215
81	Deutsches Krebsforschungszentrum Stiftung des Oeffentlichen Rechts	214
82	Yonsei University	213
83	Kyowa Hakko Kirin Co., Ltd.	212
84	Regents of the University of Minnesota	208
85	Second Military Medical University	204
86	Immunomedics, Inc.	203
86	University of Southern California	203
88	University of Massachusetts	200
89	University of Rochester	199
90	Pulike Biological Engineering Co., Ltd.	191
91	Chemo-Sero-Therapeutic Research Institute	187
91	Icahn School of Medicine at Mount Sinai	187
91	Peking University	187
94	Cedars-Sinai Medical Center	183
95	Imperial College London	181
96	Fourth Military Medical University of PLA	177
96	Rockefeller University	177
98	Yeshiva University	176
99	University of Chicago	175
100	AkzoNobel	173

RESEARCH WORKING TEAM

Project Planning

Bureau of Planning and Strategy, Chinese Academy of Sciences: Chenguang Huang, Quan Gan

National Science Library, Chinese Academy of Sciences: Huizhou Liu, Xiwen Liu, Yajuan Zhao, Xuezhao Wang, Tao Han

CAS (a division of American Chemical Society): Manuel Guzman, Matthew Toussant, Michael Dennis, Craig Stephens, Dennis McCullough, Caroline Ma

Research & Analysis

National Science Library, Chinese Academy of Sciences: Yanping Yang, Tao Han, Hualing Xie, Fang Chen, Dongqiao Li, Peijuan Chi, Xiaoli Chen

CAS (a division of American Chemical Society): Matthew Toussant, Cynthia Liu, Yingzhu Li, Yi Deng, Sunny Yu

Publicity & Promotion

National Science Library, Chinese Academy of Sciences: Lu Dong, Nan Li, Muhong Zou

Chemical Abstracts Service (CAS): Divya Ramakrishnan, Rhonda Ross, Nikki Stoughton, Dawn Fournier, Cristina Tomeo, Caroline Ma, Jinying Zhang

Consultants

Min Fang (Researcher, Institute of Microbiology, Chinese Academy of Sciences)

Haiying Hang (Researcher, Institute of Biophysics, Chinese Academy of Sciences)

Tao Jiang (Senior Engineer, Institute of Genetics and Development Biology, Chinese Academy of Sciences)

Jiandong Jiang (Director & Researcher, Institute of Materia Medica, Chinese Academy of Medical Sciences)

Xiayang Qiu (President & Chief Scientific Officer, QL Regor Inc)

Ao Zhang (Researcher, Shanghai Institute of Materia Medica, Chinese Academy of Medical Sciences)

Dongyu Wang (Post-Doc, Shanghai Institute of Materia Medica, Chinese Academy of Medical Sciences)

Wendy Zhang (Senior Director of Biologics, Bioduro)

Acknowledgement

The report has also received guidance and assistance from other experts in related fields. Siyuan Zheng and other graduate students at the National Science Library have made important contributions to the research and report writing. We would like to express our heartfelt thanks to all the help!

ABOUT CAS

CAS is the Scientific Information Solutions Division of the American Chemical Society. We partner with R&D organizations globally to provide actionable scientific insights that help them plan, innovate, protect their innovations, and predict how new markets and opportunities will evolve. Scientific researchers, patent professionals and business leaders around the world across commercial, academic and government sectors rely on our solutions and services to advise discovery and strategy. Leverage our unparalleled content, specialized technology, and unmatched human expertise to customize solutions that will give your organization an information advantage. With over 110 years' experience, no one knows more about scientific information than CAS.

CAS: WHERE SCIENCE AND STRATEGY CONVERGE.

ABOUT THE NATIONAL SCIENCE LIBRARY, CHINESE ACADEMY OF SCIENCES

The National Science Library, Chinese Academy of Sciences (NSLC) is the research library service system of CAS as well as the National Library of Sciences in the Chinese National Science and Technology Libraries (NSTL) system. NSLC functions as the national reserve library for information resources in natural sciences, inter- disciplinary fields, and high-tech fields, serving the researchers and students of CAS and researchers around the country. It also provides services in information analysis, research information management, digital library development, scientific publishing (with its 17 academic and professional journals), and promotion of sciences.

NSLC is proactively leading national efforts to build a shared National Scientific Information Infrastructure. As the key member of NSTL, it organizes research and training activities for strategic planning for STM libraries, digital library standards, knowledge organization systems, digital preservation technologies and practices, and open access policies, in addition to provision of STM information nation-wide. It collaborates with major domestic and foreign libraries for resources sharing and research in library and information services.

NSLC is credited, under the auspice of CAS University, to offer master and doctoral degree programs in library and information sciences, with a yearly enrolment of about 50. The library also hosts senior visiting scholars and organizes vocational training and continuing education programs. It is one of the most published and cited research organizations in library and information sciences in China, and the hosting institute for the China Society of Special Libraries.

NSLC, aiming at developing a world first-class information service ability and leadership in library development in the country, strives to strengthen its resources, improve its systems, and innovate its services, to best suit its users.

CAS®
A DIVISION OF
THE AMERICAN CHEMICAL SOCIETY

中国科学院文献情报中心
National Science Library, Chinese Academy of Sciences